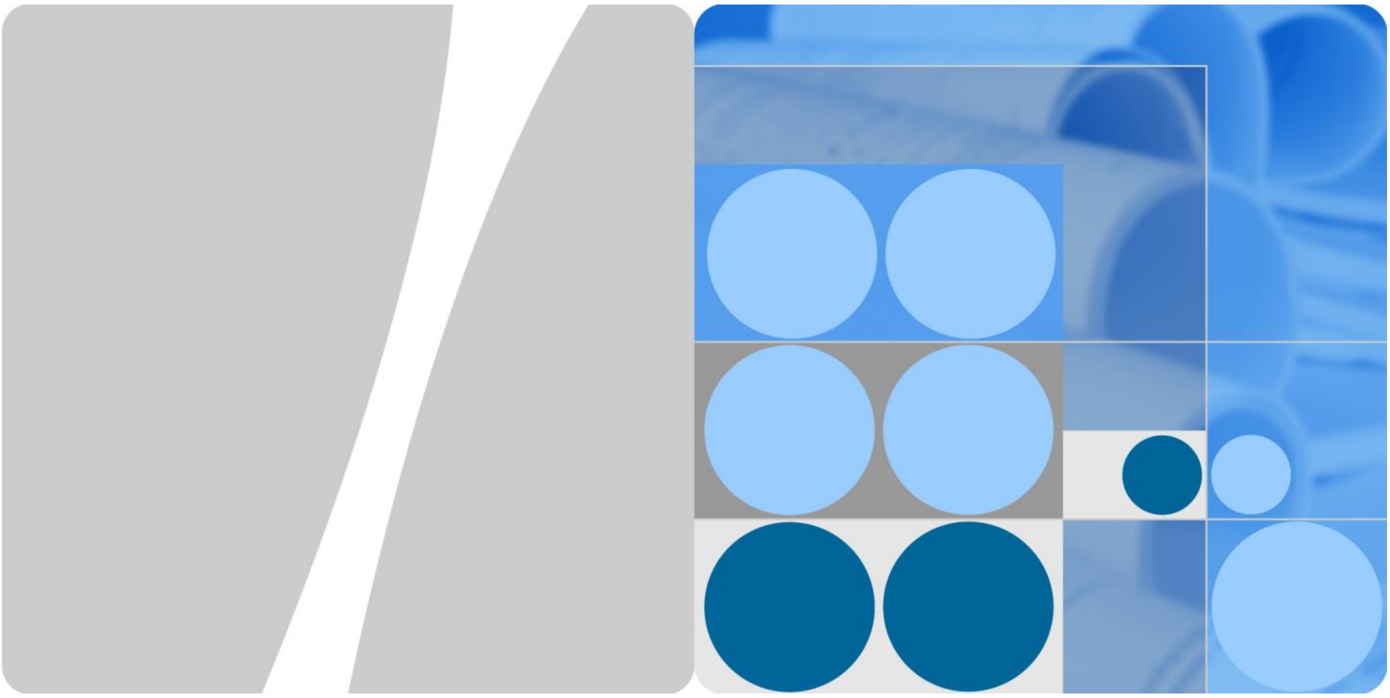


Huawei CloudFabric DCN

O&M Technology White Paper

Issue	**01**
Date	**2017-04-14**

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

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

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

Huawei Technologies Co., Ltd.

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

Website: http://www.huawei.com

Email: support@huawei.com

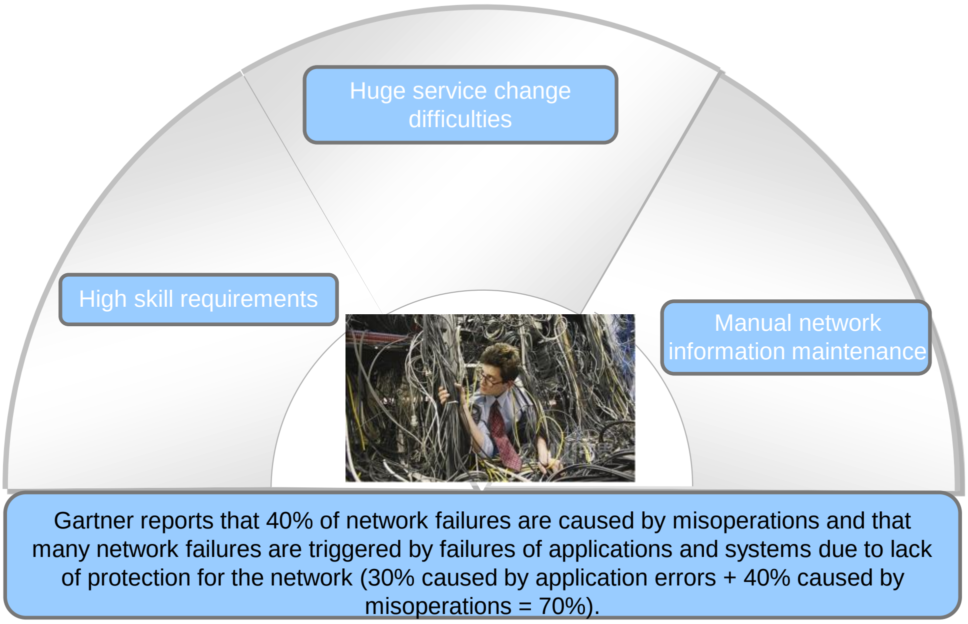
Contents

1 DCN O&M Overview

1.1 O&M Issues Faced by a Traditional DCN

In traditional data centers, computing and storage resources are fixed, so the network and traffic are also fixed accordingly. Data center networks (DCNs) are manually configured by network administrators, which results in a huge workload and is error-prone. It is difficult to locate and rectify network faults.

Figure 1-1 O&M issues faced by a traditional DCN


Huge service change difficulties
High skill requirements
Manual network information maintenance

Gartner reports that 40% of network failures are caused by misoperations and that many network failures are triggered by failures of applications and systems due to lack of protection for the network (30% caused by application errors + 40% caused by misoperations = 70%).

- High skills requirements: Network devices are provided by multiple vendors and are of various models; thousands of commands are involved; parameter settings are complex.
- Complex policy update and maintenance: Policies are manually configured, which are difficult to modify and are error-prone.
- Network information maintained manually: Application maintenance personnel maintain the mapping between applications and IP addresses, whereas network maintenance

personnel maintain IP addresses and network policies. Inter-department manual information maintenance causes complex communication and O&M.

1.2 O&M Challenges of Cloud DCNs

A cloud DCN is characterized by SDN technology. Compared with traditional DCNs, SDN-based DCNs have the following characteristics:

- Dynamic network: Logical networks can be created and deleted based on application requirements.
- Real-time response: The traditional network design is based on a minute-level slow response mechanism, for example, the SNMP mechanism used for decades of years. The slow mechanism cannot cope with SDN requirements for a high speed.
- Large scale: An SDN-based DCN has a large number of devices to be managed and faults to be processed. For example, the number of vSwitches or vRouters is 50 times the number of physical NEs.

An SDN-based DCN is composed of two layers: overlay network that carries service traffic and underlay network that supports the service traffic. The underlay networking is fixed, whereas the overlay networking is flexible. Hierarchical network technologies enable flexible and automatic service deployment and modification. However, due to the increasingly large network scale, unclear O&M responsibilities, and high skill requirements, cloud DCNs are faced by the following challenges:

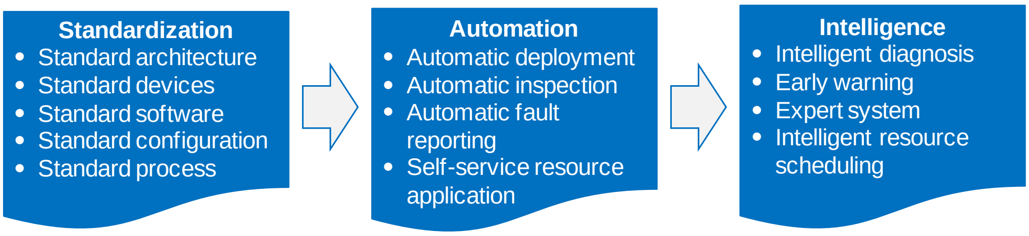
- High device virtualization: The underlay network involves numerous network devices, including egress routers of various models from different vendors and switches, firewalls, and load balancers at various layers. These physical devices use virtualization technologies to isolate logic for tenants. The status of resources on physical and logical devices and the mapping of resources to tenants must be accurately presented to implement precise O&M.
- Visualized O&M: The SDN uses the overlay technology to implement network virtualization. The underlay network uses equal-cost routes to improve the bandwidth and reliability. Therefore, there are multiple equal-cost paths for load balancing traffic. When a fault occurs, the service forwarding path cannot be quickly and accurately located.
- Resource location: After cloud-based construction of a data center, virtual machines (VMs) or physical machines are frequently created, migrated, or released. It is difficult to obtain locations of the workloads bearing services. The workloads are disassociated from physical and logical networks.
- Complex service configuration: Basic networks are manually configured or configured on the NMS in batches, and the service configuration is automatically delivered by the controller. After a fault occurs, maintenance personnel cannot quickly determine whether the cause is a manual configuration problem or an automatic configuration problem.
- Low fault diagnosis efficiency: In one aspect, network device faults and server faults are hard to demarcate. In another aspect, it takes a long time to diagnose faults between the cloud platform, controller, and network devices, and the network needs to prove its innocence.

1.3 DCN O&M Trends

Low cost, high efficiency, and high quality are three objectives of O&M. To achieve these objectives, take the following measures:

- Standardize architectures, devices, software, configurations, and management to reduce O&M objects, simplifying implementation of automation.
- Develop tools for standardized construction and maintenance scenarios to significantly improve the O&M efficiency and reduce the O&M cost.
- Introduce technologies such as machine learning, expert system, and Big Data mining for early warning and fault diagnosis, implementing intelligent O&M.

Figure 1-2 DCN O&M trends



New technologies such as cloud computing, SDN, and Big Data pose increasing difficulties for O&M management of IT infrastructures, running environments, and application systems. O&M level varies significantly by industry. The Internet industry features changing service requirements, large amounts of O&M objects, and complex IT infrastructures of data centers, and therefore dominates the data center O&M evolution direction. The DCN O&M evolution direction recognized by the industry is as follows: standardization > automation > intelligence.

2 Huawei CloudFabric DCN O&M Solution

2.1 Overview

Based on O&M features of cloud DCNs, Huawei CloudFabric DCN Solution provides various features that are developed from the perspectives of network administrators and tenant administrators and are oriented to the entire lifecycle. The lifecycle involves installation and deployment, service provisioning, fault demarcation and location, status monitoring, and network change. Table 2-1 lists the O&M features provided for various DCN O&M phases. These O&M features enable customers to build simple, efficient, open, and visualized DCNs and boost the development of cloud data center services.

Table 2-1 O&M features provided for various DCN O&M phases

Role	Installation and Deployment	Service Provisioning	Fault Demarcation and Locating	Status Monitoring	Network Change
Network administrator	• Zero Touch Provisioning (ZTP) • Management of devices by the Agile Controller-DCN • Device replacement • Capacity expansion	• Resource pool management (IP addresses, VLANs, ports, and elastic scaling) • Interconnection to traditional VLAN networks • Interconnection to PEs • Underlay network connectivity configuration • Advanced network service (dynamic routing protocol)	• Three-layer topology visibility • End port locating • Path detection • Connectivity check • Loop detection • Packet Conservation Algorithm for Internet (iPCA) • Port mirroring	• Traffic monitoring • Alarm monitoring • Performance monitoring • Log monitoring • Capacity monitoring	• Configuration backup and restoration • Device version upgrade and patch installation

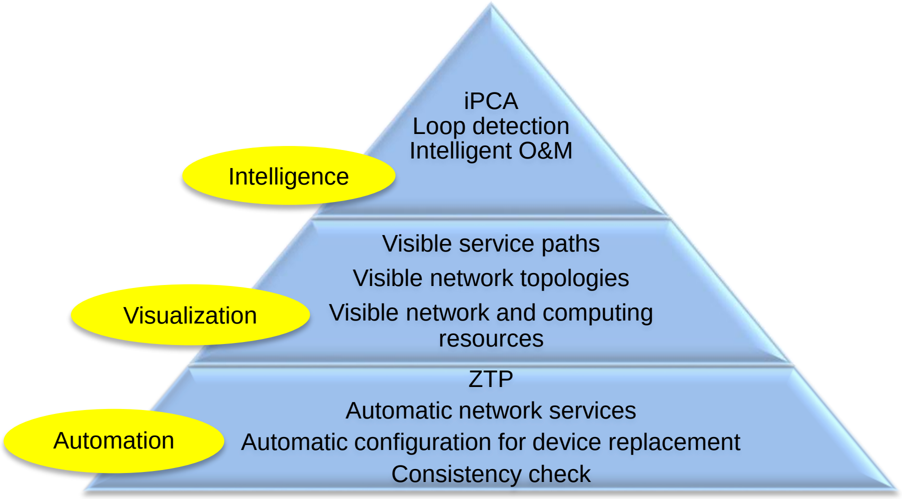
Role	Installation and Deployment	Service Provisioning	Fault Demarcation and Locating	Status Monitoring	Network Change
		• Firewall policy change			
Tenant administrator	-	• Layer 2 network • Layer 3 route • Layer 4 security rule • Layer 7 load balancing • Associated network calculation	-	• Traffic monitoring • Performance monitoring • Alarm monitoring	-

Except for service provisioning that is independent of O&M, network installation and deployment, fault diagnosis, quality monitoring, and upgrade and capacity expansion are involved in O&M. Balancing the O&M cost, efficiency, and quality has been a puzzle of the O&M field. The evolution direction ranging from standardization to automation, and then intelligence provides the roadmap for resolving the O&M puzzle. Huawei CloudFabric DCN Solution employs a large number of practices and explorations in O&M automation and intelligence.

2.2 Objectives

The overall O&M objectives of Huawei CloudFabric DCN Solution are automation, visualization, and intelligence, as shown in Figure 2-1.

Figure 2-1 Overall solution objectives


iPCA
Loop detection
Intelligent O&M
Intelligence
Visible service paths
Visible network topologies
Visible network and computing resources
Visualization
ZTP
Automatic network services
Automatic configuration for device replacement
Consistency check
Automation

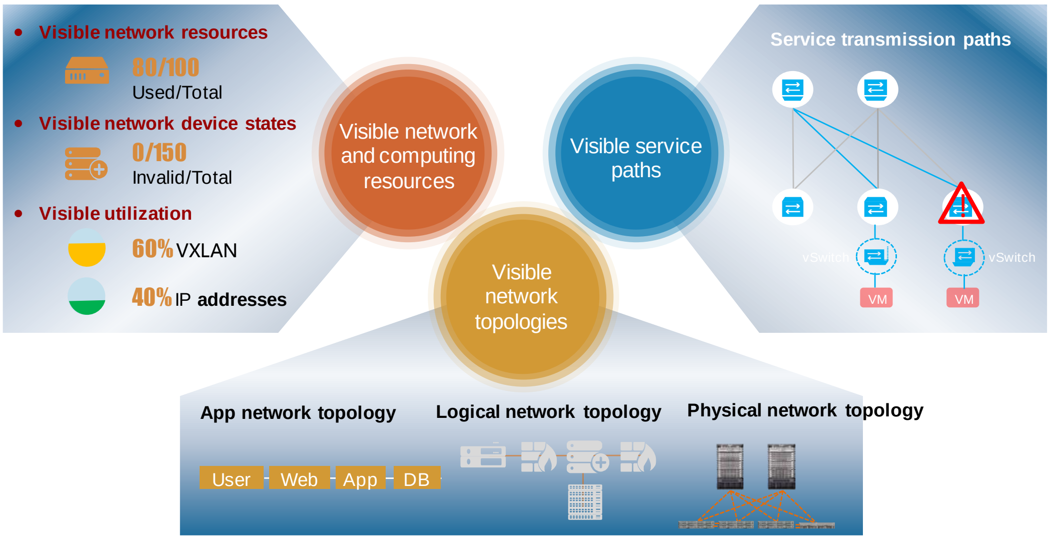
2.2.1 Automation

Automatic O&M resolves the issues of huge network service configuration workload and long online duration. Huawei CloudFabric DCN Solution provides the following functions to implement automatic O&M:

- ZTP: provides automatic installation and deployment functions to improve the deployment efficiency; scans and manages devices in batches to minimize the service provisioning preparation duration.
- Network service automation: automatically delivers Layer 2 and Layer 3 network configuration, and automatically connects and configures Layer 4 through Layer 7 value-added services (VASs).
- Automatic configuration for device replacement: automatically delivers device configurations during device replacement.
- Consistency check: automatically checks the consistency between configurations of the connection between the SDN controller and the cloud platform in the northbound direction and those of the connection between the SDN controller and devices in the southbound direction.

2.2.2 Visualization

Figure 2-2 Visualization



Visualized O&M facilitates network fault demarcation and diagnosis. Huawei CloudFabric DCN Solution provides the following functions to implement visualized O&M:

- Network and computing resource visualization: visualizes information such as the network-wide resources, logical network resources, device resources, and compute node locations.
- Network topology visualization: visualizes application network topologies, logical network topologies, and physical network topologies, and implements three-layer topology visibility.
- Service path visualization: presents actual service forwarding paths and monitors the equal-cost multiple paths between VTEPs, service connectivity (MAC/IP address ping and IP trace), and loops on network edges.

