

**Data Center Solution
V100R001C00**

DC Technical Proposal

Issue 01
Date 2011-08-31

Copyright © Huawei Technologies Co., Ltd. 2011. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions



and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: <http://www.huawei.com>

Email: support@huawei.com

Contents

1 Data Center Network Overview	1
1.1 Introduction to a Data Center	1
1.2 Overall Requirement for the DC Network	1
1.3 DC Network Solution.....	2
2 Service Requirements.....	7
2.1 Overview	7
2.2 Data Service	7
2.2.1 Overview	7
2.2.2 Network Requirements of the Data Service	8
2.3 Web Service.....	10
2.3.1 Overview	10
2.3.2 Network Requirements of the Web Service	11
2.4 Computing Service.....	13
2.4.1 Overview	13
2.4.2 Network Requirements of the Computing Service.....	13
3 DC Network Design	15
3.1 Network Architecture	15
3.2 Core Area Networking Planning	18
3.2.1 Physical Networking Planning.....	18
3.2.2 Reliability Planning	20
3.2.3 Security Planning	22
3.3 Server Area Networking Planning.....	22
3.3.1 Physical Networking Planning.....	22
3.3.2 Channel Separating on Servers	23
3.3.3 Reliability Planning	26
3.3.4 Security Planning	28
3.4 Storage Area Networking Planning.....	28
3.4.1 Physical Networking Planning.....	28
3.4.2 Reliability Planning	29
3.4.3 Security Planning	29
3.5 Interconnection Area Networking Planning.....	30
3.5.1 Physical Networking Planning.....	30

3.5.2 Internet Area	31
3.5.3 Extranet Area	32
3.5.4 Intranet Area	32
3.6 Management Area Networking Planning	33
3.6.1 Physical Networking Planning	33
3.7 VLAN Planning	35
3.7.1 VLAN Overview	35
3.7.2 Principles	36
3.7.3 Recommendation	36
3.8 IP Planning	37
3.8.1 IP Address Planning	37
3.8.2 DNS Planning	37
3.9 Route Planning	40
3.9.1 Routing Overview	40
3.9.2 IGP Design	41
3.9.3 BGP Design	42
3.10 VPN and the Service Area Planning	43
3.10.1 VPN Overview	43
3.10.2 Intranet VPN Service Isolation	43
3.11 QoS Planning	44
3.11.1 QoS Overview	44
3.11.2 QoS Planning Concerning Collaborative Computing	44
4 Desktop-Cloud Network Solution	46
4.1 Desktop-Cloud Service Overview	46
4.2 Desktop Cloud Network Structure	49
4.3 Service Network Planning	50
4.3.1 Bandwidth of the Service Network	50
4.4 Security Planning	51
5 Suggestions on Planning Multiple DCs	53
5.1 Network Architecture of Multiple DCs	53
5.2 Network Reliability Planning	55
5.2.1 Network Reliability Between Regional DCs and Global DCs	55
5.2.2 Network Reliability Between a Country/Region Branch and Regional DCs	56
5.3 Route Planning	57
5.3.1 Routing Overview	57
5.3.2 BGP Design	58
5.4 Disaster Recovery Planning	59
5.4.1 Disaster Recovery Overview	59
5.4.2 Disaster Recovery Overview	60
5.4.3 Network Planning for Disaster Recovery	63
5.4.4 Service Planning for Disaster Recovery	64

5.5 Service Distribution Planning	65
5.5.1 Service Distribution Overview	65
5.5.2 Service Distribution Planning	66
6 DC Network Maintenance Recommendations	68
6.1 Network Management	68
6.1.1 Network Routine Maintenance	68
6.1.2 Customization of Third-Party Devices	77
6.1.3 Software Upgrade and Patch Loading	81
6.2 Troubleshooting	83
6.2.1 Troubleshooting Network Devices	83
6.2.2 Troubleshooting Servers	83
6.3 Network Expansion	85
6.4 Disaster Emergency Maintenance	88
7 Recommended Products	89
7.1 S9300 Series Core Switches	89
7.1.1 Product Overview	89
7.1.2 Product Model	89
7.1.3 Product Characteristics	91
7.1.4 Specifications	92
7.2 S6700 Series Access Switches	94
7.2.1 Product Overview	94
7.2.2 Product Model	94
7.2.3 Product Characteristics	95
7.2.4 Main Specifications	98
7.3 S5700 Series Access Switches	101
7.3.1 Product Overview	101
7.3.2 Appearance	101
7.3.3 Product Characteristics	104
7.3.4 Product Specifications	107

1 Data Center Network Overview

1.1 Introduction to a Data Center

Information is key to an enterprise's competitiveness. As network and communication technologies develop at an ever increasing rate, data centers (DCs) have become the core of the information an enterprise needs to do business. A well-designed data center will improve efficiency and development of enterprises.

The DC of an enterprise is the important as it hosts key service systems, and is a center where the key data of the enterprise is managed. It controls user access, filters packets for security, processes service applications, computes information, and stores data for backup.

A DC consists of the following components:

- Equipment room
- Power supply system
- Network devices including devices on the data network, computing network, and storage network
- Servers including operating systems and application software
- Storage devices
- Security system
- Operation, administration, and maintenance (OAM) system

For enterprises, the trend is to integrate services and data in multiple DCs. This requires the enterprise network to have high level of performance and reliability.

The Huawei DC network solution provides a high performance, secure, and reliable network, which allows the DC to transmit high-quality services.

1.2 Overall Requirement for the DC Network

A DC has a large number of servers deployed and is not only the logical center of an enterprise network but also the source of services. Therefore, a DC should provide abundant bandwidth resources, secure and reliable devices, high-quality network management, and comprehensive value-added services. To create as much value as possible based on limited bandwidth when designing and constructing the DC network, focus on the following requirements:

- **Reliability**

High reliability ensures successful operations of the DC. If the user experience on enterprise services (such as e-commerce or video services) deteriorates due to DC network faults, the service expansion of an enterprise will be hindered, and users will not use the services, decreasing the profits. Reliability is an important aspect when designing an enterprise DC network.

The reliability design is achieved through redundant links, key devices, and key service modules.
- **Scalability**

Each layer of the DC uses devices with a high port density to prepare for the DC expansion.

Devices on the Internet layer, intranet layer, core layer, and aggregation layer adopt the modular design so that capacities of these devices can be expanded flexibly with the development of the DC network.

The scalability of functions enables the DC to support value-added services. The DC provides functions such as load balancing, dynamic content replication, and VLAN to support value-added service expansion.
- **Manageability**

A manageable network is the prerequisite for successful operation of the DC. The DC provides:

 - Various optimized manageable information
 - Complete QoS functions
 - Integrated SLA management system
 - Capability to manage devices of different vendors
 - Independent background management platform for the DC and users to manage the networks
- **Security**

As a concern of DC users especially e-commerce users, security is a key factor during DC construction. DC security is ensured by security control for the physical space and network. The DC provides an integrated security policy control system to ensure DC security.

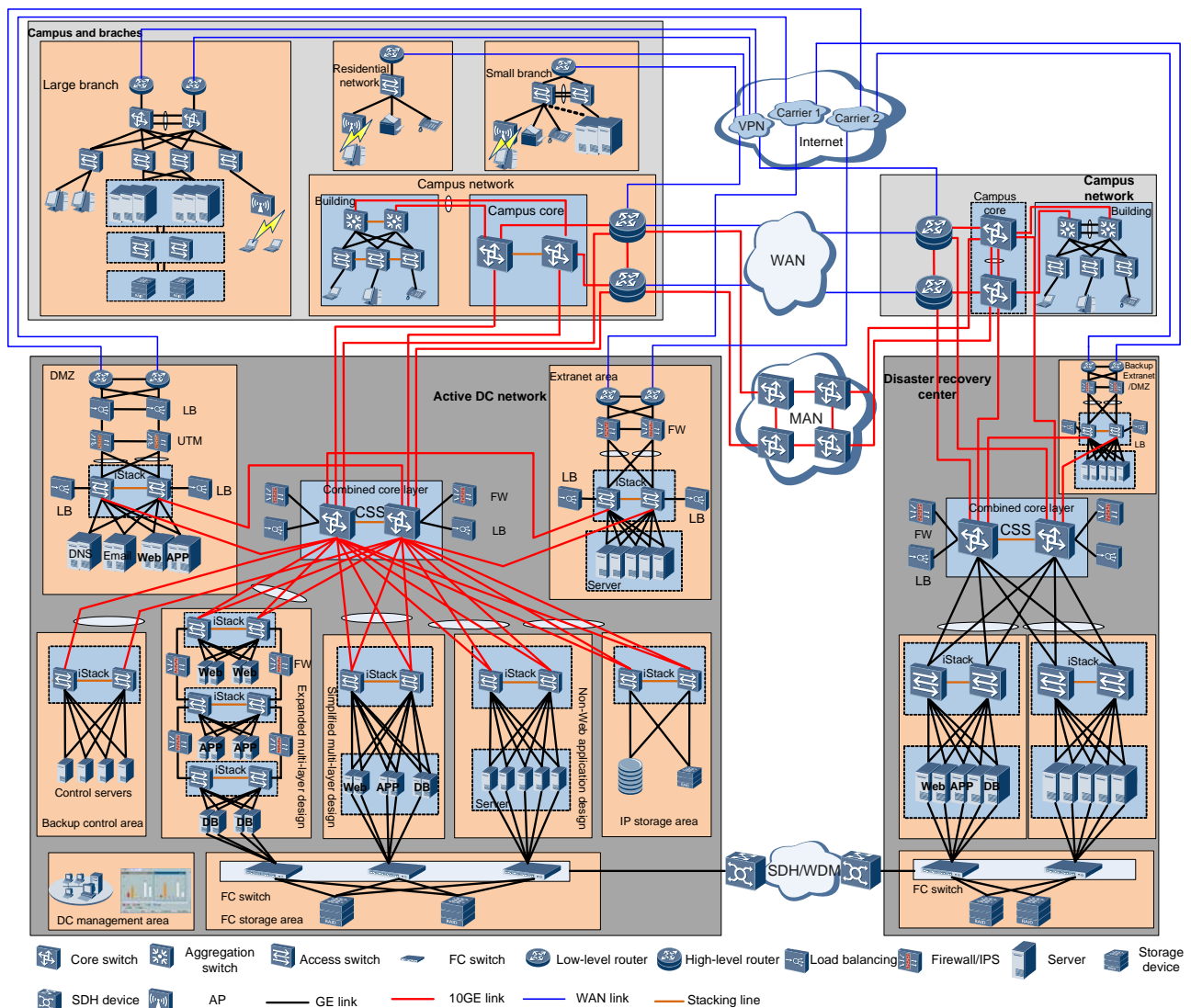
1.3 DC Network Solution

Solution Overview

The Huawei full-service DC solution has the following features:

- The architecture is modular and hierarchical.
- The service network is separated from the management network, ensuring high performance and security of the service network.
- The service bearer network is divided into service areas to provide differentiated services for users.

Figure 1-1 Networking for the DC network solution



As shown in Figure 1-1, to enhance the security, scalability, and maintainability of the network, the Huawei DC solution is divided into the service network, management network, and storage network.

- The service network consists of network access modules and server access modules.
- The management network consists of background management modules.
- The storage network consists of the storage system and the storage area network (SAN).

This technical proposal focuses on the service network and management network.

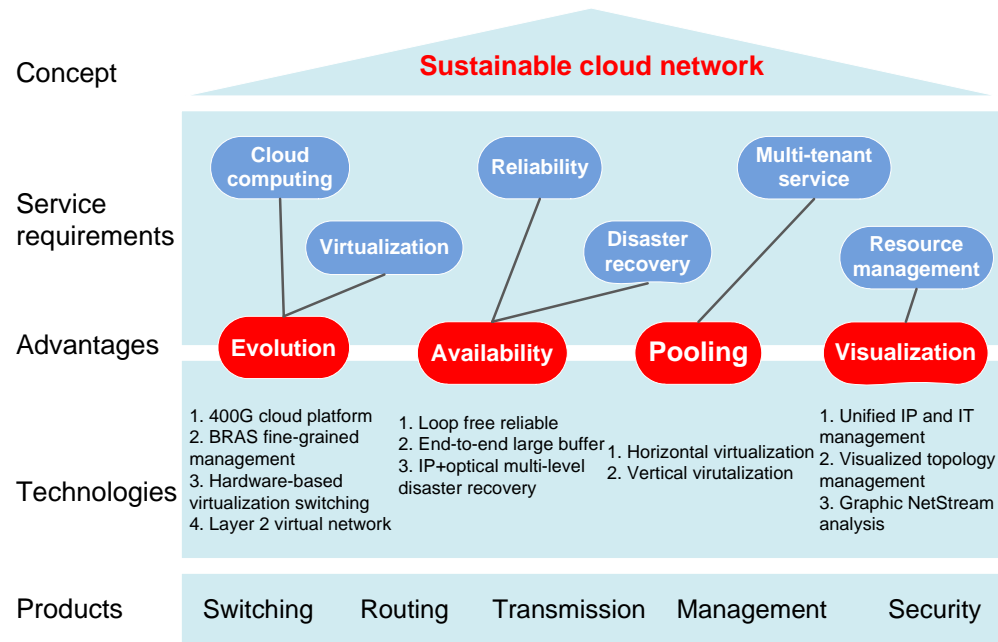
Network access modules include routers, switches, firewalls, load balancers, and unified threat management (UTM) system which contains the firewall, intrusion detection/protection system (IDS/IPS), antivirus, URL filtering, and SSL VPN. These modules provide network a high quality infrastructure with, density, availability, and security.

Server access modules are divided into different service areas based on the types and characteristics of the services provided to the user. The service areas are separated from each other logically or physically.

Advantages of the DC Network Solution

Using Cloud Network as the core concept, Huawei's DC solution is sustainable and supports evolution, availability, pooling, and visualization. Customers can use these features to systematically cope with the challenges of the cloud-computing era.

Figure 1-2 Advantages of the DC network solution



- Evolution: ready for cloud computing and virtual DCs

Cloud network platform with a rate of 400 Gbit/s: The core switches for Huawei's sustainable DC solution use the 10 Tbit/s non-blocking common lisp object system (CLOS) architecture, which can be upgraded to the 400 Gbit/s. These core switches support high-density 40*10GE service boards and 100GE ports, and are fully capable of satisfying capacity requirements of cloud-computing-based ultra-broadband DCs.

Virtualization evolution: Huawei's switches support virtual switching and policy detection defined in the IEEE 802.1Qbg VEPA standard. These functions dramatically improve performance of virtual machines (VMs), provide a clear management model and make traffic manageable and controllable. Huawei switches also support the Intermediate System-To-Intermediate System (IS-IS)-based transparent routing bridge protocols such as IEEE802.1AQ and IETF TRILL. All these enhance network evolution capabilities and make it possible to seamlessly migrate VMs on a large scale.

Desktop cloud fine-grained management: Huawei has introduced the carrier-class BRAS deployment practices to desktop cloud DCs. These desktop cloud DCs support access and management of massive desktop cloud VMs and provide fine-grained bandwidth control and SLA-based hierarchical quality of service (HQoS) for VM users and services.

- Availability: loop free reliable (LFR) Ethernet for non-stop DCs

End-to-end high-reliability architecture: Huawei's sustainable DC solution uses the end-to-end high-reliability architecture that achieves 200 ms convergence time, ensuring business and service continuity for DCs. The LFR Ethernet technology is used to form a fast-convergence loop-free network, implementing Layer 2 switching from the

aggregation layer to the access layer. Carrier-class bidirectional forwarding detection (BFD) and fast reroute (FRR) technologies are used for Layer 3 routing at the core layer and the upper layers. These technologies together with the equipment-level in-service software upgrade (ISSU) and redundant backup of key components create a continuous DC.

LFR Ethernet: Switches used in Huawei's sustainable DC solution use CSS+LAG+iStack technologies, which establish an LFR Ethernet network. This network has the reliable physical-layer hard cluster, the convergence time of 200 ms, and the cluster bandwidth of 256 Gbit/s.

Flattened no-packet loss network: High-end switches used in Huawei sustainable DC solution buffer data on 10GE/GE interfaces within 200 ms. The S12700 core switch and the S9300 switch (for EOR access) provide the following functions:

- Constitute a flattened network
- Implement end-to-end large-buffer deployment
- Bring low delay and prevent packet loss triggered by burst traffic for services such as distributed computing services.

IP+optical multi-level disaster recovery: The Huawei DC solution integrates optical transport devices and routers to provide a complete range of data- and service-level disaster recovery and backup capabilities. The optical transport network (OTN) devices provide 14 types of specialized storage interfaces such as FC, FICON, and ESCON interfaces. These interfaces support real-time hardware backup between DCs and their disaster recovery centers. NE40E routers provide flexible network interconnections and an IP SAN between DCs.

- Pooling: network resources pool for on-demand scheduling

On-demand resources scheduling: Multiple switches are virtualized into one logical switch using CSS and iStack technologies so that 100% of the network resources are shared. This is more efficient than the switch using the conventional STP technology where only 50% of the network resources are shared. A series of multi-instance technologies such as MPLS VPN and MCE ensure that resources in the network resources pool can be flexibly scheduled as required by services.

Simplified network structure: One logical switch that is virtualized from multiple switches serves as one network element (NE) on the NMS. This simplifies network architecture and reduces management and configuration workloads.

Effective service isolation: The multi-instance technologies such as MPLS VPN and MCE ensure isolation and security of DC services. In addition, access from multiple departments to DC servers can be controlled by flexibly configuring VPN access policies.

- Visualization: intelligent and visualized NMS for unified IP&IT management

Unified IP&IT management: The eSight, an intelligent NMS, can uniformly manage multiple devices and associate systems in DCs, such as network devices, servers, and enterprise application systems. It reduces costs and improves operation and maintenance efficiency. It provides open platforms that allow deep integration and wide collaboration with market-leading IT vendors such as IBM, HP, and Oracle.

For details about Huawei's OSS partners, visit the website http://www.huawei.com/partners/integrated_with_oss.do.

Visualized topology management: The eSight provides network topologies and service views, making service deployment and network configuration more visualized and convenient.

Graphical NetStream analysis: Switches and routers provide embedded NetStream boards or modules to monitor distribution of DC services in real time. Using eSight, users can obtain graphical NetStream analysis reports and also easily make service plans.

2 Service Requirements

2.1 Overview

A DC deploys various service systems in a centralized mode to integrate them. This helps to analyze services, make decisions, and maximize the information production capability.

A DC also provides Web portals, which help to establish channels with customers and improve the enterprise's brand awareness, product promotion, and customer service. With the Web portals, the enterprise can implement ecommerce and other Internet-based businesses.

In addition, a DC provides high-performance computing services, such as 3D rendering, medicine research, gene analysis, and Web search.

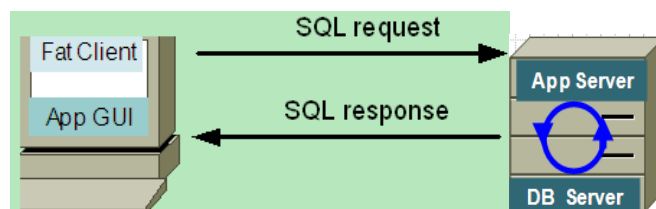
In an enterprise, a DC may provide all the preceding services concurrently. These services may be independent of each other or be integrated into a large service system. You must analyze the real situation when planning a network for the DC.

2.2 Data Service

2.2.1 Overview

The data service is the most basic service in a DC. Typical data services in an enterprise include file storage, mail service, and enterprise resource planning (ERP). The client/server (C/S) model is the basic service model.

Figure 2-1 Client/Server service mode



The C/S model consists of the following two parts:

- Client (usually a PC). A client is deployed on a campus network or an enterprise branch. SQL requests are sent from a fat client to the server and SQL responses from the server are displayed on the App GUI.
- Server. A server is deployed in a DC and stores data in a dedicated storage device. As shown in [Figure 2-1](#), a server used by the database is called DB server, a server used by applications is called App server, and data in the database is stored in a dedicated storage device (not displayed in the figure).

2.2.2 Network Requirements of the Data Service

The data service is processed as follows:

- a. The client sends a request.
- b. The server and the storage device receive and process the request.
- c. The server sends a response to the client.

The network requirements include:

- Traffic requirement

Traffic is generated by requests and responses between the client and the server. Traffic is unbalanced and becomes high during peak hours on special dates or periods. The network bandwidth must be planned to accommodate peak traffic times, and certain bandwidth must be reserved for future growth and improvements.

The number of clients and concurrent services must be also considered for network bandwidth planning. The number of concurrent services is used to configure the bandwidth convergence ratio between network devices at different layers, because no network traffic is transmitted between servers.

For example, the peak hour of each service falls on the closing date of a service or event, such as the closing date of production, a sale, or attendance services. If these closing dates fall on different dates, use the highest peak traffic rate of the three as the network bandwidth peak. If the closing dates of these services fall on the same date, use the total peak traffic rate of three services as the network bandwidth peak.

The data service has no special requirements for delay as long as the user experience is met. In most cases, the response time of a database is less than 2 seconds. The forwarding delay of the DC network is less than 1 millisecond, occupying a small proportion of the total response time. The forwarding delay of WAN is about 300 milliseconds and the time for processing data is tens of milliseconds. Some special services require short delays, for example, the stock exchange requires the network forwarding delay to be less than 5 milliseconds.

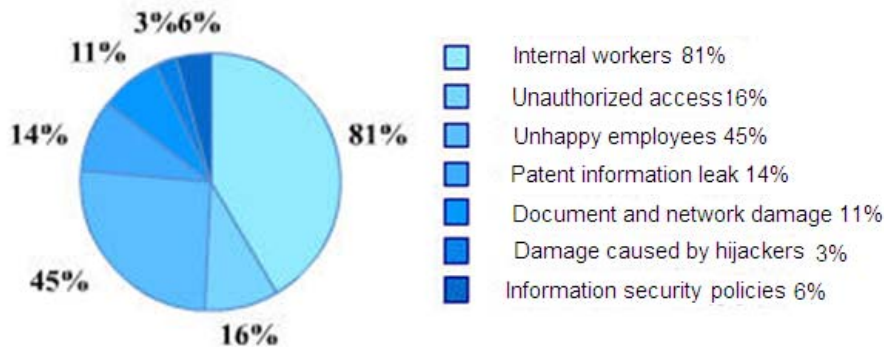
- Security requirement

A DC is an integrated IT application environment where a large amount of data is stored. It requires the highest security in the IT system.

In America, according to the survey conducted by CSI/FBI, the loss incurred by enterprises and government institutions caused by unauthorized access is more than that caused by viruses and hackers. The internal threats account for more than 80% of the total threats, as shown in [Figure 2-2](#).

In China, according to the latest statistics from the ministry of public security, the internal threats account for approximately 70% of the total threats. Approximately 80% of enterprises and organizations using computers have not established a security management system, or taken technical measures or regulations.

Figure 2-2 Factors that affect the information security in enterprises



In an enterprise, key services such as the financial service are transmitted as a data service and require high security. In addition to physical security measures, protection measures are also required on the network, including isolating different services, identifying and handling the traffic and virus attacks. Services are isolated, enabling terminals to access only servers of specified services.

- Reliability requirement

The data reliability is required and varies according to the service type (internal service and external service) on the network.

The internal service system does not require high network reliability. A fault occurring in a DC internal part recovers within 20 minutes to 30 minutes, and a fault occurring in the entire DC recovers within 4 hours to 8 hours during which services are implemented from the standby DC.

The external service system requires high network reliability. A fault occurring in a DC internal part recovers automatically or can be manually rectified within 10 minutes, while a fault occurring in the entire DC recovers within 2 hours during which services are implemented from the disaster recovery center.

The disaster recovery system is classified into the following seven tiers according to the international standard Share 78:

- Tier 0: No off-site data
- Tier 1: Data backup with no hot site
- Tier 2: Data backup with a hot site
- Tier 3: Electronic vaulting
- Tier 4: Point-in-time copies
- Tier 5: Transaction integrity
- Tier 6: Zero or near-Zero data loss
- Tier 7: Highly automated, business integrated solution

Two technical indicators are used to measure disaster recovery:

- Recovery point objective (RPO): acceptable amount of data loss
- Recovery time objective (RTO): acceptable longest duration within which services are interrupted or the shortest duration between the time when a disaster occurs and the time when services are restored

RPO measures data loss, while RTO measures service loss. RPO and RTO are not necessarily related. RTO and RPO vary according to services and enterprises, and are

calculated based on service requirements after risk analysis and service influence analysis are performed.

If the RTO is shorter and the RPO is newer, the service loss will be less. The costs in developing and building the system, however, will become higher. Both factors are an important consideration.

- Cloud-computing requirement

In most cases, service systems of the data service do not operate concurrently. To efficiently utilize the server resources, deploy multiple virtual servers on a physical server to host different service systems. This is the easiest way to apply cloud computing. When deploying multiple virtual servers on a physical server, consider the bandwidth requirement of each service to prevent one service from occupying the bandwidth of other services on the same server.

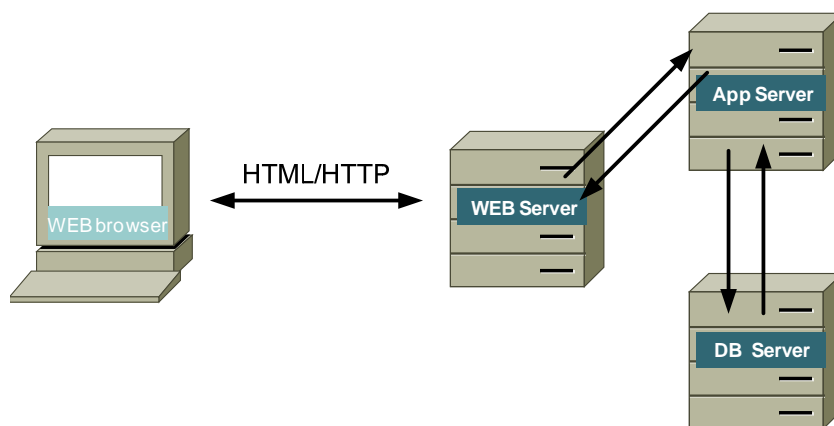
In a word, the network requirements of the data service guarantee bandwidth and security.

2.3 Web Service

2.3.1 Overview

As the Internet flourishes, the Web service takes up a larger proportion in enterprise services. The following two reasons accounts for the popularity of Web service in enterprises. The Web service provides a convenient way for users to access the information and perform the e-commerce on the Internet. The Web service also solves problems in the C/S model, such as large workload due to client software maintenance.

Figure 2-3 Web service model



As shown in [Figure 2-3](#), the Web service model adds a Web server and an App server to form a three-layer structure. Services are processed in the following process:

- The App server (App Server in Figure 2-3) processes services sent from the client on the Web browser using HTML or HTTP.
- The DB server and storage system provide DB services.
- The Web server displays information for users.

The three-layer structure enhances flexibility of the service system. You can modify the service system on the Web server, application server, or DB server. Users only need to refresh the web page on the Web browser to view the modification.

2.3.2 Network Requirements of the Web Service

Unlike the data service, the Web service requires a Web server and an application in the DC. Traffic is transmitted between the Web and application servers, and between the application server and DB server.

The network requirements include:

- Traffic requirement

The Web service traffic (such as requests and responses) is transmitted between the clients and servers, and also between the servers. The Web service traffic, however, is unbalanced just like the data service traffic.

You need to learn about deployment modes before planning bandwidth. The Web service can be deployed in layered and flattened modes, as shown in [Figure 2-4](#) and [Figure 2-5](#).

- To deploy a large number of Web servers, application servers, and DB servers in a large DC, you can deploy them in layered mode.
- To deploy a small number of servers in a small- or mid-scale DC, Huawei recommends the flattened deployment mode.

Figure 2-4 Layered deployment mode

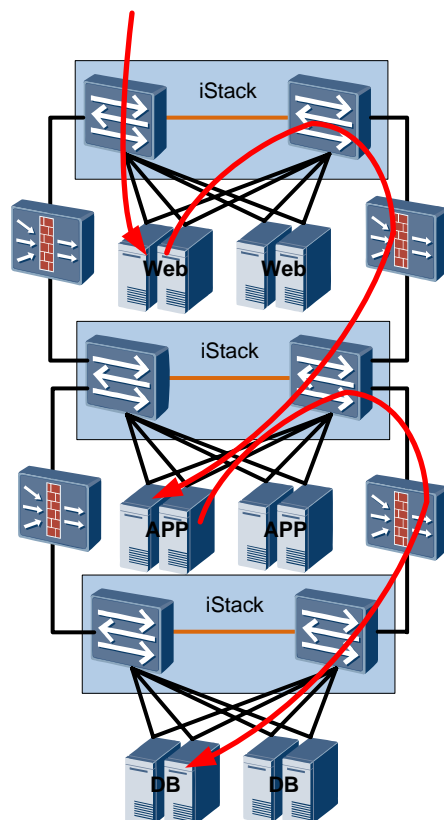
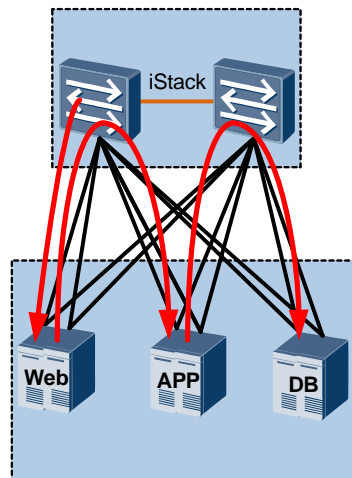


Figure 2-5 Flattened deployment mode



In the layered deployment mode, bandwidth is planned for each layer. In the flattened deployment mode, traffic between servers is aggregated to one server and bandwidth is planned based on the total traffic volume. The traffic between clients and DC is much smaller than that within the DC.

The Web service traffic is transmitted through more servers and network devices than the data service traffic. Therefore, the Web service requires a shorter network delay. The Web service interaction process is different from the data service interaction process. The Web server responds to the request from a client. The application server and DB server then process the request. Finally, request information is displayed on web pages.

Therefore, the delay of the Web server's response to the requests from clients must be short.

- Security requirement

In the Web service mode, the client and DB server are isolated by the Web server and application server. This enhances the security of the DB server and data. Traffic is transmitted among the Web server, application server, and DB server hop by hop over the network channels, which is vulnerable to hop-by-hop attacks.

Web services, especially services for Internet users, are faced with more threats because:

- The attack sources are well organized and industrialized. Attacks may come from anywhere on the Internet.
- The service system is more complex. Security holes may exist in the operating system, Web server, application server, and DB. A hole in one system may cause other systems to be corrupted one by one.
- When internal users are accessing the Internet, they may be intruded by unauthorized users and used for attacks.

- Reliability requirement

In a three-layer structure, the Web service is processed by servers at three layers together and interactions between servers are more frequent, so higher network reliability is required. The overall fault recovery time is not prolonged; however, the network reliability must improve so that the DC availability can remain unchanged in such a serial system.

The link error rate of the link between a switch and a server is $1 \text{ h}/1000 \text{ h}$. In Web service mode, a switch is connected to the Web server, application server, and DB server and three links are available. Therefore, the link error rate is $1 - (1 - 1 \text{ h}/1000 \text{ h})^3 \approx 3$

h/1000 h. If you want to keep the error rate of the entire service at 1 h/1000 h, reduce the link error rate to 20 min/1000 h.

2.4 Computing Service

2.4.1 Overview

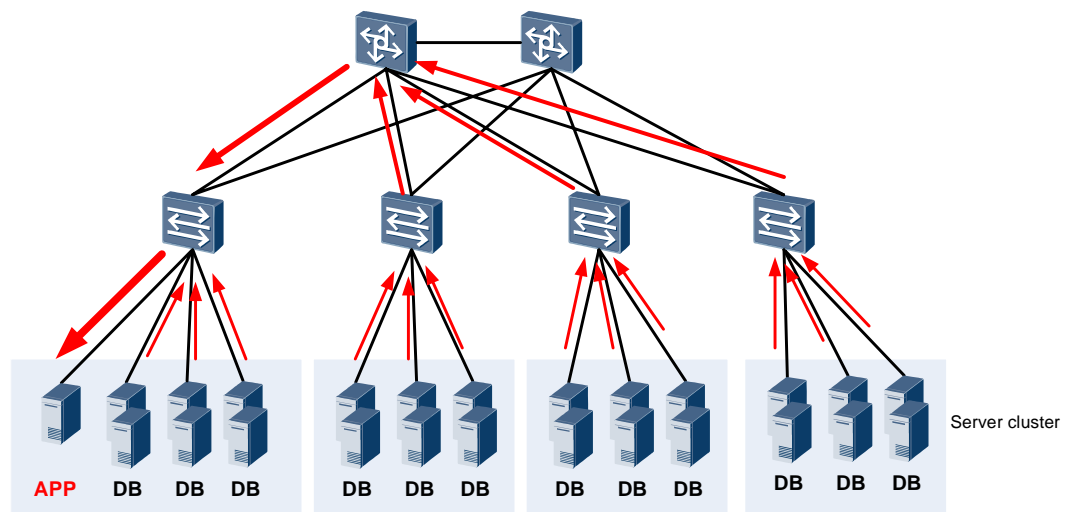
The computing service is a service requiring high computing performance, such as 3D rendering, medicine research, gene analysis, and Web search.

In the computing service mode, a large number of common servers work collaboratively as a cluster to process a computing task.

2.4.2 Network Requirements of the Computing Service

The computing service traffic is transmitted between servers, as shown in [Figure 2-6](#).

Figure 2-6 Traffic of the computing service (server cluster)

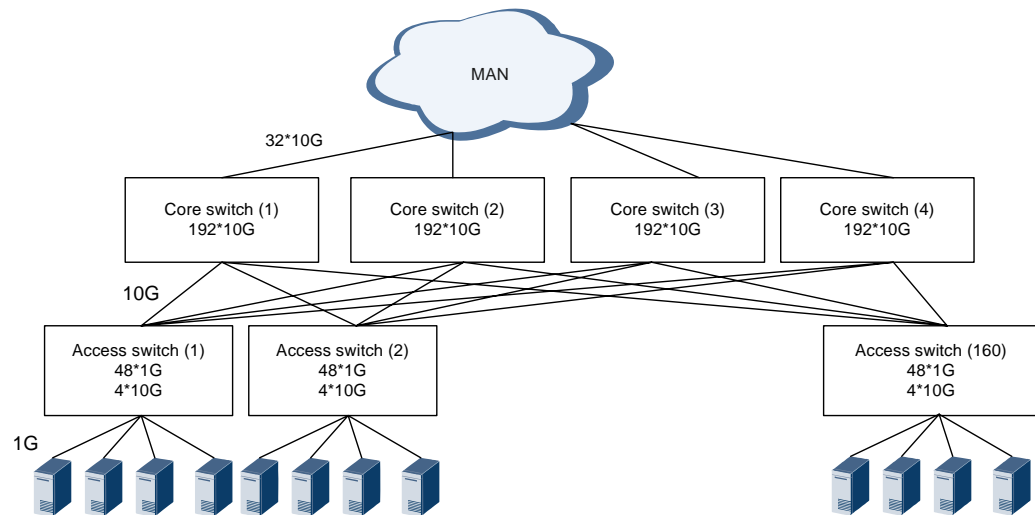


The application server distributes the computing service to a large number of DB servers, and the DB servers return the results to the application server. The network requirements include:

- Instantaneous traffic buffering capability
The application server must have a scheduling mechanism to distribute services. Otherwise, the results sent from the DB servers arrive at the application server in a short time period. The burst traffic rate exceeds the interface bandwidth on the application server. If the network cannot buffer the traffic, packets are lost and the application server cannot process all the services. This leads to more frequent interactions between the application server and DB servers and prolongs the overall processing time. Therefore, the network must be capable of buffering packets to eliminate packet loss.
- Non-blocking network
Different from the cluster model shown in [Figure 2-6](#), the cluster model shown in [Figure 2-7](#) provides interconnection between services on all servers. In this service system, servers communicate with each other in point-to-point communication mode.

Non-blocking forwarding allows any two servers can communicate services with each other so that the forwarding capability is not limited by the location.

