Enterprise Network WAN Interconnection Overview

HUAWEI

Foreword

- After years of development and evolution, the Internet has undergone significant changes. In the past, the Internet was centered on networks, and there were few Internet applications. As a major part of the network, the WAN takes the most important position on networks. However, the rise of cloud computing fully unleashes the potential of applications, and the Internet gradually becomes application-centric.
- Traditional WAN interconnection focuses on connectivity, and there is no strict requirement for QoS or SLA. How can WANs evolve to meet requirements of the application-centric Internet?
- After completing this course, you will be able to understand the development trend of WAN technologies and how to cope with the application-centric Internet.

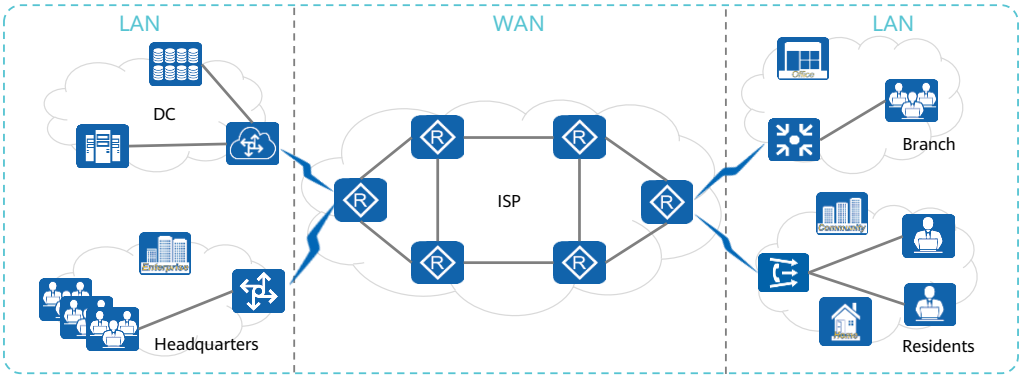
Objectives

- Upon completion of this course, you will be able to:
  - Describe the challenges of WAN interconnection in the cloud era.
  - Illustrate basic SDN concepts.
  - Explain basic concepts of SD-WAN.
  - Describe Huawei SD-WAN Solution.

Contents

What Is a WAN?

- Wide Area Network (WAN) provides interconnection services between different regions, cities, and countries. A WAN usually spans a long distance (dozens of kilometers to thousands of kilometers). To meet long-distance transmission requirements of a WAN, optical fibers are often used as the interconnection media.

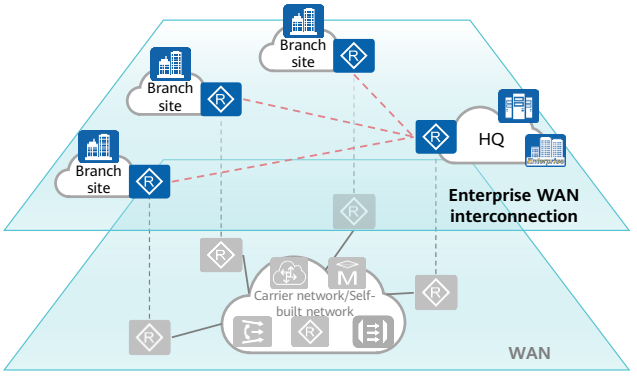


HUAWEI

- 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, data centers (DCs), branches, fixed offices, and mobile offices.
- Generally, enterprise WAN interconnection depends on a WAN built by a carrier or the self-built WAN.

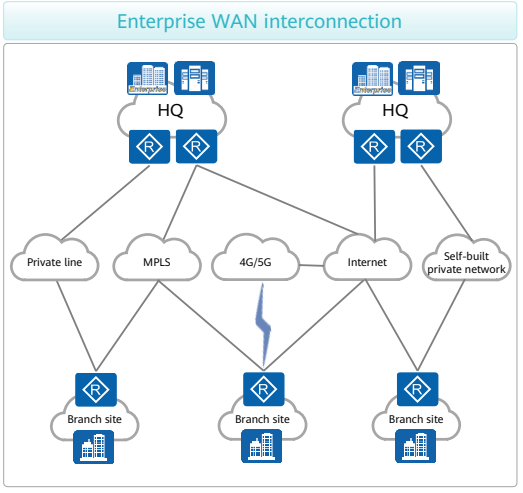


HUAWEI

- Enterprises that cannot build their own WANs usually lease ISP lines and use VPN or private line technologies to build enterprise WANs.
- Enterprises that have WAN capabilities do not need to lease ISP lines except for Internet services.

Main Enterprise WAN Interconnection Modes

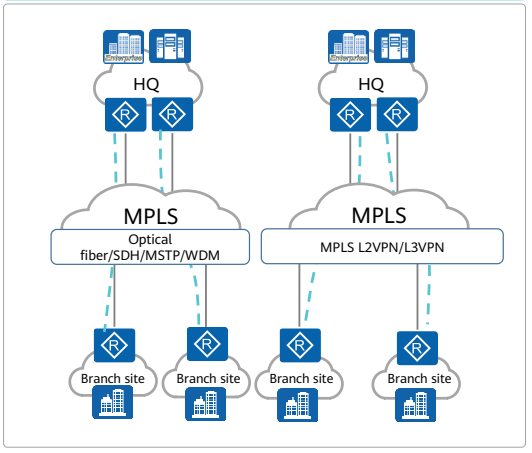
- Generally, enterprise WANs can be interconnected in the following modes:
  - Carriers' MPLS or private lines are used to connect regional networks. This mode is applicable to enterprises with high SLA requirements and is expensive.
  - The carrier Internet + VPN technology is used for connection. This 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 have network connections through self-built private lines.
- Enterprise WANs usually use a combination of the preceding connection modes.



Enterprise WAN Interconnection Technologies - MPLS and Private Line

- To ensure network reliability and security, enterprises lease MPLS or private lines from carriers when constructing enterprise WANs.
  - Private lines are expensive, but data is carried on dedicated lines, ensuring service quality and security.
  - Leasing MPLS lines from carriers is cheaper than private lines and can ensure service security. However, service reliability is not as good as that of private lines.
  - A small number of enterprises (such as transportation and electric power enterprises) have the capability of deploying optical fibers and can build their own backbone networks. For these enterprises, the cost of using MPLS or private lines is very low.

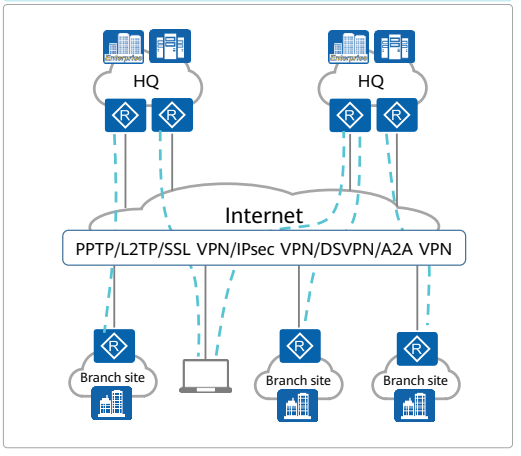
Private Line and MPLS Technologies



Enterprise WAN Interconnection Technologies - Internet and VPN

- With the development of the Internet, some enterprise services can be carried over the Internet.
- The Internet is open, 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 headquarters network.
- Internet Protocol Security (IPsec) and Generic Routing Encapsulation (GRE) technologies are used to build networks between enterprise branches or between enterprise branches and the headquarters.
- 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.

Internet and VPN technologies

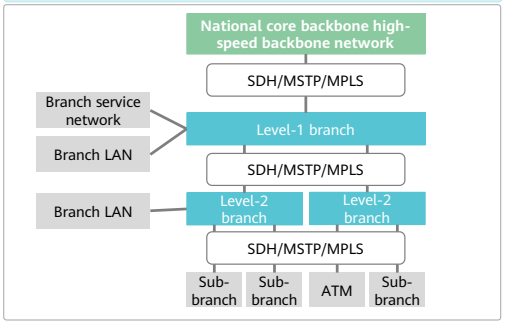


HUAWEI

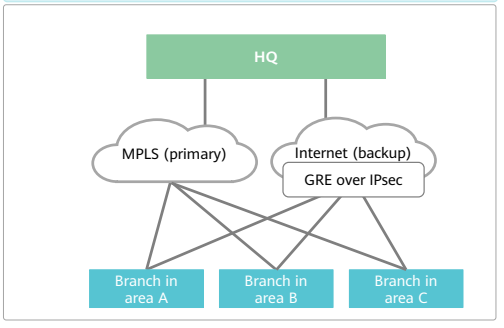
Common Application Scenarios of Enterprise WAN Interconnection

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

WAN interconnection in the financial industry



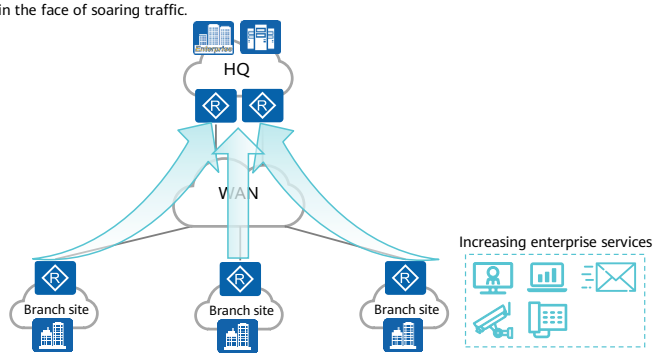
WAN Interconnection for a wine enterprise



Contents

Challenges to Enterprise WAN Interconnection Brought by Cloud Computing

- Before cloud computing is introduced, there are a few network applications, and the network service quality can be ensured only by expanding the bandwidth. Service traffic does not need to be managed in a refined manner. The Internet is mainly built based on the network.
- With the advent of cloud computing, the number of network applications is greatly increased. As a result, it is difficult for enterprises to strike a balance between line prices and service quality in the face of soaring traffic.

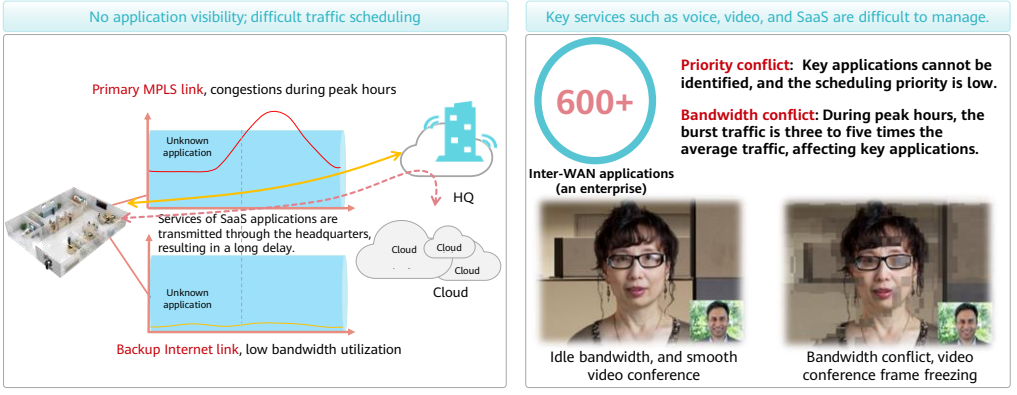


HUAWEI

- To ensure the reliability of some key services on the cloud, enterprises usually lease carriers' private lines to carry these services. However, with the increase of enterprise services, enterprises face the following problems:
  - A large amount of bandwidth is required to migrate services to the cloud. High-bandwidth private lines, especially cross-province and cross-border private lines, are expensive.
  - If the Internet is used to build VPNs, the cost can be reduced, but the reliability of key services cannot be ensured.
  - If key services need to be carried on private lines and common services need to be carried on Internet, the configuration is complex and difficult to control.

Challenges to Enterprise WAN Interconnection Brought by Multiple Services

- Enterprises have poor service traffic awareness capabilities and cannot effectively guarantee key services. In addition, the monitoring capability of service traffic is insufficient, and service traffic cannot be quickly adjusted.

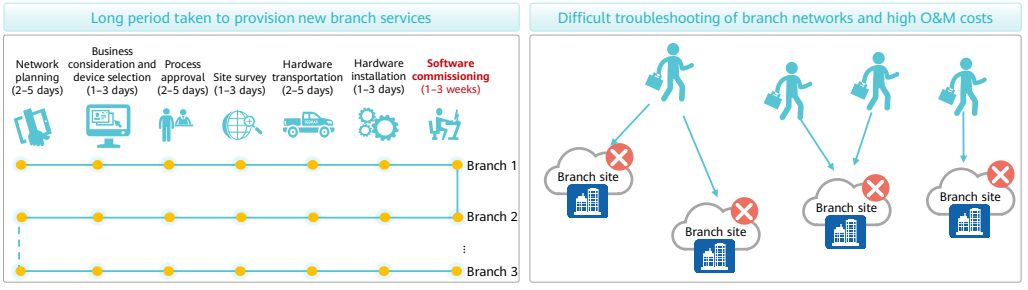


HUAWEI

- Existing traditional technologies cannot effectively monitor the actual traffic of services.
  - Traditional network management technologies (SNMP) can only monitor interface bandwidth usage.
  - The interface usage cannot be monitored based on service applications, and the quality of key services cannot be detected.
  - Service applications cannot be detected, and therefore services cannot be precisely controlled.

Challenges to Enterprise WAN Interconnection Brought by a Large Number of Branches

- With the development of companies, there will be more and more cross-city, cross-province, and cross-border branches. As a result, companies face the following problems in branch network management:
  - Too many branches result in high O&M costs.
  - It takes a long time to provision new branch services.
  - It is difficult to rectify faults on branch networks.



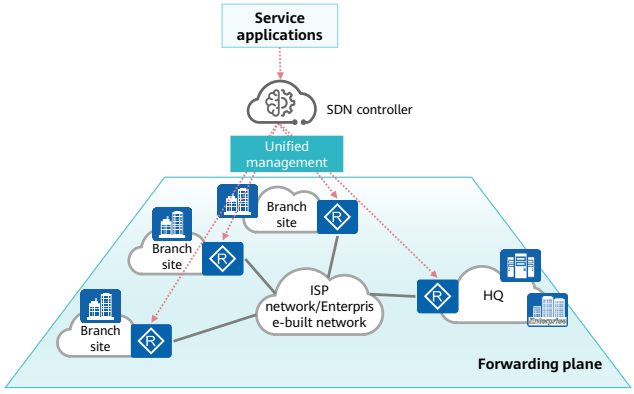
HUAWEI

- Software-defined networking (SDN) technology can be used to address the challenges brought by existing services to enterprise WANs.

Contents

What Is SDN?

- SDN decouples the forwarding plane, control plane, and service applications, allowing networks to be quickly adjusted and new services to be quickly deployed in the same way as IT applications.



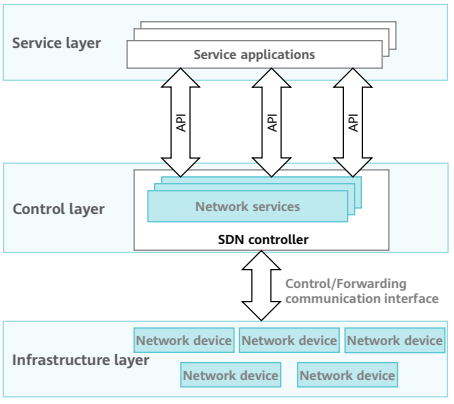
SDN Advantages

- SDN reconstructs the network architecture, and is not a new feature or function.
- SDN overcomes the disadvantages of traditional networks.

Disadvantages of traditional networks	Advantages of the SDN network
• The network architecture is distributed. A network device is a closed system consisting of hardware, an operating system, and network applications, and control and data forwarding functions are tightly coupled. • Disadvantages: ▫ Low network flexibility ▫ Complex network protocols ▫ Heavy dependency on network device vendors ▫ Difficult O&M management	• SDN provides a new network architecture that separates the network control function from the forwarding function and implements programmable control. • Advantages: ▫ Network virtualization ▫ Network automation ▫ Rapid service provisioning ▫ Openness and programmability

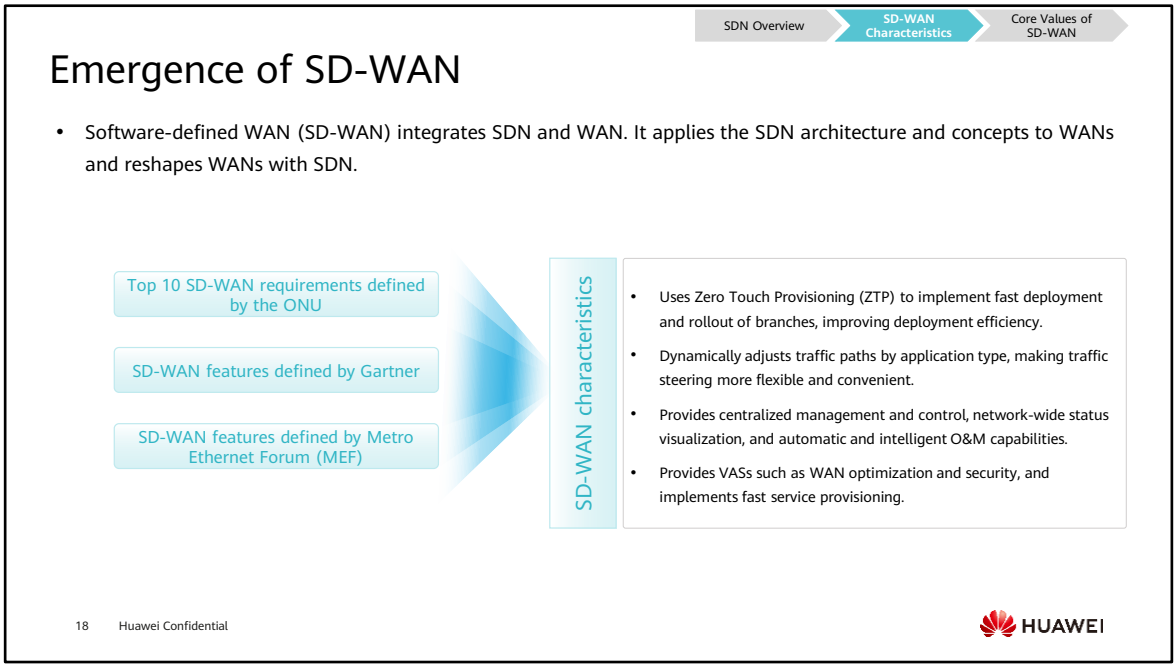
SDN Architecture

- Service layer
  - The service layer is the interaction interface at the top of the SDN architecture. It consists of various network application services and is responsible for understanding users' service requirements and orchestrating network services based on user requirements.
- Control layer
  - The control layer is the brain of SDN. It opens abstracted network functions and services to the application layer through northbound interfaces and controls the forwarding behavior of underlying network devices through southbound interfaces.
- Infrastructure layer
  - The infrastructure layer can be regarded as the core of the SDN architecture and consists of various common network devices. These network devices forward traffic based on the policies delivered by the control layer.



HUAWEI

- The SDN network architecture is provided by the Open Network Foundation (ONF).
- The SDN architecture can be divided into three layers from top to bottom: service layer, control layer, and infrastructure layer.

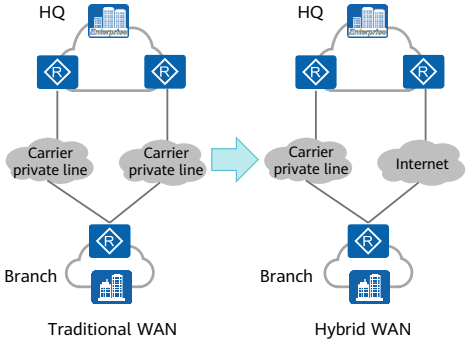


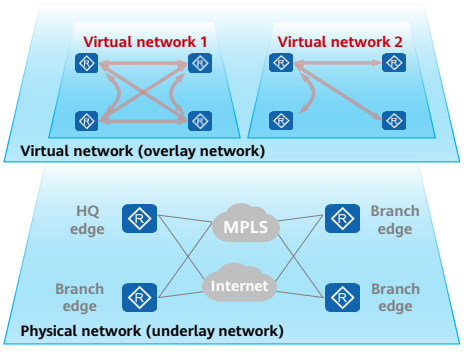
- The Open Networking User Group (ONUG) is an influential user organization led by large enterprises and mainly used by IT users. It is established by IT technology executives of well-known large enterprises in North America, and is dedicated to IT implementation and network technology transformation of large enterprises. ONUG members include large enterprises in industries such as finance, insurance, medical care, and retail. ONUG is a platform for high-end customers in North America to discuss and communicate IT requirements.
- Gartner is world's most authoritative IT research and consulting company, covering all IT industries. Gartner provides objective and fair demonstration reports and market research reports for customers in terms of IT research, development, evaluation, application, and market, and assists customers in market analysis, technology selection, project demonstration, and investment decision-making.

- MEF is a non-profit organization focusing on metro Ethernet technologies. It aims to promote wide use of Ethernet technologies as switching and transmission technologies in metro network construction. MEF aims to promote the implementation of existing and new standards, Ethernet service definitions, test procedures, and technical specifications, so Ethernet-based MANs can become carrier-class networks. MEF also provides lifecycle service orchestration (LSO)-based solutions and architectures for carriers' managed service markets, and defines northbound interfaces (NBIs) to implement multi-vendor interconnection and interworking.

Characteristics of SD-WAN: Hybrid Links

- Flexible IP overlay network based on hybrid WAN links



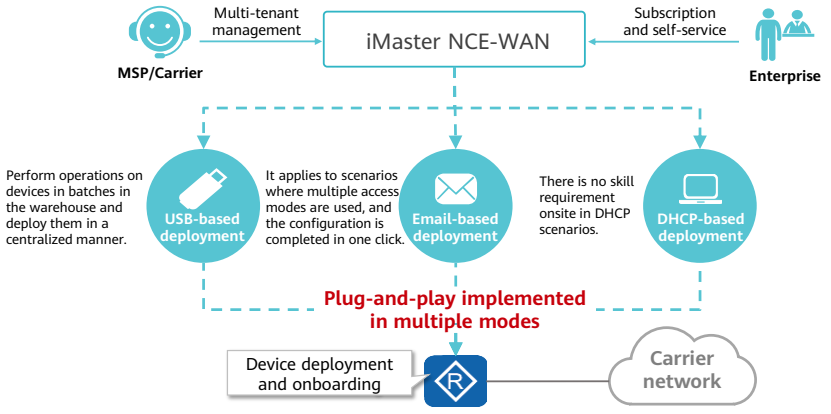


HUAWEI

- Traditional WANs are interconnected through private lines provided by carriers. With the improvement of Internet quality and expansion of coverage, the Internet can be used as a new WAN technology. In addition to MPLS private lines provided by carriers, enterprises can also use the Internet for WAN branch interconnection to implement hybrid WAN interconnection, effectively reducing WAN deployment costs and improving WAN usage efficiency.
- To implement the hybrid WAN, some key network technologies are required, that is, IP overlay and VPN technologies. IP overlay technology uses the IP tunnel encapsulation technology to encapsulate service traffic into tunnels so that the traffic can transparently traverse different WAN links. IP overlay technology cannot isolate the user overlay network from the carrier underlay network, so VPN is used. VPN adds an additional VPN field to IP packets to identify different network domains. In this way, the private network of an enterprise is isolated from the public network of a carrier, and even different private networks of an enterprise are isolated from each other. IP overlay and VPN are usually deployed together.

Characteristics of SD-WAN: Plug-and-Play

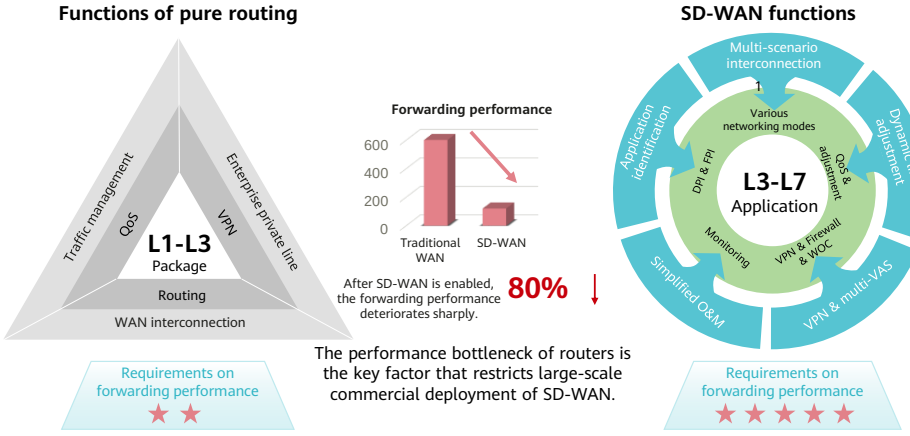
- Device plug-and-play, achieving fast service rollout



- New branches of an enterprise need to go online quickly. After devices are delivered to the site, common personnel at the site can use the plug-and-play function to provision the branch network without professional network skills. To achieve this, a centralized network control system is required to register network devices upon power-on and manage them, and to automatically and remotely deliver network configurations required by devices at the site on demand, implementing fast branch service rollout. For example, the network control system may send an email to a deployment engineer, and the email includes configuration of a site device. Deployment personnel connect and power on devices at the site, and load the configuration in the email to the devices through simple operations. Then the devices automatically register with the network control system.
- The plug-and-play mode effectively reduces skill requirements for deployment personnel and shortens the time for provisioning and adjusting branch services from several days or even months to hours, greatly improving flexible service provisioning of branch networks.

Characteristics of SD-WAN: High Performance

- High-performance branch devices are used to process all services including application-centric services.



- Based on the "multi-core CPU+NP" forwarding architecture, high-performance branch devices use professional NP chips to implement fast forwarding of L2-L4 traffic and efficient QoS processing. With open programmability of the multi-core CPU, high-performance branch devices can integrate various hardware-based intelligent acceleration engines, for example, hierarchical QoS (HQoS), application identification, and IPsec acceleration, and import them to chips to provide high-performance processing capabilities for L3-L7 services.

Characteristics of SD-WAN: Automatic Network Orchestration

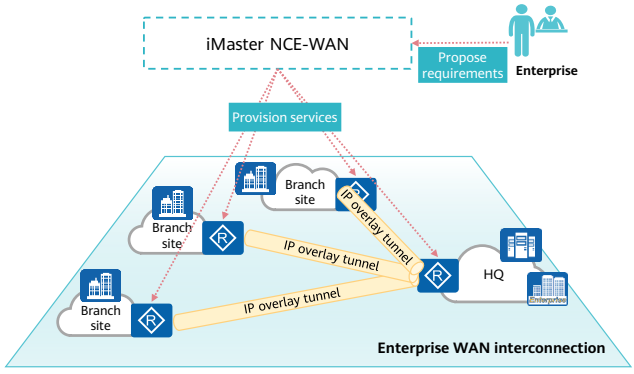
- Service- and intent-orientation, implementing network orchestration and automatic provisioning.

Traditional network

- Traditional network service provisioning requires professional network engineers to perform planning, configuration, and O&M, and then run commands or use the NMS software to configure devices one by one based on the planned services.

SDN network

- The SDN network uses a centralized network control system to abstract, orchestrate, and automatically provision network services on demand. It shields technical implementation details of the network and opens only service-oriented interfaces and parameters to users.



HUAWEI

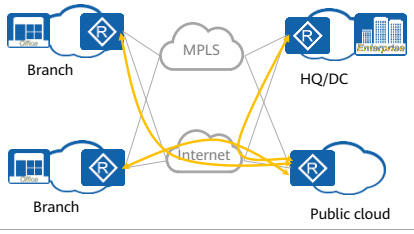
- For example, an enterprise wants to connect multiple branches through the headquarters. Traditionally, network parameters, such as interfaces, IP addresses, routing protocols, security, and VPN, are manually configured on each branch device. If there are 100 sites, devices need to be configured one by one. In addition, you need to configure special routes for the headquarters to forward all traffic. The entire process is complex and requires operators to be familiar with network technical details such as switching and routing. Service provisioning is slow and manual operations are prone to errors.
- In SD-WAN, assume that a centralized network control system serves as the control center of the network and can manage all branch devices. The network control system, as the brain of the network, displays a network connection operation primitive that can be understood by a non-network professional user. For example, a user's instruction is to enable 100 sites, and the headquarters site is used for interconnection, and the user does not need to enter network-level parameters. After receiving the request, the network control system automatically translates the original network connection requirement into traditional network configuration operations that can be understood by network devices, that is, operations such as routing and VPN that are manually performed in the traditional method. Then it further transmits the operation to the devices at the branch site, thereby implementing automatic provisioning of network services. The entire network service provisioning process is automatically completed by the network control system, which effectively lowers the network skill requirements of users and reduces network operations. In addition, automatic conversion of the network control system is not error-prone, which greatly improves the WAN service provisioning efficiency and user experience.

Characteristics of SD-WAN: Efficient Cloud Interconnection

- On-demand and efficient cloud connection

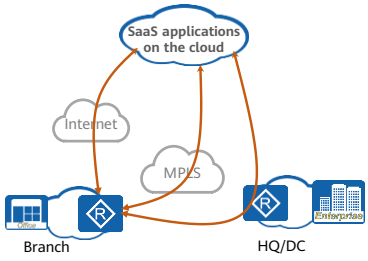
Connecting to a public cloud

- A cloud site requires an NFV-based device as a gateway to connect enterprise branches to the public cloud. Devices on the cloud remotely schedule public cloud APIs and resources through the centralized network control system to connect devices on the cloud to the branch network.



Connecting to the SaaS cloud

- Enterprises can access remote SaaS applications on the cloud through the WAN in the following ways: 1. Internet 2. Direct access through MPLS 3. Headquarters

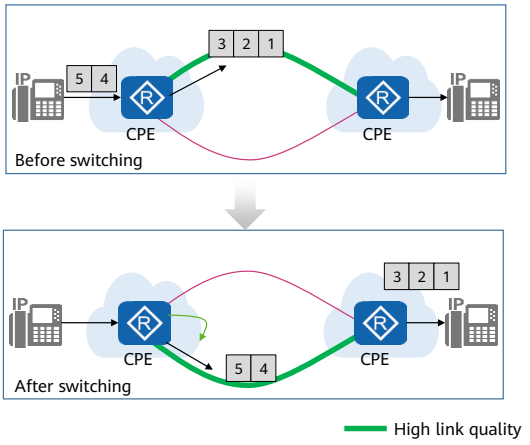


HUAWEI

- With the advent of the cloud era, enterprise WANs need to open the traditional closed network architecture and flexibly connect various cloud resources. Cloud resources closely related to enterprise WANs include IaaS basic cloud services and SaaS cloud applications. With the popularity of public clouds, more and more enterprises are considering migrating their IT systems to public clouds. An enterprise's system in the public cloud can be considered as a special branch site, that is, a cloud site. A cloud site also requires a device that functions as a gateway to connect enterprise branches and the public cloud. Because the device is deployed on the cloud, the device must be based on NFV. Devices on the cloud need to be quickly created and connected to enterprise branch sites in real time. Therefore, a centralized network control system is required to remotely schedule public cloud APIs and resources, automatically start devices on the cloud, and connect the public cloud and branch networks.
- To access SaaS applications efficiently, SD-WAN needs to have the capability of optimizing SaaS access paths. When an enterprise accesses a remote SaaS application on the cloud through the WAN, multiple paths may be available, for example, access through an Internet, an MPLS network, or the HQ. A branch needs to be capable of perceiving network Service Level Agreement (SLA) quality of each optional path in real time, with the help of the centralized network control system. It can adjust and select the optimal SaaS access path in real time.

Characteristics of SD-WAN: Intelligent Traffic Steering

- Intelligent traffic steering, ensuring application experience
  - The introduction of hybrid WAN provides multiple WAN links for enterprise service traffic.
  - Different WAN links have different network quality. High-value applications are preferentially transmitted through high-quality WAN links. In this case, route selection based on application SLA needs to be implemented.

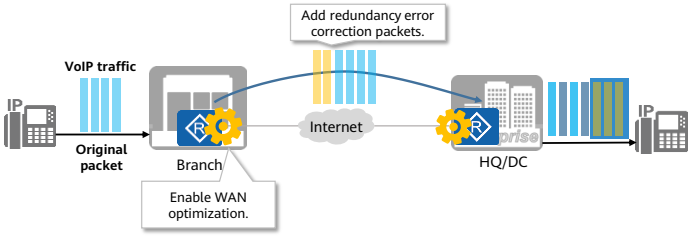


HUAWEI

- The introduction of hybrid WANs provides multiple WAN links for enterprise service traffic. Different WAN links offer varying levels of network quality. For example, the price of MPLS private lines is high and the corresponding link SLA quality is guaranteed. Although Internet links are improved greatly and can carry applications that support high bandwidth, the delay and packet loss occur frequently on Internet links and SLA is not guaranteed, making the Internet unsuitable for delay-sensitive voice and video services. The solution to this issue is to implement link selection based on the SLA quality requirements of applications. Specifically, the quality of different WAN links is measured, and the requirements of applications regarding network quality (such as packet loss rate, delay, and jitter) are defined. Among all WAN links that meet SLA requirements of applications, enterprise users can define route selection policies to preferentially transmit high-value applications through high-quality WAN links. For example, voice traffic (high-value traffic) is preferentially transmitted over high-quality MPLS links. The network has the capability of dynamically adjusting paths. When the MPLS link quality deteriorates and does not meet application SLA requirements, the network can automatically switch voice traffic to another WAN link that satisfies application SLA requirements.

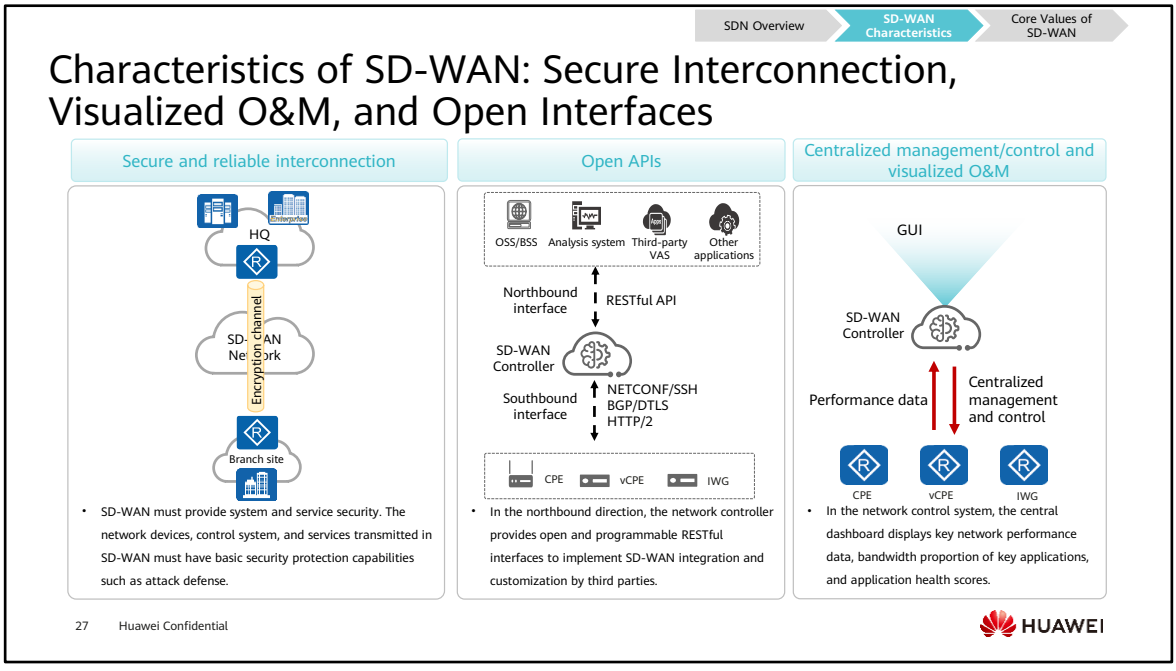
Characteristics of SD-WAN: WAN Optimization

- WAN optimization, improving WAN link quality

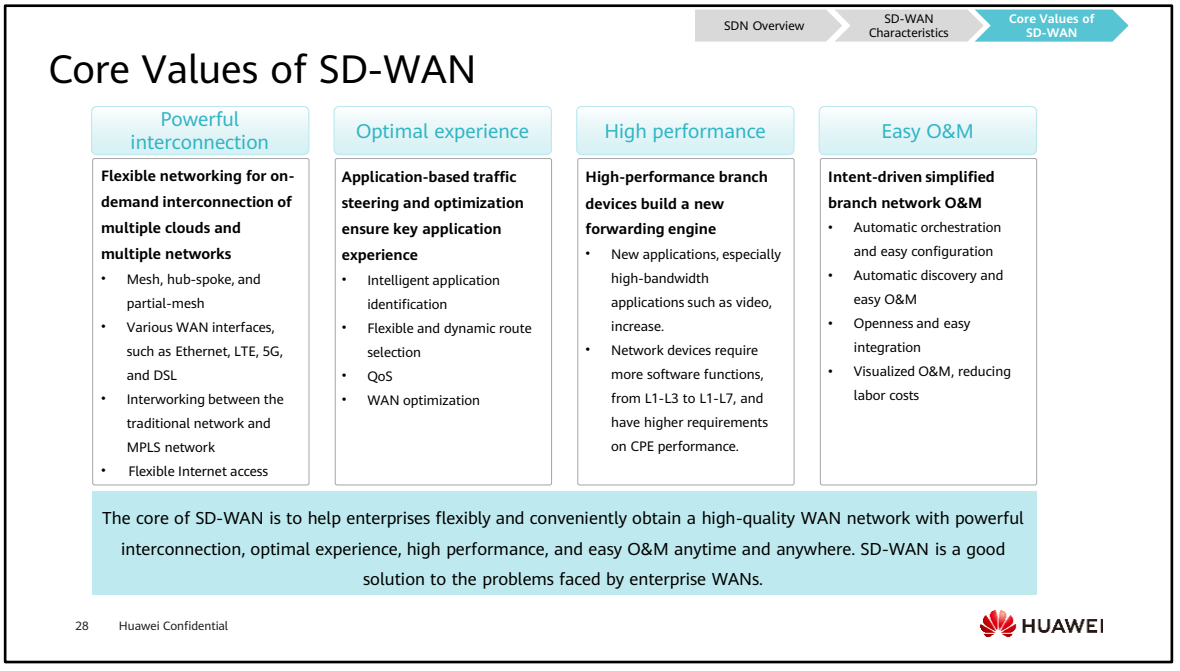


- When the quality of a WAN link deteriorates, for example, packet loss or long delay occurs, WAN optimization technologies need to be used to improve network fault tolerance and ensure data transmission quality. Common WAN optimization technologies include transmission optimization, data optimization, and packet loss concealment optimization.

- The prerequisite for application-based traffic steering is that there are WAN links that meet SLA quality requirements. When the quality of a WAN link deteriorates, for example, packet loss and long delay occur, WAN optimization technologies need to be used to improve network fault tolerance and ensure data transmission quality. Common WAN optimization technologies include transmission optimization, data optimization, and packet loss concealment optimization.
- Forward Error Correct (FEC), a common packet loss concealment technology, is used as an example to describe the value of WAN optimization. The core of FEC is to improve the fault tolerance of application data to links with deteriorated quality by reconstructing or optimizing the data transmission protocol. FEC packets are encapsulated in normal data flows to record packet digest information. During data transmission, some service packets are lost due to packet loss on the network. At the receive end, the device can restore the lost packets based on the FEC field to ensure the integrity of data transmission.
- The core of WAN optimization is to use extra compute or storage resources to improve network transmission performance and application experience. WAN optimization is an effective technical means to ensure that user service experience is not affected when the link quality deteriorates on an enterprise WAN.



- Secure and reliable interconnection
  - The transmission of enterprise WAN application data requires security assurance, including system security and service security.
    - System security includes the security of network devices and network control systems, which must be connected to the WAN and may also be connected to the Internet. Because of this, they must have basic security protection capabilities, such as attack defense.
    - For enterprise services, service data is transmitted on the WAN. Therefore, authentication and encryption are required to prevent data leakage.
- Centralized management, visualization, and easy O&M
  - Enterprises use the centralized management and O&M system. This system must be able to remotely manage branch devices, as well as display the WAN topology and remotely monitor alarms, logs, and other key event information of each branch device in real time.
  - In the centralized network control system, the central dashboard displays key network performance data, bandwidth proportion of key applications, and application health scores.
- Northbound open APIs
  - The network controller provides open and programmable northbound RESTful interfaces. Third-party BSS/OSS software interconnects with the network controller through northbound APIs.



- Powerful interconnection: flexible networking for on-demand interconnection of multiple clouds and multiple networks
  - SD-WAN flexibly uses hybrid links, such as optical fibers, DSL links, and LTE links, to quickly provision networks and reduce link costs.
  - In addition, SD-WAN provides a broad variety of networking models, including hub-spoke, full-mesh, partial-mesh, hierarchical networking, and IaaS/SaaS access, meeting requirements of different enterprise services.
- Optimal experience: Application-based traffic steering and optimization ensure key application experience
  - SD-WAN can monitor the quality of multiple links in real time, detect link connectivity, and record the packet loss rate, delay, jitter, and other real-time status information.
  - SD-WAN also provides multiple application identification methods to accurately identify application information in traffic.
- High performance: High-performance branch devices build a new forwarding engine.
  - In contrast to traditional enterprise branch devices that feature packet forwarding at Layers 1 to 3, SD-WAN branch devices deliver application-based full service processing at Layers 3 to 7. Their high-performance forwarding capabilities help enterprise branches build a new forwarding engine to allow enterprise services to operate normally.

- Easy O&M: intent-driven simplified branch network O&M
  - SD-WAN inherits the two basic design concepts of SDN: centralized control and intent-driven. This makes it possible to implement intent-based centralized management and control on the entire network. Based on centralized control, SD-WAN provides the network-wide monitoring function to obtain branch link status in real time, making the network-wide status visualized.

Contents

Architecture of Huawei SD-WAN Solution

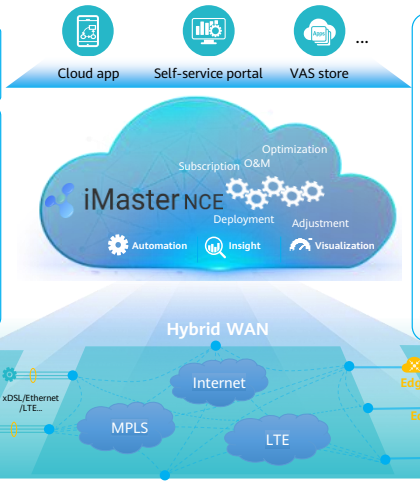
Customer benefits

Enterprise value

- Reduced O&M cost
- Improved WAN usage

Carrier value

- **Minute-level service provisioning** (location independent)
- **Improved O&M efficiency:** cloud management & automation
- **Revenue increase**
  Extended B2B service domain: VAS, connectivity
  managed LAN
- **Smooth evolution, openness, and quick integration**
  RESTful API, uCPE/vCPE



Key technologies

5G uplink: All CPE series support 5G.

- Large bandwidth: 230 Mbit/s for uplink and 2 Gbit/s for downlink
- Full frequency: 5G/4G/3G/2G
- Dual-architecture: full support for NSA/SA

High performance: no congestion during forwarding

- CPU+NP heterogeneous forwarding architecture
- High performance, meeting SD-WAN development requirements in the next five years

Optimal experience: intelligent traffic steering, ensuring experience of key applications

- Application-based intelligent traffic steering, on-demand 5G+fiber scheduling
- A-FEC ensures that no frame freezing or artifact occurs in case of 20% video packet loss.

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

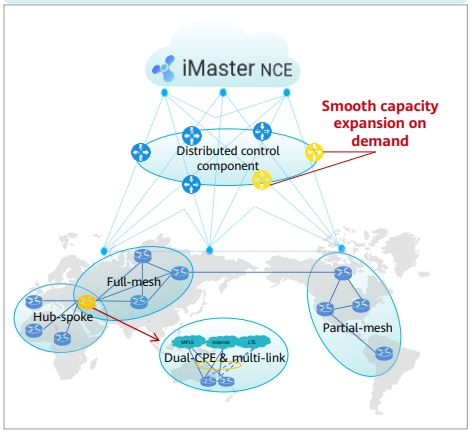
- Multiple ZTP modes, and branch network deployment in minutes
- Visibility of application, branch, device, and link status, centralized management, and simplified O&M

HUAWEI

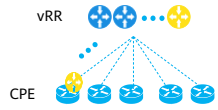
Flexible Networking and Forwarding-Control Separation

Forwarding-control separation architecture

Distributed control components are deployed on the CPE or cloud as required, simplifying the network control topology.



Large-scale networking and flexible expansion



Distributed and on-demand vRR expansion

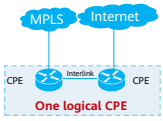
- **Two deployment modes: independent deployment/co-deployment**

Flexible networking

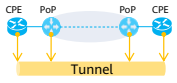
On-demand orchestration of 20+ networking models

Full-mesh, partial-mesh, hub-spoke, area-based networking, and hierarchical networking

Dual-CPE & multi-link, multi-PoP, etc.



Dual-CPE & multi-link

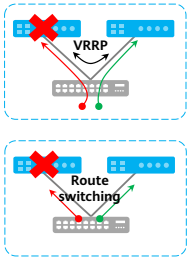


Multi-PoP E2E overlay

HUAWEI

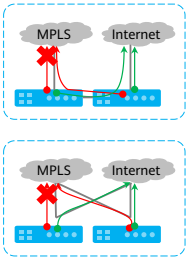
High Availability of Branch Interconnection Services

E2E reliability design ensures high availability of branch interconnection services



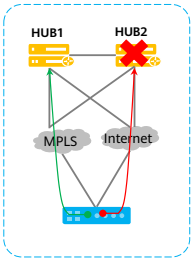
CPE redundancy

Two CPEs at a site back up each other, and they support VRRP or route switchover.



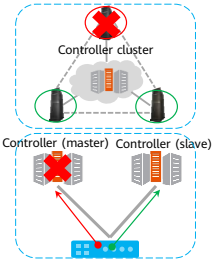
Multi-link redundancy

Underlay link redundancy is supported. When a link is faulty, services are automatically switched to other links.



Hub redundancy

The solution supports two CPEs at a single hub and dual hubs. When the hub is faulty, the site node automatically switches to the backup hub.

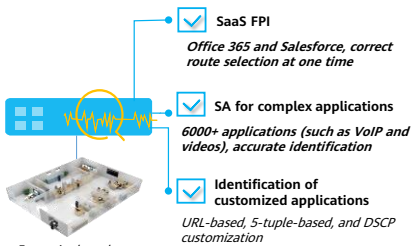


Controller redundancy

The cluster is deployed, providing high reliability, remote disaster recovery, and automatic switchover.

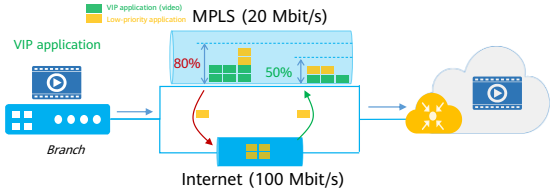
Guaranteeing User Experience in Key Applications

Three identification engines and intelligent identification of network-wide applications at layers 3 to 7



The identification rate of all services from Layer 3 to Layer 7 is doubled.

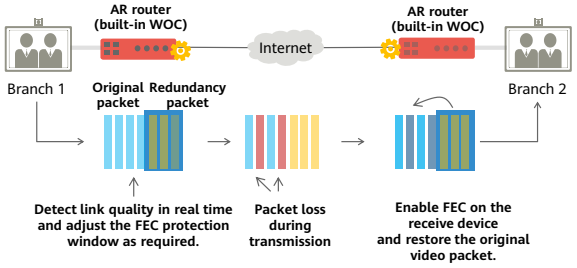
Application experience-prioritized, supporting comprehensive route selection considering bandwidth and link quality



- **Application quality-based route selection**
- **Load balancing based on link bandwidth**
- **Bandwidth utilization-based route selection**
  - If the bandwidth usage of the primary link group is higher than 80%, low-priority applications are switched to the backup link group.
  - If the bandwidth usage of the primary link group is less than 50%, low-priority applications are switched back to the primary link group.

HUAWEI

Intelligent A-FEC — 20% Packet Loss and No Frame Freezing



Networked Environment	Packet Redundancy Rate	Before Optimization	After Optimization
30 Mbit/s + 65 ms delay + 5% packet loss	7%	No frame freezing and low definition	No frame freezing and high definition
30 Mbit/s + 65 ms delay + 10% packet loss	11%	Artifact and frame freezing	No artifact or frame freezing
30 Mbit/s + 65 ms delay + 20% packet loss	22%	Serious artifact and frame freezing	No artifact or frame freezing

A-FEC: The FEC protection window and protection mode are automatically and dynamically adjusted based on the link quality.

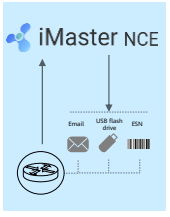
Huawei **20%** packet loss, no frame freezing, no artifact

Third party: **3%** packet loss and **artifact**

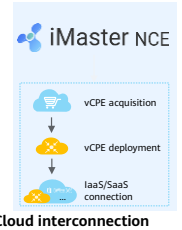
Intelligent A-FEC, Low Overhead, and High-quality Experience

- Redundancy coding is performed on historical frame information. After normal frames are sent, the corresponding redundancy frame is forwarded. The receive end can use the received redundancy frame to restore the lost packet.
- Huawei adaptive-FEC (A-FEC) uses the intelligent data analysis engine to adjust the FEC protection window and protection mode based on the link quality to achieve high recovery and low redundancy.

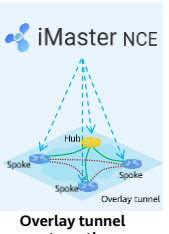
Full-Process Automation



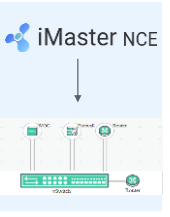
CPE deployment automation
ZTP



Cloud interconnection automation
Automatic IaaS/SaaS interconnection configuration

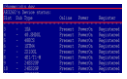


Overlay tunnel automation
Dynamic tunnel orchestration based on traffic policies



Automated service chain orchestration for VNFs
VAS provisioning in minutes

O&M automation
Network-wide visualized and automatic O&M



Traditional WAN: CLI

iMaster NCE

iMaster NCE-WAN: automation and visualization

Full-process automation of enterprise interconnection

Various built-in networking templates

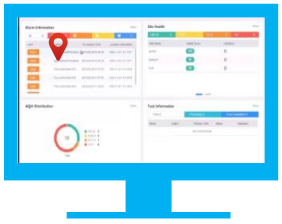
Configuration process guidance

Drag-and-drop

HUAWEI

Visualized O&M

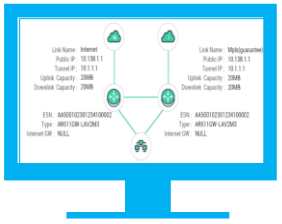
Quickly obtain abnormal traffic



Real-time alarm monitoring

- User-defined dashboard (role or preference)
- Network-wide real-time alarms (minute-level)

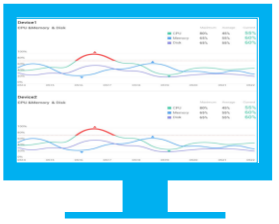
Quickly locate faulty devices or sites



Topology status visualization

- Topology display based on sites and links
- Enterprises obtain the status and performance information of sites and links in real time.

Optimize WAN investment and configuration policies



45+ user-defined views

Site/Link/Application/Device/User health view

- Site bandwidth utilization
- Sites with Top N throughput
- Top N application traffic
- Link throughput trend

Quiz

1. (Multiple-answer question) Which of the following are disadvantages of traditional WANs?
   A. Low network flexibility
   B. Complex network protocols
   C. Enterprises are highly dependent on network device vendors.
   D. Difficult O&M

- 1. ABCD

Summary

- Generally, enterprise WAN interconnection depends on WANs built by carriers or self-built WANs.
- After services are cloudified, flexible networking and fast service optimization are required.
- Huawei SD-WAN Solution uses iMaster NCE-WAN to quickly deploy networks and uses devices to detect applications and flexibly select paths.

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.

WAN Interconnection Technologies and Typical Scenarios of Enterprise Networks

HUAWEI

Foreword

- The WAN provides interconnection services for different users. Because users have various requirements, the WAN provides different interconnection technologies.
- WANs can be interconnected through transmission private lines, MPLS private lines, or the Internet. Different interconnection modes require different technologies.
- This course compares application scenarios and differences of technologies to help trainees better understand WAN interconnection modes.

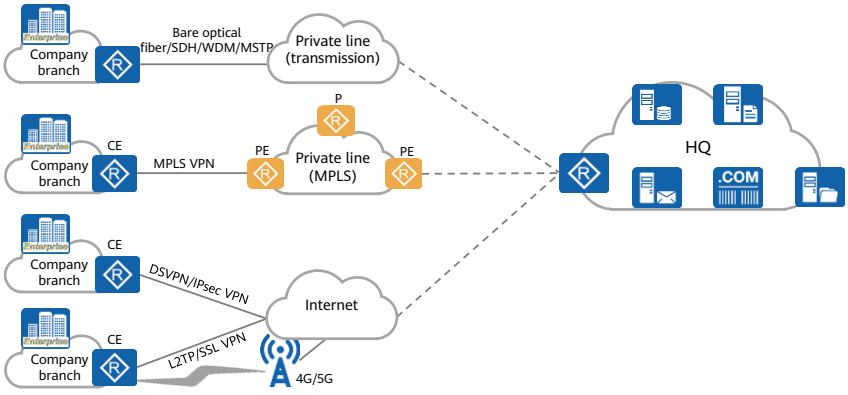
Objectives

- Upon completion of this course, you will be able to:
  - Describe the traditional WAN interconnection solution.
  - Describe technologies used on WANs.
  - Describe three application scenarios of SD-WAN.

Contents

Typical WAN Interconnection Architecture of an Enterprise Network

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



HUAWEI

Comparison Between Private Line and VPN Technologies

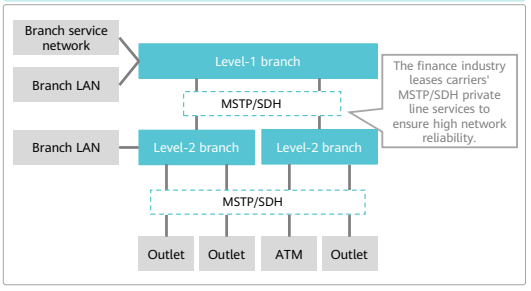
- Private lines were introduced very early. They can meet interconnection requirements of enterprises and ensure reliability and security. However, private lines are expensive.
- With the network development, VPN technologies start to occupy more market shares. However, some industries demanding high security and reliability, such as the financial industry, still want to use private line technologies.
- Selecting private line or VPN technologies depend on company's 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. Scalability is implemented only if the network 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 connected to a private line, 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.
Upgrade	Depending on the telecom department	Device upgrade is convenient.

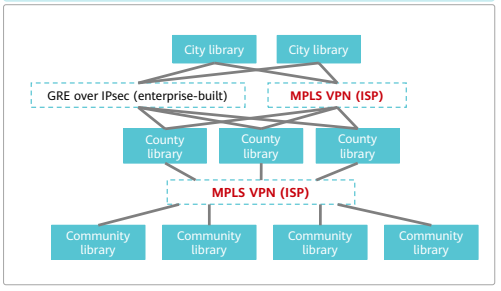
Introduction to Carriers' Private Lines

- 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 optical fibers, which are expensive but have excellent performance.
- MPLS VPN is another type of private line service provided by carriers. MPLS VPN private lines provide slightly lower performance than transmission private lines but are less expensive.

Carrier's private line (private line network in the financial industry)



Carrier's private line (MPLS VPN network of a provincial library)

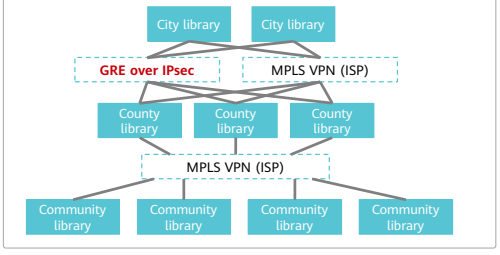


HUAWEI

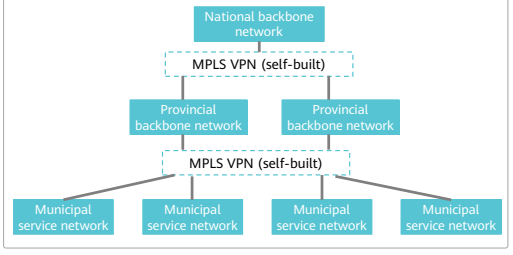
Introduction to Enterprise-Built Private Line and VPN

- Enterprises can establish VPNs, such as SSL VPN, DSVPN, and IPsec VPN, through carriers' networks.
- Some large enterprises have the capability to lay out optical fibers by themselves and can set up enterprise MPLS VPN private lines. However, few enterprises have the capability to 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

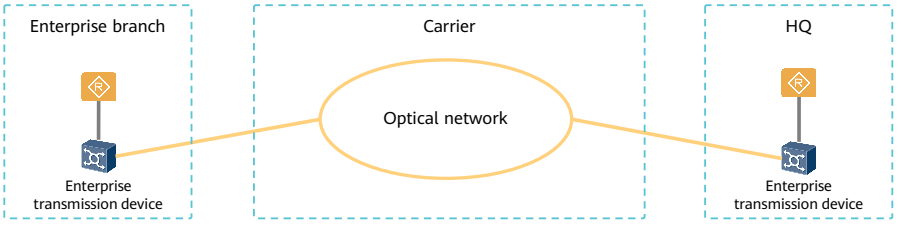
Contents

Overview of Private Line Technologies

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

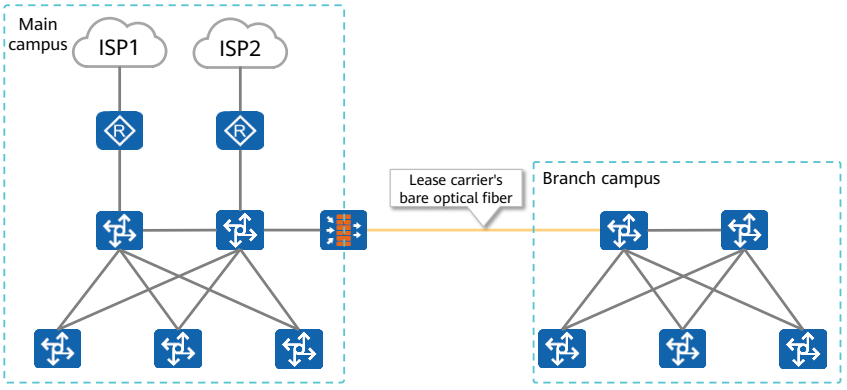
Introduction to Bare Optical Fibers

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



Application Scenario of Bare Optical Fibers

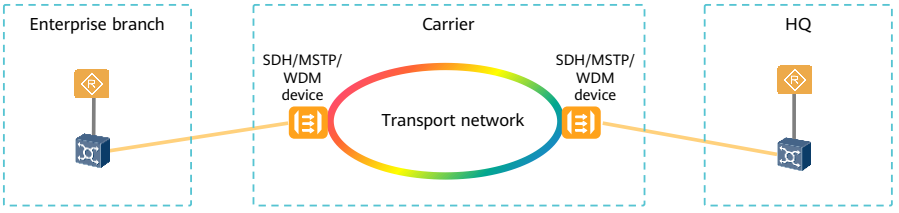
Application scenario of bare optical fibers



- Bare fibers can be leased between the main campus and branch campus in the same city to connect the two campuses, simplifying network management and access authentication management.

Introduction to SDH/MSTP/WDM Private Lines

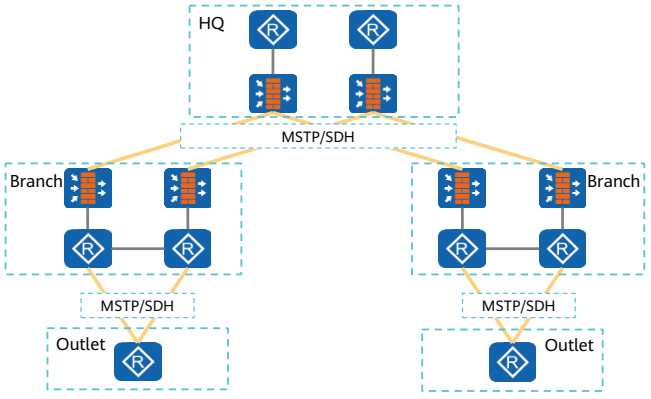
- Enterprises that require long-distance transmission and high network reliability and security can lease SDH/MSTP/WDM private lines.
- This type of private line is a transmission private line. Tenants exclusively occupy part of the bandwidth of the transmission private line. Because multiple users share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



- 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 on a single optical fiber. The transmission bandwidth of WDM devices is high, and the bandwidth on the live network can reach 8 Tbit/s.

Application Scenarios of SDH/MSTP/WDM Private Lines

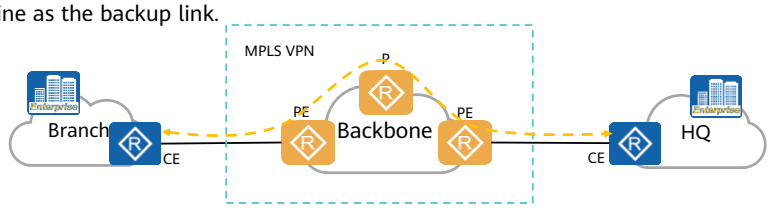
Application scenarios of SDH/MSTP/WDM private lines



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

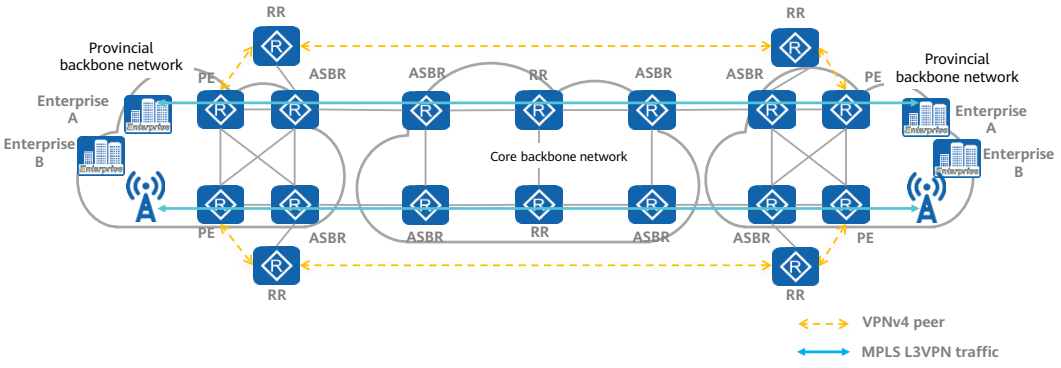
MPLS VPN Private Line

- 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 have security requirements can use MPLS VPN line as the primary link and GRE over IPsec line as the backup link.



Application Scenario of MPLS VPN

Application scenario of MPLS VPN



- The MPLS L3VPN solution is widely used on the live network. Some large-scale 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-scale inter-AS MPLS L3VPNs. Option B applies to medium- and large-scale inter-AS MPLS L3VPNs. Option C applies to large-scale or super-large-scale inter-AS MPLS L3VPNs.

Contents

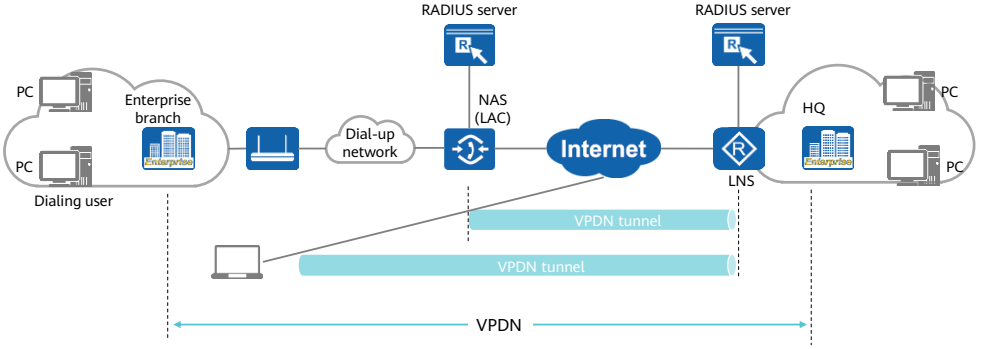
Overview of VPN Technologies

- VPN technologies are widely used in scenarios where enterprises build their own Internet.
- VPN technologies can be classified into the following three types based on the usage:
  - Access VPN (virtual private network for remote access): also called dial-up VPN or VPDN. Generally, L2TP VPN technology is used.
  - Intranet VPN (internal virtual private network of an enterprise): connects gateways and connects resources of the same company through the company's network architecture. Generally, GRE or DSVPN technology is used.
  - Extranet VPN (extended internal virtual private network of an enterprise): is used to build an extranet with the enterprise network of a partner. Generally, SSL VPN technology is used.

- By service usage:
  - Access VPN: enables mobile employees, remote office employees, and remote small-sized offices to establish private network connections with enterprise intranet and extranet through the public network. There are two types of access VPN connections: client-initiated and NAS-initiated VPN connections.
  - 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 the public network.
- According to 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

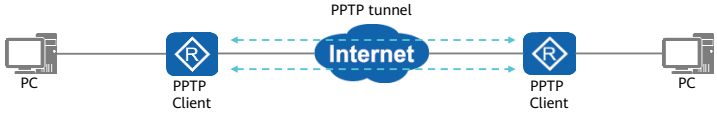
- Access VPN uses VPDN technology. VPDN is a type of VPN service based on dial-ups. It can be used for enterprise interconnection or remote access to enterprise networks.



- 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 the first VPN protocol and has been developed for more than 20 years. This protocol relies on encryption, authentication, and 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.

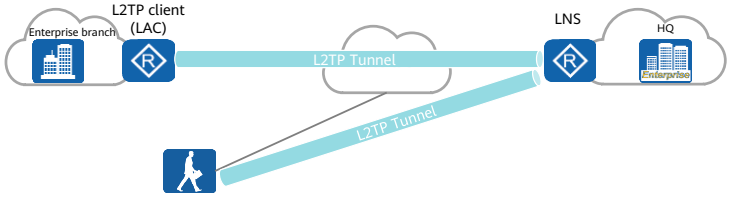


L2F Overview

- After PPTP was released, Cisco developed L2F to try to fix defects of PPTP.
- L2F encapsulates packets of link-layer protocols (such as HDLC, PPP, and ASYNC) for transmission. Therefore, the link layer of a network is independent of link-layer protocols of users.
- L2F is a secure communication tunneling protocol, but it does not provide the standard encryption method. Therefore, it has become an outdated tunneling protocol.

L2TP Overview

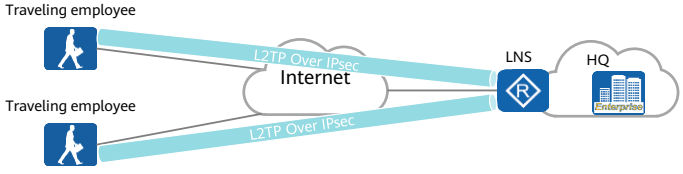
- L2TP, an open standard of IETF, combines advantages of PPTP and L2F. L2TP is especially suitable for setting up a VPN in remote access mode and has become a de facto industry standard.
- L2TP is only a tunneling protocol and does not provide encryption. Therefore, L2TP is usually used together with IPsec.
- L2TP is one of commonly used enterprise interconnection technologies. When L2TP is used, the AAA server is required. L2TP is a good choice for constructing an L2VPN.



Application of Access VPN on the Live Network

- 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.

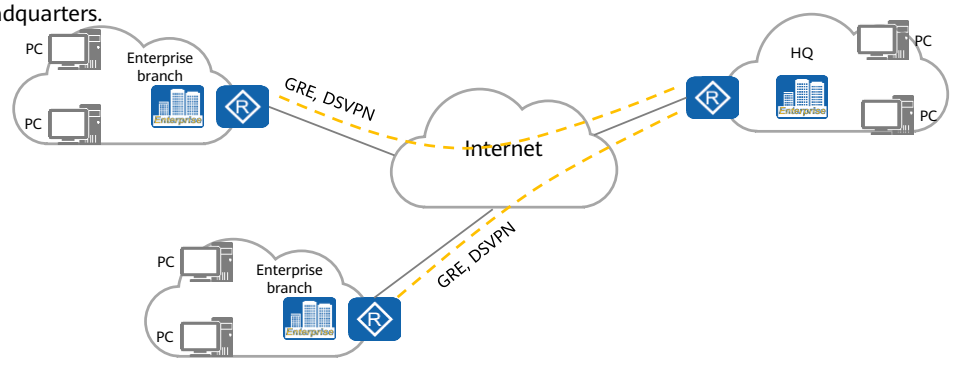
Application of L2TP over IPsec on the live network



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

Intranet VPN Overview

- Intranet VPN technology is used to construct a VPN between gateways based on the Internet. The main technologies used are GRE and DSVPN.
- GRE and DSVPN technologies are used to establish a VPs between enterprise branches and the headquarters.

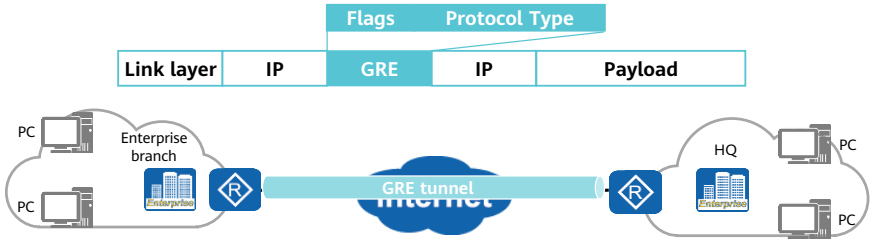


HUAWEI

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

GRE Overview

- 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.

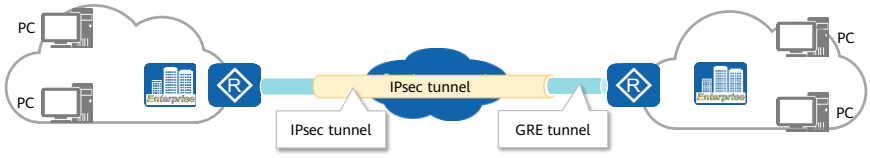


- GRE is a Layer 3 tunneling technology. A GRE tunnel is a virtual P2P connection that transmits encapsulated data packets.
- The two ends of the GRE tunnel are tunnel interfaces which encapsulate and decapsulate data packets. The tunnel interface that sends encapsulated packets is called the tunnel 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, Router_A'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:
  - Router_B receives an IP packet from its physical interface connected to the Internet and checks the destination address. If the destination is Router_B and the protocol ID in the IP packet header is 47 (indicating GRE packet), Router_B 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 module.
  - The IP protocol module forwards the packet to the enterprise headquarters.

Overview of GRE over IPsec

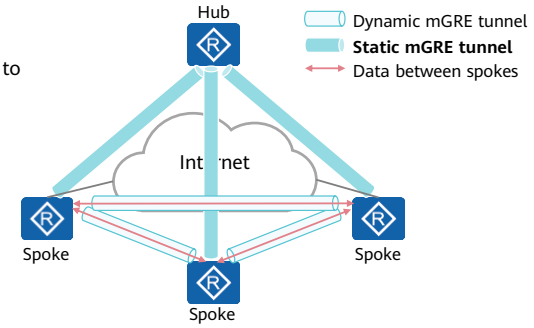
- GRE does not support encryption. IPsec supports only the IP protocol and does not support multicast.
- GRE over IPsec combines advantages of both GRE and IPsec, and offsets their disadvantages.
- GRE over IPsec is a point-to-point VPN technology commonly used by enterprises.



- 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 VPN networks.
- 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.

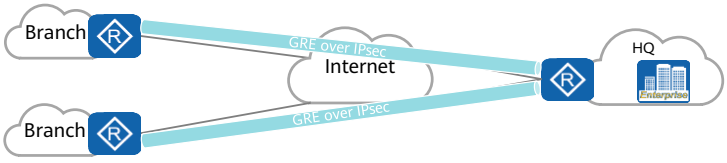


- DSVPN resolves the following defects of GRE over IPsec:
  - All traffic must pass through the hub.
  - The hub configuration needs to be modified when a site is added.
  - If spokes use dynamic addresses, problems may occur when P2P GRE is deployed.

Application of Intranet VPN on the Live Network

- On the live network, intranet VPN is mainly used for interconnection between enterprise branches and the headquarters and 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.

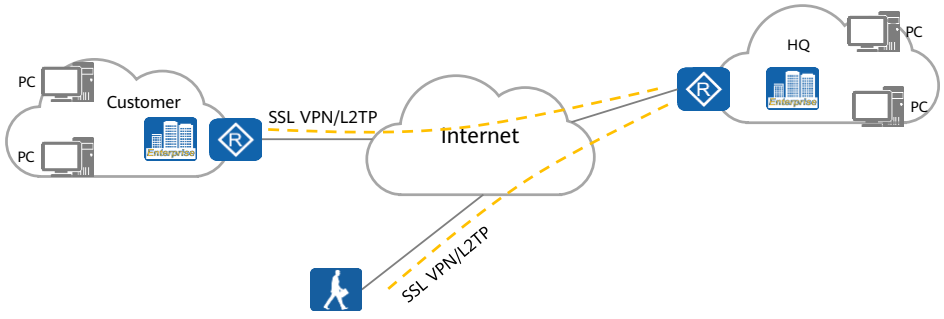
Application of GRE over IPsec



- 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 headquarters to facilitate management. 2. Enterprise O&M personnel just have a basic understanding of DSVPN, which is inconvenient for O&M.

Extranet VPN Overview

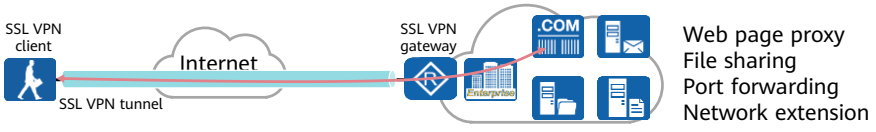
- Extranet VPN is mainly used to build secure access services between customers or suppliers. Traveling employees can also use extranet VPN to access the enterprise network.
- Extranet VPN mainly uses SSL VPN and L2TP.



- Extranet VPN mainly uses the following technologies:
  - SSL VPN
  - L2TP VPN

SSL VPN Overview

- SSL VPN is mainly used by 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.

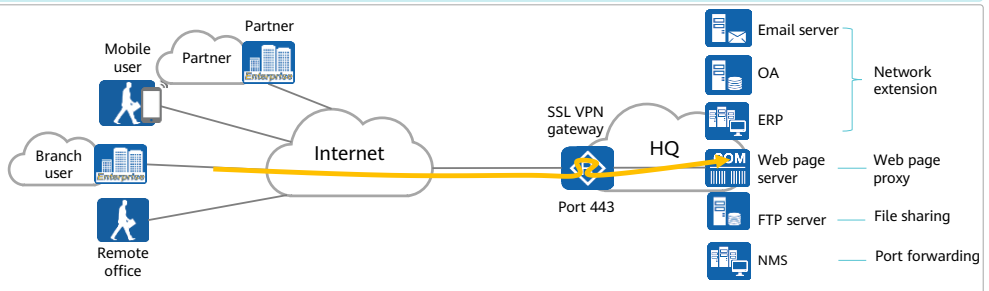


- SSL VPN is a VPN remote access technology based on SSL. Mobile users (referred to as remote users in SSL VPN) can use SSL VPN to securely and conveniently access enterprise intranets and intranet resources, improving work efficiency.
- Before SSL VPN is developed, VPN technologies such as IPsec and 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-grained control over the permissions on enterprise intranet resource access by remote users.

Application 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 installing a client. In addition, there are many SSL VPN service options, such as web page proxy, file sharing, port forwarding, and network extension.

Application of SSL VPN



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

Contents

SD-WAN Technology Review

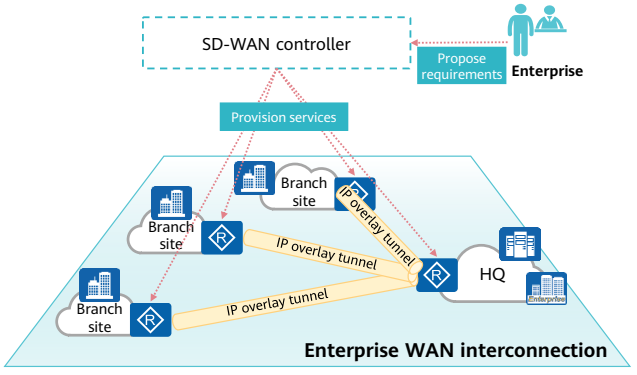
- Service- and intent-orientation, implementing network orchestration and automatic provisioning.

Traditional network

- Traditional network service provisioning requires professional network engineers to perform planning, configuration, and O&M, and then run commands or use the NMS software to configure devices one by one based on the planned services.

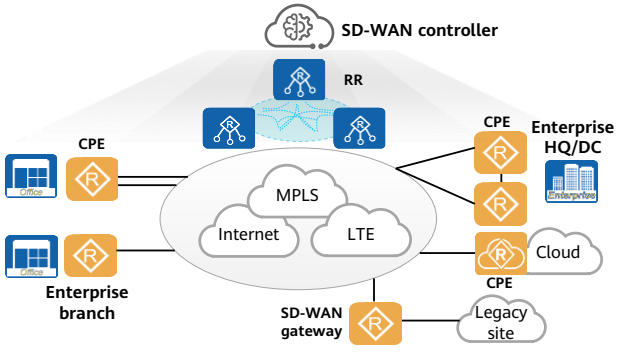
SDN

- The SDN network uses a centralized network control system to abstract, orchestrate, and automatically provision network services on demand. It shields technical implementation details of the network and opens only service-oriented interfaces and parameters to users.



HUAWEI

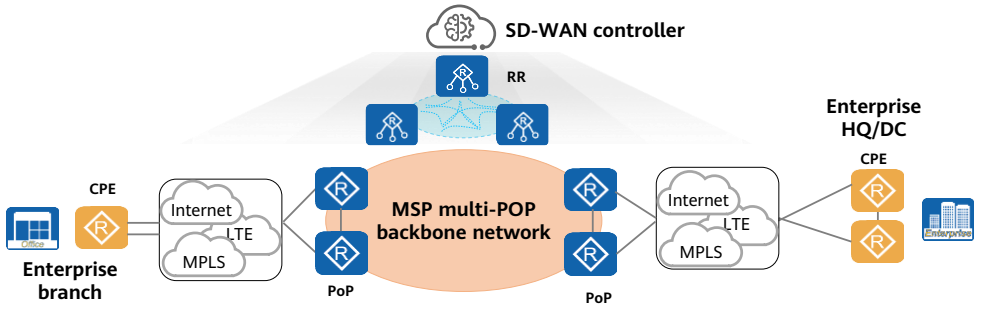
Enterprise-built SD-WAN Scenario



Device Model	Function
SD-WAN controller	Is the core of the SD-WAN solution. It uniformly manages CPEs, automatically delivers services, and implements unified control of overlay networks.
RR	Is a distributed control component, which distributes VPN routes between CPEs based on VPN topology policies.
CPE	Is an egress device of an enterprise branch, the headquarters, and a DC.
SD-WAN gateway	Is an intermediate gateway that connects an SD-WAN network to a non-SD-WAN network.

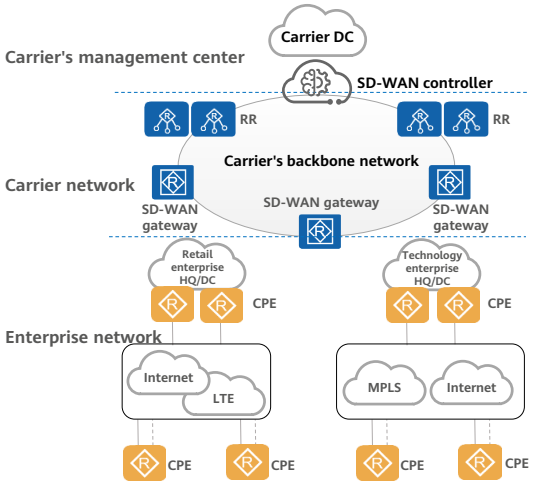
- The SD-WAN controller is deployed on the WAN to centrally manage CPEs and implement ZTP. This shortens the service provisioning time, and helps enterprises cope with challenges brought by cloud services and implement service change on demand.
- Large enterprises with a large number of branches, such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

MSP-built SD-WAN Scenario



- The management service provider (MSP) provides a unified SD-WAN controller to provide SD-WAN services for multiple enterprises. The MSP builds its own PoP backbone network and provides PoP access. Enterprises can access the nearest PoP, implementing high-quality enterprise interconnection. The PoP supports multiple tenants and provides access for multiple enterprises.
- Enterprises can serve as tenants to lease SD-WAN services provided by MSPs. An enterprise tenant can manage the SD-WAN services of all sites belonging to it. However, it cannot view the SD-WAN services of other tenants.

Carrier-built SD-WAN Scenario



- Carriers provide SD-WAN services for multiple enterprises through the SD-WAN controller.
- Enterprises can serve as tenants to lease SD-WAN services provided by carriers. An enterprise tenant can manage the SD-WAN services of all sites belonging to it. However, 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 entrust their SD-WAN services to carriers for management and control.
- SD-WAN gateways implement flexible interconnection and fast compatibility between the SD-WAN network and traditional carrier's backbone network.

Quiz

1. (Multiple-answer question) Which of the following are private line technologies?
   A. MPLS VPN
   B. SDH
   C. IPsec VPN
   D. L2TP

- 1. AB

Summary

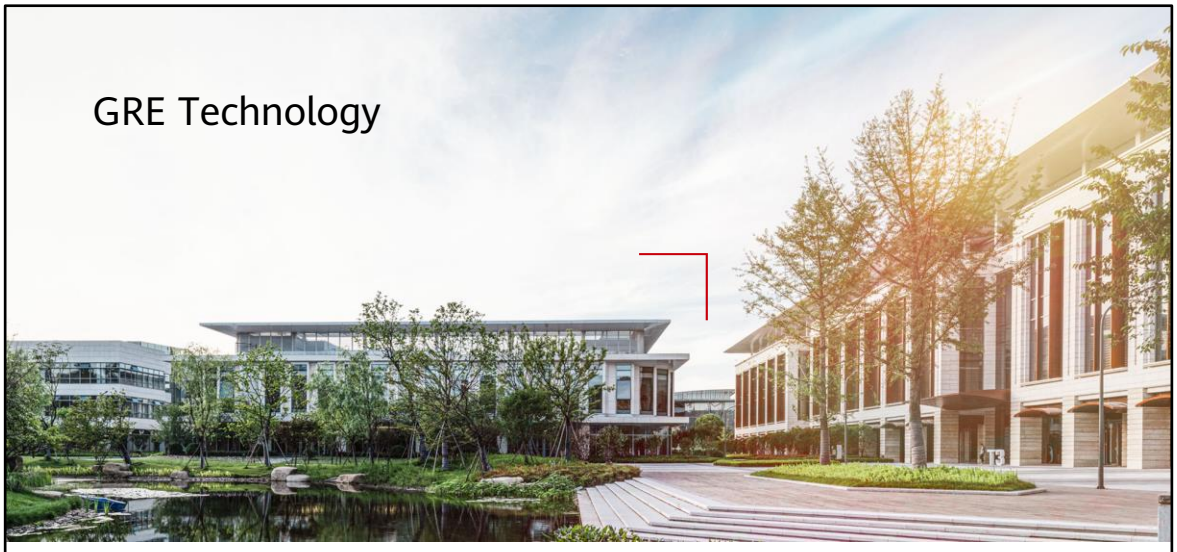
- Enterprise WANs can be connected to private lines or VPNs.
  - Private lines include bare fibers, transmission private lines (SDH and MSTP), and MPLS private lines (MPLS L2/L3 VPN).
  - VPNs include GRE, GRE over IPsec, L2TP VPN, and SSL VPN.
- The SD-WAN solution can build enterprise WANs based on private lines or VPNs.
- SD-WAN technology application scenarios are as follows:
  - Enterprise-built SD-WAN scenario
  - MSP SD-WAN Scenario
  - Carrier SD-WAN Scenario

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.

GRE Technology



HUAWEI

Foreword

- A large enterprise has a large number of branches. To enable branches to communicate with each other or with the headquarters, the private line or VPN technology needs to be used.
- The private line is expensive but has excellent performance; the VPN is cheaper than the private line but performance is lower.
- Generic Routing Encapsulation (GRE) is the most commonly used VPN technology on the live network. With GRE, an enterprise can build an intranet for the branches and headquarters at a very low cost.
- This section describes the basic concepts and fundamentals of GRE.

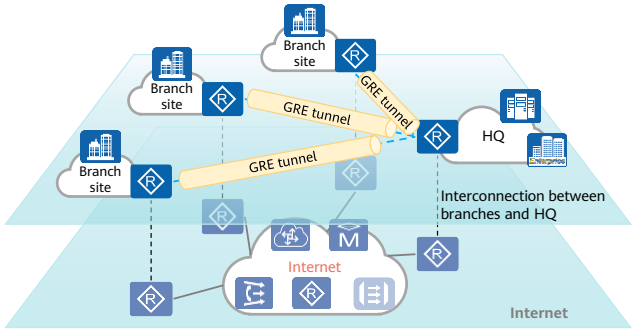
Objectives

- Upon completion of this course, you will be able to:
  - Describe basic concepts of tunnels.
  - Describe fundamentals of GRE.
  - Describe basic security mechanisms of GRE.
  - Describe application scenarios of GRE.
  - Complete basic GRE configuration.

Contents

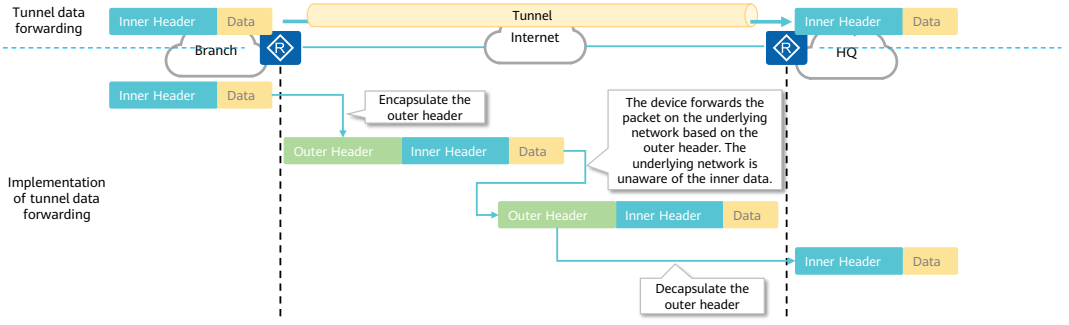
GRE Background

- With the development of enterprises, more and more enterprises need to communicate between branches and headquarters. Private lines (such as MPLS and SDH/MSTP private lines) need to be leased for communication between the headquarters and branches. However, private lines are expensive. For small- and medium-sized enterprises or cross-border companies, the cost is high.
- With the development of the Internet, the Internet has sufficient bandwidth and coverage. Therefore, it is more feasible to implement communication on the intranet between the headquarters and branches through the Internet. GRE is proposed in this background.
- Through GRE tunnels, the enterprise network can be established between the branch and headquarters based on the Internet.



Introduction to Tunneling Technologies

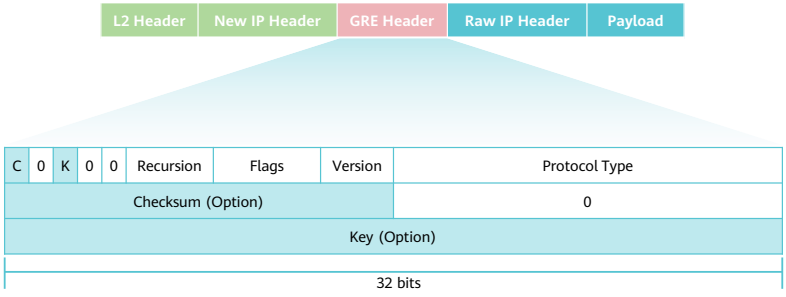
- GRE is one of tunneling technologies. A tunnel is similar to a bridge. Forwarding channels are established on the underlying network (for example, Internet). Users can establish a tunnel network by themselves without the intervention of the underlying network provider (for example, an ISP).
- There are many tunneling technologies, such as MPLS, GRE, Layer 2 Tunneling Protocol (L2TP), and Virtual Extensible LAN (VXLAN). The following figure shows the implementation of tunnel data forwarding.



HUAWEI

Basic Concepts of GRE

- As a Layer 3 tunneling technology, GRE encapsulates packets of a protocol into packets of another protocol to transparently transmit packets over GRE tunnels. This technology enables packet transmission between the HQ and branches.
- GRE tunnels can transmit IPv4/IPv6 unicast, multicast, and broadcast packets.
- GRE packet format:



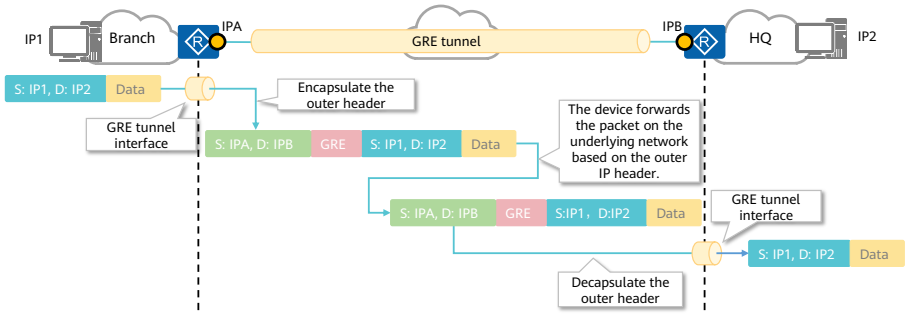
HUAWEI

- Description of fields in a GRE header:

Field	Description
C	Checksum verification bit. The value 1 indicates that the Checksum field is inserted into the GRE header. The value 0 indicates that the GRE header does not contain the checksum field.
K	Key bit. The value 1 indicates that the Key field is inserted into the GRE header. The value 0 indicates that the GRE header does not contain the keyword field.
Recursion	Number of layers where GRE packets are encapsulated. The value of this field is increased by 1 after one GRE encapsulation is complete. If the number of encapsulation layers is greater than 3, the packet is discarded. This field is used to prevent packets from being encapsulated continuously.
Flags	Reserved field. The value must be 0.
Version	Version. The value must be 0.
Protocol Type	Type of the passenger protocol. A common passenger protocol is the IPv4 protocol, with the value of 0800. The protocol number of Ethernet over GRE is 0x6558.
Checksum	Checksum of the GRE header and the payload.
Key	Key used to authenticate the packet at the receive end.

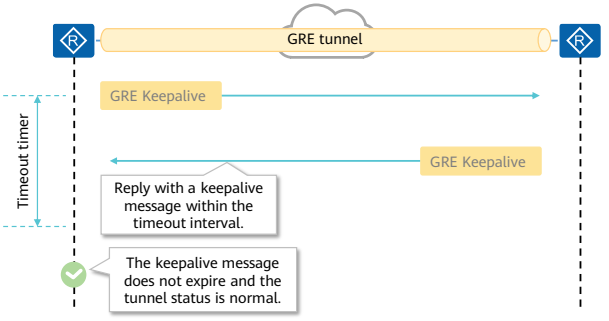
GRE Fundamentals

- The GRE tunnel is a Layer 3 tunnel and mainly carries IPv4/IPv6 packets. GRE encapsulates the outer IP header so that data can be transmitted on the public network. In this way, enterprise branches and the headquarters can communicate with each other.
- The following figure shows the process of forwarding packets over a GRE tunnel.



Keepalive Detection

- The current GRE protocol does not have the link status detection function. If the remote interface is unreachable, the GRE tunnel cannot be terminated immediately. As a result, the source continuously forwards packets to the peer. The peer, however, cannot receive packets because the tunnel is unreachable. In this case, traffic is interrupted.
- The keepalive detection function monitors tunnel status to check whether the remote end is reachable.
- Keepalive timeout interval = Sending interval (5s by default) x Retry count (3 by default)



HUAWEI

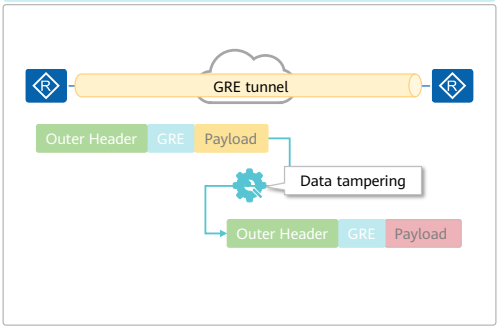
- Keepalive detection functions as follows:
  - After being enabled on the source end of a GRE tunnel, the source end starts a timer to periodically send and count keepalive messages. The number of sent keepalive messages increases by one each time a keepalive message is sent.
  - The destination end sends a response message to the source end each time it receives a keepalive message from the source end.
  - If the source end receives a reply packet before the counter value reaches the preset value, it considers the remote end reachable. If the source end does not receive any response message before the counter reaches the preset value, specifically, the retry count, the source end considers the peer end unreachable and resets the counter. Then, the source end terminates the tunnel connection. In this case, the source interface still sends Keepalive messages to the remote interface. When the remote interface becomes Up, the source interface becomes Up and sets up a tunnel with the remote interface.

Contents

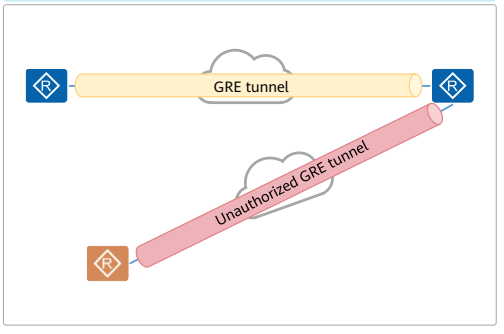
Security Threats to GRE Tunnels

- GRE tunnels are used to transmit data between branches and the HQ. Data is not encrypted and may be tampered with.
- There are potential risks in GRE tunnel establishment. Attackers can forge IP addresses to establish GRE tunnels between authorized and unauthorized devices.

GRE data tampering



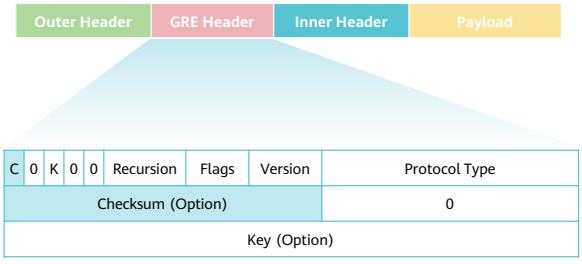
Unauthorized GRE tunnel setup

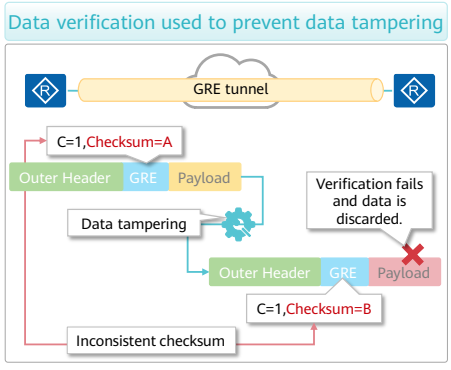


HUAWEI

GRE Data Check and Verification

- Checksum verification is an end-to-end check on encapsulated packets.
- If the C bit in the GRE header is set to 1, the checksum is valid. The sender calculates the checksum based on the GRE header and payload. Then it sends out the packet that carries the checksum. After receiving the packet, the receiver also calculates the checksum and compares the result with the checksum carried in the packet. If they are the same, the receiver further processes the packet. Otherwise, it discards the packet.



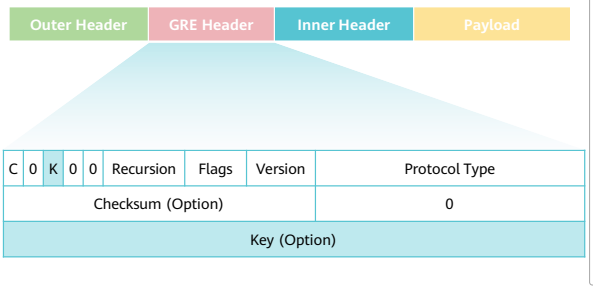


HUAWEI

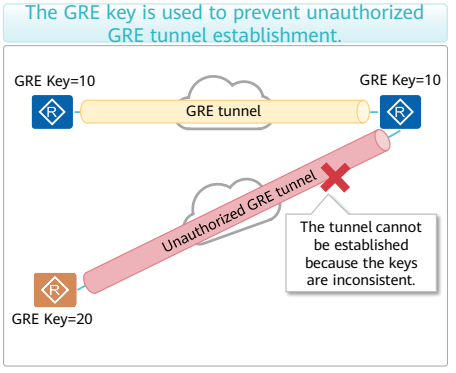
- You can enable or disable checksum verification on both ends of a tunnel in actual applications. If checksum verification is enabled on the local end and disabled on the remote end, the local end does not check checksum values of received packets, but checks checksum values of packets to be sent. If checksum verification is disabled on the local end and enabled on the remote end, the local end checks checksum values of received packets, but does not check checksum values of packets to be sent.

GRE Key

- Key authentication is used to verify validity of a tunnel interface. This security mechanism prevents tunnel interfaces on two devices at both ends of a GRE tunnel from incorrectly identifying and receiving packets from other devices.
- If the K bit in the GRE header is set to 1, a four-byte Key field is inserted into the GRE header. Both the receiver and the sender need to authenticate the key.



C	0	K	0	0	Recursion	Flags	Version	Protocol Type
Checksum (Option)								0
Key (Option)								



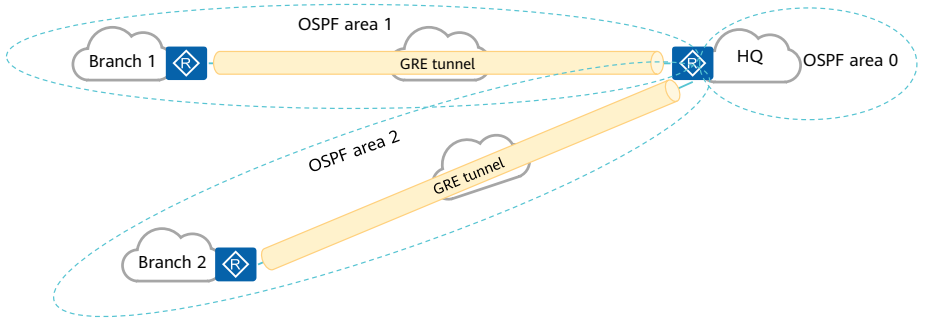
HUAWEI

- This field identifies traffic in a tunnel. Packets of the same traffic use the same key. During packet decapsulation, GRE identifies data packets of the same traffic based on the key. Packets will pass verification only when the two ends of the tunnel use the same Key field. If packets fail the verification, they will be discarded. Successful authentication requires that both ends are either configured with the same Key field or not configured with the Key field.

Contents

Using GRE to Build an Intranet Between the HQ and Branches

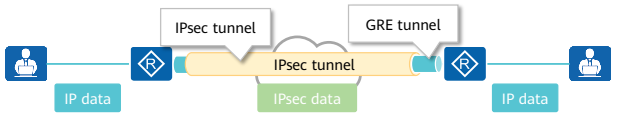
- GRE tunnels can transmit IPv4/IPv6 unicast, multicast, and broadcast packets. Dynamic routing neighbor relationships can be configured between branches and the HQ through GRE tunnels, facilitating intranet interconnection between branches and the HQ.



GRE Over IPsec

- GRE is simple. However, data is transmitted over a GRE tunnel in cleartext and can be easily obtained. On the live network, GRE is usually used together with IPsec. The GRE technology is used to establish the internal network connection between the branch and headquarters, and the IPsec technology is used to encrypt GRE tunnel packets.

GRE over IPsec data encapsulation



Contents

Basic GRE Configuration

- Context:
  - A small- or medium-sized enterprise has the headquarters (hub) and a branch (Spoke1) that are located in different areas. The enterprise requires that the hub and Spoke1 can communicate with each other.
- Configuration roadmap:
  - Ensure that the public network interfaces of the spoke and hub can communicate with each other.
  - Configure a GRE tunnel between the hub and spoke.

Creating a GRE Tunnel Between the Spoke and Hub

- The configuration on the spoke is similar to that on the hub. The configuration commands are as follows.

- The GRE tunnel needs to be configured on the devices at both ends of the tunnel.
- The configuration roadmap is as follows:
  - Ensure that the public network interfaces of the spoke and hub can communicate with each other.
  - Create GRE tunnels on the spokes and hub.
  - Configure the source and destination addresses of the GRE tunnel.

System-view

interface tunnel <interface-num> //Create a tunnel interface.

ip address <ip-address> //Set the IP address of the tunnel interface. This IP address is used as the next hop of the route for intranet communication.

tunnel-protocol gre //Set the tunnel type to GRE.

source <ip-address> //Set the source address of the tunnel, which is the same as the source IP address in the outer IP header of GRE-encapsulated packets.

destination <ip-address> //Set the destination address of the tunnel, which is the same as the destination IP address in the outer IP header of GRE-encapsulated packets.

gre key <key-num> //(Optional) Configure a GRE key that is used to check whether a GRE tunnel can be established.

Diverting Traffic to a GRE Tunnel

- There are many methods to divert traffic to a GRE tunnel, such as OSPF, static routes, and BGP.
- For details about how to configure static or dynamic routes, see the product documentation of Huawei.