2.2.3 Intelligence

Intelligent O&M is the ultimate goal of O&M. Huawei CloudFabric DCN Solution provides the following intelligent O&M functions:

- Loop detection: intelligently detects loops on networks, reducing the impact of loops on the service network.
- iPCA: intelligently calculates the network service quality. Unlike traditional detection solutions, iPCA detects network quality based on actual service flows. The detection result and fault diagnosis result are more accurate.
- Intelligent O&M: uses FabricInsight to collect traffic data and implement Big Data analysis to support functions such as traffic trend prediction, network fault self-healing, and network performance optimization.

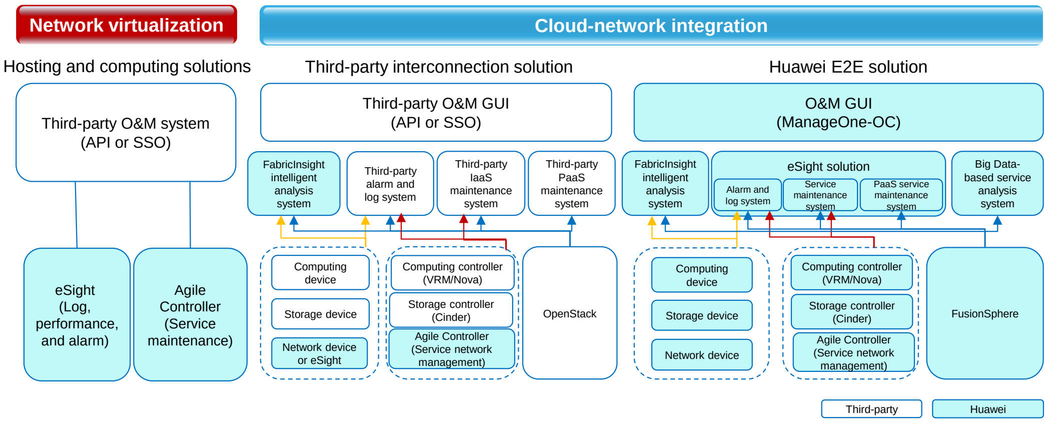
2.3 Architecture

Huawei CloudFabric DCN Solution includes the following O&M components:

- eSight: provides capabilities of managing independent NEs in data centers, including deployment management, upgrade management, alarm management, performance management, and fault management. eSight focuses on the underlay network.
- Agile Controller-DCN: provides service provisioning and O&M capabilities for data centers, including service provisioning, location query, overlay path detection, connectivity detection, real-time traffic reading, cloud service recovery, and visibility of application, logical, and physical networks.
- FabricInsight: enables proactive O&M and automatically assesses the impact of network faults on services through correlation analysis of applications and networks based on Big Data as well as high-speed collection and long-term storage of all traffic data.

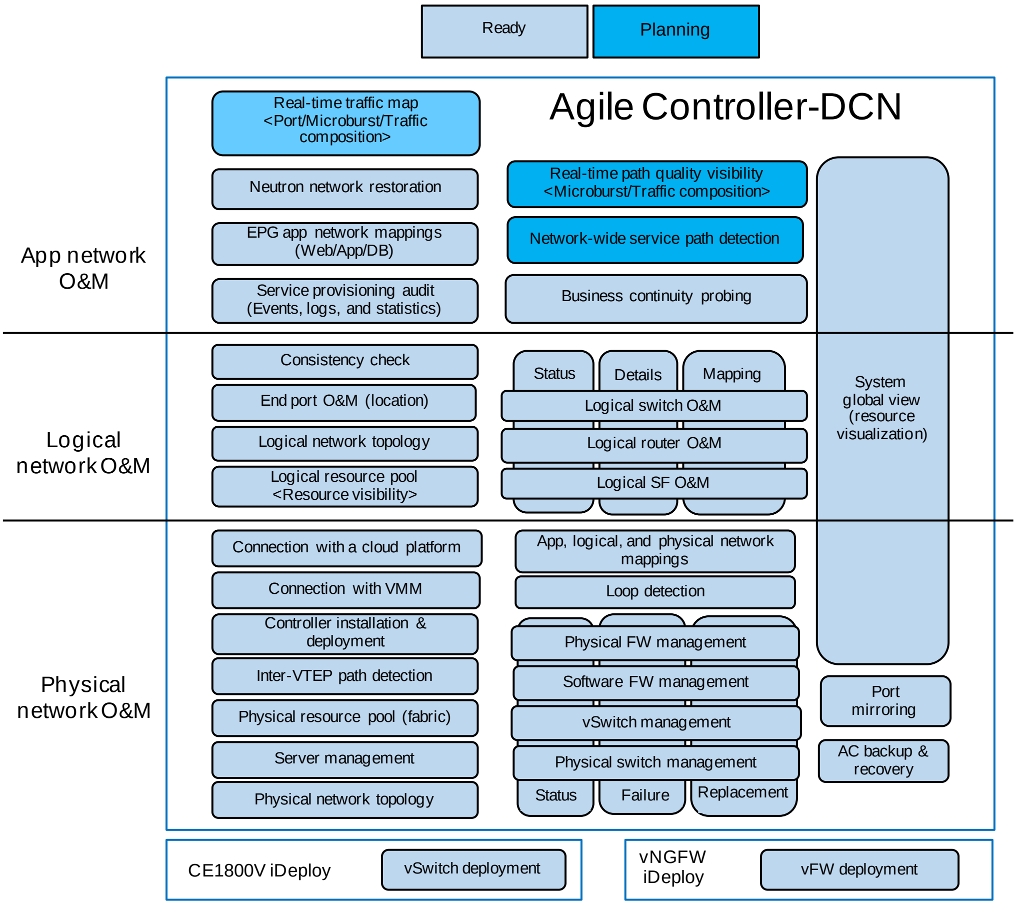
0 shows the O&M interconnection panorama of Huawei CloudFabric DCN Solution. In network virtualization and cloud-network integration scenarios, the Agile Controller-DCN, eSight, and FabricInsight can be seamlessly connected to both Huawei and third-party O&M systems.

Figure 2-3 O&M interconnection panorama of Huawei CloudFabric DCN Solution


Network virtualization
Cloud-network integration
Hosting and computing solutions
Third-party interconnection solution
Huawei E2E solution
Third-party O&M system (API or SSO)
eSight (Log, performance, and alarm)
Agile Controller (Service maintenance)
Third-party O&M GUI (API or SSO)
FabricInsight intelligent analysis system
Third-party alarm and log system
Third-party IaaS maintenance system
Third-party PaaS maintenance system
Computing device
Storage device
Network device or eSight
Computing controller (VRM/Nova)
Storage controller (Cinder)
Agile Controller (Service network management)
OpenStack
O&M GUI (ManageOne-OC)
FabricInsight intelligent analysis system
eSight solution
Alarm and log system
Service maintenance system
PaaS service maintenance system
Big Data-based service analysis system
Computing device
Storage device
Network device
Computing controller (VRM/Nova)
Storage controller (Cinder)
Agile Controller (Service network management)
FusionSphere
Third-party
Huawei

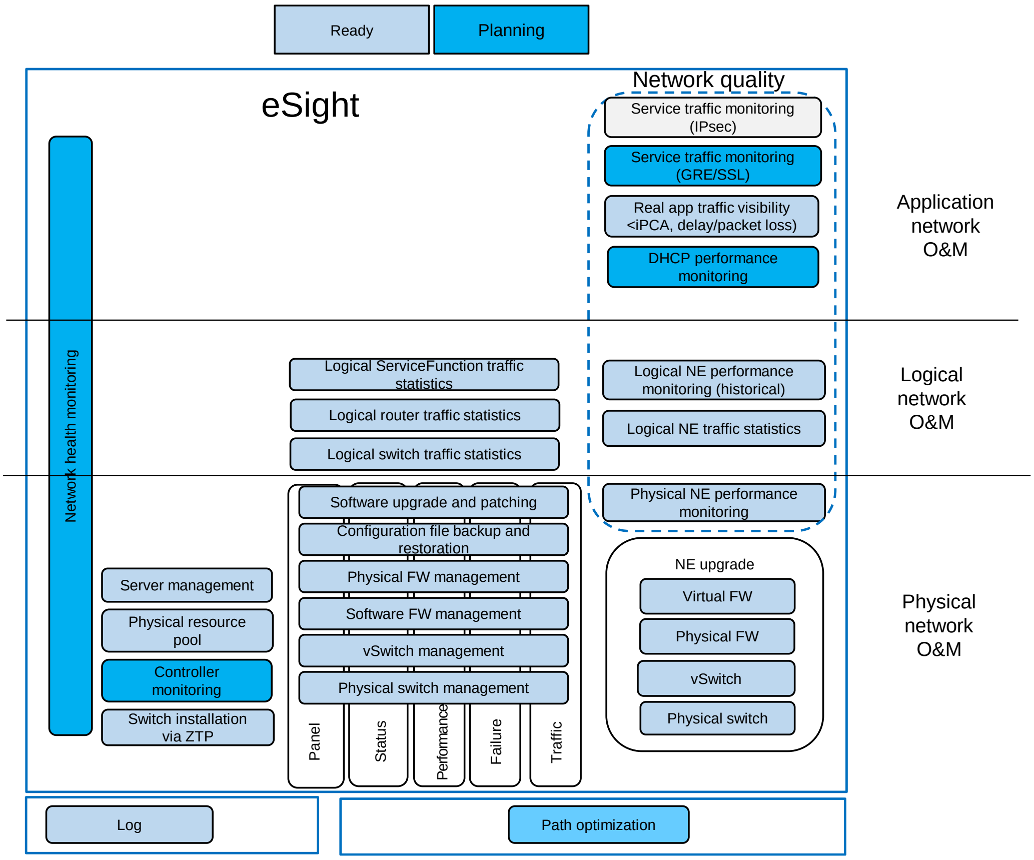
Through combination of eSight, Agile Controller-DCN, and FabricInsight, Huawei CloudFabric DCN Solution provides intelligent O&M that covers the entire data center lifecycle for the underlay and overlay networks in a collaborated manner. The following figures show the O&M functions of eSight, Agile Controller-DCN, and FabricInsight.

Figure 2-4 O&M functions of the Agile Controller-DCN


Ready
Planning
Agile Controller-DCN
App network O&M
Logical network O&M
Physical network O&M
Real-time traffic map <Port/Microburst/Traffic composition>
Neutron network restoration
EPG app network mappings (Web/App/DB)
Service provisioning audit (Events, logs, and statistics)
Real-time path quality visibility <Microburst/Traffic composition>
Network-wide service path detection
Business continuity probing
System global view (resource visualization)
Consistency check
End port O&M (location)
Logical network topology
Logical resource pool <Resource visibility>
Status
Details
Mapping
Logical switch O&M
Logical router O&M
Logical SF O&M
Connection with a cloud platform
Connection with VMM
Controller installation & deployment
Inter-VTEP path detection
Physical resource pool (fabric)
Server management
Physical network topology
App, logical, and physical network mappings
Loop detection
Physical FW management
Software FW management
vSwitch management
Physical switch management
Status
Failure
Replacement
Port mirroring
AC backup & recovery
CE1800V iDeploy
vSwitch deployment
vNGFW iDeploy
vFW deployment

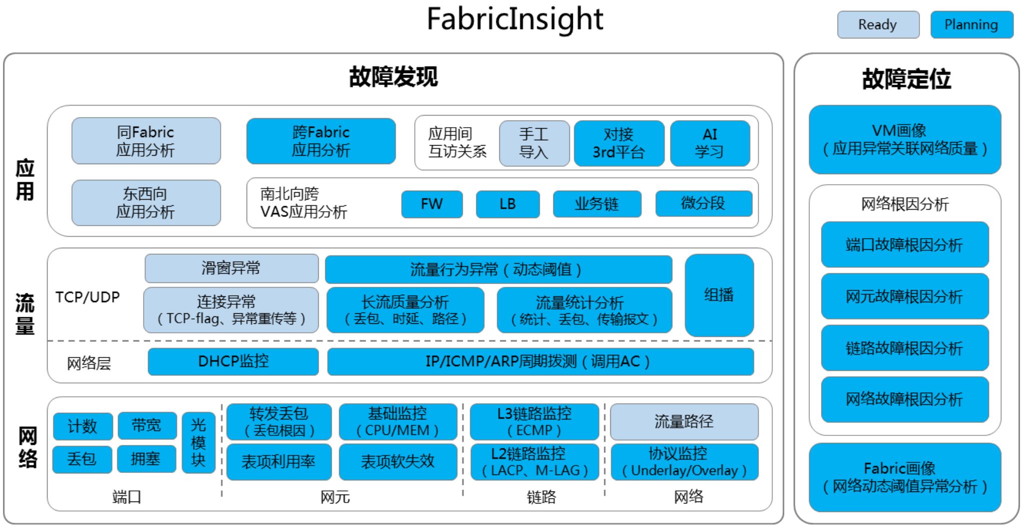
In 0, the SDN network consists of the application network, logical network, and physical network. Services provisioned by the Agile Controller-DCN are deployed on the application network and logical network, which are the key O&M layers of the Agile Controller-DCN. O&M functions on the application network and logical network include configuration consistency check, mapping between network layers, and query of logical network details. O&M on the physical network focuses on quick management of devices on the Agile Controller-DCN and the impact of the physical network on the application network and logical network.

Figure 2-5 O&M functions of eSight


Ready
Planning
eSight
Network quality
Service traffic monitoring (IPsec)
Service traffic monitoring (GRE/SSL)
Real app traffic visibility <iPCA, delay/packet loss)
DHCP performance monitoring
Application network O&M
Network health monitoring
Logical ServiceFunction traffic statistics
Logical router traffic statistics
Logical switch traffic statistics
Logical NE performance monitoring (historical)
Logical NE traffic statistics
Logical network O&M
Software upgrade and patching
Configuration file backup and restoration
Physical FW management
Software FW management
vSwitch management
Physical switch management
Panel
Status
Performance
Failure
Traffic
Physical NE performance monitoring
NE upgrade
Virtual FW
Physical FW
vSwitch
Physical switch
Server management
Physical resource pool
Controller monitoring
Switch installation via ZTP
Physical network O&M
Log
Path optimization

In Figure 2-5, eSight focuses on the traditional network management field, including physical NE management, deployment, and upgrade. On the logical network layer, eSight presents traffic statistics and historical performance data. On the application network layer, eSight provides advanced O&M functions such as iPCA, which can be used to detect the quality of application networks provided on the Agile Controller-DCN, facilitating path adjustment and troubleshooting.

Figure 2-6 O&M functions of FabricInsight



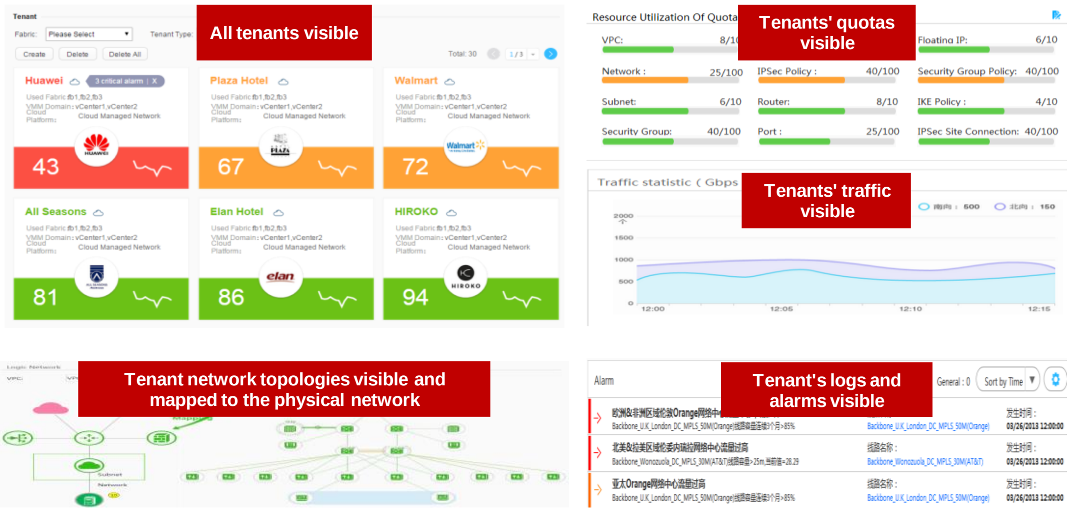
In Figure 2-6, FabricInsight collects and stores all network traffic data and analyzes network application traffic trend and faults using intelligent Big Data analysis algorithms for quick network fault identification and locating. In addition, FabricInsight is associated with the Agile Controller that provides network control capabilities to quickly isolate network faults.

2.4 Description

2.4.1 Network Resource Visualization

After a DCN is successfully deployed and user services are provisioned, administrators need to obtain information about available resources, tenant-based resource usage, and overall usage of DCN resources, and determine whether the network needs to be expanded.

