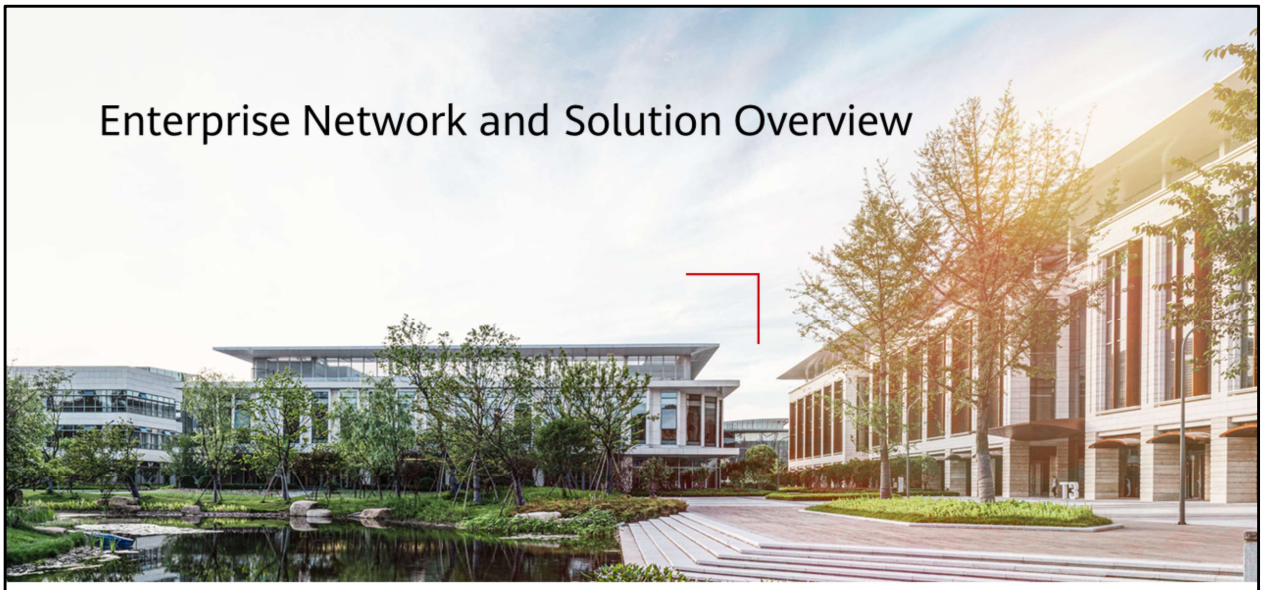
Enterprise Network and Solution Overview



HUAWEI

Foreword

- Nowadays, data has become a new factor of production, and data communications networks have become the foundation of digital society. How to design data communications networks according to enterprise service requirements is an important topic.
- Generally, the entire data communications world is divided into the following parts: wide area network (WAN), metropolitan area network (MAN), campus network, and data center network (DCN). Network security is involved in all the preceding fields and needs to be considered in an end-to-end manner.
- Enterprise WANs enable network interconnection across the country or even across the globe.
- An enterprise MAN is a medium-scale computer network that connects multiple LANs in the same area. It provides network access for office campuses, schools, hospitals, and other organizations. It is located between the LAN and WAN and covers a geographic region of the size of a metropolitan area.
- Enterprise campus networks can be classified into large and midsize campus networks and small and midsize campus networks based on the scale.
- Enterprise DCNs implement access of servers and storage devices. Typically, enterprise DCNs can be divided into financial cloud data centers and e-Government cloud data centers by industry.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Know the definition of enterprise networks
  - Understand the WAN, MAN, campus network, and DCN architecture
  - Describe typical enterprise network scenarios
  - Describe Huawei enterprise network solutions

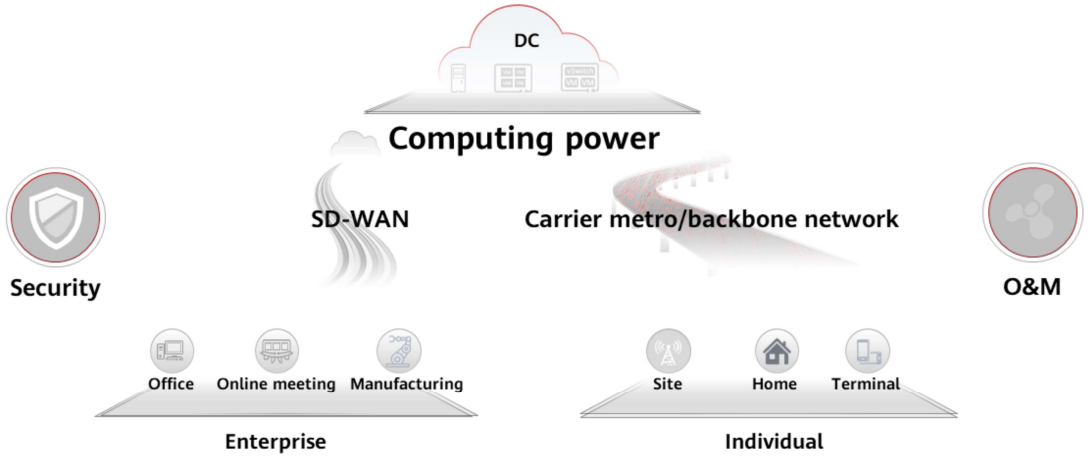
HUAWEI

Contents

1. **Enterprise Network Overview**
2. Typical Enterprise Network Scenarios
3. Trends and Challenges of Enterprise Networks
4. Huawei Enterprise Network Solutions

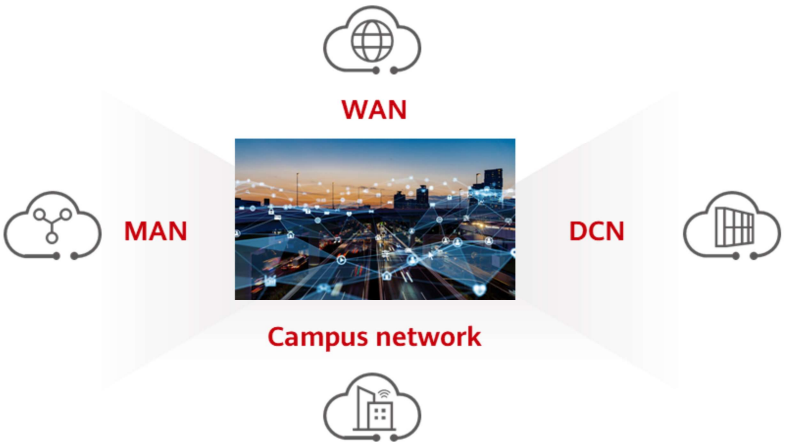
HUAWEI

Data Communications Network Builds the Foundation of Digital Society



HUAWEI

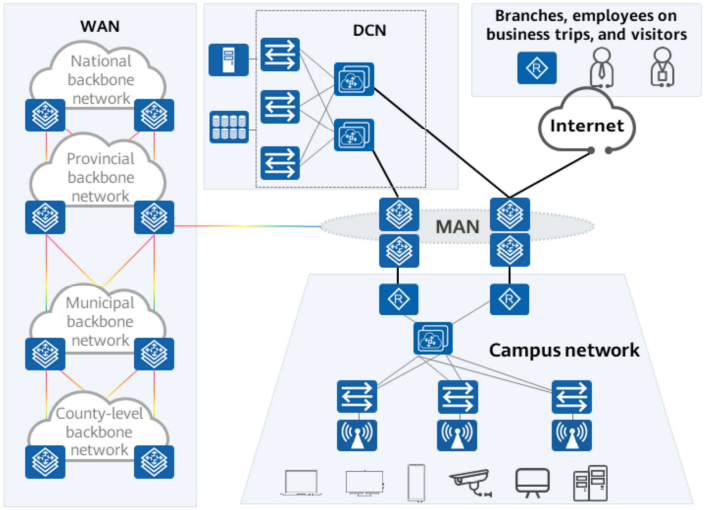
Enterprise Network



After more than 30 years of development, the network has laid a solid foundation for the intelligent society. With the acceleration of enterprise digitalization, enterprise networks need to carry more and more key services, playing an important role in the digitalization process.

HUAWEI

Enterprise Network Architecture



WAN — A network that connects nodes at different levels, such as the headquarters, branches, data centers (DCs), and mobile offices, across a large geographical area through leased carrier links or self-built links.

MAN — A medium-scale network that connects multiple LANs in the same area.

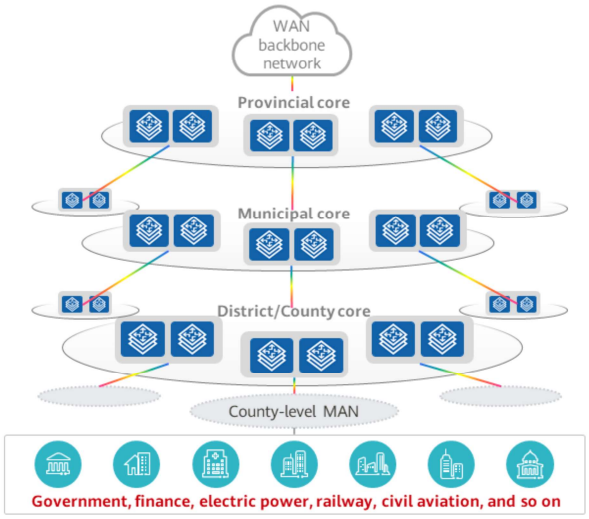
Campus network — A network used in the campus where we work and live.

DCN — A network that connects to DC resources and provides IT services for enterprises.

HUAWEI

- Backbone network: a core network that connects primary nodes for network interconnection

Enterprise WAN Overview

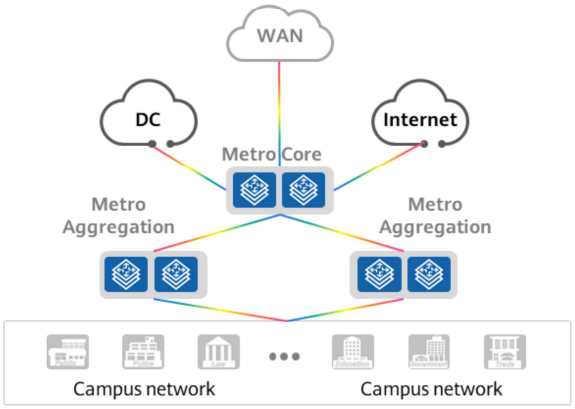


Large enterprises build their own WANs to meet the interconnection requirements of nodes at different levels, such as the headquarters, branches, DCs, and mobile offices.

Typical industry WANs include government WANs, financial WANs, electric power communication networks, railway bearer networks, and civil aviation bearer networks.

HUAWEI

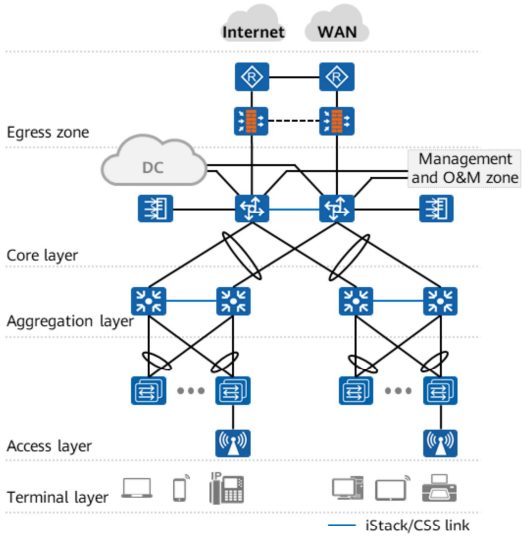
Enterprise MAN Overview



- Generally, the enterprise MAN adopts a hierarchical design. Based on the scale, the enterprise MAN is divided into the three-layer architecture (core layer-aggregation layer-access layer) or the two-layer architecture (core layer-access layer).
- The enterprise MAN provides the access function for campus networks in the same area to implement communication between the campus networks, and connects to the enterprise WAN to implement cross-region communication on the enterprise network.

HUAWEI

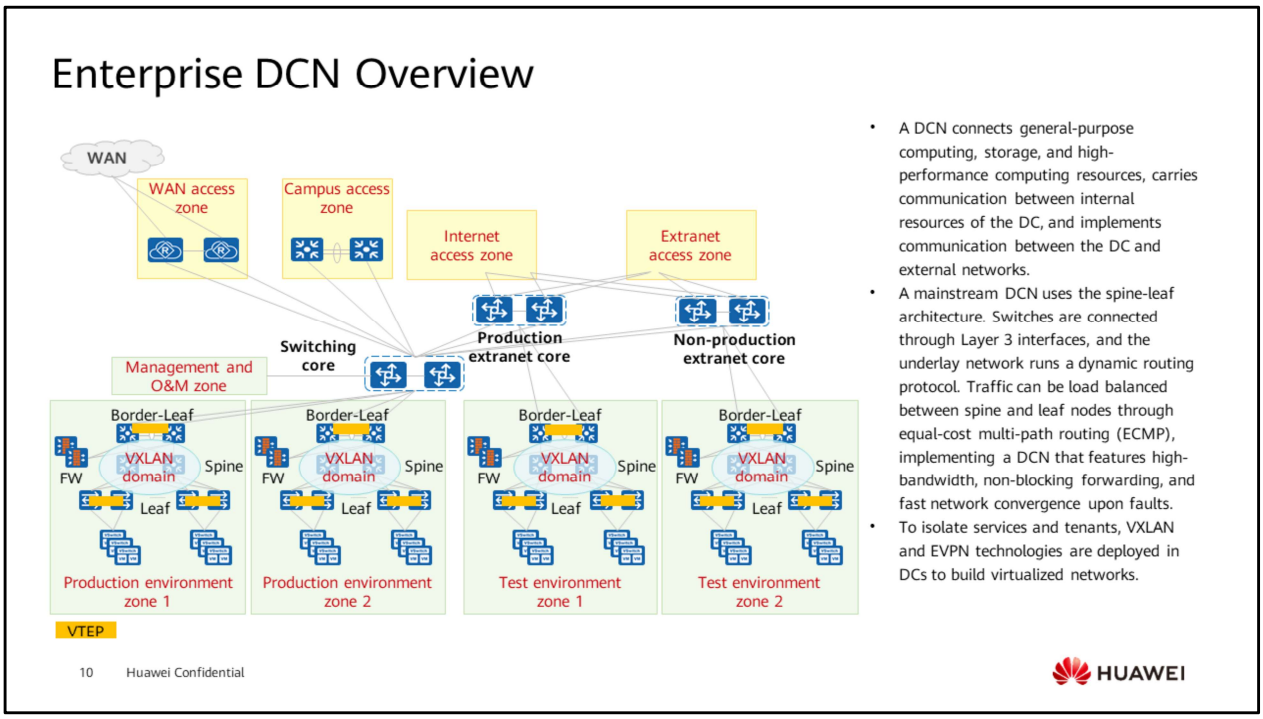
Enterprise Campus Network Overview



An enterprise campus network uses a hierarchical and modular design. Based on the number of terminal users or NEs, campus networks can be classified into small-, medium-, and large-sized ones. A typical large-sized campus network consists of the following modules:

- **Egress zone**: the border between the campus internal network and external network. It is used for internal users to access the public network and for external users to remotely access the internal network.
- **Core layer**: the core of campus data switching that connects all components of the campus network, such as the DC, management and O&M zone, and campus egress.
- **Aggregation layer**: forwards horizontal traffic between users and vertical traffic to the core layer. The aggregation layer can function as the switching core of a department or area and can connect more terminals to the network.
- **Access layer**: provides various access modes for users and is the first network layer to which terminals connect.
- **Terminal layer**: involves various terminals that access the campus network, such as PCs, printers, IP phones, mobile phones, and cameras.
- **DC**: deploys servers and application systems to provide data and application services for internal and external users of an enterprise.
- **Management and O&M zone**: deploys the network management system (such as the NMS and authentication server) and O&M system.

HUAWEI



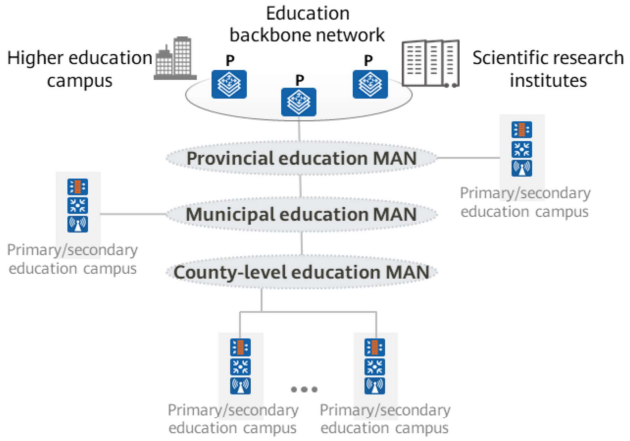
- DCNs do not have a fixed zone division mode. Different industries and enterprises have different area division modes.
- The zone division in the slide uses the financial industry as an example:
  - WAN access zone: connects to the WAN built by the enterprise.
  - Campus access zone: accesses traffic from enterprise campus users.
  - Extranet access zone: connects to external networks, such as networks of other enterprises and partners.
  - Internet access zone: accesses user traffic from the Internet.
  - Production environment zone: runs services.
  - Test environment zone: provides services for external systems informally. Before a service is brought online to the production environment zone, the service is developed, tested, and verified in the test environment zone.

Contents

1. Enterprise Network Overview
2. **Typical Enterprise Network Scenarios**
3. Trends and Challenges of Enterprise Networks
4. Huawei Enterprise Network Solutions

HUAWEI

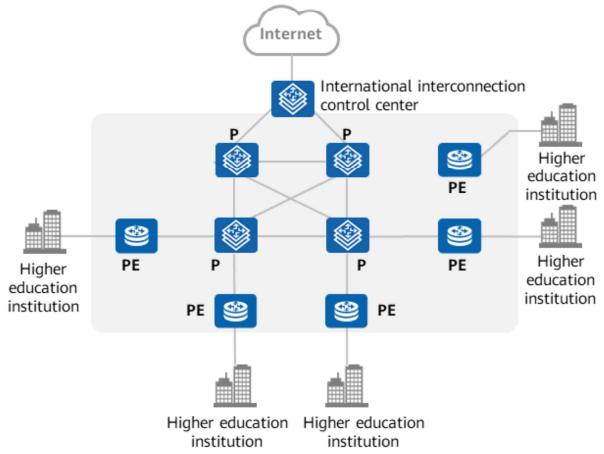
Education Network



- An education network is a basic network platform that provides services such as interconnection, resource sharing, and Internet access for education institutions, scientific research institutions, higher education campuses, and primary/secondary education campuses.
- An education network can be divided into the education backbone network, education MAN, higher education campus network, and primary/secondary education campus network.

HUAWEI

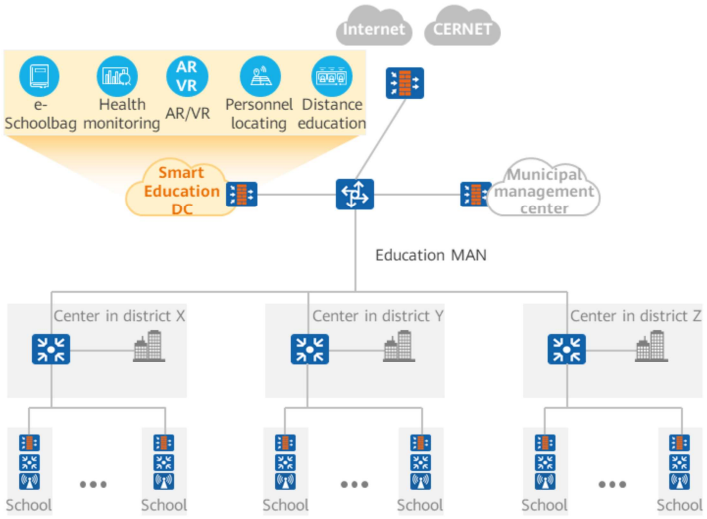
Education Backbone Network



- The education backbone network connects universities, education institutions, and scientific research institutions across the country. As the upper-level network of the education MAN, the backbone network functions as the egress for the MANs to access external networks and the hub for communication between MANs.

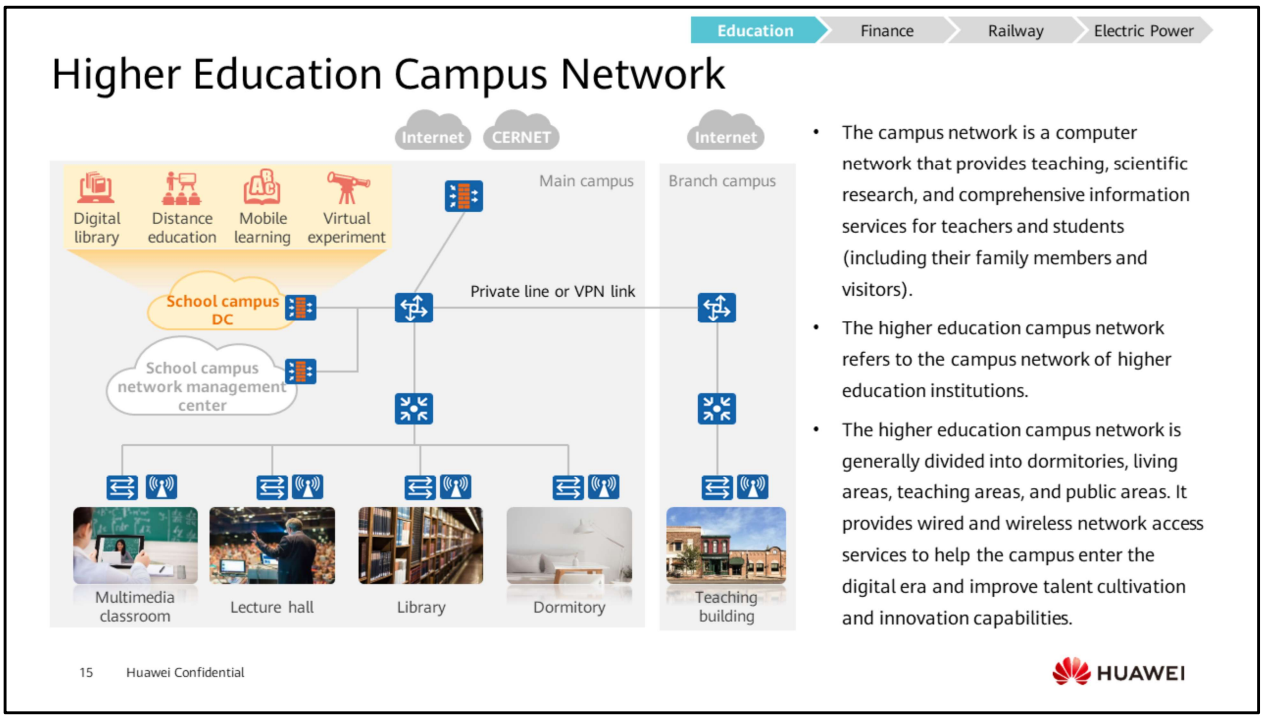
HUAWEI

Education MAN



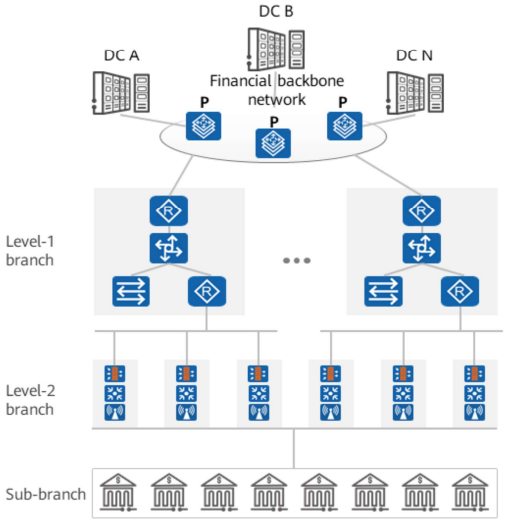
- The education MAN implements interconnection between education administrative organizations and schools. It is a regional network platform that carries functions such as education information resource sharing and education information management.
- The primary/secondary education campus network is not directly connected to the education backbone network due to its small scale. Instead, it is planned and constructed by the upper-level network management department and connected to the education MAN to implement full data convergence.

HUAWEI



- China Education and Research Network (CERNET): a nationwide education and research computer network that is funded by the Chinese government and directly managed by the Chinese Ministry of Education. It is constructed and operated by Tsinghua University and the other leading Chinese universities.

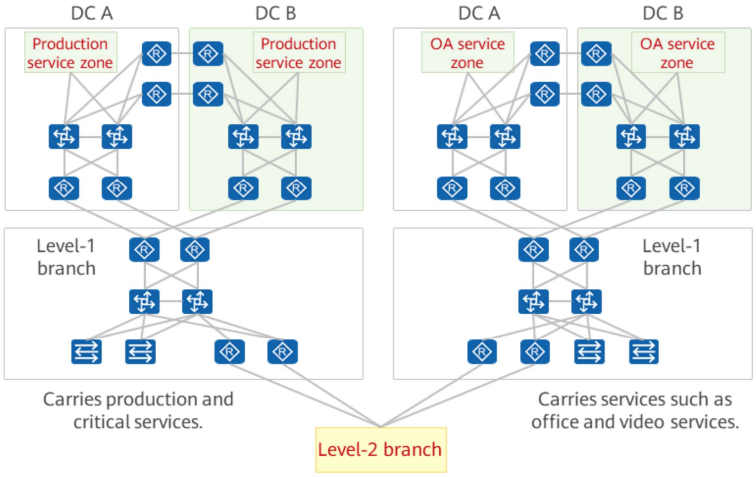
Financial Backbone Network Architecture



- To meet their own business development and operation management requirements, financial enterprises generally build their own backbone networks across the country to support the development of enterprise informatization and core services.
- The DC and its branches are connected through the backbone network. Network services are mainly centralized in the DC. Users are widely distributed inside and outside the banks. Branches access the DC of the head office through the backbone network.

HUAWEI

Financial Service Transport

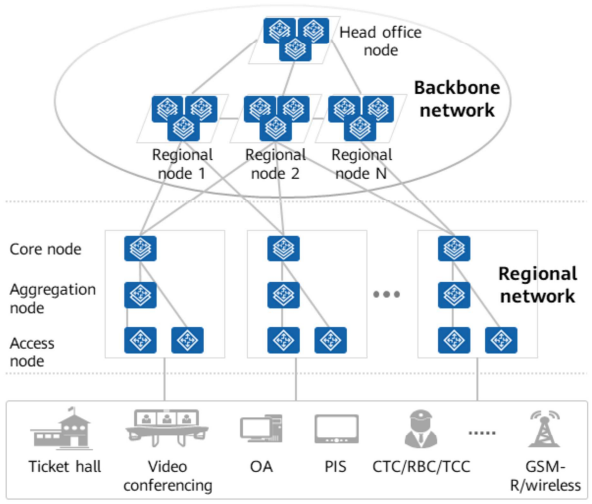


- From the perspective of service transport, the backbone network of financial enterprises has two construction models. One is two physical networks (as shown in the left figure) that carry different types of services independently, and the other is one single physical network that carries converged services.
- In the first model, networks that carry different services do not communicate with each other. A small number of services that need to be accessed across networks are accessed through firewalls.

HUAWEI

Education | Finance | Railway | Electric Power

Railway Data Communications Network



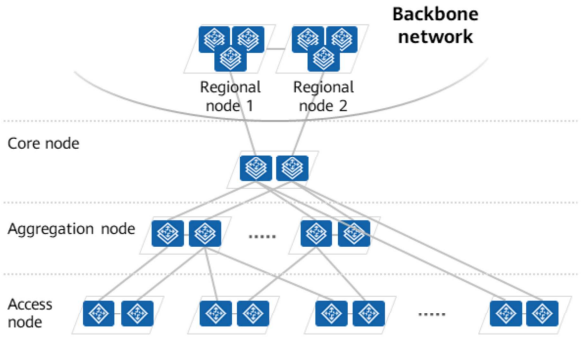
- The railway system in China is used as an example.
- In the railway system, the network provides unified transmission services for signaling, dispatching, ticketing, integrated services, and GSM-R services.
- The railway data communications network uses a two-layer architecture:
  - A backbone network consisting of multiple high-performance core routers located at regional nodes and railway bureaus.
  - A regional network set by each railway bureau. The regional network is connected to the backbone network through the backbone network's regional node of the local railway bureau to implement communication between railway bureaus.
- As an independent autonomous system, the backbone network forwards traffic between the head office and railway bureaus and between railway bureaus.
- Each railway bureau is configured with a backbone network access node and multiple backbone network forwarding nodes. The head office node functions as the service aggregation node of the entire railway system and does not forward traffic.

HUAWEI

- Generally, the regional nodes of the backbone network are deployed in the equipment rooms of each railway bureau.
- GSM-R: Global System for Mobile Communications - Railway
- CTC: Centralized Traffic Control
- RBC: Radio Block Center
- PIS: Passenger Information System
- TCC: Track Communications Center

Education Finance Railway Electric Power

Railway Regional Data Network

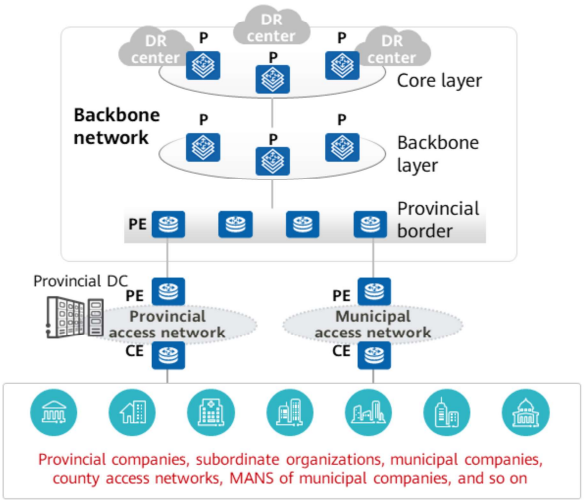


- In this example, the core node consists of two routers, which are respectively connected to the regional node of the local railway bureau and another geographically adjacent regional node.
- An aggregation node is directly connected to two routers of a core node. The aggregation node and the core node can form a mesh or partial mesh connection. Two access routers should be configured on each access node of the railway bureau; dispatch station, comprehensive maintenance center, and train detection base for the passenger dedicated line; and general packet radio service (GPRS). The access routers can be directly connected to the core nodes of the local network.
- Two routers on each access node (for a station or depot) can be directly connected to two routers of the same aggregation node or two geographically adjacent aggregation nodes.

HUAWEI

- General Packet Radio Service (GPRS).

Integrated Data Network for the Electric Power Industry



- The electric power industry in China is as an example.
- The electric power industry network is divided into two independent networks based on the carried services: dispatching network and integrated data network.
- The dispatching network, as one of the core networks for electric power informatization, mainly carries power-related service control systems.
- The integrated data network carries applications such as OA, production management, lightning weather information, and remote monitoring of unmanned substations. It is physically independent and parallel with the power dispatching network.
- The integrated data network adopts a multi-level architecture, including the national backbone network, provincial integrated data network, and municipal integrated data network. The national backbone network is divided into the core layer (consisting of multiple core nodes deployed in the country) and the backbone layer (consisting of backbone nodes distributed in provincial companies in the country).
- The data networks of provincial companies and branches are connected to the provincial border nodes of the backbone network to implement communication between provincial networks.

HUAWEI

Contents

1. Enterprise Network Overview
2. Typical Enterprise Network Scenarios
3. **Trends and Challenges of Enterprise Networks**
4. Huawei Enterprise Network Solutions

HUAWEI

Industry Digital Transformation, Improving Production Efficiency and Customer Satisfaction

Digital OA

All-wireless access, OA anytime and anywhere, one-click conference scheduling using an app, and automatic adjustment of lighting and temperature



Source: According to *Worldwide Digital Transformation 2019 Predictions* released by IDC, "at least 55% of organizations will be digitally determined by 2020."

Digital education

Change from stuffing the duck to immersion, high-quality teaching resources on demand, and real-world teaching



Source: Gartner's *2019 CIO Agenda: Higher Education Industry Insights* reveals that 86% of higher education CIOs would regard IT as a key factor in enabling education business transformation.

Digital manufacturing

Always-online mass terminals/sensors, real-time production data collection, and automatic and precise control



Source: IDC's 2018 manufacturing industry survey results shows that 84.9% manufacturing enterprises are undergoing digital transformation, driving business model innovations and reshaping the business ecosystem.

HUAWEI

In the Digital Era, How Campus Networks Support Digital Transformation Across Industries?

Ubiquitous connectivity



- Access anytime, anywhere
- High-quality service transport

On-demand services



- Quick service deployment and adjustment
- Rapid rollout of value-added applications

Perceptible experience



- Accurate measurement and evaluation of user experience
- AI-powered automatic adjustment of resource allocation

Easy O&M



- Service-based configuration delivery
- Quick and accurate fault locating

HUAWEI

Service Requirements and Challenges of Large- and Medium-Sized Campus Networks

Converged transport

Requirements:

Access terminals and services are diversified, and a converged transport network is required.

Challenges:

- Wi-Fi and IoT services are independently planned, deployed, and managed. The overall network construction cost is high.
- The workload of network management and O&M is heavy.

User experience awareness

Requirements:

Network O&M needs to be automated and intelligent to perceive user experience anytime and anywhere.

Challenges:

- Service faults cannot be detected in a timely manner.
- After a fault occurs, the fault cause is determined based on the O&M experience of professional personnel, and the fault cannot be quickly located.
- The network cannot be automatically optimized.

Network automation

Requirements:

Networks need to be automated to cope with the complexity of deployment and policies caused by the surge of applications and services.

Challenges:

- The workload is repetitive and heavy, and manual configuration is complex.
- New services need to be configured on devices one by one, which is time-consuming and costly.
- The workload of network policy deployment and adjustment is heavy.

HUAWEI

Service Requirements and Challenges of Small- and Medium-Sized Campus Networks

Plug-and-play network devices, improving deployment efficiency

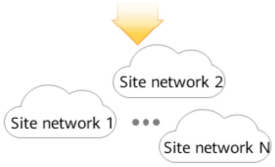
Unified management and centralized configuration



- The network configurations of multiple sites are centrally delivered, reducing onsite configuration and commissioning workload and improving deployment efficiency.
- The network is plug-and-play and able to be expanded on demand, requiring low costs for upgrades.

Centralized O&M on the cloud, simplifying O&M of multiple sites

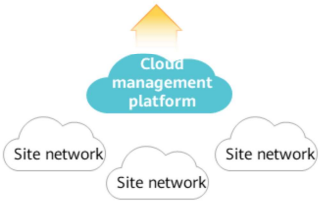
Centralized management of multiple branches on the cloud and remote automatic O&M



Scattered campus branch networks are centrally managed on the cloud through the Internet, and automation tools for troubleshooting, monitoring, and other management operations are integrated, implementing remote automated O&M.

Open APIs, accelerating integration of business applications

Open and big data analytics capability



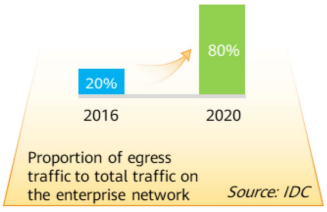
With open APIs and big data analytics capabilities, multiple management systems can be combined and interconnected to achieve unified network management. More value-added applications are available, leading enterprises into digital transformation.

HUAWEI

Development of Enterprise SD-WAN Technologies

Migration of enterprise services to the cloud

Enterprise services are migrated to the cloud, and enterprise egress traffic increases sharply.





Intelligence Cloudification Video Mobility

Emergence of Software defined networking (SDN)

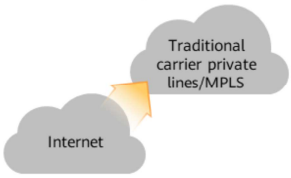
Unified management and centralized configuration



SDN introduces a new role, that is, the centralized controller, so that it can not only understand the network requirements proposed by the administrator, but also fully manage and configure the physical network, implementing automatic deployment and fast provisioning of applications on the network.

High-speed development of the Internet

Rapidly narrowing gap between the quality of the Internet and traditional private lines



The coverage and network performance of the Internet are continuously improving, and the gap between the quality of the Internet and traditional private lines is narrowing rapidly. As such, more and more enterprises and organizations are using the Internet for interconnection.

The business and technology trends related to enterprise WANs are changing rapidly, which will in turn have a profound impact on the development of enterprise WANs.

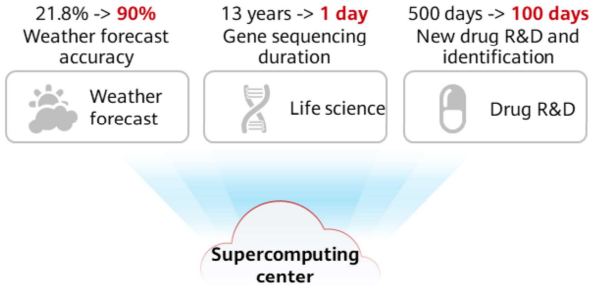
HUAWEI

Service Requirements and Challenges of Multi-Campus Network Interconnection

Branch interconnection cost	Service provisioning period	Service experience	Management and O&M	Unified management
Requirements: The cost of WAN interconnection between branches needs to be reduced. **Challenges:** Carriers' transmission private lines or MPLS VPN private lines are of high quality but expensive.	**Requirements:** The network deployment efficiency needs to be improved and the service provisioning period needs to be shortened. **Challenges:** • It takes a long time to provision traditional private lines. • Service provisioning is mainly performed by network engineers on site, which has high skill requirements and low efficiency.	**Requirements:** The application identification capability needs to be enhanced to ensure key service experience. **Challenges:** • Different enterprise applications have different requirements on link quality. • Traditional private lines cannot detect the application status and cannot guarantee key services at any time.	**Requirements:** Visualized O&M is required to simplify management and O&M. **Challenges:** • CLI-based O&M is inefficient and lacks visualized O&M methods. • If an enterprise has a large number of branches, onsite O&M will increase costs.	**Requirements:** LAN-side and WAN-side services need to be managed in a unified manner. **Challenges:** • Unified configuration management and O&M cannot be implemented.

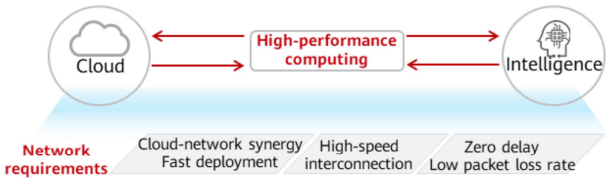
HUAWEI

Cloud + Intelligence, Accelerating HPC

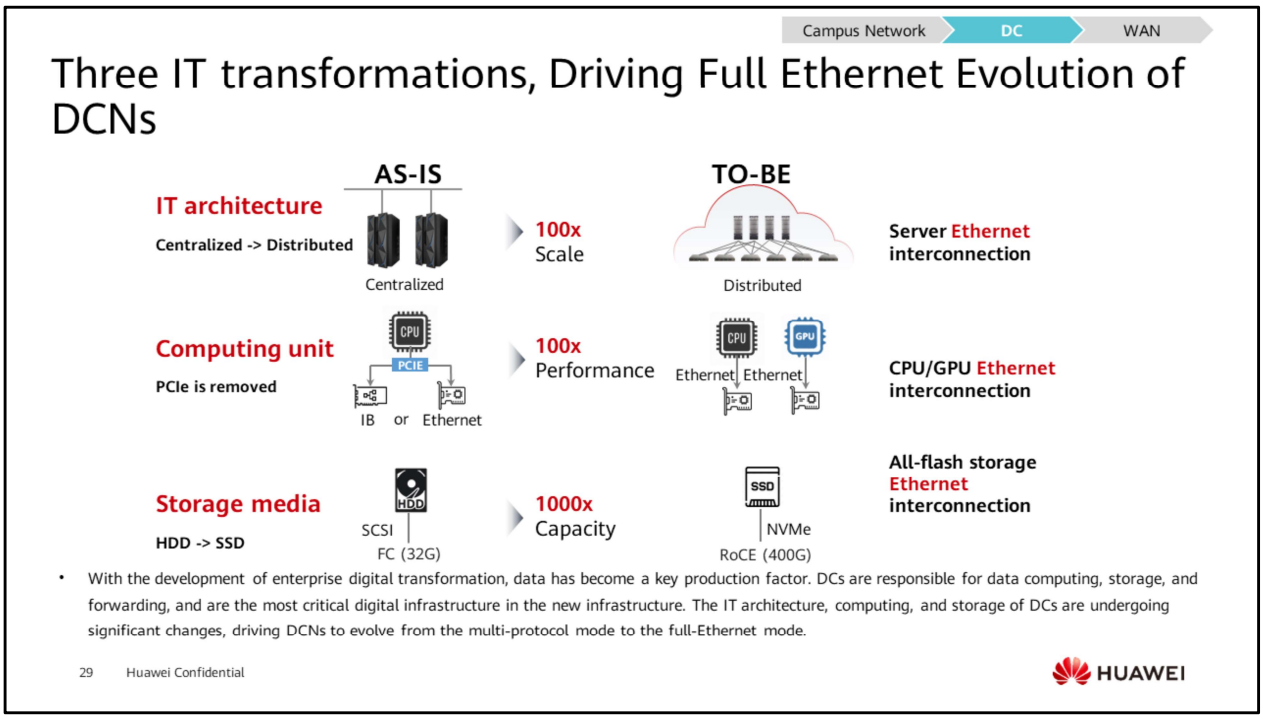


- Supercomputing — the core of high-performance computing (HPC) — provides the computing power needed to drive development of the digital economy and plays a key role in numerous scientific and technological fields.

- Speed is still the primary factor in supercomputing realizing faster and more accurate computing and simulation. This is where cloud and intelligence come in. In addition, the supercomputing of cloudification and artificial intelligence is gradually integrated, which poses higher requirements on the DCN that carries supercomputing services.



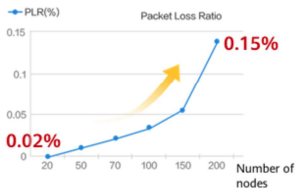
HUAWEI



- InfiniBand (IB): an input/output (I/O) switching technology that uses a central InfiniBand switch to establish a single link between remote storage devices, networks, and servers as well as to direct traffic. It has a compact structure, greatly improving system performance, reliability, and effectiveness, and relieving data traffic congestion between hardware devices.
- Remote Direct Memory Access over Converged Ethernet (RoCE): a network protocol that allows remote direct memory access (RDMA) over the Ethernet. There are two RoCE versions: RoCEv1 and RoCEv2. RoCEv1 is a data link layer protocol that allows any two hosts in the same Ethernet broadcast domain to communicate with each other. RoCEv2 is a network layer protocol and its packets can be routed.

Three Challenges Facing DCN Full-Ethernet Evolution

Zero packet loss required for high-performance computing



The packet loss rate increases exponentially with the increase of network nodes on traditional Ethernet.

Zero packet loss required for dual-active storage

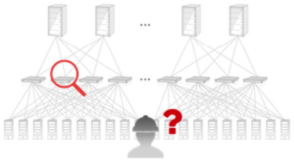
0.2–0.3% packet loss rate



The delay increases in intra-city (long distance) transmission, and it is difficult to perform flow control across DCNs on traditional Ethernet.

More complex for large-scale network O&M

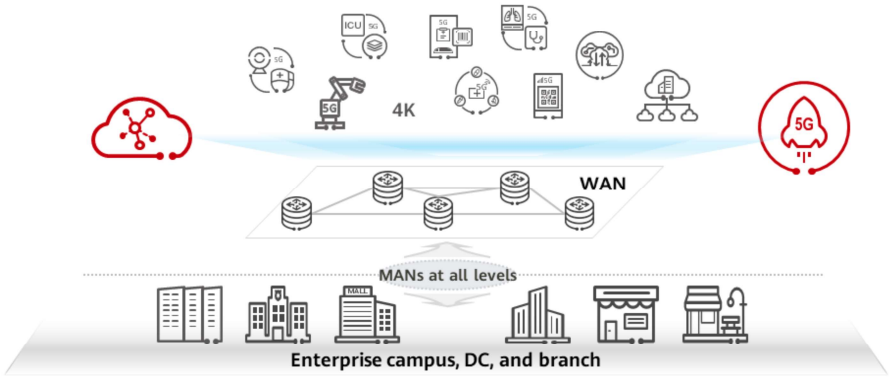
1000 nodes, **million-level** configuration



Traditional Ethernet lacks effective O&M methods, and the network is too complex to be handled manually.

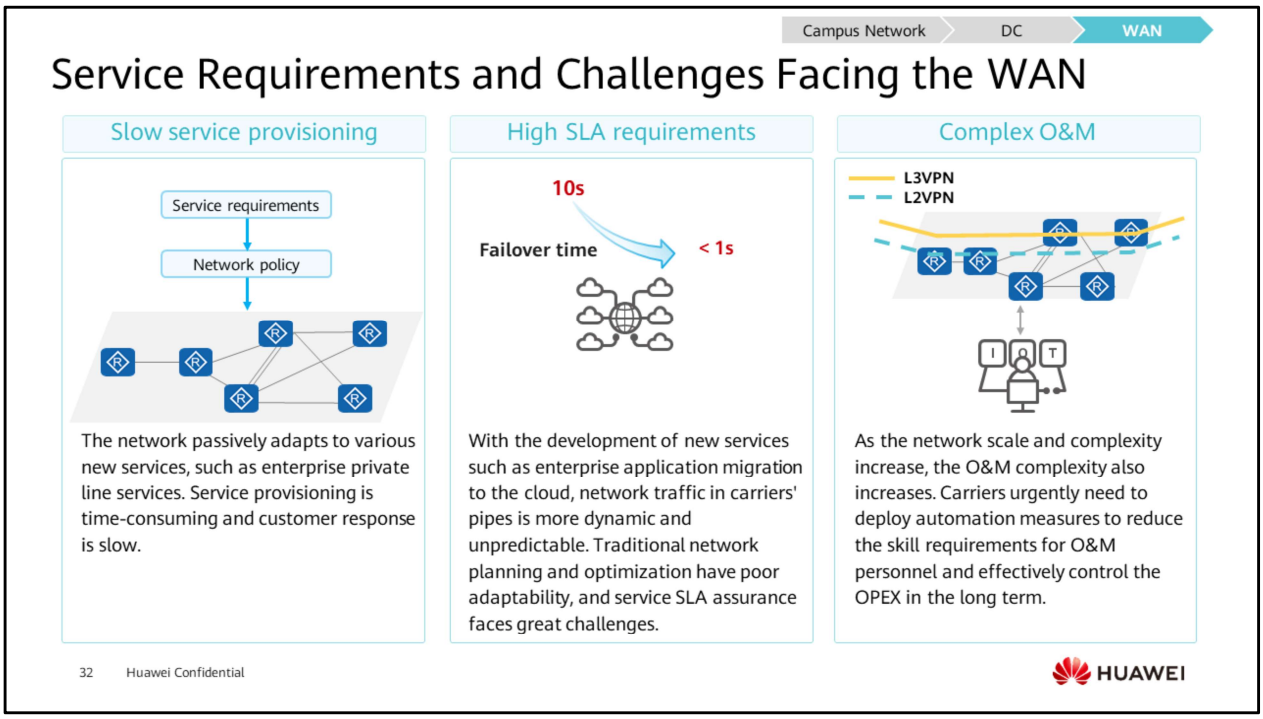
HUAWEI

5G + Cloud, Increasingly Complex WAN Transport Services



- With the advent of the 5G and cloud era, various innovative services such as VR/AR, live streaming, and autonomous driving are emerging. The traffic on the entire network increases explosively, and the dynamic complexity of services also increases the complexity of the entire network.

HUAWEI



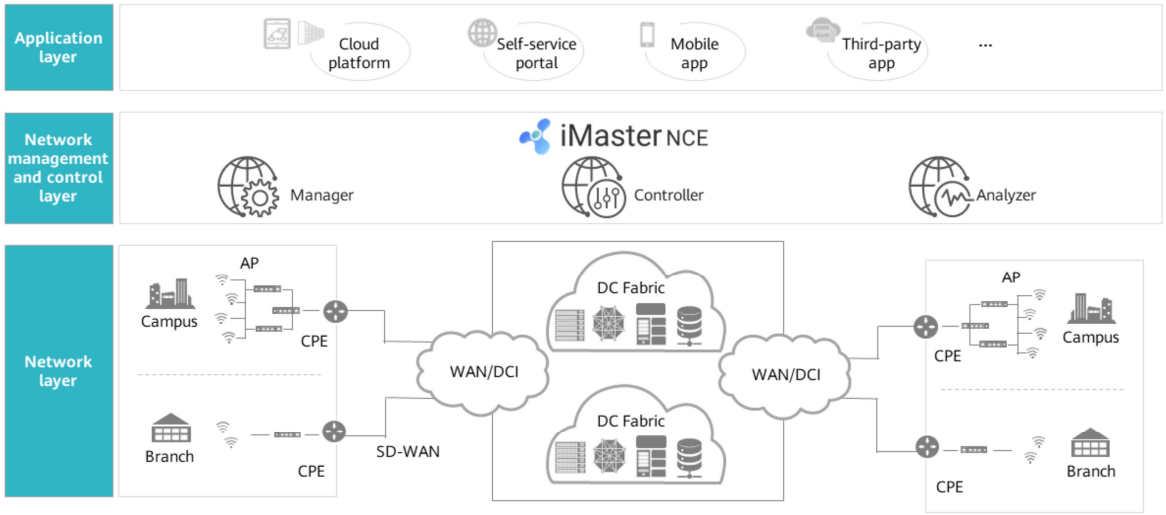
- Service Level Agreement (SLA): an agreement between a service provider and a user. The SLA defines the service type and quality provided by the service provider for the user, and the commitment to the performance and reliability of the user assurance service. For example, the SLA ensures that the service reliability is higher than 99.99%, the fault response time is within 30 minutes.
- Operating Expense (OPEX): the sum of the maintenance cost, marketing expense, labor cost, and depreciation expense during the enterprise operations.

Contents

1. Enterprise Network Overview
2. Typical Enterprise Network Scenarios
3. Trends and Challenges of Enterprise Networks
4. **Huawei Enterprise Network Solutions**

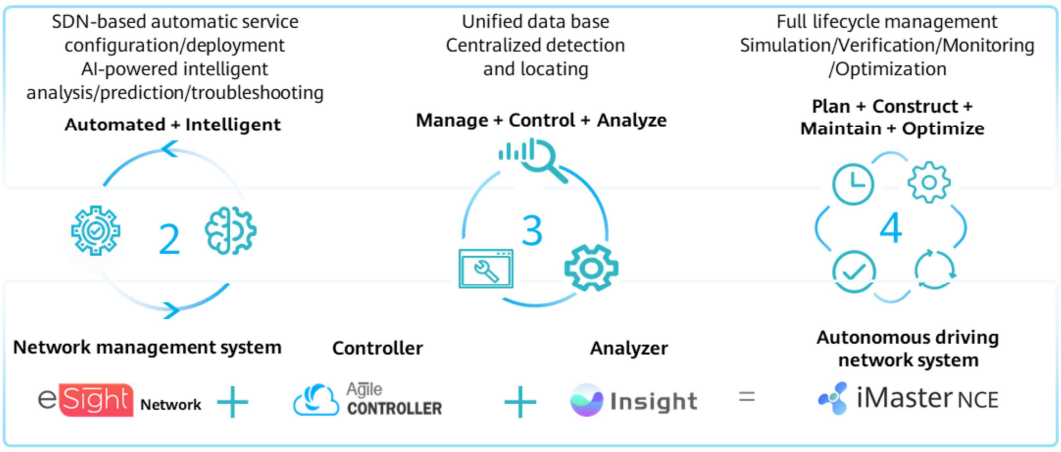
HUAWEI

iMaster NCE Builds an Autonomous Driving Network



HUAWEI

What Is iMaster NCE?



iMaster NCE is a network automation platform that integrates management, control, analysis, and AI capabilities.

HUAWEI

All-New iMaster NCE



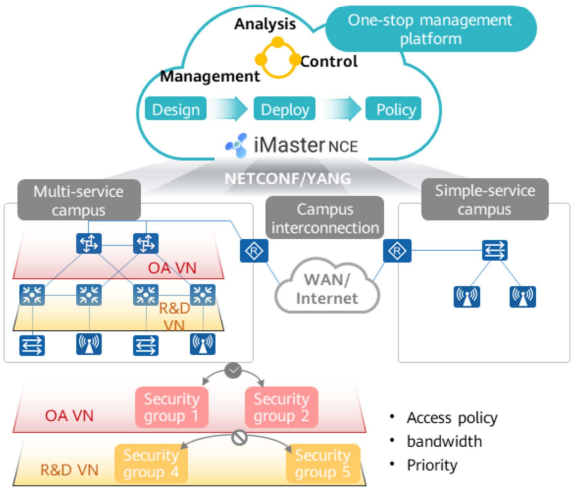
iMaster NCE

DC	iMaster NCE-Fabric
Enterprise campus	iMaster NCE-Campus
SD-WAN	iMaster NCE-WAN
IP WAN	iMaster NCE-IP
WAN transmission	iMaster NCE-T

HUAWEI

Campus Network | DC | WAN

One-Stop Autonomous Driving Solution for Campus Networks: CloudCampus



Automated network design, accurately matching service intents

- Accurate conversion between service requirements and configuration commands through the network management platform

Automated network deployment, achieving minute-level rollout

- Plug-and-play devices, automated deployment of physical networks
- Network resource pooling, one network for multiple purposes, and automatic service provisioning
- Intelligent policies (such as access policies, bandwidth, and priorities) based on virtual networks, security groups, and applications, implementing refined control

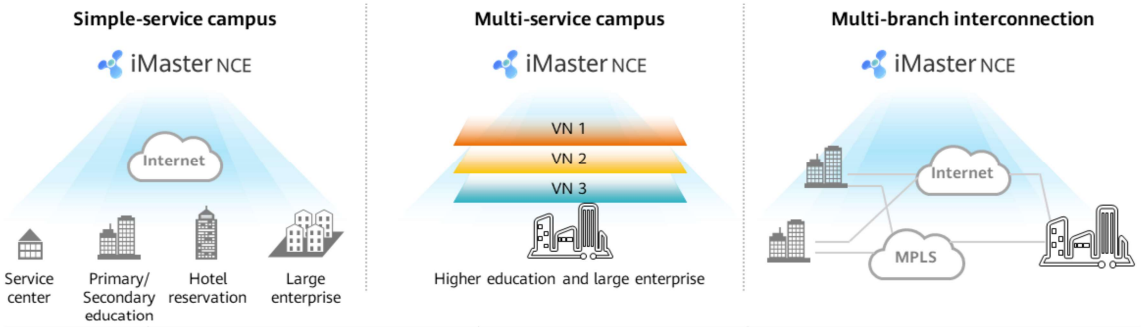
Automated intelligent O&M, improving network performance

- Real-time experience visualization: network experience visualization at any moment, for any user, and in any area
- Precise fault analysis: proactively identifies typical network problems and provides suggestions
- Intelligent network optimization: predictive optimization of wireless networks based on historical data

HUAWEI

- In the campus network scenario, iMaster NCE-Campus is used as the controller.

All-Scenario: Full Coverage from Single Campuses to Multi-Branch Interconnection Campuses

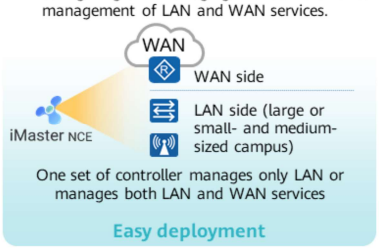


	Simple-service Campus	Multi-service Campus	Multi-branch Interconnection
Network Characteristics	Small scale, simple services Large numbers of sites, with similar models	Large scale, complex services, and coexistence of multiple services	Multiple branch sites, which need to communicate with each other through hybrid WAN links
Typical Scenarios	Multi-branch small-sized enterprise campuses, such as hotels and primary/secondary schools	higher education institutions, governments, large enterprise campuses	Large enterprises, financial services outlets

HUAWEI

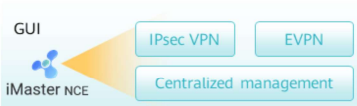
Full Convergence: One Controller Manages both LAN and WAN Services

- Huawei CloudCampus Solution integrates the configuration and management models of LAN and WAN services on the campus network. In addition to configuring and managing LAN services for campus networks, this solution can also manage WAN interconnection services, implementing integrated management of LAN and WAN services.



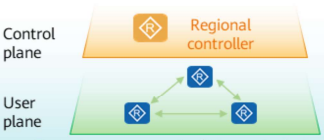
One set of controller manages only LAN or manages both LAN and WAN services

Easy deployment



GUI, flexible networking, device plug-and-play

Simplified configuration



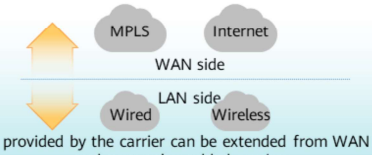
Central management of the control plane implements flexibly control while improving scalability

Forwarding-control separation



Visualized network service data for monitoring and analyzing the status of the entire network

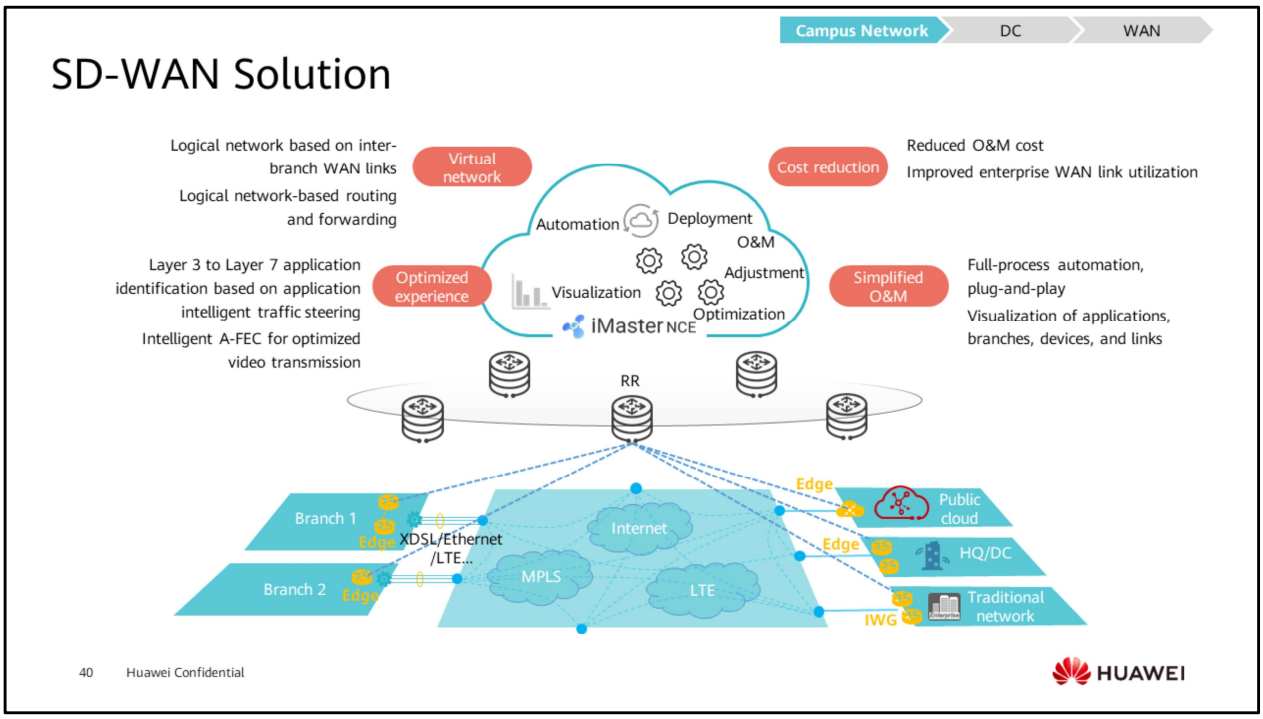
Simplified O&M



Services provided by the carrier can be extended from WAN to LAN and even value-added services

Value extension

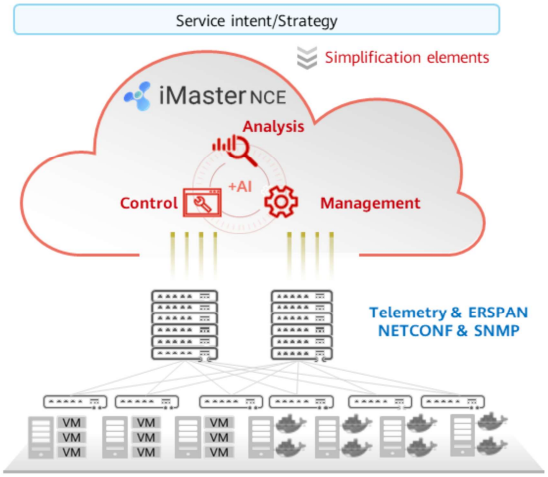
HUAWEI



- Adaptive Forward Error Correction (A-FEC): The intelligent data analysis engine is used to adjust the FEC protection window and protection mode based on the link quality to achieve high recovery and low redundancy.

Huawei CloudFabric Solution for Intelligent DCNs

- iMaster NCE-Fabric provides full lifecycle services for DCNs, including planning, construction, O&M, and optimization.



Integrated planning and construction
- Planning tool interconnected with iMaster NCE, implementing integrated planning and construction
- Zero Touch Provisioning (ZTP)

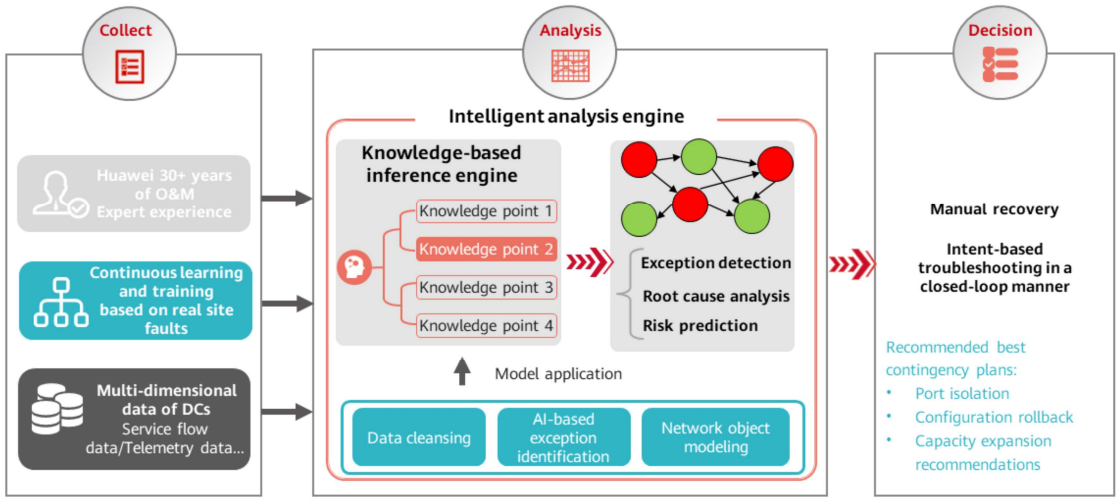
Simplified deployment
- Service intent self-understanding, conversion, and deployment
- Network change simulation evaluation, preventing human errors

Intelligent O&M
- Quick fault detection and locating based on knowledge graph and expert rules
- Fast fault rectification based on expert rules and simulation analysis

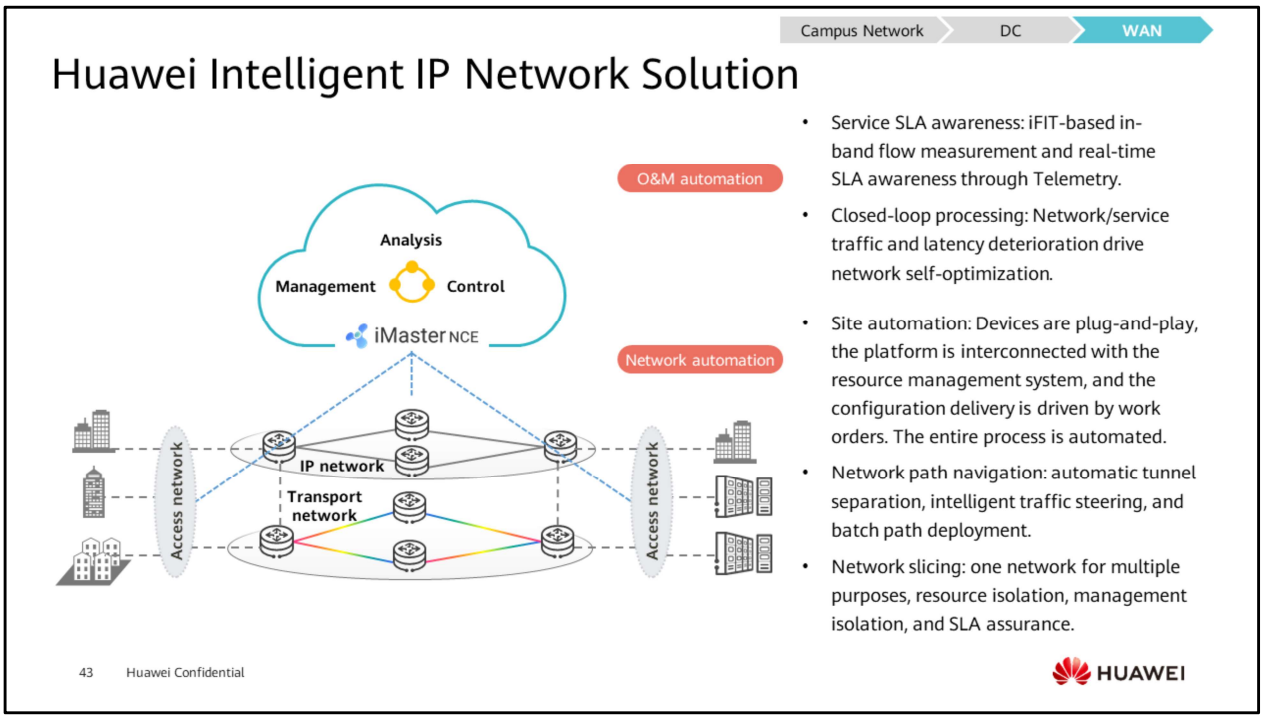
Real-time optimization
- Local inference and online model training and optimization for intelligent lossless traffic
- User behavior prediction and resource optimization suggestions

HUAWEI

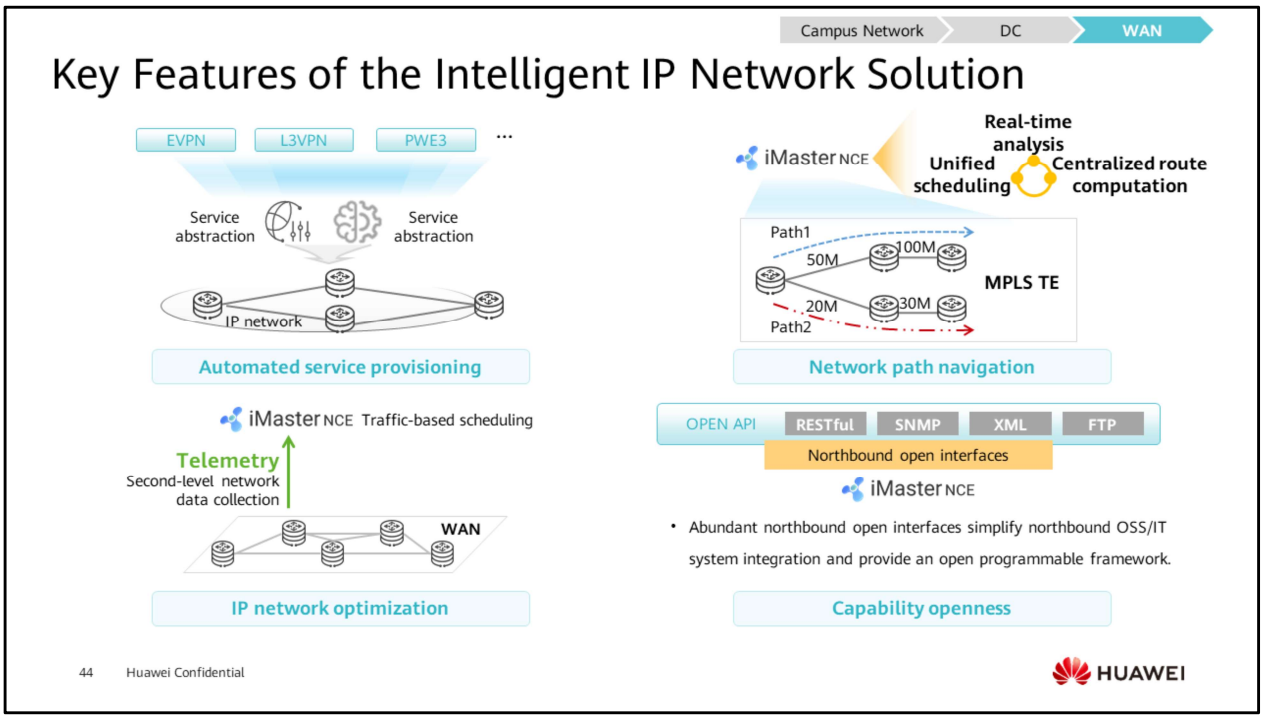
AI-based Intelligent O&M for DCNs



HUAWEI



- In this scenario, iMaster NCE-IP is used as the controller.
- In-situ Flow Information Telemetry (iFIT): a method of directly measuring service packets to obtain performance indicators of an IP network, such as the packet loss rate and delay.



- iMaster NCE automates service provisioning on IP networks through service abstraction and model-driven development. For mainstream services such as EVPN, L3VPN, and PWE3, iMaster NCE provides multiple service configuration templates, greatly reduces service provisioning parameter settings, supports batch service provisioning and automatic parameter verification, and automatically creates tunnels through service triggering, greatly shortening the service provisioning period and improving the service provisioning efficiency.
- Network path navigation: iMaster NCE performs real-time analysis, centralized path computation, and unified scheduling on the MPLS TE network to compute optimal E2E paths for MPLS TE tunnels, implementing global network traffic balancing and differentiated service management, and ensuring experience of high-priority services.
- iMaster NCE collects the link traffic bandwidth and direction of the IP network in real time. Based on the topology, traffic, and link quality of the entire network, iMaster NCE provides one-hop scheduling for cross-domain inbound and outbound traffic on the native IP network in DC egress and public network interconnection scenarios, implementing elastic bandwidth for VIP traffic, intra-city traffic localization, and egress link balancing. Optimal DC egress traffic is then provided to ensure the SLA of VIP services. In addition, iMaster NCE provides one-hop and multi-hop optimization capabilities to improve backbone network resource utilization and operation efficiency.

Quiz

1. (Single-answer question) Which of the following is the management and control platform used in Huawei's solution for data center scenarios? ( )
   A. iMaster NCE-IP
   B. iMaster NCE-Fabric
   C. iMaster NCE-WAN
   D. iMaster NCE-DCN

HUAWEI

- B

Summary

- This course describes the overall architecture of enterprise networks, typical topologies of WANs, MANs, campuses, and DCs, and typical enterprise networks in some industry scenarios.
- With the development of new technologies such as 5G and cloud and the trend of digital transformation in the industry, enterprises raise new requirements for networks. Enterprise networks are facing unprecedented transformation.
- Finally, this course briefly introduces Huawei's solutions for different enterprise network scenarios, including CloudCampus, SD-WAN, CloudFabric, and IP WAN. We will learn more about these solutions in subsequent courses.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Campus Network Architecture and Key Technology Applications



HUAWEI

Foreword

- Campuses are everywhere in a city except roads and households. Typical campus examples include factories, government buildings and facilities, shopping malls, office buildings, school campuses, and parks. According to statistics, 90% of urban residents work and live in campuses, 80% of gross domestic product (GDP) is created in campuses, and each person stays in campuses for 18 hours every day. Campus networks, as the infrastructure for campuses to connect to the digital world, are an indispensable part of campus construction and play an increasingly important role in daily working, R&D, production, and operations management.
- This course describes campus networks, common architectures and technologies of campus networks, and typical applications of these technologies. This information helps readers better understand the campus network design.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe what a campus network is like.
  - Describe the typical architecture of a campus network.
  - Understand common technologies used on campus networks.
  - Understand typical applications of campus network technologies.

HUAWEI

Contents

1. **Introduction to Campus Networks**
2. Typical Campus Network Technologies
3. Typical Applications of Campus Network Technologies

HUAWEI

Campuses Are Everywhere

90%+

of city residents work and live in campuses.

5 h & 22 h

spent in using smart terminals & staying in the campus every day



80%+

of GDP is created in campuses.

90%+

of innovations are made in campuses.

HUAWEI

Overview of a Campus Network

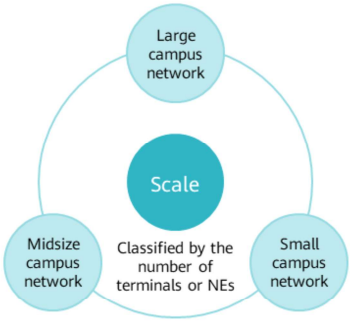


- A campus network refers to the internal network of an enterprise or organization, which is connected to the wide area network (WAN) and data center network.
- A campus network is built to ensure that key enterprise services are running more efficiently.
- Campus networks can be classified into large- and medium-sized campus networks and small- and medium-sized campus networks in terms of scale.
- Some enterprises have branches dispersed in different geographical locations. Each branch network can be considered as a single campus network.

HUAWEI

Campus Network Classification (1)

- Number of terminals > 2000
- Number of NEs > 100

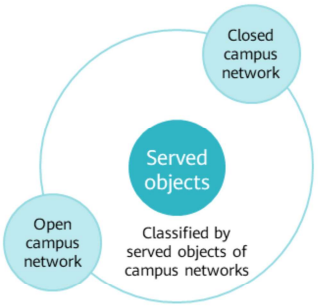


- Number of terminals: 200–2000
- Number of NEs: 25–100

- Number of terminals < 200
- Number of NEs < 25

Networks of different scales have different requirements and pain points.

- Users: internal personnel only
- Security requirements: network access control and external threat defense



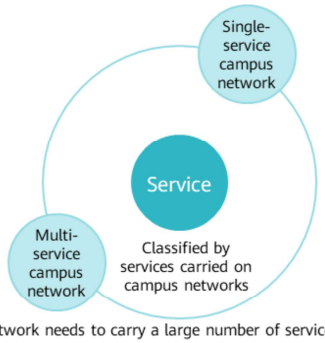
- Users: including external personnel, such as the public
- Security requirements: access control, identity identification, behavior control, security protection, etc.

A running campus network usually has both closed and open subnets.

HUAWEI

Campus Network Classification (2)

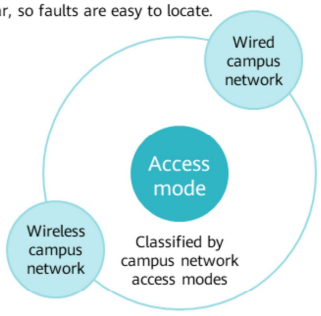
- Single services
- Simple network architecture



- The network needs to carry a large number of services, and the network scale is large. Different services need to be isolated and guaranteed.
- The campus network architecture is becoming complex and virtualized.

The complexity of the campus network architecture depends on the complexity of services carried on the campus network.

- Each device accessing a network must be connected to the preset network port through a network cable.
- The architecture is structured and hierarchical, and the logic is clear, so faults are easy to locate.



- The network is based on the 802.11 protocol (Wi-Fi) and is also called WLAN.
- AP deployment and installation affect the coverage effect. Interference and conflicts exist, so the network needs to be optimized periodically.
- Faults are difficult to locate.

Currently, most campus networks are a mix of wired and wireless networks.

HUAWEI

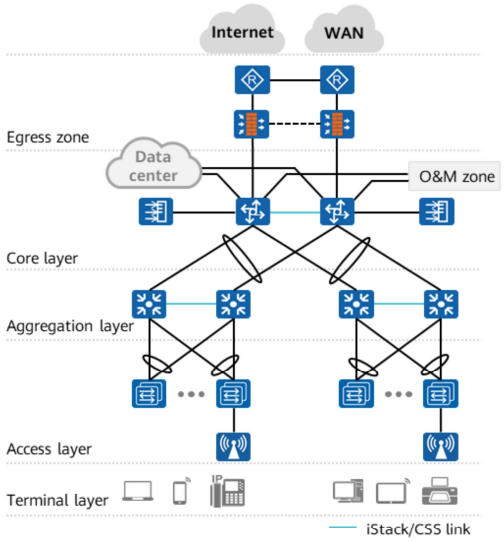
Campus Network Classification (3)

To meet requirements of different industries, the campus network architecture is designed based on the characteristics of the industries that campus networks serve, ultimately building industry-specific campus network solutions.

Enterprise campus network	School campus network	Government campus network	Business campus network
			
• It refers to the Ethernet-based enterprise office network. • The enterprise campus network focuses on network reliability and advancement, continuously improves employees' office experience, and ensures the efficiency and quality of operation and production.	• School campus networks are classified into primary/secondary education and higher education campus networks. • Higher education campus networks are complex and usually have teaching and research networks, student networks, and operational dormitory networks. • There are high requirements on network manageability and security, and specific requirements on network advancement.	• It usually refers to the internal network of a government agency. • High security is required. Generally, the internal network and external network are isolated to ensure high security of confidential information.	• This type of campus networks apply to shopping malls, supermarkets, hotels, and parks. • Such networks mainly serve consumers. In addition, they include subnets for internal office work. • Such networks provide Internet access services and help build business intelligence (BI) systems for a better user experience, lower O&M cost, higher efficiency, and value transfer.

HUAWEI

Typical Physical Architecture of a Campus Network



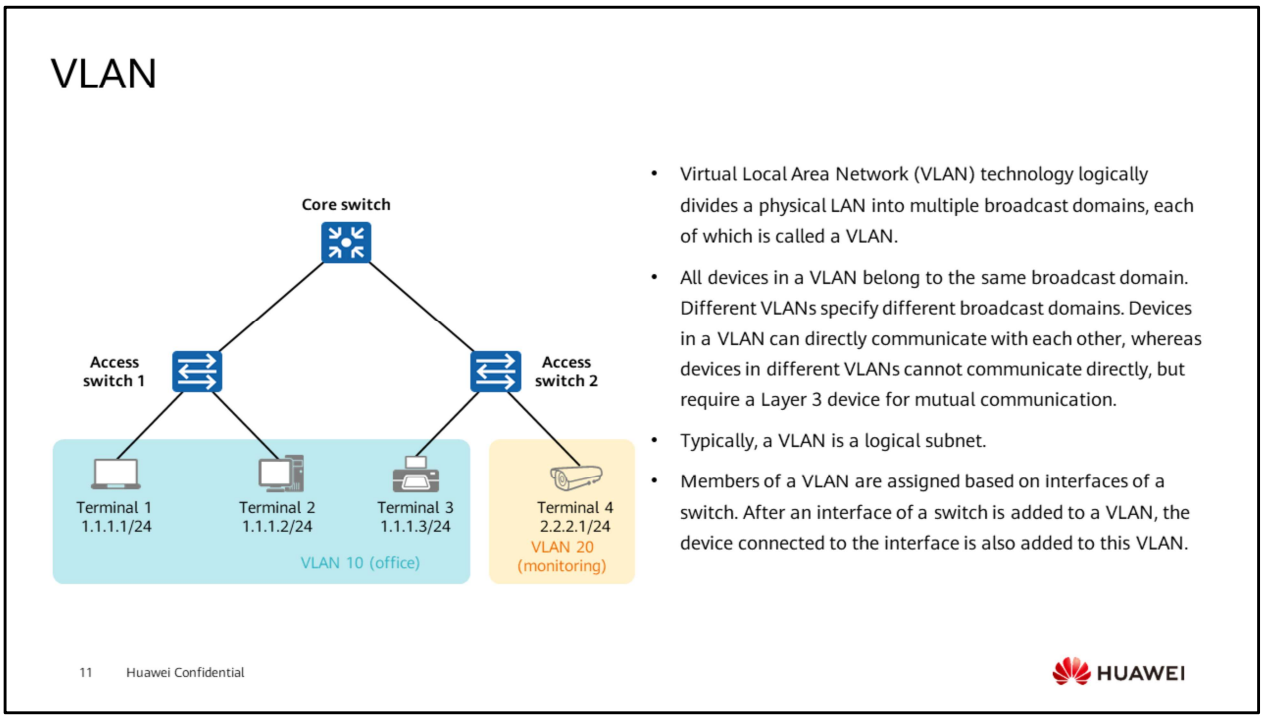
- **Egress zone**: serves as the border between the campus internal network and the external network. Through this egress zone, internal users can access the public network and external users (including customers, partners, branch users, and remote users) can access the internal network. Firewalls can be deployed at the egress zone to ensure the security of the internal network.
- **Core layer**: serves as the core of data switching on the campus network. It connects various parts of the campus network, such as the data center, O&M zone, and egress zone.
- **Aggregation layer**: forwards not only horizontal traffic between users, but also vertical traffic to the core layer. It can also function as the switching core for a department or zone and further extend the quantity of access terminals.
- **Access layer**: provides various access modes for users and is the first layer of the network to which a terminal accesses.
- **Terminal layer**: has terminals deployed to connect to the campus network. Terminals include computers, printers, IP phones, mobile phones, and cameras.
- **Data center**: has servers and application systems deployed to provide data and application services for internal and external users of the enterprise.
- **O&M zone**: manages network servers such as the NMS and authentication server.

HUAWEI

Contents

1. Introduction to Campus Networks
2. **Typical Campus Network Technologies**
3. Typical Applications of Campus Network Technologies

HUAWEI



- In addition to interface-based VLAN assignment, you can also use the following methods to assign VLANs:
  - MAC address-based assignment: assigns VLANs based on the source MAC addresses of frames. This mode applies to small-scale networks where physical locations of user terminals frequently change but their network adapters seldom change.
  - IP subnet-based assignment: assigns VLANs based on the source IP addresses of frames. This mode applies to scenarios where there are high requirements for mobility and simplified management and low requirements for security.
  - Protocol-based assignment: assigns VLANs based on the protocol (suite) types and encapsulation formats of frames. This mode applies to networks running multiple protocols.
  - Policy-based assignment: assigns VLANs based on a specified policy, which means VLANs are assigned based on combinations of interfaces, MAC addresses, and IP addresses. This mode applies to complex networks.

Voice VLAN

- Multiple types of flows are usually transmitted on a network simultaneously. Packet loss and delay seriously affect voice communication quality. Users are more sensitive to the quality of the voice service than to the quality of data and video services. Therefore, voice data flows must be preferentially transmitted when there is limited bandwidth available.
- A voice VLAN is used to exclusively transmit voice data.
- After a voice VLAN is configured on a switch, the switch can identify voice data flows, transmit the voice data flows in the voice VLAN, and provide QoS guarantee for the voice data flows.

Typical networking of voice VLAN technology



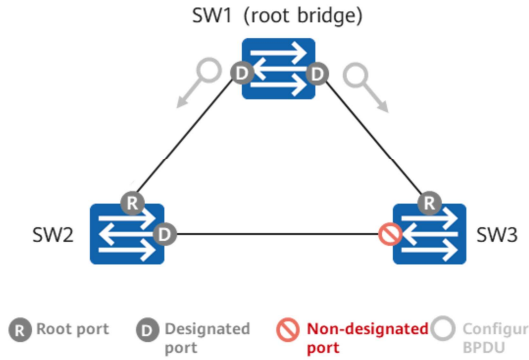
You can use a single interface of a switch to transmit both voice and data services.

An IP phone is connected to a switch independently.

HUAWEI

- A device configured with voice VLAN can identify voice flows in either of the following modes:
  - MAC address-based identification: A device identifies voice flows based on the source MAC addresses of received data packets. If a source MAC address of a packet matches the organizationally unique identifier (OUI) of a voice device, the packet is considered a voice packet. OUIs must be preconfigured and are used in scenarios where IP phones send untagged voice packets.
  - VLAN-based identification: Configuring OUIs for a large number of IP phones is time-consuming. In this case, you can configure a switch to identify voice packets based on VLAN IDs. If the VLAN ID of a received packet matches the configured voice VLAN ID, the packet is considered a voice packet. This simplifies configuration when a large number of IP phones are connected to the switch. However, the IP phones must be able to obtain voice VLAN information from the switch.

Spanning Tree Protocol (STP)



- On a switching network with physical loops, switches running STP automatically generate a loop-free working topology, which is also called an STP tree. Each node of an STP tree is a specific switch, and each branch is a specific link.
- STP transmits configuration BPDUs between switches to elect the root switch (or root bridge) and determine the role and status of each switch port.
- Each switch proactively sends configuration BPDUs during initialization. After the network topology becomes stable, only the root bridge periodically sends configuration BPDUs. Other switches send their own configuration BPDUs only after receiving configuration BPDUs from upstream devices.

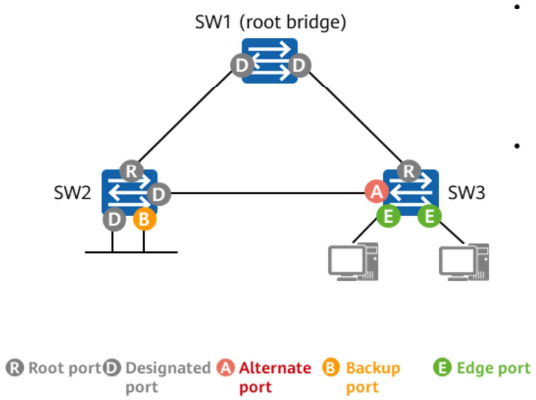
PID	PVI	BPDU Type	Flag	Root ID	RPC	Bridge ID	Port ID	Message Age	Max Age	Hello Time	Forward Delay

Format of a configuration BPDU

HUAWEI

- How is an STP tree generated?
  - STP compares four parameters: root bridge ID, root path cost (RPC), bridge ID (BID), and port ID (PID). A smaller value indicates a higher priority. All these parameters are BPDU fields.
    - Root bridge election: The device with the smallest root bridge ID is elected as the root bridge.
    - Root port election: A device compares the RPC, peer BID, peer PID, and local PID of its ports in sequence. The port with the smallest value is elected as the root port.
    - Designated port election: A device compares the RPC, local BID, and local PID of its ports in sequence. The port with the smallest value is elected as the root port.
    - After the root port and designated port are determined, all the non-root ports and non-designated ports on the switch will be blocked.

Rapid Spanning Tree Protocol (RSTP)

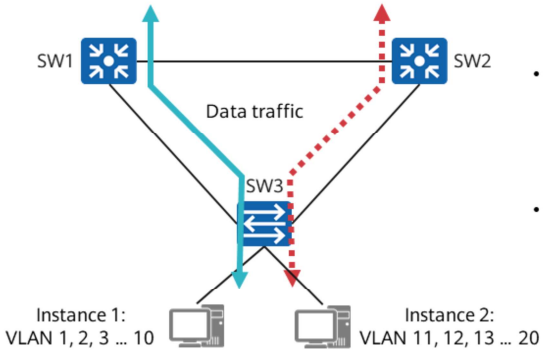


R Root port D Designated port A Alternate port B Backup port E Edge port

- As LANs are increasing in scale, the problem of slow STP topology convergence is becoming more prominent. To address this problem, the IEEE released the 802.1w standard in 2001 that defined the RSTP protocol. RSTP is an enhanced version of STP and implements rapid convergence of the switching network.
- RSTP makes the following improvements over STP:
  - Defines more port roles, making the spanning tree protocol easier to understand and configure.
  - Eliminates the listening state, reducing the number of port states from 4 to 3.
  - Uses the Flags field of STP configuration BPDUs to define port roles.
  - Processes configuration BPDUs differently from STP. For example, it allows non-root bridges to proactively generate and send configuration BPDUs, and reduces the BPDU timeout interval to three times that of the Hello Time.
  - Provides rapid convergence mechanisms, such as the Proposal/Agreement (PA) mechanism and edge port mechanism.
  - Provides protection mechanisms to ensure the stability of the switching network.

HUAWEI

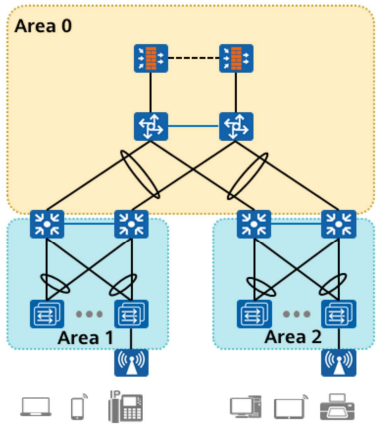
Multiple Spanning Tree Protocol (MSTP)



- Defined in IEEE 802.1s and compatible with STP and RSTP, MSTP implements fast convergence while providing multiple redundant paths for forwarding data, effectively load balancing traffic for VLANs.
- An MSTP network is composed of one or more MST regions, and each MST region contains one or more multiple spanning tree instances (MSTIs). An MSTI is a tree network that consists of switches running STP, RSTP, or MSTP.
- The switches in the same MST region share the following characteristics:
  - Have MSTP enabled.
  - Have the same region name.
  - Have the same VLAN-to-MSTI mappings.
  - Have the same MSTP revision level.

HUAWEI

OSPFv2

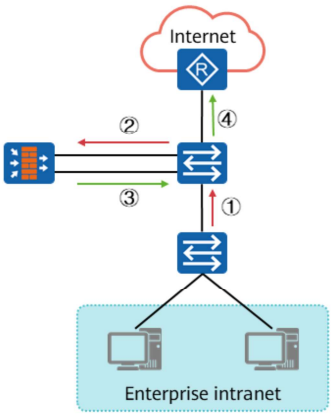


On a large-scale enterprise network, the aggregation layer may consist of Layer 3 devices, such as routers and switches. In such scenarios, static routes are not flexible because the configuration is complex and they cannot quickly respond to topology changes. Therefore, a dynamic routing protocol — typically OSPF in IGP — can be deployed on the enterprise intranet.

- The aggregation devices are divided into different non-backbone OSPF areas based on the network structure.
- Core devices, network egress devices, and uplink interfaces of aggregation devices belong to Area 0. The downlink interfaces of aggregation devices and the access devices are added to Area X.
- Redundancy links and the OSPF triggered update mechanism are used to implement backup of intranet paths.
- Aggregation devices function as area border routers (ABRs) to execute route filtering policies.

HUAWEI

Policy-Based Routing (PBR)



In the conventional routing and forwarding process, devices search their IP routing tables for routes based on packets' destination addresses and then forward the packets accordingly. However, more and more users require packet routing based on user-defined policies.

Traffic classifier

Configure a traffic classifier for matching interested data flows.

Matching criteria: VLAN ID, source or destination MAC address, Ethernet protocol type, DSCP priority, IP precedence, inbound or outbound interface, ACL rule

Traffic behavior

Redirects interested packets.

You can set the next-hop IP address or outbound interface for redirection.

Traffic policy

Traffic classifier > Traffic behavior

Apply the traffic policy

- Apply the traffic policy to the inbound direction of an interface.
- Apply the traffic policy to the incoming packets that belong to the VLAN and match the criteria defined in the traffic classifier.
- Apply the traffic policy globally or to a card.

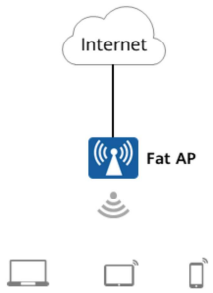
PBR can be configured using the modular QoS command line interface (MQC) or an ACL-based simplified traffic policy.

HUAWEI

- PBR is used in agile campus service orchestration, multi-egress, and anti-DDoS off-path deployment scenarios.

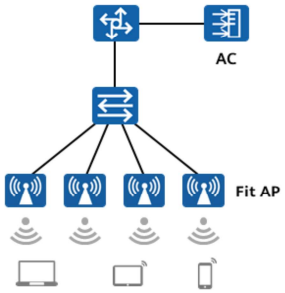
WLAN Network Architectures (1)

Fat AP



- **Networking characteristics**: A Fat AP works independently and needs to be configured separately. It provides only simple functions and is cost-effective.
- **Application scenarios**: families or mini-stores

AC + Fit AP

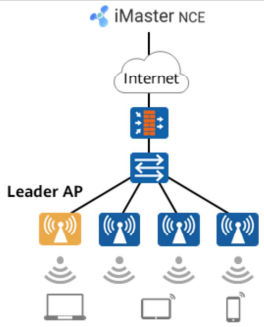


- **Networking characteristics**: A Fit AP provides a wide range of functions and is used with an AC. Fit APs are managed and configured by an AC in a unified manner, posing high requirements on skills of maintenance personnel.
- **Application scenarios**: medium- and large-sized enterprises

HUAWEI

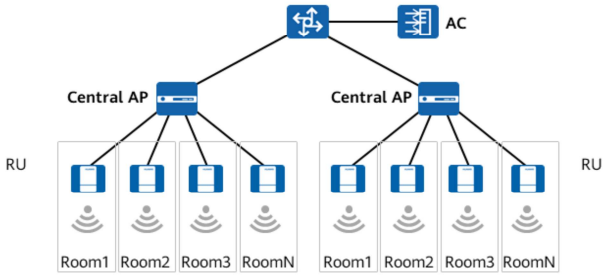
WLAN Network Architectures (2)

Leader AP



- **Networking characteristics**: A leader AP can work independently or manage a small number of common APs to implement basic roaming functions. A leader AP has a low price and low requirements for maintenance skills.
- **Application scenarios**: small and micro enterprises

Agile distributed AP

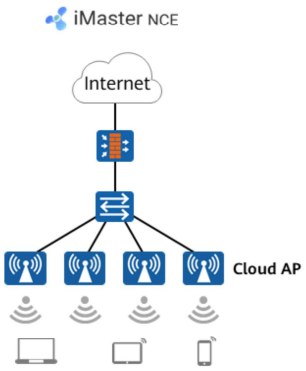


- **Networking characteristics**: The agile distributed architecture divides APs into central APs and remote units (RUs). A central AP can manage multiple RUs, and this architecture provides good coverage and reduces costs. RUs can be used in the Fat AP, AC + Fit AP, and cloud management architectures.
- **Application scenarios**: scenarios where rooms are densely distributed

HUAWEI

WLAN Network Architectures (3)

Cloud management mode

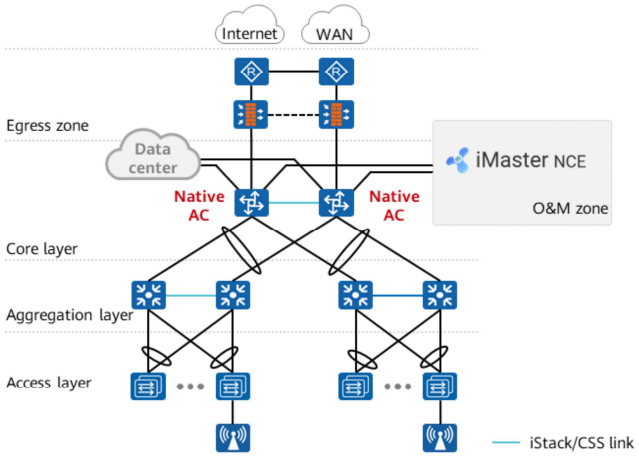


- **Networking characteristics**: APs are centrally managed and configured on iMaster NCE-Campus. They provide rich functions and support plug-and-play.
- **Application scenarios**: small- and medium-sized enterprises

HUAWEI

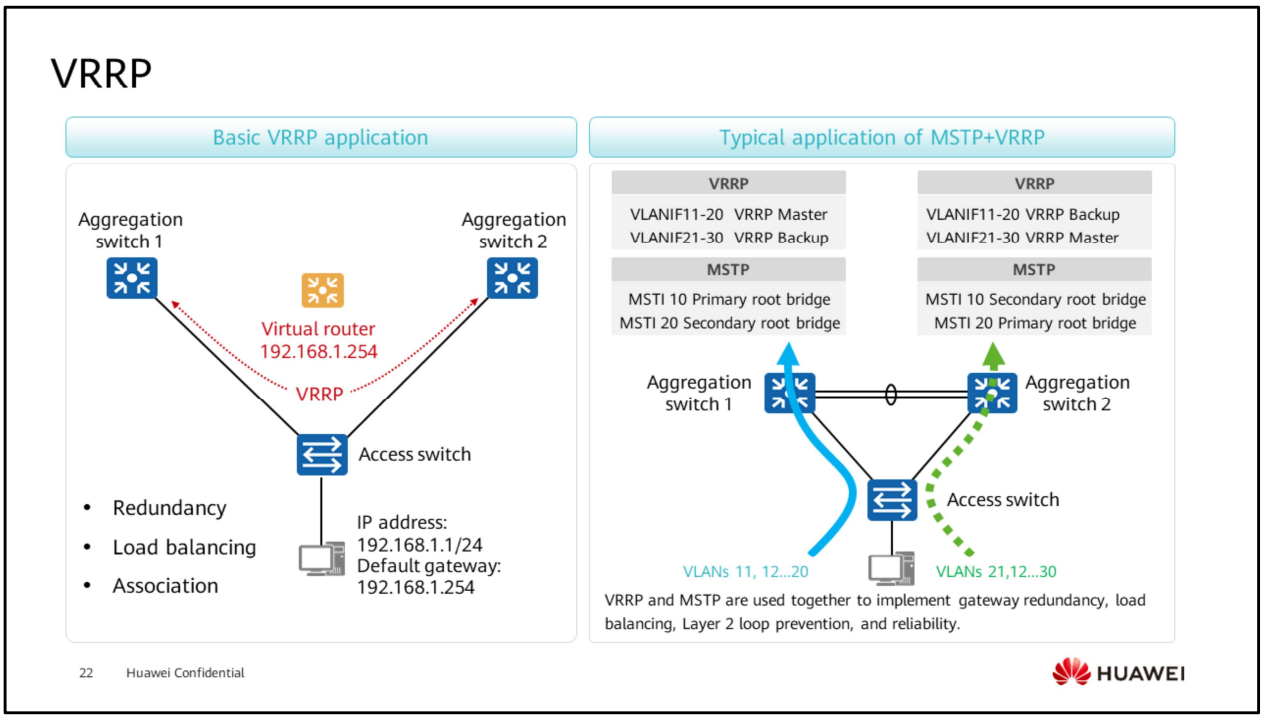
WLAN Network Architectures (4)

Native AC



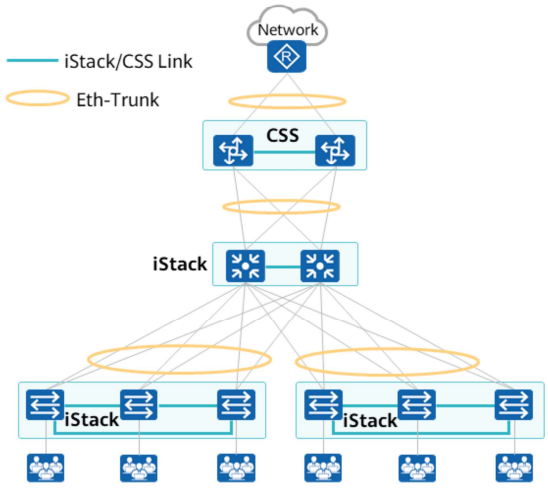
- **Networking characteristics**: This architecture uses the native AC function of switches to manage APs on a network, implementing wired and wireless network convergence. This architecture also leverages big data and AI technologies to implement simplified, smart, and secure campus networks.
- **Application scenarios**: large- and medium-sized enterprises

HUAWEI



- Generally, all hosts on the same network segment are configured with the same default route with the gateway address as the next-hop address. The hosts use the default route to send packets to the gateway, which then forwards the packets to other network segments, enabling hosts to communicate with external networks. If the gateway fails, hosts using this gateway address as the next hop of their default route cannot communicate with external networks.
- The Virtual Router Redundancy Protocol (VRRP) virtualizes several routing devices into a virtual router and uses the IP address of the virtual router as the default gateway address for the communication between users and external networks. If a gateway fails, VRRP selects another gateway to forward traffic, thereby ensuring reliable communication.
    - Redundancy: Multiple routing devices enabled with VRRP constitute a VRRP group and the VRRP group is used as the default gateway. When a single point of failure (SPOF) occurs, services are transmitted through the backup link. This reduces the possibility of network faults and ensures non-stop transmission of services.
    - Load balancing: VRRP enables multiple available routers to share the load, reducing the traffic burden on the master.
    - Association: VRRP can monitor faults on uplinks. When the uplink interface or uplink is faulty, the priority of the original master decreases, and an optimal backup becomes the master, ensuring proper traffic forwarding. Association between VRRP and BFD speeds up the active/standby switchover. To speed up the active/standby switchover in the VRRP group, configure a BFD session between the master and backup and associate the BFD session with the VRRP group. This is because BFD can fast detect faults. When the link between the master and backup becomes Down, the backup immediately switches to the master and takes over traffic.

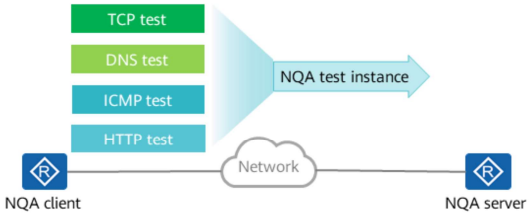
Link Aggregation, iStack, and CSS



- Ethernet link aggregation, also known as Eth-Trunk, bundles multiple physical links into a logical link to increase link bandwidth, without having to upgrade hardware.
- Intelligent stack (iStack) enables multiple stacking-capable switches to function as a single logical switch. iStack is applicable to Huawei fixed switches.
- A cluster switch system (CSS), also known as a cluster, combines two clustering-capable switches into a single logical switch. CSS is applicable to Huawei modular switches.
- Link aggregation can be used with iStack/CSS to implement link-level and device-level reliability and increase network bandwidth.

HUAWEI

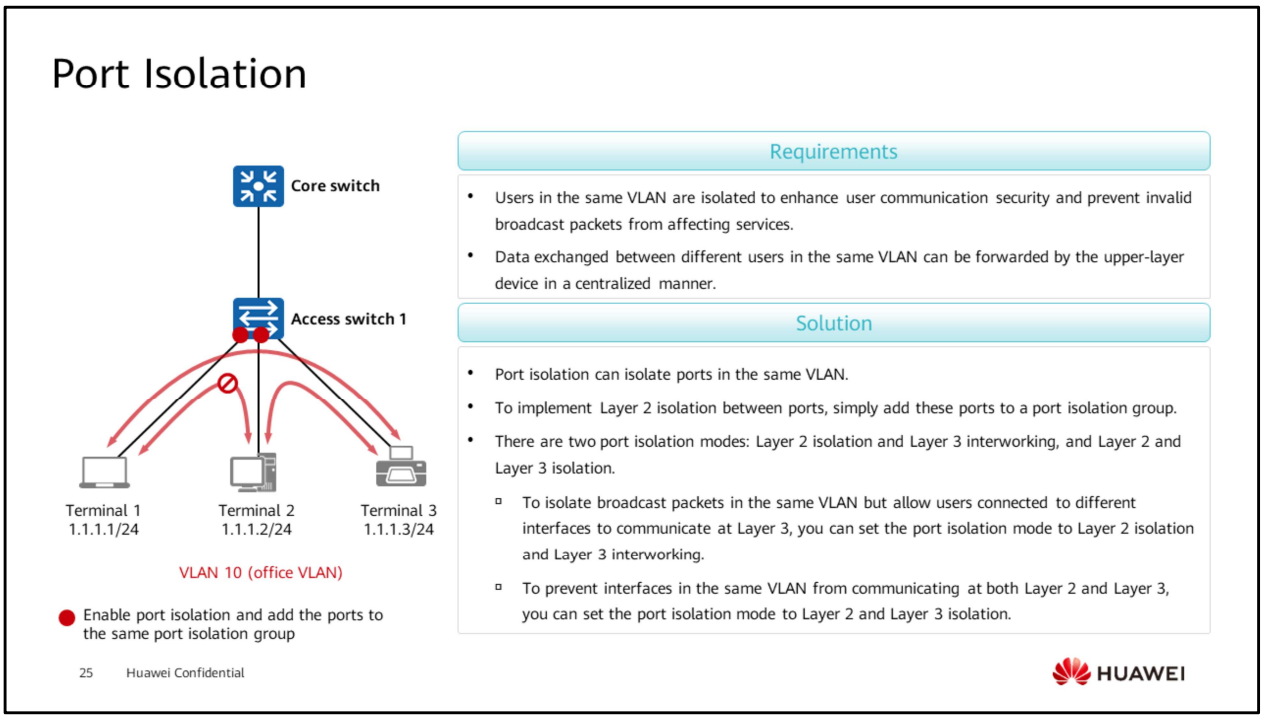
Network Quality Analysis (NQA)



- To visualize the quality of network services and allow users to check whether the quality of network services meets requirements, the following measures must be taken:
  - Enable the device to provide network service quality information.
  - Deploy probe devices to monitor network service quality.
- The preceding measures require devices to provide statistical parameters such as the delay, jitter, and packet loss ratio and require dedicated probe devices. These requirements increase investments on devices.
- NQA can precisely test the network operating status and output statistics without using dedicated probe devices, effectively reducing costs.
- NQA measures network performance and collects statistics on the delay, jitter, and packet loss ratio in real time.

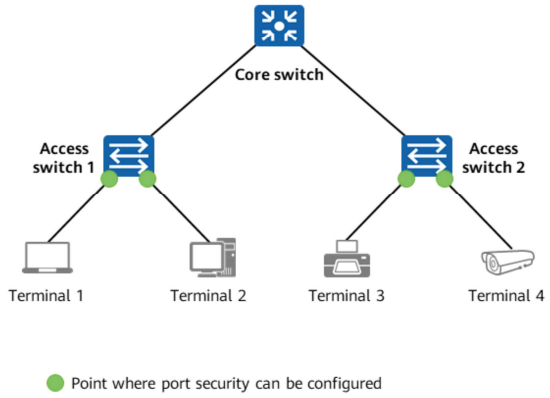
HUAWEI

- NQA measures the performance of various protocols running on networks, which helps users collect statistics about network operation indexes in real time.



- Port isolation provides more secure and flexible networking solutions.
- In some scenarios, data exchanged between terminals in a VLAN needs to be forwarded by an upper-layer device instead of an access switch. This ensures that traffic is forwarded by the upper-layer device instead of the access switch, so that traffic management and control policies can be deployed on the upper-layer device. This mode is called centralized forwarding mode. In this mode, port isolation is configured for all downstream Layer 2 devices. Therefore, proxy ARP should be configured on the gateway. In most cases, intra-VLAN proxy ARP is used.

Ethernet Port Security



Requirements

- An enterprise requires that each access switch interface connected to terminals allow only one PC to access the network so as to limit the number of MAC addresses.
- If an employee attempts to connect a small switch or hub to an interface to increase the number of Internet access interfaces, this behavior should be detected and prohibited.
- Some enterprises may require that a switch forward only data frames sent by terminals with trusted MAC addresses to the upper-layer network. Employees are not allowed to change their locations, that is, change access interfaces of the switch.

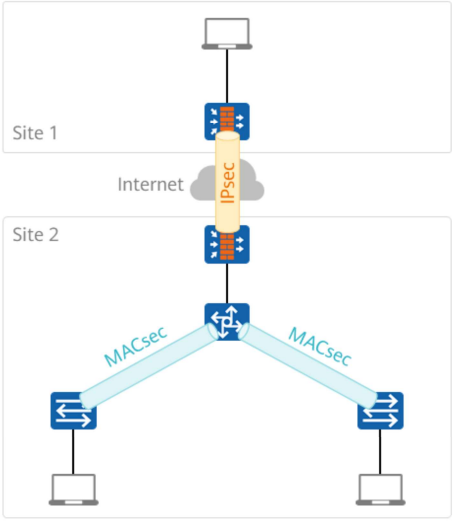
Solution

Port security converts the dynamic MAC addresses learned on an interface into secure MAC addresses (including dynamic and static secure MAC addresses, and sticky MAC addresses). This function prevents unauthorized users from using this interface to communicate with the switch, thereby enhancing device security.

HUAWEI

- Configuring port security on an interface of a switch can limit the number of MAC addresses learned by the interface. Then punishment measures can be taken when a violation occurs.
- The interface configured with port security can convert the learned MAC addresses into secure MAC addresses, preventing devices with other MAC addresses from accessing the network through the interface.

Media Access Control Security (MACsec)



Background

Most data is transmitted in plaintext on LAN links, failing to meet security requirements in scenarios demanding high security.

MACsec overview

Media Access Control Security (MACsec), in compliance with the IEEE 802.1AE standard, defines a method for secure data communication based on the Ethernet. It provides hop-by-hop encryption to secure data transmission.

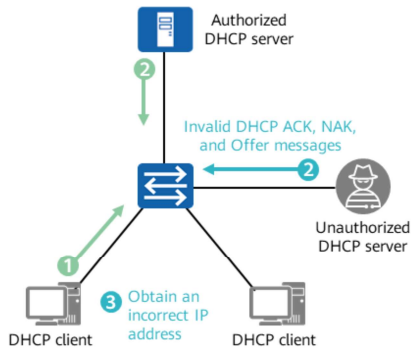
- Data integrity check
- User data encryption
- Data source authenticity verification
- Replay protection

Typical application scenarios

- MACsec is deployed between switches to ensure data security. For example, MACsec is deployed between access switches and uplink aggregation or core switches.
- When transmission devices exist between switches, MACsec can be deployed to ensure data security.

HUAWEI

DHCP Snooping

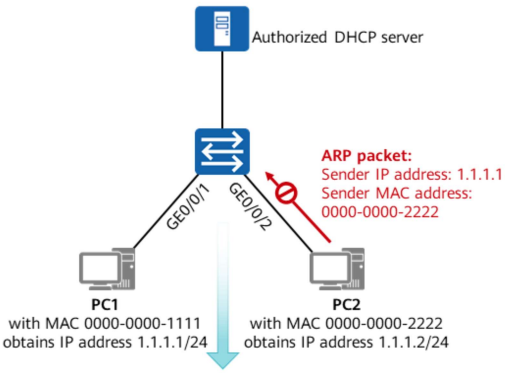


- Some attacks are launched against DHCP on the network, including bogus DHCP server attacks, DHCP server DoS attacks, and bogus DHCP message attacks.
- DHCP snooping ensures that DHCP clients obtain IP addresses from authorized DHCP servers and records mappings between IP addresses and MAC addresses of DHCP clients, preventing DHCP attacks on the network.

HUAWEI

- The DHCP snooping-enabled device forwards DHCP Request messages of users (DHCP clients) to an authorized DHCP server through the trusted interface, and then generates DHCP snooping binding entries based on the DHCP ACK messages received from the DHCP server. When receiving DHCP messages from users through the DHCP snooping-enabled interfaces, the device checks the messages against the binding table, thereby preventing attacks initiated by unauthorized users. In addition, DHCP snooping supports multiple security features, such as limiting the rate for sending DHCP messages.

Dynamic ARP Inspection (DAI)



IP Address	MAC Address	VLAN ID	Interface
1.1.1.1	0000-0000-1111	10	GE0/0/1
1.1.1.2	0000-0000-2222	10	GE0/0/2

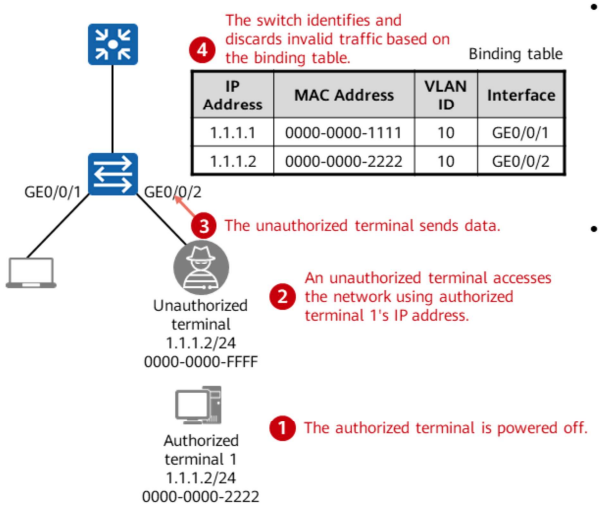
DHCP snooping binding table

- ARP attacks are frequently launched on networks. A man-in-the-middle (MITM) attack is a common ARP spoofing attack.
- To defend against MITM attacks, configure DAI on a switch.
- DAI defends against MITM attacks using a DHCP snooping binding table. When a switch receives an ARP packet, it compares the source IP address, source MAC address, VLAN ID, and interface number of the ARP packet with those in DHCP snooping binding entries.
  - If the ARP packet matches a binding entry, the switch considers the ARP packet valid and allows the packet to pass through.
  - If the ARP packet does not match any binding entry, the switch considers the ARP packet invalid and discards the packet.

HUAWEI

- An MITM attacker establishes independent connections with two parties that intend to communicate and relays messages between them. The two parties consider that they are directly communicating with each other over a private connection, but the entire conversation is in fact controlled by the attacker. In an MITM attack, the attacker can intercept all packets exchanged between the two parties and insert new ones.
- To defend against MITM attacks, configure DAI on a switch.
  - DAI defends against MITM attacks using a DHCP snooping binding table. When the switch receives an ARP packet, it compares the source IP address, source MAC address, VLAN ID, and interface number of the ARP packet with those in DHCP snooping binding entries. If the ARP packet matches a binding entry, the switch considers the ARP packet valid and allows the packet to pass through. If the ARP packet does not match any binding entry, the switch considers the ARP packet invalid and discards the packet.
  - DAI is available only when DHCP snooping is configured. A DHCP snooping-enabled switch automatically generates DHCP snooping binding entries when DHCP users go online. If a user is configured with a static IP address, you need to manually configure a static binding entry for the user.
  - When an attacker connected to the DAI-enabled switch sends bogus ARP packets, the switch detects the attack based on the binding entries and discards the bogus ARP packets. If the packet discarding alarm function is also enabled on the DAI-enabled switch, the switch will generate an alarm when the number of ARP packets discarded due to failure to match any binding entry exceeds the alarm threshold.

IP Source Guard (IPSG)

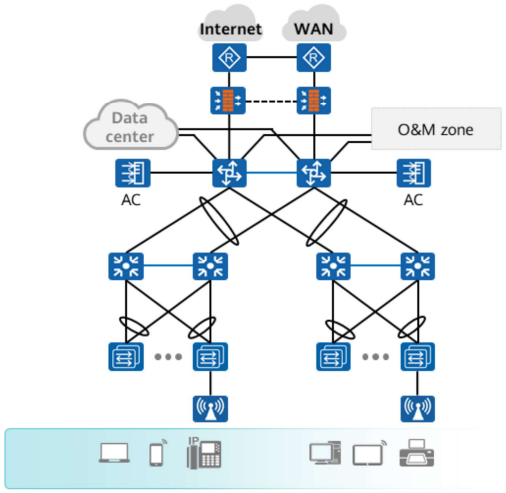


- IPSG implements source IP address filtering based on Layer 2 interfaces. IPSG prevents malicious hosts from simulating authorized hosts using forged IP addresses. In addition, IPSG prevents unauthorized hosts from accessing or attacking networks using forged IP addresses.
- IPSG checks IP packets on Layer 2 interfaces against a binding table that contains the bindings of source IP addresses, source MAC addresses, VLAN IDs, and inbound interfaces. Only packets matching the binding table are forwarded, and other packets are discarded.

HUAWEI

- As networks continue to increase in scale, a growing number of attackers are forging source IP addresses to initiate network attacks (IP address spoofing attacks). Some attackers forge IP addresses of authorized users to obtain network access rights and access networks. As a result, authorized users are unable to access networks or sensitive information may be intercepted.
- IPSG provides a mechanism to effectively defend against IP address spoofing attacks.
    - The binding table can be a static binding table or a dynamic DHCP snooping binding table.
    - IPSG only checks the IP packets from hosts. It does not check non-IP packets such as ARP packets.

Network Admission Control (NAC)



Authentication server iMaster NCE

Authentication device

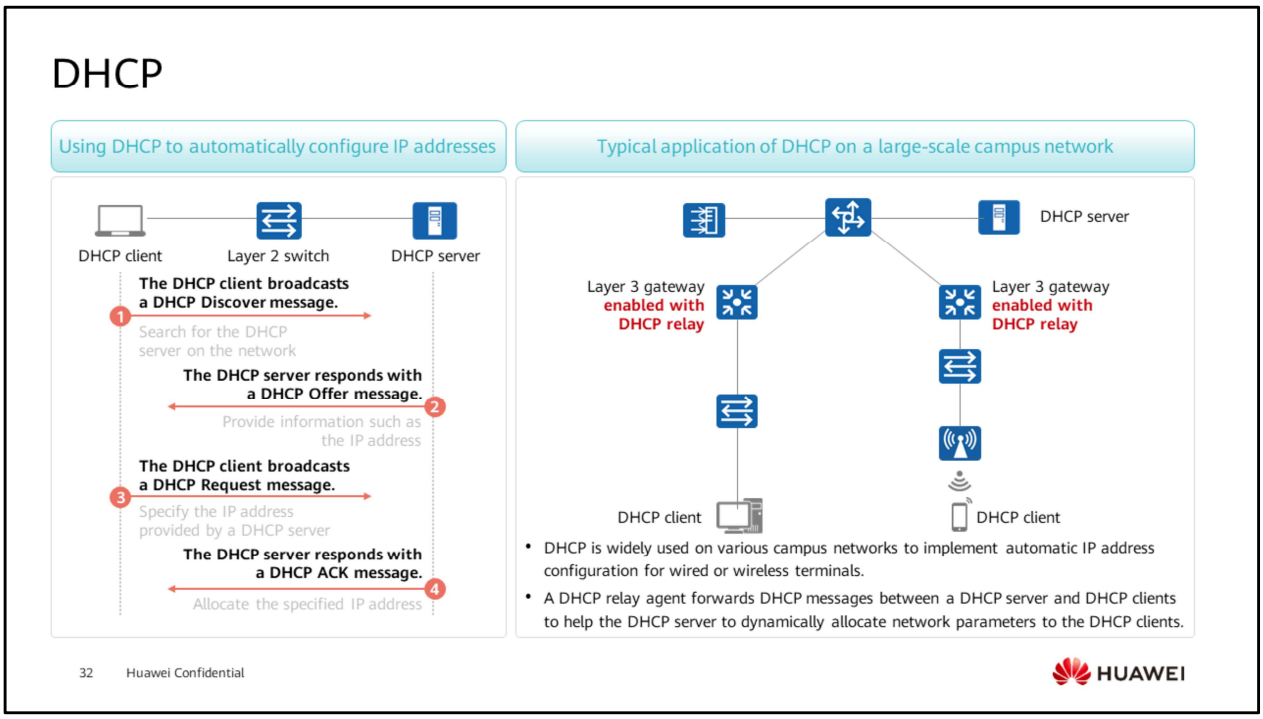
User terminal

- NAC is an end-to-end security technology that authenticates clients and users to ensure network security.
- NAC provides three authentication modes: 802.1X authentication, MAC address authentication, and Portal authentication.

- How to ensure that terminals are valid and secure?
- How to ensure that authorized terminals are authorized correctly?

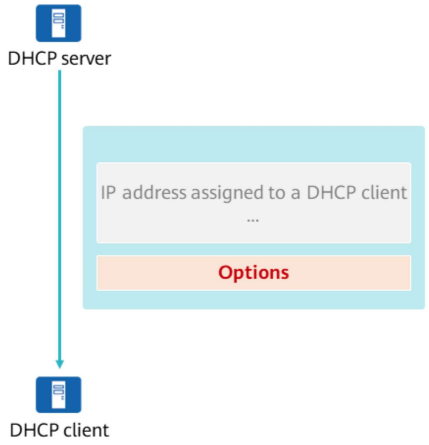
HUAWEI

- In addition to using NAC to authenticate access users and control their rights, a campus network also needs to authenticate and control rights of administrators (also called login users) who can log in to devices through FTP, HTTP, SSH, Telnet, or console ports.



- DHCP dynamically configures and uniformly manages IP addresses of hosts. DHCP is defined in RFC 2131 and uses the client/server communication mode. A DHCP client requests configuration information from a DHCP server, and the DHCP server returns the configuration information allocated to the DHCP client.
  - Instead of statically specifying an IP address for a host, DHCP enables a host to obtain an IP address dynamically.
  - DHCP can allocate other configuration parameters, such as the startup configuration file of a client, so that the client can obtain all the required configuration information by using only one message.
  - DHCP supports dynamic and static IP address allocation. A network administrator can select different address allocation modes for hosts as required.
    - Dynamic allocation: DHCP allocates an IP address with a limited validity period (known as a lease) to a client. This mechanism applies to scenarios where hosts temporarily access the network or the number of idle IP addresses is less than the total number of hosts that do not require permanent connections.
    - Static allocation: DHCP allocates fixed IP addresses to clients. Compared with manual IP address configuration, DHCP static allocation prevents manual configuration errors and enables unified maintenance and management.
- DHCP has the following benefits:
  - Reduced client configuration and maintenance cost
  - Centralized management of limited IP addresses

DHCP Options



Options field in DHCP messages

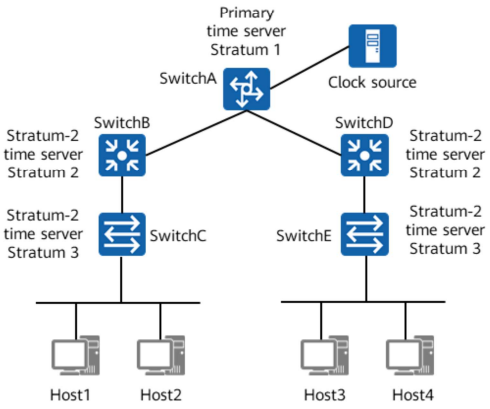
- The Options field stores control information and parameters allocated to a DHCP client.
- A DHCP message can carry multiple options.
- DHCP predefines a large number of options for various purposes, for example, setting the gateway address and DNS server address.

Typical application scenarios

- The Options field can be used to set the IP address of the NTP server.
- The Options field can be vendor-defined. For example, the CloudCampus solution uses Option 148 to implement plug-and-play of network devices. The information is used to notify the DHCP client of controller information so that the device can obtain the controller address and port number.
- Option 43 is used to register an AP with an AC that resides on a different network segment from the AP in WLAN scenarios.

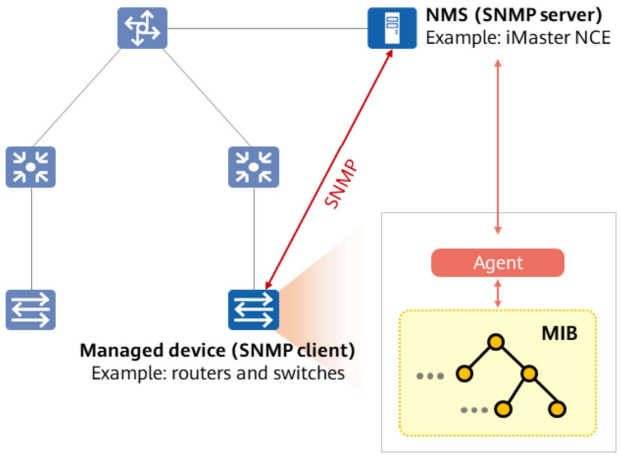
HUAWEI

Network Time Protocol (NTP)



- NTP is an application layer protocol in the TCP/IP suite that synchronizes time between time servers and clients. NTP is implemented based on IP and UDP. NTP packets are transmitted over UDP using port 123.
- Key concepts in the NTP network architecture:
  - **Synchronization subnet:** consists of the primary time server, stratum-2 time servers, PC clients, and transmission paths connecting them.
  - **Primary time server:** directly synchronizes its clock with a standard reference clock through a cable or radio. Typically, the standard reference clock is either a radio clock or the Global Positioning System (GPS).
  - **Stratum-2 time server:** synchronizes its clock with either the primary time server or other stratum-2 time servers on the network. Stratum-2 time servers use NTP to transmit time information to other hosts in a LAN.
  - **Stratum:** is a hierarchical standard for clock synchronization. It represents the precision of a clock. The value of a stratum ranges from 1 to 16. A smaller value indicates higher precision. The value **1** indicates the highest clock precision, and the value **16** indicates that the clock is not synchronized.

SNMP



Overview of SNMP

- SNMP is a network management protocol widely used on TCP/IP networks.
- SNMP provides a method for managing devices through a central computer that runs network management software — known as a network management station (NMS).
- Using in-band management, SNMP achieves efficient and batch network device management. In addition, SNMP enables unified management of different types of network devices from different vendors.

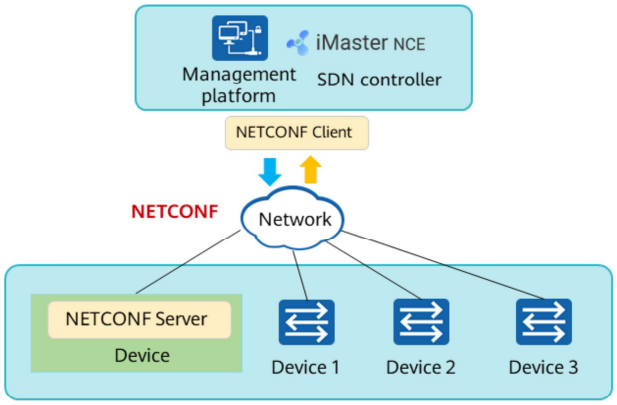
HUAWEI

Network Configuration Protocol (NETCONF)

- NETCONF provides a set of mechanisms for managing network devices. To be specific, users can use NETCONF to add, modify, and delete configurations of network devices, as well as obtain configurations and status of network devices.

NETCONF has three objects:

- NETCONF client
- NETCONF server
- NETCONF message



HUAWEI

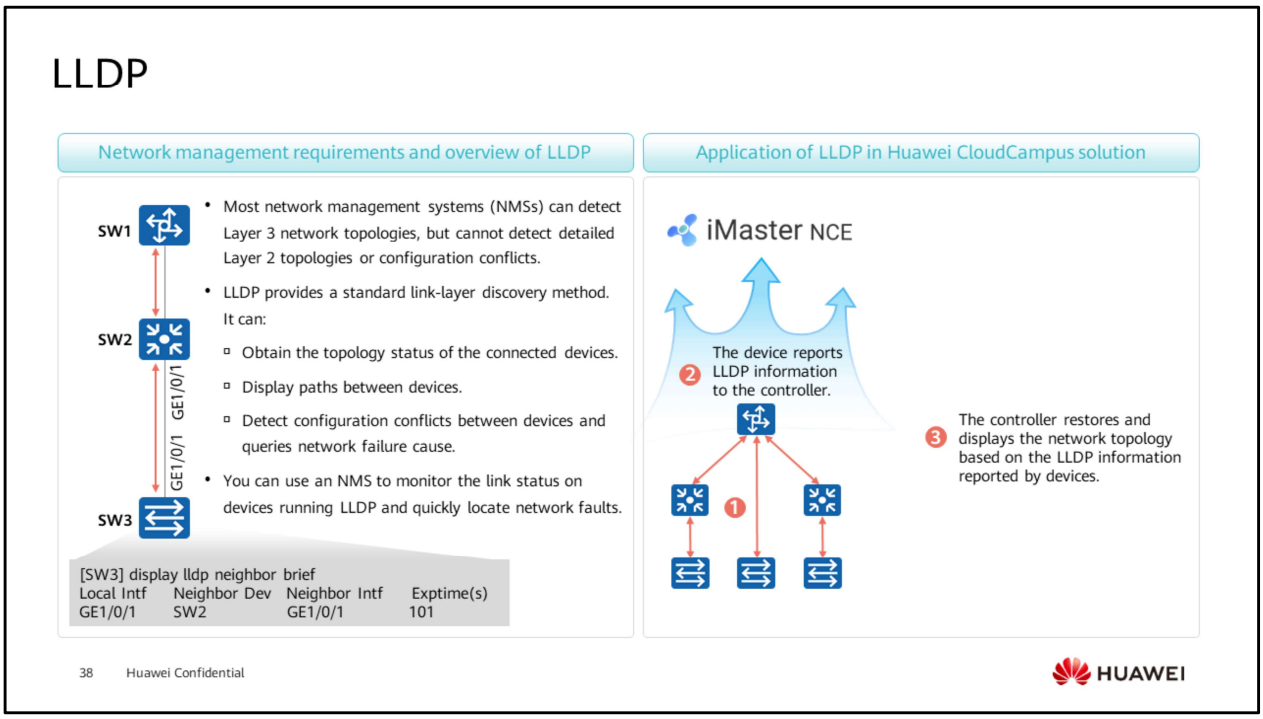
- The plug-and-play of a device is implemented by establishing a NETCONF session between the device and the controller, so that the controller can deliver configurations to the device.

Basic Operations of NETCONF

- NETCONF defines a series of operations:

Scenario	Operation	Function
Query data	<get-config>	Queries configuration data.
	<get>	Queries the current configuration and status data of the device.
Edit data	<edit-config>	Creates, modifies, or deletes configuration data.
Backup/Restoration	<copy-config>	Exports configuration data, or replaces one configuration datastore with another configuration datastore.
	<delete-config>	Deletes the configuration datastore and clears the startup configuration.
Locking/Unlocking	<lock>	Exclusively occupies the permission to modify the configuration datastore.
	<unlock>	Releases the exclusive permission to modify the configuration datastore.
Transactional operations	<commit>	Copies the candidate configuration to the device's running configuration.
	<cancel-commit>	Cancels an ongoing confirmed commit.
	<discard-changes>	Discards uncommitted changes to the candidate configuration.
	<validate>	Checks whether the syntax and semantics of the specified configuration data are correct.
Session operations	<close-session>	Normally ends the NETCONF session.
	<kill-session>	Forcibly terminates other NETCONF sessions. You must have the administrator permission to perform this operation.

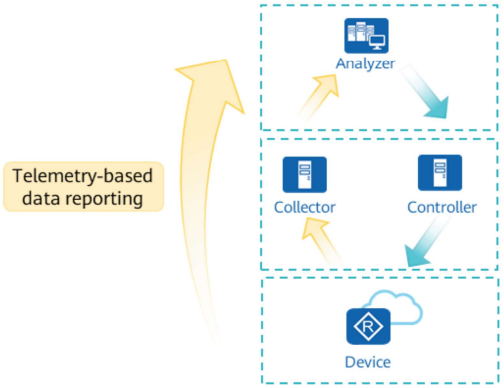
HUAWEI

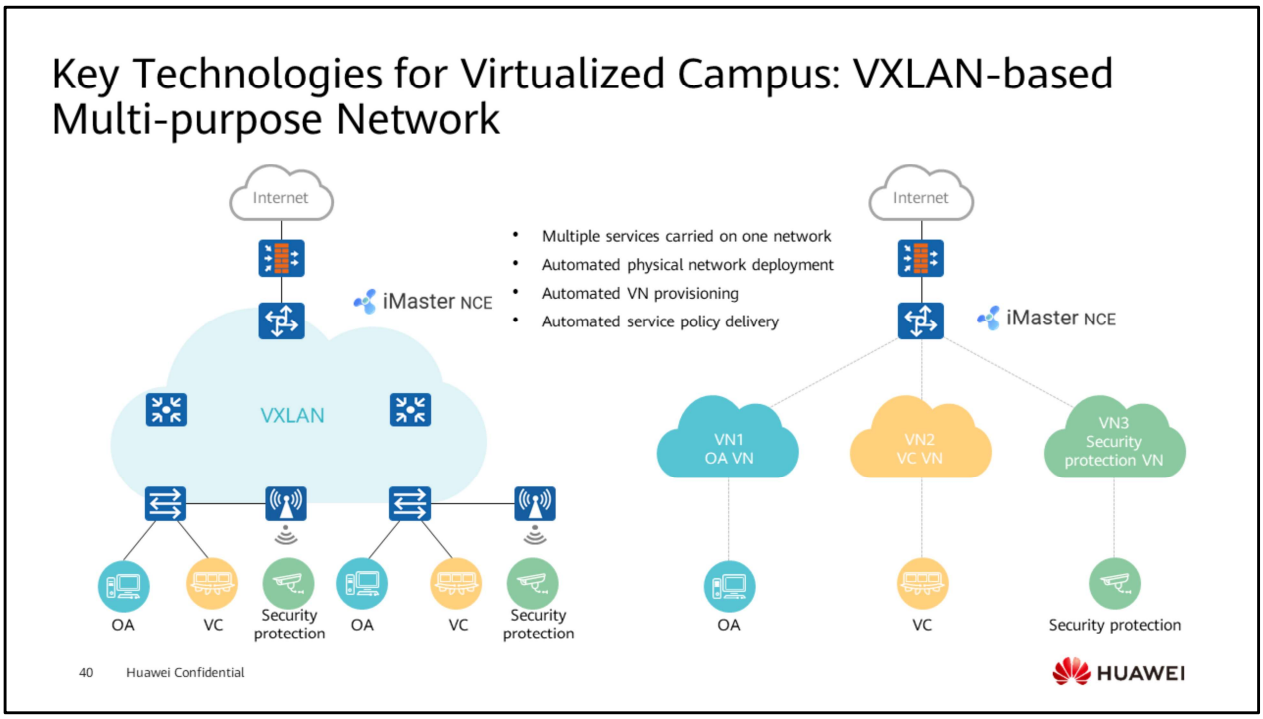


- LLDP is a standard Layer 2 topology discovery protocol defined in IEEE 802.1ab. LLDP collects local device information including the management IP address, device ID, and port ID and advertises the information to neighboring devices. Neighboring devices save the received information in their management information bases (MIBs). The NMS can query required information in MIBs to determine link status.
- As networks grow in scale, network devices are increasing in diversity with complex configurations, posing higher requirements on network management capabilities. Most NMSs can detect Layer 3 network topologies, but they cannot detect detailed Layer 2 topologies or configuration conflicts. To address the network management problems, a standard protocol is required to exchange Layer 2 information between network devices.
- LLDP provides a standard link-layer discovery method. By obtaining Layer 2 information from devices, LLDP allows users to detect the topology of neighboring devices, and display paths between clients, switches, routers, application servers, and network servers. The information also helps detect configuration conflicts between network devices and identify causes of network failures. Enterprise users can use an NMS to monitor the link status on devices running LLDP and quickly locate network faults.

Telemetry

- Telemetry, also called network telemetry, is a technology that remotely collects data from physical or virtual devices at a **high** speed. Devices periodically send information such as interface traffic statistics, CPU usage, and memory usage to collectors in push mode. Compared with the conventional pull mode (request-response interaction), the push mode provides faster and real-time data collection.

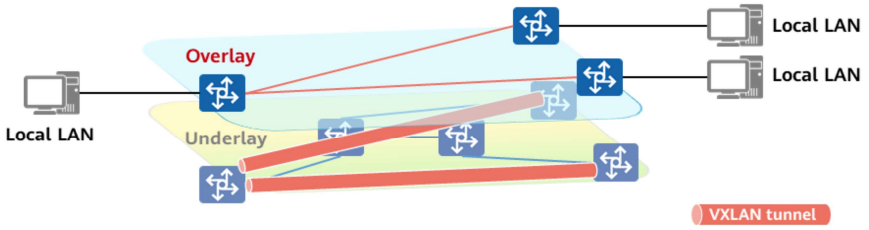




- Using virtualization technologies, multiple virtual networks (VNs) can be created on top of a physical network. Different VNs are used for different services, such as OA, videoconferencing, and security protection.

VXLAN

- VXLAN is essentially a VPN technology that is used to build a Layer 2 virtual network (overlay network) on top of any physical network (underlay network) that is reachable at Layer 3. VXLAN tunnels can be built between VXLAN gateways to implement communication within a VXLAN network as well as communication between a VXLAN network and a non-VXLAN network.
- VXLAN uses MAC-in-UDP encapsulation to extend Layer 2 networks and encapsulate Ethernet packets into IP packets. VXLAN can be deployed together with BGP EVPN and VRF to implement multi-purpose campus network, on which different VNs are isolated from each other.

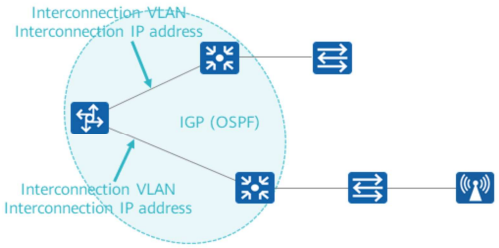


HUAWEI

Underlay and Fabric

Underlay network

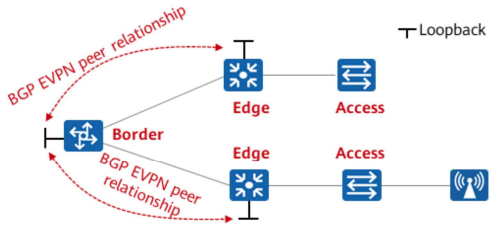
The underlay network is the foundation of the entire virtual campus network. It is a physical network consisting of physical network devices and provides interoperability for all services on a campus network.



- Uses a multi-layer architecture (core, aggregation, and access).
- Has multiple topologies, such as tree, ring, and mesh.
- Underlay information includes device interconnection VLANs, interconnection IP addresses, and IGP.

Fabric network

A campus fabric is a network resource pool that is abstracted from physical network devices on the underlay network so as to create a multiple-purpose network.

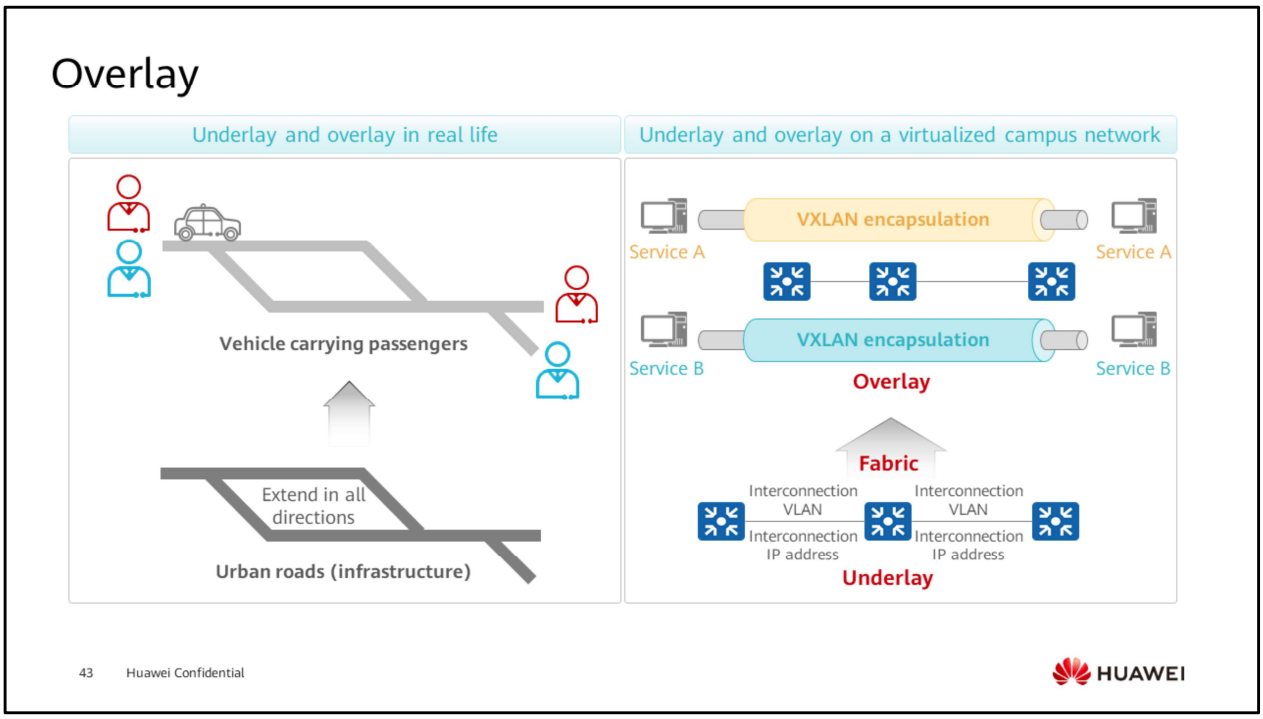


A fabric includes the following:

- Overlay network resource pool (BD IDs and VNIs) for terminal access
- VLAN ID pool for terminal access
- Access point pool (switch ports for wired access or SSIDs for wireless access) for terminals

These resource pools can be used to create multiple VNs.

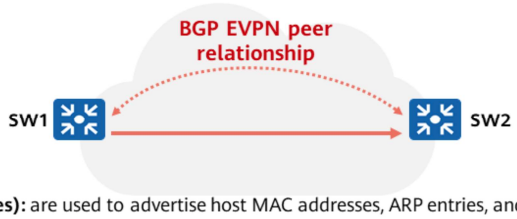
HUAWEI



- An overlay network is a logical network that is built on top of a physical network, uses the tunneling technology, and separates the forwarding plane from the control plane.

BGP EVPN

- BGP EVPN extends BGP by defining several new types of BGP EVPN routes using Network Layer Reachability Information (NLRI) in the MP_REACH_NLRI attribute.
- These BGP EVPN routes can be used to transmit VTEP addresses and host information. Therefore, when BGP EVPN is used on a VXLAN network, VTEP discovery and host information learning can be transferred from the data plane to the control plane.

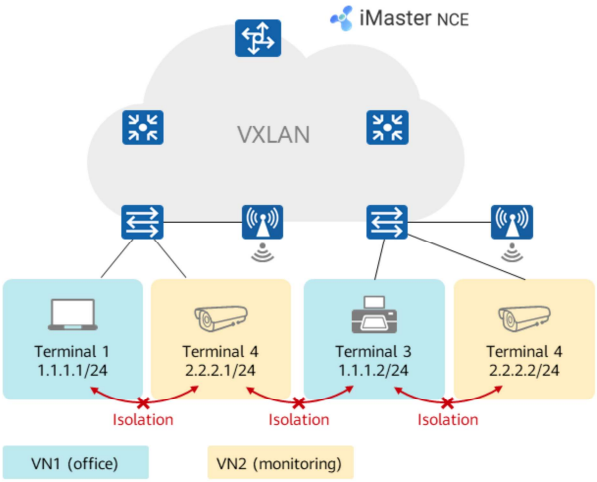


- **Type 2 routes (MAC/IP routes):** are used to advertise host MAC addresses, ARP entries, and IP routes.
- **Type 3 routes (inclusive multicast routes):** are used to transmit Layer 2 VNI and VTEP IP address information, so as to implement automatic VTEP discovery, dynamic VXLAN tunnel establishment, and BUM packet forwarding.
- **Type 5 routes (IP prefix routes):** are used to advertise host MAC addresses, ARP entries, and IP routes, as well as external network routes.

HUAWEI

- In a network virtualization overlay (NVO) scenario, BGP EVPN is used together with VXLAN and functions as the control plane protocol for VXLAN.

Using VRF for VN Isolation



Requirements

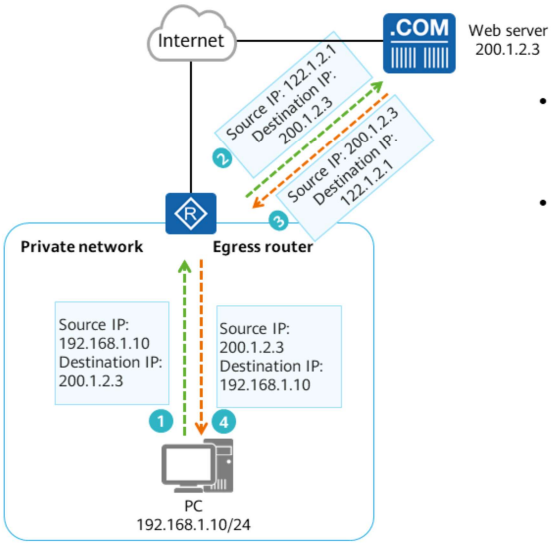
- Multiple logical networks, such as the office network and monitoring network, can be built on a simple physical campus network.
- The office network is completely isolated from the monitoring network.

Solution

- Deploying ACLs: ACLs can be used to isolate traffic between VLANs. However, ACL configuration is complex and ACLs do not apply to all scenarios.
- Build independent physical networks: Separate physical networks can be built for office and monitoring purposes, so services can be isolated. However, this increases network construction costs.
- Virtual Routing and Forwarding (VRF): It is also called a VPN instance, which is similar to a virtual device. A VPN instance uses a routing table independent of the root device to completely isolate the VPN instance from the root device, without increasing hardware costs.

HUAWEI

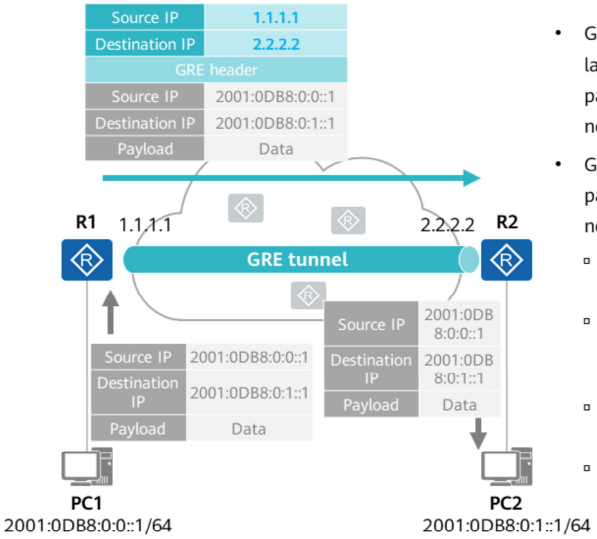
Network Address Translation (NAT)



- NAT technology translates IP addresses in IP data packets. It is widely used on live networks and is usually deployed on network egress devices, such as routers or firewalls.
- Typical application scenarios:
  - When users with private IP addresses attempt to access the Internet, NAT deployed on the egress device of the campus network can translate the source IP addresses of data packets sent from the intranet to the Internet into specific public IP addresses.
  - The campus network provides some services for Internet users. When Internet users access these services using public IP addresses, NAT can translate the destination IP addresses of data packets from the Internet.

HUAWEI

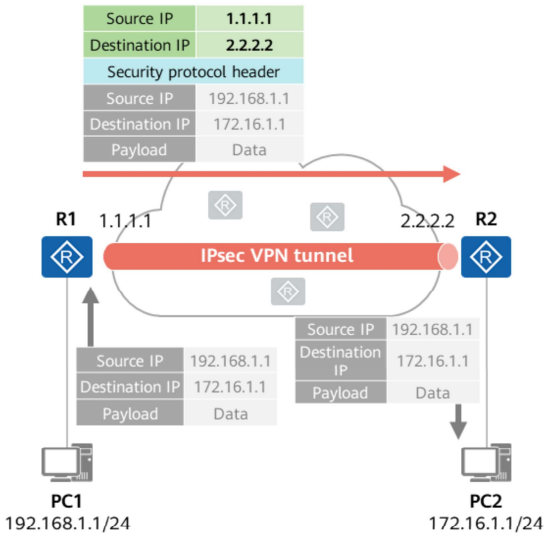
Generic Routing Encapsulation (GRE)



- GRE is a protocol that encapsulates data packets of some network layer protocols, such as IPX, IPv6, and IPv4. The encapsulated data packets can then be transmitted over a network using a different network layer protocol, such as IPv4 and IPv6.
- GRE is a Layer 3 tunneling technology that transparently transmits packets over GRE tunnels. It solves the transmission problems on the network that uses different network-layer protocols.
  - GRE is easy to implement and increases only a few loads on devices at both ends of a tunnel.
  - GRE sets up tunnels over an IPv4 network to connect networks running different protocols, reusing the original network architecture and reducing costs.
  - GRE connects non-contiguous subnets and sets up VPNs to ensure secure connections between the enterprise headquarters and branches.
  - GRE provides only simple password authentication and does not provide the encryption function.

HUAWEI

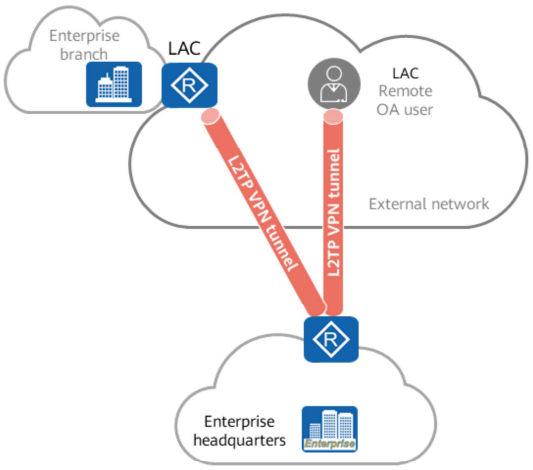
IPsec VPN



- Internet Protocol Security (IPsec) is a set of open network security protocols defined by the Internet Engineering Task Force (IETF) to secure data transmission and reduce the risk of information leakage.
- Through encryption and authentication, IPsec ensures secure service data transmission over the Internet from the following dimensions:
  - Data source authentication
  - Data encryption
  - Data integrity verification
  - Anti-replay

HUAWEI

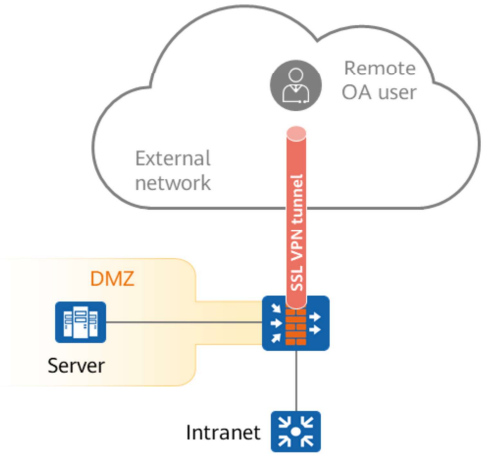
L2TP VPN



- Layer 2 Tunneling Protocol (L2TP) extends the application of the Point-to-Point Protocol (PPP). It is a VPN tunneling protocol that allows traveling employees or enterprise branches to remotely access intranet resources.
- The L2TP network architecture includes an L2TP access concentrator (LAC) and an L2TP network server (LNS).
  - An LAC provides PPP and L2TP processing capabilities, and establishes an L2TP tunnel with the LNS. LACs can be different devices in different networking environments, for example, an LAC can be a gateway or a terminal. An LAC can initiate requests for establishing multiple L2TP tunnels to isolate data flows.
  - An LNS is the peer of an LAC. That is, an L2TP tunnel is established between the LAC and the LNS. The LNS is located at the border between the private network of the enterprise headquarters and public network, and is usually the gateway of the enterprise headquarters.

HUAWEI

SSL VPN



- Secure Shell (SSL) VPN is an SSL-based remote access VPN technology. It allows mobile users (referred to as remote users in SSL VPN) to securely and conveniently access enterprise intranets and intranet resources, improving work efficiency.
- SSL VPN uses the browser/server (B/S) architecture. Remote users can directly use a web browser to access intranet resources securely and quickly, without the need to install extra client software. SSL VPN provides the following functions:
  - Fine-grained permission control based on the type of intranet resources accessed by remote users
  - Flexible identity authentication
  - Host check policies that check whether the operating systems, ports, processes, and antivirus software of remote users' devices meet security requirements. It also provides anti-nested remote desktop connection and anti-snapshot functions to eliminate security risks from remote users' devices.

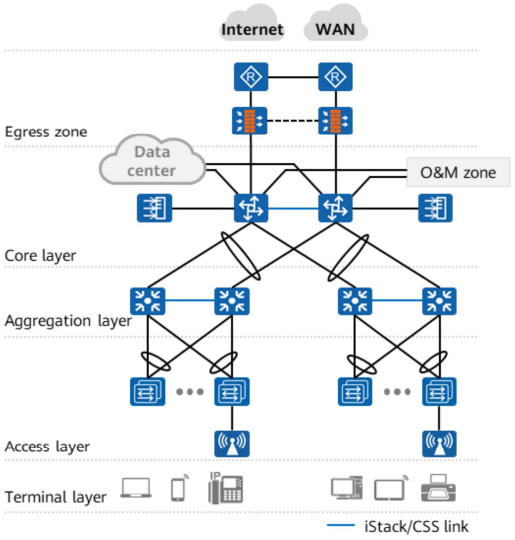
HUAWEI

Contents

1. Introduction to Campus Networks
2. Typical Campus Network Technologies
3. **Typical Applications of Campus Network Technologies**

HUAWEI

Case 1 - Traditional Campus Network



Network requirements

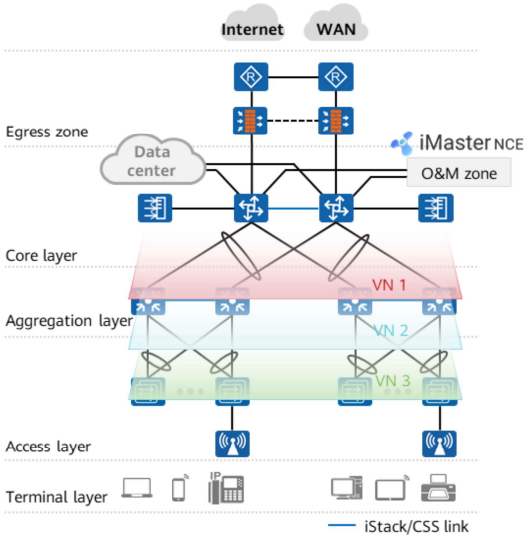
- There are about 3000 wireless and wired terminals in total, so the campus network needs to provide both wireless and wired access services.
- Network segments are divided based on departments and services. The gateways are configured on aggregation switches, and terminal addresses are automatically allocated.
- Network devices are centrally managed and monitored.
- Unified policy control is implemented for access users, and the intranet is protected from attacks from external networks.
- Network reliability must be ensured.

Solution

- Access, aggregation, and core switches are stacked or set up a CSS, and Eth-Trunks are configured on uplinks to ensure reliability.
- The AC + Fit AP wireless networking mode is used, and hot standby is configured between ACs to ensure wireless network reliability. VLANs are configured on access switches to differentiate users on different service network segments, and IP addresses are allocated to these users by the corresponding aggregation switches through DHCP. Single-area OSPF is deployed on aggregation devices and upper-layer devices.
- Access authentication is deployed for all users on the network, and corresponding policies are configured using ACLs.
- Firewalls are deployed in the egress zone and security groups are configured to control traffic.
- SNMP is enabled on all devices on the network, and the network management software is used to manage and maintain network devices in a unified manner.

HUAWEI

Case 2 - Virtualized Campus Network



Network requirements

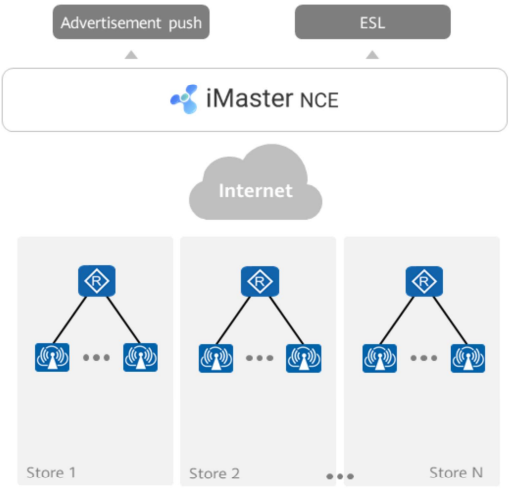
- Reliability and security requirements similar to those of traditional campus networks
- Plug-and-play and centralized management of network devices, and flexible provisioning of service configurations and policies
- Isolation of multiple services on the campus network
- Access control of terminals, and consistent network access rights and experience for users while they are moving within the campus so long as their identities remain unchanged

Solution

- iMaster NCE-Campus is used to manage network-wide devices. Protocols such as NETCONF and Telemetry are used to implement unified management and maintenance of network devices as well as quick onboarding of devices.
- VXLAN, BGP EVPN, and VRF technologies are used to divide VNs based on service requirements. iMaster NCE-Campus implements simplified deployment of the underlay and overlay networks as well as VN isolation.
- iMaster NCE-Campus is used to authenticate users on the entire network. Free mobility is deployed to define different user groups or resources on the network as different security groups and implement policy control based on security groups.

HUAWEI

Case 3 - Small- and Medium-Sized Cloud Managed Campus Network



Network requirements

- There are a large number of stores, most of which use wireless networks and have simple network topologies. In addition, onsite personnel do not have network O&M skills.
- There are many types of wireless terminals, including common wireless terminals and IoT terminals such as electronic shelf labels (ESLs).
- The stores want to push product discounts and promotion activity advertisements to customers who are connecting to the Wi-Fi network in the stores.

Solution

- Deployment by scanning barcodes achieves minute-level AP onboarding and quick provisioning of wireless network services. The cloud management platform provides one-stop management of the entire lifecycle (from planning and construction to maintenance and optimization) of all store networks.
- Deploy IoT APs and use the IoT slots built in the IoT APs to implement co-site deployment of IoT and Wi-Fi, unified planning, and shared network for data backhaul. The ESL management system connects to and interacts with the supermarket management and ERP systems to dynamically display real-time price changes and issue out-of-stock warnings.
- Customers access the network through Portal authentication, and can view customized advertisements displayed on the authentication page.

HUAWEI

Quiz

1. (Multiple-Answer Question) Which of the following technologies can be used to improve network reliability? ( )
   A. STP
   B. Eth-Trunk
   C. iStack
   D. CSS
2. (True or False) On an MSTP network, devices in the same MST region must be configured with the same VLAN-to-MSTI mappings. ( )
   A. True
   B. False

HUAWEI

1. BCD
2. A

Summary

- This course provides the definition of campus networks, common technologies used in campus networks, and application scenarios of these technologies.
- In subsequent courses, we will discuss how to design networks based on customer requirements, which requires a comprehensive and in-depth understanding of campus network technologies. Due to limited space, this document cannot list every technology used on campus networks. To better understand subsequent courses, you can read related documents for more campus network technologies and a deeper understanding of the key technologies discussed in this course.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Large- and Medium-Sized Virtualized Campus Network Design



HUAWEI

Foreword

- Huawei CloudCampus 3.0 Solution adopts a new philosophy of Intent-Driven Network (IDN), and introduces big data analytics and Artificial Intelligence (AI) technologies into cloud and software-defined networking (SDN). It helps enterprises build intelligent, simplified, converged, open, and secure networks.
- This course describes the planning and design process for large- and medium-sized virtualized campus networks that use the CloudCampus 3.0 Solution, including the underlay network design, fabric design, overlay network design, admission control and free mobility design, WLAN design, network security design, and O&M design.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the network layers and network architecture of large- and medium-sized virtualized campus networks.
  - Complete the underlay network design for a virtualized campus network based on actual requirements.
  - Complete the fabric and overlay network design for a virtualized campus network based on actual requirements.
  - Complete the admission control design, free mobility design, WLAN design, network security design, and O&M design for a virtualized campus network based on actual requirements.

Contents

1. **CloudCampus Solution and Virtualized Campus Network Overview**
2. Underlay Network Design
3. Fabric Design
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. WLAN Design
7. Network Security Design
8. O&M Design

HUAWEI

Large- and Medium-Sized Campus Networks



- Different from the wide area network (WAN) and data center network, a campus network typically refers to the internal network of an enterprise or organization. A campus network is built for more efficient running of key enterprise services.
- In terms of network scale, campus networks can be classified into large- and medium-sized campus networks as well as small- and medium-sized campus networks. A large- or medium-sized campus network typically has:
  - Over 2,000 end users
  - Over 100 network elements (NEs)
- Some enterprises have branches dispersed across different areas, and each branch network can be considered a single campus network.

HUAWEI

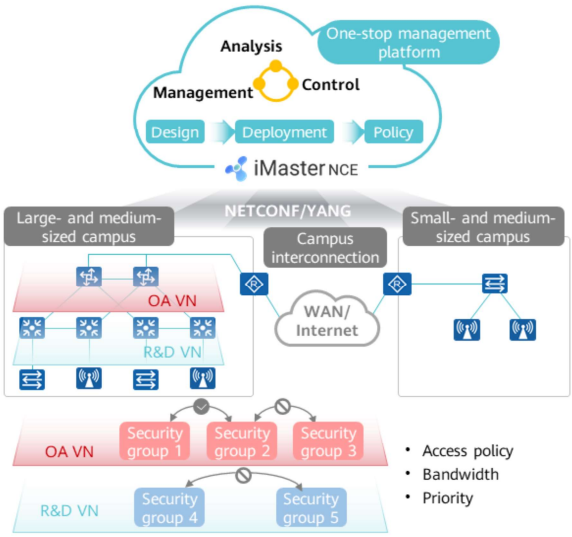
Service Requirements and Challenges of Large- and Medium-Sized Campus Networks

- Enterprises' campus networks are the cornerstone of their digital transformation. Nowadays, mobile office, cloud computing, SDN, Internet of Things (IoT), artificial intelligence (AI), and big data are gaining momentum. Driven by this, new technologies and applications are constantly emerging and are making inroads to enterprise campus networks, which poses many new challenges for the campus networks.

Converged network	Automated network deployment	User experience awareness
Requirements: Diversified access terminals and services require a converged campus network. **Challenges:** • Wi-Fi and IoT services are independently planned, deployed, and managed, resulting in high network construction costs. • The network management and O&M workload is heavy.	**Requirements:** Network deployment should be automated to address the growing complexity in deployment and policies due to the surge in applications and services. **Challenges:** • Manual configuration is repetitive, complex, and labor-intensive. • New service rollout requires configuring devices one by one, which is time-consuming and costly. • The workload of network policy deployment and adjustment is heavy.	**Requirements:** Network O&M should become automated and intelligent, with insights into user experience anytime, anywhere. **Challenges:** • Service faults cannot be detected in a timely manner. • Fault locating heavily relies on the O&M skills and experience of professionals, so faults cannot be quickly located. • The network cannot be self-optimized.

HUAWEI

Huawei CloudCampus 3.0 Solution (CloudCampus Solution)



Rapid network provisioning, improving deployment efficiency

- **Device plug-and-play**: simplified device deployment, scenario-specific guided configuration, template-based configuration
- **Simplified network deployment**: network resource pooling, multi-purpose network, automatic service provisioning

Fast service provisioning, improving user experience

- **Free mobility**: GUI-based policy configuration; access anytime anywhere, with consistent permission and experience during roaming
- **Intelligent terminal identification**: anti-spoofing for terminal access, with the accuracy of intelligent terminal identification reaching over 95%
- **Intelligent HQoS**: application-based traffic scheduling and shaping and fine-grained bandwidth management ensure service experience of VIP users

Quick intelligent O&M, improving network performance

- **Real-time experience visualization**: uses Telemetry for network experience visualization at any moment, for any user, and in any area
- **Precise fault analysis**: proactively identifies typical network issues and provides suggestions; compares and analyzes real-time data to predict faults
- **Intelligent network optimization**: predictive optimization of wireless networks based on historical data, improving overall network performance

HUAWEI

Three Deployment Modes of CloudCampus

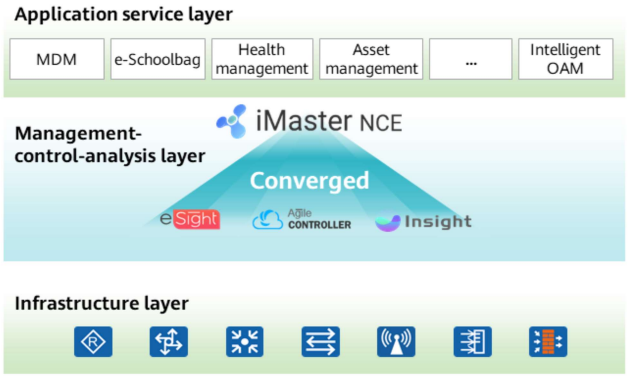
	On-Premise	Huawei Public Cloud	MSP-owned Cloud
Scenario definition	Customers purchase and own software entities, such as the controller and analyzer, which can be deployed in their data centers or on the public cloud platform.	Huawei operates the public cloud and customers do not need to purchase the controller or analyzer software. Instead, customers just purchase Huawei's cloud managed network service.	MSPs purchase software, such as the controller and analyzer, for operational purposes. The software can be deployed in their data centers or on the public cloud platform.
Target customer	Customers in industries such as government, education, large enterprise, retail, and financial services	Customers in industries such as government, education, large enterprise, retail, and financial services	MSP, carrier
Operations entity	Customer	Huawei	MSP, carrier
Software transaction mode	Perpetual license + SnS	SaaS mode	TBL subscription mode

HUAWEI

- Managed service provider (MSP): delivers and manages network-based services, applications, and devices. The serviced objects include enterprises, residential areas, and other service providers.
- Perpetual license + SnS: The perpetual license is sold together with SnS services, such as software patches, software upgrades (including new features of new releases), and remote support. In the perpetual license + SnS mode, a customer needs to pay SnS fee for a certain period of time, in addition to purchasing the license upon the first purchase. If the customer does not renew the SnS annual fee after it expires, the customer can only use functions provided in the license for the current version and cannot use the service functions covered in the SnS annual fee.
- SaaS mode: MSPs are responsible for deploying or leasing hardware infrastructure, and O&M and management of the hardware and software. Software is provided for customers as cloud services and customers need to periodically pay for the cloud services.
- Term Based License (TBL) mode: This mode differs from the perpetual license + SnS mode in that the licenses purchased by customers have limited validity periods. If a customer does not renew the subscription after the license expires, the customer can no longer use the software product.
- SnS: refers to Subscription and Support. It consists of two parts: software support and software subscription. The complete software charging mode consists of the annual software SnS fee and software license fee.
- Note: This course uses the on-premise deployment as an example.

CloudCampus Solution Components: iMaster NCE-Campus

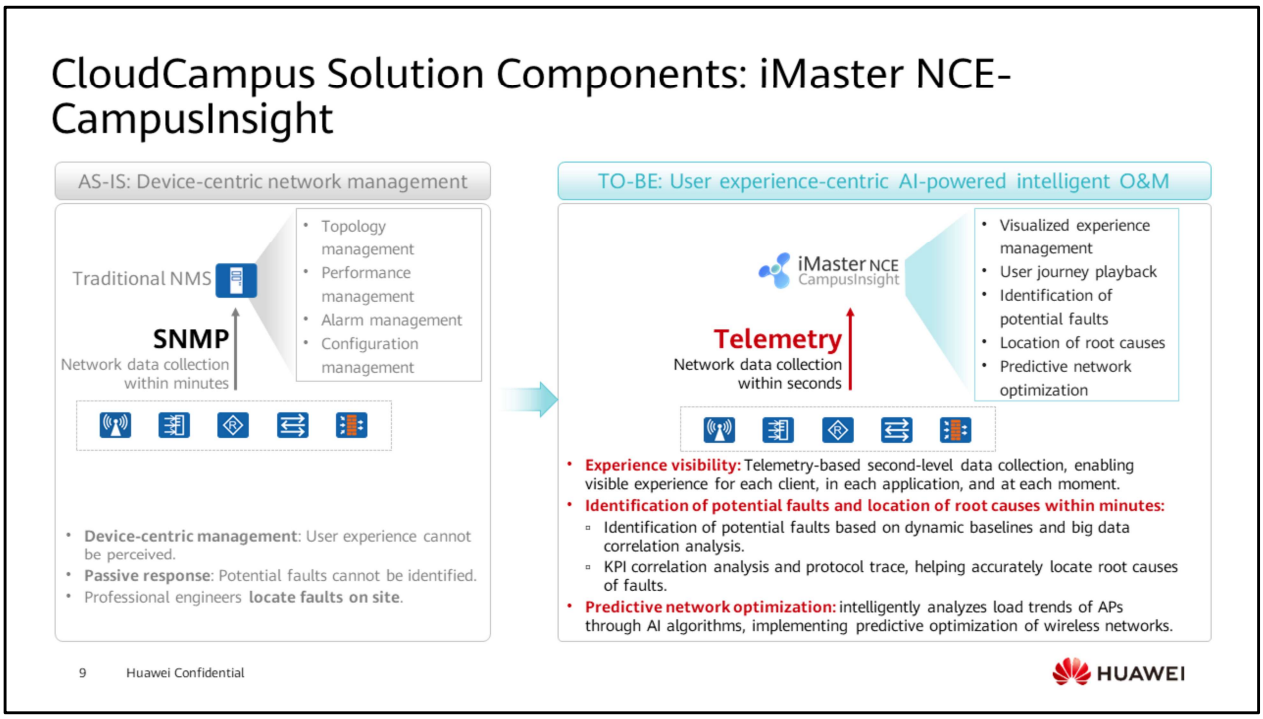
- iMaster NCE-Campus is the configuration and management platform used in the CloudCampus Solution. It provides a portal for CloudCampus service configuration, O&M, and monitoring.



Automated + Intelligent	• SDN-based automatic service configuration/deployment • AI-powered intelligent analysis/prediction/troubleshooting
Manage + Control + Analyze	• Unified data base • Centralized detection/locating/processing
Plan + Construct + Maintain + Optimize	• Full lifecycle management • Simulation/Verification/Monitoring/Optimization

iMaster NCE-Campus, an autonomous driving campus network management and control system

HUAWEI



- Using algorithms to improve efficiency and leveraging scenario-specific continuous learning and expert experience, intelligent O&M frees O&M personnel from nerve-wracking alarms and noises, making O&M automated and intelligent.

