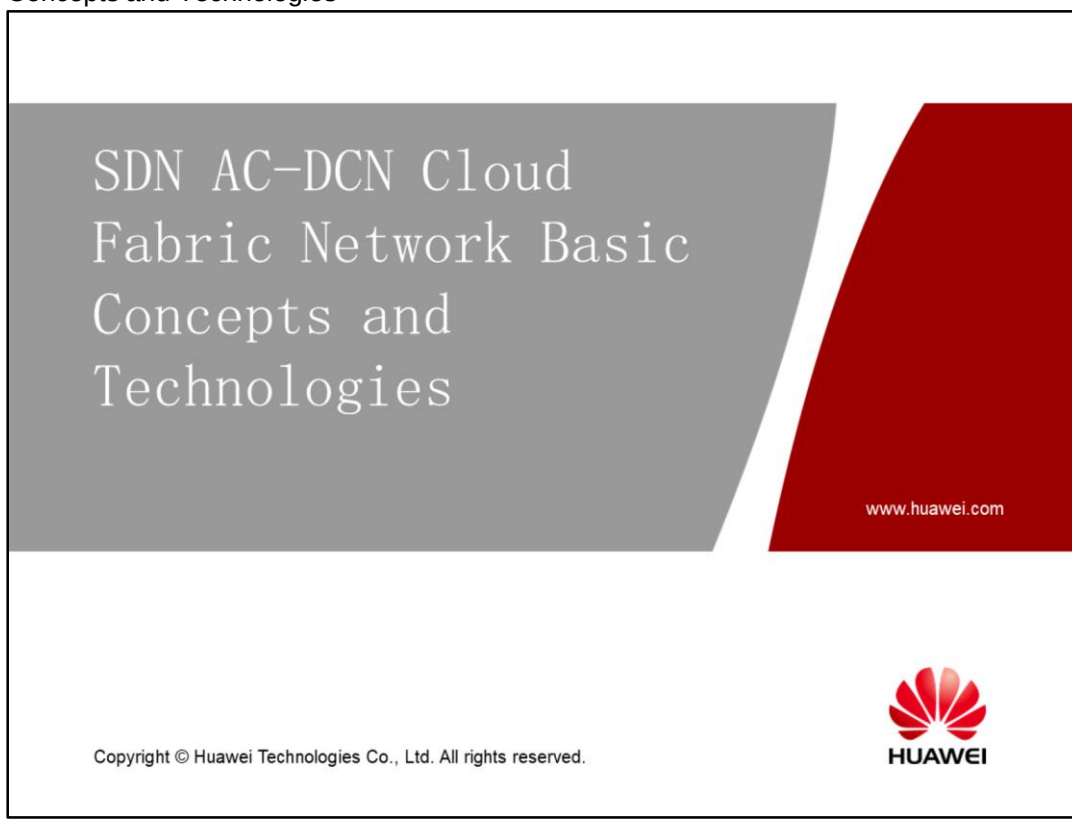
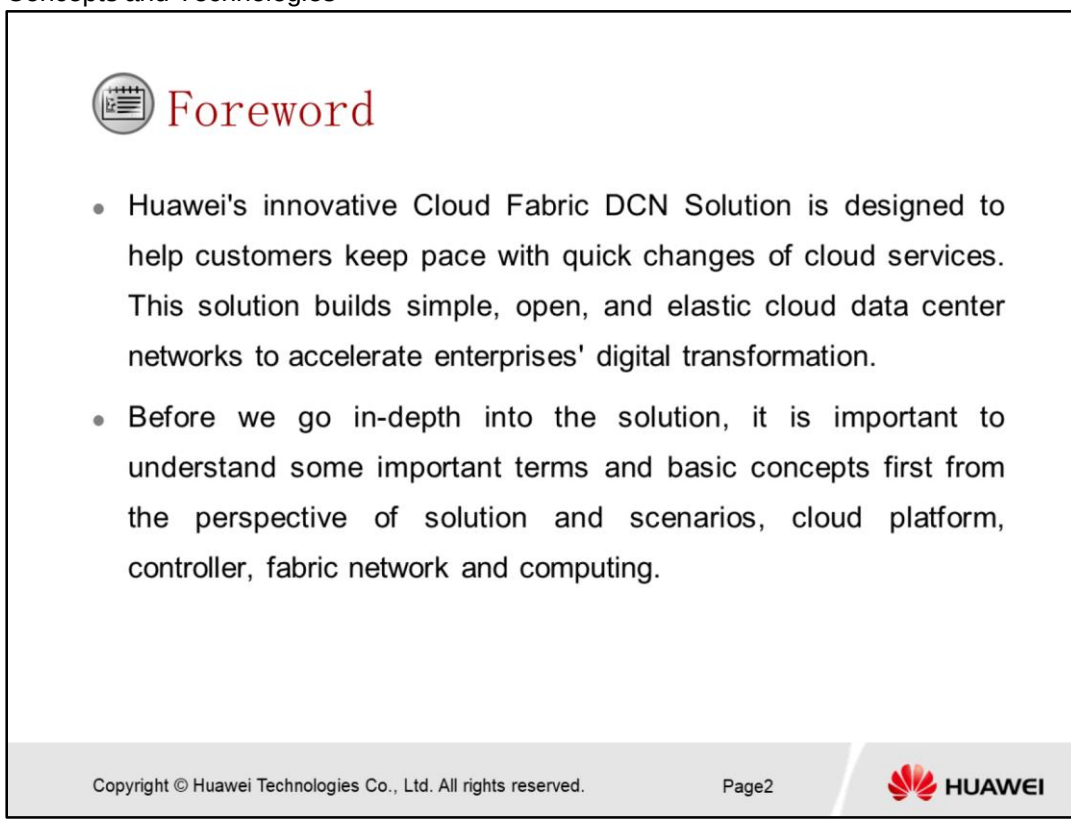
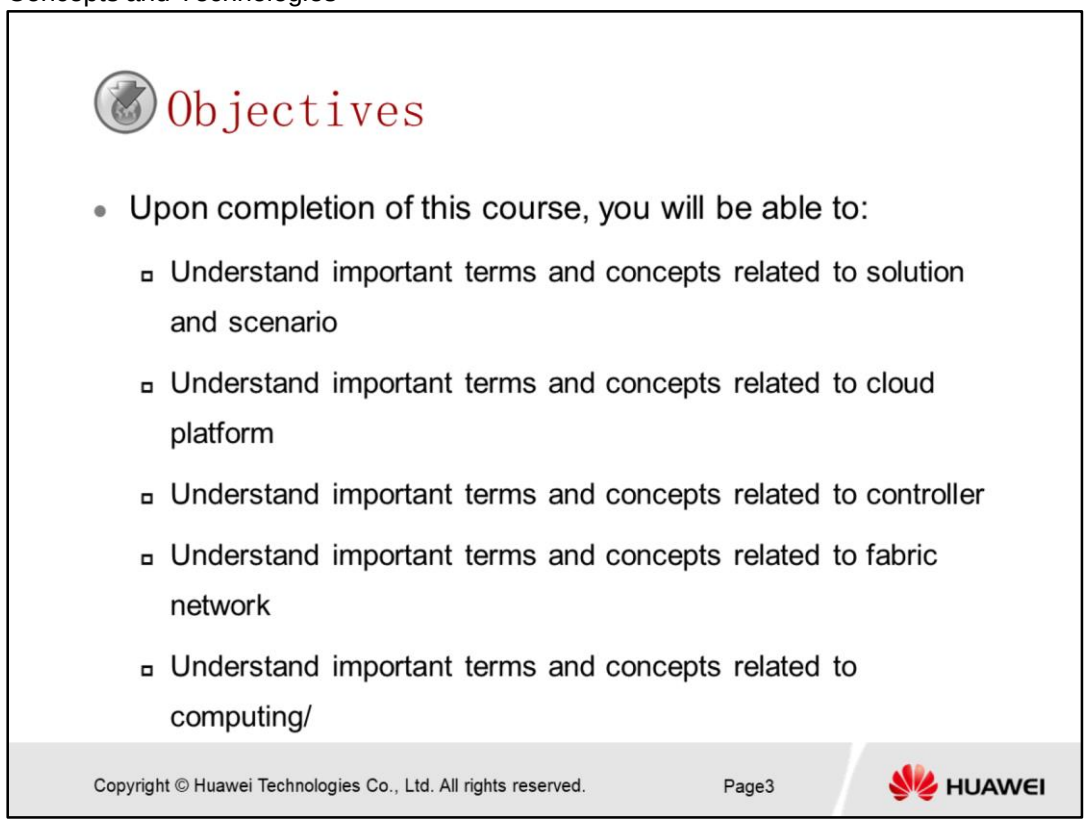
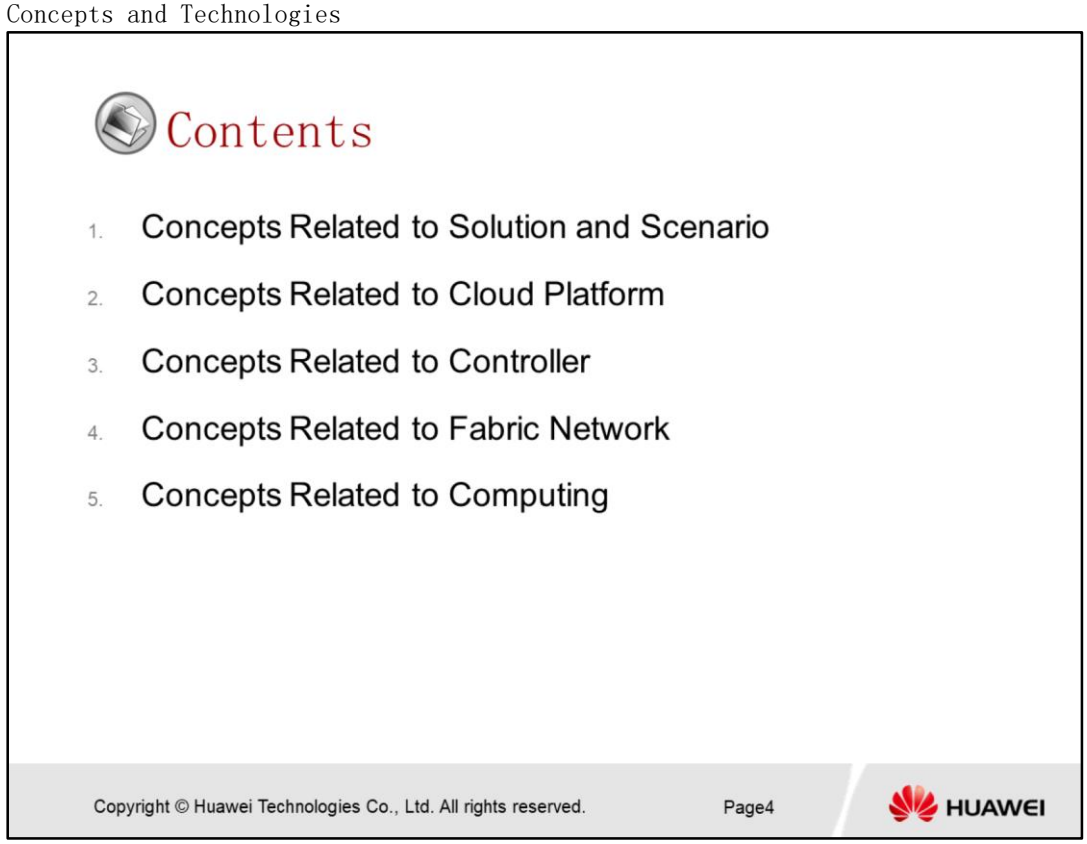
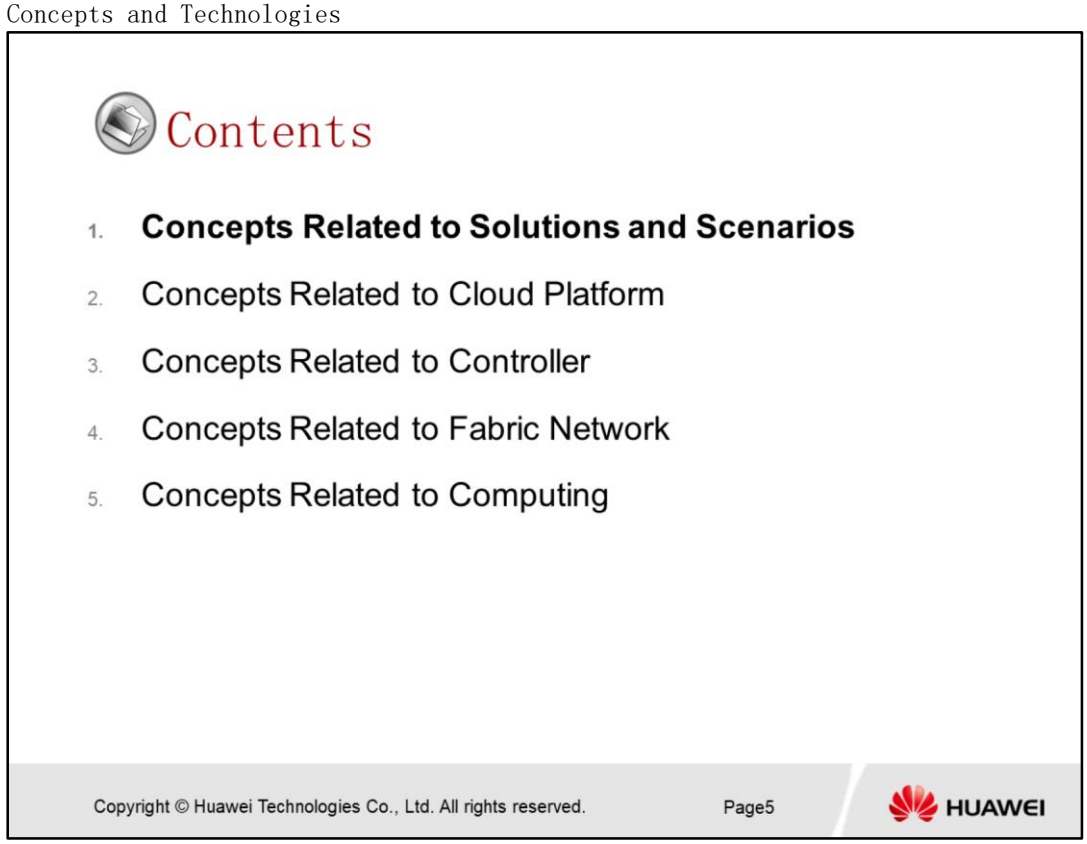