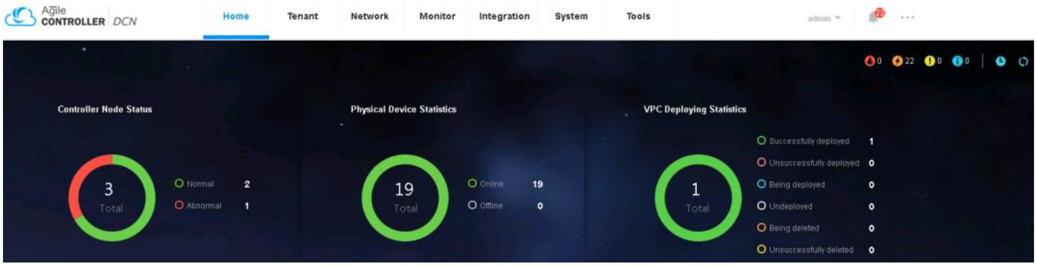
Figure 2-7 Network resource visualization panorama



2.4.1.1 Resources on the Entire Network

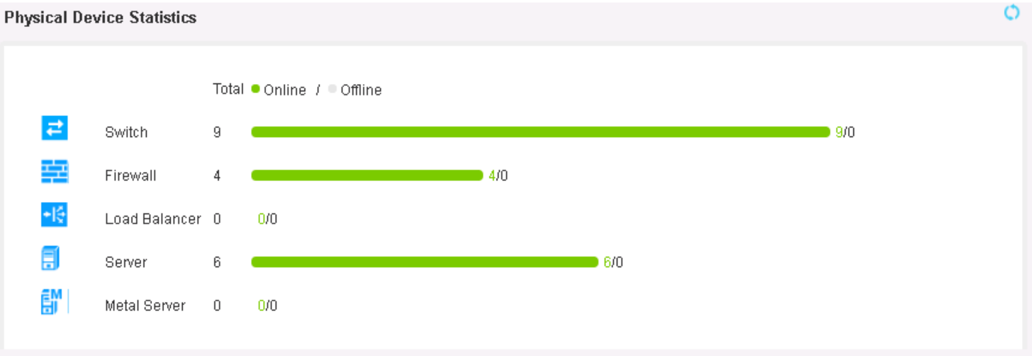
Resources on the entire network include hardware switches, virtual switches, hardware firewalls, software firewalls, and load balancers. Figure 2-8 shows statistics on physical devices and VPC deployment.

Figure 2-8 Physical device statistics (1)



The **Physical Device Statistics** page intuitively displays the total number of devices of various types and device online status. On this page, **Server** refers to virtual servers, and **Metal Server** refers to bare metal servers.

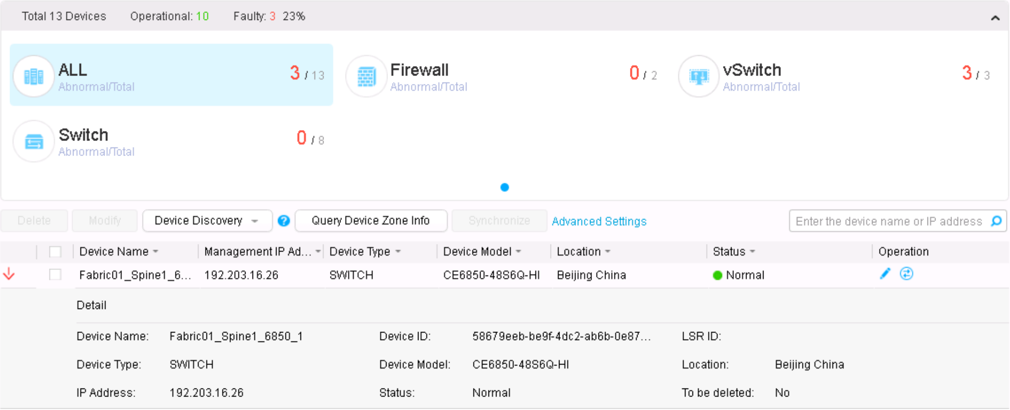
Figure 2-9 Physical device statistics (2)



Physical Device Statistics

	Total	● Online / ● Offline
Switch	9	9/0
Firewall	4	4/0
Load Balancer	0	0/0
Server	6	6/0
Metal Server	0	0/0

Figure 2-10 Detailed physical device information



Total 13 Devices Operational: 10 Faulty: 3 23%

	Device Name	Management IP Ad...	Device Type	Device Model	Location	Status	Operation
□	Fabric01_Spine1_6...	192.203.16.26	SWITCH	CE6850-48S6Q-HI	Beijing China	● Normal	

Detail

Device Name:	Fabric01_Spine1_6850_1	Device ID:	58679eeb-be9f-4dc2-ab6b-0e87...	LSR ID:	
Device Type:	SWITCH	Device Model:	CE6850-48S6Q-HI	Location:	Beijing China
IP Address:	192.203.16.26	Status:	Normal	To be deleted:	No

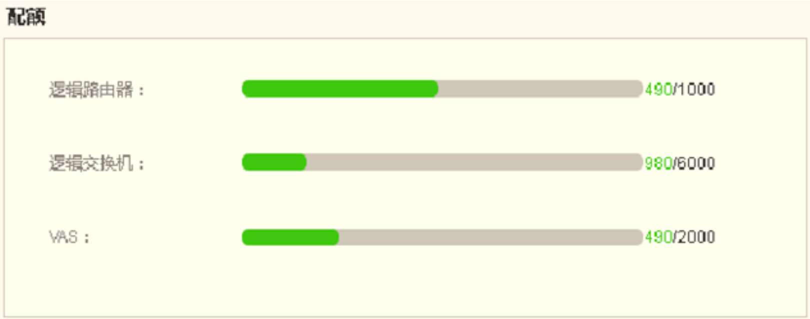
2.4.1.2 Logical Resources

After traditional data centers are cloudified, one physical device can be virtualized into multiple logical devices that concurrently provide services to multiple tenants. Tenants are logically isolated from each other.

Each tenant can control the logical resources in the quota allocated to the tenant. These logical resources include logical routers, logical switches, logical firewalls, logical load balancers (logical firewalls and load balancers are collectively called logical VAS devices), logical ports, and End Point Groups (EPGs).

When creating tenants, the network administrator allocates logical resource quotas to tenants. Tenant administrators can check the resource usage of managed tenants. When resources are insufficient, tenant administrators can apply for resource expansion.

Figure 2-11 Quota and status of a tenant's logical resources

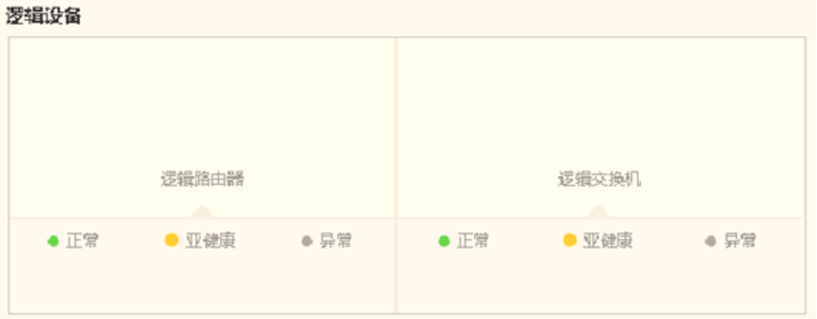


配额

逻辑路由器： 490/1000

逻辑交换机： 980/6000

VAS： 490/2000



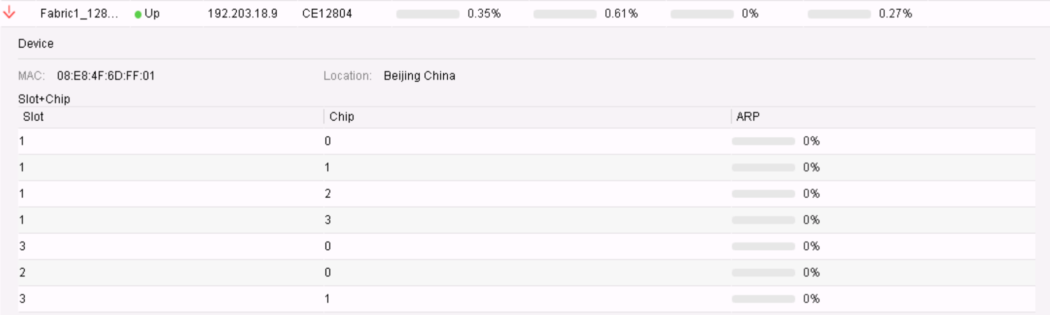
The network administrator can check the resource allocation situation of all tenants and the overall resource allocation situation of the data center. The number of logical resources that are allocated can exceed the actual physical resources. It is recommended that network administrators periodically check the logical resource view and expand the capacity in a timely manner if necessary.

2.4.1.3 Device Resources

Huawei CloudFabric DCN Solution provides the function of querying logical resources on a physical device.

The logical resource query function enables an administrator to select a fabric NE, view the overall capacity of key resources on the gateway and resource usage, and determine whether the layout of computing resources of the data center needs to be adjusted or whether the capacity needs to be expanded. Currently, resources including BDs, VRFs, ARPs, and routes on a device can be queried.

Figure 2-12 Logical resources of a device



Fabric1_128... ● Up 192.203.18.9 CE12804 0.35% 0.61% 0% 0.27%

Device

MAC: 08:E8:4F:6D:FF:01 Location: Beijing China

Slot+Chip

Slot	Chip	ARP
1	0	0%
1	1	0%
1	2	0%
1	3	0%
3	0	0%
2	0	0%
3	1	0%

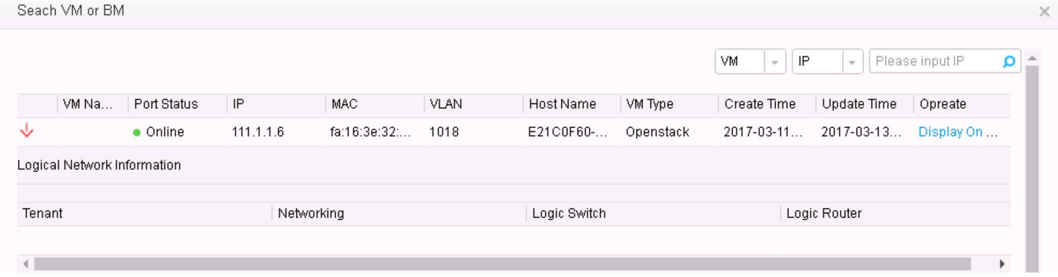
2.4.2 Computing Resource Visualization

2.4.2.1 Position Query

Data center resource pooling is one of the benefits of the cloud. Due to resource pooling, positions of workloads frequently change. For example, provisioning, reclaiming, migration, and disaster recovery of workloads lead to changes of access positions of compute nodes. When service faults occur on a workload, quickly and accurately locating the access position of the compute node will shorten the time for fault demarcation and location. In some cases, the access position identifies the fault cause, for example, access switch power failure.

Huawei CloudFabric DCN Solution enables rapid location of access positions of VM/BM resources. The IP/MAC address and computing resource type can be used to query the access position of a VM/BM. The query result includes the VM/BM name, port status, IP and MAC addresses, access VLAN, host name, VM type, and creation/modification time. Users can click **Display On Topo** to access the fabric topology page. On the displayed page, the VM/BM blinks.

Figure 2-13 VM/BM position query



Seach VM or BM

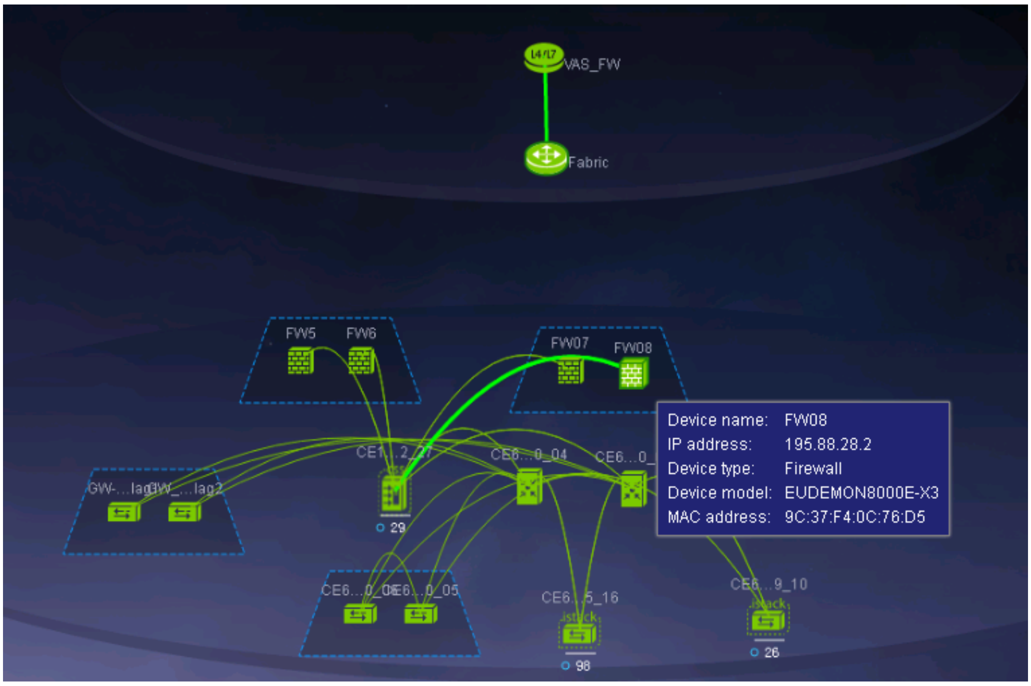
VM | IP | Please input IP

	VM Na...	Port Status	IP	MAC	VLAN	Host Name	VM Type	Create Time	Update Time	Opreate
↓		● Online	111.1.1.6	fa:16:3e:32:...	1018	E21C0F60-...	Openstack	2017-03-11...	2017-03-13...	Display On ...

Logical Network Information

Tenant	Networking	Logic Switch	Logic Router

Figure 2-14 VM/BM position topology



Device name: FW08
IP address: 195.88.28.2
Device type: Firewall
Device model: EUDEMON8000E-X3
MAC address: 9C:37:F4:0C:76:D5

2.4.2.2 Details Query

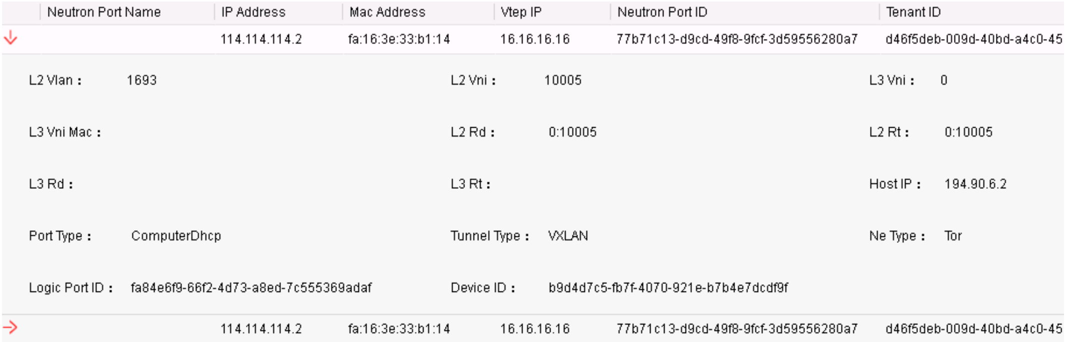
The position query function can be used to query the physical access position information about the VM/BM access network. The details query function can be used to query detailed VM/BM information.

In addition to compute node information such as the VM/BM name, IP/MAC address, port ID, port type, and host IP address, the details query function can be used to query detailed information about the logical network to which the VM/BM is connected, including the tenant ID, L2 VNI of the logical switch, L3 VNI of the logical router, EVPN RD/RT information, and VTEP IP address. The detailed information can be used to obtain the index of the logical network.

The logical switch index (L2 VNI) can be used to further query the logical switch details to obtain the panorama of the logical network to which the compute node connects.

Whether fabric NE configurations are correct can be determined based on the BGP EVPN RD/RT information and service logs.

Figure 2-15 VM/BM details view



	Neutron Port Name	IP Address	Mac Address	Vtep IP	Neutron Port ID	Tenant ID
↓		114.114.114.2	fa:16:3e:33:b1:14	16.16.16.16	77b71c13-d9cd-49f8-9fcf-3d59556280a7	d46f5deb-009d-40bd-a4c0-45

L2 Vlan :	1693	L2 Vni :	10005	L3 Vni :	0
L3 Vni Mac :		L2 Rd :	0:10005	L2 Rt :	0:10005
L3 Rd :		L3 Rt :		Host IP :	194.90.6.2
Port Type :	ComputerDhcp	Tunnel Type :	VXLAN	Ne Type :	Tor
Logic Port ID :	fa84e6f9-66f2-4d73-a8ed-7c555369adaf	Device ID :	b9d4d7c5-fb7f-4070-921e-b7b4e7dcdf9f		

→		114.114.114.2	fa:16:3e:33:b1:14	16.16.16.16	77b71c13-d9cd-49f8-9fcf-3d59556280a7	d46f5deb-009d-40bd-a4c0-45

2.4.3 Visibility of Application, Logical, and Physical Networks

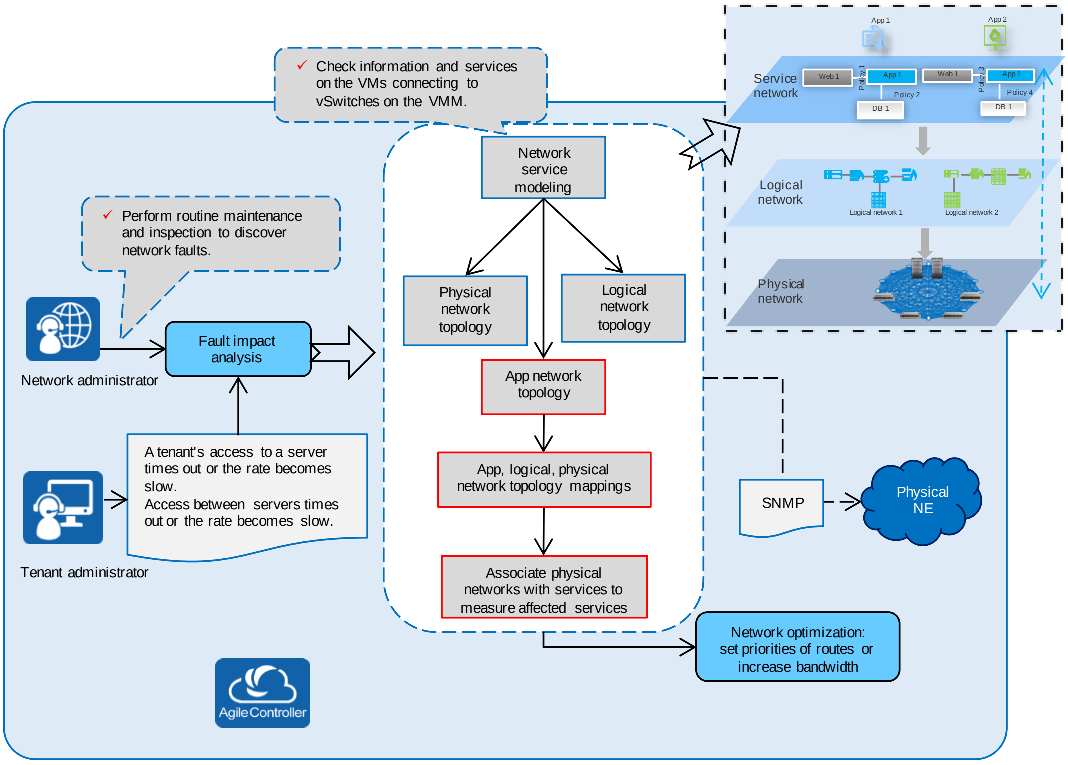
Visibility of application, logical, and physical networks includes:

- Forward direction: mapping from the application network to the logical network and mapping from the logical network to the physical network
- Reverse direction: mapping from the physical network to the logical network and mapping from the logical network to the application network

The visibility feature enables administrators to:

- Intuitively determine the tenants whose logical networks are affected by a faulty physical device.
- Intuitively determine the affected applications on the affected logical networks.
- Intuitively determine the tenants to whom an application that is unavailable provides services.
- Use O&M technologies such as multi-path detection and connectivity check on the affected logical networks to determine whether the application faults are caused by service faults or network interruptions.
- Use O&M technologies such as single-path detection and iPCA to locate the faulty NE.