Figure 2-7 Server cluster for the Internet service



3 DC Network Design

3.1 Network Architecture

Design Principles of a DC Network

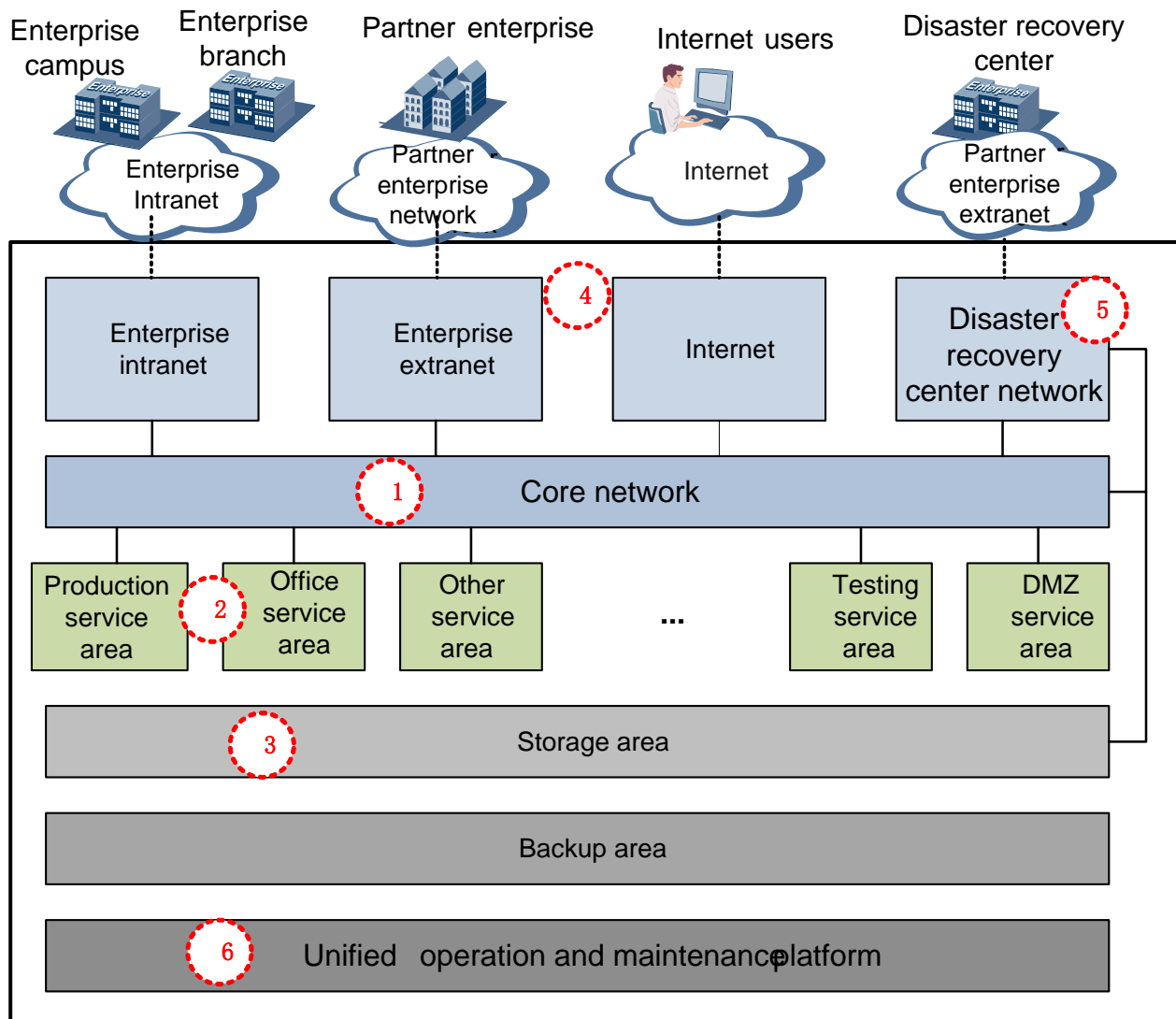
The DC network design is based on the following principles:

- **Modular architecture**
The network is deployed in modular architecture that can expand for service adjustment and development.
- **High reliability**
The network implements redundant backup of key devices and links. Highly reliable key devices are made up of hot swappable boards and modules, and support redundancy of control modules and power supplies. Network layers are reduced to simplify network architecture and enhance networking reliability.
- **Secure isolation**
The DC network adopts effective security control policies that logically isolate data based on services and rights, and uses physical isolation methods to ensure security of important service data.
Services such as server-centered services, IP storage and backup services, and management services are isolated logically. The management network is isolated from other networks physically.
- **Manageability and maintainability**
The network is highly manageable. To facilitate maintenance, use integrated products with universal modules.

Logical Architecture

Figure 3-1 shows the logical architecture of a DC.

Figure 3-1 Logical architecture of a DC



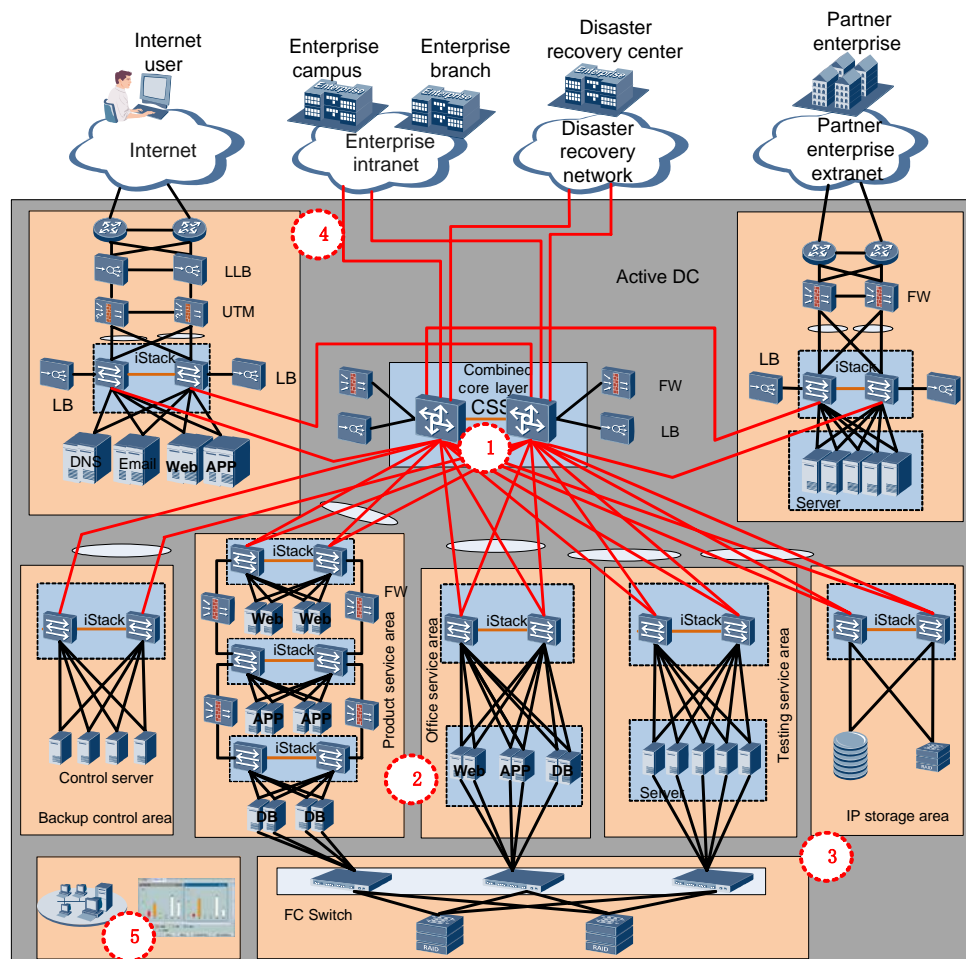
- **Core network area**
This area is the core of the DC network, and connects the inner server area, enterprise intranet, partner enterprise network, disaster recovery center, and external user network.
- **Server area**
Servers and application systems are deployed in this area. Based on security and scalability, the server area is divided into the production service area, office service area, testing service area, and the demilitarized zone (DMZ) area and other service areas.
- **Storage area**
Storage devices for the fiber channel (FC) SAN and IP SAN are deployed in this area.
- **Interconnection area**
In this area, internal and external enterprise users are connected to the DC. Based on security and scalability, the Internet area is divided into the enterprise intranet, enterprise extranet, and the Internet.

- The intranet interconnects the headquarters and branches through the enterprise campus network and the wide area network (WAN).
- The enterprise extranet connects the partner enterprise network using the metropolitan area network (MAN) and the WAN leased lines.
- The Internet allows public users, staff on a business trip, and office users without a WAN network to access the Internet safely.
- Disaster recovery center Internet area
In this area, the disaster recovery centers in the same city are interconnected by transmission devices and disaster recovery centers in different cities are interconnected by the WAN leased line.
- OAM area
The network, server, application system, and storage devices are managed in this area. The functions of the OAM area include fault management, system configuration, device performance, and data security management.

Physical Network Architecture

Figure 3-2 displays the physical architecture of the DC.

Figure 3-2 Physical network architecture of the DC



The modular DC architecture has the following features:

- Extensible architecture
 - The architecture consists of five independent extensible areas: the core area, server area, storage area, interconnection area, and management area.
 - The architecture is a star topology with the core node as the root node.
- Core area as the traffic hub
 - The core area employs core switches with a large capacity and high performance.
 - High-density 10GE ports are deployed in this area.
- Service areas and management areas
 - Service areas can be extended independently.
 - Server-centered networks for data, management, and storage can be extended independently.
- Interconnection areas
 - The four interconnection areas can be extended independently.
 - The disaster recovery interconnection network ensures that services can be smoothly migrated to other DCs.

3.2 Core Area Networking Planning

The core area is the center of the whole DC network, and connects the server area and the interconnection area. The core area transmits internal and external data traffic, and becomes the logical center for network reliability and security design.

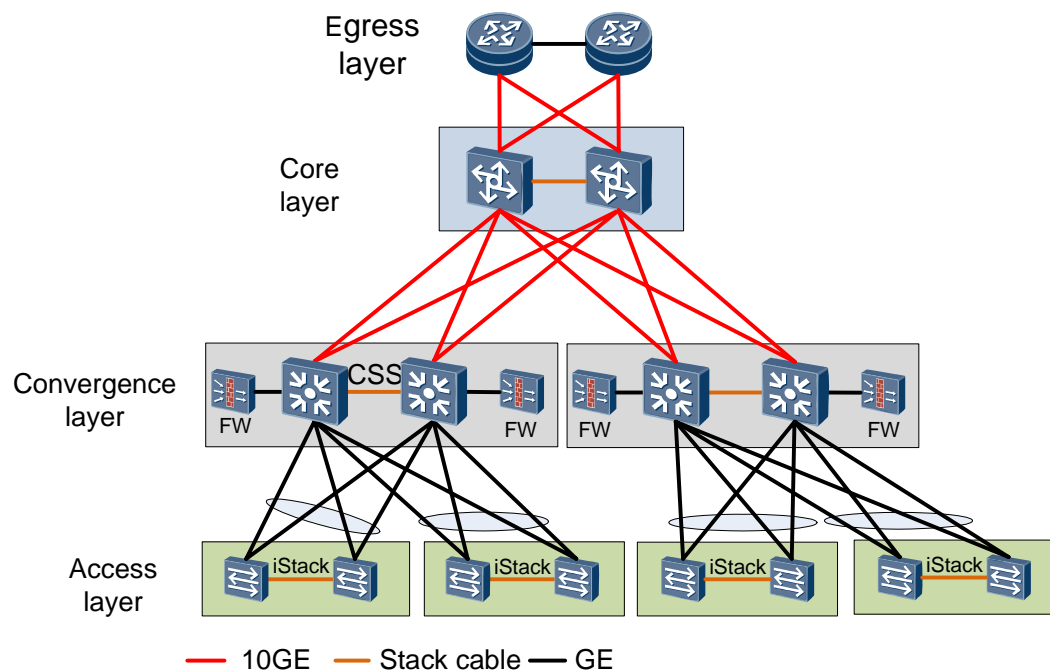
3.2.1 Physical Networking Planning

The physical network is established in the following two methods to connect the core area to the server area: one is a Layer 3 design that deploys the core layer, aggregation layer, and access layer, the other is a flattened design that integrates the core layer with the aggregation layer.

Layer 3 Networking

[Figure 3-3](#) shows the Layer 3 networking diagram. The core layer and the aggregation layer are separated in this networking. Each aggregation area has security devices such as firewalls deployed.

Figure 3-3 Layer 3 networking

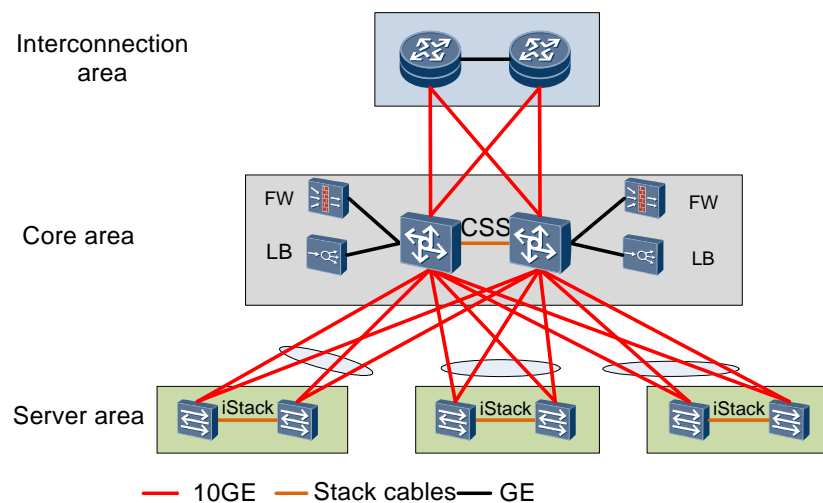


Flattened Networking

Figure 3-4 shows the flattened networking diagram. In the flattened networking, devices in the core area and the aggregation area are replaced by two large-capacity switches in a combined core area. Security devices such as firewalls of large capacities are deployed in this area.

Huawei recommends the flattened networking, which simplifies the network topology and improves data transmission efficiency.

Figure 3-4 Flattened networking in the core area



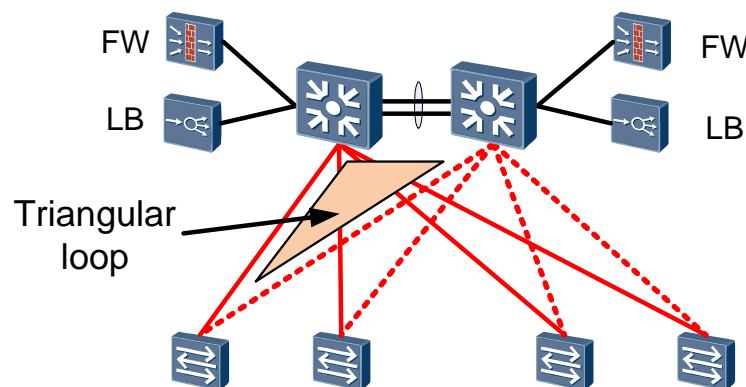
3.2.2 Reliability Planning

As shown in [Figure 3-3](#) and [Figure 3-4](#), redundancy of devices and links ensure reliability of the DC network.

If the access layer runs Layer 3 routing protocols and communicates with the core layer through Layer 3 routing, Bidirectional Forwarding Detection (BFD) and equal-cost paths are deployed to implement fast fault detection and switchover and improve usage of redundant links.

In most cases, Layer 3 routing protocols run at the core layer, which causes Layer 2 loops between the access layer and the core layer. [Figure 3-5](#) shows the design to protect the network against Layer 2 loops using Spanning Tree Protocol (STP) and Virtual Router Redundancy Protocol (VRRP).

Figure 3-5 STP networking



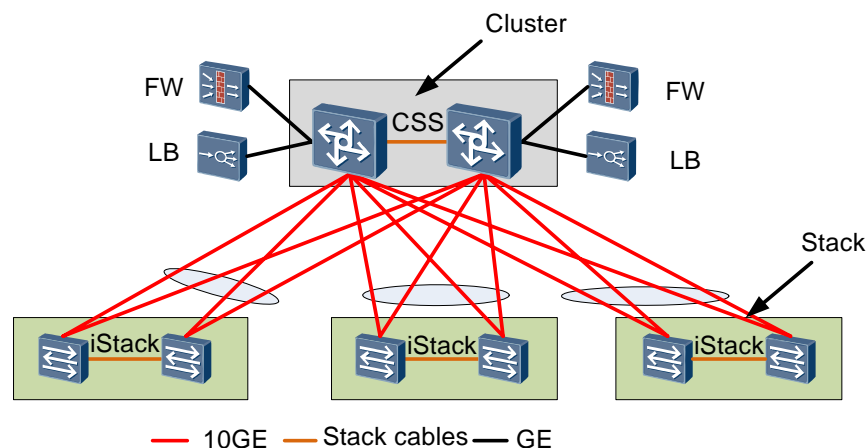
As shown in [Figure 3-5](#), dotted lines represent links that are blocked by STP. This plan uses the standard STP protocol to integrate devices from multiple vendors into a hybrid network.

The disadvantages of the plan are:

- Long convergence time
The traditional STP technology makes the network converge slowly. It takes more than 10 seconds to restore services after a fault occurs. RSTP increases the convergence speed to some extent, but the convergence still takes several seconds. A service interruption for several seconds lowers user experience.
- Low link usage
If servers in the same rack belong to the same VLAN, the bandwidth of an uplink cannot be used. In this case, the bandwidth usage is only 50%. The Multiple Spanning Tree Protocol (MSTP) optimizes the bandwidth usage based on VLANs but it cannot solve the problem completely.
- Complex configuration that is difficult to maintain, and frequently occurred faults on the network
Every access switch or aggregation switch needs to run the STP protocol. When more access switches are added to the network, the STP processing becomes more complicated, which reduces the network reliability.

Loop-free networking with cluster and stacking is used to overcome these disadvantages.

Figure 3-6 Loop-free networking



The combined core layer uses two framed switches as a cluster. The access layer uses box switches to form a stack system. Links between switches at the access layer and the combined core layer form an Eth-Trunk.

The loop-free networking design has the following advantages:

- Simplified management and configuration
The cluster and stacking networking reduces managed nodes by more than a half.
In addition, it simplifies the network topology and configuration because it does not need complex protocols such as STP, Smart Link, and VRRP.
- Fast convergence
The convergence time is less than 10 ms after a fault occurs, which significantly reduces the impact on services caused by faults on links and nodes.
- High bandwidth usage
Links form a trunk so that the bandwidth usage reaches 100%.
- Easy to expand the capacity, saving investment
When new services are provided, the enterprise can add devices directly to upgrade the network. The network capacity can be expanded without changing the network configuration, saving users' investments.

The loop-free networking improves the network reliability rate from 99.9% to 99.9999%. The fault rate on a single link is reduced from 1 hour to 3.6 seconds in 1000 hours.

Framed switches are provided in the core area to ensure network reliability in the following ways:

- The MPUs work in backup mode.
- The power supplies work in backup mode.
- Modular design of fans is provided, in which a single-fan failure does not affect system running.
- All modules are hot swappable.
- The CPU defense function is configured.
- Complete alarm functions are provided.

3.2.3 Security Planning

Firewalls are provided in the core area to ensure network security in the following ways:

- Restrict communication between server areas to isolate services.
- Restrict the communication between the enterprise campus network and server areas to ensure access security between clients and servers.
- Restrict the communication between the enterprise branch network and server areas to ensure access security between clients and servers.

3.3 Server Area Networking Planning

3.3.1 Physical Networking Planning

Access switches are placed in server racks or in independent network cabinets to provide Layer 2 switching functions. Switches in server racks are top of rack (TOR) switches, and those in independent network cabinets are end of row (EOR) switches.

The TOR access mode is applicable to high-density rack servers, and the EOR access mode is applicable to low-density cabinet servers, such as small servers. [Table 3-1](#) shows the differences between the two modes.

Table 3-1 Differences between EOR and TOR access modes

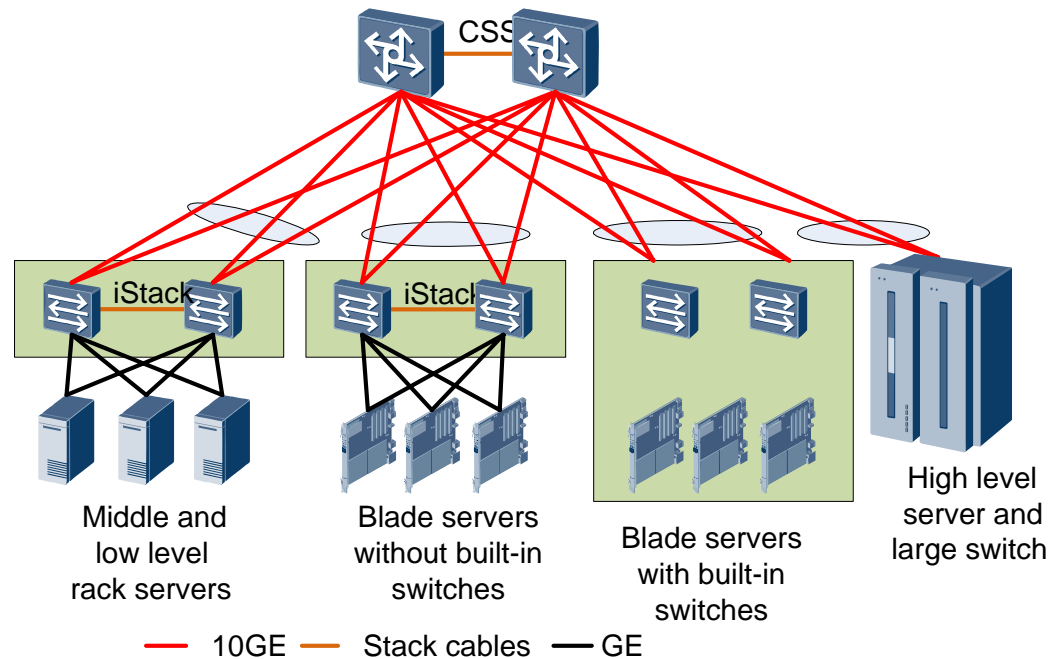
Item	EOR Access Mode	TOR Access Mode
Position of the access switch	In a network cabinet	In the same cabinet as servers
Type of switches	Framed switches Box switches	Box switches
Cabling	A large number of long cables connecting servers to the EOR switches	Short cables in a cabinet connecting servers to TOR switches A small number of cables connecting servers to core switches
Management	Low management cost (because a few EOR switches can connect to a large number of servers)	High management costs (because a large number of TOR switches need to be managed)
Application scenario	Low-density servers in the cabinet	High-density servers in the cabinet

Servers access the network in the following ways:

- A large number of middle- and low-level rack servers access the network using access switches.

- A small number of high-level servers are connected directly to core/aggregation switches to ensure bandwidth.
- Blade servers without built-in switches access the network using access switches.
- Blade servers with built-in switches directly connect to core/aggregation switches to reduce the number of network layers and improve network performance.

Figure 3-7 Access modes for servers



3.3.2 Channel Separating on Servers

The processing capacity of the CPU on a server has been significantly improved since the CPU processor has developed from single core to 128 cores. Compared with the CPU, the IO capacity is still limited. The IO development becomes a bottleneck in the network. To fully use the high-performance CPU, a server must work in multiple channels and use multiple network ports that are physically isolated.

Figure 3-8 Multiple channels on a server

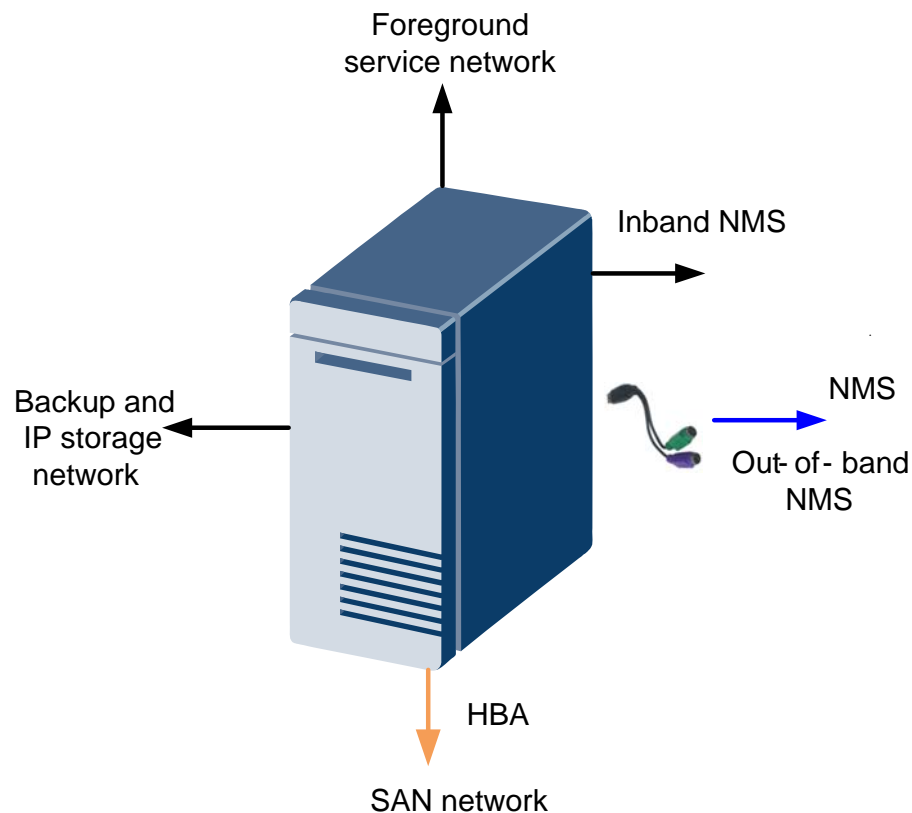


Figure 3-8 shows multiple channels on a server. A server has four types of ports that are used to access the following networks:

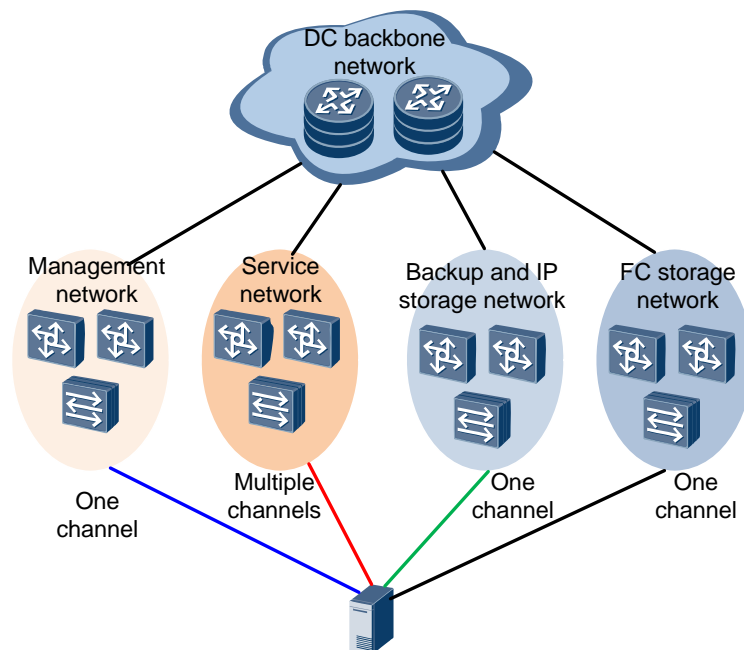
- Service network
- Network management and the keyboard video mouse (KVM) network
- SAN network
- Backup and IP storage network

A server working in multiple channels has the following advantages:

- Improves the IO capacity.
- Separates traffic of different services safely.

Figure 3-9 shows the logical networking architecture of multiple channels on a server.

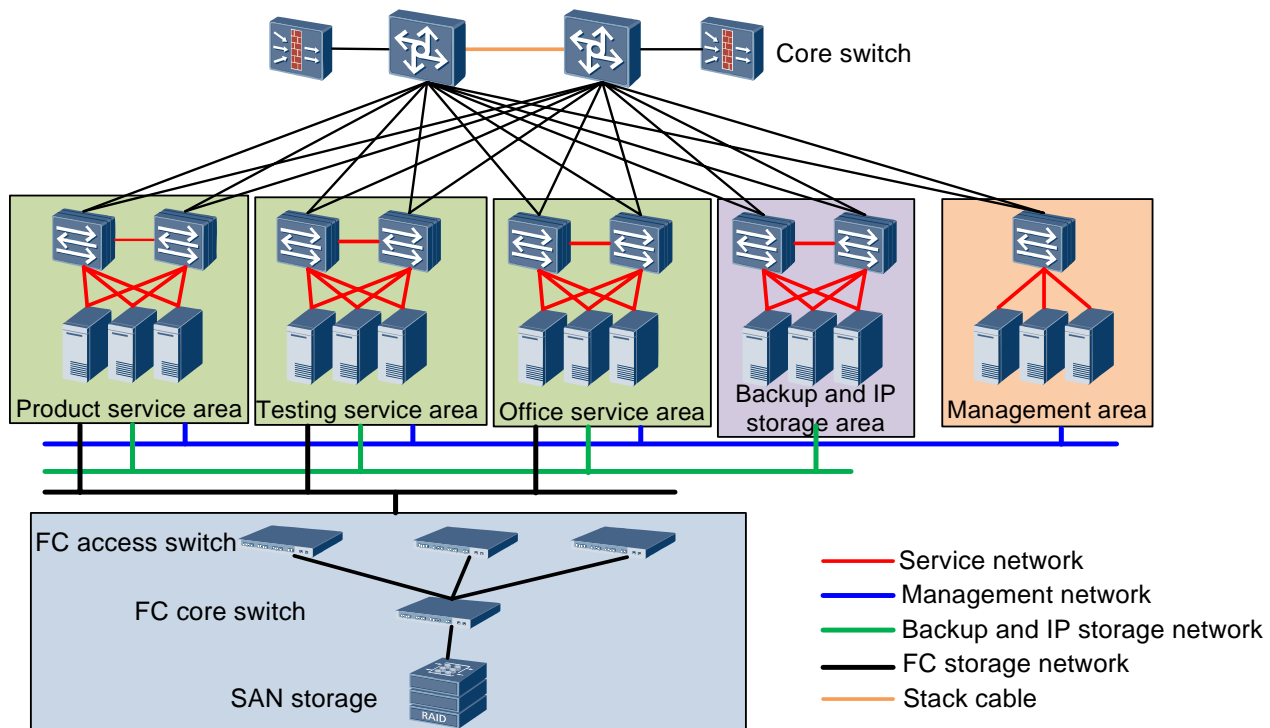
Figure 3-9 Separated networks



The server area is divided into four physically isolated networks: the service, management, storage, and backup networks. The server accesses different networks using network interface cards (NICs).

Figure 3-10 shows the physical network topology.

Figure 3-10 Physical network topology



3.3.3 Reliability Planning

Overall Reliability Planning

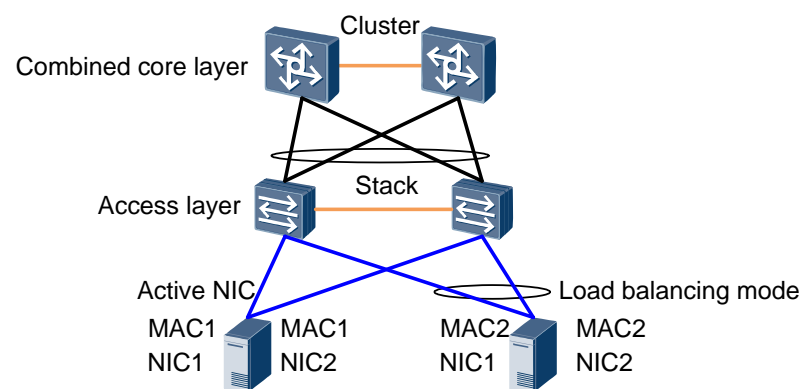
Server area reliability includes reliability of the network, devices, and servers.

- Loop-free cluster and stacking network ensures network reliability. For details, see section 3.2.2 "Reliability Planning."
- Access switches are stacked to ensure device reliability.
- Dual NICs ensure server reliability.

The network drive binds multiple NICs into a virtual NIC. The virtual NIC has a unique IP address to communicate with external devices. The server supports NIC teaming. If an NIC fails, the standby NIC shares its MAC address. The two NICs working in active/standby mode or in load balancing mode improves reliability.

Dual NICs in Active/Standby Mode

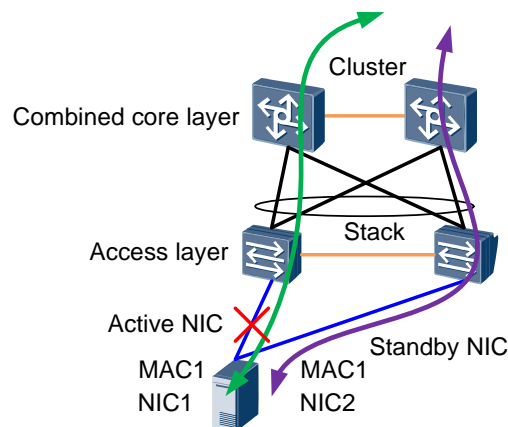
Figure 3-11 Networking for server reliability



The two NICs in active/standby mode have the same MAC address (such as MAC1 in Figure 3-11). When the active NIC fails, the server switches the traffic to the standby NIC and sends a gratuitous ARP packet from the standby NIC. Network devices must properly process gratuitous ARP packets to switch the traffic to a new directory.

Figure 3-12 shows the change of the data transmission route. Data is transmitted in the green route using the active NIC. If the active NIC fails, the data transmission route is changed from the green one to the purple one.

Figure 3-12 Change of the data transmission route using active and standby NICs



When the access switch receives a gratuitous ARP packet, it changes the outbound interface matching MAC1 to the link connected to the standby NIC. You need to add the two ports of active and standby NICs to the same VLAN and bundle the links so that the outbound interface can be updated when a switchover occurs.

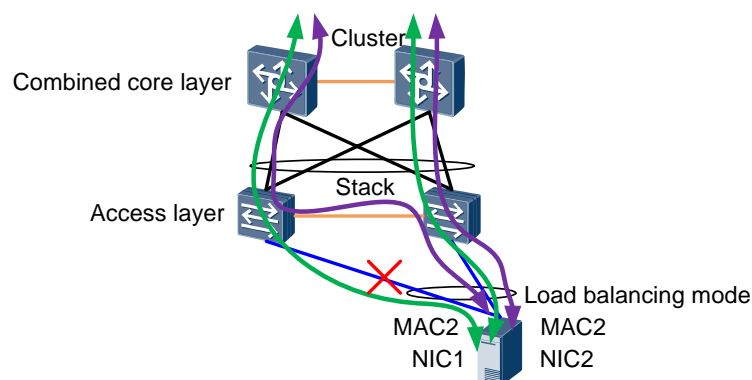
Switches at the combined core layer do not detect route changes at the access layer when receiving gratuitous ARP packets because they connect to access switches through trunk links.

Dual NICs in Load Balancing Mode

The two NICs in load balancing mode have the same MAC address (such as MAC2 in [Figure 3-13](#)). Both NICs can transmit and receive data. To shield the flapping of the MAC address between ports of switches, stack the access switches and bundle the links on ports of the active and standby NICs.

[Figure 3-13](#) shows the change in data transmission routes. Data is transmitted in the green routes using both NICs. If an NIC fails, data transmission routes are changed from green ones to purple ones.

Figure 3-13 Change in data transmission routes in load balancing NICs



Switches at the combined core layer do not detect route changes at the access layer because they are connected to access switches through trunk links. Therefore, data is still sent to the

access switch on the left, forwarded to the switch on the right through the stacking link, and then forwarded to the server.

3.3.4 Security Planning

VLANs are created to separate services in the service area. To ensure network security, use VLANs to separate the web, application, and database servers of the same service.

3.4 Storage Area Networking Planning

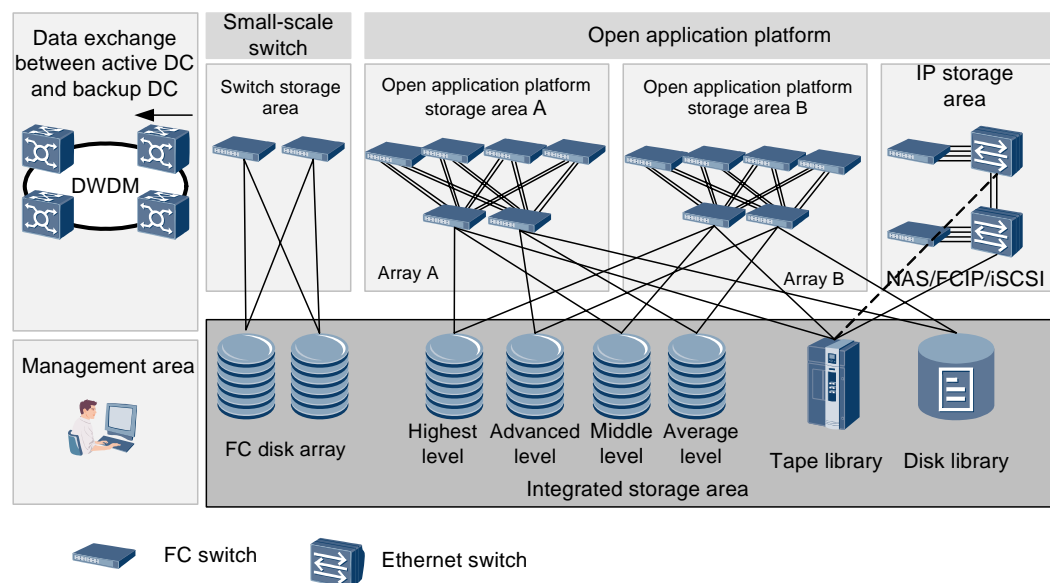
3.4.1 Physical Networking Planning

The storage area covers the IP storage network and the SAN storage network.

