Recommendations

- Huawei Learning Website
  - http://learning.huawei.com/en
- Huawei e-Learning
  - https://ilearningx.huawei.com/portal/#/portal/ebg/51
- Huawei Certification
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_31&lang=en
- Find Training
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_trainingsearch&lang=en

More Information

- Huawei learning APP

版权所有© 2019 华为技术有限公司



Unveiling DCs

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This course describes the development course of data centers (DCs), basic modules of DCs, and evolution trends of cloud DCs.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- On completion of this course, you will be able to know:
  - DC development course
  - Basic modules of a DC
  - Network structure
  - Evolution trends of cloud DCs

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

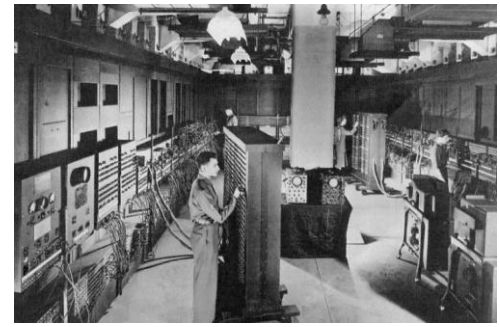
1. **DC Development Course**
2. Basic Modules of a DC
3. Evolution Trends of Cloud DCs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

DCs Took Shape - ENIAC

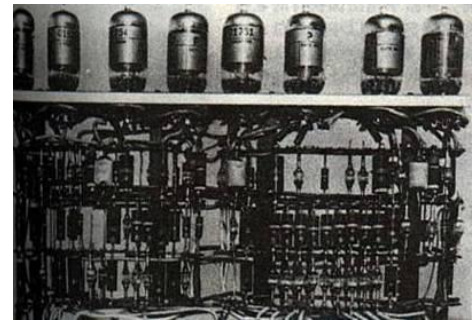
- Electronic Numerical Integrator and Computer (ENIAC) was designed in 1946 and primarily used to calculate artillery firing tables for the United States Army's Ballistic Research Laboratory.
- ENIAC contained 17,468 vacuum tubes, 7200 crystal diodes, 1500 relays, 70,000 resistors, 10,000 capacitors, 1500 relays, and 6000-plus switches. It could execute 5000 addition operations or 400 multiplication operations per second. It was one thousand times faster than electro-mechanical machines and 200,000 times faster than manual calculation.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Commercial Use of DC Virtualization Technology - TRADIC

- TRAnsistor DIgital Computer (TRADIC) was the first transistorized computer in the USA, completed in 1954. It was put into commercial use in 1960s and led the breakthrough development of mainframe computers (such as IBM System series mainframe computers). Mainframe computers were developed for government departments and militaries and have special requirements on sites and security systems. As the requirements for system performance increased, people hoped to provide additional performance support and resource sharing capabilities on the same systems. To cope with the increasing requirements, virtualization technology came into view.
- The concept of virtualization was popularized rapidly, and the multi-task processing mechanism was further improved on mainframe computers. Virtualization technology was first put into commercial use on IBM VM/370 OS in 1972. Virtualization technology, one of the most important technologies of DCs, was integrated into the development course of DCs.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Transistorized computers were the second generation of electronic computers. In 1954, Bell Labs in the United States successfully developed the first computer using transistor circuits and called it "TRADIC" with 800 transistors.
- In the first generation before the 1950s, electronic tubes were used as components in computers. Too much heat generated by the tube elements during operation, poor reliability, unpleasant operation, high cost, and large size restricted computer development. Thus, the transistor began to be used as a computer component. The transistor could not only realize the functions of electron tubes, but also featured small size, light weight, long life, high efficiency, less heat, and low power consumption. After the transistor was used, the structure of the electronic circuit was greatly improved, making high-speed electronic computers even easier to implement.

Client/Server Computing Model and the Internet

- 20 years ago, microcomputers ushered in prosperous development. Legacy PCs were replaced by network devices. Especially, the generation of the client/server model contributed to hosting and external DCs.
- In the middle of the 1990s, the Internet emerged and had great impact on the market. Additionally, it provided more options for DC deployment in the next 10-plus years. More enterprises needed to support Internet applications, network connections and collaboration services became necessary when enterprises deploy IT services. Network providers and hosting providers developed rapidly in the construction of hundreds of DCs. As a service mode, DCs have been accepted by most enterprises.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Client/Server (C/S) structure
  - The C/S structure is a well-known software system structure. It appropriately allocates tasks to the client and server, reducing the communication overhead. A client is required to perform management operations.
  - The programs running on a client are different from those running on a server. Users' programs reside on clients. The programs running on a server provide data management, data sharing, data and system maintenance, and concurrency control, and those running on a client process specific services.
  - Programs are easy to develop and operate. However, it is difficult to upgrade the applications and maintain the client programs.
- Browser/Server (B/S) structure
  - The B/S structure is a change or improved structure to the C/S structure with the emergence of Internet technology. In this structure, the user interface is implemented through the WWW browser.
  - The client does not have dedicated applications, and applications are basically on the server. Therefore, application upgrade and maintenance are performed on the server, which is convenient. Because the client uses a browser, the user interface is diversified, but functions such as data printing and output are limited. To overcome this disadvantage, the function that is difficult to implement by using a browser is developed into a control, which can be invoked by the client applications.

Energy Consumption Issue

- PC prosperity and DC appearance brought a series of problems, for example, occupying more places and increasing energy consumption. Back to 2002, DCs had consumed 1.5% power of the United States and the energy consumption increased by 10% every year. 5 million new servers were deployed in DCs and energy consumption by thousands of households increased every year.
- DC owners realized the problems and started to deploy more economical, efficient, and environment friendly infrastructure. In 2007, large-scale DC operators started to use recyclable energy technology (wind energy and solar energy) to support the daily operation of DCs.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Modular DCs

- Modular DCs are also called container DCs because the devices are deployed in containers. Sun BlackBox is one of the most famous modular DCs. In Sun BlackBox, 280 servers are deployed in 20-inch cabinets in the containers and carried to different locations worldwide.
- Although modular DCs are not as magnificent as conventional DCs, their construction cost is only 1% of conventional DCs. Additionally, modular DCs are flexible, remarkably reducing the time required for DC deployment.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Modular DCs are a new generation of DCs based on cloud computing. To cope with the server development trends, such as cloud computing, virtualization, centralization, and high density, the modular design concept is adopted to minimize the coupling of infrastructure with the equipment room environment. Subsystems, such as power distribution, cooling, cabinet, air flow control, integrated cabling, and power and environment monitoring to improve the overall operation efficiency of the DC and achieve rapid deployment, flexible expansion, and energy saving. In terms of configuration form, DCs can be divided into MDCs and CDCs.
- Modular DCs meet the urgent requirements of IT business departments for future DC infrastructure construction, such as standard design, component prefabrication, fast rollout and deployment, effective reduction of initial investment, energy pooling management in modules, high utilization of dynamic IT infrastructure resources, intelligent O&M management, and assurance of important service continuity, shared IT services (such as cross-service infrastructure, information sharing, and application sharing), quick response to service requirement changes, and green DCs.
- Advantages of modular DCs
  - Standard modules with high reliability
    - The modular DC adopts the modular, standard, and highly reliable design, which ensures the stability of the entire system. Based on customer requirements and actual conditions, the modular DC provides N+1, N+X, and 2N design solutions for core power supply and cooling devices. The design solutions are secure and reliable and meet the standards from Tier 3 to Tier 4.

Cloud DC

- Software as a service (SaaS) enables the shift from computing resource subscription based on infrastructure to on-demand subscription. Network infrastructure and DC operators work together to provide rapidly-increasing data bandwidth resources, which support a wide array of IT services.
- At the beginning, providers did not realize the rapid development of cloud DCs. However, cloud service providers such as Amazon and several other infrastructure service providers had a large number of users based on the cloud DC platform.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. DC Development Course
2. **Basic Modules of a DC**
3. Evolution Trends of Cloud DCs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

What Is a DC?

Wikipedia: A DC is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (for example, air conditioning and fire suppression) and various security devices.

Google: A DC is a multifunctional facility that can contain multiple servers and other communication devices. These communication devices have the same requirements on environments and physical security.

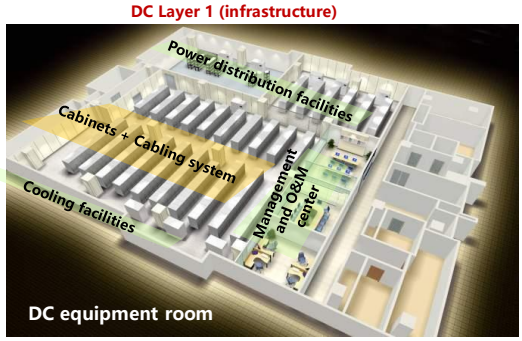
Huawei: Enterprise DCs are like reservoirs, and DC solutions are like the Dongting Lake.

- **Reservoir + Lake = DC (data computing, storage, and exchange center)**
- Pacific + Yangtze River + Yellow River = Carriers' network channels
- Tributaries + Urban pipeline networks = Enterprise network channels
- Water faucets = Network terminals
- Water = Data flows (voice/data/video)

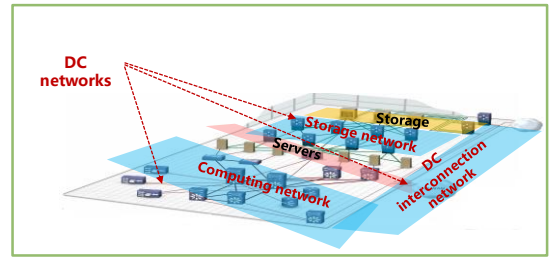
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Modules of a DC

DC Layer 1 (infrastructure)



DC Layer 2 (ICT devices)



User

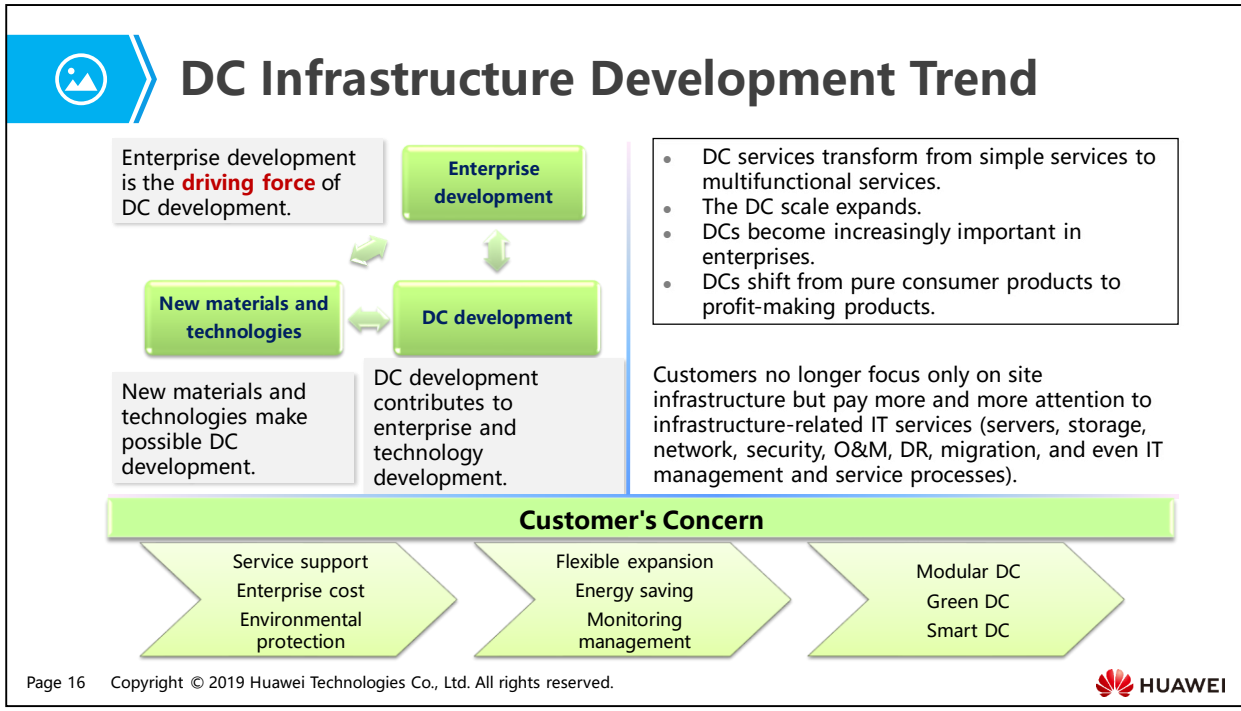
A DC is a service-oriented infrastructure. It supports the operation and growth of enterprise businesses. It consists of the following:
Secure network architecture, reliable support facilities (equipment rooms, generators, UPSs, and air conditioning systems)**, integrated servers/application platforms, centralized storage and backup, unified system management platforms, and O&M management organizations and processes** oriented to customer services.

Functions of DC networks: connects to server, storage, and cross-DC resources. The computing network is the core of the DC networks.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

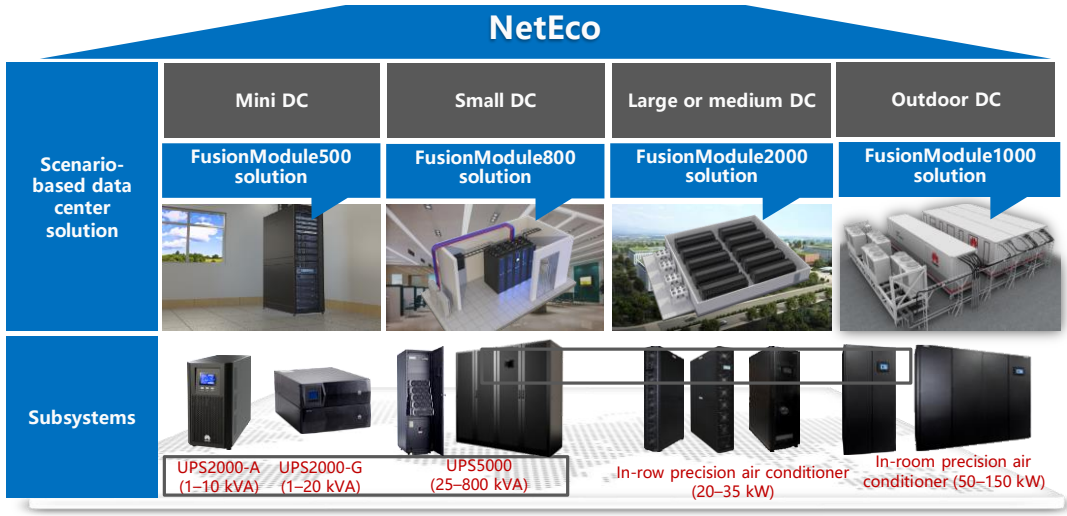
HUAWEI

- Especially Internet Data Centers (IDCs): Wikipedia (Switzerland) is in the underground shelter.
- Computing network: connecting to computing resources and users
- Storage network: connecting to storage
- DC interconnection: connecting to cross-DC resources
- L0: building: civil engineering and equipment room building.
- L1: infrastructure: equipment room auxiliaries, such as equipment room decoration, power supply, and cooling fire fighting.
- L2: IT infrastructure: infrastructure layer of the equipment room, including servers, storage devices, networks, and virtualization software.
- L3: application platform, web hosting, and PASS.
- L4: service: various services.



- As mentioned above, IT systems and informatization construction have become the driving forces behind the development of enterprises. That is to say, the development of enterprises, specifically, the development of their services, continuously raise requirements for IT systems and informatization construction, which is the driving force behind the development of DCs.
- The development of DCs — the implementation of various functions and services also promotes the further development of enterprises. Likewise, the development of DCs raises new requirements for related materials and technologies. The emergence of new technologies and materials makes the new requirements of DCs possible, promoting enterprise development.
- For example, an enterprise's business expansion makes its DC increasingly large. Daily O&M management becomes increasingly complex, and costs and risks are increasing. The demand for intelligent management of DCs is put forward. The intelligent management of DCs imposes demanding requirements on materials such as automatic data collection as well as technologies such as databases, application processing programs, and front-end exhibition platforms. With the maturity of new materials and technologies, intelligent management of DCs improves DC performance, excludes the bottlenecks of enterprise development, and promotes the further development of enterprise business.

DC Fusion Module Solution



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

Future DCs

DC		
	Environment	Environment that ensures the reliable running of information systems
	IT infrastructure	IT infrastructure on which information systems depend
	Monitoring management	Effective management and O&M of infrastructure

Customer requirements determine the future of DCs.



Modular DC

- Copes with unpredictable service and IT growth.
- Controls the initial investment and operation cost based on IT requirements.
- Modular power supplies and cooling systems enable online system expansion.



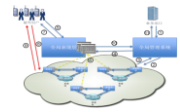
Smart DC

- Implements the unified management and monitoring of IT devices, site facilities, and IT processes.
- Supports energy management and assets management.
- Supports various technologies, such as real-time information, simulation, and remote monitoring.



Green DC

- Uses various technologies and measures to effectively reduce the OPEX, especially the Power Usage Effectiveness (PUE) value.



Cloud DC

- Supports resource sharing, flexible delivery, and dynamic resource management.
- Supports the high-density development trend.
- Supports standard expansion.
- Supports multi-DC O&M.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What will DCs in the future look like?
- For starters, in terms of enterprise requirements and DC functions, DCs in the future will be no longer of civil engineering concepts. At least the site environment, IT devices, and monitoring management will be included in the DC category. The site environment ensures the reliable and stable running of the information system. The information system operates based on IT equipment. Monitoring management includes the effective management of the site infrastructure and IT infrastructure.
- Second, we want to talk about the characteristics of DCs. In the previous slide, we mentioned the modular DC, green DC, and smart DC. Here, we want to add the concept of cloud DC.
- Modular DCs meet the requirements of unpredictable service and IT growth and control the CAPEX and OPEX based on the IT requirements at the same time. They use modular power supply and cooling, and do not require service interruptions during system expansion.
- Green DCs focus on energy conservation and consumption reduction, that is to say, using multiple technologies and means to effectively lower the OPEX (lowering PUE as the key).
- Smart DCs monitor and manage IT devices, site facilities, and IT processes in a centralized manner, manage resources and assets, and implement real-time information, simulation, and remote monitoring technologies.

Key Indicators of a Green DC

DC PUE

$$\text{PUE} = \frac{\text{Total power consumption of a DC}}{\text{Power consumption of IT equipment}}$$

The ideal PUE ranges from 1.6 to 2.0, or even lower.

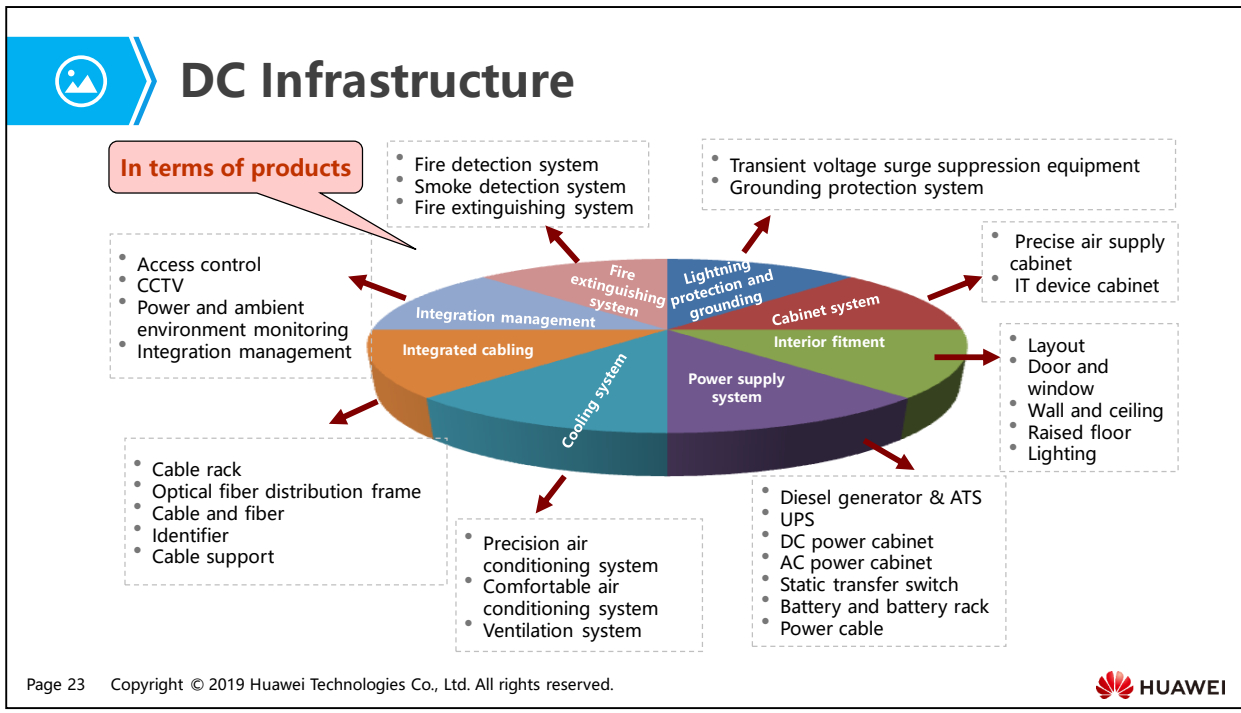
The common PUE ranges from 2.0 to 2.5, or even higher.

Key indicators of a green DC

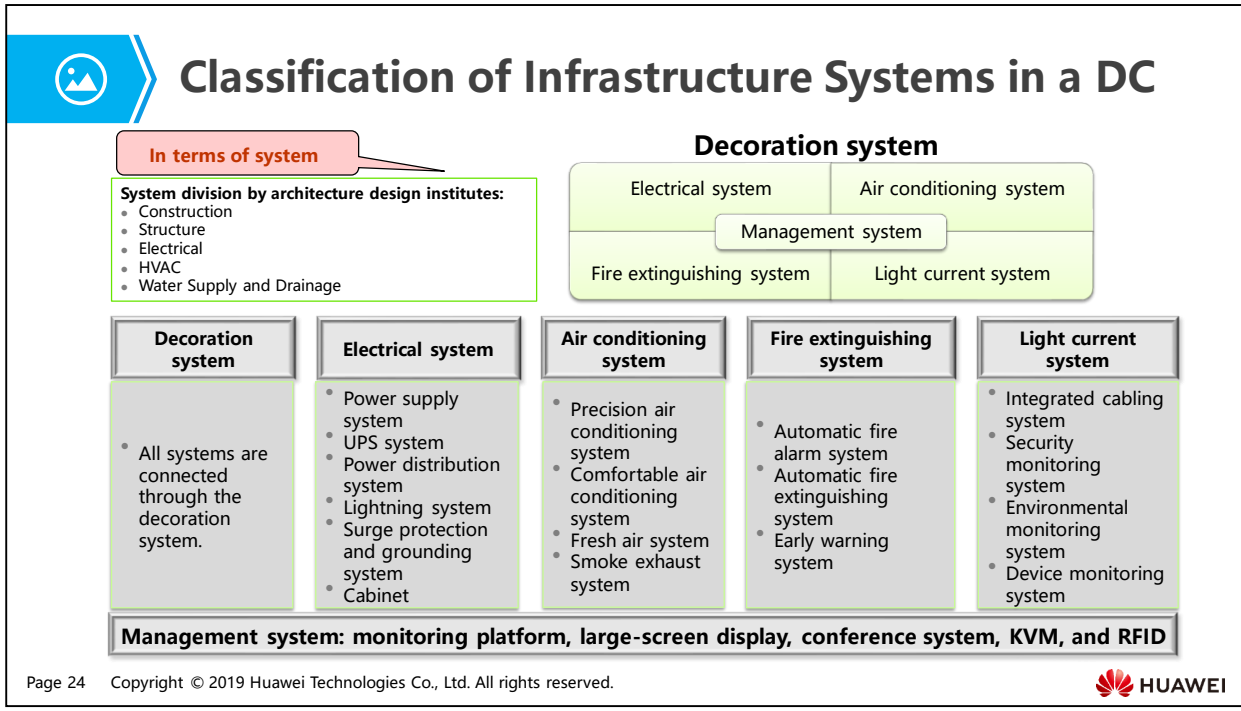
PUE	DCiE	Level of Efficiency
3.0	33%	Very Inefficient
2.5	40%	Inefficient
2.0	50%	Average
1.5	67%	Efficient
1.2	83%	Very Efficient

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The power conversion loss is about 10% when the mainstream UPS and power distribution system are used. The air conditioner cooling energy efficiency ratio (EER) is mostly between 3.5 and 5.5. The ideal PUE is between 1.6 and 2.0 and even lower when the impact of lighting, maintenance, and heat penetration (usually less than 5%) is ignored.
- In DC construction, a too low PUE may reduce the return on investment. An effective way to reduce the PUE is to optimize the heat dissipation system design.

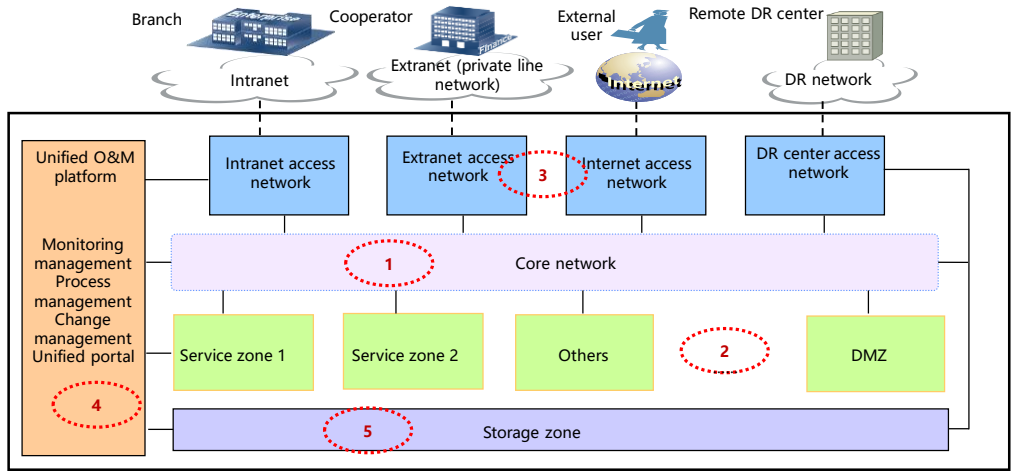


- We have talked about the function room and classifications of the infrastructure hierarchically. Next, we will talk about its systems and compositions.
- As we all know, the DC site infrastructure is an integration of many subsystems. This picture illustrates the L1 infrastructure multi-system in a Huawei DC. It can be seen that we divide the infrastructure into eight systems.
- Power supply system: It includes the DG, ATS, UPS. DC cabinet, AC cabinet, static transfer switch, storage battery, battery rack, and power cables.
- Cooling system: It includes the precision air conditioner, comfort air conditioner, and ventilation system.
- The other six systems include the interior decoration, cabinet, surge protection/grounding, fire-fighting, integrated cabling, and integration management, all of which contain some components. We will skip them here.
- We find that the eight systems defined here put much emphasis on products, for example, speaking of the power distribution system, we tend to emphasize the DG, ATS, and UPS under the system. The DC construction is a site engineering project. To meet a certain function or meet a certain requirement of a user, the DC must be considered from the perspective of the system. A single device or the stack of some devices cannot meet the requirements. We must consider the relationship between devices, connections, and devices to ensure that the system design and installation can meet the requirements.



- Next, we will talk about the construction of L1 infrastructure in the DC from the perspective of the system.
- Generally, building design institutes have several professional domains, including building, structure, electrical, heating, and water supply and drainage. Some design institutes are specially configured with hot energy and economics departments. Due to special characteristics of DCs, a DC is generally divided into decoration system, electrical system, air conditioning system, fire extinguishing system, light current system, and management system. (Some companies or design institutes also incorporate management systems into the light current systems.)
- Each system has certain functions.
- The electrical system meets the power supply requirements of equipment (including IT equipment, power equipment, and auxiliary equipment). The air conditioning system meets the air conditioning requirements in various environments. The fire extinguishing system meets the fire alarm and fire extinguishing requirements. The light current system meets various requirements, such as data communication, security defense, and environment device monitoring. These systems need to be decorated (including separation and deployment) to achieve effective operation through the management system.

Logical Zones of a Typical DC

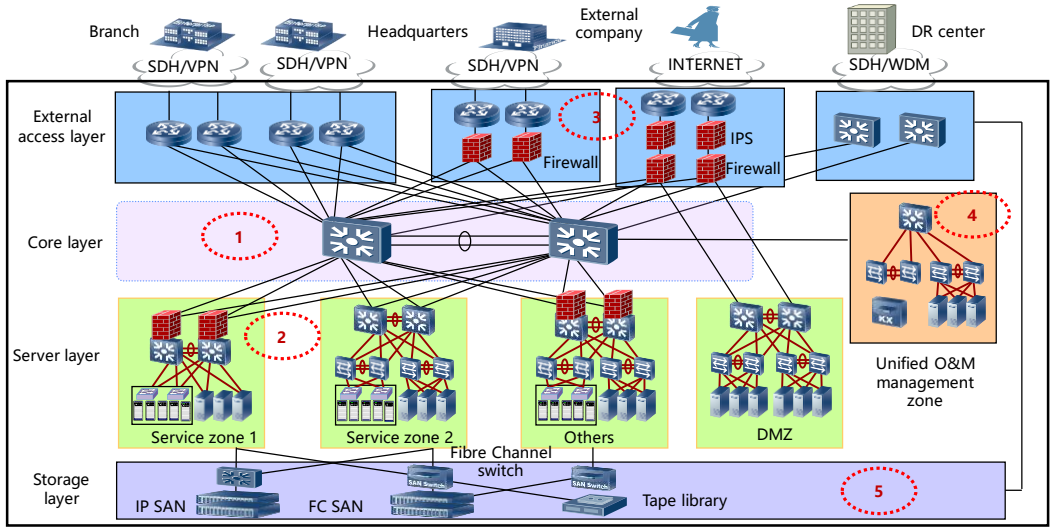


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

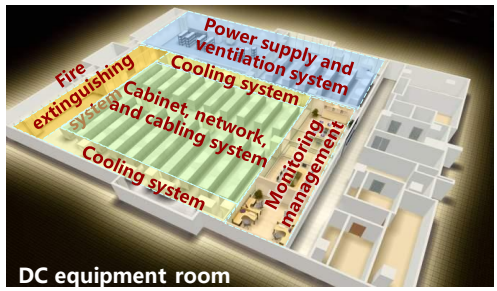
- Two large DCs: active DC in Dongguan and DR center in Nanjing
- Vertical layers: external access layer, network core layer, server access layer, and storage layer
- Horizontal zones: service zone, production zone, test zone, big data zone, and DMZ

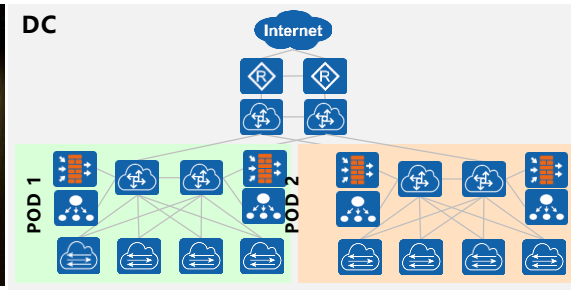
Network Architecture of a Typical DC



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Concepts About Cloud DCs





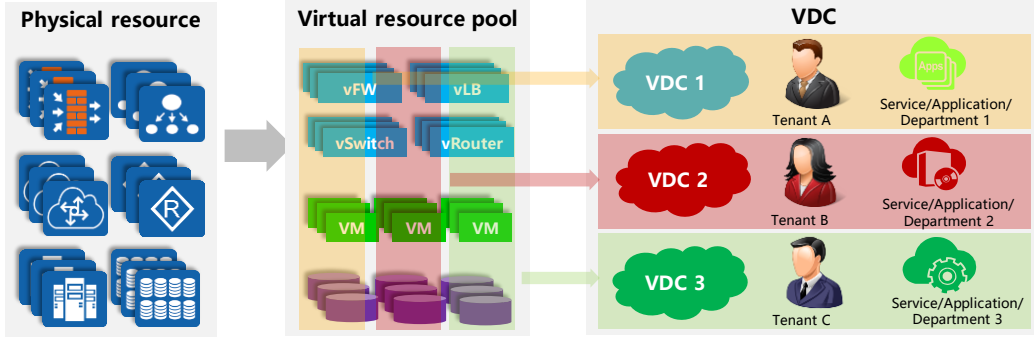
DC: A DC implements centralized data processing, storage, transmission, switching, and management in a physical space. Key devices in a DC include servers, network devices, and storage devices, and necessary DC systems include the power supply, cooling, firefighting, and monitoring systems.

Point of Delivery (POD): To facilitate resource pooling in a DC, a DC is divided into one or more physical PODs. PODs are basic deployment units of DCs. One physical device can belong to only one POD.

Availability Zone (AZ): An AZ indicates a fault isolation area. If some hosts share a power supply and network infrastructure but the infrastructure is faulty, the hosts are unavailable. During planning, AZs can be flexibly mapped to DCs based on site requirements. For example, in a large-scale public cloud, one AZ can contain multiple DCs; in one small-scale private cloud, one DC can contain one or multiple AZs.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

VDC and Tenant



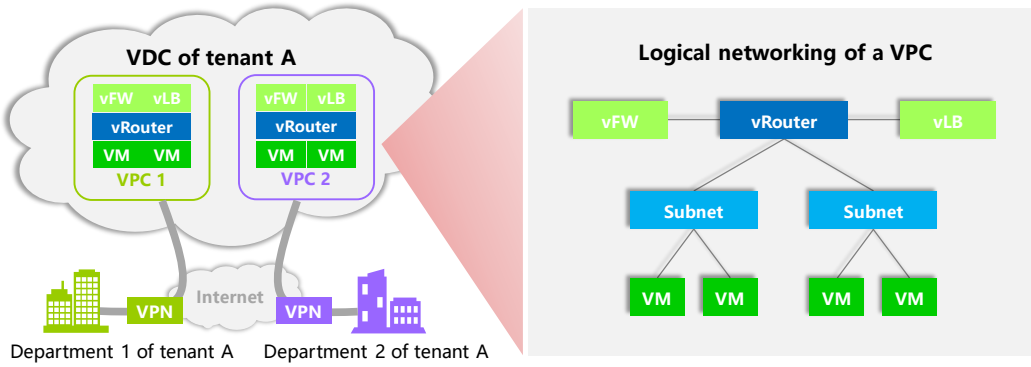
Virtual Data Center (VDC): A VDC is a collection of resources available for an organization. Such resources include computing, storage, and network resources.

Tenant: Tenants are created and allocated by system administrators. A tenant owns and manages a VDC. Different VDCs map to different tenants.

- A VDC represents a physical DC at the virtualization layer.
- In the public cloud scenario, the system administrator can define VDCs and assign the VDCs to tenants. Only the tenant of a VDC can manage resources in the VDC.
- In the private cloud scenario, VDC definition is flexible, and VDCs can be assigned to services, applications, or departments System administrators can use VDCs and resource quotas to manage different services, applications, or departments in an enterprise.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

VPC



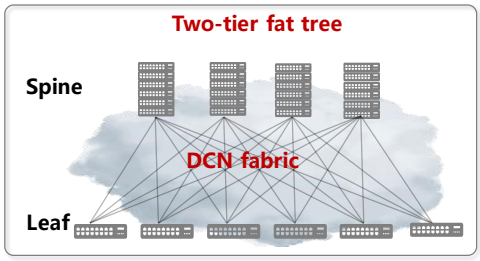
Virtual Private Cloud (VPC): VPCs use resources in VDCs. Each VPC belongs to one VDC, and each VDC can have multiple VPCs. Each VPC is a security zone, serving one service, application, or department. VPCs can provide the following functions:

- Isolated environment: VPCs provide isolated VM and network environments to meet isolation requirements of different applications and departments.
- Diversified services: Each VPC can provide separate services, such as the vFW, vLB, security group, EIP, IPsec VPN, and NAT.
- Flexible networking: VPCs provide multiple networking modes, such as direct networks, routed networks, and internal networks.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

- vRouter
  - vRouter acts as a service subnet gateway for Layer 3 communication between subnets.
- Subnet
  - A subnet is used for isolating Layer 2 broadcast domains and corresponds to one subnet segment.
  - The Layer 3 gateways of different subnets in one VPC are located on the same vRouter.
  - Subnets communicate with each other by default and different subnets are interconnected by default. You can also enable isolation through security group.
- vFW
  - As the VPC edge, vFW can control access from extranet to VPC to protect VPC resources, and provide access service from extranet to VPC.
  - Available features: FW, EIP, SNAT and IPsec VPN.
- vLB
  - vLB is used to provide load balancing between internal servers for external users.
  - A vLB can carry multiple servers. Users can apply for different servers for different services.

DCN Fabric



Term	Description
Fabric	Basic physical network topology of a DC, comprised of a group of spine and leaf nodes.
Spine	A core node on a VXLAN fabric network, which provides high-speed IP forwarding and connects to leaf nodes using high-speed interfaces.
Leaf	An access node on a VXLAN fabric network, which connects various network devices to the VXLAN network.

Fabric features

- **Physical networking**
  - CE series physical and virtual switches as fabric carrying devices
  - No difference between access nodes
  - Flat structure
  - Flexible networking
  - Easy scaling
- **Virtualization**
  - VXLAN-based fabric virtualization
  - Hardware/Hybrid overlay
- **Management**
  - SDN overlay, achieving the automatic deployment of virtual networks
  - Underlay network that supports a Layer 2 or Layer 3 network and the IGP routing protocol
  - BGP EVPN acting as the control plane of overlay
- **Services**
  - Layer 2/Layer 3 mutual communication among access nodes

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

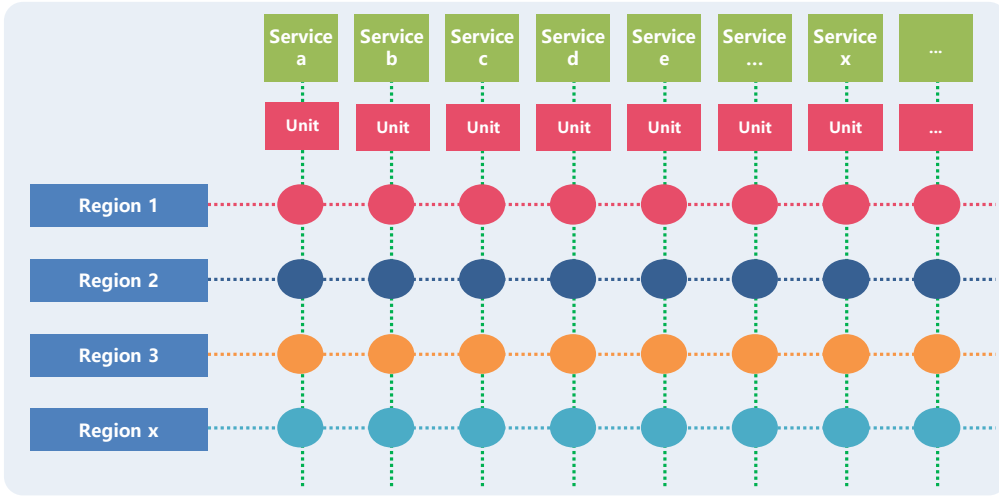
Contents

1. DC Development Course
2. Basic Modules of a DC
3. **Evolution Trends of Cloud DCs**
   - Transformation Trends
   - Huawei Cloud DCs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

IT Transformation on the Top of Enterprises' Agenda



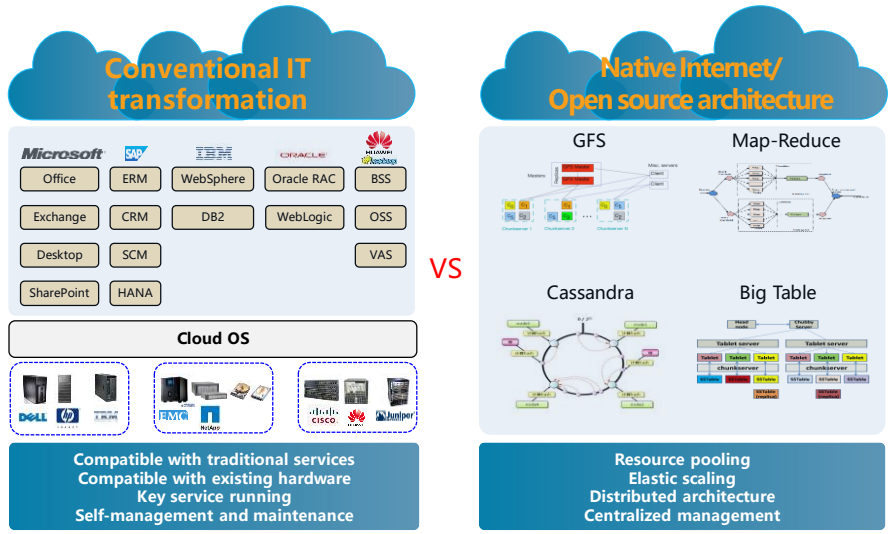
Non-unified deployments create information islands

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- IT advances expose numerous problems such as information islands in DCs, complicated service management, and low service efficiency, necessitating IT transformation in enterprises.

Technical Requirements of Enterprise IT Transformation

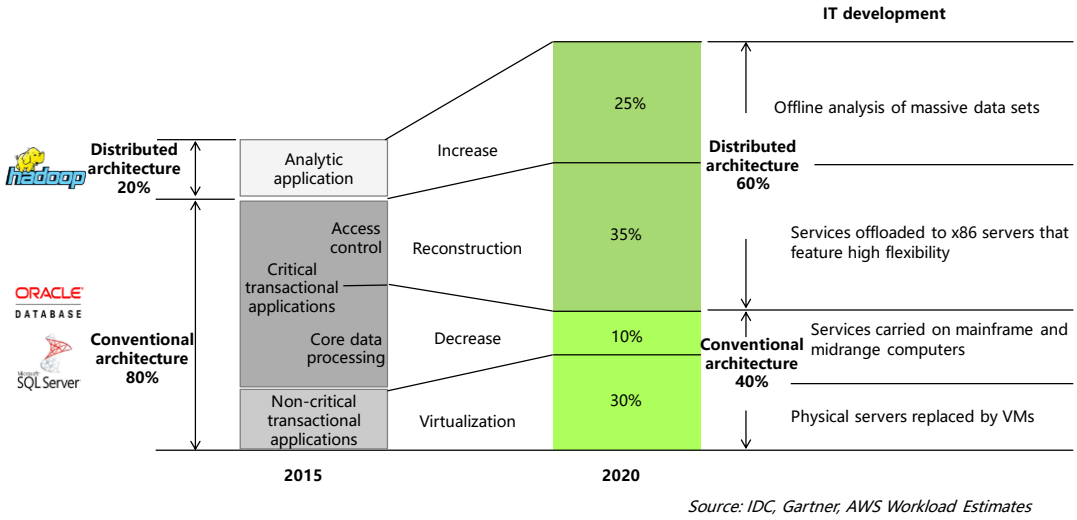


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Internet-based modes bring sweeping changes to every line of business. During enterprise IT transformation, the first concern is how to transform traditional IT systems towards cloud and manage the systems. Second, impact of Internet on business modes and technical framework must also be observed.
- Google File System (GFS) is a dedicated file system designed by Google to store massive sets of search data.
- BigTable is a distributed data storage system designed by Google. It is a non-relational database used to process massive sets of data.
- Cassandra is a third-party component and an open-source distributed NoSQL database system. It was initially developed by Facebook and used to store simple data, such as inbox. It combines Google's BigTable data model and Amazon's Dynamo distributed architecture. It was open to the public in 2008. Featuring excellent scalability, Cassandra is widely used by well-known web 2.0 websites, such as Digg and Twitter and becomes a popular distributed structured data storage solution.

Key Step: IT Systems Move to the Cloud



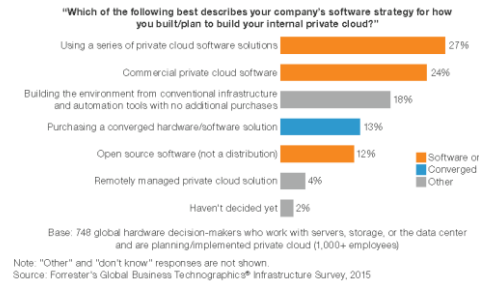
Source: IDC, Gartner, AWS Workload Estimates

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Cloud-based IT architecture and Internet-based move drive conventional IT services to the service-driven cloud-based mode. Private cloud is a key step in the current IT transformation.

Status Quo and Development Trends of Private Cloud

- Network services become mature as IaaS services develop.
- Most private clouds support OpenStack.
- Developers at different levels require increasingly demanding self-help services.
- Private clouds can easily connect to public clouds for unified management.
- Suppliers can provide higher levels of services.



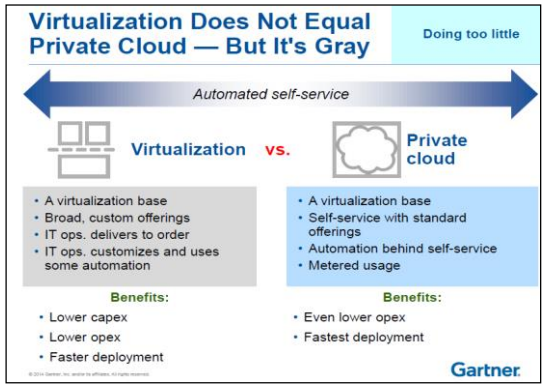


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

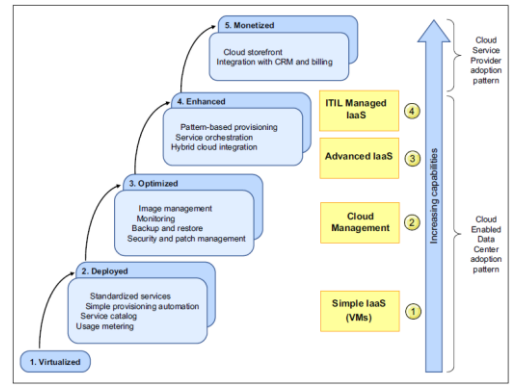
- A cloud-based process can be divided into three phases, namely cloud 1.0 (virtualization-centric), cloud 2.0 (service-centric), and cloud 3.0 (cloud-native architecture).
- With private cloud solutions stepping into cloud 2.0, IaaS+, OpenStack, and services are now oriented to developers and hybrid cloud management. Enterprises are going all out to resolve problems that they face during enterprise IT cloud transformation.

Most Enterprises Are Evolving from Virtualization to Private Cloud

- NIST believes that private cloud has five key characteristics: on-demand self-service, ubiquitous network access, location independent resource pooling, rapid elasticity, and pay per use.



Gartner holds that the path from virtualization to private cloud is gray, not absolutes.

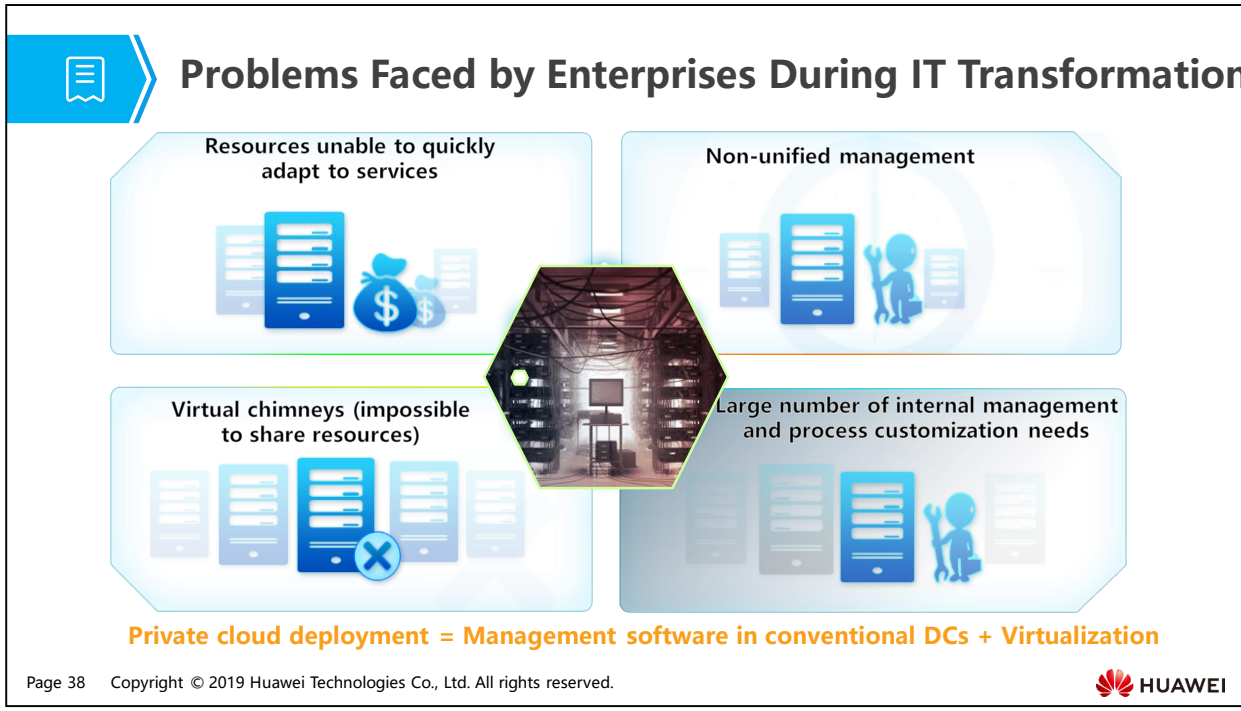


Maturity model of IBM's cloud DC

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Enterprises are evolving from virtualization to private cloud during IT transformation. However, virtualization does not equal private cloud. NIST considers private cloud to have five key characteristics while Gartner and IBM have their own definition.
- The whole transformation process is gray, which adds more uncertainties and barriers on enterprises' way to private cloud.



- According to analysis and surveys oriented to global customers, there are four major problems faced by enterprises during IT cloud transformation. The problems are virtual chimneys, resources unable to quickly match service needs, non-unified management, and large number of internal management and customized process needs.

Engines Behind Enterprise IT Transformation

Efficient



Lowered costs and enhanced management efficiency

Agile



Flexible resource scheduling and rapid service provisioning

Reliable



Rock-solid DC reliability

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Based on Huawei's operations experience and feedback from customers, engines that drive enterprise IT systems to move to the cloud are lowered costs, flexible service provisioning, and rock-solid reliability.

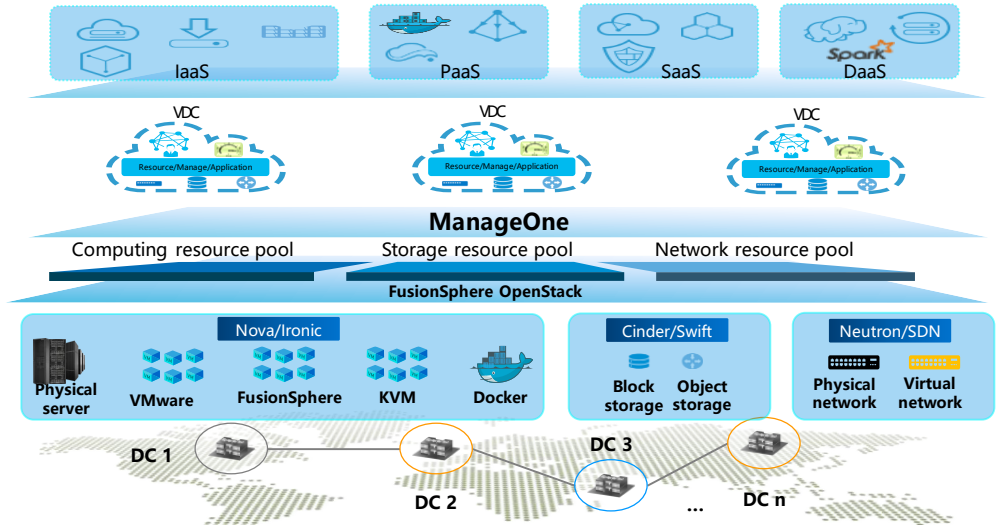
Contents

1. DC Development Course
2. Basic Modules of a DC
3. **Evolution Trends of Cloud DCs**
   - Transformation Trends
   - Huawei Cloud DCs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

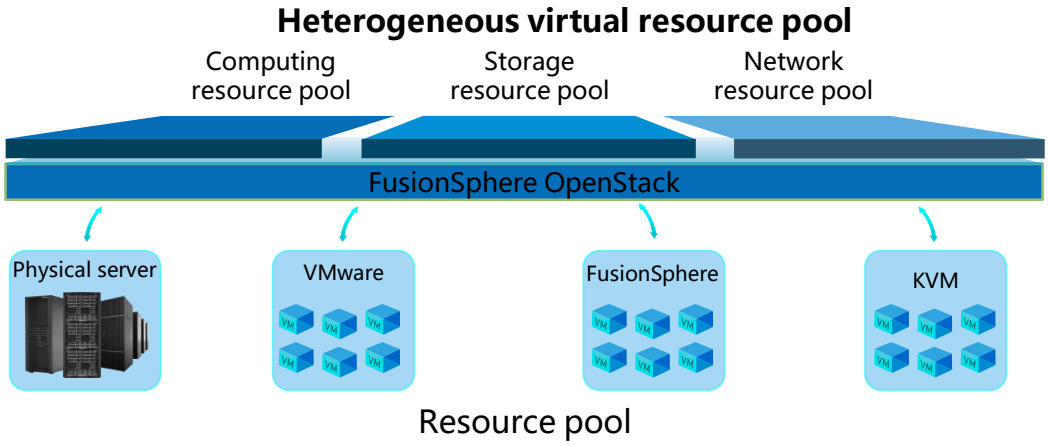
FusionCloud: Future-proof Cloud DC Architecture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Based on customers' requirements, Huawei plans a future-oriented cloud DC architecture (FusionCloud) to provide solutions for cloud services in various service scenarios, including infrastructure, cloud platform, cloud management, and cloud transformation.

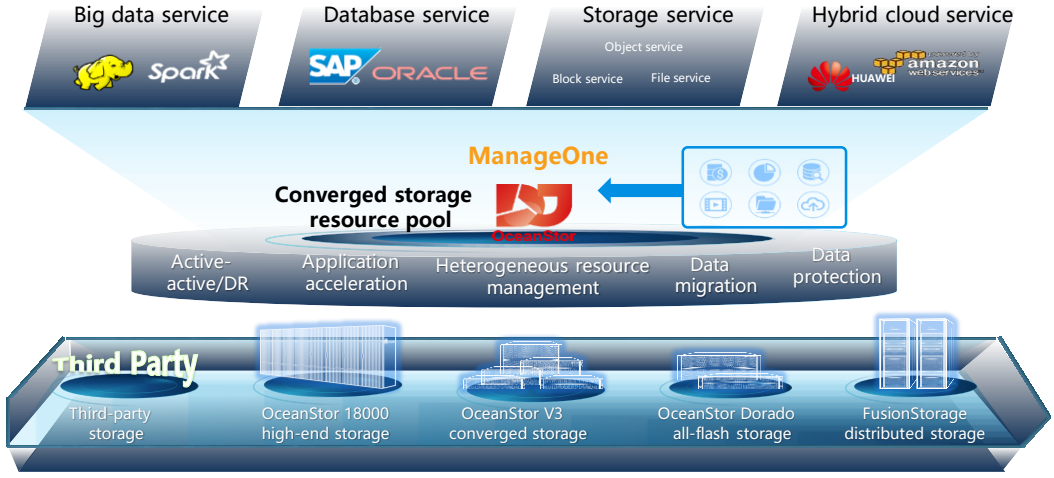
Converged Computing Resource Pool Under OpenStack Architecture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Let us look at efficient management first. Resource convergence is the first step to solve the current virtualization chimney. Huawei was elected as a Gold Member Board Director for the OpenStack 2015 Board and built a system framework based on the OpenStack architecture, which solved the problem of unified computing resource management. The system framework supports central monitoring, allocating, and recycling of physical servers, VMware, FusionSphere, and KVM (in the NFVI scenario) resources.

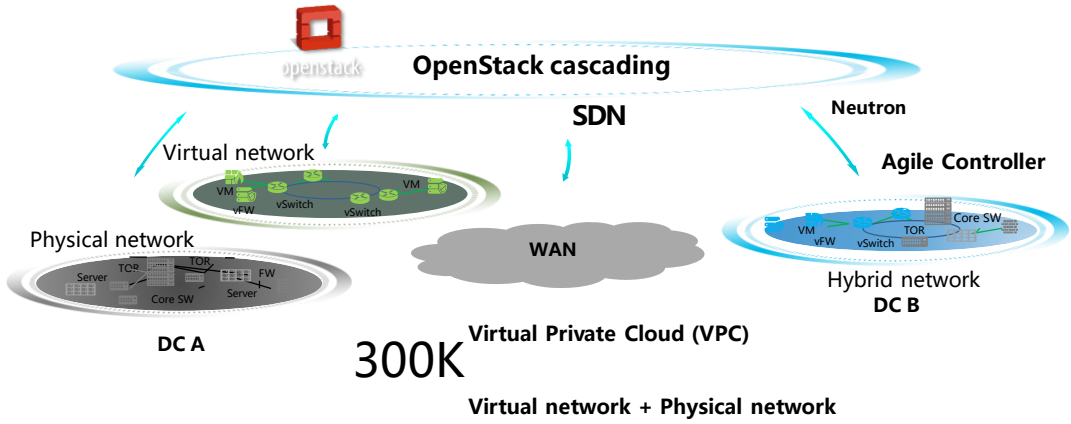
Storage as a Service Based on Heterogeneous Resource Convergence and Application Templates



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Then, I will talk about the convergence of storage resources. Storage devices of different types and from different vendors are integrated, and serviceability requirements on different storage performance indicators are met, realizing storage SLA.

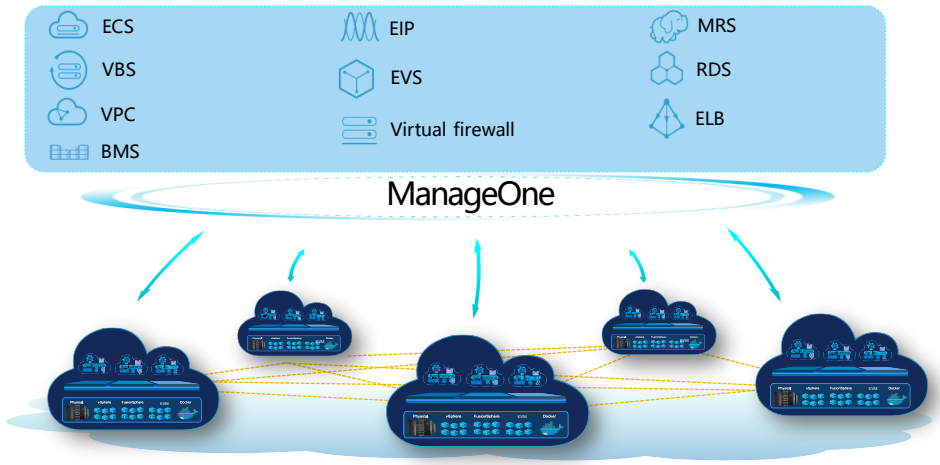
SDN Automatically and Flexibly Deploys Services Across DCs



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Network resources are streamlined with Layer-2 resources through automation, implementing automatic network configuration and cross-domain business deployment. The solutions cover physical networks, virtual networks, and WANs.

Unified Service Access Platform

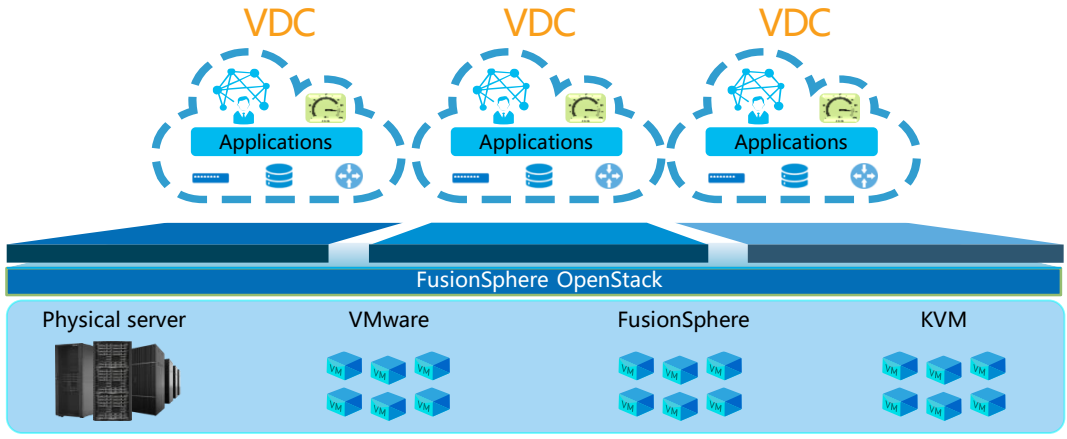


Rapid cloud service access and orchestration enable services to be automatically deployed across DCs within minutes.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The unified cloud service access platform enables cloud services to be quickly accessed and flexibly orchestrated. Distributed cloud DC management allows services to be automatically deployed across DCs within minutes.

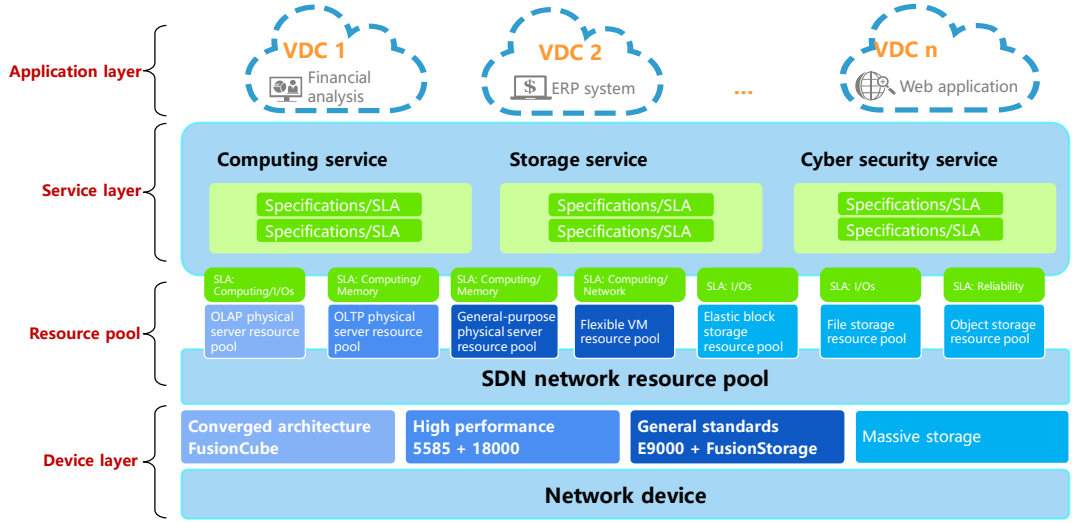
VDCs Based on Organizations and Services



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- VDC can allocate resources across resource pools and match customer organizations, to realize rights- and domain-based cloud resource management. VDC has the self-operation capability that the business department allocates cloud resources and centralizes services while the IT department maintains the cloud platform. In this mode, usage and construction is separated so that services can drive resources in a better way.

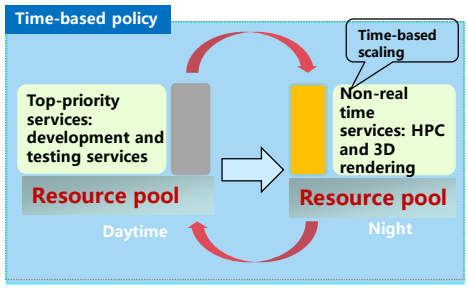
SLA Policy-based Scheduling Meets Diversified Service Needs

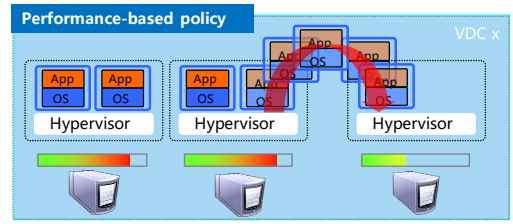


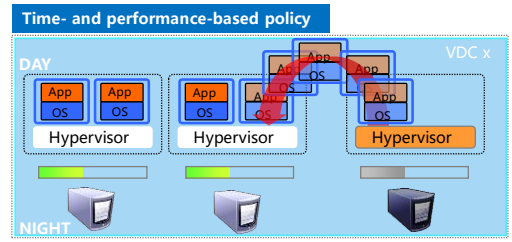
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Different services have varying requirements on resource performance and functions. With services detected, the system can allocate resources to dedicated services as required by different applications in a strategic way by categorizing the SLA levels of these resources.

Service-Driven Auto Scaling and Service Scheduling







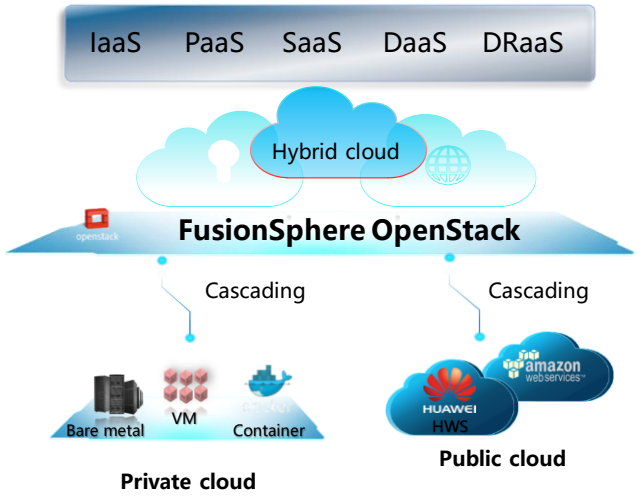
- Cross-domain auto scaling reduces redundant capacity in the local DC by 20%.
- Time- and performance-based policies intelligently and automatically schedule resources, meeting service needs during peak hours.
- Efficient resource utilization and energy saving.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- During service operation, the system can also dynamically respond to the resource performance and quantity requirements, and can initiate appropriate strategies to realize elastic expansion or scheduling of resources when FusionCloud detecting a specific service requirement.

Unified Hybrid Cloud Management



Unified:

- Lightweight hybrid cloud management
- Centralized service catalog

Open:

- Support for OpenStack cascading
- Support for AWS S3 and EC2

Flexible:

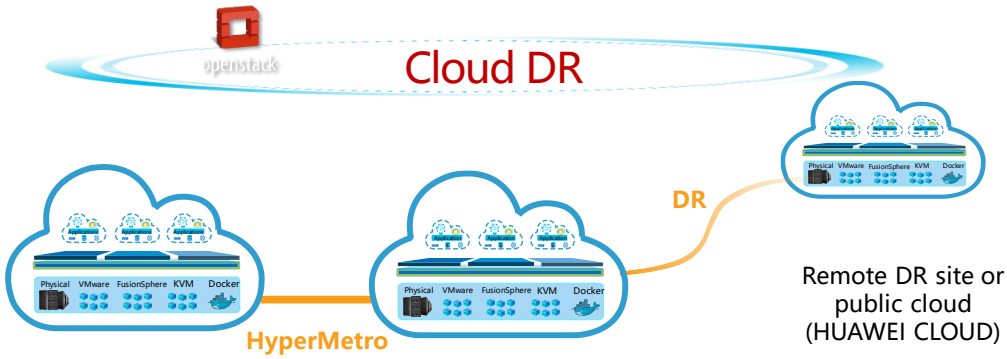
- Cross-cloud auto scaling based on service workloads
- Cross-cloud VM migration

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Unified hybrid cloud management expands the scope of resource scalability and establishes viability solutions for surge in enterprise services and service disaster recovery, ensuring seamless enterprise service expansion.

Cloud DR Solution



- Gateway-free active-active design improves active-active service performance by 30%.
- Visualized and one-click DR reduces service recovery time from 30 minutes to 10 minutes.
- Deduplication backup and public cloud backup.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.



Quiz

1. What are the modularizations of data centers?
2. What is the relationship between the VDC and the VPC in cloud data center ?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What are the modularizations of data centers?
  - A DC is a service-oriented infrastructure. It supports the operation and growth of enterprise businesses. It consists of the following: Secure network architecture, reliable support facilities (equipment rooms, generators, UPSs, and air conditioning systems), integrated servers/application platforms, centralized storage and backup, unified system management platforms, and O&M management organizations and processes oriented to customer services.
- What is the relationship between the VDC and the VPC in cloud data center ?
  - Virtual Data Center (VDC): A VDC is a collection of resources available for an organization. Such resources include computing, storage, and network resources.
  - Virtual Private Cloud (VPC): VPCs use resources in VDCs. Each VPC belongs to one VDC, and each VDC can have multiple VPCs. Each VPC is a security zone, serving one service, application, or department.

Summary

- DC development course
- Basic modules of a DC
- Network structure
- Evolution trends of cloud DCs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



AR Router Product Introduction

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- Huawei AR G3 series routers (AR for short) are the next-generation routing and gateway devices that provide routing, switching, wireless, voice, and security services. The AR G3 series include the AR1200, AR2200, AR3200, AR3600, and AR150&160&200 series routers.
- Huawei AR 500 series routers, including AR531 and AR550 series, are new-generation industrial routing gateways that are developed by Huawei under the IoT background and integrate routing, switching, wireless, and security services.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Describe AR G3 product positioning
  - Describe AR G3 hardware architecture, common cards, and common modules
  - Describe AR G3 data forwarding flows
  - Describe AR G3 usage scenarios

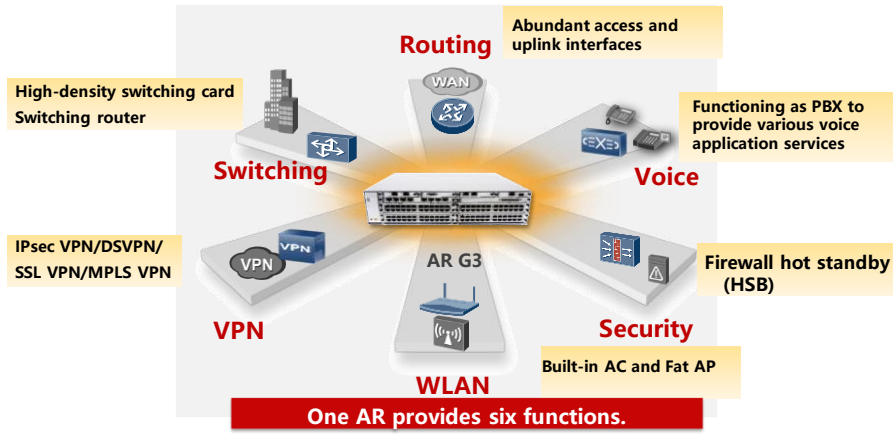
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **AR G3 Positioning**
2. AR G3 Hardware Architecture, Cards, and Modules
3. AR G3 Data Forwarding Process
4. AR G3 Feature Description and Usage Scenarios

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

AR G3 Positioning



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Huawei AR G3 series enterprise routers (AR G3) are next generation routers dedicated for enterprise customers. The AR G3 all-in-one router series integrates multiple services including; routing, switching, 3G, WLAN, voice, and security functions in one device.
- These features combine to deliver industry leading performance and extensibility, meeting customer requirements for a robust, reliable and flexible solution
- for enterprise-grade network deployments. Due to strict adherence to industry standards, the AR G3 router series are easily integrated into existing networks, accelerating multi-service network deployment while preserving existing network infrastructure investments.
- ARs are located between an internal network and a public network. The deployment of various network services over ARs reduces costs in enterprise network construction and long-term operation & maintenance (O&M).

"A-B-C" for AR G3 Routers in Cloud Era

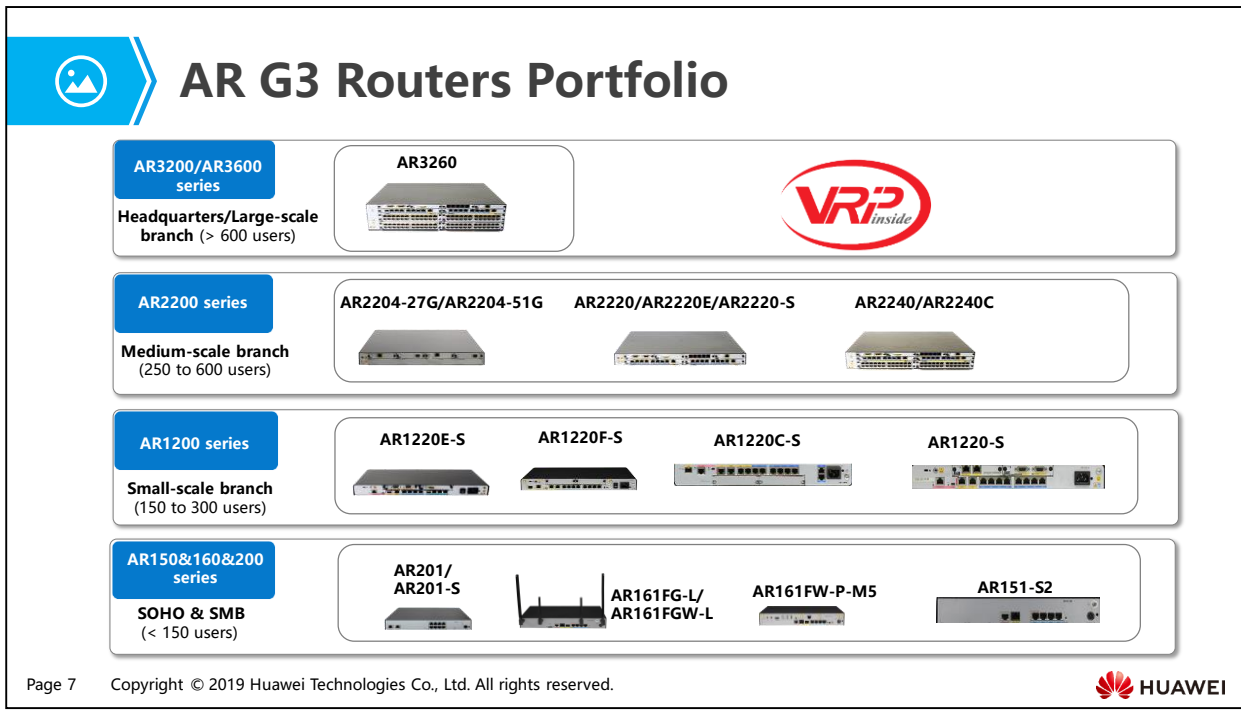


A ***A*pplications in one box**
Reduce the TCO by at least 30%.

B ***B*etter experience**
Ensure service continuity.

C ***C*ooperation platform**
Meet customized application requirements.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.



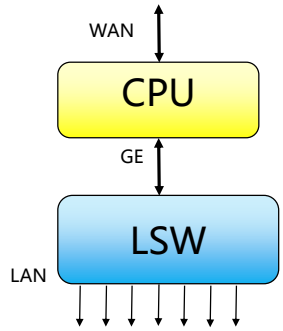
- Those models with “V” stand for supporting voice, Those models with “W” stand for supporting WIFI, Those models with “G” stand for supporting 3G upstream. AR2200 series and AR3200 series support voice function only when equipped with the DSP module.
- To provide voice services for POTS users on AR1200, AR2200 , and AR3200 series routers, 4FXS/1FXO board is required.
- To provide voice services for ISDN users on AR1200, AR2200 , and AR3200 series routers, 2BST board is required.
- For AR3670, voice services doesn’ t supported for now.

Contents

1. AR G3 Positioning
2. **AR G3 Hardware Architecture, Cards, and Modules**
   - Logical Architecture
   - SRU and Interface Card
   - Power Module and Fan Module
   - Common Interface and Cable
3. AR G3 Data Forwarding Process
4. AR G3 Feature Description and Usage Scenarios

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

AR 150&160&200 Logical Architecture



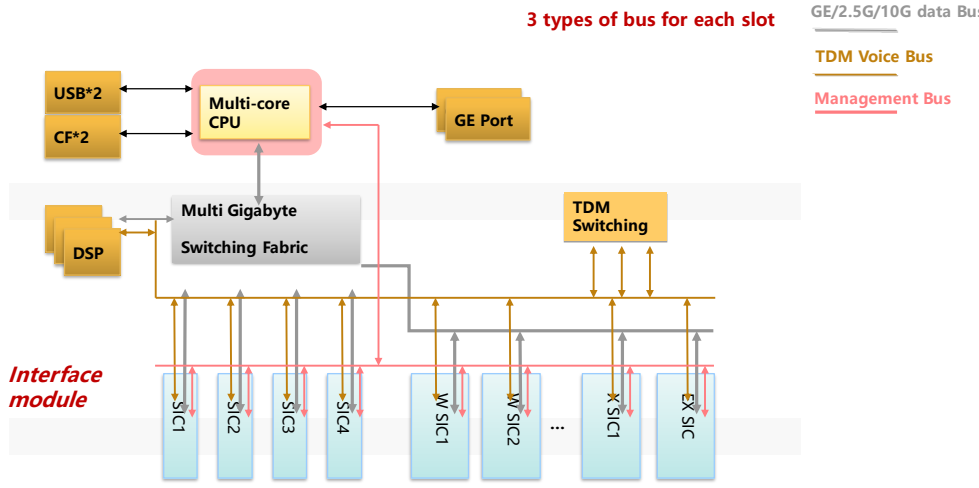
The CPU is responsible for complex calculation, it is directly connected to the WAN interface, and to the LSW with a GE bus.

The LSW (Switching module) is responsible for forwarding the L2 and L3 Ethernet traffics.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The AR150 and AR200 share the same simple logical architecture, which is consist of CPU and LSW(Switching module).
- The CPU is responsible for complex calculation, it is directly connected to the WAN interface, and to the LSW with a GE bus.
- LSW is responsible for forwarding the L2 and L3 Ethernet traffics.

AR 1200/2200/3200 Logical Architecture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The AR 1200/2200/3200 has more complex logical architecture.
- Bus interface is reserved for the pluggable cards on each slot.
- Different to the AR150&160&200, the switching fabric is added to AR1200/2200/3200, which greatly enhances the performance.

Contents

1. AR G3 Positioning
2. **AR G3 Hardware Architecture, Cards, and Modules**
   - Logical Architecture
   - SRU and Interface Card
   - Power Module and Fan Module
   - Common Interface and Cable
3. AR G3 Data Forwarding Process
4. AR G3 Feature Description and Usage Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

SRU40/SRU60/SRU80/SRU100/SRU200



SRU40/SRU60/SRU80/SRU100



SRU200

SRU specifications:

Card Name	Performance	Function
SRU40	600 MHZ CPU with 8 kernels	Does not support traffic management (TM).
SRU60	600 MHZ CPU with 8 kernels	Does not support TM.
SRU80	750 MHZ CPU with 12 kernels	Supports TM.
SRU100	750 MHZ CPU with 12 kernels	Does not support TM.
SRU200	1.2 GHz CPU with 32 kernels	Supports TM, two 10GE SFP+ interfaces, and four GE combo interfaces.

When one SRU is installed, it can be inserted into:
- Slot 11 on the AR2240
- Slot 15 on the AR3260

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The SRU40, SRU60, and SRU80 panels are identical except for having different silkscreen.
- The SRU must be installed on the AR2240 and AR3260. You can install one SRU. Two SRUs can be installed on the router.

SRUX5



SRUX5

SRU specifications:

Card Name	Performance	Function
SRUX5	2.2 GHz CPU with 6 kernels	Supports TM, and does not support the voice function.

The SRUX5 can be installed only on the AR3670.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SRU40C/SRU100E/SRU200E



SRU40C



SRU100E/200E

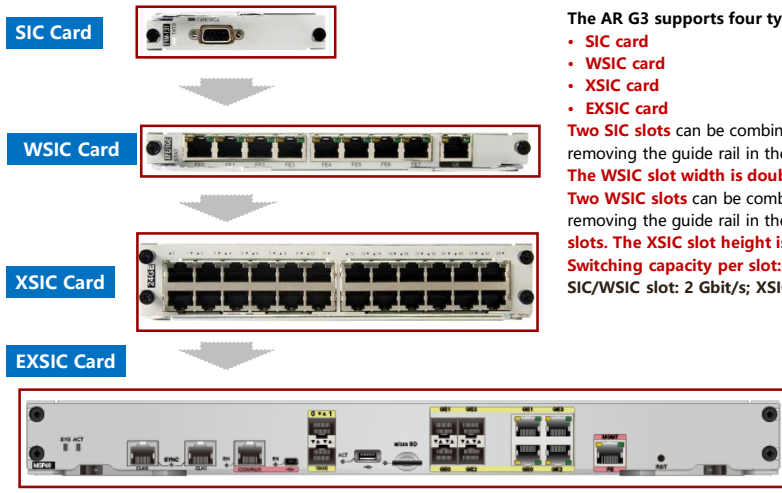
SRU specifications:

Card Name	Performance	Function
SRU40C	1.2 GHz CPU with 6 kernels	Supports TM, four GE optical interfaces, two GE combo interfaces, and four GE electrical interfaces.
SRU100E	1.2 GHz CPU with 12 kernels	Supports TM, two GE SFP interfaces, and four GE combo interfaces.
SRU200E	1.2 GHz CPU with 12 kernels	Supports TM, two 10GE SFP+ interfaces, and four GE combo interfaces.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- SRU40C only supported by AR2240C.
- SRU100E/SRU200E only supported by AR2240 and AR3260.

Interface Cards Supported by the AR G3



The AR G3 supports four types of interface cards:

- **SIC card**
- **WSIC card**
- **XSIC card**
- **EXSIC card**

Two SIC slots can be combined into **one WSIC slot** by removing the guide rail in the middle of the **two SIC slots. The WSIC slot width is doubled.**

Two WSIC slots can be combined into **one XSIC slot** by removing the guide rail in the middle of the **two WSIC slots. The XSIC slot height is doubled.**

Switching capacity per slot:

SIC/WSIC slot: 2 Gbit/s; XSIC/EXSIC slot: 20 Gbit/s

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

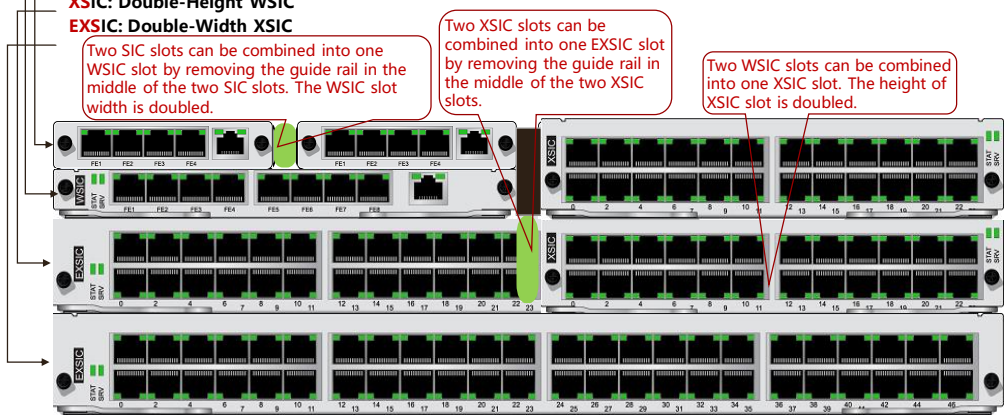
HUAWEI

- Two SIC slots can be combined into one WSIC slot by removing the guide rail.
- The two SIC slots and the WSIC slot below them can be combined into one XSIC slot by removing the guide rail.
- Two XSIC slots can be combined into one EXSIC slot by removing the guide rail.
- Slots can be combined into one, but one slot cannot be divided into multiple slots.
- After two slots are combined into one, the slot ID is the larger one between the original two slots.
- In V200R002C00, a WSIC card can be inserted into an XSIC slot with a special component. The WSIC card is in the lower side of the slot and uses the XSIC slot ID as its own slot ID.
- The AR2201-48FE and AR2202-48FE have no slot for pluggable subcards, so they do not support subcards.

Combination Mode of AR G3 Cards

- Various service cards are combined based on the standard design. The slot resources are fully used, protecting investments.

SIC: Smart Interface card
WSIC: Double-Width SIC
XSIC: Double-Height WSIC
EXSIC: Double-Width XSIC



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Slots can be combined into one, but one slot cannot be divided into multiple slots.
- The number of the new merged slot equals to the larger one of the former slots.

AR G3 WAN Interface Modules (1)

E1 SICs:



1E1/T1-M



2E1/T1-M

1/2-port-channelized E1/T1/PRI/VE1 multifunctional interface modules
These modules receive, send, and process data flows on E1 interfaces. They can be configured with VE1 interfaces to transmit voice services, data services, and fax signals.



1E1-F



2E1-F

1/2-port-channelized E1/T1 multifunctional interface modules

Ethernet SICs:



2FE

2-port 10M/100M Ethernet electrical interface module



1GEC

1-port GE Combo Ethernet interface module

Synchronous/Asynchronous SIC:



1SA

1-port enhanced synchronous/asynchronous serial interface module
Supporting V.24, V.35, X.21, RS449, and RS530 protocols.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- 1/2: indicates one or two interfaces.
- E1: indicates E1 interfaces.
- T1: indicates T1 interfaces.
- M: indicates multiflex trunks.
- PRI: indicate ISDN primary rate interfaces.
- VE1: indicates voice E1 interfaces.

AR G3 WAN Interface Modules (2)

ADSL2+ SICs:



1ADSL-A/M — 1-port ADSL over POTS interface module supporting ANNEX A/M



1ADSL-B/M — 1-port ADSL over ISDN interface module supporting ANNEX B

4G.SHDSL SIC:



4G.SHDSL — 1-port-4G.SHDSL WAN interface module

PON SIC:



1PON — 1-port-EPON/GPON interface module

CPOS SIC:



1CPOS — 1-port 155 Mbit/s CPOS interface module

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- 1CPOS-155M (1-Port Channelized POS Interface Card):
  - 1: indicates one interface. C: indicates channelized interface.
  - POS: is short for Packet Over SDH/SONET.
  - 155M: indicates a rate of 155 Mbit/s.

3G/LTE Interface Modules

- AR G3 V200R005 supports 3 types of 3G/LTE interface modules, all SIC (similar in appearance)
  - 3G-HSPA+7
  - 3G-EVDO
  - 1LTE-L



- Besides, a USB 3G/LTE modem can be installed to the fixed USB port to support 3G/LTE.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The 3G-HSPA+7 is a 3G access SIC card. It can function as the primary or backup link of an enterprise to connect to the Internet and transmit voice, video, and data services.
- The 3G-EVDO is a CDMA2000 network access module, It is installed in a SIC slot to provide high-speed wireless data transmission, enabling enterprise users to connect to CDMA2000 networks.
- The 1LTE-L is a wireless high-speed WAN access module, It is installed in a SIC slot to provide high-speed wireless data transmission, enabling enterprise users to connect to Long Term Evolution (LTE) networks.
- Only a list of USB 3G Modems are supported, you can contact Huawei TAC to get the latest list.

AR G3 Ethernet Switching Interface Modules

WSIC



8FE/1GE

8-port 100 Mbit/s+1-port 1000 Mbit/s Layer 2/Layer 3 Ethernet electrical interface module (RJ45) The module supports line-speed transmission of Layer 2 and Layer 3 packets in full-duplex or half-duplex mode. Each interface can work in auto-sensing mode.

XSIC



24GE

24-port 1000 Mbit/s Layer 2/Layer 3 Ethernet electrical interface module (RJ45) The module supports transmission of Layer 2 and Layer 3 packets in full-duplex or half-duplex mode. Each interface can work in auto-sensing mode.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The 8FE1GE can be installed in the WSIC slots of the AR1200, AR2200, and AR3260. On the AR1200 and AR2204, two SIC slots are combined into one WSIC slot.
- The 24GE can be installed into the XSIC slot on the AR2220, AR2240, and AR3260. On the AR2220, two WSIC slots are combined into one XSIC slot.

AR G3 Voice Interface Modules

Analog voice SIC



4FXS1FXO

4-FXS-port and 1-FXO-port access interface module, transmitting voice services on the Datacom network

ISDN S/T voice SICs



1BST

1-port ISDN access module, receiving, sending, and processing ISDN data flows



2BST

2-port ISDN access module, receiving, compressing, and decompressing ISDN voice signals on the Datacom network

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- An FXS interface is a simulated subscriber line interface and provides access to AT0 loop trunk of the analog phone, fax, and telephone exchange.
- An FXO interface is a loop trunk interface and provides access to the telephone exchange by using regular subscriber lines.
- The 2BST is the ISDN module on the AR routers and provides two ISDN S/T interfaces, which transmit voice service.
- The 2BST implements the ISDN BRI function and provides the bandwidth of two B channels and one D channel:
  - B channel: provides 64 kbit/s bandwidth and transmits voice service.
  - D channel: is a signaling channel and provides 16 kbit/s bandwidth.
  - The total bandwidth of two B channels and one D channel is 144 kbit/s.
- The S/T interface on the 2BST provides a rate of 192 kbit/s, including 144 kbit/s for data transmission and 48 kbit/s for maintenance information transmission.

Contents

1. AR G3 Positioning
2. **AR G3 Hardware Architecture, Cards, and Modules**
   - Logical Architecture
   - SRU and Interface Card
     - Power Module and Fan Module
   - Common Interface and Cable
3. AR G3 Data Forwarding Process
4. AR G3 Feature Description and Usage Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

AR G3 Power Modules

Product	Power Module	Module Attribute	Installation	Maintenance
AR1200	54 W AC power module in an open rack	AC power module fixed in the chassis	It is fixed in the chassis, so it does not need to be connected to the chassis using cables.	It is maintained together with the chassis.
AR2220	PWR150A	AC power module	It is inserted into the power supply slot. The device supports only one power module.	It is maintained together with the chassis.
AR2240	PWR350A	AC power module	It is inserted into the power supply slot. The device supports two power modules.	It uses front-access design and is hot swappable.
AR3260	PWR350A	AC power module	It is inserted into the power supply slot. The device supports two power modules.	It uses front-access design and is hot swappable.
AR3670	PWR700B	AC power module	It is inserted into the power supply slot. The device supports two power modules.	It uses front-access design and is hot swappable.

Redundant power module configuration:



PWR350A



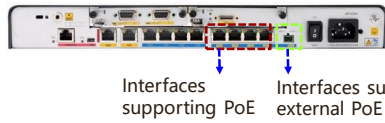
PWR700B

The AR2240/3260/3670 support power module redundancy. The AR2240/3260/3670 is equipped with a default AC power module. When power module redundancy is required, install another power module.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Only power modules of the same power can be used on an AR router.
- Powering off the AR router before removing and reinstalling power modules.
- To power off an AR router, power off all its power modules.

AR1200 PoE Power Module



The AR1200 provides 8FE LAN interfaces, among which 4FE interfaces support PoE.
To enable PoE, connect external PoE power modules through the PoE power interface on the AR1220.

PoE power module:



- 90 V AC to 264 V AC input power
- 48 V/100 W output power when no fan is used

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

AR G3 Fan Modules

The AR2200/AR3200 uses fans for heat dissipation and the air channel is left-to-right.
Air circulation through the chassis:



Fan module panel of the AR2240:



Fan module panel of the AR3260/AR3600:



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- If a single fan failed, the device will be overheated and its performance is then affected. When this occurs, replace the entire fan module immediately.

Contents

1. AR G3 Positioning
2. **AR G3 Hardware Architecture, Cards, and Modules**
   - Key Specifications and Logical Architecture
   - SRU and Interface Card
   - Power Module and Fan Module
   - Common Interface and Cable
3. AR G3 Data Forwarding Process
4. AR G3 Feature Description and Usage Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

AR G3 Ethernet Electrical Interface

- Ethernet Electrical Interface
  - Connector: RJ45
  - Cable: UTP/STP
  - Rate: 10M/100M/1G
  - Link protocol: Ethernet
  - Typical cards:
    - 8FE1GE
    - 24GE
    - 2FE



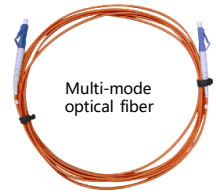
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

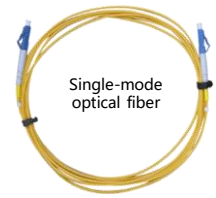
HUAWEI

- A network cable subtends devices, enables a device to communication with other network devices, and allows users to locally or remotely maintain the device.

AR G3 Ethernet Optical Interface

- Ethernet Optical Interface
  - Connector: LC or PC
  - Cable: single-mode or multi-mode cable
  - Rate: 100M/1G
  - Link protocol: Ethernet
  - Typical cards:
    - SRU
    - 1GEC







SFP optical module



LC/PC connector

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The appearances of the single-mode optical fiber and the multimode optical fiber are the same, but their colors are different. The single-mode optical fiber is yellow, and the multi-mode optical fiber is orange.
- The optical transmitting module of the multi-transverse mode is connected to the multimode fiber.
- The optical transmitting module of the single-longitudinal mode or multi-longitudinal mode is connected to the single mode fiber.

AR G3 E1/T1 Interface

- ET/T1 Interface
  - Connector: DB9
  - Cable: trunk
  - Rate: 2.048M/1.544M
  - Link protocol: PPP or HDLC
  - Typical cards:
    - 1E1T1-M
    - 2E1T1-M



E1 75-ohm unbalanced coaxial cable



E1 120-ohm or T1 100-ohm balanced twisted pair cable

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- E1 trunk cables are classified into 75-ohm unbalanced coaxial cables and 120-ohm balanced twisted pair cables. The connectors of the cables are as follows:
- 75-ohm unbalanced coaxial cable (DB9 to BNC):
  - One end provides a DB9 connector.
  - The other end provides two BNC connectors.
- 120-ohm balanced twisted pair cable (DB9 to RJ45):
  - One end provides a DB9 connector.
  - The other end provides an RJ45 connector.
- A T1 trunk cable is a 100-ohm balanced twisted pair cable. Its appearance is the same as the appearance of an E1 120-ohm balanced twisted pair cable.

AR G3 xDSL Interface

- xDSL Interface
  - Connector: RJ11 or RJ45 (4G.SHDSL)
  - Cable: PSTN cable
  - Rate: 5.696M/8M/12M/24M
  - Signal processing method: Analog Modulation and Demodulation
  - Typical cards:
    - ADSL-A/M
    - ADSL-B
    - 4G.SHDSL



4G.SHDSL cable

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The connectors of a 4G.SHDSL cable are as follows:
  - An RJ45 connector on the local end
  - Four RJ11 connector on the network side

AR G3 Synchronous/Asynchronous Serial Interface

- Synchronous/Asynchronous Serial Interf
  - Connector: DB28
  - Cable: DCE/DTE
  - Rate 64 kbit/s to 2.048 Mbit/s
  - Data processing: Protocol/Flow
  - Typical cards:
    - 1SA



V.35 DCE cable



V.35 DTE cable

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- A console cable connects the console port of the device to the serial port of an operation terminal to transmit configuration data. A shielded cable or an unshielded cable can be used according to the onsite situation.
- A console cable connects the device and terminal as follows:
- The 8-pin RJ45 connector is inserted into the console port of the device.
- The DB9 male connector is connected to an operation terminal, which is usually a PC.

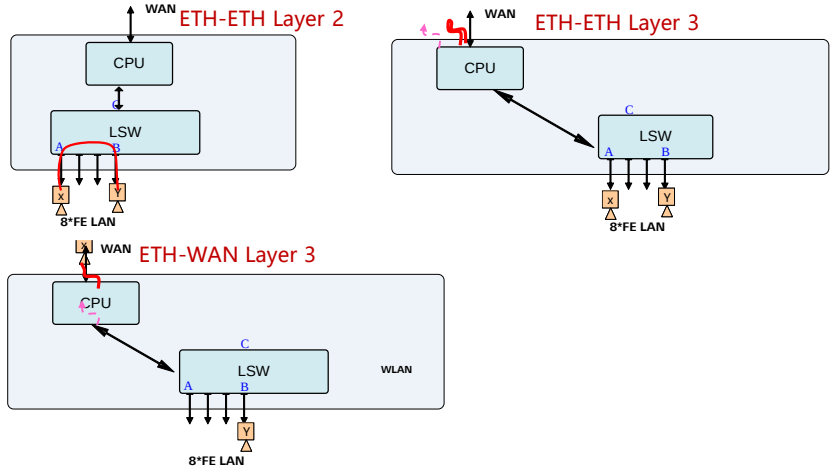
Contents

1. AR G3 Positioning
2. AR G3 Hardware Architecture, Cards, and Modules
3. **AR G3 Data Forwarding Process**
4. AR G3 Feature Description and Usage Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

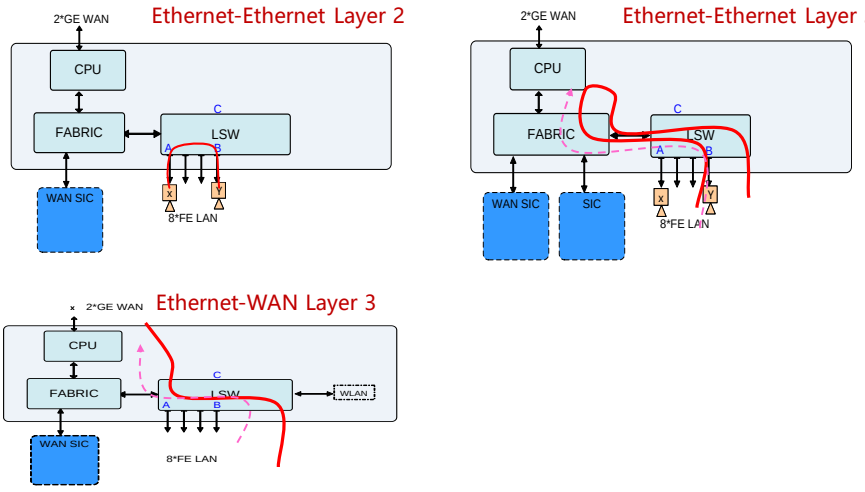
Data Forwarding Process: AR150&160&200



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- L2 traffics between LAN interfaces are forwarded through LSW.
- L3 traffics between LAN interfaces, or between LAN and WAN interfaces are forwarded through both LSW and CPU.

Data Forwarding Process: AR1220

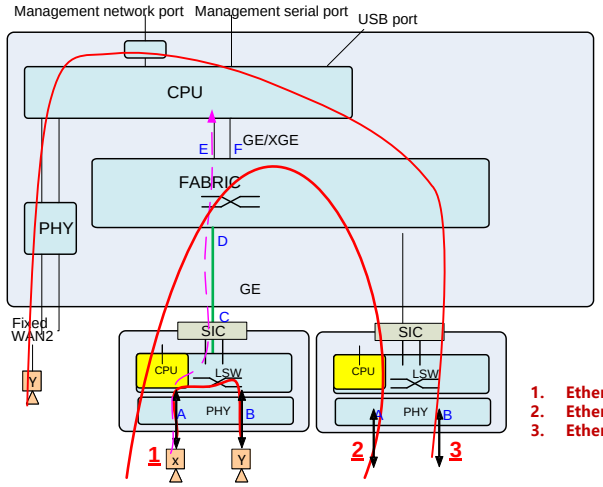


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- L2 traffics between LAN interfaces are forwarded through LSW.
- L3 traffics between LAN interfaces, or between LAN and WAN interfaces are forwarded through LSW, switching fabric and CPU.

Data Forwarding Process: AR2200/AR3200/AR3600



1. Ethernet LAN-Ethernet LAN Layer 2 (in a subcard)
2. Ethernet LAN-Ethernet LAN Layer 3 (in a subcard)
3. Ethernet LAN-fixed Ethernet WAN2 Layer 3 (in a subcard)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Ethernet LAN-Ethernet LAN Layer 2 (in a subcard): only through LSW
- Ethernet LAN-Ethernet LAN Layer 3 (in a subcard): through LSW and Fabric
- Ethernet LAN-fixed Ethernet WAN2 Layer 3 (in a subcard): through LSW, Fabric and CPU

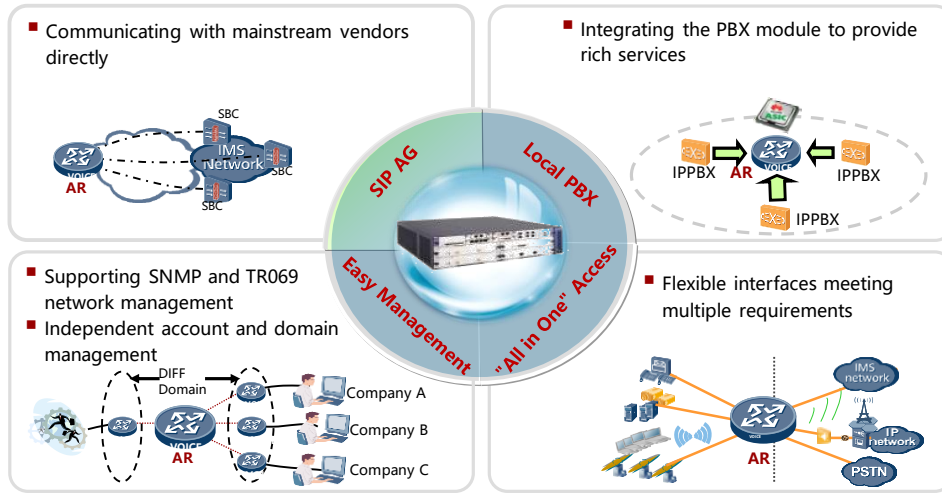
Contents

1. AR G3 Positioning
2. AR G3 Hardware Architecture, Cards, and Modules
3. AR G3 Data Forwarding Process
4. **AR G3 Feature Description and Usage Scenarios**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

AR G3 Voice Solution

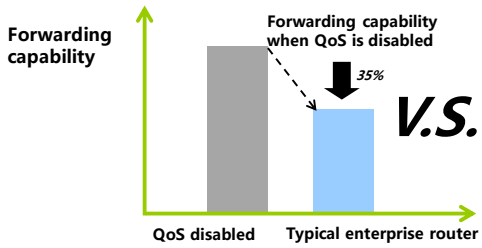


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

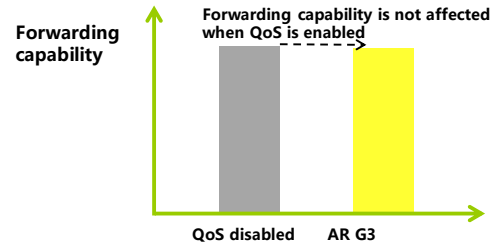
HUAWEI

- Enterprise-class voice communication is flexible and efficient, as the AR voice features integrate with data networks.
- Basic voice functions are provided by the built-in PBX, SIP server, and SIP access gateway
- Value-added voice services include multi-party communication, IVR automatic connection, ring-backtone, parallel ringing, sequential ringing, one number link you (ONLY), bill management, and subscriber management.
- Intelligent call routing enables exceptional voice service reliability.
- The AR routers can be connected with the NGN/IMS/PBX/terminal of major vendors.
- The Quality of Experience (QoE) feature monitors voice service quality in real time.
- Jitter buffer, echo cancellation, and packet loss compensation combine to deliver a superior user experience

AR G3 QoS Features: Hardware-based QoS



QoS is handled using software. When the system needs to provide differentiated services, the forwarding capability is degraded by ***30% to 40%.***



The AR G3 uses hardware-based QoS technologies, ensuring service quality without affecting the forwarding capability.

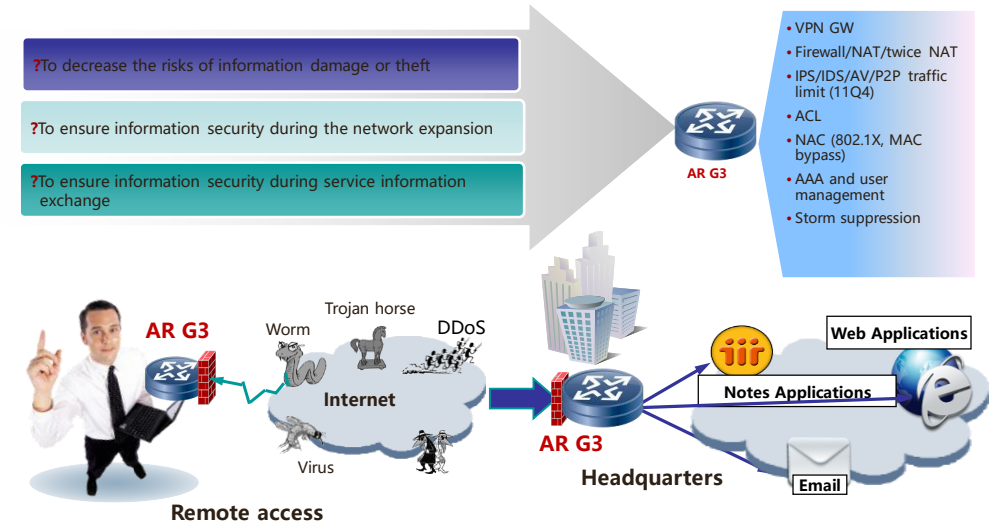


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Only SRU80 with TM card supports Hardware-based QoS, all model can support H – QoS.

AR G3 Security Feature: Comprehensive Security



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- While delivering enterprise-class network services, the AR router provides robust network security. Comprehensive security solutions include user access control, packet detection, and active attack defense.
- The built-in firewall is the first line of defense.
- Port authentication technologies include 802.1x authentication, MAC address authentication, and
- portal authentication.
- User and device authentication methods include RADIUS and HWTACACS.
- VPN technologies include IPsec VPN, GRE VPN, DSVPN, L2TP VPN and SSL VPN.

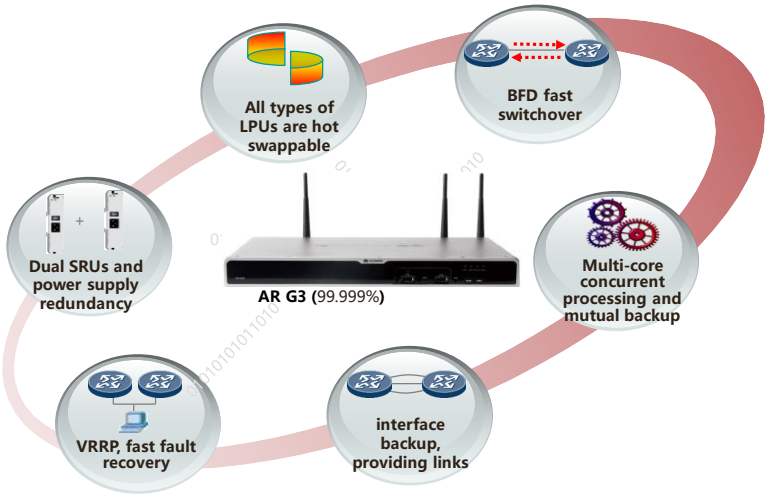
AR G3 Security Feature: IAE

- Intelligence Awareness Engine (IAE) is the collection of all in one security process framework and a series of security features or components, with a variety of security database and security intelligence center real time linkage, IAE is the security services program which makes a variety of products easy to expansion, integration and rapid release content security services. IAE based on NGE framework, also called next-generation high-performance firewall.
- In simple terms: IAE = security framework (stream + proxy) + security feature components + security database + security intelligence center real time linkage.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

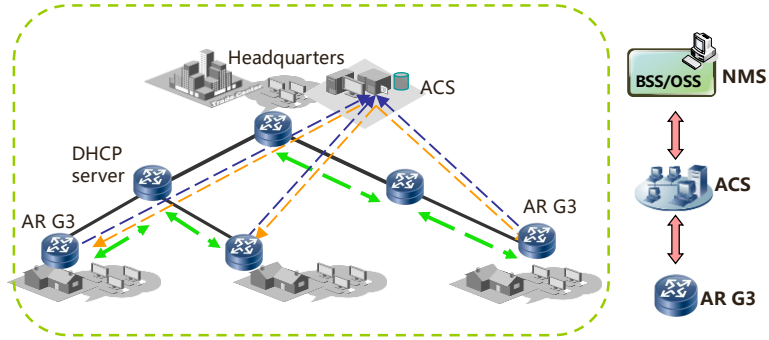
- IAE focus on solving the following new requirements of the Internet:
  - Cloud center and large enterprise network need much powerful IPS or UTM.
  - Strong demand for application identification, control and visualization.
  - Demand of integrated IPS, automatic detection and action
  - Demand of intelligent linkage (IPS, AV, URL, etc.)
  - Demand of content security virtualization (IPS, AV, URL, etc.)
  - Demand of multi content security platform sharing
- IAE core competitiveness is reflected as follows:
  - High performance
  - Merge the detection requirements of each services, the application layer only analyze and extract once. Integration strategy merging and execution, Less than 50% performance decline after enable the full threat protection feature.
  - Multi product rapid release and services expand capabilities.
  - Focus on professional content security capabilities, to achieve flexible and efficient services processing with reasonable architecture , to meet the various needs of the product.
  - Independent evolution and development of core security capability
  - SA library /IPS library /URL library /AV / reputation library, those core security capabilities can be independent development and evolution, decoupling from specific product.

AR G3 Reliability



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

AR G3 Management Feature: Free of Manual Configuration

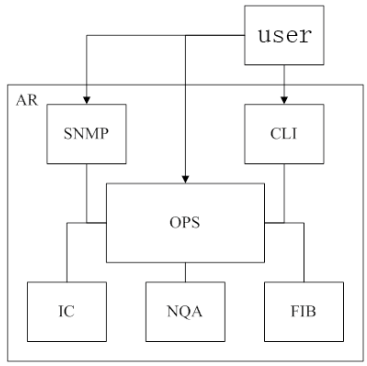


Advantages and Values

- The AR G3 supports TR069, and implements automatic configuration of "Zero Touch" by working with the ACS and NMS.
- The batch configuration and automatic upgrade save deployment and maintenance costs.
- The network status is automatically monitored using TR069, implementing simple O&M.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

AR G3 Management Feature: OPS

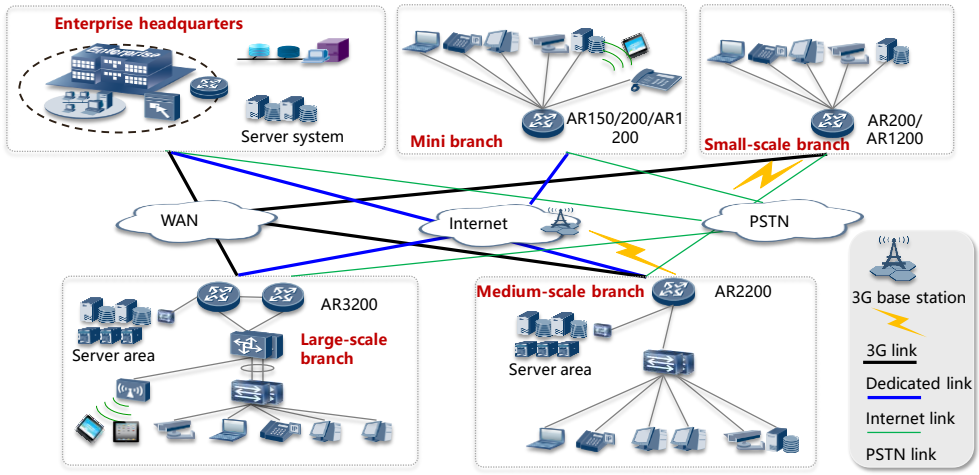


- The open programmability system (OPS) is an open platform that provides Application Programming Interfaces (APIs) to achieve programmability, allowing third-party applications to run on the platform.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OPS is the new function that being released after V2R7C00. AR1220, AR1220V, AR1220W, AR1220VW and AR1220L do not support OPS
- OPS is mainly used in following circumstance.
- Intelligence diagnosis
  - Threshold alarm: diagnostic whether equipment, memory or CPU occupancy rate is over threshold or not.
  - Interface diagnostic: diagnostic whether the interface is normal or not.
  - Route diagnostic and analysis: diagnostic whether the route of the equipment is normal or not.
  - Monitor key route change: Create notification after key routing changes.
  - Equipment diagnostic and query: diagnostic equipment is normal or not.
  - Interface flow monitoring: diagnostic whether the flow of the interface is normal or not.
- Intelligence configuration
  - To enter the configuration mode, automatically backup the configure: before starting the configuration, automatically backup the current configuration to the local and remote SSH server.
  - Configuration change add user information: after change the configuration, record the user name and IP address which change the configuration.
  - Risk warning: warning the risk before the implementation the configuration which customers consider risky.

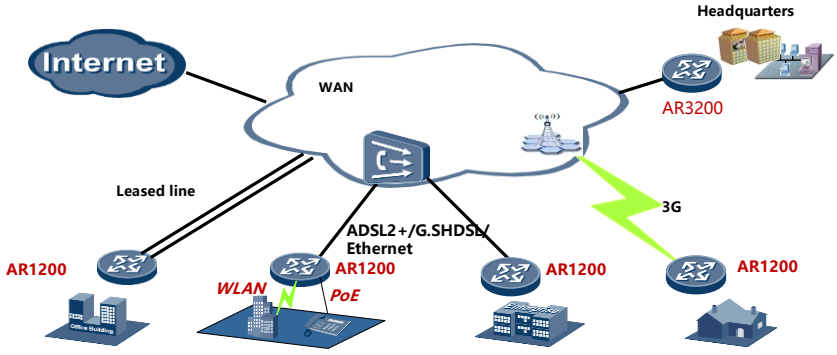
AR as the Enterprise Egress Gateway



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Application:
  - AR G3 enterprise router can be applied to
  - The enterprise headquarters and branch egress gateway to
  - Provide a cost-efficient, highly reliable, and easy-to-deploy interconnection solution.
- Benefits:
  - The AR integrates routing, switching, voice, security, and WLAN functions. You need to deploy only one device at the egress to meet multi-service requirements, which reduces the TCO and protects investments.
  - The AR supports high-density voice card 32FXS and high-density Ethernet card 24GE to connect many voice and data terminals.
  - The AR provides built-in AC, leading in industry. It provides cost-efficient WLAN access solution without deploying extra cards.
  - The AR supports dual SRUs and hot standby, ensuring nonstop service transmission.

Scenario for Enterprise Broadband Access Routers

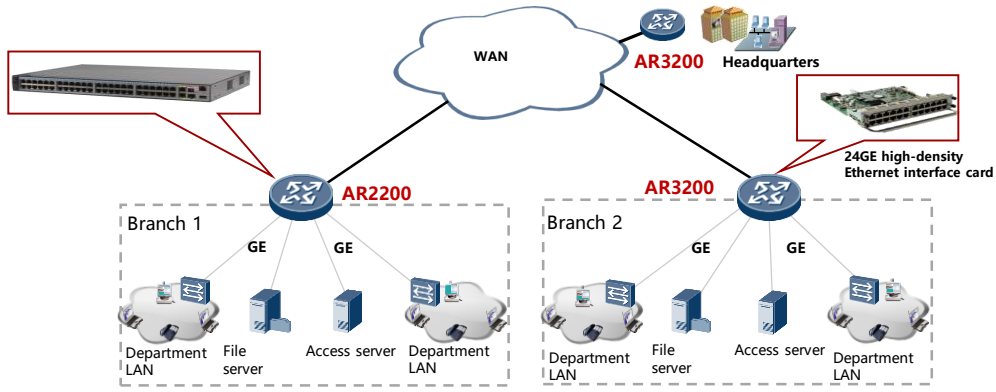


- Accessing the WAN as the enterprise egress router
- Meeting access requirements, including leased line, Ethernet, xDSL, 3G, and WLAN
- Supporting PoE in compliance with IEEE 802.3af and IEEE 802.3at

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The AR G3 routers function as the egress routers of enterprise branches and provide flexible access methods to support remote network connections.
- An AR G3 meets various access requirements, including leased line, Ethernet, xDSL, 3G, and WLAN. This saves deployment and maintenance costs and provides a large value to customers.
- The 100 Mbit/s Ethernet interfaces of an AR1220V and AR1220W (V2R1C01) support PoE in compliance with IEEE 802.3af and 802.3at; therefore, the AR1220V and AR1220W (V2R1C01) can provide power for powered devices (PDs), such as IP phones. An 802.3at interface provides higher than 30 W power, ensuring power for large-power PDs.

Integrated Routing and Switching for Enterprise Branches



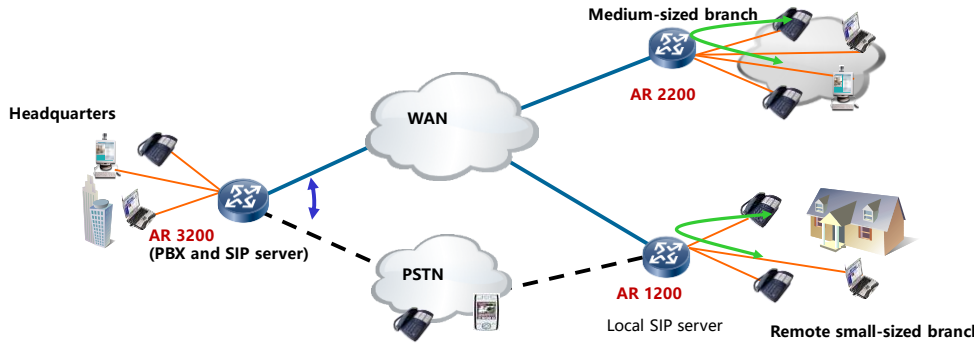
- The 8FE1GE and 24GE interface cards on the AR2200/AR3200 support inter-card VLAN switching, spanning trees, link bundling, and Layer 2/Layer 3 data exchange.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The 8FE1GECombo and 24GE interface cards on the AR2200/AR3200 support inter-card VLAN switching, spanning trees, link bundling, and Layer 2/Layer 3 data exchange.