Checking the GRE Configuration

- After the configuration is complete, run the following commands to check the configuration.

```
System-view
  display interface tunnel [interface-number]    //Check the working status of the tunnel interface.
  display tunnel-info tunnel-id [tunnel-id]      //Check tunnel information.
```

Quiz

1. (Multiple-answer question) How do we ensure GRE tunnel security?
   A. IPsec
   B. GRE data verification
   C. GRE key
   D. SSL

HUAWEI

- 1. ABC

Summary

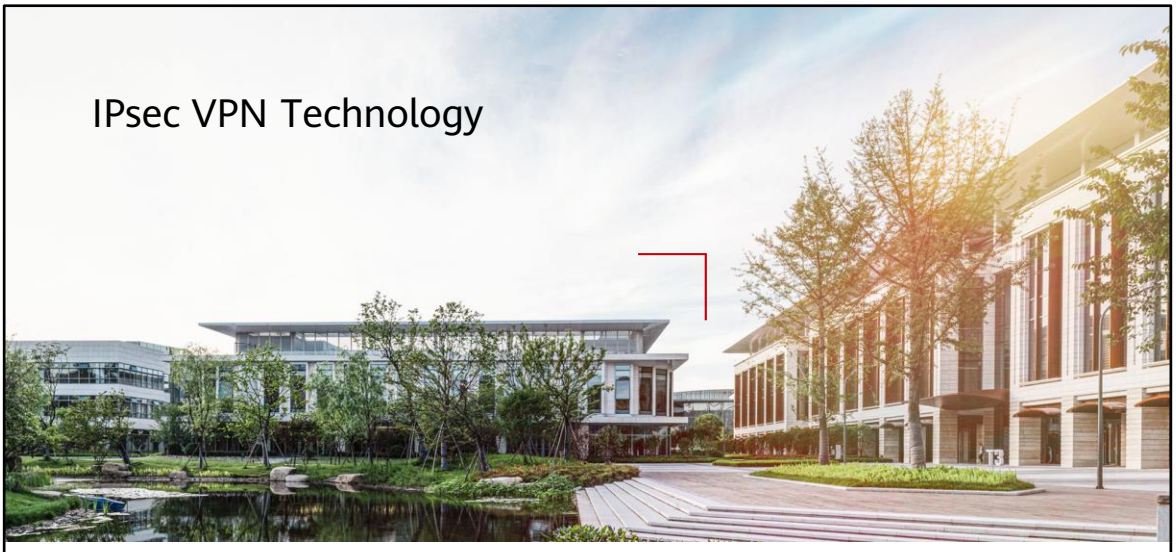
- GRE tunnels make it easy and cost-effective to establish intranet communication between enterprise branches and the HQ.
- GRE adds an outer IP header to IPv4/IPv6 data packets and builds a GRE tunnel, so that an ISP is not involved in data forwarding on the intranet.
- GRE tunnel security relies on the following technologies:
  - IPsec is used to encrypt GRE tunnel data.
  - GRE data verification is used to ensure that data in the GRE tunnel is not tampered with.
  - The GRE key is used to control setup of a GRE tunnel between sites.

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.

IPsec VPN Technology



HUAWEI

Foreword

- Most data is transmitted in clear text on the Internet, causing security risks. For example, bank accounts and passwords may be intercepted or tampered with, user information may be forged, and bank networks may be attacked. IP Security (IPsec) can address these problems by protecting the transmitted data.
- This course describes the fundamentals and application scenarios of IPsec.

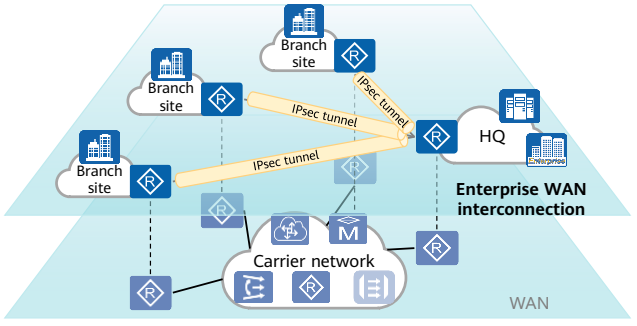
Objectives

- Upon completion of this course, you will be able to:
  - Describe the basic concepts of IPsec.
  - Understand the fundamentals of IPsec.
  - Summarize the application scenarios of IPsec.
  - Perform basic IPsec configurations.

Contents

IPsec Background

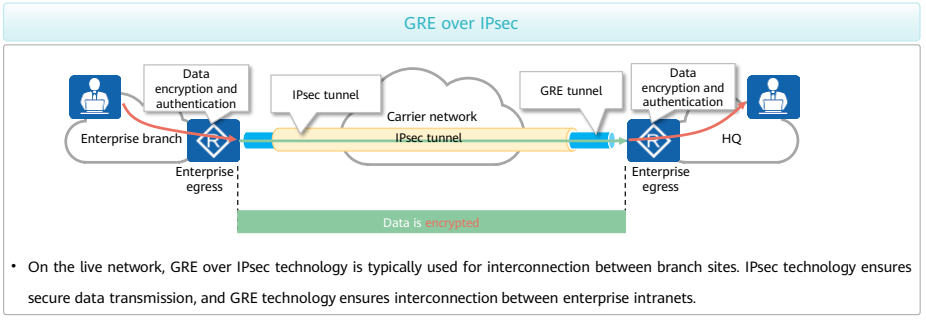
- Enterprise branches often need to communicate with each other. They can communicate using many methods, for example, using private lines or Internet links.
- Considering costs and requirements, some enterprises choose to use Internet links for interconnection. However, data may be intercepted when being transmitted on the Internet, posing security risks.
- IPsec technology encrypts data packets to secure enterprise interconnections.



HUAWEI

IPsec Overview

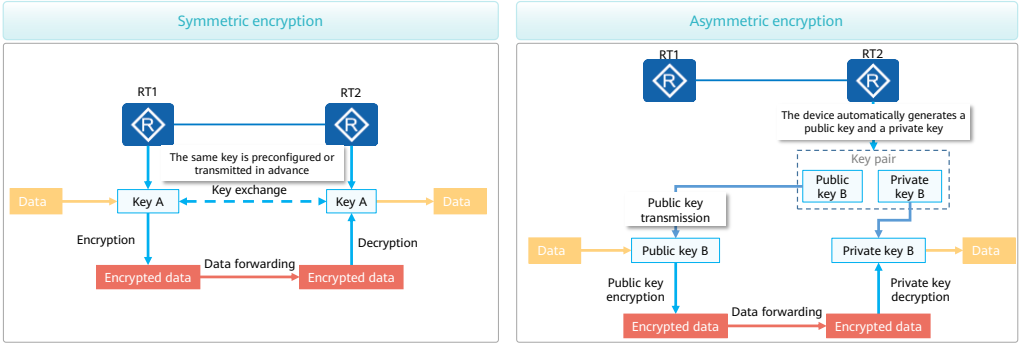
- The IPsec protocol suite is a series of security protocols developed by the Internet Engineering Task Force (IETF). It provides a cryptology-based, interoperable, and high-quality security protection mechanism for end-to-end IP packet exchange.
- IPsec encrypts and authenticates data to ensure secure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



- 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.

Data Encryption

- Data encryption prevents data from being leaked during data forwarding. Two data encryption methods are available:
  - Symmetric encryption: The same password is used for encryption and decryption, which is highly efficient. However, the key may be intercepted during key exchange.
  - Asymmetric encryption: The public key is used for encryption and the private key is used for decryption. Data security is high but the data encryption and decryption efficiency is low.

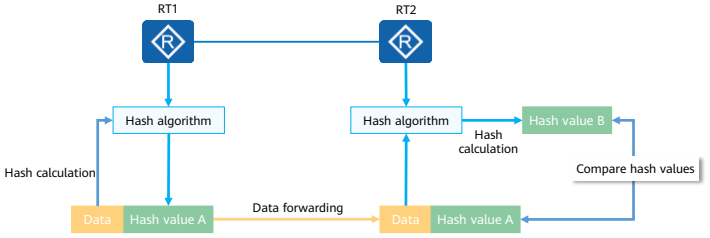


HUAWEI

- The symmetric encryption algorithm is also called traditional cryptographic algorithm, in which the encryption key can be calculated from the decryption key. The sender and receiver share the same key, which is used for both encryption and decryption. Symmetric key encryption is an effective method for encrypting a large amount of data. There are many algorithms for symmetric key encryption, and all of them aim to convert between cleartext (unencrypted data) and ciphertext. Because symmetric key encryption uses the same key for data encryption and decryption, data security depends on whether unauthorized users obtain the symmetric key. If two communicating parties want to use the symmetric key to encrypt data, they must exchange the key securely before exchanging the encrypted data.
- An asymmetric algorithm is also called public key algorithm, in which a public key is used for encryption and a private key for decryption. The two keys are mathematically related. In public key encryption, the public key can be publicly transmitted between two communicating parties or released in the public repository, but the private key is confidential. The data encrypted using the public key can be decrypted only using the private key. The data encrypted using the private key can be decrypted only using the public key.

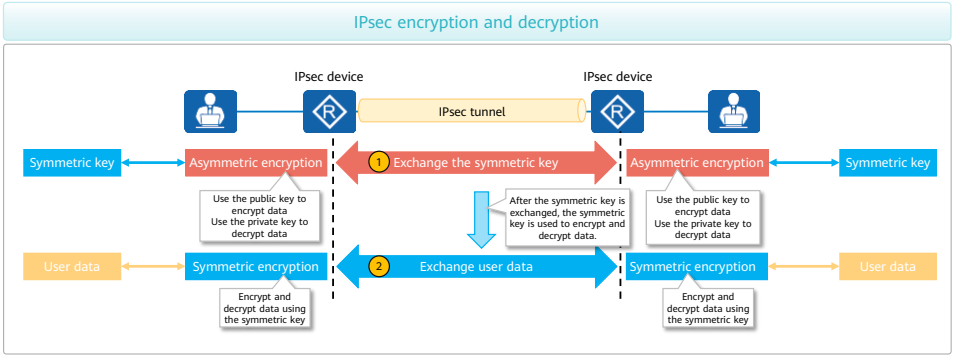
Data Authentication

- The main purpose of data authentication is to check whether data is tampered with. Data authentication is mainly based on the hash algorithm.
  - A unique hash value is calculated based on the hash algorithm and then carried in the data before being forwarded to the peer device.
  - The peer device hashes the data again to obtain the hash value. It then compares the received hash value with the calculated one. If they are the same, the data is not tampered with.



IPsec Encryption

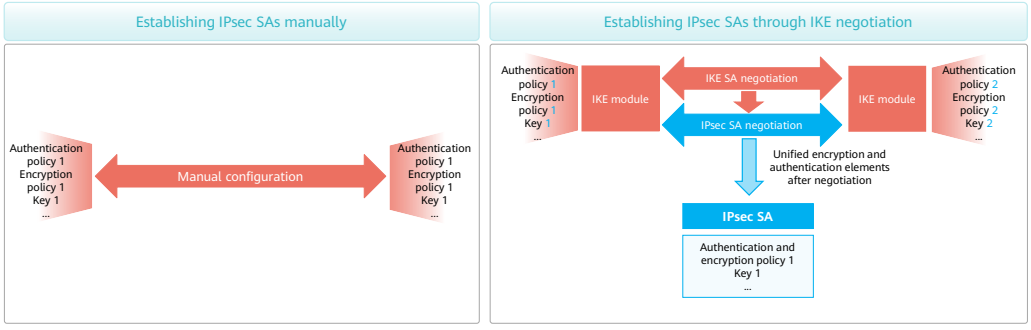
- IPsec uses both symmetric encryption and asymmetric encryption, ensuring data security and performance.
  - Uses an asymmetric algorithm to encrypt and transmit the key used for symmetric encryption.
  - Uses the exchanged symmetric key to encrypt data.



Contents

SA

- A security association (SA) is an agreement between two IPsec peers on certain elements. For example, Data Encryption Standard (DES) is used as the encryption algorithm, Message Digest Algorithm 5 (MD5) is used as the authentication algorithm, and tunnel is used as the encapsulation mode.
- An IPsec SA can be established manually or through Internet Key Exchange (IKE) negotiation.

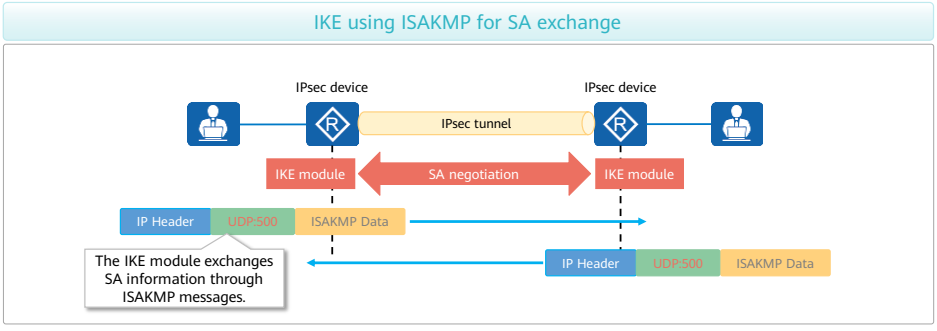


HUAWEI

- IPsec technology supports multiple data encryption, authentication, and encapsulation algorithms. When devices at both ends use IPsec for secure communication, they must use the same encryption and authentication algorithms. Therefore, a mechanism is required to help the devices negotiate these parameters.
- An IPsec SA can be established in either of the following ways:
  - Manual configuration: The management cost of manually established IPsec SAs is high. This is because the encryption and authentication modes need to be manually configured, SAs need to be manually updated, and SA information permanently exists, resulting in low security. This mode applies to small-scale networks.
  - IKE negotiation: The management cost of IPsec SAs established through IKE negotiation is low. The encryption and authentication modes are generated using the Diffie-Hellman (DH) algorithm, SA information is generated periodically, and SAs are dynamically updated. This mode applies to small-, medium-, and large-sized networks.
- An SA is uniquely identified by three parameters: security parameter index (SPI), destination IP address, and security protocol ID (AH or ESP).
- An IKE SA is used to establish a secure channel for exchanging IPsec SAs.

Key Exchange

- On the live network, the Internet Key Exchange (IKE) protocol is typically used to exchange symmetric keys.
- IKE is a UDP-based application-layer protocol. It is built upon the framework defined by the Internet Security Association and Key Management Protocol (ISAKMP). IPsec uses IKE for key auto-negotiation and IPsec SA establishment, simplifying IPsec configuration and maintenance.

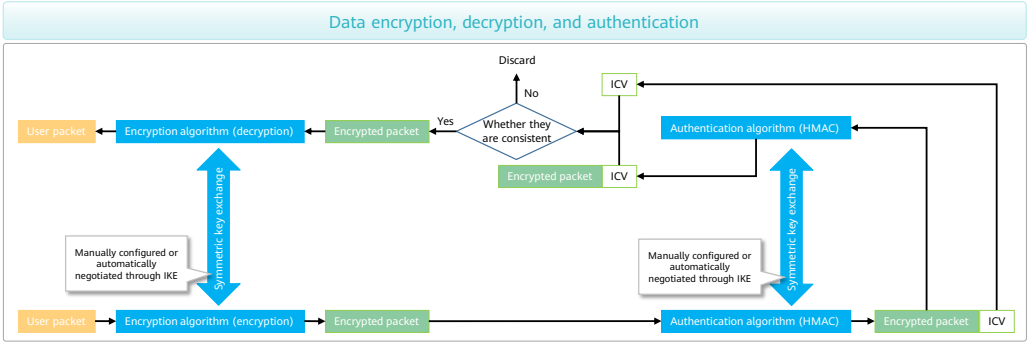


HUAWEI

- IKE supports the following authentication algorithms including MD5, Secure Hash Algorithm 1 (SHA1), SHA2-256, SHA2-384, SHA2-512, and Senior Middle 3 (SM3).
- IKE supports the following encryption algorithms: DES, 3DES, AES-128, AES-192, AES-256, SM1, and SM4.
- ISAKMP is defined in RFC 2408, which defines the procedures for negotiating, establishing, modifying, and deleting SAs and defines the ISAKMP message format. ISAKMP provides a general framework for SA attributes and the methods of negotiating, modifying, and deleting SAs, without defining the specific SA format.
- ISAKMP messages can be transmitted using UDP or TCP through port 500. In most cases, ISAKMP messages are transmitted using UDP.

Data Encryption and Authentication

- IPsec provides two security mechanisms: authentication and encryption.
  - IPsec uses symmetric encryption algorithms to encrypt and decrypt data. These algorithms require that the sender and receiver use the same key (a symmetric key) to encrypt and decrypt data.
  - IPsec uses the Hash-based Message Authentication Code (HMAC) function to compare digital signatures to check data integrity and authenticity.

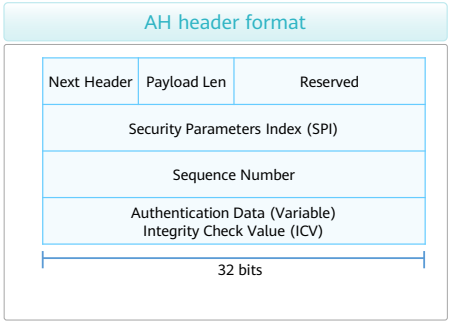


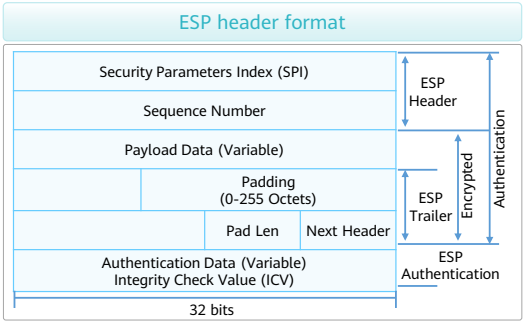
HUAWEI

- Integrity check value (ICV) is used by the receiver for integrity check. Available authentication algorithms are MD5, SHA1, SHA2, and SM3.
- Common symmetric encryption algorithms used by IPsec include Data Encryption Standard (DES), Triple Data Encryption Standard (3DES), Advanced Encryption Standard (AES), and algorithms approved by State Cryptography Administration, such as SM1 and SM4. DES and 3DES are not recommended because they are insecure and pose security risks.
- Common authentication algorithms used by IPsec include MD5, SHA1, SHA2, and SM3. MD5 and SHA1 are not recommended because they are insecure and pose security risks.
- IPsec encryption cannot verify the authenticity or integrity of information after decryption. IPsec uses the HMAC function to compare digital signatures to check integrity and authenticity of data packets. In most cases, encryption and authentication are used together. The IPsec sender uses the authentication algorithm and symmetric key to generate a digital signature for the encrypted packet and sends the IP packet and digital signature to the receiver. The receiver uses the same authentication algorithm and symmetric key to process the encrypted packet and then generates a digital signature. Then the receiver compares the received and generated digital signatures to verify the data integrity and authenticity. If the packet passes the verification, the receiver decrypts it. Otherwise, the receiver discards it.

Security Protocols

- IPsec provides two transport layer protocols for authentication or encryption: Authentication Header (AH) and Encapsulating Security Payload (ESP).
  - AH provides only authentication but no encryption capabilities.
  - ESP provides both authentication and encryption.

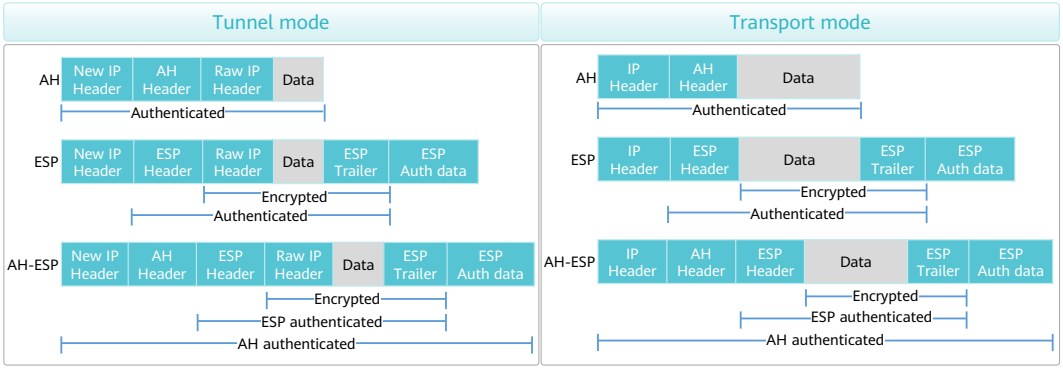




- AH provides only authentication but no encryption capabilities. According to the AH protocol, an AH header is appended to the standard IP header in each packet. The sender performs hash calculation on packets and an authentication key. After packets carrying the calculation result arrive at the receiver, the receiver also performs hash calculation and compares the calculation result with the received calculation result. Any changes to the data during transmission will make the calculation result invalid. This implements data origin authentication and integrity verification. AH provides data integrity check on an entire IP packet.
- ESP provides both authentication and encryption. An ESP header is appended to the standard IP header in each data packet, and the ESP Trailer and ESP Auth data fields are appended to each data packet. In contrast to AH, ESP encrypts the payload before encapsulating it into a data packet to ensure data confidentiality, and protects the IP header only in tunnel mode.
- Key fields:
  - Sequence Number: This field is a counter that monotonically increases from 1. It uniquely identifies a packet to prevent replay attacks.
  - SPI: This field uniquely identifies an IPsec SA.
  - Authentication Data: This field contains the Integrity Check Value (ICV) and is used by a receiver for data integrity check. Available authentication algorithms are MD5, SHA1, SHA2, and SM3.

Encapsulation Modes

- IPsec encapsulation is a process of adding AH or ESP fields to original IP packets for packet authentication and encryption. This process is implemented in transport or tunnel mode.
- On the live network, the tunnel mode is often used for encapsulation.

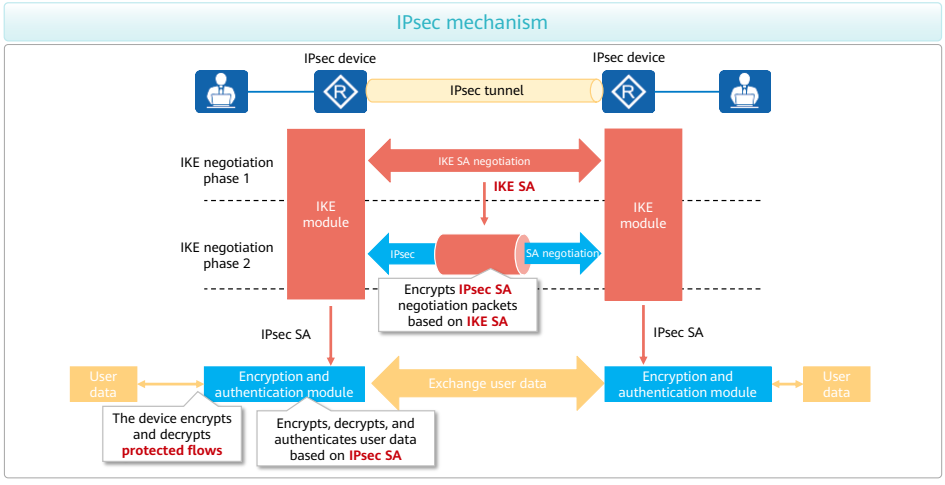


HUAWEI

- In transport mode, an AH or ESP header is added between an IP header and a transport-layer protocol (TCP, UDP, or ICMP) header to protect the TCP, UDP, or ICMP payload. As no additional IP header is added, IP addresses in the original packets are visible in the IP header of the post-encrypted packet.
- In tunnel mode, an AH or ESP header is added before the raw IP header and then encapsulated into a new IP packet with a new IP header to protect the IP header and payload.

Contents

IPsec Mechanism

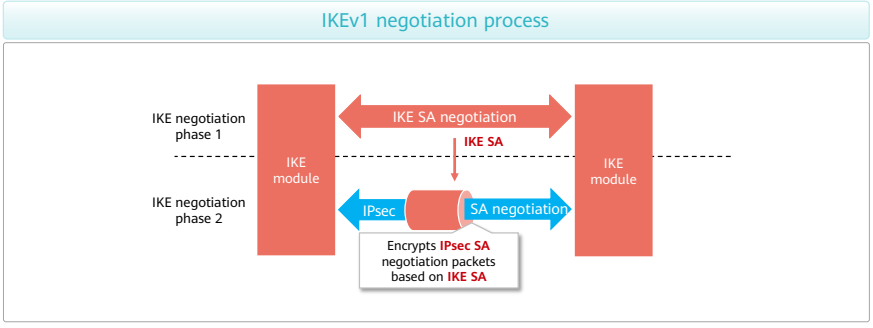


HUAWEI

- The IPsec mechanism is as follows:
  - An IKE SA is negotiated in the first phase of IKE negotiation.
  - The IKE SA is used to encrypt the packets in the second phase of IKE negotiation. That is, IPsec SAs are negotiated in the second phase of IKE negotiation.
  - IPsec SAs are used to encrypt data.

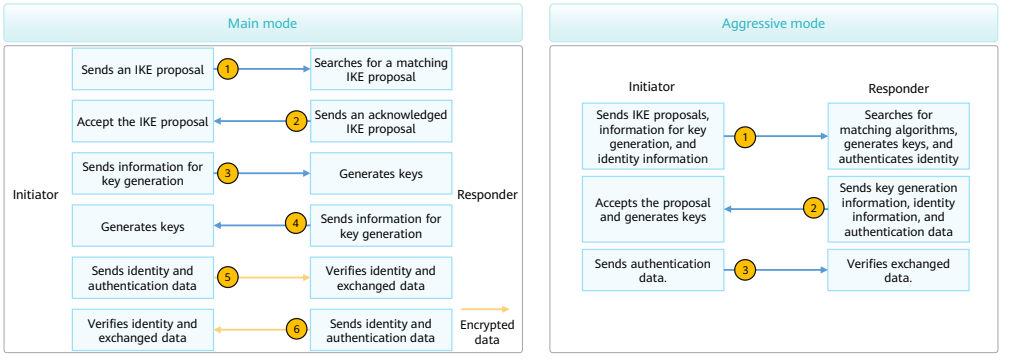
IKEv1

- IKEv1 negotiation goes through two phases: In phase 1, two IPsec peers negotiate and establish a secure tunnel (an IKE SA). In phase 2, the two IPsec peers establish a pair of IPsec SAs for secure data transmission through the secure tunnel established in phase 1.



IKEv1 Negotiation Phase (1)

- In phase 1 of IKEv1 negotiation, an IKE SA is established. After an IKE SA is established, all the ISAKMP messages transmitted between two IPsec peers will be encrypted and authenticated. The secure tunnel established in phase 1 enables IPsec peers to communicate securely in phase 2.
- Phase 1 of IKEv1 negotiation supports two negotiation modes: main mode and aggressive mode.

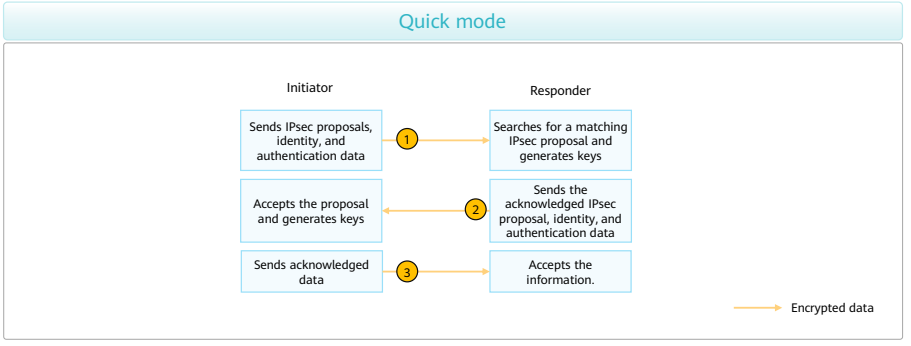


HUAWEI

- The main mode requires three exchanges between the peers, totaling six ISAKMP messages. The three exchanges are described as follows:
  - Messages 1 and 2 are used for IKE proposal exchange.
    - The initiator sends one or more IKE proposals to the responder. The responder searches for the first matching IKE proposal and then sends it to the initiator. IKE proposals of the initiator and responder match if they have the same encryption algorithm, authentication algorithm, authentication method, and DH group identifier.
  - Messages 3 and 4 are used for key information exchange.
    - The initiator and responder exchange the DH public value and nonce value to generate the IKE SA authentication key and encryption key.
  - Messages 5 and 6 are used for identity and authentication information exchange. (Both parties use the generated keys to exchange information.)
    - The initiator and responder use the generated keys to authenticate each other and the information exchanged in main mode.
- The aggressive mode uses only three messages. Messages 1 and 2 are used to negotiate IKE proposals and exchange the DH public value, mandatory auxiliary information, and identity information. Message 2 also contains the identity information sent by the responder to the initiator for authentication. Message 3 is used by the responder to authenticate the initiator.
- Compared with the main mode, the aggressive mode reduces the number of exchanged messages and speeds up the negotiation. However, the aggressive mode does not encrypt identity information.

IKEv1 Negotiation Phase (2)

- In IKEv1 phase 2, IPsec SAs need to be established and keys needs to be generated for securely transmitting data.
- This phase uses the quick mode. This mode uses the keys generated in phase 1 to verify the integrity of ISAKMP messages and identities of the initiator and responder, and to encrypt ISAKMP messages, ensuring exchange security.



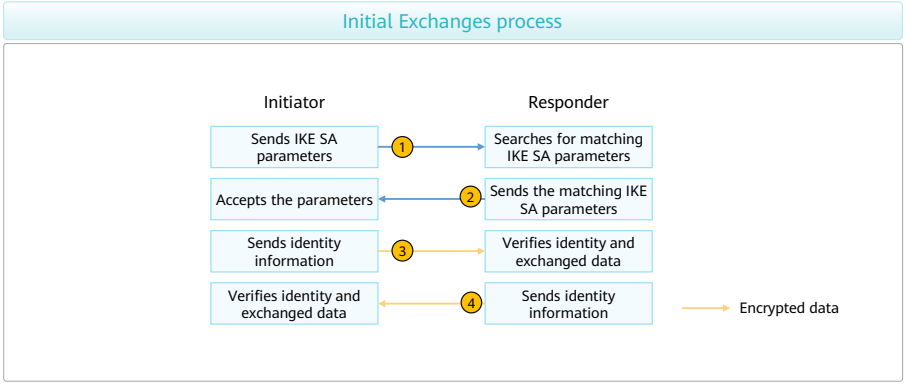
- In IKEv1 phase 2, two IPsec SAs are established through three ISAKMP messages:
  - Message 1 is used by the initiator to send local security parameters and identity authentication information to the responder.
    - Security parameters include protected data flows and parameters to be negotiated, such as an IPsec proposal. Identity authentication information includes the keys generated in phase 1 and keying materials generated in phase 2, and can be used to authenticate the peer again.
  - Message 2 is used by the responder to send acknowledged security parameters and identity authentication information, and to generate new keys.
    - The encryption key and authentication key used for secure data transmission over IPsec SAs are generated based on the keys generated in phase 1 and parameters such as the SPI and protocol. This ensures that each IPsec SA has unique encryption and authentication keys.
  - Message 3 is used by the initiator to send acknowledged information to communicate with the responder. IKEv1 negotiation then ends and IPsec SAs are established.

IKEv2

- The process of establishing SAs through IKEv2 negotiation is much simpler than that through IKEv1 negotiation. In normal cases, IKEv2 can establish a pair of IPsec SAs through only four messages in two exchanges. One additional Create_Child_SA Exchange can be used to establish another pair of IPsec SAs if required, during which only two messages are exchanged.
- IKEv2 defines three exchanges: Initial Exchanges, Create_Child_SA Exchange, and Informational Exchange.

IKEv2 Initial Exchanges

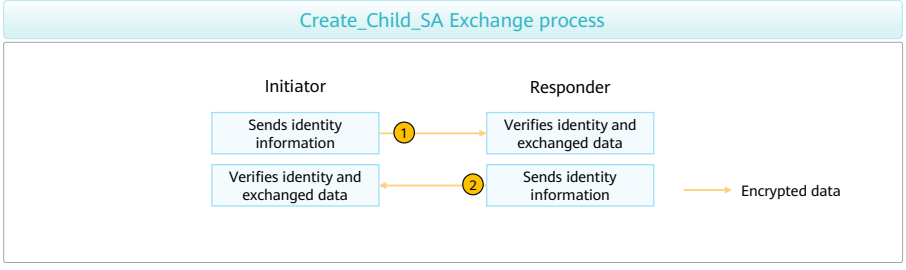
- IKEv2 establishes the first pair of IPsec SAs through Initial Exchanges. Initial Exchanges involves four messages in two exchanges.



- Messages 1 and 2 are used in exchange 1 (called IKE_SA_INIT). In exchange 1, IKE SA parameters are negotiated in plain text, including the encryption key, authentication key, random number, and DH key. After IKE_SA_INIT is complete, shared keying material is generated, from which all keys used by IPsec SAs are derived.
- Messages 3 and 4 are used in exchange 2 (called IKE_AUTH). In exchange 2, identities of the two parties and the first two messages are authenticated, and IPsec SA parameters are negotiated. IKEv2 supports Rivest-Shamir-Adleman (RSA) signature authentication, pre-shared key (PSK) authentication, and Extensible Authentication Protocol (EAP) authentication. The initiator omits the AUTH payload in message 3 to indicate that EAP authentication is required.

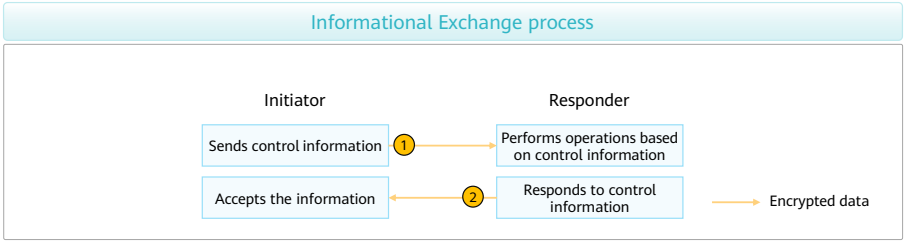
IKEv2 Create_Child_SA Exchange

- After one pair of IPsec SAs is established based on an IKE SA, Create_Child_SA Exchange can be performed to negotiate more pairs of IPsec SAs. In addition, Create_Child_SA Exchange can be performed for IKE SA re-negotiation.
- Create_Child_SA Exchange involves two messages in one exchange and corresponds to IKEv1 phase 2. The initiator in Create_Child_SA Exchange can be the initiator or responder in Initial Exchanges.



IKEv2 Informational Exchange

- IKEv2 peers perform Informational Exchange to exchange control information, including error information and notifications.
- Informational Exchange must be performed under the protection of an IKE SA. Specifically, Informational Exchange is performed after Initial Exchanges are complete. Control information may belong to an IKE SA or a child SA. Therefore, Informational Exchange must be protected by the IKE SA or the IKE SA based on which the child SA is established accordingly.



Defining IPsec-Protected Data Flows

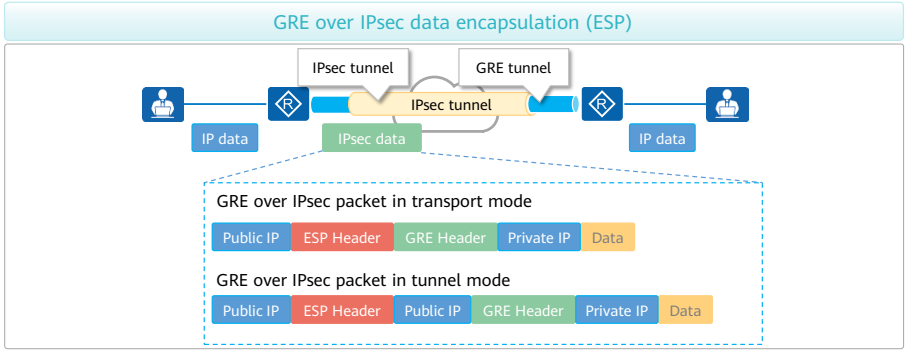
- The data flows to be protected by IPsec can be defined using either of the following methods:
  - Use ACLs.
    - ACLs can be configured to define the data flows to be protected by an IPsec tunnel. The packets matching permit clauses in the ACLs will be protected.
  - Use routes.
    - Routes can be configured to define the data flows to be protected by an IPsec tunnel established through IPsec tunnel interfaces. All packets routed to these interfaces will then be protected.
- On the live network, GRE over IPsec typically defines protected flows based on routes.

- The method of using routes has the following advantages:
  - Simplifies the IPsec configuration: IPsec-protected data flows are routed to tunnel interfaces, without the need to use ACLs to define the characteristics of traffic to be encrypted or decrypted.
  - Supports dynamic routing protocols.
  - Protects multicast traffic through GRE over IPsec.

Contents

GRE over IPsec

- Leveraging advantages of GRE and IPsec, GRE over IPsec encapsulates multicast, broadcast, and non-IP packets into ordinary IP packets and then securely transmits these IP packets through IPsec.
- GRE over IPsec encapsulates packets using GRE and then IPsec.

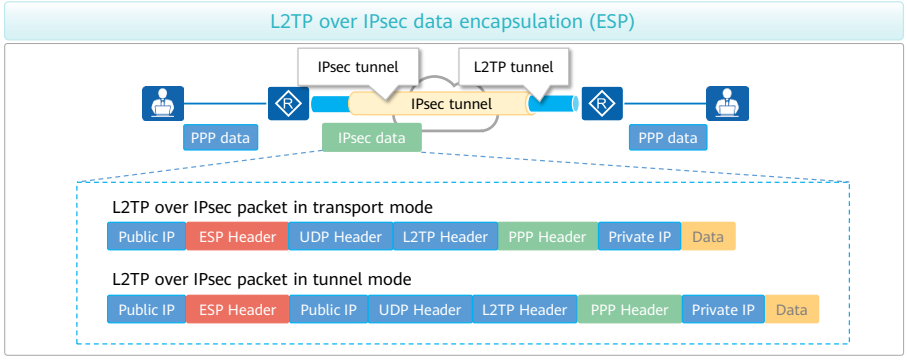


HUAWEI

- GRE over IPsec supports encapsulation in both tunnel and transport modes. An IPsec header needs to be added to packets if GRE over IPsec in tunnel mode is used, resulting in longer packets. In this case, packets are more likely to be fragmented. Therefore, GRE over IPsec in transport mode is recommended.
- In the IP header added during IPsec encapsulation, the source and destination addresses are the IP addresses of the local interface and remote interface to which an IPsec policy is applied.
- IPsec protects data flows from the GRE tunnel source to the GRE tunnel destination. In the IP header added during GRE encapsulation, the source and destination addresses are the source and destination addresses of a GRE tunnel.

L2TP over IPsec

- Layer 2 Tunneling Protocol (L2TP) over IPsec encapsulates packets using L2TP and then IPsec. It uses L2TP for user authentication and address allocation and uses IPsec for secure communication. L2TP over IPsec ensures that branches or traveling employees are securely connected to the headquarters.

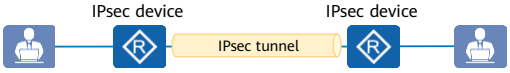


HUAWEI

- L2TP encapsulation and then IPsec encapsulation are performed on packets transmitted over an L2TP over IPsec tunnel. In the IP header added during IPsec encapsulation, the source and destination addresses are the IP addresses of the local interface and remote interface to which an IPsec policy is applied.
- IPsec needs to protect the data flows from the L2TP tunnel source to the L2TP tunnel destination. In the IP header added to packets during L2TP encapsulation, the source and destination addresses are the source and destination addresses of an L2TP tunnel. When a branch connects to the headquarters, the source address of the L2TP tunnel is the IP address of the outbound interface on the L2TP access concentrator (LAC), and the destination address is the IP address of the inbound interface on the L2TP network server (LNS).
- A public IP header is added to packets during L2TP encapsulation, and another public IP header is added to packets if L2TP over IPsec in tunnel mode is used, resulting in longer packets, which are prone to being fragmented. Therefore, L2TP over IPsec in transport mode is recommended.
- The L2TP over IPsec negotiation process and packet encapsulation process are similar when traveling employees are remotely connected to the headquarters and when branch employees are connected to the headquarters. The difference is that, L2TP and IPsec encapsulation is performed on clients when traveling employees are remotely connected to the headquarters. The L2TP tunnel source address is the private address assigned to a client and can be any address in the IP address pool configured on the LNS. The L2TP tunnel destination address is the address of the inbound interface on the LNS.

Contents

Configuring IKE



- The configurations on both ends of an IPsec tunnel are similar. The IPsec configuration roadmap is as follows:
  - Configure an IKE proposal.
  - Configure an IKE peer.
  - Define IPsec-protected data flows. In most cases, ACLs are used to define such data flows.
  - Configure an IPsec proposal, and set the encryption and authentication algorithms.
  - Configure an IPsec policy using ISAKMP or an IPsec policy template.

- Configure an IKE proposal.

```
system-view
  ike proposal [proposal-number]  //Create an IKE proposal.
    authentication-method [pre-share | rsa-signature | digital-envelope]
//Configure the IKE authentication method. By default, PSK authentication is used.
    authentication-algorithm [algorithm]    //Configure the authentication
algorithm used by IKEv1. By default, SHA2-256 is used.
    encryption-algorithm [algorithm]    //Configure the IKE encryption algorithm.
By default, AES-256 is used.
```

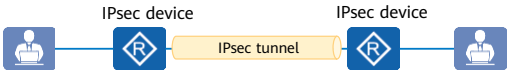
- Configure an IKE peer.

```
system-view
  ike peer [peer-name]    //Create an IKE peer.
    ike-proposal [proposal-number]    //Apply an IKE proposal.
    pre-shared-key cipher [key]  //Configure a PSK for IKE negotiation.
    remote-address [ip-address]  //Configure the remote IP address for IKE
negotiation.
```


HUAWEI

- IKE proposals and IKE peers are used to negotiate IKE SAs in the first phase of IKE negotiation.

Configuring an IPsec Proposal



- ACLs are typically used to define IPsec-protected traffic. This course describes the IPsec configuration. The ACL configuration is not described here.
- An IPsec proposal, as part of an IPsec policy or an IPsec profile, defines security parameters for IPsec SA negotiation, including the security protocol, encryption and authentication algorithms, and data encapsulation mode.

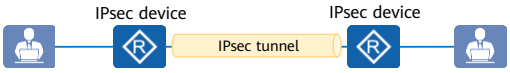
- Configure an IPsec proposal:

system-view
IPsec proposal [proposal-name] //Create an IPsec proposal.
transform [ah | esp | ah-esp] //Configure the security protocol used by IPsec.
esp authentication-algorithm [algorithm] //Configure the authentication algorithm for ESP.
esp encryption-algorithm [algorithm] //Configure the encryption algorithm for ESP.
ah authentication-algorithm [algorithm] //Configure the authentication algorithm for AH. On the live network, ESP is typically used.
encapsulation-mode [transport | tunnel] //Configure the encapsulation mode for IPsec data packets.

HUAWEI

- An IPsec proposal is used to negotiate IPsec SAs in the second phase of IKE negotiation.

Configuring an ISAKMP IPsec Policy



- An ISAKMP IPsec policy applies when the remote IP address is fixed.
- Negotiated IPsec parameters of an ISAKMP IPsec policy are defined in the IPsec policy view, and the negotiation initiator and responder must use the same IPsec parameters.

- Configure an IPsec policy:

```
system-view
  IPsec policy [policy-name] [seq-number] isakmp    //Create an ISAKMP IPsec policy.
    security acl [acl-number]    //Reference an ACL in the IPsec policy.
    proposal [proposal-name]    //Reference the IPsec proposal in the IPsec policy.
    ike-peer [peer-name]    //Reference the IKE peer in the IPsec policy.
```

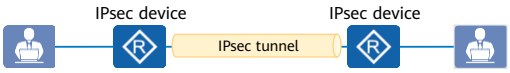
- Apply the IPsec policy to an interface:

```
system-view
  interface [interface-type interface-num]    //Enter the interface view. An IPsec policy can be applied to a common interface, sub-interface, or tunnel interface.
    IPsec policy [policy-name]    //Apply an IPsec policy to the interface.
```


HUAWEI

- After an IPsec policy group to which an IPsec policy belongs is applied to an interface, the following situations occur:
  - To modify the parameters of an IPsec proposal, unbind the IPsec policy group from the interface and then apply the IPsec policy group to the interface again.
  - If other parameters are modified, these parameters take effect during next IKE negotiation and are invalid for the IPsec tunnels that have been established through negotiation.

Configuring a Template IPsec Policy



- An IPsec policy template simplifies the configuration workload for establishing multiple IPsec tunnels. It applies to scenarios where the peer IP address is not fixed or multiple peer ends exist.
- When an IPsec tunnel is established using a template IPsec policy, the initiator determines optional parameters, and the responder accepts the parameters delivered by the initiator.

- Configure an IPsec policy:

```
system-view
  IPsec policy-template [template-name] [seq-number]   //Creating an IPsec policy template.
    security acl [acl-number]    //Reference an ACL in the IPsec policy.
    proposal [proposal-name]    //Reference an IPsec proposal in the IPsec policy.
    ike-peer [peer-name]    //Reference an IKE peer in the IPsec policy.
  IPsec policy [policy-name] [seq-number] isakmp template [template-name]
//Reference the policy template in the IPsec policy.
```

- Apply the IPsec policy to an interface:

```
system-view
  interface [interface-type interface-num]    //Enter the interface view. An IPsec policy can be applied to a common interface, sub-interface, or tunnel interface.
    IPsec policy [policy-name]    //Apply an IPsec policy to the interface.
```


HUAWEI

- Note the following points when configuring a template IPsec policy:
  - If one end (responder) of an IPsec tunnel has a template IPsec policy configured, the other end (initiator) must have an ISAKMP IPsec policy configured.
  - In an IPsec policy template, an IPsec proposal and IKE peer must be configured, while other parameters are optional. The initiator determines optional parameters in the IPsec policy template, and the responder accepts the parameters delivered by the initiator.

Quiz

1. (Multiple-answer question) Which of the following modes are supported in IKEv1 phase 1?
   A. Passive mode
   B. Aggressive mode
   C. Main mode
   D. Backup mode
2. (Multiple-answer question) What are the two IPsec data encapsulation modes?
   A. ESP mode
   B. AH mode
   C. Tunnel mode
   D. Transport mode

- 1. BC
- 2. CD

Summary

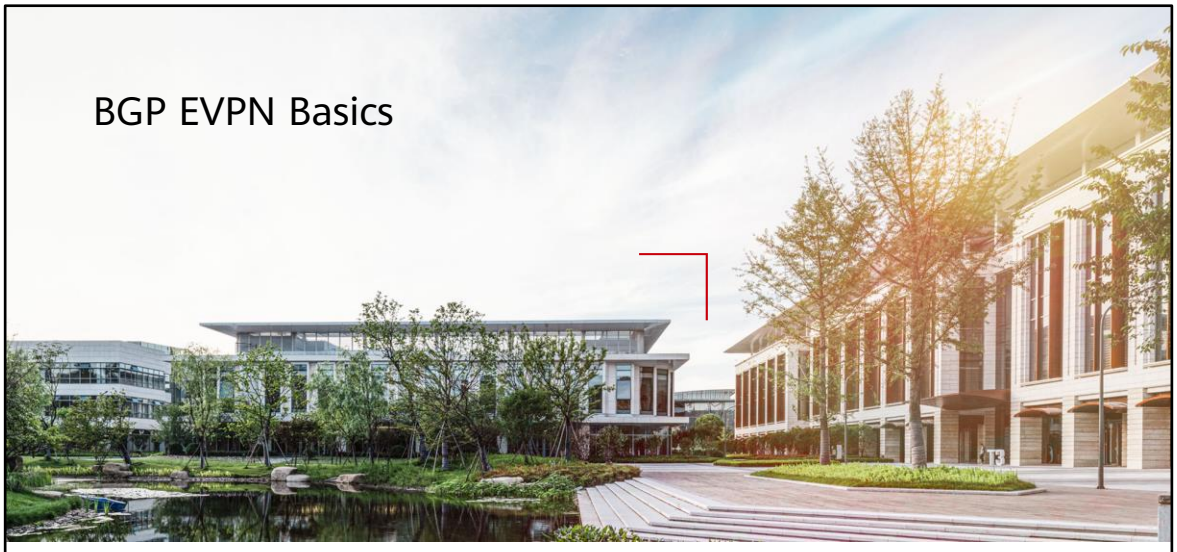
- IPsec uses IKE to transmit information required for encryption (IPsec SAs).
- To secure the transmission of security parameters by IKE, an IKE SA is established before security parameters are transmitted.
- Two IKE versions are available:
  - IKEv1
    - In IKEv1 phase 1, IKE SAs need to be negotiated. In IKEv2 phase 2, IPsec SAs need to be negotiated.
    - IKEv1 supports two modes: main mode and aggressive mode.
    - When a new IPsec tunnel needs to be established between a pair of devices, IKEv1 needs to renegotiate IKE SAs and IPsec SAs.
  - IKEv2
    - IKEv2 negotiates IKE SAs and IPsec SAs through Initial Exchanges.
    - When a new IPsec tunnel needs to be established between a pair of devices, IKEv2 can generate a new IPsec SA through Create_Child_SA Exchange, without the need to exchange IKE SAs again.

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.

BGP EVPN Basics



HUAWEI

Foreword

- Standard BGP-4 supports only IPv4 unicast addresses. To support more network layer protocols, Multiprotocol Extensions for BGP-4 (MP-BGP) (RFC 4760) was proposed as an extension to BGP-4 to allow different types of address families to be distributed in BGP at the same time. The address families include IPv4 multicast, IPv6, L3VPN, and Ethernet Virtual Private Network (EVPN) address families.
- With the development and commercial use of software-defined networking (SDN), EVPN plays an important role in various solutions, covering all scenarios, including campus networks, data centers, IP WAN transport networks, and software-defined networking in a wide area network (SD-WAN).
- This course describes the concept of MP-BGP, development history of EVPN, common EVPN route types, and EVPN usage scenarios.

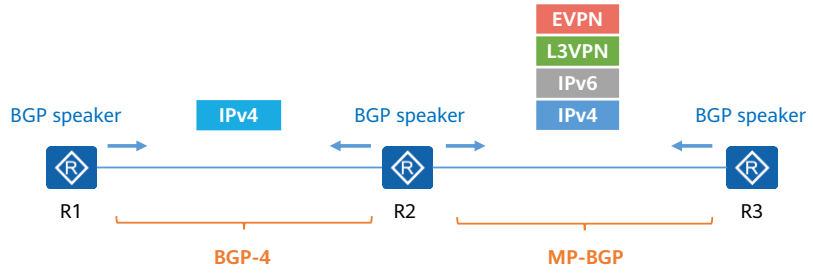
Objectives

- Upon completion of this course, you will be able to:
  - Understand basic MP-BGP concepts.
  - Understand the origin of EVPN.
  - Understand common EVPN route types.
  - Understand typical EVPN usage scenarios.

Contents

MP-BGP

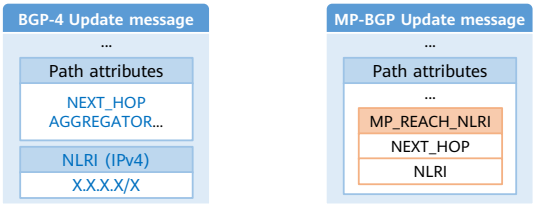
As defined in RFC 4760, MP-BGP is used to extend BGP-4 to allow BGP to carry multiple network layer protocols, such as IPv6, L3VPN, and EVPN. This extension has good backward compatibility. That is, an MP-BGP-capable router can interact with a router that supports only BGP-4.



- https://datatracker.ietf.org/doc/rfc4760/

BGP-4 Extensions

- BGP-4 has three IPv4-specific pieces of information: NEXT_HOP, AGGREGATOR, and IPv4 network layer reachable information (NLRI). To support multiple network layer protocols, BGP-4 has to provide the following abilities:
  - Ability of associating network layer protocols with next-hop information
  - Ability of associating network layer protocols with NLRIs
- The two abilities are collectively referred to as the address family (AF) defined by the Internet digital distribution agency (IANA).
- To implement forward compatibility, MP-BGP adds two new attributes: MP_REACH_NLRI and MP_UNREACH_NLRI, which are used to indicate reachable and unreachable destinations, respectively. The two attributes are optional non-transitive.



HUAWEI

- According to BGP-4, NEXT_HOP and AGGREGATOR fields are contained in Path attributes of IPv4, and the IPv4 NLRI carries IPv4 routing entries.
- The Path attributes field is added in MP-BGP. MP_REACH_NLRI is a new field of path attributes. The NEXT_HOP and NLRI fields of the corresponding network layer protocol and the NLRI belong to MP_REACH_NLRI.

MP_REACH_NLRI

- MP_REACH_NLRI is carried in a BGP Update message and provides the following functions:
  - Advertises reachable routes to BGP peers.
  - Advertises the next-hop address of a reachable route to a BGP peer.
- It contains the following fields:

MP_REACH_NLRI Format	Field Description
Address Family Identifier (2 octets)	Network layer protocol. Value 2 indicates IPv6.
Subsequent Address Family Identifier (1 octet)	This field is used together with address family identifier (AFI). Value 1 indicates unicast, and value 2 indicates IPv6 unicast.
Length of Next Hop Network Address (1 octet)	Length of a next-hop IP address.
Network Address of Next Hop (variable)	Next-hop address. The format is determined by the AFI and SAFI.
Reserved (1 octet)	All 0s.
Network Layer Reachability Information (variable)	The length of this field is variable. This field can contain reachable routes.

- In the SAFI field, value 1 indicates unicast, and value 2 indicates multicast. The value is allocated by the IANA. The allocation rules are defined in RFC 2434 (titled "Guidelines for Writing an IANA Considerations Section in RFCs").
- In this section, the AFI of EVPN is 25 (L2VPN) and the SAFI is 70 (EVPN).

MP_UNREACH_NLRI

- MP_UNREACH_NLRI is carried in BGP Update messages to withdraw unreachable routes.
- It contains the following fields:

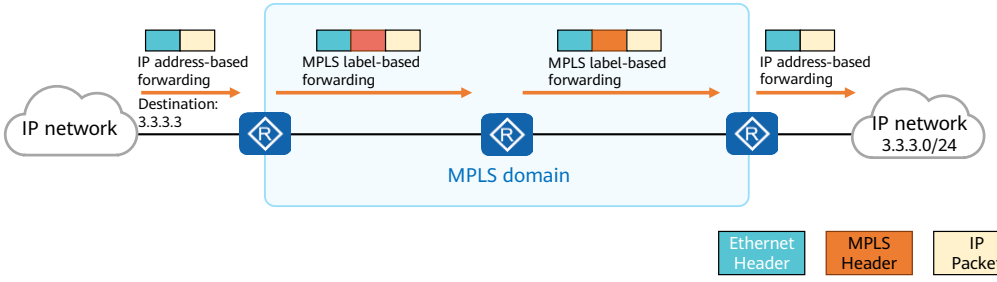
MP_UNREACH_NLRI Format	Field Description
Address Family Identifier (2 octets)	Network layer protocol. Value 2 indicates IPv6.
Subsequent Address Family Identifier (1 octet)	This field is used together with the AFI. Value 1 indicates unicast, and value 2 indicates IPv6 unicast.
Withdrawn Routes (variable)	The length of this field is variable. This field lists the routes that need to be withdrawn. The format of this field is determined by the AF and SAFI.

- The AFI of EVPN is 25 (L2VPN) and the subsequent address family identifier (SAFI) is 70 (EVPN).

Contents

MPLS Overview

- Multiprotocol Label Switching (MPLS) is located between the data link layer and the network layer in the TCP/IP protocol stack. An MPLS header is added between the two layers. Packets are forwarded based on the MPLS header. The MPLS header is also called the MPLS label.
- MPLS replaces IP forwarding with label switching to implement label-based rapid forwarding.

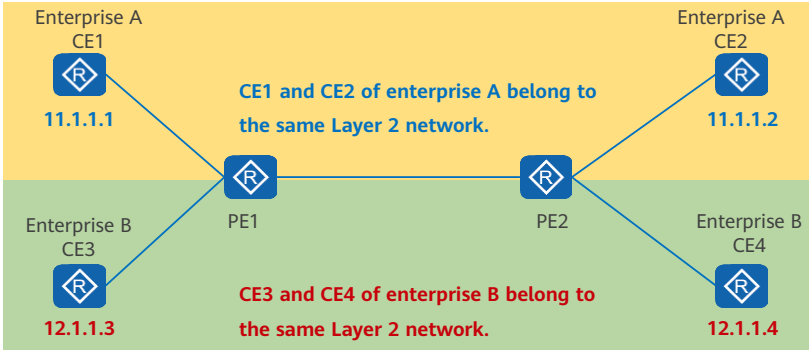


HUAWEI

- MPLS originates from IPv4 and its core technologies can be extended to multiple network protocols, including IPv6, Internet Packet Exchange (IPX), Appletalk, DECnet and Connectionless Network Protocol (CLNP). "Multiprotocol" in MPLS indicates that multiple network protocols are supported.
- MPLS replaces IP forwarding with label switching. A label is a short and fixed-length connection identifier that has only local significance. It is similar to the virtual path identifier (VPI)/virtual channel identifier (VCI) of Asynchronous Transfer Mode (ATM) and the data link connection identifier (DLCI) of Frame Relay.
- MPLS domain: An MPLS domain consists of a series of consecutive network devices that run MPLS.

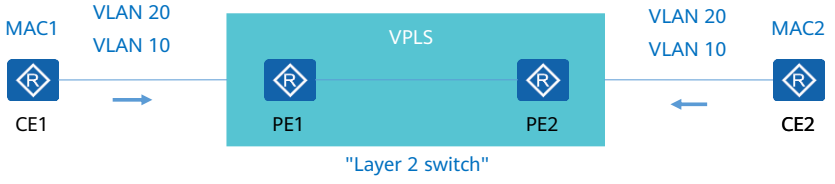
VPLS Overview

Virtual private LAN service (VPLS) is an Ethernet-based L2VPN technology. VPLS provides services similar to LAN services on an MPLS network and allows users to access the network from different locations.



Traditional L2VPN

- Traditional L2VPN services, such as VPLS, provide Layer 2 connections between remote sites. An L2VPN network is built and functions like a Layer 2 switch to transparently transmit Ethernet packets. In this example, PE1 and PE2 form a VPLS network to transparently transmit VLAN traffic between CE1 and CE2.
- In a traditional L2VPN, remote MAC addresses are learned through ARP broadcast flooding, and therefore, PEs need to carry broadcast traffic. Broadcast consumes a large amount of interface bandwidth, which is a typical issue of traditional L2VPN.

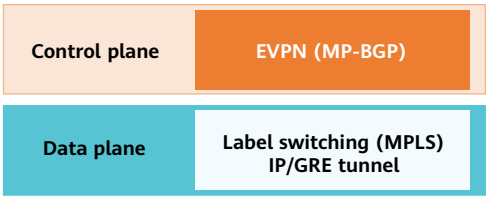


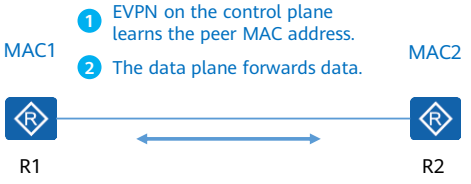
HUAWEI

- VPLS does not support all-active access or load balancing and implements slow fault convergence. For details, see materials of the HCIE-HCIE-Datacom Ethernet VPN and RFC 7209 titled "Requirements for Ethernet VPN (EVPN)."

Emergence of EVPN

- With new technologies and scenarios emerging, VPLS cannot meet the requirements of L2VPN. The industry has reviewed the requirements for Ethernet VPN (RFC 7209) and proposed a new solution, that is, EVPN.
- EVPN was first defined in RFC 7432. EVPN introduces the control plane to better control MAC address learning.
- EVPN uses MP-BGP on the control plane and supports MPLS label switched paths (LSPs) or IP/Generic Routing Encapsulation (GRE) tunneling on the data plane.



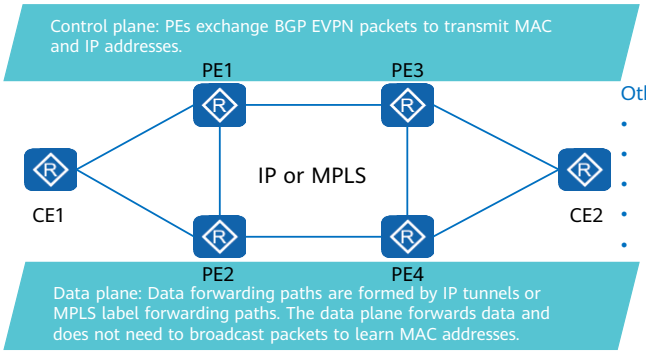


HUAWEI

- https://datatracker.ietf.org/doc/rfc7209/
- https://datatracker.ietf.org/doc/rfc7432/

Advantages of EVPN

- EVPN introduces the control plane to learn MAC and IP addresses to guide data forwarding, implementing forwarding-control separation.
- EVPN resolves typical problems in traditional L2VPNs and offers more benefits, such as active-active, rapid convergence, and simplified O&M.



Other advantages of EVPN:

- CEs can access PEs in all-active mode.
- Automatic discovery of PE members
- Loop prevention
- Broadcast traffic optimization
- Support for Equal-Cost Multi-Path (ECMP)

HUAWEI

- For more details, see the *HCIE-Datacom Ethernet VPN*.

Contents

EVPN NLRI

- EVPN defines a new type of BGP NLRI, known as EVPN NLRI, to carry all EVPN routes.
- EVPN NLRI is a new extension to MP-BGP. It is included in MP_REACH_NLRI. For the EVPN NLRI, the AFI is 25 and the SAFI is 70.

MP_REACH_NLRI format	MP_REACH_NLRI of BGP EVPN
Address Family Identifier (2 octets)	**AFI: 25**
Subsequent Address Family Identifier (1 octet)	**SAFI: 70**
Length of Next Hop Network Address (1 octet)	Length of a next-hop IP address.
Network Address of Next Hop (variable)	Next-hop IP address in an EVPN route.
Reserved (1 octet)	All 0s.
Network Layer Reachability Information (variable)	**EVPN NLRI**

EVPN Route

- The EVPN NLRI format uses the Type-Length-Value (TLV) structure, making packets highly flexible and scalable.
  - Route Type field: defines different EVPN routes. RFC 7432 defines four types of routes.
  - Length field: defines the length of a field.
  - Route Type Specific field: contains fields of a particular route type.

EVPN NLRI format

1 Ethernet A-D Route
2 MAC Advertisement Route
3 Inclusive Multicast Route
4 Ethernet Segment Route

Four common EVPN routes defined in the Route Type field

HUAWEI

- The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the EVPN NLRI (encoded as specified above).
- The EVPN NLRI is carried in BGP [RFC4271] using BGP Multiprotocol Extensions [RFC4760] with an Address Family Identifier (AFI) of 25 (L2VPN) and a Subsequent Address Family Identifier (SAFI) of 70 (EVPN). The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the EVPN NLRI (encoded as specified above).
- In order for two BGP speakers to exchange labeled EVPN NLRI, they must use BGP Capabilities Advertisements to ensure that they both are capable of properly processing such NLRI. This is done as specified in [RFC4760], by using capability code 1 (multiprotocol BGP) with an AFI of 25 (L2VPN) and a SAFI of 70 (EVPN).

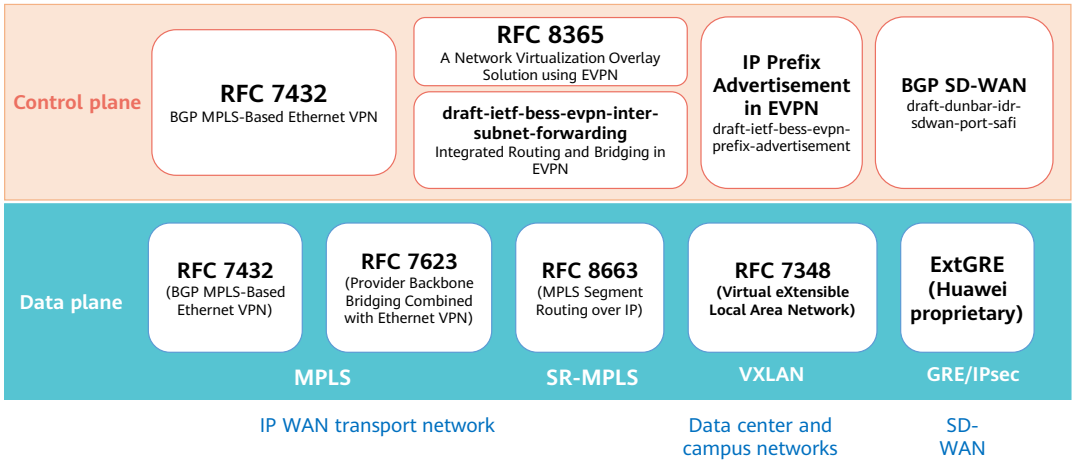
More Types of EVPN Routes and Their Functions

EVPN is not limited to L2VPN applications. With the increase of EVPN route types, more applications, such as L3VPN, are supported.

Type of Route	Function	RFC
(Type 1) Ethernet A-D Route	• Aliasing • MAC address batch withdraw • All-active flag • ESI label advertisement	RFC 7432
(Type 2) MAC/IP Advertisement Route	• MAC address learning notification • MAC/IP binding • MAC mobility	
(Type 3) Inclusive Multicast Route	Automatic discovery of multicast tunnel endpoints and multicast types	
(Type 4) Ethernet Segment Route	Automatic discovery of ES members DF election	
(Type 5) IP Prefix Route	IP prefix advertisement (support for L3VPN)	draft-ietf-bess-evpn-prefix-advertisement

- The Type 5 route (IP prefix route) related standard is in the draft phase, in draft-ietf-bess-evpn-prefix-advertisement.

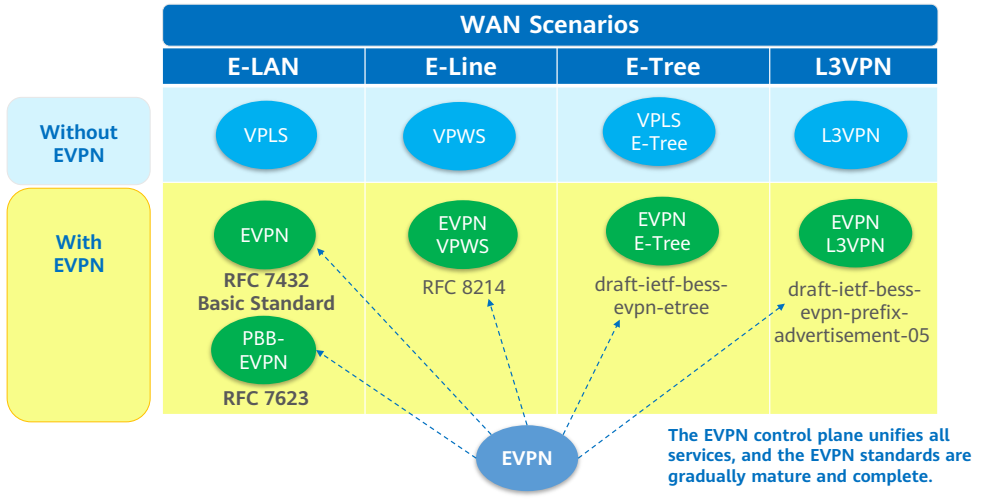
EVPN Protocol Standards



HUAWEI

Contents

EVPN on an IP WAN Transport Network

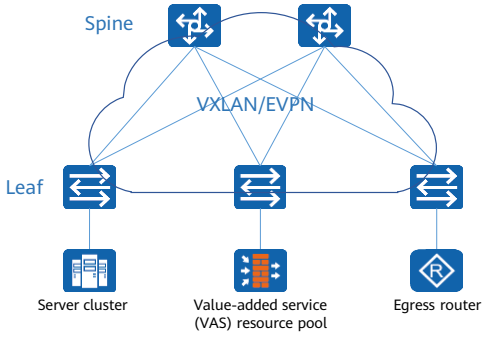


HUAWEI

- E-Line, E-Tree, and E-LAN are three types of Ethernet virtual circuits (EVCs). For details, see metro Ethernet standards at https://wiki.mef.net/display/CESG/E-Line.
- The Metropolitan Ethernet Forum (MEF) defines three types of EVCs: point-to-point EVC, multipoint-to-multipoint EVC, and root-multipoint EVC.
  - E-Line: A point-to-point EVC strictly associates two User-to-Network Interfaces (UNIs).
  - E-LAN: A multipoint-to-multipoint EVC can associate two or more UNIs. Users or carriers can add any UNIs to the EVC or delete some UNIs from the EVC without affecting other UNIs.
  - E-Tree: This EVC is similar to the hub-spoke model in L3VPN. It consists of one or more root UNIs and several leaf UNIs. The root UNI can directly communicate with all UNIs in the EVC, whereas a leaf UNI can only communicate directly with the root UNI in the EVC, and two leaf UNIs cannot communicate with each other directly.

EVPN on a Data Center Network

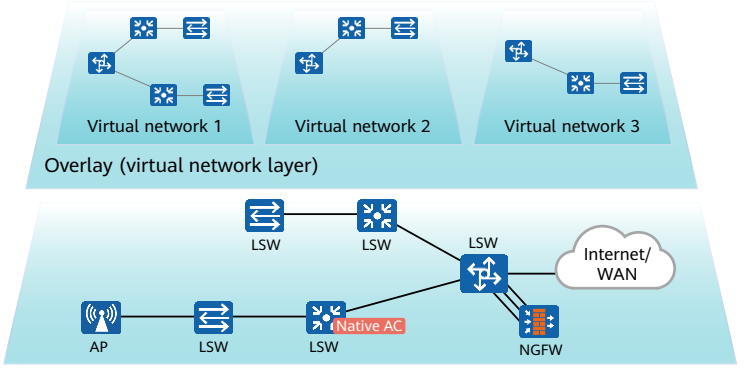
- The network virtualization overlay (NVO) solution (RFC 8365) of EVPN is used in cloud data centers.
- It is recommended that the data plane use Virtual Extensible LAN (VXLAN) encapsulation and the control plane use EVPN to construct a flexible data center overlay network.



- All services in the data center are carried by the VXLAN overlay network.
- The underlay network consisting of spine and leaf nodes performs high-speed forwarding.

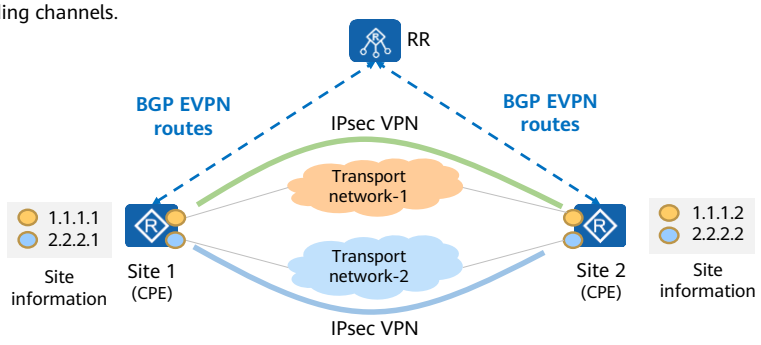
EVPN Application on a Campus Network

- The campus network virtualization solution is similar to that in the cloud data center. The EVPN NVO solution (RFC 8365) is used.
- VXLAN encapsulation and EVPN are used on different underlying networks to build a flexible overlay network.



EVPN Application in SD-WAN

- SD-WAN is a next-generation enterprise branch interconnection solution that supports intelligent dynamic traffic steering, Zero Touch Provisioning (ZTP), and visualization.
- In the SD-WAN solution, EVPN is deployed between route reflectors (RRs) and customer-premises equipment (CPE) devices to advertise SD-WAN overlay VPN routes on the control plane. IPsec VPN is used on the data plane to build secure forwarding channels.



HUAWEI

- Overlay VPN routes include site VPN route prefixes, next-hop route information, and IPsec key pairs required for data encryption of data channels between CPEs. For details, see materials of the SD-WAN course.

Quiz

1. (Essay) Please describe the principles and common route types of EVPN.
2. (Essay) Please describe usage scenarios of EVPN.

1. EVPN is an extension to MP-BGP. EVPN provides five major types of routes and is used as the control plane of Layer 2 or Layer 3 tunnels.
2. EVPN can be widely used in all enterprise scenarios, such as SD-WAN, campus networks, data centers, and WANs. In data centers and campus networks, EVPN and VXLAN are used together to construct a service overlay network. In SD-WAN scenarios, EVPN and IPsec are used together to build enterprise branch interconnection networks. On a WAN, EVPN can be used with various underlying tunneling and label technologies, such as MPLS, Segment Routing (SR), VPLS, and virtual private wire service (VPWS).

Summary

- MP-BGP's extension to BGP-4 allows different types of address families, such as IPv4 multicast, IPv6, L3VPN, and EVPN, to be distributed in BGP.
- This course describes EVPN that is used to solve the Ethernet L2VPN problems. With the increase of usage scenarios and protocol extensions, EVPN can be used in various scenarios, including WANs, data centers, campus networks, and SD-WANs.

More Information

https://datatracker.ietf.org/doc/rfc4760/

https://datatracker.ietf.org/doc/rfc7209/

https://datatracker.ietf.org/doc/rfc7432/

https://datatracker.ietf.org/doc/rfc8365/

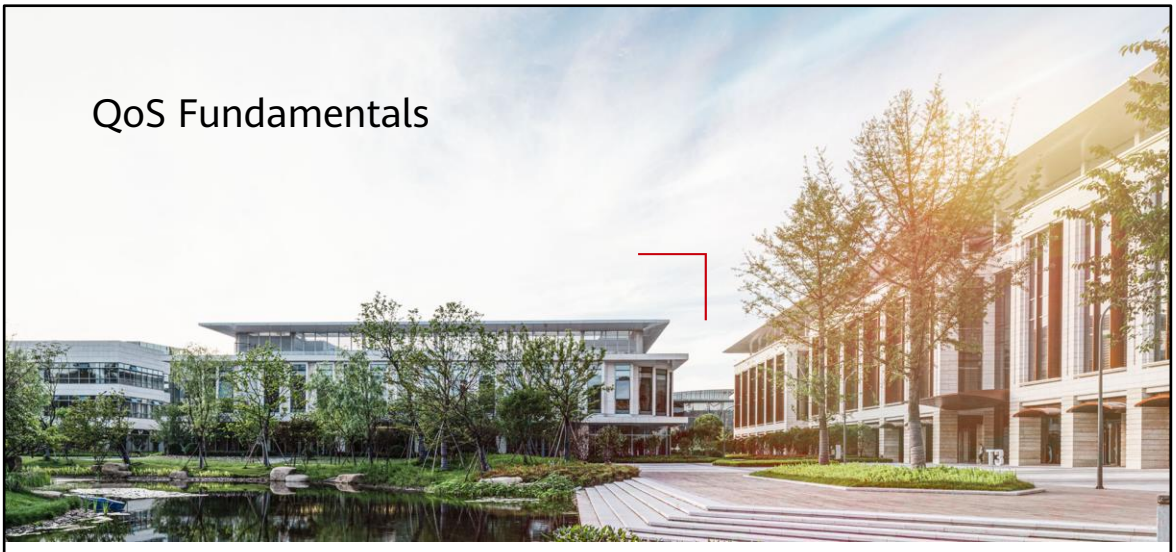
https://wiki.mef.net/display/CESG/MEF+6.3+-+EVC+Ethernet+Services+Definitions

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.

QoS Fundamentals



HUAWEI

Foreword

- With continuous development of networks, the network scale and traffic types increase continuously. As a result, Internet traffic increases sharply, network congestion occurs, the forwarding delay increases, and even packet loss occurs. In this case, the service quality deteriorates or even services are unavailable. To deploy real-time and non-real-time services on the IP network, network congestion must be resolved. The commonly used solution is to increase the network bandwidth. However, this solution is not ideal considering the network construction cost.
- Quality of service (QoS) is introduced in this situation. At limited bandwidth, QoS uses a "guaranteed" policy to manage network traffic and provides different priority services for different traffic.
- This course describes QoS fundamentals.

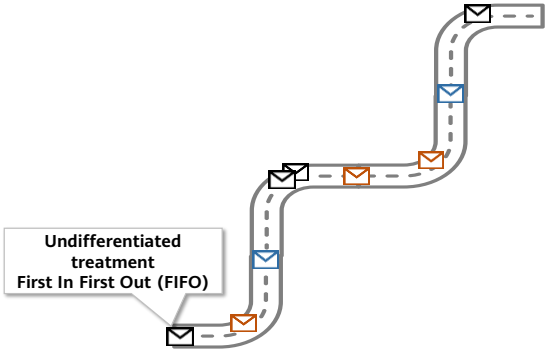
Objectives

- Upon completion of this course, you will be able to:
  - Describe the QoS background.
  - Describe QoS types.
  - Describe the implementation of the QoS DiffServ model.
  - Describe application scenarios of different QoS functions.
  - Describes basic configuration of QoS.
  - Describe HQoS fundamentals.
  - Describe basic configuration of HQoS.

Contents

"Best-Effort" Traditional Network

- When the IP network emerges, there is no QoS guarantee.
- You only know that the packets have been sent out. Whether the packets can be received and when the packets can be received are unknown.



HUAWEI

- On the traditional IP network, each network device handles all packets in an undifferentiated manner and follows the First In First Out (FIFO) rule to transmit packets. The devices transmit packets to the destination in best-effort (BE) mode, but the BE mode cannot ensure the performance such as delay and reliability.

QoS Background

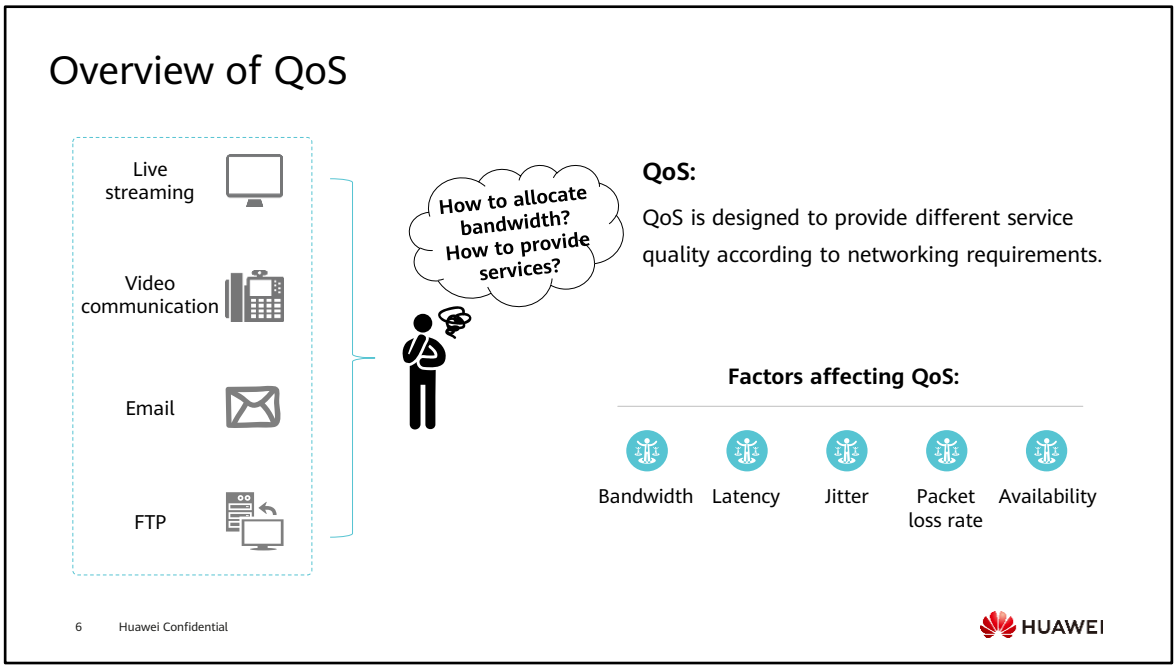
- With continuous technology improvement and fierce product competition, users have increasingly higher requirements on the network quality.

High-definition image quality and high network speed
Good signal quality

Poor image quality, low network speed, and frame freezing
Poor signal quality

HUAWEI

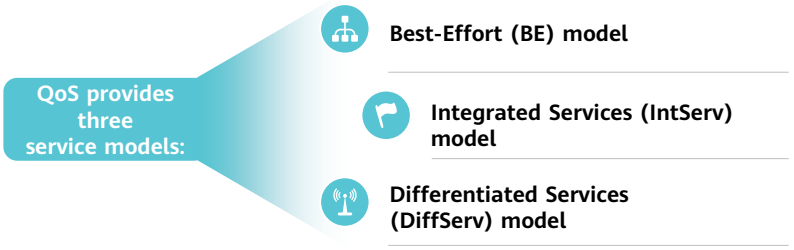
- With the emergence of new applications on IP networks, new requirements are raised to QoS of IP networks.



- To support voice, video, and data services of different requirements, the network is required to distinguish different communication types before providing corresponding QoS.
  - For example, real-time services such as Voice over IP (VoIP) demand shorter latency. A long latency for packet transmission is unacceptable. Email and the File Transfer Protocol (FTP) services are comparatively insensitive to the latency.
- To support voice, video, and data services of different requirements, the network is required to distinguish different communication types before providing corresponding QoS.
  - The BE mode of traditional IP networks cannot identify and distinguish various communication types on the networks. This distinguishing capability is the premise for providing differentiated services. The BE mode cannot satisfy application requirements, so QoS is introduced.
- What is QoS?

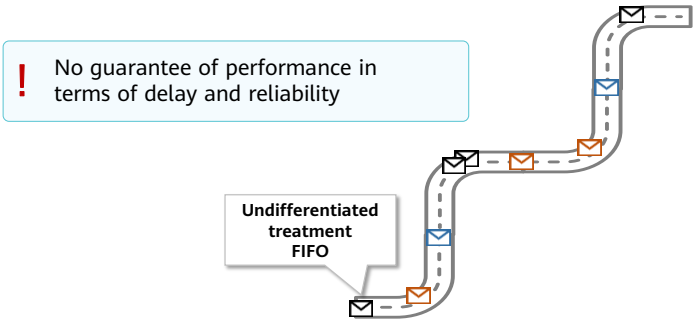
- QoS is designed to provide different service quality according to networking requirements. Example:
  - The bandwidth used by FTP on the backbone network can be limited, and database access can be given a higher priority.
  - For an ISP, its users may transmit voice, video, or other real-time services. QoS enables the ISP to differentiate these packets and provide different services.
  - QoS can provide bandwidth and low delay guarantee for time-sensitive multimedia services, and other services on the network do not affect these time-sensitive services.
- Which factors affect QoS?
  - Bandwidth: indicates the transfer speed of IP packets on a network. It can be the average value or peak value. — Bandwidth competition can be resolved by increasing the bandwidth. However, the bandwidth cannot be increased infinitely.
  - Latency: indicates the round trip time (RTT) of an IP packet between two nodes on a network. — Delay-sensitive traffic, such as video and voice traffic
  - Jitter: indicates the change in the latencies of different packets which are in the same data stream and transferred in the same direction. — It is related to the latency. If the latency is short, the jitter range is small, which has a great impact on real-time services such as voice and video services.
  - Packet loss rate: indicates the allowed maximum packet loss rate when a service is transmitted on a network. — It is used to measure the network reliability. A small number of lost packets have little impact on services, but a large number of lost packets severely affect the transmission efficiency.
  - Availability: indicates the availability of a connection between a user and the IP service, including the connection setup time and holding time.

QoS Service Models



BE Model

- An application can send any number of packets at any time.
- The network then makes the best effort to transmit the packets.

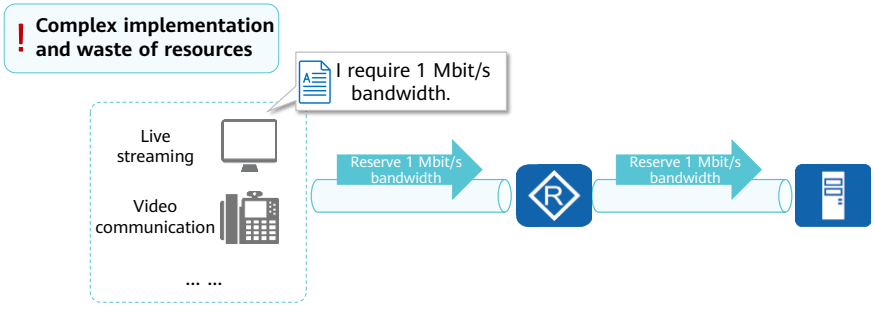


HUAWEI

- The BE model is the simplest service model in which an application can send any number of packets at any time without obtaining approval or notifying the network.
- The network then makes the best effort to transmit the packets but provides no guarantee of performance in terms of delay and reliability.
- The BE model is the default service model for the Internet and applies to various network applications, such as the File Transfer Protocol (FTP) and email. It uses FIFO queues.

IntServ Model

- Before sending packets, an application needs to apply for specific services through signaling.
- After receiving a resource request from an application, the network reserves resources for each information flow by exchanging RSVP signaling information.



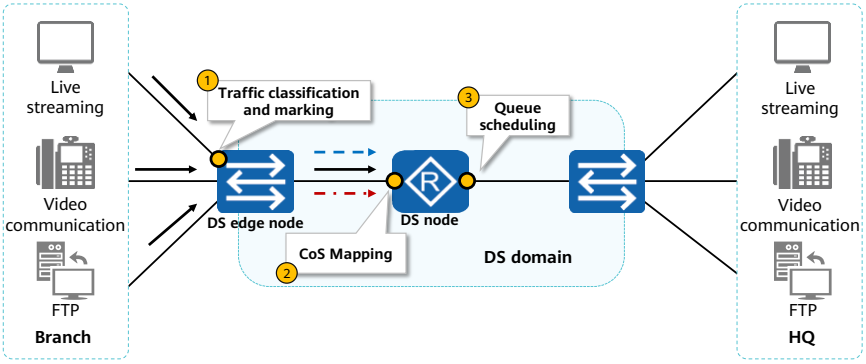
HUAWEI

- The IntServ model is a comprehensive service model to meet various QoS requirements.
- Before sending packets, an application needs to apply for specific services through signaling. This request is sent through RSVP. RSVP applies for network resources for an application before the application starts to send packets.
- Once the network determines to allocate resources to the application, the network maintains a state for each flow (determined by IP addresses, port numbers, and protocol numbers at both ends), and performs packet classification, traffic policing, queuing, and scheduling based on the state. After receiving the acknowledgment message from the network (the application confirms that the network has reserved resources for the packets of the application), the application starts to send packets. As long as packets of the application are controlled within the range described by traffic parameters, the network promises to meet QoS requirements of the application.
- Example: If you want to reserve a vehicle, you need to apply for a service in advance and reserve resources when resources are sufficient.
- However, the vehicle service vendor has to maintain a large number of booking information.
- Disadvantage: The implementation of the IntServ model is complex. When no traffic is transmitted, the bandwidth is still exclusively occupied, and the usage is low. This solution requires that all end-to-end nodes support and run the RSVP

protocol.

DiffServ Model

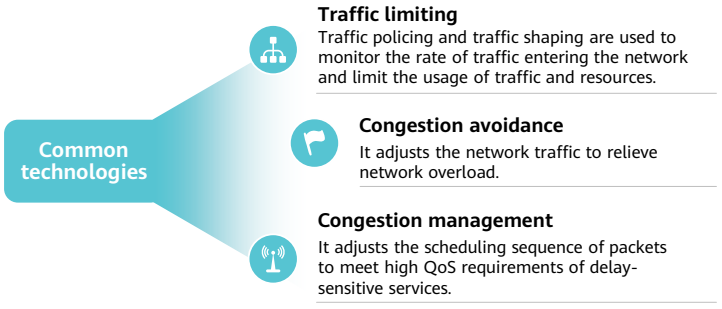
- Traffic on a network is classified into multiple classes, and a corresponding processing behavior is defined for each class, so that the traffic has different forwarding priorities, packet loss rates, and delays.



HUAWEI

- DiffServ is a multi-service model and can satisfy different QoS requirements. Currently, this model is widely used on IP networks.
- Before sending a packet, the application does not need to notify the network to reserve resources for it. In the DiffServ model, the network does not need to maintain the status of each flow. Instead, it provides specific services based on precedence fields of packets (for example, the DS field in the IP header).
- The DiffServ model classifies network traffic into multiple classes for differentiated processing. To be specific, the DiffServ model implements traffic classification first and allocates different identifiers to different classes of packets. After a network node receives these packets, it simply identifies these identifiers and processes packets based on the actions corresponding to these identifiers.
- There is an analogy between the DiffServ model and train ticket service system. A train ticket marks the service that you book: soft sleeper, hard sleeper, hard seat, or no seat. You get on a train and enjoy the specific service marked in your ticket. On an IP network, an identifier is to a packet as a train ticket is to a passenger.
- In addition to traffic classification and marking, the DiffServ model provides the queuing mechanism. When network congestion occurs on a device, the device buffers packets in queues. The device sends the packets out of queues when network congestion is relieved.

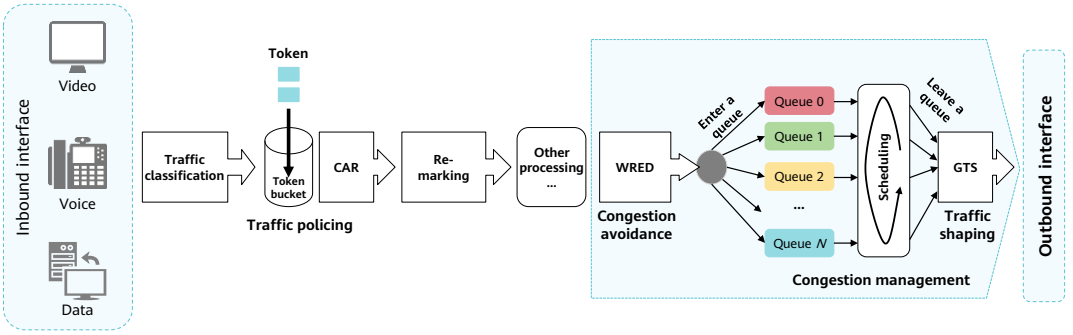
Common QoS Technologies (DiffServ Model)



HUAWEI

- Rate limiting: Traffic policing and traffic shaping monitor the rate of traffic entering the network to limit the traffic and resource usage, providing better services for users.
- Congestion avoidance and congestion management: When congestion occurs on a network, the device determines the sequence in which packets are forwarded according to a certain scheduling policy so key services are processed first. Or, the device proactively adjusts traffic to relieve network overload by discarding packets.

QoS Data Processing (DiffServ Model)



HUAWEI

- QoS technology provides the following functions:
  - Traffic classification and marking: identify objects based on certain matching rules, which is the prerequisite for implementing differentiated services. They are usually applied to the inbound direction of an interface.
  - Token bucket: is used to check whether traffic meets packet forwarding conditions.
  - Traffic policing: monitors the volume of specific data traffic that arrives at network devices, and is usually applied to incoming traffic. When the traffic volume exceeds the maximum value, traffic limiting or punishment measures are taken to protect business interests and network resources of service providers.
  - Congestion avoidance: Excessive congestion may damage network resources. Congestion avoidance monitors the usage of network resources. When congestion aggravates, congestion avoidance proactively adjusts traffic to relieve network overload by discarding packets. Congestion avoidance is generally applied to the outbound direction of an interface.
  - Congestion management: is taken to solve the problem of resource competition. Packets are buffered in queues and a scheduling algorithm is used to determine the forwarding sequence of packets. Congestion management is usually applied to the outbound direction of an interface.

- Traffic shaping: is a traffic control measure that initiatively adjusts the output speed of traffic. Traffic shaping enables the traffic to adapt to the network resources that can be provided by the downstream device to prevent packet loss and congestion. Traffic shaping is usually applied to the outbound direction of an interface.

Quiz

1. (Multiple-answer question) Which of the following service models are provided by QoS?
   A. DiffServ model
   B. IntServ model
   C. BE model
   D. Network service model

HUAWEI

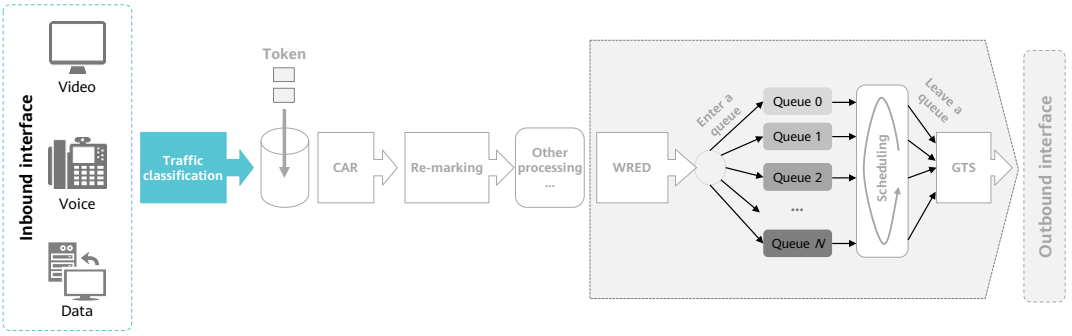
- 1. ABC

Section Summary

- QoS service models include the DiffServ, IntServ, and BE models.
- The DiffServ model is the most commonly used QoS model. It provides rate limiting, congestion avoidance, and congestion management.

Contents

QoS Data Processing

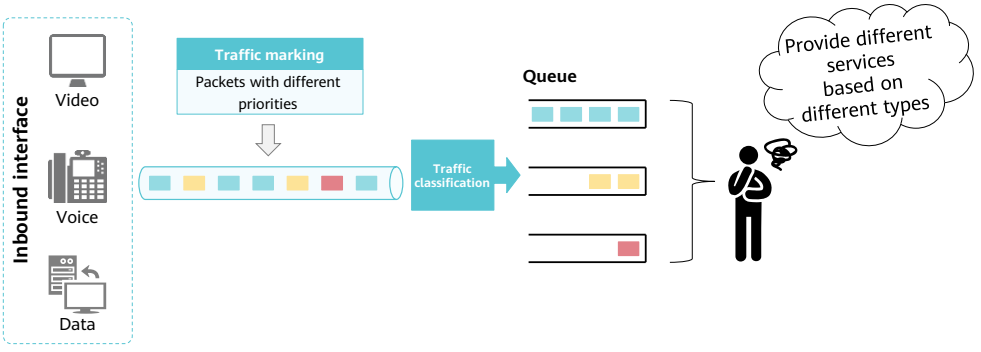


HUAWEI

- Traffic classification and marking: identify objects based on certain matching rules, which is the prerequisite for implementing differentiated services. They are usually applied to the inbound direction of an interface.

Why Are Traffic Classification and Traffic Marking Required?

- Traffic classification and marking are the basis of QoS and the prerequisite for implementing differentiated services.



HUAWEI

- To implement differentiated services, the traffic entering a DS domain needs to be classified according to certain rules, and then different services are provided for different types of traffic.
- After packets are classified at the DS domain edge, intermediate nodes provide differentiated services for classified packets. The downstream node can accept the classification result of its upstream node or classifies packets based on its own criteria.
- Traffic classification and marking are prerequisites for differentiated services.
  - Traffic classification technology classifies packets into different types, and does not modify the data packets.
  - Marking technology marks packets with different priorities, and modifies the data packets. Marking is classified into internal and external marking.
    - Internal marking

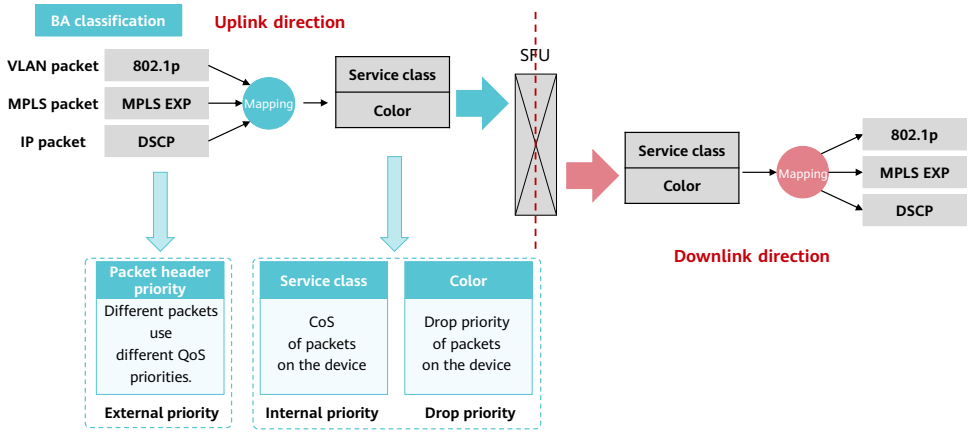
      Sets the CoS and drop precedence of packets for internal processing on a device so that packets can be placed directly in specific queues.

      Setting the drop precedence of packets is also called coloring packets. When traffic congestion occurs, packets in the same queue are provided with differentiated buffer services based on colors.

- External marking

  Sets or modifies the priority of packets so that the downstream device can provide services based on the changed priority. Modifying the packet priorities is also called re-marking.

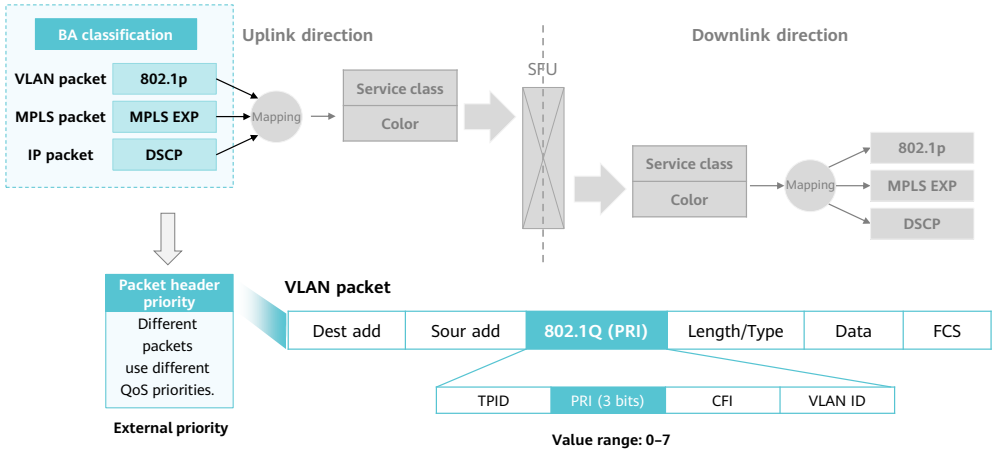
Behavior Aggregate Classification



HUAWEI

- BA classification allows the device to classify packets based on related values as follows: IP priority or DSCP value of IPv4 packets, TC value of IPv6 packets, EXP value of MPLS packets, and 802.1p value of VLAN packets. It is used to simply identify the traffic that has the specific priority or service classes for mapping between external and internal priorities.
- BA classification confirms that the priority of incoming packets on a device is trusted and mapped to the service class and color based on a priority mapping table. The service class and color of outgoing packets are then mapped back to the priority.
- Packets carry different types of precedence field depending on the network type. For example, packets carry the 802.1p value on a VLAN network, the EXP value on an MPLS network, and the DSCP value on an IP network. To provide differentiated services for different packets, the device maps the QoS priority of incoming packets to the scheduling precedence (also called service class) and drop precedence (also called color), and then performs congestion management based on the service-class and congestion avoidance based on the color. Before forwarding packets out, the device maps the service class and color of the packets back to the QoS priority, which provides a basis for other devices to process the packets.

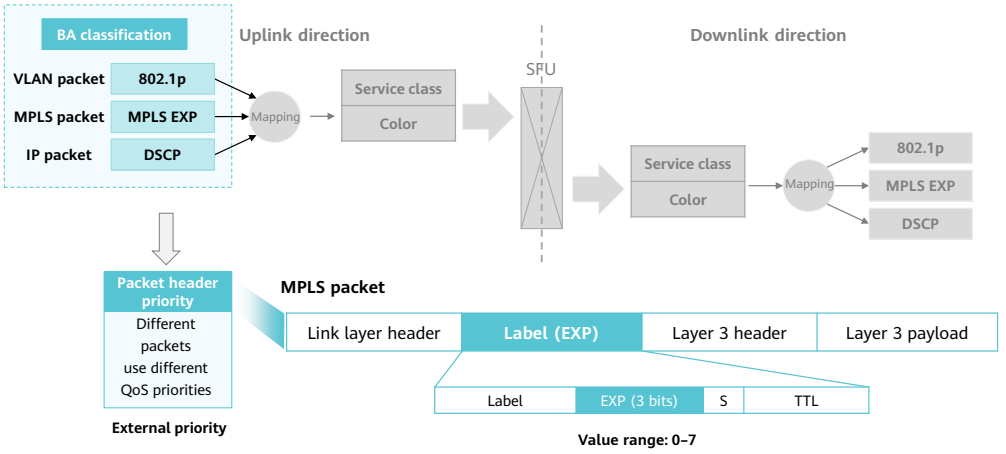
External Priority - VLAN Packet



HUAWEI

- Eight service priorities (PRIs) are defined in the VLAN tag of the Ethernet frame header.