The IP storage network transmits traffic for services that are saved in the network attached storage (NAS). The NAS transmits:

- Data traffic generated between a specified application server and the NAS
- Large amounts of network traffic generated for virtualized services

Figure 3-14 Storage area networking



Basic planning for the storage area is as follows:

- Small-scale switch area: This area is an isolated area.
- Open application platform:
 - The open application platform has two arrays, and each array has two core switches to ensure availability.
 - Edge devices and core devices are connected to one another by multiple links to prevent traffic overload.
- Integrated storage area: Devices are classified based on service class in this area.

- IP storage area:
The IP storage area is separated from other areas. Devices are deployed in this area to compress traffic transmitted in the Entire Fiber Channel Frame over IP (FCIP) channel and accelerate data transmission. Data is synchronized and saved through an IP/Multiprotocol Label Switching (MPLS) network. Implementation of virtualization speeds up data transmission between servers and storage devices.
- Management area: Operators allocate and manage storage network and storage resources in this area.
- Disaster backup in the same city:
The active DC and the disaster recovery center in the same city are connected through the dense wavelength division multiplexing (DWDM) network.

Use the following configuration to implement real-time or quasi real-time data exchange between the active DC and backup DC:

- Use the carrier's MPLS VPN or virtual leased line based on the virtual private LAN service (VPLS) to transmit data traffic between servers on the IP storage network in the active DC and backup DC.
- Use bare optical fibers or a DWDM network to transmit data between SAN storage networks in the active DC and backup DC. This implements quasi real-time data transmission at a high speed and a short delay.

Virtualization increases data exchange between servers and storage devices, so switches must access the NAS storage network through a 10G link.

3.4.2 Reliability Planning

The IP storage network uses loop-free networking with cluster and stacking to enhance reliability. For details on loop-free networking with cluster and stacking, see section [3.3 "Server Area Networking Planning."](#)

3.4.3 Security Planning

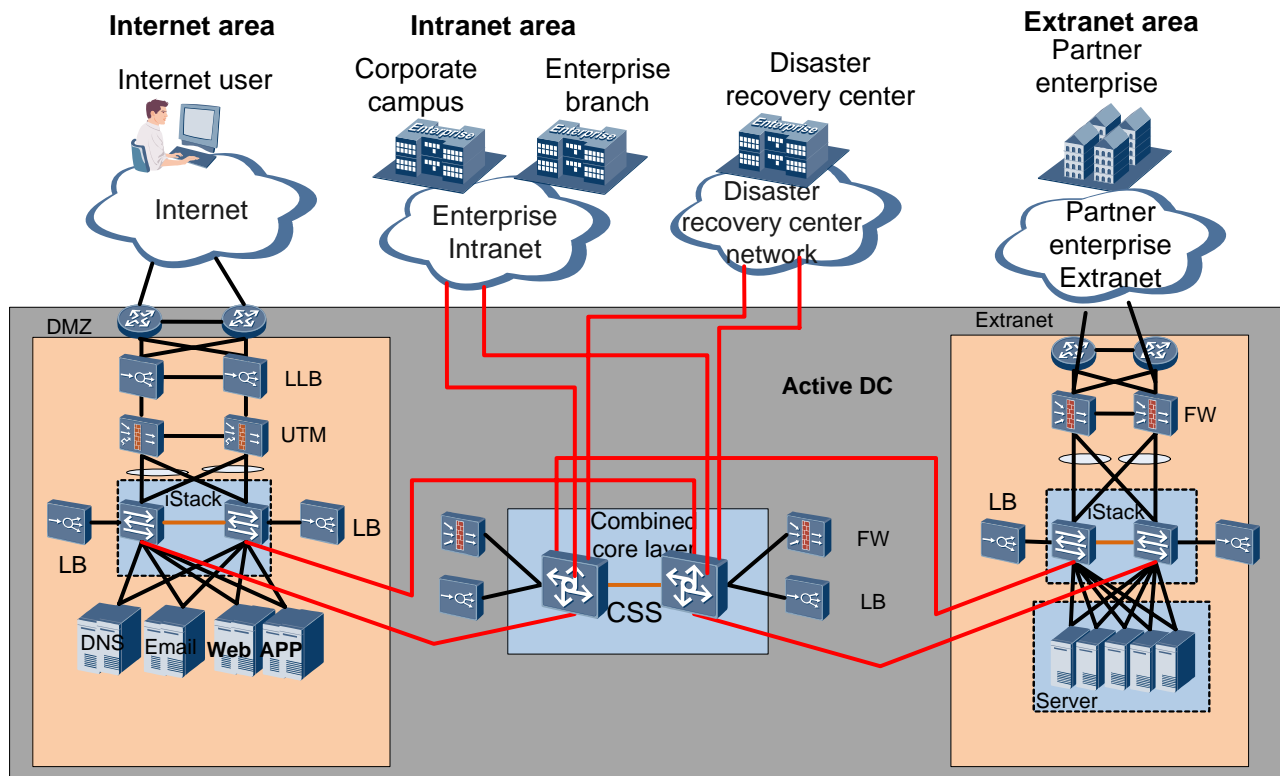
The SAN storage areas are isolated by the specialized technology.

To restrict network access, the IP storage network is divided into separate areas through VLAN or VPN technology.

3.5 Interconnection Area Networking Planning

3.5.1 Physical Networking Planning

Figure 3-15 Networking in the interconnection area



The interconnection area is divided into the following connection areas based on access modes and services:

- Intranet area
Intranet users access the DC through the WAN or the LAN.
- Internet area
External users access the DC through the Internet.
- Extranet area
Extranet users access the DC through the WAN or the LAN.



NOTE

You can assign an isolated area for the VPN users in the Internet area.

3.5.2 Internet Area

Figure 3-16 Networking in the Internet area

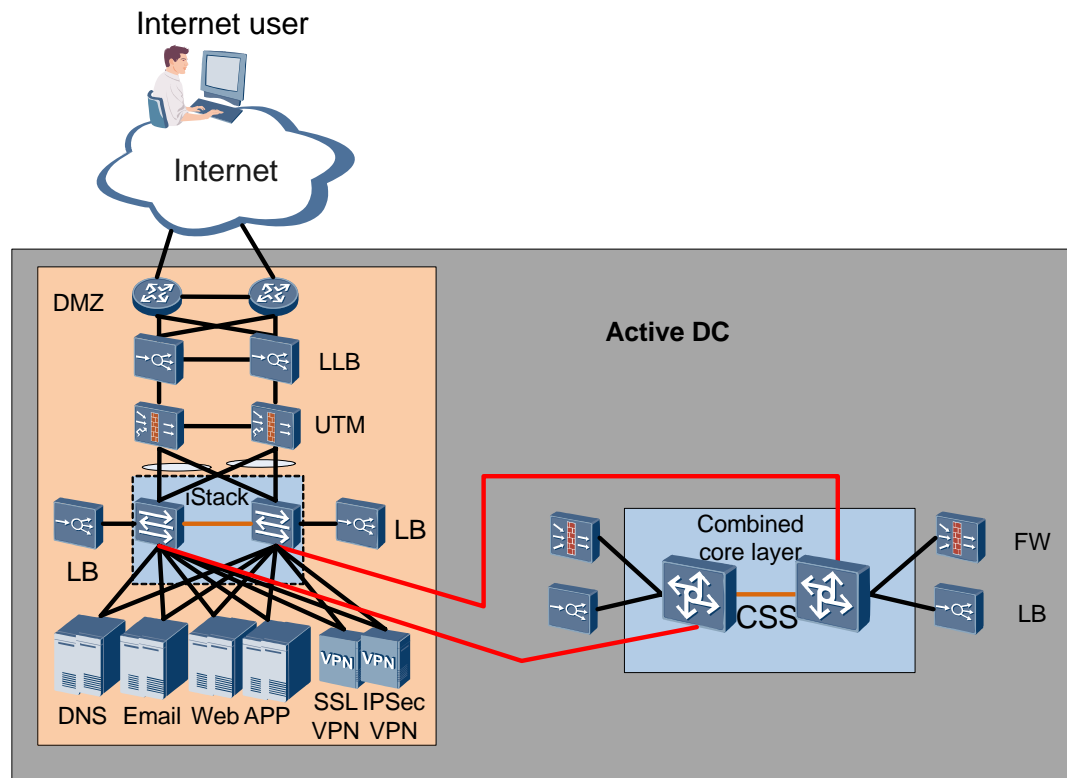


Figure 3-16 shows the Internet area devices, such as routers, link load balancers, and unified threat management (UTM) devices. The UTM devices must provide firewall and intrusion prevention system (IPS) functions.

- Load balancers are used to respond to requests from different egress routers leased by two carriers. No load balancers are required when there is only one egress router.
- The IPS detects malicious codes, attacking actions, and distributed denial of service (DDoS) attacks that are mixed in the application data stream, and takes response in real time.
- The firewall is deployed at the network layer to filter invalid traffic and protect intranet resources against attacks from the Internet.

The firewall and the IPS are important network devices, which are located at the network egress. The location and functions of the firewall and the IPS require that they should provide high reliability.

To ensure Internet area reliability, deploy devices in pairs, such as routers, link load balancers, and UTM devices (including firewalls and the IPS).

The VPN access area must provide the Internet Protocol Security (IPSec) VPN and the Secure Sockets Layer (SSL) VPN functions for secure access.

- The IPSec VPN provides site-to-site access mode.
- The SSL VPN provides client-to-site access mode.

The IPSec VPN gateway and SSL VPN gateway can be deployed independently, or connected to the network using the UTM devices.

3.5.3 Extranet Area

Figure 3-17 Networking in the extranet area

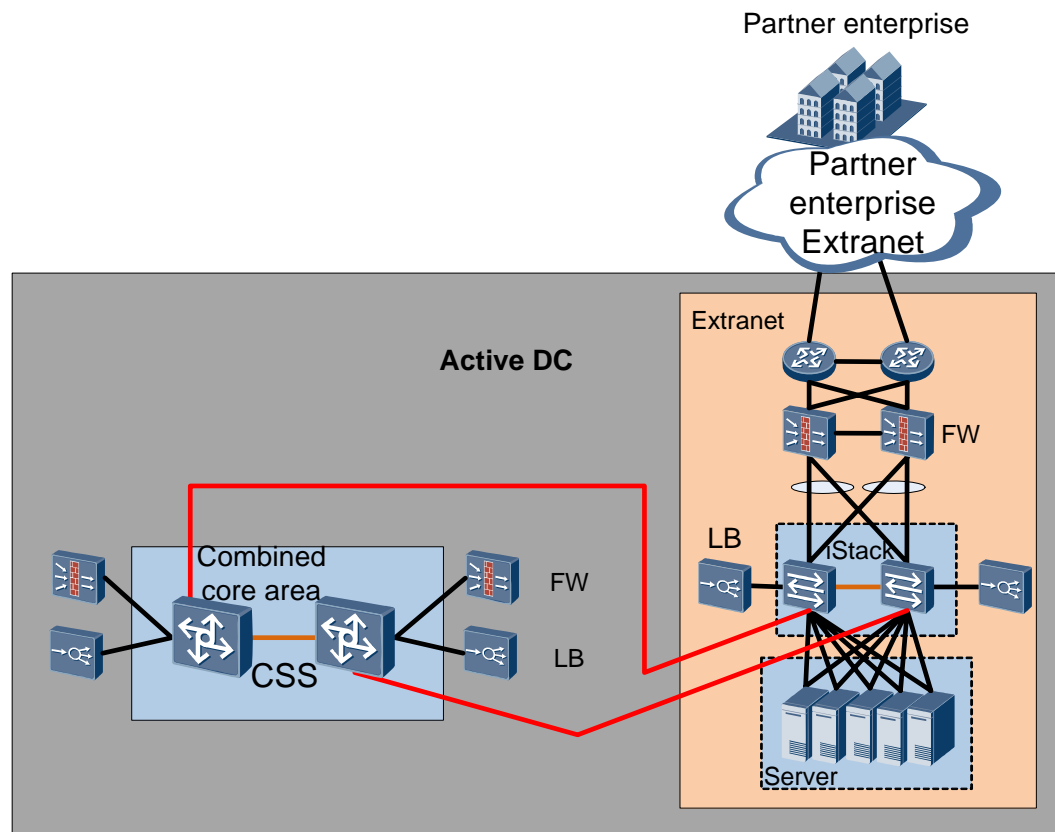


Figure 3-17 shows the networking in the extranet connection area.

Extranet users can access the extranet connection area. This area is an unreliable area, similar to the demilitarized zone (DMZ), and cannot be connected to the inner DC. Extranet users can access only the extranet connection and DMZ areas. Authority control on the Intranet must be strict.

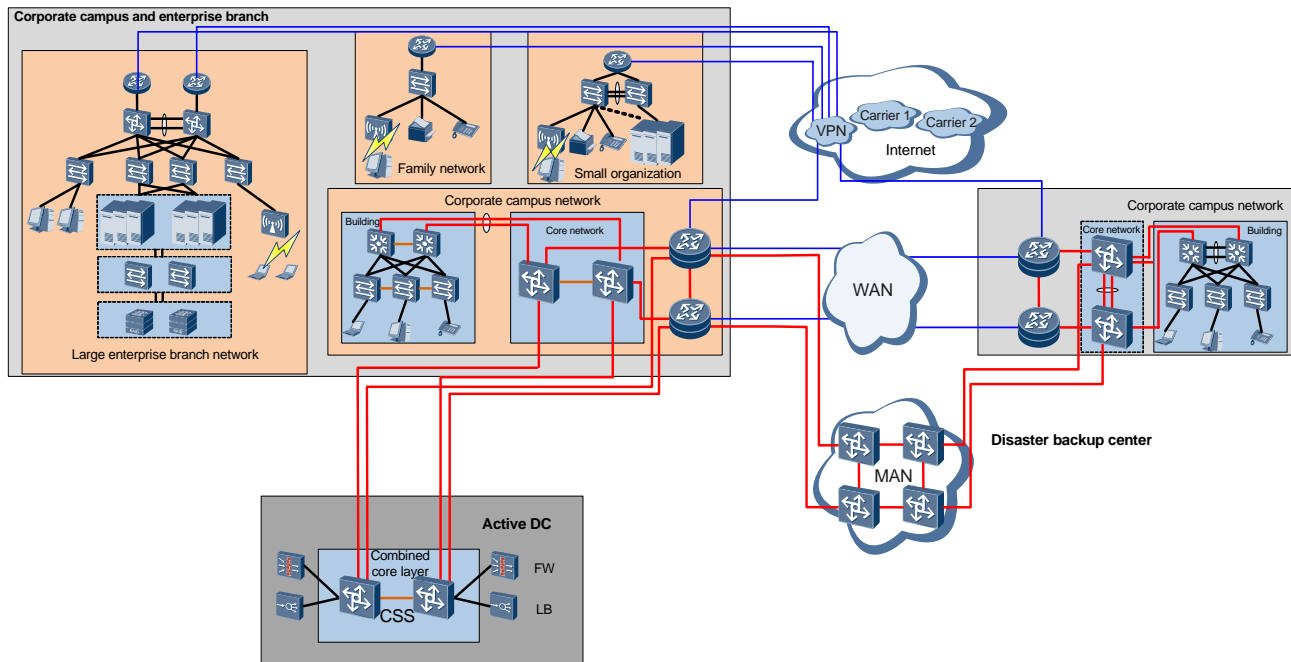
Security is provided through service isolation and the firewall:

- Service isolation: The extranet connection area provides services for extranet users like the DMZ. The extranet users cannot access the Intranet but they can access the DMZ under strict authority control.
- Firewall: The firewall is deployed at the network layer to protect Intranet resources and control access rights by shielding the network topology with the network address translation (NAT) technology.

3.5.4 Intranet Area

Intranet users access the DC through the WAN or the LAN.

Figure 3-18 Networking in the intranet area



This area uses dual-homed routes and redundancy backup of routes and devices.

Network connection reliability between branches of an enterprise is ensured through backup of multiple egress links, backup of routes, and load balancing. QoS needs to be configured on WAN link to guarantee quality of links and services.

Independent access devices and two backup devices are required to ensure device reliability.

The intranet is a safe area with low security risks which are mainly caused by intranet users who access or save data without authorization. Data access between the enterprise branch networks is restricted based on users' actual requirements.

3.6 Management Area Networking Planning

3.6.1 Physical Networking Planning

Overall requirements are:

- Out-of-band management
- Authorization-based access
- Security auditing

Figure 3-19 Management area networking

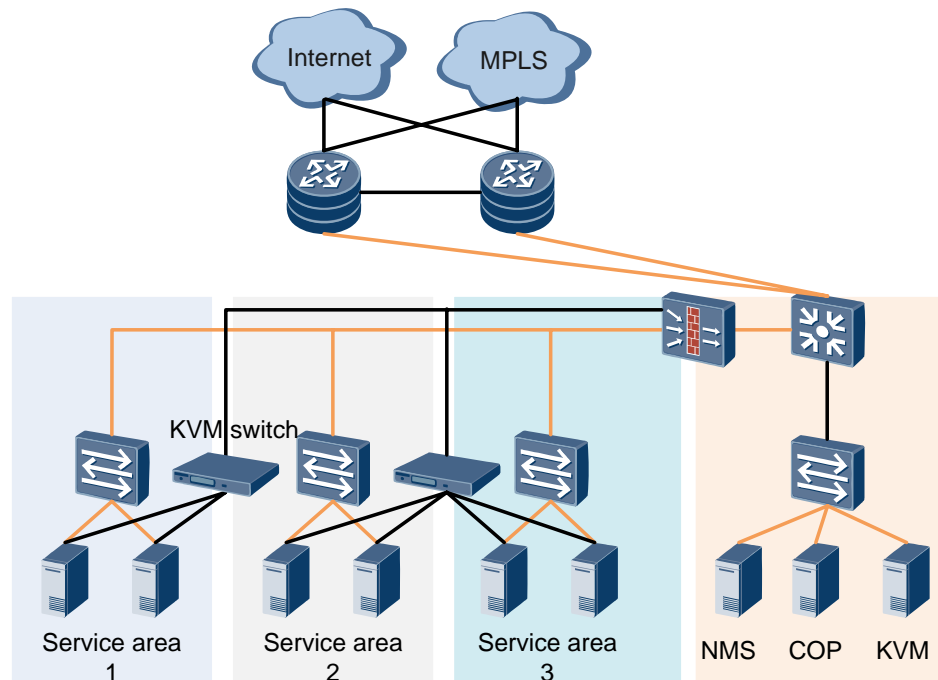
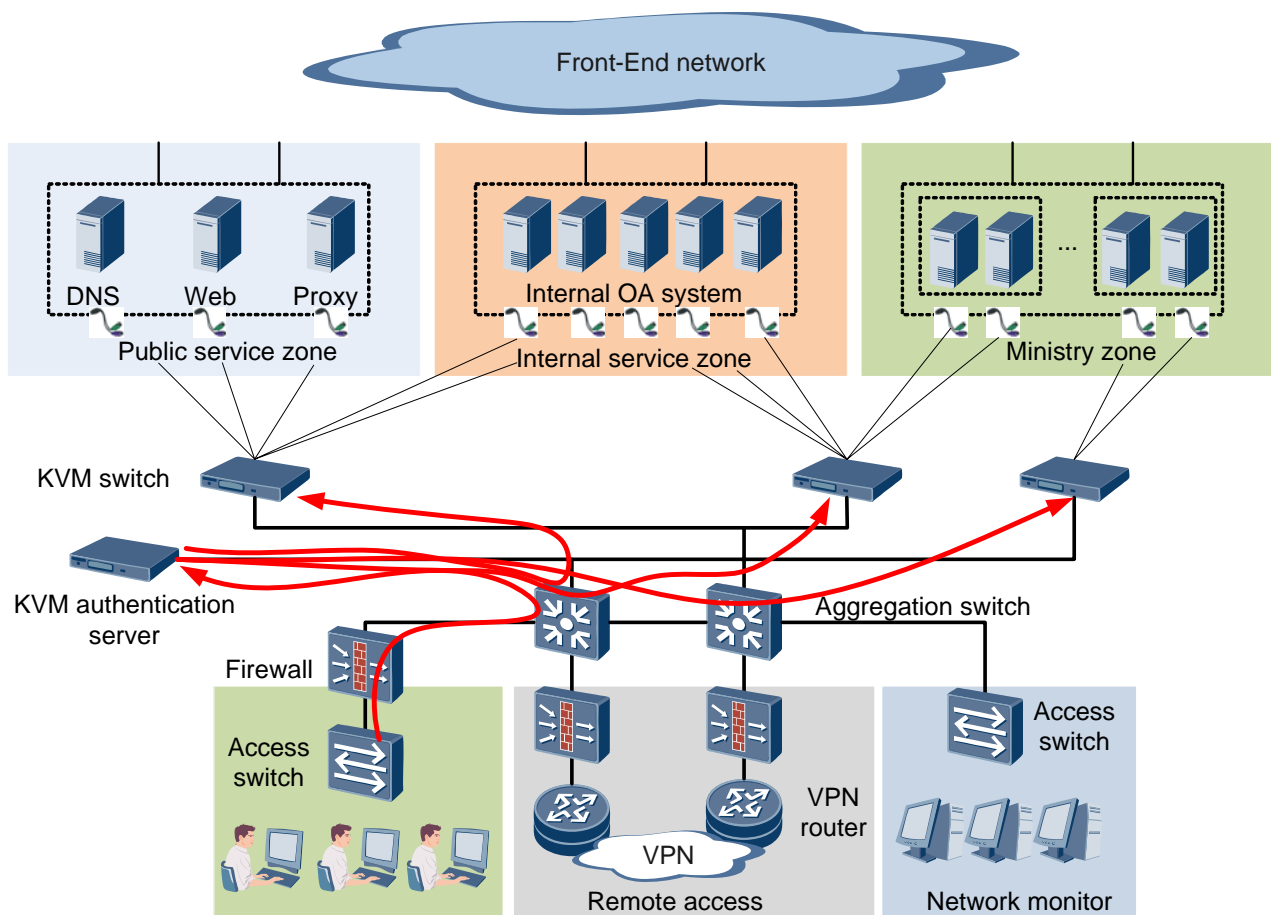


Figure 3-19 shows the networking in the management area. The management network connects all devices by the management interfaces and the KVM switches, and provides functions such as network management, data collection, and real-time surveillance.

Only administrators can access the management network that connects the inner DC using isolation measures such as VPNs and firewalls. Administrators are granted rights to access specified network devices.

Figure 3-20 KVM management network



Network management: This module manages network devices such as switches, routers, and firewalls in the aspects of the topology, configuration, asset, fault, performance, event, traffic, and report.

There are two areas of network management:

- Traffic management: This module provides functions such as traffic monitoring, traffic threshold setting, protocol analysis, and Web access behaviors audit. It works with the NetFlow analyzer to implement more refined and convenient traffic analysis.
- Application management: This module monitors the website and manages systems and upper-level applications such as the database, mail server, Web server, application server, operating system, and website surveillance.

3.7 VLAN Planning

3.7.1 VLAN Overview

Devices on a LAN are logically grouped into segments, regardless of their physical locations. VLANs isolate broadcast domains on a LAN, reduce broadcast storms, and enhance information security. As the network expands, a fault on the local network affects the entire

network. The VLAN technology can limit the network faults within a VLAN, and enhances the network robustness.

3.7.2 Principles

Observe the following principles when configuring VLANs:

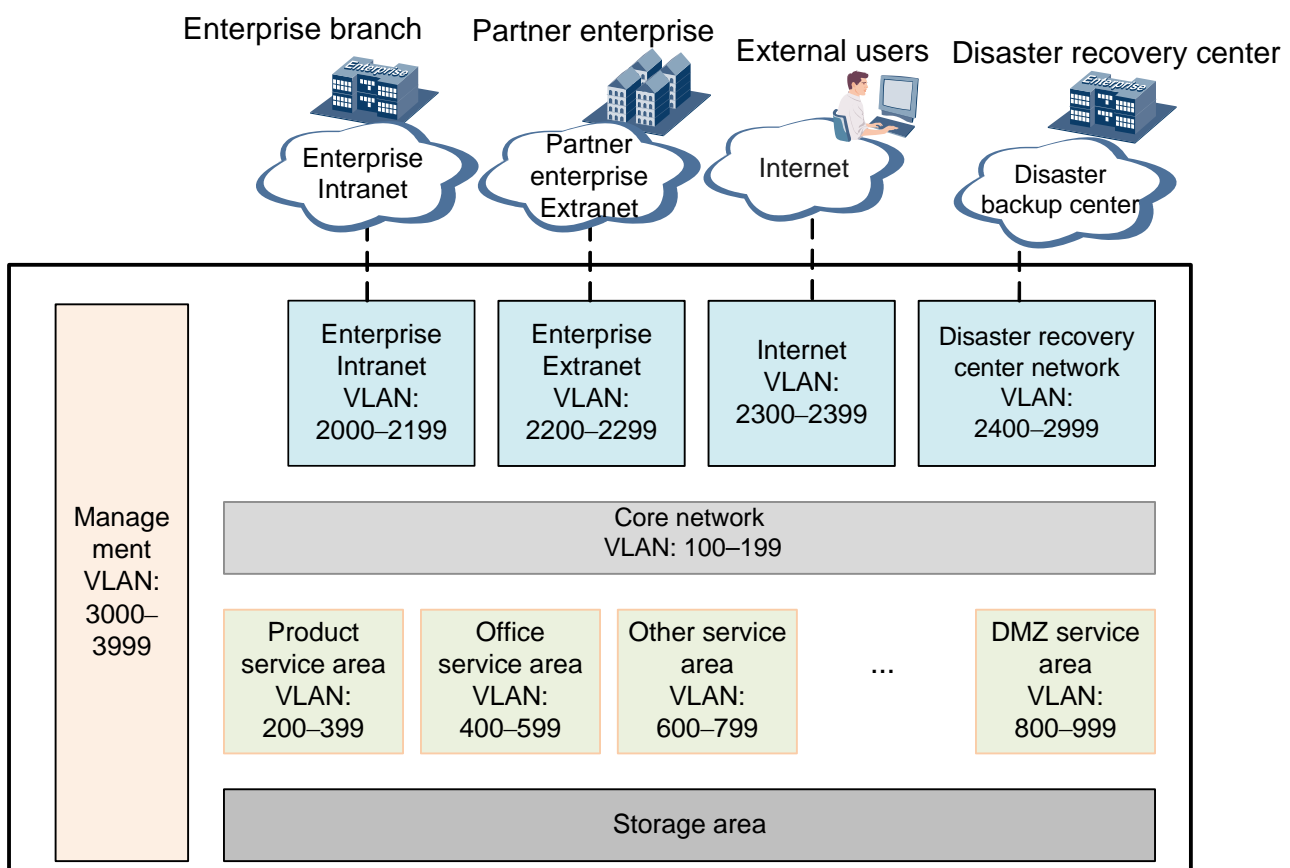
- Differentiate service VLAN, management VLAN, and interconnection VLAN.
- Add interfaces to different VLANs based on service areas.
- Add interfaces to different VLANs based on service types for the same service (such as the Web, application, and database).
- Distribute each VLAN consecutively to properly use VLAN resources.
- Reserve some VLANs for further expansion.

3.7.3 Recommendation

Configure VLAN ranges based on different areas as shown in [Figure 3-21](#).

- Core area: 100–199
- Server area: 200–999, reserved VLANs: 1000–1999
- Access network: 2000–2999
- Management network: 3000–3999

Figure 3-21 VLAN planning



3.8 IP Planning

A few devices in the Internet connection area use public IP addresses, but devices in the intranet use private IP address. IP addresses in the intranet are easy to manage because private IP address space is large, for example, 10.0.0.0 is a class-A address.

3.8.1 IP Address Planning

Plan so that the system IP address will be:

- Unique
Hosts on an IP network must use different IP addresses. Assign different IP addresses to hosts even if the MPLS/VPN technology supporting IP address overlapping is used.
- Consecutive IP addresses
Consecutive IP addresses facilitate routing aggregation on a hierarchical network, which greatly reduces the number of routing entries and improves route calculation efficiency.
- Scalable
IP addresses need to be reserved at each layer. When the network expands, IP addresses continuity is ensured.
- Meaningful
If the IP addresses are planned properly, you can identify the device that corresponds to an IP address by the IP address.

3.8.2 DNS Planning

DNS Server Roles

A domain name system (DNS) server plays the following roles in the DNS system:

- Master server
The master server manages the DNS system, and is used to add, modify, or delete a domain name. The domain information that is changed on a master server is synchronized to a slave server. One master server is deployed in the DNS system.
- Slave server
The slave server obtains the domain name information from the master server, and forms a server cluster by connecting multiple servers with hardware-based load balancers to provide DNS services. Two slave servers are deployed in the DNS system.
- Cache server
The cache server is deployed on the slave server to cache results of intranet users' DNS requests and to speed up network access.

IP Address of the DNS Server

IP addresses are allocated as follows:

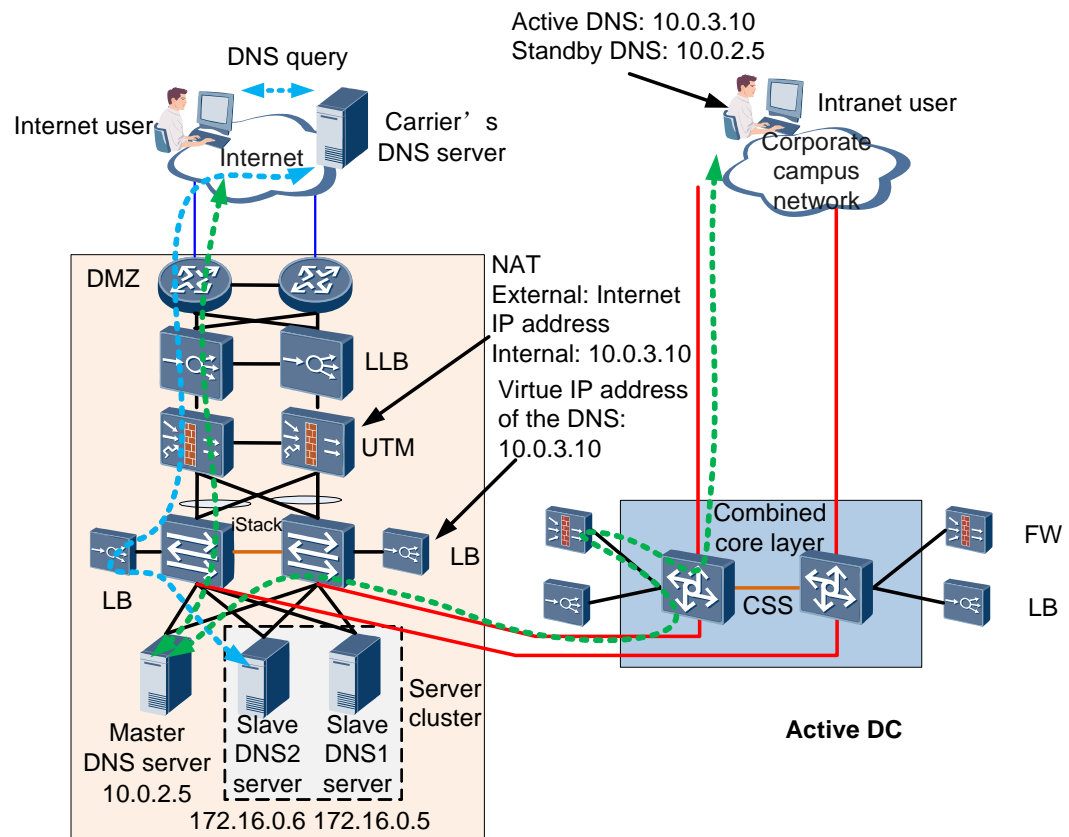
- The master server uses a private IP address.
- The slave server is allocated a private IP address, and has a virtual private address on the load balancer.

The Internet domain names and IP addresses are deployed in the following ways:

- Configure NAT mapping on the firewall to convert the virtue IP address of the slave server into a public IP address for Internet users to use for accessing the intranet.
- Provide services for Internet users using intelligent DNSs on load balancers.

Providing DNS Services for Internet Users Using the Slave Server

Figure 3-22 DNS deployment in the DC



The blue dotted line marked in [Figure 3-22](#) shows how the slave server is used to provide DNS services for Internet users.

The slave servers DNS1 and DNS2 use virtue IP addresses on the load balancer to function as master DNS servers for Internet users and slave DNS servers for intranet users.

The master DNS, slave DNS1, and slave DNS2 servers are all deployed in the DMZ area.

The process to handle DNS requests with reliable design is as follows:

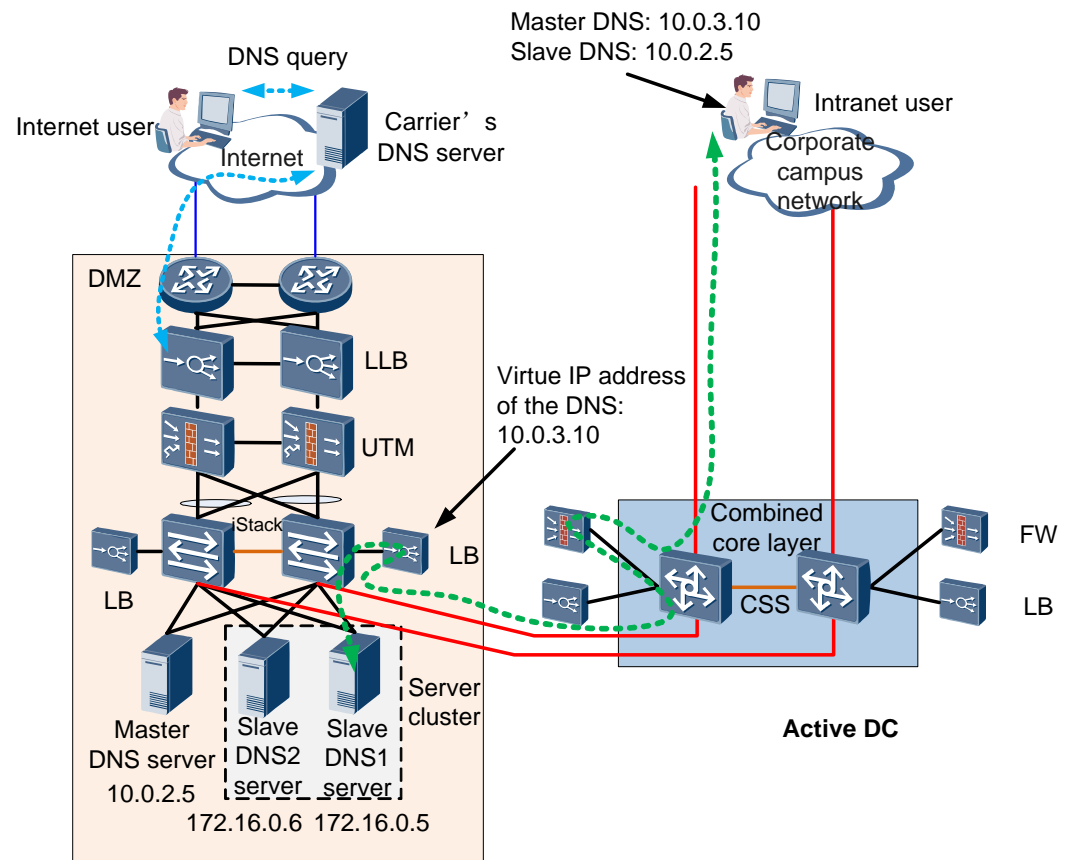
Intranet users send DNS requests to the master DNS server that communicates with the carrier's DNS server to resolve Internet domain names. If the master DNS server is faulty, the slave DNS servers provide services.

Internet users send DNS requests to the carrier's DNS server to resolve the enterprise domain name, such as Huawei.com, and relay the further resolution results, such as www.huawei.com, to the enterprise DNS server. The DNS requests are evenly distributed between slave DNS1 and slave DNS2 servers. If slave DNS 1 server is faulty, all DNS requests are sent to slave DNS2 server. If both slave DNS servers are faulty, the master DNS server provides services.

Providing Services for Internet Users Using the Intelligent DNS Server

Figure 3-23 shows how the intelligent server is used to provide DNS services for Internet users.

Figure 3-23 Intelligent DNS deployment



The Internet users send requests (such as www.huawei.com) to the carrier's DNS server to query the domain name of Huawei. The carrier's DNS server identifies the information (huawei.com), and sends the request to the DNS server in Huawei DC to resolve the domain name. The blue dotted line displays this process.

The intelligent DNS server in the link load balancer receives the request, and finishes the DNS resolution.

The intelligent DNS server recognizes user sources and resolves domain names to different IP addresses. The DNS policy resolution server resolves the domain name to the related Netcom IP address for a China Netcom user and the related Telecom IP address for a China Telecom user.

Meanwhile, the intelligent DNS server monitors carrier link quality. If a carrier's link is interrupted, the intelligent DNS server returns another carrier's IP address to ensure service continuity.