CloudCampus Solution Components: Network Hardware Products

CloudEngine S series switches



High-performance campus switches

NetEngine AR routers



Campus egress routers
Integrate routing, switching, Wi-Fi, 5G, and security functions

AirEngine Wi-Fi 6 APs



Campus APs
Models that apply to all scenarios: indoor settled, indoor wall plate, agile distributed, and outdoor AP models

HiSecEngine AI firewalls

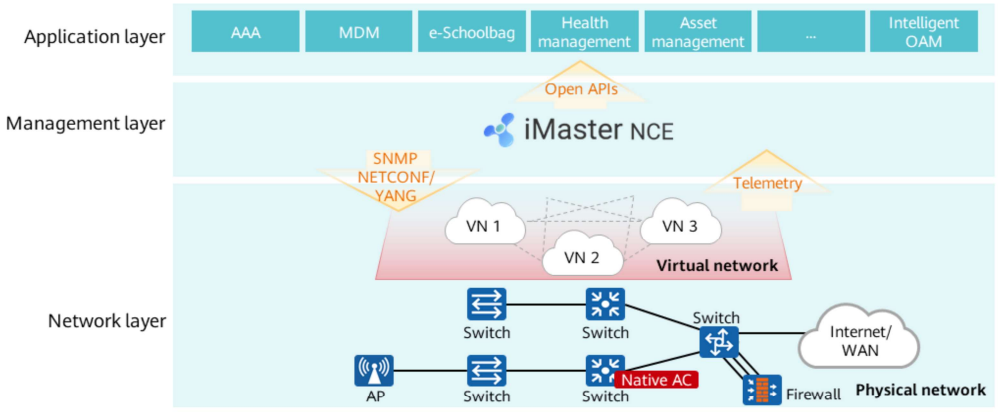


Firewalls
Provide comprehensive and integrated network security protection capabilities

HUAWEI

Network Architecture of the CloudCampus Solution (VXLAN-based Virtualized Campus Network)

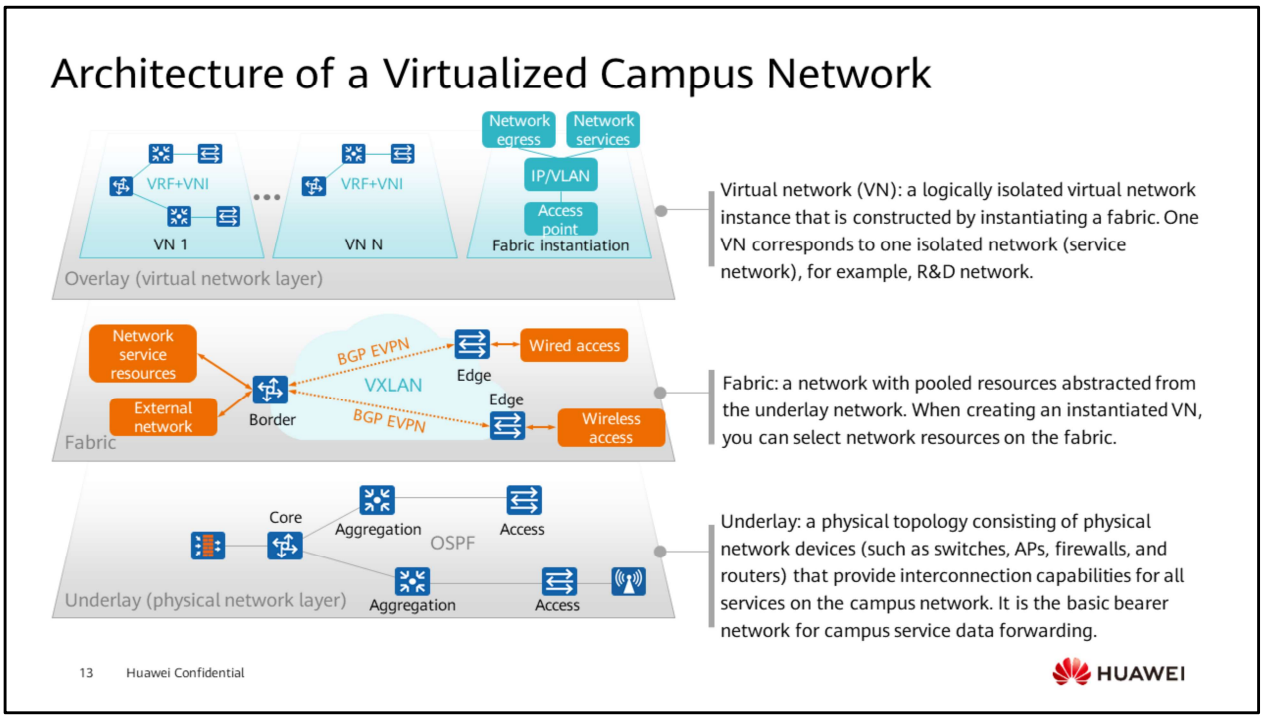
- The following figure shows the typical network architecture of the CloudCampus Solution, which consists of the network layer, management layer, and application layer.



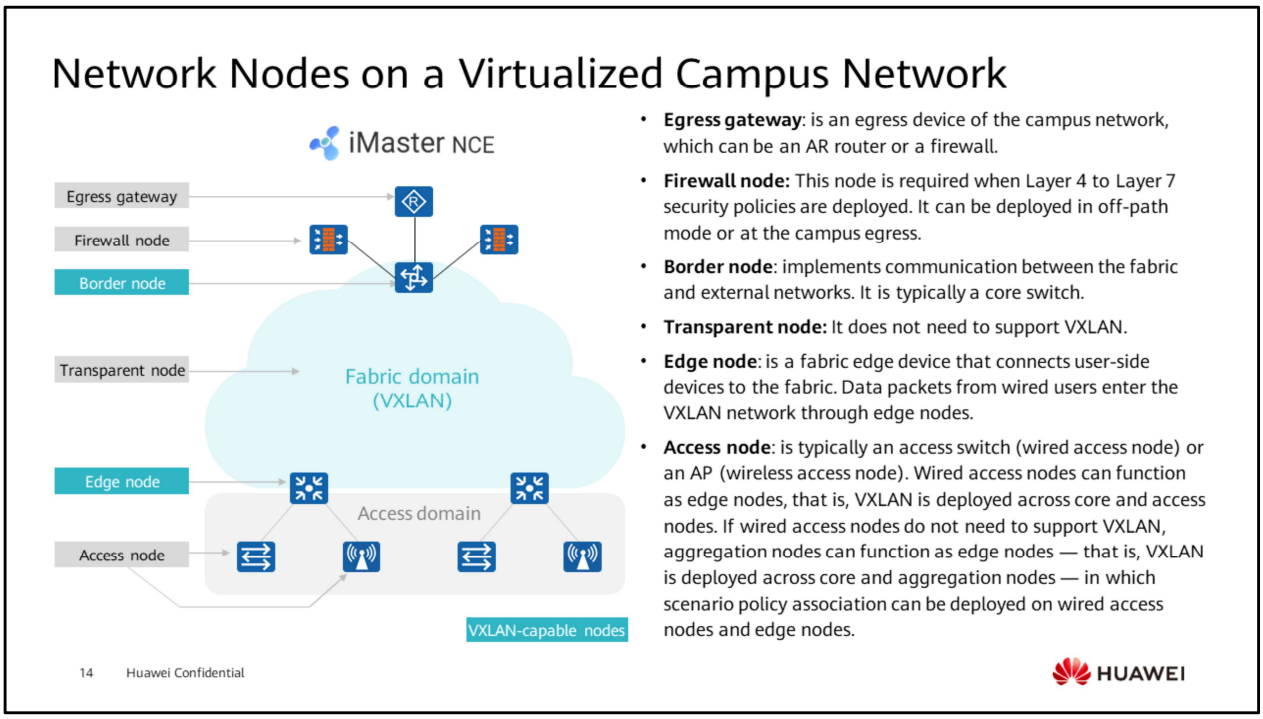
HUAWEI

- Network layer
  - Virtualization technologies are introduced to divide the network layer into a physical network and a virtual network.
    - Physical network: is also called the underlay network and provides basic connection services for a campus network. To meet access requirements of multiple types of terminals, the physical network provides converged access for three networks, allowing simultaneous access of wired, wireless, and IoT terminals.
    - Virtual network: is also called the overlay network. Virtualization technology is used to construct one or more overlay networks on top of the underlay network. Service policies are deployed on the overlay networks and are separated from the underlay network, decoupling services from networks. Multiple overlay networks can serve different services or customer segments.
- Management layer
  - The management layer provides management capabilities, such as configuration management, service management, maintenance and fault detection, as well as security threat analysis. Traditional campus networks use a network management system (NMS) for network management. Although the NMS can display the network status, it lacks flexibility and automatic management capability. Once service requirements change, the network administrator needs to re-plan services and manually modify configurations on network devices (including routers, switches, and firewalls), which is inefficient and error-prone. Therefore, maintaining network flexibility is critical in rapidly changing service environments, which requires the use of automation tools to assist in network and service management. Huawei's CloudCampus Solution uses iMaster NCE-Campus to implement automatic network and service provisioning.

    - iMaster NCE-Campus abstracts network devices and applications, and rapidly develops and automatically deploys applications through orchestration and by invoking abstract models. iMaster NCE-Campus illustrates the entire network but not independent devices (such as switches, routers, and APs) or discrete configurations (such as access control, QoS, and routing policies) on devices.
- Application layer
    - Based on iMaster NCE-Campus, Huawei CloudCampus Solution provides open standards-compliant APIs, through which various information including user identities, network resources, service quality, location information, and network topology, is opened up to upper-layer services. Third parties can use these APIs to customize innovative service applications based on service demands, meeting service requirements in multiple fields such as education, commerce, enterprise, and government.



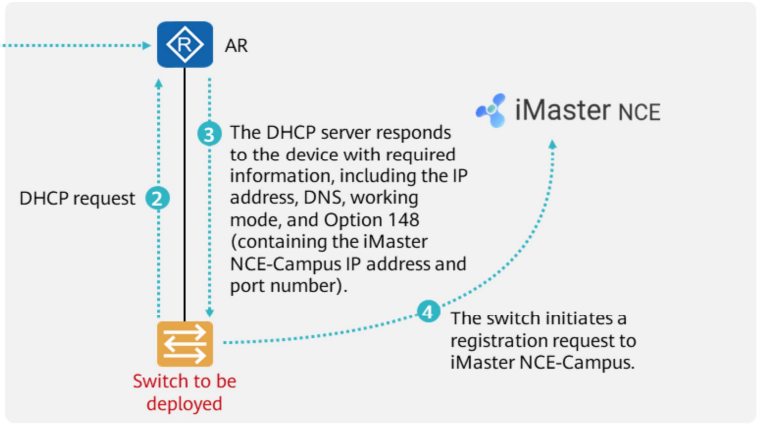
- On a large- or medium-sized campus network, the virtualization solution may need to be used to decouple services from the network, so as to build a multi-purpose network and achieve flexible, fast service deployment without changing the basic network infrastructure. Using such a solution means that the virtualized campus network architecture must be different from the traditional network architecture.
- This slide presents the virtualized campus network architecture. The underlay is the physical network layer, and the overlay is the virtual network layer constructed on top of the underlay using the Virtual Extensible LAN (VXLAN) technology.



- On a fabric network, a VXLAN tunnel endpoint (VTEP) can function as either a border or edge node:
  - Border node: It corresponds to a physical network device and forwards data between the fabric and external networks. In most cases, border nodes are VXLAN-capable core switches.
  - Edge node: It corresponds to a physical network device. User traffic enters the fabric from the edge node. Typically, edge nodes are VXLAN-capable access or aggregation switches.
- Policy association:
  - Policy association provides a solution to contradiction between policy strengths and complexity on large campus networks. In the solution, user access policies are centrally managed on the gateway devices and enforced by gateway and authentication access devices.
  - After policy association is configured, authentication access devices can transparently transmit BPDUs and report user logout and user access positions in real time. In addition, the authentication control device requests authentication access devices to enforce user access policies, thus controlling user access to the network.

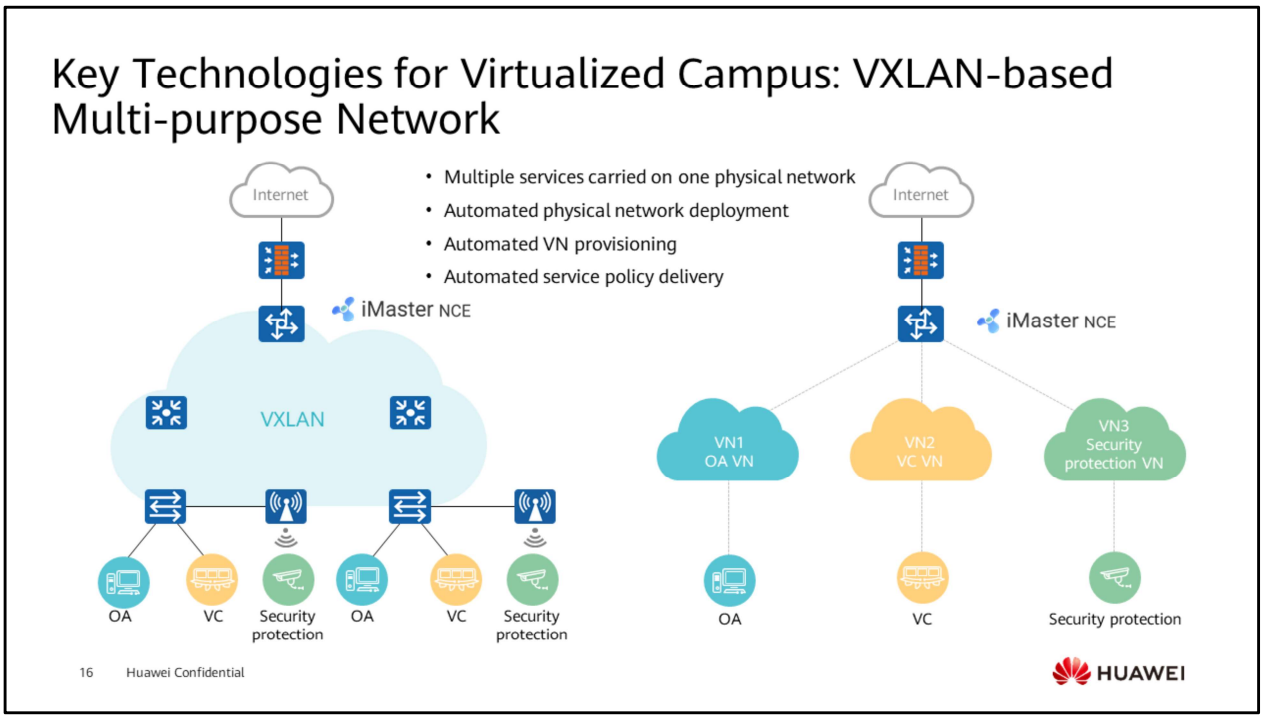
Key Technologies for Virtualized Campus: DHCP-based Plug-and-Play of Network Devices

- The network administrator has deployed the DHCP service on the network. (It is recommended that the DHCP service be deployed on the AR router.)
- In addition to delivering an IP address to the device to be deployed, the DHCP server uses DHCP Option 148 to notify the device of the iMaster NCE-Campus IP address and port number.

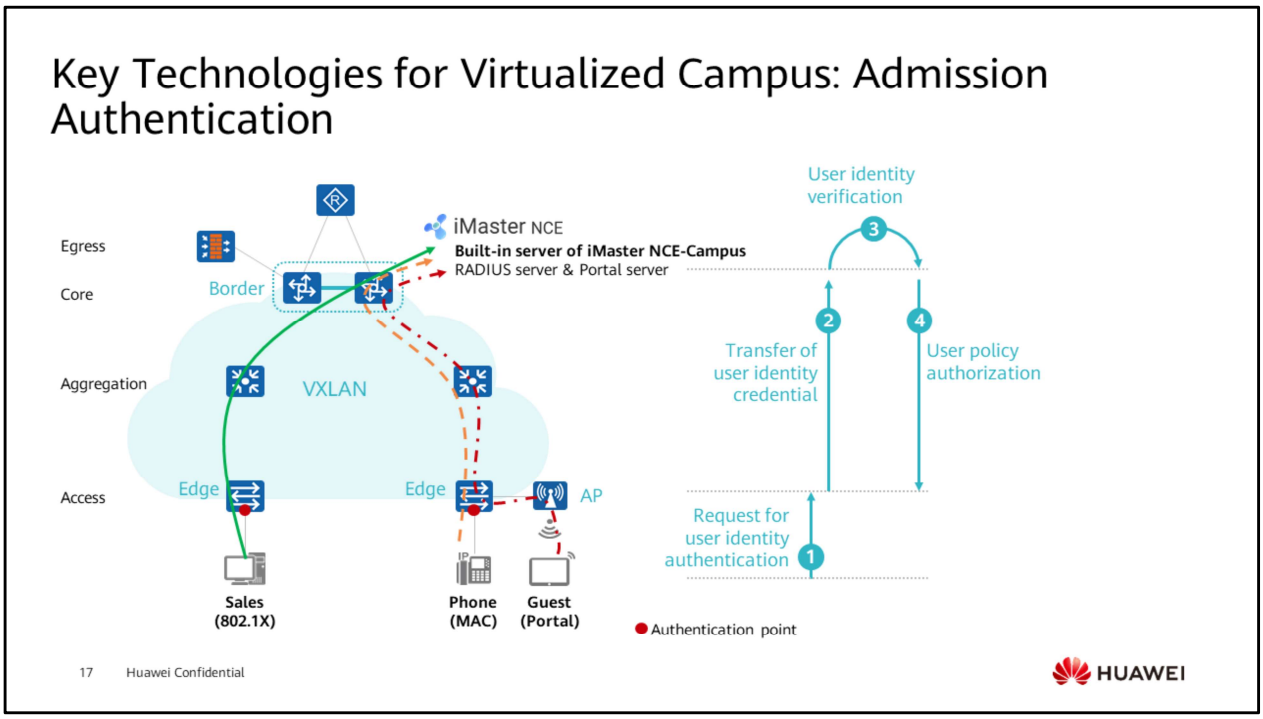


Devices supported: AR routers, switches, APs

HUAWEI

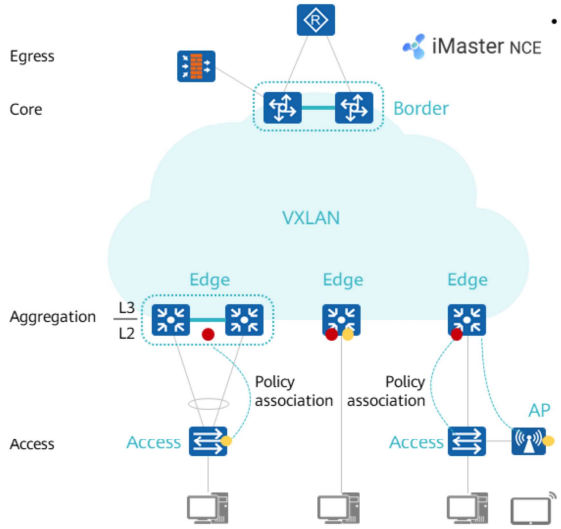


- Using virtualization technologies, multiple virtual networks (VNs) can be created on top of a physical network. Different VNs are used for different services, such as OA, videoconferencing, and security protection.
- VC: Video Conference



1. Request for user identity authentication: The terminal sends its identity credential to the admission device.
2. User identity authentication: The admission device sends the identity credential to the admission server for identity authentication.
3. User identity verification: The admission server stores user identity information and manages users. After receiving the identity credential of the terminal, the admission server verifies the identity of the terminal, determines whether the terminal identity is valid, and delivers the verification result and policy to the admission device.
4. User policy authorization: As a policy enforcement device, the admission device implements policy control over the terminal based on the authorization result provided by the admission server, for example, permitting or denying network access, or performing more complex policy control on the terminal. Complex policy control can be increasing or decreasing the forwarding priority of the terminal, or restricting the network access rate of the terminal.

Key Technologies for Virtualized Campus: Policy Association



- If access switches do not support VXLAN, policy association can be deployed between access and aggregation switches (gateways). The gateway manages user access policies in a unified manner, and the access switch enforces user access policies.
  - Authentication control points and enforcement points are connected through CAPWAP tunnels (which are the management tunnels used by policy association).
  - CAPWAP tunnels implement user association, message transmission, user authorization policy delivery, user information synchronization, and other functions.
  - After policy association is configured, authentication enforcement points can transparently transmit BPDUs, report user logout and access locations in real time, and enforce user access policies, thereby controlling user access to the network.

Authentication control point Authentication enforcement point

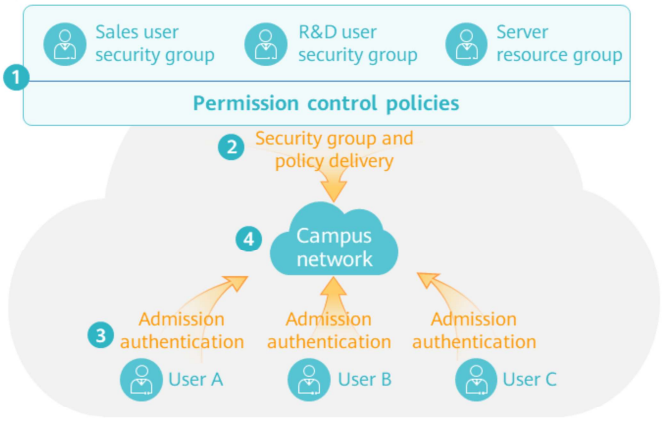
CAPWAP tunnel Link aggregation CSS/iStack

HUAWEI

- CAPWAP: Control and Provisioning of Wireless Access Points

Key Technologies for Virtualized Campus: Free Mobility

- Free mobility manages policies and controls permissions based on security groups. Regardless of user locations and IP addresses in use, this technology ensures that users can obtain consistent network permissions and corresponding policies can be enforced on the users.

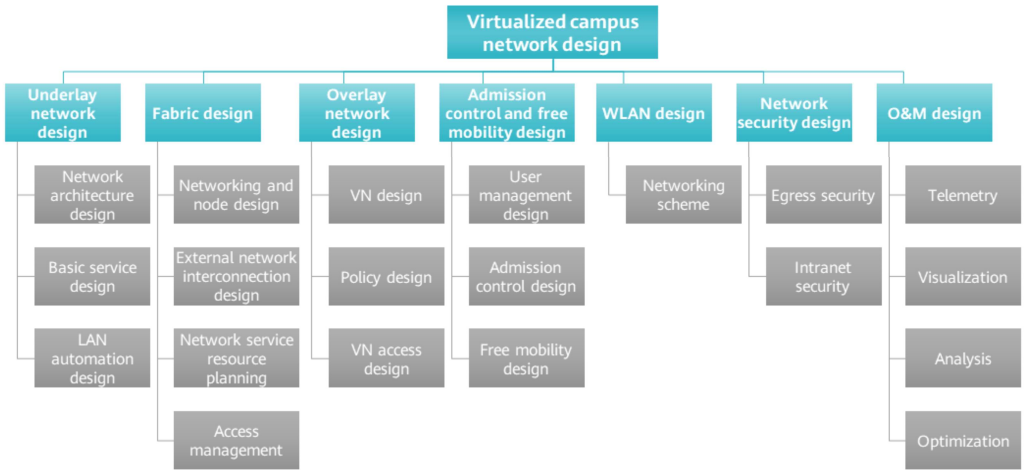


1. Create security groups. A security group is a group of users or network resources that have the same network access policies.
2. Define permission control policies based on security groups and deliver the policies to network devices.
3. Authorized security groups are assigned to the users who pass admission authentication.
4. After user traffic enters the network, network devices enforce policies based on the corresponding source and destination security groups of the traffic.

HUAWEI

- Free mobility introduces the concept of security group. Security groups are related only to user identities and are completely decoupled from network information such as user VLANs and IP addresses.
- User policy management and permission control are performed based on security groups.

Overview of Large- and Medium-Sized Virtualized Campus Network Design



HUAWEI

Contents

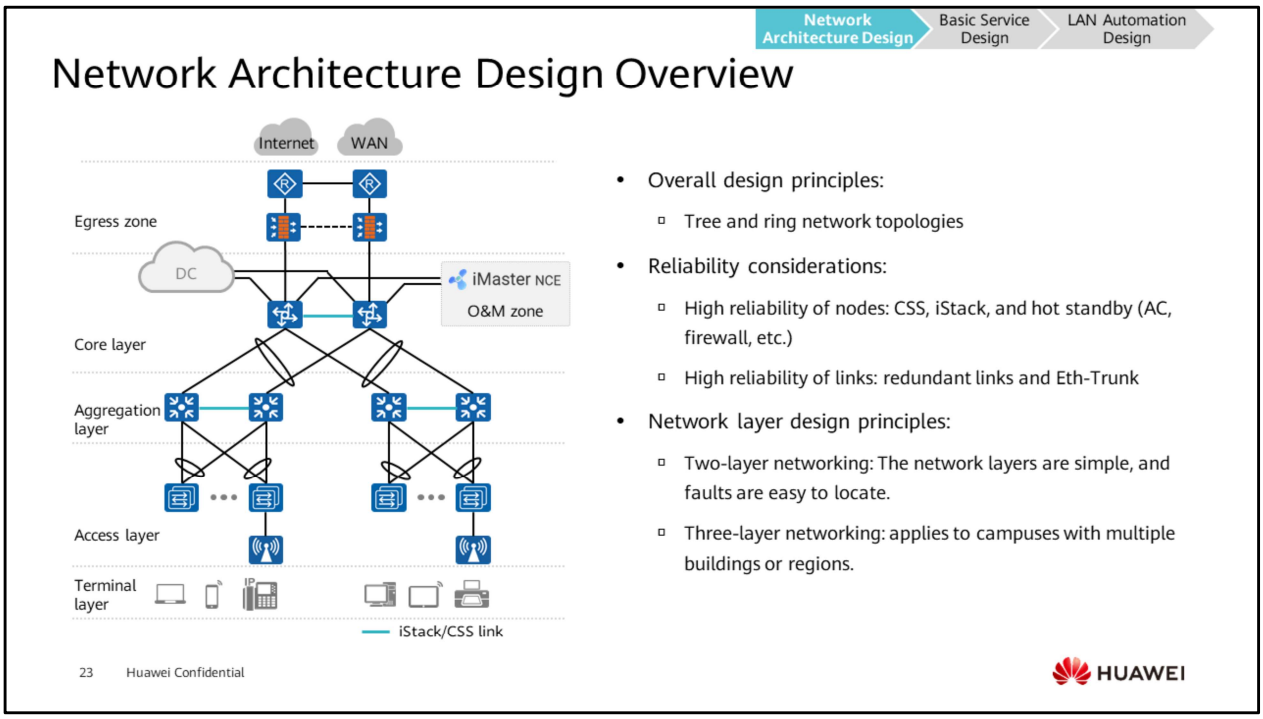
1. CloudCampus Solution and Virtualized Campus Network Overview
2. **Underlay Network Design**
3. Fabric Design
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. WLAN Design
7. Network Security Design
8. O&M Design

HUAWEI

Underlay Design Outline

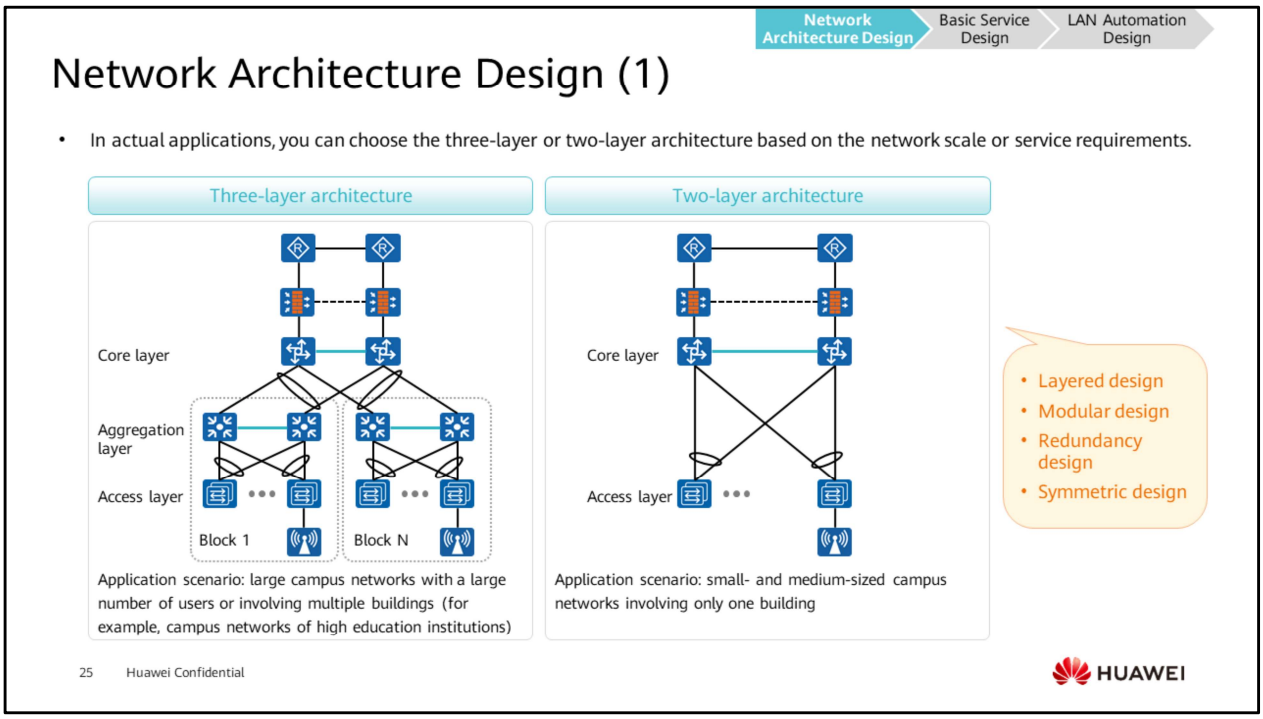


HUAWEI

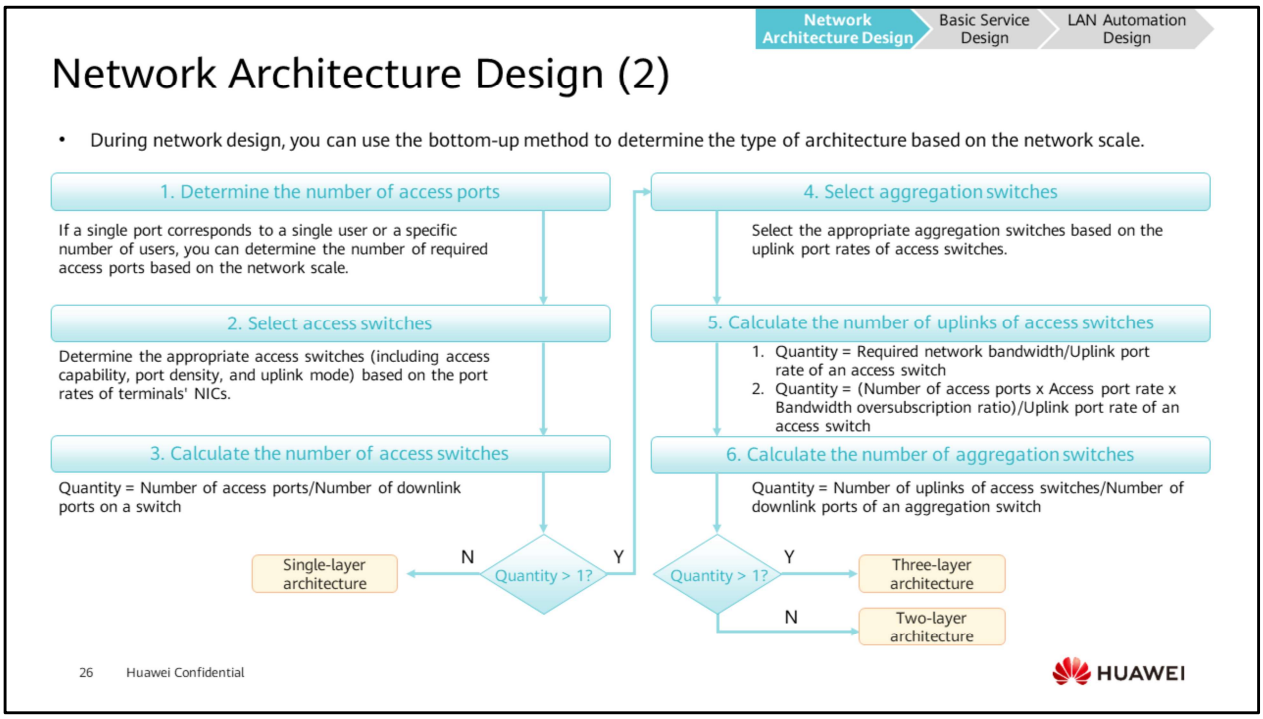


- Large- and medium-sized campus networks often use the tree topology with the core layer as the root, as shown in the figure. This topology is stable and easy to expand and maintain. A campus network can be divided into the following layers: access layer, aggregation layer, core layer, and multiple zones including the egress zone, DC zone, and O&M zone. Internal changes within a module have limited impact on other modules, facilitating fault location.
- Terminal layer
  - The terminal layer involves various types of terminals that access the campus network, such as PCs, printers, IP phones, mobile phones, and cameras.
- Access layer
  - The access layer provides various access modes for users and is the first network layer for terminals to access a campus network. The access layer is usually composed of access switches. There are a large number of access switches that are sparsely distributed in different places on the network. In most cases, an access switch is a simple Layer 2 switch. If the terminal layer has wireless terminals, the access layer provides APs that access the network through access switches.
- Aggregation layer
  - The aggregation layer connects the access layer to the core layer. The aggregation layer forwards horizontal traffic between users and forwards vertical traffic to the core layer. It can also function as the switching core for a department or zone and connect the department or zone to the exclusive server zone. In addition, the aggregation layer can further extend the quantity of access terminals.

- Core layer
    - The core layer is the core for data exchange on a campus network. It connects various components of a campus network, such as the DC zone, aggregation layer, and egress zone. The core layer is responsible for high-speed interconnection of the entire campus network. High-performance core switches need to be deployed to implement high bandwidth utilization and fast convergence upon network faults. It is recommended that the core layer be deployed if a campus has more than three departments. For a wireless network, the core layer includes ACs. After a wireless terminal accesses the network through an AP, the AP communicates with an AC using a CAPWAP tunnel.
- Egress zone
    - The campus egress is the boundary between a campus network and an external network. Internal users of the campus network can access the external network through the campus egress zone, and external users can access the internal network through the campus egress zone. The campus egress zone typically has routers and firewalls deployed. The routers enable interconnection between internal and external networks, and the firewalls provide border security protection.
- DC zone
    - In the DC zone, service servers such as the file server and email server are managed, and services are provided for internal and external users.
- O&M zone
    - In the O&M zone, network servers such as the network management system and authentication server are managed. The standard NMS interacts with network devices through the Simple Network Management Protocol (SNMP) and provides configuration, management, and maintenance functions. For example, it provides network topology and port display management, network device configuration management, network fault diagnosis and alarm, as well as network performance and status analysis. Huawei CloudCampus Solution provides intelligent O&M based on NETCONF and telemetry.
- In addition to these, a campus may have a demilitarized zone (DMZ). The DMZ contains public servers that can be accessed by guests (non-employees), therefore the access permission to this zone is strictly controlled.



- Layered design:
  - Each layer can be considered a well-structured module with specific roles and functions. This layered structure is easy to expand and maintain, reducing the design complexity and difficulty.
- Modular design:
  - Each module corresponds to a department, function, or service area. Modules can be expanded flexibly based on the network scale, and adjustment in a department or area does not affect other departments or areas, which facilitates fault locating.
- Redundancy design:
  - Dual-node redundancy design can ensure device-level reliability. Appropriate redundancy improves reliability, but excessive redundancy makes O&M difficult. If dual-node redundancy cannot be implemented, you may consider card-level redundancy, such as dual main control boards or switch fabric units (SFUs), for modular core switches or egress routers. In addition, Eth-Trunk can be deployed for important links to ensure link-level reliability.
- Symmetric design:
  - A symmetric network structure makes the topology clearer and facilitates service deployment, protocol design and analysis.

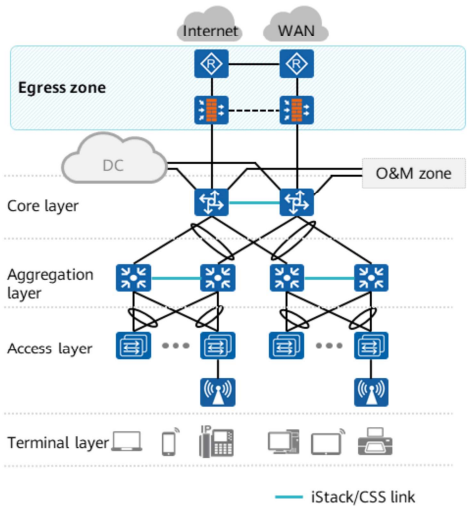


- Design method of the hierarchical model for common network architectures:
    1. Determine the number of access switch ports based on the network scale. Typically, one port corresponds to one terminal or one network access point (for example, AP).
    2. Select switches based on the port rates of terminals' NICs.
    3. Calculate the number of access switches.
        - Number of access switches required = Number of access ports/Number of downlink ports on a switch
        - If the calculation result is greater than 1, aggregation switches need to be deployed. Otherwise, use the single-layer architecture.
    4. Select aggregation switches based on the uplink port rates of access switches.
    5. Calculate the number of uplinks of access switches using either of the following methods:
        - Based on the network bandwidth: Number of uplinks = Network bandwidth/Uplink port rate of an access switch
        - Based on the network scale: Number of uplinks = (Number of access ports x Access port rate x Bandwidth oversubscription ratio)/Uplink port rate of an access switch

6. Calculate the number of aggregation switches.
    - Number of aggregation switches required = Number of uplinks of access switches/Number of downlink ports of an aggregation switch
    - If the result is greater than 1, select the three-layer architecture. Otherwise, use the two-layer architecture.

- Notes:
    - In the preceding calculations, the calculation results need to be rounded up.
    - Bandwidth oversubscription ratio = Actual downlink bandwidth of a service/Maximum physical bandwidth provided by a single switch
        - Traffic of the service may exhaust the maximum physical bandwidth. However, this probability is low in most cases. Generally, network planning focuses on service experience in common situations. That is, certain bandwidth redundancy is designed when actual possible bandwidth is met.
        - Before designing a network traffic oversubscription ratio, we need to understand the service applications and features to be deployed on the network and determine the network services and traffic models. Additionally, the volume and ratio of traffic must be thoroughly considered before designing an oversubscription ratio and selecting devices.

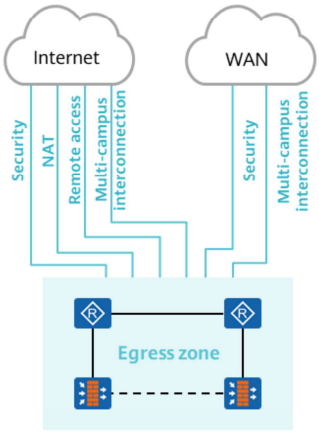
Network Egress Design Overview



- As the boundary between a campus network and external networks (including the Internet and WAN), the egress zone is responsible for communication and security protection between internal and external networks.
- The requirements for the egress zone design are as follows:
  - **Network connectivity:** Internal users can access external networks. If the campus network provides access for external users, external users should be able to access the internal network.
  - **Network security:** To ensure that the campus network is secure and controllable, especially border security, firewalls and the Intrusion Prevention System (IPS) devices need to be configured. Security components are selected based on security requirements and investment scales.
  - **Flexible access modes:** Various access modes are provided, including LAN-side and WAN-side access.
  - **Strong service control capability:** Service deployment and isolation are easy to implement, and various VPN access modes are provided, including IPsec VPN, SSL VPN, and MPLS VPN.

HUAWEI

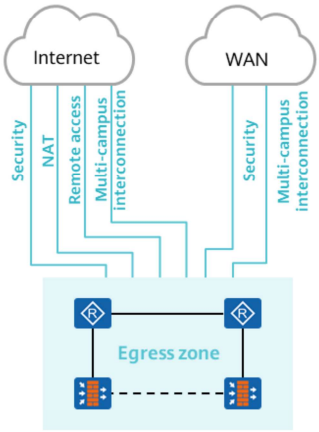
Egress Network Service Scenarios (1)



Service Scenario	Sub-scenario	Description	Applicable Network
Security	Firewall	A firewall uses security zones to divide networks and identify packet forwarding paths. When a packet is transmitted between security zones, the firewall can perform a security check and enforce corresponding security policies.	Internet and leased line
	URL filtering	URL filtering controls internal users' access to external URLs. It allows or forbids users to access certain web page resources, thereby regulating online behaviors.	Internet
NAT	Source NAT	Intranet users usually use private IP addresses to access the Internet. To ensure successful access to Internet services, private IP addresses must be translated into public IP addresses.	Internet
	NAT server	Servers deployed on the campus network, such as the file server, can be mapped to the Internet using public IP addresses or public IP addresses and port numbers through the NAT server, so that they can be accessed from the Internet.	Internet

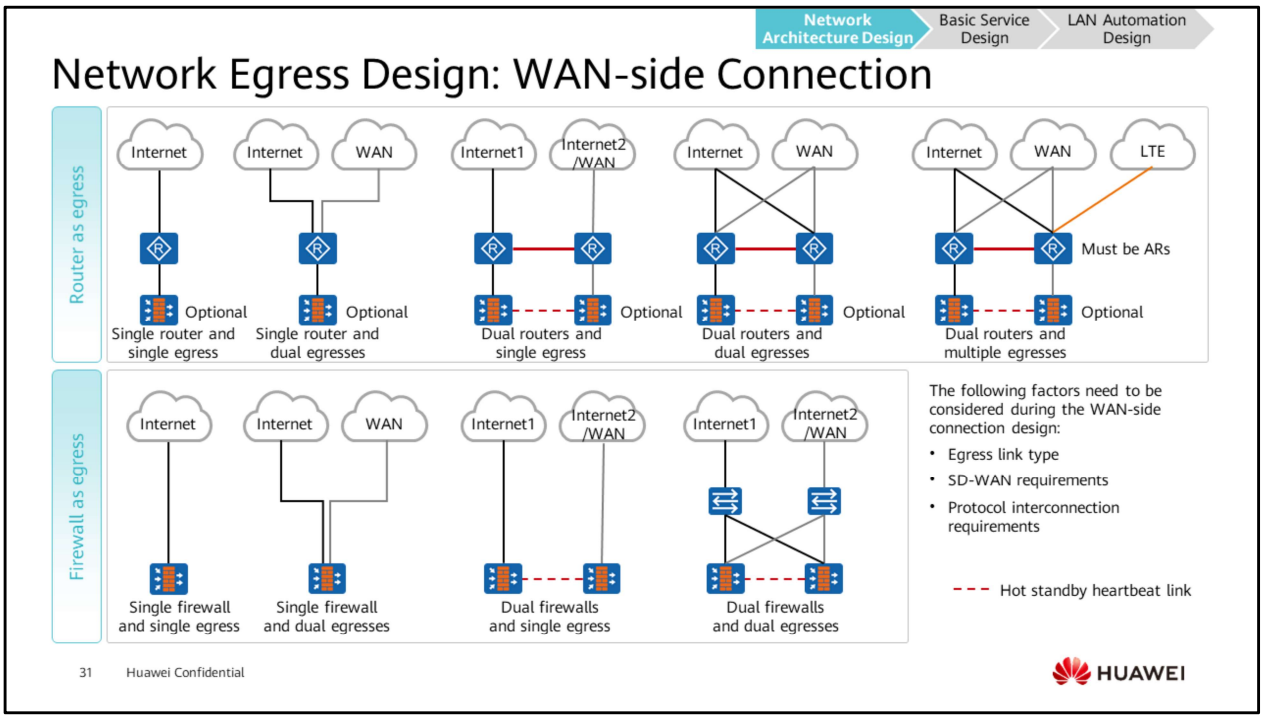
HUAWEI

Egress Network Service Scenarios (2)



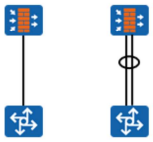
Service Scenario	Sub-scenario	Description	Applicable Network
Remote access	SSL VPN access	The SSL VPN gateway connects to the Internet and provides SSL VPN services for mobile office users (employees on business trips). After a mobile user uses a terminal (such as a laptop or smart phone) to establish an SSL VPN tunnel with the gateway, the user can remotely access intranet resources such as the web server, file server, and mail server through the SSL VPN tunnel.	Internet
Multi-campus interconnection	Traditional MPLS VPN interconnection	CE devices can be connected to the MPLS VPN network.	Leased line
	IPsec VPN interconnection	IPsec VPN tunnels are established between multiple campuses to ensure secure data transmission on a low-security bearer network.	Internet and leased line

HUAWEI

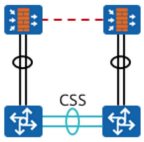


- To ensure reliability, routers and firewalls are usually deployed in redundancy mode. It is recommended that devices be deployed in redundancy mode at the egress of a large- or medium-sized campus network.
- WAN-side connection design:
  - Egress link type
    - Ethernet: Select routers (NE or AR series) or firewalls as egress devices.
    - For non-Ethernet links, such as EI, CE1, and CPOS links, select routers as egress devices.
  - SD-WAN requirements:
    - If there are SD-WAN requirements, select AR routers as egress devices.
    - If there is no SD-WAN requirement, select firewalls or NE routers as egress devices.
  - Protocol interconnection requirements:
    - If BGP runs between egress devices and external networks, it is recommended that routers be used as the egress devices. This is because routers have a larger number of routing table entries and higher performance and many routing policies need to be deployed on egress devices.

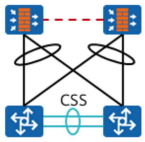
Network Egress Design: LAN-side Connection



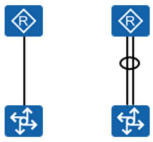
A single firewall is connected to a single core switch



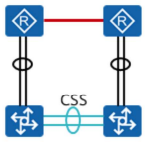
Two firewalls are connected to a CSS of core switches in square-looped mode



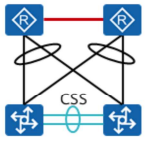
Two firewalls are connected to a CSS of core switches in dual-homed mode



A single router is connected to a single core switch



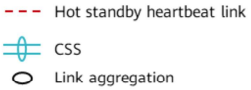
Two routers are connected to a CSS of core switches in square-looped mode



Two routers are connected to a CSS of core switches in dual-homed mode

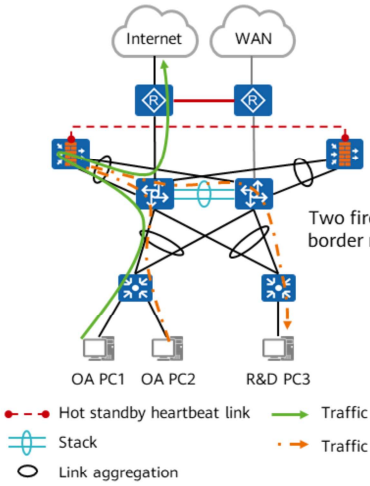
LAN-side connection design

- It is recommended that core switches be deployed in cluster or stack mode to form a logical switch, thereby ensuring reliability of core switches.
- Egress gateways are usually deployed in a two-node cluster to meet reliability requirements. Different types of gateways differ slightly in their two-node cluster deployment capabilities. For example, the link between the active and standby firewalls in a two-node cluster is used only for heartbeat and entry synchronization, and not for traffic forwarding. The link between routers functioning as gateways can be used for service interconnection and for synchronizing and backing up information between them.
- It is recommended that egress gateways and core switches be connected in dual-homed mode and link aggregation be used to enhance the link-level reliability and forwarding bandwidth between devices.



HUAWEI

Network Egress Design: Firewall Deployment in Off-Path Mode

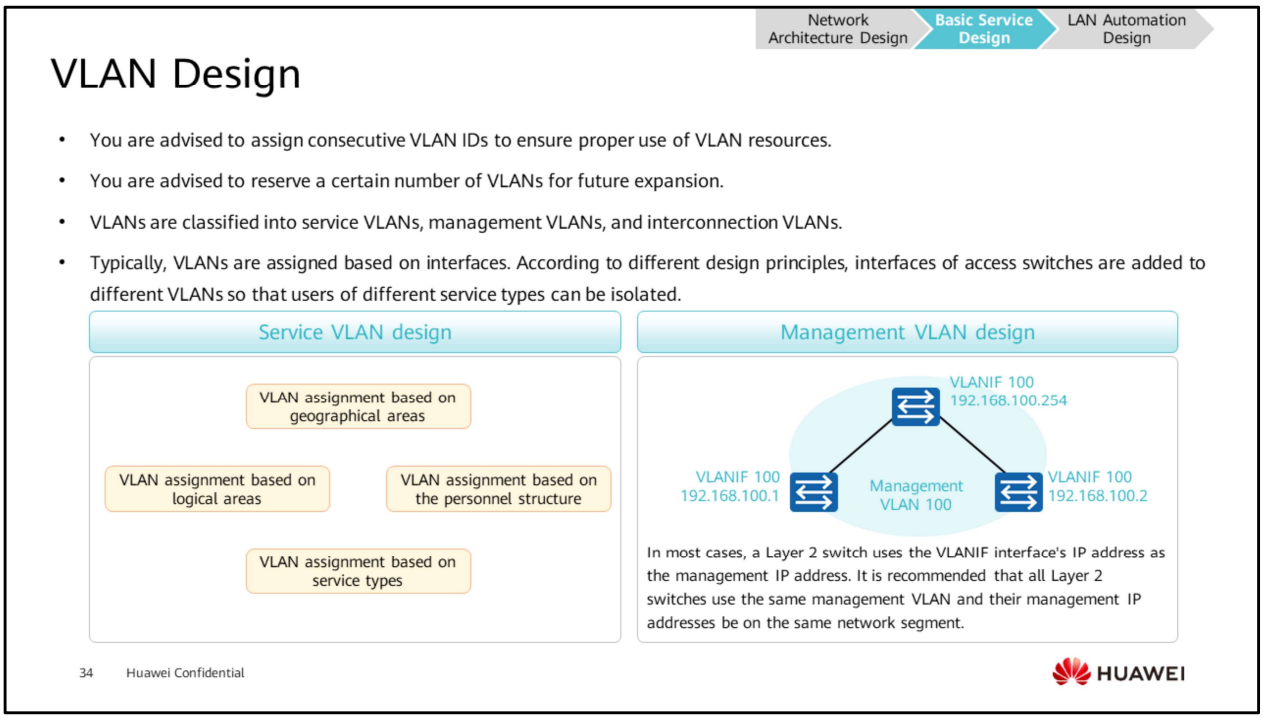


Two firewalls are connected to border nodes in off-path mode

Deploying firewalls in off-path mode

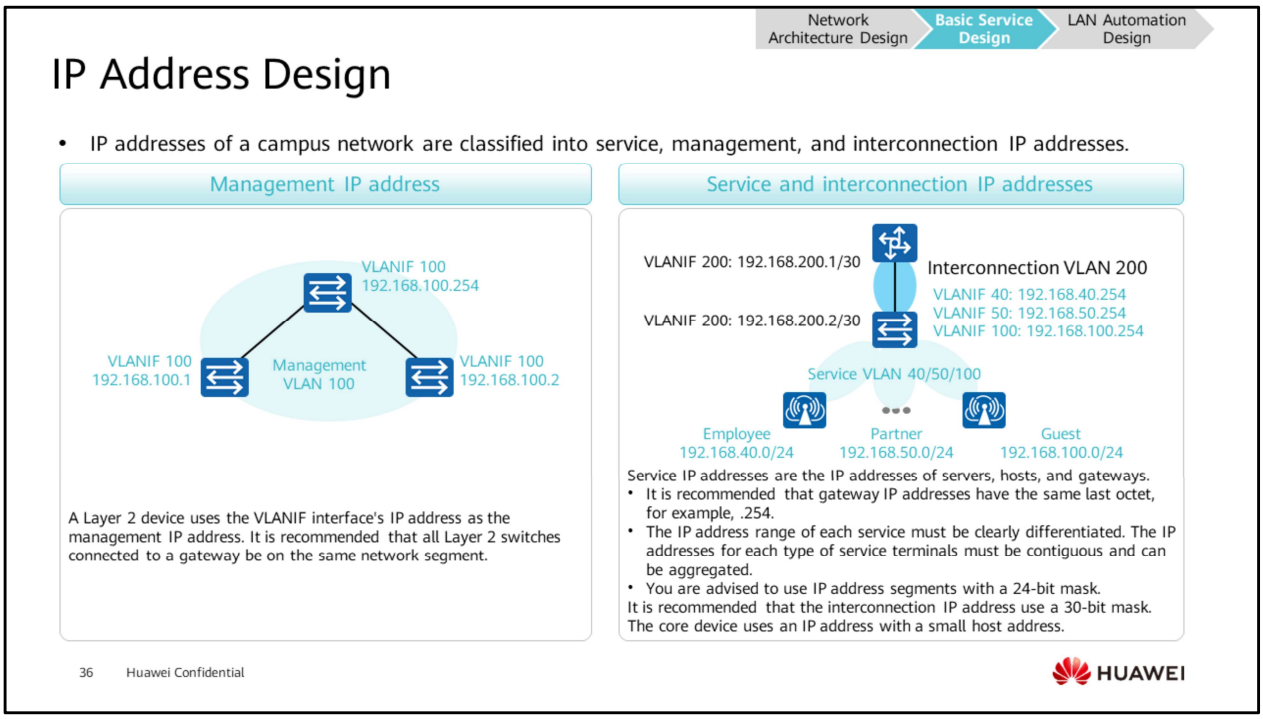
- In addition to in-path deployment, firewalls can be deployed in off-path mode. Generally, off-path deployment of firewalls is selected based on the following principles:
  - Whether traffic passes through a firewall can be flexibly defined.
  - Expansion is easy, for example, in the network reconstruction scenario.
  - A firewall fault has little impact on services.
- On a campus network:
  - If a firewall is used as the egress gateway, the firewall usually provides functions such as security policies, NAT, and VPN gateway. If the north-south traffic must pass through the firewall and if scalability is not considered, in-path deployment is recommended for the firewall.
  - If a router is used as the egress gateway and the firewall only provides high-level security functions, such as intrusion prevention, antivirus, and URL filtering, off-path deployment can be used for the firewall.

HUAWEI



- Service VLAN:
  - Assign VLANs by logical area, geographical area, personnel structure, or service type.
    - Assign VLANs by logical area. For example, VLANs 100 to 199 are used in the core network zone, VLANs 200 to 999 are used in the server zone, and VLANs 2000 to 3499 are used on the access network.
    - Assign VLANs by geographical area. For example, VLANs 2000 to 2199 are used in area A, and VLANs 2200 to 2399 are used in area B.
    - Assign VLANs by personnel structure. For example, department A uses VLANs 2000 to 2009, and department B uses VLANs 2010 to 2019.
    - Assign VLANs by service type. For example, VLANs 200 to 299 are used in the web server zone, VLANs 300 to 399 are used in the app server zone, and VLANs 400 to 499 are used in the database server zone.
  - If users are sensitive to the voice latency, the voice service must be preferentially guaranteed. It is recommended that the voice VLAN be planned for the voice service. Huawei switches can automatically identify voice data, transmit voice data in the voice VLAN, and perform QoS guarantee. When network congestion occurs, voice data can be preferentially transmitted.

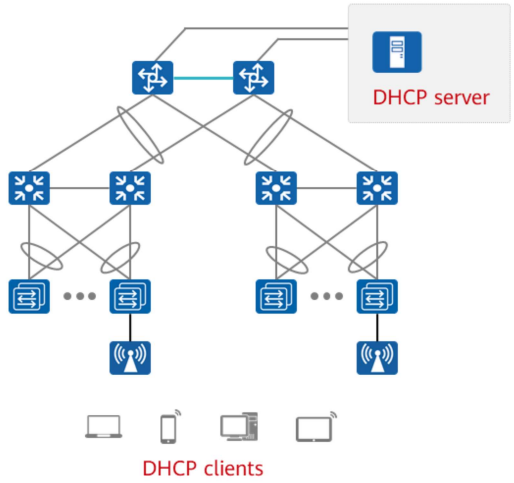
    - If different users have the same multicast data service, you are advised to plan a multicast VLAN and bind the user VLANs to the multicast VLAN. This prevents the upstream gateway from copying multicast data in multiple user VLANs.
    - Do not use VLAN 1 as a service VLAN.
- Management VLAN:
    - It is recommended that the management VLAN be planned for a Layer 2 switch and the VLANIF interface of the management VLAN be used as the management interface, through which the NMS manages the switch. It is recommended that all Layer 2 switches use the same management VLAN.
    - It is recommended that service interfaces be used as management interfaces of Layer 3 devices (gateways and their upstream devices), and no management VLAN needs to be planned for these devices.
- Interconnection VLAN:
    - An interconnection VLAN is usually configured between two Layer 3 switches or between a Layer 3 switch and a router. Create VLANIF interfaces for Layer 3 interconnection.



- IP address planning complies with the following principles:
  - Unique: Each host on an IP network must have a unique IP address.
  - Contiguous: Node addresses of the same service must be contiguous to facilitate route planning and summarization. Contiguous addresses facilitate route summarization, thereby reducing the size of the routing table and speeding up route calculation and convergence.
  - Scalable: Some IP addresses need to be reserved at each layer, so that no address segments or routing entries need to be added when the network is expanded.
  - Easy to maintain: Device and service address segments need to be clearly distinguished from each other, facilitating subsequent statistics monitoring and security protection based on address segments. IP addresses can be planned based on VLANs.

- Pay attention to the following points when designing the three types of IP addresses:
  - Service IP address:
    - Considering the scope of a broadcast domain and easy planning, it is recommended that an IP address segment with a 24-bit mask be reserved for each service. If the number of service terminals exceeds 200, another IP address segment with a 24-bit mask is reserved.
  - Management IP address:
    - It is recommended that a Layer 3 device use a Layer 3 interface for management and deployment. The interface address is used as the management IP address for local login and communication with iMaster NCE-Campus.
  - Interconnection IP address:
    - Interconnection addresses are usually aggregated before being advertised. Therefore, you need to use contiguous and aggregatable addresses during planning.
- In addition, if devices at the aggregation and access layers are interconnected, you need to plan IP addresses for policy association in some cases. The IP addresses must be planned based on the entire CloudCampus network and cannot conflict with other service IP addresses.

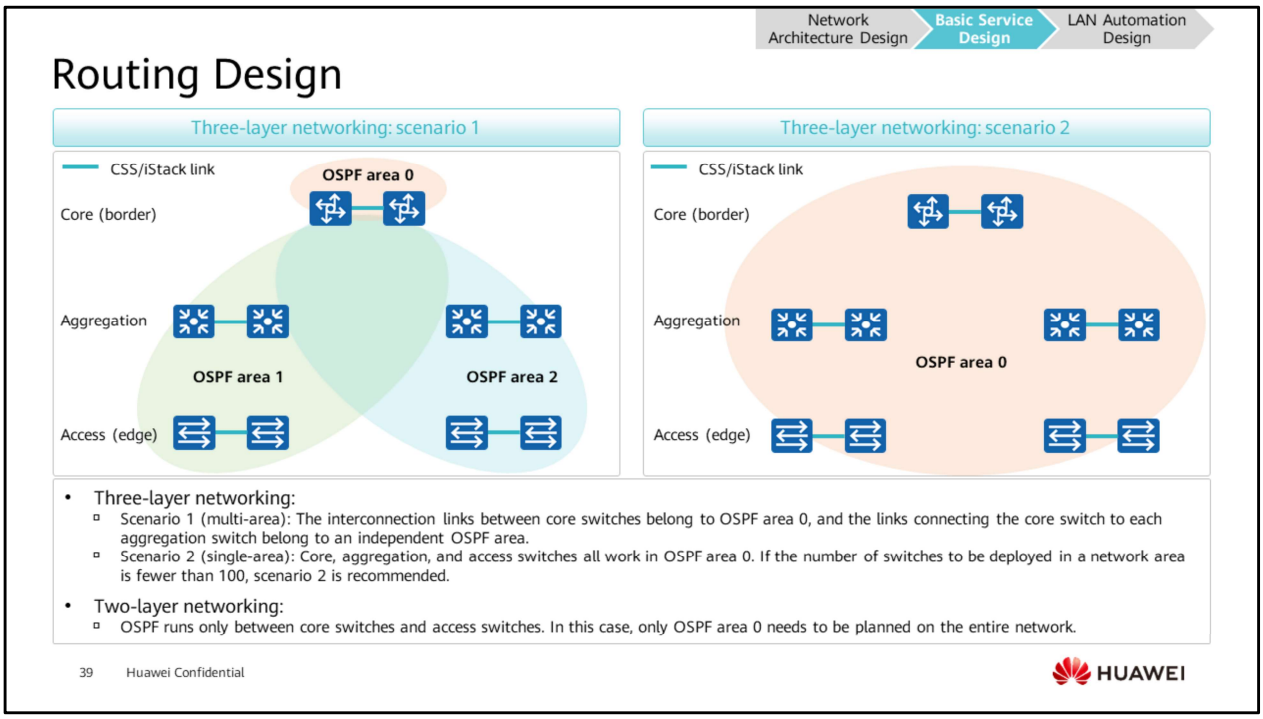
Terminal-Oriented DHCP Service Design



- Plan an independent DHCP server to allocate IP addresses to terminals.
- You are advised to configure DHCP snooping on access devices to defend against attacks.
- Network administrators can use either of the two mechanisms to assign IP addresses to hosts based on network requirements:
  - Dynamic IP address assignment: An IP address is assigned to a host with a lease. This mode applies to scenarios where hosts require temporary access or IP addresses are insufficient, for example, portable computers of traveling employees and mobile terminals in cafes.
  - Static IP address assignment: Fixed IP addresses are assigned to specified hosts or servers such as DNS servers.
- When planning an address pool, filter out static IP addresses.
- The lease needs to be planned based on the online duration of a client.
- On a large or midsize campus network, the DHCP server and hosts are usually not on the same network segment. Therefore, you need to enable the DHCP relay function on the gateway.

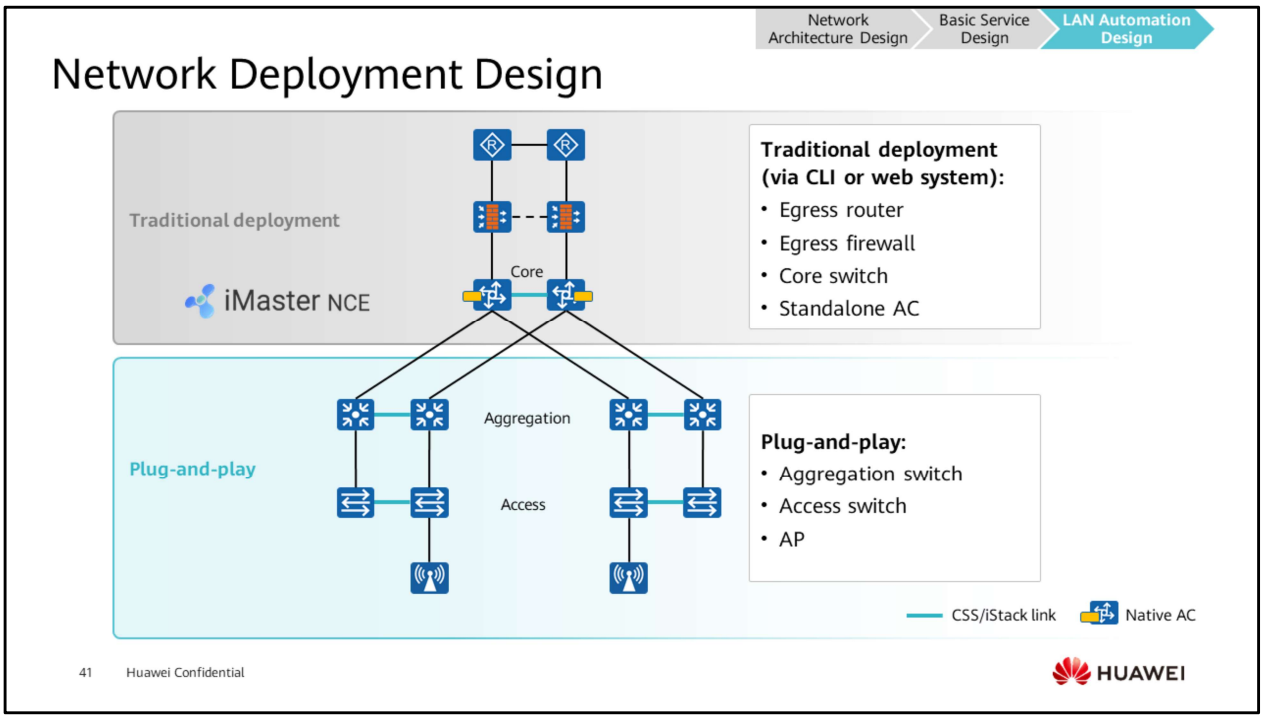
HUAWEI

- Note: It is recommended that egress devices provide the DHCP service for network deployment. For details, see the network deployment section.

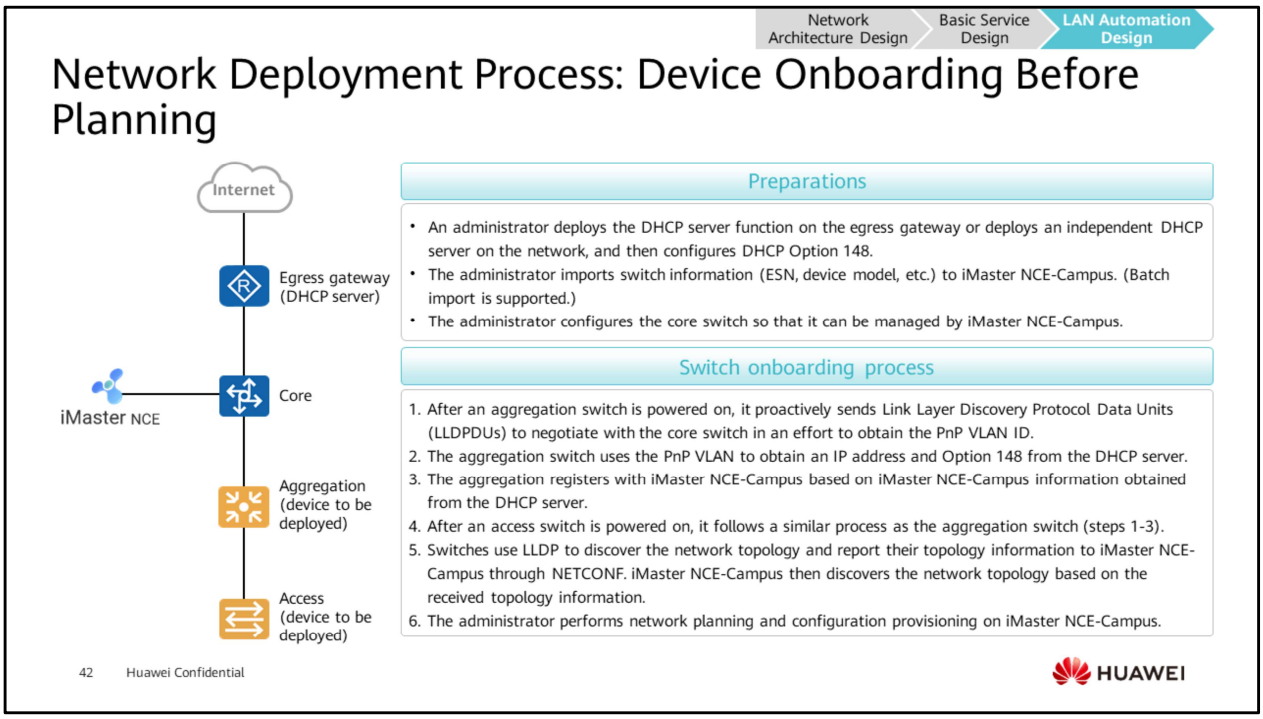


- The routing design includes internal and egress routing design for a campus network.
  - Internal routing design:
    - The routing design must support communication between devices, between terminals, and between devices and terminals on the campus network, as well as communication between these devices/terminals and the external network.
    - It is recommended that you design internal routes based on the gateway location.
      - If gateways are deployed at the aggregation layer, routes need to be deployed at core and aggregation layers. Routing tables can be dynamically updated along with network topology changes, so an Interior Gateway Protocol (IGP), such as Open Shortest Path First (OSPF), is recommended.
      - If gateways are deployed at the core layer, you only need to configure routes at the core layer. It is recommended that static routes be used preferentially.
  - Egress routing design:
    - Egress routes must enable internal terminals to access the Internet and WAN.
    - A large or medium-sized campus network usually has a large number of branches. The egress needs to support multiple links for Internet access and mutual communication between enterprise branches. For this purpose, a large number of routes need to be imported to the campus network. Therefore, you are advised to plan a dynamic routing protocol such as OSPF.

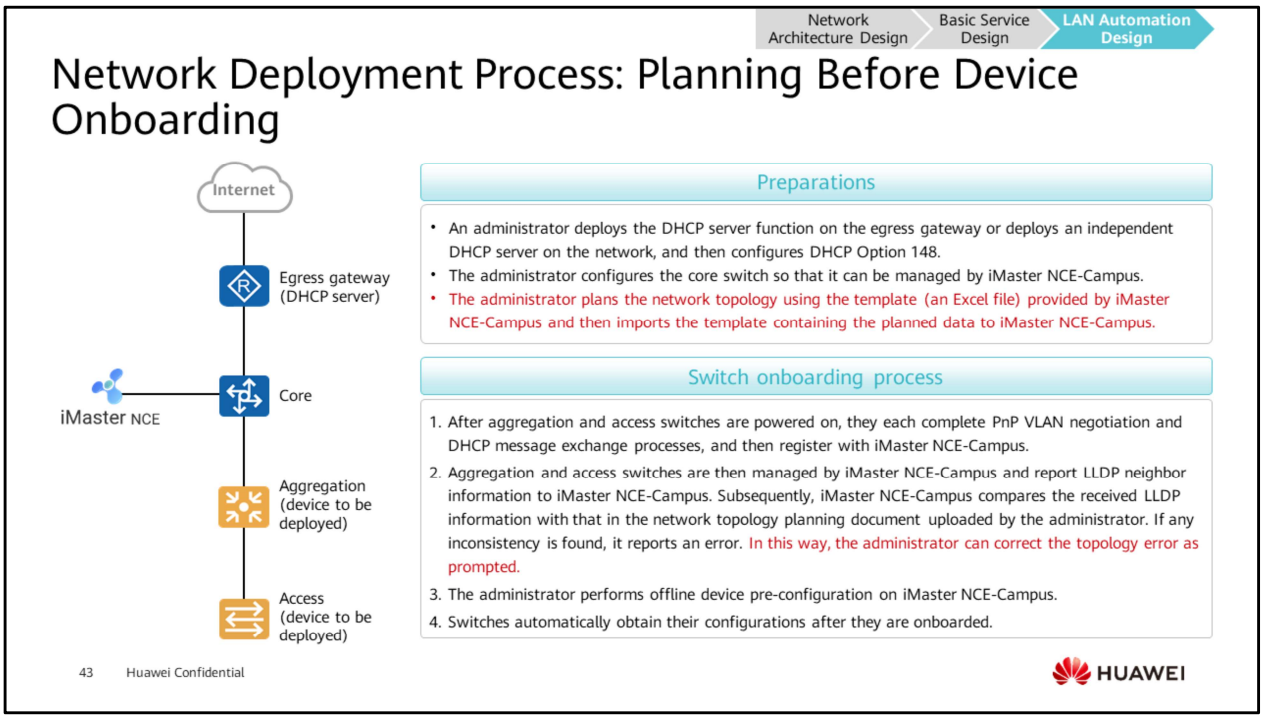
- It is recommended that OSPF be planned for a campus network. The OSPF design precautions are as follows:
  - You are advised to use the IP addresses of loopback interfaces as the router IDs.
  - Areas are divided according to the core, aggregation, and access layers. It is recommended that egress routers and core switches be deployed in the backbone area. The design of non-backbone areas depends on the geographical location and performance of devices.
- Note:
  - This slide describes the routing design in the scenario when VXLAN is deployed across core and access layers.
  - In the scenario where VXLAN is deployed across core and aggregation layers, the routing design (for example, OSPF area division) is the same as when VXLAN is deployed across core and access layers. The only difference is that the aggregation device is the border of the routing domain.



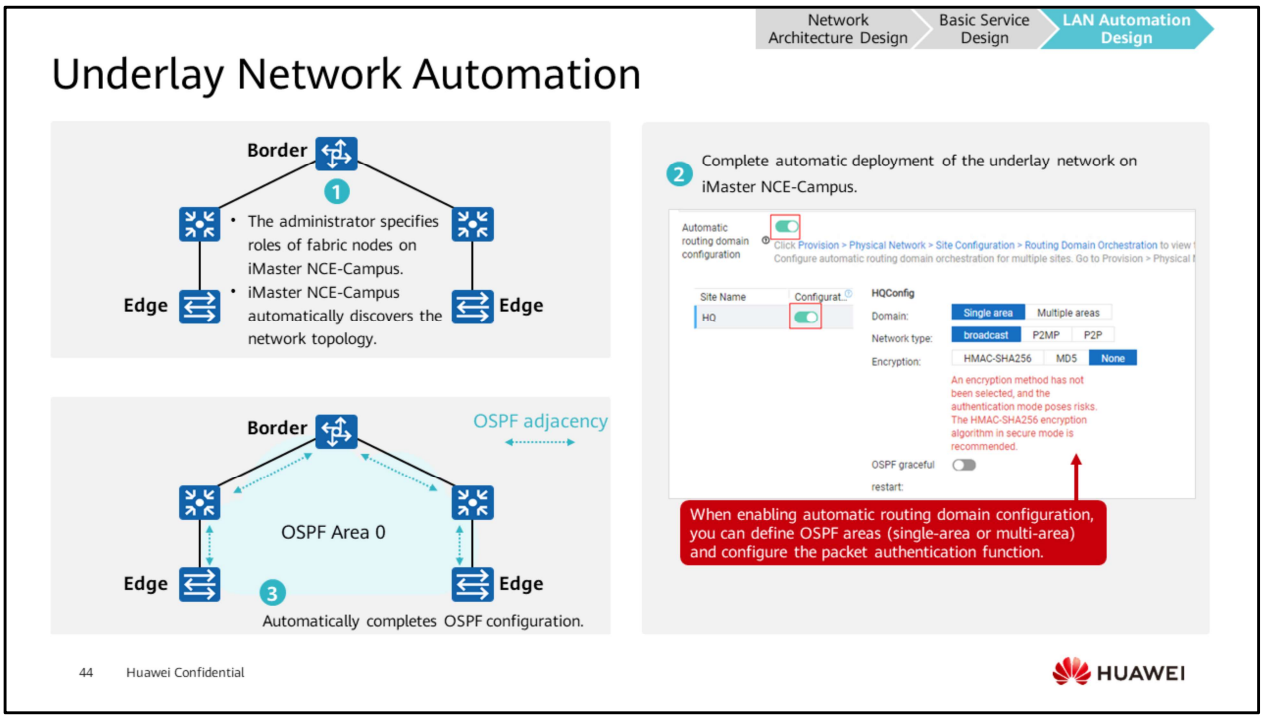
- Typically, egress and core devices on large- and medium-sized campus networks are centrally deployed in a core equipment room. Services transmitted on these devices are complex and their locations on the network are important. In most cases, network engineers need to commission devices onsite during the deployment. Therefore, you are advised to use the web system or CLI to deploy devices at the core layer and upper layers (including core devices, standalone ACs connected in off-path mode, and egress devices).
- A large number of devices (including aggregation devices, access devices, and APs) are deployed at lower layers of the core layer, and service configurations of these devices are similar. Therefore, you are advised to use the DHCP option mode to achieve plug-and-play of devices, thereby simplifying the deployment.
- Note: Core switches obtain basic configurations such as IP addresses using the CLI. Once they establish management channels with iMaster NCE-Campus, iMaster NCE-Campus will automatically deliver services to them.



- In the Huawei's CloudCampus Solution, a DHCP server helps implement plug-and-play (PnP) of switches, eliminating the need to manually enable NETCONF and configure iMaster NCE-Campus address on switches.
    - After switch starts with empty configuration, it sends a Request message containing VLAN 1 (default PnP VLAN ID) to the DHCP server.
    - Alternatively, a user-defined VLAN can be used by the device to initiate a DHCP request. This VLAN is called PnP VLAN.
- In this example, assume that a VLAN different from VLAN 1 is used as the PnP VLAN. To achieve this, you need to configure the PnP VLAN on the core switch or configure the PnP VLAN on iMaster NCE-Campus after the core switch is managed by iMaster NCE-Campus.
- In this deployment mode, devices are deployed and onboarded before the network topology is determined.
    - During network deployment, the administrator uses the recommended method to deploy and onboard devices. Link aggregation has not been configured when the devices are installed and onboarded. Therefore, if there are aggregated links between the devices, redundant links will be blocked by STP. After the devices are onboarded and register with iMaster NCE-Campus, the administrator checks the topology and deploys link aggregation and service configurations on iMaster NCE-Campus.
- This deployment process applies to scenarios with scattered installation time.



- This deployment process is applicable to scenarios with concentrated installation time. The administrator is advised to plan the network before onboarding devices. If the network cannot be planned in advance, the administrator can onboard devices first and then determine the network topology.
- In this deployment mode, the network is planned before devices are deployed and onboarded.
    - During network deployment, the administrator plans the network topology by entering device ESNs and specifying stack members and aggregated links on iMaster NCE-Campus.
    - Alternatively, the administrator can import the preceding planning information in batches using a template. Using a template to import data in batches simplifies operations and is therefore recommended.
    - Then, the administrator uses the recommended method to deploy and onboard devices.
    - After devices go online and register with iMaster NCE-Campus, iMaster NCE-Campus automatically checks whether the actual topology of the devices is the same as the planned one. If cables are incorrectly connected during installation, iMaster NCE-Campus immediately notifies the administrator.
- Note: Parameters can be planned for an Eth-Trunk interface. After a device goes online, it automatically obtains the corresponding Eth-Trunk configuration.

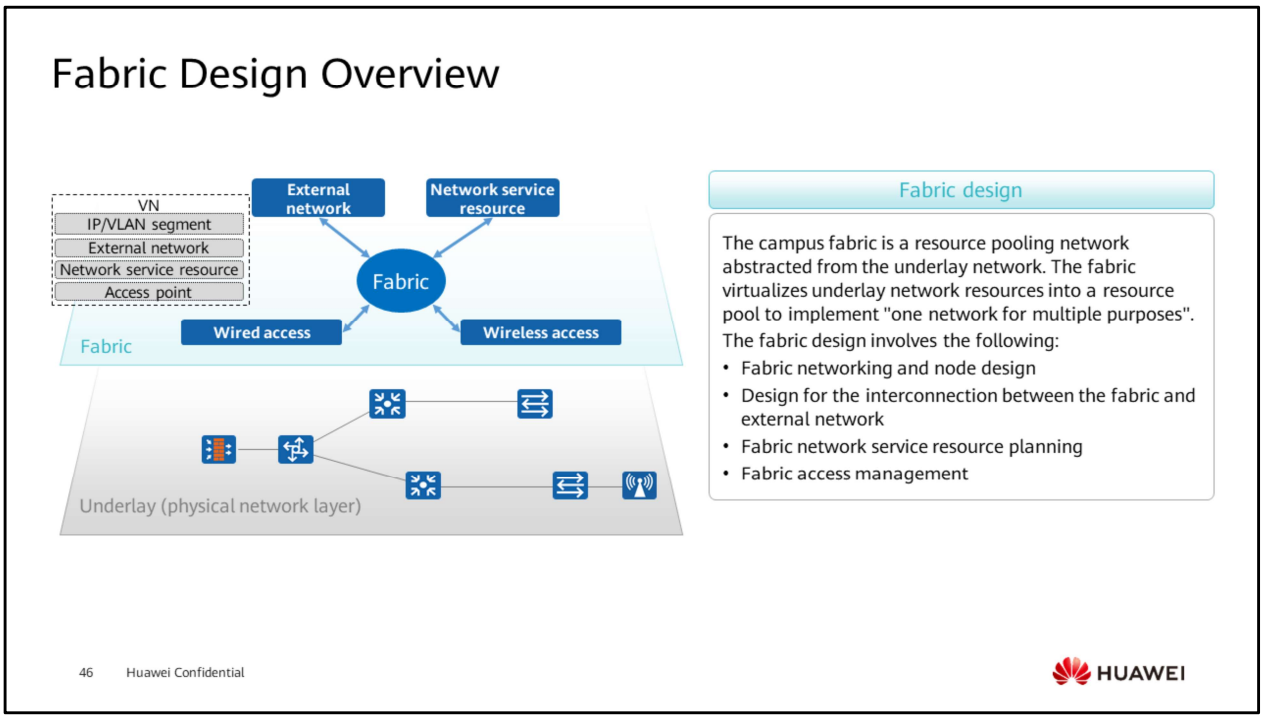


- Automatic routing domain configuration: After this function is enabled, the underlay network can be automatically configured. You can specify sites for automatic routing domain configuration and specify OSPF routing parameters. Currently, the following parameters are supported:
  - Area: In the single-area scenario, all devices belong to Area 0. In the multi-area scenario, border nodes belong to Area 0, and each edge node and its connected border node belong to an area.
  - Network type: You can set the OSPF network type to broadcast, P2MP, or P2P.
  - Encryption: You can set the encryption mode between adjacent devices to HMAC-SHA256, MD5, or none.
  - OSPF GR: You can enable OSPF GR.
- Before enabling the automatic routing domain configuration function, you need to plan network resources required for underlay network automation.
  - Underlay devices (corresponding network scope of the fabric) are connected through VLANIF interfaces at Layer 3, and you need to assign a VLAN to each interconnection link.
  - The VLANIF interfaces for connecting devices are automatically assigned interconnection IP addresses with a 30-bit subnet mask.

Contents

1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. **Fabric Design**
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. WLAN Design
7. Network Security Design
8. O&M Design

HUAWEI



- Fabric network service resource planning:
  - In the resource model design for the fabric network, network service resources are created on the border node so that service terminals on the campus network can access service resources in the network management zone, such as the DHCP server and admission server.
  - Multiple network service resources can be created, and one network service resource model can contain access addresses of multiple service resources.
  - If only a few network service resources need to be accessed in the network management zone, you are advised to plan these resources in the same network service resource model. This saves interconnection VLAN and IP address resources and simplifies route configuration for the network management zone.
- Fabric access management:
  - Configuring access management for the fabric involves configuring authentication control points and planning access point resources for VN creation.
  - The wired access point resource refers to switch interfaces to which terminals will connect, and the wireless access point resource refers to SSIDs with which terminals will associate.

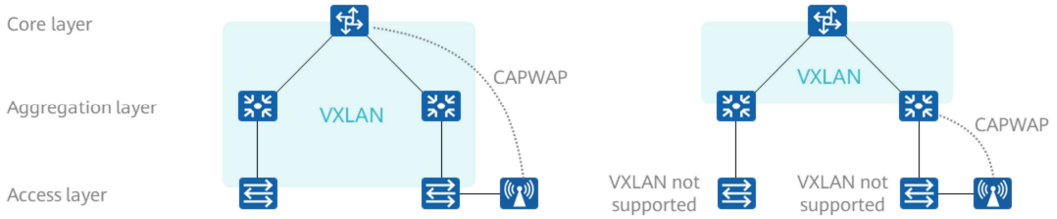
Fabric Networking Scenarios

- In Layer 2 or Layer 3 networking, border or edge nodes can function as gateways.
  - Centralized gateway: A border node functions as the gateway to implement centralized management and simplify O&M.
  - Distributed gateway: Edge nodes function as gateways to facilitate network expansion.
- Recommended networking scenario:
  - Centralized gateway, VXLAN across core and access layers, and native AC deployed at the core layer
  - Distributed gateways, VXLAN across core and aggregation layers, and native AC deployed at the aggregation layer

Networking	Border Node Location	Edge Node Location	AC Deployment Mode	Recommended
Centralized gateway	Core	Access	Native AC on a core device	Yes
			Independent AC attached to a core device in off-path mode	No
		Aggregation	Native AC on a core device	No
			Independent AC attached to a core device in off-path mode	No
Distributed gateway	Core	Aggregation	Native AC on an aggregation device	Yes
			Independent AC attached to a core device in off-path mode	No
		Access	Independent AC attached to a core device in off-path mode	No

HUAWEI

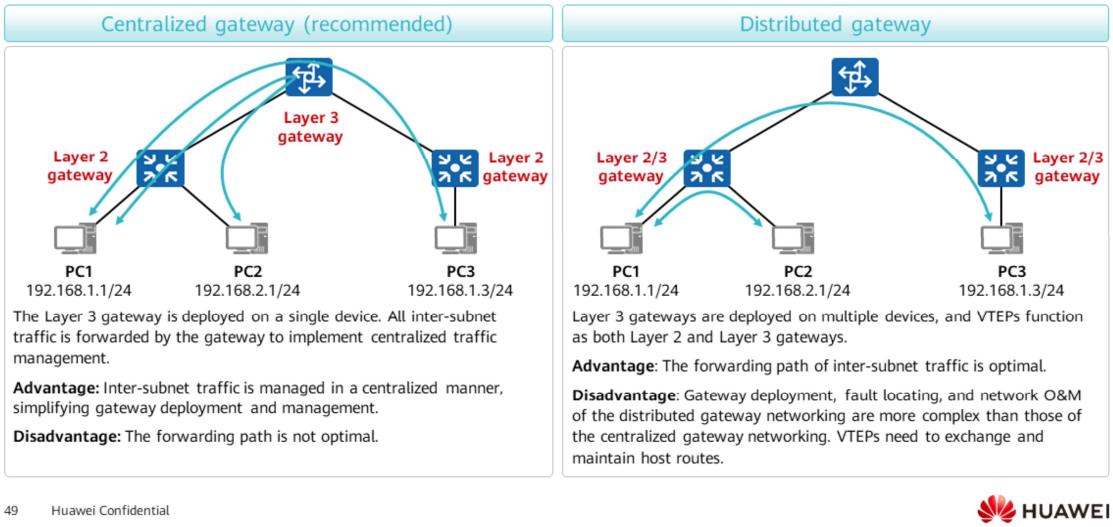
Fabric Networking Design: VXLAN Coverage Range



VXLAN range	VXLAN across core and access layers	VXLAN across core and aggregation layers
Application scenarios	• Automatic end-to-end service deployment is required on the entire network. • Recommended: Number of switch nodes < 1,000	• Existing access switches on the live network need to be reused. Alternatively, low-cost Huawei switches that do not support VXLAN need to be deployed at the access layer to reduce costs. • Recommended: Number of switch nodes < 3,000
Network construction costs	Compared with the VXLAN solution from the core layer to the aggregation layer, this solution has higher network construction costs.	Compared with the VXLAN solution from the core layer to the access layer, this solution has lower network construction costs.

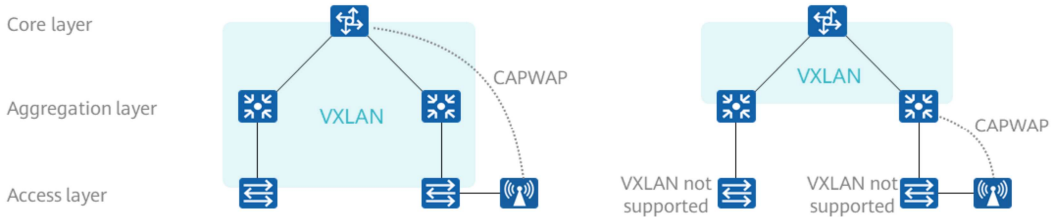
HUAWEI

Fabric Networking Design: Centralized and Distributed Gateways



- In large- and medium-sized campus networks, the virtualization solution is classified into the centralized gateway solution and distributed gateway solution based on the user gateway location. You can select a gateway solution when creating a fabric on iMaster NCE-Campus.
- In the centralized gateway solution, a border node functions as the gateway of all users, and all inter-subnet traffic is forwarded by the border node. In the distributed gateway solution, multiple edge nodes function as user gateways, and inter-subnet traffic is forwarded through these edge nodes.

Fabric Networking Design: Gateway Solution Selection

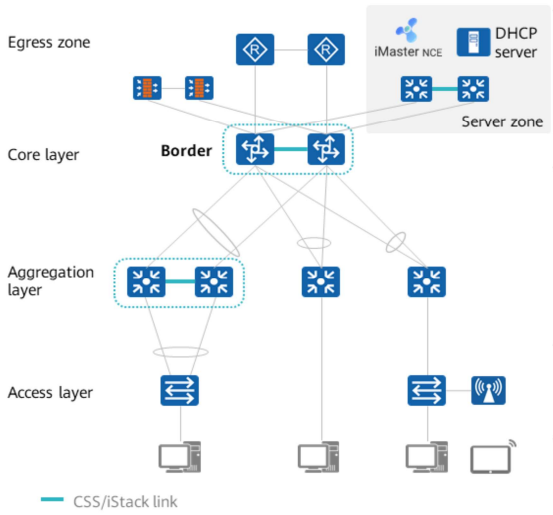


Item	Centralized Gateway	Distributed Gateway
User gateway location	Border node	Edge node
Recommended edge location	Access switch	Aggregation switch
O&M deployment	A border node functions as the gateway of all users, and the native AC function is typically enabled on the border node to configure wireless services, simplifying O&M deployment.	Multiple edge nodes function as user gateways. This ensures that only part of the network is affected upon the failure of one gateway and facilitates network expansion. The native AC function is typically enabled on these edge nodes. Wireless services need to be configured on all edge nodes, complicating O&M deployment.
Terminal scale	≤ 50,000 (calculated based on the number and specifications of centralized gateways). This solution is recommended when the number of terminals does not exceed 50,000.	≤ 100,000 (calculated based on the number and specifications of distributed gateways). This solution is recommended when the number of terminals exceeds 50,000.

HUAWEI

- When designing the virtualization solution, first determine the gateway solution to be used. After the gateway solution is determined, you can perform end-to-end design on the entire campus network based on the selected gateway solution.
- The centralized gateway solution supports only one border node, whereas the distributed gateway solution supports multiple border nodes.

Fabric Node Design: Single-Border Networking

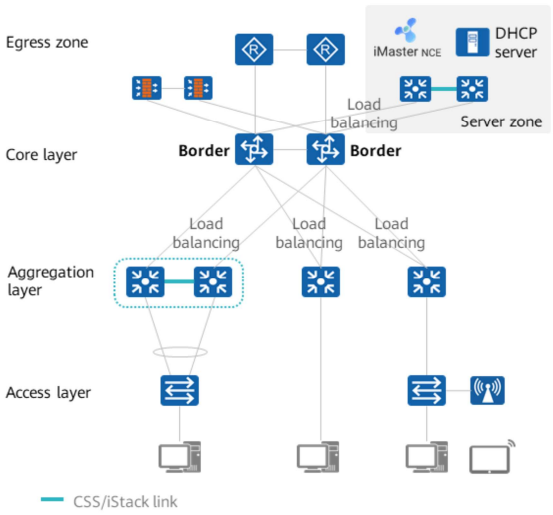


- Core layer design:
  - Deploy a cluster of core switches or a single core switch. It is recommended that large-capacity modular switches set up a cluster and S12700E series switches be used as core switches.
  - It is recommended that ENP cards be used for interconnection with aggregation switches and that Eth-Trunk be deployed.
- Aggregation layer design:
  - Deploy a stack of aggregation switches or a single aggregation switch. It is recommended that a cluster of modular switches or a stack of fixed switches be deployed. Large-capacity switches (S12700E or S6700 series switches) are recommended.
  - Eth-Trunk is recommended for interconnection with the core layer and access layer.
- Access layer design:
  - Wired access node: Each node can be a stack or a single device.
  - Wireless access node: Fit AP
- Server zone design:
  - It is recommended that iMaster NCE-Campus, analyzer, and DHCP server be connected to the campus network through switches in the server zone.

HUAWEI

- An Ethernet Network Processor (ENP) card is embedded with the Huawei-developed ENP. The card can function as a common LPU to provide data access and switching services and also as a WLAN AC to provide wireless access control functions. In this way, the card achieves wired and wireless convergence.

Fabric Node Design: Dual-Border Networking



- Core layer design:
  - Core switches are two logical devices. Large-capacity switches (S12700E series switches) are recommended. It is recommended that ENP cards be used for interconnection with aggregation switches and that Eth-Trunk be deployed.
- Aggregation layer design:
  - It is recommended that aggregation switches set up a stack to implement load balancing with core switches.
- Access layer design:
  - The design is similar to that in the single-border networking scenario.
- Server zone design:
  - It is recommended that iMaster NCE-Campus, analyzer, and DHCP server be connected to the campus network through switches in the server zone. It is recommended that aggregation switches set up a stack to implement load balancing with core switches.

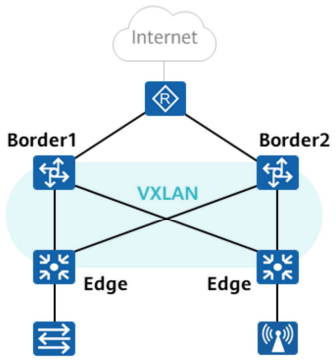
HUAWEI

Fabric Node Design: Single/Dual-Border Networking Selection

	Single-Border Networking	Dual-Border Networking
Description	The core switch is a logical device, which can be a physical standalone device or a CSS/iStack system.	Core switches are two logical devices, which can be two standalone physical devices or two CSS/iStack systems.
Characteristics	• Easy management, configuration, and maintenance. • Services may be affected during the device upgrade.	• Complex management, configuration, and maintenance. • The upgrade of core switches has little impact on services.
Application scenario	Simple network management, configuration, and maintenance are required.	• The device upgrade has little impact on services. • Core switches do not support clustering or stacking due to some reasons (such as the deployment distance).
Restrictions	There is no restriction. The single-border networking design is applicable to both centralized and distributed VXLAN gateway networking.	The dual-border networking design is applicable only to distributed VXLAN gateway networking.

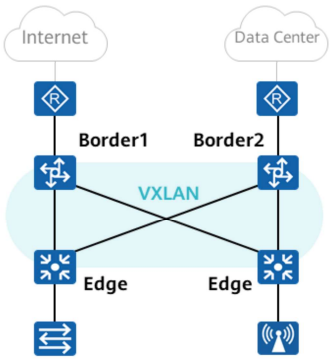
HUAWEI

Fabric Node Design: Multi-Border Networking



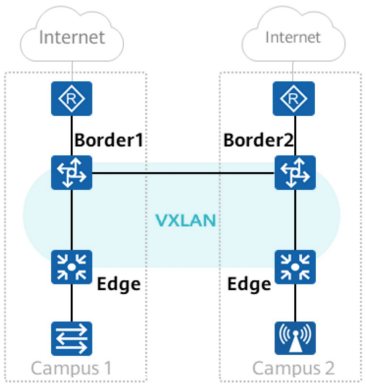
Scenario description: On a single campus network, multiple border nodes are connected to the same egress for egress redundancy.

Application scenario: A single campus network has multiple border nodes that are connected to the same external network, implementing reliability in non-stacking scenarios.



Scenario description: On a single campus network, multiple border nodes are connected to different egresses, and different services are transmitted through different border nodes.

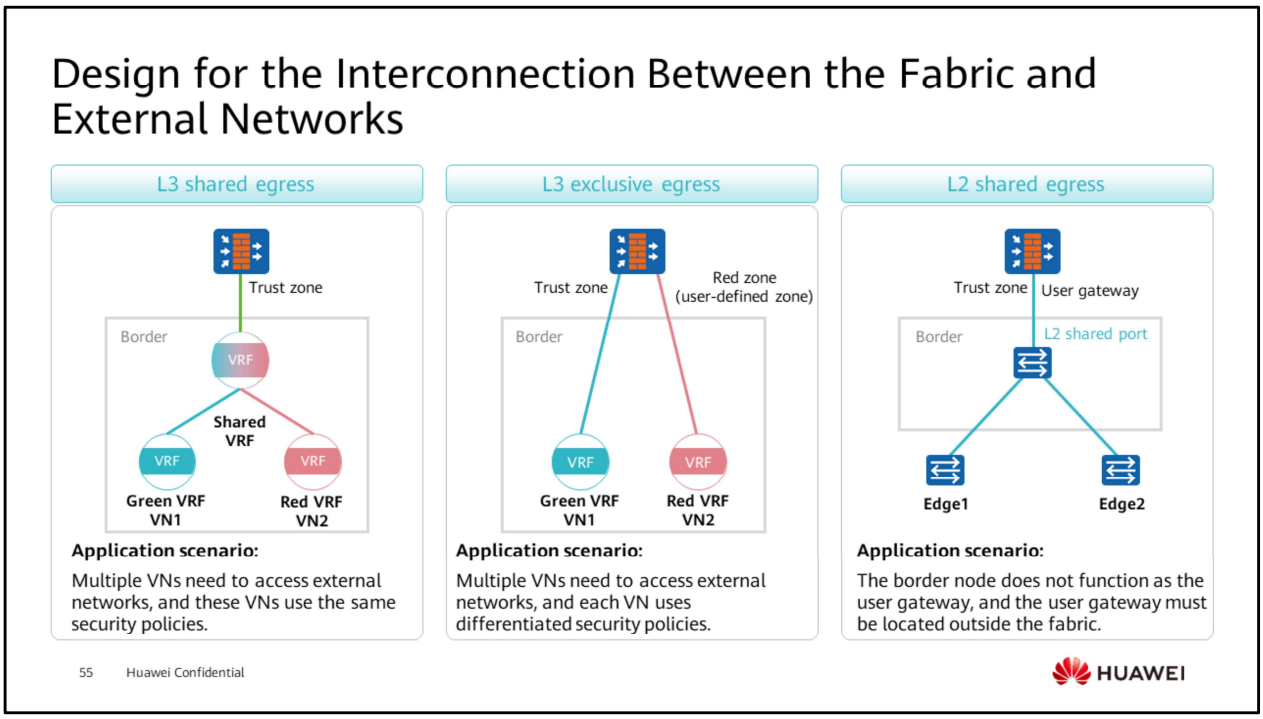
Application scenario: A single campus network is connected to multiple external networks.



Scenario description: There are multiple campuses, and each campus is connected to its external network through an independent border node.

Application scenario: Multiple campus networks belong to the same fabric, and each campus network has an independent border node and egress network.

HUAWEI



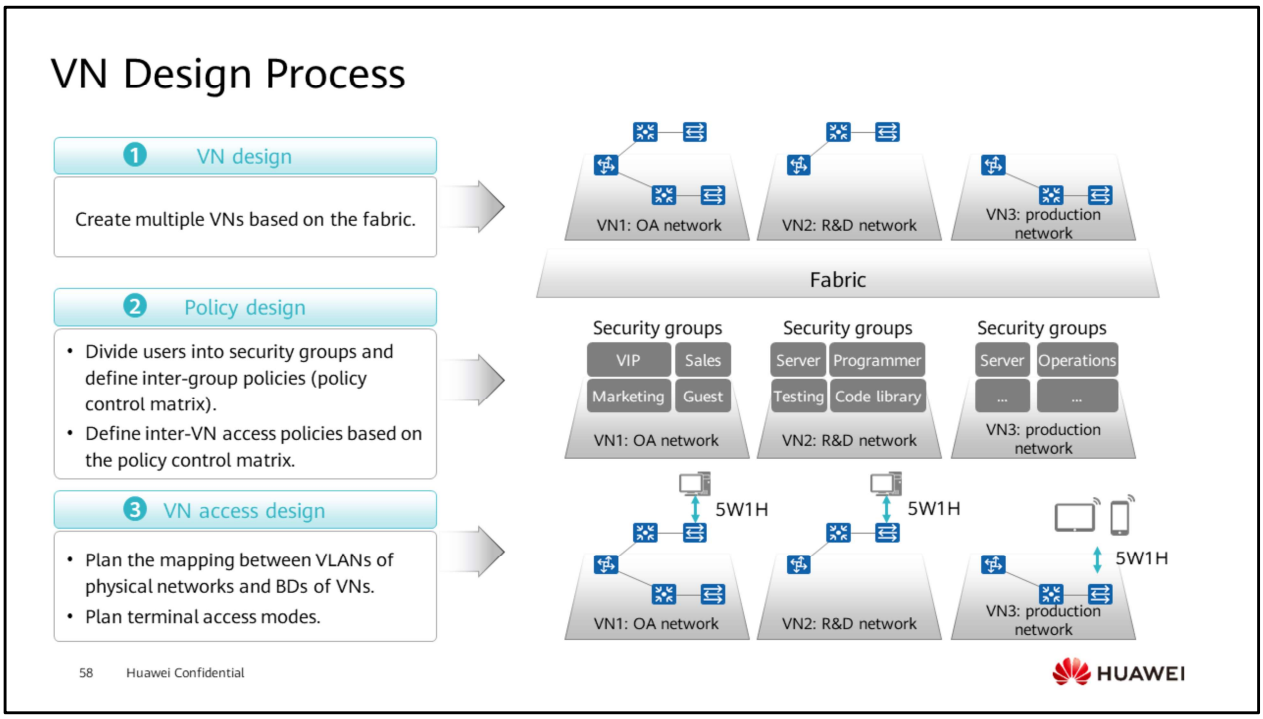
- Terminals' network access traffic inside the campus network is forwarded through VXLAN tunnels. When the campus network needs to communicate with an external network, the data must pass through the border node. For example, a campus network communicates with an external network such as the Internet, DC, or another branch through the border node.
- There are three interconnection modes between the fabric network and egress device: L3 shared egress, L3 exclusive egress, and L2 shared egress. The L3 shared egress mode is applicable to the scenario where the firewall does not need to perform security check on VNs and traffic of all VNs is transmitted in the same security zone. The L3 exclusive egress mode is suitable for the scenario where the firewall needs to perform security check on VNs and traffic of VNs is transmitted in multiple security zones. The L2 shared egress mode applies to the scenario where the user gateway must be located outside the fabric, and the border node and egress firewall are connected at Layer 2.
    - In L3 shared egress mode, all VNs shares the same VPN instance (VRF) when accessing an external network, and static routes or dynamic BGP routes can be used for route advertisement between the border node and egress firewall.

        - Notes for static route design
            - A border node learns global internal routes and it has a default route that points to the egress firewall.
            - A VPN instance on the border node is configured with a static default route that points to the interface connected to the egress firewall.
            - If there are multiple egress firewalls, you can deploy association between Network Quality Analysis (NQA) or Bidirectional Forwarding Detection (BFD) and static routes for route switching.
        - Notes for BGP route design
            - The border node establishes a BGP connection with a firewall, summarizes user routes in VPN instances, and imports the summarized routes to BGP.
    - L3 exclusive egress mode: A firewall is partitioned into security zones with the same number as VPN instances and connects to the border node through physical interfaces or sub-interfaces. On the border node, a VPN instance is configured with a static default route that points to the corresponding physical interface or sub-interface of the security zone.
    - L2 shared egress mode: The border node connects to the egress device through a Layer 2 interface, and the egress device is used as the user gateway to provide access to the external network.
- Application scenario:
    - Typically, there are only a few egress routes between the campus network and external network. Therefore, the L3 shared egress mode is recommended, and static routes are used for route advertisement between the border node and egress firewall.
    - When the egress firewall needs to perform security check on each VN based on security zones, the L3 exclusive egress mode can be used.
    - The L2 shared egress mode applies only to the scenario where the user gateway must be located outside the fabric. In higher education scenarios, if a BRAS is used as the user authentication point and PPPoE dialup is required on the network, the L2 shared egress mode can be used.

Contents

1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. Fabric Design
4. **Overlay Network Design**
5. Admission Control and Free Mobility Design
6. WLAN Design
7. Network Security Design
8. O&M Design

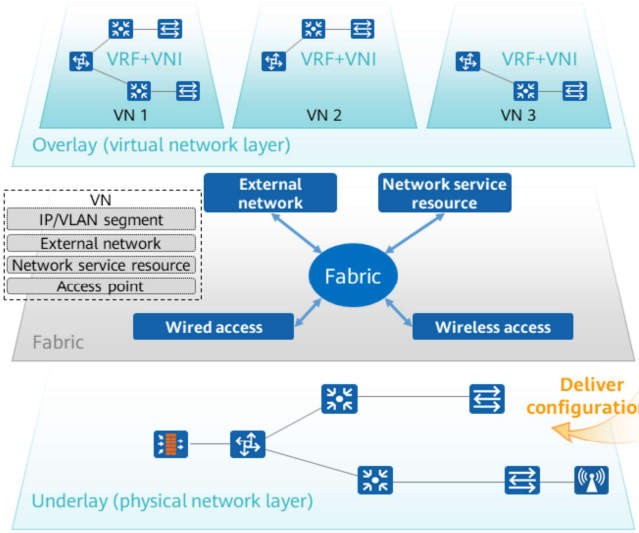
HUAWEI



- VN design:
  - VNs are generally divided based on services on a campus network. An independent service is assigned a VN, and VNs are isolated from each other by default. For example, on a school campus network, guest, teaching, IoT, and video surveillance services can each be assigned a separate VN. On an enterprise campus network, office network services, production network services, and R&D services can be allocated to different VNs.
- Policy design:
  - Some isolation requirements caused by different user roles can be fulfilled by deploying policy control technologies (such as free mobility).
  - VNs are isolated by default. To enable mutual access between user groups within a VN, additional VN interworking configuration is required.
- VN access design:
  - Service data enters from a physical network to a VN through an edge node. Service data of users enters different VNs depending on the VLANs to which users belong. Therefore, during network design, you need to plan the mappings between VLANs of physical networks and BDs of VNs, and configure VLANs for wired and wireless users.

- 5W1H:
  - Who: identity of an access user, for example, a company leader, a common employee, or a guest.
  - Where: user access location, for example, access from within the campus, or remote access.
  - What: type of the terminal used by the access user, for example, mobile phone or PC/laptop.
  - When: time when a user accesses the network, for example, whether the user accesses the network in the daytime or at night.
  - Whose: device owner, for example, whether the device is company-issued or BYOD.
  - How: access mode of a user, for example, wired or wireless access.

VN Design



1. Network service abstraction

Physical network resources are pooled through orchestration, and the network is abstracted into Fabric as a Service (FaaS). A VN is a FaaS instance and includes:

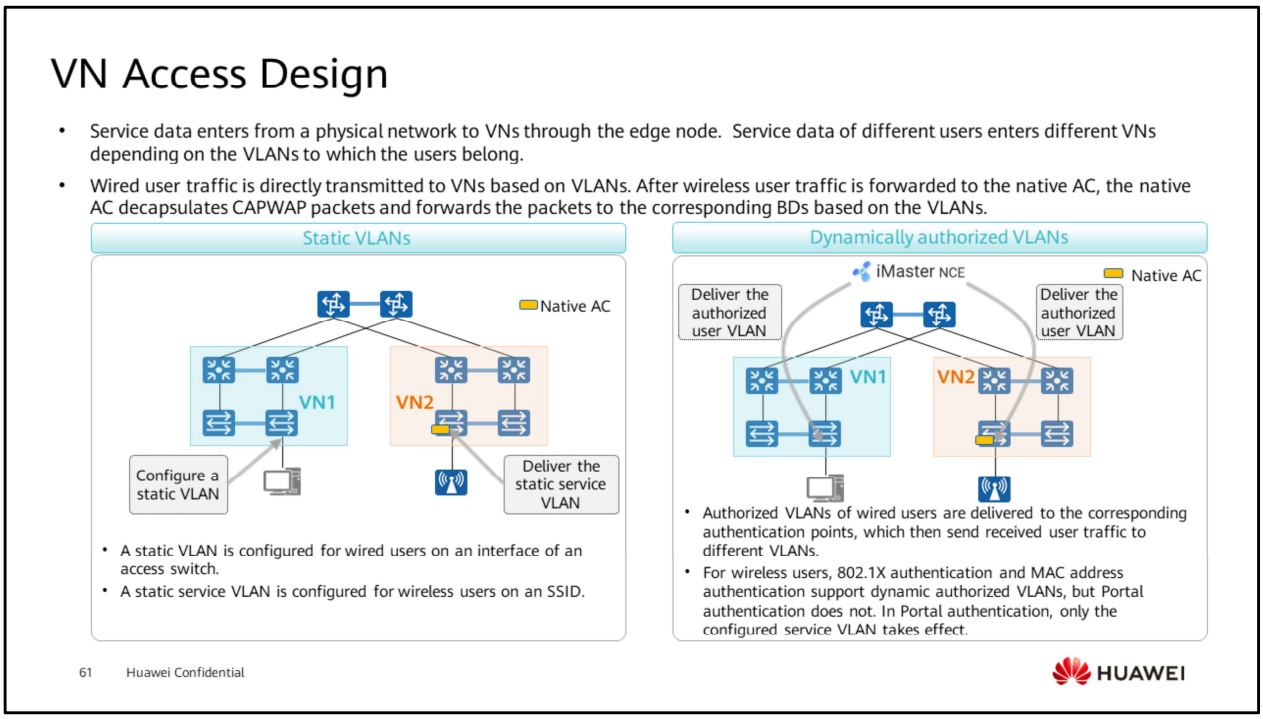
- IP/VLAN segment
- External network
- Network service resource: The IP/VLAN segment is the capability provided by the VN for clients to use network resources.
- Access point: Terminals access VNs through access points.

Deploy VNs

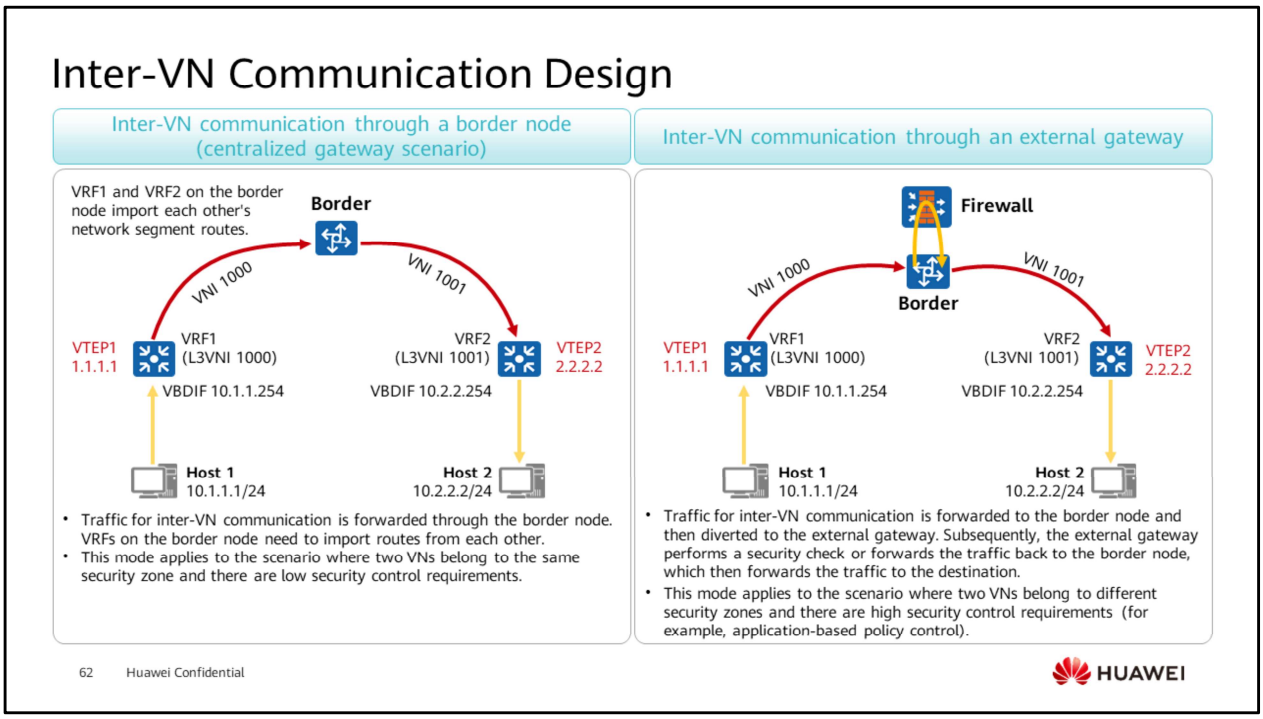
2. Network service orchestration

- Deliver the mappings between VNIs and BDs.
- Deliver the mappings between BDs and VLANs.
- Deliver the IP address segments corresponding to VBDIF interfaces.
- Deliver VRFs and bind them to VBDIF interfaces.

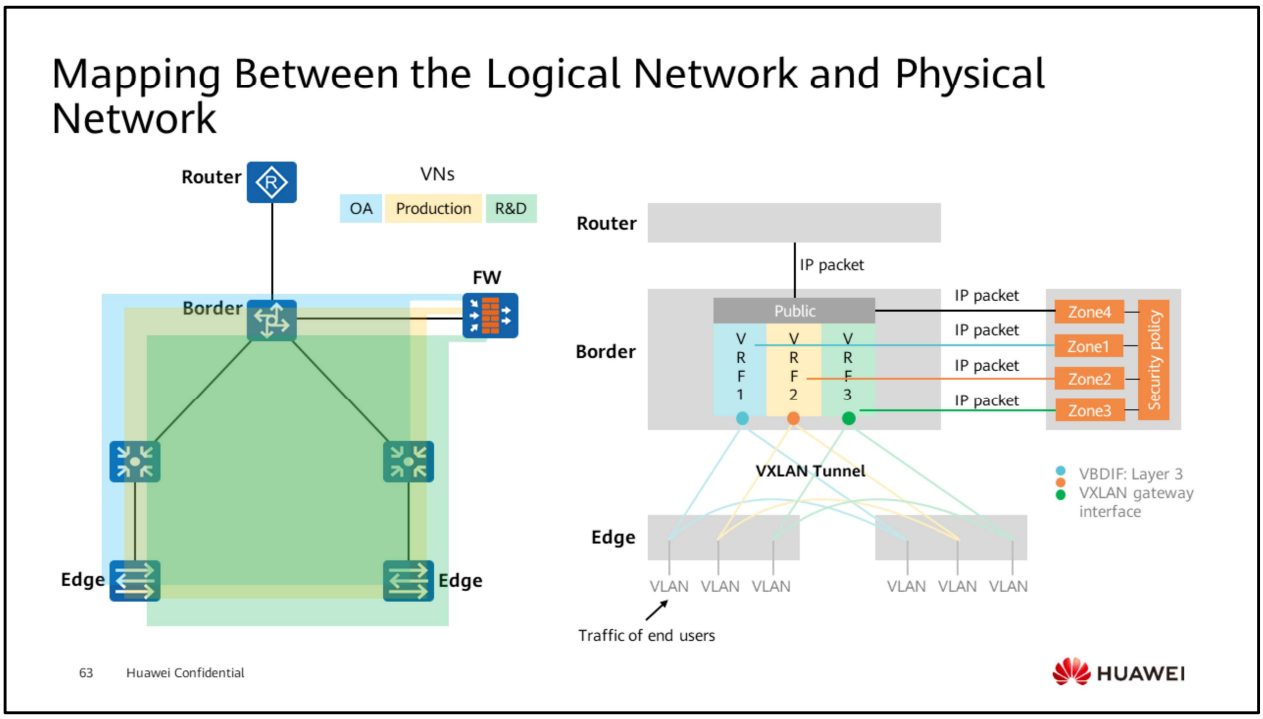
HUAWEI



- Application scenarios:
  - The static VLAN mode applies when terminals access the network at fixed locations and do not need to be authenticated. This access mode is more secure but lacks flexibility. When the locations of terminals change, you need to configure static VLANs again.
  - The dynamically authorized VLAN mode applies when terminals need to access the network from any place and need to be authenticated based on the VLAN information delivered during user authentication. This access mode is flexible and the configuration does not need to be changed when the locations of terminals change. Dynamic access is more automated, easy to manage and use, and is therefore recommended.
- VN access design for wireless users:
  - If distributed gateways are used and edge nodes have the native AC function deployed, traffic of wireless users is forwarded to the edge nodes through CAPWAP tunnels. After the edge nodes decapsulate CAPWAP packets, the decapsulated packets enter different VNs depending on the VLANs to which the wireless users belong.
  - If centralized gateways are used and border nodes have the native AC function deployed, it is recommended that traffic of wireless users be directly forwarded to the border nodes through CAPWAP tunnels. The border nodes then decapsulate CAPWAP packets and forward them to different BDs based on the VLANs. The administrator needs to configure different VLAN ranges for wired and wireless terminals. The VLANs of wired terminals are bound to BDs on the edge nodes, whereas the VLANs of wireless terminals are bound to BDs on the border nodes.



- Inter-VN communication can be implemented through a border node or an external gateway.
  - Through a border node
    - If two VNs belong to the same security zone and have low security control requirements, devices on the two VNs can directly communicate with each other through a border node. In addition, permission control can be implemented based on the free mobility policy. To implement communication between VNs, the border node needs to import their respective network segment routes that are reachable to each other.
  - Through an external gateway
    - If two VNs belong to different security zones and have high security control requirements, it is recommended that devices on the two VNs communicate through an external gateway (a firewall) and that a security zone policy be configured on the firewall for permission control.



- Each VN corresponds to a VRF.
- Each VN corresponds to multiple subnets, and each subnet corresponds to a BD and VNI.
- Traffic from a terminal enters the VN based on the VLAN to which the terminal belongs.

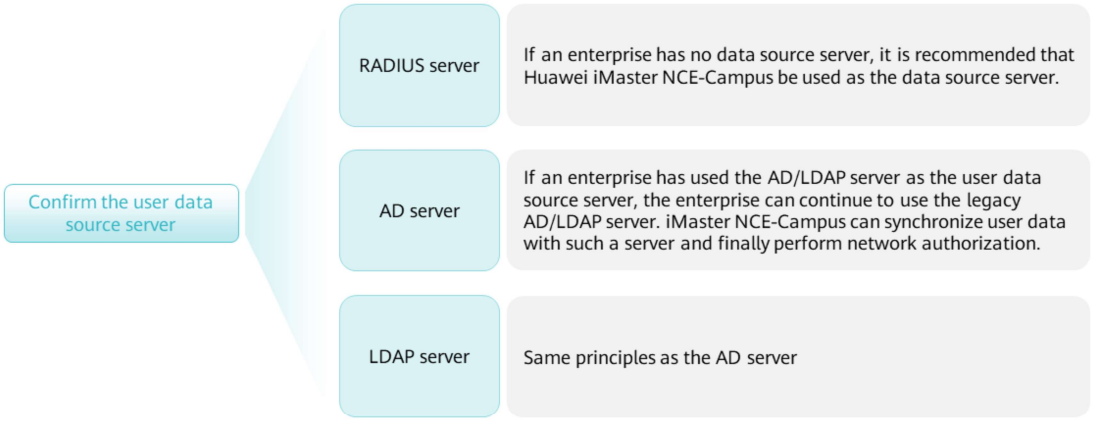
Contents

1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. Fabric Design
4. Overlay Network Design
5. **Admission Control and Free Mobility Design**
6. WLAN Design
7. Network Security Design
8. O&M Design

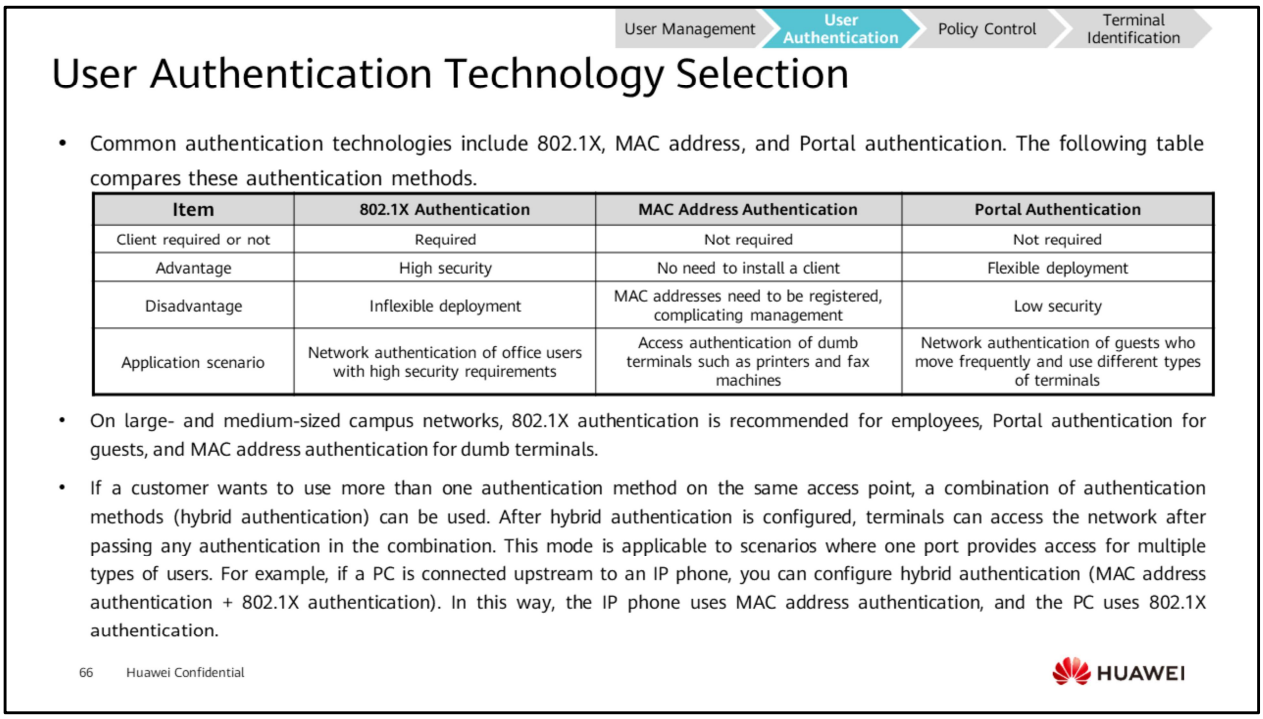
HUAWEI

User Management Solution Design

- User management solution design involves confirming the user data source server. Typical user data source servers of enterprises include RADIUS, AD, and LDAP servers.



HUAWEI



- If 802.1X or MAC address authentication (Layer 2 authentication technologies) is used, the authentication point must be on the same network segment as the user host. It is recommended that the access device function as the authentication point.
    - If a large number of access devices are deployed, the aggregation or core device needs to be deployed as the authentication point. You can also deploy policy association. The gateway functions as the authentication control point, and the access device functions as the authentication enforcement point, thereby simplifying policy deployment.
- If Portal authentication is used, the authentication point can be deployed anywhere as long as it is routable to the user host. Layer 2 Portal authentication is recommended.

Association Between Authenticated Users and VNs

1. Create VNs on iMaster NCE-Campus.

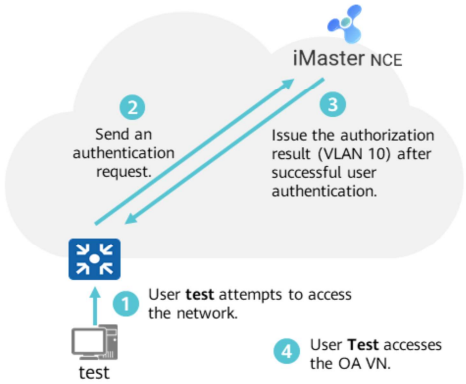
The administrator creates VNs and specifies the IP network segments and VLANs for the VNs.

OA VN
VLAN 10: 10.1.10.0/24
VLAN 20: 10.1.20.0/24

R&D VN
VLAN 30: 10.1.30.0/24
VLAN 40: 10.1.40.0/24

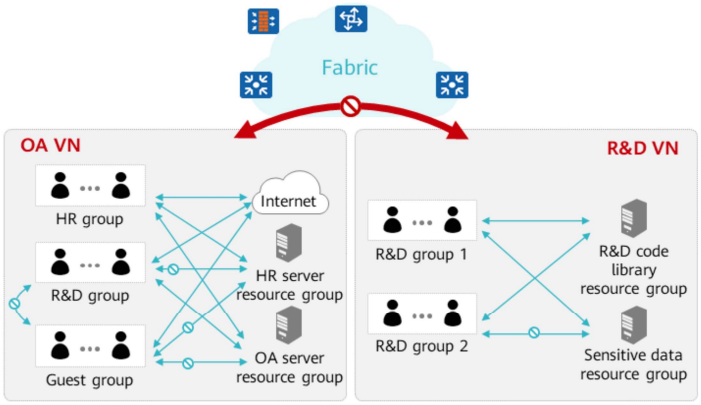
The administrator configures user authentication rules, authorization results, and authorization rules, so that users can obtain corresponding VLANs after being authenticated. For example, the administrator configures user **test** to obtain the authorized VLAN 10 after passing 802.1X authentication.

2. Authenticated users obtain authorized VLANs and are then associated with VNs based on the VLANs.



HUAWEI

Multi-Level Design for Access Policies



- Network access policies can be logically divided into two levels:
  - VN-level access policy: By default, VNs cannot communicate with each other, and service data of each VN is isolated. The CloudCampus Solution enables VNs to communicate within a fabric or through an external network.
  - Security group-level access policy: Users and network resources are divided into different security groups for efficient management and control of inter-group traffic. The free mobility policy enforcement point executes inter-group access policies.

HUAWEI

Policy Design: Create Security Groups

Dynamic security groups
Terminals and users accessing the network are divided into different security groups through NAC authorization.

- Define security groups for enterprise employees based on the enterprise organizational structure.



- Define a security group for a special group of people in the organizational structure, such as the VIP group.
- Define a security group for non-employees that access the network, such as the guest group.
- Define a security group for dumb terminals connected to the network, such as printers and access control devices.

Dynamic security groups are authorized based on the 5W1H principle.

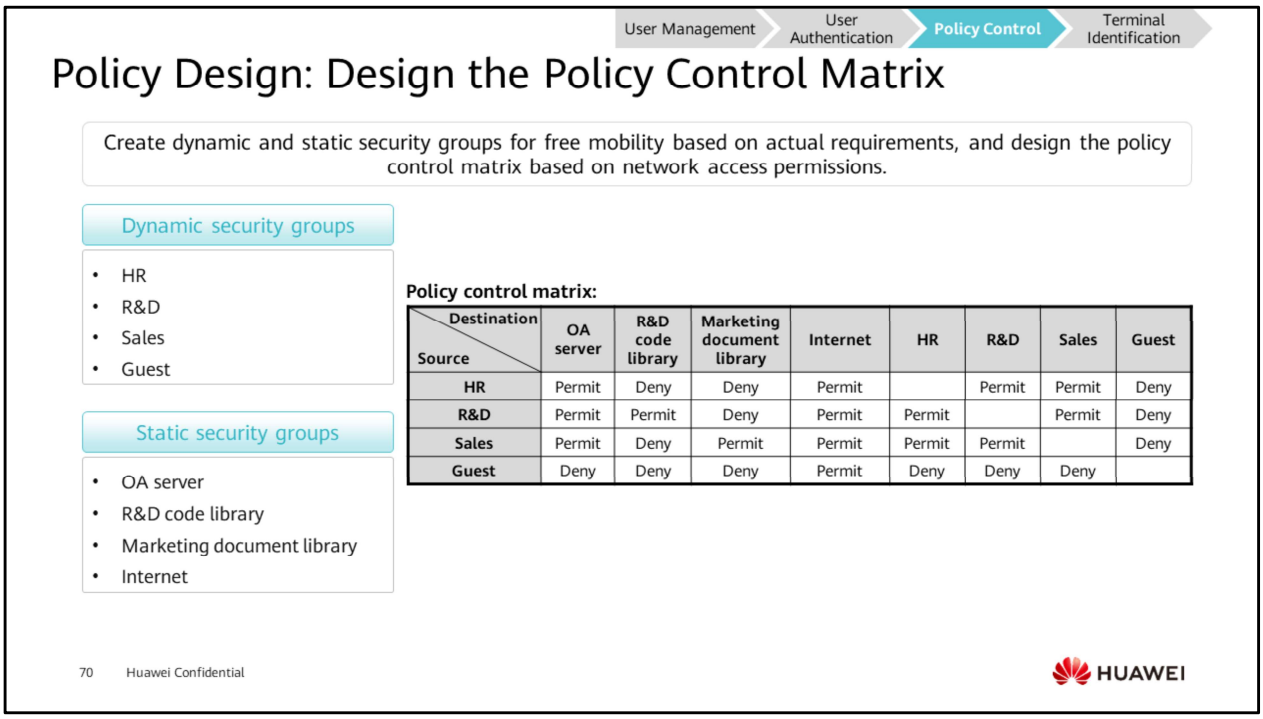
Static resource groups
Configure a group of IP address resources as a static resource group.



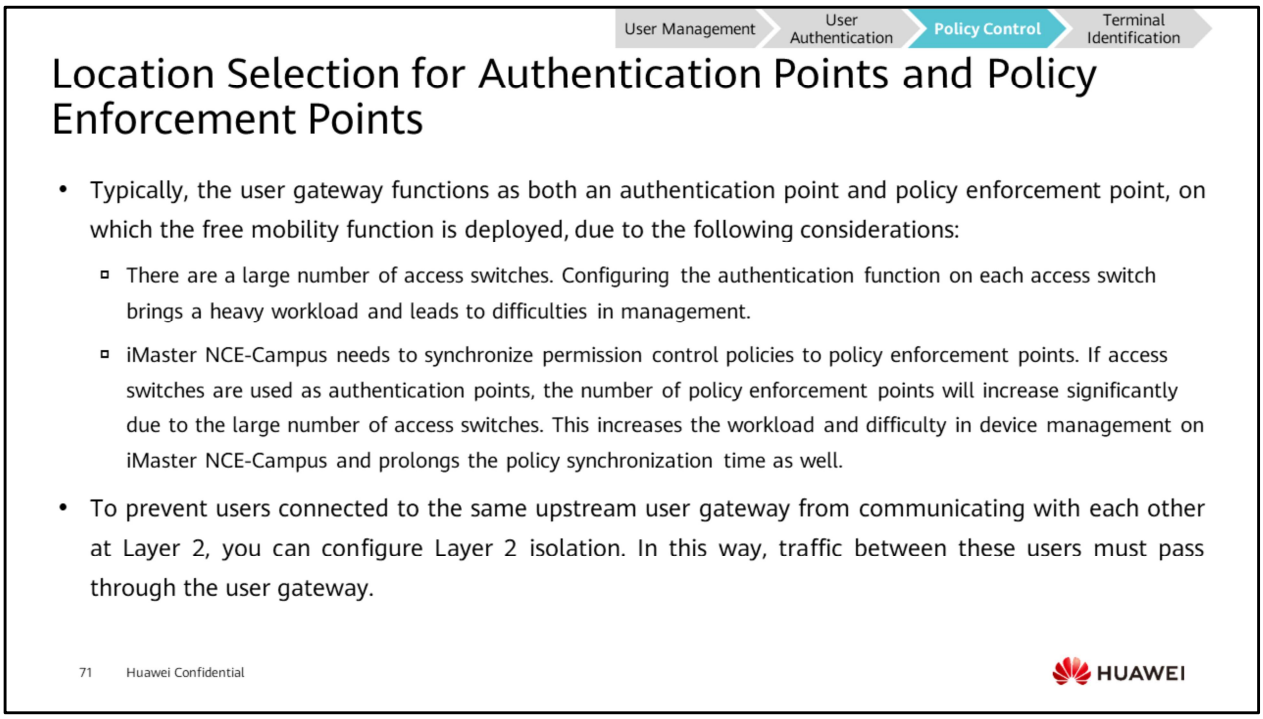
- Define security groups, such as the OA resource group and R&D resource group, for intranet resources based on security requirements.
- If Internet access control is required, define an Internet resource group.

Aggregate the IP addresses of dispersed data center applications into a resource group, thereby simplifying deployment.

HUAWEI



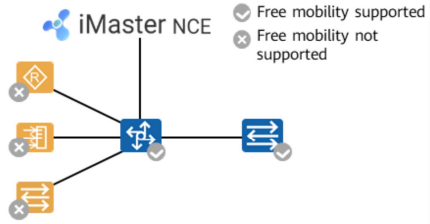
- An inter-group policy directly reflects whether two security groups can communicate with each other. When planning an inter-group policy, simply set the inter-group policy to permit or deny based on whether the two groups can communicate with each other. The administrator can configure inter-group policies through an intuitive policy control matrix on iMaster NCE-Campus. If no inter-group policy is created between source and destination security groups, the default policy between them is permit.
- Take the policy direction into account when planning inter-group policies. Typically, packets are transmitted in two directions between two terminals.
    - Huawei switches consider that traffic from A to B is unrelated to traffic from B to A, so they match policies for the two types of traffic separately and determine whether to forward the traffic based on the corresponding policies. This means that a Huawei switch enforces policies on a packet by considering only the source and destination security groups of that packet. For example, the policy "A ->B permit, B -> A deny" means that all packets sent from A to B will be permitted, whereas all packets sent from B to A will be discarded, regardless of whether A or B initiates the access. The default inter-group policy on a switch is permit.



- If an authentication point and policy enforcement point are deployed on different devices, the IP-security group entries of authenticated users need to be pushed to the specified policy enforcement point. To achieve this, the administrator needs to configure IP-security group entry subscription, that is, specify to which policy enforcement point the IP-security group entries of a certain network segment or security group need to be pushed.

IP-Security Group Entry Synchronization

Technical background

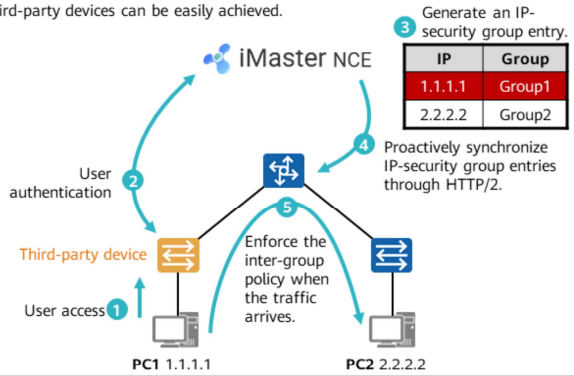


These devices do not support free mobility, so how to realize free mobility if a solution includes these devices?

- Switches that do not support free mobility
- ACs
- Routers
- Third-party devices (non-Huawei)

IP-security group entry synchronization principle

iMaster NCE-Campus synchronizes the mappings between user IP addresses and security groups to the switches functioning as policy enforcement points. In this way, authentication points and policy enforcement points can be separated, thereby implementing flexible networking. In addition, hybrid networking with third-party devices can be easily achieved.



HUAWEI

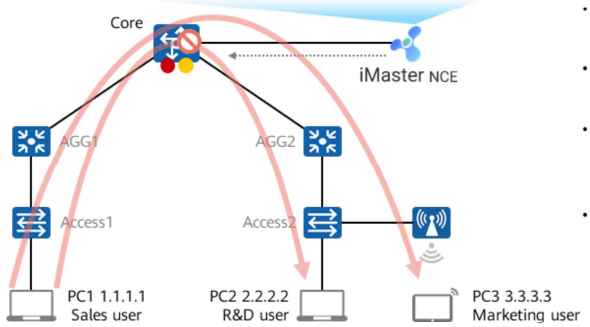
Free Mobility Solution (1)

Security group

Group Name	Group ID
Sales	1
R&D	2
Marketing	3
...	...

Security group-based policy control matrix

	Sales	R&D	Marketing	...
Sales	√	×	√	...
R&D	×	√	√	...
Marketing	√	√	√	...
...	...	...	...	...



Scenario description

- Centralized authentication point + centralized policy enforcement point.
- The authentication point and policy enforcement point are combined on the core device.
- Devices do not support VXLAN.

Scenario characteristics

- The core device functions as the centralized authentication point for both wired and wireless users on the entire network.
- The core device also works as the policy enforcement point for free mobility.
- The core device stores authentication information about all users on the network. After traffic is forwarded to the core device, it enforces policies based on the policy control matrix defined by the administrator.
- The network does not need to support or deploy VXLAN.

● Authentication point ● Policy enforcement point

Security group and policy control matrix delivery

HUAWEI

Free Mobility Solution (2)

Security group

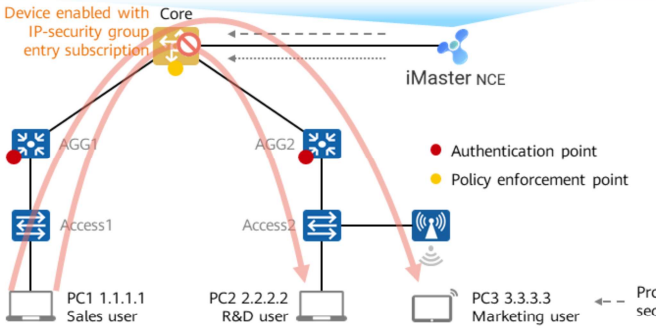
Group Name	Group ID
Sales	1
R&D	2
Marketing	3
...	...

Security group-based policy control matrix

	Sales	R&D	Marketing	...
Sales	√	×	√	...
R&D	×	√	√	...
Marketing	√	√	√	...
...	...	...	...	...

IP-security group mapping table

IP Address	Group ID
1.1.1.1	1
2.2.2.2	2
3.3.3.3	3
...	...



Scenario description

- Distributed authentication points + centralized policy enforcement point.
- Authentication points are separated from the policy enforcement point.
- Devices do not support VXLAN.

Scenario characteristics

- Multiple authentication points exist on the entire network, for example, distributed on aggregation switches.
- The core device functions as the centralized policy enforcement point for free mobility.
- iMaster NCE-Campus proactively synchronizes IP-security group entries to the core device. After traffic is forwarded to the core device, it enforces policies based on the policy control matrix defined by the administrator.
- The network does not need to support or deploy VXLAN.

HUAWEI

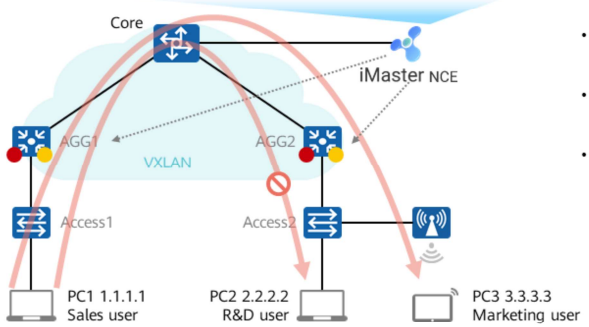
Free Mobility Solution (3)

Security group

Group Name	Group ID
Sales	1
R&D	2
Marketing	3
...	...

Security group-based policy control matrix

	Sales	R&D	Marketing	...
Sales	√	×	√	...
R&D	×	√	√	...
Marketing	√	√	√	...
...	...	...	...	...



Scenario description

- VXLAN-based virtualized campus network, with automatic service provisioning.
- Distributed authentication points + policy enforcement points.

Scenario characteristics

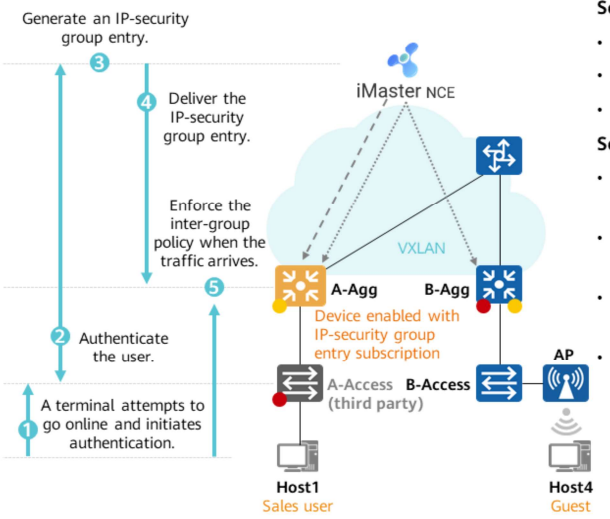
- Multiple authentication points exist on the entire network, for example, distributed on aggregation switches.
- Aggregation switches support VXLAN and function as policy enforcement points.
- The policy for traffic exchanged between users on the same aggregation switch is enforced by the aggregation switch itself.
- Traffic exchanged between users on different aggregation switches is encapsulated using VXLAN. The source security group ID is encapsulated into VXLAN-encapsulated traffic. Policies are enforced on the peer aggregation switch.

● Authentication point ● Policy enforcement point

Security group and policy control matrix delivery

HUAWEI

Free Mobility Solution (4)



Scenario description

- VXLAN-based virtualized campus network.
- Distributed authentication points + policy enforcement points.
- The third-party device functions as a user authentication point.

Scenario characteristics

- A third-party switch exists on the network and functions as a user authentication point.
- Aggregation switches support VXLAN and function as policy enforcement points.
- The policy for traffic exchanged between users on the same aggregation switch is enforced by the aggregation switch itself.
- Traffic exchanged between users on different aggregation switches is encapsulated using VXLAN. The source security group ID is encapsulated into VXLAN-encapsulated traffic. Policies are enforced on the peer aggregation switch.

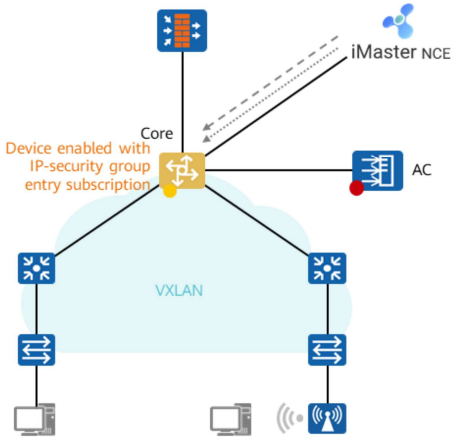
● Authentication point ● Policy enforcement point

Proactively synchronize IP-security group entries

Security group and policy control matrix delivery

HUAWEI

Free Mobility Solution (5)



Scenario description

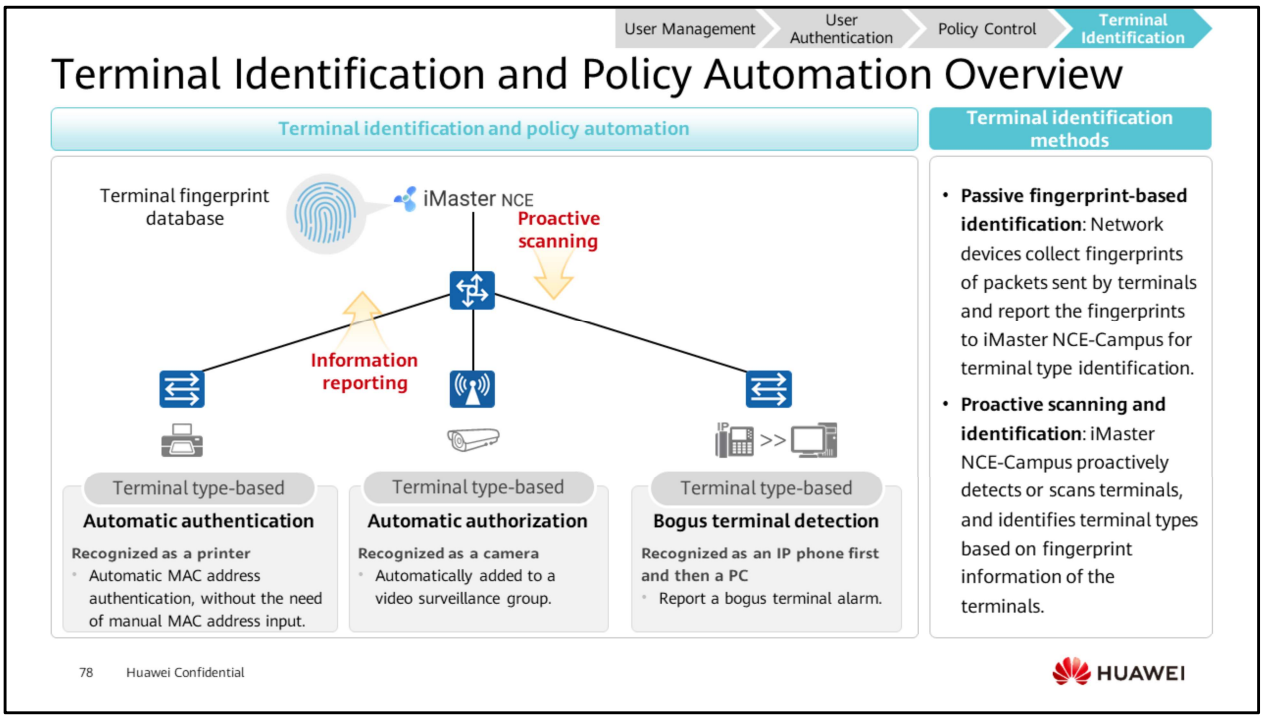
- VXLAN-based virtualized campus network
- The AC functions as the authentication point for wireless users, and the core switch functions as the gateway for wireless users.

Scenario characteristics

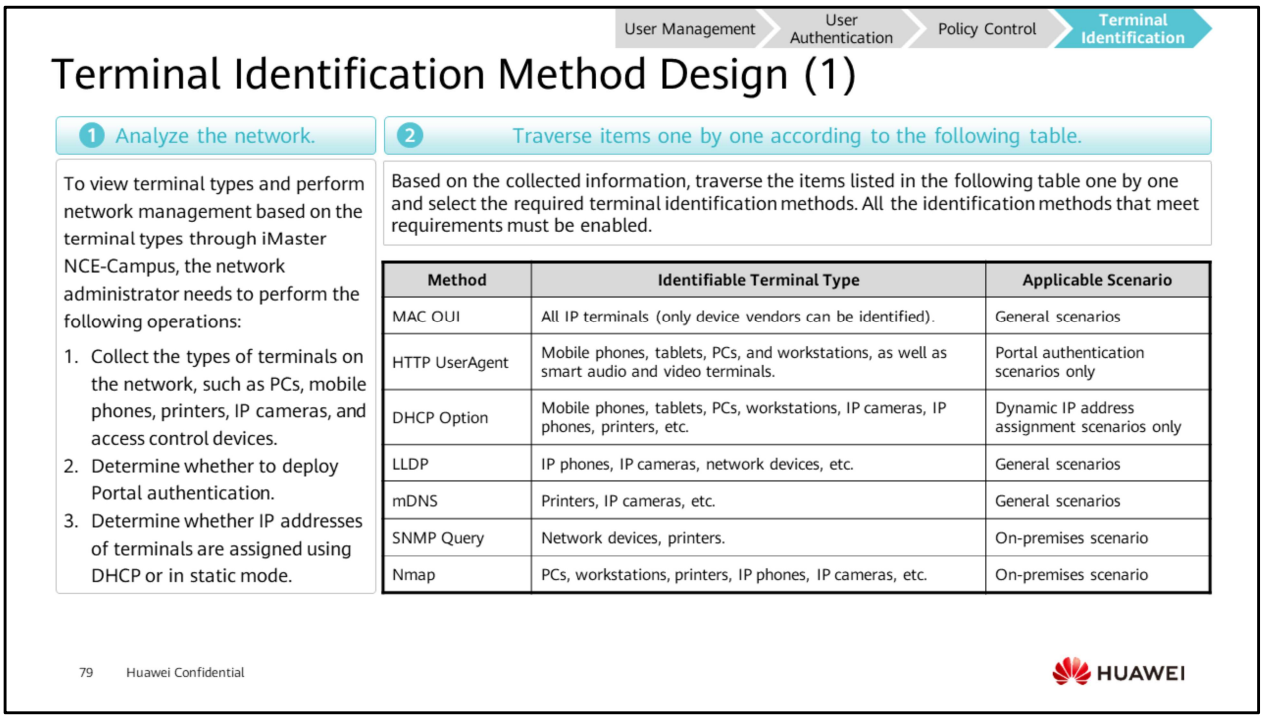
- The AC functions as the authentication point for wireless users and interacts with iMaster NCE-Campus to complete user authentication.
- The core switch functions as the free mobility policy enforcement point, to which iMaster NCE-Campus synchronizes IP-security group entries.

Authentication point
Policy enforcement point
Proactively synchronize IP-security group entries
Security group and policy control matrix delivery

HUAWEI



- On large- and medium-sized campus networks, access terminals include smart terminals (such as PCs and mobile phones) and dumb terminals (such as IP phones, printers, and IP cameras). Currently, terminal management on campus networks faces the following challenges:
    - The network management system (NMS) can only display the IP and MAC addresses of access terminals, but cannot identify the specific terminal type. As a result, the NMS cannot provide refined management for network terminals.
    - Network service configurations and policies vary according to the terminal type. Consequently, administrators need to manually configure different services and policies for each type of service terminals, complicating service deployment and operations.
- To address these challenges, Huawei provides the automatic terminal identification and policy delivery solution, which delivers the following functions:
    - iMaster NCE-Campus can display the network-wide terminal types and operating systems, for example, dumb terminals including printers, IP cameras, smart all-in-one card, and access control system. iMaster NCE-Campus can also collect statistics and display traffic by terminal type.
    - Administrators do not need to manually configure different services and policies for different types of dumb terminals such as IP phones, printers, and IP cameras on the campus network. iMaster NCE-Campus can automatically identify these terminals and deliver the corresponding admission policies and service configurations to them.



- In this slide, "general scenarios" refer to authentication, non-authentication, and dynamic/static IP address assignment scenarios.
- In non-authentication scenarios, iMaster NCE-Campus can display information about wired terminals only after the ARP snooping function is enabled on access devices.

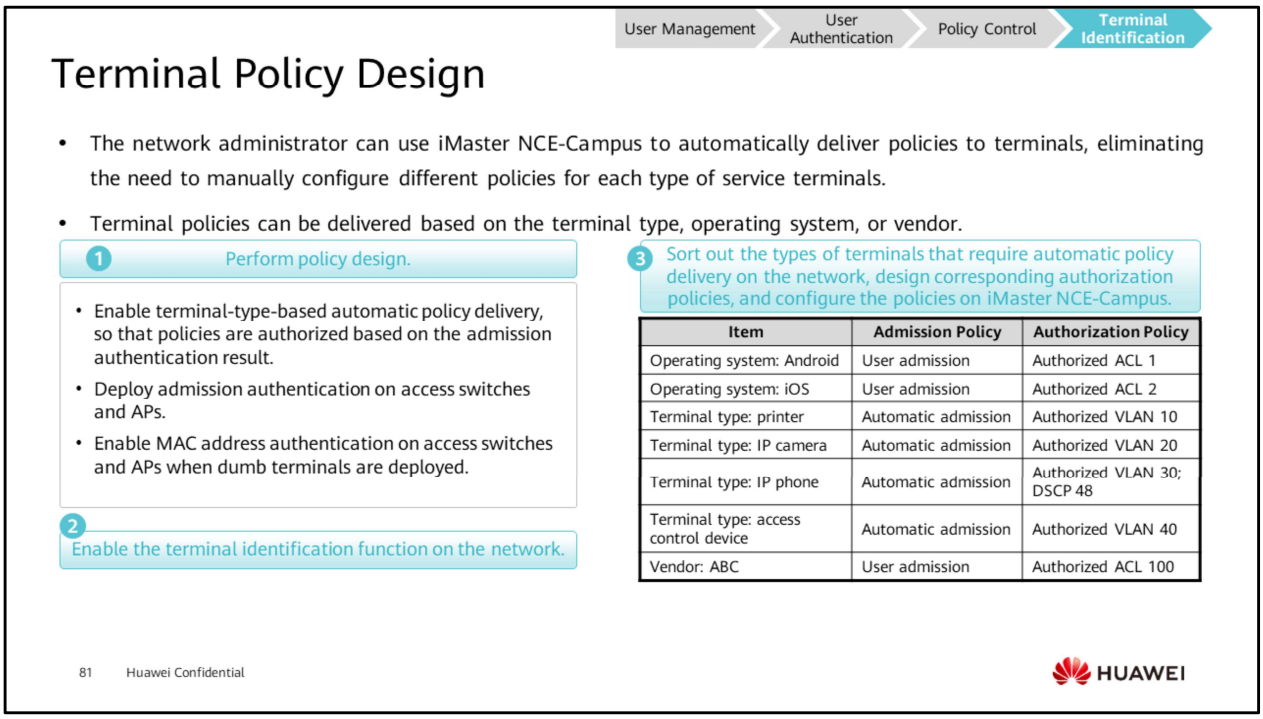
Terminal Identification Method Design (2)

3 Enable the terminal identification function.

Identification Method	Enabled On ...	Other Functions to Be Enabled
MAC OUI	Access switch and AP	-
HTTP UserAgent	Device enabled with Portal authentication	-
DHCP Option	Access switch and AP	DHCP snooping must be enabled on access switches. By default, DHCP snooping is already enabled on APs.
LLDP	Access switch and AP	-
mDNS	Access switch and AP	mDNS snooping must be enabled on access switches and APs.
SNMP Query	iMaster NCE-Campus	-
Nmap	iMaster NCE-Campus	-

- If the network administrator cannot determine which terminal identification method is to be used, the following passive fingerprint-based identification methods are recommended: MAC OUI, HTTP UserAgent, DHCP Option, LLDP, and mDNS.
- It is recommended that Nmap be disabled by default because this identification takes a long time. If the preceding passive fingerprint-based identification methods cannot meet requirements, enable Nmap.

HUAWEI



- It is recommended that admission and authorization policies be automatically delivered to dumb terminals (such as printers, IP phones, and IP cameras) based on terminal types. This helps implement automatic service provisioning and plug-and-play for dumb terminals.

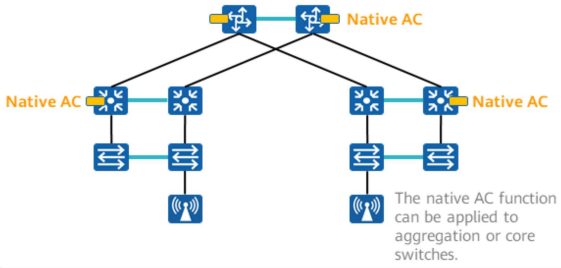
Contents

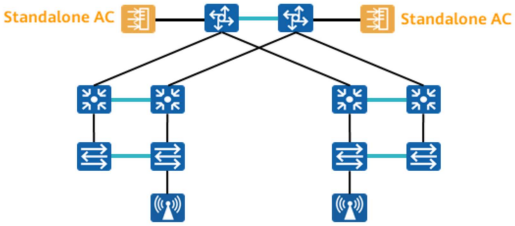
1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. Fabric Design
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. **WLAN Design**
7. Network Security Design
8. O&M Design

HUAWEI

WLAN Service Solution

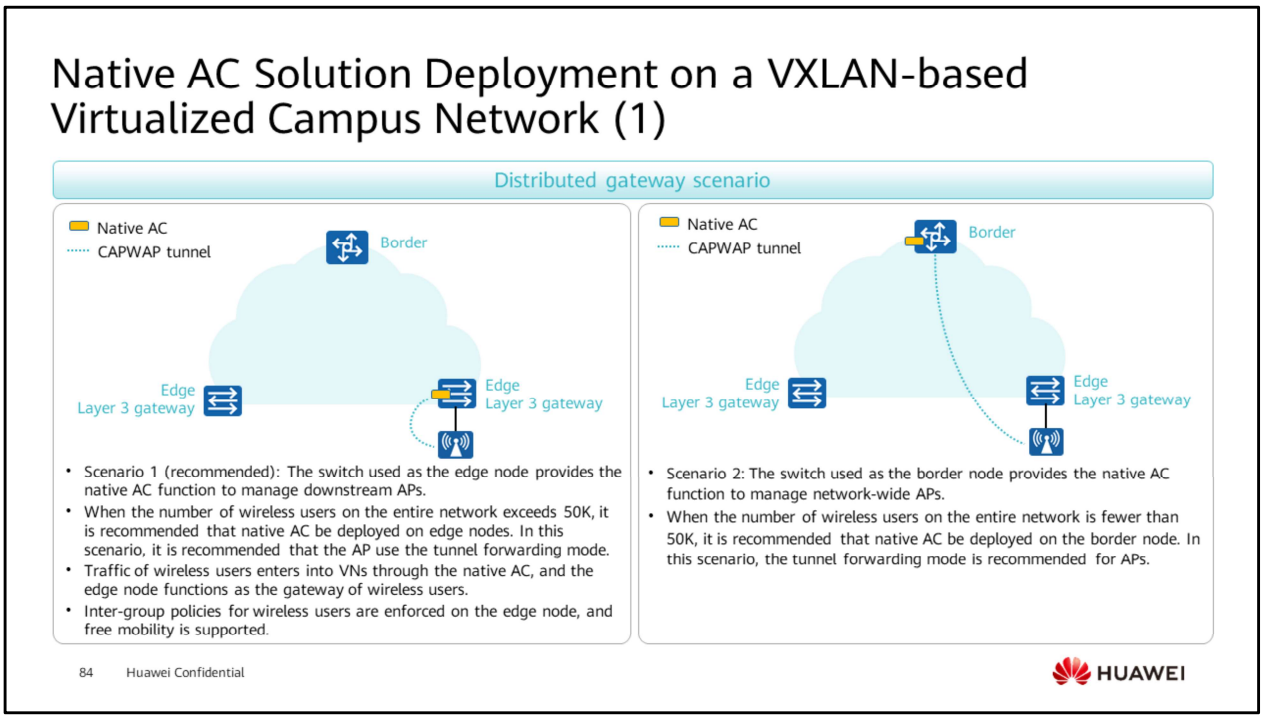
Native AC solution	Standalone AC solution
1. Switches provide the native AC function. 2. Free mobility is supported. Policies for both wired and wireless users are centrally enforced on the switches. 3. Unified O&M for both wired and wireless users.	1. Standalone AC (connected to the core switch in off-path mode). 2. The free mobility solution is supported for wireless users, but it must be used together with the IP-security group entry synchronization solution. 3. Separate O&M for wired and wireless users.





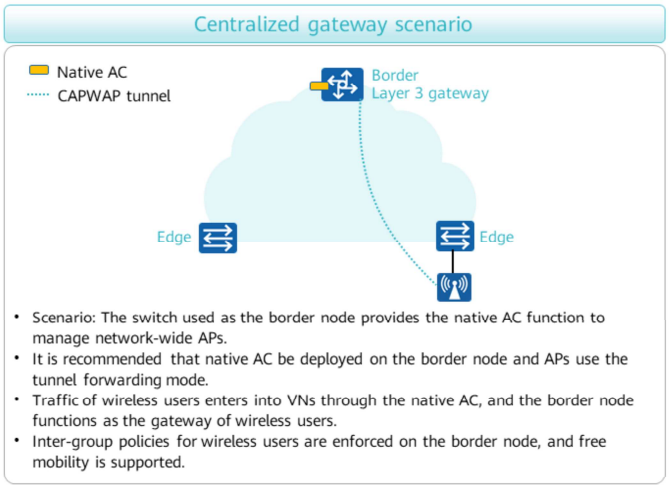
HUAWEI

- On a large- or medium-sized campus network, the WLAN typically uses the "AC + Fit AP" networking architecture. An AC can be deployed in in-path or off-path mode, depending on its location. A native AC (integrated on a switch) must be deployed in in-path mode, whereas a standalone AC can be deployed in either in-path or off-path mode (off-path mode is recommended).



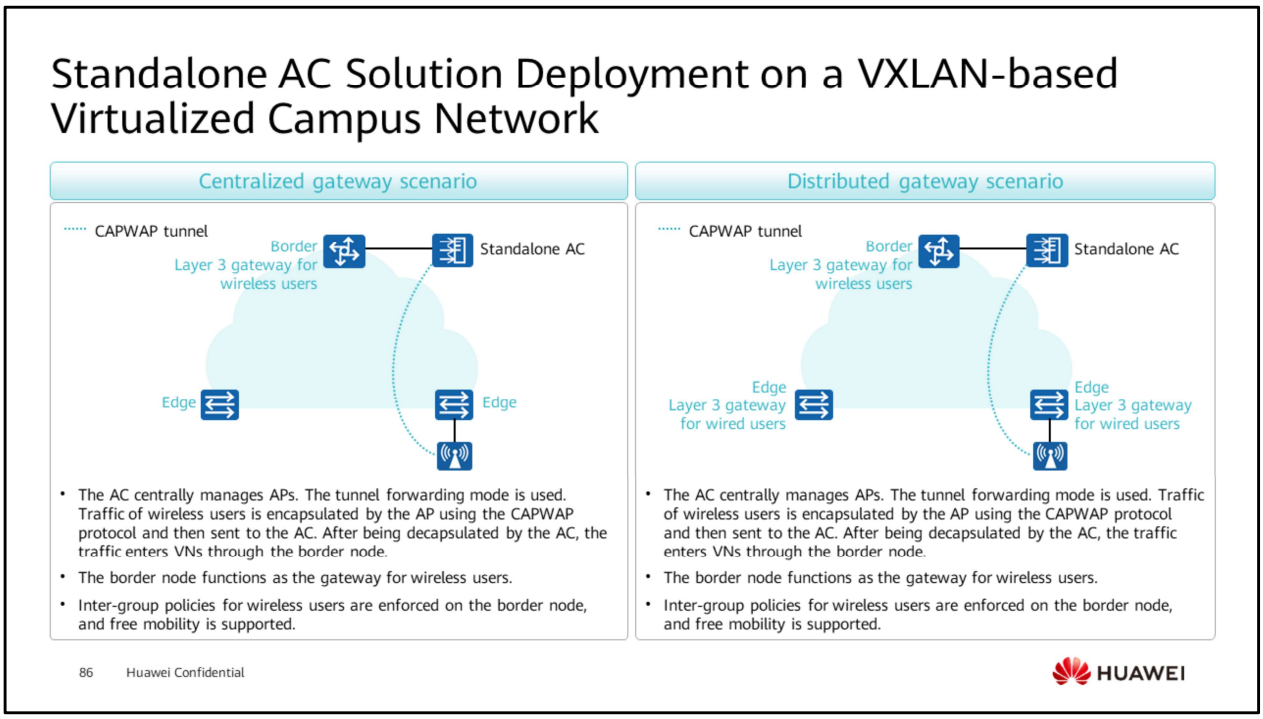
- Description of Scenario 1:
  - User gateway:
    - Edge node
  - Authentication point:
    - Wired users: edge node; Wireless users: edge node (native AC)
  - Forwarding mode:
    - Wired traffic: Traffic enters VNs through the edge node, and free mobility policies are enforced on the edge node.
    - Wireless traffic: Free mobility policies are enforced on the edge node. The tunnel forwarding mode is recommended. Traffic enters VNs through the native AC (edge node).
- Description of Scenario 2:
  - User gateway:
    - Wired users: edge node; Wireless users: border node (in tunnel forwarding mode)
  - Authentication point:
    - Wired users: edge node; Wireless users: border node (native AC)
  - Forwarding mode:
    - Wired traffic: Traffic enters VNs through the edge node, and free mobility policies are enforced on the edge node.
    - Wireless traffic: Free mobility policies are enforced on the border node. The tunnel forwarding mode is recommended. Traffic enters VNs through the border node (native AC).

Native AC Solution Deployment on a VXLAN-based Virtualized Campus Network (2)



HUAWEI

- Scenario description:
  - User gateway:
    - Border node
  - Authentication point:
    - Wired users: edge node; Wireless users: border node (native AC)
  - Forwarding mode:
    - Wired traffic: Traffic enters VNs through the edge node, and free mobility policies are enforced on the edge node.
    - Wireless traffic: Free mobility policies are enforced on the border node. The tunnel forwarding mode is recommended. Traffic enters VNs through the border node (native AC).



- Description of the centralized gateway scenario:
  - User gateway:
    - Border node
  - Authentication point:
    - Wired users: edge node; Wireless users: standalone AC
  - Forwarding mode:
    - Wired traffic: Traffic enters VNs through the edge node, and free mobility policies are enforced on the edge node.
    - Wireless traffic: Free mobility policies are enforced on the border node (the border node needs to subscribe to IP-security group entries). The tunnel forwarding mode is recommended. Traffic enters VNs through the border node (traffic is forwarded from the standalone AC to the border node and then enters VNs).
- Description of the distributed gateway scenario:
  - User gateway:
    - Wired users: edge node; Wireless users: border node (in tunnel forwarding mode)
  - Authentication point:
    - Wired users: edge node; Wireless users: AC
  - Forwarding mode:
    - Wired traffic: Traffic enters VNs through the edge node, and free mobility policies are enforced on the edge node.
    - Wireless traffic: The tunnel forwarding mode is recommended. Traffic enters VNs through the border node, and the border node enforces free mobility policies (the border node needs to subscribe to IP-security group entries).

Deployment of WLAN Deployment Solutions on a VXLAN-based Virtualized Campus Network

- The following table lists the recommended WLAN deployment solutions for centralized and distributed gateway networking.

Networking Type	AC	Gateway for Wireless Terminals	AP Forwarding Mode	Wireless Traffic Virtualization Location	Number of Terminals Supported	Applicable Scenario
Centralized gateway	Native AC on the border node	Border node	Tunnel forwarding	Border node	≤ 50,000	Recommended when the number of terminals does not exceed 50,000 and centralized forwarding is implemented for wireless terminals.
	Standalone AC connected to the border node in off-path mode	Border node	Tunnel forwarding	Border node	≤ 50,000	Recommended when the wired network is reconstructed first and centralized forwarding is implemented for wireless terminals.
Distributed gateway	Native AC on the edge node	Edge node	Tunnel forwarding	Edge node	50,000~100,000	Recommended when the number of terminals exceeds 50,000 and centralized forwarding is implemented for wireless terminals.

- In new deployment scenarios or wired and wireless network reconstruction scenarios, the native AC is recommended and tunnel forwarding mode is preferentially selected for APs.

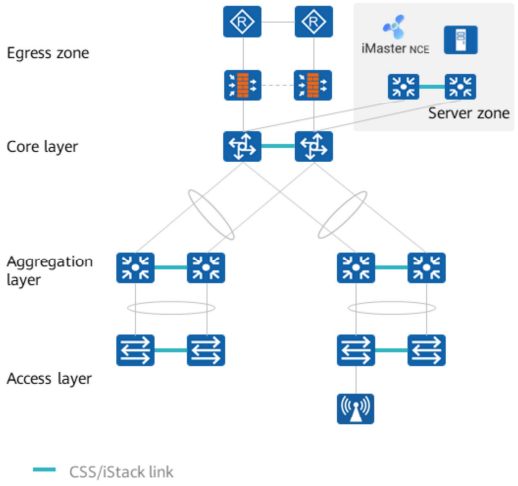
HUAWEI

Contents

1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. Fabric Design
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. WLAN Design
7. **Network Security Design**
8. O&M Design

HUAWEI

Overall Security Design



Egress security

- Campus network security design principles
  - There is no absolutely secure network but only more comprehensive protection measures.
  - Security is changing dynamically. Different networks have different security requirements. Therefore, on-demand network security design is the best choice.
  - Security is a system issue. Security design is required in all aspects of the network. In addition to egress security, campus networks also need to consider intranet security and security compliance.

Intranet security

- Campus network security system
  - Egress security
  - Intranet security: wired network security and wireless network security

HUAWEI

- On a traditional campus network, the intranet is often considered secure, and threats mainly come from external networks. Firewalls are often deployed to ensure security on campus borders. As security challenges increase, border defense at the egress cannot meet requirements. The security model needs to shift from passive into proactive and the security scope needs to be expanded from external networks to the intranet to solve security problems from the source (terminals), improving enterprise-wide information security level.

Egress Network Security Design

- Egress security requirement analysis
  - For external networks: Deploy security policies and intrusion prevention to prevent unauthorized access and attacks.
  - For the intranet: Deploy antivirus and URL filtering to implement virus detection and URL access control.
- Egress security solution design
  - Dedicated security devices, such as firewalls and intrusion prevention systems, are recommended.
  - Routers with the preceding security functions can also be used.
- Security zone division
  - The Trust zone is a security zone with a high security level. It is typically used to define the zone where intranet terminal users are located.
  - The DMZ is a security zone with a medium security level. It is typically used to define the zone where the servers that need to provide services for external networks are located.
  - The Untrust zone is a security zone with a low security level. It is typically used to define insecure networks such as the Internet.

Function	Description
Intrusion prevention	Compares traffic against the intrusion prevention signature database to prevent application-layer attacks, such as buffer overflows, Trojan horses, backdoor attacks, and worms.
Antivirus	Inspects the files transmitted on the network for viruses to protect the intranet from data breaches and system crashes caused by viruses.
URL filtering	Permits or denies access to URLs to control the online behavior of users.

Access Zone	Access Source	Trustworthiness	Recommended Security Policy
Internet	External users	Untrusted	Intrusion detection, URL filtering, and antivirus
	Employees on the go	Medium	
WAN	Enterprise branches	Medium	URL filtering and antivirus
Intranet	Enterprise employees	High	URL filtering and antivirus
	Guests	Low	

HUAWEI

Wired Network Security Design

Security Risk	Protection Measure
Broadcast storm	Enable traffic suppression and storm control.
DHCP attack	Enable DHCP snooping and configure uplink interfaces as trusted interfaces.
Impersonating legitimate users	Enable IP source guard (IPSG) and dynamic ARP inspection (DAI).
Mutual access	Enable port isolation.
CPU attack	Enable CPU attack defense.
Denial of service (DoS) attack	Enable attack source tracing.
	Enable port attack defense.
	Enable user-level rate limiting.

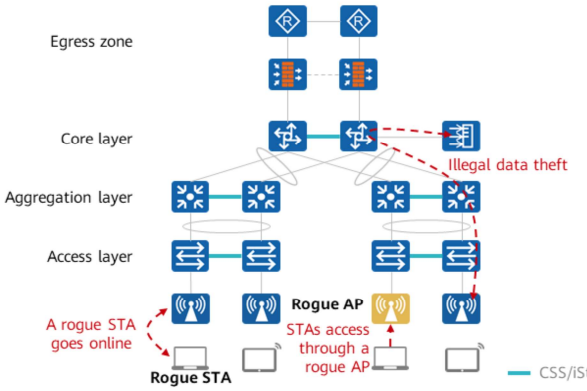
HUAWEI

- Security measures:
  - Enable traffic suppression and storm control.
    - Control broadcast, multicast, and unknown unicast packets to prevent broadcast storms. Traffic suppression limits the traffic using the configured threshold, and storm control blocks the traffic by shutting down interfaces.
  - Enable DHCP snooping and configure uplink interfaces as trusted interfaces.
    - DHCP snooping defends against bogus DHCP server attacks, DHCP server DoS attacks, bogus DHCP packet attacks, and other DHCP attacks. DHCP snooping allows administrators to configure trusted and untrusted interfaces, so DHCP clients can obtain IP addresses from authorized DHCP servers. A trusted interface forwards the DHCP packets it receives, whereas an untrusted interface discards the DHCP ACK packets and DHCP Offer packets received from a DHCP server.
    - An interface directly or indirectly connected to the DHCP server trusted by the administrator needs to be configured as a trusted interface, and other interfaces are configured as untrusted interfaces. This ensures that DHCP clients obtain IP addresses only from authorized DHCP servers and prevents bogus DHCP servers from assigning IP addresses to DHCP clients.
  - Enable IPSG and DAI.
    - IPSG prevents unauthorized hosts from using IP addresses of authorized hosts or specified IP addresses to access or attack the network.