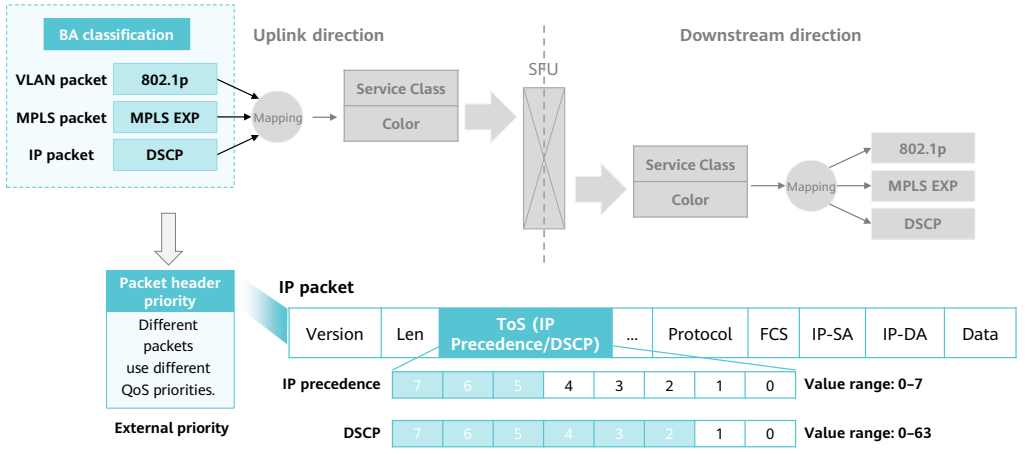
External Priority - MPLS Packet



HUAWEI

- The EXP field in the label is used as the external priority of MPLS packets to differentiate service classes of data traffic.

External Priority - IP Packet



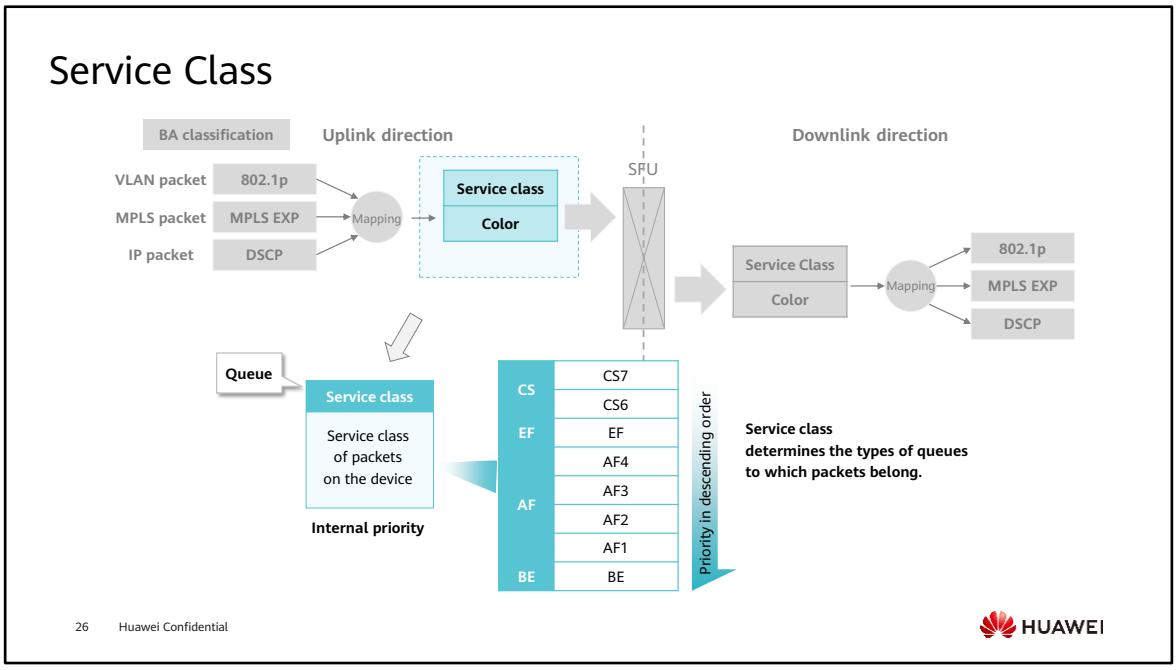
HUAWEI

- Eight IP service types are defined in the Precedence field of the ToS field in an IPv4 packet header.
- The ToS field in the IPv4 packet header is redefined as the Differentiated Services (DS) field. That is, the IP Precedence field is extended.

Mapping Between External Priorities

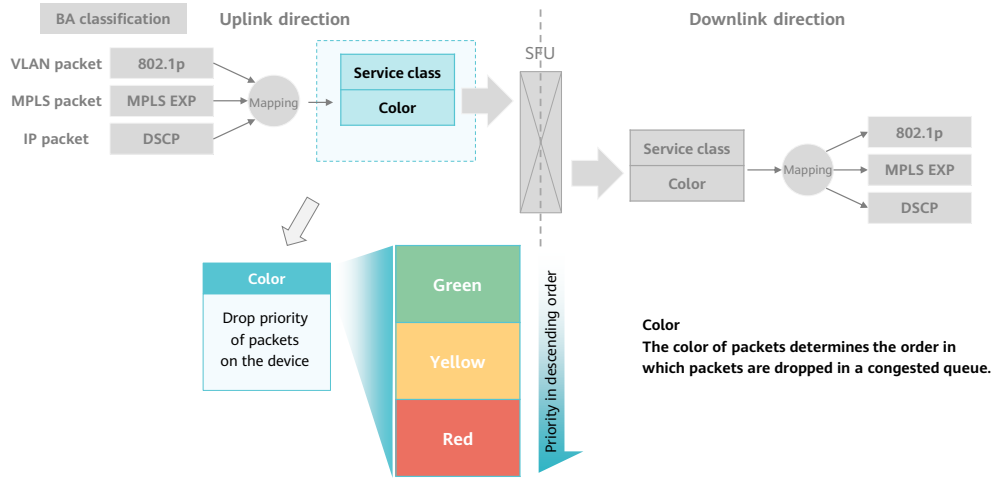
Priority in descending order

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			



- Service class, that is, queue
- Service classes refer to the internal priorities of packets. Eight service class values are available: class selector 7 (CS7), CS6, expedited forwarding (EF), assured forwarding 4 (AF4), AF3, AF2, AF1, and best-effort (BE). Service classes determine the types of queues to which packets belong.
- The priority of queues with a specific service class is calculated based on scheduling algorithms.
  - If queues with eight service classes all use priority queuing (PQ) scheduling, queues are displayed in descending order of priority: CS7 > CS6 > EF > AF4 > AF3 > AF2 > AF1 > BE.
  - If the BE queue uses PQ scheduling (this configuration is rare on live networks) but all the other seven queues use weighted fair queuing (WFQ) scheduling, the BE queue is of the highest priority.
  - If the queues of eight service classes all use WFQ scheduling, their priorities are the same.

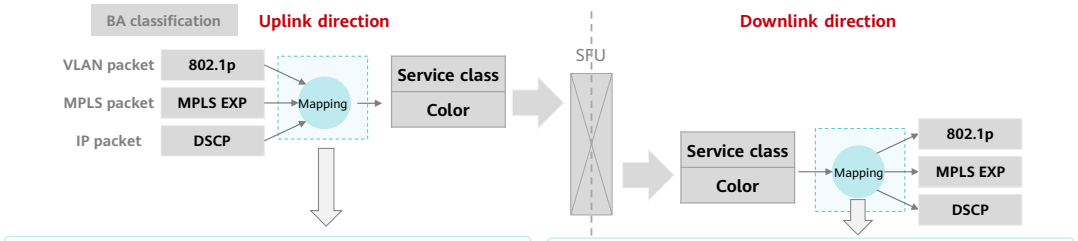
Color



HUAWEI

- Color, referring to the drop priority of packets on a device, determines the order in which packets in one queue are dropped when traffic congestion occurs.
- As defined by the Institute of Electrical and Electronics Engineers (IEEE), the color of a packet can be green, yellow, or red.
- Drop priorities are compared based on the configured parameters. For example, if a maximum of 50% of the buffer is configured to store packets colored green, whereas a maximum of 100% of the buffer is configured to store packets colored red, the drop priority of packets colored green is higher than that of packets colored red.

Mapping



Uplink mapping

DSCP				Service Class/Color			
32-39	34	36	38	AF	AF41	AF42	AF43
24-31	26	28	30		AF31	AF32	AF33
16-23	18	20	22		AF21	AF22	AF23
8-15	10	12	14		AF11	AF12	AF13

- Mapping from external priorities to internal priorities

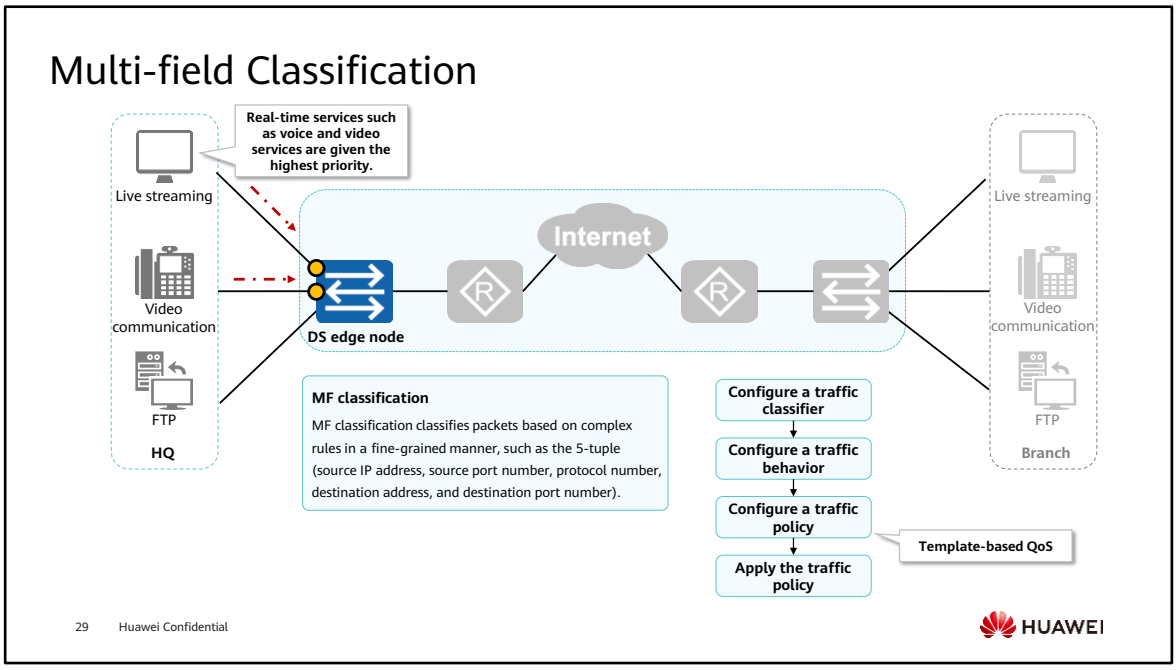
Downlink mapping

Service Class/Color				DSCP			
AF	AF41	AF42	AF43	32-39	34	36	38
	AF31	AF32	AF33	24-31	26	28	30
	AF21	AF22	AF23	16-23	18	20	22
	AF11	AF12	AF13	8-15	10	12	14

- Mapping from internal priorities to external priorities

HUAWEI

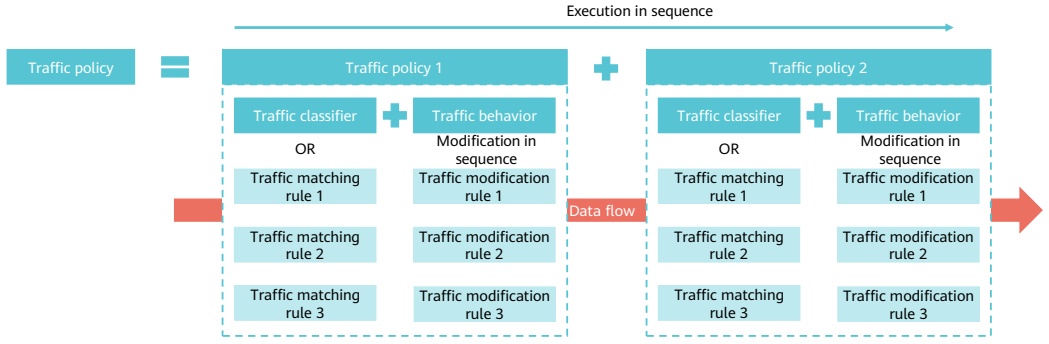
- A device maps the QoS priority to the service class and color for incoming packets and maps the service class and color back to the QoS priority for outgoing packets.



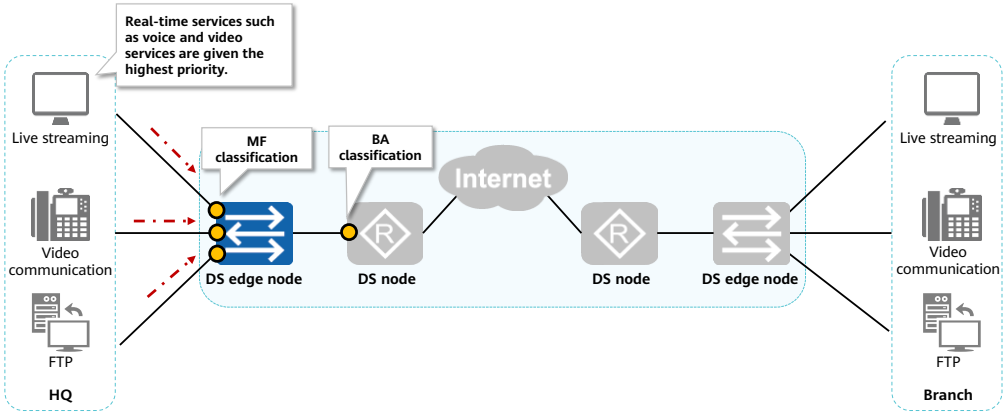
- As networks rapidly develop, services on the Internet become increasingly diversified. Various services share limited network resources. In particular, multiple services use port number 80. Because of this increasing demand, network devices are required to possess a high degree of sensitivity for services, including an in-depth parsing of packets and a comprehensive understanding of any packet field at any layer. This level of sensitivity rises far beyond what BA classification can offer. MF classification can be deployed to help address this sensitivity deficit.
- MF classification classifies packets based on complex rules in a fine-grained manner, such as the 5-tuple (source IP address, source port number, protocol number, destination address, and destination port number).

Traffic Policy Overview

- Modular QoS command line interface (MQC) uses traffic policies.
- A traffic policy is often bound to traffic classifiers and traffic behaviors. A traffic classifier is used to match data packets, and a traffic behavior is used to modify data packets.



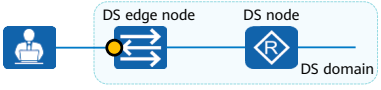
Traffic Classification Process



HUAWEI

- Requirement: The highest forwarding priority must be provided for real-time services such as voice and video services.
- Implementation: The DS edge node obtains service traffic such as voice and video traffic through MF classification and maps the traffic to the corresponding priorities. It processes the remaining traffic through BA classification.

Configuring MF Classification



- Typically, the traffic received by the DS edge node is not classified, so complex traffic classification is configured on the DS edge node. The configuration roadmap is as follows:
  - Configure a traffic classifier to match traffic.
  - Configure a traffic behavior to define an action taken on the matched traffic.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to the inbound direction of the interface on the DS edge node.

```
system-view
  traffic classifier [classifier-name]        //Create a traffic classifier.
    if-match [acl | vlan-id | .... ]        //Match traffic based on traffic characteristics.
```

```
system-view
  traffic behavior [behavior-name]        //Create a traffic behavior.
    remark [dscp-name | 8021p-value | EXP | ... ]        //Re-mark the QoS field of traffic.
```

```
system-view
  traffic policy [policy-name]     //Create a traffic policy.
    classifier [classifier-name] behavior [behavior-name]     //Bind the traffic classifier to the traffic behavior.
```

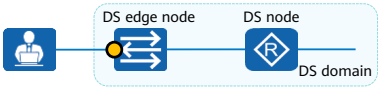
```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    traffic-policy [policy-name] [inbound | outbound]     //Apply the traffic policy to the inbound direction of an interface.
```

Checking the MF Classification Configuration

- After MF classification is configured, you can run the following commands to check the configuration.

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]    //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

(Optional) Modifying the BA Classification Configuration



- Based on the priority mapping table, BA classification maps data with the specific QoS field to the internal priority.
- The priority mapping table can be modified as required. The roadmap is as follows:
  - Specify the packet priority trusted on an interface.
  - Configure a priority mapping table.

- Specify the packet priority trusted on an interface.

```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    trust [8021p | dscp]       //Specify the priority to be trusted.
```

- Configure a priority mapping table.

```
system-view
  qos map-table [ dot1p-dot1p | dot1p-dscp | dot1p-lp | dscp-dot1p | dscp-dscp | dscp-lp ]      //Enter the priority mapping table view.
    input [input-value1] output [output-value]     //Configure mappings in the priority mapping table.
```

Checking the Priority Mapping Configuration

- After the priority mapping configuration is modified, you can run the following commands to check the configuration.

```
system-view
  display qos map-table [ dot1p-dot1p | dot1p-dscp | dot1p-lp | dscp-dot1p | dscp-dscp | dscp-lp ]
//Check the mapping between priorities.
```

Quiz

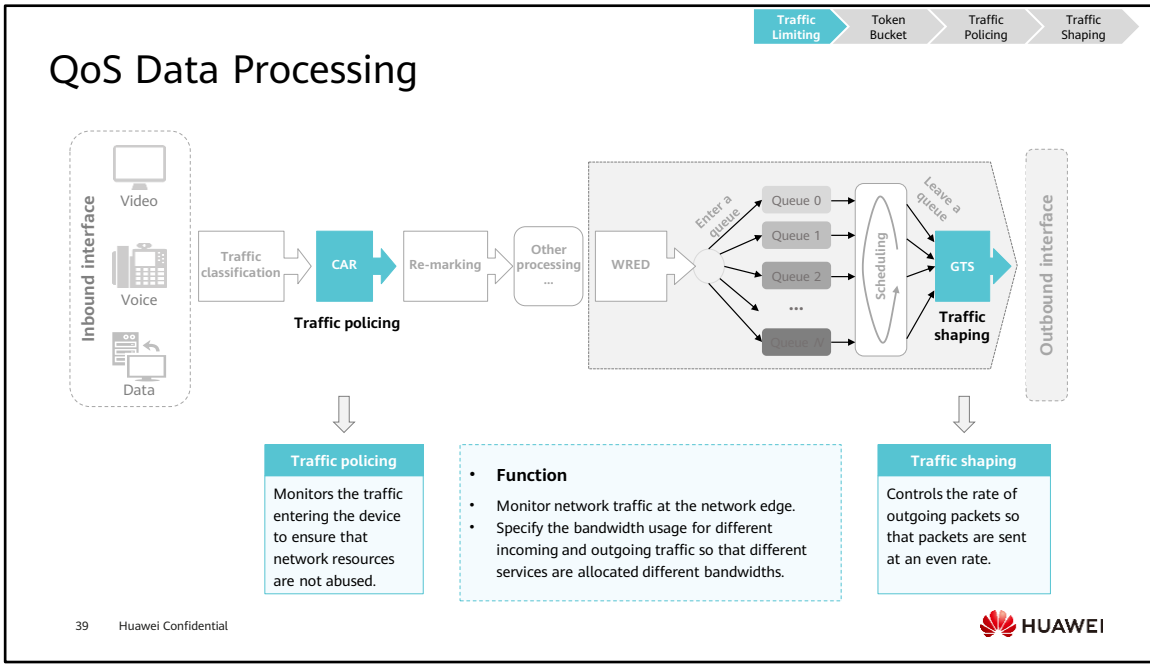
1. (True or false) MF classification is generally deployed in the inbound direction of the DS edge node.
   A. True
   B. False
2. (Multiple-answer question) Which of the following parameters are used to mark the QoS priority of data packets?
   A. EXP
   B. 802.1p
   C. DSCP
   D. IP precedence

- 1. A
- 2. ABCD

Section Summary

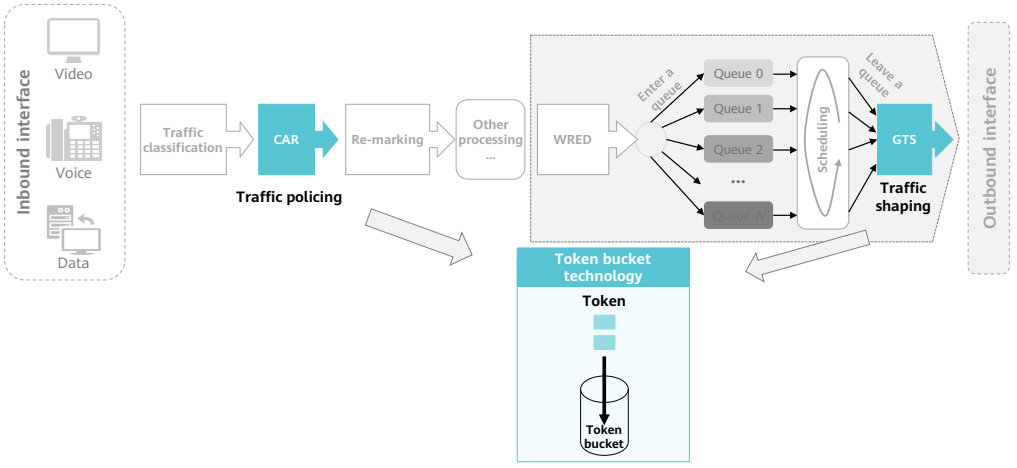
- The DiffServ model must mark packets for differentiating them. Generally, MF classification is used to mark incoming traffic on edge devices in a DS domain, and BA classification is used to mark incoming traffic on devices in a DS domain.
- Tags can be added to multiple types of data packet headers.
  - The Pri bit (802.1p priority) in the VLAN header is used to mark the QoS priority.
  - The EXP bit in the MPLS header is used to mark the QoS priority.
  - The TOS bit (DSCP/IP precedence) in the IP header is used to mark the QoS priority.

Contents



- This course describes two rate limiting technologies: traffic policing and traffic shaping.
- Traffic policing: If the traffic rate of a connection exceeds the specifications on an interface, traffic policing allows the interface to drop excess packets or re-mark the packet priority to protect network resources and protect carriers' profits. An example of this process is restricting the rate of HTTP packets to 50% of the network bandwidth.
- Traffic shaping: allows the traffic rate to match that on the downstream device. When traffic is transmitted from a high-speed link to a low-speed link or a traffic burst occurs, the inbound interface of the low-speed link is prone to severe data loss. To prevent this problem, traffic shaping must be configured on the outbound interface of the device connecting to the high-speed link.

Traffic Limiting Technology



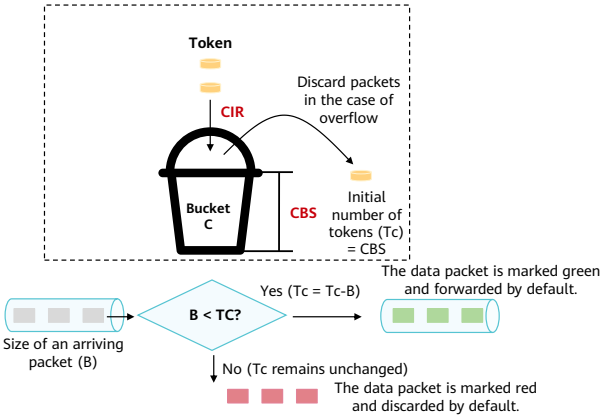
HUAWEI

- Both traffic policing and traffic shaping use the token bucket technology.
  - Token bucket: A token bucket is used to check whether traffic meets packet forwarding conditions.

Single-Rate-Single-Bucket Mechanism

- Committed Information Rate (CIR):
  - indicates the rate at which tokens are put into bucket C, in kbit/s.
- Committed burst size (CBS):
  - indicates the maximum volume of burst traffic that bucket C allows before the rate of some traffic exceeds the CIR, that is, the capacity of bucket C. The value is expressed in bytes.

- The single-rate-single-bucket mechanism does not allow burst traffic. Only committed traffic is allowed.

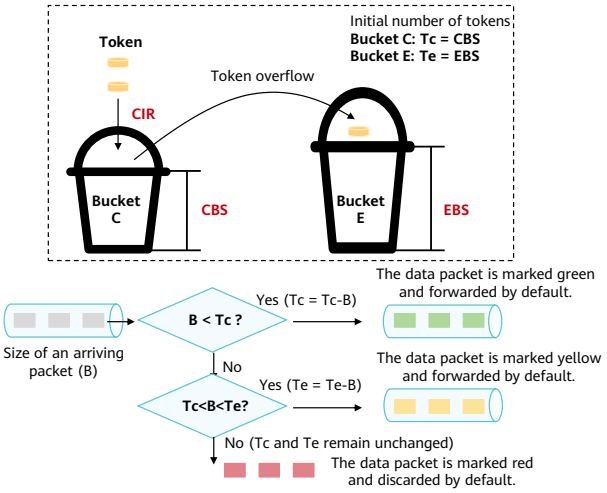


HUAWEI

- When a packet arrives, the device compares the packet with the number of tokens in the bucket. If there are sufficient tokens, the packet is forwarded (one token is associated with 1-bit forwarding permission). If there are no enough tokens, the packet is discarded or buffered.
- Tc and Te refer to the numbers of tokens in buckets C and E, respectively. The initial values of Tc and Te are respectively the CBS and EBS.
- In color-blind mode (B indicates the size of an arriving packet):
  - If B is less than or equal to Tc, the packet is marked green, and Tc decreases by B.
  - If B is greater than Tc, the packet is marked red, and Tc remains unchanged.

Single-Rate-Two-Bucket Mechanism

- CIR:
  - Indicates the rate at which tokens are put into bucket C, in kbit/s.
- CBS:
  - Indicates the maximum volume of burst traffic that bucket C allows before the rate of some traffic exceeds the CIR, that is, the capacity of bucket C. The value is expressed in bytes.
- Excess burst size (EBS):
  - Indicates the maximum volume of excess burst traffic that bucket E allows. The value is expressed in bytes.
- The single-rate-two-bucket mechanism allows transient burst traffic.



HUAWEI

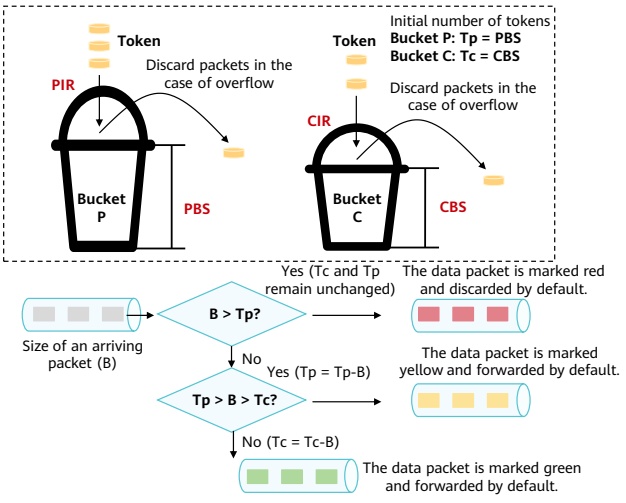
- When a packet arrives, the device compares the packet with the number of tokens in the bucket. If there are sufficient tokens, the packet is forwarded (one token is associated with 1-bit forwarding permission). If there are no enough tokens, the packet is discarded or buffered.
- Tc and Te refer to the numbers of tokens in buckets C and E, respectively. The initial values of Tc and Te are respectively the CBS and EBS.
- In color-blind mode (B indicates the size of an arriving packet):
  - If B is less than or equal to Tc, the packet is marked green and Tc decreases by B.
  - If B is greater than Tc and less than or equal to Te, the packet is marked yellow and Te decreases by B.
  - If B is greater than Te, the packet is marked red, and Tc and Te remain unchanged.

Two-Rate-Two-Bucket Mechanism

- Peak Information Rate (PIR):
  - Indicates the rate at which tokens are put into bucket P, that is, the maximum traffic rate that bucket P allows. The PIR is greater than the CIR. The value is expressed in kbit/s.
- Peak burst size (PBS):
  - Indicates the capacity of bucket P, that is, the maximum volume of burst traffic that bucket P allows. The PBS is greater than the CBS. The value is expressed in bytes.

- CIR:
  - Indicates the rate at which tokens are put into bucket C, in kbit/s.
- CBS:
  - Indicates the maximum volume of burst traffic that bucket C allows before the rate of some traffic exceeds the CIR, that is, the capacity of bucket C. The value is expressed in bytes.

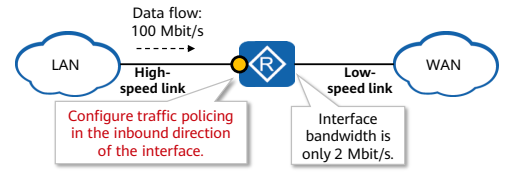
- The two-rate-two-bucket mechanism allows long-term burst traffic.



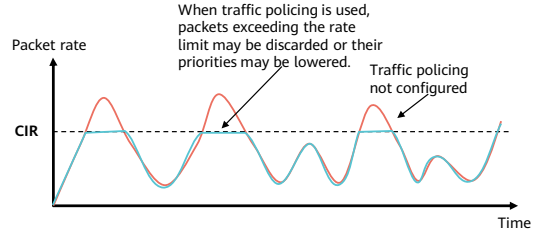
HUAWEI

- The two rate three color marker (trTCM) algorithm focuses on the traffic burst rate and checks whether the traffic rate conforms to the specifications. Therefore, traffic is measured based on bucket P and then bucket C.
- Tc and Tp refer to the number of tokens in buckets C and P, respectively. The initial values of Tc and Tp are respectively the CBS and PBS.
- In color-blind mode (B indicates the size of an arriving packet):
  - If B is greater than Tp, the packet is marked red and Tc and Tp remain unchanged.
  - If B is greater than Tc and less than or equal to Tp, the packet is marked yellow and Tp decreases by B.
  - If B is less than or equal to Tc, the packet is marked green, and Tp and Tc decrease by B.

What Is Traffic Policing?



- **Traffic policing**
- By monitoring the specifications of a certain type of traffic that enters the network, you can limit the traffic within an allowed range. If the traffic of a connection is too heavy, the packets are discarded or the priority of the packets is re-set to protect network resources. Traffic policing can be configured in inbound and outbound directions of an interface.



- **Implementation of traffic policing**
- Traffic policing uses the committed access rate (CAR) to control traffic. CAR uses the token bucket algorithm to evaluate the traffic rate and implements preset policing actions based on the evaluation result.

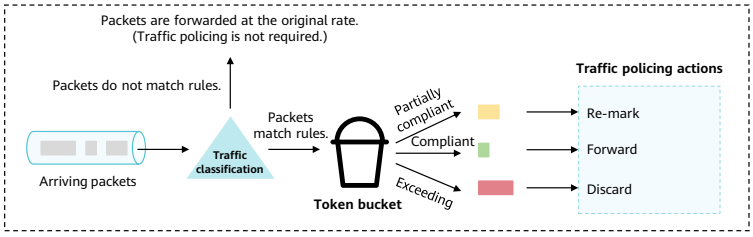
HUAWEI

- In the figure:
  - An edge network device connects a wide area network (WAN) and a local area network (LAN). The LAN bandwidth (100 Mbit/s) is higher than the WAN bandwidth (2 Mbit/s).
  - When a LAN user attempts to send a large amount of data to a WAN, the edge network device is prone to traffic congestion. Traffic policing can be configured on the edge network device to restrict the traffic rate, preventing traffic congestion.
- Characteristics of traffic policing:
  - Drops excess traffic over the specifications or re-marks such traffic with a lower priority.
  - Consumes no additional memory resources and brings no delay or jitter.
  - Packet loss may result in packet retransmission.
  - Traffic can be re-marked.

Traffic Limiting | Token Bucket | Traffic Policing | Traffic Shaping

CAR

- CAR uses token buckets to measure traffic and determines whether a packet conforms to the specification.



- **Token bucket modes**
  1. Single-rate-single-bucket
  2. Single-rate-two-bucket
  3. Two-rate-two-bucket

- **The device marks the packet red, yellow, or green based on the metering result using the token bucket.**
  1. Green indicates that the packets comply with the specifications and are directly forwarded.
  2. Yellow indicates that temporary burst traffic is allowed although it does not comply with specifications. After the traffic is re-marked, the priority is reduced and the traffic is forwarded in BE mode.
  3. Red indicates that the packet rate is high and does not comply with the specifications. Therefore, the packets are discarded.

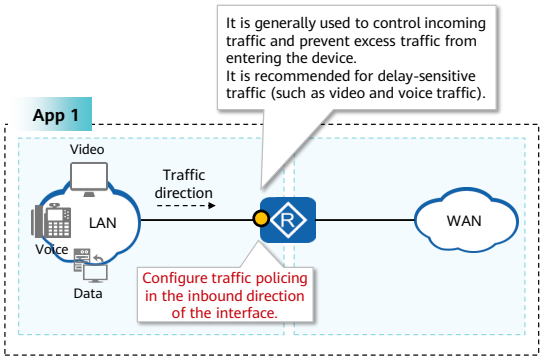
HUAWEI

- Traffic policing uses CAR to control traffic. CAR uses token buckets to measure traffic and determines whether a packet conforms to the specification.
- CAR has the following two functions:
  - Rate limiting: Only packets allocated enough tokens are allowed to pass so that the traffic rate is restricted.
  - Traffic classification: Packets are marked internal priorities, such as the service class and drop priority, based on the measurement performed by token buckets.
- CAR process:
  - When a packet arrives, the device matches the packet against matching rules. If the packet matches a rule, the device uses token buckets to meter the traffic rate.
  - The device marks the packet red, yellow, or green based on the metering result using the token bucket. Red indicates that the traffic rate exceeds the specifications. Yellow indicates that the traffic rate exceeds the specifications but is within an allowed range. Green indicates that the traffic rate is conforming to the specifications.
  - The device drops packets marked red, re-marks and forwards packets marked yellow, and forwards packets marked green.

- Three token bucket modes can be used.
  - To control the traffic rate, use single-rate-single-bucket.
  - To differentiate traffic bursts at limited bandwidth, use the single-rate-two-bucket mechanism. Note that traffic marked yellow must be processed differently from traffic marked green. Otherwise, the implementation of the single-rate-two-bucket mechanism is the same as that of the single-rate-single-bucket mechanism.
  - To control the traffic rate and check whether the traffic rate exceeds the CIR or PIR, use two-rate-two-bucket. Note that traffic marked yellow must be processed differently from traffic marked green. Otherwise, the implementation of the two-rate-two-bucket mechanism is the same as that of the single-rate-single-bucket mechanism.

Application Scenario of Traffic Policing

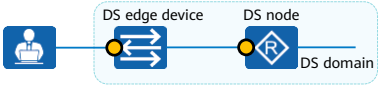
- **Application of traffic policing**
- Interface-based traffic policing
- Class-based traffic policing



HUAWEI

- Voice, video, and data services are transmitted on an enterprise network. When a large amount of traffic enters the network, congestion may occur due to insufficient bandwidth. Different guaranteed bandwidth must be provided for the voice, video, and data services in descending order of priority. In this situation, traffic policing can be configured to provide the highest guaranteed bandwidth for voice packets and lowest guaranteed bandwidth for data packets. This configuration ensures preferential transmission of voice packets during congestion.
- Interface-based traffic policing
- Interface-based traffic policing controls all traffic that enters an interface and does not identify the packet types. (based on the interface)
- Class-based traffic policing
- Class-based traffic policing controls the rate of one or more types of packets that enter an interface. (based on traffic classification)

Configuring Interface-based Traffic Policing

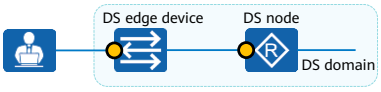


- Typically, traffic policing is performed in the inbound direction of a device. Traffic policing can be deployed on the terminal side or in the inbound direction of an egress device as required. Traffic policing can be configured based on interfaces or MQC.
- The configuration roadmap of interface-based traffic policing is as follows:
  - Set the maximum bandwidth for traffic policing on an interface, select the traffic to be policed, and adjust the behavior to be taken on the excess traffic.

- Configure interface-based traffic policing.

system-view
  interface [interface-type interface-num] //Enter the interface view.
    qos car [inbound | outbound] [acl acl-number | destination-ip-address | source-ip-address] **cir** [cir-value] [**pir** pir-value] [**cbs** cbs-value **pbs** pbs-value] //Configure traffic policing for specific traffic in the inbound or outbound direction of an interface. The CIR must be configured. The CIR indicates the maximum committed rate of traffic policing. If the PIR is not configured, it is equal to the CIR. In this case, the traffic rate cannot be higher than the CIR.

Configuring MQC-based Traffic Policing



- MQC can be used to implement traffic policing in a more refined manner. The device can control traffic and allocate different bandwidths based on the 5-tuple and QoS value of the data packet.
- The configuration roadmap is as follows:
  - Configure a traffic classifier to match traffic.
  - Configure a traffic behavior to define actions taken for packets.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to an interface.

```
system-view
  traffic classifier [classifier-name]        //Create a traffic classifier.
    if-match [acl | vlan-id | .... ]        //Match traffic based on traffic characteristics.
```

```
system-view
  traffic behavior [behavior-name]        //Create a traffic behavior.
    car cir [cir-value] [ pir pir-value ] [ cbs cbs-value pbs pbs-value ]     //Set the CIR and PIR.
```

```
system-view
  traffic policy [policy-name]     //Create a traffic policy.
    classifier [classifier-name] behavior [behavior-name]
//Bind the traffic classifier to the traffic behavior.
```

```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    traffic-policy [policy-name] [inbound | outbound]     //Apply the traffic policy to the inbound direction of an interface.
```

Checking the Traffic Policing Configuration

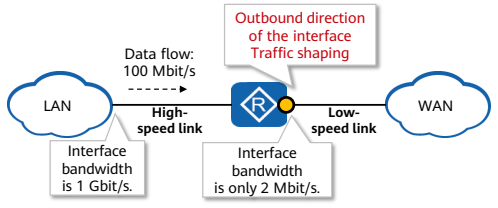
- After interface-based traffic policing is configured, you can run the following commands to check the configuration.

```
system-view
  display qos car statistics interface [interface-type interface-num] [inbound | outbound]    //Check statistics on forwarded and discarded packets on the interface.
```

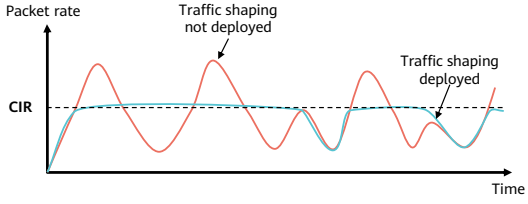
- After MQC-based traffic policing is configured, you can run the following commands to check the configuration.

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]    //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

What Is Traffic Shaping?



- **Traffic shaping**
- Traffic shaping is a measure to adjust the traffic rate sent from an interface.
- Traffic shaping is configured on the outbound interface of an upstream device so that irregular traffic can be transmitted at an even rate, preventing transient traffic congestion on the downstream device.



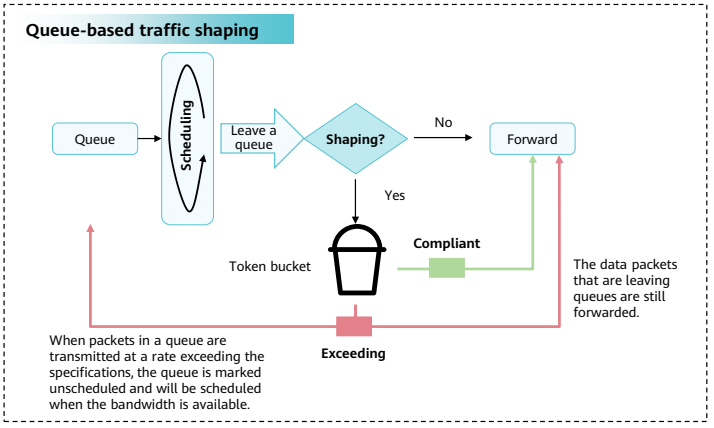
- **Implementation of traffic shaping**
- Traffic shaping is implemented using the buffer and token bucket.

- **Token bucket mode**: single-rate-single-bucket
- **Assessment result**: compliant (green), non-compliant (red)

HUAWEI

- Generic Traffic Shaping (GTS)
- When traffic is transmitted from a high-speed link to a low-speed link or a traffic burst occurs, the inbound interface of the low-speed link is prone to severe data loss. To prevent this problem, configure traffic shaping on the outbound interface of the device connecting to the high-speed link.
- When packets are sent at a high speed, they are cached and then evenly sent through the token bucket.
- Characteristics of traffic shaping:
  - Buffers excess traffic over the specifications.
  - Consumes memory resources for buffering excess traffic and brings delay and jitter.
  - Packet loss rarely occurs, so packets are seldom retransmitted.
  - Traffic re-marking is not supported.
- Token bucket mode: single-rate-single-bucket — The evaluation result can be either green or red.

Implementation of Traffic Shaping (1)

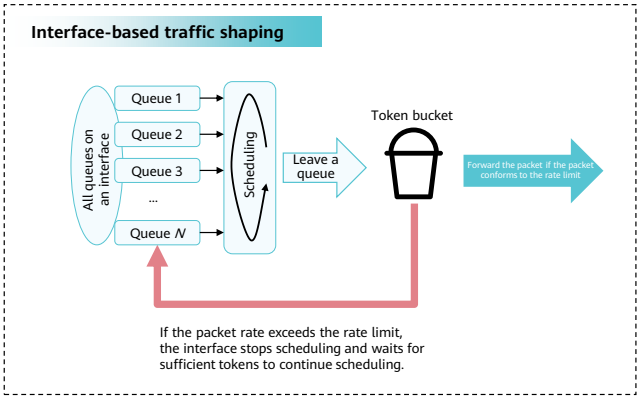


- **Queue-based traffic shaping**
- Applies to each queue on an outbound interface.

HUAWEI

- When packets leave queues, the packets that do not need to be shaped are forwarded. The packets that need to be shaped are measured against token buckets.
  - If the packet rate conforms to the rate limit, the packet is marked green and forwarded.
  - If the rate of a data packet exceeds the threshold, the data packet is still forwarded. In this case, the status of the queue where the data packet is located is changed to unscheduled, and the queue is scheduled when the token bucket is filled with new tokens. After the queue is marked unscheduled, more packets can be put into the queue, but excess packets over the queue capacity are dropped. Therefore, traffic shaping allows traffic to be sent at an even rate but does not provide zero-packet-loss guarantee.

Implementation of Traffic Shaping (2)

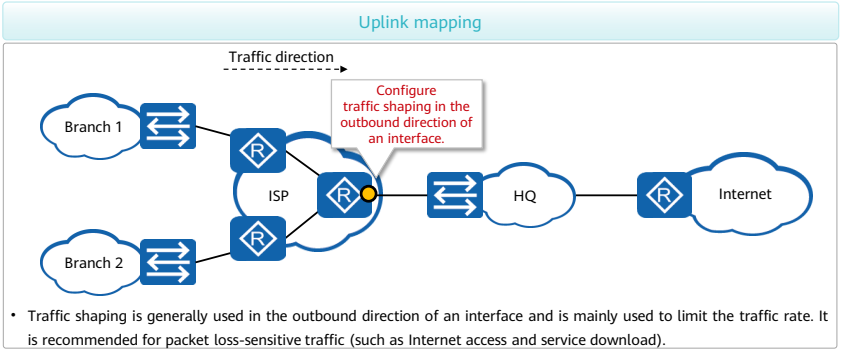


- **Interface-based traffic shaping**
- Limits the total rate of all packets sent by an interface. Traffic shaping is performed on the outbound interface, regardless of packet priorities.

- When packets leave queues, all queues are measured together against token buckets.
  - If the packet rate conforms to the rate limit, the packet is marked green and forwarded.
  - If the packet rate exceeds the threshold (that is, tokens in the token bucket are insufficient), the packet is marked red. In this case, the interface stops scheduling and continues to schedule the packets when there are sufficient tokens.

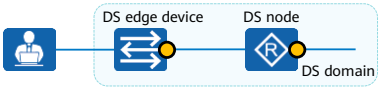
Application Scenario of Traffic Shaping

- On an enterprise network, the enterprise headquarters is connected to branches through private lines on an ISP network. Branches connect to the Internet through the headquarters.
- If all branches connect to the Internet at the same time, a large amount of web traffic sent from the headquarters to the Internet causes network congestion. As a result, some web traffic is discarded.. As shown in the figure, to prevent web traffic loss, traffic shaping can be configured before traffic sent from enterprise branches enters the enterprise headquarters.



- Traffic shaping is generally used in the outbound direction of an interface and is mainly used to limit the traffic rate. It is recommended for packet loss-sensitive traffic (such as Internet access and service download).

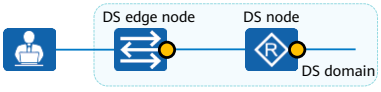
Configuring Interface-based Traffic Shaping



- Traffic shaping can be configured only in the outbound direction of a device. It falls into interface-based, queue-based, and MQC-based traffic shaping.
- Interface-based traffic shaping has a large granularity. The configuration roadmap is as follows:
  - Deploy traffic shaping in the outbound direction of an interface and configure the maximum bandwidth.
- Configure interface-based traffic shaping.

system-view
  interface [interface-type interface-num] //Enter the interface view.
    qos gts cir [cir-value] [cbs cbs-value] //Configure traffic shaping in the outbound direction of an interface. The CIR indicates the maximum traffic shaping rate. You can configure the CBS as required to control the size of the token bucket. The CIR must be configured.

Configuring Queue-based Traffic Shaping



- To shape packets in each queue on an interface, configure a queue profile and apply it to the interface.
- You can set different traffic shaping parameters for queues with different priorities to provide differentiated services. The configuration roadmap is as follows:
  - Create a queue profile.
  - Configure queue shaping.
  - Apply the queue profile to an interface.

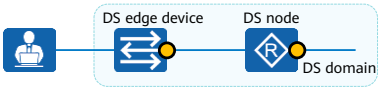
- Create a queue profile and configure queue shaping.

```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    qos queue-profile [queue-profile-name]   //Create a queue profile.
      queue [start-queue-index] to [end-queue-index ] gts cir [cir-value] [ cbs cbs-value ]      //Configure traffic shaping for a specified queue in the outbound direction and set the CIR.
```

- Apply the queue profile to an interface.

```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    qos queue-profile [queue-profile-name]     //Apply the queue profile to the interface.
```

Configuring MQC-based Traffic Shaping



- MQC-based traffic policing uses traffic classifiers to implement differentiated services.
- The configuration roadmap is as follows:
  - Configure a traffic classifier to match traffic.
  - Configure a traffic behavior to define an action for packets.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to an interface in the outbound direction.

system-view
  traffic classifier [classifier-name] //Create a traffic classifier.
    if-match [acl | vlan-id | ....] //Match traffic based on traffic characteristics.

system-view
  traffic behavior [behavior-name] //Create a traffic behavior.
    gts cir [cir-value] | **pct** [pct-value] //Configure traffic shaping based on the maximum traffic rate or the percentage of the occupied interface bandwidth.

system-view
  traffic policy [policy-name] //Create a traffic policy.
    classifier [classifier-name] **behavior** [behavior-name] //Bind the traffic classifier to the traffic behavior.

system-view
  interface [interface-type interface-num] //Enter the interface view.
    traffic-policy [policy-name] [inbound | outbound] //Apply the traffic policy to the interface in the outbound direction.

Checking the Traffic Shaping Configuration

- After queue-based traffic shaping is configured, you can run the following commands to check the configuration.

```
system-view
  display qos queue-profile [ queue-profile-name ]   //Check the queue profile configuration.
```

- After MQC-based traffic shaping is configured, you can run the following commands to check the configuration.

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]   //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

Quiz

1. (True or false) Traffic shaping caches excess traffic by default, and traffic policing discards excess traffic by default.
   A. True
   B. False
2. (Multiple-answer question) How many modes of token buckets are used to measure traffic?
   A. Single-rate-single-bucket
   B. Three-rate-two-bucket
   C. Single-rate-two-bucket
   D. Two-rate-two-bucket

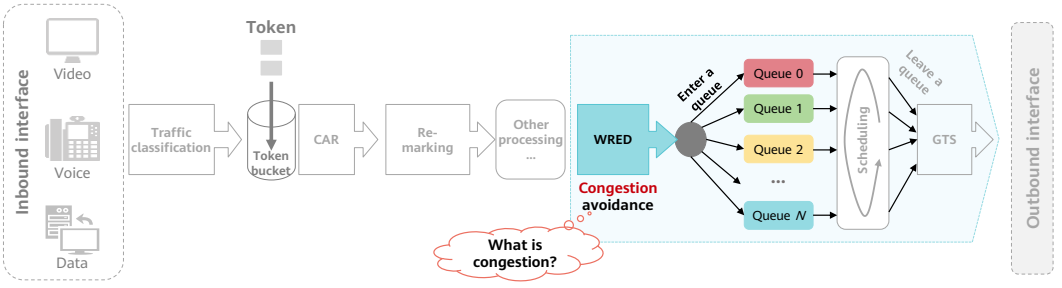
- 1. A
- 2. ACD

Section Summary

- There are two traffic limiting technologies: traffic policing and traffic shaping.
- Traffic policing discards excess traffic by default. It can be deployed in inbound and outbound directions of a device.
- Traffic shaping caches excess traffic by default. It can be deployed only in the outbound direction of a device.
- The device uses token buckets to measure traffic. There are three modes of token buckets:
  - The single-rate-single-bucket mechanism can be used together with traffic policing and traffic shaping.
  - The single-rate-two-bucket mechanism can be used only with traffic policing, and is mainly used in scenarios where burst traffic occurs occasionally.
  - The two-rate-two-bucket can be used only with traffic policing, and is mainly used in scenarios with long-term burst traffic.

Contents

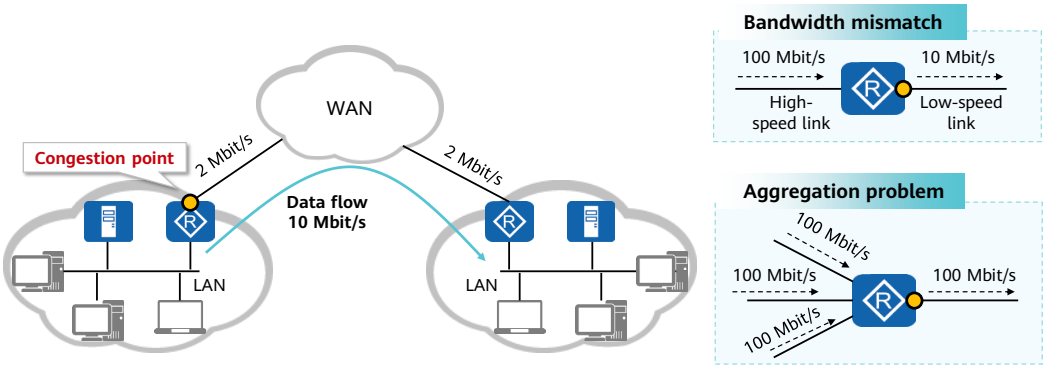
QoS Data Processing



HUAWEI

- Congestion avoidance: Excessive congestion may damage network resources. Congestion avoidance monitors the usage of network resources. When congestion aggravates, congestion avoidance adjusts traffic to relieve network overload by discarding packets. Congestion avoidance is generally applied to the outbound direction of an interface.

Background of Congestion Occurrence



HUAWEI

- Traffic congestion occurs when multiple users compete for the same resources (such as the bandwidth and buffer) on the shared network.
  - For example, a user on a LAN sends data to a user on another LAN through a WAN. The WAN bandwidth is lower than the LAN bandwidth. Therefore, data cannot be transmitted at the same rate on the WAN as that on the LAN. Traffic congestion occurs on the router connecting the LAN and WAN.
- Congestion often occurs in the following situations:
  - Traffic rate mismatch: Packets are transmitted to a device through a high-speed link and are forwarded out through a low-speed link.
  - Traffic aggregation: Packets are transmitted from multiple interfaces to a device and are forwarded out through a single interface without enough bandwidth.

Impact of Congestion

- **Traffic congestion has the following adverse impacts on network traffic:**
  1. Traffic congestion intensifies delay and jitter.
  2. Overlong delays lead to packet retransmission.
  3. Traffic congestion lowers the network throughput and damages network resources.
  4. Intensified traffic congestion consumes a large number of network resources (especially storage resources). Unreasonable resource allocation may cause resources to be locked and the system to break down.

Solution

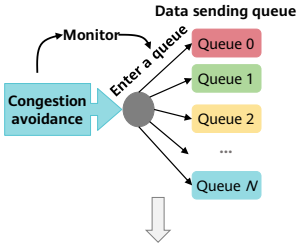
1. Congestion avoidance
2. Congestion management

HUAWEI

- Impact of congestion:
  - Congestion prevents traffic from obtaining resources immediately, which causes service deterioration. However, congestion often occurs in a complex networking environment where packet transmission and provisioning of various services are both required. Therefore, effective methods are required to avoid congestion or prevent congestion from aggravating.
- Solutions:
  - The solutions need to make full use of network resources on the premise of meeting users' requirements for service quality. Congestion management and congestion avoidance are commonly used to relieve traffic congestion.
  - Congestion management provides means to manage and control traffic when traffic congestion occurs.
  - Congestion avoidance is a flow control technique used to relieve network overload. By monitoring the usage of network resources in queues or memory buffer, a device automatically drops packets on the interface that shows a sign of traffic congestion. Congestion avoidance prevents queues from being overflowed due to network overload. The following will introduce congestion avoidance technology.

Congestion Avoidance Technology

- Congestion avoidance is a flow control technique used to relieve network overload. By monitoring the usage of network resources for queues or memory buffers, a device automatically drops packets that shows a sign of traffic congestion.

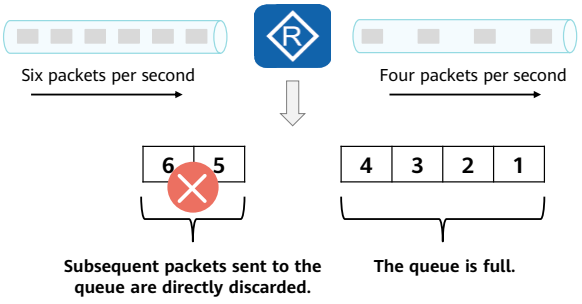


1. Tail drop: traditional processing
2. Random Early Detection (RED)
3. Weighted Random Early Detection (WRED)

- If congestion becomes severe, the packets that do not enter queues are discarded.

Policy 1: Tail Drop

- When the length of a queue reaches the maximum value, the device enabled with tail drop discards all new packets buffered at the tail of the queue.

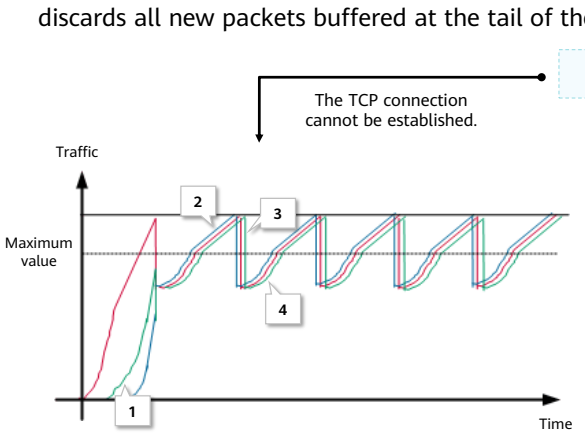


HUAWEI

- Due to the limited length of each queue, when a queue is full, the traditional processing method discards all the packets sent to the queue until the congestion is relieved. This processing method is called tail drop.

Disadvantage 1: Global TCP Synchronization

- When the length of a queue reaches the maximum value, the device enabled with tail drop discards all new packets buffered at the tail of the queue.



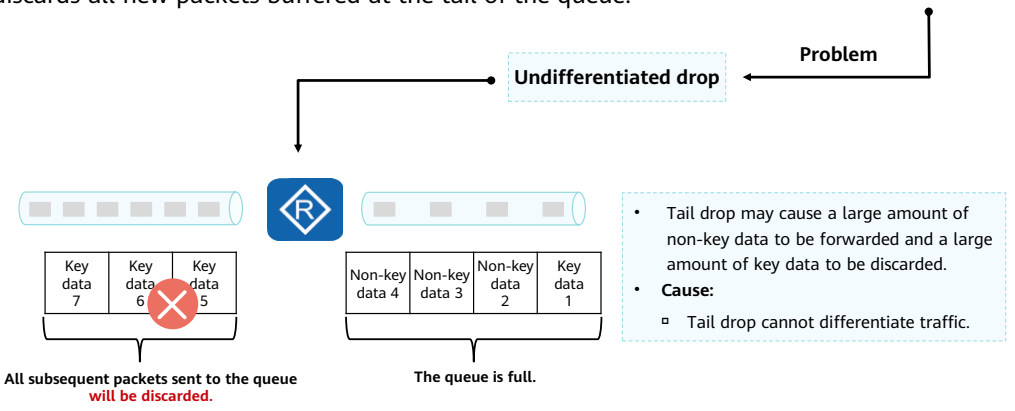
- **Process:**
  1. TCP starts.
  2. Traffic is too heavy. As a result, the queue is full and tail drop occurs.
  3. The TCP ACK packet returned by the server is discarded due to congestion. Therefore, the sender does not receive the TCP ACK packet and considers that the network is congested. In this case, the TCP sliding window size is reduced, and the overall traffic is also reduced.
  4. At this time, network congestion is eliminated, and the sender can receive the TCP ACK packet. Therefore, the sender considers that the network is not congested, and enters the TCP slow start process. This process is repeated.

HUAWEI

- As shown in the following figure, three colors indicate three TCP connections.
- Global TCP synchronization:
  - In tail drop mechanism, all newly arrived packets are dropped when congestion occurs, causing all TCP sessions to simultaneously enter the slow start state and the packet transmission to slow down.
  - When packets of multiple TCP connections are discarded in a queue, TCP connections enter the congestion avoidance and slow start state to adjust and reduce traffic. This is called TCP global synchronization. Then all TCP sessions restart their transmission at roughly the same time and then congestion occurs again, causing another burst of packet drops, and all TCP sessions enter the slow start state again. The behavior cycles constantly, severely reducing the network resource usage.

Disadvantage 2: Undifferentiated Drop

- When the length of a queue reaches the maximum value, the device enabled with tail drop discards all new packets buffered at the tail of the queue.



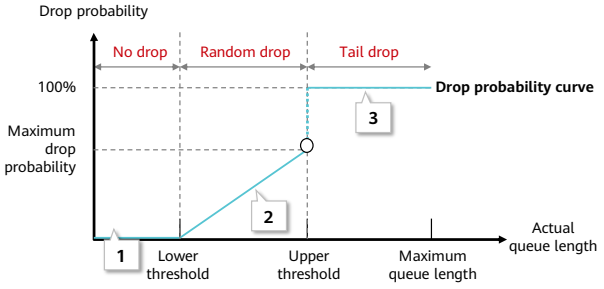
- Tail drop may cause a large amount of non-key data to be forwarded and a large amount of key data to be discarded.
- **Cause:**
  - Tail drop cannot differentiate traffic.

HUAWEI

- Tail drop cannot differentiate services and discard traffic in the same way.

Policy 2: RED

- Random early detection (RED) randomly discards data packets.



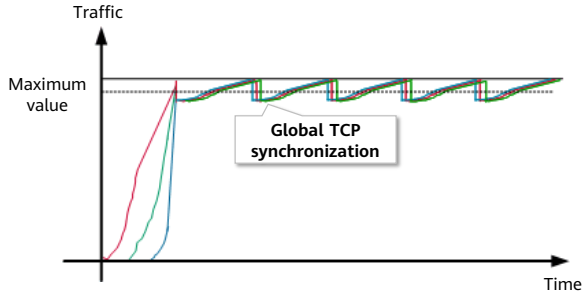
- **Process:**
  1. When the queue length is less than the lower threshold, no packets are discarded.
  2. When the queue length is between the upper threshold and the lower threshold, newly arrived packets are randomly discarded. The longer the queue is, the higher the drop probability is.
  3. All coming packets are discarded if the queue length is greater than the upper threshold.

HUAWEI

- RED defines upper and lower thresholds for the length of each queue:
  - When the queue length is less than the lower threshold, no packets are discarded.
  - When the queue length is greater than the upper drop threshold, all packets are discarded.
  - Coming packets are dropped randomly if the queue length is between upper and lower thresholds. RED generates a random number for each incoming packet and compares it with the drop probability of the current queue. If the random number is greater than the drop probability, the packet is discarded. A longer queue indicates a higher drop probability.

Relieving Global TCP Synchronization

- RED randomly discards packets so that rates of TCP connections are reduced at different times. This prevents global TCP synchronization.



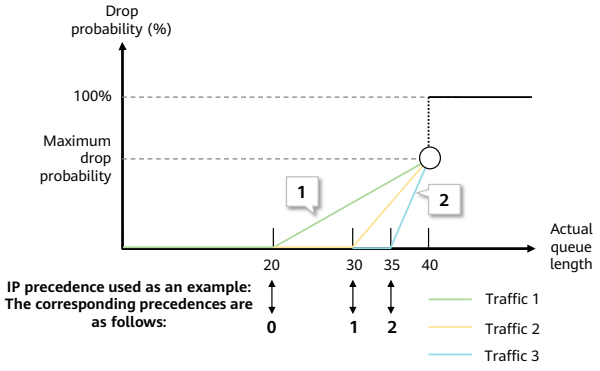
- **Symptom:**
- Global TCP synchronization may still occur, but the link usage is greatly increased.
- **Disadvantage:**
- RED cannot distinguish traffic.

HUAWEI

- RED is used to avoid global TCP synchronization that occurs with tail drop. It does this by randomly discarding packets so that the transmission speed of multiple TCP connections is not reduced simultaneously. This results in more stable rates of TCP traffic and other network traffic. — Do not adjust TCP sliding window sizes simultaneously.

Policy 3: WRED

- Weighted Random Early Detection (WRED) sets different drop policies for data packets or queues with different priorities to discard different types of traffic.

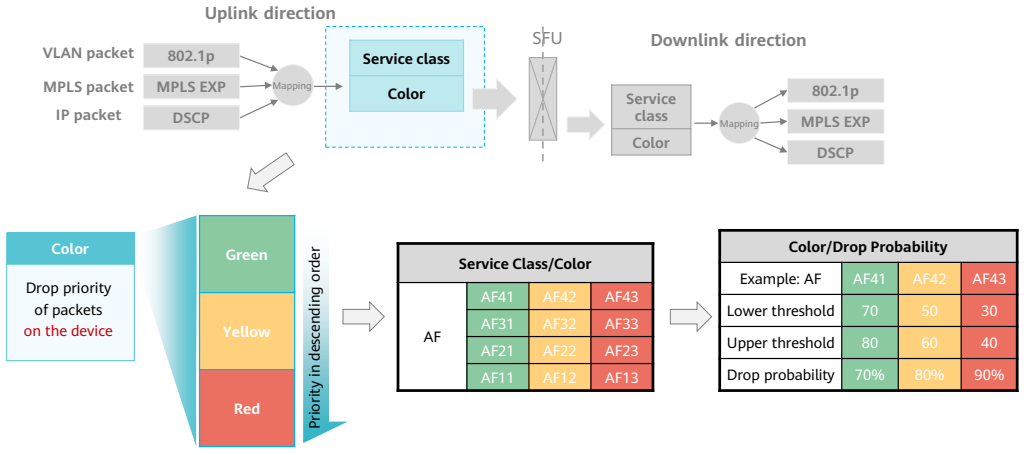


- **Example:**
  1. The lower threshold is 20 and the upper threshold is 40 for the traffic whose IP precedence is 0.
  2. The lower threshold is 35 and the upper threshold is 40 for the traffic whose IP precedence is 2. The traffic whose IP precedence is 2 is discarded later than the traffic whose IP precedence is 0.
- **Advantage:**
  1. Do not adjust TCP sliding window sizes simultaneously to avoid global TCP synchronization.
  2. Different traffic is discarded based on weights.

HUAWEI

- The device provides WRED based on RED technology.
- WRED discards packets in queues based on DSCP priorities or IP priorities. The upper and lower thresholds, and drop probability can be set for each priority. When the number of packets of a priority reaches the lower threshold, the device starts to discard packets. When the number of packets reaches the upper threshold, the device discards all the packets. As the queue length increases, the packet loss rate increases. The maximum packet loss rate does not exceed the preset packet loss rate. WRED discards packets in queues based on the drop probability, thereby preventing congestion to a certain degree. — Do not adjust TCP sliding window sizes simultaneously.

WRED Drop Priority

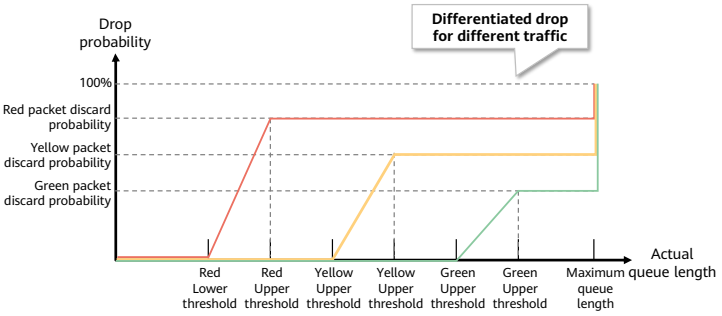


HUAWEI

- Color:
  - The color of packets determines the order in which packets are dropped in a congested queue.
- Application:
  - The WRED lower threshold is recommended to start from 50% and change with the drop priority. The lowest drop probability and highest lower and upper thresholds are recommended for green packets; a medium drop probability and medium lower and upper thresholds are recommended for yellow packets; the highest drop probability and smallest lower and upper thresholds are recommended for red packets.
  - When traffic congestion aggravates, red packets are first dropped due to the smallest lower threshold and high drop probability. As the queue length increases, the device drops green packets at last. If the queue length reaches the upper threshold for red/yellow/green packets, red/yellow/green packets start to be tail dropped.

Curve of the WRED Drop Probability

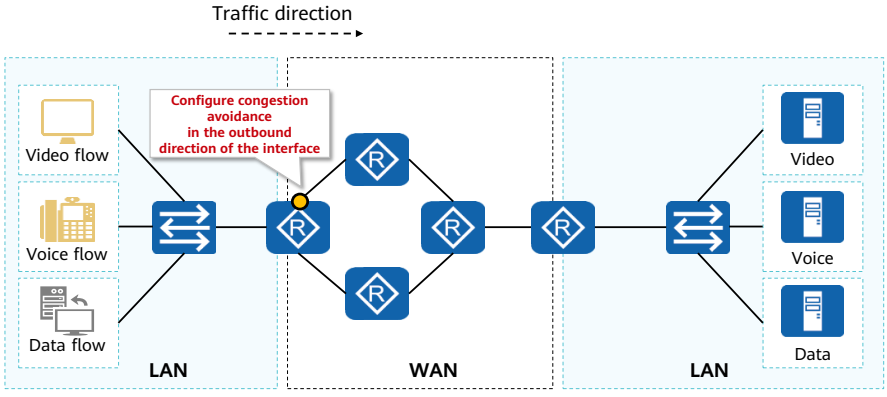
Color/Drop Probability			
Example: AF	AF41	AF42	AF43
Lower threshold	70	50	30
Upper threshold	80	60	40
Drop probability	70%	80%	90%



HUAWEI

- Color:
  - The color of packets determines the order in which packets are dropped in a congested queue.
- Application:
  - The WRED lower threshold is recommended to start from 50% and change with the drop priority. The lowest drop probability and highest lower and upper thresholds are recommended for green packets; a medium drop probability and medium lower and upper thresholds are recommended for yellow packets; the highest drop probability and smallest lower and upper thresholds are recommended for red packets.
  - When traffic congestion aggravates, red packets are first dropped due to the smallest lower threshold and high drop probability. As the queue length increases, the device drops green packets at last. If the queue length reaches the upper threshold for red/yellow/green packets, red/yellow/green packets start to be tail dropped.

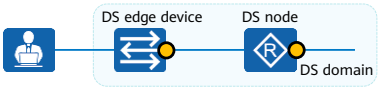
Application of Congestion Avoidance



HUAWEI

- Example:
  - Users in different LANs may upload data to the same server, so data exchanged between users and the server passes the WAN. Because WAN bandwidth is lower than LAN bandwidth, congestion may occur on the edge device between the WAN and LANs. Congestion avoidance can be configured on the edge device to discard low-priority packets such as data packets, reducing network overload and ensuring forwarding of high-priority services.

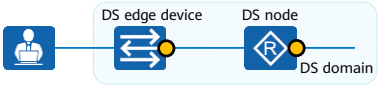
Configuring Queue-based WRED



- The device supports WRED based on DSCP priorities or IP priorities. The configuration roadmap is as follows:
  - Configure a drop profile.
  - Configure WRED parameters.
  - Reference the drop profile to a queue profile.
  - Apply the queue profile to the outbound direction of the interface.

system-view
drop-profile [drop-profile-name] //Create a drop profile.
wred [dscp | ip-precedence] //Configure a WRED drop profile based on DSCP or IP priorities.
dscp [dscp-value] **low-limit** [low-limit-percentage] **high-limit** [high-limit-percentage] **discard-percentage** [discard-percentage] //Configure WRED parameters based on DSCP priorities.
ip-precedence [ip-precedence-value] **low-limit** [low-limit-percentage] **high-limit** [high-limit-percentage] **discard-percentage** [discard-percentage] //(Optional) Configure WRED parameters based on IP priorities.
qos queue-profile [queue-profile-name] //Enter the queue profile view.
queue [queue-index] **drop-profile** [drop-profile-name] //Bind the drop profile to the specified queue in the queue profile.
interface [interface-type interface-num] //Enter the interface view.
qos queue-profile [queue-profile-name] //Apply the queue profile to the interface.

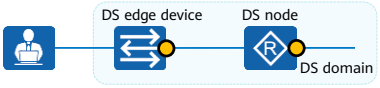
Configuring MQC to Implement Congestion Avoidance (1)



- After a drop profile is bound to a traffic behavior, associate the traffic behavior with the corresponding traffic classifier in the traffic policy and apply the traffic policy to an interface to implement congestion avoidance for traffic matching the traffic classifier. The configuration roadmap is as follows:
  - Configure a drop profile.
  - Configure a traffic classifier and a traffic behavior.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to the outbound direction of the device interface.

system-view
drop-profile [drop-profile-name] //Create a drop profile.
wred [dscp | ip-precedence] //Configure a WRED drop profile based on DSCP or IP priorities.
dscp [dscp-value] **low-limit** [low-limit-percentage] **high-limit** [high-limit-percentage] **discard-percentage** [discard-percentage] //Configure WRED parameters based on DSCP priorities.
ip-precedence [ip-precedence-value] **low-limit** [low-limit-percentage] **high-limit** [high-limit-percentage] **discard-percentage** [discard-percentage] //(Optional) Configure WRED parameters based on IP priorities.

Configuring MQC to Implement Congestion Avoidance (2)



- After a drop profile is bound to a traffic behavior, associate the traffic behavior with the corresponding traffic classifier in the traffic policy and apply the traffic policy to an interface to implement congestion avoidance for traffic matching the traffic classifier. The configuration roadmap is as follows:
  - Configure a drop profile.
  - Configure a traffic classifier and a traffic behavior.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to the outbound direction of the device interface.

system-view
  traffic classifier [classifier-name] //Create a traffic classifier.
    if-match [acl | vlan-id | ....] //Match traffic based on traffic characteristics.

system-view
  traffic behavior [behavior-name] //Create a traffic behavior.
    drop-profile [drop-profile-name] //Bind the created drop profile to the traffic behavior.

system-view
  traffic policy [policy-name] //Create a traffic policy.
    classifier [classifier-name] **behavior** [behavior-name] //Bind the traffic classifier to the traffic behavior.

system-view
  interface [interface-type interface-num] //Enter the interface view.
    traffic-policy [policy-name] **outbound** //Apply the traffic policy to the outbound direction of the interface.

Checking the Congestion Avoidance Configuration

- Checking the queue-based congestion avoidance configuration

```
system-view
  interface [interface-type interface-num]
    display this    //Check the queue profile bound to the interface.
  qos queue-profile [queue-profile-name]
    display this   //Check the drop profile bound to the queue profile.
  display drop-profile [ drop-profile-name ]   //Check the drop profile configuration.
```

- Checking the MQC-based congestion avoidance configuration

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]    //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

Quiz

1. (Multiple-answer question) Which of the following mechanisms are used by QoS to proactively discard packets?
   A. Tail drop
   B. RED
   C. MRED
   D. WRED

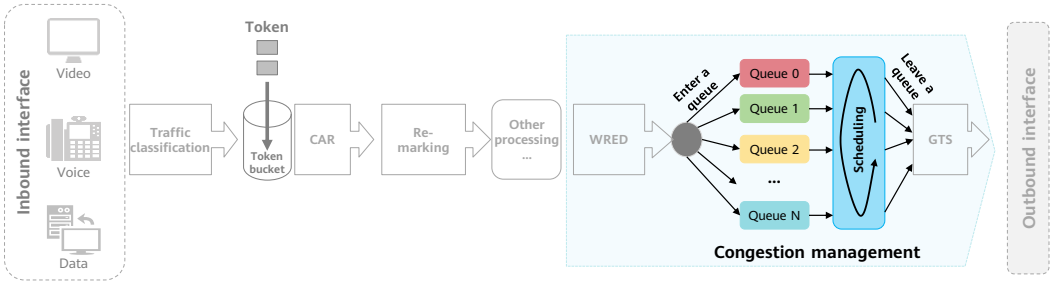
- 1. ABD

Section Summary

- Congestion avoidance technology cannot avoid congestion, but prevents problems caused by congestion. For example, tail drop causes global TCP synchronization, interface traffic is unstable, and UDP traffic preempts the bandwidth used by TCP traffic.
- RED/WRED randomly discards data packets to prevent problems such as global TCP synchronization.

Contents

QoS Data Processing

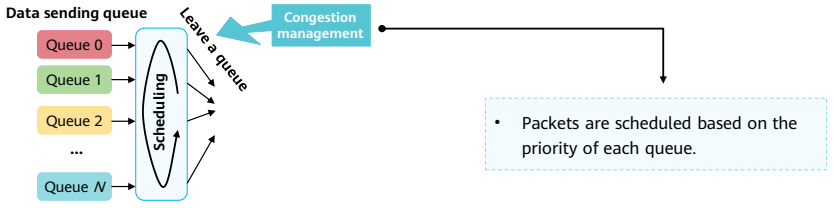


HUAWEI

- Congestion management: It is a measure that must be taken to solve the problem of resource competition. Packets are buffered in queues and a scheduling algorithm is used to determine the forwarding sequence of packets. Congestion management is usually applied to the outbound direction of an interface.

Congestion Management Technology

- Congestion management technology manages and controls different types of service traffic when network congestion occurs.
- It uses queue scheduling technology to handle traffic congestion.

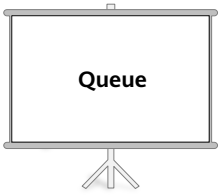


HUAWEI

- Congestion Management Technology
- Congestion management defines a policy that determines the order in which packets are forwarded and specifies drop principles for packets. The queuing technology is used.
- The queue scheduling algorithm determines the order in which packets are leaving a queue and the relationships between queues.
- Queuing technology
- Packets sent from one interface are placed into many queues which are identified with different priorities. The packets are then sent based on the priorities. Different queue scheduling mechanisms are designed for different situations and lead to varying results.

What Is a Queue?

- The queuing technology orders packets in the buffer.



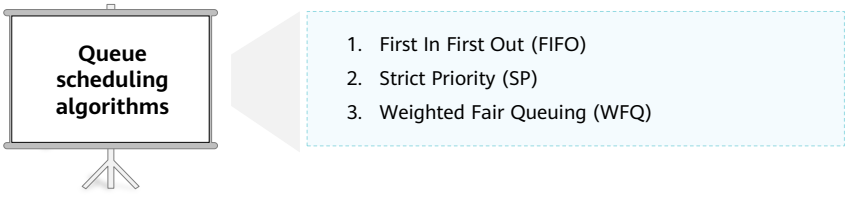
- Each interface has eight downlink queues, which are called class queues (CQs) or port queues.
- They are EF, AF1, AF2, AF3, AF4, BE, CS6, and CS7.

HUAWEI

- What is a queue?
- The queuing technology orders packets in the buffer. When the packet rate exceeds the interface bandwidth or the bandwidth configured for packets, the packets are buffered in queues and wait to be forwarded.
- Each interface on the NE20E or NE40E stores eight downlink queues, which are called CQs or port queues. The eight queues are BE, AF1, AF2, AF3, AF4, EF, CS6, and CS7.

Queue Scheduling Algorithms

- Congestion management uses the queuing technology.

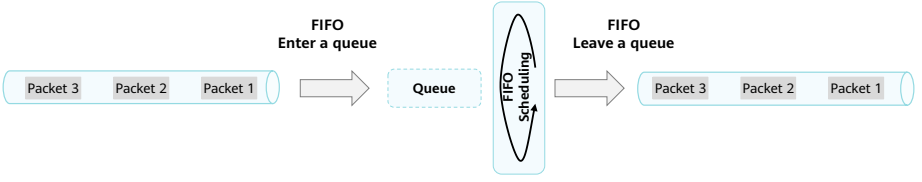


HUAWEI

- Queuing technology places packets sent from one interface into multiple queues with different priorities. These packets are then sent based on the priorities. Different queue scheduling mechanisms are designed for different situations and lead to varying results.

FIFO

- The FIFO mechanism is used to transfer packets in a queue. Resources used to forward packets are allocated based on the arrival order of packets.

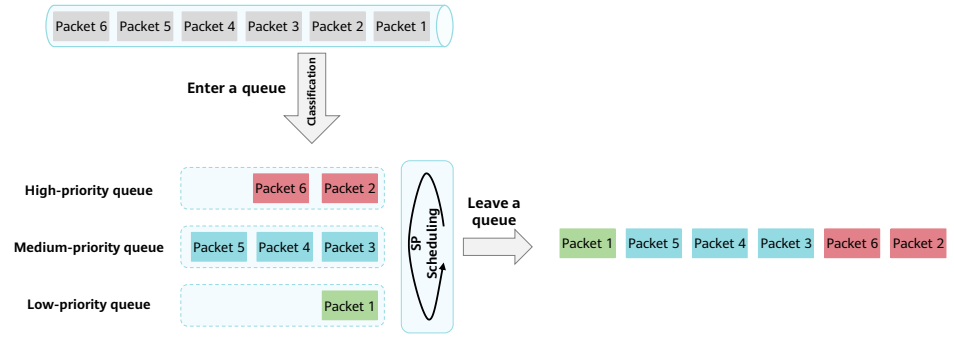


HUAWEI

- FIFO does not classify packets.
- FIFO allows the packets that come earlier to enter the queue first. On the exit of a queue, FIFO allows the packets to leave the queue in the same order as that in which the packets enter the queue.
- Characteristics:
  - Advantage: The implementation mechanism is simple and the processing speed is fast.
  - Disadvantage: Packets with different priorities cannot be processed in differentiated ways.

SP

- SP schedules packets strictly based on queue priorities.

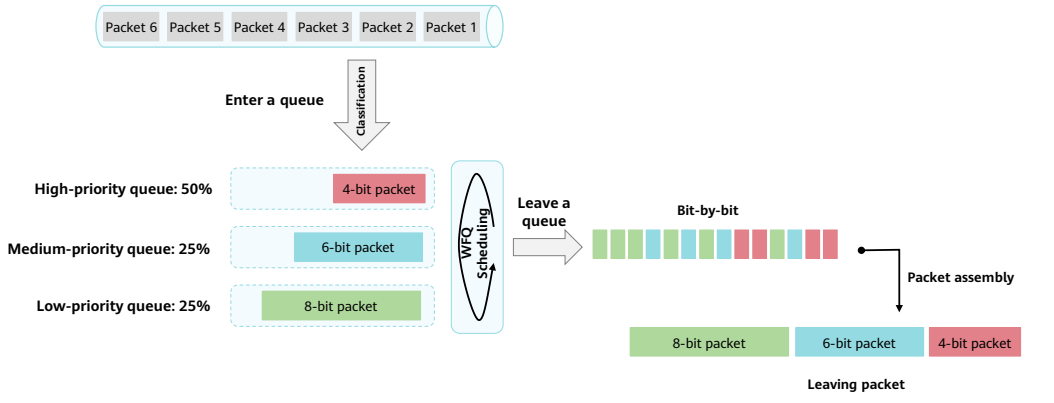


HUAWEI

- SP: Packets in queues with a low priority can be scheduled only after all packets in queues with a higher priority are scheduled.
- As shown in the figure, three queues with a high, medium, and low priorities respectively are configured with SP scheduling. The number indicates the order in which packets arrive.
- When packets leave queues, the device forwards the packets in descending order of priority. Packets in the higher-priority queue are forwarded preferentially. If packets in the higher-priority queue come in between packets in the lower-priority queue that is being scheduled, the packets in the high-priority queue are still scheduled preferentially. This implementation ensures that packets in the higher-priority queue are always forwarded preferentially. As long as there are packets in the high-priority queue, no other queue will be served.
- Characteristics:
  - Advantage: High-priority packets are preferentially forwarded.
  - Disadvantage: Low-priority queues may be starved out. That is, when congestion occurs, packets in lower-priority queues are not processed until all the higher-priority queues are empty. As a result, a congested higher-priority queue causes all lower-priority queues to starve out.

WFQ

- WFQ allocates outbound bandwidth to flows on an interface based on weights of queues.



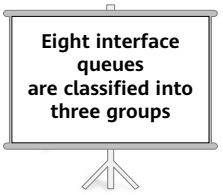
HUAWEI

- WFQ allocates bandwidths to flows based on weights of queues. In addition, to fairly allocate bandwidths to flows, WFQ schedules packets bit by bit.
- When packets leave queues, the device forwards the packets in descending order of priority. Packets in the higher-priority queue are forwarded preferentially. If packets in the higher-priority queue come in between packets in the lower-priority queue that is being scheduled, the packets in the high-priority queue are still scheduled preferentially. This implementation ensures that packets in the higher-priority queue are always forwarded preferentially. As long as there are packets in the high-priority queue, no other queue will be served.
- Characteristics:
  - Advantages:
    - Packets in different queues are scheduled fairly, and the flow delays have slight differences.
    - If many large and small packets in different queues need to be sent, small packets are scheduled first, reducing the total jitter of each flow.
    - The smaller the weight, the less the allocated bandwidth. Flows with larger weights are allocated higher bandwidth.
  - Disadvantage: Low-latency services cannot be scheduled in a timely manner. User-defined classification rules cannot be implemented.

- The bit-by-bit scheduling mode, however, is an ideal one. The NE40E performs WFQ scheduling based on a certain granularity, such as 256 bytes and 1 Kbytes. Different cards support different granularities.

Queue Scheduling Mode of an Interface

- You can configure SP scheduling or WFQ scheduling for eight queues on an interface.
- Eight queues can be classified into three groups, priority queuing (PQ) queues, WFQ queues, and low priority queuing (LPQ) queues, based on scheduling algorithms.

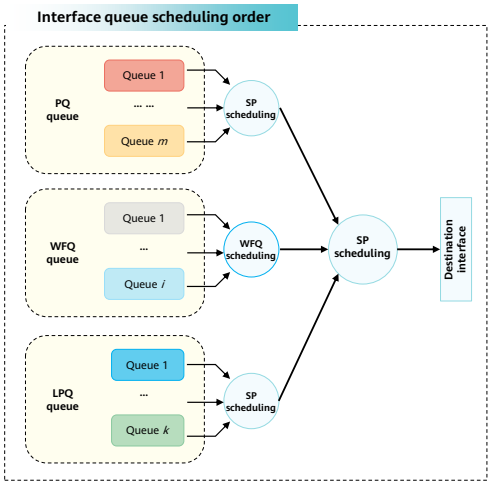


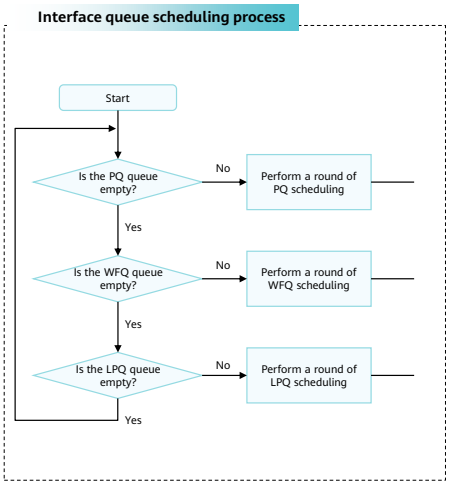
1. PQ queue: uses the SP scheduling algorithm.
2. WFQ queue: uses the WFQ scheduling algorithm.
3. LPQ queue: uses the SP scheduling algorithm.

HUAWEI

- PQ queue
  - SP scheduling applies to PQ queues. Packets in high-priority queues are scheduled preferentially. Therefore, services that are sensitive to delays (such as VoIP) can be configured with high priorities.
  - In PQ queues, however, if the bandwidth of high-priority packets is not restricted, low-priority packets cannot obtain bandwidth and are starved out.
  - Generally, services that are sensitive to delays are put into 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 that is implemented on a high-speed interface (such as an Ethernet interface). LPQ is not supported on a low-speed interface (such as a serial interface or MP-group interface).
  - SP scheduling applies to LPQ queues. The difference is that when congestion occurs, the PQ queue can preempt the bandwidth of the WFQ queue whereas the LPQ queue cannot. After packets in the PQ and WFQ queues are all scheduled, the remaining bandwidth can be assigned to packets in the LPQ queue.
  - In practice, BE flows can be put into LPQ queues. When the network is overloaded, BE flows can be limited so that other services can be processed preferentially.
- WFQ, PQ, and LPQ can be used separately or jointly for eight queues on an interface.

Scheduling Order of Three Types of Queues

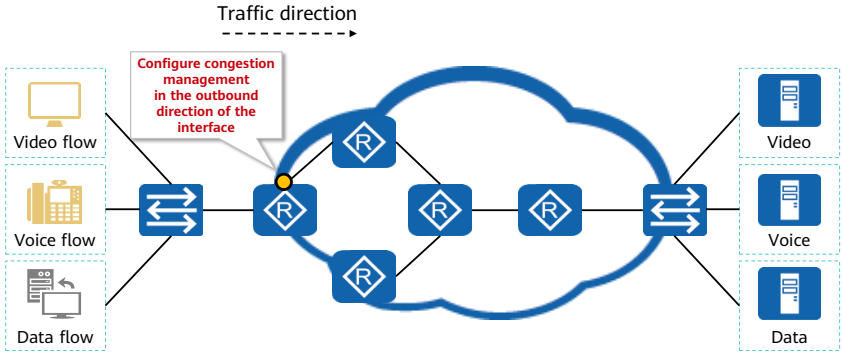




HUAWEI

- Interface queue scheduling order:
  - If PQ, WFQ, and LPQ queues use SP scheduling. PQ, WFQ, and LPQ queues are scheduled in sequence.
- Interface queue scheduling process:
  - Packets in PQ queues are preferentially scheduled, and packets in WFQ queues are scheduled only when no packets are buffered in PQ queues. When all PQ queues are empty, WFQ queues start to be scheduled. Packets in PQ queues are preferentially scheduled,
  - and packets in WFQ queues are scheduled only when no packets are buffered in PQ queues. Bandwidths are preferentially allocated to PQ queues to guarantee the PIR of packets in PQ queues.
  - Packets in LPQ queues are scheduled only after all packets in WFQ queues are sent.
- Scheduling result:
  - The PIR of PQ queues is guaranteed first, and the remaining bandwidth is allocated among WFQ queues based on weights.
  - When the PIR of all WFQ queues is guaranteed, the remaining bandwidth is allocated to LPQ queues.

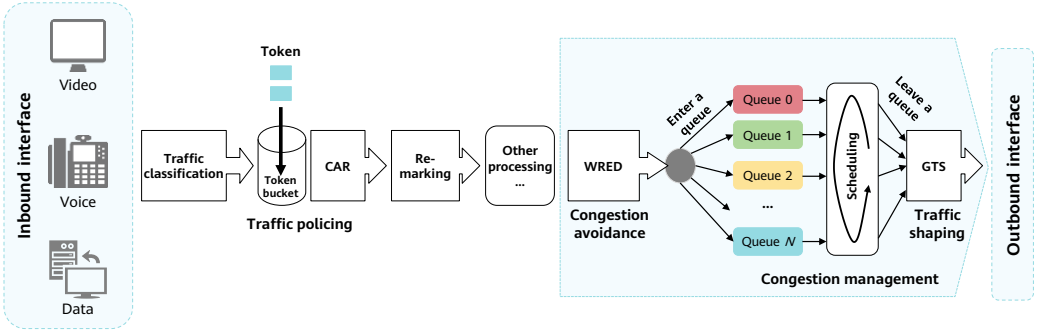
Application of Congestion Management



HUAWEI

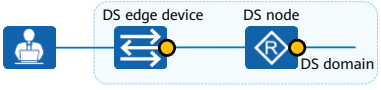
- Example:
- On a network, when multiple services compete for the same resources (such as the bandwidth and buffer), traffic congestion may occur and high-priority services may be not processed in a timely manner. Packets can be sent to different queues according to the priority mapping result, as shown in the figure. Different scheduling modes are set in the outbound direction to implement differentiated services.

QoS Processing Review



HUAWEI

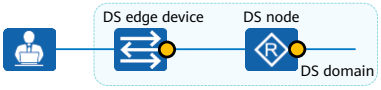
Configuring Queue-based Congestion Management



- WAN interfaces support three scheduling modes: PQ, WFQ, and PQ+WFQ. The configuration roadmap is as follows:
  - Create a queue profile.
  - Configure scheduling modes.
  - Apply the queue profile to the interface.

```
system-view
  qos queue-profile [queue-profile-name]   //Create a queue profile.
    schedule pq [queue-index] | wfq [queue-index]   //Configure scheduling modes for queues on a WAN interface.
  interface [interface-type interface-num]      //Enter the interface view.
    qos queue-profile [queue-profile-name]     //Apply the queue profile to the interface.
```

Configuring MQC to Implement Congestion Management (1)



- MQC provides three types of queues:
  - Assured Forwarding (AF) queues
  - Expedited Forwarding (EF) or LLQ queues
  - BE queues
- The configuration roadmap is as follows:
  - Configure a traffic classifier and a traffic behavior.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to the outbound direction of the device interface.

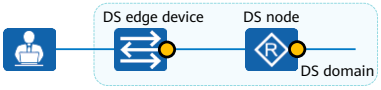
```
system-view
  traffic classifier [classifier-name]      //Create a traffic classifier.
    if-match [acl | vlan-id | .... ]      //Match traffic based on traffic characteristics.
```

```
system-view
  traffic behavior [behavior-name]      //Create a traffic behavior.
    queue af bandwidth [bandwidth | pct percentage]
//Configure the minimum bandwidth for AF queues in the traffic behavior.
    queue ef bandwidth [bandwidth | pct percentage]
//Configure the minimum bandwidth for EF queues in the traffic behavior.
    queue llq bandwidth [bandwidth | pct percentage]
//Configure the maximum bandwidth for LLQ queues in the traffic behavior.
    queue wfq queue-number [total-queue-number]
//Configure WFQ scheduling parameters for BE queues in the traffic behavior.
```


HUAWEI

- AF queue: AF queues ensure that service traffic is forwarded when the traffic rate does not exceed the minimum bandwidth.
- EF/LLQ queue: After packets matching certain rules enter EF or LLQ queues, they are scheduled in SP mode. Packets in other queues are scheduled only after all the packets in EF or LLQ queues are scheduled. In addition, EF queues can use the available bandwidth in AF or BE queues. The latency of LLQ queues is lower than that of common EF queues.
- BE queue: The remaining packets that do not enter AF or EF queues enter BE queues. BE queues are scheduled using the WFQ algorithm.
- The total bandwidth used by AF queues and EF queues cannot exceed 100% of the interface bandwidth.
- EF queues are provided with bandwidth preferentially. AF queues share the remaining bandwidth based on their weights.

Configuring MQC to Implement Congestion Management (2)



- MQC provides three types of queues:
  - AF queues
  - EF/LLQ queues
  - BE queues
- The configuration roadmap is as follows:
  - Configure a traffic classifier and a traffic behavior.
  - Bind the traffic classifier and traffic behavior to a traffic policy.
  - Apply the traffic policy to the outbound direction of the device interface.

```
system-view
  traffic policy [policy-name]    //Create a traffic policy.
    classifier [classifier-name] behavior [behavior-name]
//Bind the traffic classifier to the traffic behavior.
```

```
system-view
  interface [interface-type interface-num]      //Enter the
interface view.
    traffic-policy [policy-name] outbound     //Apply the traffic
policy to the outbound direction of the interface.
```


HUAWEI

- AF queue: AF queues ensure that service traffic is forwarded when the traffic rate does not exceed the minimum bandwidth.
- EF/LLQ queue: After packets matching certain rules enter EF or LLQ queues, they are scheduled in SP mode. Packets in other queues are scheduled only after all the packets in EF or LLQ queues are scheduled. In addition, EF queues can use the available bandwidth in AF or BE queues. The latency of LLQ queues is lower than that of common EF queues.
- BE queue: The remaining packets that do not enter AF or EF queues enter BE queues. BE queues are scheduled using the WFQ algorithm.
- The total bandwidth used by AF queues and EF queues cannot exceed 100% of the interface bandwidth.
- EF queues are provided with bandwidth preferentially. AF queues share the remaining bandwidth based on their weights.

Checking the Congestion Management Configuration

- Checking the queue-based congestion management configuration

```
system-view
  interface [interface-type interface-num]
    display this    //Check the queue profile bound to the interface.
  display qos queue-profile [queue-profile-name]   //Check the queue profile configuration.
```

- Checking the traffic classifier-based congestion management configuration

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]    //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

Quiz

1. (Single-answer question) How many queues are there on an interface?
   A. 6
   B. 7
   C. 8
   D. 9
2. (Multiple-answer question) Which of the following is a queue scheduling technology?
   A. PQ
   B. WFQ
   C. WRED
   D. FIFO

HUAWEI

- 1. C
- 2. ABD

Section Summary

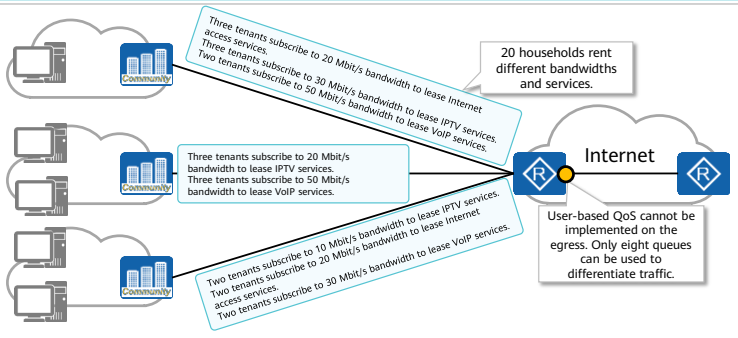
- After a data packet enters a queue, the device sends the data packet according to the queue scheduling mechanism.
- Common queue scheduling technologies include FIFO, PQ, and WFQ.
- PQ scheduling is performed before WFQ scheduling and FIFO. Queues scheduled in WFQ mode can transmit data only when queues scheduled in PQ mode have no data to transmit. The queue scheduled in FIFO mode can transmit data only when queues scheduled in PQ and WFQ mode have no data to transmit.

Contents

Limitations of QoS

- Traditional QoS distributes a flow into only eight queues for scheduling and control. Therefore, it has great limitations in multi-tenant scenarios.

QoS limitations in home broadband scenarios



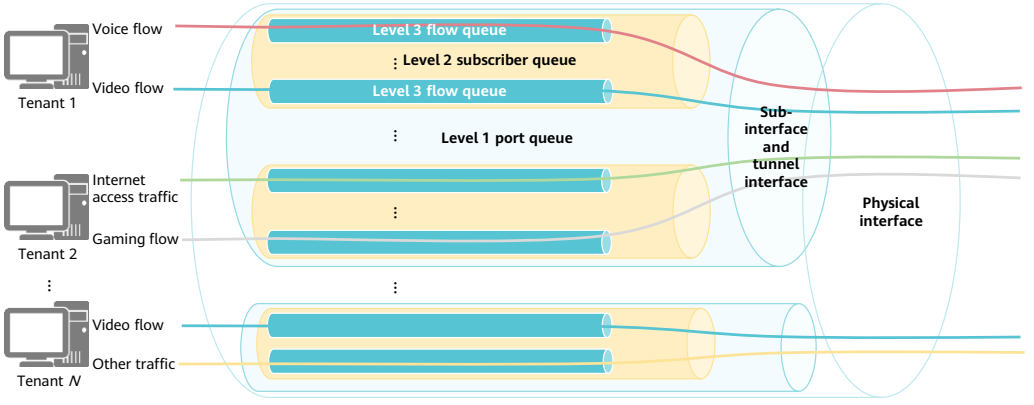
- In home broadband scenarios, different families may rent different network bandwidths and network services. Therefore, QoS cannot manage these families in a refined manner.

HQoS Overview

- Traditional QoS schedules traffic based on interfaces. An interface can only differentiate service priorities. The traffic of the same priority uses the same interface queue and competes for the same queue resources. Therefore, traditional QoS technology cannot provide differentiated services based on types of traffic and users.
- HQoS meets this requirement by implementing hierarchical scheduling based on multiple levels of queues, differentiating both services and users to provide refined QoS guarantee.
- Different devices provide different HQoS features. This section describes HQoS features supported by the CPE (AR series router).

Introduction to HQoS Queues

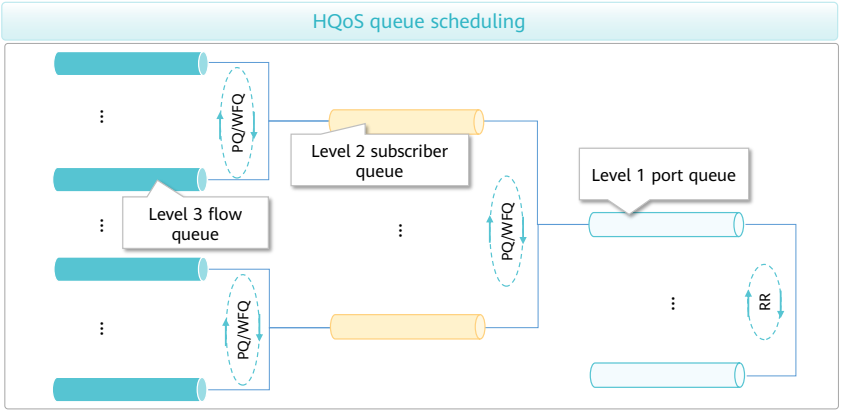
- The CPE supports three-level queues: flow queue (level 3), subscriber queue (level 2), and port queue (level 1).



- Flow queue
  - The same type of services of a user is taken as a service flow. HQoS schedules queues based on service flows. Flow queues correspond to service types and are classified into EF, AF, and BE queues. You can set scheduling modes for flow queues.
- Subscriber queue
  - Services from a user are placed into a subscriber queue. HQoS allows all services in the subscriber queue to share the bandwidth.
- Port queue
  - Each port corresponds to a queue and port queues are scheduled in RR mode. You can configure only interface-based traffic shaping, but cannot configure scheduling modes.

Introduction to HQoS Queue Scheduling

- The flow queue and subscriber queue support PQ scheduling, WFQ scheduling, and PQ+WFQ scheduling. The port queue uses RR scheduling.

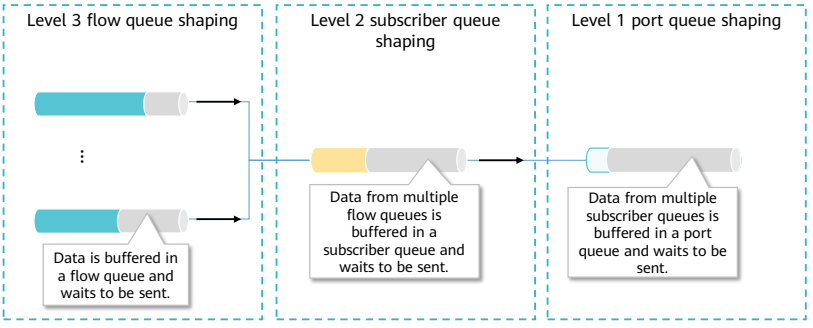


HUAWEI

- HQoS deployment for enterprise users is used as an example. Enterprise users have VoIP, video conference, and data services. Each subscriber queue corresponds to one enterprise user and each flow queue corresponds to a type of services. By deploying HQoS, the device can control the following items:
  - Traffic scheduling among three types of services of a single enterprise user
  - Total bandwidth of three types of services of a single enterprise user
  - Bandwidth allocation between multiple enterprise users
  - Total bandwidth of multiple enterprise users

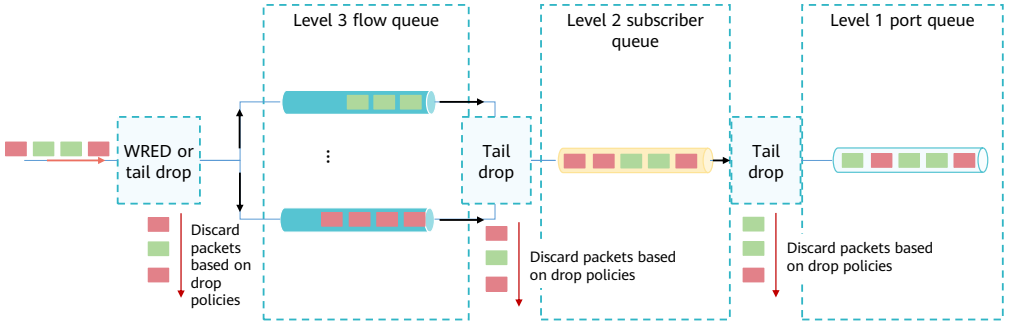
Introduction to HQoS Traffic Shaping

- The HQoS shaper buffers packets and limits the rate of packets. The device supports three levels of shapers, that is, flow queue shaper, subscriber queue shaper, and port queue shaper. After packets enter the device, the device buffers the packets in queues and sends the packets at the limited rate. Shapers can ensure the CIR and limit the maximum rate of packets by using the rate limiting algorithm.