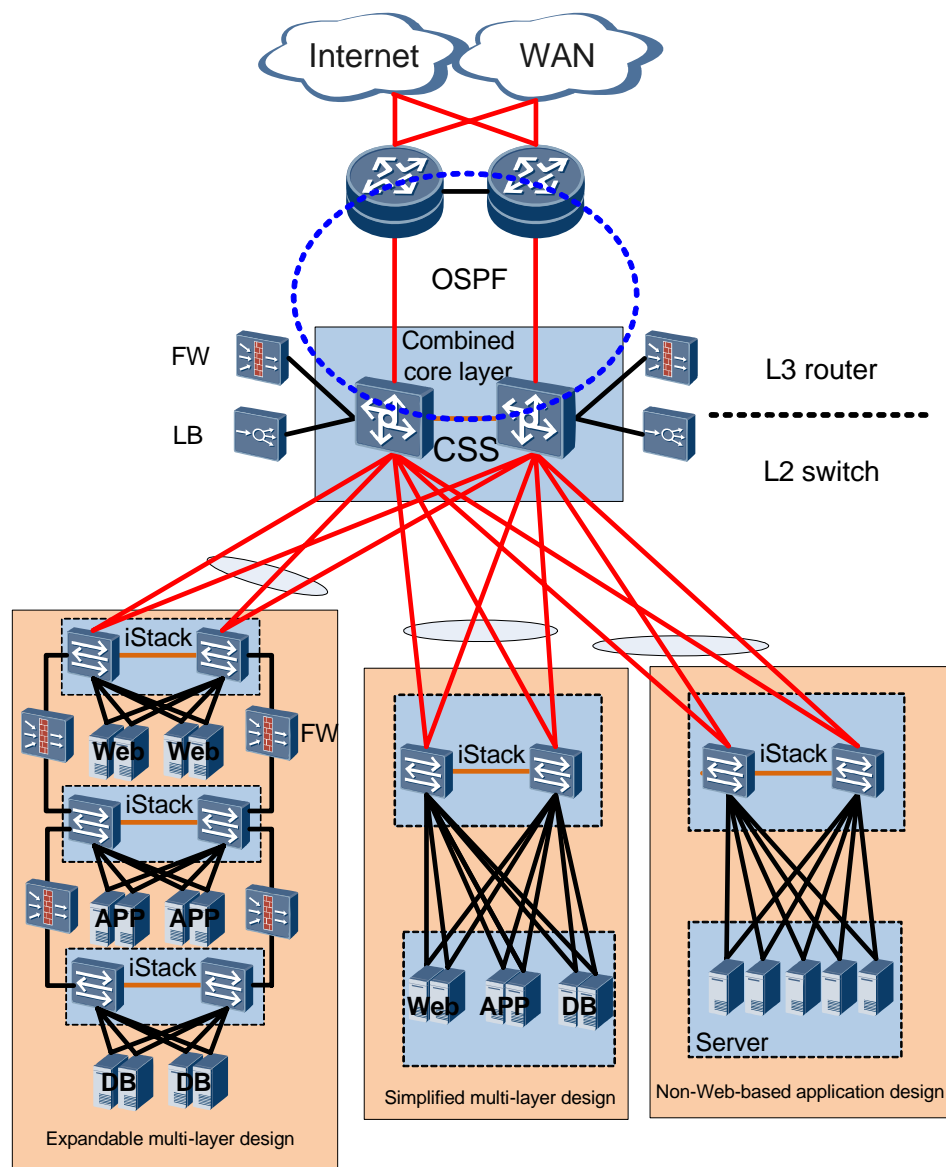
3.9 Route Planning

3.9.1 Routing Overview

Huawei recommends that the boundary between routing and switching be located on the combined core layer switches, as shown in [Figure 3-24](#).

- Layer 2 switching is used at the layer below the combined core layer.
- Layer 3 routing is used at the layer above the combined core layer.

Figure 3-24 Boundary between routing and switching



This design has the following advantages:

- Simple route configuration

Routes need to be configured only on two combined core layer switches. Access switches perform only Layer 2 switching, simplifying the configuration. Users can use the automatic configuration functions of access switches to reduce the maintenance workload.

- Scalability

You can easily increase the number of servers on a core/aggregation switch.

A new service server can be deployed in any rack. The IP address of the new server is contiguous with the IP address of the original service system.

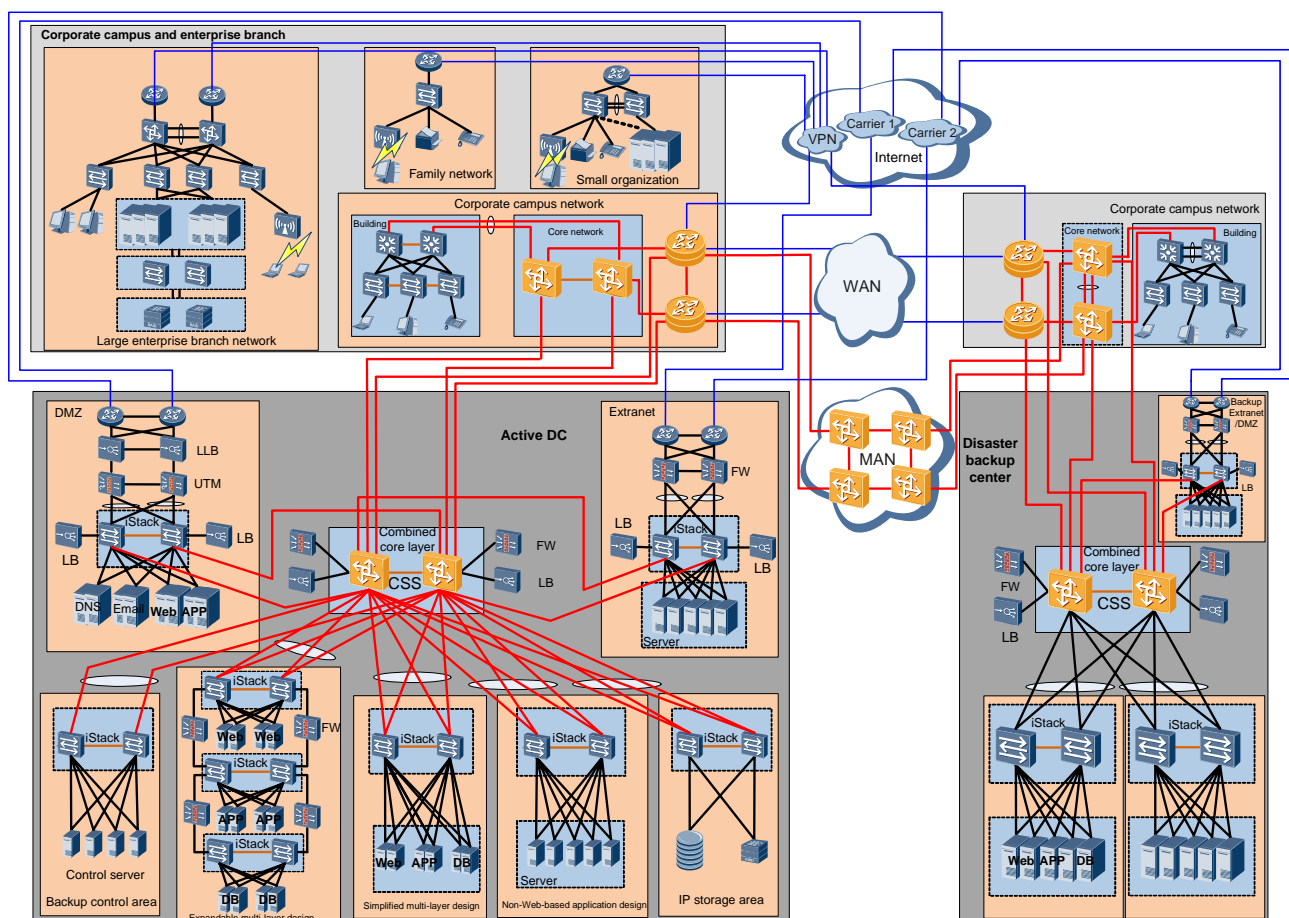
When the position of a server changes due to a service change, the carrier does not need to reconfigure the servers and the network, and the servers can be used immediately after being installed in the new position. A large Layer 2 network is needed when the next generation virtual servers are used to move servers without interrupting services.

3.9.2 IGP Design

To manage and maintain the network conveniently inside the data center, the OSPF dynamic routing protocol is recommended to ensure network stability and fast convergence of routes.

As shown in [Figure 3-25](#), the yellow-colored devices are core switches located in the backbone area, Area 0.

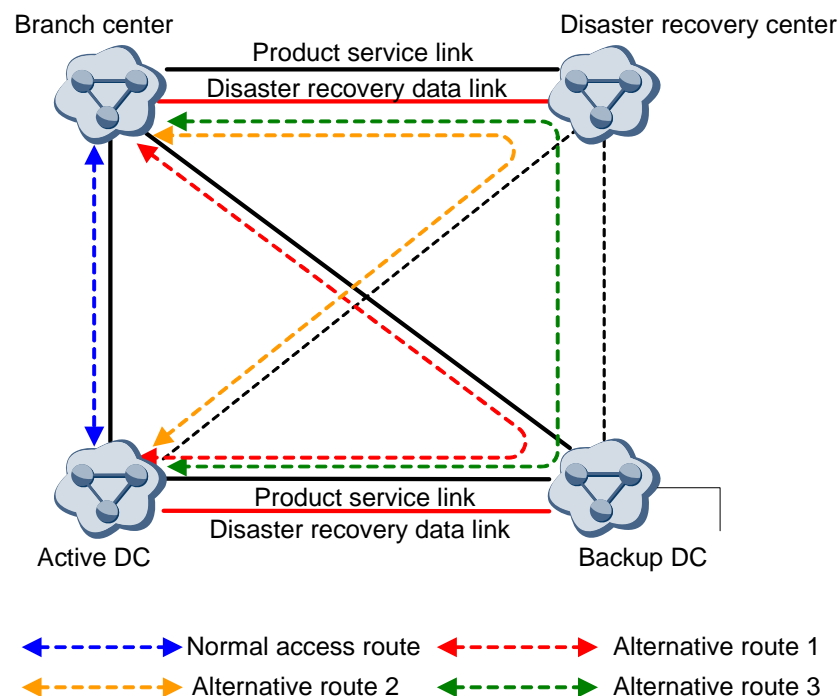
Figure 3-25 Router planning for DCs



3.9.3 BGP Design

The external Border Gateway Protocol (EBGP) is established between the branch DC and disaster recovery center using network access routers, which advertise routes to both centers. Figure 3-26 shows the network topology among the active DC, backup DC, branch DC, and disaster recovery center.

Figure 3-26 Active and standby path planning for DCs



As shown in Figure 3-26, the active DC has four paths to reach the branch DC. Priorities of four paths are as follows:

- Highest priority (normal access route): The active DC is connected to the branch DC directly.
- Second highest priority (alternative route 1): The active DC reaches the branch DC through the backup DC.
- Third highest priority (alternative route 2): The active DC reaches the branch DC through the disaster recovery center.
- Lowest priority (alternative route 3): The active DC reaches the branch DC through the backup DC and the disaster recovery center.

The priorities of the links are determined by the EBGP AS-Path and multi-exit discriminator (MED) attributes.

3.10 VPN and the Service Area Planning

3.10.1 VPN Overview

The VPN technology is used to isolate services, control network access, and implement secure isolation.

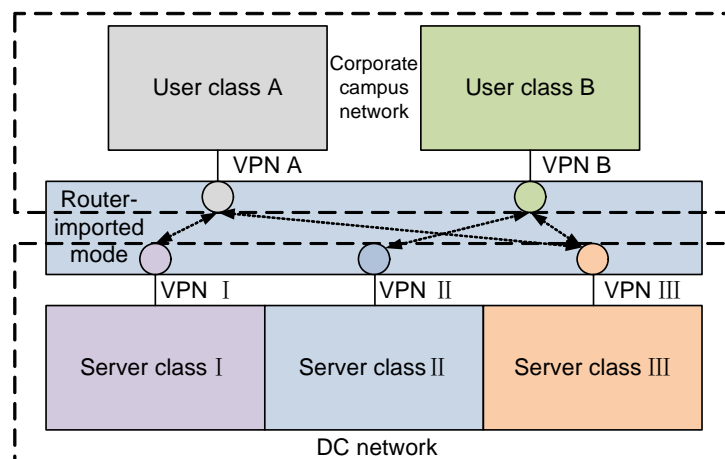
As a Layer 3 VPN, BGP/MPLS IP VPN uses the BGP to advertise VPN routes and uses MPLS to forward VPN packets on backbone networks of service providers (SPs). IP in IP VPN refers to the IP packets transmitted on VPNs.

3.10.2 Intranet VPN Service Isolation

As shown in [Figure 3-27](#), users and servers are separated and grouped into different VPNs. By default, routers for user A, user B, server I, server II, and server III are isolated so that these users and servers cannot communicate with each other.

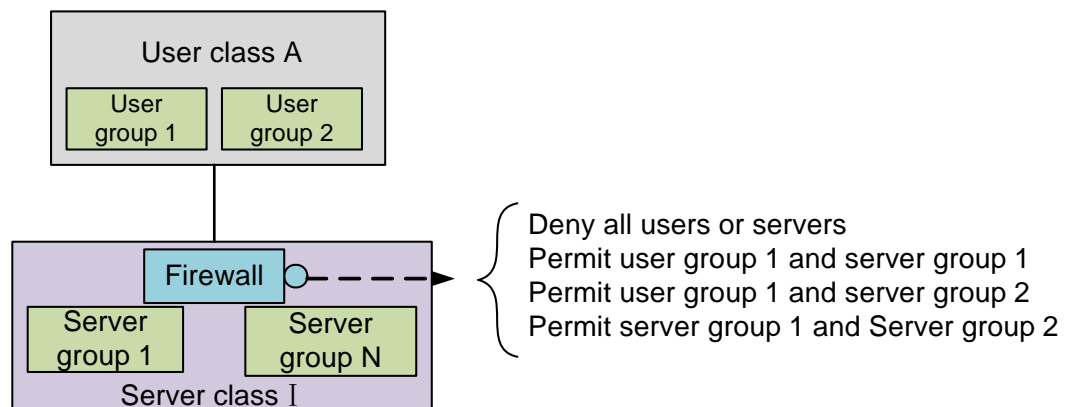
User VPN and server VPN can import routes from each other based on the user and server control policy. The imported routes among user A (VPN A), VPN I, and VPN III allow user A to access the VPN I, and VPN III servers.

Figure 3-27 Server isolation plan based on routes



As shown in [Figure 3-27](#), firewalls are used to accurately control the rights of server groups. The security policy is configured based on the table for rights of the user groups and server groups. By default, the firewalls are disabled. Users can access the server only after a security policy is configured to enable the firewalls.

Figure 3-28 Server isolation plan based on firewalls



3.11 QoS Planning

3.11.1 QoS Overview

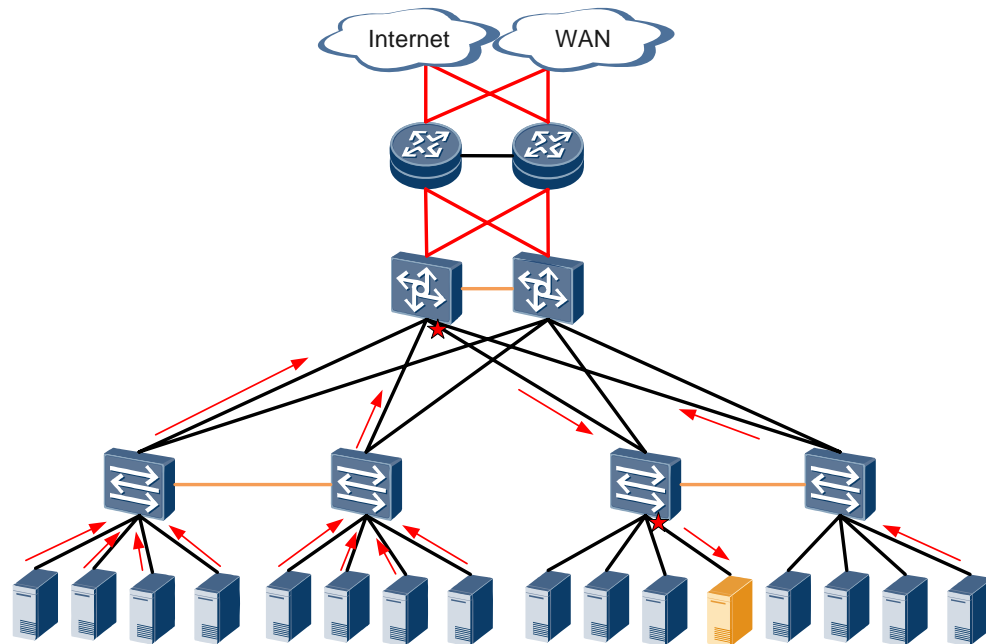
The DC planning guarantees peak-traffic services, which requires no QoS processing. The QoS planning, however, is needed in collaborative computing and Multi-Tenancy applications.

The multi-tenant applications are used to manage bandwidth and are not in the initial version. QoS planning for multi-tenancy applications can be complemented and optimized in the subsequent operations.

3.11.2 QoS Planning Concerning Collaborative Computing

Collaborative computing is used in when complicated calculations are involved. Examples of this are the computing involved for search engines, petroleum exploration, and meteorology. In collaborative computing, multiple servers may send calculation results to one server at the same time, which brings a traffic burst, which could result in data congestion on an outbound interface and packet loss.

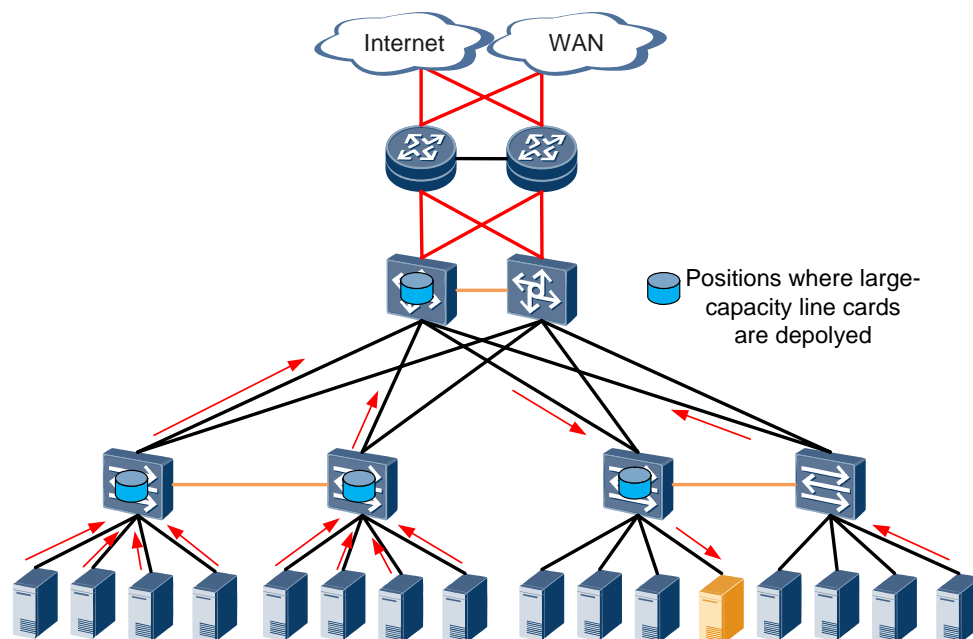
Figure 3-29 Congestion on an outbound interface when multiple servers send data to one server



As shown in [Figure 3-29](#), servers send data to the yellow-colored server and congestion occurs in the starred node. Packets are lost if queues are not sufficient in the nodes that forward data.

To solve the problem, install large-capacity line cards on the EOR switch and the core switch to cache burst data and prevent packet loss.

Figure 3-30 Large-capacity line cards on the EOR switch and the core switch to prevent packet loss



4 Desktop-Cloud Network Solution

4.1 Desktop-Cloud Service Overview

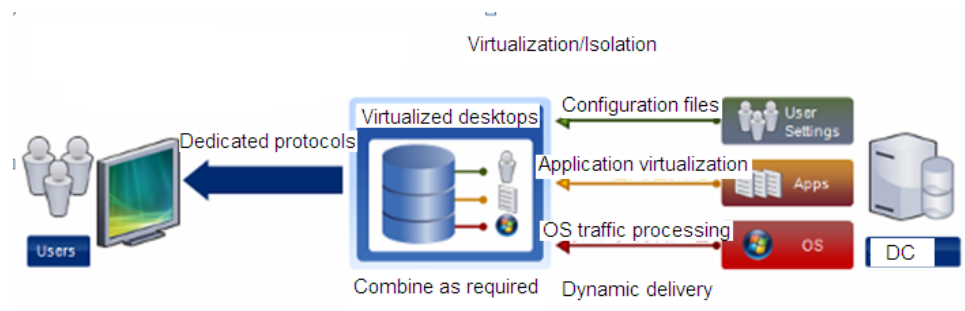
Desktop-cloud is an application model for cloud computing. Cloud computing is a new method to utilize Internet resources, providing heterogeneous and autonomous services for users. Cloud-computing resources are dynamically scalable, virtual, and provided over the Internet.

Desktop cloud is a cloud service and has the following characteristics:

- Displays services in the desktop mode
- Manages resources flexibly
- Provides services over the Internet

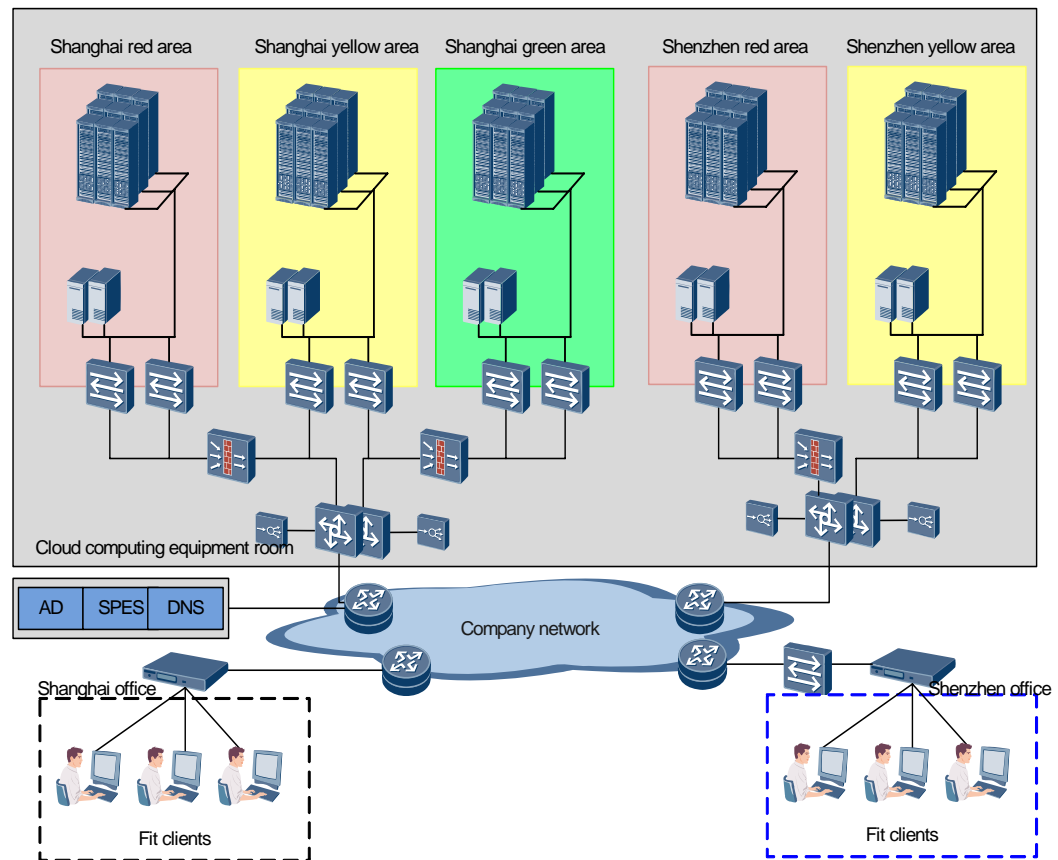
If you want to access the personal desktop and various applications hosted on the server side, you need only to deploy an appropriate client or network device and access the server through dedicated applications or browsers. The user experience is the same as if you logged in to your own personal computer. Desktop cloud separates the PC computing and storage resources from the physical PC desktop device to provide desktop services. Computing and storage resources are located on the central server to replace those located on the client. The computing and storage resources on the central server are scalable and shared, which are allocated and delivered on demand so that resource usage is increased and overall ownership cost is reduced.

Figure 4-1 Desktop cloud service



Desktop cloud connects clients (such as offices and reception areas) to the DC through the intranet or Internet. The DC adopts desktop virtualization technologies to create dozens of virtual desktops based on a physical server.

Figure 4-2 Networking for the desktop cloud DC and clients

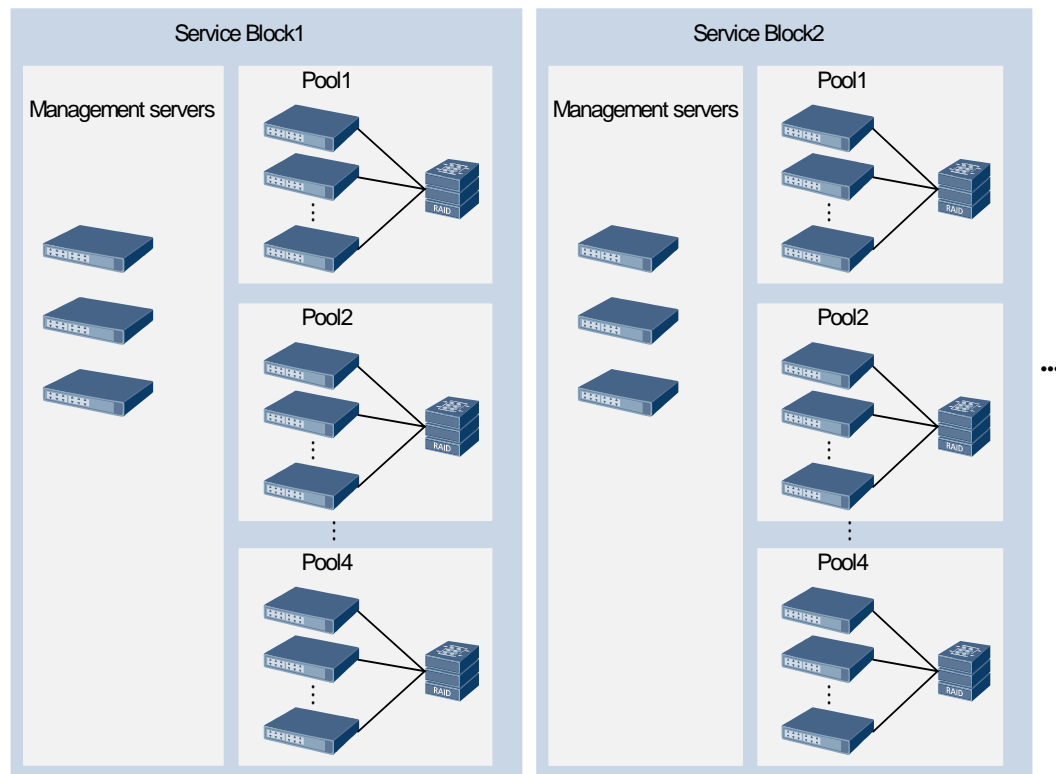


The desktop cloud DC is divided into areas based on services, security requirements, and scale limit in virtualization management.

Basic concepts related to the desktop cloud solution are as follows:

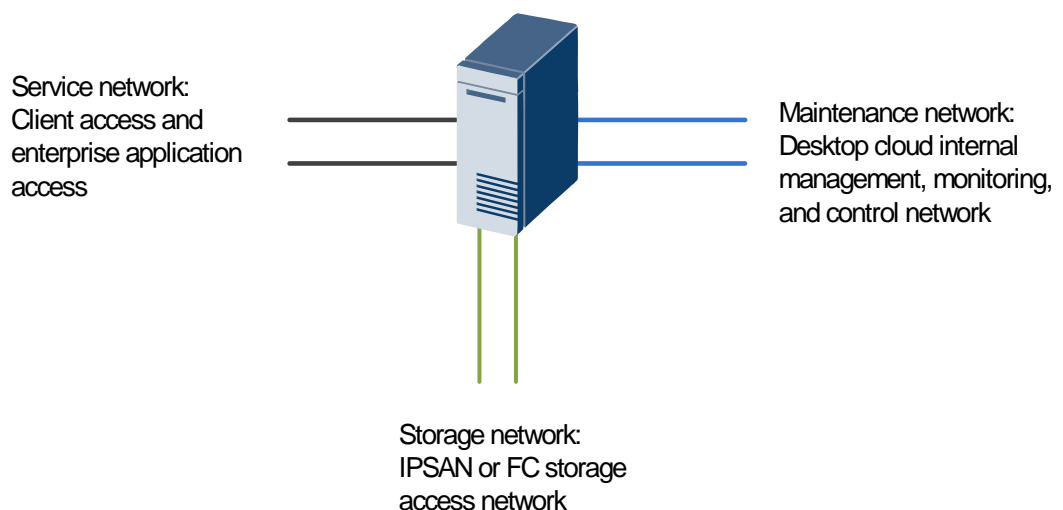
- Management server: Components such as WI, DDC, AD, and license server. Not the desktop virtual server.
- Service block: Extension unit in the solution. Each service block supports 2000 concurrent desktop users. A service block contains one to four management servers and pools.
- Pool: A pool consists of 1 to 20 servers, and one set of storage equipment (one controller subrack and several extension subracks). Each pool supports 400 to 500 concurrent desktop users.

Figure 4-3 Logic concepts of desktop cloud



As shown in [Figure 4-3](#), the blade servers serve as desktop servers. Blade servers are virtualized to create management servers and desktop virtual servers.

Figure 4-4 Server network plane



By default, a blade server is configured with service, maintenance, and storage network planes. Each plane is configured in 1+1 redundancy mode. These three network planes are independent, enhancing the network stability and availability.

4.2 Desktop Cloud Network Structure

Figure 4-5 shows a typical desktop cloud networking.

Figure 4-5 Networking for the desktop cloud DC and fit clients

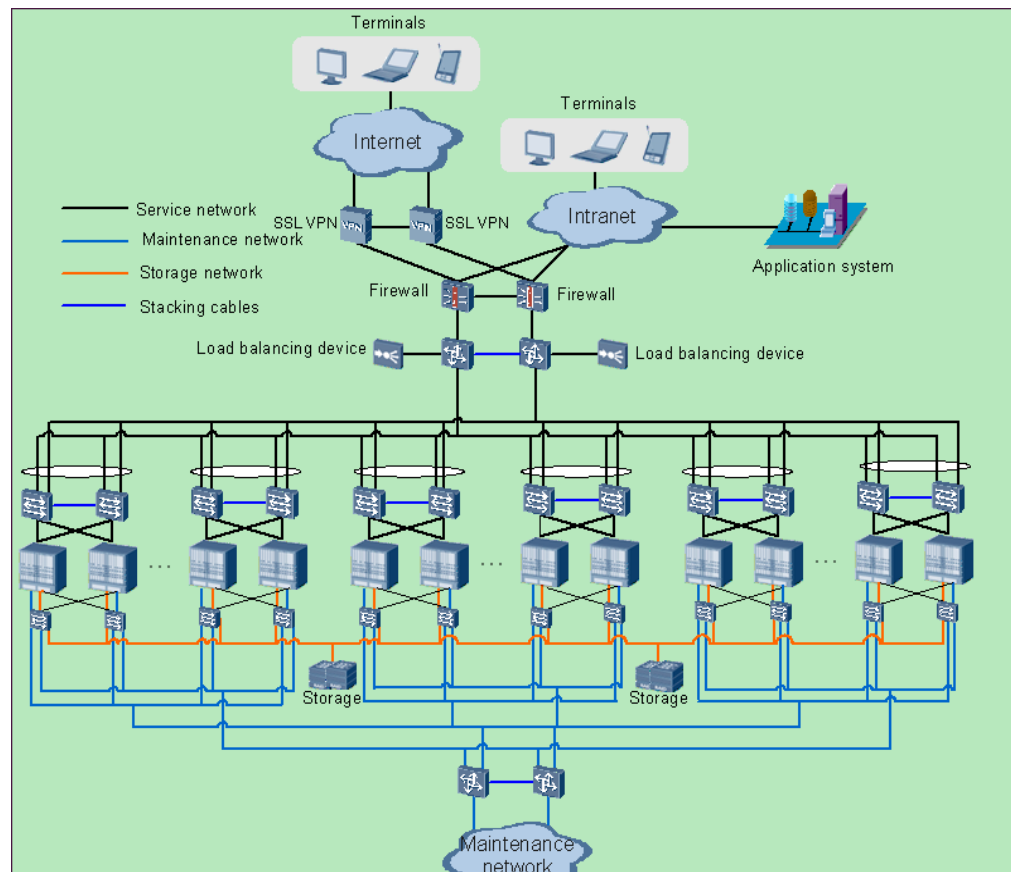


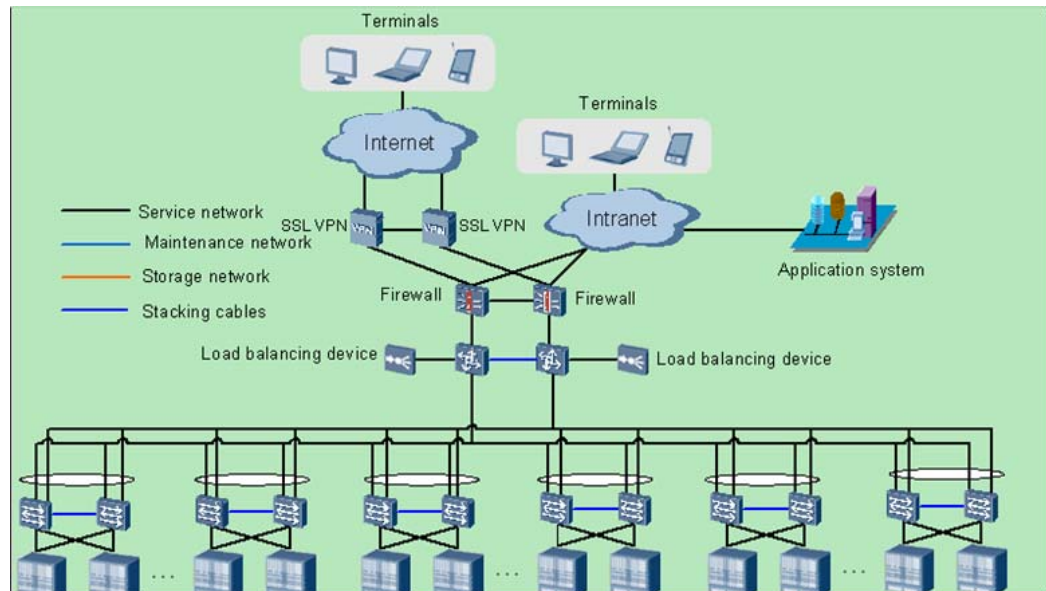
Figure 4-5 is a network configured with three service blocks. Each service block supports 2000 concurrent desktop users. The service network, maintenance network, and storage network are designed independently in the entire desktop cloud network. The service network and maintenance network consist of two layers: access layer and combined core layer. The firewall and load balancing devices are shared globally on the whole network.

The data reliability is guaranteed by the high reliability of the storage system. Therefore, no standby system or network is needed.

In the desktop cloud application system, servers are deployed in the original DC instead of the desktop cloud DC. The desktop cloud center, however, can be deployed as a large area of the original DC.

4.3 Service Network Planning

Figure 4-6 Networking of the service network in the desktop cloud DC



The service network of the desktop cloud is similar to the campus network. In the service network, desktops are deployed in the DC, each server supporting 20 virtual desktops. Therefore, box switches in the campus network are necessary to be deployed in the desktop cloud DC.

For example, in an E6000 subrack used in the desktop cloud, 10 blade servers are deployed, each supporting 20 to 23 concurrent desktops. That is, each E6000 subrack supports 200 to 230 concurrent desktops.

In addition, an E6000 subrack is configured with six NX910 modules, each providing 10 GE electrical interfaces. That is, each E6000 subrack supports 200 to 230 concurrent desktops and provides 60 GE electrical interfaces. Based on the bandwidth statistics shown in [Table 4-1](#), two upstream GE electrical interfaces can meet service requirements.

Therefore, S5700 switches are deployed at the access layer and combined core layer. Switches at the access layer are connected to the combined core layer through GE interfaces.

The performance of the load balancer and firewall is calculated using the following formula:

$$\text{Performance} = \text{Number of areas} \times \text{Number of GE interfaces}$$

4.3.1 Bandwidth of the Service Network

[Table 4-1](#) shows the bandwidth usage based on user behavior in common offices.