        - You can configure DAI to defend against man-in-the-middle (MITM) attacks and prevent theft of authorized user information. When a device receives an ARP packet, it matches the source IP address, source MAC address, VLAN ID, and interface number of the ARP packet against binding entries. If a match is found, the device considers the ARP packet valid and allows it to pass through. Otherwise, the device discards the packet.
    - Enable port isolation.
        - You are advised to configure port isolation on the interfaces connecting an access switch to terminals. This configuration secures user communication and prevents invalid broadcast packets from affecting user services.
    - Enable CPU attack defense.
        - CPU attack defense limits the rate of packets sent to the CPU so that only a limited number of packets are sent to the CPU within a certain period of time. This ensures that the CPU can properly process services.
        - Control Plane Committed Access Rate (CPCAR) is the core of CPU attack defense. CPCAR limits the rate of protocol packets sent to the control plane to ensure security of the control plane.
    - Enable attack source tracing.
        - Attack source tracing defends against DoS attacks. A device enabled with attack source tracing analyzes packets sent to the CPU, collects statistics on the packets, and allows a user to set a packet rate threshold for the packets. Packets sent at a threshold-crossing rate are considered as attack packets. The device finds the source user address or source interface of the attacker by analyzing the attack packets and generates logs or alarms to alert a network administrator. The network administrator then takes measures to defend against the attack or configure the device to discard packets sent by the attack source.
    - Enable port attack defense.
        - Port attack defense is an anti-DoS attack method. It defends against attacks based on ports and prevents protocol packets on ports from occupying bandwidth and causing other packets to be discarded.
        - By default, port attack defense is enabled on the device for common user protocol packets, such as ARP, ICMP, DHCP, and IGMP packets. If a user attack occurs, the device restricts the attack impact within the port, reducing the impact on other ports.
    - Enable user-level rate limiting.
        - User-level rate limiting identifies users based on MAC addresses, and rate-limits specified protocol packets, such as ARP, ND, DHCP Request, DHCPv6 Request, IGMP, 802.1X, and HTTPS-SYN packets. If a user undergoes a DoS attack, other users are not affected.
        - Host CAR is the core of user-level rate limiting. By default, user-level rate limiting is enabled.

Overview of Wireless Network Security Design

- On a WLAN, service data is transmitted through radio signals. Such open channels are vulnerable to service data interception and tampering during transmission, such as rogue STAs, spoofing APs, and DoS attacks of malicious terminals.
- In addition to the security risks of wired networks, wireless networks also face security risks of wireless air interfaces. Therefore, the security risks of wireless networks mainly involve the security of wireless air interfaces.



- WLAN security design involves the following:
  - **Air interface security:** identifies and defends against attacks such as rogue APs, rogue STAs, unauthorized ad-hoc networks, and DoS attacks.
  - **STA access security:** ensures the validity and security of STAs' access to the WLAN.
  - **Service security:** protects service data of authorized users from being intercepted by unauthorized users during transmission.

Wireless Network Security Design

- Wireless air interface security design:

Security Risk	Defense Measure
Wireless attack and spoofing	WIDS
	WIPS
	Attack detection
Unencrypted management frames	Protected management frame (PMF)

- STA access security design:

Security Risk	Defense Measure
Air interface encryption and access authentication	WPA/WPA2
	WAPI
Terminal access security	STA blacklist and whitelist
Brute-force cracking	Brute-force attack defense and dynamic blacklist

- Wireless service security design:
  - The wired network between APs and ACs also faces common security threats, for example, interception, tampering, and spoofing, on IP networks.
  - To improve data transmission security, CAPWAP tunnels between APs and ACs support Datagram Transport Layer Security (DTLS) encryption, including DTLS encryption of CAPWAP tunnels for management packets, DTLS encryption of CAPWAP tunnels for service data packets, encryption of sensitive information, and integrity check.

HUAWEI

- Wireless air interface security design:
  - The Wireless Intrusion Detection System (WIDS) can detect rogue and interfering APs, bridges, and STAs, as well as ad-hoc devices.
  - The Wireless Intrusion Prevention System (WIPS) can disconnect authorized users from rogue APs, disconnect rogue and interfering devices from the WLAN, and contain such devices.
  - Attack detection: The WIDS and WIPS can also detect attacks such as flood attacks, weak initialization vector (IV) attacks, spoofing attacks, brute force WPA/WPA2/WAPI pre-shared key (PSK) cracking, and brute force WEP shared key cracking in a timely manner. The two systems then record logs, statistics, and alarms to notify network administrators of such attacks. The WLAN device adds devices that initiate flood attacks and brute force key cracking attacks to the dynamic blacklist and rejects packets from such devices within the aging time of the dynamic blacklist.
  - PMF: Management frames on a WLAN are not encrypted, which may cause security problems. The PMF standard is released by the Wi-Fi Alliance based on IEEE 802.11w. It aims to apply security measures defined in WPA2 to unicast and multicast management action frames to improve network trustworthiness.

- STA access security design:
  - Four WLAN security policies are available: Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA), WPA2, and WLAN Authentication and Privacy Infrastructure (WAPI). Each security policy has a series of security mechanisms, including link authentication used to establish a wireless link, user authentication used when users attempt to connect to a wireless network, and data encryption used during service transmission.
    - WEP uses a shared key to authenticate users and encrypt service packets. Since the shared key is easy to decipher, the WEP security policy is not recommended due to its low security.
    - WLAN devices support the STA blacklist and whitelist function to filter access requests from STAs based on specified rules, allowing authorized STAs to access the WLAN and rejecting unauthorized STAs.
    - When a WLAN uses WPA/WPA2-PSK, WAPI-PSK, or WEP-Shared-Key as the security policy, attackers can use the brute force method to decrypt the password. Defense against brute-force key cracking can prolong the time needed to decrypt passwords.
  - The STA access security design is to properly plan and design the access policies of STAs, securing and facilitating STA access.
  - The following is an example of the STA access authentication policy. The following access authentication modes are recommended on enterprise wireless networks:
    - Enterprise employees: WPA/WPA2-802.1X authentication
    - Guests: WPA/WPA2-PPSK or Portal authentication
    - Dumb terminals: MAC address authentication
  - In addition, if users do not need to communicate with each other, it is recommended that user isolation be configured.
- Wireless service security design:
  - Sensitive information encryption: When sensitive information is transmitted between an AP and an AC, the information can be encrypted to ensure security. Sensitive information includes the FTP user name, FTP password, AP login user name, AP login password, and service configuration key. During inter-AC roaming, the sensitive information encryption function can also be configured to protect data transmitted between ACs.
  - Integrity check: When CAPWAP packets are transmitted between an AP and an AC, these packets may be forged or tampered with or attackers may construct malformed packets to launch attacks. Integrity check can protect CAPWAP packets between the AP and AC. If the AP and AC are both located on the intranet, this function does not need to be enabled. It is recommended that this function be enabled when the AP is connected to the AC across the Internet or the ACs are distributed across the Internet.

Contents

1. CloudCampus Solution and Virtualized Campus Network Overview
2. Underlay Network Design
3. Fabric Design
4. Overlay Network Design
5. Admission Control and Free Mobility Design
6. WLAN Design
7. Network Security Design
8. **O&M Design**

HUAWEI

Challenges Facing Campus Network O&M



Precise detection

Traditional O&M is based on SNMP and therefore data collection takes minutes. Once a fault occurs, data generated at the time when the fault occurred cannot be obtained in real time. In addition, no convenient backtracking method is available.

Key requirements: Telemetry-based precise detection



Experience perception

In traditional O&M systems, only device metrics are monitored. User experience may be poor even if the metrics are normal. In addition, traditional O&M systems do not provide correlative analysis for clients and networks.

Key requirements: client & network profiling, experience perception



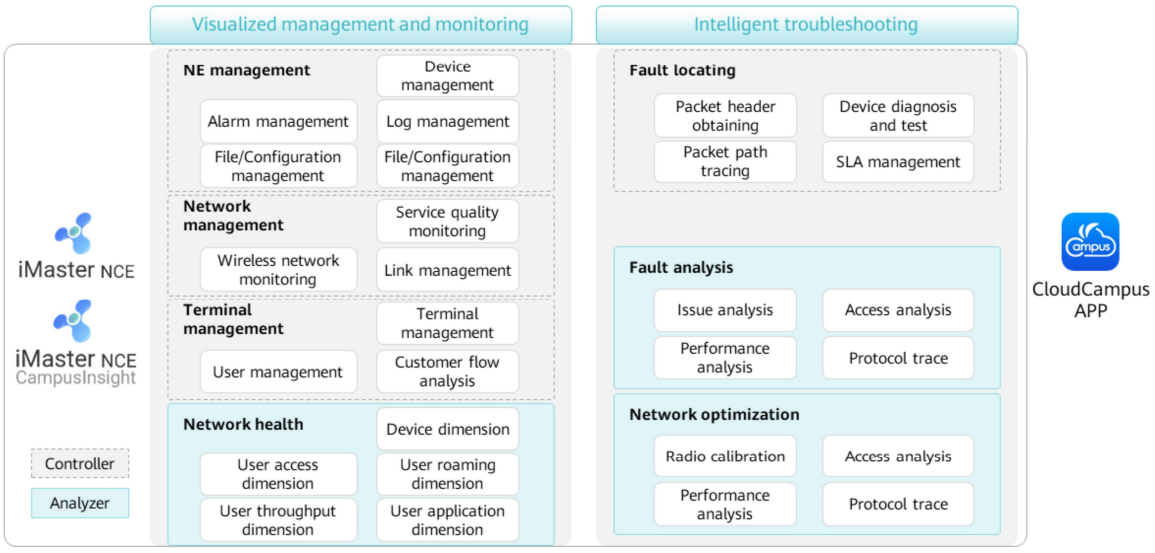
Fault identification

During traditional O&M, network faults can be detected only after receiving clients' complaints. As a result, faults cannot be effectively and proactively identified and analyzed.

Key requirements: intelligent fault identification

HUAWEI

Intelligent O&M Panorama



HUAWEI

Telemetry > Visualization > Analysis > Optimization

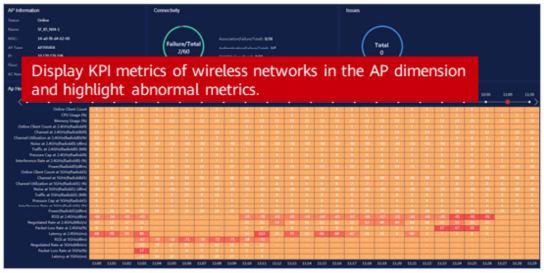
Using Telemetry to Monitor Metrics of Wireless Networks

Uses telemetry to monitor KPI metrics of wireless networks, displays the wireless network quality from the AP, radio, and client dimensions, and proactively identifies air interface performance issues, such as weak-signal coverage, high interference, and high channel utilization.

Real-time data display	Automatic issue identification
Display key metrics of wireless networks and highlight abnormal metrics.	Automatically identify air interface performance issues based on AI algorithms, correlative analysis, and exception modes.

Telemetry-based metrics collection for wireless networks

Measured Object	Measurement Metrics	Supported Device Type	Default Collection Interval
AP	CPU usage, memory usage, and number of online clients	AP	1 minute
Radio	Number of online clients, channel utilization, noise, traffic, backpressure queue, interference rate, and power	AP	1 minute
Client	RSSI, negotiated rate, packet loss rate, and latency	AP	1 minute





HUAWEI

- RSSI: Received Signal Strength Indicator

Using Telemetry to Monitor Metrics of Wired Networks

Use Telemetry to collect metrics of devices and interfaces on wired networks, and analyze the metrics, so as to proactively monitor and predict network faults.

Real-time display of KPI metrics

Display key metrics on wired networks in real time, including top N devices and historical trend.

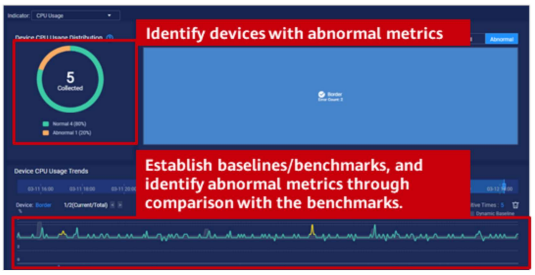
Exception detection based on dynamic baselines

Use AI algorithms to predict the baselines of KPI metrics, such as the CPU/memory usage. Identify deteriorating network metrics before service interruptions through comparison with dynamic baselines.

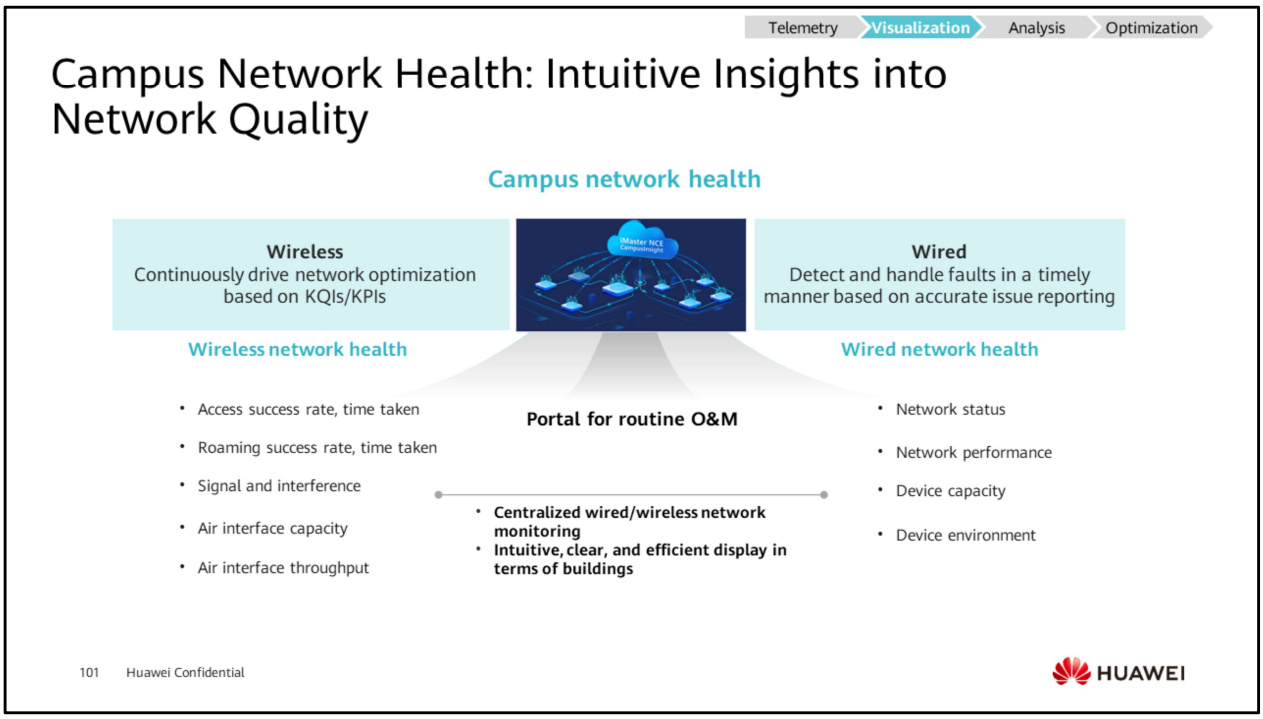
Telemetry-based metric collection for wired networks

Measured Object	Measurement Metrics	Supported Device Type	Default Collection Interval
Device/ Card	CPU usage	Switch and AC	1 minute
	Memory usage	Switch and AC	1 minute
Interface	Number of received/sent broadcast packets, number of received/sent multicast packets, number of received/sent unicast packets, number of received/sent packets that are discarded, and number of received/sent error packets	Switch and AC	1 minute

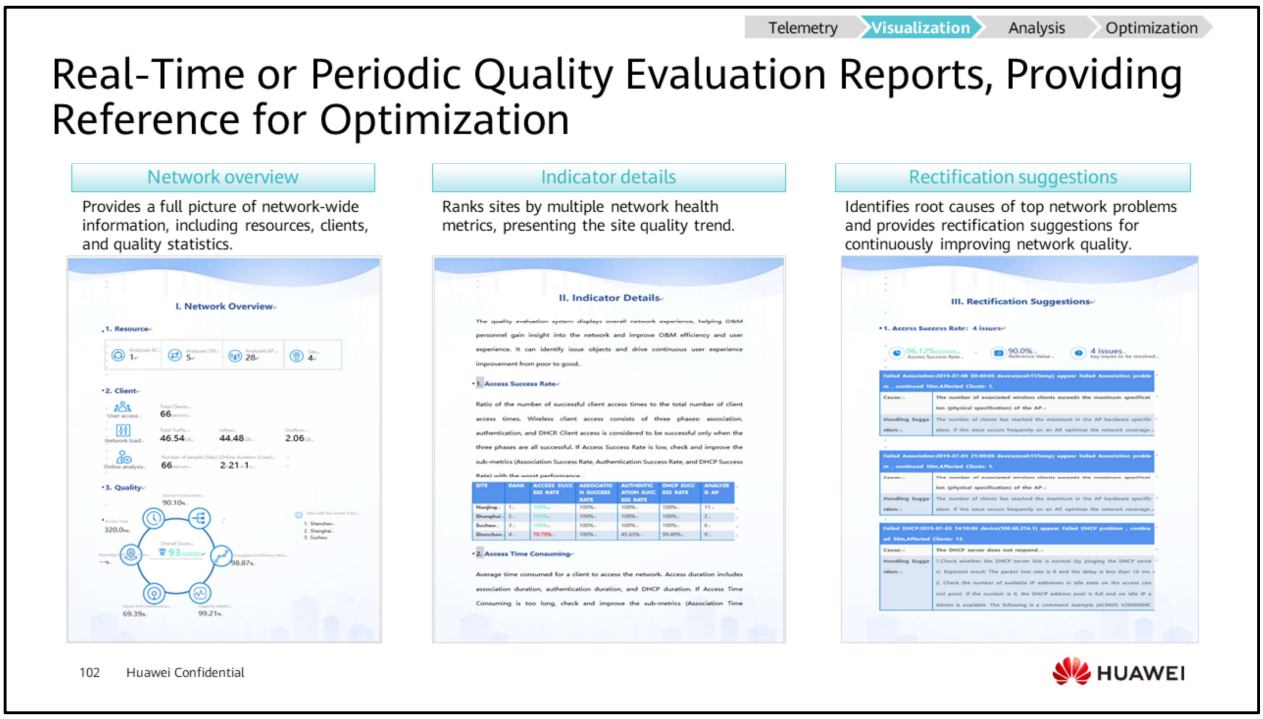




HUAWEI



- Provides insights into campus network health, including both wired and wireless network health, based on visualized topology monitoring, implementing intelligent, simplified campus network O&M.
- Network health topology as a uniform portal for rapidly handling network and device problems in a local building, simplifying network O&M.
- KQI: Key Quality Indicator



- iMaster NCE-CampusInsight provides professional evaluation report services. Specifically, it generates network quality evaluation reports including the network overview, indicator details, and rectification suggestions in real time or periodically, delivering measurable network experience.

Individual and Group Issue Analysis

During campus network O&M, administrators encounter the following types of issues:

1. **Individual issues:** for example, access failures caused by incorrect client configurations.
2. **Group issues:** for example, group authentication failures caused by authentication server faults and group weak-signal coverage issues caused by insufficient AP coverage.

Solution to individual issues



1. Per-metric: second-level data collection and exception detection.
2. Per-client: protocol-level access process analysis, client journey playback and experience analysis, and audio/video application quality perception.

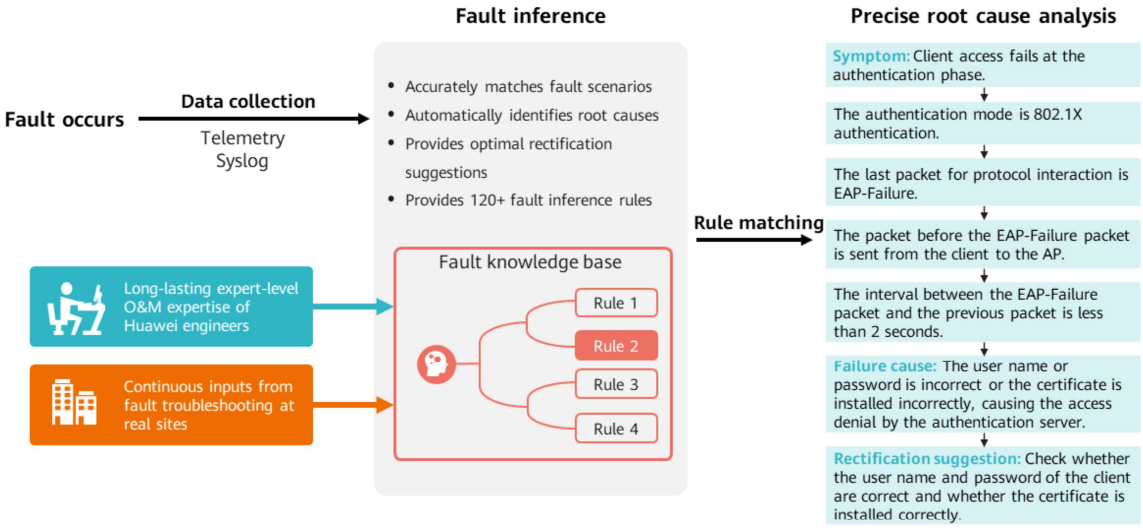
Solution to group issues



Identification of four types of issues: intelligently identifies connectivity, air interface performance, roaming, and device issues.

HUAWEI

Fault Inference



HUAWEI

AI-powered Intelligent Radio Calibration Improves Network-wide Performance

Scenario 1: manual calibration

About 20% of customers choose to manually plan channels.

However, they may face the following challenges:

- Channels planned in manual mode are not the optimal ones.
- Real-time network awareness is not supported.
  The network environment is complex and interference changes rapidly.

Real-time simulation feedback

Provide prediction and simulation tools based on real-time feedback of environment changes to drive network optimization.

Scenario 2: automatic calibration

About 80% of customers choose automatic calibration.

However, they may face the following challenges:

- Load balancing is not considered during load balancing calibration.
- Only the current network status can be detected.
  Unable to detect historical loads and interference.

Predictive Calibration

Provide the service weight balancing and optimization capability based on big data analytics and AI.

HUAWEI

Intelligent O&M Solution Deployment Design

- Huawei iMaster NCE-CampusInsight is totally different from the traditional monitoring mode that focuses on resource status. This platform applies AI to the O&M field and uses Telemetry technology to collect performance indicators and log data of network devices. By using big data, AI algorithms, and more advanced analytics technologies as well as leveraging scenario-specific continuous learning and expert experience, iMaster NCE-CampusInsight frees O&M personnel from nerve-wracking alarms and noises, making user experience visualized and O&M automated and intelligent.
- The intelligent O&M solution integrates iMaster NCE-CampusInsight, iMaster NCE-Campus, and devices. Currently, iMaster NCE-CampusInsight can manage and provide intelligent analysis for Huawei cloud switches and cloud APs.

Network bandwidth design	Deployment location design
Devices need to periodically report data to iMaster NCE-CampusInsight. Therefore, the campus network needs to reserve bandwidth for data reporting. The average bandwidth consumed by each device is 3 kbit/s.	iMaster NCE-CampusInsight and iMaster NCE-Campus can be deployed at different locations. They can collaborate with each other as long as network connectivity is available. To avoid the instability of the intermediate network, you are advised to deploy them in the same location, for example, in the same DC.

HUAWEI

Precautions for Intelligent O&M Design

Network deployment design

During network deployment, ensure that there are reachable routes between network devices and the analyzer iMaster NCE-CampusInsight so that the network devices can send KPI data and log information to iMaster NCE-CampusInsight.

During network deployment, ensure that devices are clock-synchronized with iMaster NCE-CampusInsight. You are advised to deploy an NTP server to synchronize the system clocks on the network.

Intelligent radio calibration function design

Intelligent radio calibration and traditional radio calibration cannot be both deployed on APs in the same region.

Protocol trace function design

To identify DHCP-related connectivity issues and implement protocol trace, you need to use the AC as the DHCP server or enable DHCP snooping on the AC.

Audio and video quality analysis function design

The audio and video quality analysis function requires devices to send related log information to iMaster NCE-CampusInsight. You are advised to configure the same log sending interval for switches and WLAN devices. The maximum difference between the log sending intervals cannot exceed 20s.

HUAWEI

Quiz

1. (Multiple-Answer Question) For a large-scale campus network with more than 50,000 terminals, the recommended fabric networking mode is ( ), the gateway is located on the ( ), and the ( ) functions as the edge node.
   A. Centralized gateway
   B. Distributed gateway
   C. Edge node
   D. Border node
   E. Access switch
   F. Aggregation switch

HUAWEI

- BCF

Summary

- To meet requirements of digital transformation, campus networks need to be reconstructed systematically. Based on the IDN concept, Huawei CloudCampus 3.0 Solution helps enterprises build intelligent, simplified, converged, open, and secure networks.
- This course focuses on large- and medium-sized campus networks, illustrates the solution architecture and technical solutions used for such networks, and provides suggestions on engineering design and O&M design. With this information, this course helps trainees understand Huawei CloudCampus 3.0 Solution technologies and design the optimal campus network solution.

HUAWEI

Recommendations

- CloudCampus 3.0 Solution V100R020C00 Product Documentation
  - https://support.huawei.com/hedex/hdx.do?docid=DOC1100733970&lang=en
- Design Guide of Huawei CloudCampus 3.0 Solution for Large- and Medium-Sized Network Scenario (Virtualization)
  - https://e.huawei.com/en/material/networking/campusnetwork/1720e83a99cc47cc8bd5785741421fd0

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Small- and Medium-Sized Cloud-Managed Campus Network Design



HUAWEI

Foreword

- With the development of technologies and digitalization of industries, services in small- and medium-scale scenarios, such as chain stores and branch offices, are gradually connected to the enterprise intranet through cloud-managed networks. A report of IDC reveals the radical difference between traditional products and cloud-based products in terms of annual growth rate from 2014 to 2018. For example, the annual growth rate of traditional Wi-Fi was only 5%, while that of cloud Wi-Fi was 38%. And the annual growth rate of traditional switches and access routers was just 2%, whereas that of cloud switches and access routers was as high as 61%.
- The above-mentioned data shows that cloud-managed networks have become the trend in constructing small- and medium-sized campus networks. Cloud-based solutions help enterprises improve management efficiency and build networks that can quickly support the development of new services.
- Huawei's CloudCampus Solution uses cloud computing technology to implement automatic and centralized network management, and provides data collection and analysis capabilities that are unavailable on traditional networks, so as to achieve network (LAN/WLAN) as a service (NaaS).
- This document mainly describes the design and planning methods of small- and medium-sized campus networks in terms of the solution architecture, technical solution comparison, engineering design suggestions, and O&M.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the architecture of Huawei's CloudCampus Solution for small- and medium-sized campus networks.
  - Describe typical networking schemes for small- and medium-sized campus networks.
  - Independently design the CloudCampus solution for small- and medium-sized campus networks based on user requirements, including networking design, physical network design, site deployment design, basic network service design, WLAN design, and admission control design.

HUAWEI

Contents

1. **Small- and Medium-Sized Campus Network Trends and Challenges**
2. Overview of Huawei's CloudCampus Cloud-Managed Network Solution for Small- and Medium-Sized Campus Networks
3. Design of Small- and Medium-Sized Campus Networks with Huawei's CloudCampus Cloud-Managed Network Solution

HUAWEI

Challenges Faced by Small- and Medium-Sized Campus Networks

Accelerated industry changes

Network cloudification brings rapid development in new ICT technologies, such as cloud computing, cloud security, big data, and IoT, leading to tremendous changes in all industries.

- Traditional retailers such as shopping malls and supermarkets offer free Wi-Fi as a way to attract and retain customers, and they also use wireless positioning and customer flow analytics to carry out precision marketing.
- In the education sector, e-classrooms are becoming more and more popular, and diversified multimedia teaching methods further stimulate students' interest in learning.
- Smart healthcare enables hospitals and other healthcare institutions to connect to each other through networks. This implements unified management and analysis of medical data and facilitates medical treatment.

As a pipe that carries upper-layer services and data, the network is becoming more and more complex with a growing number of nodes.

Problems in traditional deployment and management solutions

Low deployment efficiency, which slows down service provisioning

- Site survey, planning, deployment, software commissioning, configuration, and optimization must be completed onsite by professional IT personnel.

Complex network management and high OPEX

- Local professional O&M results in low O&M efficiency and high labor costs. The network management system (NMS), policy control server, charging system, and data analysis platform are deployed independently, causing high management and maintenance costs.

Poor network openness

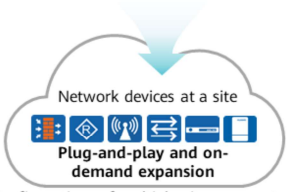
- The open data provided by multiple management systems on the traditional network needs to be further integrated. In addition, due to the incompatibility of API, the network and applications are connected at a slower speed than application development.

HUAWEI

Service Requirement Analysis for Small- and Medium-Sized Campus Networks

Plug-and-play network devices improve deployment efficiency

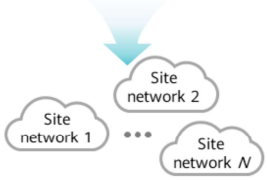
Unified management and centralized configuration



- Configurations of multiple sites are centrally delivered, reducing onsite configuration and commissioning workload and improving deployment efficiency.
- Network devices are plug-and-play and able to be expanded on demand, requiring low cost for upgrades.

Cloud-based centralized O&M simplifies multi-site O&M

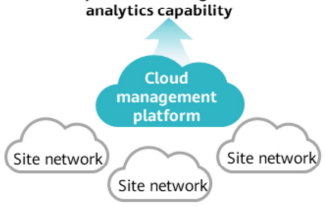
Centralized cloud management of multiple branches and remote automatic O&M



- Geographically dispersed campus branch networks are centrally managed on the cloud through the Internet.
- Troubleshooting and monitoring tools as well as many other automation tools are integrated for remote automatic O&M.

Open APIs accelerate business application integration

Openness and big data analytics capability



- With open APIs and big data analytics capabilities, the cloud management platform can interconnect with multiple management systems to achieve unified network management.
- It is also able to provide diversified value-added applications to lead enterprises into digital transformation.

HUAWEI

Small- and Medium-Sized Campus Network Trends

Network cloudification



- Due to evolution to the cloud architecture, enterprises can focus more on their mission-critical services, without paying too much attention to IT architecture construction.
- To support service cloudification, enterprises need to create a ubiquitous, intelligent, controllable, and on-demand network.
- The network needs to become more a service than a solution.

Cloud security



- Cloud security become more important than ever.
- Facing cloudification, enterprises are vulnerable to attacks that are fundamentally different from those on traditional networks when providing various services.
- Security has shifted from passive defense to proactive defense.
- Detection and response have become as important as defense.

IoT



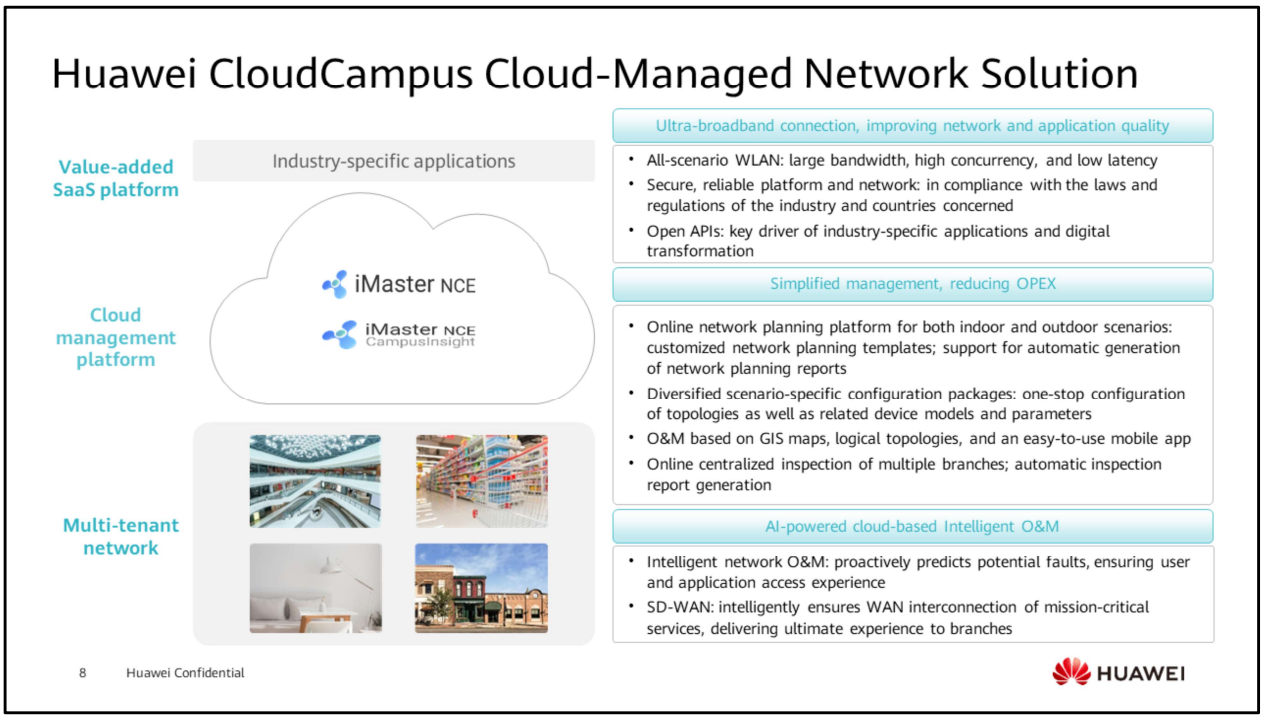
- The rise of IoT leads to a huge increase in the number and types of terminals accessing the network, and these terminals generate a large amount of data.
- Diversified IoT sensing networks need to be smoothly connected to the existing campus network.
- The types of terminals connected to the campus network are more complex than ever. As a result, the campus network becomes a converged network that accommodates multiple types of terminals and media.

HUAWEI

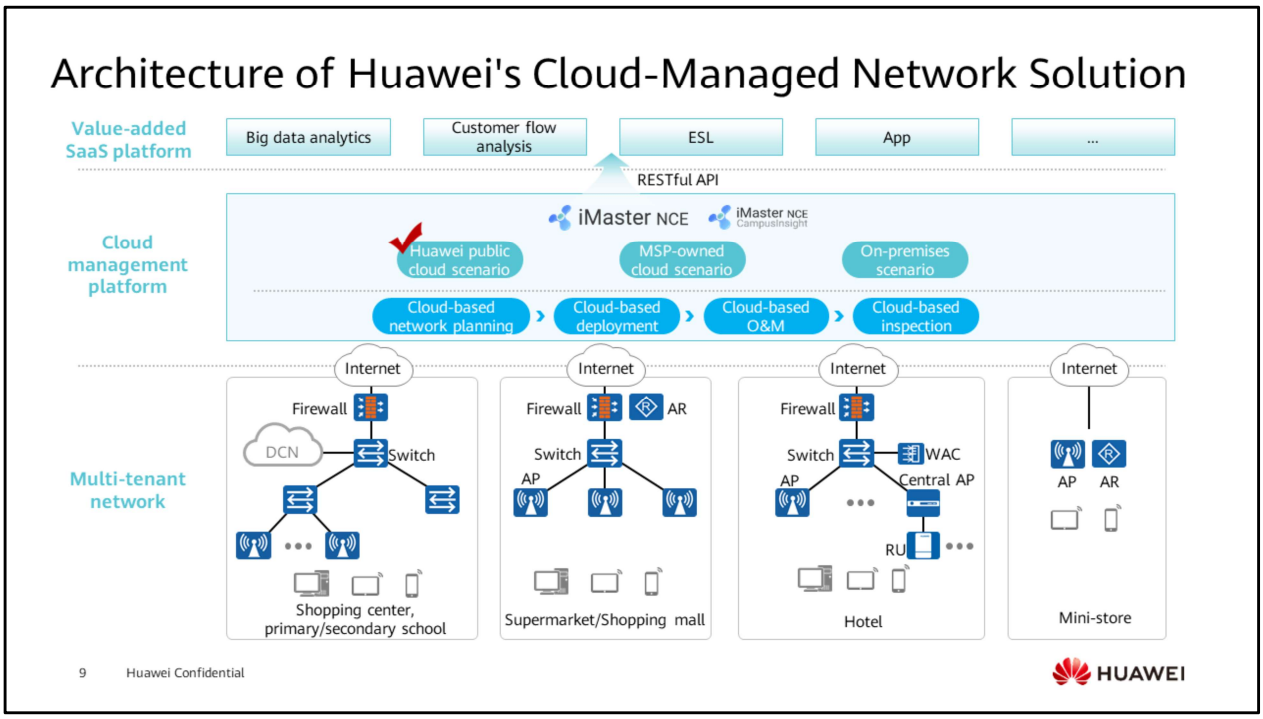
Contents

1. Small- and Medium-Sized Campus Network Trends and Challenges
2. **Overview of Huawei's CloudCampus Cloud-Managed Network Solution for Small- and Medium-Sized Campus Networks**
3. Design of Small- and Medium-Sized Campus Networks with Huawei's CloudCampus Cloud-Managed Network Solution

HUAWEI

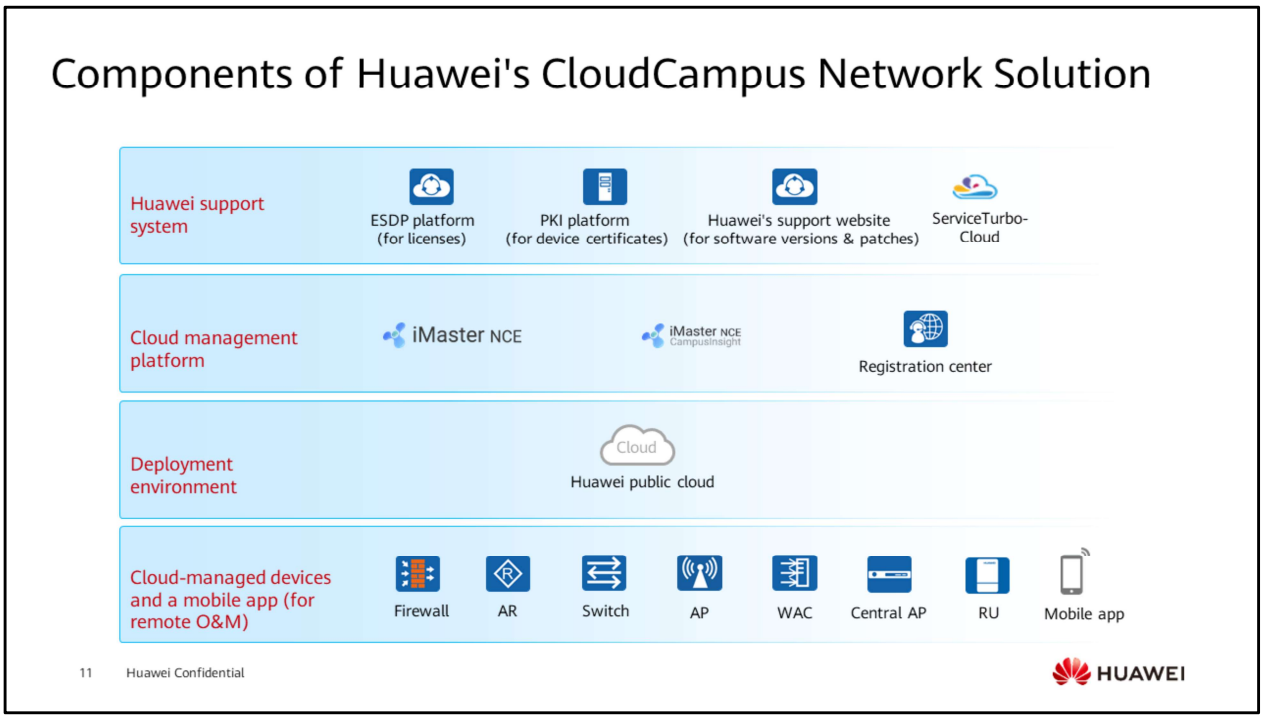


- In this slide, CloudCampus Network refers to Huawei's CloudCampus Cloud-Managed Network Solution for Small- and Medium-Sized Campus Networks.



- Huawei's CloudCampus Solution applies to three deployment scenarios: Huawei public cloud, MSP-owned cloud, and on-premises. The on-premises scenario applies to large- and medium-sized campus networks. The Huawei public cloud and MSP-owned cloud scenarios apply to small- and medium-sized campus networks. This document focuses on the Huawei public cloud scenario, which is a cloud management scenario.
- Huawei's CloudCampus Solution for small- and medium-sized campus networks uses cloud computing technology to implement automatic and centralized network management, and provides data collection and analysis capabilities that are unavailable on traditional networks, so as to achieve network (LAN/WLAN) as a service (NaaS).
- There are three layers in the architecture of Huawei CloudCampus Solution for small- and medium-sized campus networks: multi-tenant network, cloud management platform, and value-added SaaS platform.
- Multi-tenant network: It consists of over one hundred network devices, including APs, switches, firewalls, and ARs, and is deployed at the customer side to provide user access.

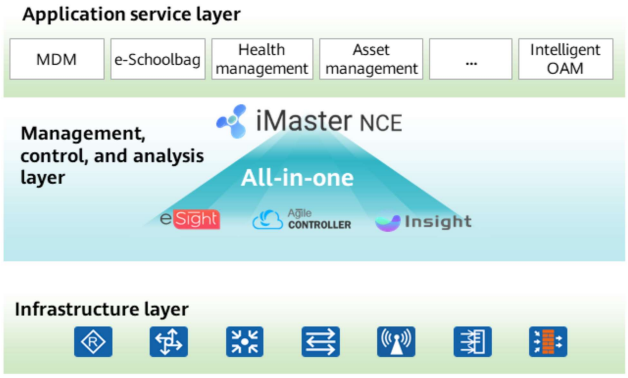
- Cloud management platform: iMaster NCE-Campus — an SDN controller — is the core component of the CloudCampus Solution. It is also a cloud-based network management, O&M, and control system. In addition to offering basic management and configuration for cloud-based devices, remote O&M and monitoring, and user admission control, iMaster NCE-Campus can implement various value-added services based on the big data platform. iMaster NCE-CampusInsight is an intelligent network analysis engine and provides intelligent O&M services for user networks. It integrates AI to the O&M. Based on data such as device performance indicators and terminal logs, iMaster NCE-CampusInsight digitalizes user experience on the network through big data analytics, AI algorithms, and more advanced analysis technologies, helping customers detect network problems in a timely manner and improve user experience.
- Value-added SaaS platform: iMaster NCE-Campus provides open interfaces to interconnect with other service systems (such as the big data platform) to offer tenants a variety of value-added application services, such as customer flow analysis, business portal push, electronic shelf label (ESL), asset management, and medical IoT.



- ServiceTurbo-Cloud (https://serviceturbo-cloud.huawei.com/) is an enterprise tool service cloud platform. It provides one-stop tool services for Huawei, partners, and customer service engineers in government and enterprise service scenarios.

Solution Component: iMaster NCE-Campus

- iMaster NCE-Campus is a main configuration and management platform in the CloudCampus Solution. It is a main portal for CloudCampus service configuration, O&M, and monitoring.

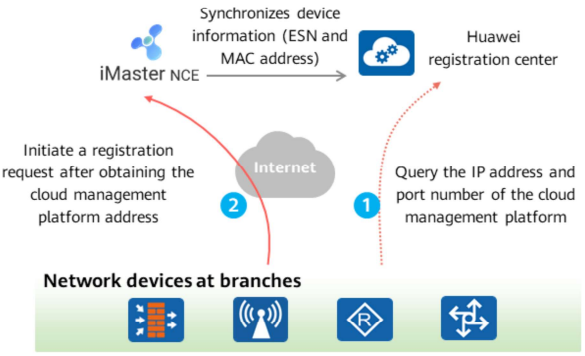


Automated + Intelligent	• SDN-based automatic service configuration/deployment • AI-powered intelligent analysis/prediction/troubleshooting
Management + Control + Analysis	• Unified data base • Centralized detection/locating/processing
Plan + Construct + Maintain + Optimize	• Full-lifecycle management • Simulation/Verification/Monitoring/Optimization

iMaster NCE-Campus, an autonomous driving campus network management and control system

HUAWEI

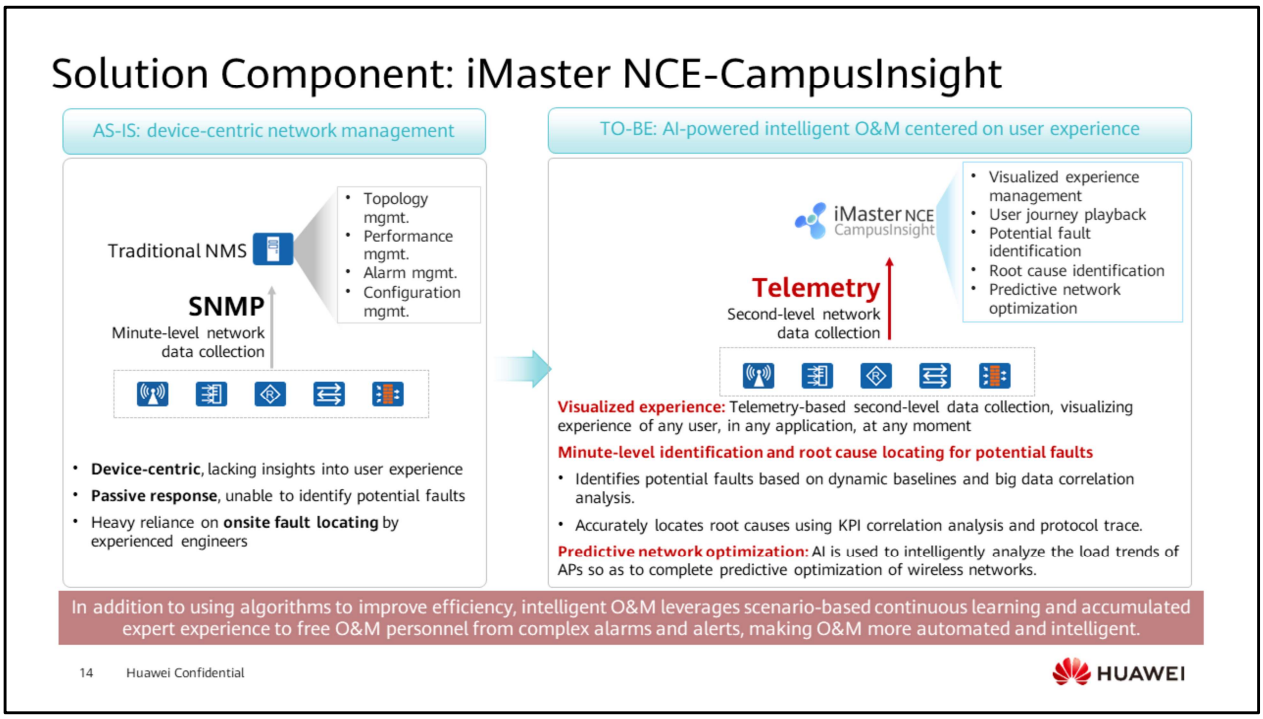
Solution Component: Registration Center



Work out of the box	Devices are already preset with the registration center domain (register.naas.huawei.com) before factory delivery. No additional configuration is required.
Secure transmission	Bidirectional authentication and encryption are performed for query channels based on HTTP/2, ensuring that user device information is not disclosed.
Wide coverage	Multiple registration centers are deployed around the world.
Quick response	The smart DNS returns the address of the registration center server nearest to devices, ensuring fast query.

At branches, devices can automatically go online on the cloud management platform without additional configuration.

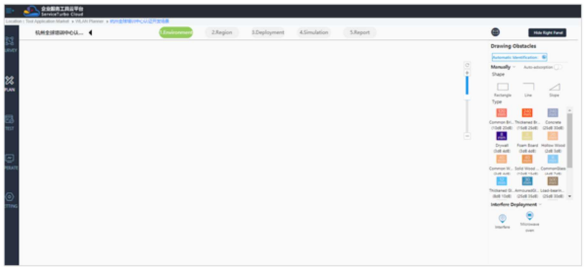
HUAWEI



- AS-IS: the current status
- TO-BE: future ideal status

Solution Component: WLAN Planner

Efficient support for WLAN network planning



Installation-free

Cloud-based tool, without software installation

All scenarios supported

Indoor, outdoor, agile distributed, and high-density scenarios

High efficiency

Higher simulation efficiency than the standalone edition

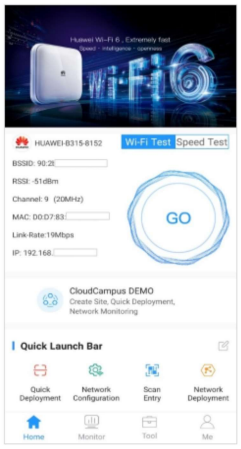
High quality

Field-proven in 10,000+ WLAN projects

The WLAN Planner is an efficient WLAN network planning tool. It enables signal simulation, helping determine AP deployment locations and signal coverage results.

HUAWEI

Solution Component: CloudCampus APP



Mobile APP that covers full lifecycle management for campus networks

Site survey	Connects to the cloud-based WLAN Planner to record photos and texts based on drawings.
Network planning	Connects to the cloud-based WLAN Planner to display network planning results, heat maps, and AP attributes anytime anywhere.
Deployment	Deployment by scanning barcode
Acceptance	One-click test, single service test, project test Deployment position acceptance based on the cloud-based network planning project
O&M	Mobile O&M, device and application monitoring AP access through Bluetooth/management VAP, AP offline diagnosis

E2E simplicity from planning and deployment, all the way to O&M

HUAWEI

Solution Component: Hardware Products

CloudEngine S series switches



High-performance campus switches

NetEngine ARs



Campus egress routers
(integrating routing, switching, Wi-Fi, 5G, and security functions)

AirEngine Wi-Fi 6 APs



Campus APs
(e.g., indoor settled APs, indoor wall plate APs, agile distributed APs, and outdoor APs)

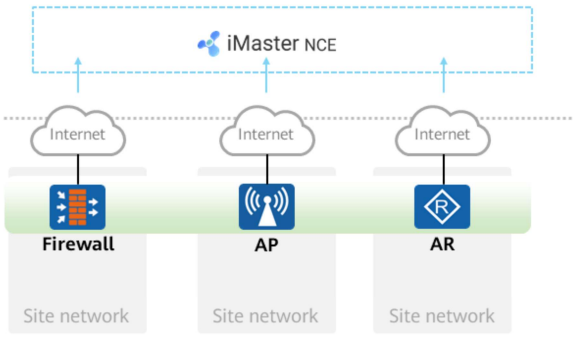
HiSecEngine AI firewalls



Firewalls
(providing comprehensive network security protection capabilities)

HUAWEI

Solution Technology: Plug-and-Play of Devices



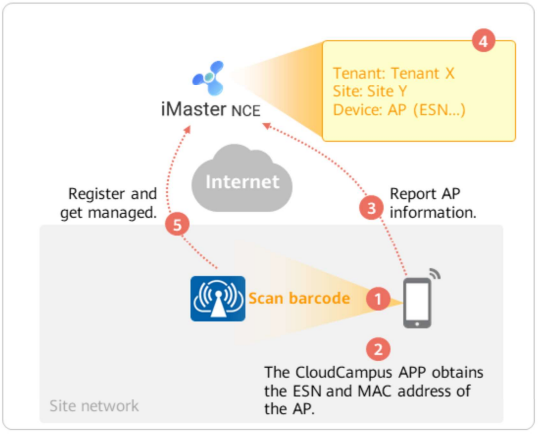
Simplified network deployment

- The time and labor costs for initial device installation and configuration as well as upgrade are reduced.
- After devices on the campus are powered on and connected to the Internet, they can obtain the IP address of the controller iMaster NCE and register with it through multiple methods. All subsequent operations can be performed on the controller, without the need of onsite visits.
- Services are deployed on the iMaster NCE in advance, greatly shortening the deployment time.
- The configuration error rate is reduced due to GUI-based operations.

HUAWEI

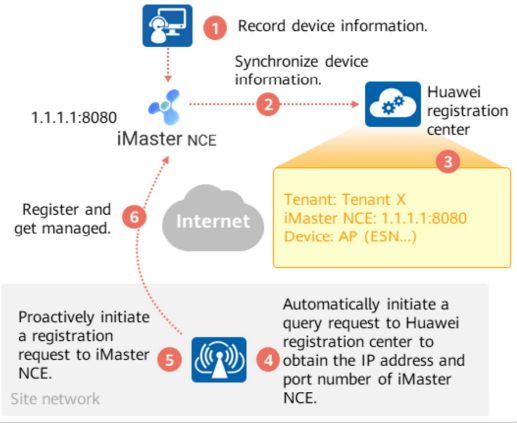
Plug-and-Play of Devices (1)

Through CloudCampus APP (barcode scanning)



Devices supported: AP

Through registration center

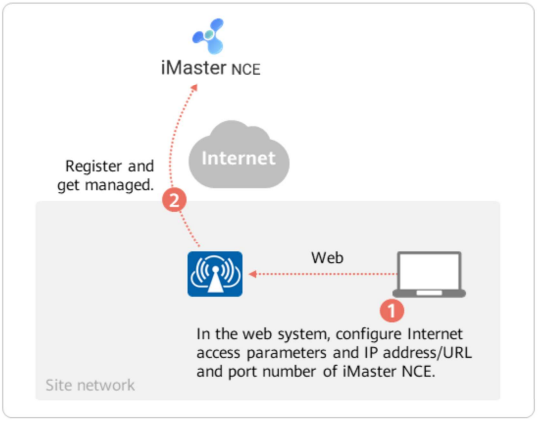


Devices supported: AR, firewall, switch, AP

HUAWEI

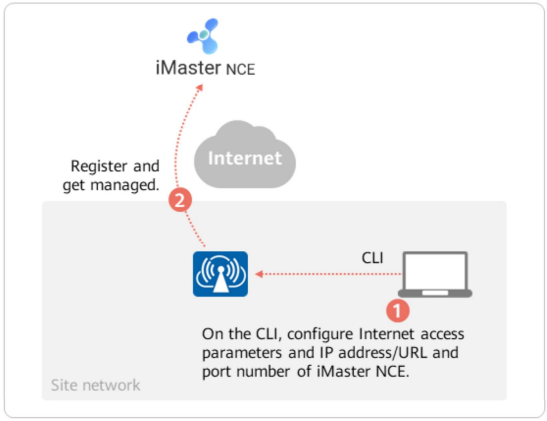
Plug-and-Play of Devices (2)

Through web system



Devices supported: AR, firewall, switch, AP

Through CLI

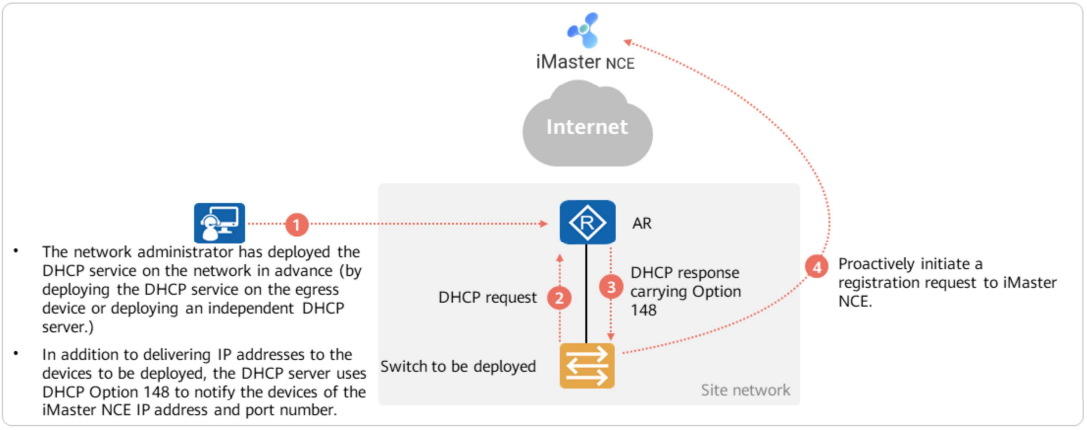


Devices supported: AR, firewall, switch, AP

HUAWEI

Plug-and-Play of Devices (3)

Through DHCP Option 148



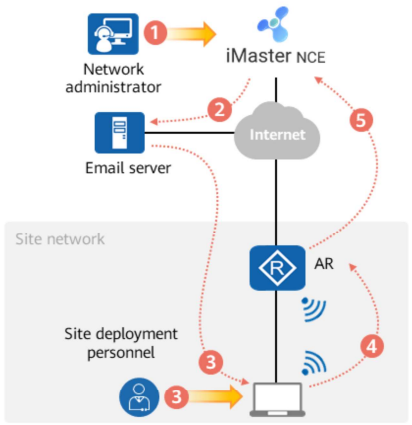
- The network administrator has deployed the DHCP service on the network in advance (by deploying the DHCP service on the egress device or deploying an independent DHCP server.)
- In addition to delivering IP addresses to the devices to be deployed, the DHCP server uses DHCP Option 148 to notify the devices of the iMaster NCE IP address and port number.

Devices supported: AR, switch, AP

HUAWEI

Plug-and-Play of Devices (4)

Email-based deployment

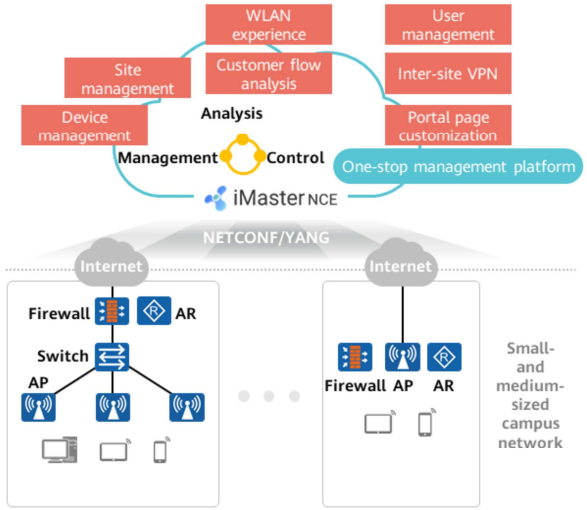


1. The network administrator configures ZTP and selects the email-based deployment mode on iMaster NCE.
2. iMaster NCE packs the deployment configuration into a series of character strings and sends them to the email address specified by the administrator.
3. The site deployment personnel log in to the email box on a PC at the site and receive the deployment email.
4. During site deployment, the site deployment personnel connect the PC to the AR in wired or wireless mode and click the hyperlink in the email body. The PC then automatically logs in to the AR, parses the parameters in the hyperlink to obtain the deployment configuration of the AR, and writes the configuration into the AR.
5. The AR connects to the Internet and automatically sends a registration request to iMaster NCE.

Devices supported: AR

HUAWEI

Solution Technology: Cloud Management



Higher management and O&M efficiency

- After devices go online, iMaster NCE provisions all required configurations and services. iMaster NCE delivers configurations to devices at the branch sites through NETCONF, without the need of CLI-based operations. This simplifies network O&M.
- iMaster NCE performs intelligent analysis and display of the operating status of branch sites, presenting intuitive insights into site health.
- iMaster NCE provides operating status monitoring, intelligent analysis, and remote O&M, so that O&M personnel do not need to visit sites. This improves O&M efficiency and reduces O&M costs.

HUAWEI

Highlights of Huawei CloudCampus Solution for Small- and Medium-Sized Campus Networks

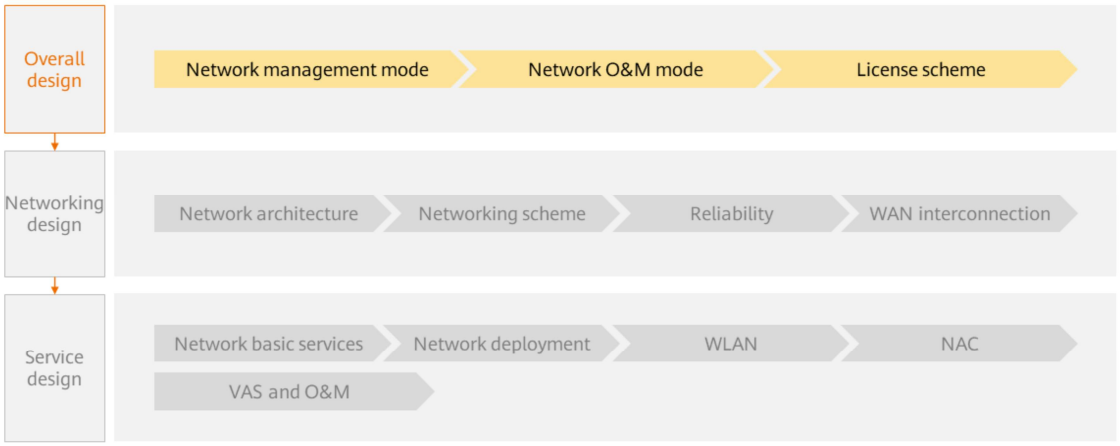
- **Automatic deployment:** Devices can be easily and quickly deployed.
- **Cloud-based network planning and mobile O&M:** WLAN design and device O&M are simplified.
- **Diversified product portfolios:** Huawei provides different product portfolios, including full series of network devices (switches, firewalls, ARs, and APs), meeting diversified network requirements of tenants.
- **Dual-working-mode:** All network devices used in this solution can work in either cloud-based or traditional management mode. Tenants can implement cloud management of traditional devices after devices are upgraded.
- **VASs**: Terminal behavior analysis is a value-added application of iMaster NCE-Campus. More VASs can be developed based on terminal behavior analysis.

Contents

1. Small- and Medium-Sized Campus Network Trends and Challenges
2. Overview of Huawei's CloudCampus Cloud-Managed Network Solution for Small- and Medium-Sized Campus Networks
3. **Design of Small- and Medium-Sized Campus Networks with Huawei's CloudCampus Cloud-Managed Network Solution**

HUAWEI

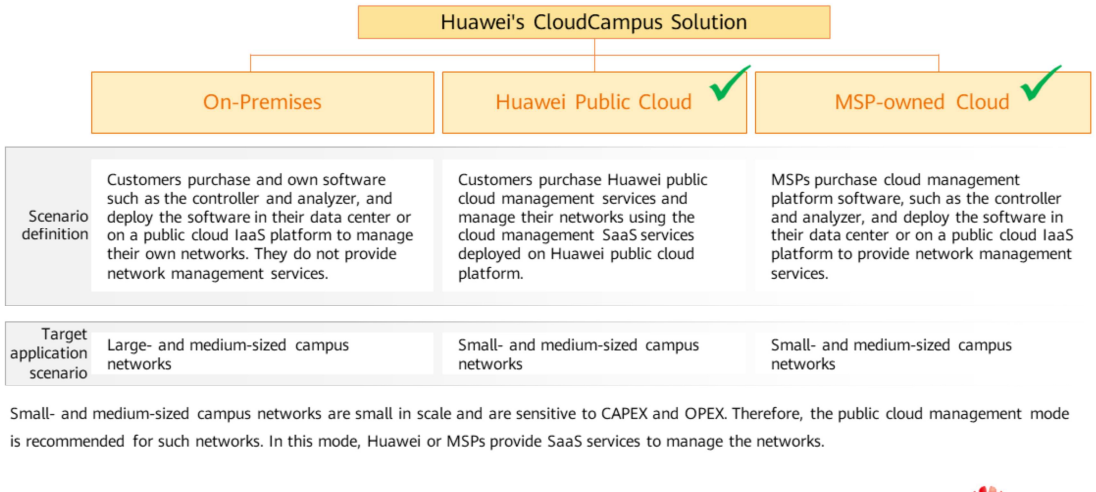
Overall Design Process



HUAWEI

Choose the Network Management Mode

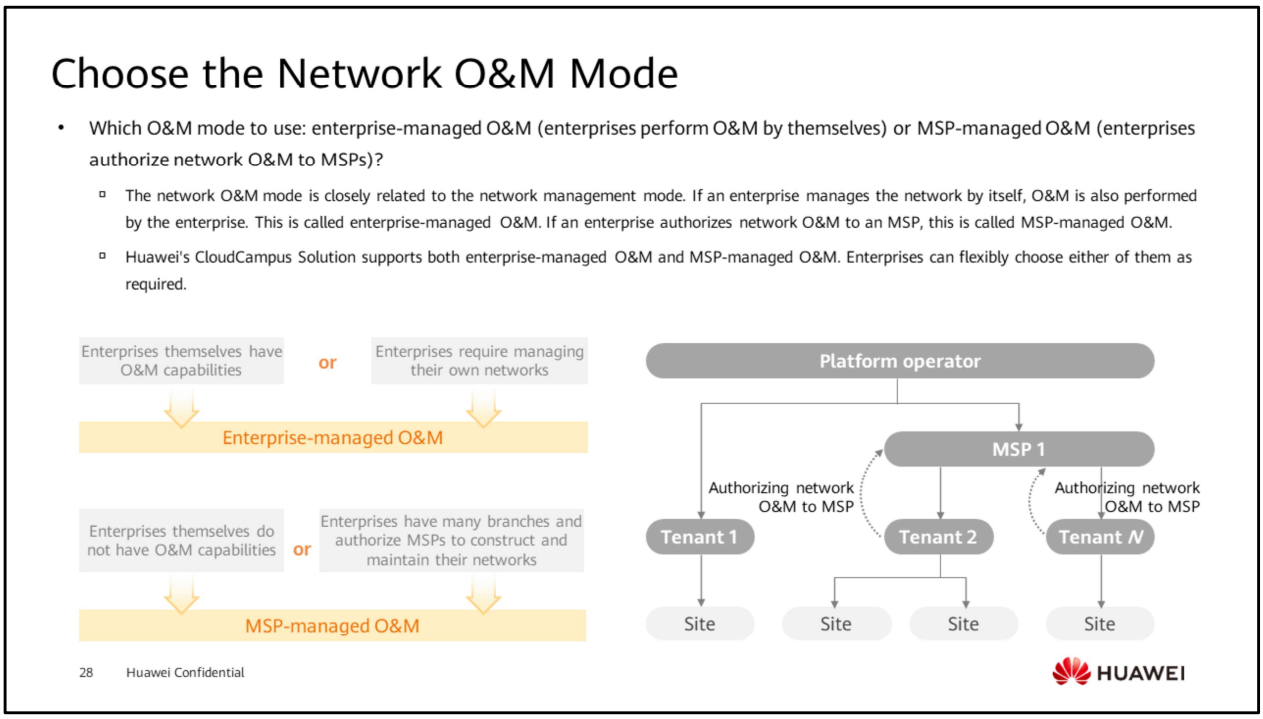
- Which management mode to use: cloud management or on-premises?



Small- and medium-sized campus networks are small in scale and are sensitive to CAPEX and OPEX. Therefore, the public cloud management mode is recommended for such networks. In this mode, Huawei or MSPs provide SaaS services to manage the networks.

HUAWEI

- The public cloud management mode can be Huawei public cloud management mode or MSP-owned cloud management mode. The two modes are essentially the same. The only difference lies in the operational entity and the provider that offers cloud management services.
- Unless otherwise specified, the Huawei public cloud management mode is used as an example in this document.



- With certain IT capabilities, a tenant administrator can deploy and maintain a campus network. This scenario is called tenant-managed construction and maintenance. The tenant administrator is the main implementer, and the MSP administrator only provides simple deployment assistance. The tenant administrator can apply to the MSP for the managed construction and maintenance services. After being authorized, the MSP constructs and maintains the campus network for the tenant. This scenario is called MSP-managed construction and maintenance, in which the MSP administrator is the main implementer.
- The following lists the differences between the deployment processes in the tenant-managed construction and maintenance and MSP-managed construction and maintenance scenarios:
- Tenant-managed construction and maintenance:
    - The tenant administrator installs cloud managed devices onsite and registers the cloud managed devices.
    - The tenant administrator logs in to iMaster NCE using their own account and deploys services.

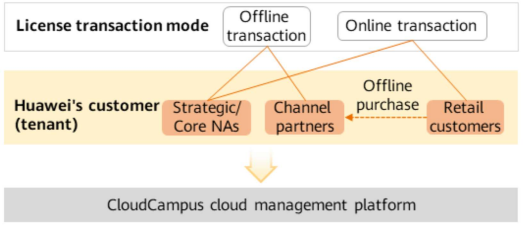
- MSP-managed construction and maintenance:
  - The MSP administrator helps the tenant install cloud managed devices onsite and register the cloud managed devices.
  - The tenant administrator logs in to iMaster NCE using their own account, chooses System > System Management > Tenant Information, enables Authorize MSP, and sets the authorization scope. Alternatively, when creating a tenant administrator, the MSP administrator can enable Authorize MSP. By default, the permission of the tenant administrator role is granted.
  - The MSP administrator logs in to iMaster NCE using their own account. In Tenant List on the home page, the MSP administrator selects a tenant who applies for MSP-managed construction and maintenance and has been authorized. The Authorization status column is displayed as Authorized.
  - Click the tenant name to access the page for MSP-managed construction and maintenance. The MSP administrator helps the tenant deploy services.

Choose the License Transaction/Purchase Mode

- Select the cloud management deployment scenario (Huawei public cloud or MSP-owned cloud), then select the license transaction or purchase mode accordingly.

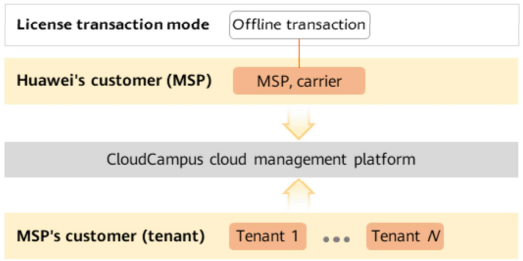
Scenario 1: Huawei public cloud

- Licenses control the available resources of tenants.
- Online transaction mode is supported. That is, licenses can be purchased online from HUAWEI CLOUD. License pooling is not supported in this mode.
- Offline transaction mode is also supported. License activation codes need to be loaded on the CloudCampus cloud management platform.



Scenario 2: MSP-owned cloud

- Licenses control the available resources of the CloudCampus cloud management platform, but not the available resources of tenants.
- Only offline transaction mode is supported. The license file needs to be loaded on the CloudCampus cloud management platform. License pooling and co-termination are supported in this mode.



HUAWEI

Choose the License Consumption and Termination Modes

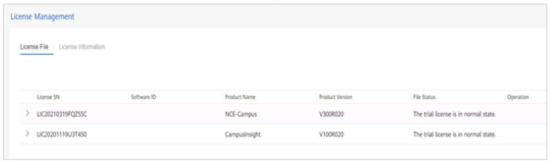
- In the Huawei public cloud scenario, both the co-termination and non-co-termination licensing models are supported.

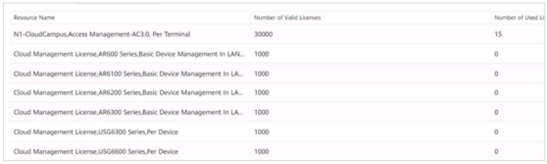
Co-termination licensing mode

- Licenses are pooled and shared by all devices, and co-terminate for all devices.
- Co-termination of licenses cannot be undone. Therefore, the license termination time must be properly designed.

Non-co-termination licensing mode

- Licenses are pooled and shared by device type. License resources are exclusive to devices of one type and not shared by devices of other types.
- Licenses for different types of devices terminate separately, while licenses for devices of the same type co-terminate.

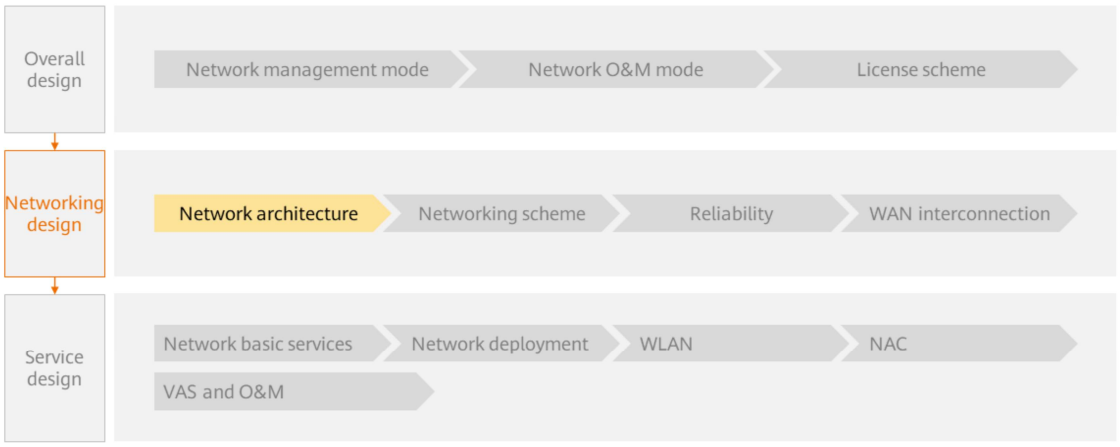




HUAWEI

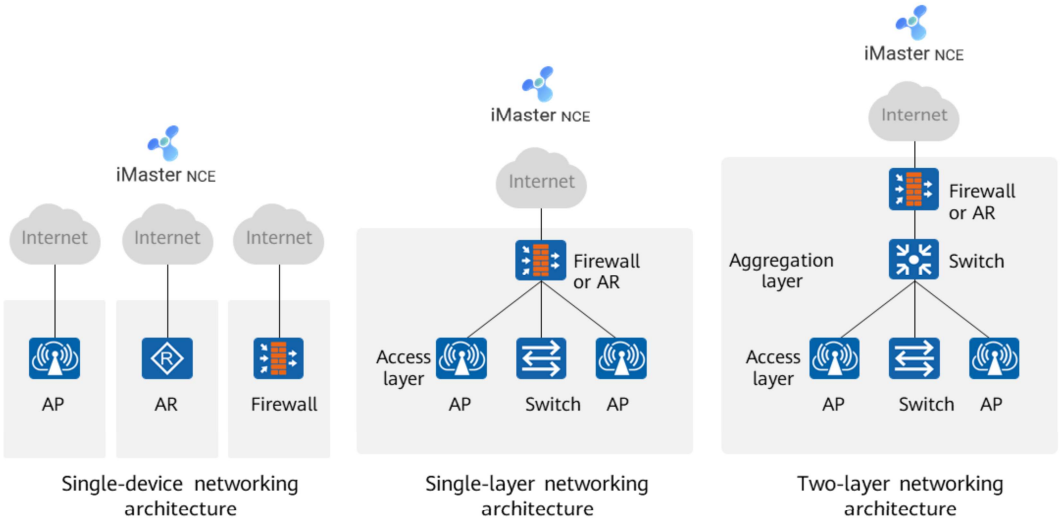
- For small- and medium-sized campus networks, the co-termination licensing model is recommended, facilitating device management and operations. If the license validity periods of devices of different types need to be precisely controlled, use the non-co-termination licensing model.

Overall Design Process



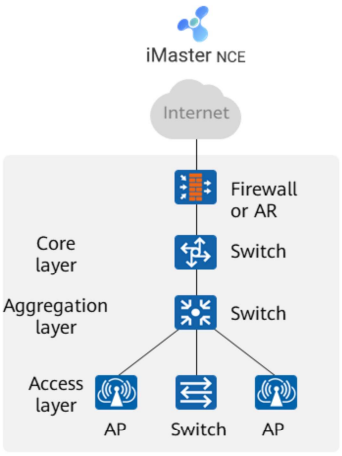
HUAWEI

Network Architecture Design: Intranet Architecture (1)

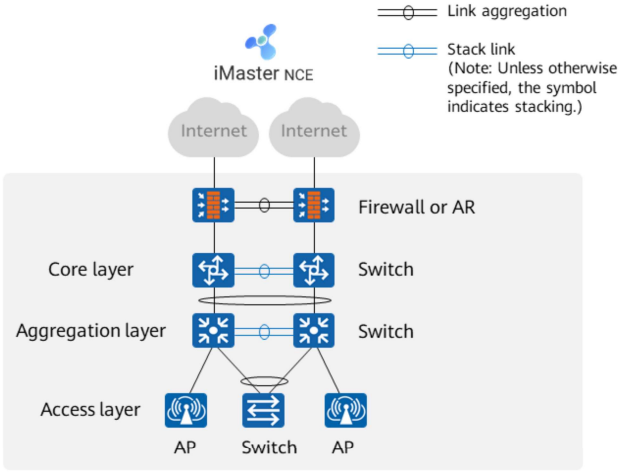


HUAWEI

Network Architecture Design: Intranet Architecture (2)

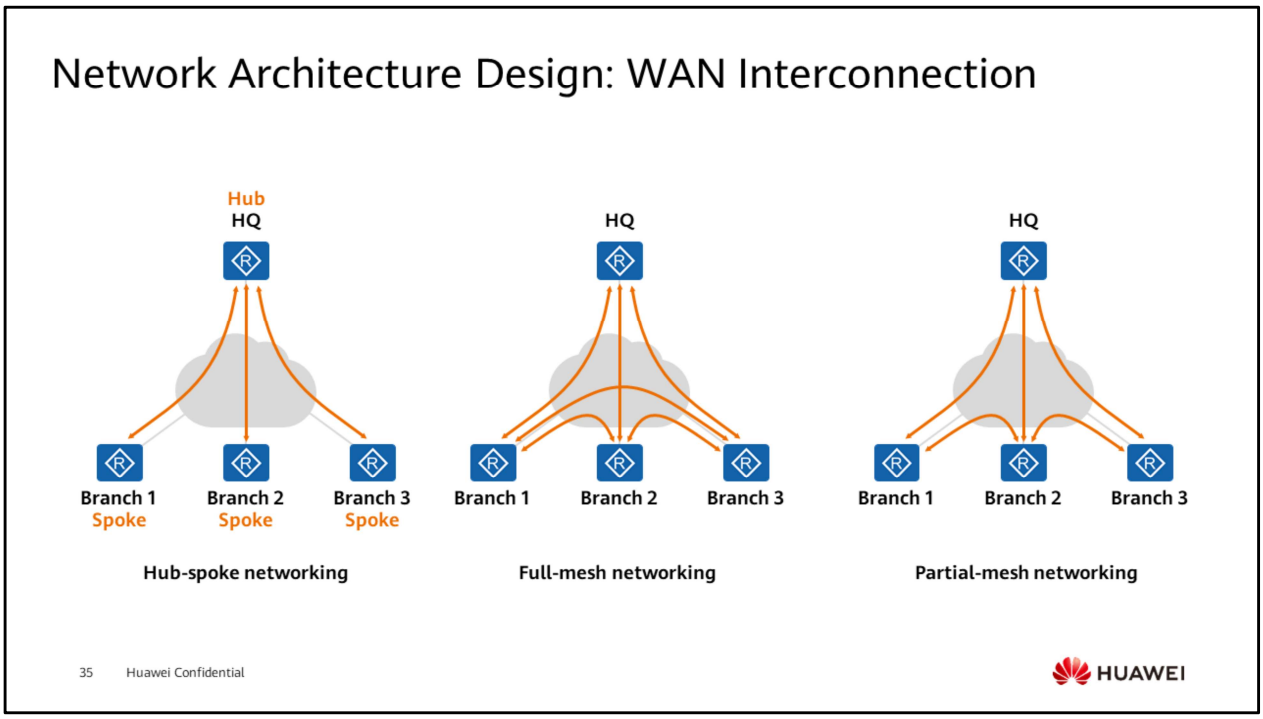


Three-layer networking architecture



Highly reliable networking architecture

HUAWEI



- Hub-spoke: Generally, the enterprise headquarters or DC functions as a hub site. Each branch site of an enterprise can communicate with the hub site and can communicate with other branch sites through the hub site. This model is applicable to scenarios where traffic between all branch sites of an enterprise must pass through the headquarters for centralized security monitoring.
- Full-mesh: All sites of an enterprise can communicate with each other. If traffic needs to be transmitted between the headquarters and branches or between branches, data is directly exchanged without traversing an intermediate node. This model is applicable to scenarios where all sites of an enterprise need to directly access each other. This model eliminates the delay caused by traffic transmission through the headquarters.
- Partial-mesh: Most sites of an enterprise can directly communicate with each other. However, the underlay WAN networks of a small number of sites cannot directly communicate with each other. For example, in the figure, branch 1 and branch 3 need to communicate with each other through the headquarters.

Network Architecture Design Considerations

1 Determine the intranet architecture

Determine the network architecture layers based on the number of access terminals or device capabilities. Determine the device model at each layer and the mapping between upper-layer and lower-layer devices.

- Single-device networking architecture
- Single-layer networking architecture
- Two-layer networking architecture
- Three-layer networking architecture

2 (Optional) Determine the WAN interconnection model

Determine the WAN interconnection model based on the communication requirements between sites.

- Hub-spoke networking
- Full-mesh networking
- Partial-mesh networking

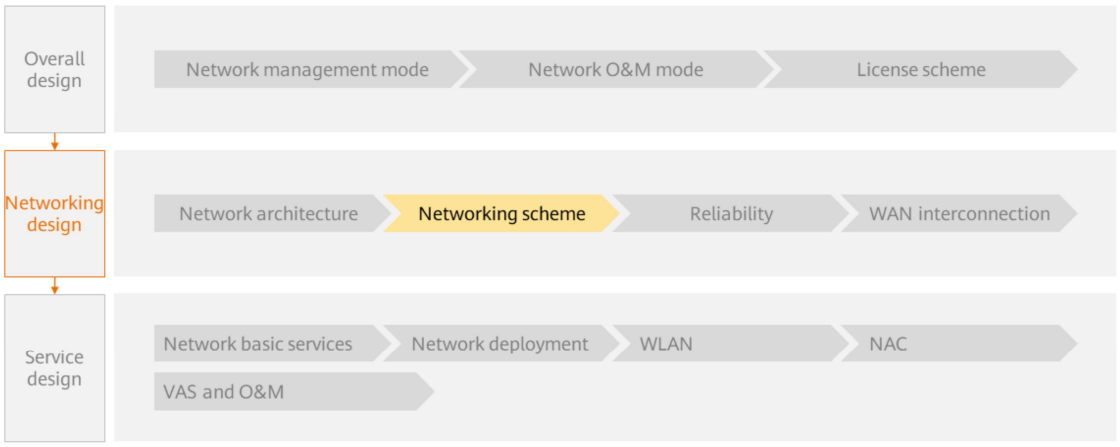
3 (Optional) Determine reliability requirements

Determine the highly reliable networking model based on the device type and capability

- Firewalls are deployed in hot standby mode
- Switch stacking
- Eth-Trunk

HUAWEI

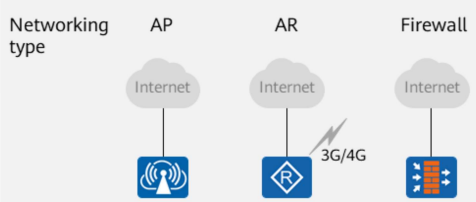
Overall Design Process



HUAWEI

Single-Device Networking

Network scale	< 50 terminals; area < 50 m^2



Overview

- Applicable to single stores and small stores (such as agent stores and gas stations)
- A single device provides gateway features, such as PPPoE dialup, DHCP, and NAT.
- An AP is deployed if only wireless access is required and only one wired Internet egress is available.
- An AR is deployed if wired access is needed and multiple uplinks, especially a 3G/4G backup link, are required.
- A firewall is deployed if security-sensitive stores, such as small logistics, office, and finance organizations, require advanced security features, including URL filtering, security protection, and antivirus.

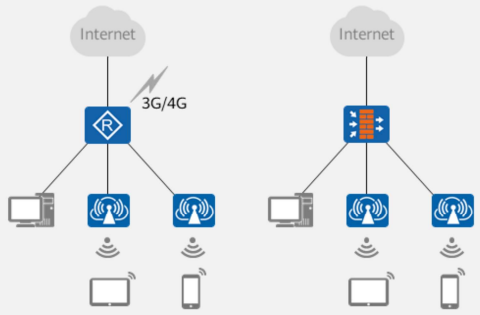
Constraints

- Wireless features such as roaming and load balancing are not supported.
- The AP supports only one Internet link.
- The AP and AR support wired user access, but do not support wired user authentication.
- The AR supports wireless user authentication in only Open, PSK, and Portal authentication modes, but not 802.1X or MAC address authentication mode. The firewall provides wired user authentication and supports only Portal authentication.
- Currently, no firewall model supports Wi-Fi and LTE, and therefore it is not recommended in actual deployments.

HUAWEI

Egress Gateway + AP

Network scale	< 200 terminals; area < 300 m²
Networking type	AR + AP / Firewall + AP



Overview

- Applies to small- and medium-sized clothing stores, supermarkets, shopping malls, etc.
- The egress gateway provides features such as PPPoE, DHCP, NAT, and LTE.
- This networking mode supports continuous coverage of multiple APs, as well as multiple uplinks, including a 3G/4G backup uplink.
- APs support mesh networking.
- A firewall is deployed at the egress for scenarios where advanced security features, including URL filtering, security protection, and antivirus, need to be met.

Constraints

- The AP and AR support wired user access, but do not support wired user authentication.
- The firewall provides wired user authentication and supports only Portal authentication.
- If the AR or firewall does not support PoE, the AP can use an external PoE power adapter.

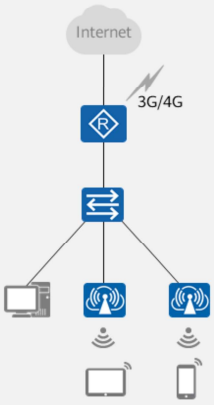
HUAWEI

Egress Gateway + Layer 2 Switch + AP

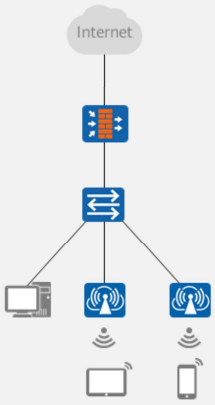
Network scale	< 2000 terminals; area < 3000 m²

Networking type

AR + Layer 2 switch + AP



Firewall + Layer 2 switch + AP



Overview

- Applies to small- and medium-sized clothing stores, retail stores, etc.
- The egress gateway provides features, such as WAN access, DHCP, and NAT.
- The Layer 2 switch provides PoE extended access and wired terminal access functions.
- APs provide access services for wireless terminals at the site.
- APs support mesh networking.
- A firewall is deployed at the egress for scenarios where advanced security features, including URL filtering, security protection, and antivirus, need to be met.

Constraints

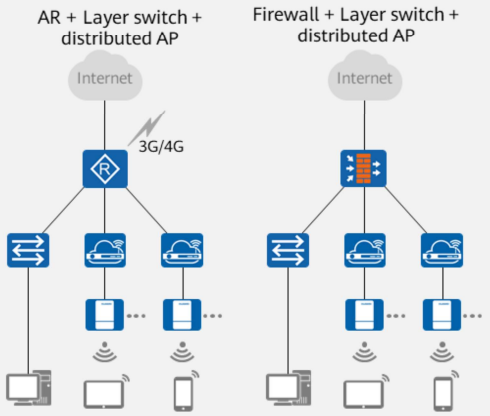
- The AR and firewall in cloud management mode do not support Eth-Trunks and cannot connect to the switch through Eth-Trunks.
- If switches need to be stacked, it is recommended to deploy multiple layers of switches.
- The AR can be deployed only in a single-node system.

HUAWEI

Egress Gateway + Layer 2 Switch + Distributed AP

Network scale	< 2000 terminals; area < 3000 m²

Networking type



Overview

- Applies to dense-room building scenarios, such as dormitories and hotels, where network planning is not required and each room transmits signals separately.
- The AR functions as the egress gateway and provides features, such as WAN access, DHCP, and NAT.
- The Layer 2 switch provides PoE extended access and wired terminal access functions.
- APs provide access services for wireless terminals at the site.
- APs support mesh networking.
- A firewall is deployed at the egress for scenarios where advanced security features, including URL filtering, security protection, and antivirus, need to be met.

Constraints

- The AR and firewall in cloud management mode do not support Eth-Trunks and cannot connect to the switch through Eth-Trunks.
- If switches need to be stacked, it is recommended to deploy multiple layers of switches.
- The AR can be deployed only in a single-node system.

HUAWEI

Summary of Typical Networking Schemes (1)

Deployment Scenario	Network Scale	Key Networking Requirements	Networking Model		Remarks
Many branches, where they are no demands for branch interconnection	< 50 terminals; area < 50 m²	Single store; wireless access only; single Internet egress	Singe-device networking	AP	
		Single store; mainly wireless access; Ethernet or LTE uplink		AR	
		Single store; wired and wireless authentication and access; multiple Internet uplinks and LTE uplinks; high security requirements (URL filtering/IPS/security protection/antivirus)		Firewall	Currently, firewalls have no Wi-Fi/LTE models.
	< 200 terminals; area < 300 m²	Wireless access only; multiple Internet uplinks and LTE uplinks	Egress gateway + AP	AR + AP	Currently, ARs and firewalls do not have PoE models. APs need to connect to external power supplies.
		Wired and wireless authentication and access; multiple Internet uplinks; high security requirements (URL filtering/IPS/security protection/antivirus)		Firewall + AP	
	< 2000 terminals; area < 3000 m²	Wired and wireless access; multiple Internet uplinks and LTE uplinks	Egress gateway + Layer 2 switch + AP	AR + Layer 2 switch + AP	
		Wired and wireless authentication and access; multiple Internet uplinks; high security requirements (URL filtering/IPS/security protection/antivirus)		Firewall + Layer 2 switch + AP	

HUAWEI

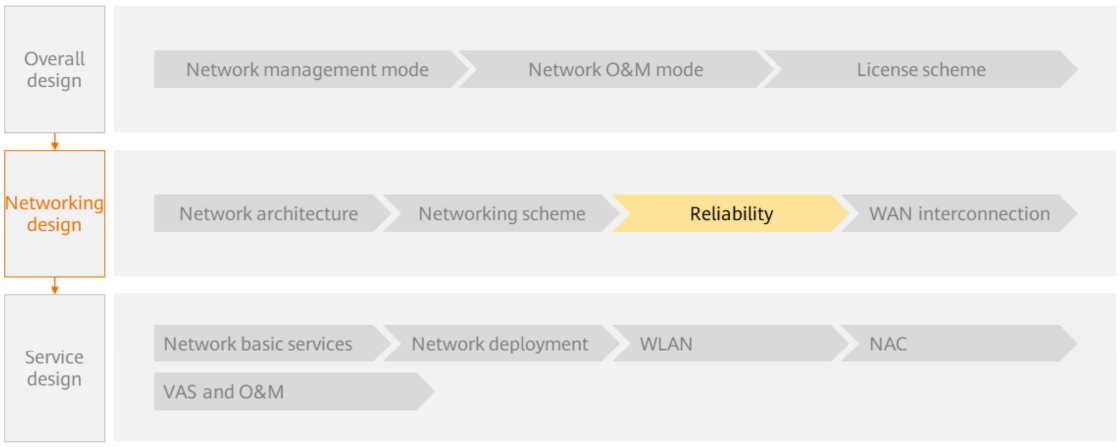
- Cloud management for small- and medium-sized campus networks applies to small- and medium-sized enterprises (SMEs) and branches, such as shopping malls/supermarkets, retail stores, hotels, educational institutions, and financial services organizations. Different networking models are used based on different network scales.

Summary of Typical Networking Schemes (2)

Deployment Scenario	Network Scale	Key Networking Requirements	Networking Model		Remarks
Many branches, where they are no demands for branch interconnection	< 2000 terminals; area < 3000 m^2	Dense-room scenarios such as hotels and dormitories; wired and wireless access; multiple Internet uplinks and LTE uplinks	Egress gateway + Layer 2 switch + distributed AP	AR + Layer 2 switch + central AP + RU	
		Dense-room scenarios such as hotels and dormitories; wired and wireless access; multiple Internet uplinks; high security requirements (URL filtering/IPS/security defense/antivirus)		Firewall + Layer 2 switch + central AP + RU	
Multiple branches, where there are demands for communication with the headquarters	Same as that in the "many branches, branches, where they are no demands for branch interconnection" scenario	Only basic communication between branches and the headquarters is required. Third-party VPN gateways may be deployed at the headquarters.	IPsec VPN interconnection	AP/AR/firewall as the egress gateway	
		Multiple links are deployed between the headquarters and branches, and intelligent traffic steering is required.	SD-WAN Interconnection	AR as the egress gateway	

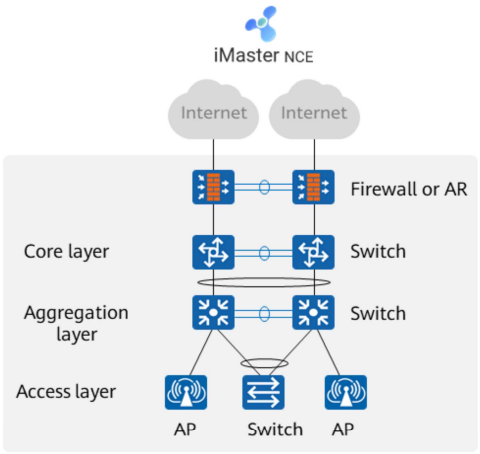
HUAWEI

Overall Design Process



HUAWEI

Network Reliability Design



Highly reliable networking architecture

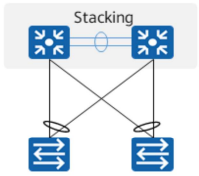
- Networking reliability refers to the reliability design of the tenant network. You can add devices or links in any typical networking scheme to enhance network reliability. Network reliability design can be classified into three levels based on the involved scope:
  - Link-level reliability design
  - Device-level reliability design
  - Network-level reliability design (which combines link and device reliability)

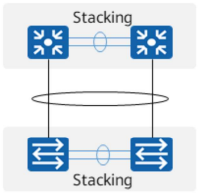
HUAWEI

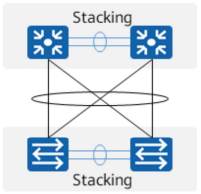
Link-Level Reliability Design (1)

Intranet link reliability

- Multiple links can be deployed between switches and bonded using Eth-Trunks to improve link reliability.
- When switches are stacked, it is recommended that Eth-Trunks be used across switches to ensure link reliability.





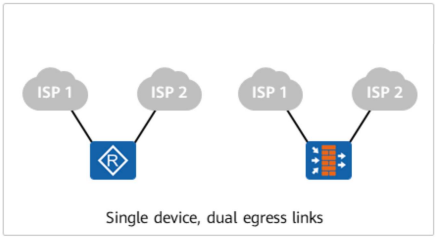


Link-Level Reliability Design (2)

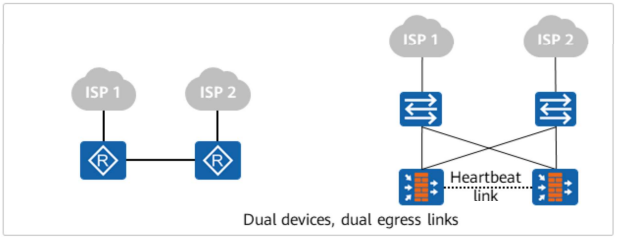
Egress link reliability

Multiple links are deployed at the campus egress and work in active/standby mode. The networking scheme is as follows:

- The "single device, multiple egress links" scheme or the "multiple devices, at least one egress link for each device" scheme can be used. The latter scheme is recommended.
- To further improve reliability, it is recommended that different egress links be connected to different carrier networks.

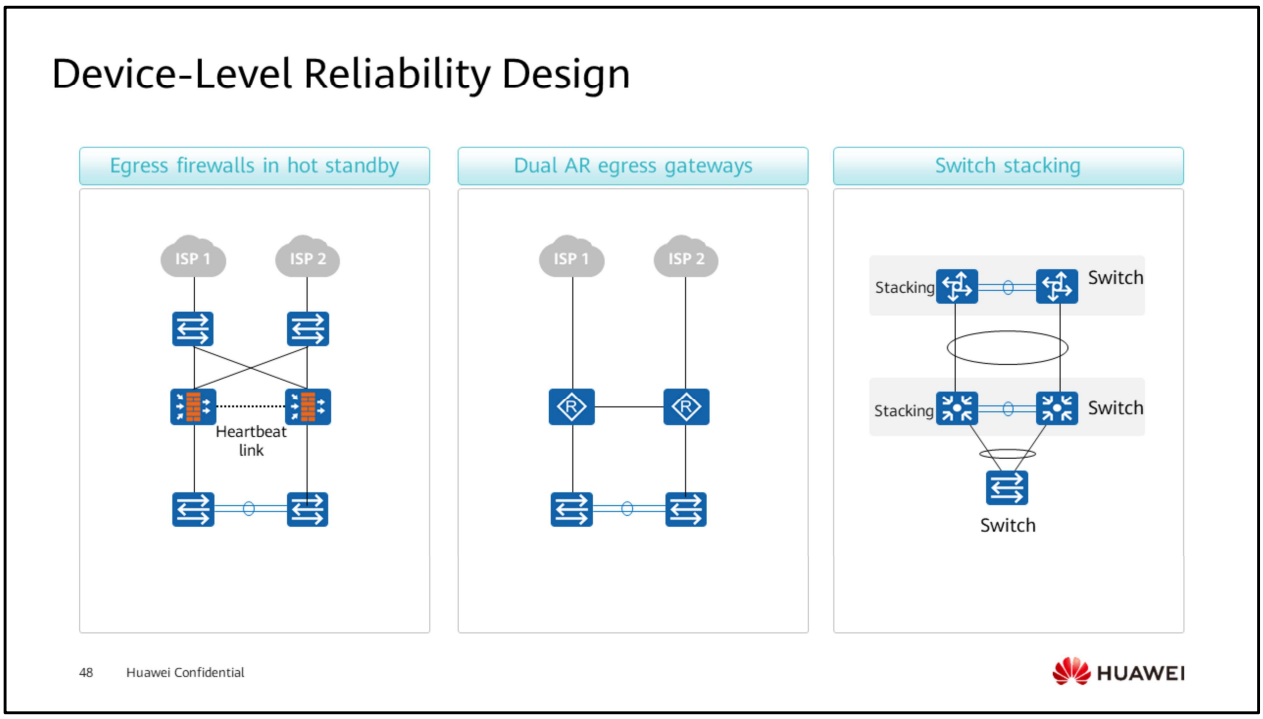


Single device, dual egress links



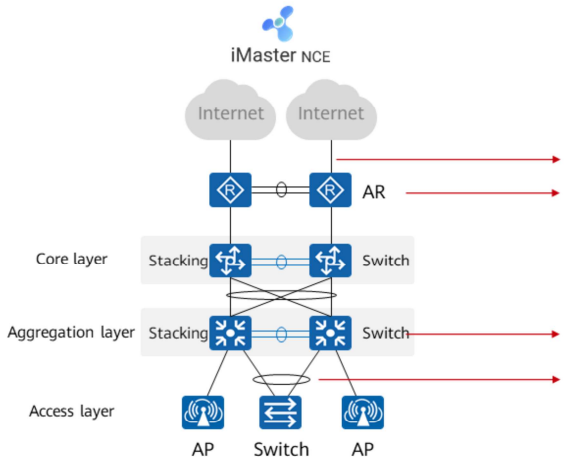
Dual devices, dual egress links

HUAWEI



- Device reliability: Fixed-configuration devices are generally used on small- and medium-sized campus networks.
- Device reliability design includes the following two aspects:
- Redundant components are used to improve device reliability.
    - In addition to selecting devices with high reliability, you can further improve the reliability of the devices by using dual power supplies or redundant components (e.g. boards).
- Joint networking is used to provide device reliability.
    - Multiple devices are deployed together to improve device reliability, for example, active/standby or stacking.
    - Two devices are deployed in the egress zone of the campus network to work in active/standby mode. Currently, in a two-node system, firewalls support only the hot standby mirroring mode. ARs support dual-CPE networking, and the two devices and two egress links can work concurrently.
    - Intranet network devices, such as LAN switches at the core or aggregation layer, are stacked to implement device-level backup. Currently, only switches can be stacked.

Network-Level Reliability Design



Typically, Eth-Trunk and stacking technologies are used in dual-device and multi-link scenarios to improve network-wide reliability.

Egress reliability design:

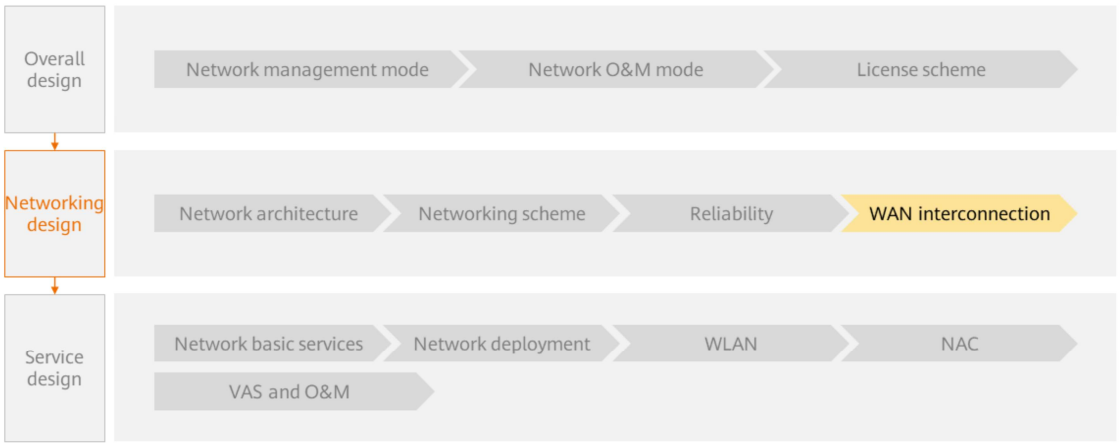
- Dual-device: active/standby, load balancing
- Multi-link, multi-carrier network

Intranet reliability design: switch networking

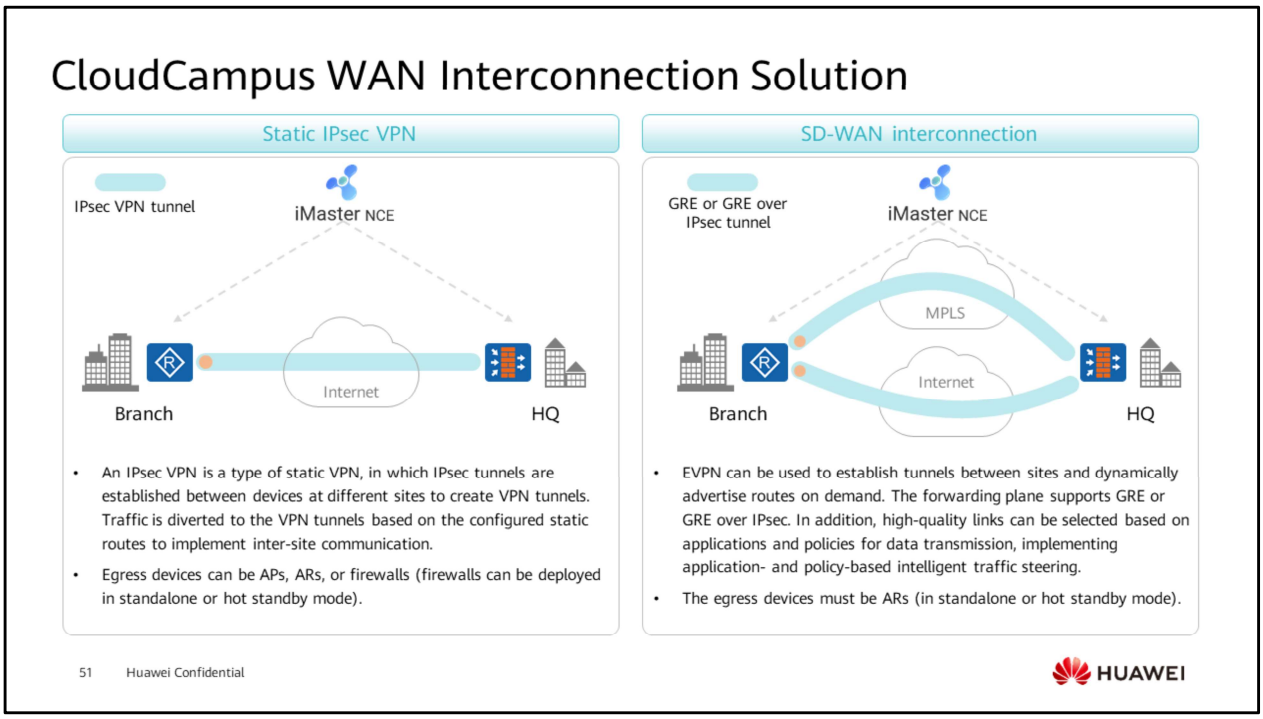
- Dual-/multi-device: stacking
- Dual-/multi-link: Eth-Trunk

HUAWEI

Overall Design Process

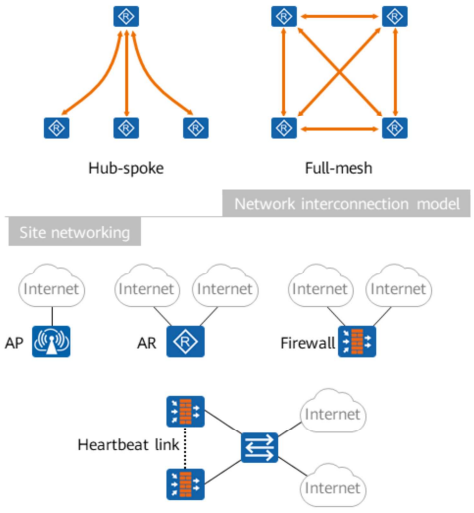


HUAWEI



- The multi-campus network interconnection solution is a sub-solution provided in the CloudCampus Solution for the interconnection between branch campuses and between branches and the HQ or DCs. With the SD-WAN solution integrated, the multi-campus network interconnection solution provides two models for WAN interconnection: static IPsec VPN and EVPN.
- An IPsec VPN is a type of static VPN, in which IPsec tunnels are established between devices at different sites to create VPN tunnels. Traffic is diverted to the VPN tunnels based on the configured static routes so that services between sites can be accessed through the VPN tunnels.
- An EVPN is a type of dynamic VPN that can establish tunnels between sites and dynamically advertise routes on demand. EVPN establishes GRE tunnels between sites to establish VPN tunnels and supports IPsec encryption on GRE tunnels to ensure tunnel encryption security. In addition, the EVPN solution offers application- and policy-based intelligent traffic steering, allowing high-quality links to be selected based on applications and policies for data transmission.

IPsec VPN Interconnection Networking Model



Network interconnection model

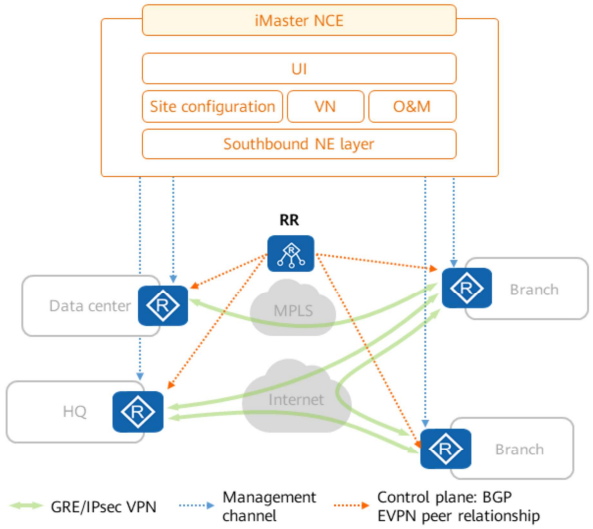
- The hub-spoke and full-mesh models are supported.
- The hub-spoke model is applicable when data traffic is mainly transmitted between branches and the HQ. In this model, branches can also communicate with each other through the HQ.
- If branches are of similar scales and a large amount of traffic is transmitted between the branches, the full-mesh model can be used. In this case, all egress devices must use public IP addresses, and only firewalls support this model. In the full-mesh model, the number of sites is limited (32 at most). Therefore, the hub-spoke model is recommended for IPsec interconnection.

Site networking

- APs, ARs, or firewalls can be used as egress devices based on site requirements.
- In some scenarios requiring high reliability, two egress links can be deployed. In scenarios requiring wireless uplinks, the egress gateway must be an LTE-capable device.
- For large campus networks, the egress gateways can be deployed in hot standby mode to ensure high reliability. Currently, only firewalls support hot standby.
- In the hub-spoke model, the hub site is usually the HQ or DC site, which requires high-performance devices, rich policies, and security configurations. Currently, the cloud management mode supports only a few security and policy features. Therefore, it is recommended that the traditional management mode be configured for devices at the hub site, and the devices work in hot standby mode.

HUAWEI

SD-WAN Solution Overview (1)



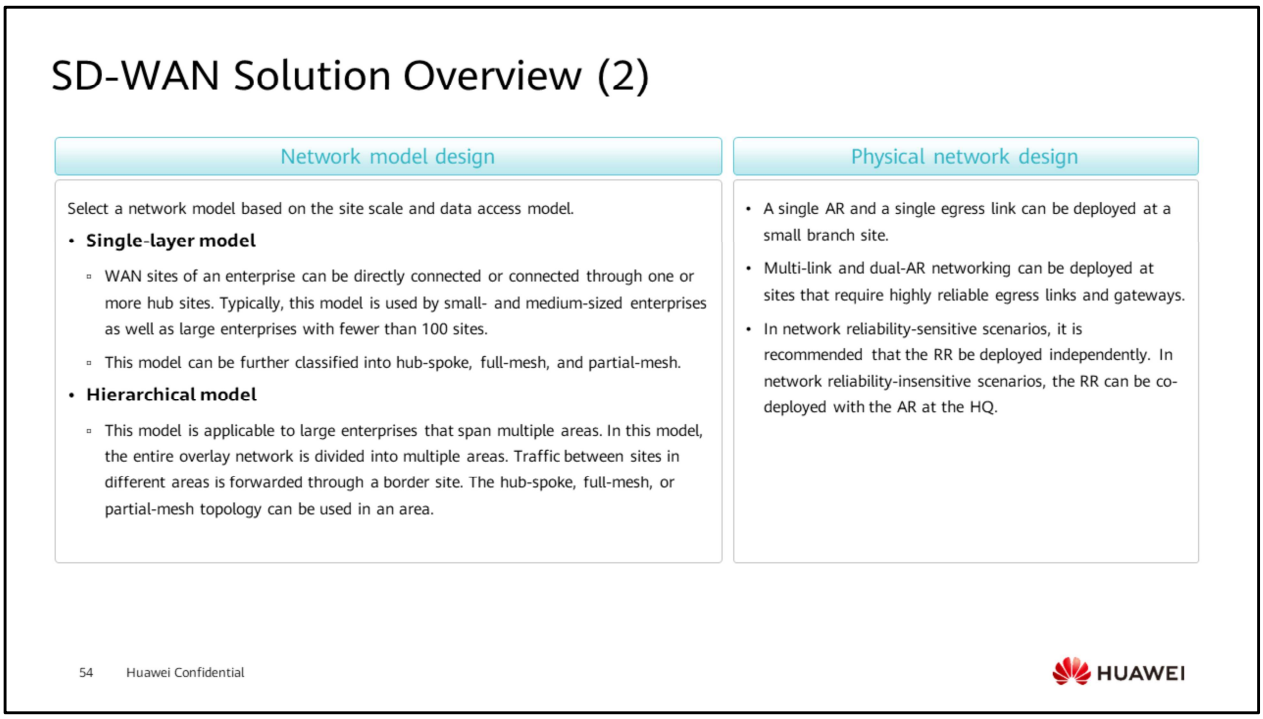
Solution overview

- The EVPN interconnection mode of the SD-WAN solution is applicable to the scenarios where a large number of branches communicate with the HQ/DC (hub site). This mode can meet the networking requirements of multiple HQs/regional centers. In this mode, the branches can communicate with the HQ through multiple links. In addition, this mode supports application- and quality-based intelligent traffic steering, QoS scheduling, and link quality display.
- The hub-spoke, full-mesh, and partial-mesh networking models are supported.
- Hub sites can work in active/standby mode to improve network access reliability at the HQ.
- Multi-VPN (VRF) and WAN-side multi-VPN interconnection are supported.
- In scenarios requiring high security, such as the finance industry, both firewalls and ARs can be deployed at sites.

Solution constraints

- ARs deployed at the HQ and branches must support SD-WAN.

HUAWEI



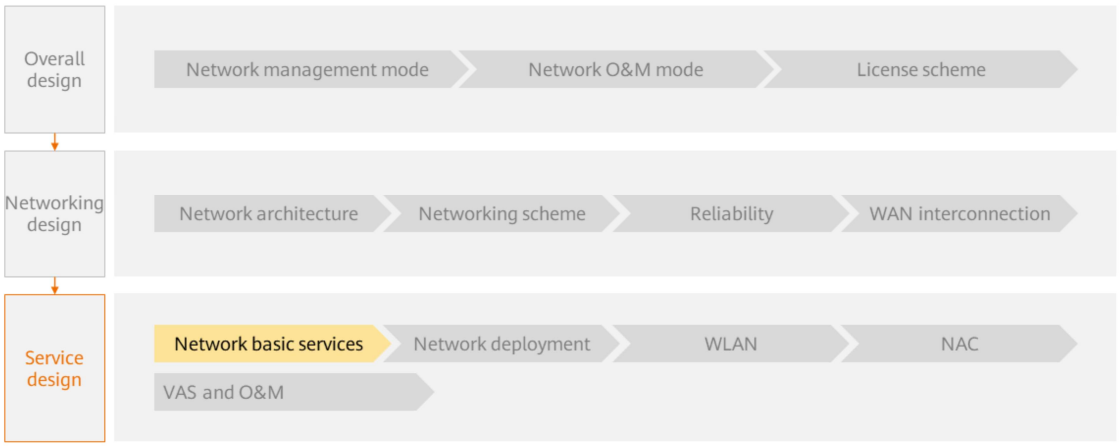
- The single-layer network model is also called the flat network model. In this model, WAN sites of an enterprise can be directly connected or connected through one or more hub sites. Typically, this model is used by small- and medium-sized enterprises as well as large enterprises with fewer than 100 sites. The single-layer network model can be further classified into hub-spoke, full-mesh, and partial-mesh.
- The hierarchical network model is applicable to large enterprises that span multiple areas. In this model, the entire overlay network is divided into multiple areas. Traffic between sites in different areas is forwarded through a border site. The hub-spoke, full-mesh, or partial-mesh topology can be used in an area.

SD-WAN Solution: Service Design

- **VPN planning**: A network can be divided into multiple VPNs based on services or departments. Policies can be independently configured for services of each VPN, and VPNs can be isolated from each other.
- **Intelligent traffic steering**: Traffic policies can be defined (based on network segments or applications) to design intelligent traffic steering, so that high-reliability links (such as MPLS links) are preferentially selected for high-priority applications (such as audio and video services) and low-priority links are selected to transmit common data.
- **QoS application**: Multi-level QoS scheduling, such as three-level QoS scheduling for ports, VPNs, and applications
- **Security**: ACL-based traffic filtering, URL filtering, firewall, and IPS are configured to ensure traffic security.
- **Site-to-Internet access**: The local Internet access, centralized Internet access, and hybrid Internet access modes are supported. If the Internet egress link is available, local Internet access can be configured. If centralized traffic policing is required, centralized Internet access will suffice.
- **Interworking with traditional networks**: The solution supports interworking with traditional networks.

HUAWEI

Overall Design Process



HUAWEI

VLAN Planning

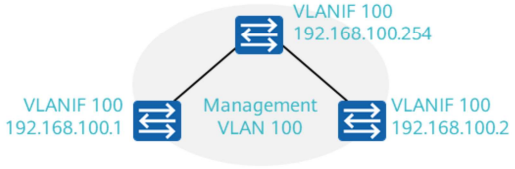
- Classify VLANs into service VLANs, management VLANs, and interconnection VLANs.
- Allocate consecutive VLAN IDs to ensure proper use of VLAN resources.
- Reserve a specific number of VLANs for future use.
- Typically, VLANs are divided based on interfaces. According to different design principles, interfaces of access switches are added to different VLANs so that users of different service types can be isolated.

VLAN Assignment	Example
VLAN assignment by logical area	• Core network area: VLANs 100–199 • Server area: VLANs 200–999 (VLANs 1000–1999 are reserved.) • Access network area: VLANs 2000–3499 • Service network area: VLANs 3500–3999
VLAN assignment by geographic area	• Area A: VLANs 2000–2199 • Area B: VLANs 2200–2399
VLAN assignment by personnel structure	• Department A in area A: VLANs 2000–2009 • Department B in area A: VLANs 2010–2019
VLAN assignment by service type	• Web server area: VLANs 200–299 • Application server area: VLANs 300–399 • Database server area: VLANs 400–499

HUAWEI

IP Address Planning (1)

Management IP address

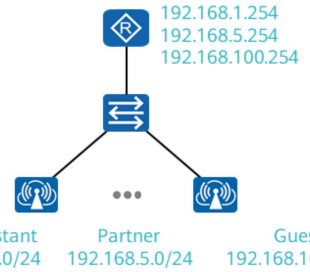


A Layer 2 device uses the VLANIF interface's IP address as the management IP address. It is recommended that all Layer 2 switches connected to a gateway be on the same network segment.

Interconnection IP address

It is recommended that the interconnection IP address use a 30-bit mask. The core device uses a smaller host IP address.

Service IP address



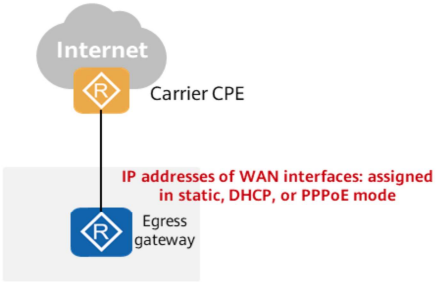
The service IP address is the IP address of a server, host, or gateway.

- It is recommended that the gateway addresses use the same rightmost three digits, such as .254.
- The IP address range of each service must be clearly distinguished. The IP addresses of each type of service terminals must be contiguous and can be aggregated.
- You are advised to use an IP address segment with a 24-bit mask.

HUAWEI

IP Address Planning (2)

Egress gateway

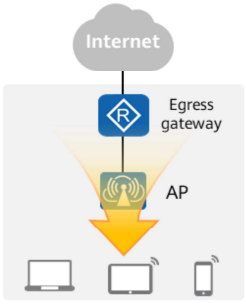


IP addresses of WAN interfaces on egress gateways are assigned by the carrier in static, DHCP, or PPPoE mode. The IP addresses of these interfaces need to be obtained from the carrier in advance.

Devices such as servers and printers

It is recommended that servers and special terminals (such as punch-card machines, printing servers, and IP video surveillance devices) use static IP addresses.

End user

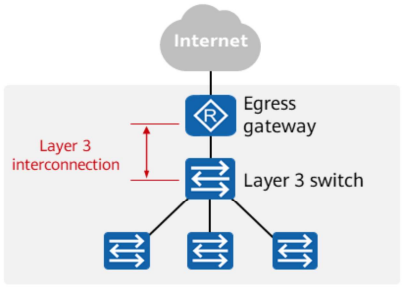


It is recommended that end users be assigned IP addresses in DHCP mode and the gateway provide the DHCP service.

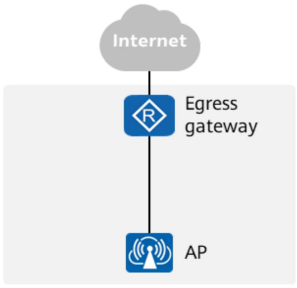
HUAWEI

IP Address Planning (3)

LAN-side devices



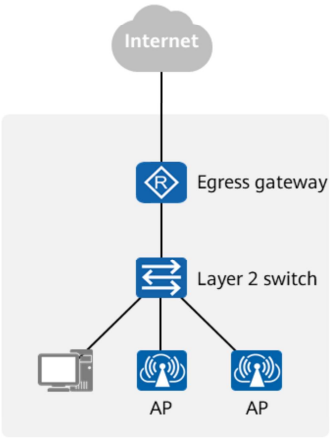
When the egress gateway interconnects with a Layer 3 switch, it is recommended that the interconnection IP addresses be manually configured in static mode.



It is recommended that the DHCP server be deployed on the gateway to dynamically allocate IP addresses to APs.

HUAWEI

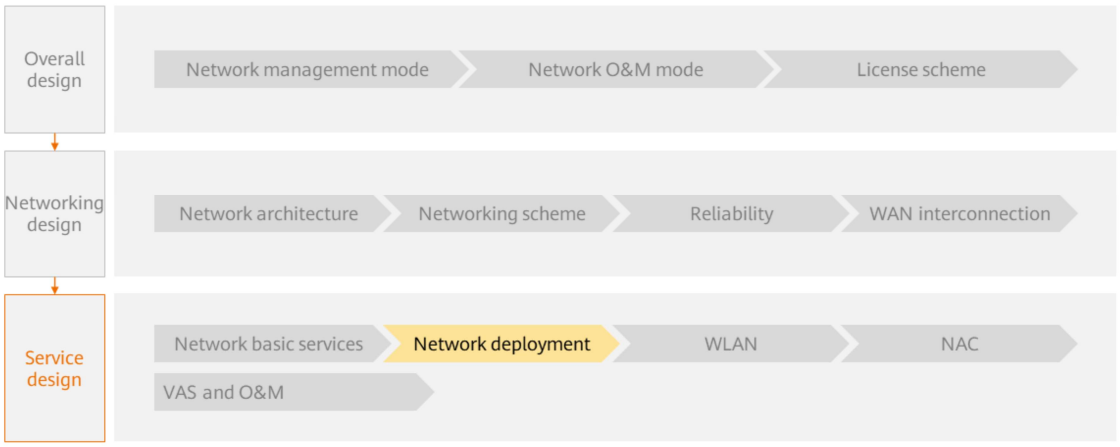
Routing Design



- **Static routes are recommended for small- and medium-sized campus networks.**
- **Internal routing design of the campus network:**
  - APs: After an AP obtains an IP address through DHCP, a default route is generated by default.
  - Switches and gateways: Static routes can be used to meet requirements. No complex routing protocol needs to be deployed.
- **Routing design for the campus egress:**
  - You are advised to configure static routes on the egress device.

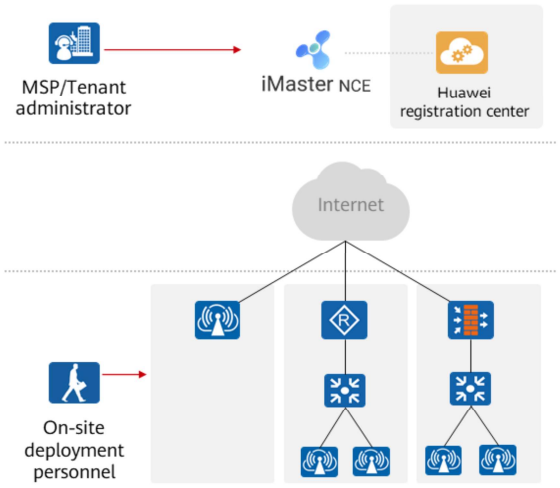
HUAWEI

Overall Design Process



HUAWEI

Network Deployment and Automated Device Registration



Automated network deployment

- Huawei's Cloud-Managed Network Solution supports device plug-and-play. If the following prerequisites are met, a device can automatically connect to a given cloud management platform after being powered on and obtain complete service configurations from the platform, achieving device plug-and-play and fast service provisioning.

Prerequisites for implementing device plug-and-play

- The administrator has registered the device ESN on the cloud management platform (by scanning the barcode or manually entering the ESN) and bound a valid license authorization package to the device.
- The administrator has pre-configured the device or device group offline on the cloud management platform.
- The device has obtained a valid IP address and DNS service through the DHCP server (another cloud managed device or deployed by the customer), and the IP address can be used to communicate with the Internet.

HUAWEI

- Traditional network deployment requires the participation of engineering installation personnel and network commissioning personnel. This results in long installation and commissioning periods and slow service provisioning, and consumes a large amount of manpower. The CloudCampus Cloud-Managed Network Solution supports automated network deployment.

Device Registration

Enterprise network scenarios are complex, and different device registration modes are required in different scenarios. Huawei devices support multiple registration modes to meet complex networking requirements. The registration model also varies with device model. The following table lists the registration modes of different devices:

Device	Registration Mode					
	Registration Center	DHCP Option	USB Flash Drive	Mobile App	Web/CLI	Email
AP	√	√		√	√	
Firewall	√		√		√	
Switch	√	√			√	
AR	√	√	√		√	√

On the same network, devices at different locations may use different registration modes. Typically, devices to be registered on a network are classified into:

- Egress gateway
- Intranet (LAN-side) device

HUAWEI

Egress Gateway Registration

- Egress line configuration of the egress gateway usually depends on the carrier network. Therefore, the registration mode is complex.
- In the CloudCampus Solution, ARs, firewalls, and APs can function as egress gateways.
- The registration modes of egress gateways vary with scenarios. For details, see the following table.
- If a DHCP server exists on the live network, the egress gateway can be registered using the DHCP option only when the DHCP server supports DHCP Option 148.
- In LAN-WAN convergence scenarios, ARs support email-based and DHCP option-based registration modes. However, in the DHCP option-based registration mode, the DHCP server must support DHCP Option 148.

Networking Scenario		Recommended Egress Gateway Registration Mode		
		AR	Firewall	AP
Huawei public cloud scenario	The egress gateway cannot automatically obtain an IP address.	Web system	Web system	CloudCampus APP
	The egress gateway can automatically obtain an IP address.	Registration center		
MSP-owned cloud scenario	The egress gateway cannot automatically obtain an IP address.	Web system	Web system	CloudCampus APP
	The egress gateway can automatically obtain an IP address.	DHCP option	Web system	DHCP option
LAN-WAN convergence scenario (where the SD-WAN function needs to be configured on the egress gateway)		Email	---	---

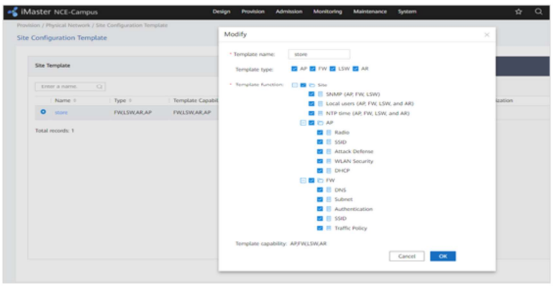
HUAWEI

Intranet (LAN-Side) Device Registration

- A LAN-side network refers to the network under the egress gateway, which is mainly built on LAN switches and APs.
- In Huawei public cloud scenarios, the Huawei registration center-based registration mode is recommended.
- If an enterprise does not want to synchronize device information to the registration center, the DHCP option-based registration mode can be used.

Networking Scenario		LAN-Side Device Registration Mode	
		LAN Switch	AP
Huawei public cloud scenario	DHCP options cannot be configured on the network.	Registration center	
	DHCP options can be configured on the network.	Registration center or DHCP option	
MSP-owned cloud scenario	DHCP options cannot be configured on the network.	Web system	CloudCampus APP
	DHCP options can be configured on the network.	Registration center or DHCP option	

Scenario-Specific Automated Deployment



Solution overview

- The scenario-specific automated deployment function is provided to further simplify network deployment.
- Network configurations, such as SSIDs, can be automatically generated based on the industry scenario selected by a customer. This means that users do not need to plan or configure the network in advance, eliminating the need for complex configuration design and planning. In addition, users can customize a scenario template to plan the site network so that the site network configuration can be quickly generated for deployment.
- Currently, templates for small- and medium-sized office scenarios are preset in the solution. Templates for other scenarios will be added in the future.

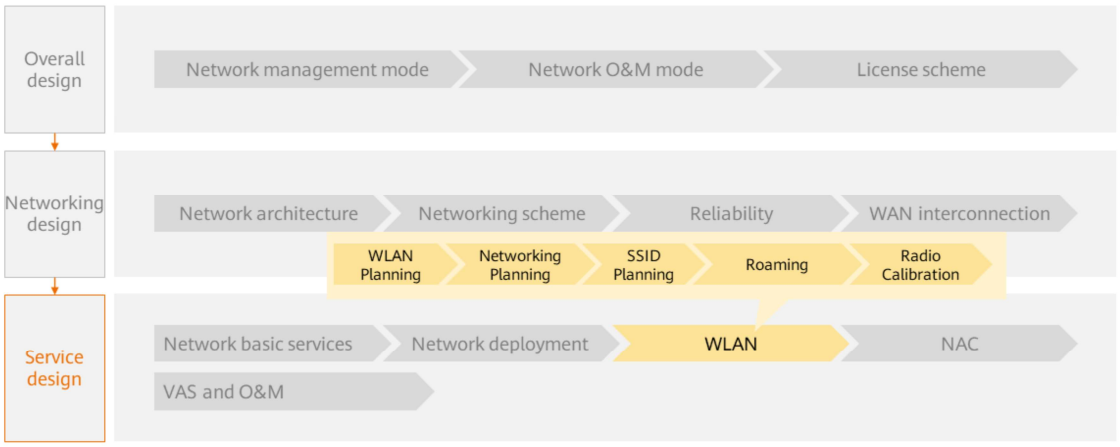
Key design points:

- Select a service scenario, that is, select a scenario template.
- Select related devices and network topologies.
- Confirm the related configuration or adjust the preset configuration.

HUAWEI

- A scenario template includes the following contents and functions related to network deployment:
  - Various networking modes: typical networking models of small- and medium-sized campus networks
  - Automated DHCP server configuration
  - Dual uplinks (active/standby mode by default) on the firewall egress for higher reliability
  - Automated Eth-Trunk configuration
  - Automated configuration of network services, including SSID, PSK password, QoS, access authentication, traffic monitoring, and application monitoring

Overall Design Process



HUAWEI

WLAN Planning: Wireless Signal Coverage (1)

Field strength reference

Coverage Area	Field Strength Requirements	Typical Scenario
Key coverage area	–40 to –65 dBm	Dormitory room, library, hotel room, lobby, conference room, office, and exhibition hall
Common coverage area	> –75 dBm	Corridor, kitchen, storeroom, and dressing room
Special coverage area	N/A	Areas that have limitations on or do not allow coverage or installation due to considerations such as service security or property management

Physical coverage reference

Scenario	Coverage Recommendation
Indoor scenario	Plan the coverage radius of 15–20 m for each AP.
Outdoor scenario	Plan the coverage radius of 50–80 m for each AP.
Indoor high-density scenario	Use small-angle directional antennas. During network planning, select AP positions and spacing based on the antenna angle.

To guarantee a good access and roaming experience, you need to plan proper signal coverage depending on different application scenarios; select proper AP models based on signal coverage demands and AP capabilities; and design a proper WLAN using the WLAN Planner.

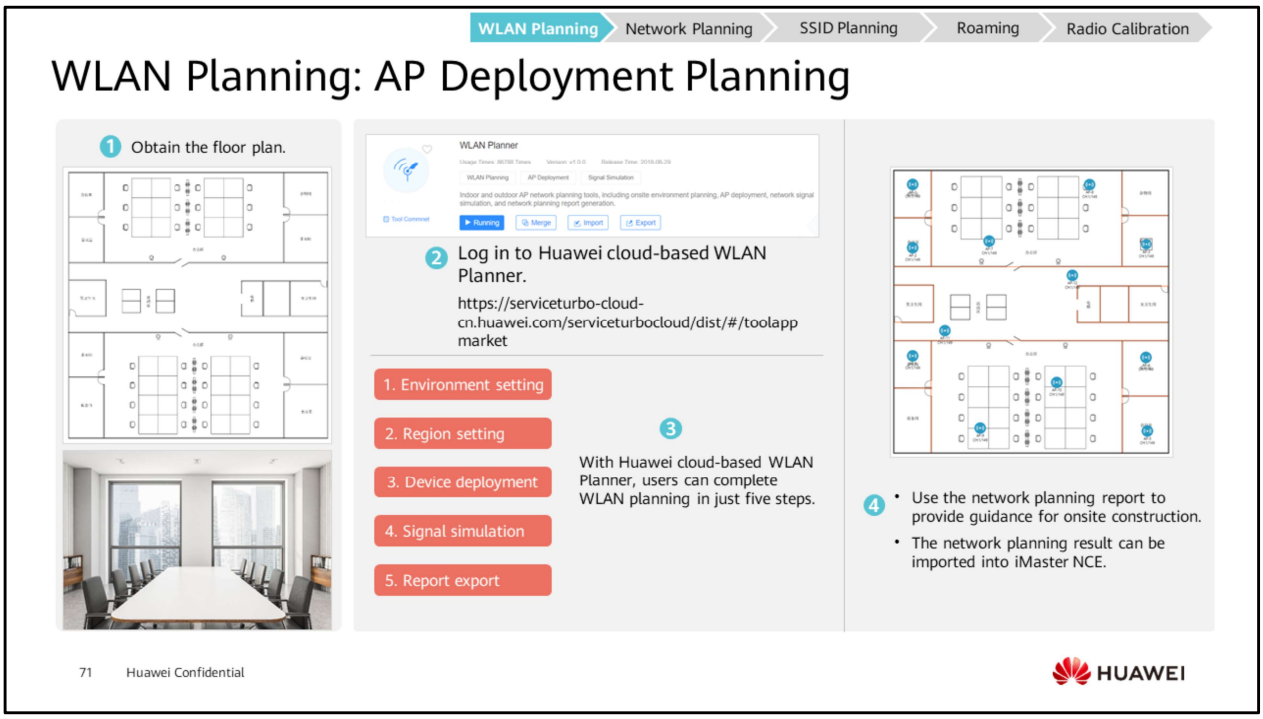
HUAWEI

WLAN Planning: Wireless Signal Coverage (2)

Observe the following rules when deploying APs:

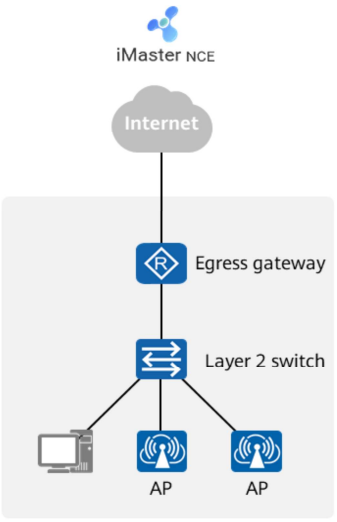
- When installing an AP, try to reduce the number of obstacles that signals pass through.
- Ensure that the front side of an AP faces the target coverage area for good coverage.
- Place APs away from interference sources, such as electronic devices. Do not deploy microwave ovens, wireless cameras, wireless phones, or other electronic devices in the coverage area.
- For areas with roaming requirements, keep a 10% to 15% signal overlap between the coverage areas of neighboring APs to ensure smooth STA roaming between APs.

HUAWEI



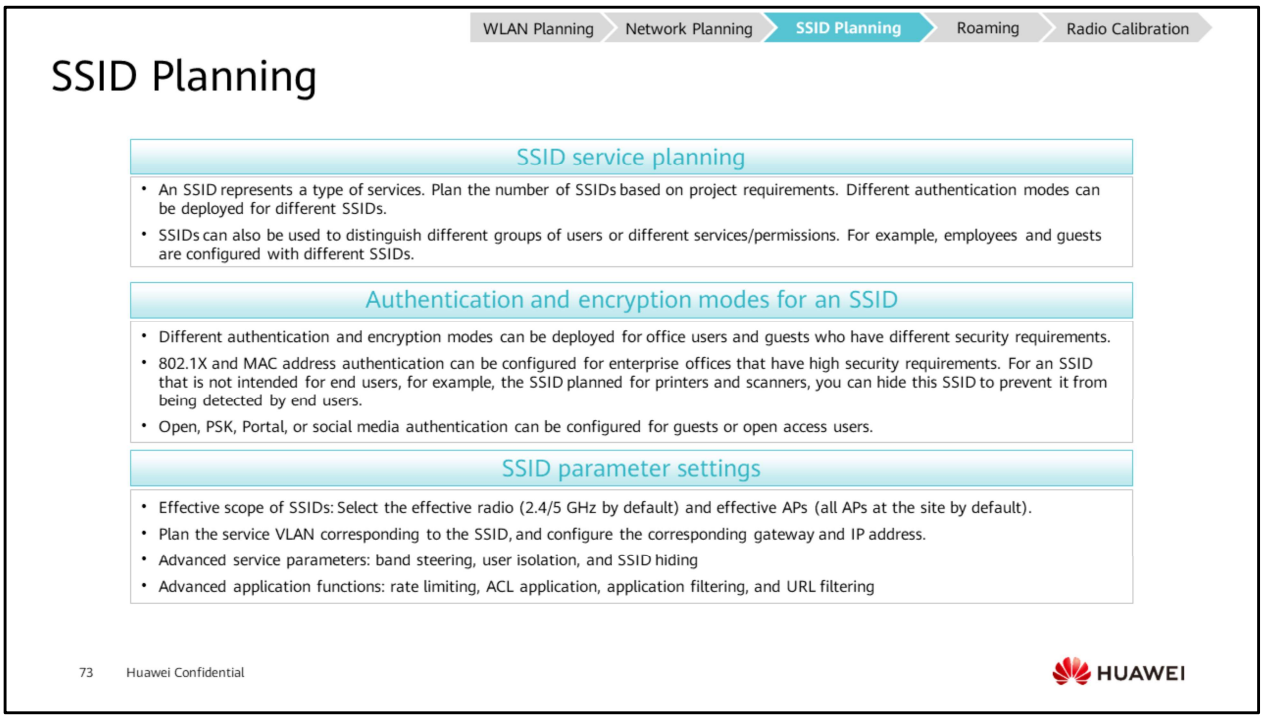
- After WLAN signal coverage requirements are determined based on scenarios, AP deployment can be planned using the WLAN Planner.
  - This tool can be used to create network planning projects based on drawings or Amap.
  - This tool supports region settings, obstacle generation, automated or manual AP deployment with one click, and WLAN signal coverage simulation.
  - This tool can generate network planning files and allows users to export the files. Users can import network planning results on the tenant management page of iMaster NCE to display AP locations and help tenants install APs.
  - CampusInsight allows users to import network planning files. After network deployment, users can view the actual signal heatmap of the network.

Network Planning

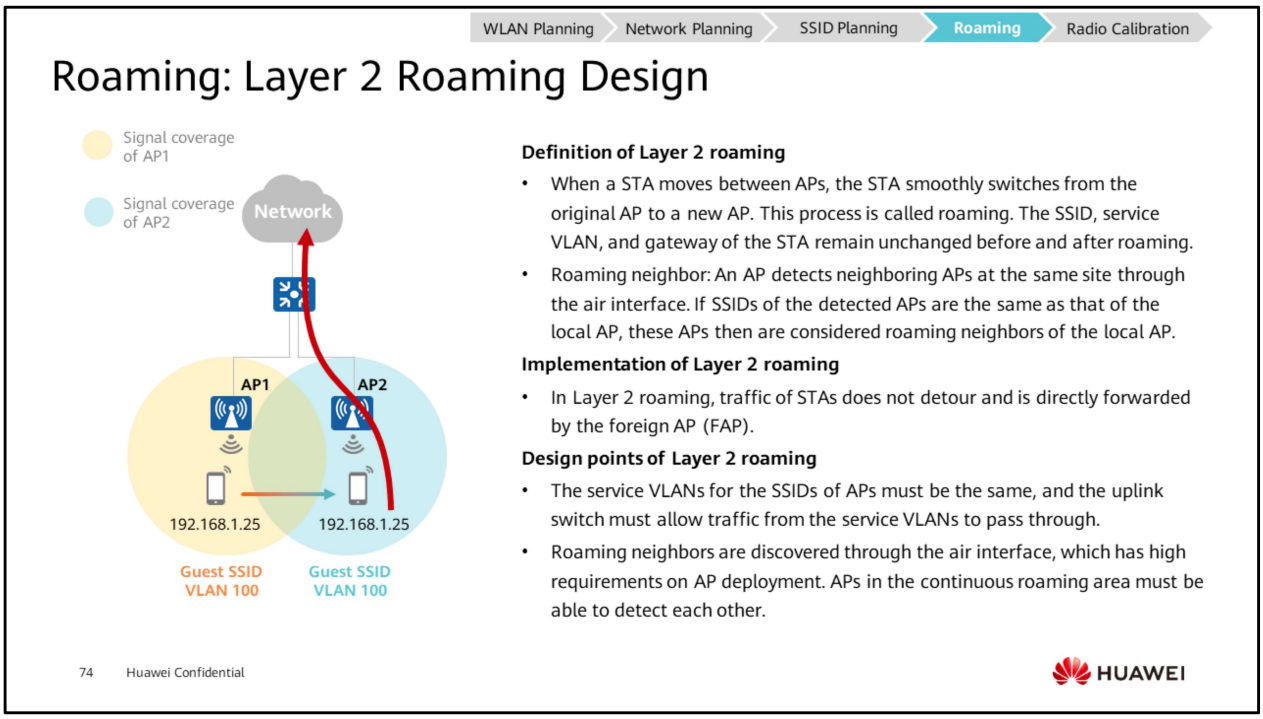


- **Number of APs**: Determine the number of APs based on the WLAN planning result.
- **Access switch selection**: PoE switches are recommended. Select switches based on the number of APs and PoE power supply requirements of APs.
- **Number of access switches**: Determine the number of access switches based on the number of APs, access relationships, and the number of ports on the selected switches.

HUAWEI



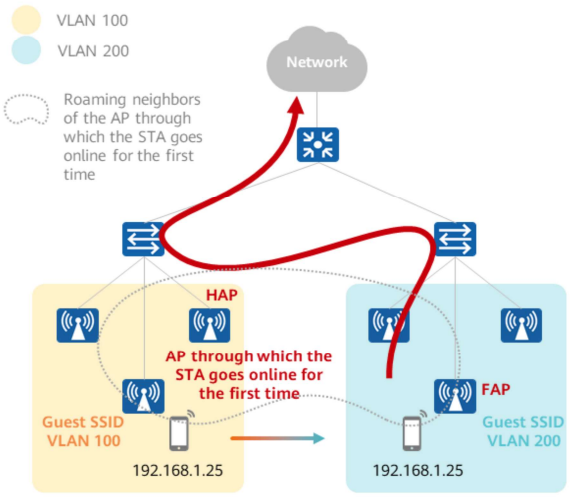
- SSIDs are used to represent WLANs. They are WLAN names displayed on a STA such as a mobile phone when you search for WLANs available for access on the STA. You can select an SSID to access the corresponding WLAN.
- Generally, SSIDs are planned based on user roles or services. Different authentication modes, network access rights, and control policies can be deployed for different SSIDs.
- Portal authentication can be used on open networks (for example, networks for external user access). PSK or PPSK authentication can be used on semi-open networks (such as guest room networks of hotels). 802.1X authentication can be used on secure networks (such as office networks). MAC address authentication can be used for dumb terminals such as printers.
- For an SSID that is not intended for end users, for example, the SSID planned for printers and scanners, you can hide this SSID to prevent it from being detected by end users.



- Definition of Layer 2 roaming: When a STA moves between APs, the STA smoothly switches from the original AP to a new AP. This process is called roaming. The SSID, service VLAN, and gateway of the STA remain unchanged before and after roaming.
- Implementation of Layer 2 roaming:
  - Roaming neighbor: An AP detects neighboring APs at the same site through the air interface. If SSIDs of the APs are the same as that of the local AP, and they can detect each other, these APs are considered roaming neighbors of the local AP. (An AP can establish a maximum of 64 roaming neighbors.).
  - Each AP establishes a wired CAPWAP tunnel (control plane tunnel) with a roaming neighbor to transmit roaming information.
  - The establishment of roaming neighbors is irrelevant to the calibration group. The calibration group is divided based on the management VLAN. Layer 2 roaming neighbors are established based on whether service VLANs are the same.

    - When a STA goes online through an AP, the AP synchronizes information such as the STA's MAC address to all roaming neighbors of the AP. If the STA roams to a new AP (FAP), the AP checks whether the STA roams from another AP, and synchronizes the STA entry from the home AP (HAP). The entry includes the VLAN, IP address, authentication result, and authorization group of the STA. The FAP then generates the STA entry. In this case, identity authentication does not need to be performed for the STA again. After roaming completes, the FAP notifies its neighboring APs to prepare for the STA's subsequent roaming. In addition, after roaming is complete, the HAP deletes the STA entry.
    - In Layer 2 roaming, traffic of STAs does not detour and is directly forwarded by the FAP.
- Design points of Layer 2 roaming:
    - The service VLANs for the SSIDs of APs must be the same, and the uplink switch must allow traffic from the service VLANs to pass through.
    - Roaming neighbors are discovered through the air interface, which has high requirements on AP deployment. APs in the continuous roaming area must be able to detect each other.

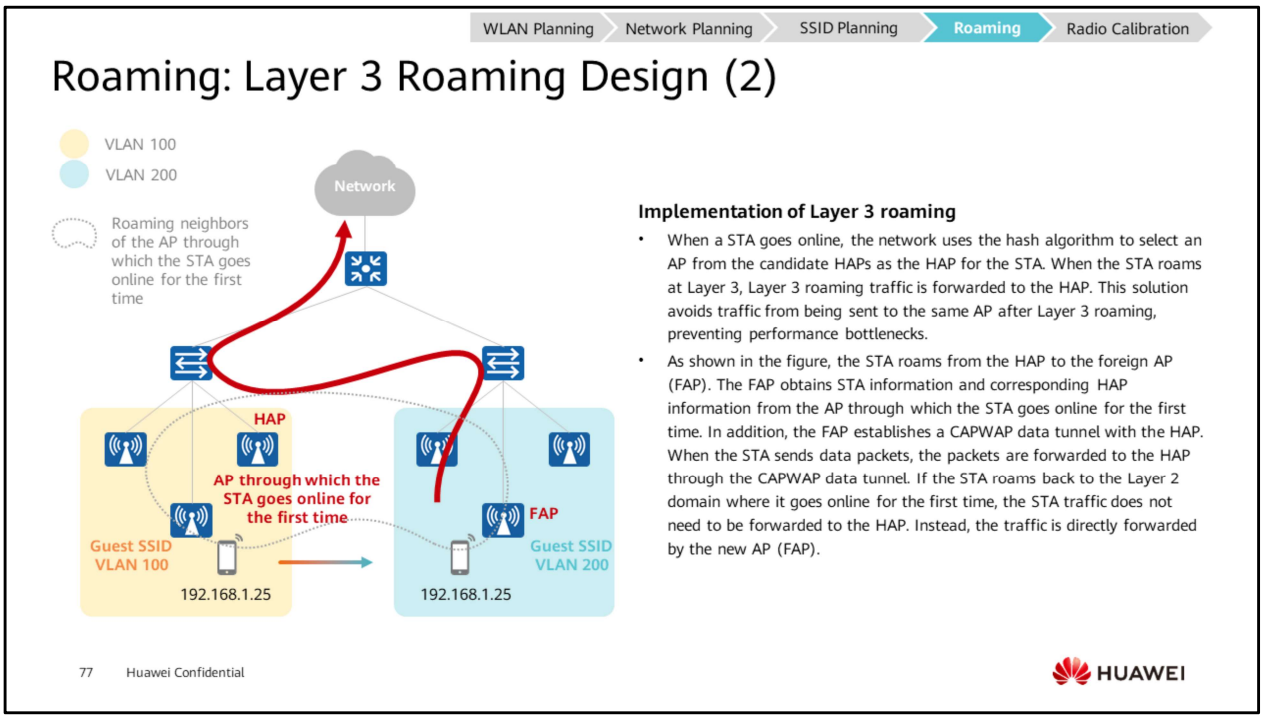
Roaming: Layer 3 Roaming Design (1)



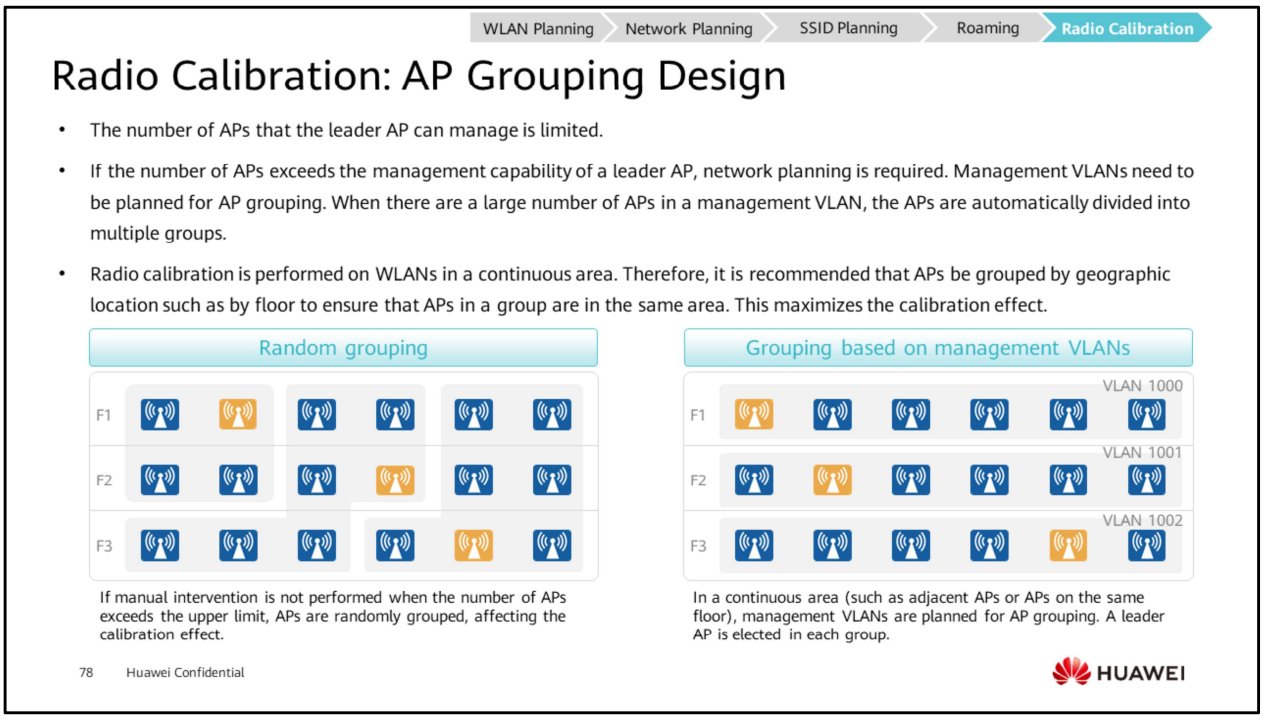
Definition of Layer 3 roaming

- **Layer 3 roaming**: The service VLANs for the SSIDs before and after STA roaming are different and correspond to different gateways. In this case, to keep the IP address of the roaming STA unchanged, the STA traffic needs to be detoured to an AP that resides on the same network segment as the AP through which the STA goes online for the first time.
- **AP through which the STA goes online for the first time**: AP that the STA first associates with in a mobility group
- **Home AP (HAP)**: is selected using the hash algorithm from the neighbors that belong to the same Layer 2 roaming domain as the AP through which the STA goes online for the first time. Once selected, the HAP forwards Layer 3 roaming traffic of the STA.

HUAWEI

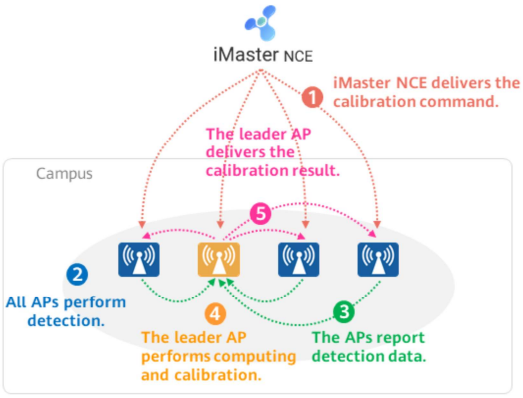


- Design points of Layer 3 roaming
  - Layer 3 roaming traffic will be detoured to the HAP. Therefore, a large Layer 2 domain needs to be planned at the entrance of the entrance hall so that the detoured traffic can be distributed to different APs during roaming.
  - Each AP supports up to 64 Layer 3 roaming STAs. When there are a large number of Layer 3 roaming STAs, the roaming fails and the STAs need to go offline and go online again. (In actual deployment scenarios, you are advised to deploy Layer 2 roaming on the same floor or in the same area based on the service VLAN planning.)



- A cloud AP is essentially a Fat AP and is independent of each other. Some WLAN services, such as radio calibration, need to be processed centrally. However, no WAC is deployed. To ensure high network reliability and performance and meet local computing requirements, a global control role similar to a WAC is required.
- A leader AP is an AP with strong capabilities in an AP group. It is responsible for global calibration of the entire group (automatically elected in a Layer 2 domain).
- The calibration region is automatically divided based on the management VLAN. On a small network, a leader AP can manage all APs (no more than 128 high-performance APs and no more than 50 low-performance APs). In this case, only one management VLAN needs to be planned for APs.
- When the number of APs exceeds the management specifications of the leader AP, you can divide multiple management VLANs to manage the network. You are advised to plan management VLANs based on floors or physical continuous coverage areas to ensure the continuity of the calibration regions.
- Port isolation cannot be configured on the uplink access switch of the cloud AP. (The calibration group is negotiated through wired-side broadcasting.)

Radio Calibration: Calibration Mode Selection



- iMaster NCE supports three radio calibration modes:
  - Automatic mode: APs periodically perform global calibration based on the calibration time and interval.
  - Manual mode: APs do not proactively perform radio calibration, and you need to manually perform global or local calibration for APs at the site on iMaster NCE.
  - Scheduled mode: APs perform global calibration at a scheduled time every day.
- APs perform calibration detection according to the configured mode and switch to other channels to scan neighboring APs. The scanning lasts for 15 minutes.
- During the detection, the APs report the detected data to the leader AP every 10s.
- The leader AP performs computing and calibration every 5 minutes and performs computing for three times to achieve algorithm convergence.
- The leader AP delivers the calibration result to each AP in the group, including the calculated channel and power.

HUAWEI

- During scheduled radio calibration, you can enable intelligent radio calibration and use the analyzer to analyze historical data of the WLAN and predict interference sources on the network. During network optimization, APs can avoid possible interference sources on the network in advance to improve the quality of the entire WLAN.
- During deployment, you are advised to perform manual calibration to automatically plan the channels and power of APs after APs are deployed and go online.

Radio Calibration: Channel and Frequency Bandwidth Selection

AP channel selection

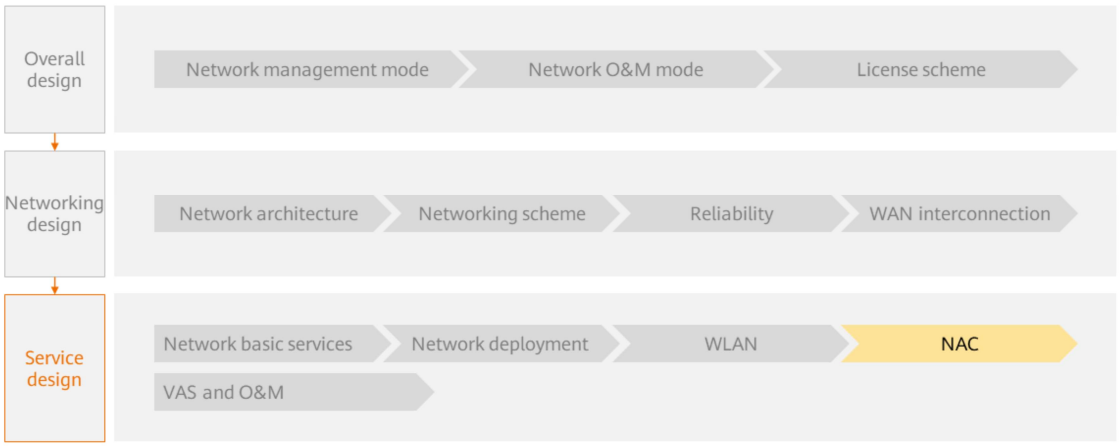
- **2.4 GHz frequency band:** Channel sets 1, 6, and 11 are recommended. If APs are densely deployed, channel sets 1, 5, 9, and 13 are recommended.
- **5 GHz frequency band:** When an AP uses a single 5 GHz radio, it is recommended that high and low frequency channels of neighboring APs be staggered. When an AP uses dual 5 GHz radios, it is recommended that two 5 GHz radios be planned at low and high frequencies respectively.

Frequency bandwidth selection

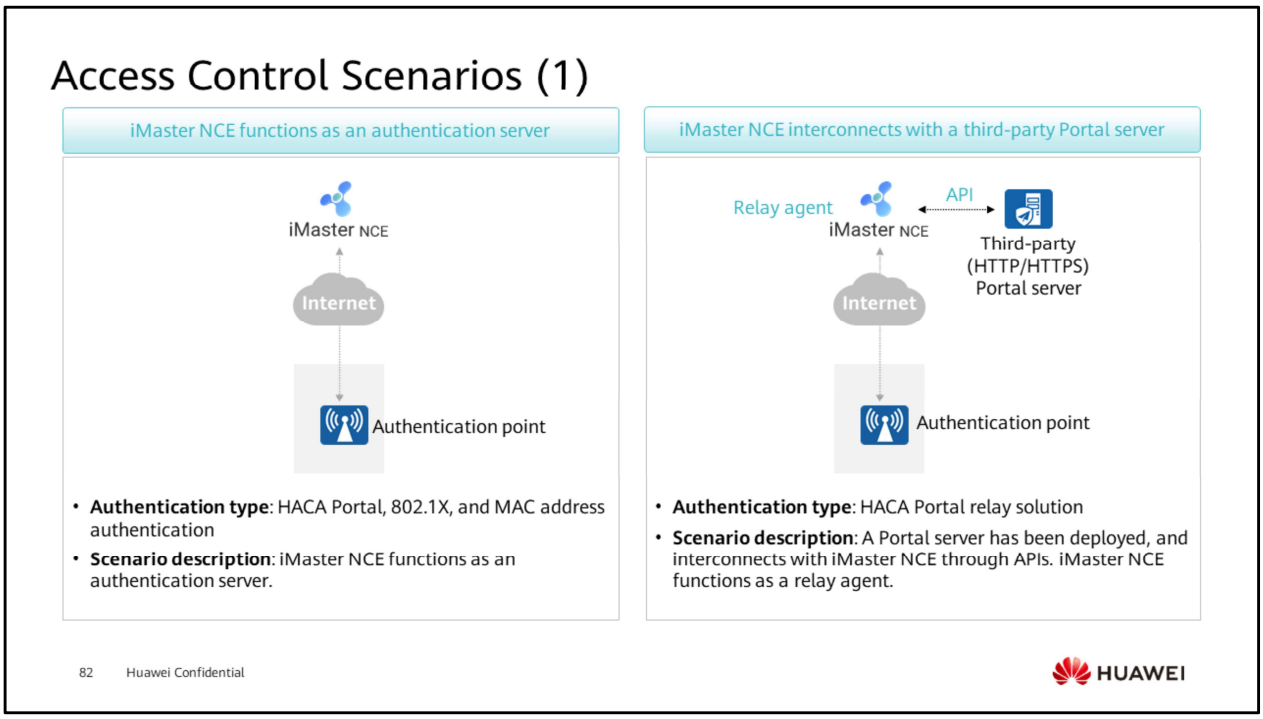
- **2.4 GHz frequency band:** Only the 20 MHz frequency bandwidth can be selected.
- **5 GHz frequency band:** The 40 MHz frequency bandwidth is recommended. The 80 or 160 MHz frequency bandwidth can be used in bandwidth-hungry scenarios.

HUAWEI

Overall Design Process



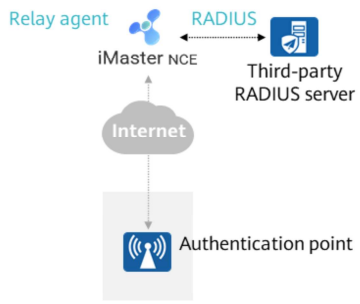
HUAWEI



- Deployment mode recommendation:
  - It is recommended that iMaster NCE functions as an authentication server to implement authentication and policy control.
  - If an enterprise has its own authentication server, the authentication server can interconnect with a third-party server in proxy mode. In addition, iMaster NCE can also implement user management and some policy management functions.
  - If an enterprise uses a third-party authentication server to provide the access authentication function, the device that functions as an authentication point can directly interconnect with the third-party authentication server.

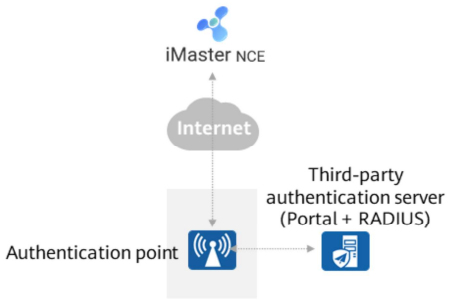
Access Control Scenarios (2)

iMaster NCE interconnects with a third-party RADIUS server



- **Authentication type**: HACA Portal-to-RADIUS solution
- **Scenario description**: A RADIUS server has been deployed, and interconnects with iMaster NCE through RADIUS. iMaster NCE functions as a RADIUS relay agent.

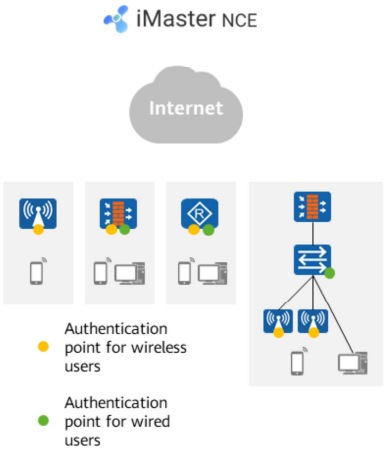
The authentication point directly interconnects with a third-party authentication server



- **Authentication type**: Portal, HTTPS Portal, 802.1X, and MAC address authentication
- **Scenario description**: The network device that functions as an authentication point directly interconnects with a third-party server. (Note: HTTPS Portal is configured using CLIs.)

HUAWEI

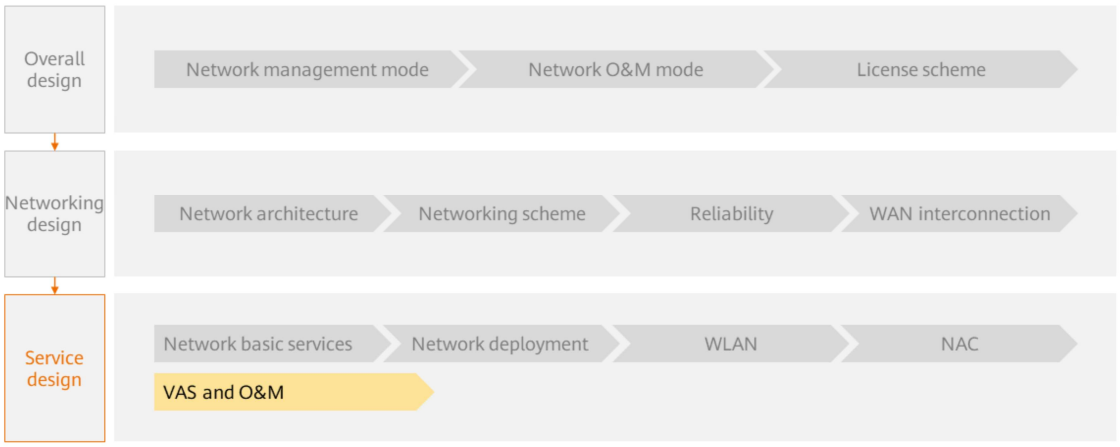
Authentication Solution Deployment



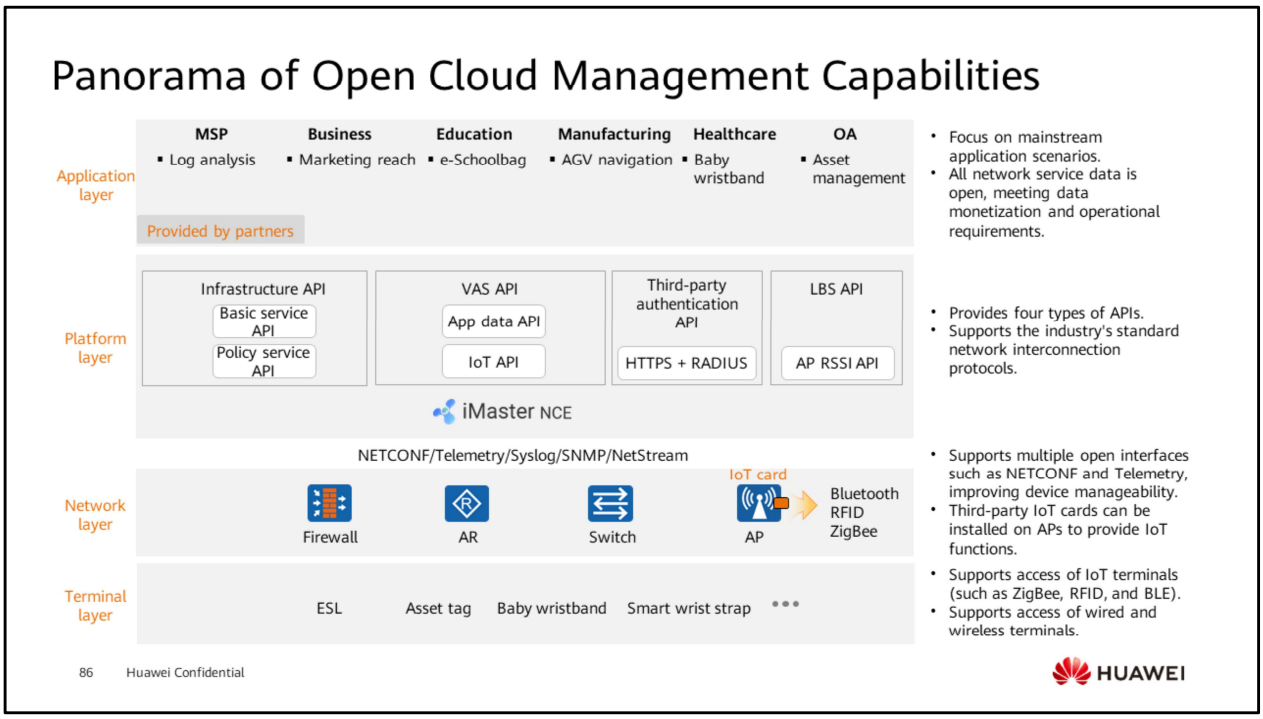
Terminal Role	Guest	Enterprise employee	Dumb terminal
Access Mode	Wireless	Wireless/Wired	Wired
Authentication Type	Portal authentication	• Portal authentication (recommended for wireless users) • 802.1X authentication (optional for both wireless and wired users)	MAC address authentication
Authentication Point	Access devices are recommended as authentication points. This has the following advantages: • Multiple access devices perform user authentication separately, reducing the authentication load. • Authentication points are closer to terminals, improving authentication security. • The configuration planning is simple. If authentication points are deployed at the upper layer, the following factors must be considered: performance specifications of the devices acting as authentication points, Layer 2 isolation at the access layer, and configuration for transparent transmission of 802.1X protocol packets at the access layer.		
Networking Scenario	Single-device networking: The authentication point is the local device, and Portal authentication is recommended. (In the cloud management scenario, ARs and firewalls do not support 802.1X or MAC address authentication.) Other networking scenarios: • APs are used as authentication points for wireless users. • Access switches are used as authentication points for wired users.		

HUAWEI

Overall Design Process

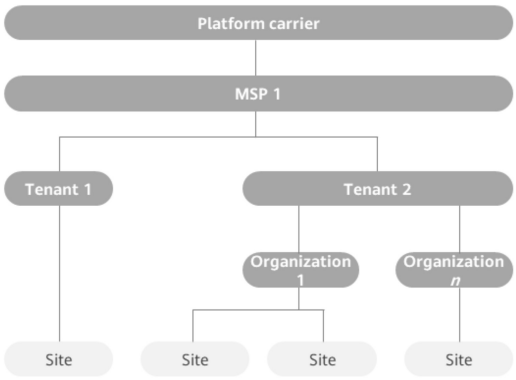


HUAWEI



- Open capabilities at all layers, for all services, and in all scenarios help MSPs, VAS application partners, and customers quickly implement system interconnection, service convergence, and data monetization.

Network Management Design: Management Level



Key points of network management design:

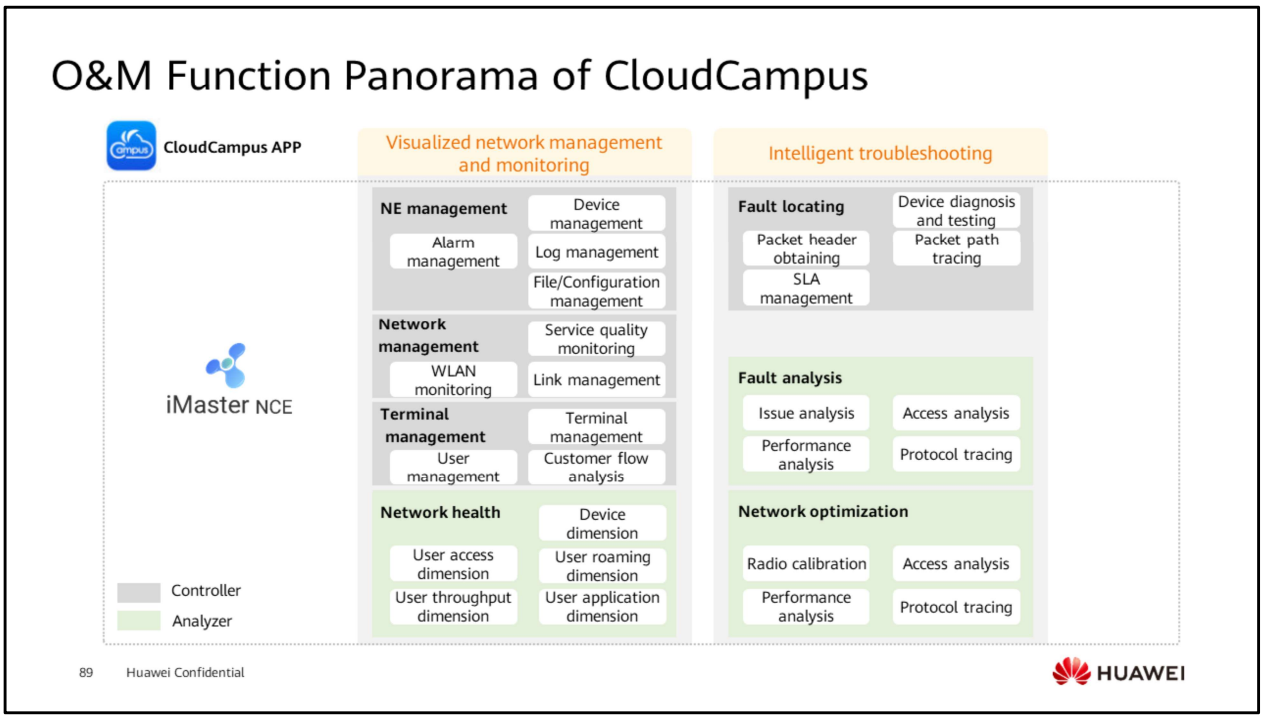
- Management level design is performed to determine whether multiple or multi-level organizations are needed to manage different sites. Huawei's Cloud-Managed Network Solution supports the design of multiple or multi-level organizations for a tenant, which meets the requirements of managing large branches. Each organization can manage multiple sites or sub-organizations.
- **(Optional) Organization planning and design**: A large organization with multiple branches needs to be managed by area. That is, the organization is divided into multiple areas to manage the branches on demand.
- **Site planning**: A network with independent network management services is managed as a site. It can be an independent campus/branch network, or a relatively independent network on a campus/branch network, for example, a network of a building or even a floor. Sites can be flexibly planned based on actual management requirements.
- Rights- and domain-based management of network administrators' rights is adopted.