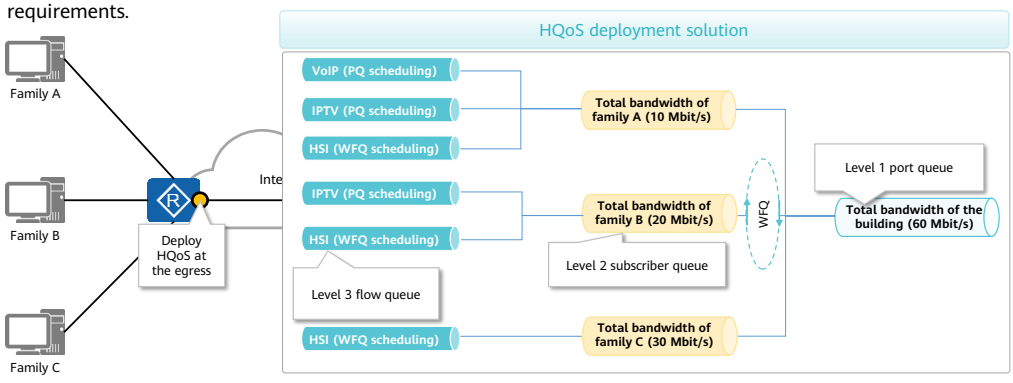
Introduction to the HQoS Dropper

- The HQoS dropper discards packets based on a drop policy before packets are sent to queues.
- The three types of queues supported by HQoS support different drop modes. The port queue and subscriber queue support tail drop; the flow queue supports tail drop and WRED.



HQoS Application Example

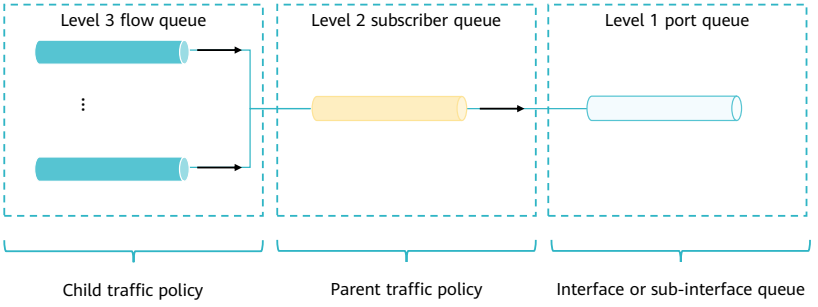
- Assume that there are three families in a building. Family A purchases 10 Mbit/s bandwidth and enables the VoIP, IPTV, and High Speed Internet (HSI) services. Family B purchases 20 Mbit/s bandwidth and enables the IPTV and HSI services. Family C purchases 30 Mbit/s bandwidth and enables only the HSI service. HQoS can meet these requirements.



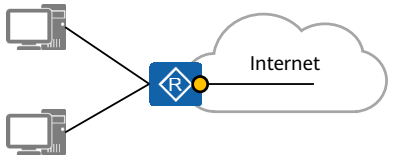
HUAWEI

HQoS Configuration Roadmap

- The HQoS configuration is complex. Generally, the MQC mode is used.
- When HQoS is configured, the policy nesting mode is used.
  - The parent traffic policy differentiates users, and the child traffic policy differentiates traffic.
  - A parent traffic policy can have multiple child traffic policies.
  - A parent traffic policy applies to an interface.



Configuring a Child Traffic Policy



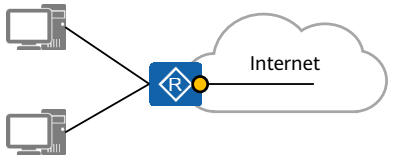
- Child traffic policies are used to differentiate services. You can configure multiple child traffic policies based on services when configuring HQoS.
- The configuration of HQoS child traffic policies is the same as that of common MQC. The configuration roadmap is as follows:
  - Configure a traffic classifier where traffic is matched based on service characteristics.
  - Configure a traffic behavior where the queue scheduling mode and queue bandwidth are defined.
  - Bind the traffic classifier and traffic behavior to a traffic policy.

system-view
traffic classifier [classifier-name] //Create a traffic classifier.
if-match [acl | vlan-id | ....] //Match traffic based on service characteristics.

system-view
traffic behavior [behavior-name] //Create a traffic behavior.
queue [af | ef | llq] bandwidth [bandwidth **| pct** percentage] //Configure AF, EF, or LLQ queue parameters in the traffic behavior.
drop-profile [drop-profile-name] //Bind the created drop profile to the traffic behavior.

system-view
traffic policy [policy-name] //Create a traffic policy.
classifier [classifier-name] **behavior** [behavior-name] //Bind the traffic classifier to the traffic behavior.

Configuring a Parent Traffic Policy



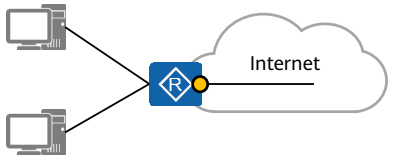
- A parent traffic policy is used to differentiate users. When configuring HQoS, you can bind multiple child traffic policies to a parent traffic policy.
- The configuration roadmap is as follows:
  - Configure a traffic classifier to match traffic based on user characteristics.
  - Configure a traffic behavior that needs to invoke a child traffic policy.
  - Bind the traffic classifier and traffic behavior to a traffic policy.

system-view
  traffic classifier [classifier-name] //Create a traffic classifier.
    if-match [acl | vlan-id | ....] //Match traffic based on user characteristics.

system-view
  traffic behavior [behavior-name] //Create a traffic behavior.
    queue [af | ef | llq] bandwidth [bandwidth | **pct** percentage] //(Optional) Configure AF, EF, or LLQ queue parameters in the traffic behavior.
    traffic-policy [policy-name] //Bind the sub traffic policy to the traffic behavior.

system-view
  traffic policy [policy-name] //Create a parent traffic policy.
    classifier [classifier-name] **behavior** [behavior-name] //Bind the traffic classifier to the traffic behavior.

Applying the Parent Traffic Policy




```
system-view
  interface [interface-type interface-num]      //Enter the interface view.
    traffic-policy [policy-name] outbound      //Apply the parent traffic policy to the outbound direction of the interface.
```

- After configuring a parent traffic policy, bind it to an interface or sub-interface.
- If the parent traffic policy is bond to a sub-interface, traffic between different sub-interfaces is sent from the physical interface in polling mode.
- The configuration roadmap is as follows:
  - Apply the parent traffic policy to the outbound direction of the interface.

Checking the HQoS Configuration

- After configuring HQoS, you can run the following commands to check the configuration.

```
system-view
  display traffic classifier user-defined [ classifier-name ]   //Check the traffic classifier configuration.
  display traffic behavior [ system-defined | user-defined ] [ behavior-name ]   //Check the traffic behavior configuration.
  display traffic policy user-defined [ policy-name ] classifier [classifier-name ]   //Check the traffic policy configuration.
  display traffic-policy applied-record [ policy-name ]   //Check the record of the specified traffic policy.
```

Quiz

1. (True or false) HQoS cannot distinguish users or services.
   A. True
   B. False
2. (Multiple-answer question) What are three types of HQoS queues?
   A. Flow queue
   B. Subscriber queue
   C. Data queue
   D. Port queue

- 1. B
- 2. ABD

Section Summary

- HQoS can ensure services with finer granularities.
- HQoS has three levels of queues: flow queue, subscriber queue, and port queue. Traffic shaping can be deployed for the three types of queues. Flow queues are scheduled in PQ+WFQ mode, subscriber queues are scheduled in PQ+WFQ mode, and interface queues are scheduled in RR mode.

Summary

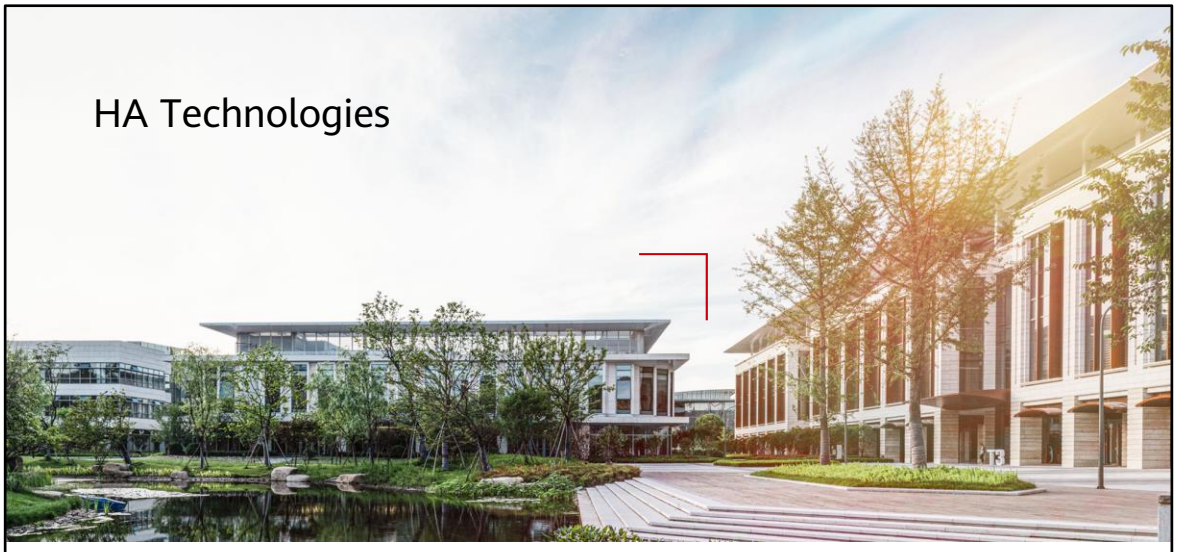
- QoS is an important means to ensure service quality. Generally, the DiffServ model is used on the live network.
- This model uses rate limiting, congestion avoidance, and congestion management.
- HQoS is used in complex scenarios with finer granularity. Flow queues, subscriber queues, and port queues can be used to distinguish different users and different services of the same user.

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.

HA Technologies



HUAWEI

Foreword

- In addition to network connectivity, network reliability is also an important indicator for measuring network quality.
- The guidelines for improving network reliability are to detect and correct faults in a timely manner.
- Various technologies, including bidirectional forwarding detection (BFD), network quality analyzer (NQA), and IP flow performance measurement (FPM), are available to detect faults in a timely manner.
- There are many technologies for correcting faults in a timely manner, such as Virtual Router Redundancy Protocol (VRRP), fast reroute (FRR), non-stop forwarding (NSF), and smart policy routing (SPR).
- Different technologies vary according to different application scenarios.
- This course introduces some commonly used high reliability (HA) technologies at the link, network, and service levels.

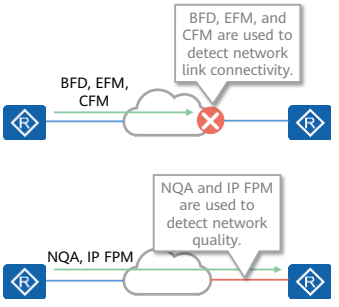
Objectives

- Upon completion of this course, you will be able to:
  - Describe the common technologies and fundamentals of link detection.
  - Describe the common technologies and fundamentals of link backup.
  - Understand the fundamentals and application scenarios of VRRP.
  - Understand the fundamentals and application scenarios of Smart Application Control (SAC).
  - Understand the fundamentals and application scenarios of SPR.

Contents

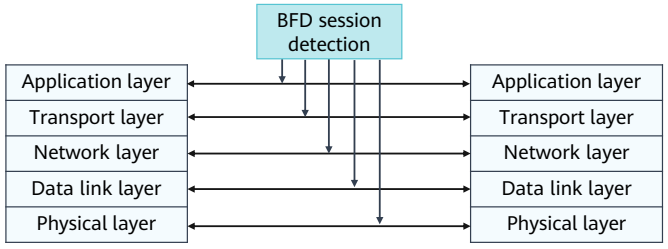
Overview of Link Detection

- Fluctuation of network link quality affects service quality. How to quickly detect link quality 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:
    - BFD
    - Ethernet in the First Mile (EFM)
    - Connectivity Fault Management (CFM)
  - Link quality detection technologies:
    - NQA
    - IP FPM
- On the live network, BFD is typically used to detect link connectivity, and NQA 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.



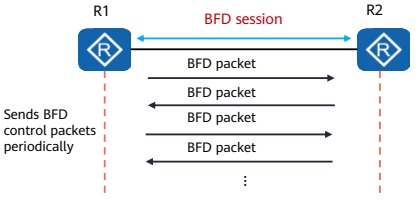
- Only one BFD session can be established in a data path. If different applications need to use different BFD parameters on the same data path, use the BFD parameters that can meet the requirements of all applications to configure a unique BFD session and enable the status changes of the BFD session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

BFD Detection

- Two systems establish a BFD session and periodically send BFD control packets along the path between them. If one system does not receive BFD control packets within a certain period, the system considers the path faulty. BFD has two modes: asynchronous mode and demand mode.

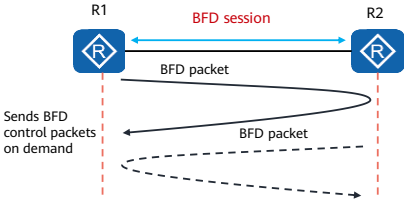
Asynchronous mode

- In asynchronous mode, two systems periodically send BFD control packets to each other. If one system does not receive any BFD control packet from the other system within the detection period, it declares the session down.



Demand mode

- To verify connectivity, one system sends a BFD control packet to the other system. If it receives a response packet from the other system within the detection period, the BFD session between the two systems remains up. Otherwise, it declares the session down.



HUAWEI

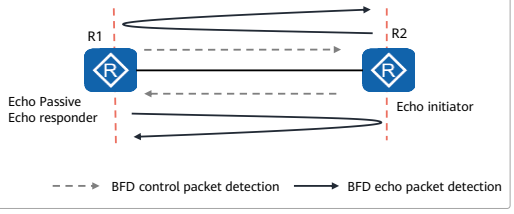
- The essential difference between the asynchronous mode and demand mode is that the detection location is different. In asynchronous mode, the local end sends BFD control packets at a certain interval, the detection location is on the remote end, and the remote end checks whether the local end periodically sends BFD control packets. In demand mode, the local end checks whether the BFD control packets sent by itself are responded.

BFD Echo

- The BFD echo function enables the local system to send BFD echo packets to the remote system, which then forwards the echo packets back along the same path in order to perform detection. This function provides rapid failure detection.
- The BFD echo function is classified into passive BFD echo and one-arm BFD echo based on application scenarios.

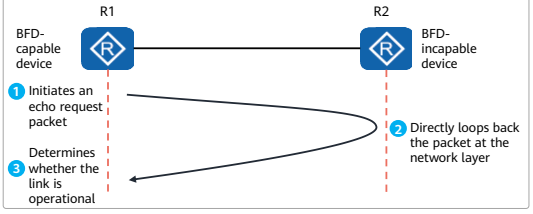
Passive BFD echo

- Passive BFD echo applies only to single-hop IP link scenarios and must work with asynchronous BFD. In asynchronous echo mode, slow detection is performed in asynchronous mode and fast detection is performed in echo mode at the same time. This function applies when the BFD detection capabilities of the two ends are inconsistent.



One-arm BFD echo

- One-arm echo applies only to single-hop IP link scenarios. A one-arm BFD echo session can be established on a device that supports BFD. After receiving a one-arm BFD echo session packet, devices that do not support BFD immediately loop back the packet, implementing fast link failure detection. This function is typically used when one system does not support BFD.



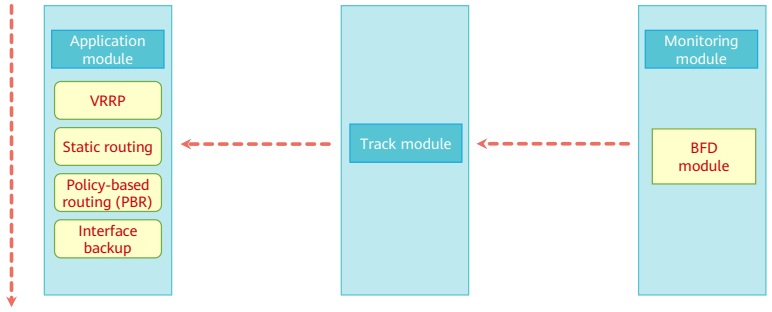
HUAWEI

- The process of establishing a passive BFD echo session is as follows:
  - R2 functions as a BFD session initiator and sends an asynchronous BFD packet to R1. The Required Min Echo RX Interval field carried in the packet is a nonzero value, requiring R1 to support BFD echo.
  - After receiving the packet, R1 finds that the value of the Required Min Echo RX Interval field in the packet is a nonzero value. If R1 has passive BFD echo enabled, it checks whether any ACL that restricts passive BFD echo is referenced. If such an ACL is referenced, only the BFD sessions that match specific ACL rules can enter the asynchronous echo mode. If no ACL is referenced, BFD sessions immediately enter the asynchronous echo mode.
  - R2 periodically sends BFD echo packets, and R1 sends BFD echo packets (with the source and destination IP addresses being the local IP address, and the destination physical address being R2's physical address) at the locally configured minimum interval for receiving BFD packets. Both R1 and R2 start a receive timer, with the receive interval being the same as the interval at which they each send BFD packets.
  - After R1 and R2 receive BFD echo packets from each other, they immediately loop back the packets at the forwarding layer. R1 and R2 also send asynchronous BFD packets to each other at an interval much smaller than that for sending echo packets.

- The one-arm echo function does not require both ends to negotiate the echo capability. That is, it applies when only one end supports BFD. The local device that has one-arm BFD echo enabled sends a special BFD packet (both the source and destination IP addresses in the IP header are the local IP address, the destination physical address is the physical address of the peer device, and the MD and YD in the BFD payload are the same). After receiving the packet, the peer device immediately loops the packet back to the local device to determine link reachability. One-arm BFD echo can be used on low-end devices that do not support BFD.

BFD Association

- BFD association involves the monitoring module, track module, and application module.



- The monitoring module monitors the link status and network performance and notifies the track module of the detection result.
- After receiving the detection result from the monitoring module, the track module changes the track status in a timely manner and notifies the application module of the change.
- The application module takes actions according to the track status, implementing association with the monitoring module and track module.

BFD Configuration

- The configuration roadmap is as follows:
  - Enable BFD on devices at both ends.
  - Specify the IP addresses of devices at both ends to establish a BFD session.
  - Configure matching local and remote discriminators on both ends to determine the BFD session.
  - Commit the configuration and start the BFD session.
  - Associate the BFD session with protocols as required. BFD sessions can be associated with a variety of protocols.

- Configure BFD.

```
system-view
  bfd       //Enable BFD globally.
  bfd [session-name] bind peer-ip [ip-address]          //Create a BFD session and specify the peer device with which the BFD session needs to be established.
  discriminator local [discr-value]       //Set the local discriminator of the BFD session. The value must be the same as the value of discriminator remote on the peer device.
  discriminator remote [discr-value]       //Set the remote discriminator of the BFD session. The value must be the same as the value of discriminator local on the peer device.
  commit       //Commit the configuration.
```

- Associate BFD with static routes.

```
system-view
  ip route-static [ip-address] [mask-length] [next-hop] track bfd-session [session-name]       //Associate the BFD session with a static route.
```


- For more configuration commands, see Huawei product manuals.

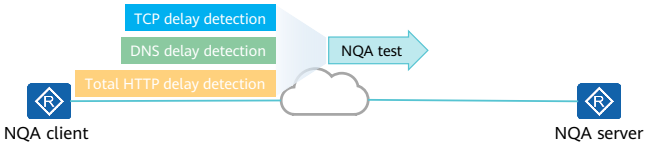
Verifying the BFD Configuration

- After BFD is configured, check information about the configured BFD session.

```
system-view
  display bfd interface [ interface]    //Check information about the BFD-enabled interface.
  display bfd session all verbose       //Check BFD session information.
  display bfd statistics        //Check global BFD statistics.
  display bfd statistics session all        //Check BFD session statistics.
```

Overview of NQA

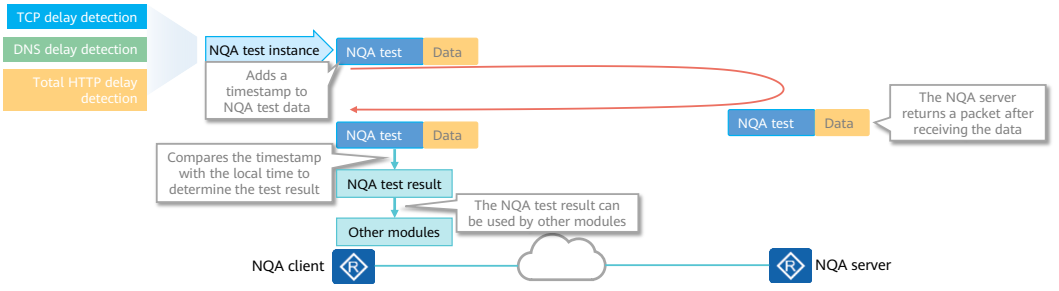
- To visualize the quality of network services and allow users to check whether the quality of network services meets requirements, the following measures must be taken:
  - Enable devices to provide network service quality information.
  - Deploy probe devices to monitor network service quality.
- The 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.
- It measures network performance and collects statistics about the response time, jitter, and packet loss rate in real time.



- 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.

NQA Fundamentals

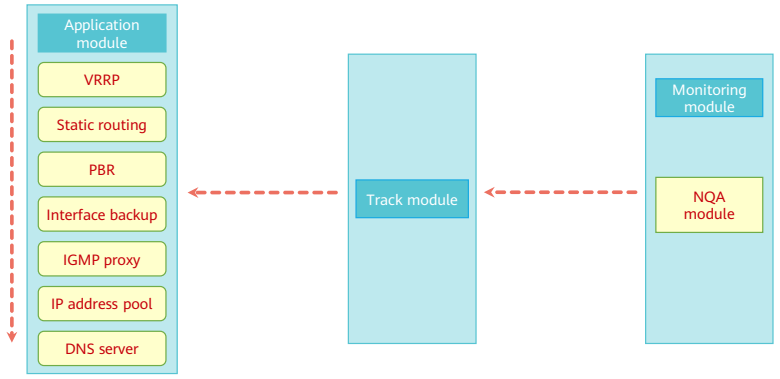
- NQA sends datagrams of a specific service to test the quality of the service on the live network.
- In an NQA test, devices on both ends are NQA client and NQA server, which are also called the source and destination. The NQA client initiates an NQA test. Test instances can be configured for the NQA client through the CLI or web system. NQA then places the test instances into test queues for scheduling.
- An NQA test can be associated with other modules. That is, NQA notifies other modules of the test result, and other modules take actions according to the test result.



- NQA can be associated with VRRP, static routing, interface backup, Internet Group Management Protocol (IGMP) proxy, IP address pool, domain name system (DNS) server, and PBR.
- The maximum delay, maximum packet loss rate, and maximum jitter can be configured for an NQA test instance. If the delay, jitter, or packet loss rate tested in an NQA test instance exceeds the configured values, the NQA test instance fails.

NQA Association

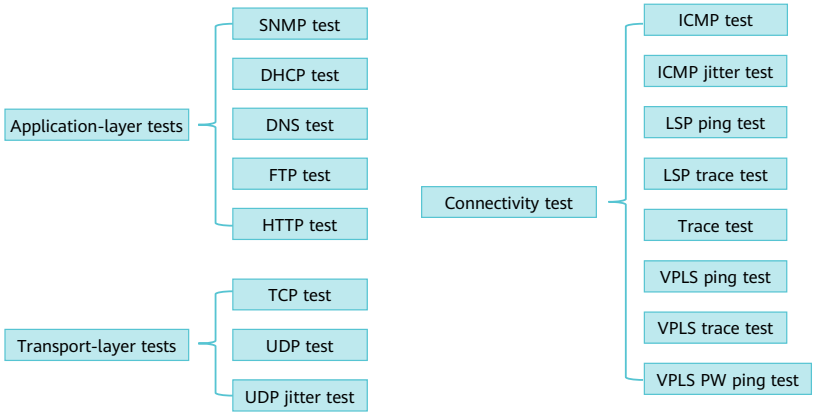
- NQA association involves the monitoring module, track module, and application module.



- The monitoring module monitors the link status and network performance and notifies the track module of the detection result.
- After receiving the detection result from the monitoring module, the track module changes the track status in a timely manner and notifies the application module of the change.
- The application module takes actions according to the track status, implementing association with the monitoring module and track module.

NQA Test Types

- NQA can be used to test the network quality of commonly used services:



NQA Server Configuration (TCP Test)

- Many NQA test types are available. The following example configures a TCP delay test.
- A server can be configured for a TCP test. In some tests, such as an HTTP test, a real HTTP server is typically used as the server.
- The configuration roadmap is as follows:
  - Configure the IP address and TCP port to be checked in the NQA TCP test.

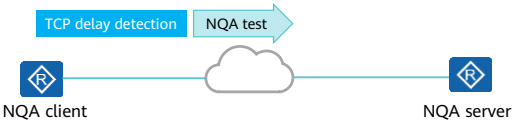
- Run the following commands on the NQA TCP test server:

```
system-view
  nqa-server tcpconnect [ip-address] [port-num]
//Configure the IP address and TCP port to be checked in the NQA TCP test.
```


HUAWEI

- The NQA server needs to be configured for some tests, including:
  - TCP test
  - UDP test
  - UDP jitter test
- Other NQA tests require a real server. For example, an HTTP test requires an NQA test request to be sent to the web server.

NQA Client Configuration (TCP Test)



- Many NQA test types are available. The following example configures a TCP delay test.
- The configuration roadmap is as follows:
  - Create an NQA test instance.
  - Set the test instance type.
  - Configure the server IP address.
  - Configure test parameters, such as the maximum delay and test period.

- Run the following commands on the NQA TCP test client:

```
system-view
  nqa test-instance [admin-name] [test-name]        //Specify the administrator name and name of the NQA test instance.
  test-type [type]        //Set the test type. If TCP services need to be tested, set the test type to TCP.
  destination-address ipv4 [ip-address]        //Configure the destination IP address for the test.
  destination-port [port-num]        //(Optional) Set the destination port for the test.
  timeout [time]        //(Optional) Set the maximum delay.
  start now        //Start the test immediately.
```

Verifying the NQA Configuration

- After an NQA test instance is configured, run the following commands to check the configuration of the test instance.
- On the client:

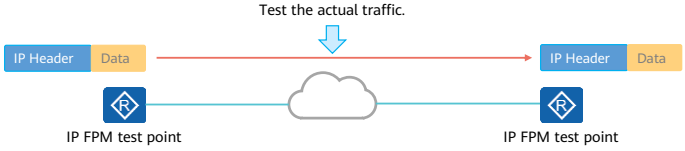
```
system-view
  display nqa application        //Check the NQA test instance type for the service.
  display nqa-parameter          //Check the parameter settings of the current test instance.
  display nqa support-server-type        //Check the server types supported by NQA.
  display nqa support-test-type        //Check the test types supported by NQA.
  display nqa-agent        //Check the NQA client status and configuration.
```

- On the server:

```
system-view
  display nqa-server        //Check NQA server information.
  display nqa-server session        //Check NQA client information on the NQA server.
```

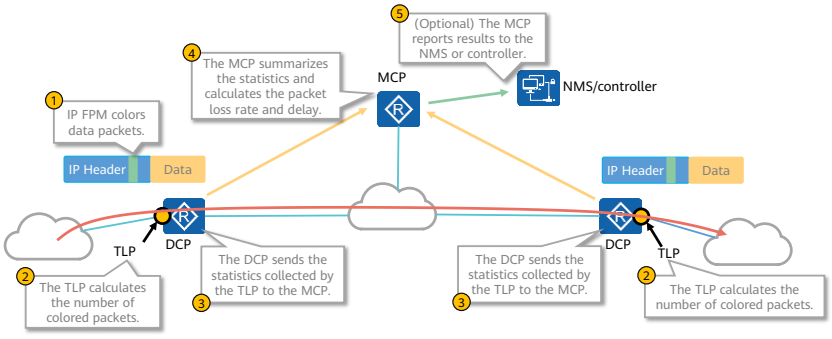
Overview of IP FPM

- 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.



IP FPM Fundamentals

- Traffic entering a network from a border device will leave the network from another border device. To determine end-to-end service performance, you only need to measure the specific traffic on these ingress and egress border devices.
- IP FPM sets the ToS or Flags field in the IP header of a packet to color the packet, helping ingress and egress devices measure the packet loss rate or jitter of specific traffic.

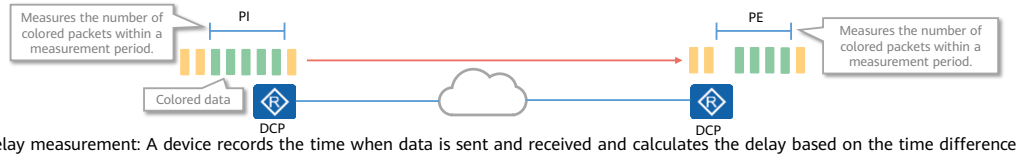


HUAWEI

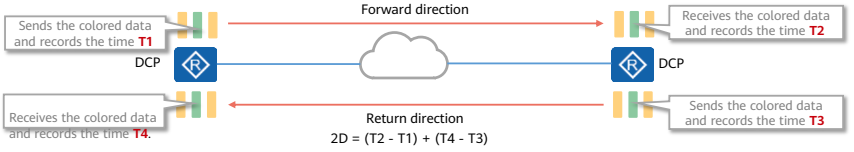
- TLP is short for target logical port.
  - TLPs are interfaces on the edge nodes of the network and provide the following functions:
  - Collect statistics about the packet loss rate and delay.
  - Generate statistics, such as the number of packets sent and received, traffic bandwidth, and timestamp.
  - 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 network management system (NMS).

IP FPM Measurement Mechanism

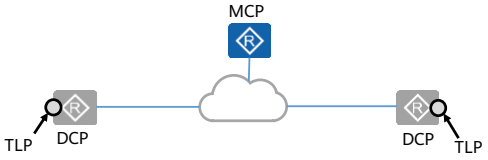
- IP FPM supports packet loss measurement and delay measurement.
- Packet loss measurement: The number of colored packets entering the network (PI) and the number of colored packets leaving the network (PE) are measured within a certain period. The packet loss rate is calculated based on the PI and PE using the following formula: Lost Packet = PI - PE.



- Delay measurement: A device records the time when data is sent and received and calculates the delay based on the time difference between the time when data is sent and the time when data is received on different devices: Two-way delay = Delay of the forward traffic + Delay of the return traffic.



Configuring an MCP

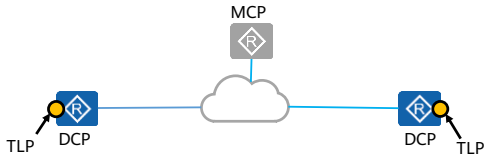


- Two roles, MCP and DCP, need to be configured for IP FPM.
- The configuration roadmap is as follows:
  - Enable MCP globally.
  - Configure an MCP ID.
  - Configure an IP FPM test instance.
  - Configure a DCP ID.

- Configure an MCP.

system-view
  nqa ipfpm mcp //Enable MCP globally.
    mcp id [id] //Configure an MCP ID. It is an IP address used for communication between the MCP and DCPs. Ensure that the MCP and DCP are reachable to each other through their IP addresses.
    authentication-mode hmac-sha256 key-id [id] **cipher** [password] //(Optional) Configure authentication for IP FPM protocol packets.
  instance [id] //Configure an IP FPM instance. The instance ID must be unique in the MCP management domain.
    dcp [id] //Configure a DCP ID. It is an IP address used for communication between an MCP and a DCP. Ensure that the MCP and DCP are reachable to each other at the IP layer.

Configuring a DCP



- Two roles, MCP and DCP, need to be configured for IP FPM.
- The configuration roadmap is as follows:
  - Enable DCP globally.
  - Configure a DCP ID.
  - Configure an MCP ID.
  - Configure an IP FPM test instance.
  - Configure a target flow to be monitored.
  - Configure a TLP.

- Configure a DCP.

```
system-view
  nqa ipfpm dcp        //Enable DCP globally.
    dcp id [id]      //Configure a DCP ID. It is an IP address used for communication between an MCP and a DCP. Ensure that the MCP and DCP are reachable to each other at the IP layer.
    authentication-mode hmac-sha256 key-id [id] cipher [password]
//(Optional) Configure authentication for IP FPM protocol packets.
    mcp [id]         //Configure an MCP ID. It is an IP address used for communication between an MCP and a DCP. Ensure that the MCP and DCP are reachable to each other at the IP layer.
  instance [id]        //Configure an IP FPM instance.
    flow forward protocol [TCP | DCP] source [ip-address/port] destination [ip-address | port]          //Configure the target flow to be monitored.
    tlp [tlp-id] [in-point ingress | out-point egress]        //Configure a TLP ID to determine whether the TLP is an inbound or outbound interface.
    loss-measure enable continual        //Configure packet loss measurement.
    delay-measure enable two-way tlp [tlp-id] continual   //Configuration delay measurement.
  interface [interface-type interface-name]       //Enter the interface view.
    ipfpm tlp [tlp-id]      //Bind the TLP ID to the interface.
```

Verifying the IP FPM Configuration

- IP FPM commands display IP FPM performance statistics, helping monitor the IP FPM running status.

```
system-view
  display ipfpm statistic-type [loss | twoway-delay ] instance [instance-id]
//Check the performance statistics of a specified IP FPM instance.
```

Quiz

1. (Multiple-answer question) Which of the following services can be checked by NQA?
   A. DNS
   B. HTTP
   C. IGMP
   D. ICMP

HUAWEI

- 1. ABD

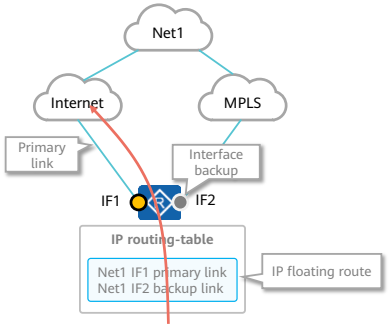
Section Summary

- BFD is mainly used to check the connectivity of the network layer. BFD can be associated with static routes, dynamic routes, and interface backup to facilitate fast network convergence.
- NQA can detect network connectivity and network quality based on services, for example, NQA can measure the TCP delay, DNS delay, and TCP jitter. NQA can be associated with static routes, dynamic routes, and interface backup to speed up network convergence.
- NQA simulates services for detection, so the test result is not accurate. IP FPM is mainly used to improve the detection accuracy. IP FPM colors packets and measures the colored packets on network border devices to check the actual service quality.

Contents

Link Backup

- To ensure reliability, a network egress is typically connected to multiple WAN links in active/standby mode. When one link is faulty (for example, an interface is faulty, the link is faulty, or the link bandwidth is insufficient), services can be immediately switched to another link. To meet this requirement, the following technologies are often used on the live network:
  - Interface backup
  - IP floating route

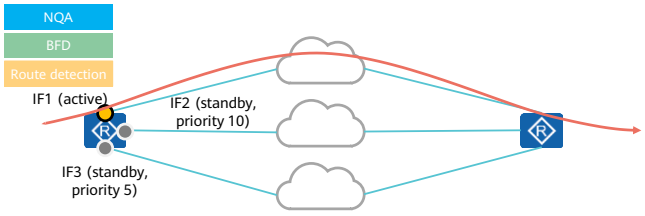


Interface Backup

- Interface backup refers to the backup between specific interfaces on the same device. When an interface is faulty or the bandwidth is insufficient, traffic can be fast switched to a standby interface. The standby interface then transmits services or load balances network traffic.
- Interface backup operates in either active/standby or load balancing mode.
  - Interface backup in active/standby mode: One interface is the active interface, and the others are standby interfaces. When the active interface fails or the network quality is poor, a standby interface transmits data.
  - Interface backup in load balancing mode: One interface is the active interface, and the others are standby interfaces. When the bandwidth of the active interface is insufficient, a standby interface is enabled. Then both the active and standby interfaces transmit data.

Interface Backup in Active/Standby Mode

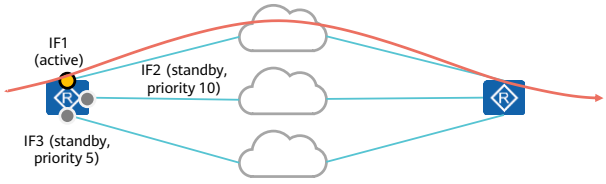
- In active/standby mode, only a single interface transmits services at any time. The mechanism of interface backup in active/standby mode is as follows:
  - When the active interface IF1 is working properly, interfaces IF2 and IF3 are in the standby state.
  - When IF1 is faulty or the link quality does not meet requirements, IF2 with the highest priority enters the forwarding state.
  - After IF1 recovers, traffic is switched back to IF1.



- Interface backup in active/standby mode can detect only faults on direct links but not the remote link status or overall link quality. In this case, interface backup in active/standby mode can be associated with NQA, BFD, or routing tables.

Interface Backup in Load Balancing Mode

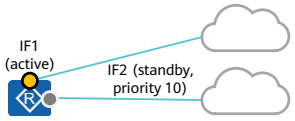
- In load balancing mode, if the bandwidth of the active interface on a device is insufficient, the device uses standby interfaces to transmit data.
  - When traffic on the active interface IF1 does not reach the upper bandwidth limit, IF1 alone forwards the traffic.
  - When the traffic reaches the upper bandwidth limit, the standby interface IF2 with the highest priority is enabled to forward traffic at the same time.
  - If one standby interface cannot meet service bandwidth requirements, the standby interface IF3 with the second highest priority is used. The rest can be deduced by analogy until the standby interface that meets service bandwidth requirements is used.
  - When the traffic volume decreases, standby interfaces are shut down in ascending order of priority.



HUAWEI

- Interface backup in load balancing mode is implemented only based on interface bandwidth usage and cannot be associated with BFD or NQA.

Configuring Interface Backup

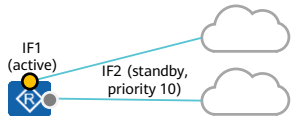


- The configuration roadmap is as follows:
  - Ensure the network connectivity of the primary and backup links.
  - Configure standby interfaces and their priorities.

- Configure interface backup:

system-view
interface [interface-type interface-num] //Enter the active interface view.
standby interface [interface-type interface-num] [Priority] //Configure a standby interface for the active interface. If multiple standby interfaces are required, you need to run this command multiple times.
standby time delay [enable-delay] [disable-delay] //Configure the interface switching delay. enable-delay specifies the delay in a switchover from the active interface to the standby interface. disable-delay specifies the delay in a switchback from the standby interface to the active interface.

Configuring Interface Backup Association



- Interface backup can be associated with NQA, BFD, and routing tables to improve service transmission reliability.
- The configuration roadmap is as follows:
  - Enable the association function when configuring a standby interface on the active interface.

- Associate interface backup with NQA:

```
system-view
  interface [interface-type interface-num]        //Enter the standby interface view.
    standby track nqa [admin-name] [test-name]
//Associate interface backup with NQA.
```

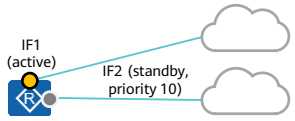
- Associate interface backup with BFD:

```
system-view
  interface [interface-type interface-num]        //Enter the standby interface view.
    standby track bfd-session session-name [name]
//Associate interface backup with BFD.
```

- Associate interface backup with routing tables:

```
system-view
  interface [interface-type interface-num]        //Enter the standby interface view.
    standby track ip route [ip-address] [mask-length]
//Associate interface backup with routing tables.
```

Configuring Interface Backup in Load Balancing Mode



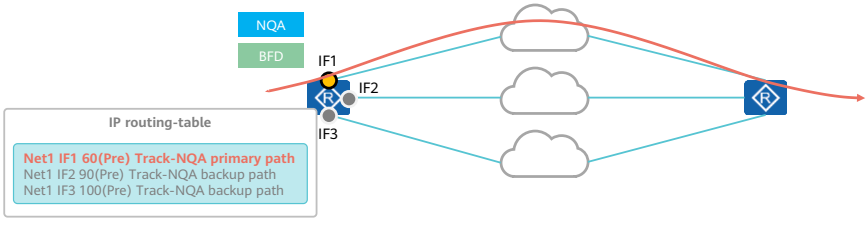
- Interface backup in load balancing mode enables multiple interfaces to transmit data at the same time.
- The configuration roadmap is as follows:
  - Configure standby interfaces on the active interface.
  - Set the percentage threshold for load balancing.

- Configure interface backup in load balancing mode:

system-view
interface [interface-type interface-num] //Enter the active interface view.
standby interface [interface-type interface-num] [Priority] //Configure a standby interface for the active interface. If multiple standby interfaces are required, you need to run this command multiple times.
standby threshold [enable-threshold] [disable-threshold] //Configure the percentage threshold for load balancing. enable-threshold specifies the threshold for enabling a standby interface, and disable-threshold indicates the threshold for disabling the standby interface.

IP Floating Route

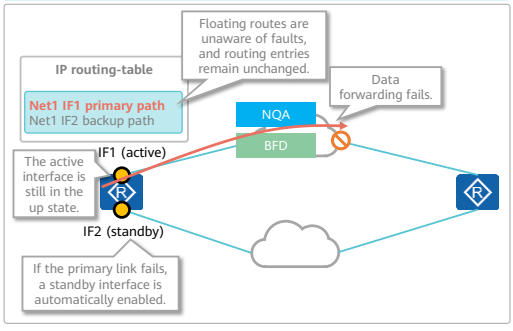
- An IP floating route is a static route. It provides a backup route when the primary route fails. A floating route is installed in the IP routing table only when the next hop of the primary route is unreachable.
- The IP floating route is implemented based on the Pre value in the IP routing table. In most cases, the Pre value of the backup route is set to be greater than that of the primary route.
- The active/standby switchover of IP floating routes is typically performed based on the interface status. Therefore, to detect the status of the entire path, NQA or BFD on the live network is often associated with IP floating routes. If NQA or BFD sessions fail, the primary route is considered ineffective.



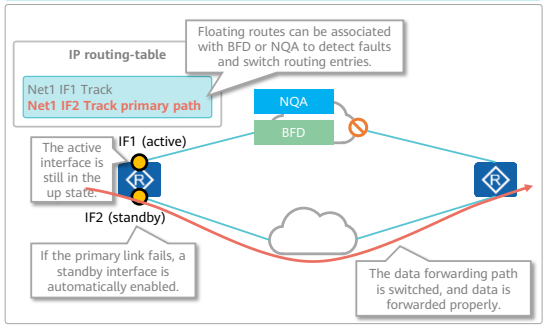
Interface Backup and IP Floating Route

- After interface backup is associated with BFD or NQA, if BFD or NQA detects that the primary link fails, a standby interface will be enabled. In this case, the active interface may still be physically up, so the related routing entries remain unchanged. As a result, data is still sent from the primary link.
- IP floating routes can be associated with BFD or NQA to solve this problem.

IP floating routes are not associated with BFD or NQA



IP floating routes are associated with BFD or NQA



HUAWEI

Quiz

1. (True or false) Interface backup can be implemented only in active/standby mode, but not in load balancing mode.
   A. True
   B. False

- 1. B

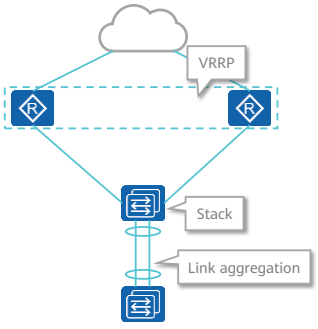
Section Summary

- Interface backup operates in either active/standby or load balancing mode. In active/standby mode, only one interface can work at a time. In load balancing mode, if the bandwidth of the active interface is insufficient, the standby interface can be used to forward traffic.
- Interface backup can be associated with NQA, BFD, and routing tables to detect link quality and determine whether to perform an active/standby interface switchover.
- Floating routes are typically used together with interface backup. Interface backup determines only whether interfaces can be enabled. Therefore, floating routes are required to guide data forwarding during Layer 3 forwarding.

Contents

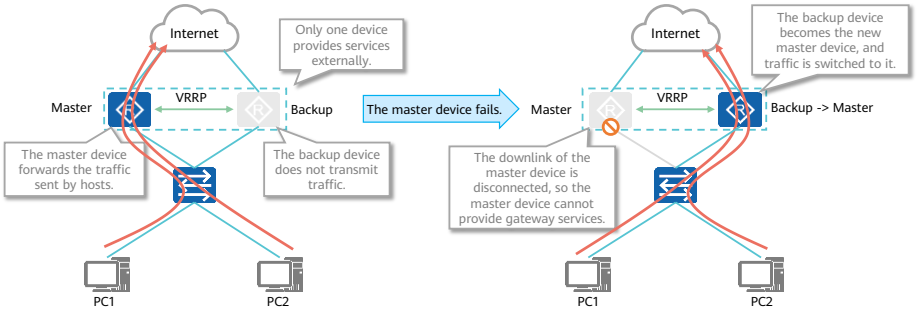
Network Reliability

- If a fault occurs on the network, the fault may not be detected or rectified in a timely manner. Therefore, 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.



Overview of VRRP

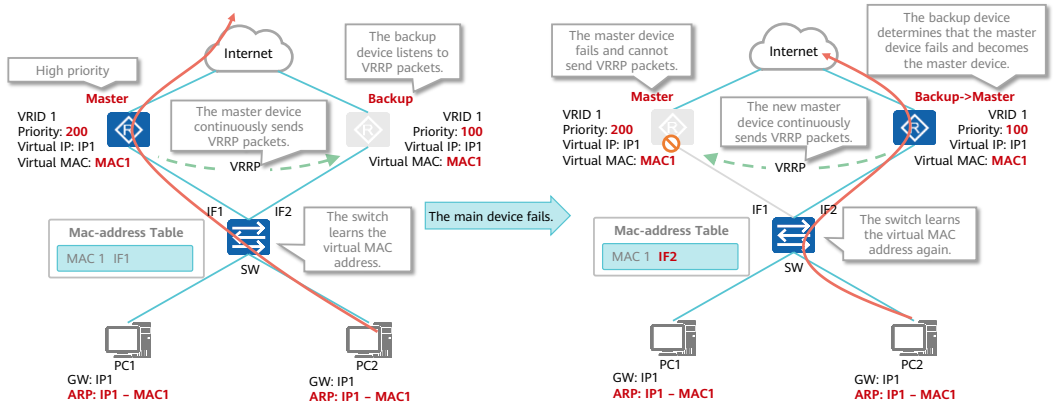
- Hosts are connected to external networks through gateways. If a single gateway fails, services will be interrupted for a long time. Adding egress gateways is a common method to improve system reliability. In this case, route selection among multiple egresses becomes essential.
- VRRP groups multiple 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

VRRP Fundamentals

- VRRP technology aims to select a master device for packet forwarding. If the master device fails, traffic can be switched to the backup device in a timely manner. The following illustrates the VRRP mechanism:



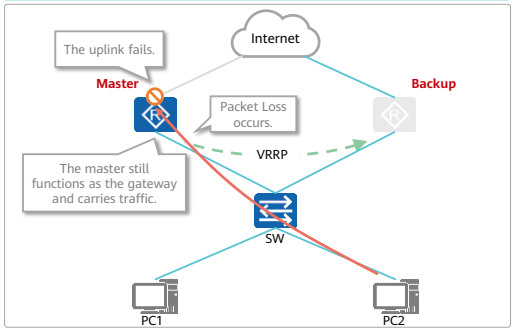
HUAWEI

- VRRP mechanism:
  - VRRP selects the master device based on the priorities of devices in a VRRP group. The master device sends gratuitous ARP packets to notify the devices or hosts that are connected to it of the virtual MAC address, and then starts forwarding packets.
  - The master device periodically sends VRRP Advertisement packets to all backup devices in the VRRP group to advertise its configurations (such as the priority) and operating status.
  - If the master device fails, the backup device with the highest priority is elected as the new master device.
  - After a master/backup switchover, the new master device immediately sends gratuitous ARP packets carrying the virtual MAC and virtual IP addresses to allow the devices or hosts that are connected to it to update corresponding MAC entries. After the update is complete, user traffic is switched to the new master device, which is transparent to users.
  - If the original master device recovers and it is the IP address owner (its priority is 255), it immediately switches to the Master state; whereas if the original master device recovers and its priority is lower than 255, it switches to the Backup state, and its original priority is restored.
  - If the priority of a backup device is higher than that of a master device, VRRP determines whether to reelect a new master, depending on the backup device's working mode (preemption or non-preemption).
    - Preemption mode: In this mode, a backup device preempts the master role if it has a higher priority than that of the current master.
    - Non-preemption mode: In this mode, a backup device does not preempt the master role even if it has a higher priority than that of the current master, provided that the current master is working properly.

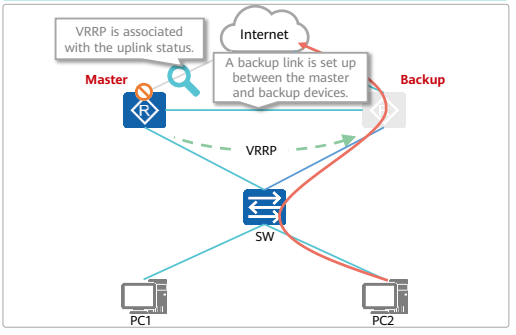
VRRP Uplink Reliability

- The master and backup devices in a VRRP group detect the master/backup status of each other. If the uplink fails, VRRP cannot detect the fault by default. As a result, packet loss occurs.
- To solve this problem, associate VRRP with the uplink or set up a backup link between the master and backup devices.

Packet loss due to a VRRP uplink failure



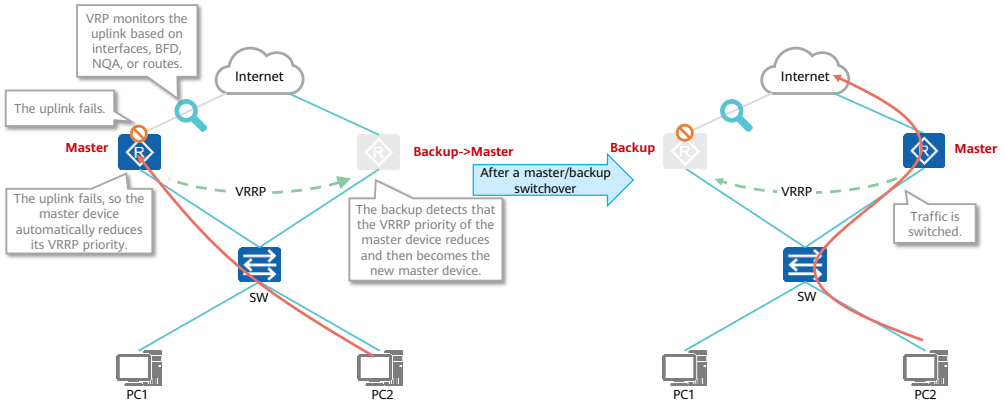
VRRP uplink reliability solution



HUAWEI

Associating VRRP with the Uplink Status

- VRRP can monitor the uplink based on interfaces, BFD, NQA, or routes. If the uplink fails, VRRP triggers a master/backup switchover to switch traffic, preventing packet loss caused by the uplink failure.



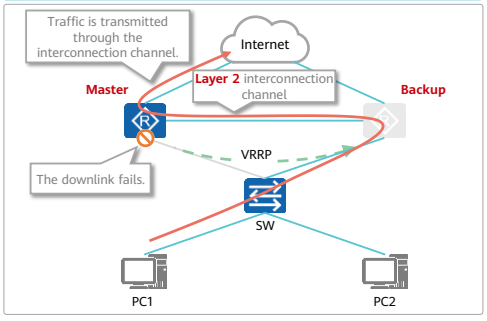
HUAWEI

- Interface-based link monitoring can monitor only the status of direct links and cannot detect the status of the remote network. Therefore, interface-based link monitoring has low accuracy but is easy to configure and does not require the peer device to support it.
- BFD/NQA/route-based uplink monitoring can monitor direct or remote links. However, BFD/NQA sessions need to be configured on both ends, which has limitations.
- After a master/backup VRRP switchover is performed, the switch learns MAC address entries again and forwards traffic to the master device based on the new MAC address table.

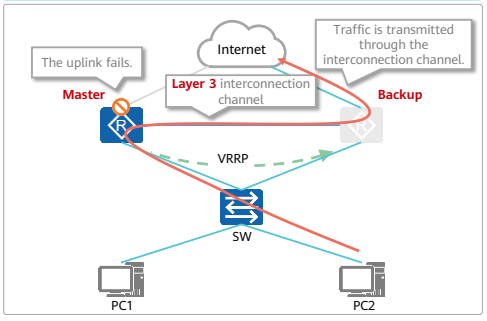
VRRP Establishing a Dual-Device Interconnection Channel

- On a live network, an additional link is deployed between the master and backup devices in a VRRP group to provide the following functions:
  - Functions as the heartbeat line between the master and backup devices to help the backup device detect the status of the master device.
  - Forwards traffic if the uplink or downlink fails.

Layer 2 interconnection channel transmitting traffic



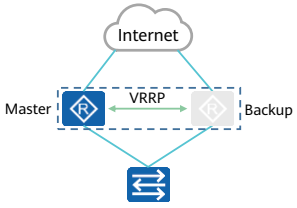
Layer 3 interconnection channel transmitting traffic



HUAWEI

- On a live network, association between the uplink status and VRRP is often deployed together with dual-device interconnection channels.
- An interconnection channel transmits traffic in the following two scenarios:
  - A Layer 2 interconnection channel is deployed between the master and backup devices. If the downlink of the master device fails but the uplink is normal, traffic is sent from the backup device to the master device and then to the Internet through the Layer 2 channel. VRRP does not trigger a master/backup switchover, improving VRRP stability.
  - A Layer 3 interconnection channel is deployed between the master and backup devices. When the downlink of the master device is normal but the uplink fails, traffic is sent to the master device, which then sends traffic to the backup device through the Layer 3 channel. The backup device forwards the traffic. In this scenario, VRRP does not trigger a master/backup switchover and uplink association is not required.

Basic VRRP Configurations

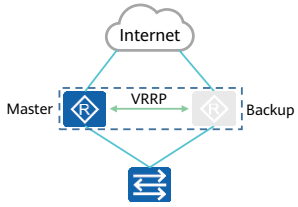


- The configuration roadmap is as follows:
  - Configure a VRRP group.
  - Set the device priority in a VRRP group.

- Perform basic VRRP configurations.

system-view
  interface [interface-type interface-num] //Enter the active interface view.
    vrrp vrid [virtual-router-id] **virtual-ip** [virtual-address] //Create a VRRP group and configure a virtual IP address for the VRRP group.
    vrrp vrid [virtual-router-id] **priority** [priority-value] //Configure the priority of the device in the VRRP group. By default, the priority value is 100. A larger value indicates a higher priority.

VRRP Association Configuration



- The configuration roadmap is as follows:
  - Configure VRRP association on an interface.

- Associate VRRP with BFD:

```
System-view
  interface [interface-type interface-name]       //Enter the active interface view.
    vrrp vrid [virtual-router-id] track bfd-session session-name [bfd-session-name]
increased | reduced [value] //Associate VRRP with BFD. If the BFD session is down, the
VRRP priority is changed.
```

- Associate VRRP with NQA:

```
system-view
  interface [interface-type interface-num]       //Enter the active interface view.
    vrrp vrid [virtual-router-id] track nqa [admin-name] [test-name] reduced [value-
reduced ]       //Associate VRRP with NQA. If the NQA session is down, the VRRP priority is
changed.
```

- Associate VRRP with the interface status:

```
system-view
  interface [interface-type interface-num]       //Enter the active interface view.
    vrrp vrid [virtual-router-id] track interface [interface-type interface-num] increased
| reduced [value]    //Associate the interface with NQA. When the interface is shut down,
the VRRP priority is changed.
```

Verifying the VRRP Configuration

- During routine maintenance, run the following commands to check VRRP packet statistics and monitor the VRRP group running status.

```
system-view
  display vrrp interface [interface-type interface-number ] [ virtual-router-id ] statistics  //Display statistics about sent and received packets of a VRRP group.
```

Quiz

1. (Single-answer question) After a master/backup VRRP switchover is performed, which of the following entries are used to forward data to the new master device?
   A. ARP entries
   B. IP routing entries
   C. MAC address entries
   D. FIB entries

- 1. C

Section Summary

- VRRP is typically deployed on gateways or egresses. Terminals are unaware of the master/backup VRRP switchover, and can access the network after the switchover.
- After a master/backup VRRP switchover is performed, the switch connected to the VRRP devices guides traffic switching.
- VRRP can be associated with multiple detection technologies to detect the uplink quality, which helps VRRP accurately perform a master/backup switchover.

Contents

Service Reliability

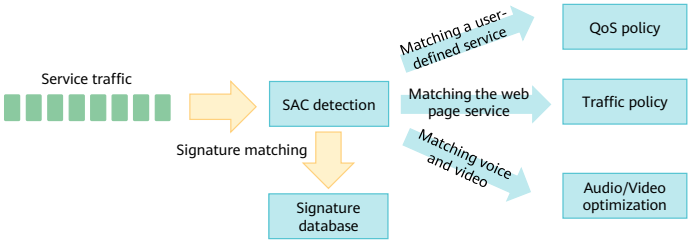
- In the cloud computing era, network reliability cannot meet user requirements. Users want to understand the live network status based on applications and adjust the network based on the application status.
- Such requirement poses the following challenges to traditional networks:
  - Traditional networks 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.

Overview of SAC

- Traditional networks are managed based on traffic. However, in the cloud computing era, services are becoming increasingly important. Networks need to be managed and monitored based on applications instead of Five-tuple information.
- 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 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.

SAC Signature Database

- Signature identification is a basic function of SA technology. Different applications typically use different protocols, and different application protocols have their own signatures. A signature that can identify a protocol is known as a signature code. The system analyzes service flows passing through a device, and compares the analysis result with the signature database on the device. It identifies an application by detecting the signature code in data packets.
- SAC signature databases include FPI and SA signature databases. FPI signatures refer to signatures for identifying FPI applications, and SA signatures refer to signatures for identifying SA applications.
- SAC working mechanism

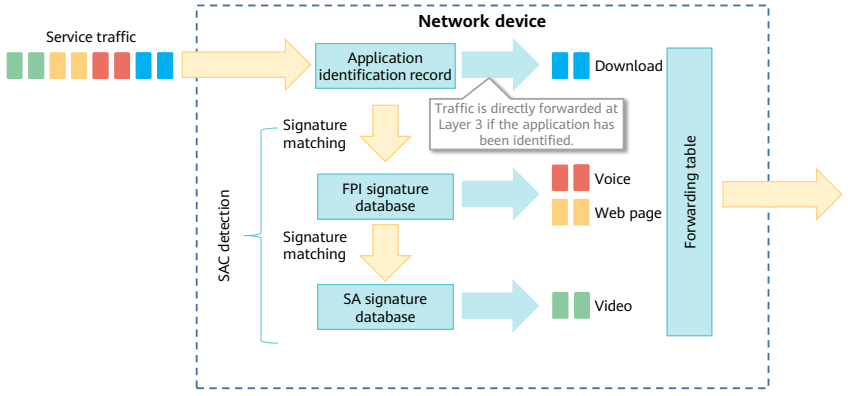


HUAWEI

- The SAC signature database file can only be updated through upgrades and cannot be manually modified.
- The SAC signature database can be updated in either of the following modes:
  - Online update: The SAC signature database can be updated through the security center platform or intranet update server.
  - Local update: The upgrade package is downloaded from the security center platform and uploaded to the device through FTP for the update of the SAC signature database.

SAC Application Identification Process

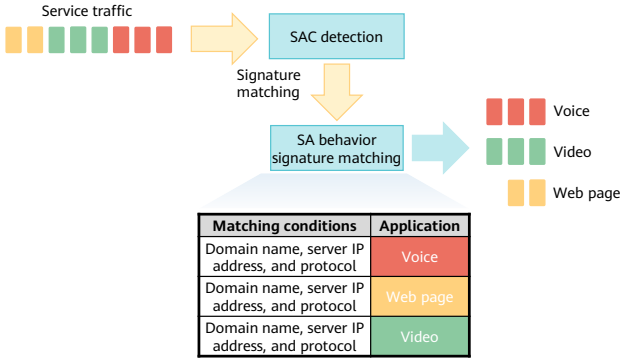
- During SAC application identification, the system checks whether an application is identified. If the application is not identified, the system checks the FPI signature database and SA signature database in sequence.



- 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. 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.

SA

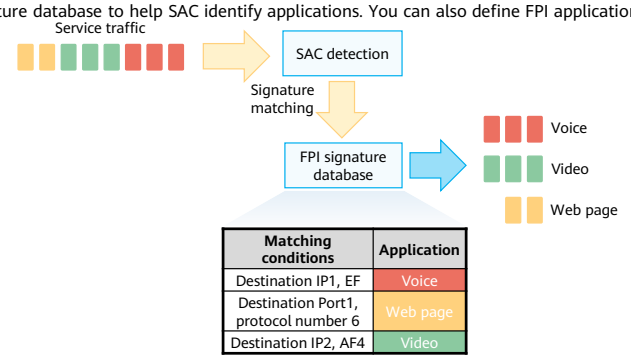
- After receiving data, the device can use service awareness (SA) technology to match applications.
- SA uses the SA signature database to detect services. The existing SA signature database is embedded with more than 6000 applications, ensuring a high identification rate for public applications. In most cases, the SA signature database can only be updated online or locally and cannot be manually modified.



Matching conditions	Application
Domain name, server IP address, and protocol	Voice
Domain name, server IP address, and protocol	Web page
Domain name, server IP address, and protocol	Video

FPI

- There is a problem in matching applications based on the SA signature code. That is, the application corresponding to the first several packets may fail to be identified based on the SA signature code. As a result, the processing on the first and subsequent packets may be inconsistent. First packet identification (FPI) enables a device to identify an application by matching the first packet of a flow.
- FPI identifies applications based on 5-tuple information, DSCP values, protocols, and DNS domain names. The system provides a predefined FPI signature database to help SAC identify applications. You can also define FPI applications to identify new applications.



Matching conditions	Application
Destination IP1, EF	Voice
Destination Port1, protocol number 6	Web page
Destination IP2, AF4	Video

HUAWEI

- FPI applications are classified into the following types:
  - Predefined and user-defined FPI applications based on the protocol and port number: These two types of applications are identified using entries that are generated based on the protocol and port number carried in packets. The difference is as follows: Packets of a predefined FPI application contain common protocols and port numbers, while packets of a user-defined FPI application contain the protocols and ports that you define.
  - Predefined and user-defined FPI applications based on the DNS domain name: These two types of applications are identified using DNS entries generated through association between FPI and DNS. The difference is as follows: Packets of a predefined FPI application contain common DNS domain names, while packets of a user-defined FPI application contain the DNS domain names that you define.
  - User-defined FPI application based on 5-tuple and DSCP information. This application is identified based on the user-defined 5-tuple and DSCP information using advanced ACL rules.
- Identification process of FPI applications based on the DNS domain name
  - FPI applications based on the DNS domain name are identified using DNS entries generated through association between FPI and DNS. The FPI signature database contains the mappings between domain names and applications. DNS response packets contain the mappings between domain names and IP addresses. Based on the mappings, a device generates DNS entries, which contain the mappings between IP addresses and applications. The device searches for DNS entries based on the IP address carried in the application protocol packets to identify the corresponding application.

Configuring SAC and Verifying the SAC Configuration

- To use the SAC function, you need to purchase the corresponding license and enable the in-depth security protection function of the device. After in-depth security protection is enabled, the system automatically loads the built-in signature database file.
- Enable SAC.

```
system-view
  engine enable      //Enable the in-depth security protection function. After the function is enabled, application identification can be performed 3 minutes later.
```

- After enabling SAC, verify the SAC configuration.

```
system-view
  display sa information      //Check the SA status. If the SA status is enabled, in-depth security protection has been enabled.
```

Quiz

1. (True or false) The signature database used by the FPI technology can be manually modified, but the signature database used by the SA technology cannot.
   A. True
   B. False

HUAWEI

- 1. A

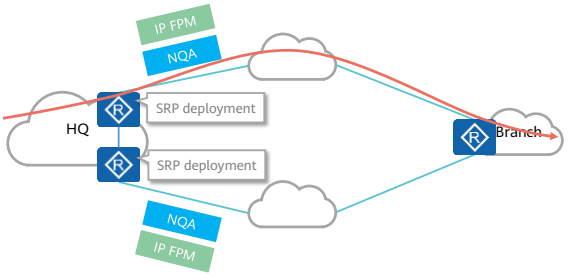
Section Summary

- The main purpose of the SAC technology is to distinguish different services in traffic and help other modules control traffic based on different services.
- SAC consists of the SA and FPI modules.
- The FPI module can identify services based on the traffic characteristics such as the 5-tuple, protocol number, and DSCP value. In addition, the FPI module supports the user-defined signature database.
- The SA module identifies services based on application behaviors and the application-layer signature codes of data packets. The SA signature database does not support user-defined signature databases.

Contents

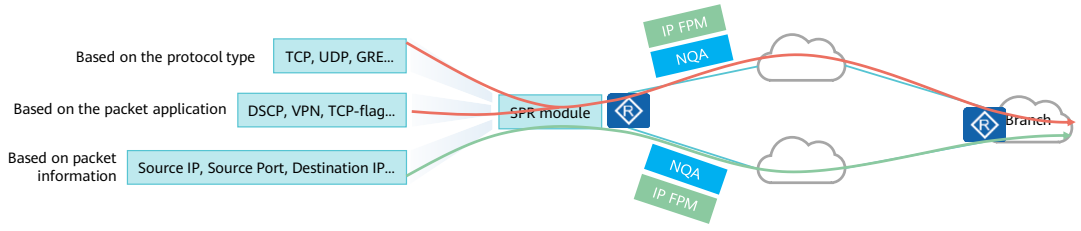
Overview of SPR

- 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.
- Smart Policy Routing (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.



SPR Service Differentiation

- SPR differentiates traffic based on the protocol type, packet application, and packet information.
- Different link quality parameter thresholds can be set for different services. You can set the delay (D), jitter (J), packet loss rate (L), and composite measure indicator (CMI).
- CMI is calculated based on the delay, jitter, and packet loss rate.
- SPR selects routes based on the CMI.

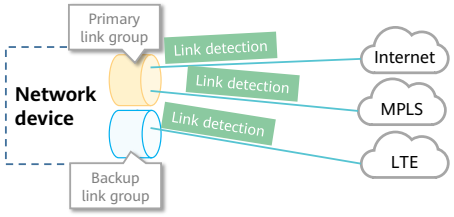


HUAWEI

- SPR classifies services based on the following attributes:
  - Protocol types: IP, TCP, UDP, GRE, IGMP, IPINIP, OSPF, and ICMP
  - Packet applications: DSCP, ToS, IP precedence, fragment, VPN, and TCP-flag
  - Packet fields: Source IP Address, Destination IP Address, Protocol, Source Port, Destination Port, Source IP Prefix, Destination IP Prefix
- When SPR selects routes for services based on the NQA detection result, the CMI is calculated using the following formula:
  - CMI = 9000 – CMI-method. The default value of CMI-method is D + J + L.
  - If NQA is used, a larger CMI value indicates better link quality.
- When SPR selects routes for services based on the IP FPM detection result, the CMI is calculated using the following formula:
  - CMI = D + J + L
  - If IP FPM is used, a smaller CMI value indicates better link quality.

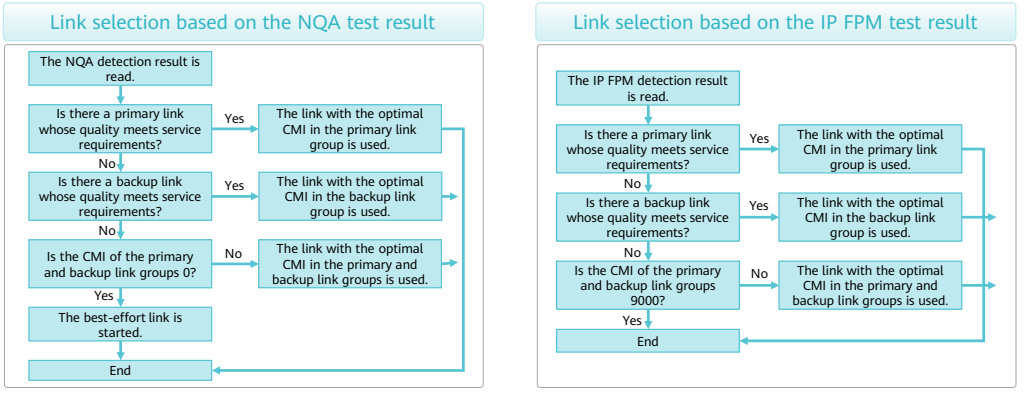
SPR Detection Link and Link Group

- SPR obtains quality indicators of detection links through probes (NQA or IP FPM) and then selects an optimal link.
- A link group can contain one or more detection links.
- SPR defines three roles for links: primary link group, backup link group, and best-effort link. When no suitable link is available in the primary and backup link groups, SPR activates the best-effort link to forward service data.



SPR Link Selection

- SPR periodically obtains the NQA or IP FPM detection result to determine whether a link meets service requirements. If the link does not meet service requirements, a link switchover is triggered.
- The SPR link selection process is as follows:

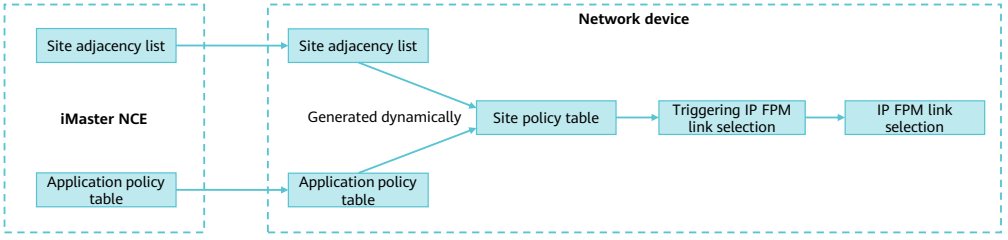


HUAWEI

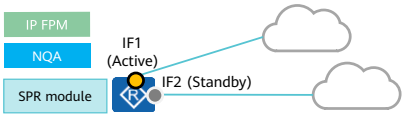
- When a network is unstable, SPR triggers link switchovers frequently, which degrades service experience. SPR provides the flapping suppression function to address this problem.
- The flapping suppression function is disabled by default, and the flapping suppression period is configurable. After traffic is switched to a new link, SPR starts the flapping suppression timer. Within a flapping suppression period, SPR does not perform a link switchover even if it does not obtain the NQA test result indicating that the link meets service requirements within a switchover period. After the flapping suppression timer expires, if SPR still does not obtain the NQA test result indicating that the link meets service requirements within a switchover period, SPR performs a link switchover. If SPR obtains the NQA test result indicating that the link meets service requirements within a switchover period, SPR retains traffic on the link without performing a link switchover.

Using iMaster NCE to Implement SPR

- During SPR deployment through iMaster NCE, to improve the site specifications on the entire network, separation between sites and application policies and traffic-triggered link selection are used together.
- iMaster NCE maintains site adjacency information and application policies, and SPR is configured on routers. Traffic-triggered link selection allows for on-demand generation of SPR configurations. This prevents a large number of configurations from being created on the device and reduces the impact of link selection (based on IP FPM) on the CPU, significantly reducing the burden on the device.
- iMaster NCE can use SAC to classify service traffic based on applications.



Setting SPR Routing Parameters

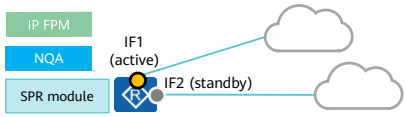


- To configure SPR, set SPR routing parameters and associate SPR with services.
- The configuration roadmap is as follows:
  - Configure an NQA or IP FPM test instance.
  - Bind the test instance to an interface group.

- Set SPR routing parameters.

system-view
  smart-policy-route //Create SPR.
    prober [interface-type interface-num] **nqa** [admin-name] [test-name] //Configure a detection link in SPR and associate it with an NQA test instance.
  nqa-ipfpm link-fail-detect interval [interval-time] **count** [detect-count] //(Optional) The device is configured to trigger an SPR-based switchover after an IP FPM session fails.
    link-group [name] //Create a link group.
      link-member [interface-type interface-num] //Add the specified interface to the link group.

Associating SPR with Services



- To configure SPR, set SPR routing parameters and associate SPR with services.
- The configuration roadmap is as follows:
  - Configure an SPR service profile.
  - Configure SPR interested traffic.
  - Configure the primary link group.

- Associate SPR with services:

```
system-view
  smart-policy-route      //Create SPR.
    service-map [name]    //Create an SPR service profile.
      match [acl | application] [acl-num | app-name]    //Create interested traffic. The interested traffic can be defined using an ACL or SAC.
      set link-group [name]       //Configure the primary link.
```


HUAWEI

Verifying the SPR Configuration

- After the SPR configuration is complete, run the following commands to check the configuration:

```
display smart-policy-route    //Check the SPR routing configuration.
display smart-policy-route service-map [ name ]    //Check the service profile configuration.
display smart-policy-route link-state [ interface-type interface-number ]    //Check the detection link status.
display smart-policy-route nqa-server link-state    //Check the NQA server link status.
```

Quiz

1. (Single-answer question) Which of the following is used by SPR to determine the optimal path?
   A. Jitter
   B. Delay
   C. Packet loss rate
   D. CMI

HUAWEI

- 1. D

Section Summary

- SPR is a type of policy-based routing. It can be associated with NQA and IP FPM so that a path can be quickly selected when a network change occurs.
- SPR selects the optimal path based on the CMI. The CMI calculation method used when NQA is used to detect link quality is different from that used when IP FPM is used to detect link quality.
- SPR can be implemented based on iMaster NCE. Services can be delivered through iMaster NCE, making SPR more convenient.

Summary

- There are many HA technologies. In the past, HA technologies focused on the high reliability of the network layer. With the development of cloud computing, there are more and more service HA requirements.
- New technologies, such as IP FPM, SAC, and SPR, are developed to meet service HA requirements.
- Traditional HA technologies, such as VRRP, BFD, and interface backup, are used together with service HA technologies to improve network reliability.

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.

Introduction to Multi-service Gateways

HUAWEI

Foreword

- Huawei's next-generation NetEngine AR routers, including the AR650, AR6100, AR6200, and AR6300 series apply to different industries, network scales, and scenarios.
- AR routers use leading hardware platforms and software architectures. They provide integrated network solutions to enterprise customers with the minimum investment; therefore, they can meet various application requirements of future business expansion and cope with IT industry development.
- This course describes WLAN and security functions of AR routers.

Objectives

- Upon completion of this course, you will be able to:
  - Describe functions and features supported by AR routers.
  - Describe WLAN service features of AR routers.
  - Describe security service features of AR routers.

Contents

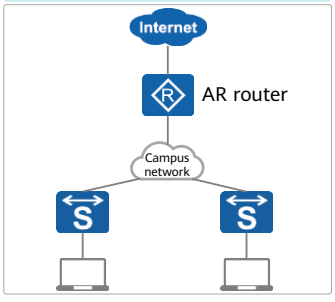
Overview of AR Routers

- AR routers are mainly oriented to enterprise users. As enterprise gateways, AR routers provide all-in-one advantages.
- AR routers provide routing, switching, security, voice, and WLAN functions, reducing initial investment of enterprise users.

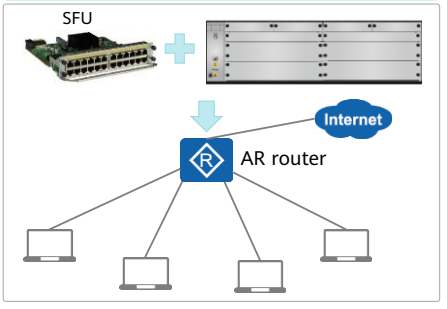
Introduction to Routing and Switching Features

- The AR router can function as the egress router of a medium-sized network. It supports multiple router protocols, MPLS, multicast routing protocols, and WAN interconnection.
- On a small-scale network, an AR router also provides the switching function. For example, after a switching card is installed on an AR router, the AR router can function as a switch and supports multiple Layer 2 technologies such as Virtual Local Area Network (VLAN) and Spanning Tree Protocol (STP).

Usage scenario of routing features

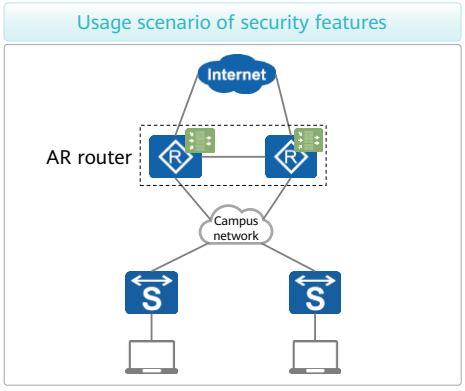


Usage scenario of switching features



Introduction to Security Features

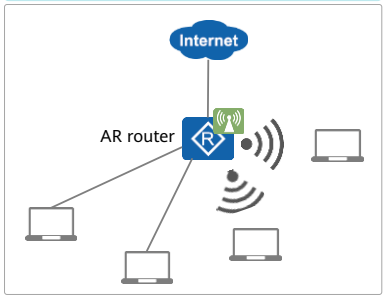
- AR routers provide access security features, such as port security and access authentication. They support network security features, such as firewall, IPS, and URL filtering. In firewall hot standby scenarios, AR routers also support firewall hot standby.



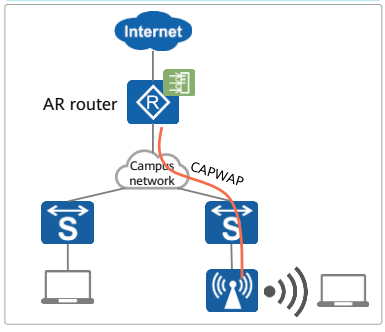
Introduction to WLAN Features

- The AR router supports two working modes: Fat access point (AP) and access controller (AC). The working mode varies according to the usage scenario.
  - An AR router functions as a Fat AP and independently provides WLAN access for stations (STAs).
  - An AR router functions as an AC and provides access to downlink Fit APs. The AR router and Fit APs together provide WLAN access for STAs.

Usage scenario of a Fat AP



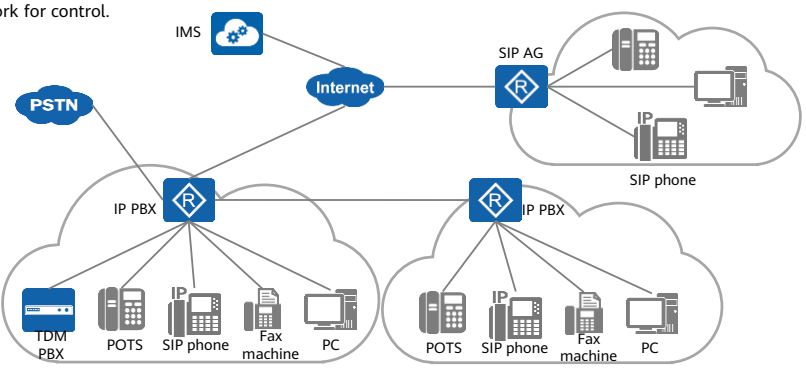
Usage scenario of AC+Fit AP



HUAWEI

Introduction to Voice Features

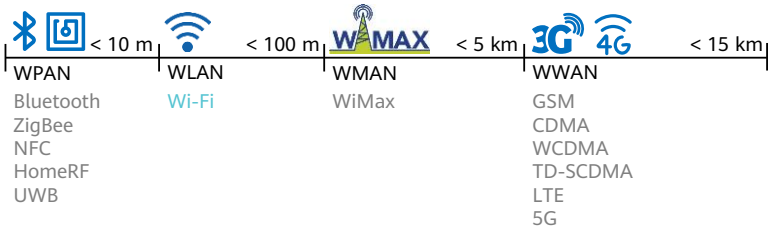
- An AR router can function as an IP PBX or a SIP gateway (SIP AG).
  - As an IP PBX, the AR router can connect to and control voice devices, and can also connect to traditional PBX devices.
  - As a SIP AG, the AR router does not provide access and control functions for voice devices, but can directly connect users to the IMS network for control.



Contents

Introduction to WLAN Technology

- With rapid development of the Internet, the communication network has developed from the traditional cabling network to the wireless network. As one of the wireless networks, the Wireless Local Area Network (WLAN) meets people's demands of mobile office.
- Wireless networks can be classified into the Wireless Personal Area Network (WPAN), WLAN, Wireless Metro Area Network (WMAN), and Wireless Wide Area Network (WWAN) based on the application scope. WLAN technology is also known as Wi-Fi.

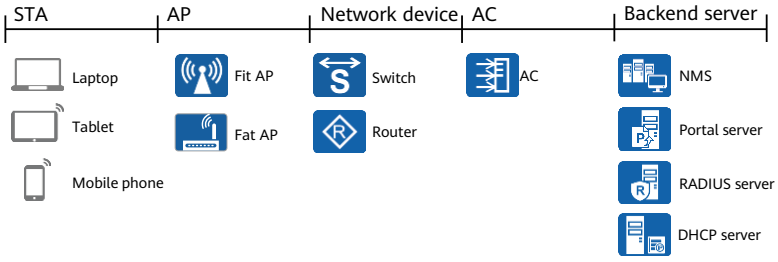


HUAWEI

- WLAN classification:
  - WPAN
  - WLAN
  - WMAN
  - WWAN

Introduction to WLAN Devices

- A WLAN consists of STAs, APs, ACs, network devices, and backend servers.
  - The AP and AC are main devices on the WLAN and are used to control and send wireless signals.
  - Backend servers are used for authentication, IP address distribution, and network monitoring.
  - Network devices are used to forward STA data.
  - The STA is a wireless terminal.



HUAWEI

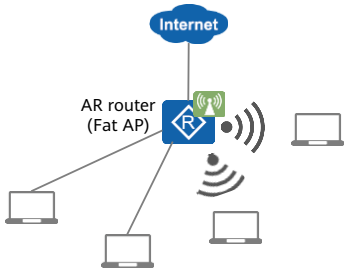
- STA
  - STAs refer to access terminals, including laptops, desktop computers with wireless NICs installed, mobile phones, and PDA.
- AP
  - APs are main devices of the WLAN and key components for wireless technologies. They provide wired connections to upstream devices and wireless access to STAs, bridging the wired and wireless networks.
  - Fat APs are traditional APs. In addition to wireless access, a Fat AP provides security, management, and performance enhancement functions. A Fat AP cannot associate with an AC.
  - Different from traditional Fat APs, Fit APs provide only reliable and high-performance wireless connections. Fit APs must work with ACs.
- AC
  - An AC controls and manages all APs on a WLAN. It can exchange with an authentication server to authenticate WLAN users.

Introduction to AR WLAN Features

- AR routers with a "W" in their name support the WLAN feature., for example, AR6120-VW or AR651W.
- WLAN-capable AR routers have built-in antennas to support Fat AP features.
- An AR router can function as a Fat AP or an AC.
  - AR routers support Fat AP without a license.
  - When the AR650 functions as an AC, no license is required. In this case, the AR650 can connect to a maximum of 16 APs.
  - When the AR6100, AR6200, or AR6300 functions as an AC, it can connect to a maximum of four APs if no license is loaded.

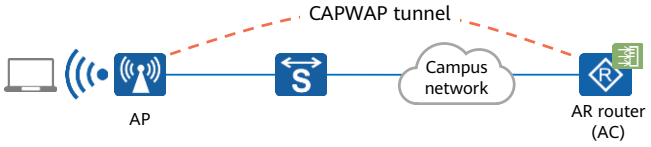
AR Router - Fat AP

- When an AR router functions as a Fat AP, it is similar to a home wireless router. All configurations are performed locally, and wireless signals are directly sent by the AR router.
- Fat APs support pre-shared key, 802.1X, and Portal authentication modes. The built-in Portal server of the AR router can be used for Portal authentication.



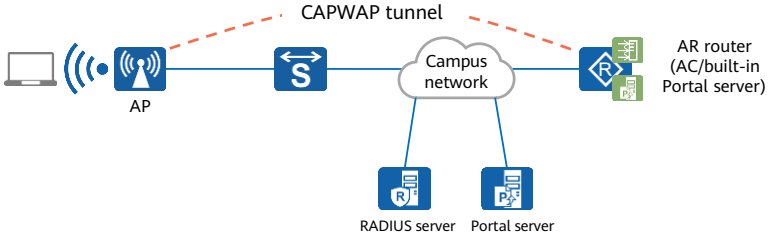
AR Router - AC

- When functioning as an AC, an AR router can connect to most Huawei APs. The working principle is the same as that of a dedicated wireless AC.
- AR routers support the following AC functions: basic wireless functions, roaming, radio management, wireless security, and wireless distribution system (WDS).
- AR routers do not support inter-AC roaming or mesh technologies.



AR Router - NAC

- An AR router provides Wi-Fi access authentication through the built-in network admission control (NAC) feature.
- NAC provides three authentication modes: 802.1X authentication, MAC address authentication, and Portal authentication. NAC needs to be used with the AAA server to implement access authentication.
- An AR router provides 802.1X authentication, MAC address authentication, and Portal authentication. Portal authentication can be implemented by an external or internal Portal server.



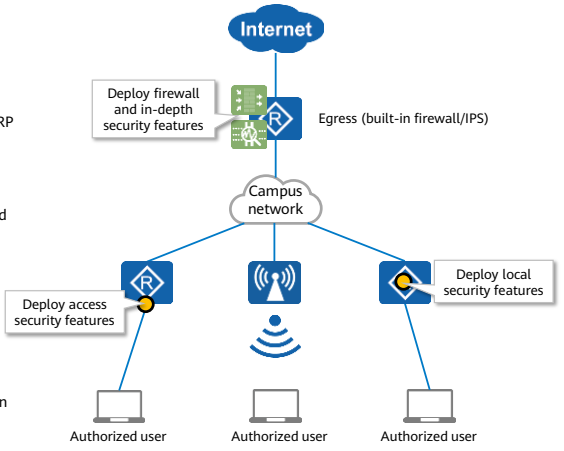
HUAWEI

- AAA servers include the RADIUS server, TACACS server, and accounting server.

Contents

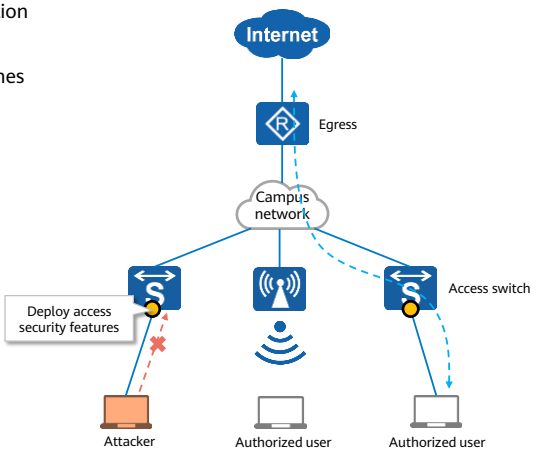
Introduction to AR Security Features

- All AR series routers support security features.
- AR routers support the following security features:
  - Access security
    - Common access security technologies include NAC, port security, ARP security, traffic suppression, and IP source guard (IPSG).
  - Local security features
    - Common local security technologies include CPU attack defense and attack source tracing.
  - Firewall features
    - Common firewall technologies include packet filtering, stateful firewall, and blacklist and whitelist.
  - In-depth security defense
    - Common in-depth security technologies include intrusion prevention system (IPS) and URL filtering.



Introduction to Access Security Features

- The network has become an important tool for information exchange. The cyber security issue needs to be resolved urgently with the wide use of network. The access switches that play an important role on the network are prone to hackers' attacks.
- Access security involves the following:
  - Confidentiality
  - Integrity
  - Availability
- Common access security technologies are as follows:
  - NAC, port security, ARP security, traffic suppression, and IPSG

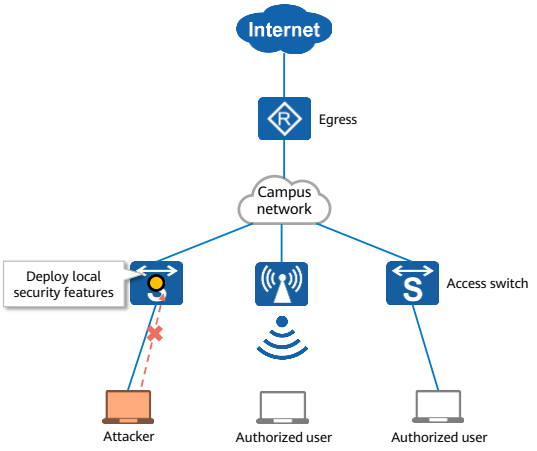


HUAWEI

- Confidentiality: When a switch stores, handles, and transmits data, the data will not be leaked to unauthorized users, entities, or procedures. Information can only be accessed by authorized users.
- Integrity: Data cannot be modified without permission. Information stored or transmitted by switches is free from corruption or loss caused by accidental or malicious deletion, modification, falsification, disordering, relocation, or insertion.
- Availability: The given functions of a switch can be executed under the specified conditions and at the specified time or time period when the required external resources are ensured. Services are always available and meet carrier-class QoS requirements.

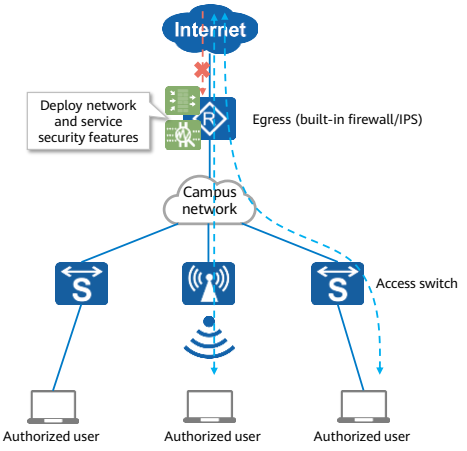
Introduction to Local Security Features

- Device CPUs need to process a large number of packets including valid packets and malicious attack packets on a network.
- To ensure that the CPU can properly process and respond to normal services, the device provides the local attack defense function.
- Common local security technologies are as follows:
  - CPU attack defense and attack source tracing



Introduction to Firewall and In-Depth Security Defense Features

- Access security features are used to protect the internal network and focus more on validity of access devices or users.
- Firewall and in-depth security defense features focus more on validity of traffic between the intranet and extranet and validity of services carried over the traffic.
- Common firewall and in-depth security defense technologies are as follows:
  - Firewall and connection control technology
  - IPS
  - URL filtering
  - Attack defense



Quiz

1. (True or false) Huawei AR routers can be used as Fat APs or Fit APs.

   A. True

   B. False

2. (True or false) Huawei AR routers can be used as firewalls, but do not support firewall hot standby.

   A. True

   B. False

- 1. B
- 2. B

Summary

- AR routers provide routing, switching, security, voice, and WLAN functions.
  - Routing and switching functions: It has similar functions of a Huawei switch.
  - Security functions: It provides similar functions of a Huawei firewall, such as IPS and URL filtering.
  - WLAN function: It can be used as a Fat AP or an AC.
  - Voice capability: It can be used as an IP PBX or SIP AG.

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.

Network Management and Maintenance

HUAWEI

Foreword

- Network management is the most important task of O&M personnel. Typical management tasks include routine network management, new site deployment, and network troubleshooting. As ICT technologies develop, traditional O&M methods are unable to meet customer requirements because they have the following disadvantages:
  - Routine management is performed based on SNMP using network management software. SNMP, however, is inflexible in the cloud computing era.
  - New site deployment is an exhausting task, which is also challenging, especially when a large number of new sites need to be deployed.
  - Network troubleshooting is challenging for most O&M personnel, and how to systematically troubleshoot faults is also an urgent problem to be addressed.
- This course describes the new protocols and methods for routine maintenance, the methods for quick site deployment, and the methods for systematic network troubleshooting.

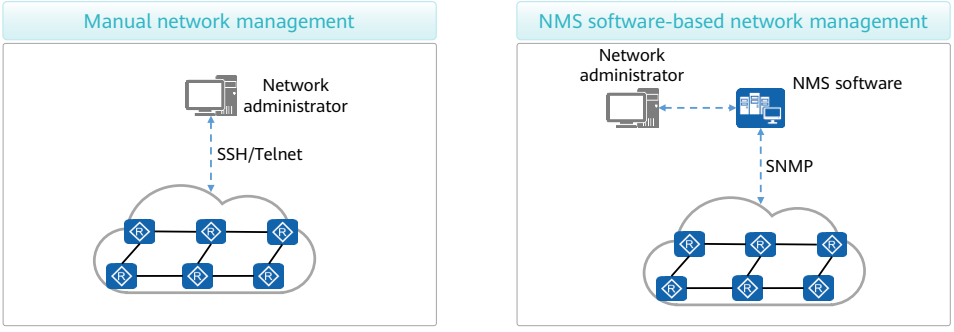
Objectives

- Upon completion of this course, you will be able to describe:
  - The disadvantages of the SNMP protocol.
  - How NETCONF flexibly controls devices.
  - The advantages of using telemetry to collect device status and performance data.
  - Northbound RESTful interfaces of the network management system (NMS) and controller.
  - Zero touch provisioning (ZTP).
  - How to troubleshoot common network faults.

Contents

Traditional Network Management

- Network management involves the use of various protocols, tools, applications, and devices to monitor and control network resources, including hardware and software, so as to meet service requirements and network objectives.
- Traditional network management can be performed manually or using NMS software.

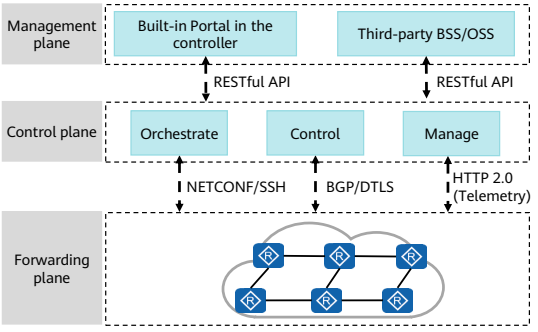


HUAWEI

- Manual network management is convenient but has the following disadvantages:
  - When managing different types of devices from different vendors, network administrators need to memorize a large number of commands and product features, resulting in high O&M costs.
  - Manual management is inefficient and does not apply to large-scale networks.
  - Problems cannot be quickly located.
- NMS software-based network management has the following advantages over manual management:
  - SNMP can be used to manage different types of devices from different vendors, lowering the requirements for network administrators.
  - NMS software can be used to manage large-scale networks.
  - Faults can be located faster.

Network Management in the Cloud Computing Era

- The advent of the cloud computing era brings great changes in network management requirements, which means network management needs to be performed based on content instead of pipes. Network management in the cloud computing era has the following disadvantages:
  - It is inefficient and complex to manage various types of products from different vendors using NMS software, and automation cannot be implemented.
  - NMS software cannot detect network performance accurately.
- Some new protocols and technologies are developed to solve these problems:
  - The NETCONF protocol simplifies network configuration.
  - Telemetry technology greatly improves the accuracy in network performance measurement.
  - RESTful APIs make network management more open.



HUAWEI

- NETCONF is developed based on the Yet Another Next Generation (YANG) model.

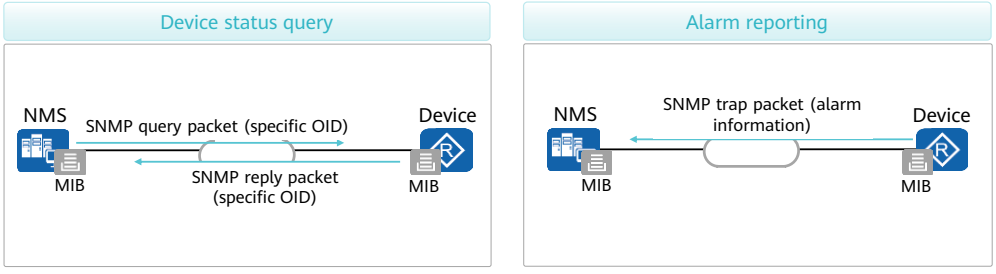
Network Management Technologies

- SNMP is the most commonly used network management protocol. It used to be very efficient, but is unable to meet the current management requirements due to the following defects:
  - SNMP cannot configure network devices efficiently and easily.
  - SNMP has a poor performance in real-time network monitoring.
- To address the preceding two problems, the following technologies are developed:
  - NETCONF: It is designed to offer standard device configuration by using the standard YANG data models. Using NETCONF, the controller can efficiently control devices and quickly deliver configurations.
  - Telemetry: SNMP uses the query-response mechanism, so it is inefficient in reporting network status. Telemetry technology uses subscription to improve efficiency in reporting device status.
- In addition, some enterprises have requirements on the openness of NMS software. To meet these requirements, the NMS software or controller provides open northbound RESTful APIs.

- The reason why SNMP cannot efficiently configure a network is that different vendors have different implementation models for the same type of access devices and the implementation models correspond to different configuration models. As a result, it is difficult to unify standards.
- The reason why SNMP has poor real-time network monitoring performance is that SNMP uses the "query-response" architecture. Frequent queries will cause high CPU usage of network devices.

SNMP Overview

- SNMP builds a data standard between the device and NMS based on the Management Information Base (MIB) data structure. The NMS uses OIDs to query specific information, such as the CPU usage and memory usage. SNMP queries use the request-response model, but subscription is not supported.
- SNMP has certain configuration delivery capabilities. However, it is difficult to implement unified configuration because different vendors have different configuration commands.
- Devices directly report locally generated alarms to the NMS.



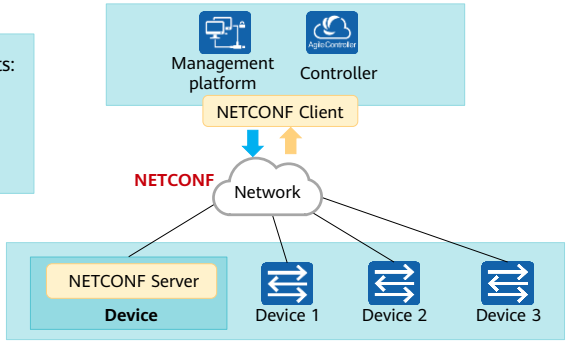
Disadvantages of the Traditional Configuration Mode

- SNMP can deliver configurations based on MIBs, but has the following defects:
  - Configuration is performed separately for each object, not for each service.
  - Devices are configured separately. Network-level configuration and multi-device configuration collaboration are not supported.
  - Binary interfaces are difficult to understand.
  - It is insecure.
- Using the CLI or CLI script for configuration is convenient, but still has the following defects:
  - Different vendors define commands differently. Therefore, scripts need to be customized for different vendors.
  - The output is unstructured data, which is complex to parse, and the CLI scripts are difficult to configure.

NETCONF Overview

- The Network Configuration Protocol (NETCONF) resolves the difficulty in configuring devices using SNMP.
- NETCONF provides a set of mechanism for managing network devices. To be specific, users can use NETCONF to add, modify, and delete configurations of network devices, as well as obtain the configurations and status of network devices.

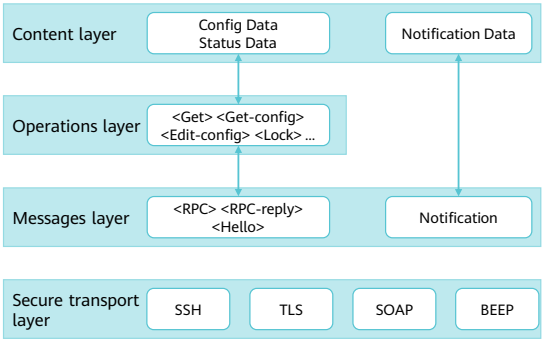
NETCONF has three objects:
- NETCONF client
- NETCONF server
- NETCONF message



- With the development of automated network O&M, automated service deployment, and SDN and NFV technologies, NETCONF and YANG have been considered the basic capabilities of network devices, and also prove to be the best choice for open programming networks.
- NETCONF sessions are carried over SSH and support the heartbeat keepalive and key mechanisms defined by SSH.

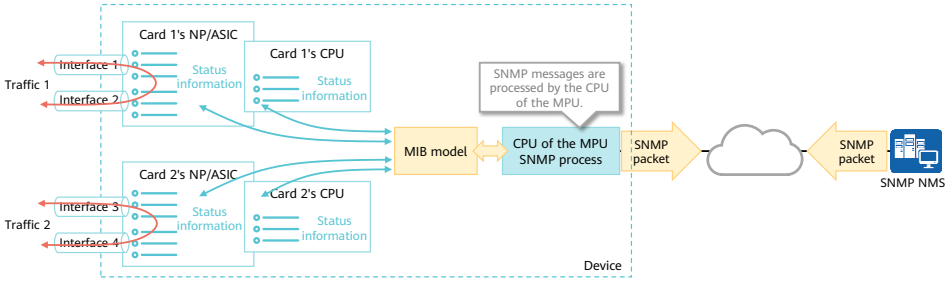
NETCONF Implementation - Protocol Architecture

- NETCONF consists of four layers, which are secure transport layer, messages layer, operations layer, and content layer from bottom to top.
  - The secure transport layer ensures protocol security. Currently, SSH is the most widely used NETCONF secure transport layer protocol.
  - Similar to SNMP, the messages layer provides a mechanism for encoding Remote Procedure Calls (RPCs) and notifications.
  - The operations layer defines the operations for obtaining and editing basic protocol configuration information.
  - The content layer consists of configuration data, status data, and notification data. Data is classified to facilitate cross-device comparison.



Disadvantages of Traditional Network Monitoring

- Traditional network monitoring uses the SNMP protocol to monitor network performance. SNMP obtains network status information in request-response mode. However, when the SNMP NMS requests a large amount of network status information, a large number of SNMP request packets need to be sent, increasing the query time.
- SNMP request and reply packets are processed by the CPU of the device's MPU. Processing a large number of SNMP packets will cause a sharp increase in the CPU usage. On the live network, the smallest SNMP query interval is 5 minutes.
- In the cloud computing era, SNMP cannot meet network performance monitoring requirements.