Table 4-1 Bandwidth usage based on user behavior

User Behavior	Receiving Bandwidth (kbit/s)	Sending Bandwidth (kbit/s)
Opening a folder	40.40	28.90

User Behavior	Receiving Bandwidth (kbit/s)	Sending Bandwidth (kbit/s)
Editing or browsing a Word file	346.51	13.51
Browsing a PPT	535.37	15.64
Browsing a web page	201.71	22.20
Playing music	1658.90	51.91
Playing a video	1293.66	36.28
Playing a high-definition video	9824.12	331.21
Performing the VoIP service	182	180

Based on the preceding table, the bandwidth of each desktop user is 150 kbit/s. The bandwidth for the desktop access varies according to service applications and user behaviors. Generally, the 200 kbit/s bandwidth is sufficient for web page applications and enterprise intranet applications.

In the service network, the total bandwidth needed is:

350 kbit/s x Number of concurrent desktops

If the service network has 2500 concurrent desktops, the required bandwidth is:

350 kbit/s x 2500 = 875 Mbit/s

Therefore, two 1GE ports are sufficient for the service network

Similarly, 1 Gbit/s throughput is sufficient for the SSL VPN gateway and load balancer.

4.4 Security Planning

In the desktop cloud DC, a virtual machine is a desktop system. The network admission control (NAC) solution for the campus network is used as security control mechanism in the desktop cloud DC. In the desktop cloud DC, mobile storage devices and external interfaces do not need to be managed, which is different from the campus network. The following security measures are needed in the desktop cloud DC:

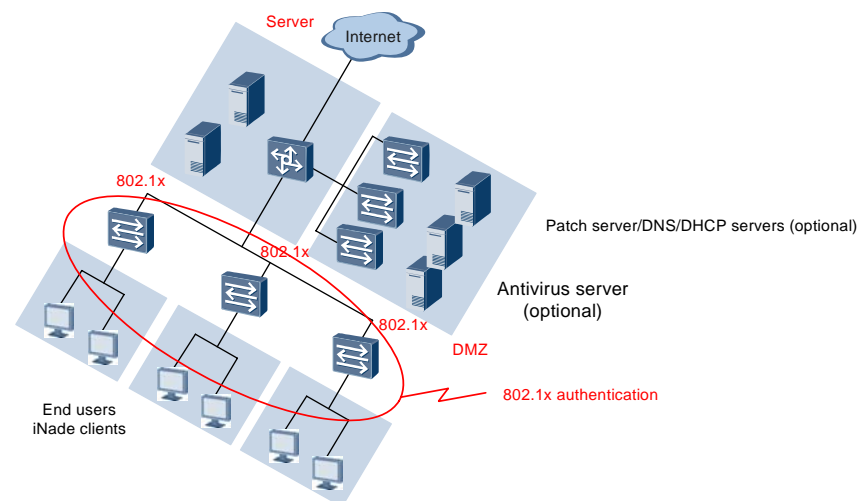
- Network behavior management
 - Controlling network traffic
 - Defending against ARP attacks
 - Controlling instant messages, stock trading, peer-to-peer software, and online games
 - Controlling Web access and IP access
- Terminal behavior management
 - Monitoring file operation
 - Forcibly running security control software
 - Preventing non-standard software from being installed

- Disabling risking services (such as the DHCP service) forcibly, or enabling necessary services

Two solutions are available to deploy policy enforcement points (PEPs) on the network devices:

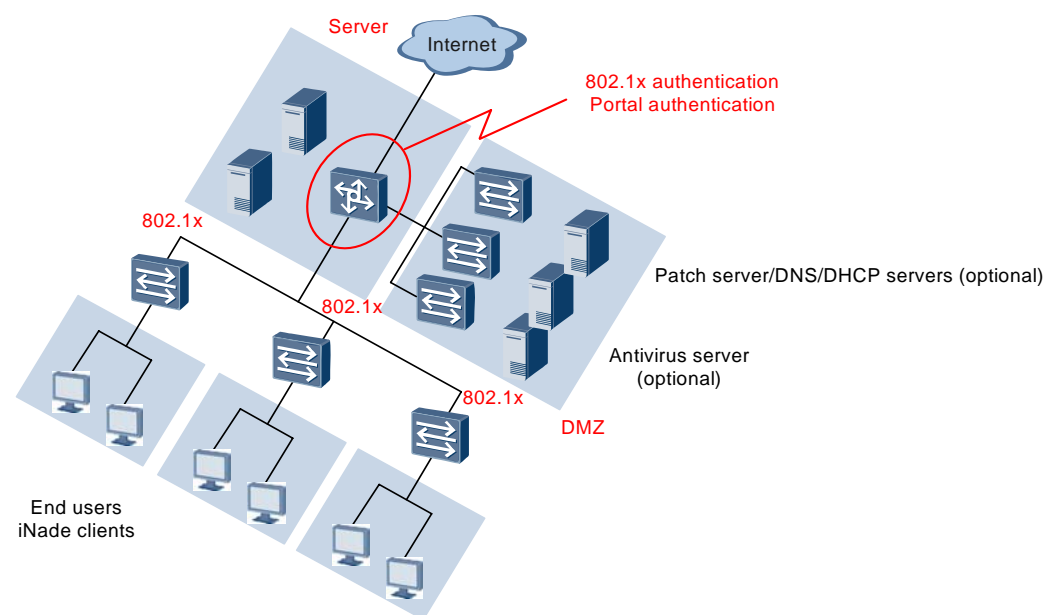
- The 802.1x authentication can be enabled on switches at the access layer. This isolates users who do not meet the security policy requirements, preventing security threats on the network.

Figure 4-7 Switches with 802.1x authentication enabled at the access layer



- The 802.1x authentication (based on the MAC address) can be enabled on switches at the aggregation layer, and security functions (such as port isolation and private VLAN) are enabled on switches at the access layer, preventing mutual influence between virtual switches that are configured on a single access switch.

Figure 4-8 Switches with 802.1x authentication enabled at the aggregation layer



5

Suggestions on Planning Multiple DCs

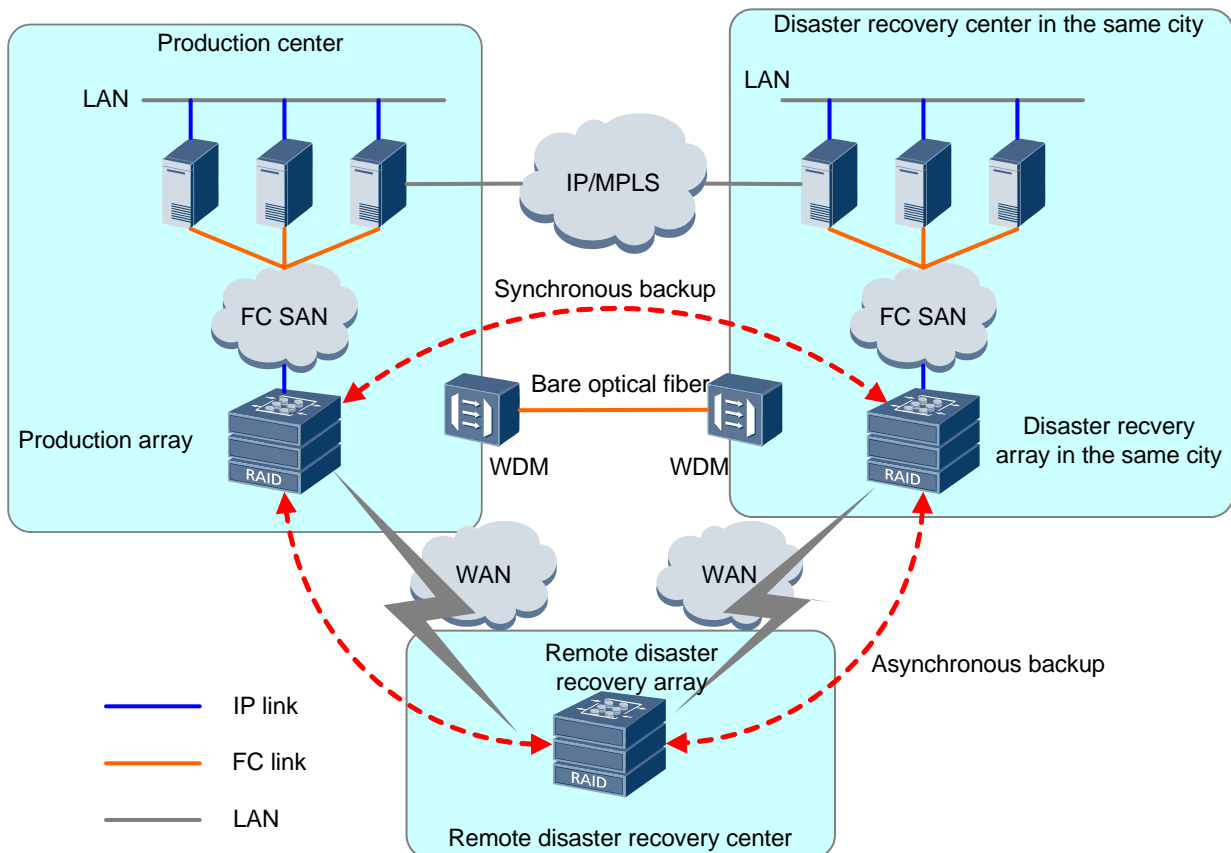
5.1 Network Architecture of Multiple DCs

A DC hosts core services and stores mass service information. It needs to provide services around the clock. A standby DC is established within the range of 50 km from the active DC in the same city to replicate service data in real time using dedicated lines or transmission devices. During data backup, some services in the active DC can be migrated to the standby DC so that both active and standby DCs are in the active state.

In case of natural disasters such as earthquakes, it is recommended that the enterprise establishes a remote disaster recovery center in another city more than 400 km away from the active and standby centers. The remote disaster recovery center is responsible for backing up data in the active and standby DCs and periodically synchronizes data between the production center and the disaster recovery center in the same city. When a disaster occurs, the remote disaster recovery center can recover services using backup data, ensuring data integrity.

[Figure 5-1](#) shows the network architecture of multiple DCs.

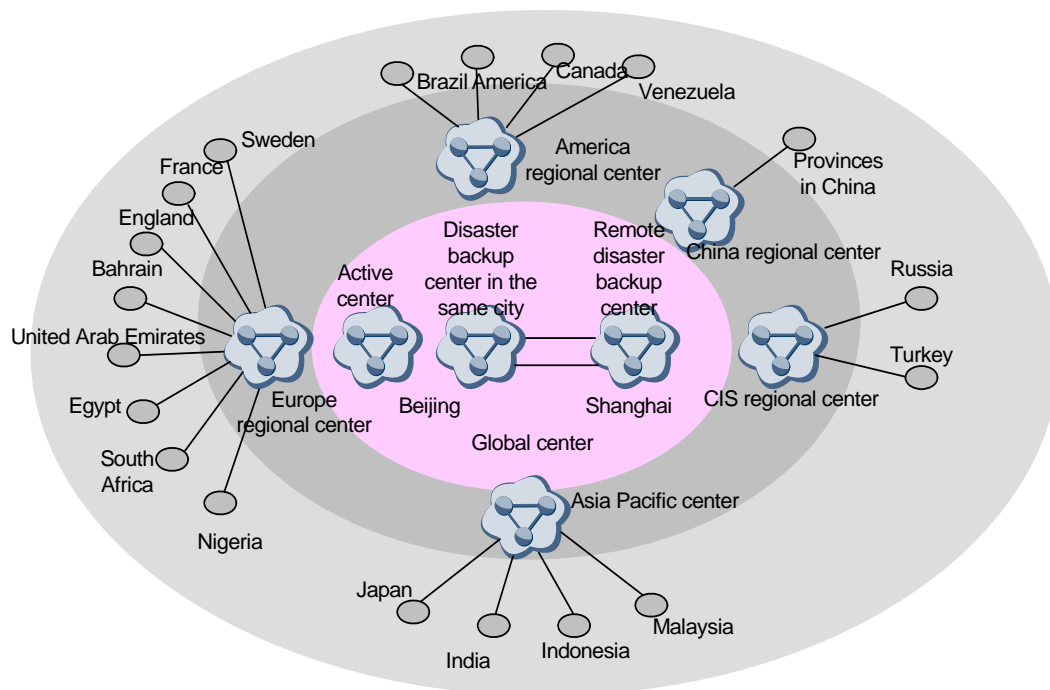
Figure 5-1 Network architecture of three centers in two areas



As more services are deployed in the enterprise, the network architecture of three centers in two areas cannot meet the requirements for service development. The architecture of multiple centers with different levels has emerged to replace the original network architecture. If DCs with different levels are established in a region, the load of global DCs is lessened, the WAN bandwidth is saved, and the response time of regional services is shortened. In addition, if a fault occurs in a region, services in other regions are not affected.

Figure 5-2 shows the network architecture of multiple centers with different levels.

Figure 5-2 Network architecture of multiple centers with different levels



5.2 Network Reliability Planning

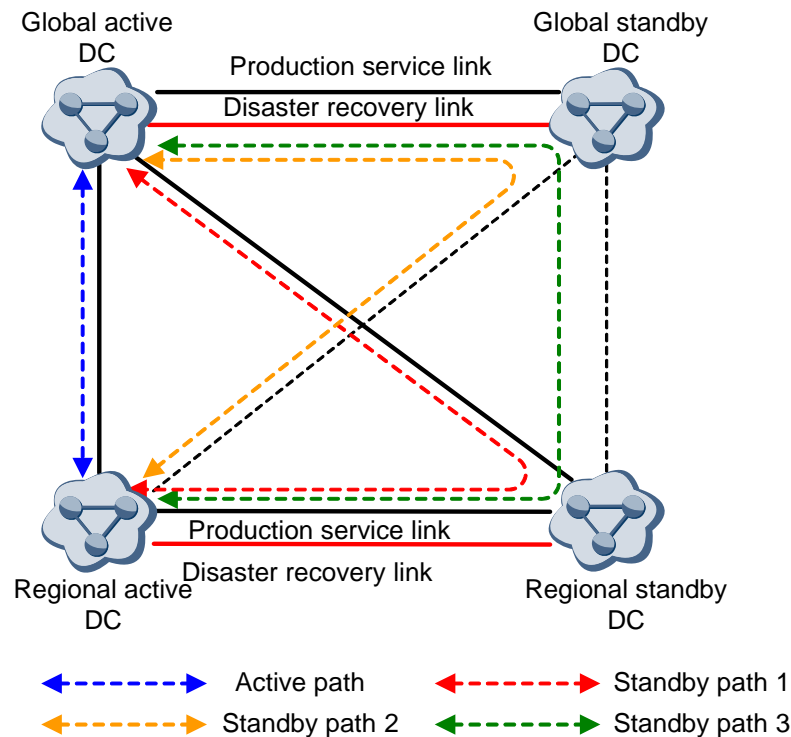
5.2.1 Network Reliability Between Regional DCs and Global DCs

The global active DC is connected to the global disaster recovery center using two independent links: a production service link and a disaster backup link. The two links are isolated to guarantee bandwidth.

The regional active DC is connected to the regional disaster recovery center using two independent links: a production service link and a disaster backup link. The regional active DC is connected to the global active DC and the regional disaster recovery center is connected to the global disaster recovery center.

[Figure 5-3](#) shows the network topology between global DCs and regional DCs.

Figure 5-3 Plan for active and standby paths connecting DCs



Four DCs are defined as four autonomous systems (ASs). They advertise routes using EBGP. As shown in Figure 5-3, the regional active DC has four paths to the global active DC. Priorities of four paths are as follows:

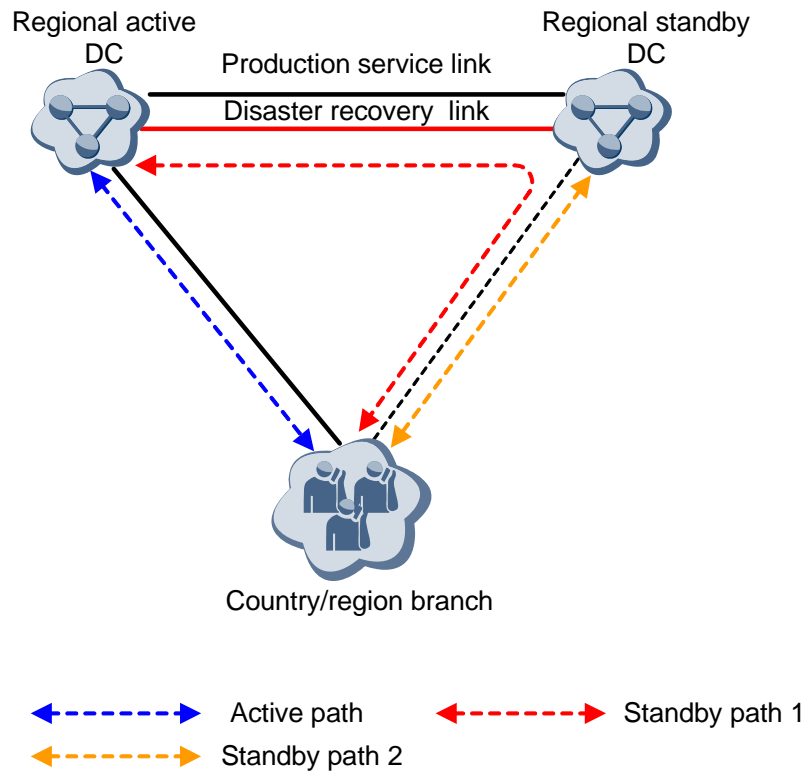
- Highest priority: active path. If the link is normal, the regional active DC is directly connected to the global active DC.
- Second highest priority: standby path 1. If the gateway or the outbound link of the regional active DC is faulty, the regional active DC is connected to the global active DC through the regional standby DC.
- Third highest priority: standby path 2. If the access device of the global active DC is faulty, the regional active DC is connected to the global active DC through the global disaster recovery center.
- Lowest priority: standby path 3. If the preceding errors occur concurrently, the regional active DC is connected to the global active DC through the regional standby DC and then global disaster recovery center.

The priorities of the links are determined by the EBGP AS-Path and MED attributes.

5.2.2 Network Reliability Between a Country/Region Branch and Regional DCs

A country/region branch is connected to the regional active/standby DCs by using active/standby links from different carrier. Figure 5-4 shows the network topology between a country/region branch and regional centers.

Figure 5-4 Country/region branch's connection to regional DCs



The active link of the country/region branch is connected to the regional active DC and the standby link to the regional standby DC. The regional active DC, regional standby DC, and country/region branch are defined as different ASs by EBGp.

- Active path. Generally, the country/region branch is directly connected to the regional active DC using the active access link.
- Standby path 1. If the active access link is faulty, the country/region branch is connected to the regional active DC through the regional standby DC using the standby access link.
- Standby path 2. If the regional active DC is faulty, the traffic is switched to the standby path 2 on the application layer using the domain name system (DNS) mechanism.

5.3 Route Planning

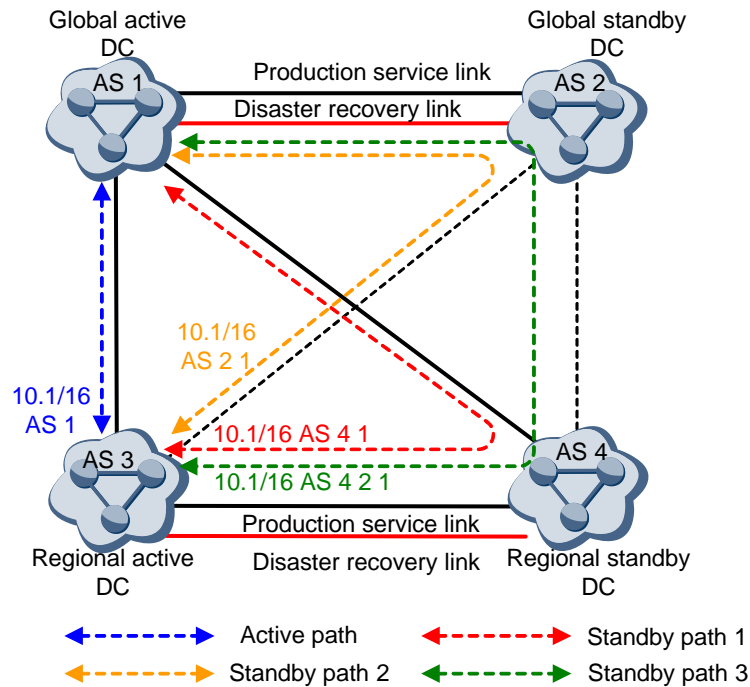
5.3.1 Routing Overview

Generally, you only need to deploy Interior Gateway Protocols (IGPs) including Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS) in the DC. To manage and maintain the network conveniently, it is recommended that you use the OSPF dynamic routing protocol to ensure network stability and fast convergence of routes. BGP is used to advertise routes between DCs. With a powerful routing control capability and abundant routing policies, BGP is applicable to interconnection between large networks.

5.3.2 BGP Design

After regional DCs and global DCs are interconnected, each DC is defined as an AS. ASs advertise their routes using EBGP. As shown in Figure 5-5, the AS-Path and MED attributes are used to control and select routes with EBGP, enhancing link reliability.

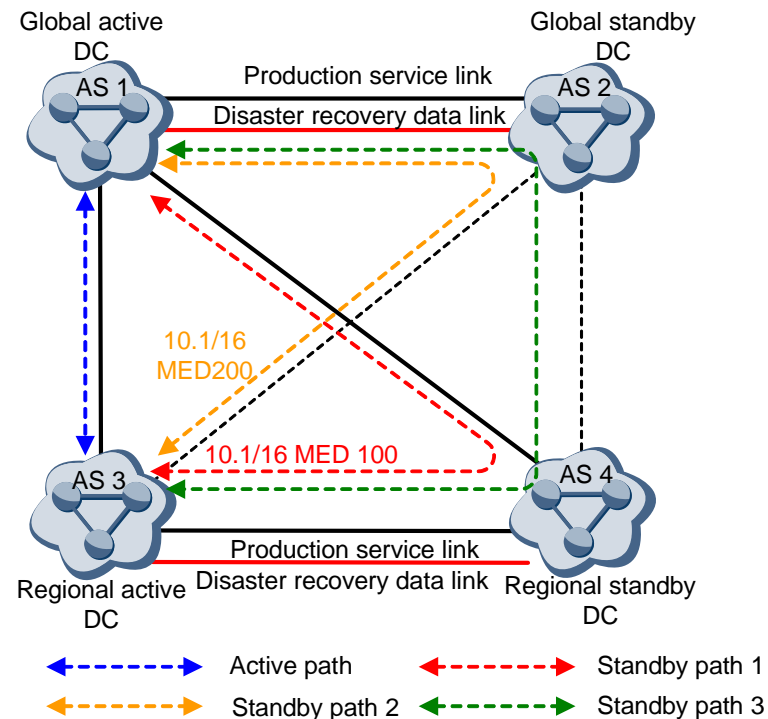
Figure 5-5 BGP AS-Path route selection



EBGP prefers the route with the shortest AS-Path. As shown in Figure 5-5, AS 3 receives information on route 10.1/16 from AS 1, AS 2, and AS 4. The AS-Paths of these routes are AS 1, AS 2 1, AS 4 1, and AS 4 2 1.

- The route advertised from AS 1 (active path) has the shortest AS-Path. Therefore, it has the highest priority and is selected.
- The route advertised from AS 4 2 1 (standby path 3) has the longest AS-Path. Therefore, it has the lowest priority.
- The routes advertised from AS 2 1 (standby path 2) and AS 4 1 (standby path 1) have the same AS-Path. The BGP MED attribute is needed to distinguish their priorities. As shown in Figure 5-6, the MED value of route 10.1/16 advertised from AS 4 is 100, smaller than that of the route advertised from AS 2. Therefore, standby path 1 has a higher priority than standby path 2.

Figure 5-6 BGP MED route selection



BGP has powerful routing control and selection capabilities. By controlling the BGP AS-Path and MED attributes, you can effectively solve the route selection and link reliability problems in multiple DCs.

5.4 Disaster Recovery Planning

5.4.1 Disaster Recovery Overview

The disaster recovery center is a computer network system established as a backup to the production center. When the production center stops working due to a disaster, the disaster recovery center takes over all or some of the services in the production center in a timely manner, which minimizes or avoids losses caused by the disaster. Therefore, the disaster recovery center can provide comprehensive and high-quality services for enterprises.

The disaster recovery system is classified into the following seven tiers according to the international standard Share 78:

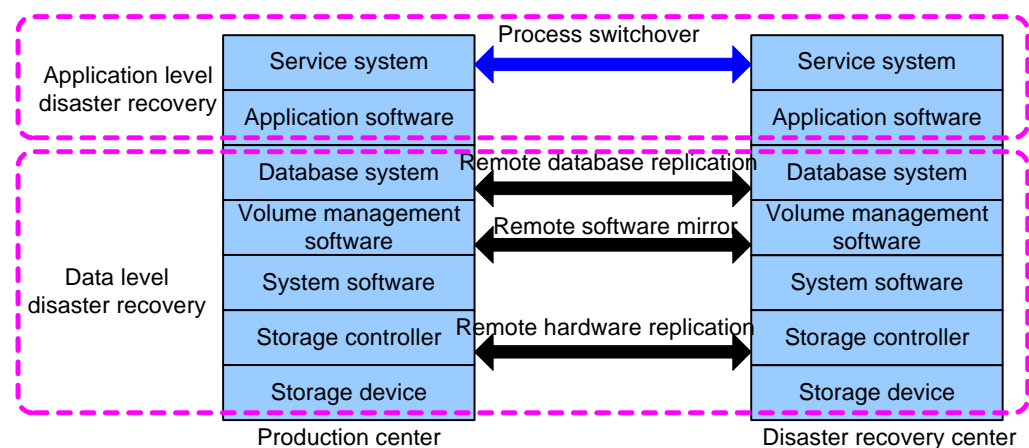
- Tier 0: No off-site data
- Tier 1: Data backup with no hot site
- Tier 2: Data backup with a hot site
- Tier 3: Electronic vaulting
- Tier 4: Point-in-time copies
- Tier 5: Transaction integrity
- Tier 6: Zero or near-Zero data loss
- Tier 7: Highly automated, business integrated solution

Based on data and service features of each tier, Huawei classifies the disaster recovery system into the following three levels:

- Tier 0 to 2: Backup level.
- Tier 3 to 5: Data level disaster recovery. A remote data system is established to replicate the key application data of the local system in real time. When a disaster occurs, the remote data system takes over services of the local system to ensure service continuity.
- Tier 6 to 7: Application level disaster recovery. A backup application system with a higher level than the data disaster recovery system is established in a remote area. The backup application system and the local application system can back up each other and work together. When a disaster occurs, the remote application system takes over services of the local application system.

Figure 5-7 shows the service frameworks of the data-level and application-level disaster recovery systems.

Figure 5-7 Disaster recovery levels



Two technical specifications are used to measure disaster recovery:

- Recovery point objective (RPO): acceptable amount of data loss
- Recovery time objective (RTO): acceptable longest duration within which services are interrupted or the shortest duration between the time when a disaster occurs and the time when services are restored

RPO measures data loss, while RTO measures service loss. RPO and RTO are not necessarily related. RTO and RPO vary according to services and enterprises, and are calculated based on service requirements after risk analysis and service influence analysis are performed.

5.4.2 Disaster Recovery Overview

In most cases, Huawei recommends that two DCs (active and standby) be built for remote disaster recovery. Applications run on the computer system of the active DC and data is stored on the storage system of the active DC. When the active DC stops working due to a disaster such as power outage, fire or earthquake, traffic is switched to network cables and PSTN lines connected to the standby DC where applications are restarted.

It takes a short time to finish the switchover. This type of recovery ensures the continuity and integrity of data in both centers.

The traditional tape backup is performed at a fixed point. If the system corrupts, data communicated from latest backup to the disaster occurrence is lost and cannot be recovered. In this backup mode, the backup speed is slow and the backup process is not performed in real time. Therefore, it cannot meet requirements for recovering a large amount of data, database continuity, and real-time performance.

The mainstream disaster recovery solution is real-time backup. A real-time data recovery can replicate updated data from the active DC to the standby DC through communications links, ensuring synchronization between the active and standby DCs. If the active DC cannot work properly, the standby DC takes over services of the active DC and maintains data integrity.

Layered Data Replication Technologies

Based on different layers in the information system, different IT technologies can be used to synchronize or replicate data. The information system is divided into six layers:

- Disk array layer
- Storage area network (SAN) layer
- Logical volume manager layer
- File system layer
- Database layer
- Application system layer

The following is the replication technologies associated with the preceding six layers:

- Mirror-based replication technology

The core of this technology is to replicate production data remotely using the storage array's disk-array-to-disk-array data block replication technology, which ensures the security of the production data in a disaster. If a disaster occurs in the active DC, data in the disaster recovery center can be used to establish an operating environment to provide IT support for services. Data in the disaster recovery center can also be used to recover the service system of the active DC to recover services quickly.

The mirror replication between disk arrays does not occupy the system CPU, memory, and I/O resources, and has little impact on the application system because it does not involve the host operating system. This is the most mature and widely used disaster recovery technology. However, it requires that the same type of storage devices from the same manufacturer be used in the production center and disaster recovery center.

Storage devices of mainstream manufacturers provide the disk array-level mirror replication technology, such as EMC DMX SRDF, EMC CX MirrorView, IBM DS8000 MetroMirror, IBM DS8000 GlobalMirror, IBM DS4000 ERM, HP XP ContinuousAccess, and HDS USP TrueCopy.

- SAN-based replication technology

This new technology has emerged in recent years. On a SAN network, a virtual storage management device is deployed in a direct or bypass manner depending on manufactures.

The SAN-based technology is applicable to heterogeneous storage devices and transparent to the host. You can use this technology when disk arrays from many manufactures exist in one DC, but it is immature and has an impact on the background I/O storage speed

The products that provide this technology now include IBM SVC, EMC Invista, and Falcon Ipstor.

- Volume manager-based replication technology

This technology functions at the volume manager layer and it mirrors or replicates disk volumes to implement disaster recovery. This technology does not require the same storage devices on both production centers and disaster recovery centers, but it occupies system CPU resources and has a great impact on the system performance. Therefore, it has poor scalability and running performance. This technology is based on the host, so unexpected unauthorized access to the protected data may occur, affecting system stability and security.

Commonly used volume replication software includes Symantec Veritas Volume Replicator.

- File system-based replication technology

This technology replicates data files from the production center to the disaster recovery center to implement data recovery. This technology functions in the file-based storage systems, such as file servers, NAS, NAS devices, or file virtualization combinations.

The file-based replication technology is widely used for backing up data. The following two reasons account for its popularity:

- This technology is easy to deploy and supports standard protocols. In addition to its own replication functions, it can work with multiple driver technologies to provide more replication functions.
- This technology provides enterprises with methods for using storage resources properly, sharing resource across media servers, and configuring storage capacity for media servers in a timely manner when the enterprises are running the block-based storage system.

- Database-based replication technology

This logical replication technology supports heterogeneous storage and operating system platforms. After analyzing redo logs of the production database, this technology generates universal or private SQL statements and transmits these statements to the backup database for application.

The replication process does not involve the lower-layer storage. The replication is performed across platforms at a high speed, but it occupies system resources, does not support some special data formats and data description language (DDL) statements, and cannot guarantee data consistency when random data is generated in the service system.

The common products that provide this technology include Oracle DataGuard, Oracle Stream, Quest SharePlex for Oracle, DSG RealSync for Oracle, and IBM DB2 HA/DR.

- Application system-based replication technology

The application system must support transaction distribution when the application system-based replication technology is used. This technology uses transaction middleware to back up online transaction concurrently in the production center and disaster recovery center, or to transmit updated data from the active DC to the standby DC, ensuring data consistency between the production center and disaster recovery center.

This technology requires low bandwidth, but existing current applications can only implement this technology after you modify these applications.

Data Backup Mode

Data can be backed up in both local and remote ends. Based on protection mechanisms of different levels, two data backup modes are available.

- Synchronous mode

Before the next write operation is performed on disks, updated data in the last write operation must be replicated to both local and remote volumes. The synchronous mode

provides the highest protection level, but application performance is affected due to the time delay caused by data transmission between arrays in local and remote ends.

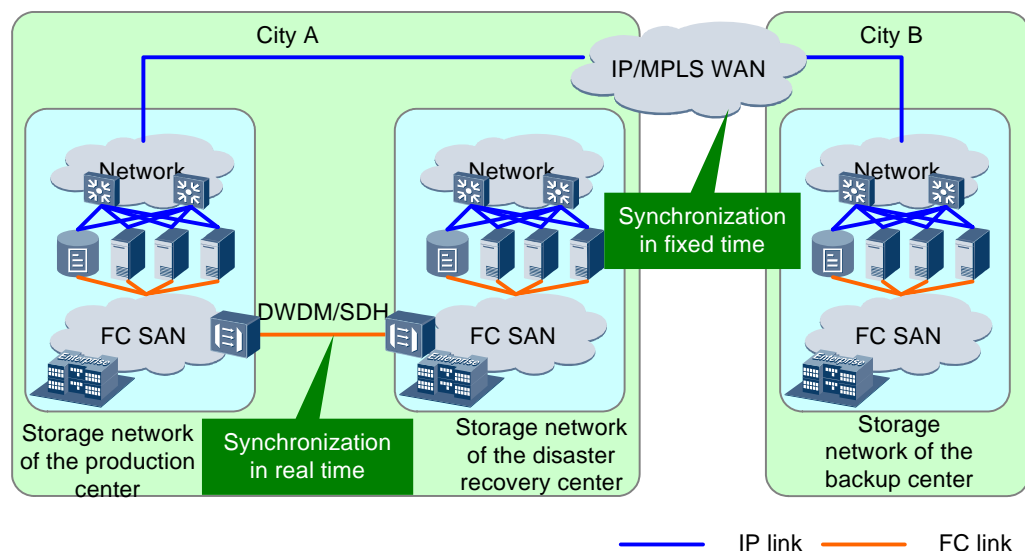
- Asynchronous mode

Local volumes can continue the write operation even if the remote volumes are not updated. Remote volumes are updated after a period of delay. This mode ensures high application performance, but data that is not updated to remote volumes will be lost if a disaster occurs.

5.4.3 Network Planning for Disaster Recovery

When you plan a network for disaster recovery, take two disaster recovery methods into consideration: real-time disaster recovery in the same city and remote disaster recovery backup. As shown in Figure 5-8, data is synchronized in real time for the disaster recovery center in the same city and in fixed time for the remote disaster recovery center.

Figure 5-8 Network planning for disaster recovery in the same/different cities



Disaster Recovery in the Same City

In the metro disaster recovery solution, Huawei recommends that core service data be backed up in synchronous or asynchronous mode based on the physical distance between the disaster recovery center and the production center.

On the FC SAN network, the Wavelength Division Multiplexing/Synchronous Digital Hierarchy (WDM/SDH) technology can be used to back up the network remotely, and the mirror-based replication technology can be used to synchronize data in real time.

If the distance between the disaster recovery center and the production center is within 100 km and two centers are connected using optical fibers, some core service data can be backed up in synchronous mode while the others in asynchronous mode with regard of transmission delay of optical fiber signals.

If two centers are connected using IP data links, the IP SAN-based communication protocols can be used to transmit data, such as Fiber Channel over IP (FCIP), Internet Fiber Channel

Protocol (iFCP), Infiniband, and Internet Small Computer System Interface (iSCSI). Huawei recommends the asynchronous mode.

Remote Disaster Recovery

In the remote disaster recovery solution, data is backed up through leased lines and on the asynchronous transfer mode (ATM) network.

If users have sufficient capital, it is recommended that users use point-to-point leased lines and WAN acceleration devices to decrease the leased WAN bandwidth, providing high-speed and efficient data backup services at minimum costs.

Data is backed up in the asynchronous mode, which meets requirements for bandwidth and transmission delay in remote disaster recovery. If the amount of data exceeds the threshold in the disaster recovery center, the overflow data is backed up to the tape library or CD-ROM library using the snapshot technology.

Data transmission delay exists between the remote disaster recovery center and local production center and varies with the adopted technologies, bandwidth, distance, and characteristics of data flows. The software-based replication technologies can easily implement queuing and resumable transmission mechanisms, ensuring data consistency if a disaster occurs.