HUAWEI

Network Management Design: Rights- and Domain-Based Management

Authorization Scope	Tenant Level	MSP Level
Rights-based: controls the functions that can be performed	• The super administrator can grant different rights to tenant administrators. • To facilitate rights assignment and management, the system presets three roles with different rights: Monitor, Open API Operator, and Tenant Administrator. • Tenant administrators authorize MSP administrators based on the preceding three roles.	• The super administrator can assign different rights to MSP administrators. • Similarly, there are three types of roles: Monitor, Open API Operator, and MSP Administrator
Domain-based: controls the range of devices that an administrator can manage	• The minimum granularity of domain-based management is site. That is, tenant administrators perform authorization by site. Different tenant administrators can manage devices at different sites. • When a tenant administrator authorizes an MSP administrator to manage its network, domain-based management is not supported. Instead, the MSP administrator can manage all sites of the tenant.	Domain-based management cannot be configured for MSPs at different levels. Once a tenant authorizes an MSP administrator, the MSP administrator can manage all devices of the tenant whatever level the MSP administrator is at.

HUAWEI



- Similar to the network management mode, the network O&M mode is classified into the following two types based on the O&M operators:
  - Self-O&M: Enterprises perform network O&M by themselves.
  - O&M by MSPs: MSPs are responsible for network O&M. After a tenant submits an application to an MSP, the MSP can directly maintain services of this tenant. In this case, the MSP uses an authorized tenant account to log in to iMaster NCE for network O&M.

Quiz

1. (Multiple-answer question) Which of the following device deployment solutions can be used on small- and medium-sized campus networks to ensure device reliability? ( )
   A. Deploy firewalls in hot standby mode.
   B. Deploy two ARs as egress gateways.
   C. Deploy switches in stacking mode.
   D. Deploy multi-chassis link aggregation.
2. (True or false) 802.1X authentication is recommended for Wi-Fi networks deployed in stores. ( )
   A. True
   B. False

HUAWEI

1. ABC. Multi-chassis link aggregation is a link reliability mechanism.
2. B. Portal authentication is recommended.

Summary

- This course describes how to use Huawei's CloudCampus Cloud-Managed Network Solution to design small- and medium-sized campus networks. Key design points include the overall design (network management mode design, network O&M mode design, and license scheme design), networking design (network architecture design, networking solution design, WAN interconnection design, and reliability design), and service design (network basic service design, network deployment design, WLAN design, NAC design, security design, QoS design, VAS design, as well as network O&M design).

HUAWEI

Recommendations

- Huawei's CloudCampus Solution V100R020C00 Product Documentation
  - https://support.huawei.com/hedex/hdx.do?docid=EDOC1100172638&lang=en
- Design Guide of Huawei CloudCampus Solution for Small and Medium Sized Campus Networks Scenario
  - https://e.huawei.com/en/material/networking/campusnetwork/5133b49714a04ab08d5851d0e44e59a1

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Key Technologies for Enterprise WAN Interconnection and Their Applications



HUAWEI

Foreword

- Wide area networks (WANs) have undergone a long history, and the technologies used on WANs have been updated many times. Enterprises have more diversified requirements over WANs than campus networks. In addition, the requirements for network connection mode, reliability, and security vary with enterprises.
- To meet the interconnection requirements of different enterprises, a variety of WAN interconnection technologies are developed.
- This course will focus on the WAN interconnection technologies and their application scenarios.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the basic architecture of a typical WAN interconnection solution.
  - Describe WAN interconnection networking technologies and their application scenarios.
  - Describe WAN interconnection reliability technologies and their application scenarios.
  - Describe WAN interconnection optimization technologies and their application scenarios.
  - Describe WAN interconnection security technologies and their application scenarios.

HUAWEI

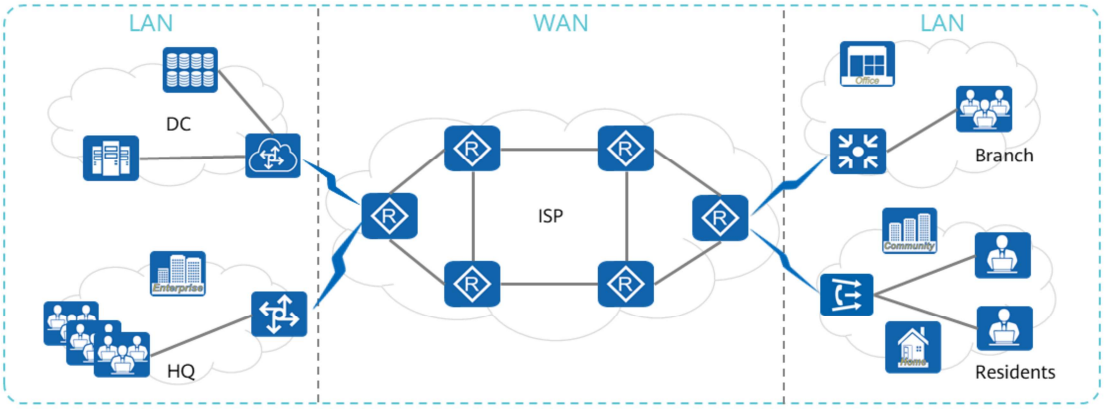
Contents

1. **Basic Architecture of a Typical WAN Interconnection Solution**
2. Networking Technologies and Their Applications of WAN Interconnection
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

What Is a WAN?

- WANs provide interconnection services across regions, cities, and countries. A WAN usually spans a long distance (dozens or even thousands of kilometers). To meet long-distance transmission requirements on a WAN, optical fibers are often used as the interconnection media.



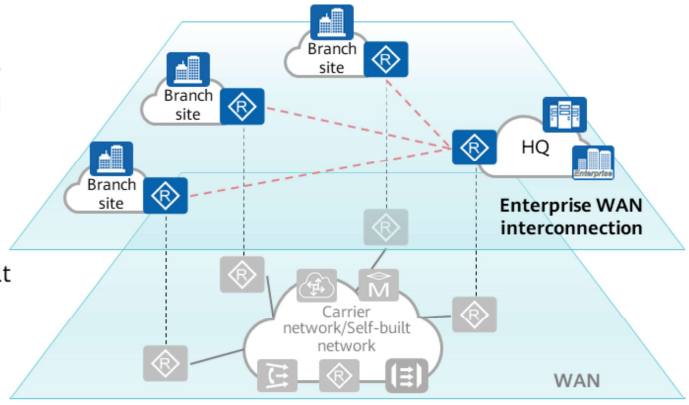
HUAWEI

- LAN
  - A local area network (LAN) is a computer network that connects computers, peripheral devices, databases, and other devices in a limited geographical area (such as a campus, factory, or organization) within thousands of meters.
- WAN
  - WANs provide wider coverage than LANs and metropolitan area networks (MANs). The communication subnet of a WAN mainly uses the packet switching technology. The communication subnet of a WAN can use the public packet switching network, satellite communication network, and wireless packet switching network to interconnect the LANs or computer systems in different areas for resource sharing.
  - The Internet is the largest WAN in the world.
- Relationship between the LAN and WAN:
  - A LAN is located in an area, whereas a WAN spans a larger area. For example, the headquarters of a large company is located in Beijing, and its branches are distributed all over the country. If the company connects all its branches through a network, a branch is a LAN, and the company network is a WAN.

- Typical WAN rates range from 56 kbit/s to 155 Mbit/s. Currently, 622 Mbit/s, 2.4 Gbit/s, and even higher rates are available. The transmission delay ranges from several milliseconds to hundreds of milliseconds (when satellite channels are used).

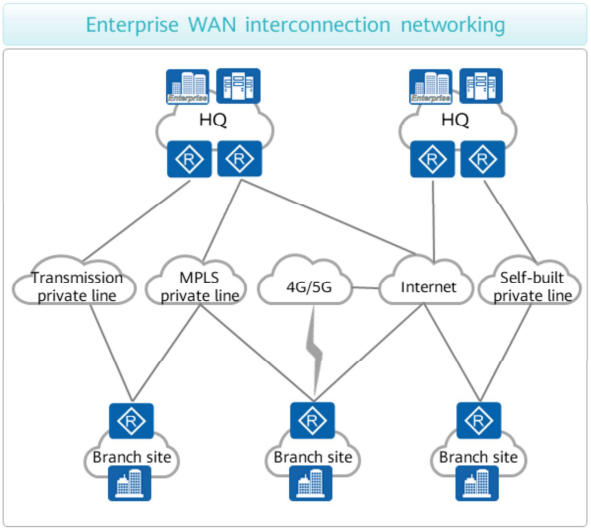
WAN and Enterprise WAN Interconnection

- Enterprise WAN interconnection refers to the interconnection between nodes at different levels, such as the headquarters (HQ), data centers (DCs), branches, fixed offices, and mobile offices within enterprises.
- Generally, enterprise WAN interconnection depends on a carrier-built or self-built WAN.



Major Enterprise WAN Interconnection Modes

- Generally, enterprise WANs can be interconnected in the following modes:
  - Carriers' transmission or MPLS private lines are used to connect regional networks. This mode is expensive and is applicable to enterprises with high SLA requirements.
  - The carrier Internet + VPN technology is used for connection. This mode is applicable to small- and medium-sized branches that do not have high SLA requirements.
  - Carriers' point-to-point (P2P) private lines are used to implement cross-city or cross-border connections. This mode is mainly used for connections between DCs, headquarters, or important branches, and is expensive.
  - Industries such as electric power and transportation build network connections through self-built private lines.
- Usually, enterprise WANs are interconnected using a combination of the preceding modes.

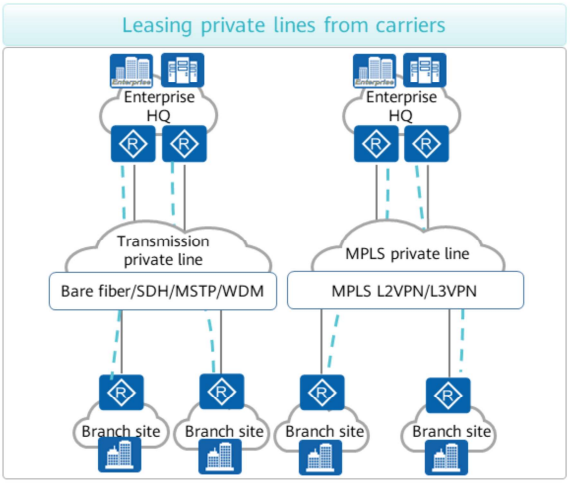


HUAWEI

- SLA is an agreement between the network service provider and the customer. It defines terms such as service type, service quality, and customer payment.
- According to the SLA requirements, the service provider uses multiple technologies and solutions to monitor and manage the network performance and traffic so as to meet the requirements defined in the SLA and generate reports about customers' network performance.

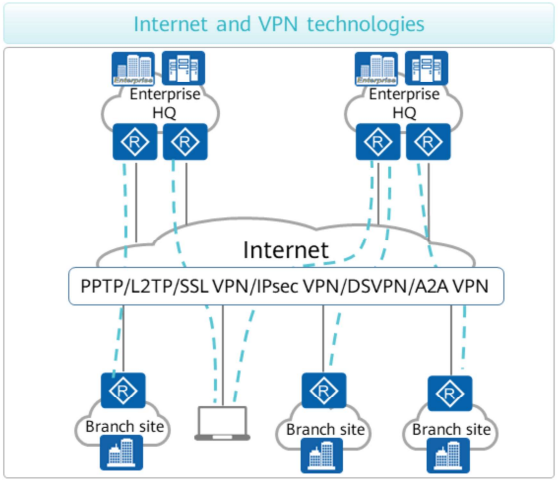
Enterprise WAN Interconnection Technologies: Carriers' Private Lines

- To ensure network reliability and security, enterprises lease transmission or MPLS private lines from carriers when constructing enterprise WANs.
  - Transmission private lines are expensive, but data is carried on dedicated lines, ensuring service quality and security.
  - MPLS private lines are cheaper than transmission private lines and can ensure service security. However, service reliability is not as good as that of transmission private lines.
  - A small number of enterprises (such as transportation and electric power enterprises) can deploy optical fibers and build their own backbone networks.



Enterprise WAN Interconnection Technologies: Internet and VPN

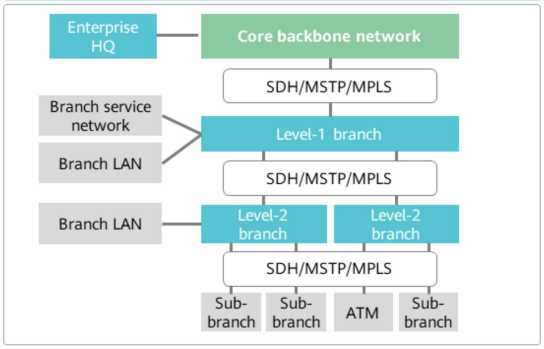
- With the development of the Internet, some enterprise services can be carried over the Internet.
- The Internet has security risks, so VPN technology is used to provide secure and reliable connections.
- Virtual Private Dial-up Network (VPDN) technologies, such as Point-to-Point Tunneling Protocol (PPTP), Layer 2 Tunneling Protocol (L2TP), and Point-to-Point Protocol over Ethernet (PPPoE), allow terminal users or branches to dial up to the carrier network or HQ network.
- Typically, Internet Protocol security virtual private network (IPsec VPN) technology is used to build network connections between enterprise branches or between enterprise branches and the HQ.
- To simplify IPsec configuration on large-scale networks, technologies such as Dynamic Smart VPN (DSVPN) and Any to Any (A2A) VPN have been developed and widely used.



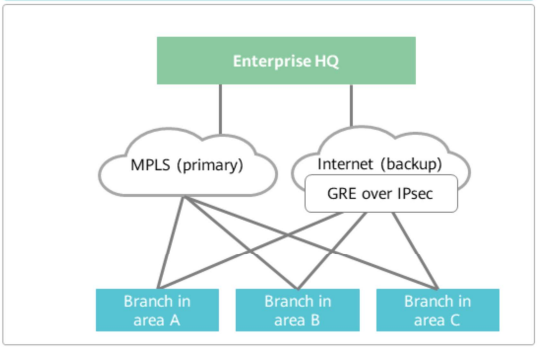
Common Application Scenarios of Enterprise WAN Interconnection Networking

- Enterprise WAN interconnection needs to be deployed based on enterprise requirements. For example, in the financial service industry, most enterprises lease transmission or MPLS private lines to guarantee network reliability and security. Considering network costs, other enterprises usually lease MPLS private lines as primary lines and Internet + VPN lines as backup lines.

WAN interconnection in the financial service industry



WAN interconnection for a wine enterprise



HUAWEI

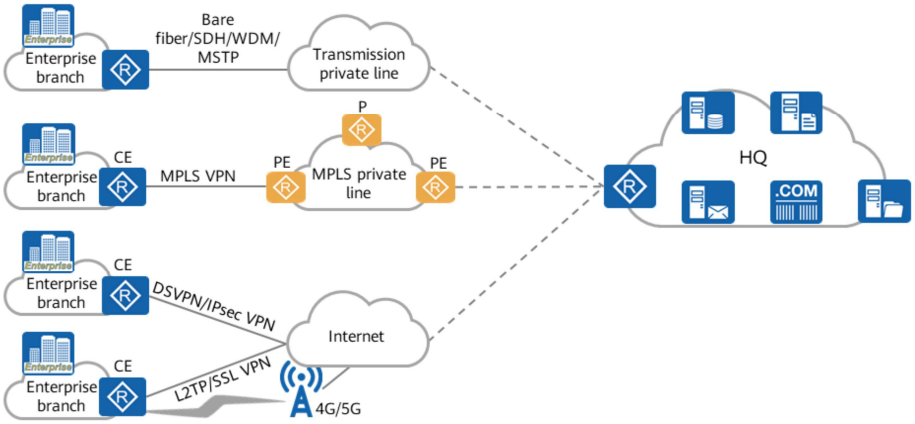
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. **Networking Technologies and Their Applications of WAN Interconnection**
   - WAN Interconnection Networking Solution
     - Private Line Technologies and Their Applications
     - VPN Technologies and Their Applications
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

Enterprise WAN Interconnection Networking Solution

- There are many interconnection modes for enterprise WANs. Generally, one or more interconnection modes are used based on different enterprise requirements.



HUAWEI

Comparison Between Private Line and VPN Technologies

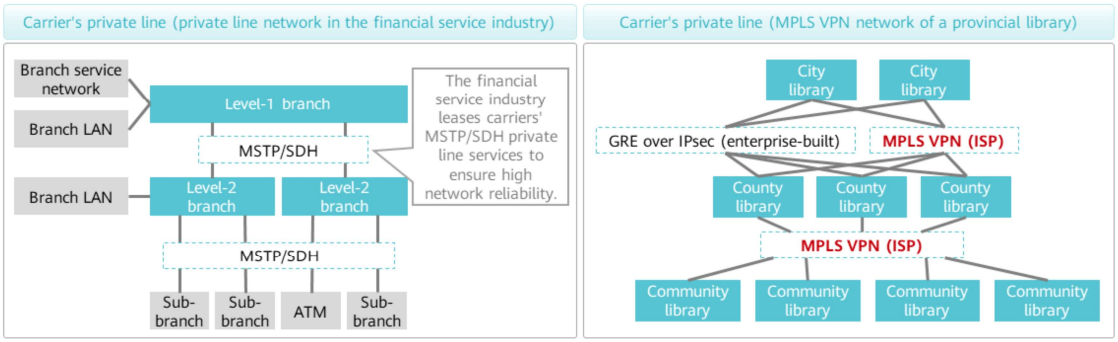
- Private lines were introduced very early. They can meet interconnection requirements of enterprises and ensure high network reliability and security. However, private lines are expensive.
- As networks develop, VPN technologies start to occupy more market shares. However, some industries demanding high security and reliability, such as the financial service industry, still prefer private line technologies.
- Selecting private line or VPN technologies depend on enterprises' services. The following table compares the two technologies.

Item	Private Line Technology	VPN Technology
Security	Relatively high: depending on ISPs	Very high: Data is encrypted before being transmitted and security control is in the hands of users.
Reliability	High: depending on ISP network reliability	Comparatively high: depending on reliability of Internet lines
Scalability	Medium: depending on ISPs	Based on TCP/IP technology, the access mode is flexible. A network can be scaled out as long as it is reachable.
Investment cost	The private line expense is very high and needs to be paid every month. In addition, the device expense needs to be invested in the initial period of network construction.	One-off device expense is invested, so there is no need to pay monthly operating expenses.
Support for mobile users	Mobile users can only connect to the network to which private lines are reachable, and internal mobile users leaving a LAN cannot access a private network.	Internal mobile users can use the Internet for secure access, eliminating geographical differences.
Transmission bandwidth	Leased bandwidth is low because of the high price.	The Internet is cheap, and the leased bandwidth is high.

HUAWEI

Overview of Carriers' Private Lines

- Carriers have a large number of line resources and launch different private line services based on different industries and scenarios.
- Carriers' high-quality transmission private line services mainly include SDH, MSTP, and bare fiber services, which are expensive but deliver excellent performance.
- MPLS VPN is another type of private line services provided by carriers. MPLS VPN private lines provide slightly lower performance than transmission private lines but are less expensive.

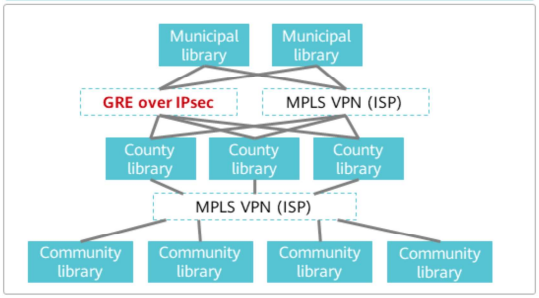


HUAWEI

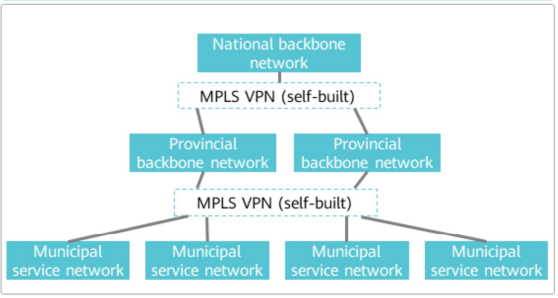
Overview of Enterprise-Built Private Line and VPN

- Enterprises can establish VPNs, such as SSL VPN, DSVPN, and IPsec VPN, through carriers' networks.
- Some large enterprises can lay out optical fibers by themselves and set up private lines. However, only few enterprises can lay out optical fibers by themselves.
- VPNs built by enterprises are more and more widely used because they are cost-effective, easy to expand, and controllable.

Enterprise-built VPN (VPN of a provincial library)



Enterprise-built private line (MPLS VPN network of the energy industry)



HUAWEI

Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. **Networking Technologies and Their Applications of WAN Interconnection**
   - WAN Interconnection Networking Solution
   - Private Line Technologies and Their Applications
   - VPN Technologies and Their Applications
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

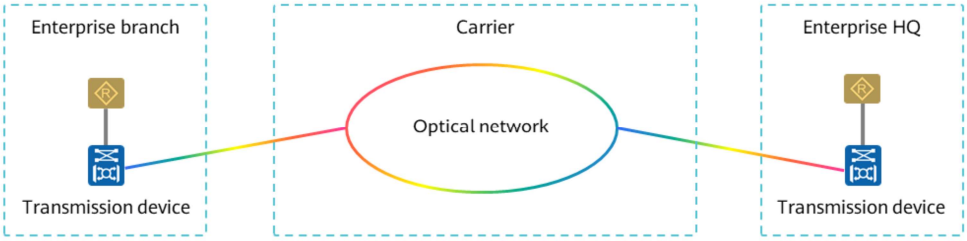
Overview of Private Line Technologies

- Private line technologies were introduced very early. As networks develop, many private line technologies, such as frame relay (FR) and ATM, are no longer used. Currently, the following private line technologies are widely used:
  - Bare fiber: Carriers provide bare fibers along which no intermediate device is deployed. Therefore, bare fibers are expensive.
  - SDH/MSTP/WDM: Transmission private lines use transmission devices to build hard pipes over optical fibers, ensuring good performance. The price of such private lines is lower than that of bare fibers.
  - MPLS VPN: MPLS private lines use Ethernet for network access and do not have hard pipes. The performance of MPLS VPN is poorer than that of transmission private lines, but MPLS VPN is the cheapest among all types of private lines.

HUAWEI

Overview of Bare Fibers

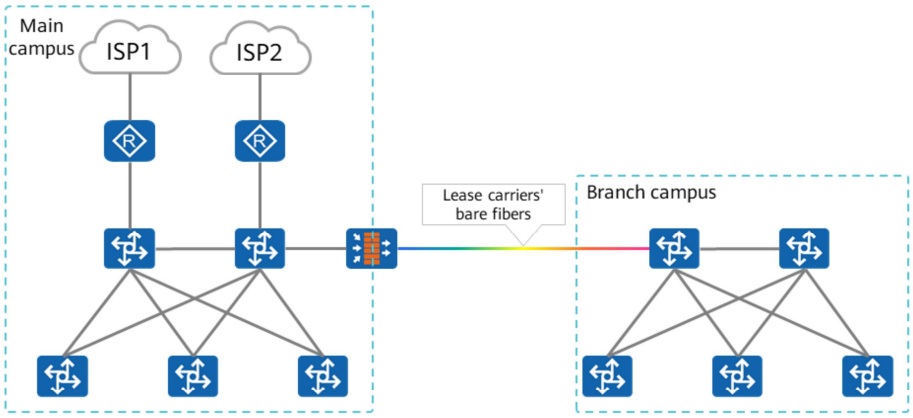
- A carrier provides a bare fiber line along with no intermediate device is deployed. The network capacity depends on the enterprise devices at both ends of the bare fiber.
- Bare fibers are charged based on the distance. A longer distance indicates a higher cost. Generally, the maximum transmission distance of a hop of an optical fiber is 300 km. If the distance between two sites exceeds 300 km, a regeneration device needs to be deployed.



HUAWEI

Exemplary Application Scenario of Bare Fibers

Application scenario of bare fibers

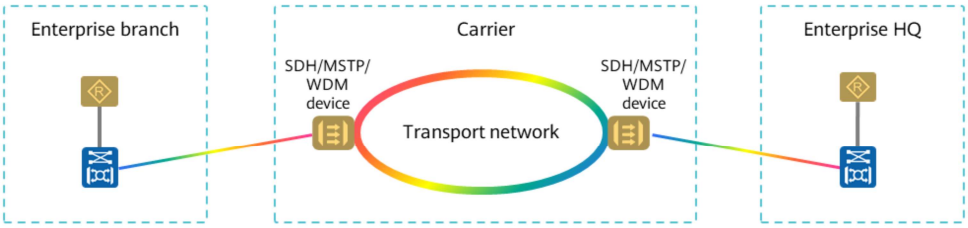


- Carriers' bare fibers can be leased to build network connections between the main campus and branch campus in the same city. This practice simplifies network management and access authentication management.

HUAWEI

Overview of SDH/MSTP/WDM Private Lines

- Enterprises that require long-distance transmission and high network reliability and security can lease SDH/MSTP/WDM private lines.
- This type of private line is a transmission private line. Tenants exclusively occupy part of the bandwidth of the transmission private line. Because multiple users share the transmission private line, its price is lower than that of bare fibers. Although transmission private lines are shared by tenants, they exclusively occupy bandwidth and use hard pipes. Therefore, they deliver high network reliability and security.
- MSTP and WDM private lines are widely used on the live network, and SDH private lines are still used in a few areas.

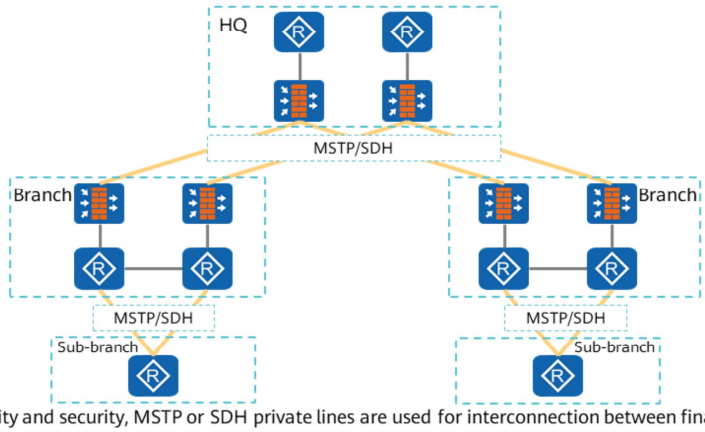


HUAWEI

- SDH is a TDM system and a traditional circuit scheduling mode.
- The multi-service transport platform (MSTP) receives, processes, and transmits TDM, ATM, and Ethernet services.
- WDM uses multiple lasers to transmit multiple beams of lasers with different wavelengths over a single optical fiber. The transmission bandwidth of WDM devices is high, and the live-network bandwidth can reach up to 8 Tbit/s.

Exemplary Application Scenario of SDH/MSTP/WDM Private Lines

Application scenario of SDH/MSTP/WDM private lines

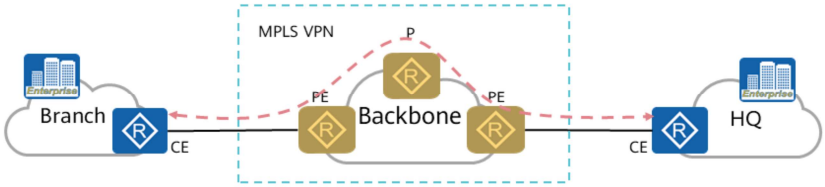


- To ensure high reliability and security, MSTP or SDH private lines are used for interconnection between financial service branches.

HUAWEI

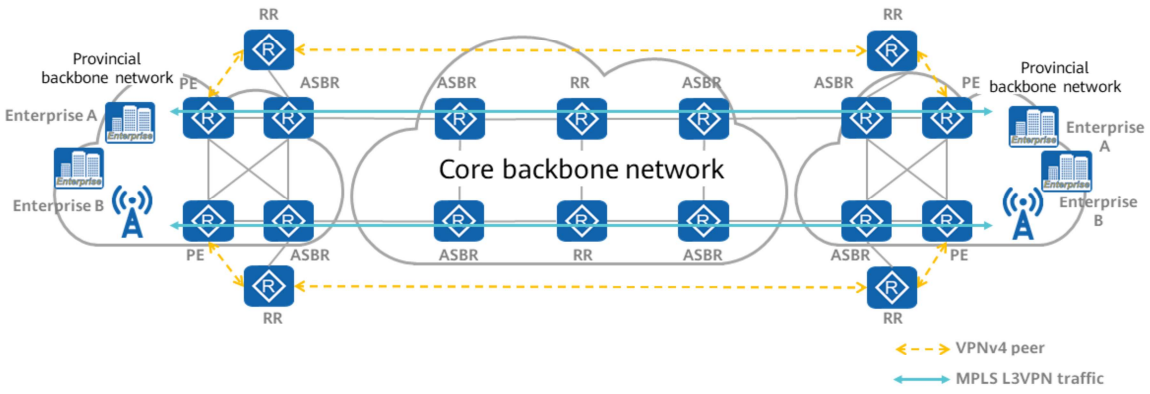
MPLS VPN Private Line

- MPLS VPN technology is widely used in enterprise interconnection scenarios. MPLS L2VPN or MPLS L3VPN can be deployed based on enterprise requirements. MPLS VPN makes a compromise between the cost and performance, so it is very popular.
- For enterprises that can build their own WANs, such as railways and electric power companies, MPLS VPN is an easy-to-manage and low-cost VPN technology. For enterprises that cannot build their own WANs, MPLS VPN is expensive.
- The enterprises that require high network reliability and security can use the MPLS VPN private line as the primary link and the GRE over IPsec line as the backup link.



Exemplary Application Scenario of MPLS VPN

Application scenario of MPLS VPN



- The MPLS L3VPN solution is widely used on the live network. Some large networks may use inter-AS MPLS L3VPN.

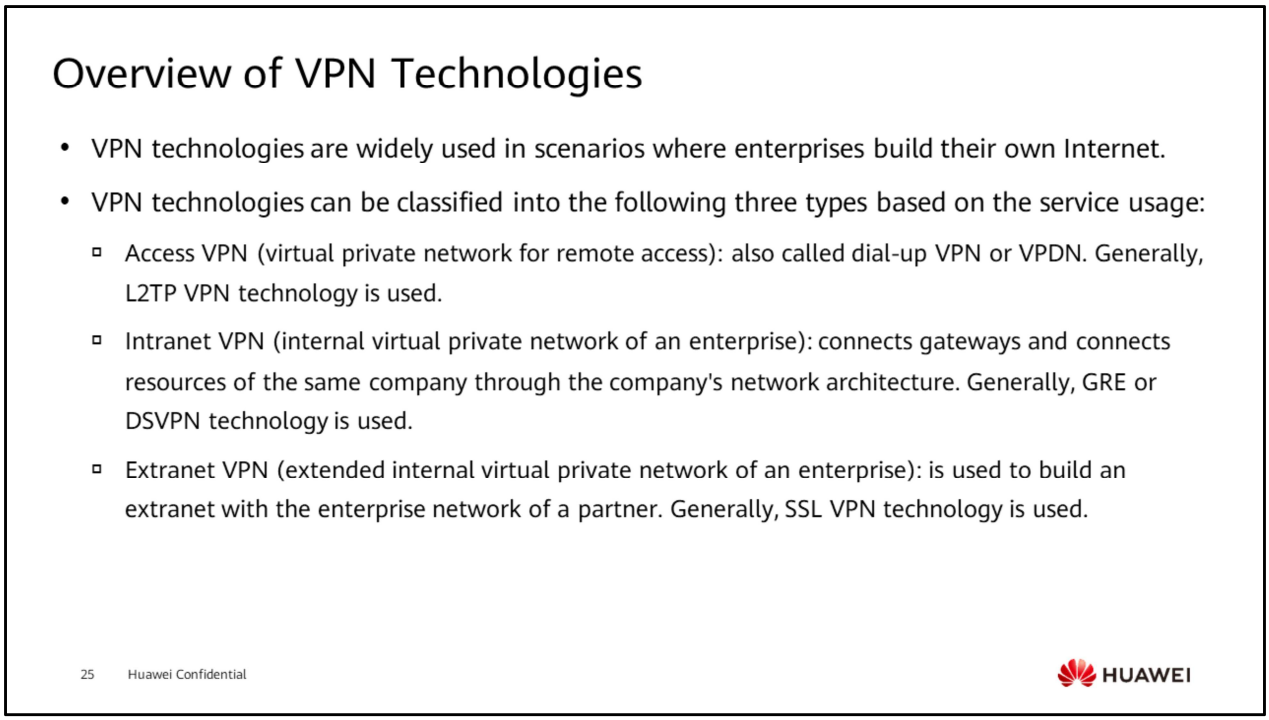
HUAWEI

- There are three types of inter-AS MPLS L3VPN solutions: Option A, Option B, and Option C.
- Option A applies to small inter-AS MPLS L3VPNs. Option B applies to midsize and large inter-AS MPLS L3VPNs. Option C applies to large or super-large inter-AS MPLS L3VPNs.

Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. **Networking Technologies and Their Applications of WAN Interconnection**
   - WAN Interconnection Networking Solution
   - Private Line Technologies and Their Applications
   - VPN Technologies and Their Applications
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

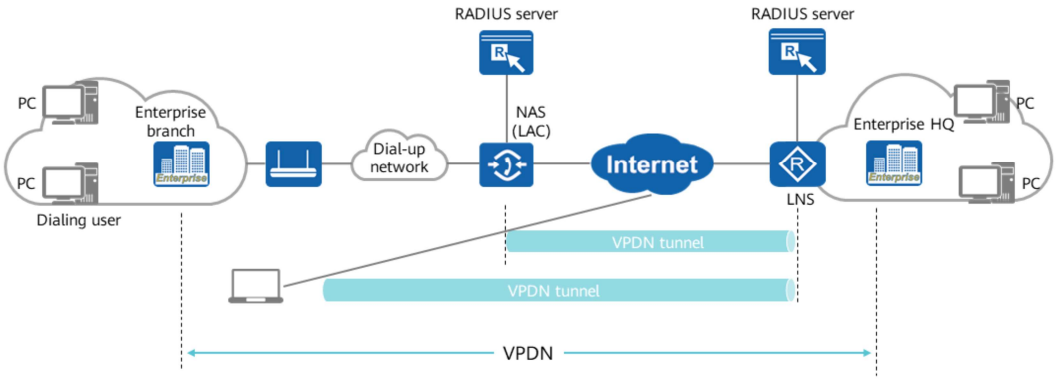
HUAWEI



- By service usage:
  - Access VPN: enables mobile employees, remote office employees, and remote small offices to establish private network connections with enterprise intranet and extranet through a public network. There are two types of access VPN connections: client-initiated and NAS-initiated.
  - Intranet VPN: Intranet VPN is an extension or replacement of traditional private lines or other enterprise networks to connect distribution points within an enterprise through a public network.
  - Extranet VPN: extends enterprise networks to suppliers, partners, and even clients over a public network.
- According to the layers of tunnels in the OSI model:
  - Layer 2 tunneling protocol: encapsulates PPP frames into a tunnel. Layer 2 tunneling protocols include the Point-to-Point Tunneling Protocol (PPTP), Layer 2 Forwarding (L2F), and Layer 2 Tunneling Protocol (L2TP).
  - Layer 3 tunneling protocol: Only Layer 3 packets are carried in a tunnel. Existing Layer 3 tunneling protocols include Generic Routing Encapsulation (GRE) and IPsec. IPsec includes the Authentication Header (AH) protocol and Encapsulating Security Payload (ESP) protocol.

Access VPN Overview

- Access VPN uses virtual private dial-up network (VPDN) technology, and is a type of VPN service based on dial-ups. It can be used for enterprise interconnection or remote access to enterprise networks.



HUAWEI

- VPDN is implemented by using a tunneling technology. That is, data of an enterprise network is encapsulated in a tunnel for transmission. On an interface between the source LAN and public network, the tunneling technology encapsulates data as a payload in a data format that can be transmitted on a public network. On an interface between the destination LAN and the public network, it decapsulates data to extract the payload. The logical path through which encapsulated data packets are transmitted on the Internet is called a tunnel. To ensure that data is encapsulated, transmitted, and decapsulated smoothly, the communication protocol is the core.
- VPDN provides three common tunneling technologies:
  - Point-to-Point Tunneling Protocol (PPTP)
  - Layer 2 Forwarding (L2F)
  - Layer 2 Tunneling Protocol (L2TP)

PPTP Overview

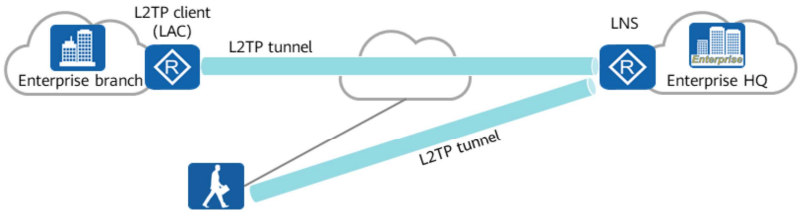
- PPTP is a type of VPN protocol and has a history of more than 20 years. This protocol relies on encryption, authentication, and Point-to-Point Protocol (PPP) for negotiation. It requires only the user name, password, and server address for connection setup.
- PPTP is fast, but has weak encryption. Among all VPN protocols, PPTP has the lowest encryption level and must be based on IP networks.



HUAWEI

L2TP Overview

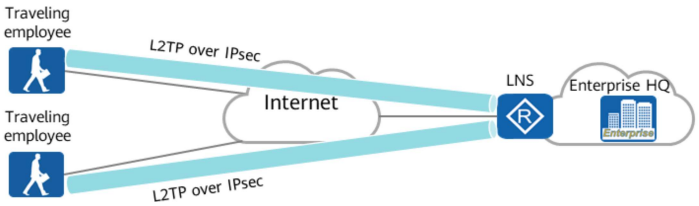
- L2TP, an open standard of IETF, combines advantages of PPTP. L2TP is especially suitable for setting up an access VPN and has become a de facto industry standard.
- L2TP is only a tunneling protocol and does not provide encryption or privacy protection. Therefore, L2TP is usually used together with IPsec.
- L2TP is one of commonly used enterprise interconnection technologies. When L2TP is used, an AAA server is required. L2TP is a good choice for constructing an L2VPN.



Live-Network Application of Access VPN

- Access VPN is mainly used for remote access of intranet users, and L2TP over IPsec is most widely used.
- PPTP requires the support of the Windows operating system. In addition, the IP address of the Windows server on the intranet needs to be mapped through NAT for extranet access. Therefore, PPTP is difficult to deploy and is seldom used.

Live-network application of L2TP over IPsec



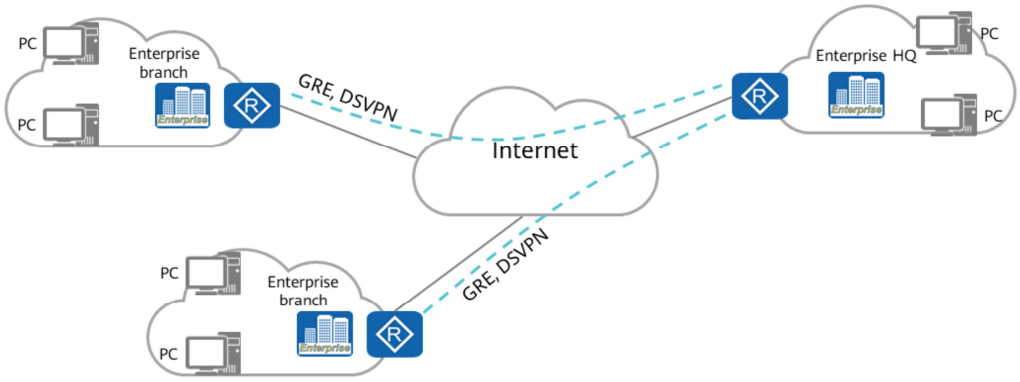
- L2TP connections can also be used between branch sites. However, L2TP cannot transmit multicast data or advertise routes between the HQ and branches. Therefore, L2TP is mainly used for remote user access on the live network.

HUAWEI

Access VPN Overview | Intranet VPN Overview | Extranet VPN Overview

Intranet VPN Overview

- Intranet VPN technology is used to construct a VPN between gateways based on the Internet. GRE and DSVPN technologies are mainly used.
- GRE and DSVPN technologies are used to establish VPNs between enterprise branches and the HQ.

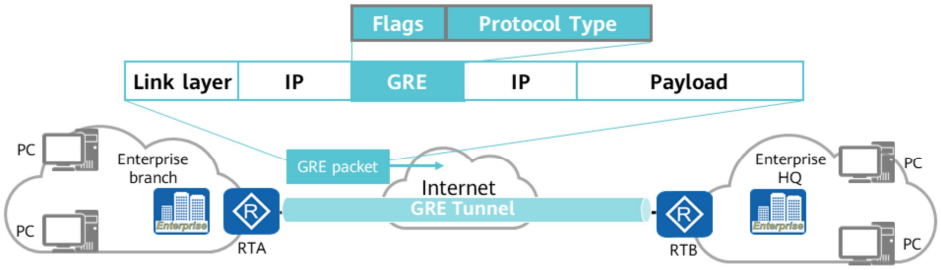


HUAWEI

- Intranet VPN mainly uses the following technologies:
  - GRE
  - GRE over IPsec
  - DSVPN
  - DSVPN IPsec

GRE Overview

- Generic Routing Encapsulation (GRE) is used to encapsulate packets of some network-layer protocols (such as IP, IPX, and AppleTalk) so that the encapsulated packets can be transmitted over the network on which another network-layer protocol is applied.
- GRE is typically used on networks with a few branch sites.



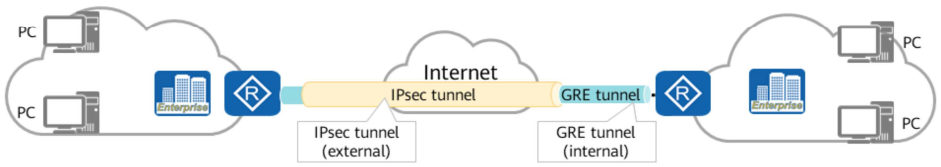
HUAWEI

- GRE is a Layer 3 tunneling technology. A GRE tunnel is a virtual P2P connection that transmits encapsulated data packets.
- The two ends of a GRE tunnel are tunnel interfaces which encapsulate and decapsulate data packets. The tunnel interface that sends encapsulated packets is called the tunnel source interface, and the one that receives these packets on the peer end is called the tunnel destination interface.
- The packet encapsulation process in the figure is as follows:
  - After receiving an IP packet, RTA's interface that connects to the enterprise branch sends a packet to the IP protocol module.
  - The IP protocol module checks the destination address in the packet header to determine how to forward this packet. If the packet is destined for the other end of the GRE tunnel, the IP protocol module sends the packet to the tunnel interface.
  - After receiving the packet, the tunnel interface encapsulates the packet using GRE and delivers the packet to the IP protocol module.

    - The IP protocol module encapsulates the GRE packet using a new IP packet header. The source address is the address of the tunnel source interface, and the destination address is the address of the tunnel destination interface. Then the IP protocol module forwards the encapsulated IP packet from the WAN interface (tunnel source interface) based on the destination address and routing table.
- As the reverse of encapsulation, the decapsulation process is as follows:
    - RTB receives an IP packet from its physical interface connected to the Internet and checks the destination address. If the destination is RTB and the protocol ID in the IP packet header is 47 (indicating GRE packet), RTB removes the IP packet header and sends the packet to the GRE module.
    - The GRE module verifies the checksum and key fields, removes the GRE header, and sends the packet to the IP protocol module.
    - The IP protocol module forwards the packet to the enterprise HQ.

GRE over IPsec Overview

- GRE does not support encryption. IPsec supports only the IP protocol and does not support multicast.
- GRE over IPsec combines advantages of both GRE and IPsec, and offsets their disadvantages.
- GRE over IPsec is a point-to-point VPN technology commonly used by enterprises.

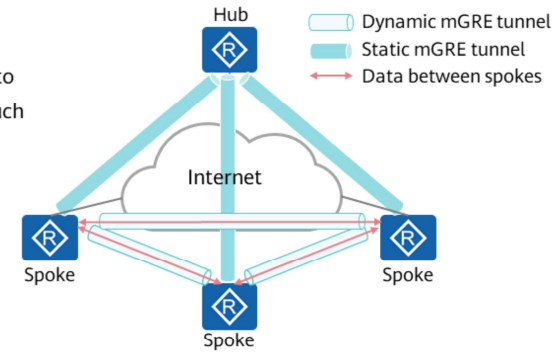


HUAWEI

- GRE encapsulates multicast data to allow data to be transmitted through GRE tunnels. Currently, IPsec can encrypt only unicast data. If multicast data, such as routing protocol, voice, and video data, needs to be transmitted over IPsec tunnels, a GRE tunnel can be established to encapsulate multicast data, and then IPsec encrypts the encapsulated packets. In this way, multicast data is encrypted and transmitted in the IPsec tunnel.
- GRE over IPsec combines advantages of both GRE and IPsec. It enables a network to support multiple upper-layer protocols and multicast packets, as well as packet encryption, identity authentication, and data integrity check.
- GRE over IPsec encapsulates packets using GRE, and then IPsec.
- GRE over IPsec supports the following encapsulation modes:
  - Tunnel mode
  - Transport mode

DSVPN Overview

- DSVPN overcomes defects of GRE over IPsec and enables enterprises with a large number of branches to easily build VPNs.
- DSVPN is a technology that dynamically establishes GRE tunnels. It uses the Next Hop Resolution Protocol (NHRP) to dynamically collect, maintain, and advertise information such as the public IP address of each spoke, allowing the source branch to obtain the public IP address of the destination branch.
- DSVPN uses mGRE technology to enable VPN tunnels to transmit multicast and broadcast packets, and a tunnel interface can establish VPN tunnels with multiple peers.
- The GRE tunnel established by DSVPN can still use IPsec technology to ensure tunnel security.



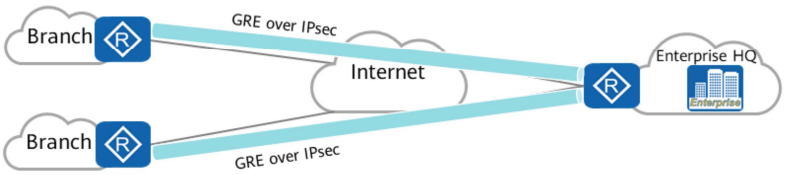
HUAWEI

- DSVPN resolves the following defects of GRE over IPsec:
  - All traffic must pass through the hub.
  - The hub configuration needs to be modified when a site is added.
  - If spokes use dynamic addresses, problems may occur when P2P GRE is deployed.

Live-Network Application of Intranet VPN

- On the live network, intranet VPN is mainly used for interconnection between enterprise branches and the HQ or between branches.
- GRE over IPsec is widely used on the live network. For enterprises with many branches, Efficient VPN can be used to simplify the branch configuration. IPsec link redundancy can be deployed to ensure GRE reliability.

Live-network application of GRE over IPsec

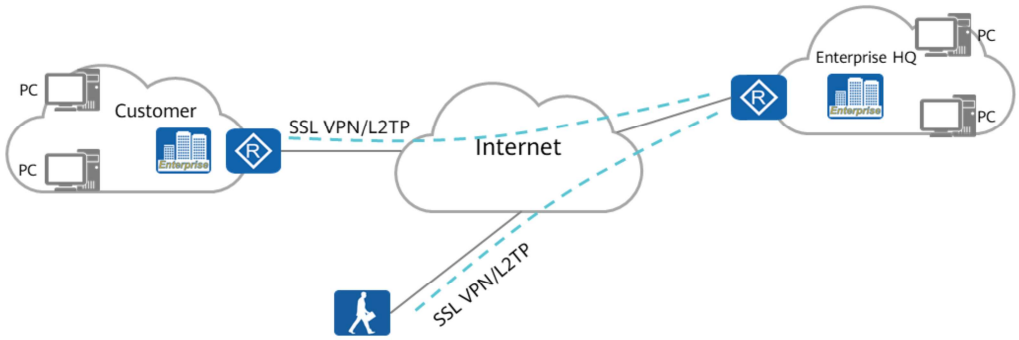


- DSVPN is seldom used on the live network due to the following reasons: 1. Many enterprises expect that traffic between branches can pass through the HQ to facilitate management. 2. Enterprise O&M personnel just have a basic understanding of DSVPN, which is inconvenient for O&M.

HUAWEI

Extranet VPN Overview

- Extranet VPN is mainly used to deliver secure network access between customers or suppliers. Traveling employees can also use extranet VPN to access the enterprise network.
- Extranet VPN mainly uses SSL VPN and L2TP.

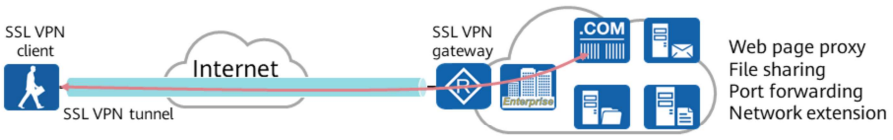


HUAWEI

Access VPN Overview | Intranet VPN Overview | Extranet VPN Overview

SSL VPN Overview

- SSL VPN is mainly used to enable traveling employees to remotely access the enterprise intranet, which is an extension of the enterprise intranet on the WAN.
- SSL VPN authenticates and controls users based on HTTP. Users do not need to configure SSL VPN, which is easy to use.



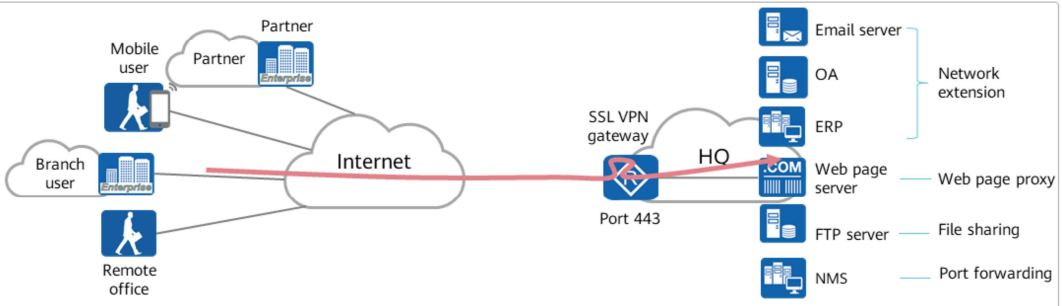
HUAWEI

- SSL VPN is a VPN remote access technology based on SSL. Mobile users (referred to as remote users in SSL VPN) can use SSL VPN to securely and conveniently access enterprise intranets and intranet resources, improving work efficiency.
- Before SSL VPN is developed, VPN technologies such as IPsec and L2TP are used to enable remote user access. However, these VPN technologies have the following disadvantages:
  - Remote users need to install specific client software on their terminals, leading to difficult network deployment and maintenance.
  - The IPsec or L2TP VPN configuration is complex.
  - Network management personnel cannot perform fine-tuned control over the remote users' access permission to enterprise intranet resources.

Exemplary Application Scenario of Extranet VPN

- It is convenient for external users to access the intranet through SSL VPN. They can directly access the intranet after passing web page authentication, without having to install a client. In addition, there are many SSL VPN service options, such as web page proxy, file sharing, port forwarding, and network extension.

Application scenario of SSL VPN



- SSL VPN is easy to use. You can use a browser to access the enterprise intranet. However, not all routers support SSL VPN. For routers that do not support SSL VPN, you can use L2TP to establish connections between the extranet and intranet.

HUAWEI

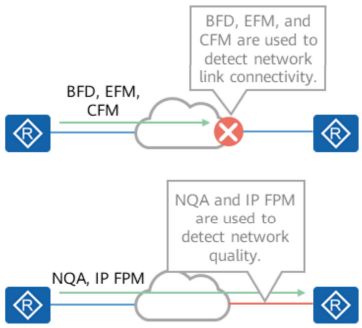
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. **Reliability Technologies and Their Applications of WAN Interconnection**
   - Link Detection Technologies and Their Applications
   - Network and Service Reliability Technologies and Their Applications
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

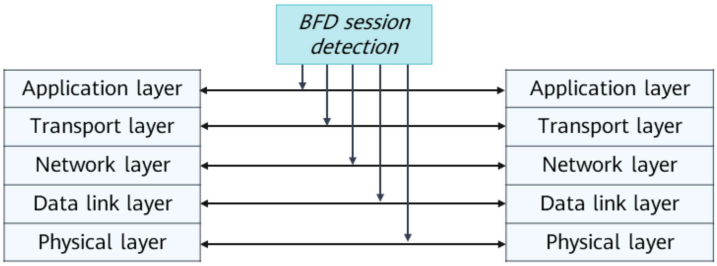
Overview of Link Detection

- Fluctuation of network link quality affects service quality. How to quickly detect the link quality fluctuation is the first step in improving link quality.
- There are many protocols and technologies for detecting link quality, which are classified into two types:
  - Link connectivity detection technologies:
    - Bidirectional Forwarding Detection (BFD)
    - Ethernet in the First Mile (EFM)
    - Connectivity Fault Management (CFM)
  - Link quality detection technologies:
    - Network Quality Analysis (NQA)
    - IP Flow Performance Measurement (IP FPM)
- On the live network, BFD is typically used to detect link connectivity, and NQA or IP FPM is typically used to detect link quality.



Overview of BFD

- BFD provides a universal, standardized, media-independent, and protocol-independent fast failure detection mechanism. It has the following advantages:
  - Provides low-overhead and fast failure detection for channels between adjacent forwarding engines.
  - Performs uniform detection for all media and protocol layers in real time.
- BFD is a simple Hello protocol. Two systems establish a BFD session channel and periodically send BFD packets to each other. If one system does not receive BFD packets from the other system within a certain period, the system considers that a fault occurs on the channel.

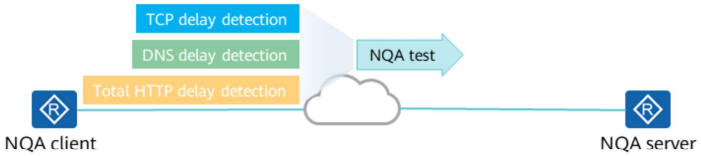


HUAWEI

- Only one BFD session can be established in a data path. If different applications need to use different BFD parameters on the same data path, use the BFD parameters that can meet the requirements of all applications to configure a unique BFD session and enable the status changes of the BFD session to be reported to all the applications bound with the BFD session.

Overview of NQA

- To visualize the quality of network services and allow users to check whether the quality of network services meets requirements, the following measures must be taken:
  - Enable devices to provide network service quality information.
  - Deploy probe devices to monitor network service quality.
- The preceding measures require devices to provide statistical parameters such as the delay, jitter, and packet loss rate and require dedicated probe devices. These requirements increase investments on devices.
- When NQA is deployed on devices, dedicated probe devices do not need to be deployed, effectively reducing costs. NQA can accurately test the network running status and output statistics.
- NQA measures network performance and collects statistics about the response time, network jitter, and packet loss rate in real time.

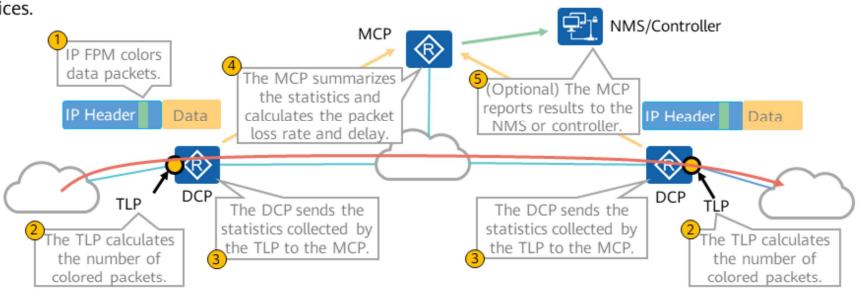


HUAWEI

- Additionally, NQA measures the performance of different protocols running on the network. This facilitates real-time collection of network performance counters, such as the total HTTP connection delay, TCP connection delay, DNS resolution delay, file transfer rate, FTP connection delay, and DNS resolution error rate.

Overview of IP FPM

- With the advent of the cloud computing era, end-to-end service performance measurement becomes essential. However, the commonly used NQA technology has the following defects in end-to-end network performance measurement scenarios:
  - NQA simulates service packet forwarding on the network by constructing service packets. Therefore, the collected performance statistics are not accurate.
  - NQA does not support end-to-end performance measurement across network layers, and cannot monitor or measure network performance in a multipath scenario of IP networks.
- IP flow performance measurement (IP FPM) can effectively solve these problems. It is a general IP network performance measurement solution. IP FPM can directly measure service packets, and the measurement data can reflect the performance of IP networks. In addition, IP FPM can monitor the changes of services carried by IP networks online and accurately reflect the running status of services.



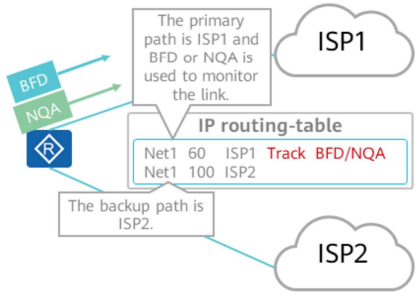
HUAWEI

- TLP is short for Target Logical Port.
  - TLPs are interfaces on the edge nodes of the network and provide the following functions:
    - Collect statistics about the packet loss rate and delay.
    - Generate statistics, such as the number of packets sent and received, volume of traffic sent and received, and timestamp.
  - An In-Point-TLP collects statistics about service flows it receives. An Out-Point-TLP collects statistics about service flows it sends.
- DCP is short for Data Collecting Point.
  - DCPs are edge nodes on the network and provide the following functions:
    - Manage and control TLPs.
    - Collect the statistics generated by TLPs.
    - Report the statistics to an MCP.

- MCP is short for Measurement Control Point.
    - MCPs are intermediate nodes on the network and provide the following functions:
        - Collect the statistics reported by DCPs.
        - Summarize and calculate the statistics.
        - Report the statistics to user terminals or the NMS.

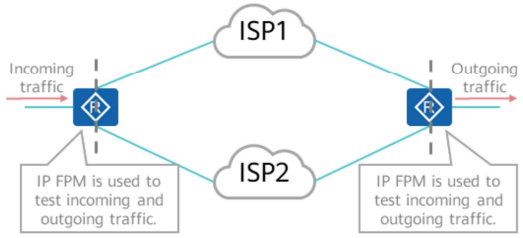
Exemplary Application Scenarios of Link Detection

Common link quality detection scenario



- NQA and BFD are usually associated with floating routes.
- When NQA or BFD detects a link fault, the high-priority route becomes invalid and the backup route is used.

E2E link quality detection



- IP FPM colors and marks data packets to help ingress and egress devices measure the packet loss or jitter of specific traffic.
- IP FPM can measure end-to-end network quality. Therefore, it can measure the actual network quality of different links in the multi-link egress scenario.

HUAWEI

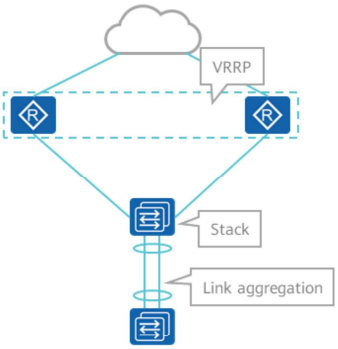
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. **Reliability Technologies and Their Applications of WAN Interconnection**
   - Link Detection Technologies and Their Applications
   - Network and Service Reliability Technologies and Their Applications
4. Optimization Technologies and Their Applications of WAN Interconnection
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

Overview of Network Reliability

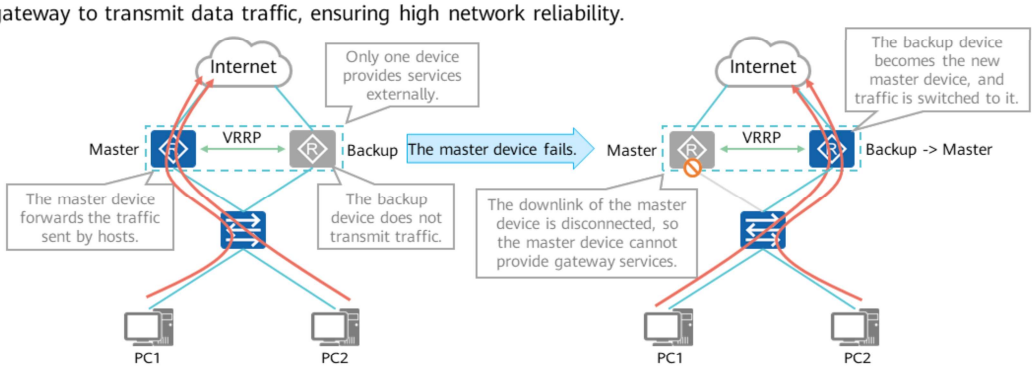
- If a fault occurs on the network, the fault may not be detected or rectified in a timely manner. To resolve this issue, redundancy technologies are required.
- Common redundancy technologies include stack, link aggregation, and VRRP.
- VRRP is the most widely used network redundancy technology on egress devices or gateways.



HUAWEI

Overview of VRRP

- Hosts are connected to external networks through gateways. If a single gateway fails, services will be interrupted for a long time. Adding egress gateways is a common method to improve system reliability. In this case, route selection among multiple egresses becomes essential.
- VRRP groups multiple routing devices into a single virtual routing device. If a gateway fails, VRRP selects a new gateway to transmit data traffic, ensuring high network reliability.



HUAWEI

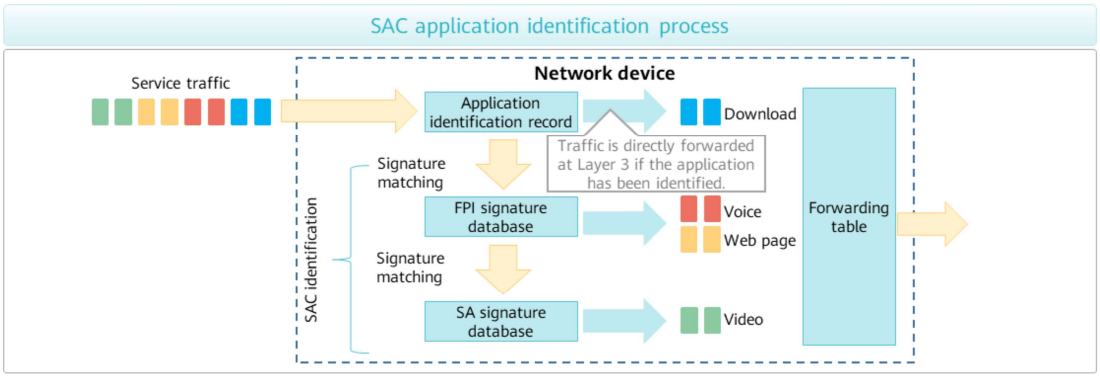
Service Reliability

- In the cloud computing era, network reliability cannot meet user requirements. Users want to understand the live network status based on applications and adjust the network based on the application status.
- Such requirements pose the following challenges to traditional networks:
  - Traditional networks cannot accurately identify applications.
  - Traditional networks cannot be adjusted based on applications.
- To cope with the challenges, two technologies are developed:
  - Smart Application Control (SAC): This technology can flexibly identify applications.
  - Smart Policy Routing (SPR): This technology can switch forwarding paths based on the network or application status.

HUAWEI

Overview of SAC

- Typically, routing and switching devices cannot identify application-layer information. Therefore, it is difficult to manage networks based on applications. SAC technology helps routing and switching devices identify classified applications.
- SAC uses service awareness (SA) and first packet identification (FPI) technologies to detect and identify Layer 4 to Layer 7 information (such as HTTP and RTP) in packets.

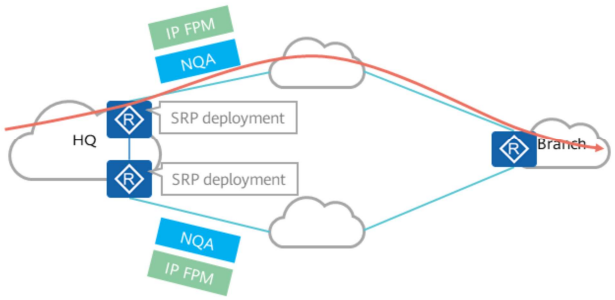


HUAWEI

- After a packet enters an SAC-enabled device, the device determines whether the corresponding application has been identified based on the 5-tuple information carried in the packet. If the application has been identified, the device forwards the packet at Layer 3 without identifying the application again. If the application has not been identified, the device performs SAC application identification. The device then processes the packet based on the SAC identification result and forwards the packet at Layer 3. The SAC application identification process is as follows: The device identifies an application based on the ACL rules defined in FPI. If the application cannot be identified, the device identifies the application based on the DNS entries defined in FPI. If the application still cannot be identified, the device identifies the application based on the protocol and port mapping table defined in FPI. If the application still cannot be identified, the device starts the SA identification process.

Overview of SPR

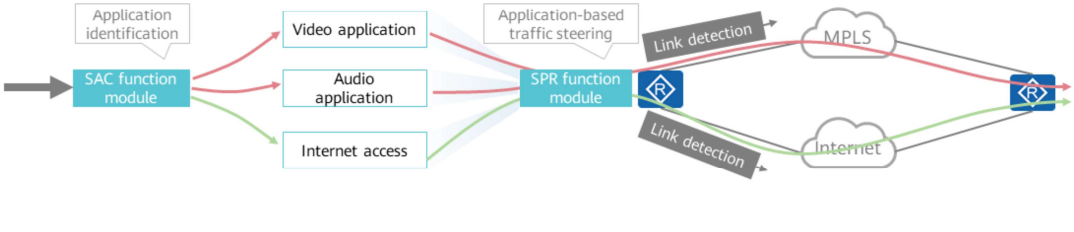
- In the cloud computing era, more users shift their attention from network connectivity to service availability, such as service response speed and service quality. However, traditional networks cannot detect link quality and service requirements, resulting in poor user experience.
- SPR addresses this problem. It actively detects the link quality and matches service requirements to select an optimal link to forward service data. SPR prevents network blackholes and flappings.



Exemplary Application Scenario of Service Reliability Technologies

- Deploying both SAC and SPR can ensure the reliability of specific services on the network.
  - SAC uses the SA signature database and FPI signature database to identify applications and group traffic.
  - SPR determines the link quality and forwarding path based on the probe packets.

Application scenario of service reliability technologies



HUAWEI

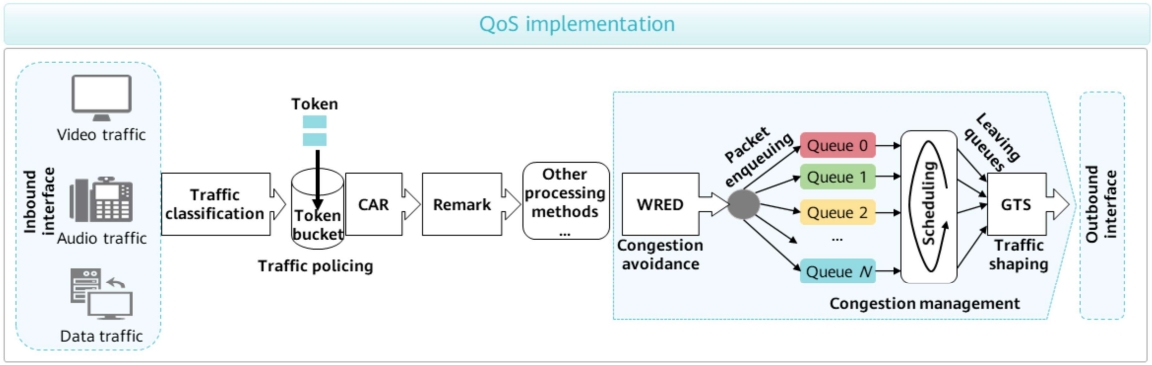
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. Reliability Technologies and Their Applications of WAN Interconnection
4. **Optimization Technologies and Their Applications of WAN Interconnection**
   - QoS Technology and Its Application
   - FEC Technology and Its Application
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

QoS Overview

- Quality of Service (QoS) defines a service provider's ability to guarantee a certain level of performance required by customers. The QoS-enabled device controls enterprise network traffic, implements congestion management and congestion avoidance, reduces the packet loss rate, and provides dedicated bandwidth for enterprise users or differentiated services (such as audio, video, and data services).



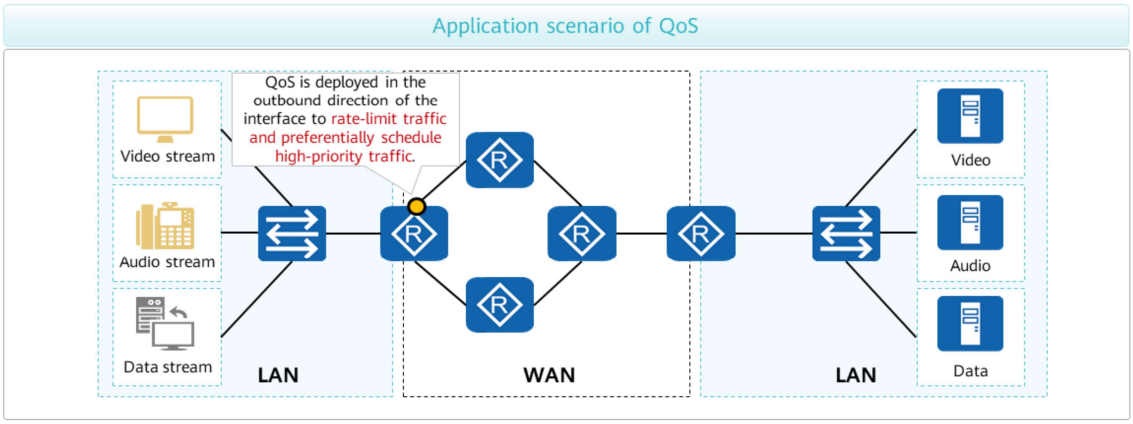
HUAWEI

- To meet SLA requirements of different services (such as audio, video, and data services), the network is required to distinguish different communication modes before providing corresponding QoS guarantee.
    - For example, real-time services such as voice over IP (VoIP) demand a shorter delay. A long delay for packet transmission is unacceptable. Email and File Transfer Protocol (FTP) services are comparatively insensitive to the delay.
    - Conventional IP networks provide the best-effect mode, which cannot help identify or differentiate the communication types on the network. The capability of differentiating communication types is the prerequisite for providing differentiated services. Therefore, the best-effect mode cannot meet the requirements of the emerging applications. This is where QoS comes in.
- QoS is designed to provide differentiated services based on application requirements. For example:
    - The bandwidth used by the FTP service on the backbone network can be limited, and a higher priority can be assigned to database access.

- An Internet service provider (ISP) can transmit real-time services such as audio or video services. With QoS technology, the ISP can differentiate the packets and provide differentiated services for users.
- High bandwidth and short delays can be guaranteed for the time-sensitive multimedia services. If other services are available on the network, these time-sensitive services are not affected.

QoS Application Example

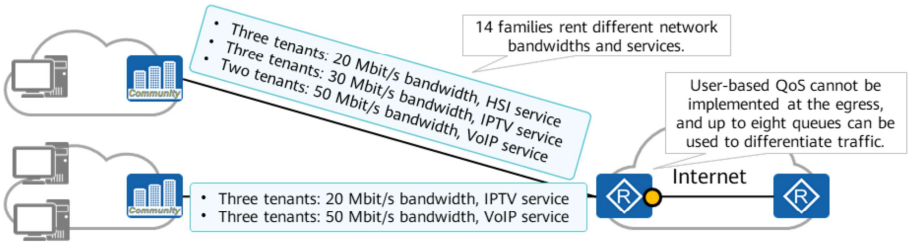
- In the enterprise WAN interconnection scenario, QoS is typically deployed on the egress link.
- QoS technology is used to ensure high network bandwidth and short delays for high-value services.



Overview of HQoS

- Conventional QoS schedules traffic based on interfaces. Interfaces themselves can only differentiate service priorities. Traffic of the same priority uses the same interface queue and compete for the same queue resources. Therefore, conventional QoS is unable to provide differentiated services based on the types of traffic and users.
- HQoS implements hierarchical scheduling based on multiple levels of queues, differentiating both services and users to provide fine-tuned QoS guarantee.

QoS limitations in multi-tenant scenarios

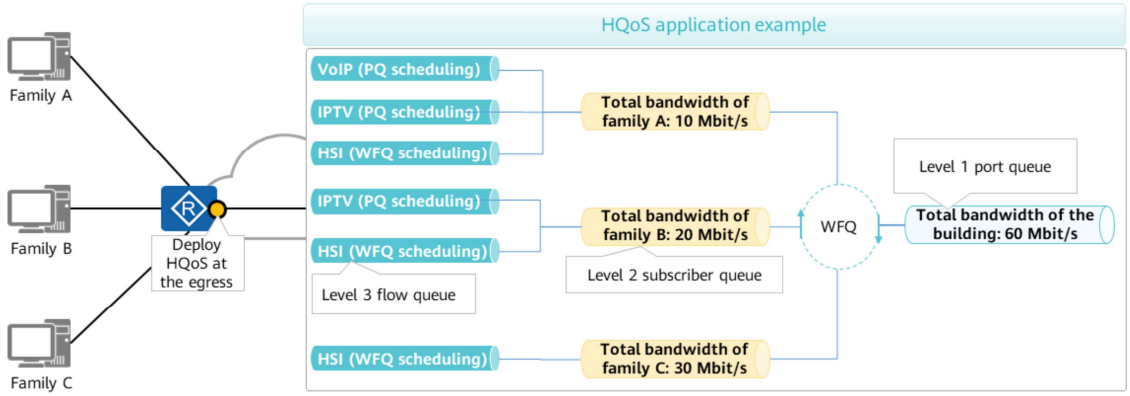


- In home broadband scenarios, different families may lease different network bandwidths and network services. QoS cannot be used to implement fine-tuned management of these families.

HUAWEI

HQoS Application Example

- HQoS is mainly used in multi-tenant scenarios. For example, there are three families in a building. Family A has purchased 10 Mbit/s bandwidth and subscribed to VoIP, IPTV, and HSI services. Family B has purchased 20 Mbit/s bandwidth and subscribed to IPTV and HSI services. Family C has purchased 30 Mbit/s bandwidth and subscribed to the HSI service. These three families have different requirements. HQoS is the best choice for this scenario.



HUAWEI

- VoIP: voice over IP service
- IPTV: Internet protocol TV service
- HSI: high-speed Internet service

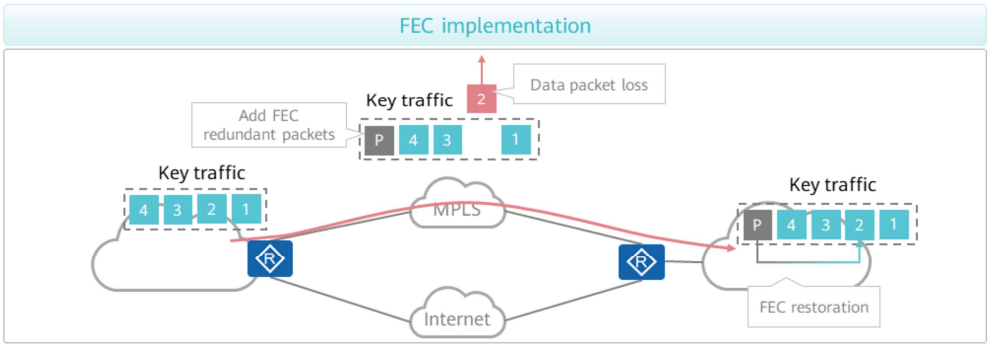
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. Reliability Technologies and Their Applications of WAN Interconnection
4. **Optimization Technologies and Their Applications of WAN Interconnection**
   - QoS Technology and Its Application
   - FEC Technology and Its Application
5. Security Technologies and Their Applications of WAN Interconnection

HUAWEI

Overview of FEC

- An IP video call often encounters image distortion or audio interruption due to packet loss. To prevent this, error control technologies are required.
- Forward error correction (FEC) is such a technology. The sender attaches FEC redundant packets to the data to be transmitted. If an error is detected, the receiver can correct the data based on the redundant packets.
- FEC is applicable to networks where random packet loss occurs or the RTT is large.

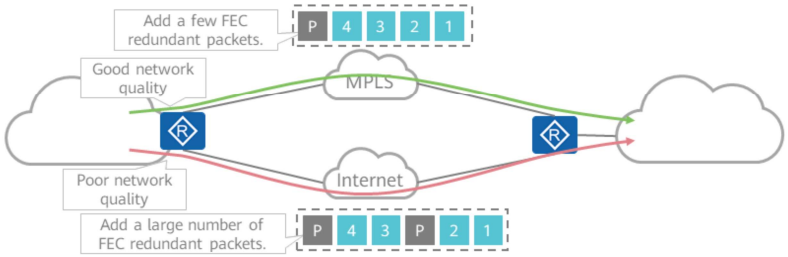


HUAWEI

- Error control technologies are classified into forward error correction (FEC) and backward error correction (BEC).
  - Automatic repeat request (ARQ) is an on-demand retransmission mechanism. The sender uses a "sending-acknowledgement" mechanism to detect whether the receiver receives data packets. If the sender does not receive any acknowledgment packet from the receiver, the sender retransmits the corresponding packet. In this error correction mode, packets need to be sent repeatedly. As a result, an extra delay is introduced. In this mode, the sender retransmits packets to implement error correction. This is why this mode is called BEC.
  - In FEC, both FEC redundant packets and data packets are sent to the receiver. If an error is found, the receiver directly restores the lost data packets by using the FEC redundant packets. In this mode, error correction is performed on the receiver. This is why this mode is called FEC.
    - The error correction capability of FEC is limited. For example, if three out of four packets are lost, FEC cannot restore these three packets based on the remaining packet.

Overview of A-FEC

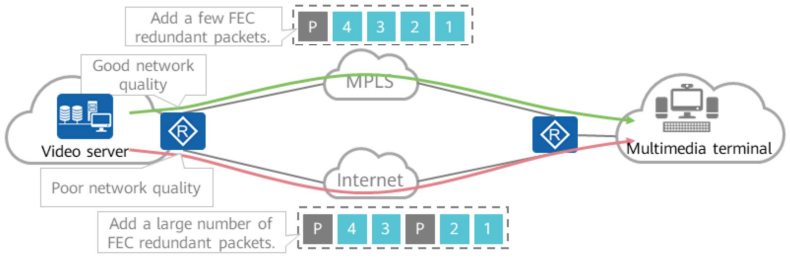
- In terms of FEC, the ratio of the original data to FEC redundant packets is fixed. If burst packet loss occurs, the original data may fail to be restored based on the redundant packets. Considering this, FEC is not applicable to the network where burst packet loss may occur. In addition, since FEC will add redundant packets, the bandwidth utilization will decrease. This is where adaptive-FEC (A-FEC) technology comes in.
- A-FEC technology can flexibly add redundant packets based on the live network quality.
  - When the network quality is good, A-FEC adds a few redundant packets or even does not add any redundant packet to improve bandwidth utilization.
  - When the network quality is poor, A-FEC adds more redundant packets to prevent data restoration failures caused by burst packet loss.



Exemplary Application Scenario of FEC/A-FEC

- FEC or A-FEC can be used to guarantee mission-critical video applications and reduce frame freezing and artifacts.

Application scenario of FEC or A-FEC



- The A-FEC algorithm is used to flexibly increase redundant packets at the transmit end based on the application and link quality to prevent continuous packet loss on the network. This helps prevent mission-critical video applications from frame freezing or artifacts even when the packet loss rate reaches up to 20%.

HUAWEI

Contents

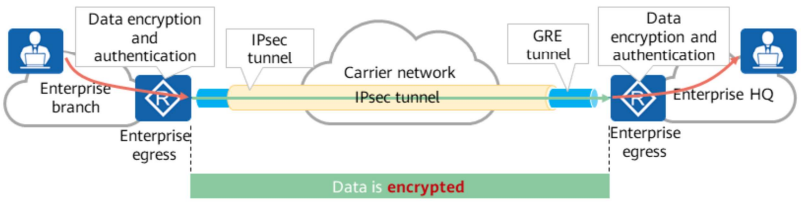
1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. **Security Technologies and Their Applications of WAN Interconnection**
   - WAN Transmission Security Technologies and Their Applications
   - WAN Service Security Technologies and Their Applications

HUAWEI

Overview of IPsec

- The IP Security (IPsec) protocol suite is a series of security protocols developed by the Internet Engineering Task Force (IETF). It provides a cryptology-based, interoperable, and high-quality security protection mechanism for end-to-end IP packet exchange.
- IPsec encrypts and authenticates data to ensure secure data transmission on the Internet.
- IPsec VPN technology can be used with multiple VPN technologies to provide flexible and secure enterprise interconnections.

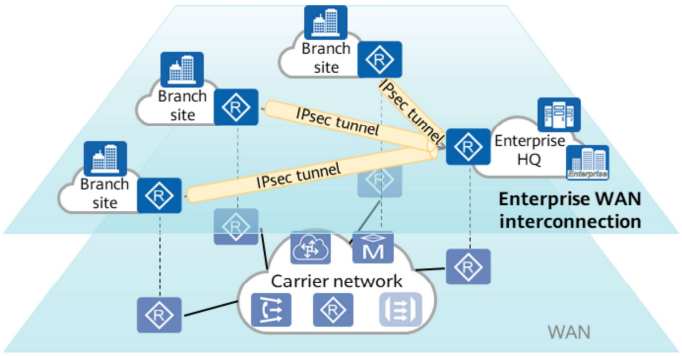
GRE over IPsec application



- On the live network, GRE over IPsec technology is typically used for interconnection between branch sites. IPsec technology ensures secure data transmission, and GRE technology ensures interconnection between enterprise intranets.

Exemplary Application Scenario of IPsec

- Enterprise branches often need to interconnect with each other. Since Internet links are prone to security risks, IPsec technology can be used to encrypt data packets before transmitting them so as to guarantee enterprise interconnection security.



HUAWEI

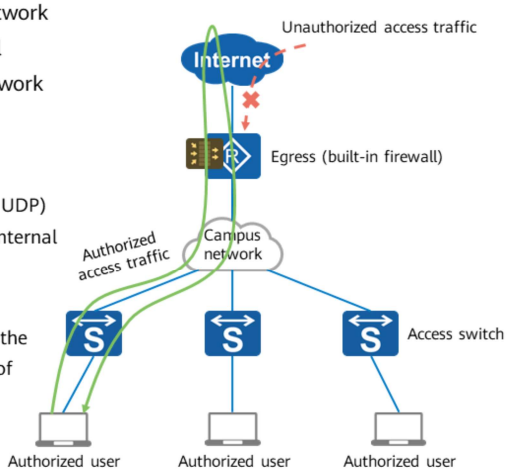
Contents

1. Basic Architecture of a Typical WAN Interconnection Solution
2. Networking Technologies and Their Applications of WAN Interconnection
3. Reliability Technologies and Their Applications of WAN Interconnection
4. Optimization Technologies and Their Applications of WAN Interconnection
5. **Security Technologies and Their Applications of WAN Interconnection**
   - WAN Transmission Security Technologies and Their Applications
   - WAN Service Security Technologies and Their Applications

HUAWEI

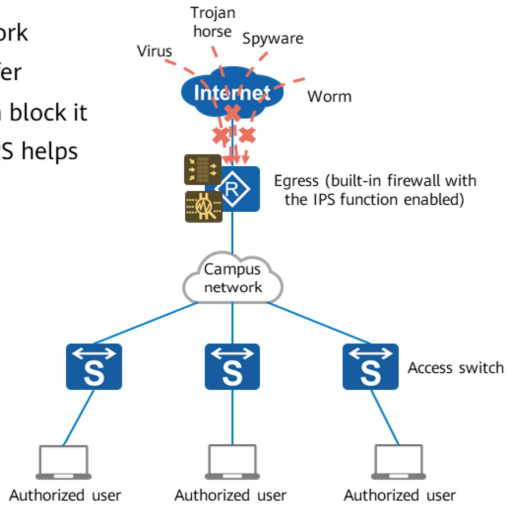
Firewall Overview

- A firewall separates an internal network from an external network to prevent external network users from accessing the internal network in an unauthorized way and protect the internal network from being intruded by unauthorized users.
- Typical functions of firewalls include:
  - Firewalls can detect transport-layer protocol information (TCP and UDP) and determine whether to allow TCP or UDP packets to enter the internal network based on the packets' source IP addresses, destination IP addresses, and port numbers.
  - Firewalls can check application-layer protocol information, such as the protocol type and port number of packets, and monitor the status of connection-based application-layer protocols.



IPS Overview

- The intrusion prevention system (IPS) examines network traffic flows to detect intrusion behavior (such as buffer overflow attacks, Trojan horses, and worms) and then block it in real time through certain response methods. The IPS helps protect enterprise information systems and network architectures against intrusions.



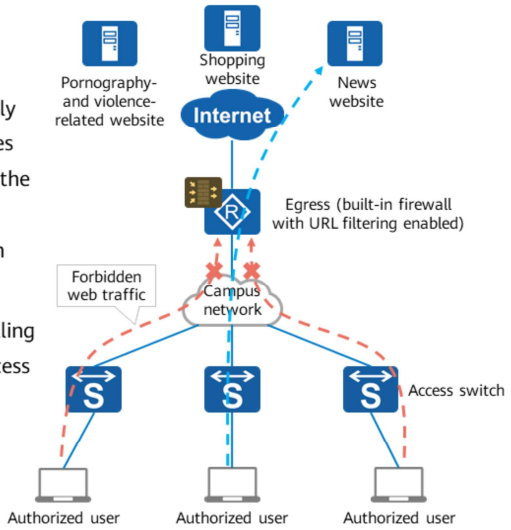
HUAWEI

- Intrusion prevention is a security protection technology that can discover and block intrusion behavior. After discovering network intrusions, the IPS can automatically discard intrusion packets or block attack sources to fundamentally prevent attacks. Advantages of intrusion prevention include:
  - Real-time attack block: A security device is deployed on a network in in-line mode. When detecting intrusion, the device blocks intrusion and attack traffic in real time, minimizing impacts of network intrusions.
  - In-depth protection: New attacks are hidden at the application layer of the TCP/IP protocol. Intrusion prevention can detect the content of application-layer packets, reassemble network data flows for protocol analysis and detection, and determine the traffic that must be blocked based on the attack type and policy.
  - All-round protection: Intrusion prevention provides protection measures against attacks such as worms, viruses, Trojan horses, botnets, spyware, adware, Common Gateway Interface (CGI) attacks, cross-site scripting attacks, injection attacks, directory traversal attacks, information leakage, remote file inclusion attacks, overflow attacks, code execution, DoS attacks, and scanning tools. All-round protection comprehensively helps defend against various attacks and protect network security.

- Internal and external protection: Intrusion prevention can protect enterprises from both external and internal attacks. The IPS can detect the traffic passing through and protect servers and clients.
- Continuous update and precise protection: The IPS signature database is updated continuously to maintain the highest security level. You can periodically update the IPS signature database of a device from the update center to ensure effective intrusion prevention.

Overview of URL Filtering

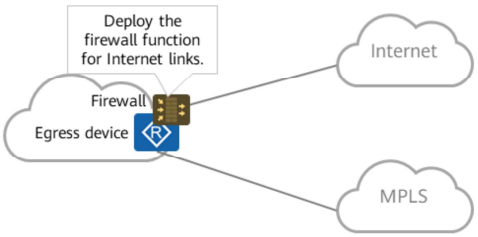
- A Uniform Resource Locator (URL) is an address of standard resources on the Internet.
- Employees' uncontrolled access to website resources will severely reduce their work efficiency, waste network bandwidth resources of enterprises, and even introduce viruses and Trojan horses to the enterprises' internal networks. In addition, a large amount of pornographic and violent information will affect people's health mentally and physically.
- The URL filtering function regulates online behaviors by controlling users' network access, thereby permitting or rejecting users' access to specified web page resources.



Security Technology Application Example

- Deploying security technologies at the egress can effectively guarantee internal network security.

Security technology application scenario



- The public network has high security risks. You are advised to deploy the firewall function for links connected to the public network to improve internal network security.
- The private network has relatively low security risks. Therefore, the firewall function does not need to be deployed.

Quiz

1. (Multiple-answer question) Which of the following are private line technologies? ( )
   A. SDH
   B. L2TP
   C. MPLS VPN
   D. IPsec VPN
2. (True or false) SRP technology can flexibly select egress links based on the link quality. ( )
   A. True
   B. False

HUAWEI

1. AC
2. A

Summary

- There are two WAN interconnection modes: private line and VPN.
- To ensure WAN reliability, the link quality needs to be detected first, and then specific technologies are used to select a proper egress or egress link based on the link quality.
- There are many WAN optimization methods. Typically, QoS is used to control egress traffic and FEC is used to ensure reliable data forwarding on WANs.
- WAN security needs to be ensured from two aspects: WAN transmission and WAN egress. Typically, IPsec is used to guarantee WAN transmission security, while technologies such as firewall are used to guarantee WAN egress security.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



Enterprise WAN Interconnection Design



HUAWEI

Foreword

- After years of development and evolution, WAN interconnection has undergone significant changes. In the past, WAN interconnection was centered on the network, and there were few applications. As the main component of WAN interconnection, the network took the most important position. However, with the rise of cloud computing, the potential of applications is fully exploited, and WAN interconnection gradually become application-centric.
- To address the challenges posed by cloud computing to WAN interconnection, Huawei launches the SD-WAN Solution.
- This course will provide insights into the functions available in the SD-WAN Solution and the design scheme for this solution.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the challenges faced by enterprise WAN interconnection.
  - Describe the architecture and functions of Huawei SD-WAN Solution.
  - Describe the SD-WAN networking design scheme.
  - Describe the SD-WAN service design scheme.
  - Describe the SD-WAN reliability design scheme.
  - Describe the SD-WAN security design scheme.

HUAWEI

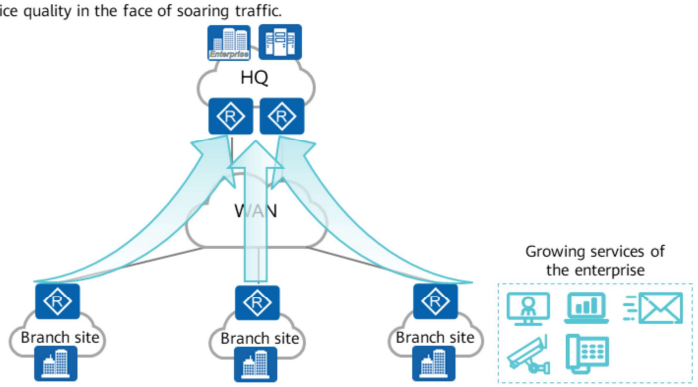
Contents

1. **Development Trends and Challenges Facing Enterprise WAN Interconnection**
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

Challenges Brought by Cloud Computing to Enterprise WAN Interconnection

- Before cloud computing is widely used, there are only a small number of network applications. The network service quality depends only on the bandwidth, and service traffic does not need to be managed in a refined manner. The Internet is built on a network-centric design.
- The advent of the cloud computing era leads to a significant increase in the number of network applications. Enterprises have difficulty in striking a balance between the line price and service quality in the face of soaring traffic.

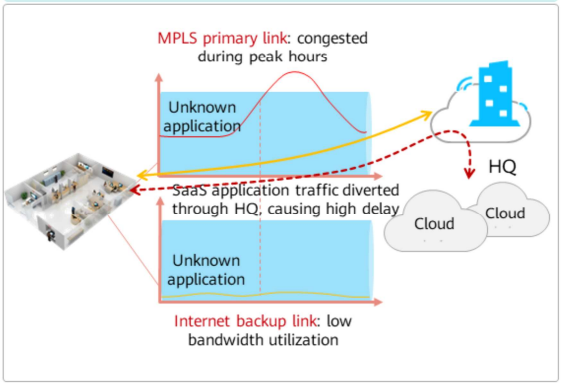


HUAWEI

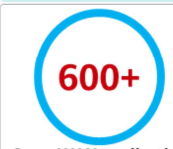
Challenges Brought by Multiple Services to Enterprise WAN Interconnection

- Enterprises are unable to detect service quality in real time and therefore cannot effectively guarantee key services. In addition, enterprises cannot monitor service traffic in real time, and therefore are unable to quickly adjust service traffic.

No application visibility, causing difficulty in traffic scheduling



Difficulty in managing key services such as voice, video, and SaaS applications



Cross-WAN application (An enterprise)

Priority conflict: Key applications cannot be identified and therefore have a low scheduling priority.

Bandwidth conflict: During peak hours, the burst traffic is three to five times the average traffic, affecting key applications.



Smooth video conference when idle bandwidth is available



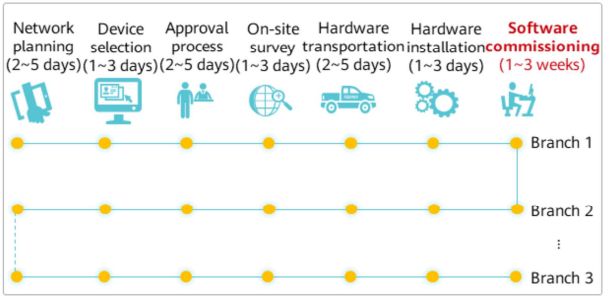
Bandwidth conflict leads to frame freezing in video conferences

HUAWEI

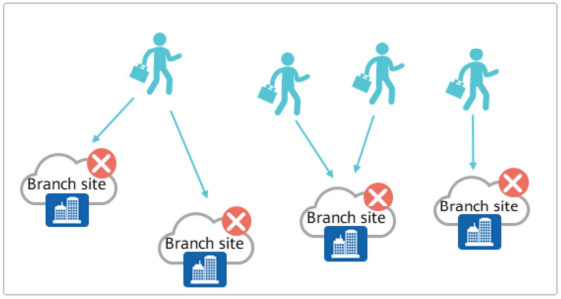
Challenges Brought by Large Numbers of Branches to Enterprise WAN Interconnection

- With the development of an enterprise, it will have more and more cross-city, cross-province, and cross-country branches, causing the following problems in managing branch site networks:
  - Too many branches result in high O&M costs.
  - It takes a long time to provision new services in branches.
  - It is difficult to rectify faults on branch networks.

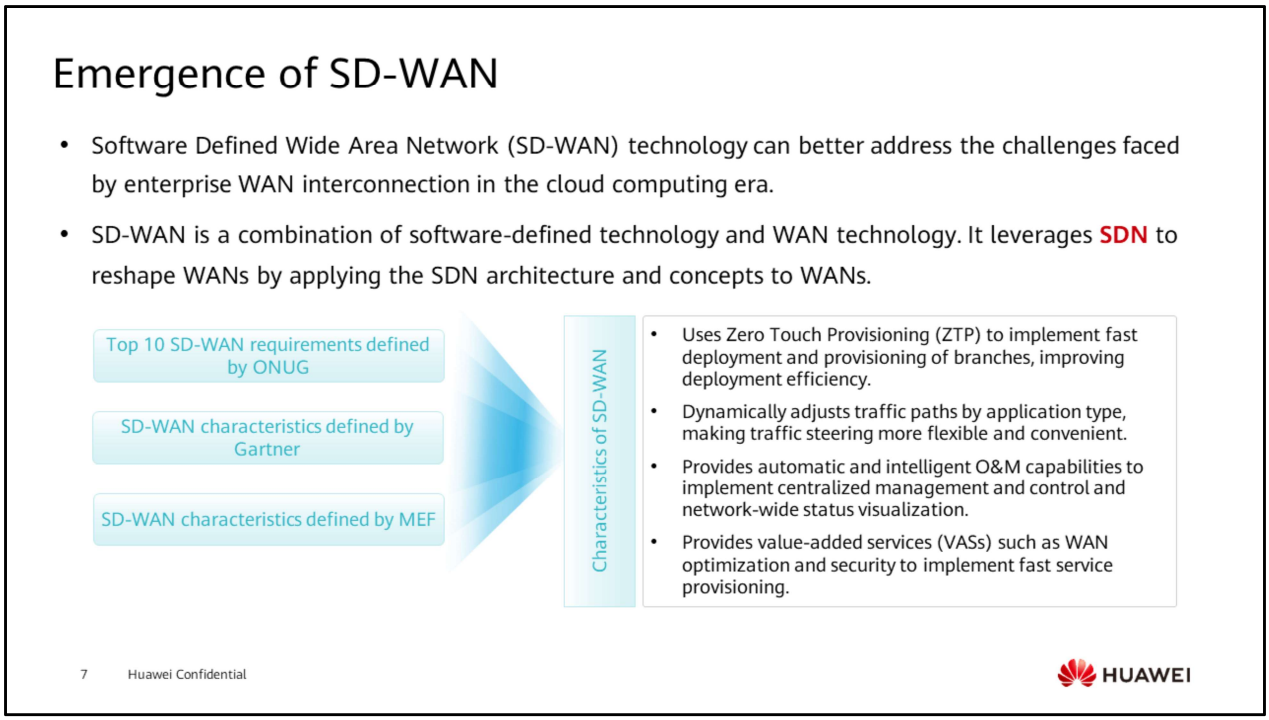
Time-consuming service provisioning in branches



Difficult troubleshooting on branch networks, resulting in high O&M costs



HUAWEI



- ONUG: refers to Open Networking User Group. ONUG is an influential user organization led by large enterprises and comprised mainly of IT users. It was founded by IT technical executives of well-known large enterprises in North America and is dedicated to driving IT implementation and network technology transformation for large enterprises. ONUG members include large enterprises in industries such as finance, insurance, medical care, and retail. ONUG serves as a platform for high-end customers in North America to discuss and communicate IT requirements.
- Gartner: It is the world's most authoritative IT research and advisory company. Its research scope covers all IT industries. It provides objective and fair demonstration reports and market research reports for customers in terms of IT research, development, evaluation, application, and market, thereby assisting customers in market analysis, technology selection, project demonstration, and investment decision-making.

- The Metro Ethernet Forum (MEF) is a non-profit organization dedicated to solving the technical problems of the Metro Ethernet. It aims to widely apply the Ethernet technology to the construction of the Metro Ethernet as switching and transmission technology. The objective of MEF is to promote the implementation of existing and new standards, Ethernet service definitions, test procedures, and technical specifications, developing Ethernet-based MANs into carrier-class networks. The major tasks of MEF also include providing LSO-based solutions (LSO is short for Lifecycle Service Orchestration) and architectures for carriers' managed service markets and defining northbound interfaces (NBIs) to enable multi-vendor interoperability.

Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. **Overview of Huawei SD-WAN Solution**
   - Huawei SD-WAN Solution
     - Major Functions of Huawei SD-WAN Solution
     - Application Scenarios of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

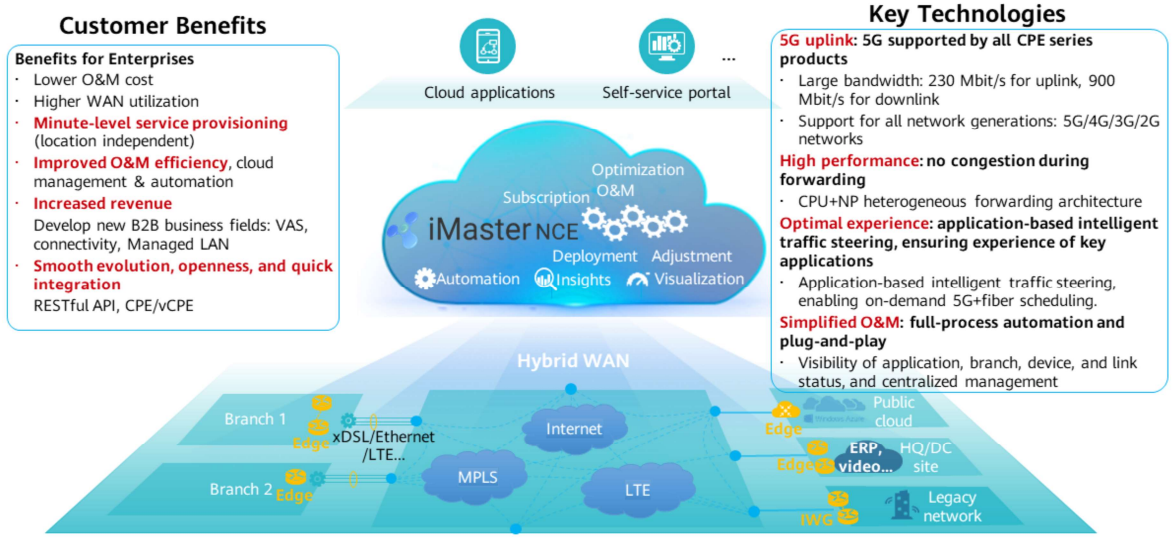
HUAWEI

Overview of Huawei SD-WAN Solution

Customer Benefits

Benefits for Enterprises

- Lower O&M cost
- Higher WAN utilization
- **Minute-level service provisioning** (location independent)
- **Improved O&M efficiency**, cloud management & automation
- **Increased revenue**
  Develop new B2B business fields: VAS, connectivity, Managed LAN
- **Smooth evolution, openness, and quick integration**
  RESTful API, CPE/vCPE



Key Technologies

5G uplink: 5G supported by all CPE series products

- Large bandwidth: 230 Mbit/s for uplink, 900 Mbit/s for downlink
- Support for all network generations: 5G/4G/3G/2G networks

High performance: no congestion during forwarding

- CPU+NP heterogeneous forwarding architecture

Optimal experience: application-based intelligent traffic steering, ensuring experience of key applications

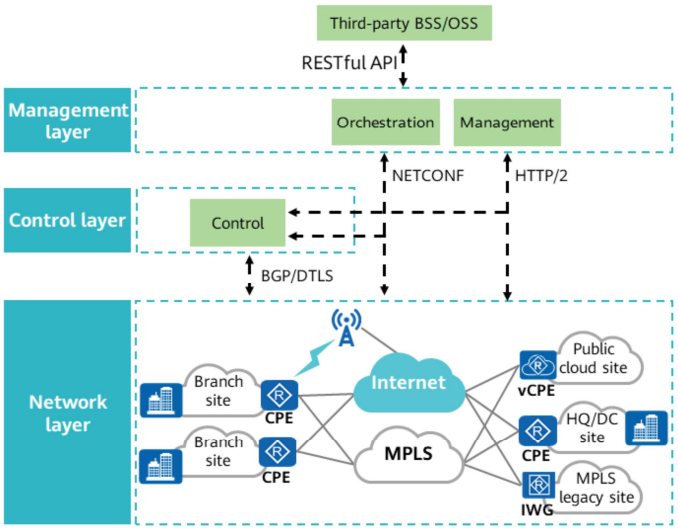
- Application-based intelligent traffic steering, enabling on-demand 5G+fiber scheduling.

Simplified O&M: full-process automation and plug-and-play

- Visibility of application, branch, device, and link status, and centralized management

HUAWEI

Architecture of Huawei SD-WAN Solution

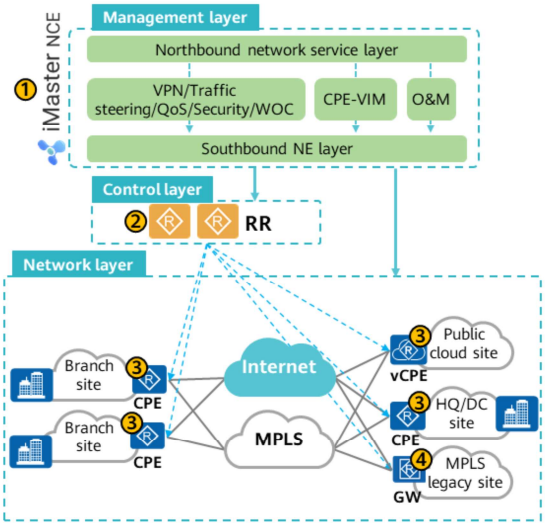


- **Management layer:** iMaster NCE-WAN manages the entire process of enterprise interconnection services in a refined manner. It uses NETCONF to manage CPEs in the southbound direction and provides standard RESTful APIs in the northbound direction to connect to third-party applications.
- **Control layer:** iMaster NCE-WAN works with the distributed control component to transmit routes between different sites in a region and implement inter-region network interconnection.
- **Network layer:** The overlay technology is used together with cost-effective CPEs to connect branches, headquarters, and the cloud platform, thereby achieving on-demand full-network connectivity by fully leveraging all types of links such as the Internet and traditional private lines.

HUAWEI

- In terms of functions, the overall architecture of the SD-WAN Solution consists of the network layer, control layer, and management layer. The layers are associated with each other through standards-compliant interfaces and communication protocols.
- iMaster NCE-WAN is the controller for Huawei SD-WAN Solution and consists of the management layer (built-in portal) and control layer.
- Enterprise sites include enterprise branches, headquarters, data centers, and IT infrastructure deployed on the cloud.

Key Components of the SD-WAN Solution

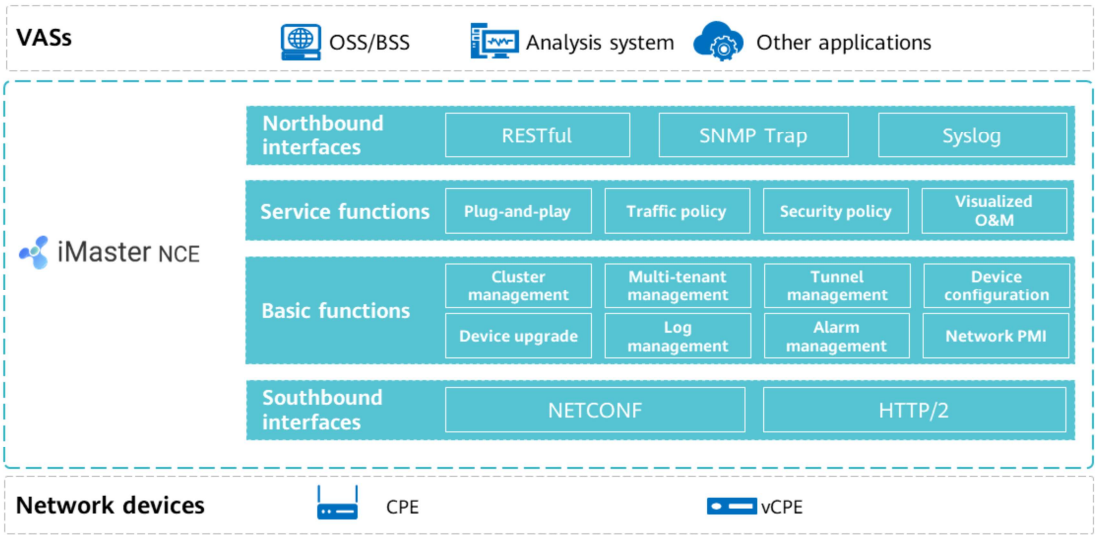


No.	Component	Functions
1	iMaster NCE-WAN	1. Network service orchestration 2. Network element (NE) control 3. Basic network O&M 4. CPE orchestration and management 5. Basic performance monitoring (provides link quality information, application quality information, traffic information, and multi-dimensional statistics for single sites and between sites)
2	RR	Distributes VPN routes and tunnel information between CPEs based on VPN topology policies.
3	CPE	Egress CPE of a site, which can be a traditional CPE or an NFV vCPE.
4	GW	Intermediate gateway device that connects the SD-WAN network and non-SD-WAN networks.

HUAWEI

- CPE: Customer Premises Equipment

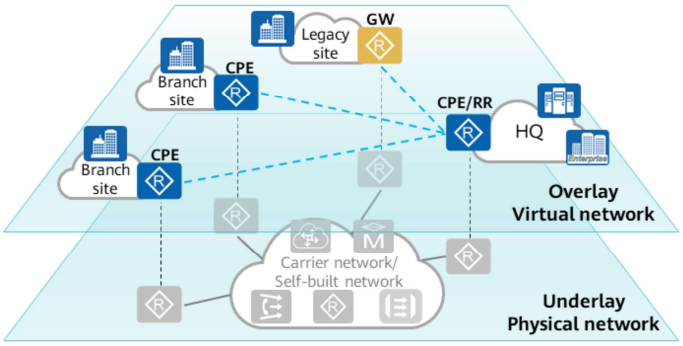
Architecture of Huawei iMaster NCE-WAN



HUAWEI

Network Layer Overview

- An enterprise SD-WAN network can be divided into two layers: physical network (underlay) and virtual network (overlay), which are completely decoupled from each other.
  - Physical network: refers to the underlay WAN provided by a carrier or built by the enterprise, including the private line network and the Internet.
  - Virtual network: also called the overlay network. Huawei SD-WAN Solution uses the IP overlay virtualization technology to build one or more virtual overlay networks on top of the physical network. Service policies are deployed on virtual networks and are decoupled from the physical network, thereby separating services from the WAN.

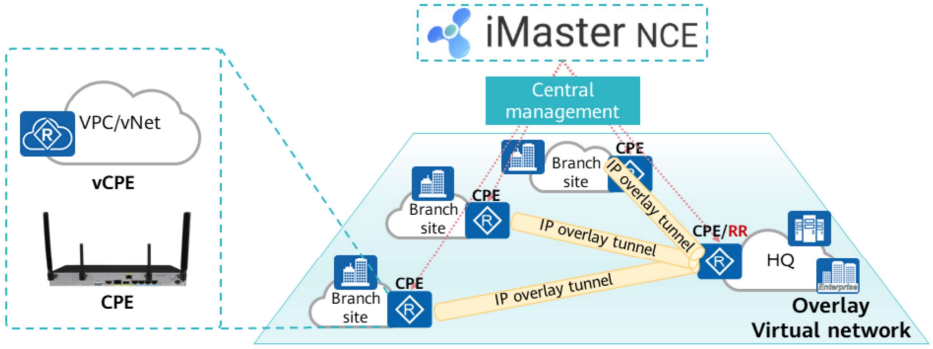


HUAWEI

- Multiple virtual networks can be deployed to provide different services for the same tenant (for example, services for multiple departments) or provide different services for different tenants.
- In terms of network device functions, the network layer of the SD-WAN Solution consists of two types of NEs: CPE and gateway (GW).

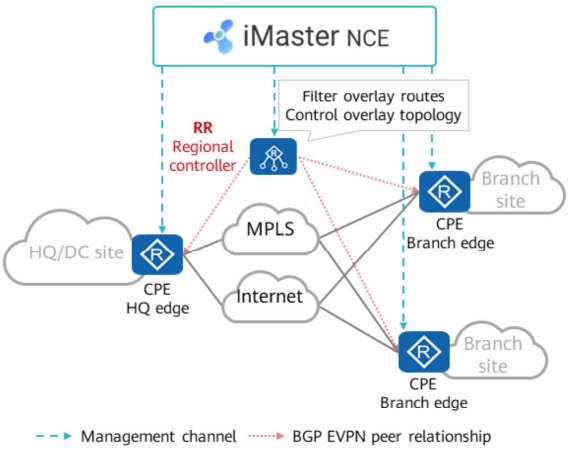
CPE Overview

- A CPE is essentially an edge node of the SD-WAN network and is also called an Edge-CPE. CPEs are interconnected using the IP overlay tunneling technology.
- Traditional hardware CPEs are typically used on the HQ and branch sites, and vCPE software virtual devices can be deployed at sites on the public cloud.
- All SD-WAN CPEs of an enterprise are centrally managed and maintained in iMaster NCE-WAN by the tenant administrator.



HUAWEI

RR Overview



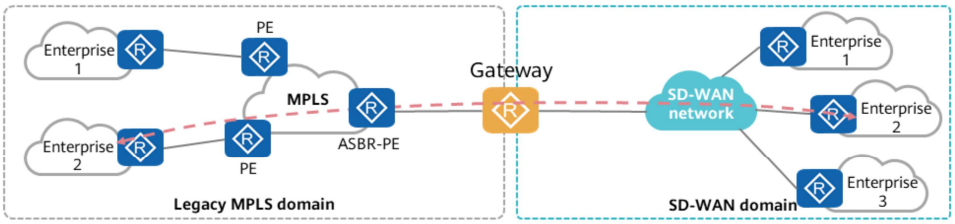
- A route reflector (RR) is used to transmit BGP routes.
- In Huawei SD-WAN Solution, RRs also control routes and network topologies. Therefore, RRs are also called regional controllers in this solution.
- Both RRs and CPEs at edge sites are managed by iMaster NCE-WAN.
- Control channels are established between RRs as well as between RRs and edge sites.
- RRs are managed by iMaster NCE-WAN and control route sending and receiving at edge sites based on the overlay network topology model. In this way, sites can communicate with each other based on the user-configured overlay topology model.

HUAWEI

- RR site: The CPE at the site functions as an RR and distributes EVPN routes between CPE gateways at edge sites based on VPN topology policy.
- If the tenant administrator assigns the role of "gateway + RR" to an egress CPE when adding the CPE, the site where the CPE resides is an RR site. If no device at a site is assigned the "gateway + RR" role, the site is an edge site.
- An edge site can establish IBGP peer relationships with two RRs that back up each other.
- Multiple RRs can be deployed for a tenant. All RRs are connected in full-mesh mode on the control plane.

Gateway Overview

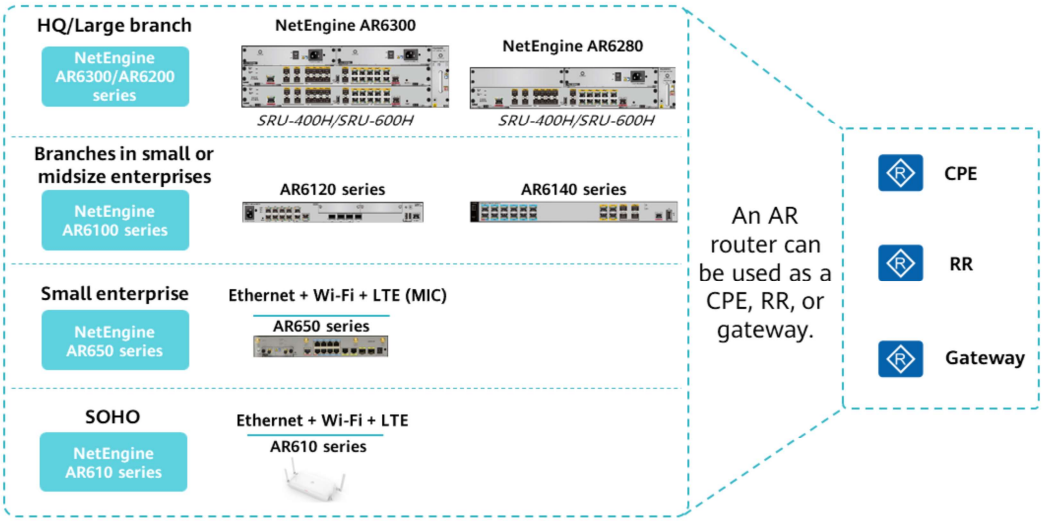
- New SD-WAN sites of an enterprise need to communicate with its legacy sites or third-party services. Some legacy sites are interconnected through MPLS VPN, and SD-WAN sites are interconnected through IP overlay tunnels. Therefore, the legacy network and SD-WAN network cannot directly communicate with each other.
- An SD-WAN gateway can connect to both the SD-WAN and legacy networks. It can function as an intermediate gateway to implement interconnection between SD-WAN and legacy networks.



HUAWEI

- A gateway has different roles in different service scenarios. For example, a gateway connected to a legacy site may be referred to as an interworking gateway (IWG), and a gateway connected to the cloud may be called a cloud gateway. These gateways can extend functions by interconnecting with each other to establish a Point of Presence (PoP) network, where these gateways are referred to as PoP gateways.

Huawei CPE Devices: NetEngine AR Routers



HUAWEI

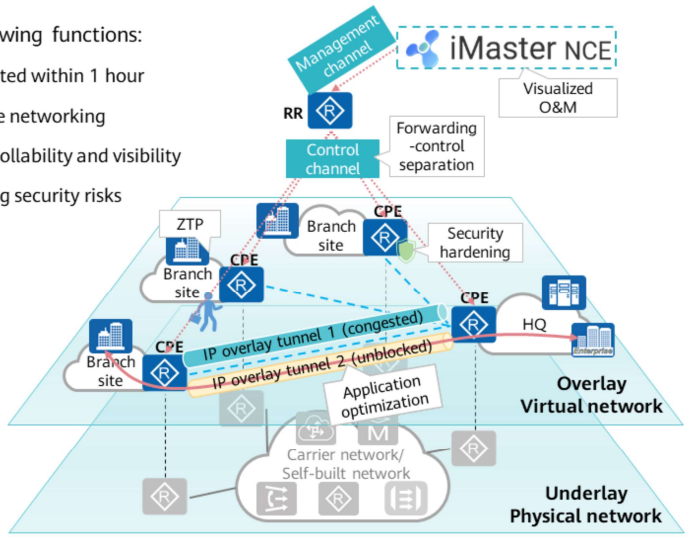
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. **Overview of Huawei SD-WAN Solution**
   - Huawei SD-WAN Solution
   - Major Functions of Huawei SD-WAN Solution
   - Application Scenarios of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

Major Functions of Huawei SD-WAN Solution

- Huawei SD-WAN Solution provides the following functions:
  - ZTP, enabling service provisioning to be completed within 1 hour
  - Forwarding-control separation, enabling flexible networking
  - Application optimization, enabling service controllability and visibility
  - Complete security protection system, eliminating security risks
  - Visualized O&M for quick fault locating

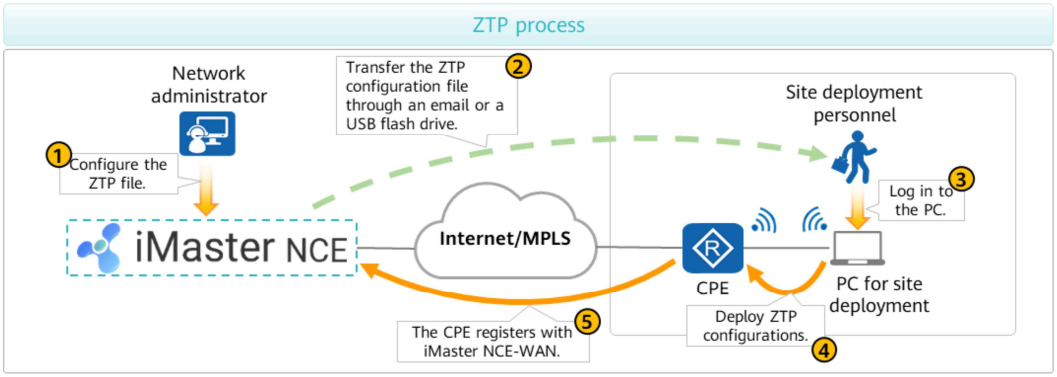


HUAWEI

- ZTP: Multiple ZTP modes are available to enable CPEs to quickly register with iMaster NCE-WAN.
- Forwarding-control separation enables flexible networking: Each CPE establishes a management channel with iMaster NCE-WAN through NETCONF, and iMaster NCE-WAN delivers configurations to CPEs to establish IP overlay tunnels between the CPEs.
- Application optimization for service controllability and visibility: The service awareness (SA) technology is used to identify applications. TCP Flow Performance Measurement (FPM) and IP FPM technologies are used to implement application-based quality measurement. The IP FPM technology can also be used for link quality measurement. Smart Policy Routing (SPR) technology provides intelligent link switchover based on the application quality.
- Complete security protection system ensures service security: Multiple VPN technologies, such as IPsec and MPLS, are leveraged to provide end-to-end protection. The firewall function is supported to provide comprehensive security protection at the hardware, pipe, and application levels.
- Visualized O&M for quick fault locating: iMaster NCE-WAN collects network-wide data and displays key indicators, helping O&M personnel quickly locate faults.

ZTP Overview

- With the development of network technologies such as SDN and cloud computing, a growing number of enterprise networks are using cloud-based management mode, but most sites still need to be deployed by technical engineers onsite, leading to high deployment costs and long deployment periods. To address these problems, Huawei develops the Zero Touch Provisioning (ZTP) function.

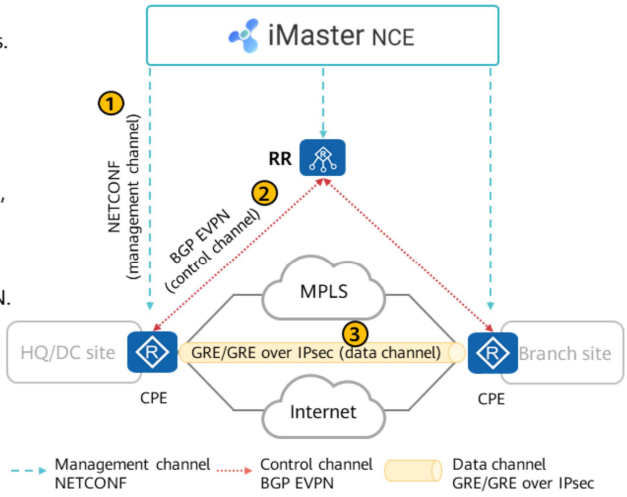


HUAWEI

- Huawei SD-WAN Solution supports the following ZTP modes:
  - Email-based deployment
  - DHCP-based deployment
  - USB-based deployment

Flexible Networking Overview

- The primary function of SD-WAN is to provide flexible and reliable networking for enterprise WANs.
- Huawei SD-WAN Solution leverages the IP overlay technology together with traditional network technologies — such as Layer 2 switching, Layer 3 routing, and VPN isolation — to achieve on-demand, flexible, and automatic connections between enterprise branches, DCs, and the cloud, with full management support provided by iMaster NCE-WAN.
- Huawei SD-WAN Solution uses the following channels to implement flexible networking:
  - Management channel
  - Control channel
  - Data channel



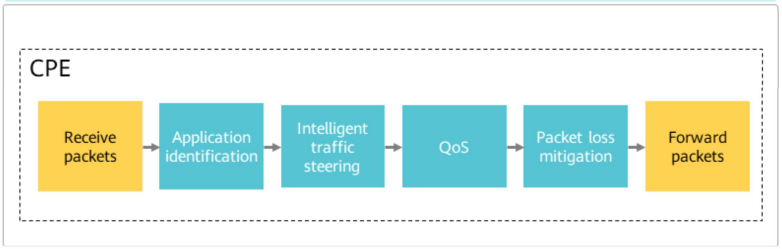
HUAWEI

- Management channel:
  - iMaster NCE-WAN sets up management channels with all devices through NETCONF, so as to manage NEs and orchestrate services on the entire network.
- Control channel:
  - CPEs set up control channels with the RR.
  - The RR centrally controls and distributes service routes between branch sites.
  - The enhanced BGP EVPN protocol is used to implement separate transmission of tenants' VPN route and next hop information, and IPsec SA negotiation.
- Data channel:
  - CPEs set up data channels with each other.
  - CPEs forward data based on GRE or GRE over IPsec tunnels. The extended GRE header carries VN IDs to differentiate tenants or departments, thereby transmitting data of multiple VNs over the same tunnel.

Application Optimization Overview

- The traditional WAN has the following disadvantages that make it unable to meet network requirements:
  - Applications of different values are carried on the same link.
  - Dynamic path selection cannot be implemented when link quality deteriorates.
  - No effective measure is available when link quality deteriorates.
- To address these problems, Huawei SD-WAN Solution provides the application experience optimization solution that offers the following functions:
  - Application identification
  - Intelligent traffic steering
  - QoS
  - Packet loss mitigation

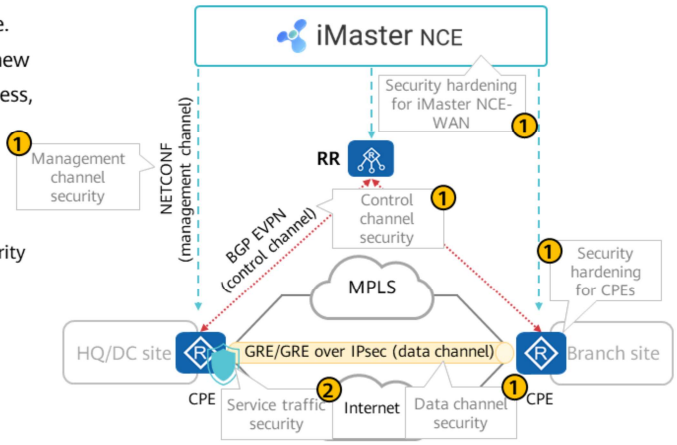
Typical application optimization process



HUAWEI

Security Overview

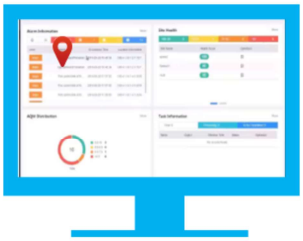
- The advent of SD-WAN stimulates the transition of the enterprise WAN architecture from the traditional closed architecture to an open one. This increases the attack surface and brings new security challenges, such as unauthorized access, data leakage, and network attacks.
- The security of Huawei SD-WAN Solution involves two aspects:
  1. System security: includes inter-component security and component security.
  2. Service security: includes firewall, IPS, and URL filtering.



HUAWEI

Visualized O&M Overview

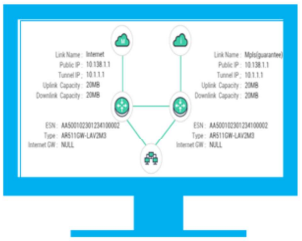
Quickly obtain abnormal traffic



Real-time alarm monitoring

- Customized dashboard (by role or preference)
- Network-wide real-time alarms (minute-level)

Quickly locate faulty devices or sites



Visualized topology status

- Displays topology based on sites and links
- Provides real-time status and performance of sites and links

Optimize WAN investment and configuration policies



45+ customized views

Site, link, application, device, and health status views

- Bandwidth utilization for a site
- Top *N* sites by throughput
- Top *N* applications by traffic
- Link throughput trend...

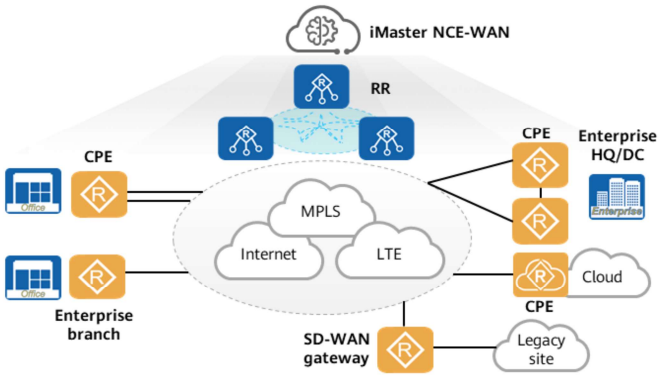
HUAWEI

Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. **Overview of Huawei SD-WAN Solution**
   - Huawei SD-WAN Solution
   - Major Functions of Huawei SD-WAN Solution
   - Application Scenarios of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

SD-WAN Business Model: Enterprise-Built

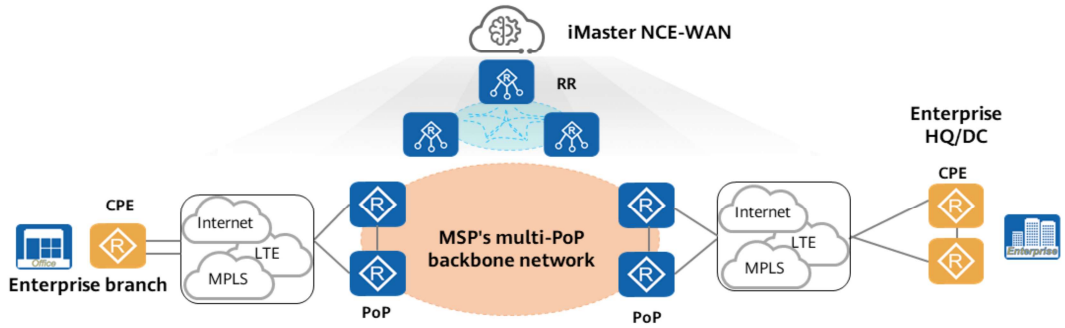


Products Involved	Functions
iMaster NCE-WAN	As the core of the SD-WAN Solution, it centrally manages CPEs, automatically delivers services, and uniformly controls the overlay network.
RR	It is the distributed control component that distributes VPN routes between CPEs based on VPN topology policies.
CPE	Egress CPEs of enterprise branches, headquarters, and DCs.
SD-WAN gateway	An intermediate gateway device that connects an SD-WAN network with a non-SD-WAN network.

- SD-WAN: iMaster NCE-WAN is deployed on the WAN to centrally manage CPEs and implement ZTP, thereby shortening the service provisioning time. This helps enterprises cope with challenges brought by cloud services and change services on demand.
- Large enterprises with a vast number of branches, such as financial institutions, retail chains, and gas stations, can deploy iMaster NCE-WAN at the headquarters to set up SD-WAN networks and manage SD-WAN services on their own.

HUAWEI

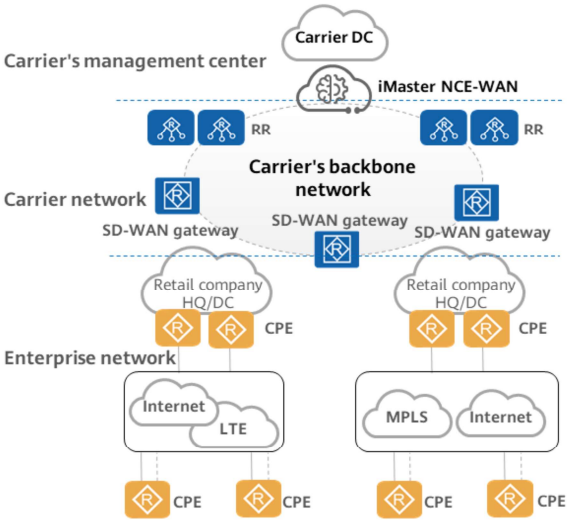
SD-WAN Business Model: MSP Resale



- A managed service provider (MSP) provides a unified SD-WAN controller (iMaster NCE-WAN) to offer SD-WAN services for multiple enterprises. The MSP builds a PoP backbone network, to which an enterprise can connect through the nearest PoP provided by the MSP, thereby achieving high-quality enterprise interconnection. A PoP supports multiple tenants and provides PoP access for multiple enterprises at the same time.
- Enterprises, as tenants, lease SD-WAN services provided by MSPs. An enterprise tenant can manage the SD-WAN services of all sites belonging to it, but it cannot view the SD-WAN services of other tenants.

HUAWEI

SD-WAN Business Model: Carrier Resale



- A carrier provides SD-WAN services for multiple enterprises through iMaster NCE-WAN.
- Enterprises, as tenants, lease SD-WAN services provided by carriers. An enterprise tenant can manage the SD-WAN services of all sites belonging to it, but it cannot view the SD-WAN services of other tenants. Enterprises either manage and control their SD-WAN services based on the tenant permissions assigned by carriers, or they can entrust their SD-WAN services to carriers for management and control.
- SD-WAN gateways are used to implement flexible interconnection and fast compatibility between SD-WAN networks and carriers' legacy backbone networks.

HUAWEI

Section Summary

- SD-WAN solves the problems of difficult management, and expensive construction and maintenance costs in multi-branch interconnection scenarios. It also guarantees the quality of key services.
- Huawei SD-WAN Solution provides the following functions:
  - ZTP, enabling service provisioning to be completed within 1 hour
  - Forwarding-control separation, enabling flexible networking
  - Application optimization, enabling service controllability and visibility
  - Complete security protection system, eliminating security risks
  - Visualized O&M for quick fault locating
- Huawei SD-WAN Solution is offered in three business models: enterprise-built, MSP resale, and carrier resale.

HUAWEI

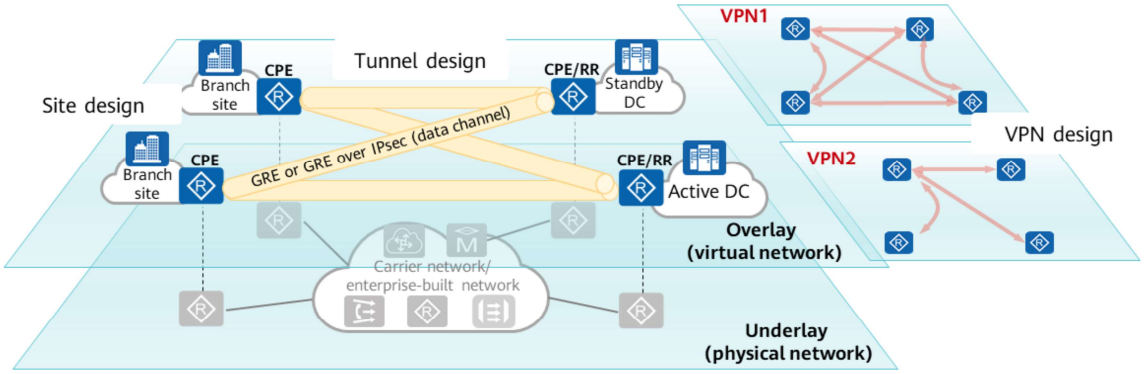
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. **Networking Design for Huawei SD-WAN Solution**
   - SD-WAN Networking Design Overview
   - SD-WAN Site Design
   - SD-WAN Tunnel Design
   - SD-WAN VPN Design
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

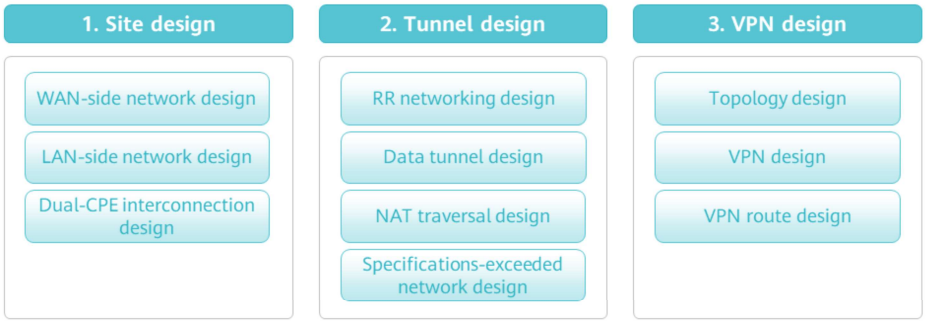
Overview of Networking Design for Huawei SD-WAN Solution

- SD-WAN networking design includes site design, tunnel design, and VPN design.



HUAWEI

Networking Design Process for Huawei SD-WAN Solution



HUAWEI

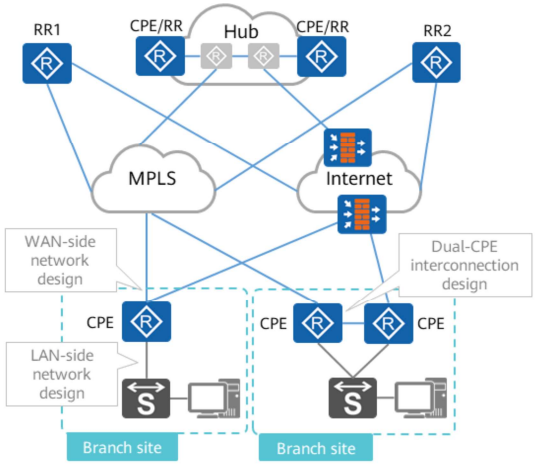
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. **Networking Design for Huawei SD-WAN Solution**
   - SD-WAN Networking Design Overview
   - SD-WAN Site Design
   - SD-WAN Tunnel Design
   - SD-WAN VPN Design
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

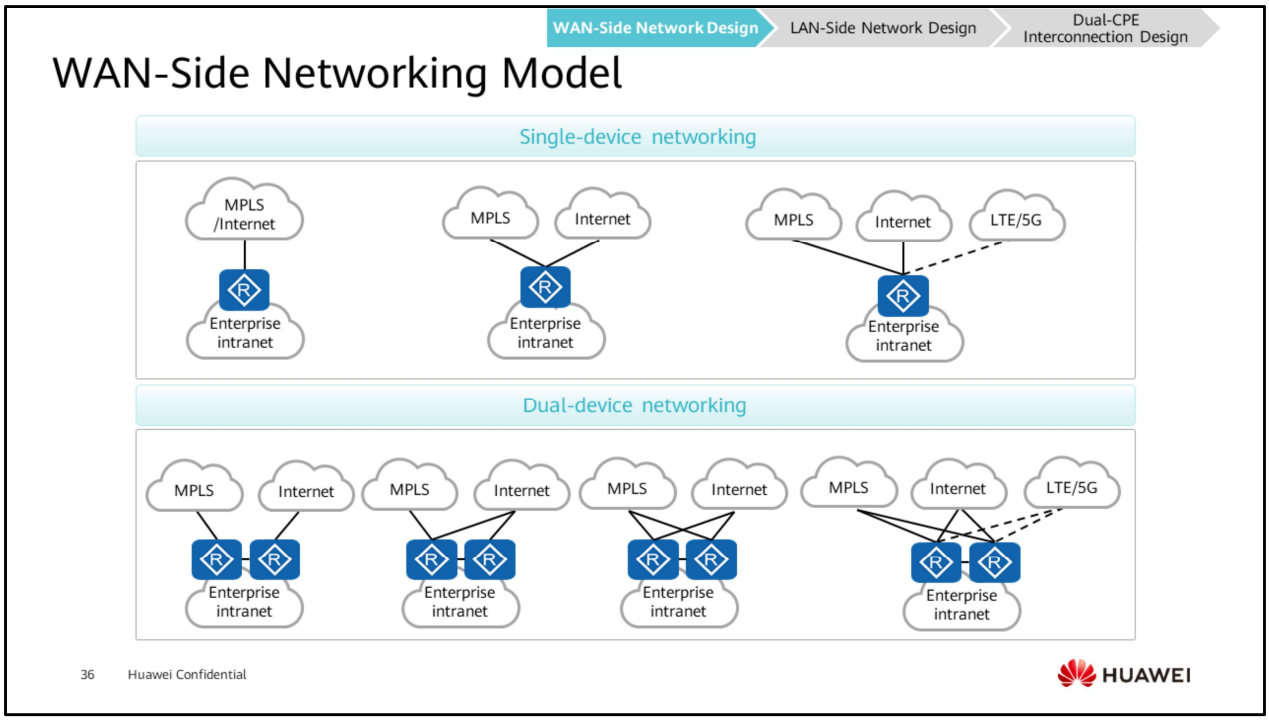
HUAWEI

Site Design Panorama

- SD-WAN sites refer to sites for which SD-WAN is deployed for interconnection, and are managed and monitored by iMaster NCE-WAN.
- Site design generally covers:
  - WAN-side network design
  - LAN-side network design
  - Dual-CPE interconnection design



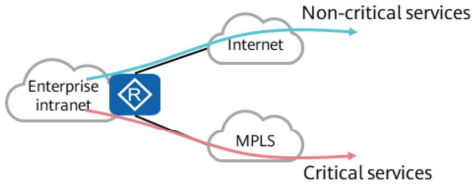
HUAWEI



- An SD-WAN site can be deployed with a single CPE or dual CPEs. For small sites, a single CPE can be deployed. For sites with high reliability requirements, dual CPEs are recommended to provide device-level redundancy.
- It is recommended that multiple links be deployed on the WAN side of a CPE. These links can back up each other, offering link-level reliability. In addition, services can easily select the active and standby links among these WAN-side links with differentiated quality.
- A maximum of 10 WAN links can be deployed for each CPE at an SD-WAN site. During actual deployments, to enhance reliability and facilitate O&M, it is recommended that a maximum of three WAN links be deployed for a single CPE at a site, and a maximum of six WAN links be deployed for a site with two CPEs.

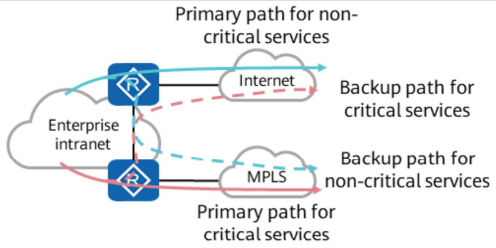
Typical WAN-Side Networking

Single-CPE multi-link networking



- For small and midsize sites or branch sites with low reliability requirements, only one CPE needs to be deployed. Multiple WAN links can be deployed and enterprise application traffic of different importance is transmitted over different links.
- Typically, one MPLS link and one Internet link are deployed on the WAN side. The MPLS link has guaranteed SLA and is used to carry enterprises' critical applications. Non-critical applications are carried on the Internet link.

Dual-CPE multi-link networking

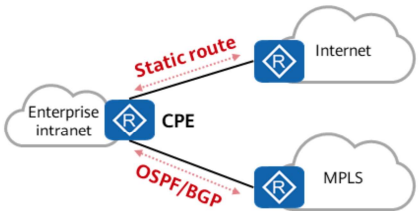


- In large sites or branch sites with high reliability requirements, dual CPEs are generally deployed. To ensure service transmission reliability, at least two WAN links are deployed for each CPE.
- Two CPEs are deployed at a branch site. The MPLS link is deployed for one CPE, and the Internet link is deployed for the other CPE. To ensure reliability, an interlink is deployed between the two CPEs to prevent data forwarding failures when a CPE fails.

HUAWEI

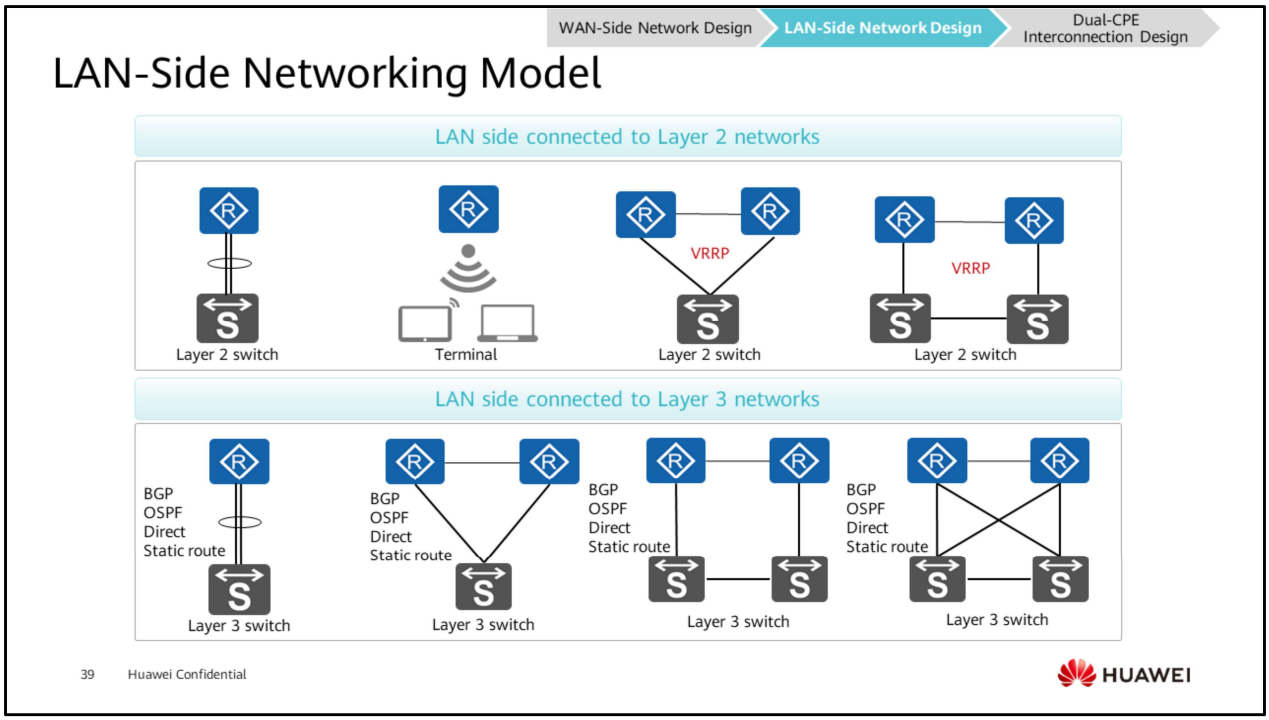
WAN-Side Underlay Route Design

BGP used to access WANs

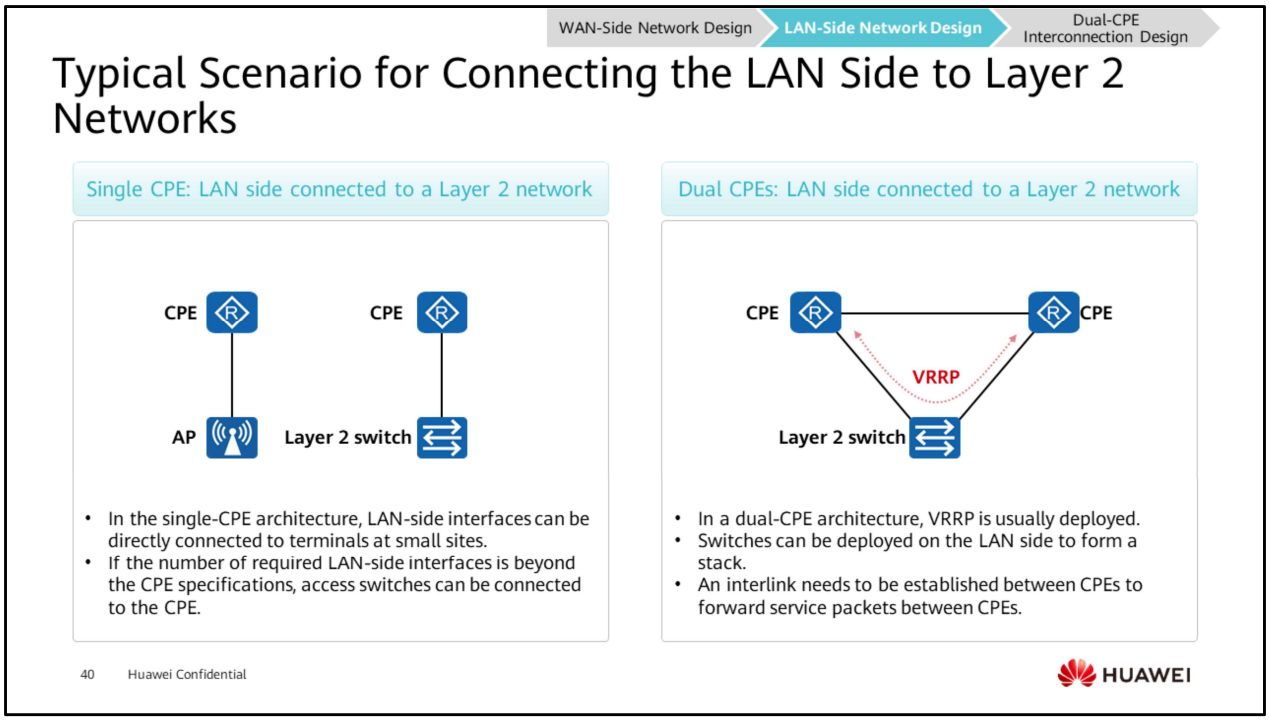


- WAN interfaces of CPEs support OSPF, EBGP, and static routing protocols. The routing protocol to be used must be the same as that of the LAN-side network device.
- On the live network, if static routes are used, services may be interrupted because WAN network faults on indirect links cannot be detected.

HUAWEI



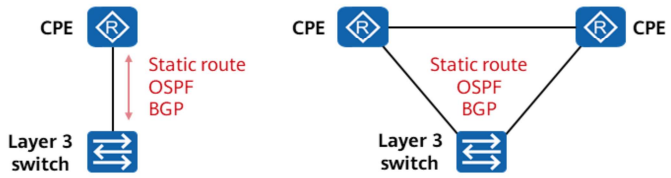
- LAN side connected to Layer 2 networks:
  - At small sites with a simple intranet structure, and CPEs typically connect to the intranet of the site at Layer 2.
- LAN side connected to Layer 3 networks:
  - In Layer 3 interconnection scenarios, SD-WAN routers support the direct connection, dual-homing, and partial-ring networking.
  - In such scenarios, routing protocols simply need to be configured for CPEs based on the requirements of the LAN device.



- In the single-CPE structure, the LAN connection is simple.
  - For small sites, for example, SOHO sites, LAN-side interfaces can be directly connected to terminals at the sites.
  - If the number of required LAN-side interfaces is beyond the CPE specifications, access switches can be connected to the CPE. In this case, access switches can be connected to the CPE in one-armed mode.
- In the dual-CPE architecture, VRRP is usually configured for the CPEs to prevent the dual-CPE architecture from affecting the LAN side.
  - Multiple switches can be deployed on the LAN side to form a stack. If two CPEs are deployed at a site, they can be interconnected directly or through the LAN.
  - CPEs can be directly connected through an Eth-Trunk link. In addition, an interlink needs to be established between the active and standby CPEs to forward service packets between the CPEs.

Typical Scenario for Connecting the LAN Side to Layer 3 Networks

Single or dual CPEs: LAN side connected to a Layer 3 network



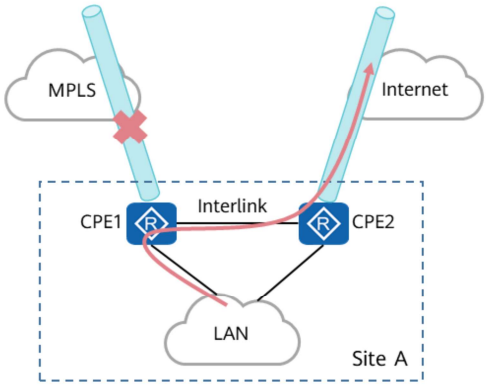
- In the Layer 3 interconnection scenario, if only one CPE is deployed, the network structure is simple. In such a scenario, only the routing protocol needs to be configured on the LAN side based on requirements of LAN-side devices.

HUAWEI

- For large enterprise sites, the network structure is complex as Layer 3 core devices are deployed on the network. Therefore, egress routers must support interconnection with Layer 3 devices. They can be connected to a Layer 3 network directly or in dual-homing mode. BGP, OSPF, and static routes are supported.
- In the Layer 3 interconnection scenario, if only one CPE is deployed, the network structure is simple. In such a scenario, only the routing protocol needs to be configured on the LAN side based on requirements of LAN-side devices.

Dual-CPE Interconnection Networking Model

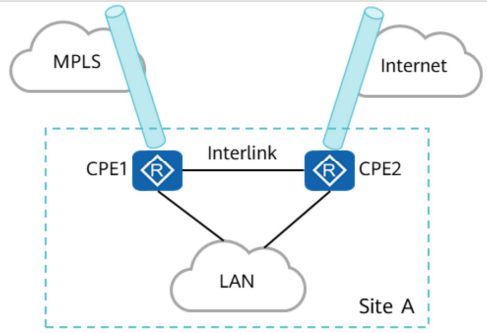
- In dual-CPE dual-link networking, if an uplink fails, the interlink between CPEs is used to transmit data.
- In addition to exchanging routes between two CPEs at a site, the interlink is also used to synchronize key information between the two CPEs, including link information, tunnel connection information, SLA information about tunnel connections, IPsec SA, and application identification result.
- With such information, resources on the two CPEs at each SD-WAN site can be regarded as a whole. Then services such as cross-device traffic steering can be implemented.



HUAWEI

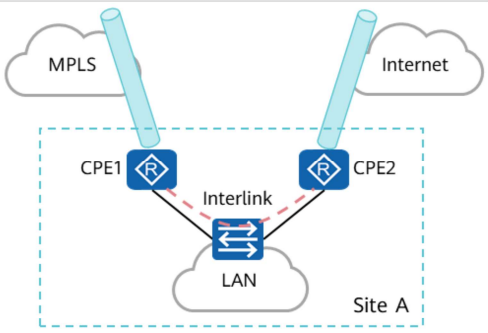
Dual-CPE Interconnection Networking Solution

Dual-CPE interconnection networking solution 1



- A direct link is used as the interlink for interconnection. If service traffic of a site is heavy and CPE interfaces are sufficient, using two interlink interfaces is recommended. iMaster NCE-WAN automatically orchestrates the two interlink interfaces into an Eth-Trunk interface to increase the interlink bandwidth and ensure interlink reliability.

Dual-CPE interconnection networking solution 2



- If CPE interface resources are limited, the interlink can be established between Layer 2 service interfaces on the LAN side of CPEs.

HUAWEI

- Solution 1 is recommended. In this solution, the interlink and service links are independent of each other. When WAN-side links are adjusted, the interlink will not be affected, and the service flow direction is clear.

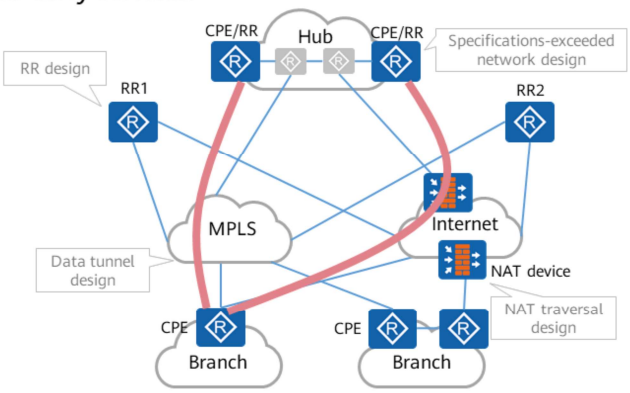
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. **Networking Design for Huawei SD-WAN Solution**
   - SD-WAN Networking Design Overview
   - SD-WAN Site Design
   - SD-WAN Tunnel Design
   - SD-WAN VPN Design
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

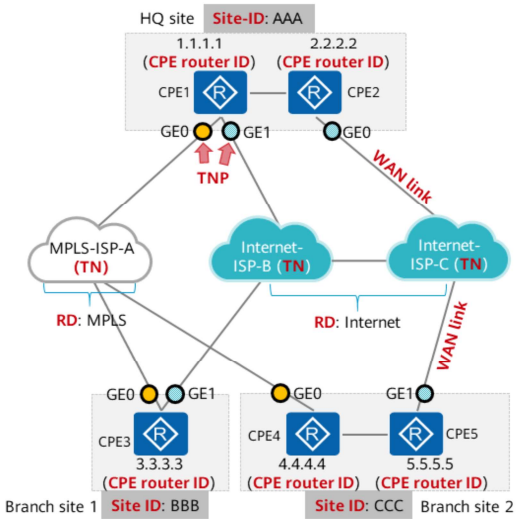
Tunnel Design Overview

- Tunnels in the SD-WAN Solution can be classified into: management tunnel, control tunnel, and data tunnel. Control tunnels are established between RRs and CPEs, and data tunnels are established between CPEs. Data tunnels carry services.
- Tunnel design generally covers:
  - Data tunnel design
  - RR design
  - NAT traversal design
  - Specifications-exceeded network design



HUAWEI

Basic Concepts of Tunnels



- **Transport network (TN)**: a carrier-provided WAN network that implements branch interconnection of enterprises over WANs.
- **Routing domain (RD)**: If different TNs are reachable to each other (for example, the Internet provided by ISP B and ISP C in the left figure can communicate with each other), they are considered to be in the same RD.
- **Site ID**: global unique identifier of a tenant site, which is allocated by iMaster NCE-WAN.
- **CPE router ID**: global unique identifier of a CPE at a site. A site can contain one or two CPEs. Generally, the device loopback address is used as the CPE router ID.
- **WAN link**: a link connecting to a WAN interface. A WAN link has a one-to-one mapping with a WAN interface. The IP address obtaining mode, link negotiation rate, and bandwidth can be configured for a WAN link.
- **Transport network port (TNP)**: WAN interface through which a CPE connects to a TN. Key information includes the site ID, CPE router ID, transport network ID, public IP address, private IP address, and tunnel encapsulation.

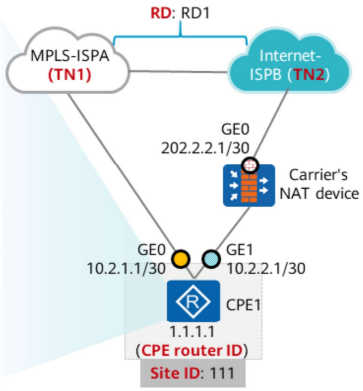
HUAWEI

- The TN and RD are used to set up overlay tunnels in enumerated mode.
- The site ID is used as the next hop of a routing entry.
- The CPE router ID is used to establish BGP peer relationships between different sites.
- TNPs are used to establish tunnels.

TNP

- A TNP mainly describes WAN link information of a site and is mainly used to establish control and data tunnels.
- The main information about a TNP is as follows:

TNP Information	Example		Description
Site ID	111		Site of the TNP.
CPE router ID	1.1.1.1		CPE to which the TNP belongs.
TNP ID	1	2	TNP ID.
Transport network	TN1	TN2	TN to which the TNP belongs.
Routing Domain	RD1	RD1	RD to which the TNP belongs.
Public IP	10.2.1.1	202.2.2.1	WAN-side public IP address (post-NAT).
Private IP	10.2.1.1	10.2.2.1	WAN-side private IP address (pre-NAT).
Encapsulation	GRE	GRE	Encapsulation mode.

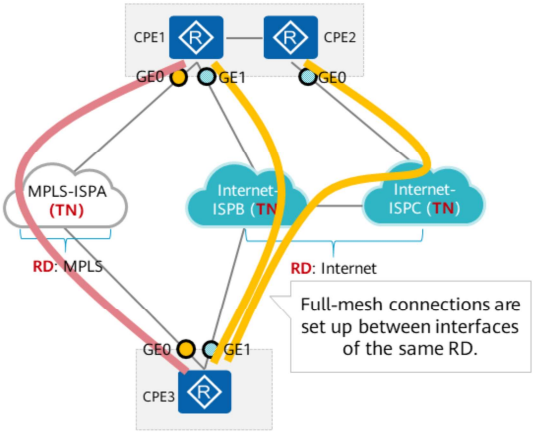


HUAWEI

- TNs and RDs are mainly used to enumerate tunnels.
  - Tunnel enumeration: All tunnels that can be established are enumerated.
- Router IDs are mainly used to establish control channels.
- A site ID is used as the next hop for data forwarding.
- A TNP ID can be considered as the interface number.
- The public and private IP addresses are used as the source or destination IP addresses of control and data channels.
  - Some CPEs are deployed behind the NAT device. To establish data channels between CPEs, you need to know the post-NAT public IP address.
    - CPEs typically use the Session Traversal Utilities for NAT (STUN) technology to detect public IP addresses.

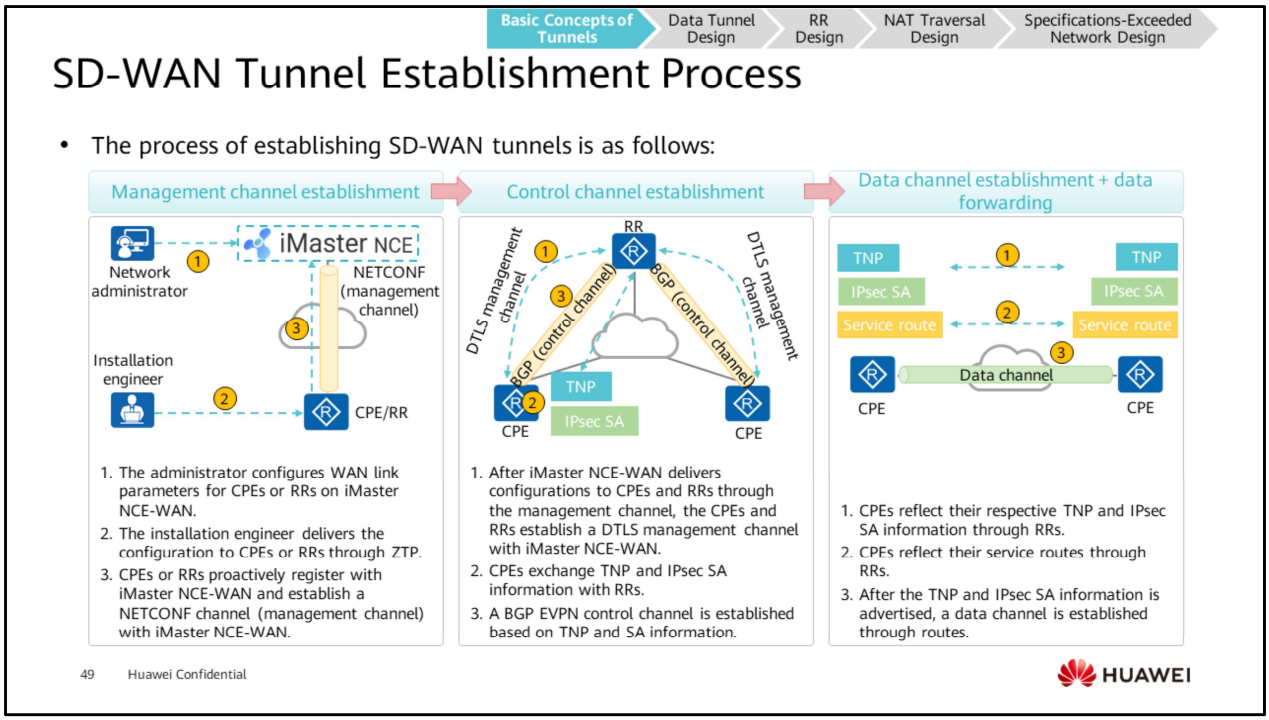
Tunnel Enumeration

- CPEs enumerate tunnels based on the **RD** and **TN** in TNP information.
- Tunnel enumeration ensures SD-WAN network reliability and guarantees service quality.
- Tunnel enumeration rules:
  - Connections can be enumerated only between TNPs in the same RD. Full-mesh connections are set up between interfaces of the same RD.
  - Connections cannot be enumerated between TNPs in different RDs.



HUAWEI

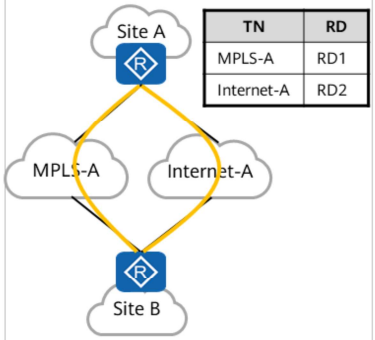
- Data tunnels are enumerated before being established to ensure that all available data tunnels are established.
- Tunnels can be enumerated only when the following conditions are met:
  - The CPE has learned service routes of the peer site.
  - The CPE has learned the TNP information of the peer site.



- The management channel is used to establish control channels and deliver basic configurations.
- Control channels are used to establish data channels.
  - Datagram Transport Layer Security (DTLS) is a security protocol for the packet transport layer, which is used to ensure TCP data security.
- Data channels are used to transfer user data.
- TNP information is exchanged twice:
  - During the control channel establishment phase, TNP information is exchanged to exchange information about channel establishment between RRs and CPEs.
  - During the data channel establishment phase, TNP information is exchanged to exchange information about channel establishment between CPEs.

Tunnel Networking Model

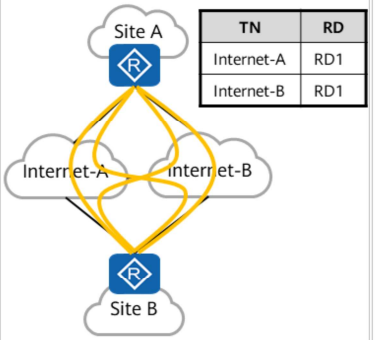
Different RDs



TN	RD
MPLS-A	RD1
Internet-A	RD2

- Sites A and B are connected through two links and are isolated on the underlay network. Different RDs are planned. Two tunnels are established between the sites.

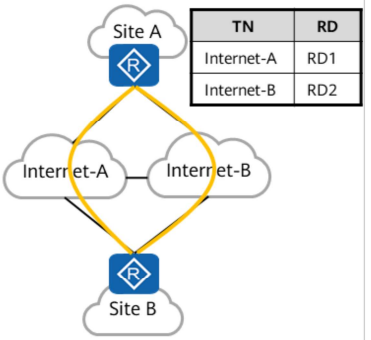
Different TNs, same RD



TN	RD
Internet-A	RD1
Internet-B	RD1

- Sites A and B are connected through two links and can communicate with each other on the underlay network. They are planned in the same RD and set up a full-mesh network.

Network in the same RD, different RDs created logically



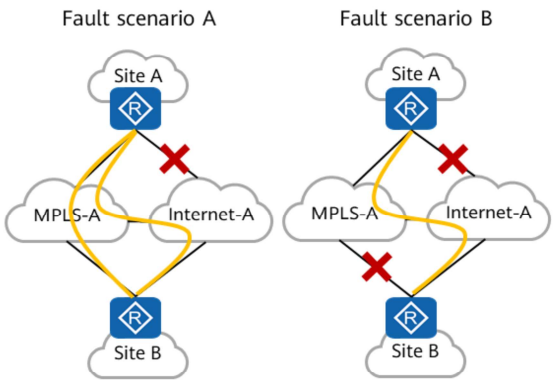
TN	RD
Internet-A	RD1
Internet-B	RD2

- Sites A and B are connected through two links and can communicate with each other on the underlay network. They are planned in different RDs. Two tunnels are established between the sites.

HUAWEI

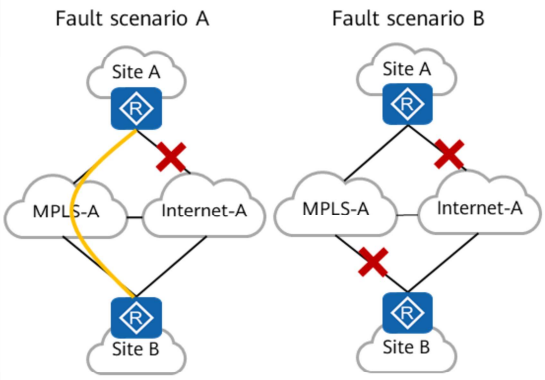
Tunnel Networking Reliability

Different TNs, same RD



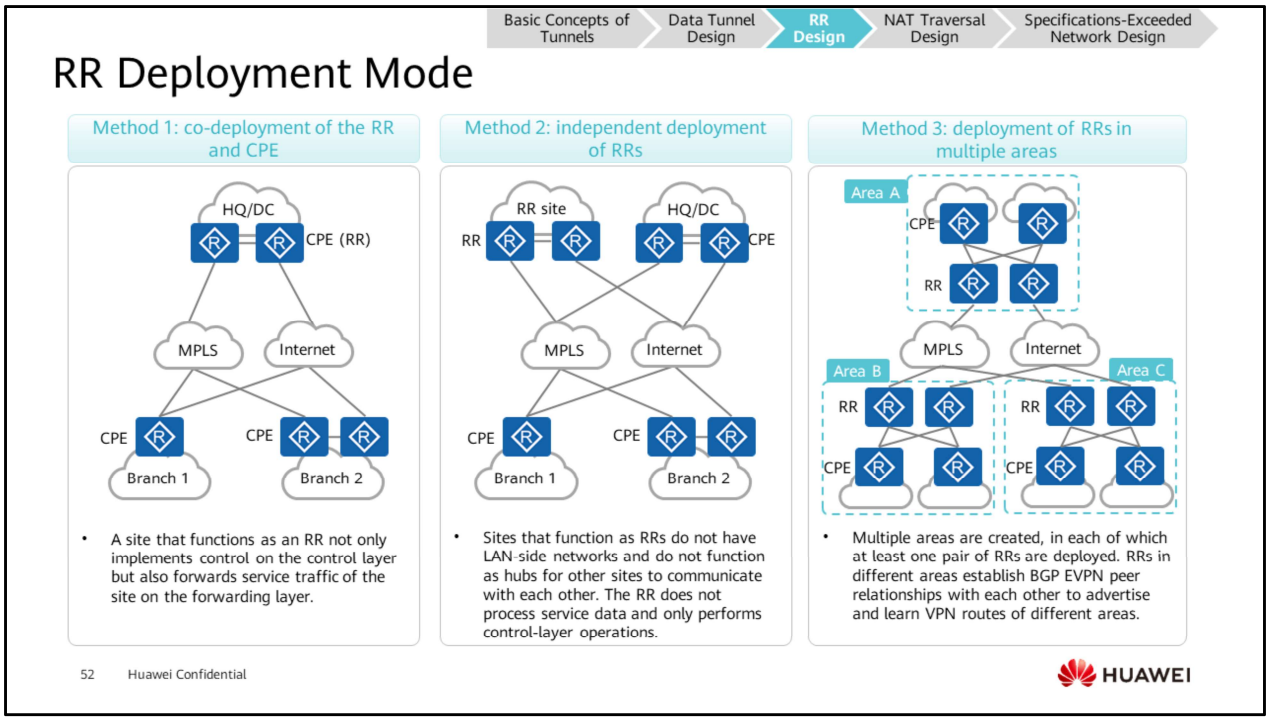
- **Same RD for different TNs**: The networking reliability is enhanced, but more virtual connection resources of the device are consumed.

Different RDs and TNs



- **Different RDs for different TNs**: The number of connections is reduced, and the networking scale is expanded. However, the networking is less reliable.

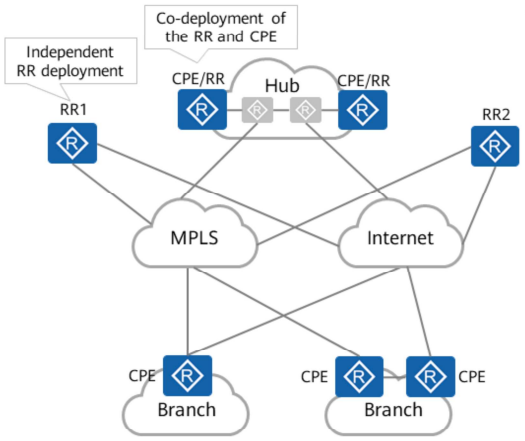
HUAWEI



- RR deployment suggestions are as follows:
  - To prevent site network adjustment from affecting the stability of RRs, you are advised to use method 2, that is, independent deployment of RRs.
  - Use high-performance devices as RRs. For details about the devices that can function as RRs, see the specifications list.
  - Configure a public IP address for an RR, or deploy a NAT device before the RR. Only 1:1 static NAT is supported.
  - An edge site can connect to a maximum of two RR sites. Two CPEs can be deployed at each RR site deployed for a tenant. If there are a large number of CPEs, multiple RRs can be deployed, and each RR serves some edge sites.
  - When one CPE is connected to two RR sites, the CPE establishes a BGP connection with each RR at the RR sites.
  - If a branch site has a standby link, for example, the branch or the RR has a standby link, and the active link is normal, no control channel is established for the standby link. When all active links from the branch site to the RR are down, the standby link is involved in the establishment of control channels.