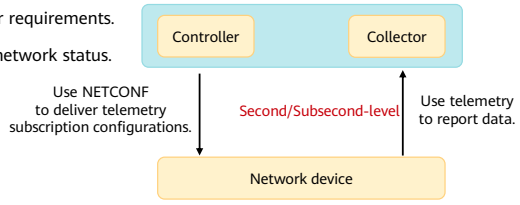


HUAWEI

SNMP | NETCONF | Telemetry | RESTful API

Telemetry Overview

- In the cloud computing era, network status needs to be monitored based on services instead of pipes. In addition, service diversity requires network monitoring to be more flexible.
- Traditional networks use SNMP to monitor network status. However, the interval for reporting network status is too long and therefore the actual network status cannot be collected. Telemetry perfectly addresses the defects of SNMP.
- Telemetry, also known as network telemetry, is a technology for network monitoring, including packet check and analysis, intrusion and attack detection, intelligent data collection, and application performance management. It has the following 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.
  - Leverages big data for data analysis and presentation.



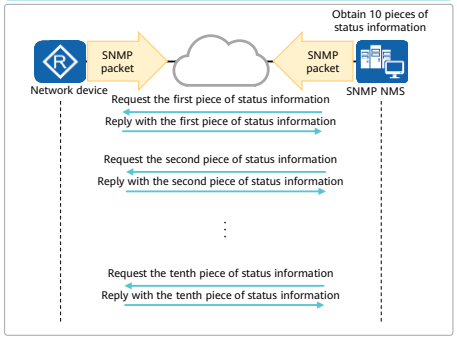
HUAWEI

- With the popularization of networks and emergence of new technologies, the network scale is growing, network deployment is increasingly complex, and users have higher requirements on service quality. To meet user requirements, network O&M must be more refined and intelligent. Network O&M are faced with the following challenges:
  - Ultra-large scale: A large number of devices need to be managed and massive amount of information needs to be monitored.
  - Quick fault locating: Users want faults to be located within seconds or even subseconds on complex networks.
  - Refined monitoring: Various types of data needs to be monitored at a finer granularity to reflect the network status completely and accurately. With the monitoring information, possible faults can be predicted, providing a sound foundation for network optimization. Network O&M involves monitoring not only traffic statistics on interfaces, packet loss on each flow, CPU usage, and memory usage, but also the latency and jitter of each flow, latency of each packet on its transmission path, and buffer usage on each device.
- 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 graphical user interface.
  - 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 need to sample and report.

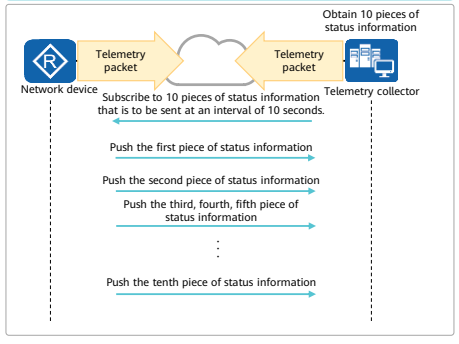
Telemetry Implementation - Subscription/Push Mechanism

- A telemetry collector subscribes to device status in static or dynamic mode, and devices periodically report status information to the collector.
- Telemetry assembles multiple pieces of status information into a packet, reducing bandwidth and hardware resource consumption.

Obtain status information through SNMP

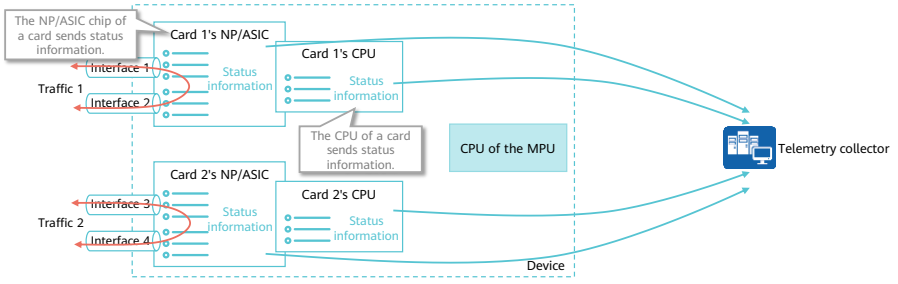


Obtain status information through telemetry



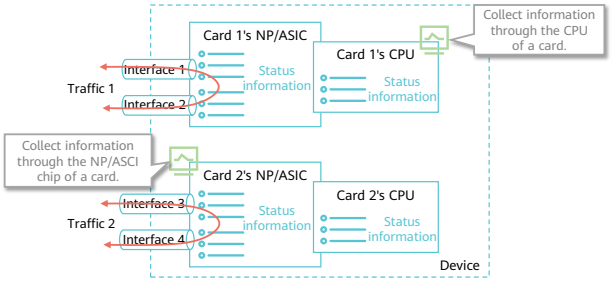
Telemetry Implementation - Distributed Processing

- Telemetry data is no longer processed by the CPU of the MPU. Instead, the device can directly send telemetry data to the collector through the NP/ASIC chip or the CPU of its card. This prevents high CPU usage of the MPU when the device needs to process a large amount of status information.



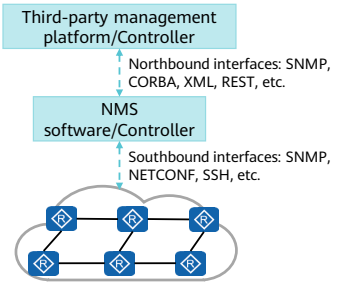
Telemetry Implementation - Data Collection by Hardware Chips

- Status information can be collected through software using the CPU of a card. In this mode, information can be collected every second.
- Status information can be collected through hardware using the NP/ASIC chip of a card. In this mode, information can be collected every 100 ms.



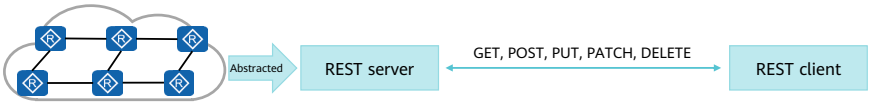
NBI Overview

- Northbound interfaces (NBIs) make network management more flexible. Some customers have self-developed network management platforms. To reduce the learning costs of maintenance personnel, network management software needs to provide open NBIs to facilitate secondary development of customers.
- There are many types of NBIs, such as SNMP, CORBA, XML, TL1 and REST NBIs, among which SNMP and REST NBIs are most widely used on the live network. However, due to many problems of SNMP, more REST interfaces will be used on SDN controllers in the future.



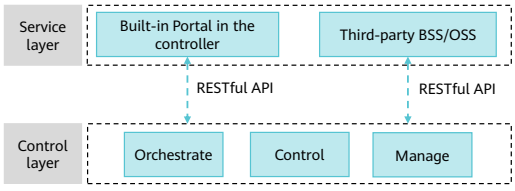
REST Overview

- The Representational State Transfer (REST) service uses the HTTP service as the communication protocol and uses HTTP primitives to express service requests. REST provides the following functions:
  - Presentational: All data generated during network O&M needs to be carried or presented in a corresponding manner. For example, all transactions on a network can be abstracted as resources in the NMS software, and each resource has a unique Uniform Resource Identifier (URI).
  - State transfer: An interaction between a client and server causes resource status change. Typically, REST implements resource status change through the HTTP get, post, put, patch, and delete operations.
- The core of REST is resources and operations. It uses URIs to locate resources and HTTP verbs (get, post, put, patch, and delete) to describe operations. It should be noted that REST is not a standard but an architecture style and design style. It provides only a set of design principles and constraints.



RESTful API Overview

- REST is a style design, based on which RESTful APIs are designed. There are no standard requirements for RESTful APIs and no requirements for a unified URL format. Therefore, RESTful APIs can be defined flexibly.
- External applications can access RESTful APIs via HTTP to implement functions such as service delivery and status monitoring. For security purposes, RESTful APIs provide only HTTPS interfaces.
- In the iMaster NCE solution, RESTful APIs are used to interconnect the control layer with the built-in portal of the controller or a third-party OSS. The interface specifications are as follows:
  - <OP> / <service-path> / <resource> ? <query>
    - OP: operation method, which can be POST, PUT, DELETE, or GET
    - service-path: REST API micro-service path
    - resource: path of a resource to be operated
    - query: parameter set, which is identified using the "name=value" pairs



Quiz

1. (True or False) NETCONF consists of four layers, which are secure transport layer, messages layer, operations layer, and content layer from bottom to top.
   A. True
   B. False
2. (True or False) In telemetry, the CPU of the MPU is used to report device status.
   A. True
   B. False

- 1. A
- 2. B

Section Summary

- SNMP has disadvantages in configuring devices and collecting device status.
- SNMP only supports configuration of each single device. It does not support network-level configuration and cannot implement configuration collaboration among devices, because these are difficult to achieve through programming.
  - NETCONF uses YANG files to translate data into XML language, making network configuration more convenient and flexible.
- The interval at which SNMP collects device status is long, and device status collection causes the CPU usage of the MPU to increase.
  - Telemetry reports device status through subscription, and device status is no longer reported through the CPU of the MPU, reducing the CPU usage of the MPU.
- In the cloud computing era, standards-compliant NBIs need to be developed for the NMS/controller. The HTTP2.0-based RESTful API is a standard open interface that can use HTTP packets to connect to the NMS/controller, simplifying control.

Contents

ZTP Overview

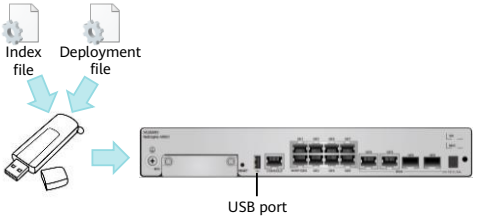
- Traditional network solutions have many pain points, such as high deployment costs and O&M difficulties. This is especially true for enterprises with a large number of geographically dispersed branches. ZTP is a powerful way to address these problems.
- Two ZTP deployment modes are available for AR routers:
  - Traditional deployment mode: used when no controller is available.
    - Auto-Config
    - Auto-Start
    - USB-based deployment
  - Deployment using iMaster NCE-WAN
    - Streamlined USB-based deployment
    - Email-based deployment
    - DHCP-based deployment

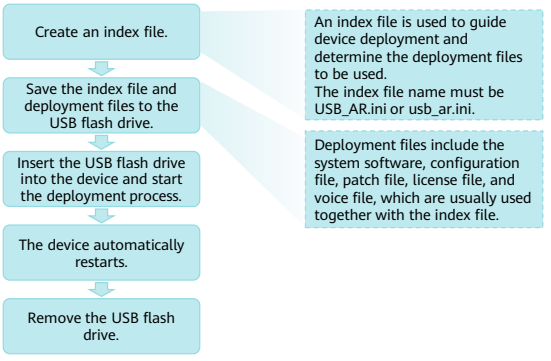
- This course describes iMaster NCE-WAN-based deployment.

USB-based Deployment | Email-based Deployment | DHCP-based Deployment

Streamlined USB-based Deployment

- USB-based deployment allows you to configure devices using a USB flash drive. Before device deployment, save the deployment files in a USB flash drive. After you connect the USB flash drive to a device, the device downloads the files from the USB flash drive to complete version and service deployment. Streamlined USB-based deployment is a type of USB-based deployment.
- After a device starts from the initial state, it automatically downloads the configuration file from the USB flash drive for deployment.



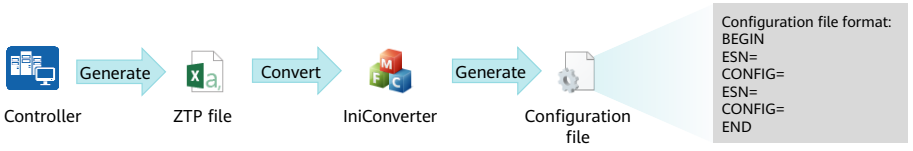


HUAWEI

- USB-based deployment is a basic feature of an AR router and is not under license control.
- Streamlined USB-based deployment can be used only in the SD-WAN solution. Devices with dual MPUs do not support streamlined USB-based deployment.
- Before using an interface on an LPU to perform streamlined USB-based deployment, ensure that the LPU has been registered. If the LPU has not been registered, restart the device before performing simplified USB-based deployment. During the deployment, determine the LPU registration status and deployment status based on the indicator status.
- After the deployment is complete, ensure that the devices are placed based on the mappings between the ESNs and sites. Otherwise, the devices may not be able to communicate with the controller.

Making a Deployment Configuration File

- The process of making a deployment configuration file is divided into two steps:
  - Generate a ZTP file through the controller.
  - Use IniConverter to convert the ZTP file into a configuration file.



Making a Deployment Index File

- Use a blank text file to make a deployment index file, which is typically named **USB_AR.ini**. The index file format is as follows:

```
BEGIN AR
[USB CONFIG]
SN=20180408.070632                  //Data change time identifier, in the format of YYYYMMDD.HHMMSS
EMS_ONLINE_STATE=NO
[UPGRADE INFO]
OPTION=AUTO
DEVICENUM=1
[DEVICE1 DESCRIPTION]
OPTION=OK
ESN=DEFAULT
MAC=DEFAULT
VERSION=DEFAULT
DIRECTORY=DEFAULT
FILENUM=1
TYPE1=SYSTEM-CONFIG-LITE
FILENAME1=ZTP.ini                   //Name of the configuration file used for deployment
END AR
```

The field values in red can be modified, and the other field values must be retained.

Checking the Streamlined USB-based Deployment Result

- Run the **display usb** *usb-id* **autoupdate state** command to check the progress of streamlined USB-based deployment.

```
<Huawei> display usb 1 autoupdate state
Info: Deployment using the USB flash drive is completed successfully.
```

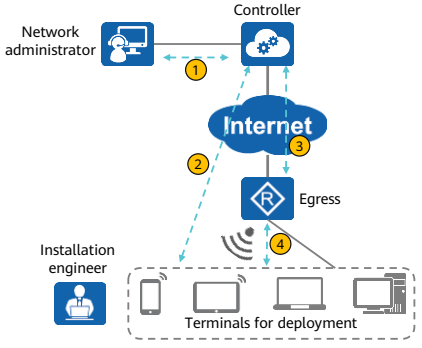
- In addition, you can observe the USB indicator on the device to determine the status of streamlined USB-based deployment.
  - If the USB indicator is **steady on**, streamlined USB-based deployment is completed.
  - If the USB indicator blinks, streamlined USB-based deployment is in progress.

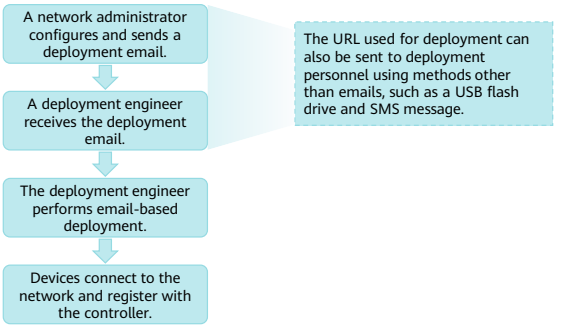
- Enable the USB-based deployment function on the device.
  - [HUAWEI]set usb autoupdate password password # Configure an authentication password for USB-based deployment.
  - [HUAWEI]autoupdate enable # Enable the USB-based deployment function.
- Before running the **autoupdate enable** command to enable the USB-based deployment function, you must run the **set usb autoupdate password** command to configure the authentication password for USB-based deployment.

USB-based Deployment | Email-based Deployment | DHCP-based Deployment

Email-based Deployment

- During email-based deployment, a network administrator specifies uniform resource locator (URL) parameters in a deployment email to configure deployment information on the controller client and then sends the deployment email to a specified deployment mailbox. A deployment engineer receives the deployment email and clicks the URL in the email to start the deployment process. Subsequently, devices automatically complete the deployment.



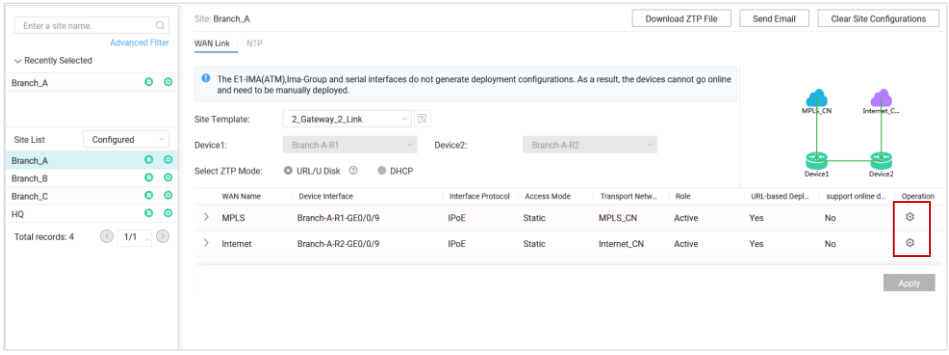


HUAWEI

- Email-based deployment can be used only in SD-WAN solutions.
- Email-based deployment applies only to devices with factory settings.
- Before email-based deployment, users must not log in to the web UI and change the password. Otherwise, the deployment will fail.
- When using the Internet Explorer for email-based deployment, you need to select **Use HTTP1.1** on the tab page displayed after you choose **Internet Options** >**Advanced**. Otherwise, the deployment will fail.

Configuring Email-based Deployment Parameters

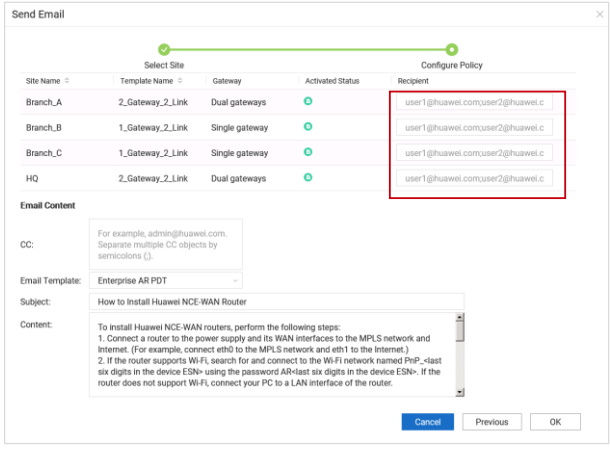
- Configure deployment parameters on the controller.



HUAWEI

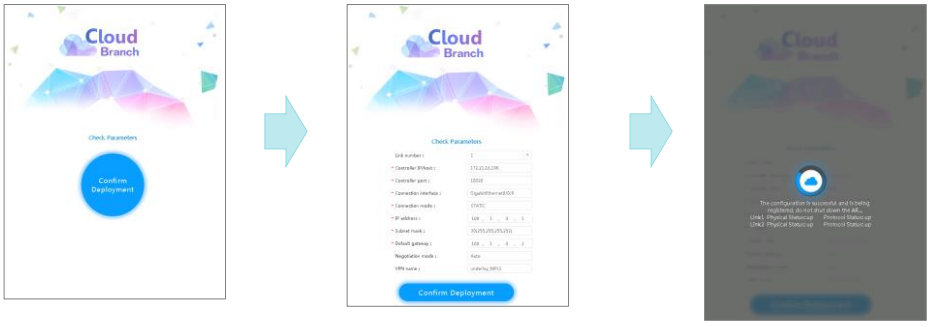
Sending the Deployment Email

- Send a deployment email and download the ZTP file.



Performing Email-based Deployment

- A URL in the following format is used for deployment:
  https://*ip*/portal?ac_host=*ac_host_value*&ac_port=*ac_port_value*...&url_pass=*url_pass_value*

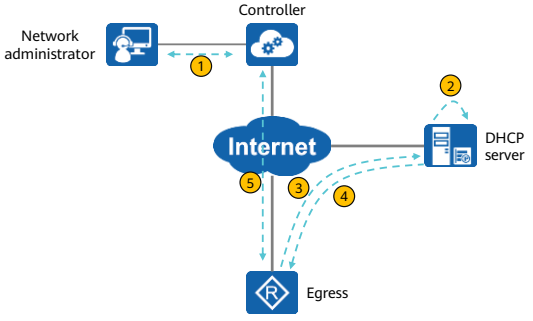


HUAWEI

- The URL in the deployment email is in the following format: https://ip/portal?ac_host=ac_host_value&ac_port=ac_port_value...&url_pass=url_pass_value
- In this format, ... indicates that multiple parameters can be configured, and ip indicates the IP address of the web system. The parameters are separated by ampersands (&). The default web system IP address and subnet mask of the device are 192.168.1.1 and 255.255.255.0, respectively.

DHCP-based Deployment

- DHCP-based deployment is implemented using the optional fields in DHCP packets. The DHCP server is configured with Option 148, which carries the controller address information. Routers obtain the information through DHCP interaction.



1. Configure device interconnection parameters on the controller.
2. Configure DHCP and set the Option 148 field.
3. Devices go online and send DHCP packets to apply for IP addresses.
4. Return DHCP packets that carry the Option 148 field.
5. Devices register with the controller based on the Option 148 field.

Option 148

- The Option 148 field is used to notify devices of the controller IP address and port number. The field format is as follows:
  - option 148 ascii agilemode=tradition;agilemanage-mode=ip;agilemanage-domain=*ip-address*;agilemanage-port=*port-number*;
    - **agilemode**: indicates the device management mode.
    - **agilemanage-mode**: indicates whether a device obtains the URL or IP address of the SD-WAN Controller.
    - **agilemanage-domain**: indicates the URL or IP address of the SD-WAN Controller.
    - **agilemanage-port**: indicates the port number used by the SD-WAN Controller.

Quiz

1. (Multiple-answer question) Which of the following ZTP modes are supported by the device?
   A. USB-based deployment
   B. Email-based deployment
   C. DHCP-based deployment
   D. FTP-based deployment

- 1. ABC

Section Summary

- ZTP makes new site deployment more convenient. ZTP can be implemented in the following ways:
  - Traditional ZTP: Auto-Config, Auto-Start, USB-based deployment
  - iMaster NCE-WAN-based ZTP: streamlined USB-based deployment, email-based deployment, DHCP-based deployment

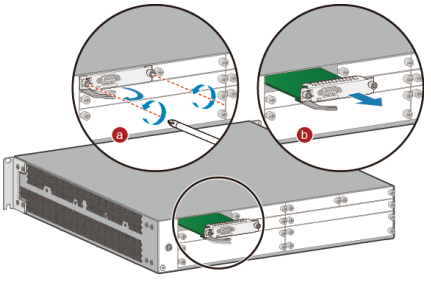
Contents

Device Maintenance Overview

- Stable operating of devices depends on a proper network plan and routine maintenance. The purpose of routine maintenance is to discover and eliminate potential threats.
- Routine maintenance is complex but mainly involves the following tasks:
  - Equipment room environment maintenance: Ensure that the temperature and humidity of the equipment room are within the required range.
  - Device hardware maintenance: Ensure that the devices are clean, with adequate heat dissipation, and the cables and labels comply with the specifications.
  - Device alarm and configuration maintenance: Ensure that the devices are running properly and the configuration files are normal.

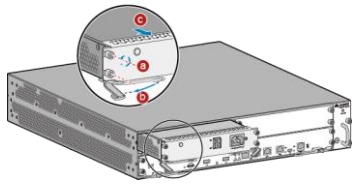
Replacing a Card

- Remove cables from the card to be replaced, and record mappings between cables and ports on the card to ensure correct cable connections on the replacement card.
- Remove the card.
  a. Loosen the captive screws on both sides of the card and rotate the ejector levers outward.
  b. Gently slide the card out of the slot along the guide rails and place it somewhere safe.
- Install the replacement card.
- Connect cables to the replacement card according to the recorded mappings between cables and ports on the replaced card.



Replacing a Power Module

- Check the installation position of the power module to be replaced.
- Turn the power switch on the power module to the OFF position.
- Uninstall the power module after removing the power cable from it.
  a. Use a Phillips screwdriver to loosen the captive screws on both sides of the power module.
  b. Rotate the ejector levers on the power module outward.
  c. Gently slide the power module out of the slot along the guide rails and place it somewhere safe.
- Insert the replacement power module into the slot and connect the power cable to the power module.
- Turn the power switch on the power module to the ON position.
- Use either of the following methods to check whether the new power module is working properly:
  - Check the STATUS indicator on the power module. If the indicator is steady green, the power module is working properly.
  - Run the **display device** command to check the running status of the new power module.

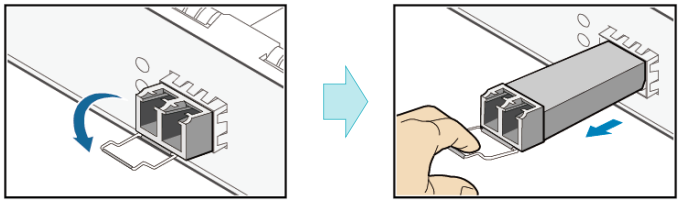


Replacing a Fan Module

- Remove a fan module.
  a. Use a Phillips screwdriver to loosen the captive screws on the fan module to be replaced.
  b. Grasp the handle of the fan module and pull out part of the fan module. Wait until the fans stop running, and then pull the fan module completely out of the slot.
- Install a new fan module.
  a. Hold the handle of the fan module with one hand and support its bottom with the other hand. Slowly slide the fan module into the slot along the guide rails, until it is completely attached to the chassis backplane.
  b. Use the Phillips screwdriver to tighten the captive screws on the fan module.

Replacing an Optical Module

- Mark the positions of the Tx and Rx optical fibers on the bores of the optical module to be replaced, and then remove the optical fibers.
- Cover the removed optical fibers with dust caps.
- Remove the optical module.
- Securely insert the replacement optical module into the optical port.
- Connect the Tx and Rx optical fibers to the optical module based on the marks.
- Check the LINK indicator on the optical port. If the indicator turns green, the new optical module works properly.



Device Alarms and Logs

- During the operating of a device, its running status and configuration change. The device generates different types of information based on these changes.
- A device generates three types of information: logs, traps, and debugging messages.
- Different types of information carry different contents:

Information Type	Content Description
Logs	Logs record user operations, system faults, and system security. There are three types of logs: user logs, security logs, and diagnostic logs. • User logs record user operations and system operating information. • Security logs record security information, including information about user account management, protocols, attack defense, and status. • Diagnostic logs record information used for locating faults.
Traps	Traps are notifications generated when the system detects faults. They record system status information. Unlike logs, traps are time-sensitive because they need to be sent to administrators in a timely manner.
Debugging message	Debugging messages show internal operating information of the system. They are mainly used to trace the device running status. Debugging messages are generated only after the debugging function of a module is enabled.

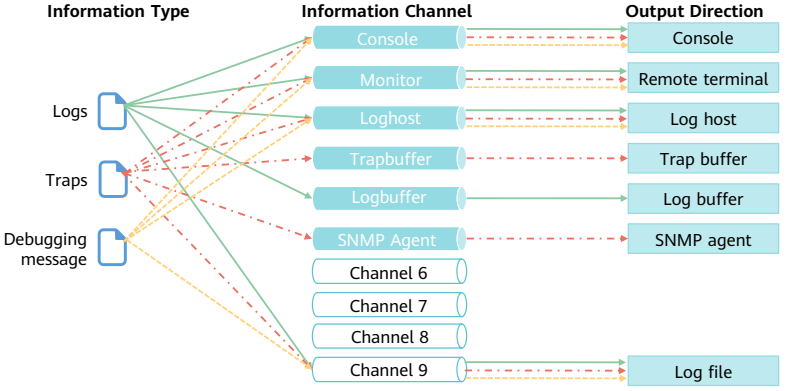
Severity Levels of Device Information

- When a large amount of information is generated, it is difficult to differentiate between information about normal operations and information about faults. By assigning severity levels to information, users can perform a rough analysis of information based on the severity levels and screen out unwanted information.
- There are eight severity levels, each identified by a number. A smaller number indicates a higher severity level.

Number Displayed	Severity Level	Description
0	Emergencies	A fault that makes the device unable to run normally unless it is restarted. For example, the device restarts due to a program exception or an error in memory usage.
1	Alert	A major fault that needs to be rectified immediately. For example, the memory usage of the system reaches the upper limit.
2	Critical	A fault that needs to be analyzed and processed. For example, the memory usage of the device falls below the lower threshold, or BFD detects that the device is unreachable.
3	Error	An incorrect operation or service processing exception that does not affect services but needs to be analyzed. For example, users enter incorrect commands or passwords, or error protocol packets are received.
4	Warning	An anomaly that occurs when a device is operating and requires attention because it may cause service processing faults. For example, a routing process is disabled, BFD detects packet loss, or error protocol packets are detected.
5	Notification	A key operation that is performed to ensure normal operation of the device. For example, an interface is shut down, a neighbor is discovered, or the protocol state machine status changes.
6	Informational	A common operation that is performed to ensure normal operation of the device. For example, the **display** command is run.
7	Debugging	Common information that is generated during normal operation of the device, which requires no attention.

Device Information Output

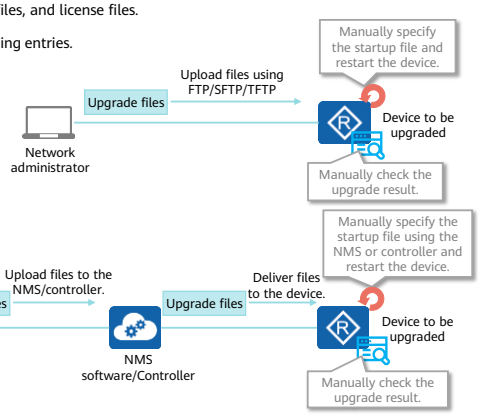
- Information generated by a device can be displayed on a remote terminal, console, log buffer, log file, or SNMP agent. To allow information to be displayed in correct places, the information center defines 10 information channels, which work independently from one another.



- To check logs recorded in the log buffer, run the **display logbuffer** command.
- To check information in log files, run the **display logfile** *[file-name]* command .
- To check traps recorded in the trap buffer, run the **display trapbuffer** *[size value]* command.

Device Software Upgrade

- Upgrading device software can fix some software bugs and add new functions to devices. Device upgrade is a high-risk operation, and you need to make the following preparations before a device upgrade:
  - Back up key files, such as system configuration files, old system software, patch files, and license files.
  - Record network status, such as the number of neighbors and the number of routing entries.
- Software can be upgraded in either of the following modes:
  - Use commands.
    - Upload files (system software, patches, and license files) using commands.
    - Specify startup files (system software, patches, and license files) using commands.
    - Restart a device and check whether the upgrade is successful and whether the network is in normal state.
  - Use the NMS software or controller.
    - Upload the files to the NMS/controller.
    - Upgrade the device system using the controller.
    - Check whether the upgrade is successful and whether the network is in normal state.



HUAWEI

- Specify the system software for the next startup:
  - startup system-software system-file [verify | signature sign-filename]
    - Verify: checks the validity of the system software content.
    - Signature: checks the validity of the digital signature file of the system software.
- Specify the patch file for the next startup.
  - startup patch patch-name

Saving Configuration Files

- During routine maintenance, network device configurations are periodically backed up to ensure that the network can be quickly restored in case of a fault. Configurations can be backed up using any of the following methods:
  - Manual backup: manually saves configuration files on the local PC.

    ```
    <Huawei> save [ all ] [ configuration-file ]
    ```

  - Automatic backup: automatically saves configuration files on the local PC.

    ```
    <Huawei> autosave interval { value | time | configuration time }          //Automatically backs up data at intervals.
    ```

    ```
    <Huawei> autosave time { value | time-value }          //Automatic backs up data at scheduled time.
    ```

  - NMS software/controller-based backup: The NMS software/controller collects device configurations and saves them locally.
- Configuration files can be stored on the local device, FTP server, NMS, or controller.

Quiz

1. (Multiple-answer question) Which of the following files are involved in device upgrade?
   A. System software
   B. Patch file
   C. Routing table file
   D. License file

HUAWEI

- 1. ABD

Section Summary

- Routine maintenance involves the following tasks:
  - Equipment room environment maintenance
  - Device hardware maintenance
  - Device alarm and configuration maintenance

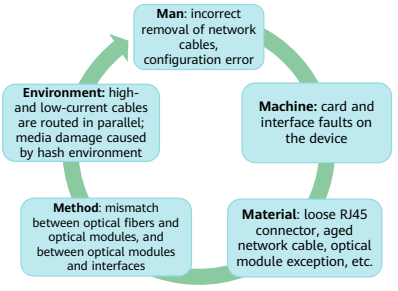
Contents

Common Network Troubleshooting Commands

- A network fault may cause packet loss or service interruption. To diagnose a network fault, you can run the following commands to identify the causes:
  - **display**: displays the device status, entries, and configurations. This command is most commonly used for locating faults.
  - **ping**: used to check network connectivity and preliminarily identify the cause of network interruption.
  - **tracert**: used to check network connectivity and preliminarily identify the cause of network interruption.
  - **debugging**: used to analyze the causes of network faults based on data packets.

Procedure for Troubleshooting an Interface Physically Down Issue

- Hardware and software failures are major causes of an interface physically Down issue.
  - Hardware failures: include failures of hardware such as cards, interfaces, optical modules, fibers, and network cables on the local and remote devices.
  - Software failures: include inconsistent configurations, such as the negotiation mode, rate, and duplex configuration of interfaces on the local and remote devices.
- To troubleshoot the interface physically Down issue, you can use the 4M1E method (Man, Machine, Material, Method, and Environment in sequence) to locate the root cause and rectify the fault.



Interface Physically Down - Man

- Check whether improper operations are performed recently, such as incorrect removal and installation of network cables, loose cable connections caused by accidental touch, and misconfigurations.
- Run the **display interface interface-type** *interface-number* command to check the status of an interface.

```
[Branch-A-1]display interface g0/0/0
GigabitEthernet0/0/0 current state : Administratively  DOWN
Line protocol current state : DOWN
Description:HUAWEI, AR Series, GigabitEthernet0/0/0 Interface
Switch Port, PVID :    1, TPID : 8100(Hex), The Maximum Frame Length is 9600
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 1c20-dbde-7020
Last physical up time   : -
Last physical down time : 2020-04-21 14:07:03
Current system time: 2020-05-13 08:00:21
Port Mode: COMMON COPPER
Speed : 1000,  Loopback: NONE
Duplex: FULL,  Negotiation: ENABLE
Mdi   : AUTO,  Clock   : -
Last 300 seconds input rate 0 bits/sec, 0 packets/sec
Last 300 seconds output rate 0 bits/sec, 0 packets/sec
Input peak rate 0 bits/sec,Record time: -
Output peak rate 0 bits/sec,Record time: -
```


- Recovery method:
  - If the **current state** value is **Administratively down**, the interface is manually shut down. In this case, run the **undo shutdown** command on the interface.
  - If the **current state** value is **Down**, check whether interfaces on both ends of the link have the same rate, duplex mode, and auto-negotiation mode.
    - If the **Negotiation** value is **ENABLE**, the interface works in auto-negotiation mode and the rate and duplex mode are negotiated by interfaces on both ends of the link. If the two interfaces have different rates or duplex modes, run the **restart** command to restart the interfaces so that they can negotiate the rate or duplex mode again. If the negotiation fails, run the **undo negotiation auto** command to configure the interface to work in non-auto negotiation mode. For the rate and duplex mode configurations, see the step below.
    - If the **Negotiation** value is **DISABLE**, the interface works in non-auto-negotiation mode. You can run the **speed { 10 | 100 | 1000 }** and **duplex { full | half }** commands in the interface view to adjust the configurations, so that the interfaces on both ends of the link can have the same rate and duplex mode.

- If the **current state** value is **ERROR DOWN (down-cause)**, check whether the interface is shut down due to an error event. You need to rectify the fault according to the **down-cause** field.

Interface Physically Down - Machine

- If a card or an interface is faulty, the physical status of the interface becomes Down. Therefore, it is important to troubleshoot hardware faults of the device.
- You can run the **display interface brief** command to check brief information about the status and configuration of all interfaces.

```
<HUAWEI> display interface brief
Interface                   PHY   Protocol InUti OutUti   inErrors  outErrors
GigabitEthernet0/0/0        down  down       0%     0%          0          0
GigabitEthernet0/0/1        up    up      0.01%     0%          0          0
GigabitEthernet0/0/2        down  down       0%     0%          0          0
GigabitEthernet0/0/3        up    up      0.01%     0%          0          0
GigabitEthernet0/0/4        down  down       0%     0%          0          0
```

- If many interfaces change from Up to Down, a card fault may occur. If a single interface changes from Up to Down, check whether the interface is damaged. If the interface is in good condition, run the **loopback internal** command to check whether a hardware fault occurs on the interface.

```
<HUAWEI> system-view
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] loopback internal
[HUAWEI-GigabitEthernet0/0/1] display this interface
GigabitEthernet0/0/1 current state : UP
Line protocol current state : UP
```


- Recovery method:
  - Multiple interfaces are physically Down.
    - If multiple interfaces on a same card become physically Down and these interfaces are connected to multiple properly-working peer devices, this card may be faulty. In this case, run the **reset slot** *slot-id* command to reset the card. If the fault persists, replace the card.
    - If multiple interfaces become physically Down and these interfaces connect the local device to the same peer device, the peer device may be faulty. You need to log in to the peer device to check whether the corresponding card is faulty. If so, reset or replace it. If an active/standby switchover is performed on the peer device, or if the peer device is in sleeping state or restarted, it is a normal situation that these interfaces become physically Down. These interfaces will automatically change to the Up state once the peer device becomes stable.
  - A single interface is physically Down.
    - If only one interface is physically Down, check whether the interfaces at both ends of the link have hardware faults such as locking tab dents. If the interface is faulty, use another idle interface.

Interface Physically Down | Ping Failure | Packet Loss | Password Recovery

Interface Physically Down - Material

- Physical interfaces can be electrical or optical. The transmission media include RJ45 network cables, optical modules, and optical fibers. When a transmission medium is aged or damaged, or when the optical module fails to transmit or receive optical signals, the interface goes Down physically.
- Perform the following steps to check whether an electrical interface is faulty:
  - Check whether cables are connected correctly.
  - Check whether the cable length and specification comply with related standards.
  - Check the status of wire pairs in a cable.
    - Run the **virtual-cable-test** command.

```
[HUAWEI] interface gigabitethernet 0/0/1
[HUAWEI-GigabitEthernet0/0/1] virtual-cable-test
Pair A length: 1meter(s)
Pair B length: 1meter(s)
Pair C length: 1meter(s)
Pair D length: 1meter(s)
Pair A state: Ok
Pair B state: Ok
Pair C state: Ok
Pair D state: Ok
```

- Perform the following steps to check whether an optical interface is faulty:
  - Check whether the optical module is Huawei-certified.
  - Check whether the optical module is in normal state.
    - Run the **display transceiver interface** command.

```
<HUAWEI> display transceiver interface gigabitEthernet 2/0/3 verbose
Diagnostic information: Temperature (° C) :39
Voltage(V):3.32
Bias Current(mA):6.91          //Current
Bias High Threshold(mA):33.34
Bias LowThreshold(mA):1.67
Current Rx Power(dBM):-4.59  //Receive power of the interface
Default Rx Power High Threshold(dBM):0.00
Default Rx Power LowThreshold(dBM):-16.99
Current Tx Power(dBM):-5.10 //Transmit power of the interface
Default Tx Power High Threshold(dBM):0.00
Default Tx Power LowThreshold(dBM):-12.50
```


HUAWEI

- Recovery method:
  - Connection media failures of electrical interfaces
    - If the network cable connected to an electrical interface is faulty, replace the network cable.
  - Connection media failures of optical interfaces
    - If the transmit power is too high or too low, replace the optical module.
    - If the receive power is too high or too low, use an optical power meter to measure the attenuation of each part of the link and fix the part where an exception is detected. If the fault persists, replace the optical module or optical fiber. Ensure that the transmission distance of the optical module and the type of the optical fiber meet the networking requirements.

Interface Physically Down - Method

- When performing operations on the connection media between devices, you need to check whether the optical module matches the optical interface, whether the optical module matches the optical fiber model, and whether the optical module is correctly connected to the optical fiber.
  - Run the **display transceiver interface** command to check whether the optical module parameters at both ends of the link are consistent.

```
<HUAWEI> display transceiver interface GigabitEthernet
0/0/1 verbose
GigabitEthernet0/0/1 transceiver information:
Common information: Transceiver
Type:1000_BASE_SX_SFP
Connector Type:LC
Wavelength(nm):850
Transfer Distance(m):500(50um),300(62.5um)
Digital Diagnostic Monitoring:YES
Vendor Name:HUAWEI
Vendor Part Number:02315204
```

  - Run the **display interface** *interface-type interface-number* command to check whether the interface is a combo interface and whether it is working in the correct mode.

```
<HUAWEI> display interface gigabitethernet 1/0/1
GigabitEthernet1/0/1 current state : DOWN
Line protocol current state : DOWN
Description:HUAWEI, Quidway Series,
GigabitEthernet1/0/1 Interface
.....
Port Mode: COMBO AUTO
Current Work Mode: COPPER
Speed : 100, Loopback: NONE
```


- Recovery method:
  - Ensure that the optical module in use matches the optical interface and optical fiber. Replace them if they do not match.
  - If the interface is a combo interface, ensure that the working mode of the combo interface is the same as the actual working mode of the interface.

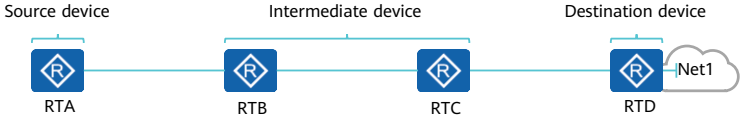
Interface Physically Down - Environment

- Network devices are used in various scenarios and can be located at any layer of a network. When a device and its connection media are abnormal due to environmental factors, its interface may also be physically Down.
- Check the temperature and humidity of the equipment room.
  - For most network devices, the operating temperature is in the range 0℃ to 45℃, and the relative humidity is in the range 5% RH to 95% RH (non-condensing).
- Check cabling.
  - Power cables and service cables are routed separately. Low-current cables must be separated from high-current cables to prevent interference. Check whether network cables are damaged or aged out because of high temperature or rack wear.

- Recovery method:
  - If the temperature in the equipment room fails to meet the requirement for a long time, repair or replace the air conditioning system. If the relative humidity of the equipment room is high, install a dehumidifier. If the relative humidity of the equipment room is low, install a humidifier.
  - If high-current and low-current cables are routed in parallel, re-route the cables to separate them. If network cables are damaged or aged out due to environmental factors, replace the cables.

Troubleshooting a Ping Failure

- A ping failure occurs when a device fails to receive any ping response packet due to reasons such as a link fault or ARP learning failure.
- The ping operation involves three types of devices: source device, intermediate device, and destination device.



- When the source device RTA fails to ping Net1, it is difficult to determine the cause of the fault. In this case, you can narrow down the fault scope. To be specific, ping RTB, RTC, and RTD one by one from RTA. If you still cannot locate the network segment where the fault occurs, ping RTC and RTD from RTB, and so forth. Stop the ping operation when you can determine the network segment where the fault occurs.

Checking Whether the Ping Command Is Correct

- The ping command provides many parameters. You can select appropriate parameters based on factors such as the detection purpose, network type, and network status.
- Check whether the **ping –f** *ip-address* command is run on the device. If the command is run, ICMP packets cannot be fragmented, so you need to check the MTU on the outbound interface of the link.
- For details about the ping command format, see the description of the ping command format in the product documentation.

Checking the Status of the Physical Link

- Check the indicator status on an interface. If the indicator is off, the interface is not connected. In this case, try another interface or network cable.
- Check whether an optical fiber or network cable is connected to correct interfaces as required in the network deployment plan. If not, connect it to the required interfaces.
- Ensure the wavelengths of optical modules used at both ends are consistent. It is recommended that Huawei-certified optical modules be used.
- For details about how to rectify the fault when a physical link goes Down, see the part "Troubleshooting the Interface Physically Down Issue".

Checking the Routing Table

- Before checking the routing table, you need to check whether the terminal is configured with the correct gateway address.
- The routing table is the most important basis for data forwarding. You can run the **display ip routing-table** *ip-address* command to check whether there are correct routes on the device.

```
<Branch-A-1>display ip routing-table
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
        0.0.0.0/0   Static  60   0           RD  202.1.2.2
GigabitEthernet0/0/9
  172.21.16.197/32  Static  60   0           RD  202.1.2.2
GigabitEthernet0/0/9
  172.21.16.198/32  Static  60   0           D   202.1.2.2
GigabitEthernet0/0/9
      202.1.2.0/30  Direct  0    0           D   202.1.2.1
GigabitEthernet0/0/9
      202.1.2.1/32  Direct  0    0           D   127.0.0.1
GigabitEthernet0/0/9
      202.1.2.3/32  Direct  0    0           D   127.0.0.1
GigabitEthernet0/0/9
```

- In addition to the routing table, policy-based routing also affects data forwarding. Policy-based routing takes priority over the routing table. You can run the **display traffic-policy applied-record** command to check the application information of traffic policies.

```
<SwitchB> display traffic-policy applied-record
-----------------------------------------------------
  Policy Name:   p1
  Policy Index:  0
    Classifier:c1     Behavior:b1
-----------------------------------------------------
 *interface GigabitEthernet1/0/2
    traffic-policy p1 inbound
slot 1    :  success
-----------------------------------------------------
  Policy total applied times: 1.
```

Checking Whether ARP Entries Are Correctly Learned

- ARP entries play a decisive role in data encapsulation. Incorrect ARP entries will result in data encapsulation error or failure.
- Run the **display arp all** command to check whether the ARP entry of the directly connected address is correctly learned.

```
<Huawei> display arp all
IP ADDRESS      MAC ADDRESS     EXPIRE(M) TYPE  VLAN INTERFACE      VPN-INSTANCE
--------------------------------------------------------------------------------
192.168.1.10    4c1f-cc17-1ca5                I -                 Vlanif10
192.168.1.11    4c1f-cc2f-3634  10            D-0                 GE1/0/1
192.168.40.1    incomplete
--------------------------------------------------------------------------------
Total:2         Dynamic:1       Static:0      Interface:1
```

- If **MAC ADDRESS** is **Incomplete** in the output, the ARP entry is temporary and the correct ARP entry is not learned. An ARP entry is learned successfully if a correct MAC address is displayed under **MAC ADDRESS**.
- Run the **display mac-address** *interface-type interface-number* command to view the MAC address entry and check whether the outbound interface of the entry is the same as that in the ARP entry. If they are different, check whether a loop or MAC address conflict occurs.

Checking Whether ICMP Packets Are Sent and Received Correctly

- Ping packets may be damaged during transmission. You can run the **display icmp statistics** command to check statistics about ICMP packets.

```
<Huawei> display icmp statistics
  Input: bad formats          0     bad checksum               0
         echo                 0     destination unreachable    0
         source quench        0     redirects                  0
         echo reply          25     parameter problem          0
         timestamp request    0     information request        0
         mask requests        0     mask replies               0
         time exceeded        0     timestamp reply            0
         Mping request        0     Mping reply                0
  Output:echo                25     destination unreachable    0
         source quench        0     redirects                  0
         echo reply           0     parameter problem          0
         timestamp request    0     information reply          0
         mask requests        0     mask replies               0
         time exceeded        0     timestamp reply            0
         Mping request        0     Mping reply                0
```


- Check whether the value of **bad checksum** before and after a ping operation. If the value increases, check whether the ICMP packets returned by the protocol stack on the remote device have the correct format.
- If the values of **echo** and **echo reply** are the same, but the ping still fails, check ICMP packet statistics to determine whether ICMP packets are successfully sent and received.
- If the values of **echo** and **echo reply** are different:
  - If the number of echo request packets sent by the local device is smaller than the number of packets sent by the ICMP module, packets are discarded by the local device.
  - If the number of echo request packets sent by the local device is larger than the number of echo request packets received by the remote device, packets are discarded during transmission.
  - If the number of echo request packets sent by the local device is the same as the number of echo request packets received by the remote device, but the number of echo reply packets sent by the remote device is smaller than the number of echo request packets received by the remote device, packets are discarded by the remote device.

Checking Whether Ping Packets Are Discarded Due to CPCAR Exceeding

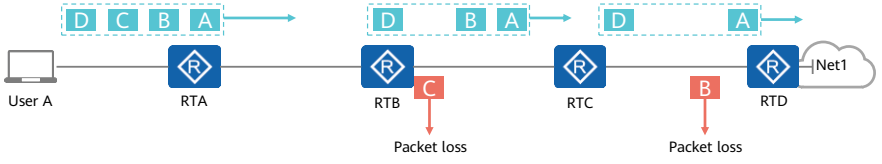
- Control Plane Committed Access Rate (CPCAR) limits the rate of protocol packets sent to the CPU based on the protocol type, preventing excess packets of a protocol from being sent to the CPU.
- Ping packets may be discarded when the packet rate exceeds the CPCAR value. You can run the **display cpu-defend statistics packet-type icmp all** command to check the number of ping packets discarded due to CPCAR being exceeded.

```
<HUAWEI> display cpu-defend statistics packet-type icmp all
Statistics on mainboard:
-------------------------------------------------------------------------------
Packet Type       Pass(Bytes)  Drop(Bytes)  Pass(Packets)  Drop(Packets)
-------------------------------------------------------------------------------
icmp                     4488            0             44             0
-------------------------------------------------------------------------------
Statistics on slot 3:
-------------------------------------------------------------------------------
Packet Type       Pass(Bytes)  Drop(Bytes)  Pass(Packets)  Drop(Packets)
-------------------------------------------------------------------------------
icmp                        0            0              0             0
-------------------------------------------------------------------------------
```


- If the **Drop** value keeps increasing during the ping operation, packets are discarded due to CPCAR being exceeded. Increase the CPCAR value and perform a ping test again to check whether the fault is rectified.

Troubleshooting Packet Loss

- When packet loss occurs, determine the location where packets are dropped, analyze the cause of the packet loss, and then rectify the fault accordingly.
- Symptoms of packet loss:
  - The Internet access speed is unstable. Web pages are displayed at a low speed, or some parts of web pages or even the entire web pages cannot be displayed sometimes.
  - Pixelation, artifacts, or frame freezing occurs when a user watches video.
  - Users are frequently logged out of instant messaging tools such as WhatsApp, or login times out.
  - Files are downloaded at a low speed.



Checking the Status of the Physical Link

- Check the indicator status on an interface. If the indicator is off, the interface is not connected. In this case, try another interface or network cable.
- Check whether an optical fiber or network cable is connected to correct interfaces as required in the network deployment plan. If not, connect it to the required interfaces.
- Ensure the wavelengths of optical modules used at both ends are consistent. It is recommended that Huawei-certified optical modules be used.
- For details about how to rectify the fault when a physical link goes Down, see the part "Troubleshooting the Interface Physically Down Issue".

Checking for CRC Errors in the Inbound Direction of an Interface

- A Cyclic Redundancy Check (CRC) is an error-detecting code commonly used on communications devices to determine if a block of data has been corrupted.
- Run the **display interface** *interface-type interface-number* command to check the CRC error count. If the CRC error count keeps increasing, packet loss is caused by a link or device fault.

```
[Huawei]display interface g0/0/0
...
Last 300 seconds input rate 0 bits/sec, 0 packets/sec
Last 300 seconds output rate 0 bits/sec, 0 packets/sec
Input peak rate 0 bits/sec,Record time: -
Output peak rate 0 bits/sec,Record time: -

Input:  0 packets, 0 bytes
  Unicast:              0,  Multicast:              0
  Broadcast:            0,  Jumbo:                  -
  Discard:              0,  Total Error:            0

  CRC:                  0,  Giants:                 0
  Jabbers:              0,  Throttles:              0
  Runts:                0,  Symbols:                0
  Ignoreds:             0,  Frames:                 0
```

Checking the Number of Discarded Packets in the Outbound Direction of an Interface

- The **Discard** field records the number of packets discarded on an interface. The most common cause of packet loss on an interface is that the interface is congested.
- Run the **display interface** *interface-type interface-number* command to view the number of discarded packets in the outbound direction of the interface.

```
[Huawei]display interface g0/0/0
GigabitEthernet0/0/0 current state : DOWN
Line protocol current state : DOWN
.....
Output:  0 packets, 0 bytes
  Unicast:              0,  Multicast:              0
  Broadcast:            0,  Jumbo:                  -
  Discard:              0,  Total Error:            0
  Collisions:           0,  ExcessiveCollisions:    0
  Late Collisions:      0,  Deferreds:              0

    Input bandwidth utilization threshold : 100.00%
    Output bandwidth utilization threshold: 100.00%
    Input bandwidth utilization :    0%
    Output bandwidth utilization :   0%
```


- If packet loss occurs, use the following methods to optimize the network:
  - Increase the port capacity.
    - If multiple flows conflict, increase the link bandwidth between devices or add more ports to the Eth-Trunk for load balancing.
  - Configure traffic rate limiting or shaping in the inbound direction.
    - Traffic burst is the major cause of unexpected packet loss. When the burst packet size exceeds the port buffer limit, service packets are discarded, affecting user services. Configuring traffic rate limiting or traffic shaping on the upstream device can relieve traffic burst or reduce the burst packet size, lowering the possibility of congestion-triggered packet loss on downstream devices.
  - Configure differentiated services on an interface, so that key services are placed in high-priority queues and processed preferentially in case of congestion.
    - Typically, an interface may carry many services, including high-priority services (such as voice and video) and low-priority services (such as Internet access). You can specify different priorities for high-priority services on the upstream device or configure priority mapping in the inbound direction of a device, so that key services enter the high-priority queues. Configuring PQ scheduling in the outbound direction can ensure that high-priority services are scheduled first.
  - If multicast services are configured on a device, adjust the packet sending mode on the multicast source server to mitigate the congestion.

Checking for Loops

- Loops are the most common and difficult-to-detect factor that leads to packet loss. You can check for loops using any of the following methods:
  - Run the **display interface brief | include up** command to check the traffic on all interfaces in Up state.

```
<Huawei> display interface brief | include up
Interface                   PHY   Protocol  InUti OutUti   inErrors  outErrors
GigabitEthernet0/0/2        up    up        76%   76%          0         0
```

  - Run the **display interface** *interface-type interface-number* command to check whether there are a large number of broadcast packets on the interface.

```
[Huawei]display interface g0/0/0
Input:  0 packets, 0 bytes
  Unicast:              0,  Multicast:              0
  Broadcast:            0,  Jumbo:                  -
  Discard:              0,  Total Error:            0
```

  - Run the **display cpu-usage** command to check whether the CPU usage exceeds 80%.

```
[Huawei]display cpu-usage
Control Plane
   CPU Usage:  5.3%   Max: 50.7%
   User:  0.0%   System:  0.0%   SoftIrq:  0.0%   HardIrq:  0.0%   Idle: 94.7%
   CPU utilization for ten seconds:  5.3%  one minute:   4.0%  five minutes:   4.0% ..
```


- For a Layer 2 network, run the **display mac-address flapping record** command to check whether MAC address flapping occurs, so as to determine whether there are loops.
- Choose from the following methods based on the loop information and networking.
  - Observe interface indicators and collect traffic statistics on interfaces to locate the interfaces experiencing broadcast storms.
  - Check the devices hop by hop according to the topology to locate the devices that cause the loop.
  - Locate the interfaces that cause the loop and shut down the interfaces to remove the loop. Shutting down interfaces can prevent loops, but you need to determine the cause of a loop after the loop is eliminated.

Checking for Attacks

- When undergoing an attack, a device is busy processing requests from the attack source, and therefore drops packets of other services.
- Common network attacks, such as ARP, ARP Miss, and DHCP attacks, can cause a high CPU usage on a device. These attacks are all initiated by sending a large number of protocol packets; therefore, packet statistics on the device show a large number of packets are sent to the CPU.
- Run the **display cpu-defend statistics** command to view statistics about the packets sent to the CPU, and determine whether too many protocol packets are discarded due to timeout.

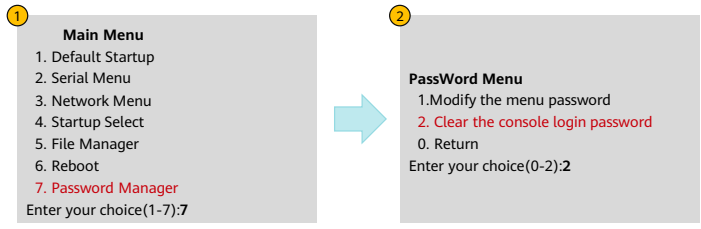
```
[Branch-A-1]display cpu-defend statistics
-------------------------------------------------------------------------------
Packet Type             Pass Packets       Drop Packets
-------------------------------------------------------------------------------
8021X                   0                  0
8021X-first             0                  0
arp-miss                1                  0
arp-reply               3189               0
arp-request             1106               0
...
```


- You can use the attack source tracing function to locate the attack source, and then create and apply a local attack defense policy based on the attack source.

Interface Physically Down | Ping Failure | Packet Loss | Password Recovery

Password Recovery

- You can log in to a device through the console port or Telnet. If login fails in either mode, use the other method and then change your password. If console- and Telnet-based login both fails, change your password through BootROM.
- Connect to the device using a serial cable, and restart the device. When the message "Press Ctrl+B to break auto startup ..." is displayed, press **Ctrl+B**, and then enter the password to access the BootROM main menu.



- If the console login password, Telnet login password, and BootROM password are all forgotten, contact Huawei technical support.

HUAWEI

- To access the BootROM menu, you must restart the device, which results in service interruption. Migrate services to a backup device and perform this operation during off-peak hours.
- If you have cleared the console login password, configure a new password immediately after login. Otherwise, you must clear the console password again to log in if the login timeout timer expires or the device restarts.
- Do not power off the device during the operation.
- In V200R003C01 and earlier versions, the default BootROM password is **huawei**. In V200R005C00 and later versions, the default BootROM password is **Admin@huawei**.

Summary

- In the cloud computing era, SNMP is no longer suitable for managing large-scale networks due to its disadvantages in configuration and status collection functions. The NETCONF protocol and telemetry technology are developed to solve these problems. NETCONF facilitates device configuration on the NMS and controller, whereas telemetry technology accelerates device status collection.
- When a large number of sites need to be deployed, the traditional deployment mode requires a large amount of manpower. ZTP is the best choice in this scenario as it can be used to quickly deploy devices.
- Routine network maintenance includes equipment room environment maintenance, device hardware maintenance, and device alarm and configuration maintenance.

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.

SD-WAN Solution Technology Overview

HUAWEI

Foreword

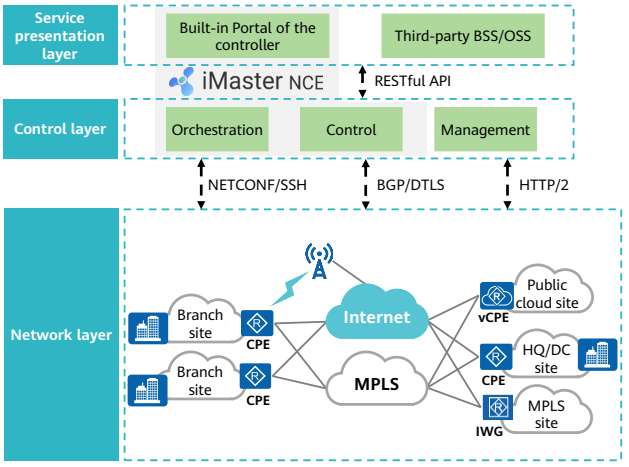
- As digital transformation of enterprises steps into the cloud era and the software-defined networking (SDN) era is coming, software-defined WAN (SD-WAN) has become a next-generation solution for enterprise branch interconnection. It is widely deployed by enterprises due to its key features such as network-agnostic, intelligent traffic steering, zero touch provisioning (ZTP), and visualization.
- There are many vendors in the SD-WAN field. According to statistics of Gartner, more than 60 vendors in the world have provided multi-layer SD-WAN solutions that support multiple business models by the end of 2019.
- SD-WAN is deemed to be a next-generation enterprise branch interconnection solution by industry analysts, which has profound impact on and preliminarily replace existing traditional MPLS VPN services. In the future, SD-WAN will continue to develop and evolve rapidly oriented to emerging technologies and solutions such as cloud, 5G, XGSPON, artificial intelligence (AI), and blockchain.

Objectives

- On completion of this course, you will be able to:
  - Describe the architecture and components of Huawei SD-WAN Solution.
  - Describe the basic implementation of Huawei SD-WAN Solution.
  - Describe the functions and features of the customer-premises equipment (CPE).

Contents

Architecture of Huawei SD-WAN Solution

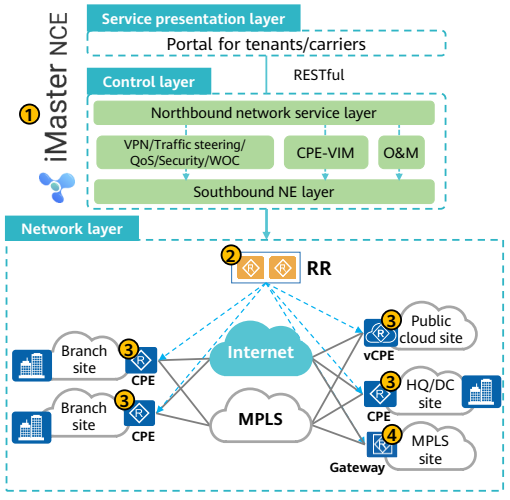


- **Service presentation layer**: presents the SD-WAN service logic.
  - Provides portal pages for carriers, MSPs, and large enterprises, on which SD-WAN services can be processed and enabled in an end-to-end manner.
  - Allows customers to customize portal pages as needed through northbound open APIs of the SD-WAN controller (iMaster NCE-WAN).
- **Control layer**: controls network layer devices. Typically, iMaster NCE-WAN is used to abstract the network model of an SD-WAN network, and pre-configure and automatically provision network services based on service templates.
- **Network layer**: is the basic physical network of an enterprise WAN, which consists of physical and virtual devices including CPEs, virtual CPEs (vCPEs).

HUAWEI

- From the perspective of functions, the overall architecture of the SD-WAN Solution consists of the network layer, control layer, and service presentation layer. These layers are associated with each other through standard interfaces and communication protocols.
- iMaster NCE-WAN is Huawei's SD-WAN controller, and consists of the service layer (built-in Portal) and control layer.
- The enterprise HQ, branches, DCs, and IT infrastructures deployed on the cloud are referred to as enterprise sites.

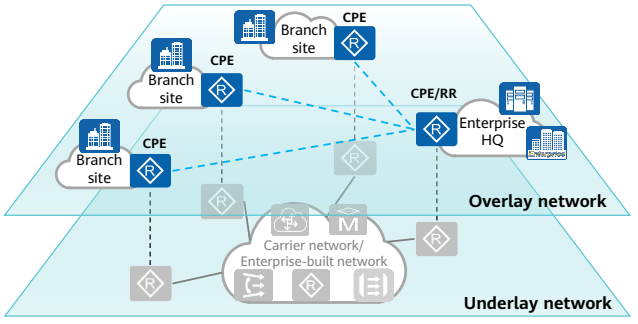
Key Components of the SD-WAN Solution



No.	Component	Functions
1	iMaster NCE-WAN	1. Network service orchestration 2. NE control 3. Basic network O&M 4. CPE orchestration and management 5. Basic performance monitoring (providing link quality information, application quality information, traffic information, as well as statistics from dimensions such as intra-site and inter-site)
2	RR	Distributes information about inter-CPE VPN routes and tunnels based on VPN topology policies.
3	CPE	Functions as the egress device of a site, which can be a traditional CPE or Network Functions Virtualization (NFV) vCPE.
4	Gateway	Connects an SD-WAN network to a non-SD-WAN network.

Network Layer Overview

- The SD-WAN network of an enterprise consists of 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 itself, including private lines and MPLS networks.
  - Virtual network: is also called an overlay network. Huawei SD-WAN Solution introduces IP overlay virtualization technology to construct one or more virtual overlay networks on a physical network. Service policies are deployed on virtual networks and decoupled from the physical network, so that service forwarding is independent of WAN interconnection.

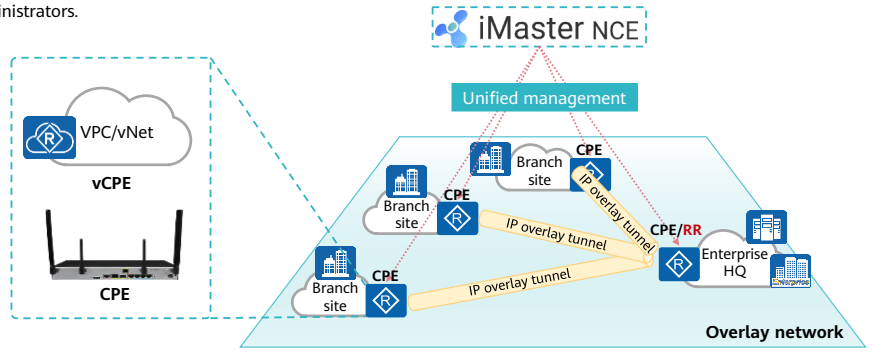


HUAWEI

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

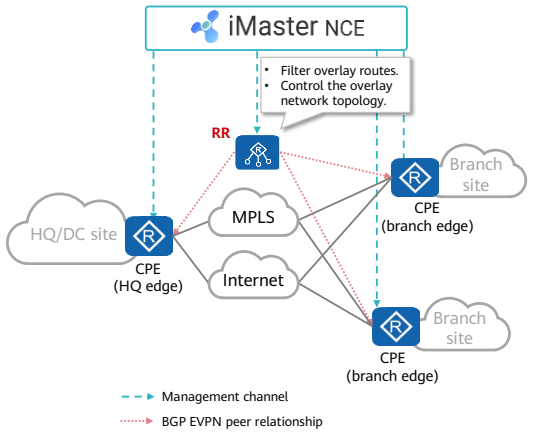
CPE Overview

- A CPE is an edge node of an SD-WAN network and also called an edge CPE. CPEs are interconnected with each other through IP overlay tunnels.
- CPEs are classified into traditional physical CPEs and vCPEs that are deployed at public cloud sites.
- All SD-WAN CPEs of an enterprise are centrally managed by iMaster NCE-WAN, and are managed and maintained by tenant administrators.



HUAWEI

RR Overview

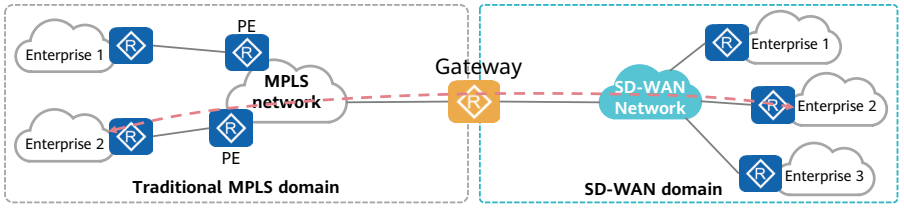


- A router reflector (RR) transfers BGP routes and reduces the number of BGP peers.
- In Huawei SD-WAN Solution, an RR also controls routes and the network topology. Therefore, an RR is also called an area controller on SD-WAN networks.
- Both CPEs at edge sites and RRs are managed by iMaster NCE-WAN.
- Control channels are established between RRs and 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 manner, sites can communicate with each other based on the designed overlay topology model.

- RR site: A CPE functions as an RR and distributes EVPN routes between CPE gateways at edge sites based on the VPN topology policy.
- If an egress CPE at a site is configured as both the gateway and RR, this site is an RR site. If no device takes the role of the gateway or RR at a site, the site is an edge site.
- One edge site can establish IBGP peer relationships with two RRs simultaneously, and the two RRs 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 often need to communicate with the enterprise's legacy sites or third-party services. Because some legacy sites are interconnected through MPLS VPN, while SD-WAN sites are interconnected through IP overlay tunnels, SD-WAN sites cannot directly communicate with legacy sites.
- An **SD-WAN gateway** can connect to both SD-WAN and traditional MPLS networks, achieving interconnection between the SD-WAN and traditional MPLS networks.

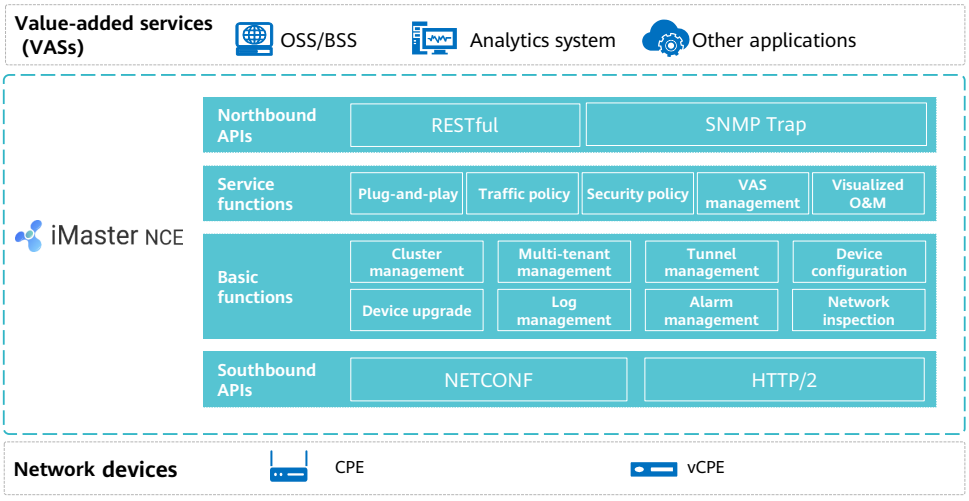


HUAWEI

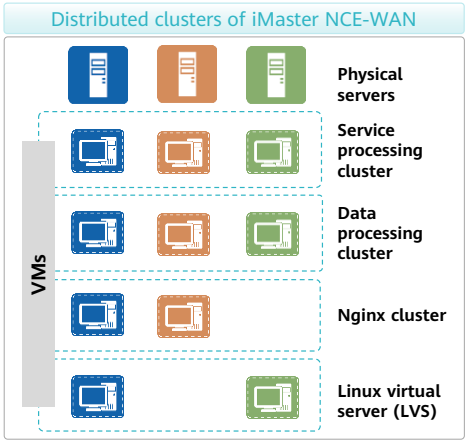
- A gateway's role name varies depending on the service scenario. For example, a gateway connecting SD-WAN sites to legacy sites is an interworking gateway (IWG), as shown in the above figure. A gateway connecting SD-WAN sites to a cloud is called a cloud gateway. In addition, functions of a gateway can be extended. A gateway that connects Point of Presence (POP) sites for building a POP network is referred to as a POP gateway.

Contents

Architecture of Huawei iMaster NCE-WAN



Distributed Cluster Deployment: Supporting Large Scale, High Reliability, and Flexible Capacity Expansion

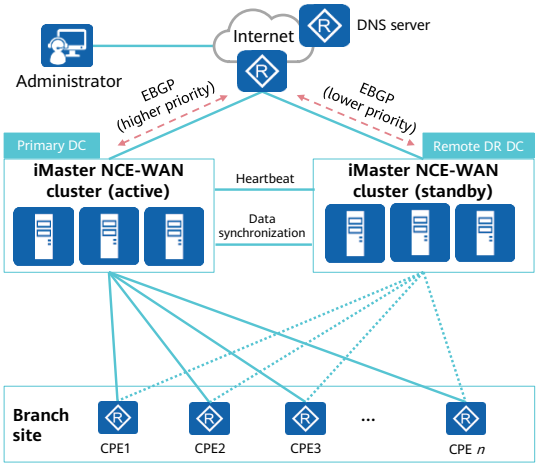


- Huawei iMaster NCE-WAN is deployed using a **distributed cluster architecture** to provide high reliability and load balancing capabilities. If one cluster node is faulty, other nodes take over services, ensuring service continuity.
- **Northbound load balancing**: External requests are distributed to all cluster nodes, instead of being processed on one node. This makes full use of cluster capabilities and improves reliability.
- **Southbound load balancing**: Cluster nodes are dynamically allocated to network devices based on the load of each cluster node.

Cluster	Functions
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	Functions as a high-performance HTTP proxy server that forwards concurrent connection requests and performs load balancing for northbound traffic at Layer 4 to Layer 7.
LVS	Functions as a load balancing component that performs load balancing for north-south traffic at Layer 1 to Layer 4.

Geographic Redundancy: Fast Switchover Ensures Normal Service Running

- Geographic redundancy supports disaster recovery and backup between two 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 northbound and southbound APIs of the controller use the same domain name or IP address. Tenants and devices access the active cluster through the same domain name or IP address. After an active/standby switchover is triggered, traffic is automatically switched to the new active cluster.
- Huawei SD-WAN Solution supports only one active controller cluster and one standby controller cluster.



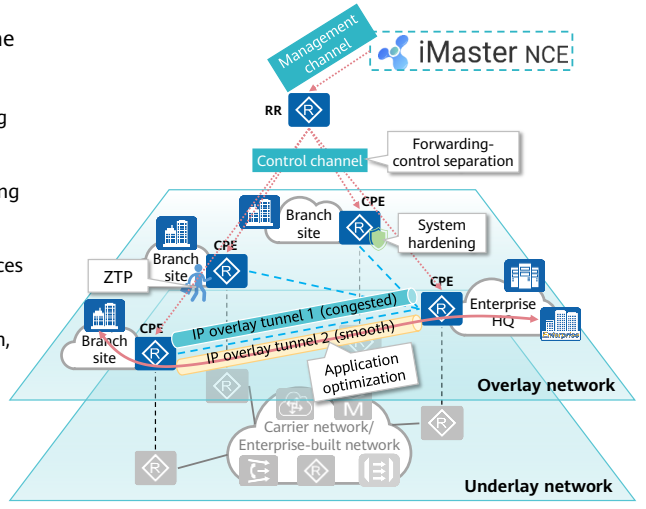
HUAWEI

- The same IP address or domain name must be configured for the southbound and northbound APIs of the active and standby controller clusters.
- Administrators access the controller through the domain name (based on the DNS record) or IP address, and the network uses BGP to control access traffic.

Contents

Main Functions of Huawei SD-WAN Solution

- Huawei SD-WAN Solution provides the following functions:
  - Zero touch provisioning (ZTP), enabling service provisioning within 1 hour
  - Forwarding-control separation, achieving flexible networking
  - Application optimization, making services controllable and visible
  - Comprehensive security defense system, ensuring service security

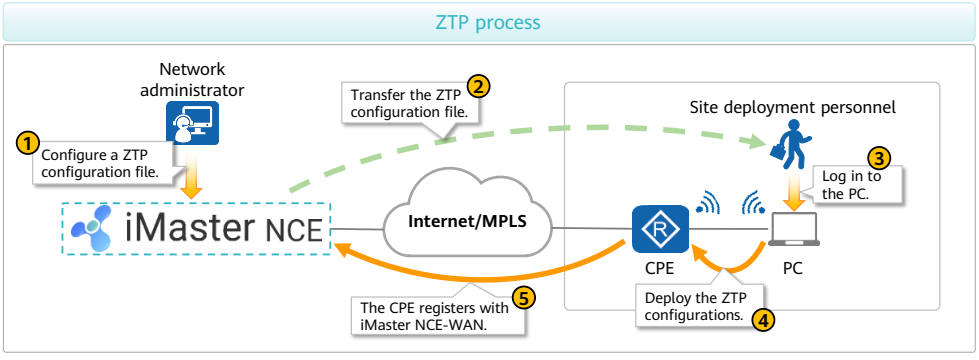


HUAWEI

- ZTP: Multiple ZTP modes are available to enable CPEs to quickly register with iMaster NCE-WAN.
- Forwarding-control separation, achieving flexible networking: CPEs establish management channels with iMaster NCE-WAN through NETCONF, and iMaster NCE-WAN delivers configurations to CPEs to establish IP overlay tunnels.
- Application optimization, making services controllable and visible: Service awareness (SA) technology is used to identify applications. TCP Flow Performance Monitor (FPM) and IP FPM technologies are used to implement application quality detection, and IP FPM technology is used to implement link quality measurement. Smart Policy Routing (SPR) technology implements intelligent link switching based on the application quality.
- Comprehensive security defense system, ensuring service security: Multiple VPN technologies, such as IPsec and MPLS, are used to provide E2E security protection. The firewall function provides comprehensive security assurance at the hardware, pipe, and application levels.

ZTP Overview

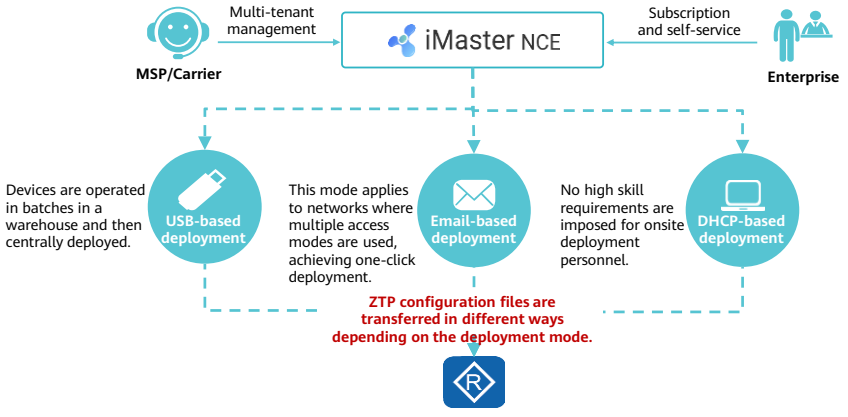
- With the development of network technologies such as SDN and cloud computing, an increasing number of enterprise networks adopt cloud-based management. However, most sites still need to be deployed by technical engineers onsite, leading to high deployment costs and long deployment period. Huawei offers ZTP to enable quick deployment.



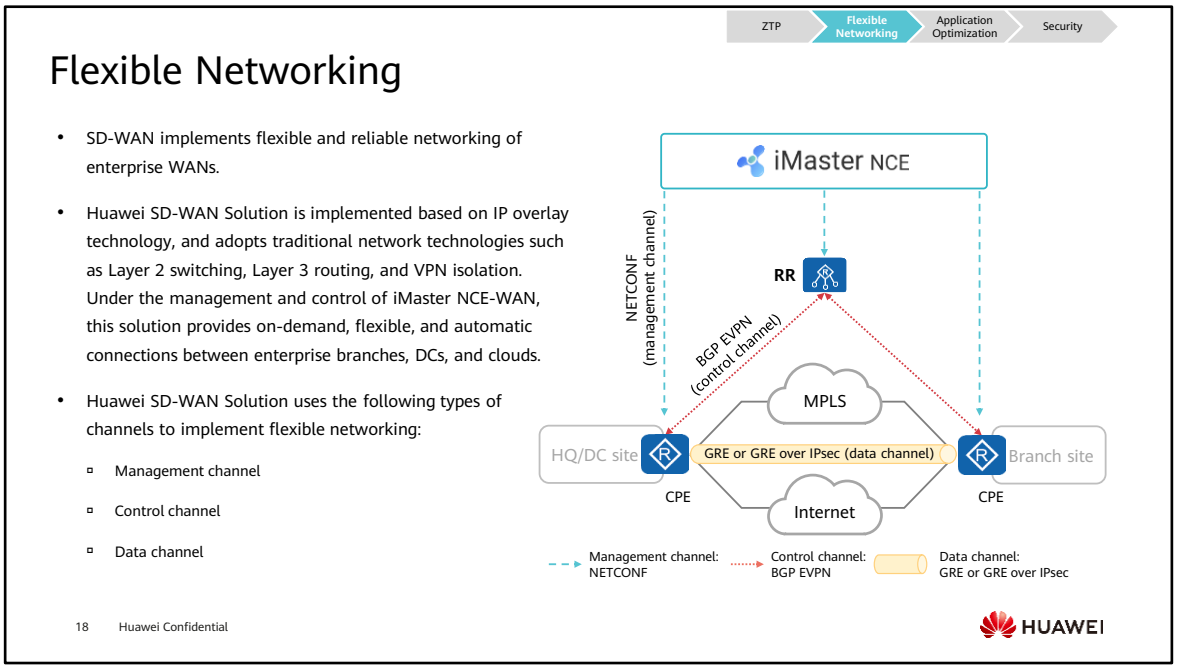
- ZTP: Zero Touch Provisioning

ZTP Modes

- Huawei SD-WAN Solution supports the following ZTP modes:



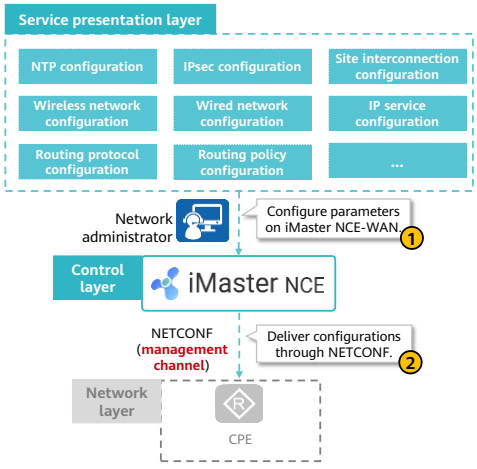
- Files for USB-based, email-based, and DHCP-based deployment can be generated through iMaster NCE-WAN.
- For details about each deployment mode, learn the course *Management and O&M*.



- Management channel:
  - iMaster NCE-WAN establishes management channels with all managed devices through NETCONF for network-wide device management and service orchestration.
- Control channel:
  - Control channels are established between CPEs and RRs.
  - RRs centrally control and distribute service routes between branch sites.
  - The enhanced BGP EVPN protocol is used to implement separate transmission of tenants' VPN routes and next-hop information, and implement IPsec SA negotiation.
- Data channel:
  - Data channels are established between CPEs.
  - Data is forwarded over GRE or GRE over IPsec data tunnels. The extended GRE header carries VN IDs to differentiate tenants or departments, thereby transmitting data of multiple VNs over the same data tunnel.

Management Channel

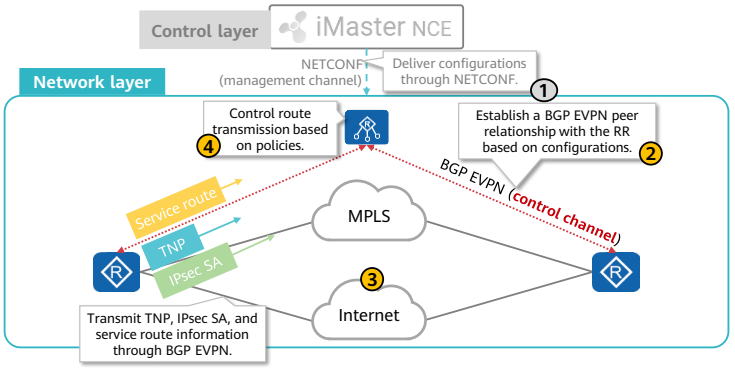
- Huawei iMaster NCE-WAN establishes management channels with CPEs through NETCONF.
- iMaster NCE-WAN delivers configurations through control channels to achieve the following functions:
  - Unified management of CPEs, automatic service delivery, and unified control of overlay networks
  - Application visualization and automatic application optimization
  - Network security services



HUAWEI

Control Channel

- After iMaster NCE-WAN delivers configurations to a CPE through the management channel, the CPE establishes a control channel with an RR through BGP EVPN.
- The control channel is used to transmit transport network port (TNP) information, IPsec SA information, and service routes.
- After the control channel is established, iMaster NCE-WAN controls route transmission and overlay topology establishment by deploying policies on the RR.

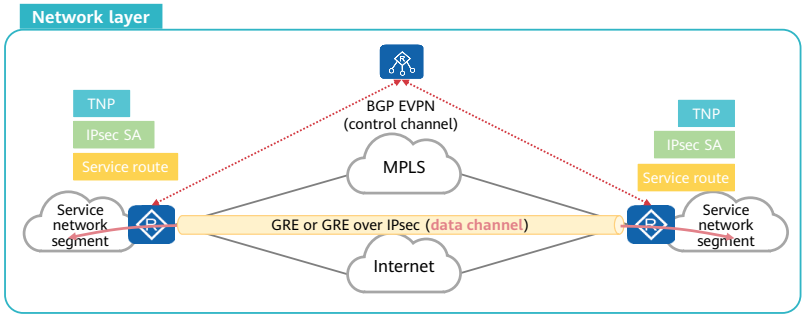


HUAWEI

- A TNP is a WAN port on a CPE used for connecting to a transport network. The key TNP information includes the site ID, CPE router ID, transport network ID, public IP address, private IP address, and tunnel encapsulation mode.

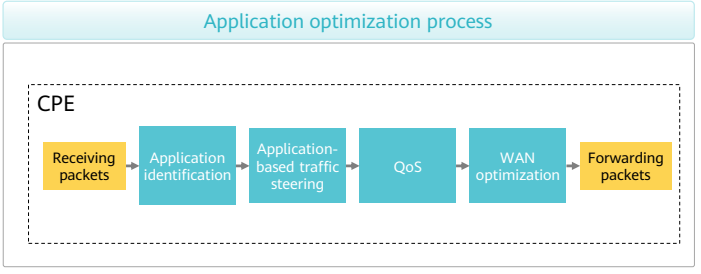
Data Channel

- Huawei SD-WAN Solution uses GRE or GRE over IPsec to establish data channels.
- CPEs establish GRE or GRE over IPsec tunnels based on the TNP and IPsec SA information transferred through BGP EVPN.
- CPEs forward data based on the service routes transferred through BGP EVPN.



Application Optimization

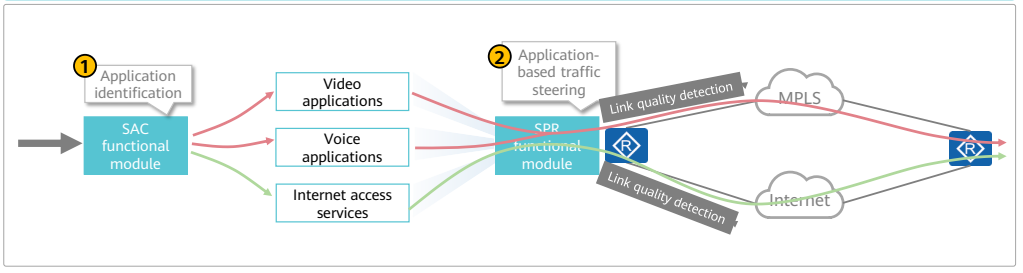
- To meet diversified requirements of enterprise applications, traditional WANs have the following issues to resolve:
  - Applications of different values are carried on the same link.
  - When link quality deteriorates, dynamic routing cannot be implemented.
  - No effective measure is available when link quality deteriorates.
- To resolve these issues, Huawei SD-WAN Solution offers enterprise experience optimization functions, including:
  - Application identification
  - Application-based traffic steering
  - QoS
  - WAN optimization



Application Identification and Traffic Steering

- Huawei SD-WAN Solution uses Smart Application Control (SAC) to identify applications and uses SPR to implement application-based traffic steering.
  - SAC enables a device to identify applications and groups application traffic through SA and first-packet identification (FPI).
  - SPR enables a device to measure the link quality based on link quality detection packets and determine forwarding paths for traffic.

Application identification and traffic steering processes

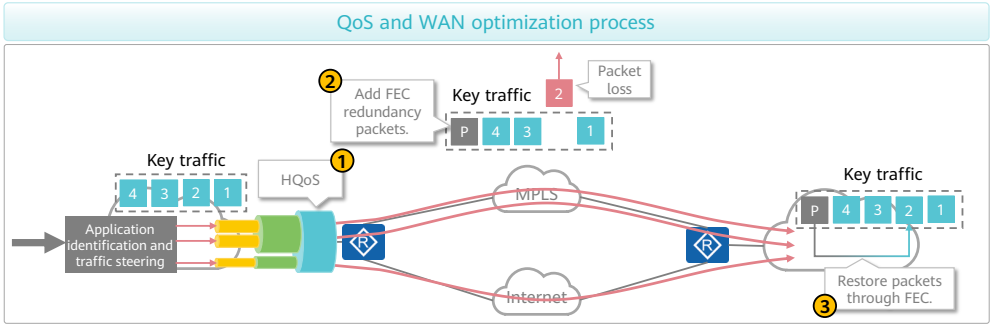


HUAWEI

- For details about SAC and SPR, learn the course *HA Technologies*.

QoS and WAN Optimization

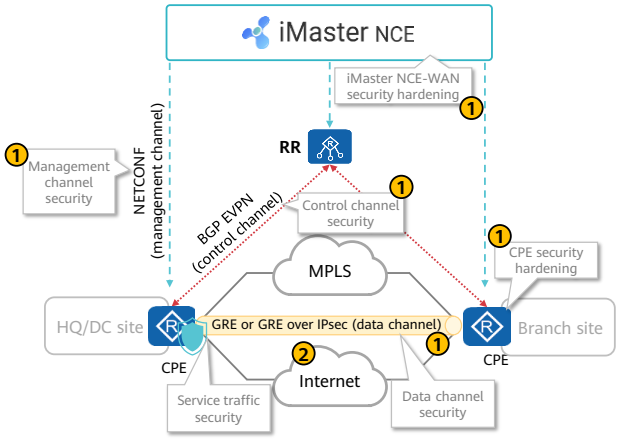
- Huawei SD-WAN Solution uses HQoS for bandwidth control and scheduling, and uses Forward Error Correction (FEC) or Adaptive FEC (A-FEC) for WAN traffic optimization.
  - HQoS implements hierarchical scheduling based on multi-level queues and differentiates services and users, implementing refined QoS.
  - FEC or A-FEC optimization enables the local device to adjust related parameters based on packet loss on the network to generate redundant packets. The peer device then verifies and reassembles the packets.



HUAWEI

Security Overview

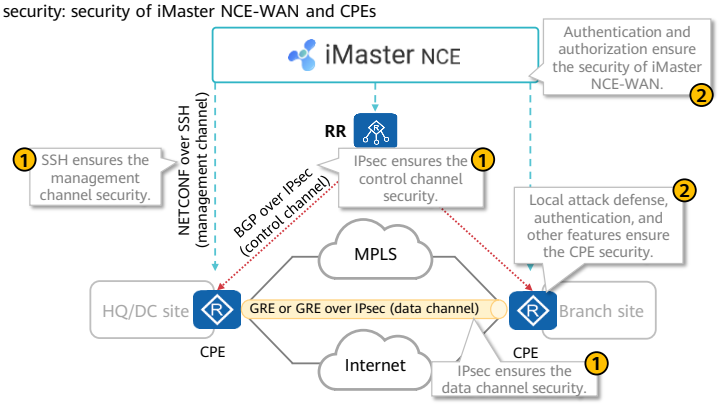
- With the emergence of SD-WAN, the traditional closed architecture of enterprise WANs is transformed into an open architecture, which enlarges the attack range and brings new security challenges, including unauthorized access, data leakage, and network attacks.
- Huawei SD-WAN Solution provides high security from two perspectives:
  - System security: component security and inter-component security
  - Service security: firewall, IPS, and URL filtering



HUAWEI

SD-WAN System Security Hardening

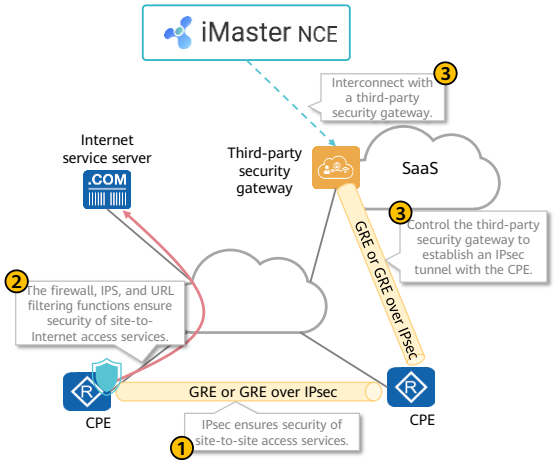
- SD-WAN system security includes:
  - Inter-component security: security of management, control, and forwarding channels
  - Component security: security of iMaster NCE-WAN and CPEs



HUAWEI

SD-WAN Service Security

- From the perspective of traffic, SD-WAN services are classified into the following types:
  - Site-to-site access service
    - IPsec ensures security of site-to-site access services.
  - Site-to-Internet access service
    - The built-in firewall, IPS, and URL filtering functions of CPEs ensure the security of site-to-Internet access services.
  - Site-to-cloud access service
    - iMaster NCE-WAN is interconnected with a third-party security gateway and controls this gateway to provide security services.



HUAWEI

Contents

Highlights of SD-WAN CPEs

Huawei NetEngine AR series routers function as CPEs and provide extensive SD-WAN features, including hybrid link access, as well as service controllability and visibility. The CPEs reduce WAN interconnection costs and improve O&M efficiency.

5G uplink

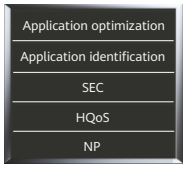
Enterprise-class 5G egress routers

The full lineup of NetEngine AR6000 routers supports 5G.

- **Large bandwidth:** 100 Mbit/s UL, 700 Mbit/s DL
- **Full frequency**: 5G, 4G, 3G, and 2G
- **Dual architectures**: standalone (SA) and non-standalone (NSA)

Ultra-high performance

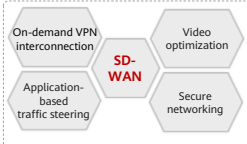
CPU + NP, high forwarding performance



5 built-in acceleration engines

Optimal experience

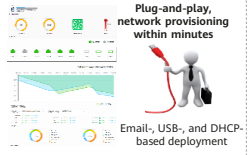
Application-based intelligent traffic steering



- On-demand **VPN** interconnection, built-in **security** functions
- Video optimization: A-FEC, ensuring no frame freezing even at **20%** packet loss rate

Simplified O&M

ZTP & visualized O&M



- Diversified O&M platforms, real-time intelligent O&M
- ZTP & visualized O&M, network provisioning in minutes

HUAWEI

Portfolio of Huawei NetEngine AR Routers

HQ/Large branches

NetEngine AR6300/AR6200 series

NetEngine AR6300

SRU-400H/SRU-600H

NetEngine AR6280

SRU-400H/SRU-600H

Small/Midsize enterprise branches

NetEngine AR6100 series

AR6120

AR6121

AR6140-9G-2AC

AR6140-16G4XG

Small enterprises

NetEngine AR650 series

Ethernet + LTE (MIC)

AR651

Ethernet + Wi-Fi + LTE (MIC)

AR651W

SOHO

NetEngine AR610 series

Ethernet + Wi-Fi

AR611W

Ethernet + Wi-Fi + LTE

AR611W-LTE4CN

Enterprise-Class 5G Branch Routers: Large bandwidth and Low Latency

The full lineup of NetEngine AR6000 series routers supports 5G uplinks.

Enterprise-class 5G branch router	Large bandwidth	Full frequency	Dual architectures
	UL 100 Mbit/s DL 700 Mbit/s * Test data	5G 3G 4G 2G	NSA (4G core; 4G, 5G; 5G AR) » SA (5G core; 5G; 5G AR)

NetEngine AR Routers: Meeting Burst Traffic Requirements of Enterprise Branches

- Huawei NetEngine AR series routers are the next-generation routing and gateway devices that provide SD-WAN, routing, switching, VPN, security, voice, and MPLS functions.

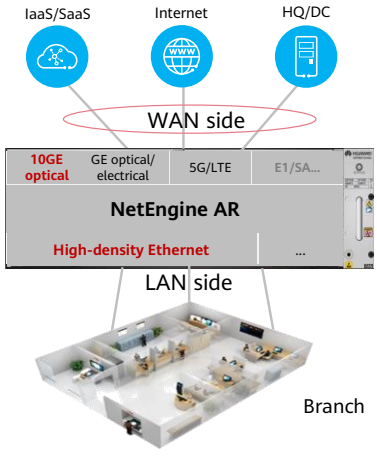
NetEngine AR

CPU + NP heterogeneous forwarding
Excellent forwarding performance

5 built-in hardware acceleration engines
Application optimization, application identification, HQoS, SEC, and NP

Ultra-fast algorithm, enabling fast forwarding
Acceleration instruction set, quick ACL and route matching

10GE Interconnection and High-Density Access: Building Wide Pipes for Branch Interconnection



- **10GE uplink, multi-link redundancy, wide pipes for WAN-side interconnection**
  - High-density 10GE ports: 14 x 10GE optical ports on the SRU-400H and SRU-600H cards; 10GE optical ports on NetEngine AR6000 series routers (except the AR6140-9G-2AC)
  - Multi-link redundancy: 10GE, GE optical/electrical, and LTE links
  - Flexible switching between LAN and WAN interfaces using commands
- **High-density access, higher LAN-side access capability**
  - SRU-400H and SRU-600H: All WAN interfaces can be switched to LAN interfaces.
- **Flexible card expansion**
  - Expansion of WSIC, XSIC, SIC, and MIC cards

Function Convergence: Simplifying Branch Interconnection

On-demand VPN interconnection

Various VPN types

Layer 2 and Layer 3 VPNs

Experience assurance for key applications

Application optimization

A-FEC, application-based intelligent traffic steering (supported only in the SD-WAN Solution)

Wi-Fi interconnection of branches

Wi-Fi integration

WAC or Fat AP

Secure interconnection of branches

Built-in security functions

Firewall, IPS, URL filtering, etc.

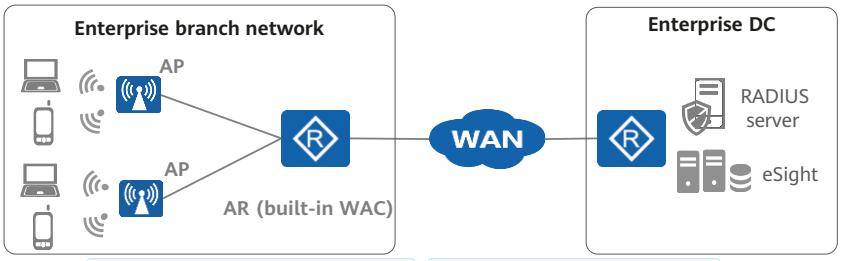
Next-generation NetEngine AR routers

Wired and wireless convergence

High-density ports, 10GE interconnection, 5G/LTE uplinks

HUAWEI

Built-in WAC: Reducing Wi-Fi Deployment Costs of Enterprises



Application scenarios	Secure and flexible access
• Small and midsize enterprises that require wired and wireless convergence • Scenarios where APs use the local forwarding mode and users are centrally authenticated by AR routers • Scenarios where both APs and WACs support Layer 2 and Layer 3 networking	• Portal authentication • 802.1X authentication • MAC address authentication • Intra-VLAN or inter-VLAN roaming on the same WAC

Wi-Fi Integration for Wired and Wireless Convergence

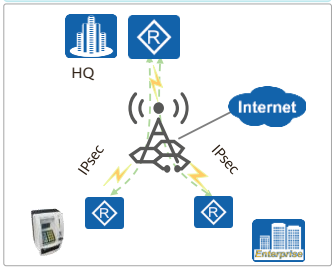
Wi-Fi integration for wired and wireless convergence

- Wi-Fi functions integrated on fixed AR routers, such as the AR651W, AR611W, and AR611W-LTE4CN
- 802.11ac/b/g/n, a maximum rate of 1167 Mbit/s
- Dual bands (2.4 GHz and 5 GHz), providing stronger wireless user access capabilities
- 2x2 MIMO

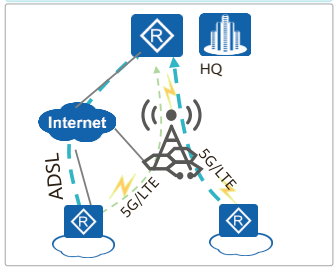
- Security: Traditional encryption and authentication ensure terminal access security.
- Network management: AR routers function as independent APs to provide functions such as user access, authentication, data security, service forwarding, and QoS.
- Networking scale: small-scale networking, requiring low costs

Wireless Uplinks: Leveraging 5G/LTE to Build Wireless Branches

5G/LTE VPN interconnection



5G/LTE link backup



5G/LTE-capable cards and devices


LTE SIC


AR611W-LTE4CN


LTE MIC


5G SIC

2 major scenarios

- **5G/LTE VPN interconnection:** wireless access scenarios such as bank ATM interconnection and mobile office
- **5G/LTE link backup**: enterprise branch interconnection, with wireless links as backup links to ensure service reliability

3 types, 4G full frequency

- **3 types of LTE modems:** LTE SIC (AR6000 series), LTE MIC (AR650 series, except the AR651C), and embedded LTE modem (AR610 series)
- **4G full frequency**, meeting network standards of different carriers

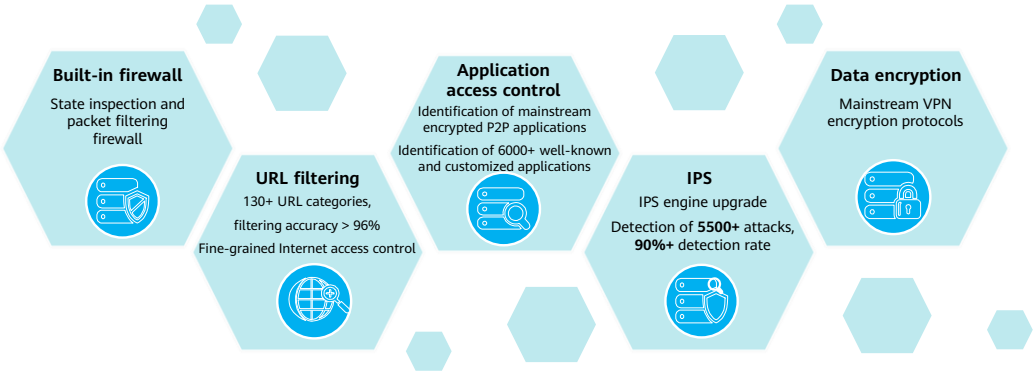
5G solution

- 5G single-chip multi-mode modem
- Sub-6G at 200 MHz
  - * DL rate: 700 Mbit/s
  - * UL rate: 100 Mbit/s

* Test data

HUAWEI

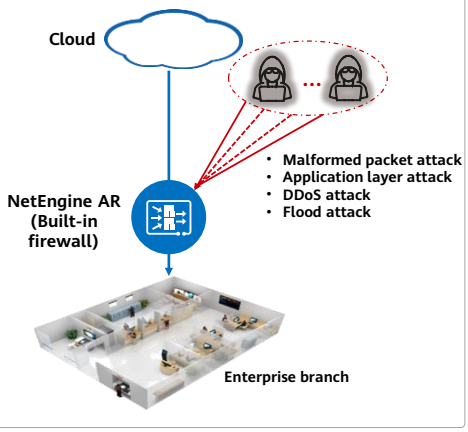
Comprehensive Security Protection



HUAWEI

Built-in Firewall: Effectively Ensuring Enterprise Network Security

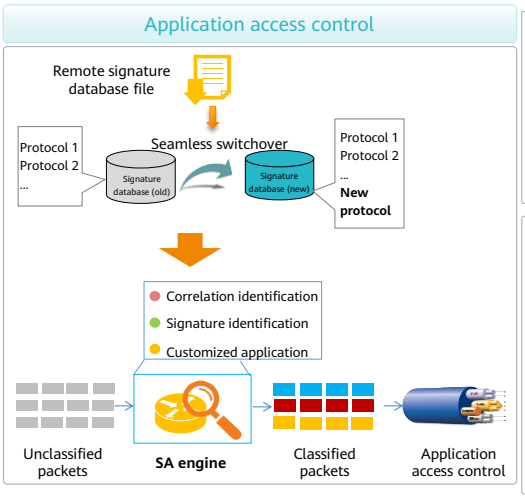
Built-in firewall, ensuring branch network security



- **Extensive firewall functions**
  - Packet filtering firewall, quickly matching ACLs and IP addresses to filter network-wide packets
  - Application specific packet filter (ASPF), detecting TCP/UDP sessions and filtering out invalid data packets
- **Comprehensive attack defense mechanism**

  Defense against the following attacks:
  - ICMP redirection attacks and ICMP unreachable attacks
  - Address scanning and port scanning attacks
  - SYN, ICMP, and UDP flood attacks
  - Land, Smurf, Fraggle, and WinNuke attacks
- **Flexible firewall policies**
  - Firewall logging, and traffic statistics and monitoring
  - Independent security policy deployment for multiple service VPNs through virtual firewalls
  - Security policy configuration based on security zones
  - Dynamic blacklist for proactive attack defense

Application Access Control: Extensive Applications, Flexible Upgrade



- **Identification of 6000+ applications**
  - Various identification modes: packet signature identification, correlation identification, behavior identification, etc.
  - Identification of mainstream protocols (such as P2P, VoIP, IM, game, and email) and 6000+ applications
  - Identification of applications customized based on URLs and 5-tuple information

- **Flexible upgrade of the SA signature database**
  - SA signature database files are released and maintained by Huawei Security Competence Center. Customized applications can also be imported.
  - Batch or scheduled upgrade of signatures in the SA signature database; periodic release of the new signature database
  - Query of the SA signature database upgrade status, including the upgrade time, countdown, progress, and result (success or failure)
  - Rollback of the SA signature database upon an upgrade failure

Comprehensive Access Control and Intrusion Prevention

URL filtering

- **Remote URL category database containing a large number of URLs**

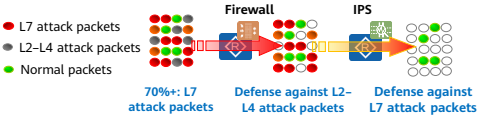
  Remote query of 100+ million URLs by category, 130+ predefined categories, category customization, and timely update and efficient query based on Huawei reputation system
- **Various URL matching modes**

  Prefix matching, suffix matching, keyword matching, exact matching, and quick matching
- **Fine-grained blacklist and whitelist**

  The blacklist and whitelist are effective supplements to help precisely define and control access to specific websites.
- **Flexible response modes**

  Various URL filtering actions can be configured to push different URL response pages.