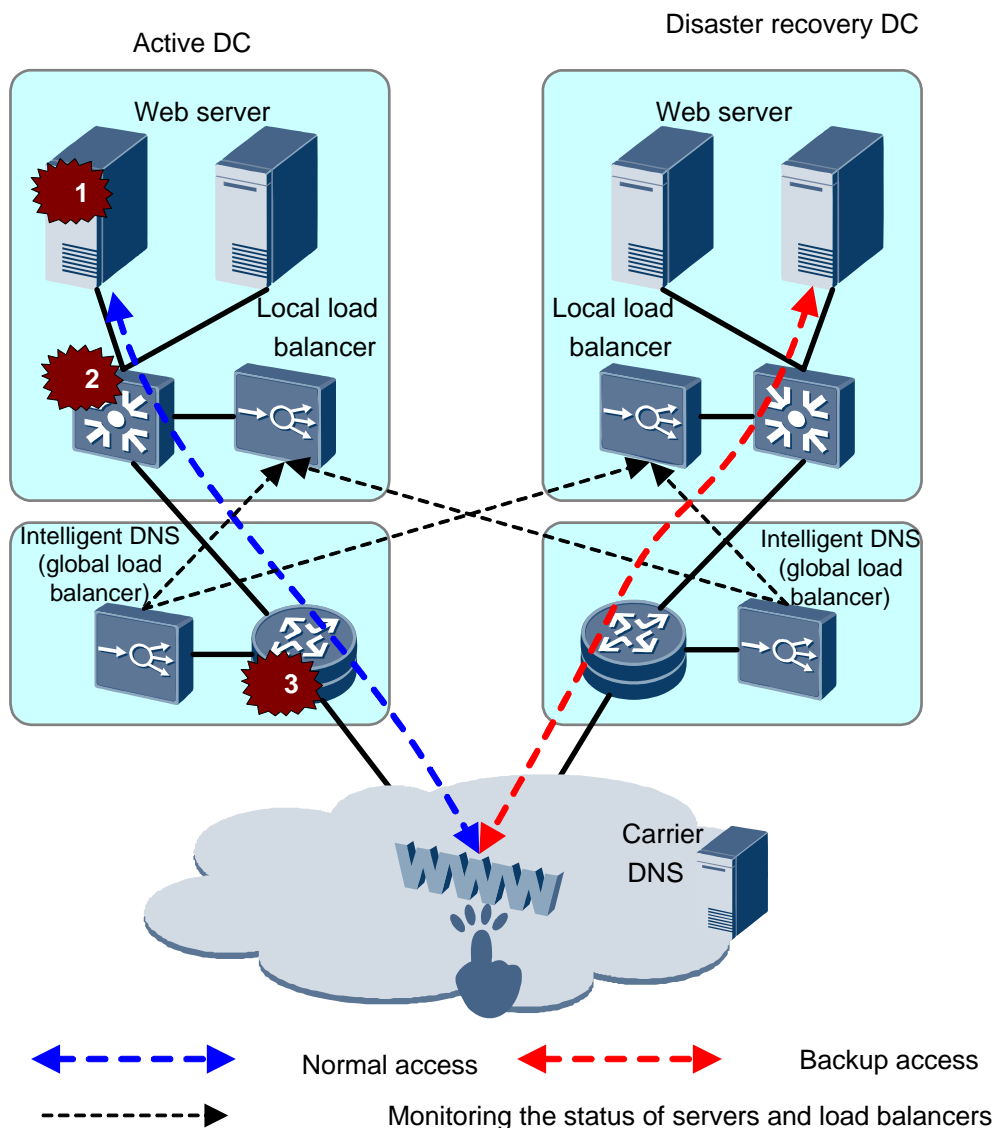
Compared with the synchronous mode, the asynchronous mode has low requirements for bandwidth and distance. It requires that all data can be replicated from the local end to the remote end within a certain period of time and does not affect application system performance. If a disaster occurs in the local production center, however, data on the remote end will be temporarily lost (if the transmission rate is low and data is not transmitted completely on WAN), but data consistency is not affected, similar to what happens if the local host is abnormally shut down.

5.4.4 Service Planning for Disaster Recovery

Based on real-time synchronization, automatic switchover and active/active load balancing can be implemented for services. As shown in [Figure 5-9](#), intelligent DNS servers (global load balancers) monitor the status of Web servers and local load balancers, and provide DNS resolution results based on the status.

If a Web server fails in the active DC, the local load balancer switches services on this Web server to the other Web server in the center. If the whole active DC fails, the global load balancer switches services in the center to the disaster recovery center.

Figure 5-9 Automatic switchover and active/active load balancing implemented based on the active/backup intelligent DNS/GSLB



The DNS service has a great impact on services in the DC, so disaster recovery for DNS servers must be taken into consideration. In multiple DCs, it is recommended that you deploy the slave DNS server in the active DC, and master DNS server in the standby DC. This guarantees the proper operation of DNS services when the whole active DC fails.

5.5 Service Distribution Planning

5.5.1 Service Distribution Overview

With the development and expansion of enterprises, the deployment mode of DCs evolves from single-center mode to three-center-in-two-area mode and multiple-center mode. Services are hosted in active DCs or regional DCs as required.

Based on user experience and service characteristics, services have different requirements for bandwidth and transmission delay. Therefore, the related DCs are deployed in different modes. For example, office automation (OA) services such as Notes and Email, are sensitive to transmission delay and require high bandwidth. Therefore, they are deployed in distributed mode, which reduces bandwidth on leased lines of regional DCs and active DCs.

5.5.2 Service Distribution Planning

Services in DCs are deployed in the centralized and distributed manner to meet operators' network requirements and increase user satisfaction. The following table lists characteristics of some application services in DCs and recommendations on their deployment modes.

Table 5-1 Centralized and distributed deployment modes of application services

Application Service	Architecture	Characteristics	Deployment Mode
OA services (such as Notes and Email)	C/S	Interactive operation: sensitive to delay Large-amount-of-data operation: sensitive to bandwidth and delay	Distributed deployment: OA services are deployed in global active DCs and regional DCs in the distributed mode.
Web service	B/S	Interactive operation: sensitive to delay Large-amount-of-data operation: sensitive to bandwidth and delay	Centralized and distributed deployment: Database servers and application servers are deployed in centralized mode. HTTP servers are deployed in distributed mode.
ERP	B/S	Sensitive to delay and error codes	Centralized deployment: ERP is distributed in global active DCs in centralized mode.
Video VoIP	-	Sensitive to bandwidth and jitter	Centralized and distributed deployment: Gatekeepers (GKs) are deployed in centralized mode. Multipoint Control Units (MCUs) are deployed in distributed mode.
Interactive production services	-	Interactive operation and low bandwidth	Interactive production services are deployed in DCs in centralized mode

Centralized and distributed deployment modes are applicable to the services in [Table 5-2](#).

Table 5-2 Services deployed in the centralized/distributed mode

Deployment Mode	Applicable To
Distributed deployment	<ul style="list-style-type: none">• Services distributed in regions• Services limited within regions• Services with heavy traffic and frequent interactions
Centralized deployment	<ul style="list-style-type: none">• Services with light traffic, such as services in the early development stage• Services of great importance and requiring surveillance by headquarters

With the global load balancing technology, distributed services meet requirements of the nearest enterprises, back up data in DCs for each other in multiple locations, and perform load balancing among multiple DCs.

- DCs provide services for the nearest enterprises.
- If the server fails or a fault occurs on the network in a DC, the remote DC will provide services for users who are not informed.
- The global load balancing technology provides the intelligent DNS function to distribute enterprises' services, ensuring load balance among servers in multiple locations.

With the global server load balancing (GSLB) technology, the redirection function can be implemented. The redirection process is as follows:

- a. A user sends a Hypertext Transfer Protocol/Real Time Streaming Protocol (HTTP/RTSP) request.
- b. GSLB servers communicate with each other to select a proper DC to provide services for the user.
- c. The GSLB server that is nearest to the user replies to the user with an HTTP/RTSP 302 redirection message which contains the virtual IP address of the selected Internet data center (IDC).
- d. User's HTTP/RTSP request is redirected to the virtual IP address of the selected IDC.

6 DC Network Maintenance Recommendations

6.1 Network Management

Huawei eSight is a new generation of NMS targeting the enterprise campus and branch. It can uniformly manage enterprise resources, services, and users.

The eSight manages all IT devices, IP devices, and third-party devices, intelligently analyzes network traffic and access users' roles, and automatically adjust network control policies to ensure enterprise network security. In addition, it provides a flexible and open platform based on which enterprises can develop their intelligent management systems.

6.1.1 Network Routine Maintenance

Overview

Routine maintenance is complex and the workload is heavy. The following tasks are involved in maintenance:

- Monitoring topology objects
- Monitoring network elements
- Configuring network elements
- Monitoring services
- Diagnosing faults
- Monitoring performance
- Checking resources
- Generating reports

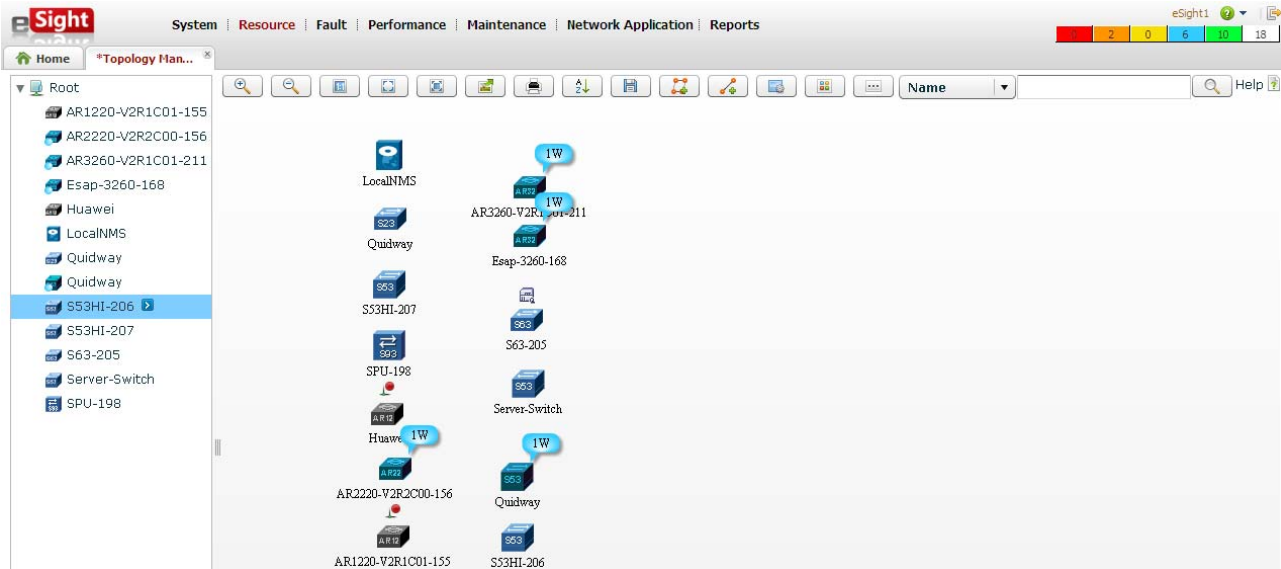
Huawei eSight can quickly and accurately provide required information for network administrators, which significantly relieves workload. The eSight provides abundant management functions and various maintenance methods for operators to implement routine maintenance easily.

Managing Topologies

As shown in [Figure 6-1](#), the eSight topology view displays the navigation tree on the left and the view on the right. The navigation tree displays the hierarchy of the network structure

while the view displays hierarchical objects in different coordinates so that users can learn about the object deployment in a clear and direct way.

Figure 6-1 Monitoring topology objects



The eSight topology view provides the following functions:

- Adding, deleting, modifying, and querying subnets, network elements (NEs), links, and virtual NEs
- Moving elements on the topology
- Displaying the alarm status and tips
- Arranging NEs, viewing NE attributes, zooming in or zooming out the NE icons, and printing the topology view.
- Providing shortcut access interfaces, such as the shortcut to accessing the NE manager or viewing device alarms

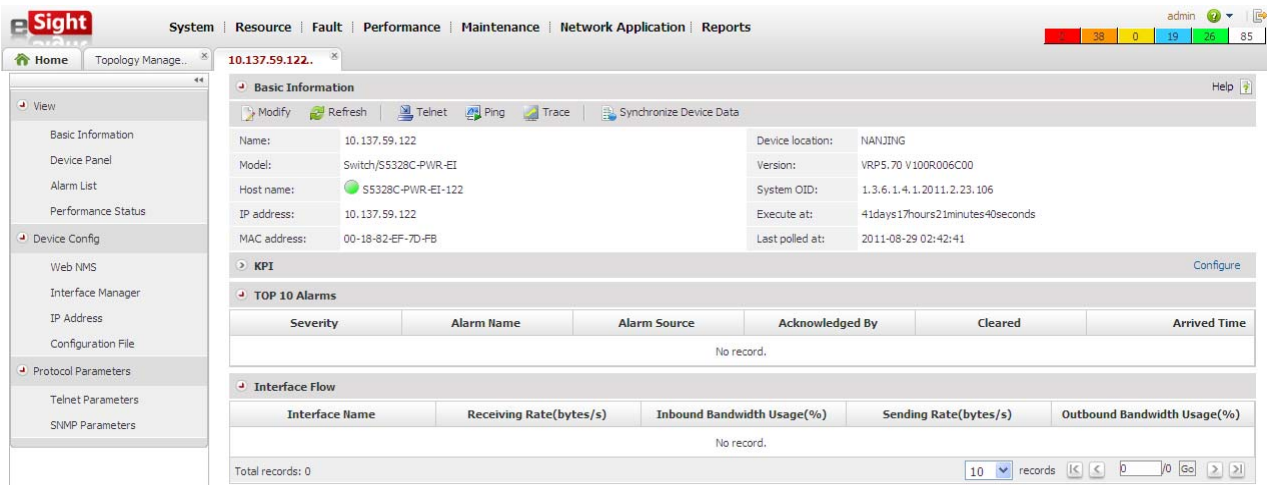
The eSight topology alarm provides the following functions:

- Coloring topology nodes to mark polling status (normal, unknown, or offline)
- Masking alarms of low severity. If multiple alarms are generated by NEs or subnets, only the alarm of the highest severity is displayed.

Monitoring NEs

The homepage of the NE manager displays basic information about NE devices, TOPN alarms, interface traffic, bandwidth usage, CPU, and memory in tables. Users can determine whether to display these performance tables as required.

Figure 6-2 NE management



As shown in Table 6-1, the eSight provides abundant NE monitoring and management functions for various devices.

Table 6-1 NE monitoring functions provided by the eSight

Device Type	Functions Supported by eSight
Huawei router and switch	<ul style="list-style-type: none">Collecting performance data and monitoring alarmsManaging the basic information about devicesSupporting device status query based on simulation images on the device panel and displaying the status of boards and interfacesSupporting the IP address query and interface information queryConfiguring and managing a single NEChecking, backing up, recovering, and comparing configuration files of devicesManaging network resources, including devices, subracks, boards, subcards, and ports
Huawei firewall	<ul style="list-style-type: none">Collecting performance data and monitoring alarms based on certain standardsManaging the basic information about devicesSupporting device status query based on simulation images on the device panel and displaying the status of boards and interfacesSupporting the IP address query and interface information queryConfiguring and managing a single NE using its Web NMSChecking, backing up, recovering, and comparing configuration files of devices

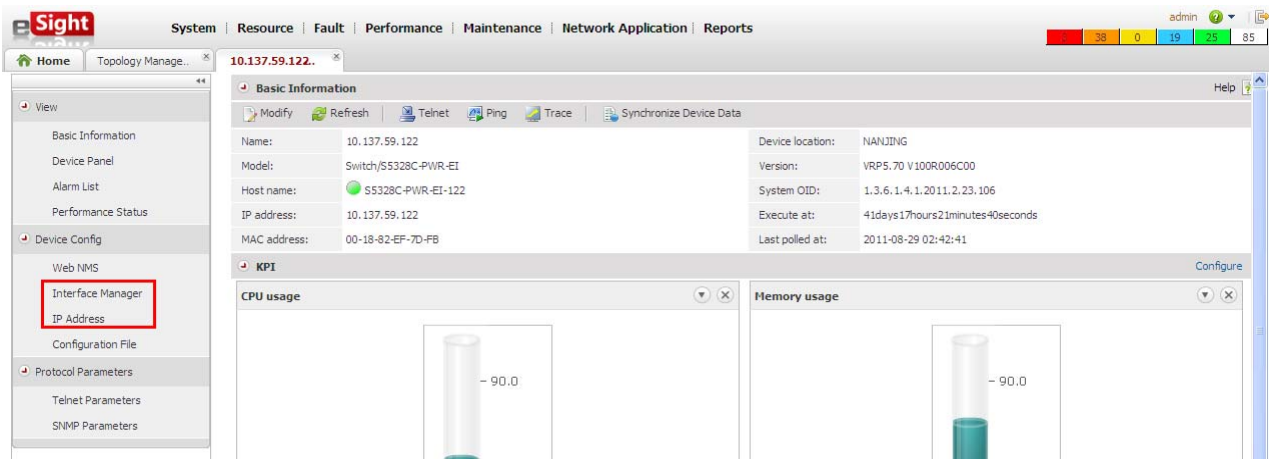
Device Type	Functions Supported by eSight
Pre-integrated mainstream Cisco and H3C devices	<ul style="list-style-type: none">• Collecting performance data and monitoring alarms based on certain standards• Managing the basic information about devices• Supporting device status query based on simulation images on the device panel and displaying the status of boards and interfaces• Supporting the IP address query and interface information query• Configuring and managing a single NE using its Web NMS• Checking, backing up, recovering, and comparing configuration files of devices• Managing network resources, including devices, subracks, boards, subcards, and ports
Non pre-integrated third-party devices	<ul style="list-style-type: none">• Collecting performance data and monitoring alarms based on certain standards• Managing the basic information about devices• Supporting device status query based on simulation images on the device panel and displaying the status of boards and interfaces• Displaying device icons, collecting the device performance, reporting alarms, and backing up configuration files by entering customized data
Servers and printers	<ul style="list-style-type: none">• Collecting performance based on certain standards• Managing the basic information about devices• Supporting device status query based on simulation images on the device panel and displaying the status of boards and interfaces• Supporting the IP address query and interface information query• Configuring and managing a single NE using its Web NMS• Managing servers and printers' storage

Configuring NEs

The eSight configures a single NE in the following ways:

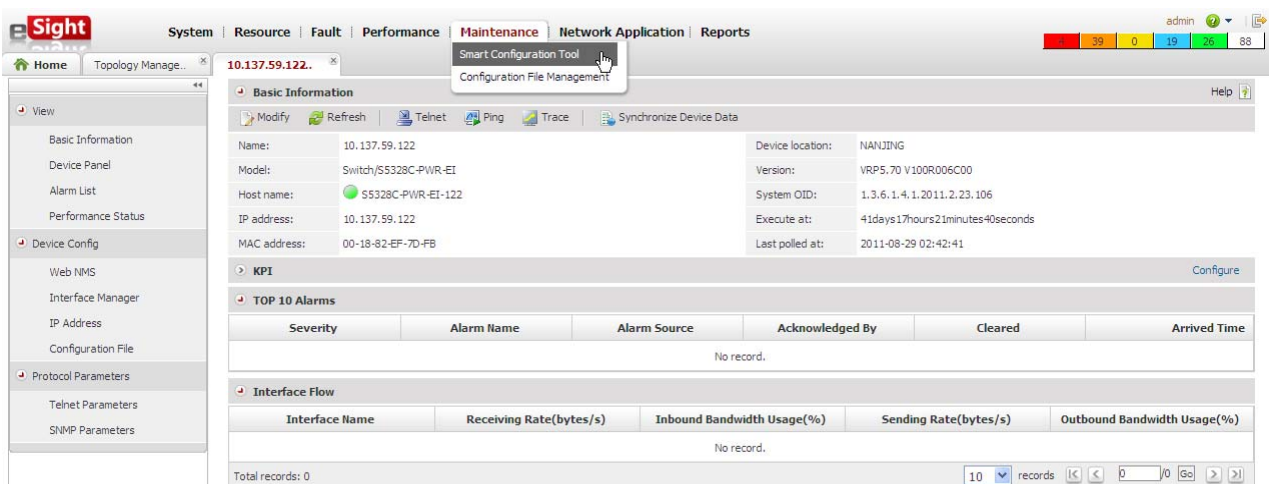
1. As shown in [Figure 6-3](#), the eSight configures interfaces and routes using the simple configuration frame.

Figure 6-3 Single NE configuration (using the simple configuration frame)

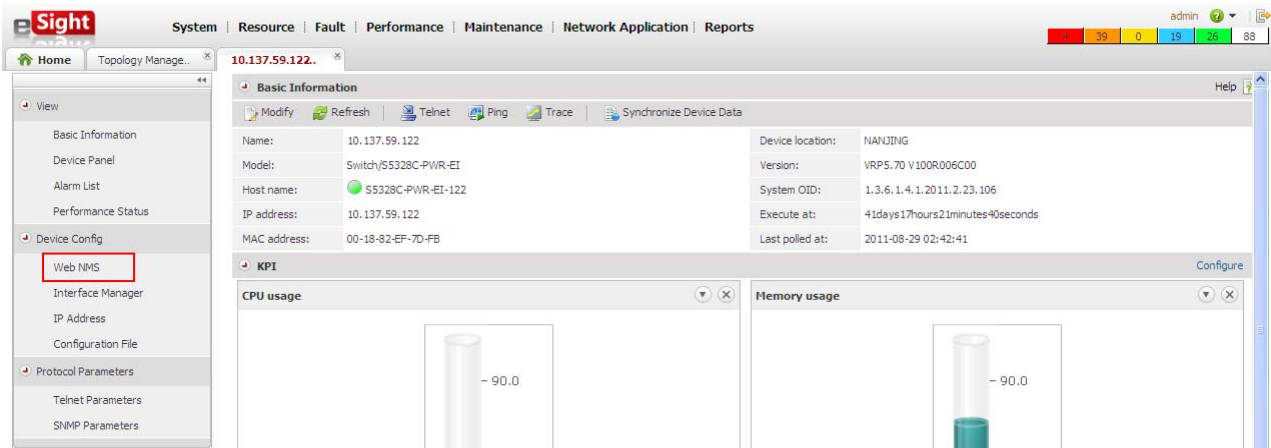


- As shown in [Figure 6-4](#), the eSight configures a single NE using the smart configuration tool.

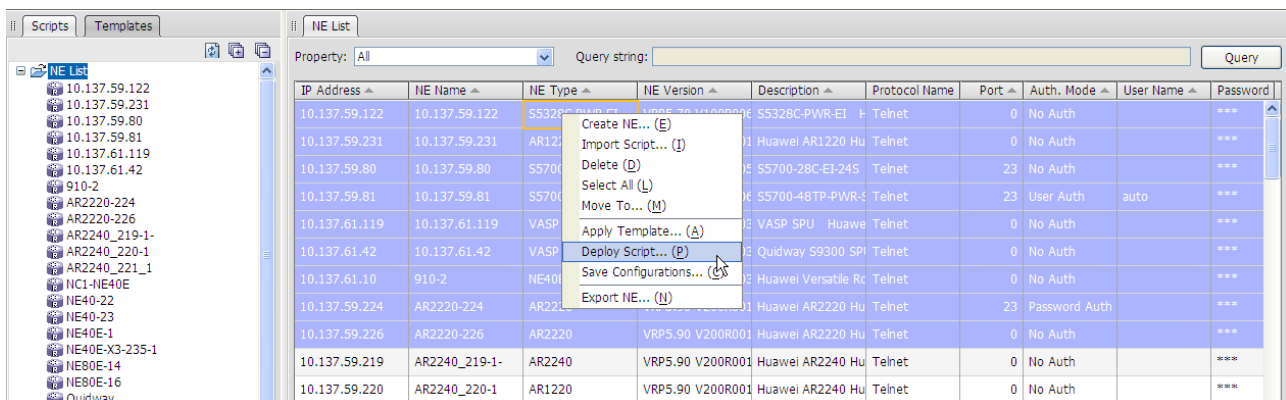
Figure 6-4 Single NE configuration (using the smart configuration tool)



- As shown in [Figure 6-5](#), the eSight configures switches, access routers, and security devices using the Web NMS.

Figure 6-5 Single NE configuration (using the Web NMS)

During new deployment and network maintenance, users need to configure services for devices deployed in centralized mode in batches. In this case, as shown in Figure 6-6, users are recommended to use the smart configuration tool to configure services for multiple devices in batches, which significantly improves operation and maintenance efficiency.

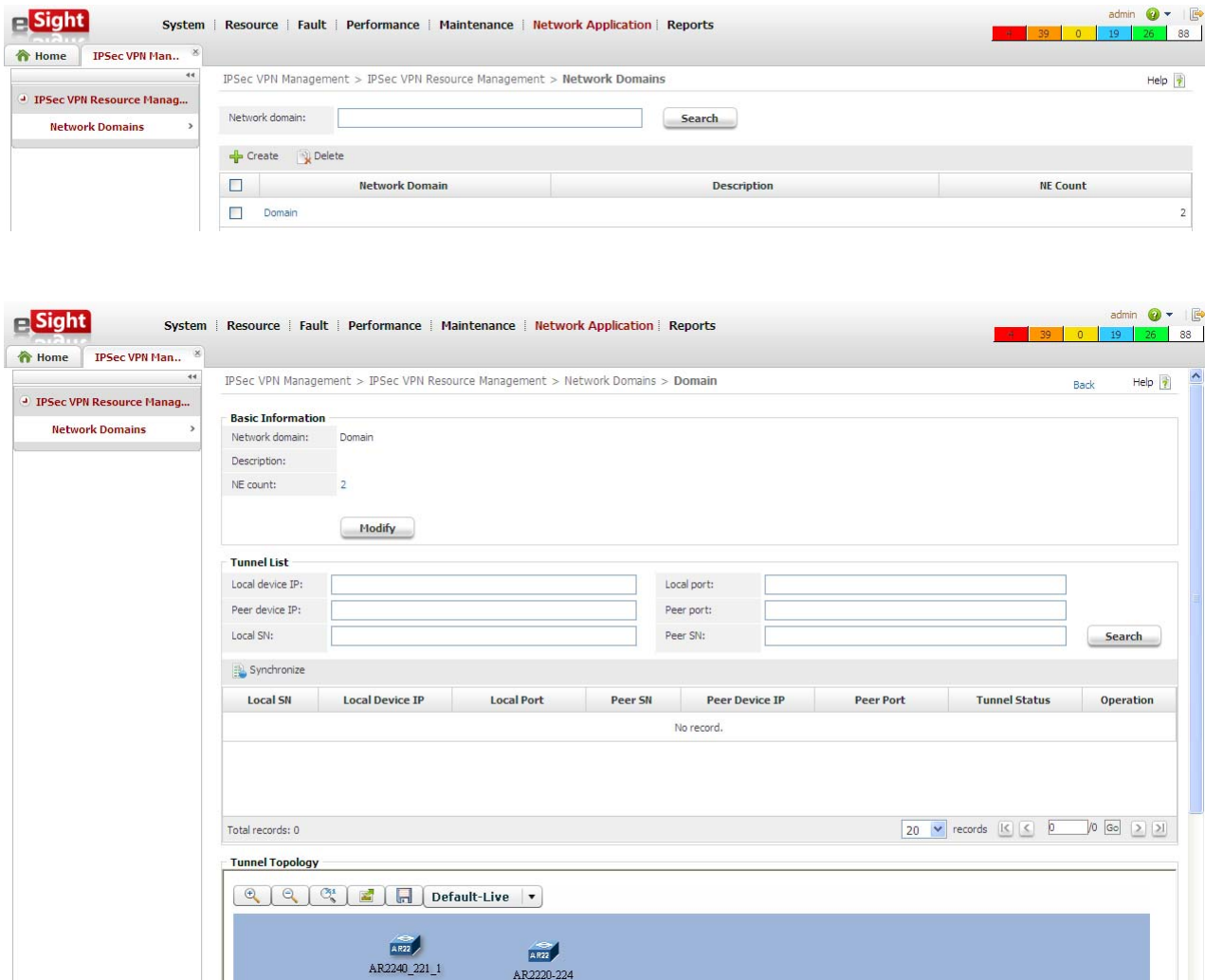
Figure 6-6 Batch NE configuration

----End

Monitoring Services

As shown in Figure 6-7, the eSight monitors services in real time and collects traffic statistics and other information based on the service type, which helps the maintenance personnel to monitor services.

Figure 6-7 Monitoring services



Monitoring Network Performance

The eSight can monitor the key performance indexes (KPIs) of a network and collects performance statistics. Users can manage network performance on the eSight graphical user interfaces (GUIs).

Users can use the performance monitoring template to manage performance indexes, configure alarm thresholds, and apply performance collection rules to multiple objects. The performance monitoring template contains the following information:

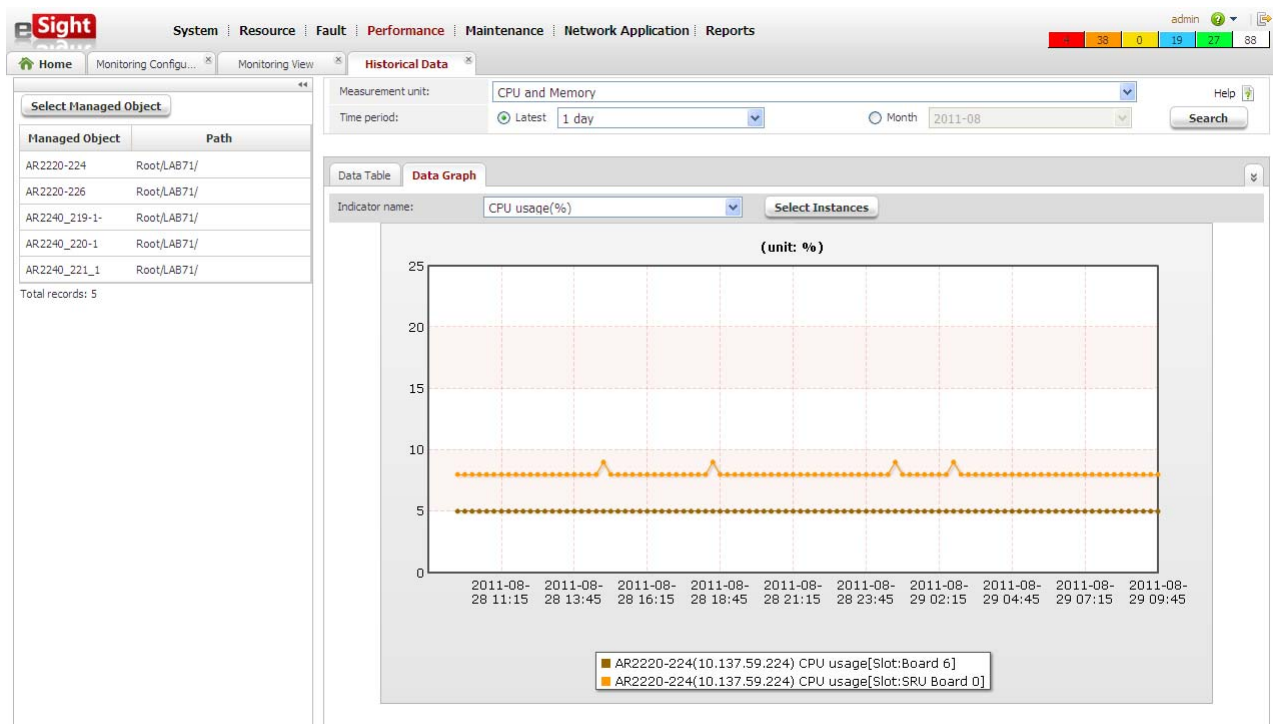
- **Performance index group**
Multiple performance indexes are integrated into a performance index group. Users can customize a performance index group to include all performance indexes in a scenario so that they can create a proper monitoring task.
- **Performance index**
The performance index defines the specific index for collecting performance.
- **Performance index collection period**
Multiple periods are available for collecting performance indexes.

- Performance threshold

A performance threshold is specified in the network. When the performance index is lower than the performance threshold, an alarm will be reported, instructing operators to take actions to prevent the network performance from deteriorating.

Performance data is collected during the performance monitoring process. As shown in [Figure 6-8](#), the performance data collected in a performance index collection period indicates the network performance in this period and provides a basis for predicting the network performance change.

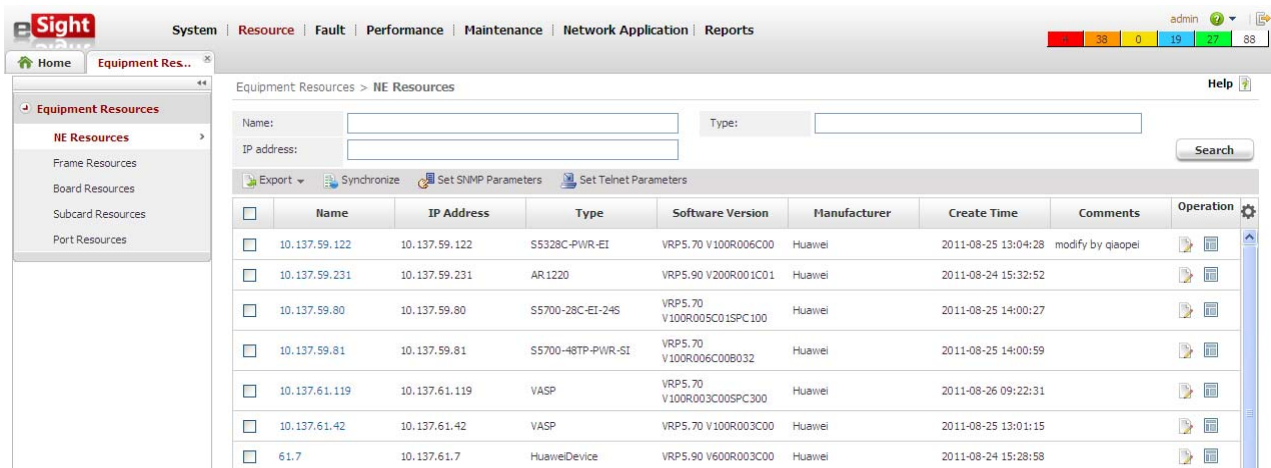
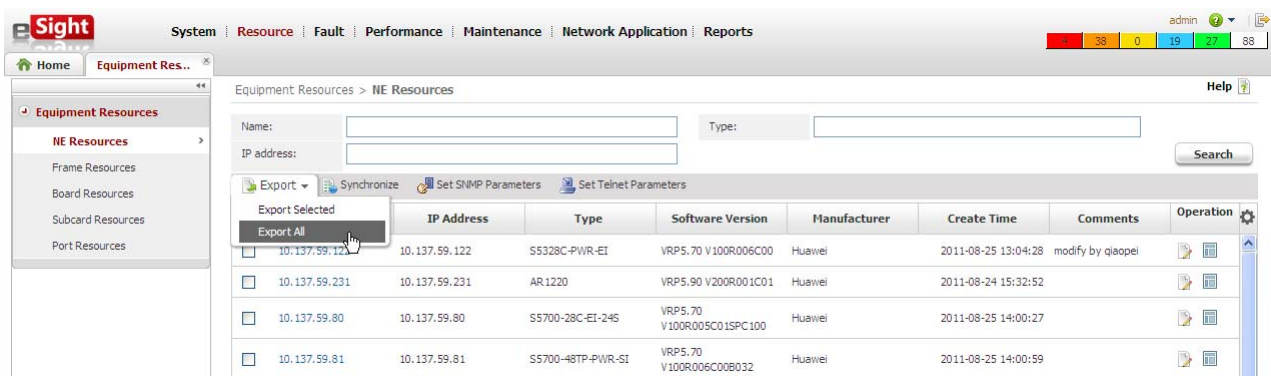
Figure 6-8 Monitoring network performance



Users can query the collected performance data displayed in GUI in the performance monitoring view to learn the network performance within a specified period and predict the network performance change.

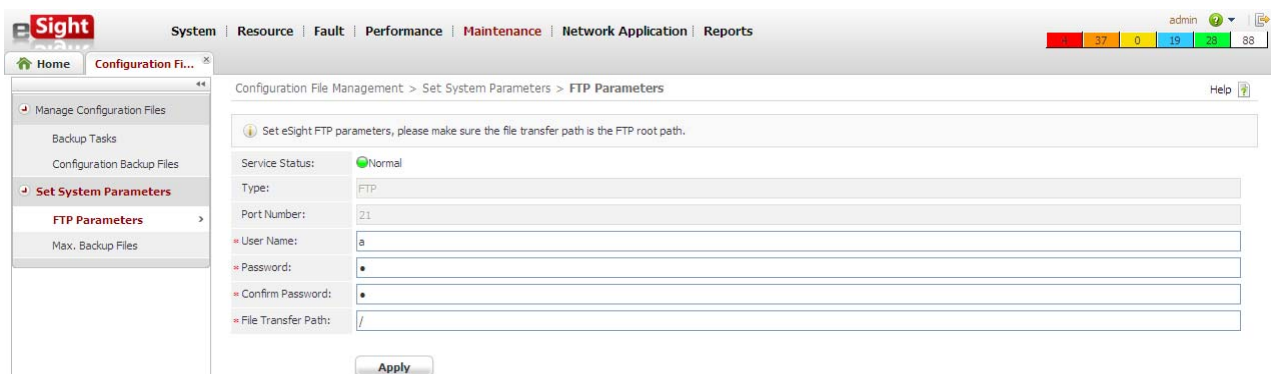
Querying Resources and Managing Reports

As shown in [Figure 6-9](#) and [Figure 6-10](#), the eSight provides various resources and predefined reports and the easy-to-use report design function so that users can design reports based on the industry features and OAM requirements.