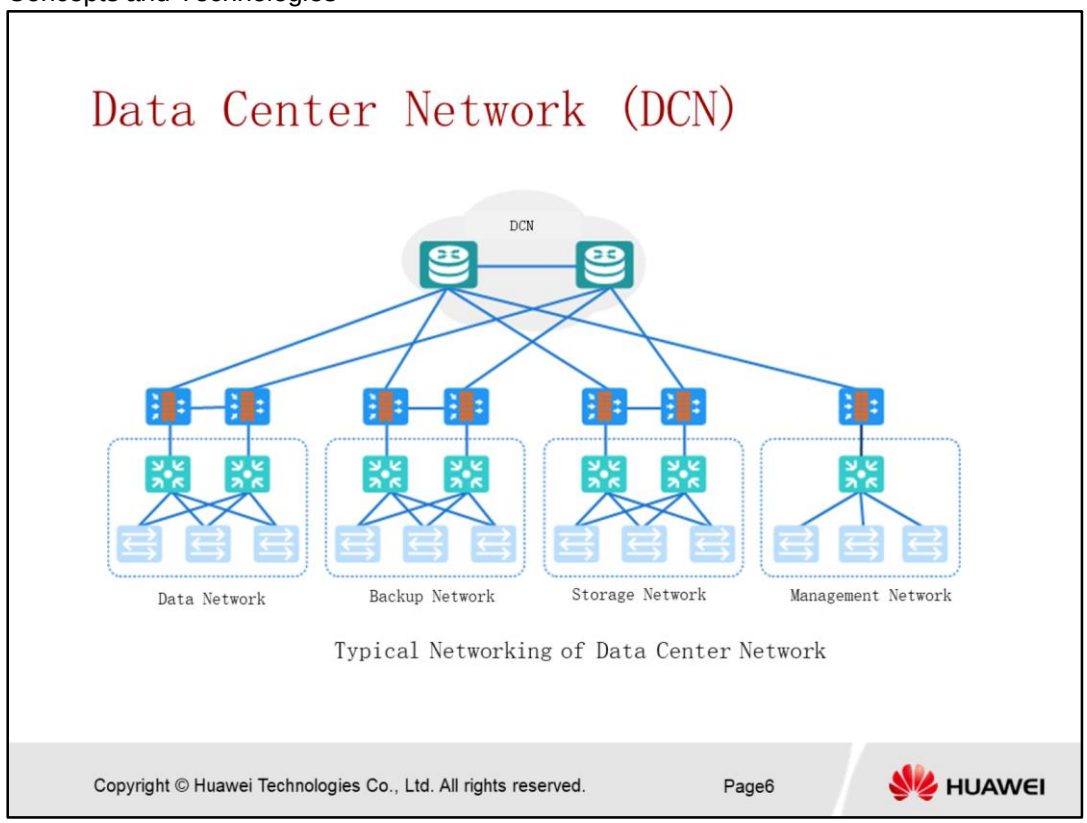




1. Concepts Related to Solution and Scenario
2. Concepts Related to Cloud Platform
3. Concepts Related to Controller
4. Concepts Related to Fabric Network
5. Concepts Related to Computing



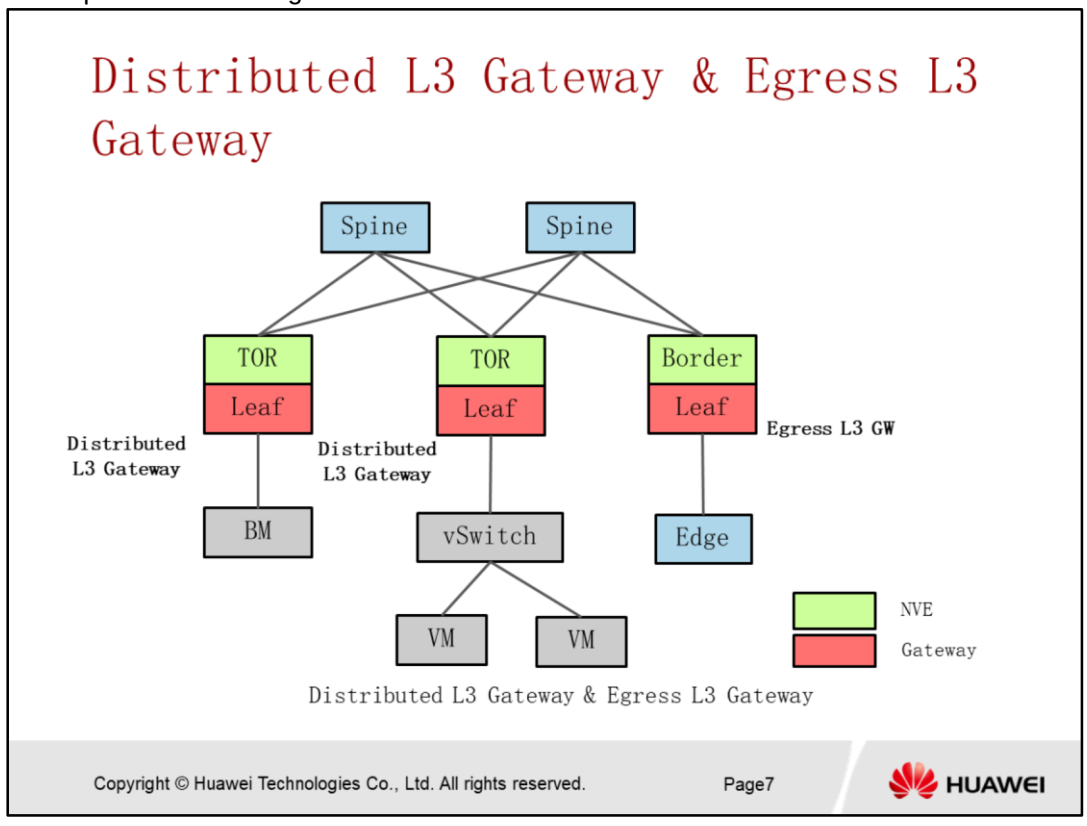




- Data Center Network (DCN) is the basic infrastructure of data center which normally comprises computing servers, storages, communication systems, intercommunication system and management systems etc.

- Multiple data center network can be located separately in geographical regions connecting branches or different organizations. Besides, data center network is also connected to interne to allow the organization or enterprise to be able to access to internet.

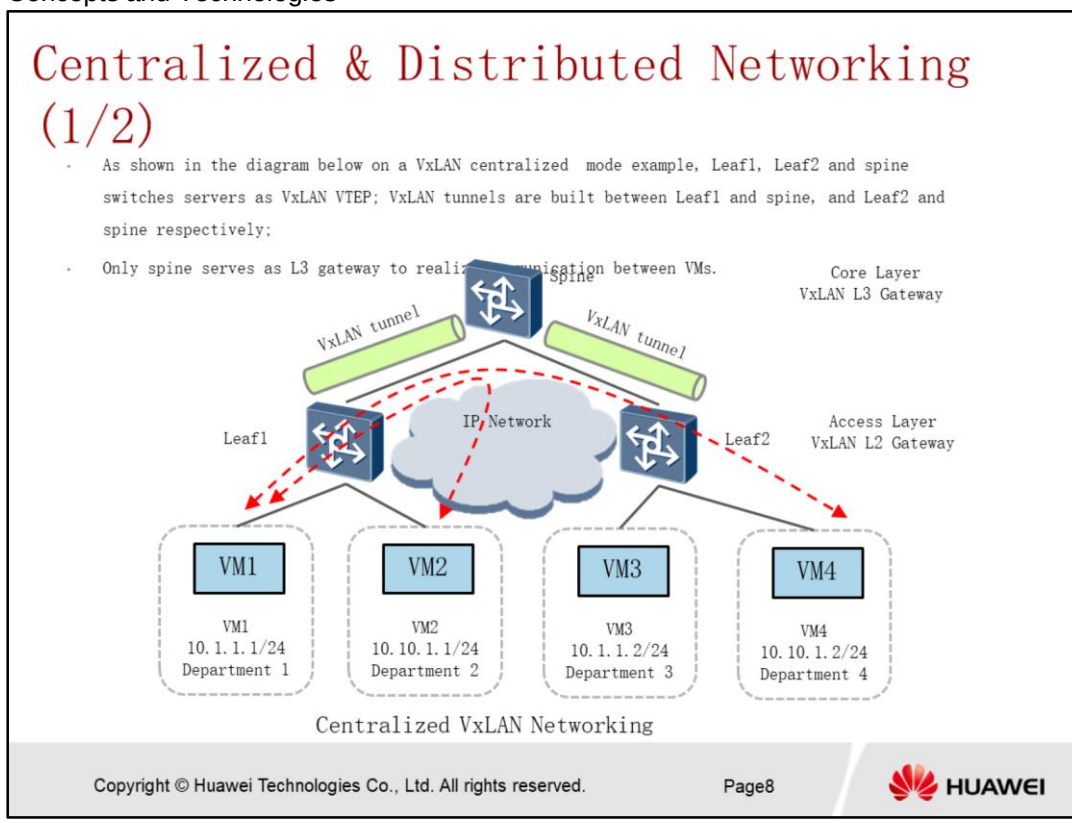
- The internal data center network traffic is typically increasing rapidly in east west traffic and thus the requirement towards data center network is mostly on high expansibility, low cost, high bandwidth, high efficiency and high requirement on server traffic isolation; thus, the conventional 3-layers hierarchical network cannot meet the growing requirement anymore; thus, the implementation of SDN and VxLAN in DCN is to achieve these requirements.



- A VXLAN distributed L3 gateway (east-west gateway) connects to servers or VMs and provides addressing for communication between VMs in a data center.

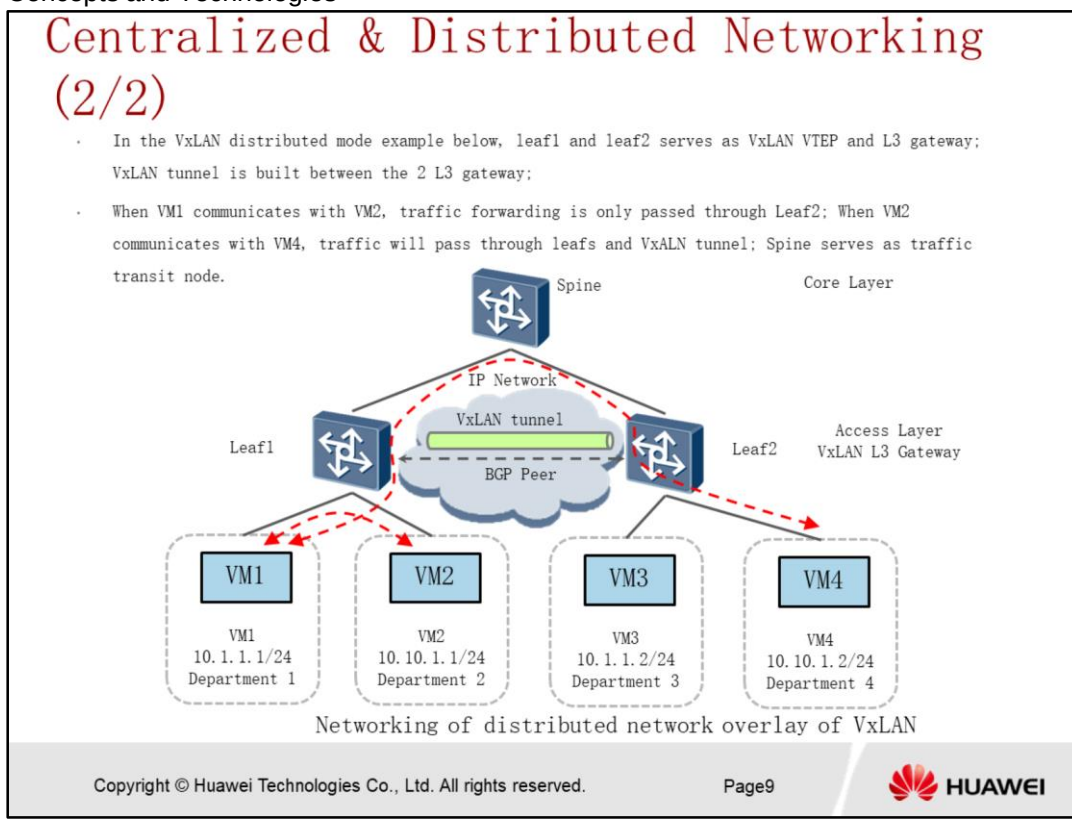
- A VXLAN egress L3 gateway (north-south gateway) connects to external edge devices (connecting to the Internet or external private networks) and provides addressing for the data center network to access the Internet or an external private network.

- Diagram above gives an example of the VxLAN L3 gateway and exit gateway in the distributed network overlay structure.

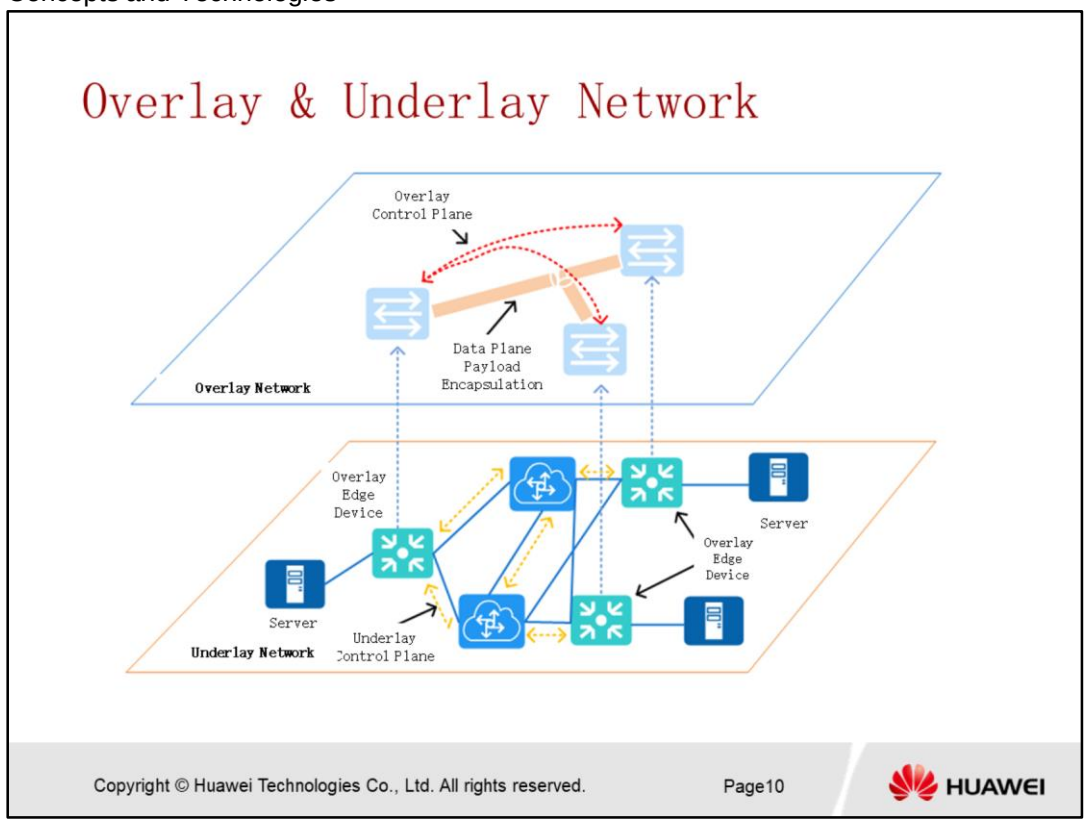


- Centralized mode
  - In VxLAN network. L3 gateway function is centralized on one or one group of switches
  - Leaf switches who is connected to firewall, load balancer, and various servers only serves as L2 gateway.

- Note
  - L3 gateway serves to allow communication between different VMs
  - L2 gateway serves to allow L2 communication in the same VxLAN network allow non-VxLAN traffic to access to a VxLAN network.