Enterprise VoIP

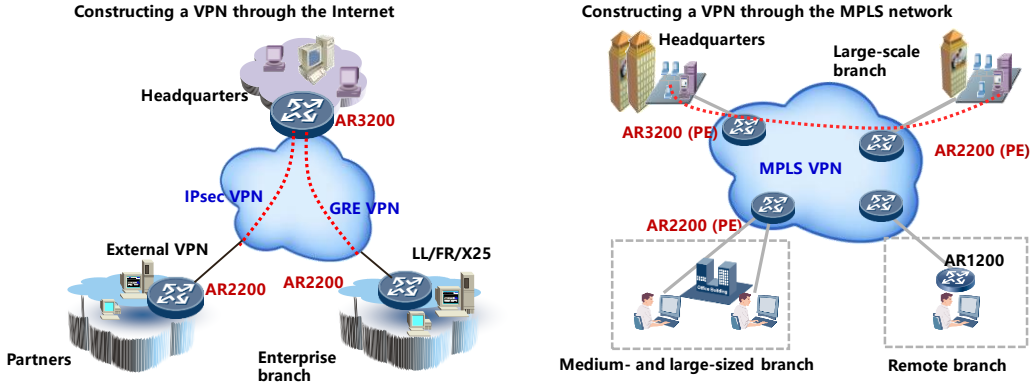


- The AR G3 provides a built-in PBX supporting voice services such as switchboard, IVR navigation, and CDR query.
- The AR G3 supports smart call routing and uses the PSTN network as a backup for calls.
- The AR G3 provides a built-in SIP server, ensuring reliability of voice services.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The AR G3 provides a built-in PBX supporting the enterprise switchboard, IVR navigation, and CDR query functions to enhance corporate image and improve enterprise communication efficiency.
- The AR G3 is located in a branch to provide the smart call routing function. When a fault occurs on the WAN, the PSTN network is used as a backup for calls.
- When the SIP server at the headquarters is unreachable, the built-in SIP server of the AR G3 implements communication between the branch and the PSTN network. This ensures reliability of voice services.
- Note: Only the AR2200/AR3200 (V2R1C01) supports the preceding functions.

Enterprise VPN Gateway

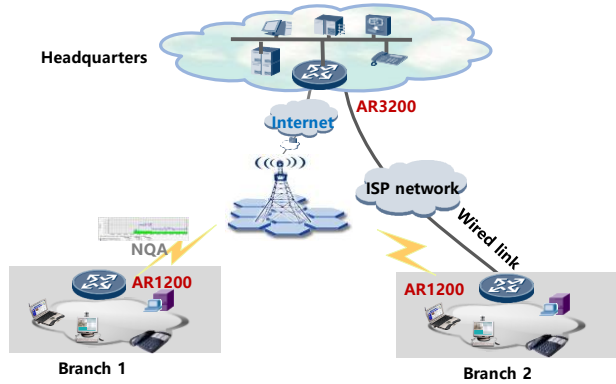


- Tunnels such as GRE VPN and IPsec VPN are established between AR G3 routers to implement secure data access and transmission. The AR G3 implements fast tunnel deployment and authentication for branches.
- As the PEs of an MPLS network, the AR G3 routers are located in the branches. Different types of services are separated by MPLS L3VPN. The AR G3 supports enterprise service operation over

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The AR G3 provides multiple security access functions such as GRE VPN tunnel and IPsec VPN tunnel, implementing secure data access and transmission. The AR G3 implements fast tunnel deployment and authentication for branches. Using a tunnel, partners can access and share enterprise resources and users are authenticated and authorized.
- As the PEs of an MPLS network, the AR G3 routers are located in the branches. Different types of services are separated by MPLS L3VPN. The AR G3 implements flexible deployment, fast distribution, and secure transmission of VPN services, and supports enterprise service operation over networks.

3G/LTE Wireless Access



- The AR G3 complies with 3G standards including CDMA2000 EV-DO, WCDMA, and TD-SCDMA.
- Users can use a 3G USB card to deploy 3G services on the AR G3, saving service card slots.
- The 3G data link can be used as a backup for wired link to protect uplinks.
- The AR G3 provides the NQA function to monitor 3G link quality, ensuring the SLA.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The AR G3 complies with 3G standards including CDMA2000 EV-DO, WCDMA, and TD-SCDMA, meeting wireless communication requirements between branches and the headquarters.
- Users can use a 3G USB card to deploy 3G services on the AR G3, saving service card slots.
- In addition, the 3G data link can be used as a backup for wired link to protect the xDSL, FE/GE, ISDN, and CPOS uplinks. The backup link improves network stability and reduces network construction costs.
- The AR G3 provides the NQA function to monitor 3G link quality, ensuring the SLA.

Quiz

1. What series are the AR G3 routers classified into?
   A. AR3200
   B. AR2200
   C. AR1200
   D. AR150&160&200
   E. AR3600
2. In terms of the appearance, what types of service cards does the AR G3 support?
   A. SIC B. DSIC C. WSIC
   D. XSIC E. EXSIC

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - ABCDE
  - ACDE

Summary

- AR G3 Positioning
  - Comprehensive Enterprise Gateway
- AR G3 Cards
  - SIC→WSIC→XSIC→EXSIC
- AR G3 Feature Description
  - Voice, QoS, Security, and Reliability
- AR G3 Networking

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- AR Product Documentation
  - http://support.huawei.com/ehedex/hdx.do?docid=DOC1000032948&lang=zh
- AR Router Typical Configuration Examples
  - http://support.huawei.com/enterprise/docinforeader.action?contentId=DOC0000707956&idPath=7919710|9856750|7923148|9858988|6078842

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei Learning Website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Library:
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



NE40E-X Series Router Introduction

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- Huawei NetEngine40E Universal Service Router (hereinafter referred to as the NE40E) is a high-end router for core and backbone networks. The NE40E is positioned as the edge or convergence router on the IP backbone network.
- This course introduces NE40E-X16, NE40E-X8 and NE40E-X3’ s hardware features.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Meanwhile, in order that you can study Huawei NE series routers in the round, we attach some contents of Huawei NE20E-X6 introduction to the end of this course especially.
- Huawei NetEngine20E-X6 High-end Service Router(hereinafter referred to as the NE20E-X6) is a high performance router designed for the following custom, such as finance, power, government, education, enterprise, carrier and so on by Huawei, in order to meet the requirement for Carrier HA of carriers and enterprise aggregation and access network.

Objectives

- Upon completion of this course, you will be able to:
  - Describe Positioning of NE40E-X Router
  - Describe Hardware Architecture of NE40E-X Router
  - List Boards of NE40E-X Router
  - Describe Highlights of NE40E-X Router

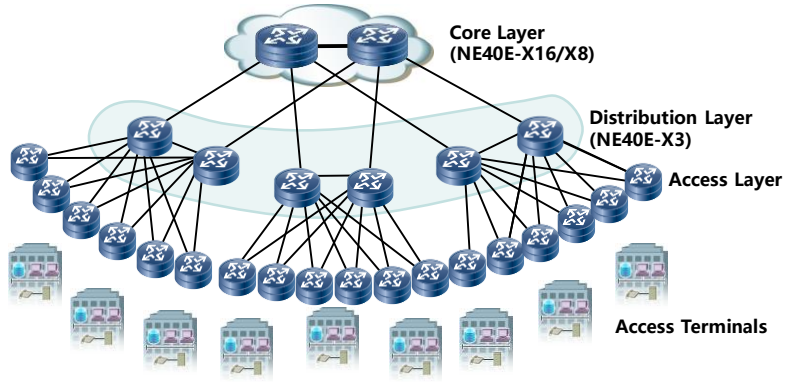
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Product Positioning of NE40E-X Router**
2. Hardware Architecture of NE40E-X Router
3. Boards of NE40E-X Router
4. Highlights of NE40E-X Router
5. Application Scenario of NE40E-X Router

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Positioning of NE Routers



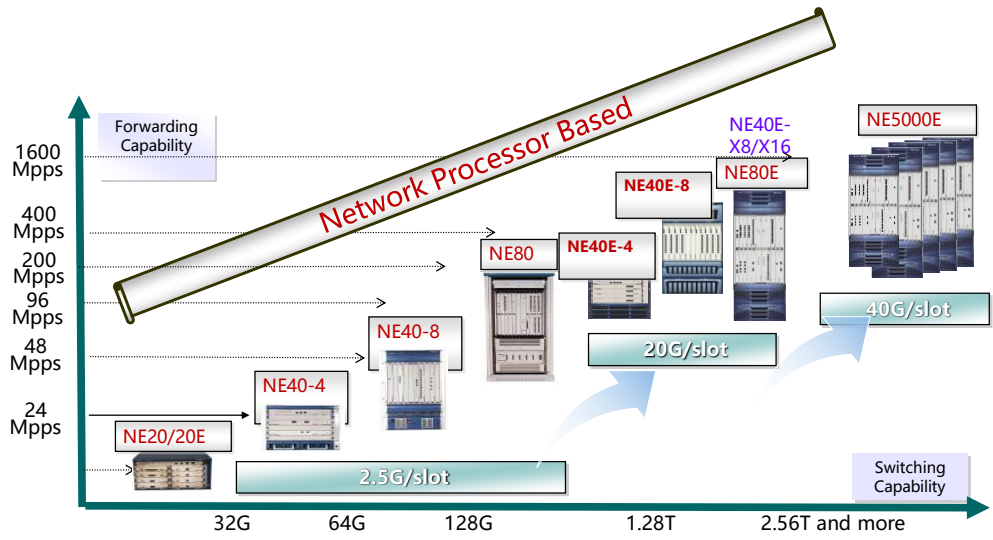
- NE40E-X16/X8 targeted at large enterprise network core or IDC export.
- The NE40E-X3 targeted at large enterprise network aggregation, or SMB network core.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Huawei HUAWEI NetEngine40E Universal Service Router (hereinafter referred to as the NE40E) is a high-end router with 10-Gbit/s interfaces designed for core and backbone networks. The NE40E is positioned as the edge or convergence router on the IP backbone network.

NetEngine Series Router



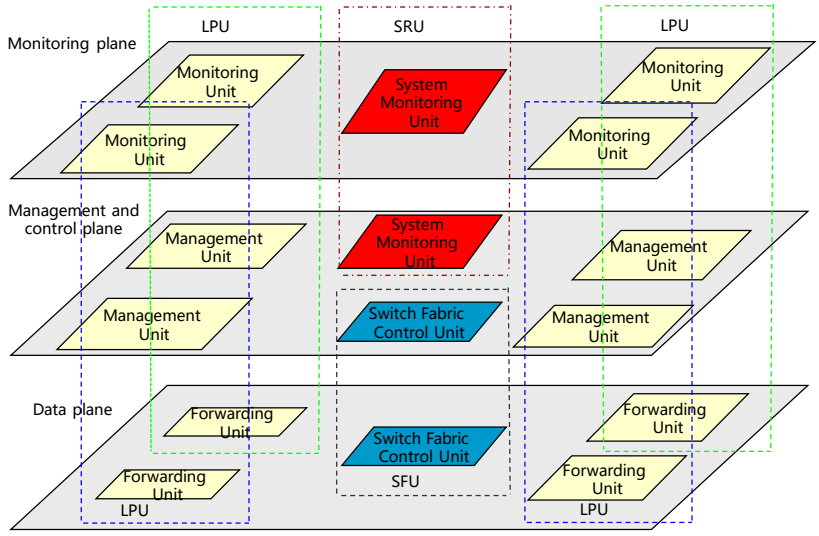
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Product Positioning of NE40E-X Router
2. **Hardware Architecture of NE40E-X Router**
3. Boards of NE40E-X Router
4. Highlights of NE40E-X Router
5. Application Scenario of NE40E-X Router

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Overview - the System Architecture

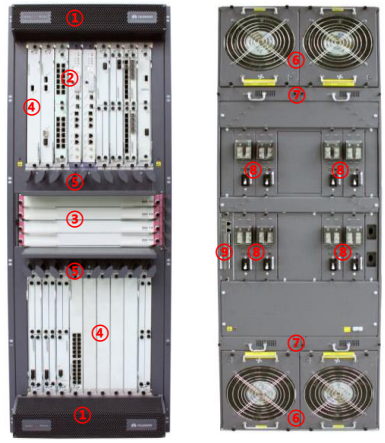


- The data plane, management and control plane are separated
- The monitoring plane and service plane are separated
- Distributed Forwarding
- Three-level switch fabric

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

- The NE40E-X adopts a system architecture as shown in Figure above. In this architecture, the data plane, management and control plane, and monitoring plane are separated. This design helps to improve system reliability and facilitates separate upgrade of each plane.

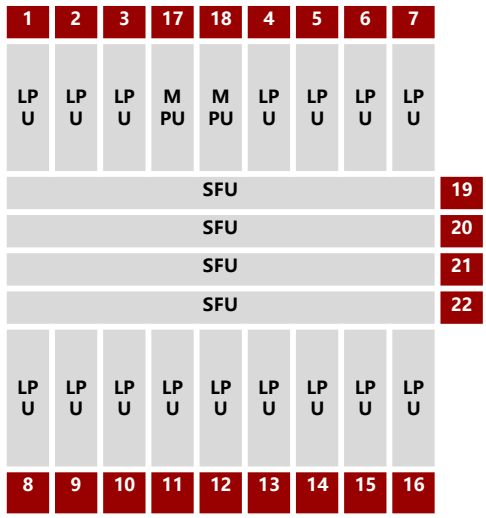
Appearance of NE40E-X16



No.	Module	Quantity
①	Air intake vent	× 2
②	MPUs	× 2
③	SFUs	× 4
④	LPUs	× 16
⑤	Cabling Area	× 2
⑥	Fan module	× 4
⑦	Filtering Box	× 4
⑧	Power Module	× 8
⑨	CMU	× 1

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Slot Layout of NE40E-X16



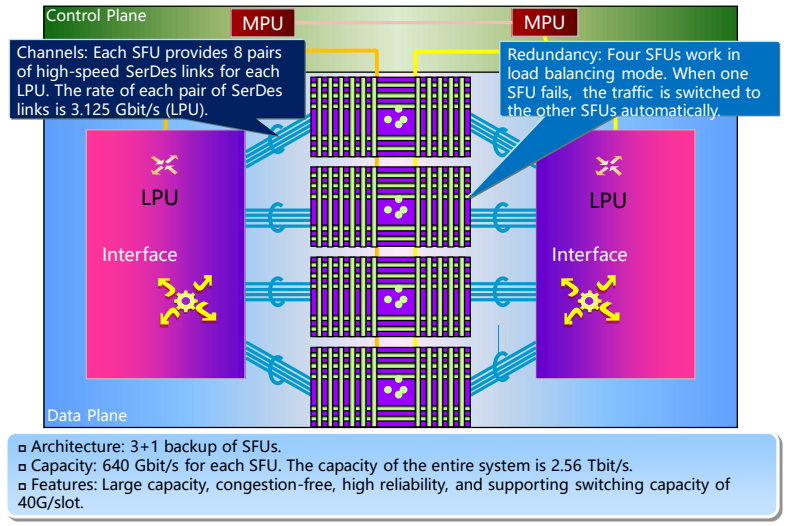
Slot No.	Remarks
1 ~ 16	Hold LPUs
17 ~ 18	Hold MPUs, working in1:1 backup mode
19 ~ 22	Hold SFUs, working in 3+1 backup mode

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- MPU: Main Processing Unit
- SFU: Switch Fabric Unit
- LPU: Line Processing Unit

Architecture of NE40E-X16

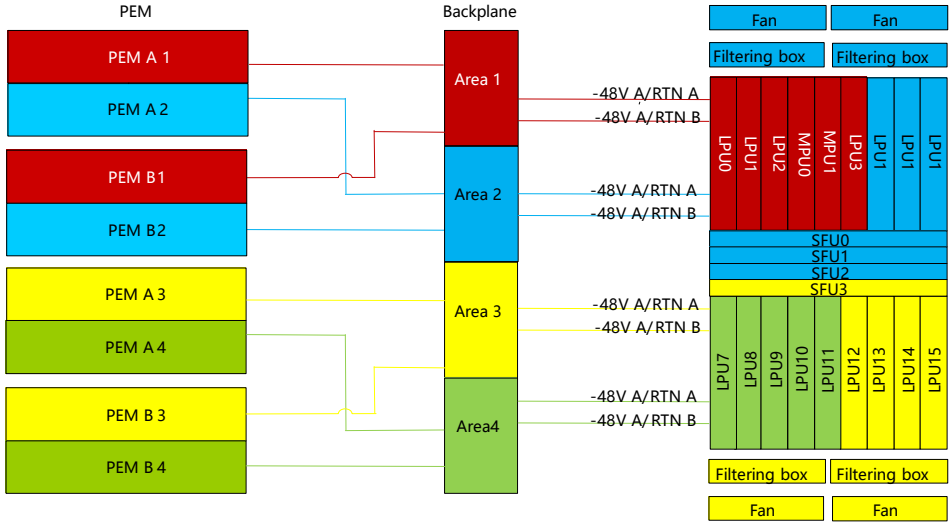


- Architecture: 3+1 backup of SFUs.
- Capacity: 640 Gbit/s for each SFU. The capacity of the entire system is 2.56 Tbit/s.
- Features: Large capacity, congestion-free, high reliability, and supporting switching capacity of 40G/slot.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

- The SFU on the NE40E-X16 switches data for the entire system at wire speed of 640 Gbit/s (320 Gbit/s for the upstream traffic and 320 Gbit/s for the downstream traffic). This ensures a non-blocking switching network.
- The NE40E-X16 has four SFUs working in 3+1 load balancing mode. The entire system provides a switching capacity at wire speed of 2.56 Tbit/s.
- The four SFUs load balance services at the same time. When one SFU is faulty or replaced, the other three SFUs automatically take over its tasks to ensure normal running of services.

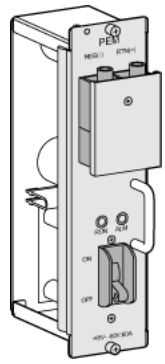
Distributed Architecture of NE40E-X16



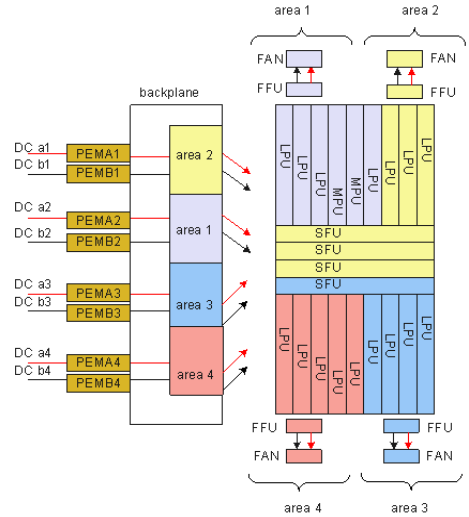
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As shown in figure above, the NE40E-X16 backplane is divided into four areas, with each area having two power inputs. These eight power inputs work in backup mode.
- The NE40E-X16 supports either DC or AC power supply.

DC Power Supply System of E40E-X16



Appearance of the NE40E-X16's PEM



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In a DC power supply system of the NE40E-X16, eight 70 A PEMs work in 4+4 backup mode.
- Figure above shows details on the DC power supply system:
  - Two -48V power inputs join on the board.
  - After the low-frequency filtering, the two -48 V power inputs for fans join inside the fan module.
  - Each DC power input contains one -48 V power inputs and one RTN inputs. Two separated RTN inputs join on the board.

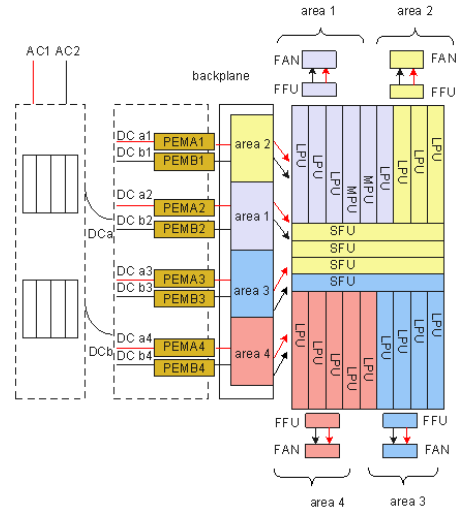
AC Power Supply System of E40E-X16



Front view of the AC rectifier module



Rear view of the AC rectifier module

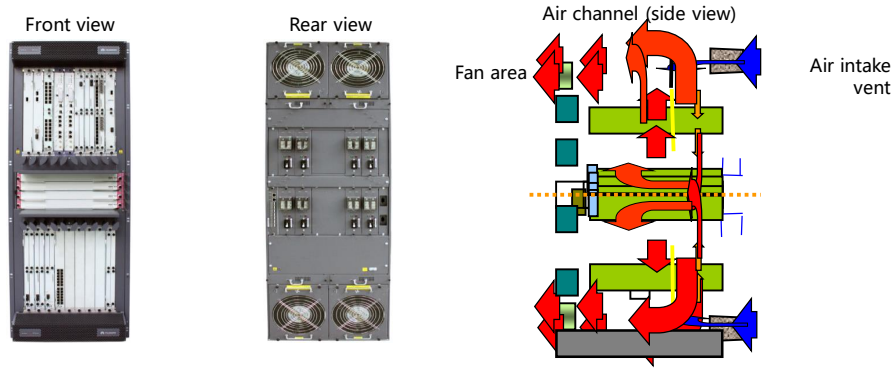


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The input AC power is converted into regulated DC power by an AC/DC converter. The resulting DC power output is connected to the PEMs through external cables to supply power for all boards and fan modules.
- Two -48V power inputs are joined on the board.
- After the low-frequency filtering, the two -48 V power inputs for fans are joined inside the fan module.

Heat Dissipation System of NE40E-X16

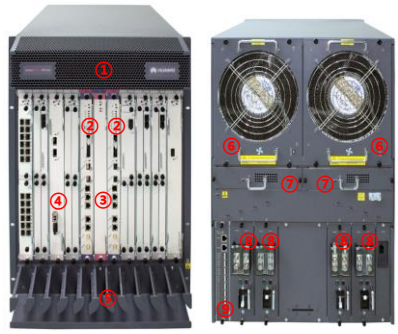


- Fan modules at the back
- 2+2 backup of fans
- U-shaped air channel for effective heat dissipation
- -X16: separate air channels for heat dissipation of the upper and lower frames

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The NE40E-X16 is divided into the upper chassis and the lower chassis, and draws air from the front and exhausts air from the rear. The air intake vent on the upper chassis resides above the board area on the front chassis; the air exhaust vent resides above the board area on the rear chassis. The lower chassis and the upper chassis are opposites. In addition, the upper chassis and the lower chassis have separate heat dissipation systems.
- The middle area of the chassis is for SFU slots. The air intake vent of this area resides on the left of the chassis. Two upper SFU slots in the area draw air from the left. When flowing to the right, the air joins the air from the upper chassis. Two lower SFU slots in the area draw air from the left. When flowing to the right, the air joins the air from the lower chassis.
- The NE40E-X16 has three air channels:
  - The upper and lower chassis have separate air channels that draw air from the front and exhausts air from the rear. The air filters at the air intake vents are vertically installed. The curved face, large area, and small windage resistance of the air filters help to improve the heat dissipation efficiency. The two air filters on the upper and lower chassis are the same.
  - The air channel in the SFU slot area is located on the left of the chassis. The air filter adopts front access. The depth of the air filter is the same as that of an SFU and the height of the air filter is four times the height of the an SFU.

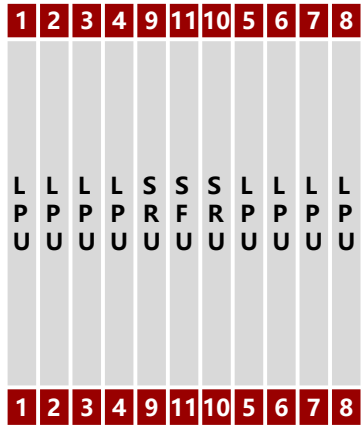
Appearance of NE40E-X8



No.	Module	Quantity
①	Air intake vent	× 1
②	SRUs	× 2
③	SFUs	× 1 (Totally 3 SFUs, 2 of which are integrated on SRUs)
④	LPUs	× 8
⑤	Cabling Area	× 1
⑥	Fan Module	× 2
⑦	Filtering Box	× 2
⑧	Power Module	× 4
⑨	CMU	× 1

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Slot Layout of NE40E-X8

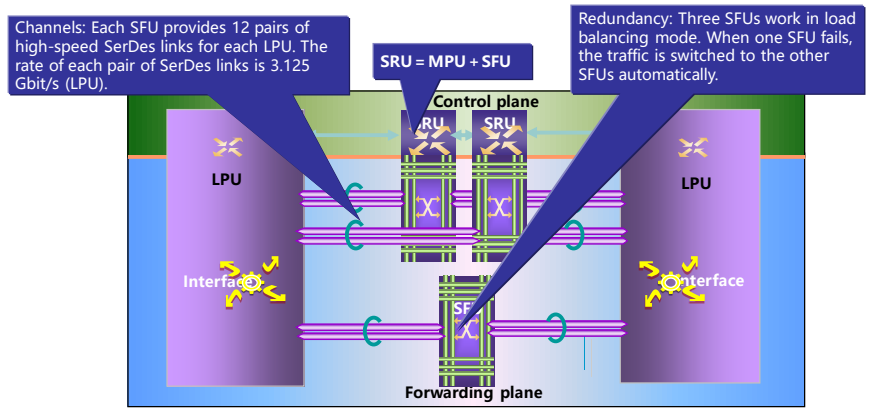


Slot No.	Remarks
1 ~ 8	Hold LPUs
9 ~ 10	Hold MPUs, working in 1:1 backup mode
11	Hold SFUs, working in 2+1 backup mode

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- SRU: Switch Router Unit

Architecture of NE40E-X8

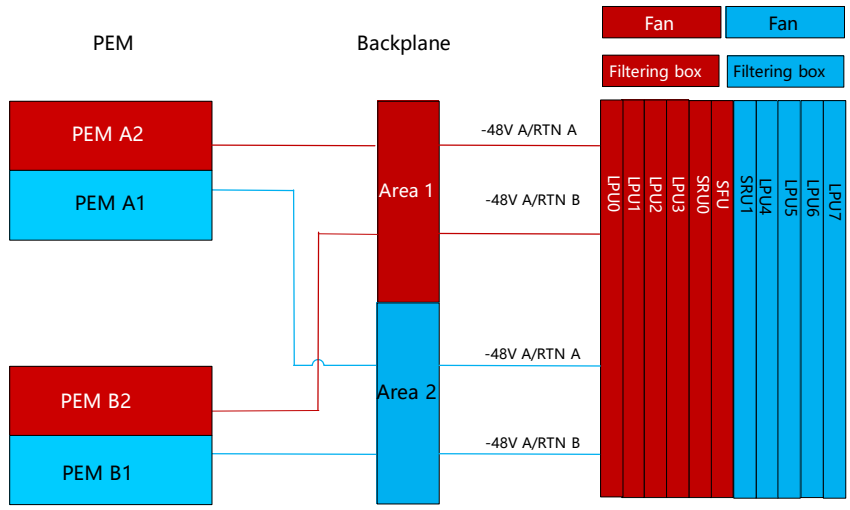


- Architecture: 2+1 backup of SFUs. Two SFUs are integrated on the SRU.
- Capacity: 480 Gbit/s for each SFU. The capacity of the entire system is 1.44 Tbit/s.
- Features: Large capacity, congestion-free, high reliability, and supporting switching capacity of 40G/slot.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The SFU on the NE40E-X8 switches data for the entire system at wire speed of 480 Gbit/s (240 Gbit/s for the upstream traffic and 240 Gbit/s for the downstream traffic). This ensures a non-blocking switching network.
- The NE40E-X8 has three SFUs working in 2+1 load balancing mode. The entire system provides a switching capacity at wire speed of 1.44 Tbit/s.
- The three SFUs load balance services at the same time. When one SFU is faulty or replaced, the other two SFUs automatically take over its tasks to ensure normal running of services.

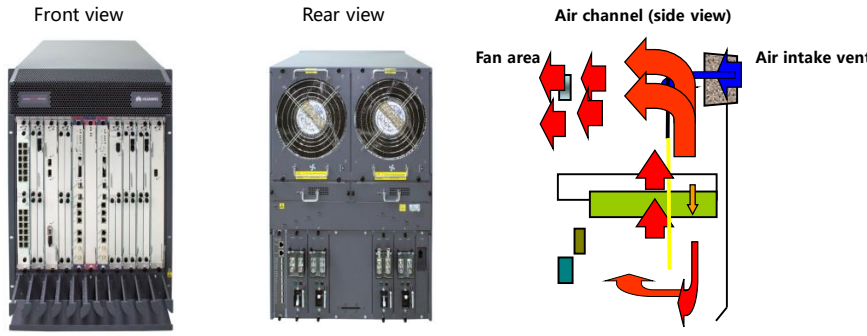
Distributed Architecture of NE40E-X8



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As shown in figure above, the NE40E-X8 backplane is divided into two areas, with each area having two power inputs. These four power inputs work in backup mode.
- The NE40E-X8 supports either DC or AC power supply.
- In a DC power supply system of the NE40E-X8, four 70 A PEMs work in 2+2 backup mode. The figure shows details on the DC power supply system:
  - Two -48 V power inputs join on the board.
  - After the low-frequency filtering, the two -48 V power inputs for fans join inside the fan module.
  - Each DC power input contains one -48 V power input and one RTN input. Two separated RTN inputs join on the board.
- In the case of an AC power supply system, an AC power frame is placed outside the chassis and installed with rectifier modules based on system power. The AC power frame is then connected to the input terminals on the DC-PEMs to supply power for the system. (In short, an external AC power frame is added to the DC power supply system to constitute an AC power supply system.)

Heat Dissipation System of NE40E-X8



- Fan modules at the back
- 1+1 backup of fan modules
- U-shaped air channel for effective heat dissipation

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The heat dissipation system is responsible for dissipating heat for the entire system. The heat generated by boards is dissipated through the heat dissipation system. In this manner, the temperature of the components on boards are controlled within a normal range, enabling the boards to work stably.
  - The heat dissipation system is composed of fan modules (one fan in each fan module), fan control boards (FCBs), temperature sensors, air filters, air intake and exhaust vents, and a system air channel.
  - When a single fan fails, the other fans automatically rotate at full speed. In this case, the heat dissipation system enables the system to work in a short period of time at ambient temperature of 40℃.
  - Temperature sensors, located on the air exhaust vent and boards, are used to monitor the temperature of the components on boards and adjust the fan speed through the command delivered by the SRU to control the temperature in a normal range.
  - The power modules of the system have two fans of their own for independent heat dissipation.
- As the figure shown above, The NE40E-X8 draws air from the front and exhausts air from the back. The air intake vent resides above the board area on the front chassis; the air exhaust vent resides above the board area on the rear chassis.
- The two fan modules of the NE40E-X8 are located side by side at the air exhaust vent, with each module containing one fan. The entire system dissipates heat by drawing air, as shown in figure above.

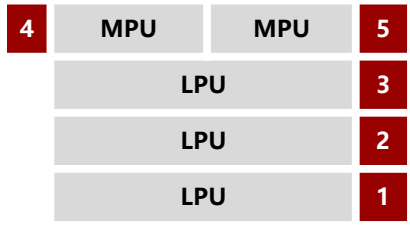
Appearance of NE40E-X3



No.	Module	Quantity
①	MPUs	× 2
②	LPUs	× 3
③	Power Module	× 2
④	Fan Module	× 1

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

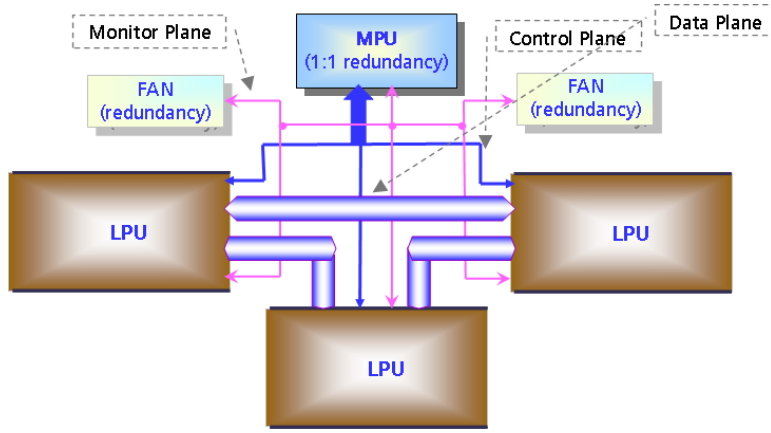
Slot Layout of NE40E-X3



Slot No.	Remarks
1 ~ 3	Hold LPUs
4~5	Hold MPUs, working in 1:1 backup mode
Notice!	No SFU is adopted on NE40E-X3

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

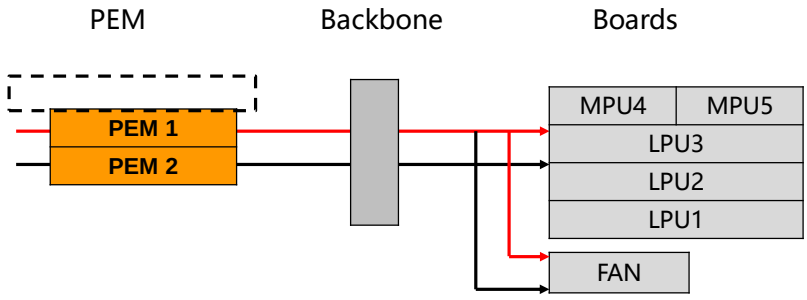
Architecture of NE40E-X3



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- With full-mesh architecture, NE40E-X3 does not need a SFU.

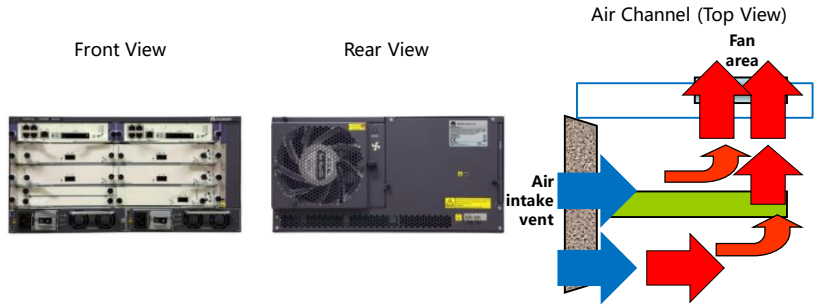
Power Supply System of NE40E-X3



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Two AC power modules or two DC power modules work in 1+1 backup mode to improve the reliability of power supply. The figure shows the diagram of the power supply system.

Heat Dissipation System of NE40E-X3



- Fan module at the back
- U-shaped air channel for effective heat dissipation

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The NE40E-X3 draws in air from the left and exhausts air from the rear. The air intake vent is located at the left side of the chassis and the air exhaust vent is located at the rear of the chassis.
- The fan module of the NE40E-X3 is located at the air exhaust vent. The system draws in air for heat dissipation.

Contents

1. Product Positioning of NE40E-X Router
2. Hardware Architecture of NE40E-X Router
3. **Boards of NE40E-X Router**
4. Highlights of NE40E-X Router
5. Application Scenario of NE40E-X Router

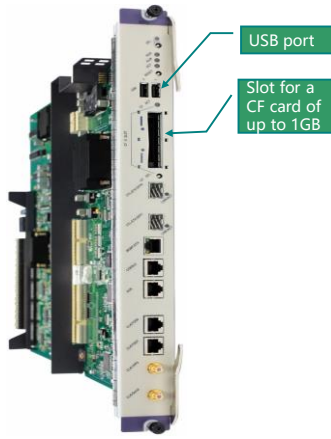
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Types of Boards of NE40E-X

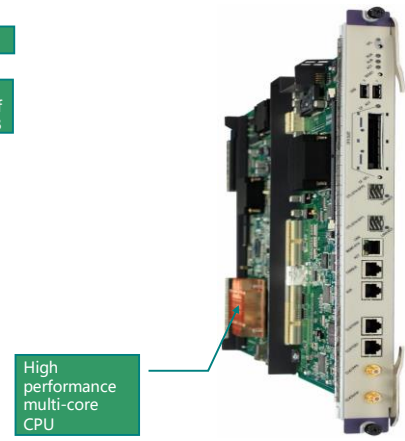
- Main Process Unit
  - X16: MPU
  - X8: SRU
  - X3: MPU
- Switch Fabric Unit
  - X16: SFUI-200-B
  - X8: SFUI-200-C
- Centralized Monitoring Unit: CMU
- Line Processing Unit

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

MPU&SRU (1)



MPU of NE40E-X16



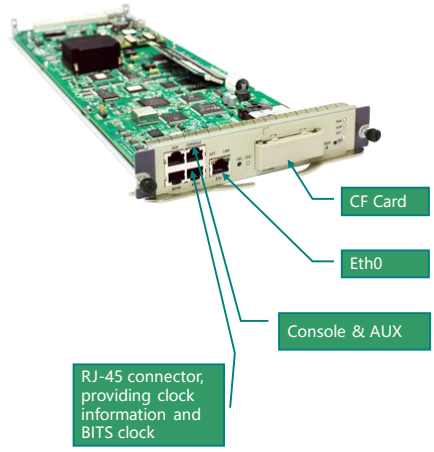
SRU of NE40E-X8(integrated with SFU)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The control plane of the NE40E-X16 adopts MPU.
- The following USB interface attributes are supported by MPU:
  - Supports the biggest USB fat32 format, and supports the memory available in the market.
  - For security reasons not allowed to write USB storage device .
  - Updates automatically, insert the USB memory without any operating.
- Highlights of the MPU
  - Two USB ports: supporting version downloading through USB devices and power supply for USB devices
  - CF card with mass storage capacity (up to 1 GB)
  - Compatible with the design of disks
  - RJ-45/SMB connector: processing Stratum-3 clock and 1588 clock; supporting input and output of 2MHz/2Mbps/1PPS clock signals
  - High performance multi-core CPU
  - The bandwidth of the control bus between the MPU and the LPU is increased to 1 Gbit/s.

MPU&SRU (2)



MPU of NE40E-X3

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The MPU of the NE40E-X3 controls and manages the system and switches data. The MPUs work in 1+1 backup mode. The MPU consists of the main control unit, system clock unit, synchronous clock unit, and system maintenance unit. The functions of the MPU are described from the following aspects.

SFU Boards of NE40E-X16/8



Appearance of the panel



Appearance of the Switch Fabric Unit B (SFUI-200-B) (Adopted NE40E-X16)



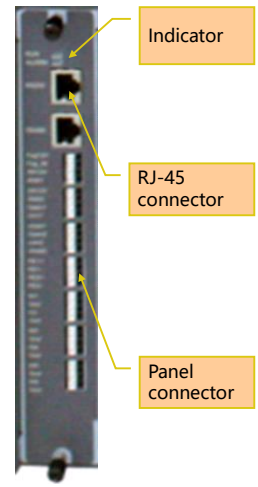
Appearance of the Switch Fabric Unit C (SFUI-200-C) (Adopted NE40E-X8)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A switching network is a key component of the NE40E and is responsible for switching data between LPUs.
- Switching Board comprises of CPU module, switching module.
- Mainly responsible for switching data between LPUs.
- NE40E-X16 has four SFUs that work in 3+1 load balancing mode.
- NE40E-X8 support 2+1 load balancing mode
- Indicators on panel include ACT indicator, RUN indicator and OFL indicator.

CMU of NE40E System (Optional)

- Extensive environment monitoring functions
  - Alarm detection of the smoke sensor
  - Detection of the ambient temperature
  - Access control management
  - Device alarm output
  - Main contact point inspection
  - One 232 and 485 serial interface
- Perform the system environment monitoring and access control functions
- Not require additional devices, saving customers' investment.

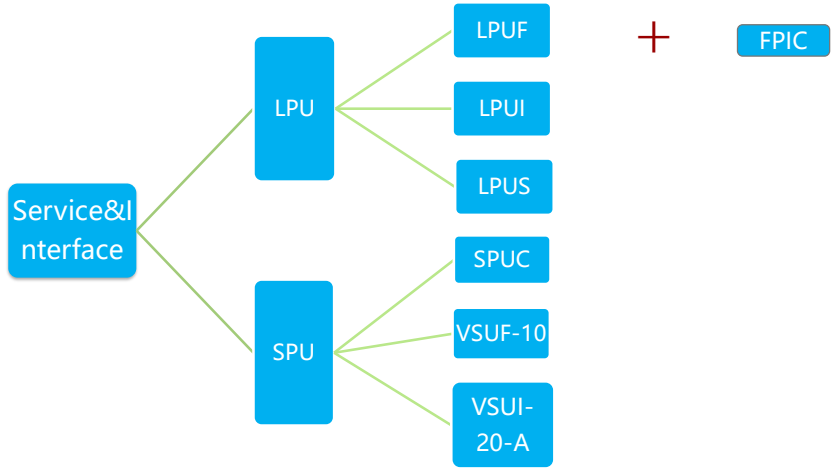


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Extensive environment monitoring functions
  - Alarm detection of the smoke sensor :Supports the connection to the smoke sensor through the panel to detect the alarm signals from the chassis or equipment room.
  - Detection of the ambient temperature :Supports the connection to the temperature sensor through the panel to detect the temperature of the chassis or equipment room.
  - Access control management :Detects whether access control is enabled through magnetic inspection and reports the inspection signal to the device. The remote unlocking function is reserved. You can instruct the CMU to enable or disable access control through the remote control function.
  - Device alarm output :The CMU provides two-level alarm output signals.
  - Main contact point inspection :The CMU can provide six main contact points to detect signal input and monitor whether the devices outside the chassis work normally.
  - One 232 and 485 serial interface :Provides an RS-232 serial interface, which is connected to the panel. You can use it to query or locate information about the CMU. In addition, the CMU provides an R-485 serial port, which is connected to the panel. You can connect an device to this interface. The interface supports full-duplex mode.

Service Interface&Boards



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As the Universal Service Router , NE40E-X series routers supply divers interfaces, such as ethernet, POS, CPOS, E1 and so on.
- Boards supplied by NE40E-X for universal service include following types:
  - LPUF(supply service interface matching corresponding FPIC)
  - LPUI
  - LPUS
  - SPU

LPU

- For NE40E-X, types of LPUs are divided into LPUF, LPUI and LPUS.Only matching corresponding FPICs, the LPUF can supply service interfaces, and LPUI and LPUS have intergrated service interfaces on them.

Board Specifications	Board Types	Corresponding Device	Initiative Version
10G	LPUF-10	NE40E-X3/X8X16	V3R3C00
20G	LPUF-20/21	NE40E-X3/X8X16	V3R3C02
	LPUS-20	NE40E-X3/X8X16	V6R1
40G	LPUF-40 LPUI-40	NE40E-X3/X8X16	V6R1C00
	LPUI-41 LPUS-41	NE40E-X3/X8X16	V6R3C00
100G	LPUF-100 LPUI-100	NE40E-X8X16	V6R3C00

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Different specifications of the LPU board is mainly reflected in the QoS.

LPUF-40 40G Board



- Provides two slots, each of which can hold a flexible plug-in card of the LPUF-40. The cards support hot swap.
- Supports a maximum of 40 Gbit/s bandwidth.
- Provides two models: LPUF-40-A and LPUF-40-B.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- High-Queue LPUF-40：Full-service Flexible Linecard, 512K flow queues, supporting features such as BRAS, 1588v2 and enhanced QoS, positioned on complicated service aggregation: BRAS and SR/PE downlink, etc.. V6R3 version will provide new linecards on LPUF-40, including 8-port 10GE oversubscribed card, 40-port GE electronic card and 4-port 10G POS card. In addition, 40G enhanced HQoS and BRAS card will be available in V6R3.
  - LPUF-40 provides two models: LPUF-40-A and LPUF-40-B
  - The LPUF-40-A supports all software features
  - the LPUF-40-B supports all software features except L3VPN, MVPN, and IPv6, and can be upgraded to support all features of the LPUF-40-A through licenses

FPICs of LPUF-40 Board

Type	Description	Remarks	Appearance
LPUF-40	4-Port 10GBase LAN/WAN-XFP 20-Port 10/100/1000Base-RJ45	P40 Flexible Card	
	2-Port 10GBase LAN/WAN-XFP-E 20-Port 1000Base-SFP-E	P40-E Flexible Card, Supports BRAS, 1588v2, Enhanced QoS	
	2-Port 10GBase LAN/WAN-XFP-V	P40-V Flexible Card	

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LPUF-100 100G Board



- Provides four slots that can hold full-height FPICs or four half-height FPICs
- An FPIC supports the following functions:
  - Hot swap
  - Automatic recovery of configurations
  - Intermixing with other types of FPICs

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- 100G linecards include two types: High-Queue LPUF-100 and Medium-Queue LPUI-100. LPUF-100 is Flexible Linecard and provide 512K flow queues, and support flexible configuration of 10GE, GE, 10G POS and 40G POS. LPUI-100 is Integrated Ethernet Linecard and provide 256K flow queues, and meet the requirement of different networking.
- 100G linecards in V6R3 can provide 8*10GE, 10*10GE, 16*10GE oversubscribed, 96*GE, 8*10G POS, 2*40G POS and 1*100GE. In industry, the interface type of NE40E 100G linecards is most abundant, and the port density of NE40E 100G linecards is highest.
- Note:
  - Provided using 100G Board, we have to switch SFU board(and SRU board on NE40E-X8) to another one with 200G, what' more, the 200G SFU board and corresponding SRU board can' t be used together with 40G SFU, LPUA, LPUB, LPUG at the same time;
  - The LPUI-100 can be used only on the NE40E-X16 and NE40E-X8

FPICs of LPUF-100 Board



5-Port 10GBase LAN/WAN-XFP Flexible Card A
Supports 1588v2



24-Port 100/1000Base-SFP Flexible Card(P100,
Occupy two sub-slots)



48-Port 100/1000Base-X-CSFP Flexible Card(P100,
Occupy two sub-slots)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LPUI-100 100G Board

- LPUI-100 board is simplified from LPUF-100 board.



10-Port 10GBase LAN/WAN-XFP Integrated Line Processing Unit (LPUI-100, Supporting 1588v2))



16-Port 10GBase LAN-SFP+ Integrated Line Processing Unit(LPUI-100)



1-Port 100GBase-CFP Integrated Line Process Unit(LPUI-100)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SPUC



- 3 running modes of SPUC：NetStream mode，Tunnel mode and NAT mode.
- In the same SPUC board, NetStream, NAT and Tunnel MVPN are incompatible.
- Fuctions achieved by SPUC need GTL License support.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- An SPUC implements the NetStream function and processes tunnel services related to GRE and NAT and multicast VPNs.
- An SPUC does not have any physical interfaces and can be inserted into any LPU slot.
  - 3 running modes of SPUC：NetStream mode，Tunnel mode and NAT mode
  - NetStream mode
  - Under the NetStream mode, the SPUC board can implement centralized NetStream mode.
  - Meanwhile, the centralized NetStream still applys License(NetStream License for SPUC), and each SPUC board need one License.
  - Tunnel mode
  - Under the Tunnel mode, the following functions can be provided:
  - Centralized multicast VPN：If running the multicast VPN in SPUC boards, We need to configurate the same number of MVPN License with SPUC amount.
  - Tunnel:SPUC board can provide centralized tunnel, including GRE tunnel and 4over6 tunnel currently。If running the tunnel in SPUC boards, We need to configurate the same number of tunnel License with SPUC amount.
  - NAT mode
  - SPUC board support NAT, NAT License must form 1:1 with the SPUC.
  - GTL:Global Trotter License

Contents

1. Product Positioning of NE40E-X Router
2. Hardware Architecture of NE40E-X Router
3. Boards of NE40E-X Router
4. **Highlights of NE40E-X Router**
5. Application Scenario of NE40E-X Router

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

NE40E Product Highlights





Large Capacity & Future Scalability

- Industry-leading 400G platform
- 1320*GE classis double average in density
- Compatibility designing & smooth upgrading



Stable, secure and reliable network

- Rich High-Availability Technologies
- Security Network with High-performance
- Business depth Monitoring

Commercial IPv6 solution

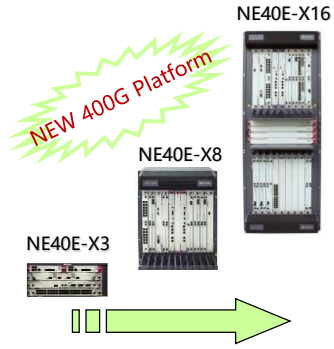


Designed for the real needs of customers

- Easy to Deploy
- H-QoS
- Video Solution

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Green 400G Platform



Green platform, unified NMS, unified platform
From 10G to 400G, compatible to all boards
Save OPEX and CAPEX, saves TCO by 40%

Greenest platform

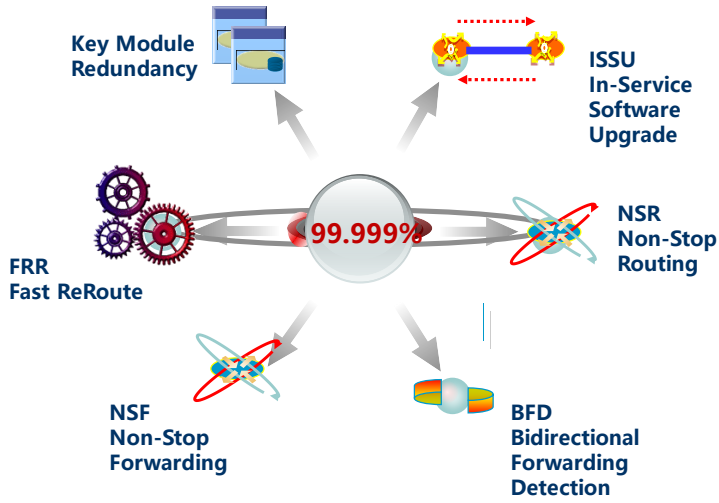
- Greenest platform: lowest power consumption, highest efficiency in heat dispersion.
- Power supply by areas, power supply mode can be unmodified.
- Lowest power consumption: Power consumption of a 10G interface is 30% lower than that of the other supplier. A device saves power of 10000 kwh each year.
- Highest efficiency in heat dispersion: innovative U-shaped air channel, greatly improves the heat dispersion system.
- Compact design, save space by 40%.

High performance and great capacity

- Most compact design: 14 U/32 U. Three chassis in a cabinet
- High interface density: up to 132*10GE
- 400G platform, backplane capacity up to 30 T

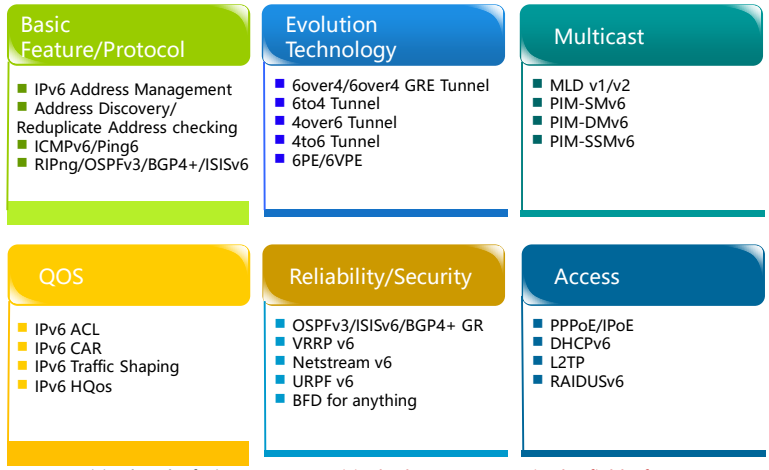
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Reliability Feature



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

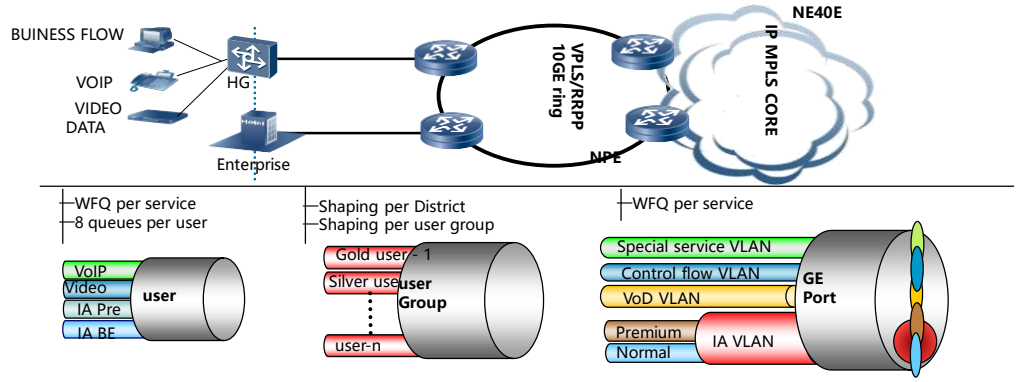
Full IPv6 Support



Huawei is ahead of Cisco, so Huawei is the best company in the field of IPv6.
- Latif Ladid (President of IPv6 Forum)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

H-QoS for Multi-Play Service



Centralized control, achieves fine service management in NPE points, more flexible service strategy.

- Gold, silver or bronze medal users have different service experience.
- The control of the Special service is more effective.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The NE40E supports entire HQoS solutions, HUAWE is the only vendor that supports HQoS, DS-TE and MPLS HQoS, the other vendors support one or two. Thus, HUAWEI can provide a entire HQoS solution to meet kinds of scenarios of carrier-class services.

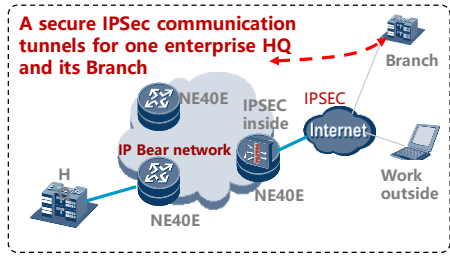
IPSec for High-speed Secure Interconnection



VSUI-20-A

\+ IPSec License → IPSec

1:1



Key Message

- High performance：line speed of **10G**/Slot(512B), **20K** concurrent tunnel
- Supports hot standby and loading-balance between boards
- Supports NAT traversal
- Entire system supports a maximum of 8

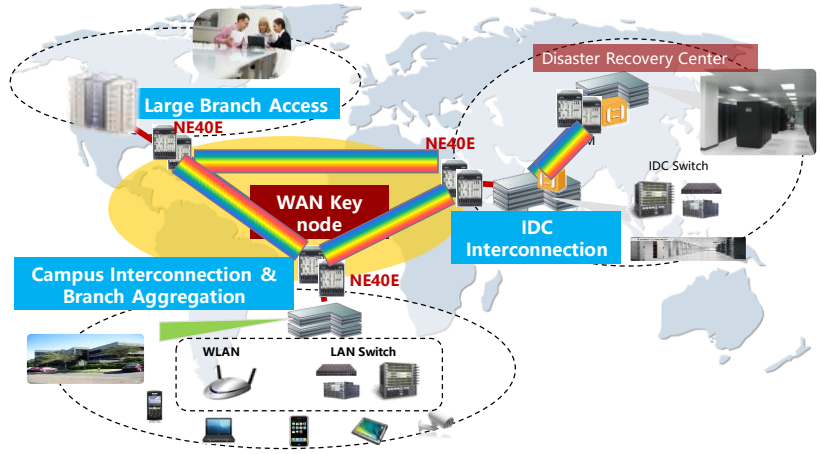
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Product Positioning of NE40E-X Router
2. Hardware Architecture of NE40E-X Router
3. Boards of NE40E-X Router
4. Highlights of NE40E-X Router
5. **Application Scenario of NE40E-X Router**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

The Main Scenario of NE Series Router



Power industry, Finance, Oil & GasPipeline, Oilfield, e-Education, Government, etc.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The main scenario of NE40E Router: Campus and IDC interconnection, Large branch access, Key nodes of WAN.

Quiz

1. Which of the following card types is supported by NE40E-X16 ?

   A. SPUC

   B. LPUF

   C. LPUI

   D. LPUS

2. Which of the following is the control plane of NE40E-X3 ?

   A. MPU B. MCU C. SRU D. SFU

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- ABCD
- A

Summary

- Huawei NE40E router located in the core / aggregation layer of a large network
- The CLOS Architecture of NE40E-X16/X8, and the Full-Mesh Architecture of X3
- Types of Boards of NE40E-X
- Highlights of NE40E-X
  - Capability, High-Reliability and HQoS
- Main Scenario of NE40E-X

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- NE40E Product Documentation
  - http://support.huawei.com/ehedex/hdx.do?docid=DOC0100595997&lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Introduction to Huawei Data Center S Series Switches

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- The Quidway S2700/S3700/S5700/S6700 Series Ethernet switches (hereinafter referred to as the SX7 fixed switch) provide the access, aggregation, and data transport functions. They are developed by Huawei to meet the requirements for reliable access and high-quality transmission of multiple services on the enterprise network.
- The S-Series modular switch applies to enterprise campus networks, helping enterprises build an Ethernet network.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Describe SX7 fixed switch product positioning
  - Describe SX7 fixed switch sub-cards and modules
  - Describe SX7 fixed switch product characteristics
  - Describe SX7 fixed switch typical application scenarios

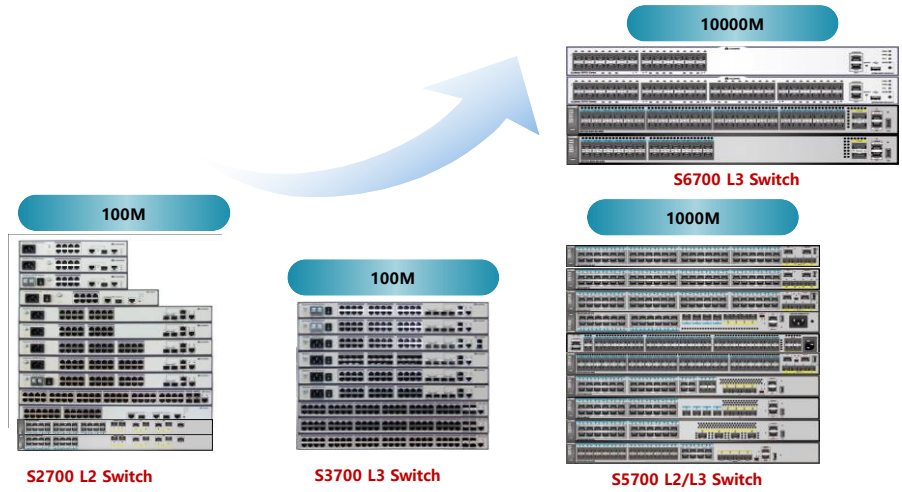
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Overview of S Series Fixed Switches**
2. Overview of S Series Modular Switches
3. Sub-cards and Modules
4. Product Features

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

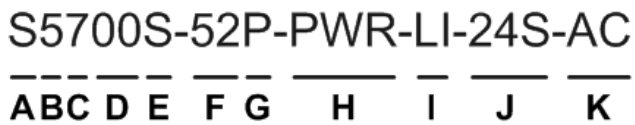
SX7 Series Ethernet Switches Family



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In terms of product functions, the S1700, S2700, S5700LI are Layer 2 switches, while the S3700, S5700 (except the S5700LI) and S6700 are Layer 3 switches.
- (Compared with Layer 2 switches, Layer 3 switches support Layer 3 features such as dynamic routing protocols in addition to Layer 2 features.
- S2700 and S3700 can support to V1R6 software version, S5710LI, S5700SI, S5700EI, S5700HI, S5710HI and S6700 can support to V2R5 software version, and the others can support to higher software version, now is V2R9.

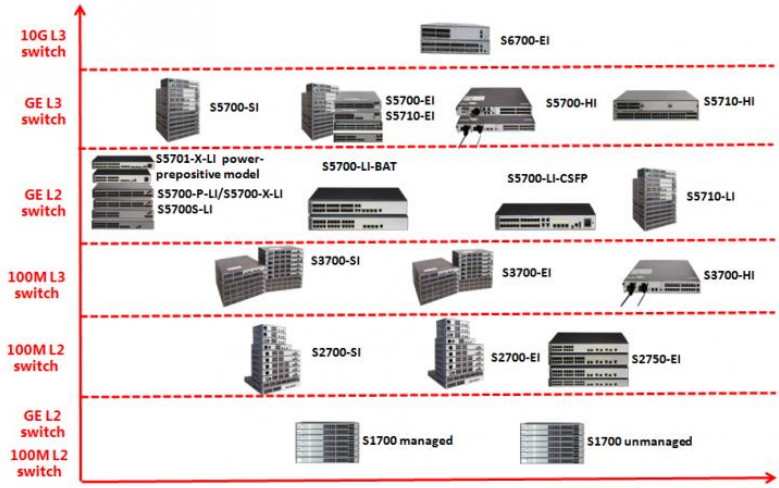
Fixed Switch Naming Conventions



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A: Switch.
- B: Series
  - 6: 10GE downlink ports
  - 5: GE downlink ports
  - 3: Layer 3 switch with 100M downlink ports
  - 2: Layer 2 switch with 100M downlink ports
- C: Enterprise series switch
- D: Product sub-series (such as 00 or 10)
- E: S: resale model
- F: Maximum number of ports
  - NOTE: On an S5710-EI switch (such as S5710-28C-EI), this field indicates the number of fixed ports on the switch.

Device Models



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- All the S1700s are Layer 2 switches, some of which provide 100M downstream ports and some provide GE downstream ports. You can distinguish these switches from their product names. The switches with a "G" in their product names have GE downstream ports, for example, S1700-52GFR-4P-AC. The switches without "G" in their product names have 100M downstream ports.
- All the S2700s are Layer 2 100M switches.
- All the S3700s are Layer 3 100M switches.
- The S5700-LI, S5700S-LI and S5710-LI series of the S5700s are Layer 2 GE switches (switches with "LI" in the name are Layer 2 switches), and the rest of the S5700s are Layer 3 GE switches.
- The S6700s are Layer 3 10GE switches.



Product Positioning

S2700	• For access layer, provides large capacity, high port density, and cost-effective Forwarding performance capabilities.
S3700	• For access layer, provides large capacity, high port density, and cost-effective Forwarding performance capabilities.
S5700	• For access layer or aggregation layer, provide all 1000M ports.
S6700	• For reliable access and high-quality transmission of multiple services on the enterprise network and the data center network. Provide all 10GE ports

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- S2700 positioned for the access layer of enterprise network.
- S5700 positioned for the access layer or aggregation layer of enterprise network.
- The Quidway S6700 Series series Ethernet switches (hereinafter referred to as the S6700) provide the access, aggregation, and data transport functions. They are developed by Huawei to meet the requirements for reliable access and high-quality transmission of multiple services on the enterprise network and the data center network.
- SX7 series switches provide large capacity, high port density, and cost-effective Forwarding performance capabilities. In addition, the SX7 swithes provide multi-service access capabilities, excellent extensibility, quality of service (QoS) guarantee, powerful multicast replication, and carrier-class security, and can be used to build ring topologies of high

Fixed Switches' Network Locations

Recommended deployment locations	S2700	S3700	S5700	S6700
SOHO/Terminal	Access	Access	-	-
Small campus	Access	Access	Aggregation layer/Access layer	-
Medium campus	-	-	Access layer	Aggregation layer
Large campus	-	-	Access layer	Access layer
Data center	-	-	-	Access layer

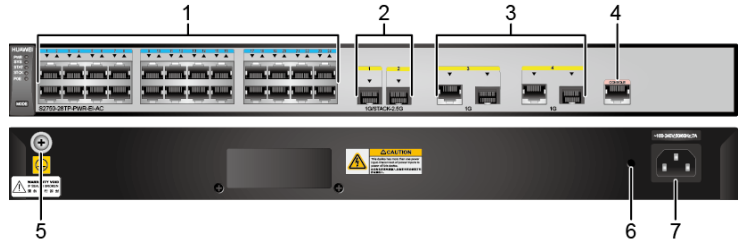
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S2700 Product Positioning

- The S2700 series Ethernet switches (S2700 for short) are next-generation energy-saving 100M Ethernet intelligent switches.
- The S2700 utilizes cutting-edge switching technologies and Huawei Versatile Routing Platform (VRP) software to meet the demand for multi-service provisioning and access on Ethernet networks. It is easy to install and maintain and supports flexible VLAN deployment, comprehensive security and QoS policies, and energy-saving technologies. These features help enterprise customers build a next-generation IT network.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S2750 Appearance and Structure



1	Twenty-four 10/100BASE-TX electrical ports	**2**	Two 1000BASE-X optical ports
3	Two combo ports (10/100/1000BASE-T + 100/1000BASE-X)	**4**	One console port
5	Ground screw NOTE: It is used with a ground cable.	**6**	Jack reserved for AC terminal locking latch.
7	AC power socket NOTE: It is used with an AC power cable.	**-**	-

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

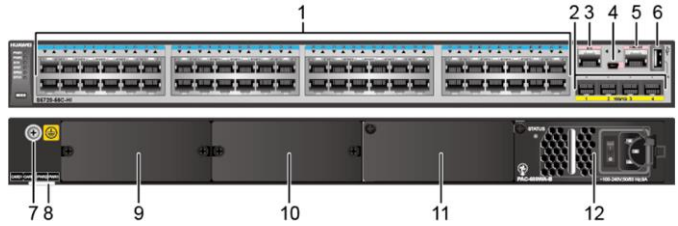
- Two 1000BASE-X ports Applicable modules:
  - GE optical module, GE-CWDM optical module, GE-DWDM optical module, GE copper module, Stack optical module, 1 m and 10 m SFP+ copper cables, 3 m and 10 m AOC cables
- Two combo ports (10/100/1000BASE-T + 100/1000BASE-X) Modules applicable to combo optical ports:
  - FE optical module, GE optical module, GE-CWDM optical module, GE-DWDM optical module
- In V200R006C10 and later versions, you can hold down this button for 6s and release it to start the web initial login mode: If the switch has no configuration file, the system attempts to enter the web initial login mode. In this mode, the status of mode indicators is as follows:
  - If the system enters the web initial login mode successfully, all mode indicators turn green and stay on for a maximum of 10 minutes.
  - If the system fails to enter the initial login mode, all mode indicators fast blink for 10 seconds and then restore to the default status.
- If the switch has a configuration file, the system cannot enter the web initial login mode. In this case, all mode indicators fast blink for 10s, and then return to the default states.

S5700 Product Positioning

- The S5700 series ethernet switches (S5700 for short) are next-generation energy-saving switches developed by Huawei to meet the demand for high-bandwidth access and Ethernet multi-service aggregation. Based on cutting-edge hardware and Huawei Versatile Routing Platform (VRP) software, the S5700 provides a large switching capacity, high reliability (double power slots and hardware Ethernet OAM), and high-density GE ports to accommodate 10 Gbit/s upstream transmissions. It also supports Energy Efficient Ethernet (EEE) and iStack. The S5700 can be used in various enterprise network scenarios. For example, it can function as an access or aggregation switch on a campus network, a gigabit access switch in an Internet data center (IDC), or a desktop switch to provide 1000 Mbit/s access for terminals.
- The S5700 is available in a lite (LI) series, a standard (SI) series, an enhanced (EI) series, and a hyper (HI) series.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S5720-HI Appearance and Structure



1	Forty-eight 10/100/1000BASE-T Ethernet electrical ports	**2**	Four 10GE SFP+ Ethernet optical ports
3	One ETH management port	**4**	One Mini USB port
5	One console port	**6**	One USB port
7	Ground screw NOTE: It is used with a ground cable.	**8**	Bar code label
9	Extended card slot 1 NOTE: This slot is reserved for a stack card.	**10**	Extended card slot 2
11	Power module slot 2	**12**	Power module slot 1

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- S5720-HI Series Agile Fixed Switches
- Fully programmable, energy-efficient Gbit/s access switches for building high-density, agile Ethernet networks.
- Innovative virtualization technology and specialized electronics greatly simplify management of converged, wired and wireless networks, provide more granular quality monitoring and error recovery, and enable rapid provisioning of new services and network features.
- Available in 24-port and 48-port models with 10 GE uplink ports enabling comprehensive services processing capabilities.

Performance Differences among S5700 Versions

- The richness of functions and features provided by these versions is in the following order: S5700LI < S5700SI < S5700EI < S5700HI.
- The following table lists their support for differences features.

Supported Feature	Switch Model			
	S5700LI	S5700SI	S5700EI	S5700HI
RIP/RIPng	N	Y	Y	Y
OSPF/BGP/PIM	N	N	Y	Y
MPLS/Netstream/Hardware-based Ethernet OAM/BFD	N	N	N	Y

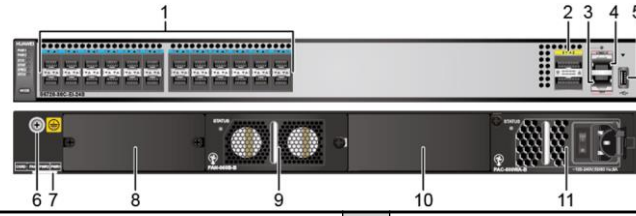
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S6700 Product Positioning

- The S6700 series ethernet switches (S6700 for short) are next-generation 10G box switches. The S6700 can function as an access switch in an Internet data center (IDC) or a core switch on a campus network.
- The S6700 has industry-leading performance and provides up to 24 or 48 line-speed 10GE ports. It can be used in a data center to provide 10 Gbit/s access to servers or function as a core switch on a campus network to provide 10 Gbit/s traffic aggregation. In addition, the S6700 provides a wide variety of services, comprehensive security policies, and various QoS features to help customers build scalable, manageable, reliable, and secure data centers.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S6720 Series Switches



1	Twenty-four 10GE SFP+ Ethernet optical ports	**2**	Two 40GE QSFP+ optical ports
3	One ETH management port	**4**	One console port
5	One USB port	**6**	Ground screw NOTE: It is used with a ground cable.
7	Equipment serial number (ESN) label	**8**	Extended card slot
9	Fan module slot	**10**	Power module slot 2
11	Power module slot 1	**-**	-

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- S6720 Series Next-Generation Enhanced 10 GE Switches
- The industry's highest-performing fixed switches, the S6720 series provides 24/48 full line-speed 10 GE ports, which are scalable to 6 x QSFP+ full line-speed ports.
- The S6720 supports long-distance stacking with up to 480 Gbit/s bidirectional stack bandwidth. It also supports 1+1 backup of AC and DC power modules that can be installed on the same device.
- These switches offer various service features, supports comprehensive security policies and QoS capabilities, and are best suited for data center servers and the core campus network.

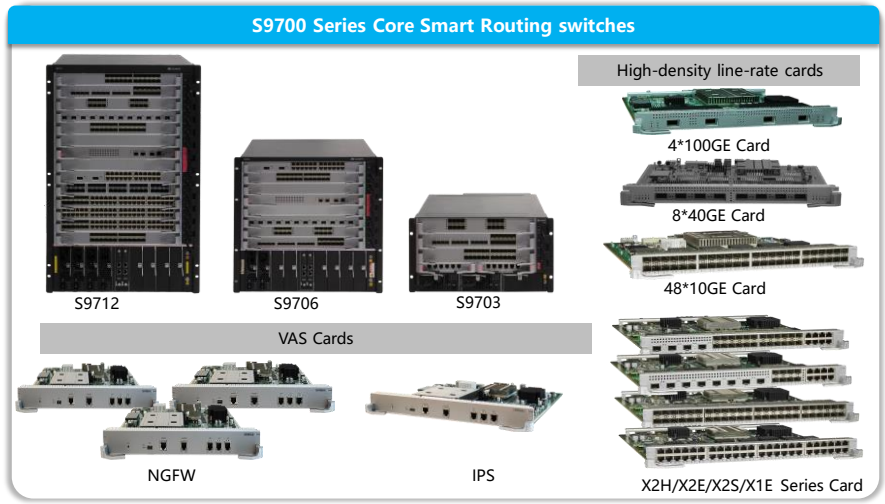
Contents

1. Overview of S Series Fixed Switches
2. **Overview of S Series Modular Switches**
3. Sub-cards and Modules
4. Product Features

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

S9700 Series Core Smart Routing switches

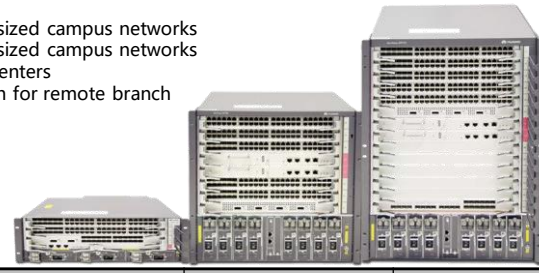


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

S9700 Series Core Smart Routing switches

Positioning

- Core switch for large-sized campus networks
- Core switch for large-sized campus networks
- Core switch for data centers
- Interconnection switch for remote branch offices



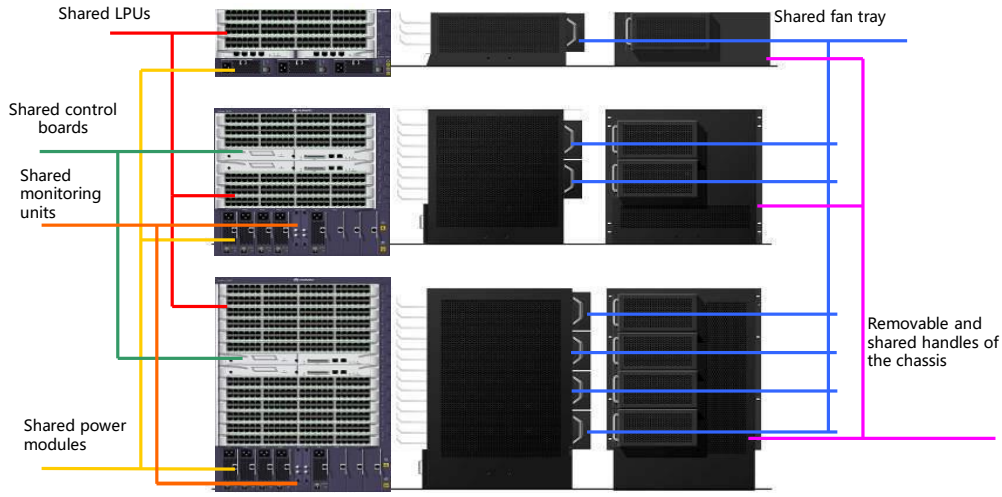
Item	S9703	S9706	S9712
Height	4U	10U	15U
MPU slots	2	2	2
LPU slots	3 horizontal	6 horizontal	12 horizontal
Maximum port density	144GE/144*10GE/24*40GE/12*100GE	288GE/288*10GE/48*40GE/24*100GE	576GE/576*10GE/96*40GE/48*100GE
System power supply	1+1	M+N	M+N

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- S9700 Series Switch is design for integrated multi-service network architecture, It is a high-end intelligent terabit routing switch.
- S9700 provides 16x10GE ports inter-board wire speed switching, and supports 40GE/100GE standards in the future.

Modular Design



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The fan trays, AC power supplies, DC power supplies, LPUs, cables, and cabinet handles can be used by all types of the switch. The handles can be removed from the cabinet.
- The SXX12 and the SXX06 shared the monitoring boards and control boards of the same type.

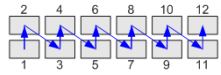
Contents

1. Overview of S Series Fixed Switches
2. Overview of S Series Modular Switches
3. **Sub-cards and Modules**
4. Product Features

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Port Numbering Conventions

- A single switch uses slot ID/subcard ID/port sequence number to identify physical ports.
  - Slot ID: indicates the slot where the switch is located. The value is 0.
  - Subcard ID: indicates the ID of a subcard.
  - Port sequence number: indicates the sequence number of a port on the switch.
- A stacked switch uses Stack ID/subcard ID/port sequence number to identify physical ports.
  - Stack ID: indicates the ID of a stacked switch. The value ranges from 0 to 8.
  - Subcard ID: indicates the ID of a subcard.
  - Port sequence number: indicates the sequence number of a port on the switch.

Port Numbering Diagram	Description
	There are two rows of service ports on the device. These ports are numbered from bottom to top and left to right, starting from 1.For example, the port on the top left is numbered 0/0/2.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Power Modules

Switch Series	Power Supply Configuration
S2700	It has a built-in power module and does not support pluggable power modules.
S5700	It uses pluggable power modules: 150/350/650W DC Power Module; 650W DC PoE Power Module; 150/600W AC Power Module; 500/580/1150W AC PoE Power Module.
S6700	It uses pluggable power modules: 170/350W DC Power Module; 170/600W AC Power Module.

- All power modules are hot swappable, but it is highly recommended that you power off a switch before removing or installing a power module in the switch to protect personal and equipment safety.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

- Before replacing a power module in a switch, make sure that the switch can be powered by the other power module after the power module is removed. Otherwise, services on the switches will be interrupted by a power failure when the power module is removed.
- Before powering off a switch, shut down all of its power supply units.
- The S5720-HI models that do not support Power over Ethernet (PoE) can use 350 W DC and 600 W AC power modules together. The S5710-HI series can use 350 W and 1150 W power modules together. The S5720-28X-PWR-SI-AC, S5720-52X-PWR-SI-AC, S5720-28X-PWR-SI-DC, S5720-52X-PWR-SI-DC, S5720-36C-PWR-EI-AC, S5720-36C-PWR-EI-DC, S5720-56C-PWR-EI-DC, and S5720-56C-PWR-EI-AC can use 500 W AC PoE and 650 W DC PoE power modules together. Other models do not allow power modules of different power values to be used in the same chassis.
- The S6720-EI can use 350 W DC and 600 W AC power modules together. Other models do not allow power modules of different power values to be used in the same chassis.

PoE Function

Series	Maximum Number of PoE Interfaces (IEEE 802.3af)	Maximum Number of PoE Interfaces (IEEE 802.3at)
S2700	8/16/24	4/8/12
S5700	8/12/16/24/48	4/6/8/12/24/26/48
S6700	Not supporting PoE	

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Switches with PWR in the name support PoE power supply, such as the S5710-52C-PWR-EI.
- PoE switches provide power for powered devices (PDs) over Ethernet electrical interfaces. All the PoE switches comply with IEEE 802.3af and 802.3at. IEEE 802.3af supports a maximum of 15.4 W power and the IEEE 802.3at supports a maximum of 30 W power. The PDs connected to a switch determine which standard the switch should comply with, and the switch is auto-sensing.
- The number of interfaces that can provide PoE power supply on a switch depends on the power module used, the corresponding standard, and the switch's own limitations. Here, I'm providing the maximum number of interfaces that each series can support theoretically. See the Hardware Description of the corresponding product for details.

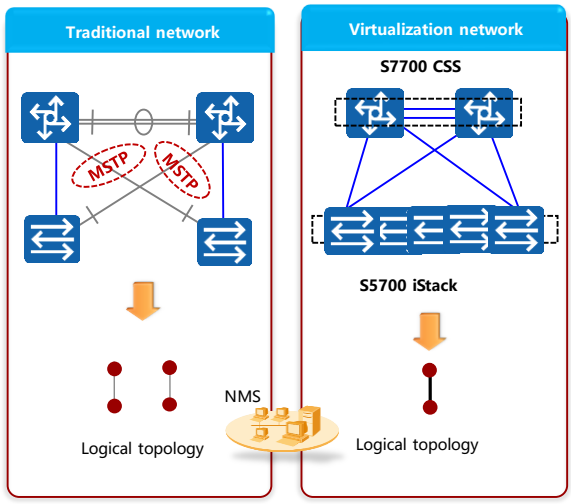
Contents

1. Overview of S Series Fixed Switches
2. Overview of S Series Modular Switches
3. Sub-cards and Modules
4. **Product Features**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

iStack Benefits



Improve bandwidth efficiency and reduce CAPEX

- Ring protection protocols such as MSTP are not required, and no link needs to be blocked.
- 100% of bandwidth is used. (Only 50% of bandwidth is used on an STP network).

Build a highly reliable loop-free network

- The convergence time is much shorter than STP.
- The system continues running if a single device fails.

Simplify O&M and reduce OPEX

- Fewer network nodes are deployed, simplifying network management.
- The master switch synchronizes the configuration file to other member switches, simplifying device operation and configuration.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Stack Card Stacking and Service Port Stacking

Stack card stacking



- **Stack card connection**

  Two situations exist:
  - Member switches are connected using **dedicated stack cards** and stack cables.
  - Stack cards are integrated on the switch' s rear panel. Member switches are connected using stack ports fixed on the rear panel and stack cables.

Service port stacking

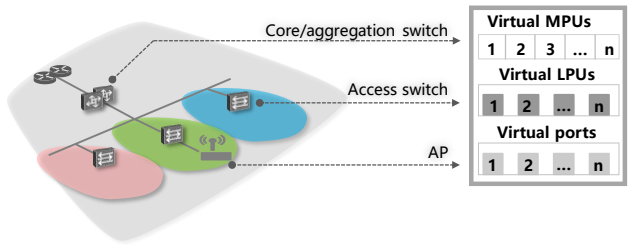


- **Service port connection**

  Member switches are connected using service ports, which are configured as physical member ports and bound to logical stack ports. This connection mode does not require stack cards.
  A logical stack port is bound to physical member ports to connect stack member switches. Each member switch supports two logical stack ports.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SVF Manages a Campus Network as One Device



Converged campus network

- Cloud structure, on-demand expansion
- Campus network virtualized into **one** device
- Access switches and APs are virtualized as extended ports on the virtual switch

Professional capabilities

- Virtualize **32** access switches, which triples industrial average
- **Unprecedentedly** support virtualization of 1K APs, simplifying network maintenance and management

- SVF manages a campus network as one device
  - Simple management: The devices on the entire network are virtualized into one, and devices are plug-and-play. Only one NE needs to be managed on the network.
  - Centralized management: Core devices manage ACL rules, QoS policies, and user security.
  - Distributed forwarding: Each device on the wired network can independently perform control and forwarding, avoiding bandwidth bottlenecks.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Flexible Ethernet Networking: Stability and Reliability



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Quiz

1. What is the meaning of each section of the switch’ s name: S5720-56C-PWR-EI-AC?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What is the meaning of each section of the switch’ s name: S3728TP-PWR-EI?
  - S: Switch
  - 57: Series
  - 20: Product sub-series
  - 56:the maximum port quantity
  - C: The product supports extended cards and its uplink ports are provided by an extended card or are fixed 10GE ports.
  - PWR: The product supports Power over Ethernet (PoE).
  - EI: enhanced version
  - AC: switch using alternating current power supply

Summary

- SX7 fixed switch product positioning
- SX7 fixed switch sub-cards and modules
- SX7 fixed switch product characteristics
- SX7 fixed switch typical application scenarios

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Introduction to Huawei Data Center CE Series Switches

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- The CloudEngine 12800 series modular switches are next-generation high-performance core switches designed for data center networks and high-end campus networks, which provide high-density 40GE/100GE line cards and support various data center features such as VXLAN, EVPN, and M-LAG.
- The CloudEngine 5800, 6800, 7800, and 8800 series fixed switches are next-generation high-performance, high-density, and low-latency Ethernet switches designed for data center networks, which use flexible front-to-rear/rear-to-front airflow design.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Understand the current development of data center switches.
  - Distinguish hardware types of Huawei CE series switches.
  - Understand networking applications of Huawei CE series switches.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Product Positioning**
2. Product Structure
3. Cards and Modules
4. Product Features and Application Scenarios

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Challenges to Cloud Data Center Networks

Big Data requires large pipes.

Every minute

1.6 million Google search requests
260 million emails are sent.
47,000 apps are downloaded.
220,000 photos are uploaded to Facebook.
660 million IP packets are transmitted.



IP traffic on data center networks
Increased by 5 times each year

Service innovation requires network agility.



320 **apps/day**

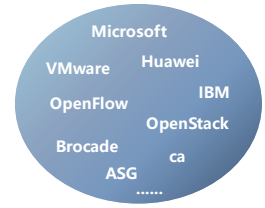


500+ **servers/week**

Data center applications
Increased by 200% every four years

Diversity requires open networks.





Diversified ecosystem
50+ mainstream SDN vendors

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Requirements for Massive Data Channels in the Cloud Era



Big Data Era

Of all the data obtained by human civilization, 90% is **generated in the past two years**. By 2020...



Data volume growth — x 50

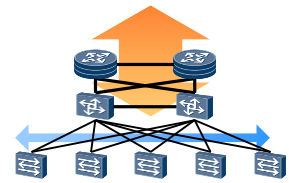


Number of servers — x 10

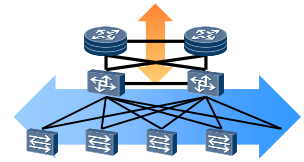


Server port rate — x 100

Evolution of servers on data center networks:
GE -> 10G -> 25G/40G -> 50G -> 100G



In a traditional data center, 80% of traffic is north-to-south traffic.



In a cloud computing data center, 70% of traffic is east-to-west traffic.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CE Data Center Switch Portfolio (1)

Core Switch

CE12800



CE12816 CE12812 CE12808 CE12804

CE12800S



CE12808S CE12804S

Virtual Switch



CE1800V

Access Switch

ToR switch with flexible cards



CE8861-4C-EI

100GE switch

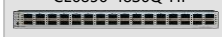


CE8850-64CQ-EI

40GE switch



CE6856-48S6Q-HI



CE7855-32Q-EI

25GE ToR switch

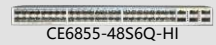


CE6865-48S8CQ-EI

10GE ToR switch



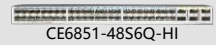
CE6856-48T6Q-HI



CE6855-48S6Q-HI



CE6855-48T6Q-HI



CE6851-48S6Q-HI



CE6860-48S8CQ-EI



CE6810-48S4Q-LI



CE6810-32T16S4Q-LI

10GE large-buffer ToR switch



CE6870-48S6CQ-EI



CE6870-48T6CQ-EI



CE6875-48S4CQ-EI

GE ToR switch



CE5855-48T4S2Q-EI



CE5855-24T4S2Q-EI

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CE Data Center Switch Portfolio (2)

Core Switch

CE12800E


CE12816E


CE12808E


CE12804E

Access Switch

10GE Switch


CE6880-48S4Q2CQ-EI


CE6880-24S4Q2CQ-EI


CE6880-48T4Q2CQ-EI


CE5880-48T4Q2CQ-EI*

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CE Switches Help Build a Next-Generation Cloud Data Center Network

Elastic: three-fold capacity of the industry average, helping construct stable network platforms for 10 years



- Industry-leading high-density line cards: **72 x 100GE and 36 x 40GE**
- **178 Tbit/s** super-large capacity, providing **11 Tbit/s bandwidth** per slot
- **24 GB** super-large buffer

Virtual: industry's highest 1-to-16 virtualization capability, improving ICT resource utilization



- Virtual system (VS): **One device can be virtualized into 16 devices.**
- Most comprehensive fabric networking in the industry: SVF, CSS, VXLAN, TRILL, and EVPN

Agile: full openness, accelerating agile innovation of cloud services



- Cooperation with the Agile Controller, implementing E2E SDN solutions
- Network automation, reducing O&M and management costs of customers
- OpenFlow and open APIs, preventing vendor lock-in

High-quality: bearing high-value services and providing high-quality experience



- Industry-leading **orthogonal architecture**
- Patented **front-to-rear airflow design** that isolates cold air channels from hot air channels
- 2-µs extra-low latency, achieving efficient forwarding
- Full hot standby for five hardware systems

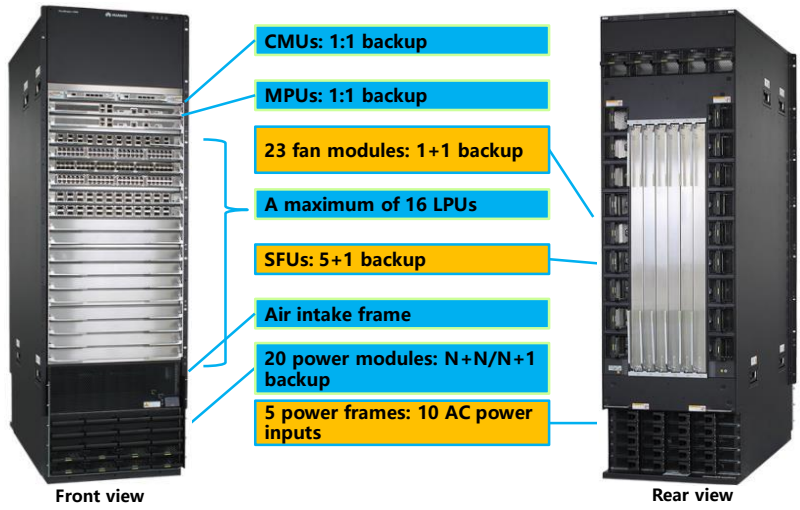
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Product Positioning
2. **Product Structure**
3. Cards and Modules
4. Product Features and Application Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

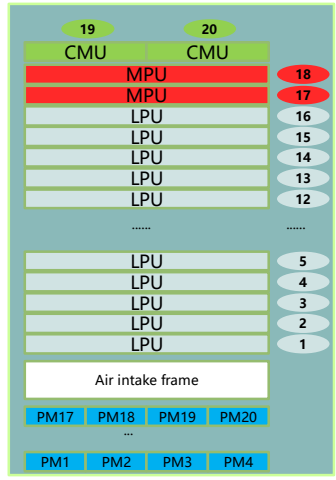
Appearance of the Core Switch CE12816



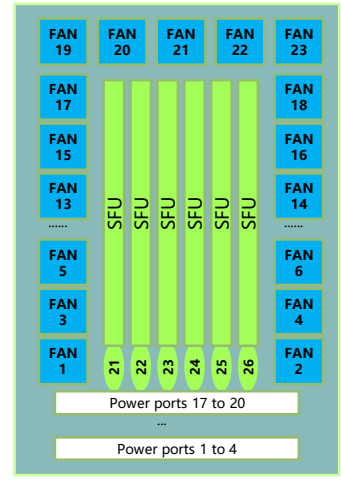
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- CMU: Central Monitoring Unit
- MPU: Main Processing Unit
- LPU: Line Processing Unit
- SFU: Switch Fabric Unit

Slot Distribution on the CE12816



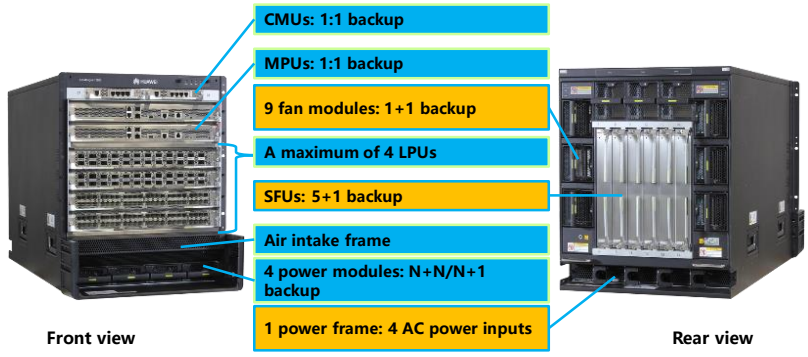
Front view



Rear view

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

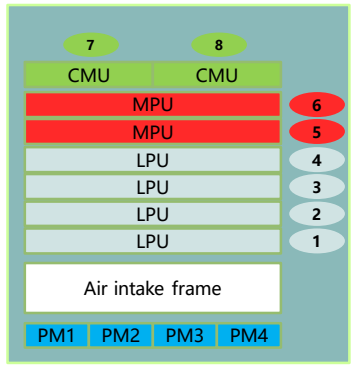
Appearance of the Core Switch CE12804



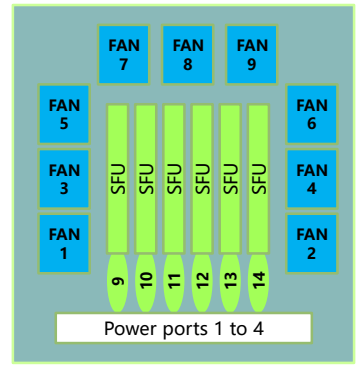
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Slot Distribution on the CE12804



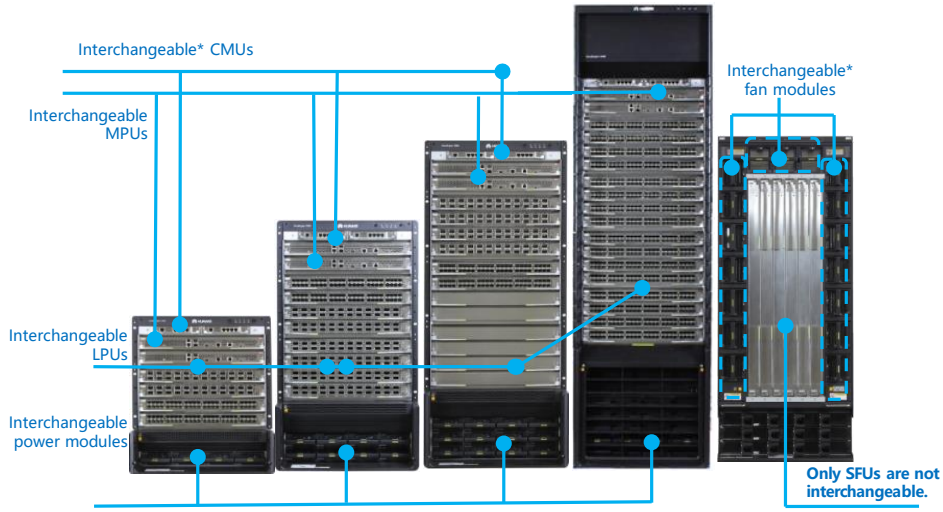
Front view



Rear view

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Systematic and Unified Design

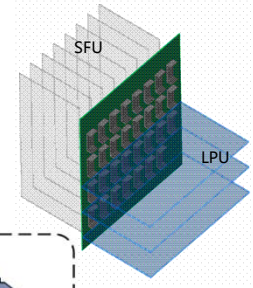


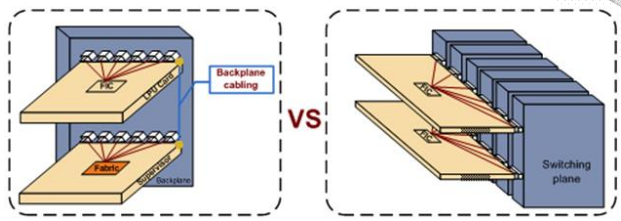
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The CE12800 series switches use the unified design, allowing interchangeable components, such as power modules, fan modules, MPUs, LPUs, and CMUs to be used on all CE12800 switches of different models. That is, all cards and modules except SFUs are interchangeable on the CE12800.
- The CE12804, CE12808, and CE12812 use the same MPUs, CMUs, LPUs, power modules, and fan modules.
- The CE12816 uses the same MPUs, LPUs, and power modules as the CE12804, CE12808, and CE12812 but has its own CMUs and fan modules.
- The CE12804S and CE12808S use the same MPUs, SFUs, and fan modules, and their LPUs and power modules are the same as those on the CE12804, CE12808, CE12812, and CE12816.

Orthogonal Architecture of LPUs and SFUs on the CE12800

Multi-level and multi-plane data switching architecture and unlimited capacity expansion, implementing large-scale non-blocking switching



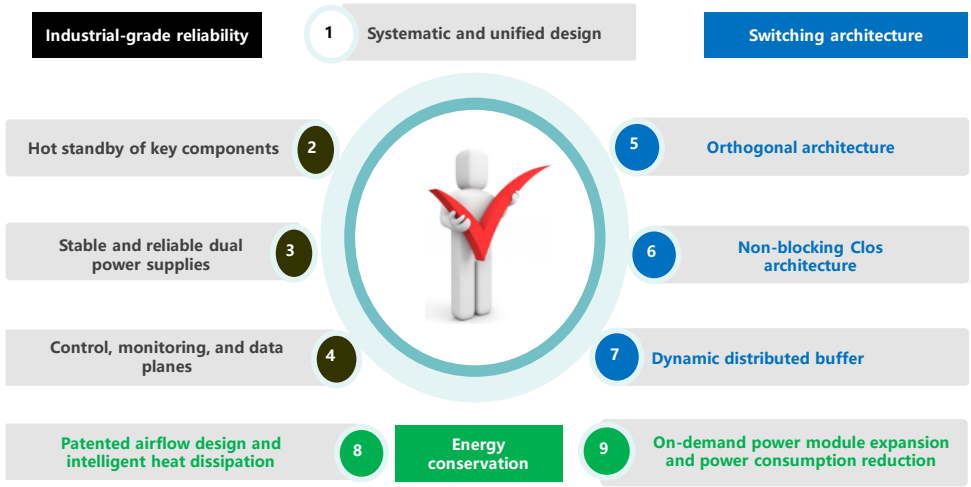


Traditional architecture **Orthogonal architecture**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- On core switches, cabling between line cards and DFUs is an important factor affecting slot bandwidth. A longer backplane cable and a higher rate indicate a greater loss.
- The CE12800 uses an orthogonal architecture, which require no wires on the backplane. This architecture greatly increases system bandwidth and improves the evolution capability. The orthogonal design (three-level Clos architecture) of LPUs and SFUs realizes multi-level and multi-plane data switching architecture and unlimited capacity expansion, implementing large-scale non-blocking switching in data centers.
- The Clos architecture has multiple levels, at each of which a switching unit is connected to all switching units at the lower level.

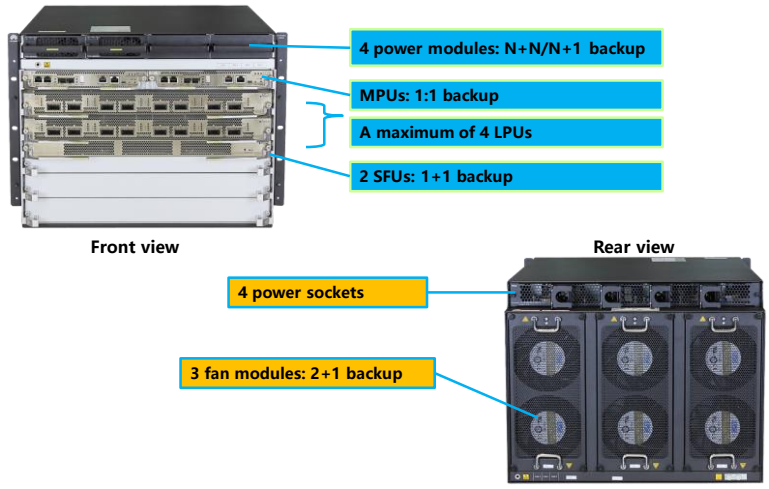
Architecture of CE12800 Series Switches



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The CE12800 provides industrial-grade reliability and supports in-service software upgrade (ISSU) to meet customer requirements for service continuity.
- All key components work in redundancy mode and all modules are hot swappable:
  - MPUs work in 1:1 backup mode.
  - SFUs work in N+M hot backup mode.
  - Power modules support N+N and N+1 backup.
  - Fan modules work in 1+1 backup mode.
  - Each fan module has two counter-rotating fans working in 1+1 backup mode.
- Independent triple-plane design: The control plane, data plane, and monitoring plane of the CE12800 are independent of each other. This design improves system reliability and ensures service continuity.
- Super large buffer on interfaces
- East-west traffic between servers predominates in cloud-computing data centers. The Map-Reduce framework used for processing a large amount of data increases the incast traffic model in which traffic bursts occur frequently. In this situation, the requirements of the incast traffic model must be met.
- The CE12800 series switches use next-generation large-buffer line cards. All service ports (100GE, 40GE, and 10GE ports) support 100 ms buffering. The distributed buffering mechanism on inbound interfaces can effectively handle incast traffic in data centers and absorb burst traffic while providing high-performance forwarding. These advantages enable the CE12800 to effectively handle traffic in the new traffic model.

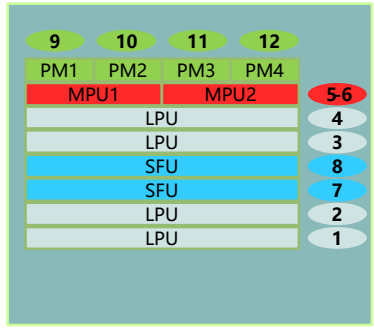
Appearance of the Core Switch CE12804S



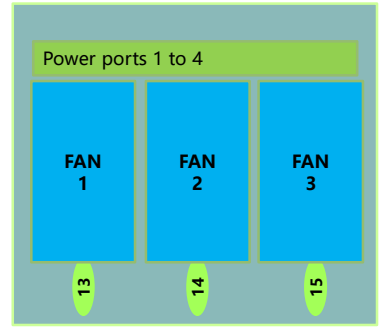
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The MPUs of the CE12800S work in 1+1 redundancy mode.
- The CE12804S supports two SFUs and the CE12808S supports four SFUs. The SFUs back up each other and load balance traffic. Removing an SFU at any time does not affect normal operations of other SFUs.
- Fan modules work in N+1 backup mode. If a fan module fails, the switch still works properly.
- Power modules support N+1/N+N backup. If a power module fails, other power modules work properly.

Slot Distribution on the CE12804S



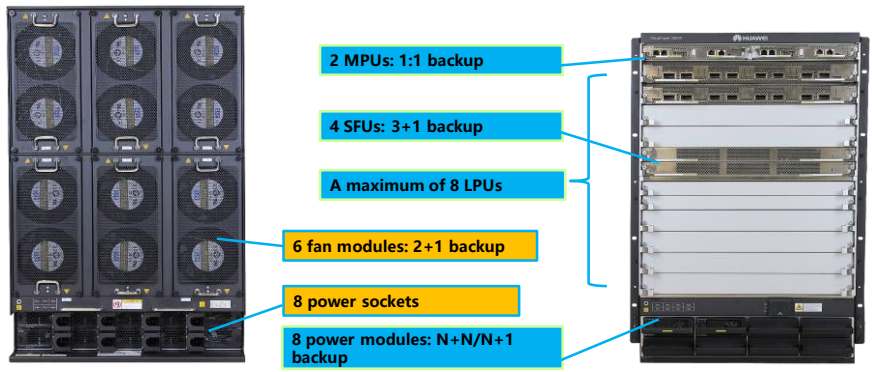
Front view



Rear view

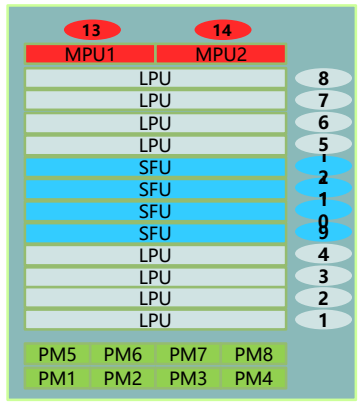
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Appearance of the Core Switch CE12808S

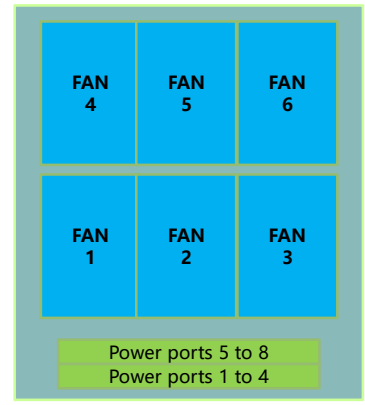


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Slot Distribution on the CE12808S



Front view



Rear view

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Dimensions of CE12800S Series Switches

Item	CE12804S		CE12808S	
Dimensions (W x D x H, mm)	442 x 751 x 352.8 (8 U)		442 x 751 x 708.4 (16 U)	
Chassis weight	Empty: 60 kg	Fully loaded: 120 kg	Empty: 100 kg	Fully loaded: 196 kg

CE12808S



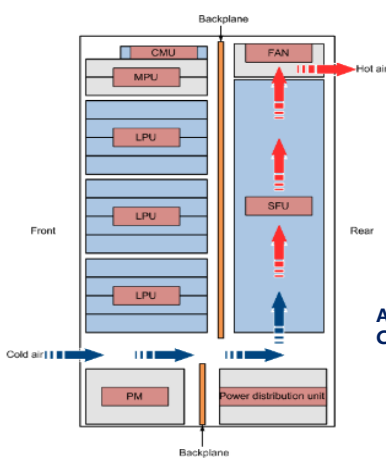
CE12804S



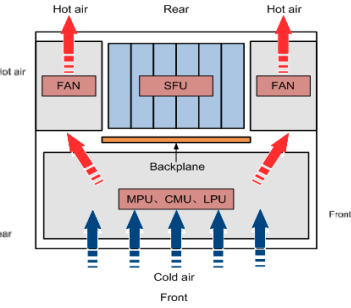
The CE12800S chassis is lighter and has lower requirements on the equipment room/cabinet.
The CE12800S can be installed in a 1 m deep cabinet.
The CE12800S is delivered with cards and packed using a cardboard box and a pallet.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

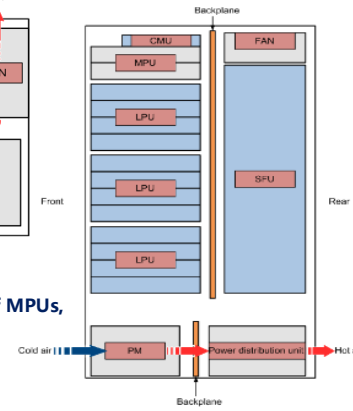
Heat Dissipation Design of the CE12800



Airflow for heat dissipation of SFUs (side view)



Airflow for heat dissipation of MPUs, CMUs, and LPUs (top view)



Airflow for heat dissipation of power modules (side view)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Cards on a switch use independent air channels for heat dissipation. SFUs use bottom-to-top airflows for heat dissipation, as shown in the left figure.
- LPUs, MPUs, and CMUs use front-to-rear airflows for heat dissipation, as shown in the middle figure.
- Fan modules at the same horizontal level is responsible for heat dissipation of corresponding cards and back up each other.

Contents

1. Product Positioning
2. Product Structure
3. **Cards and Modules**
4. Product Features and Application Scenarios

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Card and Module Naming Conventions

Card/Module	Category	Naming Convention
MPU	CE - MPU A CE - CMU A A B C A B C	A: CE series B: card category C: card version
SFU	CE - SFU 04 A A B C D	A: CE series B: card category C: product model (04/08/12/16) D: card version (versions A, B, and C in ascending order of performance)
LPU	CE – L 24 L Q - EC1 A B C D E F	See the next page
Power module	P AC - 2700W A A B C D	A: power module B: power supply type (AC: alternative current; DC: direct current) C: rated power D: power module version
Fan module	FAN - 12C A B	A: fan module B: fan module model

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LPU Naming Conventions

CE - L 24 L Q - EC1
A B C D E F

Field	Meaning	Description
A	Brand name	It is fixed as CE, representing CloudEngine.
B	LPU type	L: LPU for modular switches F: flexible card for modular switches
C	Number of ports on the LPU	For an LPU with different types of ports, this field represents the number of downlink ports. It has two digits. The first digit is 0 if the card has fewer than 10 ports.
D	Port rate	G: GE port X: 10GE port Y: 25GE port L: 40GE port C: 100GE port
E	Port type	T: Base-T port S: SFP/SFP+ port X: XFP port Q: QSFP+/QSFP28 port F: CFP/CFP2/CFP4 port C: CXP port
F	LPU specifications	BA/EA/EC/EC1: basic specifications ED/EF/EG/FD: enhanced specifications

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Main Cards on the CE12800 Series Switches

Category	Name	Overview
MPU	CE-MPUA	MPU for the CE12804/CE12808/CE12812/CE12816 chassis
	CE-MPUA-S	MPU for the CE12804S/CE12808S chassis, responsible for system control, management, and monitoring
SFU	CE-SFU04A	Applicable to the CE12804 chassis, responsible for line-rate data switching on the data plane
	CE-SFU04B	Applicable to the CE12804 chassis, responsible for line-rate data switching on the data plane
	CE-SFU16C	Applicable to the CE12816 chassis, responsible for line-rate data switching on the data plane
	CE-SFUA-S	Applicable to the CE12804S/CE12808S chassis, responsible for line-rate data switching on the data plane
	CE-SFUF-S	Applicable to the CE12804S/CE12808S chassis, responsible for line-rate data switching on the data plane
CMU	CE-CMUA	CMU for the CE12804/CE12808/CE12812 chassis
	CE-CMUB	CMU for the CE12816 chassis, responsible for device monitoring, management, and energy saving
LPU	CE-L48GT-EA	48-port 10/100/1000BASE-T interface card (EA, RJ45)
	CE-L48GS-EA	48-port 100/1000BASE-X interface card (EA, SFP)
	CE-L12XS-ED	12-port 10GBASE-X interface card (ED, SFP+)
	CE-L24XS-EC	24-port 10GBASE-X interface card (EC, SFP+)
	CE-L48XT-EC	48-port 100M/1000M/10GBASE-T interface card (EC, RJ45)
	CE-L36LQ-EG	36-port 40GE optical interface card (EG, QSFP+)
	CE-L36CQ-FD	36-port 100GE optical interface card (FD, QSFP28)
	...	...
Value-added service card	CE-FWA	40G NGFW module
	CE-IPSA	IPS module

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

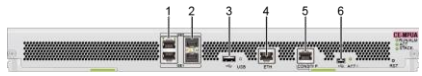
Functions of High-Performance MPUs on the CE12800/CE12800S

Function and Feature	Description
Device management and maintenance	The CE-MPUA/CE-MPUAS provides management ports (such as a console port) for operators to manage and maintain the device.
Out-of-band communication between cards	The CE-MPUA/CE-MPUAS integrates a LAN switch module that provides out-of-band communication between cards. The LAN switch module completes control, maintenance, and message exchange between CMUs, SFUs, and LPUs.
Route calculation	• The CE-MPUA/CE-MPUAS processes all routing protocol packets, which are sent from the forwarding engine. • The CE-MPUA/CE-MPUAS broadcasts and filters packets, and downloads routing policies from the policy server.
Data configuration	The CE-MPUA/CE-MPUAS stores configuration data, startup files, upgrade software, and system logs.
Data saving	The CE-MPUA/CE-MPUAS provides an embedded USB (eUSB) module as a storage device to save data files.
Device monitoring	The CE-MPUA-S integrates a monitoring module. The monitoring module provides the monitoring plane, which allows administrators to remotely power on, power off, and reset cards, upgrade firmware, monitor card temperature, voltage, and power, manage asset information, and diagnose system faults.

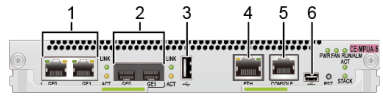
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

High-Performance MPU Ports on the CE12800/CE12800S

CE-MPUA



CE-MPUA-S

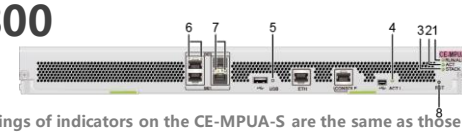


No.	Port Description
1	Two GE electrical ports
2	Two GE optical ports
3	One USB host port
4	One Ethernet port
5	One console port
6	One Mini USB port

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Indicators on the High-Performance MPU of the CE12800

CE-MPUA



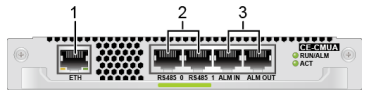
*The meanings of indicators on the CE-MPUA-S are the same as those on the CE-MPUA.

No.	Indicator/Button	Color	Description
1	RUN/ALM: running status indicator	Green	• Steady on: The card has been powered on but the system software is not running. • Slow blinking: The card is running properly. • Fast blinking: The card is loading the system software, resetting, or it is used as the standby MPU and is performing batch data backup.
		Red	Steady on: A fault that affects services has occurred and requires manual intervention, or the card has generated an alarm because the memory size is not equal to the standard specification.
		Yellow	Steady on: The card has been installed in the chassis and the CANbus has been powered on.
2	ACT: active/standby status indicator	Green	• Steady on: The card is the active MPU. • Off: The card is the standby MPU.
3	STACK: stack status indicator	Green	• Blinking: The card is not the active MPU of the stack system. • Steady on: The card is the active MPU of the stack. • Off: The stacking function is not enabled.
4	ACT: Mini USB port indicator	Green	• Steady on: The Mini USB port is active, and the console port cannot be used. • Off: The Mini USB port is inactive, and the console port can be used.
5	USB: USB-based deployment indicator	This indicator is reserved for the USB-based deployment function and will be on only when the USB-based deployment function is used.	
6	One single-color indicator for each port	Green	• Steady on: A link has been established on the port. • Off: The link on the port is disconnected.
		Yellow	• Blinking: The port is transmitting and receiving data. • Off: The port is not transmitting or receiving data.
7	Two single-color indicators for each port	Green	• Steady on: A link has been established on the port. • Off: The link on the port is disconnected.
		Yellow	• Blinking: The port is transmitting and receiving data. • Off: The port is not transmitting or receiving data.
8	RST: Reset button	This button is used to manually reset an MPU. Exercise caution when you press this button.	

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CMUs on the CE12800

Function and Feature	Description
Next-generation monitoring and management architecture	The CMU decouples the monitoring plane from the service plane. Therefore, it can still report device running status and fault recovery events when the service plane fails. With the CMU, the switch supports zero touch device management and maintenance.
Innovative and intelligent management	The CMU can work with the data center management system to realize intelligent energy allocation in a data center and energy saving for the air conditioning system. • Intelligent fan speed adjustment: The CMU monitors switch and cabinet temperature in real time and adjusts fan speeds according to the switch temperature. This reduces power consumption of fans and prevents the switch from overheating. • Intelligent power management: The CMU monitors power required by the switch and reports it to the management system in the equipment room. In this way, power supplied to the switch can be dynamically adjusted to fully use the designed capacity of the power distribution system and avoid waste of power.
Highly reliable monitoring platform	The CMU supports 1:1 hot standby.
All-round remote monitoring	• The CMU uses industry-leading monitoring system on chip (SoC) technology to provide powerful out-of-band monitoring, management, and maintenance for cards. • The monitoring plane allows administrators to remotely power on, power off, and reset cards, upgrade firmware, monitor card temperature, voltage, and power, manage asset information, and diagnose system faults.

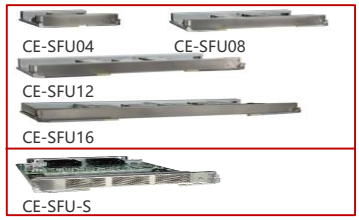


CE-CMUA

No.	Port Description
1	One Ethernet port
2	Two RS485 ports
3	Two IN/OUT ports

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SFUs on the CE12800



- Function and Feature
  - Line-rate data switching: The CE12800 uses six CE-SFUs. The CE12800S uses two or four CE-SFU-S cards to form the switching core of the data plane and provide high-speed SerDes channels for LPUs.
  - Reliability: The SFUs use a single-level multi-plane switch fabric to expand the switching capacity. The system has six switching planes, which work in load balancing mode to ensure nonstop service data transmission.



CE-SFU04

No.	Indicator	Color	Description
1	RUN/ALM: running status indicator	Green	• Steady on: The card has been powered on but the system software is not running. • Slow blinking: The card is running properly. • Fast blinking: The card is loading the system software or is resetting.
		Red	Steady on: A fault that affects services has occurred. The fault cannot be rectified automatically and requires manual intervention.
		Yellow	Steady on: The card has been installed in the chassis and the CANbus has been powered on.
2	OFL: Offline indicator	To remove an SFU, hold down the OFL button for 6s. You can remove the SFU until the OFL indicator is steady red.	
3	OFL button		

*The meanings of indicators on the CE-SFU-S are the same as those on the CE-SFU.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

GE/10GE/40GE Interface Card on the CE12800/CE12800S



CE-L48GT series (RJ45)
48-port 10/100/1000BASE-T interface card



CE-L48XT-EC (EC, RJ45)
48-port 100M/1000M/10G BASE-T interface card



CE-L48GS series (SFP)
48-port 100/1000BASE-X interface card



CE-L02LQ-EC (EC, QSFP+)
2-port 40GBASE-X interface card



CE-L12XS-ED (ED, SFP+)
12-port 10GBASE-X interface card



CE-L36LQ series (QSFP+)
36-port 40GBASE-X interface card

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

40GE Line Card on the CE12800E



New line card on the CE12800E

Name: CEL24LQED-E

Ports: 24 x 40GE (default)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

A forwarding chip is removed from this card on the basis of the CEL36LQED-E.

100GE Interface Cards on the CE12800/CE12800S



CE-L04CF series (CFP)
(4-port 100GE optical interface card)



CE-L12CF-EG (EG, CFP2)
(12-port 100GE optical interface card)



CE-L08CC-EC (EC, CXP)
(8-port 100GE optical interface card)

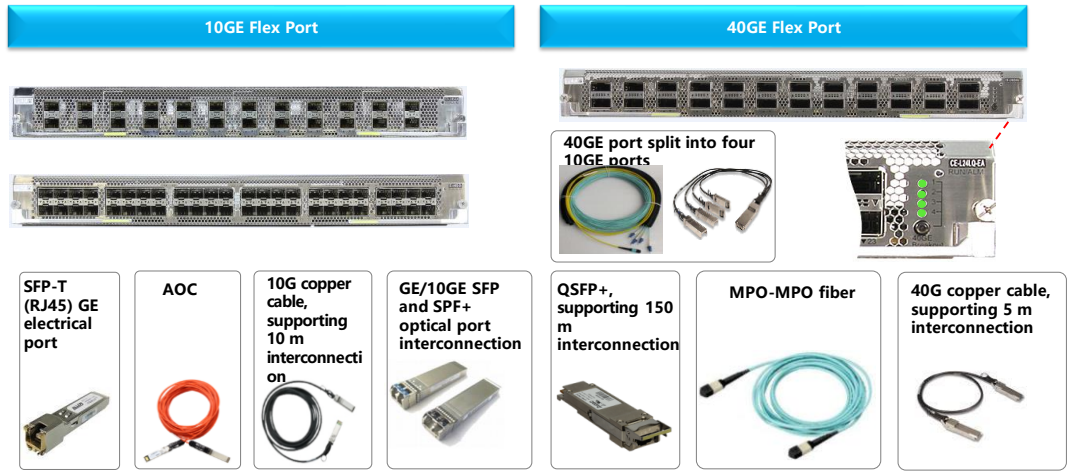


CE-L36CQ-FD (FD, QSFP28)
(36-port 100GE optical interface card)

- SFUA, SFUB, or SFUC cannot be used in the same chassis with FD or FDA series LPUs. To use FD or FDA series LPUs in a chassis, replace SFUA, SFUB, or SFUC with SFUF or SFUG.
- Due to the limitation of output power, when 2200 W DC power modules work in N+N backup mode, a CE12804 chassis can be configured with a maximum of three CE-L36CQ-FD cards, a CE12808 chassis can be configured with a maximum of six cards, and a CE12812 chassis can be configured with a maximum of 10 cards.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

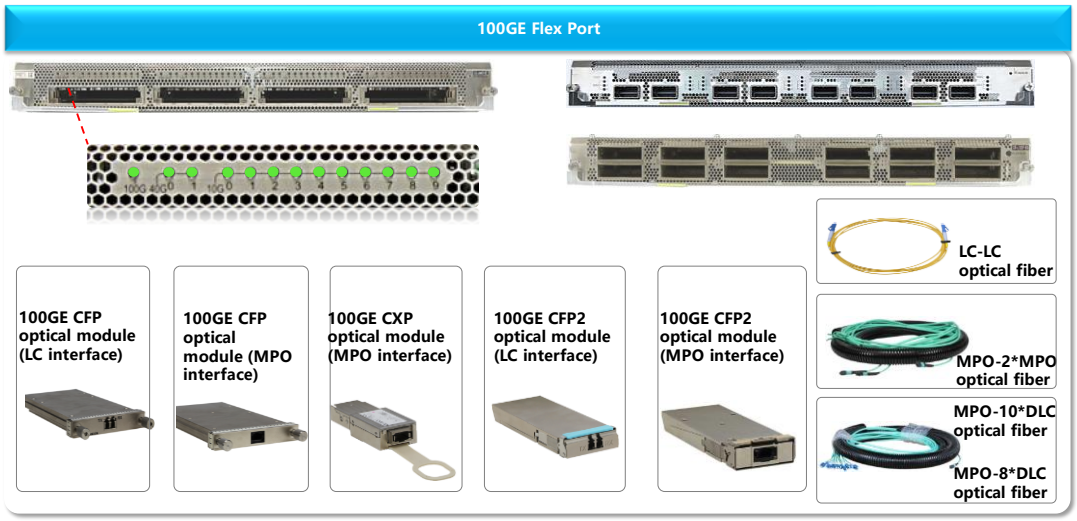
Different Types of Flex Ports



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- CE series switches support different types of ports for flexible applications.
- 10GE optical ports support GE transceivers and auto-sensing. The port speed can change automatically when a GE transceiver is installed. The 10GE optical ports can also connect to copper cables when they have copper transceivers installed.
- 40GE ports can connect to Multi-fiber Push On (MPO) optical fibers. MPO optical fibers do not differentiate Tx and Rx ports. The physical ports can prevent the MPO connectors from being reversely inserted to the ports.
- 40GE ports can connect to copper cables. A 40GE port can be split into four 10GE ports. After configuring 40GE port splitting on an LPU, you need to restart the LPU to make the configuration take effect.
- After a 40GE optical port is split into four 10GE ports, the original 40GE optical port does not work. The new 10GE ports support the same configurations and features as common 10GE optical ports, except that their numbers are different from common 10GE optical ports. The split 40GE port can be connected to the peer device using a dedicated 1-to-4 cable. After a 40GE port is split, the split ports can be used as stack ports, and the indicator shows the status of a 10GE port. The sequence number of the indicated 10GE port is identified by indicators 1, 2, 3, and 4.
- After the 40GE port is split into four 10GE ports, the 10GE ports are numbered in the 40GE x/y/N:M(10GE) format, where:
- The values of x, y, and N are the same as those in the 40GE port number.
- The value of M is 1, 2, 3, or 4.