Figure 6-9 Querying physical resources**Figure 6-10** Querying exported reports

Maintaining

As shown in [Figure 6-11](#) and [Figure 6-12](#), the eSight can manage configuration files to help users quickly save files and log in to the device. In addition, the eSight provides a tool to inspect devices periodically, lessening the workload of the maintenance personnel.

Figure 6-11 Managing configuration files

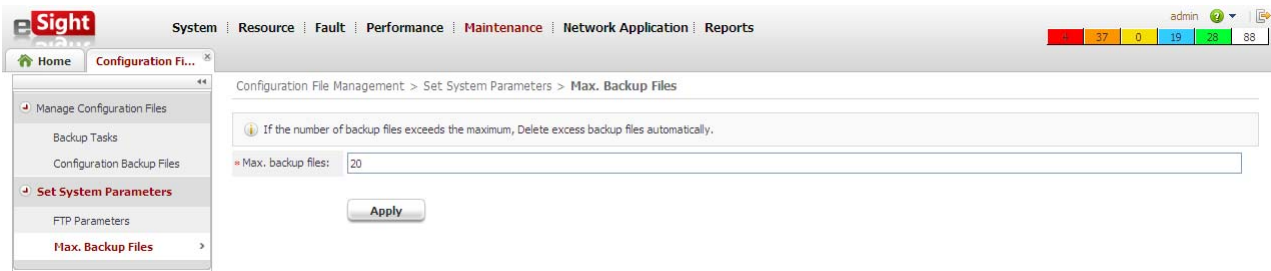
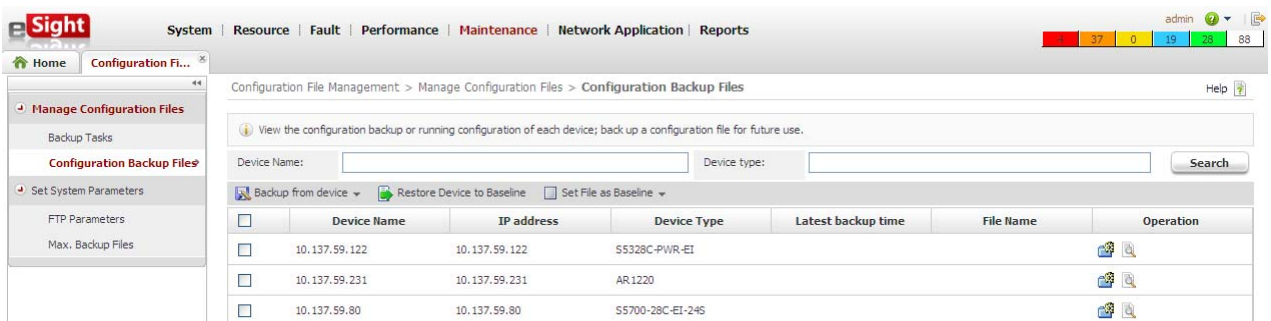
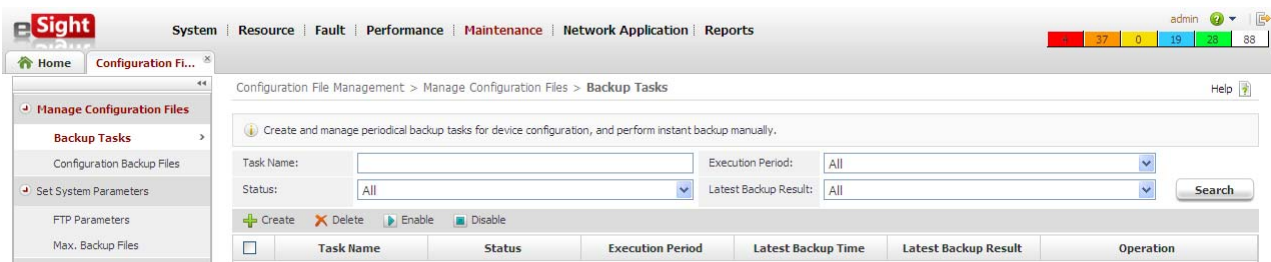


Figure 6-12 Backing up configuration files



6.1.2 Customization of Third-Party Devices

Overview

Network devices in a DC are from different manufactures and cannot be managed in a uniformly pre-integrated manner. Therefore, customization capabilities are required. If network devices are managed by their NMS, the maintenance cost will be higher and workload of the maintenance personnel will be heavier.

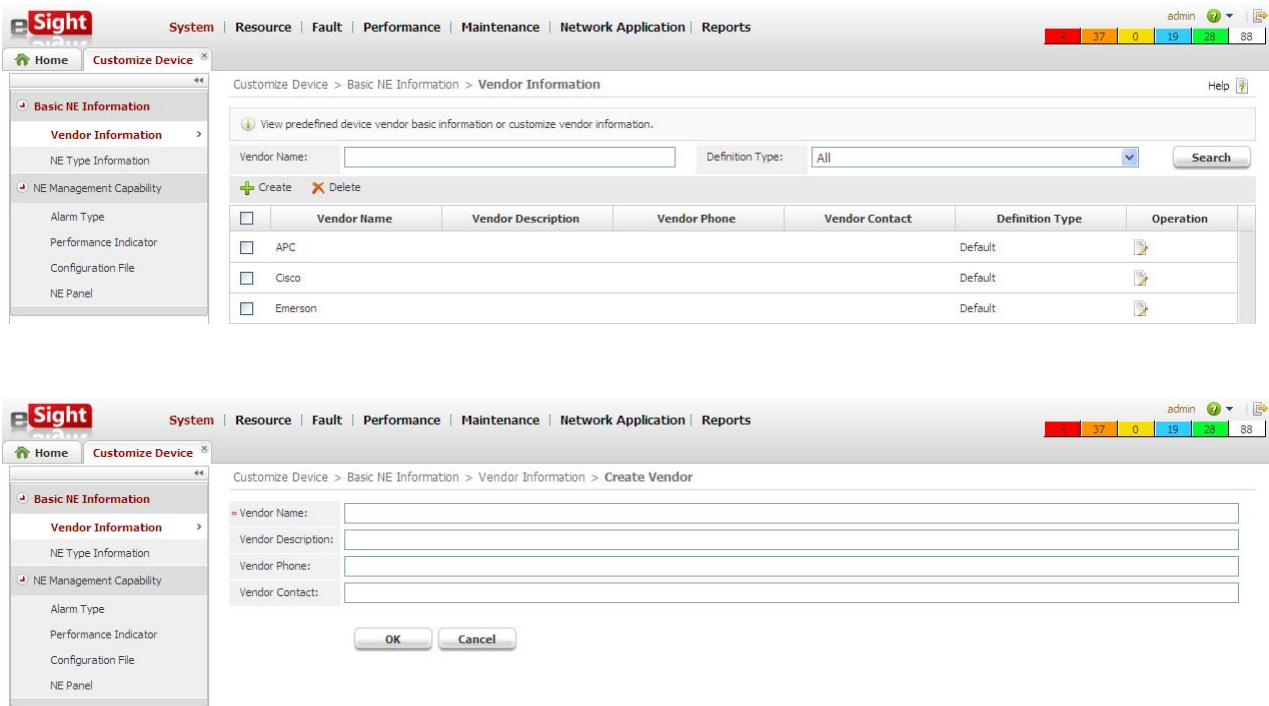
Huawei eSight provides customization capabilities for users to manage third-party devices. Users can configure the following information to manage third-party devices as required:

- Manufacturer
- Device model
- Alarm parameters
- Performance indexes
- Panel
- Configuration file management

Configuring Manufacturers

As shown in Figure 6-13, the eSight can configure the name and contact information of a manufacturer. The configured manufacturer information is used in the subsequent configuration of device models.

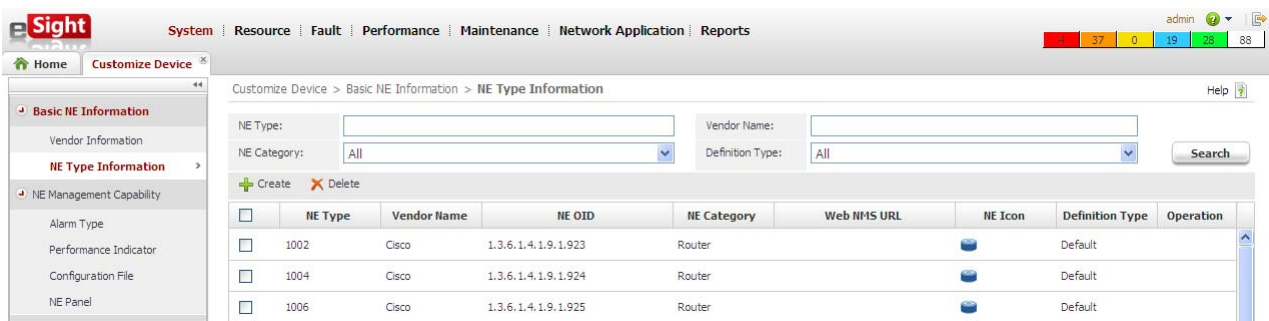
Figure 6-13 Configuring manufacturers



Configuring Device Models

As shown in Figure 6-14, the eSight can configure the description, icon, and Web link for a device model. The configured icon is displayed on the topology.

Figure 6-14 Configuring device models



Customize Device > Basic NE Information > NE Type Information > Create NE Type


NE Type:

Vendor Name:

NE OID:

NE Category:

Web NMS URL:

Current NE Icon: 

Example: S9312. Enter a user-defined NE type. A NE type indicates a device model.

Example: switch or router. Enter a user-defined device category.

Example: http://%IPAddress%/index.html. Enter the Web NMS address of a user-defined device. Access the required Web NMS, copy the address in the address box, paste it here, and change the IP address to %IPAddress%.

Customizing Alarms

As shown in Figure 6-15, the eSight can customize reported alarms. The customized alarms can be parsed and are displayed on the alarm management page.

Figure 6-15 Customizing alarms

Customize Device > NE Management Capability > Alarm Type

Alarm Name:

Vendor Name:

Alarm Severity:

SNMP Version:

Alarm Name	Vendor Name	Alarm Severity	SNMP Version	Notification Type	Alarm Index	Operation
No record.						

Customize Device > NE Management Capability > Alarm Type > Create Alarm Type

Vendor Name:

Notification Type:

Alarm Name:

SNMP Version:

Alarm Severity:

Event type:

Generic:

Enterprise ID:

Alarm OID:

Alarm Cause:

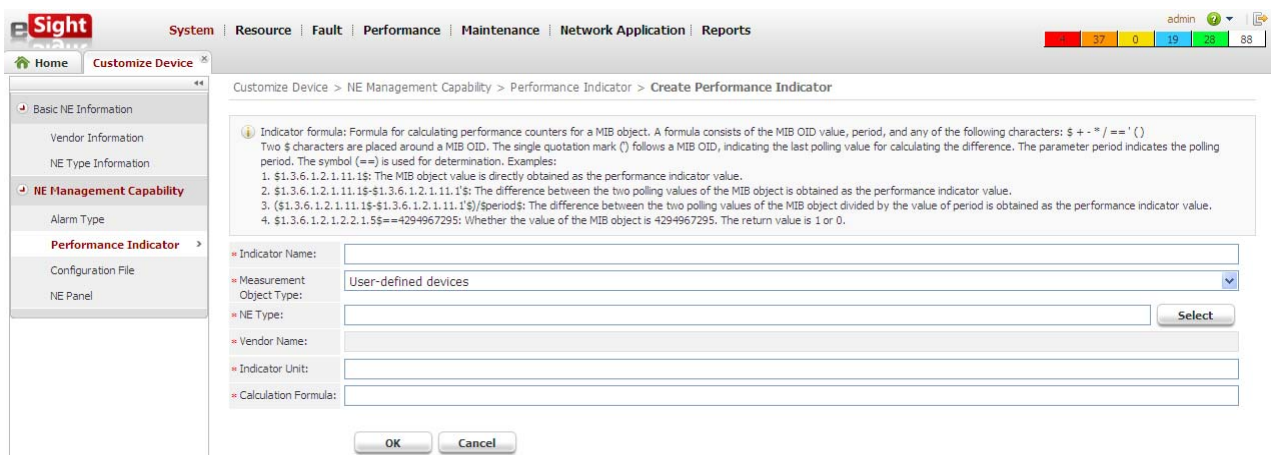
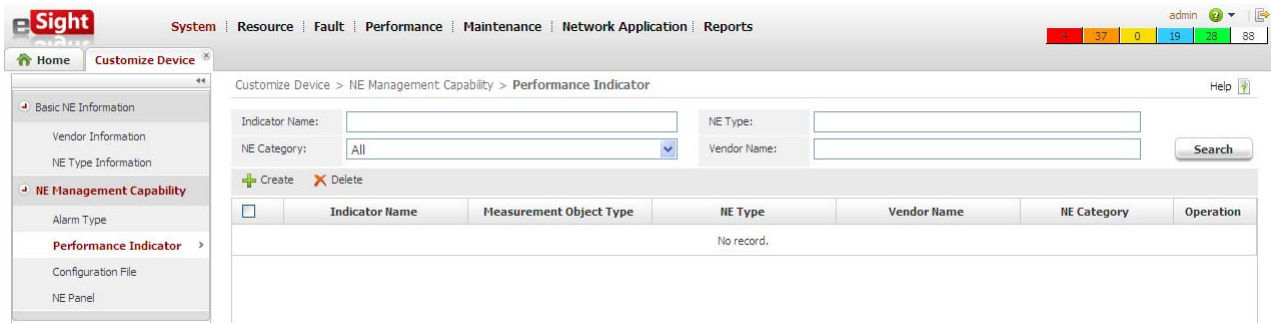
Clearance Suggestion:

Details:

Location Parameter Name	Location Parameter OID	Operation
No record.		

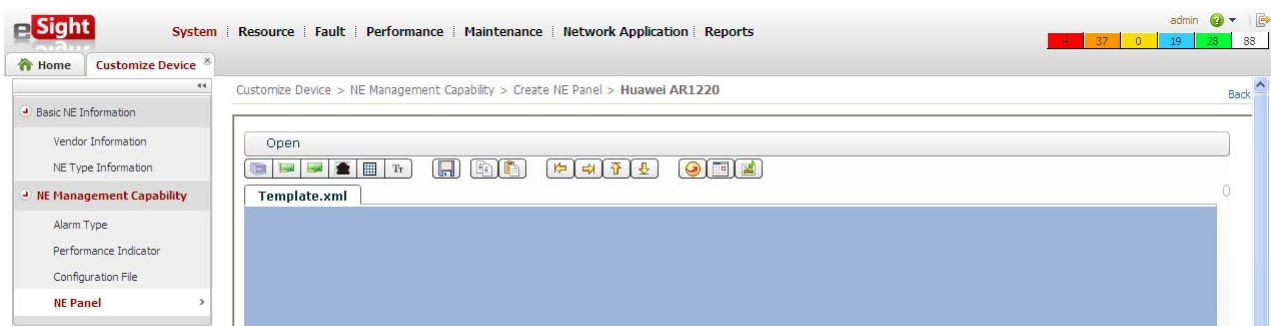
Customizing Performance Indexes

As shown in Figure 6-16, the eSight can customize performance indexes of devices. The customized performance indexes are collected by the performance statistics task and displayed on the performance page.

Figure 6-16 Customizing performance indexes

Customizing Device Panels

As shown in [Figure 6-17](#), the eSight can customize the simulation images of subracks, boards, subcards, and ports. The customized panel will display the new simulation images.

Figure 6-17 Customizing device panels

Customizing Configuration Files

As shown in [Figure 6-18](#), the eSight can customize commands to back up, restore, or restart configuration files so that configuration files can be automatically backed up.

Figure 6-18 Customizing backup and restoration of configuration files

The top screenshot shows the 'Customize Device' page in the eSight interface. The breadcrumb trail is 'Customize Device > NE Management Capability > Configuration File'. The left sidebar shows 'NE Management Capability' selected. The main area has input fields for 'NE Type', 'Vendor Name', 'NE Category' (set to 'All'), and 'Definition Type' (set to 'All'). Below these is a table with columns: 'NE Type', 'Vendor Name', 'NE Category', 'Definition Type', and 'Operation'.

NE Type	Vendor Name	NE Category	Definition Type	Operation
<input type="checkbox"/> CiscoDevice	Cisco	Unknown	Default	
<input type="checkbox"/> H3CDevice	H3C	Unknown	Default	
<input type="checkbox"/> HuaweiDevice	Huawei	Unknown	Default	
<input type="checkbox"/> WS6603	Huawei	Switch	Default	

The bottom screenshot shows the 'Create Configuration File' sub-tab. It has input fields for 'NE Type' (with a 'Select' button), 'Backup command:', 'Restore command:', and 'Restart command:'. A help text box on the right explains the command syntax and lists available variables: \$username, \$password, \$serviceIP, \$neFile, and \$networkFile.

Customizing Reports

The eSight can make report designs by modifying predefined report design files.

6.1.3 Software Upgrade and Patch Loading

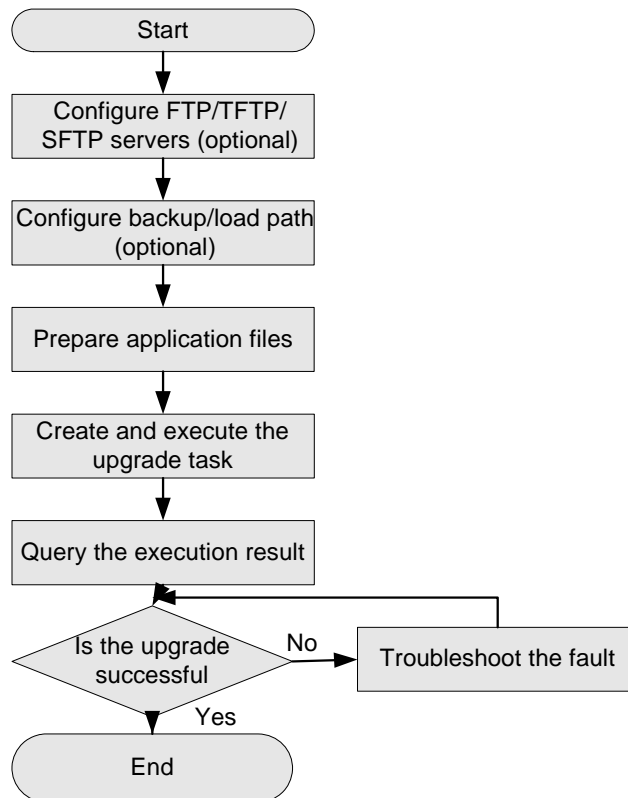
Overview

A DC has many network devices. Therefore, it is time consuming to upgrade software or load patches on these devices one by one and upgrade failures may occur due to human factors. Huawei recommends you upgrade software and load all patches remotely at one time. This method significantly lessens the workload of maintenance personnel and avoids failures caused by human factors.

Upgrading Software

The eSight provides a function to upgrade software remotely at one time. [Figure 6-19](#) displays the operation guide to upgrade devices. If the upgrade fails, the eSight provides troubleshooting methods to ensure that devices run in normal status.

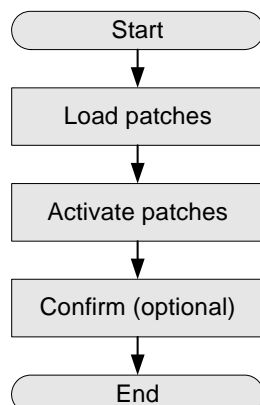
Figure 6-19 Software upgrade flowchart



Loading Patches

The eSight provides a function to load patches remotely at one time. [Figure 6-20](#) displays the operation flow to load patches. The eSight also provides the patch rollback function to restore the NE to the previous status.

Figure 6-20 Patch update flowchart



6.2 Troubleshooting

The DC network system consists of network devices, links between devices, and servers. If the network system is faulty, you can locate the fault by checking the link status, device status, or server status, or by detecting virus attacks. The upper layer application cannot work properly if any one of these components is faulty.

6.2.1 Troubleshooting Network Devices

Network devices may encounter the following faults:

- A device is down: The power indicator or other indicators on the device are off and no sound is generated.
- The CPU usage of a device is too high: The CPU usage is too high and related applications responds slowly when a user runs the monitoring software or logs in to the device.
- An error message is displayed: An error message is generated on the server when a user views the log server or logs in to the device.
- An alarm is reported: The status indicator of the device is red, indicating that an alarm is reported.

A Device Is Down

If a device is down, check the power cable and power supply in the equipment room first.

If the power cable is connected properly and the power supply is normal, call the device vendor or service provider for help immediately. If the hardware is faulty, ask the device vendor or service provider to replace parts as soon as possible.

The CPU Usage of a Device Is Too High

Report the problem to the service provider immediately. Help the technical support engineers to locate the cause. In most cases, the problem is caused by the virus attack.

An Error Message Is Displayed

Send the error message to the service provider and track the troubleshooting progress. The service provider will provide the cause to the problem after analyzing the error message. If the device has a potential fault, prepare an emergency trouble shooting scheme or replace the device.

An Alarm Is Reported

Send the alarm to the device vendor and service provider, and ask them to troubleshoot the fault or replace parts.

6.2.2 Troubleshooting Servers

Servers related to a network system are Dynamic Host Configuration Protocol (DHCP) server, access control system (ACS) server, and agent server on the external network. Faults that often occur are as follows:

- Failure to obtain an IP address.

- Failure to log in to the network device.
- Failure to access the Internet through an agent server.

Fail to Obtain an IP Address

To troubleshoot the fault, proceed as follows:

1. Perform the ping operation to check the connectivity of the DHCP server.
2. If the DHCP server is connected properly, log in to the DHCP server to check whether the DHCP service is normal. If it is normal, verify that the DHCP request times out due to the virus attack.
3. If the DHCP server fails, replace it with the backup server.
4. Configure a static IP address manually for the computer to access the network before the DHCP server recovers.

----End

Fail to Log In to the Network Device

1. Perform a ping operation to check the connectivity of the network device.
2. If the device is connected properly, log in to the ACS server to check whether the ACS service is normal.
3. If the ACS service is abnormal, log in to the network device through the console port, disable the authentication, authorization, and accounting (AAA) authentication, and enable the local authentication based on the built-in database.

----End

Fail to Access the Internet Through an Agent Server

1. Check the network connectivity by accessing other applications. Then, perform a ping operation to check the connectivity of the agent server.
2. If the agent server is connected correctly, log in to the agent server to check whether the agent service and related system services are normal. If any service is abnormal, restart the service or the agent server.
3. If the fault persists, check whether the key hardware (such as NIC) of the agent server is faulty.
4. If the hardware or system is faulty, replace the agent server with the backup agent server.
5. If there are no problems with the agent server, ping the DNS gateway and Internet service provider (ISP) gateway to check the Internet access. If the DNS gateway or ISP gateway cannot be pinged, contact the ISP to rectify the fault.
6. If the link provided by the ISP is faulty, access the Internet through the backup link.

----End

6.3 Network Expansion

Overview

With increasing expansion of services and scale of a DC, an existing network capacity cannot meet the requirements of long-term development. Therefore, network expansion is important. A smooth expansion is essential to the network expansion because services are not affected during the expansion.

The network expansion is implemented in three ways:

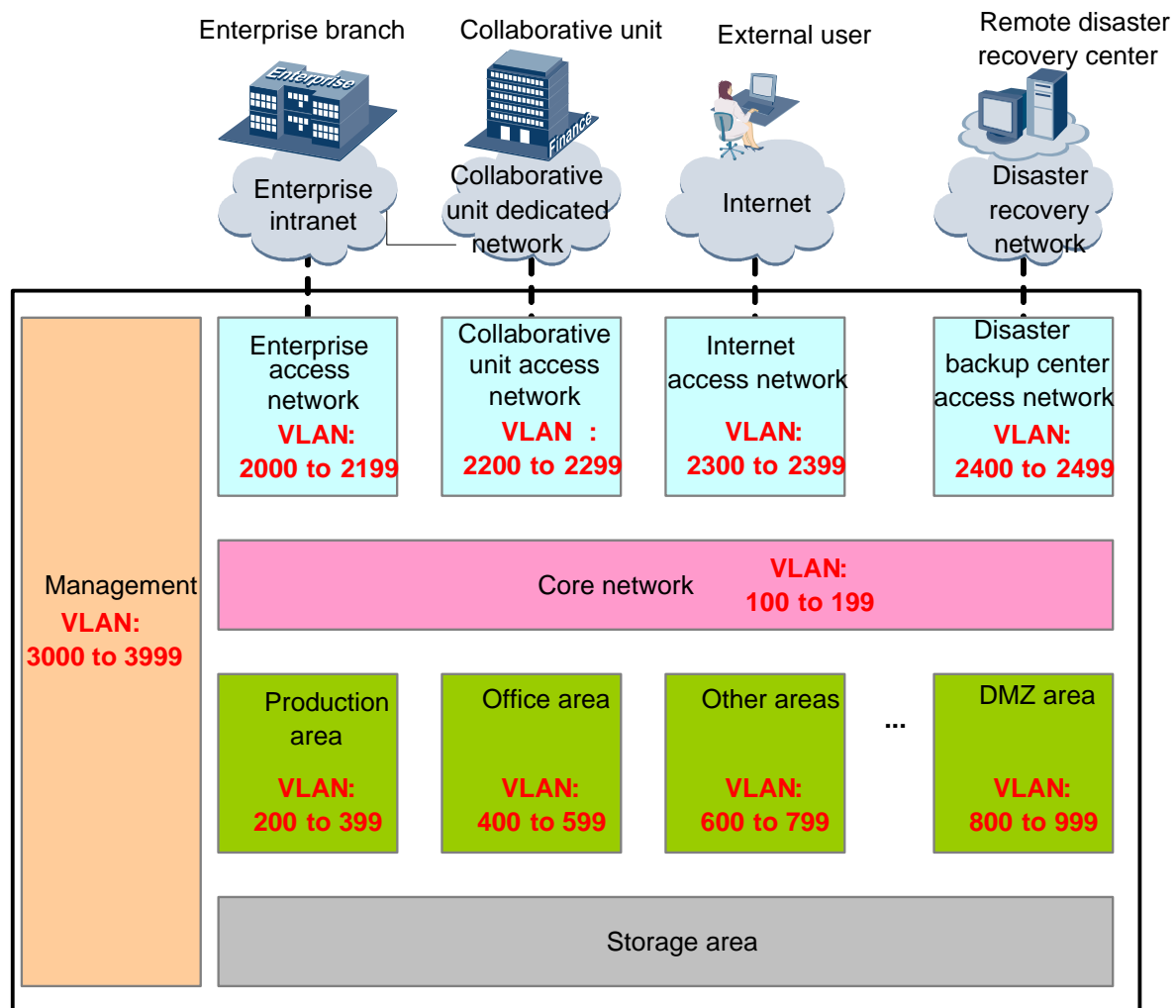
- Server expansion
- Device expansion
- Link bandwidth expansion

Use a proper expansion policy as required by expansion scenarios to expand the network capacity smoothly without affecting services.

Server Expansion

Server expansion is implemented by expanding servers in an original area or creating servers in a new area. The expansion policies of each are different.

Figure 6-21 Internal architecture of the DC



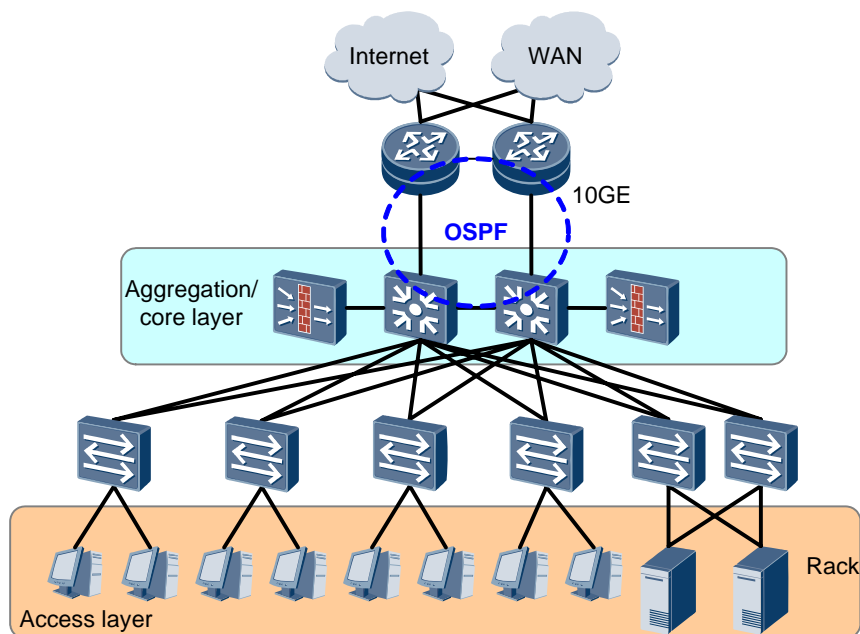
- Expand servers in an original area
With the development of production services, servers in the production area need to be expanded to meet service requirements. The servers must be expanded smoothly based on the previously planned VLANs and IP addresses, which ensures VLAN continuity, requires no change to the upstream router or firewall policy, facilitates the network maintenance, and relieves the expansion workload.
- Create servers in a new area
If the demilitarized zone (DMZ) is a newly created area, you need to allocate VLANs and IP addresses and plan a router and a firewall policy for this area. In this way, the existing services will be expanded smoothly without being affected, and the new area is easy to maintain and manage.

Device Expansion

Figure 6-22 shows common network architecture of a DC. Many ring networks exist at the access layer and aggregation layer. Once you add servers, you need to deploy routers at the access layer and connect them to the combined core layer, which makes the network more

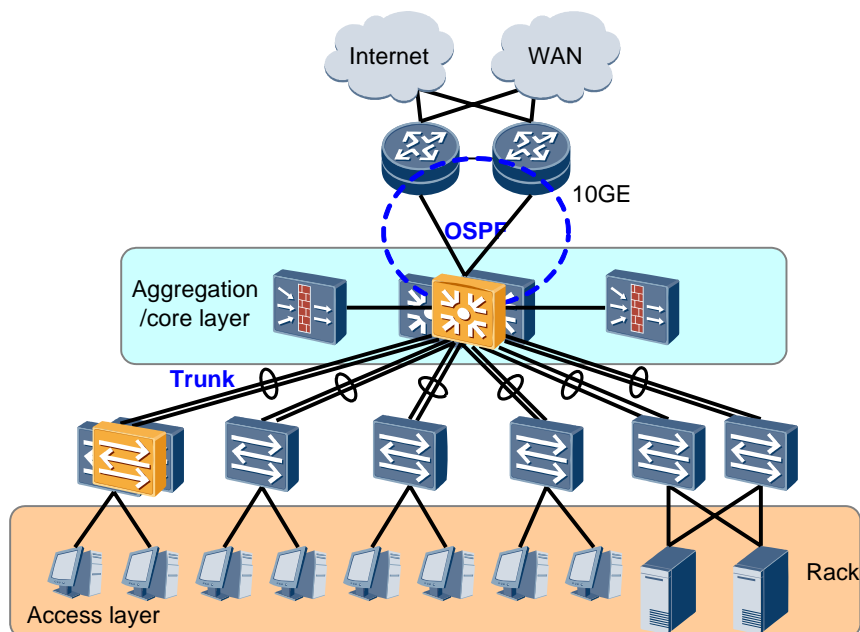
complex and requires a loop-prevention technology. Therefore, services on the existing network will be affected.

Figure 6-22 Common network architecture of a DC



To avoid affecting services while expanding, Huawei recommends cluster and stacking technologies in planning the network architecture of a DC, as shown in Figure 6-23. Cluster and stacking technologies tear down the loop prevention protocol, simplifies the network architecture, and facilitates the device expansion.

Figure 6-23 Network architecture of a DC deployed in the cluster and stacking mode



After the network is planned in the cluster and stacking mode, the network changes from a ring topology to a tree topology which is easy to maintain. When you expand devices, you only need to add new devices to the stacking system to implement smooth expansion, which has no impact on the network architecture and does not add physical links at the combined core layer.

Link Bandwidth Expansion

With the development of services, link bandwidth will increase and may become a bottleneck. You can use high-performance and high-bandwidth (for example, replace the GE board to 10GE board, or replace the 10GE board to 40GE board) boards or use the link binding technology to implement smooth expansion without affecting network services.

6.4 Disaster Emergency Maintenance

Overview

Disaster emergency maintenance requires designers to consider, during the network design, how to take emergency maintenance measures, how to recover services, and how to minimize service losses if a disaster occurs (such as an earthquake or fire) at a DC.

Suggestions on Disaster Emergency Maintenance

The three-center-in-two-area solution has taken unexpected disasters into consideration. For details about how to store data between the active center, backup center, and disaster recovery center and how to switch services if a disaster occurs, see [5 Suggestions on Planning Multiple DCs](#).

7 Recommended Products

The data center solution is made up of the following products:

- Core switch: S9300 series core switches
- Access switch: S6700 series access switches
- Access switch: S5700 series access switches
- Mini optical transport network (OTN): Optical OSN 1800
- Optical transmission platform: Optical OSN 6800

7.1 S9300 Series Core Switches

7.1.1 Product Overview

The Quidway S9300, which is a carrier-class core switch (S9300 for short), is a next-generation high-performance core routing switch developed by Huawei. The S9300 has a large switching capacity, a high port density, and can forward Layer 2 to Layer 4 packets at wire speed. In addition, the S9300 provides powerful multicast functions, a comprehensive QoS guarantee, an effective security management mechanism, and high reliability to meet the requirements of VIP users for multi-service, high reliability, large capacity, and modulation. This reduces costs in network construction and maintenance.

The S9300 can be deployed at the core and aggregation layers on various types of campus networks. It can also be used as an aggregation switch on some large campus networks that require high performance and port density.

7.1.2 Product Model

The S9300 series switches include the following models:

Table 7-1 S9300 product model

Product Model	Description
S9303	<ul style="list-style-type: none">• LPU: 3• Switch fabric capacity: 1440 Gbit/s• Backplane capacity: 3 Tbit/s• Forwarding performance: 540 Mpps
S9306	<ul style="list-style-type: none">• LPU: 6• Switch fabric capacity: 2 Tbit/s• Backplane capacity: 6 Tbit/s• Forwarding capacity: 1320 Mpps
S9312	<ul style="list-style-type: none">• LPU: 12• Switch fabric capacity: 2 Tbit/s• Backplane capacity: 12 Tbit/s• Forwarding capacity: 1320 Mpps

Figure 7-1 The S9303



Figure 7-2 The S9306



Figure 7-3 The S9312



7.1.3 Product Characteristics

Advanced Architecture, High Performance, and Flexible Configuration

The S9300 adopts the advanced and distributed architecture and the latest hardware forwarding engine technology. The services on all interfaces can be forwarded at wire speed, including IPv4 services, MPLS services, and Layer 2 forwarding services. The S9300 can use the ACL to forward packets at wire speed.

The hardware of the S9300 implements two-level packet replication to forward multicast packets at wire speed:

The SFU replicates multicast packets to the LPU.

Then the forwarding engine of the LPU replicates the multicast packets to the interfaces on the LPU.

The S9300 supports 2 Tbit/s switching capacity and various high-density boards to meet requirements for the large capacity and high-density interfaces of core and aggregation layer devices. It can meet users' increasing requirements for the bandwidth and protect and save the maximum amount of the users' investment.

Comprehensive Security Measures

The S9300 supports Authentication, Authorization, and Accounting (AAA). It performs AAA for access users based on policies. In addition, the S9300 supports 802.1x, portal, guest VLAN, and dynamic user access authentication. Therefore, it can work well with the network admission control (NAC) produced by other mainstream manufacturers.

The S9300 supports the routing protocol encryption, lawful interception, MAC address filtering, dynamic ARP detection, and ACLs to protect data for service providers and end users. Hardware-based packet filtering and sampling guarantee high performance and high scalability of the system.

The S9300 is the industry leader in integrated security solutions. It uses a 2-level CPU protection mechanism and supports 1K CPU queues, and protects the CPU by separating the data plane and control plane. In addition, the S9300 defends against DoS attacks, prevents unauthorized access, and prevents control plane overloading.

High Reliability

Huawei's carrier-class high reliability design ensures that the S9300 is 99.999% reliable, which meets and exceeds carrier-class operation requirements. The S9300 provides redundant backup for key components, including MPUs, power supply units, and fans, all of which are hot swappable. Based on distributed hardware forwarding architecture, the routing plane is separated from the switching plane to ensure service continuity.

The S9300 provides 3.3 ms hardware-based Ethernet operation, administration, and maintenance (OAM) function, which can quickly detect and locate faults. By using the Ethernet OAM technology and switchover technologies, the S9300 can provide millisecond-level protection for networks.

The service traffic can be switched between active and standby components without rebooting the equipment. The S9300 also supports the in-service software upgrade (ISSU), further reducing service interruption.