IPS



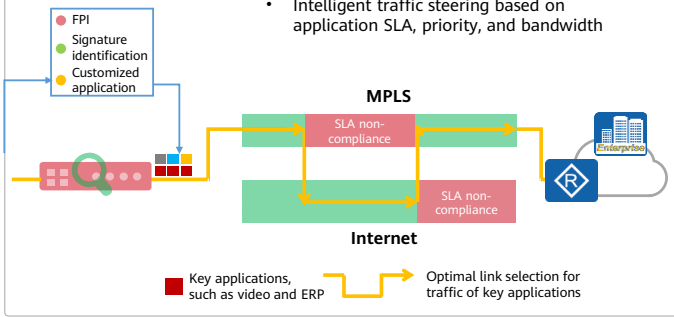
- **IPS database with a large number of signatures and a high detection rate**
  - Attack detection based on 1600+ signatures, with 90%+ detection rate
  - Extensive signatures based on network behaviors, such as Trojan horses, worms, botnets, spyware, vulnerability attacks, and web attacks
- **Flexible upgrade**
  - Online upgrade of the signature database and real-time upgrade of the IPS engine, defending against the latest intrusions
- **Converged deployment**
  - The IPS is embedded in AR routers, requiring no dedicated fault detection points while reducing operation costs.

Application-based Intelligent Traffic Steering: Ensuring Experience of Key Applications

Application-based intelligent traffic steering

Application controllability and visibility

Quick identification of key applications



- **Application-based intelligent traffic steering**, automatically switching traffic of key applications to the optimal link
- Intelligent traffic steering based on application SLA, priority, and bandwidth

Customer benefits

Automatically switching traffic of key applications to the optimal link

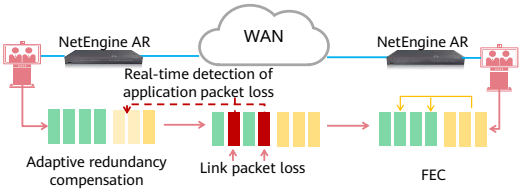
Leveraging hybrid links, such as MPLS, Internet, and LTE

A-FEC: Ensuring Smooth Video Experience Even at 20% Packet Loss Rate

A-FEC

Real-time detection of application packet loss and adaptive redundancy compensation

Optimizing experience of video conferencing, live streaming, video surveillance, and VoIP services



Traditional: frame freezing and artifacts at 2%+ packet loss rate

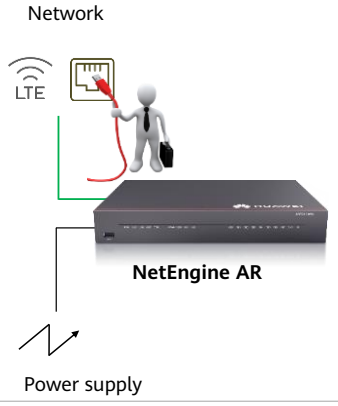
A-FEC: no frame freezing even at 20% packet loss rate

Note: A-FEC is supported only in the SD-WAN Solution.

HUAWEI

ZTP: Achieving Plug-and-Play of Devices in Multiple Scenarios

Plug-and-play and minute-level device deployment



ZTP for branch network deployment

DHCP | **USB flash drive** | **Email**

Different interfaces: Ethernet, LTE, xDSL...

Different access modes: static IP address, PPPoE, DHCP...

Different scenarios: dual-CPE, batch deployment, device replacement...

HUAWEI

Quiz

1. (Multiple choices) Which of the following layers are included in the architecture of Huawei SD-WAN Solution?
   A. Service presentation layer
   B. Control layer
   C. Network layer
   D. Device layer

- 1. ABC

Summary

- Architecture and Components of Huawei SD-WAN Solution
  - Huawei SD-WAN Solution consists of iMaster NCE-WAN, CPEs, gateways, and RRs.
- Huawei iMaster NCE-WAN
  - iMaster NCE-WAN has multiple functional modules, and supports geographic redundancy and distributed deployment.
- Implementation of Huawei SD-WAN Solution
  - Huawei SD-WAN Solution features ZTP, flexible networking, application optimization, and security hardening.
- Huawei SD-WAN CPEs
  - Huawei SD-WAN CPEs are mainly NetEngine AR series routers.

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.

SD-WAN Device Deployment

HUAWEI

Foreword

- When building a wide area network (WAN), an enterprise needs to deploy devices at each branch site. However, traditional network deployment modes take an excessively long time for large-scale enterprises.
- To accelerate deployment, many vendors have provided functions for deploying devices with zero configurations. However, it is difficult to make deployment files when these functions are used.
- The zero touch provisioning (ZTP) function of Huawei SD-WAN Solution greatly shortens the deployment time and simplifies the preparation of deployment files.
- This course describes the basic concepts and functions you need to understand during SD-WAN deployment.

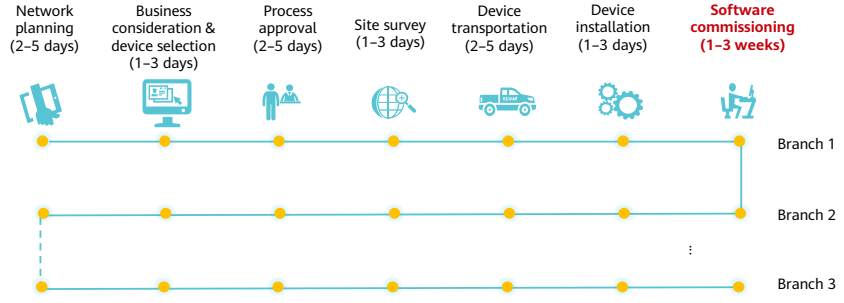
Objectives

- On completion of this course, you will be able to:
  - Describe the multi-tenant management boundary of the SD-WAN Solution.
  - Describe SD-WAN deployment modes and application scenarios.

Contents

Challenges Facing Traditional Deployment Modes

- The provisioning period of traditional private line services is as long as over 30 days, while that of international private line services even reaches 90 days. Traditional branch network deployment also faces the following challenges:
  - High technical requirements: Professional IT engineers need to deploy devices on site.
  - Low efficiency: Devices are scattered and online operations are time-consuming.
  - Misoperation risks: Errors may occur due to manual operations during initial configuration.

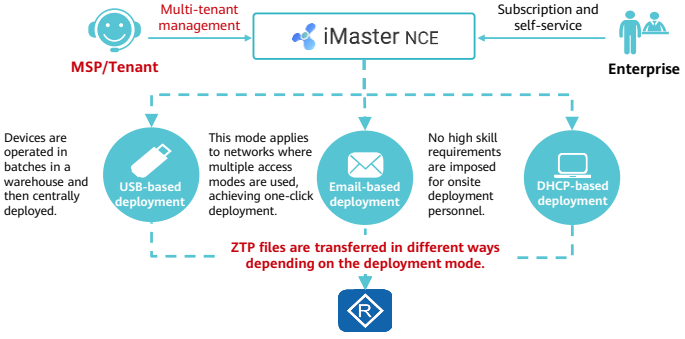


HUAWEI

- To address these challenges, the ZTP function of the SD-WAN Solution provides the following benefits:
  - Lower technical requirements of site deployment: Device deployment does not require skilled professional IT engineers.
  - E2E automatic deployment: The probability of parameter input errors during manual deployment is reduced.
  - Centralized planning of offline services for batch devices: Services are provisioned immediately after devices at sites are registered with iMaster NCE-WAN.

ZTP Modes

- ZTP greatly shortens the deployment time. Huawei SD-WAN Solution supports the following ZTP modes:
  - Email-based deployment: This mode is commonly and widely used, and has low requirements on site deployment personnel.
  - DHCP-based deployment: This mode has almost no requirements for site deployment personnel but requires support from carriers. Therefore, this mode is applicable to only a few scenarios.
  - USB-based deployment: This mode has low requirements on site deployment personnel. However, this mode involves many steps and is seldom used on live networks.



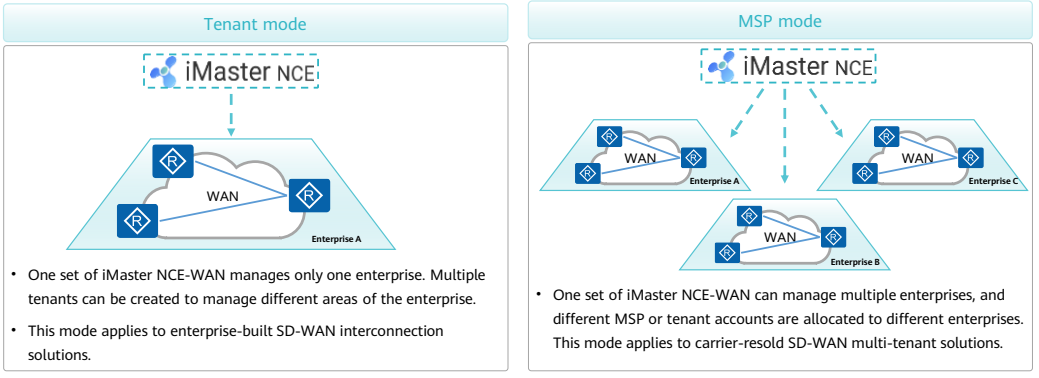
HUAWEI

- DHCP-based deployment and email-based deployment are described in this document.

Contents

Tenant Management Modes

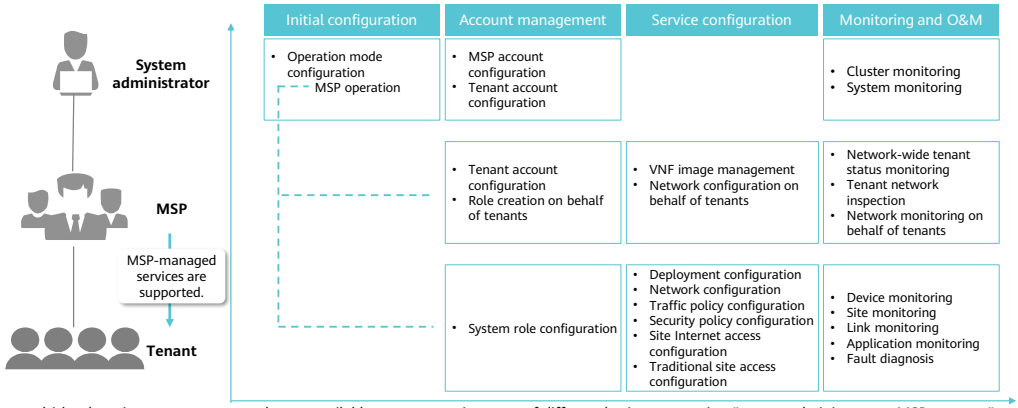
- Huawei SD-WAN Solution supports two tenant management modes based on service application scenarios.
  - MSP mode: The system administrator creates an MSP, and then the MSP creates tenants.
  - Tenant mode: The system administrator creates tenants.



HUAWEI

- Tenant mode: The system administrator creates tenants directly.
  - This mode is applicable when an enterprise wants to deploy and manage its own internal network. The system administrator can create multiple tenants to isolate and manage networks of different departments or subsidiaries. Each tenant can create an administrator to manage the tenant's network.
- MSP mode: The system administrator creates an MSP, and then the MSP creates tenants.
  - This mode is applicable when an enterprise provides network management services for external users. An MSP can be a product distributor, whereas a tenant is a customer who requires network management services.
  - The system administrator creates an MSP and specifies an MSP administrator. The specified MSP administrator can create other MSP administrators and specify their management permissions. MSP administrators create tenants and specify tenant administrators. The specified tenant administrators can create other tenant administrators and specify their management permissions.
  - A tenant administrator can authorize an MSP to manage the tenant network. In this way, the MSP administrator can maintain the tenant network.

Multi-Level Service Management



- Two multi-level service management modes are available to meet requirements of different business scenarios: "system administrator + MSP + tenant" and "system administrator + tenant".
- All tenants share one set of iMaster NCE-WAN, which differentiates tenant data based on tenant identifiers. Data is isolated between tenants, and each tenant possesses independent resources and management permissions.

HUAWEI

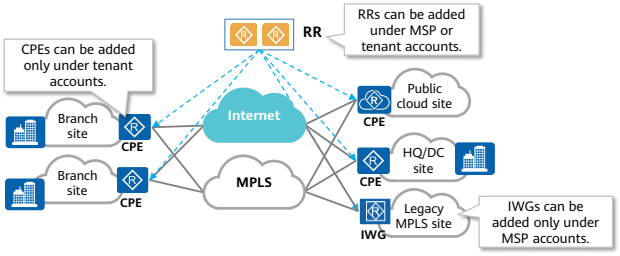
- Tenants can entrust network O&M services to MSPs so that the MSPs can maintain the tenant networks.

MSP/Tenant Roles

- MSP and tenant roles are defined to specify the access permissions of the MSPs and tenants. The SD-WAN Solution provides the following pre-defined roles:
  - Monitor: has the permission to query tenant services and related configurations.
  - Open API operator: has the permission to operate open API services and relevant configurations.
  - Tenant administrator: has the permission to operate tenant services and relevant configurations.
  - Operator: has the permission to manage system service operations.
  - CLI operator: has the permission to import device commands by using command line templates.
- New roles can be planned based on functions:
  - Management role: global administrator, who has all permissions
  - Monitoring role: global monitoring personnel who have all monitoring permissions
  - Configuration role: network configuration personnel who have the permission to configure the network, traffic policies, and security policies
  - O&M role: O&M personnel who have the permission to manage devices, files, and logs

MSP and Tenant Devices

- Devices can be added and deployed under both MSP and tenant accounts. Devices added under an MSP account can only be managed by the MSP, and devices added under a tenant account can only be managed by the tenant.
- MSP and tenant accounts are used in different service scenarios. Therefore, the roles of MSP and tenant devices are different.
  - Devices added under an MSP account can be used as IWGs or RRs but cannot be used as CPEs.
  - Devices added under a tenant account can be used as CPEs or RRs but cannot be used as IWGs.



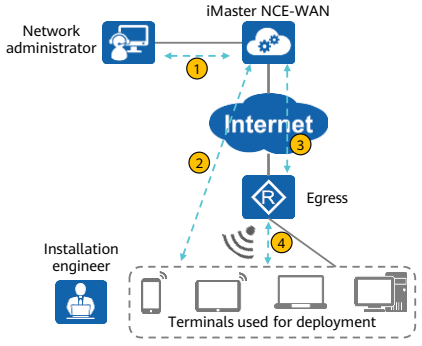
HUAWEI

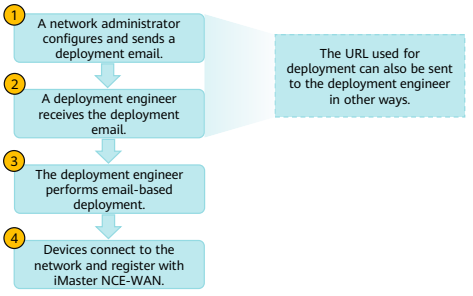
- Both MSPs and tenants can deploy devices in ZTP mode.
- If a tenant authorizes an MSP to manage its network, the MSP can also manage devices of the tenant.
- RRs added under an MSP account are shared RRs, and can be shared by multiple tenants. Shared RRs are used to reduce investment costs.
- RRs added under a tenant account are exclusive RRs, and cannot be shared by tenants. Exclusive RRs are used to improve stability.

Contents

Email-based Deployment Overview

- During email-based deployment, a network administrator configures deployment information on the iMaster NCE-WAN web UI to generate uniform resource locator (URL) parameters in a deployment email, and then sends the deployment email to a specified deployment mailbox. A deployment engineer receives the deployment email and clicks the URL in the email to start the deployment process. Subsequently, devices are automatically deployed.



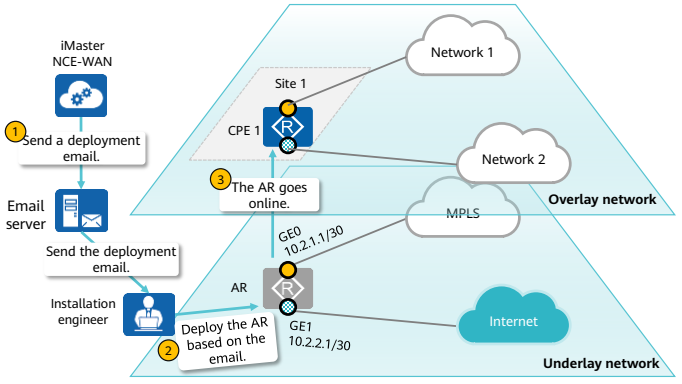


HUAWEI

- Email-based deployment can be used only in SD-WAN and CloudVPN Solutions.
- Email-based deployment applies only to devices with factory settings.
- Before email-based deployment of a device, users cannot log in to the web system of the device and change the password. Otherwise, the device will fail to be deployed.
- When using the Internet Explorer for email-based deployment, you need to select **Use HTTP1.1** on the **Internet Options** > **Advanced** tab page of the browser. Otherwise, the device will fail to be deployed.

Email-based Deployment Details

- Before performing email-based deployment, ensure that the following configurations have been completed on iMaster NCE-WAN:
  - The email server has been configured.
  - Devices have been added.
  - Global parameters have been configured.
  - Sites have been created.
  - The access mode of sites has been configured.



HUAWEI

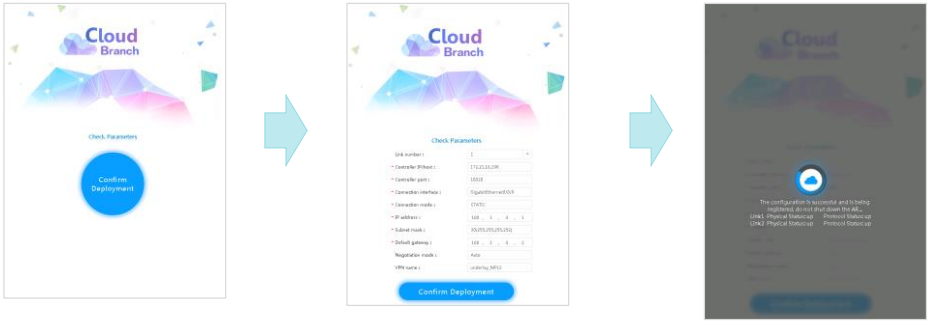
- The email server must be configured by a system administrator or MSP administrator.
  - If both the system administrator and MSP administrator have configured an email server, the email server configured by the MSP administrator is used preferentially. If the email server configured by the MSP administrator is unavailable, the email server configured by the system administrator is used.
- Devices need to be added to iMaster NCE-WAN in advance.
  - Devices can be added under an MSP account or a tenant account.
  - Devices added under an MSP account can be used as IWGs or RRs but cannot be used as CPEs.
  - Devices added under a tenant account can be used as CPEs or RRs but cannot be used as IWGs.
- The global parameter configurations include:
  - Physical network configurations: transport network, IPsec encryption, device activation security, link connectivity detection, traffic steering policy parameters, and password of the **admin** account
  - Virtual network configurations: routes, address pools, and DNS
  - Collection configurations: network traffic, application quality, and WAN link traffic
- To facilitate device management and improve service provisioning efficiency, devices on the same network of a tenant can be added to the same site.

- The access mode of a site defines information about physical links at the site, such as the interface type, interface IP address, and VPN to which the interface belongs.
  - Configuring WAN-side physical links for sites is the prerequisite for site deployment. After a site is configured or activated, WAN-side links can be added or deleted.

Performing Email-based Deployment

- Perform deployment based on the URL in the deployment email. The URL format is as follows:

  https://*ip*/portal?ac_host=*ac_host_value*&ac_port=*ac_port_value*...&url_pass=*url_pass_value*.

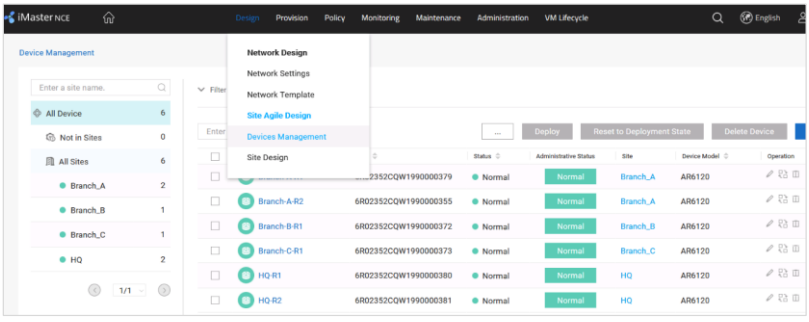


HUAWEI

- The URL in a deployment email is in the format of https://ip/portal?ac_host=ac_host_value&ac_port=ac_port_value...&url_pass=url_pass_value.
  - Multiple parameters can be configured in the URL and separated by ampersands (&).
  - **ip** indicates the IP address of the web system of the device to be deployed. The default web system IP address and subnet mask of a device are 192.168.1.1 and 255.255.255.0, respectively.

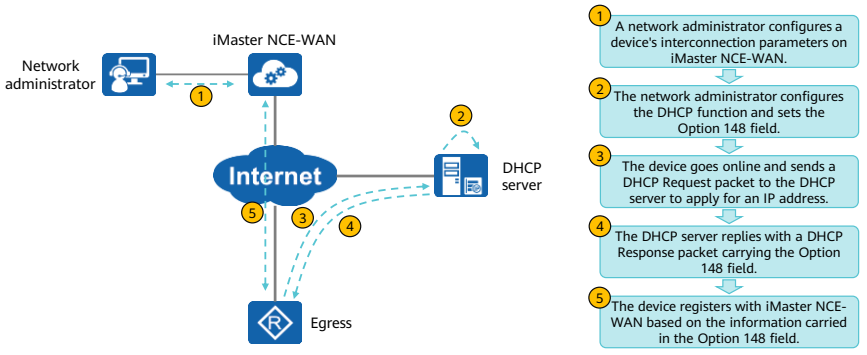
Checking the Email-based Deployment Result

- On the iMaster NCE-WAN web UI, choose **Design** > **Site Agile Design** > **Devices Management** to check whether devices are successfully deployed.
  - If the status and administrative status of the devices are normal, the devices are successfully deployed.



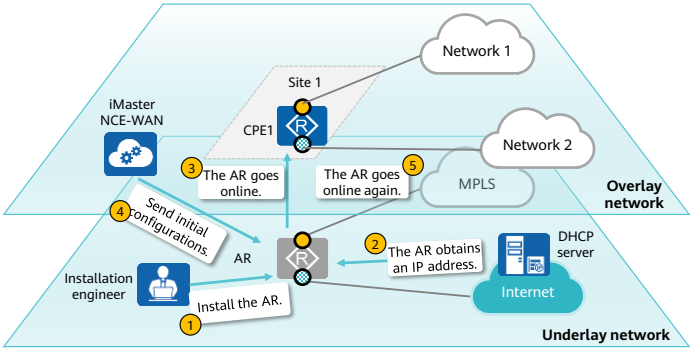
DHCP-based Deployment Overview

- DHCP-based deployment is implemented based on the option fields in DHCP packets. The DHCP server is configured with Option 148, which contains address information about iMaster NCE-WAN. Devices obtain such information through DHCP interaction.
- The DHCP Option 148 field is usually configured as follows:
  - option 148 ascii agilemode=agile-cloud;agilemanage-mode=ip;agilemanage-domain=172.21.16.198;agilemanage-port=10020;



DHCP-based Deployment Details

- After the DHCP server is configured, you need to add devices to iMaster NCE-WAN so that the devices can go online.
- Sites can be created and the global parameters and site access mode can be configured after devices go online.
- After the site access mode is configured on iMaster NCE-WAN, iMaster NCE-WAN sends initial configurations to devices and the devices go online again.

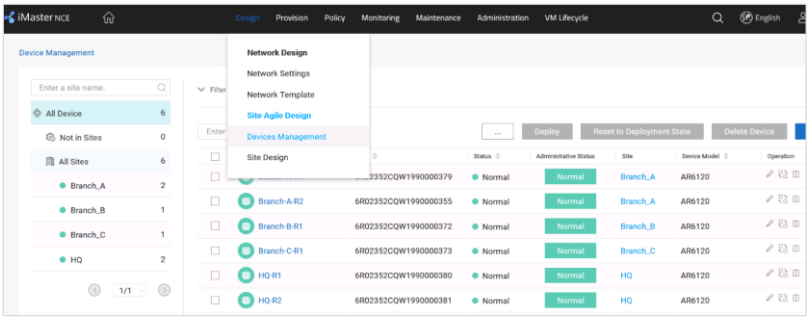


HUAWEI

- DHCP-based deployment is imperceptible to device deployment personnel, and devices are plug-and-play.

Checking the DHCP-based Deployment Result

- On the iMaster NCE-WAN web UI, choose **Design** > **Site Agile Design** > **Devices Management** to check whether devices are successfully deployed.
  - If the status and administrative status of the devices are normal, the devices are successfully deployed.



Quiz

1. (Multiple-choice question) Which of the following roles can be set for the devices added under an MSP account?
   A. CPE
   B. RR
   C. IWG
   D. Firewall
2. (True or False) In DHCP-based deployment mode, devices can register with and go online on iMaster NCE-WAN only after the site configuration is complete.

- 1. BC
- 2. False

Summary

- Huawei SD-WAN Solution supports the following user accounts:
  - System administrator account, which is used to create MSPs (in MSP mode) or tenants (in tenant mode)
  - MSP account, which is used to create tenants and add RRs and IWGs
  - Tenant account, which is used to deploy SD-WAN networks
- Huawei SD-WAN Solution supports the following ZTP modes:
  - Email-based deployment: URL parameters carrying deployment information are sent by email. This mode has low technical requirements on device installation personnel.
  - DHCP-based deployment: Devices obtain the IP address of iMaster NCE-WAN through DHCP. This mode has almost no technical requirements on device installation personnel.
  - USB-based deployment: USB flash drives are used to transmit deployment information. This mode has low technical requirements on device installation personnel.

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.

SD-WAN Networking and Planning

HUAWEI

Foreword

- To meet service development requirements, enterprises require flexible and reliable wide area networks (WANs).
- Huawei offers an SD-WAN networking sub-solution to help enterprises implement on-demand, flexible, automated networking between enterprise branches, data centers (DCs), and clouds through the SD-WAN controller (iMaster NCE-WAN).
- This course describes the implementation and design of SD-WAN flexible networking.

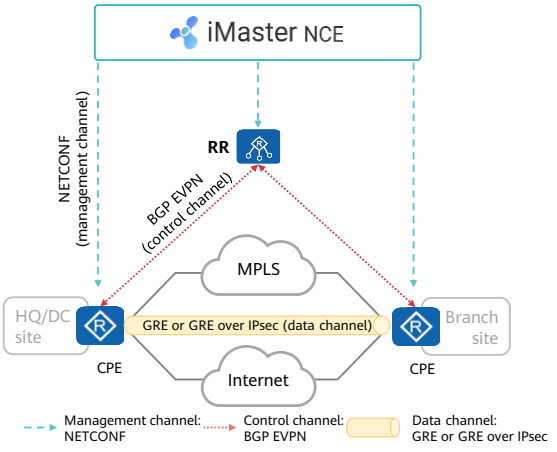
Objectives

- Upon completion of this course, you will be able to:
  - Describe basic implementations of SD-WAN flexible networking.
  - Describe the networking schemes of the SD-WAN Solution.
  - Complete the planning and design of an enterprise SD-WAN network using the design principles and methods described in this course.

Contents

Flexible Networking of Huawei SD-WAN Solution

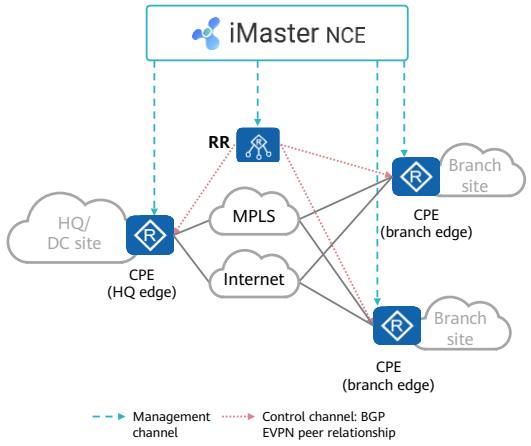
- Huawei SD-WAN Solution uses iMaster NCE-WAN to centrally manage devices and flexibly establish IP overlay tunnels, through which service traffic is forwarded.
- Huawei SD-WAN Solution implements flexible networking through the following channels:
  - Management channel: connects iMaster NCE-WAN to the control layer, so that iMaster NCE-WAN provides service portal pages for **SD-WAN service presentation** and **configuration delivery**.
  - Control channel: controls establishment of IP overlay tunnels and transmits service routes.
  - Data channel: forwards data over carrier networks or enterprise-built networks and allows for flexible traffic steering based on network Service Level Agreement (SLAs).



HUAWEI

- In Huawei SD-WAN Solution, IP overlay tunnels are mainly GRE or GRE over IPsec tunnels.
- Services on the management plane are mainly implemented by iMaster NCE-WAN.
- This course mainly describes the implementation principles of services on the control plane and forwarding plane.

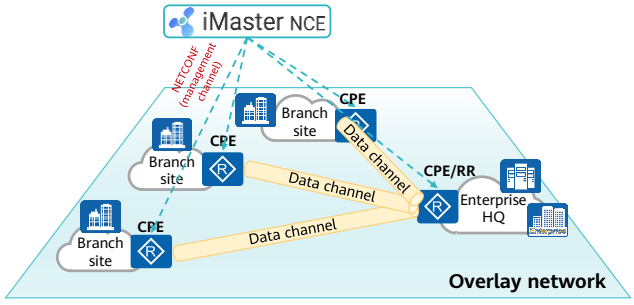
Components of Huawei SD-WAN Solution



- **iMaster NCE-WAN**: controls network elements (NEs) on the entire SD-WAN network, orchestrates network services, and performs monitoring and O&M.
- **Route reflector (RR)**: also called an area controller, distributes VPN routing and tunnel information between CPEs based on VPN topology policies.
- **CPE**: refers to an AR router deployed at the egress of a campus network.
- **Edge site**: A CPE is used as the edge router on the WAN side. A control channel is established between the edge site and RR. The RR controls route advertisement. Secure data channels are established between multiple sites. A CPE at the edge site is the source or destination of an EVPN interconnection tunnel, and can be considered as the border of an EVPN network.

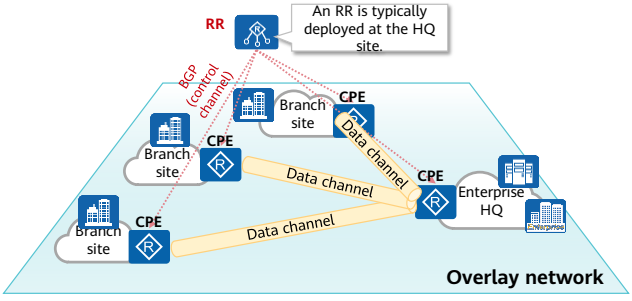
Management Channel

- iMaster NCE-WAN establishes management channels with CPEs through NETCONF.
- Management channels are used for:
  - Unified management of CPEs, automatic service delivery, and unified control of overlay networks
  - Network status monitoring and automatic application optimization
  - Device status monitoring



Control Channel

- Control channels are established between CPEs and RRs through BGP EVPN to control the establishment of data channels and transmission of service routes.
  - An RR reflects transport network port (TNP) information, IPsec SA information, and service routes. Typically, an RR is deployed at an HQ site.
  - IP overlay tunnels (data channels) are established mainly depending on TNP and IPsec SA information.
  - When service routes are reflected by an RR, the RR can control transmission of the service routes based on routing policies (typically delivered by iMaster NCE-WAN) to control the service traffic direction.
  - An RR is also called an area controller.

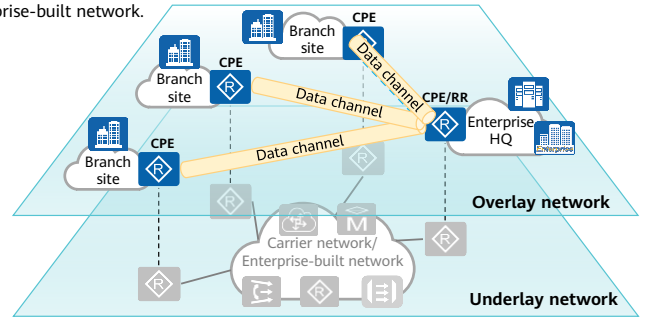


HUAWEI

- TNP information and other related information will be described in the following slides.

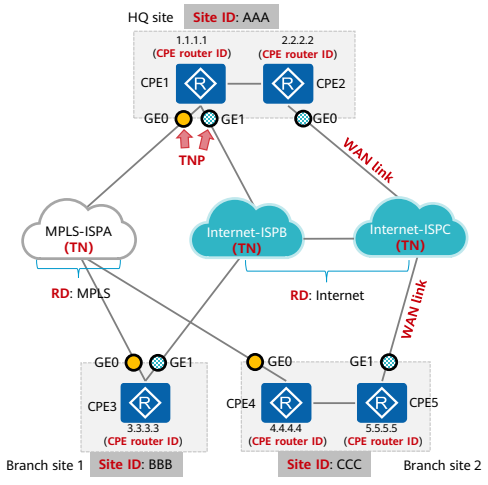
Data Channel

- Data channels are established between CPEs through GRE or GRE over IPsec. A network built by data channels is called an overlay network, which is constructed on an underlay network.
  - An overlay network is constructed based on data channels. Different overlay networks can be constructed for different services or departments.
  - An underlay network is a traditional WAN-side network and mainly carries services of overlay networks. An underlay network can be the Internet, an MPLS network, or an enterprise-built network.



HUAWEI

Basic Concepts of Huawei SD-WAN Networking



- **Transport network (TN)**: refers to a WAN access network provided by an ISP, which implements WAN interconnection of enterprise branches.
- **Routing domain (RD)**: If different TNs can communicate with each other (for example, the Internet networks provided by ISPB and ISPC in the figure on the right), they are considered to be in the same RD.
- **Site ID**: indicates the globally unique ID of a tenant site, which is allocated by iMaster NCE-WAN.
- **CPE router ID**: indicates the globally unique ID of a CPE. A site can have one or two CPEs. The router ID of a CPE is typically the CPE's loopback interface address.
- **WAN link**: refers to a link connecting to a WAN interface. The IP address allocation mode, link negotiation rate, and bandwidth can be configured for a WAN link.
- **TNP**: refers to the WAN interface on a CPE for connecting to a transport network. The key information about a TNP includes the site ID, CPE router ID, transport network ID, public IP address, private IP address, and tunnel encapsulation mode.

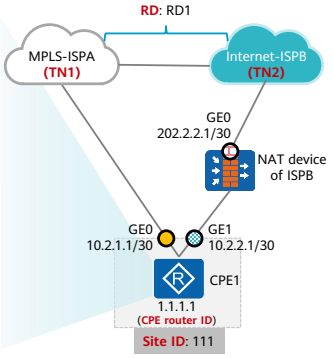
HUAWEI

- TN and RD information are used for enumerating and establishing overlay tunnels.
- A site ID is used as the next hop for addressing and data forwarding during user routing.
- CPEs' router IDs are defined for establishing BGP peer relationships between sites.
- TNPs are defined for establishing tunnels.

TNP

- A TNP mainly describes WAN link information about a **site**, depending on which **control and data channels are established**.
- The following table describes TNP information.

TNP Information	Example		Description
Site ID	111		Site to which the TNP belongs
CPE Router ID	1.1.1.1		CPE to which the TNP belongs
TNP ID	1	2	TNP number
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

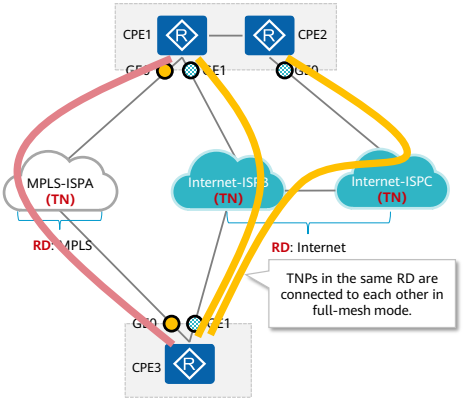


HUAWEI

- TN and RD information are used for enumerating overlay tunnels.
  - Tunnel enumeration ensures that all available tunnels are established.
- A router ID is used for establishing a control channel.
- A site ID specifies the next hop for data forwarding.
- An interface number can be used as a TNP ID.
- The public and private IP addresses refer to the source and destination IP addresses of the control and data channels.
  - To establish a data channel between CPEs behind NAT devices, the CPEs need to learn the post-NAT IP addresses (public IP addresses).
    - Typically, a CPE learns public IP addresses through STUN technology.

RD and TN

- A CPE enumerates data tunnels based on the **RD** and **TN** IDs in the TNP information.
- Tunnel enumeration ensures the reliability of SD-WAN networks and improves service awareness.
- Tunnel enumeration rules:
  - Only TNPs in the same RD can be connected with each other in enumeration mode and fully meshed.
  - TNPs in different RDs cannot be connected in enumeration mode.



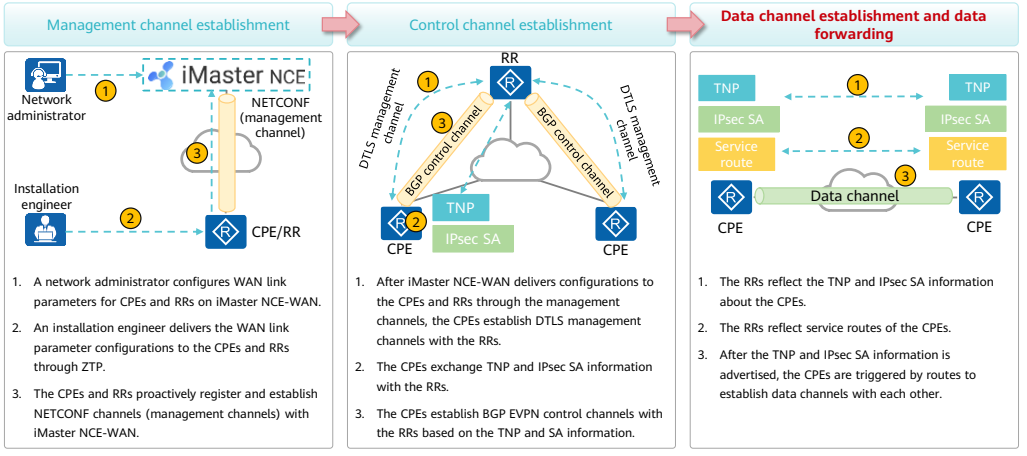
HUAWEI

- Tunnel enumeration is performed before data channel establishment to ensure that all available data channels are established.
- Prerequisites for tunnel enumeration:
  - CPEs have learned service routes from the peer site.
  - CPEs have learned TNP information from the peer site.
- The details about how to learn service routes and TNP information will be described in the following slides.

Contents

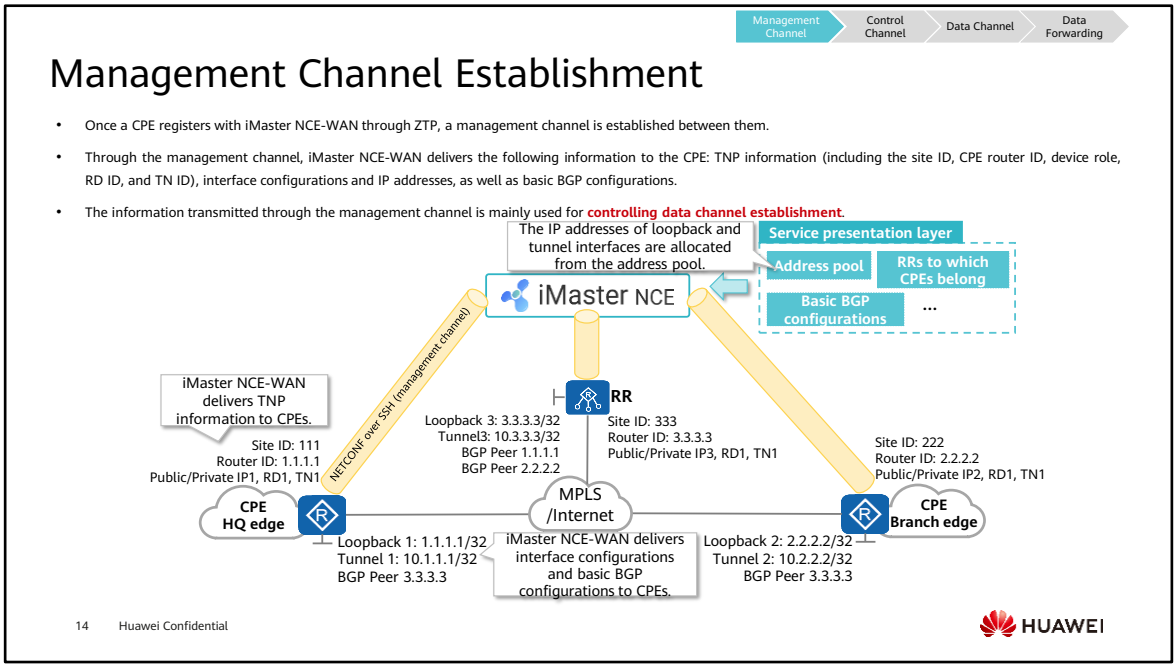
SD-WAN Networking Process

- The SD-WAN networking process is as follows:



HUAWEI

- A management channel is used for establishing control channels and delivering basic configurations.
- Control channels are used to control establishment of data channels.
- Data channels transmit user data.
- TNP information is exchanged twice as follows:
  - TNP information is exchanged for the first time for establishing control channels between RRs and CPEs.
  - TNP information is exchanged for the second time for establishing data channels between CPEs.

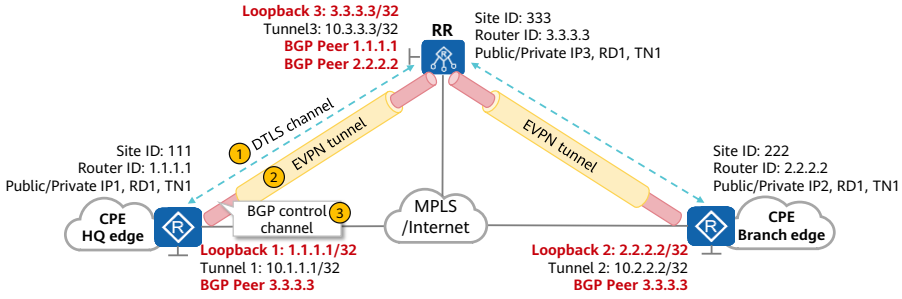


- For details about ZTP and management channel establishment, learn the course *SD-WAN Deployment*.
- An RR is typically deployed at the same site as CPEs of the HQ. For ease of understanding, an RR is deployed independently of CPEs of the HQ in the above figure.
- All configurations on the CPEs and RR are delivered by iMaster NCE-WAN.
- The IP address of a CPE's loopback interface is used as the CPE's router ID.
- Tunnel interfaces are used for establishing management channels, which use the GRE over IPsec encapsulation mode.
- iMaster NCE-WAN also delivers basic BGP configurations to instruct CPEs to establish BGP peer relationships with the RR through loopback interfaces.
- Site IDs are generally allocated to CPEs in ascending order based on the sequence in which the CPEs go online.

Management Channel | Control Channel | Data Channel | Data Forwarding

Control Channel Establishment

- Control channels use BGP to transmit data channel parameters. To ensure the control channel security, BGP is carried over EVPN tunnels.
- Control channels mainly transmit TNP information, IPsec SA information, and service routes. The TNP and IPsec SA information is used for establishing data channels. Service routes are used for traffic steering during data forwarding.



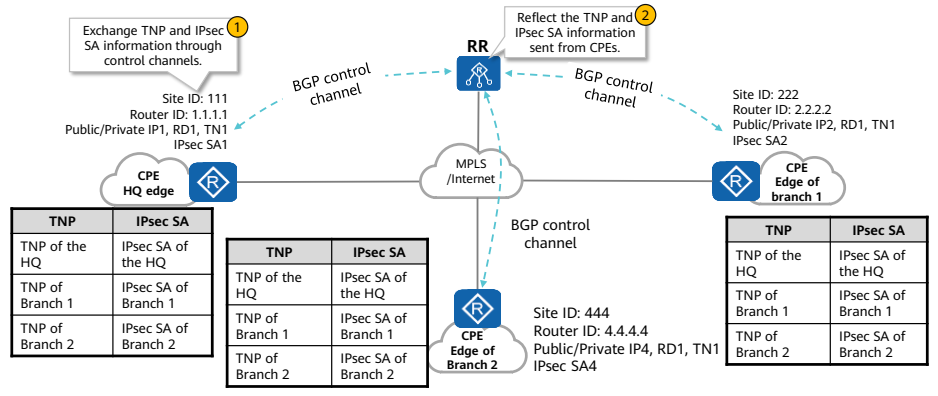
HUAWEI

- A control channel is established as follows:
  1. A CPE establishes a DTLS tunnel with an RR based on the TNP information delivered by iMaster NCE-WAN.
     - The DTLS tunnel is established to ensure security of TNP information exchanged between the CPE and RR.
  2. The CPE and RR exchange TNP and IPsec SA information through the DTLS tunnel, and establish an EVPN tunnel based on the TNP and IPsec SA information.
     - In Huawei SD-WAN Solution, EVPN tunnels use the GRE over IPsec encapsulation mode.
     - The source and destination IP addresses of an EVPN tunnel are determined by the TNP information about the CPE and RR.
  3. The CPE and RR establish a BGP peer relationship through loopback interfaces.
     - All CPEs are BGP RR clients of the RR.
     - The CPE and RR send BGP packets to each other to exchange the TNP and IPsec SA information required for establishing data channels.

Management Channel | Control Channel | Data Channel | Data Forwarding

TNP and IPsec SA Information Exchange

- An RR can reflect TNP and IPsec SA information about different CPEs through control channels.
- After TNP and IPsec SA information is exchanged, CPEs **cannot establish data channels immediately with each other**. Instead, data channel establishment between them is **triggered by service routes**.



TNP	IPsec SA
TNP of the HQ	IPsec SA of the HQ
TNP of Branch 1	IPsec SA of Branch 1
TNP of Branch 2	IPsec SA of Branch 2

TNP	IPsec SA
TNP of the HQ	IPsec SA of the HQ
TNP of Branch 1	IPsec SA of Branch 1
TNP of Branch 2	IPsec SA of Branch 2

TNP	IPsec SA
TNP of the HQ	IPsec SA of the HQ
TNP of Branch 1	IPsec SA of Branch 1
TNP of Branch 2	IPsec SA of Branch 2

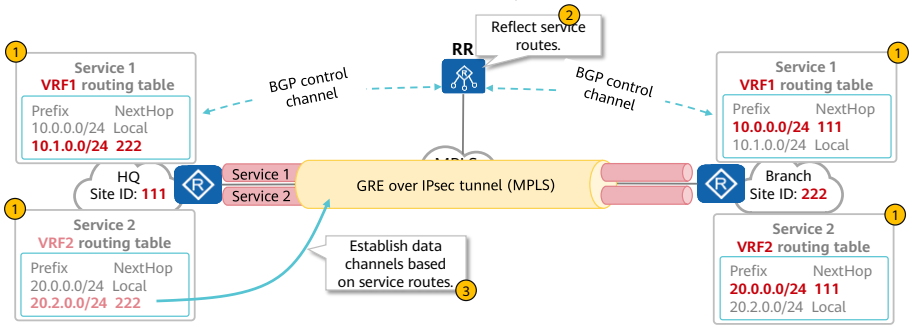
HUAWEI

- TNP and IPsec SA information is exchanged as follows:
  - Through the BGP control channel, a CPE sends the local TNP and IPsec SA information to the RR through a BGP route.
    - The RD ID and TN ID in the TNP information are used for enumerating and establishing data tunnels.
    - The public and private IP addresses are used as the source and destination IP addresses of a data tunnel.
    - IPsec SA information is used to encrypt the data tunnel.
    - The site ID is used for traffic steering. The functions of a site ID will be detailed in the following slides.
  - The RR sends the BGP route received from the CPE to all CPEs associated with the RR.

Management Channel | Control Channel | Data Channel | Data Forwarding

Service Route Transmission and Data Channel Establishment

- CPEs' service routes are transmitted through **BGP control channels**.
- Based on the next hop of a service route as well as the TNP and IPsec SA information, a CPE is triggered to create a data channel with the peer CPE.
  - The next hop of an SD-WAN service route is determined by the **site ID**.
  - The CPE searches the connection table for TNP information based on the site ID, and then **enumerates data channels** based on the TNP information.



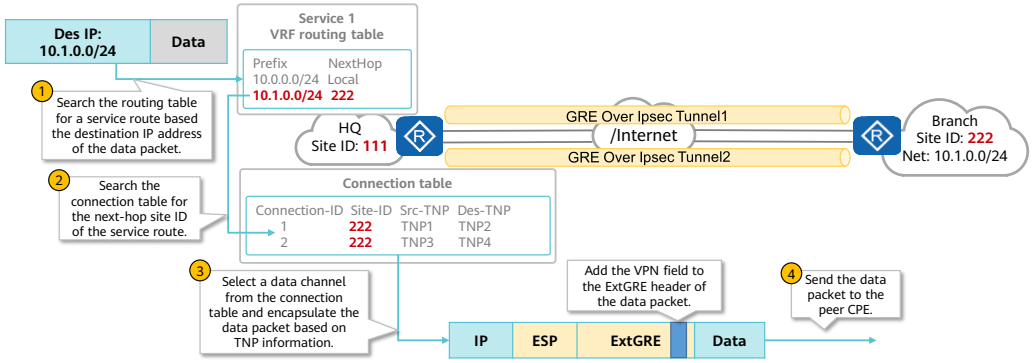
HUAWEI

- Data channels use the GRE over IPsec encapsulation mode.
- VPN technology is used to isolate routes of different services.
  - Different service routes belong to different VPNs. Therefore, different service routes are placed in different VRF routing tables.
- Multiple links may be established between different CPEs. If different routes of a service are in the same RD, the CPEs enumerate data channels.
  - Each data channel is also called an EVPN connection.
  - Multiple data channels are carried in one GRE over IPsec tunnel.

Management Channel | Control Channel | Data Channel | Data Forwarding

Service Data Forwarding Process

- After a data channel is established between CPEs, the CPEs will select data channels based on the routing and connection tables for data forwarding.

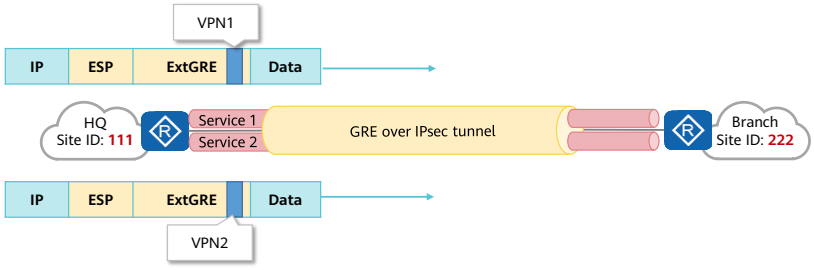


HUAWEI

- There may be multiple data channels destined for the same site ID. Therefore, a CPE needs to select a data channel for data forwarding.
  - Typically, a data channel is selected based on the priority. However, there are various data channel selection policies. For details, learn the course *SD-WAN Application Experience*.
- Data transmitted over a data tunnel will be re-encapsulated before being sent.
  - The data encapsulation mode is GRE over IPsec.
  - The VPN field is added to the ExtGRE header to identify the service (VPN) to which the data belongs during data forwarding.

Service Traffic Isolation

- In Huawei SD-WAN Solution, routes of different services are placed in different VRF routing tables for isolation.
- On the forwarding plane, the VPN field in the ExtGRE header is used to differentiate service traffic.

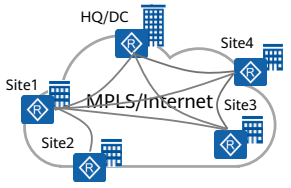


Contents

Challenges Facing Flexible Networking

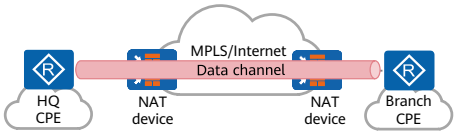
- SD-WAN flexible networking faces the following challenges on the live network:
  - How to flexibly control the overlay topology?
  - How to establish data channels between CPEs behind NAT devices?

Diversified overlay topology requirements



- Different services may use different overlay topologies.
- In addition to hub-spoke networking and full-mesh networking, partial-mesh networking may be used.

Establishment of data channels between CPEs through NAT traversal

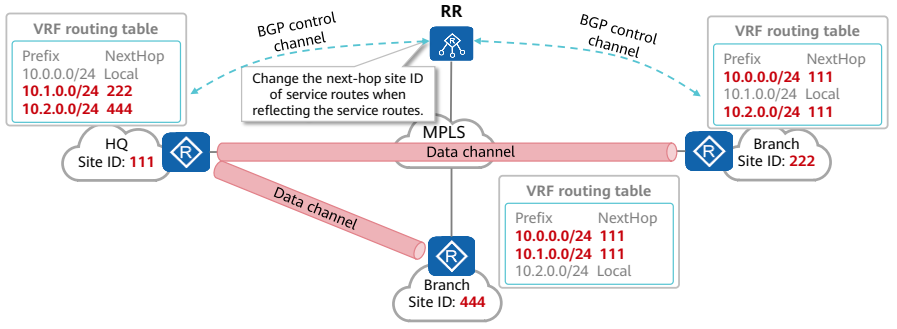


- On the live network, some CPEs are deployed behind NAT devices.
- The CPEs behind NAT devices cannot directly establish data channels with each other.

HUAWEI

Overlay Topology Control

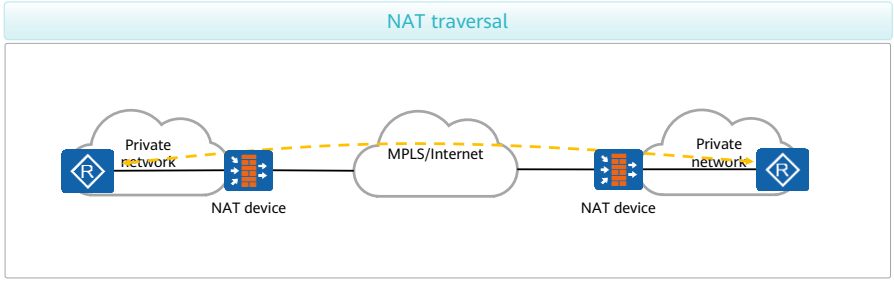
- Data channels of CPEs constitute an overlay topology and are established mainly depending on service routes.
  - CPEs establish data channels based on the next-hop site ID of service routes and TNP information.
  - Establishment of a data channel can be controlled by controlling the next-hop site ID of a service route.
  - When establishment of data channels is controlled, the overlay topology is controlled.



- Different services may use different overlay topologies.
- Because an RR reflects service routes of all CPEs associated with it, the RR can change the next-hop site ID of service routes to control the overlay topology.
- When the hub-spoke topology is used, the RR only needs to change the next-hop site ID of service routes to the site ID of the hub site.
- When the full-mesh topology is used, a full-mesh network can be built without the need to change the next-hop site ID of service routes.
- When the partial-mesh topology is used, the next-hop site ID of only some service routes needs to be changed.
- Because an RR can control the overlay topology, it is also called an area controller.

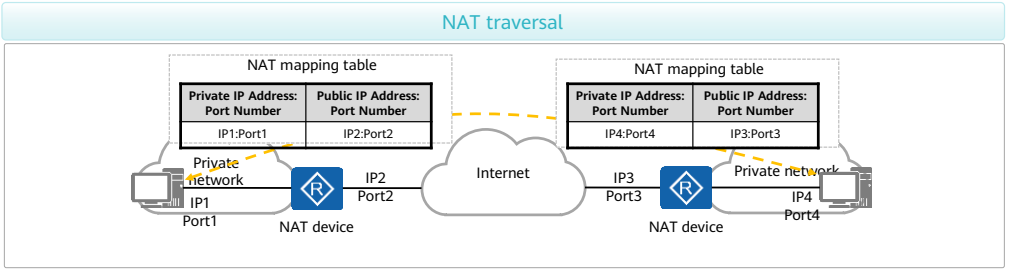
NAT Traversal

- Although NAT allows private network users to access a public network, it has the following defects:
  - Private IP addresses are hidden, making it difficult for external network devices to access private network devices.
- To allow CPEs deployed behind NAT devices to directly establish data channels with each other, Huawei SD-WAN Solution adopts Session Traversal Utilities for NAT (STUN) technology.



Overview of STUN

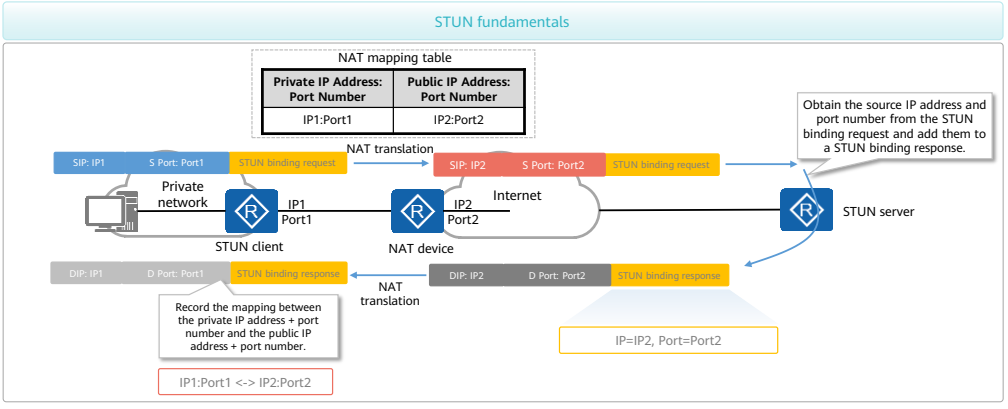
- By leveraging the cone NAT feature, NAT traversal technology is used to create NAT mapping entries on NAT devices and then perform hole punching on the NAT devices. In this manner, connections can be established between private networks based on the NAT mapping entries.
- STUN is mainly used to obtain hole punching information (namely, the mapping between pre-NAT private IP addresses and port numbers and post-NAT public IP addresses and port numbers) on NAT devices. CPEs establish data channels traversing NAT devices using other technologies.



- STUN is defined in RFC 3489 and is a complete NAT traversal solution.
- In RFC 5389, the STUN protocol is positioned as a tool used to allow packets to traverse NAT devices, rather than a complete solution. RFC 5389 supports TCP traversal, which is not supported in RFC 3489.

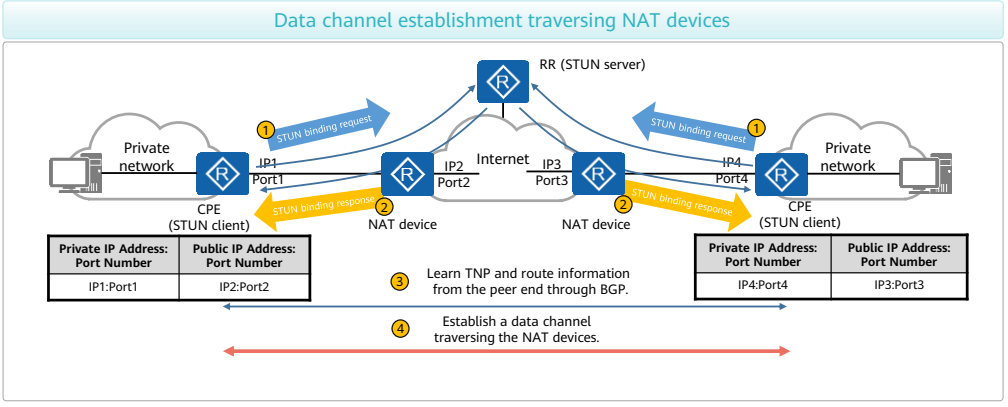
STUN Fundamentals

- STUN is implemented as a client-server protocol. Through packet exchange with a STUN server, a STUN client can detect a NAT device and determine the IP address and port number allocated by the NAT device.



Data Channel Establishment Traversing NAT Devices

- In SD-WAN scenarios, STUN can be used to allow CPEs deployed behind NAT devices to establish data channels traversing NAT devices.

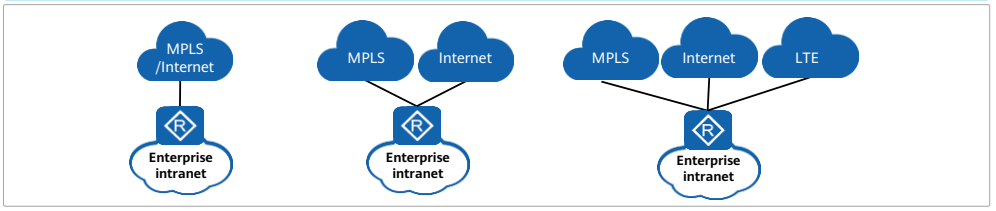


- Each STUN client sends a binding request to the STUN server.
- The STUN server obtains the source IP address and port number from the binding request, and sends a binding response to each STUN client.
- The STUN client obtains an IP address and a port number from the binding response, and compares the obtained IP address and port number with the source IP address and port number carried in the binding request. If they are different, a NAT device is used between the STUN client and STUN server.
- STUN clients learn each other's TNP information (including the pre-NAT and post-NAT IP addresses and port numbers) through BGP routes.
- After the preceding STUN messages are exchanged, a data channel is established between the STUN clients traversing the NAT devices based on the hole punching mechanism.

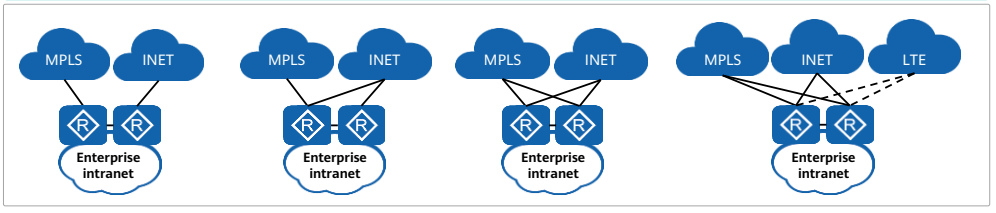
Contents

Typical WAN-side Networking Modes of Sites

Single-site single-CPE networking (a single CPE supporting a maximum of 3 links)

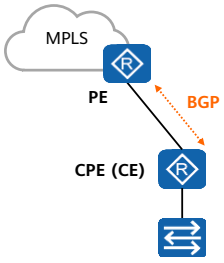


Single-site dual-CPE networking (a maximum of 6 links, with each CPE supporting a maximum of 3 links)



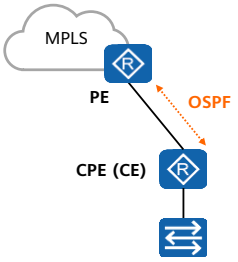
Underlay WAN-side Routing: BGP and OSPF

WAN connection using BGP



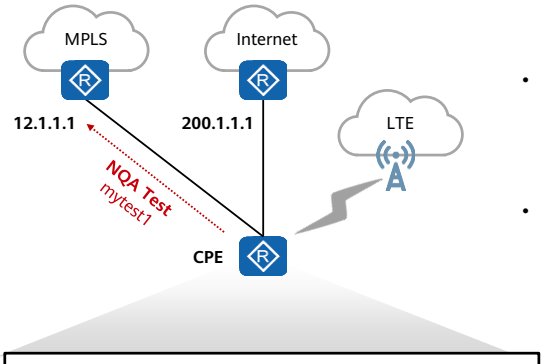
- If a CPE is connected to an MPLS WAN network and the PE uses BGP to exchange routes, the CPE typically needs to use BGP to exchange routing information with the PE.
- iMaster NCE-WAN can configure route filtering rules based on IP network segments to control the advertisement and receiving of BGP routes.

WAN connection using OSPF



- When a CPE is connected to an MPLS WAN network and the PE uses OSPF to exchange routes, OSPF needs to be deployed on the CPE.
- iMaster NCE-WAN can configure OSPF priorities and configure blacklist- or whitelist-based route filtering policies to control the advertisement and receiving of OSPF routes.

Underlay WAN-side Routing: Static

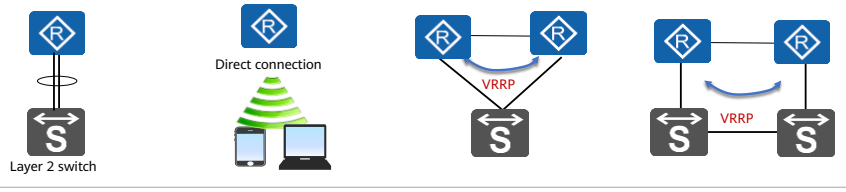



```
[CPE] ip route-static 10.1.0.0 16 12.1.1.1 track nqa admin mytest1
[CPE] ip route-static 10.1.0.0 16 200.1.1.1 preference 80
```

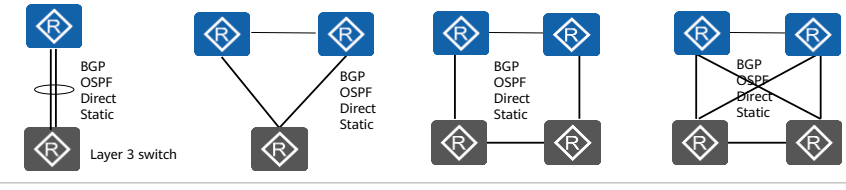
- Static routes are used in many scenarios. For example, default static routes are configured for Internet access, and blackhole routes are configured to prevent loops.
- On the live network, static routes do not involve protocol interaction and cannot detect faults on indirectly connected links of the WAN. This may cause service interruption.
- When configuring a static route on a CPE, you can enable the track function and use an NQA test instance to detect the link quality. If the detection fails, the CPE considers that the WAN is faulty and automatically selects the backup link for data forwarding, preventing long-time service interruption.

Typical LAN-side Networking Modes of Sites

Interconnection with a Layer 2 network or hosts on the LAN side (through a CPE that functions as a gateway)

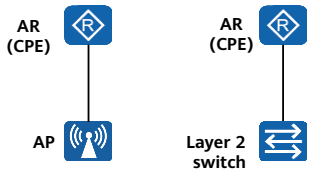


Interconnection with a Layer 3 network on the LAN side through BGP, OSPF, or static routes



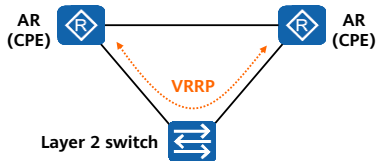
LAN-side Site Model: Layer 2 Interconnection

Single-CPE Layer 2 interconnection



- If a site has only one CPE and a small scale, the CPE can be directly connected to terminals at the site through LAN-side interfaces.
- If the CPE has insufficient LAN-side interfaces, an access switch can be connected to the CPE.

Dual-CPE Layer 2 interconnection

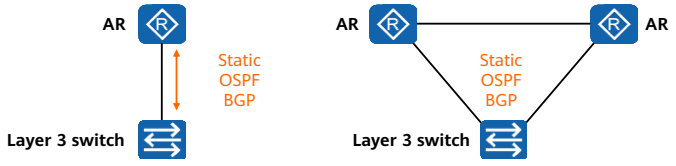


- When a site has two CPEs, VRRP is usually deployed on the CPEs.
- LAN-side switches can be stacked.
- An interlink needs to be established between the CPEs to forward service packets between them.

- If a site has only one CPE, LAN-side connections are simple.
  - If the site has a small scale (for example, a SOHO site), the CPE can be directly connected to terminals at the site through LAN-side interfaces.
  - If the CPE has insufficient LAN-side interfaces, an access switch can be connected to the CPE through one-armed routing.
- If a site has two CPEs, VRRP is deployed on the CPEs to prevent the dual-CPE architecture from affecting the LAN-side network.
  - 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-side network.
  - If the two CPEs are directly interconnected, an interlink needs to be established between them to forward service packets. The interlink can also be an Eth-Trunk.

LAN-side Site Model: Layer 3 Interconnection

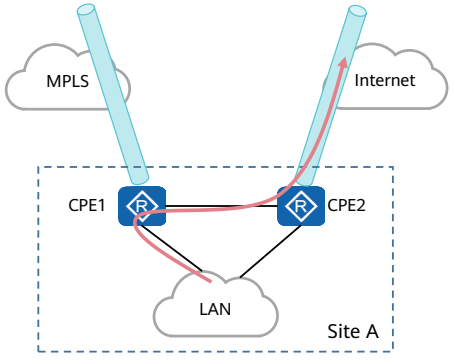
Single- or dual-CPE Layer 3 interconnection



- In the Layer 3 interconnection scenario, if only one CPE is deployed at a site, only a routing protocol needs to be configured on the LAN side based on requirements of LAN-side devices.
- In the Layer 3 interconnection scenario, a CPE can only interconnect with a standalone device or a stack on the LAN side. Therefore, when a CPE needs to interconnect with multiple devices, these devices must be stacked.

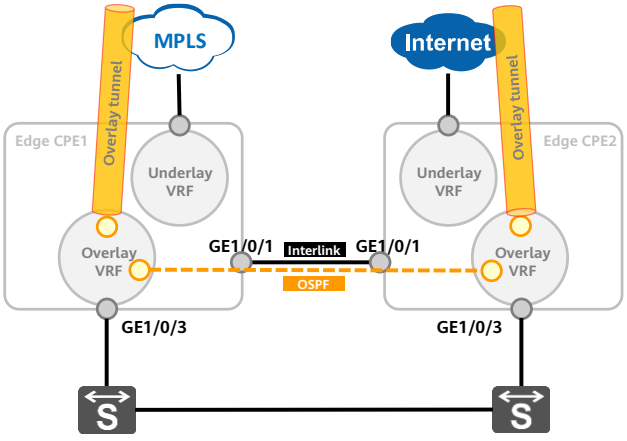
- For a large site, the site network has a complex structure and complex network facilities (for example, Layer 3 core devices). Therefore, the egress routers must support direct connection or dual-homing to Layer 3 devices. BGP, OSPF, and static routing are supported.
- In the Layer 3 interconnection scenario, if only one CPE is deployed at a site, only a routing protocol needs to be configured on the LAN side based on requirements of LAN-side devices. If a CPE needs to interconnect with two LAN-side devices, the LAN-side devices must be stacked to function as a whole.

Dual-Gateway Networking Scenario



- Traffic routing for a dual-gateway site
  - Two CPEs are deployed at a branch site as gateways, and each gateway has at least one WAN link, such as an MPLS or Internet link.
  - In this deployment model, to support application-based traffic steering and ensure reliability, the two CPEs need to be considered as a whole so that users' service traffic can be routed across the CPEs.
  - When LAN-side services need to access a remote site, the local site can select one from the two WAN links through the interlink between the CPEs.
  - CPE1 can not only check the local link status, but also check the link status and SLA status of CPE2. In this way, CPE1 can select a WAN link based on a pre-configured application-based traffic steering policy.

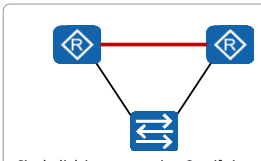
Dual-Gateway Interlink Solution



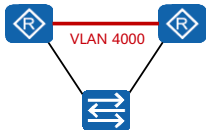
- **Dual-gateway design**
  - In the SD-WAN Solution, services are processed on a per-site basis, and the links on the two gateways (CPEs) at a site are regarded as a whole and all participate in traffic steering.
- **Interlink design**
  - The two gateways establish an interlink with each other through interconnection interfaces and forward service (VPN) traffic between them through sub-interfaces.
  - Information, such as OSPF routing information, is synchronized between the gateways through interlink sub-interfaces.

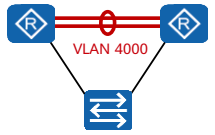
Interfaces Connecting the Two Gateways

- For a dual-gateway site, configure a Layer 3 or Layer 2 link for interconnection between the two CPEs (gateways). By default, a Layer 3 link is used.

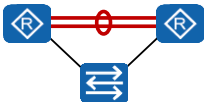


Single-link interconnection: Specify interfaces on the CPEs for interconnecting with each other.

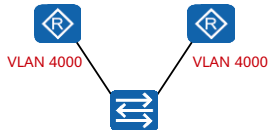




If only one interface is specified on each CPE, the two CPEs are interconnected with each other through a single link. If two interfaces are specified on each CPE, the system automatically binds the two interfaces into an Eth-Trunk, and the two CPEs communicate with each other through VLANIF interfaces.



Eth-Trunk interconnection: Specify two interfaces on each CPE for interconnecting with each other. The system automatically bundles the two interfaces into an Eth-Trunk.



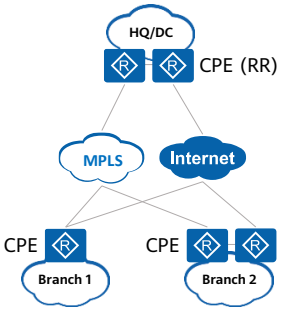
A LAN-side Layer 2 link is used as the data forwarding channel between the CPEs. Data between the CPEs and data from the LAN side to the CPEs are isolated through VLANs, without affecting each other.

- To improve the reliability of egress links, multiple links are usually provided, that is, one active link and one standby link are used. This design is simple and reliable. The standby link is in backup state and does not forward network traffic in normal cases. Therefore, enterprise customers need to pay extra fees for reliability.
- The SD-WAN Solution provides link backup. In this solution, multiple uplinks of a site are active at the same time and services can be load balanced among the links according to a preconfigured traffic scheduling policy. If a link is faulty, the link fault or quality deterioration can be detected within sub-seconds. Then, services can be switched from the faulty link to an operational link. This mechanism ensures link reliability and makes full use of enterprises' link resources, providing high access bandwidth and facilitating interconnection between enterprise sites.

Contents

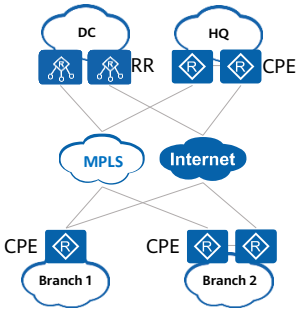
RR Deployment Modes

Combined deployment of a site and RR



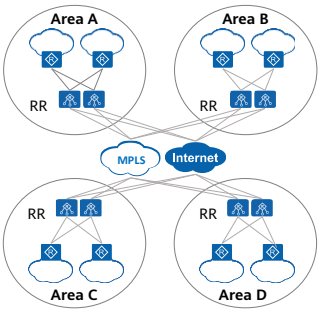
A site that functions as an RR not only undertakes tasks on the control plane but also forwards service traffic of other sites on the forwarding plane.

Independent deployment of RRs



A site that functions as an RR does not have a LAN network, does not function as a hub for communication between other sites, and does not transmit service data. Such a site just undertakes tasks on the control plane.

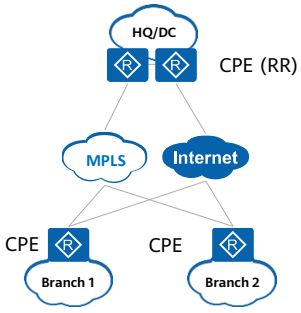
Deployment of RRs in different areas



At least one pair of RRs is deployed in each area. RRs in different areas establish BGP EVPN peer relationships with each other to advertise and learn VPN routes in different areas.

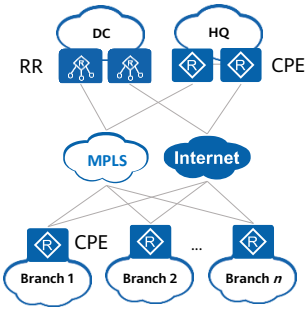
RR Deployment Scenario: Enterprise-built Network

Combined deployment of a site and RR



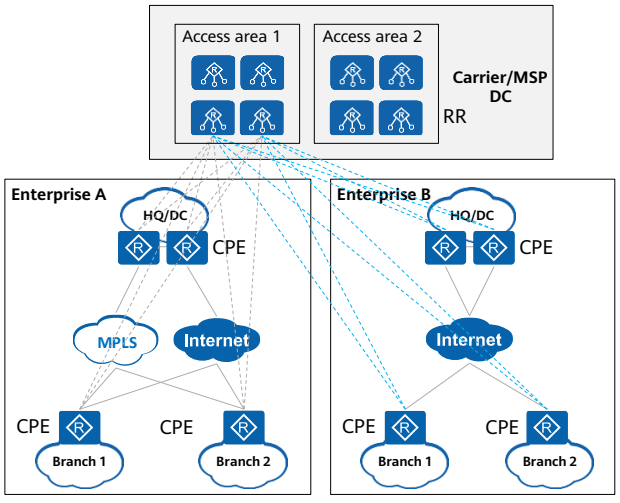
For small and midsize enterprises with a small number of branch sites and no heavy traffic between branches and the HQ/DC, it is recommended that the hub sites function as RRs and undertake tasks on the control plane.

Independent deployment of RRs



For large-scale enterprises with a large number of sites and high network reliability requirements, it is recommended that the RRs be deployed independently.

RR Deployment Scenario: Carrier/MSP Resale Scenario



- **Deployment of MSP RRs**

  In carrier/MSP resale scenarios, it is recommended that RRs be deployed independently. An MSP administrator creates RRs, divides access areas, and allocates the RRs to the corresponding access areas.

- **MSP RR shared by multiple tenants**

  For small and midsize enterprises with a small number of sites, multiple tenants share one RR. Sites of different enterprise tenants are connected to the same multi-tenant RR, and the RR undertakes tasks on the control plane.

- **Independent allocation of MSP RRs**

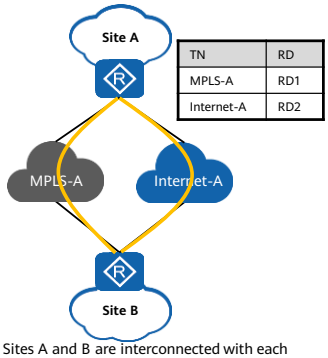
  For large enterprises with a large number of sites, RRs can be allocated independently. MSPs allocate RRs in access areas to the enterprises.

- **Deployment of tenant RRs**

  For new enterprise customers, if the live network is a carrier's private line network and the carrier/MSP DC is unreachable, the hub site of a tenant can act as an RR.

Overlay Network Planning and Connection Management

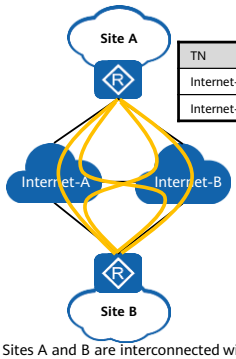
TNs are in different RDs.



TN	RD
MPLS-A	RD1
Internet-A	RD2

Sites A and B are interconnected with each other through two isolated TNs on the underlay network. Two TNs are planned to be in different RDs and the two sites establish two tunnels with each other.

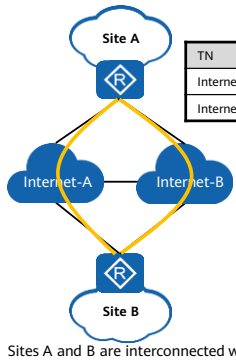
TNs are in the same RD.



TN	RD
Internet-A	RD1
Internet-B	RD1

Sites A and B are interconnected with each other through two interconnected TNs on the underlay network. Two TNs are planned to be in the same RD and the two sites are interconnected with each other in full-mesh mode.

TNs are physically interconnected in different RDs.



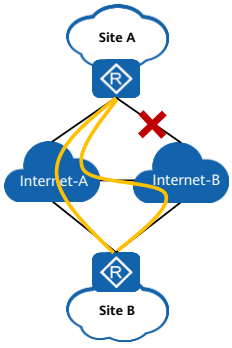
TN	RD
Internet-A	RD1
Internet-B	RD2

Sites A and B are interconnected with each other through two interconnected TNs on the underlay network. Two TNs are planned to be in different RDs and the two sites establish two tunnels with each other.

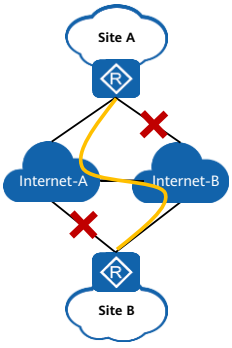
Reliability Comparison Between Overlay RD Planning Modes

Same RD, different TNs

Fault scenario A



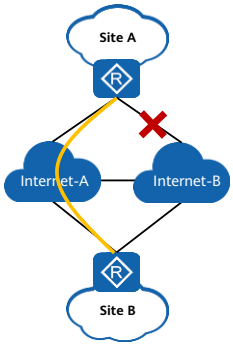
Fault scenario B



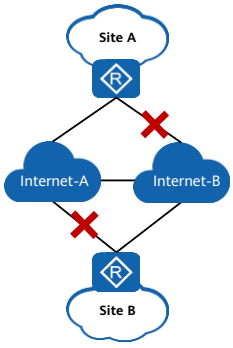
Different TNs in the same RD: higher networking reliability, but more virtual connection resources consumed

Physically interconnected TNs, different RDs

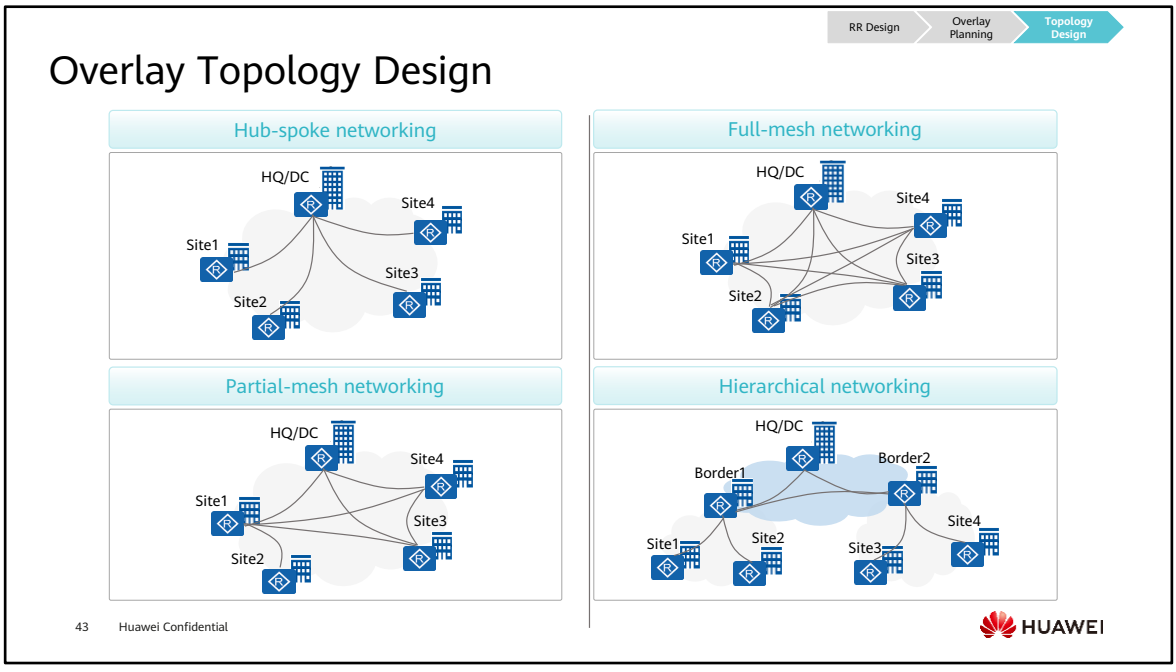
Fault scenario A



Fault scenario B

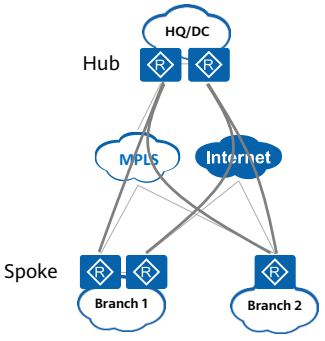


Different TNs in different RDs: smaller number of connections and larger networking scale, but lower networking reliability



- Hub-spoke networking
  - The HQ or DC of an enterprise functions as a hub site, and enterprise branches function as spoke sites and access servers deployed at the HQ or DC through WANs.
- Full-mesh networking
  - Different branches of an enterprise can directly communicate with each other, without the need to divert traffic through intermediate nodes.
- Partial-mesh networking
  - A partial-mesh network can be considered as a type of special full-mesh network. If direct underlay network connections are available between two sites, traffic is directly transmitted between the sites. Otherwise, traffic between the sites is forwarded through a redirect site, to which both sites are connected.
- Hierarchical networking
  - The hierarchical network model can be considered as a combination of single-layer network models. A WAN is divided into multiple areas, which are interconnected through a centralized backbone area. In this way, sites can communicate with each other across areas.

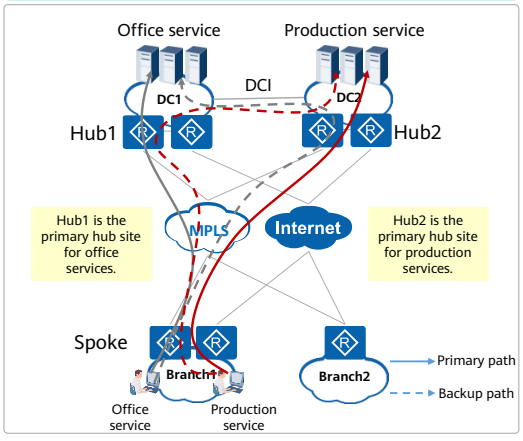
Hub-Spoke Networking Scenario



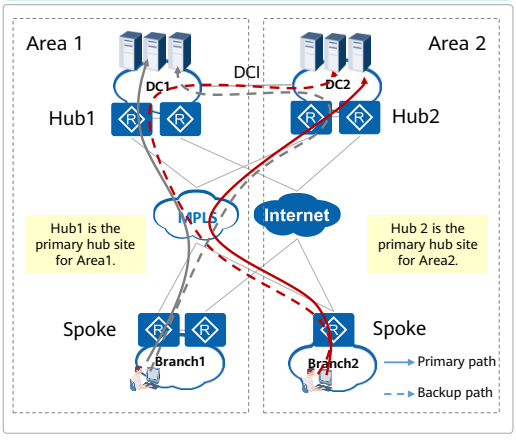
- **Solution description**
  - The HQ or DC of an enterprise functions as a hub site, and enterprise branches function as spoke sites and access servers deployed at the HQ or DC through WANs.
  - If enterprise branches need to communicate with each other, traffic needs to be forwarded by the hub site. External traffic of all branch sites must be first sent to the hub site.
- **Application scenarios**
  - Enterprises whose major service traffic is sent from branch sites to hub sites and applications are centrally stored on servers at the HQ and DC sites
  - Enterprises with almost no traffic transmitted between branch sites, for example, chain enterprises. The major service traffic of chain enterprises is sent from chain branches to the HQ or DC, and there is almost no traffic between chain branches.

Dual-Hub Active-Active Networking Scenario

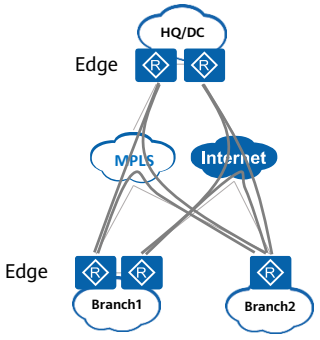
Scenario 1: active-active hub sites based on service network segments



Scenario 2: active-active hub sites based on spoke sites

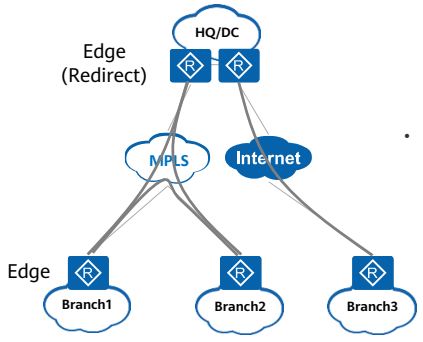


Full-Mesh Networking Scenario



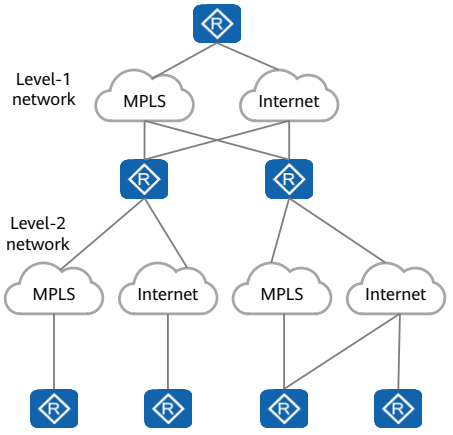
- **Solution description**
  - Different branches can directly communicate with each other, without the need to divert traffic through intermediate nodes.
- **Application scenarios**
  - Small enterprises with a small number of sites or large enterprises whose branches need to communicate with each other
  - Large enterprises with collaborative services (for example, high-value applications such as VoIP and video conferencing). These services have high requirements on network performance such as the packet loss rate, latency, and jitter, and require direct communication between branch sites.
  - The full-mesh networking is simple and features high service access efficiency but average scalability. It is applicable to networks with 10 to 100 branch sites.

Partial-Mesh Networking Scenario



- **Solution description**
  - A partial-mesh network can be considered as a type of special full-mesh network. If direct underlay network connections are available between two sites, traffic is directly transmitted between the sites. Otherwise, traffic between the sites is forwarded through a redirect site, to which both sites are connected.
- **Application scenarios**
  - The underlay networks of the sites that need to directly communicate with each other cannot be directly interconnected. In this case, the sites can communicate with each other through the redirect site.
  - In the full-mesh networking scenario, the underlay networks of the sites that need to directly communicate with each other can be directly interconnected. For reliability purposes, a redirect site is deployed. In this manner, when the underlay networks of the sites are faulty and the sites cannot directly communicate with each other, traffic of the sites can be forwarded by the redirect site.

Hierarchical Networking Scenario

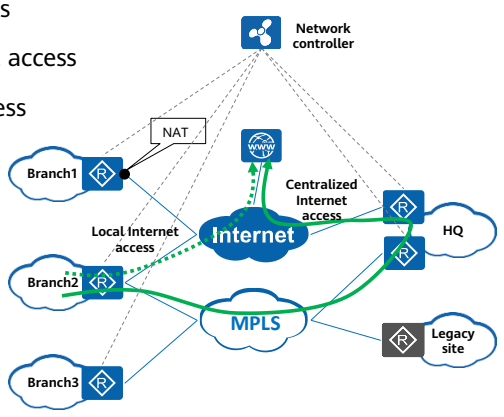


- **Solution description**
  - The hierarchical network model can be considered as a combination of single-layer network models. A WAN is divided into multiple areas, which are interconnected through a centralized backbone area. In this way, a large number of sites can communicate with each other across areas.
  - For example, the network of a multinational enterprise is divided into three areas (China, Europe, and America) based on the enterprise's management structure. Each area uses a single-layer network topology (hub-spoke or full-mesh) with one or more sites as border sites. The border sites constitute the backbone area (level-1 backbone network), which implements interconnection of the areas. Border sites of an area are connected to both the level-2 area network and level-1 backbone network.
- **Application scenarios**
  - The hierarchical network model features a clear network structure and excellent scalability, and is therefore applicable to enterprises with a large number of sites or multinational enterprises with widely distributed sites.

Contents

Internet Access Solution

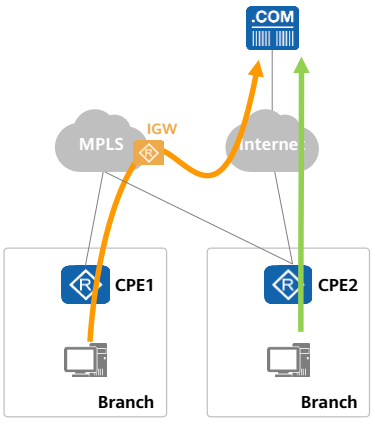
- The SD-WAN Solution involves the following Internet access scenarios:
  - Local Internet access
  - Centralized Internet access
  - Hybrid Internet access



HUAWEI

- Local Internet access
  - Internet access traffic of a site is directly routed out from local CPEs.
- Centralized Internet access
  - Internet access traffic of all sites is diverted to the centralized Internet access site and then to the Internet.
- Hybrid Internet access
  - Local Internet access + centralized Internet access for all traffic: By default, all Internet access traffic is routed out through the local Internet access interface. If the local Internet access interface is faulty, Internet access traffic is diverted to the centralized Internet access site and then to the Internet.
  - Centralized Internet access + local Internet access (for specified traffic): By default, Internet access traffic is routed out through the centralized Internet access site. Traffic of some specified services (for example, Office365) is directly routed out through local WAN links.
- Note: When users in multiple departments access the Internet at the same time, ensure that the users' IP addresses in different departments do not overlap.

Local Internet Access



Application scenarios

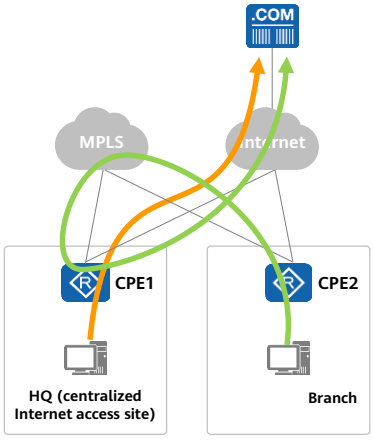
- Local Internet access is applicable to small-scale enterprises or scenarios where Internet access traffic does not require centralized security control and links for accessing the Internet are available on the WAN side.

Solution description

- Internet access traffic of a site is routed out from the local Internet link.
- Local Internet access policies can be configured on a per-department or per-site basis.
- Local Internet access can be implemented in either of the following modes based on traffic classification:
  - All Internet access traffic is routed out from the local site.
  - Only Internet access traffic of specified applications is routed out from the local site.
- Outbound interfaces must be configured for local Internet access. A maximum of three outbound interfaces can be configured. If multiple outbound interfaces are configured, Internet access link backup is implemented based on the priorities of the outbound interfaces.
- In local Internet access mode, NAT is provided. You can determine whether to enable the NAT function based on the outbound interface. Currently, NAT in Easy IP mode is provided. That is, the IP address of the outbound interface is used as the post-NAT public IP address.

- Internet access traffic cannot be load balanced among multiple links. Only priority-based link backup is supported.
- Local Internet access for specified applications must be enabled together with centralized Internet access, and is implemented through policy-based routing (PBR).
- When local Internet access is enabled, the default routes on the underlay WAN need to be configured separately. The default routes can be static routes (for Internet access through Internet interfaces) or BGP/OSPF routes (for Internet access through MPLS network interfaces).

Centralized Internet Access



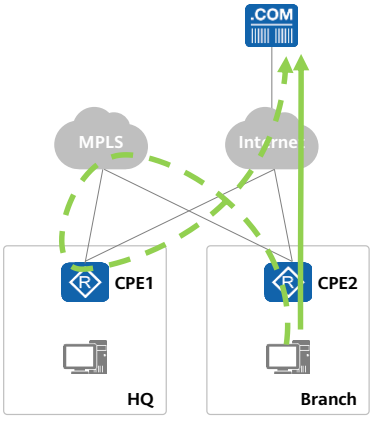
Application scenarios

- Centralized Internet access is applicable to scenarios where sites do not have links for accessing the Internet or Internet access traffic requires centralized security control. In this mode, a centralized Internet access gateway is configured. Traffic from other sites is forwarded to the centralized Internet gateway through the overlay network and then to the Internet.

Solution description

- Internet access traffic of all sites of a tenant is routed out through a centralized Internet access site.
- The centralized Internet access site can use either of the following methods for Internet access:
  - The site has an Internet egress on the LAN side, through which all Internet access traffic is routed out. In this case, you need to configure default routes on the LAN side, or configure a dynamic routing protocol so that the default routes can be learned from the LAN side.
  - The site can also access the Internet through WAN-side interfaces, through which all Internet access traffic is routed out. (Note that local Internet access must be enabled for the centralized Internet access site.)

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



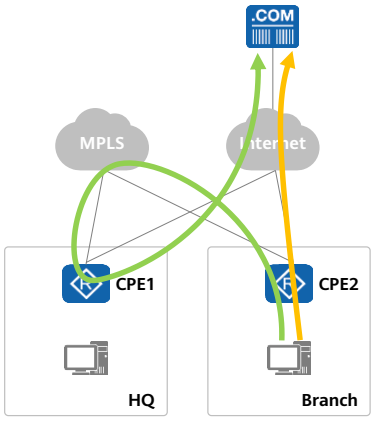
Application scenarios

- This Internet access mode is applicable to small-scale enterprises or scenarios where Internet access traffic does not require centralized security control, sites have WAN-side links for accessing the Internet, and high Internet access reliability is required.

Solution description

- Both local Internet access and centralized Internet access are enabled on the SD-WAN network.
- All Internet access traffic of sites is routed out through local WAN-side links.
- If a site has one local Internet access path and one centralized Internet access path, the default priority of the local Internet access path is higher.
- By default, all Internet access traffic is routed out through the local Internet access interface. If the local Internet access interface is faulty, Internet access traffic is diverted to the centralized Internet gateway and then to the Internet.

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



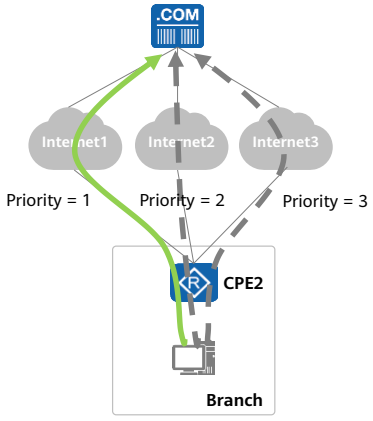
Application scenarios

- This Internet access mode is applicable to the scenarios where a site has links for accessing the Internet, most Internet access traffic requires centralized security control, and Internet access traffic of some applications can be routed out from the local site according to SLA requirements.

Solution description

- Both local Internet access and centralized Internet access are enabled on the SD-WAN network.
- At the centralized Internet access site, local Internet access is enabled and all Internet access traffic is routed out from this site.
- At other sites, local Internet access is enabled and Internet access traffic of specified applications is routed out from the local site.
- If a site has one local Internet access path and one centralized Internet access path, the default priority of the local Internet access path is higher.
- By default, Internet access traffic is routed out through the centralized Internet access site. Internet access traffic of specified experience-sensitive applications is routed out through local WAN-side links.