Multiple 100GE Cards Support Different Types of Ports



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- After the 100GE port is split into 40GE ports, the 40GE ports are numbered in the 100GE x/y/N:M(40GE) format, where:
- The values of x, y, and N are the same as those in the 100GE port number.
- The value of M is one of the following:
  - If the 100GE port is split into two 40GE ports, M is 1 or 2.
  - If the 100GE port is split into three 40GE ports, M is 1, 2, or 3.
- After the 100GE port is split into 10GE ports, the 10GE ports are numbered in the 100GE x/y/N:M(10GE) format, where:
- The values of x, y, and N are the same as those in the 100GE port number.
- The value of M is one of the following:
  - If the 100GE port is split into eight 10GE ports, M is an integer in the range from 1 to 8.
  - If the 100GE port is split into ten 10GE ports, M is an integer in the range from 1 to 10.
  - If the 100GE port is split into twelve 10GE ports, M is an integer in the range from 1 to 12.

Power Modules on the CE12800 and CE12800S



AC: 2700 W

DC: 2200 W

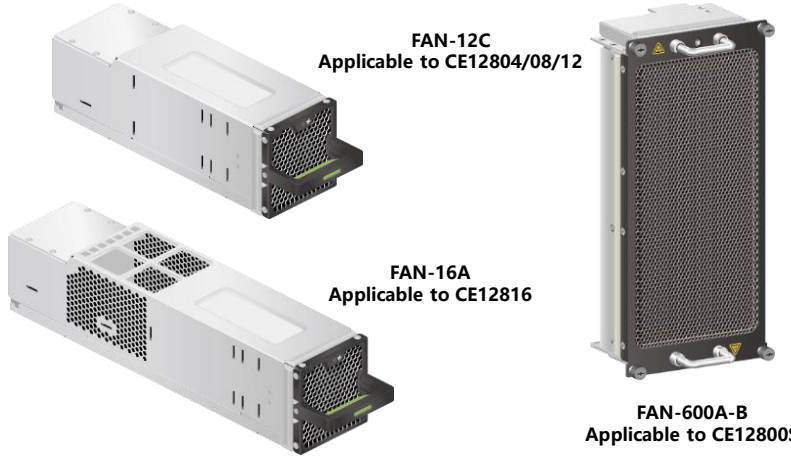
High-voltage DC: 3000 W (Supports both AC and DC power inputs)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- A 2700 W AC power module receives 110 V AC/220 V AC input power and provides 53.5 V DC/2700 W output power.
- A 2200 W DC power module receives -48 V DC/-60 V DC input power and provides 48.5 V DC/2200 W output power.
- A 3000 W high-voltage DC power module receives 220 V AC/110 V AC or 240 V DC input power and provides 53.5 V DC/3000 W output power.
- All the power modules provide input undervoltage protection, input overvoltage protection, input overcurrent protection, output overvoltage protection, output current limiting protection, output short-circuit protection, and overtemperature protection.
- All the power modules support hot swap and heat dissipation using fans.

CE12800 and CE12800S Fan Modules

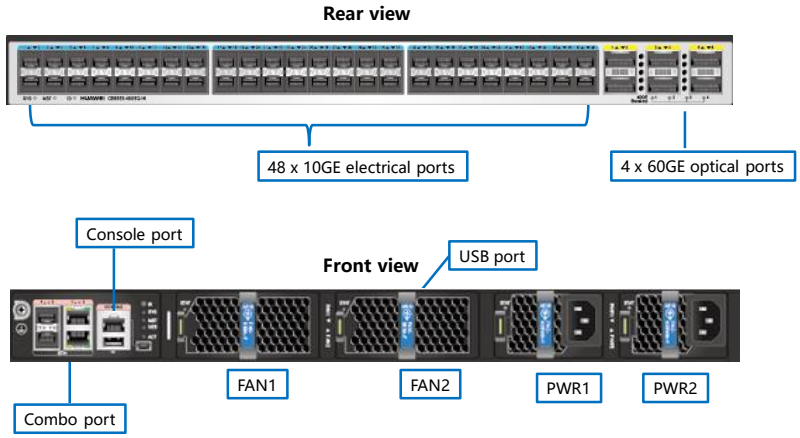


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Fan modules are installed at the rear of the CE12804S/CE12808S chassis to cool the chassis.
- Fan modules have the following functions:
  - Noise reduction: When the fan modules are powered on, they rotate at 40% of the full speed for at most 90s. After the fan modules communicate normally with the MPU, the MPU controls running of the fan modules.
  - Automatic fan speed adjustment: After the fan modules communicate normally with the MPU, the MPU controls the speed of fans according to temperature of cards in the chassis.
  - Alarm reporting: The fan modules can report alarms on loss of communication, fan failures, abnormal fan speeds, and blocking of a single fan.
  - Electronic label: The MPU reads and loads electronic labels of the fan modules through I2C buses.

Panel of the CE6800 ToR Switch (1)

CE6855-48T6Q-HI: 48 x 10GE electrical + 4 x 40GE optical

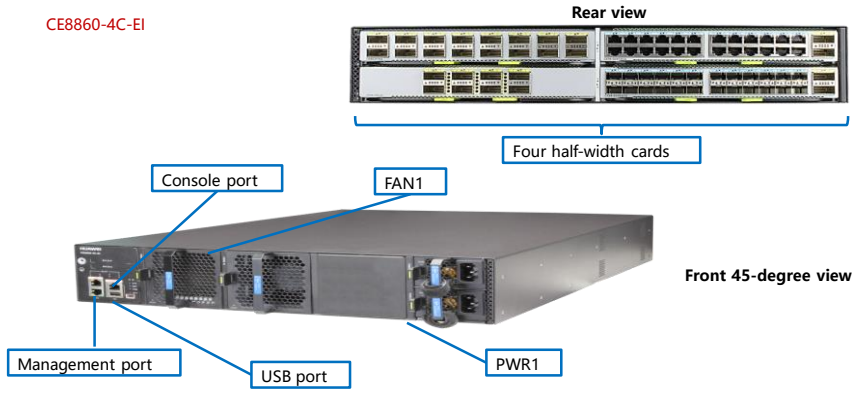


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Each CE6800 switch provides a total switching capacity of 1.28 Tbit/s, which is the industry's highest performance (in a 1 U ToR). The switch has a 960 Mpps total forwarding performance and supports L2/L3 line-speed forwarding. The CE6800 provides a maximum of 64*10GE ports, which is the industry’s highest 10GE port density (in a 1 U TOR) and meets the requirement for high-density 10GE server access. The CE6800 has 4*40GE QSFP+ ports. Each of the QSFP+ ports can be used as 4*10GE ports, allowing flexible network deployment. Through the 40GE QSFP+ ports, CE6800 switches can work with CE12800 switches to build a non-blocking network platform.

Panel of the CE8860 ToR Switch

CE8860-4C-EI



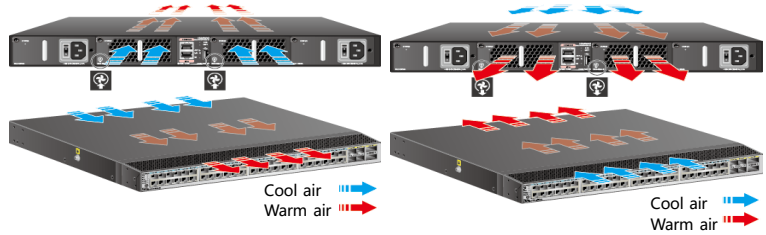
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cards of Huawei CE8860 ToR Switch

Card	Port
CE88-D24S2CQ	24 x 10GE/25GE SFP28 + 2 x 40GE/100GE QSFP28 
CE88-D24T2CQ	24 x 10GE RJ45 + 2 x 40GE/100GE QSFP28 
CE88-D16Q	16 x 40GE QSFP 
CE88-D8CQ	8 x 40GE/100GE QSFP28 

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Heat Dissipation Design for CE5800/CE6800/CE7800/CE8800 Series Switches



Front-to-back airflow **Back-to-front airflow**

Select fan modules based on the heat dissipation design characteristics of data center equipment rooms.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

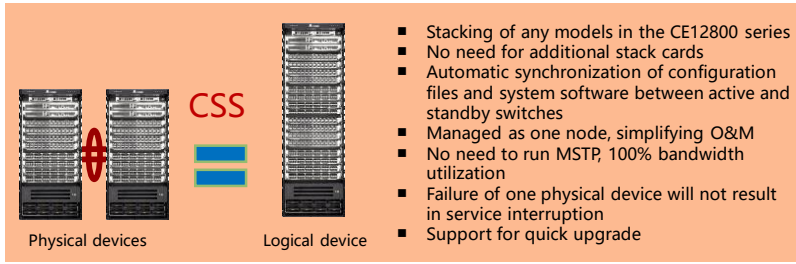
- The cooling systems of the CE5800, CE6800, and CE7800 series switches have front-to-back or back-to-front airflow depending on the airflow direction of the power modules and fan modules used.
- The airflow direction of the power modules and fan modules depends on the direction in which the CE5800, CE6800, and CE7800 series switches are installed in cabinets. Typically, cabinets in a data center have cold air flowing in from the front and hot air exhausted from the back. If switches are installed with the power supply side facing the front, you are advised to use fan modules and power modules with front-to-back airflow in the switches.

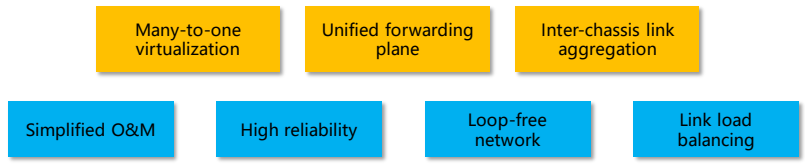
Contents

1. Product Positioning
2. Product Structure
3. Cards and Modules
4. **Product Features and Application Scenarios**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CSS: Cluster Switch System

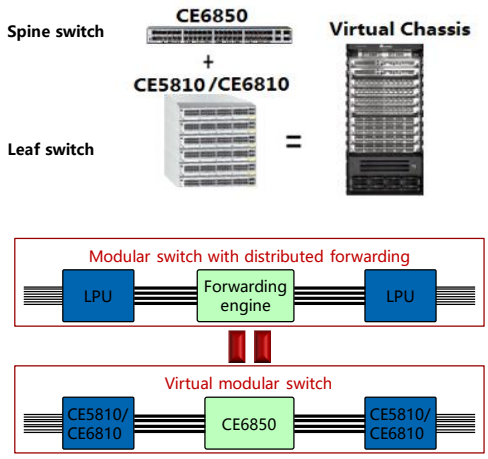




 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Cluster Switch System (CSS) virtualizes multiple switches into a single, high-performing logical switch.
- CSS offers the following features:
  - Many-to-one virtualization: CSS virtualizes multiple switches into one logical switch that has a unified control plane and provides unified management.
  - Unified forwarding plane: CSS uses a unified forwarding plane that shares and synchronizes forwarding information.
  - Inter-chassis link aggregation: Links between physical switches in a CSS are aggregated to one trunk link for interconnection with downstream devices.
- CSS simplifies network topology and greatly improves network performance by offering the following features:
  - Simplified O&M: A CSS functions as one logical switch, simplifying O&M and reducing OPEX.
  - High reliability: When one switch in a CSS fails, another switch in the CSS takes over the control and forwarding of packets to prevent services from being influenced by single-point failures.
  - Loop-free network: CSS supports inter-chassis link aggregation to prevent loops. Therefore, the deployment of complicated protocols, such as MSTP, is unnecessary.
  - Link load balancing: CSS supports equal cost multiple path (ECMP) across switches, making full use of network links and bandwidths.

SVF: Super Virtual Fabric



Vertical Stacking

- Vertical stacking at the network layer: Aggregation and access switches form a stack system.
- There are two types of devices: spine switch (also called parent switch) and leaf switch.
- Spine switch: is the core of the stack system and is responsible for control and forwarding of the entire stack system.
- Leaf switch: is used as a line card of the spine switch. The control plane is moved to the spine switch, and the leaf switch only forwards traffic locally.

Advantages

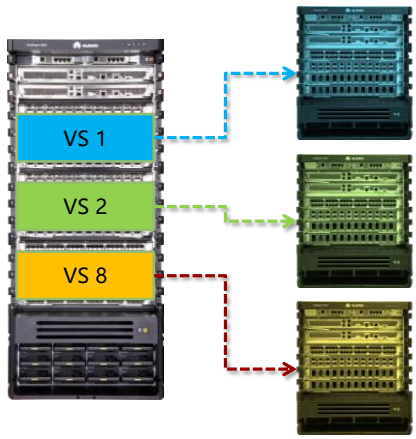
- Simplifies network deployment and maintenance.
- Simplifies cabling and reduces costs.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Super virtual fabric (SVF) is a vertical virtualization technology that virtualizes access and aggregation switches into one device to simplify network configuration and management.
- Compared with traditional access layer networking, SVF networking has the following advantages:
  - Lower network construction costs: Low-cost switches are used as access switches, so network construction costs are reduced.
  - Simpler configuration and management: SVF virtualizes multiple devices into one, reducing the number of nodes to manage. You do not need to deploy complex loop prevention protocols, simplifying network configuration and management.
  - Higher scalability and more flexible deployment: When more access ports are required on the network, you only need to add low-cost fixed switches to the network. Moreover, these low-cost switches are deployed near servers, making network deployment more flexible.

VS: Virtual Switch

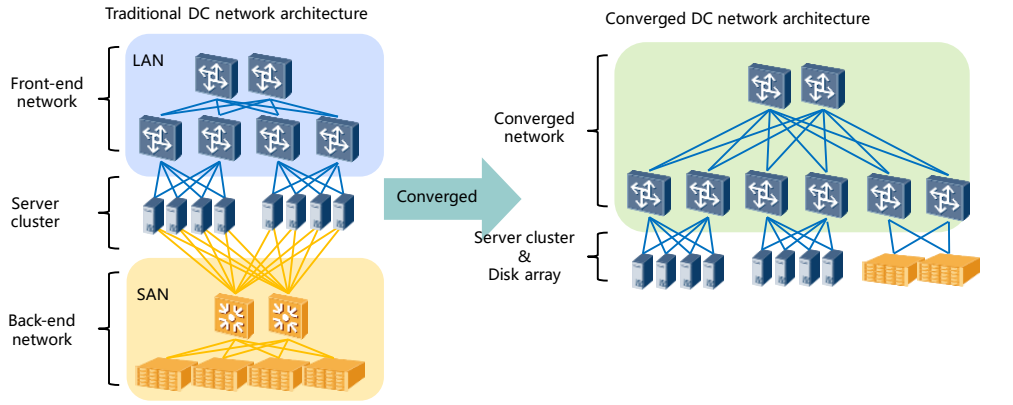


- VS definition
  - VS technology virtualizes one physical device into multiple logical devices, which are isolated from each other.
- VS characteristics
  - 1:N virtualization: One physical switch is used as multiple logical switches.
  - VS isolation: Each VS is assigned exclusive system resources and runs independent network services. VSs are isolated and do not affect one another.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A virtual machine in a data center switch removes barriers between physical devices, changing physical device resources into logical and manageable resources. These logical resources run transparently on a physical device platform, enabling isolation and on-demand distribution of resources.
- As a key feature of Huawei’ s CloudFabric Solution, Huawei VS provides the technical architecture of network device virtualization, dividing multiple logical or virtual systems on physical devices. Each VS is a virtual machine on a network device and can be independently configured, managed, and maintained. In addition, each VS is isolated from other VSs, running and processing network services independently. Data center networks process various services and serve various user groups using the VS on physical devices, which enables the following:
  - Improvements in service isolation, network reliability and security
  - Increase in device use efficiency
  - Reduction in user investments
  - Isolation between and management of user groups
  - Simplification of network O&M

FCoE: Fibre Channel over Ethernet

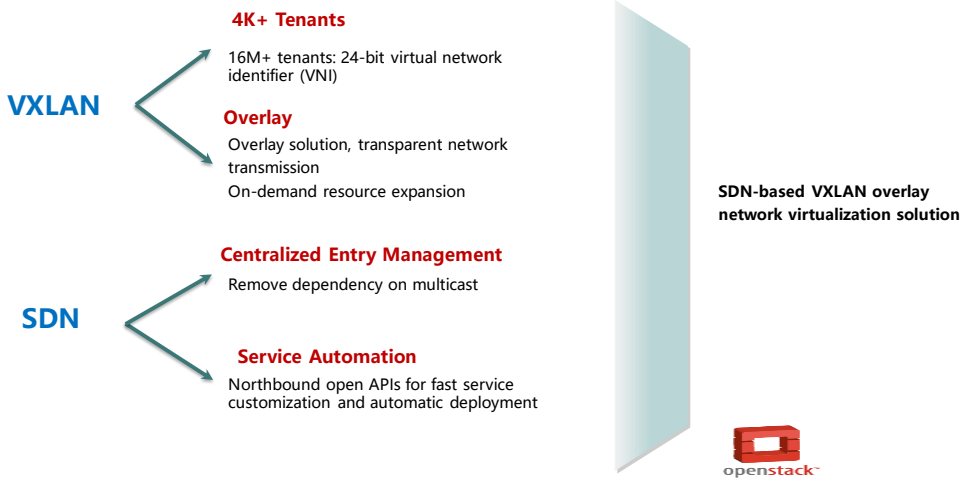


- Problems of traditional DC network architecture
  - Complex network, independent LAN and SAN, difficult to expand
  - Low energy efficiency, at least 4 to 6 network adapters in each server
- Converged network
  - Simplified network, converged LAN and SAN, unified switching
  - Reduced TCO, converged network adapters (CNAs) in servers

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A traditional data center has three independent networks: data network, storage area network (SAN), and high-performance computing (HPC) network.
  - The front-end network is a high-speed data network, on which interfaces are connected using an Ethernet links.
  - The back-end storage network is a NAS or FC SAN.
  - Parallel computing among servers is mostly implemented using an infiniband or Ethernet network.
  - Each server needs four to six network adapters, two connected to the FC SAN, two connected to the Ethernet network, and two for distributed computing connection.
  - Different servers use different operating systems, software structures, interfaces, and data formats.
- After networks are converged, a data center requires only one network: LAN.
  - FC switches provide access for FC storage, and storage data is forwarded over the Ethernet network (LAN).
  - Servers can implement HPC over the Ethernet network (LAN).
  - Each server needs only one converged network adapter (CNA).

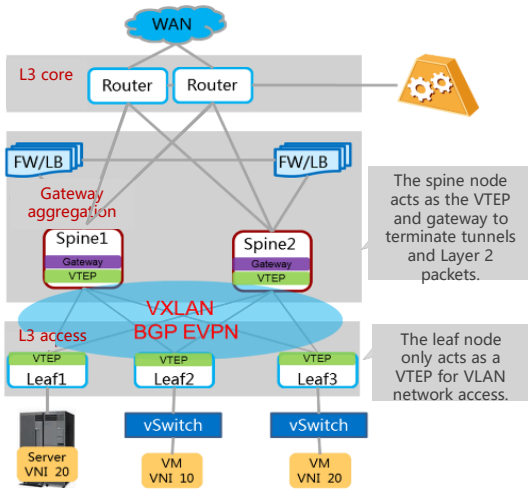
VXLAN Allows for More Than 4K Tenants



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- After servers are virtualized, a data center has multiple times more VMs than previous physical hosts, and the number of MAC addresses for virtual NICs of VMs also increases accordingly. This brings great impact on the MAC address table of a single ToR switch. In large-scale data centers and public clouds, VLAN technology can no longer meet the network isolation requirements, because only 4094 VLAN IDs can be used. Moreover, VMs need to migrate over a wide range in a cloud data center. On a VLAN network, VMs can only migrate within a Layer 2 network. To support VM migration, VLANs need to be pre-configured on the Layer 2 network. This wastes VLAN IDs, affects broadcast domain isolation, and reduces the network efficiency.
- Virtual eXtensible Local Area Network (VXLAN) is a tunnel encapsulation technology that enables large Layer 2 network expansion. This technology has been made into an IETF standard draft and used as one of mainstream network virtualization technologies.

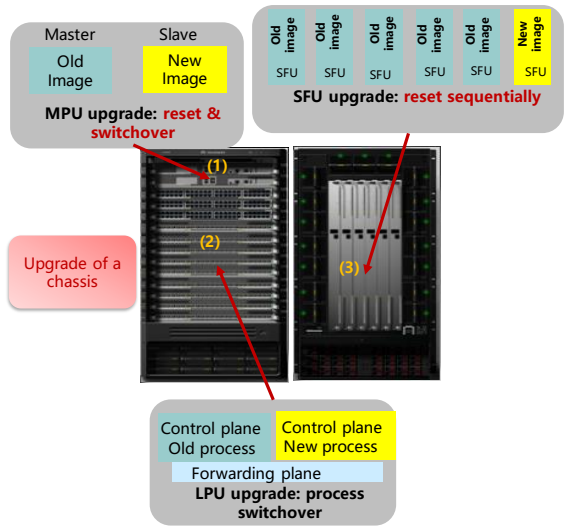
EVPN: Standard VXLAN Control Plane



- Dynamic VXLAN tunnel setup
- MAC routes learned using EVPN
- Host routes learned using EVPN
- Subnet routes learned using EVPN

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

ISSU: In Service Software Upgrade



ISSU ensures uninterrupted services during software upgrades.

Upgrade process: MPU -> LPU -> SFU

- MPU: The slave MPU loads the new image and resets. After a master/slave switchover, the original master MPU loads the new image and resets.
- LPU: A new process is created on the control plane and loads the new image. After services are switched to the new process, the forwarding plane updates forwarding entries.
- SFU: Perform N+1 backup, load new image, and reset sequentially.

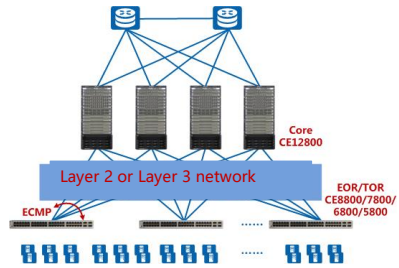
Upgrade scenarios

- Single CE12800 chassis
- CSS without specific requirements on networking
- CSS with dual-homing networking
- SVF system of fixed switches

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In most networks, network devices need to restart after software upgrades, and service running and traffic forwarding will be interrupted during a restart. To relieve the impact of system upgrades on services, you can set up multiple equal-cost links. Services can then be switched to the backup links during a software upgrade. In this case, network configurations must be modified, which increases the error probability and upgrade time. Additionally, services may still be interrupted because some links may be too busy to transmit services.
- ISSU is a mechanism that enables a device to be upgraded while services are running. This mechanism minimizes the service interruption time during system upgrades.

Core and Aggregation Layers in DCs: Fat-Tree Networking

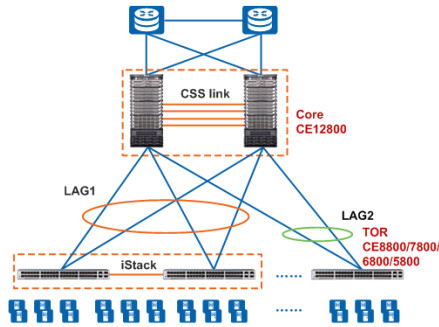


- In the fat-tree networking model, servers are dual-homed or single-homed to ToR switches through GE/10GE/25GE ports. ToR switches connect to multiple (two or four) core switching planes through 10GE/25GE/40GE uplinks. Each core switching plane has one core switch, which connects to egress routers through 10GE/25GE/40GE/100GE uplinks. In specific scenarios, core switches can also function as egress routers.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- North-south traffic between external clients and internal servers predominates in the traffic model of a traditional data center. A data center egress has a strict bandwidth limit. Therefore, traditional data centers often use a three-layer architecture consisting of access, aggregation, and core layers. In this architecture, oversubscription ratio between these layers is large.
- The wide use of 10GE servers and VM load balancers rapidly increases the volume of east-west traffic in DCs. The traditional three-layer networking architecture cannot meet the requirements of the new traffic model. Therefore, the fat-tree networking based on the CLOS architecture was developed to reduce the oversubscription ratio.
- A Layer 2 or Layer 3 network can be deployed between ToR switches and core switches according to service requirements and network design. If a Layer 3 network is deployed, the IGP (OSPF for example) protocol can run between ToR and core switches. Traffic from ToR switches can be load balanced to different core switches using IP-based equal-cost multi-path (ECMP).
- If a Layer 2 network is deployed, Multiple Spanning Tree Protocol (MSTP) can run between ToR and core switches to prevent loops. However, MSTP brings risks of link congestion, bandwidth waste, and slow convergence. For improved link efficiency and reliability of the DC network, the Transparent Interconnection of Lots of Links (TRILL) protocol can be used to build a large-scale Layer 2 loop-free network. ToR switches use uplink ports to connect to different core switches through ECMP.

Core and Aggregation Layers in DCs: CSS Networking



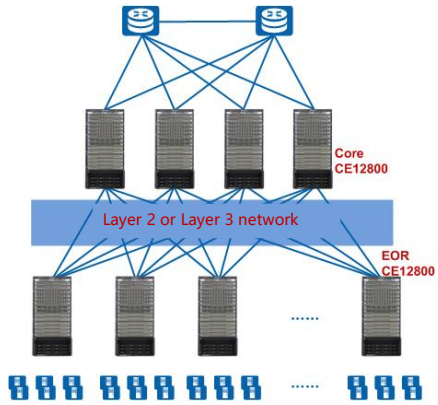
- Servers are single-homed or dual-homed to ToR switches through GE/10GE ports. You can configure ToR switches one by one or configure multiple ToR switches to form a stack using the iStack function. ToR switches (or the ToR stack) connect to the CSS through the LAG consisting of 10GE or 40GE links. The CSS has two CE12800 switches, which establish a CSS link between service ports (10GE or 40GE ports) to provide inter-chassis connections for horizontal traffic of core switches.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- CSS virtualizes multiple physical core switches into one logical core switch. This virtualization combines multiple independent links used when ToR switches are dual-homed or multi-homed to multiple devices into a link aggregation group (LAG) of a logical device. Subsequently, ToR switches and the logical core switch establish a point-to-point connection. In this situation, MSTP and other complex protocols do not need to be deployed. Compared to fat-tree, multi-plane networking and TRILL networking, CSS networking is a flexible, lightweight deployment mode, which applies to small- and medium-sized DCs.
- In CSS networking, ToR switches (or ToR stack) connect to the CSS in a similar manner to how two switches connect to each other. Therefore, a Layer 2 or Layer 3 network can be flexibly deployed.

Data Center Access Layer: EoR Application

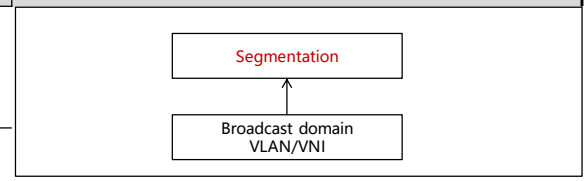
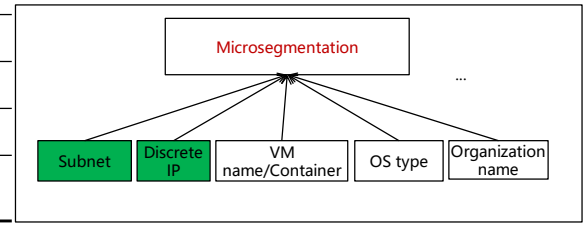


- Servers connect to EoR access switches (CE12800) through GE or 10GE uplinks, and EoR switches connect to core switches (CE12800), forming fat-tree or CSS networking.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Depending on server deployment density and data center cabling mode, an access switch can be deployed in two modes: top of rack (ToR) mode and end of row (EoR) mode. In ToR mode, an access switch is placed on the top of the server rack. Servers directly connect to the access switch of the local or neighboring cabinet. This access switch is often a fixed switch, which is called a ToR switch. For example, CE6800 series switches can function as ToR switches. In EoR mode, an access switch is placed in a cabinet at the end of a row of servers. This access switch can be a fixed or modular switch, which is called an EoR switch.
- Generally, the mainstream 1-U fixed switch with 48 GE or 10GE ports functions as an access witch. When an EoR switch requires more than 48 ports or must have the active/standby switchover capability, the CE12800 can function as the EoR switch, which directly connects to GE or 10GE servers.
- A Layer 2 or Layer 3 network can be deployed between EoR switches and core switches, depending on service requirements and network design.

Financial Tenant-Level Security: Microsegmentation

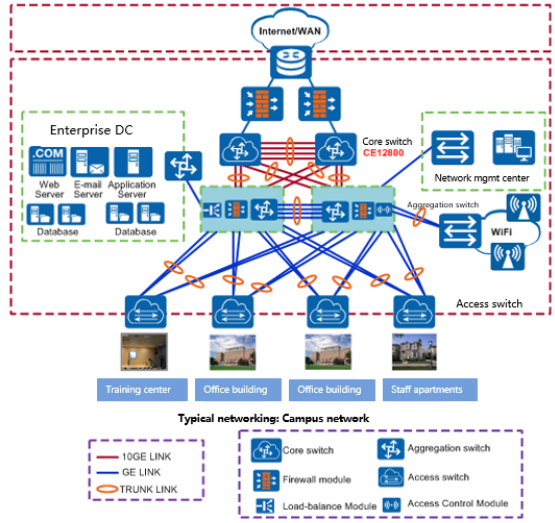
Function	Description	Remarks
Overall	As the source, microsegmentation integrates solutions including microsegmentation, GBP, security group, service chain, and third-party VAS.	
Interconnection cloud	Support for GBP model and OpenStack security group model	
Group granularity	Based on IP addresses and network segments	
Inter-group isolation	Inter-group isolation and communication	
Intra-group isolation	Intra-group isolation and communication	
Isolation mode	Stateless isolation and communication based on devices	

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- 1. Microsegmentation has been deployed in the data center of China UnionPay Shanghai.
- 2. Microsegmentation is implemented using ENP, delivering tenant-level security. ENP has strong specifications: masks support inclusion relationships, and the number of masks is not limited (only three for TD3 chips).
- 3. For other vendors, microsegmentation can be implemented only through TD3 ACL. The specifications are small (only 256), and the service capability requiring ACL resources together is poor.

Core Layer of a Campus Network: High-End Core Switch



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The CE12800 series switches are next-generation core switches with high forwarding performance, low power consumption, and high-density 10GE ports. When the campus aggregation layer needs to connect to the core layer through 10GE uplinks, the CE12800 series switches can be used as core switches.
- In typical campus networking shown in the figure, access switches connect various terminals to the network and connect to aggregation switches through GE uplinks, and aggregation switches connect to core switches through 10GE uplinks. In most cases, link aggregation group (LAG) and dual-homing mode are used to ensure link reliability.
- When functioning as core switches on a campus network, two CE12800 switches are often deployed to provide redundant connections or set up a CSS.

Quiz

1. Which components on the CE12800 are replaceable?
2. What will happen if you remove the active MPU without performing an active/standby switchover when replacing an MPU?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Which components on the CE12800 are replaceable?
  - The CE12800 supports replaceable cards, power modules, fan modules, optical modules, cables, and air filter sponge.
- What will happen if you remove the active MPU without performing an active/standby switchover when replacing an MPU?
  - When the standby MPU works properly, data synchronization between the active and standby MPUs requires a period of time. If the active MPU is removed online, the latest data on the active MPU cannot be completely backed up to the standby MPU even if the system performs an active/standby switchover. As a result, system statistics are incorrect or data is lost.
  - If the active MPU is removed online when the standby MPU does not work normally, all services on the related modules are interrupted. Partial or all services in the system are blocked.

Summary

- Product Positioning of CE Series Data Center Switches
- Structure of CE Series Data Center Switches
- Cards and Modules of CE Series Data Center Switches
- Functions and Application Scenarios of CE Series Data Center Switches

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei official websites
  - Huawei enterprise business: http://e.huawei.com/en/
  - Huawei technical support: http://support.huawei.com/enterprise/en/index.html
  - Online learning: http://learning.huawei.com/en/
- Documentation tool
  - HedEx Lite
- Simulator
  - eNSP

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei e-Learning website:
  - https://support.huawei.com/learning/en/newindex.html
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



USG6000 Series Product Introduction

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Understand the USG6000 series products
  - Be familiar with the interface cards of the USG6000 series products
  - Understand the hard disk of the USG6000 series products
  - Understand the power supply of the USG6000 series products

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **USG6000 Series Product Appearance and Panel Views**
   - USG6000 Series Portfolio and Product Appearance
   - USG6000 Series Hardware Architecture
2. Expansion Cards for USG6000 Series
3. Hard Disks for USG6000 Series
4. Power Supply for USG6000 Series

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

USG6000 Features

- New 10-Gigabit Multi-Core Hardware Platform
- Professional Content Security Defense
- Integration of Security, Routing, and VPN Services
- Refined Management by Application and User
- Visualized Management and Diversified Logs and Reports
- Carrier-Class Reliability
- Flexible Scalability

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

USG6000 Series Portfolio

- USG6300/6500/6600 series has **17 models**
- Delivers **1G to 40G application-layer** performance and **20G full-featured threat prevention** performance
- Provides a minimum of 6*GE interfaces and a maximum of **64*GE+14*10GE**


USG6570, 9Gbps FW + application identification, 4G full-featured threat prevention, 1U,8GE+4SFP


USG6390, 8Gbps FW + application identification, 4G full-featured threat prevention, 1U,8GE+4SFP


USG6380, 6Gbps FW + application identification, 3G full-featured threat prevention, 1U,8GE+4SFP


USG6550, 5Gbps FW + application identification, 3G full-featured threat prevention, 1U,8GE+4SFP


USG6370, 4Gbps FW + application identification, 2G full-featured threat prevention, 1U,8GE+4SFP


USG6530, 3Gbps FW + application identification, 1.5G full-featured threat prevention, 1U, 4GE+2Combo


USG6360, 3Gbps FW + application identification, 1.5G full-featured threat prevention, 1U, 4GE+2Combo


USG6350, 2Gbps FW + application identification, 1G full-featured threat prevention, 1U, 4GE+2Combo


USG6330, 1Gbps FW + application identification, 500M full-featured threat prevention, 1U,4GE+2Combo


USG6680, 40Gbps FW + application identification,20G full-featured threat prevention 3U, 4*10GE+16GE+8SFP


USG6670, 35Gbps FW + application identification,18G full-featured threat prevention , 3U, 4*10GE+16GE+8SFP


USG6660, 25Gbps FW + application identification, 13G full-featured threat prevention 3U, 2*10GE+8GE+8SFP


USG6650, 20Gbps FW + application identification, 10G full-featured threat prevention 3U, 2*10GE+8GE+8SFP


USG6630, 16Gbps FW + application identification, 8G full-featured threat prevention, 1U, 8GE+4SFP


USG6620, 12Gbps FW + application identification, 6G full-featured threat prevention, 1U, 8GE+4SFP

Expansion cards


WSIC-2SFP+&8GE


WSIC-8GE


SAS-300GB


WSIC-8SFP


WSIC-4GE-BYPASS

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **USG6000 Series Product Appearance and Panel Views**
   - USG6000 Series Portfolio and Product Appearance
   - USG6000 Series Hardware Architecture
2. Expansion Cards for USG6000 Series
3. Hard Disks for USG6000 Series
4. Power Supply for USG6000 Series

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (1)

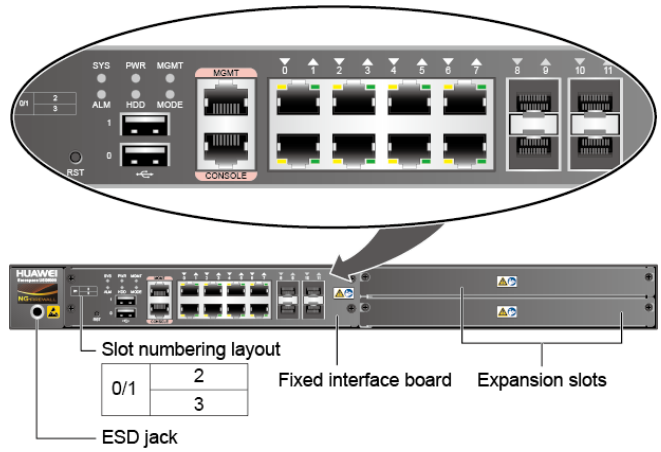


- USG6600 series products are NGFW products designed for large and medium-sized enterprises and data centers.
- USG6600 series uses industry-leading hardware and software architecture and is able to provide security and bandwidth management based on the awareness of network environment information, such as application, user, content, threat, time, and location. USG6600 series also provides IPS, antivirus, and data loss prevention (DLP) functions based on application identification to comprehensively protect the information security of enterprises.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (2)

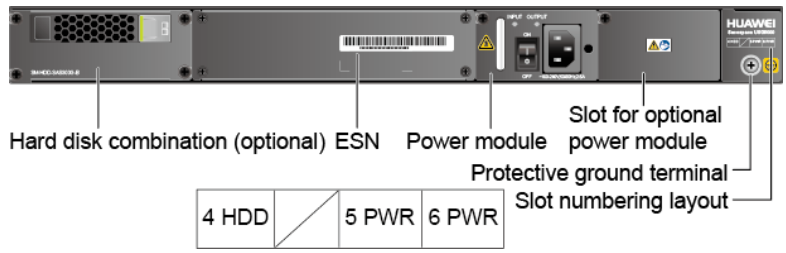
- USG6620/6630 Front Panel



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (3)

- USG6620/6630 Rear Panel

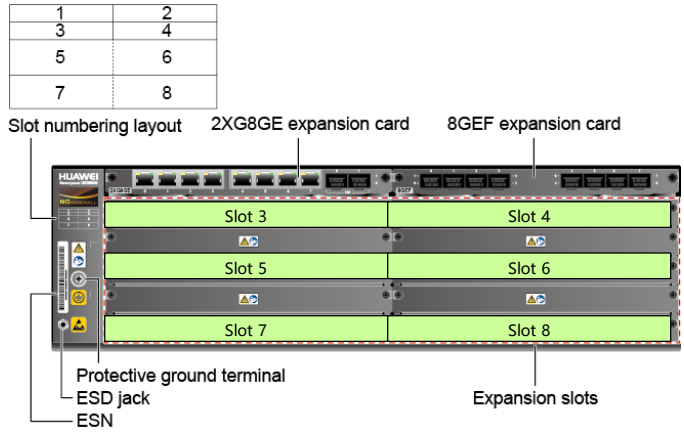


If no hard disk is installed, a filler panel must be installed on slot HDD4 to ensure normal air flow and keep out dust.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (4)

- USG6650/6660 Front Panel

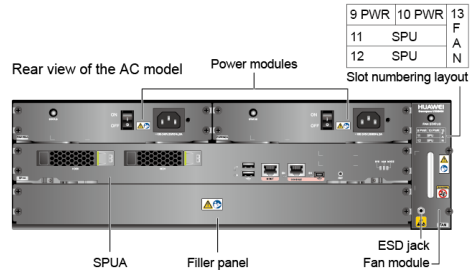


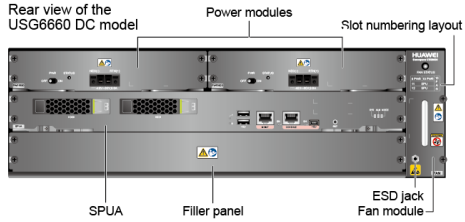
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (5)

- USG6650/6660 Rear Panel

  - The USG6650/6660 supports two 350W power modules for 1+1 redundancy. USG6650 supports only AC power modules, and USG6660 supports both AC and DC power modules for 1+1 power redundancy so that if one power module is faulty, it can be hot-swapped.



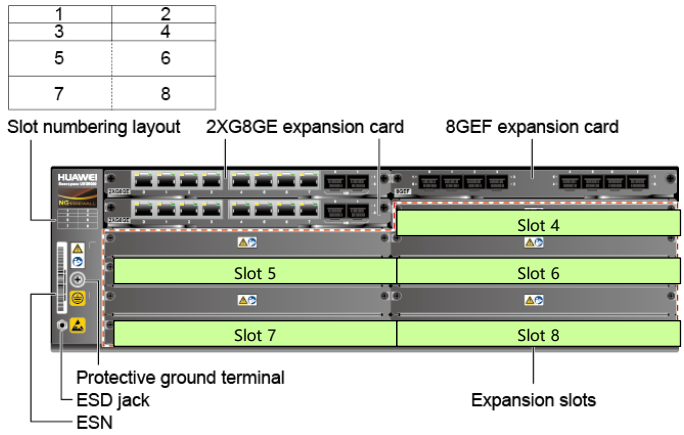


  - The USG6650/6660 provides a dedicated fan module for heat dissipation. The fan module supports hot-swapping. However, to prevent overheating, do not operate the device without a functioning fan module for more than one minute.
  - The USG6650/6660 provides Hard disk slots, support 2.5-inch SAS hard disks. You can install two hard disks to form a RAID-1 array and store log and report data.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (6)

- USG6670 Front Panel

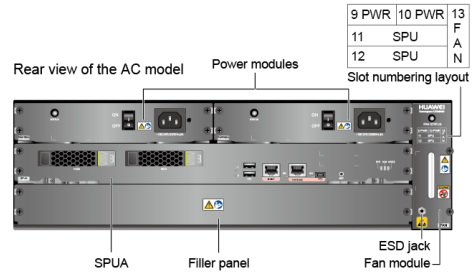


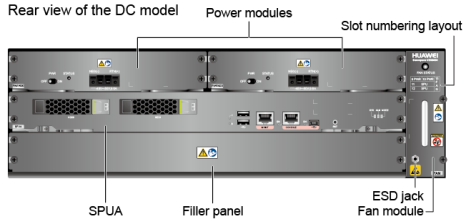
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (7)

- USG6670 Rear Panel

  - The USG6670 supports two AC or DC 350W power modules for 1+1 redundancy so that if one power module is faulty, it can be hot-swapped.



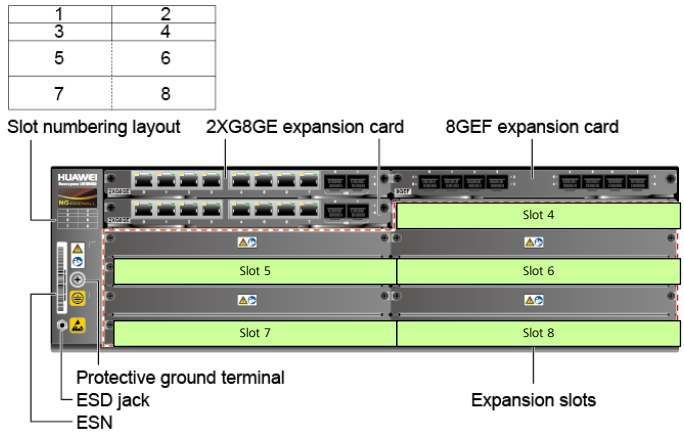


  - The USG6670 provides a dedicated fan module for heat dissipation. The fan module supports hot-swapping. However, to prevent overheating, do not operate the device without a functioning fan module for more than one minute.
  - The USG6670 provides Hard disk slots, support 2.5-inch SAS hard disks. You can install two hard disks to form a RAID-1 array and store log and report data.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (8)

- USG6680 Front Panel

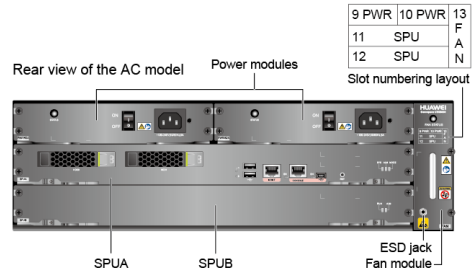


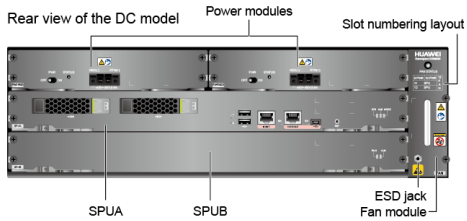
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Appearance of the USG6600 Series (9)

- USG6680 Rear Panel

  - The USG6680 supports two AC or DC 350W power modules for 1+1 redundancy so that if one power module is faulty, it can be hot-swapped.
  - The USG6680 provides a dedicated fan module for heat dissipation. The fan module supports hot-swapping. However, to prevent overheating, do not operate the device without a functioning fan module for more than one minute.





  - USG6680 provides an SPUB board to share the service processing workload with the MPU.
  - The USG6680 provides Hard disk slots, support 2.5-inch SAS hard disks. You can install two hard disks to form a RAID-1 array and store log and report data.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

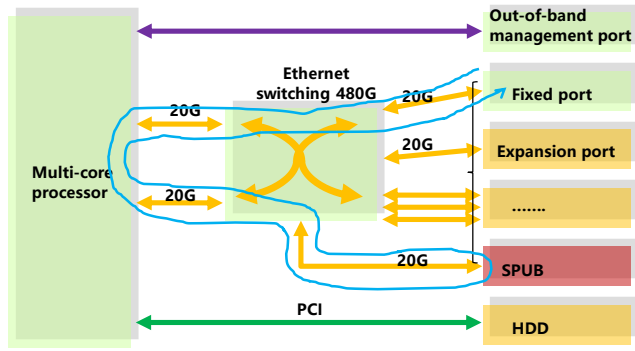
Contents

1. **USG6000 Series Product Appearance and Panel Views**
   - USG6000 Series Portfolio and Product Appearance
   - USG6000 Series Hardware Architecture
2. Expansion Cards for USG6000 Series
3. Hard Disks for USG6000 Series
4. Power Supply for USG6000 Series

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

USG6000 Series Hardware Architecture



Multi-core CPU+Switch architecture
- The CPU and switching chip both provide two 20G high-speed service ports.
- The switching capacity is high, so the upstream bandwidth in the expansion slot does not have any bottleneck.
- Provides various cards for flexible configuration.
- The USG6680 supports expansion of SPUB, providing 1+1 CPU processing capability.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The USG6000 employs the multi-core CPU+Switch architecture. The traffic from all interfaces must go through the switching chip to be processed by the CPU. After CPU processing, the traffic is forwarded from the switching chip. On the USG6680, some traffic will be forwarded to SPUB for processing.

Contents

1. USG6000 Series Product Appearance and Panel Views
2. **Expansion Cards for USG6000 Series**
3. Hard Disks for USG6000 Series
4. Power Supply for USG6000 Series

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Expansion Card Appearance



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The USG6680 supports a limited number of expansion cards. The number of supported expansion cards is determined by the power. On the USG6680 V1R1C20, the power is updated to 700W, and the number of supported expansion cards is not limited.
- Capacity expansion of the USG6680 is as follows:
- 2XG8GE : 1
- 4GE-BYPASS: 2
- 2XG: 1
- 8GE: 2
- 8GEF: 1
- The USG6600 provides both Wide Service Interface Card (WSIC) and Extended Service Interface Card (XSIC, 2 times higher than WSIC cards) slots. The lower half of the XSIC slot houses a WSIC card. The upper half does not house any card and is reserved for an XSIC card in the future.
- 8GE WSIC interface card: provides eight RJ45 GE ports.
- 2XG8GE WSIC interface card: provides eight Gigabit RJ45 and two 10G SPF+ ports.
- 8GEF WSIC interface card: provides eight Gigabit SFP ports.
- 4GE-BYPASS WSIC card: provides two electrical bypass links.
- Do not replace the expansion card during power-on, because the expansion card does not support hot swap; otherwise, the expansion card may be damaged.
- Replacing expansion cards will interrupt services. Please replace expansion cards during off-peak hours.
- Wear the ESD wrist strap while working on the USG to avoid possible damages to the USG and expansion cards.

WSIC-8GEF Interface Card Panel and Indicators



Indicator	Color	Description
Indicators 0 through 7	Green	Steady on: The link of the port is connected. Blink: Data is being sent or received through the port. Off: The link of the port is disconnected.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

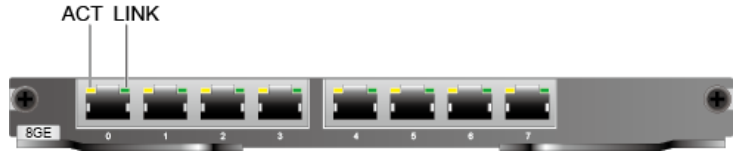
WSIC-2XG8GE Interface Card Panel and Indicators



Indicator	Color	Description
LINK indicator	Green	Off: The link is disconnected. Steady on: The link is connected.
ACT indicator	Yellow	Blink: Data is being sent or received. Off: No data is being sent or received.
Indicator 0	Green	Steady on: The link of SFP+ port 0 is connected. Blink: Data is being sent or received through SFP+ port 0. Off: The link of SFP+ port 0 is disconnected.
Indicator 1	Green	Steady on: The link of SFP+ port 1 is connected. Blink: Data is being sent or received through SFP+ port 1. Off: The link of SFP+ port 1 is disconnected.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

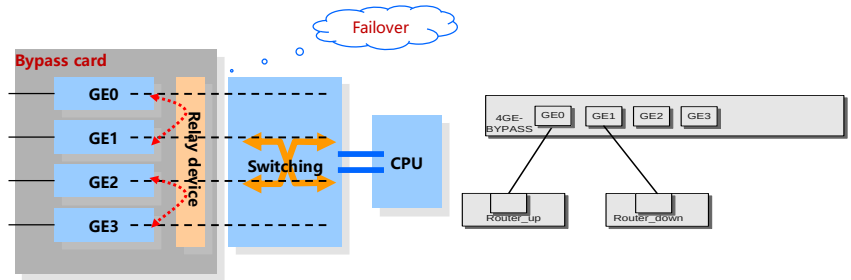
WSIC-8GE Interface Card Panel and Indicators



Indicator	Color	Description
LINK indicator	Green	Off: The link is disconnected. Steady on: The link is connected.
ACT indicator	Yellow	Blink: Data is being sent or received. Off: No data is being sent or received.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

WSIC-4GE-BYPASS Introduction



- The 4GE-BYPASS card provides two pairs of interfaces to allow traffic to bypass the device in case of a failure, such as power failure and unexpected restart.
- The interfaces of the 4GE-BYPASS card can serve as the service interfaces of the NGFW.
  - When the NGFW is working normally, traffic from Router_up goes to the NGFW through GE0. After the traffic is processed by the NGFW, the traffic is sent to Router_down out of interface GE1.
  - When the NGFW fails or is powered off, traffic from Router_up goes to GE0, and then to GE1, and then to Router_down. It is equivalent that Router_up is directly connected to Router_down.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Normally, bypass GE0/GE1 is one pair interfaces and GE2/GE3 is other pair. Bypass card detect heart information between CPU and itself, to estimate whether normal status. When CPU happened exception occurs, Bypass card change to bypass status, relay device connect GE0 and GE1, or GE2 and GE2. Upstream and Downstream pass through the device in bypass, until CPU have recovered normal status.

WSIC-4GE-BYPASS Panel and Indicators



Indicator	Color	Description
LINK indicator	Green	Off: The link is disconnected. Steady on: The link is connected.
ACT indicator	Yellow	Blink: Data is being sent or received. Off: No data is being sent or received.
Indicator 0-1	Green	Status indicator of the Layer 2 bypass link formed by GE0 and GE1 Steady on: The card is powered on. The bypass formed by GE0 and GE1 is in protection state. Blink: The card is powered on. The bypass formed by GE0 and GE1 is in working state. Off: The card is powered off. The bypass formed by GE0 and GE1 is in protection state.
Indicator 2-3	Green	Status indicator of the Layer 2 bypass link formed by GE2 and GE3 Steady on: The card is powered on. The bypass formed by GE2 and GE3 is in protection state. Blink: The card is powered on. The bypass formed by GE2 and GE3 is in working state. Off: The card is powered off. The bypass formed by GE2 and GE3 is in protection state.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

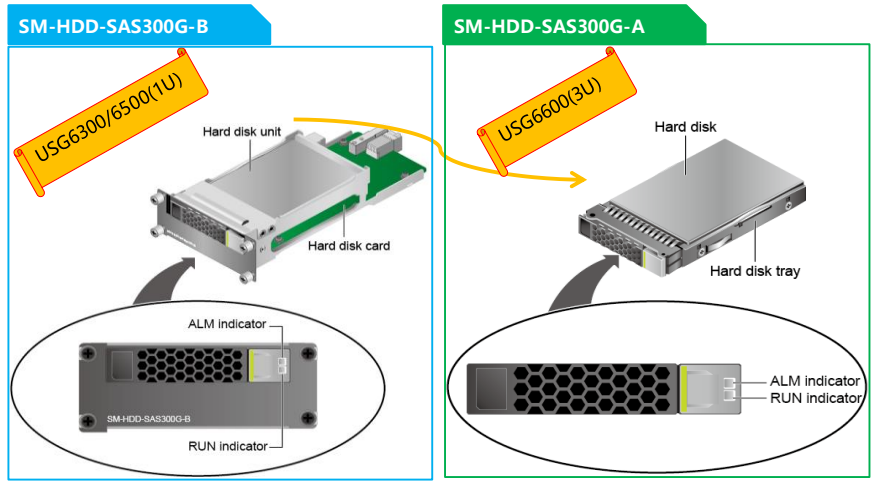
Contents

1. USG6000 Series Product Appearance and Panel Views
2. Expansion Cards for USG6000 Series
3. **Hard Disks for USG6000 Series**
4. Power Supply for USG6000 Series

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

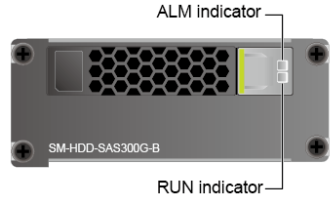
Hard Disk Appearance



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The SM-HDD-SAS300G-A hard disk has the following functions:
  - Stores log and report data.
  - Can function with another hard disk to form RAID1 and provide reliable backup for user data. Once the working hard disk is faulty, the system automatically reads data from the mirroring hard disk, ensuring user services.
  - Supports "hot replacement", that is, replacing the faulty hard disk without powering it off. After the replacement, you only need to restore data from the mirroring hard disk.
- The SM-HDD-SAS300G-B hard disk group has the following functions:
  - Stores log and report data.
  - Hard disk SM-HDD-SAS300G-A is hot swappable, but the hard disk card is not hot swappable.

SM-HDD-SAS300G-B Panel and Indicators



Indicator	Color	Description
ALM indicator	Red	Steady on: The hard disk fails. Off: The hard disk is running properly.
RUN indicator	Green	Steady on: The hard disk is running. Blink twice every second (2 Hz): Data is being read from or written to the hard disk. Off: The hard disk is not detected.

Hard disk unit SM-HDD-SAS300G-A is hot-swappable, but the hard disk card is not.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SM-HDD-SAS300G-A Panel and Indicators



Indicator	Color	Description
ALM indicator	Red	Steady on: The hard disk fails. Off: The hard disk is running properly.
RUN indicator	Green	Steady on: The hard disk is running. Blink twice every second (2 Hz): Data is being read from or written to the hard disk. Off: The hard disk is not detected.

Works with another hard disk unit to form a RAID1 for reliable user data backup. Once the working hard disk fails, the system automatically reads data from the mirror hard disk, ensuring non-stop services. Supports hot swap. If a hard disk fails, you can replace it without powering off the NGFW. After the replacement, you can restore data from the mirror hard disk.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

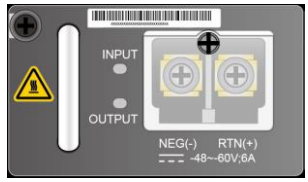
Contents

1. USG6000 Series Product Appearance and Panel Views
2. Expansion Cards for USG6000 Series
3. Hard Disks for USG6000 Series
4. **Power Supply for USG6000 Series**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

DC Power Module



Item	Description
Dimensions (H^b x W^a x D)	40mm×69mm×195mm
Weight	0.82kg
Maximum input voltage	-40V DC ~ -72V DC
Rated input voltage	-48V/-60V
Maximum output voltage	12V DC
Maximum output current	14.2A
Maximum output power	170W



Item	Description
Dimensions (H^b x W^a x D)	38.5mm×201mm×260.5mm
Weight	1.28kg
Maximum input voltage	-40V DC ~ -72V DC
Rated input voltage	-48V/-60V
Maximum output voltage	12V DC
Maximum output current	29.2A
Maximum output power	350W

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

AC Power Module





Item	Description
Dimensions (H^b x W^a x D)	38.5mm×201mm×260.5mm
Weight	1.45kg
Maximum input voltage	90V ~ 264V
Rated input voltage	100V ~ 240V
Maximum output voltage	12V DC
Maximum output current	29.2A
Maximum output power	350W

Item	Description
Dimensions (H^b x W^a x D)	40mm×69mm×195mm
Weight	0.82kg
Maximum input voltage	90V ~ 264V
Rated input voltage	100V ~ 240V
Maximum output voltage	12V DC
Maximum output current	14.2A
Maximum output power	170W



Item	Description
Dimensions (H^b x W^a x D)	38.5mm×201mm×260.5mm
Weight	1.28kg
Maximum input voltage	90V ~ 264V
Rated input voltage	100V ~ 240V
Maximum output voltage	12V DC
Maximum output current	58.4A
Maximum output power	700W

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.



Quiz

1. What are the features of the USG6000 series?
2. What is the role of the WSIC-4GE-BYPASS card?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- What are the features of the USG6000 series?
  - New 10-Gigabit Multi-Core Hardware Platform
  - Professional Content Security Defense
  - Integration of Security, Routing, and VPN Services
  - Refined Management by Application and User
  - Visualized Management and Diversified Logs and Reports
  - Carrier-Class Reliability
  - Flexible Scalability
- What is the role of the WSIC-4GE-BYPASS card?
  - Normally, bypass GE0/GE1 is one pair interfaces and GE2/GE3 is other pair. Bypass card detect heart information between CPU and itself, to estimate whether normal status.When CPU happened exception occurs, Bypass card change to bypass status, relay device connect GE0 and GE1, or GE2 and GE2. Upstream and Downstream pass through the device in bypass, until CPU have recovered normal status.

Summary

- USG6000 series products
- The interface cards of the USG6000 series products
- The hard disk of the USG6000 series products
- The power supply of the USG6000 series products

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Routine Maintenance of DC Network Devices

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Display objective of routine maintenance
  - Describe the usage scenarios of maintenance methods.
  - Describe tasks and methods of the routine maintenance of Sx7 series switches.
  - Execute periodic routine maintenance in the network made of NE and AR routers.
  - Use maintenance methods to locate USG normal faults.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Routine Maintenance Items Introduction**
2. Maintain Commonly Used Methods and Equipment Environmental Checks
3. Methods for Handling Sx7 Switch Password Loss
4. CE Switches Parts Replacement
5. AR and NE Routers Common Maintenance Commands
6. USG Common Troubleshooting

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives of Routine Maintenance

- Routine maintenance is a preventive measure.
  - It is carried out regularly during the normal running of a device. Routine maintenance is to detect and remove the defects or potential hazards in time. Therefore it is essential for secure, stable, and reliable running of the device in a long period.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Stable running of devices depends on proper network planning, routine maintenance, and monitoring on the devices.

Routine Maintenance Principles

- Equipment room environment requirements
  - The environment in an equipment room must comply with national or industrial standards and regulations. An equipment room must be clean and tidy, and prevent dust, moist, rats, and insects.
- Personnel requirements
  - Before maintaining devices, maintenance personnel must receive training on maintenance principles and procedures, learn necessary maintenance knowledge about devices, and master basic operation skills about devices and emergency handling procedures. In addition, they must strictly abide by operating regulations and security rules when maintaining devices.
- Device maintenance regulations
  - The regulations involve maintenance rules, periodic check, password management regulations, operation regulations, data modification regulations, and electrostatic discharge (ESD) standards.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Device maintenance regulations
  - Maintenance rules
    - Maintenance personnel must attach importance to routine maintenance for fault prevention, and establish regulations to ensure proper management and maintenance of the equipment.
  - Periodic check
    - Maintenance personnel should perform the routine checks or tests and record the results periodically based on the suggestions in this document.
  - Password management regulations
    - System administrators must keep administrative-level passwords well and change the passwords periodically. The passwords need to be classified into different levels and be allocated based on the operation rights of the maintenance personnel and workstations. This is to ensure secure functioning of local maintenance terminals (LMTs).
  - Operation regulations
    - Maintenance personnel must record all the major operations, such as swapping cards, restarting the system, and loading software. Before performing a major operation, maintenance personnel must verify that the operation is feasible, back up related configurations or data, and work out emergency and security measures.

Classification - Daily Routine Maintenance

- Daily routine maintenance is carried out by the ordinary maintenance personnel.
  - Alarm system
  - Operation status of the device
- Daily routine maintenance is performed to:
  - Locate alarms or other problems and solve them promptly. This helps in the stable functioning of the device and lowers the failure rate.
  - Find out the faults on the links or connections quickly and solve them to ensure normal provision of services.
  - Have real-time information about the operation of the device and the network. This helps to improve the efficiency of troubleshooting.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Classification - Periodic Routine Maintenance

- Periodic routine maintenance is relatively complex. It must be carried out by the maintenance personnel who have undergone professional trainings.
  - Power supply
  - Grounding resistance
  - Dust on the device
- Periodic maintenance is done to:
  - Ensure the secure and stable operation of the device for a long period.
  - Check for problems such as device aging, function failures, and performance degradation. You can do this by regularly checking, testing and cleaning the devices and creating a backup of the data.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Routine Maintenance Items

Maintenance Cycle	Maintenance Item
Daily	Logs and alarms
	Temperature and voltage
	Fan and power supply
	CPU usage and memory usage
Monthly	Configuration files
	Available space on the CF card
	License information
	System time
	Interface traffic
	Board running status
	OSPF/IS-IS/BGP neighbor status
	Routing information
	Management-level user control
	Telnet control
	Anti-attack detect
	FTP password control
	Change of the login password
	Backup of configuration files and log files
	Temperature and humidity in the equipment room
Quarterly	Cleaning of air filters
Yearly	Cleaning of fan frames, and boards

NOTE:
•The installation tools, meters, and related devices are not delivered with the equipment.
•Meters must be calibrated before use.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The routine maintenance items are the references and suggestions for maintenance of the running environment, software and hardware of the CE switch
- In routine maintenance, you can maintain the CE switch according to the maintenance items. In addition, you can add or delete the routine maintenance items according to the actual networking and environment of the equipment room.
- Routine maintenance and operations please refer to the CE switch product manual.

Routine Maintenance Suggestions

- Stable operation of equipment on the one hand depends on the perfect network planning, on the other hand, through routine maintenance and monitoring equipment to found run risks is also very necessary.
- Network maintenance staff can organize a inspects Checklist used for regular equipment inspection. Inspection need to focus on:
  - Equipment running environment
  - Device basic information
  - Device running status
  - Interfaces information
  - Services

No.	Check Item	Criteria and Remarks	Result
1	Position of the switch	The switch should be placed steadily in a flat position in a ventilated and dry environment. No sundries exist around the switch.	□ Satisfied □ Unsatisfied □ Not involved
2	Ambient temperature in the equipment room	The long-term ambient temperature in the equipment room should range from 0°C to 50°C; the short-term ambient temperature should **NOTE:** Short-term operation means that the continuous working time does not exceed 48 hours and the accumulated time per year does not exceed 15 days.	□ Satisfied □ Unsatisfied □ Not involved

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- inspects Checklist in the product manual "Maintenance and Fault Management" section.
  - Equipment running environment
    - normal operating environment is the precondition to ensure the normal operation of equipment.
  - Device basic information check
    - The basic device information includes software version, patch information, and system time.
  - Checking the Device Running Status
    - The running status of a switch includes the running status of subcards, resetting information, and device temperature.
  - Checking the Interface Information
    - The interface information includes the negotiation mode, configurations, and status of the interface.
  - Checking Services
    - All the services on the switch must run normally.

Contents

1. Routine Maintenance Items Introduction
2. **Maintain Commonly Used Methods and Equipment Environmental Checks**
3. Methods for Handling Sx7 Switch Password Loss
4. CE Switches Parts Replacement
5. AR and NE Routers Common Maintenance Commands
6. USG Common Troubleshooting

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Methods (1)

- Analyzing the indicator status
  - This method is used to check the running status of each card and user-side device and determine whether a card or user-side device is damaged.
- Analyzing alarm logs
  - This method is to check the current and historical alarms on a network management system (NMS) terminal to determine whether the system runs properly. If faults occur, the alarm logs help to locate the faults. After a fault is rectified, the corresponding alarm will be cleared.
- Replacing a component or device
  - When a fault cannot be located using the preceding methods, you can replace a card, line, or user-side device that is possibly faulty with a proper card, line, or user-side device to locate the fault.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Methods (2)

- Ping
  - If a fault occurs on the service network or management network, you can ping IP address of each node to locate the fault.
- Observing
  - Some faults can be determined based on observation, which is the first method that maintenance personnel use when a fault occurs. Correctly determining the fault is the key to analyze and rectify the fault.
- Removing and reinserting connectors
  - When a circuit board is faulty, you can remove and reinsert the connector to check whether the fault is caused by improper connection or a processor exception.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Methods (3)

- Isolating a faulty part
  - When the system is partially faulty, you can isolate the faulty part from related devices to check whether the fault is caused by effects of other devices.
- Self-check
  - After the system or a circuit board is powered on again, it performs self-check. When a device is performing self-check after powered on again, the indicators on the panel blink regularly. You can check the indicator status to determine whether the circuit board is faulty.
- Tightly reconnecting a chip or connector
  - You can tightly reconnect a chip or cable connector to check whether a fault is caused by improper connection.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Checking the Device Environment (1)

Recommended Maintenance Period	Check Item	Evaluation Criteria and Description	Result
Day	Air conditioners in the equipment room	Air conditioners keep running steadily so that the temperature in the equipment room is within an acceptable range.	□ Pass □ Fail □ N/A
	Connection to the power supply	The power cable is correctly and securely connected to the specified position of the device. The power supply indicator on the device should be steady green.	□ Pass □ Fail □ N/A
Week	Ambient temperature in the equipment room	The long-term ambient temperature in the equipment room should range from 0°C to 50°C. The short-term ambient temperature should range from –5°C to 55°C. Note: Short-term operation means that the continuous working time does not exceed 48 hours and the accumulated time per year does not exceed 15 days.	□ Pass □ Fail □ N/A
	Ambient humidity in the equipment room	The ambient humidity in the equipment room should range from 10% RH to 90% RH.	□ Pass □ Fail □ N/A

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Checking the Device Environment (2)

Recommended Maintenance Period	Check Item	Evaluation Criteria and Description	Result
Month	Device position	The device is placed stably in a flat position in a ventilated and dry environment. No sundries exist around the device.	□ Pass □ Fail □ N/A
	Grounding and the ground resistance	The working grounding, protection grounding, and surge protection grounding should be arranged separately in the equipment room. Joint grounding can be used if the equipment room is under limited conditions. Grounding is important especially for outdoor devices because outdoor devices are possibly damaged by lightning.	□ Pass □ Fail □ N/A
	Power supply system	The power supply system should run stably. The DC rated voltage ranges from –48 V DC to –60 V DC. The AC rated voltage ranges from 100 V AC to 240 V AC.	□ Pass □ Fail □ N/A

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Routine Maintenance Items Introduction
2. Maintain Commonly Used Methods and Equipment Environmental Checks
3. **Methods for Handling Sx7 Switch Password Loss**
4. CE Switches Parts Replacement
5. AR and NE Routers Common Maintenance Commands
6. USG Common Troubleshooting

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Restoring the Console Port Password - STelnet/Telnet (1)

- Method 1: Log in to the switch using STelnet or Telnet and change the console port password.
  - Log in to the switch using STelnet. Ensure that your user right is level 3 or higher.
  - Run the display users command to display all the users that have logged into the switch. The item marked with a plus (+) indicates your user account, which corresponds to VTY1.

```
<HUAWEI> display users
  User-Intf    Delay    Type   Network Address     AuthenStatus    AuthorcmdFlag
  129 VTY 0   00:23:36  TEL    10.135.18.67          pass          no
+ 130 VTY 1   01:20:36  TEL    10.135.18.91          pass          no
  131 VTY 2   00:00:00  TEL    10.135.18.54          pass          no
```

  - Run the display user-interface command to display user rights of all users. VTY1 corresponds to the user right level 15; therefore, you have the rights to change the console port password.

```
<HUAWEI> display user-interface
  Idx  Type     Tx/Rx      Modem Privi ActualPrivi Auth  Int
  0    CON 0    9600         -     15    -          P     -
+ 129  VTY 0                 -     15    15         P     -
+ 130  VTY 1                 -     15    15         P     -
+ 131  VTY 2                 -     15    -          P     -
  132  VTY 3                 -     15    15         P     -
......
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Restoring the Console Port Password - STelnet/Telnet (2)

- Change the console password. The following example changes the authentication mode to password authentication and the password to huawei@123.

```
<HUAWEI> system-view
[HUAWEI] user-interface console 0
[HUAWEI-ui-console0] authentication-mode password
[HUAWEI-ui-console0] set authentication password cipher huawei@123
[HUAWEI-ui-console0] return
```

- Save the configuration to prevent configuration loss after a restart.

```
<HUAWEI> save
The current configuration will be written to the device.
Are you sure to continue?[Y/N]y
Now saving the current configuration to the slot 0.
Save the configuration successfully.
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Deleting Password Through the BootROM Menu (1)

- Use a serial cable to connect a PC to the switch and restart the switch. Press Ctrl+B or Ctrl+E and enter the password (admin@huawei.com by default) to display the BootROM menu.
- Delete the password for login through the console port.

```
BootROM  MENU
   1. Boot with default mode
   2. Enter serial submenu
   3. Enter startup submenu
   4. Enter ethernet submenu
   5. Enter filesystem submenu
   6. Modify BootROM password   //Modify BootROM password is displayed if the version is V200R006
or earlier. Enter password submenu is displayed if the version is V200R007 or later.
   7. Clear password for console user
   8. Reboot
   (Press Ctrl+E to enter diag menu)
Enter your choice(1-8): 7
Note: Clear password for console user? Yes or No(Y/N): y
Clear password for console user successfully. Choose "1" to boot, then set a new password.
Note: Do not choose "8. Reboot" or power off the device, otherwise this operation will not take effect.
```

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- When the message "Press Ctrl+B to enter BootROM menu..." (if the version is V200R002 or V200R003) or "Press Ctrl+B or Ctrl+E to enter BootROM menu..." (if the version is V200R005 or later) is displayed, press Ctrl+B or Ctrl+E and enter the password to enter the BootROM menu. The default password is Admin@huawei.com; however, if the version is V100R006C03 or earlier, the default password may be huawei.
- You can use the BootROM menu of a switch to clear the lost password for console port login. Then the system can start and load all configurations normally, except that it does not prompt you to enter the console password. After the switch starts, change the console port password and save the configuration.

Deleting Password Through the BootROM Menu (2)

- Select 1 on the BootROM menu to start the switch.
- After the switch starts, you can log in to the switch through the console port without entering the password only this time. After logging in to the switch, configure a new console port login password immediately. The following example changes the authentication mode to password authentication and the password to huawei@123.

```
<HUAWEI> system-view
[HUAWEI] user-interface console 0
[HUAWEI-ui-console0] authentication-mode password
[HUAWEI-ui-console0] set authentication password cipher huawei@123
[HUAWEI-ui-console0] return
```

- Save the configuration to prevent configuration loss after a restart.

```
<HUAWEI> save
The current configuration will be written to the device.
Are you sure to continue?[Y/N]y
Now saving the current configuration to the slot 0.
Save the configuration successfully.
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Deleting the Configuration File Through the BootROM Menu (1)

- Use a serial cable to connect a PC to the switch, restart the switch, and enter the BootROM menu.
- Clear the startup configuration file. The switch will start with no configuration.

```
BootROM  MENU
   1. Boot with default mode
   2. Enter serial submenu
   3. Enter startup submenu
   4. Enter ethernet submenu
   5. Enter filesystem submenu
   6. Modify BootROM password   //Modify BootROM password is displayed if the version is
V200R006 or earlier. Enter password submenu is displayed if the version is V200R007 or later.
   7. Clear password for console user
   8. Reboot
   (Press Ctrl+E to enter diag menu)
Enter your choice(1-8): 3
      Startup Configuration  Submenu
   1. Display startup configuration
   2. Modify startup configuration
   3. Return to main menu


Enter your choice(1-3): 2
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Deleting the Configuration File Through the BootROM Menu (2)

```
Note: startup file field can not be cleared
'.'=clear field; '^D'=quit; Enter=use current configuration

startup type(1: Flash)
  current: 1
  new    :

Flash startup file (can not be cleared)
  current: HUAWEI-v200r002c00.cc
  new    :

saved-configuration  file
  current: vrpcfg.zip
  new    : .          //Clear the current value.

patch package
  current:
  new    :
      Startup Configuration  Submenu
    1. Display  startup  configuration
    2. Modify  startup  configuration
    3. Return to main  menu

Enter your choice(1-3): 3
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Deleting the Configuration File Through the BootROM Menu (3)

- On the BootROM menu, select 1 to restart the switch.
- After the switch starts, factory settings are restored. When you log in to the switch through the console port, the system asks you to set the console port login password. The following example uses the password huawei@123.

```
An initial password is required for the first login via the console.
Continue to set it? [Y/N]:y
Set a password and keep it safe. Otherwise you will not be able to login via the
console.

Please configure the login password (8-16)
Enter Password:     //Enter huawei@123.
Confirm Password:     //Enter huawei@123 again.
```

- The factory configuration of the switch is restored. To restore the original configuration without saving the console port password, download the original configuration file to the PC and delete the console port configuration. Then upload the file to the switch, specify the file for next startup, and restart the switch.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Restoring the Telnet Login Password - AAA Mode

- You can reconfigure a password for your old login account. For example, if your old login user name is huawei, you can reset the password to huawei@123 and the user privilege level to 2 as follows:

```
<HUAWEI> system-view
[HUAWEI] user-interface vty 0 4
[HUAWEI-ui-vty0-4] protocol inbound telnet
[HUAWEI-ui-vty0-4] authentication-mode aaa
[HUAWEI-ui-vty0-4] quit
[HUAWEI] aaa
[HUAWEI-aaa] local-user huawei password irreversible-cipher huawei@123
[HUAWEI-aaa] local-user huawei service-type telnet
[HUAWEI-aaa] local-user huawei privilege level 2
```

- After completing the configuration, you can enter the user name huawei and password huawei@123 to log in to the switch. If you do not remember your old user name, you can create an account and set the user name to huawei and password to huawei@123 using the same method.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Telnet can be used to manage and maintain a device remotely. If the password of a Telnet account is lost, use another method to log in to the device (for example, log in to the device through the console port) and configure a new password.
  - AAA mode: Enter a user name and password to log in.
  - Password mode: Enter only the password to log in.
- The preceding example configures the same password for VTY user interfaces 0 to 4.

Restoring the BootROM Password (1)

- In any view, restore the default BootROM password.

```
<HUAWEI> reset boot password
The password used to enter the boot menu by clicking Ctrl+B or Ctrl+E will be restored to the
default password, continue? [Y/N]y
Info: Succeeded in setting password of boot to "Admin@huawei.com".
```

- The default password has a low security level. Change the default password to a password that is easy for you to remember.
  - In the system view, run the **bootrom password change** command to change the BootROM password.

```
<HUAWEI> system-view
[HUAWEI] bootrom password change
Old Password:                  //Enter the old password of the account.
New Password(6 to 79 chars):       //Enter a new password.
Confirm Password(6 to 79 chars):   //Enter the new password again.
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Restoring the BootROM Password (2)

- Change the BootROM password in the BootROM menu.
  - Run the reboot command to restart the switch and then enter the BootROM menu.

```
BOOTROM  MENU
  1. Boot with default mode
  2. Enter serial submenu
  3. Enter startup submenu
  4. Enter ethernet submenu
  5. Enter filesystem submenu
  6. Modify  BOOTROM password
  7. Clear password for console user
  8. Reboot
Enter your choice(1-8):6 //Select 6 to change the BootROM password.
Old password:  //Enter the old BootROM password (Admin@huawei.com by default).
New password:  //Enter a new BootROM password.
Verify:       //Enter the new BootROM password again.
Save password to Flash...OK!
Save backup password to Flash...OK!
```

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- If the version is V200R007 or later, select 1. Modify BOOTROM password on the BootROM menu to change the BootROM password. The output is displayed as follows:
- BootROM MENU
  - 1. Boot with default mode
  - 2. Enter serial submenu
  - 3. Enter startup submenu
  - 4. Enter ethernet submenu
  - 5. Enter filesystem submenu
  - 6. Enter password submenu
  - 7. Clear password for console user
  - 8. Reboot
  - (Press Ctrl+E to enter diag menu)
- Enter your choice(1-8): 6 //Select 6 to enter the password submenu.
  - PASSWORD SUBMENU
  - 1. Modify BootROM password
  - 2. Reset BootROM password
  - 3. Return to main menu

Contents

1. Routine Maintenance Items Introduction
2. Maintain Commonly Used Methods and Equipment Environmental Checks
3. Methods for Handling Sx7 Switch Password Loss
4. **CE Switches Parts Replacement**
5. AR and NE Routers Common Maintenance Commands
6. USG Common Troubleshooting

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

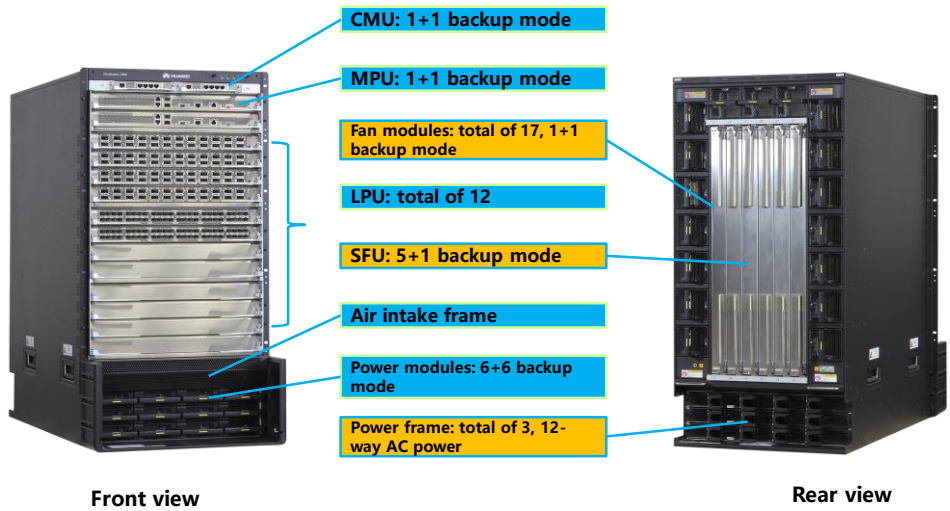
Parts Replacement Overview

- The following two scenarios require parts replacement:
- Device Maintenance
  - Routine maintenance
    - Device components need to be maintained periodically. For example, an air filter sponge needs to be cleaned periodically.
  - Troubleshooting
    - When alarms or indicators show that components (such as cards or cables) have faults that affect services, the components must be replaced immediately.
- Components Upgrade
  - A component needs to be upgraded when new functions become available.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- To ensure a successful card upgrade, follow all steps described in this document when you remove, reinstall, and restore configuration of a card.

CE12812 Core Switch



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

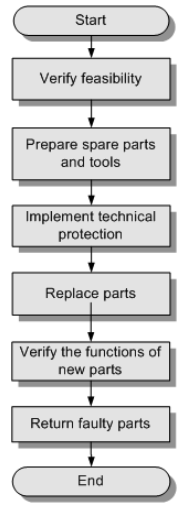
- CMU: Centralized Monitoring Unit
- MPU: Main Processing Unit
- LPU: Line Process Unit
- SFU: Switch Fabric Unit

Replaceable Parts

Parts	Impact on Services
Card	If no backup card is available, services will be interrupted.
	If cards such as SRUs work in backup mode, replacing one card may interrupt services.
Power supply	If no backup power supply is available, services will be interrupted.
	If a backup power supply is available, replacing one power supply will not interrupt services.
Fan module	If no backup fan module is available, replacing the existing fan within 2 minutes will not interrupt services.
	If a backup fan module is available, replacing one fan module will not interrupt services.
Optical module	Services will be interrupted.
Cable	Replacing an internal power cable (single-cable) in a cabinet will interrupt services on the device.
	Replacing an optical fiber or a cable will interrupt services over that fiber or cable.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

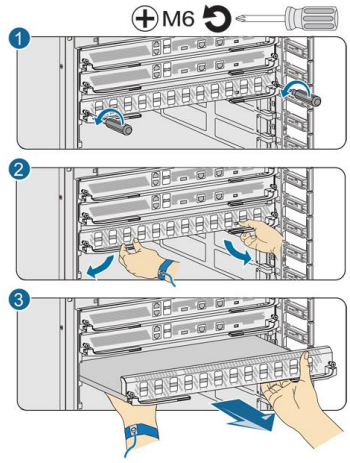
Parts Replacement Process



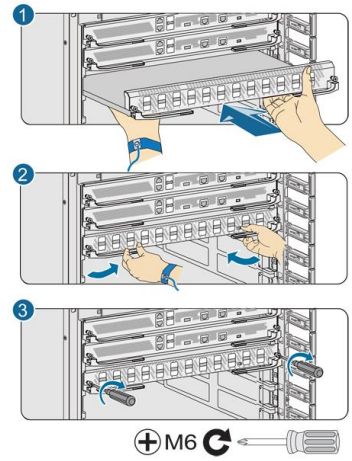
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Determine the feasibility of parts replacement.
- Before replacing a part for equipment troubleshooting or maintenance, determine the following aspects of operation feasibility:
  - Ensure that the spare part is available in the storehouse. If the spare part is not available, contact Huawei for technical support.
  - Ensure that the maintenance personnel have the following qualifications:
    - Be trained or certified in accordance with local safety regulations.
    - Know the functions of each part on the CE series switches.
    - Understand the operations of parts replacement.
    - Have skills relevant to parts replacement.
  - Ensure that the risks associated with parts replacement are controllable. Before replacing a part, conduct a thorough risk assessment. You should assess whether the risks can be controlled by taking protective measures without powering off the device. Parts replacement must be performed only when the risks are controllable. If the risks cannot be controlled, contact Huawei for technical support.

Common Card Replacement Methods



Removing a card



Installing a card

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Wear an ESD wrist strap and insert the ground terminal into the ESD jack on the cabinet, or wear ESD gloves.
- Select a spare card. The new card must be of the same type as the old card. If their types are different, ensure that the cards are compatible. Make sure that the components on the new card are not damaged or missing and record the bar code on the new card.
- Record the location of the cables and check whether the labels on the cables are correct and clear. If the labels are hard to identify, attach new labels to the cables.
- Remove cables from the card to be replaced.
- Remove the card.
  - Loose the captive screws at both ends of the card with a screwdriver, as shown in (1) in left figure.
  - Raise the ejector levers to separate the card from the backplane, as shown in (2) in left figure.
  - Grasp the ejector levers and pull out the card smoothly and slowly along with the guide rail of the slot, as shown in (3) in left figure.
  - CAUTION: When removing the card, do not touch the components on other cards.