- **Distributed Mode:**
  - In Network Overlay distributed VxLAN network, all leaf physical switches are equipped with L3 gateway functions; Spine functions as traffic forwarding node, and does not function as VTEP;
  - In Hybrid Overlay distributed VxLAN network, certain leaf physical switches and all vSwitches are equipped with L3 gateway functions; Spine functions as traffic forwarding node, and does not function as VTEP;

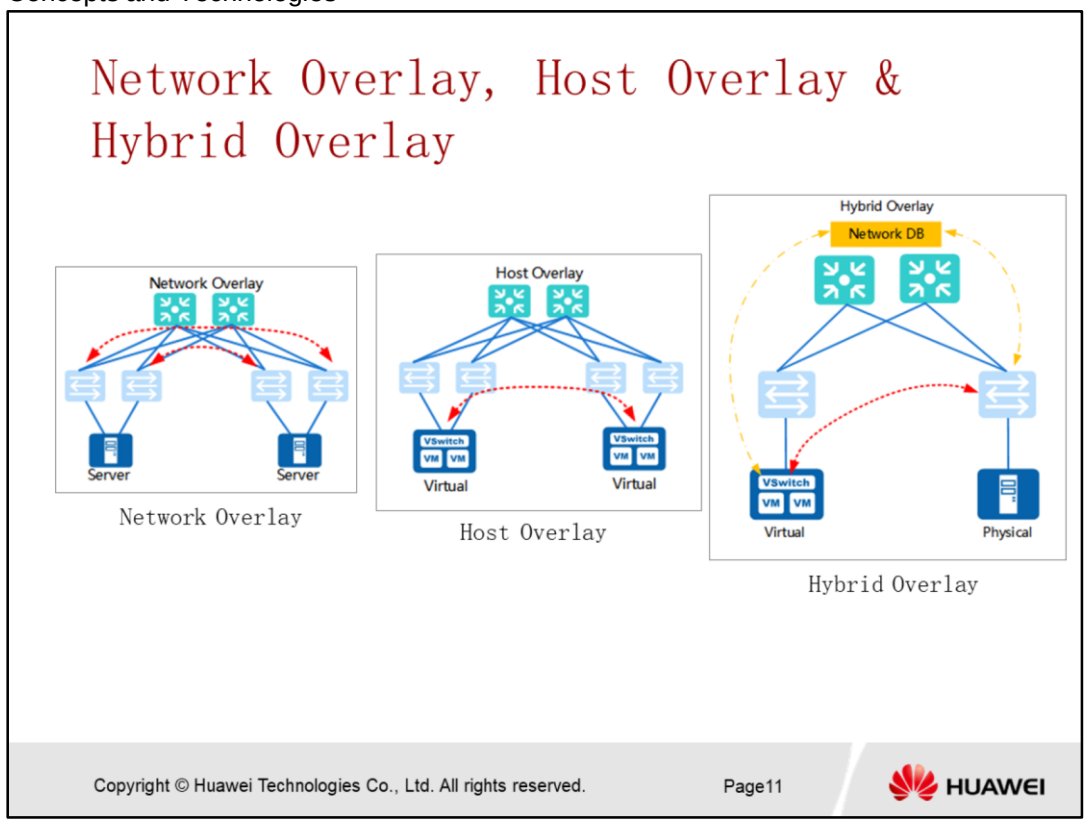


- Overlay can be understood as stacking concepts onto the physical network, which means that a logical network is defined by using overlay concepts to overcome network problem encountered in physical network. Overlay is a network technology which is used to carry a L2 data over a conventional L3 packet to be forwarded through L3 network.

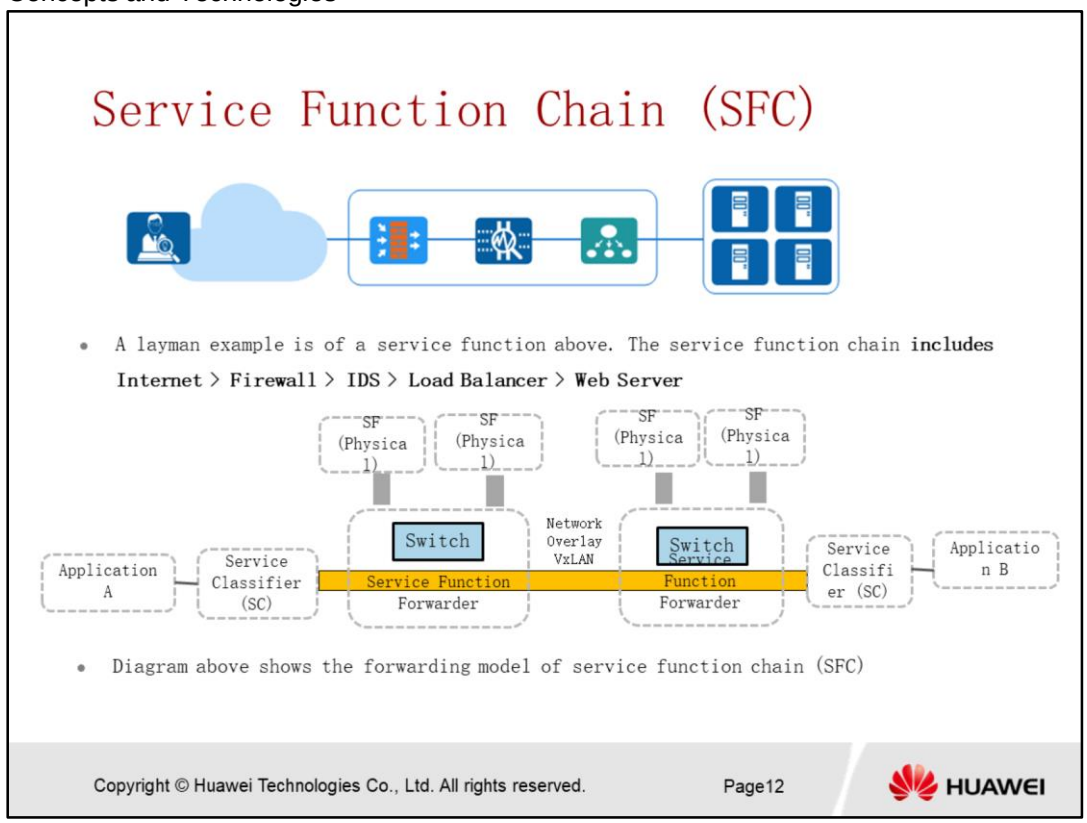
- To realize overlay technology, various overlay encapsulation technology is introduced, including VxLAN and NVGRE. The basic concept is evolved around the concept of encapsulating a L2 frame and forward transparently in the existing L3 network.

- Underlay is a physical bearer network which is consisting of multiple devices such as TOR switches, aggregation switches, core switches, load balancer, and firewall equipments etc. After overlay concept is realized, a logical network is formed on top of a underlay network.

- Overlay network consists of logical links and logical nodes. It has independent control and forwarding plane.

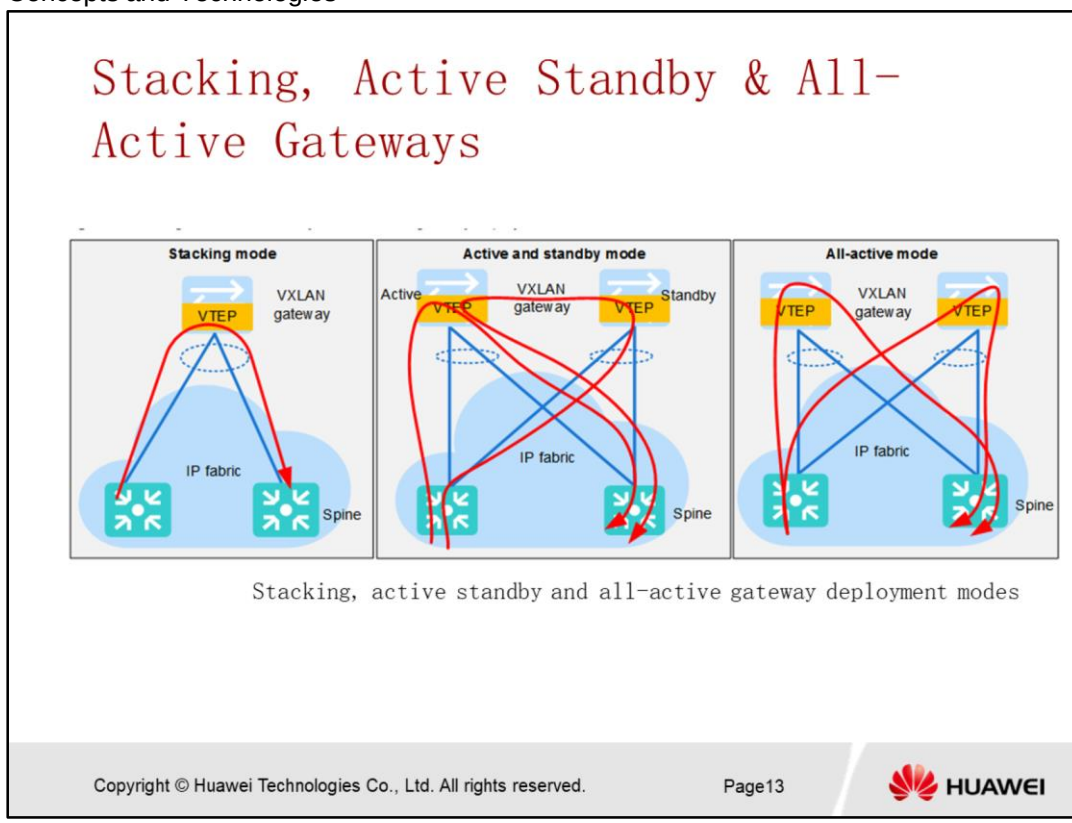


- Network Overlay – VxLAN Overlay tunnel VTEP is located on physical switches.
- Host Overlay – VxLAN Overlay VTEP is configured on vSwitches (vSwitches are installed on servers)
- Hybrid Overlay – VxLAN Overlay VTEP might be configured on physical switches and vSwitches.



- Service chain is a sequential flow model of a service function, which is used to ensure specific service flows based on these service function nodes.

- There are a few concepts related to SFC, as per listed below:-
  - SF (Service Function) – A network function that is used to perform specific checking and handling on data packet such as data filtering and NAT etc; DCN network normally defines it as VAS (value-added service); it can be achieved or realized by adding certain modules on physical server or virtual instance. It is normally done by LB or firewall devices.
  - SC (Service Classifier) – Perform service flow classification, setting service identifier and perform encapsulation
  - SFF (Service function forwarder) – Connect service function node and perform service forwarding based on the service classification done; This is performed by TOR, gateway, vSwitch etc.



- In Huawei CloudFabric Solution, gateways can be deployed in stacking, active and standby, and all-active modes.

- **Stacking mode**
- When deployed in stacking mode, two gateway devices set up a CSS or an iStack, forming a unified control plane to act as a VXLAN Layer 3 gateway.
- The gateways function as a single device. The configuration is simple and the O&M GUI is clear.
- Traffic on the gateways can be load balanced between the stack members.

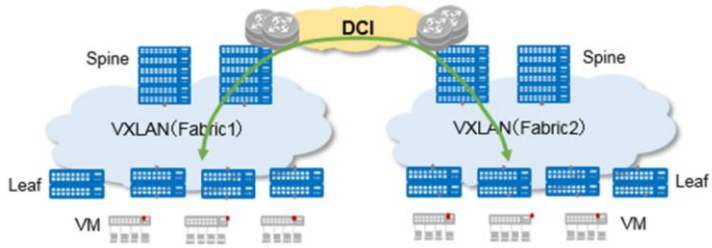
- **Active standby mode**
- When deployed in active and standby mode, two independent gateway devices set up a Virtual Router Redundancy Protocol (VRRP) group and are configured with different tunnel end point addresses. The active and standby gateways use the same VRRP virtual IP address as the VXLAN Layer 3 gateway address and act as one logical gateway for servers.
- Two gateways have independent control planes and back up each other.

- **All-active mode**
- When deployed in all-active mode, two or four independent gateways set up a DFS group. They have the same gateway address and tunnel end point address, and act as one logical gateway for servers.

- Two gateways have independent control planes and back up each other.
- Traffic on the gateways can be load balanced between the DFS group members.

Data Center Interconnection (DCI)

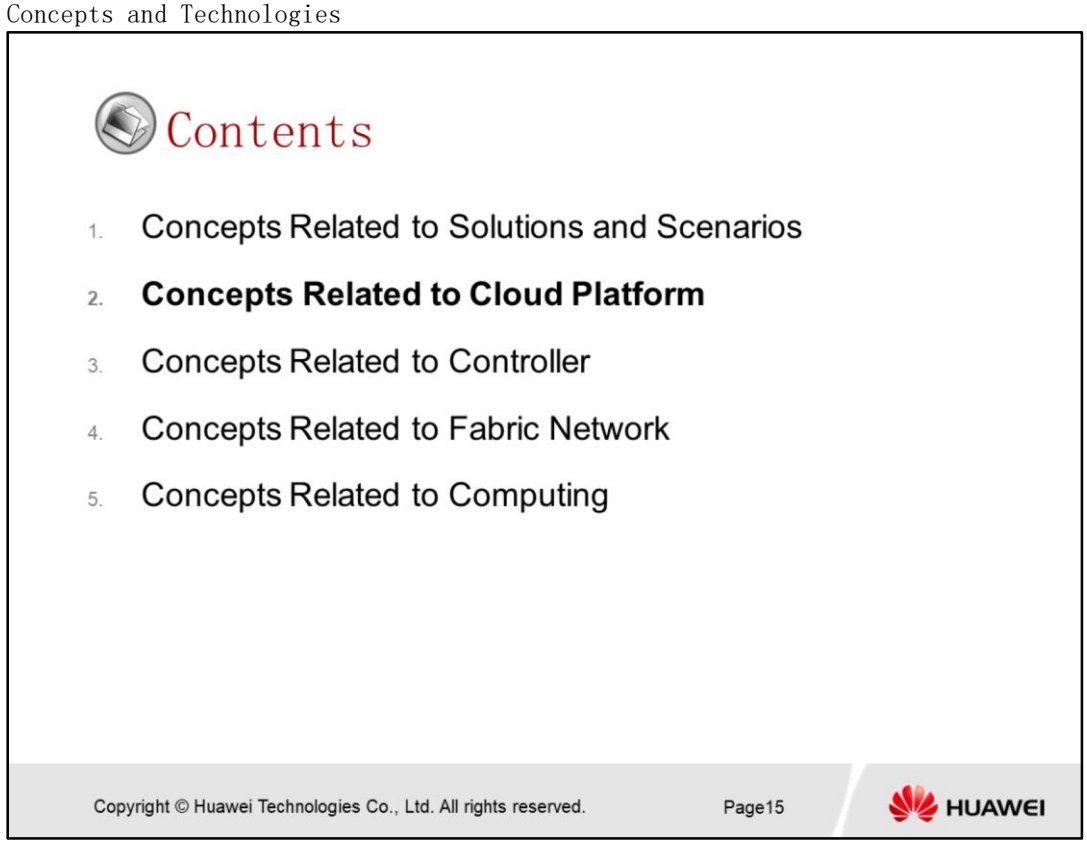
- DCI (Data Center Interconnection) is referring to the interconnection between 2 data center network to realize data communication and service migration between different data center network.