Figure 2-16 Visibility of application, logical, and physical networks


✓ Check information and services on the VMs connecting to vSwitches on the VMM.
✓ Perform routine maintenance and inspection to discover network faults.
Network administrator
Fault impact analysis
A tenant's access to a server times out or the rate becomes slow.
Access between servers times out or the rate becomes slow.
Tenant administrator
Network service modeling
Physical network topology
Logical network topology
App network topology
App, logical, physical network topology mappings
Associate physical networks with services to measure affected services
Network optimization: set priorities of routes or increase bandwidth
SNMP
Physical NE
AgileController
App 1
App 2
Service network
Web 1
Policy 1
App 1
Policy 2
DB 1
Web 1
Policy 3
App 1
Policy 4
DB 1
Logical network
Logical network 1
Logical network 2
Physical network

2.4.3.1 Physical Network

The Agile Controller-DCN scans NEs such as switches, vSwitches, firewalls, and load balancers using SNMP, automatically perceives the connections between NEs using LLDP, draws a physical topology view based the connections, and presents the physical topology view to users.

Physical network topology management has the following characteristics:

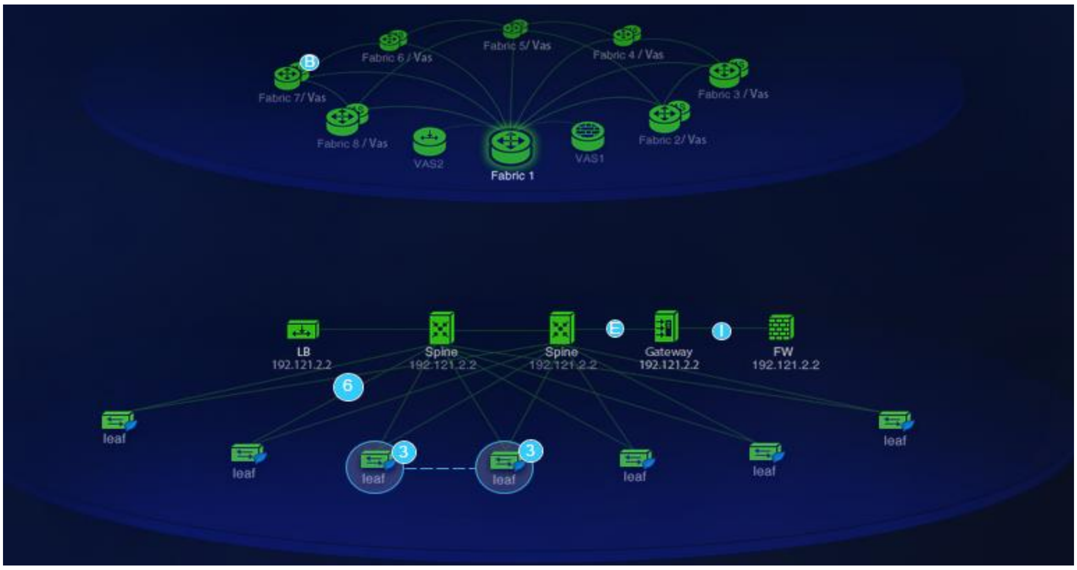
- Based on the actual network deployment, the physical network topology displays the physical connections between data center network devices. Users can check whether the network topology is consistent with actual physical deployment.
- The physical network topology displays subnets by layer and uses different colors to identify the alarm event status of devices.

A physical topology is composed of the following objects:

- Fabric: an autonomous system of a group of homogeneous networks. For example, a VXLAN network is a fabric that provides Layer 2 and Layer 3 connections for external interfaces.
- VAS pool: a group of device pools providing Layer 4 to Layer 7 VASs. The Agile Controller-DCN selects devices from the VAS pool to provide VASs to users.
- Device group: a group of devices that are associated through technologies such as iStack, CSS, M-LAG, SVF, VRRP, and HRP (firewalls).
- SF device: a device providing Layer 4 to Layer 7 VASs, for example, a firewall or load balancer.

- Fabric device: a network device that provides Layer 2 and Layer 3 basic forwarding services, for example, a CE series switch or CE1800V switch.
- Host: a server, which may be a bare metal BM server or a virtual server (VM on which the Hypervisor is deployed).
- Link: a Layer 2 link, which may be a link between the TOR switch and the host, link between devices, or link connecting to an external network.

Figure 2-17 Physical topology view



The physical topology presents the actual networking of the data center. Before data center service provisioning, information displayed on the physical topology can be used to determine whether network deployment is normal. Data center services can be provisioned only when the physical topology is consistent with the designed network topology.

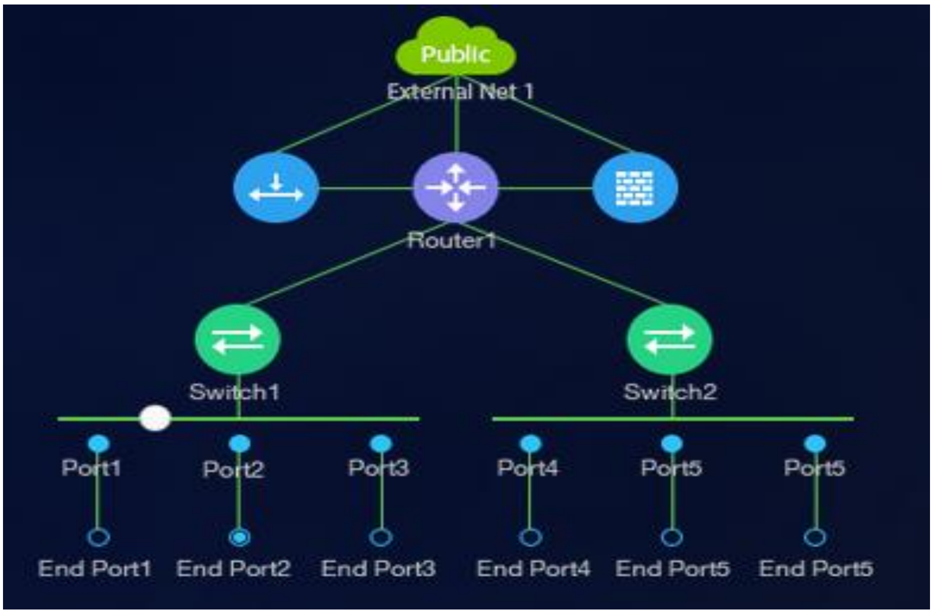
2.4.3.2 Logical Network

Logical networks are exclusively occupied by tenants and are isolated for tenants by allocating logical resources on physical devices. For CE series switches, logical resources refer to VRFs, VNIs, VLANs, and routes. For the Next-Generation Firewall (NGFW), logical resources refer to the vSYS, NAT, and policies. For load balancers, logical resources refer to routing domains.

The preceding logical resources can be orchestrated into logical networks to provide Layer 2 to Layer 7 network services for tenants.

- Logical router: abstraction of a Layer 3 VXLAN gateway, mapped to a VRF.
- Logical switch: abstraction of a Layer 2 broadcast domain in a VXLAN, mapped to a VNI.
- Logical port: abstraction of an end node port on a fabric device.
- SF: abstraction of Layer 4 to Layer 7 VASs.
- End port: port outside of the fabric edge, such as ports on the VM, BM, external device, and firewall.
- Logical link: link connecting a logical port and an end port.

Figure 2-18 Logical topology view



2.4.3.3 Application Network

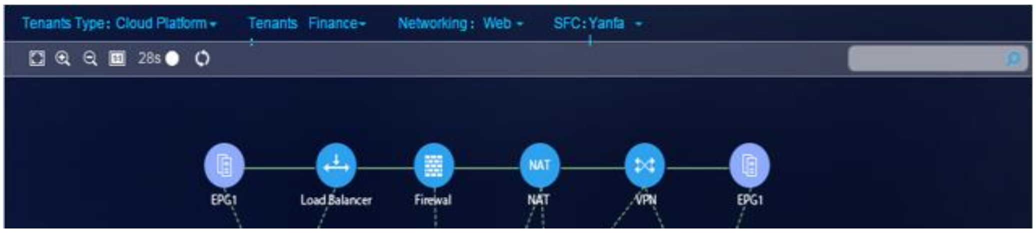
The definition of the application topology on the Agile Controller-DCN varies according to scenario.

Web-App-DB is a typical orchestration mode for tenants' logical networks. The three-layer architecture describes the deployment model of a service. The web, app, and DB can be defined on different networks (subnets). Then the subnets can be associated with EPGs to generate the application topology related to tenant services.

- App: aggregation of multiple EPGs, providing comprehensive network and computing resources.
- SFC: standard service chain model in the industry, including two EPGs and a group of sequenced abstraction service functions.
- EPG: collection of end ports and logical NEs with the same security policy.

The service chain is used to connect EPGs and SFs to implement service-based refined control.

Figure 2-19 Application topology view



Tenants Type: Cloud Platform
Tenants Finance
Networking: Web
SFC: Yanfa

2.4.3.4 Three-Layer Network Visibility

Three-layer network visibility includes:

- Forward direction: visibility of application > logical > physical network topologies
- Reverse direction: visibility of physical > logical > application network topologies

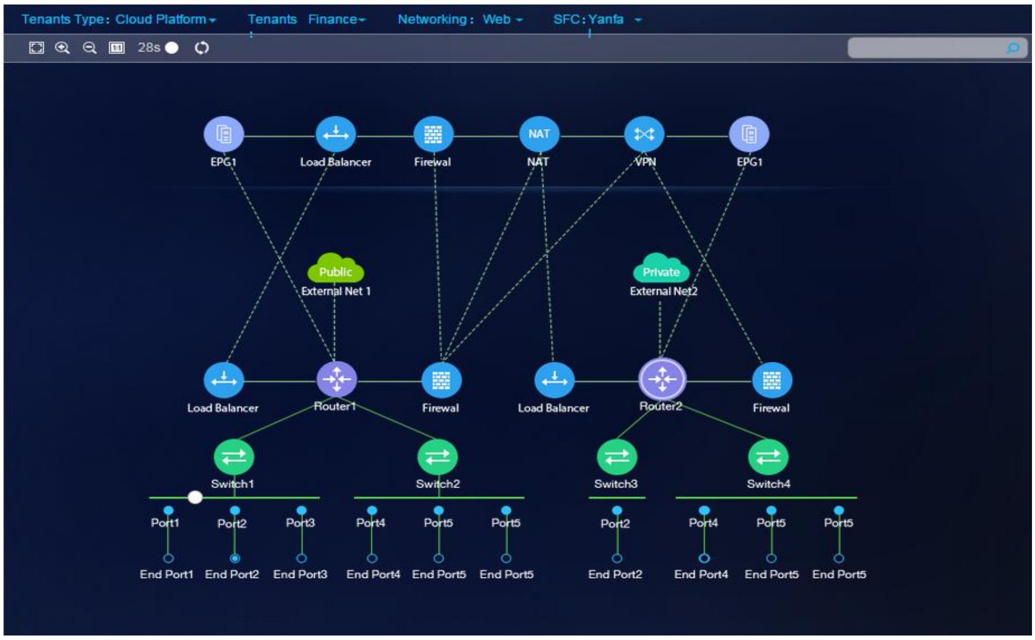
Forward Direction: Visibility of Application > Logical > Physical Network Topologies

- System administrators can view logical and physical resources used by all tenants' application networks and perform routine O&M, monitoring, and detection based on the mappings. If the usage of physical resources reaches the specified threshold, system administrators can purchase, expand, or upgrade resources.
- During resource preparation and service provisioning, system administrators can query physical resources corresponding to provisioned logical resources and adjust resource allocation as required.
- Logical and physical networks change when the application network changes.
- During service provisioning based on the logical topology, network administrators can query physical resources based on the logical resources that are mapped to the physical resources.

Reverse Direction: Visibility of Physical > Logical > Application Network Topologies

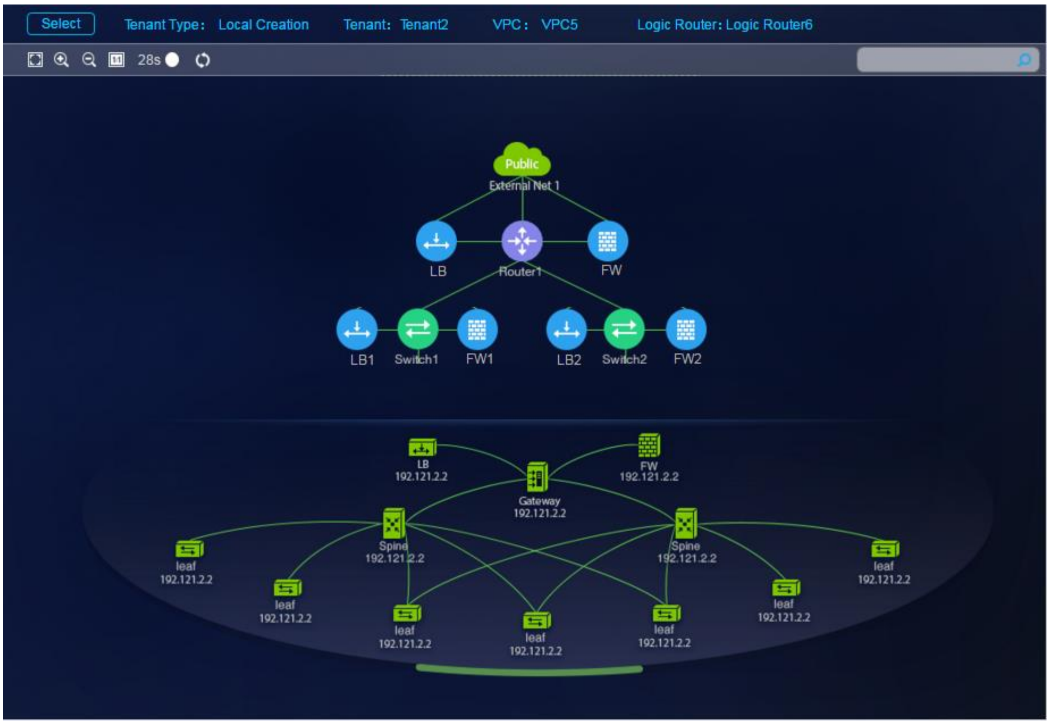
- Administrators and tenant administrators can determine the impact scope of faulty physical devices and ports based on Physical > Logical > Application mapping in the reverse direction, and isolate or suspend application services accordingly.
- The Agile Controller-DCN can simulate the impacts of device replacement, enabling administrators to know the affected services and suspend or migrate the services accordingly before device replacement or migration.
- Logical and application networks change when the physical network changes.