Querying Card Information

- Log in to the switch.
- Run the display device command to view the type and status of the new card. View the Type field to check whether the card type is correct. If the Register field is displayed as Registered, the card is registered successfully. If the Alarm field is displayed as Normal, the card is running properly.
- Run the display version command to view the card software version, hardware type, and information about the MPU and LPUs.

```
<HUAWEI> display device
CE12804's Device status:
--------------------------------------------------------------------------------
Slot  Sub    Type          Online    Power  Register      Alarm    Primary
--------------------------------------------------------------------------------
3     _      CE-L24XS-EA   Offline   -      Unregistered  -    NA
4     _      CE-L24XS-EA   Present   On     Registered    Normal   NA
5     _      CE-MPUA       Present   On     Registered    Normal   Master
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Replacing the MPU (Single MPU)

- The MPU faults are classified into two types:
- Complete fault: Services are interrupted and the CLI operation cannot be performed.
  - Remove the MPU and install the new one. For details about card installation, see Common Card Replacement Methods.
- Incomplete fault: Services are still available and the CLI operation can be performed.
  - Install the new MPU into the standby MPU slot.
  - Run the slave switchover command to perform an active/standby switchover.
  - After removing cables from the MPU to be replaced, remove the MPU.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

- Fault Type 1: Services Are Interrupted
  - CAUTION: Install the component slowly and horizontally to prevent it from colliding with other cards, and do not touch the components.
  - Log in to the BIOS menu of the MPU through the serial port and copy the configuration and license files from the MPU.
- The operation procedure is as follows:
- Wear an ESD wrist strap and insert the ground terminal into the ESD jack on the cabinet, or wear ESD gloves.
- Take out the new MPU from the package box, and make sure that the components on the new MPU are not damaged or missing.
- Record the cable locations on the MPU and check whether the labels on the cables are correct and clear. If the labels are hard to identify, attach new labels to the cables.
- Remove cables from the MPU to be replaced.
- Remove the MPU and install the new one. For details about card installation, see Replacing Cards. NOTE: The system software version on the new MPU must be the same as that on the MPU to be replaced.
- View the RUN/ALM indicator status of the new MPU. NOTE: The new MPU automatically starts and registers. This process lasts 5 minutes.

Replacing the MPU (Dual MPUs)

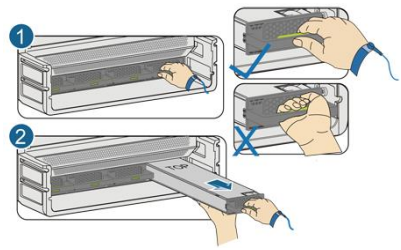
- Before replacing an active MPU, perform an active/standby switchover. Before replacing a standby MPU, you do not need to back up service data.
- Run the slave switchover command to perform an active/standby switchover if the active MPU needs to be replaced.
- Run the display switchover state command to check the switch status. Continue your operations until the Switchover State field displays as Ready.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

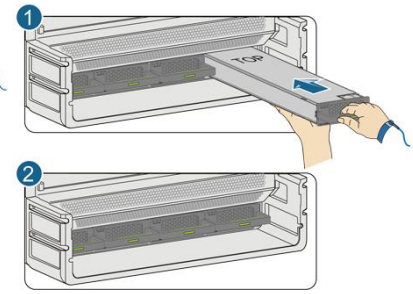
- CAUTION:
  - Before replacing an active MPU, perform an active/standby switchover. Before replacing a standby MPU, you do not need to back up service data.
  - Install the component slowly and horizontally to prevent it from colliding with other cards, and do not touch the components.
  - Data has been loaded to the flash of the new MPUs when they leave the factory. After being installed into the chassis, the new MPU automatically synchronizes the system software from the active MPU.
- Procedure
- Wear an ESD wrist strap and insert the ground terminal into the ESD jack on the cabinet, or wear ESD gloves.
- Take out the new MPU from the package box, and make sure that the components on the new MPU are not damaged or missing.
- Run the slave switchover command to perform an active/standby switchover if the active MPU needs to be replaced.
- Run the display switchover state command to check the switch status. Continue your operations until the Switchover State field displays as Ready.

Replacing Power Modules

- Before replacing a power module, switch off the corresponding circuit breaker on the power distribution frame to power off the power module.
- Exercise caution when replacing a power module to prevent scald.



Removing a power module



Installing a power module

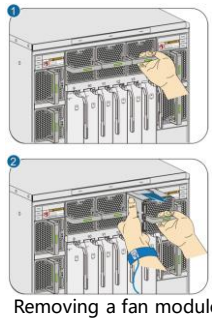
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

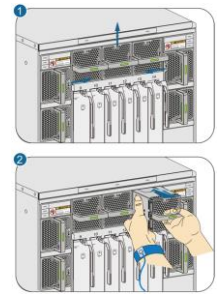
- Procedure
- Check the position of the power module to be replaced.
- Before removing a power module, find the cabinet and chassis where the power module resides. Then attach a label to the panel of the power module to identify it.
- Wear an ESD wrist strap and connect the ground terminal to the ESD jack on the chassis.
- Switch off the corresponding circuit breaker on the power distribution frame to power off the power module.
- Remove the power module from the chassis.
  - Remove the cables from the power module.
  - Unlock the power module. Use three fingers to press the release button at the interior of the handle, as shown in (1) of the left figure.
  - Remove the power module. Slowly pull out the power module with one hand and hold the power module with the other hand, as shown in (2) of the right figure.
- Insert the spare power module into the chassis.
  - Identify the top and bottom of the power module. Keep the top of the power module (marked with TOP) facing up, as shown in (1) of the right figure.

Replacing a Fan Module

- Do not remove the active and standby fan modules simultaneously; otherwise, the device temperature will increase quickly, severely affecting the safety and stability of the device.
- Exercise caution when replacing a fan module to prevent scald.



Removing a fan module



Installing a fan module

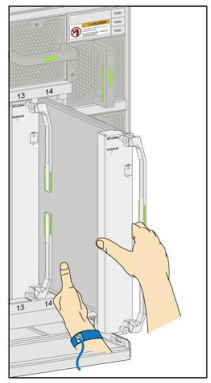
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

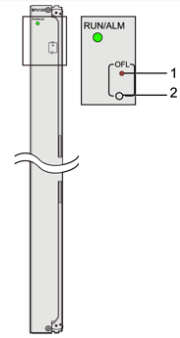
HUAWEI

- Procedure
- Check the location of the fan module to be replaced.
- Before removing a fan module, find the cabinet and chassis where the fan module is located. Then attach a label to the panel of the fan module to identify it.
- Wear an ESD wrist strap and connect the ground terminal to the ESD jack on the chassis.
- Remove the fan module from the chassis.
  - Unlock the fan module. Use three fingers to press the release button at the interior of the handle, as shown in (1) of the left figure.
  - Remove the fan module. Slowly pull out the fan module with one hand and hold the power module with the other hand, as shown in (2) of the left figure.
- Install the spare fan module into the chassis.
  - Identify the top and bottom of the fan module. The plane closest to the status indicator of the fan module is the top, and the opposite plane is the bottom. (1) of the right figure shows the installation position of a fan module in different fan slots. The top of a fan module faces different directions when the fan module installed in different fan slots of a chassis:

Replacing an SFU

The CE-SFUs of different series cannot be used in the same chassis.





1、OFL indicator 2、OFL button

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Before replacing an SFU, read the following requirements:
  - The CE-SFUs of different series cannot be used in the same chassis. For example, CE-SFU04As, CE-SFU04Bs, and CE-SFU04Cs cannot be installed on the same CE12804 chassis. A CE12804 chassis can have only one type of these CE-SFUs installed. NOTE: You can determine which series an SFU belongs to according to the name of the SFU.
  - To install or remove an SFU, hold the bottom of the SFU with one hand and hold the front panel with the other hand. Gently push or pull the SFU, as shown in the left Figure.
  - Install the component slowly and horizontally to prevent it from colliding with other cards, and do not touch the components.
  - CAUTION: To ensure that services are not affected when an SFU is replaced, hold down the OFL button before removing a running SFU. Remove the SFU when the SFU is isolated from the system (its OFL indicator turns red). You can directly replace an unregistered SFU and do not need to press the OFL button. When an SFU is running properly, its RUN/ALM indicator blinks green once every 2s (0.5 Hz) and its OFL indicator is off. the right figure shows the OFL indicator and OFL button.

Precautions for Replacing Optical Modules

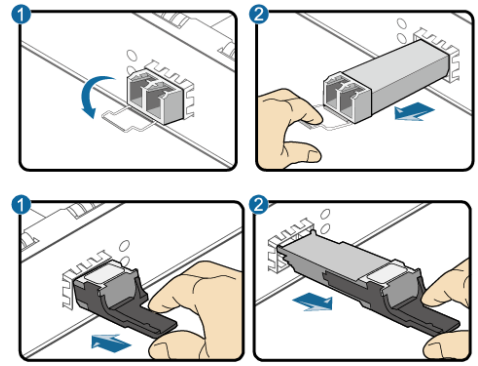
- When replacing an optical module, do not look directly at optical port without eye protection. The laser emitted from the optical port can injure your eyes.
- WARNING:
  - Be careful when you remove or insert an optical fiber to prevent damage to the fiber connector.
  - An optical module is an electrostatic sensitive device. Always take ESD protection measures when replacing an optical module to prevent the optical module from being damaged.
  - An optical module cannot be inserted inversely. If you cannot completely insert an optical module into the port, do not force it. Instead, reverse it and insert it into the port again.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Follow the rules to replace an optical module:
  - Ensure that the new optical module and the optical module to be replaced have the same center wavelength and support the same standards.
  - Install dust-proof caps on ports when removing optical fibers from optical modules.
  - Before replacing an optical module, remove all fibers from it. Exercise caution when removing the optical module to prevent damage.

Replacing Optical Modules

- Hold the handle to pull out the optical module, as shown in the figure.
- On a two-fiber bidirectional optical port, if the LINK indicator is off, swap the two fibers.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Configuration Limitations
- The copper module, high speed cable, and optical module have different configuration limitations. After you replace a module with a module of a different type, the interface configuration may be modified. Therefore, you need to confirm the interface configuration after installing a new module.
  - Note the following when an electrical module, a high speed cable, or an optical module is installed on a 10G card. When a 1000 Mbit/s copper module is installed on the 10G optical port, the port supports 1000 Mbit/s auto-negotiation and full duplex mode. The port cannot be set to work in non-auto negotiation mode.
  - When a 1000 Mbit/s optical module is installed on the 10G optical port, the port supports 1000 Mbit/s auto-negotiation and full duplex mode. The port can also be set to work in non-auto negotiation mode.
  - When a 10 Gbit/s optical module or a high speed cable is installed on the 10G optical port, the port works at a rate of 10 Gbit/s and supports non-auto negotiation and full duplex mode. The port cannot be set to work in auto-negotiation mode.

Precautions for Replacing Cables

- Take ESD protection measures before replacement.
- When replacing the ground cable, be sure to take electrical and mechanical safety precautions.
- When removing a cable, do not pull other cables. Prevent the cables from being twisted.
- When installing or maintaining an optical interface card or optical fiber, do not look at the optical interface or fiber connector without eye protection.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Power cables
- Precautions
  - Back up service data before replacing the cable.
  - Cut off the input power before replacing the cable.
  - Insulate the power cable terminals and other exposed electrical parts.
  - Ensure correct polarity when connecting power cables.
- Operation Suggestions
  - Before the replacement, attach labels to the switches that need to be operated.
  - Attach labels to the power switches that are not allowed to be operated.
  - Check the labels on power cables and ensure that the new power cables are connected in the same sequence as the replaced ones.
- Cable
- NOTE: The cables include Ethernet cables and high speed cables.
- Precautions
  - During replacement of a cable, the services transmitted over the cable are interrupted.

Precautions for Replacing an Air Filter Sponge

- To ensure good heat dissipation and prevent dust accumulation on an air filter sponge, periodically clean and replace the air filter sponge on the air filter door. It is recommended that an air filter sponge be cleaned at least once every three months and be replaced once every year. When an air filter sponge is used in a dusty environment, it needs to be cleaned and replaced more frequently. When an over temperature alarm is generated on a switch, check whether its air filter sponge is blocked. If the air filter sponge is blocked, clean or replace it immediately.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Routine Maintenance Items Introduction
2. Maintain Commonly Used Methods and Equipment Environmental Checks
3. Methods for Handling Sx7 Switch Password Loss
4. CE Switches Parts Replacement
5. **AR and NE Routers Common Maintenance Commands**
6. USG Common Troubleshooting

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Commands (1)

Command	Function
dir	Displays the information on the specified file or directory in the storage device of the S7700.
display alarm { slot-id | all }	Displays alarms.
display bgp peer	Displays the information on BGP peers.
display clock	Displays the current date and time of the S7700.
display cpu-usage	Displays the information on CPU usage.
display current-configuration	Displays the current valid parameters of the S7700.
display fan	Displays the status of the fan.
display fib	Displays the total number of the FIB entries.
display device [pic-status | slot-id]	Displays the basic information on the routers.
display ftp-server	Displays the parameters of the current FTP server.
display interface	Displays the operating status and statistics of an interface.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Commands (2)

Command	Function
display ip interface [interface-type interface-number]	Displays the detailed IP-related configurations and statistics of an interface. The packets, bytes, and multicast packets transmitted and received, and broadcast packets received, sent, and discarded are included.
display ip routing-table	Displays the brief information of the IPv4 routing table.
display isis peer	Displays the IS-IS peer relationship.
display logbuffer	Displays the record in the log buffer.
display memory-usage	Displays the CPU usage of the routers.
display ospf [process-id] brief	Displays the brief information of OSPF.
display ospf [process-id] peer	Displays OSPF neighbors.
display rip process-id neighbor	Displays RIP neighbors.
display patch-information	Displays the status of the patch.
display power	Displays the status of the power supply.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Common Maintenance Commands (3)

Command	Function
display saved-configuration	Displays the configuration files for next startup of the routers.
display startup	Displays the system software and the configuration file names related to the current and next startup.
display switchover state	Displays the backup status of the active and standby boards.
display temperature	Displays the status of the slot temperature sensor.
display trapbuffer	Displays the record in the alarm buffer.
display voltage slot slot-id	Displays the status of voltage sensor in the specified slot.
display version	Displays the version of system software.
ping	Displays IP network connection.
tracert	You can check whether the host is reachable.
	Tests the gateways the packets pass through from the host to the destination.
	You can check the network connection to locate the faults.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Routine Maintenance Items Introduction
2. Maintain Commonly Used Methods and Equipment Environmental Checks
3. Methods for Handling Sx7 Switch Password Loss
4. CE Switches Parts Replacement
5. AR and NE Routers Common Maintenance Commands
6. **USG Common Troubleshooting**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

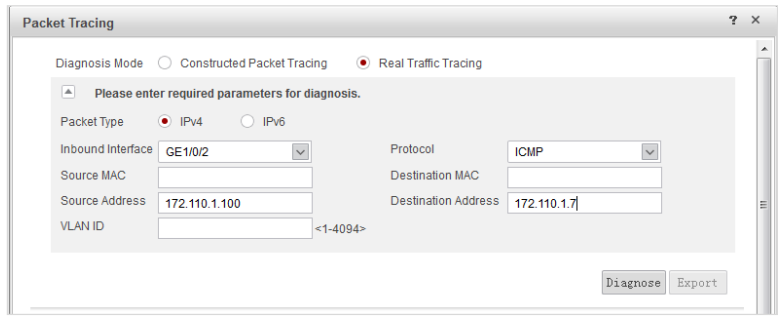
Packet Tracing Principle

- Packet tracing displays key path information during packet forwarding. Based on this information, the administrator understands how packets are processed and why packets are discarded.
- Currently, packet tracing falls into the following situations:
  - Real packet tracing
  - Constructed packet tracing

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Real Packet Tracing

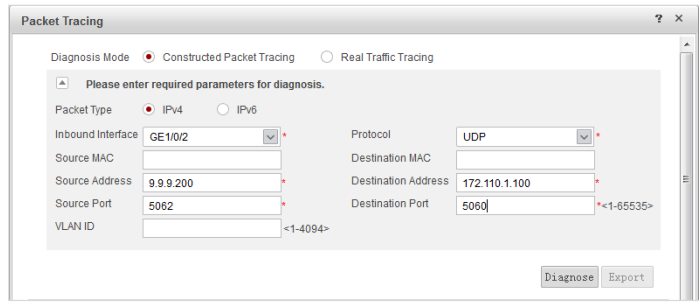
- Key paths of real service traffic are traced based on a configured quintuple.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

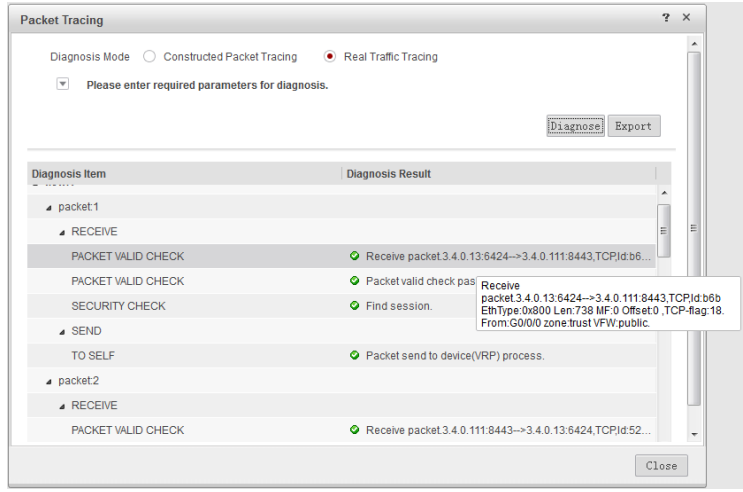
Constructed Packet Tracing

- Service packets are constructed as if they were forwarded on the NGFW. Constructed packet tracing helps check whether the current configuration applies to actual service situations. Only simple packets (Layer-2 and Layer-3 headers) are constructed.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Packet Tracing Result



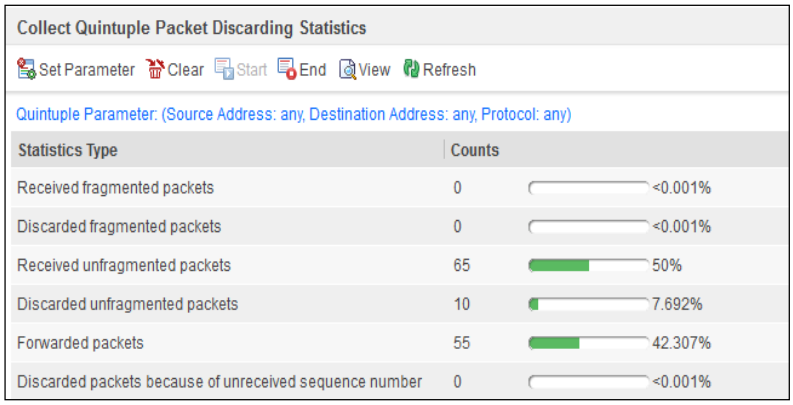
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quintuple Packet Discarding Statistics

- In addition to ACL statistics, quintuple packet discarding statistics can be displayed on the web UI, including statistics on packets that are not received. The statistics displayed on the web UI sum up statistical data of each flow and show fragment statistics. Statistics on packets that are not received can be collected based on a specified quintuple to monitor the continuity of TCP packet sequence numbers.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Packet Statistical Dimensions

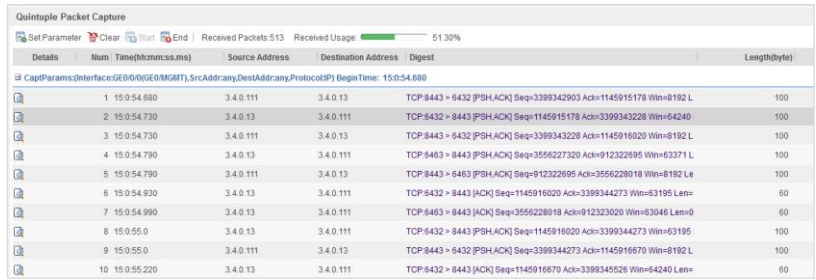


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Received fragmented packets: indicates the number of fragments entering the NGFW based on a configured quintuple.
- Discarded fragmented packets: indicates the number of fragments discarded due to attack defense, rate limiting, lack of routes, or ARP MISS based on a configured quintuple.
- Received unfragmented packets: indicates the number of packets entering the NGFW based on a configured quintuple.
- Discarded unfragmented packets: indicates the number of packets discarded due to attack defense, rate limiting, lack of routes, or ARP MISS based on a configured quintuple.
- Forwarded packets: indicates the number of forwarded packets (including fragments and non-fragmented packets) based on a configured quintuple.
- Discarded packets because of unreceived sequence number: indicates the number of packets discarded before entering the NGFW based on a configured quintuple used to monitor the continuity of TCP packet sequence numbers.

Quintuple Packet Capture

- Quintuple packet capture enables an NGFW to copy passing packets and save or display them in a certain format.

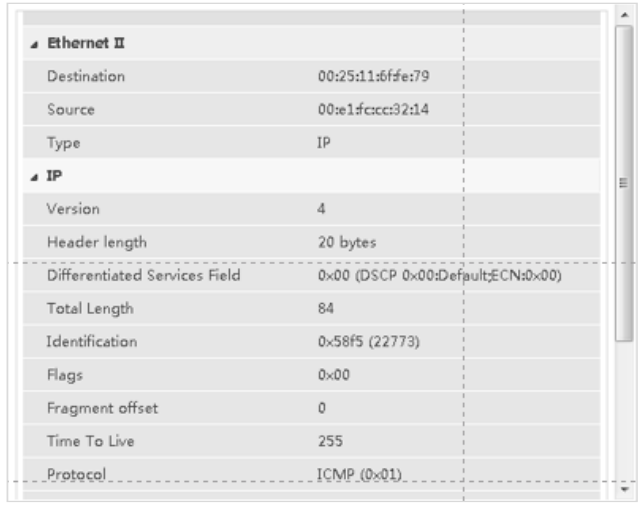


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- If the NGFW or service is faulty and the fault cannot be located after you check the configuration and statistics, you can enable quintuple packet capture for the NGFW to capture headers (payloads are not captured or displayed to prevent sensitive information leaks) of a specified flow for fault analysis.

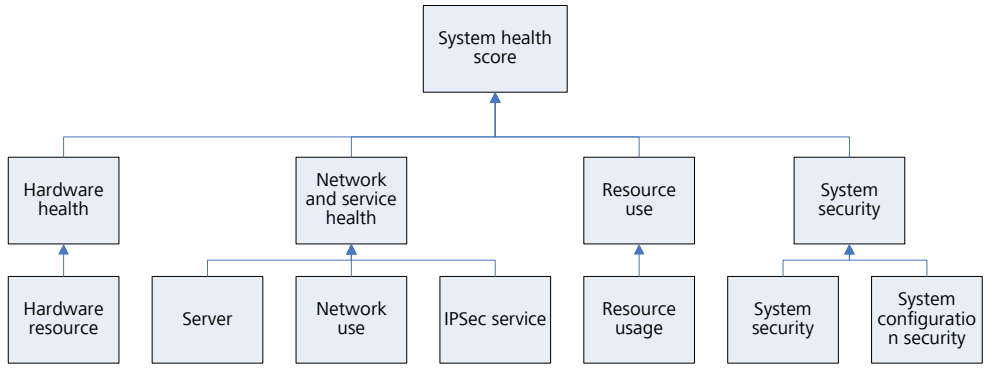
Quintuple Packet Capture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Device Health Check (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The health score represents the conditions of the network where the NGFW and its server reside. A higher score indicates a healthier condition and a better server performance. If the score is low, service efficiency is reduced or a fault occurs. The health score is calculated based on scores in four dimensions (hardware health, network and service health, resource use, and system health) and weights. Metrics in each dimension are classified based on features, as shown in the above figure.

Device Health Check (2)

- If the device health score is low, find the cause and remediation suggestions. The total score is calculated using the following formula: S = (S1 + S2 + S3 + S4)/4 x W. In this formula, S represents the total score; S(i) indicates the score of each dimension; W means the weight of the dimension with the lowest score. The weight of each dimension can be 1 for 100 (score), 0.9 for 90, 0.8 for 70, and 0.7 for 50.
- For example:

Example	Formula	Health Score
1	(100+100+100+90)/4 x 0.9	88
2	(100+100+70+70)/4 x 0.8	68
3	(90+90+90+70)/4 x 0.8	68
4	(90+90+70+70)/4 x 0.8	64
5	(90+70+70+70)/4 x 0.8	60
6	(100+100+100+50)/4 x 0.7	61

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Hardware Health

	Dimension	Metric
Hardware health	Hardware resource	**None (100):** The power and fan modules work properly. The temperature is normal. **Warning (90):** single power module **Minor (70)**: The ambient temperature exceeds the threshold. **Major or higher (50):** The fan module is faulty.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Network and Service Health

	Dimension	Metric
Network and service health	Server	**None (100):** All servers work properly. **Warning (90):** N/A **Minor (70):** One server does not work properly. **Major or higher (50):** Configured servers become invalid. Note: Servers refers to NAT , DNS, RADIUS, and AD servers.
	Network use	**None (100):** All indicators are normal. **Warning (90):** The uplink direction has error packets. There are packets whose TTL is 0. **Minor (70):** One uplink interface is Down. **Major or higher (50):** The bandwidth usage of any uplink interface is higher than 95% or all upstream interfaces are Down.
	IPSec service	**None (100):** All indicators are normal. **Warning (90):** 0 < IPSec service packet discarding rate < 50% Maximum IPSec connection rate < 50% **Minor (70):** 50% < IPSec service packet discarding rate < 80%; **Major or higher (50):** 80% < IPSec service packet discarding rate

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Resource Use

	Dimension	Metric
Resource use	Resource usage	**None (100):** Usage of all resources is within the threshold. **Warning (90):** The resource usage does not have this grade. **Minor (70):** The session usage exceeds 95% of the specifications; the memory usage exceeds 90%; the CPU usage exceeds 90%. **Major or higher (50):** The CPU usage exceeds 98%; the memory usage is 100%; the session usage reaches the specifications. Note: The memory and CPU usage refers to that on the data plane. This data keeps consistent with that on the web UI.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

System Health

	Dimension	Metric
System health	System security	**None (100):** All metrics are normal. **Warning (90):** 0 < DDoS attack traffic proportion < 50% 0 < AV/IPS alarm or blocking events **Minor (70):** 50% <= DDoS attack traffic proportion Note: AV/IPS blocking and alarm events are read from the Engine.
	System configuration security	**None (100):** Upgrades are properly performed. **Warning (90):** No AV/IPS database is loaded. The license is not activated. The license is activated but not loaded. The AV database has been not updated for more than 10 days; The IPS database has not been updated for more than 15 days; the SA database has not been updated for more than 30 days. Attack defense is not configured. The default security policy is in use. **Minor (70):** The AV database has been not updated for more than 20 days; The IPS database has not been updated for more than 45 days; the SA database has not been updated for more than 60 days.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. What is the purpose of routine maintenance?
2. What are the common methods of maintenance?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What is the purpose of routine maintenance?
  - It is carried out regularly during the normal running of a device. Routine maintenance is to detect and remove the defects or potential hazards in time. Therefore it is essential for secure, stable, and reliable running of the device in a long period.
- What are the common methods of maintenance?
  - Analyzing the indicator status
  - Analyzing alarm logs
  - Replacing a component or device
  - Ping
  - Observing
  - Removing and reinserting connectors
  - Isolating a faulty part
  - Self-check
  - Tightly reconnecting a chip or connector

Summary

- Routine Maintenance Items Introduction
- Maintain Commonly Used Methods and Equipment Environmental Checks
- Methods for Handling Sx7 Switch Password Loss
- CE Switches Parts Replacement
- AR and NE Routers Common Maintenance Commands
- USG Common Troubleshooting

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Huawei Data Center Servers Introduction and Maintenance

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- Huawei FusionServer V5 series rack servers are new-generation rack servers that can be widely used for the Internet, Internet Data Center (IDC), cloud computing, enterprise market, and telecom service applications.
- V5 rack servers combine low power consumption with high scalability and reliability, and easy deployment and management, and are 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.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be proficient in:
  - RH series servers.
  - Installation of key components for RH series servers.
  - Know about the routine maintenance and troubleshooting procedures for servers.
  - Know about the roadmap for server fault diagnosis.
  - Know about collecting logs from servers.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

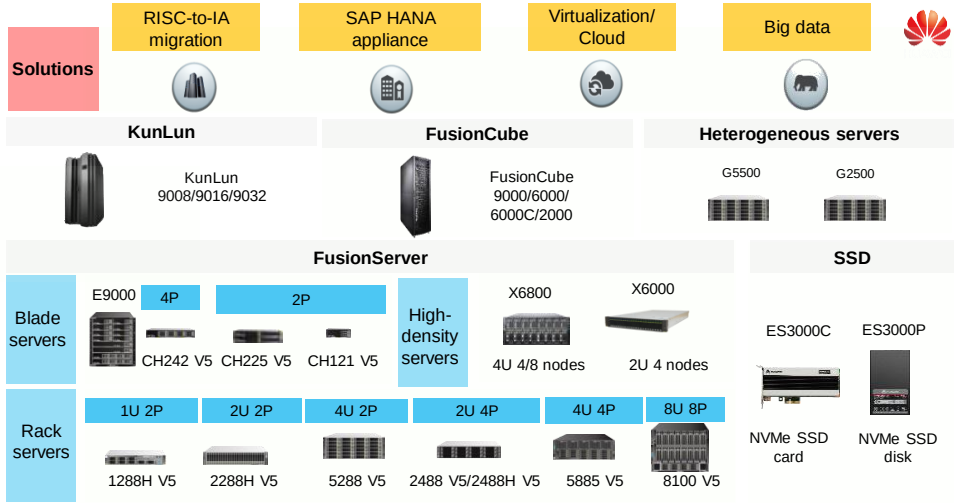
Contents

1. **RH Series Server Overview and Positioning**
2. RH Series Server Description
3. RH Series Server Routine Maintenance

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

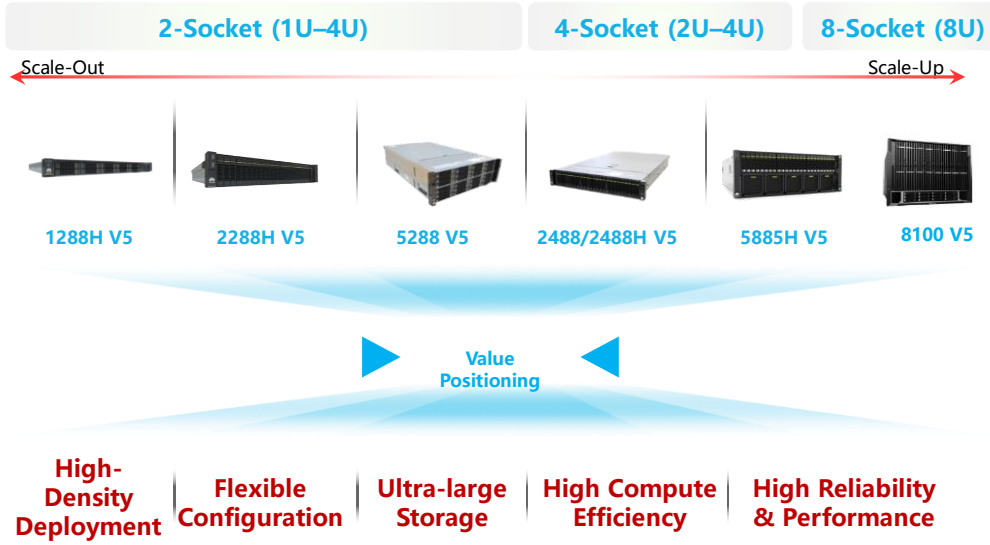
HUAWEI

Huawei V5 Server Portfolio



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

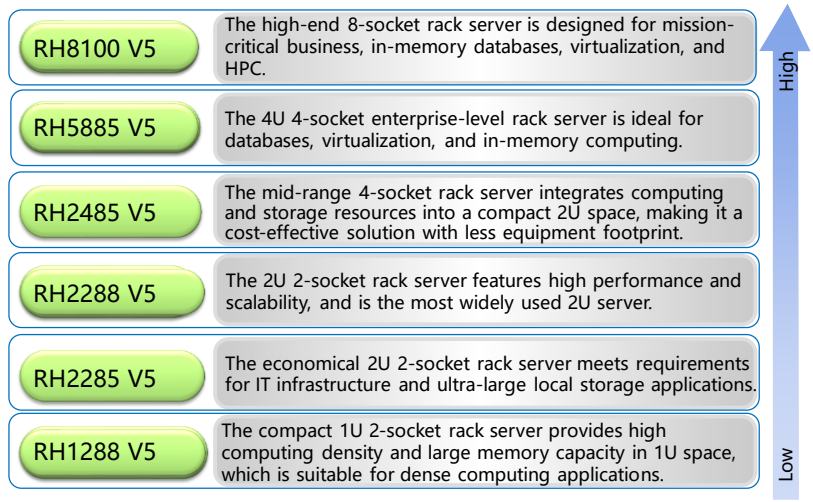
Huawei FusionServer V5 Rack Server Portfolio



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Huawei FusionServer V5 Rack Server Overview
- Form factor: 1U-8U, 2S-8S
- The servers are categorized as follows to meet different service requirements:
  - High-density deployment
  - Flexible configuration
  - Ultra-large storage
  - Computing efficiency-focused
  - High reliability and performance

Server Positioning



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

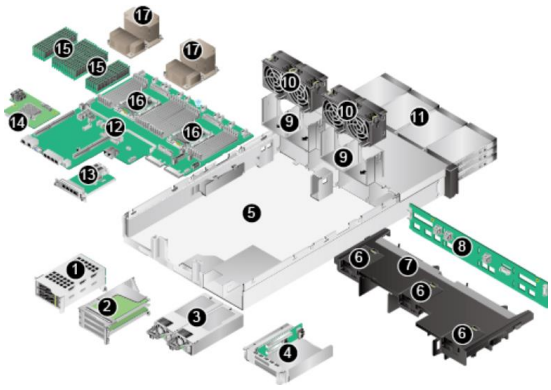
Contents

1. RH Series Server Overview and Positioning
2. **RH Series Server Description**
3. RH Series Server Routine Maintenance

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

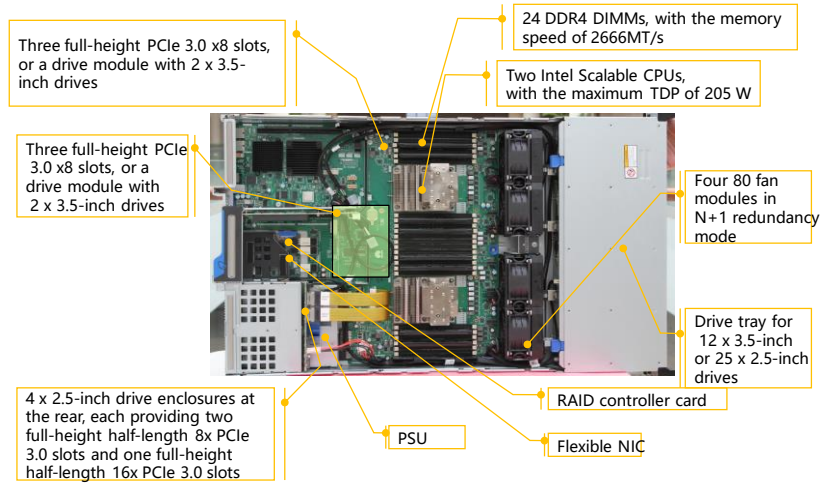
2288H V5 Physical Structure



No.	Name	No.	Name
1	IO module 1	2	IO module 2
3	Power supply unit	4	IO module 3
5	Chassis	6	Supercapacitor tray
7	Air duct	8	Front drive backplane
9	Fan module bracket	10	Fan module
11	Front drive	12	Mainboard
13	Flexible NIC	14	RAID controller card
15	DIMM	16	CPU
17	Heat sink	—	—

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

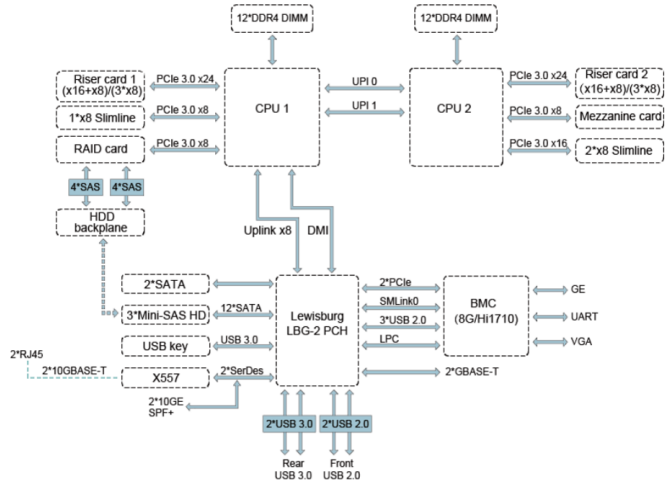
2288H V5 Internal Structure



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

2288H V5 Logical Structure

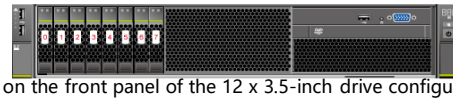


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

2288H V5 Specifications - Front View

- The 2288H V5 supports four types of configuration, 8-drive, 12-drive, 24-drive, and 25-drive.

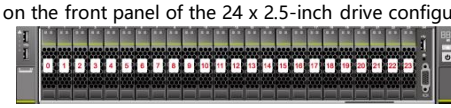
Components on the front panel of the 4 x 3.5-inch drive configuration



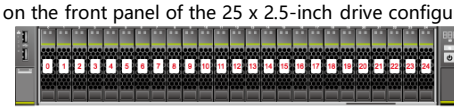
Components on the front panel of the 12 x 3.5-inch drive configuration



Components on the front panel of the 24 x 2.5-inch drive configuration



Components on the front panel of the 25 x 2.5-inch drive configuration



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

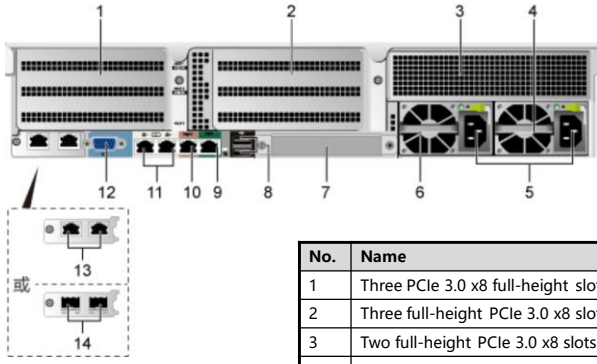
2288H V5 Specifications - Front View (8-drive Configuration)



No.	Name	No.	Name
1	USB 2.0 ports	2	Built-in DVD-RW drive (optional)
3	USB 3.0 ports	4	VGA port
5	Label (including SN)	6	Drive

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

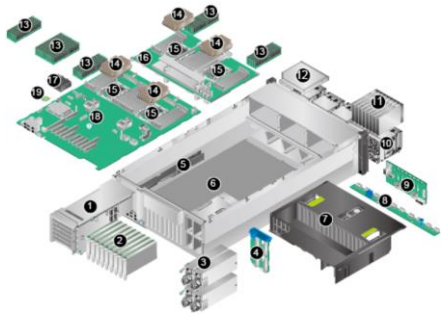
2288H V5 Specifications - Rear View (No Drives)



No.	Name	No.	Name
1	Three PCIe 3.0 x8 full-height slots	8	Two USB 3.0 ports
2	Three full-height PCIe 3.0 x8 slots	9	RJ45 serial port
3	Two full-height PCIe 3.0 x8 slots	10	RJ45 management network port
4	PSU	11	RJ45 GE service network ports
5	PSU sockets	12	VGA port
6	PSU	13	Two 10GE RJ45 electrical ports
7	PCIe 3.0 x8 flexible NIC	14	Two 10GE SFP+ optical ports

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

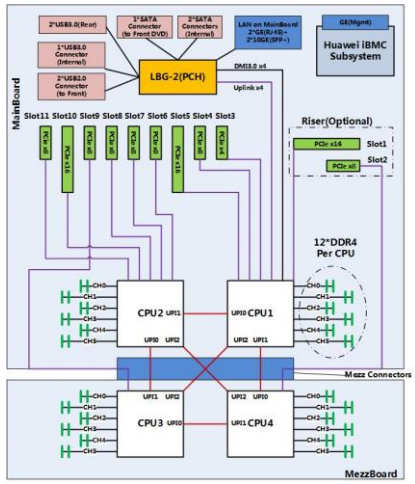
5885H V5 Physical Structure



No	Name	No.	Name
1	Riser card	2	PCIe cards
3	PSUs	4	PSU backplane
5	Cable guide	6	Chassis
7	Air duct	8	Fan adapter board
9	Drive backplane	10	Fan modules
11	Drives	12	DVD drive (or LCD)
13	DIMMs	14	Heat sinks
15	CPUs	16	Daughter board
17	Supercapacitor	18	Mainboard
19	TPM	—	—

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

5885H V5 Logical Structure



- One mainboard, supporting two CPUs, 24 DIMMs, PCIe riser cards, PCIe slots, and BMC modules
- One daughter board, supporting two CPUs and 24 DIMMs
- Drive configuration at the front:
  - 8 x 2.5-inch drives + one DVD drive
  - 25 x 2.5-inch drives
  - 16 SAS + 8 NVMe 2.5-inch drives
  - 24 x 2.5-inch drives
- Backplane and front VGA/DVD drive connecting to the RAID controller card and mainboard using cables
- Rear I/O module, providing six standard PCIe 3.0 slots, including three x16, seven x8 slots, and one x4 slots (Slots 1 and 3 are on the riser card.)
- Two AC PSUs, installed vertically and connected to the mainboard through the PSU backplane
- Only a plug-in PCIe RAID controller card is supported. The position of the RAID controller card is restricted in slot 4 for 8-drive, 24-drive, 25-drive, and 16 SAS + 8 NVMe drive configurations and in slot 5 for 8 NVMe configuration.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

5885H V5 Specifications - Front View

- The 5885H V5 supports three types of configuration, 8-drive, 24-drive, or 25-drive.

Components on the front panel of the 8 x 2.5-inch drive configuration



Components on the front panel of the 24 x 2.5-inch drive configuration

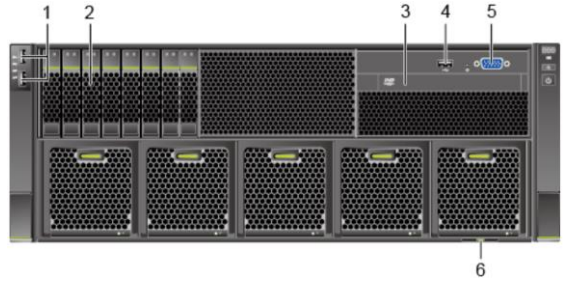


Components on the front panel of the 25 x 2.5-inch drive configuration



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

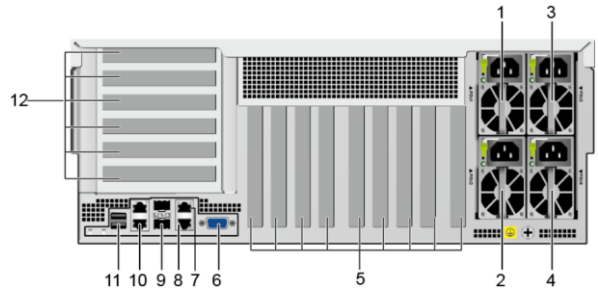
5885H V5 Specifications - Front View (8-drive Configuration)



No.	Name	No.	Name
1	USB 2.0 ports	2	SAS/SATA drives
3	Built-in DVD drive or LCD	4	USB 3.0 prots
5	VGA port	6	Label (with the ESN label)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

5885H V5 Specifications - Rear View

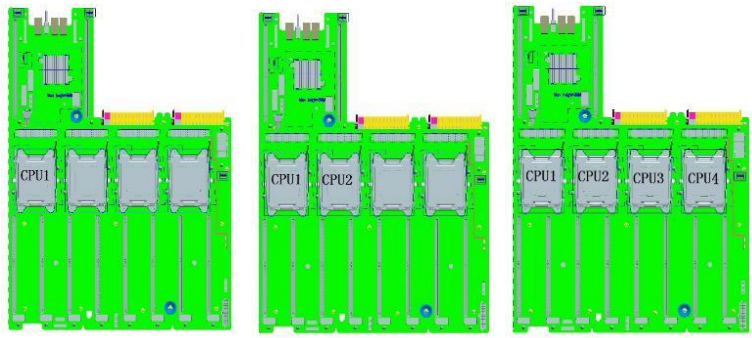


No.	Name	No.	Name	No.	Name	No.	Name
1	PSU 1	2	PSU 2	3	PSU 3	4	PSU 4
5	PCIe slots 7 to 15 (from left to right)	6	VGA port	7	Serial port	8	Management network port
9	10GE optical port	10	GE electrical port	11	USB 3.0 port	12	PCIe slots 1 to 6 (from top to bottom)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

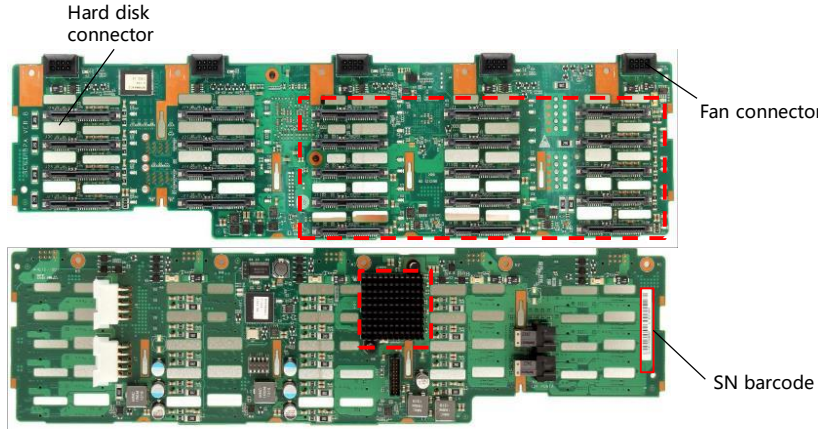
Processor Installation Rules

- The RH5885H V5 supports one, two, or four processors.
- If only one processor is to be configured, install it in socket CPU1.
- If two processors are to be configured, install them in sockets CPU1 and CPU2.
- If four processors are to be configured, install them in sockets from CPU1 to CPU4.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

RH5885(H) V5 Hard Disk Backplane



The server does not provide the components in the red rectangles.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

RH5885(H) V5 LCD

- The LCD on the front panel provides the following functions:
  - Monitors the installation status and running status of server components.
  - Queries alarms and fault information to locate faults.
  - Sets an IP address for the server.

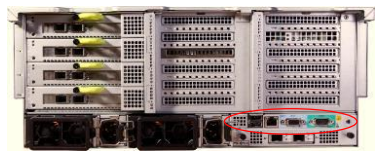




Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

BMC Card

- The BMC card is installed on the mainboard and provides the following functions:
  - Manages and monitors the server.
  - Supports the KVM function.
  - Supports the SOL redirection function.
  - Supports the online upgrades of the BIOS, FPGA, and CPLD.
  - Provides one VGA port, two USB ports, one 10/100Base-T management network port, and one serial port (system serial port by default).

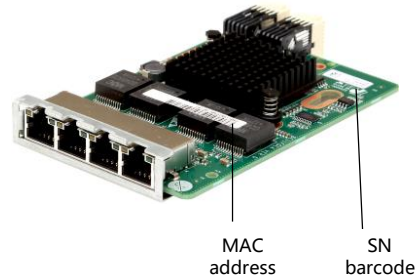




Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LOM

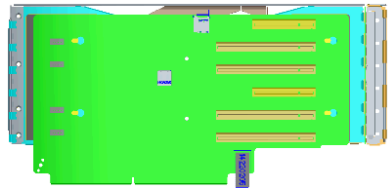




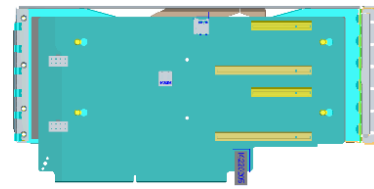
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Riser Card





PCIe x8 riser card

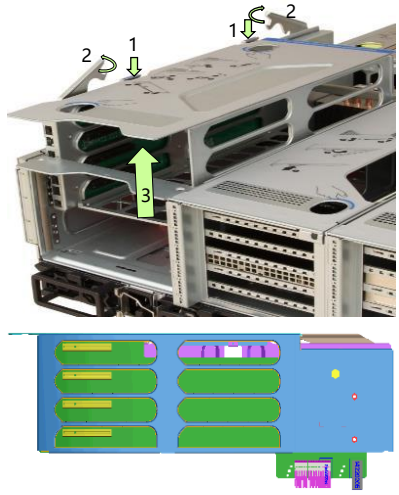


PCIe x16 riser card

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A PCIe x8 riser card provides two PCIe x4 slots (x8 connectors), four PCIe x8 slots (x16 connectors), and one USB 2.0 port.
- A PCIe x16 riser card provides two PCIe x16 PCIe slots (x16 connectors), two PCIe x4 slots (x8 connectors), and one USB 2.0 port.

Hot-Swappable Riser Card





 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A hot-swappable riser card is used for I/O expansion. It provides four hot-swappable 8x PCIe slots. The riser card supports maintenance without opening the chassis, which improves maintenance efficiency.

Contents

1. RH Series Server Overview and Positioning
2. RH Series Server Description
3. **RH Series Server Routine Maintenance**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Routine Inspection - Basic Principle

- Use unique IDs or names to identify devices.
- Keep records of the rectification of identified issues.
  - Make one change at a time and record the change result.
- Use the tools, resources, and software provided by Huawei
  - Be familiar with the updates to operating systems and application software
- Make reliable backup plan.
- Keep spare parts on site for replacements in case of anomalies.
- Save the latest network topology for later troubleshooting.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Basic principles for routine maintenance are as follows:
- Use unique ID or names to identify devices.
  - Maintain a device name list to avoid name or ID conflicts.
- Keep records of the rectification of resolved issues.
  - Make one change at a time and record the change result, including error messages and other extra information.
- Use the tools, resources, and software provided by Huawei to avoid possible faults or issues.
  - Be familiar with the updates to operating systems and application software and update operating system and application software as required.
- Make reliable backup plan.
  - Back server data at a regular basis based on server operations.
  - Implement in-time data backup if the data changes frequently.
  - Save backed up libraries based on information storage requirements.
  - Implement regular checks to ensure the correctness of saved data.
- Keep spare parts on site for replacements in case of anomalies.
  - Replenish spare parts in time.
  - Check for potential conflicts before adding new devices.
- Save the latest network topology for later troubleshooting.

On-site Inspection (1)

- On-site inspection covers equipment room environment and device running status.
- The following table lists the safety indicators that are commonly used in the inspections on equipment room environment:

Indicator	Description
	Indicates that the marked device is a dangerous power device. To avoid electric shocks, Do not open the cover of the device. Warning: All devices with this label may cause electric shocks, and no maintainable components are inside.
	Indicates that the marked device is a dangerous power device. Such a device may cause electric shocks. Do not uncover the device. No maintainable components are inside. Warning: To prevent an electric shock, do not remove the cover of the device!
	Indicates that the temperature of the surface of the marked device is high. Warning: To avoid burns, do not touch the surface until it cools down!
	Indicates that the marked device is a dangerous power device. Any operations on the marked device may cause damages to the device or operator.
	Indicates the external ground point of a device. The two ends of each power cable are connected to different devices. To ensure device operation and safety of operators, connected devices must be safely grounded using the ground point.
	Indicates the internal ground point of a device. The two ends of each power cable are connected to different devices. To ensure device operation and safety of operators, connected devices must be safely grounded using the ground point.
	It indicates an electrostatic sensitive area. Do not touch the device with bare hands. When you operate the device in this area, take antistatic measures, such as wearing an ESD wrist strap or ESD gloves.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

On-site Inspection (2)

- The equipment room environment checks involve the checks over the temperature, humidity, and power supplies within the room.

No.	Technical Indicator	Reference Value	Check Result
1	Operating temperature	10℃ to 35℃ (41℉ to 95℉)	
2	Storage temperature	-40℃ to +65℃ (-40℉ to 149℉)	
3	Temperature change rate	15℃/h (59℉/h)	
4	Operating humidity	8% RH to 90% RH (non-condensing)	
5	Storage humidity	5% RH to 95% RH (non-condensing)	
6	Altitude	≤3000m	
7	PSU	AC input voltage: 100V AC to 240V AC, 50/60Hz DC power supply: -48V DC, nominal voltage fluctuation range-38.4V to 38.4V DC	

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

On-site Inspection (3)

- To connect or remove cables, obtain the permissions from the client. The following table lists the inspections about the cable layout for servers:

No.	Item	Remarks	Result
1	General cable layout	The strong current cables and weak current cables must be respectively routed from the two sides of cabinets. That is, service cables are separated from power cables. Check whether any optical fiber is bent strongly or extended forcedly.	
2	Power cable layout	The cable layout is neat and in order, with the cabling style being the same as that in other racks in the equipment room. Do not bend power cables or wrap the power cables in a loop.	
3	Service cable layout	The cable layout is tidy and in order, with the cabling style being the same as that in other racks in the equipment room. Do not bend power cables or wrap the power cables in a loop.	
4	PGND	Servers must be properly grounded.	
5	Cable labels	The characters on the label are clear, marks are clear and correct, and labels are firmly attached to cables.	
6	Power plugs	Power cables are securely plugged into the power socket.	
7	Signaling plugs	Check the signals between server devices and switches, and check whether the data link in between is firmly connected.	

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

On-site Inspection (4)

- To inspect server running status, see to items lists in the following table:

No.	Inspection	Remarks
1	Indicators	On the front and rear panels of the servers, there are UID buttons/indicators, HEALTHY indicators, network port indicators, and power buttons/indicators. You can know the current status of the server based on the status of these indicators. For details on indicator status, refer to the server product documentation.
2	iBMC health inspection	Use the on-site management network to perform the inspection. If no management network is available on site, use a network cable to connect the portable computer to the iBMC. Then, log in to the iBMC web UI to perform the health inspection. For details on alarms, see the iBMC Alarm Reference.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Fault Information Collection (1)

- If a server becomes faulty, collect the following information for fault diagnosis:
  - Basic fault information (including the basic information about the customer, Equipment model and configuration, fault symptom).
  - Server hardware logs (using the iBMC to collect information about the server hardware) for system fault identification.
  - Service plane logs (operating system logs and service software logs) for the analysis of software-level issues.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Fault Information Collection (2)

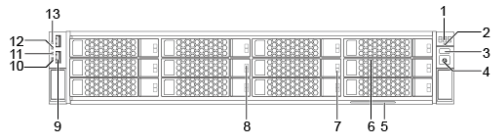
- Collect the basic fault information using the following table:

Basic Server Fault Information			
Trouble Ticket Number	For example, 123456	Ticket Submission Time	YYYY-MM-DD HH:MM:SS
Customer	Xxx	Address	Xx county/town/street, xx city, xx province
Name	Jack	Contact	Telephone number / Email address
Server Model	For example, RH2285 V5	ESN	2102310XXXXX
Hardware Configuration	Such as CPU, memory, RAID, and NIC model	OS and Service Software Version	For example, SUSE11 SP1 64 bits, Oracle 10u2
Fault Occurs on	YYYY-MM-DD HH:MM:SS		
Symptom	For example, the system automatically restarts during the installation.		
Operation Before the Fault Occurs	For example, changed the BIOS hardware watchdog.		
Operations After the Fault Occurs and Consequences	For example, powered off the server and then powered on the server, and the fault still persisted. Used another CD-ROM, and the fault still persisted. ...		

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Colleting Indicator Status

- The indicators on the server panel shows the server’ s status.



1	Fault diagnosis digital tube	2	Health indicator
3	UID button/indicator	4	Power button/indicator
5	Label (including ESN label)	6	Hard disk (numbered 0 to 11 from top to bottom and from left to right)
7	Hard disk active indicator	8	Hard disk fault indicator
9	USB 2.0 port	10	Ethernet port indicator 4
11	Ethernet port indicator 3	12	Ethernet port indicator 2
13	Ethernet port indicator 1		

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Indicators and Buttons

- You can observe the indicators to determine the current status of the server.

Silk Screen	Meaning	Color	State Description
	Fault diagnosis digital tube	-	•---: The server is operating properly. •Error Code: A fault occurs in server hardware.
	Power button/indicator	Yellow and green	•Off: The server is not powered on. Blinking yellow: The iBMC is being started. Steady yellow: The system is in the standby state. Steady green: The system is properly powered on. NOTE: You can hold down the power button for 6 seconds to power off the server.
	UID button/indicator	Blue	Identify and locate a server in a rack. Manually pressing the UID button or remotely running a command on the iBMC CLI to turn on or off the UID indicator. You can hold down the UID button for 4 to 6 seconds to reset iBMC.
	Health indicator	Red and green	Steady green: The server is operating properly. Blinking red at 1 Hz: A major alarm is generated. Blinking red at 5 Hz: A critical alarm is generated.
	NMI button	None	The NMI button triggers a server to generate a non-maskable interrupt. You can press this buttton or control it remotely through the iBMC WebUI. NOTICE: Click the NMI button only when the OS is abnormal. Do not click this button when the server is operating properly. Click the NMI button only for internal commissioning. Before clicking this button, ensure that the OS has the handler for NMI interrupt. Otherwise, the OS may crash. Exercise caution when clicking this button.
-	Hard disk active indicator	Green	Off: The hard disk is not detected or is faulty. Blinking green: Data is being read from, written to the hard disk, or synchronized between hard disks. Steady green: The hard disk is inactive.
-	Hard disk fault indicator	Yellow	Off: The hard disk is operating properly or hard disks cannot be detected in the RAID group. Blinking yellow: The hard disk is being located, or the RAID is being reconstructed. Steady yellow: The hard disk is not detected or is faulty.
1 3 2 4	Network port link status indicator	Green	Steady green: The port is properly connected. Off: The port is not in use. NOTE: If the NIC provides two network ports, they correspond to network port indicators 1 and 2 on the front panel.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Alarm Syntax

- A standard alarm consists of five fields separated by commas. The content of each field is the string after each colon. The alarm attributes are described as follows:
  - Time
    - Time when an alarm is generated, for example, Time:Wed Sep 19 09:28:11 2012.
  - Sensor
    - Name of the sensor where an alarm is generated, for example, Sensor:CPU 1 Status.
  - Event
    - Details of an alarm, for example, Description:Configuration error.
  - Assertion
    - Severity of an alarm, for example, Severity:Assertion Critical.
  - Event code
    - Event code that corresponds to an alarm, for example, Code:0x0705ffff.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Rules for Fault Locating

- Rules for Fault Locating
  - Ensure that all your operation will not cause any data losing.
  - Check the equipment running environment first then check the equipment.
  - Do the easy checking first.
  - For example, both HDD and HDD back plane can cause data can not be accessed. But checking HDD status is much easier than checking a HDD back plane.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Method for Fault Locating

- Analyse all the information you have collected
- Using fault diagnosis tools
- Referring to cases
- Some usefull method
  - Minimum system
  - Switch the component and Compare
  - Add or Remove the Component one by one
- Contact Huawei TAC

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Component Replacement Procedures and Precautions (1)

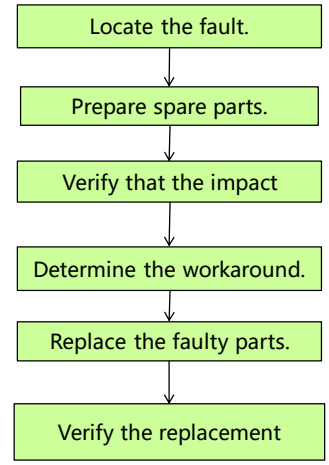
- The server parts to be replaced in route maintenance are as follows:
  - CPU/memory/Hard drive
  - Main board
  - Power supply backplane/power module
  - RAID card (storage controller)
  - RAID card battery (capacitor)
  - Fan module
  - Riser card
  - PCI-E card
  - Hard drive disk backplane
  - SAS cable
  - IO rack
  - Switching module

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Note: For detailed replacement operations, see the server maintenance guide
- Enter the following address into the address box of a browser: http://support.huawei.com/enterprise/productsupport?lang=zh&pid=9856522&idAbsPath=7919749|9856522, and access the directory of a desired server model and search for the maintenance guide.

Component Replacement Procedures and Precautions (2)

- Notes:
  - Prevent static electricity.
  - Ground the rack properly.
  - Follow the operation procedure in replacing server parts.
  - Carefully hold server parts during the replacement.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Component Replacement Procedures and Precautions (3)

- Check the warranty
  - Server SN
- Apply the Component
  - Part SN





Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Component Replacement Procedures and Precautions (4)

- When replacing server components, pay attention to the following items:
  - Service interruption occurs during the replacement of non-hot-swappable components. Ask for the permission before the replacement.
  - Wear ESD gloves for the replacement of PCB boards and cards.
  - Wear ESD wrists for the replacement of CPUs. Do not wear ESD gloves.
  - Carefully handle the CPU. Install or remove the CPU by nudging it down or lifting it up vertically. Do not move the CPU horizontally, preventing CPU pins from bending over.
  - Before the replacement of main board, verify related software versions (especially, iBMC/BIOS versions) .

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- For details on the replacement procedure, refer to the maintenance manual of the desired server model. To obtain the maintenance manual, enter http://support.huawei.com/enterprise in the address box of a browser, choose server product and access the directory of a desired server model, and search for the maintenance manual.

Quiz

1. Huawei RH2285 V5 and RH2288 V5 series servers are 2U 2-Socket servers.(Ture or False)
2. The rear panel of Huawei RH2285 V5 series servers are equipped with two 2.5-inch hard disks for system area, reducing the capacity of the occupied data area and improving data security and reliability.(Ture or False)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- T
- T

Summary

- RH series server overview and positioning
- RH series server description, including the structures, and components
- Typical applications of RH series servers
- RH V5 series server routine maintenance
- RH V5 series server fault information collection
- RH V5 series server fault diagnosis and alarm handling

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Huawei Data Center Storage System Product Introduction and Maintenance

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This chapter introduces Huawei data center storage products, Including familiarizing with the OceanStor V5 product positioning, hardware architecture, and daily operation and maintenance.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- After completing this course, you will be able to know:
  - Positioning of OceanStor V5 converged storage systems.
  - Hardware architecture of OceanStor V5 converged storage systems.
  - Features of OceanStor V5 converged storage systems.
  - The basic daily operation and maintenance operations.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Product introduction**
2. Hardware introduction
3. Daily operation and maintenance

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Storage Products

- Storage solution
  - OceanStor 18000 V5
  - OceanStor 6800 V5
  - OceanStor 5300/5500/5600/5800 V5
  - OceanStor Dorado 2100 G2/5100



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Positioning

OceanStor V5 converged storage systems

Centralized storage	High-density virtualization	Tiered storage	Data disaster recovery (DR)



OceanStor

- **Brand-new hardware architecture delivering industry-leading performance and specifications:** The V5 converged storage systems employ 16 Gbit/s Fibre Channel and 10 Gbit/s FCoE host ports and provide up to 768 GB/s system bandwidth, 16 TB cache capacity, and 27 PB storage capacity.
- **Convergence of SAN and NAS:** SAN and NAS are converged to provide elastic storage, simplify service deployment, improve storage resource utilization, and reduce total cost of ownership (TCO).
- **Outstanding scalability and reliability:** Up to eight controllers and linear growth of performance and capacity are supported. Loads are balanced among controllers that serve as hot backup for each other, achieving higher reliability. Resources are centrally stored, simplifying resource management.

Up to sixteen controllers

IP Scale-out and load balancing

Online deduplication and compression

Higher storage resource utilization

Wide channel

Latest 16 Gbit/s Fibre Channel, 12 Gbit/s SAS, and PCIe 3.0

Virtualization

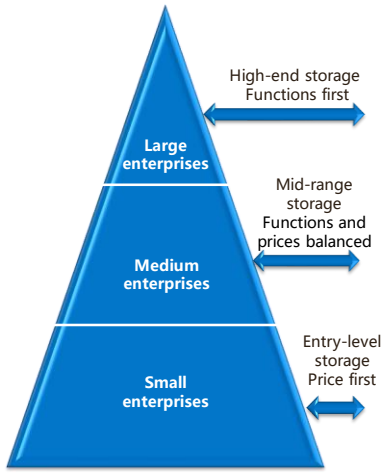
Block-level virtualization, heterogeneous virtualization, and computing virtualization

High specifications

Large capacity, high cache speed, and large number of ports

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Positioning



Type	Model	Feature	Application Scenario
High-end storage	18800 V5 6800 V5	Unified storage High performance Large capacity and flexible scalability Maximum capacity optimization High efficiency Flash memory optimization	Large-scale consolidation Layer 1 application virtualization Mixed workloads Multiple application programs High-performance application programs
Mid-range storage	5800 V5 5600 V5 5500 V5 5300 V5	Unified storage Stable performance Large capacity and flexible scalability Maximum capacity optimization High efficiency Flash memory optimization	Enterprise application programs (Oracle databases/emails/SAP) Storage consolidation Server virtualization Advanced storage tiering Data protection File sharing
Entry-level storage	2600 V5 2200 V5	Good performance and sufficient capacity Ease-of-use Cost-effective	Basic consolidation Microsoft application Programs Entry-level server virtualization iSCSI SAN Video surveillance

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Application scenario

- High-Performance Applications
  - On-Demand System Performance Boost
  - Dynamic Storage Tiering for Hotspot Data
- High-Availability Applications
  - In-Service Routine Maintenance
  - Tolerance of Single Points of Failures
  - Resilience Against Disasters
- High-Density and Multi-Service Applications
  - High-Density Virtual Machine Applications
  - Multi-Service Applications

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Product positioning
2. **Hardware introduction**
   - Hardware
     - Interface modules
3. Daily operation and maintenance

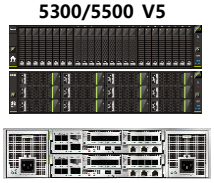
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

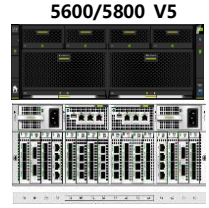
Product Form

- OceanStor V5 converged storage systems adopt brand-new hardware platform.
  - 2200/2600/5300/5500 V5: disk and controller integration (2 U).
  - 5600/5800 V5: disk and controller separation (3 U independent engine).
  - 6800 V5: disk and controller separation (6 U independent engine).
  - Active-active controllers.

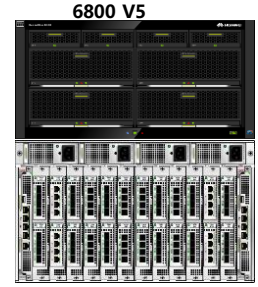
2200/2600 V5
5300/5500 V5



5600/5800 V5



6800 V5



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Features

- **High performance**
  - PCIe 3.0 high-speed bus and SAS 3.0 high-speed I/O channel
- **Flexible scalability**
  - Hot-swappable I/O interface modules
  - Support for 4 interface modules and 2 onboard interface modules (2 U)
  - Support for 16 interface modules (3 U)
  - Support for 24 interface modules (6 U)
- **Robust reliability**
  - Full redundancy design
  - Built-in BBUs+data coffer
  - Various data protection technologies
- **Energy saving**
  - Intelligent CPU frequency control
  - Delicate fan speed control

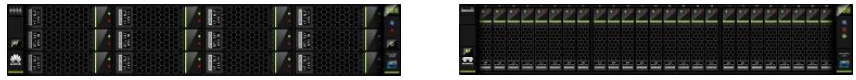
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

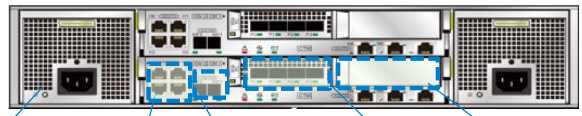
Common Storage Advanced Technology

- SmartThin: Storage space is allocated when it is written. SmartThin allows a host to allocate real space to a Thin LUN when writing data to a Thin LUN.
- SmartTier: An intelligent data storage tiering feature can help select an appropriate storage tier for each data block based on the data activity level. This feature improves the storage system performance and lowers the total cost of ownership.
- SmartQoS: SmartQoS is an intelligent QoS control feature developed by Huawei that dynamically allocates a storage system's resources to meet specific performance goals of certain applications.
- SmartPartition: It allows you to assign cache partitions of different sizes to different applications. The system provides the cache capacity specified by a partition for the corresponding applications, thereby ensuring the performance of the applications in the partition.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

OceanStor 2200 V5 Controller Enclosure





Power-fan module
- 1+1 redundancy
- AC or 240 V high-voltage DC

SAS expansion port
- Each controller provides two SAS expansion ports.

Onboard port
- Four GE ports on each OceanStor 2200 V5 controller

Filler panel

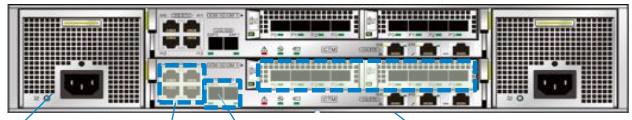
Port module
- One hot-swappable I/O module slot only on the front end
- Port type: 8 x 8 Gbit/s Fibre Channel, GE, 10GE electrical port, SmartIO card (16 Gbit/s Fibre Channel, 8 Gbit/s Fibre Channel, 10 Gbit/s FCoE (VN2VF), and 10 Gbit/s ETH (optical port))

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

OceanStor 2600 V5 Controller Enclosure







Power-fan module
- 1+1 redundancy
- AC, -48 V DC, or 240 V high-voltage DC

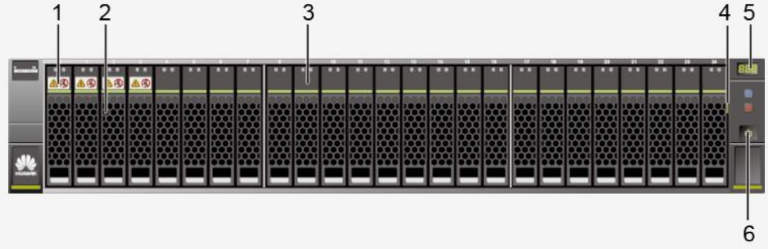
SAS expansion port
- Each controller provides two SAS expansion ports.

Onboard port
- Four GE ports on each OceanStor 2600 V5 controller

Port module
- Two hot-swappable I/O module slots only on the front end
- Port type: 8 Gbit/s Fibre Channel (4 x 8 Gbit/s Fibre Channel or 8 x 8 Gbit/s Fibre Channel), GE, 10GE electrical port, SmartIO card (16 Gbit/s Fibre Channel, 8 Gbit/s Fibre Channel, 10 Gbit/s FCoE (VN2VF), and 10 Gbit/s ETH (optical port)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

5300/5500 V5 Controller Enclosure (25 disk slots)

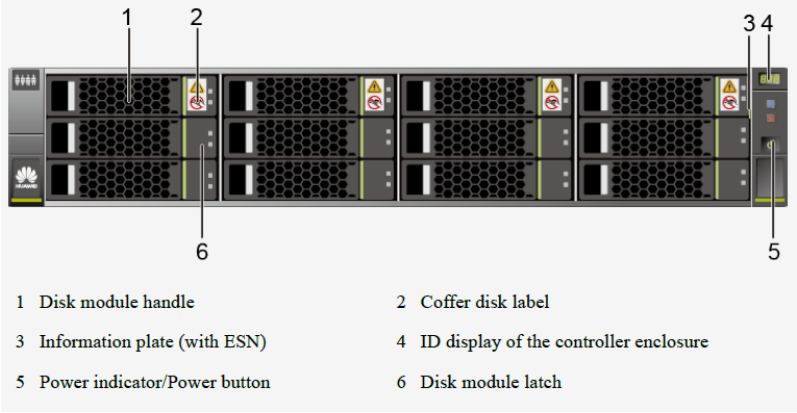


1 Coffer disk label	2 Disk module handle
3 Disk module latch	4 Information plate (with ESN)
5 ID display of the controller enclosure	6 Power indicator/Power button

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Note: The disk slots of a 2 U 25-disk-slot controller enclosure are numbered from 0 to 24 from left to right. Coffer disks are in slots 0 to 3.

5300/5500 V5 Controller Enclosure (12 disk slots)

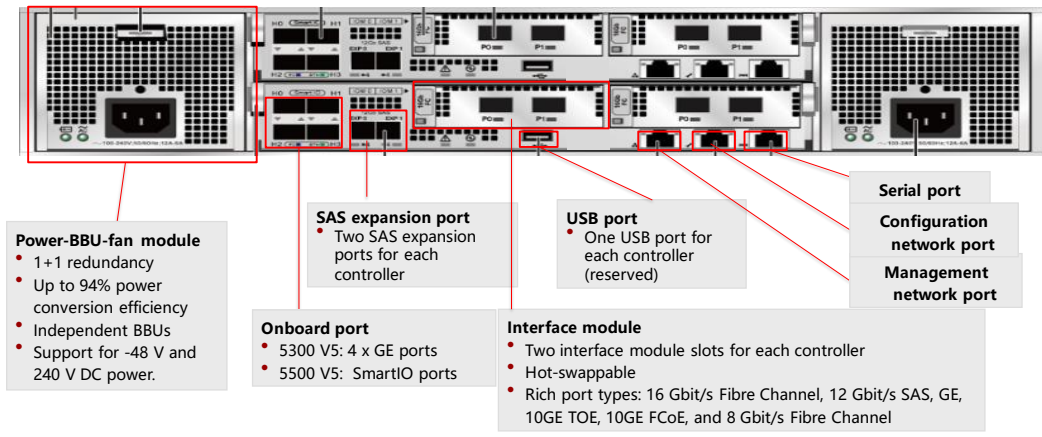


1 Disk module handle	2 Coffer disk label
3 Information plate (with ESN)	4 ID display of the controller enclosure
5 Power indicator/Power button	6 Disk module latch

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Note: The disk slots of a 2 U 12-disk-slot controller enclosure are numbered from 0 to 11 from left to right and then from top to bottom. Coffer disks are in slots 0 to 3.

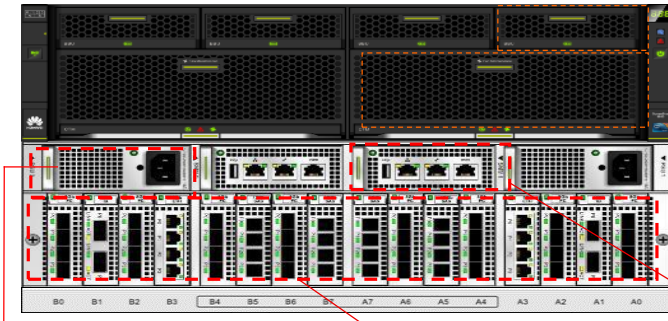
5300/5500 V5 Rear Panel



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Onboard GE ports are supported by the OceanStor 5300 V5 and onboard SmartIO ports are supported by the OceanStor 5500 V5.
- A controller enclosure supports 8 Gbit/s Fibre Channel interface modules, GE electrical interface modules, 10GE electrical interface modules, 16 Gbit/s Fibre Channel interface modules, 10 Gbit/s TOE, 10 Gbit/s FCoE (two ports), 10 Gbit/s FCoE (four ports), 56 Gbit/s IB interface modules, SmartIO interface modules, High-Density 8 Gbit/s Fibre Channel interface modules and 12 Gbit/s SAS expansion modules. The following figure uses a 16 Gbit/s Fibre Channel interface module of OceanStor 5500 V5 as an example.
- When the maintenance network port is used for management and maintenance, the maintenance network port can only be used by Huawei technical support for emergency maintenance and cannot be connected to the same network with the management network port. Otherwise, a network loopback may occur, causing a network storm. The initial value for the IP address of the maintenance network port is 172.31.128.101 or 172.31.128.102. The default subnet mask is 255.255.0.0. You are advised to only connect the management network port to the network.

5600/5800 V5 Controller Enclosure



BBU module
- 5600 V5: 1+1 redundancy (remaining two slots inserted by filler BBU modules); 5800 V5: 2+1 redundancy (remaining one slot inserted by a filler BBU module)
- DC/AC power failure protection

Controller module
- Support for dual controllers and four controllers
- Mainstream server platform
- Automatic frequency control and energy saving
- Fan module (integrated into a controller module but maintained independently)0

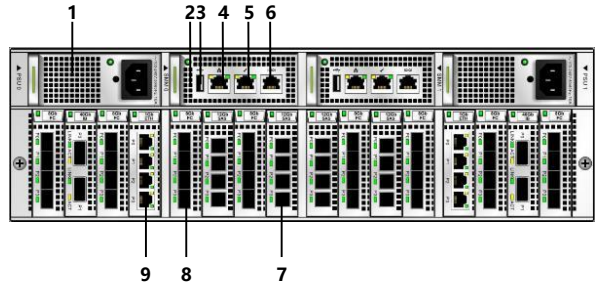
Power module
- 1+1 redundancy
- Up to 94% power conversion efficiency
- 240 V high-voltage DC

Interface module
- 16 interface module slots
- A maximum of 1250 disks
- Hot-swappable
- Rich port types: 16 Gbit/s Fibre Channel, 12 Gbit/s SAS, GE, 10GE TOE, 10GE FCoE, and 8 Gbit/s Fibre Channel

Management module
- 1+1 redundancy
- Hot-swappable
- Support for multi-controller Scale-out interconnection with a heartbeat mechanism

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

5600/5800 V5 Ports



1	Power module	2	Management board	3	USB port (reserved)	4	Management network port
5	Maintenance network port	6	Serial port	7	12 Gbit/s mini SAS HD	8	8 Gbit/s Fibre Channel port
9	GE port						

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

6800 V5 Controller Enclosure (1)

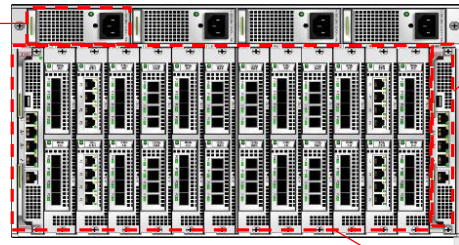


BBU module
- 3+1 redundancy
- DC/AC power failure protection

Controller module
- Support for dual controllers and four controllers
- Mainstream server platform
- Automatic frequency control and energy saving
- Fan module (integrated into a controller module but maintained independently)

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

6800 V5 Controller Enclosure (2)



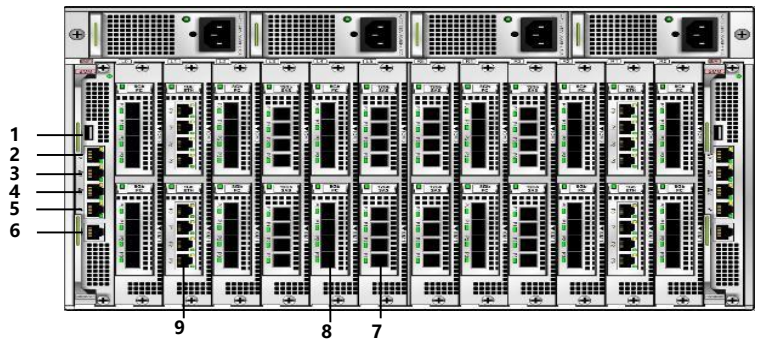
Management module
- 1+1 redundancy
- Hot-swappable
- Support for multi-controller Scale-out interconnection with a heartbeat mechanism

Interface module
- Dual controllers: 12 interface module slots; four controllers: 24 interface module slots
- A maximum of 2000 disks
- Hot-swappable
- Rich port types: 16 Gbit/s Fibre Channel, 12 Gbit/s SAS, GE, 10GE TOE, 10GE FCoE, and 8 Gbit/s Fibre Channel

Power module
- 1+1 redundancy
- Up to 94% power conversion efficiency
- 240 V high-voltage DC

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

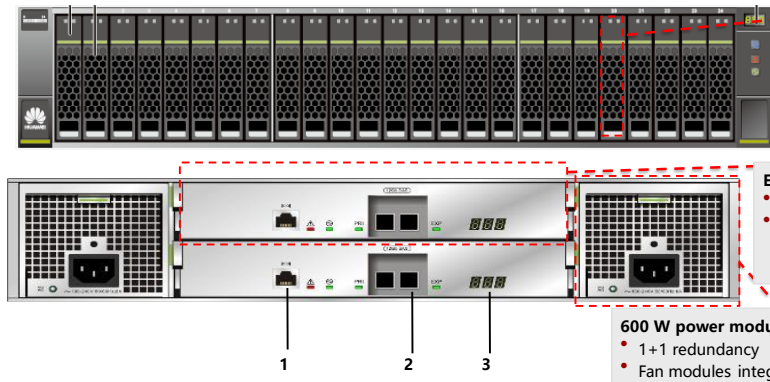
6800 V5 Ports



1	USB port (reserved)	2	Management network port 0	3	Management network port 1	4	Management network port 2		
5	Maintenance network port	6	Serial port	7	12 Gbit/s mini SAS HD	8	8 Gbit/s Fibre Channel port	9	GE port

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

2 U 2.5-inch Disk Enclosure



2.5-inch disk unit
- No disk connector
- Support for 6 Gbit/s SAS disks, NL-SAS disks, and 6 Gbit/s SAS SSDs

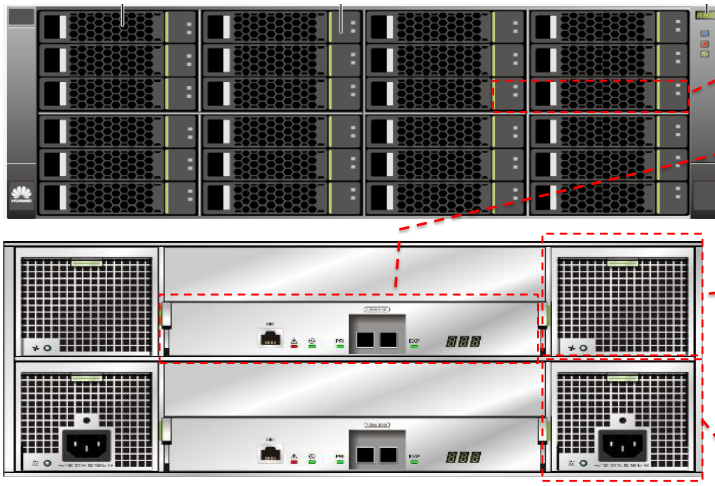
Expansion module
- Dual expansion modules
- 12 Gbit/s SAS uplink and downlink

600 W power module
- 1+1 redundancy
- Fan modules integrated into a disk enclosure (1+1 fan redundancy, namely, each power module equipped with one fan module)
- DC/AC power supplies

1 Serial port 2 mini SAS HD expansion port 3 Disk enclosure ID display

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

4 U 3.5-inch Disk Enclosure



3.5-inch disk unit
- No disk connector
- Support for 6 Gbit/s NL-SAS disks

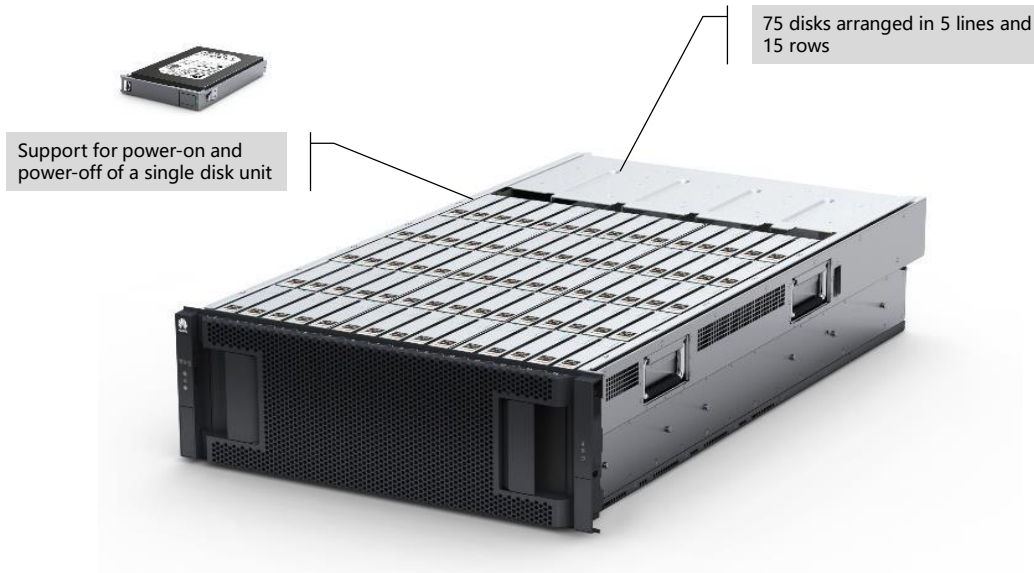
Expansion module
- Dual expansion modules
- 12 Gbit/s SAS uplink and downlink

Fan module
- 2 fan modules (2 fan modules and 2 power modules compose a heat dissipation module.)
- 5+1 redundant fan modules in a disk enclosure
- High-speed fans

600 W power module
- 1+1 redundancy
- Fan modules integrated into a disk enclosure
- Support for DC (-48 V)/AC (90 V to 240 V)
- Support for 240 V high-voltage DC

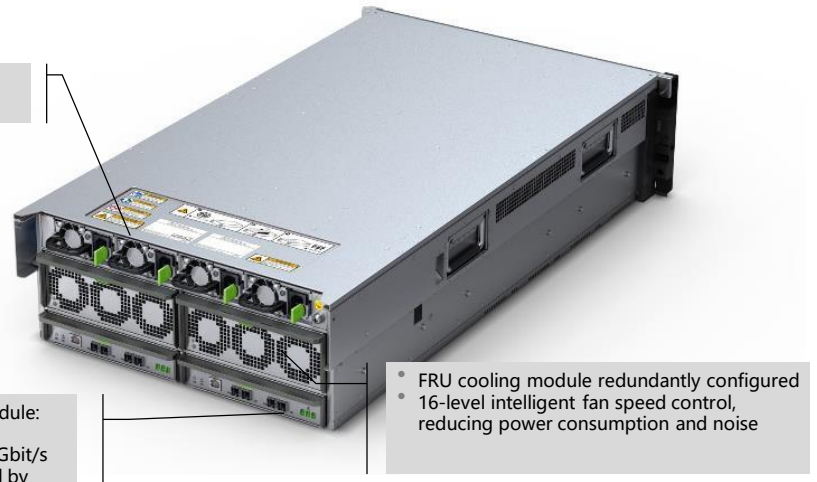
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

4 U 3.5-inch High-density Disk Enclosure (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

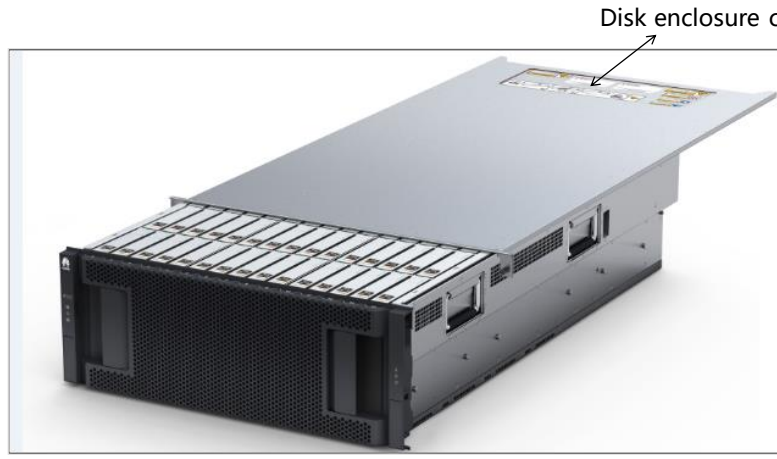
4 U 3.5-inch High-density Disk Enclosure (2)



- FRU PSU: 2+2 redundancy
- FRU SAS expansion module: 1+1 redundancy
- Four 4-channel wide 6 Gbit/s SAS data links provided by each I/O module; mini SAS HD port
- FRU cooling module redundantly configured
- 16-level intelligent fan speed control, reducing power consumption and noise

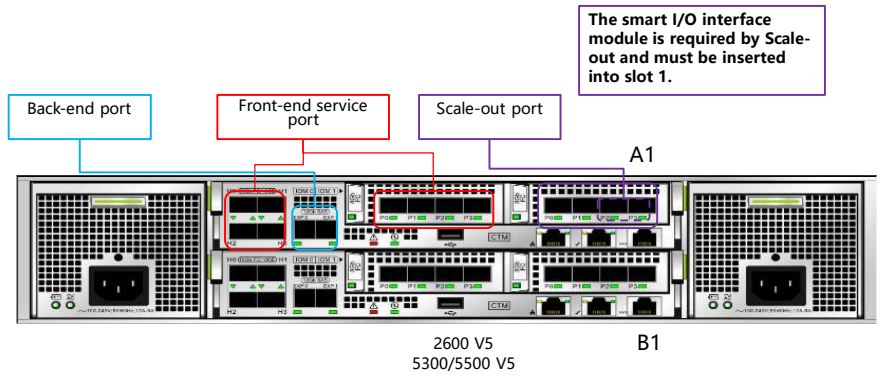
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

4 U 3.5-inch High-density Disk Enclosure (3)



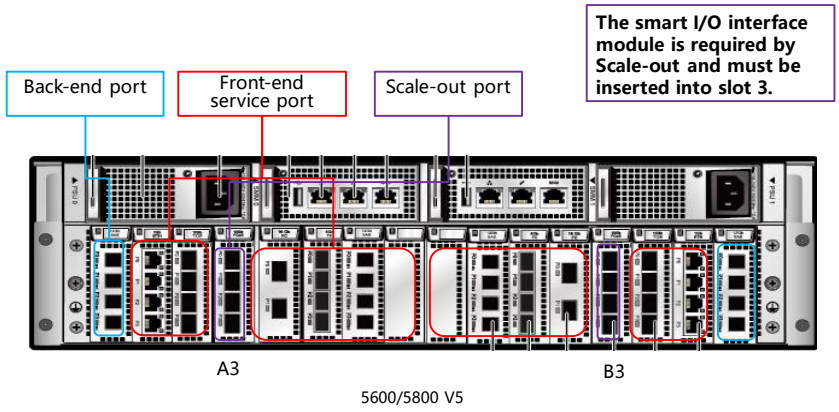
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

IP Scale-out (2 U)



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

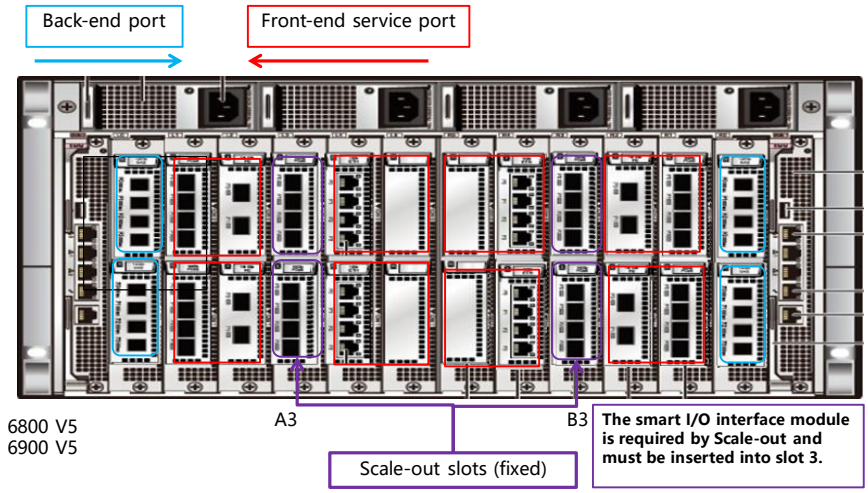
IP Scale-out (3 U)



5600/5800 V5

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

IP Scale-out (6 U)



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

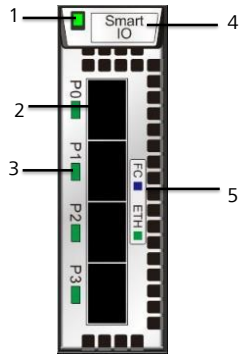
Contents

1. Product introduction
2. **Hardware instruction**
   - Hardware
   - Interface modules
3. Daily operation and maintenance

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Smart I/O Interface Module

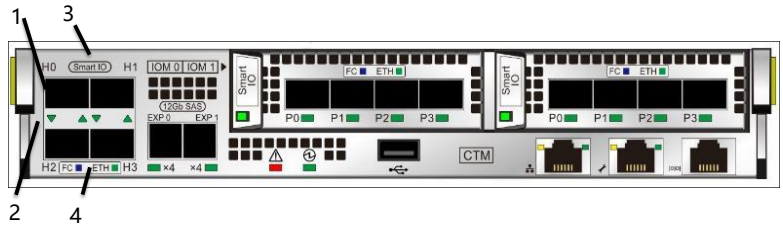


1	Power indicator/Hot Swap button
2	16 Gbit/s Fibre Channel/8 Gbit/s Fibre Channel/FCoE/iWARP(Scale-Out)
3	Port indicator (Link/Active/Mode indicator)
4	Module handle
5	Port working mode silkscreen

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Power indicator:
  - Green on: The module is working properly.
  - Blinking green: The module needs to be hot-swapped.
  - Red on: The module is faulty.
  - Off: The module is not powered on.
- Port indicator (Link/Active/Mode indicator):
  - Blinking blue slowly: The module is working in FC mode with link down.
  - Blinking blue quickly: The module is working in FC mode with link up and data is being transmitted.
  - Steady blue: The module is working in FC mode with link up but no data is being transmitted.
  - Blinking green slowly: The module is working in FCoE/iWARP mode with link down.
  - Blinking green quickly: The module is working in FCoE/iWARP mode with link up and data is being transmitted.
  - Steady green: The module is working in FCoE/iWARP mode with link up but no data is being transmitted.

Onboard Smart I/O Interface Module

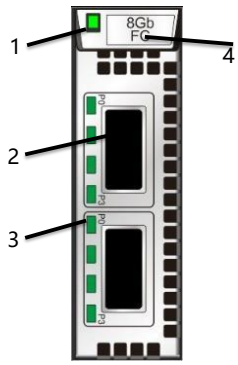


1	16 Gbit/s Fibre Channel/8 Gbit/s Fibre Channel/FCoE
2	Port indicator (Link/Active/Mode indicator)
3	Module handle
4	Port working mode silkscreen

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Port indicator (Link/Active/Mode indicator):
  - Blinking blue slowly: The module is working in FC mode with link down.
  - Blinking blue quickly: The module is working in FC mode with link up and data is being transmitted.
  - Steady blue: The module is working in FC mode with link up but no data is being transmitted.
  - Blinking green slowly: The module is working in FCoE mode with link down.
  - Blinking green quickly: The module is working in FCoE mode with link up and data is being transmitted.
  - Steady green: The module is working in FCoE mode with link up but no data is being transmitted.

8 Gbit/s Fibre Channel High-density Interface Module (1)



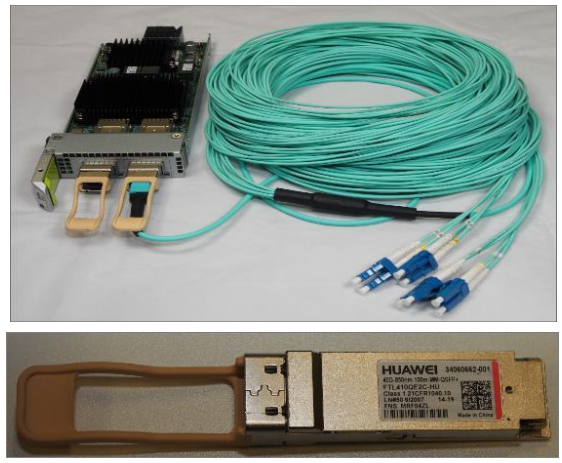
- The interface module provides two 4 x 8 Gbit/s Fibre Channel physical ports. The two ports are converted into eight 8 Gbit/s Fibre Channel optical ports using dedicated optical cables (no optical module connected to the eight ports). The eight ports are used as service ports to connect application servers to storage systems and used to receive data exchange commands from application servers.
- The interface module contains two 4 x 10 Gbit/s QSFP+ optical modules.
- The interface module is equipped with 15 m four-in-one dedicated optical cables.

1	Power indicator/Hot Swap button
2	8 Gbit/s Fibre Channel port
3	Port Link/Active indicator
4	Module handle/Silkscreen

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

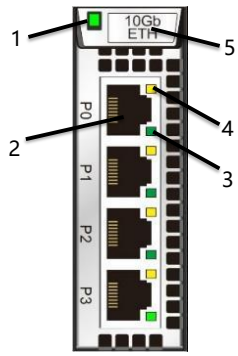
- Power indicator:
  - Green on: The module is working properly.
  - Blinking green: The module needs to be hot-swapped.
  - Red on: The module is faulty.
  - Off: The module is not powered on or can be hot-swapped.
- Port Link/Active indicator:
  - Blue on: Data is being transmitted at the rate of 8 Gbit/s between the storage system and application servers.
  - Blinking blue: Data is being transmitted.
  - Green on: Data is being transmitted at the rate of 4 Gbit/s or 2 Gbit/s between the storage system and application servers.
  - Blinking green: Data is being transmitted.
  - Red on: A port fault occurs.
  - Off: The port link is down.

8 Gbit/s Fibre Channel High-density Interface Module (2)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

10 Gbit/s ETH Electrical Interface Module

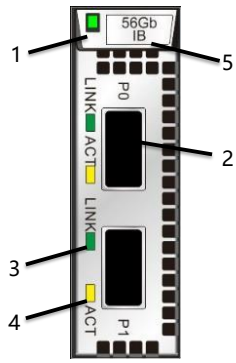


1	Power indicator/Hot Swap button
2	10 Gbit/s Ethernet port
3	Port Link/Active indicator
4	Port Speed indicator
5	Module handle

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Power indicator:
  - Green on: The module is working properly.
  - Blinking green: The module needs to be hot-swapped.
  - Red on: The module is faulty.
  - Off: The module is not powered on.
- Port Link/Active indicator:
  - Green on: The connection between the interface module and application servers is normal.
  - Blinking green: Data is being transmitted.
  - Off: The connection between the interface module and application servers is abnormal.
- Port Speed indicator:
  - Orange on: Data is being transmitted at the rate of 10 Gbit/s between the storage system and application servers.
  - Off: Data is being transmitted at the rate of lower than 10 Gbit/s between the storage system and application servers.

56 Gbit/s IB Interface Module

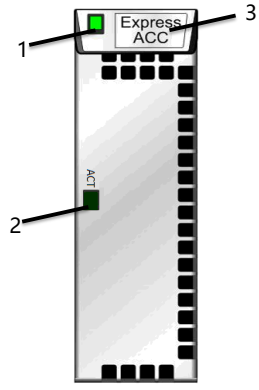


1	Power indicator/Hot Swap button
2	4 lane x 14 Gbit/s IB electrical port
3	Port Link indicator
4	Port Active indicator
5	Module handle/Silkscreen

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Power indicator:
  - Green on: The module is working properly.
  - Blinking green: The module needs to be hot-swapped.
  - Red on: The module is faulty.
  - Off: The module is not powered on or can be hot-swapped.
- Port Link indicator:
  - On: The connection between the module and application servers is normal.
  - Off: The port link is down.
- Port Active indicator:
  - Blinking: Data is being transmitted.
  - Steady on: No data is being transmitted.

Deduplication and Compression Acceleration Interface Module (ACC)



1	Power indicator/Hot Swap button
2	Port Active indicator
3	Module handle/Silkscreen

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Power indicator:
  - Green on: The module is working properly.
  - Blinking green: The module needs to be hot-swapped.
  - Red on: The module is faulty.
  - Off: The module is not powered on or is removable.
- Port Active indicator:
  - Green on: No data is being processed.
  - Blinking green: Data is being processed.
  - Off: The module is not working properly.

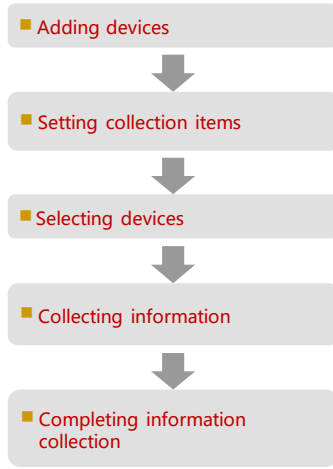
Contents

1. Product introduction
2. Hardware introduction
3. **Daily operation and maintenance**
   - Information collection tool
   - Inspection tool
   - Upgrade tool
   - FRU replacement tool

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

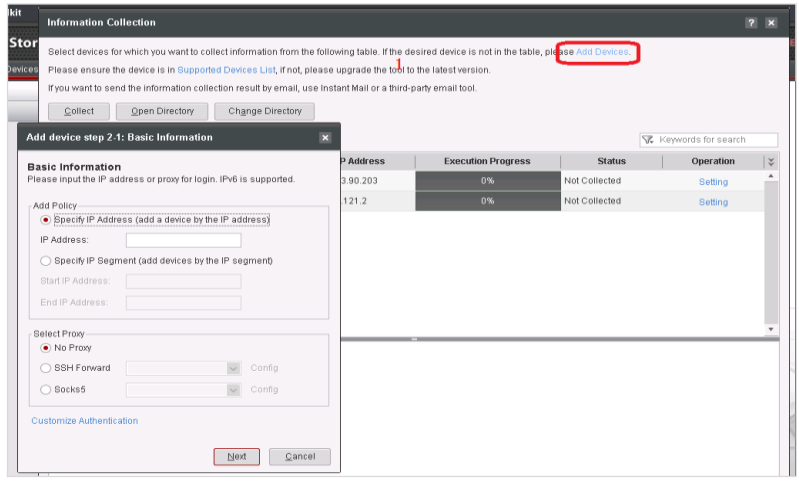
Working Process



- Adding devices: Add devices whose information you want to collect.
- Setting collection items: Select desired collection items.
- Selecting devices: Select devices whose information you want to collect.
- Collecting information: Run the commands to collect information and package collected information.
- Information collection completed: After the information collection is complete, you can click Open Directory to view collected information.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

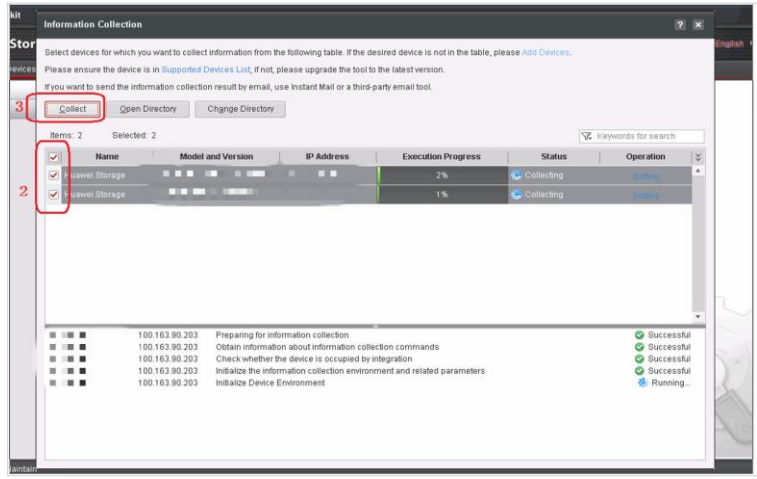
Adding Devices



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Collecting Information



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

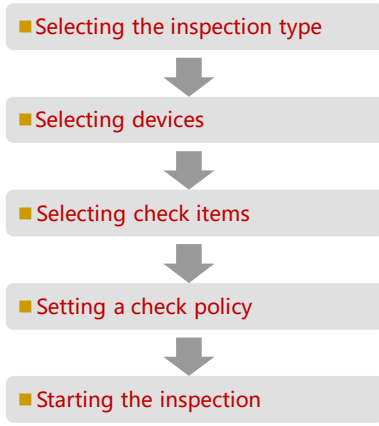
Contents

1. Product introduction
2. Hardware introduction
3. **Daily operation and maintenance**
   - Information collection tool
   - Inspection tool
   - Upgrade tool
   - FRU replacement tool

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

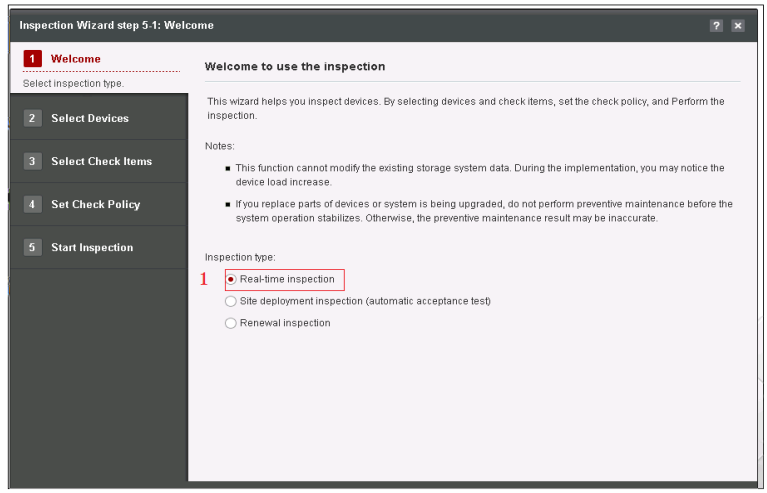
Working Process



- Selecting the inspection type: Select the inspection type for specific scenarios.
- Selecting devices: Select devices that you want to inspect.
- Selecting check items: Select desired check items.
- Setting a check policy: Set a location for saving inspection reports.
- Starting the inspection: Run the commands to start the inspection.

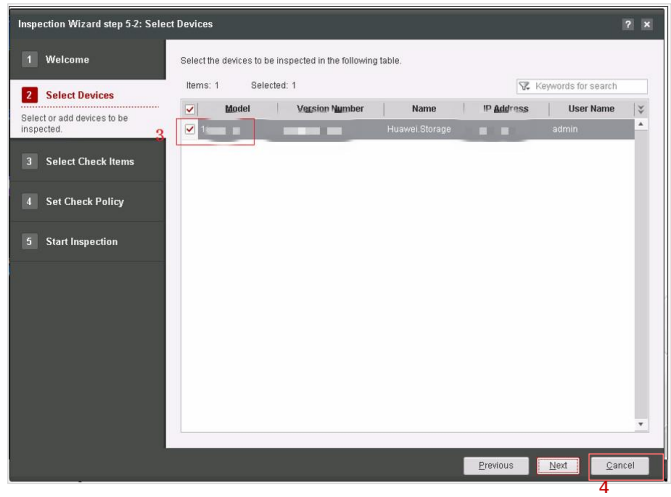
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Selecting the Inspection Type



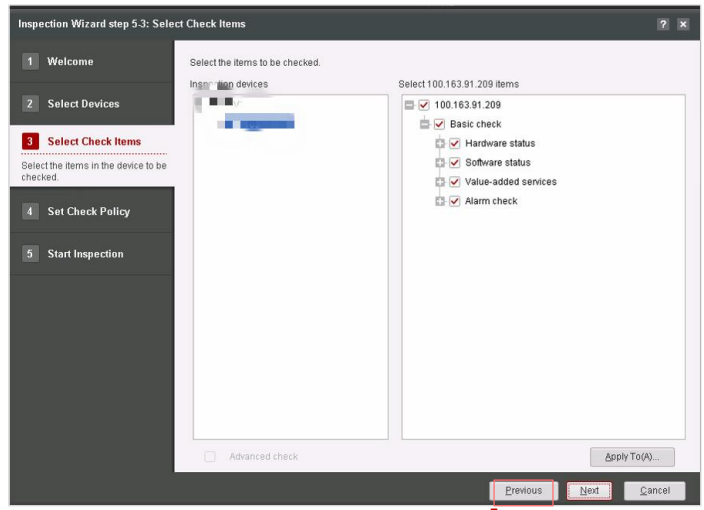
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Selecting Devices



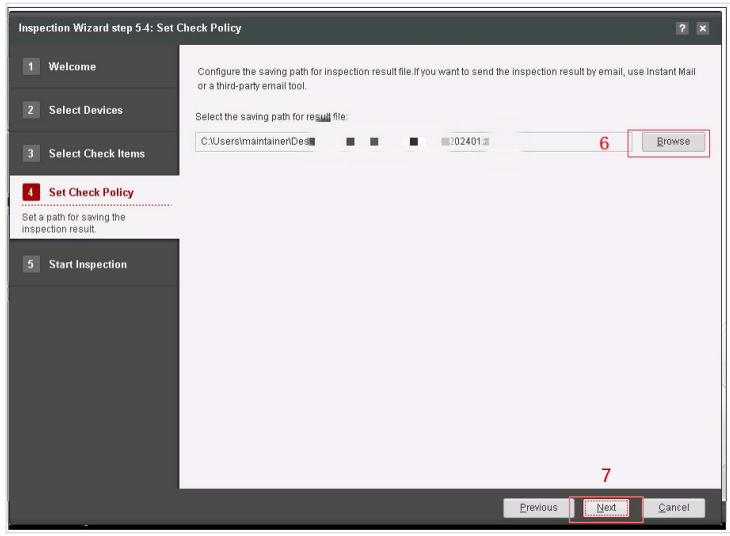
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Selecting Check Items



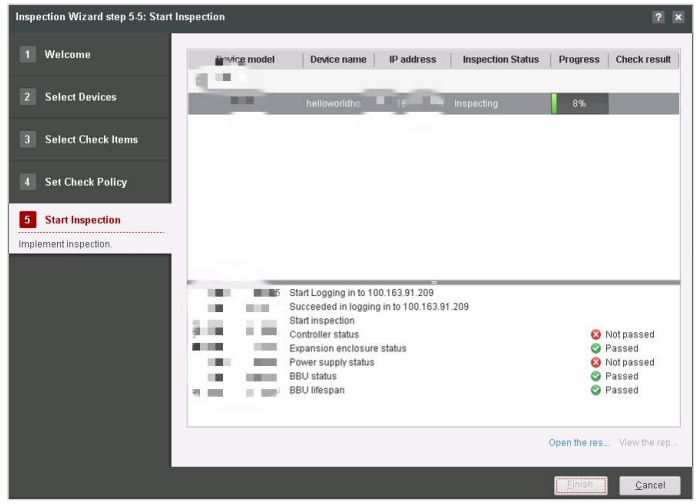
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Setting a Check Policy



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Starting Inspection



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

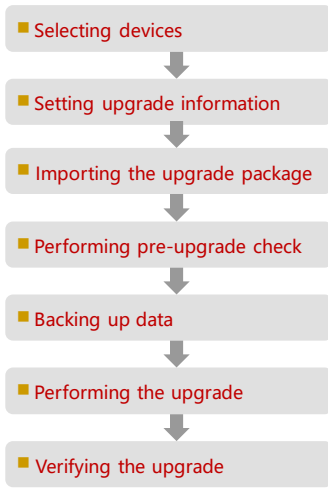
HUAWEI

Contents

1. Product introduction
2. Hardware introduction
3. **Daily operation and maintenance**
   - Information collection tool
   - Inspection tool
   - **Upgrade tool**
   - FRU replacement tool

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

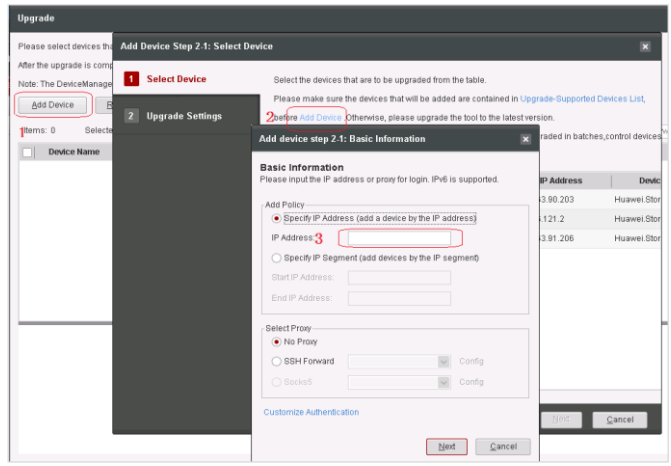
Working Process



- Selecting devices: Select devices that you want to upgrade.
- Setting upgrade information: Set a path for saving the upgrade package, a path for saving backup data, and an upgrade mode (online or offline).
- Importing the upgrade package: Import the upgrade package into the storage array.
- Performing pre-upgrade check: Check whether issues that may affect the upgrade exist.
- Backing up data: Back up the database data.
- Performing the upgrade: Notify the storage array to perform the upgrade.
- Verifying the upgrade: Check whether issues exist.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

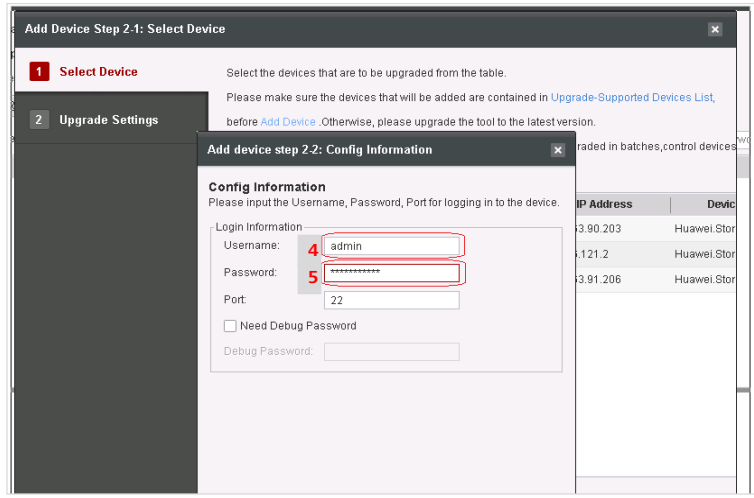
Selecting Devices (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

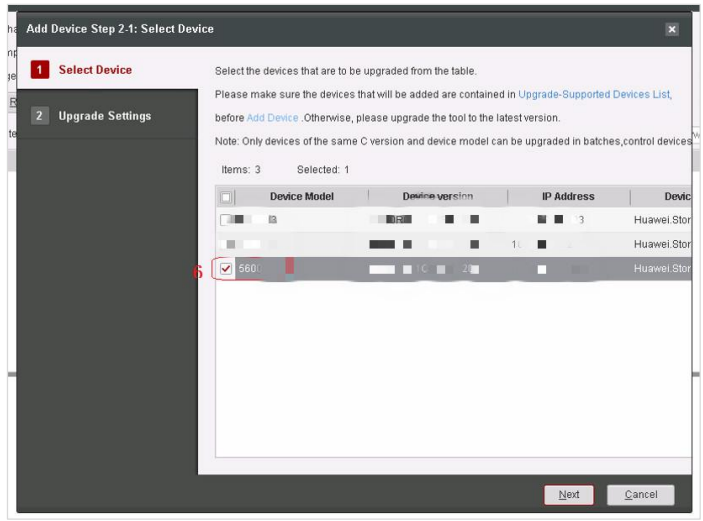
Selecting Devices (2)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

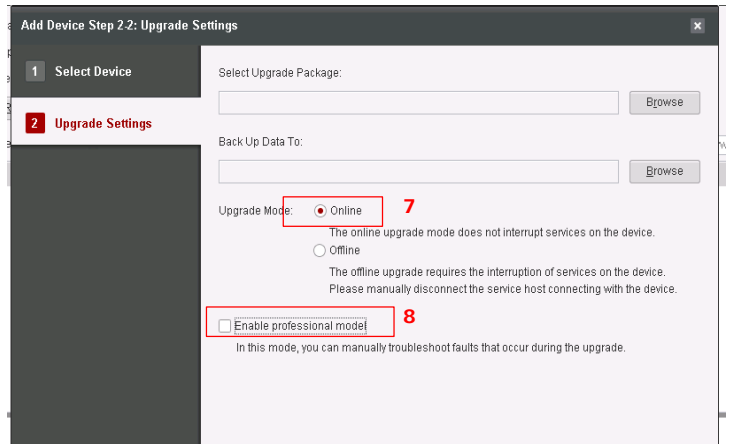
HUAWEI

Selecting Devices (3)



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

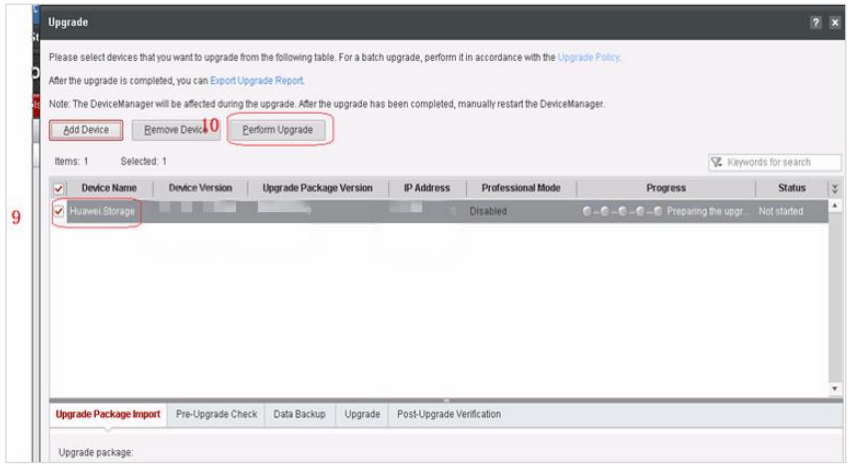
Setting Upgrade Information



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

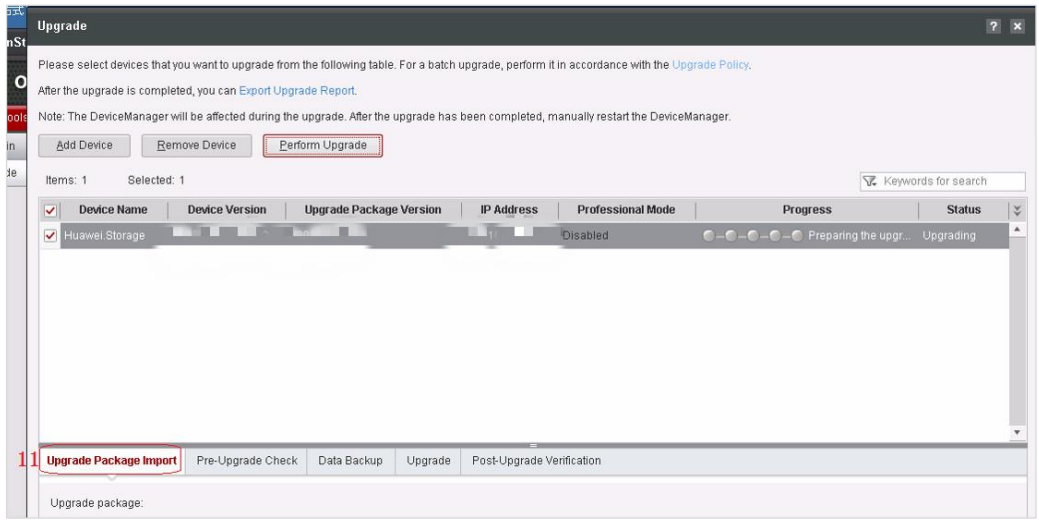
HUAWEI

Clicking Perform Upgrade



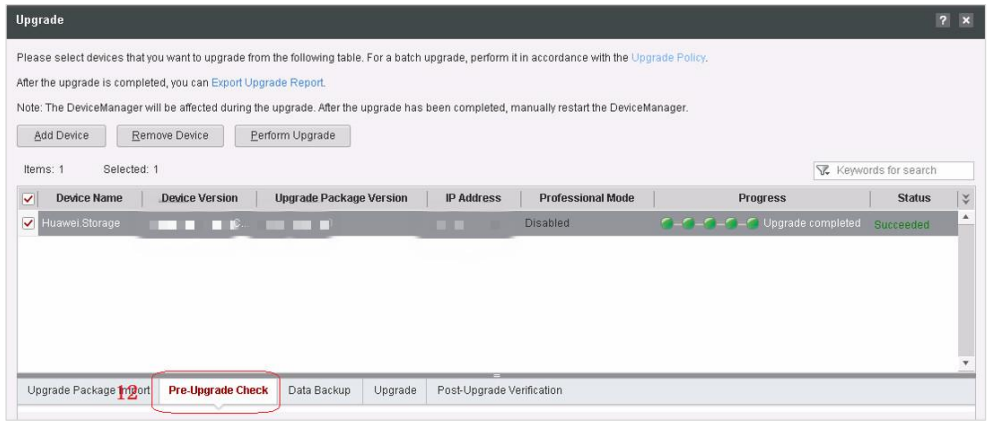
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Importing the Upgrade Package



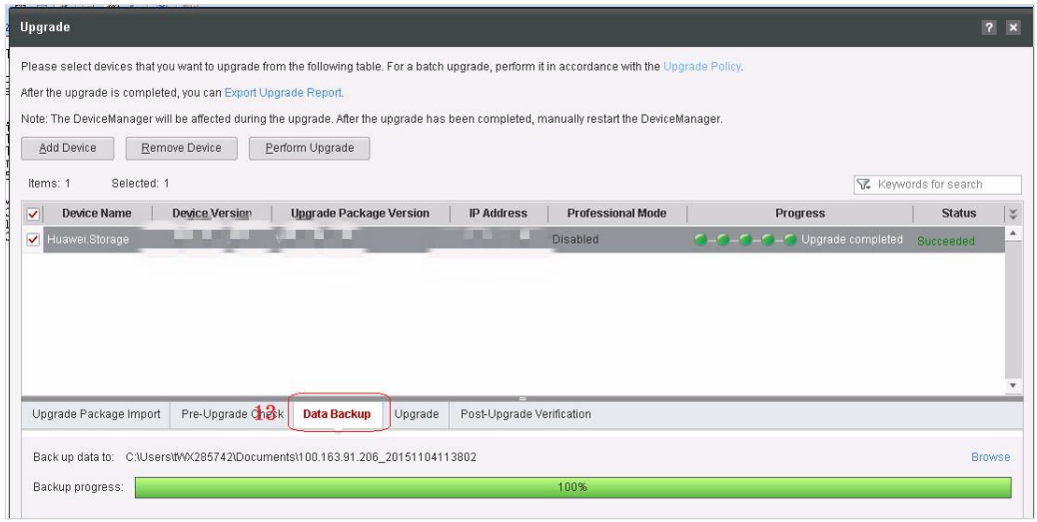
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Performing Pre-upgrade Check



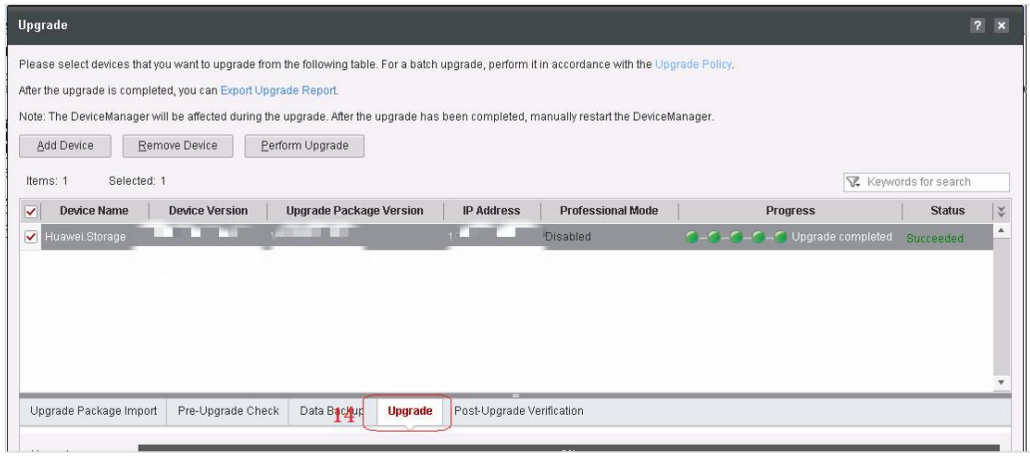
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Backing Up Data



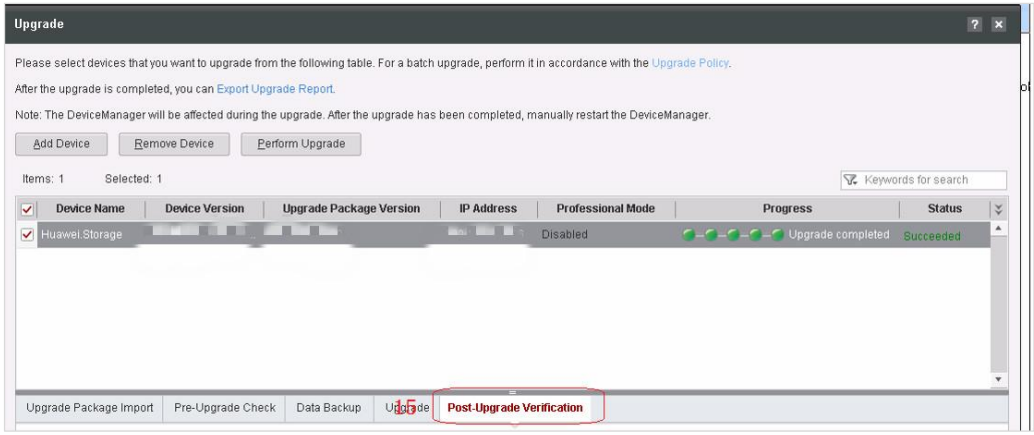
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Upgrading the Devices



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Verifying the Upgrade



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

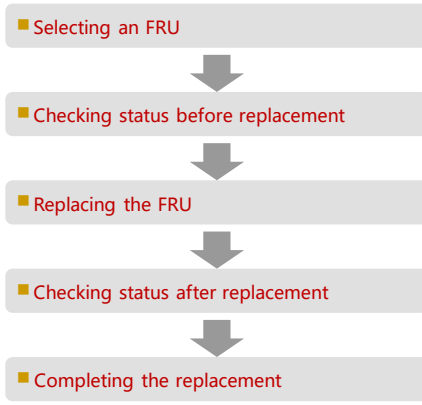
Contents

1. Product introduction
2. Hardware introduction
3. **Daily operation and maintenance**
   - Information collection tool
   - Inspection tool
   - Upgrade tool
   - FRU replacement tool

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Working Process

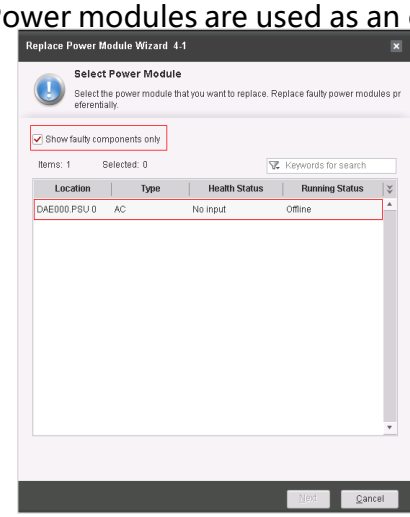


- Selecting an FRU: Select an FRU that you want to replace. Replace faulty FRUs first.
- Checking status before replacement: Check whether the FRU needs to be replaced. If the FRU does not need to be replaced, an error will be reported.
- Replacing the FRU: Replace the FRU following the procedure described in the replacement guide.
- Checking status after replacement: Check whether the new FRU is working correctly. If the new FRU is working incorrectly, an error will be reported.
- Replacement completed: A message is displayed indicating that the replacement is completed.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Selecting an FRU

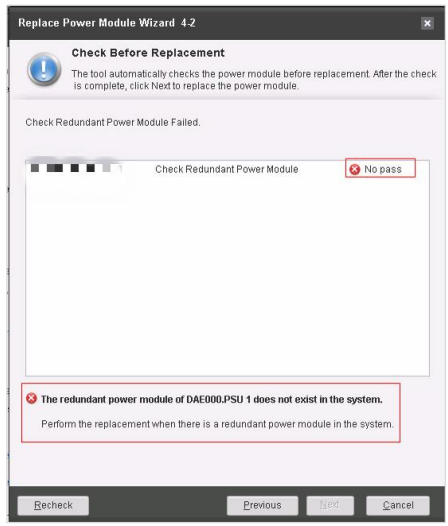
- Power modules are used as an example.



- By default, only faulty components are displayed. If you want to select other components, deselect **Show faulty components only**.
- Component silkscreens such as **CTE0.PSU0** are used to distinguish FRUs.
- You can click **Next** to start the pre-replacement check only after selecting the FRUs that you want to replace.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

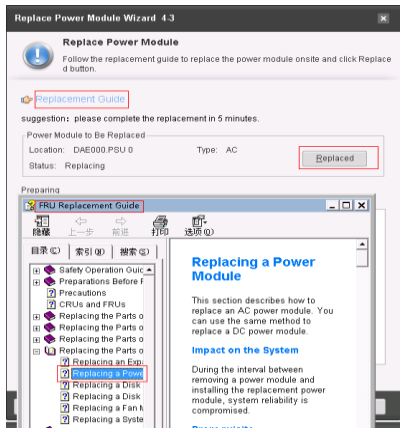
Checking Status Before Replacement



- The number of check items varies with the impact scope of FRUs.
- If the check result of one item is negative, causes and recommended actions are displayed. After troubleshooting, you can click Recheck to start the check from the item.
- You can click Next to replace FRUs only after the results of all check items are positive.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

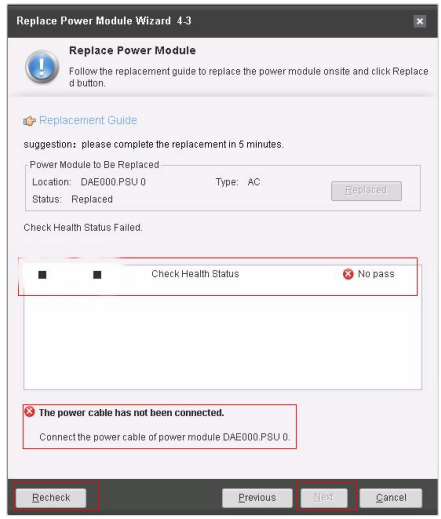
Replacing the FRU



- You can click **Replacement Guide** to read the replacement guide of the FRU to be replaced.
- Strictly follow instructions in the *Replacement Guide* to replace the FRU.
- After the replacement is complete, you can click **Replaced** to start the post-replacement check.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

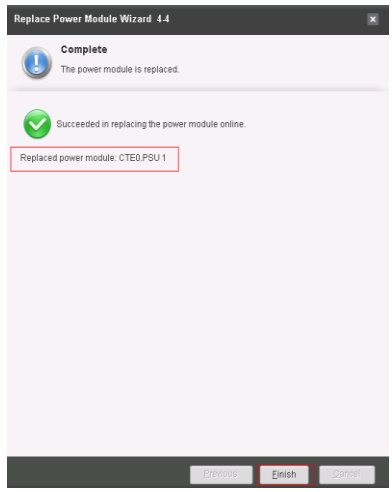
Checking Status After Replacement



- The number of check items varies with the impact scope of FRUs.
- If the check result of one item is negative, causes and recommended actions are displayed. After troubleshooting, you can click **Recheck** to start the check from the item.
- You can click **Next** to complete the FRU replacement only after the results of all check items are positive.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Completing the Replacement



- A message is displayed indicating that the FRU replacement is complete.
- Click **Finish** to exit.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. What are the common storage technologies used in Huawei's storage solutions?
2. What is the product positioning of OceanStor V5 storage?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What are the common storage technologies used in Huawei's storage solutions?
  - SmartThin
  - SmartTier
  - SmartQoS
  - SmartPartition
- What is the product positioning of OceanStor V5 storage?
  - Brand-new hardware architecture delivering industry-leading performance and specifications; Convergence of SAN and NAS; Outstanding scalability and reliability; Up to sixteen controllers, etc.

Summary

- Product positioning
- Hardware introduction
- Routine Operation and Maintenance of the Information Collection Tool.
- Routine Operation and Maintenance of the Inspection Tool.
- Routine Operation and Maintenance of the FRU Replacement Tool.
- Routine Operation and Maintenance of the Upgrade Tool.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Cloud DC solution
  - http://e.huawei.com/cn/solutions/business-needs/data-center

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Data Center Network Technology Application Basis

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This course introduces the most basic networking technologies in the data center.
- The principle of VLAN (Virtual Local Area Network) technology, and the application of VLAN technology in the network.
- The basic principles and application scenarios of link aggregation.
- The basic principles and applications of OSPF unicast routing protocols.
- The basic principles and application scenarios of BGP.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As the network scale expands increasingly, users propose increasingly high requirements on Ethernet backbone network bandwidth and reliability. Originally, to increase the bandwidth, users use high-speed cards or devices supporting high-speed interface cards to replace old interface cards or devices. This solution, however, is costly and inflexible.
- Link aggregation helps increase bandwidth by bundling a group of physical interfaces into a single logical interface, without having to upgrade hardware. In addition, link aggregation provides link backup mechanisms, greatly improving link reliability.

Objectives

- Upon completion of this course, you will be able to:
  - Describe the principle of VLAN technology
  - Describe the different port link types and characteristics
  - Describe the functions and purpose of link aggregation
  - Understand the basic principles of LACP
  - Understand the basic principles of OSPF
  - Describe the route calculation process of link state algorithm
  - Describe the establishment of BGP neighbor relationships

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

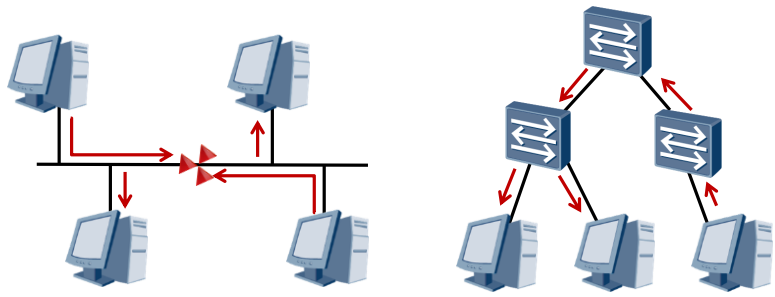
Contents

1. **VLAN technology**
   - VLAN overview
     - VLAN principle description
     - Inter-VLAN communication
2. Link aggregation technology
3. OSPF protocol
4. BGP protocol

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LAN Limitations

- No broadcast domain to manage expanding local networks.

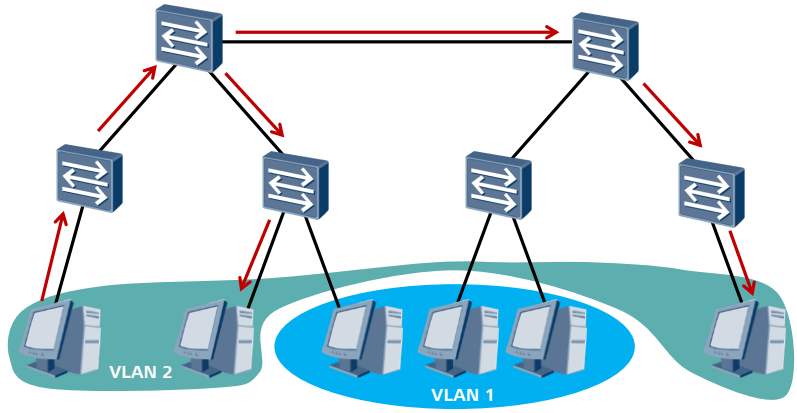


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As local networks expand, traffic increases and broadcasts become more common. There are no real boundaries within such an expanding network, causing interrupts and growing traffic utilization to occur. Traditionally, the alternative option was to implement a layer three device within the local network to generate broadcast domains, however in doing so additional expense was incurred and the forwarding behavior of such devices did not provide as efficient throughput as found with switches, leading to bottlenecks at transit points between broadcast domains.

VLAN Technology

- A VLAN enables logical isolation of traffic at the data link layer.

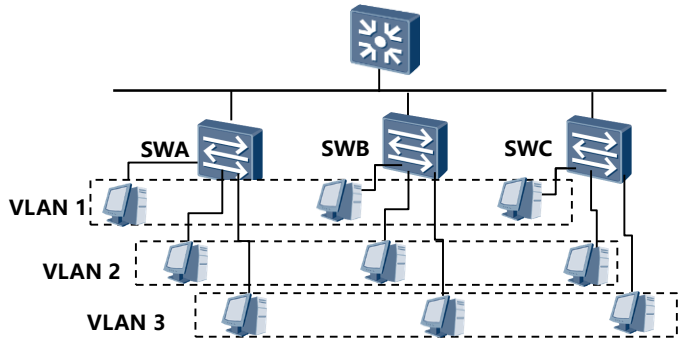


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The principle of VLAN technology was introduced that enabled traffic isolation at the data link layer. VLAN technology has the added advantage of traffic isolation without the limitation of physical boundaries. Users can be physically dispersed but still be associated as part of a single broadcast domain, logically isolating users from other user groups at the data link layer. Today VLAN technology is applied as a solution to a variety of challenges.

What is VLAN

- The virtual local area network (VLAN) technology groups devices on a physical LAN into different logical LAN segments.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

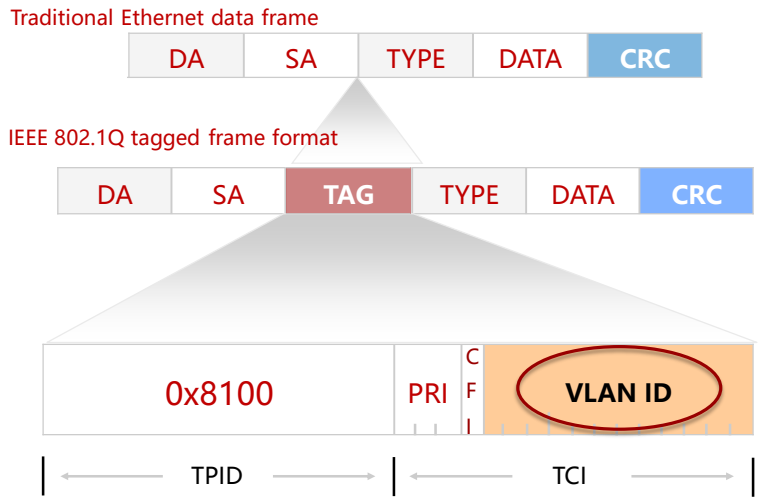
- Devices in a VLAN can directly communicate with each other, whereas devices in different VLANs cannot directly communicate with each other. Broadcast packets are forwarded within a VLAN, improving network security. For example, if different enterprise users located in the same business building construct their LANs separately, the investment is high; if these enterprise users share the same LAN of the business building, the information security cannot be ensured. If the VLAN technology is used, these enterprises can share the LAN resources without risks to their private information.
- The preceding networking diagram shows a typical VLAN application network. SwitchA, SwitchB, and SwitchC are placed in different floors of a building. Each switch is connected to three PCs that belong to different VLANs.

Contents

1. **VLAN technology**
   - VLAN overview
   - VLAN principle description
   - Inter-VLAN communication
2. Link aggregation technology
3. OSPF protocol
4. BGP protocol

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Format of the VLAN Frame

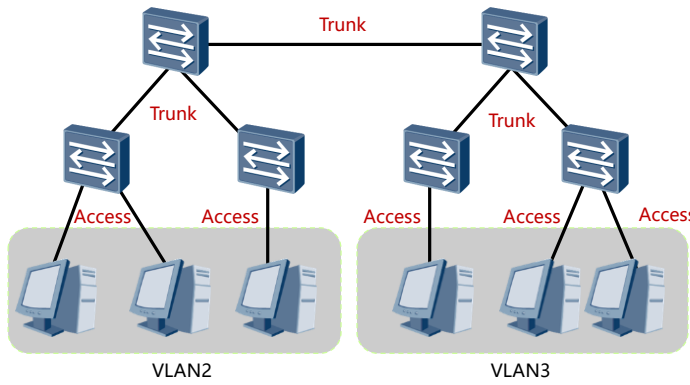


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The 802.1Q tag including the following fields:
- TPID
- Tag Protocol Identifier (TPID), indicating the frame type. The value 0x8100 indicates an 802.1Q-tagged frame. An 802.1Q-incapable device discards the 802.1Q frames.
- IEEE 802.1Q protocol defines the value of the field as 0x8100. However, manufacturers can define their own TPID values and users can then modify the value to realize interconnection of devices from different manufacturers.
- PRI
- Priority. The length is 3 bits. The value ranges from 0 to 7. A larger value indicates a higher priority. If congestion occurs, the switch sends packets with higher priorities first.
- CFI
- Canonical Format Indicator (CFI). The length is 1 bit. This field indicates whether the MAC address is in the canonical format. If the CFI field is 0, the MAC address is in the canonical format. If the CFI field is 1, the MAC address is in the non-canonical format. This field is used to identify the Ethernet frame, fiber distributed digital interface (FDDI) frame, and token ring frame. The CFI field is 0 in the Ethernet frame.
- VID
- VLAN ID. The field is 12 bits. This field indicates to which VLAN the frame belongs. On S series mid-range and low-end switch, the value of the field ranges from 0 to 4095. VLAN IDs 0 and 4095 are reserved and unavailable to users.

Link Types

- A trunk represents a backbone for the transmission of VLAN traffic between switches.



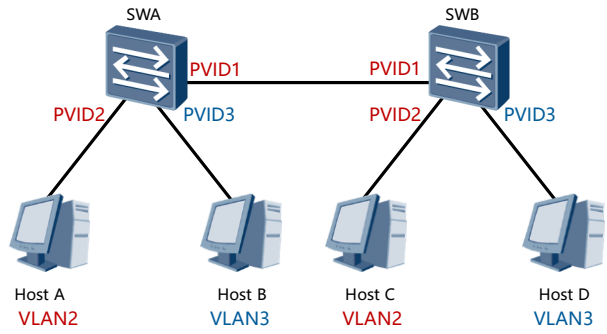
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Ethernet links fall into the following types, depending on the number of allowed VLANs:
  - Access link: An access link can transmit data frames of only one VLAN. It connects a switch to a user terminal, such as a host, server, and simplified Layer 2 switch. Generally, user terminals do not need to know the VLANs to which they belong and cannot identify tagged frames; therefore, only untagged frames are transmitted along an access link.
  - Trunk link: A trunk link can transmit data frames from multiple VLANs. It connects a switch to another switch or a router. Frames on a trunk link must be tagged so that other network devices can correctly identify VLAN information in the frames.

PVID

- PVID represents the default VLAN for each interface.
- The PVID is set to VLAN 1 for all ports by default.

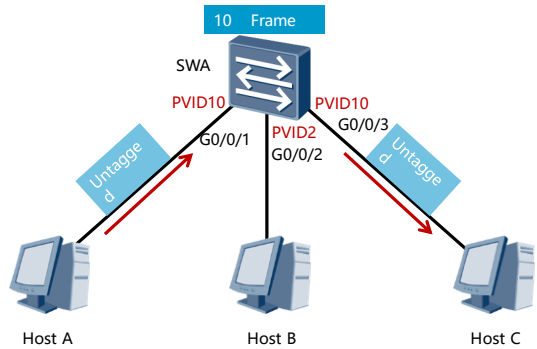


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The default VLAN ID of an interface is called the port default VLAN ID (PVID). Frames processed in a switch all carry VLAN tags. When the switch receives an untagged frame, it adds a VLAN tag to the frame according to the default VLAN of the interface that receives the frame. The PVID is used in the following scenarios:
  - When an interface receives an untagged frame, the interface adds a tag with the PVID to the frame and sends the frame to the switch for processing. When an interface receives a tagged frame, the switch does not add a tag with the PVID to the frame.
  - When an interface sends a frame in which the VLAN ID is the same as the PVID, the switch removes the tag from the frame before sending it out from the interface.

Port Types - Access

- An access interface often connects to a user terminal.
- Access interfaces can only receive and send untagged frames, and can add only a unique VLAN tag to untagged frames.

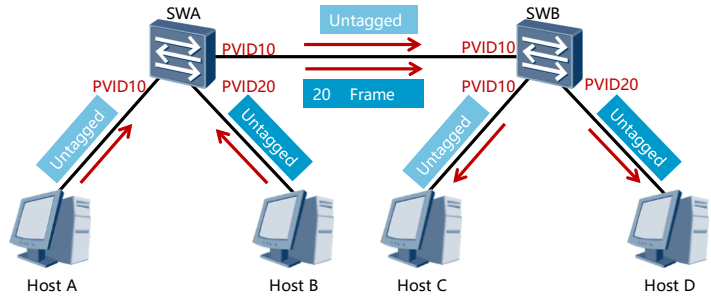


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- An access interface often connects to a user terminal such as a user host or server that cannot identify VLAN tags, or is used when VLANs do not need to be differentiated. Access interfaces can only receive and send untagged frames, and can add only a unique VLAN tag to untagged frames.
- An access interface has the following characteristics:
  - Only the frame whose VLAN ID is the same as the PVID of the access interface is allowed..
  - When the access interface receives an untagged frame, the switch adds the PVID of the interface to the frame.
  - The access interface sends only untagged frames to the peer device.
- In the example，upon receiving the frame, the switch will associate the frame with VLAN 10 based on the PVID of the interface. The switch is able to identify at the port interface the PVID and make a decision as to whether the frame can be forwarded. In the case of Host C the PVID matches the VLAN ID in the VLAN tag, for which the tag is removed and the frame forwarded. For Host B however the frame and the PVID differ, and therefore the frame is restricted from being forwarded to this destination.

Port Types - Trunk

- A trunk interface often connects to a switch, router, AP, or voice terminal.
- It allows tagged frames from multiple VLANs and untagged frames from only one VLAN.

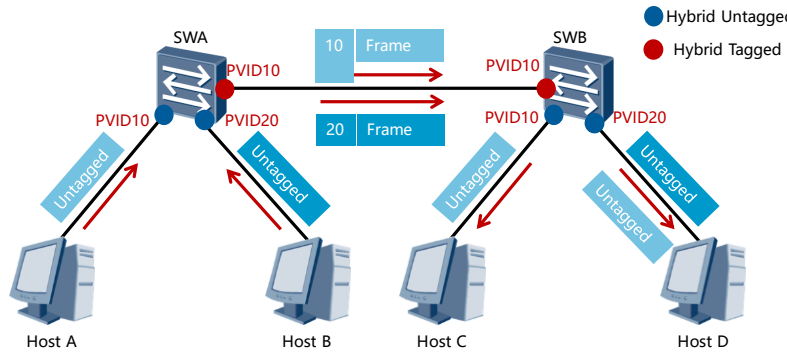


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A trunk interface often connects to a switch, router, AP, or voice terminal that can receive and send tagged and untagged frames simultaneously. It allows tagged frames from multiple VLANs and untagged frames from only one VLAN.
- A trunk interface has the following features:
  - A trunk interface allows tagged frames from multiple VLANs to pass through.
  - If a tagged frame sent out through a trunk interface carries a VLAN ID that is the same as the default one of the interface, the switch removes the VLAN tag from the frame. It is because the PVID of each interface is unique. Only in this case, frames forwarded by the trunk interface are untagged.
  - If a tagged frame sent out through a trunk interface carries a VLAN ID that is different from the default one of the interface, the switch directly forwards the frame without performing any other actions.
- The example demonstrates a trunk interface assigned with a PVID of 10, for which it should be assumed that all VLANs are permitted to traverse the trunk link. Only frames associated with VLAN 10 will be forwarded without the VLAN tag, based on the PVID. For all other VLAN frames, a VLAN tag must be included with the frame and be permitted by the port before the frame can be transmitted over the trunk link. Frames associated with VLAN 20 are carried as tagged frames over the trunk link.

Port Types - Hybrid

- A hybrid interface can connect to not only a user terminal or network device.
- It allows tagged frames from multiple VLANs.

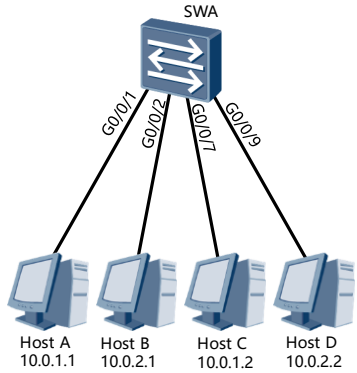


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- A hybrid interface can connect to not only a user terminal (such as a user host or server) or network device (such as a hub or simplified Layer 2 switch) that cannot identify tags, but also a switch, router, voice terminal, or AP that can receive and send tagged and untagged frames. It allows tagged frames from multiple VLANs. Frames sent out from a hybrid interface are tagged or untagged according to the VLAN configuration.
- Ports which are considered untagged will generally receive untagged frames from end systems, and be responsible for adding a tag to the frame based on the Port VLAN ID (PVID) of the port. One of the key differences is in the hybrid port' s ability to selectively perform the removal of VLAN tags from frames that differ from the PVID of the port interface. In the example, Host D is connected to a port which specifies a Port VLAN ID of 20, whilst at the same time is configured to allow for the removal of the tag from frames received from VLAN 10, thereby allowing Host D to receive traffic from both VLANs 10 & 20.
- Hybrid Ports that are tagged will operate in a similar manner as a regular trunk interface, however one major difference exists. VLAN frames that both match the PVID and are permitted by the port will continue be tagged when forwarded.

VLAN Assignment Methods



Assignment Method	VLAN 5	VLAN 10
Port based	G0/0/1, G0/0/7	G0/0/2 G0/0/9
MAC based	00-01-02-03-04-AA 00-01-02-03-04-CC	00-01-02-03-04-BB 00-01-02-03-04-DD
IP Subnet based	10.0.1.*	10.0.2.*
Protocol based	IP	IPX
Policy based	10.0.1.* + G0/0/1+ 00-01-02-03-04-AA	10.0.2.* + G0/0/2 + 00-01-02-03-04-BB

- Five methods of VLAN assignment are possible.
- Port based VLAN assignment is the default assignment method.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- VLAN assignment can be implemented based on one of five different methods, including Port based, MAC based, IP Subnet based, Protocol based and Policy based implementations. The port based method represents the default and most common method for VLAN assignment. Using this method, VLANs are classified based on the port numbers on a switching device. The network administrator configures a Port VLAN ID (PVID), representing the default VLAN ID for each port on the switching device. When a data frame reaches a port, it is marked with the PVID if the data frame carries no VLAN tag and the port is configured with a PVID. If the data frame carries a VLAN tag, the switching device will not add a VLAN tag to the data frame even if the port is configured with a PVID.
- Using the MAC address assignment method, VLANs are classified based on the MAC addresses of network interface cards (NICs). The network administrator configures the mappings between MAC addresses and VLAN IDs. In this case, when a switching device receives an untagged frame, it searches the MAC-VLAN table for a VLAN tag to be added to the frame according to the MAC address of the frame. For IP subnet based assignment, upon receiving an untagged frame, the switching Device adds a VLAN tag to the frame based on the IP address of the packet header.

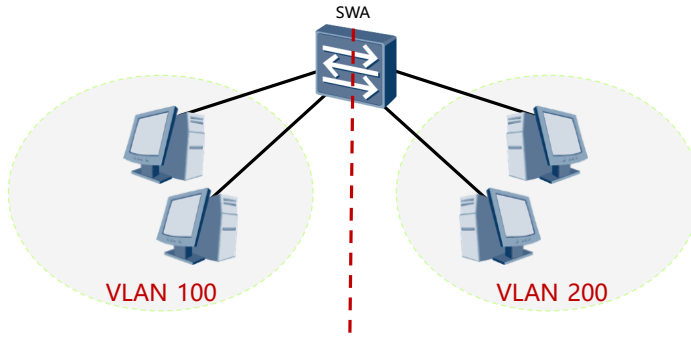
Contents

1. **VLAN technology**
   - VLAN overview
   - VLAN principle description
   - Inter-VLAN communication
2. Link aggregation technology
3. OSPF protocol
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

VLAN Disadvantages

- Attempts to limit broadcast domain size through VLAN implementation isolates users.

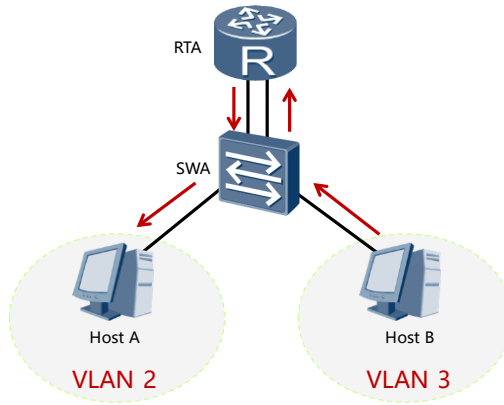


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The general principle of VLAN implementation is to isolate networks as a means of minimizing the size of the existing broadcast domain, however in doing so, many users are cut off from other users within other VLAN domains and require that layer three (IP) communication be established in order for those broadcast domains to re-establish communication through reachable routes. The implementation of a layer three switch offers an ideal means for supporting VLAN routing whilst reducing operating costs. One of the constraints however of VLAN routing is the need for strict IP address management.
- Generally however the VLAN routing principle is applicable to small scale networks on which users belong to different network segments and IP addresses of users are seldom changed.

Inter-VLAN Communication

- Flows between different VLANs cannot directly cross VLAN boundaries, we can use routers so that messages can be forwarded from one VLAN to another VLAN.



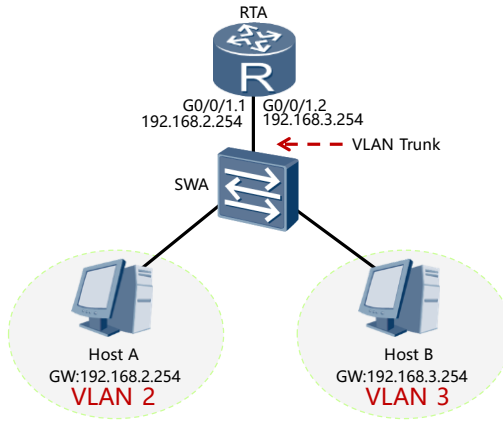
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Flows between different VLANs cannot directly cross VLAN boundaries, and route is needed to forward packet from one VLAN to another VLAN.

Route Selection in Inter-VLAN Communication

- A single trunk supports VLAN routes by using sub-interfaces.

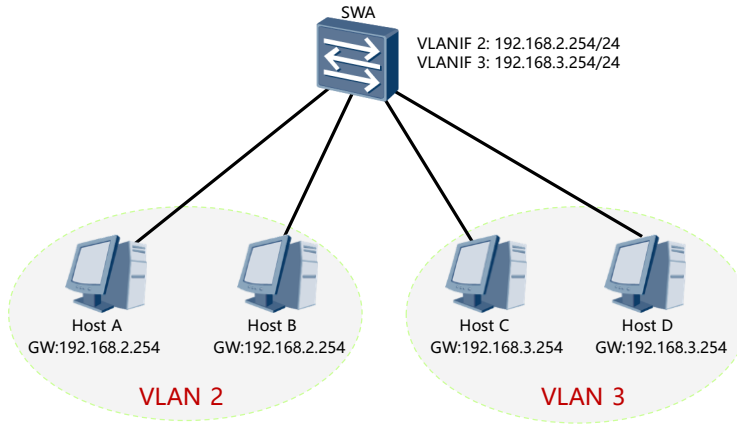


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In order to allow communication over a single trunk interface, it is necessary to logically segment the physical link using sub-interfaces. Each sub-interface represents a logical link for the forwarding of VLAN traffic before being routed by the router via other logical sub-interfaces to other VLAN destinations. Each sub-interface must be assigned an IP address in the same network segment as the VLAN that it is created for as well as 802.1Q encapsulation to allow for VLAN association as traffic is routed between VLANs.
- It is also necessary to configure the type of the Ethernet port of the switch that connects to the router as either a Trunk or Hybrid link type, and allow frames of the associated VLANs (VLAN 2 & VLAN 3 in this case) to pass.

L3 Switch Based VLAN Routing

- VLANIF are used by each VLAN as a route gateway.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The implementation of L3 switches brings about benefits to the process of VLAN routing that are not possible through the use of a router. One of those features is the ability to forward VLAN traffic with very little delay due to support of what is known as line speed forwarding as a result of bottom layer ASIC chips that allow traffic to be forwarded based on hardware rather than software. Along with this is the fact that a single device is used with no trunk link that may otherwise face congestion under heavy traffic loads. VLAN routing when using a layer 3 switch relies on the implementation of VLAN interfaces (VLANIF). If multiple users on a network belong to different VLANs, each VLAN requires a VLANIF that acts as the VLAN gateway and so must associate with an IP address relevant to the network of the VLAN. If a large number of VLANs exist however, this can tally up to a large number of IP addresses being required to support each VLANIF, as well as the hosts that are part of the VLAN with which the VLANIF is associated. Through the VLANIF, routing between different VLANs can be supported.

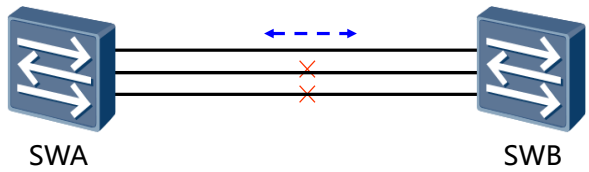
Contents

1. VLAN technology
2. **Link aggregation technology**
   - Basic concepts of link aggregation
   - Link aggregation modes
   - Application environments of link aggregation
3. OSPF protocol
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Link Aggregation Background

- If there are multiple links which are in the same broadcast domain between 2 switches, what will happen?
- If STP is enabled, will these links be in forwarding state?

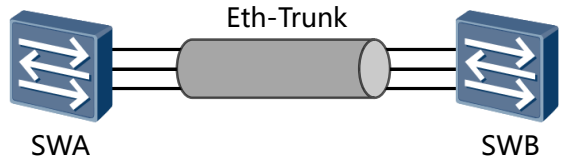


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Multiple links between 2 switches will cause the broadcast storm.
- If STP is enabled, there will be only one link in forwarding state.

Basic Concepts of Link Aggregation

- Ethernet link aggregation, also called Eth-Trunk, bundles multiple physical links to form a logical link to increase link bandwidth.
- The bundled links back up each other, increasing reliability.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- LAG and LAG interface
  - A link aggregation group (LAG) is a logical link bundled by multiple Ethernet links.
  - Each LAG corresponds to a logical interface, that is, link aggregation interface or Eth-Trunk. The Eth-Trunk can be used as a common Ethernet interface. The only difference between the Eth-Trunk and common Ethernet interface is that the Eth-Trunk needs to select one or more member interfaces to forward traffic.
- Active and inactive interfaces and links
  - There are two types of interfaces in an LAG: active interface that forwards data and inactive interface that does not forward data.
  - The link connected to an active interface is the active link, whereas the link connected to an inactive interface is the inactive link.

Advantages of Link Aggregation

- Increased bandwidth
  - The bandwidth of the link aggregation interface is the sum of bandwidth of member interfaces.
- Higher reliability
  - When an active link fails, traffic on this active link is switched to another active link, improving reliability of the link aggregation interface.
- Load balancing
  - In a link aggregation group (LAG), traffic is load balanced among active links of member interfaces.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Before deploying link aggregation, if we want networks connected by 100 Mb/s interfaces to achieve faster transmission rate, we need to replace the physical hardware(like 1000 Mb/s interfaces). But this method may cost high which doesn’ t apply to medium and small enterprise networks. We can use link aggregation to aggregate several interfaces together. In this way we can reduce the cost and fulfill the requirements of increasing interface bandwidth.
- Meanwhile, there are more than one interface members in one aggregation link. When there is one member down, it will not cause the whole link to get down, which largely increase the reliability of the network.

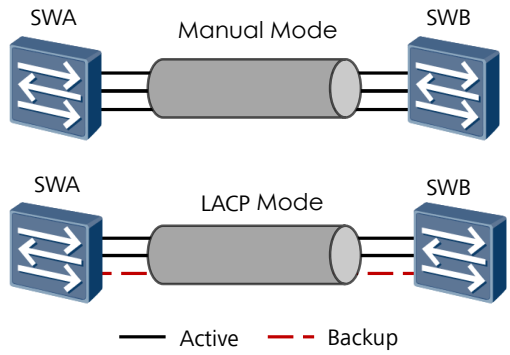
Contents

1. VLAN technology
2. **Link aggregation technology**
   - Basic concepts of link aggregation
   - Link aggregation modes
   - Application environments of link aggregation
3. OSPF protocol
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Link Aggregation Modes

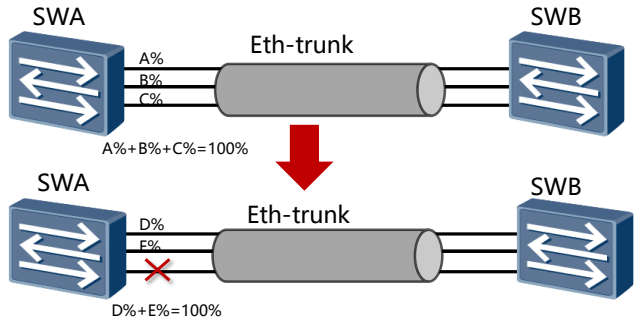
- Based on whether LACP is used or not, link aggregation can be classified into manual mode and LACP mode.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Link Aggregation in Manual Mode

- An Eth-Trunk is created between SWA and SWB. In manual mode, three active links participate in data forwarding and load balance traffic.

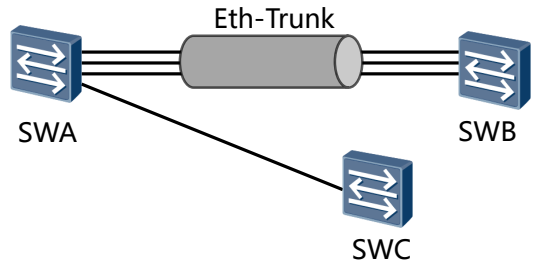


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In manual mode, you must manually create an Eth-Trunk and add member interfaces to the Eth-Trunk. In this mode, LACP is not required. The manual mode applies to the scenario where a high link bandwidth between two directly connected devices is required but the remote device does not support the LACP protocol. This mode can increase bandwidth, enhance reliability, and implement load balancing.
- As shown in the above figure, an Eth-Trunk is created between SWA and SWB. In manual mode, three active links participate in data forwarding and load balance traffic. When one link becomes faulty, the remaining two links load balance traffic.

Limitation of Manual Mode

- In manual mode, can SWA detect the fault when an member interface on SWA is incorrectly connected to an interface on SWC?

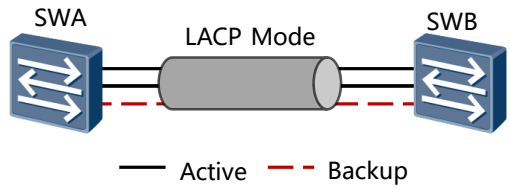


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- An Eth-Trunk in manual mode can increase the bandwidth. However, the manual mode can only detect member link disconnections, but cannot detect other faults such as link layer faults and incorrect link connections.
- For example, in the above figure, four interfaces on SWA are bundled into an Eth-Trunk and the Eth-Trunk is connected to the corresponding interfaces on SWB. Because an interface on SWA is incorrectly connected to an interface on SWC, SWA may incorrectly send data destined for SWB to SWC. However, the Eth-Trunk in manual mode cannot detect this fault in a timely manner.
- If LACP is enabled on SWA and SWB, the Eth-Trunk correctly selects active links to forward data after negotiation. Data sent by SWA can reach SWB.

Link Aggregation in LACP Mode

- The Link Aggregation Control Protocol (LACP) can improve fault tolerance of the Eth-Trunk, provide backup, and ensure high reliability of member links.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- LACP uses a standard negotiation mechanism for a switching device so that the switching device can create and start the aggregated link based on its configuration. After the aggregated link is created, LACP maintains the link status. If an aggregated link's status changes, LACP adjusts or removes the link.
- In LACP mode, LACP is used to negotiate parameters to determine active links in an LAG. This mode is also called the M:N mode, where M refers to the number of active links and N refers to the number of backup links. This mode guarantees high reliability and allows traffic to be load balanced among M active links.
  - As shown in the above figure, M+N links with the same attributes (in the same LAG) are set up between two devices. When data is transmitted over the aggregated link, traffic is load balanced among M active links and no data is transmitted over N backup links. Therefore, the actual bandwidth of the aggregated link is the sum of the M links' bandwidth, and the maximum bandwidth of the aggregated link is the sum of the M+N links' bandwidth.
  - If one of M links fails, LACP selects a link from N backup links to replace the faulty link. The actual bandwidth of the aggregated link is still the sum of M links' bandwidth, but the maximum bandwidth of the aggregated link is the sum of the (M+N-1) links' bandwidth.

Implementation of LACP Mode

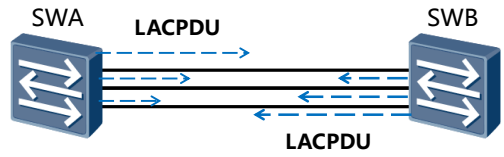
- LACP, as specified in IEEE 802.3ad, implements dynamic link aggregation and de-aggregation.
- LACP allows both ends to exchange Link Aggregation Control Protocol Data Units (LACPDUs).

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- After member interfaces are added to an Eth-Trunk in LACP mode, each end sends LACPDUs to inform its remote end of its system priority, MAC address, member interface priorities, interface numbers, and keys. The remote end then compares this information with that saved on itself, and selects which interfaces to be aggregated. The two ends perform LACP negotiation to select active interfaces and links.
- Active and inactive interfaces and links
  - There are two types of interfaces in an LAG: active interface that forwards data and inactive interface that does not forward data.
  - The link connected to an active interface is the active link, whereas the link connected to an inactive interface is the inactive link.
  - Only when active interfaces fail, can inactive interfaces become new active interfaces.
- LACP system priority
  - LACP system priorities are set on devices at both ends of an Eth-Trunk. In LACP mode, active member interfaces selected by both devices must be consistent; otherwise, an LAG cannot be established. To keep active member interfaces consistent at both ends, set a higher priority for one end so that the other end selects active member interfaces based on the selection of the end with a higher priority. The smaller the LACP system priority value, the higher the LACP system priority.

Eth-Trunk Setting Up Process in LACP Mode (1)

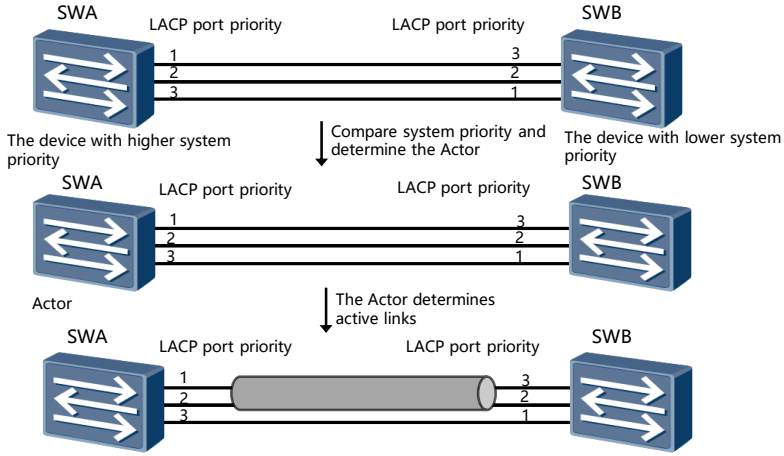
- Devices at both ends send LACPDUs to each other.
  - As shown in the figure, you need to create an Eth-Trunk in LACP mode on SWA and SWB and add member interfaces to the Eth-Trunk. Then the member interfaces are enabled with LACP, and devices at both ends can send LACPDUs to each other.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Eth-Trunk Setting Up Process in LACP Mode (2)

- Devices at both ends determine the Actor and active links.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As shown in the figure, devices at both ends receive LACPDUs from each other. For example, when SWB receives LACPDUs from SWA, SWB checks and records information about SWA and compares system priorities. If the system priority of SWA is higher than that of SWB, SWA acts as the Actor. If SWA and SWB have the same system priority, the device with a smaller MAC address functions as the Actor.
- After devices at both ends select the Actor, they select active interfaces according to the priorities of the Actor's interfaces. Then active interfaces are selected, active links in the LAG are specified, and load balancing is implemented among these active links.

Comparisons Between Link Aggregation Modes

Item	Manual Mode	LACP Mode
Definition	You must manually create an Eth-Trunk and add member interfaces to the Eth-Trunk.	An Eth-Trunk is created based on LACP. LACP provides a standard negotiation mechanism for a switching device.
Data forwarding	All links are active links. All active links participate in data forwarding.	Some links are active links. All active links participate in data forwarding. If an active link fails, the system selects a link among inactive links as the active link.
Fault detection	This mode can only detect member link disconnections.	This mode can detect member link disconnections and other faults such as link layer faults and incorrect link connections.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Configuration Notes Before an Eth-Trunk Is Configured

- Member interfaces cannot be configured with some services or static MAC address entries.
- An Eth-Trunk cannot be added to another Eth-Trunk.
- Member interfaces of an Eth-Trunk must use the same Ethernet type and rate.
- Both devices of the Eth-Trunk must use the same number of physical interfaces, interface rate, duplex mode, and flow control mode.
- Both devices of an Eth-Trunk must use the same link aggregation mode.
- When the number of active interfaces falls below the lower threshold, the Eth-Trunk goes Down.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

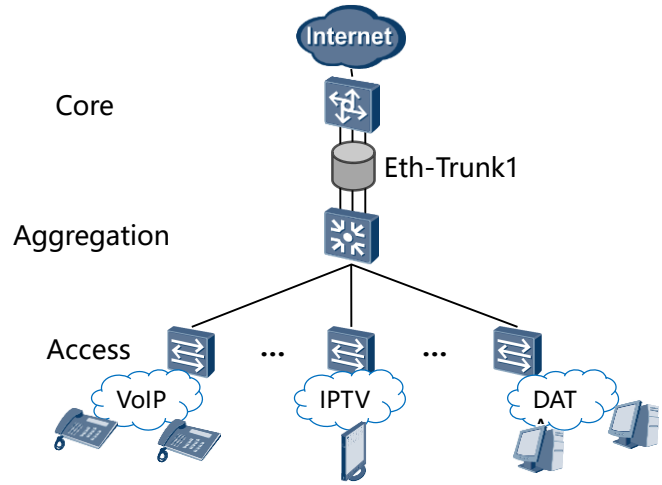
- Member interfaces cannot be configured with some services or static MAC address entries. For example, when an interface is added to an Eth-Trunk, the interface must use the default link type.
- Interfaces that use different Ethernet types and rates cannot join the same Eth-Trunk. For example, GE and FE interfaces cannot join the same Eth-Trunk, and GE electrical and optical interfaces can join the same Eth-Trunk.
- If an interface of the local device is added to an Eth-Trunk, an interface of the remote device directly connected to the interface of the local device must also be added to an Eth-Trunk so that the two ends can communicate.

Contents

1. VLAN technology
2. **Link aggregation technology**
   - Basic concepts of link aggregation
   - Link aggregation modes
     - Application environments of link aggregation
3. OSPF protocol
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Switches Are Directly Connected Using Link Aggregation

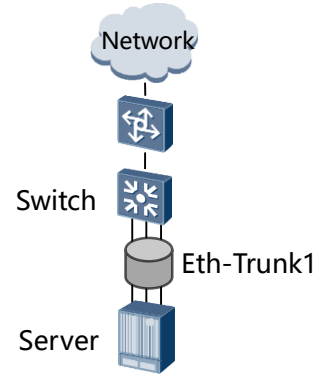


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As shown in the above figure, traffic of services with different priorities is sent to the Internet through the access, aggregation and core layer devices. To ensure the bandwidth and reliability of the link between the aggregation and core layer devices, an LAG, Eth-Trunk 1, is established.
- You can determine the working mode for the Eth-Trunk according to the following situations:
  - If devices at both ends of the Eth-Trunk support LACP, the LACP mode is recommended.
  - If the device at either end of the Eth-Trunk does not support LACP, you must use the manual mode.
- QoS can be implemented on an Eth-Trunk as a common interface. At both ends of Eth-Trunk 1, traffic shaping, congestion management, and congestion avoidance can be performed for outgoing traffic, ensuring that packets of high priorities are sent in a timely manner.

A Switch Connects to a Server Using Link Aggregation

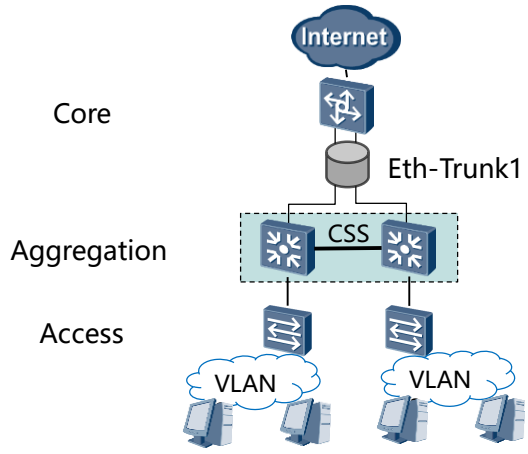
- To improve the server bandwidth and reliability, two or more network adapters of the server are aggregated to form a network adapter group to implement load balancing or redundancy.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Network adapters of the server must use the same type.
- The link aggregation modes on the server and access device must be consistent.Intel network adapter is used as an example. A server often uses static or IEEE 802.3ad dynamic link aggregation. When the server uses static link aggregation, the access device must use the manual mode. When the server uses IEEE 802.3ad dynamic link aggregation, the access device must use the LACP mode.

A Switch Connects to a Stack Using Link Aggregation



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As shown in the above figure, the switch connects to a stack using link aggregation, and the Eth-Trunk is enabled to preferentially forward local traffic. Preferentially forwarding local traffic ensures reliable transmission, reduces the bandwidth burden between CSS devices, and improves the forwarding efficiency.

Contents

1. VLAN technology
2. Link aggregation technology
3. **OSPF protocol**
   - OSPF protocol overview
     - Basic OSPF concepts
     - OSPF neighbor and adjacency relationship
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

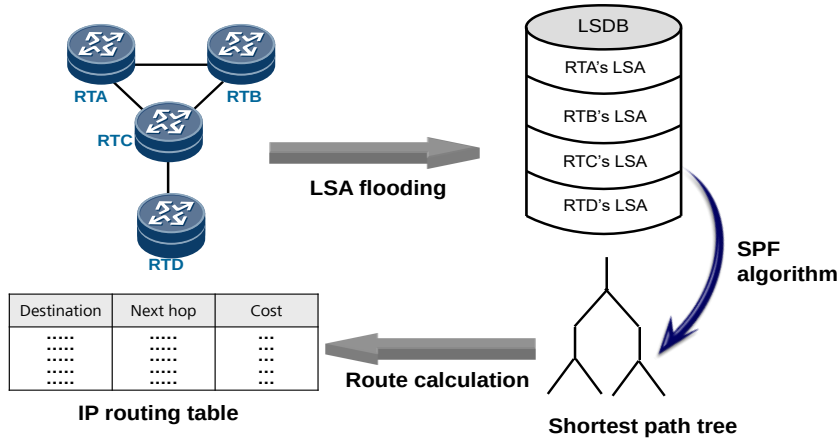
Basic Characteristics of OSPF

- Supporting Classless Inter-Domain Routing (CIDR)
- Free of routing loops
- Fast convergence
- Sending and receiving protocol data by using IP multicast
- Supporting equal-cost routes
- Supporting protocol packet authentication

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Supporting Classless Inter-Domain Routing (CIDR)
- As a routing protocol specially developed for TCP/IP environments, OSPF explicitly supports CIDR and Variable-Length Subnet Mask (VLSM).
- Free of routing loops
- OSPF calculates routes based on detailed link state information, namely, network topology information, to generate a shortest path tree (SPT) rooted on the local router. Therefore, the routes calculated by OSPF are loop-free.
- Fast convergence
- OSPF is based on triggered updates. Once a topological change occurs, the new link state information is flooded. Therefore, OSPF is sensitive to topological changes and converges quickly.
- Sending and receiving protocol data by using IP multicast
- OSPF routers send and receive protocol data by using multicast and unicast. Therefore, the network traffic generated is very low.
- Supporting equal-cost routes
- When multiple equal-cost paths exist to the same destination address, the traffic is evenly distributed on these equal-cost paths.
- Supporting protocol packet authentication
- All the packets exchanged between OSPF routers are authenticated. This ensures network security at the protocol level.

Route Calculation Using the Link State Algorithm



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OSPF route calculation is based on the link state algorithm. Before studying the OSPF routing protocol, you need to understand the process of route calculation using the link state algorithm.
- 1. Adjacency establishment
- OSPF adjacencies are established among neighboring OSPF routers. The routers exchange LSAs only after adjacencies are established among them.
- 2. LSDB synchronization
- After adjacencies are established, each router sends LSAs to its neighbors, receives LSAs from its neighbors, and sends the LSAs received from other routers to its neighbors. Each router stores received LSAs. All the LSAs build an LSDB.
- 3. Route calculation
- After LSDB synchronization, each OSPF router, with itself as the root, runs the SPF algorithm to generate an SPT rooted on the router.
- 4. Routing table generation
- According to the SPF tree, each router calculates routing information respectively and adds the routing information to its routing table.

Contents

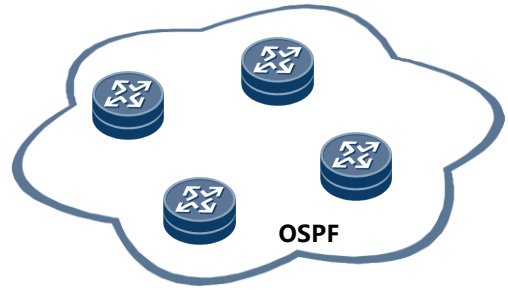
1. VLAN technology
2. Link aggregation technology
3. **OSPF protocol**
   - OSPF protocol overview
   - Basic OSPF concepts
   - OSPF neighbor and adjacency relationship
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

AS (Autonomous System)

- In OSPF, an AS refers to a group of routers that exchange routing information by running the same routing protocol.



- In this example, all the routers run OSPF and belong to the same AS.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- An IGP runs within a certain scope. And this scope is named as AS (Autonomous System). As an IGP, OSPF runs within ASs.
- Autonomous System (AS)is a set of routers that exchange routing information using the same routing protocol.

Router ID

- A router ID is a 32-bit integer that uniquely identifies an OSPF router in an AS. The format of a router ID is the same as that of an IP address.



- In this example, run the router id 1.1.1.1 command to configure the router ID of the router as 1.1.1.1. After the router ID is configured, run the display router id command. The command output information shows that the router ID of the router is 1.1.1.1.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

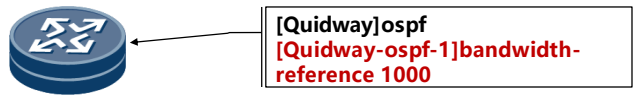
- During OSPF route calculation, each OSPF router needs to save the link state information about all the routers on the network. To distinguish the link state information about different routers in an LSDB, each router on the network is uniquely identified by a route ID in the LSDB.
- A router ID can be configured manually. If no router ID is specified by using a command, the system automatically selects one of the existing interface IP addresses as the router ID.
- The principle of selecting a router ID is as follows:
- The highest IP address among the loopback addresses is preferentially selected as the router ID.
- If no loopback interface is configured, the highest IP address among the physical interface addresses is selected as the router ID.
- A new router ID is selected only after the interface IP address selected as the router ID is deleted.
- It is recommended to use loopback 0 address as the router ID of a router.

Cost

- An OSPF cost is a 16-bit unsigned integer ranging from 1 to 65535.
  - By default, the cost of an interface is calculated as 108/BW(bps)
  - The cost of an OSPF interface can be configured manually



  - Bandwidth reference values are configurable. When configuring bandwidth reference values, ensure that the bandwidth reference values of all the routers running the OSPF process are the same.

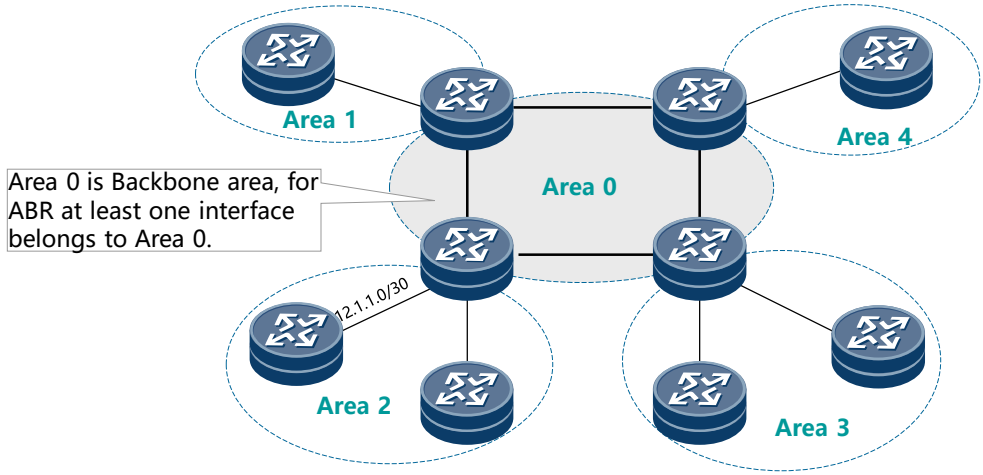


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- If no cost is directly configured for an OSPF interface, OSPF automatically calculates the cost of the interface according to the bandwidth of the interface. The formula for calculating an interface cost is: interface cost = bandwidth reference value/interface bandwidth. The integer part of the result is used as the interface cost (if the result is less than 1, the interface cost is 1). You can change the interface cost by changing the bandwidth reference value.

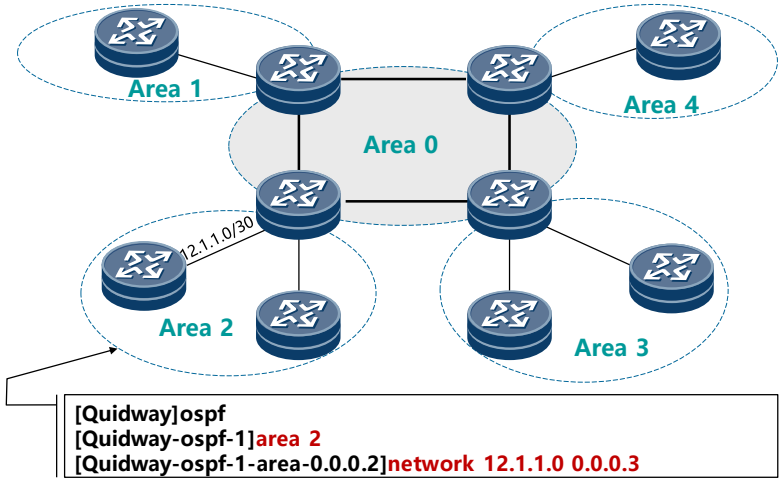
OSPF Area (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- If OSPF is enabled on all the routers on a large network and the number of routers on the network keeps increasing, the following problems occur:
- 1.All the routers generate LSAs respectively and the LSDBs become very large. Therefore, LSDB synchronization takes long and occupies much memory space.
- 2.The running of the SPF algorithm is more complicated and occupies more CPU resources.
- 3.When the network size grows, the probability of topological changes also increases. As a result, a large number of OSPF packets are transferred on the network. This lowers the bandwidth utilization of the network.
- 4.Moreover, each change causes route recalculations on all the routers on the network.
- OSPF solves these problems by dividing an AS into areas.
- An area is a logical group of routers and is identified by an area ID. A network segment (link) belongs to only one area. In other words, the area to which an OSPF interface belongs must be specified.

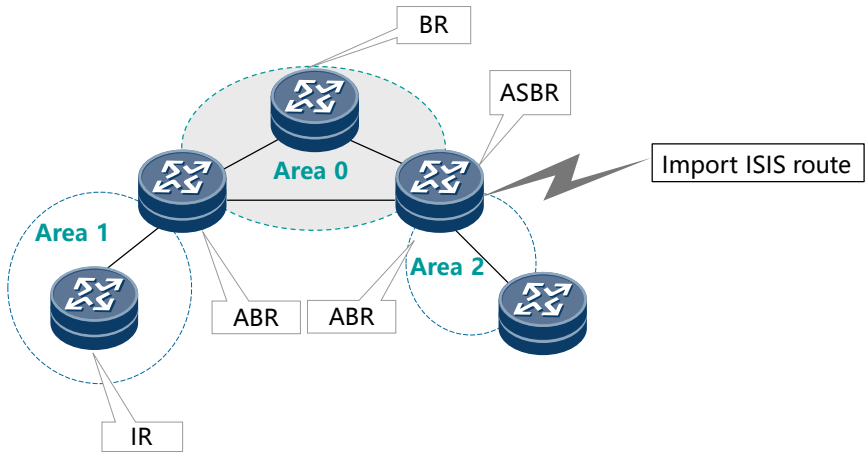
OSPF Area (2)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- 1.Area 0 is the backbone area. The backbone area is responsible for advertising the routing information (not detailed link state information) summarized by area boarder routers (ABRs) between non-backbone areas.
- 2.To prevent inter-area routing loops, OSPF disallows direct inter-area routing information advertisement between non-backbone areas. Therefore, an ABR must have at least one interface to the area 0. That is, each non-backbone area must be connected to the backbone area.
- 3.Each area has an LSDB unique to the area. A router maintains a separate LSDB for each area to which the router is connected. Detailed link state information is not advertised outside any area. Therefore, LSDB sizes are greatly reduced.

OSPF Router Roles



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Internal router (IR)
- An IR is a router with all directly connected networks belonging to the same area. IRs that belong to the same area maintain the same LSDB.
- Area border router (ABR)
- An ABR is a router directly connected to multiple areas. An ABR maintains an LSDB for each area to which the ABR is directly connected.
- Backbone router
- A backbone router is a router that has at least one interface (or virtual link) to the backbone area. Backbone routers include all the ABRs and the routers with all their interfaces directly connected to the backbone area.
- AS boundary router (ASBR)
- An ASBR is a router that exchanges routing information with routers belonging to other ASs. An ASBR advertises AS-external routes throughout the entire AS.
- An ASBR can be an IR or ABR. An ASBR can belong to or does not belong to the backbone area.

Contents

1. VLAN technology
2. Link aggregation technology
3. **OSPF protocol**
   - OSPF protocol overview
   - Basic OSPF concepts
     - OSPF neighbor and adjacency relationship
4. BGP protocol

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

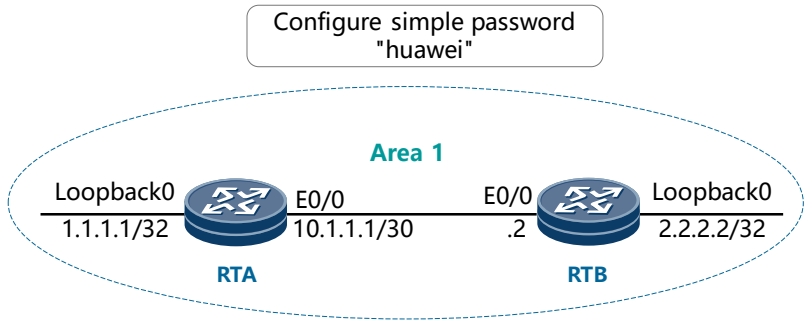
OSPF Packet Types

Type	Packet name	Packet function
1	Hello	Discover /maintain neighbor
2	Database Description	Summarize database contents
3	Link State Request	Database download
4	Link State Update	Database update
5	Link State Ack	Flooding acknowledgement

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- There are five types of OSPF packets. By exchanging protocol packets, OSPF routers establish neighbor relationships among them and exchange link state information to complete route calculation. This section describes the functions of OSPF packets.
- Hello packets are used to discover neighbors and maintain neighbor relationships.
- Database description (DD) packets summarize link states by carrying LSA header information.
- Link state (LS) request packets are used to request the LSAs that are discovered by receiving DD packets but not available on the local router.
- Detailed LSAs are sent in LS Update packets to synchronize LSDBs.
- LS Ack packets are flooded to guarantee reliable routing information exchange.

OSPF Packet Header Authentication-Interface Authentication



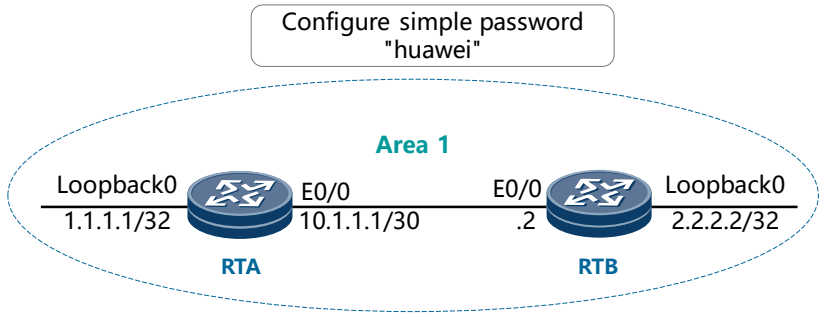