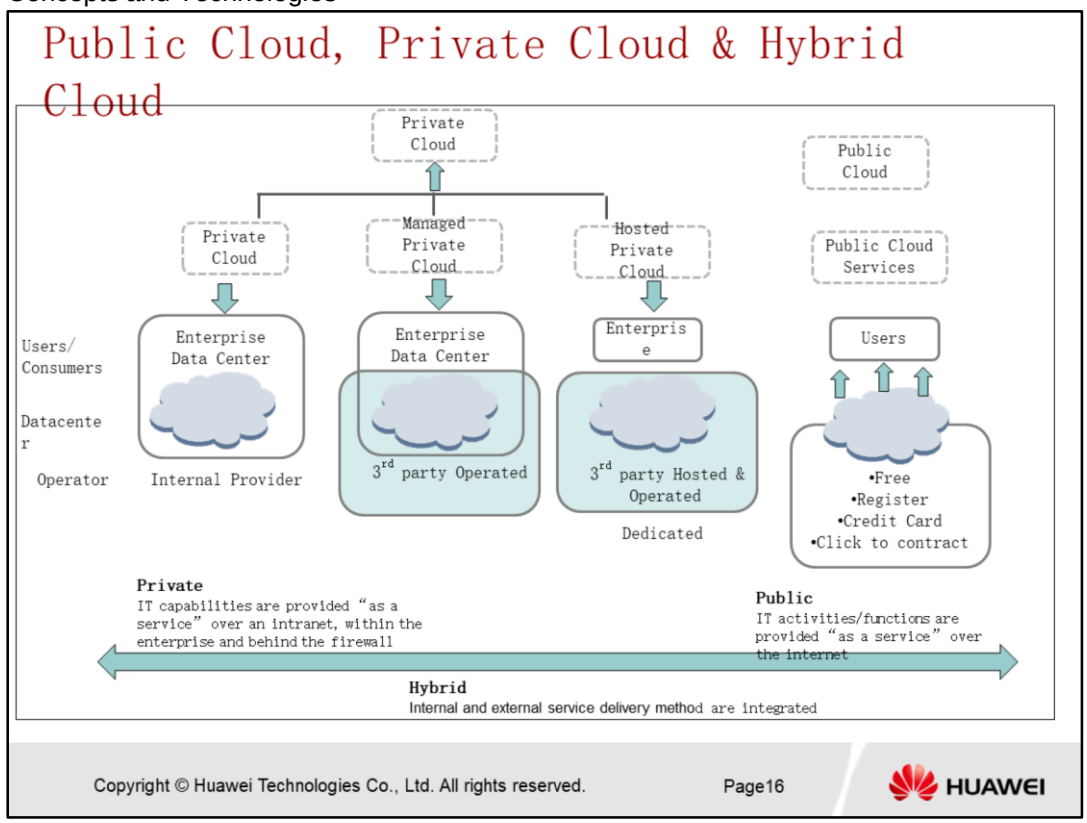


DCI Networking

- DCI (Data Center Interconnection) is referring to the interconnection between 2 data center network to realize data communication and service migration between different data center network.

- As for Overlay network, BGP EVPN (Ethernet Virtual Private Network) VxLAN technology is used to allow L2 communication between different DCN.

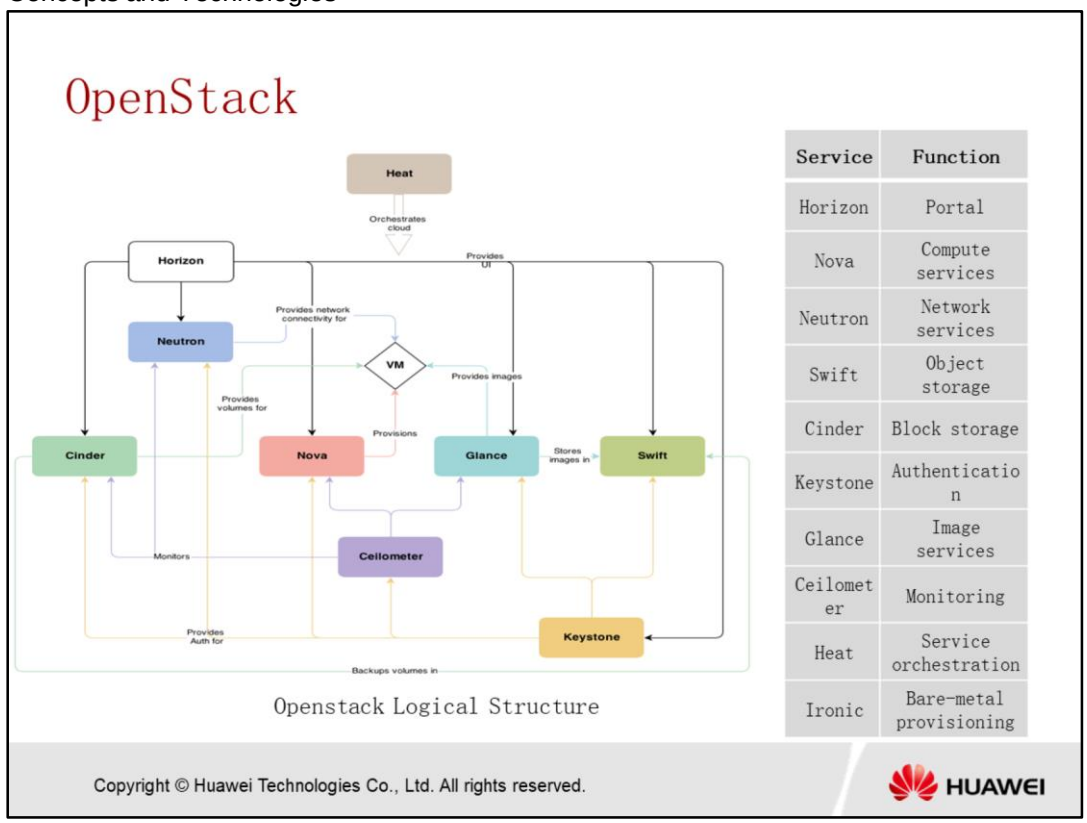




- **Private cloud**: The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

- **Public cloud**. The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

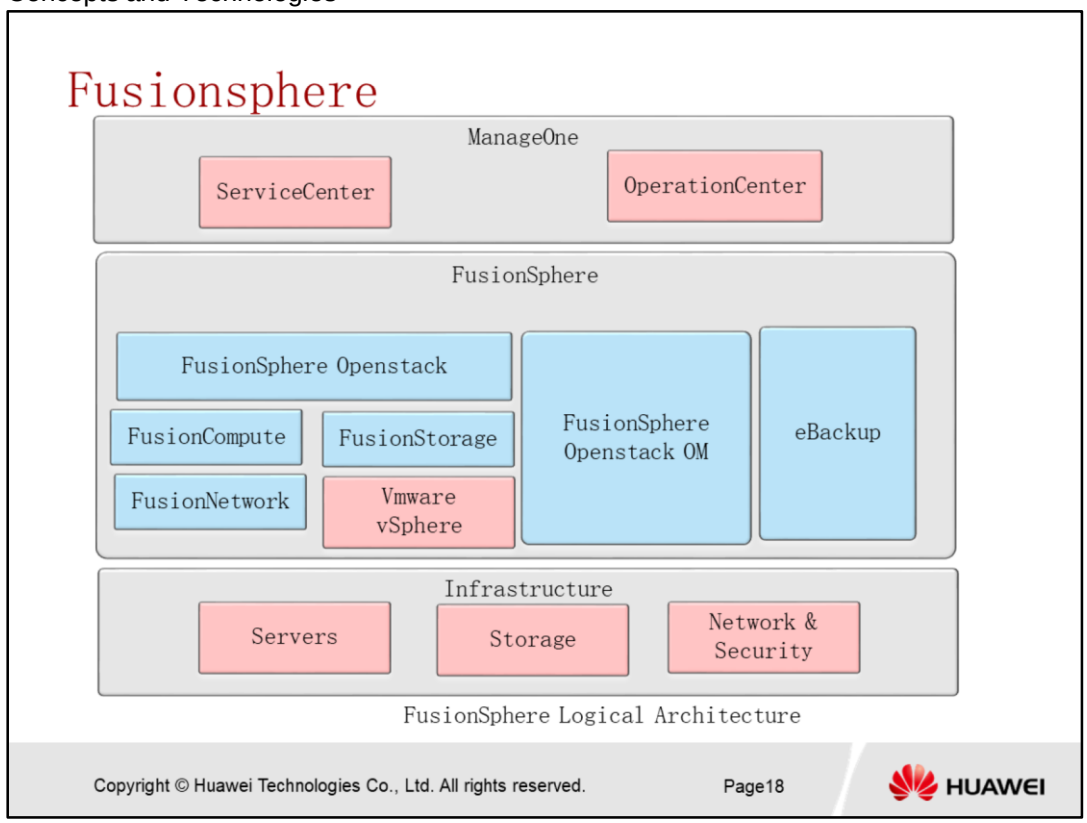
- **Hybrid cloud**. The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).



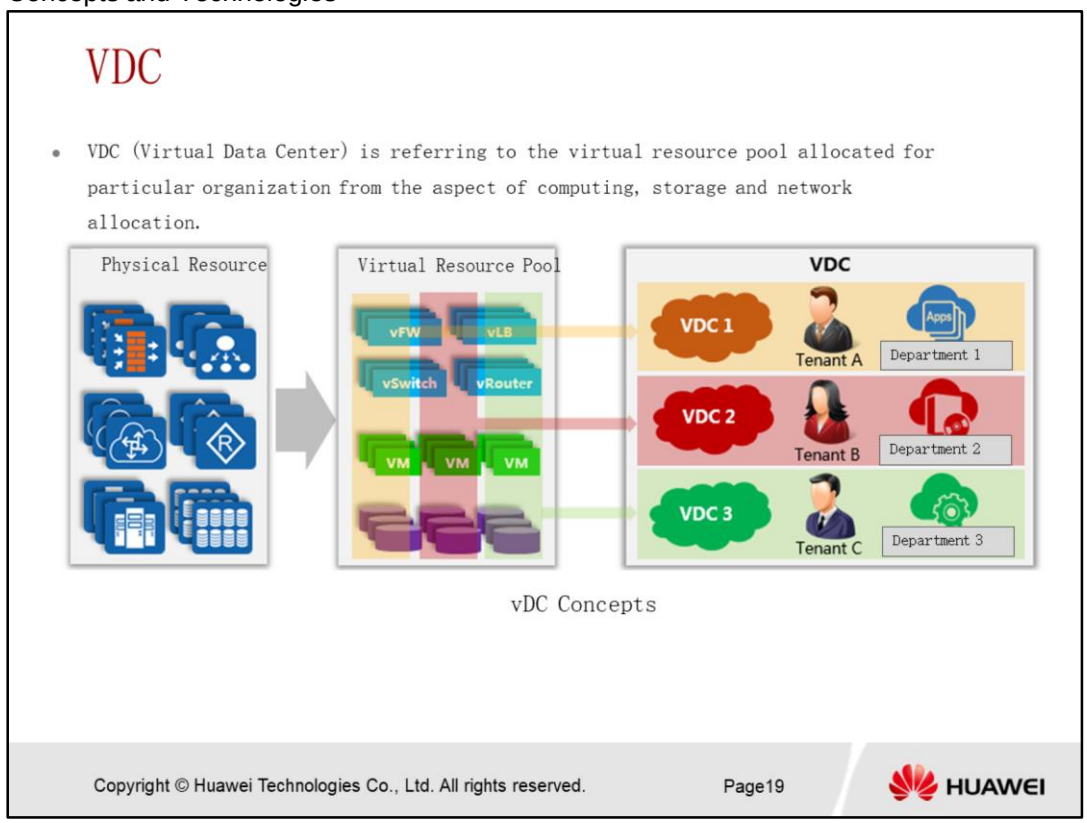
Service	Function
Horizon	Portal
Nova	Compute services
Neutron	Network services
Swift	Object storage
Cinder	Block storage
Keystone	Authentication
Glance	Image services
Ceilometer	Monitoring
Heat	Service orchestration
Ironic	Bare-metal provisioning

- **OpenStack** is an open source platform for creating and managing large groups of virtual private servers in a cloud computing environment. The platform supports interoperability between cloud services and allow businesses to build and deploy private cloud services in their own data centers. The National Aeronautics and Space Administration (NASA) worked with Rackspace, a managed hosting and cloud computing service provider, to develop OpenStack. RackSpace donated the code that powers its storage and content delivery service and production servers. NASA contributed the technology that powers their high performance computing, networking and data storage cloud service.
- OpenStack has a modular architecture that currently has eleven components:

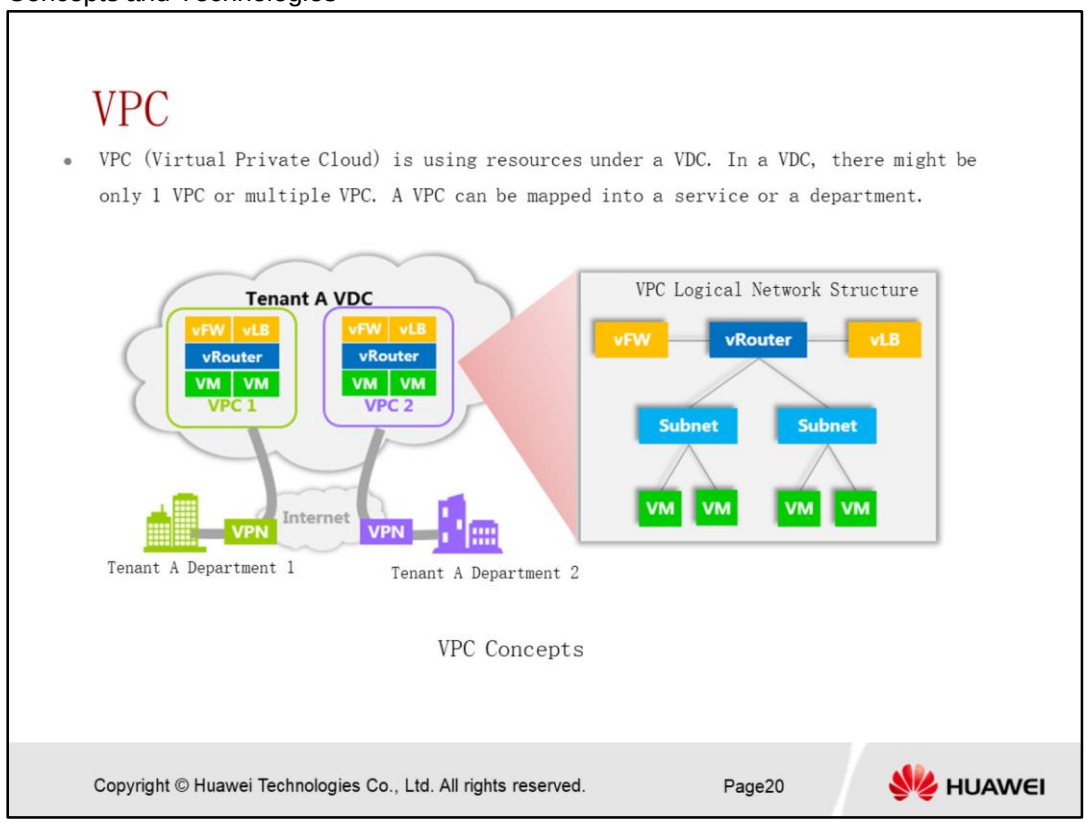
1. **Horizon** - provides a modular web-based user interface (UI) for OpenStack services.
2. **Nova** - provides virtual machines (VMs) upon demand.
3. **Neutron** - provides network connectivity-as-a-service between interface devices managed by OpenStack services.
4. **Swift** - provides a scalable storage system that supports object storage.
5. **Cinder** - provides persistent block storage to guest VMs.
6. **Keystone** - provides authentication and authorization for all the OpenStack services.
7. **Glance** - provides a catalog and repository for virtual disk images.
8. **Ceilometer** - provides a single point of contact for billing systems.
9. **Heat** - provides orchestration services for multiple composite cloud applications.
10. **Trove** - provides database-as-a-service provisioning for relational and non-relational database engines.
11. **Ironic –** provides API interface for bare metal provisioning



- Fusionsphere is Huawei cloud platform solution which is designed based on Openstack architecture; it is used in the enterprise cloud computing data center scenario widely as it provides strong virtualization function and resource management and optimization capabilities, as well as provide open API for integration.