- It is not recommended that standby links be established at RR sites.
- If the number of CPEs on the network exceeds the RR control specification, you are advised to assign sites to different areas. An independent RR is deployed in each area, and BGP peer relationships are established between RRs in each area to exchange routes between areas.
- When deploying RRs, consider the total number of routes on the LAN side of each site to prevent the route specifications of the RR from being exceeded.
- The BGP peer relationship is established between the CPE and RR using the loopback address. If multiple tunnel connections are available between the CPE and RR, a random connection is used to establish the BGP peer relationship. When a WAN link fails, BGP automatically switches services to another link, and services between sites are not affected. When all RRs are faulty, the connections with CPEs are unavailable, interrupting services between sites. Therefore, RR reliability must be ensured during actual deployment. For example, deploy two CPEs at an RR site, specify two RR sites for the edge, and use distributed RR sites (each RR site is responsible for some CPEs). This prevents network-wide service interruption caused because the only RR deployed on the network is faulty.

RR Deployment Principles



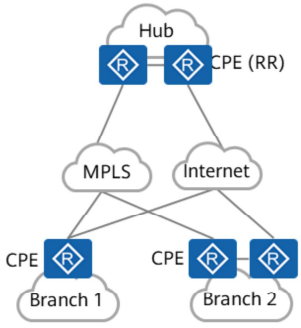
- Considering the importance of RRs, follow the principles below when deploying RRs:
  - Redundancy must be implemented for RRs. That is, at least two RRs must be deployed on the live network. This prevents network-wide services from being interrupted because the only RR deployed on the network fails.
  - An edge site must be connected to two RRs to implement egress backup.
  - It is recommended that RRs be independently deployed to ensure stability.
  - If RRs cannot be deployed independently, select routers at core positions as RRs and ensure that the device performance can meet the requirements.
  - Use the RR models recommended in the specifications list.

HUAWEI

- A CPE can be connected to a maximum of two RR sites (four RRs).
- For small networks (for example, a network with fewer than 50 sites), RRs and hubs can be deployed in co-located mode.
- RRs require strong BGP connection capabilities (number of BGP peers), large number of EVPN connections, and high route reflection capability and efficiency. In actual deployments, select the RR models recommended in the specifications list, for example, AR6300/AR6280.

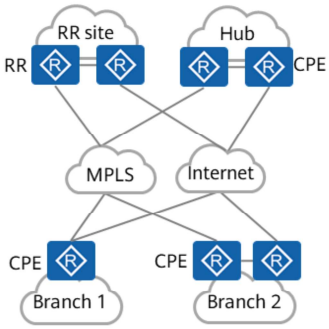
Typical RR Networking: Enterprise-Built

Co-deployment of the RR and CPE



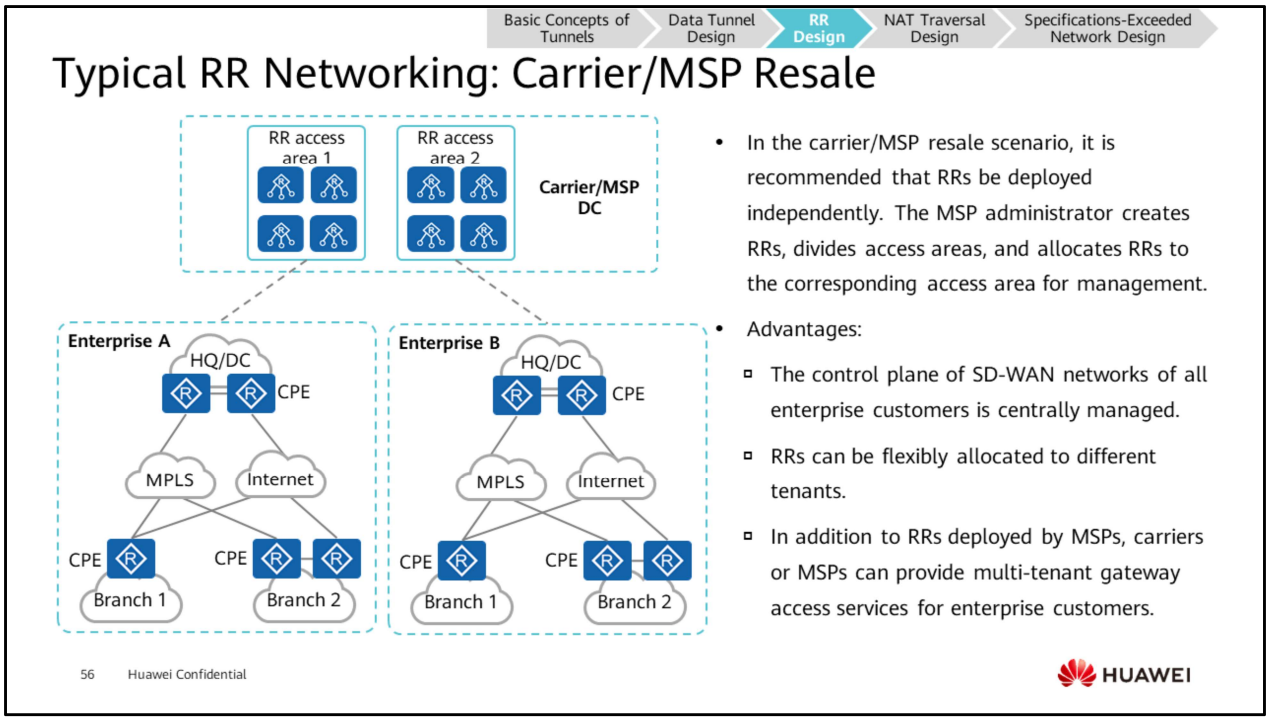
- Small and midsize enterprises have a small number of branch sites. If the traffic between branches and the HQ/DC is not heavy, it is recommended that the RR be co-deployed with the CPE. The hub site functions as the RR to carry the control plane.

Independent RR deployment



- Large enterprises have a large number of sites. If these enterprises have high network reliability requirements, independent RR deployment is recommended.

HUAWEI

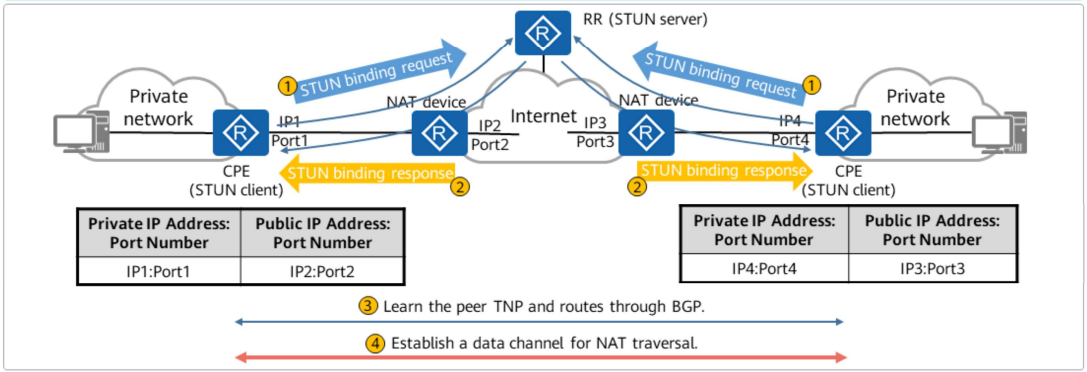


- An MSP administrator deploys independent RRs as a service provided by carriers or MSPs for enterprise users to access. Two RR service modes are available: sharing and exclusive. In sharing mode, one RR is shared by multiple tenants. In exclusive mode, one RR is exclusively used by a tenant.
    - MSP RR sharing by multiple tenants: For small and midsize enterprises, it is recommended that one RR be shared by multiple tenants, if there are a small number of sites. Sites of different enterprise tenants are connected to the same multi-tenant RR, which carries the control plane.
    - MSP RR exclusive to a tenant: Carriers can specify an independent RR in an access area for a large enterprise with a large number of sites to use exclusively. Based on actual scenarios, carriers or MSPs can deploy RRs in co-deployment mode, independently, and based on areas for tenants.

NAT Traversal

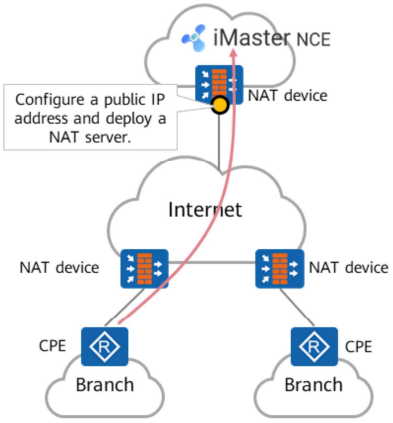
- On the live network, some enterprise branches use home broadband to access the network.
  - In this case, traffic needs to pass through the NAT device. Therefore, tunnels cannot be directly established for such enterprise branches.
- In SD-WAN scenarios, NAT traversal technology is required for interconnection between such enterprise branches.

Establishing a data channel for NAT traversal



HUAWEI

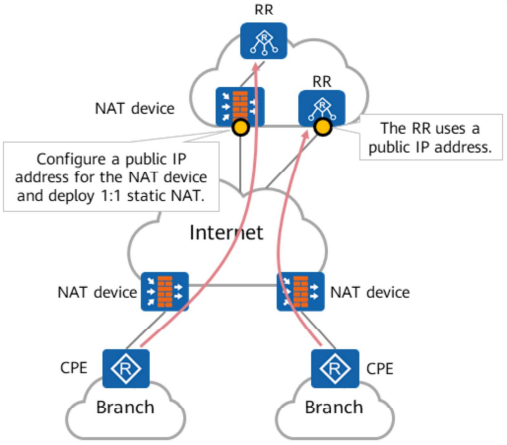
NAT Traversal: Connection Between CPEs and iMaster NCE-WAN



- A CPE proactively sends a registration request to iMaster NCE-WAN. To implement this, the southbound IP address of iMaster NCE-WAN must be reachable.
- On the live network, iMaster NCE-WAN may be deployed behind a NAT device. In this case, the NAT Server needs to be deployed on the NAT device and the following ports need to be mapped:
  - 10020: southbound port, which is used by devices to register with iMaster NCE-WAN.
  - 10031: southbound port, which is used by devices to connect to iMaster NCE-WAN through HTTP.
  - 18008: northbound port, which is used for web login.
  - 80: northbound port, which is used for web login.
  - 18018: northbound port, which is used by PCs to upload files to iMaster NCE-WAN.
  - 18021: file server port, which is used by iMaster NCE-WAN to update files.
- The NAT Server can be deployed. Therefore, NAT traversal is not involved in this scenario.

HUAWEI

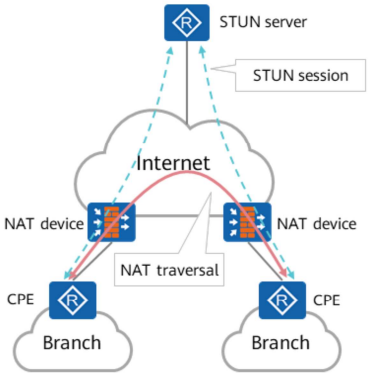
NAT Traversal: Connection Between CPEs and RRs



- An RR must have a public IP address or be configured with 1:1 static NAT. Therefore, NAT traversal is not involved in communication between CPEs and RRs.

HUAWEI

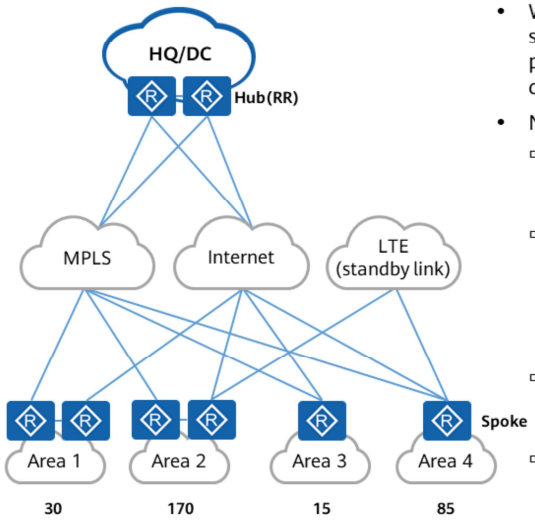
NAT Traversal: Connection Between CPEs



- A data channel is established between CPEs, and traffic between sites is transmitted through the data channel. If a CPE is deployed on the private network behind the NAT device, NAT traversal is required for communication with another CPE, especially in scenarios where two CPEs at two sites are deployed on the private network behind the NAT device.

HUAWEI

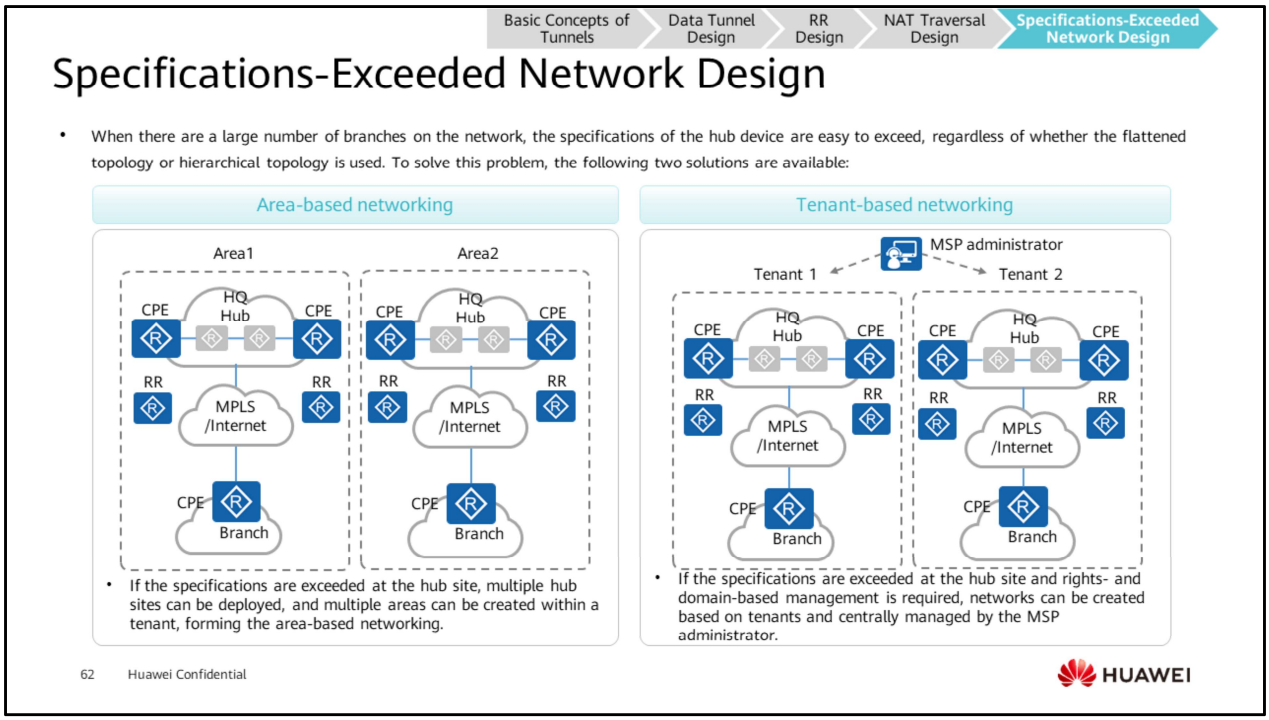
Networking Specification Calculation



- When planning a network, consider the device performance specifications. In the SD-WAN Solution, the number of BGP peers, number of tunnels, and bandwidth also need to be considered.
- Networking specification calculation
  - Number of BGP peers of hubs/RRs = Number of sites with dual gateways x 2 + Number of sites with a single gateway
    - Number of BGP peers = 200 x 2 + 100 = 500
  - Number of hub tunnels = Total number of tunnels of each gateway at a hub site
    - The two hubs at a dual-hub site share the tunnel specification. For example, if a single CPE supports 1000 tunnels, the two hubs at a dual-hub site share the tunnel specification.
    - Number of hub tunnels = 300 x 4 = 1200
  - Network bandwidth of a hub site = Number of branch sites x Bandwidth required by sites
    - Total network bandwidth required by the hub = 300 x 10 Mbps = 3 Gbps
  - iMaster NCE-WAN can manage a large number of sites. In most cases, you do not need to pay attention to the networking specification calculation.

HUAWEI

- Different device models have different BGP peer specifications and tunnel specifications.
- For details about the product specifications, see the product documentation.



- Area-based networking:
  - Multiple areas are created under a tenant, and multiple hub sites are deployed in the HQ/DC. Each area is associated with one or two hub sites.
  - Branch sites are added to corresponding hub sites by area.
  - RRs can be independently deployed. Each pair of RRs are associated with sites in an area.
  - Traffic between inter-area sites is transmitted through the LAN side of the hub.
- Tenant-based networking:
  - The MSP administrator creates multiple tenants, and multiple hub sites are deployed in the HQ/DC. Each tenant is associated with one or two hub sites.
  - Branch sites are grouped by their geographical areas and are added to different tenants.
  - RRs can be independently deployed. Each pair of RRs are associated with sites in an area.
  - Traffic between inter-area sites is transmitted through the LAN side of the hub.

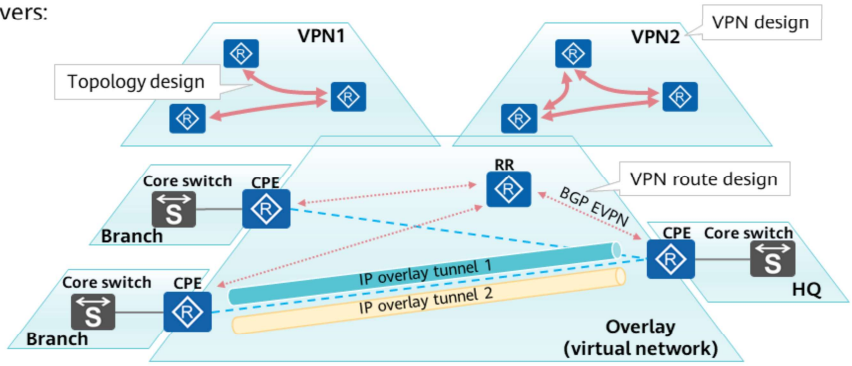
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. **Networking Design for Huawei SD-WAN Solution**
   - SD-WAN Networking Design Overview
   - SD-WAN Site Design
   - SD-WAN Tunnel Design
   - SD-WAN VPN Design
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

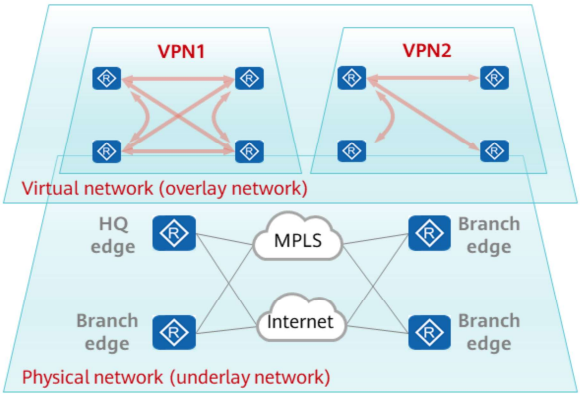
VPN Design Panorama

- The SD-WAN Solution uses VPNs to isolate services for multiple departments of a single tenant. In this case, the network of each department under a tenant is an independent service VPN.
- VPN design generally covers:
  - VPN design
  - Topology design
  - VPN route design



HUAWEI

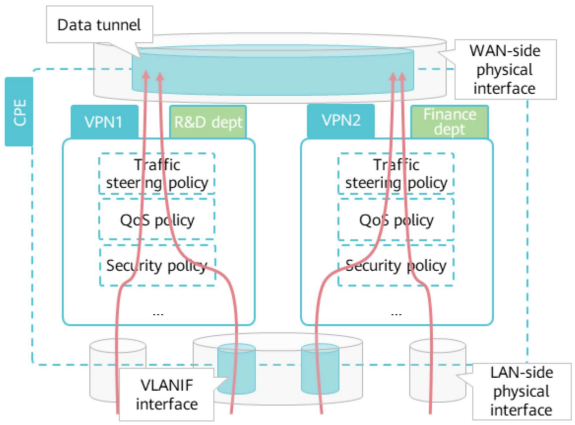
SD-WAN VPN



- VPN is short for virtual private network on SD-WAN. Each VPN is an independent IP Layer 3 private network. Multiple VPNs are logically isolated from each other, from the tunnel established at a site to the CPE at the site. Therefore, these VPNs cannot directly communicate with each other.
  - Each VPN can use an independent topology model, including hub-spoke, full-mesh, partial-mesh, and hierarchical networking.
  - Service policies, such as traffic steering policies and QoS policies, can be independently configured for each VPN.

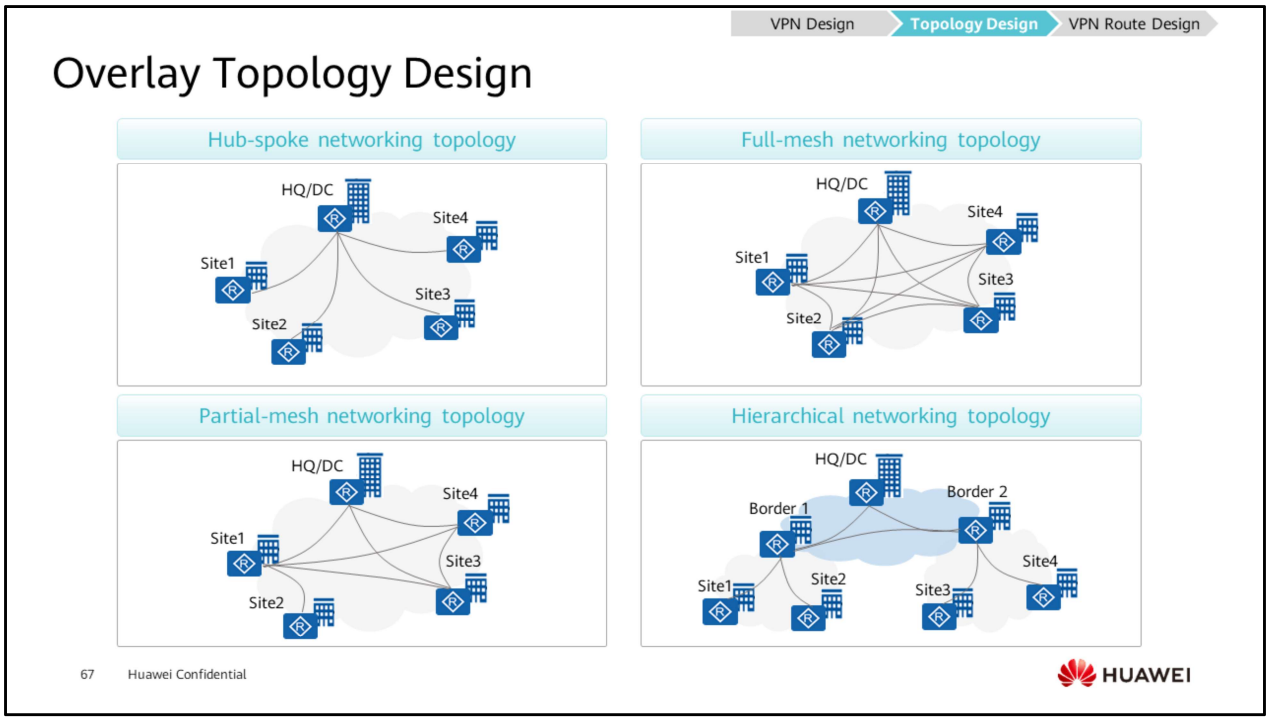
HUAWEI

VPN Planning and Design



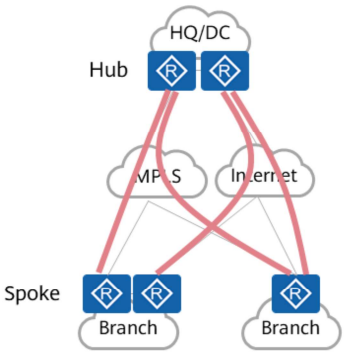
- VPN planning suggestions:
  - Determine the number of departments, for example, R&D department, finance department, and marketing department, whose services need to be isolated based on service isolation requirements.
  - Assign LAN-side physical interfaces and VLANs to different departments. One department can have an independent LAN-side physical interface, or all departments share one physical interface and are isolated through VLANs.
  - Plan initial policies for different departments, including traffic steering policies, Internet access policies, QoS policies, legacy site access policies, and URL, IPS, and firewall policies.

HUAWEI



- Hub-spoke:
  - Generally, the enterprise HQ/DC functions as a hub site, and enterprise branches function as spoke sites. Server applications deployed in the HQ/DC are accessed through the WAN in a centralized manner.
- Full-mesh:
  - In the topology, different branches can directly communicate with each other, without the need to divert traffic through intermediate nodes.
- Partial-mesh:
  - Partial-mesh can be considered as a special type of full-mesh networking. If an underlay network is available between two sites, traffic can be directly transmitted between the two sites. Otherwise, the two sites communicate with each other through a redirect site.
- Hierarchical networking:
  - The hierarchical networking model can be considered as the combination of the single-layer networking model. A WAN is divided into multiple areas, which are interconnected through a centralized backbone area to implement inter-area communication between sites.

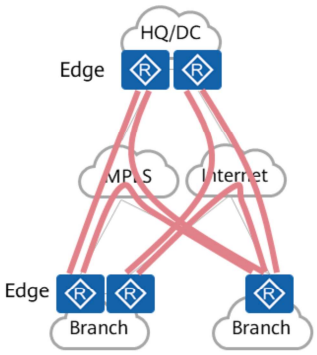
Hub-Spoke Networking



- **Solution Overview**
  - Generally, the enterprise HQ/DC functions as a hub site, and enterprise branches function as spoke sites. Server applications deployed in the HQ/DC are accessed through the WAN in a centralized manner.
  - If branches of an enterprise need to communicate with each other, traffic between them is transmitted through the hub site. All external access traffic of branch sites is first sent to the hub site.
- **Application Scenario**
  - This mode is generally applicable to enterprises where traffic is mainly sent from branch sites to the hub site.
  - Enterprise applications are centrally stored on servers at the HQ/DC. The main service traffic is from branch sites to the hub site.
  - Only small traffic is transmitted between branch sites.
  - A typical example is chain stores. The major traffic of a chain store is destined to the HQ/DC, and there is almost no traffic between chain stores.

HUAWEI

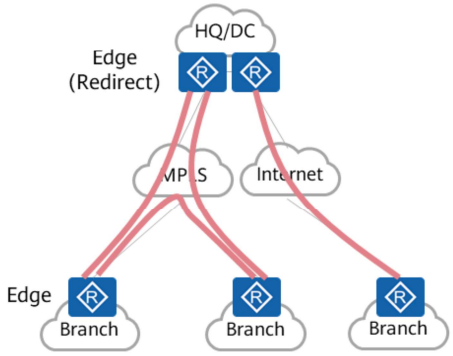
Full-Mesh Networking



- **Solution Overview**
  - In the full-mesh topology, branches can directly communicate with each other, without the need to divert traffic through intermediate nodes.
- **Application Scenario**
  - The full-mesh topology is applicable to small enterprises with a small number of sites or large enterprises whose branches need to collaborate with each other.
  - Collaborative services, for example, high-value applications including VoIP and video conferencing, have stringent requirements on network performance such as the packet loss rate, delay, and jitter. To meet requirements of such services, branches are recommended to directly communicate with each other.
  - The full-mesh topology is simple, and service communication is efficient. However, the full-mesh topology has average network scalability and applies to a network with 10 to 100 branch sites.

HUAWEI

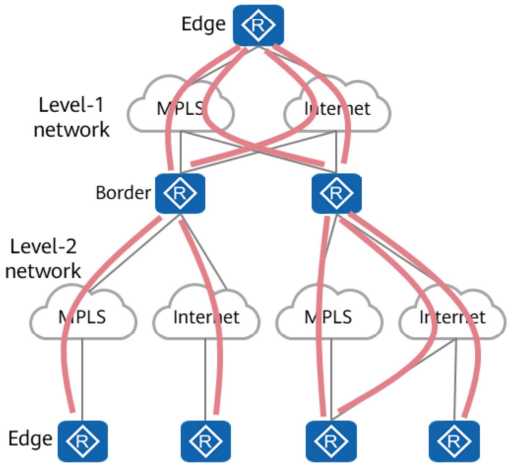
Partial-Mesh Networking



- **Solution Overview**
  - A partial-mesh network can be considered a type of special full-mesh network. If direct underlay network connections are available between sites, traffic is directly transmitted between the sites. Otherwise, traffic between sites is forwarded through a redirect site, to which both sites are connected.
- **Application Scenario**
  - Branch sites cannot directly communicate with each other over the underlay network and use a redirect site for communication instead.
  - In full-mesh networking, branch sites directly communicate with each other over the underlay network. To enhance reliability, a redirect site can be deployed. When the underlay network is faulty and sites cannot directly communicate with each other, traffic between branch sites is transmitted through the redirect site.

HUAWEI

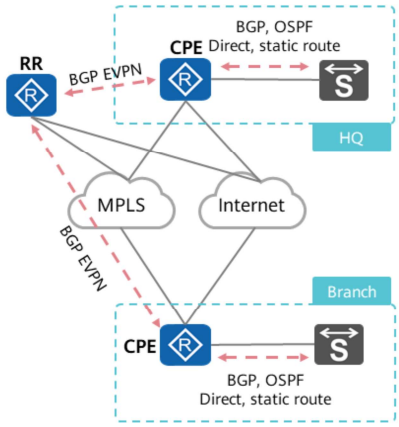
Hierarchical Networking



- **Solution Overview**
  - The hierarchical networking model can be considered as the combination of the single-layer networking model. A WAN is divided into multiple areas, which are interconnected through a centralized backbone area to implement inter-area communication between a large number of sites.
  - For example, a multinational enterprise can be divided into multiple areas (China, Europe, America, etc.) based on its management structure. Each area uses a single-layer networking model, which can be hub-spoke or full-mesh. In addition, each area uses one or more sites as their border sites. The border sites of each area form the backbone area for interconnection between areas, that is, the level-1 network. In this way, border sites connect to both the level-2 area network and level-1 network.
- **Application Scenario**
  - The hierarchical networking is applicable to enterprises with a large number of sites or multinational enterprises with widely distributed sites. The hierarchical networking has a clear network structure and good network scalability.

HUAWEI

VPN Route Design



- The VPN routes of the local CPE are collected using the LAN-side routing protocol and then sent to the peer CPE through the BGP EVPN protocol of the WAN.
  - LAN-side routes: LAN-side interfaces of CPEs support OSPF, BGP, and static routing protocols. The routing protocol to be used must be the same as that of the LAN-side network device.
  - WAN-side overlay routes: Generally, BGP EVPN is deployed to establish BGP peer relationships with RRs and advertise overlay network domain routes. The configuration is automatically orchestrated by the controller.
  - LAN- and WAN-side routes can be filtered using routing policies.

HUAWEI

Section Summary

- SD-WAN networking design includes site design, tunnel design, and VPN design.
  - Site design: includes WAN/LAN networking design and dual-CPE design.
  - Tunnel design: includes data tunnel design, RR design, NAT design, and specifications-exceeded network design.
  - VPN design: includes VPN design, topology design, and VPN route design.

HUAWEI

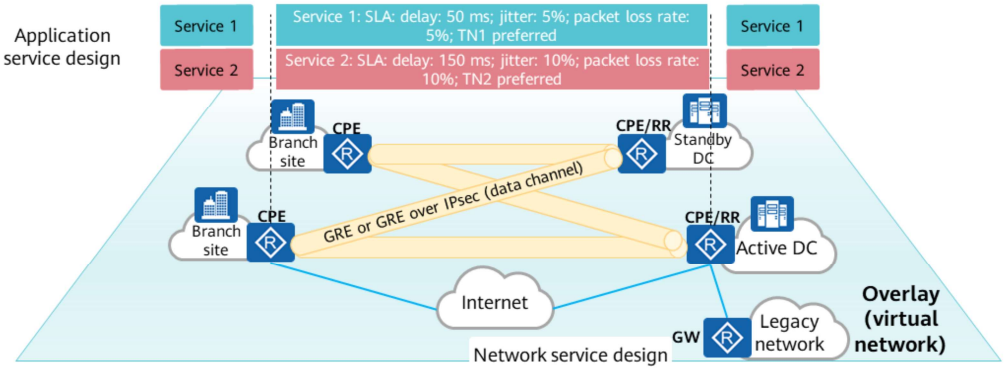
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. **Service Design for Huawei SD-WAN Solution**
   - SD-WAN Service Design Overview
   - SD-WAN Application Service Design
   - SD-WAN Network Service Design
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

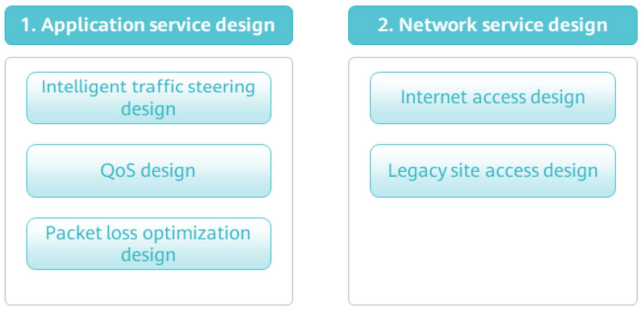
Overview of Service Design for Huawei SD-WAN Solution

- SD-WAN service design includes application service design and network service design.



HUAWEI

Service Design Process for Huawei SD-WAN Solution



HUAWEI

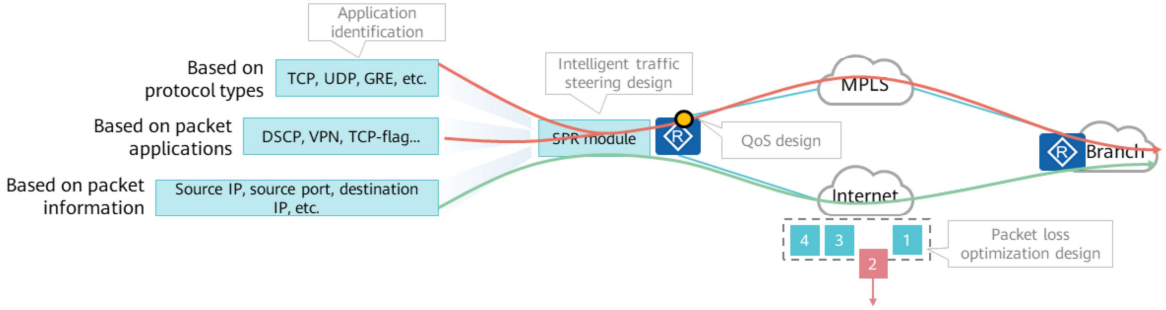
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. **Service Design for Huawei SD-WAN Solution**
   - SD-WAN Service Design Overview
   - SD-WAN Application Service Design
   - SD-WAN Network Service Design
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

Application Service Design Panorama

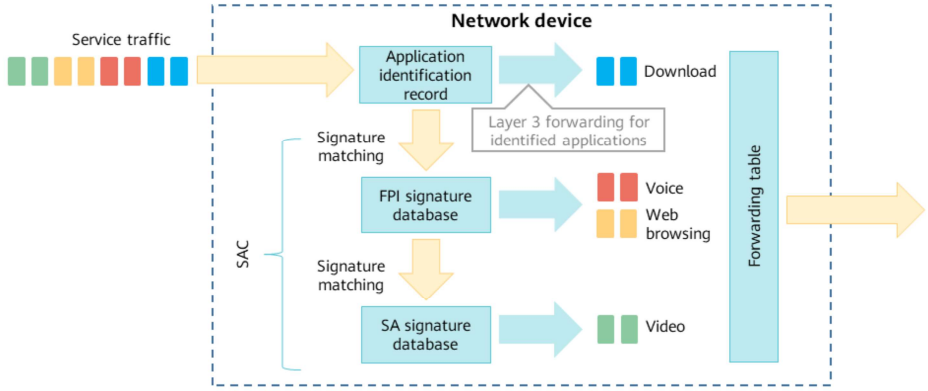
- Enterprises have diverse applications, such as production, coordination, and entertainment applications. Different applications have different requirements on bandwidth and link quality.
- Application service design generally covers:
  - Intelligent traffic steering design
  - QoS design
  - Packet loss optimization design



HUAWEI

Application Identification

- Traditional routing and switching devices cannot identify application-layer information. Therefore, it is difficult to manage networks based on applications. Smart Application Control (SAC) technology helps routing and switching devices identify and classify applications.
- SAC first checks whether an application has been identified. If an application has not been identified, SAC checks the first-packet inspection (FPI) signature database and service awareness (SA) signature database in sequence.

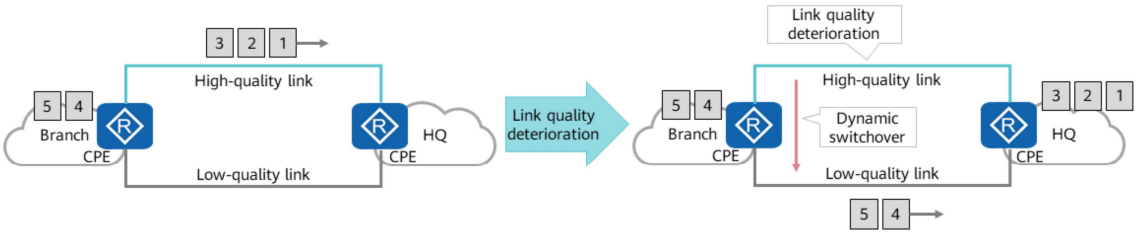


HUAWEI

- After a packet enters the device, the device determines whether the corresponding application has been identified based on the 5-tuple information carried in the packet. If the application has been identified, the device forwards the packet at Layer 3 without identifying the application again. If the application has not been identified, the device performs the SAC application identification process. The device then processes the packet based on the SAC identification result and forwards the packet at Layer 3.
- SAC identifies an application as follows: SAC first identifies an application according to ACL rules in FPI. If the application fails to be identified, the application is identified according to the DNS association table in FPI. If the application fails to be identified, the device identifies the application based on the protocol port table in the FPI. If the application still fails to be identified, the device attempts to identify the application using SA.

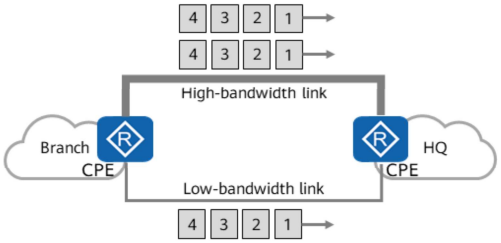
Intelligent Traffic Steering Design: Link Quality-based Traffic Steering

- Different applications have different requirements on link quality. For example, voice and video services have stringent requirements on the delay and packet loss rate. Generally, the delay must be within 150 ms and the packet loss rate must be less than 1% for these services.
- To guarantee voice and video services, the MPLS link that offers good quality can be configured as the active link for the services, with the Internet link functioning as the standby link. In addition, you can configure SLA requirements for the services to implement traffic steering based on the link SLA.



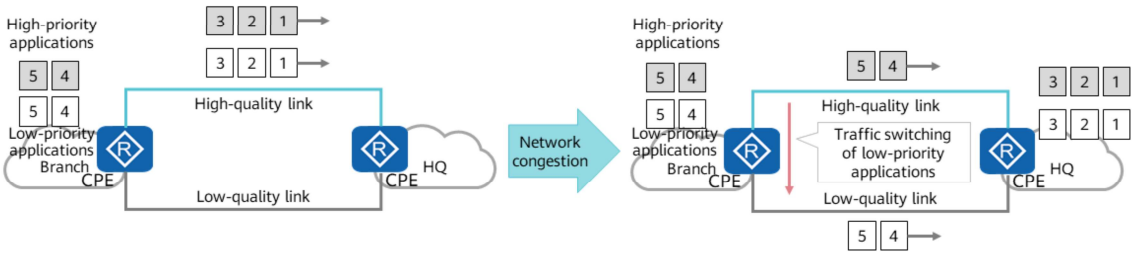
Intelligent Traffic Steering Design: Load Balancing-based Traffic Steering

- If an enterprise with multiple links wants to fully utilize the link bandwidth resources and steer traffic in load balancing mode based on the link bandwidth, load balancing-based traffic steering can be configured.
- For example, the two MPLS links can be configured as primary links for the voice service. If the quality of both links meets the SLA requirements of the voice service, voice service flows can be carried over the two MPLS links in load balancing mode. Through real-time bandwidth utilization monitoring, the bandwidth utilization can reach 85%, fully utilizing the link bandwidth.

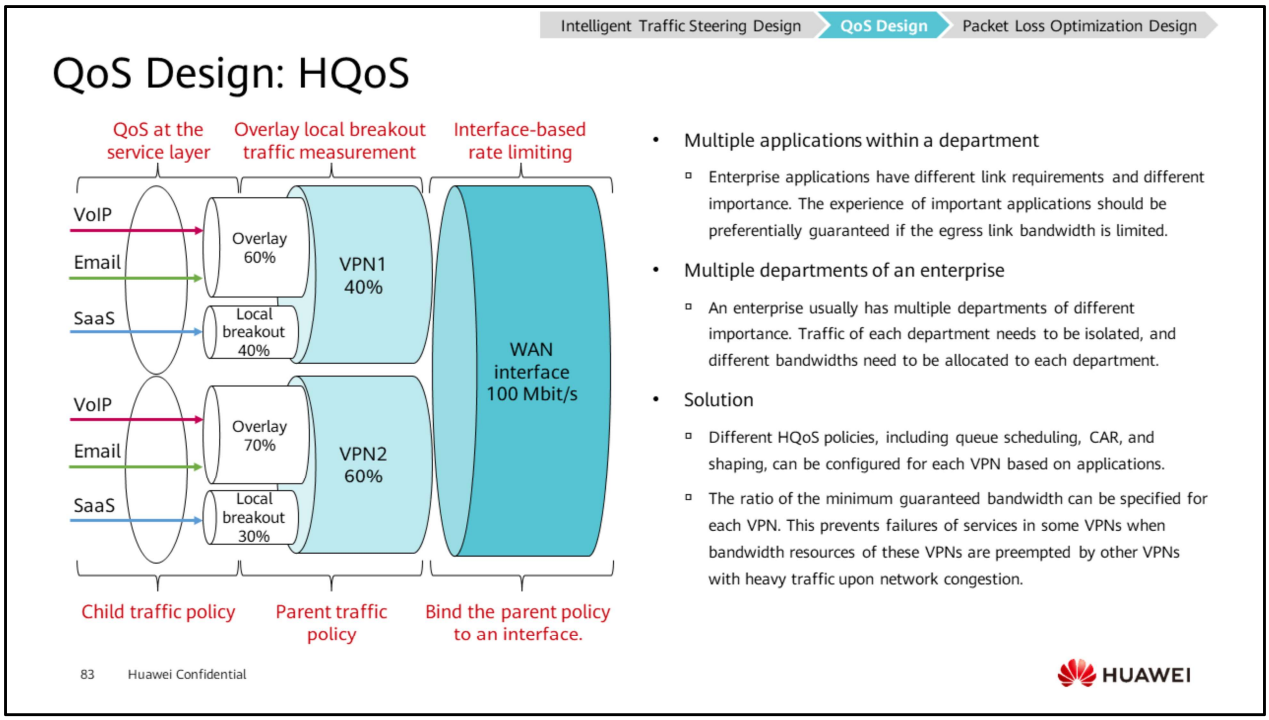


Intelligent Traffic Steering Design: Application Priority-based Traffic Steering

- If multiple types of service packets are transmitted on the same link, application priority-based traffic steering can be used, so that traffic of high-priority applications is preferentially processed when congestion occurs, ensuring experience of high-priority applications.
- For example, voice, video, and file transfer services are carried on an MPLS link. If the link bandwidth is insufficient, the experience of the voice and video services is preferentially guaranteed.



HUAWEI



- An enterprise usually has multiple departments of different importance. Traffic of each department needs to be isolated, and different bandwidths need to be allocated to each department.
    - A specified bandwidth quota is assigned to each department to meet its service requirements.
    - If some departments do not fully use their bandwidth quotas, the idle bandwidth resources can be used by other departments with insufficient bandwidth.
    - Internet access traffic and traffic for communication with legacy sites needs to be controlled separately.

Packet Loss Optimization Design

- Forward error correction (FEC) technology can mitigate packet loss. During data transmission, FEC adds redundant packets carrying check information at the transmit end and performs check at the receive end. If a packet is lost or damaged on the network, the redundant packet can be used to recover it.
- The following FEC modes are supported:
  - Adaptive forward error correction (A-FEC): A CPE automatically detects the packet loss rate and adjusts the number of redundant packets in real time based on the packet loss rate. This mode reduces the number of redundant packets and saves bandwidth when the packet loss rate is low. When the packet loss rate increases, redundant packets are increased to improve the packet loss mitigation performance. However, if the packet loss rate suddenly increases greatly, it takes a period of time for the CPE to detect the event and adjust the number of redundant packets. As a result, packet loss occurs for a short period of time.
  - Determined forward error correction (D-FEC): The number of redundant packets is calculated based on the configured packet loss rate. This mode can recover the lost packets within the configured maximum packet loss rate. Although the redundancy of this mode is greater than that of A-FEC, packet loss does not occur when the packet loss rate suddenly changes.
- After FEC is deployed and the network runs for a period of time, you can view the packet loss mitigation effect on iMaster NCE-WAN to determine whether the FEC design is proper.

HUAWEI

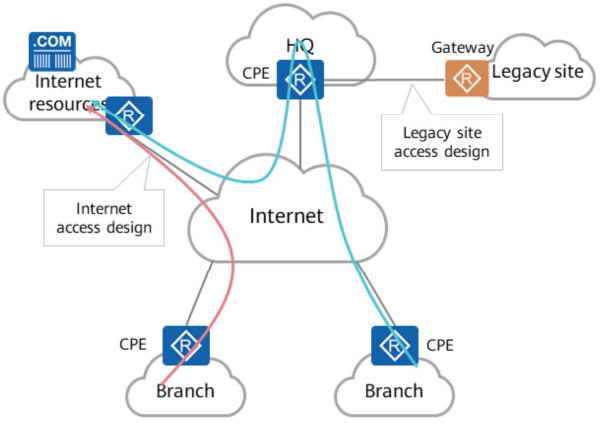
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. **Service Design for Huawei SD-WAN Solution**
   - SD-WAN Service Design Overview
   - SD-WAN Application Service Design
   - SD-WAN Network Service Design
5. Reliability and Security Design for Huawei SD-WAN Solution

HUAWEI

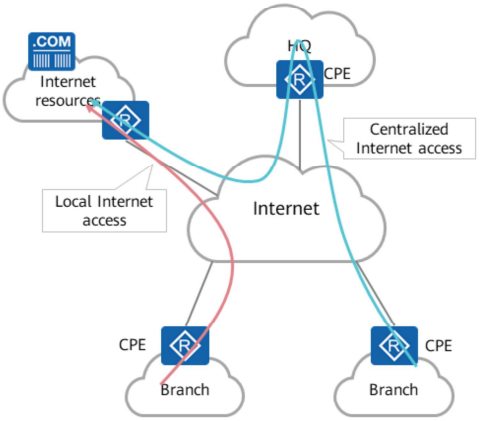
Network Service Design Panorama

- SD-WAN sites may also need to access the Internet or legacy sites. Therefore, specific network services need to be deployed.
- Network service design generally covers:
  - Internet access design
  - Legacy site access design



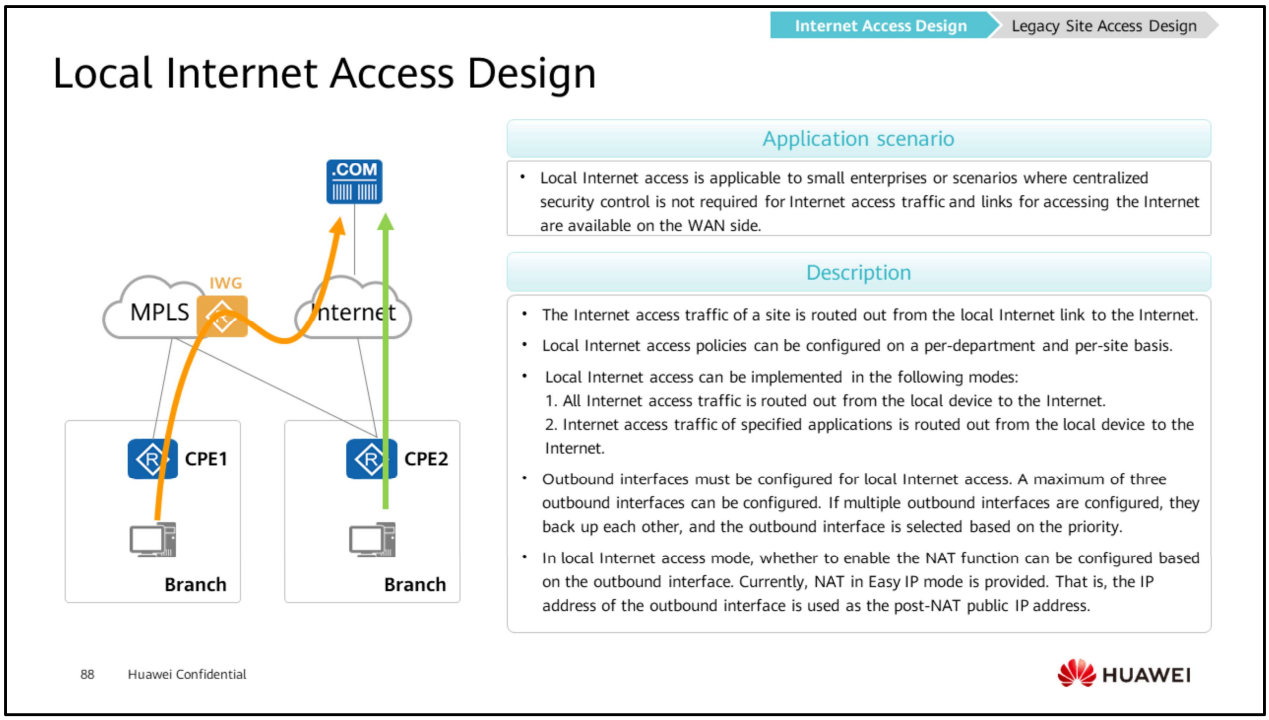
HUAWEI

Internet Access Solution



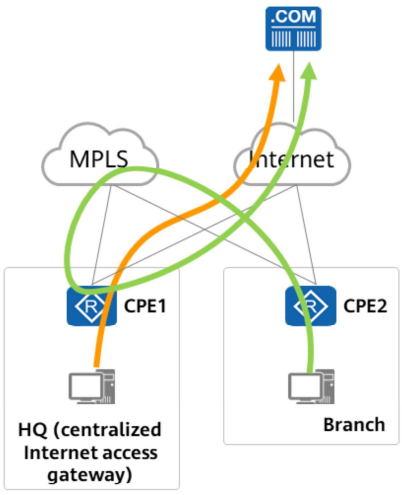
- The SD-WAN Solution provides the following Internet access modes:
  - Local Internet access: The Internet access traffic of a site is routed from the local CPE to the Internet.
  - Centralized Internet access: The Internet access traffic of all sites is diverted to the centralized Internet access site and then to the Internet.
  - Hybrid Internet access: combination of local Internet access and centralized Internet access.
    - Local Internet access (default) + centralized Internet access: By default, all Internet access traffic is routed out from the local device. When the local Internet access interface is faulty, Internet access traffic is forwarded through the centralized gateway.
    - Centralized Internet access (default) + Local Internet access for specified traffic: By default, Internet traffic is routed out through the centralized Internet access site. Traffic of some specified services is directly routed to the Internet through the local WAN-side link.

HUAWEI



- Traffic cannot be transmitted to the Internet through multiple links in load balancing mode. The links can work only in active/standby mode based on their priorities.
- If local Internet access is configured for specified application traffic and centralized Internet access is also configured, local Internet access for specified application traffic is implemented by orchestrating policy-based routing (PBR).
- If local Internet access is enabled, the default route on the underlay WAN needs to be configured. The default route can be a static route (mainly for Internet access through the Internet network interface) or BGP/OSPF route (mainly for Internet access through the MPLS network interface).

Centralized Internet Access Design



Application scenario

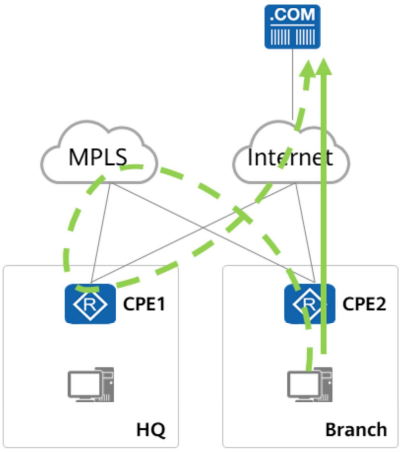
- This mode is applicable to scenarios where no link is available for a site to access the Internet or centralized security control is required for Internet access traffic. A centralized Internet access gateway is configured so that traffic from other sites can be forwarded to the centralized Internet access gateway through the overlay network.

Description

- All sites of a tenant access the Internet through a centralized Internet access site.
- Either of the following solutions can be used for Internet access through the centralized gateway site:
  - If the centralized Internet access site has the Internet egress on the LAN side, all Internet access traffic is routed out through the LAN side of the centralized Internet access site. In this mode, you need to configure a default route or dynamic routing protocol on the LAN side so that the default route can be learned from the LAN side.
  - If the centralized Internet access site accesses the Internet through the WAN-side interface, all Internet access traffic is transmitted to the Internet through the WAN side of the centralized Internet access site. (Note: For the site that functions as a centralized gateway, the local Internet access function must also be enabled.)

HUAWEI

Hybrid Internet Access Design: Local Internet Access (Default) + Centralized Internet Access



Application scenario

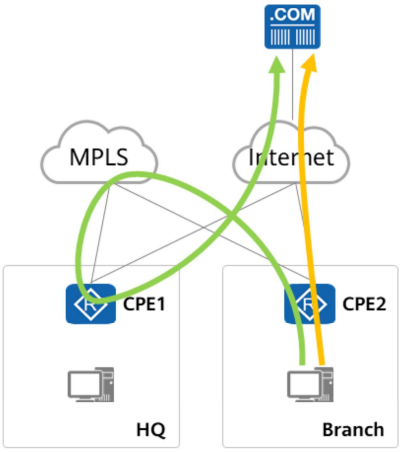
- This mode applies to small enterprises or scenarios where Internet access traffic does not need to be centrally managed.
- The WAN side of the site has links to access the Internet, and the site has high requirements on Internet access reliability.

Description

- Both local Internet access and centralized Internet access are configured on the SD-WAN network.
- Local Internet access is configured for all traffic at a site.
- If a site has two Internet access links, that is, one for local Internet access and one for centralized Internet access, the link for local Internet access has a higher default route preference.
- By default, all Internet access traffic is routed out from the local Internet access interface. When the local Internet access interface is faulty, Internet access traffic is diverted to the centralized gateway and then routed out.

HUAWEI

Hybrid Internet Access Design: Centralized Internet Access (Default) + Local Internet Access for Specified Traffic



Application scenario

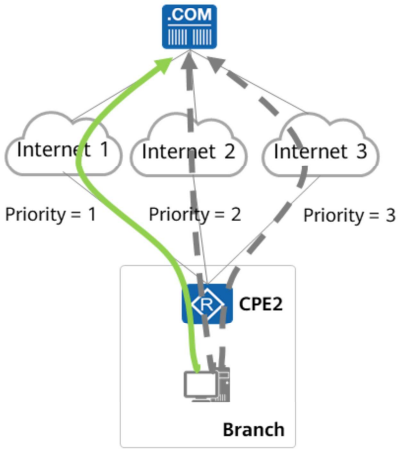
- This mode applies to the scenario where a site has Internet access links, most Internet access traffic needs to be centrally managed, and traffic of some applications can be directly routed out to meet SLA requirements.

Description

- Configure both local Internet access and centralized Internet access on the SD-WAN network.
- Local Internet access is configured at the centralized Internet access site, where all traffic is routed out in local Internet access mode.
- Local Internet access is enabled at other sites and is configured for specified applications.
- By default, Internet access traffic is routed out through the centralized Internet access site, and specific experience-sensitive service traffic is directly transmitted to the Internet through the local WAN-side link.

HUAWEI

Internet Access Link Reliability



- **Standby link solution design:**
  - Links for Internet access are configured on a per-site basis. A maximum of three WAN links can be configured for Site-to-Internet access.
  - A priority is configured for each WAN link. Currently, each link has a unique priority. That is, only one WAN link is used as the active link for accessing the Internet at a time.
  - When a WAN link with the highest priority fails, traffic to the Internet is automatically switched to the link with the second highest priority.
  - The standby link and hybrid Internet access mode can be used together.

HUAWEI

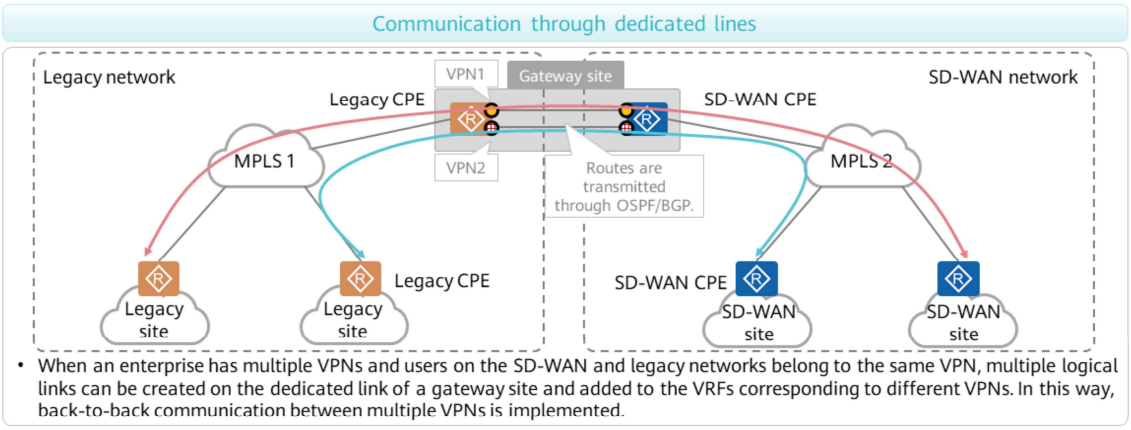
Legacy Site Communication Solution

- Before deploying the SD-WAN Solution, enterprises may have legacy sites that are interconnected through traditional WAN private lines (such as MPLS). After the SD-WAN Solution is deployed, SD-WAN sites and legacy sites may need to communicate with each other. To implement this, a solution for communication with legacy sites must be deployed.
- Communication between SD-WAN sites and legacy sites involves the following scenarios:
  - Communication through enterprise sites: Enterprise sites are directly connected to legacy networks. Two communication solutions are available in this scenario:
    - Communication through dedicated lines
    - Local access
  - Communication through the IWG: Carriers use this mode to provide communication with legacy sites for multiple enterprises. Generally, three communication solutions are available in this scenario:
    - Inter-AS Option B
    - Inter-AS Option A (Layer 3 VXLAN)
    - Inter-AS Option A (Layer 3 VLAN)

HUAWEI

Communication Through Enterprise Sites: Communication Through Dedicated Lines

- When the underlay network connected to an SD-WAN site cannot communicate with a legacy MPLS network, select an SD-WAN CPE and a legacy CPE, which are logically considered as a gateway site.



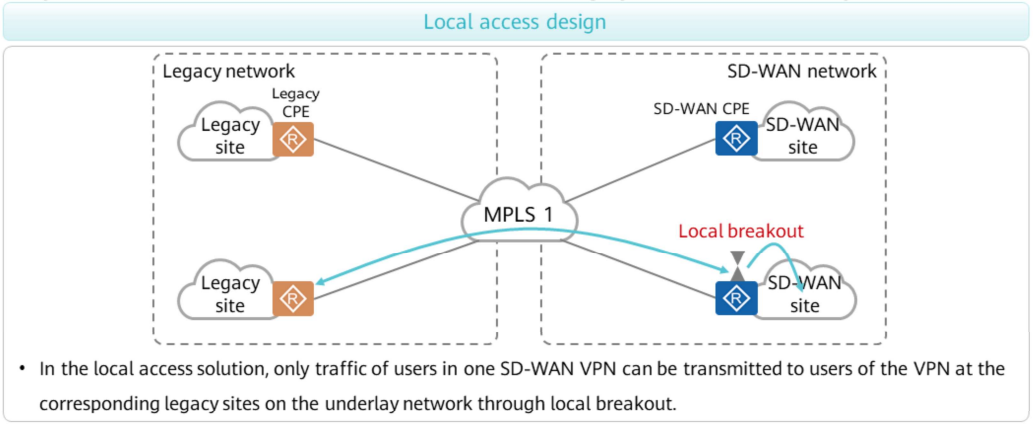
- When an enterprise has multiple VPNs and users on the SD-WAN and legacy networks belong to the same VPN, multiple logical links can be created on the dedicated link of a gateway site and added to the VRFs corresponding to different VPNs. In this way, back-to-back communication between multiple VPNs is implemented.

HUAWEI

- A dedicated link is established between user-side interfaces on both the legacy CPE and SD-WAN CPE. The dedicated link runs a protocol such as BGP or OSPF to exchange routes between the legacy MPLS network and SD-WAN network. In this way, users on the two networks can communicate with each other through the dedicated link.

Communication Through Enterprise Sites: Local Access

- When the underlay network connected to an SD-WAN site is the legacy MPLS network or can communicate with the legacy MPLS network, local breakout technology can be used to transmit user traffic of the SD-WAN site to the underlay MPLS network. Then the traffic is transmitted to the legacy sites over the underlay network.



- In the local access solution, only traffic of users in one SD-WAN VPN can be transmitted to users of the VPN at the corresponding legacy sites on the underlay network through local breakout.

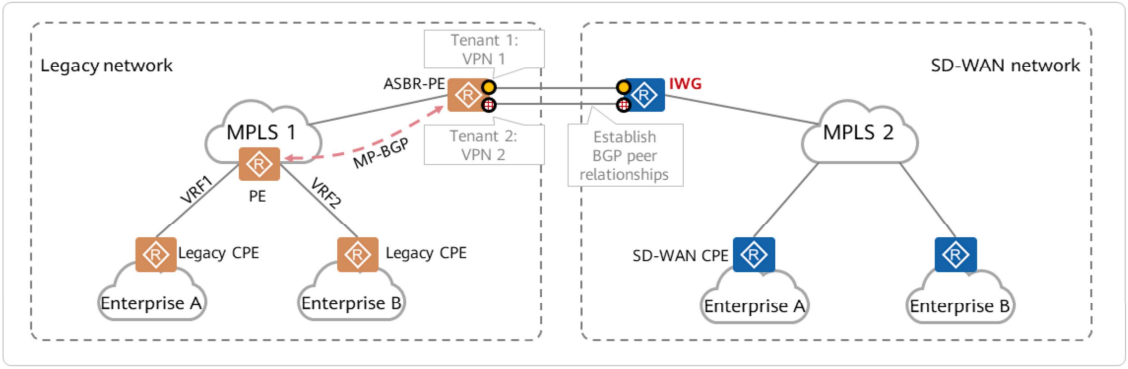
HUAWEI

- Multiple traffic models are supported in this scenario, and you can choose one based on your service requirements.
  - Distributed local access: This model applies if all SD-WAN sites can access legacy sites over the underlay MPLS network through local breakout. In this model, traffic of each site is directly forwarded through the local site, without the need of being forwarded through overlay tunnels.
  - Centralized local access: If some SD-WAN sites cannot access legacy sites through local breakout, you can configure a site that can communicate with the legacy sites as the centralized access site. Traffic from other SD-WAN sites is sent to the centralized access site through overlay tunnels, and then forwarded to the legacy sites through local breakout.
  - Hybrid local access: The SD-WAN Solution enables multi-link sites using the distributed local access model to use local access preferentially, with centralized local access as a backup. This enhances reliability. Traffic from a site that uses the distributed local access model is preferentially transmitted to a legacy site through local breakout. If the MPLS link for local access fails, traffic is automatically switched to the overlay tunnel of another link and transmitted to the centralized access site. The centralized access site then forwards the traffic to legacy sites.

IWG

- The IWG can connect both SD-WAN sites and legacy MPLS VPN sites for multiple enterprise tenants. Each time a tenant is added, only one MP-EBGP peer needs to be configured to interconnect with the peer ASBR-PE.s

Solution design for multiple tenants to communicate with legacy sites

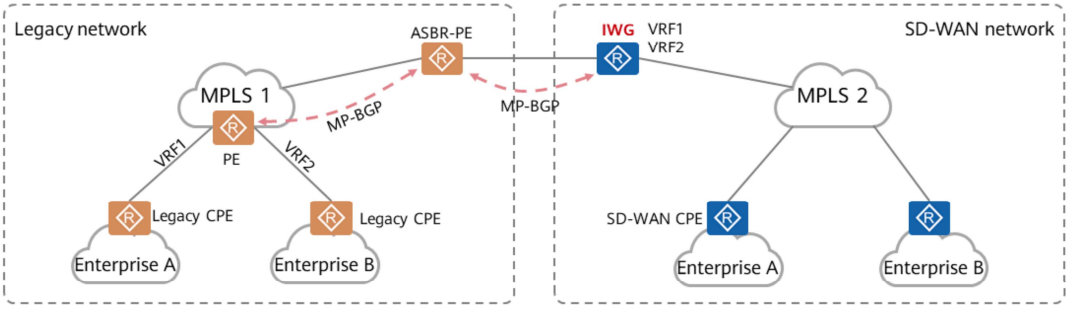


HUAWEI

Communication Through the IWG: Inter-AS Option B

- Similar to the inter-AS Option B solution in MPLS VPN, in the inter-AS Option B solution in SD-WAN, MP-EBGP is used between the IWG and ASBR-PE on legacy networks to exchange labeled VPN-IPv4 routes received from PEs in their respective ASs.

Inter-AS Option B

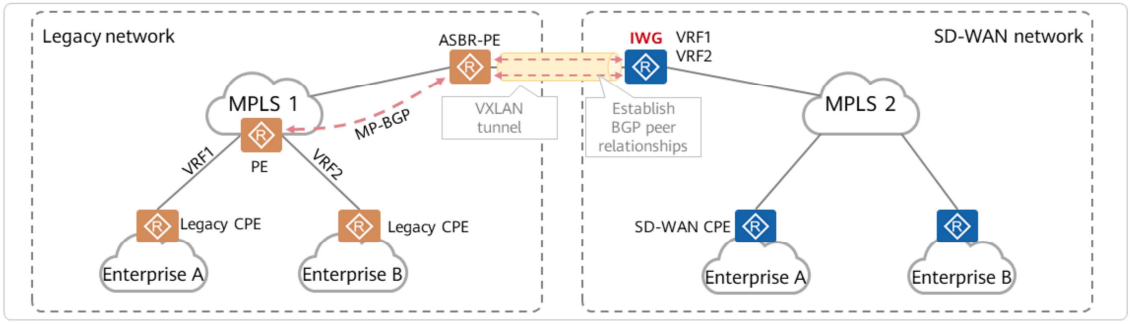


- A pair of public interfaces are configured on the IWG and ASBR-PE, and MPLS is enabled. MP-EBGP is used to exchange labeled VPN-IPv4 routes. BGP is used to transmit inter-AS labels. Therefore, LDP is not required between the IWG and ASBR-PE.

Communication Through the IWG: Inter-AS Option A (Layer 3 VXLAN)

- Similar to the inter-AS Option A solution in MPLS VPN, in the inter-AS Option A solution for SD-WAN, IWGs are interconnected with ASBR-PEs through Layer 3 VXLAN tunnels, and the EBGP or IBGP peer relationship is established at both ends of VXLAN tunnels to exchange routes between SD-WAN sites and legacy sites.

Inter-AS Option A (Layer 3 VXLAN) solution design

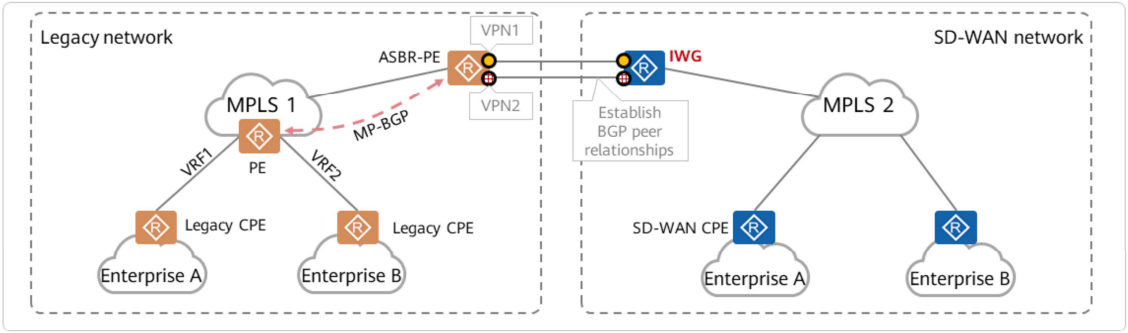


HUAWEI

Communication Through the IWG: Inter-AS Option A (Layer 3 VLAN)

- Between the IWG of the SD-WAN network and the PE of the legacy MPLS network, a pair of specific interfaces (VLANIF interfaces or Ethernet sub-interfaces) are configured for each user VPN and are added to the corresponding VRF. In addition, EBGP is configured based on the VPN instance to exchange routes between the SD-WAN and MPLS domains.

Inter-AS Option A (Layer 3 VLAN) solution design



HUAWEI

Section Summary

- SD-WAN service design includes application service design and network service design.
  - Application service design: includes intelligent traffic steering design, QoS design, and packet loss optimization design.
  - Network service design: includes Internet access design and legacy site access design.

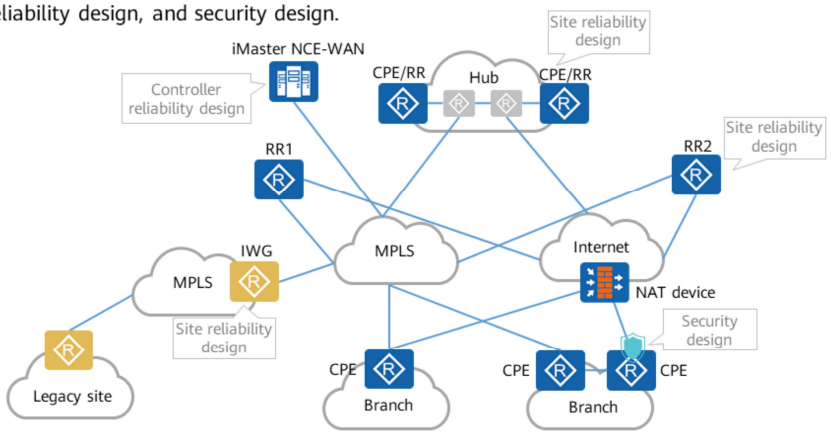
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. **Reliability and Security Design for Huawei SD-WAN Solution**
   - Reliability and Security Design Overview
   - Site Reliability Design
   - Controller Reliability Design
   - Security Design

HUAWEI

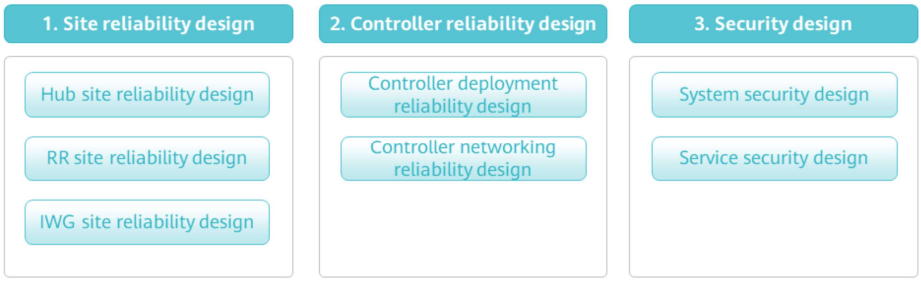
Overview of Reliability and Security Design for Huawei SD-WAN Solution

- Reliability and security design for Huawei SD-WAN Solution mainly includes three parts: site reliability design, controller reliability design, and security design.



HUAWEI

Reliability and Security Design Process for Huawei SD-WAN Solution



HUAWEI

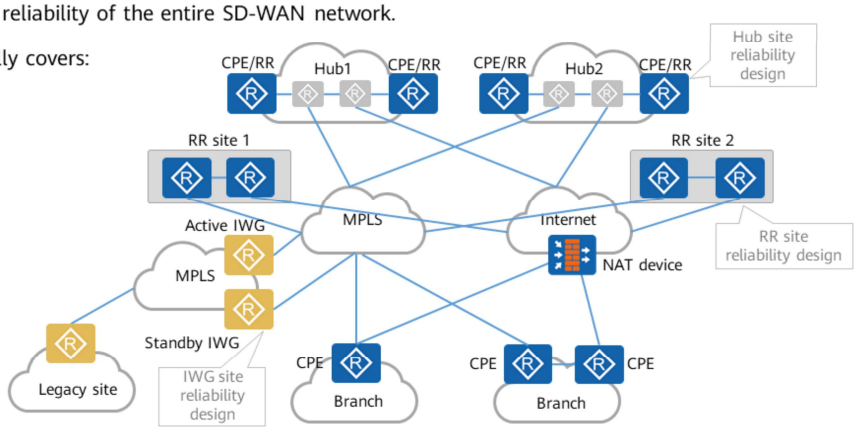
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. Reliability and Security Design for Huawei SD-WAN Solution
   - Reliability and Security Design Overview
   - **Site Reliability Design**
   - Controller Reliability Design
   - Security Design

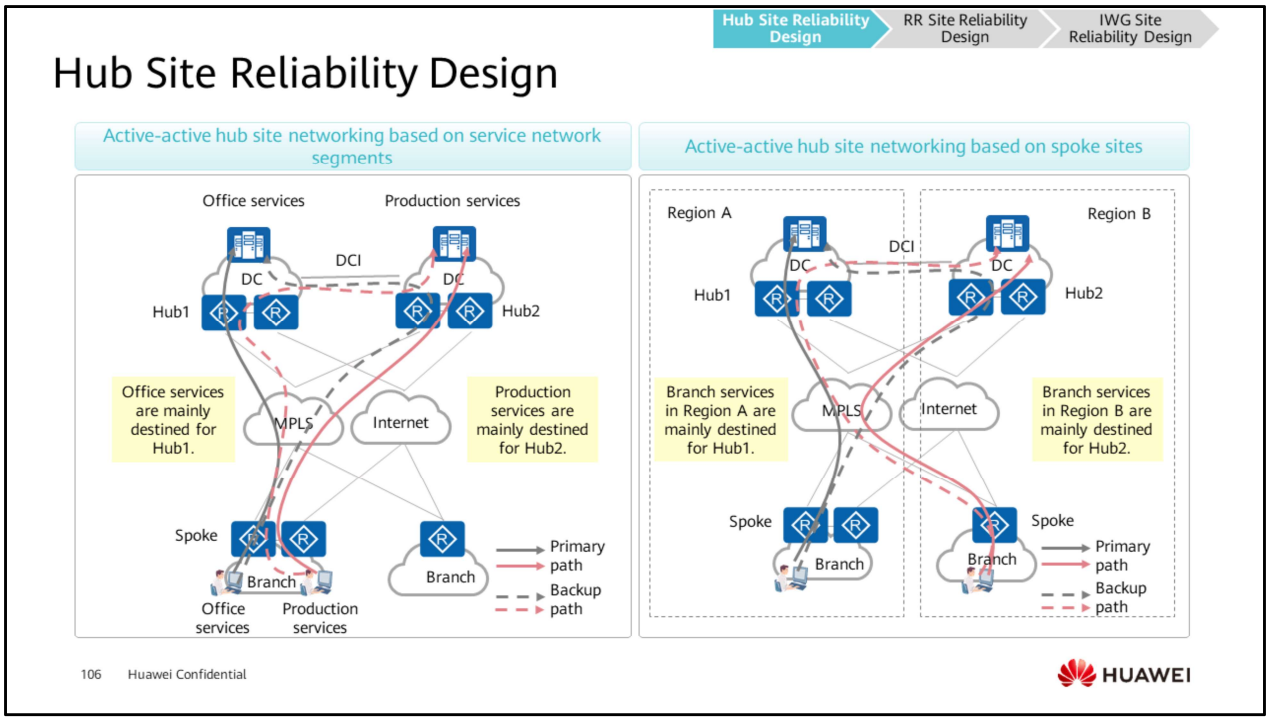
HUAWEI

Site Reliability Design Overview

- SD-WAN sites are typically classified into edge sites, hub sites, and RR sites, as well as gateway sites connected to traditional networks. In Huawei SD-WAN Solution, hub sites, RR sites, and gateway sites are key sites, and their reliability greatly affects the reliability of the entire SD-WAN network.
- Site reliability design generally covers:
  - Hub site reliability design
  - RR site reliability design
  - IWG site reliability design



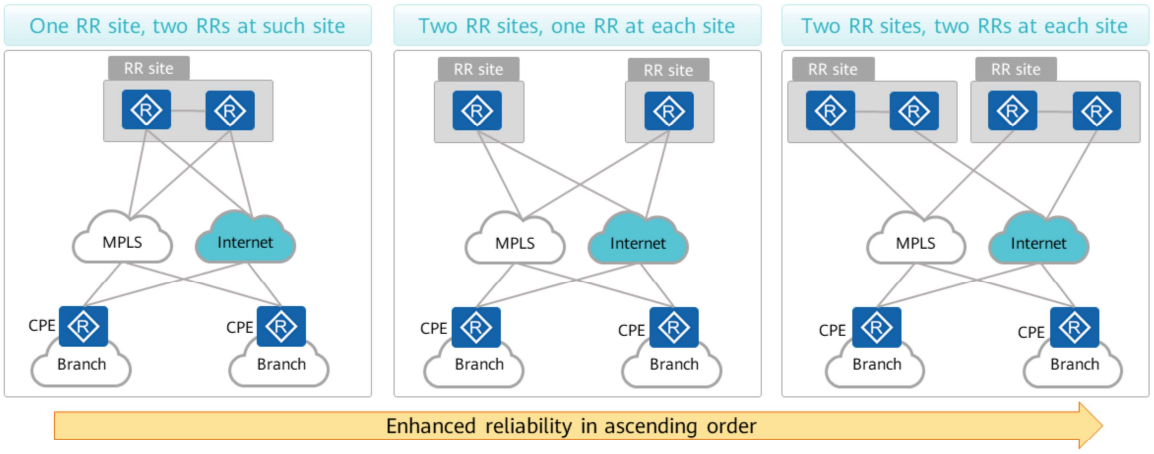
HUAWEI



- Active-active hub site networking based on service network segments:
  - If branch sites' access to DCs can be distinguished based on service network segments, active-active hub sites can be implemented based on the service network segments. Each of the hub sites functions as the active hub site for one service or standby hub site for another different service.
- Active-active hub site networking based on spoke sites:
  - If an enterprise's WAN spans a large physical distance and the enterprise's two DCs are far away from each other, active-active hub sites based on spoke sites are recommended. Each spoke site connects to both hub sites deployed in the DCs and preferentially accesses the nearer hub site for a better DC service access experience. The nearer hub site is the primary one, while the farther hub site is the backup one.

RR Site Reliability Design: RR Redundancy

- One CPE can be connected to two RR sites (four RRs) at most. The following figures from left to right show the RR site deployment modes in three scenarios, where reliability is enhanced in ascending order.

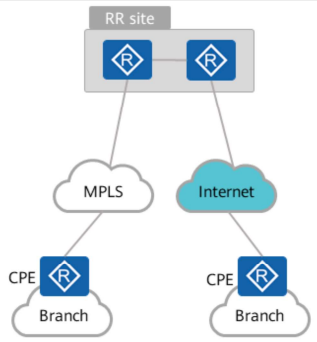


HUAWEI

RR Site Reliability Design: RR Link Redundancy

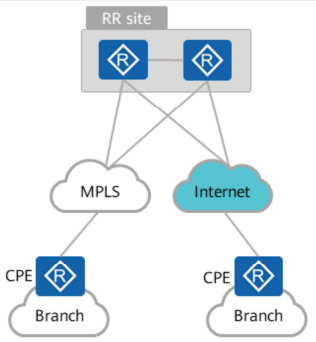
- The edge site needs to be dual-uplinked to the RR site.

Single link



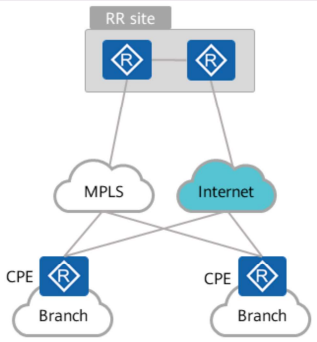
- Low reliability: If one RR is faulty, some branch sites will lose connection with the RR site, causing service interruptions.

Multi-link (topology 1)



- Multiple links are deployed on the RR side. This ensures reliable connections between CPEs and RRs.

Multi-link (topology 2)

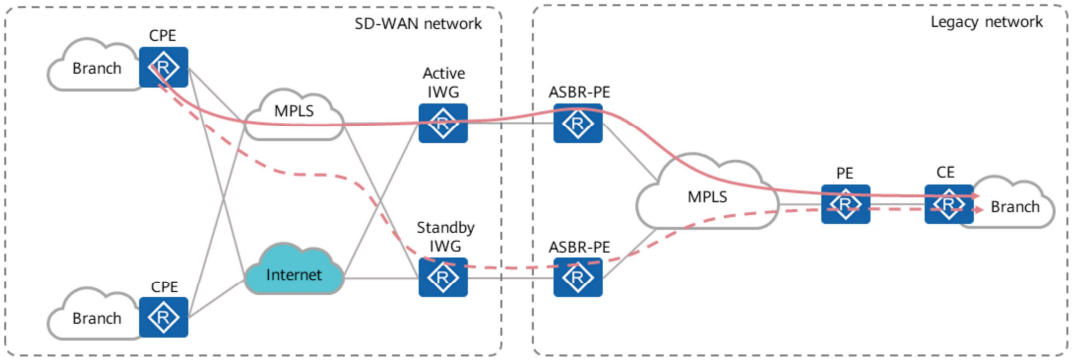


- Multiple links are deployed at the branch site. This ensures reliable connections between CPEs and RRs.

HUAWEI

IWG Reliability Design

- Branch sites are specified with active and standby IWGs to establish IWG tunnels. On the active and standby IWGs, BGP routing policies are automatically orchestrated to control the preference of BGP routes originated from the legacy network.
- The preference of BGP routes on the active IWG is higher than that of BGP routes on the standby IWG.



HUAWEI

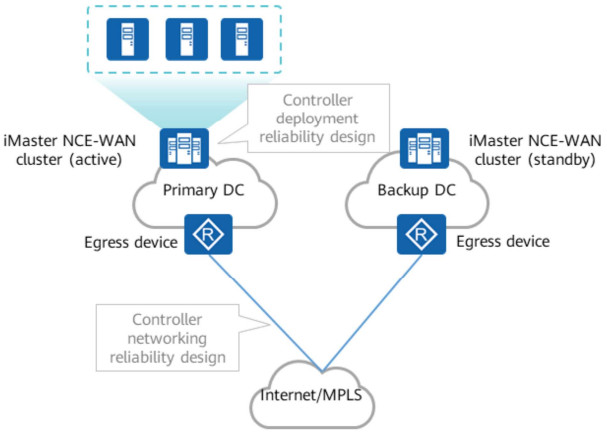
Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. **Reliability and Security Design for Huawei SD-WAN Solution**
   - Reliability and Security Design Overview
   - Site Reliability Design
   - Controller Reliability Design
   - Security Design

HUAWEI

Controller Reliability Design Overview

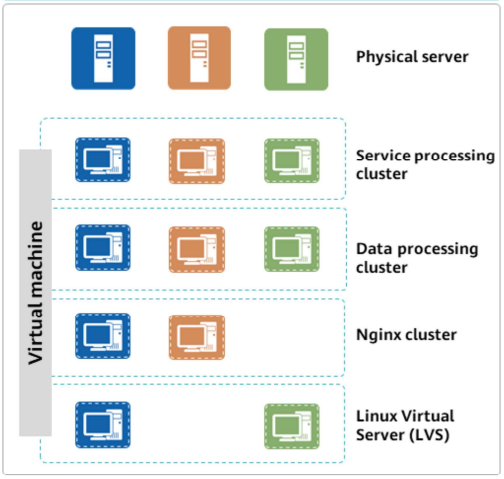
- The controller iMaster NCE-WAN is the core component of Huawei SD-WAN Solution. All services in Huawei SD-WAN Solution are orchestrated through the controller. Therefore, controller reliability directly affects network services.
- Controller reliability design generally covers:
  - Controller deployment reliability design
  - Controller networking reliability design



HUAWEI

Controller Deployment Design in a Site

Controller deployment in distributed cluster architecture



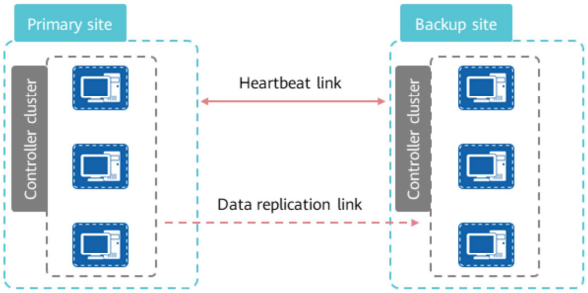
- The controller is deployed in **distributed cluster architecture** to provide high reliability and load balancing capabilities. When a node in the cluster is faulty, other nodes take over services, without affecting services.
- **Northbound load balancing:** External requests are distributed to all nodes in the cluster instead of being centrally processed on just one node. Doing so fully utilizes cluster capabilities and improves reliability.
- **Southbound load balancing:** Controller nodes are dynamically assigned to network devices based on the loads of each controller node.

Cluster	Function Description
Service processing cluster	Provides service processing capabilities, such as CPE management, overlay network configuration delivery, and traffic policy configuration.
Data processing cluster	Provides functions such as CPE performance data storage and data aggregation.
Nginx cluster	Serves as a high-performance HTTP proxy server that forwards concurrent connection requests and implements L4-L7 load balancing for northbound traffic.
LVS	Works as a load balancing component and implements L1-L4 load balancing for southbound and northbound traffic.

HUAWEI

Controller DR Deployment Design

- The geo-redundant disaster recovery (DR) system consists of two sets of iMaster NCE-WAN, which are deployed at two geo-redundant sites.



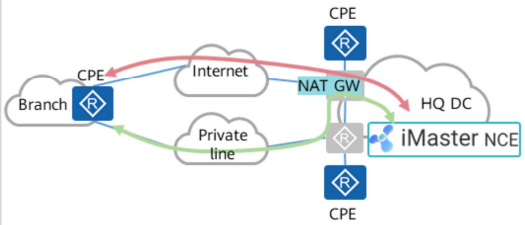
- Geo-redundant DR ensures backup between two controller clusters. The number of nodes in the active cluster must be the same as that in the standby cluster.
- The active and standby controller clusters are both running. However, only the active cluster can provide services, while the standby cluster does not provide services. Data in the active cluster is synchronized to the standby cluster in real time to ensure data consistency.
- The same domain name solution is used for northbound and southbound interfaces of the active and standby controller clusters. Tenants and devices access the active controller cluster through this unified domain name. After active/standby switchover, this domain name is mapped to the new active controller cluster.
- Active/standby switchover is performed manually and can be completed within 10 minutes. In the case of component-level faults in the controller cluster, services are still guaranteed through reliability technologies, including clustering and fault tolerance.
- Geo-redundant DR network requirements: latency ≤ 20 ms; bandwidth ≥ 125 Mbit/s

HUAWEI

Reliability Design for Single-Controller Networking

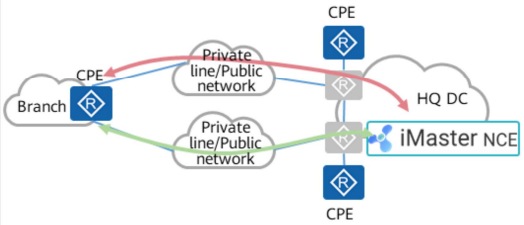
- To ensure the reliability of SD-WAN management channels, dual gateways and dual links generally need to be deployed for highly reliable communication between CPEs/RRs and the controller.

Hybrid networking (public and private network)



- The controller is deployed inside the HQ DC, and uses public IP addresses to provide services for CPEs.
- 1:1 static NAT is deployed on the public network egress device at HQ.
- NAT-related public IP addresses need to be advertised into the private line network.
- Communication between CPEs and the controller through the public network or private line must traverse the NAT device on the public network.

Single-network-type networking



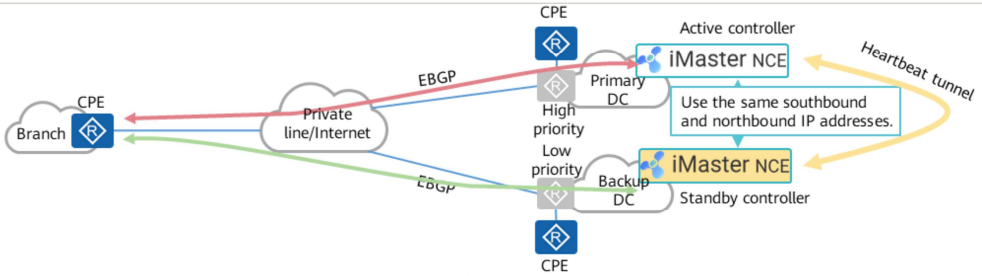
- The controller is deployed inside the HQ DC.
- The private line network needs to send the network segment where the controller resides to VPN networks.
- On the public network, 1:1 static NAT must be deployed on the public network egress device. Static NAT uses the same public IP address.

HUAWEI

Reliability Design for Active and Standby Controller Networking

- In active and standby controller scenarios, communication between CPEs and the active and standby controllers is regulated by controlling route priorities.

Single-network-type networking



- On a pure private line network, the southbound and northbound IP addresses of the active and standby controllers need to be advertised to the private line network through EBGP peers. Additionally, routing policies are used to ensure the route destined for the active controller is preferentially selected.
- On a pure public network, 1:1 static NAT needs to be deployed for southbound and northbound IP addresses, and the NAT configurations on the two egress gateways must be the same. Routes related to NAT addresses are advertised to the public network through EBGP, and routing policies are used to ensure the NAT-related route destined for the primary DC is preferentially selected.

HUAWEI

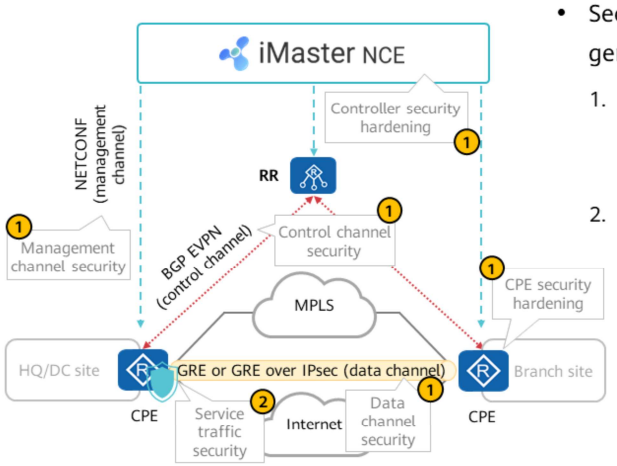
- The active controller is deployed in the primary DC, and the standby controller in the backup DC. A heartbeat tunnel is established between the active and standby controllers to synchronize data and verify the controller status.
- The active and standby controllers use the same southbound and northbound IP addresses. In public network scenarios, the same NAT address must also be configured.

Contents

1. Development Trends and Challenges Facing Enterprise WAN Interconnection
2. Overview of Huawei SD-WAN Solution
3. Networking Design for Huawei SD-WAN Solution
4. Service Design for Huawei SD-WAN Solution
5. **Reliability and Security Design for Huawei SD-WAN Solution**
   - Reliability and Security Design Overview
   - Site Reliability Design
   - Controller Reliability Design
   - Security Design

HUAWEI

Security Design Overview

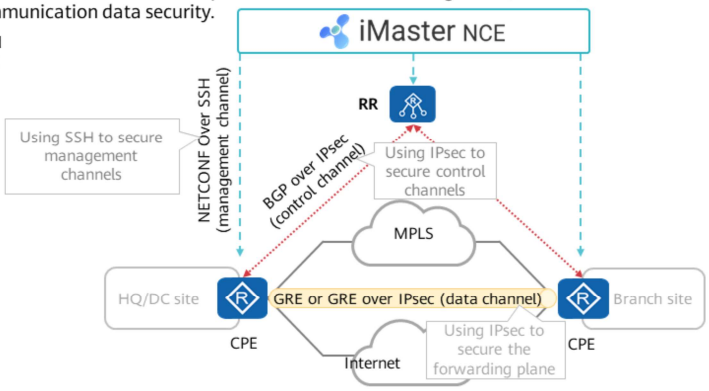


- Security design for Huawei SD-WAN Solution generally covers:
  1. System security: The security requirements of the SD-WAN Solution must be met. In this way, the entire system can run securely and stably.
  2. Service security: The security requirements of services carried by the SD-WAN Solution must be met. Appropriate security protection measures are flexibly selected as required, thereby ensuring secure and reliable running of user services.

HUAWEI

System Security Design: Inter-Component Communication Security

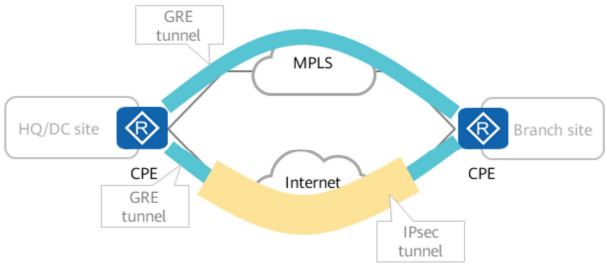
- The communication channels between components of Huawei SD-WAN Solution are classified into three types: management channels, control channels, and data channels. CPEs/RRs communicate with the controller through management channels, CPEs communicate with RRs through control channels, and CPEs communicate with each other through data channels.
- The components of Huawei SD-WAN Solution use secure communication protocols to establish management, control, and data channels, thereby ensuring inter-component communication data security.
  - Security mechanisms for management channels and control channels are enabled by default. You do not need to configure them.
  - Security mechanisms for data channels need to be designed flexibly as required.



HUAWEI

System Security Design: Data Channel Security

- Data channels are overlay tunnels established between CPEs. IPsec is used on such tunnels to ensure data confidentiality and integrity during transmission.
  - IPsec must be deployed when data channels are established over the Internet.
  - IPsec can be deployed as required when data tunnels are established over MPLS or private networks.



System Security Design: Administrator Authentication and Authorization

- The system supports local and remote authentication for administrators. Strict identity authentication ensures that only authorized administrators can log in to the system.
- When a tenant administrator logs in to iMaster NCE-WAN, two-factor authentication (account/password + SMS verification code) is supported. The tenant administrator can be authenticated in either of the following modes:
  - Local authentication: When a tenant administrator logs in to iMaster NCE-WAN, iMaster NCE-WAN authenticates the tenant administrator.
  - LDAP server authentication: iMaster NCE-WAN connects to an LDAP server (such as a general-purpose LDAP server or Windows AD server). When a tenant administrator logs in to iMaster NCE-WAN, the LDAP server authenticates the tenant administrator.
- When a device administrator logs in to a CPE through the CLI, the device administrator can be authenticated in either of the following modes:
  - Local authentication: The tenant administrator can set passwords for device administrators on iMaster NCE-WAN, which then delivers the passwords to CPEs. When a device administrator logs in to a CPE, the CPE authenticates the device administrator.
  - TACACS authentication: A CPE connects to a TACACS server through the underlay network. When a device administrator logs in to the CPE, the TACACS server authenticates the device administrator.

HUAWEI

System Security Design: CPE Security

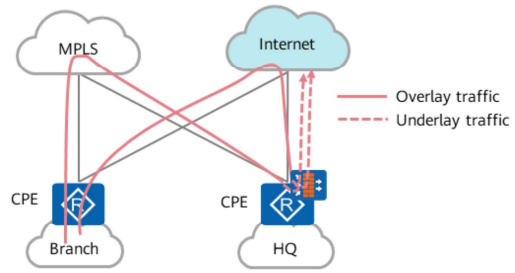
- A CPE must have a secure system architecture and support multiple security protection measures for defending against various security threats.
- A CPE must have multiple security protection capabilities on the control plane, management plane, and forwarding plane, including but not limited to the following:
  - Physical security: A CPE must be able to disable the service ports, serial ports, and services that are not in use, thereby preventing attacks.
  - Data security: A CPE must be able to encrypt sensitive information, such as service data, user names, and passwords, to prevent sensitive information leakage. A CPE must also control data access permissions to prevent unauthorized access to data.
  - Authentication: A CPE must provide system permission control and account permission management functions to implement strict identity authentication and permission control on login behaviors. It also must support security mechanisms such as account/password protection, password complexity check, and anti-brute force cracking of passwords.
  - Attack defense: A CPE must be able to defend against various network attacks, such as IP flood attacks, ICMP flood attacks, malformed packet attacks, and fragment attacks.
  - Security audit: A CPE must have a comprehensive log system to log all configuration operations and abnormal status during system running for future audit.

HUAWEI

Service Security Design: Internet Access Security

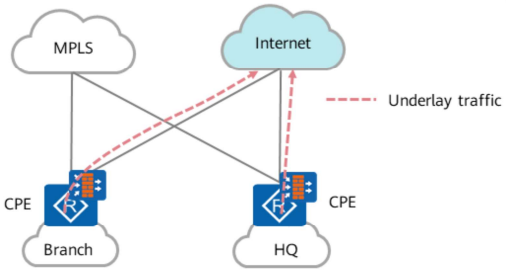
- Huawei SD-WAN Solution uses the built-in security capabilities of CPEs — including ACL filtering, firewall, IPS, and URL filtering — to ensure service security, thereby flexibly meeting service security requirements in different scenarios.

Centralized Internet access



- Internet access traffic from all branch sites is diverted to the HQ site and then goes to the Internet.
- Enable the firewall function on the CPE at the HQ site to isolate internal and external networks.

Local Internet access



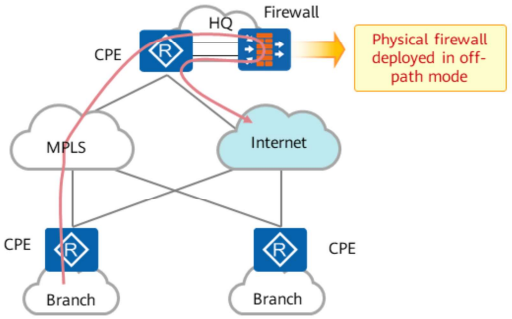
- Internet access traffic from the branch site and HQ site directly goes to the Internet through their respective local CPE.
- Enable the firewall function on the CPEs at the branch site and HQ site to isolate internal and external networks.

HUAWEI

Service Security Design: VAS Advanced Security

- In the centralized Internet access scenario, when a site accesses the Internet through a WAN-side outbound interface, a physical firewall can be deployed in off-path mode to provide VAS-based advanced security protection functions.

VAS advanced security



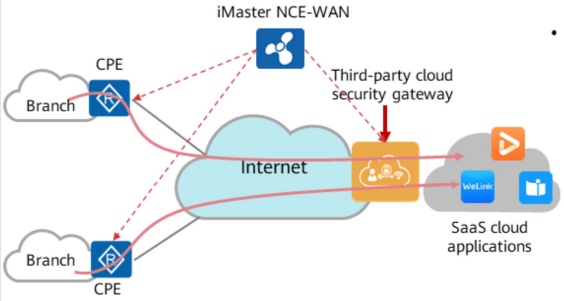
- A physical firewall is deployed in off-path mode. It provides advanced security protection functions for centralized Internet access traffic from the site.
- After VASs are deployed, centralized Internet access traffic is diverted to the physical firewall on the LAN side of the HQ CPE. After being processed by the firewall, centralized Internet access traffic then goes to the Internet through the underlay network.

HUAWEI

Service Security Design: Third-Party Cloud Security

- Huawei SD-WAN Solution can connect to third-party cloud security gateways (Zscaler and Forcepoint). In this way, enterprises' public cloud access traffic and SaaS traffic are sent to such gateways. Third party cloud security gateways perform security check for the traffic and provide access control, threat prevention, and data protection functions, implementing security protection.

Application scenario



- To improve network reliability, the third-party cloud security gateway generally provides two access points to establish tunnels with the CPE, and the CPE uses GRE tunnels to connect to such access points.
  - Single-gateway scenario
    One CPE establishes two GRE tunnels (active and standby) with the two access points of the third-party cloud security gateway.
  - Dual-gateway scenario
    Two CPEs establish four GRE tunnels (active and standby) with the two access points of the third-party cloud security gateway.

Section Summary

- Reliability and security design for Huawei SD-WAN Solution covers site reliability design, controller reliability design, and security design.
  - Site reliability design: hub site reliability design, RR site reliability design, and IWG site reliability design.
  - Controller reliability design: controller deployment reliability design and controller networking reliability design.
  - Security design: system security design and service security design.

HUAWEI

Quiz

1. (Multiple-answer question) Which of the following overlay network topology types are supported by Huawei SD-WAN Solution? ( )
   A. Hub-spoke
   B. Full-mesh
   C. Partial-mesh
   D. Hierarchical networking
2. (True or false) The intelligent traffic steering function can dynamically adjust data forwarding links based on link quality. ( )
   A. True
   B. False

HUAWEI

1. ABCD
2. A

Summary

- Huawei SD-WAN Solution efficiently addresses the challenges facing today's WANs.
- Huawei SD-WAN Solution provides abundant functions, including ZTP, flexible networking, intelligent traffic steering, and service security.
- Design for Huawei SD-WAN Solution mainly covers three parts:
  - Networking design: site design, tunnel design, and VPN design
  - Service design: application service design and network service design
  - Reliability and security design: site reliability design, controller reliability design, and security design

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



Data Center Network Architecture and Key Technology Application Overview



HUAWEI

Foreword

- Data centers (DCs) have been playing an important role since the information age. They have strong or weak connections with enterprises of various scales in all industries including Internet, finance, and education.
- With the development and transformation of IT technologies, such as virtualization, cloud computing, container, and artificial intelligence (AI) technologies, DCs, as the basis for carrying enterprise services, have evolved from traditional DCs to cloud DCs.
- This course describes the data center network (DCN) technology and DC solution.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the definitions of DC and DCN
  - Describe the development phases of the DCN technology
  - Describe key IT technologies of DCs
  - Illustrate key network technologies of DCs

HUAWEI

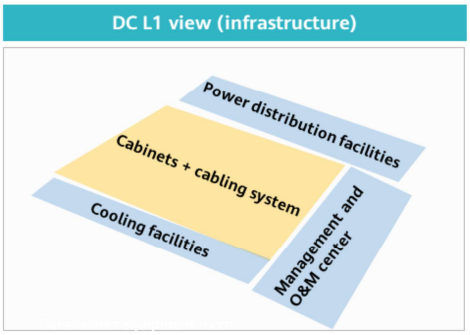
Contents

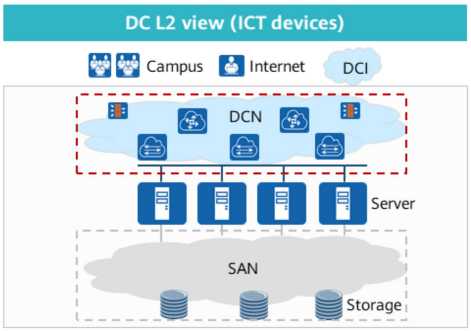
1. **DC and DCN Overview**
2. Terms and Key Technologies of DCs
3. CloudFabric Solution and Related Technologies

HUAWEI

What Is a DC?

- A DC is a complex set of facilities. It consists of the computer system and related devices (such as communications and storage devices), as well as redundant data communication connections, environmental control devices, monitoring devices, and various security devices.
- A DC uses a layered architecture. Generally, L1 refers to the DC infrastructure, and L2 refers to the ICT devices in the DC.



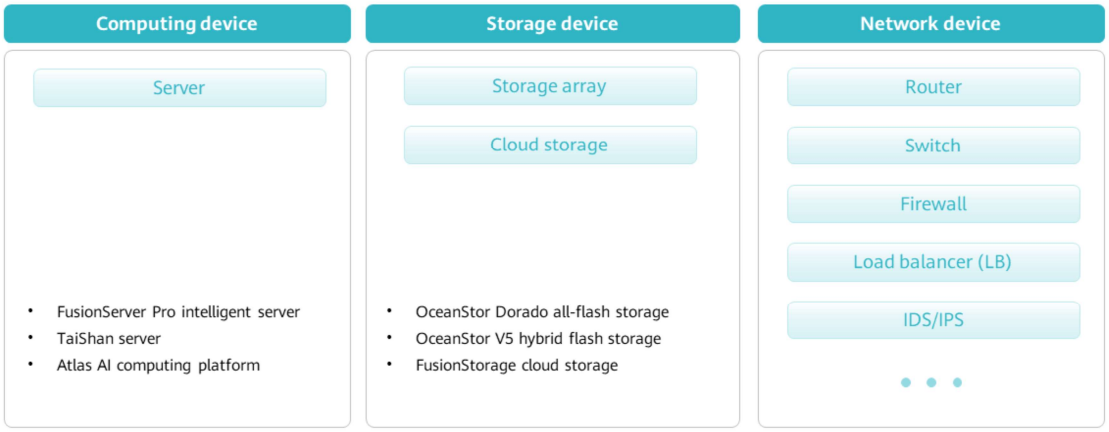


HUAWEI

- Generally, a DC can be divided into the following layers:
  - L0: civil engineering and equipment room building.
  - L1: infrastructure, including equipment room decoration, power supply, cooling, and fire fighting.
  - L2: IT infrastructure, including servers, storage devices, networks, and virtualization software.
  - L3: application platform.
  - L4: various services.
- The concepts, topology, and architecture of DCs in this course are based on L2 and above.

What Does a DC Have?

- DC devices are classified into computing devices, storage devices, and network devices, excluding DC infrastructure.

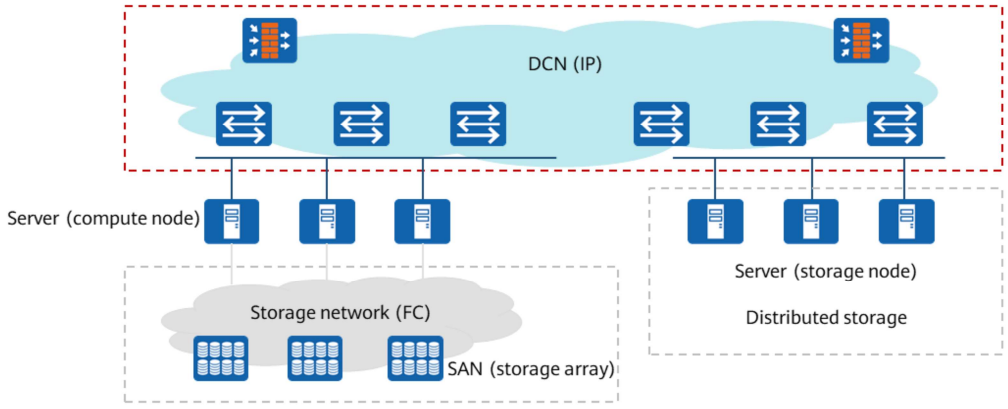


HUAWEI

- This course uses Huawei devices as an example.

Typical DC Networking

- Typical DC networking consists of the DCN, storage area network (SAN), and servers.

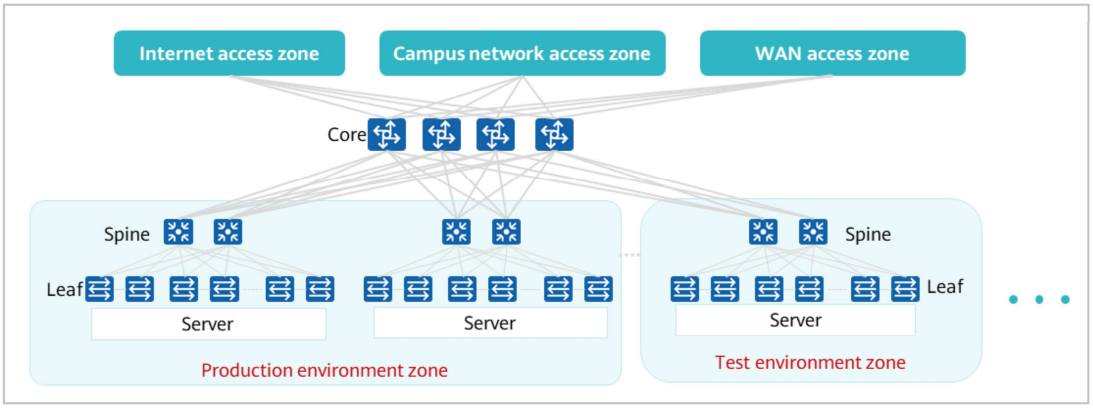


HUAWEI

- A DCN provides interconnection between computing units in a DC and interconnection between computing units in a DC and an external egress.
- A SAN consists of storage arrays and Fibre Channel (FC) switches and is used to provide block storage. A storage network that uses the FC protocol is called an FC SAN, and a storage network that uses the IP protocol is called an IP SAN.
- Distributed storage: The deployment mode of distributed storage is different from that of storage arrays. It stores data on multiple independent servers (storage nodes), which is also used for cloud storage.
- Server (compute node): a server that provides computing services.

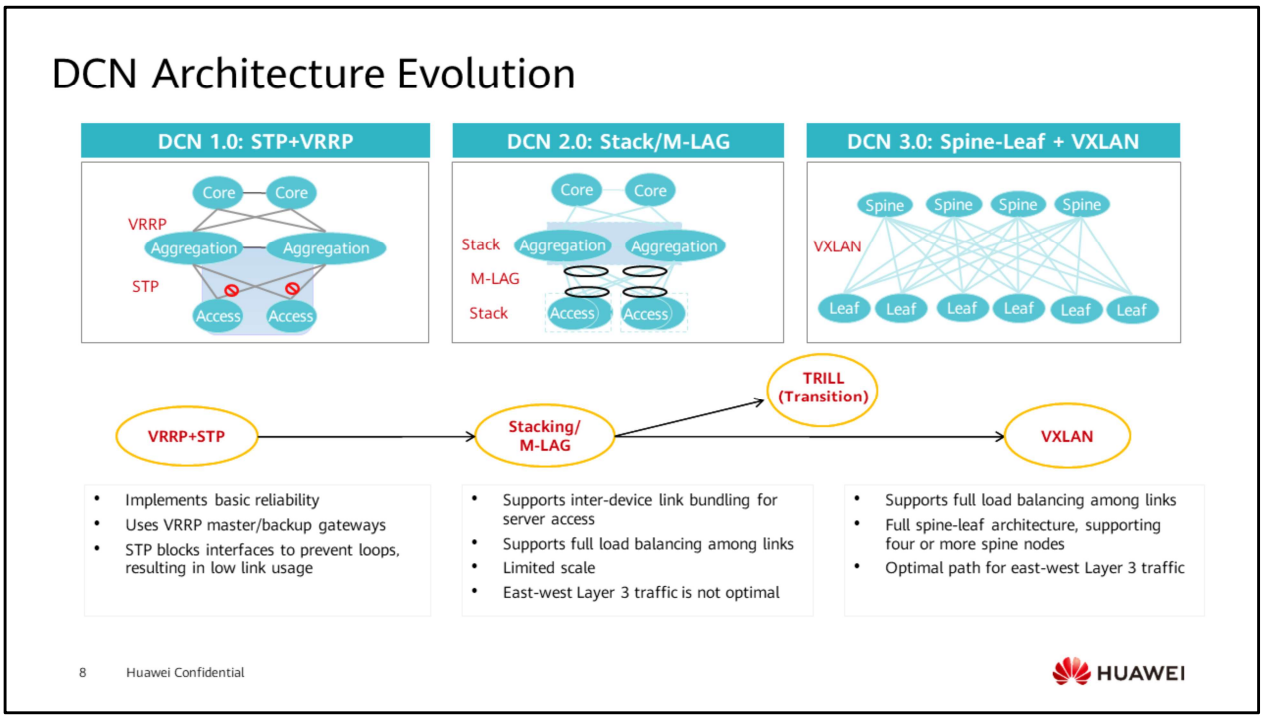
DCN

- A DCN provides interconnection in a DC and interconnection between a DC and an external egress.
- A DCN consists of a series of network devices and is divided into different zones based on service functions.



HUAWEI

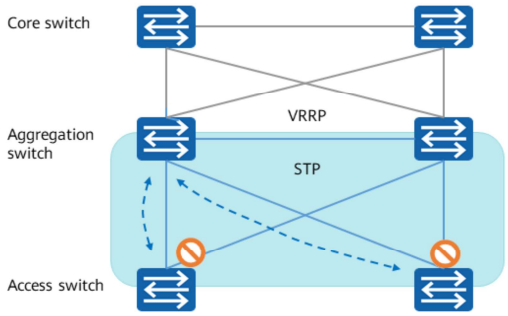
- Zone division for a DCN varies depending on industries and enterprises. For example, a financial DC is divided into production zone 1, production zone 2, test zone 1, test zone 2, big data zone, and operation management zone.
- In this example:
  - Internet access zone: is used for Internet user traffic access.
  - Campus network access area: is used for enterprise campus user traffic access.
  - WAN access area: connects to the WAN built by the enterprise.
  - Production environment zone: connects to the production environment network.
  - Test environment zone: connects to the test environment network.



- Transparent Interconnection of Lots of Links (TRILL) is a large Layer 2 technology. However, TRILL is seldom used because it is difficult to deploy.

DCN 1.0: STP+VRRP

- At the early stage of DCN development, no independent DCN solution is available. In this case, the typical STP+VRRP solution is used. The key points of the networking are as follows:
  - Aggregation switches and access switches form a Layer 2 network, and STP is used to eliminate Layer 2 loops.
  - Gateways are deployed on the aggregation switches, and VRRP is used to provide redundancy.

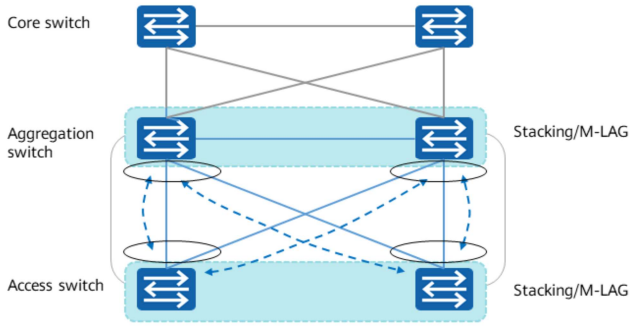


HUAWEI

- This solution solves the basic reliability problems of Layer 2 and Layer 3 switching. However, with the increase of the DC scale and service bandwidth, the following problems become increasingly prominent:
  - The network throughput is low and cannot meet the increasing traffic requirements:
    - The STP blocking mechanism causes low network link usage.
    - The VRRP single-active gateway mechanism fails to fully unleash the device forwarding performance.
    - Layer 2 traffic is forwarded along a non-shortest path. The root bridge easily becomes a bandwidth bottleneck and causes a long forwarding delay.
  - The STP network scale is limited, and the convergence performance is poor.
  - There are a large number of managed nodes, and the logical topology and maintenance are complex.

DCN 2.0: Stacking/M-LAG

- DCN 2.0 uses the stacking and M-LAG technologies to solve the STP+VRRP network problems in the DCN 1.0 phase. Both stacking and M-LAG are network virtualization technologies that implement many-to-one virtualization of devices and support inter-device link bundling. The key points of the networking are as follows:
  - Aggregation layer: Devices form an M-LAG or are stacked, which function as the Layer 3 gateways. The aggregation layer and access layer are interconnected through inter-device link bundling.
  - Access layer: Devices form an M-LAG or are stacked to meet the requirement that server NICs are dual-homed to two physical devices in NIC binding mode.

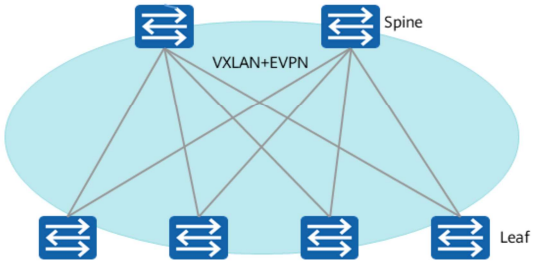


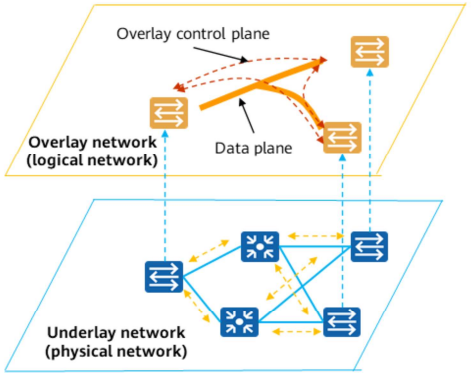
HUAWEI

- Regardless of whether the stacking or M-LAG technology is used, DCN 2.0 has the following characteristics:
  - Traffic forwarded at Layer 2 and Layer 3 is load balanced, making full use of all links.
  - Inter-device link aggregation is deployed to support protection against physical node faults.
  - Compared with STP, the Layer 2 forwarding path is shorter and the delay is shorter.
  - A larger Layer 2 network can be constructed.

DCN 3.0: Spine-Leaf + VXLAN

- DCN 3.0 uses the spine-leaf architecture and combines VXLAN and Ethernet VPN (EVPN) technologies.
- VXLAN uses MAC-in-UDP packets to implement data plane encapsulation and construct IP-based overlay virtual local area networks (VLANs). Spine and leaf nodes are fully meshed.





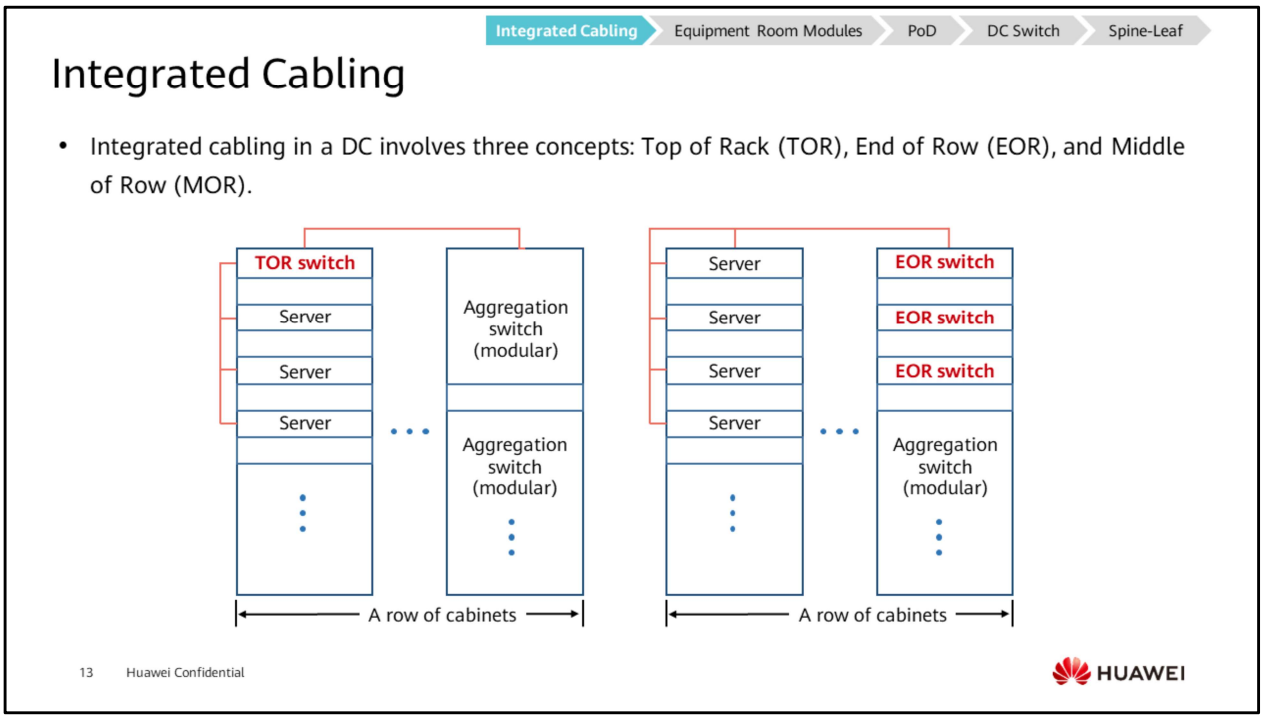
HUAWEI

- Compared with DCN 2.0, DCN 3.0 greatly improves the scalability and flexibility of service networks. DCN 3.0 has the following advantages:
  - 10GE NICs are mainly used on servers, and 10GE access and 40GE interconnection are used on the network.
  - Network-wide resource sharing and access resource pooling in a data center are supported.
  - VXLAN and EVPN technologies are used to improve link utilization.
  - The spine-leaf architecture has strong horizontal scalability.
  - Layer 3 ECMP networking is formed through full mesh connections to implement a non-blocking, highly reliable, high-performance, and low-latency network.
  - The M-LAG technology can be used to support load balancing of Layer 3 gateways.
  - The EVPN distributed gateway function is used to achieve optimal Layer 3 traffic paths.
  - Outer packets are encapsulated using the common IP protocol, which facilitates network extension and reuse of assets and tools.
- DCN 3.0 is recommended for DCN planning and design.

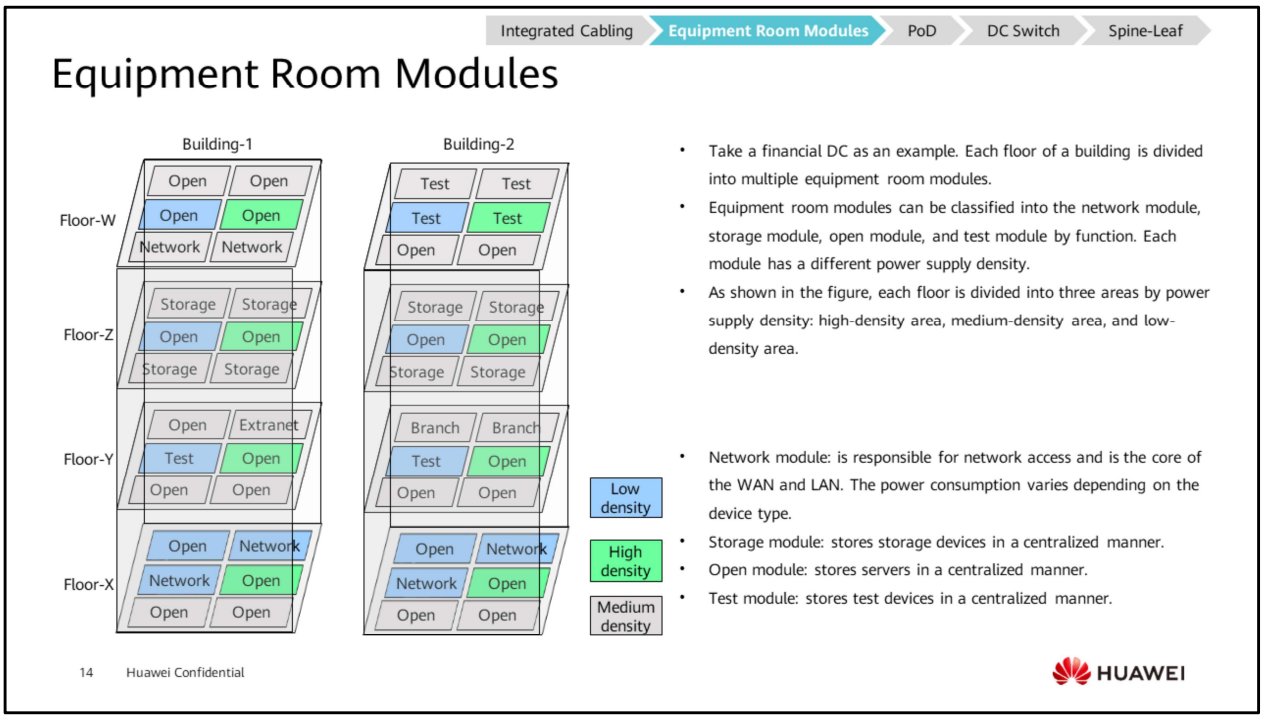
Contents

1. DC and DCN Overview
2. **Terms and Key Technologies of DCs**
   - Common Terms of DCs
   - Key Services of DCs
3. CloudFabric Solution and Related Technologies

HUAWEI



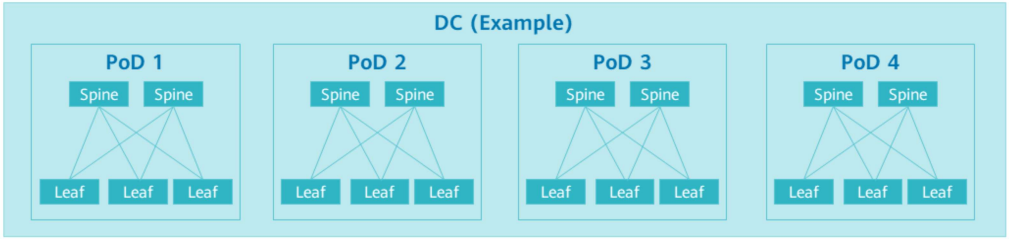
- A common standard cabinet in a DC is 42 U, and the height of each U is 4.445 cm.
- TOR switches are deployed on the top of a cabinet. Servers in the cabinet are connected to the TOR switches through optical fibers or network cables. The TOR switches are connected to the upper-layer aggregation switches. This distributed access mode is applicable to the scenario where there are a large number of access devices or the density of devices in a single cabinet is high. This mode reduces the number of cables between server cabinets and network cabinets, simplifying connection management. However, access switches are distributed in multiple cabinets, which is inconvenient for centralized maintenance and management of switches. This mode is common in cloud DCs.
- EOR switches are deployed in one or two cabinets at the end of each row of cabinets. All servers in the row are connected to the EOR switches through horizontal cabling. This mode is common in traditional DCs. In this mode, a large number of cables are aggregated from multiple server cabinets to the network cabinet, which is difficult to manage. However, switches can be managed and maintained in a centralized way.
- MOR switches are deployed in one or two cabinets in the middle of each row of cabinets. In this mode, the cabling from the server cabinet to the network cabinet is simpler than that in the EOR mode. In addition, switches are managed in a centralized manner, which is a compromise between the TOR and EOR modes.



- Network module: is responsible for network access and is the core of the WAN and LAN. It connects to carriers' networks and connects to branches across the country. Internally, it functions as the LAN core of the DC. The power consumption varies depending on the device type. For example, large core network devices consume high power, and LBs, firewalls, switches, and routes consumer low power.
- Storage module: stores storage devices, including NAS and SAN storage devices, tape libraries, and virtual tape libraries.
- Open server module: stores standard servers in a centralized manner. The servers include PC servers, blade servers, and midrange servers. Servers feature high standardization and high density.
- Test module: stores test devices in a centralized manner. The test module is flexible and has lower security requirements than the production module. It can be adjusted based on the test requirements at any time. The management requirements are not high.
- In the DC of a financial institution, several special types of equipment room modules can be planned for some special purposes.
- Internet module: is mainly responsible for connecting to Internet applications such as e-banking, website, and e-commerce applications. It usually needs to be connected to the Internet, and is vulnerable to an attack. Therefore, the security requirements of this module are high.
- Extranet module: connects to third-party companies, third-party organizations, and partner companies that have service relationships. There are a large number of interconnection private lines.
- Host module: is used to deploy a host system. The device type and deployment mode of a host system are different from those of common IT devices. Therefore, a dedicated host module can be planned.
- Visit modules and hosting modules can also be planned.

- Power supply density of equipment room modules:
- Based on the power supply density, equipment room modules are classified into low-density modules (0.4–0.9 kW/m^2), medium-density modules (0.9–2.5 kW/m^2), and high-density modules (> 2.5 kW/m^2).
    - Low-density modules can be used to deploy network devices. They can be used to deploy a single cabinet for large-scale network devices (6 kW) and meet the requirements for low average power consumption of network modules.
    - Medium-density modules can well meet the deployment requirements of conventional PC servers and midrange computers. A 42 U cabinet can be used to deploy 12 PC servers of 2 U (5–7 kW) or 8 PC servers of 4 U (6–7 kW) to make full use of power resources.
    - High-density modules are suitable for blade servers. Three blade servers can be deployed in a 42 U cabinet. The power consumption is up to 14.3 kW, fully utilizing power resources. If PC servers are deployed in a high-density module, the maximum power utilization can reach only about 50%.
- Suggestion: The area of each equipment room module is about 400 to 500 square meters. The designed area of a medium-density module is 1–2 kW/m^2. The number of such modules accounts for more than 70%. The number of low-density and high-density modules does not exceed 30%.

PoD



- To facilitate resource pooling and management in a DC, a DC can be divided into one or more physical partitions. Each physical partition is called a Point of Delivery (PoD). It is a modular entity that integrates network, storage, and computing.
- PoDs can be defined based on actual service requirements:
  - For a large DC, the entire equipment room can be defined as a PoD.
  - For a medium DC, two or more rows of cabinets can be defined as a PoD.
  - For a small DC, one or more cabinets can be defined as a PoD.

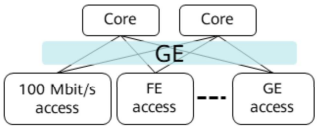
HUAWEI

- The PoD definition varies depending on the enterprise. For example, in the DC of a large enterprise, a PoD may be considered smaller than an equipment room. In another enterprise, a PoD consists of 48 TOR switches and four spine switches.

DC Switch

- DC switches generally refer to hardware switches. Its development goes through the following three phases: FE/GE access and GE uplink; GE/10GE access and 10GE/40GE uplink; 10GE/40GE access and 100GE uplink.

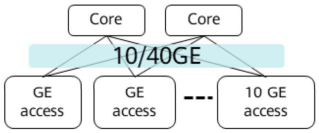
Phase 1: High-density FE/GE access and GE uplink



About 2000 to 2006

DC servers mainly use FE/GE access. Most DC switches mainly provide 100 Mbit/s or 1000 Mbit/s interfaces, which can meet the requirements of a DCN consisting of a small number of servers.

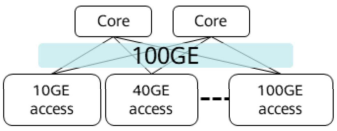
Phase 2: High-density GE/10GE access and 10GE/40GE uplink



About 2006 to 2012

DC servers mainly use GE/10GE access. DC switches provide mainly GE interfaces and a few 10GE interfaces to meet the requirements of a small or medium DCN that consists of GE servers.

Phase 3: High-density 10GE/40GE access and 100GE uplink



About 2012 to 2020

DC servers mainly use 10GE/40GE access. Therefore, core switches in the DC must provide high-density 10GE/40GE interfaces to meet server access requirements.

HUAWEI

DC Switches for the AI Era

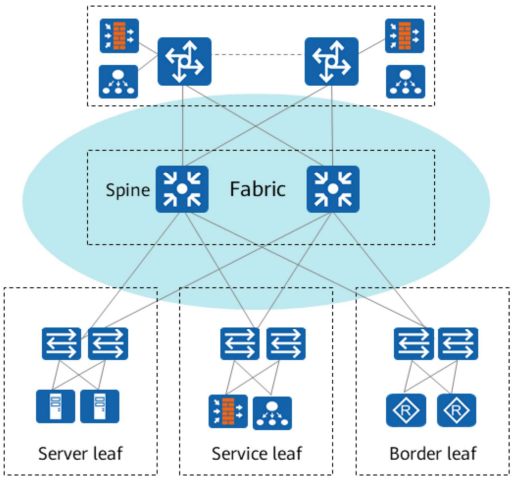


CloudEngine 16804 CloudEngine 16816 CloudEngine 16808

- Huawei CloudEngine 16800 series switches (CE16800 for short) are DC switches built for the AI era.
  - Ultra-large capacity: The CE16800 provides 734 Tbit/s switching capacity, scalable to 3096 Tbit/s. It supports 400GE smooth evolution to meet the demand for surging digital traffic in the future.
  - Intelligent engine: The iLossless algorithm is used to learn and train network-wide traffic in real time, implementing network adaptation and self-optimization. Zero packet loss, E2E µs-level latency, and high throughput are achieved.
  - Intelligent O&M: The CE16800 provides telemetry-based data collection in milliseconds, building a data basis for the intelligent O&M platform. The edge AI processing capability is supported to implement local fault inference.

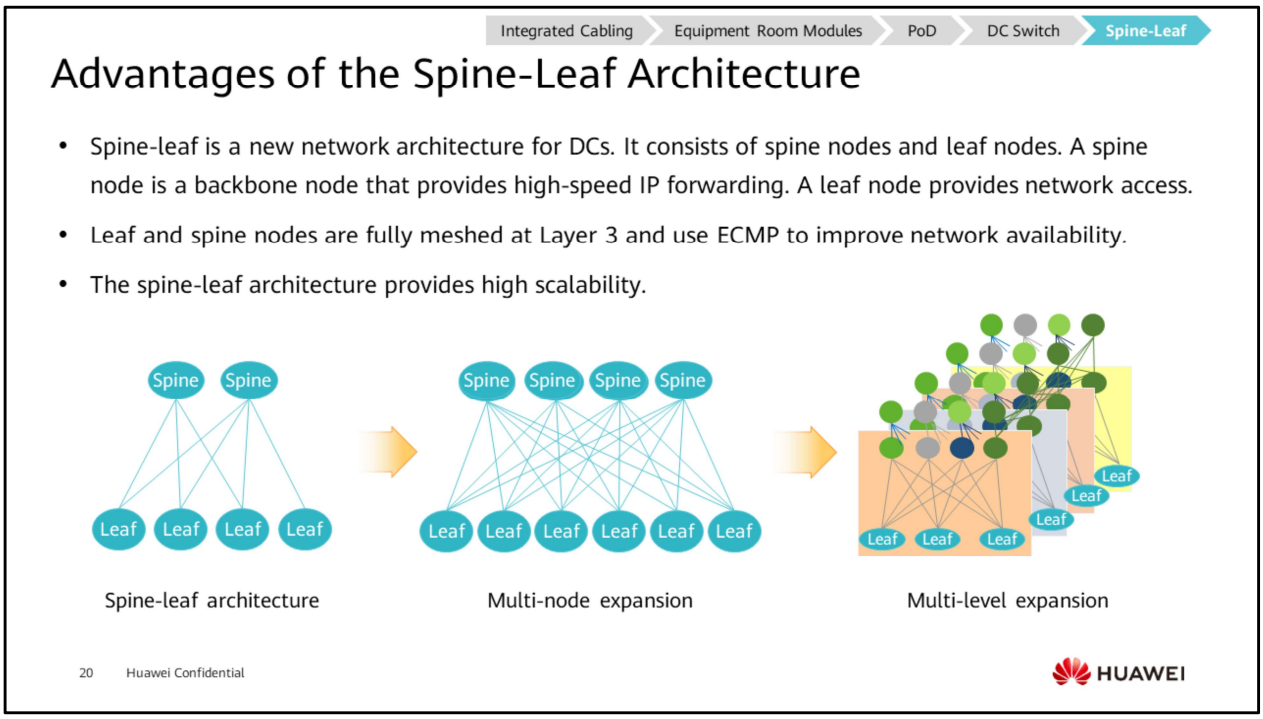
HUAWEI

Basic Concepts of the Spine-Leaf Architecture



Concept	Description
Spine	Core node on a VXLAN fabric, which provides high-speed IP forwarding and connects to leaf nodes using high-speed interfaces.
Leaf	Access node on a VXLAN fabric, which connects various network devices to the VXLAN network.
Fabric	Basic physical network topology of a DC, comprised of a group of spine and leaf nodes.
Service leaf	Functional node that connects to L4-L7 value-added service (VAS) devices, such as firewall and LBs, to the VXLAN fabric.
Server leaf	Functional node that connects virtual and physical servers to the VXLAN fabric.
Border leaf	Functional node that connects routers or transmission devices outside a DC to the VXLAN fabric of the DC.

HUAWEI

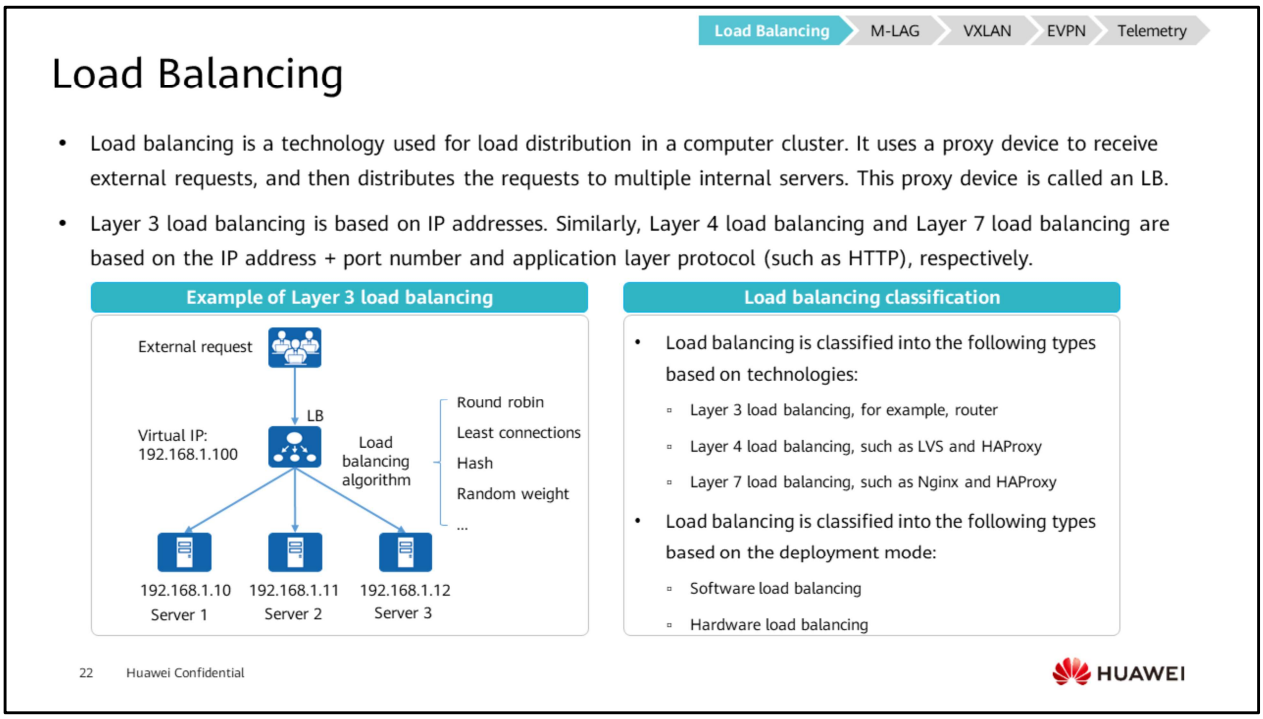


- Characteristics of the spine-leaf architecture:
  - Each leaf node (low level) is fully meshed to all spine nodes (high level).
  - Nodes at the same level are not connected to each other.
  - In a typical application scenario, the spine-leaf architecture functions as a logical modular switch. A leaf node functions as an LPU of the modular switch and is responsible for external traffic access. A spine node functions as the SFU of the modular switch and is responsible for traffic forwarding between leaf nodes.
- The spine-leaf architecture provides high scalability:
  - The number of spine nodes can be expanded to 4 or more. The maximum number of spine nodes depends on the number of uplink interfaces on leaf nodes.
  - A two-level spine-leaf architecture can be expanded to a three-level spine-leaf architecture to implement high-speed data exchange between more leaf nodes.

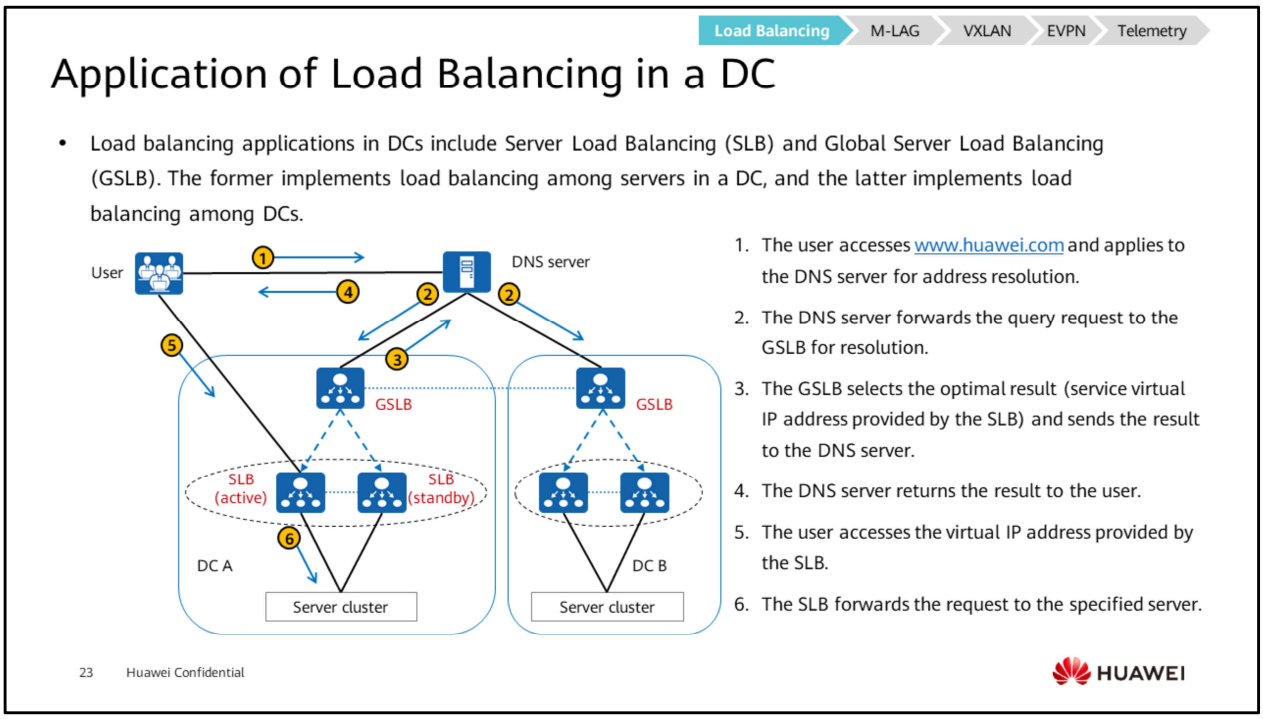
Contents

1. DC and DCN Overview
2. **Terms and Key Technologies of DCs**
   - Common Terms of DCs
   - Key Technologies of DCs
3. CloudFabric Solution and Related Technologies

HUAWEI



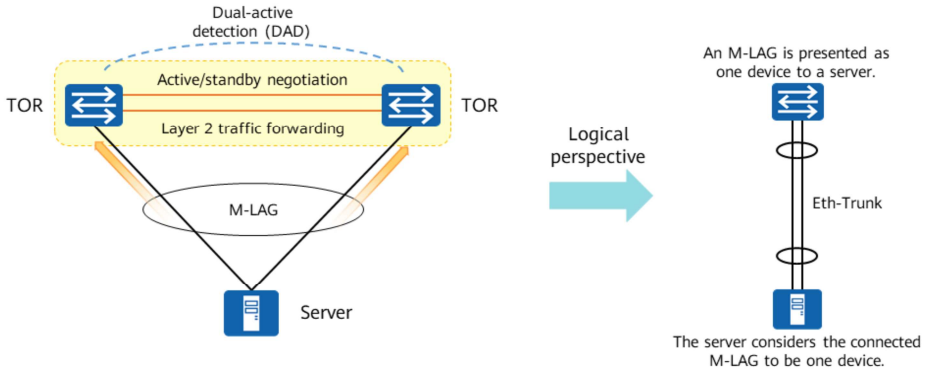
- The access IP address provided by the LB is called the virtual IP address in this example and is called the floating IP address (FIP) in OpenStack.
- The load balancing algorithm determines the health servers to be selected. The common algorithms are as follows:
  - Round robin: The LB distributes each request sequentially around the server list.
  - Least connections: The LB preferentially selects the server with the least connections.
  - Hash: The LB calculates a hash value based on the source IP address of a request and determines the server to which the request is distributed based on the hash value.
  - Random: The LB randomly distributes traffic based the weight.



- GSLB is used when an enterprise has multiple DCs in different locations. Users can access the nearest DC based on their locations. There are multiple GSLB solutions. This example describes the most common DNS-based GSLB solution used in DCs.
- In the GSLB solution, the domain name service provider adds a GSLB with the intelligent DNS resolution function to the Name Server (NS) record of a domain name. Then the GSLB resolves the domain name. If GSLBs are deployed in multiple locations, all of them should be added to the NS record to ensure high availability. The GSLB performs health check on backend servers and public IP addresses of other DCs. Health check results are synchronized between GSLBs in different DCs using proprietary protocols. The GSLB selects the optimal address based on the global load balancing policy and sends the address to the DNS server. The DNS server then sends the optimal address to the user.
- The SLB in the DC distributes requests to application servers based on the information contained in the requests. In addition, the SLB ensures that the system selects the optimal server for processing based on the predefined policy. This improves the availability and scalability of applications to some extent.

M-LAG

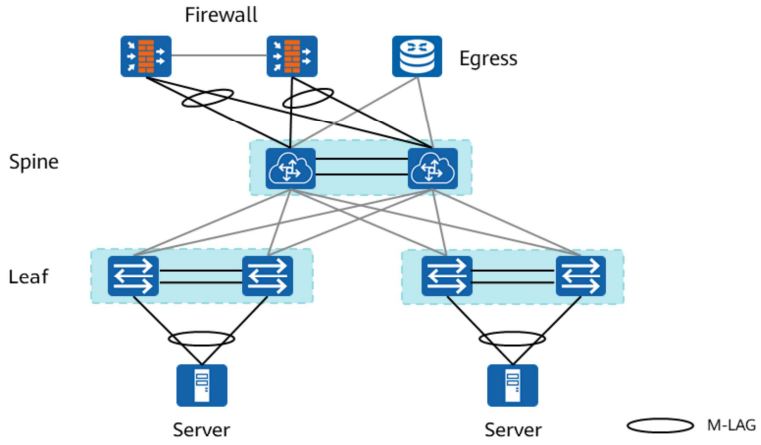
- M-LAG is a technology that implements inter-device link aggregation. It provides device-level link redundancy. M-LAG implements load balancing as well as backup protection.
- In a DC, two TOR switches negotiate to form an M-LAG in active/standby mode to connect to other devices (servers or firewalls).



HUAWEI

Application of M-LAG in a DC

- Currently, a DCN uses the spine-leaf architecture. To meet high reliability requirements, M-LAG is recommended for server or firewall access.

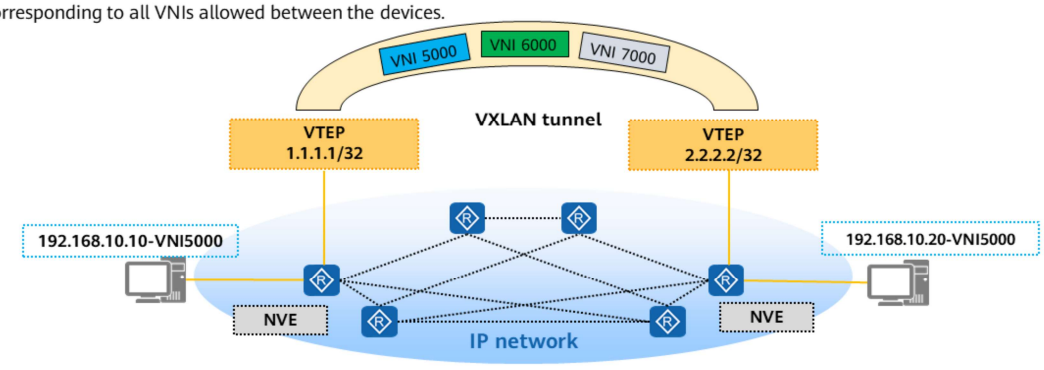


HUAWEI

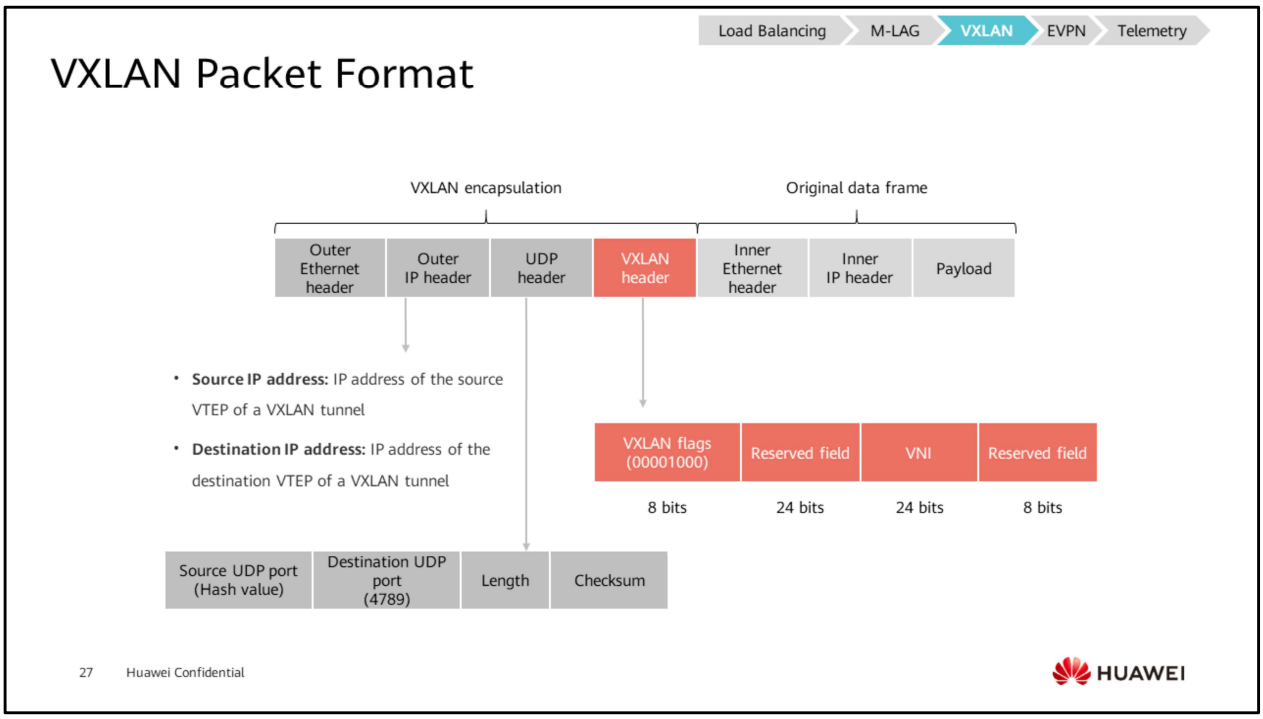
- In this example, the spine node, border leaf node, and service leaf node are combined.

VXLAN

- VXLAN is a point-to-point Layer 2 tunneling technology that can build a Layer 2 virtual network over a physical network with reachable routes. The underlay routing network on which the VXLAN depends is not limited by the network structure and provides scalability.
- A VXLAN packet contains a VXLAN Network Identifier (VNI) field, which is similar to a VLAN ID and is used to differentiate VXLANs. Generally, there is only one VXLAN tunnel between two devices. Similar to a trunk link, a VXLAN tunnel is used to carry traffic corresponding to all VNIs allowed between the devices.



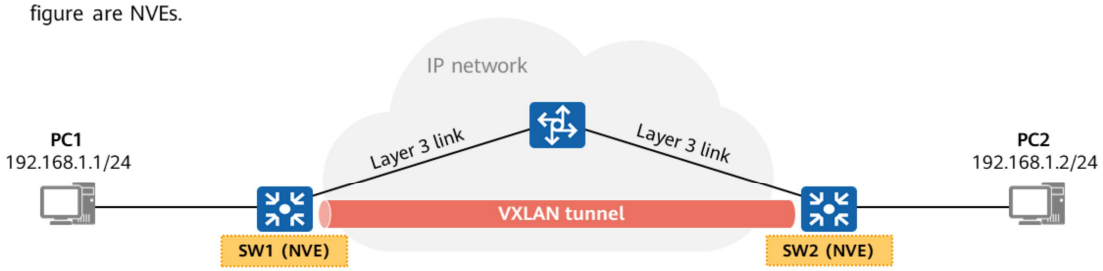
HUAWEI

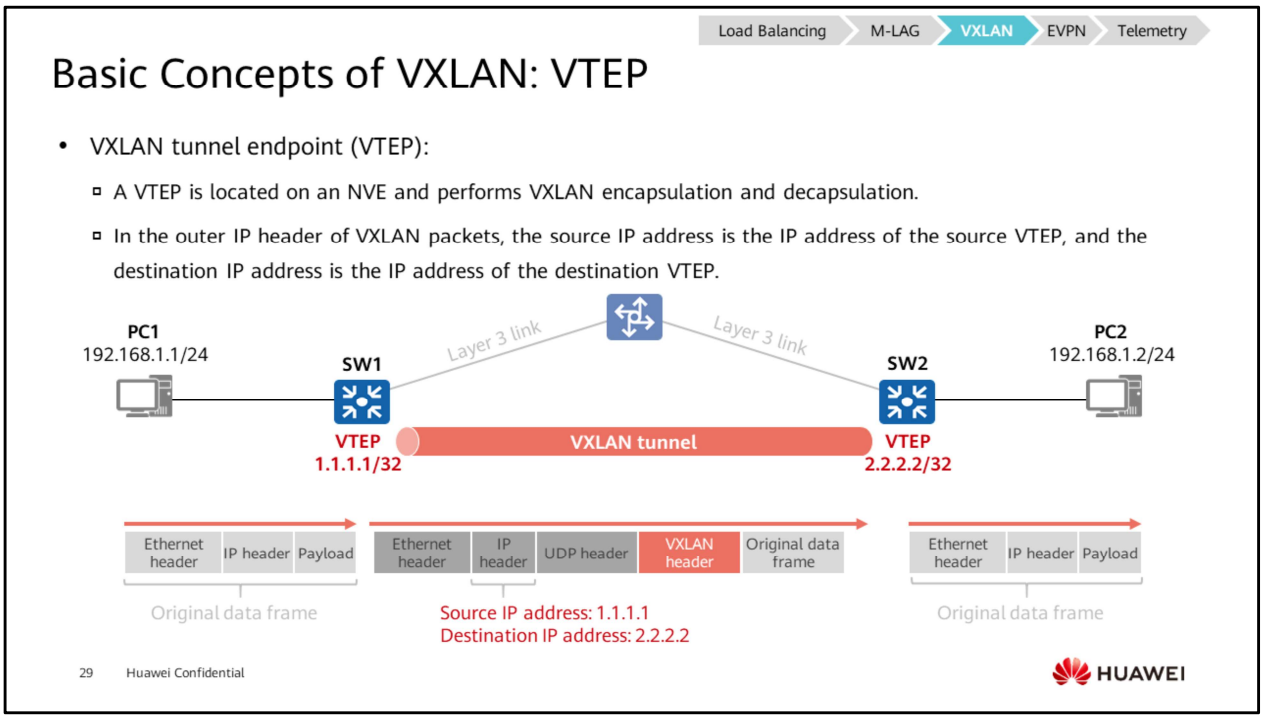


- The preceding figure shows the standard VXLAN packet format. The packet format of Huawei CloudEngine S series switches is different from the standard VXLAN packet format. The reserved fields in the VXLAN packet format are user-defined.

Basic Concepts of VXLAN: NVE

- Network Virtualization Edge (NVE):
  - An NVE is a network entity that implements network virtualization functions. A hardware or software switch can work as an NVE.
  - NVEs run VXLAN and construct a Layer 2 virtual network over a Layer 3 network. SW1 and SW2 in the figure are NVEs.

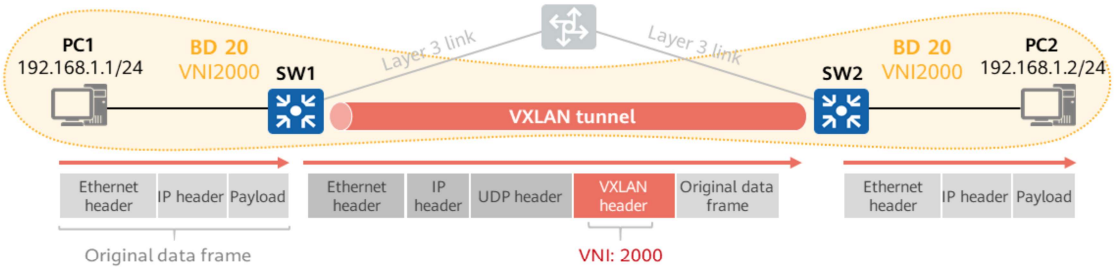




- A VXLAN tunnel is defined by a pair of VTEPs.
- The source VTEP encapsulates packets and sends the encapsulated packets to the destination VTEP through the VXLAN tunnel. After receiving the encapsulated packets, the destination VTEP decapsulates the packets.
- Generally, the IP address of a loopback interface on a device is used as the VTEP address.

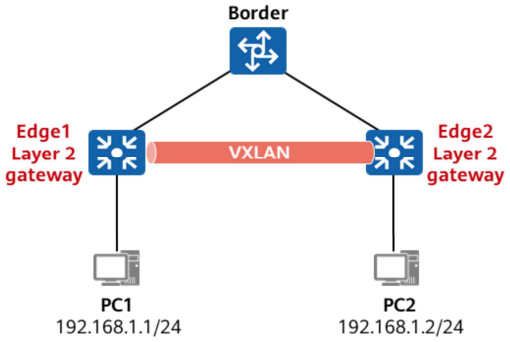
Basic Concepts of VXLAN: VNI and BD

- VNI:
  - VNIs are similar to VLAN IDs and are used to differentiate VXLANs. VMs in different VXLANs cannot communicate with each other at Layer 2.
  - A tenant can have one or more VNIs. The VNI length is 24 bits, and a maximum of 16M tenants are supported.
- Bridge domain (BD):
  - VLANs are used to divide broadcast domains on a traditional network. Similarly, BDs are used to divide broadcast domains on a VXLAN network. A BD identifies a large Layer 2 broadcast domain on a VXLAN network.
  - VNIs are mapped to BDs in 1:1 mode. Terminals in the same BD can communicate with each other at Layer 2.

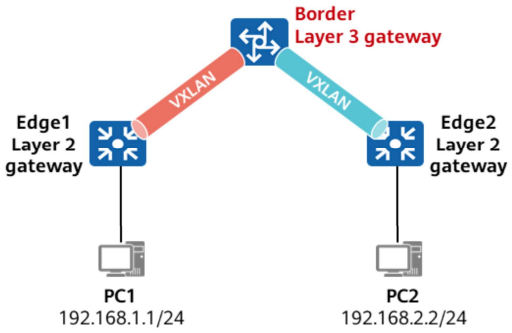


HUAWEI

Basic VXLAN Concepts: Layer 2 and Layer 3 VXLAN Gateways



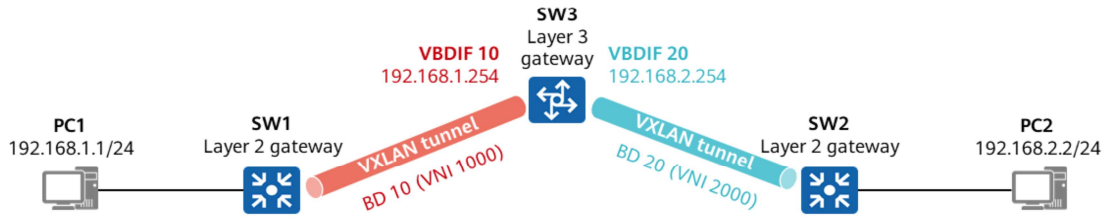
Layer 2 gateway: allows traffic to enter a VXLAN network or is used for intra-subnet communication between terminals on the same VXLAN network.



Layer 3 gateway: enables inter-subnet communication between terminals on a VXLAN network and allows terminals to access external networks (non-VXLAN networks).

HUAWEI

Basic Concepts of VXLAN: VBDIF Interface

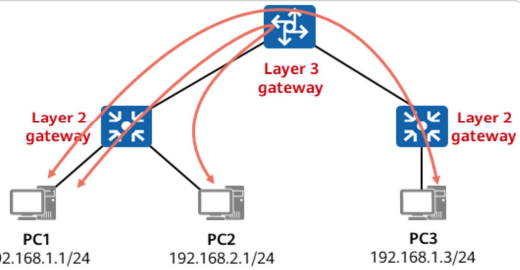


- VLANIF interfaces are used for communication between broadcast domains on a traditional network. Similarly, VBDIF interfaces are used for communication between BDs on a VXLAN network.
- A VBDIF interface is a Layer 3 logical interface created for a BD on a Layer 3 VXLAN gateway.
- VBDIF interfaces allow users on different network segments to communicate with each other through VXLAN networks, implements communication between VXLAN and non-VXLAN networks, and connects a Layer 2 network to a Layer 3 network.

HUAWEI

Basic VXLAN Concepts: Distributed and Centralized Gateways

Centralized gateway

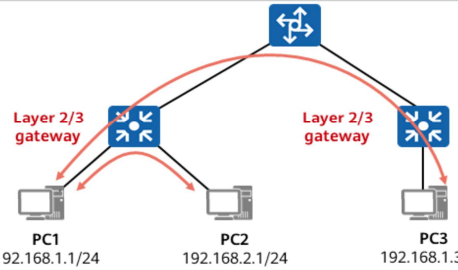


A Layer 3 gateway is deployed on a device. All inter-subnet traffic is forwarded through the Layer 3 gateway, implementing centralized traffic management.

Advantage: Inter-subnet traffic is managed in a centralized manner, simplifying gateway deployment and management.

Disadvantage: The forwarding path is not optimal. The ARP entry specification is a bottleneck. The centralized gateway needs to maintain ARP entries of a large number of terminals connected to the network through VXLAN.

Distributed gateway



VTEPs function as both Layer 2 and Layer 3 gateways. Non-gateway nodes are unaware of VXLAN tunnels and only forward VXLAN packets.

Advantage: VTEPs only need to learn ARP entries of terminals connected to them. Therefore, the number of ARP entries supported is no longer a bottleneck on distributed VXLAN gateways, and the network scalability is improved.

Disadvantage: The deployment, configuration, and implementation of distributed gateways are complex, and the deployment workload is heavy.

HUAWEI

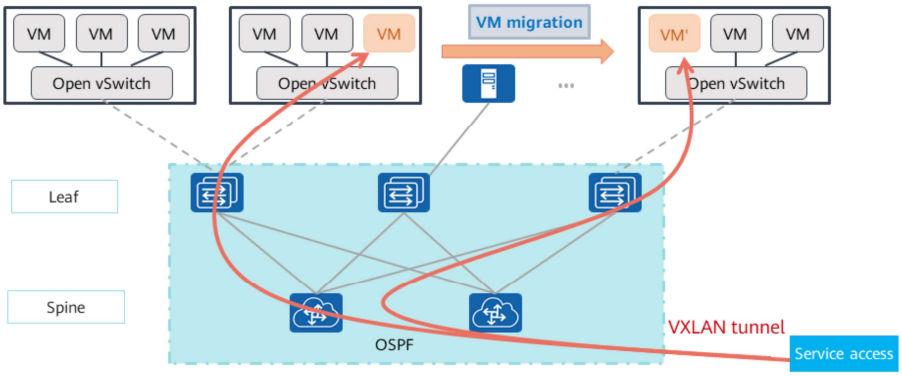
VXLAN Tunnel Establishment Modes

- A VXLAN tunnel is identified by a pair of VTEP IP addresses. Packets are encapsulated on a VTEP and then transmitted in the VXLAN tunnel via routing. A VXLAN tunnel can be established successfully as long as the VTEPs at both ends of the tunnel have reachable routes to each other's IP address at Layer 3.
- VXLAN tunnels are classified into the following types based on the VXLAN tunnel creation mode:
  - Static tunnel: The local and remote VNIs, VTEP IP addresses, and ingress replication list (head-end peer-list) are configured manually.
  - Dynamic tunnel: VXLAN tunnels are dynamically established using BGP EVPN. The local and remote VTEPs establish a BGP EVPN peer relationship and then exchange BGP EVPN routes to learn the VNIs and VTEP IP addresses from each other.

HUAWEI

Application of VXLAN in a Cloud DC

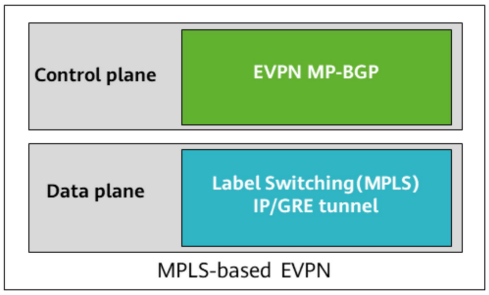
- In a cloud DC, a large number of services are deployed on VMs, which can be live migrated in a server cluster. When a VM is migrated to a server connected to another leaf node, the IP address and MAC address of the VM remain unchanged. In this way, service access is not interrupted.
- Spine-leaf + VXLAN is the best practice in this scenario.



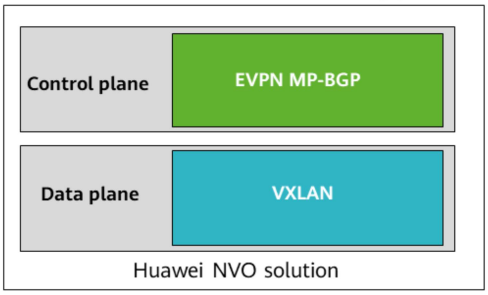
HUAWEI

EVPN

- EVPN is initially defined in RFC 7432. The MPLS-based EVPN meets the requirements for high bandwidth and complex QoS.
- The virtualization technology is introduced to the cloud DC. A host can carry multiple VMs that belong to different tenants. This poses new requirements on the network, and the Network Virtualization Overlay (NVO) solution is used to solve this problem.
- Huawei NVO solution uses BGP EVPN as the control plane and VXLAN tunnels as the data forwarding plane.



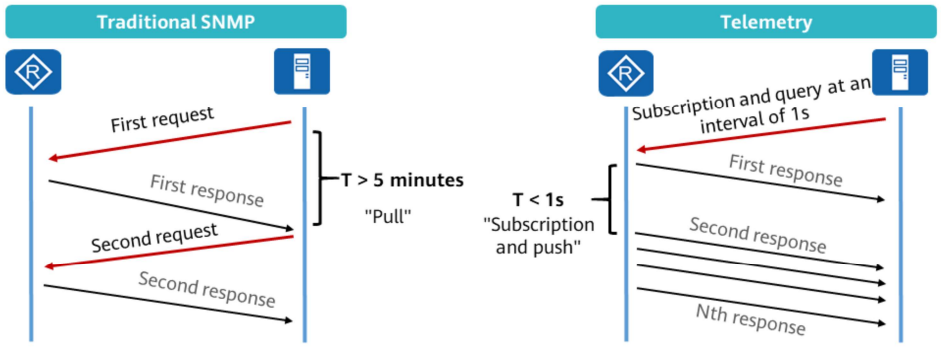
MPLS-based EVPN



Huawei NVO solution

Telemetry

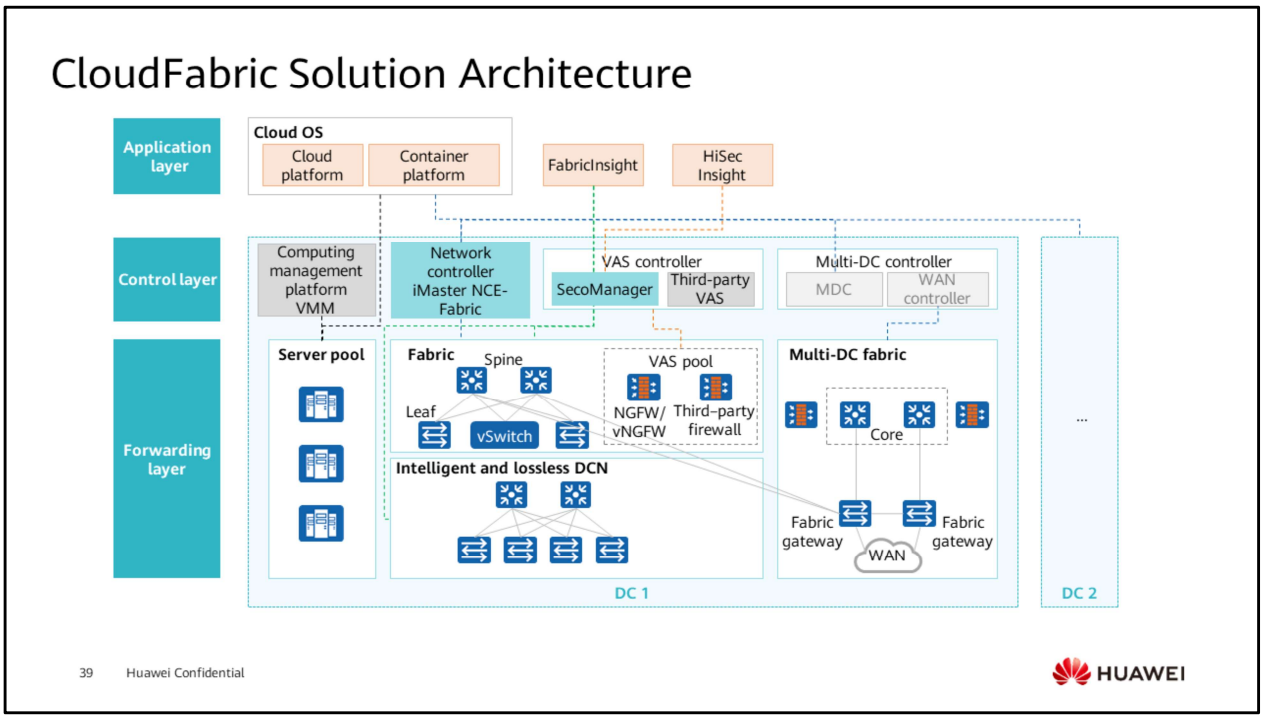
- Telemetry is also called network telemetry, which is a technology that remotely collects data from physical or virtual devices at a high speed.
- Compared with SNMP, telemetry has a subsecond-level collection interval. Devices periodically report information in push mode, providing real-time and high-speed data collection.