```
[RTA-Ethernet0/0]ospf authentication-mode simple plain huawei
```

```
[RTB-Ethernet0/0]ospf authentication-mode simple plain huawei
```

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OSPF supports packet header authentication to guarantee OSPF security. OSPF packets are accepted only after being authenticated. If the OSPF packet authentication fails, no neighbor relationship can be established.
- The versatile routing platform (VRP) supports two authentication modes: area authentication and interface authentication. When area authentication is used, the authentication modes and passwords of all the routers belonging to the same area must be the same in the area. For example, for all the routers belonging to area 0, the authentication mode is configured as simple authentication and the password as abc. If both area authentication and interface authentication is configured, interface authentication is used preferentially.

OSPF Packet Header Authentication-Area Authentication

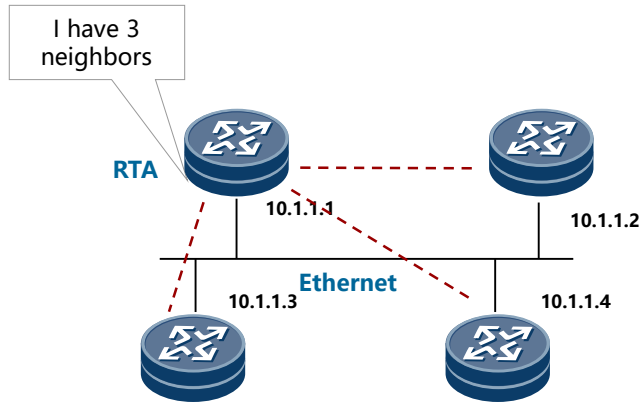



```
[RTA-ospf-1-area-0.0.0.1]authentication-mode simple plain huawei
```

```
[RTB-ospf-1-area-0.0.0.1]authentication-mode simple plain huawei
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

What Is Neighbor? What Is Adjacency?

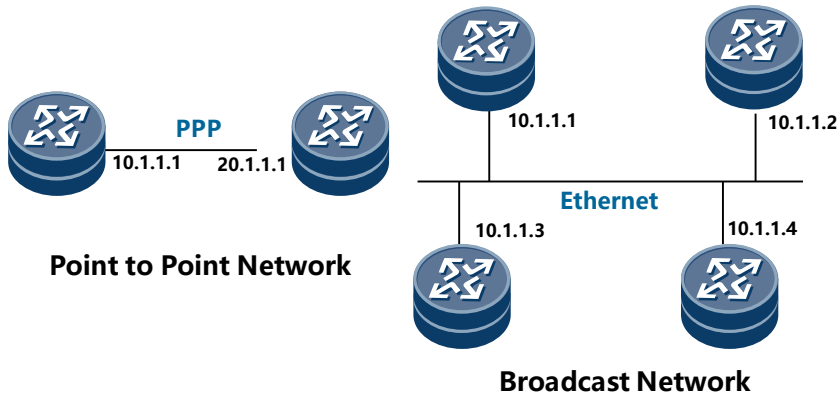


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OSPF is a dynamic routing protocol. Before two OSPF routers can exchange link state information and routing information, a neighbor relationship and adjacency need to be established between the routers.
- Neighbor
- Two routers directly connected to the same network segment are neighbors. Neighbor relationships are maintained by using OSPF Hello packets.
- Adjacency
- An adjacency is a neighbor relationship selected to exchange routing information.

- Not all the neighbor relationships can become adjacencies. Whether an adjacency is established also varies with network types.
- In this example, RTA and the other three routers are directly connected to the same network segment. As shown in the preceding figure, OSPF runs on all the interfaces of all the routers. According to the above-mentioned definitions, RTA establishes neighbor relationships with the other three routers. According to the principle of running OSPF on Ethernet, RTA establishes adjacencies with only the designated router (DR) and backup designated router (BDR).
- The following topic describes the types of networks on which OSPF runs and the principles of establishing neighbor relationships and adjacencies on different types of networks.

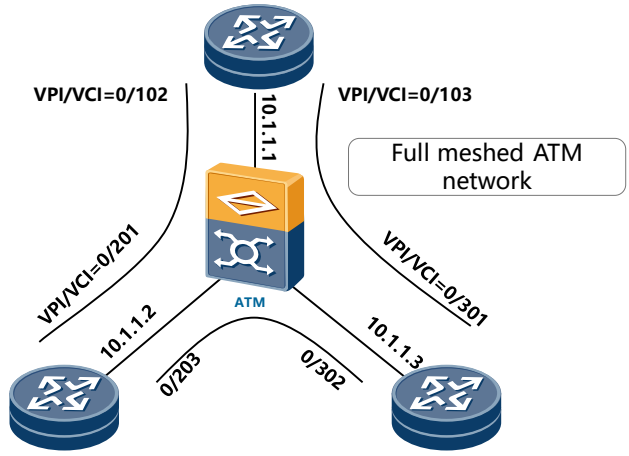
What Are the OSPF Network Types? - P2P and Broadcast



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- As mentioned in the preceding section, not all the neighbor relationships become adjacencies to allow the exchange of link state information and routing information. Whether an adjacency is formed depends on the network type. A network type refers to the Layer 2 link type of the network segment running OSPF.
- In OSPF, four network types are defined: point-to-point (P2P), broadcast, non-broadcast multi-access (NBMA), and point-to-multipoint (P2MP).
- P2P: A P2P network is a network where two routers are directly interconnected.
- Broadcast: A broadcast network is a network that supports the interconnection of more than two routers and has broadcast capabilities.
- NBMA: An NBMA network is a network that supports the interconnection of more than two routers but does not have any broadcast capability. On an NBMA network, OSPF simulates the operations performed on a broadcast network, but the neighbors of each router need to be configured manually. All the routers on an NBMA network must be fully-meshed.
- P2MP: An entire non-broadcast network is considered as a group of P2P networks. The neighbors of each router can be discovered by using a lower-layer protocol, for example, inverse address resolution protocol (ARP).

What Are the OSPF Network Types? - NBMA



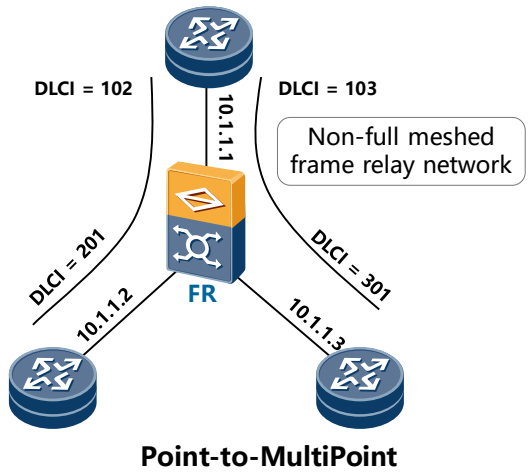
Non-Broadcast Multiple Access (NBMA)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- NBMA: An NBMA network is a network that supports the interconnection of more than two routers but does not have any broadcast capability. On an NBMA network, OSPF simulates the operations performed on a broadcast network, but the neighbors of each router need to be configured manually. All the routers on an NBMA network must be fully-meshed.

What Are the OSPF Network Types? - P2MP



Point-to-MultiPoint

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- P2MP: An entire non-broadcast network is considered as a group of P2P networks. The neighbors of each router can be discovered by using a lower-layer protocol, for example, inverse address resolution protocol (ARP).

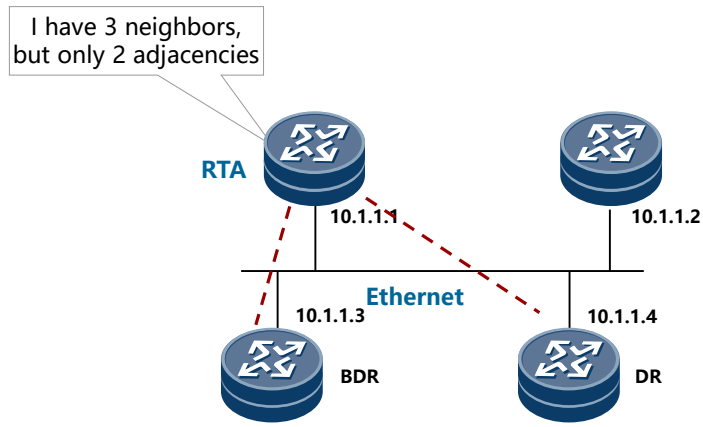
Default Network Types Corresponding to Common Link Layer Protocols

Network type	Common data link layer protocols
Point-to-point	PPP, LAPE, HDLC
Broadcast	Ethernet
NBMA	Frame Relay, ATM

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The table list default network types for common data link layer protocols.
- Point-to-MultiPoint is not a default network type.

DR and BDR

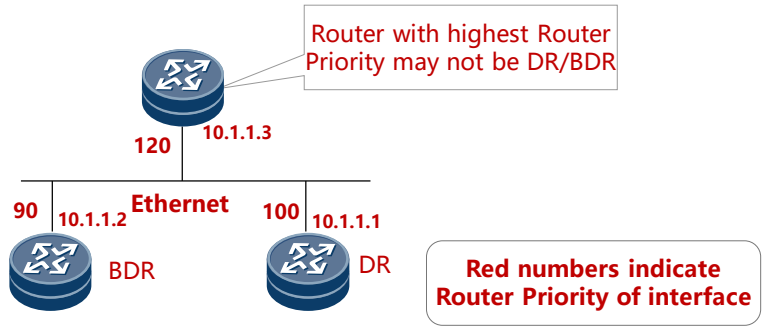


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In a broadcast network or NBMA network, routing information needs to be transferred between any two routers. If n routers exist in the network, n ? (n-1)/2 adjacencies need to be established. As a result, a route change on any router needs to be transferred for multiple times and bandwidth resources are wasted. To solve this problem, DR is defined in the OSPF protocol and all the routers only need to send information to the DR for broadcasting the network link states.
- If the DR fails due to a fault, all the routers in the network must re-elect the DR and be synchronized to the new DR. During this process, which takes quite long, route calculation may be incorrect. To shorten this process, the BDR concept is defined in OSPF.
- DR: A DR is the router that maintains adjacencies with all the other OSPF routers on the same network segment and exchanges LSAs with these routers.
- BDR: A BDR is a backup DR.
- DR Other: A router that is neither a DR nor BDR is a DR Other. DR Others do not form adjacencies between themselves or exchange routing information. Therefore, the number of adjacencies formed between the routers on the broadcast network or NBMA network is reduced.

DR Election and BDR Election

- Instead of being manually specified, the DR and BDR are elected among all the routers on the local network segment. The DR priority of a router interface determines the eligibility of the interface in the DR election and BDR election.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Process Description
- The DR and BDR are elected by the Hello protocol. The election procedure is as follows:
- Each router writes the DR it votes for into a Hello packet advertised to other routers on the same network segment.
- When two routers on the same network segment declare themselves the DR, the router with a higher DR priority wins.
- If the DR priorities are the same, the router with a larger router ID wins.
- A router with the priority 0 is not elected as DR or BDR.
- Note the following points:
- DR is elected only on broadcast or NBMA interfaces. No DR is elected on P2P or P2MP interfaces.
- DR is based on the network segment and relative to a router interface. A router that functions as the DR on an interface may be a BDR or DR Other on another interface.
- If the DR and BDR are elected, a newly added router, regardless of its DR priority, does not become the DR of the network segment immediately.
- The DR is not necessarily the router with the highest DR priority. Likewise, the BDR is not necessarily the router with the second highest DR priority.
- On the Ethernet shown in the preceding figure, the DR is 10.1.1.1 and the BDR is 10.1.1.2. If a router is added to the network, configure the priority of the added router as 120, which is greater than the priority of the original DR, 100, and the priority of the original BDR, 90. The added router does not become the new DR though it has the highest priority. This maintains the network stability.

DR Election and BDR Election

- DR election and BDR election takes time and affects the speed of OSPF route convergence. During the actual OSPF application, the broadcast network and NBMA network types are often changed to the P2P type to prevent the election of the DR or BDR. The following command is used to change the network type of an OSPF network interface:

```
[RTB-Ethernet0/0] ospf network-type p2p
```

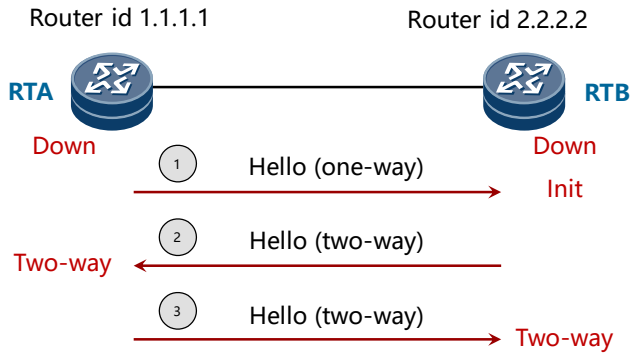
- The following figure describes whether adjacencies are established with neighbors for different network types:

Network Type	Establish Adjacency with Neighbor or not
Point-to-Point	Neighbors always become adjacent
Point-to-MultiPoint	Neighbors always become adjacent
Broadcast NBMA	DR is always adjacent to all the other routers including BDR; BDR is always adjacent to all the other routers including DR; Routers whose interface state is DR, Other is adjacent to only DR and BDR.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Establishment of Neighbor Relationship (1)

- Neighbor relationship establishment on a broadcast network:



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In OSPF, whether neighbor relationships are established depends on network types. This topic describes the procedure of establishing neighbor relationships on P2P networks and broadcast networks, which are the most commonly used.
- When the OSPF state becomes “Two-way” on a broadcast network, it means this Router’s neighbor relationship is established.
- Process Description
- 1.Neighbor relationship establishment on a broadcast network.
- 2.Neighbor relationship establishment on a P2P network.

Establishment of Neighbor Relationship (2)

- The difference between the 1-way and 2-way states is as follows: If a Hello packet carrying the router's router ID has recently been seen from the neighbor, it indicates the 2-way state. Otherwise, it indicates the 1-way state.

```
[Quidway]display ospf peer

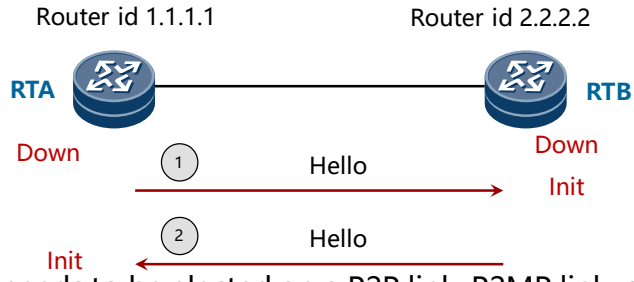
            OSPF Process 1 with Router ID 4.4.4.4
                    Neighbors

 Area 0.0.0.1 interface 10.1.1.4(Ethernet0/0)'s neighbor(s)
 RouterID: 1.1.1.1       Address: 10.1.1.1
     State: 2 Way  Mode: None    Priority: 1
     DR: 10.1.1.2  BDR: 10.1.1.3
     Dead timer expires in 37s
     Neighbor has been up for 00:00:00
```

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Establishment of Neighbor Relationship (3)

- Neighbor relationship establishment on a P2P network:



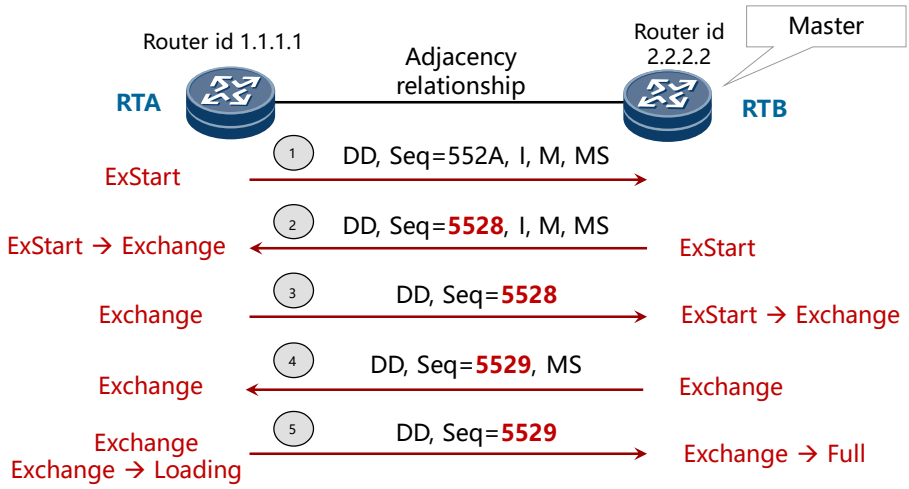
- No DR or BDR needs to be elected on a P2P link, P2MP link, or virtual link.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- No DR or BDR needs to be elected on a P2P link, P2MP link, or virtual link.
- After the interface goes up, the interface states from DOWN state to point-to-point state and attempts to establish a neighbor relationship with its neighbor.
- After the interface receives a Hello packet, the router enters the Init state. This process is different from that on a broadcast link or NBMA link.

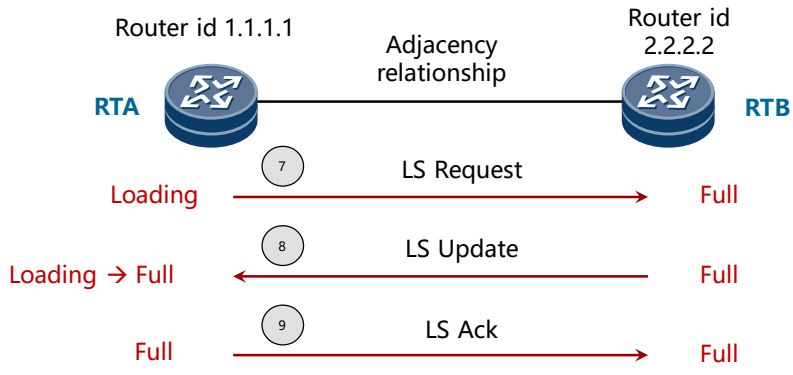
Adjacency Establishment Process (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Process Description:
- 1.After the neighbor state becomes ExStart, RTA sends RTB the first DD packet in which the DD sequence number is set to 552A (assumed). The Init bit is set to 1, indicating that this packet is the first in the sequence of DD packets. The More bit is set to 1, indicating that more DD packets are to follow. The Master/Slave bit is set to 1, indicating that the router is the master during the database exchange process.
- 2.After the neighbor state becomes ExStart, RTB sends RTA the first DD packet in which the DD sequence number is set to 5528 (assumed). As the router ID of RTB is larger than that of RTA, RTB should function as the master. After the router ID comparison is complete, RTA generates a NegotiationDone event. Therefore, the RTA neighbor state transitions from ExStart to Exchange.
- 3.After the neighbor state becomes Exchange, RTA sends a new DD packet carrying the LSDB summary information. The DD sequence number is set to that used by RTB in step 2. The More bit is set to 0, indicating that no more DD packet is needed to describe the LSDB. The Master/Slave bit is set t 0, indicating that RTA asserts itself as the slave. On receiving the DD packet, a NegotiationDone event is generated on RTB. Therefore, the state of RTB changes to Exchange.
- 4.After the neighbor state changes to Exchange, RTB sends a new DD packet that carries LSDB description information, with the DD sequence number set to 5529 (the previously used DD sequence number increase 1).

Adjacency Establishment Process (2)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- 5.RTA does not need any new DD packet to describe its LSDB. Functioning as the slave, however, RTA needs to acknowledge each DD packet sent by RTB, which is the master. Therefore, RTA sends RTB a new and empty DD packet whose sequence number is 5529.
- 6.After sending the last DD packet, RTA generates an ExchangeDone event to change the neighbor state to Loading. After RTB receives the last DD packet, if the LSDB of RTB is the most recent and RTB does not need to send any update request to RTA, RTB transits to the Full state.
- 7.After the neighbor state becomes Loading, RTA starts sending LSRs to RTB, asking for the link state information that is discovered by DD packets in the Exchange state but is not found in the local LSDB.
- 8.After receiving the LSR, RTB sends an LSU carrying the detailed information about the requested link state to RTA. On receiving the LSU, RTA changes the neighbor state from Loading to Full.
- 9.RTA sends LS Ack packets to RTB to ensure reliable information transmission. LS Ack packets are used to flood the acknowledgments of received LSAs.
- 10.The neighbor state becomes Full, indicating that an adjacency is fully established.

Check Neighbor State of OSPF Router

```
[RTA]display ospf peer

        OSPF Process 1 with Router ID 1.1.1.1
                Neighbors
Area 0.0.0.0 interface 10.1.1.1(Ethernet0/0)'s neighbor(s)
RouterID: 2.2.2.2       Address: 10.1.1.2
  State: Full  Mode: Nbr is Master  Priority: 1
  DR: 10.1.1.1  BDR: 10.1.1.2
  Dead timer expires in 35s
  Neighbor has been up for 04:35:02
```


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

LSDB Synchronization

- Synchronized LSDBs are the basis for proper OSPF route calculation. When an adjacency is established, an OSPF router completes the LSDB synchronization with the adjacent router.
- After an adjacency is established and LSDB synchronization is completed, LSAs are updated in the following ways:
  - Periodic update: Every 30 minutes, an OSPF router floods the LSAs in the LSDB of the current area to the adjacent routers in the corresponding areas.
  - Triggered update: When a network topology change occurs, the router generates new LSAs and floods them so that the topology information about the network remains correct and consistent.

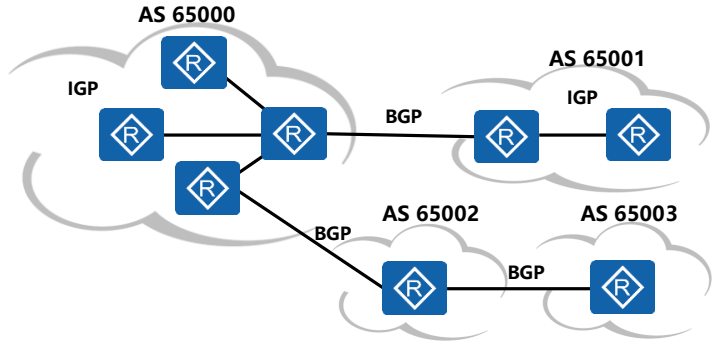
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. VLAN technology
2. Link aggregation technology
3. OSPF protocol
4. **BGP protocol**
   - BGP Overview
     - BGP Neighbor Relationship Establishment and Configuration
     - BGP Route Advertisement Rules and Route Processing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic BGP Functions

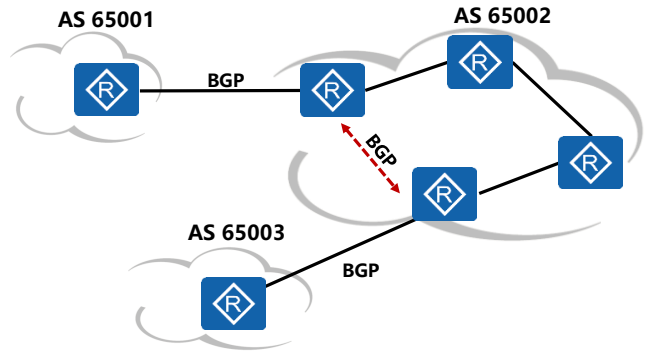


- IGPs, such as OSPF, IS-IS, and RIP, are used within an AS to calculate and discover routes.
- BGP is used between ASs to transmit and control routes.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- EGP, the predecessor of BGP, is simple in design and can only transmit routing information between ASs and cannot select optimal routes or prevent routing loops between ASs. Therefore, EGP was replaced by BGP.
- Compared with EGP, BGP has the following advantages:
  - Discovers neighbors and establishes neighbor relationships.
  - Selects optimal routes and advertises routes.
  - Prevents routing loops, efficiently transmits routes, and maintains a large amount of routing information.
  - Provides various route control capabilities between ASs that are not fully trusted.
- Using BGP to transmit routing information, a routing domain functions as a whole to exchange routing information with another routing domain. This routing domain is an AS. An AS is a set of routers and networks that consist of these routers. These routers are under a single technical administration and use the same routing policies.
- An AS is uniquely identified by an AS number, which is assigned by the Internet Assigned Numbers Authority (IANA). Before January 2009, only 2-byte AS numbers can be used, which range from 1 to 65535. AS numbers 1 to 64511 are public AS numbers, and AS numbers 64512 to 65534 are private AS numbers. After January 2009, the IANA decided to use 4-byte AS numbers, which range from 65536 to 4294967295.

BGP Characteristics



- In the figure, two BGP routers can establish a neighbor relationship across multiple routers.
- To implement on-demand route control and selection, various BGP attributes are designed and carried in routes.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

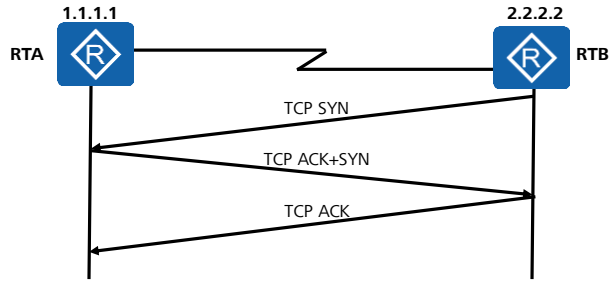
- To ensure reliable data transmission between ASs, BGP uses TCP to establish connections. Therefore, BGP can establish a neighbor relationship across multiple routers, while IGP can only establish a neighbor relationship hop by hop.
- Routers between ASs do not completely trust each other. To implement on-demand route control and selection, various BGP attributes are designed.

Contents

1. VLAN technology
2. Link aggregation technology
3. OSPF protocol
4. **BGP protocol**
   - BGP Overview
   - BGP Neighbor Relationship Establishment and Configuration
   - BGP Route Advertisement Rules and Route Processing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

BGP Neighbor Discovery

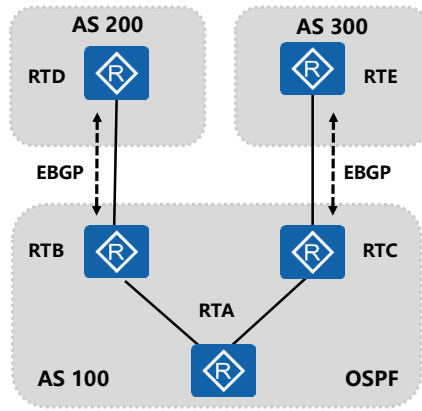


- The device that starts BGP first initiates a TCP connection. In the figure, RTB first starts BGP and uses a random port number to initiate a TCP connection with port 179 of RTA.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- BGP is designed to run between ASs to transmit routers. There are WAN links between ASs, and unpredictable link congestion or packet loss may occur during packet transmission on WANs. Therefore, BGP uses TCP as the transport protocol to ensure reliability.
- BGP uses TCP port 179 to establish neighbor relationships, and TCP establishes connections in unicast mode. Therefore, unlike RIP and OSPF, BGP does not discover neighbors in multicast mode. Establishing connections in unicast mode requires neighbors to be manually specified in BGP.

BGP Neighbor Type - EBGP

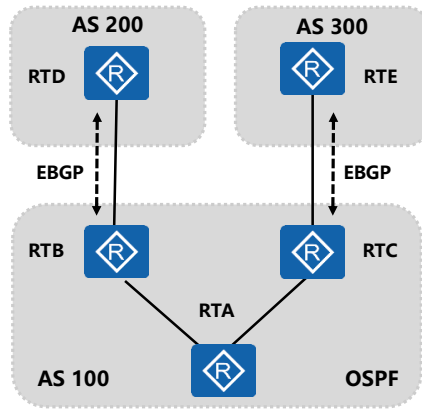


- BGP routers in different ASs establish EBGP neighbor relationships.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- EBGP transmits routes only between different ASs. In the figure, RTB and RTC in AS 100 can learn different routes from AS 200 and AS 300 respectively. How to transmit routes of AS 200 and AS 300 within AS 100?
- To meet this requirement, on RTB and RTC, import BGP routes into IGP (OSPF in the figure) and then import IGP routes back into BGP.
- However, this method has the following disadvantages:
  - There are a huge number of BGP routes on the public network. After these BGP routes are imported into IGP, IGP cannot support these BGP routes.
  - When BGP routes are imported into IGP, strict control is required. This complicates the configuration and maintenance.
  - When BGP attributes carried in BGP routes are imported into IGP, these attributes may be lost because they cannot be identified by IGP.
- To overcome these disadvantages, BGP needs to be designed to transmit routes within an AS.

BGP Neighbor Type - IBGP

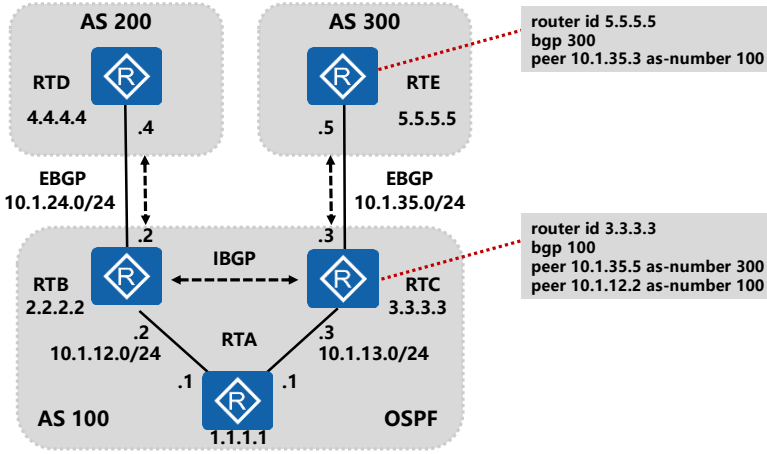


- BGP routers in the same AS establish IBGP neighbor relationships.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- BGP uses TCP as the transport protocol. Therefore, BGP can establish neighbor relationships across multiple devices. In the figure, RTB and RTC establish an IBGP neighbor relationship and transmit the routes learned from other ASs to each other so that BGP routes can be transmitted within an AS.

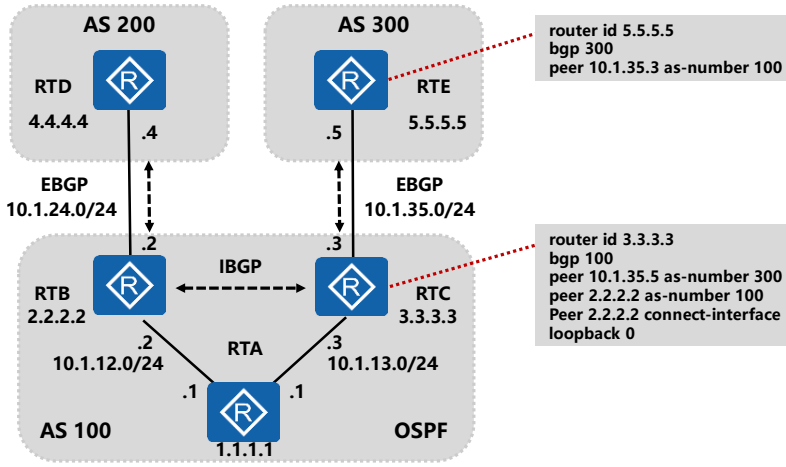
BGP Neighbor Relationship Configuration



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Configuration procedure:
  - Configure a router ID to identify a router.
  - Configure an EBGP neighbor relationship to transmit routes between ASs.
  - Configure an IBGP neighbor relationship to transmit routes within an AS.
- Description:
  - If no router ID is configured for a BGP router, it automatically selects a router ID according to the following rules:
    - Selects the highest IP address among all loopback interfaces.
    - Selects the highest IP address among all physical interfaces if it does not have loopback interfaces.
    - Configuration command: router id X.X.X.X
  - BGP neighbor relationship type is identified by the configured AS number. The parameter following the peer keyword indicates the interface IP address of the neighbor, and the parameter following the as-number keyword indicates the AS number of the neighbor. If two routers have the same AS number, they establish an IBGP neighbor relationship. If they have different AS numbers, they establish an EBGP neighbor relationship.

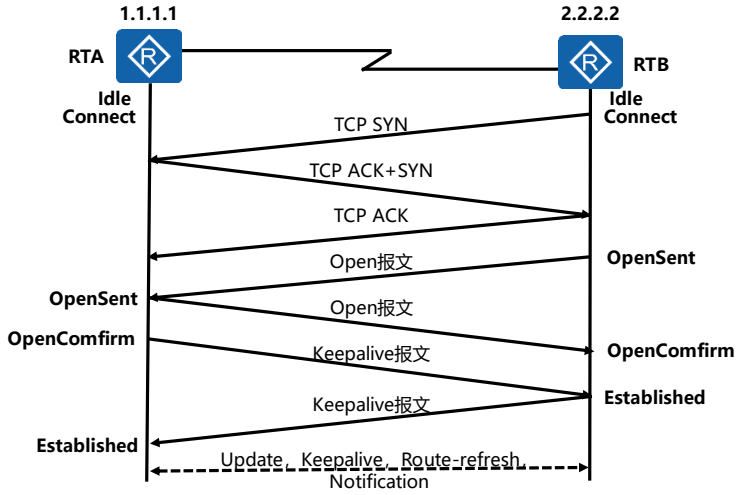
BGP Neighbor Relationship Configuration Optimization



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Directly connected interface IP addresses are often used to establish EBGP neighbor relationships, and loopback interface IP addresses are often used to establish IBGP neighbor relationships.

BGP Neighbor Relationship Establishment



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- BGP routers exchange BGP messages to establish neighbor relationships and update routing information. BGP messages are classified into Open, Update, Notification, Keepalive, and Route-refresh messages.
  - Open message: is the first message sent after a TCP connection is established. It is used to establish a BGP connection between neighbors. After a BGP neighbor receives an Open message and negotiation succeeds, the neighbor sends a Keepalive message to confirm and retain the connection. Then BGP neighbors can exchange Update, Notification, Keepalive, and Route-refresh messages.
  - Update message: is used to exchange routing information between BGP neighbors. It can advertise multiple reachable routes with the same route attributes and withdraw multiple unreachable routes.
    - An Update message can advertise multiple reachable routes with the same route attributes. These routes can share a group of route attributes. All the route attributes carried in a specific Update message apply to all the destinations (specified by IP prefixes) of the Network Layer Reachability Information (NLRI) field in this Update message.
    - An Update message can withdraw multiple unreachable routes. Each withdrawn route identified by a destination address (an IP prefix) is the route that was advertised between BGP routers.

Contents

1. VLAN technology
2. Link aggregation technology
3. OSPF protocol
4. **BGP protocol**
   - BGP Overview
   - BGP Neighbor Relationship Establishment and Configuration
   - BGP Route Advertisement Rules and Route Processing

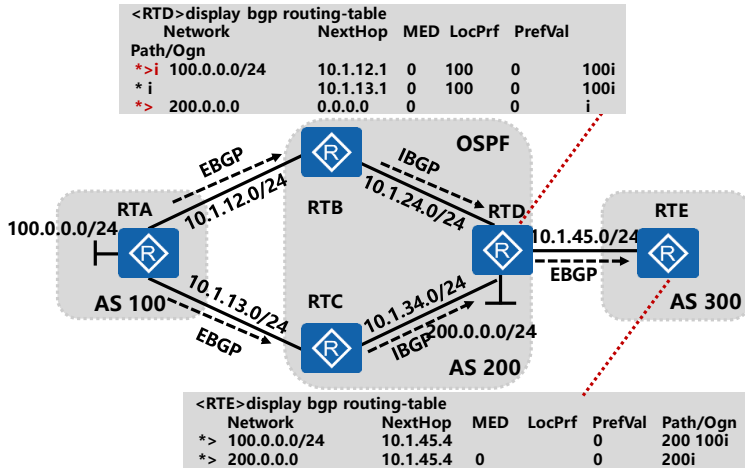
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

BGP Update Message

- BGP routes are generated in either network or import mode. They are encapsulated in Update messages and advertised to neighbors. BGP advertises routing information only after a neighbor relationship is established.
- Update messages are used to advertise reachable routes and withdraw unreachable routes. An Update message contains the following information:
  - Network Layer Reachability Information (NLRI): advertises the IP prefix and prefix length.
  - Path attribute: provides loop detection and controls optimal route selection.
  - Withdrawn route: describes the prefix and prefix length of the unreachable withdrawn route.
- BGP route advertisement must follow specific rules to prevent potential problems.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

BGP Route Advertisement Rule (1)

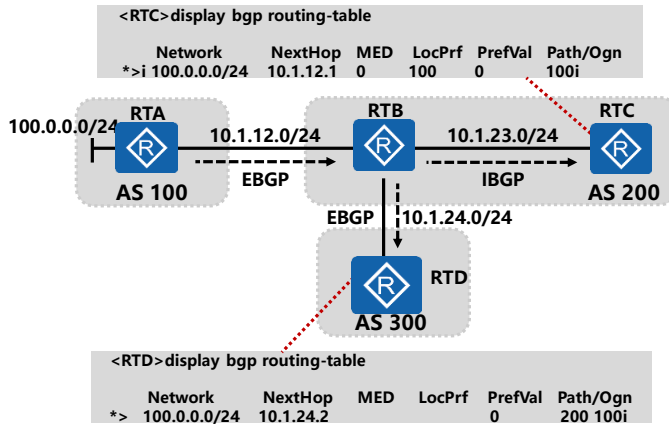


- BGP Route Advertisement Rule 1: Advertise Only the Optimal Route to Neighbors

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- When multiple valid routes exist, a BGP router advertises only the optimal route to its neighbor.
  - RTD can learn the route 100.0.0.0/24 from two BGP neighbors (RTB and RTC) and RTD advertises its directly connected route 200.0.0.0/24 into BGP. Run the display bgp routing-table command on RTD. The following command output is displayed:
  - Run the display bgp routing-table command on RTE. The following command output is displayed. You can view that RTD has advertised the optimal route marked valid to its BGP neighbor RTE.
- Fields in a BGP routing table include:
  - Status codes: * - valid, > - best, d - damped, h - history, i - internal, s - suppressed, S - Stale
  - Origin : i - IGP, e - EGP, ? – incomplete
  - Network: network address
  - NextHop: next-hop address
  - MED: route metric
  - LocPrf: local preference
  - PrefVal: protocol preferred value
  - Path/Ogn: AS_Path and Origin attribute
  - Community: Community attribute information

BGP Route Advertisement Rule (2)



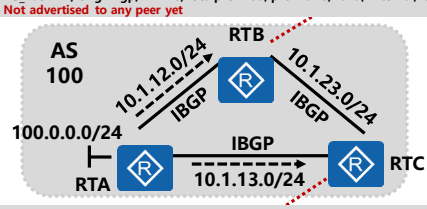
- BGP Route Advertisement Rule 2: Advertise the Optimal Route Obtained Through EBGP to All BGP Neighbors.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A BGP router advertises the optimal route obtained through EBGP to all BGP neighbors (including EBGP neighbors and IBGP neighbors).
  - In the figure, RTA has a user network segment 100.0.0.0/24 and advertises this network segment to a BGP neighbor RTB through EBGP. After RTB receives this route from its EBGP neighbor, it advertises this route to its IBGP neighbor RTC and EBGP neighbor RTD.

BGP Route Advertisement Rule (3)

```
<RTB>display bgp routing-table 100.0.0.0
BGP local router ID : 2.2.2.2
 Local AS number : 100
 Paths:   1 available,  1 best, 1 select
 BGP routing table entry information of 100.0.0.0/24:
 From: 10.1.12.1 (1.1.1.1)
 Route Duration: 00h01m39s
 Relay IP Nexthop: 0.0.0.0
 Relay IP Out-Interface: GigabitEthernet0/0/0
 Original nexthop: 10.1.12.1
 Qos information : 0x0
 AS-Path Nil, origin igp, MED 0, localpref 100, pref-val 0, valid, internal, best, select, active, pre 255
 Not advertised to any peer yet
```



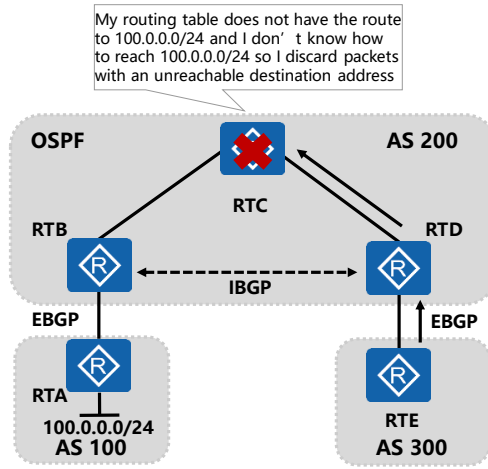

```
<RTC>display bgp routing-table
      Network         NextHop        MED        LocPrf    PrefVal  Path/Ogn
*>i   100.0.0.0/24    10.1.13.1      0          100       0        i
```

- BGP Route Advertisement Rule 3: Do Not Advertise the Optimal Route Obtained Through IBGP to Other IBGP Neighbors.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A BGP router does not advertise the optimal route obtained through IBGP to other IBGP neighbors.
  - In the figure, RTA has a user network segment 100.0.0.0/24. RTA, RTB, and RTC are IBGP neighbors. RTA advertises the route 100.0.0.0/24 to RTB and RTC through IBGP, but RTB does not advertise the received IBGP route to its IBGP neighbor RTC.
  - This design prevents routing loops within an AS. As defined, when a BGP route is transmitted within an AS, its AS_Path attribute remains unchanged. In the figure, when RTA advertises the route 100.0.0.0/24 to RTB, the AS_Path attribute of this route remains unchanged and is empty. If RTB can advertise this IBGP route to RTC, RTC may also advertise this route to RTA because the AS_Path attribute of the route is still empty, and RTA will not reject this route. As a result, a routing loop occurs. Therefore, this route advertisement rule can prevent routing loops within an AS.

BGP Route Advertisement Rule (4)

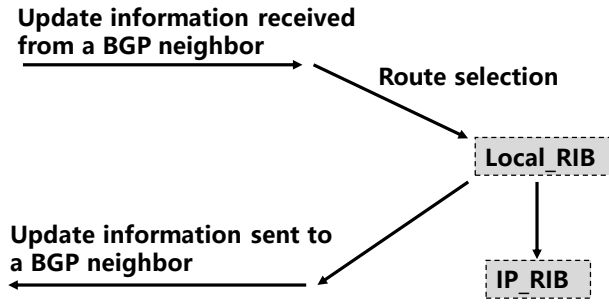


- BGP Route Advertisement Rule 4: Synchronize BGP and IGP.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- RTA has a user network segment 100.0.0.0/24 and advertises it to RTB through EBGP. RTB and RTD establish an IBGP neighbor relationship. RTD learns this BGP route through IBGP and advertises it to the EBGP neighbor RTE.
- When RTE accesses the network segment 100.0.0.0/24, it examines its routing table, finding that the next hop of the route to 100.0.0.0/24 is RTD. After RTE finds the outbound interface, it sends a packet to RTD. RTD receives the packet and examines its routing table, finding that the next hop of the route is RTB and the outbound interface is the interface connected to RTC and sends the packet to RTC. RTC receives the packet and examines its routing table, finding that there is no route to 100.0.0.0/24 and discards this packet. In this situation, the routing blackhole problem occurs.
- BGP route advertisement rule: Before a BGP router advertises a route learned from an IBGP neighbor to another BGP neighbor, IGP must know this route. That is, BGP must synchronize with IGP.

BGP Routing Information Processing



- When receiving an Update message from a BGP neighbor, a BGP router uses the route selection algorithm to determine the optimal route for each prefix.
- The router stores the selected optimal route to the local BGP routing table (Local_RIB) and then submits it to the local IP routing table (IP_RIB) to determine whether to install it.
- The router encapsulates the selected valid optimal route in an Update message and sends it to the BGP neighbor.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved. HUAWEI

- IP routing table (IP_RIB): global routing information database, including all IP routing information.
- BGP routing table (Local_RIB): BGP routing information database, including routes selected by the local BGP router, neighbor table, and neighbor list.
- After receiving an Update message from a BGP neighbor, a BGP router uses the route selection algorithm to determine the optimal route for each prefix and stores the selected optimal route to the local BGP routing table (Local_RIB).
- If multipath is enabled on a BGP router, it submits the optimal route and all equal-cost routes to IP_RIB to determine whether to install them. In addition to the optimal route received from BGP neighbors, Local_RIB also includes the routes injected by the router. These routes are called locally originated routes.
- In Local_RIB, a router encapsulates only the optimal prefix in an Update message and advertises it to BGP neighbors.

Quiz

1. What is the principle of VLAN?
2. What is the value of link aggregation technology?
3. What is the core idea of the OSPF routing protocol?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What is the principle of VLAN?
  - The virtual local area network (VLAN) technology groups devices on a physical LAN into different logical LAN segments.。
- What is the value of link aggregation technology?
  - Increased bandwidth
    - The bandwidth of the link aggregation interface is the sum of bandwidth of member interfaces.
  - Higher reliability
    - When an active link fails, traffic on this active link is switched to another active link, improving reliability of the link aggregation interface.
  - Load balancing
    - In a link aggregation group (LAG), traffic is load balanced among active links of member interfaces.
- What is the core idea of the OSPF routing protocol?
  - The most notable feature of OSPF is the use of link state algorithms to maintain routing tables.

Summary

- VLAN technology
- Link aggregation technology
- OSPF protocol
- BGP protocol

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei's official websites:
  - Enterprise business: http://e.huawei.com/en/
  - Technical support: http://support.huawei.com/enterprise/en/
  - Online learning: http://learning.huawei.com/en/
- Document tool
  - HedEx Lite
- Simulator
  - eNSP

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei E-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Data Center SDN Network Basics

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This course describes data center SDN network basics, helping you learn about challenges faced by cloud data centers and the industry's SDN development and technologies.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Describe new challenges for cloud data center networks.
  - Describe the industry's SDN development and technical roadmap.
  - Understand Huawei's CloudFabric Solution.

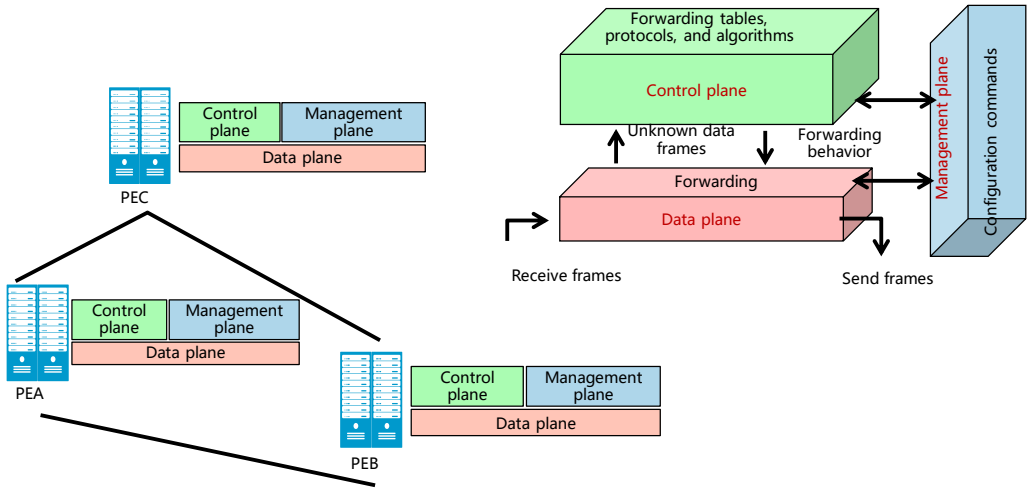
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Limitations of Traditional Networks**
2. SDN Overview
3. SDN Value
4. SDN Standards Organizations
5. SDN Applications

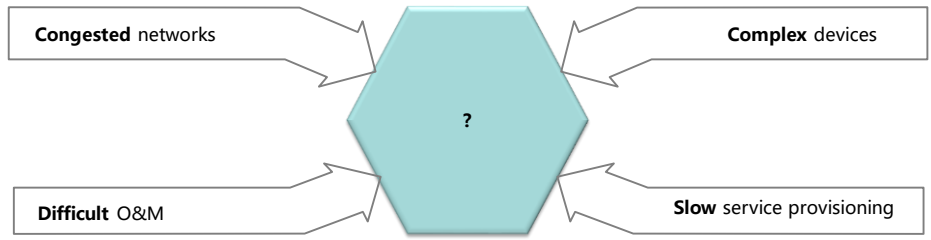
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Traditional Distributed Network



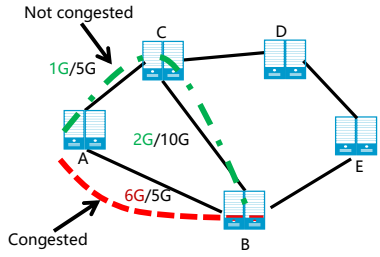
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Problems Faced by Traditional Networks



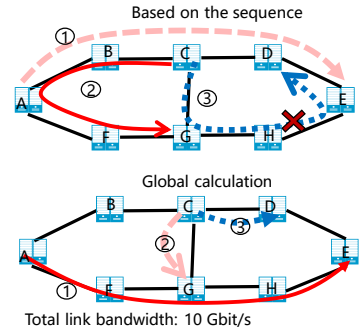
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Congested Networks



Occupied bandwidth/Total bandwidth

Question 1: The link between router A and router B is the shortest path. Packet loss is going to occur on this link. However, other links are idle. Why not adjust traffic to the path from router A to router C and then to router B?



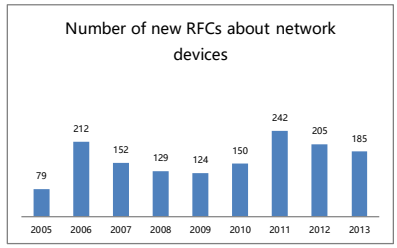
Total link bandwidth: 10 Gbit/s

Service requirements:
1. A->E: 6 Gbit/s; 2. C->G: 4 Gbit/s; 3. C->D: 8 Gbit/s. (Bandwidth of each link is 10 Gbit/s. The above topology shows the service setup sequence.)
Question 2: In traditional mode, services are deployed based on the sequence, and service 3 fails to be set up. Why not use global calculation to ensure that all services are successfully set up?

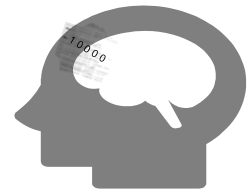
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Complex Network Devices from Multiple Vendors

- If you want to become an IP expert, you will probably have to read more than 2,500 RFCs about network devices — that would take over six years if you read one a day. And that only accounts for a third of all RFCs, the number of which continues to grow.



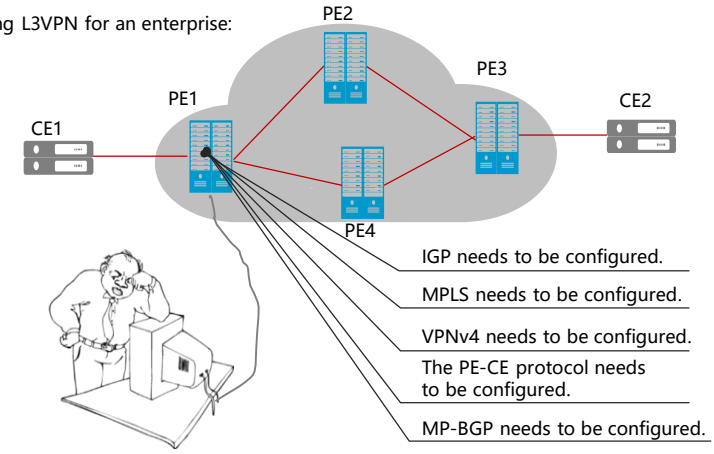
- If you want to have expert knowledge of a vendor's devices, you will need to master more than 10,000 commands — a number that is constantly growing.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Numerous Network Protocols, Difficult Network Management and O&M

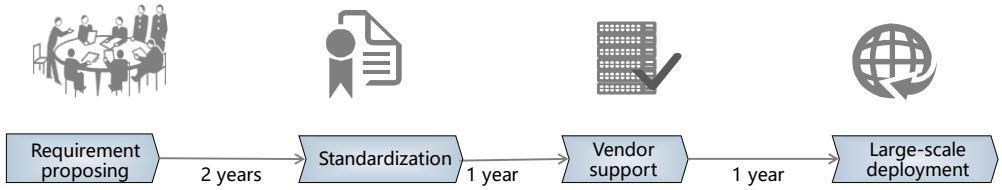
Example for deploying L3VPN for an enterprise:

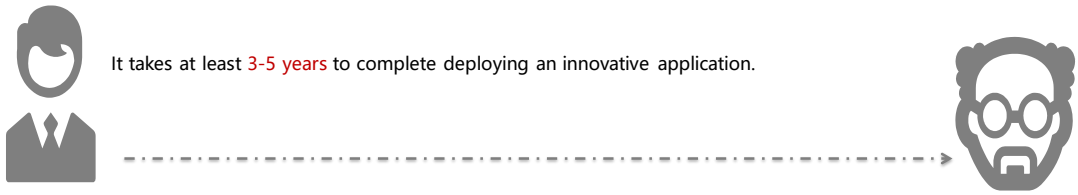


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Approximately 50 configuration commands for a single PE:
- 1 Configure IGP
- [~PE1] interface loopback 1
- [~PE1-LoopBack1] ip address 1.1.1.9 32
- [~PE1] interface gigabitethernet3/0/0
- [~PE1-GigabitEthernet3/0/0] ip address 172.1.1.1 24
- [~PE1] ospf
- [~PE1-ospf-1] area 0
- [~PE1-ospf-1-area-0.0.0.0] network 172.1.1.0 0.0.0.255
- [~PE1-ospf-1-area-0.0.0.0] network 1.1.1.9 0.0.0.0
- [~PE1-ospf-1-area-0.0.0.0] quit
- 2 Configure MPLS
- [~PE1] mpls lsr-id 1.1.1.9
- [~PE1] mpls
- [~PE1] mpls ldp
- [~PE1] interface gigabitethernet 3/0/0
- [~PE1-GigabitEthernet3/0/0] mpls
- [~PE1-GigabitEthernet3/0/0] mpls ldp

Slow Provisioning of Innovative Network Services





Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

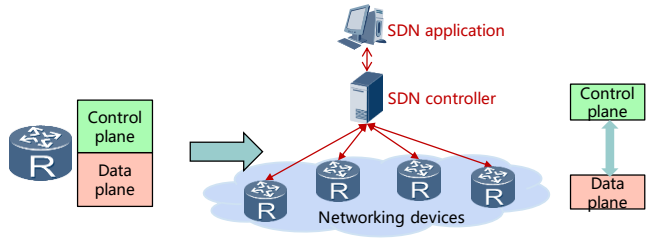
Contents

1. Limitations of Traditional Networks
2. **SDN Overview**
3. SDN Value
4. SDN Standards Organizations
5. SDN Applications

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Origin of SDN

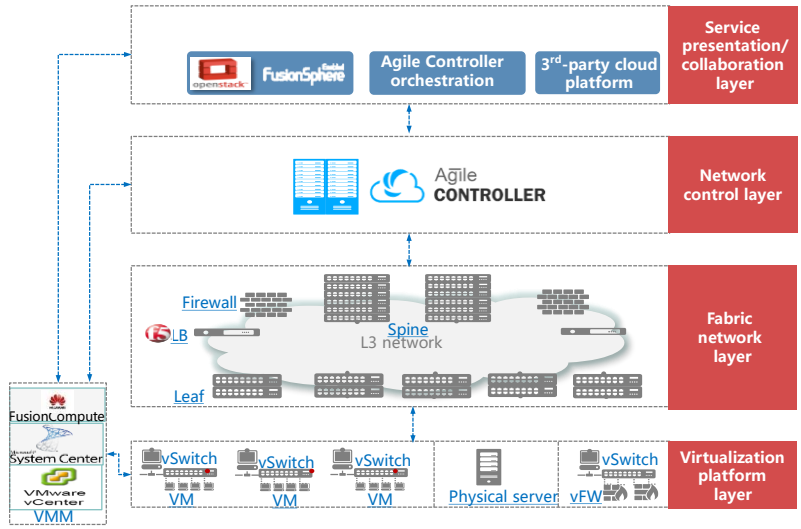
- Software-Defined Networking (SDN) is a new innovative network architecture that was proposed in the Clean Slate program at Stanford University.
- As the core technology of SDN, OpenFlow separates network devices' control plane from the data plane to implement flexible network traffic control. SDN provides a well designed platform for core network and application innovations.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- SDN is oringal of data plane and control plane separation, but is not limited to it.
- With the development of technology, overlay and traffic optimization solutions are also treated as SDN.

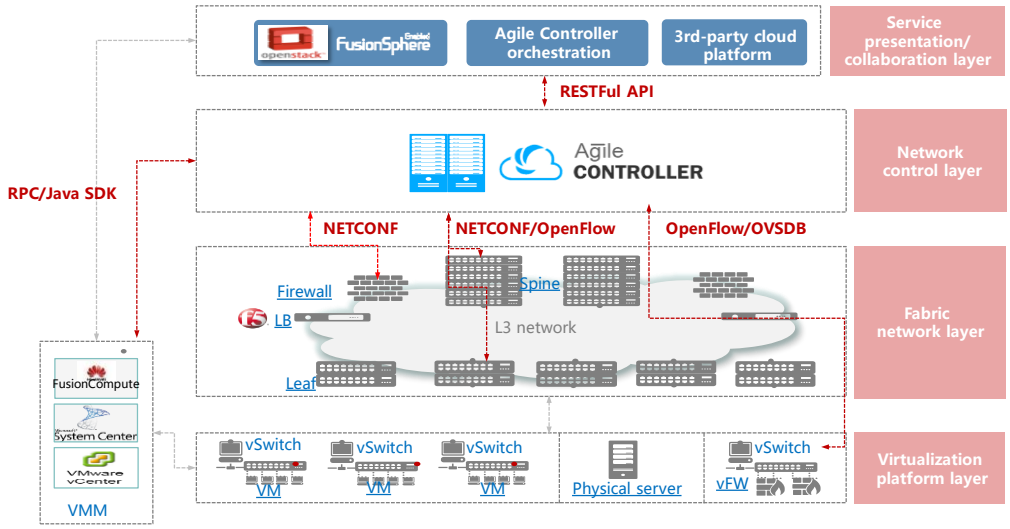
Architecture of SDN Data Center Networks



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Overlay solution is the best practice in data center.
- Service presentation/collaboration layer
  - Provides service portals for carriers, enterprises, VPCs, and RSPs.
  - Provides customizable service portals.
- Network control layer
  - The network control platform is the SDN controller (Agile Controller), which implements network modeling and network instantiation.
  - Northbound open APIs can be deployed for rapid customization and automatic provisioning of services.
  - Southbound OpenFlow/NETCONF interfaces can be deployed for centrally managing and controlling physical and virtual networks.
- Fabric network layer
  - Physical devices set up a basic physical network to support the overlay network.
  - The hardware VXLAN gateway improves service performance.
  - The fabric network is compatible with traditional VLANs.
- Virtualization platform layer
  - vSwitches manage network configurations and policies for local VM access.
  - vFWs enforce security policy control and load balancing based on software firewalls.

Interface Relationship Between Components



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

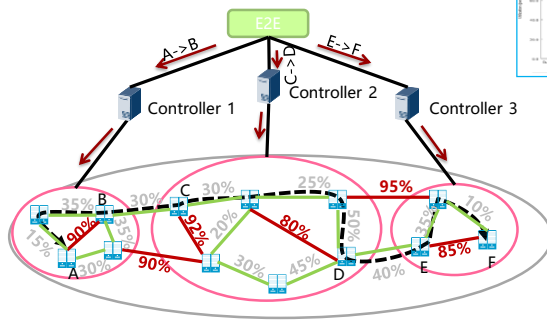
Contents

1. Limitations of Traditional Networks
2. SDN Overview
3. **SDN Value**
4. SDN Standards Organizations
5. SDN Applications

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Optimizing Network Paths and Traffic







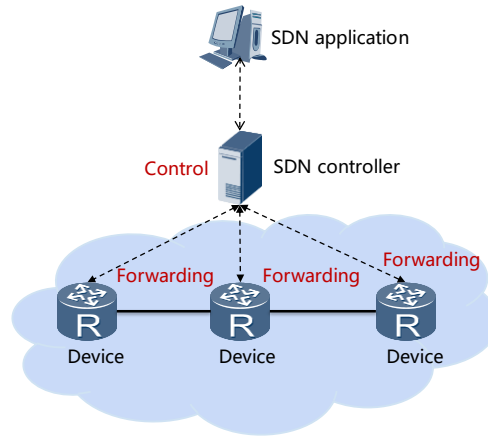
Google increased its average link utilization from 30% to almost 95% through SDN.

- Traffic optimization solution is preferred in WAN network.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Simplifying Networks

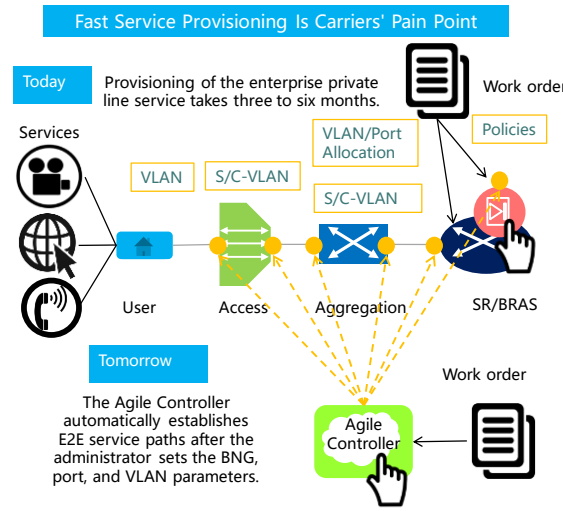
- By centralizing network control and separating the forwarding and control planes, SDN simplifies networks and eliminates deployment of many IETF protocols, lowering the costs of learning and O&M, and increasing the service deployment speed.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Fast Network Innovation

- The programmability and openness of SDN accelerate service development and innovation.
- When deploying new services on a network, you can modify or enhance SDN software to achieve fast network programmability and provisioning and verification of new services.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

White-Box Trend of Network Devices

- If interfaces between controllers and forwarders are standardized based on the SDN architecture (for example, the OpenFlow protocol matures gradually), the white-box trend of network devices is possible.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Limitations of Traditional Networks
2. SDN Overview
3. SDN Value
4. **SDN Standards Organizations**
5. SDN Applications

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Different Opinions on SDN

... separated control and forwarding planes, OpenFlow ...

ONF

... operate the network like a "programmable patch panel"...

Guido Appenzeller, Former CEO **of Big Switch**

... separated control and forwarding planes and open programmable interfaces ...

Gartner



... overlay virtualized network ...

VMware

... open programmable interfaces and the resulting service agility ...

Light Reading network expert

Definitions of SDN differ. ... What matters is the value that SDN brings.

Richard Huang, CEO of QingCloud

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Major Schools of Thought in SDN

Each of the schools is trying to establish dominance in the industry chain while SDN is flourishing, and for that reason it is difficult in the short term to create uniform industry standards.

Revolutionary School: Based on OpenFlow

ONF OPEN NETWORKING FOUNDATION

Characteristics:
1. Separated forwarding and control planes.
2. Standardized forwarding plane.
3. OpenFlow standardization.

Members:
1. Google
2. Ericsson
3. NEC

Challenges:
1. Costs in unifying forwarding devices.
2. Difficulty in evolving the live network.

Reformist School: Improve the Existing System

IETF

Characteristics:
1. Open capabilities of network devices.
2. Standardized APIs.
3. Emphasis on smooth device evolution.

Members:
1. Cisco
2. Juniper

Challenges:
1. Insufficient programmability.

IT School: Based on COTS

Characteristics:
1. Software-based network functions.
2. Unified hardware platform.

NFV ISG

ETSI World Class Standards

Members:
1. VMware

Challenges:
1. Performance problem of software-based network functions.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SDN Open Source Communities

OpenDaylight Open Source SDN Controller

Characteristics:

1. Open source SDN controller.
2. SAL is compatible with multiple protocols, including OpenFlow, BGP, PCEP, I2RS, SNMP, and NETCONF.
3. Declared orientation to carriers and data centers.

Challenges:

1. De facto standards are controlled by some vendors.
2. Sustainability of the open source system.

ONOS Open Source SDN Controller

Characteristics:

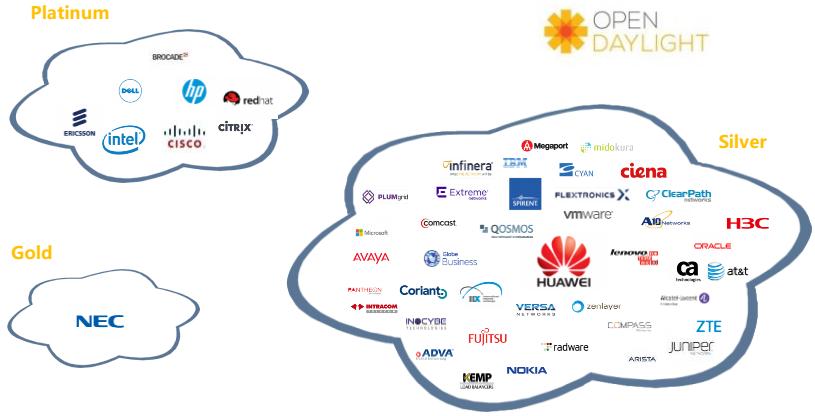
1. Open source SDN controller.
2. Southbound interfaces of the controller mainly support OpenFlow.
3. Declared orientation to carriers.

Challenges:

1. Implementation of OpenFlow.
2. Sustainability of the open source system.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

OpenDaylight Members



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

ONOS Members

- Partners

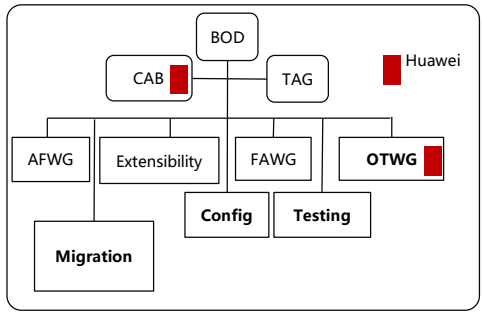


  - Alcatel-Lucent, AT&T, China Unicom, Ciena, Cisco, Ericsson, Fujitsu, Huawei, Intel, NEC, NTT, SK Telecom, Verizon.
- Collaborators
  - AARNET, Adara, Airhop Communications, Akamai, AmLight, BlackDuck, BTI Systems, Beijing University of Posts and Telecommunications, Cavium, ClearPath Networks, CNIT, CREATE-NET, Criterion Networks, CSIRO, ECI Telecom, ETRI, Consortium GARR, GEANT, Happiest Mind, Internet2, KAIST, KREONET, KISTI, NAIM Networks, NetCracker, OpenFlow Korea, Oplink Communications, ONF, Postech, Radisys, SRI International.

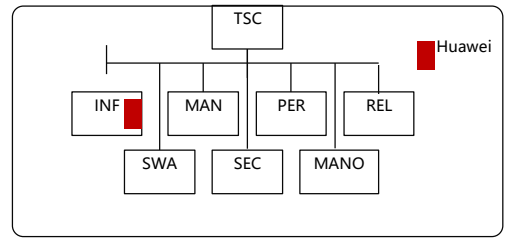
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Huawei Actively Promotes SDN Standardization (1)

Chair/Vice Chair of Multiple ONF WGs

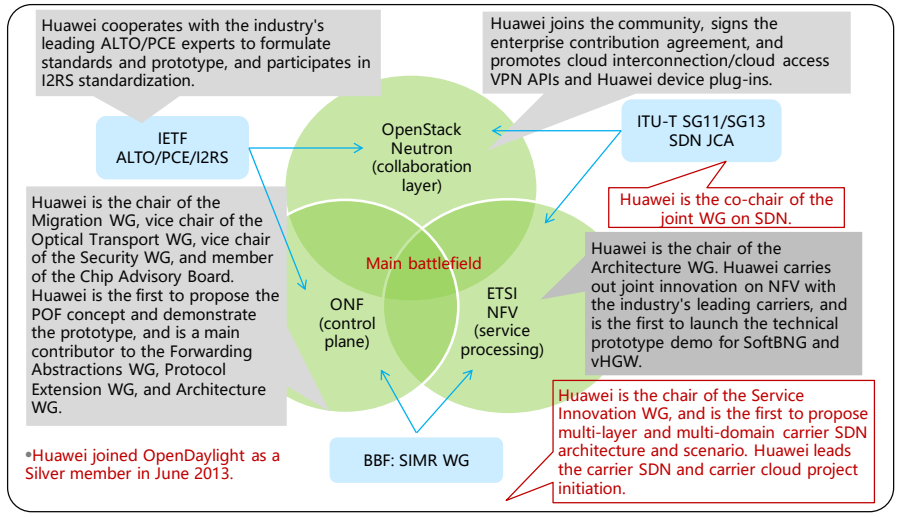


Vice Chair of the NFV Architecture WG, Contributing the Most Articles



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Huawei Actively Promotes SDN Standardization (2)



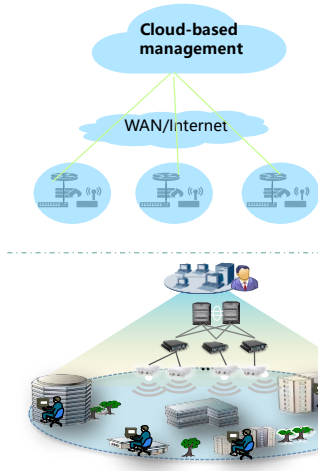
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Limitations of Traditional Networks
2. SDN Overview
3. SDN Value
4. SDN Standards Organizations
5. **SDN Applications**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cloud-based Management of Government/Enterprise Campus Networks



Cloud-based management of campus networks:

- **High efficiency**: Networks are managed on the cloud.
- **Human labor saving**: Professional maintenance tools and capabilities are deployed on the cloud.
- **Quick provisioning**: New services and functions are provided on the cloud.
- **Maintenance-free**: Devices support plug-and-play deployment, and multiple service modes are enabled.

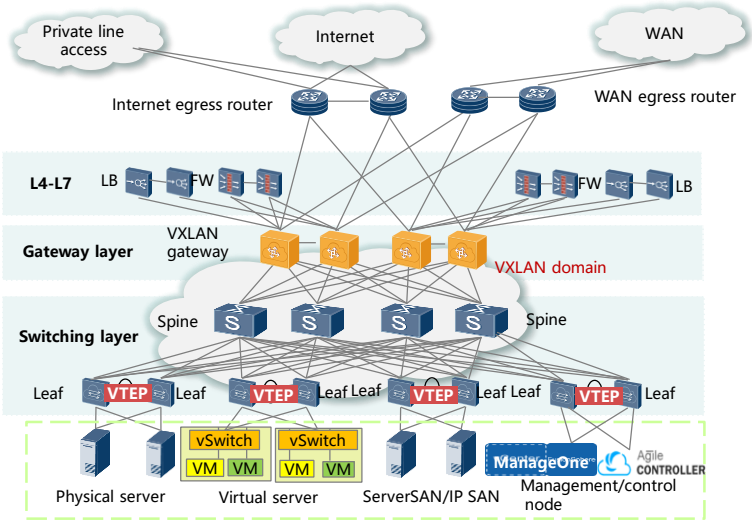
Management of traditional campus networks:

- Networks are complex, requiring professional maintenance.
- Heavy manual configuration workloads lead to slow service provisioning.
- Branches rely on on-premises maintenance.
- OPEX accounts for up to 70% of the total cost.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Application Scenarios
  - Medium-sized enterprises who expect to focus on their core business but are unwilling and have no capability to build an independent IT team
  - Branches and edge nodes of large enterprises
- Benefits to Customers
  - Reduce OPEX and save approximately 83% of human labor costs.
  - Increase IT-based innovative services by over 40%.
- Solution Description
  - Network planning: efficient and professional network planning and visualized and predictable WLAN planning without coverage holes
  - Network deployment: plug-and-play deployment of devices and cross-NAT management
  - Service provisioning: device group management and quick template configuration to enable fast service provisioning
  - Network monitoring: GIS-based visible network status and mobile O&M app to enable network monitoring anytime, anywhere
  - Fault location: multiple online location tools and one-click fault diagnosis
- Highlights
  - 3D cloud-based network planning, cloud-based PMI, and mobile O&M

Data Center Architecture Reconstruction

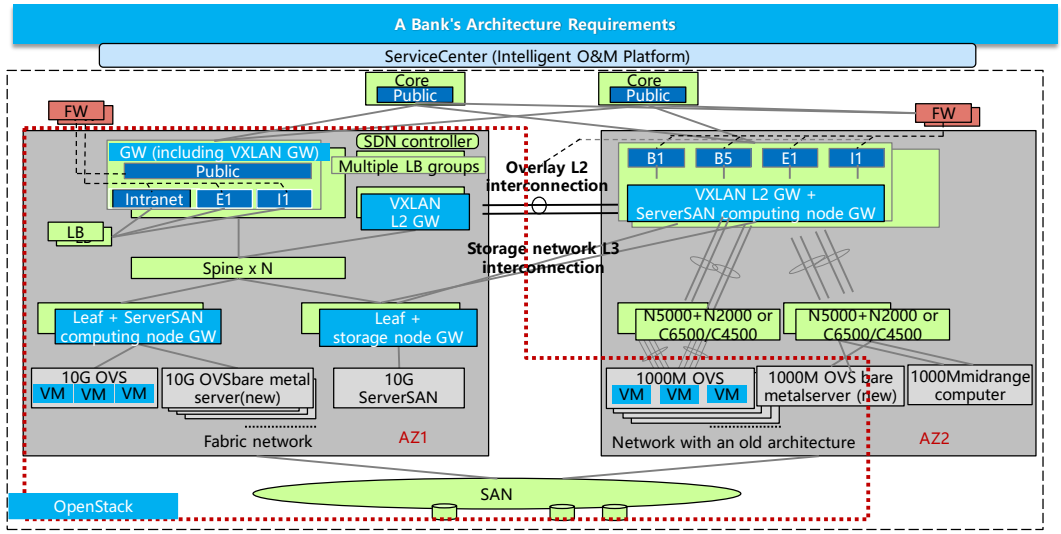


Flat topology structure

Network traffic trend transferring

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

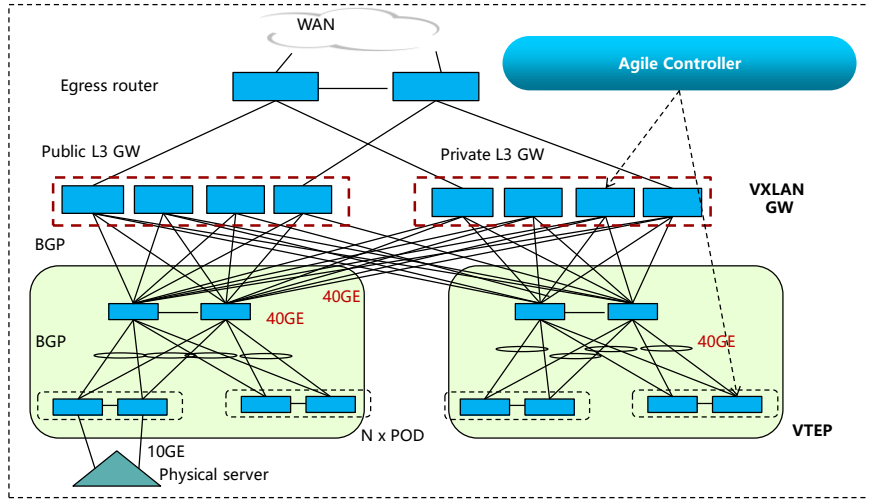
Network and Service Automation for a Bank



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Service Automation Capability
  - The ServiceCenter delivers services to the OpenStack cloud platform and Agile Controller, achieving E2E automated delivery of network services.
- SDN+VXLAN Solution
  - The solution uses the overlay hierarchical network technology, supports large L2 communication between the new and old networks, resource sharing, and flexible scheduling, and achieves converged carrying of data access, management, and storage services with the fabric architecture.
- SDN-based Security Service Automation Solution
  - The ServiceCenter uniformly delivers security policies, achieving automated provisioning of security services.
- Current status and problems:
- 17 cabinets, 4,570 servers, 27 physical network partitions, and 107 firewalls
- Long service TTM: The network is highly coupled with applications. The service TTM is as long as 90 days.
- Low resource utilization: Data center resources are distributed in different physical partitions, resulting in low resource utilization.
- Difficult security management: 20,000 security policies and 10,000 changes each year.

Highly Efficient Network Operations for an e-Commerce Platform



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- An e-Commerce Platform's Cloud Network Development Requirements
  - Fast service delivery: Efficient provisioning of new services must be supported.
  - Highly reliable, stable, and flexible network: Smooth service running must be ensured.
  - Automated network O&M: Manually upgrading network devices takes at least one year.
- Value of Huawei's SDN Solution
  - Controller level: The Agile Controller is responsible for underlay network control and collaboration and achieves automated network deployment.
  - Fabric level: The solution provides high-performance physical VXLAN L3 gateways, implements collaboration of the virtual overlay network and physical underlay network, and simplifies O&M.

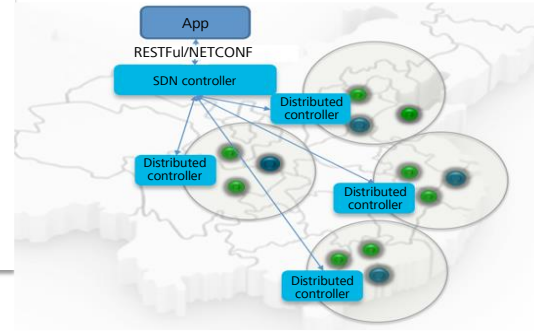
Traffic Scheduling Between Multiple Data Centers for an Internet Company

An Internet Company's Network Development Requirements

10+ cities, 60+ IDC clusters, 300,000+ servers, 20,000 Gbit/s+ IDC bandwidth, and 5,000 GB+ CDN traffic

Problems:

- The overall private line utilization was low. However, local capacity expansions frequently occurred.
- The business department suffered from insufficient WAN bandwidth and poor flexibility.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Solution Description
  - The SDN controller and routers run IS-IS/SNMP. The controller collects network topology and bandwidth information.
  - The SDN controller completes path computation based on information transferred using IS-IS/SNMP and delivers path computation results to routers using PCEP.
- Solution Value
  - The service provisioning efficiency is increased 10-fold, and the network-wide DCI link utilization is increased to 80%.
  - The service scheduling time is decreased from minutes to seconds.
  - Centralized path computation reduces resource occupancy and accelerates network convergence.



Quiz

1. What are the routes from SDN to development and commercial application?
2. Why is overlay technology a best practice in data center networks?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- What are the routes from SDN to development and commercial application?
  - SDN originated from the laboratory openflow control separation, commercial applications developed a variety of routes such as traffic tuning, overlay and so on.
- Why is overlay technology a best practice in data center networks?
  - The overlay is compatible with traditional network technologies, and the evolution is smooth. It is better to implement service separation and rapid deployment of service networks to cope with daily network changes.

Summary

- Limitations of Traditional Networks
- SDN Overview
- SDN Value
- SDN Standards Organizations
- SDN Applications

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei's official websites:
  - Enterprise business: http://e.huawei.com/en/
  - Technical support: http://support.huawei.com/enterprise/en/
  - Online learning: http://learning.huawei.com/en/
- Document tool
  - HedEx Lite
- Simulator
  - eNSP

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei E-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Data Center Storage Technology

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- In this lesson, we will first learn why enterprises need data and then introduce how data is generated and stored in the enterprise. Next, we will learn about the Redundant Array Of Independent Disks (RAID). Then, we will learn five important ICT infrastructures: Direct Attached Storage (DAS), Storage Area Network (SAN), Network Attached Storage (NAS), distributed storage, and object storage.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- After completing this module, you will be able to:
  - Explain the importance of data for an organization.
  - Understand the difference between structured and unstructured data.
  - Know about common Redundant Array Of Independent Disks (RAID) types.
  - Describe the features of Direct Attached Storage (DAS), along with its advantages and disadvantages.
  - Identify and describe the main Storage Area Network (SAN) components.
  - Understand the role of Network-attached Storage (NAS) in Information And Communications Technology (ICT) architecture.
  - Master NAS network topology.
  - Understand the principles behind distributed and object storage.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **Data Management**
2. Traditional RAID Technology
3. Introduction to DAS Technology
4. Introduction to SAN Technology
5. Introduction to NAS Technology
6. Huawei Distributed Storage
7. Object Storage

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Data Management

- SNIA definition: Data is the digital representation of anything in any form.
- Three key principles of data management:
  - Companies use or produce large amounts of data in the course of running their business.
  - Data should be available as long as the business needs it.
  - When data is no longer needed it must/can be destroyed.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

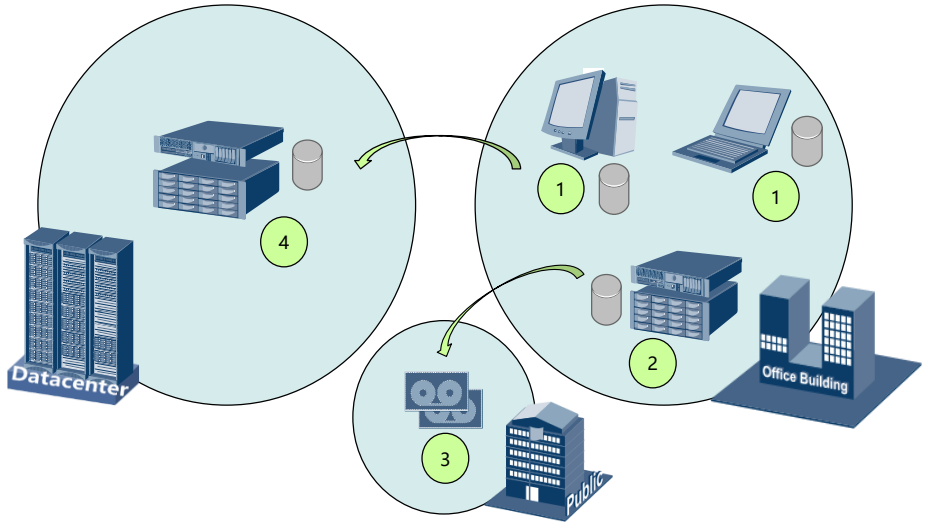
- Storage Networking Industrial Association (SNIA) defines data as the digitalized representation of everything. This definition sounds relatively abstract, but when you look at the operation of today's enterprises, you will find that an enterprise generates a large amounts of data every day. All this data is used to maintain business operation and profitability.