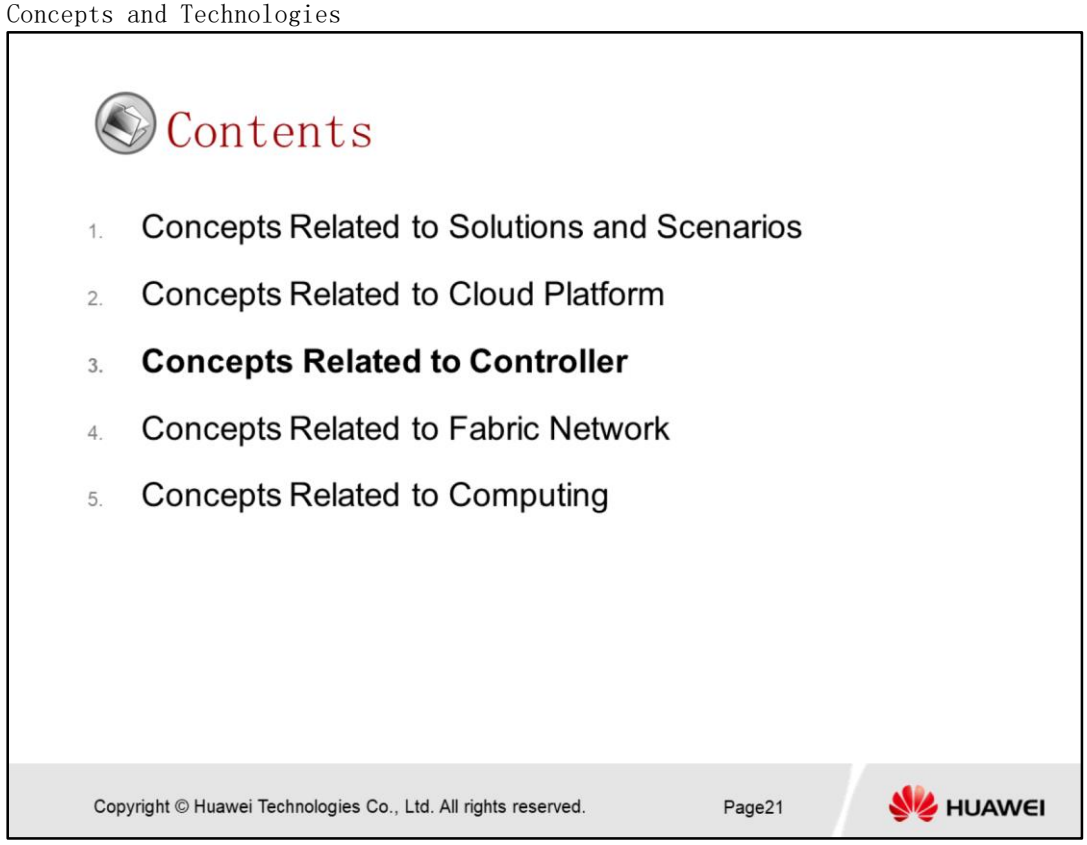


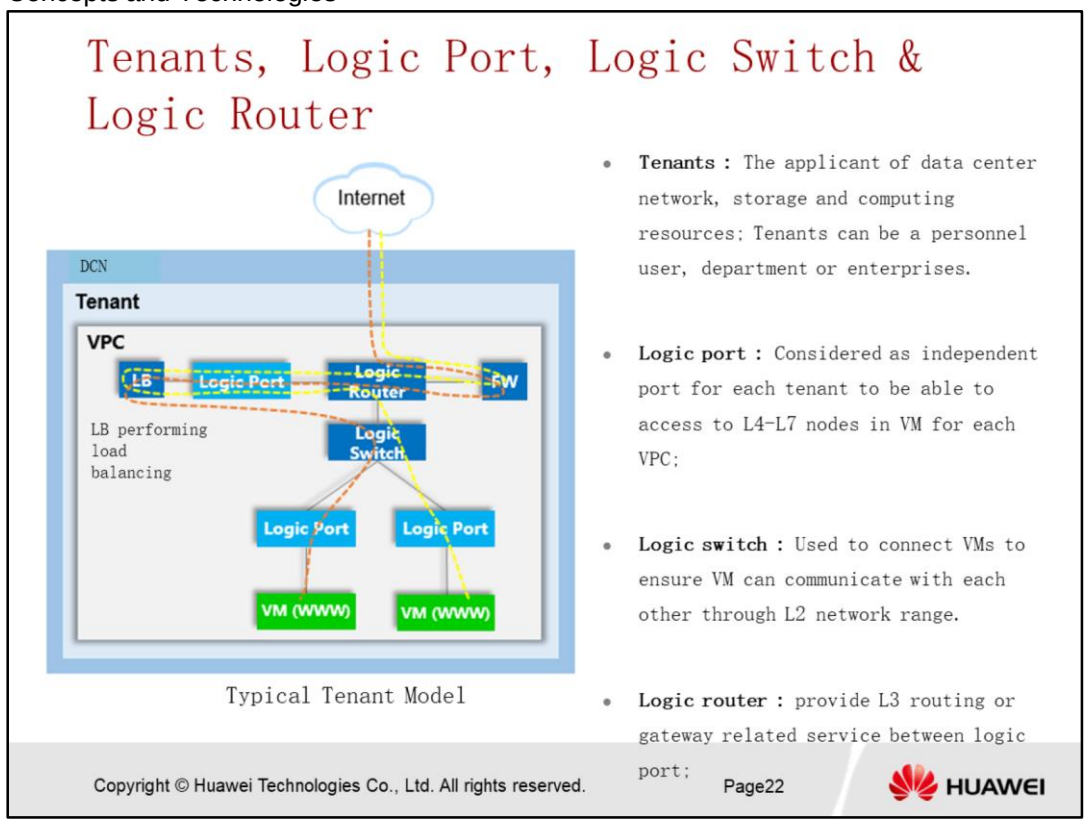
- VDC (Virtual Data Center) is referring to the virtual resource pool allocated for particular organization from the aspect of computing, storage and network allocation.
- For public cloud scenario, an administrator can define a VDC and allocate tenants for that VDC. Only tenants in the VDC can manage VMs under the VDC allocated.
- For private cloud scenario, VDC can be defined flexibly and be allocated for a service, application or department. Administrator can allocate different resources for different services, applications and departments inside an enterprise through VDC



- VPC (Virtual Private Cloud) is using resources under a VDC. In a VDC, there might be only 1 VPC or multiple VPC. A VPC can be mapped into a service or a department.

- The advantages of VPC are listed below:-
  - Network isolation: VPC provides isolated VMs and network environment to achieve network isolation requirements of different departments.
  - Rich in services: Each VPC can provide independent virtual firewall, elastic IP addressed, security group, SuperVLAN, Ipsec VPN, NAT gateway related service etc.
  - Flexible networking: Supports directly connected network and various types of routing network.





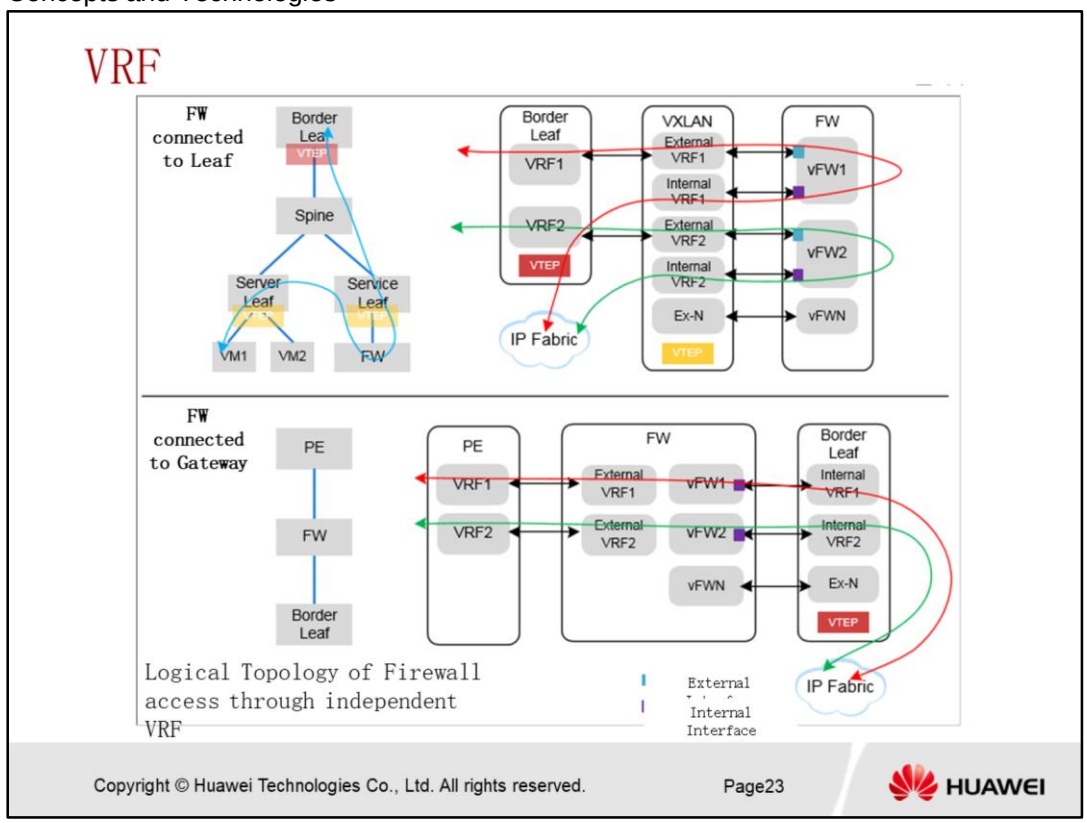
- **Tenants** – It is referring to the applicant of data center network, storage and computing resources; Tenants can be a personnel user, department or enterprises. The administrator of a DCN can allocate specific amounts of resources to different users so as to provide differentiated services.

- **Logic port** – is considered as independent port for each tenant to be able to access to L4-L7 nodes in VM for each VPC; From physical point of views, there might be multiple tenants sharing 1 physical port in the network.

- **Logic switch** – is used to connect VMs to ensure VM can communicate with each other through L2 network range. One network node is able to virtualise multiple logic switch to allow multiple tenants to use even they are sharing one physical network node physically. A tenant will see the logic switch as a independent and real switch equipped with software resources; 2 logic switches are totally isolated from each others. Communication between Logic switch is completed through L2 switching network.

- **Logic router** – provide L3 routing or gateway related service between logic port; 1 VPC can have only 1 logic router. Logic router can be realized by using virtualized software; it is used to connect VMs through L3 network. It allows one physical network

node can be shared by multiple tenants which treats them like independent router.



- VRF (Virtual Routing and forwarding), allows multiple vpn instance routing table to be existing in one router; each VPN instance and virtual routing table is independent from each other; Repeated IP addresses range can be used in respective vpn instance because each VPN instance is totally isolated from each other.
- In AC-DCN solution, network isolation among multiple tenants can be realized by VxLAN gateway and VRF technology.
- For the scenario firewall is connected to leaf, traffic from IP fabric flows into firewall from the internal VRF of service leaf, and flow from the external interface of firewall into external VRF of service leaf, and lastly be forwarded to external network through VRF in border leaf. Diagram above represents 2 different user VRF traffic, marked in green and red color.
- In the scenario firewall is directly connected to gateway, traffic from IP fabric flows into internal interface of firewall through the internal VRF of border leaf. Internal firewall integrates external VRF functions and flow traffic to external network.

Node Discovery

Node Type	Node Model	Discovery Method
Load Balancer	F5 Load Balancer	3rd party device import
Physical Firewall	USG9500/E8000E	Automatic Discovery Batch Import
Software Firewall	vNGFW	Automatic Discovery Batch Import
	Checkpoint Firewall/ Fortinet FIrewall	3rd party device import
Physical Switch	NE/CE12800/CE6850HI/CE6860/6870/ CE6880/CE7850/CE8860	Automatic Discovery Batch Import
Software Switch (vSwitch)	CE1800V	Device registration
	EVS/OVS	
Server	ExSi Host	Automatic link discovery & node discovery
	Hyper-V Host	
	KVM Host	
	Bare metal server	Automatic Discovery
	Physical server	Manual addition

- To allow AC to be able to deploy configure and services to devices such as switches, firewall and servers, users are required to configure the corresponding protocol parameter and configurations on the device, and thus device can be discovered on the AC GUI.
- There are 4 main types of node discovery methods, including automatic discovery, batch import, device registration and 3rd party device import.

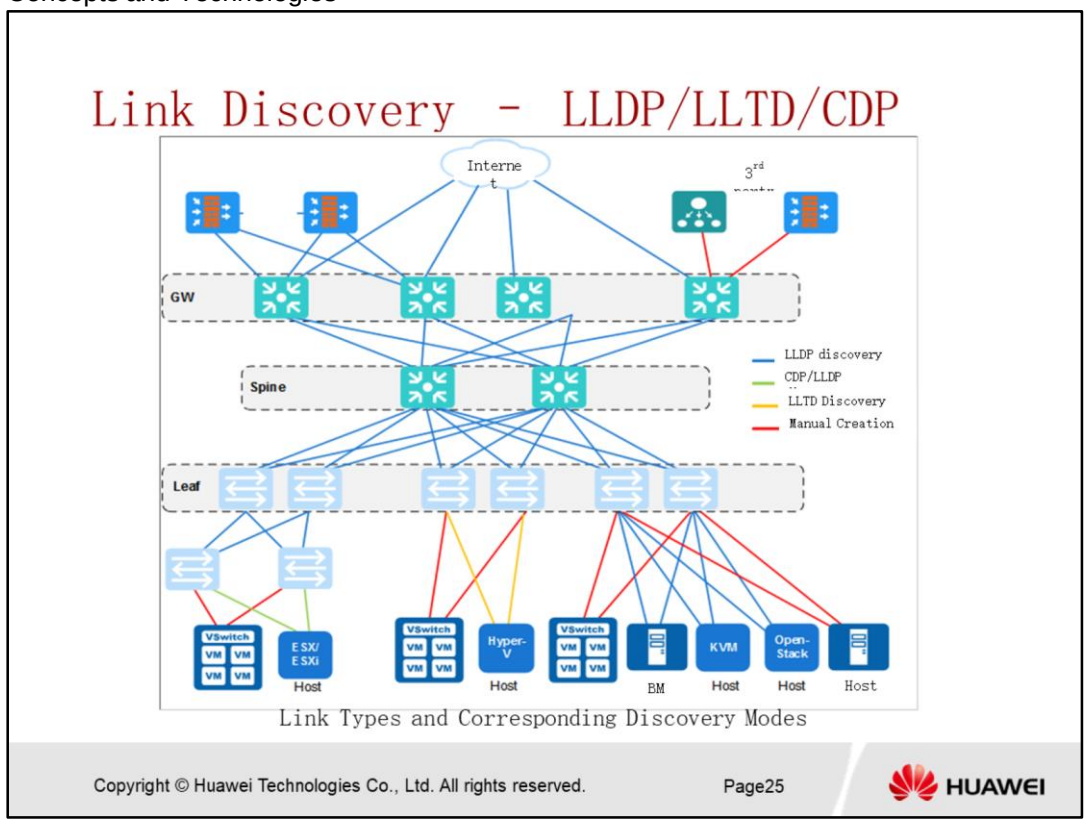
1. **Automatic discovery:** Huawei hardware switch and Huawei firewall supports SNMP protocols, and the IP address is normally concentrated on certain IP range; thus this method is recommended to be used to discover the devices on AC

2. **Batch import:** When Huawei hardware switch or firewall supports SNMP protocol but the IP address configured are relatively scattered, it is recommended to import devices into AC using batch import template.