The S9300 supports the link aggregation defined in IEEE 802.3ad, the IEEE 802.1s/w standard, and Virtual Router Redundancy Protocol (VRRP). In addition, it supports various millisecond switchover technologies, such as Rapid Ring Protection Protocol (RRPP), Smart Link, IP fast reroute (FRR), traffic engineering (TE) FRR, and virtual private network (VPN) FRR. These features improve the reliability of data transmission.

7.1.4 Specifications

The following table lists the specifications of the S9300 series switches.

Table 7-2 Main specifications of the S9300 series switches

Specifications	S9303	S9306	S9312
Backplane capacity (Tbit/s)	1.2	2.4	4.8
Service slot	3	6	12
GE port density	144	288	576
10G port density	120	240	480
VLAN	<ul style="list-style-type: none">• Access, trunk, and hybrid interfaces• Default VLAN• VLAN switching• QinQ and enhanced selective QinQ		
MAC address	<ul style="list-style-type: none">• Automatic learning and aging of MAC addresses• Static, dynamic, and blackhole MAC address entries• Packet filtering based on source MAC addresses• MAC address learning limitation based on interfaces and VLANs		
STP	<ul style="list-style-type: none">• STP, RSTP, and MSTP• Bridge protocol data unit (BPDU) protection, root protection, and loop protection• BPDU tunnels		

Specifications	S9303	S9306	S9312
IP routing	<ul style="list-style-type: none"> • IPv4 dynamic routing protocols, such as, RIP, OSPF, IS-IS, and BGP • IPv6 dynamic routing protocols, such as, RIPng, OSPFv3, ISISv6, and BGPv4 		
Multicast	<ul style="list-style-type: none"> • IGMP snooping • IGMP fast leave • Multicast traffic control • Multicast queries • Suppression on multicast packets • Multicast ACL 		
MPLS	<ul style="list-style-type: none"> • Basic MPLS functions • MPLS OAM • MPLS traffic engineering (TE) • MPLS VPN, VLL, and VPLS 		
Clock	<ul style="list-style-type: none"> • Synchronous Ethernet clock • IEEE 1588v2 		
QoS	<ul style="list-style-type: none"> • Traffic classification based on the Layer 2 protocol header, Layer 3 protocol, Layer 4 protocol, and 802.1p priority • Actions such as ACL, CAR, remark, and schedule • Queue scheduling styles such as PQ, WRR, DRR, PQ+WRR, and PQ+DRR • Congestion avoidance mechanisms such as Weighted Random Early Detection (WRED) and tail drop • Traffic shaping 		
Configuration and maintenance	<ul style="list-style-type: none"> • Terminal services such as Console, Telnet, and SSH • Network management protocols such as SNMPv1/v2/v3 • Uploading and downloading of files using FTP and TFTP • BootROM upgrade and remote online upgrade • Hot patches • User operation logs 		
Security and management	<ul style="list-style-type: none"> • 802.1x authentication and portal authentication • RADIUS and HWTACACS authentication for login users • Hierarchical protection for commands to prevent unauthorized users from accessing the device • Protection against DoS attacks, SYN flood attacks of TCP, UDP flood attacks, broadcast storms, and large-traffic attacks • CPU channel protection • Ping and traceroute • RMON 		

Specifications	S9303	S9306	S9312
Chassis dimension (H x W x D)	442 mm x 476 mm x 175 mm	442 mm x 476 mm x 441.7 mm	442 mm x 476 mm x 663.95 mm
Chassis weight (empty)	< 22 kg	< 42 kg	< 70 kg
Working voltage	DC power supply: -38.4 V to -72 V AC power supply: 90 V to 264 V		
Typical power consumption	180 W	< 350 W	< 650 W
Power supply capability of the device (PoE not included)	800 W	1600 W	1600 W

7.2 S6700 Series Access Switches

7.2.1 Product Overview

The Quidway S6700 (S6700) is a next-generation 10GE box-shaped switch developed by Huawei. The S6700 can serve as access switches in the data center to access the 10GE server, aggregation switches on a metropolitan area network (MAN), and core switches on a campus network.

As one of the class-A switches in the industry, the S6700 provides a maximum of 24 or 48 10GE interfaces at wire speed, which enables the high-density 10GE access and high-density 10GE aggregation on the campus network. The S6700 provides rich service features, a comprehensive security control policy, and various QoS mechanisms to meet the requirements for extensibility, reliability, manageability, and security of the data center.

7.2.2 Product Model

The S6700 series switches include two models.

- S6748-EI: provides 48 GE small form-factor pluggable (SFP)/10GE small form-factor pluggable plus (SFP+) ports, two slots for power supplies, and a USB port.
- S6724-EI: provides 24 GE SFP/10GE SFP+ ports, two slots for power supplies, and a USB port.

Figure 7-4 The S6748-EI



Figure 7-5 The S6724-EI



7.2.3 Product Characteristics

High-Density 10GE Flexible Access

With the increasing bandwidth required by the clients, the 10GE network interface cards on the server are widely used. The switch in the data center provides higher forwarding performance and 10GE interface extensibility. Compared with other similar switches in the industry, the S6700 box-shape switch has the highest 10GE port density and largest switching capacity. An S6700 can support packet forwarding at wire speed on a maximum of 48 10GE interfaces.

The GE/10GE interfaces support flexible access and can automatically identify the type of an installed optical module. The S6700 can access the optical/electrical interfaces on the GE server. This saves the users' investments and ensures flexible usage of the S6700.

To meet the requirements for heavy traffic and non-blocking transmission, the S6700 provides large buffer capacity and uses advanced buffer scheduling mechanisms to maximize the effective usage of buffer capacity.

Comprehensive Security Measures

The S6700 provides various security measures. It can defend against Denial of Service (DoS) attacks, attacks to networks, and attacks to users. DoS attacks include SYN Flood attacks, Land attacks, Smurf attacks, and ICMP Flood attacks. Attacks to networks refer to STP BPDU/root attacks. Attacks to users include bogus DHCP server attacks, man-in-the-middle attacks, IP/MAC spoofing attacks, DHCP request flood attacks, and DoS attacks by changing the CHADDR field of packets.

The S6700 listens to information about the MAC or IP address of an access user, IP address lease, VLAN ID, and interface by establishing and maintaining a DHCP snooping binding table. The S6700 directly discards invalid packets such as ARP spoofing packets and packets with bogus IP addresses that do not match binding entries. In this manner, hackers or attackers are prevented from carrying out the man-in-the-middle attacks by using ARP packets on campus networks. The trusted interface feature of DHCP snooping ensures the validity of the DHCP server.

The S6700 supports strict learning of ARP entries to prevent ARP spoofing attackers from exhausting ARP entries so that authorized users can access the Internet. The S6700 supports IP source check to prevent DoS attacks caused by MAC address spoofing, IP address spoofing, and MAC/IP spoofing. Unicast reverse path forwarding (URPF) provided by the S6700 can reverse check packet transmission path to authenticate packets, which can protect the network against increasing source address spoofing attacks.

The S6700 supports the integrated MAC address authentication and 802.1x authentication. User information, such as the user name, IP address, MAC address, VLAN ID, access interface, and a flag indicating whether antivirus software is installed on the client, can be bound statically or dynamically, and policies (VLAN, QoS, and ACL) can be delivered dynamically.

The S6700 can limit the number of MAC addresses learned on an interface to prevent attackers from exhausting MAC address entries by using bogus source MAC addresses. In this way, MAC addresses of authorized users can be learned and flooding is prevented.

High Reliability

The S6700 supports dual power supplies for backup and can use an AC power supply and a DC power supply at the same time. Users can select a single power supply or dual power supplies to improve device reliability. The switch provides two built-in fans to improve operating stability and has a long mean time between failure (MTBF).

Enhancing STP, RSTP, and MSTP, the S6700 supports the MSTP multi-process that greatly increases the number of sub-ring instances. It supports enhanced Ethernet technologies such as Smart Link and RRPP to implement millisecond-level protective switchover, improving network reliability. Smart Link and RRPP both support multi-instance to implement load balancing among links, further improving bandwidth usage.

The S6700 supports enhanced trunk (E-Trunk). When a client edge (CE) is dual homed to a VPLS, VLL, or PWE3 network, an E-Trunk can be configured to protect the links between the CEs and provider edges (PEs) and implement backup between PEs. The E-trunk can implement link aggregation across devices to upgrade the link reliability to device level.

The S6700 supports Smart Ethernet Protection (SEP) protocol, a ring network protocol applied to the link layer of an Ethernet network. SEP is applicable to open ring networks and can be deployed on upper-layer aggregation devices to provide millisecond-level switchover without interrupting services. Huawei devices have implemented Ethernet link management using SEP. SEP features simplicity, high reliability, high switchover performance, convenient maintenance, and flexible topology and enables users to conveniently manage and plan networks.

The S6700 supports VRRP to keep the communication continuity and reliability, ensuring a stable network. Multiple equal-cost routes can be configured on the S6700 to implement route redundancy. When the active uplink route is faulty, traffic is automatically switched to a backup route. This feature implements multi-level backup for uplink routes.

Rich QoS Capabilities

The S6700 can implement complex traffic classification based on information such as the 5-tuple, IP preference, ToS, DSCP, IP protocol type, ICMP type, TCP source port, VLAN, the protocol type of an Ethernet frame, and CoS. The S6700 supports inbound and outbound ACLs. The S6700 supports the flow-based two-rate and three-color CAR. Each interface supports eight priority queues, multiple queue scheduling algorithms such as WRR, DRR, SP, WRR+SP, and DRR+SP, and WRED congestion avoidance mechanism, which ensures the quality of network services such as voice, video and data services.

High Extensibility

The S6700 supports long-distance intelligence stacking (iStack). A common interface can be configured as a stack interface at the CLI, enabling flexible interface usage. The optical fibers can be used for stacking, greatly increasing the distance between stacked devices. Compared with a single device, intelligent stacking features powerful extensibility, reliability, and performance.

When customers need to expand the device or replace a single faulty device, they can add new devices without interrupting services. Compared with chassis switches, the performance and port density of intelligent stacking are not restricted by the hardware architecture. Multiple stacked devices can be considered as a logical device, which simplifies the network management and configuration.

Convenient Operation and Maintenance

The S6700 supports automatic configuration, plug-and-play, deployment from USB devices, and batch remote upgrade. Upgrade and delivery of the S6700 can be completed at one time, which simplifies management and maintenance. Maintenance costs are greatly reduced.

The S6700 supports diversified management and maintenance modes such as SNMPv1/v2/v3, CLI, Web network management, Telnet, and Huawei Group Multicast Protocol (HGMP), which makes device management more flexible. In addition, the S6700 supports NTP, SSHv2.0, TACACS+, RMON, multi-log host, interface-based traffic statistics, and NQA, which helps to better deploy and adjust networks.

The S6700 supports the GARP VLAN Registration Protocol (GVRP). The GVRP technology implements dynamic configuration of VLANs. In a complicated networking environment, GVRP can simplify VLAN configuration and reduce network communication faults caused by incorrect configuration of VLANs. This reduces the manual configurations of network managers and ensures correct VLAN configurations.

The S6700 supports MUX VLAN. The MUX VLAN function is used to isolate Layer 2 traffic between interfaces on a VLAN. Subordinate VLANs can communicate with the MUX VLAN but cannot communicate with each other. MUX VLAN is usually used on enterprise intranets. With this function, a user interface can communicate with a server interface but cannot communicate with other user interfaces. MUX VLAN prevents communication between network devices connected to some interfaces or interface groups but allows these devices to communicate with the default gateway. This function ensures resource sharing and secure communication in an enterprise.

The S6700 supports BFD and provides millisecond-level detection for protocols such as OSPF, IS-IS, VRRP, and PIM to improve network reliability. Complying with IEEE 802.3ah and 802.1ag, the S6700 supports point-to-point Ethernet fault management. It can detect faults on user links. Ethernet OAM improves the network management and maintenance capabilities on the Ethernet and ensures a stable network.

Rich IPv6 Features

The S6700 supports IPv4/IPv6 protocol stack and can be smoothly upgraded. The S6700 hardware supports the IPv4/IPv6 protocol stack, IPv6 over IPv4 tunnels (including manual tunnels, 6to4 tunnels, and ISATAP tunnels), and Layer 3 wire-speed forwarding. Therefore, the S6700 can be deployed on IPv4 networks, IPv6 networks, and networks that run IPv4 and IPv6 simultaneously. This makes the networking flexible and meets the requirements for the network transition from IPv4 to IPv6.

The S6700 supports various IPv6 routing protocols including RIPng and OSPFv3. It uses the IPv6 Neighbor Discovery Protocol (NDP) to manage packets exchanged between neighbors. It also provides the Path MTU Discovery (PMTU) mechanism to select a proper MTU on the path from the source to the destination, optimizing network resources and obtaining the maximum throughput.

7.2.4 Main Specifications

Table 7-3 Main specifications of the S6700 series products

Item	S6724-EI	S6748-EI
Port description	24 GE SFP/10GE SFP+ ports	48 GE SFP/10GE SFP+ ports
Forwarding performance (PPS)	358 Mbit/s	715 Mbit/s
Interface switching capacity (bit/s)	480 Gbit/s	960 Gbit/s
MAC address table	<ul style="list-style-type: none">• Capacity of 128K MAC addresses• Automatic learning and aging of MAC addresses• Static, dynamic, and blackhole MAC address entries• Packet filtering based on source MAC addresses	
VLAN	<ul style="list-style-type: none">• 4K VLANs• Guest VLANs and voice VLANs• VLANs based on MAC addresses, protocols, IP subnets, policies, and interfaces.• 1:1 and N:1 VLAN switching• QinQ and selective QinQ	
IPv4 route	<ul style="list-style-type: none">• Static route, RIPv1, RIPv2, ECMP, and URPF• OSPF, IS-IS, and BGP• VRRP• Policy-based routing• Routing policy	
IPv6 route	<ul style="list-style-type: none">• Static route• RIPng• Manual tunnels• Six-to-four tunnels• ISTAP tunnels	

Item	S6724-EI	S6748-EI
IPv6 features	<ul style="list-style-type: none"> • Neighbor Discovery (ND) • PMTU • IPv6 Ping, IPv6 Tracert, and IPv6 Telnet • Six-to-four tunnels, ISATAP tunnels, and manually configured tunnels • ACLs based on the source IPv6 address, destination IPv6 address, Layer 4 interface, or protocol type • MLDv1/v2 snooping 	
Multicast	<ul style="list-style-type: none"> • Static Layer 2 multicast MAC address • MAC address-based multicast forwarding • IGMP snooping and IGMP fast leave • Multicast VLAN • MLD snooping • IGMP proxy • Controllable multicast • Interface-based multicast traffic statistics • IGMP v1/v2/v3 • PIM-SM, PIM-DM, and PIM-SSM • MSDP 	
QoS/ACL	<ul style="list-style-type: none"> • Rate limit on packets sent and received by an interface • Packet redirection • Port-based traffic policing and two-rate and three-color CAR • Eight queues on each port • WRR, DRR, SP, WRR+SP, and DRR+SP queue scheduling algorithms • WRED • Re-marking of the 802.1p priority and DSCP priority of packets • Packet filtering on Layer 2 to Layer 4, filtering out invalid frames based on the source MAC address, destination MAC address, source IP address, destination IP address, port number, protocol, and VLAN ID • Queue-based rate limit and port-based traffic shaping 	
Reliability	<ul style="list-style-type: none"> • STP, RSTP, and MSTP • BPDU protection, root protection, and loop protection • RRPP topology and RRPP multi-instance • Smart Link tree topology, Smart Link multi-instance, and the millisecond-level protection • SEP • BFD for OSPF, IS-IS, VRRP, and PIM • Enhanced trunk (E-trunk) 	

Item	S6724-EI	S6748-EI
Security	<ul style="list-style-type: none"> • Hierarchical user management and password protection • DoS attack defense, ARP attack defense, and ICMP attack defense • Binding of the IP address, MAC address, interface, and VLAN • Interface isolation, interface security, and sticky MAC addresses • Blackhole MAC addresses • Limit on the number of learned MAC addresses • IEEE 802.1x authentication and limit on the number of users on an interface • Multiple authentication methods including AAA, RADIUS, TACACS+, and NAC authentication • SSH v2.0 • Hypertext Transfer Protocol Secure (HTTPS) • CPU protection • Blacklist and whitelist 	
Management and maintenance	<ul style="list-style-type: none"> • Stack function on service interfaces • MAC forced forwarding (MFF) • Virtual cable detection (VCT) • Ethernet OAM (IEEE 802.3ah and 802.1ag) • Local port mirroring, remote switched port analyzer (RSPAN) and the packet forwarding on observing ports • Remote configuration and maintenance using Telnet • SNMPv1/v2/v3 • RMON • Network management system (NMS) and Web NMS • HGMP • System logs and multi-level alarms • GVRP • MUX VLAN • 802.3az Energy Efficient Ethernet (EEE) 	
Working environment	<ul style="list-style-type: none"> • Working temperature: 0°C to 45°C (long term); -5°C to 50°C (short term); relative humidity: 10% to 90% (non-condensing) 	
Input voltage	AC power supply <ul style="list-style-type: none"> • Rated voltage: 100 V to 240 V, 50/60 Hz • Maximum voltage: 90 V to 264 V, 50/60 Hz 	
Dimensions (H x W x D)	43.6 mm x 442 mm x 420 mm	
Power consumption	165 W	237 W

7.3 S5700 Series Access Switches

7.3.1 Product Overview




The Quidway S5700 (S5700) is a next-generation GE switch developed by Huawei to meet the requirements for high-bandwidth access and Ethernet multi-service aggregation, providing powerful Ethernet functions for carriers and enterprise customers. Based on the next-generation high-performance hardware and Huawei Versatile Routing Platform (VRP) software, the S5700 features large capacity and high-density GE interfaces, and provides 10 Gbit/s uplinks for customers. The S5700 can meet the requirements of multiple scenarios such as service aggregation on campus networks and enterprise networks, GE access to IDC, and the GE desktop access to the enterprise network.






The S5700 is a box-shaped device with a chassis of 1 U high, providing a limited version (LI), a standard version (SI), an enhanced version (EI), and an advanced version (HI). LI provides various Layer 2 functions while SI supports Layer 2 functions and basic Layer 3 functions. EI supports all routing protocols and service features. In addition to the functions of EI, HI supports some advanced functions such as MPLS and hardware OAM.







7.3.2 Appearance



The following table lists models of the S5700.

Table 7-4 Models of S5700

Model	Appearance	Description
S5706TP-LI		<ul style="list-style-type: none">• Four 10/100/1000Base-T ports• Two 1000 Mbit/s combo ports• AC power supply
S5724TP-SI		<ul style="list-style-type: none">• 20 10/100/1000Base-T ports• Four 100/1000Base-X 1000M combo ports• AC/DC power supply• RPS 12 V power supply backup• USB port
S5724TP-PW R-SI		<ul style="list-style-type: none">• 20 10/100/1000Base-T ports• Four 100/1000Base-X 1000M combo ports• Pluggable dual AC power supplies• PoE• USB port

Model	Appearance	Description
S5748TP-SI		<ul style="list-style-type: none"> • 44 10/100/1000Base-T ports • Four 100/1000Base-X 1000M combo ports • AC/DC power supply • RPS 12 V power supply backup • USB port
S5748TP-PWR-SI		<ul style="list-style-type: none"> • 44 10/100/1000Base-T ports • Four 100/1000Base-X 1000M combo ports • AC power supply • PoE • USB port
S5728C-SI		<ul style="list-style-type: none"> • 24 10/100/1000Base-T ports • Four 100/1000Base-X 1000M combo ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable power supplies • USB port
S5728C-PWR-SI		<ul style="list-style-type: none"> • 24 10/100/1000Base-T ports • Four 100/1000Base-X 1000M combo ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable AC power supplies • PoE • USB port
S5752C-SI		<ul style="list-style-type: none"> • 48 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable power supplies • USB port

Model	Appearance	Description
S5752C-PW R-SI		<ul style="list-style-type: none"> • 48 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable AC power supplies • PoE • USB port
S5728C-EI		<ul style="list-style-type: none"> • 24 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable power supplies
S5728C-PW R-EI		<ul style="list-style-type: none"> • 24 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, or two 10GE SFP+ subcards • Dual pluggable AC power supplies • PoE
S5728C-EI-2 4S		<ul style="list-style-type: none"> • 24 100/1000Base-X ports • Four 10/100/1000Base-T GE combo ports, two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable power supplies
S5752C-EI		<ul style="list-style-type: none"> • 48 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards • Dual pluggable power supplies
S5752C-PW R-EI		<ul style="list-style-type: none"> • 48 10/100/1000Base-T ports • Two 10GE XFP uplink ports, four 1000Base-X SFP uplink ports, or two 10GE SFP+ subcards • Dual pluggable AC power supplies • PoE

Model	Appearance	Description
S5728C-HI		<ul style="list-style-type: none">• 24 10/100/1000Base-T ports• Four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards• Dual pluggable power supplies
S5728C-HI-24S		<ul style="list-style-type: none">• 24 100/1000Base-X ports• Four 1000Base-X SFP uplink ports, two 10GE SFP+ uplink ports, or four 10GE SFP+ subcards• Dual pluggable power supplies

7.3.3 Product Characteristics

High Extensibility

The S5700 supports intelligent stacking (iStack). Multiple S5700s constructs a virtual switch automatically after being connected by stacking cables.

Compared with a single device, intelligent stacking features powerful extensibility, reliability, and performance. When customers need to expand the device or replace a single faulty device, they can add new devices without interrupting services. Compared with chassis switches, the performance and port density of intelligent stacking are not restricted by the hardware architecture. Multiple stacked devices can be considered as a logical device, which simplifies the network management and configuration.

Powerful Service Support

The S5700 supports the enhanced selective QinQ to add outer VLAN tags to packets, without occupying ACL resources, which meets requirements for multi-service provisioning.

The S5700 supports IGMPv1/v2/v3, IGMP snooping, IGMP filter, IGMP fast leave, and IGMP proxy. It supports wire-speed multicast VLAN replication, multicast load balancing in an Eth-Trunk, and controllable multicast. These multicast features provide high-quality video services for users.

The S5700 supports multi-VPN-instance CE (MCE) to isolate users on different VPNs on a device, ensuring the user's data security and saving the user's investments.

The S5700HI switches are cost-effective box-shaped MPLS switches. They support basic MPLS and VLL functions and can be used as high-quality access devices to provide leased line services for enterprises. The S5700HI can help customers to construct an MPLS edge network.

The S5700 provides multiple devices that support PoE and comply with IEEE802.3af and 802.3at (POE+) standards. By using the Ethernet, the S5700 can supply power to standard PD devices such as the IP Phone, WLAN AP, and Bluetooth AP. Each interface provides 30 W power. This reduces the power cable layout and management cost for terminal devices. The S5700 can also be configured to provide power for PDs at specified times as required.

High Reliability

The S5700 supports dual power supplies for backup and can use an AC power supply and a DC power supply at the same time. Users can select a single power supply or dual power supplies to improve device reliability. The switch provides three built-in fans to improve stability and has a long MTBF.

Enhancing STP, RSTP, and MSTP, the S5700 supports the MSTP multi-process that greatly increases the number of sub-ring instances. It supports enhanced Ethernet technologies such as Smart Link and RRPP to implement millisecond-level protective switchover, improving network reliability. Smart Link and RRPP both support multi-instance to implement load balancing among links, further improving bandwidth usage.

The S5700 supports E-Trunk. When a CE is dual homed to a VPLS, VLL, or PWE3 network, an E-Trunk can be configured to protect the links between the CEs and PEs and implement backup between PEs. The E-trunk can implement link aggregation across devices to upgrade the link reliability to device level.

The S5700 supports SEP, a ring network protocol applied to the link layer of an Ethernet network. SEP is applicable to open ring networks and can be deployed on upper-layer aggregation devices to provide millisecond-level switchover without interrupting services. Huawei devices have implemented Ethernet link management using SEP. SEP features simplicity, high reliability, high switchover performance, convenient maintenance, and flexible topology and enables users to manage and plan networks conveniently.

The S5700 supports VRRP to keep the communication continuity and reliability, ensuring a stable network. Multiple equal-cost routes can be configured on the S5700 to implement route redundancy. When the active uplink route is faulty, traffic is automatically switched to a backup route. This feature implements multi-level backup for uplink routes.

Rich Security Measures and QoS Policies

The S5700 provides various security measures. It can defend against DoS attacks, attacks to networks, and attacks to users. DoS attacks include SYN Flood attacks, Land attacks, Smurf attacks, and ICMP Flood attacks. Attacks to networks refer to STP BPDU/root attacks. Attacks to users include bogus DHCP server attacks, man-in-the-middle attacks, IP/MAC spoofing attacks, DHCP request flood attacks, and DoS attacks by changing the CHADDR field of packets.

The S5700 listens to information about the MAC or IP address of an access user, IP address lease, VLAN ID, and interface by establishing and maintaining a DHCP snooping binding table. The S5700 directly discards invalid packets such as ARP spoofing packets and packets with bogus IP addresses that do not match binding entries. In this manner, hackers or attackers are prevented from carrying out the man-in-the-middle attacks by using ARP packets on campus networks. The trusted interface feature of DHCP snooping ensures the validity of the DHCP server.

The S5700 supports strict learning of ARP entries to prevent ARP spoofing attackers from exhausting ARP entries so that authorized users can access the Internet. The S5700 supports IP source check to prevent DoS attacks caused by MAC address spoofing, IP address spoofing, and MAC/IP spoofing. URRF provided by the S5700 can reverse check the packet transmission path to authenticate packets, which can protect the network against increasing source address spoofing attacks.

The S5700 supports the integrated MAC address authentication and 802.1x authentication. User information, such as the user name, IP address, MAC address, VLAN ID, access interface, and a flag indicating whether antivirus software is installed on the client, can be

bound statically or dynamically, and policies (VLAN, QoS, and ACL) can be delivered dynamically.

The S5700 can limit the number of MAC addresses learned on an interface to prevent attackers from exhausting MAC address entries by using bogus source MAC addresses. In this way, MAC addresses of authorized users can be learned and flooding is prevented.

The S5700 can implement complex traffic classification based on information such as the 5-tuple, IP preference, ToS, DSCP, IP protocol type, ICMP type, TCP source port, VLAN, the protocol type of an Ethernet frame, and CoS. The S6700 supports inbound and outbound ACLs. The S5700 supports the flow-based two-rate and three-color CAR. Each interface supports eight priority queues and multiple queue scheduling algorithms such as WRR, DRR, SP, WRR+SP, and DRR+SP, which ensures the quality of network services such as voice, video and data services.

Convenient Operation and Maintenance

The S5700 supports automatic configuration, plug-and-play, deployment from USB devices, and batch remote upgrade. Upgrade and delivery of the S5700 can be completed at one time, which simplifies management and maintenance. Maintenance costs are greatly reduced. The S5700 supports diversified management and maintenance modes such as SNMPv1/v2/v3, CLI, Web network management, Telnet, and HGMP, which makes device management more flexible. In addition, the S5700 supports NTP, SSHv2.0, TACACS+, RMON, multi-log host, interface-based traffic statistics, and NQA, which helps to better deploy and adjust networks.

The S5700 supports GVRP. The GVRP technology implements dynamic configuration of VLANs. In a complicated networking environment, GVRP can simplify VLAN configuration and reduce network communication faults caused by incorrect configuration of VLANs. This reduces the manual configurations of network managers and ensures correct VLAN configurations.

The S5700 supports MUX VLAN. The MUX VLAN function is used to isolate Layer 2 traffic between interfaces on a VLAN. Subordinate VLANs can communicate with the MUX VLAN but cannot communicate with each other. MUX VLAN is usually used on enterprise intranets. With this function, a user interface can communicate with a server interface but cannot communicate with other user interfaces. MUX VLAN prevents communication between network devices connected to some interfaces or interface groups but allows these devices to communicate with the default gateway. This function ensures resource sharing and secure communication in an enterprise.

The S5700 supports BFD and provides millisecond-level detection for protocols such as OSPF, IS-IS, VRRP, and PIM to improve network reliability. Complying with IEEE 802.3ah and 802.1ag, the S5700 supports point-to-point Ethernet fault management. It can detect faults on user links. Ethernet OAM improves the network management and maintenance capabilities on the Ethernet and ensures a stable network.

The S5700HI and the S5706 provide 3.3 ms hardware-based Ethernet OAM function and Y.1731, which can quickly detect and locate faults. By using the Ethernet OAM technology and switchover technologies, the S5700 can provide millisecond-level protection for networks.

Rich IPv6 Features

The S5700 supports IPv4/IPv6 protocol stack and can be upgraded smoothly. The S5700 hardware supports the IPv4/IPv6 protocol stack, IPv6 over IPv4 tunnels (including manual tunnels, 6to4 tunnels, and ISATAP tunnels), and Layer 3 wire-speed forwarding. Therefore, the S5700 can be deployed on IPv4 networks, IPv6 networks, and networks that

simultaneously run IPv4 and IPv6. This makes the networking flexible and meets the requirements for the network transition from IPv4 to IPv6.

The S5700 supports various IPv6 routing protocols including RIPng and OSPFv3. It uses the IPv6 NDP to manage packets exchanged between neighbors. It also provides the PMTU mechanism to select a proper MTU on the path from the source to the destination, optimizing network resources and obtaining the maximum throughput.

7.3.4 Product Specifications

Table 7-5 Main specifications of the S5700 series products

Item	S5706TP-LI	S5700-SI	S5700-EI	S5700HI
Extended slot	<ul style="list-style-type: none">• The S5706 has no extended slot.• The 5700TP provides a stacking extended slot.• The S5700C provides two extended slots. One supports subcards and the other supports stacking cards.• The S5700HI provides an extended slot that supports subcards.			
Forwarding performance (PPS)	<ul style="list-style-type: none">• S5706: 9 Mbit/s• S5724TP-SI/S5724TP-PWR-SI: 36 Mbit/s• S5748TP-SI/S5748TP-PWR-SI: 72 Mbit/s• S5728C-SI/S5728C-PWR-SI/S5728C-EI/S5728C-PWR-EI/S5728C-EI-24S/S57HI: 96 Mbit/s• S5752C-SI/S5752C-PWR-SI/ S5752C-EI/S5752C-PWR-EI: 132 Mbit/s			
Interface switching capacity (bit/s)	<ul style="list-style-type: none">• S5706: 12 Gbit/s• S5724TP-SI/S5724TP-PWR-SI: 48 Gbit/s• S5748TP-SI/S5748TP-PWR-SI: 96 Gbit/s• S5728C-SI/S5728C-PWR-SI/S5728C-EI/S5728C-PWR-EI/S5728C-EI-24S/S57HI: 128 Gbit/s• S5752C-SI/S5752C-PWR-SI/ S5752C-EI/S5752C-PWR-EI: 176 Gbit/s			
Backplane switching capacity	256 Gbit/s			
MAC address table	<ul style="list-style-type: none">• LI/SI: 16K; EI/HI: 32K• Automatic learning and aging of MAC addresses• Static, dynamic, and blackhole MAC address entries• Packet filtering based on source MAC addresses			
VLAN	<ul style="list-style-type: none">• 4K VLANs• Guest VLANs and voice VLANs• VLANs based on MAC addresses, protocols, IP subnets, policies, and interfaces.• 1:1 and N:1 VLAN switching• QinQ and selective QinQ			

Item	S5706TP-LI	S5700-SI	S5700-EI	S5700HI
MPLS features	Not supported.	Not supported.	Not supported.	<ul style="list-style-type: none"> • Support basic MPLS functions. • Support MPLS VLL.
IPv4 route	Static route	Static route, RIPv1, RIPv2, ECMP, and URPF	<ul style="list-style-type: none"> • OSPF, IS-IS, and BGP • VRRP • Policy-based routing • Routing policy • The same as those of the SI 	Same as those of the EI
IPv6 route	Static route	<ul style="list-style-type: none"> • RIPng • Manual tunnels • 6to4 tunnels • ISTAP tunnels 	<ul style="list-style-type: none"> • OSPFv3 • The same as those of the SI 	Same as those of the EI
IPv6 features	<ul style="list-style-type: none"> • ND • PMTU • IPv6 Ping, IPv6 Tracert, and IPv6 Telnet • 6to4 tunnels, ISATAP tunnels, and manually configured tunnels • ACLs based on the source IPv6 address, destination IPv6 address, Layer 4 interface, or protocol type • MLDv1/v2 snooping 			
Multicast	<ul style="list-style-type: none"> • Static Layer 2 multicast MAC address • MAC address-based multicast forwarding 	<ul style="list-style-type: none"> • IGMP snooping and IGMP fast leave • Multicast VLAN • MLD snooping • IGMP proxy • Controllable multicast • Interface-based multicast traffic statistics 	<ul style="list-style-type: none"> • IGMP v1/v2/v3 • PIM-SM, PIM-DM, and PIM-SSM • MSDP • The same as those of the SI 	Same as those of the EI

Item	S5706TP-LI	S5700-SI	S5700-EI	S5700HI
QoS/ACL	<ul style="list-style-type: none"> • Rate limit on packets sent and received by an interface • Packet redirection • Port-based traffic policing and two-rate and three-color CAR • Eight queues on each port • WRR, DRR, SP, WRR+SP, and DRR+SP queue scheduling algorithms • WRED (supported by the S5706 and the S5700HI) • Re-marking of the 802.1p priority and DSCP priority of packets • Packet filtering on Layer 2 to Layer 4, filtering out invalid frames based on the source MAC address, destination MAC address, source IP address, destination IP address, port number, protocol, and VLAN ID • Queue-based rate limit and port-based traffic shaping 			
Reliability	<ul style="list-style-type: none"> • STP, RSTP, and MSTP • BPDU protection, root protection, and loop protection • RRPP topology and RRPP multi-instance • Smart Link tree topology, Smart Link multi-instance, and the millisecond-level protection • SEP • BFD for OSPF, BFD for IS-IS, BFD for VRRP, and BFD for PIM (supported by the S5700EI/HI series) • E-trunk 			
Security	<ul style="list-style-type: none"> • Hierarchical user management and password protection • DoS attack defense, ARP attack defense, and ICMP attack defense • Binding of the IP address, MAC address, interface, and VLAN • Interface isolation, interface security, and sticky MAC addresses • Blackhole MAC addresses • Limit on the number of learned MAC addresses • IEEE 802.1x authentication and limit on the number of users on an interface • Multiple authentication methods including AAA, RADIUS, TACACS+, and NAC authentication • SSH v2.0 • CPU protection • Blacklist and whitelist 			

Item	S5706TP-LI	S5700-SI	S5700-EI	S5700HI
OAM	<ul style="list-style-type: none"> • Hardware implementation • EFM OAM • CFM OAM • Y.1731 performance test supports hardware-level delay and jitter detection 	Software implementation	Software implementation	<ul style="list-style-type: none"> • Hardware implementation • EFM OAM • CFM OAM • Y.1731 performance test supports hardware-level delay and jitter detection
Management and maintenance	<ul style="list-style-type: none"> • Intelligent stacking (excluding the S5700HI and the S5706) • MFF • Virtual cable test • Ethernet OAM (IEEE 802.3ah and 802.1ag) • Local port mirroring, remote switched port analyzer (RSPAN) and the packet forwarding on observing ports • Remote configuration and maintenance using Telnet • SNMPv1/v2/v3 • RMON • Network management system (NMS) and Web NMS • HGMP • System logs and multi-level alarms • Dying gasp power-off alarm (supported only by the S5706) • GVRP • MUX VLAN • HTTPS • 802.3az EEE (supported only by the S5700HI and the S5706) 			
Working environment	<ul style="list-style-type: none"> • Working temperature: 0°C to 50°C (long term); -5°C to 55°C (short term) • Relative humidity: 10% to 90% (non-condensing) 			
Input voltage	<p>AC power supply</p> <ul style="list-style-type: none"> • Rated voltage: 100 V to 240 V, 50/60 Hz • Maximum voltage: 90 V to 264 V, 50/60 Hz <p>DC power supply</p> <ul style="list-style-type: none"> • Rated voltage range: -48 V to -60 V • Maximum voltage: -36 V to -72 V <p>Note: Models supporting the PoE supply do not use DC power supplies.</p>			

Item	S5706TP-LI	S5700-SI	S5700-EI	S5700HI
Dimensions (H x W x D)	<ul style="list-style-type: none"> • S5706: 250 mm x 180 mm x 43.6 mm • S5724TP-SI/S5724TP-PWR-SI/S57HI: 442 mm x 220 mm x 43.6 mm • Others: 43.6 mm x 442 mm x 420 mm 			
Power consumption	S5706: < 40 W	<ul style="list-style-type: none"> • S5724TP-SI: < 40 W • S5724TP-PWR-SI: < 455 W • S5748TP-SI: < 64 W • S5748TP-PWR-SI: < 907 W • S5728C-SI: < 56 W • S5728C-PWR-SI: < 891 W • S5752C-SI: < 78 W • S5752C-PWR-SI: < 917 W 	<ul style="list-style-type: none"> • S5728C-EI: < 60 W • S5728C-PWR-EI: < 472 W • S5728C-EI-24S: < 63 W • S5752C-EI: < 88 W • S5752C-PWR-EI: < 930 W 	S57HI: < 93 W