Figure 2-20 Mapping from the application topology to the logical topology



Tenants Type：Cloud Platform
Tenants Finance
Networking：Web
SFC：Yanfa
28s

Figure 2-21 Mapping from the logical topology to the physical topology



Select
Tenant Type：Local Creation
Tenant：Tenant2
VPC：VPC5
Logic Router：Logic Router6
28s

2.4.4 Overlay Connectivity Check

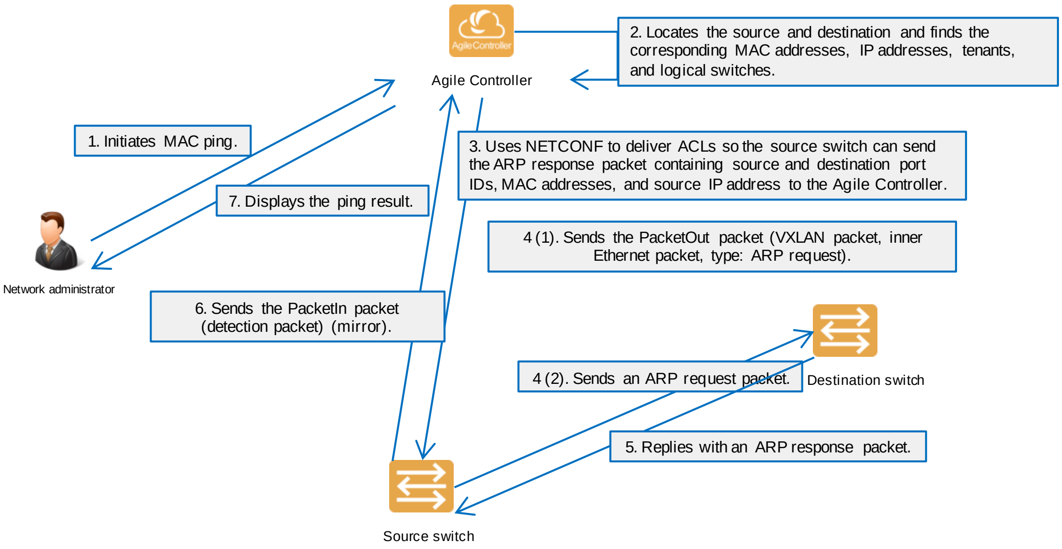
Huawei CloudFabric DCN Solution introduces the **ping** and **trace** commands that are used to demarcate faults on traditional networks into the overlay network to demarcate and locate logical network faults for tenants. It provides the following overlay connectivity check functions:

- MAC ping
- IP ping

2.4.4.1 MAC Ping

MAC ping is also called ARP ping. Encapsulated ARP packets are sent to a specific interface in the same broadcast domain to check whether the packets can be forwarded at Layer 2. If the packets fail to reach the destination interface, a fault occurs on the fabric network. During MAC ping, MAC forwarding entries of switches are checked. Currently, the MAC ping operation is supported between a VM and a physical server or between two VMs. Figure 2-22 shows the MAC ping process.

Figure 2-22 MAC ping process


Agile Controller
2. Locates the source and destination and finds the corresponding MAC addresses, IP addresses, tenants, and logical switches.
1. Initiates MAC ping.
3. Uses NETCONF to deliver ACLs so the source switch can send the ARP response packet containing source and destination port IDs, MAC addresses, and source IP address to the Agile Controller.
7. Displays the ping result.
4 (1). Sends the PacketOut packet (VXLAN packet, inner Ethernet packet, type: ARP request).
Network administrator
6. Sends the PacketIn packet (detection packet) (mirror).
4 (2). Sends an ARP request packet.
Destination switch
5. Replies with an ARP response packet.
Source switch

If the ping between VMs/BMs fails, check whether the network is abnormal due to the following factors:

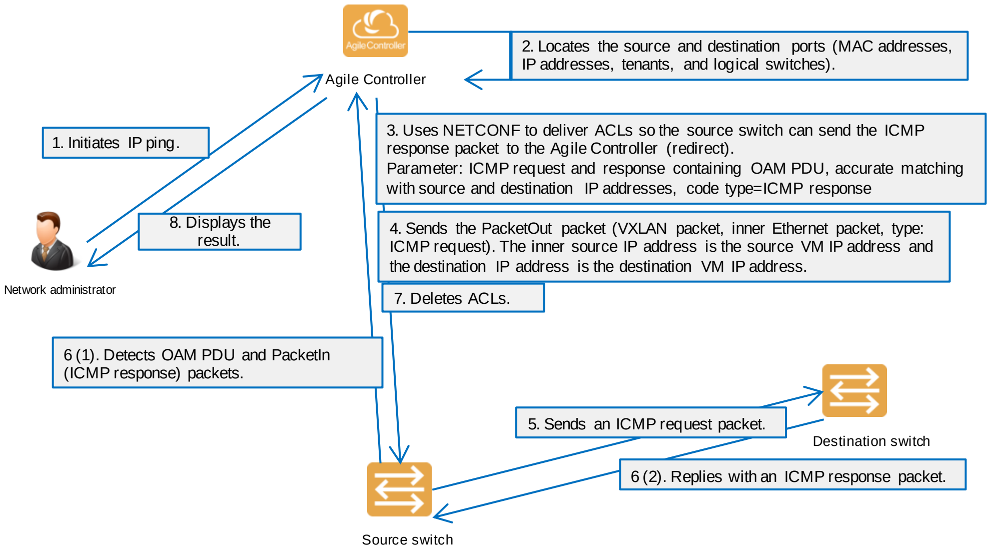
- The connection between the source and the gateway is abnormal.
- Layer 2 forwarding is abnormal in scenarios where the source and destination VMs are located on a Layer 2 network.
- IP packets are filtered due to firewall settings but the Layer 2 network is functional.

When performing the MAC ping operation, the Agile Controller-DCN constructs an ARP request packet that is originated from the source VM/BM and destined to the destination IP address (on the same network segment) and sends the packet to the NVE connected to the source VM/BM to simulate the process in which the source VM/BM queries the MAC address of the destination IP address. The Agile Controller-DCN listens for ARP response packets that the source NVE node receives from the destination IP address. If the source NVE node receives an ARP response packet from the destination IP address, the Layer 2 network is reachable. The cause needs to be further analyzed.

2.4.4.2 IP Ping

IP ping uses encapsulated Internet Control Message Protocol (ICMP) packets to check whether a specific interface in the same routing domain is reachable. If the packets fail to reach the destination interface, a fault occurs on the fabric network. During IP ping, MAC, ARP, and route forwarding entries of switches are checked. Currently, the IP ping operation is supported between a VM and a physical server or between two VMs. Figure 2-23 shows the IP ping process.

Figure 2-23 IP ping process



1. Initiates IP ping.

2. Locates the source and destination ports (MAC addresses, IP addresses, tenants, and logical switches).

3. Uses NETCONF to deliver ACLs so the source switch can send the ICMP response packet to the Agile Controller (redirect).
Parameter: ICMP request and response containing OAM PDU, accurate matching with source and destination IP addresses, code type=ICMP response

4. Sends the PacketOut packet (VXLAN packet, inner Ethernet packet, type: ICMP request). The inner source IP address is the source VM IP address and the destination IP address is the destination VM IP address.

5. Sends an ICMP request packet.

6 (1). Detects OAM PDU and PacketIn (ICMP response) packets.

6 (2). Replies with an ICMP response packet.

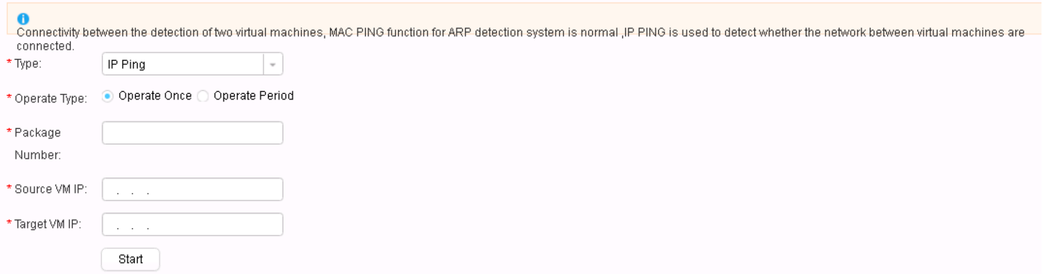
7. Deletes ACLs.

8. Displays the result.

If a VM or BM fails to ping an IP address, network maintenance personnel can perform the IP ping operation on the Agile Controller-DCN to confirm the fault. When performing the IP ping operation, the Agile Controller-DCN constructs an ICMP request packet that is originated from the source VM/BM and destined to the destination IP address and sends the packet to the NVE connected to the source VM/BM to simulate the process in which the source VM/BM pings the destination IP address. The Agile Controller-DCN listens for ICMP response packets that the source NVE node receives from the destination IP address. If the source NVE node receives an ICMP response packet from the destination IP address, the Layer 2 network is reachable.

If the ping fails, the source VM/BM may be faulty. If the source NVE node does not receive an ICMP response packet from the destination IP address, a network fault may occur and needs to be further diagnosed.

Figure 2-24 IP/MAC ping operation page



Connectivity between the detection of two virtual machines, MAC PING function for ARP detection system is normal ,IP PING is used to detect whether the network between virtual machines are connected.

* Type: IP Ping

* Operate Type: ● Operate Once ○ Operate Period

* Package Number:

* Source VM IP: . . .

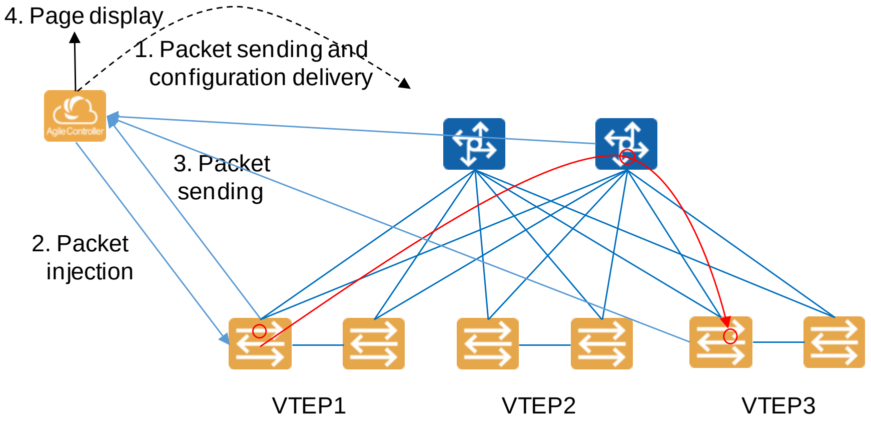
* Target VM IP: . . .

2.4.5 Overlay Single Path Detection

If the overlay connectivity check technology demarcates that the fault is caused by the fabric network, the overlay single path detection technology can be used to locate the NE that causes the fault. Overlay single path detection is also called overlay IP trace, which is referred to as IP trace in this document.

The IP trace function detects service packet forwarding paths hop by hop to accurately locate the place where packets are unavailable. In addition, the IP trace function can accurately predict the services affected by forwarding congestion based on the complete service packet forwarding path and network congestion situation.

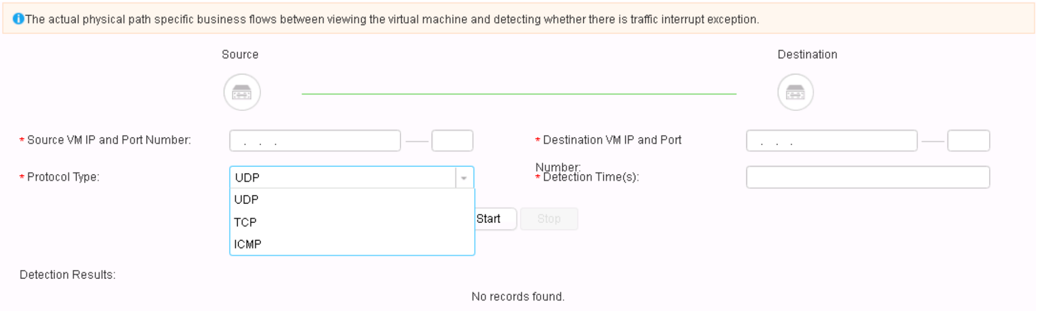
Figure 2-25 Process of overlay single path detection



Based on the encapsulated packet type, IP trace is classified into ICMP Trace, UDP Trace, and TCP Trace.

If two IP addresses of a data center are unreachable, it is recommended that IP ping be performed on the Agile Controller-DCN to preliminarily check for network faults. If IP ping fails on the Agile Controller-DCN, it is preliminarily determined that the network is faulty. In this case, it is recommended that ICMP trace be performed to check the forwarding paths. Then the Agile Controller-DCN can accurately detect the forwarding paths of ICMP request and response packets and display the traversed devices in the network topology. In this way, network O&M personnel can easily locate the faulty device.

Figure 2-26 IP trace operation page



The actual physical path specific business flows between viewing the virtual machine and detecting whether there is traffic interrupt exception.

Source

Destination

* Source VM IP and Port Number: . . . —

* Destination VM IP and Port Number: . . . —

* Protocol Type: UDP

* Detection Time(s):

Detection Results:

No records found.

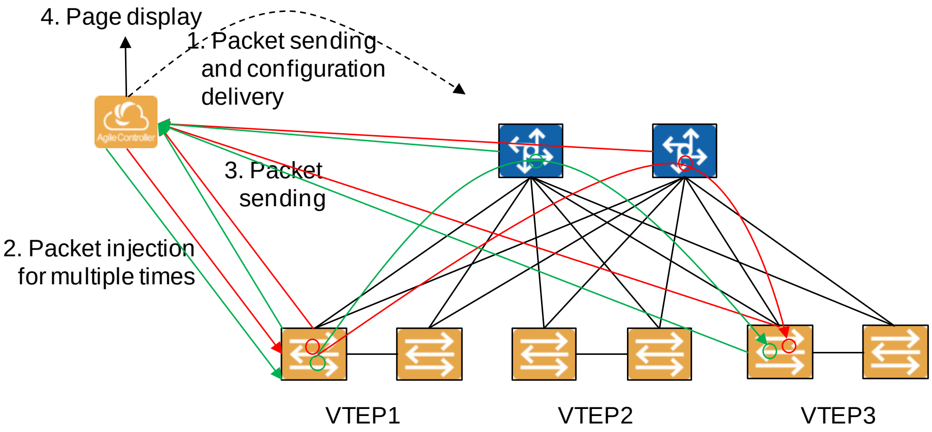
The TCP Trace and UDP functions can be used to locate the forwarding path of service packets on the network. If two nodes can ping each other and only some services are affected, UDP/TCP packets can be used to locate the faulty network device. Services that will be affected upon congestion during full-configuration forwarding or upon network upgrade can be predicted based on the service packet forwarding path located using TCP/UDP packets. Then measures can be taken accordingly to minimize the impact on the services.

2.4.6 Underlay Multi-Path Detection

Under network virtualization, underlay networks are usually IP equal-cost multi-path routing (ECMP) networks, on which VTEP tunnels carry traffic in load balancing mode through ECMP. If service access to two or more VTEPs is interrupted sometimes, or some services between two VMs or BMs are abnormal, some packets may be lost along paths between the VTEPs due to traffic congestion.

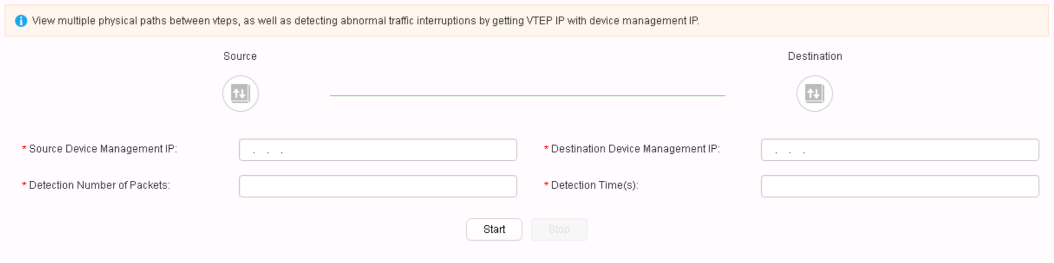
To detect as many forwarding paths as possible, Huawei CloudFabric DCN Solution provides the multi-path detection function for VTEPs on the underlay network. The Agile Controller-DCN sends a batch of detection packets to detect possible forwarding paths between VMs or BMs.

Figure 2-27 Process of underlay multi-path detection


4. Page display
1. Packet sending and configuration delivery
Agile Controller
3. Packet sending
2. Packet injection for multiple times
VTEP1
VTEP2
VTEP3

IP ECMP uses the hash algorithm to select routes on an underlay network. If there are only limited packets, some forwarding paths may fail to be detected. Therefore, underlay multi-path detection can only be used as a supplementary.

Figure 2-28 Operation page of underlay multi-path detection



View multiple physical paths between vteps, as well as detecting abnormal traffic interruptions by getting VTEP IP with device management IP.

* Source Device Management IP:	. . .	* Destination Device Management IP:	. . .
* Detection Number of Packets:		* Detection Time(s):	

Differences Between Overlay Single Path Detection and Underlay Multi-Path Detection

- Overlay single path detection requires information such as the IP address, protocol, Layer 4 port, which specifies a unique service packet forwarding path. Underlay multi-path detection detects multiple paths for forwarding service packets of various types between VTEPs.
- Overlay single path detection only needs to send a small number of packets to specify the forwarding path, which takes a short time and causes low pressure on the Agile Controller-DCN. Underlay multi-path detection requires a large number of packets to detect as many forwarding paths as possible, which is time-consuming and causes transient high pressure on the Agile Controller-DCN.
- During overlay single path detection, the VM/BM packets simulated by the Agile Controller-DCN carry special identifiers. When receiving the packets, the source VTEP queries the forwarding entries to obtain the destination VTEP, adds the VXLAN header to the packets, and then sends the packets. During underlay multi-path detection, only the source and destination VTEPs are added to the original packets sent by the Agile Controller-DCN, and no VXLAN header is added. When receiving the packets, the source VTEP queries forwarding entries and directly forwards the packets on the underlay network without modifying them.

Similarities of Overlay Single Path Detection and Underlay Multi-Path Detection

Before overlay path detection and underlay path detection, the Agile Controller-DCN delivers related commands to devices that packets may traverse. When forwarding packets with special identifiers, the devices process the packets and send them to the Agile Controller-DCN. Then the Agile Controller-DCN calculates forwarding paths hop by hop based on the received packets.