Internet Access Link Reliability



- **Backup link solution design**
  - Site-to-Internet access is configured by site. A maximum of three WAN links can be configured for each site as the site-to-Internet links.
  - Each WAN link has a unique priority when it functions as a site-to-Internet link. That is, only one WAN link can function as the active site-to-Internet link at a time.
  - If the WAN link with the highest priority is faulty, site-to-Internet traffic is automatically switched to the link with the second highest priority.
  - The backup link and hybrid Internet access can be enabled at the same time.

Quiz

1. (Multiple choices) Which of the following parameters are required for establishing a control tunnel in the SD-WAN Solution?
   A. TNP
   B. IPsec SA
   C. Service route
   D. Loopback interface of the CPE or RR

- 1. ABCD

Summary

- SD-WAN flexible networking involves three steps:
  - Management channel establishment, control channel establishment, and data channel establishment
- The overlay topology can be controlled by controlling the next hop of service routes on the RR.
- If a NAT device is deployed on a network, CPEs can use STUN technology to establish data channels with each other traversing the NAT device.
- SD-WAN supports both single-CPE and dual-CPE networking modes.
- SD-WAN supports four topology types:
  - Hub-spoke, full-mesh, partial-mesh, and hierarchical topologies
- SD-WAN supports three Internet access modes:
  - Local Internet access, centralized Internet access, and hybrid Internet access

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.

SD-WAN Application Experience

HUAWEI

Foreword

- Enterprises generally deploy diversified types of applications, which have different requirements on bandwidth and link quality. For example, real-time video conferencing has low tolerance for packet loss, delay, and jitter of links. If packet loss occurs on links, frame freezing and artifacts occur. The applications such as email and FTP-based file transfer are relatively insensitive to packet loss but require high bandwidth to ensure quick transfer.
- To meet diversified requirements of applications, traditional WANs have the following issues to resolve:
  - Applications of different values are carried on the same link.
  - When link quality deteriorates, dynamic traffic steering cannot be implemented.
  - No effective measure is available when the link quality deteriorates.
- To resolve these issues, Huawei SD-WAN Solution provides experience assurance for enterprise applications.
- This course describes SD-WAN application experience assurance, involving application-based traffic steering, hierarchical quality of service (HQoS), and WAN optimization.

- This course is based on Huawei SD-WAN Solution.

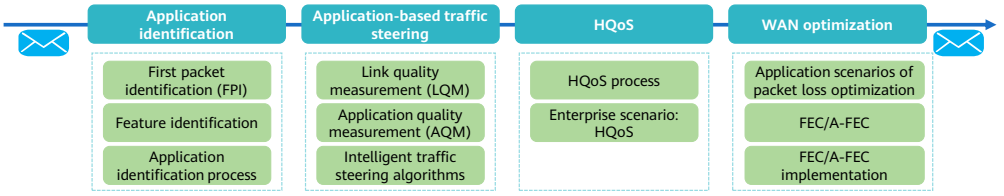
Objectives

- On completion of this course, you will be able to:
  - Describe the overall application experience implementation process.
  - Describe benefits of link quality measurement (LQM).
  - Describe application scenarios of HQoS.
  - Differentiate traffic steering policies.

Contents

Application Experience Solution Overview

- Huawei SD-WAN application experience solution consists of four parts:
  - Application identification: identifies the enterprise applications to which network traffic belongs based on traffic characteristics.
  - Application-based traffic steering: continuously monitors the status of multiple WAN links based on link quality requirements of enterprise applications, and selects the optimal link for traffic transmission.
  - HQoS: provides differentiated services for diversified enterprise applications based on QoS. In addition, HQoS provides service quality guarantee by service, department, and site.
  - WAN optimization: Packet loss compensation technology prevents frame freezing and artifacts of videos even if the link quality deteriorates and a large number of packets are lost. In addition, transmission and data optimization technologies are provided to improve data transmission efficiency.



FEC: forward error correction
A-FEC: adaptive FEC

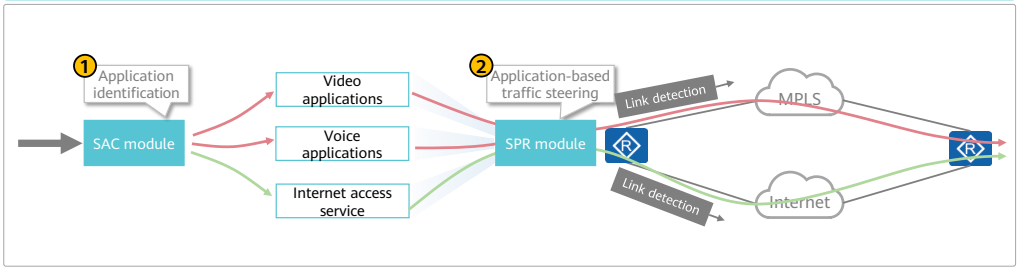
HUAWEI

Contents

Application Identification and Traffic Steering

- Huawei SD-WAN Solution uses Smart Application Control (SAC) technology to implement application identification and Smart Policy Routing (SPR) technology to implement application-based traffic steering.
  - SAC enables a device to identify applications and group traffic based on the service awareness (SA) and FPI signature databases.
  - SPR enables a device to determine the link quality based on link detection packets and then select the traffic forwarding path.

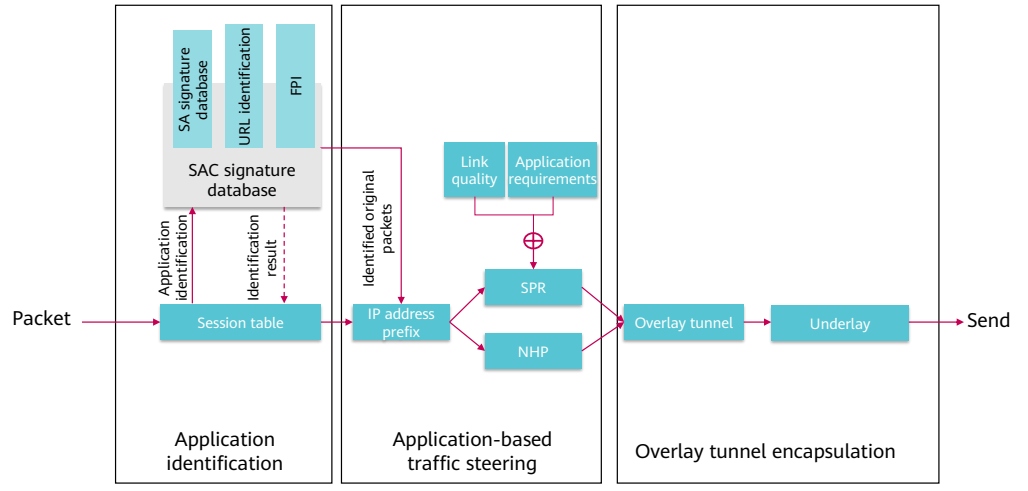
Application identification and traffic steering processes



HUAWEI

- For details about SAC and SRP, see *HA Technologies*.

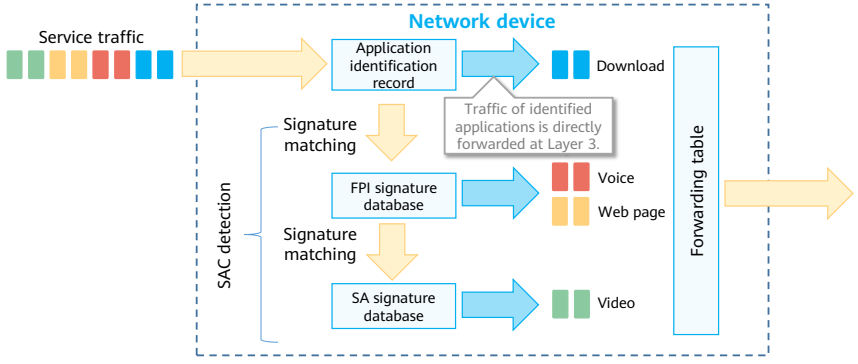
Intelligent Traffic Steering Overview: Key Technologies



- After receiving a service packet, a CPE processes the packet as follows:
  - Identifies the application.
    - If no session table exists, the CPE identifies the service type of the packet based on the SA signature database, URL identification, or FPI, performs application-based traffic steering, and sets up a session table.
    - If a session table exists, the CPE directly performs application-based traffic steering.
  - Performs application-based traffic steering.
    - If intelligent traffic steering is not configured, the CPE searches the routing table for a route to forward the packet.
    - If an intelligent traffic steering policy is configured, the CPE performs traffic steering based on the link or application quality.
  - Finally, the CPE performs overlay and underlay tunnel encapsulation for the packet, and sends it out.

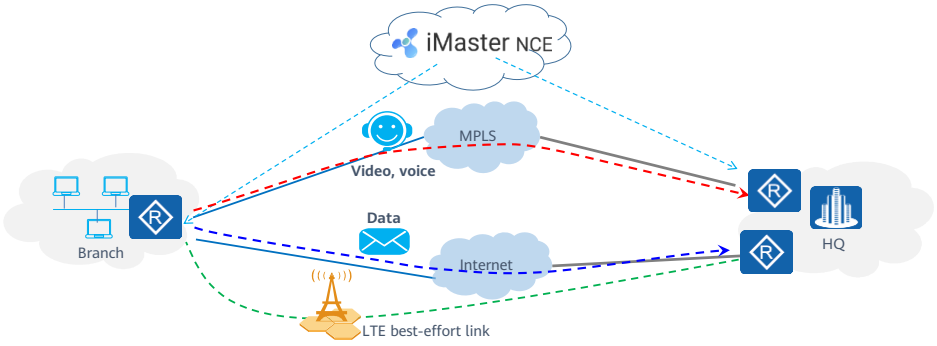
SAC Application Identification Process

- During SAC application identification, a network device checks whether an application is identified. If not, the network device matches the application against the FPI signature database and SA signature database in sequence.



- For details about application identification, see *HA Technologies*.

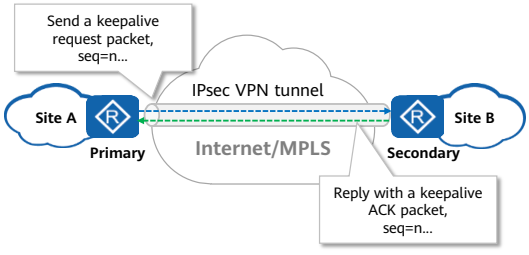
What Is Intelligent Traffic Steering?



- Intelligent traffic steering monitors link quality in real time based on application requirements and then selects the optimal link for traffic transmission.
- Different applications have different service level agreement (SLA) requirements. Intelligent traffic steering can be deployed to dynamically and automatically select network paths that meet applications' SLA requirements among multiple WAN links with different network quality, while ensuring WAN network efficiency.

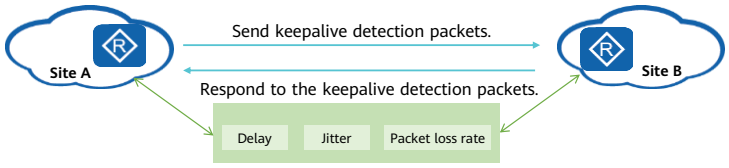
Contents

LQM



- LQM is the basis of application-based traffic steering and relies on the keepalive mechanism.
- Keepalive interaction is implemented between a primary device and a secondary device. The primary device automatically sends keepalive request packets based on a timer to control sessions. The secondary device only needs to respond to the keepalive request packets.
- keepalive interaction is unidirectional. After the primary device receives the instruction of enabling detection delivered by the controller, it constructs and encapsulates packets and the timer is triggered for sending packets.
- The primary and secondary devices are negotiated based on the following rules:
  - An RR does not function as the primary device.
  - An IWG does not function as the primary device.
  - The device with a smaller site ID functions as the primary device.

Keepalive Detection Mechanism



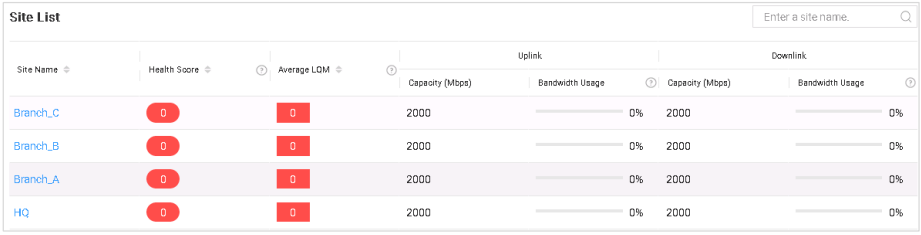
- **Keepalive detection action taken by the primary device**
  - By default, the primary device sends a keepalive detection packet every 1000 ms for six times in a detection period.
  - If the primary device does not receive any keepalive response packet from the secondary device within 6 seconds, keepalive detection fails.
  - The default keepalive detection parameters can be modified on iMaster NCE-WAN.
- **Keepalive detection action taken by the secondary device**
  - After receiving a keepalive detection packet, the secondary device responds immediately.
  - If the secondary device does not receive any keepalive detection packet within a detection period, keepalive detection fails.

Viewing LQM Information on iMaster NCE-WAN

- LQM
  - The LQM of a link is calculated based on the packet loss rate, delay, and jitter. The LQM is calculated as follows by default:

    LQM = F (Delay) x Weight (25%) + F (Jitter) x Weight (25%) + F (Packet loss rate) x Weight (50%)
  - The parameters in the LQM calculation formula can be modified.
  - The LQM of a site is the average of the LQM values of all links at the site.
  - The LQM between sites is the average of the LQM values of all links between the sites.

Viewing the Site Health on iMaster NCE-WAN

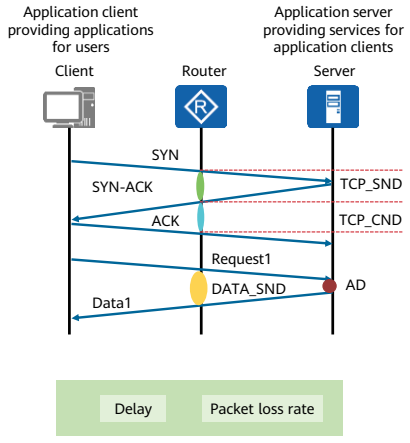


- Site health calculation method
  - Single-gateway site:
    - Site health score = Site LQM x 10 x 0.1 + Gateway health score x 0.9
  - Dual-gateway site:
    - Site health score = Site LQM x 10 x 0.1 + (Health score of gateway 1 + Health score of gateway 2) x 0.9/2

Note: The health score of a gateway is obtained by calculating the CPU usage and memory usage.

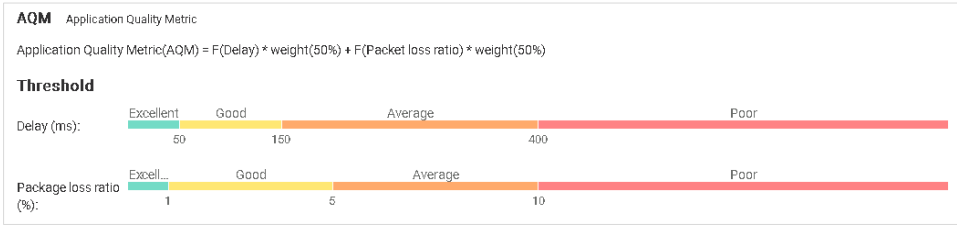
Contents

AQM Implementation



- When an application client sends a connection request to an application server through a router, the router forwards the request to the application server, identifies the SYN packet in the first handshake, obtains the source and destination IP addresses, creates a bidirectional flow table, and records the timestamp.
- After the application server receives the connection request, the router identifies the SYN-ACK packet in the second handshake by querying the flow table, records the timestamp, forwards the response packet sent by the application server to the application client, and calculates the network delay of the application server (TCP_SND).
- When the response packet reaches the application client, the router identifies the ACK packet in the third handshake by querying the flow table, records the timestamp, and calculates the network delay of the application client (TCP_CND).
- DATA_SND: response delay, that is, the difference between the time when the client sends a connection request and the time when the first response packet is received
- Packet loss rate = Number of retransmitted packets/Total number of sent packets

Viewing AQM Information on iMaster NCE-WAN



- AQM
  - The AQM is calculated based on the packet loss rate and delay of an application.
  - By default, AQM = F (Delay) x Weight (50%) + F (Packet loss rate) x Weight (50%)
  - The parameters in the AQM calculation formula can be modified.

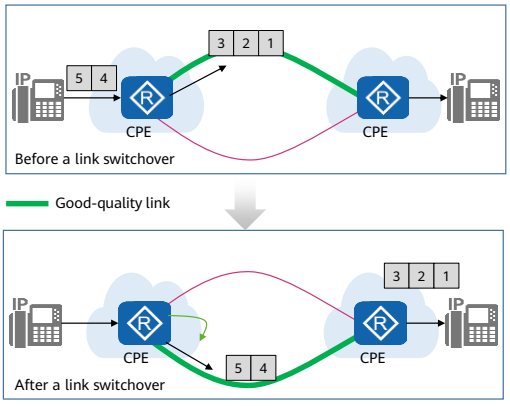
Contents

Intelligent Traffic Steering Algorithms

Category	Traffic Steering Algorithm	Description	Mandatory or Optional
Basic traffic steering	Link quality–based traffic steering	Traffic steering is implemented based on link quality. If the quality of one link does not meet SLA requirements (including the delay, jitter, and packet loss rate), a link switchover is triggered. If one of the switchover conditions (delay, jitter, and packet loss rate) exceeds the specified threshold, a link switchover is triggered. The system pre-defines the thresholds of the switchover conditions for four typical services: voice, real-time video, low-delay data, and large-volume data services. You can directly select a specific service type to implement traffic steering. You can also customize the thresholds of the switchover conditions based on service requirements.	Mandatory
Advanced traffic steering	Bandwidth-based traffic steering	You can set the upper and lower thresholds for the link bandwidth utilization to implement bandwidth-based traffic steering. • If the link bandwidth utilization is less than the lower threshold, the existing and subsequent traffic is forwarded over the current transport network. • If the link bandwidth utilization is between the lower and upper thresholds, the existing traffic is still forwarded over the current transport network, and subsequent traffic is switched to another transport network. • If the link bandwidth utilization is greater than the upper threshold, some existing traffic and all subsequent traffic are switched to another transport network.	Optional
	Load balancing–based traffic steering	When there are multiple links, per-flow load balancing can be performed for application traffic over these links.	Optional
	Application priority–based traffic steering	If multiple types of service packets are transmitted on the same link, traffic of high-priority applications is preferentially processed when congestion occurs, delivering optimal user experience.	Optional

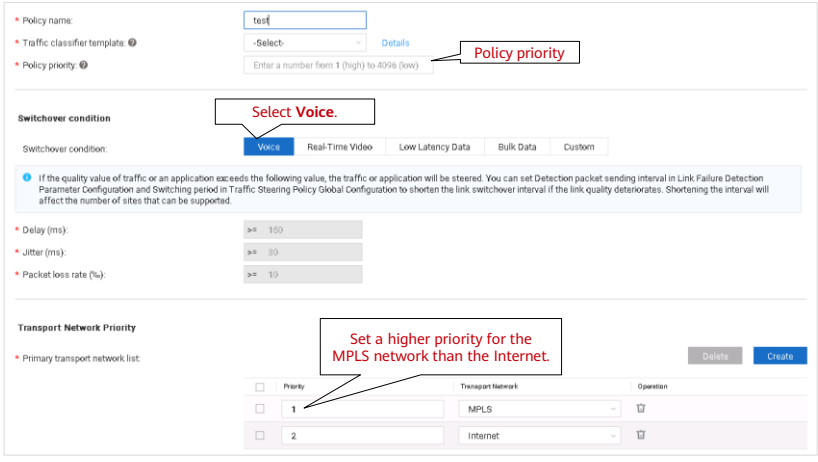
- Each site can be configured with primary and secondary transport networks.
  - Primary transport network: You can configure multiple primary transport networks for a site and specify priorities of them. A smaller value indicates a higher priority. You can set the same priority for multiple primary transport networks.
  - Secondary transport network: A secondary transport network provides escape links. Application traffic is switched to a secondary transport network only when all the primary transport networks are unavailable.

Intelligent Traffic Steering Algorithm: Link Quality-based Traffic Steering



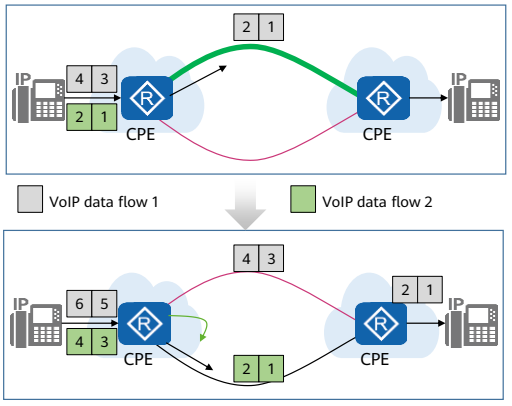
- Traffic steering is implemented based on link quality. If the quality of one link does not meet SLA requirements (including the delay, jitter, and packet loss rate), a link switchover is triggered.
- If one of the switchover conditions (delay, jitter, and packet loss rate) exceeds the specified threshold, a link switchover is triggered.
- Application scenario: Different applications have different SLA requirements for links. When a link does not meet the SLA requirements, traffic needs to be dynamically switched to another link that meets the SLA requirements.
- Voice services have low tolerance for the delay and packet loss rate. A CPE performs periodic link detection to monitor the delay, jitter, and packet loss rate of a link in real time, over which application traffic is transmitted. If one of the three switchover conditions exceeds the specified threshold, a link switchover is triggered.

intelligent traffic steering Algorithm: Link Quality-based Traffic Steering Configuration



Link Quality-based Traffic Steering | **Bandwidth-based Traffic Steering** | Load Balancing-based Traffic Steering | Application Priority-based Traffic Steering

Intelligent Traffic Steering Algorithm: Bandwidth-based Traffic Steering



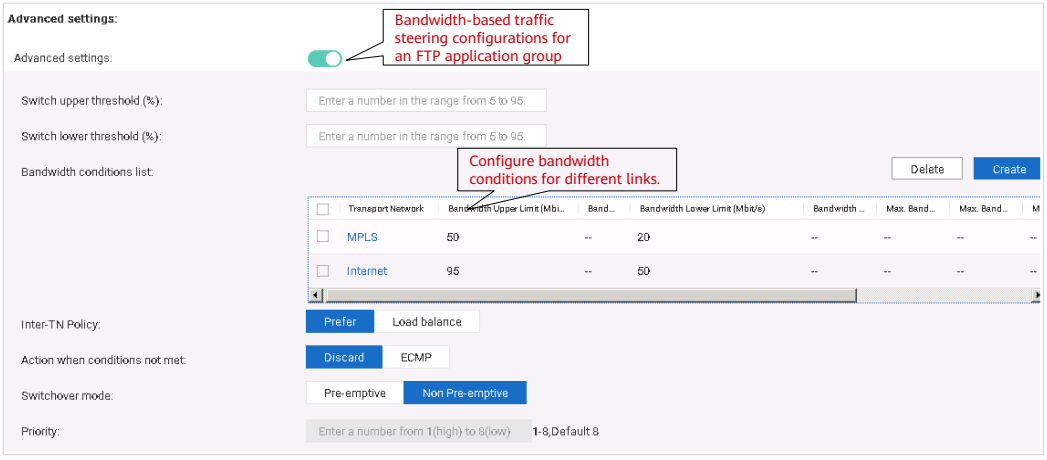
- As the bandwidth utilization of one link increases, VoIP data flow 2 is transmitted over another link.

- When the bandwidth of a link reaches a given threshold or its remaining bandwidth is lower than a given threshold, the link cannot be selected for forwarding subsequent traffic of certain applications, so as to prevent application or link quality deterioration.
- The link bandwidth of high-priority applications is preferentially guaranteed. To prevent high-priority applications from occupying all the link bandwidth, you can configure bandwidth limits for the applications. In this way, when the bandwidth utilization of an application reaches a given threshold, subsequent traffic of the application will be switched to another link.
- Application scenarios: Bandwidth resources are reserved for high-priority applications to ensure their user experience while high-value links are fully utilized.

HUAWEI

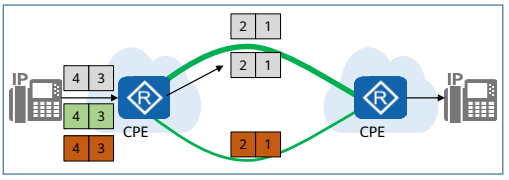
- MPLS links are high-value links, which are expected to be fully utilized. To ensure user experience of high-priority applications, it is recommended that MPLS links not be used to transmit traffic of low-priority applications when the link bandwidth utilization exceeds 70%.
- A link cannot be selected for bandwidth-demanding services (such as the backup service) when its remaining bandwidth is insufficient. In this case, bandwidth-based traffic steering can be configured.
- For example, an enterprise leases MPLS, Internet1, and Internet2 links. It expects to reserve certain bandwidth resources for high-value VoIP services to ensure user experience of VoIP services, while fully utilizing the MPLS link. In this case, bandwidth-based traffic steering can be configured. In addition, bandwidth conditions can be configured for low-value applications (such as email and FTP services) to select the MPLS link. For example, when the bandwidth utilization of the MPLS link exceeds 50%, new traffic is not transmitted over the MPLS link; when the bandwidth utilization of the MPLS link exceeds 70%, existing traffic on the MPLS link needs to dynamically switch to other links.

Intelligent Traffic Steering Algorithm: Bandwidth-based Traffic Steering Configuration



- Bandwidth-based traffic steering is applicable to FTP applications, with an MPLS link as the primary link and an Internet link as the secondary link. When the bandwidth utilization of the MPLS link exceeds 50%, a primary/secondary link switchover is triggered.
- Bandwidth-based traffic steering is not applicable to VoIP services. VoIP services use an MPLS link as the primary link and an Internet link as a secondary link.

Intelligent Traffic Steering Algorithm: Load Balancing-based Traffic Steering

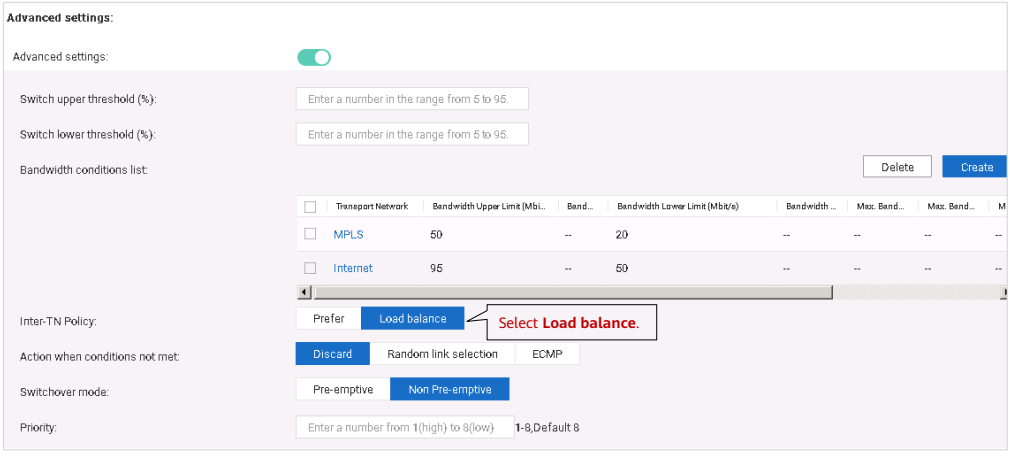


VoIP data flow 1 VoIP data flow 2 VoIP data flow 3

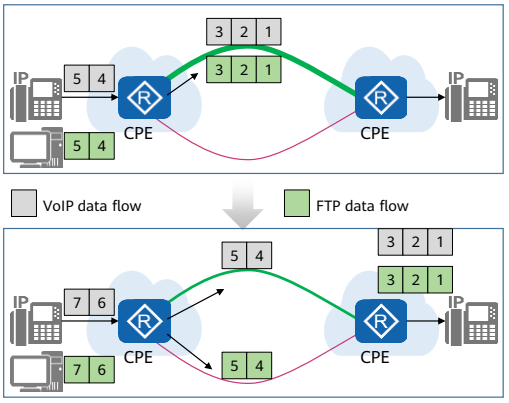
- As shown in the figure above, an enterprise purchases two MPLS links (one with the 100 Mbit/s rate and the other with the 50 Mbit/s rate) from different carriers, which can be added to the primary link group of the VoIP service. If the quality of both links meets the SLA requirements of the VoIP service, VoIP service flows can be carried over the two MPLS links in load balancing mode. Through real-time bandwidth utilization monitoring, the bandwidth utilization can be higher than 85%, ensuring full bandwidth utilization of the links.

- When load balancing–based traffic steering is deployed, per-flow load balancing can be performed among multiple links with the same priority. Bandwidth proportion–based load balancing can be performed among multiple links that have the same priority and meet SLA requirements.
- Application scenarios: When an enterprise has multiple links and expects to fully utilize the link bandwidth, load balancing–based traffic steering can be deployed.

Intelligent Traffic Steering Algorithm: Load Balancing-based Traffic Steering Configuration

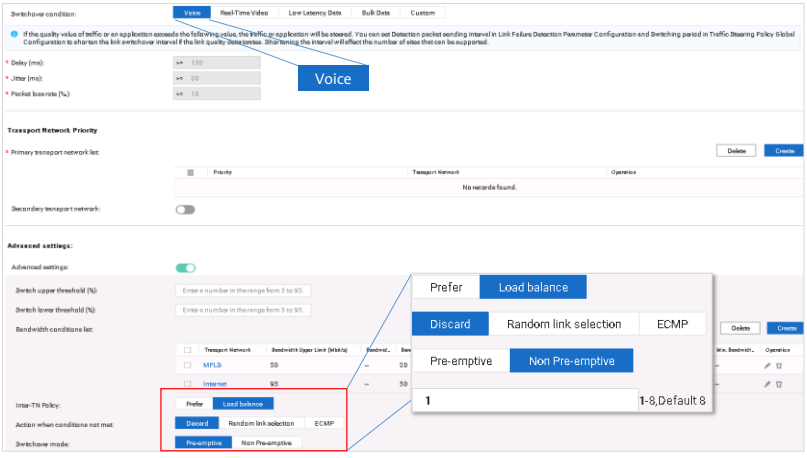


Intelligent Traffic Steering Algorithm: Application Priority-based Traffic Steering

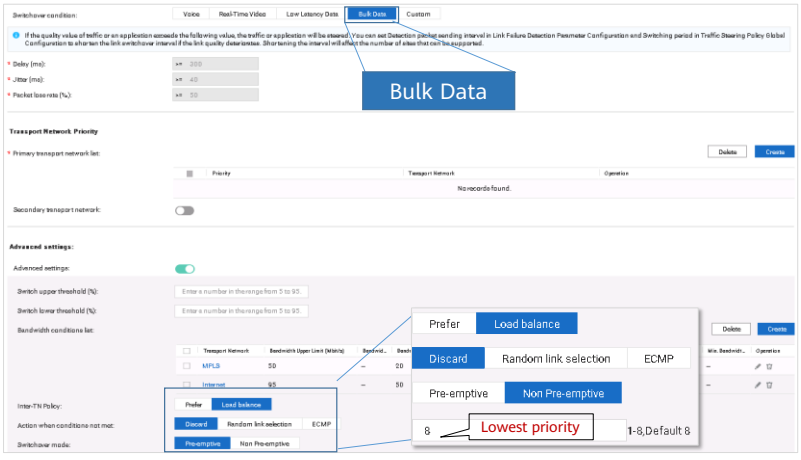


- For example, an enterprise leases an MPLS link and an Internet link from a carrier, and expects to fully utilize the MPLS link. To achieve this, the enterprise sets highest priorities for high-value applications and lower priorities for low-value applications (such as email and FTP applications). In addition, the enterprise sets the MPLS link as the primary link and the Internet link as the secondary link. In this way, when the MPLS link is congested, the CPE preferentially schedules traffic of low-priority applications to the Internet link based on application priorities to prevent quality deterioration of the MPLS link from affecting high-value applications.
- Application scenarios: High-value links (such as MPLS links) need to be preferentially selected and fully utilized. When link congestion occurs, traffic of low-priority applications needs to be preferentially scheduled to low-value links (such as Internet links).

Intelligent Traffic Steering Algorithm: Application Priority-based Traffic Steering Configuration (VoIP)

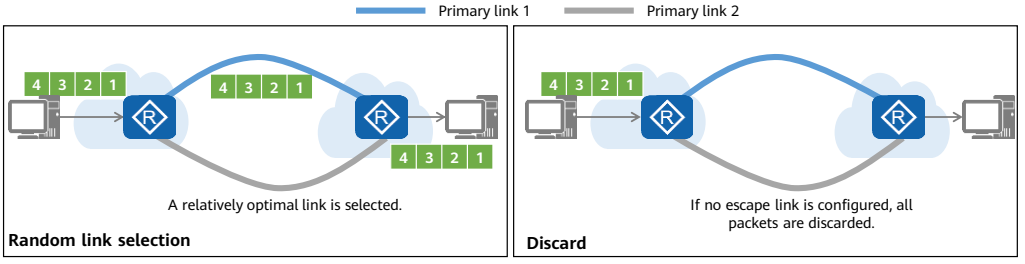


Intelligent Traffic Steering Algorithm: Application Priority-based Traffic Steering Configuration (FTP)



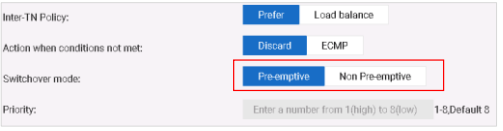
Intelligent Traffic Steering Implementation: Action Taken When SLA Requirements Are Not Met

- If no link on the primary transport network meets SLA requirements or the link bandwidth utilization exceeds a given threshold, traffic steering is performed in the following modes:
  - **Random link selection**: When **Inter-TN Policy** is set to **Prefer**, a relatively optimal link is selected among the links on the primary transport network.
  - **ECMP**: When **Inter-TN Policy** is set to **Load balance**, the CPE searches the routing table for a route to forward packets.
  - **Discard**: If an escape link is configured, packets are forwarded through the escape link. If no escape link is configured, all packets are discarded, which may cause service interruption.

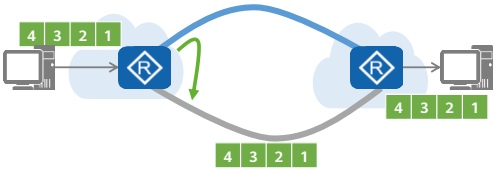


Intelligent Traffic Steering Implementation: Link Switchback

- The **Switchover mode** parameter specifies whether to allow traffic to switch back to the original link if the quality of the original link recovers. Link switchovers include the switchover between primary transport networks with different priorities and the switchover between primary and secondary transport networks.

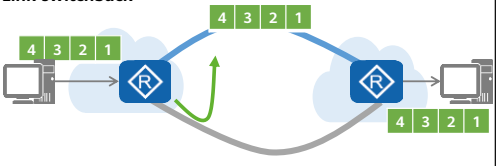


Link switchover



When the quality of primary link 1 does not meet requirements, traffic is switched to primary link 2.

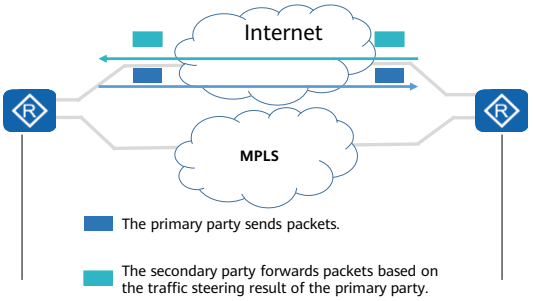
Link switchback



After the quality of primary link 1 recovers, traffic is switched back to primary link 1.

HUAWEI

Intelligent Traffic Steering Implementation: Scheduling Following



Scenario	Primary Party	Secondary Party
1	Selects the Internet link preferentially.	Selects the Internet link too.
2	Switches traffic to the MPLS link when detecting that the inbound traffic rate of the Internet link exceeds a given threshold.	Switches traffic to the MPLS link too.

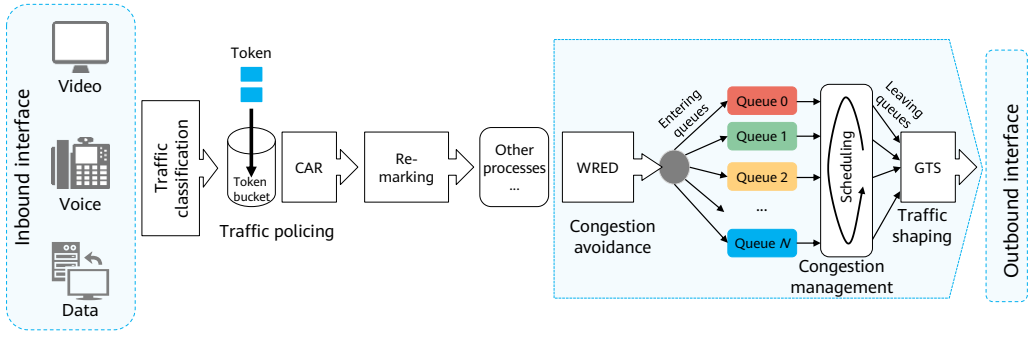
- When incoming and outgoing paths are inconsistent, the inbound link may become congested. The scheduling following function solves this problem. After the primary and secondary parties are determined based on certain rules and the primary party performs traffic steering, the secondary party follows the traffic steering result of the primary party.
- The primary party is determined as follows:
  - Between dual-gateway sites: The site with a smaller ID is selected as the primary party.
  - Between a single-gateway site and a dual-gateway site: The single-gateway site is selected as the primary party.
  - Between single-gateway sites: The site with a smaller WAN uplink bandwidth is selected as the primary party. If the sites have the same WAN uplink bandwidth, the site with a smaller ID is selected as the primary party.
  - Between a tenant site and an IWG (POP): The tenant site is selected as the primary party.

HUAWEI

- The scheduling following function is supported only when **Inter-TN Policy** is set to **Load balance**. Dynamic primary/secondary switchover is not supported.

Contents

QoS Data Processing Procedure (DiffServ Model)



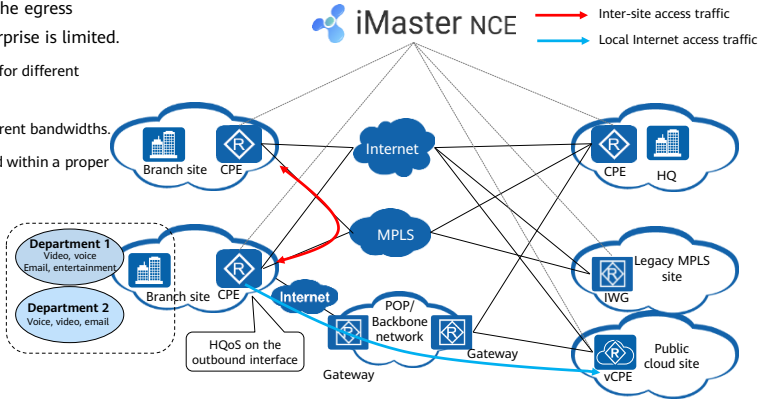
HUAWEI

- QoS provides the following functions:
  - Traffic classification and marking: helps to identify objects based on certain matching rules, which is the prerequisite for implementing differentiated services. Traffic classification and marking are typically applied to the inbound interface.
  - Token bucket: is used to check whether traffic meets packet forwarding conditions.
  - Traffic policing: monitors the volume of specific data traffic that arrives at a network device, and is typically applied to the inbound interface. When traffic exceeds a given amount, traffic limiting and punishment are performed to protect network resources and enterprise users' interests.
  - Congestion avoidance: Severe congestion causes damage to network resources. The congestion avoidance mechanism enables a device to monitor the utilization of network resources, and discards packets when network congestion becomes worse, thus to regulate network traffic and prevent network overload. Congestion avoidance is typically applied to the outbound interface.
  - Congestion management: must be implemented to solve the problem of resource competition. Packets are buffered in a queue and a scheduling algorithm is adopted to determine the packet forwarding sequence. Congestion management is typically applied to the outbound interface.

- Traffic shaping: is a traffic control measure that initiatively adjusts the output speed of traffic. Traffic shaping enables traffic to adapt to the network resources that can be provided by downstream devices to prevent packet loss and congestion. Traffic shaping is typically applied to the outbound interface.

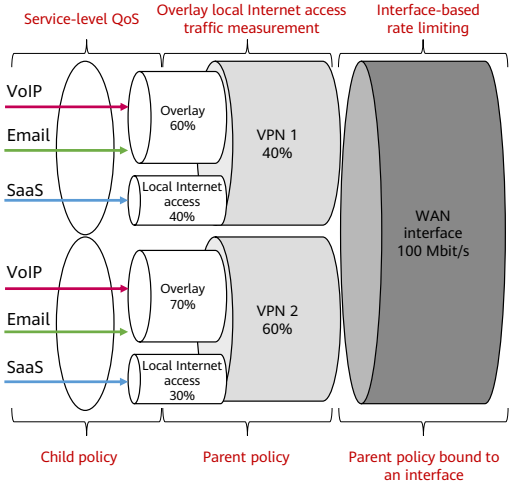
HQoS Service Scenario: Enterprise Scenario

- Objective: HQoS ensures that traffic of important applications, departments, and services is preferentially transmitted when the egress bandwidth purchased by an enterprise is limited.
  - Different priorities are guaranteed for different applications.
  - Different departments occupy different bandwidths.
  - The Internet access traffic is limited within a proper range.



HUAWEI

Enterprise Scenario: HQoS

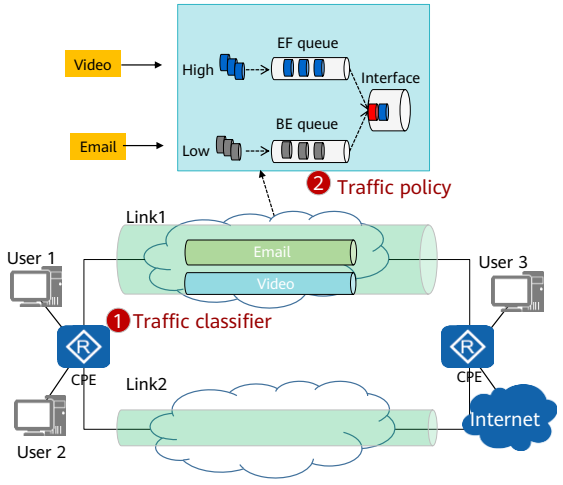


- Multiple applications in a department
  - Enterprise applications have different link requirements and different importance. The experience of important applications should be preferentially guaranteed when the egress link bandwidth is limited.
- Multiple departments of an enterprise
  - An enterprise usually has multiple departments of different importance, which require traffic isolation and differentiated bandwidths.
- Solution:
  - Different HQoS policies, such as queue scheduling, CAR, and traffic shaping, can be configured for each VPN based on applications.
  - The minimum guaranteed bandwidth proportion can be specified for each VPN. This prevents some VPNs from preempting the bandwidth of other VPNs in the case of network congestion.

HUAWEI

- An enterprise usually has multiple departments of different importance, which require traffic isolation and differentiated bandwidths.
  - A specified bandwidth quota is assigned to each department to meet its service requirements.
  - If some departments do not fully use their bandwidth quotas, idle bandwidth resources can be used by other departments with insufficient bandwidth.
  - The bandwidth for Internet access or legacy site access needs to be limited separately.

Enterprise Scenario: Intra-VPN HQoS Policy



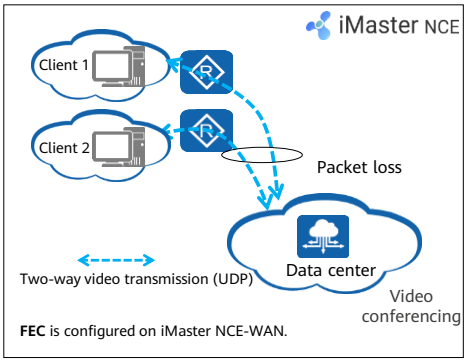
- Create traffic classifiers.
  - When traffic of VPN users (for example, voice, data, office application, and Internet access traffic) needs to be classified, a traffic classifier must be created for each type of traffic object.
- Configure actions in HQoS traffic policies.
  - Create HQoS policies, select a corresponding traffic classifier for each policy, and specify the HQoS action to be taken for the traffic object corresponding to each policy.

HUAWEI

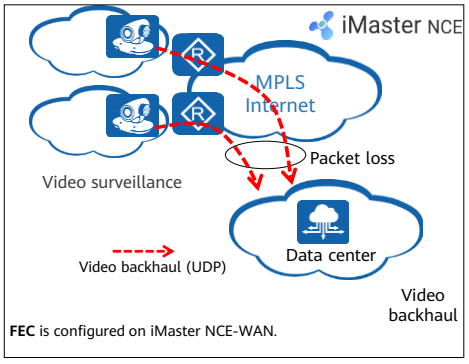
- Traffic can be classified based on one or a combination of the following:
  - Five-tuple
  - Applications or application groups
  - DSCP values
- Currently, the following traffic actions are supported:
  - Priority-based queue scheduling
  - Bandwidth limiting: CAR & shaping
  - DSCP re-marking

Contents

Application Scenario of Packet Loss Optimization



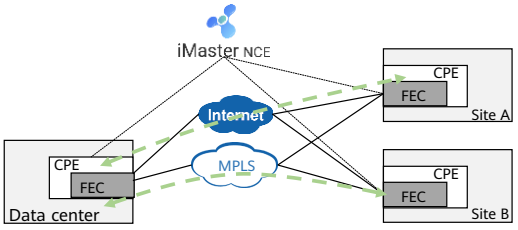
- Video conferencing
  - Characteristics: from multiple sites to one site, two-way
  - Key issue: Network packet loss affects video quality.
  - Deployment locations: egresses of the video terminal site and data center



- Video surveillance
  - Characteristics: from multiple sites to one site, one-way
  - Key issue: Network packet loss affects video quality.
  - Deployment locations: egresses of the camera area and data center

HUAWEI

FEC/A-FEC

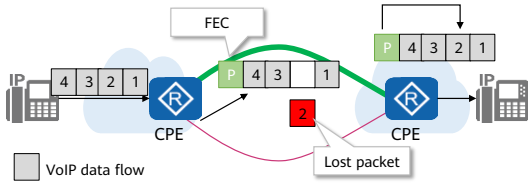


- Scenario:
  - An employee initiates a video conference at a branch site and connects to the data center. Because video conferencing is sensitive to packet loss and the traffic needs to go through the Internet that has poor quality, the packet loss and delay requirements cannot be met. As a result, the video quality deteriorates, and frame freezing and artifacts occur.
- Solutions:
  - FEC
  - A-FEC

HUAWEI

- FEC optimization technology is used to optimize packet loss by specifying data flows based on 5-tuple information through an agent. The agent obtains specified data flows, adds verification information to packets, and performs verification at the receive end. If a packet is lost or damaged on the network, it can be recovered based on the verification information.
- On the basis of FEC, A-FEC can automatically adjust the FEC redundancy rate to save bandwidth when the packet loss rate is low. When the packet loss rate increases sharply in a short period of time, the redundancy rate can be increased adaptively to offset the impact of packet loss on the network.
- Huawei FEC/A-FEC has the following advantages:
  - Different from the TCP-based retransmission mechanism, FEC does not require packet retransmission and has high real-time performance.
  - Compared with the simple exclusive OR algorithm, the RS algorithm provides the packet recovery capability upon burst packet loss.
  - A-FEC helps to dynamically adjust the redundancy rate to reduce bandwidth waste and prevent continuous packet loss.

FEC/A-FEC Implementation



- FEC implementation process:
  - A CPE at the transmit end receives and optimizes original packets. Meanwhile, it accumulates original packets and constructs a packet group.
  - The CPE at the transmit end performs FEC encoding on the original packets in the group, and generates FEC redundancy packets.
  - After receiving the original packets and redundancy packets, the CPE at the receive end detects the packet loss information, re-constructs a packet group, and recovers the lost packets.
  - The CPE at the receive end performs FEC decoding on the received original packets and redundancy packets in the same group to obtain the original packets.
  - The CPE at the receive end sends the original packets to the user.
- A-FEC implementation process:
  - The CPE at the receive end sends an FEC-ACK packet to the CPE at the transmit end. The packet contains real-time information about the packet loss rate and continuous packet loss on the network. The CPE at the transmit end adjusts the FEC redundancy rate according to the received FEC-ACK packet, alleviating or even eliminating the impact of packet loss on data transmission.

Quiz

1. (Multiple-answer question) Which of the following traffic steering policies are supported in Huawei SD-WAN Solution?
   A. Link quality–based traffic steering
   B. Bandwidth-based traffic steering
   C. Load balancing–based traffic steering
   D. Application priority–based traffic steering
2. (True or False) The scheduling following function is used to solve the problem of inconsistency between incoming and outgoing paths. The secondary party performs traffic routing, and the primary party follows the traffic steering result of the secondary party.

- 1. ABCD
- 2. False

Summary

- Enterprises usually deploy multiple egress links to ensure network reliability. The SD-WAN intelligent traffic steering solution improves link utilization, properly utilizes each link, and preferentially ensures experience of high-value services.
- A network device identifies the service type of traffic through FPI or SA, and applies the traffic steering policy matching the service type. Each service type has different link quality requirements. When the quality of a link cannot meet requirements, traffic is automatically switched to the secondary link.
- Huawei SD-WAN Solution supports intelligent traffic steering based on the link quality, bandwidth, load balancing, and application priority. You can select a traffic steering policy based on actual situations. HQoS and WAN optimization can be deployed to improve user experience upon link congestion.

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.

SD-WAN Security

HUAWEI

Foreword

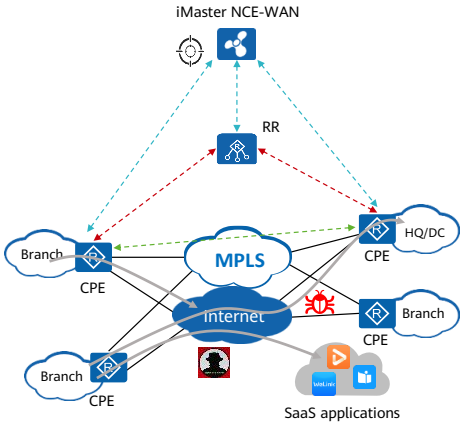
- On a conventional enterprise WAN, enterprise branches connect to the headquarters, for example, through MPLS private lines, to access the headquarters or data center for service operations or Internet access. Security policies can be implemented at the headquarters, for example, deploying firewalls and other security devices, to manage and control branch-to-Internet access behaviors. The relatively closed network architecture of conventional enterprise WANs ensures security to some extent.
- With the emergence of SD-WAN, the conventional closed architecture of enterprise WANs transforms to an open architecture, which enlarges the attack surface and brings new security challenges such as unauthorized access, data leakage, and network attacks.
- This course describes how to cope with these security challenges to meet the confidentiality, integrity, availability, and traceability requirements of SD-WAN networks and provide a secure, reliable, and stable service running environment.

Objectives

- On completion of this course, you will be able to:
  - Describe the basic concepts and design principles of system security and service security in the SD-WAN Solution.
  - Understand the security protocols and mechanisms used for communication between components of the SD-WAN Solution.
  - Describe the basic principles and application scenarios of service security functions provided by the SD-WAN Solution.

Contents

SD-WAN Security Risks

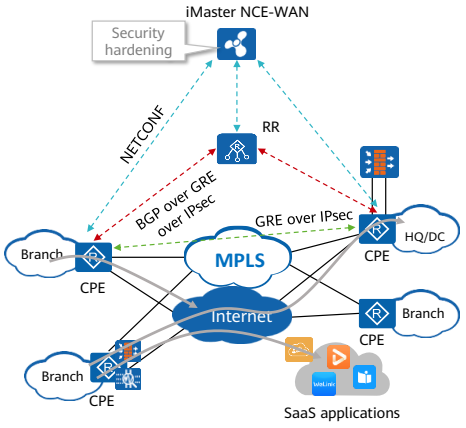


- The SD-WAN Solution is built on a public network, and therefore its components are vulnerable to attacks and the communication between components faces security risks. In addition, user services carried by the SD-WAN Solution, such as site-to-site, site-to-Internet, and site-to-SaaS application access services, are threatened by various security risks. The SD-WAN Solution faces the following security challenges:
  - Identity spoofing
  - Data leakage
  - Network attacks
  - Service security requirements

HUAWEI

- Identity spoofing: A component may be spoofed when it is registered and goes online. In this case, an identity authentication mechanism is required to ensure secure access of components and prevent identity spoofing.
- Data leakage: Inter-component communication data and user service data are transmitted over a public network such as the Internet. Such data may be intercepted or tampered with, which affects system running.
- Network attacks: Components provide interfaces for external interaction, and therefore are vulnerable to various attacks and intrusions such as flood attacks and application-layer attacks, which affect system availability.
- Service security requirements: User services carried by the SD-WAN Solution have security requirements, such as filtering specific packets, blocking intrusion behaviors, and limiting access to URLs.

SD-WAN Security Solution



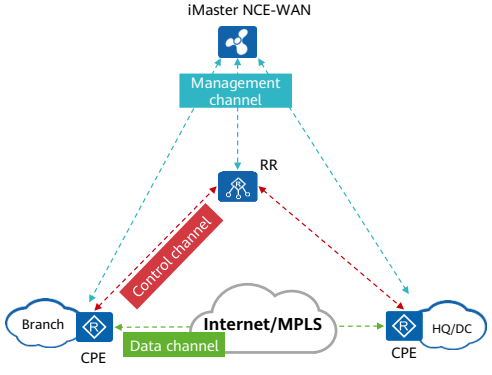
- To help enterprises better cope with the security challenges facing the SD-WAN Solution and take measures accordingly, Huawei defines security of the SD-WAN Solution from two aspects:
  - System security: refers to the security of the SD-WAN Solution itself. Huawei SD-WAN Solution provides system security to ensure secure and reliable system running.
  - Service security: refers to the security of services carried by the SD-WAN Solution. Huawei SD-WAN Solution provides security protection measures for enterprises to flexibly select based on their actual requirements to ensure secure and reliable running of user services.

HUAWEI

- System security covers the following aspects:
  - Inter-component communication security
  - Component security
- Service security covers the following aspects:
  - Site-to-site access security
  - Site-to-Internet access security
  - Site-to-SaaS application access security
- Huawei SD-WAN Solution provides the following service security functions:
  - IPsec
  - Firewall
  - IPS
  - URL filtering
  - Advanced security VAS
  - Third-party cloud security

Contents

Inter-Component Communication Security

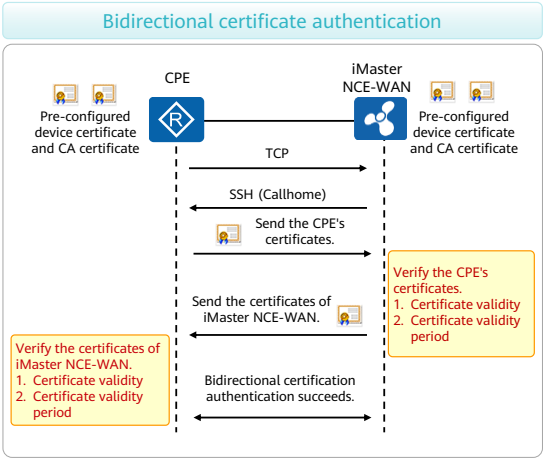


- Key components of the SD-WAN Solution include iMaster NCE-WAN, route reflectors (RRs), and customer-premises equipment (CPEs). The components use secure communication protocols to establish management, control, and data channels with each other to ensure communication security.
- Management channel
  - Management channels are established between CPEs or RRs and iMaster NCE-WAN.
  - The management channel security mechanisms include identity authentication and secure data transmission.
- Control channel
  - Control channels are established between CPEs and RRs.
  - The control channel security mechanisms include identity authentication and secure data transmission.
- Data channel
  - Data channels are established between CPEs.
  - The data channel security mechanism ensures secure transmission of service data.

Management Channel Security (1)

- After a CPE is powered on and registers with iMaster NCE-WAN, iMaster NCE-WAN configures and delivers services to the CPE. In this process, the CPE and iMaster NCE-WAN may be spoofed because they reside on a public network.
- Bidirectional certificate authentication is used to enable the CPE and iMaster NCE-WAN to verify the certificates of each other to ensure their validity.

- Certificate Authority (CA): issues certificates.
- CA certificate: refers to a certificate issued by a CA.
- Device certificate: refers to a certificate issued by a CA to a device. Here, devices refer to the CPE and iMaster NCE-WAN.

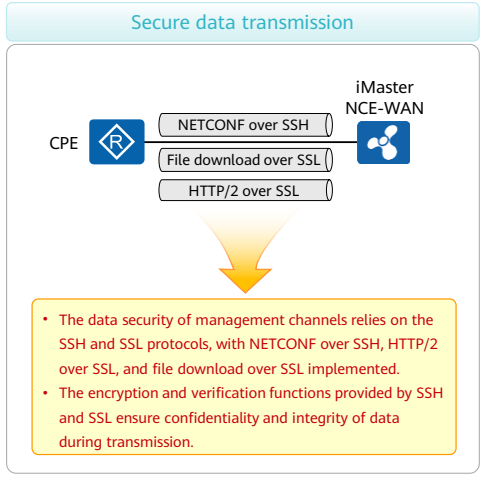


HUAWEI

- iMaster NCE-WAN verifies the validity and validity period of a CPE's certificates.
- A CPE verifies the validity and validity period of the certificates of iMaster NCE-WAN.
- Callhome: proactive device registration function

Management Channel Security (2)

- Because CPEs and iMaster NCE-WAN reside on a public network, data transmitted between them may be leaked or tampered with.
- The Network Configuration Protocol (NETCONF) provides a mechanism for managing network devices, which can be used to add, delete, modify, and obtain network device data. NETCONF uses the Secure Shell (SSH) protocol as the secure transport layer to provide a communication path for interaction between the client and server.
- Secure Sockets Layer (SSL) is a security protocol that provides security and data integrity for network communication. It can be used to encrypt network connections between the transport layer and application layer.

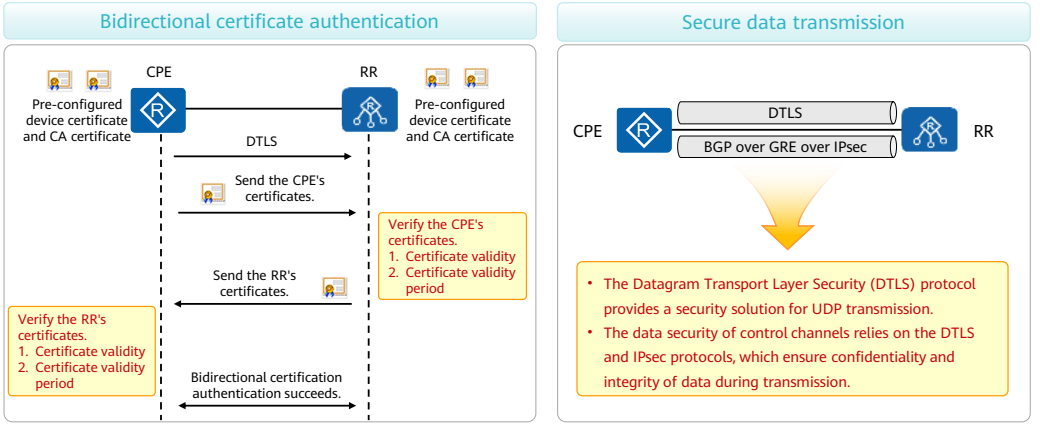


HUAWEI

- Management channels contain NETCONF connections, file download connections, as well as HTTP/2 connections for performance data reporting.
- SSH provides secure remote login and other security network services over insecure networks.

Control Channel Security

- Because CPEs and RRs reside on a public network, control channels established between them may be spoofed. After control channels are established, data may be intercepted or tampered with during transmission in the channels.

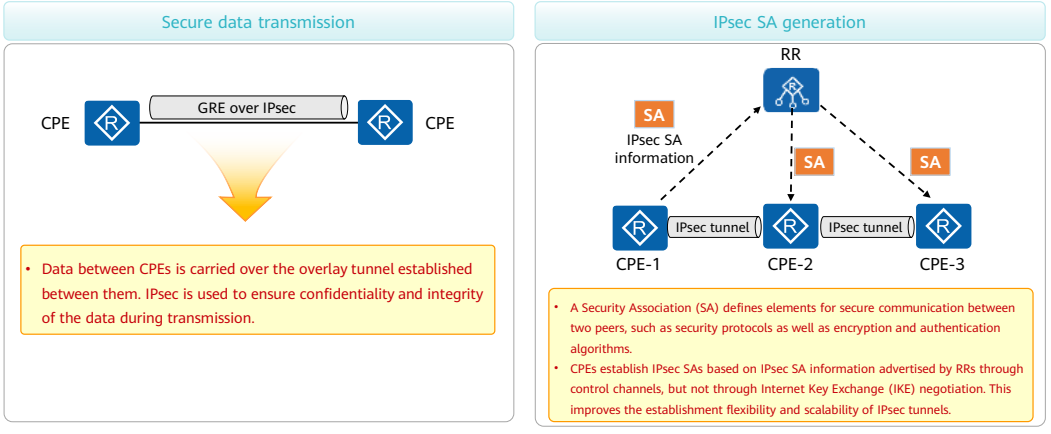


HUAWEI

- Control channels contain DTLS connections and BGP connections that are derived from DTLS connections. A CPE exchanges IPsec tunnel information with an RR through a DTLS connection and then establishes a BGP connection with the RR under the protection of the IPsec encryption mechanism.
- The RR verifies the validity and validity period of the CPE's certificates.
- The CPE verifies the validity and validity period of the RR's certificates.

Data Channel Security

- Inter-site data is transmitted across public networks, and may be leaked or tampered with during transmission. In this case, encryption is required to ensure security of data during transmission.



HUAWEI

- IPsec encryption can be enabled or disabled based on VPNs (departments).
- The IKE protocol provides the mechanisms of automatic key negotiation and IPsec SA establishment to simplify IPsec configuration and maintenance.

Contents

Component Security (1)

- Security of iMaster NCE-WAN
  - iMaster NCE-WAN is a key component of the SD-WAN Solution, and its security directly determines the reliability and availability of the entire network. iMaster NCE-WAN must be deployed in a firewall-protected area and provide comprehensive security functions to mitigate security risks. The security functions include but are not limited to those in the following tables.

Category		Measures
Authentication and permission control	Identity authentication	• Local and remote authentication (LDAP) • Two-factor authentication (user name/password + SMS verification code)
	Permission control	• Role-based permission control • Tenant- and domain-based permission control
Data protection	Key management	• Secure encryption algorithms • Hierarchical key management
	Data storage	• Data access control • Encrypted data storage
	Data transmission	• Secure communication protocols • Encrypted data transmission
Security detection and response	Attack detection	Attack detection from multiple dimensions, including ports, web pages, and operating systems
	Intrusion prevention	Defense against various intrusion behaviors
	Anti-DoS/DDoS	Defense against common traffic attacks

Category		Measures
Privacy protection		Strict access permission control
Security management	Security audit	Comprehensive logging and event recording functions
	Secure upgrade and patch installation	Administrator authentication (Only authenticated administrators can perform upgrade and patch installation operations.)
System protection		Integrity protection and signature verification for software and patches
Security deployment		Zone planning, hierarchical deployment, and firewall deployment for isolation

- The Lightweight Directory Access Protocol (LDAP) is used for accessing online directory services based on TCP/IP.

Component Security (2)

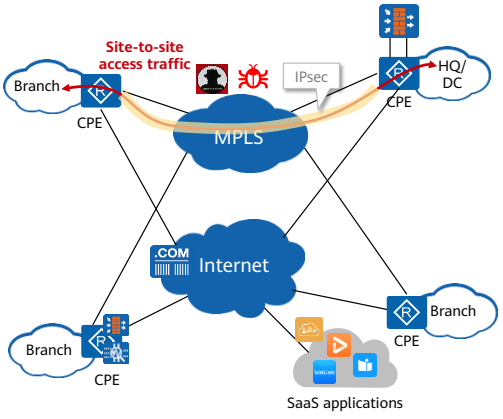
- Security of CPEs and RRs
  - The security of the physical and network environments where CPEs and RRs reside must be ensured, so that the CPEs and RRs can run securely, reliably, and stably. The system architecture of CPEs and RRs complies with the three-layer and three-plane security isolation mechanism defined by ITU-T X.805, in which the management, control, and forwarding planes are isolated. This ensures that attacks on any of the planes do not affect other planes. In addition, CPEs and RRs themselves must have multiple security protection capabilities, including but not limited to those in the following table.

Category	Measures
Physical security	The service ports, serial ports, and services that are not in use are disabled to prevent attacks to devices through them.
Data security	Sensitive data, such as service data, user names, and passwords, are encrypted to prevent leakage. Data access permissions are controlled to prevent unauthorized access to data.
Authentication	Identity authentication and permission control are performed on user login behaviors. Local authentication and remote authentication (HWTACACS) are supported. User names and passwords are strictly protected, password complexity check is performed, and the anti-brute force cracking mechanism is implemented.
Attack defense	CPEs and RRs can defend against various network attacks, such as IP flood attacks, ICMP flood attacks, malformed packet attacks, and packet fragment attacks.
Security audit	A log system records all the system configuration operations and the exceptions that occur during running of the system, facilitating post-event auditing.

- The ITU-T X.805 standard defines a security architecture for end-to-end communication systems. It defines three security layers (infrastructure, service, and application security layers) and three security planes (control, management, and forwarding planes). In addition, this security architecture protects each security plane at each security layer from eight security dimensions.
- Huawei Terminal Access Controller Access Control System (HWTACACS) is an enhancement of TACACS (defined in RFC 1492), and is a centralized information exchange protocol using the client/server architecture. It uses TCP for transmission, and the TCP port number is 49.
- The authentication, authorization, and accounting services provided by HWTACACS are independent of each other and can be implemented on different servers.

Contents

Site-to-Site Access Security

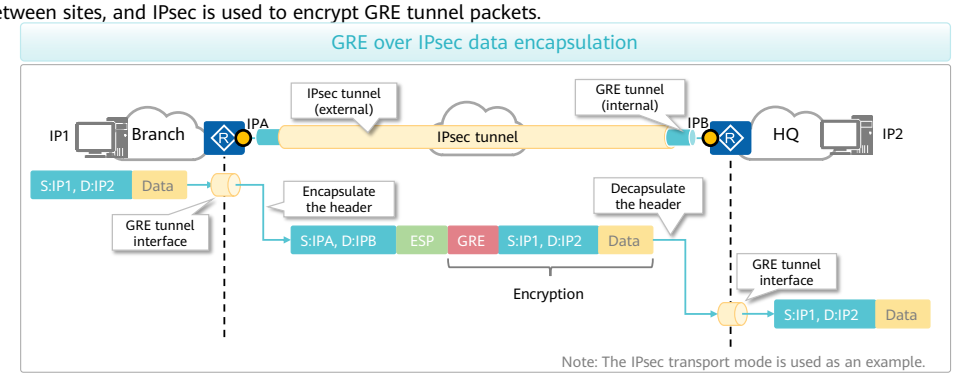


Service traffic between sites is transmitted over a public network (such as the Internet), and may be leaked or tampered with during transmission. In this case, IPsec is required to protect data. Site-to-site access security is also a type of data channel security.

HUAWEI

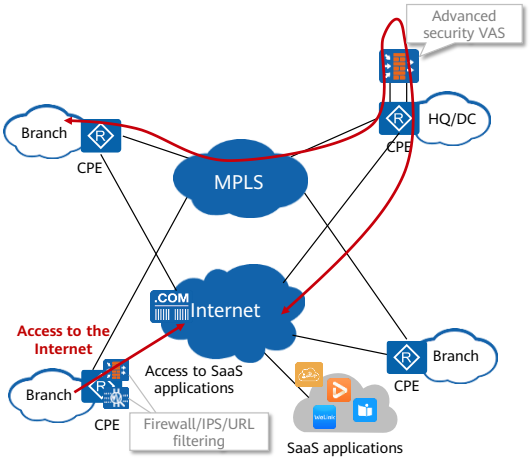
GRE over IPsec for Secure Data Transmission

- Site-to-site access traffic is first encapsulated by GRE. GRE is simple. However, data is transmitted over GRE tunnels in clear text and prone to interception.
- Typically, GRE is used together with IPsec on the live network. GRE is used to establish interconnection channels between sites, and IPsec is used to encrypt GRE tunnel packets.



Contents

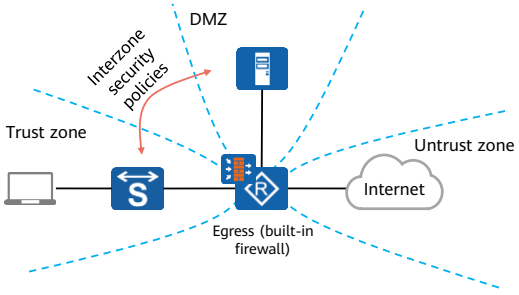
Site-to-Internet Access Security



Direct site-to-Internet access faces various security risks. In such access scenario, CPEs need to provide certain service security protection capabilities, such as firewall, intrusion prevention system (IPS), and URL filtering. In addition, Value-Added Service (VAS) functions can be deployed on SD-WAN networks to provide advanced security protection through physical firewalls deployed in off-path mode.

Basic Firewall Concepts: Zone

- Security zones (or zones) are defined on a firewall. A security zone is a collection of networks connected through one or more interfaces. The firewall considers that data flows within a single security zone are trustful and no security policy is required. The firewall checks data and carries out security policies only when the data flows from one zone to another.



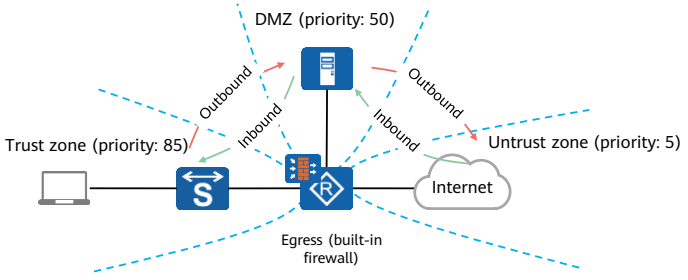
Zone	Default Security Priority
Untrust zone	5 (low security level)
Demilitarized zone (DMZ)	50 (medium security level)
Trust zone	85 (high security level)
Local zone	A local zone is a device itself, including interfaces on the device. The local zone of a device is the highest-level security zone with a security priority of 100.

- The names of default security zones on a firewall contain only lowercase letters and are case-sensitive. The security zones include:
  - Untrust zone: defines insecure networks such as the Internet.
  - DMZ: defines the zone where intranet servers reside. Intranet servers are frequently accessed by extranet devices but are not allowed to proactively access the extranet. Therefore, intranet servers face security risks, and need to be deployed in a security zone with a priority lower than the trust zone and higher than the untrust zone.
    - A DMZ is originally a military term, referring to a partially controlled area between a military control area and a public area. A DMZ configured on a firewall is logically and physically separated from intranets and extranets.
    - The servers such as WWW servers and FTP servers that provide network services for external devices are deployed in a DMZ. If these servers are deployed on an intranet, malicious users may exploit security vulnerabilities of some services to attack the intranet. If they are deployed on an extranet, their security cannot be ensured.
  - Trust zone: defines the zone where intranet terminals reside.
  - Local zone: defines a device itself, including the interfaces on the device. All packets constructed on and proactively sent from the device are considered to be sent from the local zone, and the packets to be responded and processed by the device (not only detected or forwarded) are considered to be received by the local zone. Local zone configurations cannot be modified. For example, interfaces cannot be added to the local zone.

- Due to the particularity of the local zone, a security policy needs to be configured to permit packet exchange between the local zone and the security zone of a peer in scenarios where a device is required to send and receive packets.

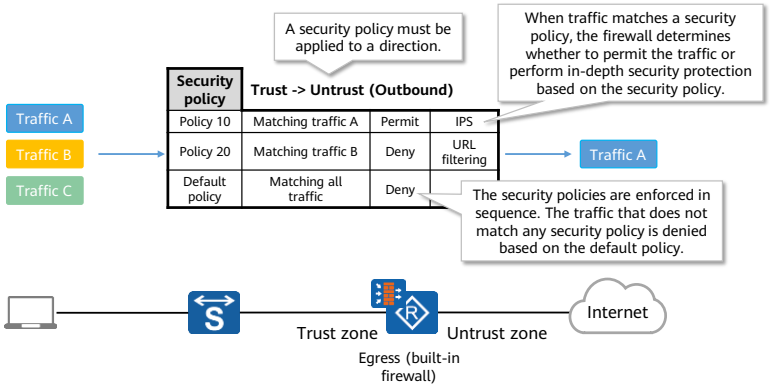
Basic Firewall Concepts: Interzone Direction

- The traffic passing through a firewall is directional. An interzone security policy takes effect only when it is applied to the correct direction.
- The direction of traffic on a firewall is determined by zone priorities.
  - Inbound direction: refers to the direction of traffic flowing from a low-priority zone to a high-priority zone.
  - Outbound direction: refers to the direction of traffic flowing from a high-priority zone to a low-priority zone.



Basic Firewall Concepts: Security Policy

- Security policies can be used to control traffic between security zones on a firewall.

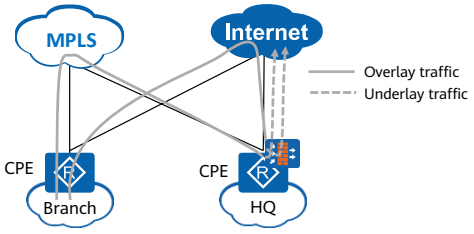


HUAWEI

- A firewall basically protects a network from being attacked by any untrusted network while permitting legitimate communication between two networks. Security policies are used to check data flows passing through a firewall. Only the data flows that match the security policies with the action of Permit are allowed to pass through the firewall.
- Security policies on a firewall can control the access permissions of intranet users to the extranet and control the access permissions between the subnets of different security levels on the intranet. In addition, security policies can control the access to a firewall itself, for example, restricting the IP addresses that can be used to log in to the firewall through Telnet and the web system and controlling communication between the NMS/NTP server and the firewall.
- Security policies define rules for processing data flows on a firewall. The firewall processes data flows according to the rules. Therefore, the core functions of security policies are to filter the traffic passing through the firewall according to the defined rules and determine the next operation performed on the filtered traffic based on keywords.
- Security policies on a firewall are a basic means for providing secure network access to the data flows passing through the firewall, and determine whether subsequent application data flows are processed. An NGFW analyzes traffic and retrieves traffic attributes, including the source security zone, destination security zone, source IP address, source region, destination IP address, destination region, user, service (source port number, destination port number, and protocol type), application, and time range.

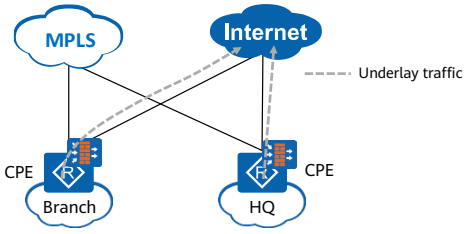
Application Scenarios of the Firewall Function

Centralized Internet access scenario



- Internet access traffic of all branch sites is diverted to the HQ site and then to the Internet. The firewall function needs to be enabled on the CPE at the HQ site to isolate the intranet and extranet.

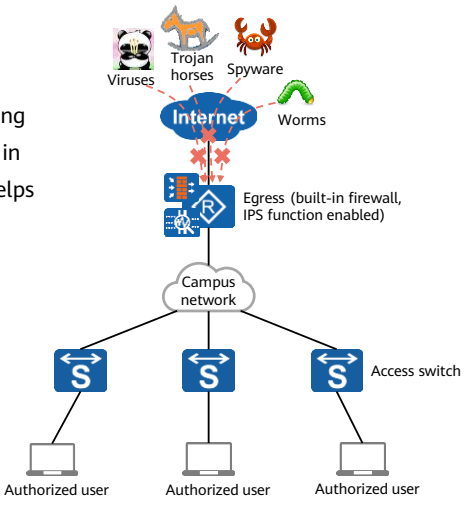
Local Internet access scenario



- Internet access traffic of the branch site and HQ site is directly transmitted to the Internet from the local CPEs. The firewall function needs to be enabled on the CPEs at the branch site and HQ site to isolate the intranet and extranet.

IPS Overview

- The IPS is a network security mechanism.
- It is used to detect intrusion behaviors (such as buffer overflow attacks, Trojan horses, and worms) by analyzing network traffic, and then block the intrusion behaviors in real time through certain response methods. The IPS helps to protect enterprise information systems and network architectures against intrusions.
- The IPS has the following advantages:
  - Real-time attack blocking
  - In-depth protection
  - All-round protection
  - Internal and external prevention
  - Constant update for up-to-date protection

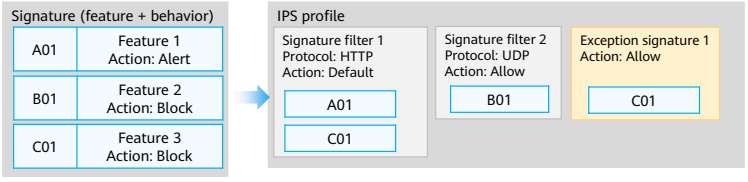


HUAWEI

- The IPS can detect and block network intrusions in real time. After detecting network intrusions, the IPS can automatically discard intrusion packets or block attack sources to fundamentally prevent attacks. The IPS has the following advantages:
  - Real-time attack blocking: When the IPS deployed on a network in in-path mode detects network intrusions, it can block the intrusions and attack traffic in real time, minimizing impacts of the intrusions.
  - In-depth protection: New attacks are hidden at the application layer of the TCP/IP protocol. The IPS can detect the content of application-layer packets, analyze and 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: The IPS provides protection measures against a variety of 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 leaks, remote file inclusion attacks, overflow attacks, code execution, DoS attacks, and scanning attacks, comprehensively protecting network security.
  - Internal and external protection: The IPS can protect enterprises from both external and internal attacks. The IPS can detect the traffic passing through and protect servers and clients.
  - Constant update for up-to-date protection: The IPS signature database is constantly updated to maintain the highest security level. You can periodically update the IPS signature database from the update center to ensure effective intrusion prevention.

Basic Concepts of IPS

- The IPS needs to identify intrusion traffic before controlling the intrusion traffic. Intrusion traffic identification and control are implemented by the following functional modules:
  - IPS signature database: defines features of various common intrusion behaviors and assigns a unique intrusion behavior ID for each kind of intrusion behavior feature.
  - IPS signature: describes the features of an intrusion behavior on the network and the action to be taken for the intrusion behavior. IPS signatures can be pre-defined or user-defined.
  - Signature filter: is a collection of signatures that meet specified filtering conditions. You can add multiple signatures to a signature filter and redefine the action to be taken for the traffic matching any signature in the signature filter.
  - Exception signature: Some signatures in a signature filter can be configured with actions different from that of the signature filter.
  - IPS profile: contains multiple signature filters and exception signatures. A CPE processes traffic based on an IPS profile.

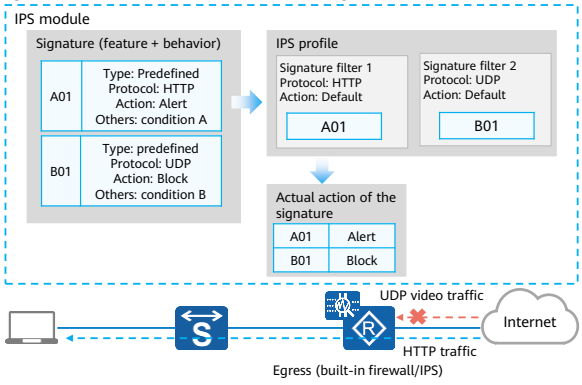


- By default, a CPE has multiple default IPS profiles for different application scenarios. The default IPS profiles can be viewed, copied, and referenced by security policies, but cannot be modified or deleted.
  - **strict**: It contains all signatures and the action is block. It is applicable to all protocols and all threat categories. This profile applies to scenarios where all packets that match signatures need to be blocked.
  - **web_server**: It contains all signatures and the default actions are used. It is applicable to the DNS, HTTP, and FTP protocols, and all threat categories. This profile applies to the scenarios where the CPE is deployed in front of a web server.
  - **file_server**: It contains all signatures and the default actions are used. It is applicable to the DNS, SMB, NetBIOS, NFS, SunRPC, MSRPC, file transfer, and Telnet protocols, and all threat categories. This profile applies to the scenarios where the CPE is deployed in front of a file server.
  - **dns_server**: It contains all signatures and the default actions are used. It is applicable to the DNS protocol and all threat categories. This profile applies to the scenarios where the CPE is deployed in front of a DNS server

- **mail_server**: It contains all signatures and the default actions are used. It is applicable to the DNS, IMAP4, SMTP, and POP3 protocols, and all threat categories. This profile applies to the scenarios where the CPE is deployed in front of a mail server.
- **inside_firewall**: It contains all signatures and the default actions are used. It is applicable to all protocols and all threat categories. This profile applies to the scenarios where the CPE is deployed behind a firewall.
- **dmz**: It contains all signatures and the default actions are used. It is applicable to all protocols except NetBIOS, NFS, SMB, Telnet, and TFTP, and all threat categories. This profile applies to the scenarios where the CPE is deployed in front of a DMZ.
- **outside_firewall**: It contains all signatures and the default actions are used. It is applicable to all protocols and all threats except Scanner. This profile applies to the scenarios where the CPE is deployed in front of a firewall.
- **ids**: It contains all signatures and the action is alert. It is applicable to all protocols and all threat categories. This profile applies to the scenarios where the CPE is deployed in off-path mode as an IDS device.
- **default**: It contains all signatures and the default actions are used. It is applicable to all protocols and all threat categories. This profile applies to the scenarios where the CPE is deployed in in-path mode as an IPS device.

IPS Implementation

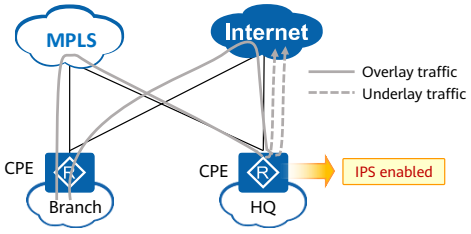
- When a data flow matches a security policy that contains an IPS profile, the CPE sends the data flow to the IPS module to match the signatures referenced by the IPS profile one by one.
- If the data flow matches a signature, the action defined for the signature, such as block or alert, will be taken for the data flow.



- When a data flow matches multiple signatures, the actual action for the data flow is as follows:
  - If the actions defined for all the matched signatures are alert, the action for the data flow is alert.
  - If the action defined for any of the matched signatures is block, the action for the data flow is block.
- When a data flow matches multiple signature filters, the action defined for the signature filter with the highest priority will be taken for the data flow.

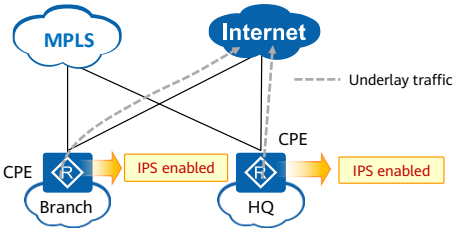
Application Scenarios of IPS

Centralized Internet access scenario



- Internet access traffic of all branch sites is diverted to the HQ site and then to the Internet. The IPS function needs to be enabled on the CPE at the HQ site to block various intrusions from the Internet.

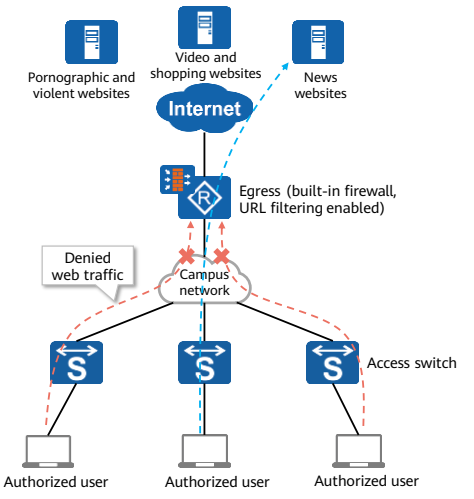
Local Internet access scenario



- Internet access traffic of the branch site and HQ site is directly transmitted to the Internet from the local CPEs. The IPS function needs to be enabled on the CPEs at the HQ site and branch site to block various intrusions from the Internet.

Overview of URL Filtering

- URLs, provided for manual operations on clients, open the first door for web attacks.
- Uncontrolled access of employees to website resources severely reduces work efficiency, wastes network bandwidth resources of enterprises, and introduces threats such as viruses and Trojan horses from malicious sites to the intranet. In addition, a great deal of pornographic and violence information affects people's physical and psychological health.
- URL filtering regulates users' online behaviors by controlling URLs accessible to users and permitting or denying users' access to some web resources.



HUAWEI

- URL filtering regulates users' online behaviors by controlling their HTTP requests and permitting or denying users' access to certain network resources.

URL Matching

- Each web page on the Internet has a unique identifier, that is, the URL.
- The URL format is as follows:

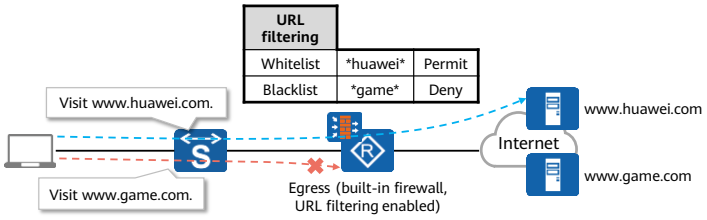
- A regular expression is typically used for URL matching. The following table lists commonly used matching modes.

Matching Mode	Definition	Example	Matching Result
Prefix matching	All URLs that start with a specified character string are matched.	www.test.com*	All URLs that start with **www.test.com** are matched, for example, **www.test.com/index.html**.
Suffix matching	All URLs that end with a specified character string are matched.	*aspx	All URLs that end with **aspx** are matched, for example, **www.test.com/news/solutions.aspx**.
Keyword matching	All URLs that contain a specified character string are matched.	*sport*	All URLs that contain **sport** are matched, for example, **sports.test.com/it/**.
Exact matching	A URL is first matched against a specified character string. If the URL is not matched, the last directory in the URL is removed, and the remaining part is matched against the character string. If the URL is still not matched, the last directory is removed. This process continues until the URL matches the character string.	www.example.com	Based on matching rules, the following URLs match **www.example.com**: • www.example.com • www.example.com/news • www.example.com/news/en/

- Each web page on the Internet has a unique identifier, that is, the URL.
- URLs fully describe the addresses of web pages or other resources on the Internet. To put it simply, a URL is a web address.
- The URL format is **protocol://hostname[:port]/path[? query]**
  - **protocol**: Used application protocol. HTTP is the most commonly used protocol. You do not need to enter **http://** when the protocol is HTTP.
  - **hostname**: DNS host name or IP address of the web server.
  - **port**: (Optional) Communication port. Each transmission protocol has a default port number.
  - **path**: Directory or file address on the web server.
  - **query**: (Optional) This field is used to transmit parameters to dynamic web pages.

Implementation of URL Filtering

- When an HTTP request matches a URL, a CPE functioning as a gateway processes the HTTP request according to the URL filtering mode and URL filtering process. The following URL filtering modes are supported:
  - Blacklist- or whitelist-based URL filtering: A CPE filters received HTTP requests based on the URL whitelist or blacklist configured on it.
  - Category-based URL filtering: After receiving an HTTP request, a CPE queries the URL category in the predefined URL category database. After the URL category is queried, the CPE processes the HTTP request according to the action defined for the URL category.

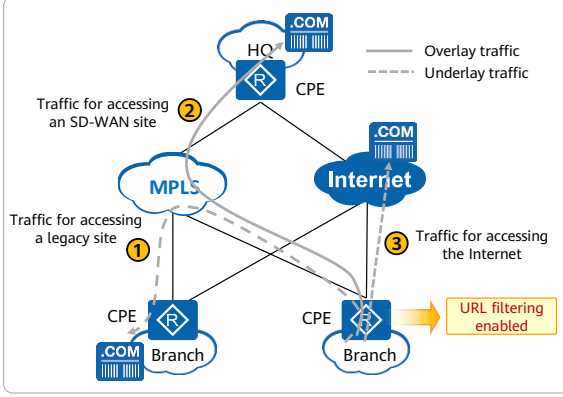


HUAWEI

- On a gateway, URL filtering is implemented as follows:
  - After the gateway receives an HTTP GET or POST request from a user, it checks the validity of the request based on the configured policies.
    - If the URL is valid, the HTTP request is permitted and the user can browse the website.
    - If the URL is invalid, the gateway pushes an alarm page and blocks the HTTP connection.
- URL categories fall into user-defined and predefined URL categories.
  - User-defined categories are configured and maintained by administrators. Administrators can perform more refined control over URLs in user-defined categories than over URLs in predefined categories.
  - Predefined categories contain common URLs in the system. Unlike user-defined categories, predefined categories enable administrators to easily control accessible and inaccessible URL categories.

Application Scenarios of URL Filtering

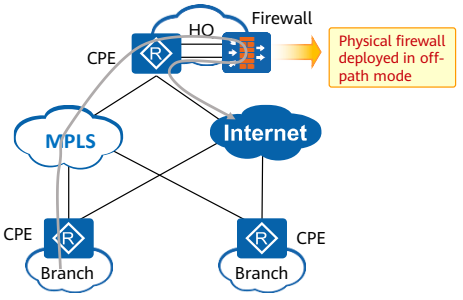
Application scenarios of URL filtering



- Randomly visiting various websites at branch sites may bring security risks. URL filtering regulates users' online behaviors by controlling URLs accessible to users and permitting or denying users' access to some web resources.
- URL filtering applies to the following scenarios:
  1. Site-to-legacy site access
  2. Site-to-SD-WAN site access
  3. Site-to-Internet access

Application Scenarios of Advanced Security VAS

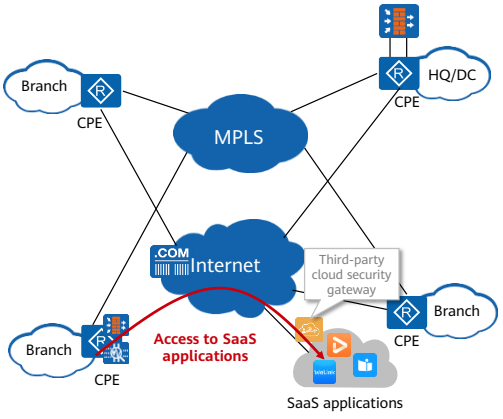
Application scenarios



- A physical firewall is deployed in off-path mode and provides advanced security protection for centralized Internet access traffic of sites.
- After the VAS function is deployed, the centralized Internet access traffic of sites is diverted to the LAN-side physical firewall connected to the CPE at the HQ site in off-path mode. After being processed by the firewall, the traffic is transmitted to the Internet through the underlay network.

Contents

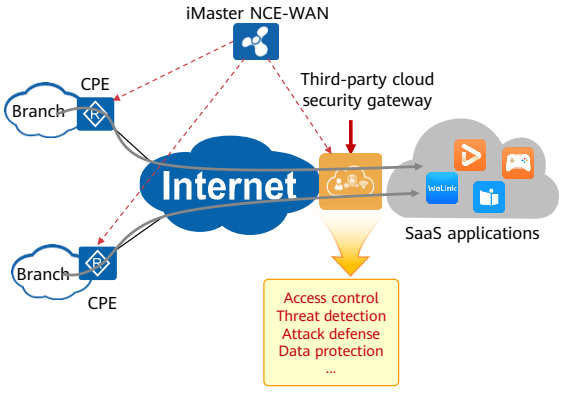
Site-to-SaaS Application Access Security



As traditional enterprise applications are gradually migrated to the cloud, branch sites of enterprises access applications on the cloud through the Internet, which may face security risks. In such service scenarios, a third-party cloud security gateway can be deployed to implement security protection on the cloud.

Application Scenarios of Third-Party Cloud Security

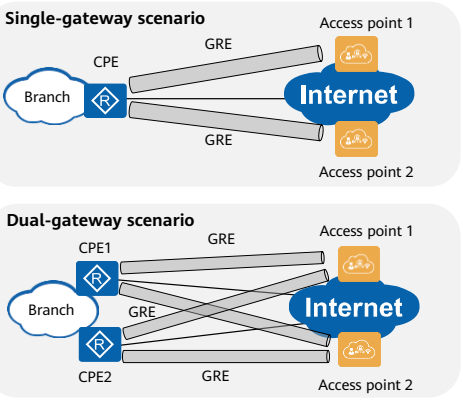
Application scenarios



- As traditional enterprise applications are gradually migrated to the cloud, branch sites of enterprises can directly access applications on public clouds through the Internet, which faces security risks. In such service scenarios, security protection can be implemented on the cloud.
- The SD-WAN Solution supports interconnection with a third-party cloud security gateway, so that traffic for accessing SaaS applications on the cloud can be sent to the third-party cloud security gateway for security protection.

Deployment Modes of Third-Party Cloud Security Gateways

Deployment modes



- To improve network reliability, a third-party cloud security gateway usually provides two access points to establish GRE tunnels with a CPE.
- Single-gateway scenario
  - A CPE establishes one active GRE tunnel and one standby GRE tunnel with the two access points of the third-party cloud security gateway.
- Dual-gateway scenario
  - Two CPEs establish four GRE tunnels (two pairs of active and standby GRE tunnels) with the two access points of the third-party cloud security gateway.

HUAWEI

- Policy-based routing (PBR) is configured on CPEs to divert the traffic for accessing SaaS applications to GRE tunnels. In addition, CPEs can use network quality analysis (NQA) to detect network reachability and perform an active/standby tunnel switchover.

Quiz

1. (Multiple-answer question) Which of the following are included in the inter-component communication security of the SD-WAN Solution?
   A. Management channel security
   B. Data channel security
   C. Control channel security
   D. Connection channel security
2. (True or False) The built-in firewall function enables CPEs to control the traffic between different security zones. For example, when the priorities of zone 1 and zone 2 are 20 and 60 respectively on a firewall, the traffic from zone 1 to zone 2 is outbound traffic.

1. ABC
2. False

Summary

- SD-WAN security includes system security and service security.
- System security includes the security of components (iMaster NCE-WAN, CPEs, and RRs) as well as the security of management, control, and data channels established between the components.
- Service security includes the site-to-site, site-to-Internet, and site-to-SaaS application access security. The firewall, IPS, URL filtering, advanced security VAS, and third-party cloud security functions can effectively defend against various security threats.

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.

SD-WAN Intelligent O&M

HUAWEI

Foreword

- When maintaining an SD-WAN network, network engineers need to quickly locate and rectify faults. In traditional CLI-based O&M mode, however, global traffic paths cannot be detected, and fault locating and rectification are time-consuming and inefficient and require high skills for network engineers. Therefore, CLI-based O&M is not applicable to quickly deployed large-scale networks.
- iMaster NCE-WAN provides visualized monitoring on network topologies and services, helping O&M personnel learn about the network running status and adjust the network accordingly. It also offers the alarm function to facilitate fault locating and rectification and the log function to facilitate fault source tracing.
- This course describes the monitoring, alarm, and log functions provided by iMaster NCE-WAN for intelligent O&M from the perspective of tenants.

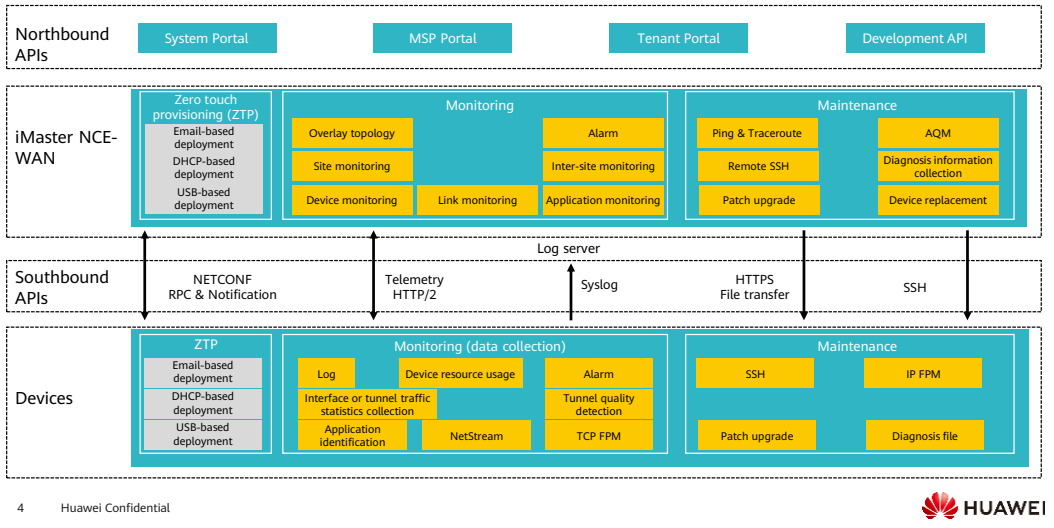
- This course is based on Huawei SD-WAN Solution.

Objectives

- On completion of this course, you will be able to:
  - Describe visualized monitoring functions provided by iMaster NCE-WAN.
  - Describe fault diagnosis methods provided by iMaster NCE-WAN.

Contents

Overall Intelligent O&M Architecture of the SD-WAN Solution



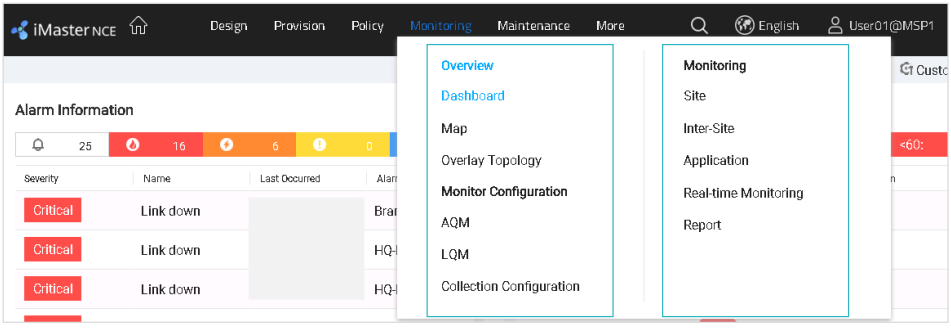
HUAWEI

- The main APIs between O&M components include:
  - NETCONF:
    - iMaster NCE-WAN delivers service configurations to devices through NETCONF.
    - Devices send alarms to iMaster NCE-WAN through NETCONF notification.
    - iMaster NCE-WAN uses NETCONF RPC messages to remotely maintain devices.
  - Telemetry:
    - Telemetry is used to encrypt HTTP/2 connections between devices and iMaster NCE-WAN.
    - Devices use Telemetry to periodically send collected performance data to iMaster NCE-WAN for performance monitoring.
    - Historical trend analysis provides high-precision monitoring data sources.
  - Syslog:
    - Devices can send logs to a third-party log server using the Syslog protocol.
  - HTTPS:
    - Devices and iMaster NCE-WAN use HTTPS for download and management of files such as software packages and patches.
  - SSH:
    - iMaster NCE-WAN allows O&M personnel to remotely log in to devices through SSH for O&M and diagnosis.

Contents

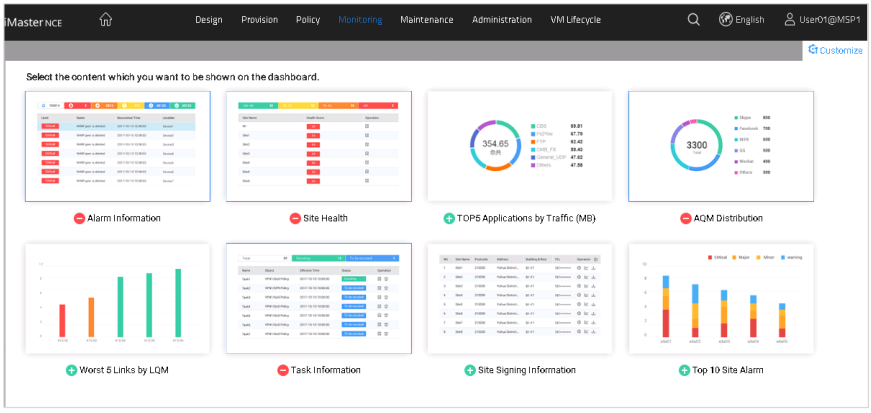
Monitoring Overview

- Monitoring is the basis of visualized O&M. iMaster NCE-WAN can monitor network and service status and statistics, providing basis for subsequent fault diagnosis, fault analysis, and policy application.
- iMaster NCE-WAN provides multi-dimensional monitoring capabilities, including the dashboard, site monitoring, service (link and application) monitoring, and alarm monitoring. The monitoring data can be sorted, searched, and filtered by date.



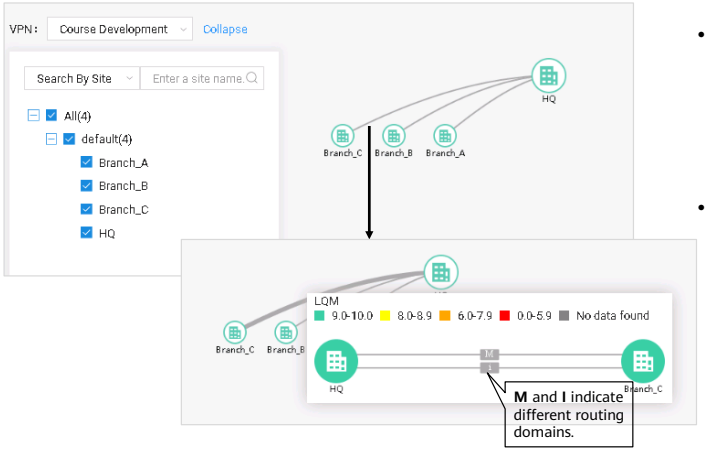
Customizing Information Displayed on the Dashboard

- Information displayed on the dashboard, such as alarm information, site health status, top worst links, and task information, can be customized based on customer requirements.