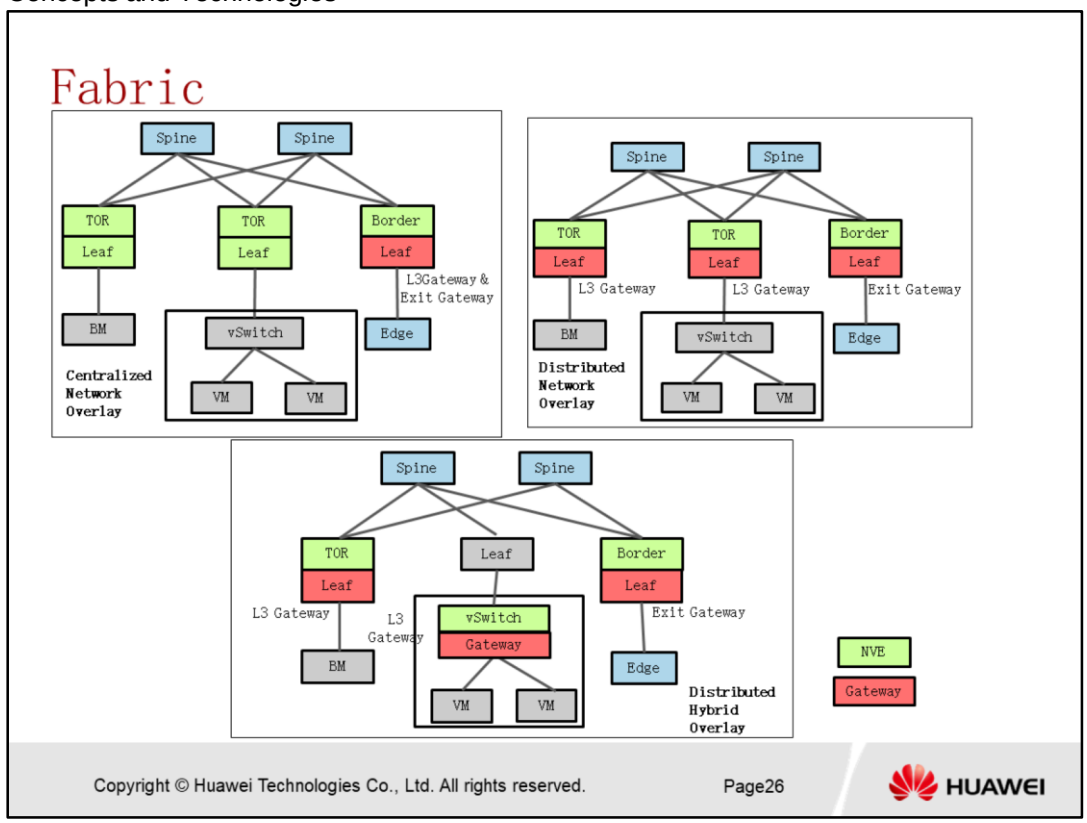
3. **Device registration:** Applied for AC to manage Huawei vSwitch CE1800V; First, equipment information is registered on AC UI, and initiate CE1800V to request for connection establishment. The registered CE1800V can be discovered through AC.

4. **3rd party device import:** Even though AC does not function to mange 3rd party device, there might be necessity to show 3rd party device into AC UI; This method

can be used to import 3rd party device into AC UI.



- AC can discover links through automatic discovery and manual creation.
- For the links between devices, protocol that can be used to discover devices include:-
- **LLDP – Link Layer Discovery Protocol**
  - By using LLDP, local device management IP address, device identifier and interface information can be advertised to peer; upon receiving these information, the peer is going to keep these in MIB for fast L2 network discovery later.
- **LLTD – Link Layer Topology Discovery Protocol**
  - Is a type of link layer topology discovery protocol which is able to discover LLTP compatible devices and links.
- **CDP – Cisco Discovery Protocol**
  - A link layer discovery protocol implemented by Cisco which is widely used on Cisco devices. Through CDP, cisco devices can share information such as system version, IP address etc to the directly connected peer.



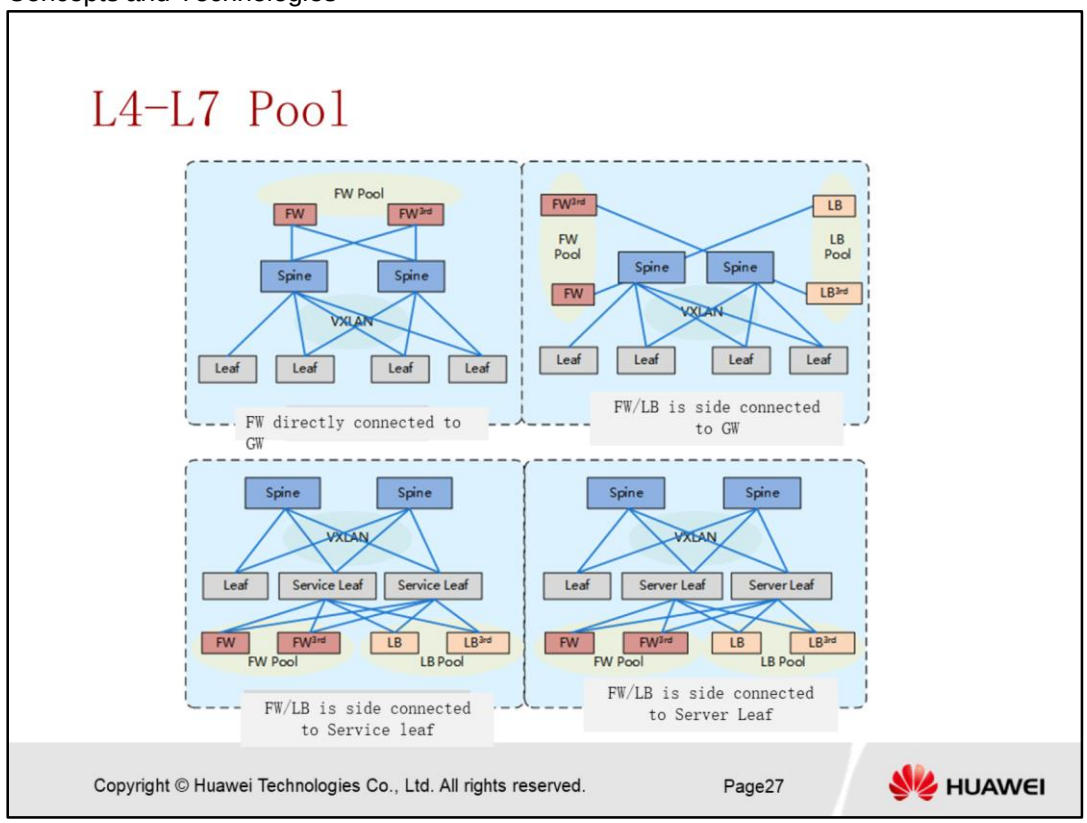
- Fabric network comprises of a group of the interconnection between spine and leaf nodes, to provide connection of different nodes in order to achieve the objective of simultaneous usage of a physical server for multiple tenants; This is to save cost and increase resource utilization rate.

- Fabric network can be divided into 3 types as shown in the diagram above:-

1. **Centralized network overlay:** VxLAN VTEP consists of only physical switches. (Border leaf is physical server) VxLAN gateway is centralized on 1 single edge device, which serves as L3 gateway and exit gateway:

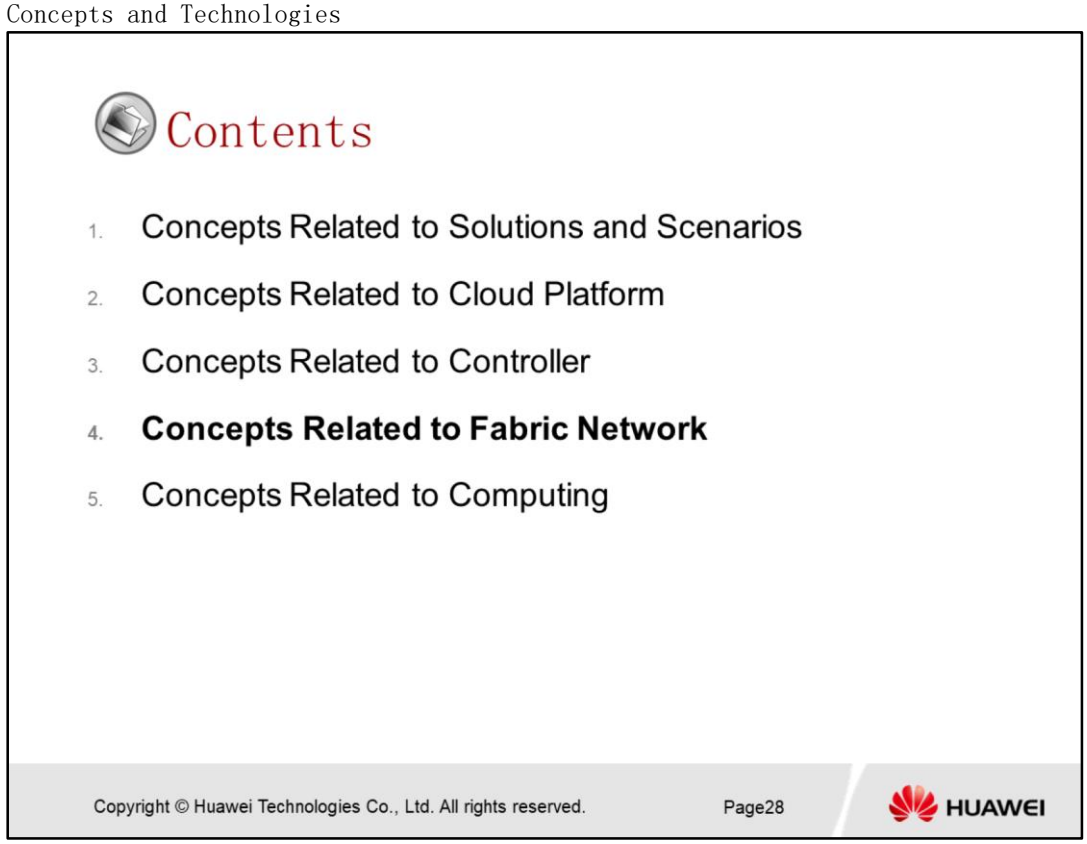
2. **Distributed network overlay:**VxLAN VTEP consists of only physical switches. (Border leaf is physical server) VxLAN gateway is distributed on different devices.

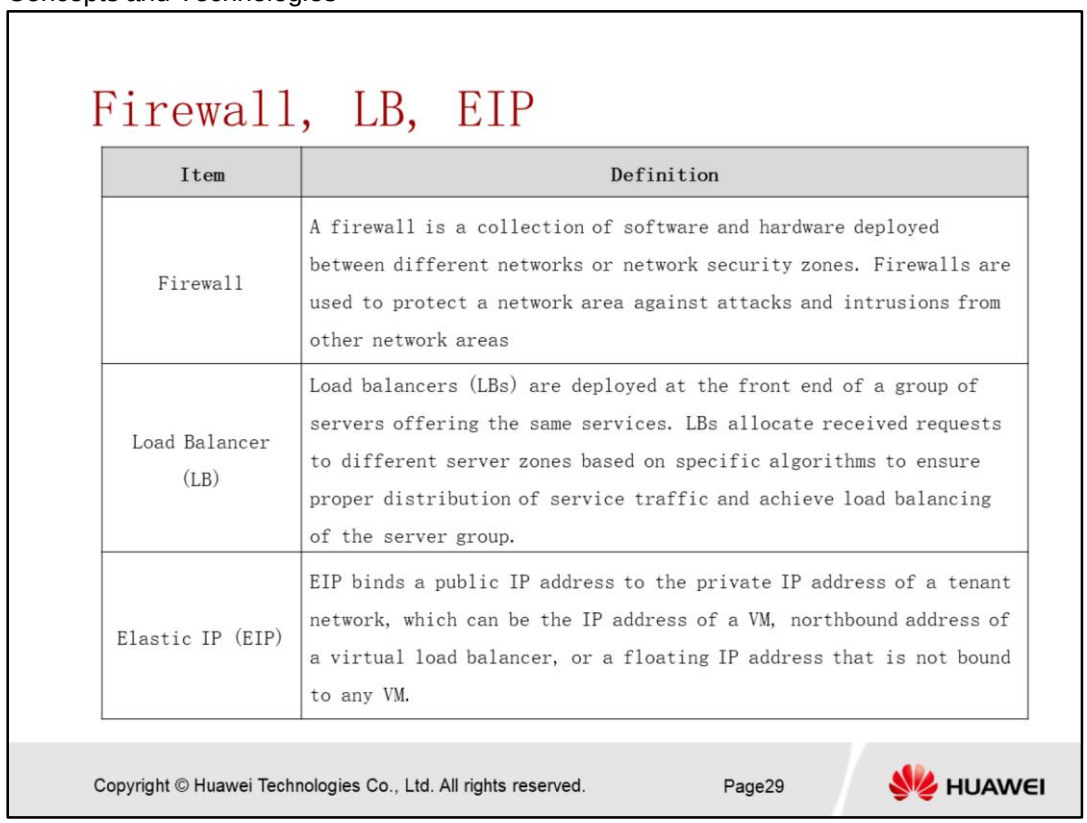
3. **Distributed hybrid overlay:** VxLAN VTEP consist of both physical and virtual switches. VxLAN gateway is distributed on different devices.



- L4-L7 device is referring to service provided by load balancer (LB) and firewall (FW), such as ACL, load balancing, NAT, Qos, IPSec VPN etc. To increase the utilization rate of L4-L7 device and for ease of management for L4-L7 pool, firewall and load balancer that is discovered by AC-DCN should be classified into specific L4-L7 pool, to allow AC-DCN allocation for different tenants .