Definition of Information

- Information is data that is processed to meet user decision-making requirements. Information is extracted from collected data.
- Information can:
  - provide a company with marketing information and information about customer behavior.
  - help to run the business more effectively.
  - help to determine risk factors.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

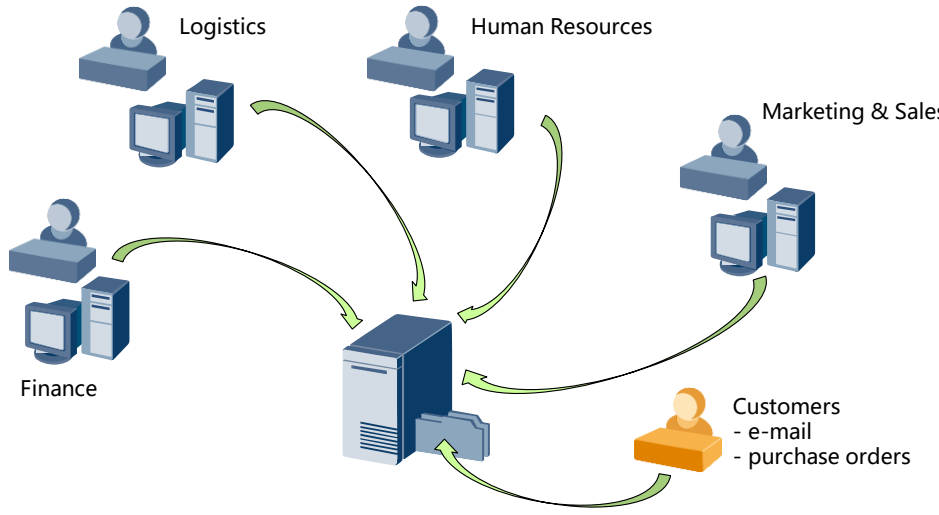
Where Is the Data?



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The following figure shows various data flows. 1, 2, 3, and 4 indicate data generation, local storage, data archiving, and remote backup, respectively.
- Large enterprises usually cover branch offices, which are distributed in different cities or countries. Even a small company may have multiple offices, and each location generates data from tools such as e-mail programs or Word app. Some companies may use their homegrown professional software to generate data.
- For Production-oriented enterprises often use graphic design software (for example, computer-aided design software) to design products, or logistical software to track orders — the transport of parts, materials, tools, and goods. All the data must be saved and secured.
- If data is saved only on employees' laptops or PCs, it is difficult to prevent data loss caused by human errors or hardware faults. This is why most enterprise data is stored in the Main Equipment Rooms (MERs) or data centers. Generally, an MER must provide sufficient air flow and temperatures, as well as power facilities, for all equipment. In a well-equipped data center, in addition to sufficient air conditioners and power facilities, backup power supplies such as diesel generators are also required so that the power supply of the entire data center can still be ensured even in the event of a power abnormality.
- However, even well-equipped data centers require data protection measures. For example, if a storage hardware fault occurs, or if a disaster —flood, fire, or earthquake — affects the data center or MER, the enterprise data is expected to be complete and available. To prevent loss to data or services, a disaster recovery data center or data backup is required to ensure data security in addition to the original data center.

Who Creates or Uses the Data?



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Enterprises generate structured and unstructured data. Structured data refers to data that is easy for machines and staff to handle, digest, and analyze, making it the most common form of data. Unstructured data is that generated from documents, pictures, web pages, and videos, which is harder to process and extremely time- and resource-consuming to process and mine.

Information and Data

- Every company needs information to be able to do business.
- Information is extracted from both structured as well as unstructured data.
- All data is now generated in a digital form.
- Data should be accessible to multiple employees.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Information Lifecycle Management

- Issues to be addressed in information lifecycle management:
  - What data is needed for every person in the organization?
  - In which format should data be presented/kept?
  - How long should data be retained?
  - What needs to be done with the data that is no longer required?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Information can be classified into market data, sales information, along with product, human resource, logistics, and R&D cost. Enterprises must ensure that their employees obtain correct information at the right time.
- The information/data saving format is even as important as the ownership of information. For example, if an employee sends an email containing important information to another employee, the recipient needs to have a terminal, email program, and email account to receive the email. If I receive a file, but my application cannot open it, then I can't get the information.
- But this leads to another question: how long should data be stored? This is related to specific business and data, but generally government regulations specify how long an enterprise should store the information. Some information may even need to be stored for decades. For example, a construction company that built a bridge needs to keep the architect drawings and designs in the time the bridge remains in use.

Information Lifecycle Management: Hard Copy

- What is the format in which data should be available to the organization?
- 1-Hard copy formats:
  - Online data or paper-based data
  - Read only/Eyes only/Not reproduceable
  - Version control
  - Environment requirements for saving hard copies

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Information Lifecycle Management: Digital Information

- What is the format in which data should be available to the organization?
- 2-Digital information:
  - What software needs to be used to browse or modify data?
  - Text storage format (PDF, ODF, or DOC...)
  - Picture storage format (JPG, TIFF, DWG, or PNG...)
  - Is there any standard that needs to be complied with (ODF, CALS, or BASEL...)?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Digital files have their limits. For example, they need the corresponding software to open, read, or modify them. When creating a digital file, you need to select a file format that allows all related users to access and record data.
- There are three typical file types: Text files, Bitmap, and Vector graphic.
  - Text file contains characters (letters or/and numbers) and a small amount of images, such as word documents, spreadsheets, and database files.
  - Bitmap is where all related graphical elements (or pixels) are stored separately. This means a large amount of storage space is required for storing photos and scanned copies in the bitmap format.
  - Vector graphic images are defined using mathematics and formulas. Typical examples are documents generated by drawing software such as AutoCAD.
- It is best to select a file saving format that allows common software rather than a company's specific software to open the file. Plus, the file format should be supported by common software for many years. For example, a bitmap in the TIFF format, a vector graph in the IGES format, or a file in the SGML format. For text files, the Open Document Format (ODF) format is increasingly used.
- Bitmap information is often compressed for storage. Although compression reduces the storage space of files, it also means loss of some information. Sometimes, the file is lossless saved to prevent information loss. TIFF is a lossless format, and the JPG format is

not lossless because of the internal compression technology.

Information Lifecycle Management: Hardware

- What is the format in which data should be available to the organization?
- 3-How can data be stored?
  - Does the access to information need to be authorized or permitted?
  - Does data need to be read-only (without being modified)?
  - Can multiple employees access the same data at the same time?

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Information Lifecycle Management: Retention Period

- Factors that determine the retention period are:
  - based on the business requirements of the organization itself.
  - based on the general rules for your type of business.
  - based on various regulatory regimes that may differ from country to country and region to region.
- How to arrange for digital information to be stored for many years?

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Information Lifecycle Management: Disposing of Obsolete Information

- Who is responsible for data?
  - SOX, JSOX, EuroSOX
- Information on physical media is destroyed by:
  - Shredding
  - Burning
- Information on digital media is destroyed by:
  - Erasing data from the operating system disk.
  - Secure erasing
  - Disk shredding

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Data that exceeds its retention period is no longer useful, and enterprises no longer need to save it. So, how can the data be destroyed?
- Some regulations stipulate this type of data regulations. For example, the US-based Sarbanes-Oxley (SOX) Act stipulates that a company has the responsibility to preserve or destroy the data generated or used by itself and ensure that no one can copy or bring data out of the enterprise.
- If the information is archived in paper form, the paper must be destroyed (e.g. shredded) or burned.
- Digital information is not so easy to destroy. Traditional methods such as disk formatting are not very secure, because there may be technical methods to restore the data that you want to destroy. Instead, some software can erase the data on the medium and write random data to where the old data is written (multiple times of rewrite operation may be required). For some government agencies, data erasure or disk crash is required to ensure that no one can reuse the disks.

Value of Data

- Recovery Point Object (RPO): 15 minutes
  - When a disaster or emergency occurs, data can be restored to a previous time point. This in turn determines how much data is lost when a disaster occurs. In this case, 15 minutes worth of data is lost.
- Recovery Time Objective (RTO)
  - Refers to the length of time it takes for a system to resume normal operation after a breakdown.
- Cost Of Downtime (COD)
  - Refers to the cost per hour of service disruption.

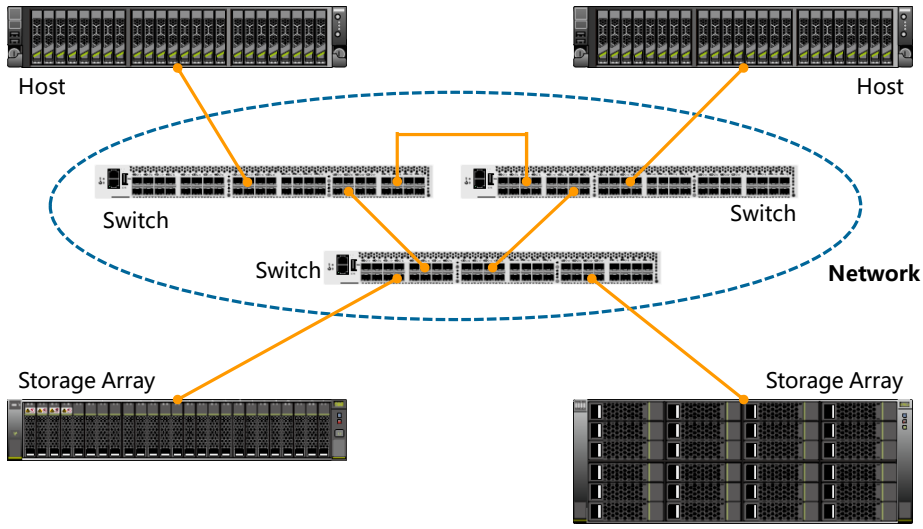
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Components in the ICT Infrastructure

- The following are common terms used in ICT infrastructure:
  - Host: A computer system that has a disk, disk subsystem, or file system for data access and storage.
  - Storage array: A set of disks or tape subsystems with control software available for access.
  - Network: A group of nodes that are interconnected for communication between each other.
  - A network can provide an exclusive channel for any two network nodes that connect to itself.

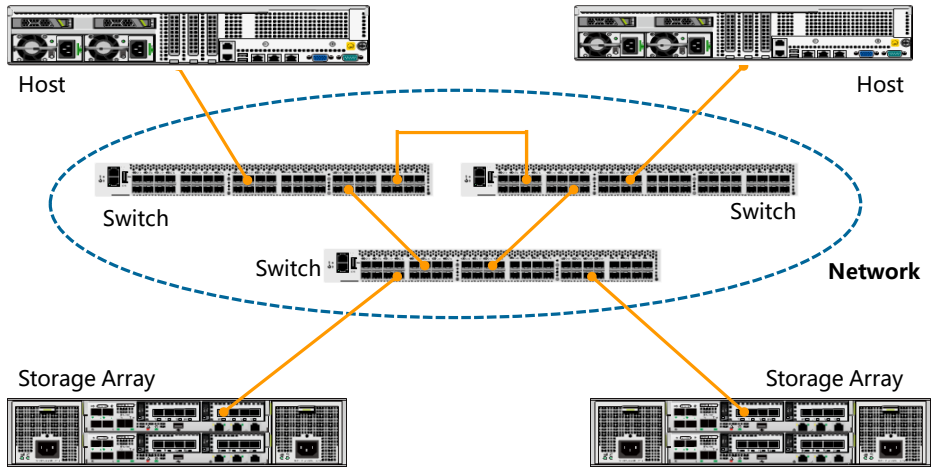
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Front View of ICT Infrastructure Components



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

Rear View of ICT Infrastructure Components



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Data Management
2. **Traditional RAID Technology**
3. Introduction to DAS Technology
4. Introduction to SAN Technology
5. Introduction to NAS Technology
6. Huawei Distributed Storage
7. Object Storage

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Concepts and Implementation Methods of RAID

- RAID: short for redundant array of independent disks, also referred to as a disk array.



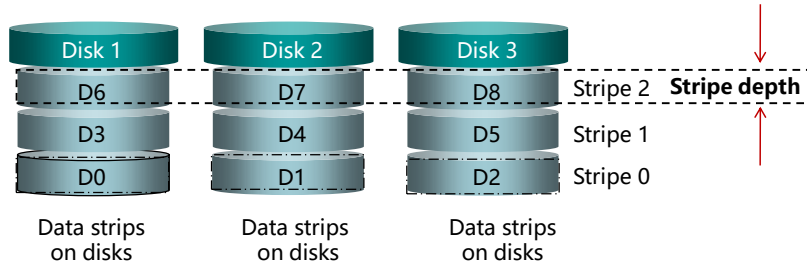
- Implementation methods:
  - Hardware RAID
  - Software RAID

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- RAID technology was originally used to combine multiple small-capacity disks to form a larger storage space, but now if often used for data protection. That is, if a physical device fails, RAID can be used to prevent data loss.
  - By stripping the data on the disk to achieve block access of data and reading the data simultaneously from few drives in the array, RAID reduces the mechanical seek time of the drive and improves the data access speed.
  - By mirroring or storing parity check information, it achieves redundant data protection.
- RAID has many forms, but only a few are still in use. In this chapter, we will discuss the most common types of RAID and learn the other related functions, such as data protection, as well as differences in performance or cost.
- In a storage device, the RAID function can be implemented in the two ways: hardware RAID and software RAID.
  - Hardware RAID uses dedicated RAID adapters, disk controllers, or storage processors. The RAID controller has its own processor, I/O processing chip, and memory, which improves resource utilization and data transmission speed. The RAID controller manages routes and buffers to control data streams between the host and RAID. Hardware RAID is usually used on servers.
- Software RAID does not have its own processor or I/O processing chip, but is completely dependent on the host processor. Therefore, a low-speed CPU cannot meet the RAID implementation requirements. Software RAID is usually used on enterprise-class storage devices.

Data Organization Units of RAID

- Strip: One or more consecutive sectors in a disk form a strip, which is the minimum unit for reading and writing data on a disk. Strips are the elements that form a stripe.
- Stripe: Indicates the strips on the same "location" (or with the same number) of multiple disk drives in the same disk array.

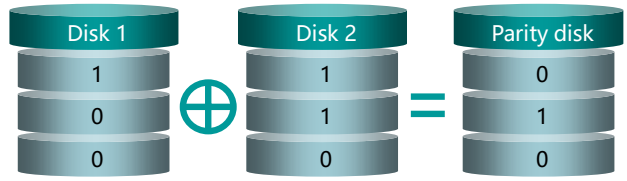


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Stripe width refers to the number of member disks in a stripe, whereas stripe depth refers to the storage capacity of a stripe

RAID Data Protection Methods

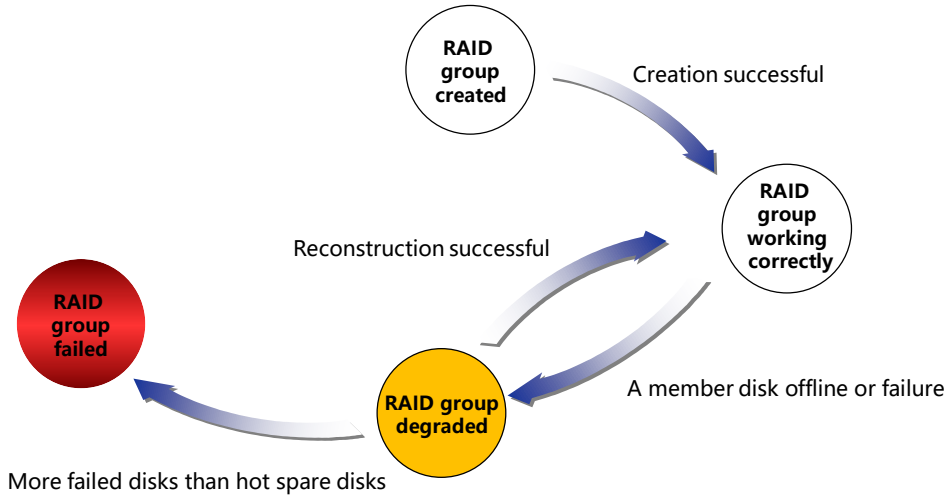
- Method 1: Copies of data that are stored on another redundant disk
- Method 2: Parity check algorithm (XOR)
  - XOR operation is widely used in digital electronics and computer science.
  - XOR: true whenever both inputs differ and false whenever both inputs are the same
  - $0 \oplus 0 = 0$; $0 \oplus 1 = 1$; $1 \oplus 0 = 1$; $1 \oplus 1 = 0$



XOR redundancy backup

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

RAID Status

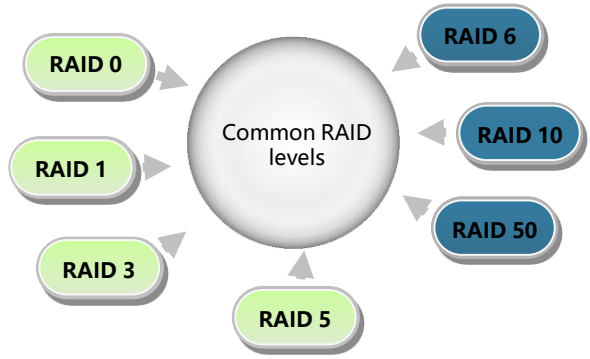


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

- The RAID technology combines multiple physical disks into a RAID group (sometimes referred to as a disk set). This RAID group maintains its own status.
  - When all disks in a RAID group work properly, the RAID group is normal state.
  - If a certain number of disks fail, but the entire RAID group can still prevent data loss and the data recovery process is not started, this state is called degradation.
  - If the faulty disks are replaced or hot spare disks are available in the system, and data is being restored to the new or hot spare disks, this state is called reconstruction.
  - Once reconstruction is successfully complete, the RAID group is back in normal state.
  - If a large number of disks are faulty and the number of faulty disks is greater than the number of redundant disks supported by the RAID type, the data recovery function becomes invalid. This state is called RAID group failure.
- Complete data reconstruction depends on the RAID type, the number of faulty disks, and the availability of new disks.
- Hot spare disks are used to replace failed member disks of a RAID array. Their tasks are to bear data of the replaced disks.

Common RAID Levels and Classification Criteria

- RAID technology combines multiple independent physical disks into a logical disk in different modes. A number of RAID levels exist, corresponding to these modes. This mechanism improves the read/write performance of disks while increasing data security.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Comparison of Common RAID Levels

RAID Level	RAID 0	RAID 1	RAID 5★	RAID6★	RAID 10★	RAID 50
Fault tolerance	No	Yes	Yes	Yes	Yes	Yes
Redundancy type	No	Replication	Parity check	Parity check	Replication	Parity check
Hot spare disk	No	Yes	Yes	Yes	Yes	Yes
Read performance	High	Low	High	High	Medium	High
Random write performance	High	Low	Low	Low	Medium	Low
Sequential write performance	High	Low	Low	Low	Medium	Low
Min. number of disks	2	2	3	4	4	6
Available capacity (Capacity of a single disk)	N x	1/N x	(N - 1) x	(N - 2) x	N/2 x	(N - 2) x

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- From the table above, we can see that there is no RAID type that can meet all requirements. RAID types depend on site speed, security, or cost requirements.
- Each RAID group should not contain too many physical disks because as the number of RAID groups increases (the number of disks increases), the number of disk failures increases accordingly. RAID 5 supports up to 12 disks, whereas, RAID 6 supports up to 42 disks.

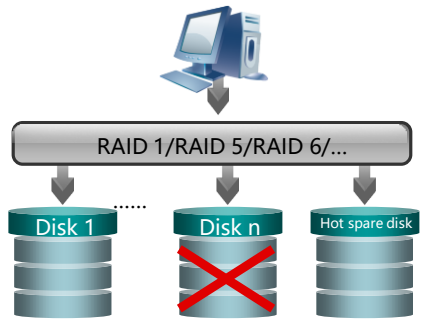
Typical Application Scenarios of RAID

RAID Level	Application Scenario
RAID 0	A scenario requiring fast reads and writes but not high security, such as graphic workstations
RAID 1	A scenario featuring random writes and requiring high security, such as servers and databases
RAID 5	A scenario featuring random transfer and requiring medium security, such as video editing and large databases
RAID 6	A scenario featuring random transfer and requiring high security, such as mail servers and file servers
RAID 10	A scenario involving large amounts of data and requiring high security, such as certain banking or finance applications
RAID 50	A scenario involving random data transmission, security requirements, and concurrency requirements, such as mail servers, and web servers

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Hot Spare Disk

- Hot spare = When one of the disks in a RAID group fails and an idle or standby disk immediately replaces the failed disk, this disk is known as the hot spare.
- Hot spare disks are classified as global hot spare disks or as dedicated hot spare disks.

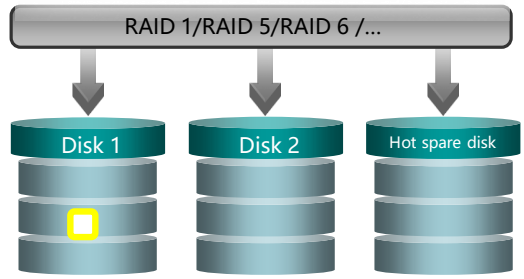


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In most storage solutions, a disk array may have many disks that come in different types. Each disk type has its specific parameters in terms of capacity, rotational speed, access speed, and reliability. By creating multiple RAID groups, you can configure RAID levels and allocate storage capacity using correct specifications. Assume that four RAID groups are in use. How can the administrator configure hot spare disks? How many disks are required for hot spare? It all depends on the actual situation. In normal cases, each RAID group has its own hot spare disk. When one disk fails, a standby disk is available. In four RAID groups, if only one disk is faulty at a time, this hot spare disk can also prevent data loss.
- There are global and local hot spare disks.
  - A hot spare disk shared by different RAID groups is called a global hot spare disk. It will replace any invalid disk in any disk group. The hot spare disk must have equal or greater capacity than the failed disk, and be the same type.
- The hot spare disk that is used by only one RAID group is a local hot spare disk. If a disk in another RAID group fails, this hot spare disk will not take effect.

Pre-Copy

- Pre-copy: When the system detects that a member disk in a RAID group is about to fail, data on that disk is copied onto a hot spare disk, reducing the risk of data loss.

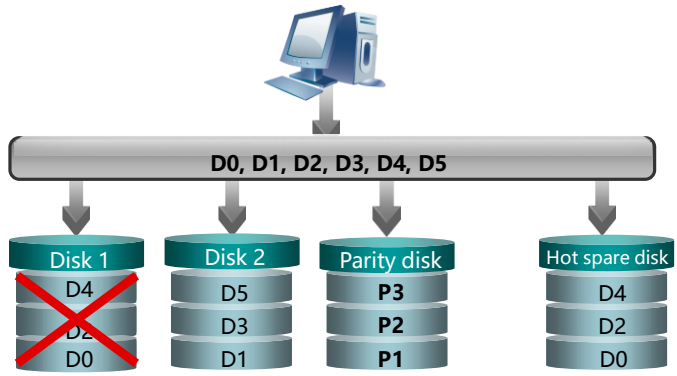


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Pre-copy is another data protection mode of a storage array. For storage device and service maintenance personnel, pre-copy is a good function that makes maintenance easier (or effortless). Most enterprise-class disk devices are provisioned with the SMART tool for disk self-monitoring, analysis, and reporting. This tool enables the disks to monitor their own health by checking the rotation speed and the condition of the magnetic surface.
- With appropriate tools, we can receive messages from smart disks and take corresponding measures. If a smart disk reports that it is not well, it means it is not invalid for the time being, but may fail later.
- When the tool receives a SMART message, the device starts to copy data to the hot spare disk from the disk that is about to fail. When the drive fails later, most of its data has been migrated to the hot spare disk. This reduces the reconstruction time. Copying data to the hot spare disk from the disk that is about to fail is pre-copy.

Reconstruction

- Reconstruction: When a disk in a RAID group is faulty, the system recalculates all data (user data and parity data) on the faulty disk based on the RAID algorithm and other normal member disks, and writes the data to the hot spare disk or new disk that replaces the faulty disk.

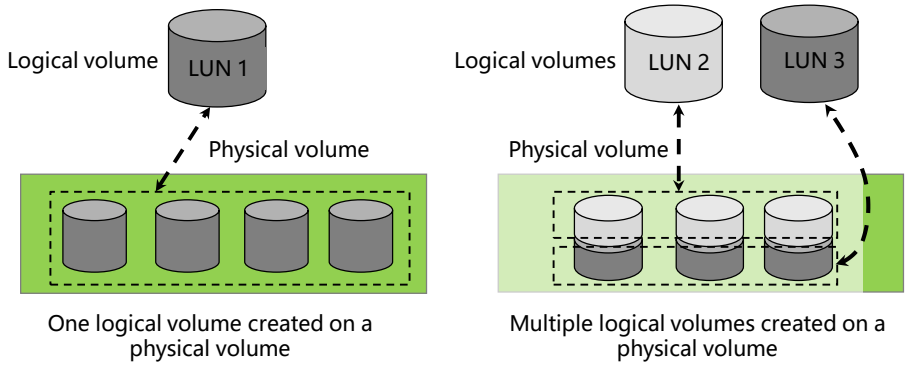


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Reconstruction is another data protection mode of RAID. Hot spare disks can automatically reconstruct failed disks to enhance data protection. Reconstruction does not affect other read and write operations of the RAID group at the same time, nor interrupt host services. The following conditions should be met for optimal reconstruction:
  - The hot spare disk is working properly and not used by other RAID groups.
- The RAID level used is RAID 1, 3, 5, 6, 10, or 50.

Relationship Between RAID and LUNs

- RAID is like a large physical volume composed of multiple disks.
- We can create one or multiple logical units of a specified capacity on the physical volume. Those logical units are referred to as LUNs. They are the basic block units that can be mapped to hosts.



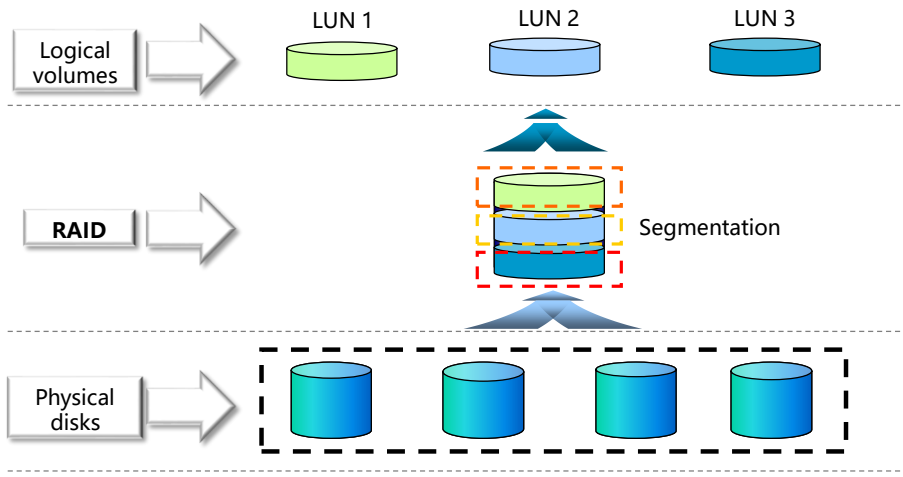
One logical volume created on a physical volume

Multiple logical volumes created on a physical volume

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Data is stored as files on volumes "visible" to the operating system. The Windows operating system uses the drive C:, F:, and others to indicate the volume in use. If the operating system is UNIX or Linux, mount points are used instead. The relationship between the drive letter (or mount point) and the physical disk is as follows:
  - Physical disks form a RAID group.
  - A RAID group has a specific RAID type associated with itself.
- A LUN is composed of (or a section of) storage capacity of a RAID group. LUNs are mapped to a host to serve as the storage space that can be used by the operating system.

Creating RAID Groups and Logical Volumes



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- For example:
  - There are four physical disks, each of which is 300 GB. If the four physical disks are placed in one RAID group, the total disk capacity is 4 × 300 GB = 1.2 TB. If the disk group is set to RAID 5 for data protection, the actual available space is 3 x 300 GB = 900 GB. The resources of one disk is "drained" because a quarter of the disk capacity is needed to store parity information.
  - A storage administrator can create one 900-GB LUN as a whole or divide it into multiple smaller LUNs.
- The data protection level of each LUN is RAID 5.

Contents

1. Data Management
2. Traditional RAID Technology
3. **Introduction to DAS Technology**
4. Introduction to SAN Technology
5. Introduction to NAS Technology
6. Huawei Distributed Storage
7. Object Storage

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

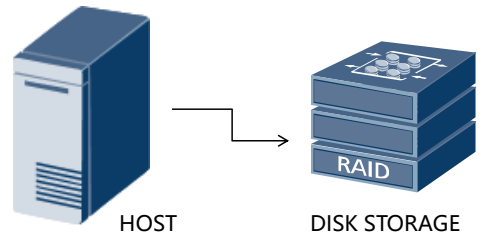
Building an ICT Infrastructure

- An ICT infrastructure is the physical solution that allows users to access the digital information they need.
- Components of an ICT infrastructure include:
  - Host: Computer systems for data access and storage
  - Software: Operating System software and application software
  - Network devices: Devices that provide channels for interconnection between devices.
  - Storage device: Device where user data is saved and can be quickly accessed.
  - Backup device: Separate device where backup data is saved.
- Three major infrastructural designs can be used: DAS, NAS, and SAN.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

DAS Concepts

- An ICT infrastructure is the physical solution that allows users to access information they require.
- The first ICT infrastructures were based on a simple concept we now refer to as "direct attached storage" (DAS).
- DAS definition: One or more dedicated storage devices connected to one or more servers. These storage devices provide block data access service for servers.
- Based on the location between storage devices and servers, DASs are classified into internal DASs and external DASs.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- DAS is an architecture where a storage device is directly connected to a server. It provides block-level storage services (not file system-level ones) for servers and is used for disks inside servers, tape libraries that are directly connected to servers, and external disk enclosures that are directly connected to servers. Internal DAS and external DAS can be used based on the location between storage devices and servers.
  - For internal DAS, a storage device is connected to a servers via the serial or parallel bus inside the server chassis. However, due to the distance limitation of the physical bus, internal DAS only supports short-distance, high-speed data transmission. Additionally, there are also limitations on the quantity of devices that can be connected to the internal bus, and storage devices can take up large amounts of space within the servers. This makes maintenance of other parts of the servers difficult.
- For external DAS, a server is directly connected to an external storage device. In most situations, they communicate to each other through FC or SCSI protocols. In comparison to internal DAS, external DAS has overcome the limitation of short distance and devices limit faced by internal DAS. Additionally, external DAS can even provide centralized management of the storage devices, making storage device management easier.

Benefits of DAS

- Ideal solution for local data supply
- High reliability
- Rapid deployment for small-scale environments
- Simple deployment
- Low complexity
- Small investment

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Benefits of DAS

- ATA (IDE) and SATA
  - Mainly used for internal DAS
- SCSI
  - Parallel SCSI (mainly used for internal DAS)
  - Serial SCSI (mainly used for external DAS)
- FC
  - A high-speed network interconnection technology

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The DAS protocols used for communication between the server and the storage device include IDE/ATA, SATA, SAS, SCSI, and FC protocols. These protocols require the disk controller support. A storage device is usually named by the protocol it supports.
- Integrated Drive Electronics, or IDE, is a disk drive that integrates a disk controller and a disk body. Integrating the disk body with the controller reduces the quantity and length of cables connected to the disk interfaces. This enhances data transmission reliability, makes disk manufacturing easier, and improves compatibility with different manufacturers. Generally, the IDE interface is also called the Advanced Technology Attachment (ATA) interface. This interface was developed by several companies, including Compaq and Western Data, in 1986, and was applied to desktop systems at the beginning of the 1990s.

DAS Management

- Internal DAS
  - The host provides:
    - Disk partition (volume management)
    - File system layout
  - DAS devices are managed through servers and Operating Systems.
- External DAS
  - Array-based management.
  - The TCO of management data and storage architecture is relatively low.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Challenges Faced by DAS

- Poor scalability
  - The number of ports that can be connected to the host is limited.
  - The number of addressable disks is limited.
  - Distance restrictions
- The system needs to be powered off while the internal DAS is being maintained.
- Difficulties with resource sharing
  - The array's front-end ports and storage space cannot be shared.
  - Resource silos: A DAS with insufficient storage space cannot share the remaining space of a DAS with surplus storage space.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- DAS has a poor scalability. The quantity of available ports on a DAS storage device is often few. Therefore, the number of hosts that can be connected to the storage device is limited. In addition, a DAS device has lower bandwidth and therefore its I/O performance is not high. If the host connected to the DAS device requires high I/O performance that reaches its upper threshold, the availability of the host service is affected. This also generates a chain reaction — that is, the performance of all hosts connected to the DAS device is affected.

Contents

1. Data Management
2. Traditional RAID Technology
3. Introduction to DAS Technology
4. **Introduction to SAN Technology**
5. Introduction to NAS Technology
6. Huawei Distributed Storage
7. Object Storage

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Ideal ICT Infrastructure

- Scalable in capacity
- Capable of being extended across the entire world
- Highly reliable
- Highest possible transportation speeds
- Easy to manage and flexible
- Interoperable with heterogeneous hardware

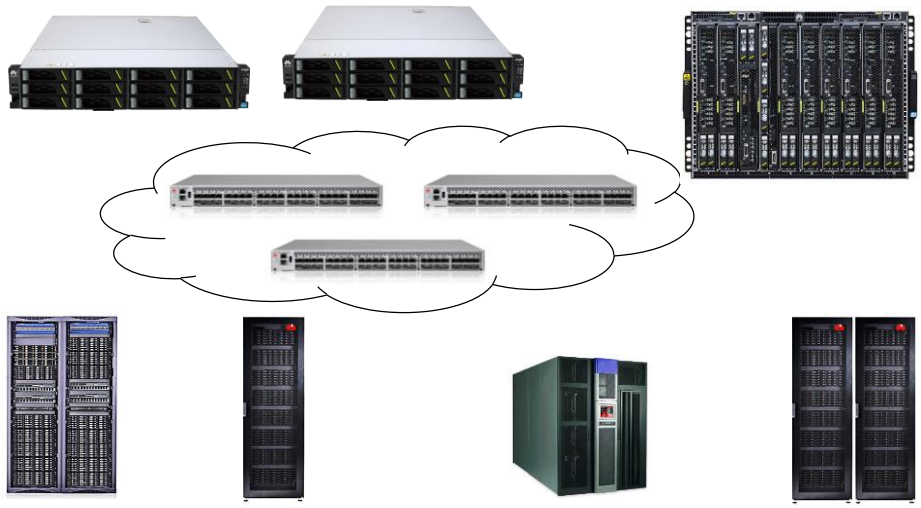






Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SAN Concepts



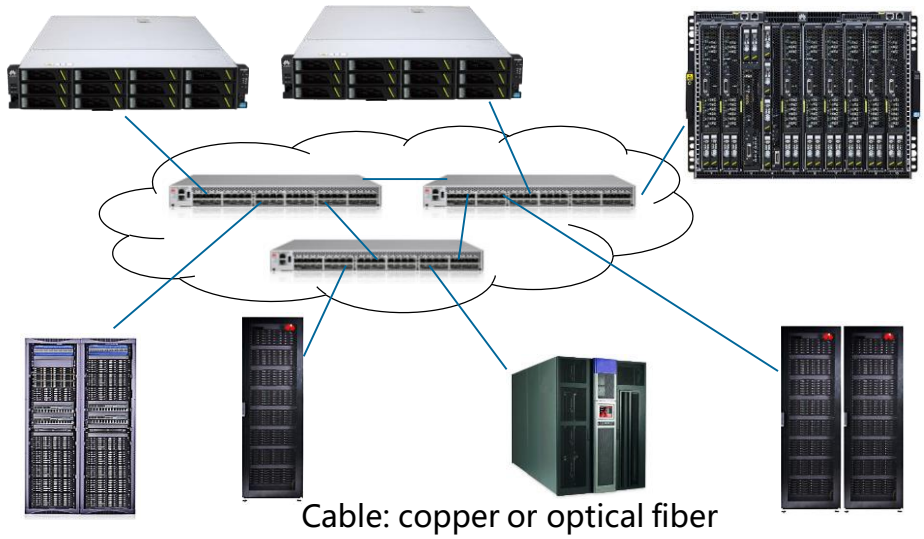
SAN Components: hosts, storage devices, and switches.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Storage Area Network (SAN) is a high-performance and dedicated storage network used between servers and storage resources. It is optimized for transferring mass original data. It can be considered as the advanced SCSI protocol for long-distance data transmission. SCSI and Fiber Channel (FC) are the typical SAN protocol suite. FC is especially suitable for this application, because it supports long-distance and large-block data transmission. SAN applies to high-end, enterprise-class storage applications, which have demanding requirements for performance, redundancy, and data availability. Components such as storage arrays and backup devices are all storage devices.

SAN Cabling

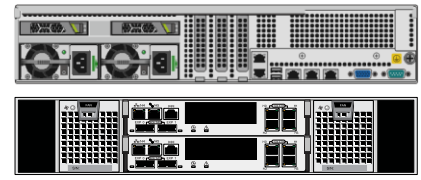


Cable: copper or optical fiber
Protocol: FC, iSCSI, or FCoE

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SAN Components

- Components in a SAN are:
  - Servers/Hosts
  - Interconnect devices:
    - switches; routers
  - Storage devices:
    - Disk arrays; backup devices
- Cables are used to connect them all together.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Differences Between DAS and SAN

Item	DAS	SAN
Protocol	SCSI protocol	Multiple protocols: FC, iSCSI, FCoE
Application scenario	Small- and medium-sized organizations that have a relatively small number of servers	Mid-range and high-end storage environments such as key databases, centralized storage, backup, and disaster recovery
Advantages	Easy deployment, small investment	High availability, performance, and scalability, powerful compatibility, centralized management
Disadvantages	Poor scalability, waste of resources, management difficulties, performance bottlenecks	Comparatively large investment

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

SAN Storage Applications

Centralized deployment of storage devices enables application servers to access and share data in a cost-effective manner.

Storage resources are divided into blocks that are mapped to application servers to implement storage resource sharing.



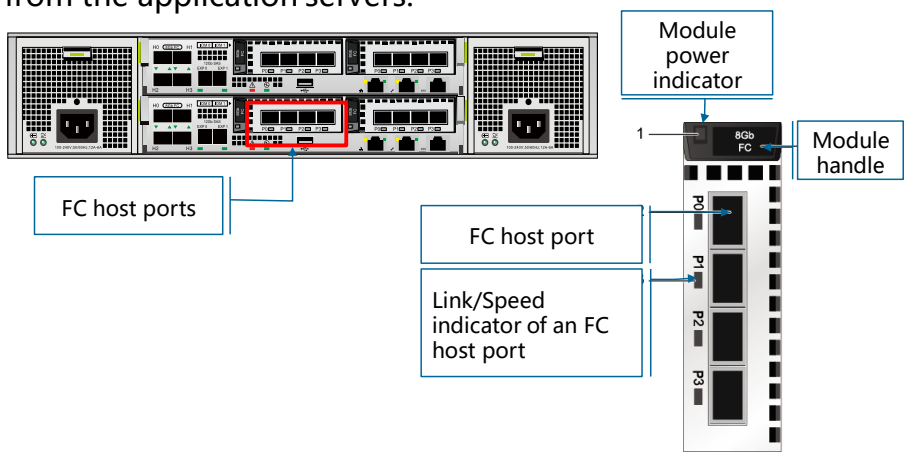
Data backup uses a SAN independent from the service network, making backup possible for diverse forms of data across heterogeneous servers.

SANs employ multiple mechanisms for automatic data backup, allowing data to be immediately recovered after a disaster occurs.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Device with FC Interface

- The Fibre Channel (FC) interface modules on a storage device provide service interfaces for connecting to application servers and receiving data exchange requests from the application servers.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FC Switch

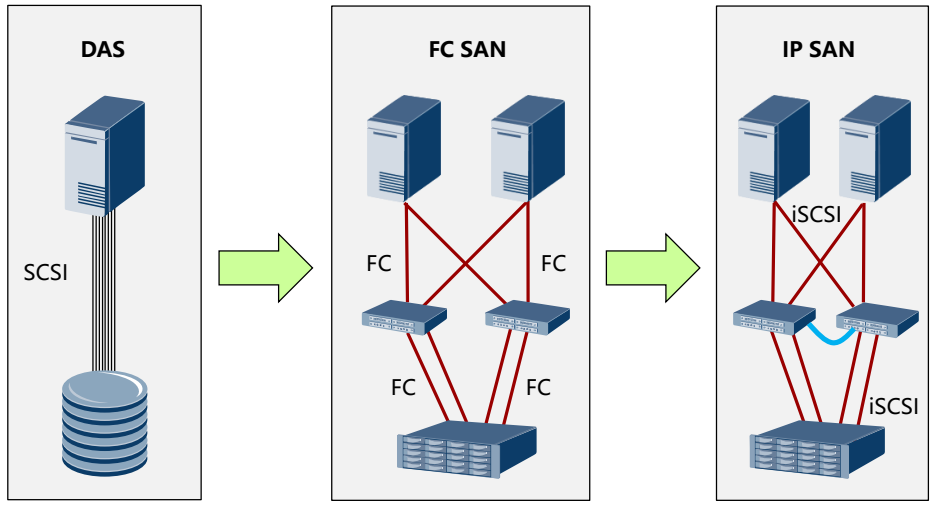
- Direct connection to an FC network.
- Direct connection to an initiator and a target.
- Exclusive use of all optical bandwidths.
- Switch-based zoning.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Like DAS, SAN transmits data in SCSI blocks from a storage device to a server. There is a significant difference between SAN and DAS — DAS cables support up to a 25-meter transmission distance, while SAN cables supports connection up to hundreds or thousands of kilometers.
- SAN encapsulates SCSI blocks (user data or load data) into a data packet or frame and transmits the data packet to a longer distance.
- There are many other approaches (that is, protocols) to send SCSI blocks to a cross-SAN connection. Each protocol describes the transmission mode for processing SCSI blocks in a unique way.
- The following three protocols are commonly used in a SAN architecture:
  - FC: The SAN architecture using this protocol is called FC SAN.
  - Internet SCSI: The SAN architecture using this protocol is called IP SAN.
  - FCoE: Fibre Channel over Ethernet.
- The FC and iSCSI protocols are used in the modern SAN architecture, while the FCoE protocol is used more often when the server needs to integrate SAN and LAN services.

FC SAN and IP SAN

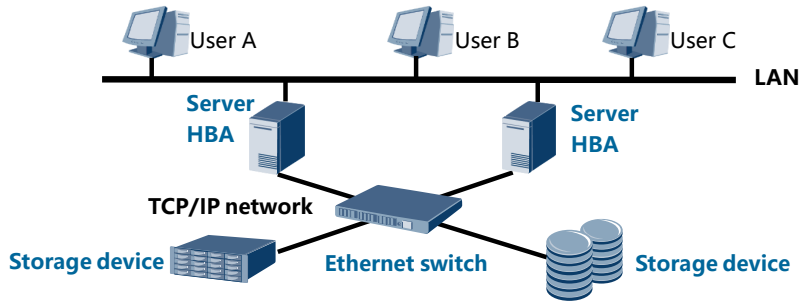


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

- FC switches support direct connection through FC network routes. Initiators can directly connect to targets using the routing software on the FC switches, exclusively using all available bandwidth.
- An FC switch is the SAN core that connects a host to a storage device. FC switches are classified into entry, workgroup, and core levels. Workgroup FC switches are mostly adopted in small-scale SAN. A large-sized and virtual switch that supports distributed processing and long-distance transmission can be set up by cascading several workgroup FC switches. Core FC switches (also called directors) are located at the center of a large-scale SAN and support the non-FC protocols such as InfiniBand, along with advanced FC services such as security, trunk, and frame filtering. Core FC switches usually use blade-based, hot-swappable circuit boards.
- FC switches using the FC protocol are used to build FC networks, while Ethernet switches using the TCP/IP protocol are used to build Ethernet networks.
- To isolate some devices, FC switches offers the zoning function. Similar to VLAN of Ethernet switches, zoning logically group devices (hosts and storage devices) in a SAN to different zones. Devices in different zones cannot communicate with each other.

What Is an IP SAN?

- An IP SAN is an approach to using the Internet Protocol in a storage area network, usually over Gigabit Ethernet.
- The typical protocol that implements an IP SAN is Internet SCSI (iSCSI), which defines the encapsulation mode of SCSI instruction sets in IP transmission.



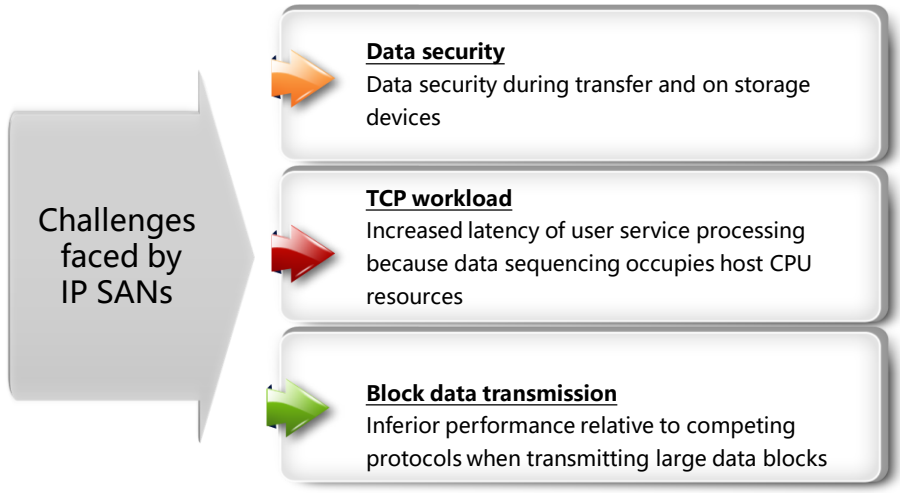
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Advantages of IP SANs

Standard access	• IP SANs do not need dedicated HBAs or FC switches. Common NICs and switches for connecting storage devices to servers can be used.
Long transmission distance	• IP SANs are available wherever IP networks exist, and IP networks are now the most widely used networks in the world.
Enhanced maintainability	• Enterprises using IP SANs can draw on both the large pool of professionals skilled in IP network technologies, and a range of powerful IP network maintenance tools.
Scalable bandwidth	• With the development of the 10 Gbit/s Ethernet, a single port can now support transmission of 10 Gbit/s over IP SAN.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Challenges Faced By IP SANs



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

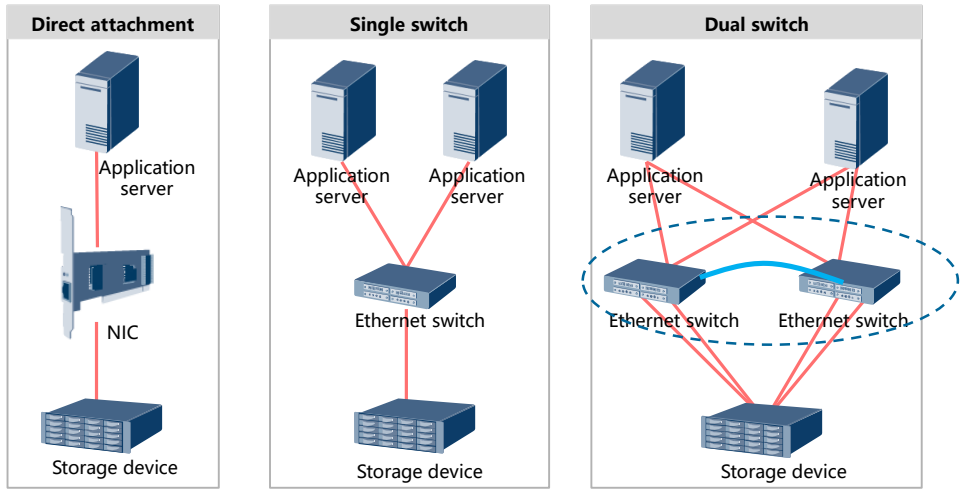
- Because IP SAN requires only a small amount of common hardware for network configuration, it is much cheaper than FC SAN. Most hosts have been configured with proper NICs and switches. These NICs and switches are also suitable for iSCSI-based transmission. High-performance IP SAN requires dedicated iSCSI Host Bus Adapters (HBAs) and high-end switches.

FC SAN vs. IP SAN

Indicator	Fibre Channel SAN	IP SAN
Transmission speed	4 Gbit/s, 8 Gbit/s, 16 Gbit/s	1 Gbit/s, 10 Gbit/s, 40 Gbit/s
Network architecture	Dedicated FC networks and Host Bus Adapters (HBAs)	Existing IP networks
Transmission distance	Limited by the maximum transmission distance of optical fibers	Unlimited, in theory
Management and maintenance	Complicated technologies and management	As simple as operating IP devices
Compatibility	Poor	Compatible with all IP network devices
Performance	High transmission and read/write performance	1 Gbit/s (mainstream) and 10 Gbit/s
Cost	High purchase cost (of FC switches, HBAs, FC disk arrays, and so on) and maintenance cost (of staff training, system configuration and supervision, and so on)	Lower purchase and maintenance costs and higher Return On Investment (ROI) than FC SANs
Disaster recovery	High hardware and software costs for disaster recovery (DR)	Local and remote DR available on existing networks at a low cost
Security	High	Medium/Low

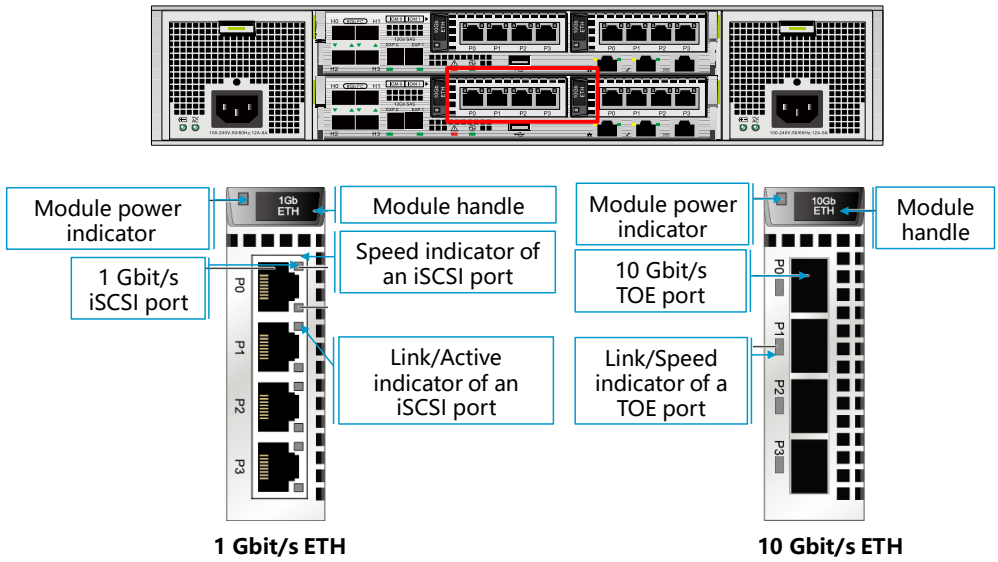
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Networking in IP SANs



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Huawei IP SAN Storage Devices

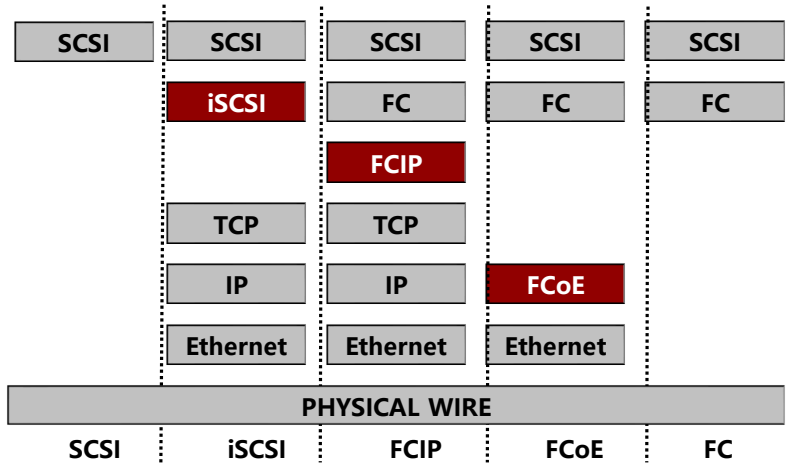


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Typical IP SAN networking modes are direct connection, through an Ethernet switch, or dual-switch:
  - Direct connection: The host connects to the storage device through a NIC, TOE card, or iSCSI HBA. This networking mode is simple and cost-effective, but the storage resources cannot be sufficiently shared by other hosts.
  - Through an Ethernet switch: The host that has a NIC, TOE card, or iSCSI HBA installed connects to the storage device through an Ethernet switch. This networking mode features high scalability and allows multiple hosts to share the storage resources offered by the same storage device. However, once the switch fails, the storage resources become unavailable.
- Dual-switch: A host connects to a storage device through more than one path. This networking mode features high scalability and prevents a single point of switch failure.

FC and TCP Convergence

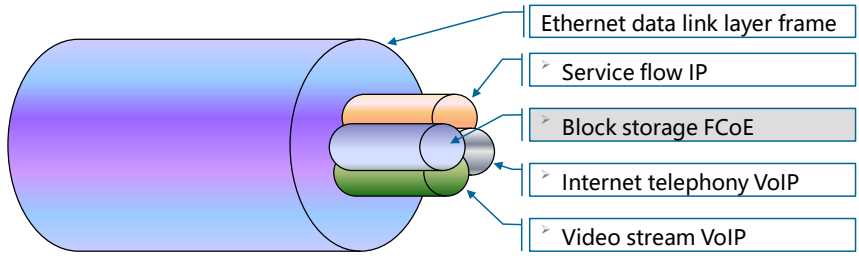
- Ethernet and FC technologies are both developing quickly. IP SANs and FC SANs currently coexist and will continue to serve as complements to each other for the foreseeable future.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FCoE Protocol

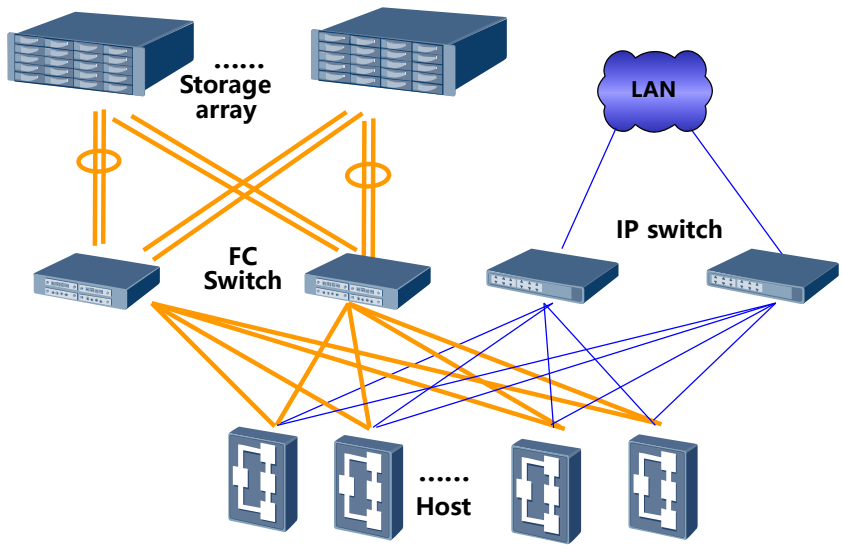
- The FCoE protocol is used to transmit FC signals over a lossless enhanced Ethernet.
- FCoE encapsulates FC data frames into Ethernet packets and allows service traffic on a LAN and a SAN to be concurrently transmitted over the same physical interface.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Convergence in this slide indicates that a system uses the FC and iSCSI protocols at the same time. The first way encapsulates iSCSI data packets into FC frames, and the other encapsulates FC frames into Ethernet frames.
- FC and TCP are converged in various ways. Fibre Channel over Ethernet (FCoE) is most widely used. FCoE is an increasingly popular technology that allows simultaneous use of FC and Ethernet technologies. In fact, FCoE, as a low-cost solution, can use one switch to transmit FC and IP information at the same time.

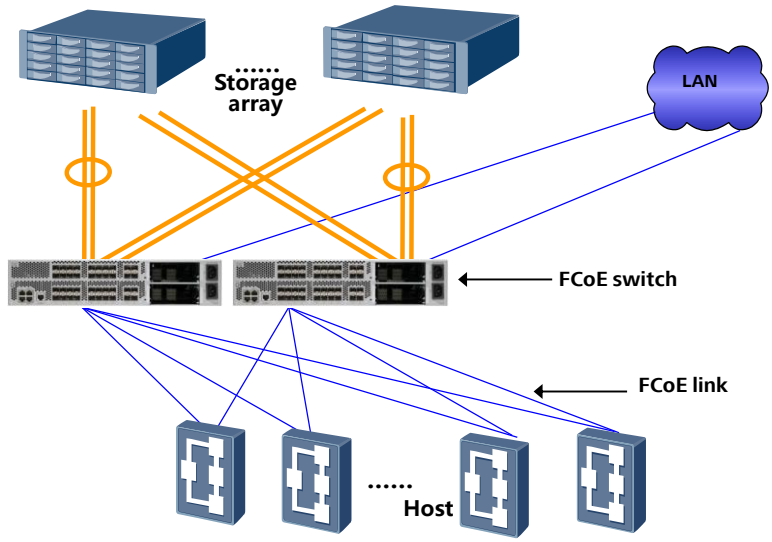
Networking Before the Use of FCoE



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- FCoE offers standard FC services, including discovery, global naming, and zoning. These services run in the same way as the original FC services with low latency and high performance.
- FCoE bears FC on a new type of link, that is, Ethernet Layer 2 link. Note that the Ethernet must be enhanced lossless Ethernet to meet the link-layer transmission requirements of the FC protocol.
- VoIP stands for Voice over IP. It is a method for transmitting audio and video digital technologies over the Ethernet.

Networking Simplified By the Adoption of FCoE



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Data Management
2. Traditional RAID Technology
3. Introduction to DAS Technology
4. Introduction to SAN Technology
5. **Introduction to NAS Technology**
6. Huawei Distributed Storage
7. Object Storage

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

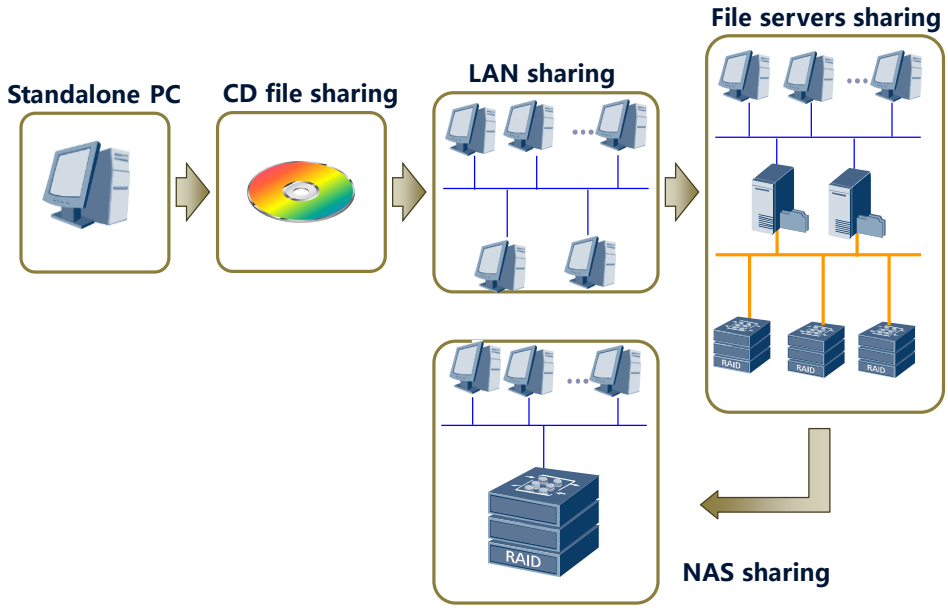
- The FCoE switch sends FC SAN traffic to the storage array and transmits Ethernet service data to the LAN client. When FCoE is used, networking is simplified, and a quantity of network devices is reduced. Plus, cooling, management, and overall maintenance costs of a data center are reduced.

File Sharing Environment

- The file system is used to store and organize data structures.
- File sharing
  - Data is accessed through network storage.
  - The file system can be used only after being mounted.
- In the traditional C/S model, remote file sharing is implemented through the file sharing protocol.
  - FTP
  - DFS

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

File Sharing Technology Evolution

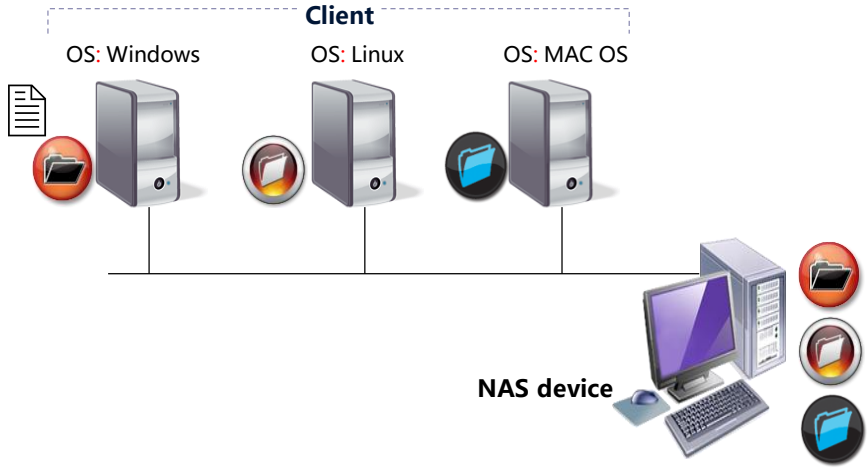


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- File sharing involves data storage and access. In a file sharing environment, a user who creates a file can specify other users' read, write, execute, add, delete, and list permissions and control file changes. When files are shared, a protection method is required to maintain data integrity if multiple users access the same file at the same time. The C/S model uses the file sharing protocol and Distributed File System (DFS). The following are some instances:
  - Although FTP can transmit data on the network, it is a standard file transfer protocol. The TCP protocol is used for transmitting data between the server and client. FTP data is not encapsulated during transmission, so data transmission is not secure. FTP over Secure Shell (SSH) adds security specifications to the FTP protocol.
- DFS can be distributed on multiple hosts. Any host can access the entire file system. DFS provides efficient data management and data security assurance.

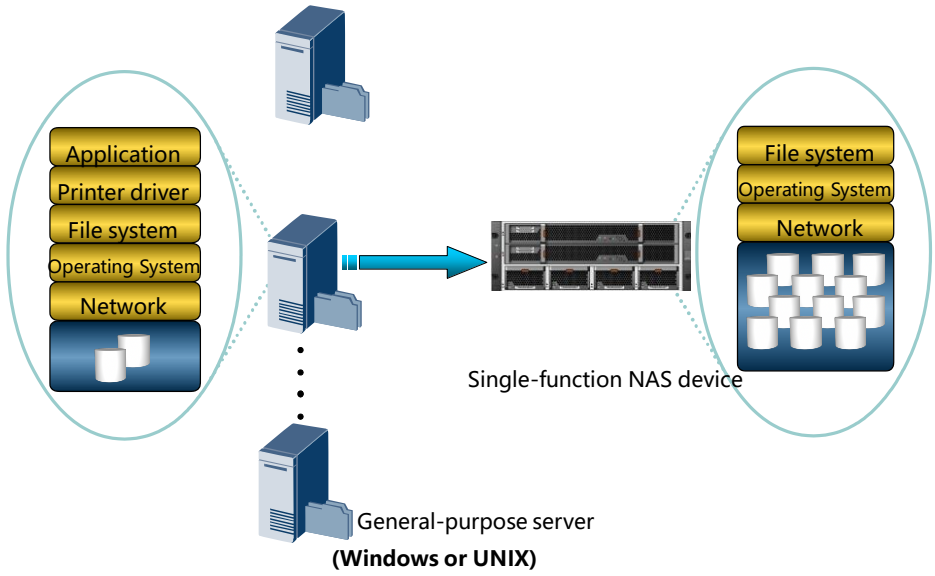
What Is NAS?

- NAS is a storage device shared through the network.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Universal Servers and NAS Devices



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A Network Attached Storage (NAS) is an IP-based file sharing device connected to a LAN. It provides storage resources through file-level data access and sharing, enabling customers to quickly share files with the minimum storage management overhead. You do not need to set up multiple file servers. NAS also helps eliminate bottlenecks when users access common servers, and uses network and file sharing protocols to archive and store data. These protocols include TCP/IP for data transmission and CIFS and NFS for remote file services.
- The two common data sharing modes are Network Access Server (NAS) and File Transfer Protocol (FTP). For NAS on UNIX systems, use Network File System (NFS), and use Common Internet File System (CIFS) for Windows. NAS can be expanded to offer high performance and reliability demanded by enterprises to access data. NAS devices are dedicated file services and storage systems with high performance, high speed, and single purpose. NAS clients communicate with servers through IP networks. Most NAS devices support multiple interfaces and networks. . An NAS device executes file I/O better than a general purpose server and can connect to more clients than a traditional server. Therefore, a NAS device can integrate traditional servers.

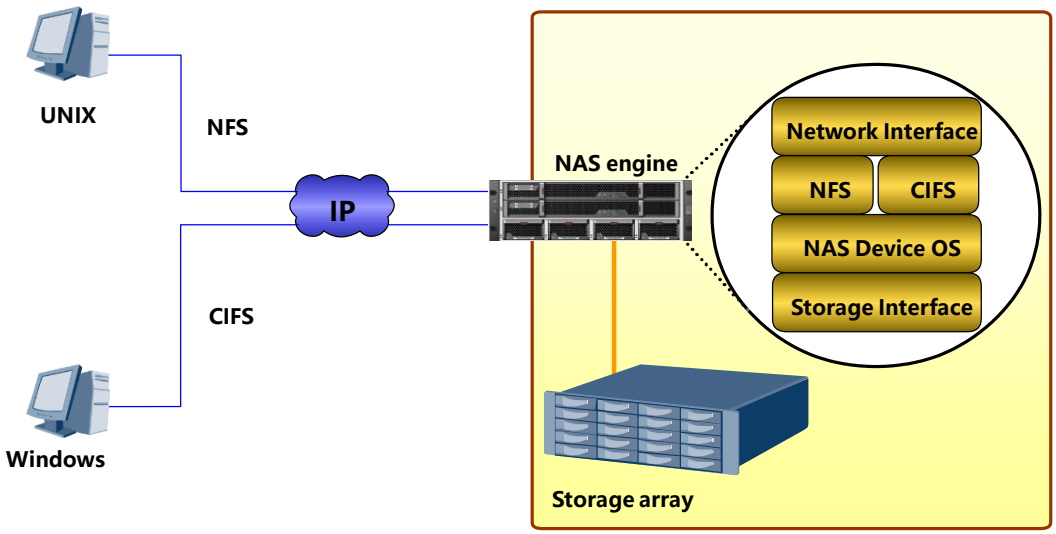
Highlights of NAS

- Information can be obtained in a comprehensive manner
- Improved efficiency
- Improved flexibility
- Centralized storage
- Simplified management
- Higher scalability
- High availability achieved by using local clusters
- Security integration environment (using user authentication)

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- NAS devices are optimized based on common servers, in terms of file service functions, storage, retrieval, and access to application client files.
- As shown in the figure, a general-purpose server runs a general-purpose operating system for any app. A NAS device is dedicated to file services and provides file sharing services for other operating systems by using open standard protocols. To improve high availability of NAS devices, some NAS vendors also support the NAS cluster function.

NAS Components



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- NAS provides the following advantages:
  - Full acquisition of information: NAS implements efficient file sharing and supports many-to-one and one-to-many configuration. The many-to-one configuration allows the NAS device to be accessed by multiple clients at the same time. One-to-many configuration allows a client to connect to multiple NAS devices at the same time.
  - Improved efficiency: NAS devices provide dedicated operating systems to provide file services, reducing file service operations on common file servers, improving the usage of common servers, and eliminating the bottleneck of file access by common servers.
  - Improved flexibility: NAS uses industry standard protocols and is compatible with UNIX and Windows clients. Different types of clients can provide services flexibly when accessing the same resource.
  - Centralized storage: Centralized data storage reduces duplicate data on client workstations, simplifies data management, and provides higher data protection.
  - Simplified management: A console centrally manages file systems in an effective way.
  - Strong scalability: Provides expansion with high performance and low latency based on different utilization configurations and service applications.
  - High availability: NAS devices can use the local cluster technology for failover. NAS provides multi-connection options with redundant network components. The replication function and recovery options are comprehensive, helping to achieve high data availability.
- Security: Data security is ensured by authenticating users, locking files, and using the industry standard security architecture.

NFS File Sharing Protocols

- Two universal NAS file sharing protocols:
  - CIFS – Common Internet File System protocol
    - The file sharing protocol in the traditional Microsoft environment is based on the Server Message Block Protocol (SMB protocol).
  - NFS – Network File System protocol
    - Traditional UNIX environment file sharing protocol.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A NAS device consists of the following components:
  - NAS engine, such as a CPU and memory
  - One or more network interface cards (NICs) — gigabit Ethernet (GE) and 10GE network adapters — for network connections
  - An operating system optimized for NAS function management
  - NAS file sharing protocols, CIFS and NFS
  - Disk resources using industry standard storage protocols, such as ATA, SCSI, and FC
- The NAS environment includes the clients that access NAS devices through IP networks that use standard protocols.

About NFS

- C/S applications
- TCP-based Remote Procedure Call (RPC) mechanism
- Accesses the remote file system through the mount point of the local file system
- Accesses the mount point under permission control

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

NAS File Sharing: CIFS

- Universal Internet file system
  - Developed by Microsoft in 1996
  - Enhanced version of the SMB protocol
  - Status protocol
    - When the client application has the connection recovery function, the connection can be automatically restored and the interrupted file can be opened again.
    - If the client application does not have the automatic connection function, the user must reestablish the CIFS connection.
  - Application/Presentation layer working in the OSI model
  - It is commonly used in the Microsoft Operating System but is not related to the platform.
  - CIFS runs on TCP/IP and uses Domain Name Service (DNS) for name resolution.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- NFS is a client/server application that uses Remote Procedure Call (RPC) to communicate between computers. Users store and update files on the remote NAS device, just like on their own computers.
- The user's system requires an NFS client to connect to an NFS server. The NFS server and client use TCP/IP to transfer files. Therefore, TCP/IP must be installed on both the server and client.
- A user or system administrator can leverage NFS to mount all file systems, along with a part of any directory or subdirectory tree. The mounted file systems are controlled with access permissions (for example, read-only or read-write).

Comparison Between CIFS and NFS

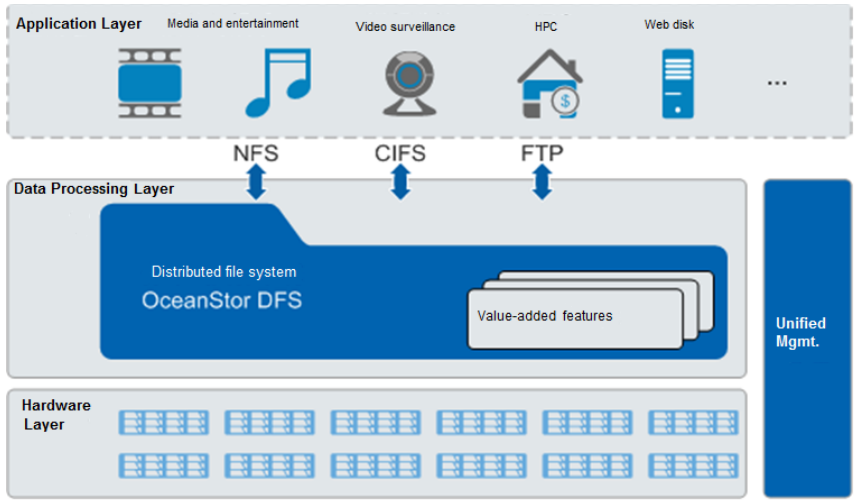
- If a file system is already set to:
  - CIFS share, the file system can subsequently be set to read-only NFS share.
  - NFS share, the file system can subsequently be set to read-only CIFS share.

Protocol	Transmission Protocol	Client	Fault Impact	Efficiency	Supported Operating Systems
CIFS	TCP/IP	Integrated operating system without the need for additional software	Large	High	Windows
NFS	TCP or UDP	Additional software required	Small: The interaction process can be automatically resumed.	Low	Unix

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Example of Unified NAS

- Huawei OceanStor 9000



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- CIFS is a network-based sharing protocol that has high requirements on network transmission reliability. Therefore, CIFS normally uses TCP/IP. NFS is used for independent transmission, so it uses TCP or UDP.
- For NFS the client must be equipped with dedicated software. CIFS is integrated into the operating system, without additional software support.
- NFS is a stateless protocol and CIFS is a stateful protocol. NFS can automatically restore connection once the connection fails. However, CIFS cannot. CIFS provides low redundancy, so it has higher transmission efficiency than NFS.
- Both protocols need to convert the file format.

Contents

1. Data Management
2. Traditional RAID Technology
3. DAS Technology Introduction
4. SAN Technology Introduction
5. NAS Technology Introduction
6. **Huawei Distributed Storage**
7. Object Storage

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OceanStor 9000 is also a unified NAS system. However, unlike a dedicated storage architecture, OceanStor 9000 uses multiple universal x86 servers to build a NAS cluster, with powerful performance, capacity, and scalability.

About Server SAN

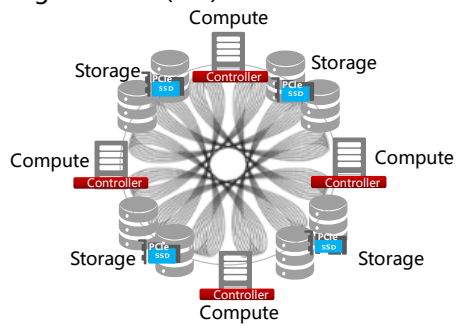
- Concept
  - Server SAN is a storage resource pool that consists of storage units on multiple independent servers. It incorporates both compute and storage resources.
- Features
  - Purpose-built devices converted to general-purpose ones
  - Linear scalability for compute and storage
  - Simple management and low TCO

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Decoupled from the vendor's dedicated hardware: A traditional storage system highly integrates software and hardware. A Server SAN product is decoupled from hardware. Therefore, it is not bound to specific hardware.
- Convergence of storage and compute: Server SAN is built on universal x86 servers, and integrates compute and storage.

Huawei Server SAN Family: FusionStorage

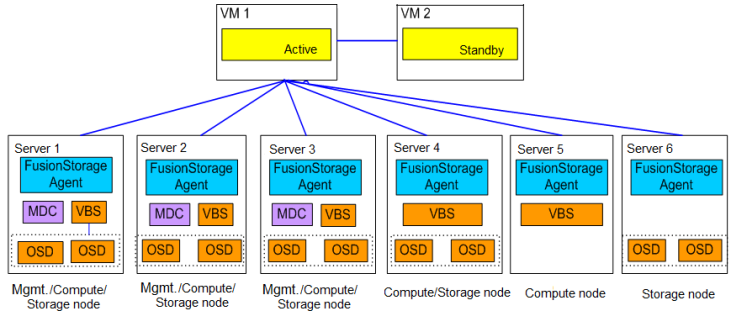
- Distributed block storage software
- Organizes local storage media, such as Hard Disk Drives (HDDs) and Solid-State Drives (SSDs), of general x86 servers into a large-scale storage resource pool using distributed computing technologies
- Provides industry standard SCSI and iSCSI interfaces for upper-layer applications and Virtual Machines (VMs) in non-virtual environments.
- Open Application Programming Interface (API)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Typical use cases of traditional block storage: Various service applications, such as SQL, Oracle RAC, Web, along with industry applications.
- Integration with main cloud platforms, such as Huawei FusionSphere, VMware vSphere, and open-source OpenStack; storage resources are allocated on demand.
- Huawei's FusionStorage is the only PB-scale Server SAN product in commercial use.

FusionStorage: Logical Architecture



- FusionStorage Manager (FSM): FusionStorage management module. It provides O&M functions including alarm management, service monitoring, operation logging, and data configuration. In most cases, FSM-enabled nodes are deployed in active/standby mode.
- FusionStorage Agent (FSA): Performs FusionStorage agent functions. It is deployed on each node to enable the node to communicate with the FSM module. An FSA consists of three processes: MetaData Controller (MDC), Virtual Block System (VBS), and Object Storage Device (OSD). Depending on system configuration requirements, different combinations of processes are used for different nodes to implement specific functions.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- MetaData Controller (MDC): Controls the status of distributed clusters, data distribution rules, and data reconstruction rules. MDCs are deployed on ZooKeeper disks of three nodes to form an MDC cluster.
- Virtual Block System (VBS): Manages volume metadata and provides the access point service for distributed clusters so that compute resources can access distributed storage resources through the VBS. A VBS process, by default, is deployed on each node so that these VBS processes can form a cluster. Multiple VBS processes can be deployed on a single node to improve I/O performance.
- Object Storage Device (OSD): Performs I/O operations. Multiple OSD processes are deployed on each server. By default, one disk corresponds to one OSD process. When Solid-State Drives (SSDs) serve as the main storage, multiple OSD processes can be deployed on one SSD to maximize the SSD usage and performance. For example, one 2.4 terabyte SSD supports a maximum of six OSD processes, and each OSD process manages 400 gigabytes of SSD space.

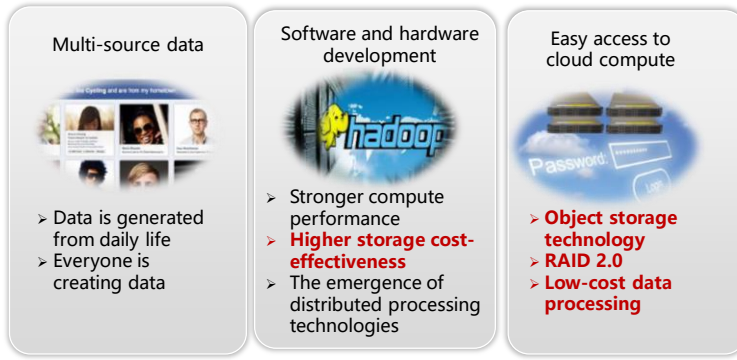
Contents

1. Data Management
2. Traditional RAID Technology
3. DAS Technology Introduction
4. SAN Technology Introduction
5. NAS Technology Introduction
6. Huawei Distributed Storage
7. **Object Storage**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

About Object Storage

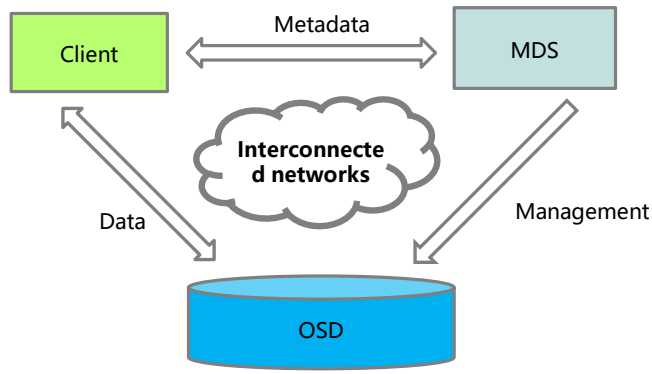
- For enterprises, big data analysis is becoming increasingly important. As big data applications employ ever larger data volumes, enterprises demand larger storage capacity. The primary objective is to store large quantities of data.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Multi-source data
  - Today's data comes from more fields than before. The GPS system installed on cars and people's mobile phones are all data sources.
- Software and hardware development
  - The compute speed of computers increases every year according to Moore's Law. Storage media density soars so the storage capacity cost-per-unit decreases year by year.
  - A large amount of unstructured data is a major feature that distinguishes big data from traditional data. Traditional relational databases cannot process the rapidly increasing unstructured data. The Hadoop-based distributed data processing technology makes unstructured data processing impractical.
- Easy access to cloud compute
- Big data needs to be processed on a powerful software and hardware platform. Building such a platform is expensive. Some small enterprises may not be able to bear the expense. The IaaS and PaaS services of cloud compute can provide pay-on-your-terms software and hardware platform services for enterprises. These services enable small enterprises to process big data.

Object Storage: Technical Structure



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In an object storage system, the MetaData Server (MDS) is responsible for mappings between files and Object Storage Devices (OSDs), as well as organizing files and directories. The MDS provides all name domain operations of file systems, including file search and creation, along with file and directory property processing. In terms of clients, the MDS acts as a logical file window, while the OSD is a physical file window. When you retrieve a file, the file system obtains the address where the file is stored from the MDS and accesses the file on the OSD. There is no need to access the MDS for the follow-up I/O operations. This reduces loads of the MDS and enables system scalability.

Object Storage: Composition

- Object
  - It contains file data and related attribute information for self-management. The size of objects, including entire data structures such as files and database entries, varies.
- OSD
  - Each OSD is an intelligent device that has its own storage media, processor, memory, and network system. As the core of an object storage system, the OSD manages local objects.
- File system
  - It runs on the client to transfer file system requests from applications to the MDS and OSD. The MDS provides metadata and cache consistency services.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. What do RTO and RPO refer to?
2. What is the relationship between RAID and LUN?
3. What are the differences between DAS, SAN, and NAS?

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- What do RTO and RPO refer to?
  - Recovery Point Object (RPO):When a disaster or emergency occurs, data can be restored to a previous time point. This in turn determines how much data is lost when a disaster occurs
  - Recovery Time Objective (RTO):Refers to the length of time it takes for a system to resume normal operation after a breakdown.
- What is the relationship between RAID and LUN?
  - RAID is like a large physical volume composed of multiple disks.
  - We can create one or multiple logical units of a specified capacity on the physical volume. Those logical units are referred to as LUNs. They are the basic block units that can be mapped to hosts.
- What are the differences between DAS, SAN, and NAS?
  - DAS:One or more dedicated storage devices connected to one or more servers. These storage devices provide block data access service for servers.
  - SAN:Storage Area Network (SAN) is a high-performance and dedicated storage network used between servers and storage resources. It is optimized for transferring mass original data. It can be considered as the advanced SCSI protocol for long-distance data transmission.
  - NAS:A Network Attached Storage (NAS) is an IP-based file sharing device connected to a LAN. It provides storage resources through file-level data access and sharing, enabling customers to quickly share files with the minimum storage management overhead.

Summary

- Information data management
- Traditional RAID technology
- Introduction to DAS technology
- Introduction to SAN technology
- Introduction to NAS technology
- Introduction to object storage

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei official websites:
  - Enterprise business: https://e.huawei.com/en/
  - Technical support: https://support.huawei.com/enterprise/en/index.html
  - Online learning: http://learning.huawei.com/en/
- Document tool:
  - HedEx Lite
- Simulator:
  - eNSP

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei Learning website:
  - https://support.huawei.com/learning/en/newindex.html
- Huawei Support Case Library:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



DC Cloud Computing Basics

Huawei Data Center Series of Courses

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This course introduces data center (DC) cloud computing from the perspectives of DC development, cloud computing development, and cloud computing advantages and deployment modes. After learning this course, you are supposed to have a clear understanding of the entire framework of DC cloud computing.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- On completion of this course, you will be able to know:
  - DC development and evolution
  - Virtualization of cloud computing
  - Benefits of cloud computing
  - Deployment modes of cloud computing

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

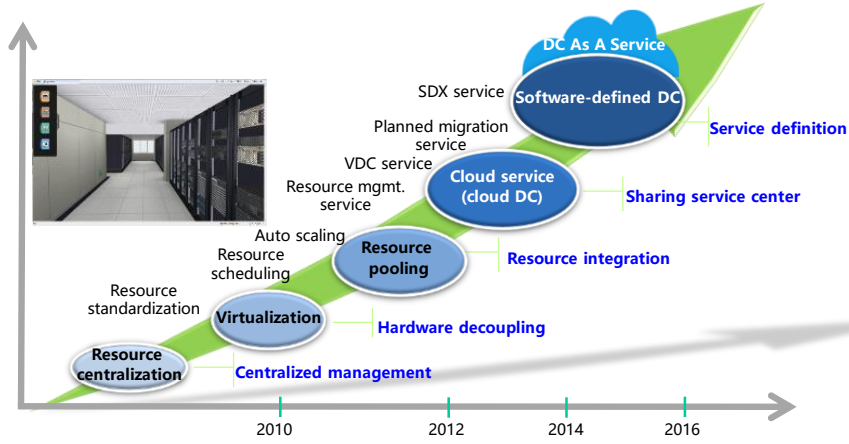
Contents

1. **DC Development**
2. Cloud Computing Development
3. Benefits of Cloud Computing
4. Deployment Modes of Cloud Computing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

DC Development

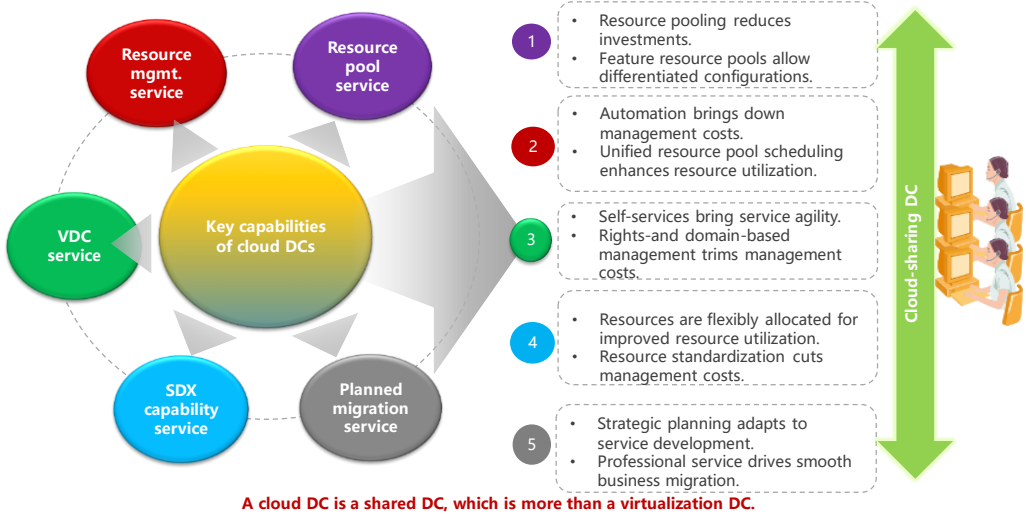


- **Service capability is a key indicator for weighing DC maturity. Cloud DCs are the future of DCs.**

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A DC involves three types of hardware: server, storage, and network. All of them are defined by software. A software-defined DC is basically an application-defined DC for offering all user-demanded applications.

Key Capabilities of Cloud Computing DCs

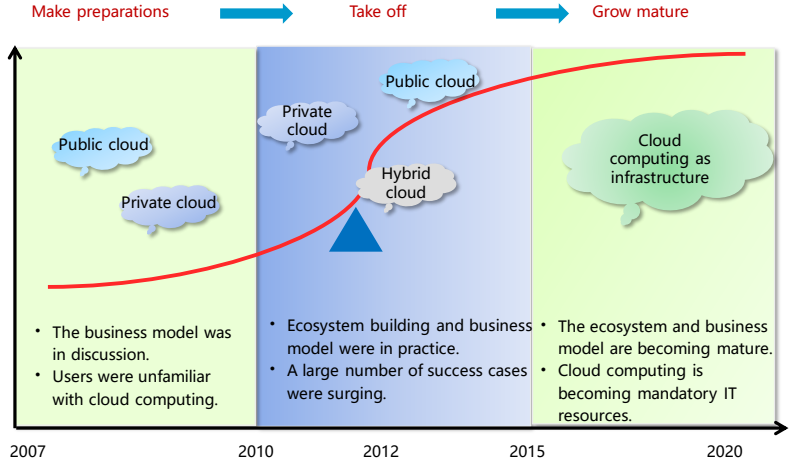


A cloud DC is a shared DC, which is more than a virtualization DC.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

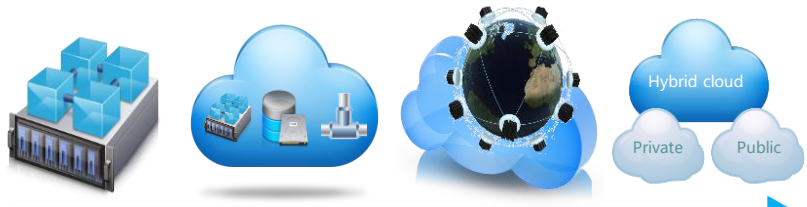
- This slide introduces key cloud DC capabilities: resource pooling, unified resource management, flexible resource definition (SDX, such as SDS and SDN. For details, see the remarks on slide 5), rights- and domain-based management (physical resource pools can be divided into VDCs by department or enterprise), and scheduled service migration.

Rapidly Developing Cloud Computing Is a Preferred Solution to IT Construction



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

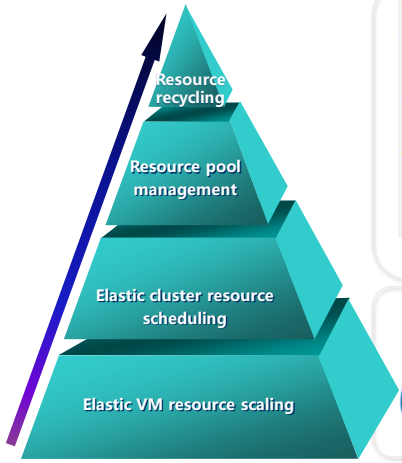
Cloud Computing Development

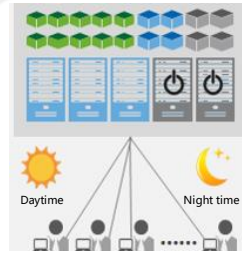


Virtualization	Private cloud	Multi-DC consolidation	Hybrid cloud
• Computing virtualization • Storage virtualization • Network and security virtualization	• Automatic management • Elastic resource scheduling • Large cluster-based HA and DRS	• Multi-level backup and DR • SDN virtualization	• OpenStack/AWS APIs • Hardware/Virtual resource pool management • Software-defined DCs

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Private Cloud: Automation Helps Simplify Management





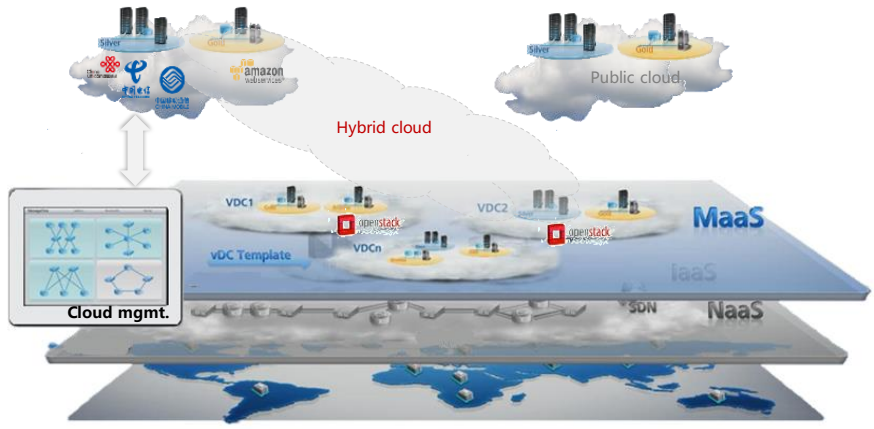
- Resource recycling: Elastic scaling and time-based resource reuse of application resources achieve resource sharing among applications.
- Resource pool management: Resource pool quota management and priority management ensure that prioritized VIP applications always obtain required resources.
- DPM: Energy management policies, VM migration, and automatic power-off of idle servers improve energy efficiency.



- DRS: Automatic VM migration and load balancing policies ensure high computing performance.
- Storage DRS: Automatic storage migration ensures high I/O performance.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

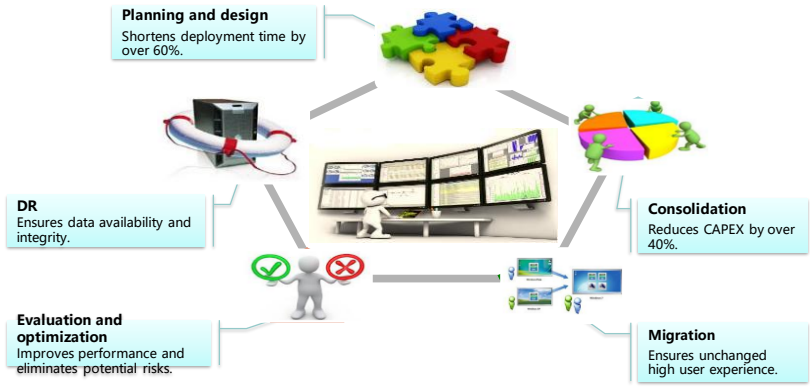
Hybrid Cloud: Makes IT as a Service



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Cloud DC Service System

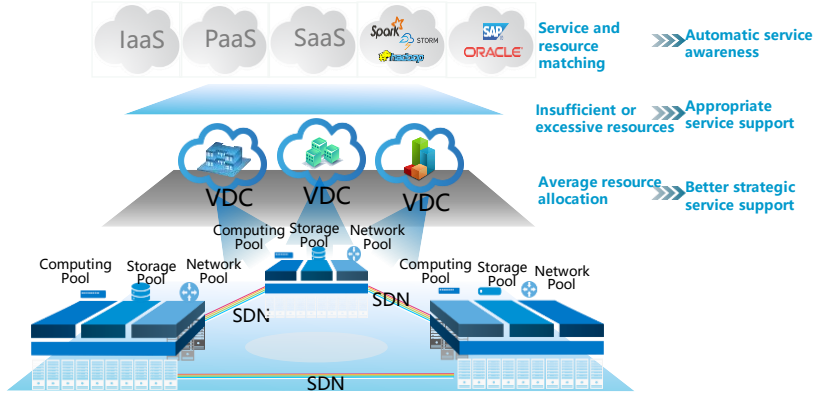


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Service-driven Distributed Cloud DC Allows On-Demand IT Resource Usage

Service-**D**riven **D**istributed **C**loud **D**ata **C**enter
SD-DC²: rebuilding DCs based services



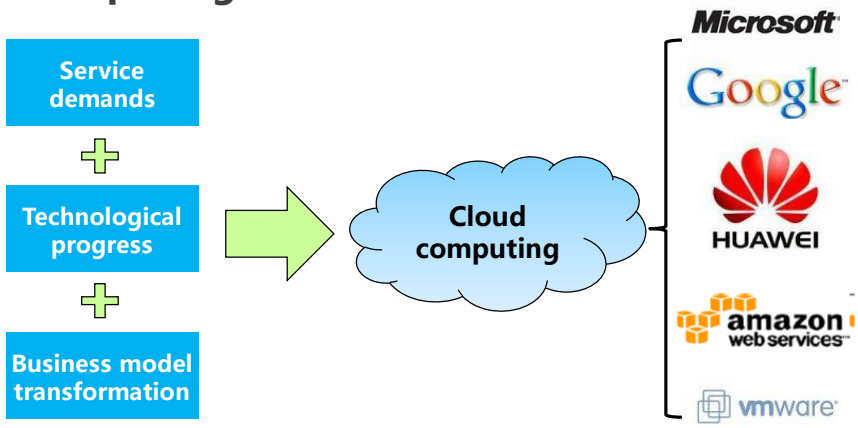
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

Contents

1. DC Development
2. **Cloud Computing Development**
   - Concept
     - Virtualization
3. Benefits of Cloud Computing
4. Deployment Modes of Cloud Computing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

What Contributes To the Emergence of Cloud Computing?



Service demands, technological progress, and business model transformation contribute to the emergence of cloud computing.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Service demands, technological progress, and business model transformation contribute to the emergence of cloud computing.
- Service demands
  - Governments and enterprises are keen for digitalization that requires low costs but generates high performance.
  - Individual users have strong requirements for Internet and mobile Internet applications and constantly pursue compelling user experience.
- Technological progress
  - Virtualization, distributed and parallel computing, and mature Internet technologies make it possible to provide Internet-driven IT infrastructures, development platforms, and software applications.
  - As broadband technology and users develop, the Internet-based service mode becomes the mainstream.

IT Development History



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Mainframe centralized mode: processes large-scale applications centrally and is confined to DCs.
- Client/Server mode: enables applications to be no longer restricted by DCs but to be used by common users.
- PC: Personal Computer.
- Cloud computing mode: responds to the explosive growth of information and the demand for a dynamic flexible architecture.

What Is Cloud Computing?

- National Institute of Standards and Technology (NIST): Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
- Wikipedia: Cloud computing is a type of computing technology based on Internet. By using this technology, users can provide shared software and hardware resources and information for computers and other devices on demand. It relies on sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over a network.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Key Characteristics of Cloud Computing

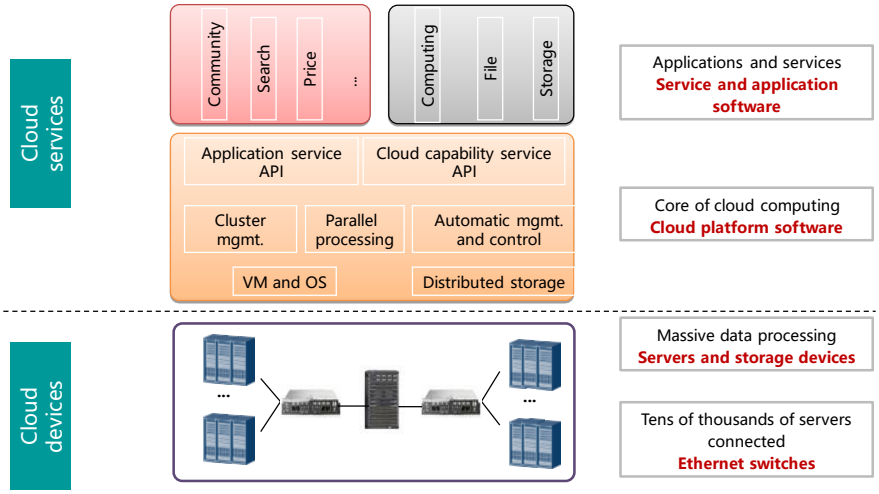
- On-demand self-service
- Ubiquitous network access
- Location independent resource pooling
- Rapid and elastic
- Pay-per-use



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- On-demand self-service: Consumers can deploy processing capabilities based on service requirements on servers and network storage, and do not need to communicate with each service provider.
- Ubiquitous network access: Various capabilities can be obtained over the Internet. The Internet can be accessed using the standard mode through various clients, such as mobile phones, laptops, and PDAs.
- Location independent resource pooling: Computing resources of service providers are centralized for customers to rent. Physical and virtual resources can be dynamically allocated to address customer requirements. Customers generally cannot control or know the exact location of the resources. The resources include storage, processors, memory, network bandwidth, and VMs.
- Rapid elastic: Resources can be rapidly and elastically provided to users. Users can also rapidly expand or reduce resources. A user can rent unlimited resources at any time.
- Pay per use: The service is charged based on pay-per-use or advertisement to optimally use resources. For example, a user can be charged per month based on the actually used storage, bandwidth, and computing resources. The cloud used by the same organization in a company can be charged between departments.

Cloud Computing Network = Computing and Storage Network



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

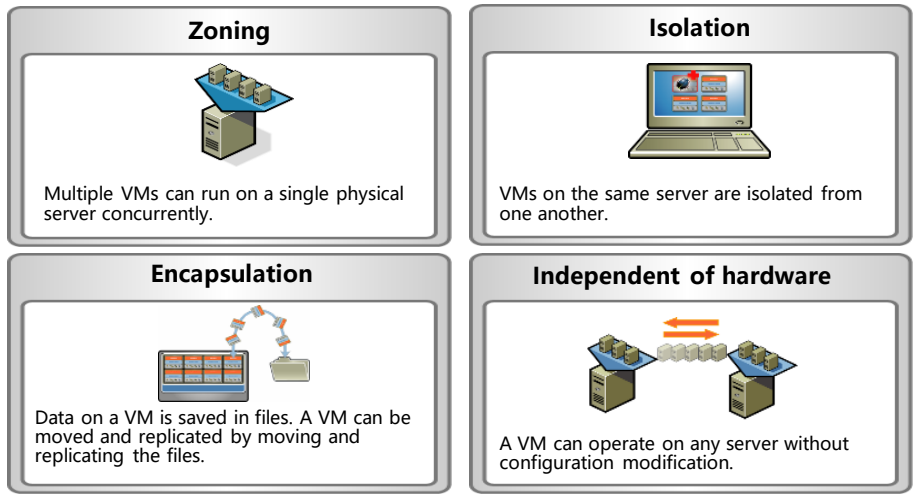
- From the technical prospective, cloud computing includes cloud devices and cloud services.
  - Cloud devices include the servers used for data computing and processing, the devices used for data storage, and the switches used for data communications.
  - Cloud services include the cloud platform software used for managing the physical resource virtualization scheduling and the application platform software used for providing services for users.
- Development of cloud computing
  - Numerous low-cost servers replace dedicated mainframes, minicomputers and high-end servers.
  - Distributed software replaces the traditional single-host OS.
  - Automatic management and control software replaces the traditional centralized management.

Contents

1. DC Development
2. **Cloud Computing Development**
   - Concept
   - Virtualization
3. Benefits of Cloud Computing
4. Deployment Modes of Cloud Computing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

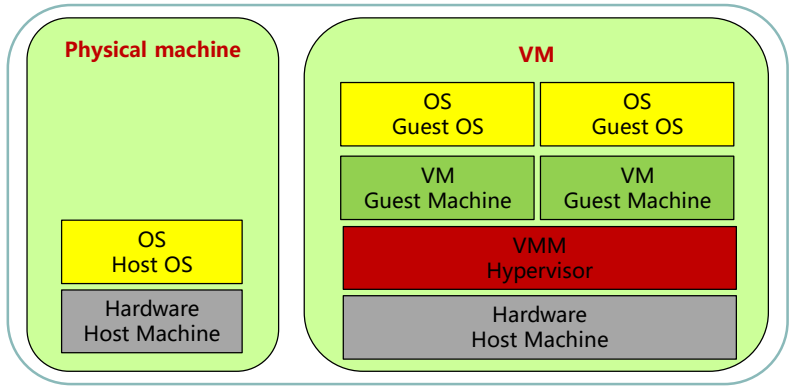
Essence of Virtualization



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Zoning: indicates the VMM's capability of allocating server resources to multiple VMs. Each VM can run an independent OS (same as or different from the OSs running on other VMs) so that multiple applications can run on one server. Each OS gains access only to its own virtual hardware (including the virtual NICs, CPUs, and memory) provided by the VMM.
- Isolation: VMs that run on the same server are isolated from each other.
  - Even if one VM cracks down or fails due to an OS failure, application breakdown, or driver failure, other VMs on the same server can still run properly.
  - If a VM is infected with worms or viruses, it will be isolated from other VMs. It seems that each VM is located at an independent physical machine.
  - Resource control helps achieve performance isolation. Specifically, you can specify the maximum and minimum resource usage for each VM to ensure that one VM does not use all resources, leaving no available resources for other VMs in the same system.
- Encapsulation: All VM data, including the hardware configuration, BIOS configuration, memory status, disk status, and CPU status, is stored into a small group of files that are independent of physical hardware. This enables users to copy, save, and migrate VMs by copying, saving, and migrating several files.

Important Concepts About Virtualization



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Host machine
  - Indicates customers' resources.
- Guest machine
  - Indicates virtualized resources.
- Guest and host OSs
  - If one physical machine is virtualized into multiple virtual machines, the physical machine is a host machine, and the virtual machines are guest machines. The OS installed on the physical machine is a host OS, and that installed on the virtual machine is a guest OS.
- Hypervisor
  - Through virtualization by the virtualization layer, upper-layer software regards VMs as real machines. The virtualization layer is called virtual machine monitor (VMM) or hypervisor.

Virtualization Technology Classification

- **By application scenario**
  - OS virtualization
    - VMware: vSphere
    - Microsoft: Hyper-V
    - Citrix: XenServer
    - Huawei: FusionSphere
  - Desktop virtualization
    - Microsoft: Hyper-V
    - Citrix: XenDesktop
    - VMware: VMware View
    - Huawei: FusionAccess
- **By hardware resource allocation mode**
  - Full virtualization
    - Virtual OSs are separated from underlying hardware, and the hypervisor located between them converts the codes sent by the virtual guest OSs to invoke the underlying hardware. Full virtualization requires no changes to the guest OS and boasts excellent compatibility.
    - Typical applications include VMware WorkStation, early-stage ESX servers, and Microsoft virtual servers.
  - Paravirtualization
    - Dedicated virtualization instructions are added to virtual guest OSs for invoking hardware resources through the hypervisor, relieving the performance overhead introduced by the hypervisor. The typical application is Xen.
  - Hardware-assisted virtualization
    - A new instruction set and processor running mode are added to the CPU for virtual OSs to directly invoke hardware resources.
    - The typical applications are Intel VT and AMD-V.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

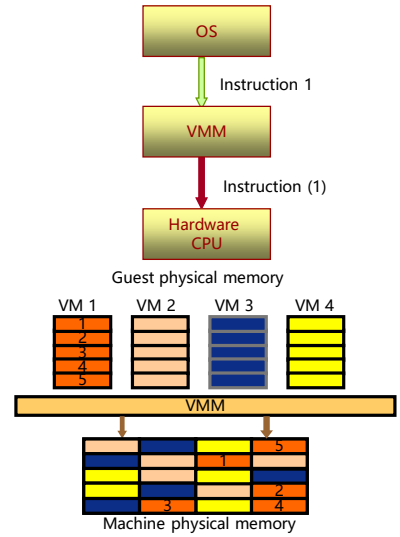
Introduction to Virtualization Products



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Computing Virtualization Principles

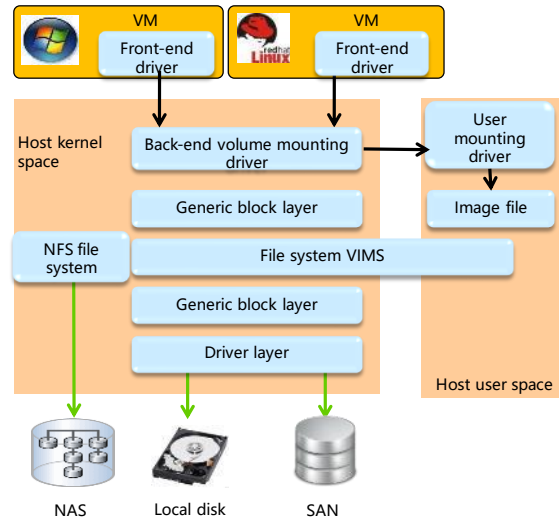
- **CPU virtualization**
  - When timer interrupt, the similar mechanism as that of the traditional OS, is triggered, VMM works to schedule VM disks based on the preset scheduling mechanism. Virtual disks can be SSDs, dynamic disks, and differential disks.
- **Memory virtualization**
  - The physical memory of physical machines is centrally managed and dynamically allocated to VMs.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

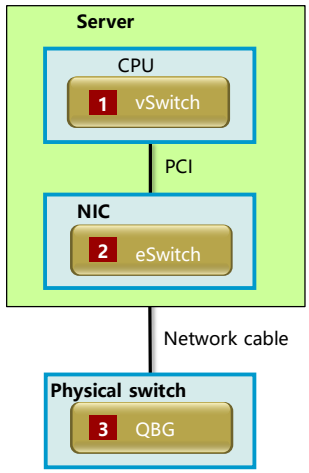
Storage Virtualization Principles

- Storage devices provide different performance and use different interface protocols. To address these differences, Huawei storage virtualization formats storage devices and converts various storage resources to centrally managed data storage resources. These data storage resources can be used to store information about VM disks, VM configurations, and snapshots. Users can apply the similar way to manage storage, improving storage management efficiency.
- A VM disk can be used only after it is mounted to a VM using disk driver. Disk driver is categorized as user driver, kernel back-end driver, and kernel front-end driver. The process of writing VM I/O data into disk files is as follows: The front-end driver captures all VM I/O operations and forwards them to the back-end driver. Then, the back-end driver converts these operations to I/O operations in the user driver and writes the data into disk files.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

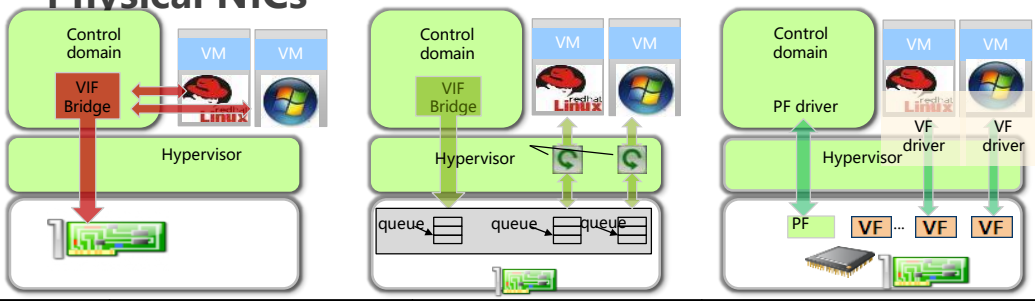
Three Typical Network Virtualization Methods



- **Server CPU-based virtual switching**
  - Advantage: flexibly extends functions.
  - Disadvantage: consumes server CPUs and provides low performance.
- **NIC-based virtual switching**
  - Advantage: provides high performance and conserves server CPU resources.
  - Disadvantage: depends on special NIC hardware.
- **Physical switch-based virtual switching**
  - Advantage: inherits functions of Layer 2 switches.
  - Disadvantages: features small specifications and poor scalability and does not support hypervisors.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

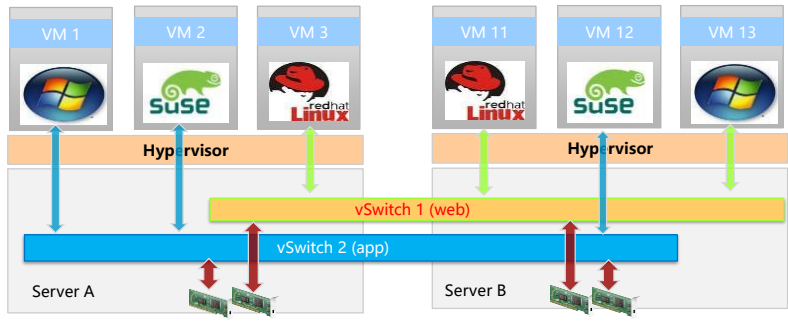
Relationship Between Network Virtualization and Physical NICs



	Common Virtual NIC	VMDq	SR-IOV
Difference	•Dom0 bridge queue	•Independent VM message queue	•Address translation enabled by the SR-ROV hardware technology
	•One-time data copy	•Hypervisor consumes a small amount of resources for each address translation.	•Hypervisor does not require address translation, thereby reducing calculation consumption.
Characteristic	•High host CPU overhead and low VM density •Smooth VM migration and snapshot	•Little host CPU overhead •Smooth VM migration and snapshot	•Little host CPU overhead •Affected VM migration and snapshot

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

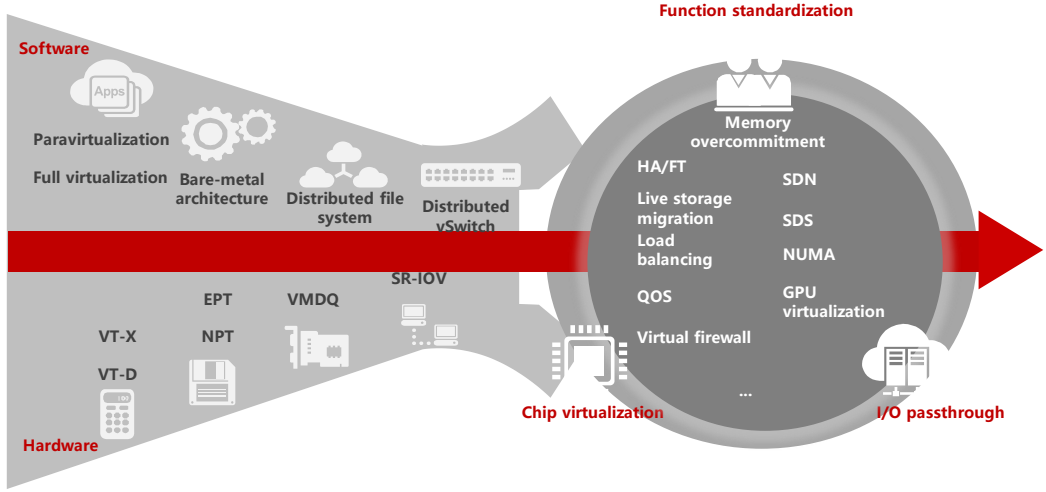
Distributed Virtual Switches



- **Technical features**
  - Provides a unified interface for configurations and management.
  - Supports VLANs, Layer 2 security policies, and bandwidth control.
- **Benefit**
  - Supports flexible access control over networks

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Development of Cloud Computing: Virtualization Functions Are to Be Standardized



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Virtualization is one of important cloud computing technologies.
- Software architectures are trend to become the same.
- Hardware-assisted virtualization narrows the software performance gap.
- I/O passthrough transfers the virtualization software functions to hardware boards.
- The rapid development of Microsoft Windows Server 2012 has changed the monopoly position of VMware.
- Software-defined storage (SDS)
  - SDS is a data storage mode, in which all control operations related to storage are performed by the external software. The software acts as a part of a server, OS, or hypervisor but not firmware in the storage device.
- SDN: According to the data released by InfoWorld in November 2011, SDN ranked second in the 10 new technologies that will affect the world in the coming 10 years. In July 2012, SDN vendor Nicira, a company focusing on SDN and network virtualization, was acquired by VMware for $1.26 billion. After that, Google pronounced that they had deployed SDN on their 10 IDC networks worldwide. The two cases made SDN draw strong attention.

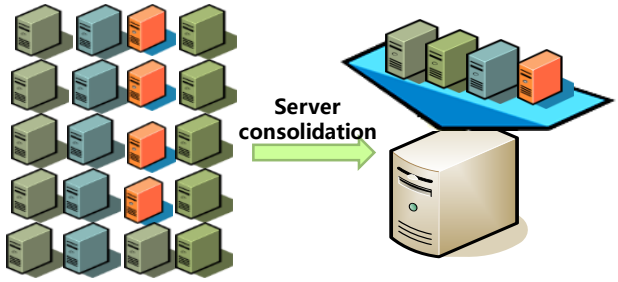
Contents

1. DC Development
2. Cloud Computing Development
3. **Benefits of Cloud Computing**
4. Deployment Modes of Cloud Computing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Resource Consolidation for Improved Resource Utilization

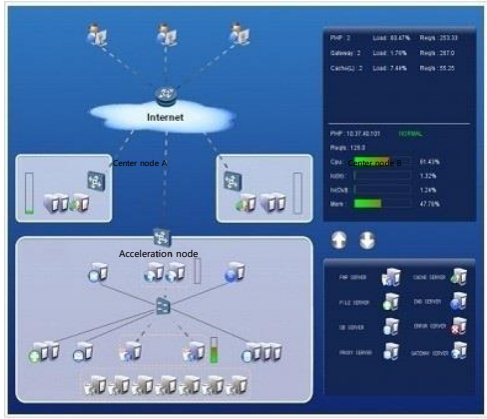


Benefits of resource consolidation
Shared hardware resources
Flexible adjustment of CPUs and memory of virtual servers
Memory overcommitment, improving resource utilization
Server consolidation, improving server utilization
Server consolidation, saving equipment room space

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Virtualization supports elastic scaling of resources.
- Each server is virtualized into multiple VMs, avoiding exclusive server occupation by a specific service.
- VM specifications (CPUs and memory) can be flexibly changed to increase or reduce the number of VMs, meeting computing resource requirements of services.
- With the virtualization computing technology, physical memory resources can be virtualized into more virtual memory resources, and thus more VMs can be created.

Rapid Deployment and Elastic Capacity Expansion



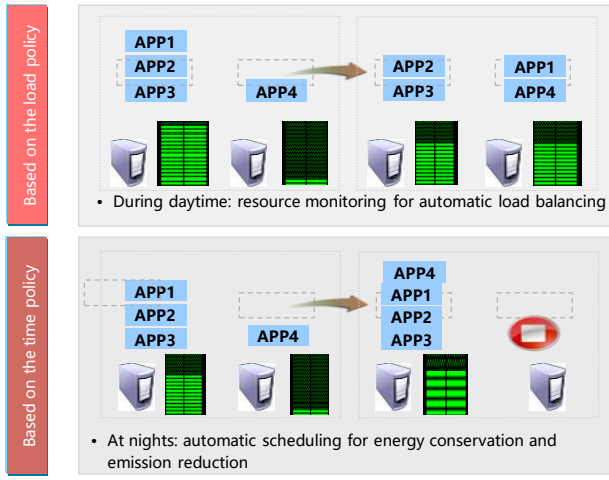
- Batch VM deployment
- Elastic resource addition/reduction
- Shortening the deployment period of a new service

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- At the early stage of deployment, deploy fewer servers because the business scale is small. Scale capacity later if necessary. Scaling is very simple. You only need to use PXE or ISO to install several compute nodes and then add them to the system through the O&M portal.
- The cloud-based service system adopts batch VM deployment.
- Large-scale resources are deployed in short time, fast responding to service requirements and saving a large amount of deployment time.
- Elastic resource addition/reduction adapts to shifting service requirements.
- Automatic deployment with minimal manual intervention is promoted.
- Customers will no miss market opportunities due to slow service deployment.
- The traditional service deployment cycle is month-based whereas the cloud-based service deployment period is shortened to minute or hour levels.