Different from the **trace** command that controls the TTL to detect the interface address hop by hop on traditional networks, the hop-by-hop path detection technology delivers packet capturing commands and sends packets with special identifiers to detect paths. Therefore, both overlay single path detection and underlay multi-path detection are applicable only to Huawei CE series switches. These functions can be used on non-Huawei CE switches but the forwarding paths cannot be calculated.

NOTE

The preceding restraints affect only path detection. When the connectivity check function is used, VTEPs must be Huawei CE series switches but no requirements are imposed on devices on the underlay network.

2.4.7 Performance Monitoring (eSight)

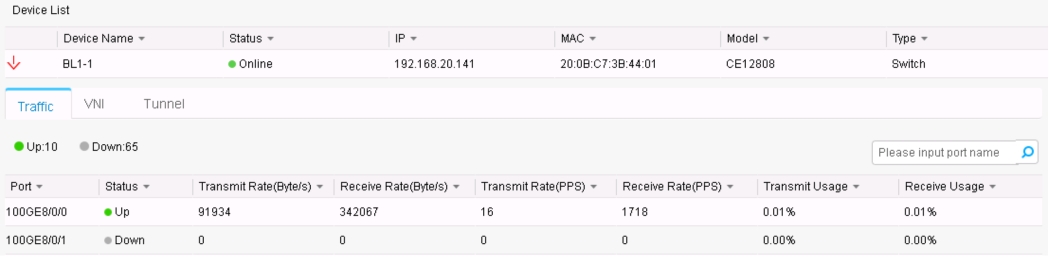
During normal operation, the performance of a network may deteriorate due to internal or external factors, triggering network faults. Network efficiency such as network disconnection rate and utilization needs to be planned, monitored, and measured to ensure sufficient network performance in low costs and meet future requirements on the network. Performance management helps detect and resolve network performance deterioration, preventing network faults.

In Huawei CloudFabric DCN Solution, eSight provides detailed performance monitoring schemes and related features such as performance monitoring templates, performance monitoring setting, real-time performance query, and historical performance query. Users can select required function modules based on actual requirements.

In addition, the Agile Controller-DCN in Huawei CloudFabric DCN Solution provides the following overlay traffic statistics functions:

- Port traffic statistics: collects statistics on the real-time traffic and bandwidth usage on ports of hardware switches, as shown in Figure 2-29.

Figure 2-29 Port traffic statistics



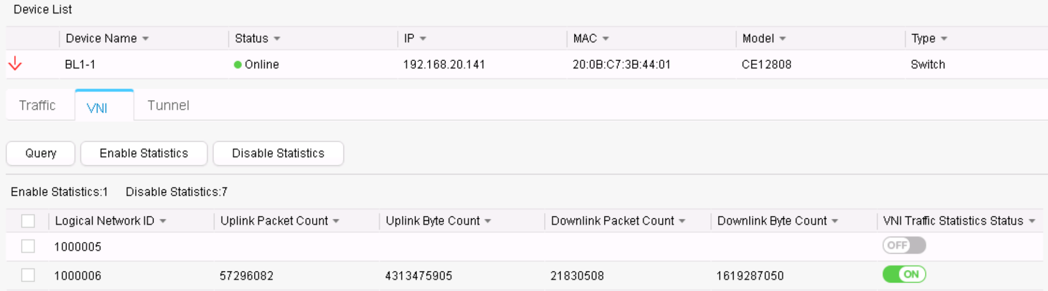
Device List

	Device Name	Status	IP	MAC	Model	Type
↓	BL1-1	● Online	192.168.20.141	20:0B:C7:3B:44:01	CE12808	Switch

Port	Status	Transmit Rate(Byte/s)	Receive Rate(Byte/s)	Transmit Rate(PPS)	Receive Rate(PPS)	Transmit Usage	Receive Usage
100GE8/0/0	● Up	91934	342067	16	1718	0.01%	0.01%
100GE8/0/1	● Down	0	0	0	0	0.00%	0.00%

- VNI traffic statistics: collects statistics on the number of uplink and downlink packets and bytes of a specific VNI on physical switches, as shown in Figure 2-30.

Figure 2-30 VNI traffic statistics



Device List

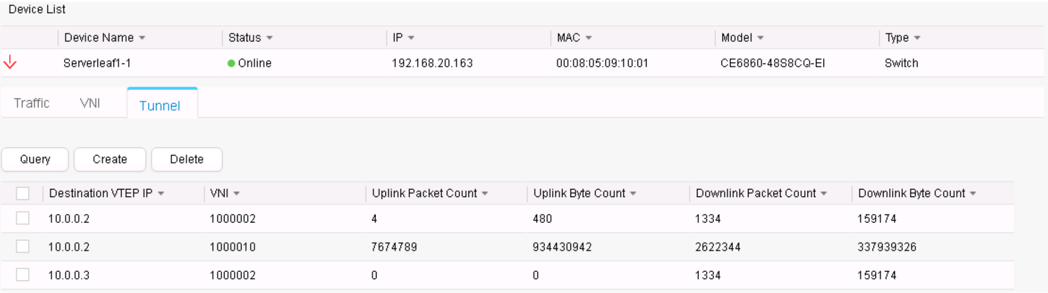
	Device Name	Status	IP	MAC	Model	Type
↓	BL1-1	● Online	192.168.20.141	20:0B:C7:3B:44:01	CE12808	Switch

Enable Statistics:1 Disable Statistics:7

□	Logical Network ID	Uplink Packet Count	Uplink Byte Count	Downlink Packet Count	Downlink Byte Count	VNI Traffic Statistics Status
□	1000005					OFF
□	1000006	57296082	4313475905	21830508	1619287050	ON

- Inter-VTEP tunnel traffic statistics: collects statistics on the traffic on VXLAN tunnels between two VTEPs (physical switches), as shown in Figure 2-31.

Figure 2-31 Inter-VTEP tunnel traffic statistics



Device List

	Device Name	Status	IP	MAC	Model	Type
↓	Serverleaf1-1	● Online	192.168.20.163	00:08:05:09:10:01	CE6860-48S8CQ-EI	Switch

□	Destination VTEP IP	VNI	Uplink Packet Count	Uplink Byte Count	Downlink Packet Count	Downlink Byte Count
□	10.0.0.2	1000002	4	480	1334	159174
□	10.0.0.2	1000010	7674789	934430942	2622344	337939326
□	10.0.0.3	1000002	0	0	1334	159174

The preceding overlay traffic statistics functions need to be enabled on the GUI and support only real-time traffic statistics on the fabric network. eSight needs to be deployed to provide advanced performance monitoring functions such as historical traffic analysis.

2.4.8 Alarm Management (Agile Controller-DCN + eSight)

In Huawei CloudFabric DCN solution, the Agile Controller-DCN and eSight implement the alarm function. The Agile Controller-DCN presents alarms on the overlay network, and eSight presents alarms on the underlay infrastructure network. Users can determine whether to deploy eSight based on actual requirements.

Alarms Displayed on the Agile Controller-DCN

Alarms displayed on the Agile Controller-DCN can be classified into the following types:

- Service alarm: ARP attack alarm, firewall role switching alarm, alarm reported when logical service resources reach an alarm threshold, alarm of disconnection from managed devices, and alarm of interruption of links between devices
- Software alarm: alarms reported when the CPU or memory usage exceeds a threshold, licenses expire, resources are exhausted, the controller cluster service is offline, or a cluster switchover is triggered

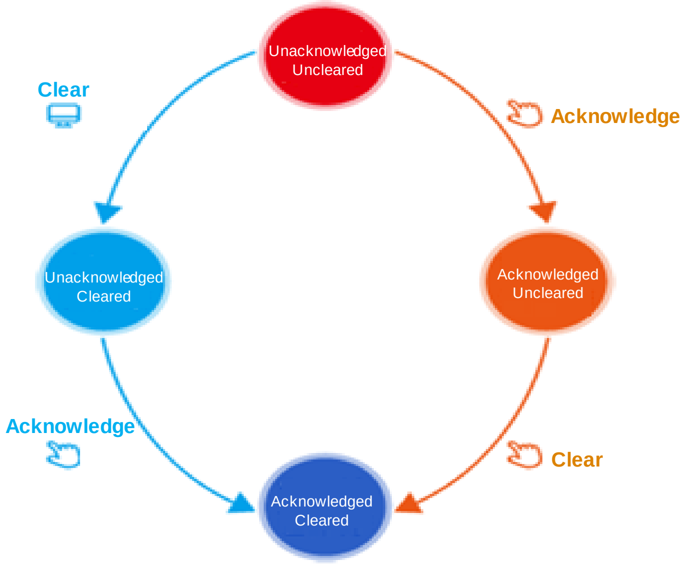
Alarms Displayed on eSight

eSight, as a gateway software component, provides powerful alarm management capabilities. On eSight, network maintenance personnel can centrally monitor NE or eSight alarms and locate and rectify faults in a timely manner, ensuring proper network running.

If the network scale is large, it is recommended that eSight features such as alarm consolidation and alarm correlation be used to optimize the capacity to locate the specific alarm and rectify network faults.

Alarms have different states and severities. Based on the alarm and event urgency, the following alarm severities are defined: warning, minor, major, and critical. Users can associate alarm types with the corresponding alarm severity. Figure 2-32 shows the alarm states.

Figure 2-32 Alarm clearance process



2.4.9 Port Mirroring

During network maintenance, you may need to obtain and analyze packets in some conditions. For example, if network maintenance personnel detect suspected attack packets, they need to obtain and analyze the packets without affecting packet forwarding.

The mirroring function copies packets on a mirrored port to an observing port without affecting packet processing on devices. Users can analyze the copied packets using the data monitoring device for network monitoring and troubleshooting.

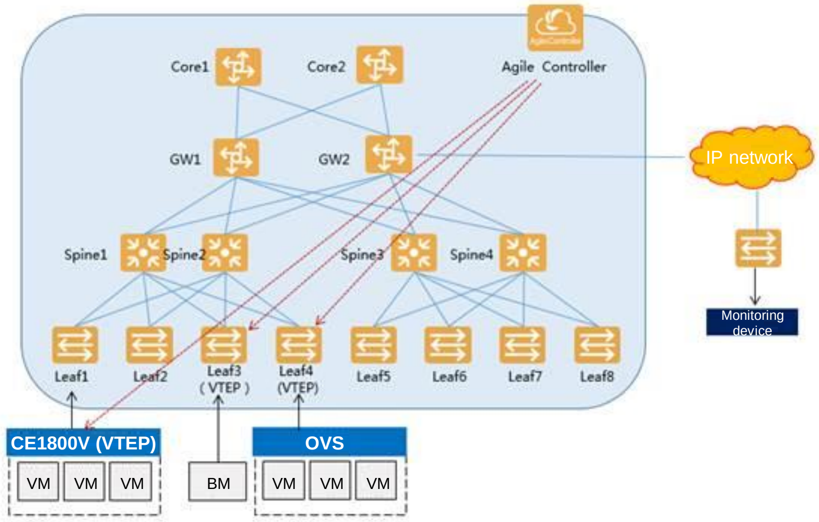
Mirroring is classified into the following types:

- Local port mirroring: copies packets normally forwarded on a switch port to another monitoring port on the switch. The packets are sent in original packet format.
- Remote port mirroring: copies packets normally forwarded on a switch port to another monitoring port on the switch. The original packets are added with a VLAN tag.
- Tunnel-based remote port mirroring: copies packets normally forwarded on a switch port to another monitoring port on the switch. The original packets are encapsulated using GRE.

Cloud data centers use the SDN solution. The network is composed of the overlay network and underlay network. The underlay network usually uses full-mesh networking. To prevent the forwarding capacity loss due to loop avoidance protocols such as STP, IP ECMP is deployed on the underlay network. This enables multi-link protection between devices and maximizes the oversubscription ratio, providing larger forwarding capacity.

Local port mirroring and remote port mirroring cannot traverse the underlay IP network. Tunnel-based remote port mirroring is recommended in the SDN solution.

Figure 2-33 Application of tunnel-based remote port mirroring



Tunnel-based remote port mirroring involves the following types of traffic:

- The CE1800V functions as the VTEP to remotely mirror packets passing through the ports connecting to VMs.
- The CE switch directly connecting to a physical server functions as the VTEP to remotely mirror packets passing through the physical port of the CE switch.

- The CE switch connecting to VMs through third-party OVS functions as the VTEP to remotely mirror packets passing through the CE switch based on flows (by matching the VM IP address).

2.4.10 Overlay Loop Detection

Although fewer loops occur on data center networks than on campus and enterprise networks, loops still exist. Loops occur on a data center network in the following situations:

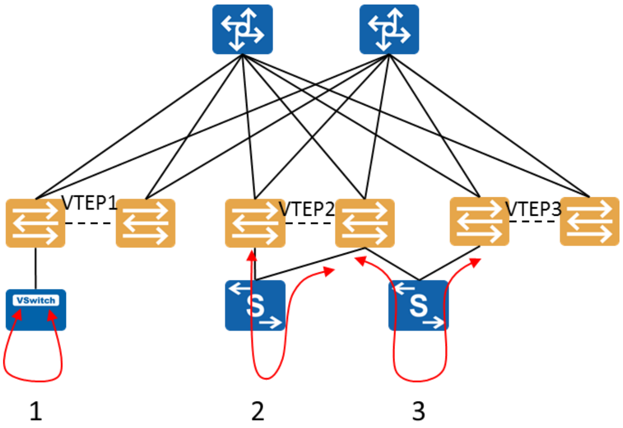
- In the network overlay scenario where virtual servers connect to the network, loops may occur among vSwitch ports on virtual servers.
- In the network overlay hosting scenario, loops may occur among ports on tenants' gateway switches or Layer 2 switches.

Huawei CloudFabric DCN Solution provides the loop detection function to solve the preceding problems. This function detects possible loops on data center networks to reduce impacts on service networks.

The loop detection function can detect the following types of loops:

- Loop on a single port in the single VTEP node scenario
- Loop among multiple ports in the single VTEP node scenario
- Loop among multiple ports in the multi-VTEP node scenario

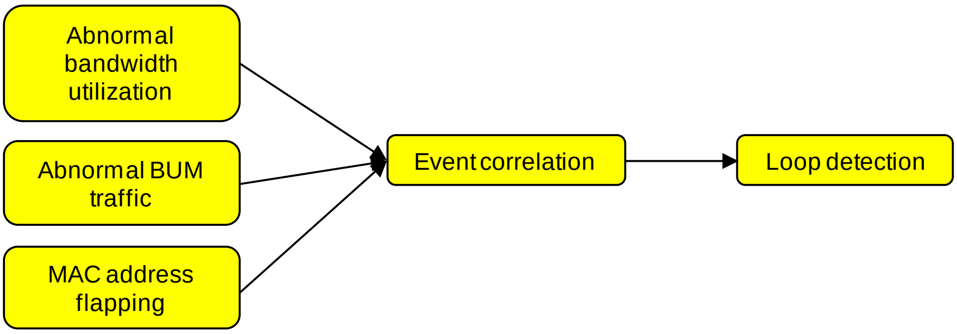
Figure 2-34 Loop scenarios



In the preceding loop scenarios, devices automatically detect exceptions, encapsulate alarm information into packets, and inform the Agile Controller-DCN of the exceptions. The Agile Controller-DCN isolates the fault source based on policies to reduce impacts on services.

Loops occur on a network beyond the control scope of the Agile Controller-DCN. The Agile Controller-DCN can only isolate faults, but cannot automatically rectify them. Administrators need to manually rectify the faults after receiving alarms.

Figure 2-35 Loop detection implementation



2.4.11 OpenStack Service Recovery

In the cloud-network integration scenario, computing, storage, and network resources are uniformly managed and delivered by the cloud platform (OpenStack). For network resources, multiple tenants share the same physical network but each tenant has an independent logical network. The logical networks of tenants are isolated through logical resources and do not affect services of each other.

Tenant administrators allocate logical networks on the cloud platform. Their capabilities of operating and maintaining logical networks are weak. Therefore, how to locate problems on logical networks of tenants is an urgent matter.

As the core NE for network management and control, the Agile Controller-DCN connects to the cloud platform in the northbound direction and connects to fabric NEs in the southbound direction. The Agile Controller-DCN can analyze cloud platform services and convert the service language into the language used by fabric NEs. The Agile Controller-DCN displays logical networks on the GUI and provides services, O&M, and fault location that are not supported by other cloud platforms.

Figure 2-36 and Figure 2-37 show implementation of service recovery.