- There are 4 scenarios for firewall and LB connection:-
  - FW directly connected to gateway – supports Huawei firewall and 3rd party firewall connected to gateway
  - FW/LB side connected to gateway - supports Huawei firewall and 3rd party firewall, Huawei LB and 3rd party LB;
  - FW/LB side connected to service leaf – supports Huawei firewall and 3rd party firewall, Huawei LB and 3rd party LB;





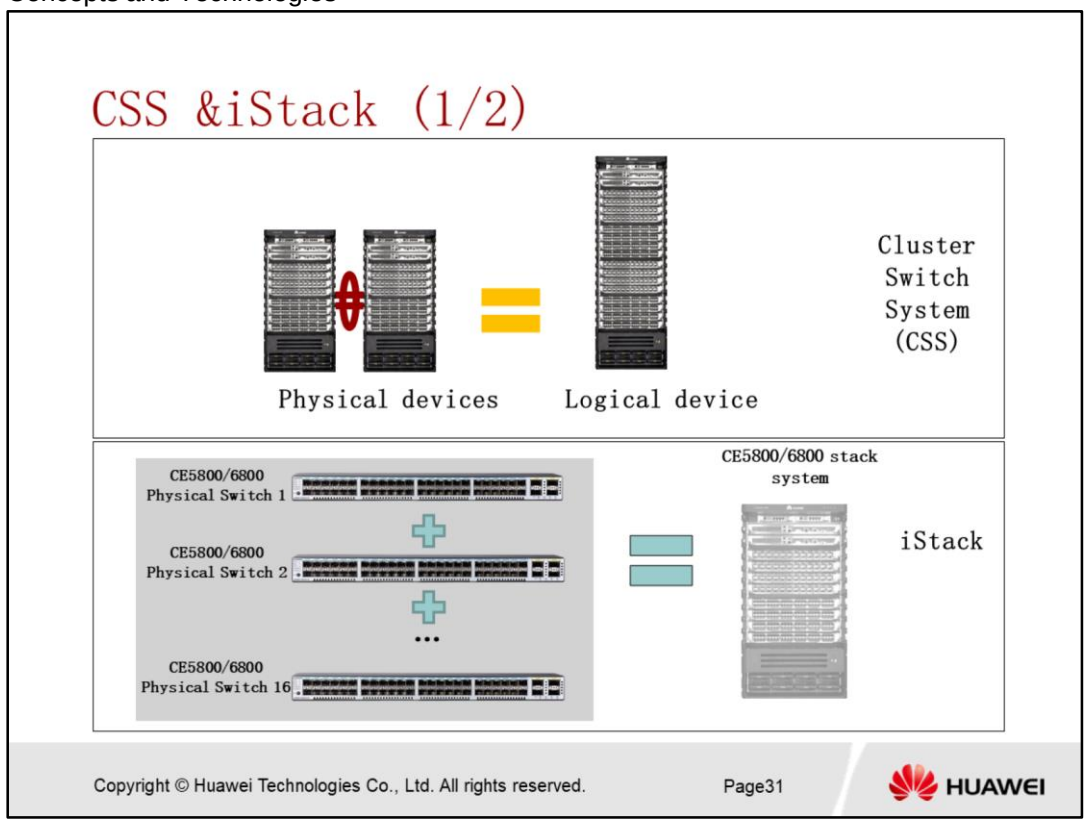
- **Firewall**
  - A firewall is a collection of software and hardware deployed between different networks or network security zones. Firewalls are used to protect a network area against attacks and intrusions from other network areas.
  - Firewalls can isolate and mitigate attacks, and can be deployed at the network border or used for subnet isolation. For example, they can be configured as an enterprise network egress, subnet isolation in a large-sized network, or DCN border to control access behavior in inbound and outbound traffic. Defense is the core attribute of firewalls.

- **Load Balancer (LB)**
  - Load balancers (LBs) are deployed at the front end of a group of servers offering the same services. LBs allocate received requests to different server zones based on specific algorithms to ensure proper distribution of service traffic and achieve load balancing of the server group.
  - The LB provides a virtual IP address for external traffic to access the server group. When a new task reaches the virtual IP address, the LB uses load evaluation algorithms to find a server with a light load and assigns the task to it. Therefore, the performance (including service throughput, average response speed, and system resource usage) of the server group can be kept at an

optimal level.

- In addition, the LB checks the running status of each server in the server group according to health status standards. When a server is faulty, the LB allocates new tasks to the other servers.
  - LBs provide the following functions: virtual IP configuration, self-IP configuration, listener, Secure Sockets Layer (SSL) certificate, traffic statistics and charging, health status standards, and session hold policies.
- In Huawei CloudFabric Solution, LBs are L4-L7 components and provide server load balancing externally.

- **Elastic IP (EIP)**
  - Elastic IP (EIP) allows external networks to access tenant networks.
  - EIP binds a public IP address to the private IP address of a tenant network, which can be the IP address of a VM, northbound address of a virtual load balancer, or a floating IP address that is not bound to any VM. An external user can use the public static IP address to access resources in a tenant network.



- Both CSS and iStack are stacking technology. Frame switch stacking is called CSS while box switch stacking is called iStack. In AC-DCN solution, Gateway devices are normally Huawei CE devices and normally 2 CE devices are stacked into 1 logical switch.

- CSS features
  - Many-to-one virtualization: Multiple switches are virtualized into one logical switch that has only one control plane and provides unified management.
  - Unified forwarding plane: CSS uses a unified forwarding plane that shares and synchronizes forwarding information.
  - Multi-chassis link aggregation group : Links between an upstream and a downstream CSS are aggregated to one Eth-trunk link.

- CSS virtualizes multiple switches into one logical device and supports inter-chassis link aggregation. It simplifies network topology and greatly improves network performance:

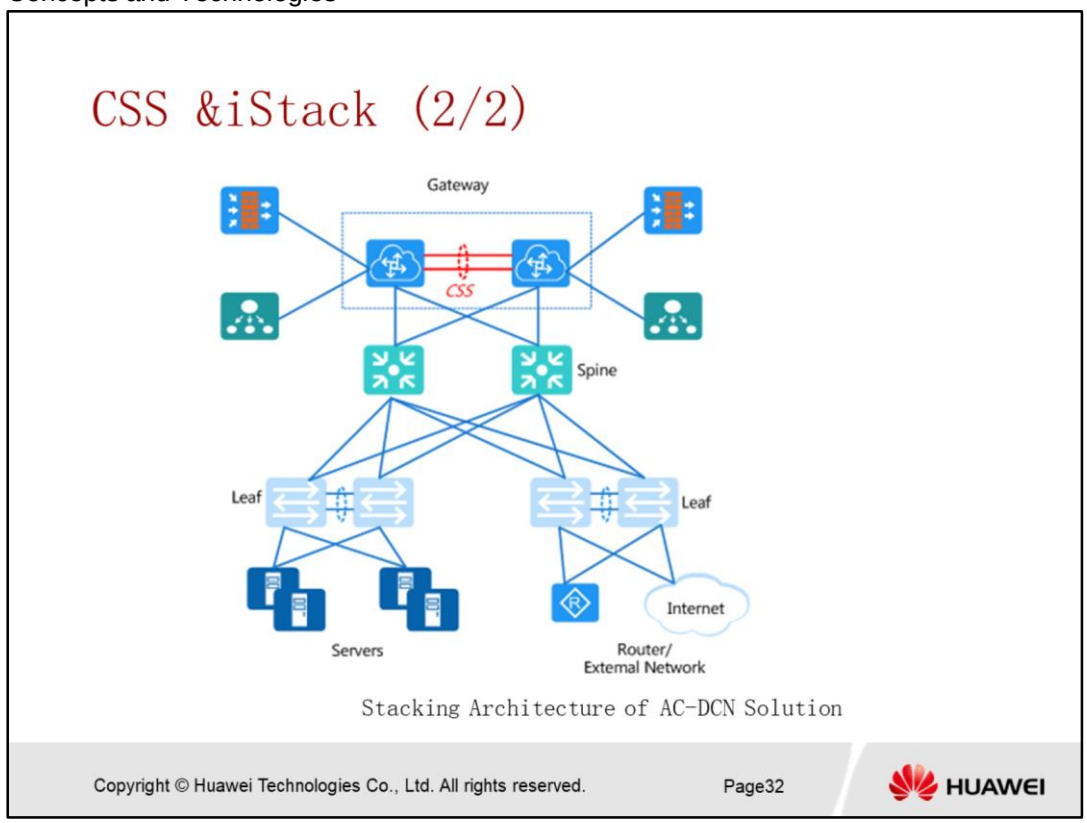
  ❑ Simplified operation and maintenance: CSS functions as one logical switch, simplifying operation and maintenance and reducing OPEX.

  ❑ High reliability: When one switch fails, another switch in the CSS takes over the control and forwarding of packets, so that services are not influenced by single-point failure.

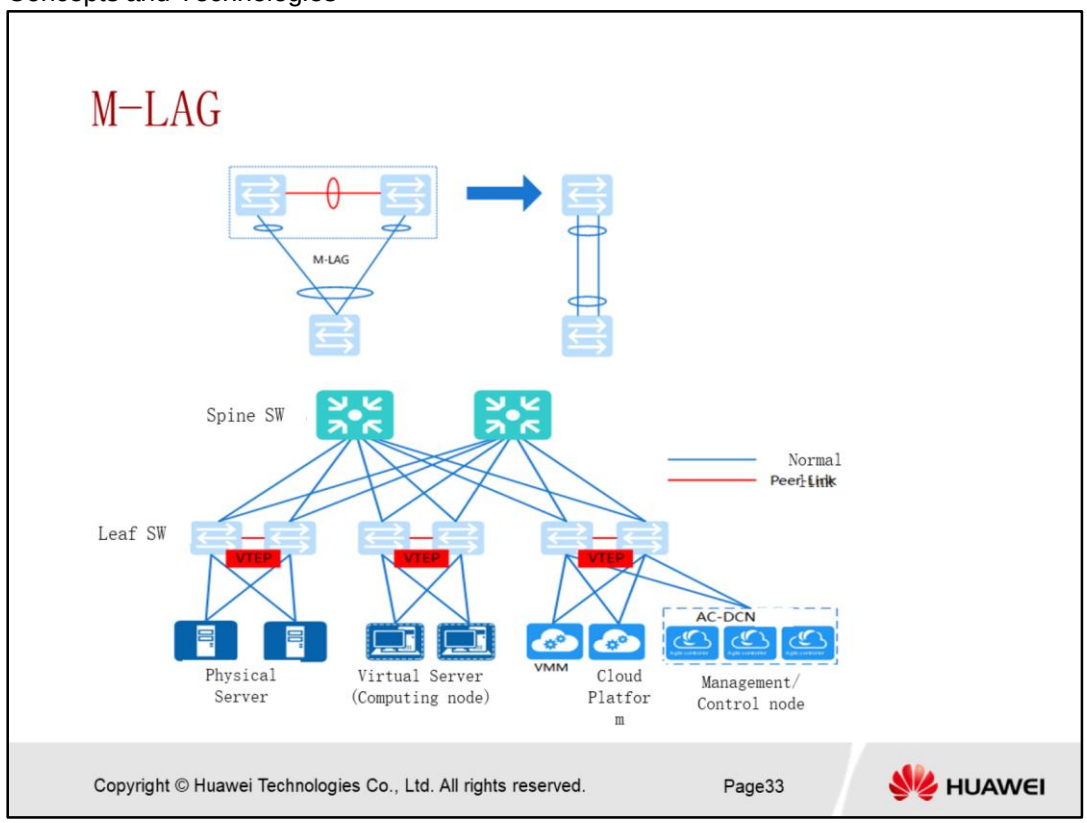
  ❑ Loop-free network: CSS supports multi-chassis link aggregation group to prevent loops.

  ❑ Load balancing of links: CSS supports equal cost multiple path (ECMP) across switches, making full use of network links and bandwidths.

- CSS simplifies network architecture, improves forwarding performance, and does not lead to the loss of any network functions. CSS has all functions of the physical switches in the cluster and provides better performance. CSS gains wide customer recognition and becomes a preferred solution for simple and efficient network deployment.



- In AC-DCN solution, Gateway devices are normally Huawei CE devices and normally 2 CE devices are stacked into 1 logical switch



- M-LAG, Multi-chassis Link Aggregation Group is a type of link aggregation mechanism which can be realized between 1 single device to 2 links on 2 different devices, so as to improve the redundancy and protection level from link and board level to equipment level protection. It is like an enhanced version of LAG.

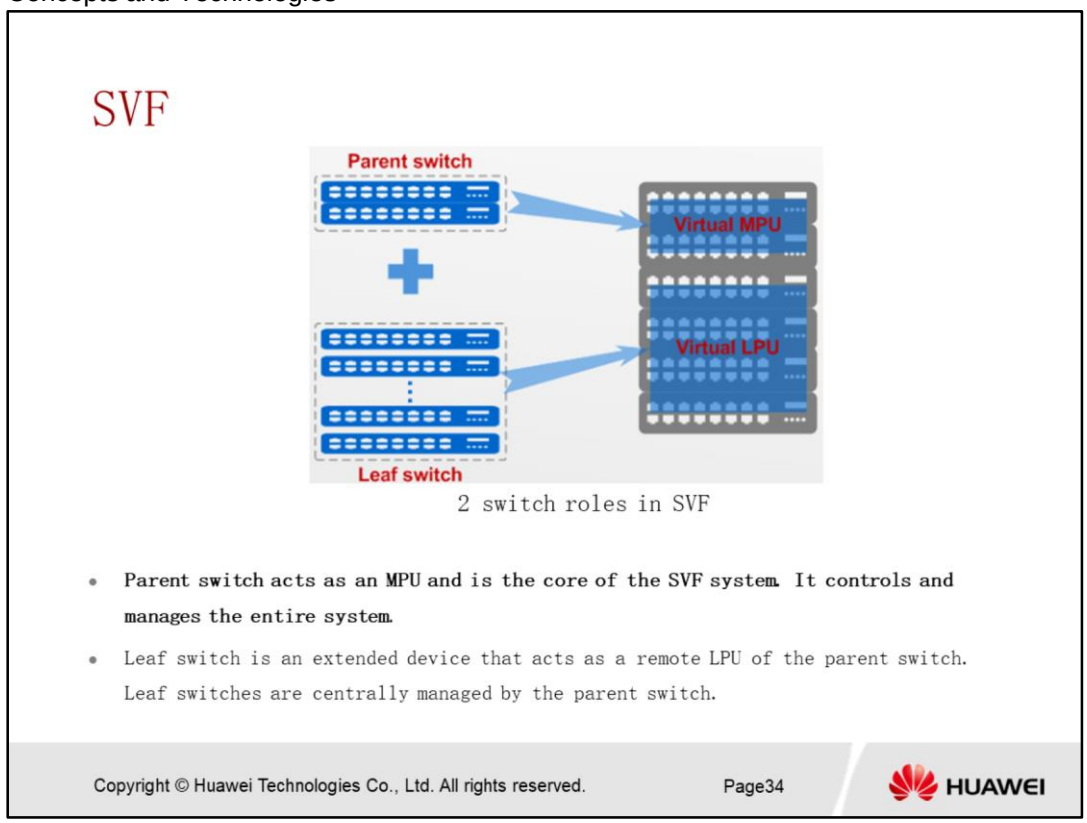
- In AC-DCN solution, M-LAG is normally deployed on the leaf switches to improve the HA level in the solution.

- The advantages of M-LAG include:-

1. **Higher reliability** – when one equipment or one link is faulty, the traffic can be automatically switched to the other equipment or link to prevent service interruption.

2. **Simplify network and configuration** – M-LAG can be considered as a type of simple virtualization technology to make the dual-homed devices to act like only 1 logical device in the network topology.

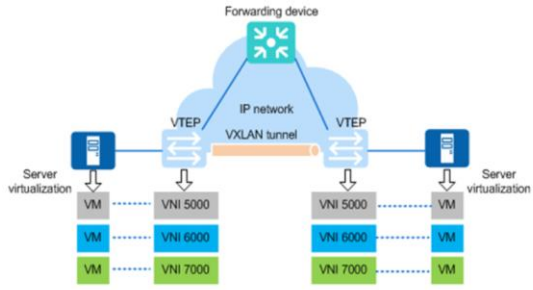
3. **Independent upgrade** – both devices can be upgraded separately; during upgrade of one device, the other device will ensure the traffic forwarding is working properly.