Contents

1. DC and DCN Overview
2. Terms and Key Technologies of DCs
3. **CloudFabric Solution and Related Technologies**

HUAWEI

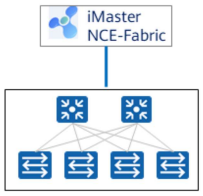


- Application layer
  - Cloud OS
    - OpenStack-based cloud platform (open-source OpenStack or Huawei FusionSphere): collaboratively manages computing, storage, and network resources.
    - Container provisioning platform: creates and provisions containers.
  - Analyzer: iMaster NCE-FabricInsight (FabricInsight for short) is used to analyze the network and application health status of DCNs. HiSec Insight, an advanced threat analysis system, can effectively collect massive basic network data, such as traffic on the network and network status and security logs of various devices. Through real-time and offline big data analysis, HiSec Insight can effectively detect potential threats on the network based on machine learning technologies, expert reputation, and intelligence drive, implementing network-wide security situational awareness.
- Control layer
  - Computing management platform: The Virtual Machine Management (VMM) implements virtualization and resource management at the computing layer.
  - Network controller: iMaster NCE-Fabric is used to centrally manage and control cloud DCNs. It provides automatic mapping from applications to physical networks, resource pool deployment, and visualized O&M, helping customers dynamically schedule service-centric network services.

    - VAS controller: The SecoManager implements centralized security policy management and control for firewalls, monitors events in real time, comprehensively analyzes security events such as attacks, and provides statistical reports in different formats. These benefits help customers understand the network security status at any time.
    - Multi-DC controller (MDC) and WAN controller (DCI link optimization)
- Forwarding layer
    - Network device
        - Physical switches use CloudEngine (CE for short) series. They are high-performance cloud switches designed for next-generation DCs and support extensive DC and campus service features.
        - The virtual switch uses Huawei CloudEngine 1800V (CE1800V for short). The CE1800V is an intelligent software switch designed for virtualization environments of enterprises and industry DCs.
    - VAS device
        - Huawei next-generation firewall (NGFW) or virtual NGFW (vNGFW) provides NAT, VPN, virtualization, and multiple security features for DCs.
        - The load balancer (LB) flexibly load balances services for DCs.
        - Third-party VAS devices can be deployed in DCs, and traffic can be diverted to them.

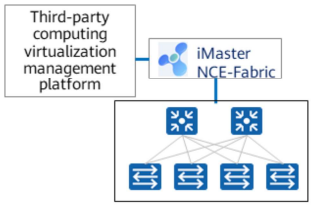
Huawei CloudFabric Solution Scenarios

Scenario 1: Network Virtualization - Hosting



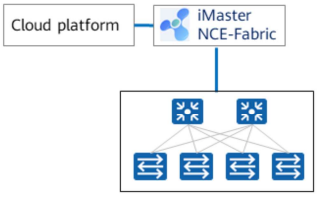
- No computing platform or cloud platform is available. The network administrator uses iMaster NCE-Fabric to provision network services.

Scenario 2: Network Virtualization - Computing



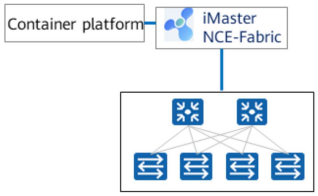
- The network administrator uses iMaster NCE-Fabric to configure the network.
- After detecting that a VM goes online, the network automatically configures a logical port for connecting to the VM.

Scenario 3: Cloud-network integration



- The cloud platform provides a unified interface for managing computing, storage, and network resources.
- The cloud platform delivers network services through iMaster NCE-Fabric.

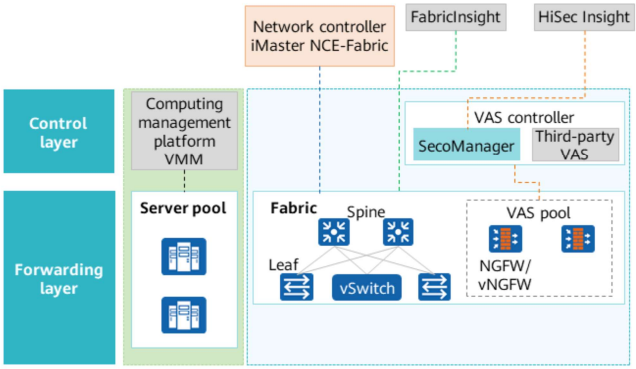
Scenario 4: Container network



- The container platform provides a unified interface for managing computing and network resources.
- The container platform delivers network services through iMaster NCE-Fabric.

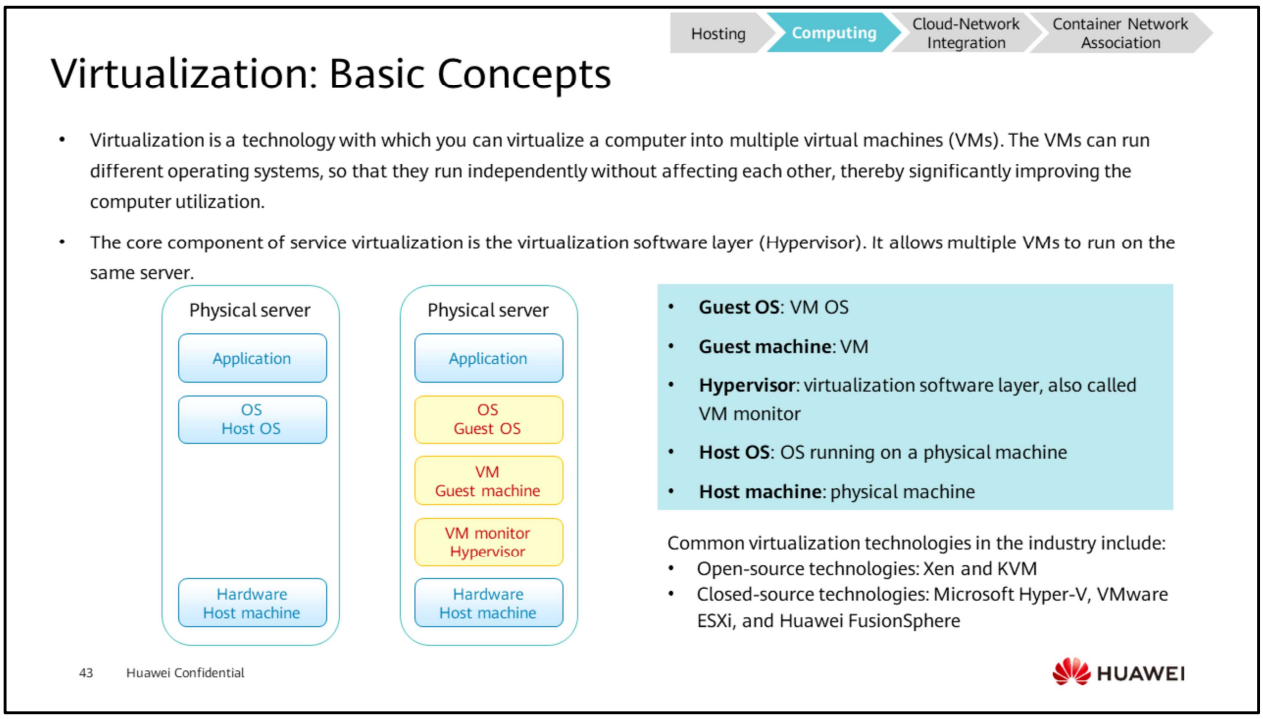
HUAWEI

Huawei CloudFabric Solution: Network Virtualization - Hosting Scenario



- In this scenario, no computing virtualization platform is connected, and network administrators directly deliver network resources and services.
- Network Virtualization - Hosting has the following characteristics:
  - Service automation: iMaster NCE-Fabric automatically delivers network configurations and provides multi-path detection, single-path detection, network connectivity detection, and pre-orchestrated emergency plans to implement fast fault locating and isolation.
  - Self-service: iMaster NCE-Fabric provides service operation portals and open APIs. Users can develop service portals and provide self-service capabilities for tenants to quickly deploy and adjust services.
  - Large Layer 2 network: The VXLAN overlay technology is used to build a large Layer 2 network, improving resource utilization.

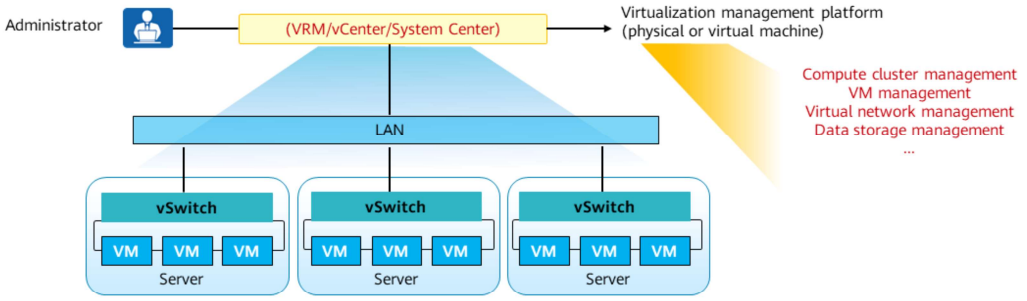
HUAWEI



- Multiple virtualization technologies including computing, storage, and network virtualization are required to implement virtualization:
  - Computing virtualization: includes CPU, memory, and I/O virtualization. CPU virtualization technologies include Intel VT-x and AMD-V. Memory virtualization technologies include shadow page table and extended page table (EPT). I/O virtualization technologies include QEMU simulation, Virtio, and SR-IOV.
  - Storage virtualization: For example, Huawei FusionCompute implements storage virtualization using the high-performance virtual cluster file system VIMS.
  - Network virtualization: is implemented through virtual switches. For example, Huawei FusionCompute implements network virtualization using distributed virtual switches (DVSs).

Virtualization: Virtualization Management Platform

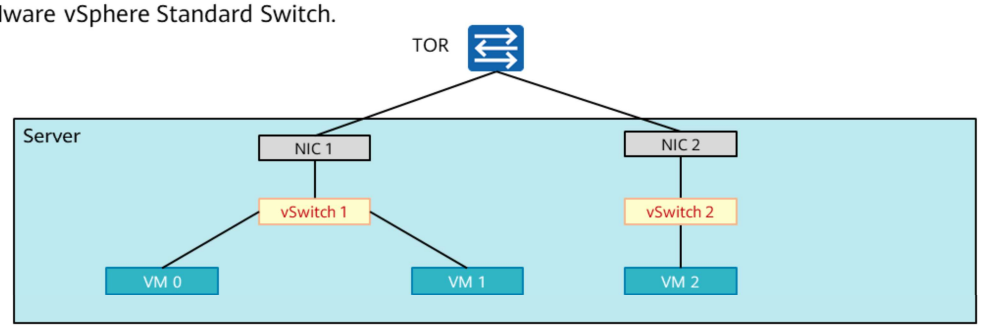
- When there are multiple virtualized servers, a platform is needed to centrally manage and control them, that is, virtualization management platform.
- A virtualization management platform manages virtual resources, such as VMs, virtual switches (vSwitches), and cluster file systems, and provides GUIs for performing operations on these resources, such as creating VMs and vSwitches and adding data stores.
- Mainstream virtualization platforms in the industry include Huawei VRM, VMware vCenter, and Microsoft System Center.



HUAWEI

Virtual Network: vSwitch

- vSwitch is a software switch deployed on a server. It is mainly used to implement communication between virtual interfaces on a server.
- Common vSwitches in the industry include the open-source Open vSwitch, Huawei CE1800V, and VMware vSphere Standard Switch.

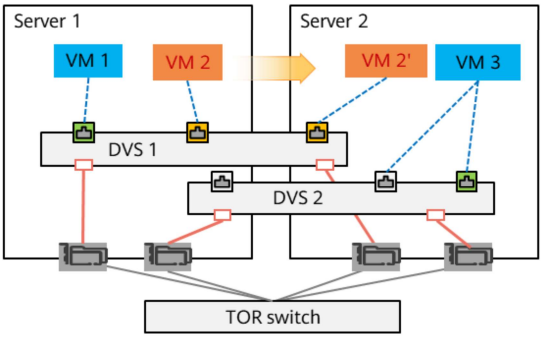


HUAWEI

- Open vSwitch (OVS) is an open-source project of Apache 2.0. It is a virtual switch running on a virtualization platform (such as KVM and Xen) and is the mainstream virtual switch in the industry. Strictly speaking, OVS is a DVS.
- vSwitches provide various functions. For example, OVS provides Layer 2 switching for dynamically changed endpoints to control access policies, isolate networks, and monitor traffic on virtual networks.

Virtual Network: DVS

- DVS provides virtual network functions across hosts to meet VM migration requirements. VM migration requires that the network status (such as the VLAN, IP address, and network policy) remain unchanged after the migration. Common vSwitches cannot meet this requirement.
- In this example, after VM 2 is migrated to VM 2', its network status remains unchanged.



DVS provides two types of virtual interfaces:

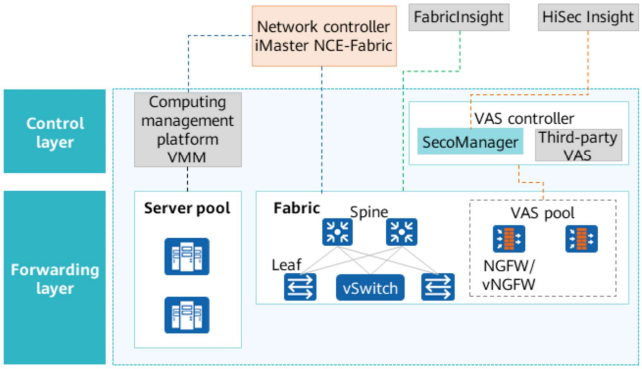
- Port group: is used to bind VM NICs and add different VLAN tags to VM traffic.
- Uplink: is used to select a physical network port to be bound to guide traffic forwarding.

Port group
Uplink

HUAWEI

- DVS is the distributed virtual switch on the Huawei virtualization platform VRM.
- In vCenter, vSphere Distributed Switch (VDS) is used.
- Huawei CloudFabric Solution supports interconnection with mainstream virtualization platforms.

Huawei CloudFabric Solution: Network Virtualization - Computing Scenario



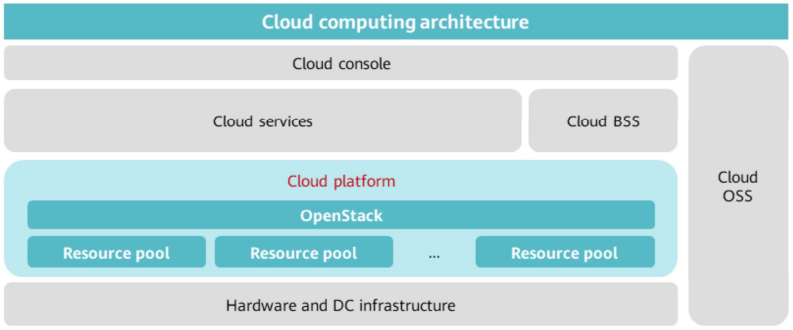
- In this scenario, a computing virtualization platform is connected and no cloud platform is connected. iMaster NCE-Fabric and the computing management platform deliver services.
- Network Virtualization - Computing has the following characteristics:
  - Service automation: Network devices are managed by the controller and are interconnected with the VMM to implement automatic network service deployment. This shortens the deployment period of traditional networks to minutes.
  - Elastic resource pooling: VXLAN overlay network virtualization and SFC technologies are used to virtualize network forwarding resources and service resources into pools. This improves the flexibility and scalability of the network framework and allows the framework to evolve.
  - Refined O&M: The application, logical, and physical networks are visible, the service quality is visible and controllable, and the network status is visible. The solution provides multiple fault location and O&M methods, such as multi-path detection, single-path detection, network connectivity detection, and pre-orchestrated emergency plan. It implements fast fault location and isolation based on iMaster NCE-FabricInsight.

HUAWEI

Hosting | Computing | Cloud-Network Integration | Container Network Association

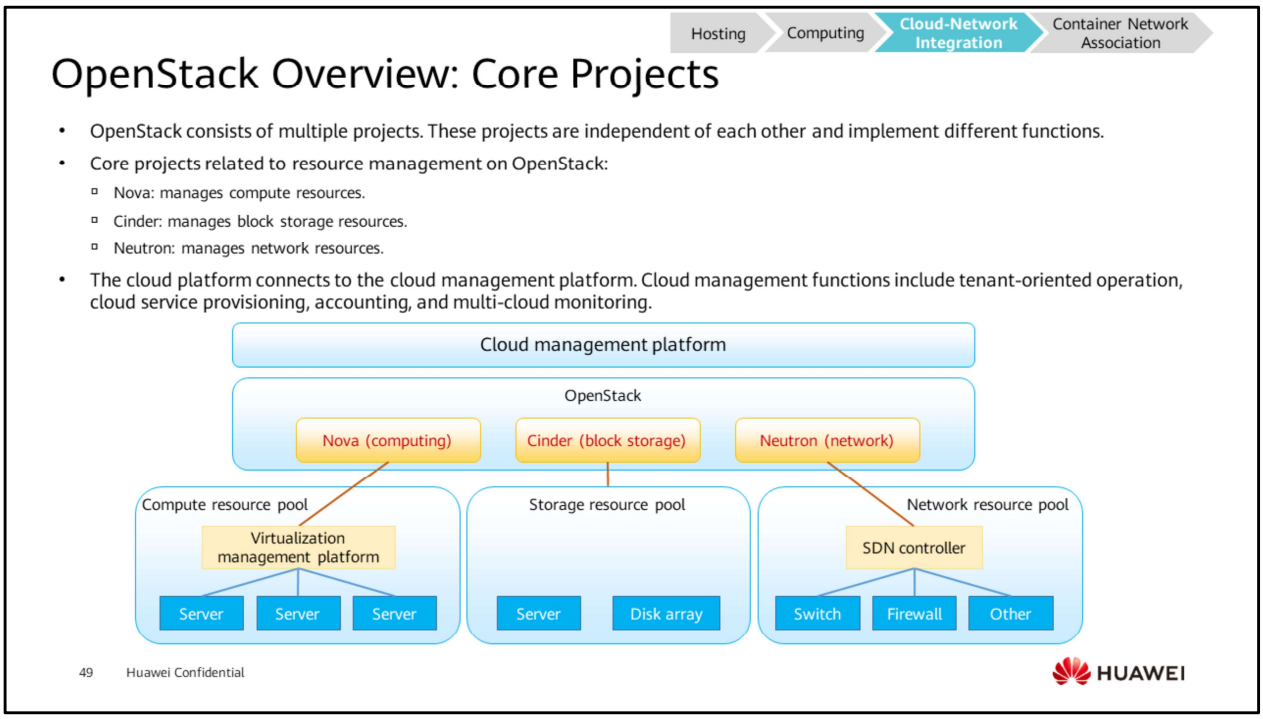
OpenStack Overview

- OpenStack is an open-source cloud operating system that controls large pools of computing, storage, and network resources throughout a DC. After OpenStack is deployed, users can manage resources through the web UI, CLI, or APIs.
- OpenStack is not a cloud computing platform, but a key component of cloud computing. OpenStack is designed for resource management and can be used to manage computing, storage, and network resource pools of heterogeneous vendors.



HUAWEI

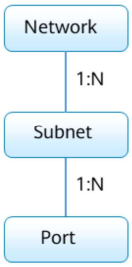
- FusionSphere OpenStack is Huawei's commercial release with a built-in Huawei KVM virtualization engine.

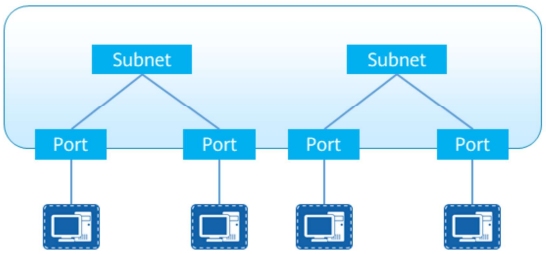


- In the HUAWEI CLOUD Stack (HCS) architecture, ManageOne functions as the cloud management platform, which provides refined operation, intelligent O&M, and centralized management.
- FusionSphere OpenStack is Huawei's commercial release with a built-in Huawei KVM virtualization engine.
- Nova: is a core component of OpenStack. It provides virtual compute resources for OpenStack and manages the lifecycle of VMs, including creating, modifying, and deleting VMs.
- Cinder (block storage): is an indispensable component in the virtual infrastructure and is the basis for storing VM image files and data used by VMs.
- Neutron: is a core component of OpenStack and provides network services for it. Neutron enables OpenStack to flexibly divide physical networks and provide an independent network environment for each tenant in multi-tenancy scenarios.
- In the Huawei CloudFabric Solution, Neutron interconnects with the SDN controller to manage hardware network resources.

Neutron's Core Network Models

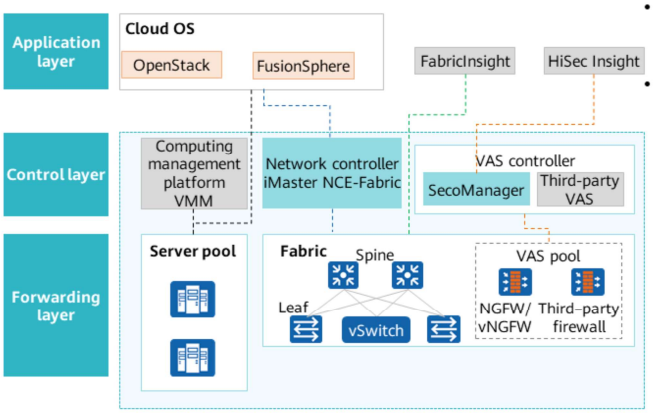
- Neutron provides network services for OpenStack and has three core network models: network, subnet, and port.
- Network: indicates a Layer 2 network. It describes basic network information, such as the name and creation time.
- Subnet: indicates a network segment, for example, 192.168.10.0/24. A subnet is associated with a network.
- Port: indicates a virtual port. A MAC address and an IP address are defined on the port. The IP address is obtained from the subnet.





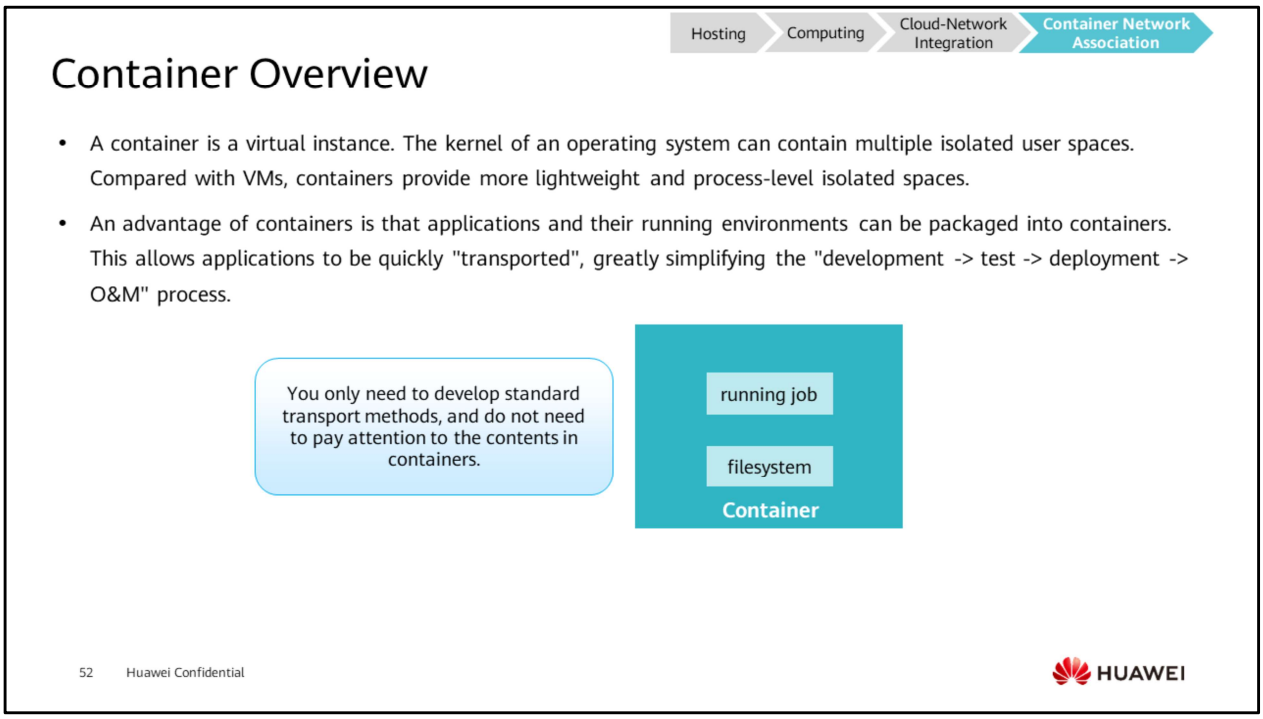
HUAWEI

Huawei CloudFabric Solution: Cloud-Network Integration Scenario



- In this scenario, the cloud platform provides a unified computing and network management interface. iMaster NCE-Fabric interconnects with the cloud platform.
- Cloud-network integration has the following characteristics:
  - Service automation: Network devices are managed by the controller and are interconnected with the cloud platform and VMM to implement automatic network service deployment. This shortens the deployment period of traditional networks to minutes.
  - Elastic resource pooling: VXLAN overlay network virtualization and SFC technologies are used to virtualize network forwarding resources and service resources into pools. This improves the flexibility and scalability of the network framework and allows the framework to evolve.
  - Refined O&M: The application, logical, and physical networks are visible, the service quality is visible and controllable, and the network status is visible. The solution provides multiple fault location and O&M methods, such as multi-path detection, single-path detection, network connectivity detection, and pre-orchestrated emergency plan. It implements fast fault location and isolation based on iMaster NCE-FabricInsight.

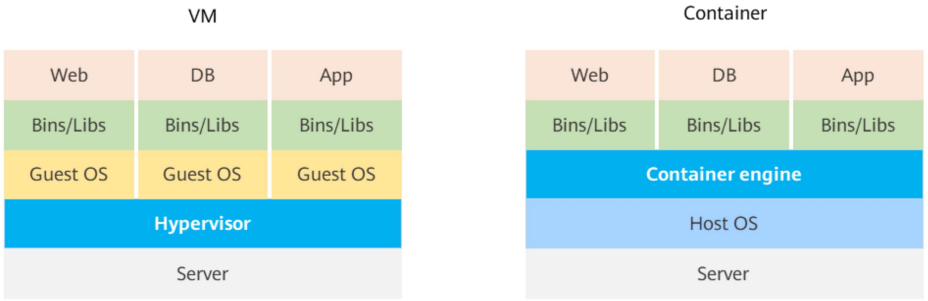
HUAWEI



- On Linux, the kernel space is distinguished from the user space. The Linux OS and drivers run in the kernel space, and applications run in the user space.
- A container is a runtime instance of an image.
- Container image
  - An application and its dependencies (including all files and directories of the complete OS) can be packaged into an image.
  - The image contains all dependencies required for application running. You only need to run the image in the isolated sandbox without any modification or configuration.
  - Images focus on packaging applications and their runtime environments in a unified format. This ensures high consistency between the local environment and the cloud environment.

Comparison Between Containers and VMs

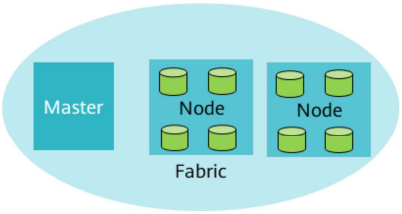
- Compared with VMs, containers share the kernel of an operating system, and therefore are less independent, and only process-level isolation is provided.
- The startup speed, running performance, and server resource usage of containers are higher than those of VMs.



HUAWEI

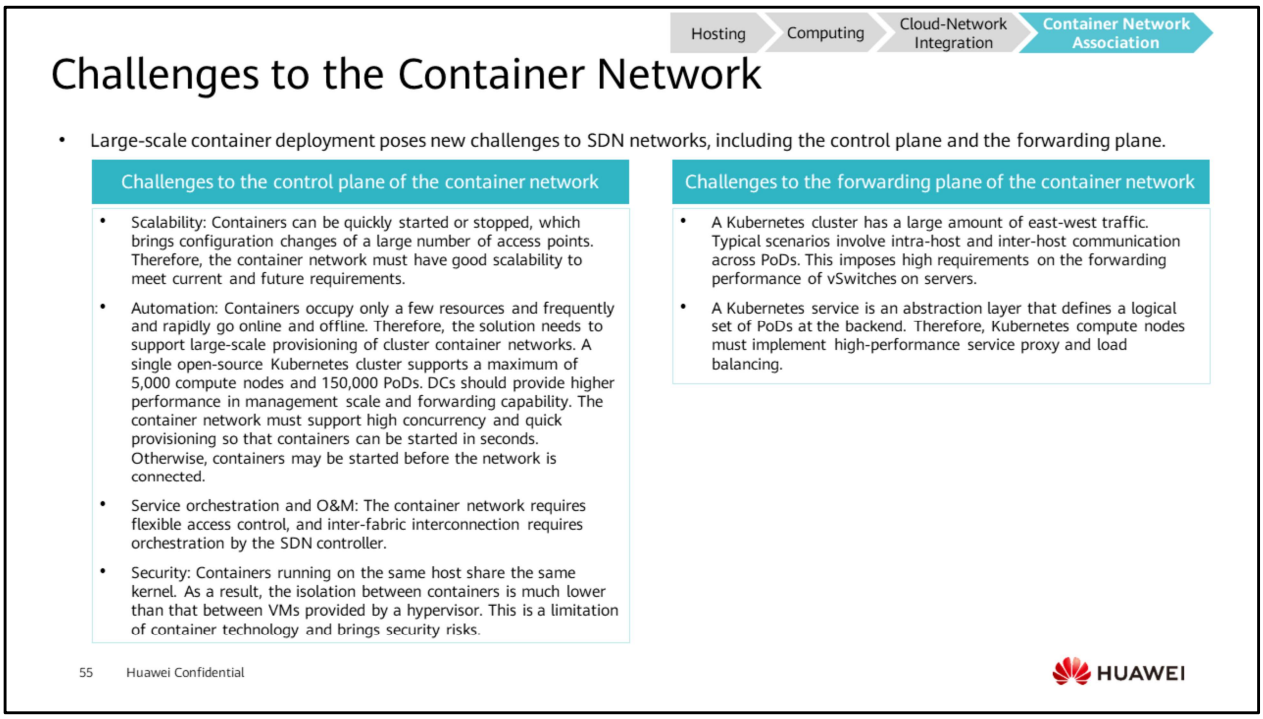
Container Orchestration: Kubernetes

- Container orchestration is used to manage a series of operations for defining, creating, and configuring a container. Technically, large-scale applications of containers need to be centrally managed and scheduled by the container orchestration engine. Commercially, users are willing to pay for the tools with platform capabilities, because users' services or applications need to be deployed on a platform.
- Kubernetes is an open-source project initiated by Google. It aims to manage containers across hosts and provide basic deployment, maintenance, and application scaling. A Kubernetes cluster consists of the master node (master) and compute nodes (node).
  - The master node provides the cluster's control plane and makes global decisions about the cluster, such as scheduling. Generally, user containers do not run on the master node.
  - There are many compute nodes in a cluster, which are used to run containerized applications. The master node installs kubelet on each compute node as the agent for managing the node.



HUAWEI

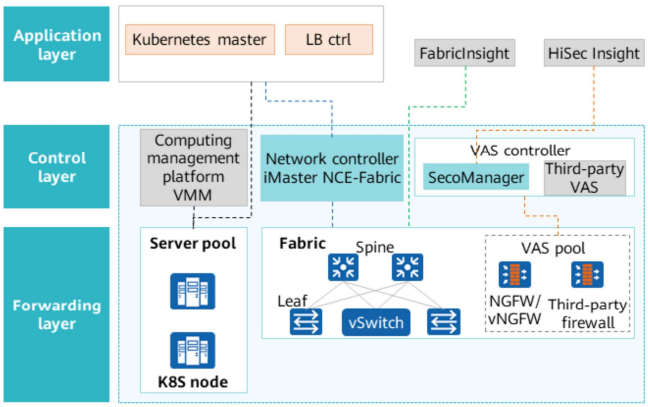
- Kubernetes is an open-source container scheduling platform across host clusters. It can manage various underlying containers. It is similar to OpenStack, which manages a large number of virtual nodes at the bottom layer to form a large cluster to provide cloud services. Kubernetes provides a container-centric architecture by managing the underlying container nodes and automating deployment, scaling, and operations of application containers.
- In terms of container technology development, the major orchestration platforms include Docker's Swarm, Apache's Mesos (initially developed by AMPLab of the University of California, Berkeley), and Google's Kubernetes. Kubernetes (K8S) has become the standard of container management platforms.



- A container network can be deployed independently or associated with a physical network.

Hosting > Computing > Cloud-Network Integration > Container Network Association

Huawei CloudFabric Solution: Association Between the Container Network and the Physical Network

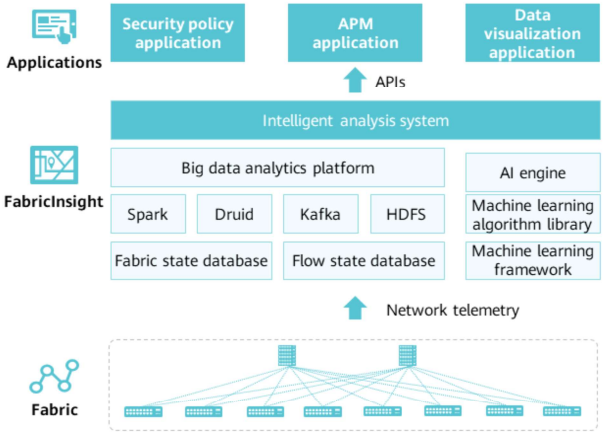


- In this scenario, iMaster NCE-Fabric and the container orchestration platform deliver services.
- Association between the container network and the physical network has the following characteristics:
  - Service automation: The containers on the container network can be quickly started and the network has good scalability, meeting future requirements.
  - Refined O&M: The application, logical, and physical networks are visible, the service quality is visible and controllable, and the network status is visible. The solution provides multiple fault location and O&M methods, such as multi-path detection, single-path detection, network connectivity detection, and pre-orchestrated emergency plan. It implements fast fault location and isolation based on iMaster NCE-FabricInsight.

HUAWEI

- The LB ctrl plug-in is deployed on the Kubernetes master node to listen to the service and ingress object events of the API server and deliver the L4-L7 service orchestration configuration to LBs.

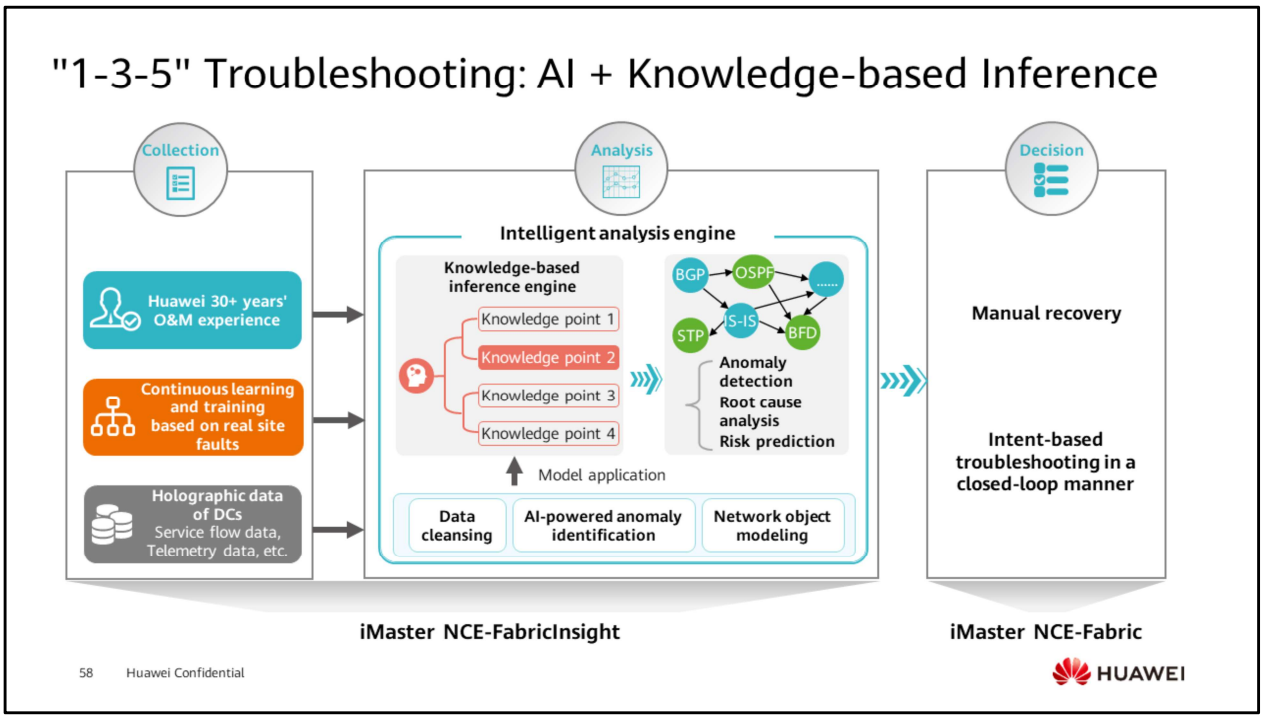
iMaster NCE-FabricInsight Intelligent O&M Solution



- Based on Huawei-developed big data analytics platform, iMaster NCE-FabricInsight (FabricInsight for short) receives data from network devices in telemetry mode and uses intelligent algorithms to analyze network data.
- The overall architecture of FabricInsight consists of three parts: network devices, FabricInsight collector, and FabricInsight analyzer. In standard clustering deployment of FabricInsight, the analyzer and collector are combined. That is, no independent collector server needs to be deployed. FabricInsight uses the microservice architecture. Each service is deployed in multi-instance mode, which features high reliability and scalability. The system management capacity can be expanded by adding instance nodes.

HUAWEI

- Network devices
  - Network devices include Huawei CloudEngine switches and NetEngine routers. For details about the models and supported specifications, see the specification list of the corresponding version.
  - These devices use telemetry to report performance metrics such as interface traffic based on Google Remote Procedure Call (gRPC). Devices are connected to FabricInsight as gRPC clients. Users can run commands to configure the telemetry sampling function on the devices. The devices then proactively establish a gRPC connection with the target collector and send data to the collector. The current version supports the following sampling metrics: CPU and memory usage of devices and boards; number of sent and received bytes, number of discarded sent and received packets, and number of sent and received error packets of interfaces; number of congested bytes of queues; packet loss behavior data. For details about metrics and device models, see the product specification list.
- FabricInsight collector
  - The FabricInsight collector receives data reported by network devices in telemetry mode, including performance metric data reported through gRPC. The collector parses the metrics, combines and compresses the metric data, and reports the data to the analyzer.
- FabricInsight analyzer
  - The FabricInsight analyzer receives performance metrics from switches. In addition, the FabricInsight analyzer establishes dynamic baselines for some performance metrics based on the AI algorithm, detects exceptions, and displays the analysis result on the GUI.



- "1-3-5" troubleshooting: Faults can be detected within 1 minute, located within 3 minutes, and rectified within 5 minutes.

Quiz

1. (Short-answer question) What is a DC?
2. (Single-answer question) Which of the following protocols is used as the control plane in the Huawei CloudFabric Solution? ( )
   A. VLAN
   B. VXLAN
   C. EVPN
   D. VRRP

HUAWEI

1. A DC is a complex set of facilities including buildings. It not only includes the computer system and related devices (such as communications and storage devices), but also includes redundant data communication connections, environmental control devices, monitoring devices, and various security devices.
2. C

Summary

- DCN has evolved from DCN 1.0 and DCN 2.0 to DCN 3.0. Currently, the spine-leaf architecture is typically used together with VXLAN and EVPN.
- DCs carry many mission-critical IT services, including virtualization, cloud computing, distributed storage, and AI applications. Knowing these services helps you better understand DCN technologies.
- Key DCN technologies include load balancing, M-LAG, VXLAN, EVPN, telemetry, and intelligent O&M.
- This course provides an overview of the preceding contents to help you get a preliminary understanding of DCs.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Data Center Network Design

Financial Data Center Network Design



HUAWEI

Foreword

- The Huawei CloudFabric 3.0 Solution provides customers with intelligent, lossless, and ultra-broadband infrastructure networks. The solution supports one-click automatic deployment, artificial intelligence (AI)-based intelligent O&M, and on-demand self-service customization, helping customers quickly complete the planning, design, and deployment of industry data center network (DCN) solutions.
- Based on the CloudFabric 3.0 solution, this course uses the financial DCN as an example to describe the typical DCN planning and design process, including the DCN overview, design principles and methods, and design cases.
- The design cases cover design for the DCN logical architecture, physical architecture, key network partitions, and key network features.

HUAWEI

Objectives

- Upon completion of this course, you will be able to:
  - Describe the architecture and key components of Huawei CloudFabric 3.0 solution.
  - Describe the DCN design principles and methods.
  - Complete the underlay and overlay network design for a DCN based on actual requirements.
  - Complete the high reliability design, network security design, and network O&M design for a DCN based on actual requirements.

HUAWEI

Contents

1. **CloudFabric Solution and DCN Overview**
2. DCN Planning and Design Methods
3. Typical DCN Design Cases

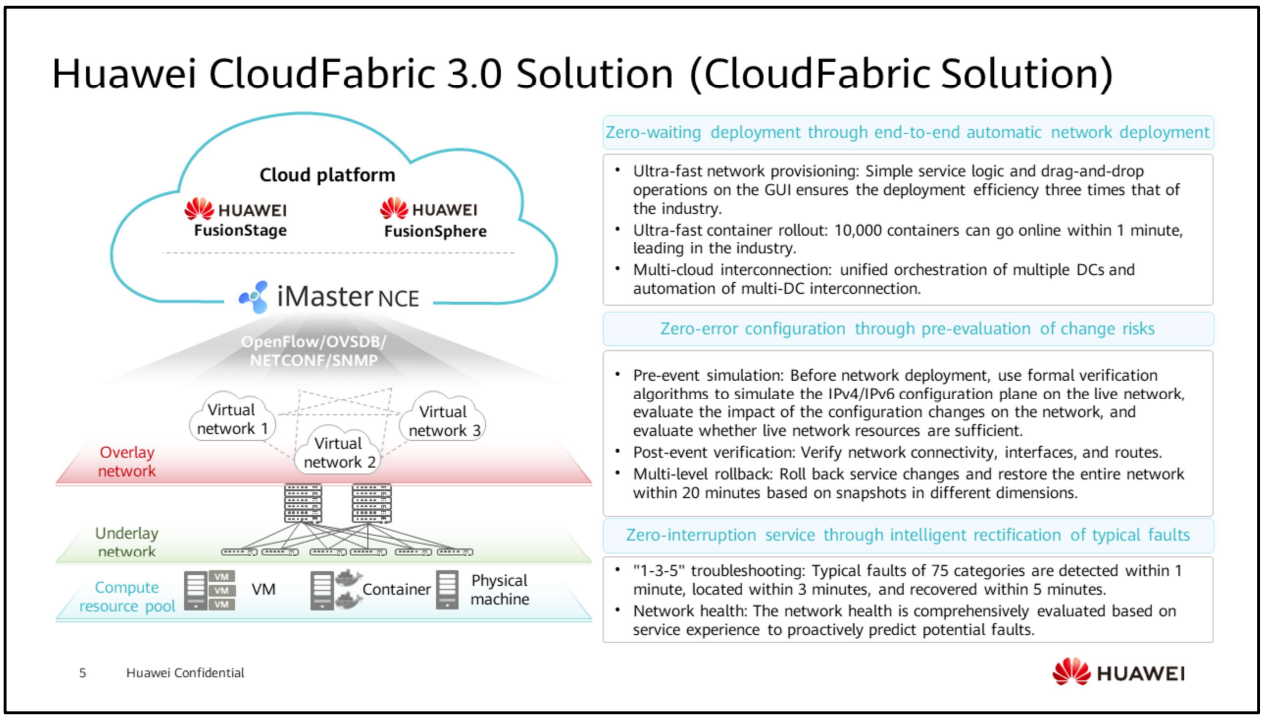
HUAWEI

Challenges Faced by DCNs

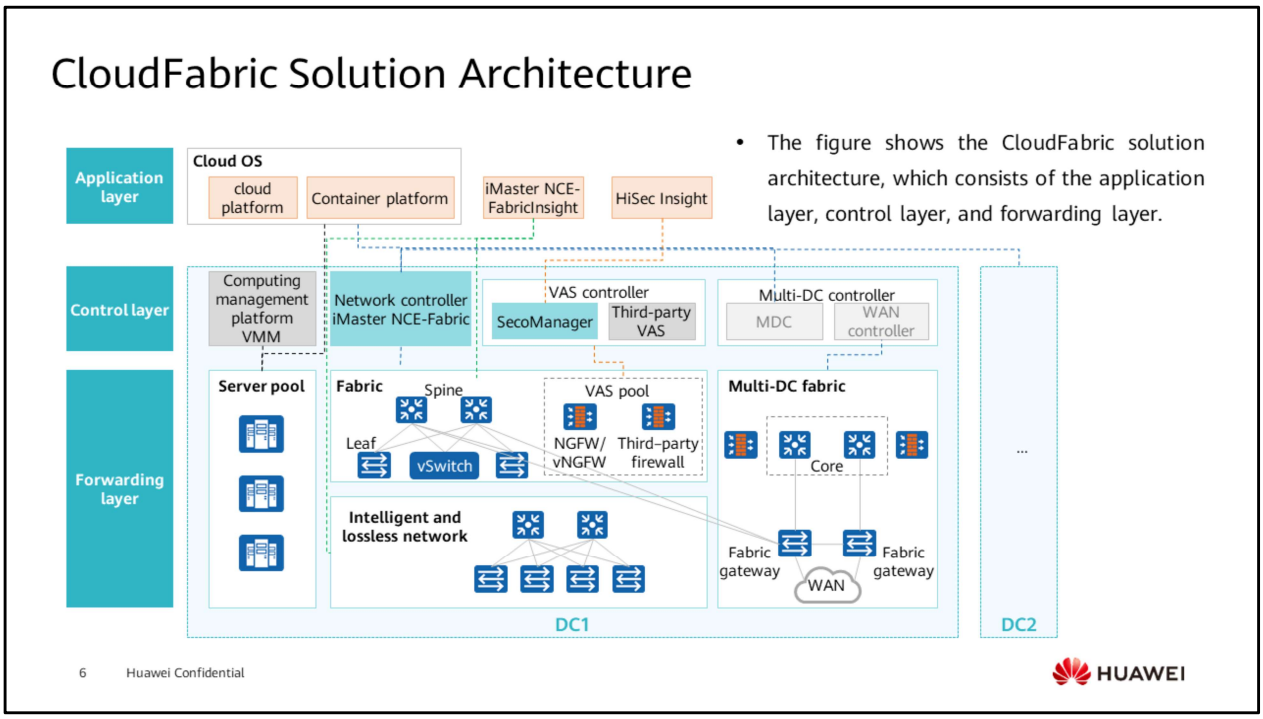
- Currently, more and more enterprises and organizations use data centers (DCs) to provide various services. In addition, cloud computing is a new business model, and migrating various applications to the cloud has become a mainstream trend. Typical features of application migration to the cloud include on-demand self-service, extensive access, resource pooling, fast and elastic resource scaling, and measurable services.
- Under this context, services in cloud computing scenarios are having increasingly demanding requirements on infrastructures of DCNs. To meet these requirements, traditional DCNs must address the following challenges:

Network pooling and deployment automation	Efficient DCN O&M	Configuration and adjustment of massive security policies
• Various types of services need to be brought online and adjusted frequently. The configuration mechanism of traditional networks cannot meet such requirements. • Resource pools need to be set up in a DC and the resource pools of multiple DCs need to be managed in a unified manner. • Traditional networks do not support decoupling of application deployment from physical locations.	• As the DC scale expands, the fault domain also expands. • Virtual networks extend to servers (such as vSwitches), blurring network and IT O&M boundaries. • To improve user experience, administrators need to locate application faults in real time.	• A large number of tenants and applications in a DC require massive security isolation policies. • Security policy adjustment is time- and labor-consuming and error-prone. • Performance of value-added service (VAS) devices varies, and resource pooling and unified management for these devices are required.

HUAWEI



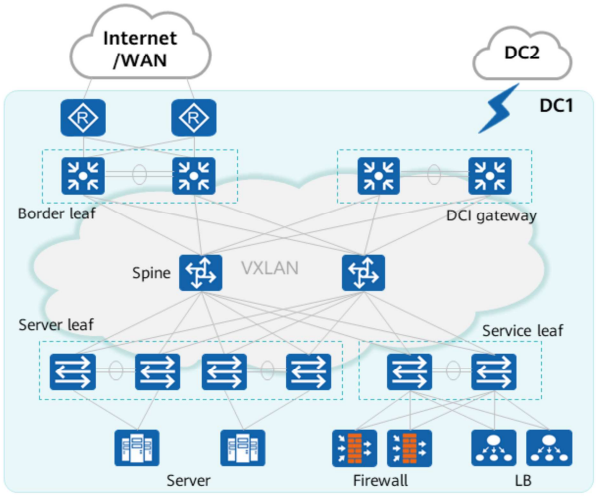
- Software-Defined Networking (SDN) network architecture:
  - Underlay network: It is a physical network or infrastructure network. Any two points on the physical network must be reachable to each other.
  - Overlay network: It is a virtual network built on the underlay network to carry applications on the physical network. Virtual networks of different services are isolated from each other.
- VM: virtual machine



- Application layer:
  - Cloud OS:
    - OpenStack-based cloud platform (open-source OpenStack or Huawei FusionSphere): collaboratively manages computing, storage, and network resources.
    - Container provisioning platform: creates and provisions containers.
  - Analyzer:
    - iMaster NCE-FabricInsight: analyzes the health of DCNs and applications to quickly detect and locate faults.
    - HiSec Insight: security situational awareness system, which proactively detects various security threat events in real time, helping users handle threat events in the entire process.
- Control layer:
  - Computing manager (Virtual Machine Management (VMM)): implements virtualization and resource management at the computing layer. Common VMMs include vCenter and System Center.
  - Network controller: iMaster NCE-Fabric is used to centrally manage and control cloud DCNs. It provides automatic mapping from applications to physical networks, resource pool deployment, and visualized O&M, helping customers dynamically schedule service-centric network services.

    - VAS controller: SecoManager implements centralized security policy management and control for firewalls, monitors events intuitively in real time, comprehensively analyzes security events such as attacks, and provides statistical reports in different formats. It helps customers understand the network security status at any time.
    - Multi-data center controller (MDC): orchestrates services across multiple DCs. It applies to the scenario where a customer has multiple DCs that need to be uniformly orchestrated and managed.
- Forwarding layer:
    - Network devices:
        - Physical switches use CloudEngine (CE for short) series. They are high-performance cloud switches designed for next-generation DCs and support extensive DC and campus service features.
        - Virtual switches are Huawei CloudEngine 1800V (CE1800V for short). The CE1800V is an intelligent software switch designed for virtualization environments of enterprises and industry DCs.
    - VAS devices:
        - Huawei firewalls are used to provide network address translation (NAT), virtual private network (VPN), virtualization, and multiple security features for DCs.
        - Load balancers (LBs) flexibly load balance services for DCs.
        - Third-party VAS devices can be deployed in DCs, and traffic can be diverted to them.

CloudFabric DCN Architecture



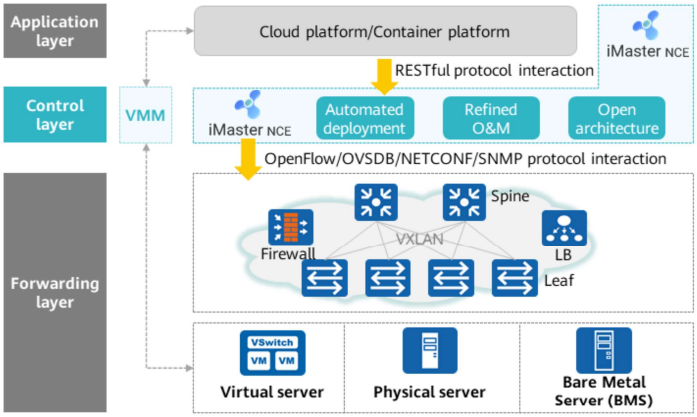
- A DCN is an infrastructure for carrying DC services.
- Multiple DCNs can connect to branches of an enterprise or organization in different areas. In addition, DCNs can connect to the Internet or WAN.
- In the CloudFabric solution, the spine-leaf architecture is recommended for underlay networking.

HUAWEI

- Leaf and spine nodes are fully meshed so that equal-cost multi-path routing (ECMP) paths are available to ensure high availability of the network.
    - Generally, spine nodes are two or more large-capacity switches, and connect to leaf nodes at Layer 3 to establish a fabric network.
    - Leaf nodes can be access switches and are fully meshed to spine nodes. Leaf nodes are located at the border between Layer 2 and Layer 3 on the underlay network, and connect to spine nodes at Layer 3. Generally, a large number of top of rack (TOR) switches are deployed as leaf nodes. It is recommended that leaf nodes use iStack, M-LAG, or SVF to connect to servers.
- A leaf node is an access node on a VXLAN fabric network, which connects various network devices to the VXLAN network. Leaf nodes are classified into the following types based on their access objects:
    - Border leaf node: connects to routers or transmission devices to forward traffic from external networks to the VXLAN fabric network of a DC.
    - Server leaf node: connects compute resources such as virtual and physical servers to the VXLAN fabric network.
    - Service leaf node: connects L4-L7 VAS devices, such as firewalls and LBs, to the VXLAN fabric network.
- A spine node is a core node on a VXLAN fabric network, which provides high-speed IP forwarding and connects to leaf nodes using high-speed interfaces.

CloudFabric Solution Component: iMaster NCE-Fabric

- iMaster NCE-Fabric is an autonomous driving management and control system launched by Huawei for DCN scenarios. It integrates management, control, analysis, and AI functions. It can centrally manage and control cloud DCNs, and provide automatic mapping from applications to physical networks, resource pool deployment, and visualized O&M.

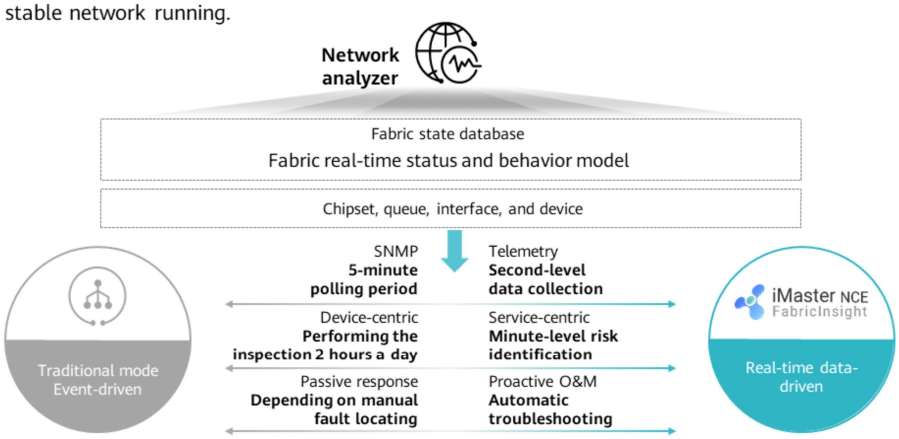


HUAWEI

- iMaster NCE-Fabric uses standard network protocols to manage network resources, and interconnects with compute resources through standardized southbound interfaces to implement collaboration between compute and network resources. It can independently carry out application services and seamlessly interconnect with mainstream cloud platforms or container platforms through standardized open northbound interfaces. Customers can then flexibly deploy and schedule network resources based on service development, enabling DCNs to agilely provide cloud or container services.
- The figure shows the panorama of the CloudFabric solution. iMaster NCE-Fabric is located at the control layer. When iMaster NCE-Fabric does not connect to the cloud platform in the northbound direction and provisions and manages services independently, iMaster NCE-Fabric works at the application layer. Based on an open and highly reliable distributed cluster architecture, iMaster NCE-Fabric can seamlessly interconnect with the industry's mainstream cloud platforms, computing management platforms, and container platforms to provide automatic deployment, refined O&M, and multi-DC disaster recovery capabilities.

CloudFabric Solution Component: iMaster NCE-FabricInsight

- Huawei iMaster NCE-FabricInsight is an intelligent analysis platform oriented to DCNs. Based on big data analytics technologies, it displays the network status in real time and helps users quickly identify, locate, and rectify faults, ensuring stable network running.



HUAWEI

- iMaster NCE-FabricInsight consists of the collector and analyzer.
- Features
  - Visibility of applications and networks, identifying faults in seconds
    - Display service flows and network-wide KPIs through Telemetry in seconds, implementing correlation analysis of services, network paths, and network devices and visualizing network health.
  - AI + knowledge-based inference, locating faults within minutes
    - Train the knowledge-based inference engine based on machine learning to analyze root causes of dozens of faults within minutes, implement edge intelligence based on software and hardware, and perform full packet analysis on TCP and UDP flows.
  - Predictive maintenance based on machine intelligence
    - Construct dynamic baselines, identify device, queue, and port exceptions, and proactively predict traffic and optical module faults.

CloudFabric Solution Component: Network Hardware Products

CloudEngine data center switches



Huawei CloudEngine (CE for short) series switches are high-performance cloud switches designed for next-generation DCNs and high-end campus networks.

USG firewall



USG9500 series



HiSecEngine USG6600E series



USG6600 series

USG6700E series

Huawei firewalls provide NAT, VPN, virtualization, and multiple security features for DCs.

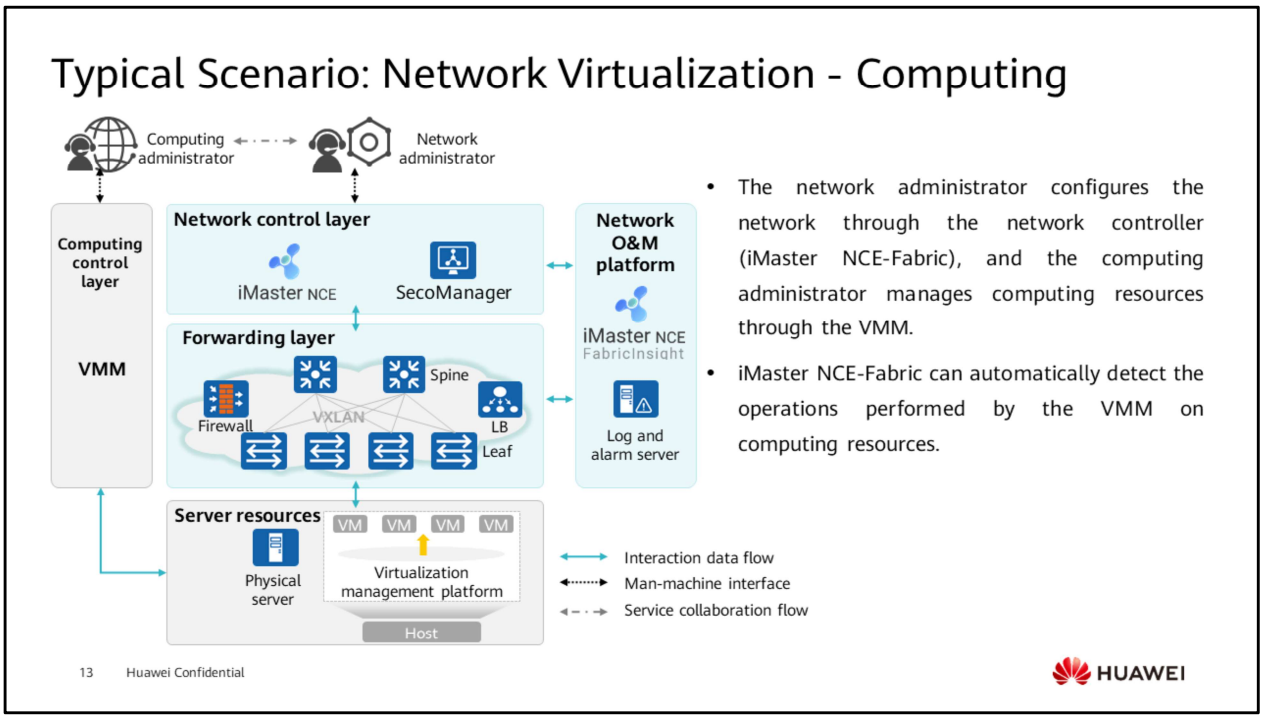
HUAWEI

CloudFabric Software Deployment

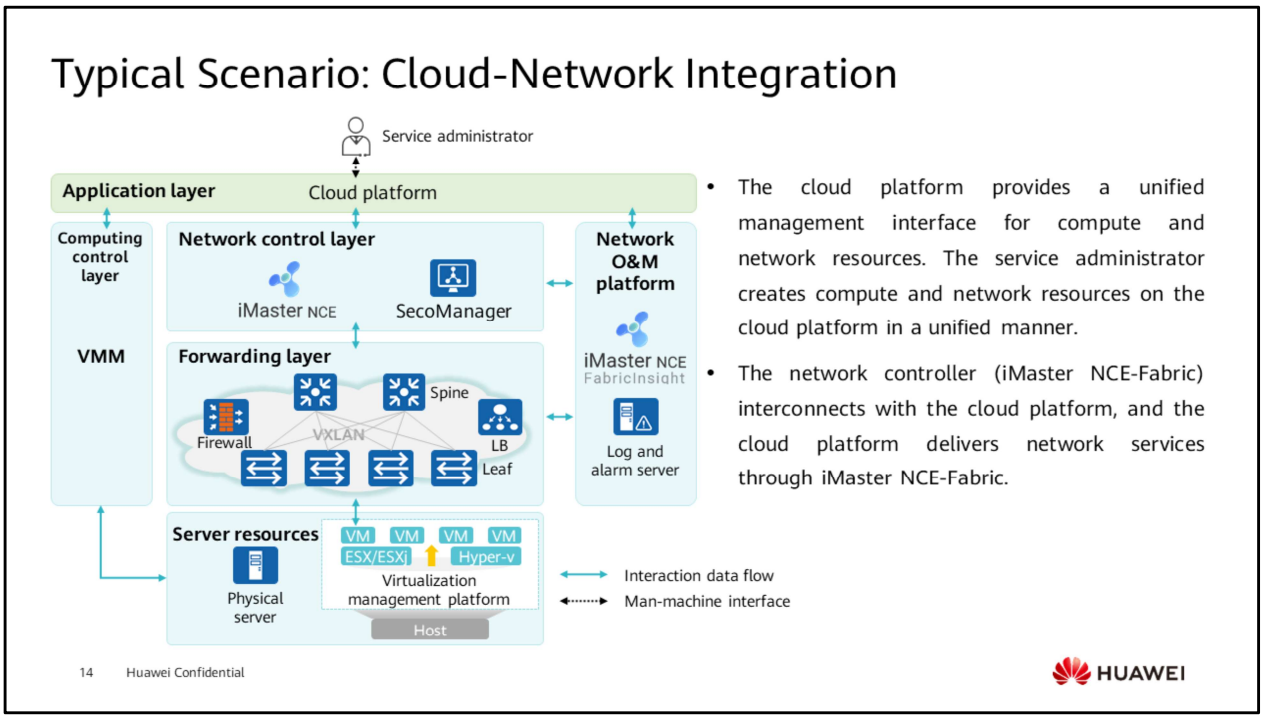
Component	Mandatory or Not	Deployment	Description
iMaster NCE-Fabric	Mandatory	Cluster	The controller cluster consists of *N* nodes. The controller can be deployed on physical servers or VMs.
iMaster NCE-FabricInsight	Optional	Single-node system	In single-node system deployment of iMaster NCE-FabricInsight, the collector and analyzer are combined. Only one server needs to be connected to the leaf node. A maximum of 100 CE switches can be managed.
		Standard cluster	In standard clustering deployment of iMaster NCE-FabricInsight, the analyzer and collector are combined. That is, no independent collector server needs to be deployed.
		Advanced cluster	In advanced clustering deployment of iMaster NCE-FabricInsight, the collector and analyzer are deployed separately. It is recommended that the collector and analyzer be connected to independent leaf nodes to prevent link congestion caused by traffic pressure on the service link.
SecoManager	Optional	Independent deployment	Deploy SecoManager on a server or VM as independent software.
		Integrated into iMaster NCE-Fabric	Integrate SecoManager into iMaster NCE-Fabric as a service.

HUAWEI

- To ensure high availability of network services, iMaster NCE-Fabric supports cluster deployment and small-scale deployment for commercial use.
- In the cluster deployment solution (recommended), a copy of iMaster NCE-Fabric consists of multiple nodes and has the following advantages:
  - Load balancing of services across multiple cluster nodes ensures high reliability and performance.
  - Continued operation of the entire cluster is ensured, even if a single node fails, delivering improved reliability.
  - Flexible expansion enhances the performance of the entire cluster, leading to advanced scalability.
- In the small-scale deployment solution, a copy of iMaster NCE-Fabric consists of only one node. This solution can save hardware resources to the maximum extent, but is only applicable in network overlay networking.

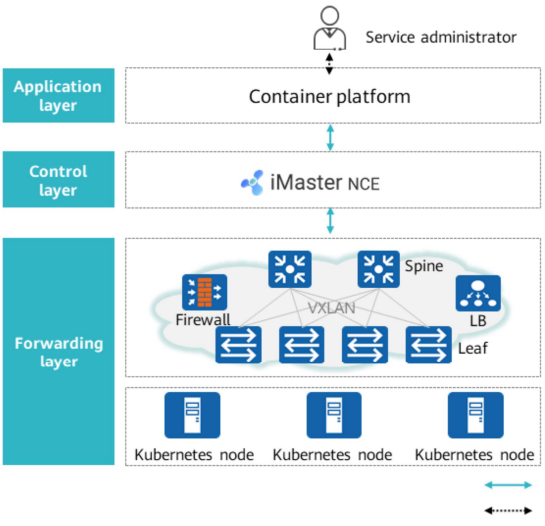


- Characteristics: The DCN connects to the computing virtualization platform and does not connect to the cloud platform. The controller and computing virtualization platform are responsible for service delivery to implement collaborative provisioning of compute and network resources.
- Application scenario: When the computing service management system is complex, or computing management and network management are not integrated, and a unified cloud platform cannot be constructed, the computing scenario can be used.
- Service provisioning:
  - Network service provisioning: The network administrator uses the network controller to allocate network resources to specified services or applications. The network controller automatically delivers service configurations of the overlay network and access configurations of virtual machines (VMs) and physical machines (PMs).
  - Computing service provisioning: The computing administrator creates, deletes, and migrates compute and storage resources on the VMM.



- Background: To solve the problems of low service deployment efficiency, low resource utilization, and complex O&M and management in traditional DCs, DCs need to evolve to the cloud computing architecture.
- Service provisioning:
  - The service administrator allocates network resources to specified services or applications on the cloud platform. The cloud platform delivers the service provisioning instruction to the network controller. The network controller then automatically delivers the configurations to devices without manual configuration.
  - The service administrator creates, deletes, and migrates compute and storage resources on the cloud platform. The cloud platform, network controller, network devices, and servers interact with each other without manual intervention.

Typical Scenario: Container Network



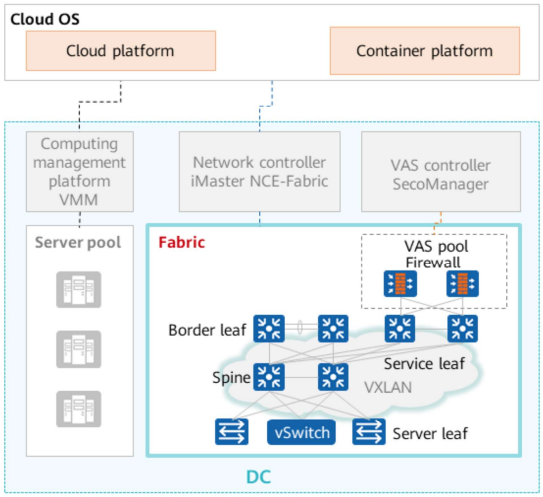
- The container platform provides a unified management interface for compute and network resources. The service administrator creates compute and network resources on the container platform in a unified manner.
- The container platform delivers network services through the network controller (iMaster NCE-Fabric).

Interaction data flow
Man-machine interface

HUAWEI

- In this scenario, the container network is divided into the following layers:
  - Application layer:
    - It is the Kubernetes cluster management entrance. It provides the Kubernetes cluster management function through the northbound RESTful API. It monitors the API server in the southbound direction and delivers Layer 4 to Layer 7 service orchestration configurations to LBs.
  - Control layer:
    - iMaster NCE-Fabric completes network modeling and instantiation, automatically orchestrates and delivers network configurations, and diverts traffic to VAS devices.
  - Forwarding layer:
    - Network resources:
      - CloudEngine switches function as hardware NVE nodes, VXLAN gateways, and distributed routers.
      - Leaf nodes support stacking and M-LAG, and border leaf nodes support stacking and multi-active.
      - BGP is used for communication with external routers.
      - Hardware firewalls are connected to border leaf nodes in inline or bypass mode.
      - LBs are connected to service leaf nodes in bypass mode.
    - Compute resources:
      - Containers are connected to CE1800Vs at Layer 2 or Layer 3.
      - In Layer 3 mode, servers and server leaf nodes run BGP.

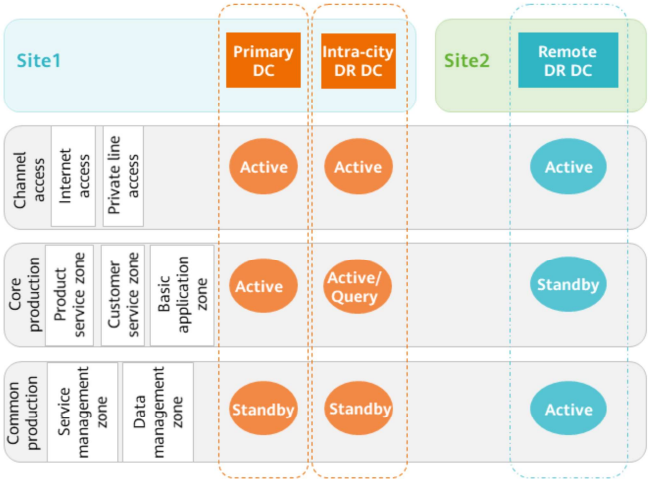
Key Concepts: Single DC and Fabric Network



- Single DC:
  - A single DCN consists of one or more fabric networks, the access zone (Internet, private line, and WAN), and the O&M management zone. Services of different security levels are deployed in different fabric networks to meet high security isolation requirements.
- Fabric network:
  - A fabric network is a switching network that uses the spine-leaf architecture. The network contains spine, server leaf, service leaf, and border leaf nodes, and a VAS resource pool. The spine, border leaf, and service leaf nodes can be combined.
  - The fabric network connects to the cloud platform through iMaster NCE-Fabric to implement automatic network configuration.
  - VXLAN can be deployed in the fabric network to implement network resource pooling.

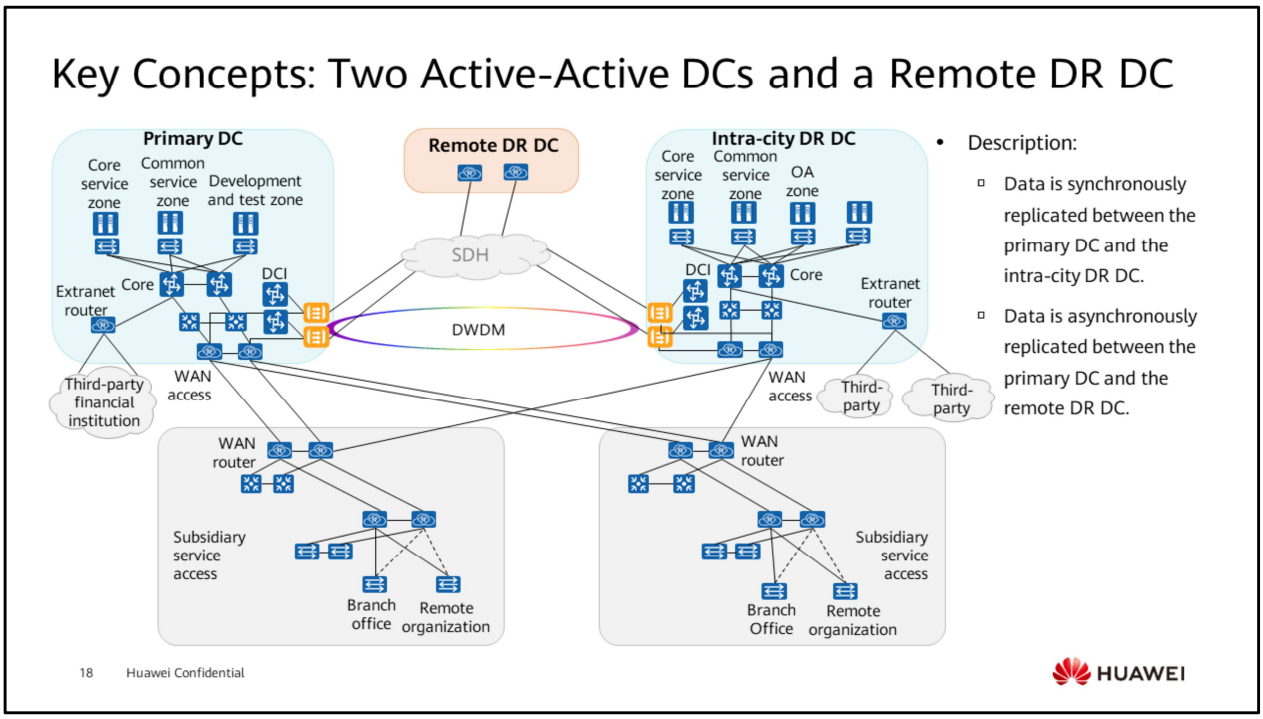
HUAWEI

Key Concepts: Active-Active DCs

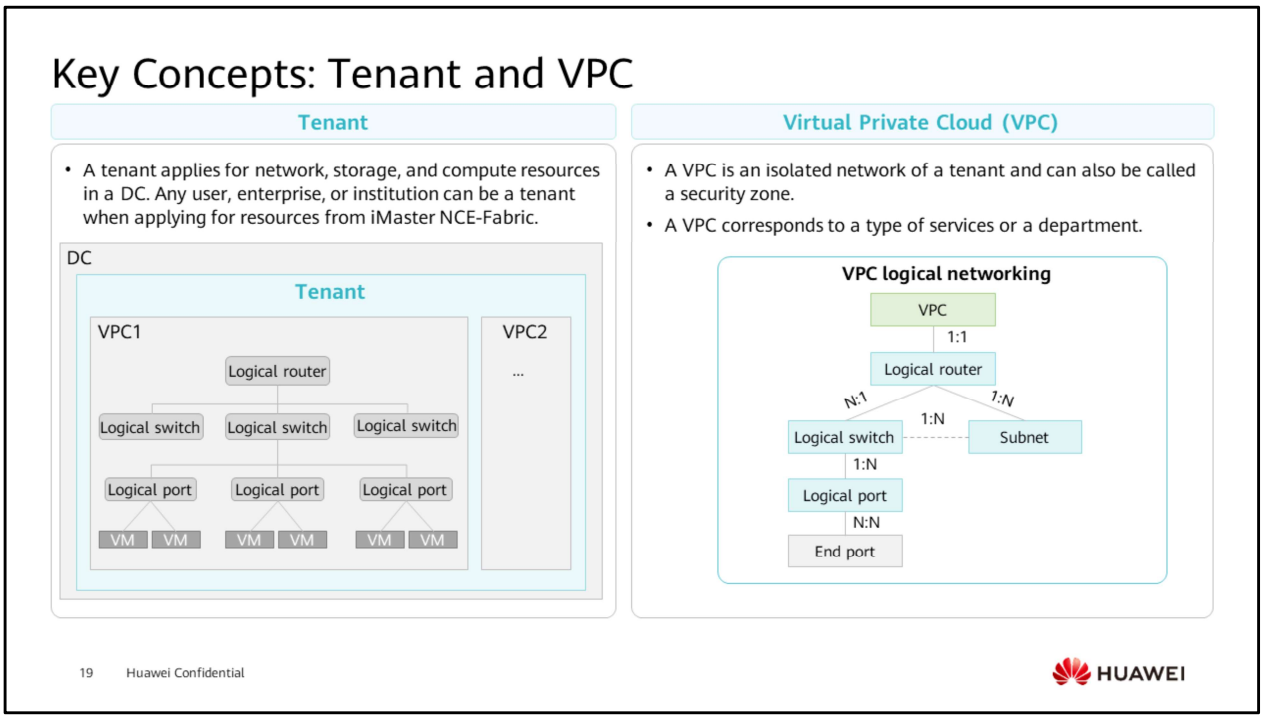


- Description:
  - Active-active DCs indicate that two DCs are running and can carry production services at the same time, improving the overall service capability and system resource utilization of the DCs.
  - Different services use different deployment modes:
    - Channel access services (such as Internet access and private line access): multi-active in multiple regions
    - Core production services (such as product services, customer services, and basic applications): intra-city active-active and remote backup
    - Common production services (such as service management and data management): remote primary DC, implementing interactive DR

HUAWEI



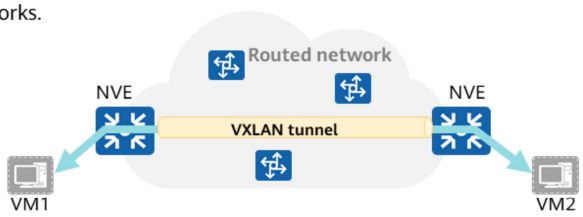
- As shown in the figure:
  - A large enterprise group has a core backbone network. The core backbone network can be a BGP/MPLS IP VPN network or a bare IP network.
  - The DCs in the same city are directly connected through DWDM devices. The core switches and aggregation switches (border leaf nodes on the fabric network) in the DCs are connected to the DWDM devices. In this way, the Layer 2 and Layer 3 networks are directly connected without detouring the WAN.
  - Subsidiaries are connected to two or three DCs through the carrier's private line.



- VPC advantages:
  - Isolated environment: VPCs provide isolated VMs and network environments for different departments.
  - Rich services: Each VPC can provide independent virtual firewall, elastic IP address, security group, super VLAN, IPsec VPN, and NAT gateway services.
- Note:
  - Logical router: A logical router is virtualized on a network device where virtualization software is running, and connects to VMs on different networks, so that VMs can communicate with each other across a Layer 3 network. A network device can be virtualized into multiple logical routers for different tenants to use. That is, multiple tenants share the same network device. A logical router is mapped to a VRF of a network device.
  - Logical switch: A logical switch connects to different VMs and enables them to communicate with each other at Layer 2. A network device can be virtualized into multiple logical switches for different tenants to use. That is, multiple tenants share the same network device.
  - Subnet: On iMaster NCE-Fabric, VXLAN EVPN-based distributed Layer 3 gateways are abstracted as subnets and mapped to VBDIF interfaces and IP addresses of network devices. It is recommended that one logical switch be associated with only one subnet, and one subnet be associated with only one logical switch. (A logical switch can be associated with multiple subnets and exists as a sub-IP address on a physical device.)
  - Logical port: A physical port on a network device can be virtualized into multiple logical ports for different tenants to use. In this way, multiple tenants can share the same physical port to access the network. For each tenant, a logical port is equivalent to an independent and real port.
  - End port: An end port refers to a server-side VM that is connected to the VPC. Each end port corresponds to a VM or a device (such as an LB) that is connected to the fabric network in the similar way as a VM.

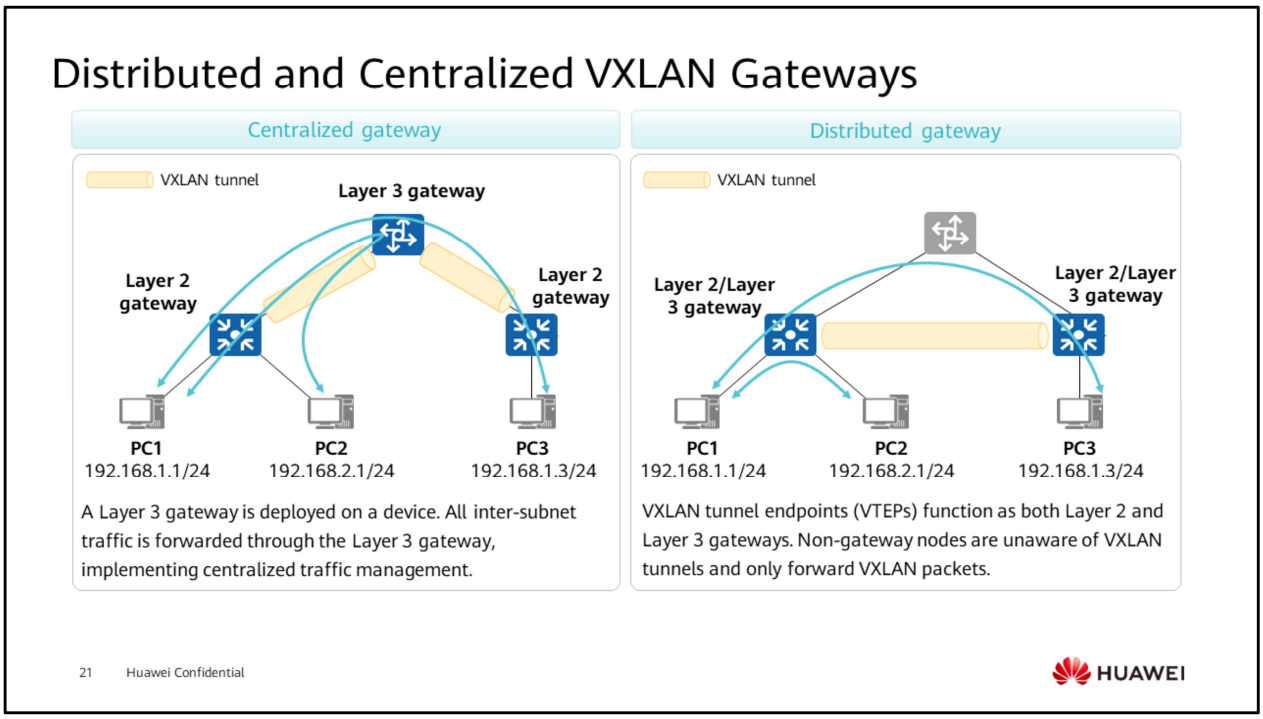
Key Technologies: VXLAN and EVPN

- VXLAN is a tunneling technology that can overlay a Layer 2 virtual network on any network with reachable routes. VXLAN gateways can be used to implement communication within a VXLAN network and communication with traditional non-VXLAN networks.



- In the distributed network overlay solution, the VXLAN control plane uses **BGP EVPN**. The working process of BGP EVPN is as follows:
  - The controller uses NETCONF to deliver BGP EVPN and other configurations to NVE nodes.
  - Devices use BGP EVPN to dynamically establish VXLAN tunnels.
  - The entries (MAC addresses, host routes, and subnet routes) locally learned by devices are globally synchronized through BGP EVPN.

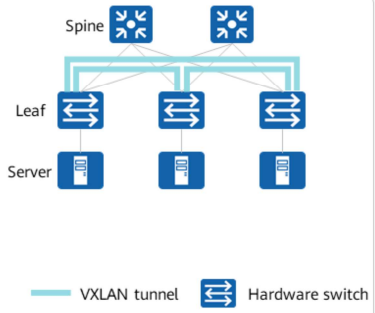
HUAWEI



- Layer 2 and Layer 3 VXLAN gateways:
  - Layer 2 VXLAN gateway: transmits traffic to the VXLAN network and enables intra-subnet communication between terminals on the same VXLAN network.
  - Layer 3 VXLAN gateway: enables inter-subnet communication between terminals on a VXLAN network and allows terminals to access external networks (non-VXLAN networks).

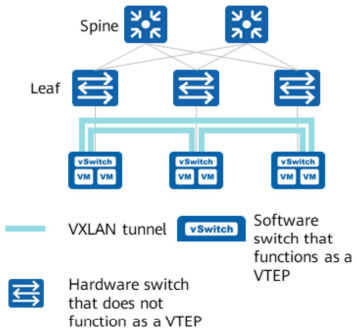
Network Overlay/Host Overlay/Hybrid Overlay

Network overlay



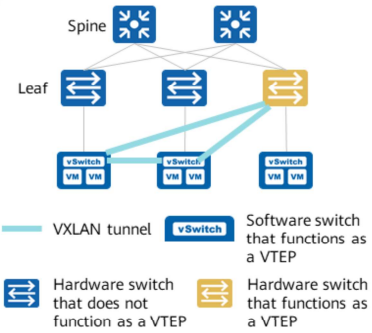
In network overlay, all VTEPs are deployed on hardware switches.

Host overlay



In host overlay, all VTEPs are deployed on software switches (installed on servers).

Hybrid overlay

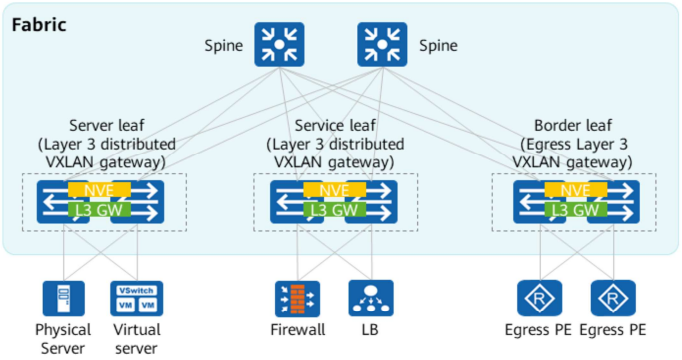


In hybrid overlay, VTEPs are deployed on both hardware and software switches. That is, the ingress and egress of a VXLAN tunnel can be a hardware or software device.

HUAWEI

Distributed Network Overlay

- Distributed network overlay is a type of VXLAN networking. "Distributed" indicates that Layer 3 VXLAN gateways are deployed on all leaf nodes. "Network" indicates that all VXLAN NVE nodes are hardware switches.



- In distributed network overlay networking, NVE nodes are edge nodes of the overlay and responsible for mapping between VLAN IDs and VNIs.
- NVE nodes are interconnected to form a Layer 2 VXLAN network of the overlay. The Layer 2 network is isolated by VNIs of the VXLAN.

HUAWEI

- In the distributed network overlay solution, BGP EVPN is deployed as follows:
  - BGP single-process deployment solution:
    - Deploy an Interior Gateway Protocol (IGP) such as the Open Shortest Path First (OSPF) for the underlay network, and deploy the Internal Border Gateway Protocol (IBGP) for the overlay network.
    - Deploy the External Border Gateway Protocol (EBGP) for both the underlay and overlay networks.
  - When IBGP is deployed, it is recommended that a route reflector (RR) be deployed on a CE switch.
- Common networking of the distributed network overlay solution:
  - Spine nodes are deployed independently and function as RRs to establish BGP EVPN peer relationships with leaf nodes in different groups. BGP EVPN functions as the VXLAN control plane.
  - When border leaf and service leaf nodes are combined, the firewall is connected to the egress Layer 3 gateway in bypass mode to filter traffic and implement secure access control.

Contents

1. CloudFabric Solution and DCN Overview
2. **DCN Planning and Design Methods**
3. Typical DCN Design Cases

HUAWEI

Introduction to the Project Lifecycle

- The project lifecycle consists of project survey (survey status, requirements, objectives, and scope), solution design (including HLD and LLD), delivery, and maintenance.
- In the planning and design phase, focus on project survey and solution design.



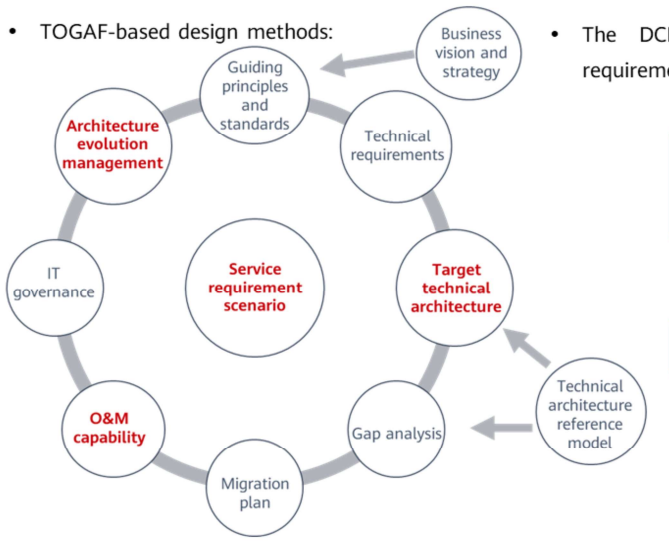
Project survey **Solution design (HLD and LLD)** **Delivery** **Maintenance**

HUAWEI

- High level design (HLD): HLD clarifies the network scope and function deployment solution based on the network architecture. It includes the interaction process of important services within the scope of the contract, function allocation of all services on each NE, peripheral interconnection relationships of NEs, interconnection requirements and rules, and quantitative calculation of interconnection resources.
- Low level design (LLD): LLD refers to the activities of further detailing the design content, providing data configuration rules, and guiding data planning based on the network topology, signaling channel routing, and service implementation solution developed in the network planning phase.
- Note: This course focuses on HLD.

Design Methods

- TOGAF-based design methods:



- The DCN design method is based on TOGAF, is requirement-centric, and focuses on architecture and O&M.



HUAWEI

- The Open Group Architecture Framework (TOGAF) is an industry standard architecture framework.

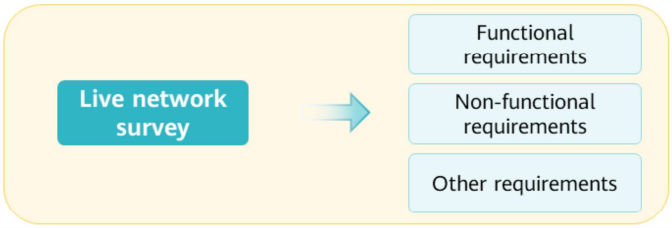
Design Principles

High reliability	▪ Perform high reliability design from multiple dimensions such as the network, controller, and device to meet requirements of network running around the clock. ▪ Dual links between devices are highly reliable, and inter-card links of modular devices are bundled. ▪ Consider the scale of the network fault domain to avoid architecture risks caused by an excessively large fault domain, control the scope of unknown unicast, broadcast, and multicast, and prevent loops.
High security	▪ Multi-layer and multi-dimensional security control, two-layer heterogeneous firewalls, Anti-DDoS and IPS deployed in the access zone, and IDS deployed in the resource pool zone. ▪ Deploy firewalls based on physical partitions, and conform to rules of on-demand deployment and nearby protection to ensure comprehensive protection. ▪ Implement network-wide collaboration automation and firewall resource pooling through firewall policy deployment.
Maintainability Easy management	▪ Use standardized architectures and deployment solutions to gradually improve automation capabilities and reduce the workload of service deployment. ▪ Construct an integrated network O&M system based on standard northbound interfaces.
High performance	▪ Meet the requirements of service applications for connection at any time, real-timeness, interaction, and intelligence. Provide high bandwidth, low latency, large buffer, 10GE/25GE access, and 40GE/100GE interconnection. ▪ Meet requirements for high network scalability by large-scale server clusters, distributed storage, big data services, and other services. ▪ Deploy firewall clusters to meet requirements for large throughput and high concurrency and support service development.
Self-service Foresight	▪ Support network SDN and cloud computing platforms to implement IaaS, PaaS, and SaaS automation. ▪ Based on IaaS automation, implement self-service, flexible resource pool expansion, and agile delivery.

HUAWEI

Requirement Analysis > Design Objectives > Solution Design

Requirement Analysis



Live Network Survey	Description
Platforms	Currently used service platforms: virtualization, cloud computing, container, big data, and AI
Service scale	Current network scale, compute cluster scale, and storage scale
Device deployment	• Network architecture and device selection (fixed and modular devices) • Server type: rack server, blade server, high-density server, and HPC, as well as NIC quantity and type • Storage: block storage device, file storage, and object storage, such as FC SAN, NAS, and distributed storage
Service evaluation	• Estimated service scale: number of servers and storage devices • Service network requirements: oversubscription ratio, latency, and bandwidth

HUAWEI

- Note:
  - FC SAN: is a type of storage area network (SAN) that uses the Fibre Channel (FC) protocol to connect to disk arrays.
  - NAS: network attached storage
  - Oversubscription ratio = Total bandwidth of southbound interfaces/Total bandwidth of northbound interfaces

Requirement Analysis: Functional Requirements

Functional Requirements	Description
Network functions and performance	Whether specified features need to be supported: • Solution requirements, for example, whether SDN is supported • Device requirements, such as VXLAN, EVPN, load balancing, port density, port rate, NIC model (GE, 10GE, 25GE, electrical port, and optical port), NIC mode (active/standby and load balancing), and network oversubscription ratio (1:1 and 3:1)
Security	Whether the following security requirements are specified: physical isolation, logical isolation, and firewall and its function and performance requirements
Storage	Whether IP SAN, NAS, and distributed storage services are used
Platform functional requirements	Whether the cloud platform is associated with the network, whether the VMM platform is associated with the network, and whether there are RoCE network and other requirements

HUAWEI

- Note:
  - IP SAN: is a type of SAN that uses the IP protocol, that is, a technology that uses IP channels to connect servers and storage devices.

Requirement Analysis: Non-Functional Requirements

Non-Functional Requirements	Description
High reliability	• Whether remote disaster recovery and backup are required • Whether local high reliability and backup are required • High availability and fault tolerance of platforms, controller, devices, and links
High scalability	• Single-fabric scalability, multi-fabric scalability, and cross-DC scalability • Fabric interconnection, intra-city DC interconnection, and remote DC interconnection
Maintainability	• O&M tool requirements and fault detection, location, and recovery requirements • Northbound interconnection with the unified O&M platform

HUAWEI

Requirement Analysis: Other Requirements

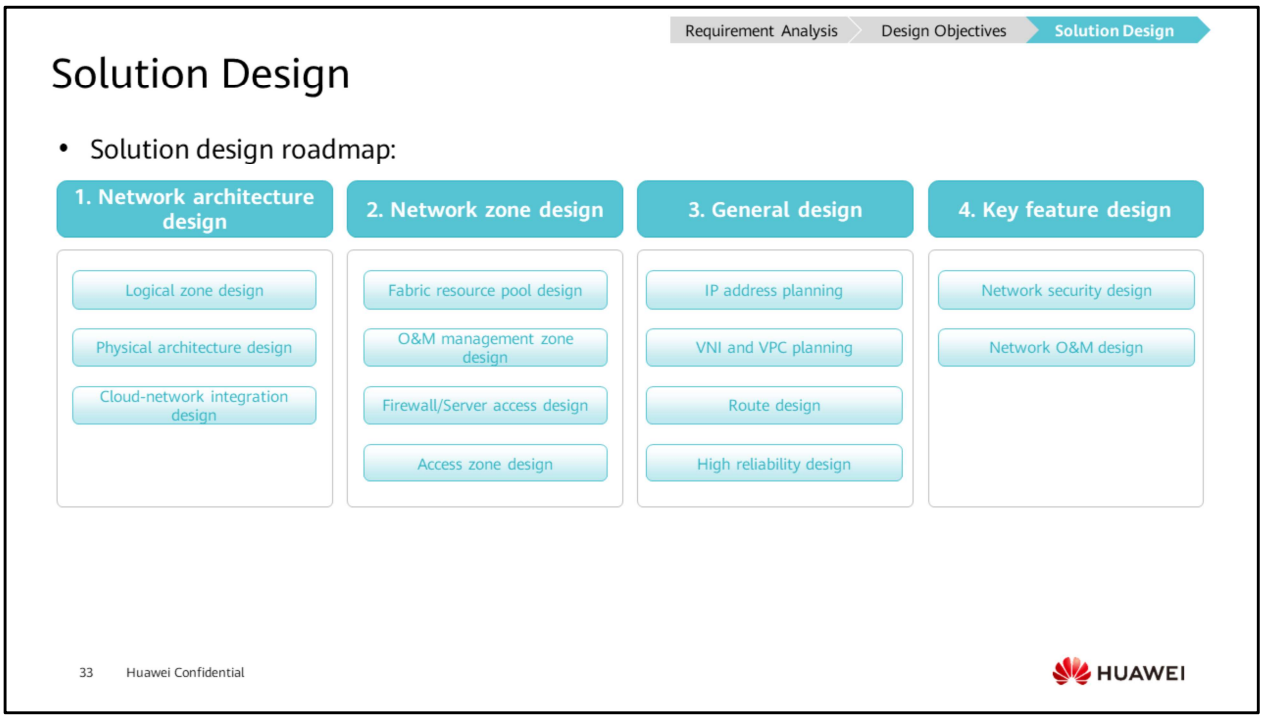
Other Requirements	Description
Physical requirements	Equipment room size, power supply density (medium-density, low-density, or high-density), POD scale, rack capacity, device selection, and layout
Cabling	TOR, EOR, MOR, single-mode, multimode, LC or MPO
Other special requirements	...

HUAWEI

Design Objectives: Solution Selection

- Select the DC architecture based on the construction objectives:
  - Single DC, active-active DCs, two active-active DCs and a remote DR DC, and multiple DCs in multiple cities
- Select the following solutions based on requirement analysis:
  - Underlay solution
  - Solution for interconnection between the underlay network and the cloud platform
  - Network Virtualization - Computing solution
  - Cloud-network integration solution
  - Container network solution
  - iLossless solution

HUAWEI



- Note: This course uses this solution design roadmap as an example to explain the DCN design process in financial scenarios. The design, deployment parameters, and device quantity involved in this course are all examples. You can consider the specific deployment based on the actual services on the live network.

Contents

1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

HUAWEI

Analysis of Network Functional Requirements

Functional Requirements	Description
Big data requirements	• A large buffer is required because burst traffic exists. • The network is required to provide a good oversubscription ratio. Typically, the oversubscription ratio of servers to the access layer is 3:1. • The network is required to provide sufficient access bandwidth (10GE access and 40GE interconnection), and the network latency is required to be less than 1 ms.
Network service deployment requirements	• The controller interconnects with mainstream cloud platforms in the industry through standard APIs. This implements unified network service provisioning and orchestration on the cloud platform and achieves end-to-end automatic provisioning of network services. • The controller can automatically discover network devices and add them to the controller system, and implement automatic deployment of the underlay network. • A fabric network can be logically partitioned, the default isolation policy can be used, and different isolation domains in the fabric network cannot communicate with each other. • When the VM status changes, network service resources are adjusted on demand based on the scheduling of the cloud platform. • The controller provides the GUI for visualized network service provisioning and orchestration. Network services should be orchestrated in terms of overlay, tenant, and application. • The controller can control the interconnection between the fabric network and external networks through the GUI. • When a fabric network is logically partitioned, the controller supports association with L4-L7 devices such as firewalls, LBs, SSL uninstallation devices, and encryptors to securely manage and control east-west traffic of tenants and between tenants. • A fabric network can be divided into multiple logical networks to carry multiple homogeneous or heterogeneous cloud platforms.
Basic network function requirements	• EVPN-based VXLAN is supported. • Distributed VXLAN gateways are supported.

HUAWEI

Analysis of Network Non-Functional Requirements

Non-Functional Requirements	Description
Security requirements	• Different security protection application systems need to communicate with each other through firewalls. • The production network and non-production network are isolated through different fabric networks. The intranet and DMZ are isolated through different fabric networks. • The controller supports hierarchical rights-based user management. • The controller, spine nodes, and leaf nodes support AAA.
High reliability requirements	• The high reliability of the network system is required to prevent Layer 2 loops (such as broadcast storms and unicast flooding). • Broadcast storm suppression is supported. • The controller cluster supports load balancing. When a controller is faulty, the fault can be automatically detected and the upper-layer service system is unaware of the fault. • When all SDN controllers are faulty, traffic forwarding on spine and leaf nodes is not affected.
High scalability requirements	• The horizontal scalability of the network is improved and the reuse efficiency of compute and storage resources in physical and virtual environments is maximized. • The SDN controller can manage at least 200 leaf switches. • The SDN controller supports access and management of at least 50,000 VMs.
Maintainability requirements	• The SDN controller automatically deploys software and policies on all switches without manual operations. • The controller can record operations of administrators. • Internal traffic of a fabric network can be forwarded to physical ports of each device. • Various fault locating and troubleshooting functions are provided to quickly locate faulty devices, components, ports, and service communication paths in a fabric network.

HUAWEI

DCN Overall Design Objectives

VXLAN resource pool

- Build EVPN-based VXLAN resource pools so that servers can be flexibly deployed.

Network-wide SDN

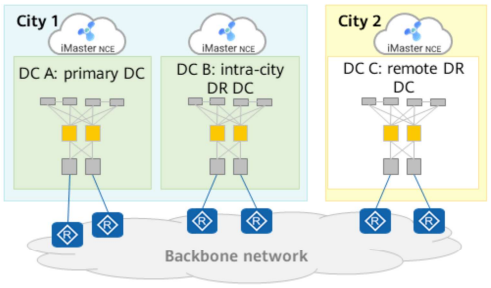
- Deploy SDN on the entire network to implement network infrastructure automation and fast service rollout.

Cloud-network integration

- Implement unified provisioning of computing, storage, and network resources based on the cloud platform.

HUAWEI

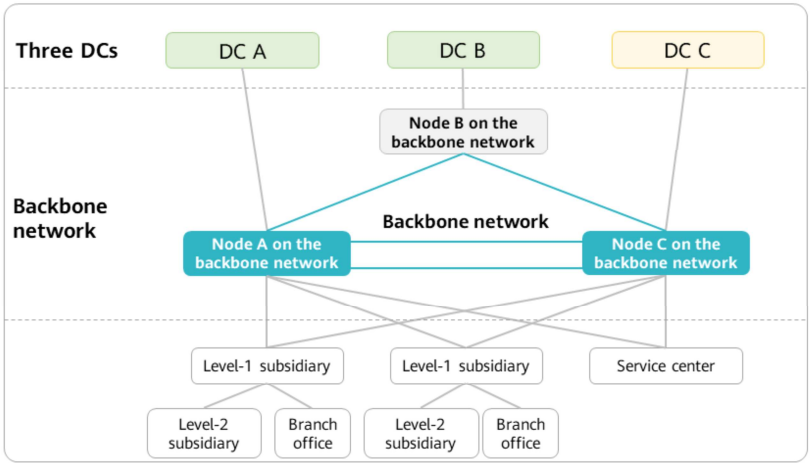
Solution Selection: Cloud-Network Integration Solution (Two Active-Active DCs and a Remote DR DC)



- **DC A**: is the primary production DC.
  - It has the highest availability.
  - Most service systems (transactions) mainly access the production DC. The information system of the production DC uses load balancing and two-node cluster in hot standby mode, and defines a complete emergency switchover scheme, meeting the local HA design and running requirements. The organization architecture and process of the production operation are improved, and the DC is responsible for operation monitoring and routine management of multiple DCs in multiple cities.
- **DC B**: is the intra-city DR DC.
  - In the case of a disaster, it can fast take over services of the primary DC.
  - It also functions as the emergency site for some centralized operation services. When a regional disaster occurs, it supports emergency switchover of DC operation services.
- **DC C**: is the remote DR DC.
  - It provides comprehensive coverage for important service systems.

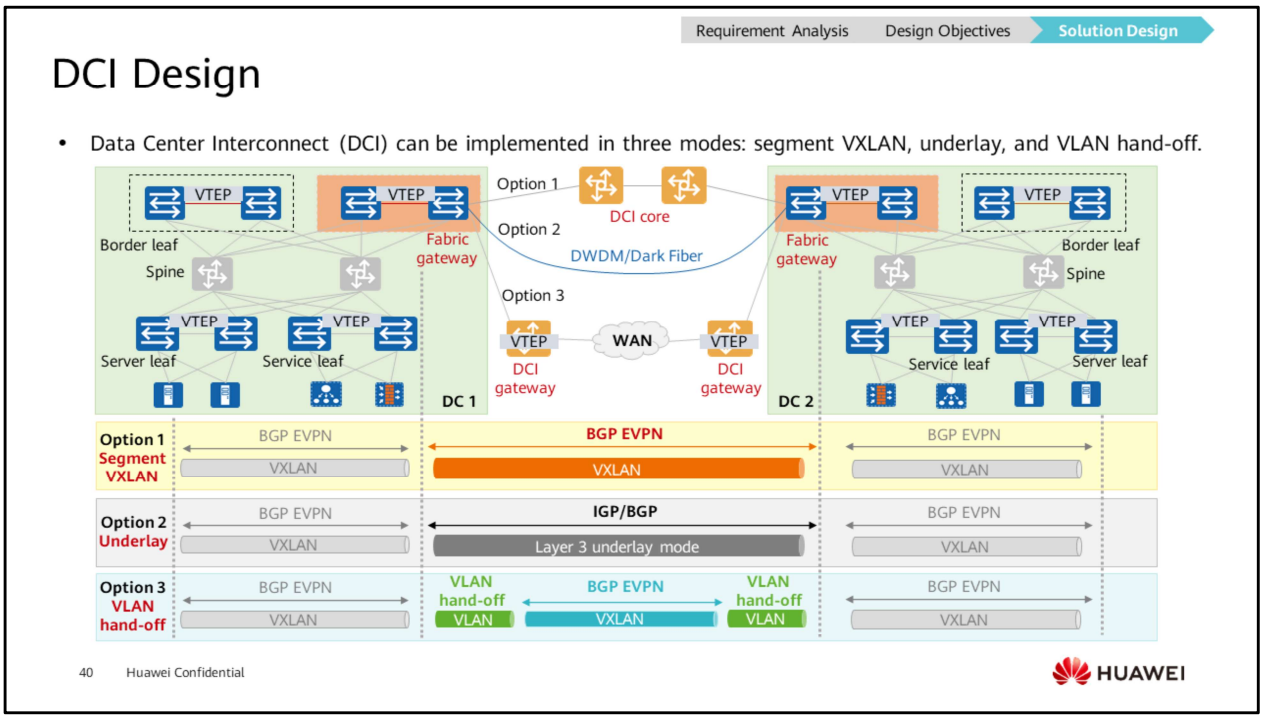
HUAWEI

Logical Network Architecture of Two Active-Active DCs and a Remote DR DC



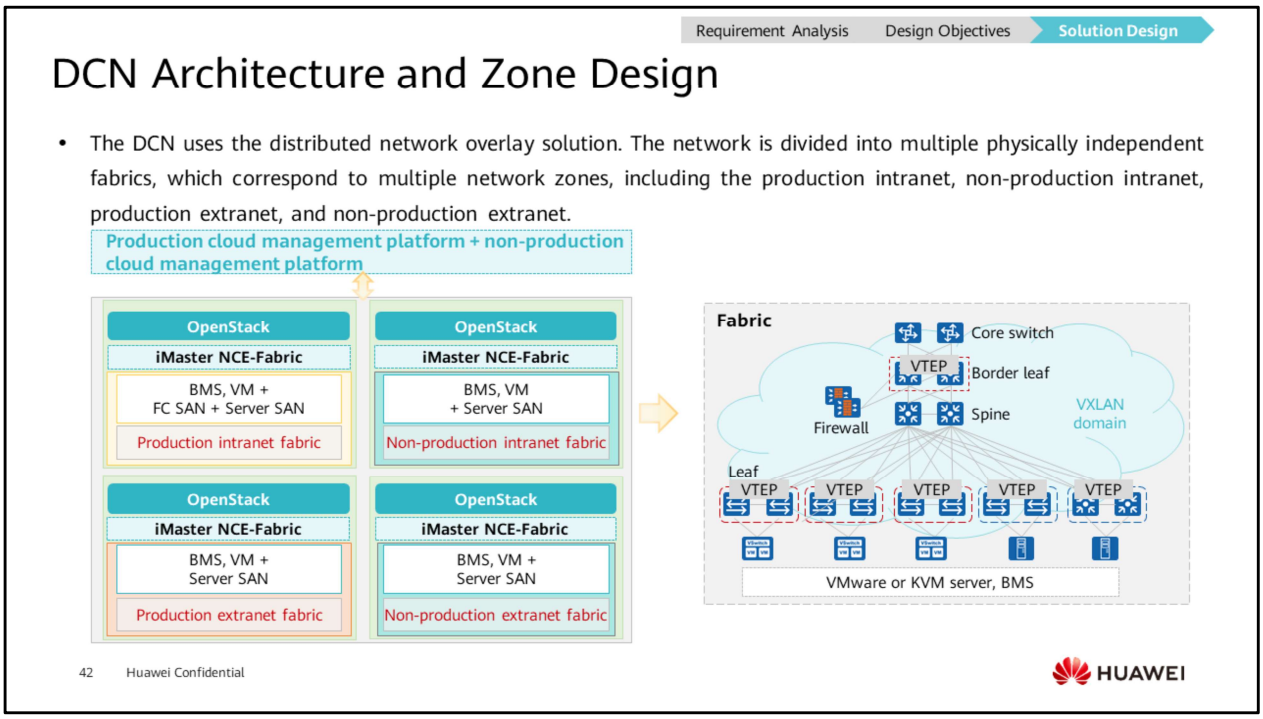
- Overall network architecture:
  - Two Active-Active DCs and a Remote DR DC + backbone network
  - Data of multiple nodes is transmitted to the backbone network for service convergence.
  - Level-1 subsidiaries, branches, and service centers are connected to the backbone network, and level-2 subsidiaries and branches are connected to level-1 subsidiaries.

HUAWEI



- Option 1 segment VXLAN mode:
  - This mode is applicable to the scenario where many sites are connected through DCI core switches.
  - BGP EVPN is configured on fabric gateways (border leaf nodes) to create VXLAN tunnels. The VXLAN packets received from a DC are decapsulated, re-encapsulated, and then tunneled to the peer DC. This process implements end-to-end transmission for cross-DC VXLAN packets. Therefore, only reachable underlay routes between fabric gateways are required on the DCI physical network.
  - The segment VXLAN mode supports Layer 2 and Layer 3 interconnection across DCs.
- Option 2 underlay mode:
  - This mode is applicable to the scenario where a small number of sites are deployed and fabric gateways of two DCs are directly connected through bare optical fibers or DWDM devices.

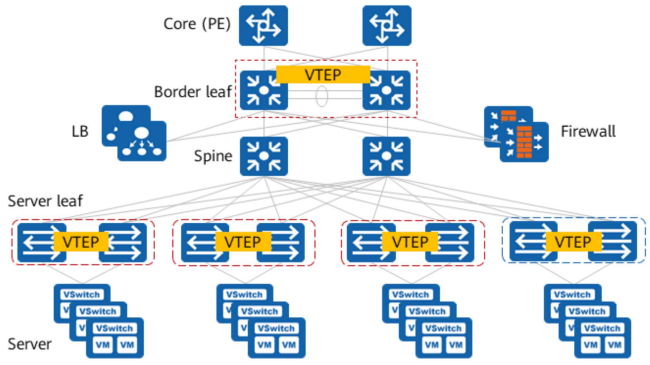
    - An IGP or BGP is deployed between fabric gateways that are configured with VRFs. The fabric gateways function as VTEPs in a fabric, and Layer 3 interconnection between fabrics is implemented through an IGP or BGP. VXLAN packets are decapsulated on the fabric gateways and forwarded based on the IGP or BGP routes on the underlay network. Therefore, a DCI physical network needs to transmit private network routes of services.
- Option 3 VLAN hand-off mode:
    - This mode is applicable to the scenario where many sites are connected through the WAN.
    - A VXLAN-capable device is added to each DC as a DCI gateway. BGP EVPN is deployed between DCI gateways in each DC to establish VXLAN tunnels. BGP EVPN is deployed in a fabric. A fabric gateway functions as a VTEP in the fabric and is connected to a DCI gateway through a common VLAN, and VLAN is configured on the fabric gateway and DCI gateway for access to a VXLAN network. VXLAN packets across DCs are decapsulated on a fabric gateway and converted into common Ethernet packets. After receiving the packets, a DCI gateway encapsulates the packets into VXLAN packets and sends them to the peer DC through a VXLAN tunnel. Therefore, only reachable underlay routes between fabric gateways are required on the DCI physical network.
    - The VLAN hand-off mode supports Layer 2 and Layer 3 interconnection across DCs.



- Network architecture description:
  - A micro-module is built for two rows of racks. Each micro-module corresponds to one fabric. Alternatively, a fabric is built based on the equipment room.
  - Spine-Leaf architecture, VXLAN, and SDN architecture are deployed in the cloud DC fabric resource pool. Border leaf nodes and spine nodes can be combined.
  - iMaster NCE-Fabric interconnects with OpenStack (FusionSphere) to implement network configuration automation. iMaster NCE-Fabric can manage firewalls.
- Note:
  - BMS is short for Bare Metal Server.
  - SAN is short for Storage Area Network.
  - Server SAN is a storage resource pool consisting of multiple independent servers.

Requirement Analysis | Design Objectives | Solution Design

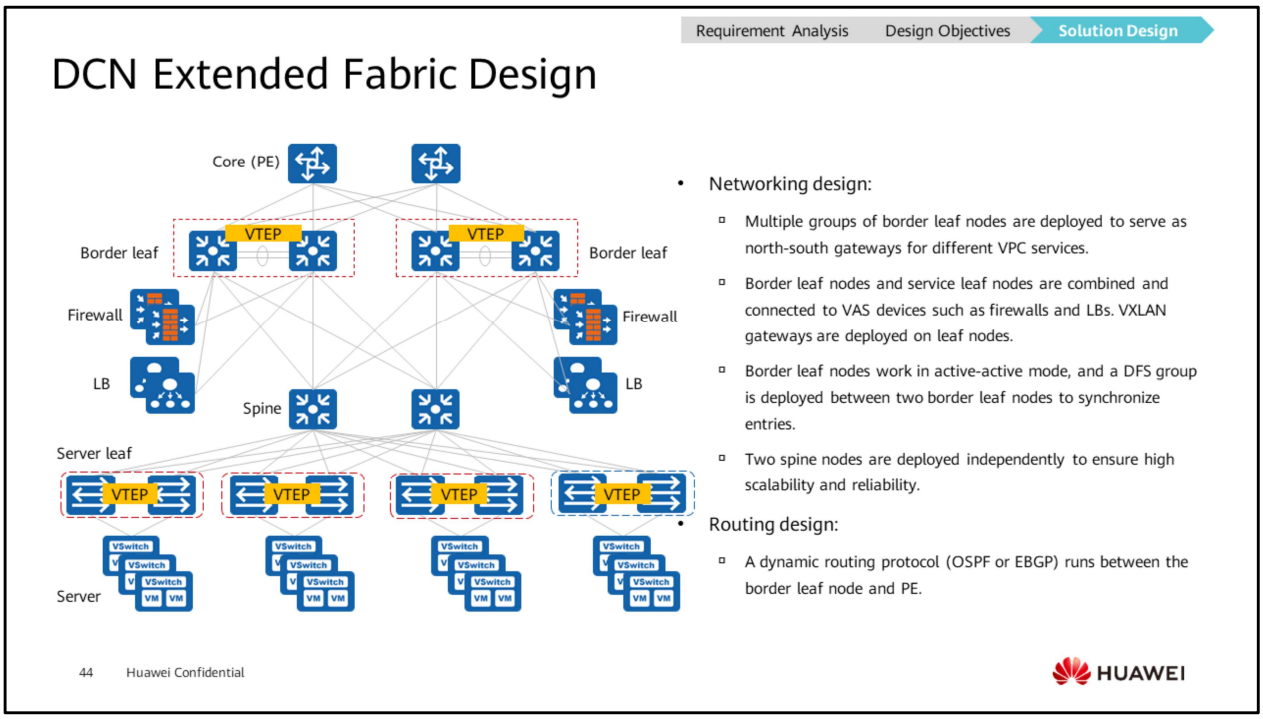
DCN Typical Fabric Design



- Networking design:
  - Border leaf nodes and service leaf nodes are combined and connected to VAS devices such as firewalls and LBs. VXLAN gateways are deployed on leaf nodes.
  - Border leaf nodes work in active-active mode, and a DFS group is deployed between two border leaf nodes to synchronize entries.
  - Two spine nodes are deployed independently to ensure high scalability and reliability.
  - The oversubscription ratio of server leaf nodes is 3:1. Each spine node uses two 40GE uplinks. Each border leaf node provides two 10GE interfaces to connect to two core switches, respectively.
- Routing design:
  - A dynamic routing protocol (OSPF or EBGP) runs between the border leaf node and core node (or PE device).

HUAWEI

- Application scenario:
  - The network scale is small and the number of servers is fewer than 1,500, for example, big data zone in an equipment room.
- DFS group is short for Dynamic Fabric Service Group. A DFS group is used for pairing between M-LAG devices. M-LAG devices use the DFS group protocol to synchronize information such as the interface status and entries.

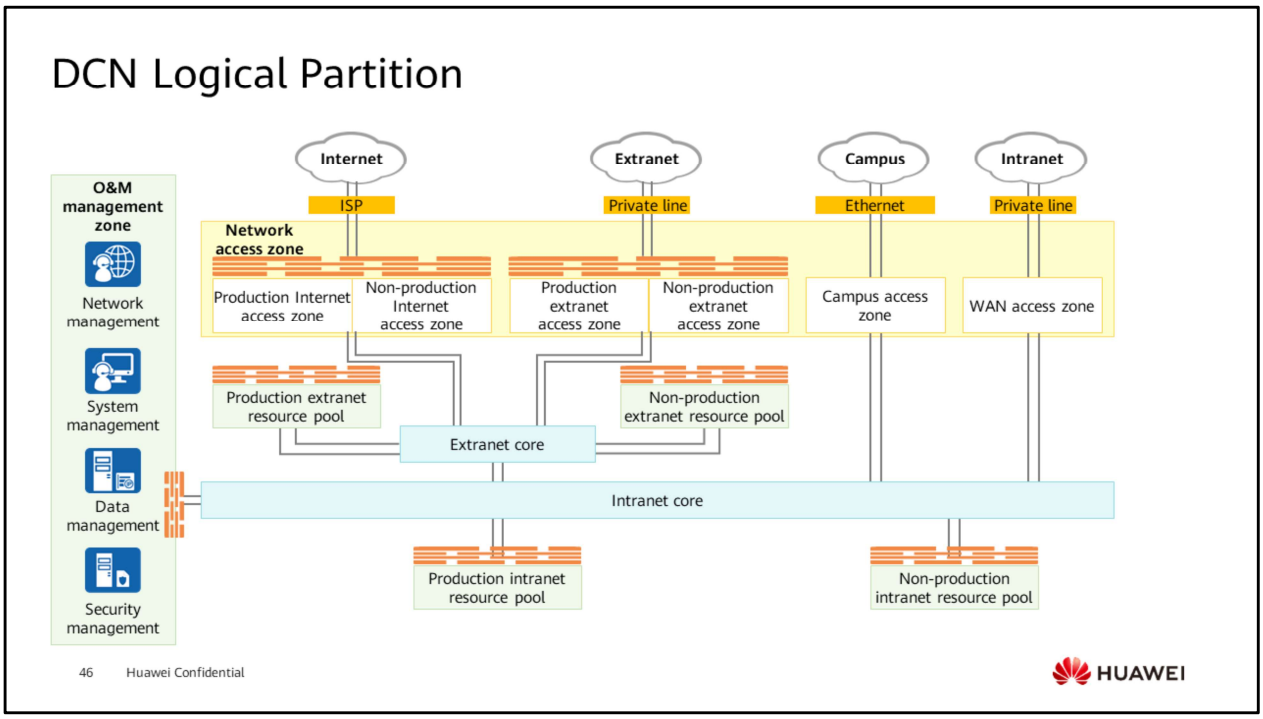


- Characteristics:
  - Firewalls and LBs are connected to border leaf nodes to control secure access of north-south traffic. This facilitates fault locating and is consistent with existing O&M habits.
  - Border leaf nodes and service leaf nodes are combined, reducing device investment.
  - Multiple groups of border leaf nodes are deployed, meeting requirements of high availability, high security, and high performance.
- Application scenario:
  - Different types of services are deployed in a fabric, and different border leaf nodes must be used as egresses.
  - Cases:
    - The DMZ of a bank's DC contains e-banking DMZ border leaf nodes, extended DMZ border leaf nodes, and third-party DMZ border leaf nodes.
    - The VM resource pool of a bank has intranet border leaf nodes and DMZ border leaf nodes.

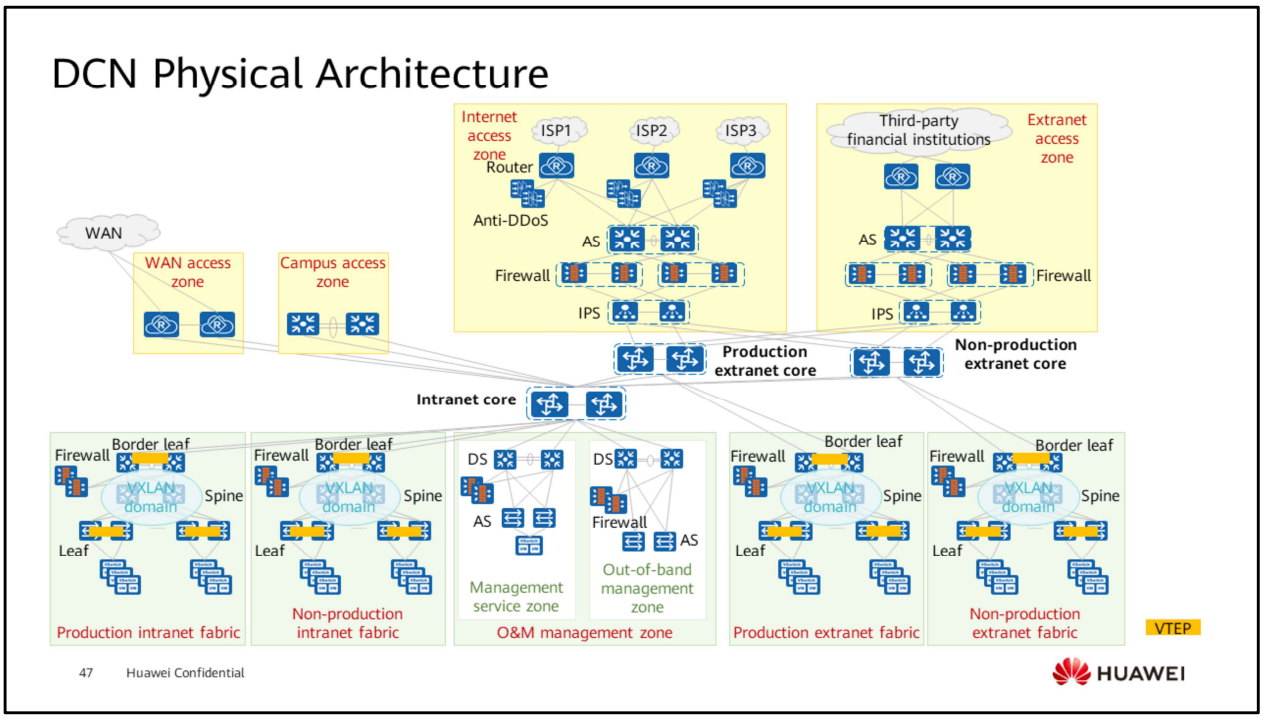
Contents

1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

HUAWEI

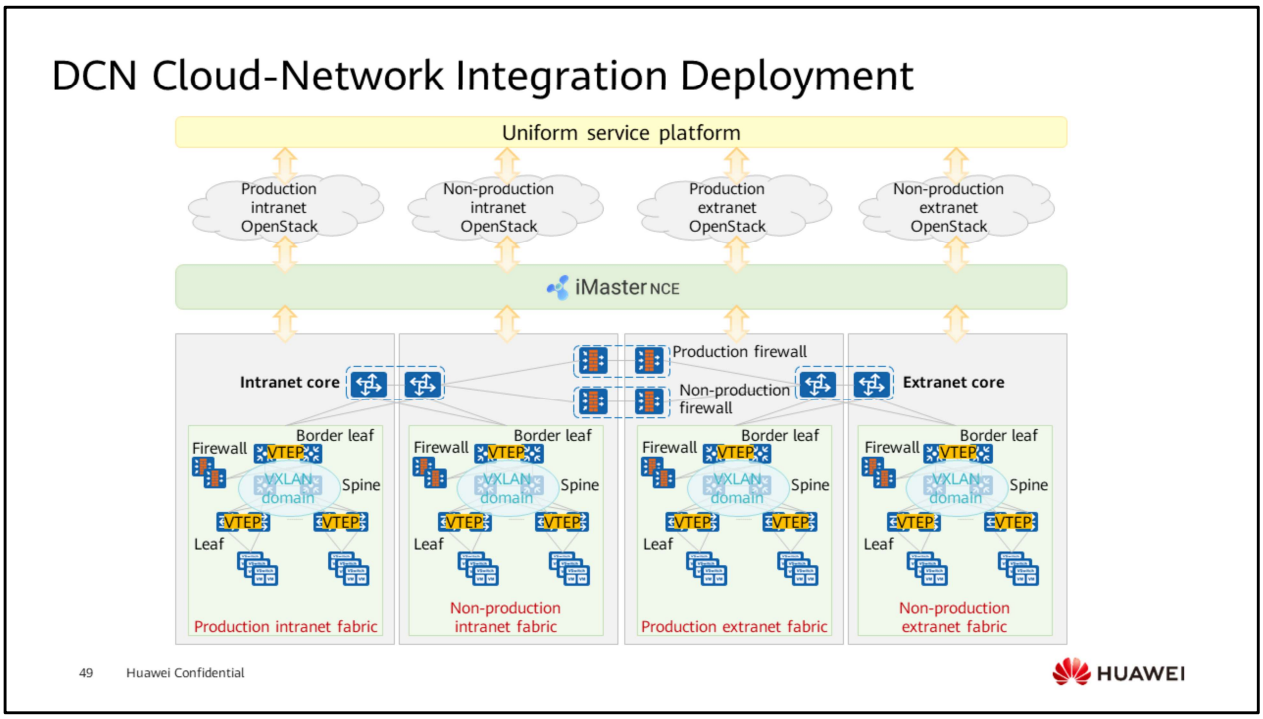


- Partition description:
  - The DCN consists of the production Internet access zone, non-production Internet access zone, production extranet access zone, non-production extranet access zone, campus access zone, WAN access zone, O&M management zone, production extranet zone, non-production extranet zone, production intranet zone, and non-production intranet zone, and core switching zone.
  - The out-of-band management network is deployed in each zone.
  - Firewalls are connected to the border of each zone in bypass mode for isolation.
- Note:
  - Out-of-band management: iMaster NCE-Fabric connects to the out-of-band management network ports on network devices through an independently deployed out-of-band management switch, and manages and controls the network devices through an independent out-of-band network.
  - In-band management: No independent management switch and network are configured. iMaster NCE-Fabric directly connects to the service network through a service switch, and manages and controls network devices through the underlay layer of the service network.



- In the DCN, VXLAN is deployed on the production intranet, non-production intranet, production extranet, and non-production extranet to build a fabric resource pool.
- Switch positioning:
  - In the resource pool zone, border leaf nodes and service leaf nodes are combined.
  - Extranet core switches are connected to the Internet access zone, extranet access zones, extranet resource pool zone, and DC core switches to control the advertisement of intranet routes to external networks.
- Firewall positioning:
  - Firewalls are deployed at the border of the resource pool zone and are connected to the border leaf switches in bypass mode to perform access control on all traffic entering and leaving the zone. The cloud platform drives the SDN controller to automatically deliver the traffic diversion policy of the border firewalls of the resource pool zone.
  - Firewalls in the Internet access zone and extranet access zone meet the two-layer heterogeneous deployment requirements and perform access control on all traffic entering and leaving the access zones.
  - All incoming and outgoing traffic of the O&M management zone passes through the border firewalls.

    - The WAN access zone and campus access zone are trusted zones, where no firewall needs to be deployed.
    - Traffic between the intranet and external network passes through the border firewalls twice.
- Note:
    - AS: is an access switch at the leaf layer.
    - DS: is a core switch at the spine layer.
    - IPS is short for Intrusion Prevention System.
    - Anti-DDoS is short for anti-distributed denial of service.



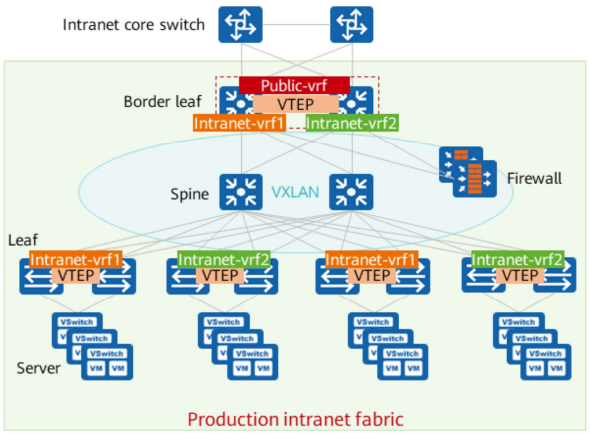
- Solution:
  - Four fabrics and four VXLAN domains correspond to the production intranet, non-production intranet, production extranet, and non-production extranet, respectively.
  - Based on service attributes, one copy of iMaster NCE-Fabric is deployed to interconnect with four sets of OpenStack cloud platforms to process services on the production intranet, non-production intranet, production extranet, and non-production extranet, respectively.
  - As the service entrance, the unified service platform collaborates with four sets of OpenStack cloud platforms' computing, storage, and network components to implement agile service provisioning and resource pool sharing.
  - The cloud platform and controller are deployed in in-band mode.
- Note: In the preceding figure, border leaf nodes also function as service leaf nodes.

Contents

1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

HUAWEI

Production Intranet Fabric Network Design

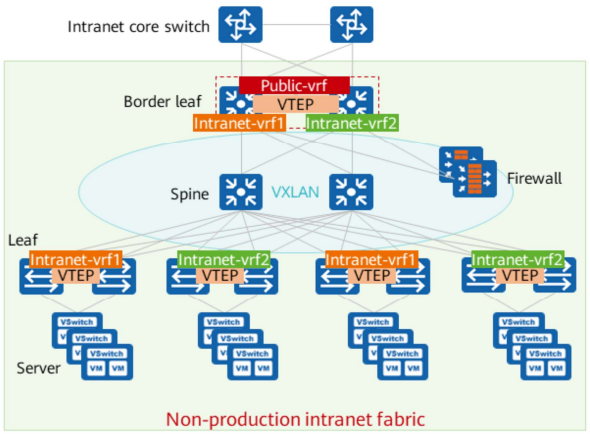


- Networking design:
  - The production intranet fabric network connects to intranet core switches through an independent physical network. The distributed network overlay is used, and gateways are deployed on leaf switches.
  - Border leaf nodes and spine nodes are deployed independently, while border leaf nodes and service leaf nodes are combined to connect to firewalls and IDSs.
  - M-LAG is deployed between border leaf nodes, and each border leaf node provides two 40GE ports to connect to two core switches.
  - The oversubscription ratio of leaf nodes is 3:1. Two leaf nodes function as TOR switches, and two 40GE links are configured as a peer-link to connect them. Each leaf node uses two 40GE links to connect to two spine nodes.
  - Firewalls perform access control on all incoming and outgoing traffic of the fabric. Each firewall is connected to two border leaf nodes through two 10GE links.
  - The fabric uses VPCs to isolate services. Intra-VPC service access traffic does not pass through firewalls, and inter-VPC service access traffic passes through firewalls by default.
  - iMaster NCE-Fabric automatically creates VRFs on leaf nodes to isolate different VPC services. Border leaf nodes connect to intranet core switches through Public-vrf.

HUAWEI

- Routing design:
  - The underlay network uses OSPF, the overlay network uses IBGP EVPN, and spine nodes function as RRs.
- Note:
  - IDS is short for Intrusion Detection System.

Non-Production Intranet Fabric Network Design

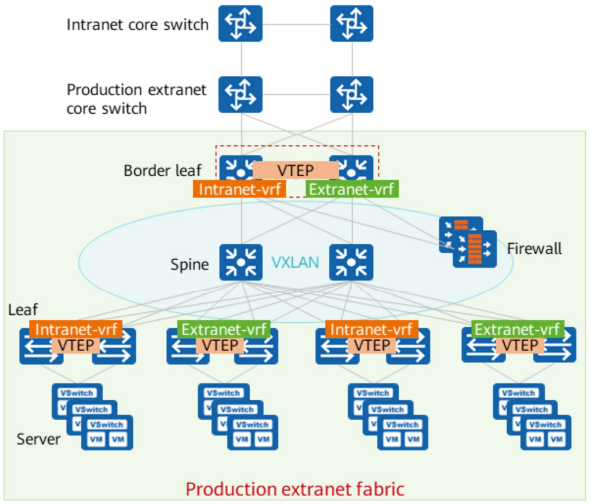


- Networking design:
  - The non-production intranet fabric network connects to intranet core switches through an independent physical network. The distributed network overlay is used, and gateways are deployed on leaf switches.
  - Border leaf nodes and spine nodes are deployed independently, while border leaf nodes and service leaf nodes are combined to connect to firewalls and IDSs.
  - M-LAG is deployed between border leaf nodes, and each border leaf node provides two 40GE ports to connect to two core switches.
  - The oversubscription ratio of leaf nodes is 3:1. Two leaf nodes function as TOR switches, and two 40GE links are configured as a peer-link to connect them. Each leaf node uses two 40GE links to connect to two spine nodes.
  - Firewalls perform access control on all incoming and outgoing traffic of the fabric. Each firewall is connected to two border leaf nodes through two 10GE links.
  - The fabric uses VPCs to isolate services. Intra-VPC service access traffic does not pass through firewalls, and inter-VPC service access traffic passes through firewalls by default.
  - iMaster NCE-Fabric automatically creates VRFs on leaf nodes to isolate different VPC services. Border leaf nodes connect to intranet core switches through Public-vrf.

HUAWEI

- Routing design:
  - The underlay network uses OSPF, the overlay network uses IBGP EVPN, and spine nodes function as RRs.

Production Extranet Fabric Network Design

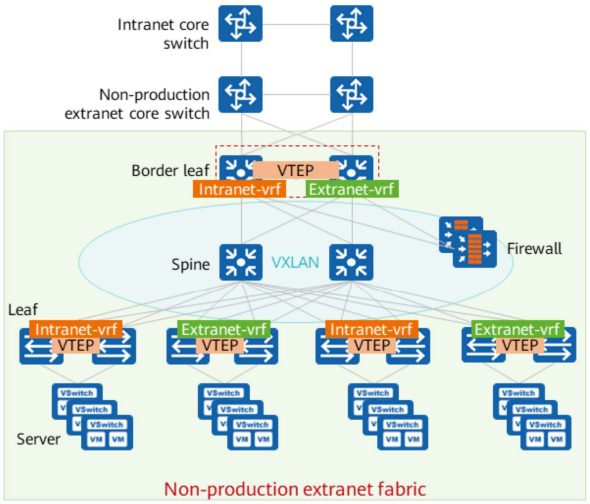


- Networking design:
  - The production extranet fabric network connects to production extranet core switches through an independent physical network. The distributed network overlay is used, and gateways are deployed on leaf switches.
  - Production extranet core switches are connected to intranet core switches through 40GE links.
  - Border leaf nodes and spine nodes are deployed independently (combining border leaf nodes and spine nodes is recommended for small-scale networks). Border leaf nodes and service leaf nodes are combined to connect to firewalls and IDSs.
  - M-LAG is deployed between border leaf nodes. Each border leaf node provides two 40GE ports to connect to two production extranet core switches.
  - The oversubscription ratio of leaf nodes is 3:1. Two leaf nodes function as TOR switches, and two 40GE links are configured as a peer-link to connect them. Each leaf node uses two 40GE links to connect to two spine nodes.
  - Firewalls are connected to border leaf nodes in bypass mode to perform access control on all incoming and outgoing traffic of the fabric. Each firewall is connected to two border leaf nodes through two 10GE links.
  - On the production extranet fabric, VPCs (VRFs) are used to isolate intranet traffic from extranet traffic.

HUAWEI

- Routing design:
  - The underlay network uses OSPF, the overlay network uses IBGP EVPN, and spine nodes function as RRs.

Non-Production Extranet Fabric Network Design

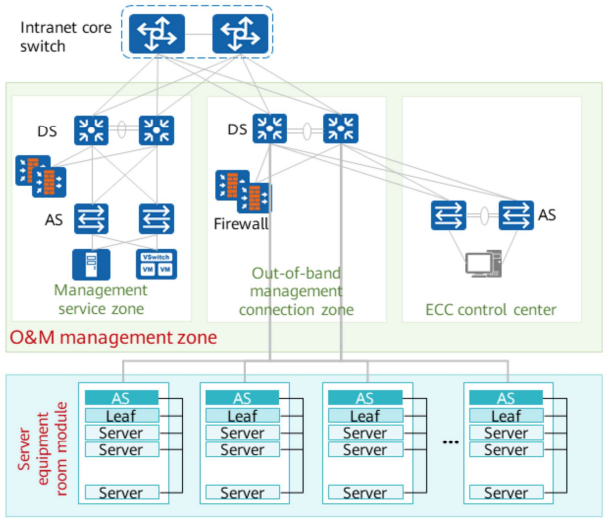


- Networking design:
  - The non-production extranet fabric network connects to non-production extranet core switches through an independent physical network. The distributed network overlay is used, and gateways are deployed on leaf switches.
  - Non-production extranet core switches are connected to intranet core switches through 40GE links.
  - Border leaf nodes and spine nodes are deployed independently (combining border leaf nodes and spine nodes is recommended for small-scale networks). Border leaf nodes and service leaf nodes are combined to connect to firewalls and IDSs.
  - M-LAG is deployed between border leaf nodes, and each border leaf node provides two 40GE ports to connect to two non-production extranet core switches.
  - The oversubscription ratio of leaf nodes is 3:1. Two leaf nodes function as TOR switches, and two 40GE links are configured as a peer-link to connect them. Each leaf node uses two 40GE links to connect to two spine nodes.
  - Firewalls are connected to border leaf nodes in bypass mode to perform access control on all incoming and outgoing traffic of the fabric. Each firewall is connected to two border leaf nodes through two 10GE links.
  - On the non-production extranet fabric, VPCs (VRFs) are used to isolate intranet traffic from extranet traffic.

HUAWEI

- Routing design:
  - The underlay network uses OSPF, the overlay network uses IBGP EVPN, and spine nodes function as RRs.

O&M Management Zone Design

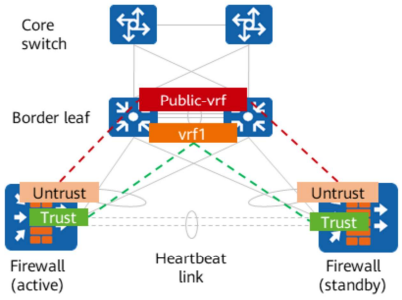


- Networking design:
  - The O&M management zone consists of the ECC control center, management service zone, and out-of-band management connection zone.
  - Management service zone:
    - The management service zone consists of DSs and ASs:
    - Two DSs are deployed in M-LAG mode and are cross-connected to two intranet core switches through four 40GE uplinks. Two firewalls are connected to the DSs in bypass mode and work in active/standby mode.
  - Out-of-band management connection zone
    - The out-of-band management connection zone consists of two DSs and several ASs. The ASs are deployed in server cabinets and connect to the out-of-band management interfaces of servers and out-of-band management interfaces of network devices.
    - Two DSs are deployed in M-LAG mode and are cross-connected to two intranet core switches through four 40GE uplinks. Two firewalls are connected to the DSs in bypass mode and work in active/standby mode.
  - The ASs in the ECC are connected to uplink out-of-band management DSs.

HUAWEI

- Routing design:
  - The DSs in the management service zone and out-of-band management connection zone are bound to VRFs, default routes pointing to firewalls are configured based on VRFs, and static routes are imported into OSPF on the DSs.
- Note: ECC is short for Emergency Command Center.

Firewall Access Design

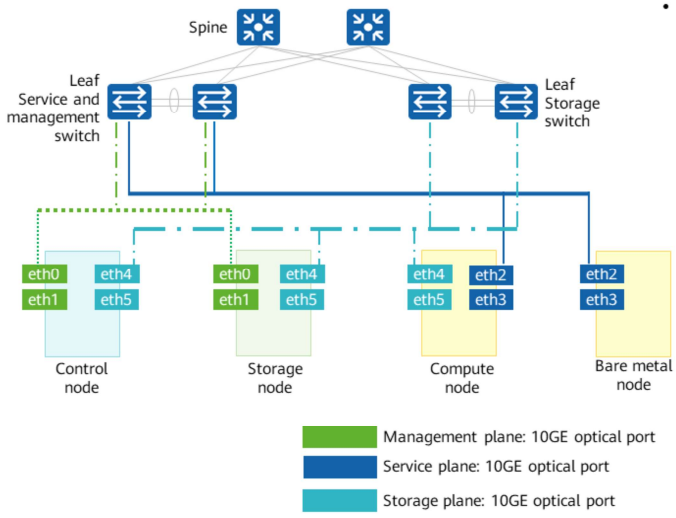


- Networking design:
  - Firewalls are connected to border leaf nodes in bypass mode and are logically connected between the service VRF and Public-vrf.
  - Firewalls are deployed in active/standby mode. Each firewall is connected to two border leaf nodes through two 10GE links. M-LAG is deployed on border leaf nodes to connect to the active and standby firewalls.
  - An Eth-Trunk is deployed on the firewalls, and Eth-Trunk sub-interfaces are used to distinguish traffic of the Trust zone from that of the Untrust zone.
  - Two 10GE links are deployed between the active and standby firewalls as heartbeat synchronization links.
  - Firewalls perform access control on all incoming and outgoing traffic of the fabric.

HUAWEI

- Routing design:
  - Border leaf nodes connect to core switches through Public-vrf, run OSPF between Public-vrf and core switches, and use service VRFs to isolate VPC routes.

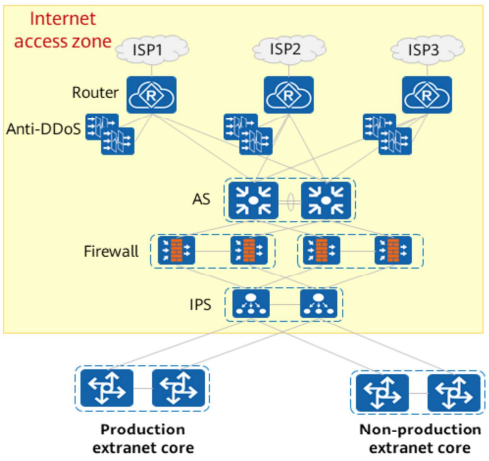
Server Access Design



- Design solution:
  - Servers in adjacent cabinets share two groups of leaf nodes, which are connected to a server's service & management NIC and storage NIC, respectively.
  - The two groups of leaf nodes are connected to spine nodes.
  - M-LAG is deployed between leaf nodes for connecting to servers, two 40GE links are configured as a peer-link, and each group of leaf nodes uses four 40GE links to connect to two spine nodes.
  - Server SAN storage servers are deployed in a storage node cabinet, and traffic between Server SAN nodes is heavy. It is recommended that storage nodes be connected to independent storage switches.
  - Network configurations need to be performed for the management and storage planes in advance through iMaster NCE-Fabric. A network administrator needs to plan access interfaces for the service and storage planes and connect storage NICs to storage switches.

HUAWEI

Internet Access Zone Design

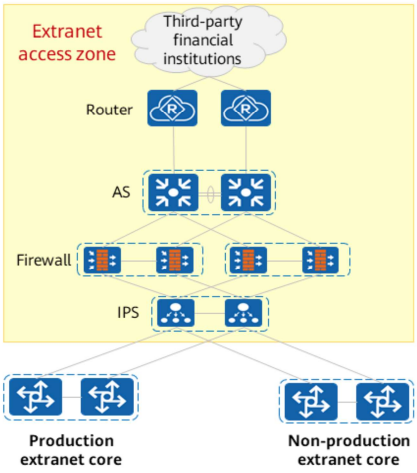


- Networking design:
  - Three access routers are connected to the links of three ISPs and use static routes to connect to the carrier networks.
  - Each router is connected to two anti-DDoS devices in bypass mode. One anti-DDoS device detects DDoS traffic, and the other cleans the traffic.
  - The access routers are connected to access switches (ASs) to aggregate traffic of the production Internet and non-production Internet. Two ASs are connected through two 40GE links.
  - There are two groups of extranet firewalls, each of which are connected to ASs through two 10GE uplinks in square-looped topology and are connected to two IPS devices through two 10GE downlinks in square-looped topology.
  - Two IPS devices are connected to two groups of firewalls through four 10GE uplinks, and are connected to core switches on production and non-production extranets through 10GE downlinks in square-looped topology. The IPS devices are deployed in transparent mode.

HUAWEI

- Routing design:
  - Static routes are used between extranet firewalls and ASs, and extranet firewalls perform NAT.
  - Policy-based routing (PBR) is deployed on access routers to divert traffic to production or non-production firewalls.

Extranet Access Zone Design

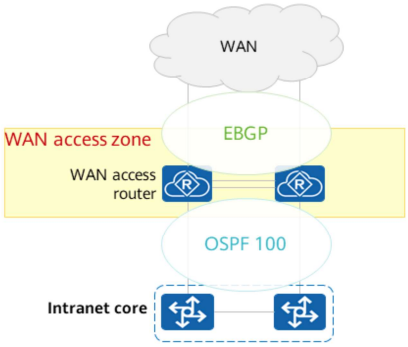


- Networking design:
  - Access routers are connected to third-party extranet private lines.
  - ASs are connected to access routers in square-looped topology to aggregate traffic from production and non-production third-party devices. Two ASs are connected through two 40GE links.
  - Two groups of extranet firewalls perform access control on production and non-production extranet traffic. Each group of extranet firewalls are connected to two ASs through two 10GE ports in square-looped topology, and connected to two IPS devices through two 10GE ports in square-looped topology.
  - The IPS devices are deployed in transparent mode and connected to the core switches on the production and non-production extranets through 10GE links in square-looped topology.

HUAWEI

- Routing design:
  - Access routers are connected to the third-party extranet through static routes. OSPF runs between two access routers. Static routes are used between ASs and access routers. Static routes are used between extranet firewalls and ASs.
  - Extranet firewalls perform NAT.
  - Routing policies are configured on access routers to divert traffic to production or non-production firewalls.
  - Return routes are deployed on ASs to divert traffic to production or non-production firewalls.

WAN Access Zone Design

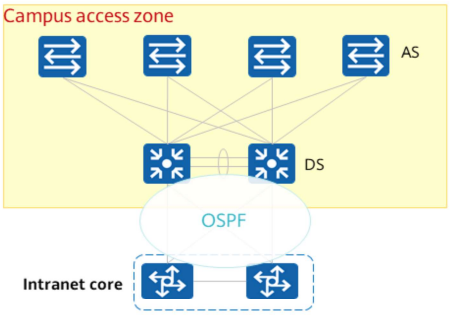


- Networking design:
  - Intranet core switches are connected to two WAN access routers through two 10GE links in square-looped topology.
  - Two WAN access routers are connected through two 10GE links.

HUAWEI

- Routing design:
  - OSPF is deployed between intranet core switches and WAN access routers.
  - EBGP is deployed between WAN access routers and WAN routers.
  - Summarized OSPF routes are imported into BGP on WAN access routers.
  - Routing policies are deployed on WAN access routers to import only intranet service routes.

Campus Access Zone Design



- Networking design:
  - The campus network uses the traditional standard design model, core/aggregation layer and access layer.
  - Aggregation switches (DSs) are modular switches, and access switches (ASs) are fixed switches. ASs are connected to two DSs through two 10GE links.
  - The DSs are cross-connected to two intranet core switches through 40GE uplinks.
  - All the interconnected links between devices are 10GE links.

HUAWEI

- Routing design:
  - OSPF is deployed between the DSs and intranet core switches.

Contents

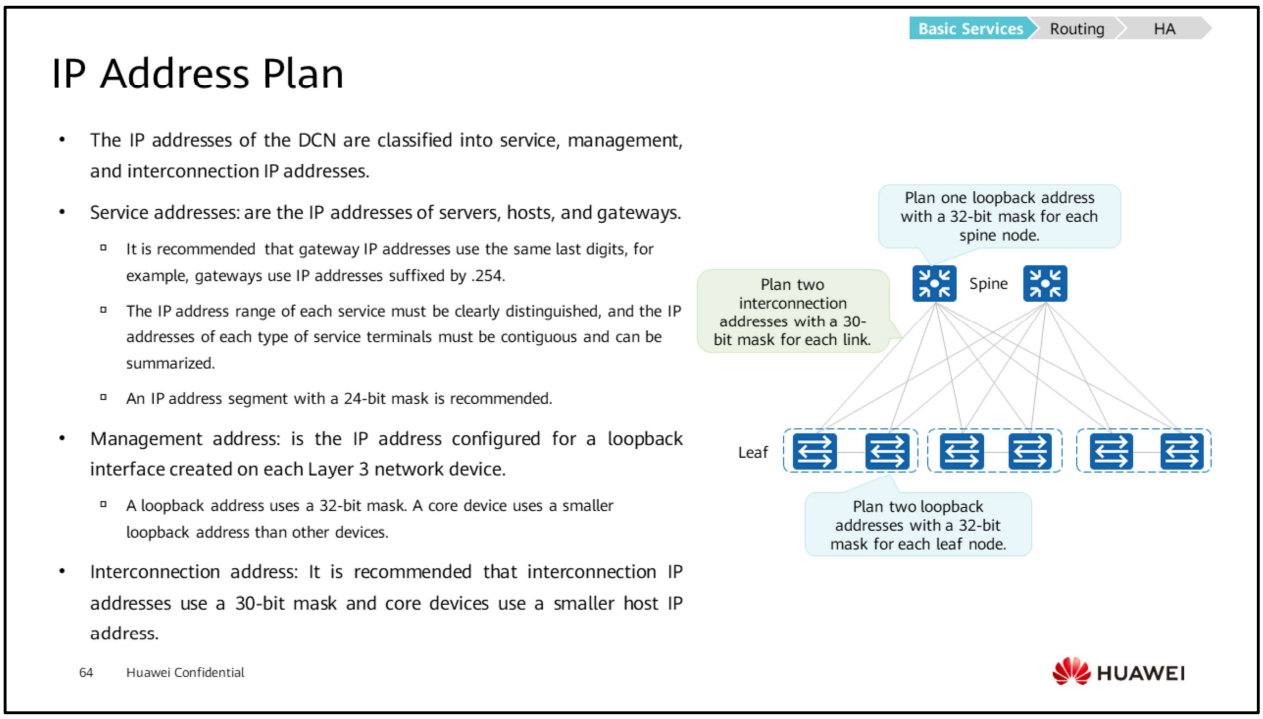
1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

HUAWEI

DCN General Design



HUAWEI



- IP address planning principles are as follows:
  - IP addresses on the entire network must be managed and allocated uniformly.
  - IP address allocation should be simple and easy to manage, reflect network layers, simplify network management and network expansion, and be visualized.
  - Extensibility of IP address planning must be ensured. That is, some IP addresses should be reserved at each layer so that IP addresses to be summarized can be contiguous during network expansion.
  - IP addresses should be contiguous. The routes with contiguous addresses can be summarized easily on the hierarchical network. This reduces the routing table size and speeds up route calculation and route convergence.
  - IP address allocation must be flexible to allow optimization of various traffic, security, and routing policies and make full use of the address space.
- IP address allocation principles:
  - The network ID with variable length and host address mask are used for IP address allocation. Some IP addresses need to be reserved based on the number of hosts on network segments. This ensures that IP addresses can be summarized and prevents the waste of IP addresses.
  - To facilitate route summarization, assign IP addresses on the same network segment to devices in the same network area. If the preceding requirements cannot be met due to limitations, ensure that routes can be summarized at the aggregation layer.

VNI and VPC Plan

VNI design

VNI:
- Unique ID of a DC
- Unique ID of an equipment room module in a DC
- Unique ID of OpenStack in an equipment room module of a DC

VNI range:
- Primary DC: VNIs from 1000000 to 1999999, a total of 1 million VNIs
- Intra-city DR DC: VNIs from 2000000 to 2999999, a total of 1 million VNIs
- Remote DR cloud DC: VNIs from 3000000 to 3999999, a total of 1 million VNIs

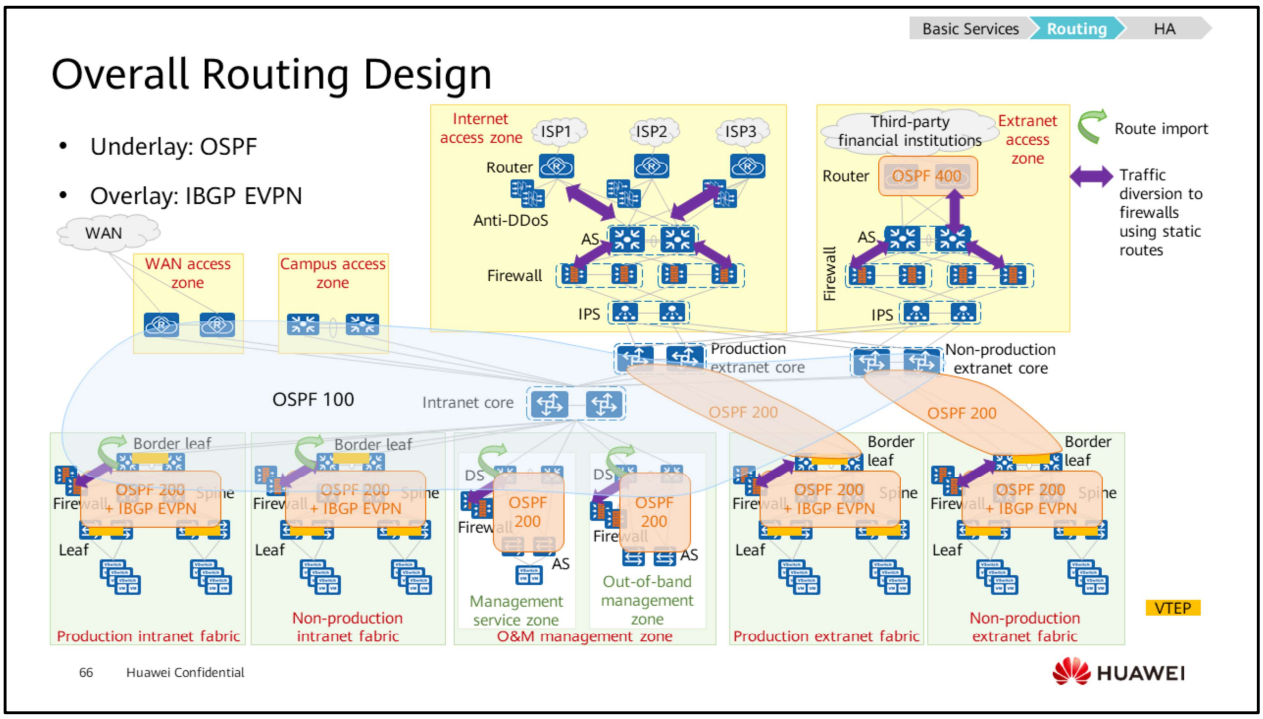
VPC design

Fabric-VPC:
- VPC-Srv-DMZ: production extranet
- VPC-Mgt-DMZ: service assurance extranet
- VPC-Srv-Intranet: production intranet
- VPC-Mgt-Intranet: service assurance intranet
- VPC-Mgt: in-band management

VNI allocation example:
- DC ID (leftmost first digit): The value ranges from 1 to 9. It can be 1 (primary DC), 2 (intra-city DR DC), 3 (remote DR DC), or 4 to 9 (reserved).
- Equipment room ID (leftmost second and third digits): The value ranges from 01 to 99.
- OpenStack ID (leftmost fourth digit): The value ranges from 1 to 9.
- Subnet ID (leftmost fifth, sixth, and seventh digits): The value ranges from 001 to 999. That is, each OpenStack can use 999 subnets.

VNI allocation example: If the primary DC equipment room is the first enabled cloud network equipment room module, the possible VNI in the fabric is 1011001.

HUAWEI

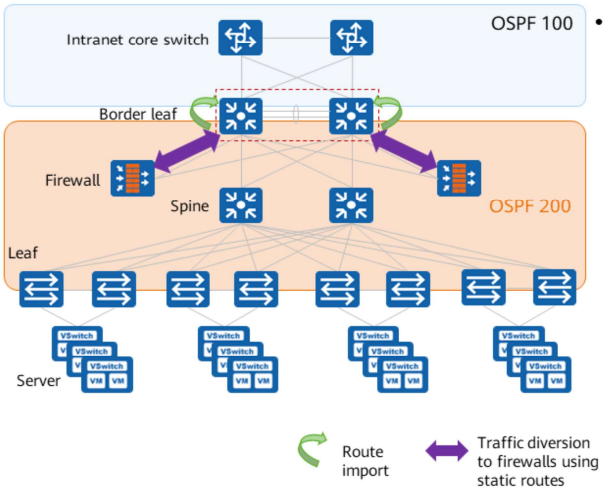


- OSPF 100 is deployed between interfaces that connect intranet core switches to production extranet core switches and non-production extranet core switches, the product intranet, non-product intranet, the WAN access zone, campus access zone, and the O&M management zone.
- OSPF 200 is deployed between extranet core switches and the production extranet. OSPF 100 imports routes of OSPF 200, and OSPF 200 imports the summary routes of OSPF 100.
- OSPF 200 is deployed between non-extranet core switches and the non-production extranet. OSPF 100 imports routes of OSPF 200, and OSPF 200 imports the summary routes of OSPF 100.
- OSPF 200 is deployed on the underlay network of the production intranet, static routes are deployed between border leaf nodes and firewalls, and static routes are imported into OSPF 100 on the border leaf nodes. IBGP EVPN is deployed on the overlay network of the production intranet.
- OSPF 200 is deployed on the underlay network of the non-production intranet, static routes are deployed between border leaf nodes and firewalls, and static routes are imported into OSPF 100 on the border leaf nodes. IBGP EVPN is deployed on the overlay network of the non-production intranet.

- OSPF 200 is deployed on the underlay network of the production extranet, static routes are deployed between border leaf nodes and firewalls, and static routes are imported into OSPF 100 on the border leaf nodes. IBGP EVPN is deployed on the overlay network of the production extranet.
- OSPF 200 is deployed on the underlay network of the non-production extranet, static routes are deployed between border leaf nodes and firewalls, and static routes are imported into OSPF 100 on the border leaf nodes. IBGP EVPN is deployed on the overlay network of the non-production extranet.
- OSPF 200 is deployed in the O&M management zone. Static routes pointing to firewalls are configured on DSs. The DSs import static routes into OSPF 100.
- In the campus access zone, static routes pointing to firewalls are configured on DSs, and the DSs import static routes into OSPF 100.
- Static routes are used between firewalls in the Internet access zone and core switches on the production and non-production extranets. In the Internet access zone, static routes are used between the ASs and access routers.
- Static routes are used between firewalls in the extranet access zone and core switches on the production and non-production extranets. In the extranet access zone, static routes are used between the ASs and access routers. OSPF 400 is deployed between two access routers.

Basic Services > Routing > HA

Underlay Routing Design for the Resource Pool

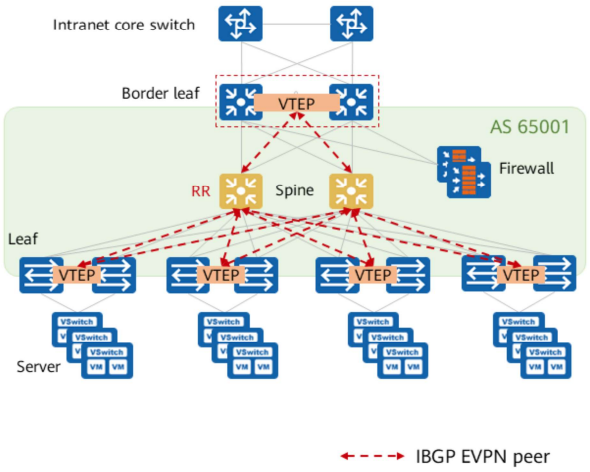


- Deployment design:
  - OSPF 100 is enabled between Public-vrf of border leaf switches and intranet core switches.
  - OSPF 200 is deployed on the interconnection interfaces between border leaf switches and leaf switches.
  - Each switch in the OSPF routing domain is configured with an independent loopback address as the router ID and OSPF is enabled on the loopback interface using the network command. Devices are interconnected at Layer 3. The OSPF network type can be set to P2P.
  - Static routing is deployed between firewalls and border leaf switches.
  - On the border leaf switch, a default route pointing to the firewall is configured for the service VRF and imported to OSPF 100.

HUAWEI

- Route convergence acceleration design:
  - Configure BFD for OSPF to speed up route convergence.
  - Configure the OSPF route calculation interval and intervals for updating and receiving LSAs to speed up route convergence in case of OSPF network faults.

Overlay Routing Design for the Resource Pool

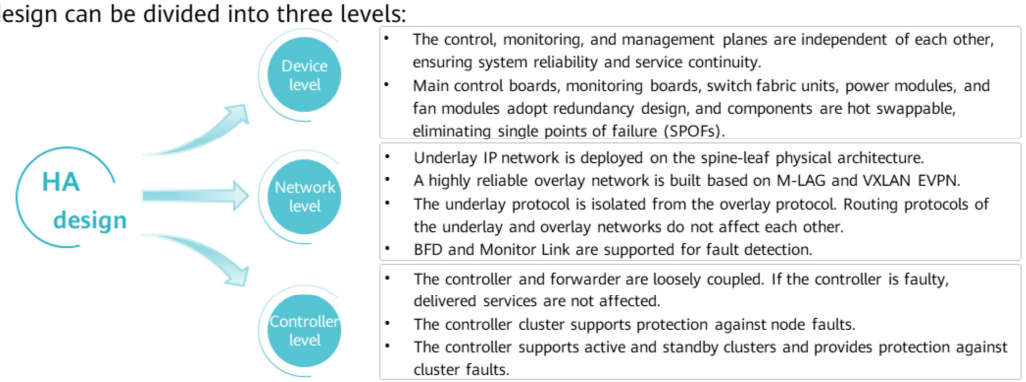


- Deployment design:
  - The distributed network overlay is used. Each leaf node functions as a Layer 3 gateway, and traffic is forwarded along the shortest path.
  - IBGP EVPN is deployed, and each fabric has an independent autonomous system (AS).
    - When IBGP is deployed, loopback addresses are used to establish IBGP peer relationships, and ECMP can be used to improve reliability.
    - IBGP RRs reduce the number of IBGP peers and device resource consumption. Two spine switches are used as RRs.

HUAWEI

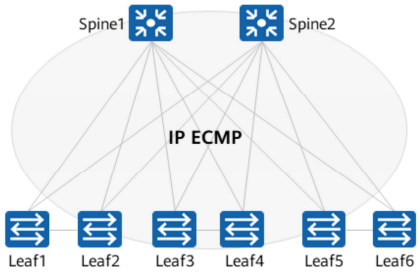
HA Design Overview

- As the core department of the customer's IT infrastructure, the DC stores various data, runs a variety of services, and provides services to external networks. DC faults will cause great loss in every year. Therefore, stable and reliable running of DCs is critical.
- HA design can be divided into three levels:



HUAWEI

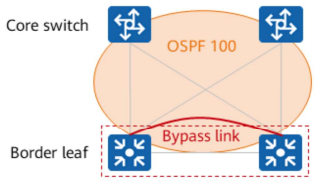
Spine HA Design



- Link redundancy: Spine nodes are fully meshed with all leaf nodes to form a full-mesh architecture.
- Device redundancy: Multiple high-density DC switches are deployed to implement device-level reliability. In a typical deployment scenario, two or four spine nodes are deployed. The number of deployed spine nodes depends on the live network scale.
- Network redundancy:
  - In the spine-leaf networking architecture, multiple spine nodes are deployed to construct an IP ECMP load balancing network, implementing network-level reliability.
  - Routing protocols (BGP/OSPF) are deployed to achieve high-reliable architecture.
- Fault detection and tolerance design:
  - When the link of a spine node fails, leaf nodes quickly switch traffic to a normal link through ECMP routes on the underlay network.
  - If a spine node fails, leaf nodes quickly switch traffic to other spine nodes through ECMP routes on the underlay network.

HUAWEI

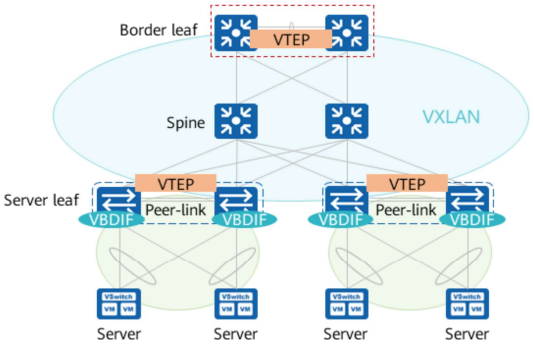
Border Leaf HA Design



- Link redundancy:
  - It is recommended that border leaf nodes be connected to different core switches and spine nodes or to the same core switch and spine node through multiple links across boards.
  - Eth-Trunk needs to be deployed as a peer-link through at least two links between border leaf nodes, and inter-board deployment is recommended.
- Device redundancy: Active-active device groups are deployed on border leaf nodes to implement device-level reliability.
- Network redundancy:
  - It is recommended that border leaf nodes be cross-connected to core switches and be fully meshed with spine nodes to build an IP ECMP network, eliminating Layer 3 loops.
  - Link aggregation protocols (LAG or M-LAG) and routing protocols (BGP or OSPF) are used to ensure high reliability of the architecture.
- Fault detection and tolerance design:
  - A bypass link is deployed between border leaf nodes to prevent traffic interruptions caused by the failure of all uplink interfaces. The bandwidth must be at least equal to the total bandwidth of the uplinks of a single device. (Optional. The bypass link is required in square-looped topology.)

HUAWEI

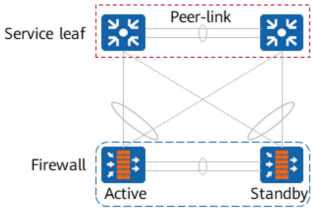
Server Leaf HA Design



- Link redundancy: Eth-Trunk needs to be deployed as a peer-link through at least two links between server leaf nodes, and inter-board deployment is recommended.
- Device redundancy: Server leaf nodes are deployed as an M-LAG active-active device group to implement device-level reliability.
- Network redundancy:
  - Server leaf nodes are fully meshed with spine nodes to form an IP ECMP network, eliminating Layer 3 loops.
  - Link aggregation protocols (LAG or M-LAG) and routing protocols (BGP or OSPF) are used to ensure high reliability of the architecture.
- Fault detection and tolerance design:
  - If possible, configure dual-active detection (DAD) based on out-of-band management network ports. Otherwise, configure DAD based on service network ports.
  - Deploy a Monitor Link. If an uplink fails, the associated downlink also goes Down, preventing traffic interruptions.
  - Configure storm suppression of broadcast, unknown unicast, and multicast packets (BUM packets) on the downlink ports of leaf nodes. When VLAN 1 is not used, packets from this VLAN are denied to prevent loops.

HUAWEI

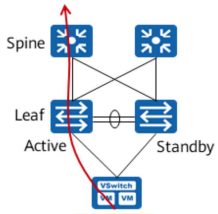
Firewall HA Design



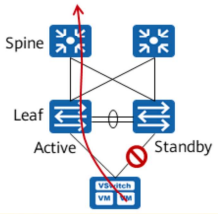
- Link redundancy: Each firewall is dual-homed to the M-LAG active-active service leaf device group through link bundling.
- Device redundancy: Firewalls are deployed in active/standby mode and support hot standby.
- Fault detection and tolerance design:
  - A heartbeat link is deployed between the active and standby firewalls to check the status of the peer firewall and the connectivity of the heartbeat link.
  - Each firewall can be connected to service leaf nodes in one-armed or two-armed mode. In one-armed mode (for example, a firewall is logically connected to a service leaf node in one-armed mode), if the uplink logical port is faulty, the downlink logical port is also faulty, and vice versa. Therefore, fault detection is faster.
  - Link fault: If any link between the active firewall and a service leaf node fails, the M-LAG member interface goes Down, and the dual-homing networking changes to single-homing networking, without affecting traffic forwarding on the active firewall.
  - Firewall fault: After an active/standby firewall switchover is performed, the new active firewall sends gratuitous ARP packets to trigger the update of forwarding entries on the service leaf nodes and divert traffic to the new active firewall.

HUAWEI

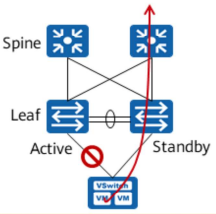
Examples of HA Design Fault Scenarios



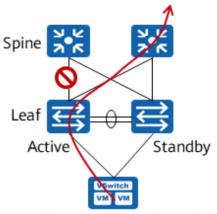
Normal



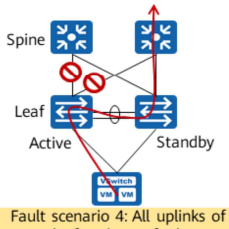
Fault scenario 1: The standby NIC or standby link is faulty.



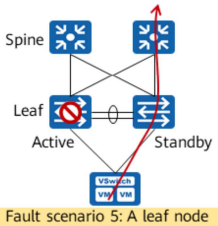
Fault scenario 2: The active NIC or active link is faulty.



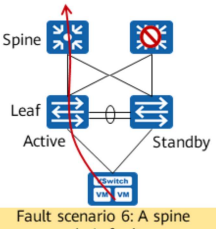
Fault scenario 3: The uplink of a leaf node is faulty.