Figure 2-36 OpenStack logical network model

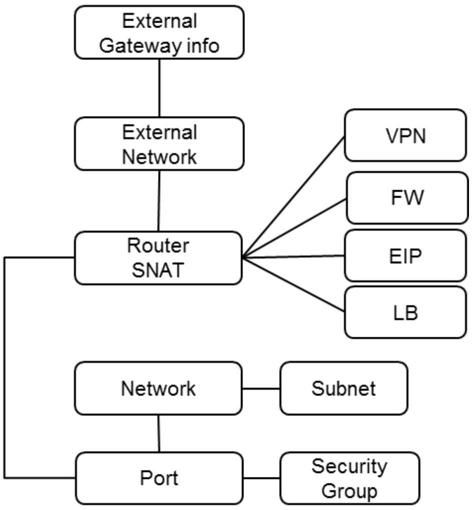
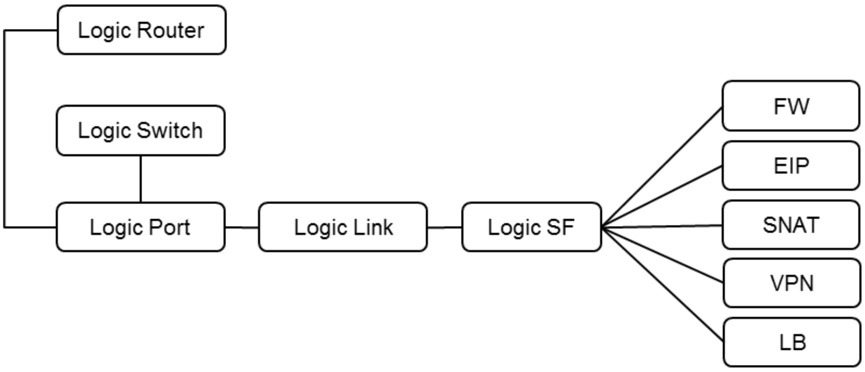


Figure 2-37 FaaS network model of the Agile Controller-DCN



The logical network of the OpenStack is mapped to the FaaS model of the Agile Controller-DCN using the following mapping algorithm:

- A router on the cloud platform is mapped to a logical router, which is similar to the VRF of the physical switch.
- A network is mapped to a logical switch, which is similar to a VLAN or VNI of the physical switch.
- A port is mapped to a logical port, which is similar to a sub-interface of the physical switch or a VM port of the virtual switch.
- A firewall is mapped to a logical service function, which is similar to a vSYS of the physical firewall.
- A load balancer is mapped to a logical service function, which is similar to a router domain of the physical load balancer.
- VASs such as VPN, EIP, and SNAT are mapped to corresponding features of a logical service function, such as IPSec tunnel, NAT server, and NAT pool.

OpenStack service recovery is performed to recover logical networks of tenants on the GUI of the Agile Controller-DCN, enable advanced functions such as mapping between three layer networks, path detection, and connectivity check based on the logical networks, locate faults on logical networks, and improve O&M efficiency.

2.4.12 ZTP and Device Replacement

In data centers that are newly constructed or expanded, a large number of access switches need to be deployed. If all access switches are configured manually, configuration errors are prone to occur and are difficult to locate, which will prolong provisioning of data center services. In the data center expansion scenario, inappropriate configuration may occur and negatively affect services that have been deployed.

Therefore, in scenarios where a large number of access switches are deployed, it is recommended that the ZTP function of the Agile Controller be used to automatically deliver device programs and startup files to access switches in batches, optimize the deployment process, and speed up service provisioning.

ZTP improves efficiency of device deployment, routine maintenance, and troubleshooting, and reduces labor costs. After a device plan is complete, the network administrator does not need to commission software of devices on site. After unconfigured devices are powered on, they can automatically connect to specified management devices and load system files including the configuration file, system software package, and patch file. ZTP implements fast device deployment.

2.4.12.1 Scenarios

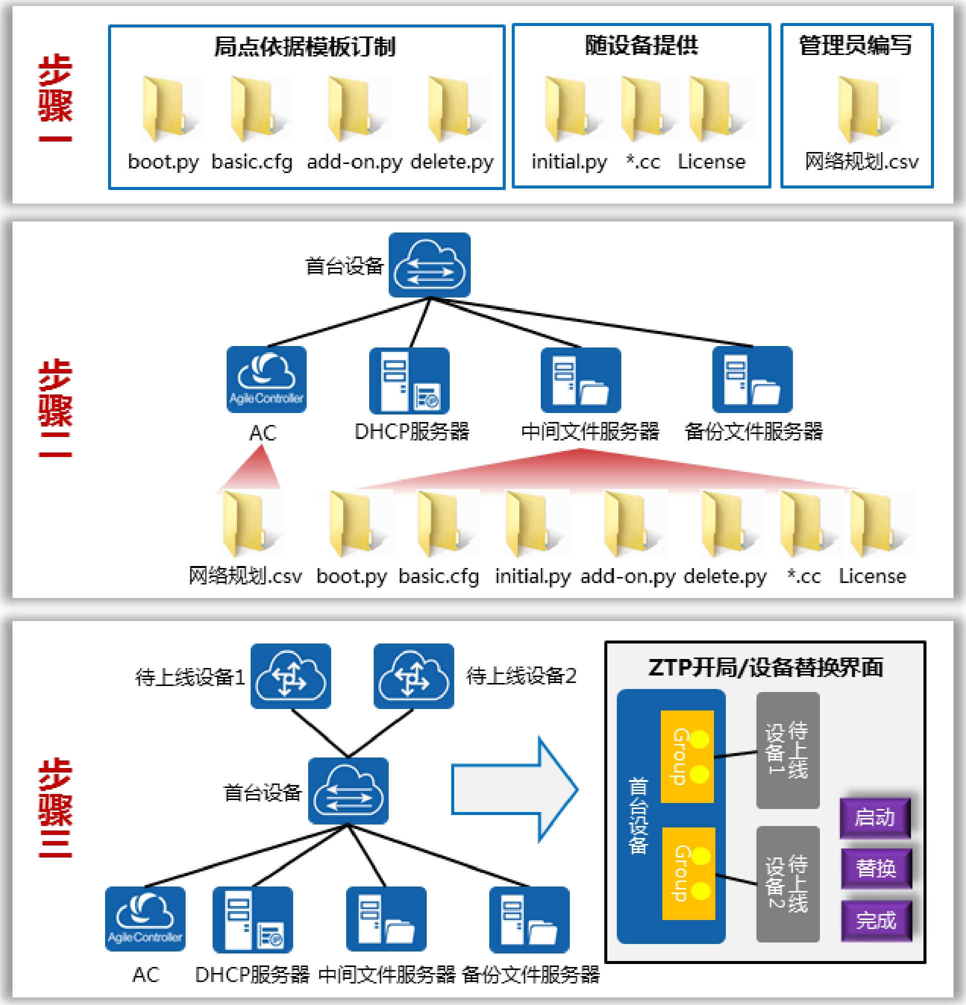
ZTP includes two scenarios where customers have different requirements on planning of network parameters of the devices to be deployed (such as the management IP address, device name, Layer 3 interconnection IP address, and routing protocol parameters).

- Automatic parameter allocation scenario (non-precise planning scenario): The Agile Controller automatically obtains network parameters from the resource pool and delivers them to the devices to be deployed. This scenario applies to data center sites where users have high requirements on deployment automation but low requirements on network parameter and topology planning.
- Manual parameter planning scenario (precise planning scenario): The network administrator manually plans the network parameters for devices to go online. This scenario applies to MAN sites where users have low requirements on deployment automation but high requirements on network parameter and topology planning.

2.4.12.2 Procedure

2.4.12.2.1 ZTP

Figure 2-38 ZTP process



The ZTP process is as follows:

1. Prepare deployment files.

   Deployment files can be classified into the following two types:

- Fixed files: Contents of the files are fixed and do not need to be modified. The files include the following:

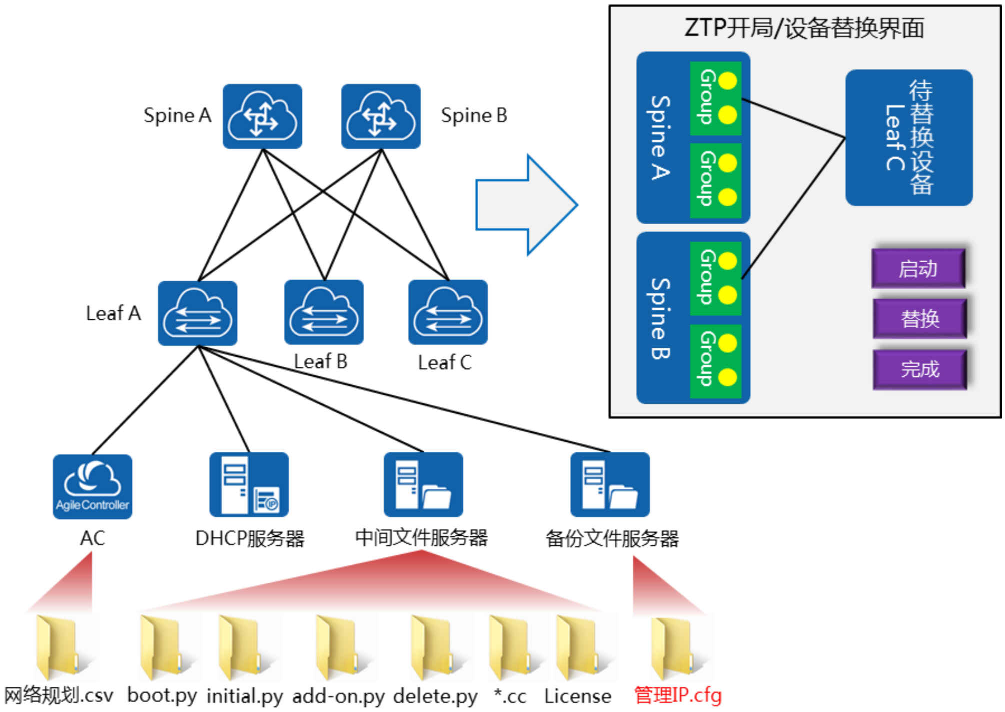
  - **initial.py**: initial configuration script file
  - ***.cc**: software version file
  - **License**: device license file
- Files customized for sites: Contents of the files vary by site and are customized based on templates. The files include the following:
  - **basic.cfg**: basic configuration file
  - **boot.py**: script file used for device startup
  - **add-on.py**: additional configuration script file
  - **delete.py**: configuration clearing script file
  - **网络规划.csv**: network plan file (not required in the automatic parameter allocation scenario)

2. Prepare a deployment environment.
   a. An administrator installs the Agile Controller, DHCP server, and intermediate file server and connects these servers to the first online device. In in-band management mode, the first online device must be a leaf device, and in out-of-band management mode, it is recommended that a spine device be the first online device.
   b. The administrator manually configures and enables Leaf A. If out-of-band management mode is used, a management switch needs to be manually configured and enabled.
   c. Install the devices to be online and connect cables. If out-of-band management mode is used, correctly connect the cables between the devices and the management switch.
   d. The administrator configures the DHCP Option field on the DHCP server.
   e. The administrator uploads deployment files to servers, as shown in Figure 2-38.
3. Perform automatic deployment.
   f. The administrator configures the Agile Controller to manage the first online device.
   g. The administrator creates a ZTP zone and configures the management network type, information about the intermediate file server, backup file server, and DHCP server, planning type (precise or non-precise), and resource pool (of management IP addresses, interconnection IP addresses, and VTEP IP addresses).
   h. After planning the network topology, the administrator clicks **启动** on the Agile Controller to start the ZTP process.
   i. The administrator powers on the devices to be online. After a device is started, the ZTP process is automatically started. After the automatic deployment is complete, the Agile Controller manages the device.
   j. After all devices on the network are automatically deployed, the administrator clicks **完成**.

2.4.12.2.2 Device Replacement

Device replacement relies on the deployment files, environment, and networking for ZTP.

Figure 2-39 Device replacement process



The device replacement process is as follows:

1. An administrator enters the IP address, user name, and password of the backup file server on the Agile Controller. (The device configuration file is named **管理IP.cfg** and stored on the server.)
2. The administrator selects the device to be replaced and clicks **替换** to start the device replacement process.
3. The administrator uninstalls the old device, installs the new device, connects cables (the cable connection must be the same as that of the old device), and powers on the new device. After the new device is started, a device replacement process is automatically started and then the new device is managed by the Agile Controller.
4. After the new device successfully goes online, the administrator clicks **完成**.

2.4.12.3 Deployment Description

Table 2-2 Description of ZTP and device replacement

Item	Description
Device model	Only CE switch hardware and versions used in CloudFabric V300R002C00 are supported. Switches of earlier versions are not supported.

Item	Description
Management type	The following management types are supported: • In-band management: No independent management switch is deployed. The management network and service network share service network interfaces. • Out-of-band management (MEth interfaces): An independent management switch is deployed. The management network and service network are separately deployed. The MEth interfaces are bound to device management IP addresses. • Out-of-band management (VPN isolation): An independent management switch is deployed. The management network and service network are separately deployed, and common GE interfaces are connected to the management VPN network.
Networking type	• ZTP applies to single-node, stack, and M-LAG systems but not the SVF system. ZTP operations can be performed for only one system type each time. • A member switch in a stack cannot be separately replaced. To replace a member switch in a stack, the administrator needs to manually configure a member ID on the new device and connect the new device to the stack. Based on the stack capabilities, the new device will then synchronize configuration from the old device.
Deployment of the file server and DHCP server	• The SFTP server software for the intermediate and backup file servers and DHCP server software need to be provided by users. Third-party open-source software is recommended. • The Agile Controller, intermediate and backup file servers, and DHCP server can be deployed in different VMs on the same physical server to reduce the number of physical servers. The intermediate and backup file servers can be deployed in the same VM.
File transmission protocol	The intermediate and backup file servers support SFTP.
Capability requirement of the DHCP server	The DHCP server must support the following Option fields: • Option 1: contains the temporary IP address range used by a device to go online. • Option 3: contains the egress gateway IP address of a DHCP client. • Option 66: contains the IP address of the intermediate file server. • Option 67: contains the path and name of the **boot.py** file.
Limitation on the number of layers during ZTP	ZTP implements layered deployment, and theoretically has no limitation on the number of layers. Generally, two to three layers are involved during actual deployment.
Limitation on the number of devices that concurrently go online during ZTP	During ZTP, devices at each layer go online in distributed mode. The number of devices that can go online concurrently is not limited by the Agile Controller but depends on the file delivery capabilities of the file servers.

Item	Description
Limitation on the number of concurrently replaced devices	Only one device can be replaced at a time. Replacing devices in batches is not supported.
Requirements on configuration file backup and management	Device replacement requires network management software that provides capabilities of automatically uploading and managing configuration files. (eSight is recommended.)

2.4.13 Service Quality Visibility (eSight)

In-line network service quality visibility is a technology that calculates network quality based on iPCA. iPCA is a new network quality detection solution. Unlike traditional detection solutions, iPCA detects network quality based on actual service flows. The detection result and fault diagnosis result are more accurate.

Huawei CloudFabric DCN Solution supports detection on service packet loss and transmission delay. The solution uses CE6880EI and eSight to detect network-grade packet loss and delay.

iPCA can detect a specific service flow (based on 5-tuple including IP addresses, protocol number, and port numbers, or 5-tuple+overlay information). It can detect the device that discards packets, time when packets are discarded, and type of discarded packets.

The following formula can be used to check whether packet loss occurs:

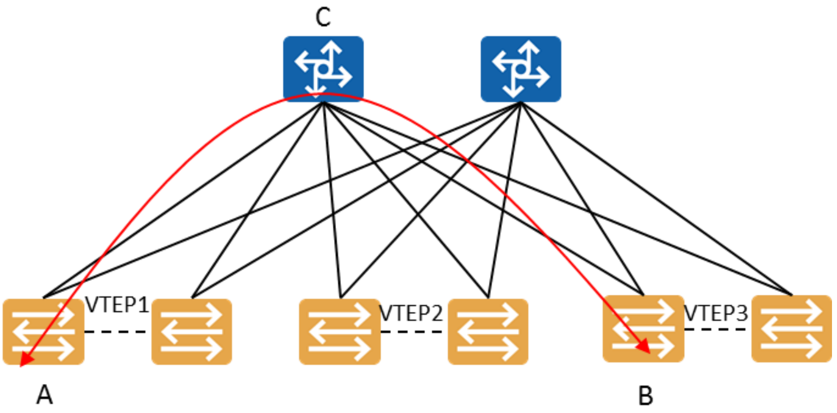
Packets received by the system + Packets generated in the system = Packets sent from the system + Packets absorbed by the system

If the preceding condition is met, the system runs normally; otherwise, packet loss occurs in the system.

2.4.13.1 iPCA Implementation

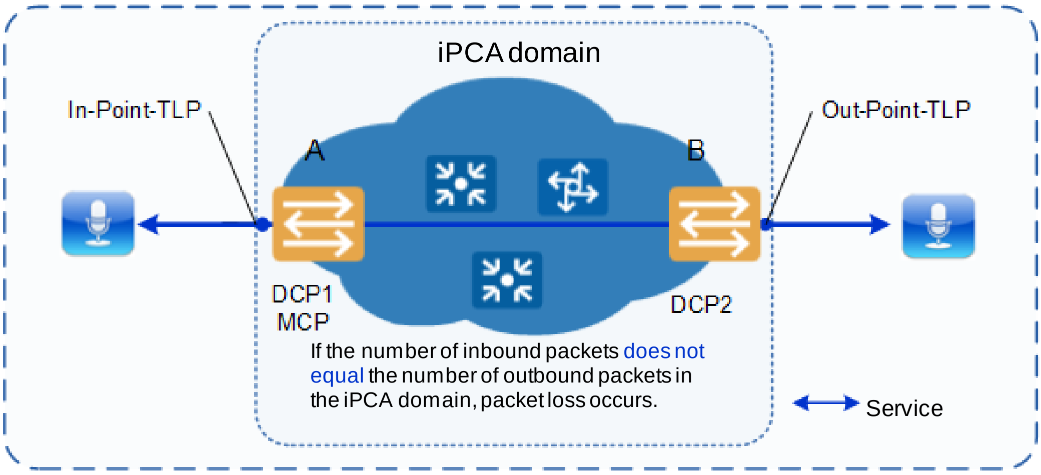
If services are transmitted between device A and device B on a fabric network and the services are interrupted intermittently, data center O&M personnel can create a detection task between device A and device B on eSight to check whether packet loss occurs on the network where the services are transmitted. If packet loss is detected, O&M personnel can carry out fault diagnosis to determine the device on which packets are lost and then rectify the fault.