Automatic Scheduling for Energy Conservation and Emission Reduction



- Automatic and intelligent resource scheduling:
  - During daytime, resources are monitored according to the load policy to implement automatic load balancing and efficient heat management.
  - At nights, resources are adjusted according to the time policy to minimize the power consumption by powering off unused servers.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Policy-driven intelligent and automatic resource scheduling enables resources to be automatically allocated on demand and intelligently balanced. This reduces power consumption and emission.
  - During daytime, resources are monitored according to the load policy to implement automatic load balancing and efficient heat management.
  - At nights, resources are adjusted according to the time policy to minimize the power consumption by powering off unused servers.
- Dynamic Power Management (DPM) helps reduce power consumption in DCs. With DPM, when detecting that the load on a cluster reaches the predefined energy saving policy, the system consolidates the VMs in the cluster to a smaller number of hosts and then power off the idle hosts where no VMs are running. When detecting that VMs require more resources, the system powers on some hosts to provide resources to the VMs.

Lowered Temperature and Reduced Noise Create Green Offices

< 55 dB → 1 dB



Note:
1 dB is the lowest sound that people can hear. If the sound is below 20 dB, the environment is quite; if the sound is between 20 to 40 dB, it is like people softly whispering; if the sound is between 40 to 60 dB, people can carry on normal conversations; if the sound is above 60 dB, the environment is noisy.

35°C to 45°C → ≈ Room temperature



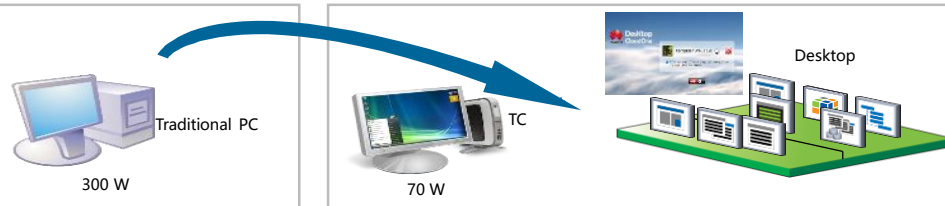
300 W → 70 W

Note:
Generally, the ambient temperature of a PC ranges between 35°C and 45°C. In a dense office environment, the ambient temperature is close to the PC temperature.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Hosts of physical PCs are replaced with TCs for greatly reduced heat and improved office environment.
- A physical host processes resources locally with powerful CPUs, and disks, and fans configured, thereby generating noise pollution. Compared with the physical host, a local TC only receives instruction inputs and displays interfaces but does not process resources locally (computing resources are processed at the remote DC), which produces little noise and optimizes the office environment.
- PC is short for Personal Computer.
- TC is short fort Thin Client. It allows access and use of virtual desktops in a cloud computing environment.

Efficient O&M for Reduced Costs



Maintenance

- Maintenance process: report a fault > assign personnel to handle > locate the fault > rectify the fault
- The PC maintenance process takes about two to four hours, resulting in a long downtime and high labor cost.
- Maintenance personnel: account for about 3% to 5% of the total number of employees.

Power consumption

- Traditional PCs are power consuming and drive high expenses.

Maintenance

- Maintenance-free frontend
- Maintenance process: failure (crash) > self-service restart > completed
- The maintenance process only takes about 3 minutes and the service downtime is short.
- Each maintenance personnel can maintain an average of 1000 desktops. The number of maintenance personnel is greatly reduced, which significantly cuts maintenance cost.

Power consumption

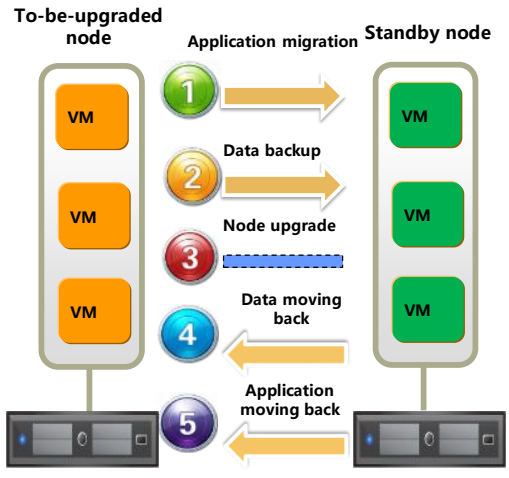
- Each TC consumes power from several watts to 20+ W, following the policy of energy conservation and emission reduction policy while trimming costs.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Traditional PCs may bring the following troubles from PC model selection, procurement, warehousing, provisioning to maintenance:
  - It will take a long time from project initiation to project launching.
  - The power consumption of traditional PCs is high, driving high costs.
  - The repair time of a traditional PC is long, affecting office efficiency.
  - Traditional PCs cannot be recycled but can only be replaced by new ones every three years.
  - Traditional PCs are spread across all offices, thereby requiring a large number of maintenance personnel and pushing up labor costs.
- In desktop cloud office scenarios, the number of resources to be processed is small and is centrally processed in DCs, eliminating office concerns.
- PC is short for Personal Computer.
- TC is short fort Thin Client. It allows access and use of virtual desktops in a cloud computing environment.
- IT is short for Information Technology. It is used in building enterprise DCs.

Hitless Capacity Expansion



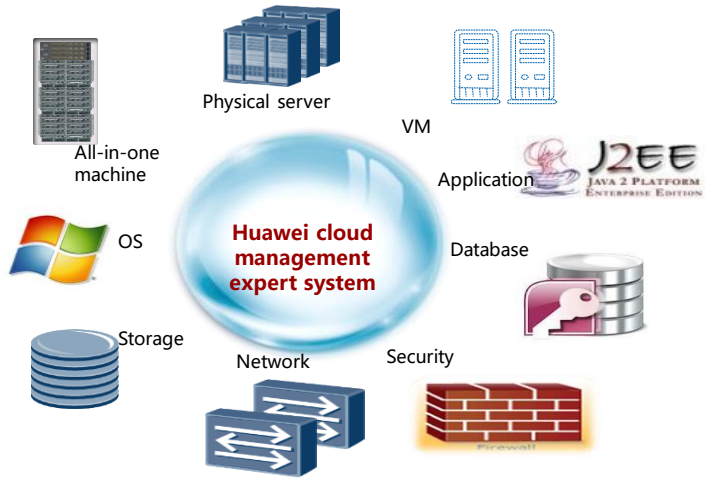
- Technical features
  - Live migration, automatically obtaining the standby node space
  - Automatic pipeline operation mode, easy to use
- Benefits
  - On-demand capacity adjustment, cutting maintenance costs
  - Reduced upgrade costs and risks, ensuring business continuity

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Management node upgrade: There are active and standby management nodes. You can upgrade one first. After performing an active/standby switchover, upgrade the other.
- Compute node upgrade: Before upgrading a compute node, migrate its VMs to another node. After the upgrade is completed, move the VMs back.

Unified Hardware and Software Management



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Huawei cloud solutions support a unified management over all-in-one machines, servers, storage devices, network devices, security devices, VMs, OSs, databases, and application software.
- Huawei cloud solutions support mainstream heterogeneous servers and storage devices.
- Huawei cloud solutions can integrate both FusionCompute and VMware virtualization software.
- According to the preceding description, it can be seen that the unified software and hardware management can facilitate management and reduce the procurement costs of management systems and the labor costs.
- It is assumed that the devices purchased by a company have to be managed by multiple different management systems, then:
  - The company has to purchase multiple different management systems.
  - Maintenance personnel have to learn more skills in using the management systems.
  - System upgrades, skill training, and personnel mobility will drive higher costs.

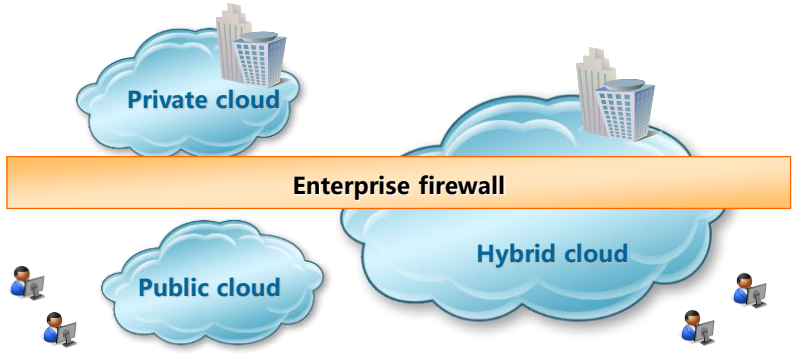
Contents

1. DC Development
2. Cloud Computing Development
3. Benefits of Cloud Computing
4. **Deployment Modes of Cloud Computing**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

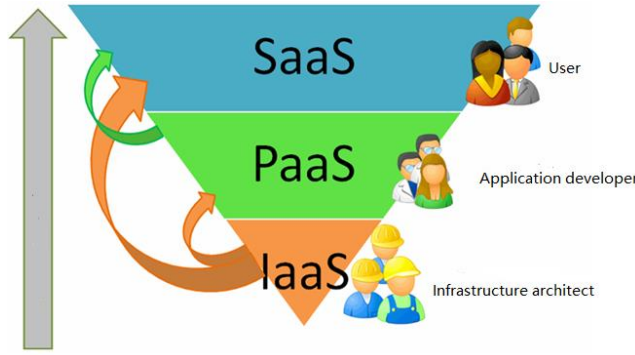
Deployment Modes of Cloud Computing



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Deployment models of cloud computing include private cloud computing, public cloud computing, and hybrid cloud computing.
  - Private cloud computing: It is used and operated solely by a single organization. For example, Huawei DCs adopt this model. Huawei serves both as its operation party and its user.
  - Public cloud computing: It is similar to a switch, which is owned by a telecom carrier and used by general public.
  - Hybrid cloud: It emphasizes that infrastructure consists of two or more clouds, but it is presented as a complete entity. For example, an enterprise can store its critical data, like finance data, to the private cloud and non-critical data to the public cloud. The private and public clouds together form a hybrid cloud. Another example goes like this: An on-line store has its steady volume of business and it is enough to have a private cloud to handle daily transactions. In Christmas promotion season, it can rent public cloud servers from carriers to cope with the temporary sales surge. All resources in the private and public clouds can be scheduled in a unified manner

Business Modes of Cloud Computing (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Infrastructure as a Service (IaaS): The infrastructure is provided to end users as a service. The infrastructure includes computing, storage, network, and other computing resources. Users can deploy and run any software, including OSs and applications. Typical applications are VM leasing and web disks.
- Platform as a Service (PaaS): The re-developed platform is provided to end users as a service. Customers do not need to manage or control the underlying cloud computing infrastructure and only need to control the deployed application development platform. The typical application is the Microsoft Visual Studio development platform.
- Software as a Service (SaaS): Applications running in the cloud computing infrastructure are provided to end users as a service. The typical application is the enterprise office system.

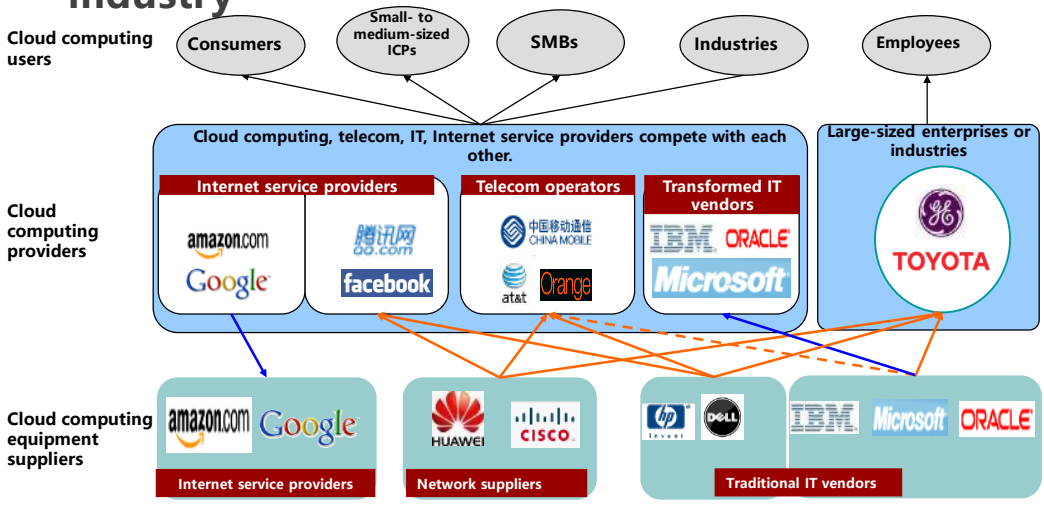
Business Modes of Cloud Computing (2)

Service Type	Service Content	Profit Model	Instance
SaaS	• Internet Web 2.0 application • Enterprise application • Teleservice	• Provides services meeting end-users' requirements on a pay-per-use basis.	Salesforce: CRM
PaaS	• Provides the application running and development environment. • Provides application development components (such as databases).	• Leases IT resources, universal web capabilities, and communication capabilities as a package to application developers and operators on a pay-per-use basis.	Microsoft: Azure Visio Studio
IaaS	• Leases IT resources such as computing, storage, and networks.	• Pay per use. • Makes profits by mass sales.	Amazon: EC2 cloud host

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Relative to Web 1.0, Web 2.0 is a general term of the new type of Internet applications. The key feature of Web 1.0 is that users obtain information through browsers. Comparatively, Web 2.0 focuses on interaction with the users. Specifically, uses are able to browse and edit web pages, which means users are no longer just readers but also editors of the Internet. In this context, the user participation pattern shifts from "only input" to "input + output + co-construction". Users start to proactively create Internet information instead of receiving only, making the service more personalized.
- A web hosting service is a type of Internet hosting service that allows individuals and organizations to provide their own website accessible via the World Wide Web.
- CRM is short for Customer Relationship Management.

Intense Competition in the Cloud Computing Industry



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The cloud computing industry involves cloud computing equipment suppliers, cloud computing service providers, and cloud computing end users.
- Cloud computing equipment suppliers refer to the suppliers who provide software and hardware required for building a cloud computing environment. Hardware suppliers refer to the server, storage device, switch, security device, and TC suppliers. Software suppliers refer to the cloud virtualization platform, cloud management platform, cloud desktop access, and cloud storage software suppliers.
    - Internet service providers are pioneers of cloud computing and leaders of advanced technologies and innovative business modes. They mainly provide cloud computing-driven low-cost information processing services. Such providers include Google and Amazon.
    - Telecom carriers use cloud computing to resolve practical problems, such as improving telecom network capabilities (massive computing and storage) and reducing costs. Such carriers include BT and AT&T.

Summary

- DC development
- Concepts of cloud computing
- Benefits of cloud computing
- Deployment modes of cloud computing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. Cloud computing is a type of computing technology based on Internet. By using this technology, users can provide shared software and hardware resources and information for computers and other devices on demand. (True or False)
2. (Multiple Choice Answer) What benefits does cloud computing bring to enterprise IT? ( )

A. Resource reuse, improving resource utilization

B. Unified maintenance, cutting down maintenance costs

C. Fast and flexible deployment

D. Centralized data management, enhancing information security

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Answers
  - True or False: T
  - Multiple Choice Answer: ABCD

More Information

- Huawei Official Websites
  - Enterprise business: http://e.huawei.com/us/
  - Technical support: http://support.huawei.com/enterprise/en/index.html
  - Online learning: http://learning.huawei.com/en/
- Documentation Tool
  - HedEx Lite

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei Learning Website
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei Support Case Library
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Introduction to Huawei KVM Virtualization

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Foreword

- Kernel-based Virtual Machine (KVM) is a virtualization module in the Linux kernel that allows the kernel to function as a hypervisor.
- This slide describes the virtualization technology, as well as the architecture and typical features of Huawei's KVM virtualization.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- Upon completion of this course, you will be able to:
  - Understand the definition, characteristics, and categorization of virtualization technologies.
  - Understand the KVM architecture.
  - Understand the relationship between KVM and QEMU and libvirt.
  - Understand how KVM virtualizes the CPU, memory, and I/O.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

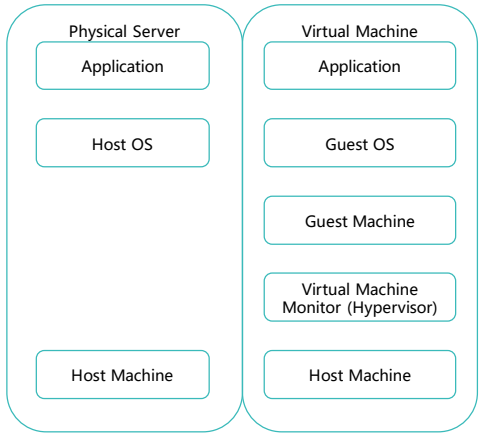
Contents

1. Introduction to Virtualization
2. KVM Background and Architecture
3. KVM Implementation Principles

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Important Concepts in Compute Virtualization



Guest OS:

Operating system running in a virtual machine (VM)

Guest Machine:

Virtual machine created through virtualization

Hypervisor:

Virtualization software layer, or

Virtual Machine Monitor (VMM)

Host OS:

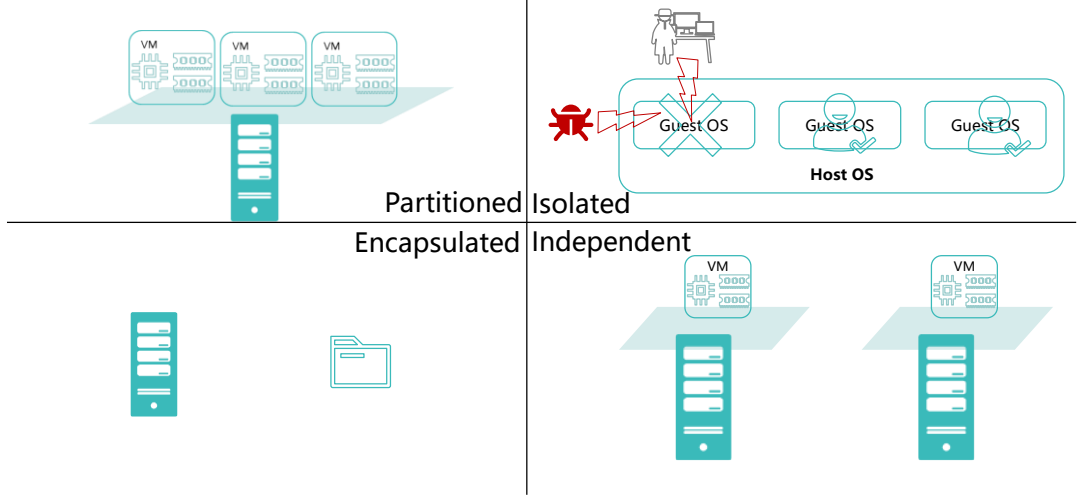
Operating system running in a physical machine

Host Machine:

Physical machine

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Characteristics of Virtualization



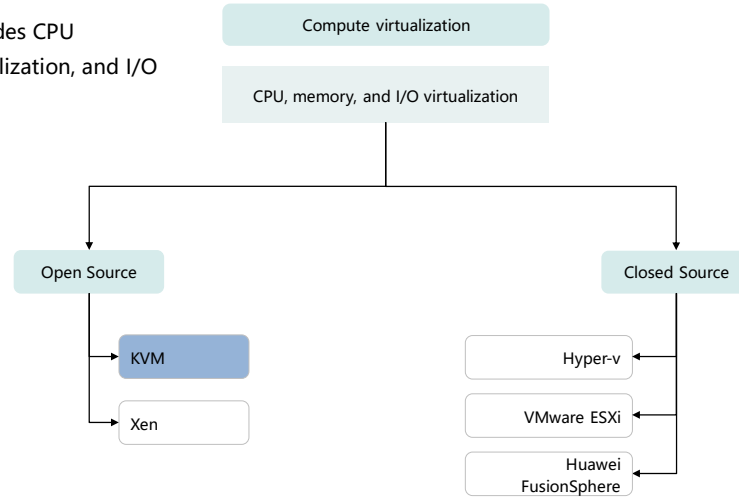
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Partitioned: indicates the VMM's ability to allocate server resources to multiple VMs. Each VM runs an independent OS (same as or different from the OSs running on other VMs on the same server) so that multiple applications can coexist on one server. Each OS has access only to its own virtual resources (including the virtual NIC, virtual CPUs, and virtual memory) provided by the VMM.
- Isolated: VMs that run on the same server are isolated from each other.
  - Even if one VM breaks down or experiences an OS failure, application breakdown, or driver failure, other VMs can still run properly.
  - If one VM is infected with worms or other viruses, other VMs will not be affected. It's like each VM is running on an independent physical machine.
  - Resources can be managed to provide performance isolation. Specifically, you can specify the minimum and maximum resources available for each VM to ensure that one VM does not use all resources, leaving no resources for other VMs residing on the same host.
  - Multiple workloads, applications, or OSs can run concurrently on one physical server, thus avoiding problems that may occur on the x86 server, for example, application program conflicts or DLL conflicts.
- Encapsulated: All VM data, including the hardware configuration, BIOS configuration, memory status, disk status, and CPU status, is stored into a group of files that are independent of the physical hardware. This enables users to clone, save, and migrate a VM simply by copying, saving, and migrating these files.

- Independent: VMs run on top of the virtualization layer. Therefore, only virtual hardware provided by the virtualization layer is visible. The VMs are independent of the underlying physical resources. In this way, the VMs can run on any x86 server (IBM, Dell, and HP) without any modification. This breaks the coupling between OSs and hardware and between applications and OSs/hardware.
- Therefore, these functions are significant for creating and running VMs.

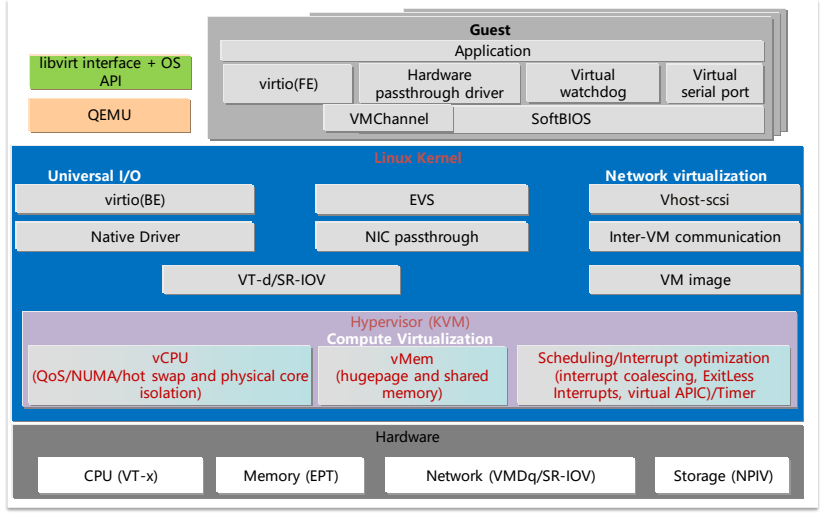
Compute Virtualization Technologies

- Compute virtualization includes CPU virtualization, memory virtualization, and I/O virtualization.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Huawei UVP Architecture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Guest (VM): The client system, including the vCPU, memory, console, NIC, storage device, is run by KVM in a restricted CPU mode.
- KVM: runs in the kernel space to virtualize the CPU and memory and to intercept I/Os of the guest OS. I/Os intercepted by KVM are processed by QEMU. KVM provides /dev/kvm device interfaces to QEMU.
- QEMU: is the modified native QEMU code used by the KVM VMs, runs in the user space, and provides I/O virtualization; Communicates with KVM through the ioctl function.

Contents

1. Introduction to Virtualization
2. KVM Background and Architecture
3. KVM Implementation Principles

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

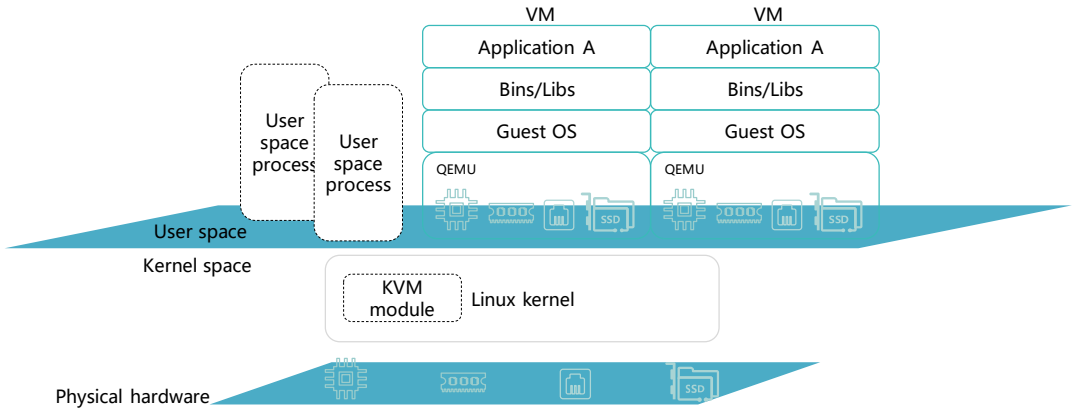
KVM Background

- KVM
  - Open source full virtualization solution
    - Architecture support
      - x86 (32-bit or 64-bit), IA-64, PowerPC, and S390
    - Relies on x86 hardware support: Intel VT-x/AMD-V
    - Kernel module, which turns the Linux kernel into a hypervisor
  - Cloud platform support for KVM
    - OpenStack, Eucalyptus, and AbiCloud support both KVM and Xen.
- KVM history
  - In October 2006, Israel company Qumranet released KVM.
  - In December 2006, KVM was integrated into the Linux kernel (Linux 2.6.20rc).
    - In February 2007, Linux 2.6.20 was officially released.
  - In September 2008, Red Hat acquired Qumranet for US$107 million.
  - In September 2009, RHEL 5.4 started to support KVM, as well as Xen.
  - Since November 2010, RHEL 6.0 or later version supported only KVM.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

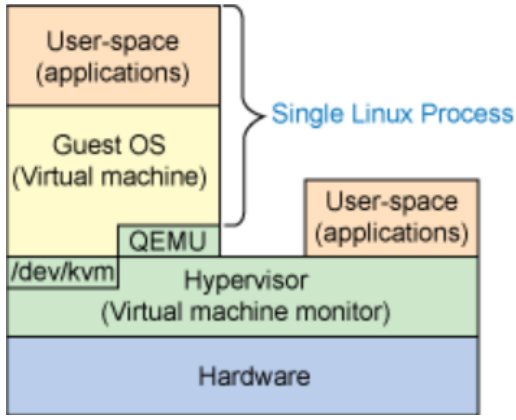
What Is KVM?

- KVM is a virtualization module embedded in the standard Linux kernel. KVM consists of a kernel module, kvm.ko, that manages the virtual CPU and memory.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

What Is KVM?



Guest: The guest OS, including the CPU (vCPU), memory, and drivers (console, NIC, and I/O device drivers), is run by KVM in a restricted CPU mode.

KVM: Runs in the kernel space to virtualize the CPU and memory and to intercept I/Os of the guest OS. I/Os intercepted by KVM are processed by QEMU.

QEMU: is the modified native QEMU code used by the KVM VMs, runs in the user space, and provides I/O virtualization. Enables communication between dev/kvm and KVM through the ioctl function.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

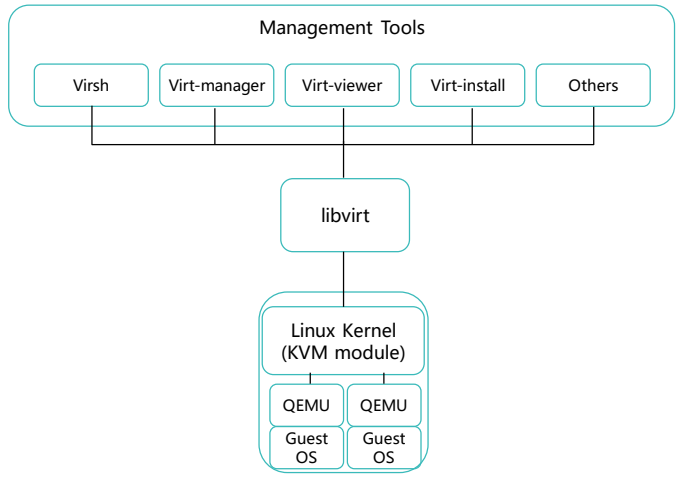
HUAWEI

- KVM can convert a standard Linux kernel into a VMM. The Linux standard kernel embedded with the KVM module can load guest OSs through KVM tools. Therefore, in such an OS platform, the VMM virtualization layer resides directly on the physical hardware layer of the host, and no independent host OS layer is provided. In this case, the VMM functions as the host OS. CPU instructions of the guest OS are executed directly, rather than through QEMU. This greatly improves the speed. KVM exposes the necessary APIs through /dev/kvm. User-mode programs can access these APIs by calling the ioctl function.
- The KVM kernel module provides only CPU and memory virtualization. Therefore, it must be combined with QEMU to form a complete virtualization solution, that is, qemu-kvm.
- As a hypervisor, KVM focuses on VM scheduling and memory management. I/O peripheral tasks are handled by the Linux kernel and QEMU. With I/O virtualization, storage and network resources are virtualized by the Linux kernel and QEMU.
- By integrating KVM, QEMU uses its kernel module to process CPU instructions by invoking the /dev/kvm interface through ioctl. KVM is responsible only for CPU and memory virtualization. QEMU emulates I/O devices (such as NICs and disks). Server virtualization is jointly implemented with KVM and QEMU, so QEMU is called qemu-kvm. QEMU is an emulator that emulates the CPU and other hardware needed by the guest OS. The guest OS believes that it communicates with the hardware directly. In fact, it interacts with the hardware simulated by QEMU. QEMU translates and sends these instructions to the real hardware. The

performance is compromised because all the instructions need to be translated by QEMU.

- QEMU also emulates other hardware, such as NICs and disks, which also affects the performance of these devices. To address this, pass-through and paravirtualization techniques, such as virtio_blk and virtio_net, are used to improve device performance.

KVM Architecture

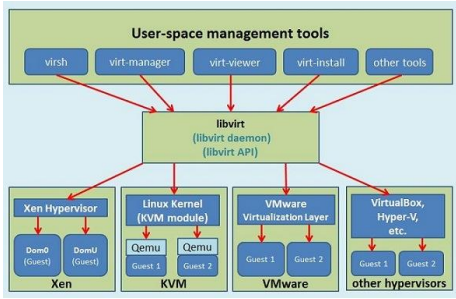


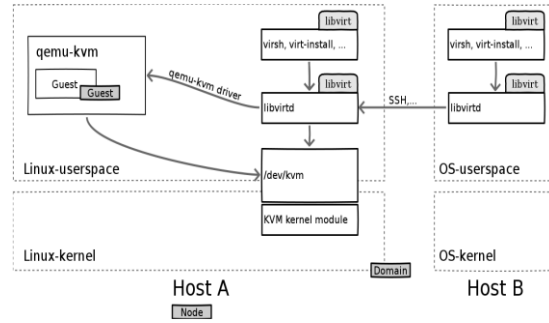
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

KVM and libvirt

- libvirt is the most widely used multi-platform VM management tool and API (multi-language).
- In KVM scenarios, libvirtd invokes qemu-kvm to manage VMs.
- Both QEMU and libvirtd trigger the kernel module through the character device /dev/kvm.
- VM management processes, such as virsh, use libvirt to query VMs.





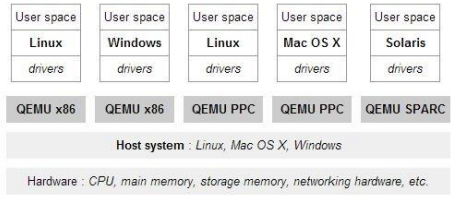
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

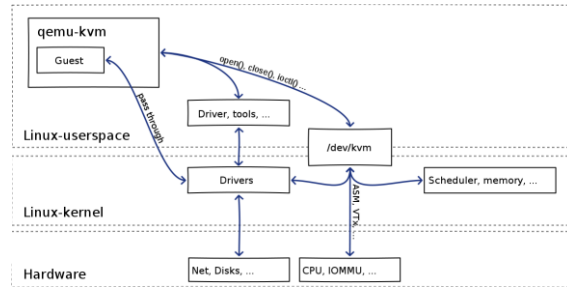
HUAWEI

- libvirtd is a daemon process that can be invoked by virsh locally or remotely.
- qemu-kvm invokes the kernel module of KVM to implement CPU virtualization.

KVM and QEMU

- QEMU is an emulator that emulates CPUs of a variety of hardware platforms, but at a low efficiency.
- By integrating KVM, QEMU leaves the handling of CPU instructions to the kernel module by calling KVM interfaces through ioctl. This is called qemu-kvm (qemu-system-XXX).
- QEMU also emulates other hardware, such as NICs and disks, which also affects the performance of these devices. To address this, pass-through and paravirtualization techniques, such as virtio_blk and virtio_net, are used to improve device performance.





 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Note: qemu-kvm in the figure is different from the native QEMU that does not support hardware acceleration.

Contents

1. Introduction to Virtualization
2. KVM Background and Architecture
3. KVM Implementation Principles
   - CPU Virtualization
   - Memory Virtualization
   - I/O Virtualization

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

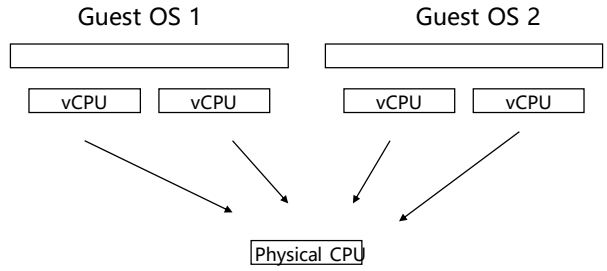
CPU Virtualization Challenges

- For native OSs, the following CPU requirements must be met:
  - CPU resources are always ready.
  - The OS has the highest level of privilege on the CPU.
- CPU virtualization leads to the following issues:
  - Multiple VMs share CPU resources.
  - Some commands can be used only by hypervisors.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CPU Virtualization - CPU Sharing

- How to enable CPU sharing between multiple VMs:
  - VMs use vCPUs, and the hypervisor schedules vCPUs to physical CPUs, thus implementing time-based sharing of physical CPU resources.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CPU Virtualization - Rights Management

- Sensitive instruction processing
  - Typically, Privilege Deprivileging and Trap-and-Emulation are used. Specifically, the guest OS runs on the non-privilege level (Deprivileging) and VMM on the highest privilege level, that is, the VMM fully controls system resources. While de-privileged, the guest OS can still execute most of the instructions on hardware. Only the privileged instructions are trapped and emulated by the VMM.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

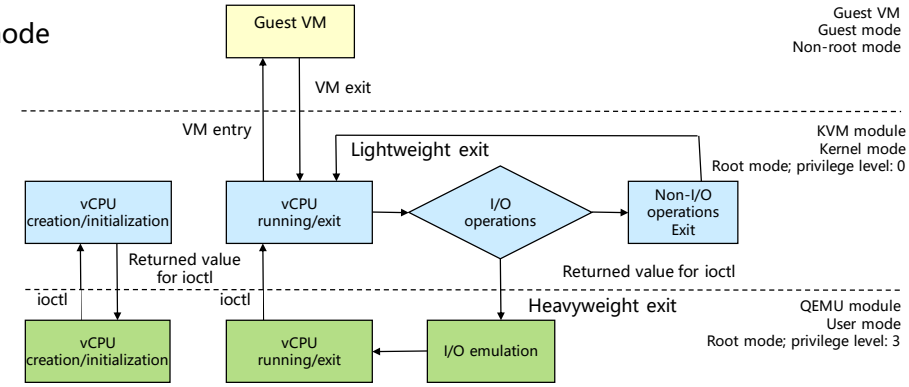
CPU Virtualization - VT-x

- VT-x: Intel hardware-assisted virtualization technology, which quickly implements vCPU context switching using mechanisms such as VMCS and VMEntry/VMExit
- With Intel VT-x, there are two distinct modes of CPU operation: root mode and non-root mode
  - Root mode: used to run hosts.
  - Non-root mode: used to run VMs. VM entry is a VT-x instruction that can switch from the root mode to the non-root mode.
- All x86 instructions can run in both root mode and non-root mode. The difference is that when a privileged instruction runs in non-root mode, it is switched to the root mode through VM exit, and the KVM module processes the privileged instruction through simulation. That is, privilege instructions cannot be handled in non-root mode.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

CPU Virtualization Modes with KVM

- User mode
- Kernel mode
- Guest mode



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Introduction to Virtualization
2. KVM Background and Architecture
3. KVM Implementation Principles
   - CPU Virtualization
   - Memory Virtualization
   - I/O Virtualization

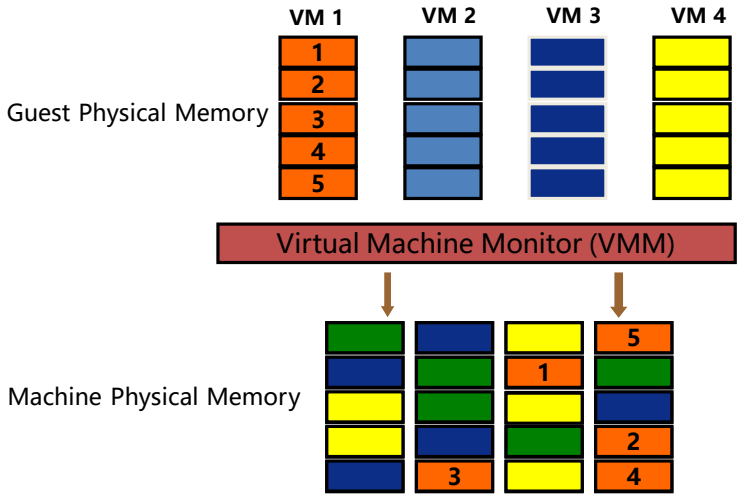
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Memory Virtualization Challenges

- For native OSs, the following memory requirements must be met:
  - The memory starts from physical address 0.
  - Consecutive memory blocks are allocated.
- Virtualization leads to the following issues:
  - **Starting from physical address 0:** There is only one physical address 0. The memory space of every VM cannot all start from 0.
  - **Consecutive addresses:** Although consecutive physical addresses can be allocated, this way of memory allocation leads to poor efficiency and flexibility.
- Glossary
  - GVA: Guest Virtual Address
  - GPA: Guest Physical Address
  - HPA: Host Physical Address

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Memory Virtualization Challenges



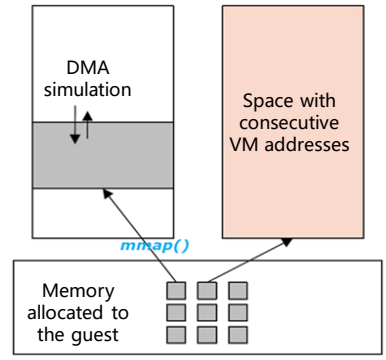
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The VMM needs to handle the following:
  - Maintains the physical address mapping between the guest and host.
  - Intercepts VM access requests, and replaces the addresses in these requests with the physical addresses of the host according to the mapping relationship.

Memory Virtualization with KVM

- KVM provisions the VM memory by calling the mmap interface. In the virtual address space of the QEMU main thread, a space of consecutive addresses is declared for the physical memory mapping of each guest machine.
- The physical addresses exposed to the VM are consecutive VM addresses in the space.

KVM memory mapping

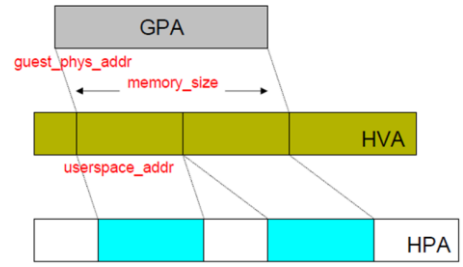


Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Memory Virtualization with KVM - MMU Virtualization

- VM memory allocation
  - Guest physical addresses are located in the virtual address space of the kvm-qemu process.
  - The guest physical memory page is allocated by the host on demand.
- Two types of MMU virtualization
  - Software: Shadow page table
  - Hardware: EPT (Intel)

GPA: Guest Physical Address
EPT: Extended Page Table
MMU: Memory Management Unit

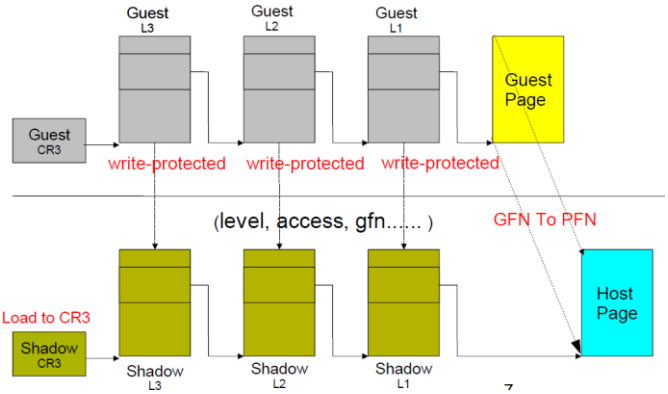


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- MMU: Memory management unit
- KVM uses the existing kernel mechanism to manage memory pages and allocate resources on demand. The guest physical memory (GPA) addresses are located in the virtual address space of the kvm-qemu process, and the virtual address space of the guest is mapped to the GPA.
- Two ways of MMU virtualization are available: shadow page table and hardware-assisted. Currently, the hardware-assisted mode (Intel EPT/AMD NPT) implements two-level address translation through the hardware, accelerating GPA-HPA conversion and reducing VM exits.
- Address space
  - Physical address space: hardware address space, for example, 4 GB.
  - Virtual address space: Each process has its own virtual address space.

KVM Memory Virtualization - Shadow Page Table

- Guest Page Table is set to write-protected to ensure consistency between the guest pages table and the shadow page table.
- If the dirty bit of guest PTE is not set, the writable bit of shadow PTE is deleted.

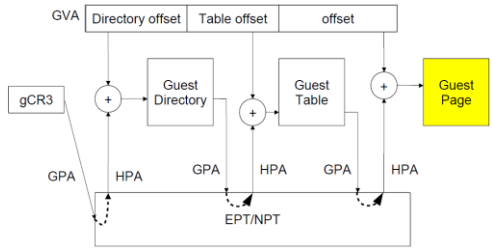


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

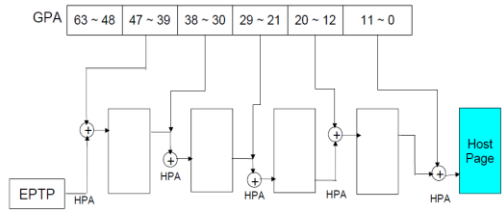
- The shadow page table is used when there is no EPT. However, CPUs launched by Intel since 2007 and 2008 support EPT.

KVM Memory Virtualization - EPT/NPT



Advantage: VM exits are reduced, outperforming shadow page tables.

EPT/NPT



Converting GPA into HPA

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- EPT: Intel adds EPT based on VT-x as an extension of VT-x. EPT is dedicated to memory virtualization. EPT implements memory virtualization access using hardware.
- After the KVM module creates the GPA-HPA page table of the VM, memory access can be completed by only hardware, without the need of simulation and interception.

Contents

1. Introduction to Virtualization
2. KVM Background and Architecture
3. KVM Implementation Principles
   - CPU Virtualization
   - Memory Virtualization
   - I/O Virtualization

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

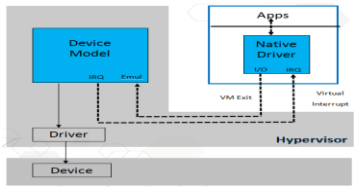
I/O Virtualization Challenges

- I/O virtualization handles the following:
  - Device discovery
    - Controls devices accessible to VMs.
  - Access interception
    - Accesses devices through I/O ports or MMIO.
    - Devices exchange data with the memory through DMA.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

KVM I/O Virtualization - Full Simulation

- Use software to fully simulate a specific device.
  - Keep the same software ports, for example, PIO, MMIO, DMA, and interrupt.
  - Simulate virtual devices that are different from physical devices in the system.
- Multiple context switches are required for each I/O operation.
  - VM and hypervisor
  - QEMU and hypervisor
- Device simulated by software does not affect the software stack of VMs.
  - Native drivers



handle guest io instructions code path:

vcpu_enter_guest

| *vmexit*

kvm_x86_ops->run ---> kvm_x86_ops->handle_exit(*vmx_handle_exit*)

|

kvm_vmx_exit_handlers[exit_reason](*handle_io* in vmc.c)

|

kvm_emulate_pio (x86.c)

|

/ \

return 0 kernel_pio(x86.c) --> complete_io

| |

return (userspace) vcpu_enter_guest

| *mode switch@ qemu-kvm*

kvm_handle_io (kvm_run() kvm_all.c)

|

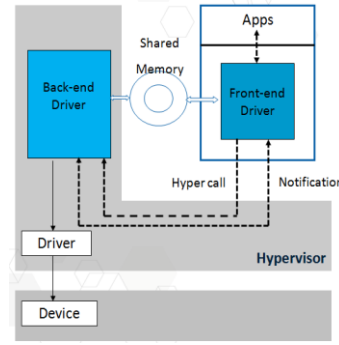
ioctl(KVM_RUN)

Full simulation involves a lot of data copying. For devices that require frequent I/O operations, such as disks and NICs, full simulation delivers poor performance.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

KVM I/O Virtualization Optimization - virtio

- Virtualization of special devices
  - Special device drivers, including the frontend drivers on VMs and the backend drivers on the hosts
  - Efficient communication between the frontend and backend drivers
- Reduced data transmission overhead between VMs and hosts
  - Shared memory (Virt RING)
  - Batched I/O
  - Asynchronous event notification mechanism (wait and notify) between Eventfd lightweight processes
- Efficient and standard PV drivers
  - PCI-compatible: device discovery, configuration, and IRQ
  - Support multiple virtualization platforms, such as KVM and Lguest.
  - Device types:
    - virtio-blk, virtio-net, virtio-balloon, virtio-console, virtio-scsi, and virtio-9p

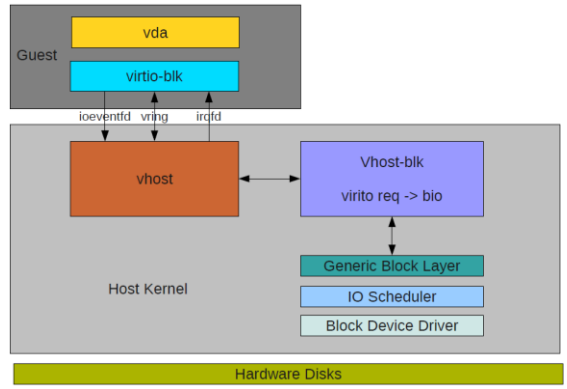


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- virtio, to be specific, the frontend and backend driver model, delivers high simulation efficiency without kernel modification.
- virtio uses vring to store requests and responses. vring can be accessed by both the frontend and backend. The frontend notifies the backend of request receipt through PIO, and the backend notifies the frontend of request completion through interrupt injection. The vring space is requested by the frontend which writes the addresses to the PCI configuration space. As all the memory of the VM is simulated by QEMU, QEMU stores the relationship between the GPA and the HVA. Therefore, the frontend and backend can easily access the memory pointed to by vring. vring is also called virtual object interface. The number of virtual object interfaces varies according to devices. For example, the network driver uses two virtual queues, one for receiving and the other for sending, while the block device driver uses only one virtual queue.

KVM I/O Virtualization Optimization - vhost

- vhost optimization covers vhost-block, vhost-net, and vhost-scsi.
- The I/O requests from the VM are directly mapped to bio on the host, reducing I/O execution on the VM and delivering better performance. The frontend is fully compatible with vhost.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- virtio, to be specific, the frontend and backend driver model, delivers high simulation efficiency without kernel modification.
- virtio uses vring to store requests and responses. vring can be accessed by both the frontend and backend. The frontend notifies the backend of request receipt through PIO, and the backend notifies the frontend of request completion through interrupt injection. The vring space is requested by the frontend which writes the addresses to the PCI configuration space. As all the memory of the VM is simulated by QEMU, QEMU stores the relationship between the GPA and the HVA. Therefore, the frontend and backend can easily access the memory pointed to by vring. vring is also called virtual object interface. The number of virtual object interfaces varies according to devices. For example, the network driver uses two virtual queues, one for receiving and the other for sending, while the block device driver uses only one virtual queue.

Quiz

1. Which CPUs are supported by KVM? ()

   A. Inter VT-x

   B. AMD-V

   C. Power PC

   D. s390

2. Which of the following statements are true? ()

   A. KVM does not perform any hardware emulation, which is implemented by QEMU.

   B. Since Linux 2.6.20, KVM has become a module included in the Linux kernel.

   C. To use KVM, CPUs need to support the virtualization function.

   D. KVM is completely open-source.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Answer:
  1. ABCD
  2. ABCD

Summary

- Virtualization Technology
- KVM Background and Architecture
- Implementation Principles

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Huawei official websites:
  - Enterprise business: http://e.huawei.com/en/
  - Technical support: http://support.huawei.com/enterprise/en/
  - Online learning: http://learning.huawei.com/en/
- Document tool:
  - HedEx Lite

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei E-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Huawei FusionSphere OpenStack Cloud Platform

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Foreword

- This course describes the organizational architecture, functions, and features of OpenStack and product features of Huawei FusionSphere OpenStack.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- After finishing this course, you will be able to:
  - Understand the background of OpenStack.
  - Describe the organizational structure of OpenStack.
  - Master functions and features of OpenStack.
  - Understand Huawei FusionSphere OpenStack enhancements.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. **OpenStack Background**
2. OpenStack System Architecture
3. Functions and Features of OpenStack
4. Huawei FusionSphere OpenStack Enhancements

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Open + Stack = OpenStack

- OpenStack is a cloud computing platform project jointly developed by the National Aeronautics and Space Administration (NASA) of the United States and Rackspace and is an open-source project released under the terms of the Apache license. It helps service providers and enterprises to achieve cloud infrastructure services similar to Amazon EC2 and S3.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- OpenStack is an open-source cloud computing management platform project. It consists of several main components. OpenStack supports almost all types of cloud environments and aims to provide a rich, standard, manageable, and scalable cloud computing management platform. OpenStack provides the Infrastructure as a Service (IaaS) solution through various complementary services, each of which provides an API for integration.
- OpenStack is an open-source project aimed at providing software for constructing and managing public and private clouds. The OpenStack community has more than 130 enterprises and 1350 developers. These organizations and individuals use OpenStack as the general frontend for IaaS resources. The primary object of OpenStack is to simplify the cloud deployment process and provide high scalability.
- Five open-source protocols (BSD, Apache, GPL, LGPL, MIT)
- BSD open-source protocol(original BSD license and free BSD license)
- Apache License is a protocol used by the famous non-profit open-source organization Apache. Similar with BSD, Apache License encourages code sharing, respects copyright of the original author, allows code modification and re-release (as open-source or commercial software).
- Apache License is also a friendly license for commercial applications. Users can modify code as required and release or sell them as open-source or commercial products.

Open + Stack = OpenStack

- The main objective of OpenStack is to manage resources in the data center and simplify resource allocation. OpenStack manages the following types of resources:
  - Compute resource: OpenStack can plan and manage a large number of virtual machines (VMs), allowing enterprises or service providers to provide compute resources on demand. Developers can use the APIs to access compute resources to create cloud applications. Administrators and users can use a web browser to access these resources.
  - Storage resource: OpenStack can provide required objects and block storage resources for cloud services and cloud applications. Due to the requirements for performance and the price, many organizations are not satisfied with traditional enterprise-level storage technologies. OpenStack can provide configurable object storage and block storage functions based on customer requirements.
  - Network resources: Data centers now have a large number of devices such as servers, network devices, storage devices, and security devices, and these devices will be divided into more virtual devices or virtual networks. This causes the explosive increase of IP addresses, route configuration, and security rules. Traditional network management technologies cannot manage the next-generation networks with high scalability and high automation. OpenStack provides plugin, scalable, and API-driven network and IP address management.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

OpenStack Participants



In 2013, Huawei was officially accepted by the OpenStack Foundation as a Gold Member.

In 2017, Huawei was officially accepted by the OpenStack Foundation as a Platinum Member.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Typical Business Models and Vendors of OpenStack

- Integrated solution providers
  - Representatives: HP, Huawei, Mirantis, IBM, and Oracle
- Public cloud carriers
  - Representatives: Rackspace and Huawei
- Distributions vendors
  - Representatives: Red Hat, Canonical, and SUSE
- System software vendors
  - Representatives: VMware
- Hardware device vendors
  - Representatives: Dell, EMC, and Cisco
- Application and management component vendors
  - Representatives: Tesora and Parallels

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

- Nebula computing platform of NASA
- The following large-scale hardware vendors support OpenStack: IBM, AMD, Intel, and Dell.
- In October 2010, Microsoft announced to support the integration of OpenStack and Windows Server 2008 R2.
- In February 2011, Cisco officially joined the OpenStack project, focusing on development of OpenStack network services.
- Ubuntu will develop cloud network plans focusing on stacks in future.
- In April 2012, IBM announced to join the OpenStack project and serve as a major sponsor.
- In October 2012, the Viacloud interconnection cloud platform joined the OpenStack project to develop OpenStack-based public and private cloud platforms.
- In 2013, IBM announced to provide OpenStack-based private cloud services and related applications in the IBM Pulse conference.

Enterprise-Level OpenStack Requirements

- OpenStack is an ideal foundation for enterprise-level private clouds and will necessarily become a new-generation cloud operating system (OS) kernel. However, it is not a complete cloud OS.
- At present, OpenStack faces challenges in several key areas. To deal with these challenges, OpenStack is delivered in robust enterprise-level products. These products provided in the industry support technical support, quick installation, and routine management. Without vendors providing these products, OpenStack would never be widely used.
- OpenStack is not MySQL. It is similar to Linux kernel which needs a complete OS to run. What does enterprise-level OpenStack exactly need? There are six key factors as follows:
  - API availability of 99.999% and scalable control plane
  - Robust management and security model
  - Open architecture
  - Hybrid cloud compatibility
  - Scalable resilient architecture
  - Comprehensive support and services

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

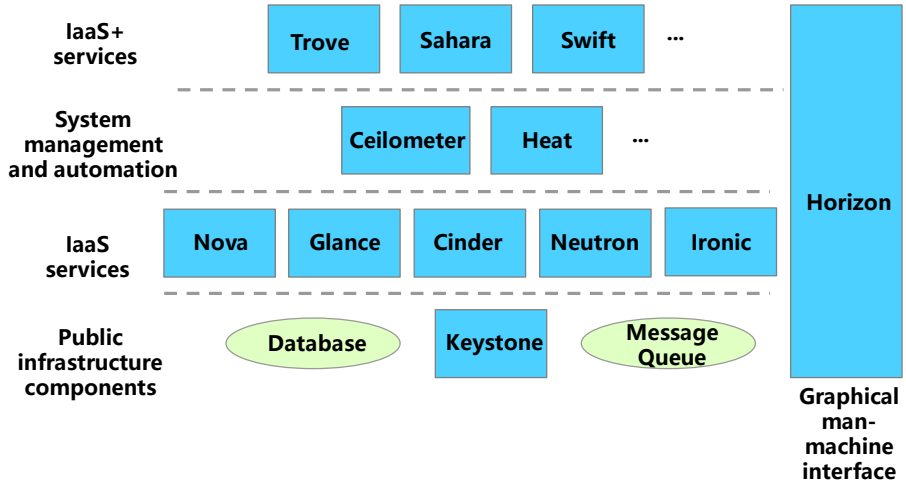
Contents

1. OpenStack Background
2. **OpenStack System Architecture**
3. Functions and Features of OpenStack
4. Huawei FusionSphere OpenStack Enhancements

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

OpenStack Layers



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- By now, the OpenStack project covers common service types at the IaaS layer, part of system management and automation services, and some important IaaS+ services.

OpenStack Architecture and Core Projects

Service	Project Name	Description
Console	Horizon	A user can use Horizon to interact with various OpenStack services, such as starting VM instances, assigning IP addresses, and configuring access control.
Compute	Nova	A user can use Nova to allocate and manage VMs on demand.
Network	Neutron	Nova uses Neutron to manage the connection between network settings. Neutron allows end users to create and add network interfaces and supports a large number of network devices and technologies with plugins.
Storage services		
Object storage	Swift	Swift can be used to store files, but it cannot be used to mount files.
Block storage	Cinder	Cinder can be used to provide the block storage service for persistent storage.
Shared services		
Identity authentication	Keystone	Keystone provides authentication and authorization for OpenStack.
Image service	Glance	Glance provides the VM image registration service. At the same time, Nova uses Glance to dispatch instances.
Metering/Monitoring service	Ceilometer	Ceilometer provides functions such as charging, benchmark tests, and data statistics.
High-level services		
Orchestration service	Heat	Heat can use its HOT template or AWS CloudFormation template and REST APIs of each OpenStack service to organize component resources to cloud applications.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

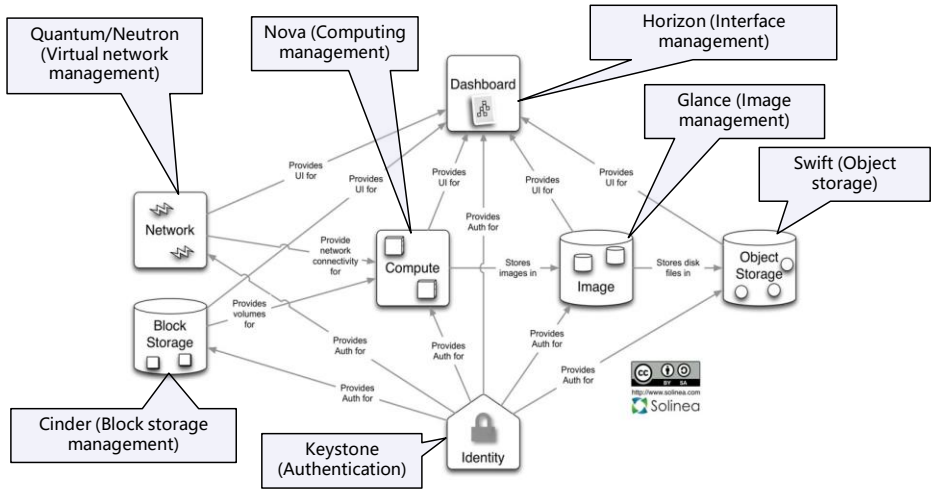
Contents

1. OpenStack Background
2. OpenStack System Architecture
3. **Functions and Features of OpenStack**
4. Huawei FusionSphere OpenStack Enhancements

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

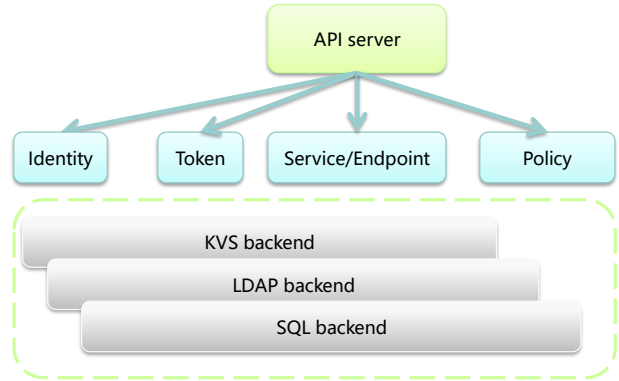
Introduction to OpenStack Modules



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Keystone Overview

- Keystone provides authentication and access policy services for all OpenStack components. Depending on its REST (identity API) system, Keystone provides authentication and authorization for mainly (but not limited to) Swift, Glance, and Nova. In fact, Keystone authenticates the request from action and message sources.
  - User
  - Tenant
  - Role
  - Service
  - Endpoint



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- A user refers to a person or program that can access system services using Keystone. Users are authenticated by Keystone using their credentials, such as their passwords and API keys.
- A tenant is a collection of resources that can be accessed in each service. For example, a tenant can be machines in Nova, image storage in Swift and Glance, and network resources in Quantum. Users are always bound to certain tenants by default.
- A role indicates resource rights that a group of users can access, such as VMs in Nova and images in Glance. Users can be added to any global or tenant's role. For a global role, the role permission of a user is applicable to all tenants, that is, the user can execute the rights specified by the role on all tenants. For a tenant's role, the user can execute the rights specified by the role on only the tenant.
- A service can be Nova, Glance, and Swift. According to the preceding definitions (user, tenant, and role), a service can be used to check whether the current user has the rights to access the service resources. If a user attempts to access a service of its tenants, the user must know whether the service exists and how to access it. In this case, different names are used to indicate different services. The role mentioned above can be bound to a service. For example, if the administrator rights are required for Swift to create an object, it is not necessary to access Nova as an administrator for the same role. To achieve this goal, we should create two independent administrators roles. One is bound to Swift, and the other is bound to Nova. In this way, the administrator access to Swift does not affect Nova and other services.
- An endpoint can be understood as an access point of a service. If you want to access a service, you must know its endpoint. Keystone contains an endpoint template that provides endpoints of all existing services. You can see the endpoint template in the conf folder during Keystone installation.

Examples

Parameter	Example
User	Hotel guests
Credentials	Room key
Token	Special key
Tenant	Hotel
Service	Service types provided by the hotel such as diet and entertainment services
Endpoint	Services in detail such as barbecue and badminton
Role	Higher VIP level, higher permission

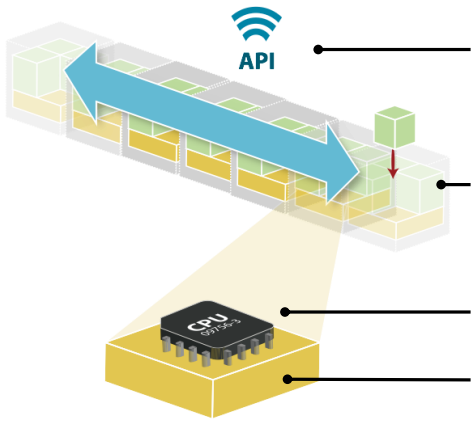
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Nova Overview

- Nova is a core component of OpenStack. Many other OpenStack components are separated from the Nova project and serve it. All activities in the OpenStack VM instance life cycle are processed by Nova. This makes Nova a scalable platform to manage compute resources, networks, and authentication. However, Nova does not provide any virtualization capability. Instead, it uses the libvirt APIs to interact with supported hypervisors (Xen and KVM). Nova provides services externally through the web services APIs which are compatible with Amazon Web Services (AWS) EC2 APIs and supports message-based asynchronous communication.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Nova - Compute Virtualization (1)



1. Based on the REST APIs
   Friendly service access mode
2. Supports large-capacity horizontal expansion.
   The quantity of supported VM instances linearly increases with the quantity of nodes.
3. Independent from hardware and supports multiple types of standard hardware.
   No customization requirements for dedicated hardware devices
4. Independent from hypervisor and supports multiple types of hypervisor.
   KVM, LXC, QEMU, UML, ESX, Xen, PowerVM, Hyper-V
   Supports all mainstream hypervisors and does not depend on a specified vendor.

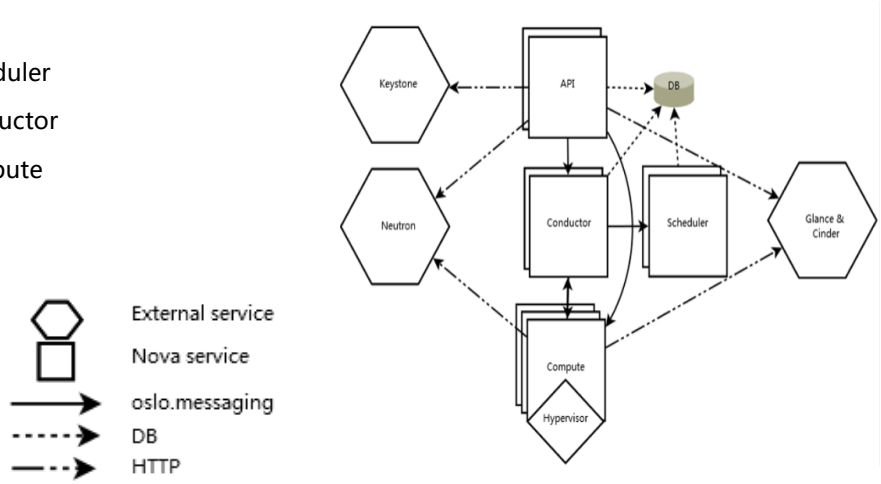
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Functions and features:
- Instance lifecycle management
- Compute resource management
- Network and authorization management
- REST APIs
- Asynchronous continuous communication
- Supports various hosts: Xen, XenServer/XCP, KVM, UML, VMware vSphere, and Hyper-V

Nova - Compute Virtualization (2)

Nova consists of the following components:

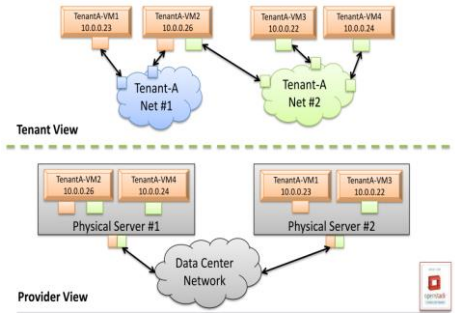
- Nova-api
- Nova-scheduler
- Nova-conductor
- Nova-compute



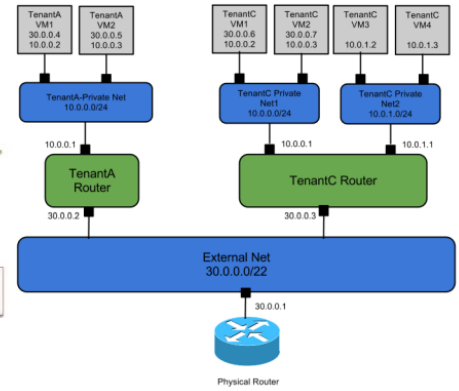
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The API server provides an interface for the cloud infrastructure to interact with the outside. It is the only channel for external users to manage the cloud. The API of each EC2 is invoked using the web service, and then the API server sends the request to the target object in the cloud through the message queue for processing. As a substitute for EC2-API, the native API of OpenStack, which is called "OpenStack API", can also be used.
- OpenStack uses the message queue for communication based on the Advanced Message Queuing Protocol (AMQP). Nova asynchronously invokes the request response. After the request is received, Nova triggers a callback immediately. Because of asynchronous communication, no action of users is set to the waiting status for a long time of period. For example, the process of starting an instance or uploading an image is time-consuming. The API invocation waits for the returned result without affecting other operations. This asynchronous communication plays an important role in making the system efficient.
- The scheduler invokes the nova-API to the target. The scheduler runs as the daemon process named nova-schedule and selects a computing server from the available resource pool according to the scheduling algorithm. Many factors may affect the scheduling result, such as the load, memory, distance between sub-nodes, and CPU architecture. Nova-scheduler uses the pluggable architecture.
- The main task of nova-compute is to manage the full life cycle of the instance. Nova-compute receives and executes the request through the message queue, and performs various operations on the instance. In a typical production environment, many nova-computes are deployed. According to the scheduling algorithm, an instance can be deployed on any available nova-compute.

Neutron - Network Virtualization



Mappings between the physical network and virtual network

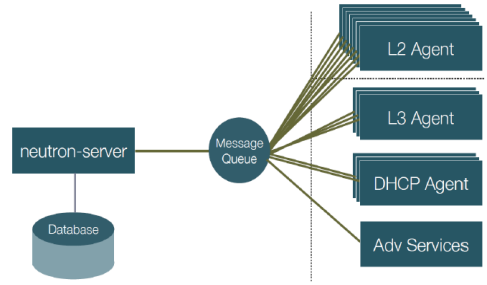


Multi-tenant virtual network

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Logical Architecture of Neutron: Components

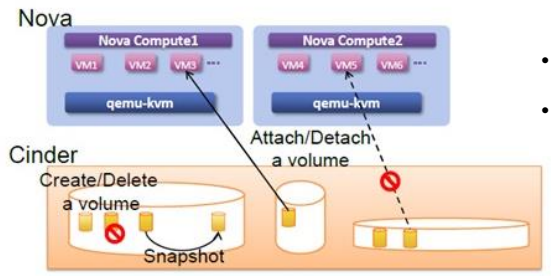
- Neutron-Server
- Core plugin
- Various advanced service plugins
  - L3 service plugin
  - LB service plugin
  - Firewall
  - VPN
- Various agents
  - L2 (ovs-agent)
  - L3 agent
  - DHCP agent
  - Metadata agent
  - Agile Controller-DCN



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cinder Overview

- The storage resource management system is responsible for providing persistent block storage resources for VMs.
- Encapsulates the backend storage resources and provides a unified API externally.
- The main core is volume management, allowing operations on the volume, volume type, and volume snapshot.



- Create, delete and snapshot a volume.
- Attach and detach a volume.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cinder

Type	Block Storage		Object Storage
Name	Temporary storage	Block storage Cinder	Object storage Swift
Purpose	VM root disk and other disks	Mounted to the VM and provides additional disk space to the VM.	Provides backup and archiving data storage space for the VM. Used to store VM images.
Similar Windows OS	Root disk: Driver C Disk: D disk	USB	NAS archive server
Life cycle	Storage is created during VM start and exists in the VM life cycle. Its size depends on the flavor (ECS type).	Storage is created by users, limited by the user quota. It exists until the user is deleted.	

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cinder Architecture

- Three major components
  - Cinder-api externally provides the Cinder REST API.
  - Cinder-scheduler allocates storage resources.
  - Cinder-volume encapsulates driver. Different drivers control different backend storage systems.
- RPC between components is achieved using the message queue.
- Cinder development is concentrated in the scheduler and driver to provide more scheduling algorithms, more functions, and more backend storage systems.
- The volume metadata and status are saved in the database.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Functions of Cinder

No.	Object	Action
1	Volume	Create a volume.
2		Use an existing volume to create a volume (clone).
3		Expand a volume.
4		Delete a volume.
5	Volume – VM	Mount a volume to a VM.
6		Detach a volume from a VM.
7	Volume – snapshot	Create a volume snapshot.
8		Use an existing volume snapshot to create a volume.
9		Delete a snapshot.
10	Volume – mirror	Create a volume from an image.
11		Create an image from a volume.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Ceilometer Overview

- In OpenStack, Ceilometer can collect operation records and the system running status at the IaaS layer to generate metering data. In FusionSphere OpenStack, Ceilometer provides monitoring and alarm data for upper-layer O&M components. In public cloud scenarios, Ceilometer can be used for charging.
- Ceilometer has a flexible architecture, supports distributed deployment, and has high scalability.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Objective: In metering, Ceilometer aims to provide unified resource usage data collection function for upper-layer charging, settlement, and monitoring applications.

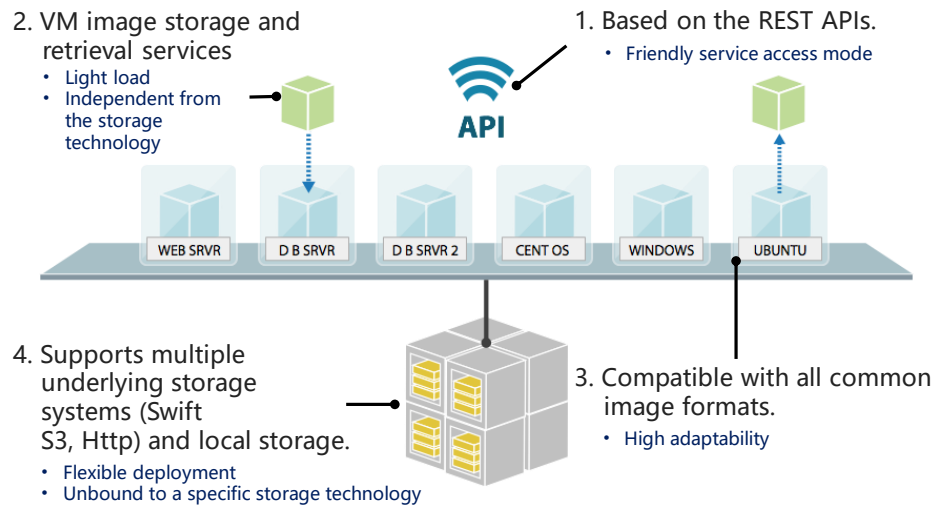
Glance Overview

- Glance is an image service component of OpenStack. It provides VM image discovery, registration, and access to services.
- Glance provides the RESTful APIs to query VM image metadata and obtain the image.
- VM images made available through Glance can be stored in a variety of locations from simple filesystems to object-storage systems like the OpenStack Swift project.
- Glance provides the REST APIs to support the following image operations:
  - Querying
  - Registering
  - Uploading
  - Obtaining
  - Deleting
  - Access right managing

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Glance has two versions of REST APIs: REST API V1 and V2. The two versions are different from each other.
- REST API V1 provides only basic image and member operation functions: Create, delete, and download an image; Query and update the list and detailed information; Create, delete, and list tenant members.
- REST API V2 supports all functions of V1 and the following functions:
  - Add, delete, and modify an image location.
  - Perform operations on the metadata namespace.
  - Perform operations on the image tag.
- Both REST API V1 and V2 support the image store.
- By default, Glance Cli and Horizon use the REST API V1.

Glance - Image Management



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The principle and implementation of Glance are easier and more direct than other components.
- Image management in multiple data centers
- Unified service interface and private and open images
- Supports multiple storage backends through the driver and supports isolation of image storage space between tenants (only Swift).
- Supports comprehensive image formats.
- New features such as local cache of images, multiple data center (MDC) replication, multiple locations, inremental images, and Cinder-Volume backend simplify image management during large-scale MDC deployment, improve the storage space usage and VM provisioning efficiency, and support image backup.

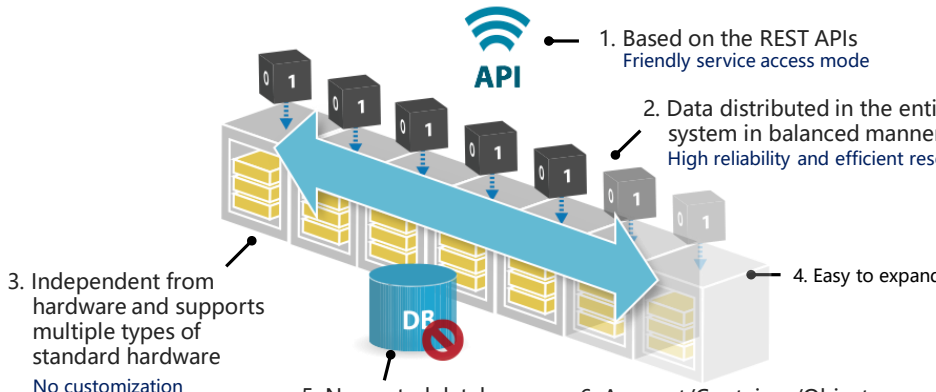
Swift Overview

- Swift was originally a high-availability (HA) distributed object storage service developed by Rackspace. In 2010, Swift was contributed to the OpenStack open-source community as the one of the first core sub-projects providing the VM image storage service for the Nova sub-project. Swift is constructed on the standard hardware storage infrastructure that is cheap without redundant array of independent disks (RAID). Swift achieves HA and scalability by using consistency hash and data redundancy at the software layer and compromising a certain degree of data consistency. It supports the multi-tenant mode, container, and object read/write operations, which is suitable for resolving unstructured data storage problems in Internet application scenarios.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- This project is developed based on Python. It uses the Apache 2.0 license and can be used to develop commercial systems.

Swift Storage Virtualization - Object Storage



1. Based on the REST APIs
   Friendly service access mode
2. Data distributed in the entire system in balanced manner
   High reliability and efficient resource utilization
3. Independent from hardware and supports multiple types of standard hardware
   No customization requirements for dedicated hardware devices
4. Easy to expand
5. No central database.
   No single-point performance bottleneck or single point of failure (SPOF) risks
6. Account/Container/Object
   The three-level storage structure does not need the file system and has N ($N \geq 3$) copies.
   High data reliability

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

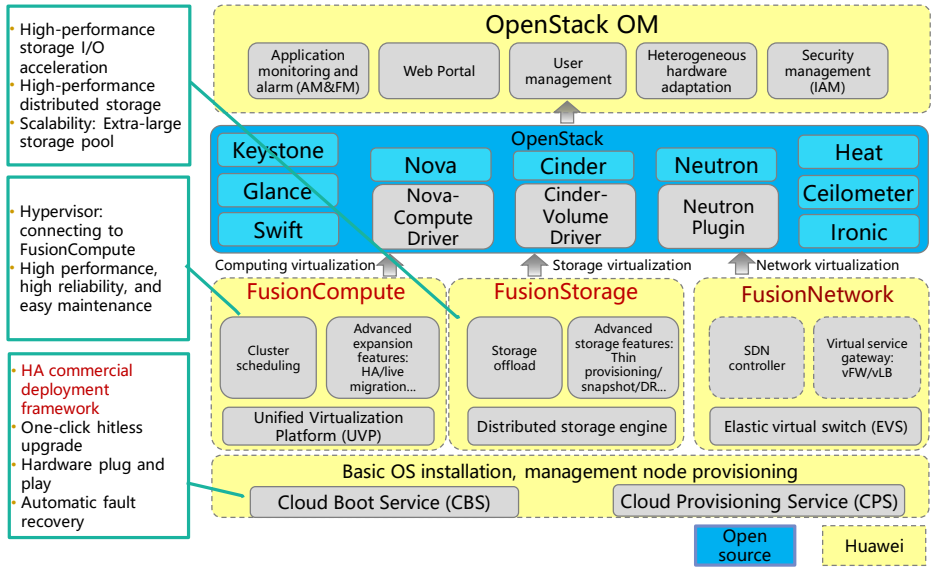
Contents

1. OpenStack Background
2. OpenStack System Architecture
3. Functions and Features of OpenStack
4. **Huawei FusionSphere OpenStack Enhancements**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

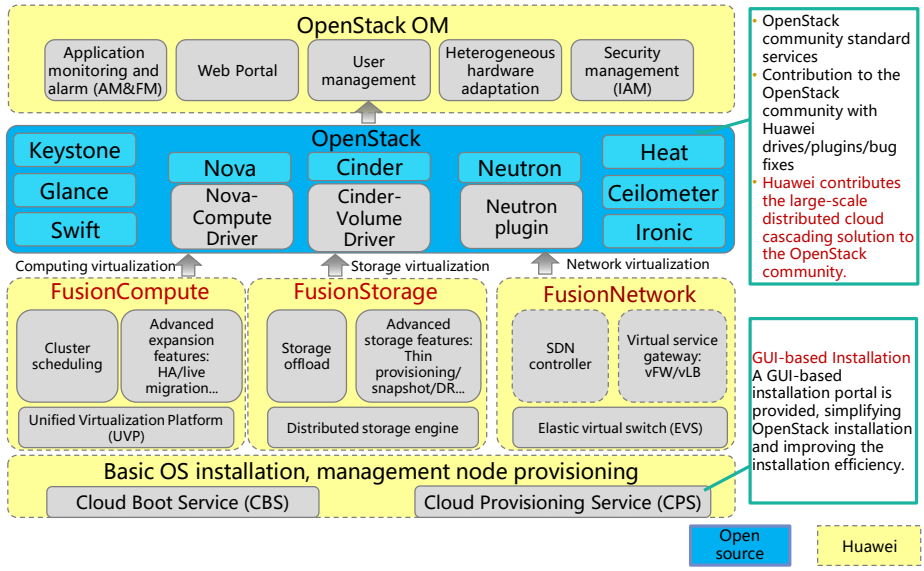
FusionSphere OpenStack Commercial Enhancements (1)



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The FusionSphere cloud platform solution is constructed based on OpenStack community. With the OpenStack plugin mechanism, FusionCompute, FusionStorage, and FusionNetwork plugins can seamlessly interconnect with native OpenStack. FusionSphere is an OpenStack-based commercial cloud platform.

FusionSphere OpenStack Commercial Enhancements (2)



- OpenStack community standard services
- Contribution to the OpenStack community with Huawei drives/plugins/bug fixes
- Huawei contributes the large-scale distributed cloud cascading solution to the OpenStack community.

GUI-based Installation
A GUI-based installation portal is provided, simplifying OpenStack installation and improving the installation efficiency.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FusionSphere OpenStack Commercial Enhancement Features

- Reliability
  - System reliability includes the reliability of the entire system, a single device, and data. The cloud platform employs the distributed architecture, which improves the reliability of the entire system and lowers reliability requirements for a single device.
- Availability
  - System availability is represented by such features as redundancy, high-availability clusters, and loose coupling between applications and underlying devices. This solution employs various measures, including hardware redundancy, link redundancy, and application fault tolerance (FT), to ensure system availability.
- Security
  - System security complies with the industry security specifications and is designed to ensure the security of data centers. It focuses on the security of networks, hosts, virtualization, and data.

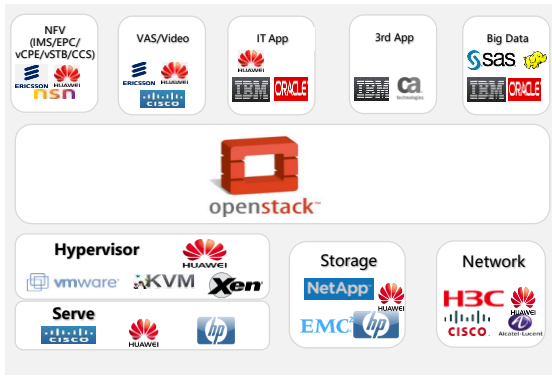
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FusionSphere OpenStack Commercial Enhancement Features

- Maturity
  - FusionSphere OpenStack uses the architecture solution, hardware, and software that are tested in large-scale commercial practices and adopts the IT management solution that complies with the Information Technology Infrastructure Library (ITIL) standards to ensure the solution maturity.
- Advancement
  - Customer benefits are highlighted using the advanced cloud computing technology and idea. Advanced technologies and modes such as virtualization and dynamic resource deployment are used with services, ensuring the validity and applicability of advanced technologies and modes.
- Scalability
  - DC resources must be flexibly adjusted to meet actual service load requirements, and the IT infrastructure must be loosely coupled with service systems. Therefore, users only need to add IT hardware devices when service systems require capacity expansion.
- Openness
  - Built upon the mainstream open-source cloud platform, FusionSphere OpenStack, the solution embraces the industry ecosystem and minimizes the investments on resource pools. With close cooperation with ISVs in the industry, the solution fully unleashes the power of cloud-based applications.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FusionSphere: OpenStack-based Open Cloud Service and Cloud Management Platform



Platinum Member in OpenStack Community

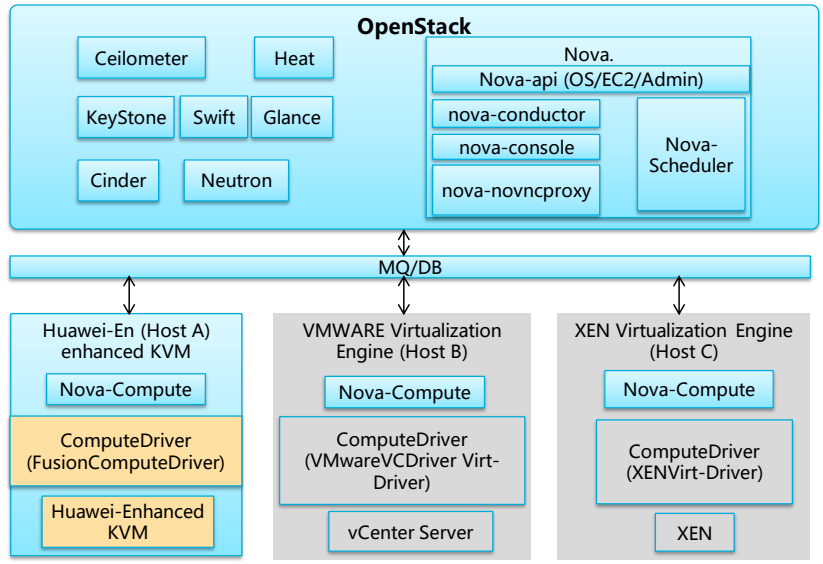
- **Standard OpenStack APIs**
  - Developed based on native OpenStack APIs.
  - Quickly adapts to new OpenStack releases.
- **Support for third-party vendors**
  - Compatible with the OpenStack ecosystem chain.
  - Supports heterogeneous hypervisors and hardware devices.
- **SOA-based loosely coupled architecture**
  - Compute, storage, and network resources are decoupled from each other.
  - Supports hybrid networking with multi-vendor resource pools.

- Huawei, a platinum member of OpenStack community, acquired a seat in the BOD in early 2016.
- Huawei ranked second in terms of contribution to OpenStack community in 2019.
- With the help from the OpenStack and KVM communities, Huawei cloud platform has been widely accepted by major hardware and software vendors in the OpenStack ecosystem chain, and tends to be compatible with more products.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

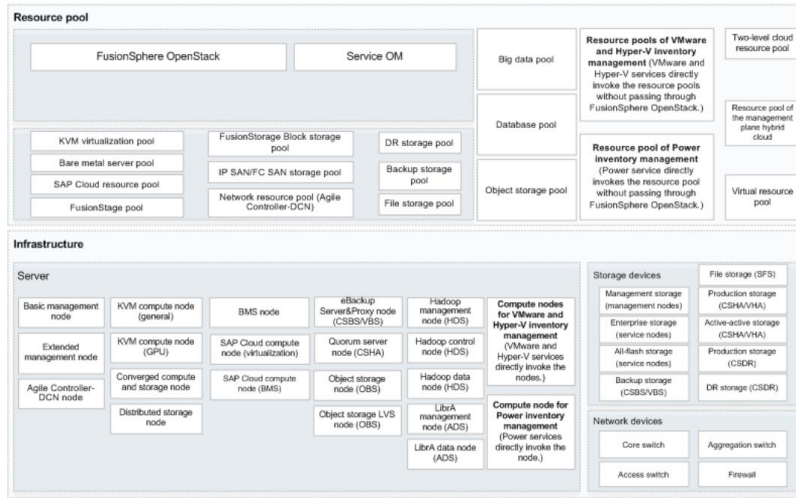
- This slide focuses on the fact that FusionSphere is developed based on standard OpenStack APIs. It is open and compatible with all OpenStack-based products, rather than a closed commercial solution.

OpenStack-based Plugin Enhancements



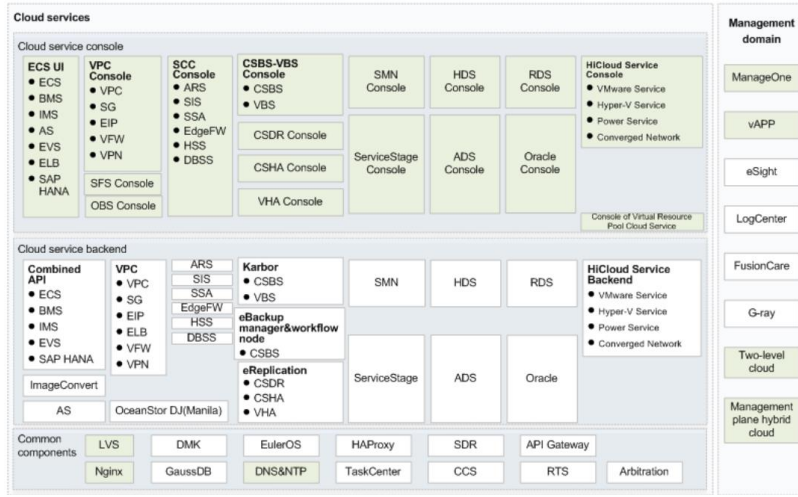
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

FusionSphere OpenStack Architecture



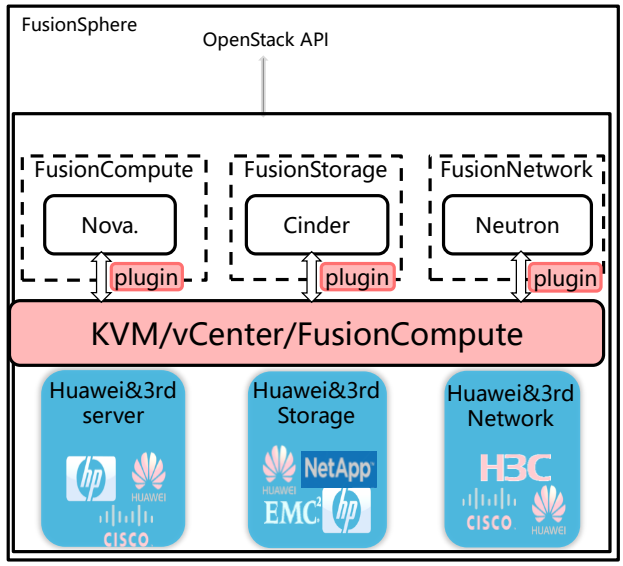
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cloud Service Architecture for FusionCloud



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Open Architecture Supports Heterogeneous Virtualization



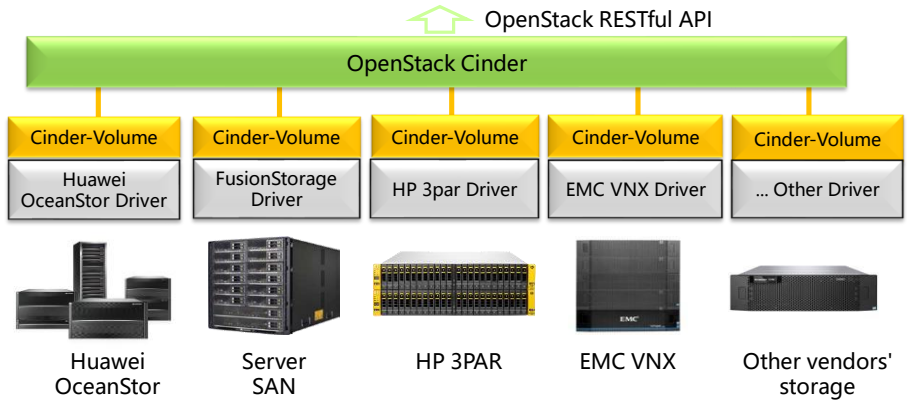
FusionSphere provides a full set of OpenStack APIs, modular deployment interfaces, and infrastructure hardware management interfaces to communicate with northbound components.
Southbound heterogeneous compatibility:

- Supports compute devices, including general-purpose IT hardware and IT appliances.
- Huawei's hypervisor is compatible with industry-leading virtualization software such as KVM, vCenter, and FusionCompute.
- Storage devices of different vendors, which can connect to FusionSphere using each vendor's own driver. FusionSphere also supports local storage, IP SAN, and distributed storage modes.
- Network and security devices, which can connect to FusionSphere using each vendor's own neutron plugin and driver.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Open Architecture Supports Heterogeneous Storage Devices

- Cinder provides persistent block storage services that provision resources on demand through unified interfaces (similar to Amazon EBS).
- Various backend storage devices (local storage, network storage, FC SAN, and IP SAN) can be accessed using drivers.
- Northbound APIs: OpenStack Cinder APIs support centralized management of storage resources.
- Southbound APIs: Different Cinder-Volume drivers are compatible with storage devices of different vendors to prevent vendor lock-in.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Summary

- OpenStack Background
- OpenStack System Architecture
- Functions and Features of OpenStack
- Huawei FusionSphere OpenStack Enhancements

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. True or False

   Heat of OpenStack provides the resource orchestration capability. ( )

2. Multiple Choice Question

   Which of the following are true of OpenStack? ( )

   Open-source project

   Ceilometer provides network services.

   Glance provides the image service.

   Horizon provides the UI service in OpenStack.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - 1.T
  - 2.ACD

More Information

- Huawei's official website
  - Enterprise business: http://e.huawei.com/en/
  - Technical support: http://support.huawei.com/enterprise/en/cloud-computing/fusionsphere-openstack-pid-21100528
  - Online learning: http://learning.huawei.com/en/
- Documentation tool
  - HedEx Lite

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei Learning website
  - http://support.huawei.com/learning/en/newindex.html
- Huawei Support case library
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com



Huawei ManageOne Solution

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Foreword

- This course describes ManageOne in terms of its application scenarios, architecture, and features. After learning this course, you will have a general knowledge of ManageOne.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objectives

- After learning this course, you will be able to:
  - Understand the ManageOne solution architecture.
  - Understand the ManageOne system networking.
  - Describe the typical ManageOne features.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

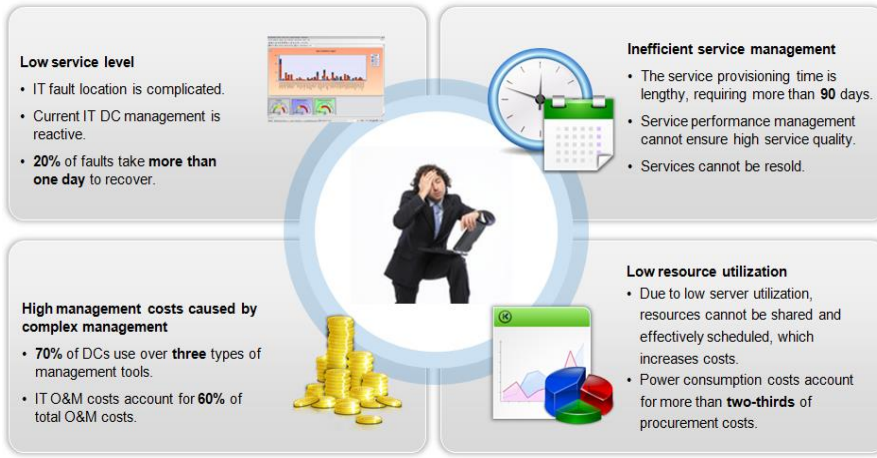
Contents

1. **Trends and Challenges of Cloud Data Center Management**
2. ManageOne Cloud Management Solution
3. ManageOne Cloud Management Platform Features

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

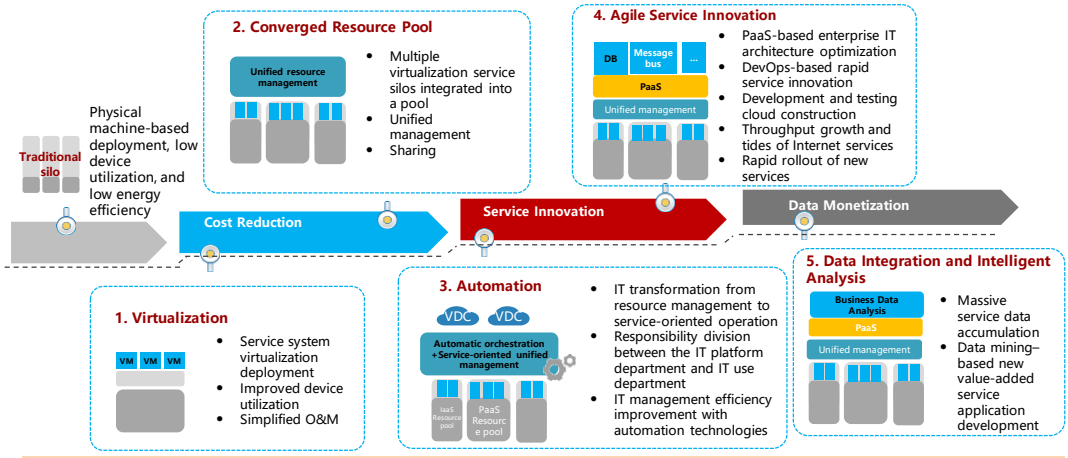
Challenges to DC Management



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Poor service quality:
  - IT problems are difficult to be located. For more than 20% IT problems, over one day is required for locating each problem.
  - Traditional DCs have no unified and open management platform. As a result, resources cannot be centrally allocated to support diversified applications.
- Inefficient service management
  - Services for traditional DCs are deployed from the bottom layer, which requires a long hardware installation phase, complex basic configuration, and more than 30 days for service rollout. This situation cannot meet requirements of rapid service development.
- Complex management and high management costs
  - Traditional DCs have various standards and planning methods. As a result, hardware resources cannot be centrally managed or shared.
  - Network systems become increasingly complex. Therefore, a large number of professional O&M personnel are required to meet customer requirements.
  - System maintenance consumes a lot of resources. According to statistics, more than 70% of IT budgets are used for system maintenance, leaving insufficient investment for deploying new IT systems.
- Low resource utilization
  - Resource usage in traditional DCs is generally less than 20%, which indicates that a large number of resources are wasted. Meanwhile, servers in the idle state are consuming power and increasing customers' cost.

Evolution and Demands of Enterprise IT Cloud Migration



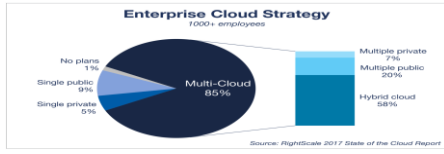
As a unified platform for customers to use, manage, and operate their cloud resource pools, ManageOne needs to consider the coexistence of existing and new IT infrastructure and different operation modes during enterprise IT cloud migration.

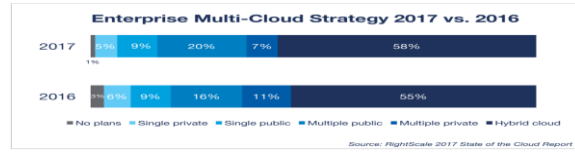
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Trends of Cloud Data Center Management

Trend 1: Enterprise multi-cloud management.

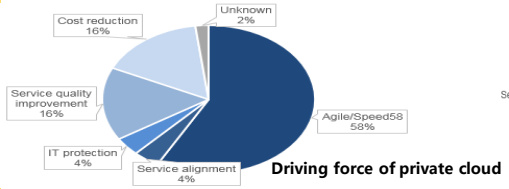
Multi-cloud management, especially the hybrid cloud, becomes a new growth point.



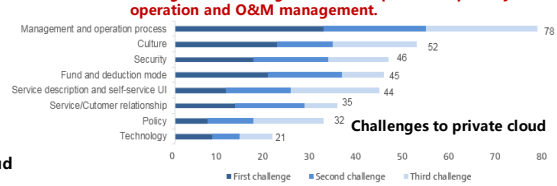


Trend 2: The private cloud has entered the implementation phase, which focuses on O&M.

Key factors of private cloud implementation are cloud management and organization adaptation, especially operation and O&M management.



Driving force of private cloud



Challenges to private cloud

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Cloud Management Concepts of Huawei



Fine-grained Operations

- Diverse Cloud Services
- Multi-Level VDCs
- Metering & Charging
- Application Orchestration



Intelligent O&M

- Unified Monitoring
- Intelligent Fault Locating
- Visualized O&M
- Automated O&M



Centralized Management

- One Cloud Multi-Pool
- Unified Configuration Operation Center
- Hybrid Cloud
- Unified Multi-Cloud Monitoring

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

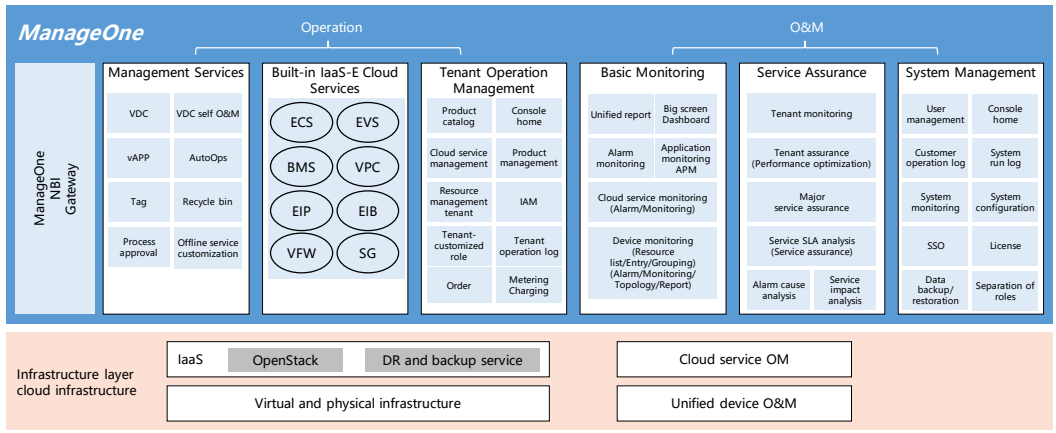
Contents

1. Trends and Challenges of Cloud Data Center Management
2. **ManageOne Cloud Management Solution**
3. ManageOne Cloud Management Platform Features

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

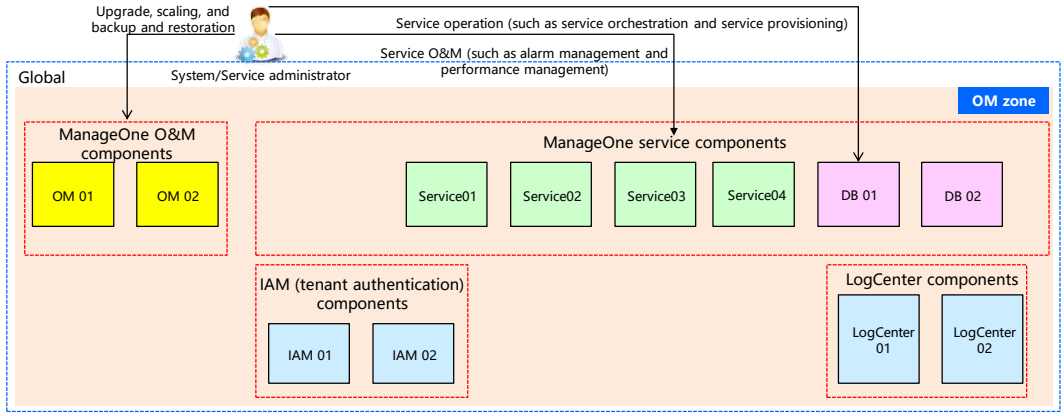
HUAWEI

ManageOne Functional Architecture



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Typical Networking Schemes of ManageOne

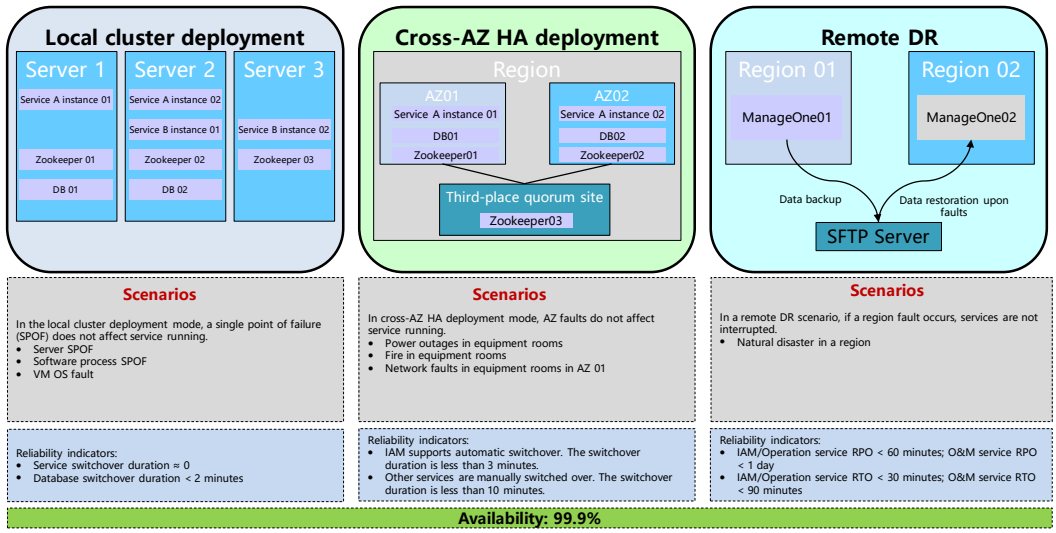


Networking description:

1. ManageOne is deployed in the Global OM zone. A total of 12 VMs are required, including two O&M components, six ManageOne service components, two IAM components, and two LogCenter components.
2. Management scale levels:
   Micro-scale: VMs ≤ 200; small-scale: 200 < VMs ≤ 1000; medium-scale: 1001 < VMs ≤ 2000; large-scale: 2001 < VMs ≤ 5000; super large-scale: 5001 < VMs ≤ 10,000
3. IAM management specifications: 1000 tenants, 20 concurrencies, and concurrency calling interval of 1 minute
4. LogCenter specifications: 10 logs per second per node per service

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

ManageOne Deployment Modes



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- ManageOne instances can be deployed in three modes: local cluster deployment, cross-AZ HA deployment, and remote DR.

Contents

1. Trends and Challenges of Cloud Data Center Management
2. ManageOne Cloud Management Solution
3. **ManageOne Cloud Management Platform Features**
   - Fine-grained Operations
   - Intelligent O&M
   - Centralized Management

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Fine-grained Operations

- Diverse Cloud Services
  - A complete catalog of cloud services (IaaS, PaaS, and SaaS)
  - Consistent operation experience for self-developed and third-party cloud services
- Multi-Level VDCs
  - Up to five levels of VDCs, well suited to complex enterprises
  - Multi-level approval process
  - Fine-grained authorization, precise control of user rights
- Application and Automation
  - Graphical template orchestration
  - One-click application deployment



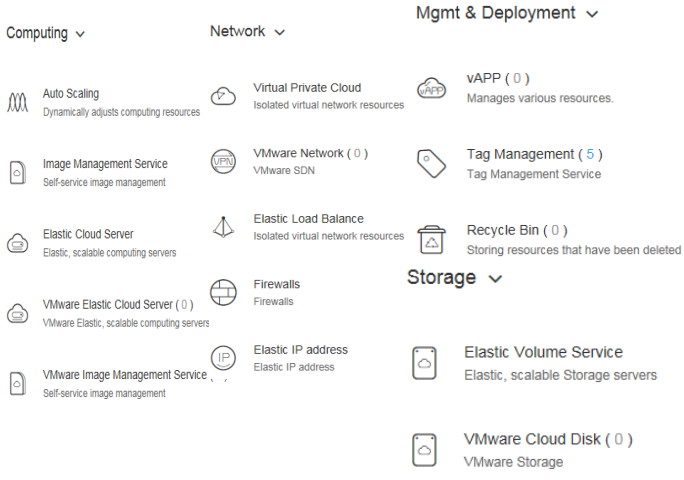
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Fine-grained Operations

1. Diverse Cloud Services
2. Multi-Level VDCs
3. Application and Automation

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Various Cloud Services



Customer Scenarios (Why)

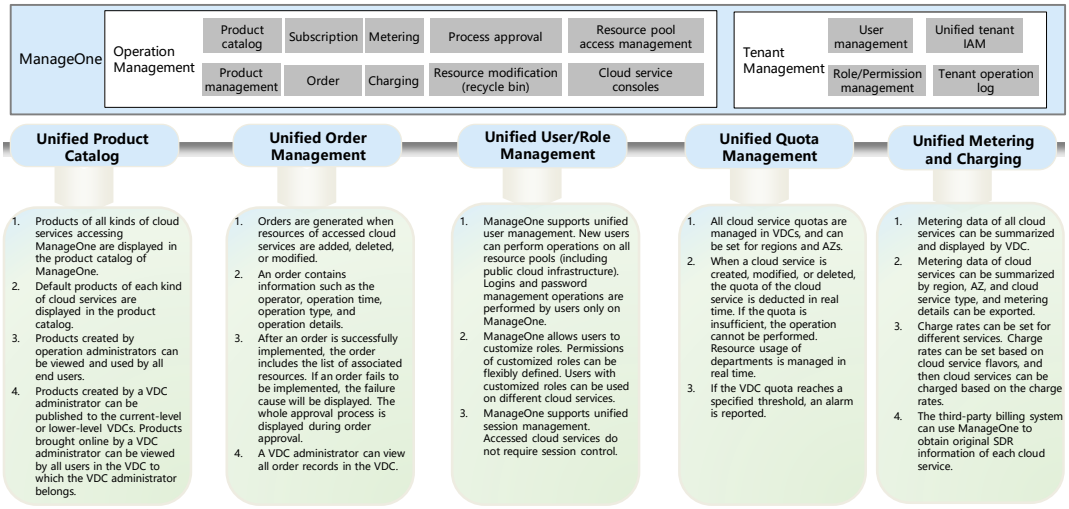
- Various cloud services are provided to meet the configuration, application, and usage requirements of all service-oriented instances.

Customer Benefits (How)

- All computing, network, and storage cloud computing features and resources are provided in the form of services.
- Resource application, use, and recycling are convenient.
- Monitoring, auditing, and service orchestration (vAPP) services are provided.
- Heterogeneous VMware virtualization service provisioning is supported.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Various Cloud Services — Unified Operation



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

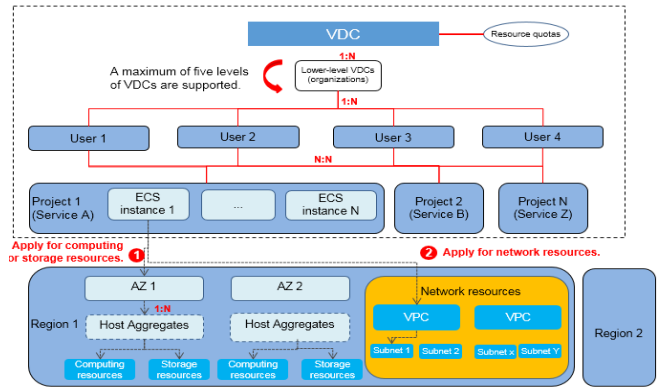
Fine-grained Operations

1. Various Cloud Services
2. Multi-Level VDCs
3. Application and Automation

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs

- VDC (Virtual Data Center)
  - A virtual data center (VDC) is a resource allocation unit that matches the structure between an enterprise and its organizations. In VDCs, user management, quota management, project management, product definition, resource provisioning, and service assurance are supported. Multi-level VDCs can be created to meet the requirements for multi-level operation scenarios. For example, a group includes multiple subsidiaries, and each subsidiary includes multiple lower-level departments. In the e-Government Cloud scenario, there are multiple government offices.



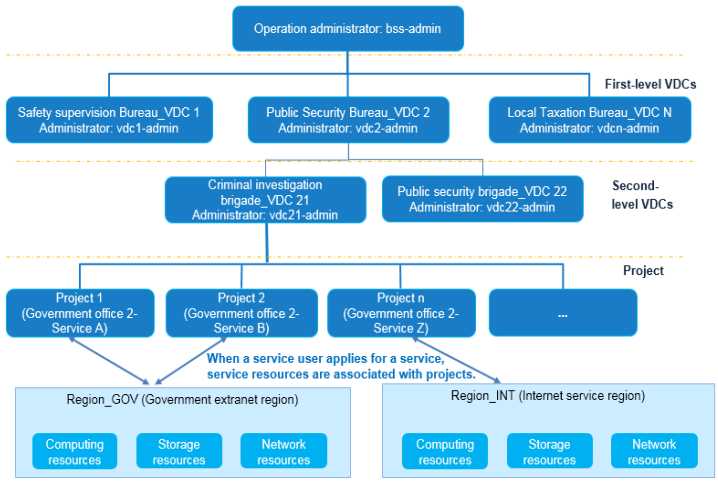
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs - Roles and Scenarios in the Operation Field

	Objectives	Focuses	Management Scope
 Operation administrator	■ Accesses resource pools. ■ Manages tenants, including creating, approving, and setting first-level VDCs. ■ Manages all products. ■ Manages system users and roles, and configures the system information (such as charge rates and third-party authentication).	Resource pool management System user management Global product management Tenant/First-level VDC management	■ Resource pool management: Access resource pools change the resource pool status, and perform other operations to manage underlying resources. ■ Product management: Define and publish all products, bring all products online, and perform othe product management operations. ■ System configuration: Configure system information, such as users, roles, charge rates, and tags.
 VDC administrator	■ Manages VDCs, including creating, modifying, deleting VDCs, and managing VDC quotas, metering data, and roles. ■ Manages products, including applying for, modifying, extending, and deleting products in VDCs. ■ Manages orders, including creating, querying, and exporting orders, and subscribing to emails. ■ Manages approval processes. ■ Manages VDC application templates, including orchestrating and debugging application templates and publishing the templates to VDC service catalogs. ■ Applies for, modifies, extends, deletes services, and maintains obtained resources.	Current-level and lower-level product management Current-level and lower-level VDC management Service life cycle VDC Self O&M	■ Product management: Define and publish current-level and lower-level products, bring current-level and lower-level products online, and perform other product management operations. ■ VDC management: Define current-level and lower-level VDCs, and manage resource quotas, projects, users, roles, metering data, and logs. ■ Order management: Apply for, modify, extend services in current-level and lower-level VDCs, and perform other management operations.
 VDC operator	■ Applies for, modifies, extends, and deletes services. ■ Queries, operates, and maintains obtained resources. ■ Queries application orders.	Life cycles of own services Management of own orders Life cycles of services in the same VDC	■ Order management: Apply for, modify, extend own services, and perform other management operations. ■ Service life cycle: Query, operate, and maintain own projects and projects in the same VDC.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs - Operation Organization Division

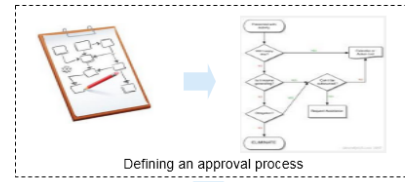


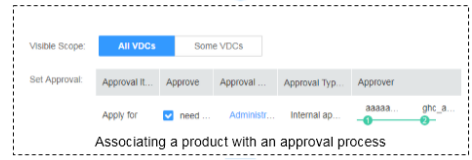
- VDC planning:
  - The government offices want to delegate the operation administrator to manage operation. The operation administrator allocates a tenant (create a first-level VDC by default) for each government office.
  - A first-level VDC administrator determines whether to create a lower-level VDC based on the organization's structural requirements. Quotas can be set separately for VDCs at each level.
  - VDC administrators can view provisioned resources and manage and maintain resources in the VDCs they belong to, and their lower-level VDCs.
- Project planning:
  - Projects are created in VDCs at each level and managed by VDC administrators.
  - A user can be associated with projects in different lower-level VDCs of a first-level VDC.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs - Multi-Level Approval

- Independent approval processes can be defined on ManageOne. Operation administrators and VDC administrators can define approval processes. Approval processes published by operation administrators are globally visible. Approval processes published by VDC administrators are visible to the VDCs to which the VDC administrators belong and their lower-level VDCs.
- An approval process supports a maximum of five levels of approvals. Multiple approvers can be set at each level.
- An approval process defined on ManageOne can be associated with a third-party work order system. That is, after the approval process is started, ManageOne sends an approval request to the third-party work order system.


Defining an approval process



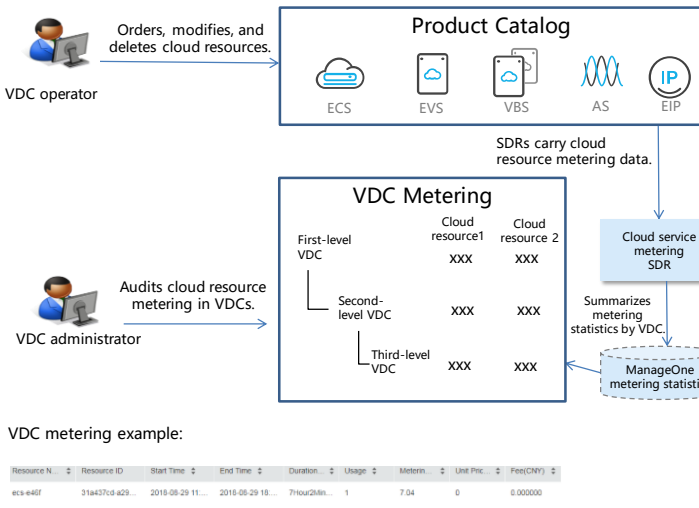
Associating a product with an approval process



Applying for and approving products

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs - VDC Metering



VDC metering example:

Resource N...	Resource ID	Start Time	End Time	Duration	Usage	Meterin...	Unit Pric...	Fee(CNY)
ecs-e46f	31a437cd-a29...	2018-08-29 11:...	2018-08-29 18:...	7Hour2Min...	1	7.04	0	0.000000

Customer Scenarios (Why)

VDC metering data is provided to facilitate audits and control of cloud resources consumed by services.

Application Scenarios (What)

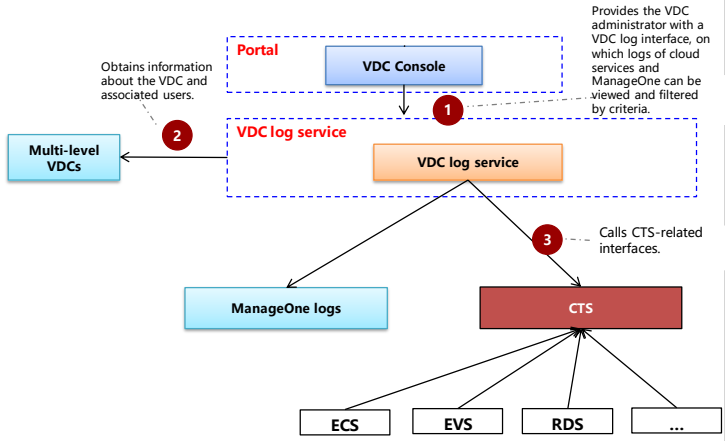
- Enterprises allocate VDCs to their internal organizations, and IT resources usage statistics are collected by VDC for internal settlement.
- In the e-Government scenario, VDCs are leased to enterprise tenants, and pay-per-use pricing is used.

Customer Benefits (How)

- Metering files are generated for each cloud service based on the resource life cycle. They are stored centrally, and displayed on the management platform.
- Cloud resource metering data generated by cloud services is summarized by VDC.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Level VDCs - VDC Logs



Customer Scenarios (Why)

Operation logs can be queried based on VDCs, which helps VDC administrators audit behavior of users in an organization.