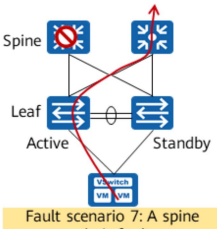
Fault scenario 4: All uplinks of a leaf node are faulty.



Fault scenario 5: A leaf node is faulty.



Fault scenario 6: A spine node is faulty.



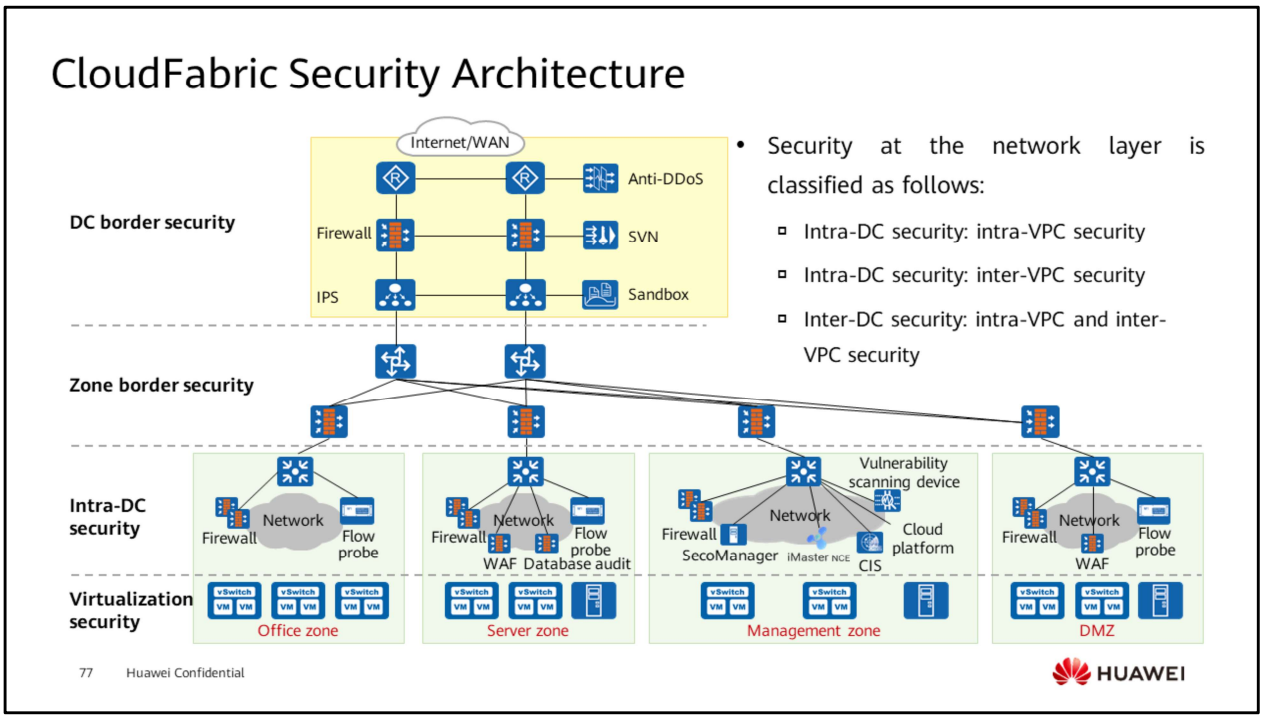
Fault scenario 7: A spine node is faulty.

HUAWEI

Contents

1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

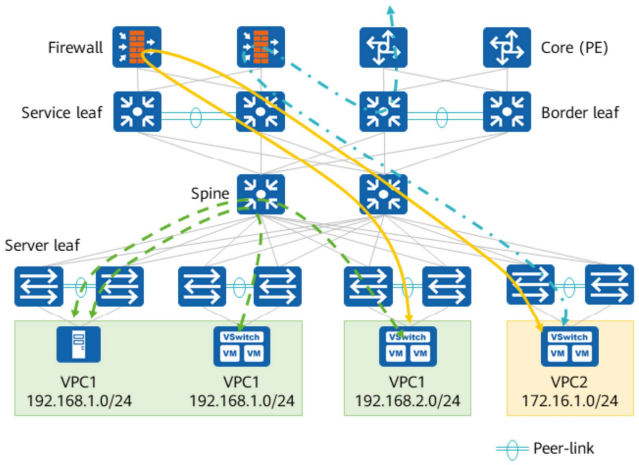
HUAWEI



- Virtualization security:
  - Security groups of the cloud platform are used to protect VMs. The cloud platform adds VMs that require security control to different security groups through orchestration and defines security policies between security groups for access control.
- Intra-DC security:
  - Office zone:
    - Access control policies are configured on firewalls to protect security of mutual access between intranet office users.
    - The CIS flow probe can be deployed to collect traffic in the office zone for in-depth threat detection.
  - Server zone:
    - Firewalls or web application firewalls (WAFs) are used to protect intranet servers.
    - Database audit is used to protect database servers.
    - The CIS flow probe can be deployed to collect traffic in the server zone for in-depth threat detection, preventing intranet threats from being spread.

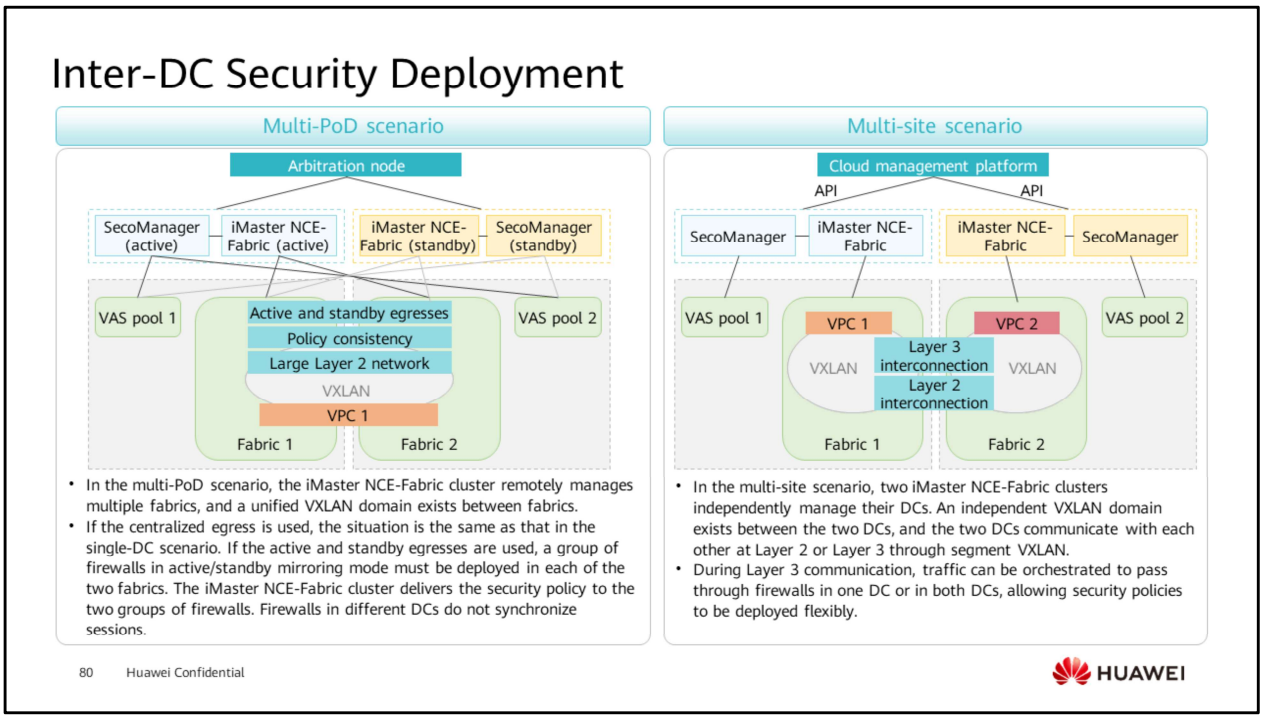
    - Management zone: Vulnerability scanning devices, cloud platform, network controller (iMaster NCE-Fabric), security controller (SecoManager), and security analyzer (CIS) are deployed to provide core security services for the entire network.
    - DMZ:
        - Firewalls or WAFs are deployed to defend against external threats and ensure security of servers in the DMZ.
        - The CIS flow probe can be deployed to collect traffic in the DMZ for in-depth threat detection.
- Zone border security:
    - Security policies are configured on firewalls to protect access traffic between zones.
- DC border security:
    - Professional anti-DDoS devices are used to defend against external DDoS attacks.
    - Secure Sockets Layer VPN (SVN) access devices are deployed to provide secure access based on IPsec and SSL VPN.
    - Professional Intrusion Prevention System (IPS) devices are deployed to detect and protect north-south traffic.
    - The sandbox is deployed to prevent malicious emails, web traffic, scanning, and penetration.
    - Firewalls are deployed to protect incoming and outgoing traffic passing through a DC border.

Intra-DC Security Deployment



- Intra-VPC traffic can be directly forwarded through VXLAN. Alternatively, SFC can be used to redirect traffic on the same subnet or different subnets to VAS devices to implement in-depth security access control for traffic on the subnets.
- Inter-VPC traffic is forwarded across subnets and security isolation is required. Therefore, the traffic needs to pass through the firewall for security access control.
- Traffic sent from an external user outside a DC to a server in a VPC inside the DC needs to pass through VAS devices such as the IPS/firewall and LB for security access control or load balancing.

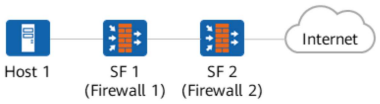
HUAWEI



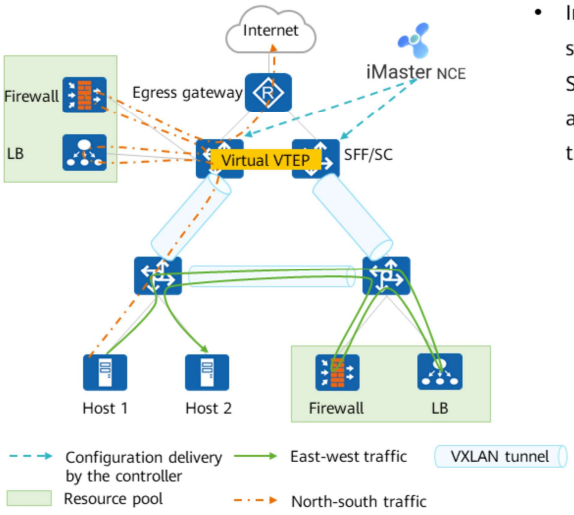
- A point of delivery (PoD) is a minimum service unit in a DC, which is a collection of resources, including switches, firewalls, load balancers, and servers. The resources collaborate with each other to provide network services and applications.

SFC

- Service Function Chaining (SFC) technology logically connects services on network devices to provide an ordered service set for the application layer.

Basic concepts	SFC implementation
• SFC domain: A domain where SFs are deployed is called an SFC domain. • Service classifier (SC): identifies the service flows to be diverted to the SFP. • Service Function (SF): provides value-added services. SFs include firewalls, LBs, application accelerators, lawful interception (LI) devices, and NAT devices. • Service Function Forwarder (SFF): is a device that connects to SFs. It can identify and forward service flow information by SFs. • Service Function Path (SFP): An SFP is calculated based on configurations and precisely defines the location of each SF.	• SFC can be implemented in Policy-based Routing (PBR) mode. • PBR mode: ▫ PBR selects routes based on user-defined policies. ▫ PBR is used to redirect service traffic based on the modular QoS command (MQC). In this way, traffic is forwarded based on a specified path instead of the routing table. **SFC** 

SFC Application Scenarios

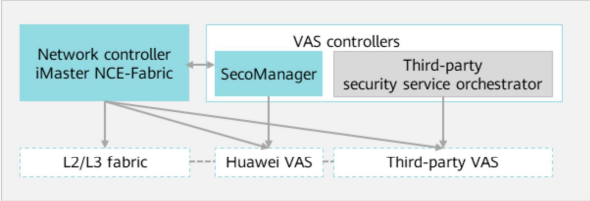


- In addition to SFC deployment in single-node mode, the switch can cooperate with iMaster NCE-Fabric to implement SFC. iMaster NCE-Fabric orchestrates SFs, configures SFPs, and delivers the SFP configuration to the SC and SFFs through NETCONF interfaces.
  - East-west traffic: When servers of different security levels need to access each other, service nodes in the resource pool through which east-west access traffic passes can be flexibly defined based on user requirements for security protection. In this way, service processing requirements can be implemented flexibly, easily, efficiently, and securely regardless of the topology.
  - North-south traffic: Security protection between external networks and DCNs is the core for network security. SFC flexibly orchestrates external north-south traffic on different SFs to implement various functions such as NAT and security filtering.

HUAWEI

Security Services

- Security service architecture:

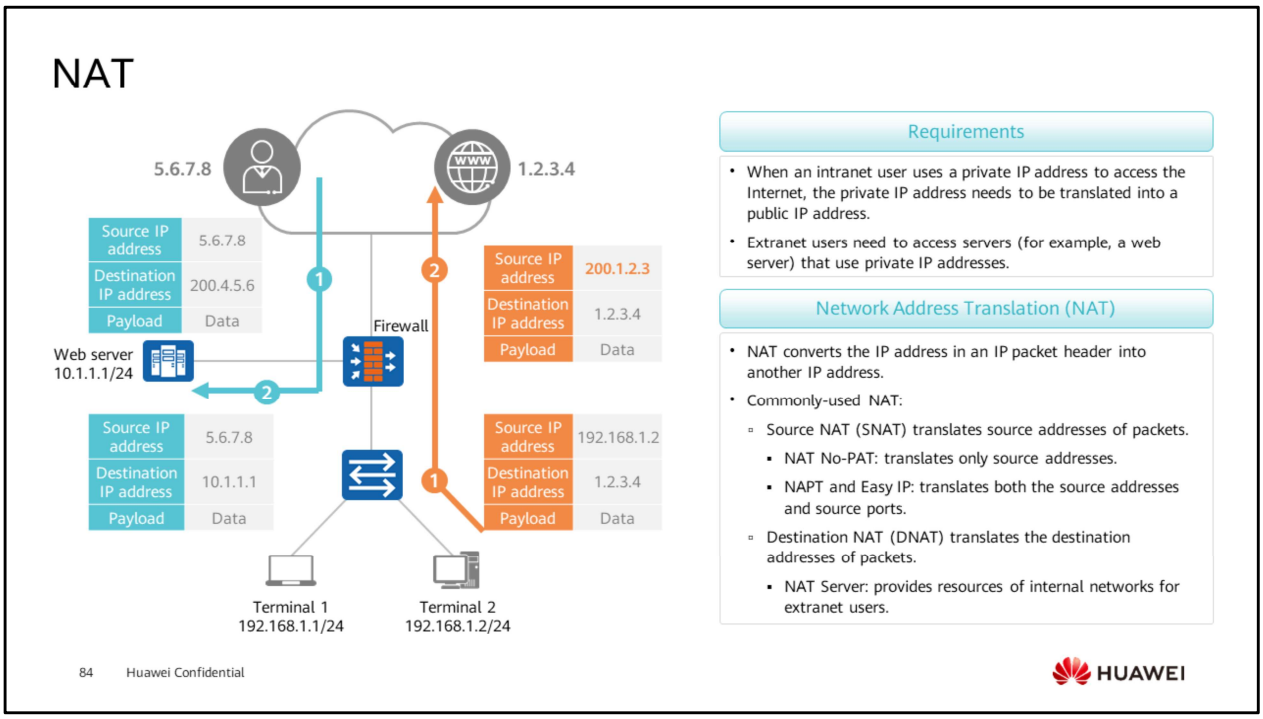


- iMaster NCE-Fabric: provisions logical network services, orchestrates the bidirectional interconnection network between Huawei VAS devices and Layer 2 or Layer 3 fabric, and manages Huawei CE switches and delivers network configurations to them.
- SecoManager: orchestrates services for Huawei VAS devices, manages Huawei VAS devices, and delivers network configurations to them.
- Huawei VAS device: provides service functions such as ACL, EIP, SNAT, IPsec VPN, and content security detection.

The CloudFabric solution provides the following security services:

- IPsec VPN
- SNAT
- Elastic IP
- Firewall-based bandwidth management
- Anti-DDoS
- Security policy
- Content security detection
- Virtual system

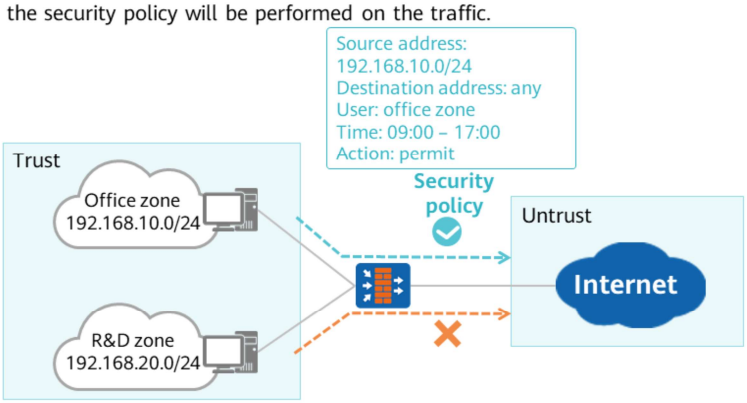
HUAWEI



- When an internal user accesses an extranet resource, the firewall processes packets as follows:
  - Upon receiving a packet destined from a private network user to the extranet resource, the firewall translates the source IP address of the packet from the private IP address to a public IP address.
  - Upon receiving the return packet, the firewall translates the destination IP address of the packet from the public IP address into a private IP address.
- When an extranet user accesses an internal web server, the firewall processes packets as follows:
  - Upon receiving a packet destined from the extranet user to the internal web server, the firewall translates the destination IP address of the packet from the public IP address into a private IP address.
  - Upon receiving the return packet, the firewall translates the source IP address of the packet from the private IP address into a public IP address.

Security Policy

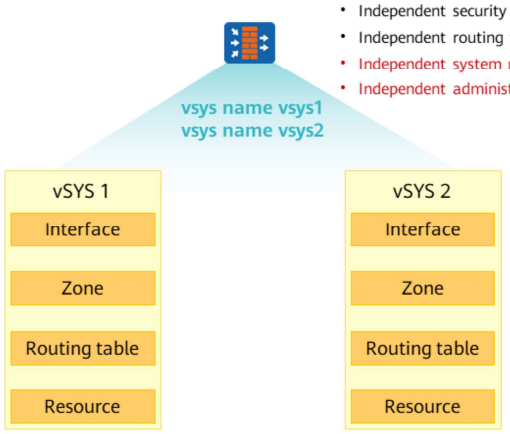
- A security policy controls traffic forwarding and performs content security detection on traffic.
- A firewall identifies the attributes (such as the quintuple information, user, and time range) of received traffic and matches the attributes against the conditions in a security policy. If the traffic matches the conditions, the action specified in the security policy will be performed on the traffic.



HUAWEI

- Actions of a security policy:
  - If the action is permit, the device processes the traffic as follows:
    - If content security detection is not configured, the device permits the traffic.
    - If content security detection is configured, the device determines whether to permit the traffic based on the content security detection result.
  - If the action is deny, the device denies the traffic.

Virtual System



A virtual system (vSYS) has:

- Independent interfaces
- Independent security zones
- Independent routing tables
- Independent system resources
- Independent administrator accounts

Requirements

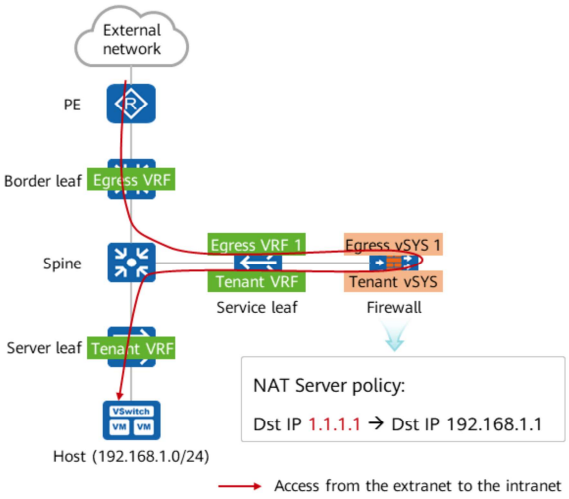
- vSYSs are independent logical devices created on a physical device.
- A firewall can be logically divided into multiple vSYSs. Every vSYS functions as a real device and has its own interfaces, address sets, users or user groups, routing entries, and policies. It can be configured and managed by the vSYS administrator.

Solution

- The traffic of different vSYSs is isolated to ensure security.
- Different vSYSs can still communicate with each other if needed.
- vSYS technology reduces hardware investment, power consumption, and equipment footprint.

HUAWEI

Elastic IP (EIP)



EIP

- EIP binds a public IP address to a specific tenant network's private IP address, so that various tenant network resources provide services for external systems by using the fixed public IP address.

Unidirectional access scenario

- Internal servers in a DC provide services externally using the floating IP address (1.1.1.1 in the figure), and external hosts access internal services of the DC using the floating IP address.
- Communication process
  - An external host accesses the floating IP address 1.1.1.1.
  - The packet is forwarded to the tenant vSYS, and the tenant vSYS translates the destination address into the actual address 192.168.1.1 of the internal host according to the delivered NAT Server policy.
  - The tenant vSYS forwards the packet to the internal host.

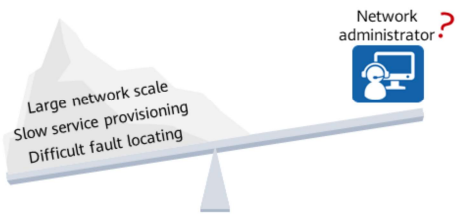
HUAWEI

Contents

1. CloudFabric Solution and DCN Overview
2. DCN Planning and Design Methods
3. **Typical DCN Design Cases**
   - Overview
   - DCN Overall Architecture Design
   - DCN Zone Design
   - DCN General Design
   - DCN Security Design
   - DCN O&M Design

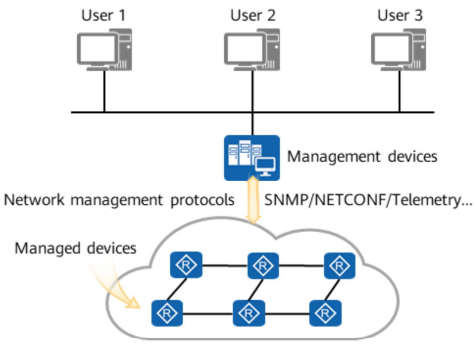
HUAWEI

Network O&M Overview



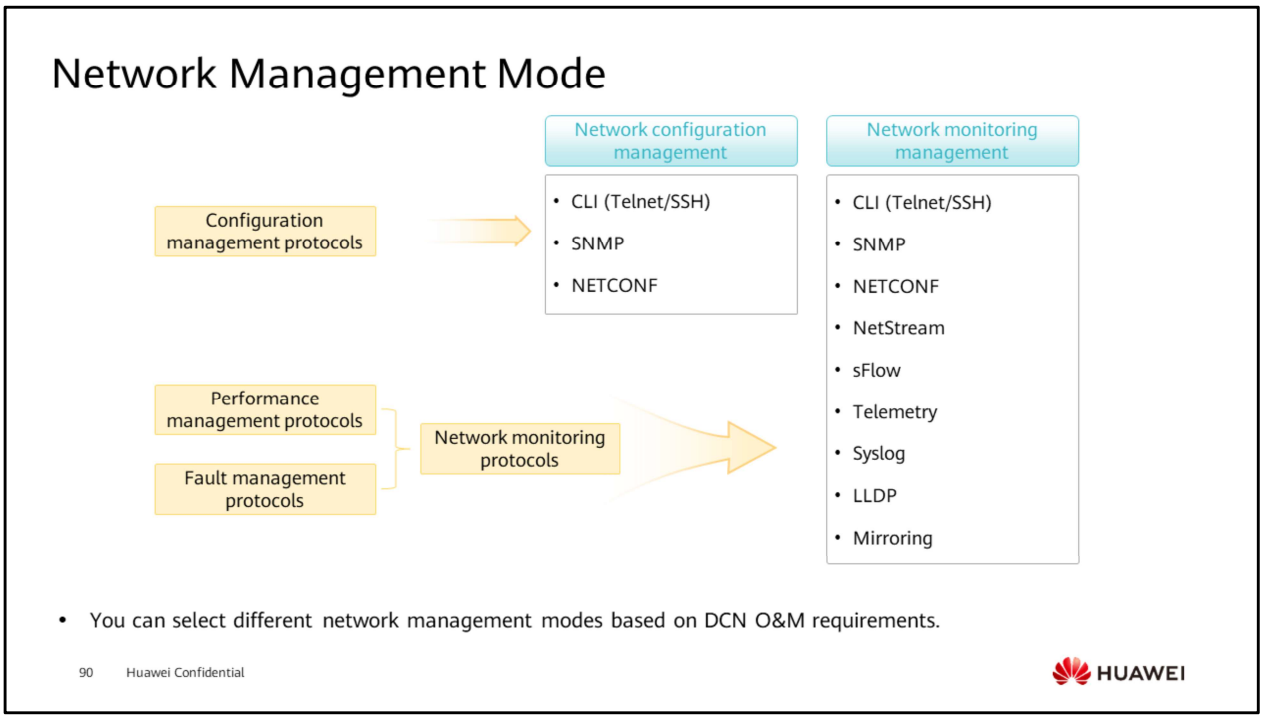
- As the new technologies such as artificial intelligence, big data, and cloud computing are developing rapidly, industries will undergo digital transformation and enterprise services will become diversified as digital transformation is implemented.
- Digitalization brings changes to the network model, and the traditional network management mode can no longer meet the new requirements of digital services. Against this backdrop, new intelligent O&M solutions are developed.

Typical architecture of network management systems



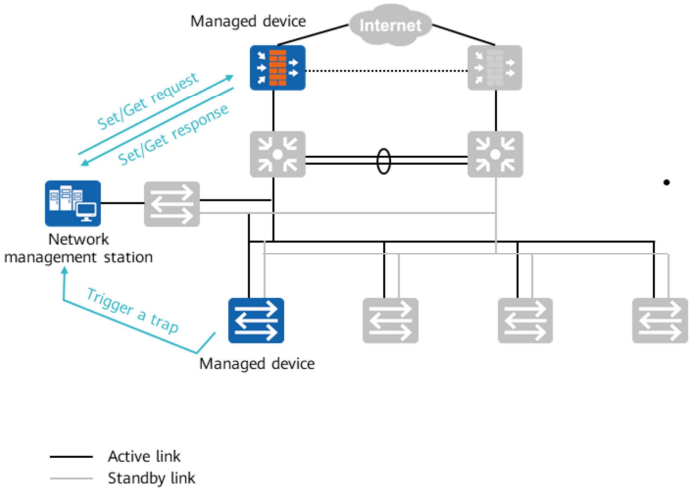
HUAWEI

- Network management systems usually have the same basic architecture, which consists of two key elements:
  - Managing device, also called a network management station
  - Managed device, also called an agent



- The command-line interface (CLI) supports both network configuration management and network monitoring management.
- For SNMP, the Set function is used for network configuration management, and the Trap function is used for network monitoring management.
- For NETCONF, The Edit function is used for network configuration management, and the Get function is used for network monitoring management.

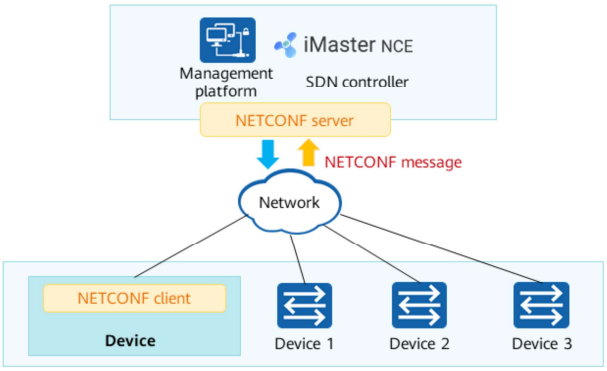
SNMP



- You can configure the SNMP management program in the network management station, enable the agent program on the managed device, and configure the SNMP protocol on the network.
- After SNMP is configured:
  - The network management station can obtain or change device information through the agent to implement remote monitoring and management.
  - The agent can report device status to the network management station in a timely manner.

HUAWEI

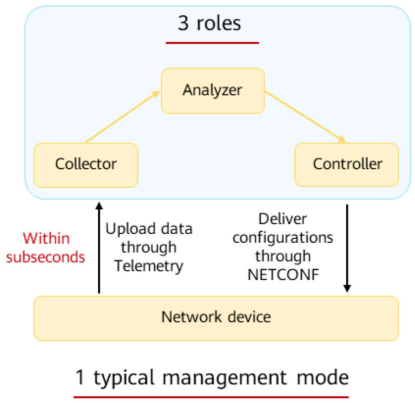
NETCONF



- NETCONF provides a set of mechanism for managing network devices. With this mechanism, users can add, modify, delete, back up, restore, lock, and unlock network device configurations. In addition, NETCONF provides transaction and session operation functions to obtain network device configuration and status information.
- NETCONF has three objects:
  - NETCONF client
  - NETCONF server
  - NETCONF message

HUAWEI

Telemetry



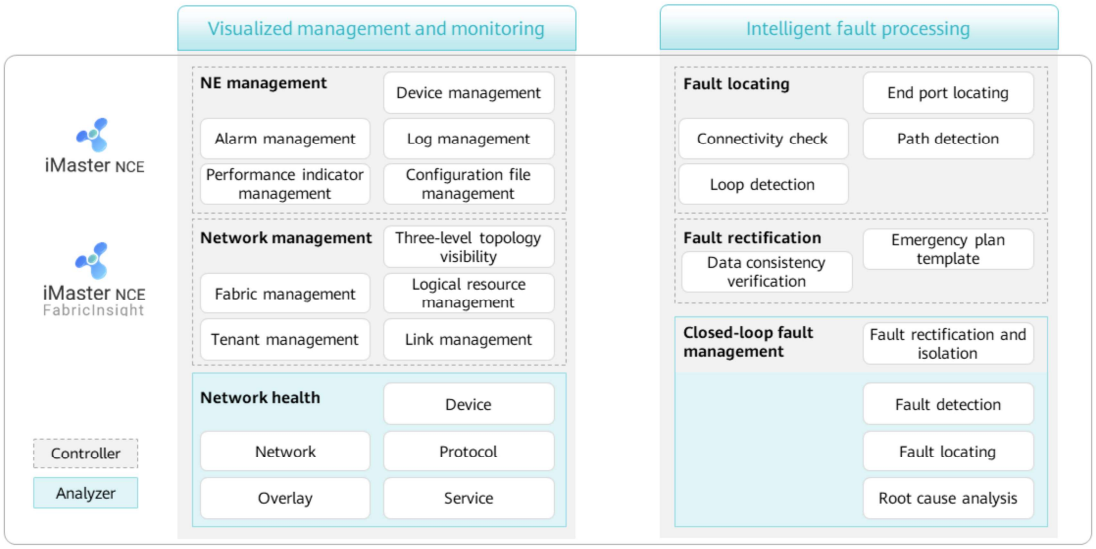
1 typical management mode

- Telemetry, also known as network telemetry, is used to monitor networks, including packet check and analysis, intrusion and attack detection, intelligent data collection, and application performance management.
- Advantages:
  - Supports multiple implementation modes, meeting diversified user requirements.
  - Collects a wide variety of data with high precision to fully reflect network status.
  - Continuously reports data with only one-time data subscription.
  - Locates faults rapidly and accurately.

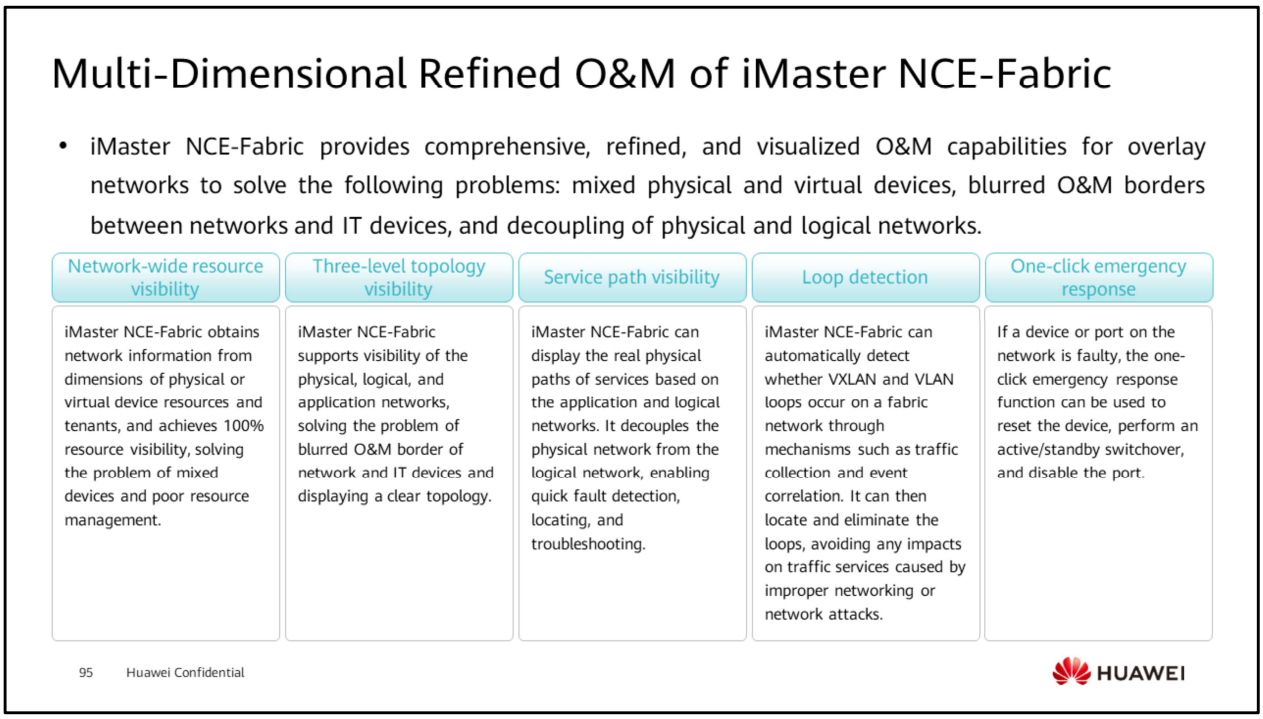
HUAWEI

- The collector, analyzer, and controller are components of the network management system.
  - The collector receives and stores monitoring data reported by network devices.
  - The analyzer analyzes the monitoring data received by the collector and processes the data, for example, displays the data on the GUI.
  - The controller uses NETCONF to deliver configurations to devices, so as to manage network devices. The controller can deliver configurations to network devices based on the analysis data provided by the analyzer and adjust the forwarding behavior of network devices. It also controls the data that the network devices sample and report.

CloudFabric O&M Function Overview



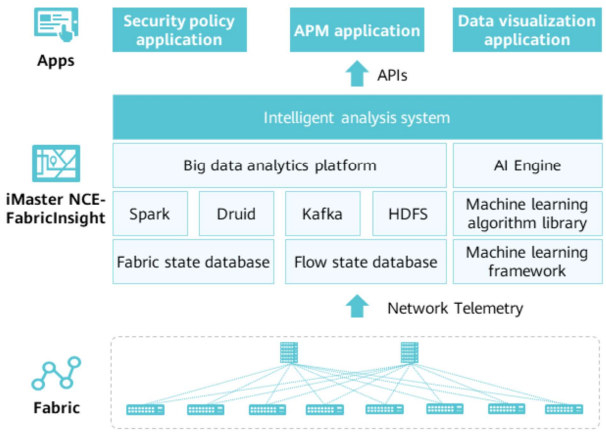
HUAWEI



- Network-wide resource visibility:
  - iMaster NCE-Fabric provides unified management of physical and virtual resources, and monitors the resource status of physical and virtual network devices on the entire network, enabling users to gain insights into the running status of all NEs in a timely manner.
  - iMaster NCE-Fabric monitors the network running status from the tenant dimension and displays all tenants, tenant quotas, and tenant traffic, enabling users to gain insights into the running status of the tenants.
- Three-level topology visibility:
  - iMaster NCE-Fabric can display the logical and physical network resources used by application networks, facilitating capacity expansion or reduction.
  - iMaster NCE-Fabric can display information including the tenant and application carried by a physical resource such as a device, link, or port. When a physical network changes, iMaster NCE-Fabric rapidly identifies the affected tenants and applications and sends notifications to them.
  - iMaster NCE-Fabric can display physical network resources and applications used and carried by the logical network and centrally control the services and networks.
  - When a physical network resource on the underlay network changes (device restart or link disconnection), iMaster NCE-Fabric automatically synchronizes the changes to logical and application networks.

- Service path visibility: This function supports the filtering of path information based on 5-tuple information and the display of hop-by-hop path information, enabling users to view service path information as needed.
    - Single-path detection traces the actual physical paths between VMs and checks whether service flows are interrupted.
    - Multi-path detection traces multiple physical paths between NVE devices to check whether service flows are interrupted.
    - Connectivity detection detects connectivity between VMs, VMs and BMs, as well as VMs/BMs and the Internet.
- One-click emergency response (currently, the following actions are supported):
    - VMM isolation and recovery
    - Cloud platform isolation and recovery
    - Interface shutdown and enabling on CE switches
    - Switch reset
    - Active/standby firewall switchover

Intelligent O&M Solution of iMaster NCE-FabricInsight



- Based on Huawei-developed big data analytics platform, iMaster NCE-FabricInsight receives data from network devices through Telemetry and uses intelligent algorithms to analyze network data.
- The overall architecture of iMaster NCE-FabricInsight consists of three parts: network devices, iMaster NCE-FabricInsight collector, and iMaster NCE-FabricInsight analyzer.

HUAWEI

- Network device
  - Network devices include Huawei CE switches and NetEngine routers (NE routers for short). For details about the models and supported specifications, see the specification list of the corresponding version.
  - These devices use Telemetry to report performance metrics such as interface traffic based on Google Remote Procedure Call Protocol (gRPC). Devices are connected to FabricInsight as gRPC clients. Users can run commands to configure the Telemetry sampling function on the devices. The devices then proactively establish a gRPC connection with the target collector and send data to the collector. The current version supports the following sampling metrics: CPU and memory usage of devices and boards; number of sent and received bytes, number of discarded sent and received packets, and number of sent and received error packets of interfaces; number of congested bytes of queues; packet loss behavior data. For details about metrics and device models, see the product specification list.
- FabricInsight collector
  - The FabricInsight collector receives data reported by network devices through Telemetry, including performance metric data reported through gRPC. The collector then parses the metrics, combines and compresses the metric data, and reports the data to the analyzer.
- FabricInsight analyzer
  - The FabricInsight analyzer receives performance metrics from switches. In addition, the FabricInsight analyzer establishes dynamic baselines for some performance metrics based on the AI algorithm, detects exceptions, and displays the analysis result on the GUI.

One-stop Health Management of DCNs

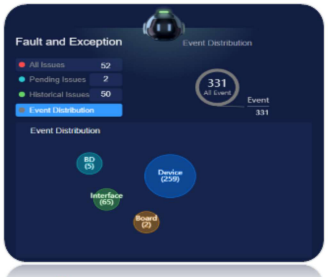
- The health function evaluates the overall health status of the current network based on identified network issues, helping users quickly and accurately identify and rectify faults.



Health check

Real-time health awareness

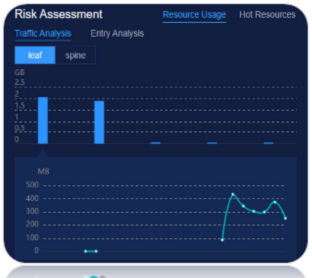
Comprehensively evaluates whether key health items on the entire network are normal based on the five-layer model and compares the health quality trend with that of the previous day.



Health diagnosis

Proactive exception notification

Proactively reports abnormal health items, checks root causes for common faults, and checks the distribution objects and fault propagation relationships for potential risks.



Prevention

Intelligent risk assessment

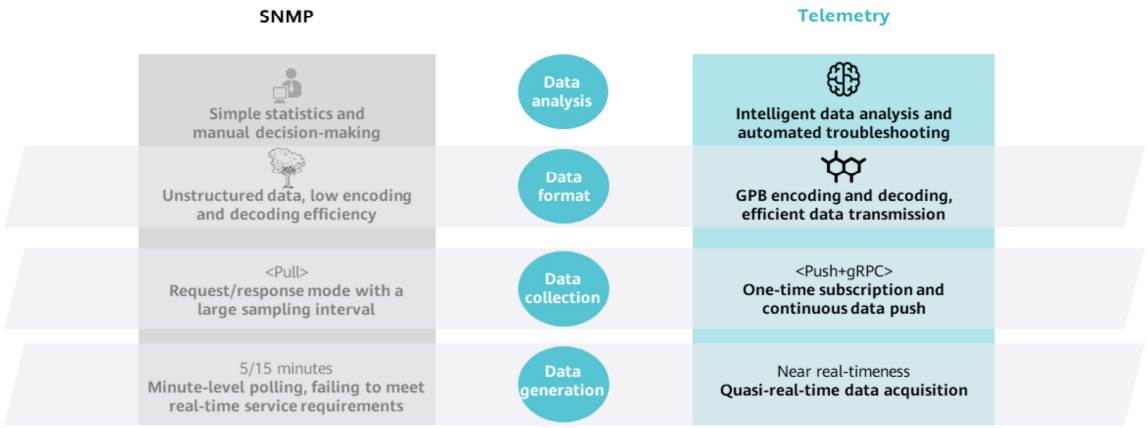
Analyzes the traffic and threshold-crossing entry capacity based on the current hotspot resources to detect risks in advance.

HUAWEI

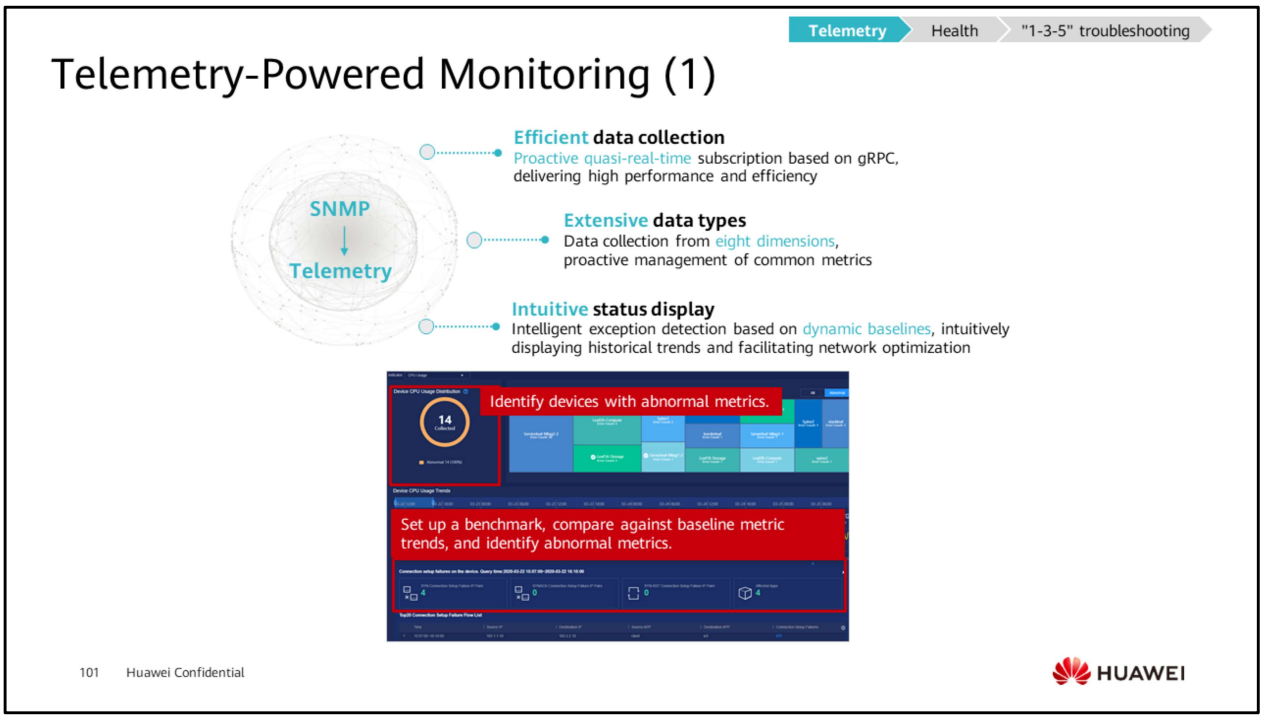
- The health function can identify the following issues:
  - Performance
    - Switch CPU threshold exceeded, switch memory threshold exceeded, service affected by switch port congestion, firewall CPU or IPv4 session threshold exceeded, abnormal switch CPU usage increasing, abnormal switch memory usage increasing, abnormal firewall CPU usage increasing, and abnormal firewall memory usage increasing
  - Capacity
    - Switch ARP entry threshold exceeded, switch ND entry threshold exceeded, switch MAC entry threshold exceeded, switch storage space threshold exceeded, switch ACL entry threshold exceeded, switch SFU forwarding performance insufficiency, number of routes received from a BGP peer exceeded the limit, abnormal switch ARP entry increasing, and abnormal switch ND entry increasing

- Status
  - Switch LPU exception, repeated switch LPU exception, switch MPU exception, switch SFU exception, switch fan exception, switch power supply exception, link interface status flapping, unidirectional link connectivity exception on the network side of a switch, switch port error-down, BGP peer relationship flapping, IP address conflict on VXLAN network access side, suspicious Layer 2 loop, repeated switch restarts, switch fault, switch disconnection, two master switches in M-LAG, VXLAN tunnel interruption, license file expired, OSPF router ID conflict, change of OSPF neighbor status, change of BGP peer status, stack fault, host IP address conflict, IP address conflict on network side, and access-side port blocked by STP
- Policy
  - ARP attack and ND attack
- Intent
  - Link port metric inconsistency, route loop on entire network, route blackhole on entire network, service reachability intent verification failure, and service isolation intent verification failure

Telemetry-Powered Real-Time Analysis



HUAWEI



- The Telemetry module displays metric data of devices, boards, chips, interfaces, queues, optical links, and RoCE. You can view and analyze these metrics of devices, boards, chips, interfaces, queues, optical links, and RoCE, which help users identify the status of the devices, boards, chips, interfaces, queues, optical links, and RoCE.

Telemetry-Powered Monitoring (2)

- Real-time monitoring of key metrics in eight dimensions, gaining deep insights into network status.

Measurement Object	Metric	Default Interval
Device	CPU usage and memory usage	1 minute
Board	CPU and memory usage, and FIB/MAC entry usage	1 minute
Chip	Rule usage, meter usage, counter usage, and slice/bank usage	1 minute
Interface	Numbers of received/sent packets and bytes, lost packets, error packets, broadcast packets, multicast packets, unicast packets, and ECN packets, as well as bandwidth usage	1 minute
	ARP and ND attack source tracing	
Queue	Buffer size	100 ms
	Numbers of received/sent PFC backpressure frames, PFC deadlock monitoring times, PFC deadlock recovery times, used headroom buffer, and used guaranteed buffer	1 minute
Optical link	Receive/transmit optical power, current, voltage, and temperature	1 minute
Packet loss behavior	Packet loss due to forwarding and congestion-triggered packet loss	1 minute
Entry	FIB/ARP/ND entry details	Dynamic subscription

HUAWEI

Health

Key assurance and network SLA

Display key metrics such as network-wide service objects, intent verification result, transmission delay, and packet loss rate.



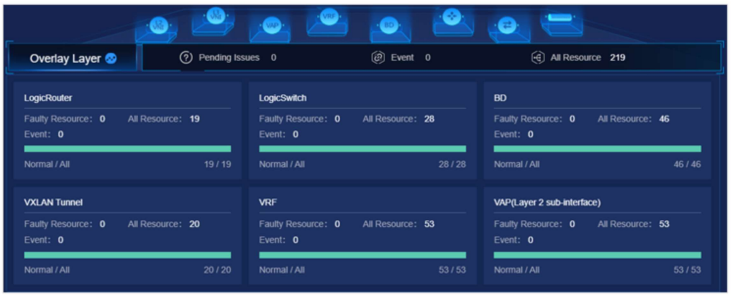
Network-wide resource status check

View the network-wide underlay/overlay resources and collected metric data, and compare the data collected yesterday with that collected today.



Five-layer health evaluation

Analyze the network health from dimensions of device, network, protocol, service, and overlay and determine the health status.

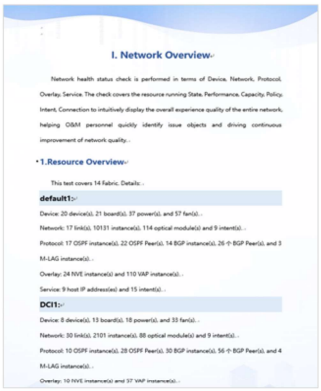


HUAWEI

Real-Time or Periodic Push of Health Reports, Providing References for Optimization

Network overview

Intuitively display the resource overview, client overview, and quality overview across the entire network.



Metric details

Identify network quality issues based on the five dimensions of the network health evaluation system.

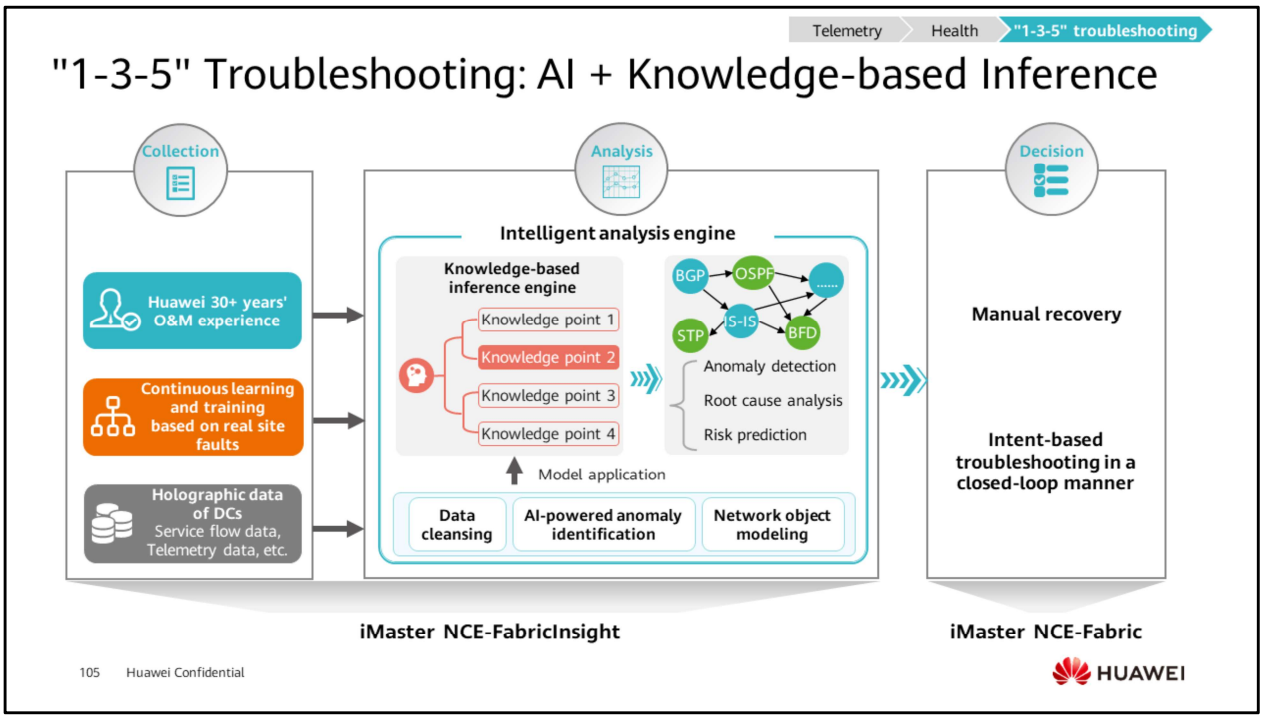


Report details

Display reports in multiple dimensions, and identify abnormal monitoring items, facilitating troubleshooting and network optimization.



HUAWEI



- "1-3-5" troubleshooting: Faults can be detected within 1 minute, located within 3 minutes, and rectified within 5 minutes.

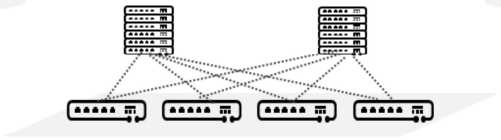
Fault Detection from Five Dimensions

Device

13 types of monitored objects | **36** metrics | **40 types** of issues

Service

6 types of monitored objects | **18** metrics | **19 types** of issues



Network

6 types of monitored objects | **18** metrics | **19 types** of issues

Overlay

3 types of monitored objects | **10** metrics | **5 types** of issues

Protocol

3 types of monitored objects | **8** metrics | **7 types** of issues

HUAWEI

Quiz

1. (Multiple-answer question) In a distributed network overlay scenario, if only one BGP process is created on the network, which of the following BGP EVPN deployment policies are supported? ( )

   A. Deploy OSPF on the underlay network and IBGP on the overlay network.

   B. Deploy OSPF on the underlay network and EBGP on the overlay network.

   C. Deploy EBGP on the underlay network and IBGP on the overlay network.

   D. Deploy EBGP on both underlay and overlay networks.

HUAWEI

1. AD

Quiz

2. (Multiple-answer question) Which of the following deployment modes can be used to ensure high reliability of border leaf nodes? ( )

A. Deploy border leaf nodes in active-active mode.

B. Connect border leaf nodes to core devices in dual-homed mode in the uplink direction.

C. Connect border leaf nodes to spine nodes in full mesh mode in the downlink direction.

D. Deploy a bypass policy between border leaf nodes.

HUAWEI

2. ABCD

Summary

- This document describes the network architecture, solution components, and typical scenarios of Huawei CloudFabric 3.0 Solution.
- This document describes the common methods, principles, and reference elements for typical DCN planning and design. On this basis, this document describes the DCN architecture design, including the logical architecture, physical architecture, DCI solution, key zone design, key attribute design, network security design, and O&M design.
- Upon completion of this document, you will be able to understand typical DCN design methods and plan and design DCNs.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Bearer WAN Architecture and Key Technologies



HUAWEI

Foreword

- The IP bearer WAN, which usually covers a country, is a data communication network that provides interconnection between multiple LANs or branch networks across regions.
- This course introduces the concepts, principles, and applications of the enterprise bearer WAN's typical architecture, bearer technologies, VPN services, traffic optimization, SLA, reliability, and network management and analysis. To introduce these key aspects, this course uses a large enterprise with three data centers in two cities and multiple branches in different regions as an example.

HUAWEI

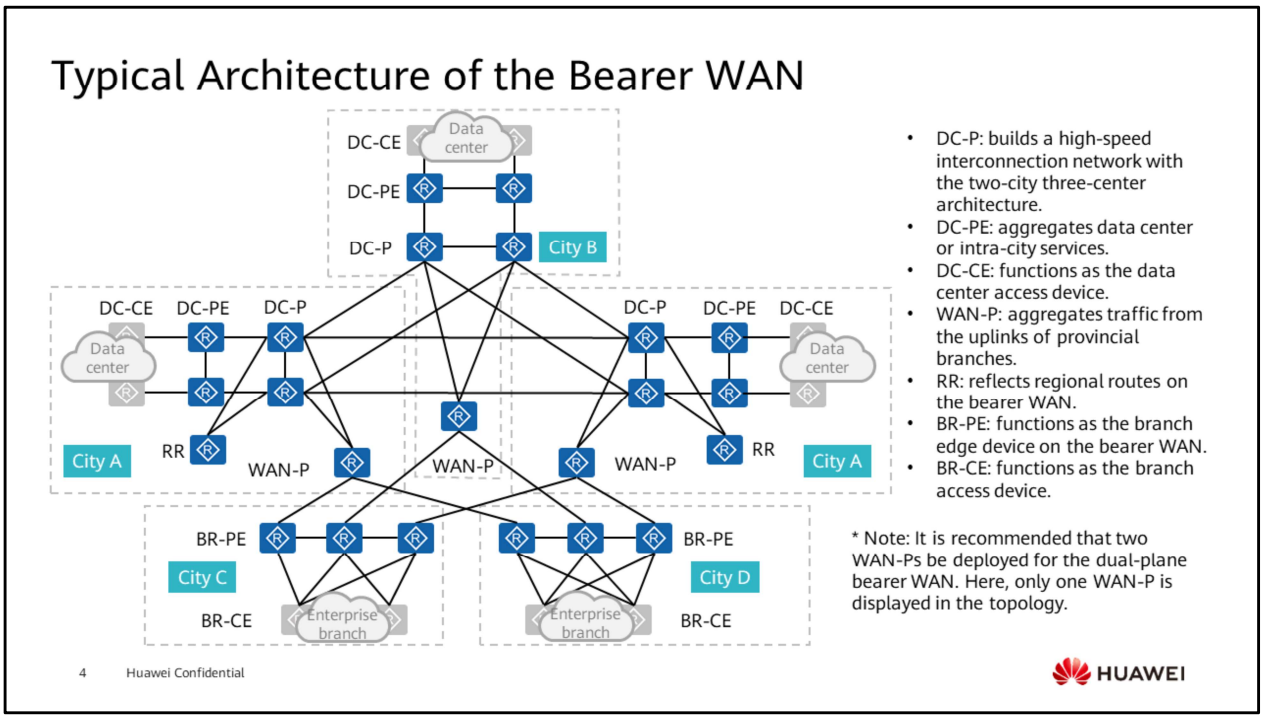
Objectives

- Upon completion of this course, you will be able to:
  - Describe the typical architecture of the bearer WAN.
  - Describe the basic concepts of Multiprotocol Label Switching (MPLS), Segment Routing-Multiprotocol Label Switching (SR-MPLS), and Segment Routing IPv6 (SRv6).
  - Describe the principles of WAN traffic optimization.
  - Describe the SLA model and packet processing flow.
  - Describe the network reliability design.
  - Describe the key protocols for network management and analysis.

Contents

1. **Bearer WAN Architecture**
2. Bearer WAN Basics
3. VPN Service
4. Network Traffic Optimization
5. SLA
6. Network Reliability
7. Network Management and O&M

HUAWEI

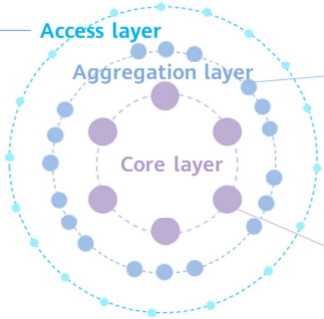


- The Ps for intra-city data center are directly connected using WDM or bare optical fibers. The link bandwidth can reach 10 Gbit/s. To reduce costs, consider connecting local data centers with remote data centers through carriers' MSTP links.

Typical Architecture of the Bearer WAN

- The typical architecture of a bearer WAN is divided into three layers: access layer, aggregation layer, and core layer.

Access layer (CE-PE):
- Provides access for data centers in different cities.
- Provides access for branches in different cities.
- Provides access for external services in different cities.

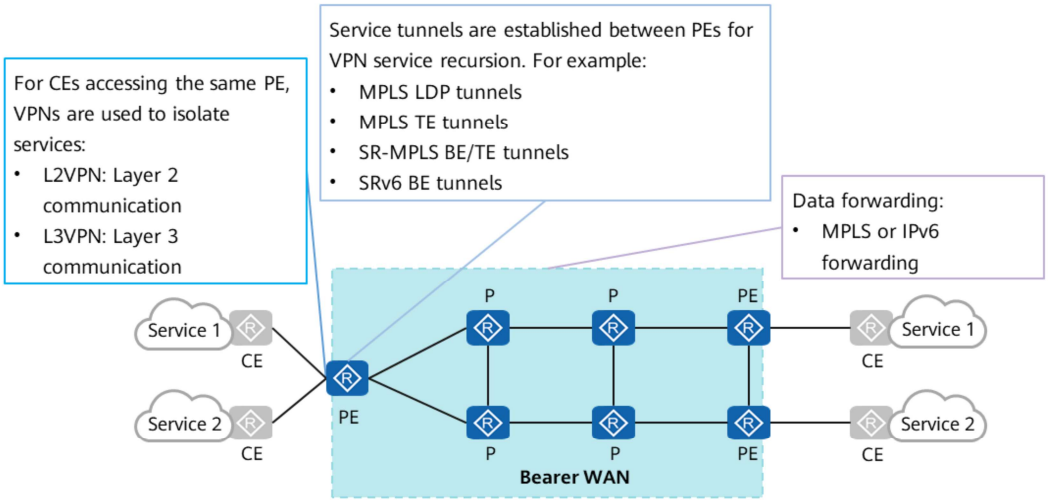


Aggregation layer (PE-P):
- Aggregates and transmits different types of services to the bearer WAN based on physical locations.

Core layer (P):
- Generally adopts the full-mesh + dual-plane architecture.
- Provides stable, reliable high-speed traffic forwarding between regions.

HUAWEI

Technologies Used by the Typical Bearer WAN Architecture



HUAWEI

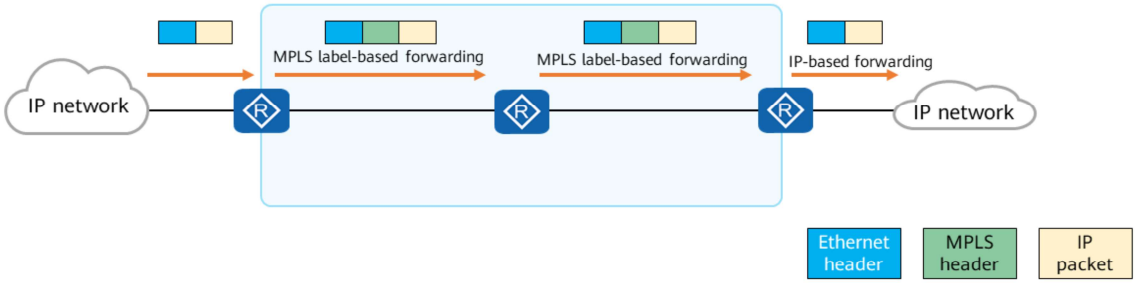
Contents

1. Bearer WAN Architecture
2. **Bearer WAN Basics**
3. VPN Service
4. Network Traffic Optimization
5. SLA
6. Network Reliability
7. Network Management and O&M

HUAWEI

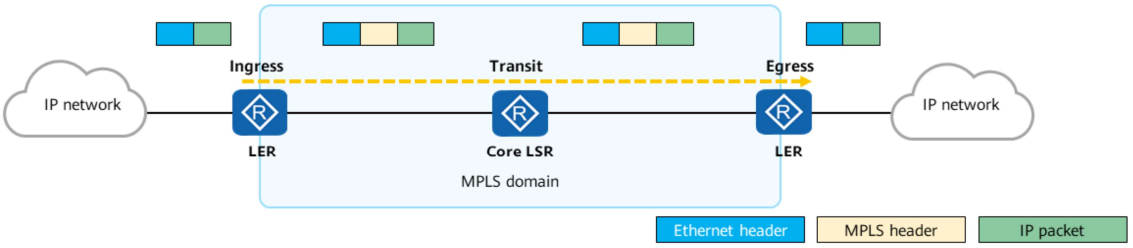
MPLS Overview

- MPLS is located between the data link layer and the network layer in the TCP/IP protocol stack and can provide services for all network layers.
- An MPLS header is added between a data-link-layer header and a network-layer header, and data can be forwarded quickly based on the MPLS header.



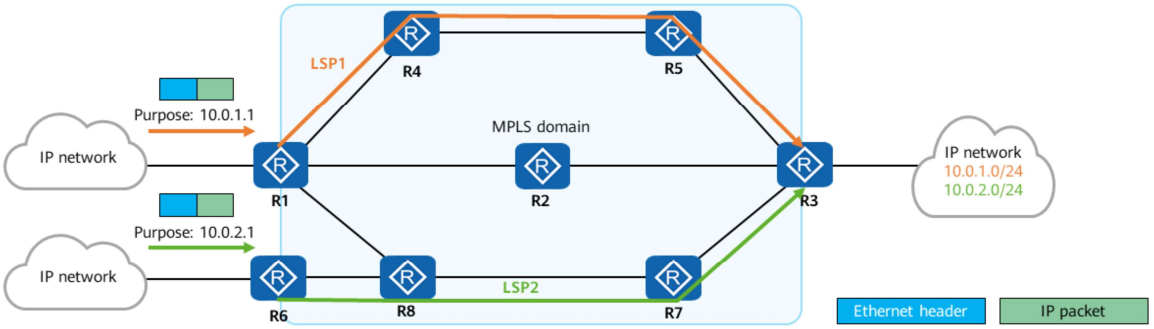
MPLS Terms (1)

- MPLS domain: consists of a series of consecutive network devices that run MPLS.
- Label switching router (LSR): a routing device, such as a router or switch, that runs MPLS. An LSR that resides at the edge of an MPLS domain and connects to a non-MPLS network is called a label edge router (LER). An LSR that resides inside an MPLS domain is called a core LSR.
- The path that MPLS packets take in an MPLS network is called a label switched path (LSP). The LSP is a unidirectional path that transmits traffic from the ingress to the egress.
- The start node of an LSP is called the ingress, an intermediate node of the LSP is called the transit node, and the end node of the LSP is called the egress. An LSP has one ingress, one egress, and zero, one, or multiple transit nodes.



MPLS Terms (2)

- Forwarding equivalence class (FEC): a set of packets with similar or identical characteristics and forwarded in the same way by LSRs. In traditional IP forwarding that uses the longest match algorithm, all packets that match the same route belong to the same FEC.
- An LSP is composed of an ingress LSR, an egress LSR, and a variable number of transit LSRs. Therefore, an LSP can be considered as an ordered set of these LSRs.
- An LSP must be established before a packet is forwarded; otherwise, the packet fails to traverse an MPLS domain.
- An LSP is a unidirectional path from the start point to the end point. If bidirectional data communication is required, an LSP for return traffic needs to be established between the two ends.



HUAWEI

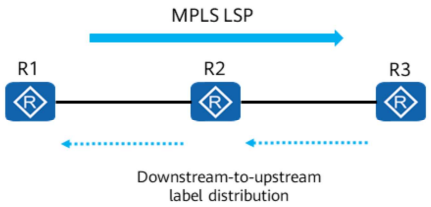
- For more information about MPLS, see HCIP-Datacom-Advanced Routing & Switching Technology.

MPLS Terms (3)

- LSPs can be statically configured or dynamically established.
- There are two common protocols for dynamically establishing LSPs: LDP and Resource Reservation Protocol-Traffic Engineering (RSVP-TE).

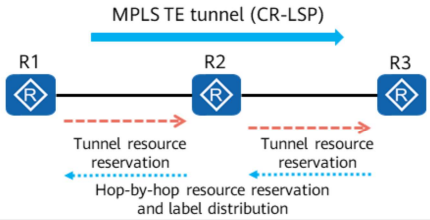
LDP: used to establish common LSPs

- LDP distributes labels from downstream routers to upstream routers based on the routing table to set up common LSPs.
- A common LSP is usually the shortest path calculated by an IGP, which does not factor in aspects such as bandwidth, tunnel protection, and traffic optimization.



RSVP-TE: used to establish CR-LSPs

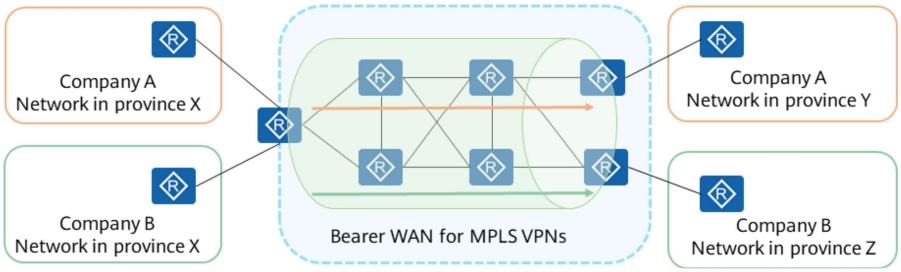
- RSVP-TE establishes tunnels by applying for and reserving tunnel resources end to end.
- An LSP that is set up based on bandwidth or path constraints is called a constraint-based routed label switched path (CR-LSP).



HUAWEI

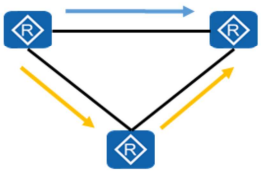
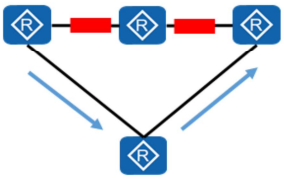
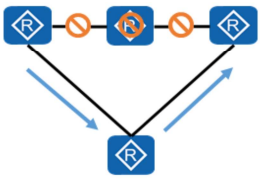
MPLS LDP Overview

- LDP is a control protocol of MPLS and provides functions such as FEC classification, label distribution, and LSP establishment and maintenance. LDP defines the messages used in label distribution as well as the message processing procedures.
- LDP is easy to configure and maintain and is widely used to create LSPs in BGP/MPLS IP VPN scenarios. As shown in the figure, the carrier builds a bearer WAN for MPLS VPNs to provide inter-provincial L3VPN services for customers.
- MPLS LDP LSPs are established based on the shortest IP paths and do not support the planning of tunnel forwarding paths.



MPLS TE Overview

- MPLS TE, as its name suggests, is a combination of MPLS and TE. It provides functions such as path planning, traffic optimization, and fault protection for MPLS VPN services.
- Compared with MPLS LDP, MPLS TE enhances VPN traffic control and protection.

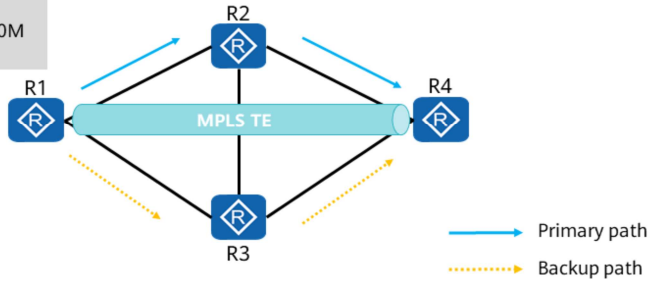
Path planning	Traffic optimization	Fault protection
		
• Different paths are planned for different services.	• Traffic is evenly distributed to idle links when unbalanced traffic distribution occurs due to major events, such as live broadcast.	• Fast protection switching is performed in the case of a device or link fault.

HUAWEI

MPLS TE Tunnel

- MPLS TE often associates multiple LSPs with a virtual tunnel interface, and such a group of LSPs is called an MPLS TE tunnel.
- An MPLS TE tunnel provides SLA assurance, but requires complex configuration and manual planning.

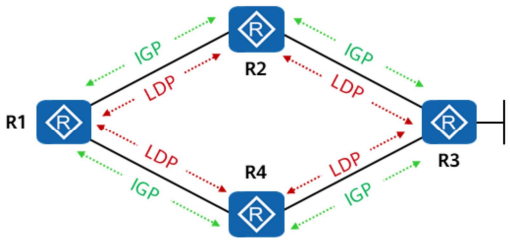
```
[R1] interface tunnel1
[R1-Tunnel1] ip address ...
[R1-Tunnel1] tunnel-protocol mpls te
[R1-Tunnel1] destination R4
[R1-Tunnel1] mpls te bandwidth 50M
...
```



HUAWEI

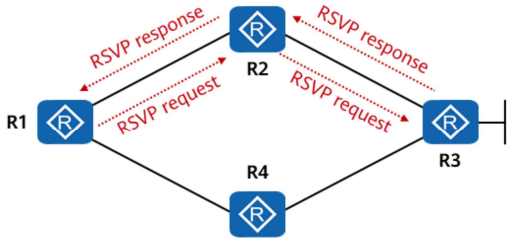
Issues with MPLS LDP and RSVP-TE

MPLS LDP



- LDP itself does not have the path computation capability and requires an IGP for path computation.
- Both the IGP and LDP need to be deployed for the control plane, and devices need to exchange a large number of packets to maintain neighbor relationships and path states, wasting link bandwidth and device resources.
- If LDP-IGP synchronization is not achieved, data forwarding may fail.

RSVP-TE



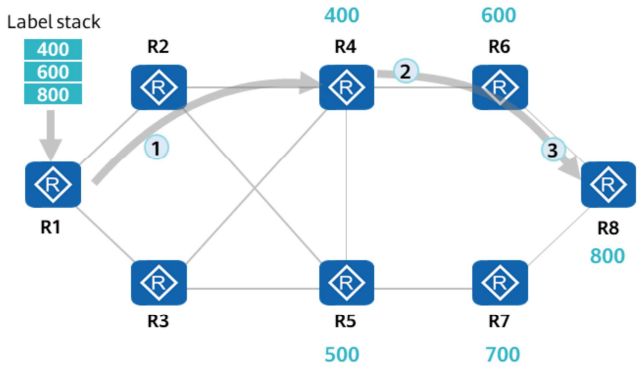
- RSVP-TE is complex to configure and does not support load balancing.
- To implement TE, devices need to exchange a large number of RSVP packets to maintain neighbor relationships and path states, wasting link bandwidth and device resources.
- RSVP-TE uses a distributed architecture, so that each device only knows its own state and needs to exchange signaling packets with other devices.

How can we simplify the control plane?

HUAWEI

SR-MPLS Overview

- SR is designed to forward data packets on a network using the source routing model.
- SR-MPLS, as its name suggests, is SR based on MPLS label forwarding.



- Source routing:
  - The source node selects a path and pushes an ordered label stack into the packet.
  - Other nodes on the network forward the packet according to the label stack encapsulated into the packet.
- SR has the following characteristics:
  - Extends existing protocols (e.g. IGP) to facilitate network evolution.
  - Supports both centralized controller-based control and distributed forwarder-based control, providing a balance between the two control modes.
  - Enables networks to quickly interact with upper-layer applications through the source routing technology.

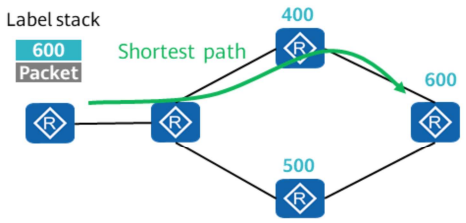
HUAWEI

SR-MPLS BE and SR-MPLS TE

- The SR-MPLS tunneling technology can be implemented in either SR-MPLS BE or SR-MPLS TE mode.

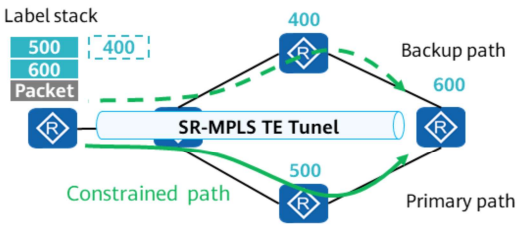
SR-MPLS BE

- Forwarding path: Similar to an MPLS LDP LSP, an SR-MPLS BE LSP is calculated using the IGP shortest path first (SPF) algorithm. An SR-MPLS BE LSP has only one label layer (destination node).
- In the production environment, SR-MPLS BE is generally used as the DR solution for SR-MPLS TE. For example, if a controller fault causes a tunnel delivery failure, the IGP can be used to generate forwarding tunnels.



SR-MPLS TE

- Forwarding path: An SR-MPLS TE path is created using SR based on TE constraints. An SR-MPLS TE tunnel generally uses multiple layers of labels to implement path control and supports primary and backup paths.
- SR-MPLS TE is usually used with a controller. After the controller globally computes a path, it delivers a label stack to the corresponding ingress.



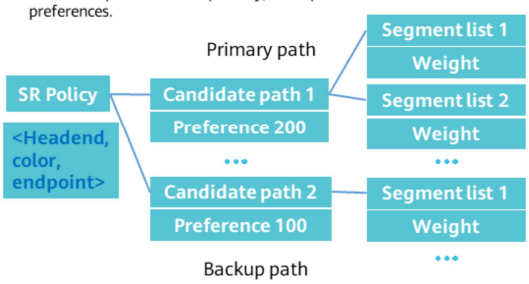
HUAWEI

SR-MPLS Policy

- The SR-MPLS Policy, also called the SR-MPLS TE Policy, is one of the mainstream SR-MPLS implementation modes.
- As defined in the corresponding RFC, an SR Policy is identified by <headend, color, endpoint> and contains multiple candidate paths.

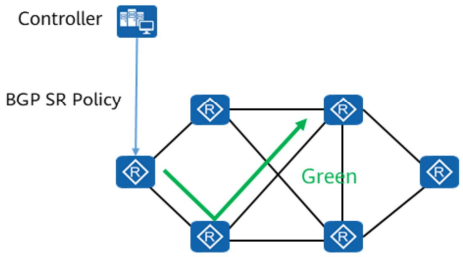
SR Policy model

- A candidate path can contain multiple segment lists, and load balancing can be implemented among these segment lists based on weights.
- Candidate paths work in primary/backup mode based on their preferences.



SR Policy application scenario

- The controller collects information such as the global topology, network bandwidth, and link delay. After computing an SR Policy path based on service requirements, the controller uses BGP to deliver the path to the ingress.

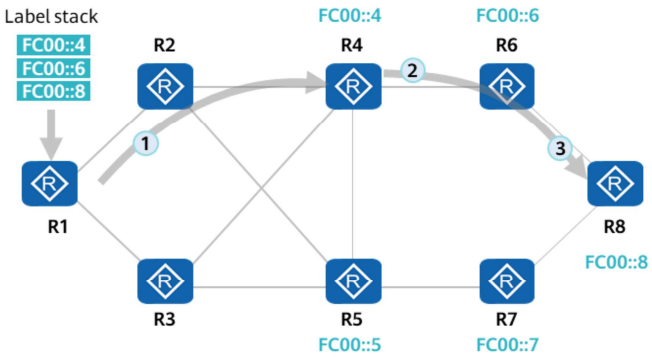


HUAWEI

- Based on MPLS and IPv6 forwarding technologies, SR Policies can be classified into SR-MPLS and SRv6 Policies.

SRv6 Overview

- SRv6 is designed to forward data packets on an IPv6 network using the source routing model.
- Both SRv6 and SR-MPLS comply with the SR architecture. Their main difference lies in data plane instructions. The former is based on the IPv6 network and uses IPv6 addresses as instructions. In contrast, the latter is based on the MPLS network and uses MPLS labels as instructions.



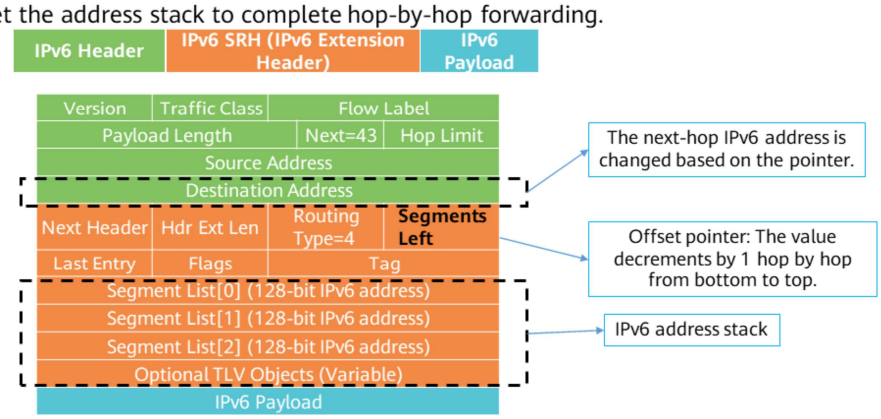
Source routing:

- The source node selects a path and pushes an ordered label stack into the packet.
- Other nodes on the network forward the packet according to the label stack encapsulated into the packet.

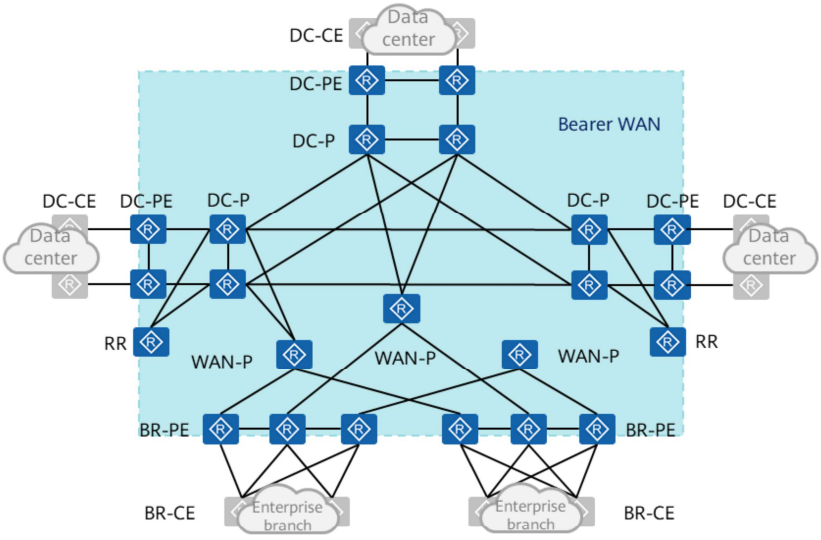
HUAWEI

SRv6 Extension Header

- SRv6 adds a segment routing header (SRH) to IPv6 packets. The SRH contains an explicit IPv6 address stack. During the forwarding process, SRv6 nodes continuously update the destination address and offset the address stack to complete hop-by-hop forwarding.



Summary: WAN Bearer Technologies



Enterprise bearer WANs can be roughly classified into two types: MPLS network and IPv6 network.

- MPLS network:
  - MPLS LDP
  - MPLS TE
  - SR-MPLS
- IPv6 network:
  - SRv6

HUAWEI

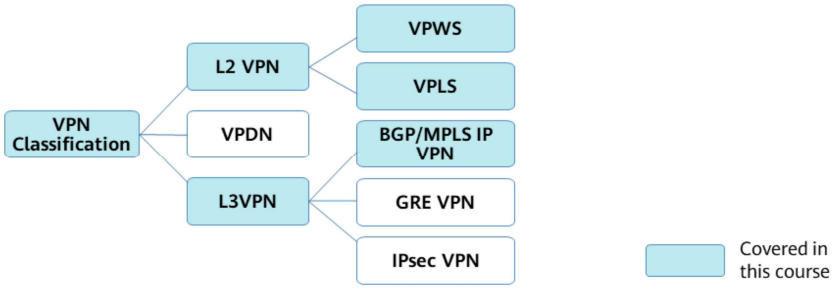
Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. **VPN Service**
   - WAN VPN Overview
     - Tunnel Management Overview
4. Network Traffic Optimization
5. SLA
6. Network Reliability
7. Network Management and O&M

HUAWEI

VPN Classification

- The VPN technology is widely used as a virtual private tunneling technology. VPN can be classified into various types from different perspectives. For example, VPN can be classified into Layer 3 VPN (L3VPN), Layer 2 VPN (L2VPN), and Virtual Private Dial-up Network (VPDN) by implementation layer.
- VPWS, VPLS, and BGP/MPLS IP VPN are more widely used on bearer WANs.
- GRE VPN and IPsec VPN, which are mainly used on the Internet, are beyond the scope of this course.



HUAWEI

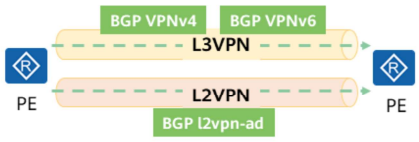
- Traditional switching networks, such as asynchronous transfer mode (ATM) and frame relay (FR) networks, are integrated with IP or MPLS networks. As a result, Layer 2 virtual private network (L2VPN) emerges. L2VPN includes Virtual Pseudo Wire Service (VPWS) and Virtual Private LAN Service (VPLS):
  - VPWS is a P2P L2VPN technology that emulates the basic behaviors and characteristics of services such as ATM and frame relay.
  - VPLS provides P2MP L2VPN services so that sites are connected as if they were on the same LAN.
- Virtual Private Dial-up Network (VPDN) is a virtual private network constructed on the public network. It uses a dedicated network encryption communication protocol to provide access services for international organizations and mobile workforce of enterprises. There are multiple VPDN tunneling protocols, among which Layer Two Tunneling Protocol (L2TP) is the most widely used. Strictly speaking, L2TP is also a type of L2VPN, but its network structure and protocol design are quite different from those of other types of L2VPN. In addition, L2TP uses the dial-up mode. Therefore, L2TP is classified as VPDN.
- L3VPN is also called Virtual Private Routing Network (VPRN), including RFC 2547-based BGP/MPLS IP VPN as well as IPsec VPN and GRE VPN carried over IPsec or GRE tunnels.

WAN VPN Service Overview

- An enterprise establishes a bearer WAN to provide wide-area interconnection for its internal and external services, such as production, office, external connection, and test services. These services are logically isolated but share the same physical network resources. Therefore, these services are called WAN VPN services.
- WAN VPN can be classified into L2VPN and L3VPN.

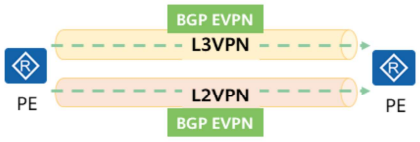
Traditional L2VPN and L3VPN technologies

- Traditional L2VPN includes VPLS and VPWS, which can use LDP or BGP to establish virtual links.
- Traditional L3VPN uses VPN instances to isolate services and uses BGP VPNv4/v6 to transmit Layer 3 information.



EVPN

- EVPN provides both Layer 2 and Layer 3 capabilities and functions as a control plane protocol to transmit Layer 2 and Layer 3 information (MAC/IP addresses). EVPN can be used together with the traditional VPN technology.
- EVPN supports EVPN VPLS, EVPN VPWS, EVPN L3VPN, etc.

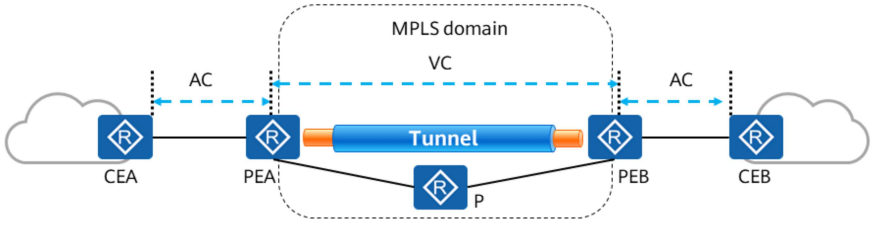


HUAWEI

- A traditional L2VPN does not have any control plane and does not transmit service route information (MAC addresses). It uses BGP as the signaling protocol to establish VCs.
- For details about VPN classification, see the book SRv6 Network Programming: Ushering in a New Era of IP Networks.

Traditional WAN L2VPN Overview

- The traditional WAN L2VPN is based on the MPLS network. VPWS provides a point-to-point Layer 2 network, and VPLS provides a point-to-multipoint Layer 2 network.
- The basic MPLS L2VPN architecture is composed of the attachment circuit (AC), VC, and tunnel.
  - AC: independent physical or virtual circuit connecting a CE and a PE. An AC interface can be either a physical or logical interface.
  - VC: logical connection between two PEs. A VC is established using a signaling protocol, such as BGP AD.
  - Tunnel: used to transparently transmit service data. Typical tunnels include MPLS LDP tunnels and MPLS TE tunnels.
- VPLS and VPWS are widely used on carrier networks to provide MPLS Layer 2 private line services for enterprises.

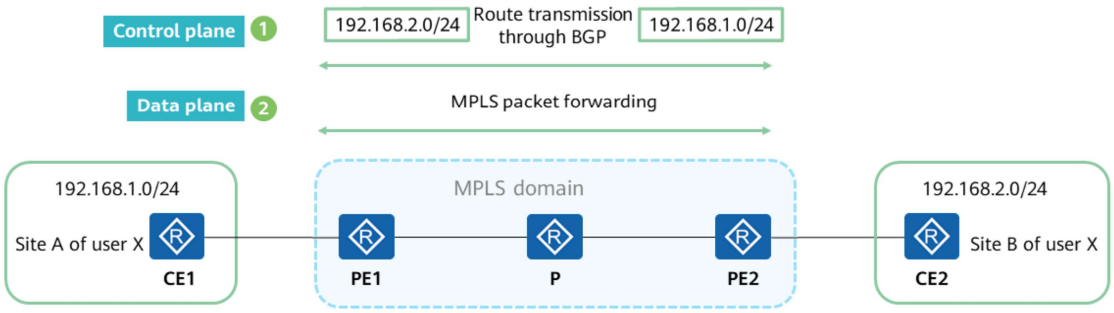


HUAWEI

- VCs are also called pseudo wires (PWs) in some documents.

Traditional WAN L3VPN Overview

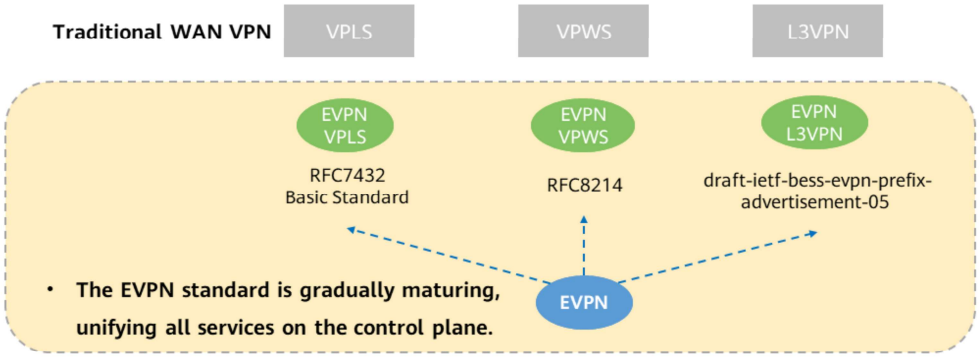
- Traditional WAN L3VPN generally refers to BGP/MPLS L3VPN.
- It distinguishes the control plane from the data plane. The control plane uses MP-BGP to advertise VPN routes, and the data plane uses MPLS LSPs to forward VPN packets.
- BGP/MPLS L3VPN is widely used on carrier networks to provide MPLS Layer 3 private line services for enterprises.



HUAWEI

WAN EVPN Overview

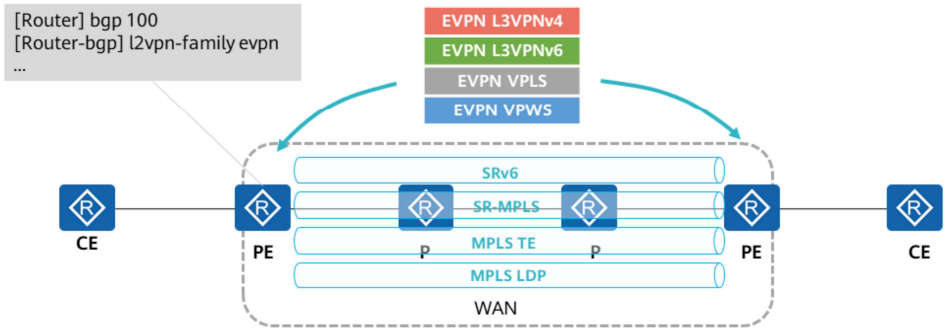
- EVPN was initially designed as an L2VPN technology based on BGP extensions. With the development of protocol extensions, EVPN can also support L3VPN now.
- EVPN can well serve as the control plane protocol for WAN VPN. It can be used with traditional VPN technologies to provide EVPN VPLS, EVPN VPWS, and EVPN L3VPN.



HUAWEI

WAN EVPN Application

- On a WAN, EVPN can be used with multiple tunneling technologies to support multiple application scenarios.
- EVPN, as a control plane protocol, can work with MPLS LDP, MPLS TE, SR-MPLS, and SRv6 tunnels, as shown in the figure.
- On Huawei devices, EVPN L2VPN and L3VPN services share the same address family.



HUAWEI

Summary: WAN VPN

- The VPN technology is widely used in various enterprise scenarios. VPNs can be built over either the Internet or a private network.
- L2VPN (VPLS and VPWS), L3VPN (BGP/MPLS IP VPN), and L2/L3 EVPN can be deployed on a bearer WAN built by an enterprise.
- EVPN, as a control plane protocol, can work with different bearer technologies (such as MPLS LDP, MPLS TE, SR-MPLS, and SRv6) to provide integrated and unified VPN services for enterprises.
- When multiple tunneling technologies are deployed on an enterprise bearer WAN, VPNs must recurse to tunnels based on tunnel policies.

HUAWEI

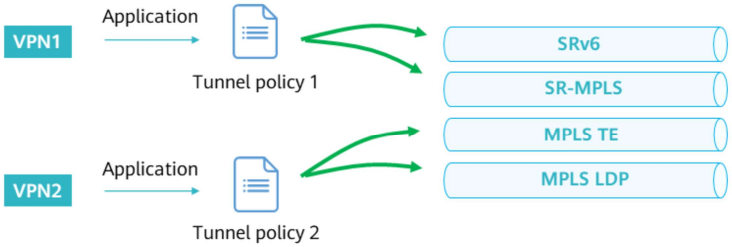
Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. **VPN Service**
   - WAN VPN Overview
   - Tunnel Management Overview
4. Network Traffic Optimization
5. SLA
6. Network Reliability
7. Network Management and O&M

HUAWEI

Tunnel Management

- Huawei devices use a tunnel management (TNLM) module to manage tunnels. It selects a certain tunnel for an application according to specific configurations and notifies the application of the tunnel's status.
- Common VPN tunnels include LSPs (MPLS LDP), MPLS TE tunnels, GRE tunnels, SR-MPLS Policies, and SRv6 Policies.
- Tunnel management configuration includes two parts: configuring a tunnel policy and applying a tunnel policy to a VPN.



HUAWEI

- GRE: GRE can be applied to both L2VPN and L3VPN. Generally, the bearer WAN for MPLS VPN uses LSPs as public network tunnels. If the bearer WAN (P devices) has only IP functions but not MPLS functions, and the PEs at the network edge have MPLS functions, the LSPs cannot be used as public network tunnels. In this case, GRE tunnels can be used to replace LSPs to provide L3VPN or L2VPN solutions on the bearer WAN.
- SR-MPLS Policy: a type of SR-MPLS tunnel.
- SRv6 Policy: a type of SRv6 tunnel.
- You can configure tunnel policies or tunnel policy selectors for tunnel management. This course uses tunnel policy configuration as an example. Tunnel policy selectors apply to inter-AS VPN scenarios. For details, see the product documentation for NetEngine products.

Configuring a Tunnel Policy

- A tunnel policy determines the types and sequence of tunnels to be selected.
- By default, a VPN service recurses only to one LSP. If multiple LSPs are available, a tunnel policy can be used for load balancing among these LSPs.
- In this example, to implement load balancing between MPLS LDP LSPs and TE tunnels, you need to configure a tunnel policy for the VPN and apply the tunnel policy to the VPN. The tunnel policy policy1 requires tunnels to be selected in the sequence of first CR-LSPs and then LSPs, and the number of tunnels for load balancing is 2. The system preferentially selects two CR-LSPs. If only one or no CR-LSP is available on the network, the system selects one or two LSPs, respectively, for service transmission. In the scenario where only one CR-LSP is available, it works together with the selected LSP.

```
[PE1] tunnel-policy policy1
[PE1-tunnel-policy-policy1] tunnel select-seq cr-lsp lsp load-balance-number 2
```



HUAWEI

- The configuration of tunnel policy parameters involves many details. For example, CR-LSP-based tunnels include RSVP-TE tunnels and SR-MPLS TE tunnels. The system determines the priorities of these tunnels based on their up time. For details, see "VPN Tunnel Management Configuration" in the product documentation for Huawei NetEngine routers.

Applying a Tunnel Policy to a VPN

- After being configured, a tunnel policy needs to be applied to a VPN. The mode in which a tunnel policy is applied to a VPN varies according to the VPN type.
- This example shows how to apply a tunnel policy to an L3VPNv4 instance.

```
[PE1] ip vpn-instance vpn1
[PE1-vpn-instance-vpnb] ipv4-family
[PE1-vpn-instance-vpnb-af-ipv4] tnl-policy policy1
```

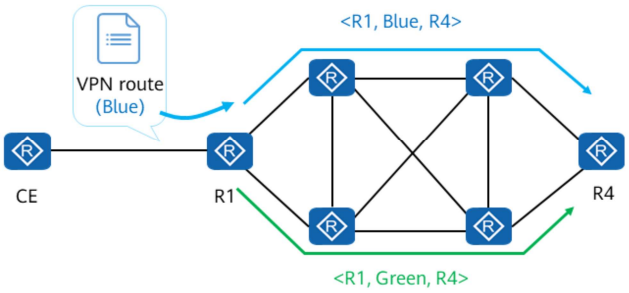


HUAWEI

- For the application of other types of VPN, such as VPNv6, L2VPN, and EVPN, see the product documentation for NetEngine routers.

Extension: SR Policy-based Traffic Diversion

- Tunnel policies are designed by Huawei for VPN service recursion in the MPLS era. They effectively decouple tunnel establishment from tunnel selection. In this way, the traffic of a VPN can be directed to multiple tunnels for load balancing. After the SR technology is introduced, the implementation mode changes. SR Policies integrate tunnel establishment (SR forwarding path) and tunnel policies (color-based traffic diversion by default). SR Policies cannot be selected together with other types of tunnels.
- An SR Policy is identified by <headend, color, endpoint> and can contain multiple forwarding paths. A VPN service selects an SR Policy based on the color attribute.



- There are two SR Policies (blue and green) between R1 and R4.
- R1 queries its VPN routing table after receiving traffic from the CE.
- If the color of the corresponding VPN route is Blue, R1 selects an SR Policy with the same color for route recursion.

HUAWEI

- For details, see "SR Policy" in 2. Terminology in RFC 8402.
- SR Policy traffic diversion can be based on the binding SID, color, and DSCP value. Details are not provided here.

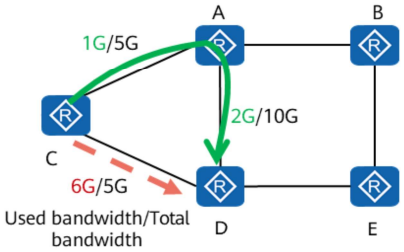
Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. VPN Service
4. **Network Traffic Optimization**
5. SLA
6. Network Reliability
7. Network Management and O&M

HUAWEI

Network Congestion Background

Drawback of fixed bandwidth-based route selection and related solution

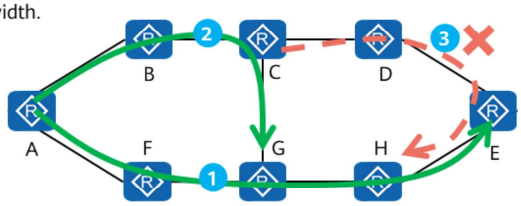


Used bandwidth/Total bandwidth

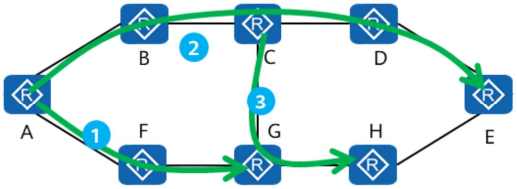
The network computes forwarding paths based on bandwidth. The link from Router C to Router D is the shortest forwarding path. The rate of service traffic from Router C to Router D exceeds the link bandwidth, and packet loss occurs. Although other links are idle, the algorithm still selects the shortest path for traffic forwarding. From a global perspective, the optimal traffic forwarding path is C -> A -> D in this situation.

Drawback of sequenced tunnel establishment and related solution

Tunnels are established in the following sequence: A-E > A-G > C-H. Tunnel C-H, however, fails to be established due to insufficient bandwidth.



Global path computation for optimal tunnel path adjustment:



HUAWEI

Network Traffic Optimization Overview

- Network traffic optimization is to perform global analysis on network congestion, obtain the path computation result based on a proper optimization policy (ensuring the SLA of critical services), and apply the computation result to the network for congestion elimination.
- Network traffic optimization can be divided into three phases: network information collection, path computation for network traffic optimization, and optimization result delivery.

1. Network information collection

The controller collects global network information, including:

- Network topology
- Network bandwidth
- Link delay
- Network traffic and other information

2. Path computation for network traffic optimization

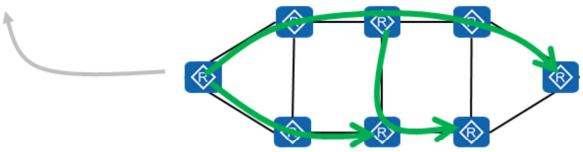
The controller computes paths based on the optimization target. Optimization targets include:

- Least path cost
- Shortest path delay
- Maximum link bandwidth utilization
- ...

3. Optimization result delivery

The controller delivers the computation result to network devices in any of the following modes:

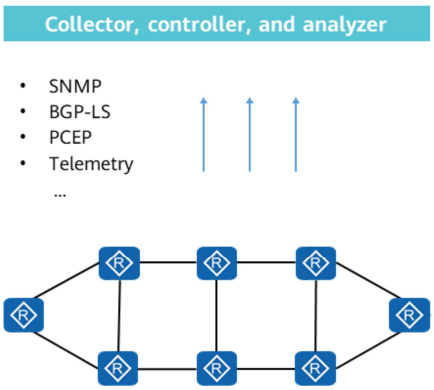
- NETCONF
- PCEP
- BGP SR Policy
- ...



HUAWEI

Network Information Collection

- The collection of network information, including the network topology, interface bandwidth, link delay, and traffic statistics, is the prerequisite for network traffic optimization.



Network topology and interface bandwidth collection

- In the industry, SNMP is generally used to collect basic network topology and device information.
- BGP-LS is used to collect IGP and TE topology information (including interface bandwidth).
- PCEP and BGP SR Policy are used to collect TE tunnel information.

Link delay collection

- The link delay is collected using TWAMP, flooded in the IGP domain, and then reported to the controller through BGP-LS.

Traffic statistics collection

- To determine bandwidth sufficiency and perform traffic optimization, the controller needs to collect interface and tunnel traffic statistics in real time.
- Mainstream traffic statistics collection technologies include SNMP, telemetry, and NetStream.

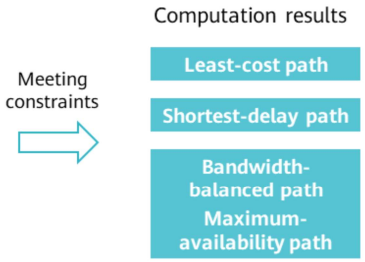
HUAWEI

- Different collection protocols may be used in different solutions. For example, PCEP is used to collect TE tunnel information on Huawei MPLS networks, and BGP SR Policy is used to collect TE tunnel information on SRv6 networks.

Network Optimization Computation

- Network optimization computation refers to the computation of global or local optimal paths based on service requirements through the corresponding algorithms.
- When computing paths, the controller needs to ensure that computed paths meet the related constraints. If the constraints cannot be met, the controller retains the original paths. The following table lists some constraints.

Constraint	Description
Priority	This constraint specifies the priorities of different types of tunnels and enables a tunnel with a higher priority to preempt the bandwidth resources of a tunnel with a lower priority.
Bandwidth	This constraint requires paths to be computed based on tunnel bandwidth requirements.
Hop count	This constraint requires paths to be computed based on hop count requirements. For example, the path length of an SR-TE tunnel is limited by the maximum stack depth (MSD) of the ingress node.
Explicit path	This constraint requires paths to be computed in either strict or loose mode. You can specify the links or nodes that are to be included or excluded.
Delay threshold	This constraint requires paths to be computed within the threshold range specified for path computation.
Affinity	This constraint supports the include-all, include-any, and exclude modes.

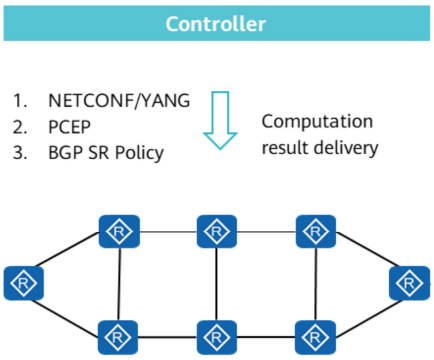


HUAWEI

- Bandwidth-balanced path: path with more remaining bandwidth among all paths that meet the constraints and have the same cost.
- Maximum-availability path: path with the maximum availability among all paths that meet the constraints.

Optimization Result Delivery

- After the controller computes a network path, you can choose whether to apply the computation result to the network. There are multiple implementation modes:



1 NETCONF/YANG

- The controller delivers the computation result to network devices as configurations.
- The YANG model is standardized and provides good compatibility with different vendors.

2 PCEP

- The controller delivers PCEP messages to create or update LSPs.
- The PCEP standard does not define a tunnel model, and vendor-specific protocols cannot interoperate with each other.

3 BGP SR Policy

- The controller uses BGP extensions to deliver tunnels.
- The RFC defines the tunnel model and data packet structure in a unified manner, facilitating product interoperability between different vendors.

HUAWEI

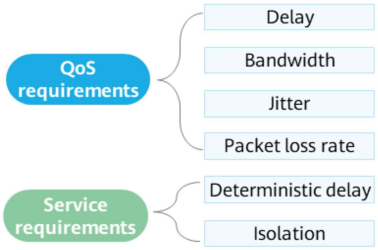
Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. VPN Service
4. Network Traffic Optimization
5. **SLA**
6. Network Reliability
7. Network Management and O&M

HUAWEI

New Requirements for Bearer Network SLA Assurance

- An SLA is a formal commitment between a service provider and a customer. In the WAN service field, in addition to basic connectivity requirements, the SLA also focuses on deterministic delay, bandwidth, reliability, and isolation (security).
- Generally, a bearer path carries the traffic of multiple types of services. When different types of service traffic are transmitted on the same path, differentiated bearer needs to be provided based on SLA requirements. Traditional QoS uses statistical multiplexing to set different priorities for specific services to ensure smooth experience of high-priority services. However, QoS falls short of meeting the isolation and deterministic delay requirements.
- Network slicing can divide a bearer network into virtual networks of different service levels and provide dedicated logical channels for services with high quality and security requirements.

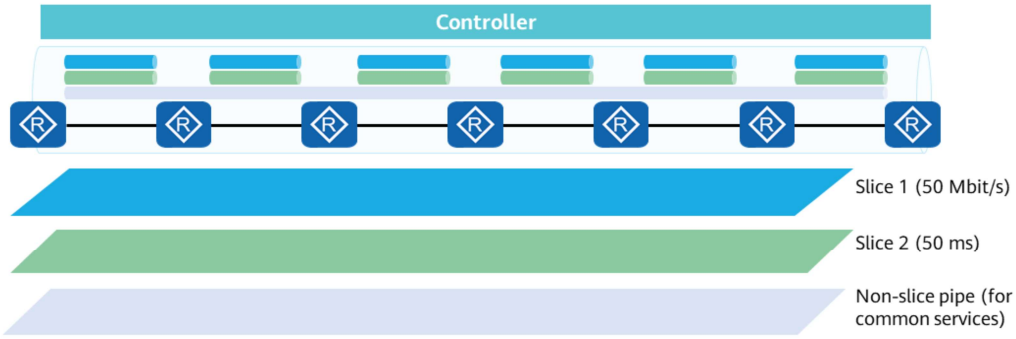


- For example, fund settlement services between financial enterprises require high security and stable delay; the file transfer service within an enterprise requires high bandwidth; and enterprise voice services require low jitter.

HUAWEI

Slice-based Bearer Network: One Network for Multiple Purposes (Carrying Multiple Services), Lowering Costs

- The IP network uses statistical multiplexing to greatly improve network utilization and reduce per-bit transmission costs. However, statistical multiplexing brings uncertainty to the quality assurance levels of different services. Moreover, it is inappropriate to prepare resources based on the highest SLA requirements to meet the requirements of all types of private lines and customers. The converged bearer network needs to balance multi-service isolation and statistical multiplexing to meet the SLA requirements of each service.
- Slice resource reservation technologies, such as FlexE, channelized sub-interfaces, and QoS queues, can be used to direct services to respective service slices. These slices are isolated from each other and do not affect each other, providing different SLA levels.



HUAWEI

QoS Overview

- QoS provides differentiated service quality for different applications.
- Generally, QoS provides three service models: best-effort service, integrated service (IntServ), and differentiated service (DiffServ).
- DiffServ is the most widely used QoS model on IP networks.



Best-effort service

- Applications can send any number of packets at any time.
- The network tries its best to send packets and does not provide delay or reliability guarantee.

IntServ

- An application needs to apply for specific services before sending packets.
- The network reserves network resources for the application using a protocol such as RSVP.

DiffServ

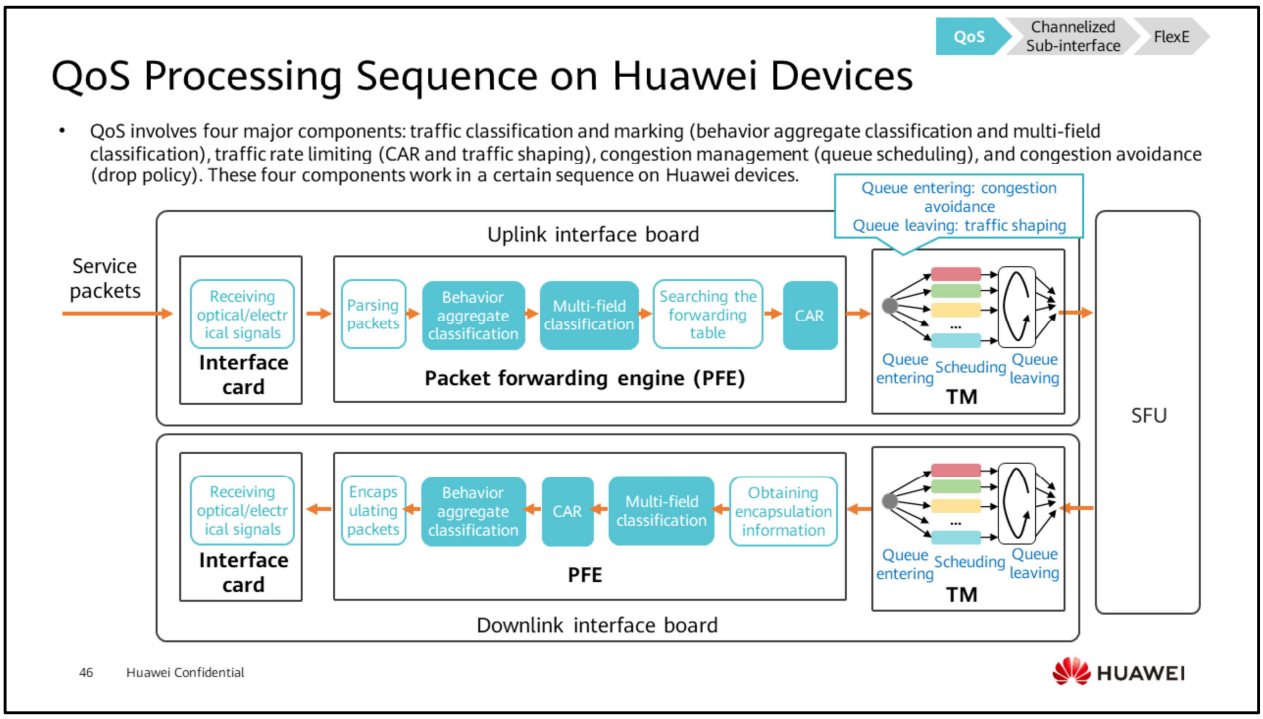
- Traffic on the network is divided into multiple classes. A processing behavior is defined for each class, so that each class has a different forwarding priority, packet loss rate, and delay.
- It is most widely used on IP networks.

HUAWEI

DiffServ Model-based QoS Components

- DiffServ model-based QoS consists of four components: traffic classification and marking, traffic policing and shaping, congestion management, and congestion avoidance.
  - Traffic classification and marking: Dividing data packets into different classes or configuring different priorities for packets is the prerequisite for implementing differentiated services. This component classifies data packets into different types, with each type of traffic being a traffic class. Traffic classification does not change the original data packets. In comparison, marking, which sets different priorities for data packets, changes the original data packets.
  - Rate limiting (traffic policing and shaping): This component limits the rate of service traffic. It does this by discarding excess traffic when the service traffic exceeds the rate limit. Traffic policing controls the traffic receiving rate, and traffic shaping controls the traffic sending rate.
  - Congestion management: This component buffers packets in queues upon network congestion and determines the forwarding order using a specific scheduling algorithm.
  - Congestion avoidance: This component monitors network resource use. When congestion becomes severe, some packets are discarded to prevent network overload.
- Traffic classification and marking is the prerequisite and foundation for implementing differentiated services.
- Traffic policing, traffic shaping, congestion management, and congestion avoidance are used to control network traffic from different aspects.

HUAWEI

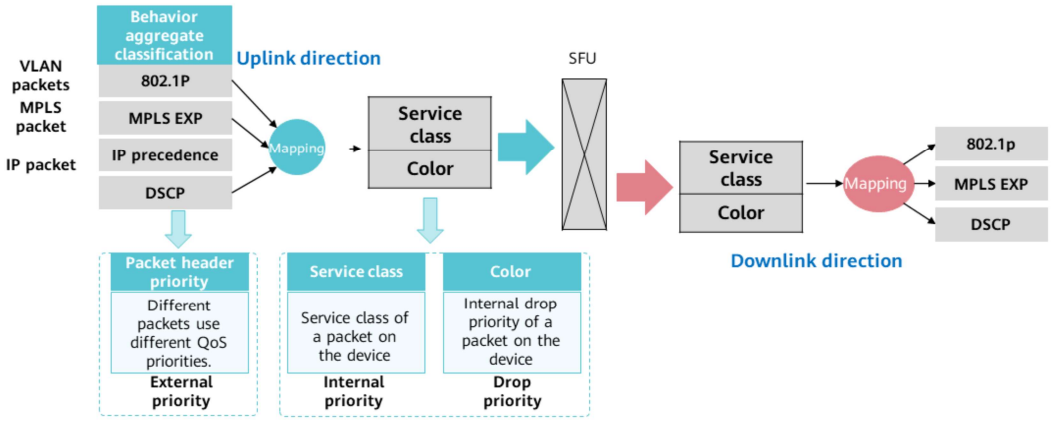


- For the basic packet forwarding process, see HCIP-Datacom-Core Technology-01 Introduction to Network Devices.
- Behavior aggregate classification: Packets are roughly classified based on the IP precedence or DSCP value of IP packets, TC value of IPv6 packets, EXP value of MPLS packets, and 802.1p value of VLAN packets to identify traffic with different priorities or service levels and implement internal-external priority mapping for the traffic.
- Multi-field classification elaborately classifies packets based on complex rules, such as the 5-tuple (source address, source port number, protocol number, destination address, and destination port number).
- Packet Forwarding Engine (PFE): After a router is powered on, it runs a routing protocol to learn the network topology and generate a routing table. If the interface board registers successfully, the main control board can generate forwarding entries according to the routing table and deliver entries to the interface board. In this manner, the router can forward packets according to the forwarding table. The component that forwards data packets is a chip located on an interface board and is called a packet forwarding engine (PFE).

- Traffic Management (TM): The interface board of a Huawei high-end router has a TM chip, which has a high-speed cache. In the case of congestion, data packets are temporarily stored in the cache in the form of queues. Then, the TM schedules the data packets out of the queue according to certain rules and sends the data packets to the switched network. If the cache capacity is exceeded, packets are discarded based on certain rules.
- Committed access rate (CAR): The CAR ensures that the traffic does not exceed the bandwidth allowed by the inbound or outbound interface. Excess packets are directly discarded. Currently, CAR is used for traffic policing. CAR is implemented by the PFE. CAR can be implemented on the uplink PFE to ensure that the traffic rate does not exceed the bandwidth of the inbound interface. You can also perform CAR on the downlink PFE to ensure that the traffic rate does not exceed the bandwidth of the outbound interface.
- Obtaining encapsulation information: On the downlink interface board, the obtained encapsulation information varies according to the forwarding technology. For example, the destination MAC address is obtained for IP forwarding, and the label information is obtained for MPLS forwarding.

Behavior Aggregate Classification

- Behavior aggregate classification allows a device to roughly classify packets based on simple rules.



HUAWEI

DSCP/IP Precedence/ 802.1p/EXP Value Mapping

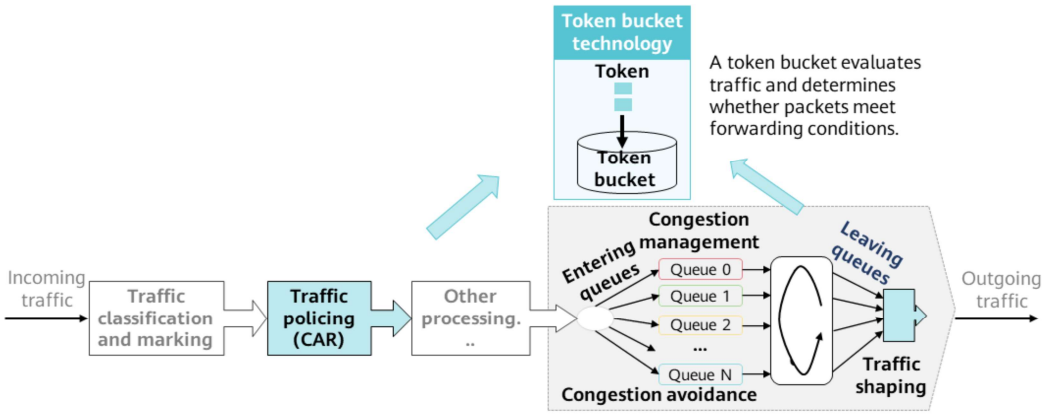
Priorities in descending order

802.1p	MPLS EXP	IP Precedence	DSCP	DSCP Name				
7	7	7	56-63	CS	CS7			
6	6	6	48-55		CS6			
5	5	5	40-47	EF	EF			
4	4	4	32-39	AF	AF4	AF41	AF42	AF43
3	3	3	24-31		AF3	AF31	AF32	AF33
2	2	2	16-23		AF2	AF21	AF22	AF23
1	1	1	8-15		AF1	AF11	AF12	AF13
0	0	0	0-7	BE	BE			

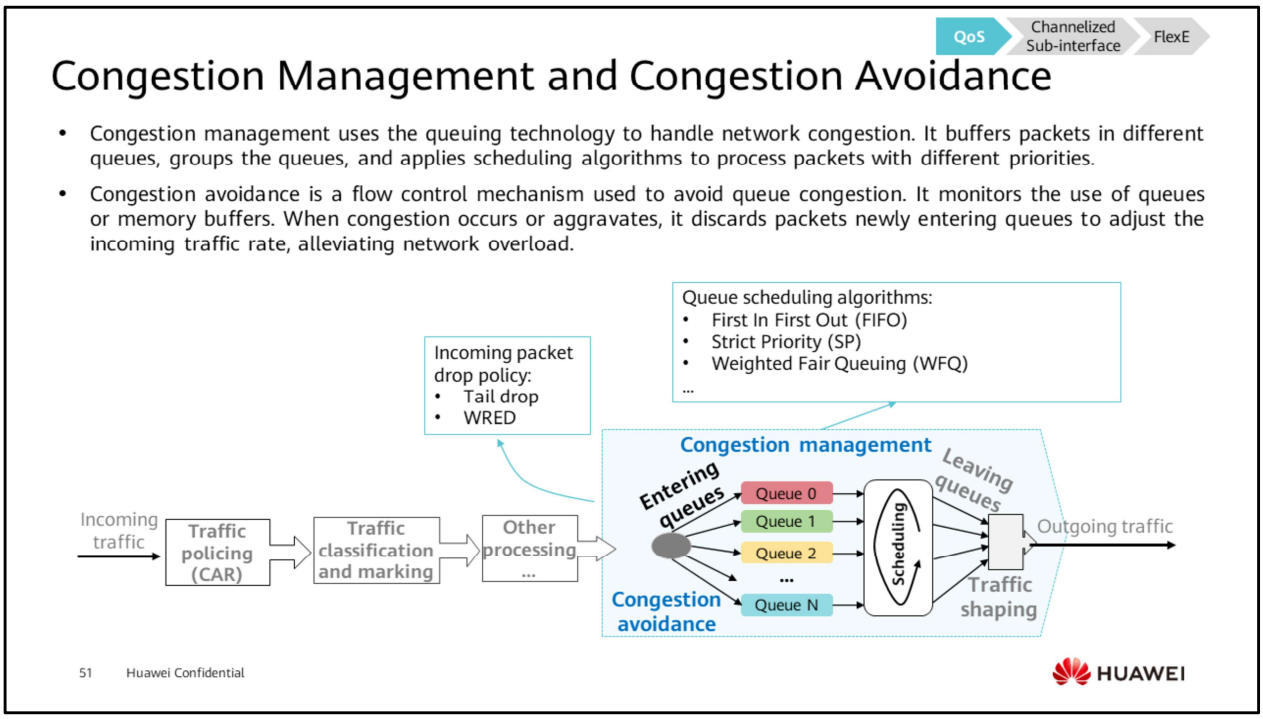
HUAWEI

Traffic Policing and Traffic Shaping

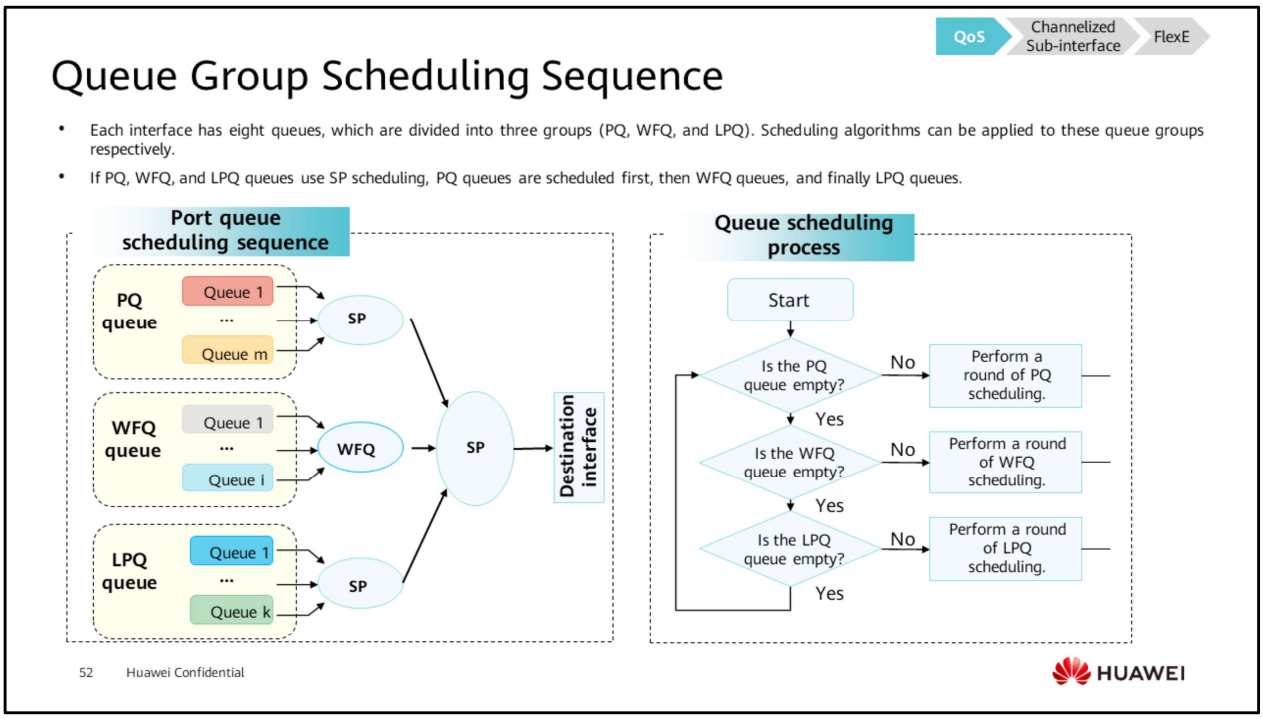
- Both traffic policing and traffic shaping are traffic rate limiting technologies. The former monitors the incoming traffic of a device and limits the incoming traffic rate to a permitted range. If the traffic rate is too high, excess packets are discarded or the packet priorities are re-set. Traffic shaping controls the rate of outgoing packets, so that packets can be sent at an even rate.
- Both traffic policing and traffic shaping use the token bucket algorithm to evaluate the traffic rate.



HUAWEI

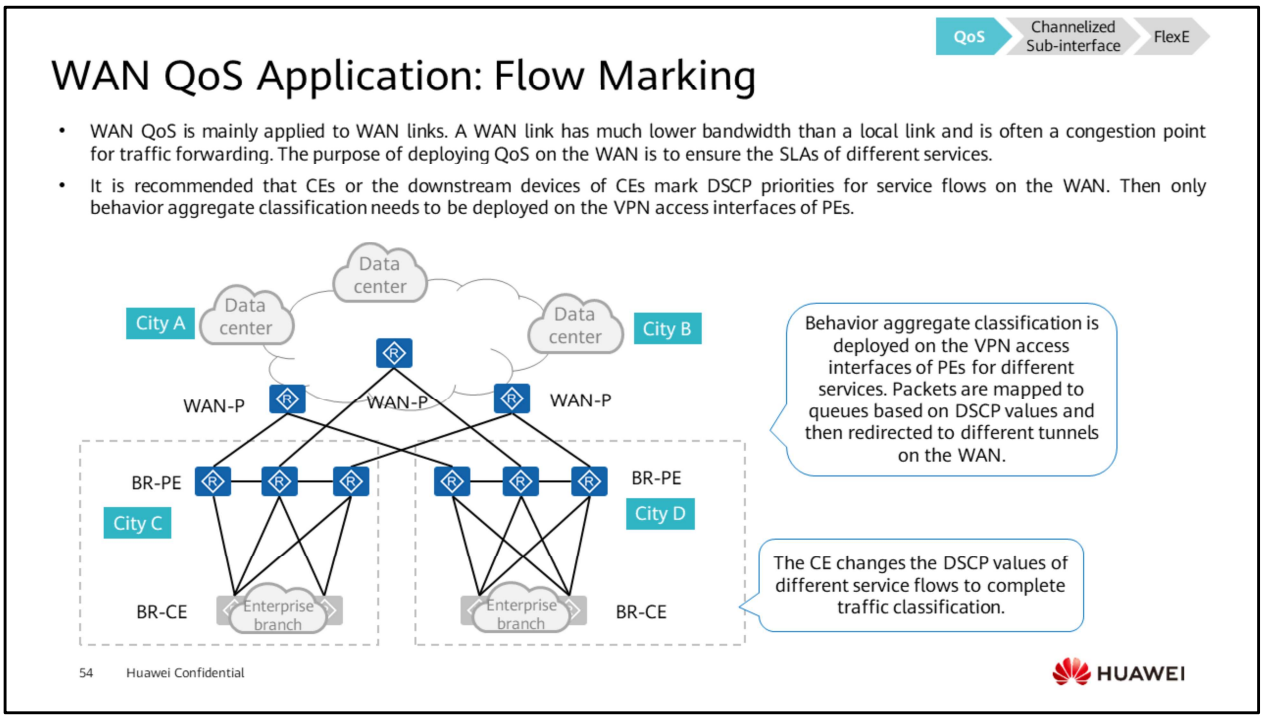


- Weighted Random Early Detection (WRED): The system discards packets based on the drop policies configured for data packets or queues with different priorities. WRED is a congestion avoidance mechanism used to discard packets to prevent queues from being congested. For details, see the product documentation for Huawei NetEngine products.
- FIFO: FIFO does not classify packets. FIFO allows packets to be queued and forwarded in the same order as they arrive at an interface.
- SP: Queues are scheduled strictly according to their priorities. Packets in queues with a low priority can be scheduled only after all packets in queues with a higher priority are scheduled.
- WFQ: The egress bandwidth is allocated to each flow according to the queue weight.
- Other scheduling algorithms, such as RR polling, WRR weighted polling, and DRR differential polling, are not described here.



- PQ queue
  - PQ queues use the SP scheduling algorithm. That is, the packets in the queue with the highest priority are scheduled first. In this way, an absolute priority can be provided for different service data, the delay of delay-sensitive applications such as VoIP can be guaranteed, and the use of bandwidth by high-priority services can be absolutely prioritized.
  - Disadvantage: If the bandwidth of high-priority packets is not limited, low-priority packets may fail to obtain bandwidth and be scheduled.
  - Generally, only delay-sensitive services enter PQ queues.
- WFQ queue
  - WFQ queues are scheduled based on weights. The WFQ scheduling algorithm can be used to allocate the remaining bandwidth based on weights.

- LPQ queue
    - LPQ is a queue scheduling mechanism implemented on a high-speed link (Ethernet) interface. Low-speed links (such as serial and MP-Group links) do not support LPQ queues.
    - Similar to PQ queues, SP scheduling is also used between LPQ queues. The difference lies in that PQ queues can preempt the bandwidth of WFQ queues in the case of congestion, whereas LPQ queues cannot do so. After PQ and WFQ queue scheduling is complete, the remaining bandwidth is allocated to LPQ queues.
    - In actual applications, BE flows can be scheduled into LPQ queues. In this way, when the network load is heavy, BE flows can be completely contained to preferentially meet the requirements of other services.
- All the eight queues on an interface can be configured as WFQ queues, PQ queues, LPQ queues, or a combination of WFQ and LPQ queues.



- If the CE or the downstream device of the CE does not have the traffic marking capability, deploy multi-field classification on the ingress PE on the bearer WAN to mark traffic for queuing. This, however, affects the forwarding performance of the bearer WAN.

WAN QoS Application: Flow Scheduling

- The WAN's forwarding bottleneck lies in the egress WAN link in each region. For example, remote sites are interconnected through a carrier's MSTP private line, which has limited bandwidth. To prevent service flows from being discarded due to rate limiting after entering WAN links on the carrier network, deploy traffic shaping on egress WAN links to ensure that the traffic entering the carrier's MSTP private line does not exceed the rate limit.
- Properly design QoS policies for enterprises' internal services. For example, deploy high-priority PQ scheduling for core production services or services with high QoS requirements and WFQ scheduling for services insensitive to packet loss and delay (only basic bandwidth for service continuity is needed in this case).

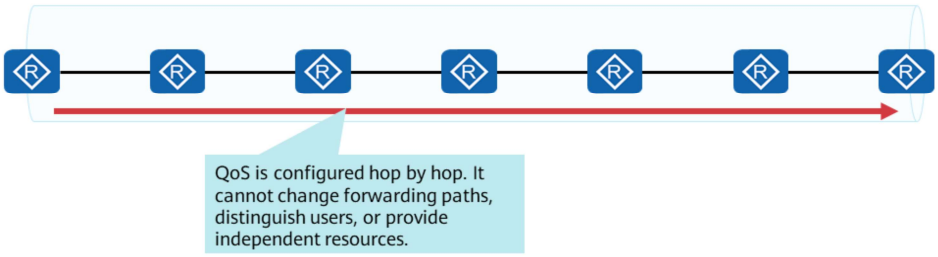
QoS design for an enterprise

Service Type	Importance	Priority	Scheduling Mode
Protocol packet	High	CS7/CS6	PQ
Core service	High	EF	PQ
Video service	High	AF4	PQ
External service	High	AF3	WFQ
Office and test services	Medium	AF2	WFQ
Others	Low	BE	WFQ

HUAWEI

QoS Limitations

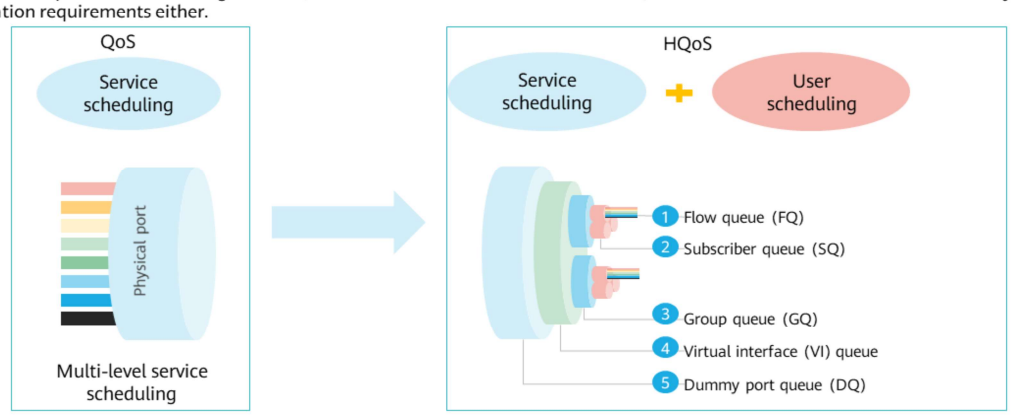
- QoS itself cannot solve the congestion problem or provide isolation and deterministic delay assurance for services:
  - QoS involves only single-hop behaviors and does not change the network topology.
  - QoS does not change service behaviors. If the bursty traffic of a single flow is too heavy, congestion still occurs.
  - The number of QoS queues is small, and SLA assurance cannot be provided for specific users. As a result, deterministic delay assurance cannot be provided.
  - The QoS mechanism is an experience system for resource management and cannot provide independent resources for users.



HUAWEI

HQoS: Providing Scheduling and Bandwidth Guarantee for Users

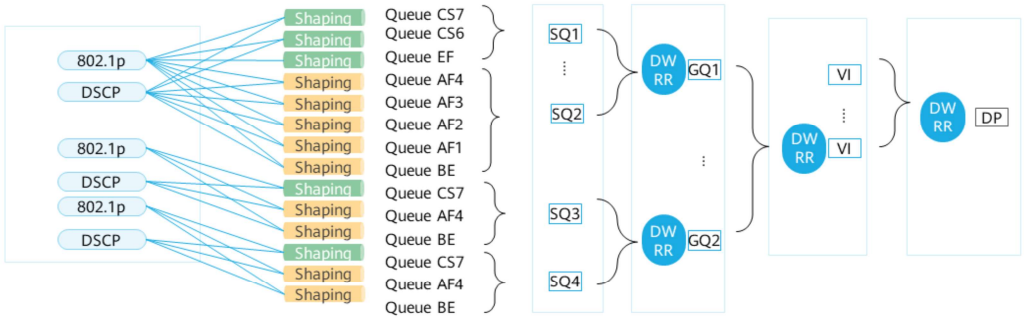
- On the basis of QoS, hierarchical quality of service (HQoS) provides finer-grained and more hierarchical scheduling and management of interface resources to implement fine-grained allocation and management of interface resources. Every scheduling queue can provide bandwidth guarantee, but the entire mechanism is based on QoS and cannot meet the deterministic delay and isolation requirements either.



HUAWEI

HQoS: Providing User-Level Bandwidth Guarantee When the CIR of Each Scheduler Does Not Exceed the Upper Limit

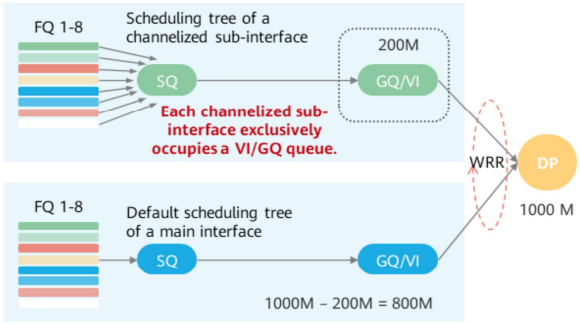
- HQoS adds hierarchical schedulers such as SQ, GQ, and VI. Each scheduler has two attributes (CIR and PIR) and uses SP scheduling for flows whose rates are between the CIR and PIR. The CIR is preferentially guaranteed.
- HQoS guarantees bandwidth through strict CIR deduction and shaping. The sum of the CIRs of all schedulers is less than the interface bandwidth.



HUAWEI

Channelized Sub-interfaces Providing Management Entities on the Basis of HQoS to Provide SLA Assurance

- Bandwidth guarantee: Channelized sub-interfaces use HQoS-based hierarchical scheduling to implement flexible and refined resource management. Each channelized scheduling tree has independent buffer resources and bandwidth resources, providing bandwidth guarantee.
- Delay guarantee: Because the resources for channelized sub-interface are strictly guaranteed, the delay can be guaranteed within a certain range.

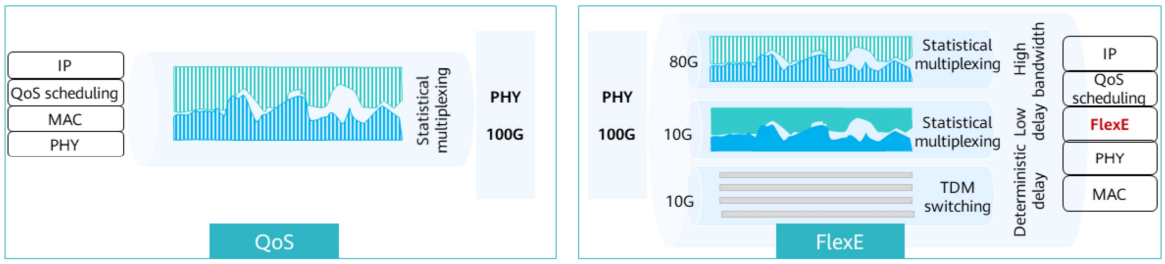


- Channelized sub-interfaces use the HQoS mechanism and exclusively occupy the HQoS VI/GQ scheduling trees and bandwidth to implement strict scheduling isolation.
- Remaining bandwidth of a main interface = Total interface bandwidth – Total bandwidth of all channelized sub-interfaces. Bandwidth is automatically deducted during the channelized sub-interface enabling process, simplifying the HQoS application model.
- Channelized sub-interfaces provide management entities and can work with the controller for resource management and E2E resource reservation, meeting the SLA assurance requirements of P2MP services.

HUAWEI

FlexE: Delivering Isolation and Deterministic Delay Assurance Based on Independent Resources

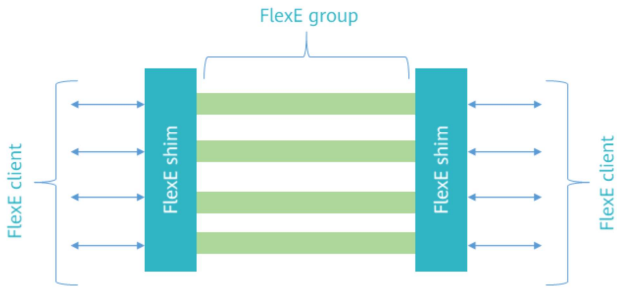
- Flexible Ethernet (FlexE) is a new technology introduced between the MAC and PHY layers. It is a lightweight enhanced technology for IP networks and is compatible with the existing Ethernet standard and QoS capabilities.
- It provides isolated FlexE interface links, enabling one network to carry different types of services.
- It provides an independent P2P time division multiplexing (TDM) tunnel for each service, meeting isolation and deterministic delay requirements.



HUAWEI

Basic FlexE Architecture

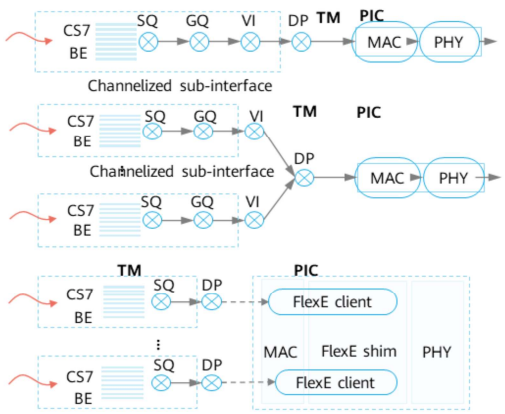
- FlexE decouples the MAC layer from the PHY layer by introducing the FlexE shim layer on the basis of IEEE 802.3, thereby providing flexible rates.
- FlexE adopts the client/group architecture and allows client interfaces at different rates to coexist in a FlexE group.



Client: corresponds to various user interfaces on a network. Each FlexE client can flexibly apply for bandwidth from the resource pool of a FlexE group, adjust the bandwidth, and transfer data flows to the FlexE shim layer as 64B-/66B encoded bit streams.
Group: consists of various Ethernet PHY layers defined in IEEE 802.3 and divides the PHY bandwidth into 1G timeslots (supported by Huawei devices) or 5G timeslots.
Shim: an extra logical layer inserted between the MAC and PHY layers of the traditional Ethernet architecture. It implements key FlexE functions through calendar timeslot distribution.

HUAWEI

Comparison of Slicing Technologies



QoS: All traffic shares eight queues. QoS schedules resources in a unified manner to maximize the statistical multiplexing capability. It cannot differentiate users, and so cannot provide independent resource reservation for different users.

Channelized sub-interface: Queue resources are isolated. Hierarchical scheduling is used to implement flexible and refined management of interface resources, provide bandwidth guarantee, and work with the controller to provide E2E resource reservation.

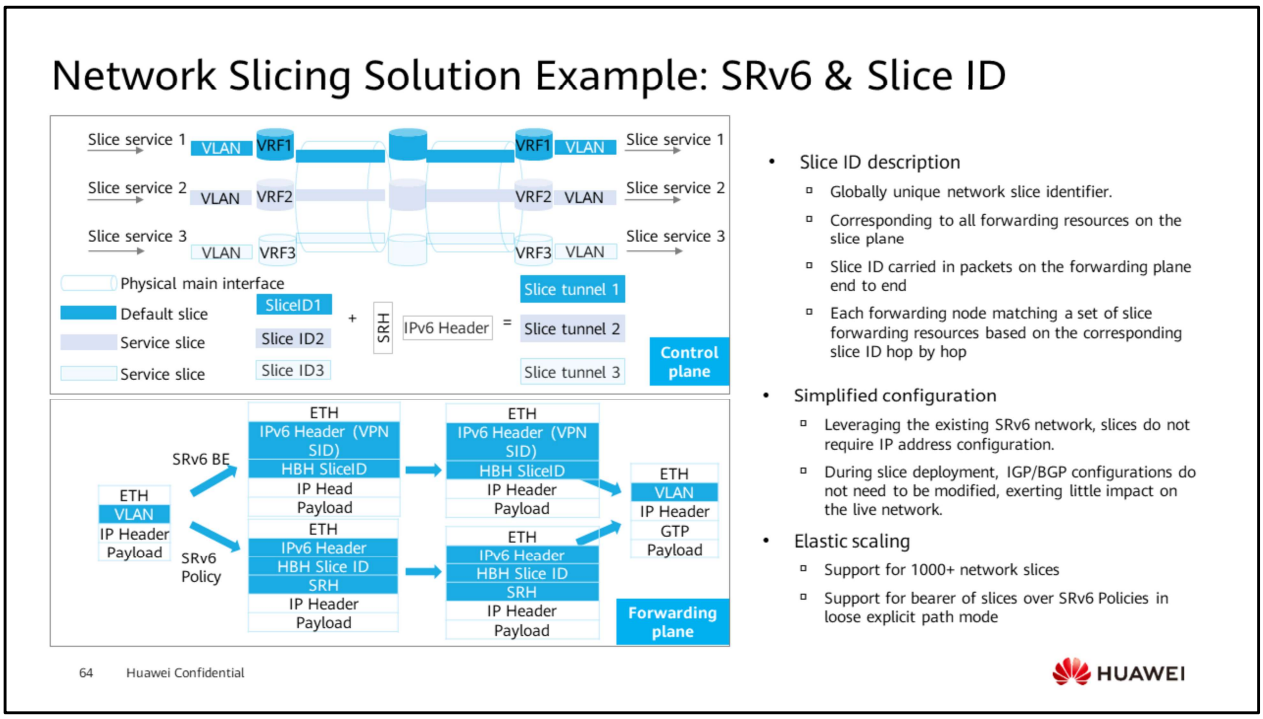
FlexE: Queue and interface resources are isolated. Every resource is divided by TDM timeslot. This meets the requirements for exclusive resource use and resource isolation and provides flexible and refined management of interface resources.

HUAWEI

Slicing Technology Implementation Modes

Interface Name	Usage Scenario and Feature Description
FlexE sub-interface (also called client interface)	A physical interface in standard Ethernet mode has fixed bandwidth. FlexE, however, can enable one or more physical interfaces to work in FlexE mode and add them to a group. The total bandwidth of this group can be allocated on demand to logical interfaces in the group. The group to which physical interfaces are added is referred to as a FlexE group. The logical interfaces that share bandwidth of the physical interfaces in the FlexE group are called FlexE interfaces (also referred to as FlexE service interfaces). FlexE interface bandwidth varies, which allows services to be isolated. Compared with traditional technologies, FlexE technology permits bit-level interface bundling, which solves uneven per-flow or per-packet hashing that challenges traditional trunk technology. In addition, each FlexE interface has a specific MAC address, and forwarding resources between interfaces are isolated. This prevents head-of-line (HOL) blocking that occurs when traditional logical interfaces such as VLAN sub-interfaces are used for forwarding. FlexE interface technology especially fits scenarios in which high-performance interfaces are required for converged bearer, such as mobile bearer, home broadband, and private line access. Services of different types are carried on different FlexE interfaces and are assigned bandwidth based on FlexE interfaces. In this way, FlexE achieves service-specific bandwidth control, meeting network slicing requirements in 5G scenarios.
VLAN channelized sub-interface	A channelized interface can strictly isolate interface bandwidth. A VLAN channelized sub-interface is a channelization-enabled sub-interface of an Ethernet physical interface. Different types of services are carried on different channelized sub-interfaces and assigned bandwidth based on channelized sub-interfaces. This implementation strictly isolates bandwidth among different channelized sub-interfaces on the same physical interface and achieves service-specific bandwidth control, preventing bandwidth preemption among different sub-interfaces.
Ethernet sub-interfaces	An Ethernet sub-interface is a virtual interface configured on a main interface and has Layer 3 features. You can configure an IP address for an Ethernet sub-interface to implement inter-VLAN communication. The main interface can be either a physical interface or a logical interface. The sub-interface inherits the physical layer parameters of the main interface but has its own link layer and network layer parameters. You can activate or deactivate the sub-interface, without affecting the performance of the main interface. The change of the main interface status, however, affects the sub-interface.

HUAWEI



- Different VLANs are used for service access. Logical interfaces correspond to VPN instances VRF1, VRF2, VR3... on the network slice.
- The ingress PE encapsulates the VPN SID and SRv6 Policy information into the service flow on the network slice with the slice ID being 2, and inserts an extension header with the Hop By Hop Slice ID being 2 between the IPv6 header and SRH of each packet.
- Each transit node queries the SRv6 SID in the SRH hop by hop to obtain the physical outbound interface, and then queries the specific "resource reservation" sub-interface of the physical outbound interface based on the slice ID. The Hop By Hop Slice ID remains unchanged throughout this process.
- The egress PE pops the Hop By Hop extension header and forwards the packet to the AC interface of the corresponding VPN instance based on the VPN SID.
- By default, the slice ID is 0, and the IPv6 Hop By Hop extension header does not need to be inserted. The packet format on the forwarding plane is the same as that of traditional L3VPN over SRv6 Policy.

Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. VPN Service
4. Network Traffic Optimization
5. SLA
6. **Network Reliability**
7. Network Management and O&M

HUAWEI

WAN Reliability Overview

- WAN reliability covers two parts: device reliability and network reliability.
- Device reliability: includes router reliability and controller reliability.
  - Controller reliability is implemented through cluster deployment and disaster recovery (DR) system deployment.
  - Router reliability can be implemented using device features such as non-stop routing (NSR). These features are beyond the scope of this course.
- Network reliability: reduces the impact of link and node faults on services through fast detection and convergence mechanisms at each layer.

HUAWEI

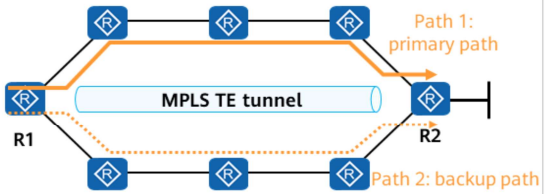
- Because different WAN VPN technologies use different terms, this section briefly describes various protection mechanisms, but does not describe specific protection technologies.

TE Tunnel Protection Technology Basics

- MPLS TE tunnel protection can be provided from two perspectives: local protection and E2E protection. TE tunnels, including MPLS TE tunnels, SR-MPLS TE tunnels, and SR-MPLS Policies can all be protected from the two perspectives, but their technical implementation is slightly different.

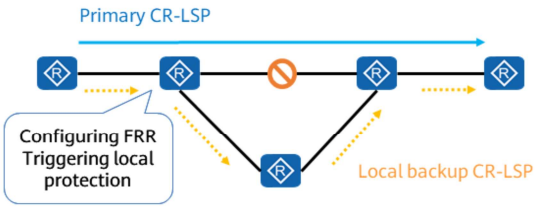
E2E protection

- E2E protection: establishes multiple CR-LSPs between the ingress and egress of a tunnel, with each CR-LSP traversing a different path. When detecting that the primary CR-LSP is faulty, the ingress switches traffic to the backup CR-LSP.
- In E2E protection, path switching is slow. Therefore, E2E protection needs to be used with a detection mechanism for fast path switching.



Local protection

- Fast reroute (FRR), also called local protection, is a temporary protection measure. If a transit node/link fails, local protection is triggered. A backup CR-LSP is then established locally for traffic switching. In addition, the tunnel ingress is instructed to recompute a path and switch traffic to the backup path in a timely manner.



HUAWEI

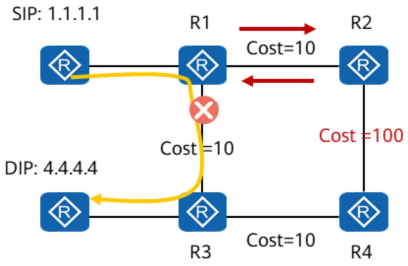
- MPLS TE E2E protection is classified into HSB and ordinary backup. In HSB protection, the backup path and primary path are created at the same time.
- Segment Routing adopts Topology Independent-Loop Free Alternate (TI-LFA), an enhancement of FRR, for local protection.
- Fast detection mechanism: Fast detection mechanisms represented by BFD support fast detection of communication faults between devices.

TI-LFA FRR

- TI-LFA FRR provides link and node protection for SR tunnels. If a link or node fails, traffic is rapidly switched to the backup path.

Limitations of the traditional LFA algorithm

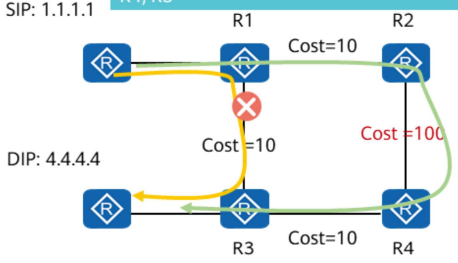
- The traditional LFA algorithm has topological limitations. As shown in the figure, SIP traffic is forwarded to the DIP through R1. If the R1-R3 link fails, R1 forwards the traffic to R2. However, no backup path can be formed before R2 detects the failure.



TI-LFA algorithm

- Using the source routing capability of SR, TI-LFA computes a backup path on each node to protect the failure point. When a node detects a failure, traffic is rapidly switched to the backup path.

Primary R1-R3 path: 4.4.4.4; segment list: R1, R3
Backup R1-R3 path: 4.4.4.4; segment list: R1, R2, R4, R3

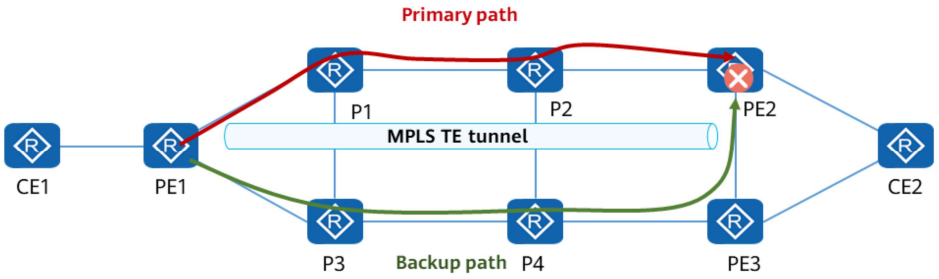


HUAWEI

- In a distributed network architecture, each device independently calculates paths, and there is no consensus on the shortest path when a fault occurs. As a result, traditional LFA cannot form a backup path.

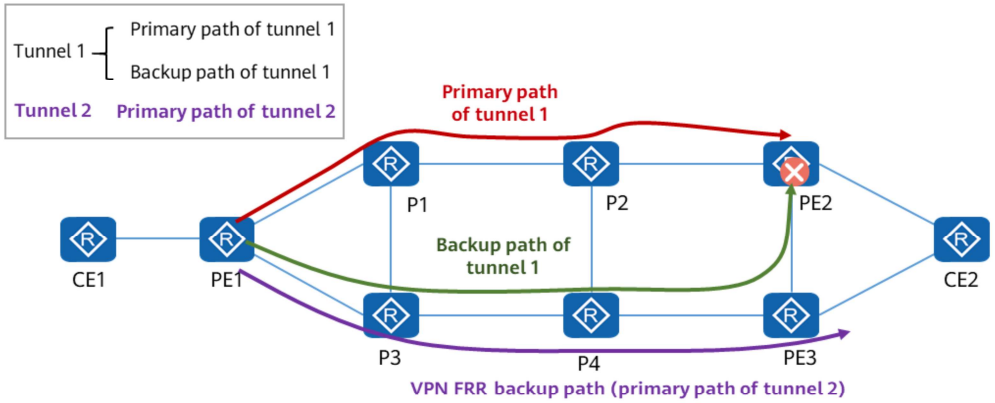
Limitations of E2E Protection

- Technologies such as HSB can protect E2E paths, but cannot rectify faults on PEs.
- In this example, a TE tunnel with both primary and backup paths is established between PE1 and PE2. If PE2 is faulty, the backup path cannot solve the problem.



VPN FRR

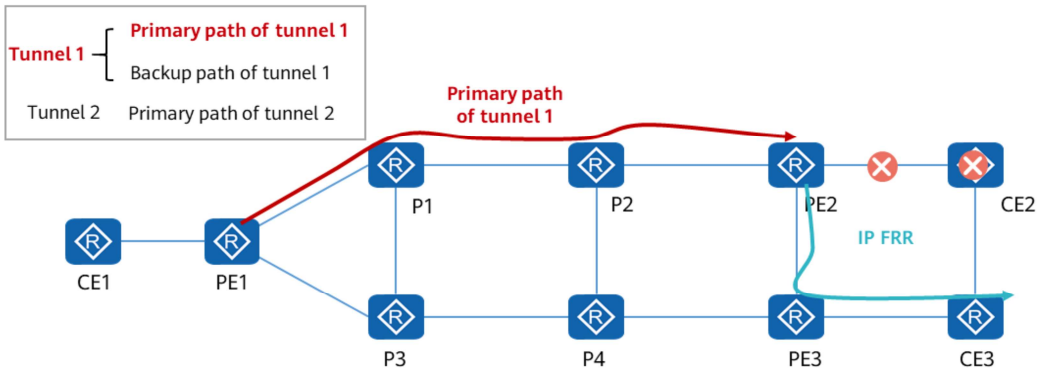
- VPN FRR sets the primary and backup forwarding paths pointing to the active and standby PEs on the remote PE in advance. It works with fast PE fault detection to accelerate fault-triggered E2E service convergence in scenarios where a CE is dual-homed to two PEs.



HUAWEI

IP FRR

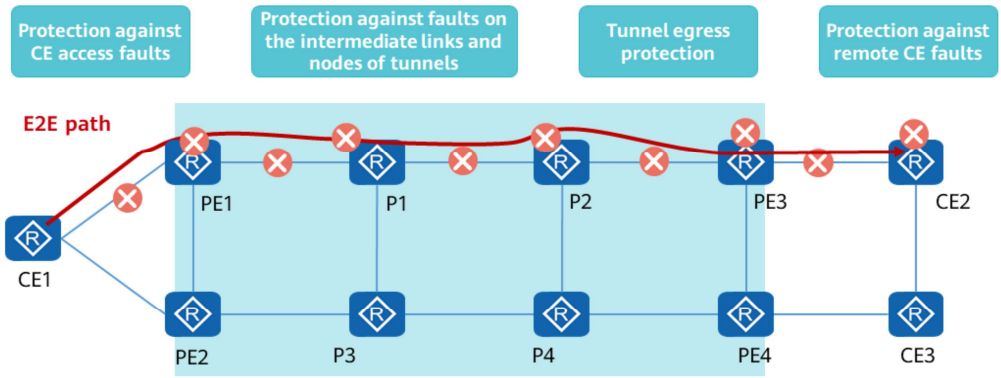
- If the link between PE2 and CE2 fails but PE2 still functions properly, the tunnel between PE1 and PE2 is still available. In this case, E2E tunnel switching is not required. PE2 selects PE3 as the backup next hop. If the link between PE2 and CE2 or CE2 fails, IP FRR is implemented to rapidly switch IP traffic.
- IP FRR is applicable to the IP network between CEs and PEs.



HUAWEI

Summary: Multi-Level Network Protection

- To sum up, different protection measures are taken to ensure tunnel reliability based on the locations of faults on E2E paths.
- From the perspective of E2E forwarding paths, it is recommended that multi-level protection be used:



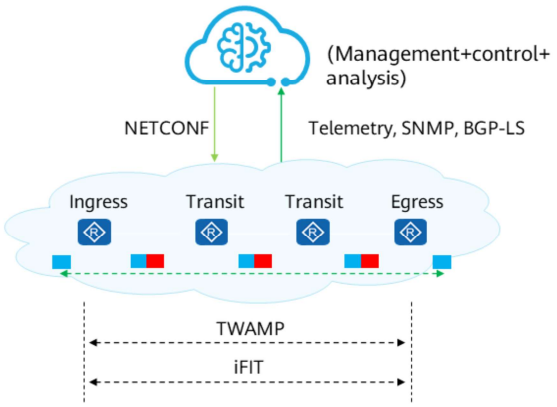
Contents

1. Bearer WAN Architecture
2. Bearer WAN Basics
3. VPN Service
4. Network Traffic Optimization
5. SLA
6. Network Reliability
7. **Network Management and O&M**

HUAWEI

WAN Management and O&M

- Among mainstream WAN solutions provided by different vendors, the SDN solution is preferred. The SDN controller centrally manages and delivers WAN services to forwarders.
- In Huawei's solution, the controller not only provides "control" functions, but also provides "management" and "analysis" functions.



Network analysis

- The AI algorithm is used to obtain network analysis results based on massive amounts of network performance and monitoring data.

Network collection/management

- Multiple protocols and channels are used to collect network configuration data, performance data, and monitoring data.
- Efficient network configuration management is provided.

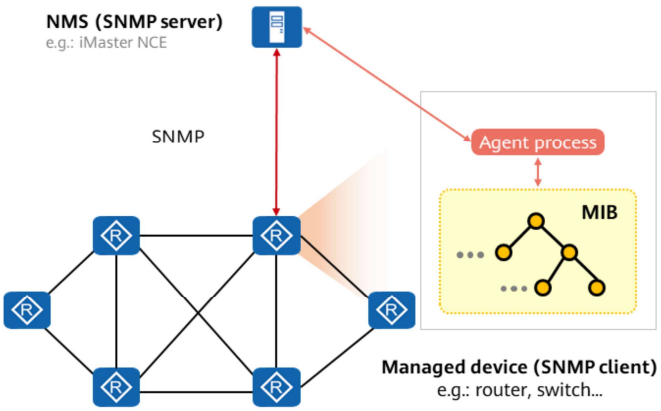
Network measurement

- Network performance indicators, such as the delay, jitter, and packet loss rate, are measured.

HUAWEI

- This figure does not show protocols related to tunneling and traffic statistics collection.

SNMP

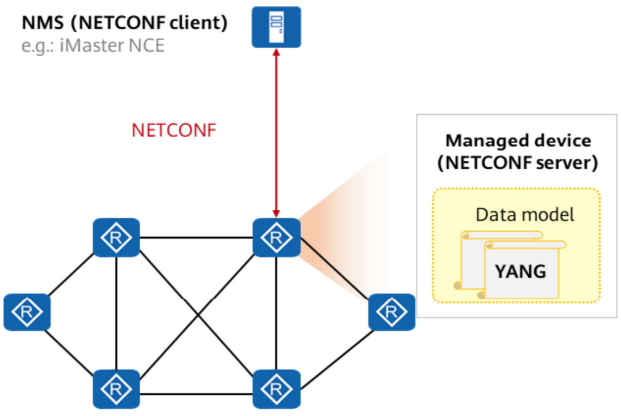


SNMP overview

- Simple Network Management Protocol (SNMP) is a network management standard widely used on TCP/IP networks.
- It provides a method for managing devices through a central computer that runs network management software — known as a network management station (NMS).
- By employing the "network management over networks" mode, SNMP implements efficient and batch network device management. In addition, SNMP enables unified management of network devices of different types and from different vendors.

HUAWEI

NETCONF



SNMP's drawbacks

SNMP is not a configuration-oriented protocol. On a large-sized network with a complex topology, SNMP cannot meet network management requirements, especially the configuration management requirements.

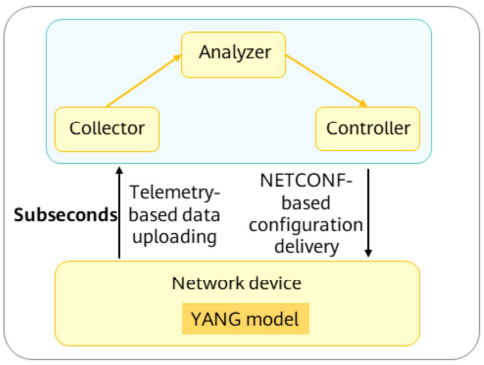
NETCONF overview

- The Network Configuration Protocol (NETCONF) provides a mechanism for the NMS to communicate with network devices.
- To be specific, the network administrator can use this mechanism to add, modify, and delete the configurations of network devices as well as obtain the configurations and status of network devices.
- NETCONF uses a YANG file to describe the data model of a device. A YANG file has a more hierarchical structure than a MIB file.

HUAWEI

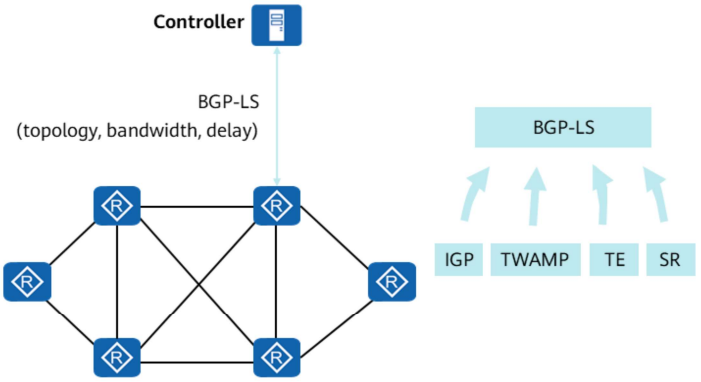
Telemetry

- Telemetry, also known as network telemetry, is mainly used to monitor networks, including packet check and analysis, intrusion and attack detection, intelligent data collection, and application performance management. Generally, it is used together with NETCONF. The analyzer analyzes the data collected by telemetry and then instructs the controller to automatically modify device configurations based on analysis results.
- Advantages of telemetry:
  - Is developed based on the YANG model.
  - Collects a wide variety of data with high precision to fully reflect network status.
  - Continuously reports data with only one-time data subscription.
  - Locates faults rapidly and accurately.



BGP-LS

- BGP-LS introduces new NLRI into BGP. The NLRI carries link, node, topology prefix, and other information, and is also referred to as the link state NLRI.
- BGP-LS can aggregate network-layer topology, bandwidth, delay, and other information and send the information to the controller for path computation.



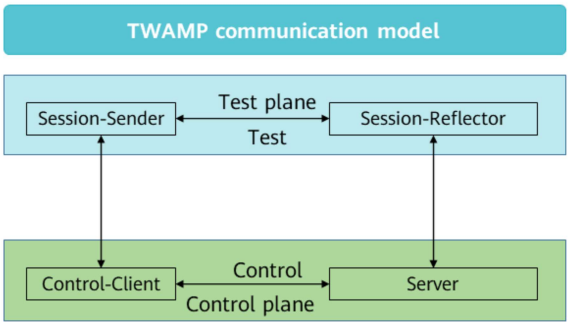
BGP-LS aggregates information collected by various network layer protocols, including:

- IGP: IGP information of each AS
- TWAMP: measurement information, such as interface delay
- TE: TE information, such as bandwidth
- SR: SR information, such as SR labels

HUAWEI

TWAMP

- Two-Way Active Measurement Protocol (TWAMP) measures the two-way delay, jitter, and packet loss rate between devices on an IP network. It performs negotiation over a TCP connection and uses UDP data packets as measurement packets.



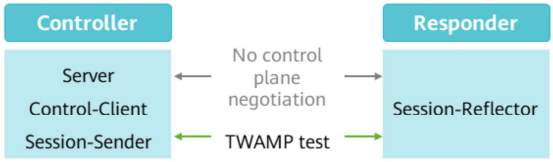
- Control-Client: establishes, starts, and stops a test session and also collects statistics.
- Server: responds to the Control-Client's request for establishing, starting, or stopping a test session
- Session-Sender: proactively sends probes for performance statistics after being notified by the Control-Client.
- Session-Reflector: replies to the probes sent by the Session-Sender with response probes after being notified by the Server.

HUAWEI

- TWAMP Light is a lightweight version of TWAMP defined based on standard protocols. Compared with TWAMP, TWAMP Light simplifies the control protocol used to establish performance measurement sessions.

TWAMP Light

- As the light version of TWAMP, TWAMP Light eliminates the necessity of the TWAMP-Control protocol, and moves the control plane from the Responder to the Controller so that TWAMP control modules can be centrally deployed on the Controller. Therefore, TWAMP Light greatly relaxes its requirements on the Responder performance, allowing the Responder to be rapidly deployed.
- The Control-Client, Server, and Session-Sender are deployed on the same host and function as the Controller. Therefore, the control session establishment process in the standard architecture can be ignored in the performance detection process. The Session-Reflector is deployed on another host as the Responder.



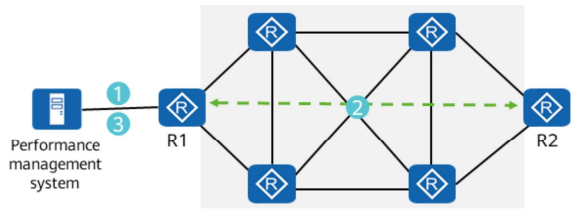
1. The Controller sends test packets to the Responder, and the Responder functions as the Session-Reflector to reflect the test packets.
2. The Session-Reflector does not need to learn Session status. After receiving the test packets, the Session-Reflector copies the necessary information, generates a test packet sequence number and timestamp, and returns the test packets to the Controller.
3. Upon receiving the reflected test packets, the Controller collects the indicators in both path directions.

HUAWEI

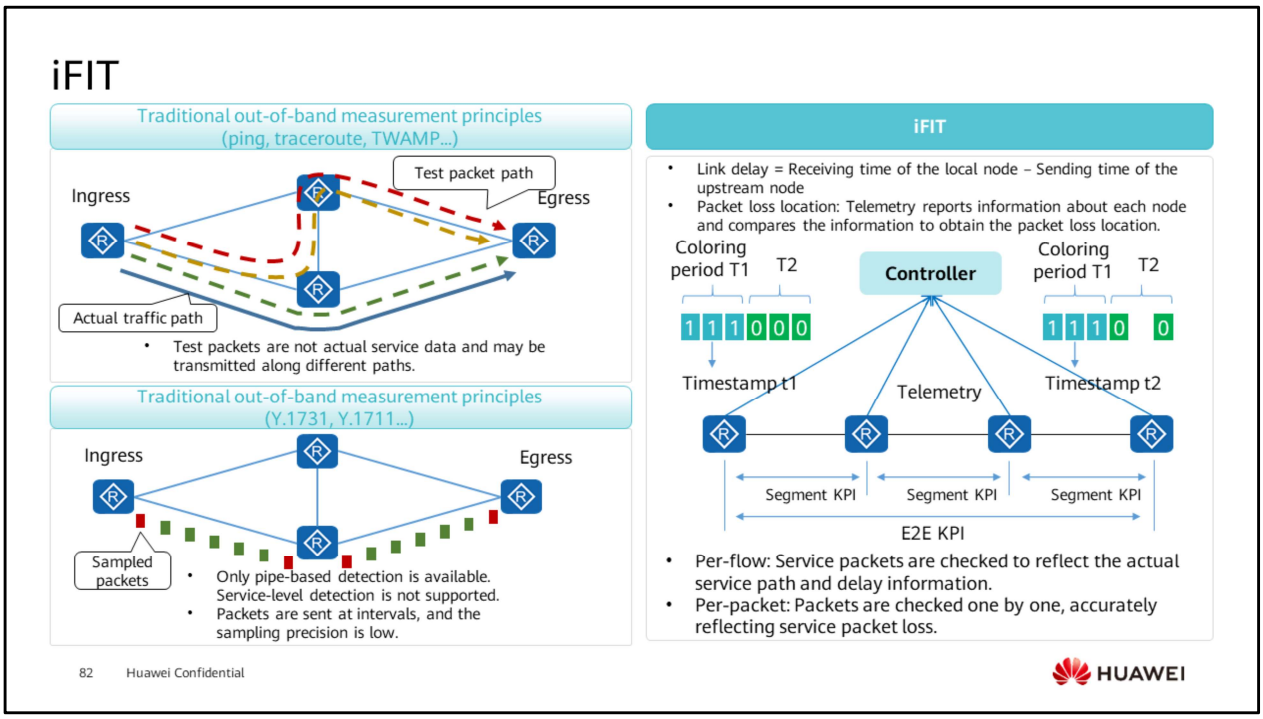
- The standard TWAMP version uses TCP for control plane negotiation, and test packets are based on UDP. The reflector needs to know the session status so that devices of different vendors can communicate with each other.
- TWAMP Light does not involve control plane negotiation, and test packets are also based on UDP. The implementation and configuration are simple, and the reflector does not need to know the session status.

TWAMP Application Scenarios

- With TWAMP, NEs do not need to generate or maintain IP network performance statistics. The performance management system can easily obtain statistics about the entire network by managing only the TWAMP clients initiating statistics collection requests. In this way, IP performance statistics are collected quickly and flexibly.
- TWAMP is used on enterprise WANs to measure the delay, jitter, and packet loss rate between any two nodes, providing a reference for troubleshooting and traffic optimization.