HUAWEI

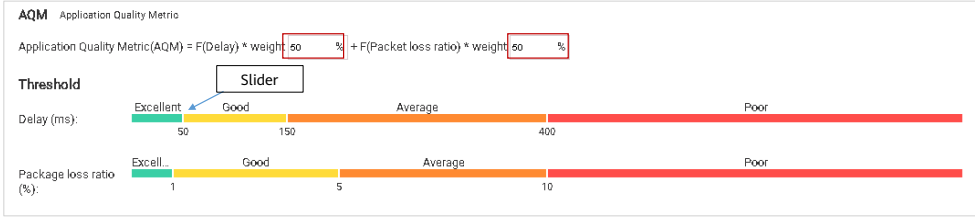
Overlay Topology Monitoring



- With the topology monitoring function, O&M administrators can view the overlay network status and check whether the overlay deployment result is correct.
- On the overlay topology, O&M administrators can also check the quality of inter-site links through link quality measurement (LQM) and discover inter-site link problems in advance.

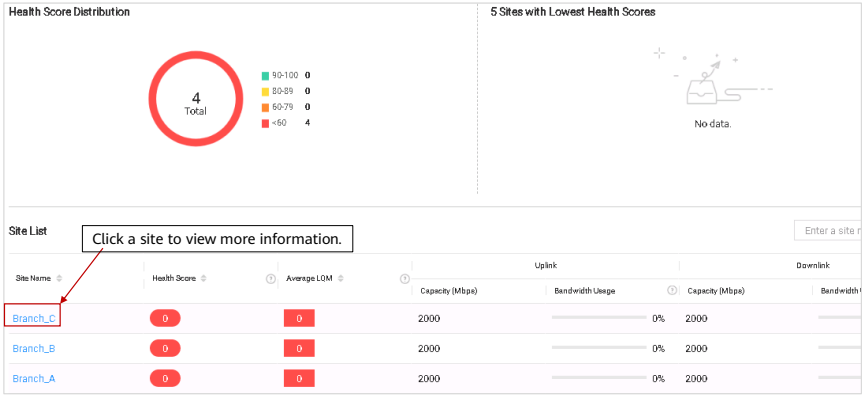
Customizing AQM and LQM

- O&M administrators can customize the AQM and LQM calculation rules and modify the thresholds of factors in the calculation formulas.
  - Take AQM as an example. Choose **Monitoring** > **Monitor Configuration** > **AQM** from the main menu.
  - Click **Modify** to adjust the weights of the delay and packet loss rate in the AQM calculation formula.
  - Drag the slider to adjust the delay and packet loss rate threshold ranges for **Excellent**, **Good**, **Average**, and **Poor**.



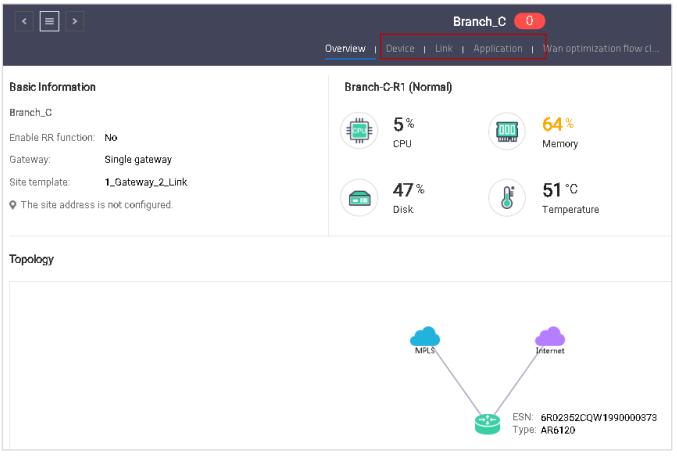
Site Monitoring (1)

- Choose **Monitoring** > **Monitoring** > **Site** from the main menu.



HUAWEI

Site Monitoring (2)

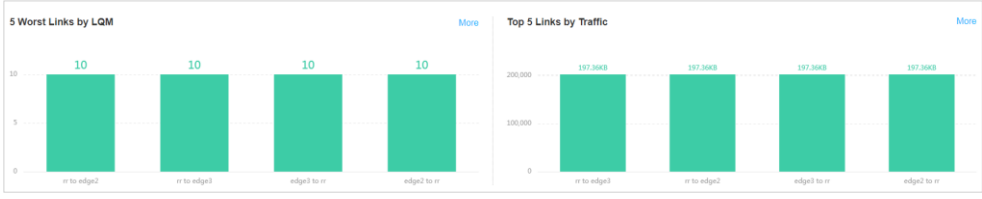


- **Overview**
  - View basic information about a site.
  - View the topology of a site.
  - View IP resources of a site.
  - View statistics on a site's visitor quantity, average AQM, bandwidth utilization, and throughput.
- **Device**
  - Monitor information about a single device, such as the CPU usage, memory usage, hard disk usage, and temperature.
- **Link**
  - Monitor links of a single site.
- **Application**
  - Monitor applications of a single site.

HUAWEI

Inter-Site Monitoring

- Choose **Monitoring** > **Monitoring** > **Inter-Site** from the main menu.
- View the worst 5 inter-site links by LQM.
- View the top 5 inter-site links by traffic.

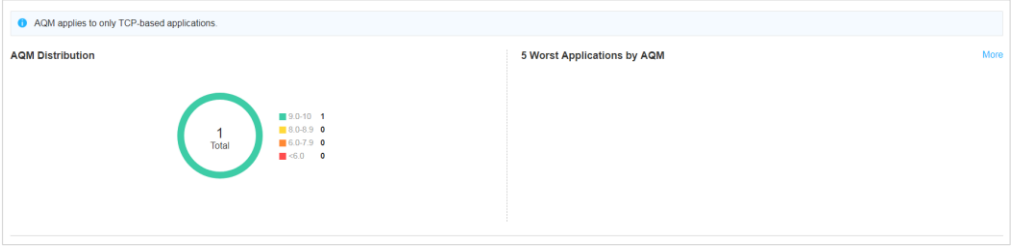


HUAWEI

Application Monitoring

O&M administrators can view statistics on the communication quality and visits of network-wide applications.

- Choose **Monitoring** > **Monitoring** > **Application** from the main menu.
- View AQM distribution information about network-wide applications and worst 5 applications by AQM.

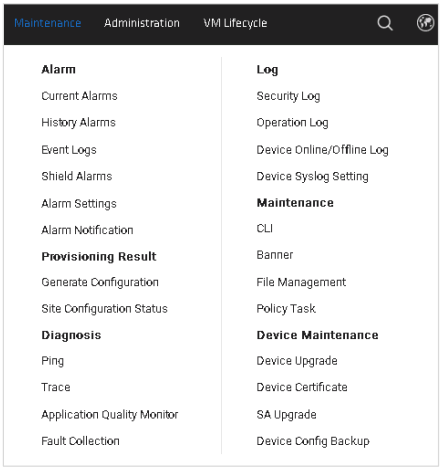


HUAWEI

Contents

Alarm | Configuration Result | Diagnosis | Log | Device Maintenance

Overview of Maintenance Tools



- iMaster NCE-WAN provides the following maintenance tools:
  - Alarm
  - Diagnosis
  - Log
  - Maintenance
  - Device maintenance

HUAWEI

- When iMaster NCE-WAN itself, devices, services, or systems managed by it, or its connections with peripheral systems are faulty or have potential risks, alarms are generated. With the alarm management function, O&M administrators can view alarm information in real time and troubleshoot faults based on alarm details and handling suggestions in a timely manner, ensuring normal running of services.
- On iMaster NCE-WAN, O&M administrators can set device alarm thresholds, for example, CPU and memory usage alarm thresholds. When a threshold is reached, an alarm is reported to iMaster NCE-WAN.
- When iMaster NCE-WAN is running, it can record system management operation logs and run logs, which facilitate audit and fault locating.

Alarm Classification

Alarm Category	Description	Administrator
Controller alarm	Alarms generated by iMaster NCE-WAN, including: • Alarms about services managed by iMaster NCE-WAN, such as topology change alarms • Alarms generated upon exceptions of iMaster NCE-WAN, such as license, cluster, or system status exceptions • Alarms generated due to disconnections between iMaster NCE-WAN and peripheral systems	System administrator
Device alarm	Alarms and events reported by devices to iMaster NCE-WAN, for example, device restart, user logout, interface protocol change, and smart policy change	Tenant administrator

Classification by alarm source

Classification by alarm status

Alarm Category	Description
Current alarm	Include uncleared and unacknowledged alarms, acknowledged and uncleared alarms, and unacknowledged and cleared alarms.
Historical alarm	Include alarms that have been cleared and acknowledged.
Masked alarm	Alarms that do not need to be handled can be masked. Masked alarms are displayed in the masked alarm list. If such alarms are generated later, they will not be displayed in the current alarm list.
Event	Events are alarms of the lowest severity and indicate that certain events happen. Events do not need to be handled.

Alarm Severities and Status

Alarm Severity	Color	Description	Handling Policy
Critical		Services are affected. Corrective measures must be taken immediately.	Rectify the fault immediately. Otherwise, services may be interrupted or the system may break down.
Major		Services are affected. If the fault is not rectified in a timely manner, serious consequences may occur.	Rectify the fault in a timely manner. Otherwise, key services will be affected.
Minor		The fault affects services slightly currently, but needs to be rectified to avoid more severe faults.	Find out the cause of the alarm and rectify the fault.
Warning		A potential or imminent fault that affects services is detected, but services are not affected currently.	Handle warning alarms based on the network and NE running status.

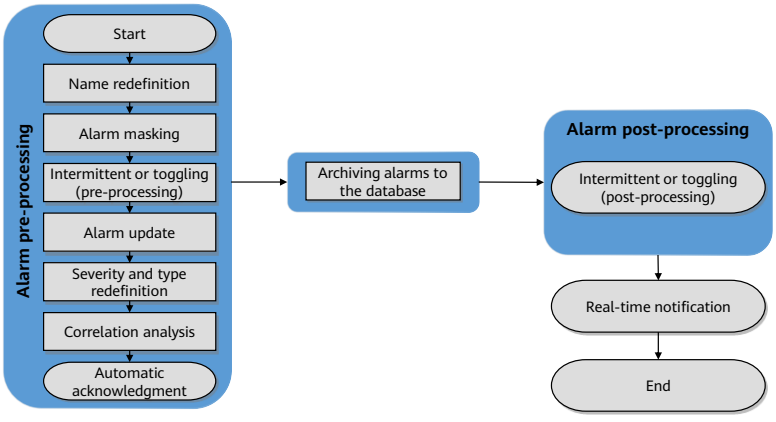
Status Name	Alarm Status	Description
Acknowledgment status	Acknowledged and unacknowledged	The initial acknowledgment status is **Unacknowledged**. A user who views an unacknowledged alarm and plans to handle it can acknowledge the alarm. When an alarm is acknowledged, its status changes to **Acknowledged**. Acknowledged alarms can be unacknowledged. When an alarm is unacknowledged, its status is restored to **Unacknowledged**. O&M administrators can also configure auto acknowledgment rules to automatically acknowledge alarms.
Clearance status	Cleared and uncleared	The initial clearance status is **Uncleared**. When the fault that causes an alarm is rectified, a clearance notification is automatically reported to the alarm management system and the clearance status changes to **Cleared**. For some alarms, clearance notifications cannot be automatically reported to the alarm management system. O&M administrators need to manually clear these alarms after corresponding faults are rectified. The background color of cleared alarms is green.

Alarm and Event Types

Alarm or Event Type	Description
Communications alarm	Alarms caused by failures of the communications in an NE, between NEs, between an NE and a management system, or between management systems, for example, device communication interruption alarms.
Service quality alarm	Alarms caused by service quality deterioration, for example, device congestion alarms.
Equipment alarm	Alarms caused by physical resource faults, for example, card fault alarms.
Operation alarm	Alarms generated when the required services cannot run properly due to problems such as service unavailability, faults, or incorrect invocation, for example, alarms caused by service rejection, service exit, and procedural errors.
Security alarm	Alarms generated when security issues are detected by a security service or mechanism, for example, alarms caused by authentication failures, confidential disclosure, and unauthorized access.
Object creation	Events generated when a managed object (MO) instance is created.
Object deletion	Events generated when an MO instance is deleted.
Relationship change	Events generated when relationship attributes of an MO change.
State change	Events generated when status attributes of an MO change.
Route change	Events generated when routes change.
File transfer status	Alarms or events reported when file transfer succeeds or fails.
...	...

Internal Alarm Handling Process

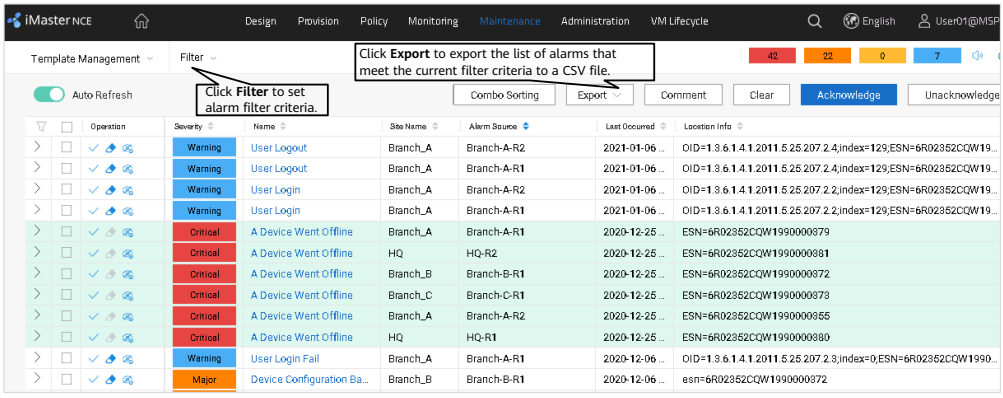
- After alarms are reported to the alarm management system, the system processes the alarms, including masking alarms, performing correlation analysis, and redefining severities.



Description of the Internal Alarm Handling Process

Operation	Description
Name redefinition	After receiving alarms, the alarm management system matches the alarms against the name redefinition rule and changes the names of the alarms that meet the name redefinition rule.
Alarm masking	The alarm management system discards the alarms that meet the masking rules, that is, the alarms are not archived to the database, or records the alarms in the masked alarm data table.
Intermittent or toggling (pre-processing)	The alarm management system records the alarms that meet the intermittent or toggling handling rules in the intermittent or toggling data table.
Alarm update	The alarm management system updates the information of current alarms, such as clearing alarms and changing the severities, based on the reported alarm changes.
Severity and type redefinition	The alarm management system redefines the alarms that meet the severity and type redefinition rules.
Correlation analysis	The alarm management system marks the alarms that meet the correlation rules as root alarms or correlative alarms, and handles the root alarms or correlative alarms based on the actions in the rules.
Automatic acknowledgment	The alarm management system automatically acknowledges the alarms that meet the auto acknowledgment rules. The alarms that are automatically acknowledged are recorded in the historical alarm data table.
Archiving alarms to the database	The alarm management system archives the remaining alarms to the database. Post-processing is not performed on the alarms that are masked or moved to historical alarms during alarm pre-processing. The information on the alarms is updated in real time.
Intermittent or toggling (post-processing)	The alarm management system analyzes the alarms in the intermittent or toggling data table and handles the alarms that meet the intermittent or toggling policies.
Alarm merging	The alarm management system merges the alarms that meet the merging conditions.
Real-time notification	The alarm management system updates the alarm information on the alarm interface in real time.

Alarm Management on iMaster NCE-WAN



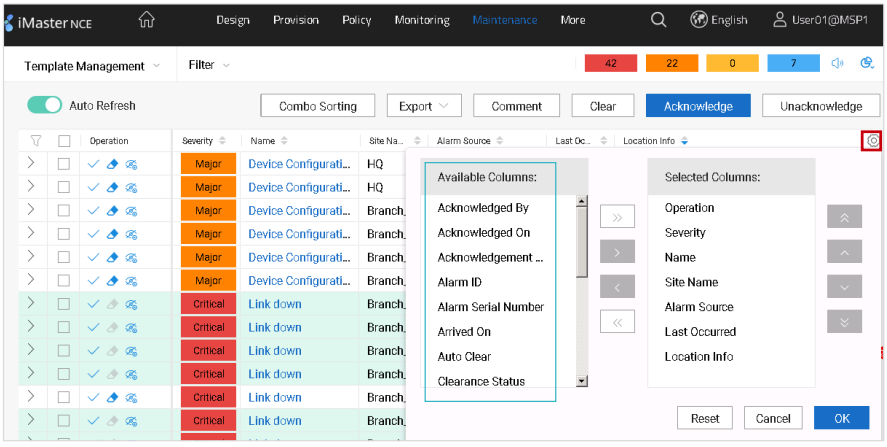
Click to clear an alarm.

Click to acknowledge an alarm.

Click to set a masking rule.

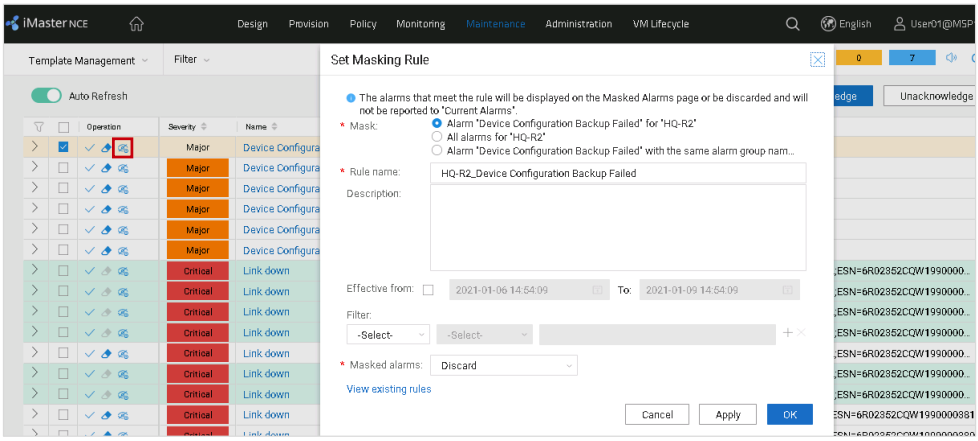
Customizing Alarm Items on iMaster NCE-WAN

Alarm items, such as the alarm ID and clearance time, can be customized on iMaster NCE-WAN.



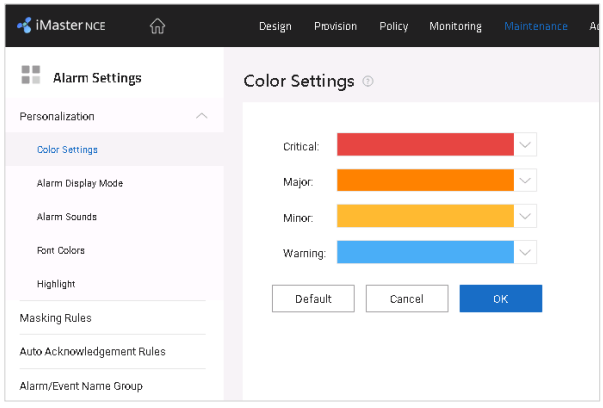
HUAWEI

Setting Alarm Masking Rules on iMaster NCE-WAN

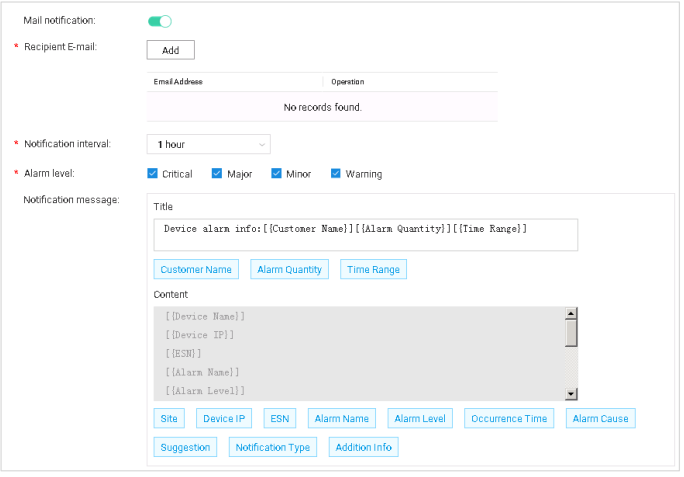


Alarm Settings on iMaster NCE-WAN

- Choose **Maintenance** > **Alarm** > **Alarm Settings** from the main menu. Alarm colors, display modes, and sounds can be set.



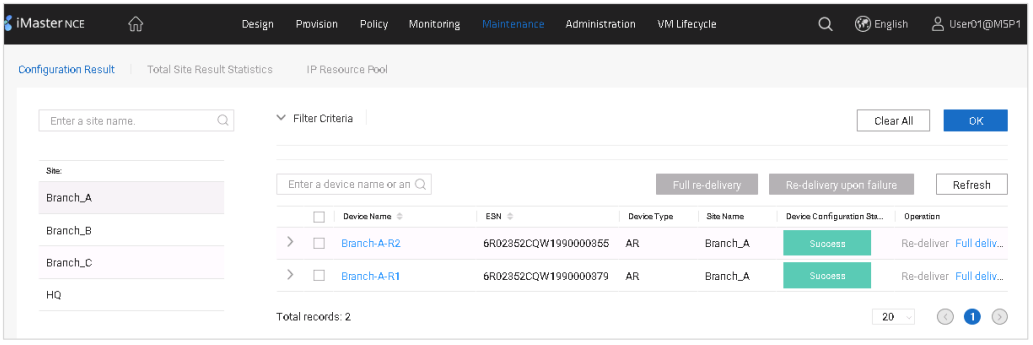
Configuring Alarm Notification on iMaster NCE-WAN



Choose **Maintenance** > **Alarm** > **Alarm Notification** from the main menu. Enable **Mail notification** and set related parameters, so that alarm notifications will be sent to the specified recipient by email.

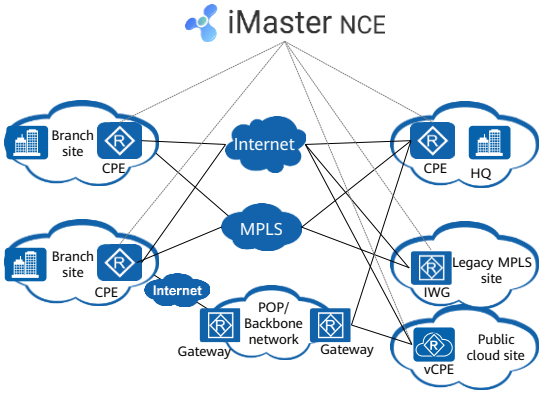
Viewing Configuration Results

Choose **Maintenance** > **Provisioning Result** > **Site Configuration Status** from the main menu. Click the **Configuration Result** tab and select a site to view the site's configuration result.



HUAWEI

Diagnosis Overview



iMaster NCE-WAN provides the following diagnosis functions:

- Ping

  On iMaster NCE-WAN, an administrator delivers a **ping** command to a CPE to check the connectivity, packet loss rate, and average delay of the underlay and overlay networks between the CPE and the destination host.

- Trace

  On iMaster NCE-WAN, an administrator delivers a **traceroute** command to a CPE to track the route packets take from the CPE to the destination host. iMaster NCE-WAN then provides the visualized diagnosis result for the administrator.

- Application quality diagnosis

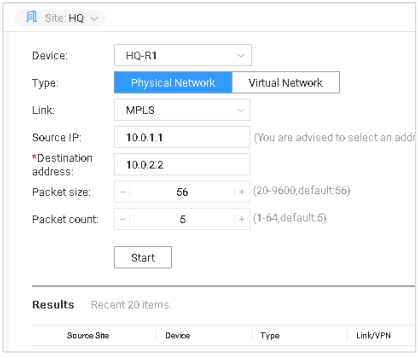
  An administrator delivers a command for checking the quality of a specified application to a CPE. The CPE then uses IP FPM technology to check the quality of the specified application (only the packet loss rate can be checked).

- Fault information collection

  If a CPE or network fault occurs but the root cause cannot be located using the preceding methods, O&M administrators can collect the CPE's diagnosis information with one click and send the information to Huawei technical support for fault locating.

Diagnosis: Ping and Trace

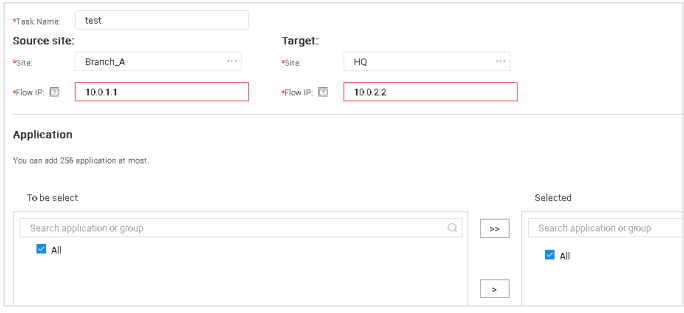
- Choose **Maintenance** > **Diagnosis** > **Ping** from the main menu.
- On the **Ping** page that is displayed, configure parameters for checking the network connectivity, including the source site, device, network type, link or virtual private network, source IP address, destination IP address/URL, packet size, and number of packets.
- Click **Start** to start the network connectivity check.



- The trace operation is similar to the ping operation.

Diagnosis: AQM

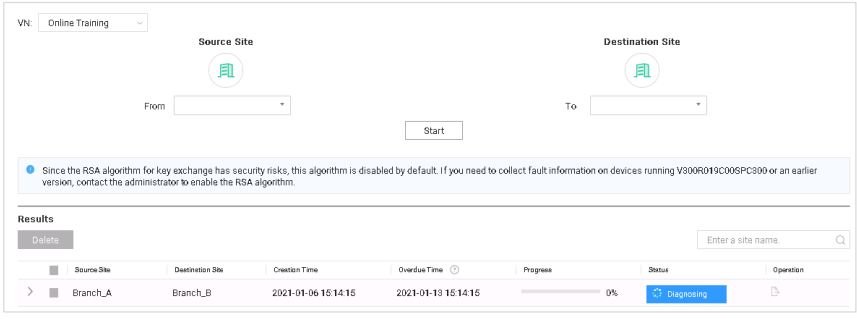
- Choose **Maintenance** > **Diagnosis** > **Application Quality Monitor** from the main menu. On the **Application Quality Monitor** page that is displayed, click **Create** to create an application traffic detection task.
- Set the task name, source site, and destination site, select the application to be diagnosed from the application list, and click >. Then, click **OK**. The task starts automatically.
- To diagnose an application again, select the target diagnosis task and click **Start** to start diagnosis. After sufficient data is collected, click **Stop**.



HUAWEI

Diagnosis: Fault Information Collection

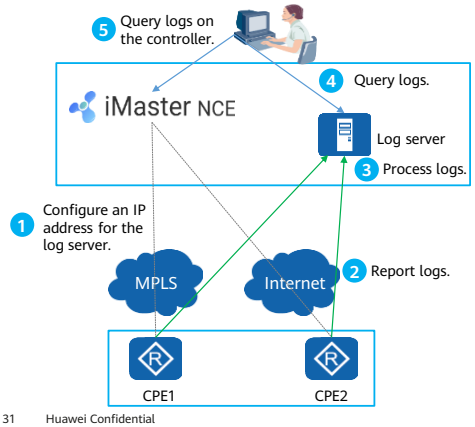
- Choose **Maintenance** > **Diagnosis** > **Fault Collection** from the main menu. The **Fault Collection** page is displayed.
- Two methods are available for downloading fault diagnosis logs from iMaster NCE-WAN.
  - To export fault diagnosis logs of all devices in a collection task, click in the **Operation** column of the task.
  - To export fault diagnosis logs of a single device, click › next to the device to display the fault information collection details. Then click in the **Operation** column to export fault diagnosis logs.



HUAWEI

Log Management Overview

- When iMaster NCE-WAN is running, it can record system management operation logs and run logs, which facilitate audit and fault locating.
- Tenant administrators can configure devices to report device logs to a third-party log server and view device logs on the log server, facilitating device management and maintenance.



1. A tenant administrator configures log server parameters on a CPE through iMaster NCE-WAN.
2. Logs generated by the CPE are reported through the syslog channel.
3. The log server receives and displays log information.
4. The tenant administrator queries the device logs saved on the log server.
5. The tenant administrator queries logs on the iMaster NCE-WAN web UI.

HUAWEI

- When iMaster NCE-WAN is running, it can record system management operation logs and run logs, which facilitate audit and fault locating.

Log Types

Log Type	Description
Operation logs	Record all add, delete, and modify operations triggered by users or iMaster NCE-WAN to facilitate audit.
Security logs	Record operations related to user accounts, such as login to and logout of iMaster NCE-WAN as well as password change, to facilitate audit.
Device login and logout logs	Record device login and logout information to facilitate audit.
Device logs	Because the storage space of iMaster NCE-WAN is limited, device logs are not saved. Therefore, tenant administrators cannot view and audit historical logs of devices. To resolve this problem, they need to configure a third-party log server to which devices report logs, and then can log in to the log server to view the logs for device management and maintenance.

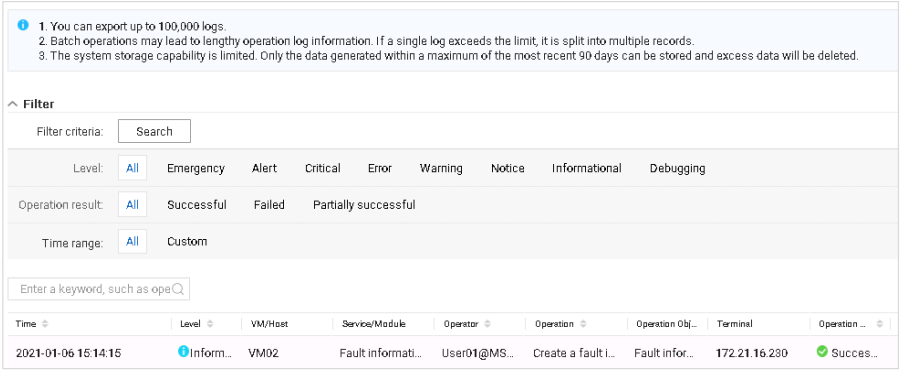
Tenant administrator

System administrator

Log Type	Description
Operation logs	Record all add, delete, and modify operations triggered by users or iMaster NCE-WAN to facilitate audit.
Security logs	Record operations related to user accounts, such as login to and logout of iMaster NCE-WAN as well as password change, to facilitate audit.
Run logs	Record status information generated during the running of iMaster NCE-WAN and execution of tasks. Run logs can be used to diagnose iMaster NCE-WAN faults.

Log: Operation Logs

- On the **Log** page, O&M administrators can manage security logs, operation logs, device login and logout logs, and device system logs.

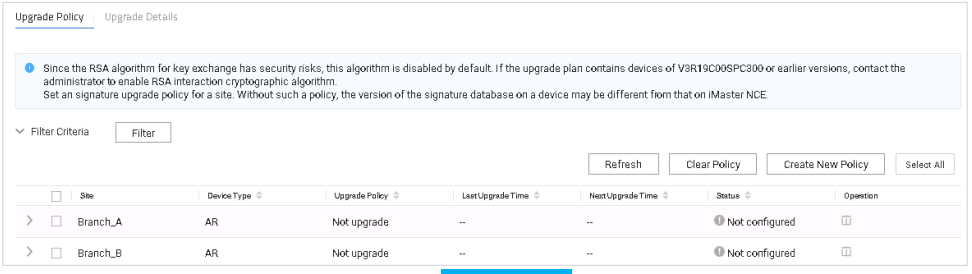


Operation logs

HUAWEI

Device Maintenance

- iMaster NCE-WAN provides the device upgrade, device certificate management, SA upgrade, and device configuration backup functions on the **Maintenance** > **Device Maintenance** page.



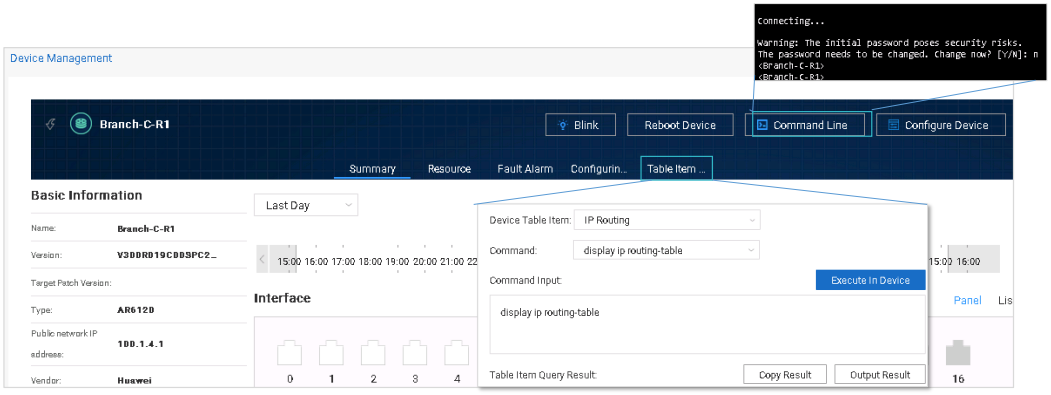
SA upgrade

- iMaster NCE-WAN also provides multiple device management functions, such as remote device login, indicator blinking, and one-click device restart.

HUAWEI

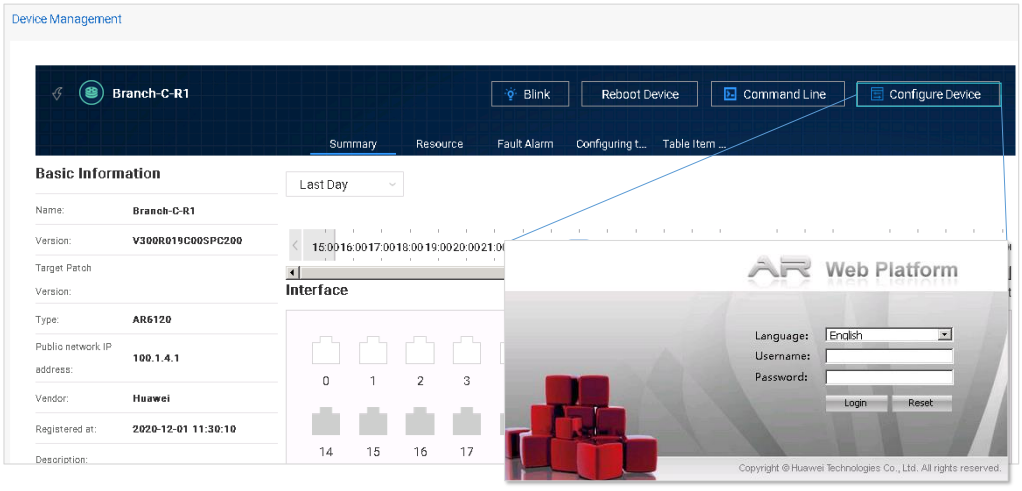
Device Maintenance: Remote Login to a Device

- Choose **Design** > **Site Agile Design** > **Devices Management** from the main menu.
- Click the name of a device to access the device details page.
- Click **Command Line** in the upper right corner to remotely log in to the device's CLI. Select the device administrator authentication mode.
- Obtain the device's routing table and ARP table with the device entry query function provided by iMaster NCE-WAN.



HUAWEI

Device Maintenance: Web Management System



HUAWEI

Quiz

1. (Single-answer question) Which protocol does iMaster NCE-WAN use to deliver service configurations to devices?
   A. NETCONF
   B. Telemetry
   C. Syslog
   D. SSH
2. (Multi-answer question) Which of the following device information can be viewed on the iMaster NCE-WAN web UI?
   A. CPU usage
   B. Memory usage
   C. Hard disk usage
   D. Temperature

- 1. A
- 2. ABCD

Summary

- iMaster NCE-WAN provides multi-dimensional monitoring capabilities, including the dashboard, site monitoring, service (link and application) monitoring, and alarm monitoring. The monitoring data can be sorted, searched, and filtered by date.
- iMaster NCE-WAN provides various O&M tools, such as alarm management, diagnosis, log management, and device maintenance. With the alarm management function, users can learn about alarm information in real time, and rectify faults based on alarm details and handling suggestions to ensure normal service running. They can perform ping and trace operations to check network connectivity, and view operation logs, security logs, as well as device login and logout logs. In addition, they can upgrade the SA signature database or software versions of devices based on actual requirements, and log in to CPEs through SSH.
- These O&M tools are used together to implement intelligent O&M.

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.

SD-WAN Design Practice (FSI Scenarios)

HUAWEI

Foreword

- In today's world, people's production and consumption activities are closely related to finance, and digitization of the financial services industry (FSI) becomes more important than ever.
- Ever-changing service types and external environments, such as cloud and lightweight services, pose new requirements on financial services.
- Financial services constantly change, raising new challenges to networks. In response, Huawei offers the SD-WAN Solution to meet new requirements of financial WANs.
- This course describes typical networking modes and design schemes of Huawei SD-WAN Solution in the FSI (banks).

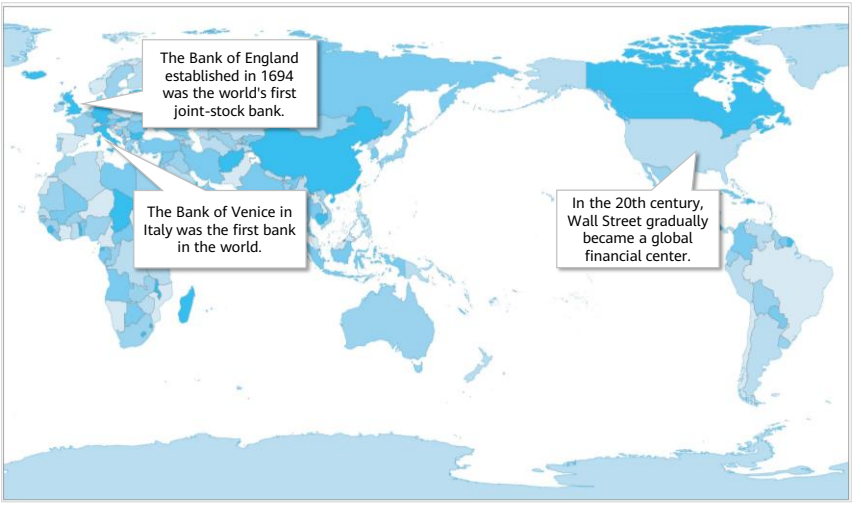
- This course is based on Huawei SD-WAN Solution.

Objectives

- On completion of this course, you will be able to:
  - Describe typical services in the FSI.
  - Describe the ICT development trend of the FSI (banks).
  - Understand the network architecture of the FSI (banks).
  - Describe how to design SD-WAN networks in the FSI (banks).

Contents

Brief History of the FSI



HUAWEI

- The FSI began with treasure deposits and loans provided by temples of Babylon as early as 2000 BC and Greek temples in the 6th century BC.
- In the years from the 3rd century BC to the 3rd century AD, silver coinage merchants and bank-like commercial organizations came into being in ancient Athens and Rome.
- In Europe, modern banks were originated from currency exchange and goldsmith services, the first of which was the Bank of Venice in Italy (established in 1580).
- The world's first joint-stock bank, the Bank of England, was established in 1694, which determined a basic form of organizations for the modern FSI.

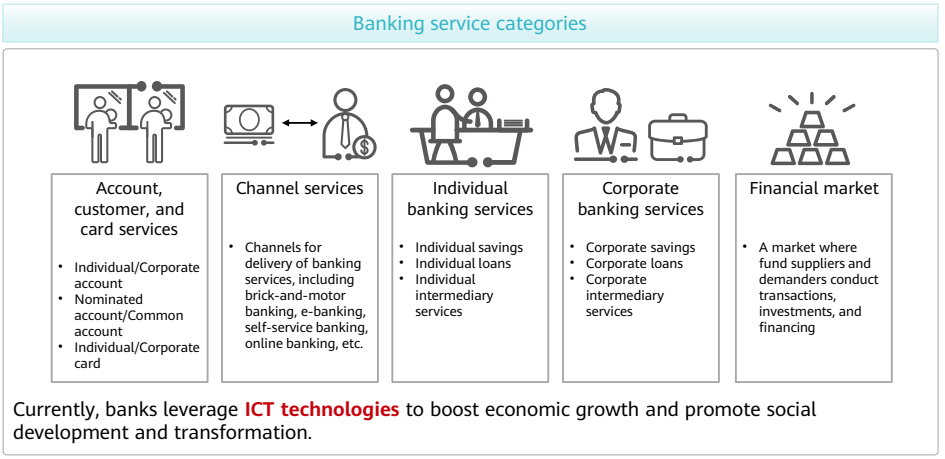
Classification and Functions of the FSI

HUAWEI

- The FSI connects all aspects of national economy. Meanwhile, financial means, such as interest rate, exchange rate, credit, and settlement, have direct influence on micro-economic entities.
- Finance is related to the economic sovereignty and wealth control of a country, and plays an important role in maintaining economic growth and national interests as well as serving the real economy and citizens.
- Finance is not only the core of modern economy, but also the core of modern politics and modern society.
- Banking: Banks are financial institutions that offer financial services, such as deposits, loans, remittances, and savings, and also serve as credit intermediaries.
- Insurance: is a means of protection from financial loss. It is a form of risk management, primarily used to hedge against the risk of a contingent or uncertain loss.
- Trust: A trust company is a legal entity that acts as a fiduciary, agent, or trustee on behalf of a person or business for a trust. A trust company is typically tasked with the administration, management, and the eventual transfer of assets to beneficiaries.

- Securities: The securities sector is engaged in securities issuance and transaction, and consists of stock exchanges, securities companies, securities associations, and other related financial institutions.
- Lease: A lease is a contractual arrangement calling for the lessee (user) to pay the lessor (owner) for use of an asset.

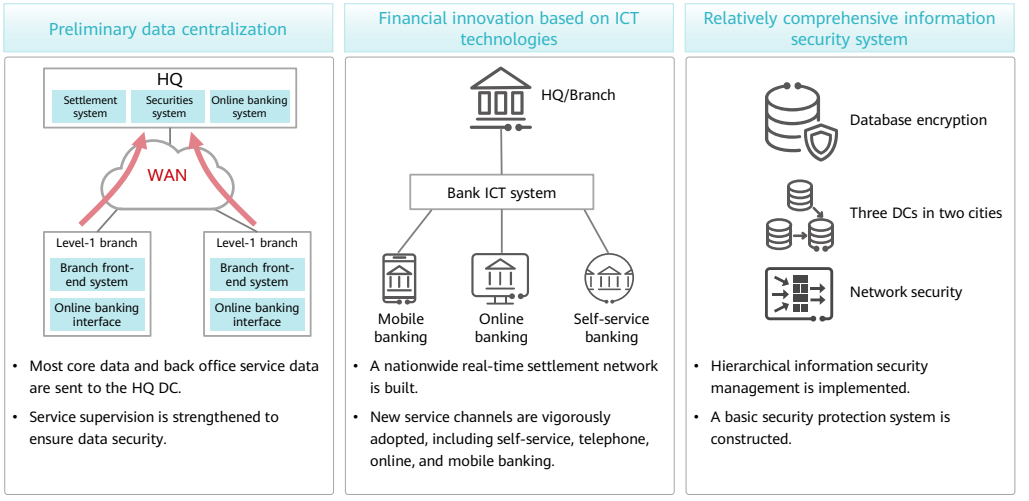
Overview of Banking Services



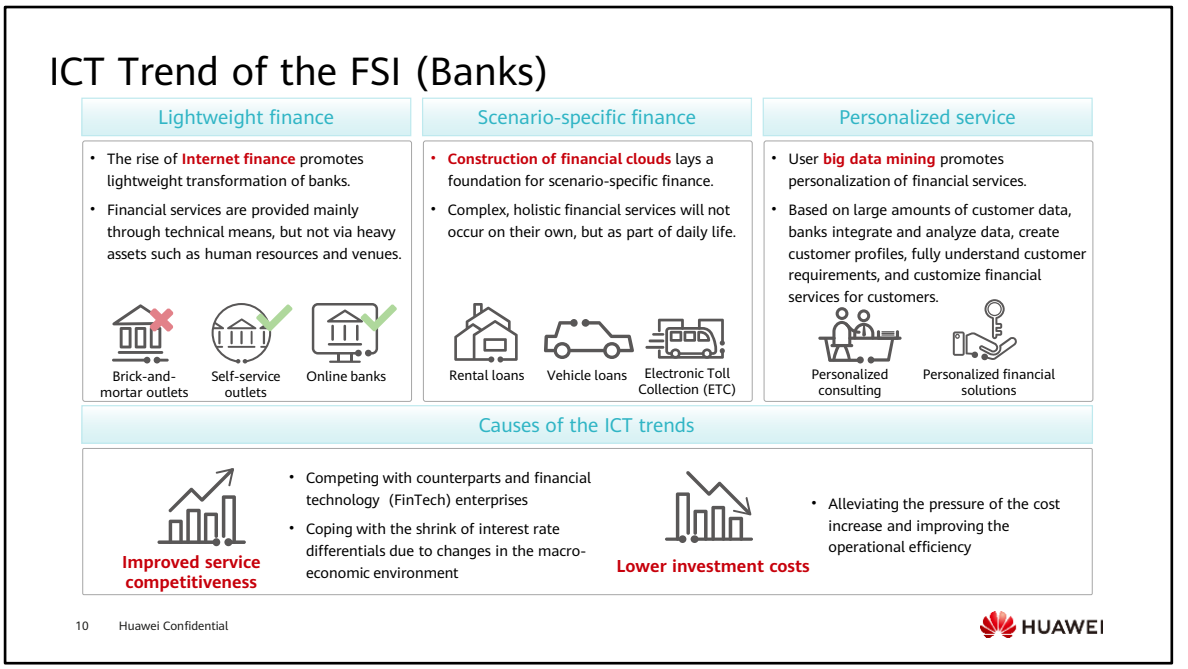
- This course uses the banking sector as an example to describe ICT construction requirements of the FSI.
- Based on bank functions and architecture, a bank is broken up into three parts: front office, middle office, and back office.
  - The front office is responsible for service development. It is directly oriented to customers and provides one-stop and all-round services for customers. Bank tellers, account managers, and lobby managers are all front office personnel.
  - The middle office is responsible for formulating service development policies and strategies by analyzing the macro market environment and internal resources, providing professional management and guidance for the front office, and controlling risks.
  - The main responsibilities of the back office are to support and process services and transactions, including accounting treatment, IT support, and call center. It is also responsible for centralized loans approval.

Contents

Current ICT Situation of the FSI (Banks)

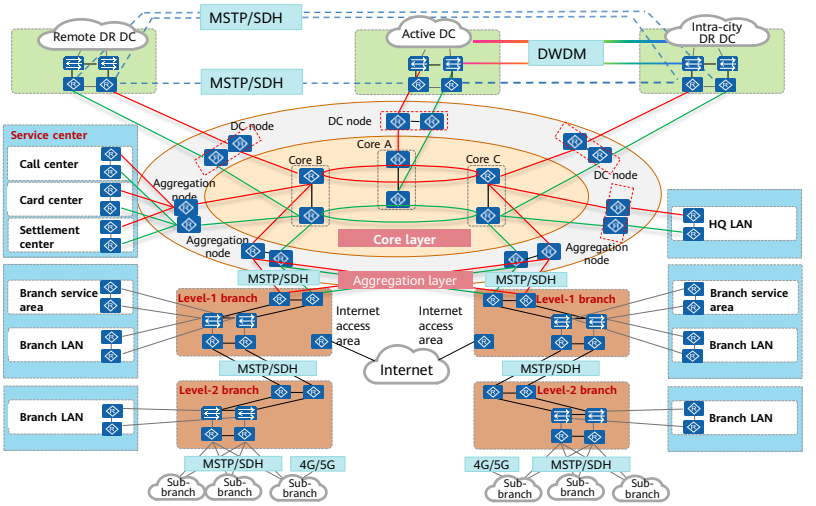


- The goals of ICT construction of banks are more than to computerize service operations. More importantly, there are two other goals. One is to build and improve a financial risk control mechanism by integrating technological transformation with institutional transformation, and the other is to reshape service models and processes by adopting ICT technologies.



- FinTech is the technology and innovation that aims to compete with traditional financial methods in the delivery of financial services.
- Based on the FinTech revolution, digital banking and mobile finance that focus on services and experience gradually change the service model of banks, create new growth points of digital finance, and play an increasingly important role in bank services.
- Big data, AI, IoT, and cloud computing technologies also provide new technical engines for bank outlets to implement full-link evolution from perspectives of customer management, process reconstruction, risk prevention and control, open ecosystem, and channel convergence.
- The ICT trends of finance and the causes of the trends pose great challenges to financial WANs.

Typical WAN Architecture of the FSI (Banks)

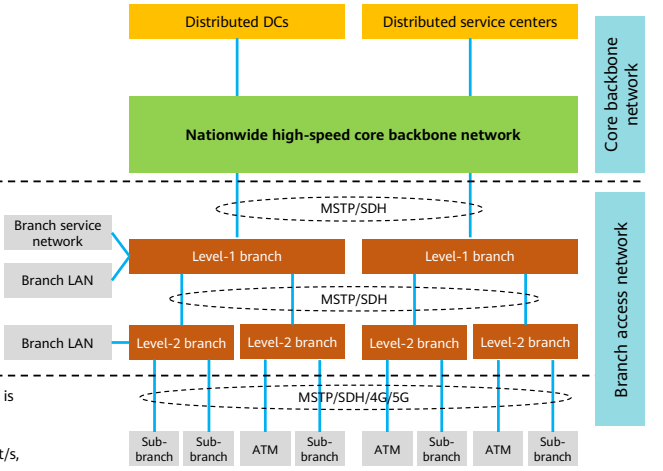


HUAWEI

- Multi-DC disaster recovery (DR)
  - Hierarchical design of the physical network ensures stability of the core backbone network.
  - Multiple DCs in multiple places are interconnected with each other over the cloud through the core backbone network.
- Hierarchical WAN networking
  - The tree network structure is used, and requires hierarchical network construction and level-by-level traffic aggregation.
  - East-west traffic diversion is prevented and link utilization is improved.
  - Network O&M responsibilities are clarified to avoid cross-area maintenance.
- Flat networking of outlets
  - Abundant line resources allow densely located outlets in cities to directly connect to branches, forming a flat network.
  - Network construction and maintenance costs are reduced.
  - The impact of the increase in line leasing costs on the overall cost needs to be comprehensively considered.

Current Situation of Financial (Bank) WANs

- The backbone network connects branches/sub-branches to DCs.
- The traffic includes north-south traffic and only a small amount of east-west traffic.
- One physical network carries all the office, production, and security protection services.
- Dual-device or single-device dual-uplink networking is used, achieving traffic load balancing.
- Dual-device or single-device dual-uplink networking is used, achieving traffic load balancing.
- The uplink bandwidth is 20 Mbit/s, 10 Mbit/s, 6 Mbit/s, or 4 Mbit/s, and the bandwidth utilization is 60%.



HUAWEI

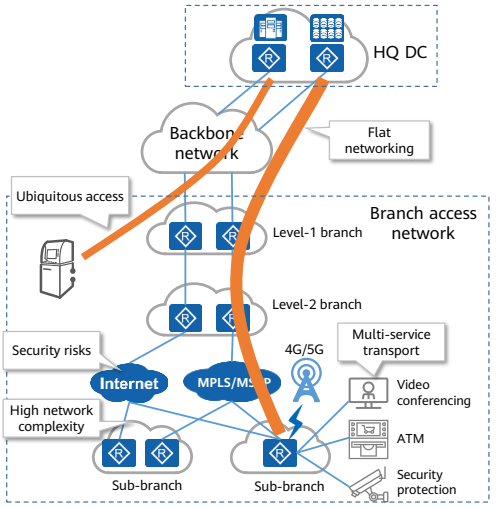
- A financial (bank) WAN consists of the branch access network and core backbone network .
- The branch access network transmits traffic from branches and sub-branches to the core backbone network.
- The core backbone network transmits traffic from branches to the DCs or HQ, and is also responsible for DC interconnection.

Current Situation of Financial (Bank) Branch Access Networks

Category	Network Situation
Network bandwidth	1. The private line bandwidth is insufficient. The network bandwidth difference between level-2 branches is large, typically, in the range from 4 Mbit/s to 32 Mbit/s. Currently, the network bandwidth of sub-branches or outlets is 20 Mbit/s, 10 Mbit/s, 4 Mbit/s, or 2 Mbit/s. The average bandwidth utilization exceeds 60%. 2. Burst of heavy traffic affects key services. Video conferences and learning materials occupy a large high bandwidth in a short period of time. 3. Security protection occupies private line bandwidth. Banks need to view HD surveillance videos of outlets in real time (2 Mbit/s bandwidth is required for one channel of videos). The bandwidth for transmitting a maximum of two channels of videos must be ensured. 4. The patch and virus library need to be updated periodically for terminal access, which occupies a large number of link resources. 5. Value-added service (VAS) traffic affects mission-critical services: The facial recognition service will also occupy private line link resources.
Network policy	1. It is difficult to adjust policies in IP-based management mode. To adjust the service scope, banks need to manually modify policies at multiple control points. 2. CLI-based manual configuration results in complex policy deployment, and frequent network changes cause heavy configuration adjustment workloads. 3. Fragmented policy control is used, and E2E policy orchestration and streamlining from access points to DCs cannot be implemented.
Network O&M	1. No dedicated network maintenance personnel are available for a large number of branches. There is a small number of network maintenance personnel of level-1 and level-2 branches. 2. Banks lack in fault locating methods. Terminals cannot access the Internet or the Internet access speed is slow. No quick fault locating method is available. O&M personnel have to use the **ping** command to locate network faults segment by segment.

Challenges Facing Branch Access Networks Under New Financial Trends

- Branch access networks of banks face the following challenges under new financial trends:
  - Networks become flat.
  - Cloud services reshape bank WANs.
  - Diversified banking services make multi-service transport a must.
  - Network complexity increases, and O&M costs need to be reduced.
  - Network security risks increase.



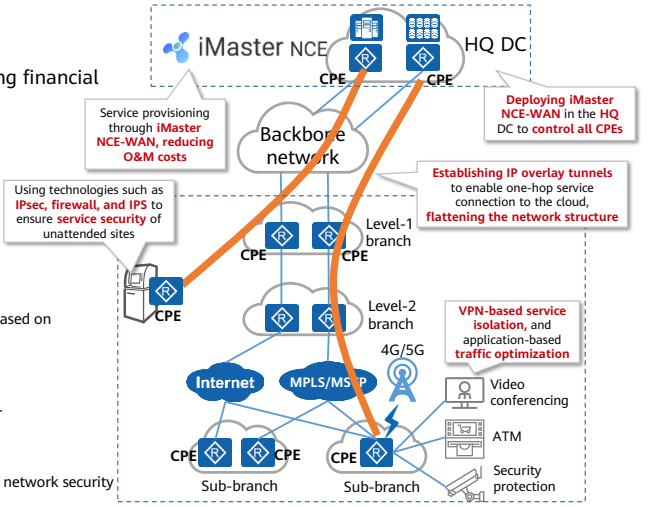
HUAWEI

- Currently, the business growth of brick-and-mortar bank branches is slowing down, and banks' requirements for coverage of these branches are greatly reduced. Brick-and-mortar branches seem to become operation burdens of banks.
- Bank branches are undergoing an intelligent, digital, and lightweight transformation. Driven by ICT technologies, brick-and-mortar branches break financial service barriers, and well converge online and offline omni-channel financial services. They gradually transform from purely transaction settlement service nodes to service, experience, expansion, and marketing centers that provide marketing and investment consulting services.
- Under this trend, banks' branch access networks face the following challenges:
  - Networks become flat.
    - Centralization of production services demands for flat networking.
    - Public cloud private lines are introduced, requiring one-hop connection to the cloud.

    - Cloud services reshape bank WANs.
        - Distributed deployment of banking services and emergence of multi-cloud DCs enable the DCI network topology of bank WANs to evolve to mesh interconnection.
        - The development of cloud services requires bank WANs to provide ubiquitous connections instead of covering only bank branches.
        - Cloud-and-network synergy drives network revolution towards SDN.
    - Diversified banking services make multi-service transport a must.
        - Mixed service operations and cloud access of branches require the access networks to provide multi-service transport and isolation capabilities.
        - Diversified services promote multi-network integration to implement one network with two domains (financial and non-financial service domains).
        - Remote counter services based on video conferencing develop, increasing network traffic.
- Network complexity increases, and O&M costs need to be reduced.
    - Multiple types of links are introduced to carry multiple links, improving the cost-effectiveness of purchasing links.
    - The flattening trend leads to an increase in link fees, and the cloudification leads to an increase in ubiquitous connections.
    - The SDN technology introduced by the development trend of cloud-network synergy implements unified O&M on the entire network.
    - Network security risks increase.
    - Increased intranet security risks
    - IoT security risks introduced by smart devices and IoT devices
    - Cloud-network-security synergy is difficult.

Addressing Challenges Facing Financial Branch Access Networks

- Using SD-WAN to address the challenges facing financial branch access networks:
  - Network flattening
    - Establishing E2E IP overlay tunnels
  - Cloud services, reshaping bank WANs
    - One-hop connection to the cloud
  - Multi-service transport
    - Multi-VN isolation, and intra-VN traffic optimization based on applications
  - High O&M costs
    - Centralized management and control by the controller
  - Network security risks
    - Technologies, such as IPsec, firewall, and IPS, ensuring network security

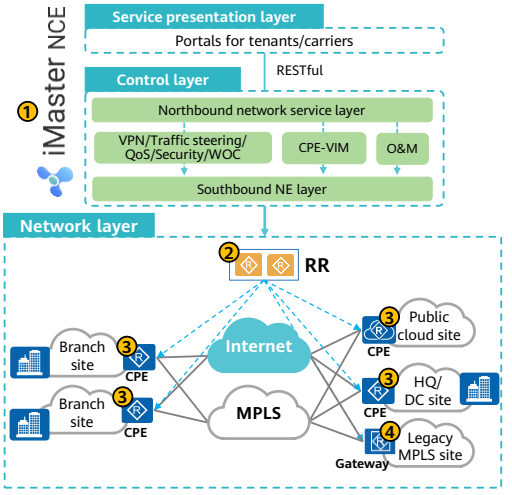


HUAWEI

- Financial networks have the following characteristics:
  - Multiple DCs for disaster recovery
    - At least two DCs are built.
    - They communicate with each other at Layer 3, and learn routes from and advertise routes to each other.
    - They also provide data services and work in active/standby mode.
  - Large number of outlets
    - Generally, more than 1000 outlets on the live network need to be centrally managed.
    - Especially, large-scale enterprises generally have tens of thousands of outlets.
  - High performance of aggregation nodes
    - With the increasing density of outlets in cities, carriers' line resources are continuously enriched, and the fees are continuously reduced. Outlets can be directly connected to branches.
    - In this case, aggregation devices are required to provide forwarding performance.

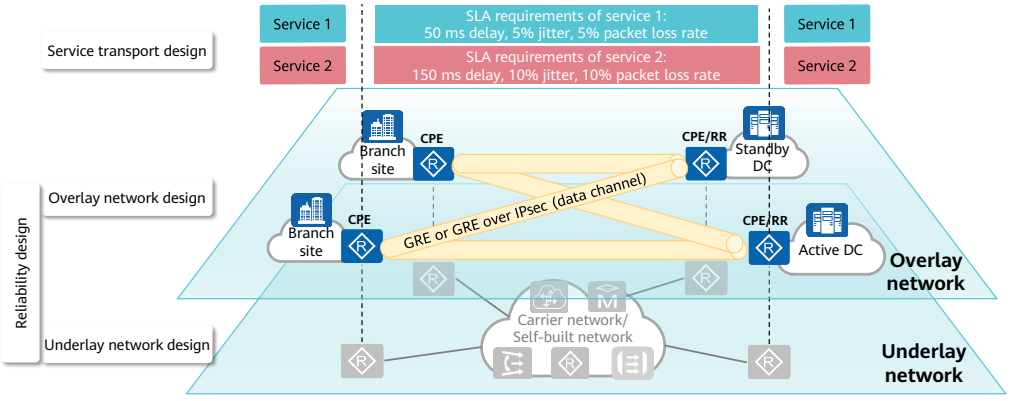
Contents

Key Components of the SD-WAN Solution



No.	Component	Functions
1	iMaster NCE-WAN	1. Network service orchestration 2. NE control 3. Basic network O&M 4. CPE orchestration and management 5. Basic performance monitoring (providing link quality information, application quality information, and traffic information, as well as statistics from dimensions such as intra-site and inter-site statistics)
2	RR	Distributes VPN routes and tunnel information between CPEs based on VPN topology policies.
3	CPE	Functions as the egress device of a site.
4	Gateway	Connects an SD-WAN network to a non-SD-WAN network.

Overall Design Roadmap of a Financial SD-WAN Network

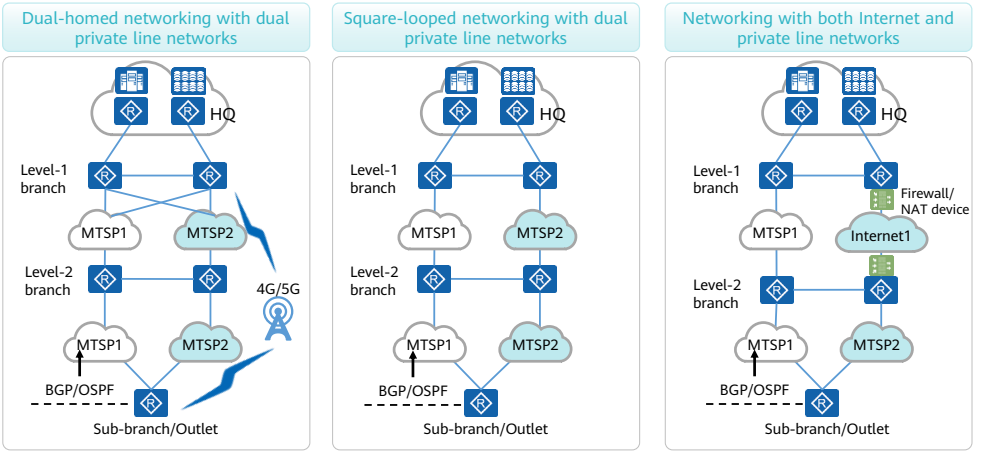


HUAWEI

- An SD-WAN network can be designed from the following perspectives:
  - Underlay network design: includes the WAN-side networking design and LAN-side networking design.
  - Overlay network design: includes the topology design and VN design.
  - Service transport design: includes the transport network selection and security design.
  - Reliability design: is involved in the physical topology, logical topology, and service transport design, and includes the link reliability design, CPE reliability design, controller reliability design, and RR reliability design.
- This course describes the underlay network design, overlay network design, and reliability design.

Contents

Typical Underlay Networking (1)

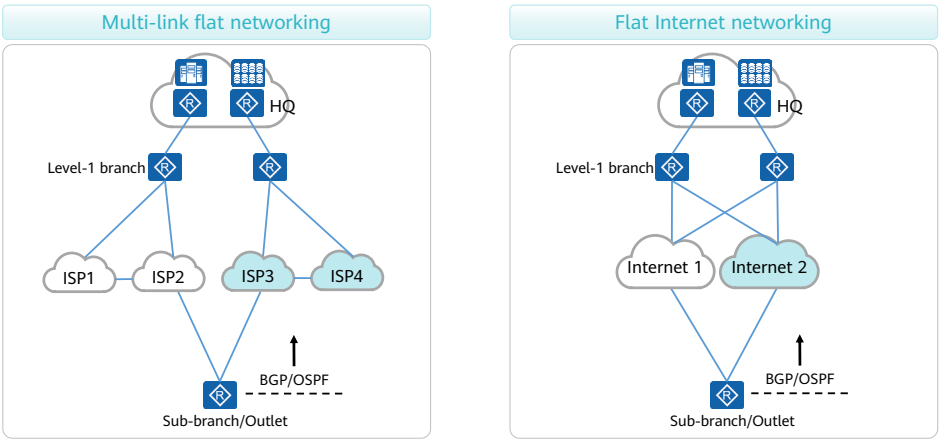


HUAWEI

- On the underlay WAN side, IP addresses and IGP routing protocols are configured based on the networking habits of the live network. Generally, a bank network uses a dynamic routing protocol at the egress to connect to upper-level branch devices.

Scenario	Dual private line cross-connection	Double private line square-shaped	Private line + Internet private line
Level 1 branch	Cross-province dual private line interconnection	Cross-province dual private line square-shaped interconnection	Cross-province leased line + IPSec tunnel
Level 2 branch	Upstream: cross-province dual private line square-shaped Downstream: intra-province dual private line square-shaped	Upstream: cross-province dual private line square-shaped Downstream: intra-province dual private line square-shaped	Upstream: cross-province private line + square-shaped Internet Downstream: square-shaped dual private lines in the province
outlets	Dual private line + 4G/5G backup	Dual private networks	Intra-province dual private line

Typical Underlay Networking (2)

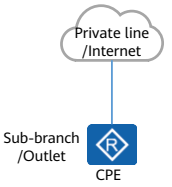
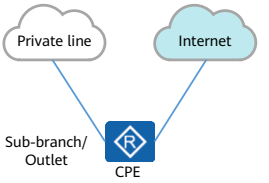
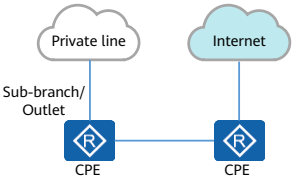


HUAWEI

- On the underlay WAN side, IP addresses and IGP routing protocols are configured based on the networking habits of the live network. Generally, a bank network uses a dynamic routing protocol at the egress to connect to upper-level branch devices.

Scenario	Multi-link flattening	Internet flattening
Level 1 branch	Multi-ISP private line access	Multi-Internet Egress Access
Level 2 branch	None	None
outlets	Dual private line	Single link or dual Internet

Underlay Network Design for Sub-Branches/Outlets

Single-gateway single-homed networking	Single-gateway dual-homed networking	Dual-gateway dual-homed networking
		
• Advantages: low deployment cost • Disadvantages: risks of single device and single link failures	• Advantages: no risk of single link failures, moderate deployment cost, high reliability • Disadvantages: risks of single device failures	• Advantages: no risk of single device or single link failures, high reliability, exclusive link bandwidth occupation • Disadvantages: high deployment cost, a large number of port resources required

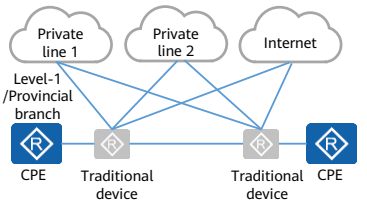
- Financial enterprises are concerned about network reliability but are not sensitive to ICT construction costs. Therefore, the dual-gateway dual-homed networking is recommended.

HUAWEI

- Sub-branches and outlets generally use the dual-homed networking mode.

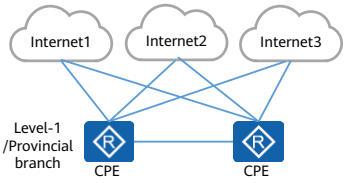
Underlay Network Design for Level-1/Provincial Branches

Networking with dual gateways deployed in off-path mode and multiple private line networks



- Advantages: no risk of single device or single link failures, high reliability, support for smooth SD-WAN network upgrade
- Disadvantages: CPEs need to be connected to traditional devices in off-path mode.
- Applicable to banks

Networking with dual gateways and multiple Internet networks



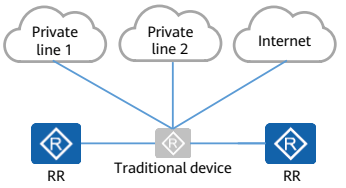
- Advantages: no risk of single device or single link failures, high reliability
- Disadvantages: The SD-WAN network upgrade is complex and requires replacement of original devices.
- Applicable to securities and insurance enterprises

HUAWEI

RR Networking Design

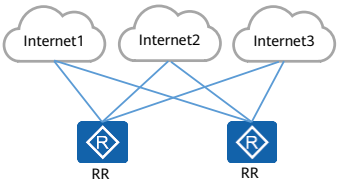
- Financial enterprises, especially banks, have many sub-branches and outlets. To improve the stability of RRs, dual standalone RRs are typically deployed.

Networking with dual standalone RRs deployed in off-path mode



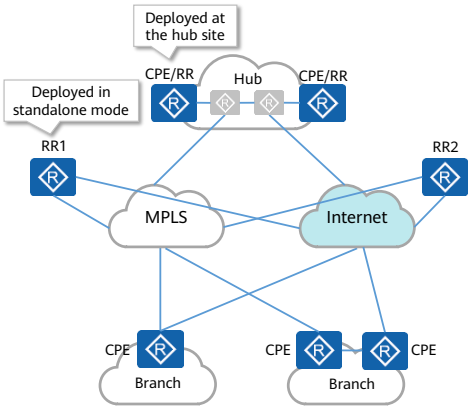
- Advantages: no risk of single device or single link failures, high reliability, support for smooth SD-WAN network upgrade
- Disadvantages: The RRs need to be connected to a traditional device in off-path mode. If CPEs need to connect to the RRs through a public network, 1:1 NAT needs to be configured.
- Applicable to banks

Networking with dual standalone RRs deployed in in-path mode



- Advantages: no risk of single device or single link failures, high reliability
- Disadvantages: The SD-WAN network upgrade is complex and requires replacement of original devices. If CPEs need to connect to the RRs through a public network, 1:1 NAT needs to be configured.
- Applicable to securities and insurance enterprises

RR Deployment Rules



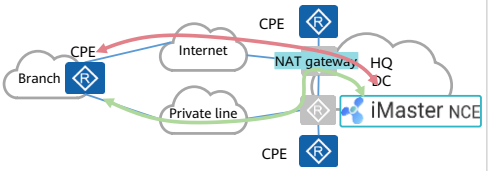
- RRs must be deployed in redundancy mode. At least two RRs must be deployed on the live network.
- Each CPE should be dual-homed to two RRs to implement egress backup.
- It is recommended that RRs be deployed in standalone mode to ensure reliability.
- If standalone RRs cannot be used, configure CPEs at hub or border sites as RRs.
- Use the RR models recommended in the specification list.

- A CPE can be connected to a maximum of two RR sites (four RRs).
- On small-scale networks, for example, with fewer than 50 sites, RRs can be deployed at hub sites.
- RRs need to support large numbers of BGP peers and EVPN connections and provide strong high route reflection capabilities and efficiency. In actual deployments, use the RR models recommended in the specification list, for example, AR6300 and AR6280.

Underlay Network Design with One Controller Deployed

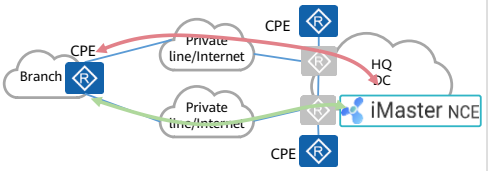
- To ensure the reliability of SD-WAN management channels, dual gateways and dual links need to be deployed to ensure that CPEs/RRs at branches can communicate with iMaster NCE-WAN.

Networking with both public and private networks



- iMaster NCE-WAN is deployed in the HQ DC and uses a public IP address to provide services for CPEs.
- 1:1 static NAT is deployed on the egress devices of the HQ connecting to the Internet.
- NAT-related public IP addresses need to be advertised to the private line network.
- Data between branch CPEs and iMaster NCE-WAN passes through the NAT device no matter whether the branch communicates with the HQ over the Internet or a private line network.

Single-type networking



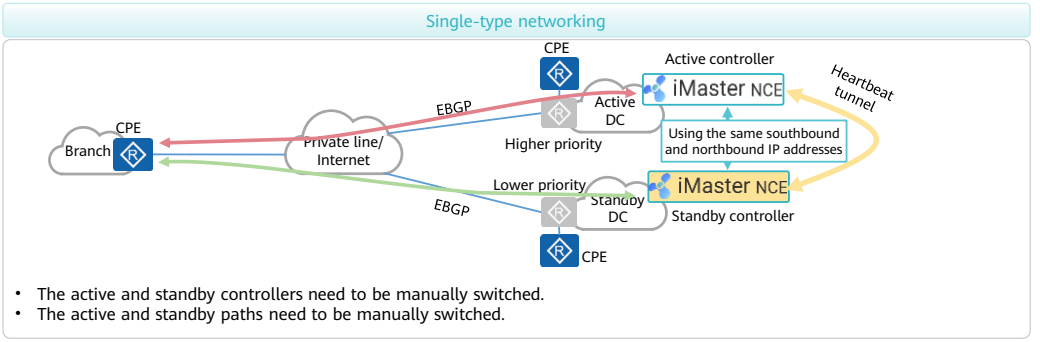
- iMaster NCE-WAN is deployed in the HQ DC.
- If branches communicate with the HQ through a private line network, the network segment where iMaster NCE-WAN resides needs to be sent to the private line network.
- If branches communicate with the HQ through the Internet, 1:1 static NAT needs to be deployed on egress devices of the HQ connecting to the Internet. The egress devices of the HQ use the same public IP address.

HUAWEI

- Generally, 1:1 static NAT is deployed in the system view on the egress device of the public network at the HQ.

Underlay Network Design with Controllers Deployed in Active/Standby Mode

- When iMaster NCE-WAN is deployed in active/standby mode, the route priorities are specified to steer communication traffic between CPEs and the active and standby controllers.



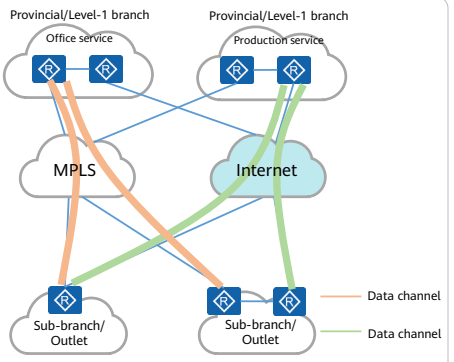
HUAWEI

- The active and standby controllers are deployed in the active and standby DCs, respectively. A heartbeat tunnel is established between the active and standby controllers to synchronize data and detect the controller status.
- The active and standby controllers use the same southbound and northbound IP addresses. When branches are connected to the DCs through a public network, the same NAT address must be configured for the active and standby controllers.
- When branches are connected to the DCs through a private line network, the controllers' southbound and northbound IP addresses need to be advertised to the private line network through EBGP, and a routing policy needs to be configured to ensure that the route to the active controller is preferentially selected.
- When branches are connected to the DCs through a public network, 1:1 static NAT needs to be deployed for the southbound and northbound IP addresses. The NAT configurations must be the same on the two egress gateways. NAT related routes need to be advertised to the public network through EBGP and a routing policy needs to be configured to ensure that NAT related routes to the active DC are preferentially selected.

Contents

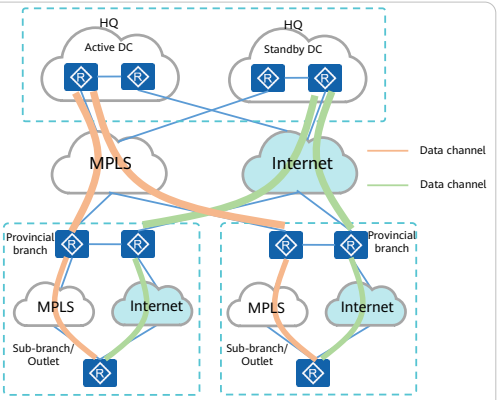
Common Overlay Topologies of Financial SD-WAN Networks

Flat topology



Data tunnels are established directly between sub-branches/outlets and provincial/level-1 branches.

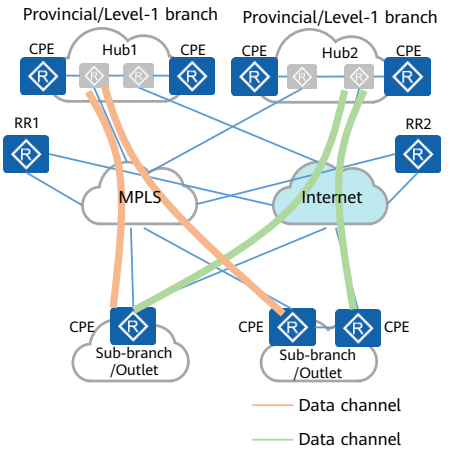
Hierarchical topology



Data channels are established between sub-branches/outlets and provincial/level-1 branches as well as between provincial/level-1 branches and the HQ.

HUAWEI

Flat Topology of Financial SD-WAN Networks

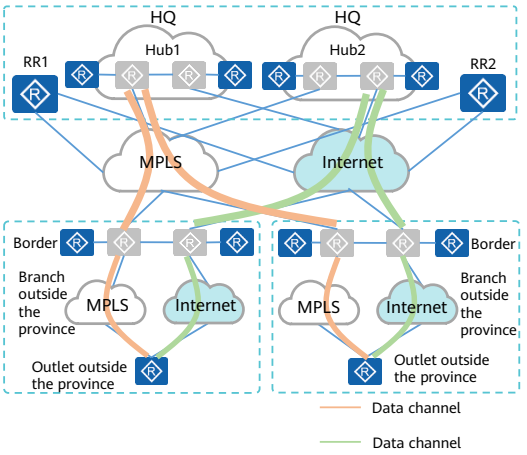


- RR deployment
  - RRs' IP addresses must be advertised to each ISP network, so that the RRs are reachable on the underlay network.
  - RRs can be deployed independently (for high reliability) or at hub sites. The RRs need to work in active/standby mode.
- Hub deployment
  - CPEs at hub sites can be connected to traditional egress routers in off-path mode to support smooth evolution to an SD-WAN network.
  - Two CPEs are deployed at each hub site to ensure intra-site reliability.
  - Hub sites are deployed in active-active mode based on site requirements or services to ensure inter-site reliability.
- For a large-scale network, it is recommended that multiple tenant networks be planned based on administrative areas on the live network and managed independently.

HUAWEI

- The flat topology is applicable to large banks and insurance enterprises. Network O&M is implemented by administrative area, and networks in each province are managed by MSPs.
- Provincial branches, level-1 branches, and hubs at the HQ are connected to all ISP networks.
- The controller is usually deployed in a DC.
- Generally, CPEs at outlets are dual-homed to RRs to ensure reliability.
- North-south traffic is a majority of traffic.
- East-west traffic between outlets needs to pass through the hubs.

Hierarchical Topology of Financial SD-WAN Networks

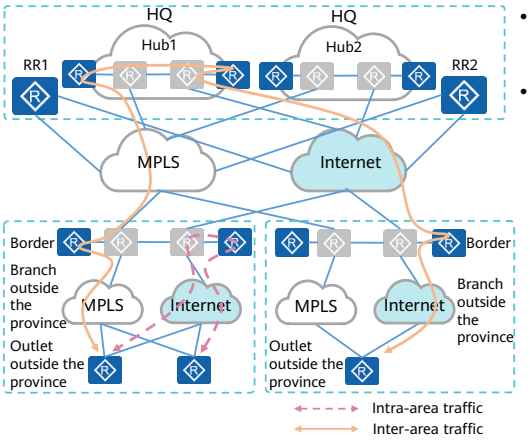


- RR deployment
  - RRs' IP addresses must be advertised to each ISP network, so that the RRs are reachable on the underlay network.
  - RRs can be deployed independently (for high reliability) or at hub sites. The RRs need to work in active/standby mode.
- Hub deployment
  - CPEs at hub sites can be connected to traditional egress routers in off-path mode to support smooth evolution to an SD-WAN network.
  - Two CPEs are deployed at each hub site to ensure intra-site reliability.
  - Hub sites are deployed in active-active mode based on site requirements or services to ensure inter-site reliability.
- Border deployment
  - CPEs at border sites can be connected to traditional egress routers in off-path mode to support smooth evolution to an SD-WAN network.
  - CPEs at border sites connect downstream devices to upstream devices. The one-way forwarding performance of CPEs depends on the total bandwidth of the downstream outlets and the total bandwidth of the upstream hub sites.

HUAWEI

- The hierarchical topology is applicable to small or midsize rural commercial banks. Outlets outside the province are connected to the HQ through the Internet, while those in the province are directly connected to the HQ.
- Provincial branches, level-1 branches, and hubs at the HQ are connected to all ISP networks.
- The controller is usually deployed in a DC.
- Outlets are not directly connected to level-1 branches or HQ. Instead, traffic of outlets is aggregated to level-2 branches or branches outside the province and then sent to level-1 branches or HQ.
- Branches outside the province are interconnected with level-1 branches through private lines.
- Sub-branches and outlets are usually connected to upper-level branches through multiple uplinks to achieve high reliability.
- North-south traffic is a majority of traffic.
- East-west traffic between outlets needs to pass through border sites in the province or hub sites of the HQ.

Hierarchical Topology of Financial SD-WAN Networks: Inter-Site Communication

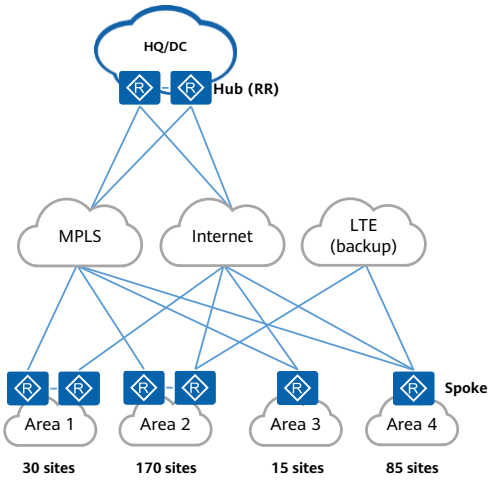


- Access requests of all branches in an area are forwarded by the border nodes in the area.
- Branches in different areas communicate with each other through the active hub of the HQ.

HUAWEI

- Inter-area traffic is forwarded as follows:
  - Traffic of an outlet is first forwarded to the border nodes of the area where the outlet resides.
  - The border nodes forward the traffic to the active hub of the HQ based on the configured traffic steering policy.
  - The hub then forwards the traffic to the border nodes in the destination area based on the routing policy.
  - The border nodes in the destination area search local routing tables for routes and then forward the traffic to the destination outlet.

Specification Calculation for Financial SD-WAN Networks



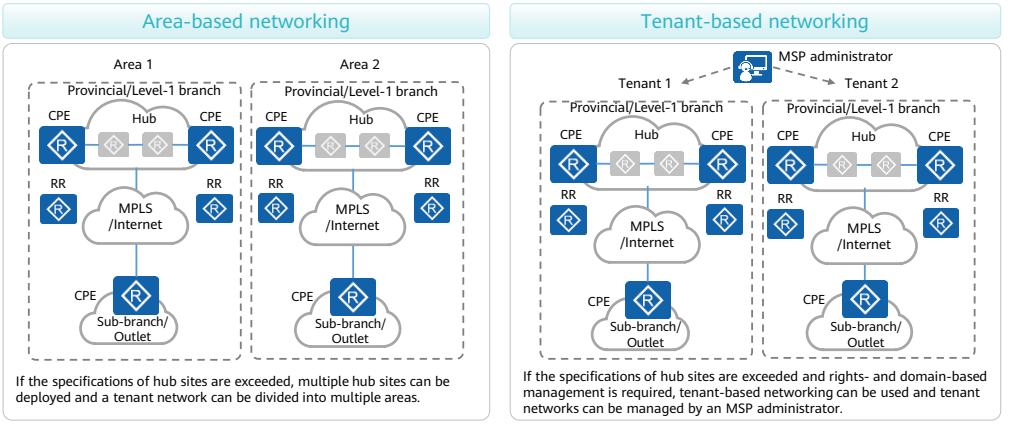
- Background
  - Two hubs are deployed at the HQ of a bank and also function as RRs. The bank has a total of 300 sites in different areas, including 200 dual-gateway sites and 100 single-gateway sites.
  - Each hub (RR) supports a maximum of 1000 BGP peers and 2000 data tunnels.
  - LTE links are used as backup. When primary links are available, no LTE link is established.
- Calculation of networking specifications
  - Number of BGP peers supported by the hubs (RRs) = Number of dual-gateway sites x 2 + Number of single-gateway sites
    - In this example, the total number of BGP peers is 500 (200 x 2 + 100), which meets the specification.
  - Number of data tunnels supported by the hubs (RRs) = Total number of data tunnels established between the hubs and gateways
    - The two hubs share the data tunnel specification. For example, if the data tunnel specification of each hub is 1000, the data tunnel specifications of the two hubs is 1000.
    - In this example, the total number of data tunnels is 1200 (300 x 4 = 1200), which meets the specification.
  - The network bandwidth must meet the customer requirements.

HUAWEI

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

Beyond-Specification Design for Financial SD-WAN Networks

- A financial network generally has a large number of branch nodes, and the requirements of a financial enterprise may beyond the specifications of devices at hub sites regardless in the flat or hierarchical topology. There are two solutions to this problem.



HUAWEI

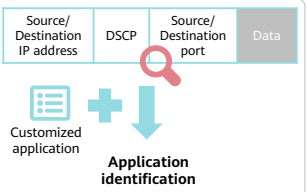
- Area-based networking
  - A tenant network is divided into multiple areas, multiple hub sites are deployed in the HQ/DC, and each area is associated with one or two hub sites.
  - Branch sites are added to the corresponding hub sites based on areas.
  - RRs can be deployed independently, and each pair of RRs is associated with sites in the corresponding area.
  - Sites in different areas are interconnected through hub sites on the LAN side.
- Tenant-based networking
  - An MSP administrator creates multiple tenants, multiple hub sites are deployed in the HQ/DC, and each tenant is associated with one or two hub sites.
  - Branch sites are added to the corresponding tenant tenants based on geographical areas.
  - RRs can be deployed independently. Each pair of RRs is associated with sites in an area.
  - Sites on different tenant networks are interconnected through hub sites on the LAN side.

Contents

Service Transport Design

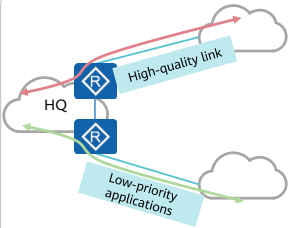
- Financial enterprises generally deploy various types of services, which have different network requirements.
- Most services of financial enterprises are private network services. Therefore, general feature databases cannot be used to identify financial services.
- The financial SD-WAN service transport design covers the following aspects.

Application identification



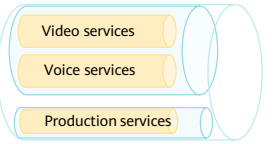
- Application identification based on 5-tuple information or DSCP values
- Identification of customized applications

Intelligent traffic steering



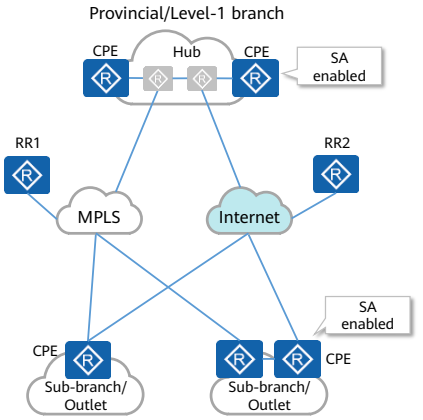
- Link quality–based traffic steering
- Application priority–based traffic steering

HQoS



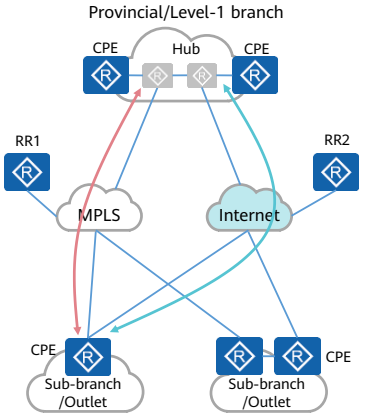
- Service-based HQoS

Application Identification Design



- The main objective of application identification is to distinguish traffic for subsequent processing.
- The service awareness (SA) function is enabled on the devices at HQ sites and outlets to implement fine-grained identification of services on the live network.
- Application identification based on 5-tuple information or DSCP values
  - Applications of traffic with specified source and destination IP addresses can be identified based on 5-tuple information or DSCP values.
  - Generally, traffic classifiers are configured to identify such applications.
- Identification of customized applications
  - Enterprises that need to monitor customized applications can use this application identification mode.
  - Customized applications are defined based on the destination IP address, protocol number, and signature.

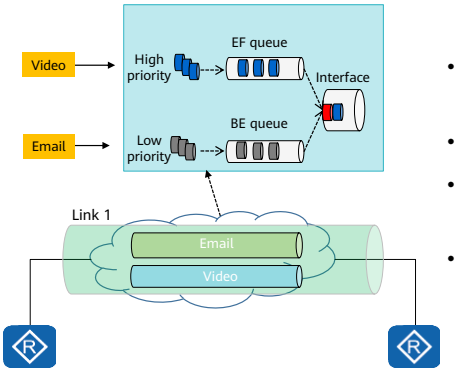
Intelligent Traffic Steering Design



- In the SD-WAN Solution, intelligent traffic steering ensures optimal experience of different services.
- With this function enabled, the network quality can be monitored in real time, and the network paths that meet SLA requirements of applications can be dynamically and automatically selected among multiple WAN links with different network quality.
- Link quality–based traffic steering
  - This mode is applicable when enterprises have different link quality requirements for different services.
  - Multiple traffic steering policies can be configured to enable transmission of different services over different links.
- Application priority–based traffic steering
  - This mode is applicable when enterprises want to use high-value lines (private lines) to guarantee experience of high-value applications.

HUAWEI

HQoS Design



- Bank networks have various applications, including voice, video conferencing, file transfer, email, and software as a service (SaaS) applications.
- Different applications have different link quality requirements. Therefore, different HQoS policies need to be deployed.
- Traffic is differentiated by application.
- High priorities can be set for preferential scheduling of delay-sensitive or mission-critical traffic.
- Larger bandwidth is allocated to bandwidth-demanding traffic.

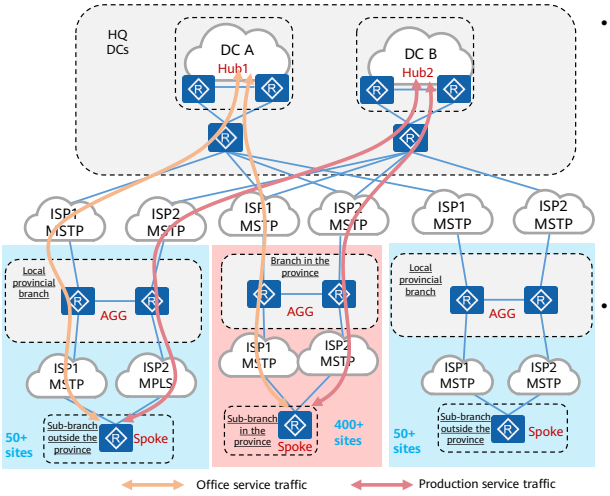
Contents

Project Requirements of a Financial Enterprise

- A financial enterprise deploys the SD-WAN Solution to meet the following requirements:
  - Replace existing MSTP lines with MPLS or Internet lines.
  - Implement efficient intelligent traffic steering at the egress.
  - Simplify network management and O&M through email-based deployment.
  - Provide high link-level, device-level, and inter-site reliability.

Current network environment	Two DCs are built in the same city and work in active-active mode to provide services concurrently. The DCs are interconnected at Layer 3. Each equipment room is connected to the Internet through a single private line.
Current network services	• Production, office, and video surveillance service traffic exists on the live network. **Production traffic between outlets needs to be isolated from other types of traffic.** • **Uplink traffic is load balanced based on service types, and Internet access sites are strictly specified for downlink traffic based on service types.** For example, office service traffic of an outlet is transmitted over link 1 and terminated at hub 1, and hub 2 functions as the standby node. Production service traffic is transmitted over link 2 and terminated at hub 2, and hub 1 functions as the standby node.
Customer requirements	Unified network management is required to lower O&M costs.

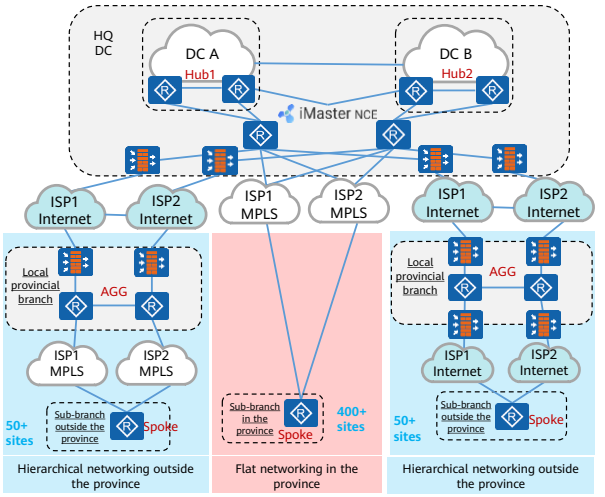
Current Network Architecture of the Financial Enterprise



- Current network architecture
  - The financial enterprise has 400+ sub-branches in the province and 100+ sub-branches outside the province. The sub-branches use the single-device or dual-device dual-uplink networking mode at egresses.
  - Sub-branches in the province are connected to branches in the province through MSTP private lines, which are connected to the HQ through MSTP private lines.
  - Sub-branches outside the province are connected to local provincial branches through MSTP private lines, which are connected to the HQ through MSTP private lines.
- Current traffic model
  - Traffic of all sub-branches in and outside the province is first aggregated to branches and then centrally forwarded to the HQ.
  - Office service traffic of sub-branches in and outside the province is sent to DC A.
  - Production service traffic of sub-branches in and outside the province is sent to DC B.

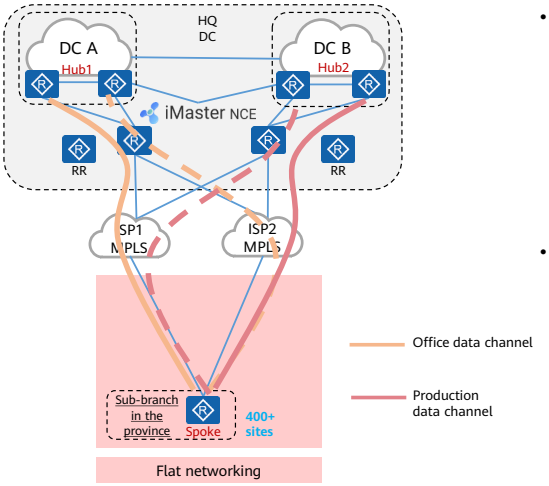
HUAWEI

Overall SD-WAN Network Reconstruction Design



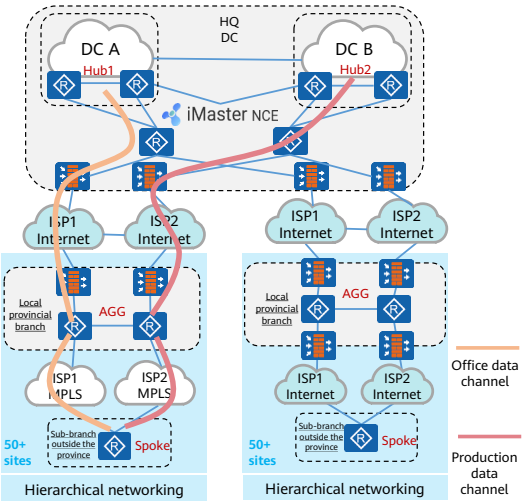
- Underlay network design
  - The single-device or dual-device dual-uplink networking mode is used.
  - RRs are deployed independently, and iMaster NCE-WAN is deployed in active/standby mode.
  - Sub-branches in the province are connected to the HQ through MPLS private lines.
  - Sub-branches outside the province are connected to local provincial branches through MPLS private lines or the Internet, which are connected to the HQ through the Internet.
- Overlay network design
  - Two VNs are planned based on services to carry office and production services respectively.
  - Sub-branches in the province directly establish tunnels with the HQ, implementing flat networking.
  - Sub-branches outside the province establish data tunnels with local provincial branches, which establish data tunnels with the HQ, implementing hierarchical networking.

SD-WAN Network Design in the Province



- Underlay network design
  - The financial enterprise expects to reduce O&M costs and centrally manage branch networks.
  - MPLS private lines in the province have a moderate price and better performance than Internet links. Therefore, MPLS private lines are used to replace MSTP lines to reduce line costs.
  - Sub-branches in the province use the flat networking and are directly connected to the HQ through MPLS private lines.
- Overlay network design
  - After the flat networking is implemented for sub-branches in the province, the sub-branches send traffic directly to the HQ and establish data channels directly with the HQ.
  - The overlay network uses the flat topology.
  - Dual uplinks of each device at a sub-branch back up each other, providing high reliability for office and production services.

SD-WAN Network Design Outside the Province



- Underlay network design
  - Cross-province MPLS and MSTP private lines are expensive. Therefore, the Internet is used instead of MSTP private lines to reduce line costs.
  - The prices of the Internet and MPLS private lines are similar in local provinces. Links are selected based on the site requirements.
  - For service security purposes, firewalls must be deployed at both ends of each Internet link.
- Overlay network design
  - Sub-branches outside the province do not use the flat networking. Therefore, traffic of these sub-branches is first aggregated to local provincial branches and then centrally sent to the HQ.
  - The overlay network uses a hierarchical topology.
  - Dual uplinks of each device at a sub-branch back up each other, providing high reliability for office and production services.

HUAWEI

SD-WAN Device Specification Design

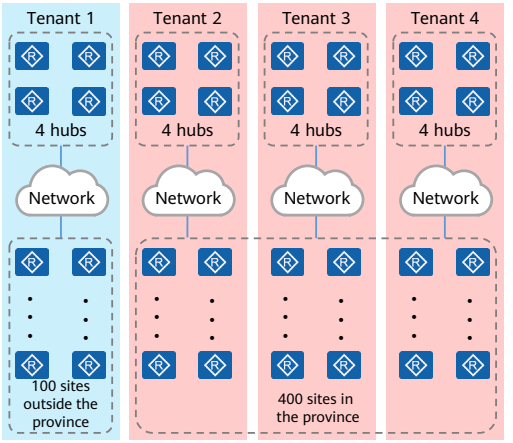
- The financial enterprise has 500 sites, including 350 dual-gateway sites and 150 single-gateway sites. Each site requires 10 Mbit/s bandwidth.
- The numbers of BGP peers and tunnels are calculated as follows:
  - Number of BGP peers at RR sites: 850 (350 x 2 + 150). It is recommended that AR6280 or AR6300 series routers be used as RRs and two RR sites (with two RRs each) be deployed in active/standby mode.
  - Number of BGP peers at branch sites: 4 (two RRs at the active and standby RR sites each). AR650 series routers are recommended as RRs at small-sized sites and AR6100 series routers at medium-sized sites.
  - Number of data tunnels at hub sites: 2000 (500 x 4), with two hubs at the active and standby hub sites each. **AR6280 or AR6300 series routers equipped with the SRU400H are recommended as the hubs.**
- The traffic specification is calculated as follows:
  - One-way traffic at hub sites: 5000 Mbit/s (500 x 10 Mbit/s). The number of required hubs is calculated based on the tunnel forwarding capability of devices. If AR6280 or AR6300 series routers equipped with the SRU400H are used as hubs, **4 groups of hubs are needed**. If each group has 4 hubs, 16 hubs are required.

Note: The specifications apply only to the SD-WAN Solution V100R019C00.

HUAWEI

- AR6280 and AR6300 series routers support a maximum of 1000 BGP peers.
- AR6280 and AR6300 series routers equipped with the SRU400H support a maximum of 3000 data tunnels. The maximum one-way bandwidth of each data tunnel is 1.5 Gbit/s.
- For details about the device specifications, contact Huawei technical engineers.

Beyond-Specification SD-WAN Network Design



- Limited by device specifications, the financial enterprise needs to deploy four groups of hubs.
- To facilitate management and distribute traffic to the four groups of hubs, four tenants are planned on the SD-WAN network, and each tenant has different VNs.
- The management scope of the four tenants is planned as follows:
  - Tenant 1: manages a total of 100 sites outside the province.
  - Tenants 2, 3, and 4: manage a total of 400 sites in the province based on the site requirements.

HUAWEI

- The requirements of the financial enterprise are beyond the specifications of a single device. To address this issue, use the tenant-based networking as the financial enterprise has branches outside the province and requires rights-based management.

SD-WAN Service Transport Design

- The financial enterprise has two main types of services:
  - Delay-sensitive, mission-critical production services (such as transaction services): requiring 6 Mbit/s bandwidth
  - Non-production services, including:
    - Delay-sensitive services (such as video security services): requiring 3 Mbit/s bandwidth
    - Delay-insensitive services (such as office services): no requirement on assured bandwidth
- Service transport design

Service Category	Application Classification Mode	Intelligent Traffic Steering Configuration	QoS Configuration
Production services	Application classification based on source and destination IP addresses	The **Prefer** scheduling mode is used. Higher-priority link 1 is preferentially selected, and **Low Latency Data** is selected as the switchover condition.	The highest-priority queue is specified, and 6 Mbit/s bandwidth is allocated.
Delay-sensitive non-production services	Application classification based on source and destination IP addresses	The **Prefer** scheduling mode is used. Higher-priority link 1 is preferentially selected, and **Real-Time Video** is selected as the switchover condition.	A high-priority queue is specified, and 3 Mbit/s bandwidth is allocated.
Delay-insensitive non-production services	Application differentiation based on source and destination IP addresses	The **Prefer** scheduling mode is used. Lower-priority link 2 is preferentially selected, and **Bulk Data** is selected as the switchover condition.	No queue is specified and no bandwidth is allocated.

- The maximum egress bandwidth of each branch is 10 Mbit/s.
- If no queue is specified for a service, service traffic enters the BE queue.

Quiz

1. (Single-answer question) Which of the following underlay networking modes is recommended for sub-branches?
   A. Single-gateway single-homed networking
   B. Single-gateway dual-homed networking
   C. Dual-gateway dual-homed networking
   D. Three-gateway triple-homed networking

- 1. C

Summary

- This course describes the SD-WAN network design for the FSI, including underlay network design, overlay network design, service transport design, and reliability design.
- Typical underlay networking modes for the FSI include dual-homed networking with dual private line networks, single-homed networking with dual private line networks, networking with both private line and Internet networks, and multi-link flat networking.
- Typical overlay networking modes for the FSI include flat networking and hierarchical networking.

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.