Application Scenarios (What)

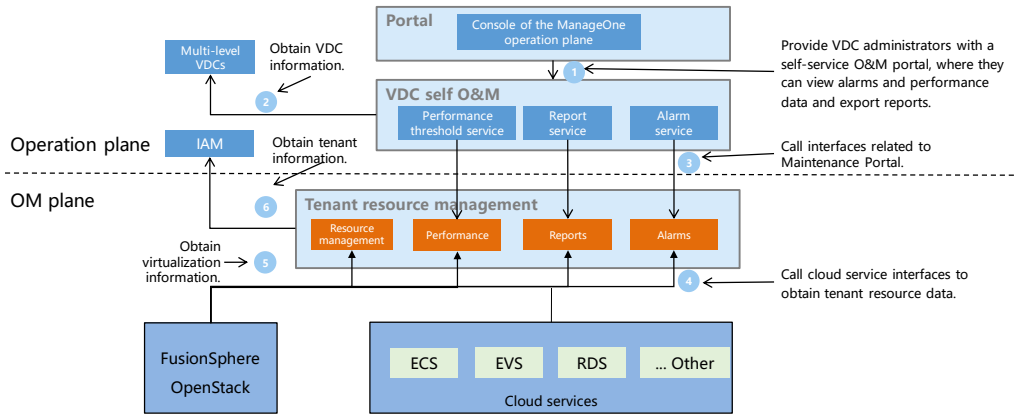
- Trace users' operations on the cloud resources in an organization.
- Assist location and eliminate faults of cloud resources in the organization.

Customer Benefits (How)

- VDC logs are generated by ManageOne and cloud services in CTS.
- The VDC log service classifies and summaries logs by VDC.
- Only VDC administrators can view VDC logs.
- VDC logs can be stored for 1 year.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

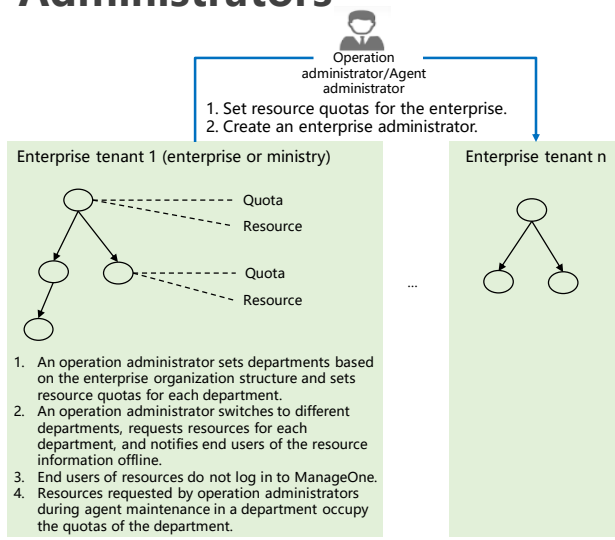
Multi-Level VDCs - VDC Self O&M



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- ManageOne Maintenance Portal obtains all O&M data and monitors resources by tenant.
- Self O&M allows VDC administrators to set and monitor resources, alarms, and performance thresholds for VDCs to which they belong as well as their lower-level VDCs.
- The O&M data and basic functions are provided by ManageOne Maintenance Portal. The self-service O&M function of ManageOne Operation Portal only displays service O&M data by VDC and provides the portal for managing the data.

Multi-Level VDCs - Agent Maintenance by Administrators



1. An operation administrator sets departments based on the enterprise organization structure and sets resource quotas for each department.
2. An operation administrator switches to different departments, requests resources for each department, and notifies end users of the resource information offline.
3. End users of resources do not log in to ManageOne.
4. Resources requested by operation administrators during agent maintenance in a department occupy the quotas of the department.

Customer Scenarios (Why)

In the e-Government Cloud scenario, operation administrator applies for resources for resource users on ManageOne. Resource users do not log in to ManageOne. An administrator does not need to repeatedly change accounts to perform the agent maintenance. Therefore, administrators' workload can be reduced.

Application Scenarios (What)

An operation administrator creates departments of an enterprise based on the enterprise organization structure, and sets resource quotas for each department.

Customer Benefits (How)

- Operation administrators or agent administrators must have required permissions to perform operations on all VDCs or certain VDCs.
- When applying for a resource, an agent administrator needs to determine the organization to which the resource belongs based on the project of the resource and deduct the quota of the organization.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

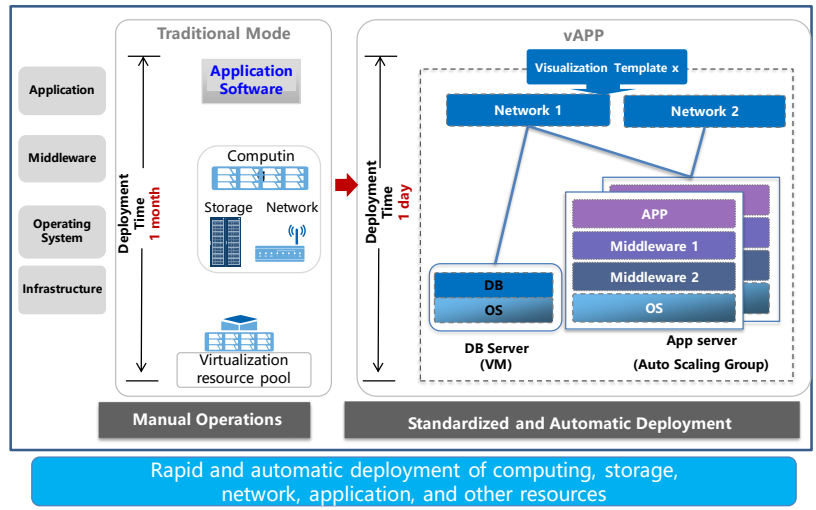
Fine-grained Operations

1. Various Cloud Services
2. Multi-Level VDCs
3. Application and Automation

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Application and Automation - Application Orchestration (vAPP)

The vAPP service allows users to drag diagram elements on the visualized orchestration interface to quickly and automatically deploy compute, storage, network, and application resources and serves users as a whole, helping quick service rollout. In addition, it allows users to define AS policy to automatically scale ECSs, achieving load balancing and full resource utilization.

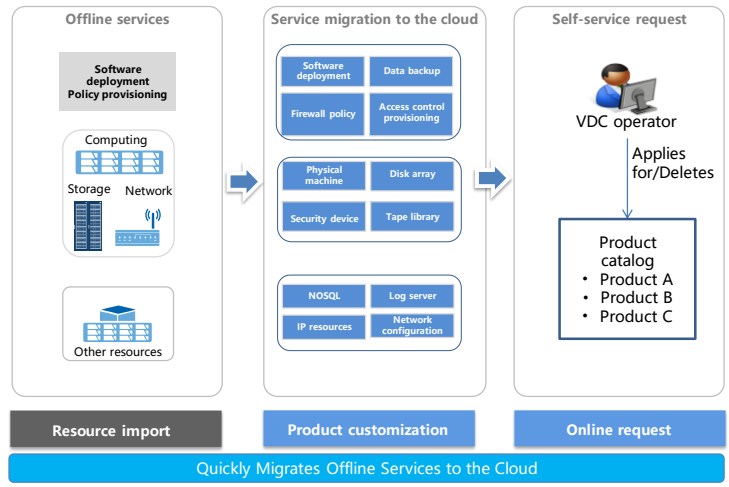


Rapid and automatic deployment of computing, storage, network, application, and other resources

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

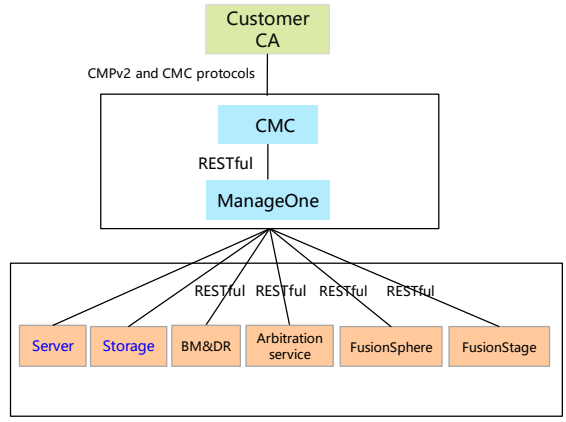
Application and Automation - Offline Service Customization

Offline Service Customization allows customers to customize services, migrate their offline services to the cloud, and request, approve, and provision services implemented offline.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Application and Automation - Unified Certificate Management



Cloud services interconnect with the ManageOne certificate management module in unified authentication mode that is based on the 10 unified principles, and implement service interaction through RESTful interfaces.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Trends and Challenges of Cloud Data Center Management
2. ManageOne Cloud Management Solution
3. **ManageOne Cloud Management Platform Features**
   - Fine-grained Operations
   - Intelligent O&M
   - Centralized Management

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent O&M

- Unified Monitoring
  - Provides all-round O&M monitoring from physical devices to cloud resource pools, from cloud services to big data, and from system resources to tenant resources.
- Intelligent Fault Locating
  - Displays alarms precisely, compresses 80% of repeated alarms and correlative alarms.
  - Centrally processes faults based on scenarios, improving fault locating efficiency.
- Visualized O&M
  - Scenario-specific preset dashboards and reports
  - Various customization capabilities



Fine-grained Operations

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

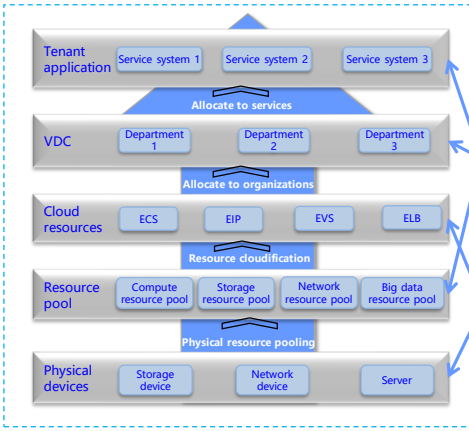
Intelligent O&M

1. Unified Monitoring
2. Intelligent Fault Locating
3. Visualized O&M

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring

- Unified Monitoring can be used to monitor objects such as physical devices, resource pools, cloud resources, VDCs, and tenant applications, proactively monitor the status of cloud DCs, and transform from single-device monitoring to service-based analysis. This function helps enterprises reduce IT costs and improve O&M efficiency.



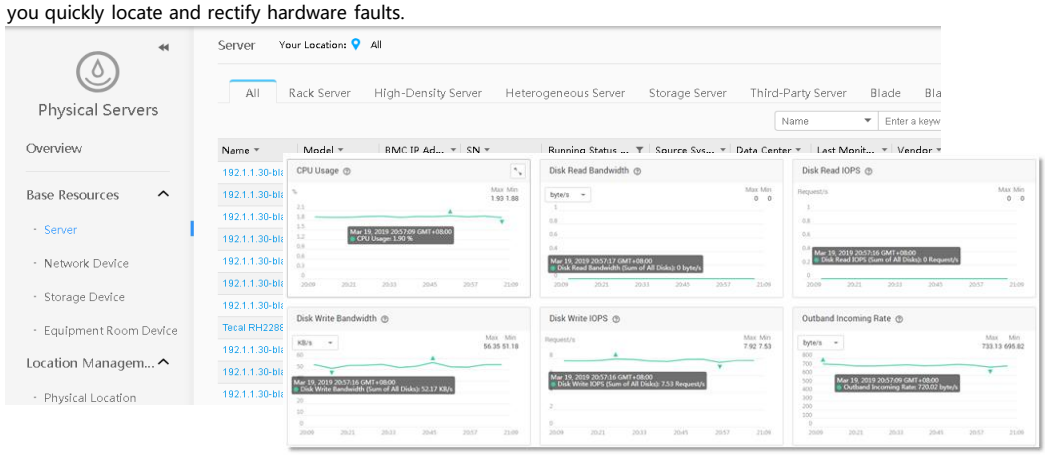
Overview of the monitoring system:

1. **Physical device monitoring:** Centrally monitor and manage hardware devices such as data center servers, storage devices, and network devices; and provide comprehensive monitoring, including alarms, resources, topologies, and performance, helping users quickly locate and rectify hardware faults.
2. **Resource pool monitoring:** Take advantage of unified monitoring and analysis of computing, storage, network, and big data resource pools to help you efficiently use resources, identify potential risks and problems, and provide improvement measures or suggestions.
3. **Cloud resource monitoring:** Enjoy comprehensive monitoring of cloud resource alarms, resources, topologies, and performance, helping you quickly demarcate and locate cloud resource assurance problems.
4. **VDC monitoring:** VDC-based comprehensive analysis and evaluation capabilities help you use resources appropriately and improve resource usage.
5. **Tenant application monitoring:** Monitor resources from the perspective of applications, continuously evaluate application resource usage from aspects such as capacity and load, and provides all-round assurance for key services.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Physical Device Monitoring

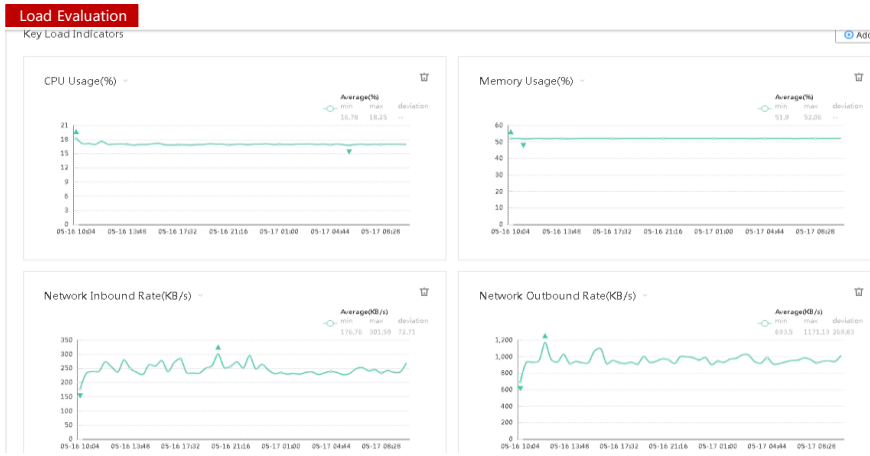
- Centrally monitor and manage hardware devices such as data center servers, storage devices, and network devices; and provide comprehensive monitoring capabilities. Monitor alarms, resources, topologies, and performance, helping you quickly locate and rectify hardware faults.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

Unified Monitoring - Resource Pool Monitoring

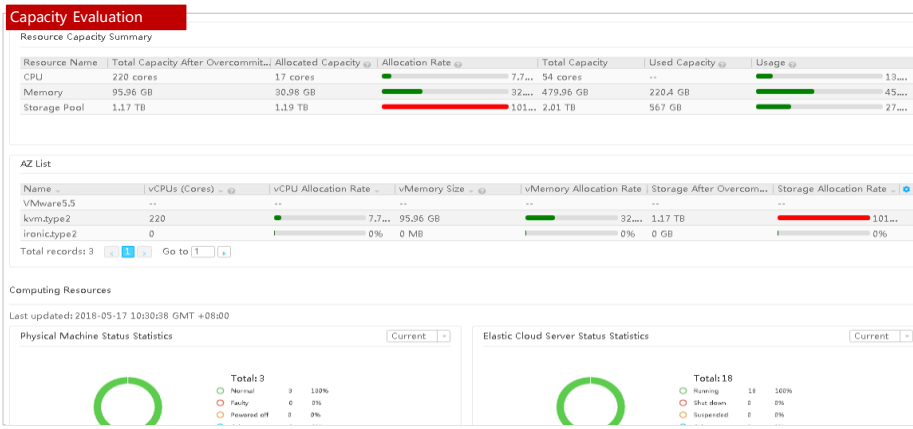
- Resource Pool Monitoring continuously evaluates the resource pool load based on KPIs, and provides root causes of high loads.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Resource Pool Monitoring

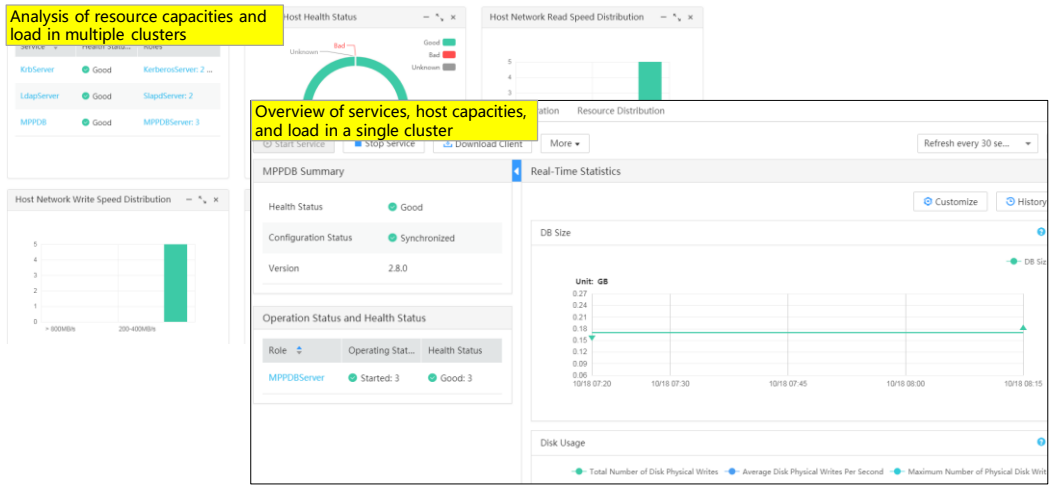
- Evaluate the capacity of compute, storage, and network resource pools at multiple layers, such as region, resource pool, AZ, and cluster, to instruct administrators to plan capacity and capacity expansion, thereby improving resource utilization.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Big Data Resource Pool Monitoring

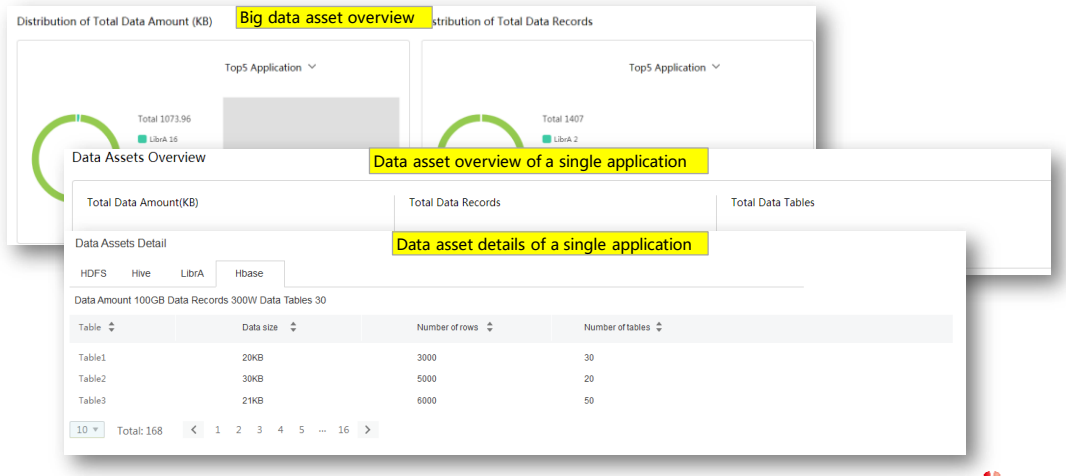
- Multiple big data clusters are analyzed.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Big Data Resource Pool Monitoring

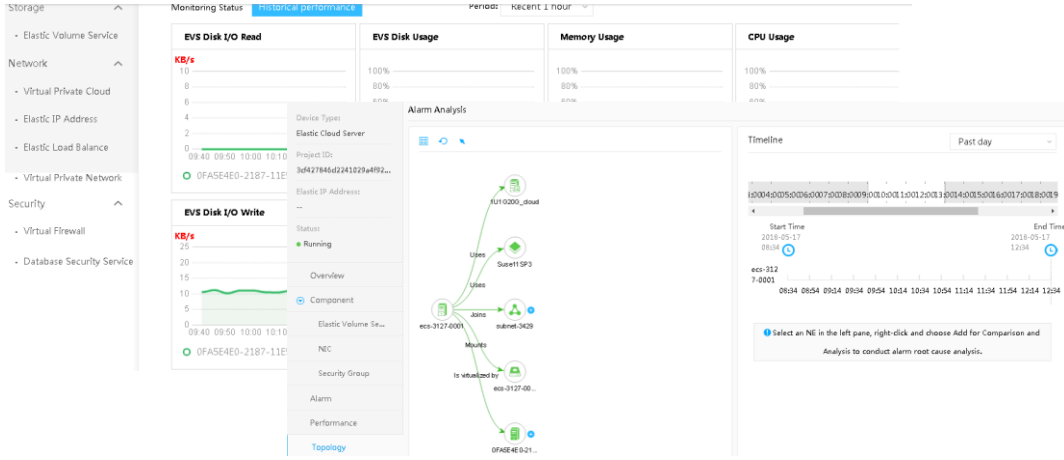
- Data assets are displayed from the perspective of applications based on a logical hierarchy of cluster > physical resource > component (service) > tenant. Users get a big-picture view of overall data consumption.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Cloud Resource Monitoring

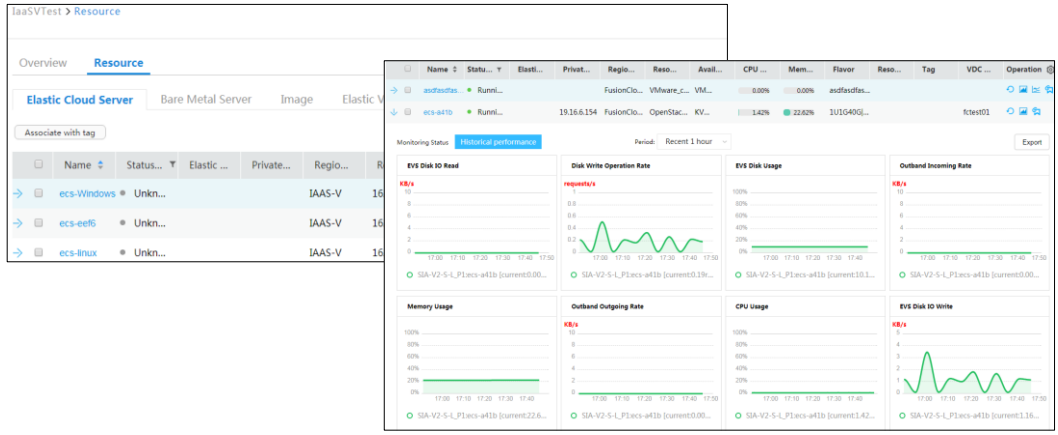
- Cloud Resource Monitoring provides comprehensive monitoring of cloud resource alarms, resources, topologies, and performance, helping users quickly demarcate and locate cloud resource assurance problems.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - VDC Monitoring

- VDC Monitoring provides comprehensive VDC-based analysis and evaluation to help users use resources more appropriately and improve resource usage.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Monitoring - Tenants' Big Data Application Monitoring

This function:

- Monitors resources from the application perspective.
- Continuously evaluates application resource usage from various aspects such as capacity and load.
- Globally controls the overall load of tenants' applications.
- Clearly displays the status of key applications.
- Provides all-round assurance for applications in terms of load, resource consumption, fault, and associated topology.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

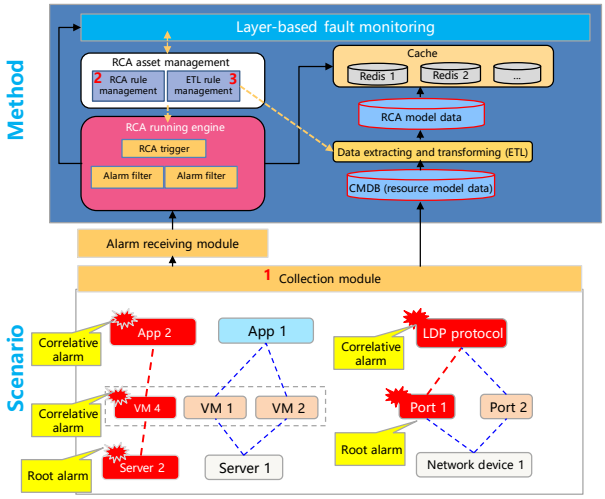
Intelligent O&M

1. Unified Monitoring
2. Intelligent Fault Locating
3. Visualized O&M

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent Fault Locating - Alarm Management

- Multiple methods are provided to compress alarms in different scenarios, making fault locating more accurate and O&M more efficient.



Application Scenarios and Benefits

- Devices are deployed from a vertical box-shaped to a cross-layer structure, including infrastructures, VMs, and virtual NEs. All layers are closely connected, and seamless integrated and maintained, which brings great challenges for fault demarcation and locating.
- Hierarchical decoupling creates risks, complicates O&M, and increases fault demarcation and locating complexity.

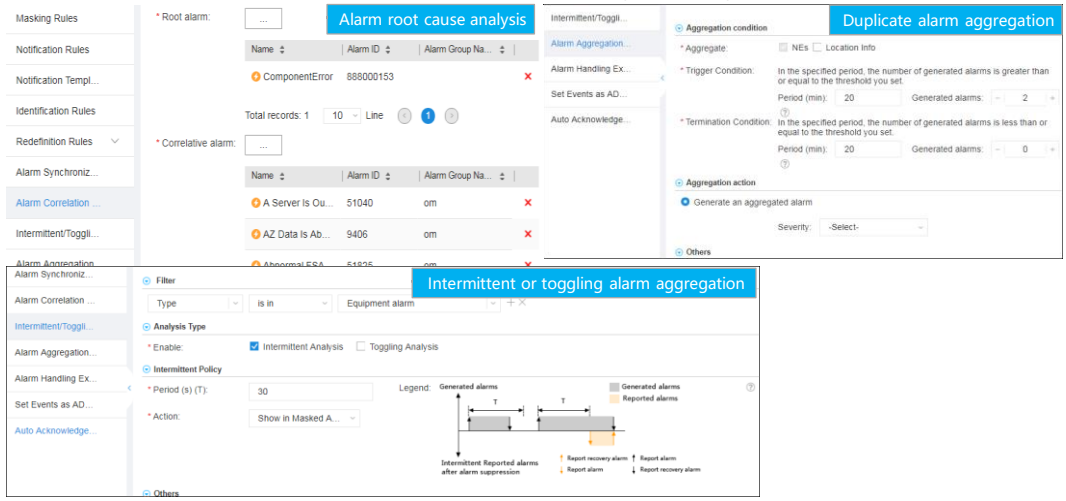
Key Technologies and Specifications

- Intermittent or toggling alarm: O&M personnel can discard intermittent or toggling alarms to improve O&M efficiency.
- Alarm aggregation: After an alarm aggregation rule is set, the system automatically aggregates the repeated alarms reported within the specified period into one alarm to improve O&M efficiency.
- Alarm correlation analysis: An alarm correlation rule identifies the root alarm and the correlative alarms. When monitoring or viewing alarms, you can set an alarm correlation rules to filter out correlative alarms and focus on root alarms only.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent Fault Locating - Alarm Management

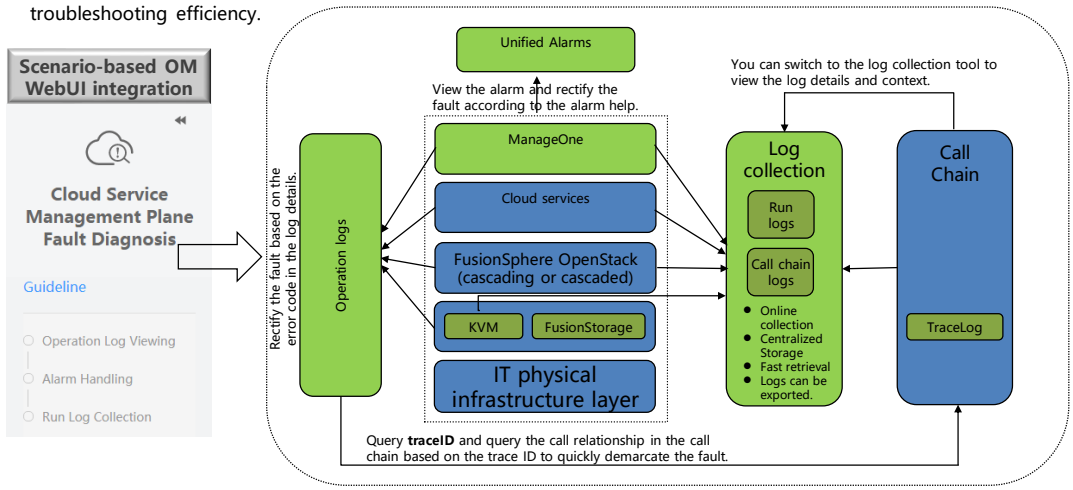
- Multiple methods are provided to compress alarms in different scenarios, making fault locating more accurate and O&M more efficient.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

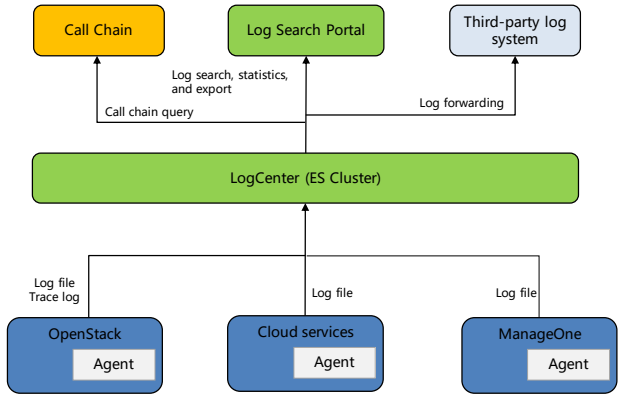
Intelligent Fault Locating - Unified Scenario-based Troubleshooting

- Integrated OM WebUI based on typical scenarios used for unified troubleshooting. A variety of O&M tools improve troubleshooting efficiency.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent Fault Locating - Log-assisted Fault Demarcation and Locating



Fault locating

- Run logs of different services on each node can be quasi-real-time collected and stored centrally.
- Logs in a specific fault scenario can be searched for by keyword.
- Logs can be exported.
- Logs can be forwarded to third-party log analysis systems and storage systems.

Auxiliary fault demarcation

- Call chain logs can be collected and log search is provided for call chains to facilitate call chain fault demarcation.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent O&M

1. Unified Monitoring
2. Intelligent Fault Locating
3. Visualized O&M

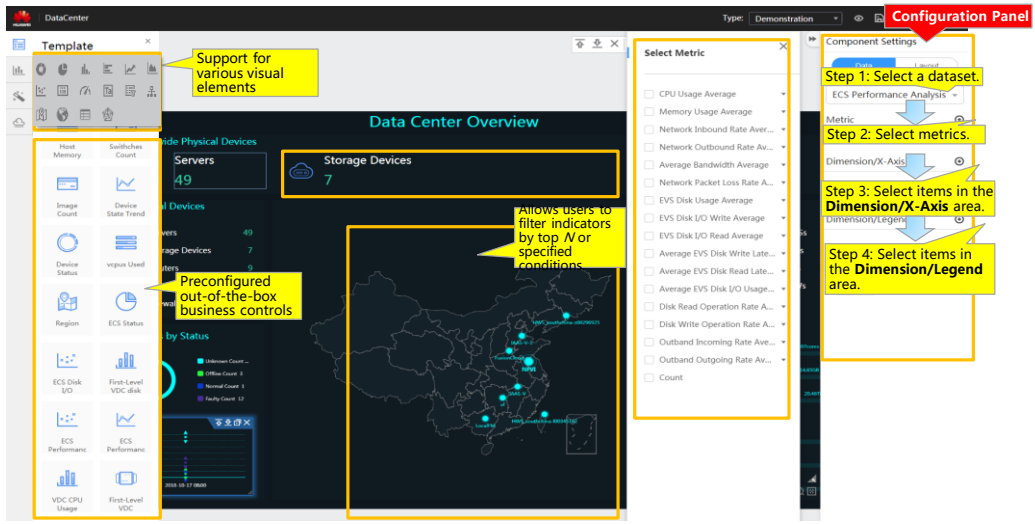
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Visualized O&M - Preconfigured Scenario-based Dashboards

- **Data Center Overview**: collects statistics on physical resources and resource provisioning.
- **Resource Pool Overview**: analyzes resource pool resource usage, capacity, and load.
- **VDC Resource Overview**: analyzes and collects statistics on VDC resource usage and capacity.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Visualized O&M - Flexible Customization of Dashboards



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. Trends and Challenges of Cloud Data Center Management
2. ManageOne Cloud Management Solution
3. **ManageOne Cloud Management Platform Features**
   - Fine-grained Operations
   - Intelligent O&M
   - Centralized Management

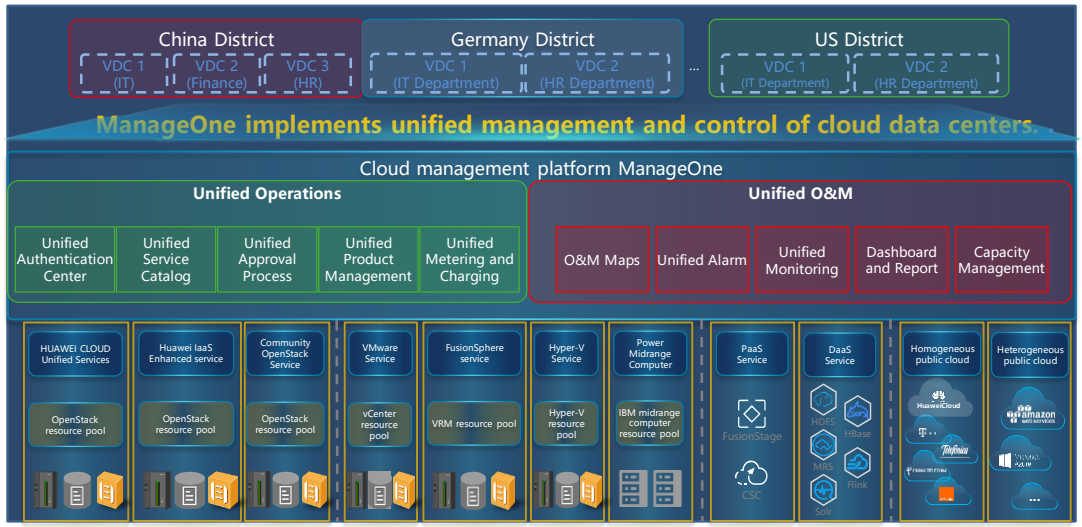
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent O&M

1. One Cloud Multi-Pool
2. Unified Configuration Center
3. Hybrid Cloud
4. Multi-Cloud Management

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

One Cloud Multi-Pool



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

One Cloud Multi-Pool

- If a customer already has a VMware resource pool or specifically requires Huawei build a VMware resource pool, VMware Service can be used to take over the VMware cloud services and rent them to tenants, including VMware ECS, VMware EVS, VMware IMS, and VMware snapshot. In this way, the customer can centrally manage their new and inventory VMware resources.
- If a customer already has an SCVMM (Hyper-V) resource pool or specifically requires Huawei to build an SCVMM (Hyper-V) resource pool, Hyper-V Service can be used to take over the SCVMM (Hyper-V) resources and rent services to tenants, including Hyper-V ECS, Hyper-V EVS, Hyper-V IMS, and Hyper-V snapshot. This way, the customer can centrally manage their new and inventory SCVMM (Hyper-V) resources.
- If a customer already has a Power resource pool or specifically requires Huawei to build a Power resource pool, Power Service can be used to take over the Power midrange servers, allowing tenants to apply for computing services of high-performance Power VMs, and provisions them to tenants. In this way, the customer can centrally manage their new and inventory Power resources.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Intelligent O&M

1 One Cloud Multi-Pool

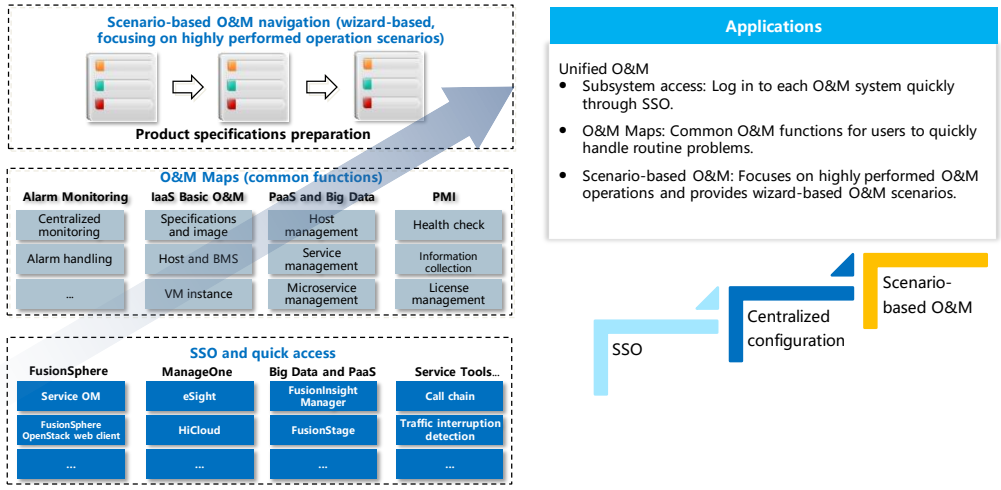
Unified Configuration Center 2

3 Hybrid Cloud

Multi-Cloud Management 4

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

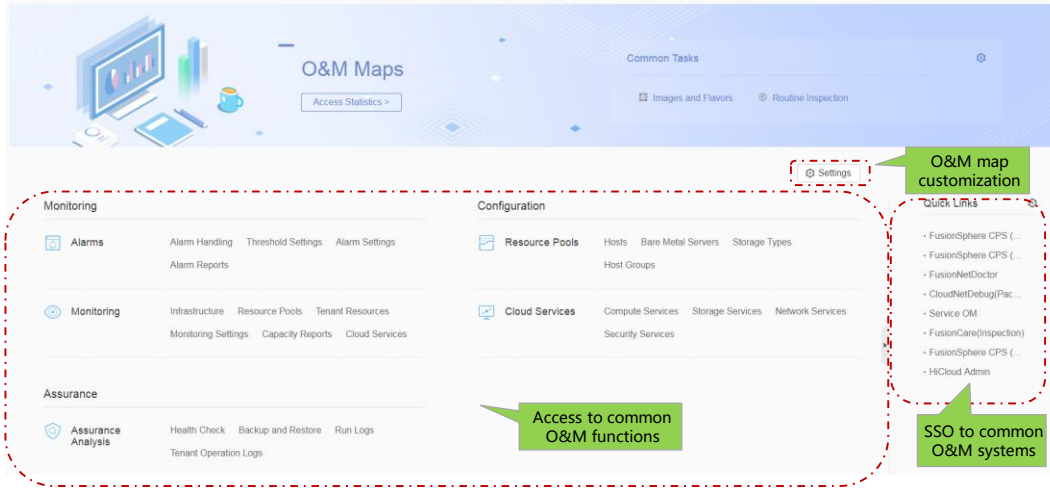
Unified Configuration Center - Unified O&M GUI



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

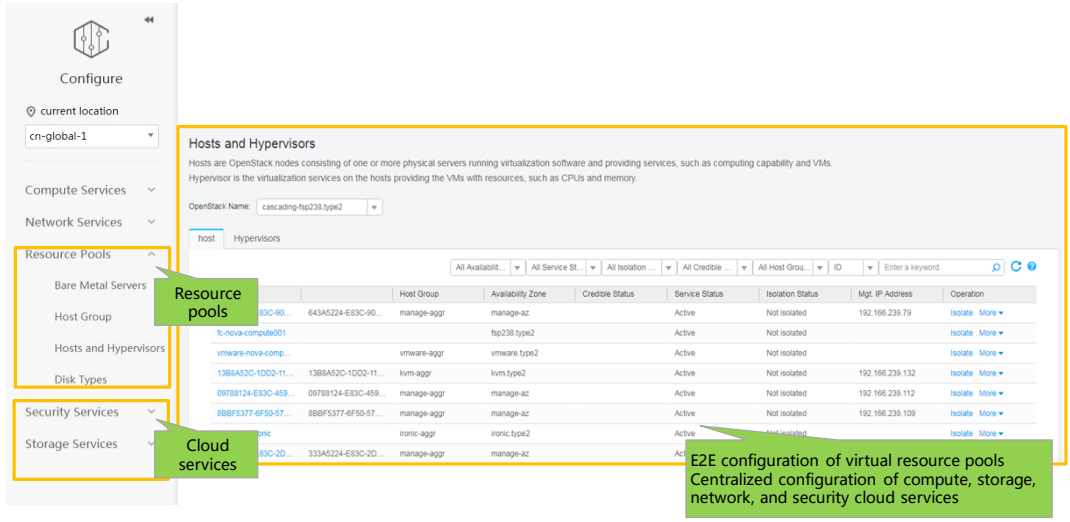
- ManageOne provides a unified O&M portal to resolve issues such as many O&M entries and no E2E scenario, helping users complete O&M based on wizards.
- ManageOne improves operation efficiency and user experience. Only one O&M portal is reserved for customers.

Unified Configuration Center - O&M Maps



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Unified Configuration Center - Scenario-based Configuration



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- ManageOne provides scenario-specific configuration capabilities for resource pools and cloud services in data center scenarios.

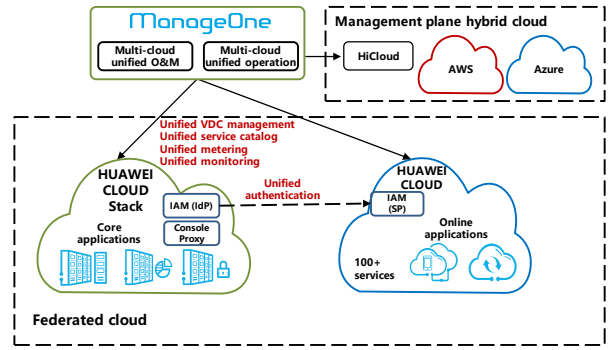
Intelligent O&M

1. One Cloud Multi-Pool
2. Unified Configuration Center
3. Hybrid Cloud
4. Multi-Cloud Management

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Hybrid Cloud - Overall Architecture

- Hybrid Cloud with HUAWEI CLOUD:
  - A federated cloud allows offline enterprise customers to use a broad set of service catalogs of HUAWEI CLOUD. Resources are available globally.
  - A federated cloud rapidly integrates with the public cloud service catalogs through federation authentication.
  - A federated cloud provides operation and O&M functions such as unified VDC management and unified monitoring.
- Hybrid Cloud with AWS and Azure:
  - Management plane hybrid cloud is implemented through APIs, and it provides services such as ECS, EVS, VPC, and EIP to meet the requirements of customers outside China for hybrid heterogeneous public cloud.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

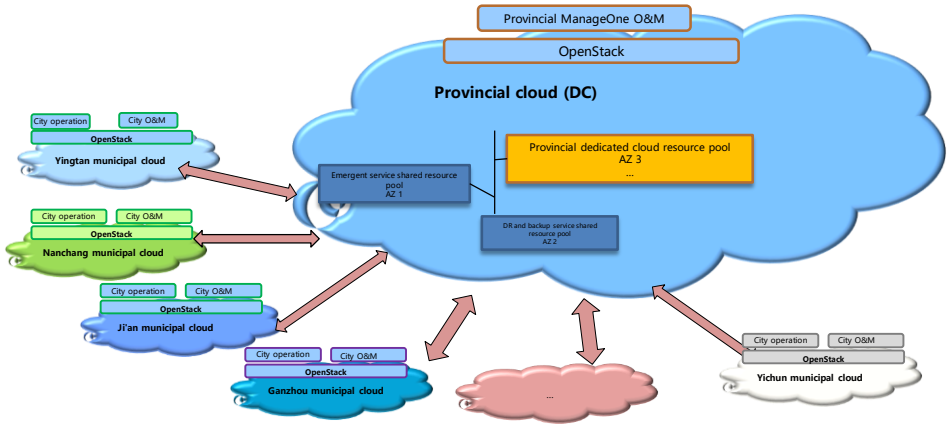
Intelligent O&M

1. One Cloud Multi-Pool
2. Unified Configuration Center
3. Hybrid Cloud
4. Multi-Cloud Management

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Cloud Management - Unified Multi-Cloud Monitoring

- Multi-Cloud Monitoring is a unified O&M monitoring function for provincial, municipal, and multi-cloud systems. It provides global cloud resource query and statistics.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Multi-Cloud Management - Unified Multi-Cloud Monitoring

- Have a good command of global resource usage and horizontal comparison and analysis of resource usage of multiple clouds.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Summary

- ManageOne Overview
- ManageOne Network Architecture
- ManageOne Feature Description

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. True or False

   Services created by the system administrator and organization administrator are authorized to the organization or specified VDCs in the organization. ( )

2. Multiple-Answer Question

   A VDC is the encapsulation and boundary definition of the virtual resources used by a department. It is a collection of virtual resources, including ( )

   A. Compute resources

   B. Storage resources

   C. Network resources

   D. Database resources

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Reference answer:
  - True or False: F
  - Multiple-Answer Question: ABC

More Information

- Huawei official websites
  - Enterprise service: http://e.huawei.com/en/
  - Technical support: http://support.huawei.com/enterprise/en/cloud-computing/manageone-sc-pid-21270651?category=product-documentation
  - Online learning: http://learning.huawei.com/en/
- Documentation tool
  - HedEx Lite

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei E-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support cases:
  - http://support.huawei.com/enterprise/servicecenter?lang=en

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank you
www.huawei.com



Introduction to eSight

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Introduction

- eSight is an integrated convergent O&M management solution oriented to enterprise data centers, campuses, branch networks, unified communications, videoconferencing, and video surveillance.
- eSight centrally manages servers, storage devices, virtualization, switches, routers, WLANs, firewalls, eLTE devices, eNodeBs, service engines, equipment room facilities, UC, telepresence, video surveillance, and application systems. eSight enables automatic deployment, visualized fault diagnosis, and intelligent capacity analysis for enterprise ICT devices. With these functions, eSight effectively helps enterprises improve O&M efficiency and resource usage at lower O&M costs, ensuring reliable operations of ICT systems.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Objective

- Upon completion of this course, you will understand:
  - eSight overview
  - eSight architecture
  - eSight functions
  - eSight deployment

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

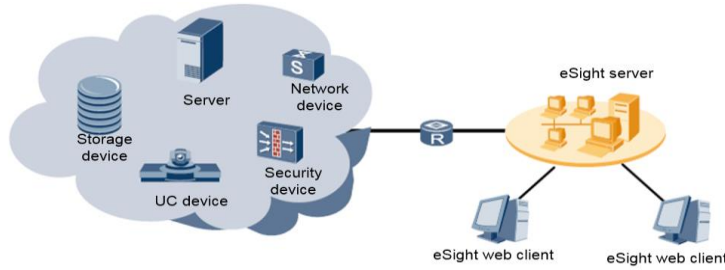
Contents

1. **eSight Overview**
2. eSight Architecture
3. eSight Functions
4. eSight Deployment

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Positioning

- eSight is a new-generation comprehensive operation and maintenance solution developed by Huawei for network infrastructure management, unified communications, telepresence conferencing, video surveillance, and data centers of enterprises. eSight supports unified monitoring and configuration management over devices of various types and from various vendors, monitors and analyzes network and service quality, and implements unified management and intelligent association for enterprise resources, services, and users.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Positioning

Traditional O&M		eSight
Independent device management with multiple maintenance systems used Associated troubleshooting failure caused by multi-person maintenance and management	Manpower	Unified management of devices by one maintenance system Reduced maintenance costs and associated troubleshooting
Time-consuming and costly site deployment Time-consuming and labor-intensive troubleshooting	Time	Automatic deployment with higher efficiency Visualized diagnosis with reduced downtime
Losses caused by device faults and service interruption Repeated capital investment caused by lack of planning basis	Capital	Full lifecycle management of installation, deployment, maintenance, optimization, and upgrade Intelligent capacity analysis with increased resource value

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Features

- Lightweight and web-based clients
  - eSight uses the B/S architecture and requires no other plug-ins.
  - With the distributed feature, eSight allows users to perform operations like querying and browsing anywhere anytime.
- Large-scale management capability
  - A maximum of 20,000 NEs can be managed.
  - A maximum of 100 clients can be online at the same time.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Features

- Support for multiple types of operating systems
  - Windows
  - SUSE Linux
- Support for multiple types of databases
  - Oracle
  - MySQL
  - SQL Server

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Product Features

- Capability of managing devices from multiple vendors
  - Huawei devices: switches, routers, UC devices, telepresence devices, video surveillance devices, servers, and storage devices
  - Non-Huawei devices: devices from H3C, Cisco, and ZTE, and IT devices from IBM, HP, and SUN
- Multiple service management components
  - eSight uses a component-based architecture and provides various components. Customers can select the required components based on the site requirements.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- eSight can manage mainstream devices from H3C, Cisco, and ZTE by default. For other non-Huawei devices, eSight allows users to customize the management method.
  - Users can use eSight to manage non-Huawei devices that support standard management information bases (MIBs) (including RFC1213-MIB, Entity-MIB, SNMPv2-MIB, and IF-MIB) through user-defined settings.
  - Users can use eSight to manage non-Huawei devices that do not support standard MIBs through NE adaptation packages.

Product Features

- Support for multiple types of southbound interfaces, including:
  - SNMP
  - Telnet/STelnet
  - FTP/SFTP/FTPS
  - TR069
  - Huawei Man-Machine Language (MML)
  - SMI-S
  - Modbus
  - HTTPS
- System reliability
- Architecture scalability
- Ability to be integrated

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

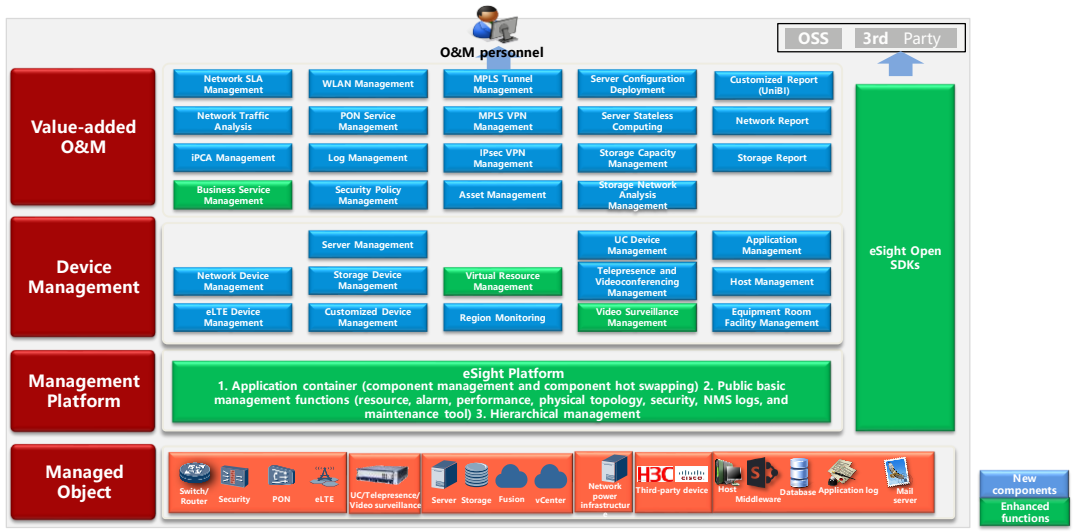
Contents

1. eSight Overview
2. **eSight Architecture**
3. eSight Functions
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

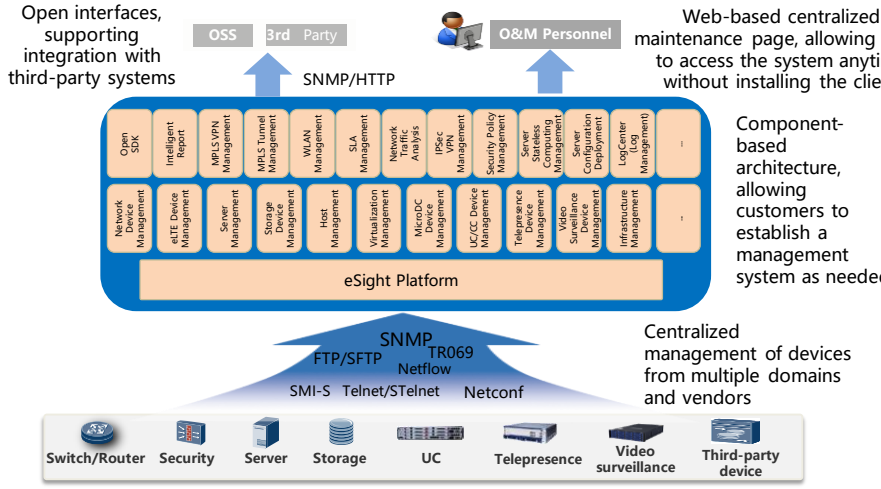
HUAWEI

Overall eSight Solution



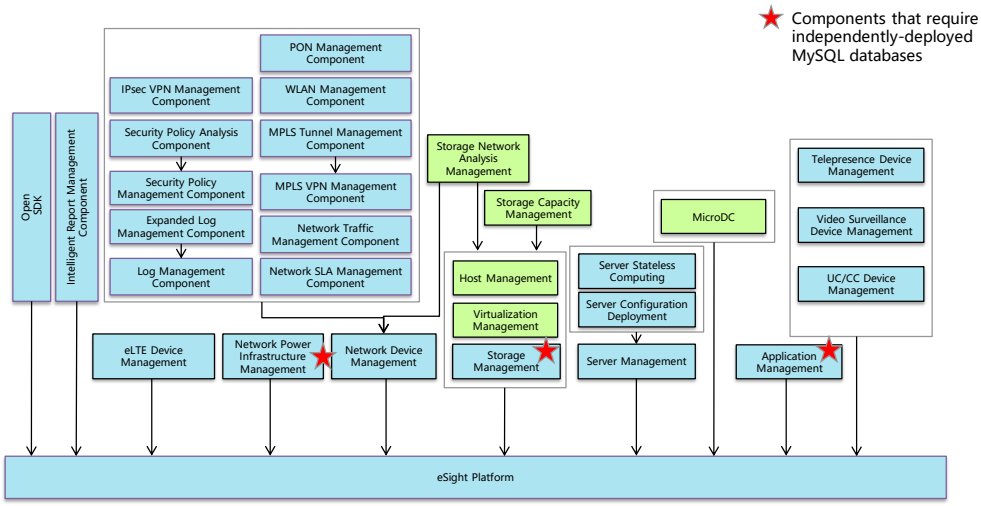
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

eSight Architecture



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Dependencies Between eSight Components



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.
HUAWEI

Contents

1. eSight Overview
2. eSight Architecture
3. **eSight Functions**
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

eSight Functions

- Basic management functions
  - Log management
  - Resource management
  - Alarm management
  - Performance management
  - License management

Server management

MicroDC management

Storage management

Equipment room facility management

Network device and service management

Host management

eLTE management

Unified communications and collaboration management

Computing virtualization management

Application management

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. eSight Overview
2. eSight Architecture
3. **eSight Functions**
   - Basic Management Functions
   - Server Management
   - Storage Management
   - Network and Security Management
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Basic Management Functions: Resource Management

- Resource management includes adding and managing devices and subnets.
- Devices can be added to eSight in any of the following ways: automatic discovery, manual creation, and batch import.
- Multiple protocols are supported, including SNMP, SNMP+Telnet/STelnet, HTTPS, IPMI, MML, REST, SMI-S, SOAP, SSH, TLV, TR069, and WMI.

The SNMPv1 and SNMPv2c protocols have security risks. SNMPv3 is recommended. The Telnet protocol has security risks. STelnet is recommended.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

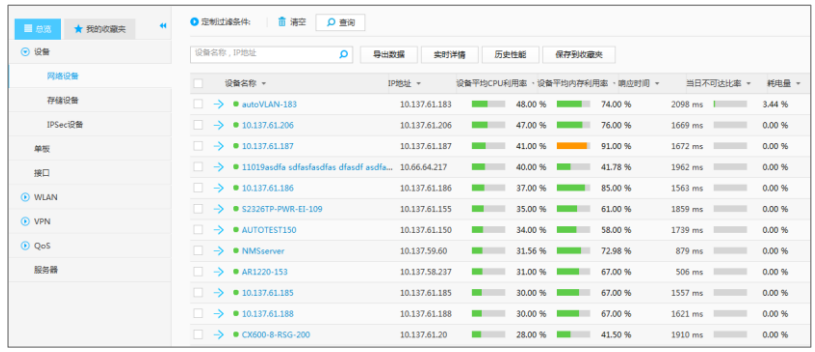
Basic Management Functions: Alarm Management

- eSight needs to promptly notify maintenance personnel of network exceptions so that the maintenance personnel can take proper measures to recover the network.
- Alarm management includes the following functions:
  - Monitors network-wide alarms and remotely sends alarm notifications to notify maintenance engineers in a timely manner, ensuring troubleshooting efficiency.
  - Blocks alarms and provides maintenance experience library to improve the efficiency and accuracy of alarm handling.
  - Synchronizes alarms to ensure alarm reliability.
  - Provides customized functions such as alarm filtering and alarm severity redefinition to meet requirements in various scenarios.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Performance Management

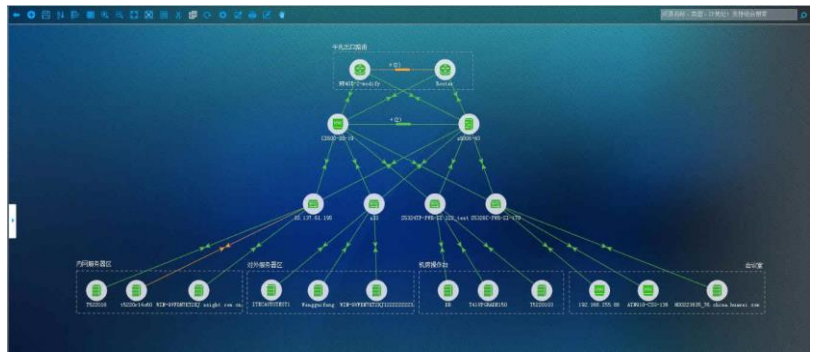
- The network performance may deteriorate due to internal or external factors, causing network faults. To ensure that the current network runs properly at a low cost and prepare for future network performance requirements, the network efficiency, such as the disconnection rate and usage, needs to be monitored. Performance management allows users to detect the deterioration trend in advance and eliminate these risks before faults occur.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Topology Management

- Topology management displays the managed NEs and their connections and status in the topology view. Users can check the topology view to better understand the hierarchy and running status of devices on the entire network.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Log Management

- Log management records important user operations and allows users to query the log list, view detailed log information, and export operation, system, and security logs to a CSV file.
- eSight logs mainly include security logs, system logs, operation logs, and background run logs.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Report Management

- Network report: eSight displays performance and alarm reports of network devices from multiple dimensions, helping users with analysis, optimization, and decision-making.
- Storage report: eSight displays performance and capacity analysis reports of storage devices, hosts, and virtualization servers from multiple dimensions, helping users analyze performance bottlenecks and plan capacity usage.
- Resource statistics report: Resource statistics reports collect statistics on the number of new resources managed by eSight and the total number of resources managed by eSight from dimensions such as time, region, category, type, and vendor. In addition, the report can display the trend and status of resource quantity.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Asset Management

- eSight manages networks, servers, and storage assets throughout their life cycles from asset stock-in to online running to returning.
- Bench sheet management
  - The bench sheet management function allows users to import, modify, and view assets such as devices and accessories, and provides asset modification functions such as deploying, removing, transferring, repairing, and returning assets.
- Returned assets
  - Users can view the returned device and accessory assets and export them to an Excel file.

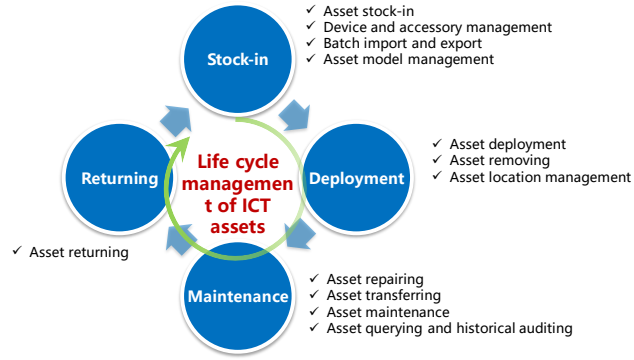
- Location and rack management
  - Before recording an asset to eSight, you need to define the location model, location instance, and cabinet information of the asset in the location and rack management component. In this way, the asset can be specified to a specific geographical location and cabinet when it is imported into the database.
- Model management
  - An asset management model is preconfigured in the asset management component. If the model cannot meet the management requirements, you can add asset attribute fields based on the built-in model to meet the asset management requirements.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Asset Management

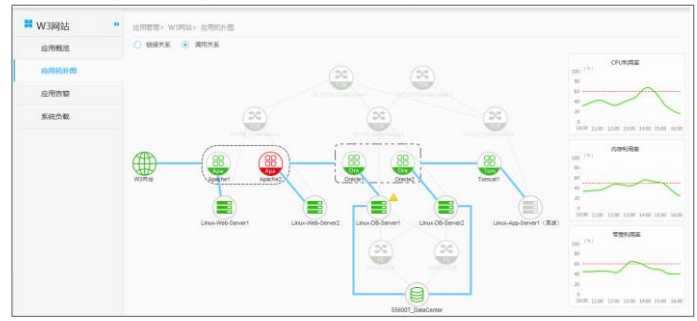
- By managing devices, accessories, locations, changes, and models, eSight can manage data center assets such as networks, servers, and storage devices throughout their life cycles from stock-in, deployment, maintenance, to returning. This helps enterprises comprehensively understand asset distribution and running status and make correct investment decisions to lower the cost and enhance O&M efficiency.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

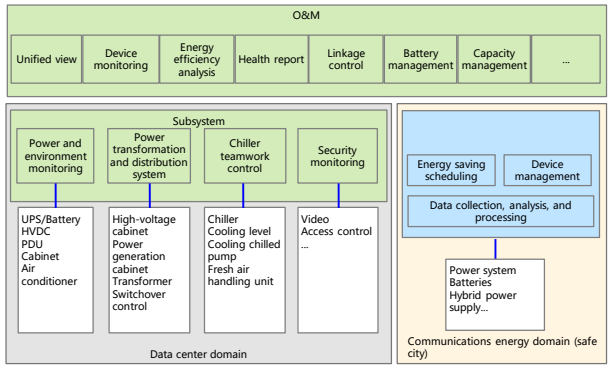
Basic Management Functions: Service Management

- The core value of service management lies in the ability to connect service processes to IT services and underlying applications and infrastructure components, providing insight into service quality from the service user perspective. The key to establishing service management is to connect service impacts with service models of IT assets and resources, including service definition, dependency discovery and mapping, service availability indicator definition for end users, and service model maintenance throughout the service life cycle.
- Based on the comprehensive infrastructure management capability, eSight can create a visualized model between services and IT infrastructure and application components (web service, application service, middleware, database, operating system, FC switch, storage device, switch, and router). In addition, eSight provides a 360-degree view for each service and displays service running status from various aspects including external service status, service application logical topology, application alarm, and system load.
- IT personnel only need to focus on specific IT infrastructure resources of the business service and solve related problems, greatly decreasing the time required for troubleshooting, quickly recovering services, and reducing the service interruption time.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Management Functions: Infrastructure Management



Simple O&M

- Unified management of the data center and site power supply
- Integration of devices and subsystems and joint positioning
- Comparison and analysis of the power supply component reliability to quickly identify outdated devices
- Visualization and WYSIWYG

Power supply reliability assurance

- Linkage control and fault isolation
- Device health check and report
- Dynamic configuration and evaluation
- Device reliability management (battery management, capacitor life evaluation, bus temperature monitoring, and fan blocking)

Maximizing resource utilization

- Recommendation of resource maximization policies and actions
- Power grid quality management: load matching
- Identification of cold and hot islands and cooling optimization
- Visualization of KPI (PUE and SPUE) power consumption
- Energy consumption statistics and analysis by domain
- Electricity cost analysis

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

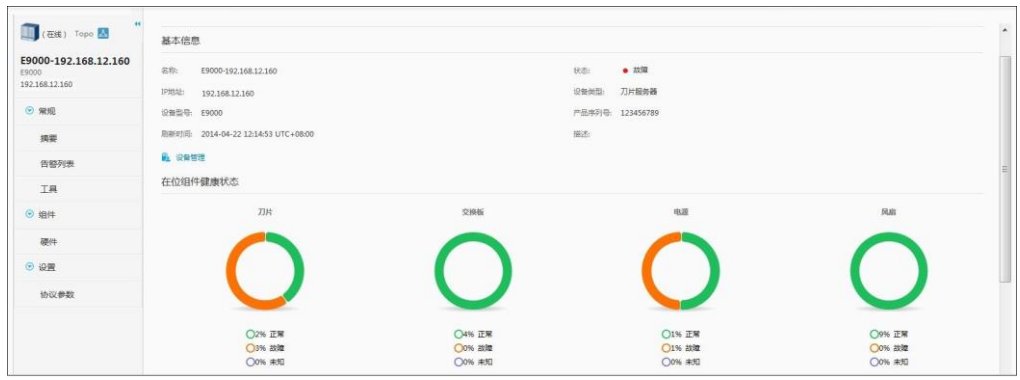
1. eSight Overview
2. eSight Architecture
3. **eSight Functions**
   - Basic Management Functions
   - Server Management
   - Storage Management
   - Network and Security Management
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Server Management: Device Management

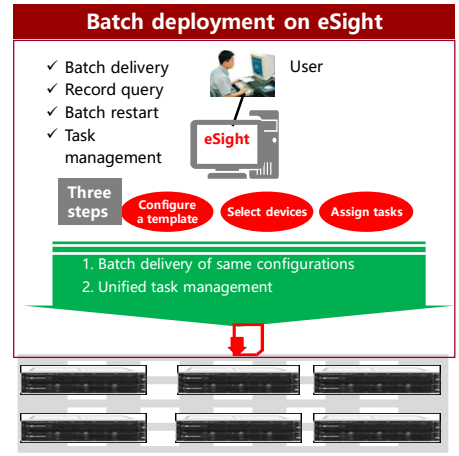
- Server management: eSight provides centralized server fault monitoring, performance analysis, and virtual media integration tool, which greatly improve O&M efficiency and reduce the O&M cost.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Server Management: Configuration Deployment

- Supports batch power-on, power-off, and restart of servers.
- Supports batch configuration of management network ports.
- Supports the function of configuring iBMC (iMana).
- Supports batch configuration of BIOSs.
- Supports batch configuration of RAID groups.
- Supports batch configuration of HBAs.
- Supports batch configuration of CNAs.
- Supports partition configuration when an operating system is installed.
- Supports batch operating system installation and deployment.
- Supports software distribution and automatic installation.
- Supports the function of configuring switch boards.



- Import and export of configuration templates, and template sharing among multiple NMSs
- Batch configuration and automatic deployment

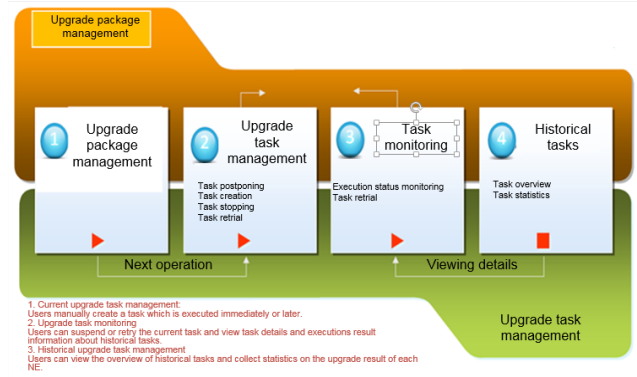
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Server Management: Firmware Upgrade

- Remote firmware upgrade greatly reduces manpower, travel expenses, and service costs, improves maintenance efficiency, and shortens the upgrade duration, quickly providing new service experience for customers and enhancing product competitiveness.

Prepare the upgrade package:

1. Manually download the firmware package from Huawei support website.
2. Firmware upgrade package management automatically downloads an upgrade package with a version later than that of the connected device from Huawei support website.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

1. eSight Overview
2. eSight Architecture
3. **eSight Functions**
   - Basic Management Functions
   - Server Management
   - Storage Management
   - Network and Security Management
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Storage Management: Storage Device Discovery and Batch Import

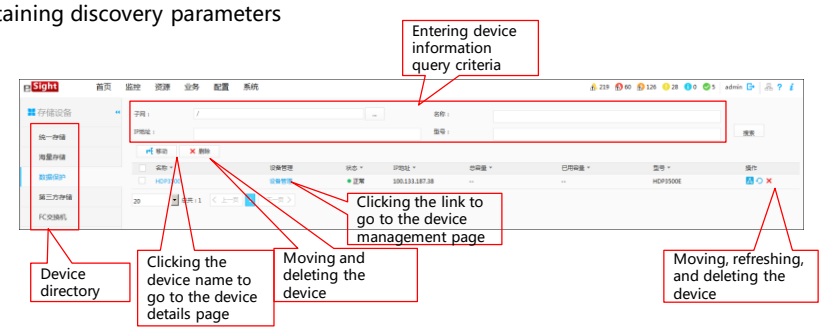
- Huawei-developed storage devices:
  - Unified storage devices
  - Massive storage devices
  - Data protection devices
- Third-party storage devices and FC switches
- The discovery modes are as follows:
  - Single addition
  - Automatic discovery
  - Batch import

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

1. Single addition indicates that one device is added to eSight at a time. The device can be a unified storage device, massive storage device, data protection device, third-party storage device, or FC switch.
2. Automatic discovery indicates discovery of devices through an IP address segment and related discovery parameters. This mode can easily add devices with the same user name and password to eSight.
3. Batch import means that users enter device IP addresses and discovery parameters in an Excel template and import the devices to eSight at a time. This mode can easily add devices whose IP addresses and discovery parameter information are known.

Storage Management: Storage Device Management

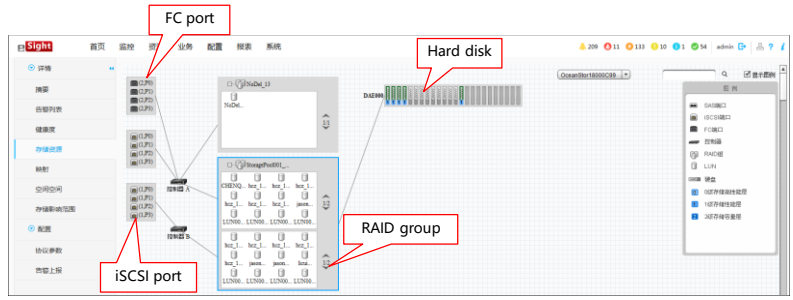
- Storage device management includes:
  - Displaying the status and attribute of devices and their components.
  - Managing device alarms.
  - Maintaining discovery parameters



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Management: Storage Visualization

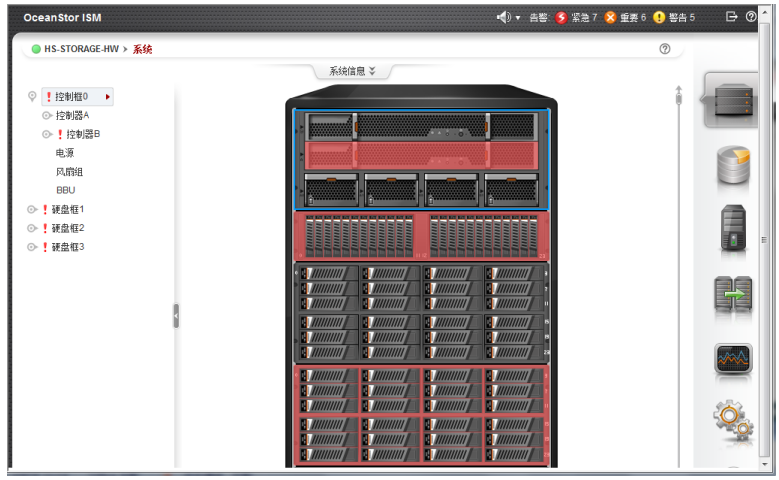
- Monitors the whole storage process from the front-end host port, controller, LUN, storage pool, to the hard disk.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Management: Storage Visualization

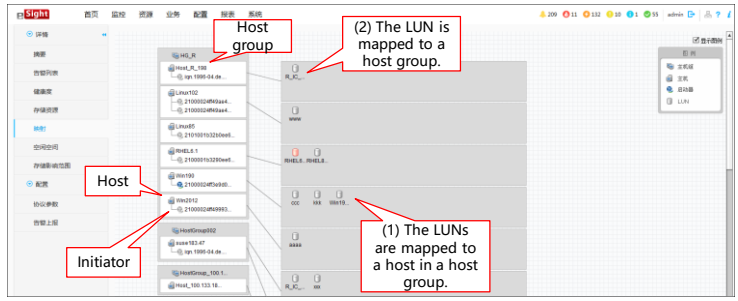
- The integrated Storage Device Manager displays the device rack diagram.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Management: Storage Device Mapping View

- Displays the mappings between arrays and unified storage devices.
  - LUNs can be mapped to host groups or hosts. One host group can contain multiple hosts, and one host can be configured with multiple initiators.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Management: Storage System Health Evaluation

- eSight can assess the overall health of a device from the workload, exception, and alarm statistics dimensions.
  - Health score
  - Health trend in the last 24 hours
  - Load and exceptions in the last hour
  - Workload of each component
  - Exception statistics of each component
  - Alarm statistics



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Storage Management: Cloud Service

- Cloud Service automatically sends device or eSight information to the O&M center through emails.
- The O&M center analyzes the running status of eSight and devices, detects faults in a timely manner, and reminds users of handling the faults.
- The backhaul information includes the health check report and event log package of the storage device, which can reflect the hardware and software configurations and fault status of the device.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Cloud Service refers to the built-in function of eSight Storage Manager. The function provides an independent Cloud Service software for remote notification of alarms reported by Huawei storage devices.

Contents

1. eSight Overview
2. eSight Architecture
3. **eSight Functions**
   - Basic Management Functions
   - Server Management
   - Storage Management
   - Network and Security Management
4. eSight Deployment

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Network Management

- eSight network management provides the following functions:
- IP topology management
- Link management
- Single-NE feature management
- Terminal resource
- VLAN management
- Smart configuration tool
- Configuration file management
- MIB management
- Device software management
- SLA management
- iPCA management
- QoS management
- Network traffic analysis
- IPsec VPN management
- Secure Center (security policy management)
- SVF management
- Zero touch provisioning (ZTP)
- eSight Mobile

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Security Management

Protection Mechanism	Mechanism Description	Security Policy
System security	System security ensures that the operating system, database, and middleware run properly.	Patch policy, hardening policy, password policy, authentication, data encryption, security log, minimum permission principle, and file permission management
Network security	Network security includes the normal running of network devices such as switches, routers, and firewalls.	• Network isolation: The LAN is isolated from the external network through routers to enhance data communication security. • Network firewalls are set up in the system to ensure system network security. • Permission on services that can be accessed externally is controlled and managed.
Data security	Data security includes the storage, transmission, and management security of user identity information, configuration information for normal system running, system run logs, and database data.	• Encryption policy: Sensitive data is encrypted for storage and transmission. • User management policy: minimum authorization • Backup/Restoration policy: Key data is periodically backed up. • Data storage security: The HA two-node cluster switchover mechanism is supported to restore the system in a timely manner.
Operation and maintenance security	eSight provides security mechanisms for users, applications, and auditing to ensure O&M security.	• Group- and permission-based access mechanism • Login access control policy: password policy, login lock and unlock, and authentication policy • Audit logs: security logs, operation logs, and system logs • Automatic client logout mechanism • Application software security mechanism: Provides password and identity authentication, and uses high-strength data encryption algorithms to encrypt sensitive user information for storage. The system allocates a password to each user and verifies the password when providing services for the user, securing user information.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

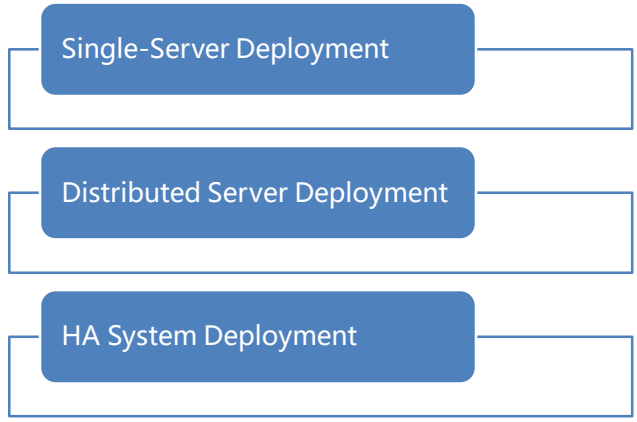
Contents

1. eSight Overview
2. eSight Architecture
3. eSight Functions
4. **eSight Deployment**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

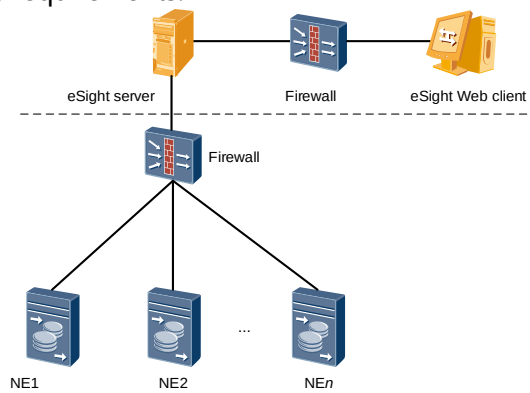
Deployment Modes



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Single-Server Deployment

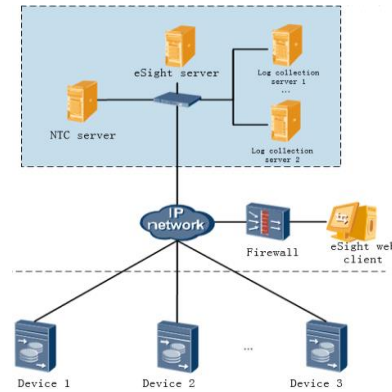
- eSight is deployed in a local single-node system. This mode applies to scenarios with low security requirements.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Distributed Server Deployment

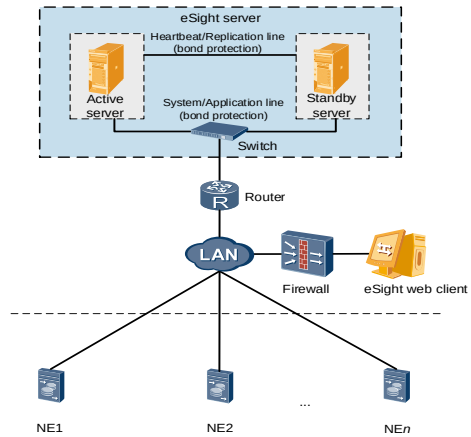
- One eSight server and one or more distributed collector servers are required.
- This mode applies to large-scale network management.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Local HA System Deployment

- eSight is deployed in a local two-node cluster and uses local disaster recovery (DR).



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

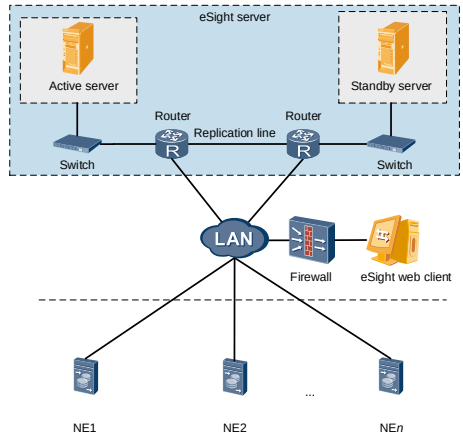
- The eSight local HA system consists of one active server and one standby server.

  The eSight software is installed on both the active and standby servers. Data on the active and standby servers are synchronized through a dedicated replication line. When the active server is faulty, services are automatically switched to the standby server to ensure that the eSight system runs properly.

- You can set a floating IP address between the active and standby servers. In this case, devices do not need to reconnect to eSight after the active-standby switchover.

Remote HA System Deployment

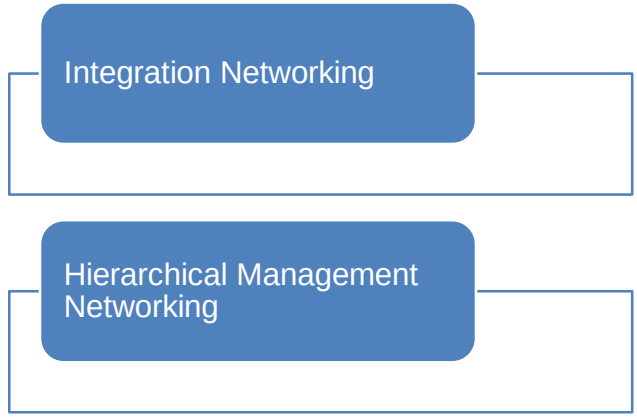
- eSight is deployed in a remote two-node cluster and uses remote DR.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The eSight remote HA system consists of one active server and one standby server. The eSight software is installed on both the active and standby servers. The two servers can be deployed in geographically-dispersed places. In case of a fault on the active server, services are automatically switched to the standby server. Data between the active and standby servers is synchronized through a dedicated replication line, which ensures normal running of the eSight system.
- The two eSight servers in the remote HA system use different IP addresses. After the remote HA system is deployed, the IP addresses of both the active and standby servers must be set on the managed devices. After the active-standby switchover, information such as the alarm information on the devices is automatically sent to the standby server to ensure normal device monitoring and management.

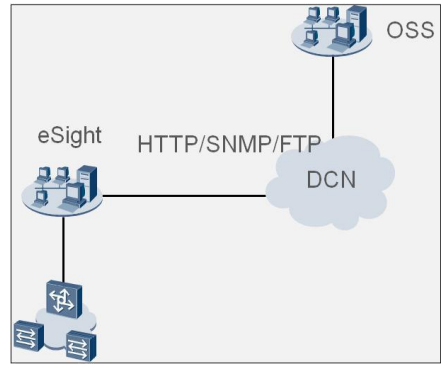
Networking Modes



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Networking Mode: Integration Networking

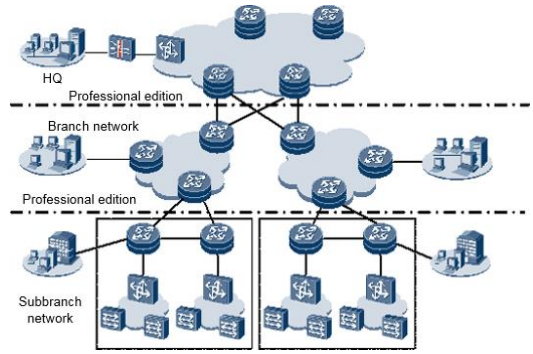
- eSight can be integrated into third-party systems such as upper-layer OSSs. Third-party systems can obtain network resources and alarms managed by eSight through SNMP or HTTP interfaces.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Networking Mode: Hierarchical Management Networking

- eSight supports hierarchical management, allowing enterprise HQs to monitor and manage networks in different regions.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- In the hierarchical management networking mode, upper-layer NMSs can add lower-layer NMSs to the system and provide links for accessing the lower-layer NMSs. When a user clicks a link for accessing a lower-layer NMS, a new browser window is displayed and the login page of the lower-layer NMS is displayed in the new browser window.

Standards and Protocols

- SNMP and MIB-II standards for interfaces between eSight and devices
  - RFC1155: structure and identification of management information for TCP/IP-based Internet
  - RFC1157: simple network management protocol
  - RFC1213: management information base for network management of TCP/IP-based internet (MIB-II) •XML 1.0
- ITU-T X.733: fault management specification
- JSR-286 Portlets specifications: Java Portlet specification v2.0
- HTTP/1.0|HTTP/1.1: Hypertext Transfer Protocol
- HTTPS: Hypertext Transfer Protocol Secure
- Session Initiation Protocol (SIP) RFC3261
- Transfer Control Protocol (TCP) RFC0872
- TCP and User Datagram Protocol (UDP) RFC1356
- Storage Management Initiative – Specification (SMI-S)
- Modbus Protocol

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Summary

- eSight Overview
- eSight Architecture
- eSight Functions
- eSight Deployment

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- eSight product:
  - http://e.huawei.com/cn/products/software/mgmt-sys/esight
- eSight Demo:
  - https://122.112.233.209:31943/

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Learning Materials

- Huawei e-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank you
www.huawei.com



Introduction to the Agile Controller-DCN

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Foreword

- This course describes the positioning, functions, application scenarios, standards, and protocols of the Agile Controller-DCN.

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Traditional Network:
  - CLI configuration
  - Scattered platform
  - Device management one by one
  - Independent scheduling of physical and virtual networks
  - Global invisibility
- SDN network
  - Automatic management:
  - GUI-based configuration
  - Open architecture
  - Application-based management and control
- Global visibility

Objectives

- Upon completion of this course, you will be able to:
  - Understand basic information about the Agile Controller-DCN.
  - Understand functions of the Agile Controller-DCN.
  - Be familiar with application scenarios of the Agile Controller-DCN.
  - Be familiar with standards and protocols with which the Agile Controller-DCN complies.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

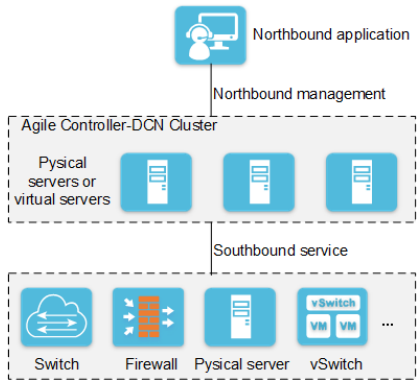
Contents

1. **Introduction to the Agile Controller-DCN**
2. Functions of the Agile Controller-DCN
3. Application Scenarios of the Agile Controller-DCN
4. Standards and Protocols of the Agile Controller-DCN

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Physical Deployment Architecture



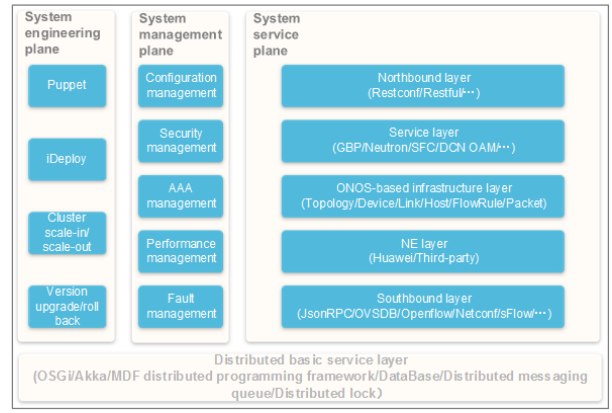
Module Type	Function Description	Deployment
Northbound proxy	Provides unified northbound access through a floating IP address, and forwards northbound requests to different service cluster nodes for load balancing.	In a cluster, only two nodes provide the northbound proxy function. To improve the system reliability, Northbound proxy is deployed on two cluster nodes in active/standby mode.
Cluster management	Uses the floating IP address for management of all Agile Controller-DCN cluster nodes, such as the internal cluster configuration and maintenance, cluster configuration, startup, stop, and process protection.	In a cluster, only two nodes provide the cluster management function. Cluster management is deployed in active/standby mode on the cluster nodes where northbound proxy is deployed.
Service processing	Indicates service management nodes that are used to process the Agile Controller-DCN services, such as NE management, topology management, and service provisioning, and send processing results to southbound forwarding devices.	The cluster deployment mode is used and at least 3 service processing nodes are required. If the cluster performance reaches the bottleneck, you can add nodes to improve the cluster performance.
Distributed lock	Provides capabilities of locking global resources for services.	The components use the cluster deployment mode and can be deployed independently or on the cluster nodes where service processing is deployed.
Distributed queue	A message sender sends the message to a message queue. Multiple receivers obtain the data and process their services in parallel, accelerating service processing.	
Database	Saves all service settings for real-time service query and configuration restoration after restart.	The component uses the cluster deployment mode and can be deployed independently or on the cluster nodes where service processing is deployed.

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- To ensure system performance and reliability, the Agile Controller-DCN must be deployed in cluster mode. Cluster technology has the following advantages:
  - Load balances service processing across multiple servers to ensure performance.
  - Ensures normal running of the entire cluster even if a single node fails, improving reliability.
  - Supports flexible expansion to enhance the performance of the entire cluster and features good scalability.

Logical Architecture

- Basic service layer of the distributed system: This plane provides the basic middleware service and Model Driven Framework (MDF) programming framework for SDN distributed programming.
- System engineering plane: This plane provides functions such as the Agile Controller-DCN cluster installation, deployment, scale-in, scale-out, and upgrade.
- System management plane: This plane provides system management capabilities for SDN services, including configuration management, security management, Authentication, Authorization, and Accounting (AAA) management, service performance monitoring, and fault management.
- System service plane: This plane is the key for Agile Controller-DCN service implementation. It collects network resources in the southbound and abstracts them for unified display and provides open northbound interfaces to provision SDN network services.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Basic service layer of the distributed system:
  - It provides basic middleware services for distributed SDN programming, including the Open Services Gateway Initiative (OSGi) container, Akka cluster management, distributed caching, distributed database storage, and distributed locking services. The OSGi container is provided by the ONOS platform, and the Akka cluster management service is provided by the ODL platform. Commercial functions of other distributed basic services are enhanced based on mainstream open source components in the industry, and fully meet the reliability, performance, and security requirements.
  - The distributed model driven framework (MDF) provides a modular service architecture based on ODL MD-SAL to ensure separated running and scheduling of processes and threads of various service protocols. This framework is compatible with MD-SAL interfaces to support enhanced functions, such as synchronous/asynchronous RPC encapsulation, routed RPC performance optimization, and high-performance DOM storage. The MDF framework integrates Kafka-based distributed messaging service bus and distributed event management capability, providing the reliability and performance.

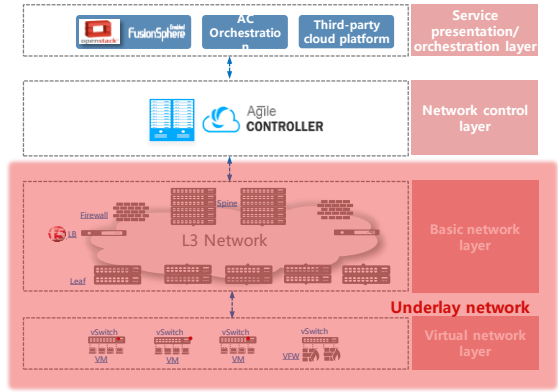
Contents

1. Introduction to Agile Controller-DCN
2. **Functions of Agile Controller-DCN**
   - Basic Concepts of Underlay
   - Basic Concepts of Overlay
   - Other Functions and Features
3. Application Scenarios of Agile Controller-DCN
4. Standards and Protocols of Agile Controller-DCN

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Concepts of Underlay

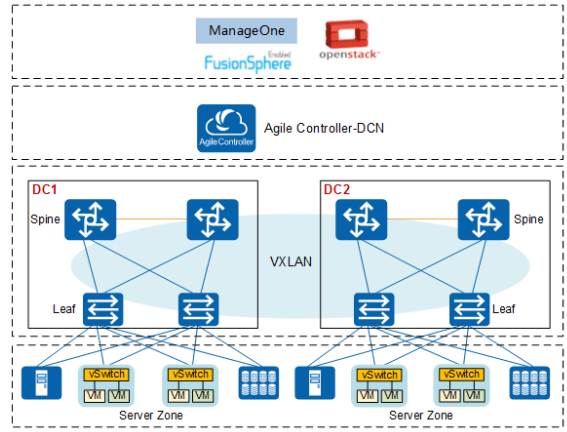
- An underlay network is a bearer network consisting of physical devices, such as TOR switches, aggregation switches, core switches, LBs, and firewalls. In Huawei CloudFabric Solution, the underlay network mainly refers to the basic network layer.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Underlay Network Features of Huawei CloudFabric Solution

- **High bandwidth:**
  - High-density 10GE
  - Evolution from 10GE to 25GE/40GE/100GE
- **Large Layer 2 network:**
  - Spine-leaf high scalability architecture
  - VXLAN
- **High reliability:**
  - CSS/iStack
  - M-LAG



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

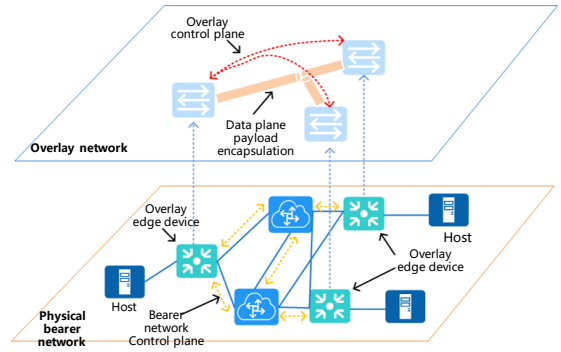
1. Introduction to Agile Controller-DCN
2. **Functions of Agile Controller-DCN**
   - Basic Concepts of Underlay
   - Basic Concepts of Overlay
   - Other Functions and Features
3. Application Scenarios of Agile Controller-DCN
4. Standards and Protocols of Agile Controller-DCN

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Basic Concepts of Overlay

- An overlay network (large Layer 2 network) is constructed at Layer 3 or Layer 4 on a traditional network and encapsulates data into Layer 3 or Layer 4 packets for transmission.
- Overlay technology is a tunnel encapsulation technology, including VXLAN and Network Virtualization Using Generic Routing Encapsulation (NVGRE). It encapsulates Layer 2 packets over tunnels, transparently transmits the encapsulated packets, and finally decapsulates the packets to obtain the raw packets after the packets arrive at the destination. That is, a large Layer 2 network is built on the existing network.



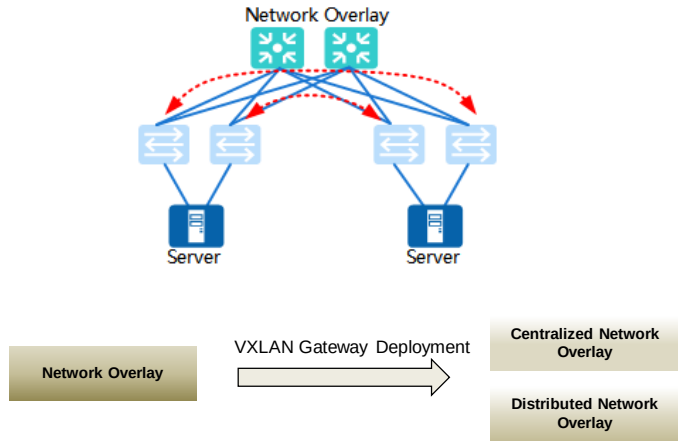
Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Basic Concepts of Overlay

- In Huawei CloudFabric Solution, an overlay is built using VXLAN technology. Service packets are transmitted on the VXLAN overlay that is decoupled from the physical bearer network.
- Based on the VXLAN NVEs, VXLAN-based overlays are classified into:
  - Network Overlay
  - Host Overlay
  - Hybrid Overlay

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Overlay Networking - Network Overlay



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

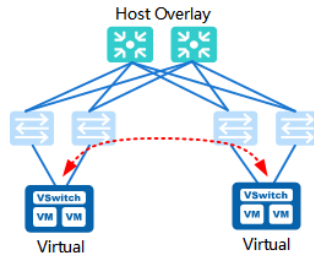
- KVM can convert a standard Linux kernel into a VMM. The Linux standard kernel embedded with the KVM module can load guest OSs through KVM tools. Therefore, in such an OS platform, the VMM virtualization layer resides directly on the physical hardware layer of the host, and no independent host OS layer is provided. In this case, the VMM functions as the host OS. CPU instructions of the guest OS are executed directly, rather than through QEMU. This greatly improves the speed. KVM exposes the necessary APIs through /dev/kvm. User-mode programs can access these APIs by calling the ioctl function.
- The KVM kernel module provides only CPU and memory virtualization. Therefore, it must be combined with QEMU to form a complete virtualization solution, that is, qemu-kvm.
- As a hypervisor, KVM focuses on VM scheduling and memory management. I/O peripheral tasks are handled by the Linux kernel and QEMU. With I/O virtualization, storage and network resources are virtualized by the Linux kernel and QEMU.
- By integrating KVM, QEMU uses its kernel module to process CPU instructions by invoking the /dev/kvm interface through ioctl. KVM is responsible only for CPU and memory virtualization. QEMU emulates I/O devices (such as NICs and disks). Server virtualization is jointly implemented with KVM and QEMU, so QEMU is called qemu-kvm. QEMU is an emulator that emulates the CPU and other hardware needed by the guest OS. The guest OS believes that it communicates with the hardware directly. In fact, it interacts with the hardware simulated by QEMU. QEMU translates and sends these instructions to the real hardware. The

performance is compromised because all the instructions need to be translated by QEMU.

- QEMU also emulates other hardware, such as NICs and disks, which also affects the performance of these devices. To address this, pass-through and paravirtualization techniques, such as virtio_blk and virtio_net, are used to improve device performance.

Overlay Networking - Host Overlay

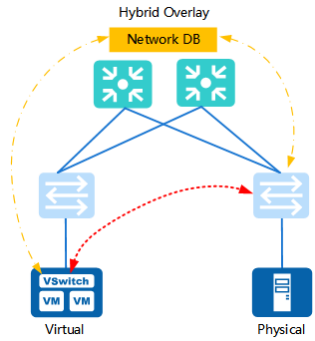
- On a host overlay network, all VXLAN overlay tunnel end points are deployed on software switches (installed on servers). That is, both the ingress and egress of a VXLAN tunnel are software switches.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Host Overlay is based on CE1800V

Overlay Networking - Hybrid Overlay



Differences	**In a network overlay scenario, all overlay devices are physical devices and a VXLAN tunnel is encapsulated on a physical switch.**
	In a hybrid overlay scenario, overlay devices include physical network devices and virtual network devices. A VXLAN tunnel is encapsulated on a physical switch or a virtual switch where the host server is located.
Advantages	**Uses the high performance of physical network devices for forwarding. Uses the existing physical network devices and overlay of physical servers to improve performance. Provides more flexible networking.**

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Contents

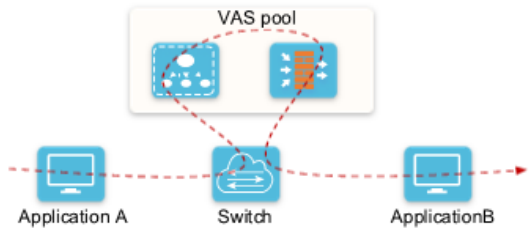
1. Introduction to Agile Controller-DCN
2. **Functions of Agile Controller-DCN**
   - Basic Concepts of Underlay
   - Basic Concepts of Overlay
   - Other Functions and Features
3. Application Scenarios of Agile Controller-DCN
4. Standards and Protocols of Agile Controller-DCN

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Flexible Service Orchestration

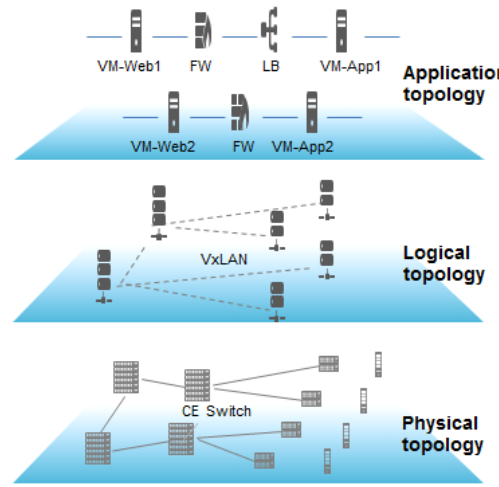
- SFC makes service orchestration more flexible. It has the following features:
  - Supports decoupling of service functions from fabric network devices, implementing flexible deployment and on-demand expansion and breaking the restrictions of the physical topology.
  - Provides GUIs to simplify operations and improve the orchestration efficiency, and provides personalized services for each tenant on demand without affecting other tenants.
  - Supports VAS pooling, flexible scalability, and on-demand allocation, realizing sharing and expansion of VAS capabilities in multiple services.



Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Refined O&M

- Network resource visualization
- Visibility of the physical network, logical network, and application network
- Network path detection
  Single-path detection
  Multi-path detection
- Connectivity detection
- Loop detection



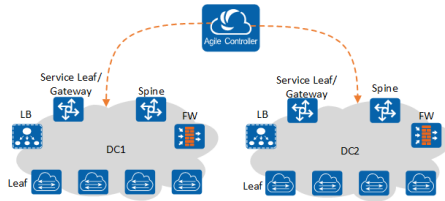
 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

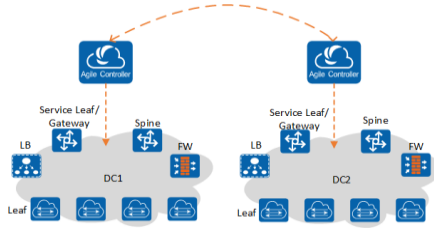
- Entire network resource visibility
  - The Agile Controller-DCN obtains network information from the dimension of physical or virtual device resources and tenants, and supports resource visibility, solving the problem of mixed devices and poor resource management.
  - The Agile Controller-DCN uniformly manages physical and virtual resources, and monitors the resource status of physical and virtual network devices on the entire network as well as the NE running status.
  - The Agile Controller-DCN monitors the network running status from the tenant dimension and displays the tenants, tenant quotas, and tenant traffic.
- Visibility of the physical network, logical network, and application network
  - The Agile Controller-DCN supports the visibility of the physical network, logical network, and application network.
  - Supports mapping of the logical network topology and the physical network topology.
  - Displays logical network resources used by application networks and physical network resources used by logical networks (that is, mapping from the top to the bottom).

Multi-DC Pooling

- Remote management of the Agile Controller-DCN: An Agile Controller-DCN cluster manages multiple DCs and centrally delivers the DC configuration on the overlay network.



- Cluster federation: The Agile Controller-DCN clusters are deployed independently for different DCs. The DCs exchange service routing information on the overlay network through BGP-EVPN to implement service communication at Layer 3.

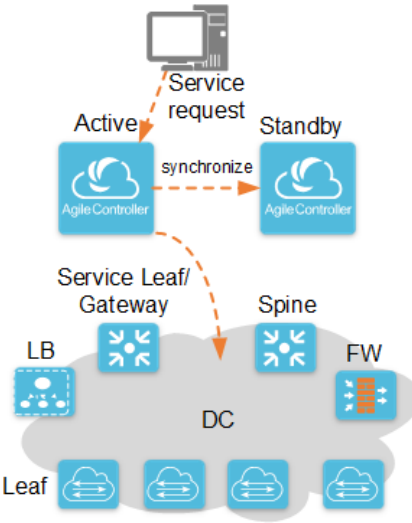


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The Agile Controller-DCN can manage both a single data center (DC) and multiple DCs in different regions, expanding the size and scope of DC services and breaks the physical distance limitations of traditional DCs. In this case, customers can share DC network resources in different regions, implementing flexible resource scheduling and improving resource utilization. The Agile Controller-DCN also supports geographic redundancy deployment of active and standby clusters. When the active cluster fails, a geographic redundancy switchover is triggered automatically or manually, ensuring smooth running of services and improving the DC reliability.
- The Agile Controller-DCN manages DCs in the following modes: remote management of multiple DCs using a single Agile Controller-DCN cluster, active/standby Agile Controller-DCN management, independent deployment of multiple sets of the Agile Controller-DCN.
    - Remote management of multiple DCs using a single Agile Controller-DCN cluster: A set of the Agile Controller-DCN manages multiple DCs and uniformly delivers the DC configuration on the overlay. This scenario supports inter-DC cluster and elastic resource scaling.
- Independent deployment of multiple sets of the Agile Controller-DCN: An independent Agile Controller-DCN cluster is deployed for each DC. DCs exchange the service routing on the overlay through BGP-EVPN and streamline services through upper-layer applications, implementing service communication at Layer 3 and elastic scalability.

Geographic Redundancy Deployment

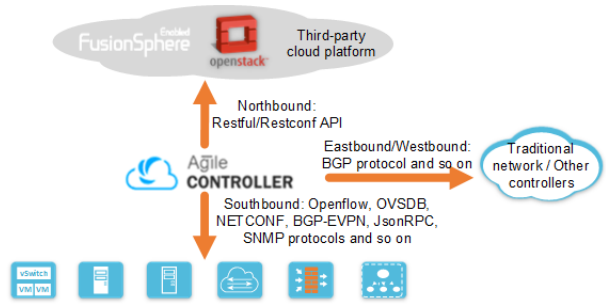


 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- The Agile Controller-DCN can manage both a single data center (DC) and multiple DCs in different regions, expanding the size and scope of DC services and breaks the physical distance limitations of traditional DCs. In this case, customers can share DC network resources in different regions, implementing flexible resource scheduling and improving resource utilization. The Agile Controller-DCN also supports geographic redundancy deployment of active and standby clusters. When the active cluster fails, a geographic redundancy switchover is triggered automatically or manually, ensuring smooth running of services and improving the DC reliability.
- The Agile Controller-DCN manages DCs in the following modes: remote management of multiple DCs using a single Agile Controller-DCN cluster, active/standby Agile Controller-DCN management, independent deployment of multiple sets of the Agile Controller-DCN.
    - Remote management of multiple DCs using a single Agile Controller-DCN cluster: A set of the Agile Controller-DCN manages multiple DCs and uniformly delivers the DC configuration on the overlay. This scenario supports inter-DC cluster and elastic resource scaling.
- Independent deployment of multiple sets of the Agile Controller-DCN: An independent Agile Controller-DCN cluster is deployed for each DC. DCs exchange the service routing on the overlay through BGP-EVPN and streamline services through upper-layer applications, implementing service communication at Layer 3 and elastic scalability.

Southbound and Northbound Openness

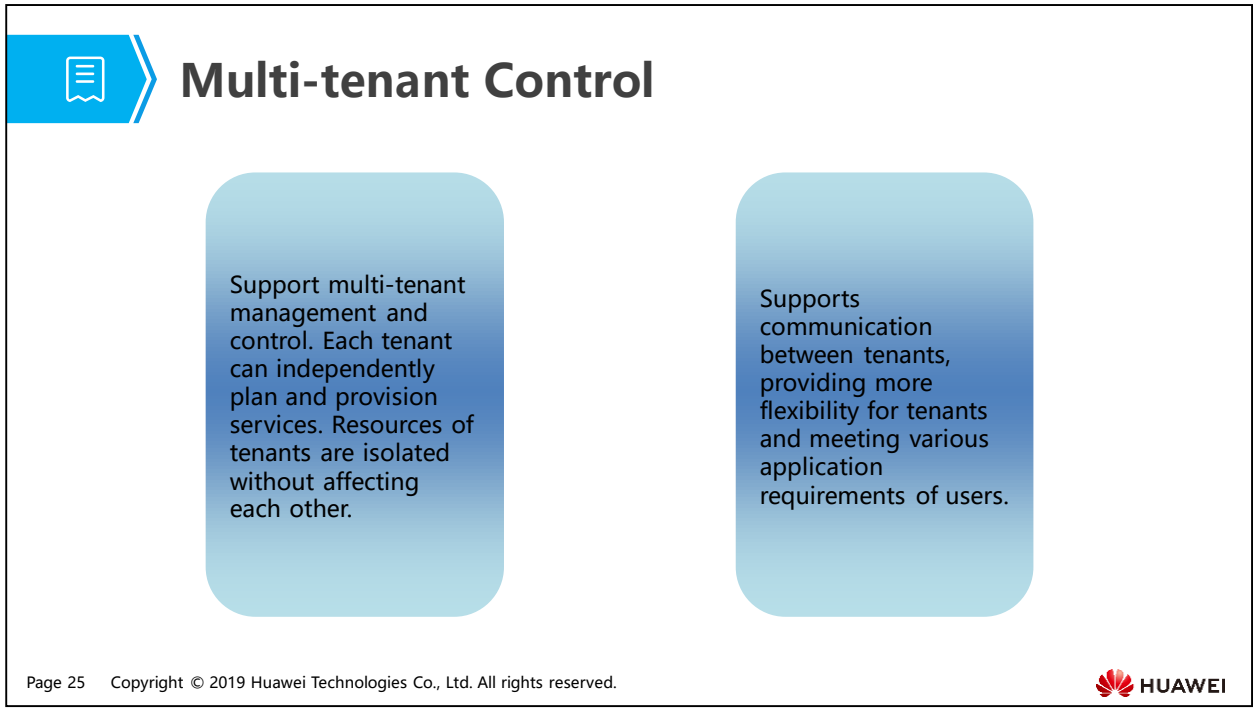
- The Agile Controller-DCN is based on an open software platform and has an architecture with loosely coupled components. The Agile Controller-DCN can provide extensive northbound API capabilities and southbound interface capabilities to control network devices and computing resources.



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The Agile Controller-DCN interconnects with the open-source OpenStack cloud platform, Huawei FusionSphere cloud platform, and applications seamlessly using the northbound standard RESTful interface.
- The Agile Controller-DCN manages the physical and virtual devices using the southbound standard OpenFlow, OVSDB, NETCONF, BGP-EVPN, JsonRPC, and SNMP protocols.
- The Agile Controller-DCN communicates with traditional networks or other controllers through the eastbound and westbound protocols such as BGP.



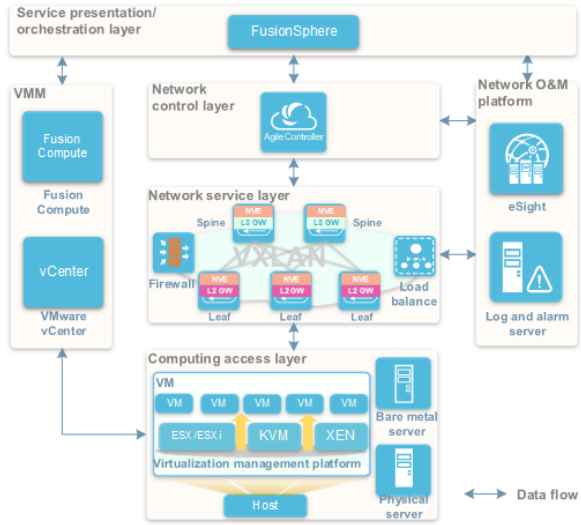
- The Agile Controller-DCN supports multi-tenant management and control during service provisioning or collaboration with other cloud platforms to provision services. Each tenant can independently plan and provision services. Resources of tenants are isolated without affecting each other.
- Supports communication between tenants under certain control, providing more flexibility for tenants and meeting various application requirements of users.

Contents

1. Introduction to the Agile Controller-DCN
2. Functions of the Agile Controller-DCN
3. **Application Scenarios of the Agile Controller-DCN**
4. Standards and Protocols of the Agile Controller-DCN

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

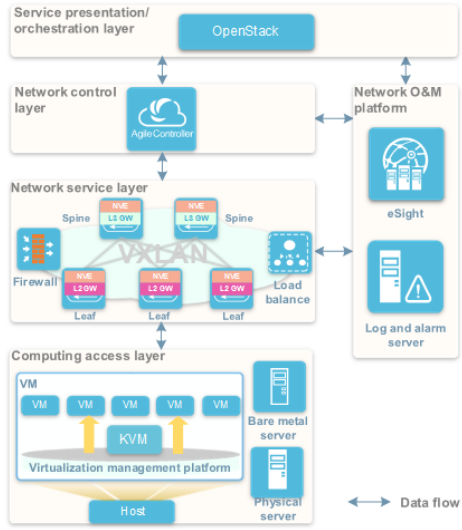
Cloud-Network Integration - FusionSphere



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Traditional data centers have low service provisioning efficiency, low resource usage, and O&M difficulties. Therefore, the traditional data center architecture needs to be evolved to the cloud computing architecture. In the cloud-network integration-FusionSphere scenario, computing and network services are uniformly provisioned based on FusionSphere. FusionSphere provides the unified page to manage computing and network resources. The Agile Controller-DCN interconnects with FusionSphere to implement flexible and convenient resource management, service provisioning, and migration.
- Service presentation/orchestration layer
  - The service presentation layer is oriented to data center users. The cloud platform at this layer provides GUIs for service, network, and tenant administrators, implementing service management, automatic service provisioning, as well as resource and service guarantee.
  - The service orchestration layer consists of Nova, Neutron, and Cinder components of the cloud platform. This layer controls and manages resources through the components to implement virtualization and pooling of computing, storage, and network resources. The components interoperate to realize collaboration of resources.

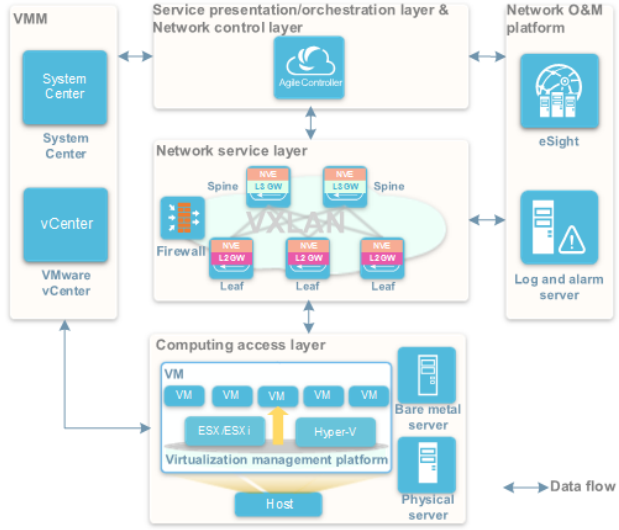
Cloud-Network Integration - OpenStack



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Traditional data centers have low service provisioning efficiency, low resource usage, and O&M difficulties. Therefore, the traditional data center architecture needs to be evolved to the cloud computing architecture. In the cloud-network integration-FusionSphere scenario, computing and network services are uniformly provisioned based on FusionSphere. FusionSphere provides the unified page to manage computing and network resources. The Agile Controller-DCN interconnects with FusionSphere to implement flexible and convenient resource management, service provisioning, and migration.
- Service presentation/orchestration layer
  - The service presentation layer is oriented to data center users. The cloud platform at this layer provides GUIs for service, network, and tenant administrators, implementing service management, automatic service provisioning, as well as resource and service guarantee.
  - The service orchestration layer consists of Nova, Neutron, and Cinder components of the cloud platform. This layer controls and manages resources through the components to implement virtualization and pooling of computing, storage, and network resources. The components interoperate to realize collaboration of resources.

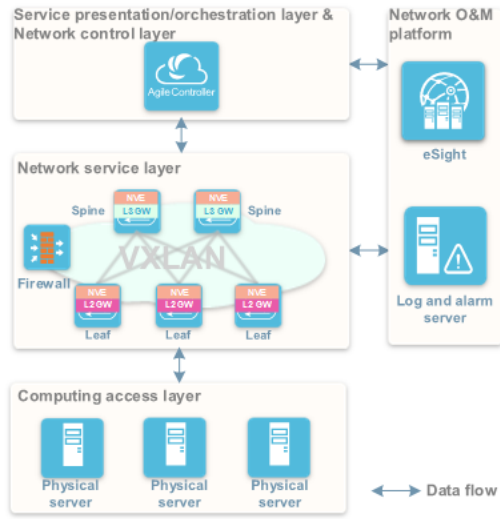
Network Virtualization - Computing



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Different from cloud-network integration scenarios, the network virtualization – computing scenario realizes service provisioning without a cloud platform. The Agile Controller-DCN directly provides a separate management page to manage network resources. The computing system can be associated with network system. If a unified cloud platform cannot be built due to the complexity of the computing service management system or inadequate convergence between computing management and network management, the network virtualization–computing scenario is recommended. In this scenario, the Agile Controller-DCN interconnects with a computing virtualization platform to implement automatic network configuration and collaborative provisioning of computing and network resources and provide flexible and convenient resource management. This scenario is applicable to data center construction without a cloud platform.
- The service presentation layer is oriented to data center users. The Agile Controller-DCN provides GUIs for service, network, and tenant administrators, implementing service orchestration, policy provisioning, automatic service provisioning, as well as O&M.

Network Virtualization - Hosting



 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- In a network virtualization - hosting, the Agile Controller-DCN but not the cloud platform is used to provision services. However, the Agile Controller-DCN does not interconnect with a virtualization platform. The network administrator provisions the network resource services. This scenario is applicable to independent network service provisioning. In this case, the Agile Controller-DCN provides an independent GUI to uniformly manage physical and virtual network resources. In the network virtualization - hosting scenario, a lessor leases equipment rooms and cabinet space to tenants and provides the following types of services to tenants:
  - Basic services: including leasing of equipment space and racks, and network access services such as access bandwidth and outbound bandwidths.
  - VASs: including Layer 4 to Layer 7 VASs such as the security, load balancing, VPN, and NAT services.
- The network virtualization - rack leasing scenario includes the following two scenarios according to whether tenants have gateways:
  - Gateways provided by tenants: Devices hosted by tenants include servers, Layer 2 switches, gateways, and firewalls. Tenant devices access the network of the lessor at Layer 3.

Contents

1. Introduction to the Agile Controller-DCN
2. Functions of the Agile Controller-DCN
3. Application Scenarios of the Agile Controller-DCN
4. **Standards and Protocols of the Agile Controller-DCN**

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Standards Compliance

- Southbound interface protocol
  - Southbound interface protocols include the NETCONF, OpenFlow, and SNMPv3

Standard No.	Description
RFC6241	Network configuration protocol (NETCONF)
RFC3414	Simple network management protocol v3 (SNMPv3), based on a user authentication model
OpenFlow Switch Specification Version 1.3.4	OpenFlow

- Northbound interface protocol
  - Northbound interface protocols include HTTPS and HTTP 1.1

Standard No.	Description
RFC2818	HTTP, based on TLS
RFC2616	Hypertext Transfer Protocol 1.1

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Quiz

1. Single-Answer Question

   Which of the following is not a southbound interface protocol of Agile Controller-DCN?

   A. NETCONF
   B. SNMP
   C. OpenFlow
   D. OSPF

2. Multiple-Answer Question

   Which of the following are overlay networking modes in Huawei CloudFabric Solution?

   A. Host overlay
   B. Physical overlay
   C. Network overlay
   D. Hybrid overlay

 Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

- Answer:
  1. D
  2. ACD

Summary

- Introduction to the Agile Controller-DCN
- Functions of the Agile Controller-DCN
- Application scenarios of the Agile Controller-DCN
- Standards and protocols of the Agile Controller-DCN

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

More Information

- Agile Controller-DCN
  - http://support.huawei.com/enterprise/zh/sdn-controller/agile-controller-dcn-pid-21481886

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.

Recommendations

- Huawei e-Learning website:
  - http://support.huawei.com/learning/Index!toTrainIndex
- Huawei support case library:
  - http://support.huawei.com/enterprise/servicecenter?lang=zh

Copyright © 2019 Huawei Technologies Co., Ltd. All rights reserved.


Thank You
www.huawei.com

Recommendations

- Huawei Learning Website
  - http://learning.huawei.com/en
- Huawei e-Learning
  - https://ilearningx.huawei.com/portal/#/portal/ebg/51
- Huawei Certification
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_31&lang=en
- Find Training
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_trainingsearch&lang=en

More Information

- Huawei learning APP

版权所有© 2019 华为技术有限公司