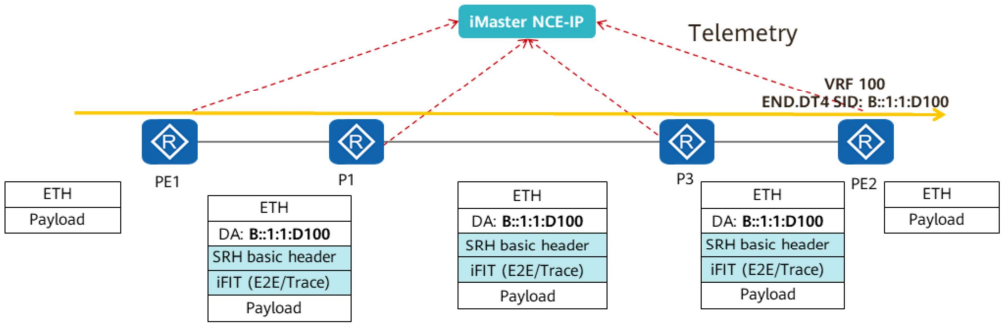
1. The Controller instructs network devices to establish performance measurement sessions. For example, R1 functions as the Control-Client to initiate IP performance measurement.
2. R1 and R2 initiate a TWAMP measurement session.
3. R1 reports the measurement result to the Controller.



- iFIT measures E2E service packets to obtain performance indicators of an IP network, such as the packet loss rate and delay. iFIT adds a color flag to the packet header in the service flow. Telemetry is used to periodically collect information. Features such as E2E delay measurement and packet loss measurement are supported.

iFIT for VPN over SRv6 Scenario

- Basic SRv6 scenario:
  - The ingress inserts the SRH between the IPv6 header and payload. The SRH carries the SRH basic header and iFIT extension header.
  - SRv6-capable nodes can report iFIT statistics in either E2E or hop-by-hop mode.
  - A node that does not support SRv6 but supports IPv6 forwarding can properly forward service packets carrying iFIT information.

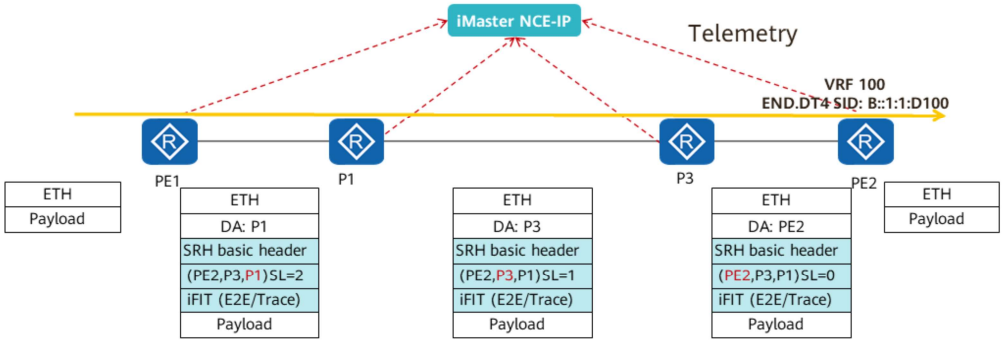


HUAWEI

- An End.DT4 SID (PE endpoint SID) identifies an IPv4 VPN instance on a network.
- For MPLS packets, the iFIT header is inserted between the MPLS label and MPLS payload.

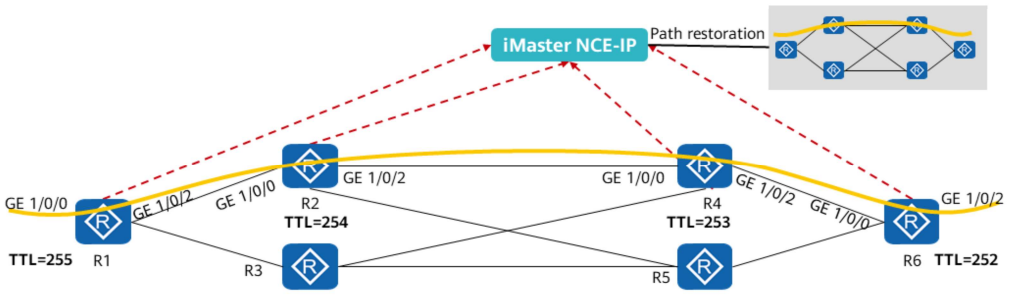
iFIT for SRv6 Policy Scenarios

- In an SRv6 Policy scenario:
  - The iFIT extension header is encapsulated into the Optional TLV field of the SRH.
  - SRv6-capable nodes can report iFIT statistics in either E2E or hop-by-hop mode.
  - A node that does not support SRv6 but supports IPv6 forwarding can properly forward service packets carrying iFIT information.



iFIT-based Service Path Display

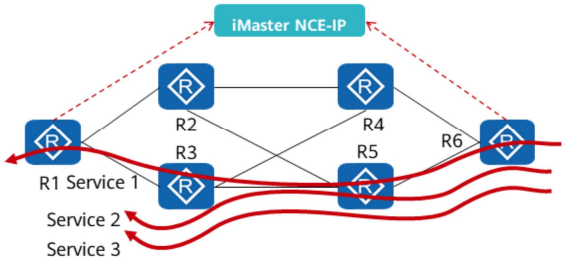
- In each measurement period, a device reports the flow direction, interface number, and TTL information when reporting packet statistics.
- iMaster NCE-IP restores the path information of the flow based on the information reported by each node.
- The implementation is independent of the tunnel type (SRv6/SRv6 Policy/SR-MPLS TE/SR-MPLS BE/MPLS...).



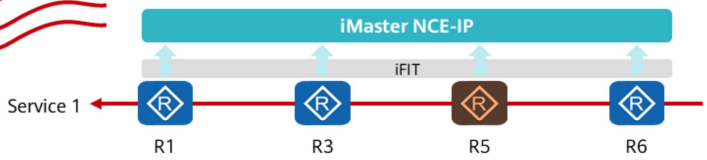
HUAWEI

iFIT-based Fault Locating

- Path aggregation is performed based on the physical topology, end nodes of poor-QoE services, and the built-in AI algorithm of iMaster NCE-IP to determine the minimum area that causes poor service quality, helping further locate and demarcate faults.



1. Find out poor-QoE services, such as services 1 to 3, on the entire network.
2. Intelligently compute the common paths of the three services.
3. Select a service for iFIT-based fault detection. Service 1 is used as an example here.
4. Determine the faulty device R5 based on per-hop iFIT measurement.



HUAWEI

Quiz

1. (True or False) An SR Policy is identified by <headend, color, endpoint> and contains multiple candidate paths. ( )
   A. True
   B. False
2. (Multiple-answer question) Which of the following are WAN bearer technologies? ( )
   A. SR-MPLS
   B. SRv6
   C. MPLS LDP
   D. MPLS TE

HUAWEI

1. A
2. ABCD

Summary

- This course introduces the concepts and principles of the enterprise bearer WAN's typical architecture, bearer technologies, VPN services, traffic optimization, SLA, reliability, and network management and analysis. To introduce these key aspects, this course uses a large enterprise with three data centers in two cities and multiple branches in different regions as an example.
- On a real production network, engineers need to determine the network architecture and technical applications based on the live network conditions and enterprises' services.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI

Enterprise Bearer WAN Design



HUAWEI

Foreword

- The enterprise bearer WAN is also called the enterprise backbone WAN, which is used for interconnection between enterprise branches, between enterprise branches and clouds, and between clouds.
- In the past, the enterprise bearer WAN only needs to forward traffic. However, as growing enterprise services depend increasingly on the network, service SLA assurance requirements are imposed on the network.
- To meet service requirements, the enterprise bearer WAN has employed various new technologies, which complicates the O&M and design of the bearer WAN.
- This course describes the technologies and network design roadmap of the bearer WAN based on Huawei's CloudWAN solution.

HUAWEI

- This course is based on Huawei's CloudWAN solution.

Objectives

- Upon completion of this course, you will be able to:
  - Describe the challenges faced by the enterprise bearer WAN.
  - Describe the basic functions of Huawei's CloudWAN solution.
  - Describe the basic design roadmap of the enterprise bearer WAN.
  - Describe the tunnel and VPN design roadmap of the enterprise bearer WAN.
  - Describe the SLA and reliability design roadmap of the enterprise bearer WAN.
  - Describe the optimization and O&M design roadmap of the enterprise bearer WAN.

HUAWEI

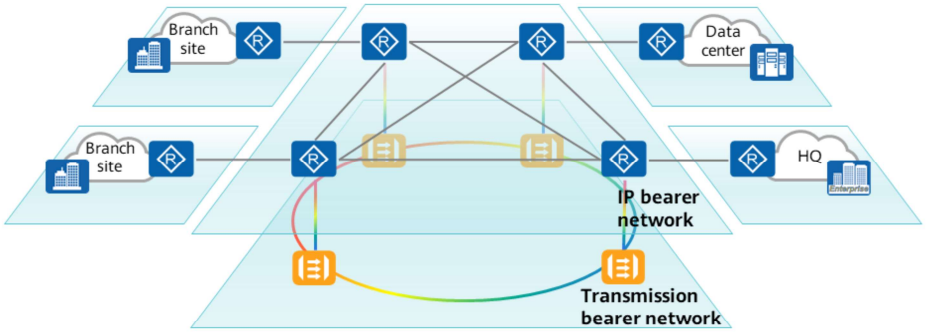
Contents

1. **Current Situation and Challenges of the Enterprise Bearer WAN**
2. Huawei CloudWAN Solution Overview
3. Basic Design for the Enterprise Bearer WAN
4. Tunnel and VPN Design for the Enterprise Bearer WAN
5. SLA and Reliability Design for the Enterprise Bearer WAN
6. Optimization and O&M Design for the Enterprise Bearer WAN

HUAWEI

Enterprise Bearer WAN Overview

- Bearer WANs can be classified into IP bearer networks and transmission bearer networks by network layer. For most enterprises, transmission bearer networks cannot be built by themselves (optical fiber layout requires qualification certificates). Therefore, they usually lease transmission lines from carriers. An enterprise can build its own IP bearer network over the transmission bearer network.
- The IP bearer network built by an enterprise is called the enterprise bearer WAN, which is also called the enterprise backbone network or enterprise core network.
- The enterprise bearer WAN carries the enterprise's internal cross-region interconnection services, such as synchronization services between data centers and voice services between enterprise branches and the HQ.

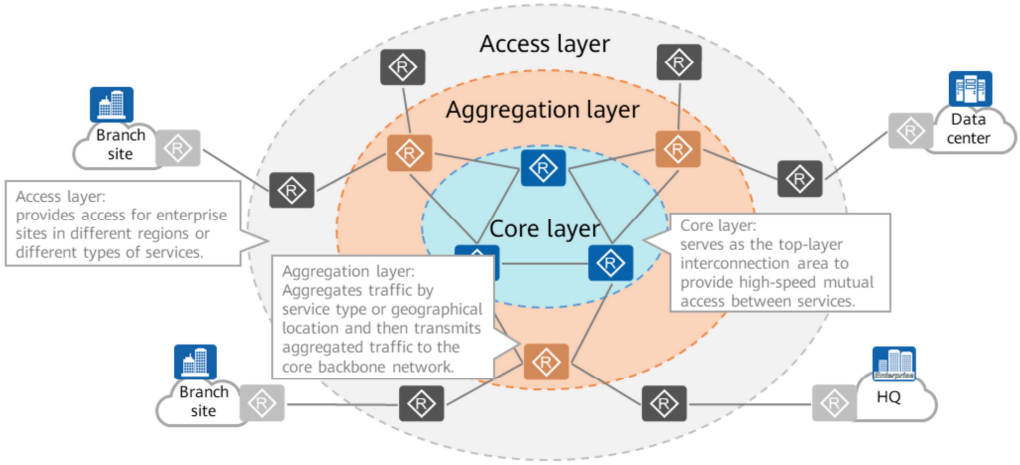


HUAWEI

- The bearer WAN mentioned in this course mainly refers to the IP bearer WAN.

Enterprise Bearer WAN Architecture

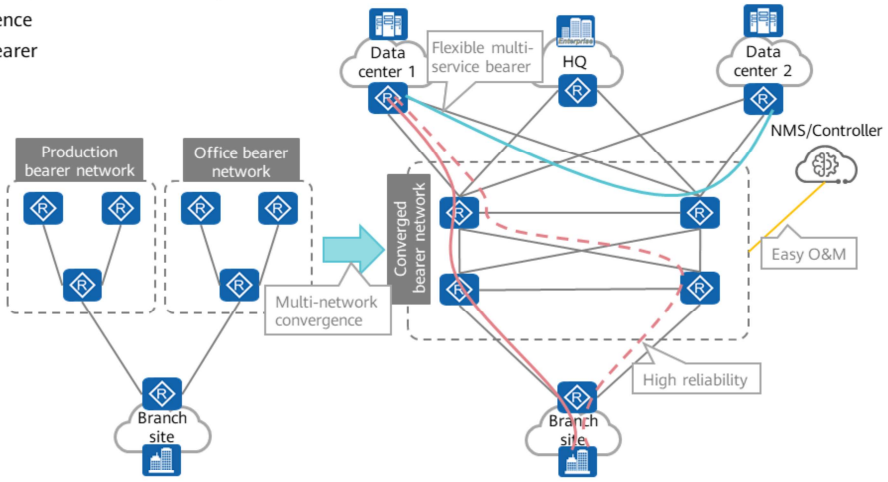
- While the bearer WANs of different enterprises vary greatly in scale and actual topology, their architecture can be generally divided into the access, aggregation, and core layers.



HUAWEI

Development Trend of the Enterprise Bearer WAN

- The emergence of new technologies such as cloud computing and big data promotes the development of enterprise services and poses new requirements on enterprise bearer WANs.
  - Multi-network convergence
  - Flexible multi-service bearer
  - High reliability
  - Easy O&M



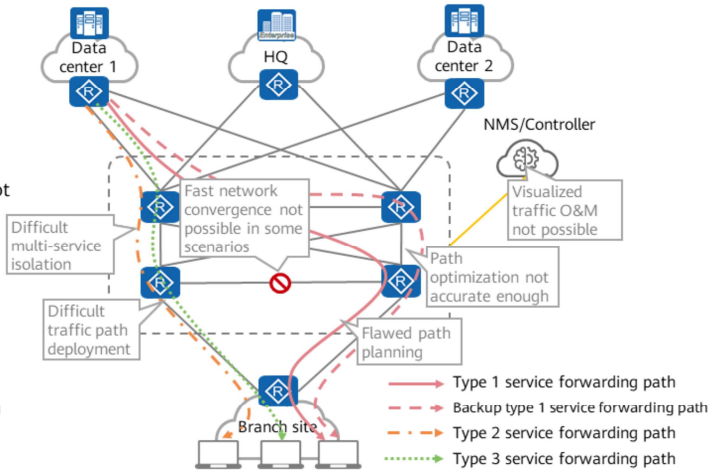
HUAWEI

- Multi-network convergence:
  - The bearer network uses one physical network to carry production, office, and other services. Powerful BGP routing policies are used to control traffic on the bearer network. Traffic diversion policies are deployed based on service attributes so that different types of services, such as production and office services, can run on different paths based on the customized policies.
- Flexible multi-service bearer:
  - As multiple services are migrated to the cloud and multiple networks are converged, issues such as low network resource utilization and unbalanced traffic distribution become more prominent. The segment routing technology can be used to flexibly plan paths to carry traffic, improving network utilization.
- High reliability:
  - Reliability is classified into network reliability and service reliability.
  - Network reliability is ensured from aspects such as network architecture and network stability.
  - Service reliability is mainly ensured by optimizing network bandwidth, delay, and packet loss rate based on service requirements.

- Easy O&M:
  - Topologies, events, configurations, service provisioning, and link quality are managed in a unified manner.
  - The import of traffic into tunnels, tunnel bandwidth, paths, and node traffic are dynamically controlled.
  - Application data flows are visible, analyzable, and predictable.

Challenges Faced by Enterprise Bearer WANs

- The development trend of the enterprise bearer WAN brings many challenges.
  - Multi-network convergence
    - Challenge: It is difficult to isolate multiple services.
  - Flexible multi-service bearer
    - Challenge: Network resource planning is not precise enough, and traffic path planning and deployment are difficult.
  - High reliability
    - Challenge: Fast convergence cannot be implemented in all scenarios, and service SLA cannot be effectively guaranteed.
  - Easy O&M
    - Challenge: Precise traffic path optimization and visualized O&M cannot be provided.



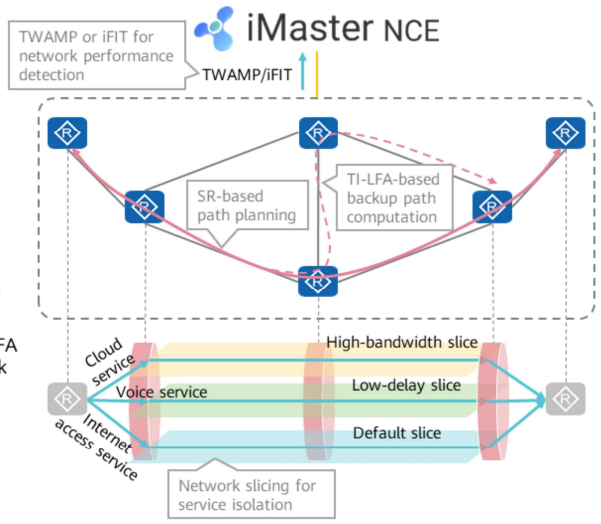
HUAWEI

- Multi-service isolation:
  - Due to the convergence of multiple networks, the originally physically isolated production and office networks are now carried on the same network. The service isolation requirements of some enterprises can be met using the VPN technology. However, for other enterprises, services need to be isolated using hard pipes.
- Network resource and traffic path planning:
  - To make full use of network resources (mainly bandwidth), network traffic needs to be planned. Traditional traffic engineering (mainly MPLS-TE) has many disadvantages, such as complex configuration and flawed path computation mechanism. Therefore, it cannot effectively solve the problem in multi-service bearing.
- Fast network convergence and service SLA assurance:
  - Traditional networks usually use FRR to shorten the convergence time. However, FRR based on the LFA or RLFA algorithm has restrictions in application scenarios and may fail to meet the high reliability of some networks.

- Traffic path optimization and visualized O&M:
  - Traditional enterprise WANs mainly use the NQA technology to measure network quality. However, the NQA technology cannot precisely reflect the actual network situation. As a result, traffic paths cannot be effectively optimized.
  - Traditional network management software cannot perform end-to-end traffic monitoring based on services, and therefore cannot implement visualized O&M for traffic.

Solutions to Challenges Faced by the Enterprise Bearer WAN

- To address challenges faced by the enterprise Bearer WAN, Huawei provides the CloudWAN solution.
  - Difficult multi-service isolation
    - Network slicing technology is deployed to carry different services over different slices.
  - Difficult traffic path planning and deployment
    - Based on SR, the controller plans paths globally and automatically deploys paths.
  - Lack of fast network convergence and service SLA assurance
    - TI-LFA is used for backup path computation, and TI-LFA works with FRR and HSB technologies for fast network convergence.
  - Difficult traffic path optimization and lack of visualized O&M
    - TWAMP and iFIT are used to precisely detect network quality, facilitating accurate traffic optimization.

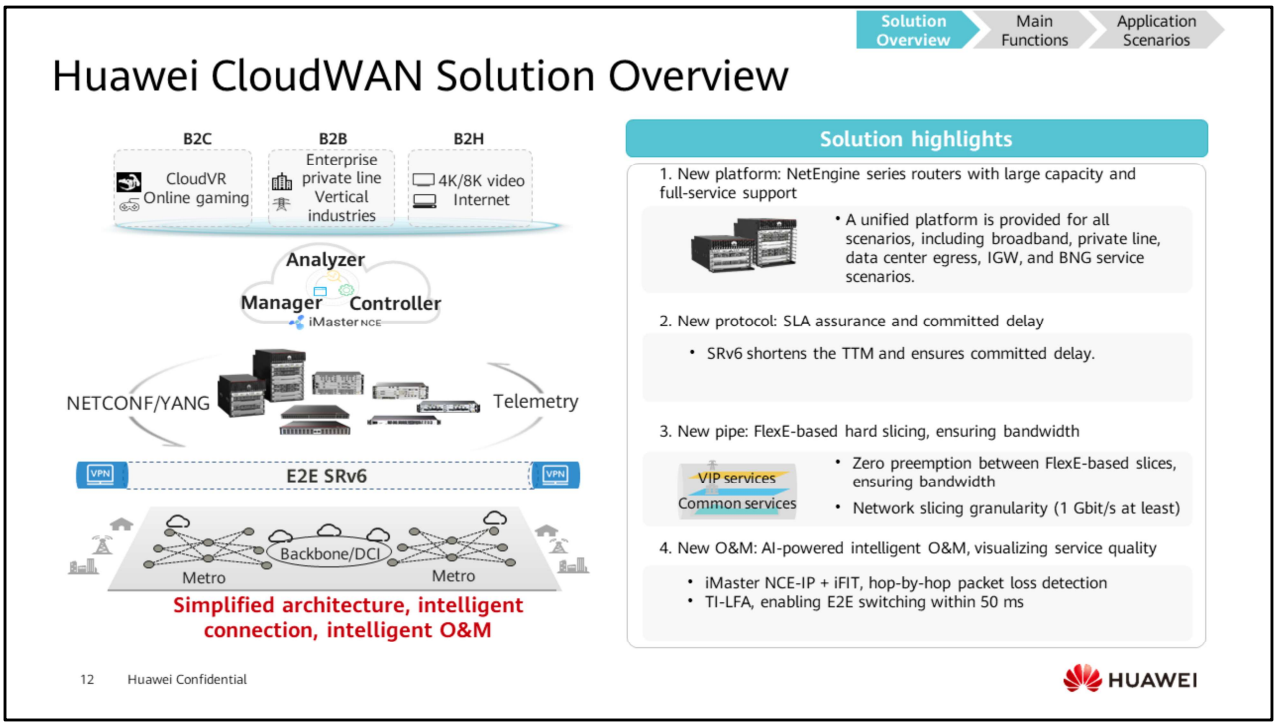


HUAWEI

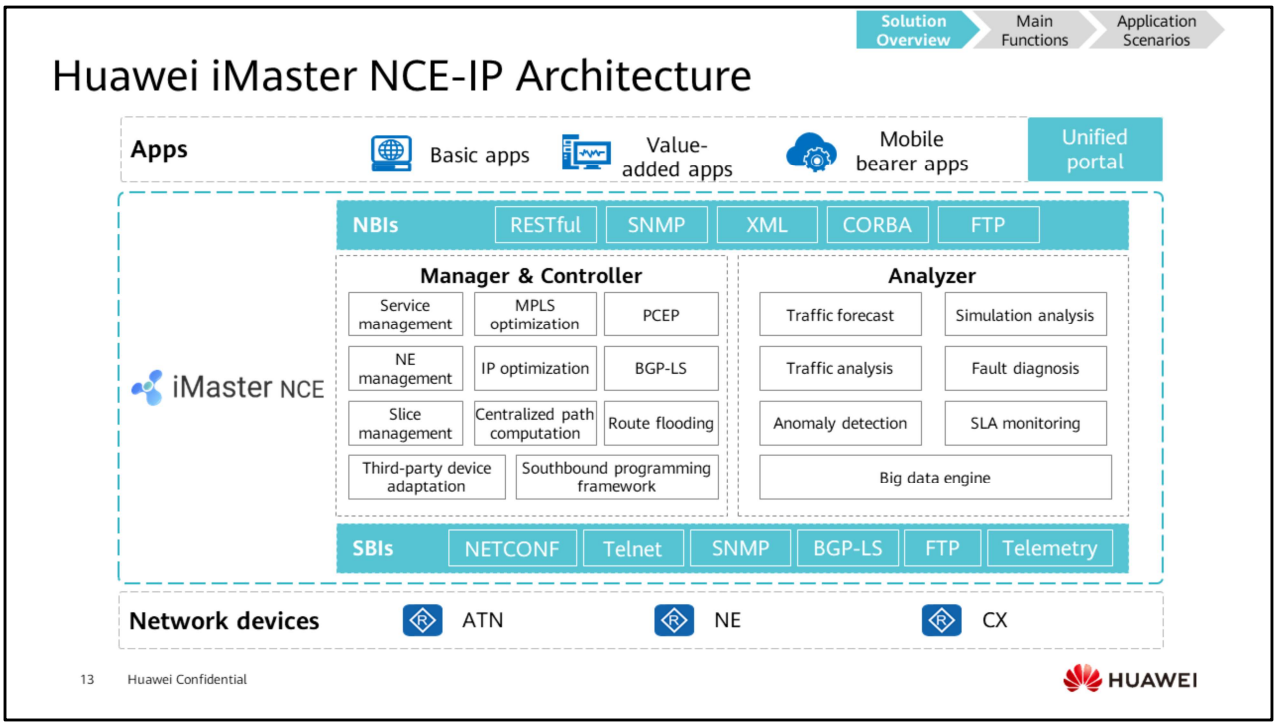
Contents

1. Current Situation and Challenges of the Enterprise Bearer WAN
2. **Huawei CloudWAN Solution Overview**
3. Basic Design for the Enterprise Bearer WAN
4. Tunnel and VPN Design for the Enterprise Bearer WAN
5. SLA and Reliability Design for the Enterprise Bearer WAN
6. Optimization and O&M Design for the Enterprise Bearer WAN

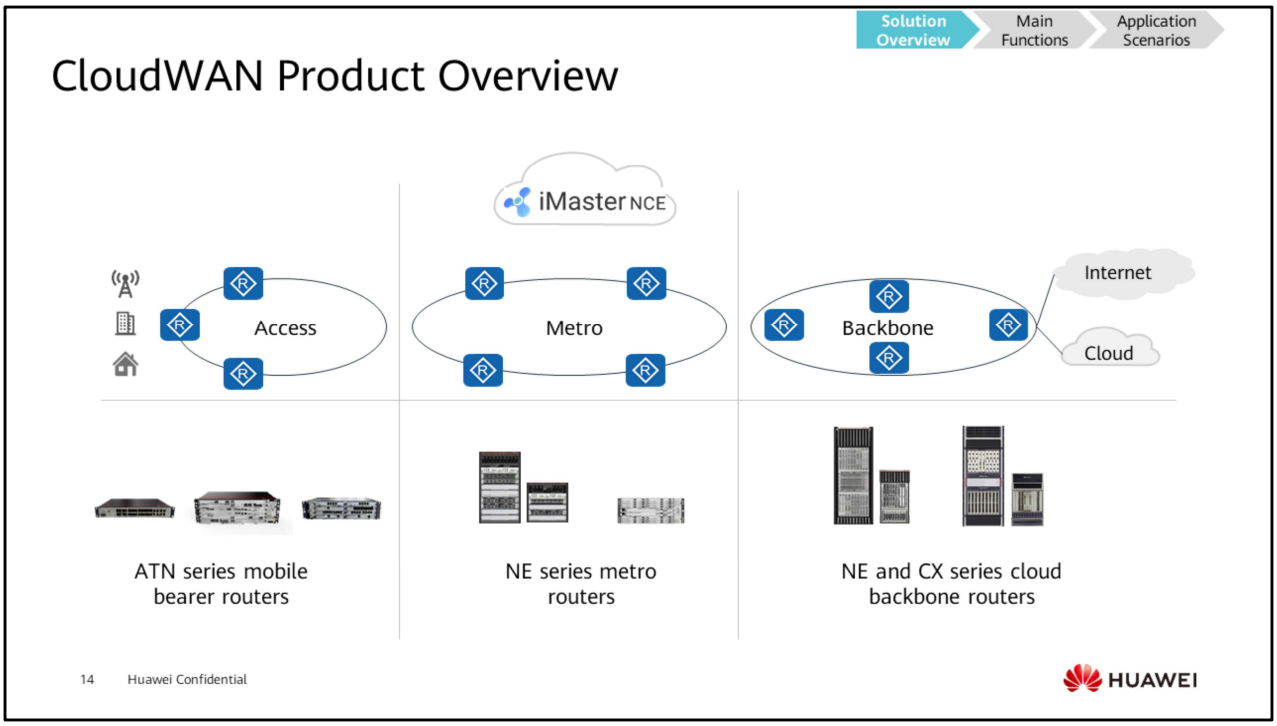
HUAWEI



- TI-LFA: Topology-Independent Loop-free Alternate
- BNG: broadband network gateway
- TTM: time to market



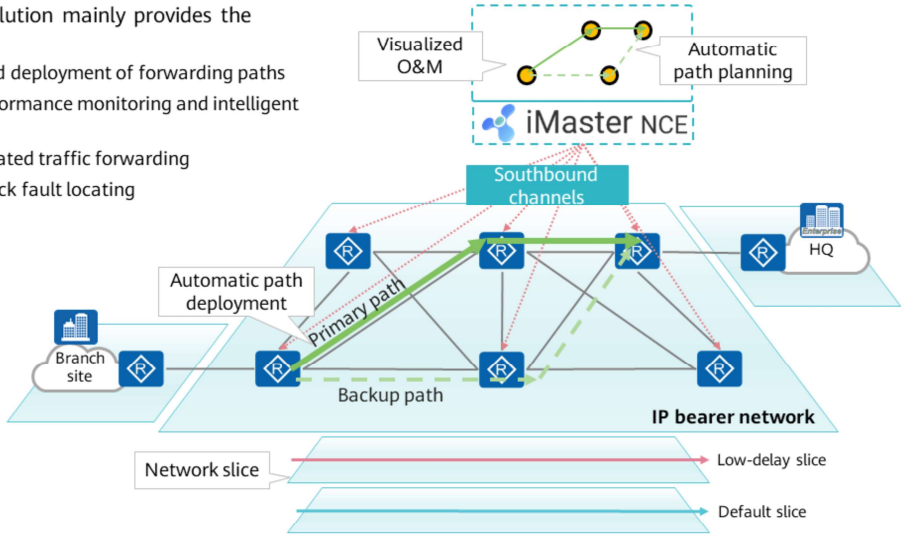
- Based on the cloud platform, iMaster NCE-IP provides three logical modules (Manager, Controller, and Analyzer) and various scenario-specific applications to achieve flexible modular deployment based on customer requirements.



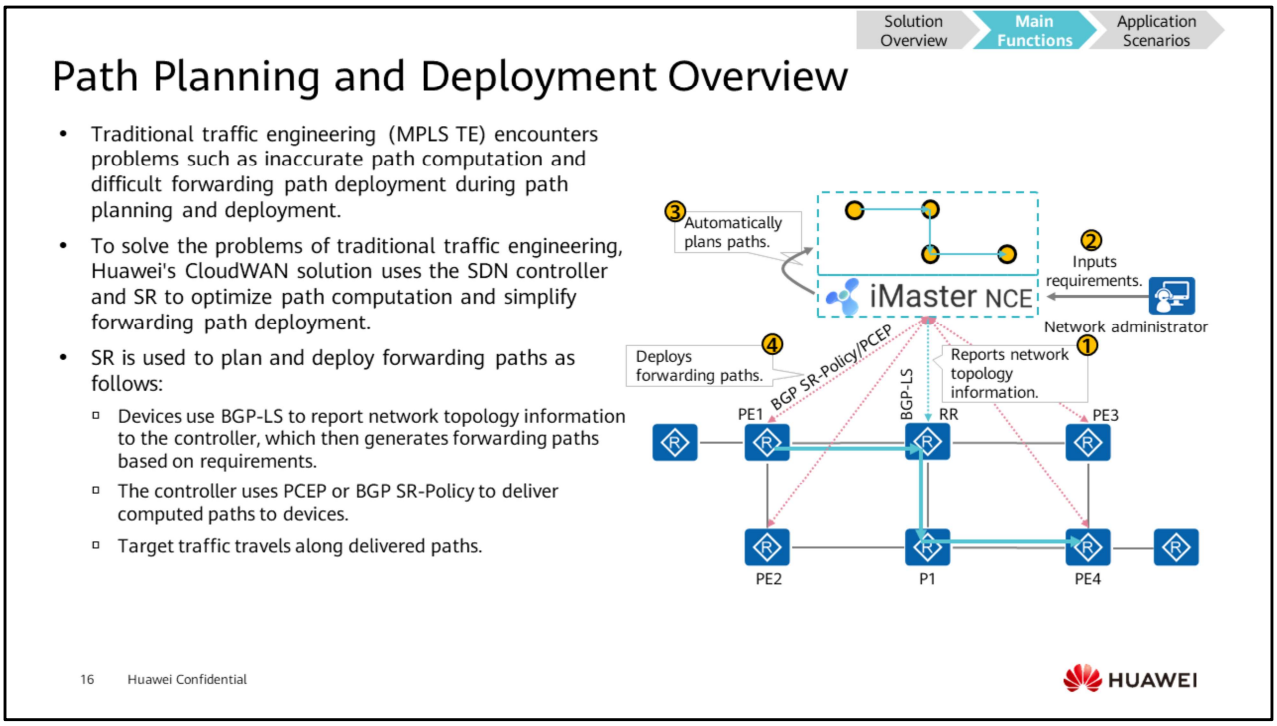
- Huawei has all-scenario SRv6 product capabilities and can provide access, metro, and backbone network routers for carriers and enterprises.

Main Functions of Huawei's CloudWAN Solution

- Huawei's CloudWAN solution mainly provides the following functions:
  - Automatic planning and deployment of forwarding paths
  - Real-time network performance monitoring and intelligent traffic optimization
  - Network slicing for isolated traffic forwarding
  - Visualized O&M for quick fault locating



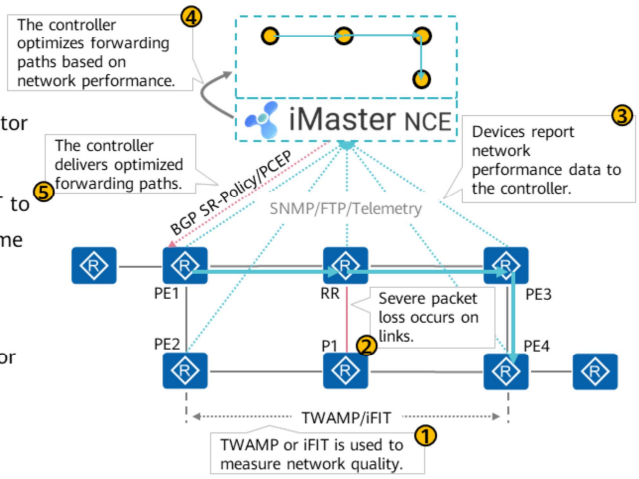
HUAWEI



- BGP-LS is a new method of collecting network topology information, which makes topology information collection simpler and more efficient. Using BGP-LS to report topology information has the following advantages:
  - Requirements on the computing capability of the upper-layer controller are lowered, and there is no requirement on the IGP capability of the controller.
  - BGP summarizes the topology information of each process or AS and directly sends the complete topology information to the controller, facilitating path selection and calculation.
  - All topology information on the network is sent to the controller through BGP, unifying the protocols for sending topology information.
- BGP SR-Policy delivers data forwarding path information to the headend through the BGP route. The headend then directs traffic to a specific SR Policy. Segment lists in SR Policies are used to guide traffic forwarding. A segment list is calculated based on a series of optimization objectives and constraints, such as delay, affinity, and SRLG.
- SR Policy is the mainstream SR implementation mode.

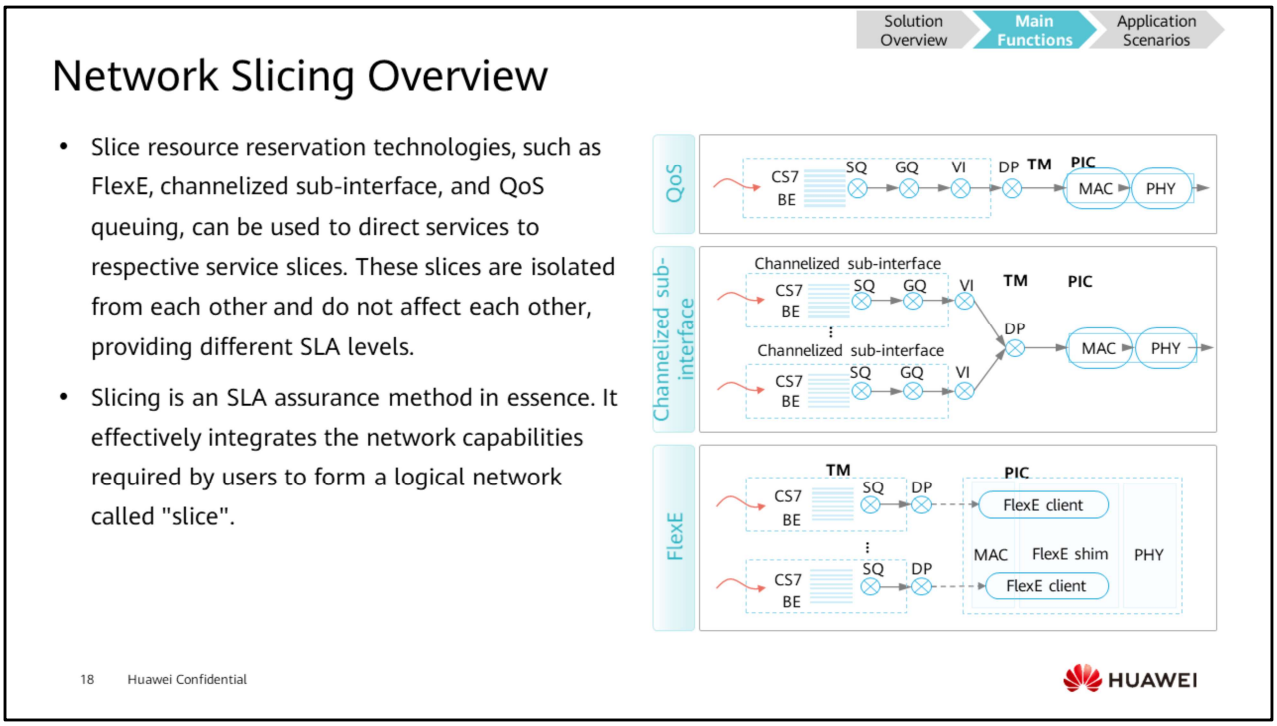
Network Performance Monitoring and Optimization

- As networks carry increasingly more enterprise services, they need to meet increasingly higher requirements. As such, a fast, flexible IP network performance measurement tool is required to monitor network performance in a timely manner.
- Huawei's CloudWAN solution uses TWAMP and iFIT to precisely and quickly measure historical and real-time network performance.
- Based on real-time network performance, the controller can optimize the network through PCEP or BGP SR-Policy.



HUAWEI

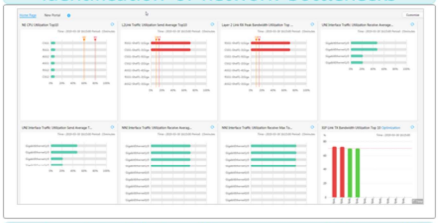
- TWAMP is a standard protocol and can be deployed on IP, MPLS, and L3VPN networks. TWAMP is easy to obtain and deploy and does not require clock synchronization.
- iFIT measures the packet loss rate and delay of service packets transmitted on an IP network to determine network performance. It is easy to deploy and provides an accurate assessment of network performance.



- Channelized sub-interfaces provide a mechanism to isolate different types of services. Different types of service traffic can be forwarded to different VLAN channelized sub-interfaces that use different dot1q encapsulation modes. Each channelized sub-interface can implement independent HQoS scheduling to isolate different types of services.
- FlexE is an Ethernet-based bearer technology for multi-rate sub-interfaces on multiple PHY links. FlexE supports bundling, channelization, and sub-rating.
- SQ: subscriber queue
- GQ: group queue
- VI: virtual interface
- DP: data plane
- TM: traffic manager
- PIC: physical interface card

O&M Overview

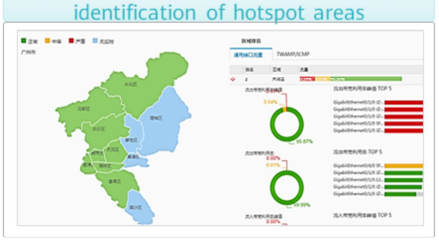
Network top N, enabling quick identification of network bottlenecks



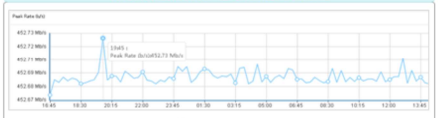
Performance topology, clearly showing quality deterioration



Regional traffic/quality map, enabling quick identification of hotspot areas



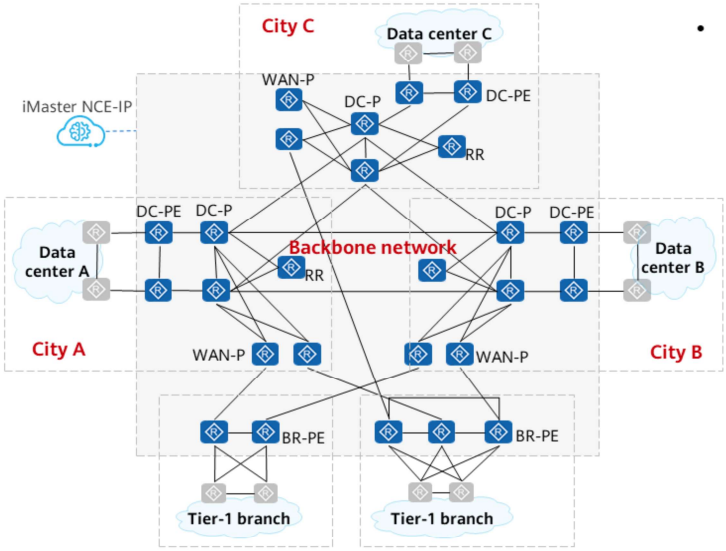
Second-level peak traffic, actually reflecting bandwidth bottlenecks



- A second-level traffic burst is detected.
- The peak traffic refers to the highest rate of traffic measured over a 15-minute period at a sampling interval of 5 seconds.

HUAWEI

Financial CloudWAN Scenario



- With the deployment of distributed cloud data centers, the cloud backbone network is elastically expanded by adding core and aggregation nodes to implement interconnection between these data centers. In addition, technologies such as SDN and SRv6 are introduced to implement functions such as intelligent network management and traffic optimization.
  - Fast service provisioning, centralized path computation, intelligent path selection
  - Intelligent network visualization, enabling real-time awareness of service status
  - Evolution to the distributed multi-city multi-center architecture
  - Capability openness (the controller provides open southbound and northbound interfaces)

HUAWEI

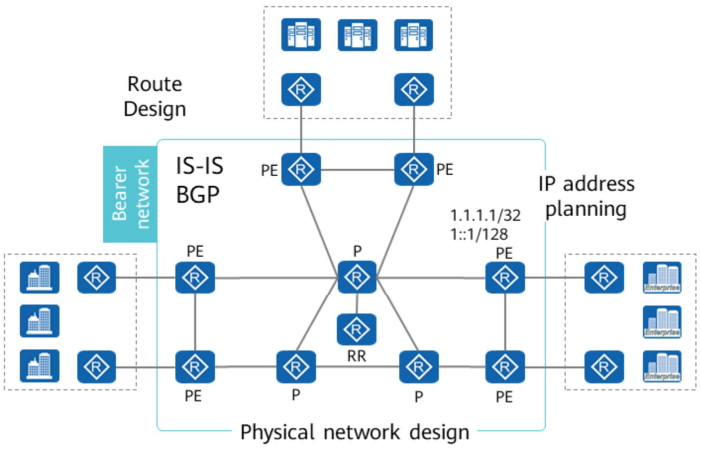
Contents

1. Current Situation and Challenges of the Enterprise Bearer WAN
2. Huawei CloudWAN Solution Overview
3. **Basic Design for the Enterprise Bearer WAN**
4. Tunnel and VPN Design for the Enterprise Bearer WAN
5. SLA and Reliability Design for the Enterprise Bearer WAN
6. Optimization and O&M Design for the Enterprise Bearer WAN

HUAWEI

Basic Design Overview for the Enterprise Bearer WAN

- Basic design for the enterprise bearer WAN includes physical network design, IP address planning, and routing design.



HUAWEI

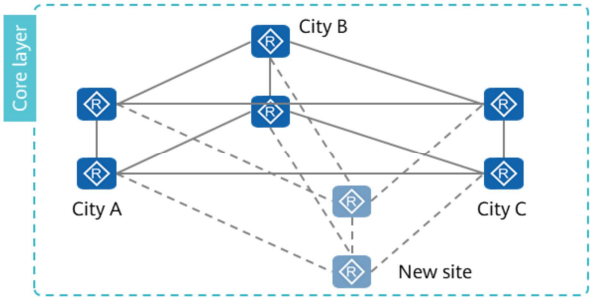
Physical Network Design for the Enterprise Bearer WAN



HUAWEI

Core Layer Design

- The core layer, as the backbone of the entire network, aggregates and forwards various service traffic. When selecting core nodes, consider the following factors:
  - Service volume: Consider the current service volume and expected service growth.
  - Physical location: Ensure that core nodes are secure, easy to obtain, and easy to maintain, as they are the key infrastructure of the bearer network.
  - Number of nodes: The core layer usually adopts the full-mesh + dual-plane architecture to ensure stability.

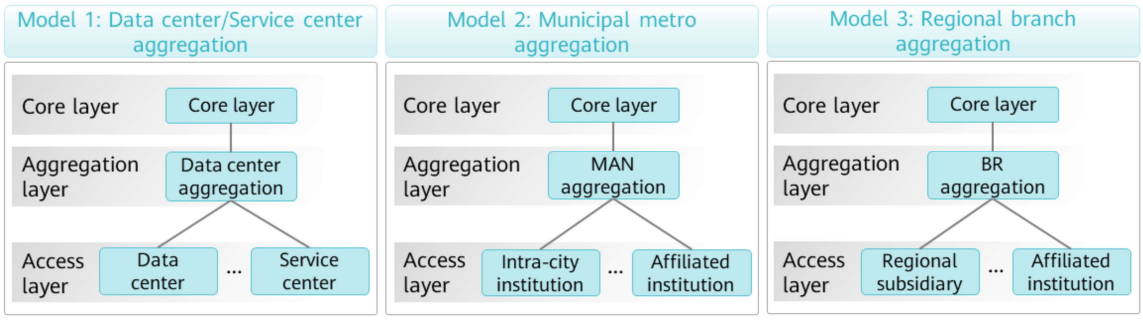


HUAWEI

- The service volume involves two aspects. The first is the service flow direction, that is, where the services concentrate. The second is the service volume size. The two aspects are complementary to each other. Generally, a greater concentration indicates a larger service volume. The core nodes must be nodes where services concentrate and the service volume is large.
- Core nodes in the same city are interconnected through WDM, and core nodes in different cities are interconnected through inter-provincial or inter-metro carrier private lines. The number of core nodes must be comprehensively considered and cannot be too large.

Aggregation Layer Design

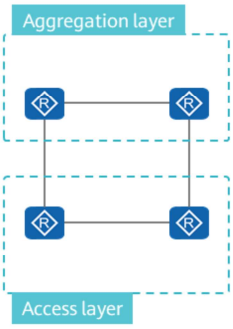
- Aggregation layer design must take into account the service types and scale on each aggregation node. In the initial phase of bearer network planning and construction, the aggregation layer can be planned based on existing enterprise services. The following three aggregation modes are available:
  - Data center aggregation (DC-PE): aggregates traffic from service units that provide services for the HQ in an enterprise. Such service units include data centers and service centers.
  - Metro aggregation (MAN-PE): aggregates the metro services of intra-city institutions and affiliated institutions.
  - Branch aggregation (BR-PE): aggregates the services of branches, provincial metro institutions, and affiliated institutions.

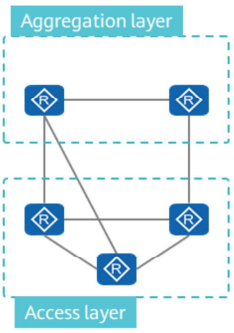


HUAWEI

Access Layer Design Factors

- When designing the access layer, you need to consider factors such as the bandwidth required by access services, required private lines, and private line prices. Moreover, you need to consider access reliability and traffic load balancing. For example:
  - You can use single-homed networking at the access layer to form dual planes working in active/standby mode. This helps improve network reliability.
  - On the active and standby planes, you can load-balance WAN traffic through service planning or routing policies.

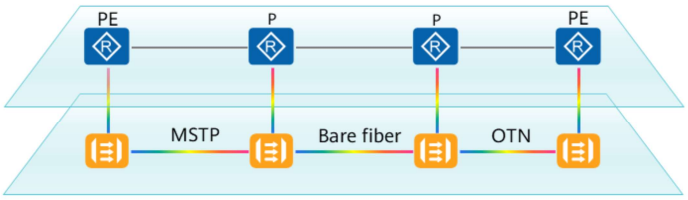




HUAWEI

Bearer Network Private Line Selection

- Only a few enterprises have the capability to build their own transmission private lines. Most enterprises need to purchase transmission private lines from carriers and build their own bearer networks over these transmission private lines.
- Currently, the transmission private lines provided by carriers are mainly MSTP and OTN private lines:
  - MSTP private lines or MPLS VPNs can be used to transmit services from branches to the access layer of the bearer network.
  - MSTP private lines can be used between the access and aggregation layers of the bearer network.
  - If the aggregation layer and core layer are in different equipment rooms, MSTP/OTN private lines can be used between them.
  - MSTP/OTN private lines can be used between core-layer devices. DWDM private lines can also be used between core-layer devices if bare fibers are available.



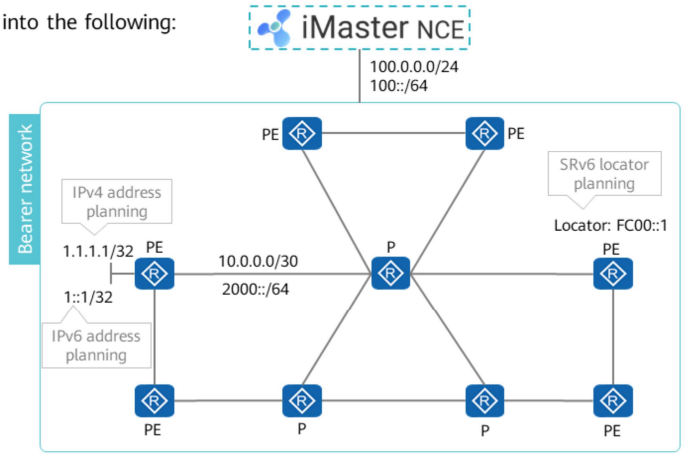
IP Address Planning Roadmap for the Enterprise Bearer WAN



HUAWEI

IP Address Planning Overview

- With the development of services, increasingly more networks are deployed as IPv4/IPv6 dual-stack networks. Dual-stack address planning is essential to dual-stack networks.
- IP address planning is generally classified into the following:
  - IPv4 address planning
  - IPv6 address planning
  - SRv6 locator planning



HUAWEI

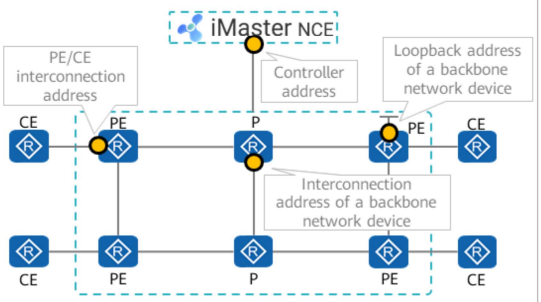
IPv4 Address Design Rules

- For the convenience of service inheritance and network management, the backbone network retains interconnection IPv4 addresses. IPv4 address allocation must comply with the existing IPv4 address allocation specifications of the customer.

IPv4 address design rules

- **Uniqueness**
  - Hosts on the backbone network must use unique IP addresses. Try to allocate a different address to each host even if they support VPN address overlapping.
- **Contiguity**
  - Routes with contiguous addresses can be easily summarized on a hierarchical network, reducing the routing table size and accelerating route calculation.
- **Scalability**
  - Addresses need to be reserved at each layer to ensure contiguity of addresses when the network is expanded.
- **Meaningfulness**
  - A well-planned IP address denotes the device to which the IP address belongs.

IPv4 address allocation range on the bearer WAN

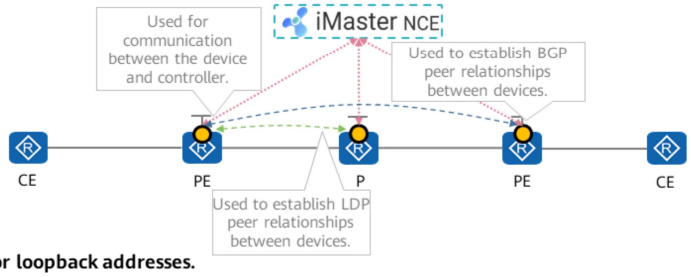


HUAWEI

Loopback Address Design Rules

- The loopback address of each router plays an important role in the normal running of the entire network. Therefore, a unified, dedicated address space must be used for the allocation and management of loopback addresses on each router.

Loopback address design rules and application scenarios



- **Use 32-bit masks for loopback addresses.**
  - Allocate loopback addresses based on physical locations and reserve sufficient address space.
  - Allocate loopback addresses for the same geographical location by plane. If there are two planes, allocate IP addresses to planes 1 and 2 in sequence.

HUAWEI

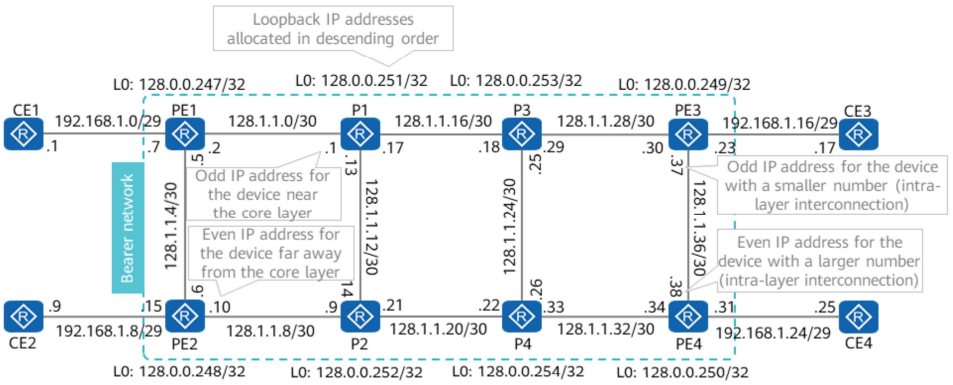
IPv4 Address Planning Suggestions

- It is recommended that the bearer WAN use a dedicated subnet. The allocated addresses include the management addresses and internal interconnection addresses of devices and access addresses of the bearer network. The following table describes the address allocation plan:
  - Internal interconnection addresses: Use a 30-bit mask for IP addresses used for internal interconnection between devices on the bearer network.
  - Access addresses: The access-layer device interfaces on the bearer network need to connect to the original network. Therefore, it is recommended that the IP addresses of these interfaces be allocated based on the original planning. For example, use a 29-bit mask for these IP addresses.
  - Address allocation sequence: Allocate interconnection addresses in ascending order and loopback addresses in descending order.
  - Interconnection between devices of the same layer: Assign an odd address to the device with a smaller number and an even address to the device with a larger number.
  - Interconnection between devices at different layers: Assign an odd address to the device close to the network core and an even address to the device far away from the network core.

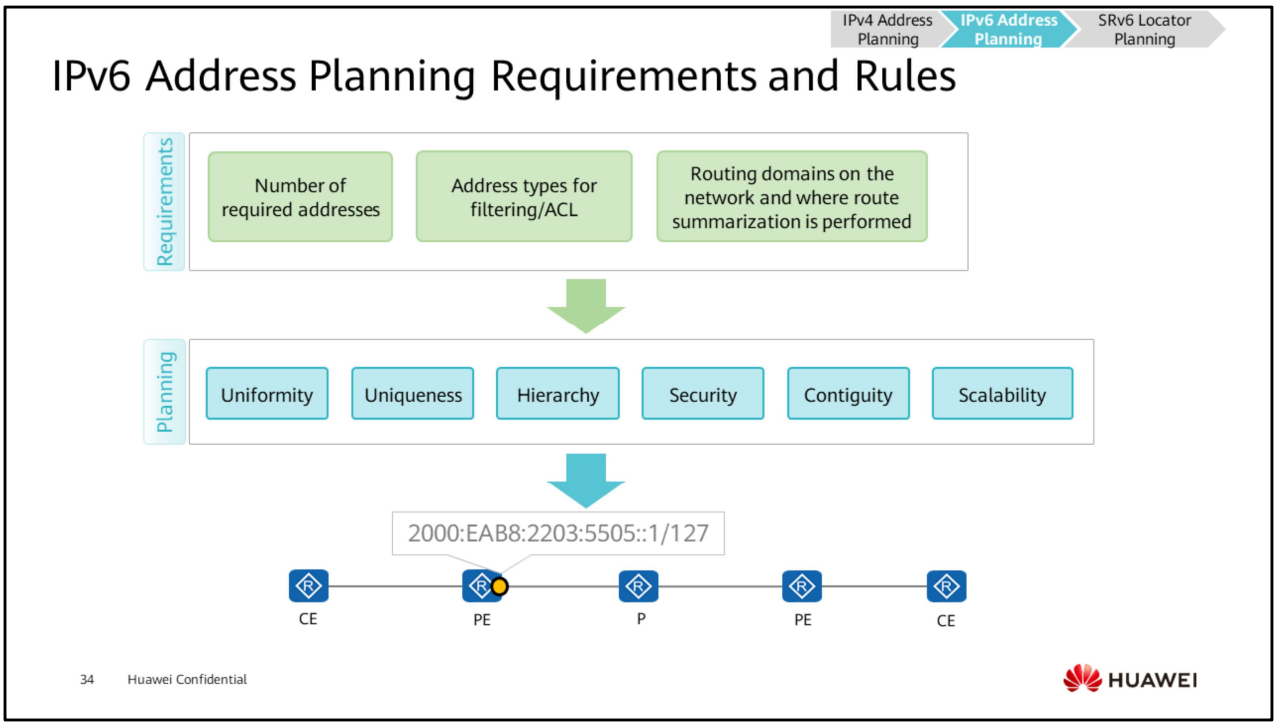
HUAWEI

Typical IPv4 Address Planning

- IPv4 addresses can be planned based on planning suggestions or the customer's planning rules.



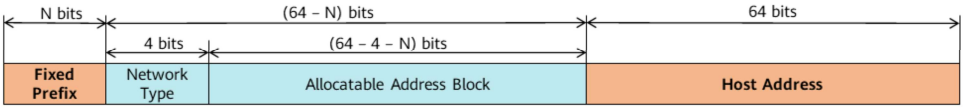
HUAWEI



- Uniformity: All IP addresses on the entire network are planned in a unified manner, including service addresses, platform addresses, and network addresses.
- Uniqueness: Each address is unique throughout the entire network.
- Hierarchy: The massive IPv6 address space poses higher requirements on the route summarization capability. The primary task of IPv6 address planning is to reduce network address fragments, enhance the route summarization capability, and improve the network routing efficiency.
- Security: Services with shared attributes have the same security requirements. Mutual access between services needs to be controlled. Services with shared attributes are allocated with addresses in the same address space, which facilitates security design and policy management.
- Contiguity: IPv6 addresses in an IPv6 address segment must be contiguous to prevent address wastes.
- Scalability: IP addresses must be planned and allocated based on network development requirements to reserve space for future capacity expansion. The addition of a small number of subnets does not require large-scale architecture or policy adjustment.

Structured IPv6 Address Planning

- Structured planning: The massive IPv6 address space poses higher requirements on the route summarization capability. Hierarchical IPv6 address allocation is recommended, which helps reduce the sizes of routing tables through route summarization. In addition, the subnet field of an IPv6 address can be divided into several independent fields to make the address more readable. (If possible, it is recommended that the subnet field be divided by 4 bits.) The following is a planning example:

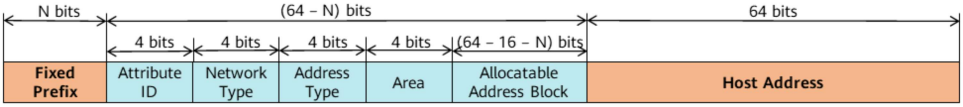


  - Fixed Prefix: indicates the fixed prefix applied for.
  - Network Type: indicates the type of a network. For example, 0 indicates a backbone network, 1 indicates a data center, and 2 to F are reserved.

HUAWEI

IPv6 Address Planning Suggestions

- The user-defined part of an IPv6 address needs to be planned based on its summarization characteristics. Fields that are easy to summarize, such as the address space and area, should be placed in the left-most part. Excessive layering results in strong coupling between services and address planning, which hinders subsequent service development and reduces address space utilization. Therefore, you need to determine the number of fields based on site requirements.

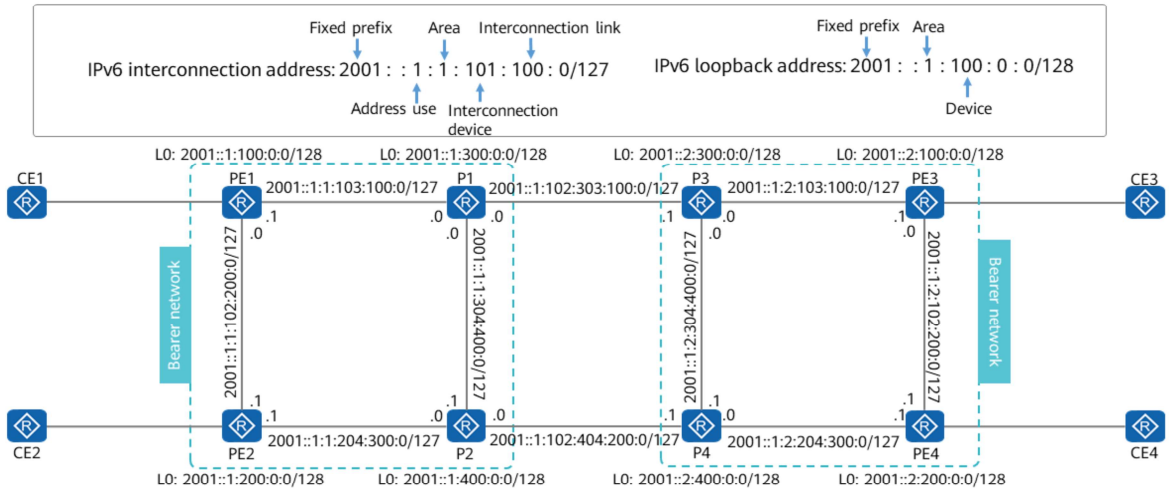


  - Fixed Prefix: indicates a fixed-length prefix applied for by an enterprise from an address allocation organization.
  - Subnet:
    - Attribute ID: is used to distinguish address types. It is used for level-1 address classification.
    - Network Type: identifies the type of a network.
    - Address Type: identifies the type of an address on the network.
    - Area ID: identifies an area on the network.
    - Allocatable Address Block: is reserved for future address allocation.
  - Interface Address: indicates the last 64 bits of an IP address. It is equivalent to the host ID in an IPv4 address.

HUAWEI

Typical IPv6 Address Planning

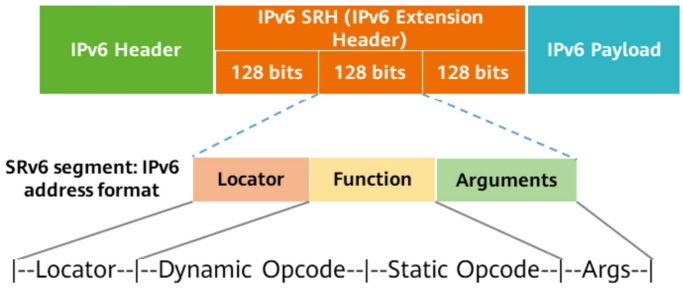
- IPv6 addresses can be planned based on planning suggestions or the customer's planning rules.



HUAWEI

SRv6 SID Overview

- An SRv6 SID usually consists of three fields: Locator, Function, and Arguments. The SRv6 SID is expressed in the *Locator:Function:Arguments* format. Note that the total length (Locator + Function + Arguments) is less than or equal to 128 bits. If the total length is less than 128 bits, the reserved bits are padded with 0s.
- If the Arguments field does not exist, the format is *Locator:Function*. The Locator field occupies the most significant bits of an IPv6 address, and the Function field occupies the remaining part of the IPv6 address.

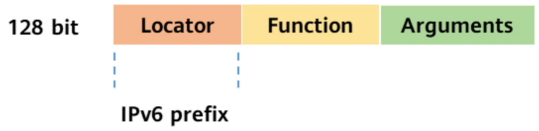


HUAWEI

- The locator is an IPv6 network segment. All IPv6 addresses in this network segment can be allocated as SRv6 SIDs. After a locator is configured for a node, the system generates a locator route. The node can be located based on the locator route. In addition, all SIDs advertised by the node can reach the node through the locator route.
- The Function field is also called opcode, which can be dynamically allocated using an IGP or statically configured using the **opcode** command. When configuring a locator, you can use the **static** *static-length* parameter to specify the length of the static segment, which determines the number of static opcodes that can be configured in the locator. When an IGP dynamically allocates opcodes, it applies for opcodes outside of the static segment range to ensure that SRv6 SIDs do not conflict.
- The Args field is determined by the **args** *args-length* parameter. The **Args** field is optional in SRv6 SIDs and is determined by the command configuration.

SRv6 Locator Overview

- The Locator field identifies the location of a network node, and is used for other nodes to route and forward packets to this identified node so as to implement network instruction addressing.
- A locator has two important characteristics: routable and aggregatable. After a locator value is configured for a node, the system generates a locator route and propagates the route throughout the SR domain using an IGP. A device advertising a locator route can be located by other devices on the same network based on the received locator route information, and all the SRv6 SIDs advertised by the device can be reached over the route.
- It is recommended that the same mask be used for the locators of all NEs on the entire network. After the mask is set, locators can be allocated in ascending order for the core, aggregation, and access layers in sequence.

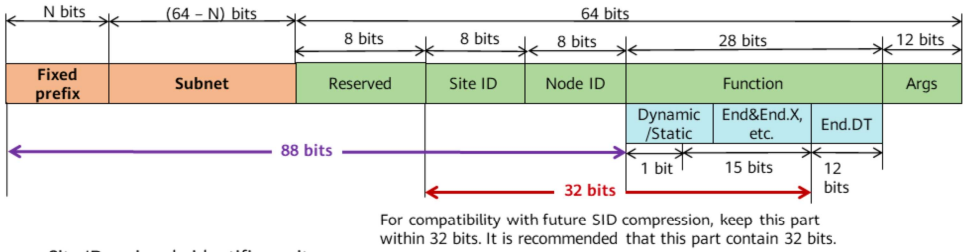


HUAWEI

IPv4 Address Planning | IPv6 Address Planning | SRv6 Locator Planning

SRv6 Locator Planning Suggestions

- SRv6 locator addresses are allocated from the IPv6 host address field. Using hierarchical address allocation and reserving some address spaces are recommended for future expansion. An example is provided as follows:



For compatibility with future SID compression, keep this part within 32 bits. It is recommended that this part contain 32 bits.

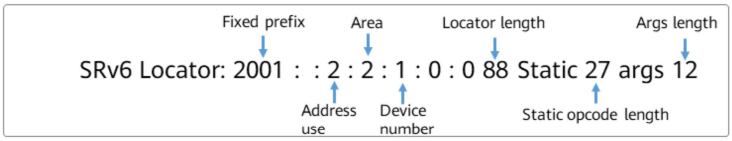
- Site ID: uniquely identifies a site.
- Node ID: uniquely identifies a device at a site.
- Function: indicates the Function field in an SRv6 SID.
  - Dynamic/Static: indicates whether a SID is a dynamic or static SID. 0 indicates a static SID (it is recommended that the End, End.X, and OAM-related SIDs be statically allocated). 1 indicates a dynamic SID (if many VPNs exist or VPNs change frequently, it is recommended that service SIDs, such as End.DT SIDs, be dynamically allocated).
  - End&End.X, etc: indicates the type of a SID. 0x000[1–F] indicates an End SID, and 0X1[peer site ID]X indicates an End.X SID.
- Args: indicates the parameter field in an SRv6 SID.

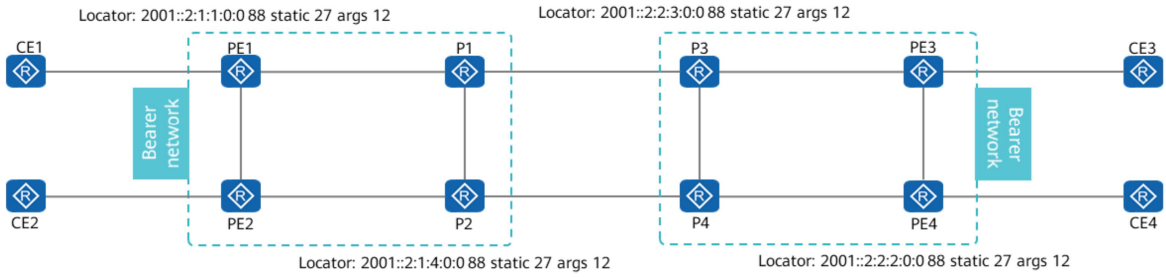
HUAWEI

- End.DT SIDs can be classified into End.DT4 SIDs and End.DT6 SIDs.
  - An End.DT4 SID (PE endpoint SID) identifies an IPv4 VPN instance on a network.
  - An End.DT6 SID (PE endpoint SID) identifies an IPv6 VPN instance on a network.
- In SRv6, End.OP SIDs are used to implement operation, administration and maintenance (OAM).
  - End.OP SIDs are mainly used in ping/tracert scenarios.

Typical SRv6 Locator Planning

- In SRv6 Locator Planning, it is recommended that the locator length not exceed 88 bits. Moreover, try to statically configure SRv6 SIDs, so that SRv6 SIDs remain unchanged after devices go online again. This facilitates management.





HUAWEI

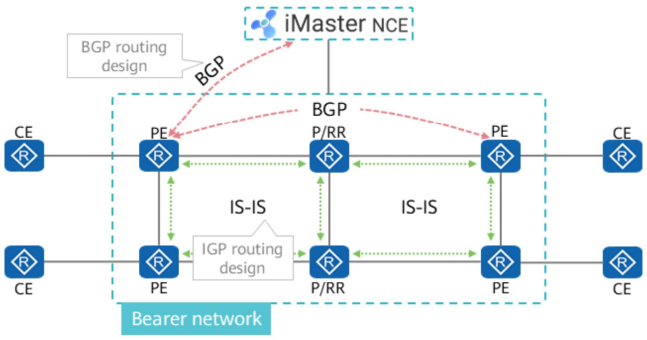
Routing Design Roadmap for the Enterprise Bearer WAN



HUAWEI

Routing Design Overview

- Routing protocols are generally classified into IGP and BGP. These two types of protocols have different functions and therefore are designed for different purposes.
- Routing design can be classified into the following types:
  - IGP routing design
  - BGP routing design



HUAWEI

IGP Overview

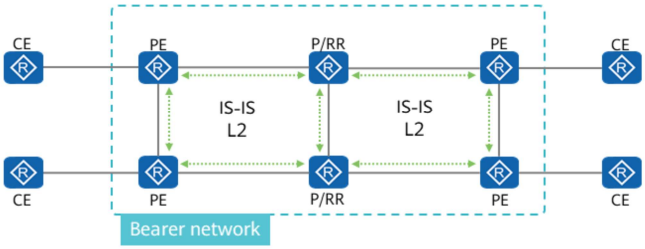
- On a bearer WAN, an IGP functions as a basic support protocol to collect and flood Layer 3 topology information on the entire network, and works with protocols such as TWAMP and iFIT to collect network status information, such as link delay.
- Generally, OSPF or IS-IS can be used on the backbone network for route reachability. However, the application scenarios of the two protocols are different to some extent.

	IS-IS	OSPF	Remarks
Network scale	Large	Small	Some devices support a maximum of 4K IS-IS routes and a maximum of 2K OSPF routes in a single area (with a single level).
IPv6 support	TLV-based packets are used, and no extra protocol needs to be independently deployed.	OSPFv3 needs to be independently deployed.	When OSPF supports both IPv4 and IPv6, OSPFv2 and OSPFv3 neighbor relationships need to be configured separately. The exchange of large numbers of protocol packets consumes a lot of resources.
SRv6 support	Related standards or drafts are available.	Related standards or drafts are available.	The standardization process of OSPF lags behind that of IS-IS in terms of SRv6 support. Not all vendors support SRv6-oriented OSPF extensions.
SR-MPLS support	Supported. The related functions are complete.	Supported. The related functions are basically complete.	

HUAWEI

IGP Route Planning

- Compared with OSPF, IS-IS supports SR-MPLS/SRv6 in a more comprehensive and scalable manner. Considering network expansion brought by enterprise service growth, large enterprise bearer networks generally start to use IS-IS. Therefore, it is recommended that IS-IS be preferentially used as the IGP for bearer networks.



- IS-IS: One IS-IS process is configured on the entire network, and an IS-IS level-2 area is configured in E2E mode.

HUAWEI

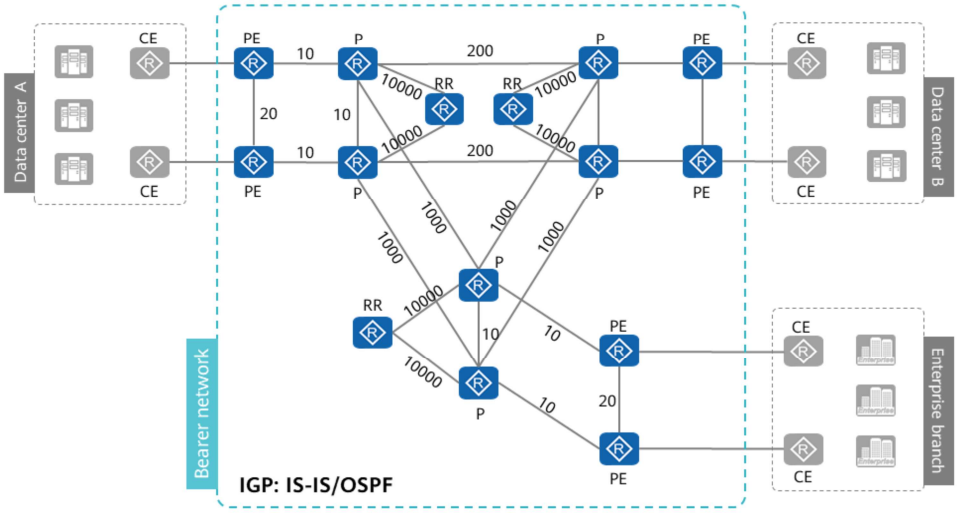
- It is recommended that some IGP parameters be set as follows:
  - IGP process ID: It is recommended that the IS-IS or OSPF process ID of a device on the backbone network be the same as the BGP AS number.
  - IS-IS NET: The recommended format is aa.bbbb.cccc.dddd.00. The loopback0 address of the device is used for NET derivation. For example, if the loopback0 address is 21.231.232.1, then the derived NET is 21.0231.0232.0001.00.
  - OSPF router ID: The global router ID is used. Generally, the router ID is the same as the loopback0 address.
  - Interface type: To speed up convergence, all interfaces are of the P2P type.
  - Route advertisement: IGP is mainly used to ensure the reachability of internal addresses on the WAN. Therefore, IGP advertises only interconnection interface addresses and device management addresses.

IGP Metric Planning

- The bearer network transmits different types of service traffic. When deploying an IGP on the bearer network, properly plan route metric to maximize bandwidth utilization, improve service quality, and ensure service reliability.
- IGP metric design rules:
  - Ensure that the metric of access-layer links is lower than that of aggregation-layer links.
  - Ensure that the metric of aggregation-layer links is lower than that of core-layer links.
  - Ensure that the metric of links between data centers is lower than that of WAN links between branches and data centers.
  - Ensure that inter-plane traffic between data centers preferentially traverses across planes through core nodes.
  - If a standalone RR is deployed, ensure that the metric of the link between the RR and core P is set to the maximum value (the RR only reflects routing information and does not forward data).

HUAWEI

IGP Metric Planning Suggestions

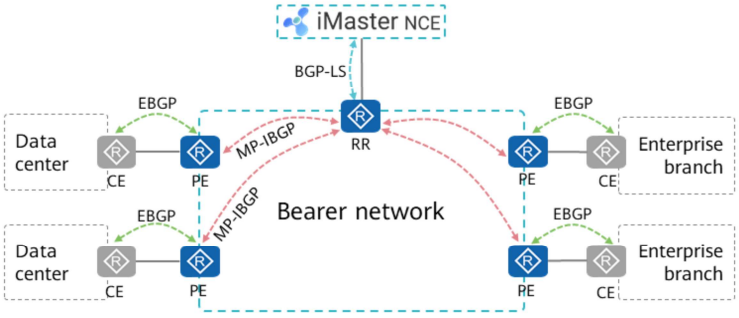


HUAWEI

IGP Route Planning | BGP Route Planning

Overall BGP Planning

- After an enterprise bearer WAN is used to carry services that were used to be carried by multiple networks, it needs to use BGP/MPLS IP VPN to isolate these services.
- Generally, IBGP runs on the bearer network, and EBGP runs between the bearer network and other ASs (such as data centers and enterprise branches). PEs use BGP policies to control the transmission of VPN routes between ASs, achieving complex access control.
- If a controller is deployed on the bearer WAN, BGP-LS must be deployed between the controller and RR.

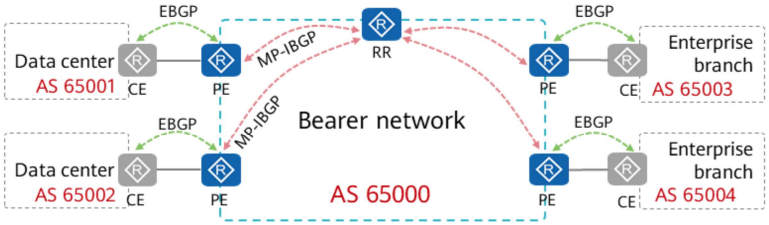


HUAWEI

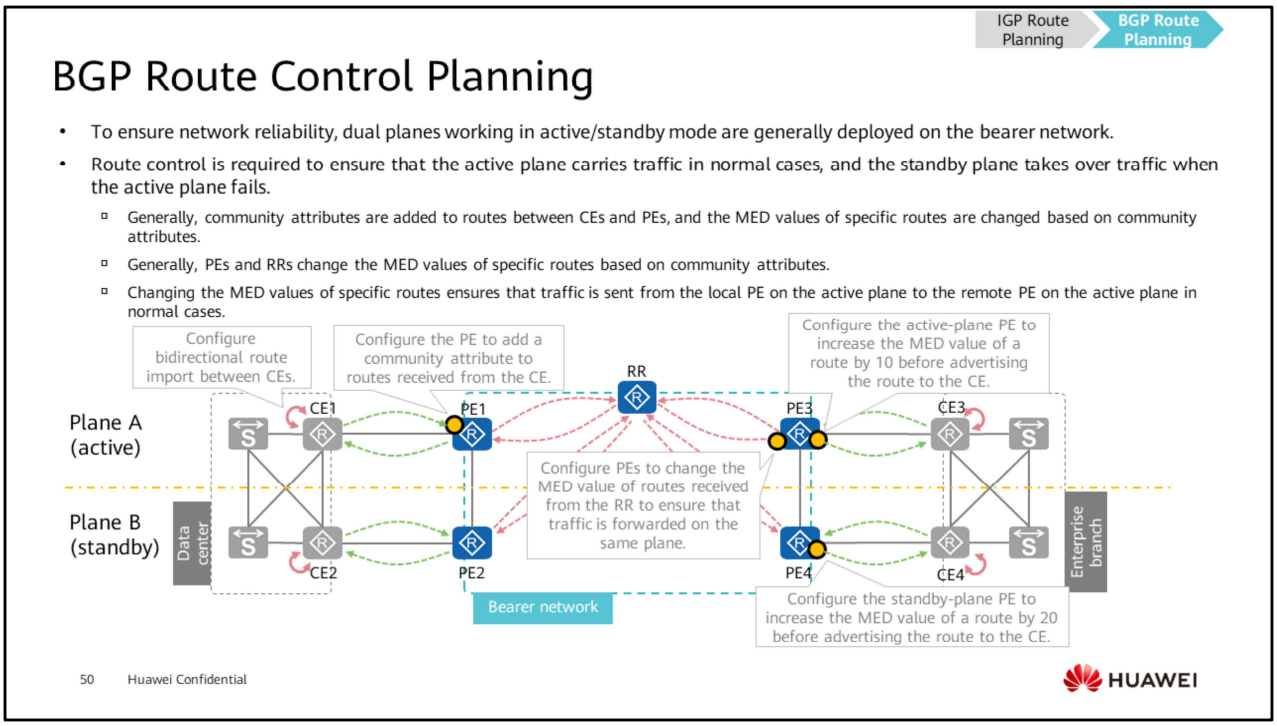
- In BGP peer relationship establishment, IBGP peer relationships are established using Loopback0 addresses, and EBGP peer relationships are established using interface addresses.
- AS: The bearer WAN can be classified as an independent AS.
- IBGP: PEs use loopback addresses to establish IBGP peer relationships with all RRs and use MP-IBGP to exchange VPN routes.
- EBGP: PEs use interface IP addresses to establish EBGP peer relationships with CEs. In inter-AS VPN route exchange scenarios, Option A is generally used.
- BGP-LS: The controller establishes BGP-LS peer relationships with all RRs to collect logical topology information on the backbone network.
- Deploy independent RRs and establish IBGP peer relationships for RRs on the backbone network.
- In addition to EBGP, IGPs such as OSPF, IS-IS, and RIP can also be used between PEs and CEs on the bearer network. Static routes can also be used to meet the requirements of flexible access in various scenarios.

BGP AS Planning

- An enterprise network usually uses a private AS number ranging from 64512 to 65534 during BGP deployment.
- It is recommended that one AS be deployed as the high-speed forwarding core of the entire bearer network, independent ASs be deployed for data centers and enterprise branches in different regions, and EBGP peer relationships be established between these ASs and the bearer network AS.



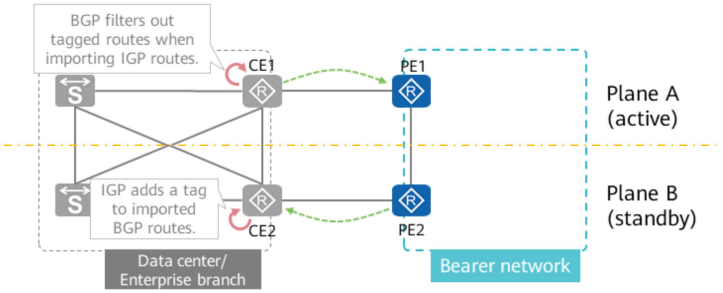
HUAWEI



- PE3 changes the MED value to 100 for the route whose next hop is PE1 (a PE on the same plane) and changes the MED value to 200 for the route whose next hop is PE2 (a PE on a different plane).
- PE4 changes the MED value to 100 for the route whose next hop is PE2 (a PE on the same plane) and changes the MED value to 200 for the route whose next hop is PE1 (a PE on a different plane).
- If PE1 and PE2 on the left learn the same VPN route and advertise the route to PE3 and PE4 on the right through the RR, PE3 and PE4 preferentially select the VPN route on the same plane as them. After the route is advertised to the CE, traffic from the CE preferentially travels along the route advertised by PE3 (because the MED value of the route advertised by PE3 is only increased by 10).

Design for BGP Routing Loop Prevention and Sub-optimal Route Prevention

- Generally, data center networks and branch networks do not run BGP. Their IGPs need to import routes learned by EBGP.
- When an IGP imports EBGP routes, it needs to add a tag to the routes, so that the IBGP peer can filter out imported BGP routes based on tags when importing local IGP routes to BGP.



HUAWEI

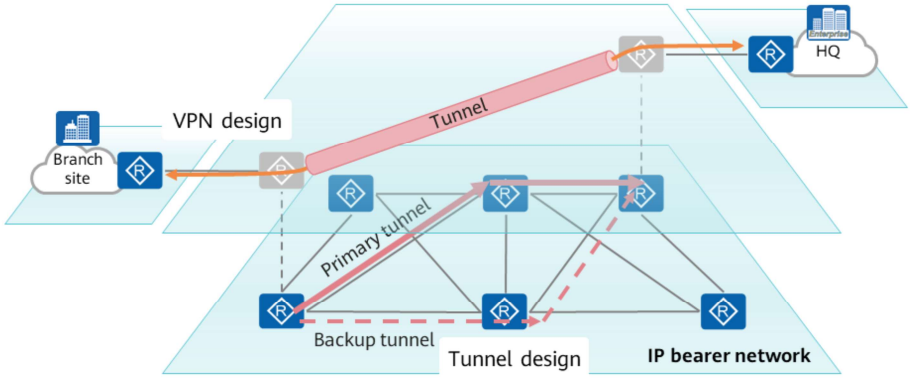
Contents

1. Current Situation and Challenges of the Enterprise Bearer WAN
2. Huawei CloudWAN Solution Overview
3. Basic Design for the Enterprise Bearer WAN
4. **Tunnel and VPN Design for the Enterprise Bearer WAN**
5. SLA and Reliability Design for the Enterprise Bearer WAN
6. Optimization and O&M Design for the Enterprise Bearer WAN

HUAWEI

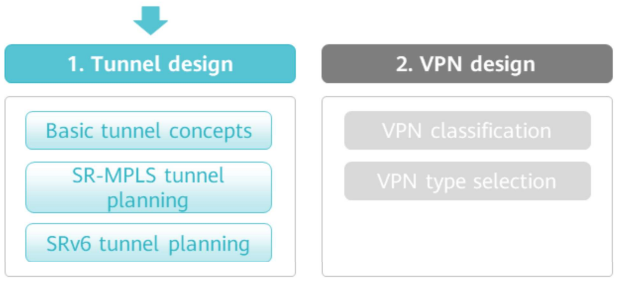
Tunnel and VPN Design Overview for the Enterprise Bearer WAN

- Enterprises usually use VPN to isolate services and SR to establish tunnels for traffic optimization and path planning.
- VPN traffic is carried over tunnels to isolate enterprise services while ensuring service quality.



HUAWEI

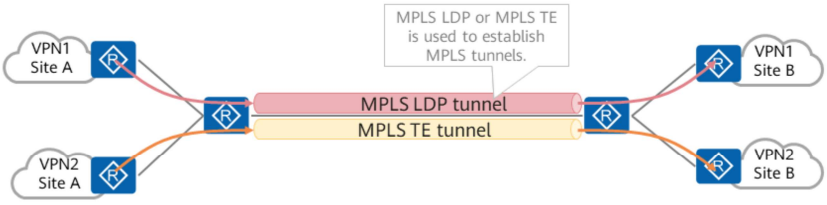
Tunnel Design Roadmap for the Enterprise Bearer WAN

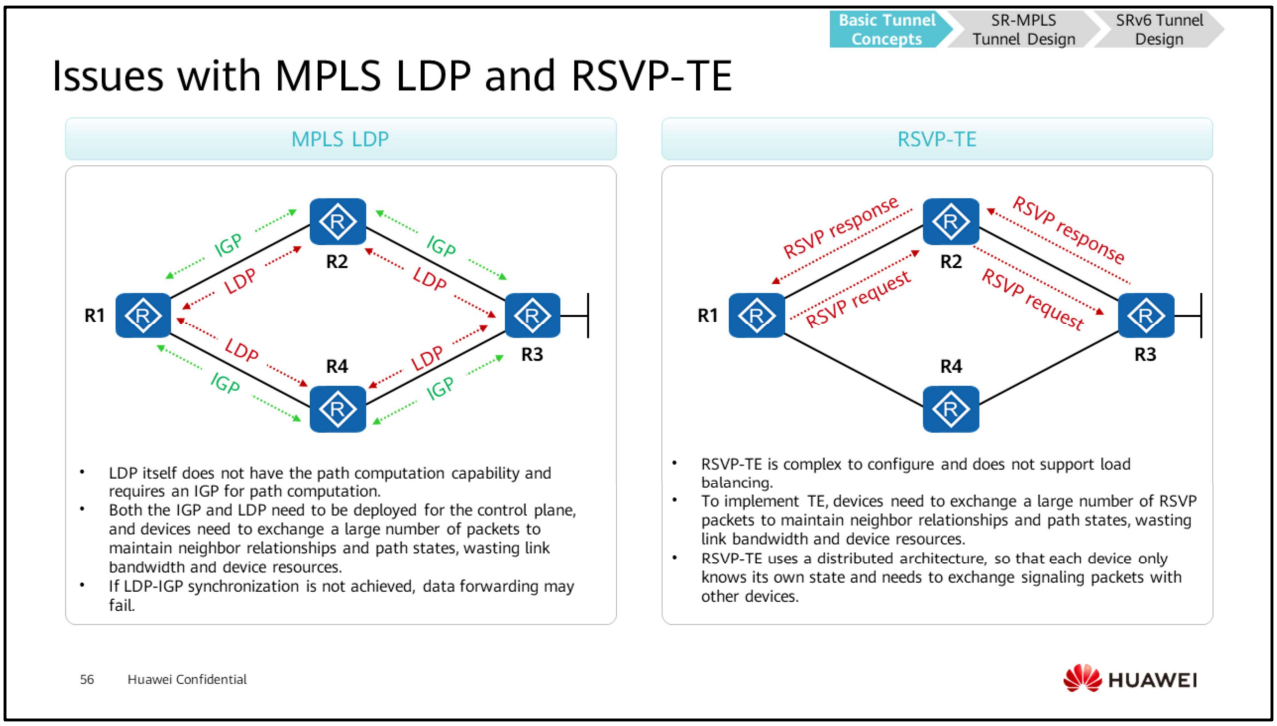


HUAWEI

Tunneling Technology Overview

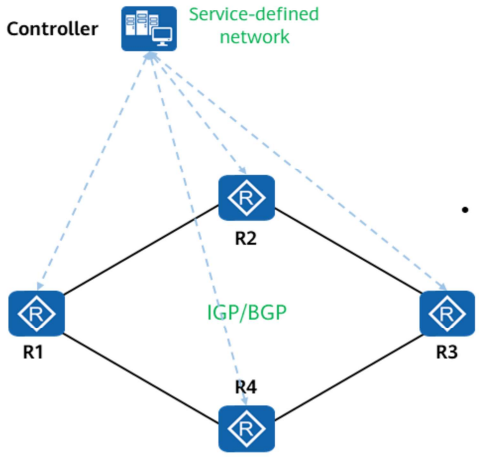
- The tunneling technology uses one protocol to encapsulate the packets of another protocol, and the carrier protocol itself can be encapsulated or carried by other protocols.
- For example, BGP/MPLS IP VPN commonly used on enterprise networks uses MPLS LDP or MPLS TE tunnels to carry VPN services.





- MPLS is a tunneling technology that guides data forwarding in essence and has complete tunnel creation, management, and maintenance mechanisms. The preceding mechanisms are driven by network operation and management requirements, not by applications.

Solution Provided by SR



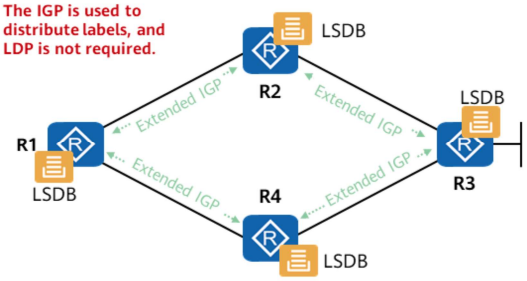
- Simplifying protocols and extending existing protocols
  - The extended IGP/BGP supports label distribution. Therefore, LDP is not required on the network, achieving protocol simplification. In addition, devices require only software upgrades instead of hardware replacement, protecting investments on the live network.
  - The source routing mechanism is introduced. The forwarding policy is instantiated into a segment list on the ingress to control the forwarding path of service traffic.
- Enabling networks to be defined by services
  - After an application raises requirements (e.g. delay, bandwidth, and packet loss rate), the controller collects information (e.g. network topology, bandwidth usage, and delay) and computes an explicit path according to the requirements.

HUAWEI

SR-MPLS Classification

SR-MPLS BE

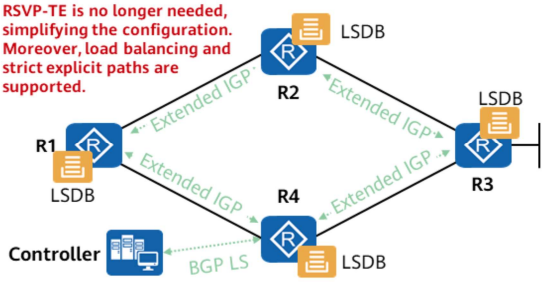
The IGP is used to distribute labels, and LDP is not required.



- An extended IGP has the label distribution capability. An LSDB can be formed based on the SRGB, prefix SID, link state, and other information distributed through the IGP, and the SPF algorithm can be used to compute the shortest forwarding paths based on labels. Therefore, LDP is no longer required on the network.
- Because only the IGP/BGP is used, the network does not have traffic blackholes and does not need to maintain LDP peer relationships.

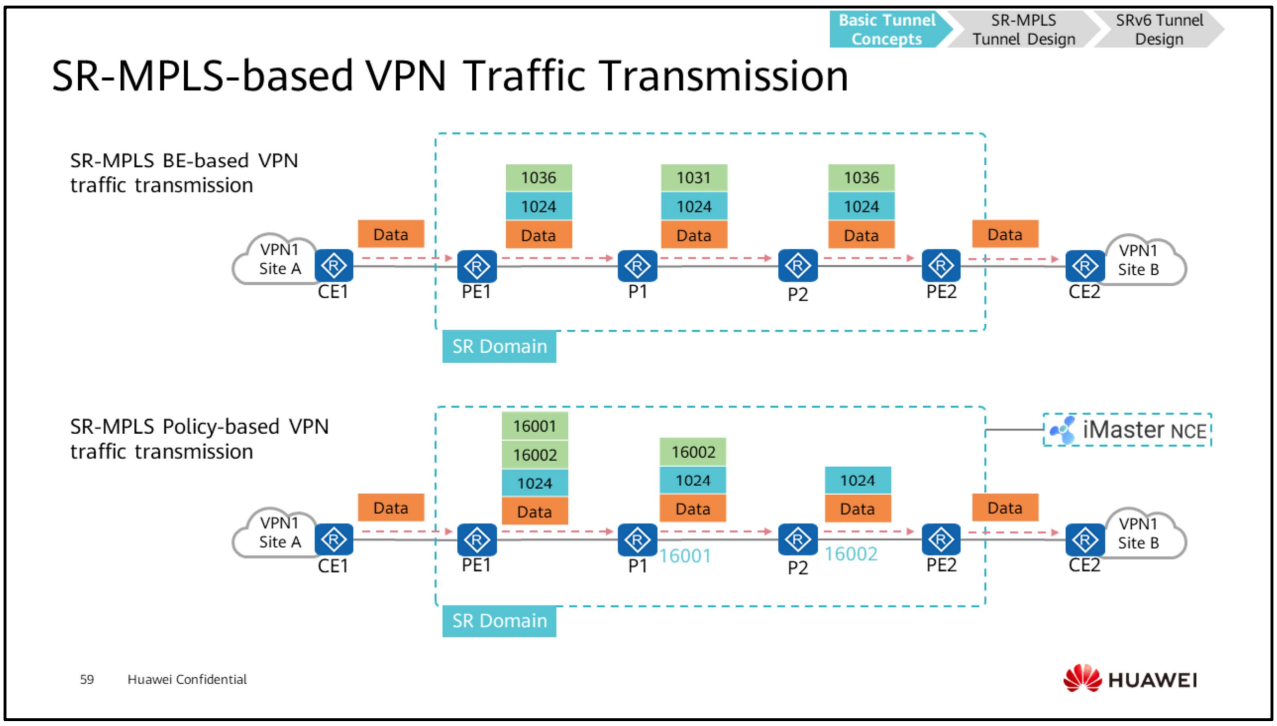
SR-MPLS Policy

RSVP-TE is no longer needed, simplifying the configuration. Moreover, load balancing and strict explicit paths are supported.



- Extended BGP/IGP sends the LSDB that carries label information to the controller, which then globally computes paths. There is no need to use RSVP-TE packets to apply for paths or maintain path status.
- Node SIDs and adjacency SIDs are used to implement load balancing and strict explicit paths.
- Path computation and service delivery are performed by the controller, simplifying configuration.

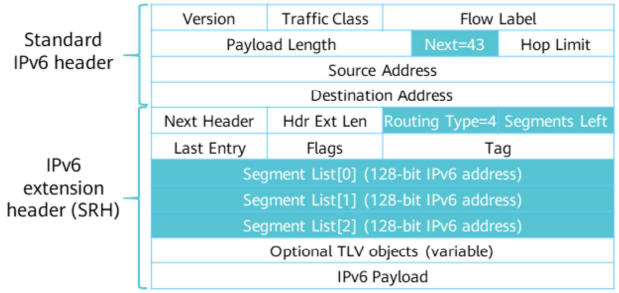
HUAWEI



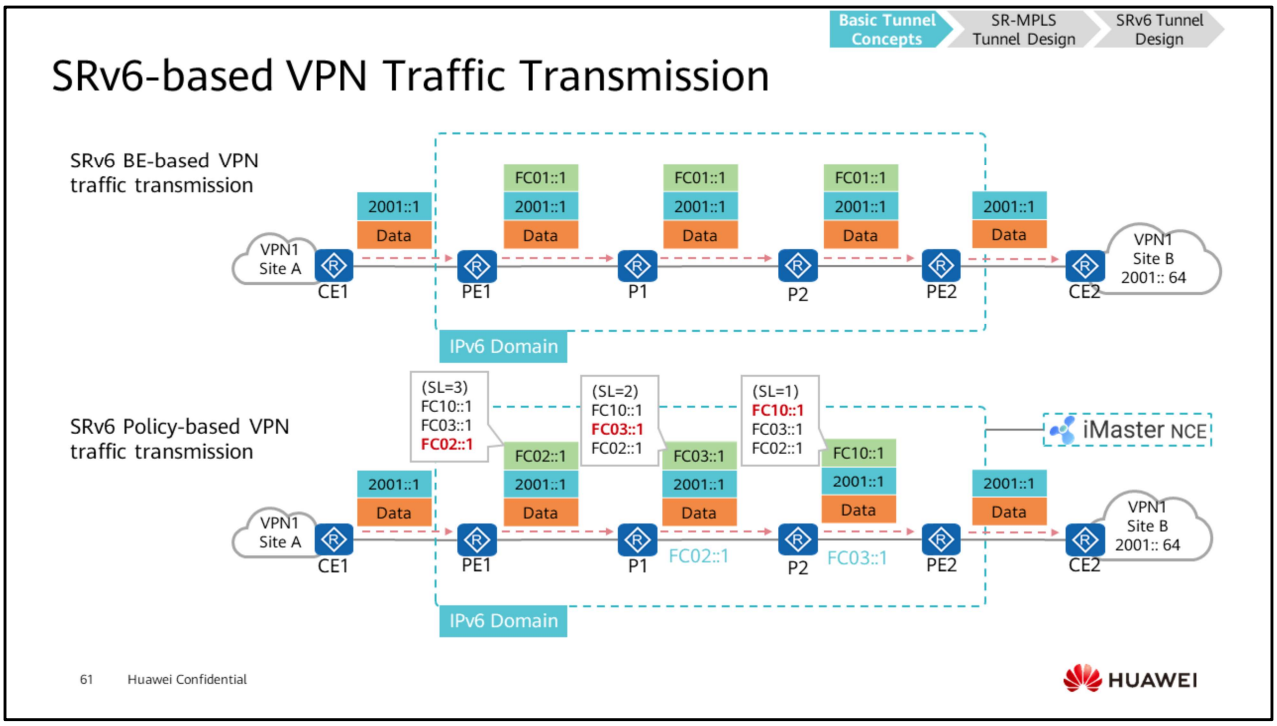
- The process of forwarding VPN traffic based on SR-MPLS BE is similar to the process of forwarding BGP/MPLS IP VPN traffic based on LDP.
    - After PE1 receives a VPN packet from CE1, PE1 searches the routing table and pushes two layers of labels into the packet. The outer label is a public network label, and the inner label is a private network label.
    - PE1 then sends the packet to P1, which swaps the outer label of the packet based on the SR-MPLS BE tunnel entry and sends the packet to P2. The process on P2 is similar to that on P1.
    - Upon receipt of the packet, PE2 sends the packet to a specific VPN site based on the inner label (PHP is not considered in this case).
- When an SR-MPLS Policy is used to carry VPN traffic, the forwarding path must be pre-computed and delivered to the ingress (PE1) as a segment list.
    - After receiving a VPN packet from CE1, PE1 searches the corresponding table and pushes the related segment list into the packet.
    - PE1 then sends the packet to P1, which determines the forwarding path based on the outer label, pops out the outer label, and sends the packet to P2.
    - After receiving the packet, P2 determines the forwarding path based on the outer label, pops out the outer label, and sends the packet to PE2.
    - PE2 sends the packet to the specified VPN site according to the inner label.

SRv6 Overview

- SRv6, or Segment Routing IPv6, is designed to forward IPv6 data packets on a network based on the source routing paradigm.
- SRv6 forwarding is no longer based on MPLS, simplifying the forwarding plane. SRv6 implements hop-by-hop forwarding by adding a Segment Routing header (SRH) into IPv6 packets, encapsulating an explicit IPv6 address stack into the SRH, and continuously updating destination addresses on transit nodes.



HUAWEI



- When SRv6 BE is used to carry VPN traffic, data packets carry two layers of IPv6 headers. The outer IPv6 header address is used to identify the VPN to which the data belongs, and the inner IPv6 header identifies the actual destination address of the data.
    - The outer IPv6 address is generated by the locator of PE2 and advertised to PE1 through BGP. PE2 advertises the locator to other devices in the form of a route.
    - After PE1 receives a packet destined for the destination network segment (2001:: 64), PE1 encapsulates the packet with an outer IPv6 header and forwards the packet based on the routing table.
    - Ps (P1 and P2) forward the packet based on the outer IPv6 header.
    - After receiving the packet, PE2 matches the packet with the corresponding VPN instance based on the outer IPv6 header and forwards the packet based on the routing table.
- When SRv6 Policy is used to carry VPN traffic, data packets carry two layers of IPv6 headers. The outer IPv6 header address is replaced by each hop based on the SRH information, and the inner IPv6 header identifies the actual destination address of the data.
    - Upon receipt of a packet destined for 2001:: 64, PE1 adds an outer IPv6 header (including the SRH) to the packet and sends the packet to the next hop based on the header.

- After receiving the packet, P1 replaces the outer IPv6 header based on the SRH information and forwards the packet. P2 processes the packet in a similar way.
- After receiving the packet, PE2 determines the VPN to which the packet belongs based on the outer IPv6 header and forwards the packet based on the routing table.

SR-MPLS Tunnel Design Principles

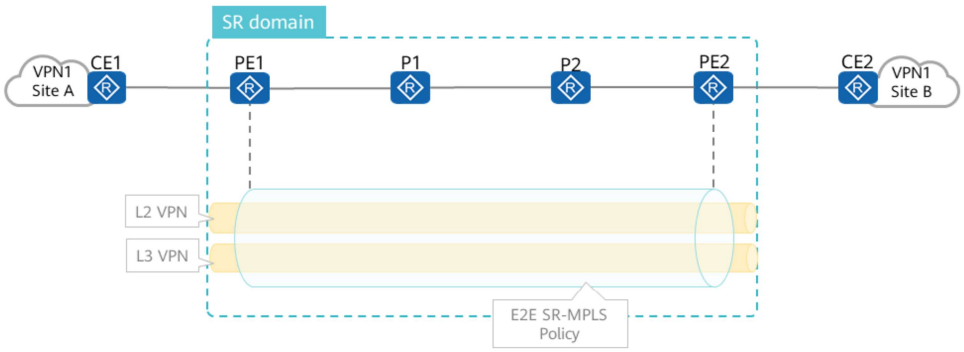
- Tunnels are intended to better serve services and focus on optimization requirements under the prerequisite that VPN bearer requirements are met. Comply with the following principles to achieve a balance among aspects such as service path visualization, service protection, network maintainability, and network scalability:
  - Service path visualization: Associate service traffic with tunnels to achieve some degree of path visualization.
  - Maintainability: Keep the total number of tunnels at an appropriate level to reduce live network maintenance pressure and shorten the optimization time.
  - Ease of optimization: Ensure that the traffic on each tunnel is not too heavy. Otherwise, bandwidth optimization will be difficult.
  - Reliability: Ensure that main services are under protection, and key services can be quickly converged.
  - Scalability: Consider possible network expansion in the future.

HUAWEI

- SR-MPLS BE tunnels are similar to LDP tunnels. Tunnel establishment depends on IGP design. Therefore, after IGP design is complete, SR-MPLS BE design is complete.
- This section mainly focuses on SR-MPLS Policy design.

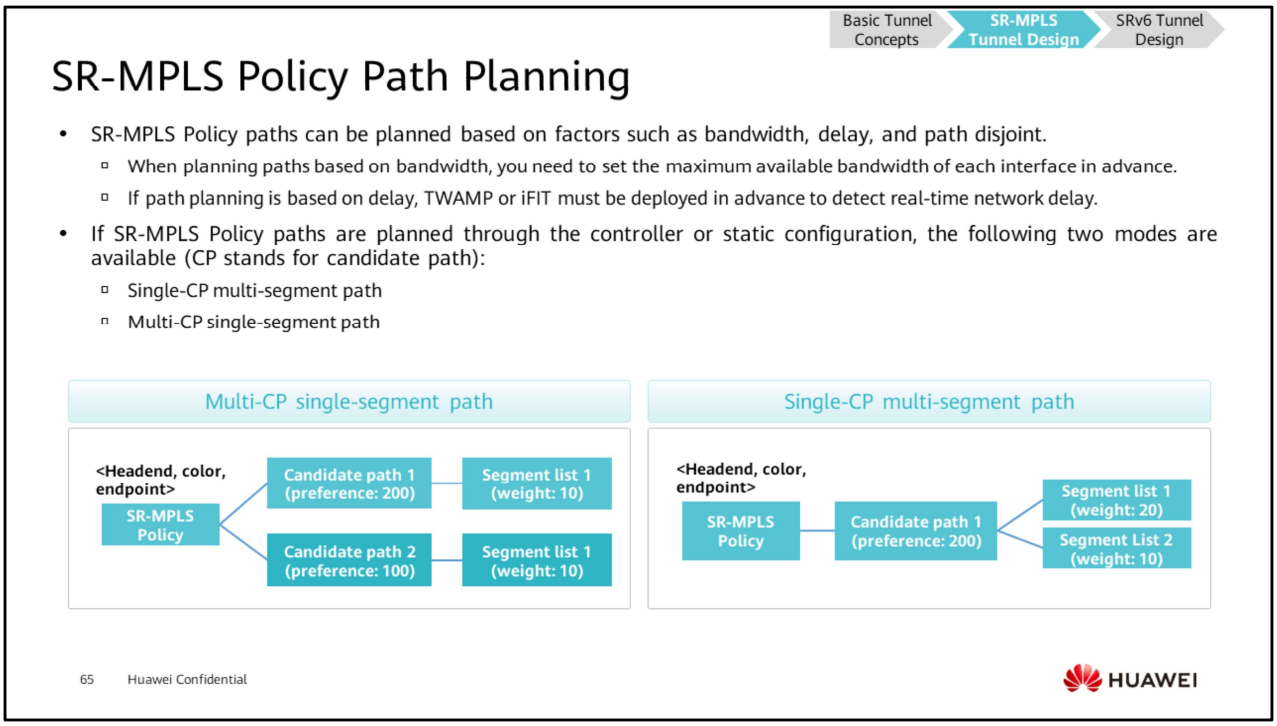
SR-MPLS Policy Deployment Design

- Generally, an enterprise bearer network belongs to an independent AS. Therefore, E2E SR-MPLS Policies can be deployed.
- An E2E tunnel can be deployed between the ingress (PE) where service traffic enters the bearer network and the egress (PE) where service traffic leaves the backbone network.

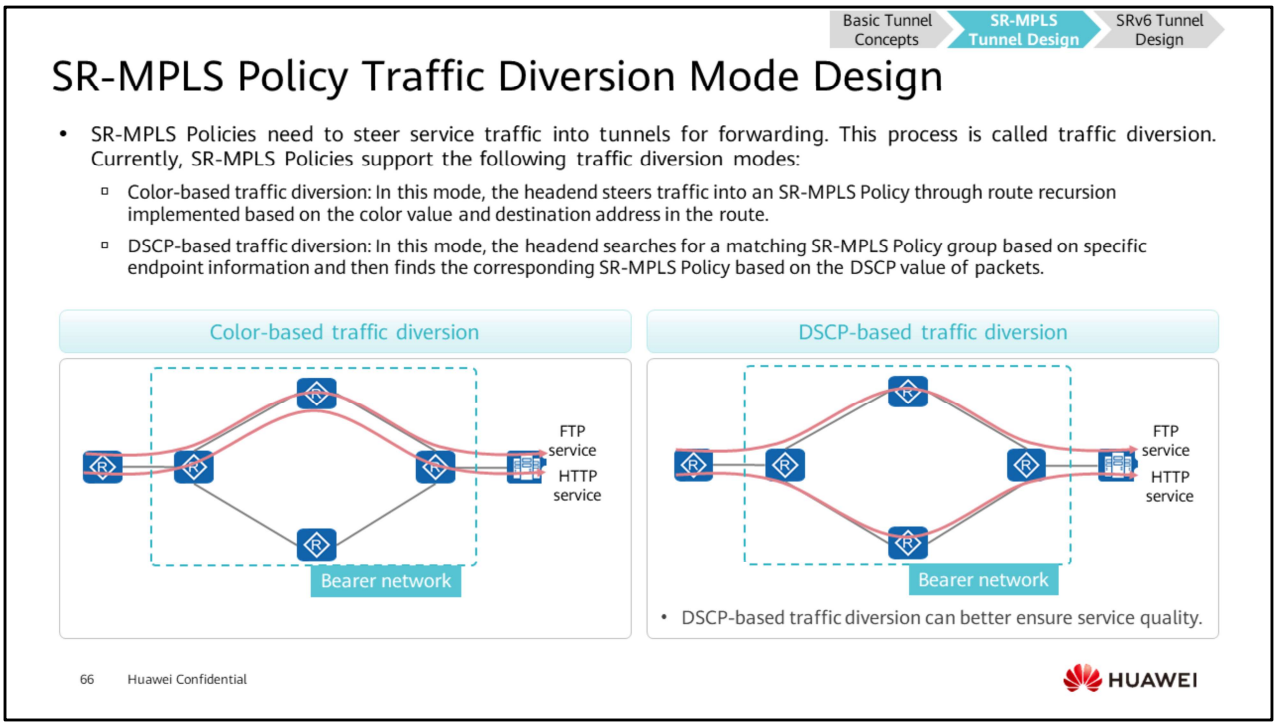


HUAWEI

- End-to-end deployment has the following characteristics:
  - Strong path control can be implemented through the planning of end-to-end paths (especially explicit paths).
  - Path visualization is relatively good.
  - If the network is large (for example, a network with multiple data centers and dozens of branches) and has 5,000 to 10,000 TE tunnels, the maintenance workload is heavy.
  - The scalability is fair. If a new data center or branch is added, a large number of end-to-end tunnels need to be added.



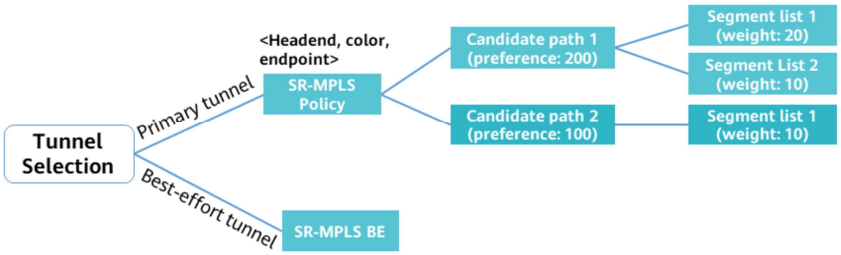
- If iFIT is used to measure the network delay, 1588v2 must be enabled on the entire network. Therefore, there are restrictions on application scenarios.
- TWAMP requires only NTP in network delay measurement.
- For a tunnel planned based on bandwidth, the actual traffic volume of the tunnel cannot be limited on devices after the tunnel is delivered. The traffic volume of a tunnel needs to be limited on the ingress, and the QoS or network slicing technology needs to be used.



- In color-based traffic diversion, different tunnels (including primary and backup tunnels) can only be selected based on endpoints. If different service traffic (such as HTTP and FTP traffic) is destined for the same address, color-based traffic diversion diverts the traffic to the same tunnel. As a result, the quality of some services deteriorates.
- In DSCP-based traffic diversion, different tunnels can be selected based on endpoint + DSCP information. If different service traffic (such as HTTP and FTP traffic) is destined for the same address, DSCP-based traffic diversion will steer different services into different tunnels based on the configuration, thereby ensuring the service quality.

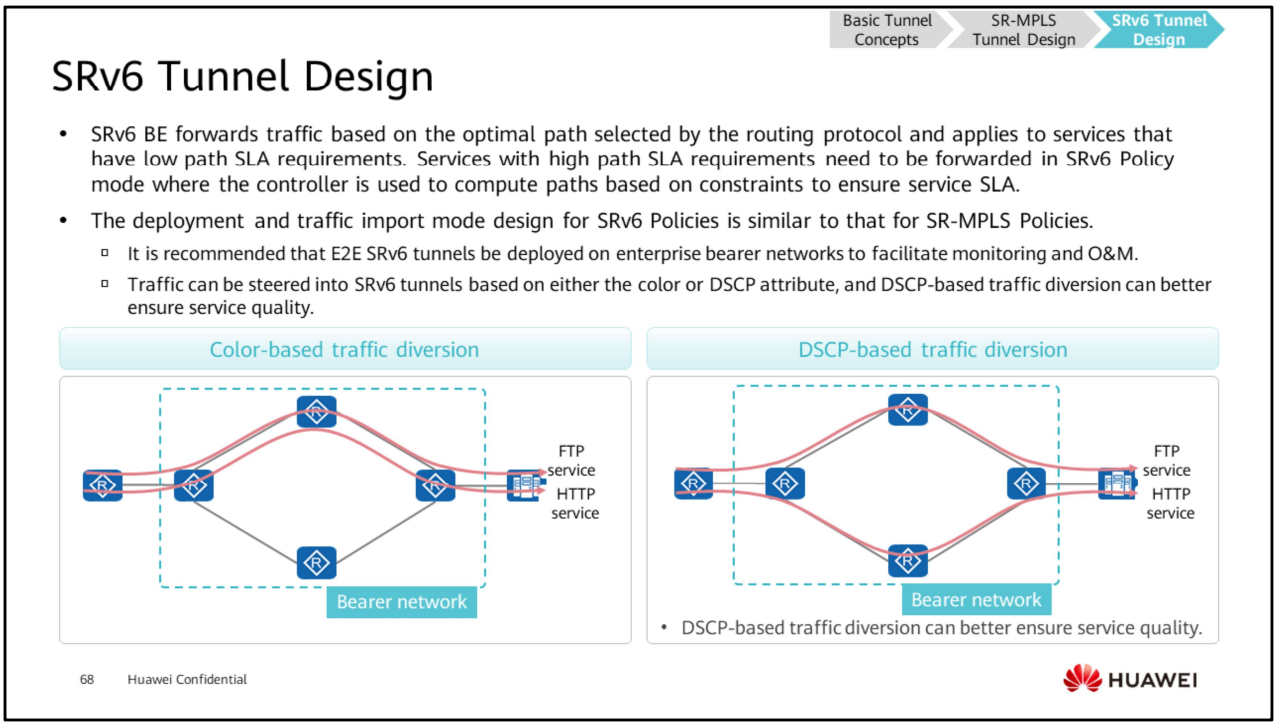
Best-Effort Forwarding Design for SR-MPLS Tunnels

- VPN services can be carried over either SR-MPLS Policies or SR-MPLS BE tunnels.
- An SR-MPLS Policy can contain multiple candidate paths with the preference attribute. The valid candidate path with the highest preference functions as the primary path of the SR-MPLS Policy, and the valid candidate path with the second highest preference functions as the HSB path.
- If all SR-MPLS Policies fail, VPN services can be carried over SR-MPLS BE tunnels.
- It is recommended that both SR-MPLS Policies and SR-MPLS BE tunnels be deployed. SR-MPLS Policies are preferentially used, and SR-MPLS BE tunnels serve as their backup.

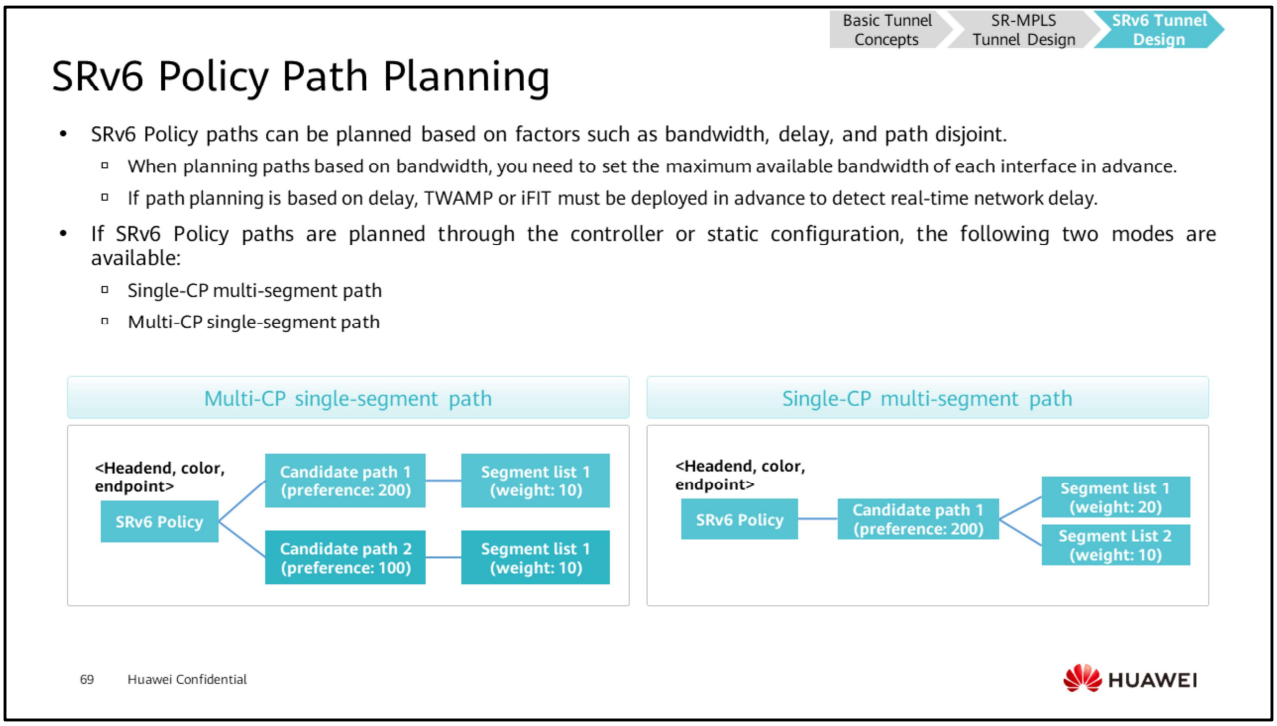


HUAWEI

- An SR-MPLS Policy can have multiple candidate paths, such as CP1 and CP2. Each path is uniquely identified by a 3-tuple <protocol, origin, discriminator>.
- CP1 is the activated path because it is valid and has a higher priority. The two SID lists (also called segment lists) of CP1 are delivered to the forwarder, and traffic is balanced between the two tunnel paths based on weight. For example, traffic along the SID list <SID11, SID12> is balanced based on W1/(W1+W2). In the current mainstream implementation, a candidate path has only one segment list.
- If a controller is used to generate SR-MPLS Policies, only primary and backup tunnels can be established, and load balancing cannot be implemented for the primary tunnel.



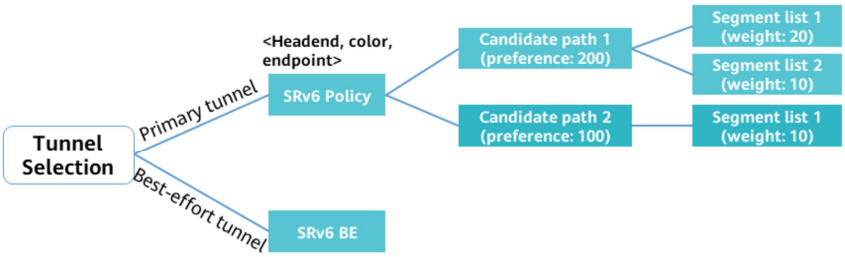
- Similar to LDP tunnels, SRv6 BE tunnels are calculated based on IGP/BGP optimal paths. Unlike label-featured MPLS, SRv6 BE uses the shortest path first (SPF) algorithm to calculate forwarding paths based on SRv6 SIDs in an IGP domain. SRv6 BE requires only one segment to identify a forwarding path and the carried services. Traffic forwarding along paths depends on cost planning. Traffic is forwarded based on the least-cost route.
- This section describes SRv6 TE Policy design.



- If iFIT is used to measure the network delay, 1588v2 must be enabled on the entire network. Therefore, there are restrictions on application scenarios.
- TWAMP requires only NTP in network delay measurement.
- For a tunnel planned based on bandwidth, the actual traffic volume of the tunnel cannot be limited on devices after the tunnel is delivered. The traffic volume of a tunnel needs to be limited on the ingress, and the QoS or network slicing technology needs to be used.

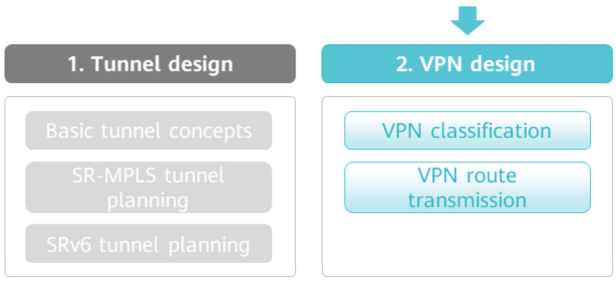
Best-Effort Forwarding Design for SRv6 Tunnels

- Similar to SR-MPLS, SRv6 can use either SRv6 BE tunnels or SRv6 Policies to carry VPN services.
- SRv6 BE tunnels are generally used as best-effort tunnels. If all SRv6 Policies fail, SRv6 BE tunnels are used to carry services.
- It is recommended that SRv6 BE tunnels and SRv6 Policies both be deployed. SRv6 Policies are preferentially used, and SRv6 BE tunnels serve as their backup.



HUAWEI

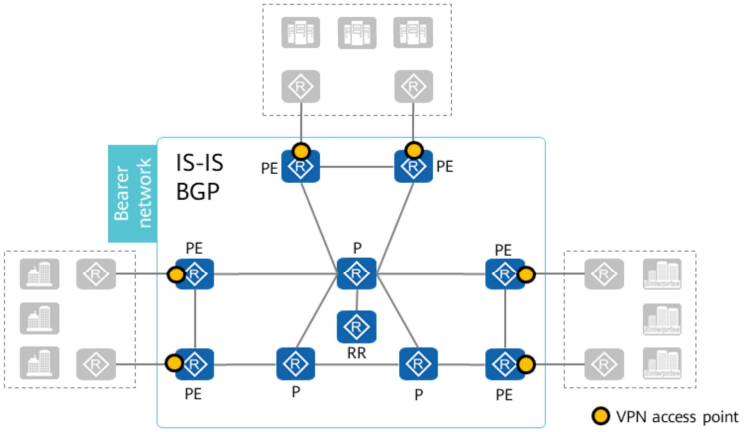
VPN Design Roadmap for the Enterprise Bearer WAN



HUAWEI

VPN Classification

- Different enterprises have different requirements for VPN division. Most enterprises use one VPN for major services (including production and office services), one VPN for non-major services, one VPN for external services, and one VPN for Internet services.



HUAWEI

IPv4 VPN Type Selection

- When SR-MPLS is used to carry IPv4 L3VPN traffic, the BGP VPNv4 or BGP EVPN address family can be used to transmit VPN routes.
- When SRv6 is used to carry IPv4 L3VPN traffic, it is recommended that the BGP EVPN address family be used to transmit VPN routes, so that IPv6 and Layer 2 services can be carried in the same manner.
- The L3VPN capabilities and implementation processes of VPNv4 and EVPN address families are basically the same.

	VPNv4	EVPNv4	Remarks
Forwarding plane	MPLS/SRv6	MPLS/SRv6	
VPN route learning	BGP	BGP	The BGP route formats are different.
Whether RR-based route reflection is supported	Supported	Supported	
Failover	Supported	Supported	VPN FRR, IP FRR, TE-HSB (MPLS), CBTS, TI-LFA (SRv6), Mirror-SID

HUAWEI

IPv6 VPN Type Selection

- When SR-MPLS is used to carry IPv6 L3VPN traffic, the MP-VPNv6 or BGP EVPN address family can be used to transmit VPN routes.
- When SRv6 is used to transmit IPv6 services, it is recommended that the BGP EVPN address family be used to transmit VPN routes.

	VPNv6	EVPNv6	Remarks
Forwarding plane	MPLS/SRv6	MPLS/SRv6	
VPN route learning	BGP	BGP	The BGP route formats are different.
Whether RR-based route reflection is supported	Supported	Supported	
Failover	Supported	Supported	VPN FRR, IP FRR, TE-HSB, CBTS, TI-FLA, Mirror-SID

HUAWEI

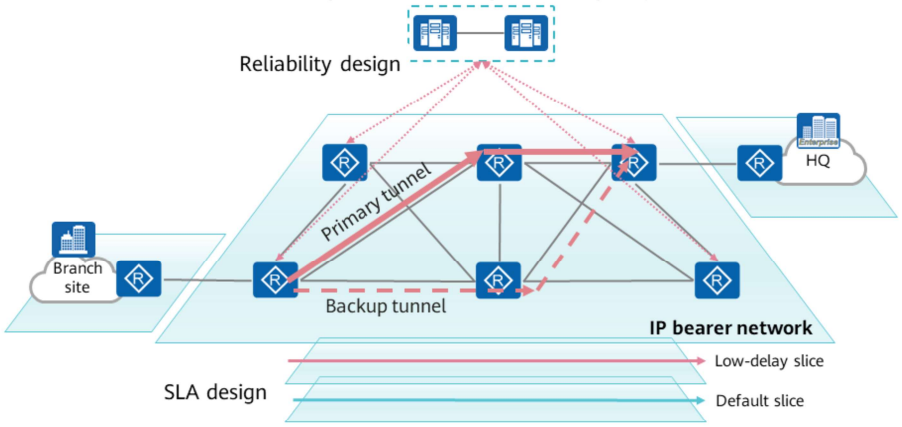
Contents

1. Current Situation and Challenges of the Enterprise Bearer WAN
2. Huawei CloudWAN Solution Overview
3. Basic Design for the Enterprise Bearer WAN
4. Tunnel and VPN Design for the Enterprise Bearer WAN
5. **SLA and Reliability Design for the Enterprise Bearer WAN**
6. Optimization and O&M Design for the Enterprise Bearer WAN

HUAWEI

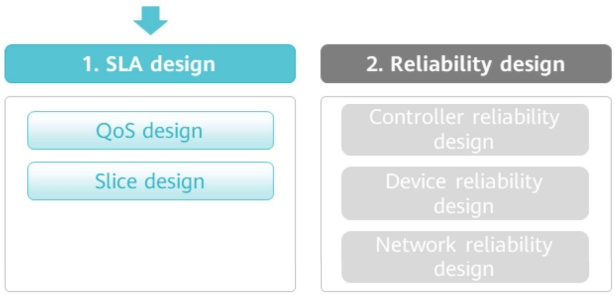
SLA and Reliability Design Overview for the Enterprise Bearer WAN

- The enterprise bearer WAN uses reliability and SLA technologies to effectively ensure the quality of carried services. Therefore, reliability and SLA design are very important.



HUAWEI

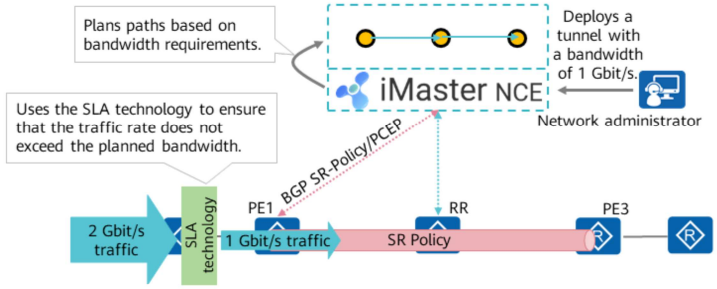
SLA Design Roadmap for the Enterprise Bearer WAN



HUAWEI

SLA Technology Overview

- In Huawei's CloudWAN solution, iMaster NCE-IP can compute paths based on bandwidth requirements and deliver forwarding paths (SR Policies) to network devices.
- Although the controller can compute paths based on bandwidth requirements, the delivered path information does not contain any traffic rate limiting policy. As a result, the traffic rate on the forwarding path (SR Policy) may exceed the planned bandwidth.
- To ensure that the traffic rate does not exceed the planned bandwidth, SLA technologies need to be deployed on the network to limit traffic bandwidth.
- SLA technologies mainly include QoS and network slicing.



QoS Planning Principles

- The enterprise bearer network needs to provide differentiated service quality assurance for different services. QoS planning ensures that various services are properly forwarded on the bearer network. QoS planning mainly complies with the following four principles:
  - Reasonableness: Resources must be allocated appropriately based on the importance of services.
  - Consistency: QoS planning involves various behaviors (such as service classification, marking, scheduling, and rate limiting), which must be consistent on the entire network.
  - Scalability: Current QoS policies must take into account future service expansion.
  - Maintainability: Because services change rapidly in real-world situations, QoS policies may be frequently adjusted during routine maintenance. Ensure that QoS policies can be easily adjusted and maintained.

HUAWEI

QoS Design Suggestions

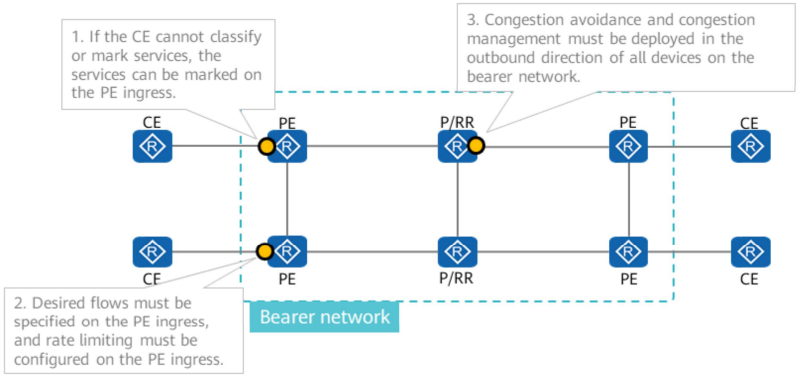
- When the network is normal, a different link can be planned for each service of an enterprise to ensure that these services do not affect each other. However, when a link is faulty, the affected services are switched to other links, and bandwidth competition may occur. QoS policies need to be deployed for key services in a unified manner.
- If there are more than eight types of enterprise services, they need to be properly classified and combined. The following is an example of common enterprise service planning.

Priority	Service	Scheduling Mode	Weight	Traffic Shaping	Drop Method
CS6/CS7	Protocol packet	PQ	NA	NA	Tail drop
EF	Production service	PQ	NA	NA	Tail drop
AF4	VoIP, video conference	PQ	NA	Configure rate limiting to prevent bandwidth starvation of low-priority services.	Tail drop
AF3	Video surveillance	WFQ	Determined based on live network conditions.	NA	WRED
AF2	Office service	WFQ	Determined based on live network conditions.	NA	WRED
AF1	OA service	LPQ	NA	NA	Tail drop
BE	Other workloads	LPQ	NA	NA	Tail drop

HUAWEI

QoS Deployment Design

- QoS planning requires that QoS features be properly deployed at different locations on the network. The following figure shows the deployment of different features on the network:

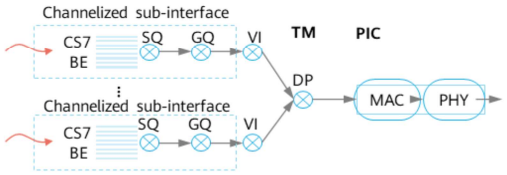


HUAWEI

Network Slicing Overview

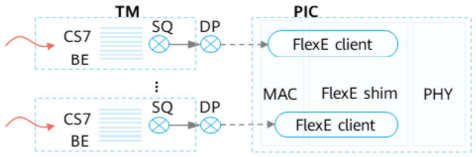
- Network slicing can be used to allocate dedicated network resources on a network to carry high-value service traffic.
- Network slicing and SR tunnels apply to different network layers. Network slicing reserves resources on the Layer 1.5 or Layer 2 network and can be used together with SR tunnels.
- Network slicing is generally implemented based on channelized sub-interfaces or FlexE.

Channelized sub-interface



- Channelized sub-interface: Queue resources are isolated. Hierarchical scheduling is used to implement flexible and refined management of interface resources, provide bandwidth guarantee, and work with the controller to provide E2E resource reservation.

FlexE

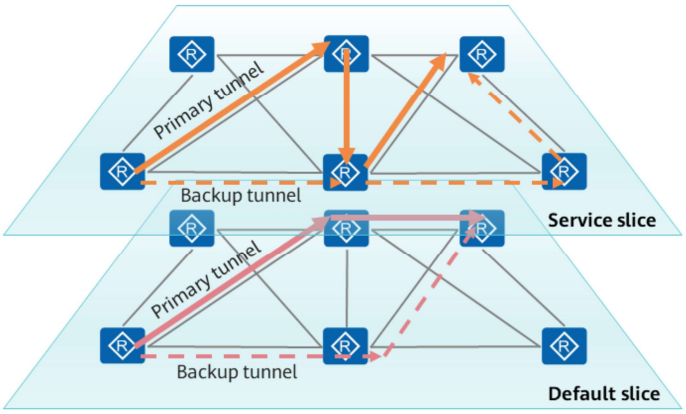


- FlexE: Queue and interface resources are isolated. Every resource is divided by TDM timeslot. This meets the requirements for exclusive resource use and resource isolation and provides flexible and refined management of interface resources.

HUAWEI

Network Slicing Design

- Currently, existing devices mainly implement network slicing based on bandwidth.
- A slice with the corresponding bandwidth is created based on actual service requirements. An SR tunnel is then bound to the slice for bearing.

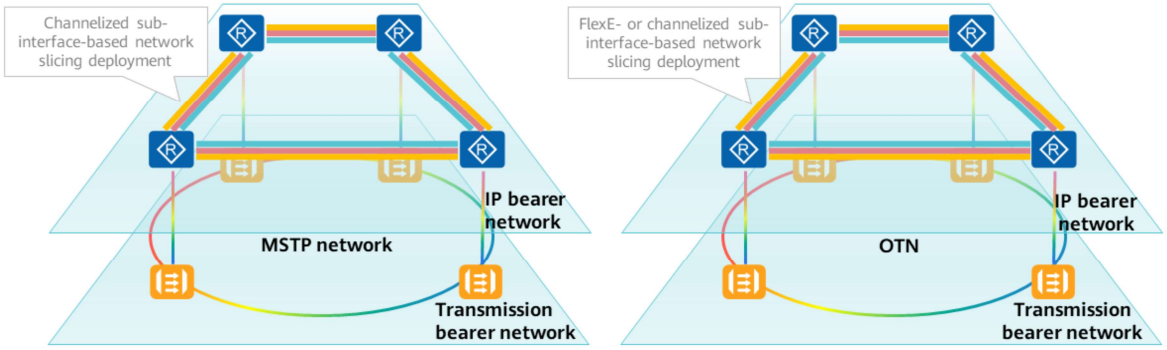


- SR BE tunnels or SR Policies can be deployed in the default slice.
- SR Policies are deployed in the service slice. BFD is deployed to detect tunnel status, implementing HSB protection and VPN FRR within slices.
- iMaster NCE-IP computes tunnels over slices based on affinity attributes.
- iMaster NCE-IP optimizes traffic within slices.

HUAWEI

Network Slicing Application Scenarios

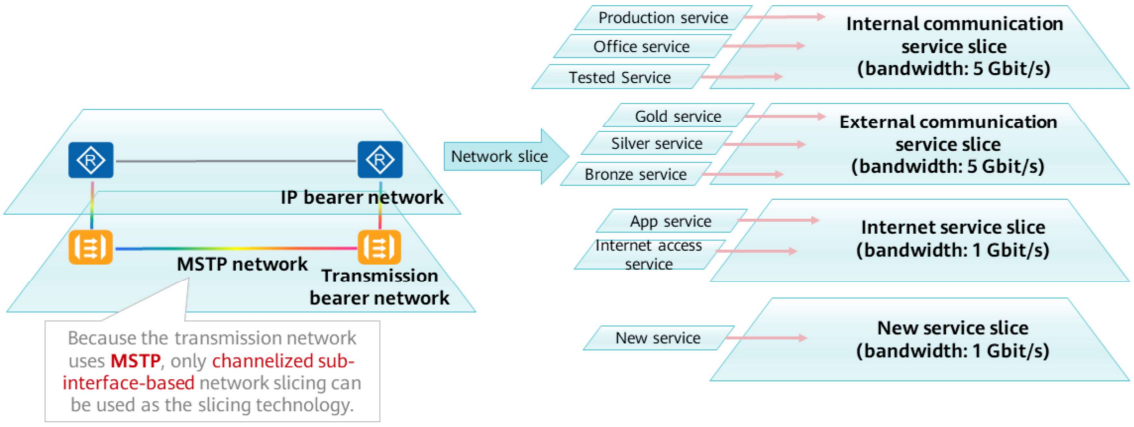
- FlexE-based network slicing and channelized sub-interface-based network slicing have different application scenarios.
  - It is recommended that FlexE be used to reserve resources for 50GE and higher-speed interfaces and channelized sub-interfaces be used to reserve resources for lower-speed interfaces.
  - Only channelized sub-interface-based network slicing can be deployed across MSTP devices.
  - FlexE- or channelized sub-interface-based network slicing can be deployed across OTN devices as required.



HUAWEI

Typical Network Slicing Planning

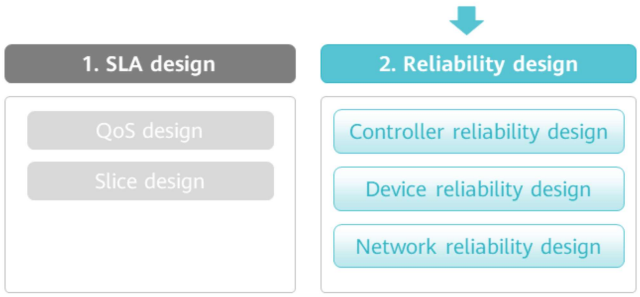
- Services can be placed into different slices based on application types. Moreover, different slicing technologies can be used based on transmission lines (MSTP or OTN).



Because the transmission network uses **MSTP**, only channelized sub-interface-based network slicing can be used as the slicing technology.

HUAWEI

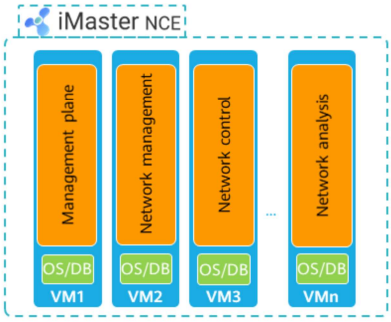
Enterprise Bearer WAN Reliability Design



HUAWEI

Controller Local Reliability Design

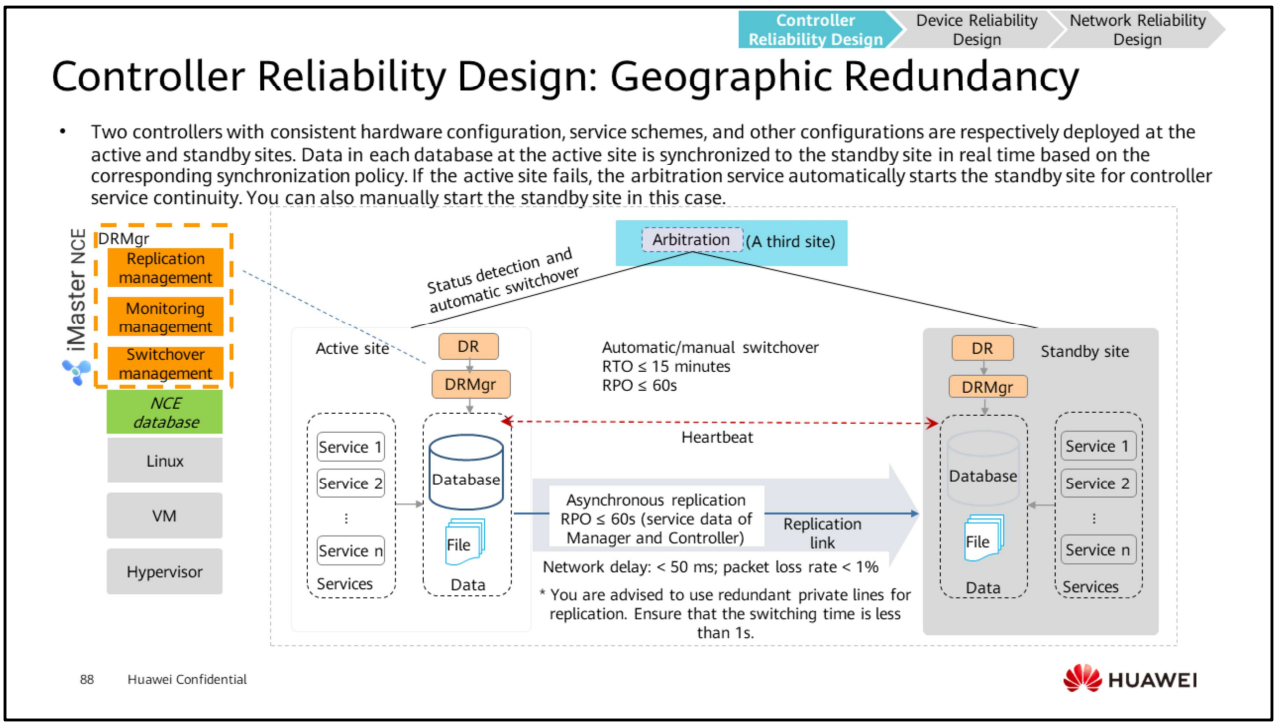
- iMaster NCE-IP has two high reliability modes:
  - Active/standby protection mode: Only the services on the active node are in the running status. If a service process on the active node encounters a fault, the controller automatically starts the service process on the standby node to provide services.
  - Cluster protection mode: When running properly, all cluster nodes are in the all-active state. If one node is faulty, other nodes share the load of the faulty node and continue to provide services evenly.



Protection Layer	Component	Remarks
Application layer	Manager	Protection mechanism: application active/standby protection Performance indicators: RPO: 0s; RTO: ≤ 5 minutes
	Controller	Protection mechanism: application active/standby protection Performance indicators: RPO: 0s; RTO: ≤ 60s
	Analyzer	Protection mechanism: application active/standby protection Performance indicators: RPO: 0s; RTO: ≤ 5 minutes
Database	Database	Protection mechanism: database active/standby protection Performance indicators: RPO: 60s; RTO: ≤ 60s
Virtualization layer	FusionCompute	N/A
Server	TaiShan/RH/E9000 server	Management module: 1+1 backup; power supply: 1+1 backup Network port: 1+1 backup; hard disk: RAID1/RAID10

HUAWEI

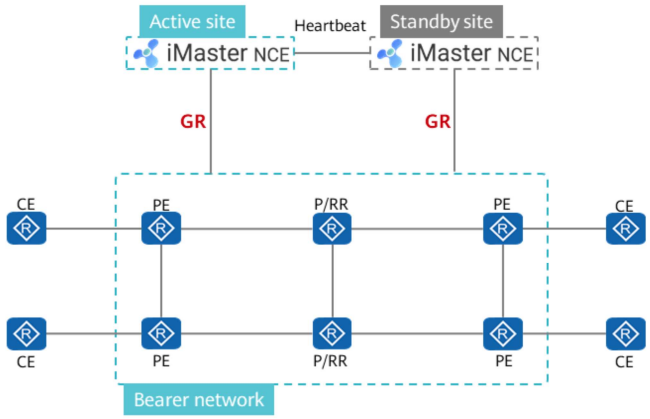
- RPO: recovery point objective
- RTO: recovery time objective



- During the operation of the DR system, NCE monitors the association status of the primary and secondary sites over the heartbeat link and synchronizes data over the replication link. If the heartbeat or replication link between the primary and secondary sites is abnormal, the controller reports an alarm. The fault can be either manually rectified or automatically processed by the arbitration service.
- The DRMgr service is developed by Huawei. Customers do not need to purchase third-party HA management software.

Control Network Reliability Design for the Controller

- GR needs to be deployed on both iMaster NCE-IP and devices, so that policy entries on forwarders can be retained for a longer time if the controller is faulty, maintained, or upgraded, or an active/standby controller switchover occurs.



HUAWEI

- Currently, the active/standby switchover of iMaster NCE-IP can be completed within 10 minutes. Therefore, the GR time cannot be less than 20 minutes.
- GR: graceful restart

Device Reliability Design

- High device reliability is essential to the effective running of the network and needs to be guaranteed through the hardware, software, and protection mechanisms.
- Device reliability deployment:
  - 1+1 active/standby protection for main control boards (NSR is recommended for smooth active/standby switchovers)
  - Load balancing among multiple microengines for interface boards
  - Maximum backup for power supplies
  - Redundancy protection for other key components (such as fan modules)



HUAWEI

- NSR: non-stop routing

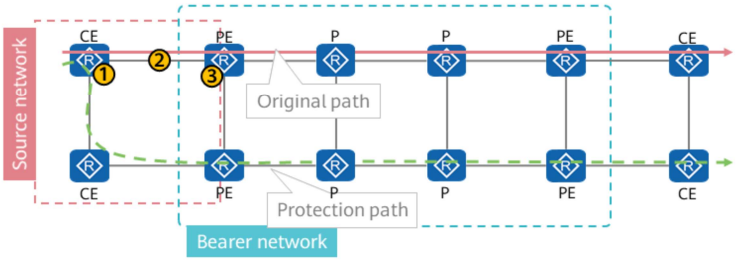
Network Reliability Solution Overview

- Multiple reliability technologies can be deployed on the bearer network to implement E2E reliability protection.
- Network reliability technologies include:
  - BFD/SBFD: is mainly used to check network connectivity.
  - IP FRR: is mainly used to provide link and node protection for transit networks and can work with the LFA/RLFA/TI-LFA algorithm.
  - Anycast FRR: is mainly used to provide protection for specific nodes, including transit and egress nodes.
  - HSB: is mainly used to provide E2E tunnel protection.
  - Mirror SID: is mainly used to provide protection for egress nodes.
  - Microloop avoidance: is mainly used to prevent temporary loops caused by inconsistent route convergence time on the entire network.

HUAWEI

Source Network Reliability Design

- The source network mainly includes source CEs, source PEs, and links between source CEs and source PEs. The source network may encounter the following faults:
  1. Source CE fault
  2. Link fault between a source CE and a source PE
  3. Source PE fault
- These three types of faults can be quickly detected through BFD, and IP FRR (mainly based on the LFA/RLFA algorithm) can be used to compute backup links, so that services can be quickly switched to the protection paths.

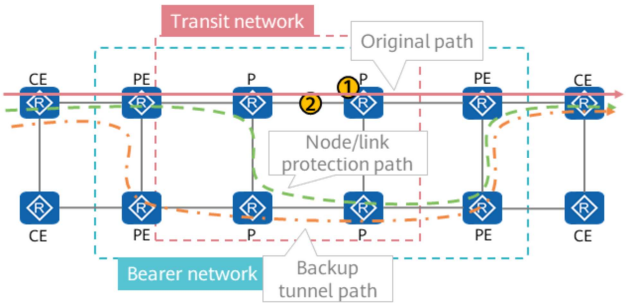


HUAWEI

Controller Reliability Design | Device Reliability Design | Network Reliability Design

Transit Network Reliability Design

- The transit network mainly includes Ps, links between Ps and source PEs, and links between Ps.
- The transit network may encounter the following faults:
  1. P fault
  2. Link fault between a P and a PE or between Ps
- The following method can be used to protect the network against these two types of faults:
  - Deploy BFD to quickly detect network link faults.
  - Use IP FRR (mainly based on the TI-LFA algorithm) to compute backup paths.
  - If the transit network cannot meet service requirements due to a fault, use the tunnel HSB technology to switch traffic to the backup path.

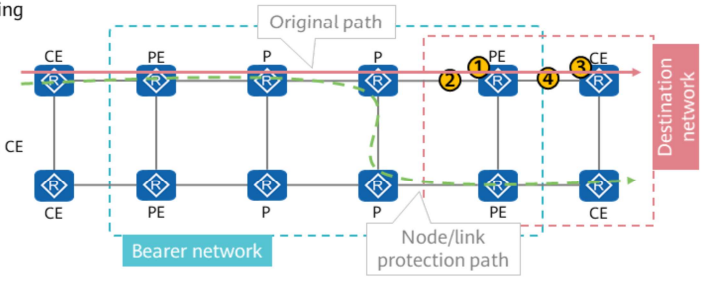


HUAWEI

- The transit network may fail to meet service requirements due to insufficient bandwidth or long delay. To detect the network bandwidth or delay, network quality detection technologies such as TWAMP or iFIT need to be deployed.

Destination Network Reliability Design

- The destination network mainly includes destination PEs, destination CEs, links between destination PEs and Ps, and links between destination PEs and destination CEs.
- The destination network may encounter the following faults:
  1. Destination PE fault
  2. Link fault between a destination PE and a P
  3. Destination CE fault
  4. Link fault between a destination PE and a destination CE
- The following method can be used to protect the network against the type 1 and type 2 faults:
  - Deploy BFD to quickly detect network link faults.
  - Use mirror SID or anycast FRR to switch traffic to the standby destination PE.
- The following method can be used to protect the network against the type 3 and type 4 faults:
  - Deploy BFD to quickly detect network link faults.
  - Compute backup paths through IP FRR (mainly based on the LFA/RLFA algorithm).



HUAWEI

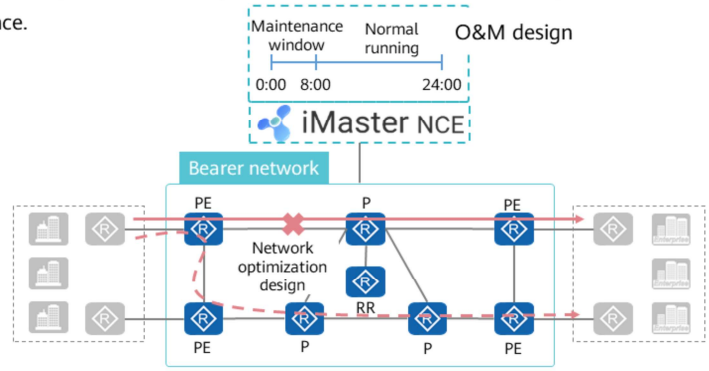
Contents

1. Current Situation and Challenges of the Enterprise Bearer WAN
2. Huawei CloudWAN Solution Overview
3. Basic Design for the Enterprise Bearer WAN
4. Tunnel and VPN Design for the Enterprise Bearer WAN
5. SLA and Reliability Design for the Enterprise Bearer WAN
6. **Optimization and O&M Design for the Enterprise Bearer WAN**

HUAWEI

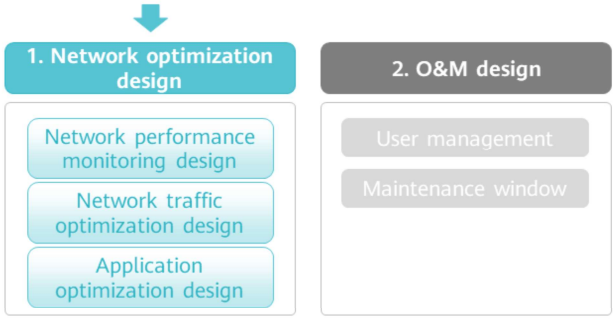
Optimization and O&M Design Overview for the Enterprise Bearer WAN

- Network operation becomes the new focus after the initial stage of network construction is complete. Network optimization and O&M are essential to smooth network operation.
- To better support subsequent network optimization and O&M, network optimization and O&M design must be performed in advance.



HUAWEI

Optimization Design Roadmap for the Enterprise Bearer WAN



HUAWEI

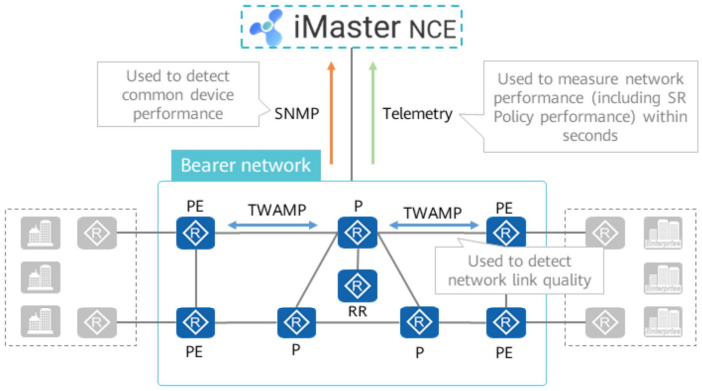
Network Performance Monitoring Overview

- Network performance monitoring can be implemented in multiple ways, one of the most common being SNMP. However, as services impose increasingly more requirements on networks, new technologies are required to quickly obtain network performance indicators.
- Network performance monitoring technologies commonly used on live networks include:
  - SNMP: During performance monitoring, SNMP obtains network performance information in a query-reply manner. When frequently obtaining information, SNMP imposes heavy pressure on the device.
  - Telemetry: Telemetry mainly obtains network performance information in a subscription-reporting manner. When frequently obtaining information, telemetry does not impose much pressure on the device.
  - NQA: NQA is mainly used to detect network quality. NQA uses simulated traffic to test the network environment. Therefore, the test result is not so accurate.
  - TWAMP: TWAMP is mainly used to detect network quality. Compared with NQA, TWAMP has a unified detection model and packet format, and is easy to deploy.

HUAWEI

Application Scenarios of Network Performance Monitoring

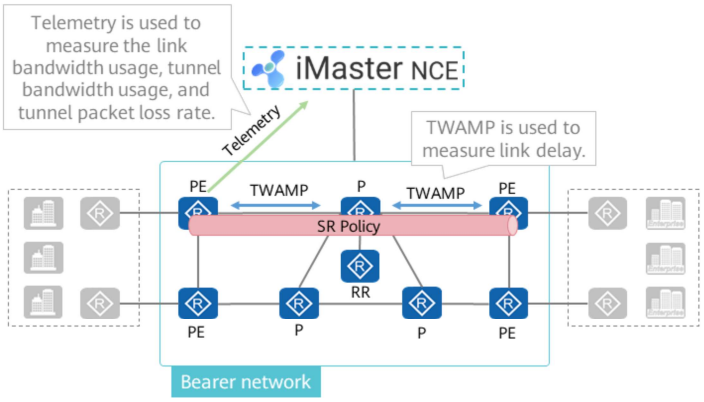
- In Huawei's CloudWAN solution, SNMP is used to collect common device performance information, telemetry is used to collect information that needs to be measured within seconds, and TWAMP is used to collect link quality-related performance data (including packet loss, delay, and jitter).



HUAWEI

Network Performance Monitoring Design

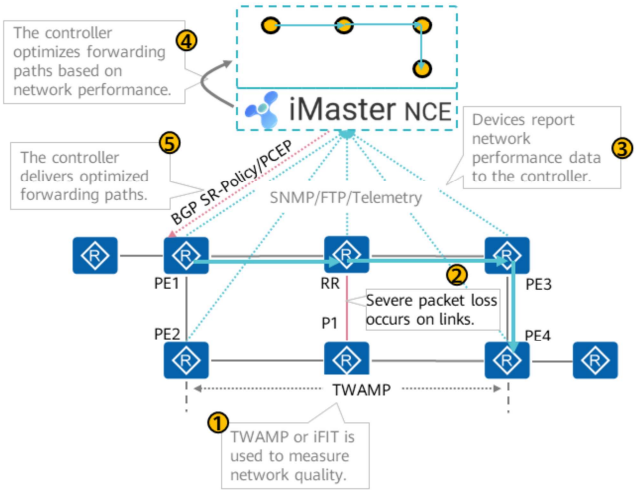
- The major performance indicators for WAN services include:
  - Link bandwidth usage (measured through telemetry)
  - Link delay (measured through TWAMP)
  - Tunnel bandwidth usage (measured through telemetry)
  - Tunnel packet loss rate (measured through telemetry)



HUAWEI

Network Traffic Optimization Overview

- MPLS TE uses the Constrained Shortest Path First (CSPF) algorithm to compute paths. This distributed computation and management mode, however, cannot manage network bandwidth resources in a coordinated manner.
- Huawei's CloudWAN solution uses BGP-LS (SRv6 Policy) or BGP-LS + SR-MPLS Policy (PCEP) to obtain tunnel status and adjusts tunnel paths based on the network status.

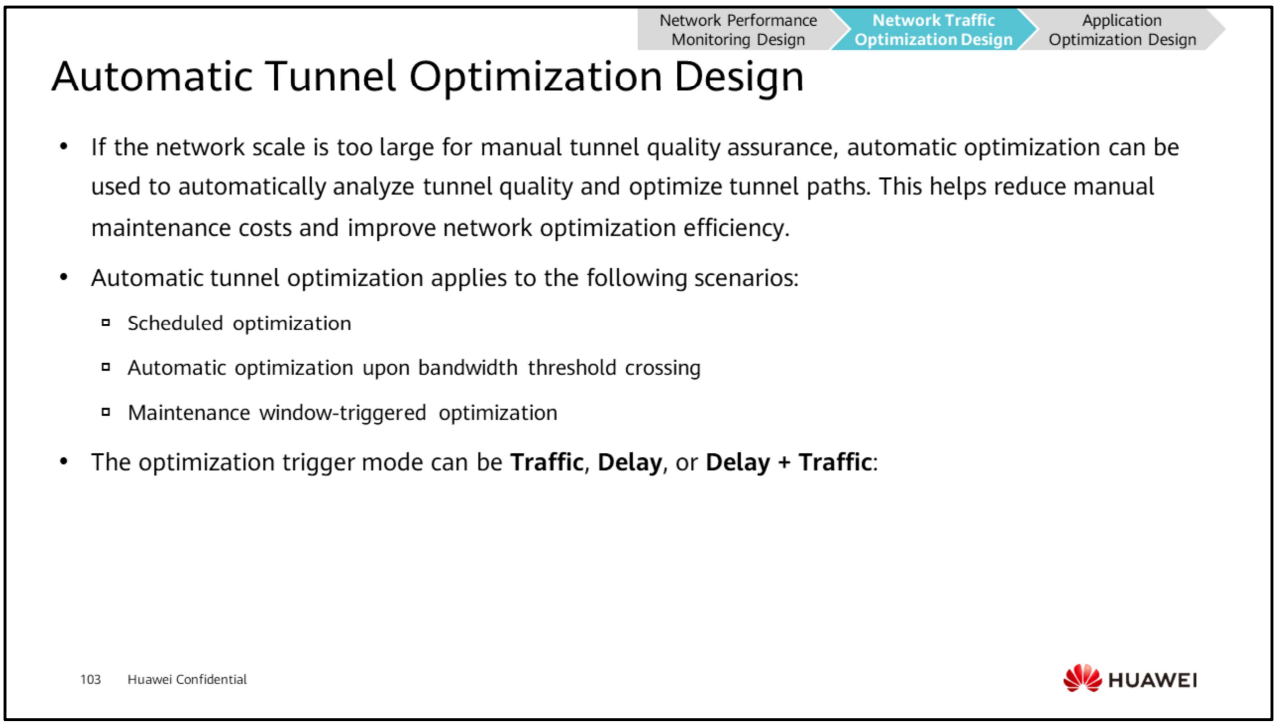


HUAWEI

Network Traffic Optimization Design

- Network optimization design focuses on two aspects: what to optimize and how to optimize.
- What to optimize
  - Simply put, optimization is to optimize service paths (LSPs). An LSP is a logical path equivalent to a tunnel, and a link is a physical path. One link can carry multiple tunnels, and one tunnel (LSP) corresponds to one service. Therefore, service paths can be changed based on either links or tunnels.
    - Link optimization: When one or more links are selected for optimization, all LSPs carried by the selected links are involved in path computation.
    - Tunnel optimization: When one or more tunnels are selected for optimization, the LSPs corresponding to the selected tunnels are involved in path computation.
- How to optimize
  - Optimization can be performed either automatically or manually:
    - In an automatic optimization scenario, traffic is automatically analyzed, and the optimization is automatically performed at scheduled times.
    - In a manual optimization scenario, traffic is manually optimized on demand based on network conditions.

HUAWEI



- Automatic optimization:
  - Scheduled optimization: You can set the interval for automatically optimizing network paths to 5 minutes or longer to ensure that the current service paths are optimal.
  - Automatic optimization upon bandwidth threshold crossing: You can set the link threshold. Then, when the bandwidth usage of a link exceeds the threshold, the system automatically adds tunnels over the link to the path computation queue and performs optimization when the optimization period arrives.
  - Maintenance window-triggered optimization: You can maintain a node or link. During the maintenance, the node or link is unavailable. After the maintenance starts or ends, the controller automatically recomputes the paths of tunnels that pass through the node or link.
  - The automatic optimization interval is an integer multiple of 5 minutes, for example, 10 minutes or 15 minutes.
- Optimization policy:
  - Traffic-based optimization: In this mode, the link threshold must be set. If the bandwidth usage of links exceeds the threshold, the controller determines whether to perform local or global optimization based on the number of threshold-crossing links.

- Delay-based optimization: In this mode, the controller traverses and compares the configured delay of all tunnels with the accumulated delay collected from forwarders, and performs local optimization on all the tunnels whose accumulated delay exceeds the configured delay.
- Delay+traffic-based optimization: If either of the preceding trigger conditions is met, traffic optimization is performed.

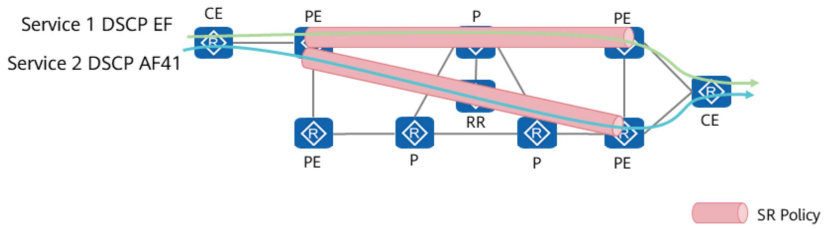
Manual Tunnel Optimization Design

- In manual optimization, the operator analyzes the network status and then adjusts the network manually. Manual optimization can be performed on demand at any time.
- Manual optimization is typically performed in the following scenarios:
  - If the current traffic paths do not meet requirements, you can modify link constraints and then manually trigger optimization to change traffic paths.
  - After configuring a service, you can manually perform network re-optimization to ensure that all tunnel paths are optimal.
  - When the link quality deteriorates (or the bandwidth usage is high but does not reach the automatic optimization threshold) or the bandwidth usage of links is uneven (some links have high bandwidth usage while others are idle), you can manually perform local/global optimization.

HUAWEI

Application Optimization Overview

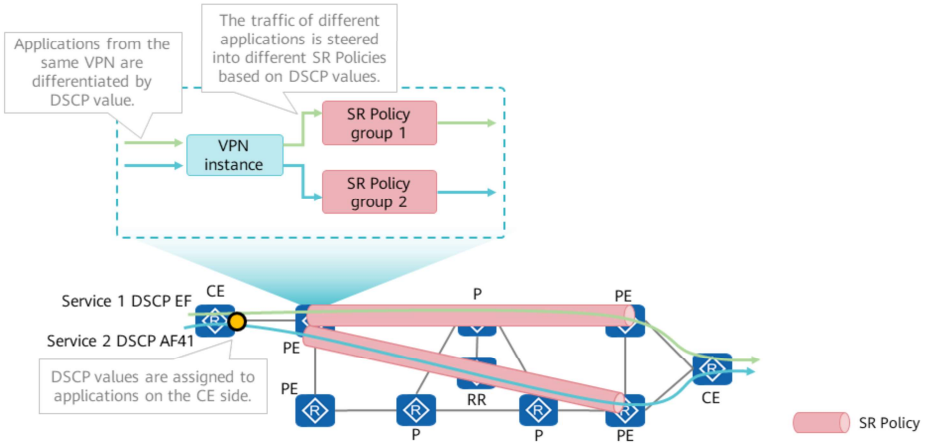
- On a network where VPN is deployed and application-based differentiated SLA path assurance is required, applications can be classified into different types and identified based on differentiated services code point (DSCP) values. The DSCP value corresponds to the color of a tunnel. Application traffic can be steered into the corresponding tunnel based on the DSCP value.



HUAWEI

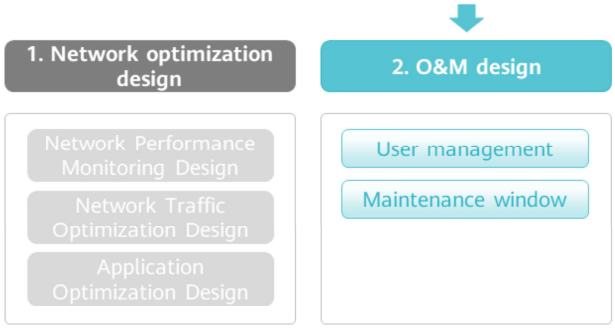
Application Optimization Design

- A different DSCP value must be assigned to each application prior to application optimization. Generally, DSCP values are assigned to applications on the CE side.
- To enable different types of applications to be carried over different tunnels, the SR Policy group+DSCP mode is generally used.



HUAWEI

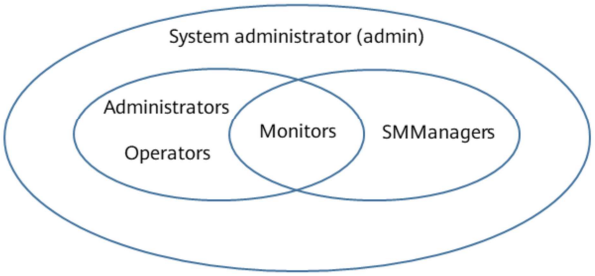
Optimization Design Roadmap for the Enterprise Bearer WAN



HUAWEI

User Type Overview

- Operations that can be performed by a user on the cloud WAN vary according to the user type. iMaster NCE-IP provides the following default user roles:
  - Administrators: Users attached to this role have permission to perform operations except managing users, querying security logs, querying personal security logs, and viewing online users.
  - SMManagers: Users attached to this role have permission to manage users, query security logs, and view online users.
  - Operators: Users attached to this role have permission to perform non-security-related operations.
  - Monitors: Users attached to this role have permission to view non-security-related functions, but do not have permission to perform the corresponding operations.



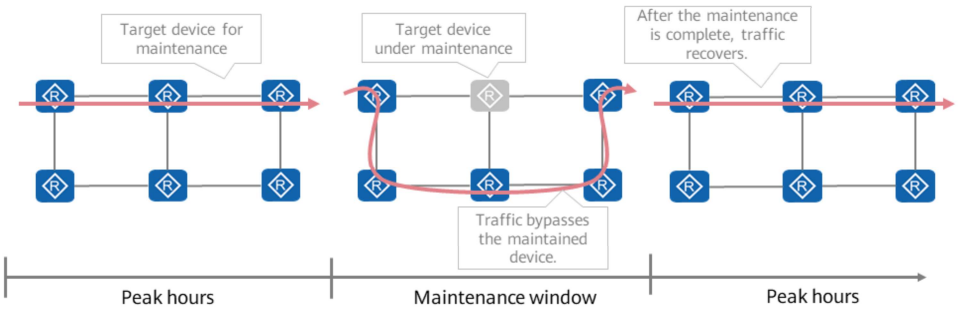
User Monitoring Overview

- By monitoring user sessions, the security administrator can learn information such as online users in the user system, IP addresses used by these users to access the system, access time, and roles of these users. When detecting that a user is attempting to perform unauthorized operations, the security administrator can send an instant message to the user or forcibly deregister the user.
  - A user session refers to the connection between a user and the system. A session starts when a user logs in and ends when the user deregisters or logs out. A user can generate multiple sessions.
  - The maximum number of online sessions allowed for a user is specified by the **Max. online sessions** parameter.
  - The function of monitoring user sessions does not involve users' personal information.

HUAWEI

Maintenance Window Overview

- When cutting over or maintaining devices and links, you can configure maintenance policies for them. The Network Path Navigation app allows you to configure maintenance policies for NEs and links.
- After the maintenance starts, the system generates new service tunnel paths that bypass the maintained devices and links. After the maintenance is complete, the system triggers path computation for traffic optimization.
- During the maintenance window, you can set the maintenance time and maintain the target devices. During the maintenance, the traffic does not pass through the maintained devices.



Quiz

1. (Multiple-answer question) Which of the following tunnels can be used to carry IPv6 services? ( )
   A. SR-MPLS BE tunnel
   B. SRv6 BE tunnel
   C. SR-MPLS Policy
   D. SRv6 Policy

HUAWEI

- ABCD

Summary

- When designing the enterprise bearer WAN, you need to design the infrastructure network first, including the physical network, IPv4/IPv6 addresses, and IGP/BGP.
- Tunnels, including SR-MPLS BE tunnels, SR-MPLS Policies, SRv6 BE tunnels, and SRv6 Policies, can be established on the infrastructure network. During tunnel design, pay attention to the tunnel path planning, traffic diversion, and best-effort forwarding mode.
- After a tunnel is established, it can carry VPN services. VPN planning must be based on enterprise service types and requirements.
- Reliability and SLA assurance are also very important for services. QoS and network slicing design helps ensure service SLA, and high reliability design for controllers, devices, and networks helps ensure service reliability.
- A network needs optimization and maintenance after it is constructed. Optimization design needs to cover both performance monitoring and traffic optimization. The maintenance scope needs to be divided based on user roles, and the network needs to be maintained at a proper time.

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.

Copyright©2021 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