Figure 2-40 Actual service quality detection solution



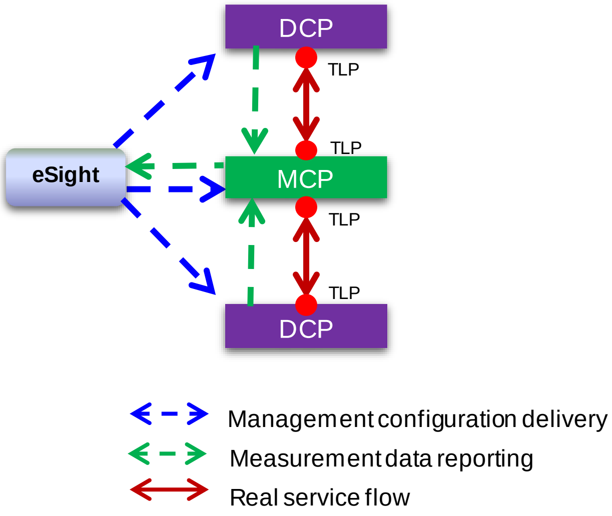
During the detection process, the network where the services are transmitted can be regarded as an iPCA domain. Figure 2-41 shows the simplified service quality detection solution.

Figure 2-41 Simplified service quality detection solution


iPCA domain
In-Point-TLP
Out-Point-TLP
A
B
DCP1
MCP
DCP2
If the number of inbound packets does not equal the number of outbound packets in the iPCA domain, packet loss occurs.
Service

After abstraction, the simplified network service solution involves Target Logical Ports (TLPs), Data Collecting Point (DCP) NEs, and Measurement Control Point (MCP) NEs. Figure 2-42 shows their logical relationship.

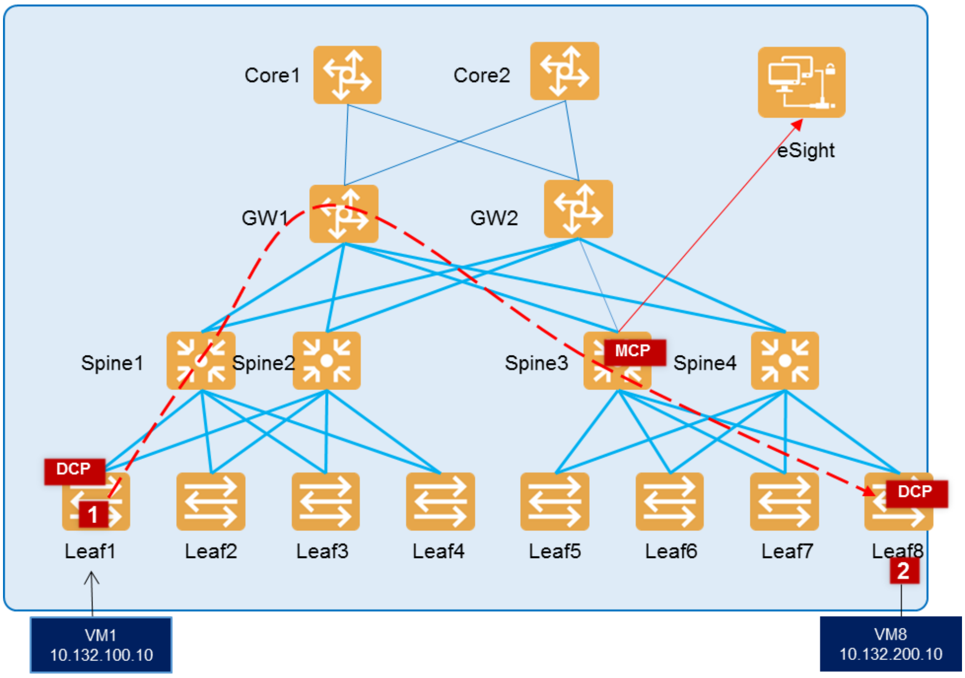
Figure 2-42 Abstracted service quality detection solution



- TLP: network device port that functions as an observing point located at the edge of a network. A TLP collects statistics on packet loss, and generates statistics on the numbers of sent and received packets, sent and received traffic, and timestamps. For a target flow in an iPCA domain, the upstream TLP is called In-Point-TLP, and the downstream TLP is called Out-Point-TLP.
- DCP: data collecting device in the iPCA system. A DCP manages and controls TLPs, collects statistics generated on TLPs, and sends statistics to the MCP.
- MCP: control device in the iPCA system. The MCP summarizes and calculates statistics sent from DCPs, and reports calculation results to eSight.
- eSight: network management software. eSight summarizes iPCA data sent from the MCP and displays detection results to users in graphs.

2.4.13.2 iPCA Detection Process

Figure 2-43 iPCA detection process



The iPCA detection process is described as follows:

1. Enable the NTP time synchronization function on the entire network to ensure detection result accuracy.
2. Use VM/BM location function implemented using the computing resource visualization function to detect ports of the source and destination leaf nodes (VTEPs). The two leaf nodes function as DCPs.
3. Select the MCP. It is recommended that a network device that detection packets may pass through be selected.
4. On eSight, inject detection packets to the source DCP (VTEP) through the Agile Controller-DCN. The detection packets are transmitted to the destination DCP (VTEP) through the fabric network. The source and destination DCPs generate detection information and aggregate the information to the MCP. The MCP then generates a report and sends the report to eSight. eSight displays the report to users.

2.4.14 Intelligent O&M (FabricInsight)

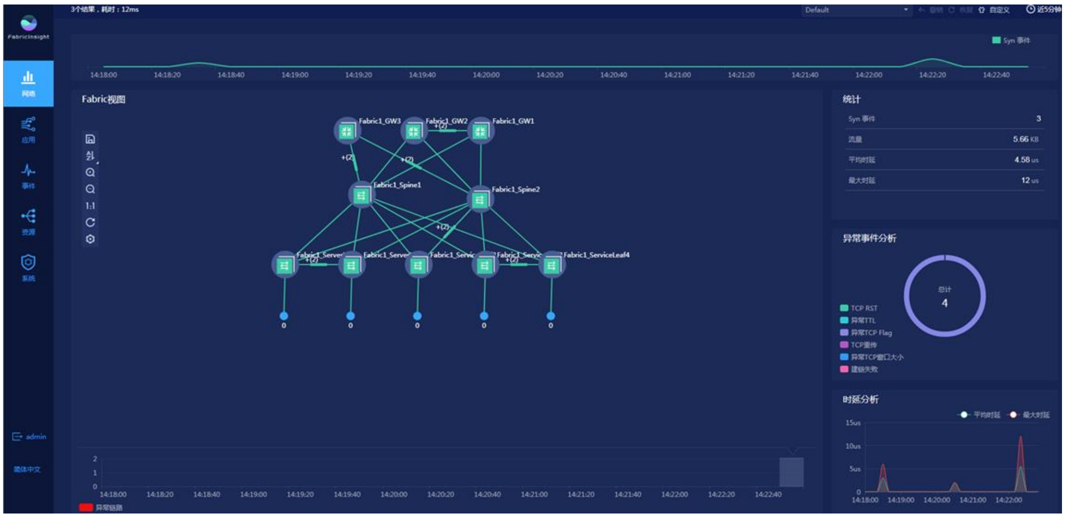
FabricInsight provides intelligent O&M to address the following problems:

- Reproduction and source tracing of service traffic errors are unavailable as network traffic is invisible. Therefore, when the service quality (including delay and jitter) deteriorates, the network connectivity may be abnormal. FabricInsight supports network visualization. It provides high-speed collection and long-term storage of all traffic data to enable network administrators to discover and rectify network faults before they receive complaints on services. Additionally, FabricInsight also analyzes trends of and potential faults in application traffic on networks, facilitating application optimization and proactive O&M.

- Users' access to services triggers multiple east-west traffic flows, and any abnormal east-west traffic flow deteriorates user experience. When a fault occurs, streamlining service flows (classification and sorting the sequence by time) is a huge challenge for network O&M personnel. FabricInsight supports application visualization. It classifies VMs by application using VM templates manually imported or obtained from a third-party system or through machine learning. It also analyzes traffic correlations and automatically generates diagrams showing the time sequence of service flows between components. If a network fault occurs, FabricInsight can automatically assess the impact of the fault on services.

2.4.14.1 Network Visualization

Figure 2-44 Network visualization page



统计
Syn 事件 3
流量 5.66 KB
平均时延 4.58 us
最大时延 12 us

Table 2-3 Description of network visualization functions

Function	Description
Summary statistics	Displays statistics about the SYN event quantity, traffic volume, and average and maximum delays within a specified time period.
Abnormal event analysis	Displays abnormal TCP events occurring on the network within a specified time period, including the TCP RST, TCP retransmission, TCP Flag packet exception, TCP window exception, and abnormal TTL packet.
Delay analysis	Displays results of comparison and analysis of the trends of the average and maximum delays of TCP events on the network within a specified time period.

Function	Description
Fabric topology display	• Collects the number of abnormal links with a delay exceeding the threshold within a specified time period. • Displays trend distribution of the number of abnormal links (by delay) within a specified time period. • Displays the current Fabric network topology. • Displays the abnormal links with a delay exceeding the threshold in the topology over a specified time period. • Displays information about a link within a specified time period, including the total number of TCP event packets through the link (bidirectional), statistics about traffic through the link (bidirectional), and average and maximum forwarding delays of the link (bidirectional). • Displays information about a device within a specified time period, including the device name and IP address, total number of TCP event packets passing through the device, and statistics about traffic passing through the device. • Collects and displays the number of active IP addresses used to connect to leaf switches within a specified time period.
Flow trace	Performs flow trace on a link.

2.4.14.2 Application Visualization

Figure 2-45 Application visualization page

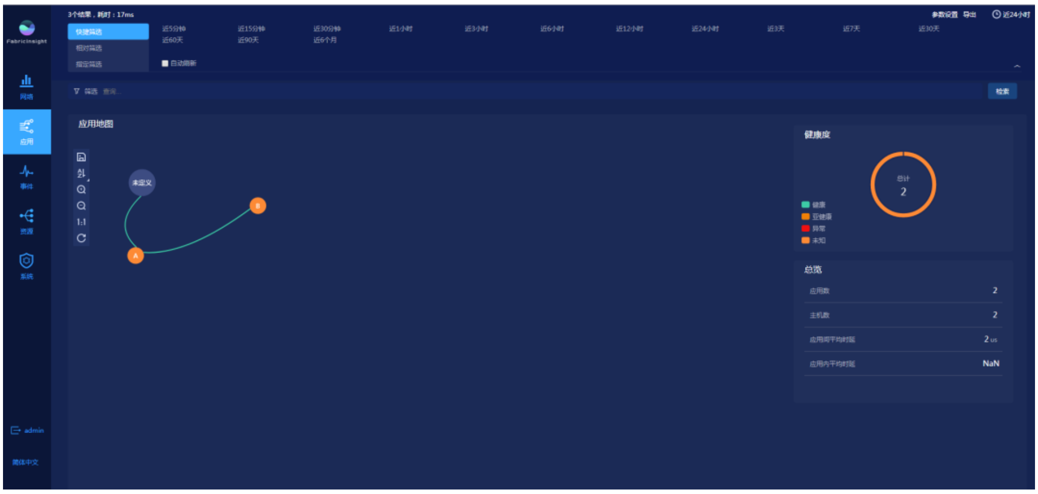


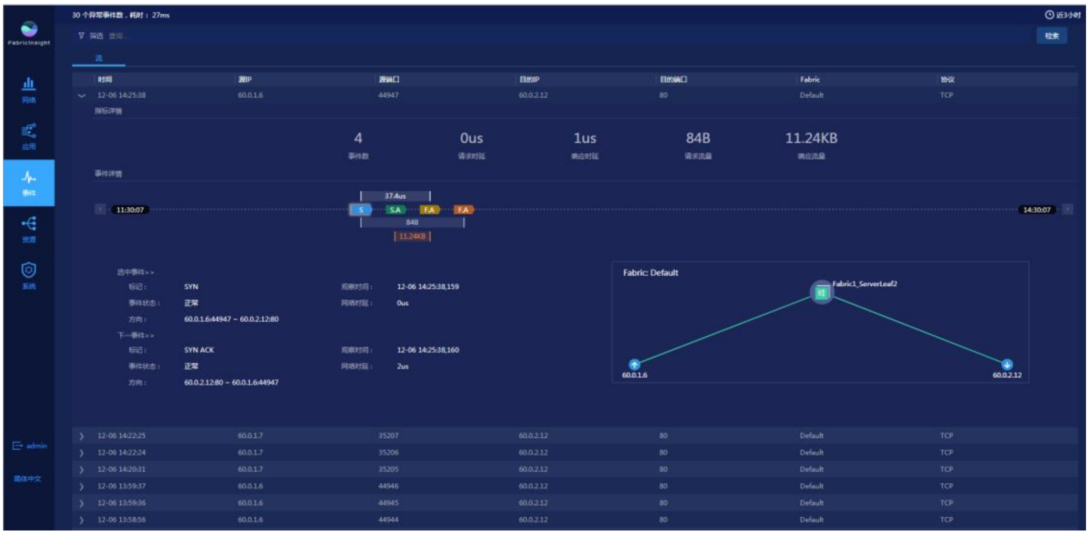
Table 2-4 Description of application visualization functions

Function	Description
Application health analysis	• Defines the application health threshold based on the delay. • Collects application health data within a specified time period and displays distribution of the data.

Function	Description
Inter-application interaction analysis	• Displays interaction relationship between applications in a topology (application map). Only the applications involved in flow interaction in the current query period are displayed. • Displays application information on pages. The information includes the application name, number of clusters, number of hosts, total traffic volume, number of flow events, average delay, and maximum delay. • Displays connections between applications as well as indicators of application access traffic in different directions, including the delay, traffic volume, and number of flow events. • Queries inter-application access data based on the following dimensions: application name, application health status, and host (VM) IP address. The following data can be queried: number of application clusters, average delay, maximum delay, and traffic volume. Fuzzy search is also supported. • Displays summary statistics about inter-application access, including the number of applications involved in flow event interaction within a specified time period, number of hosts, average inter-application interaction delay, and average intra-application interaction delay.
Intra-application interaction analysis	• Displays the number and trend abnormal logical interaction links between hosts over a specified time period. • Displays an intra-application interaction topology (interaction relationships between clusters and between hosts of an application) within a specific time period (a bar in a bar graph) on the timeline. • Displays statistics about hosts in an application cluster on a tab page, including the host IP address, listening port (port provided when a host functions as the provider), total number of interaction IP addresses, number of IP addresses involved in abnormal interaction, and total number of connections (number of flow events). • Displays connections between hosts as well as indicators of host access traffic in different directions, including the delay, traffic volume, and number of flow events. • Displays correlations between applications and networks, including details about the flow events occurring on a logical link within a specified time period and paths that a flow traverses. • Queries intra-application access data based on the host IP address or listening port. The following data can be queried: number of flow events, number of abnormal logical links, delay, and traffic volume. • Displays summary statistics about intra-application access, including the number of link setup events (SYN/SYNACK), number of link disconnection events (FIN/RST), intra-application interaction traffic, and average and maximum delays within a specified time period. • Collects information about the intra-application average network delay over a specified time period and displays the information in a curve.

2.4.14.3 Event Analysis

Figure 2-46 Event analysis page



30 个异常事件数，耗时：27ms

时间	源IP	源端口	目的IP	目的端口	Fabric	协议
12-06 14:25:38	60.0.1.6	44947	60.0.2.12	80	Default	TCP

指标详情

4	0us	1us	84B	11.24KB
事件数	请求时延	响应时延	请求流量	响应流量

事件详情

选中事件>>

标记：SYN　观察时间：12-06 14:25:38,159

事件状态：正常　网络时延：0us

方向：60.0.1.6:44947 ~ 60.0.2.12:80

下一事件>>

标记：SYN ACK　观察时间：12-06 14:25:38,160

事件状态：正常　网络时延：2us

方向：60.0.2.12:80 ~ 60.0.1.6:44947

时间	源IP	源端口	目的IP	目的端口	Fabric	协议
12-06 14:22:25	60.0.1.7	35207	60.0.2.12	80	Default	TCP
12-06 14:22:24	60.0.1.7	35206	60.0.2.12	80	Default	TCP
12-06 14:20:31	60.0.1.7	35205	60.0.2.12	80	Default	TCP
12-06 13:59:37	60.0.1.6	44946	60.0.2.12	80	Default	TCP
12-06 13:59:36	60.0.1.6	44945	60.0.2.12	80	Default	TCP
12-06 13:58:56	60.0.1.6	44944	60.0.2.12	80	Default	TCP

Table 2-5 Description of event analysis functions

Function	Description
TCP flow analysis	Collects and displays details about abnormal flows within a specified time period. • Collects statistical data of TCP flows based on 5-tuple information (source IP address, source port, destination IP address, destination port, and transmission protocol). • Displays measurement data of a single flow by time, source IP address, source port, destination IP address, destination port, or flow status. The measurement data includes the number of TCP events, request delay, response delay, request traffic volume, and response traffic volume. Data in a table cannot be sorted.
TCP flow event analysis	Displays a timeline of TCP events in a flow with abnormal events within a specified time period as well as event details. • Displays details about a single TCP event (SYN or SYNACK), including the TCP flag, occurrence time, direction (IP-to-IP), and network delay. • Displays the topology of the paths that a flow traverses as well as the hop-by-hop delay. • Displays request traffic, response traffic, and interval between link setup and link disconnection (SYN and FIN) on the timeline of TCP events. Restriction: Paths cannot be computed for ERSPAN packets reported by service loopback interfaces for which VXLAN Layer 3 gateways are configured.