- Super Virtual Fabric (SVF) is a vertical virtualization technology that virtualizes switches at different layers into one logical switch. SVF satisfies high-density access requirements of DCs and simplifies network topologies and management.
- Two switch roles exist in SVF: parent switch and leaf switch.
  - Parent switch acts as an MPU and is the core of the SVF system. It controls and manages the entire system.
  - Leaf switch is an extended device that acts as a remote LPU of the parent switch. Leaf switches are centrally managed by the parent switch.
- SVF can be divided into two types based on the model of the parent switch:
  - SVF system consisting of fixed switches: The parent switch and leaf switches are all fixed switches.
  - SVF system consisting of modular and fixed switches: A modular switch is deployed as the parent switch and fixed switches are deployed as leaf switches.
- In Huawei CloudFabric Solution, SVF is deployed at the access layer. CE series switches supporting VXLAN are deployed as parent switches and cost-effective CE series switches that do not support VXLAN are deployed as leaf switches. SVF increases access port quantity without the need to purchase more VXLAN-capable devices.
- SVF has the following advantages:
- **Lower network construction costs**
  - Cost-effective switches are used as access switches, so network construction costs are reduced.
- **Simplified configuration and management**
  - SVF virtualizes multiple devices into one, reducing the number of nodes to be managed. Complicated ring protection protocols are not required; therefore, the network configuration and management are much simpler.
- **Higher scalability and more flexible deployment**
  - When more access ports are required on the network, you only need to add cost-effective fixed

switches to the network. Moreover, these cost-effective switches are deployed near servers, making network deployment more flexible.

VXLAN / NVE / VTEP / VNI

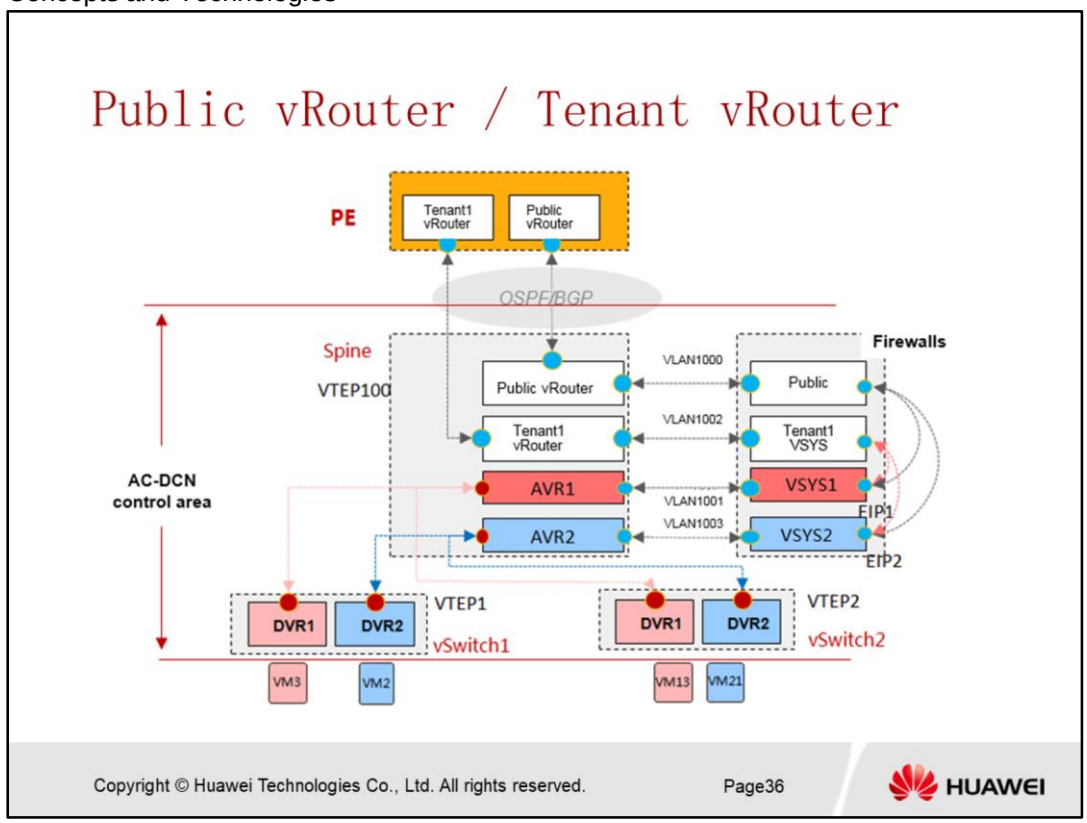


VxLAN Network Architecture

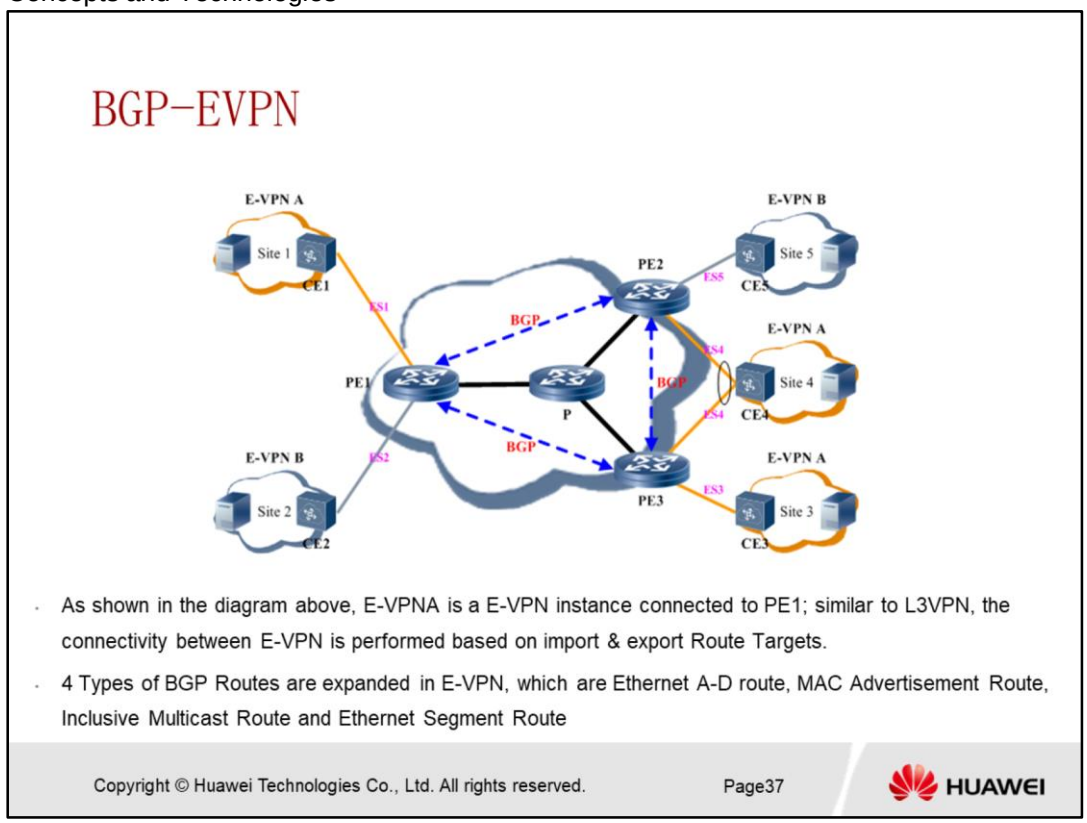
Item	Definition
Network virtualization edge (NVE)	Network entities that implement network virtualization
VXLAN tunnel end points (VTEPs)	VTEP is an end point of a VXLAN tunnel on an NVE device and is used to encapsulate and decapsulate VXLAN packets.
VXLAN network identifier (VNI)	VNIs are similar to VLAN IDs and are used to identify VXLAN segments. A VNI represents a tenant. VMs with different VNIs cannot communicate with each other at Layer 2.
VXLAN tunnel	A VXLAN tunnel is set up between two VTEPs and is a virtual tunnel that transmits VXLAN packets.

- Virtual Extensible Local Area Network (VXLAN) is a Network Virtualization over Layer 3 (NVO3) standard defined by the Internet Engineering Task Force (IETF). VXLAN uses MAC-in-UDP encapsulation to enable L2 forwarding over an L3 network. This technology allows VMs to migrate over a large L2 network and isolates tenants in a data center.
- **Network virtualization edge (NVE)**
  - NVEs are network entities that implement network virtualization. In Huawei CloudFabric Solution, CE series switches and vSwitches can act as NVEs.
- **VXLAN tunnel end points (VTEPs)**
  - A VTEP is an end point of a VXLAN tunnel on an NVE device and is used to encapsulate and decapsulate VXLAN packets.
- **VXLAN network identifier (VNI)**
  - VNIs are similar to VLAN IDs and are used to identify VXLAN segments. A VNI represents a tenant. VMs with different VNIs cannot communicate with each other at Layer 2. The VNI field in a VXLAN packet header has enough bits to support a massive number of tenants.
- **VXLAN tunnel**
  - "Tunnel" is a logical concept. A VXLAN tunnel is set up between two VTEPs and is a virtual tunnel that transmits VXLAN packets.
- VXLAN has the following advantages:

1. Supporting a large number of tenantsA VXLAN supports a maximum of 16M VXLAN segments with 24-bit VXLAN VNIs, so that a data center can accommodate numerous tenants.
2. Improving device performanceVXLAN reduces the number of MAC addresses that network devices need to learn and enhances network performance because only devices at the edge of the VXLAN network need to identify VM MAC addresses.
3. Reducing network management difficultiesVXLAN extends Layer 2 networks using MAC-in-UDP encapsulation and decouples physical and virtual networks. Tenants are able to plan their own virtual networks, without being limited by the physical network IP addresses or broadcast domains. This greatly simplifies network management.

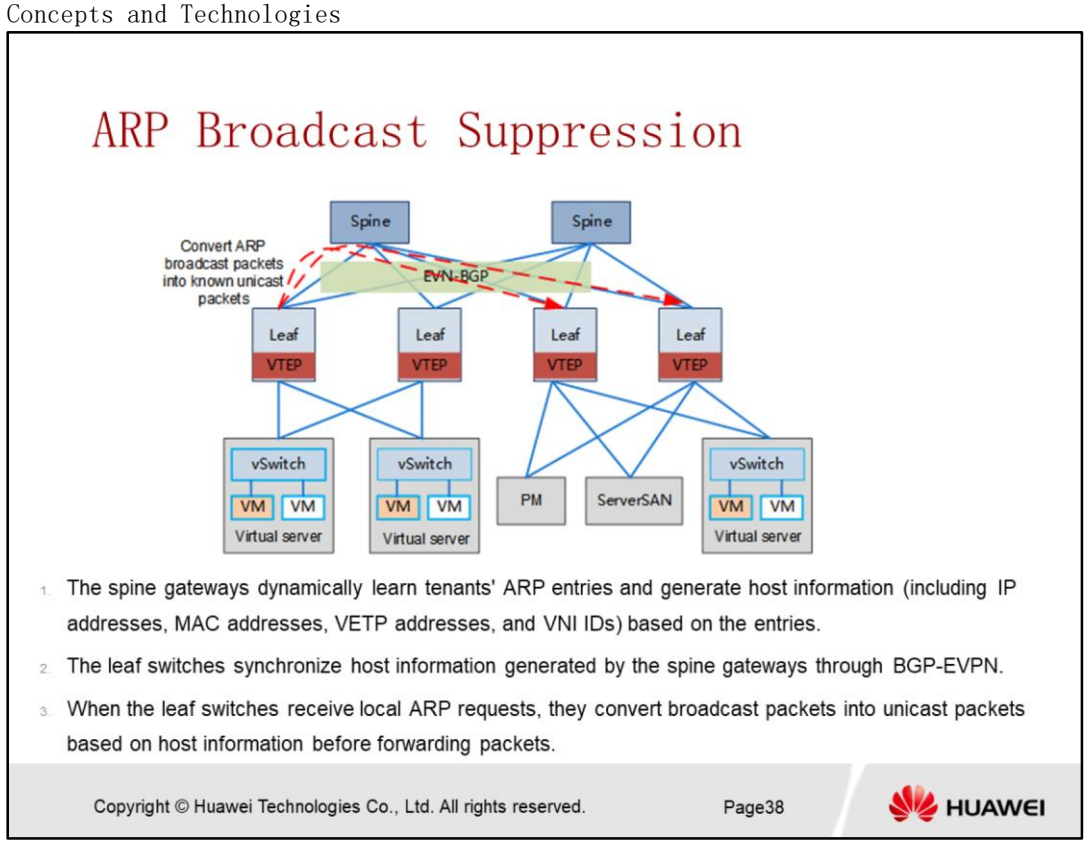


- Public vRouter, tenant vRouter, and virtual systems of firewalls are important logical elements in Huawei CloudFabric Solution. Understanding these concepts helps you better understand service traffic forwarding of the solution.
- Public virtual router (public vRouter):
  - Public vRouter is a dedicated VPN Routing and Forwarding (VRF) created on a gateway node to connect the gateway to public vFWs.
  - Public vRouter dynamically advertises EIP addresses allocated to VPCs to PEs, adds static routes for the EIP addresses, and sets the next hops of the static routes to the IP addresses of public vFWs to be connected.
- Tenant virtual router (tenant vRouter) is a dedicated VRF on a gateway node created for a tenant. A tenant vRouter connects tenant vFWs and PEs and applies to the scenario where firewalls are deployed in bypass mode. The default route of a tenant vRouter is destined for the VRF connected to PEs.
- The virtual systems of firewalls are classified into two types:
- **Public virtual firewall (public vFW):** is a special virtual system that exists on firewalls by default. After the virtualization function is enabled, the public vFW inherits the previous configurations on the firewalls. In Huawei CloudFabric Solution, the public vFW serves as the summary routing domain of the vFWs of all VPCs. The default next hop of the vFWs of all VPCs is the public vFW, as shown in Figure 5-8.
  - The default next hop of the public vFW is the public vRouter.
  - When an EIP address has been allocated to a VPC, the public vFW sets up a static route to the vFW of the VPC. The destination IP address of the static route is the EIP address of the VPC, and the next hop of the static route is the IP address of a virtual interface on the vFW
- **Virtual system (VSYS)**: is a logical device divided on a firewall. Each VSYS works independently. In Huawei CloudFabric Solution, the tenant VSYS serves as the summary routing domain of the VSYSs of all VPCs of a tenant for accessing remote intranets outside the data center. The next hop of the VSYSs of all VPCs of a tenant is the tenant VSYS, as shown in Figure 5-8.
  - All VPCs of a tenant share one tenant VSYS.
  - The default next hop of the tenant VSYS is the tenant vRouter.



- Border Gateway Protocol-Ethernet Virtual Private Network (BGP-EVPN) defines a new BGP Network Layer Reachability Information (NLRI), called the EVPN NLRI.
- EVPN can function as the VXLAN control plane by using inclusive multicast routes (IMRs) carried in the EVPN NLRI. VTEP IP addresses are stored in the Originating Router's IP Address field of an IMR.
- BGP-EVPN applies to VXLAN networks. In Huawei CloudFabric Solution, BGP-EVPN is widely used in centralized and distributed VXLAN scenarios.
- BGP-EVPN provides the following functions:

- **Automatically establishes VXLAN tunnels.**
  - VXLAN standards do not define protocols for setting up tunnels. Manual configuration is inefficient and prone to errors. BGP-EVPN enables automatic information exchange for establishing VXLAN tunnels between devices.
- **ARP broadcast suppression**
  - You can use BGP-EVPN to advertise ARP routes to NVEs. After the configuration, when the gateway receives an ARP request, it will first check whether host information exists according to the destination IP address. If so, the gateway replaces the broadcast MAC address in the ARP request packet with a unicast MAC address, and converts the broadcast ARP packet into a unicast ARP packet.

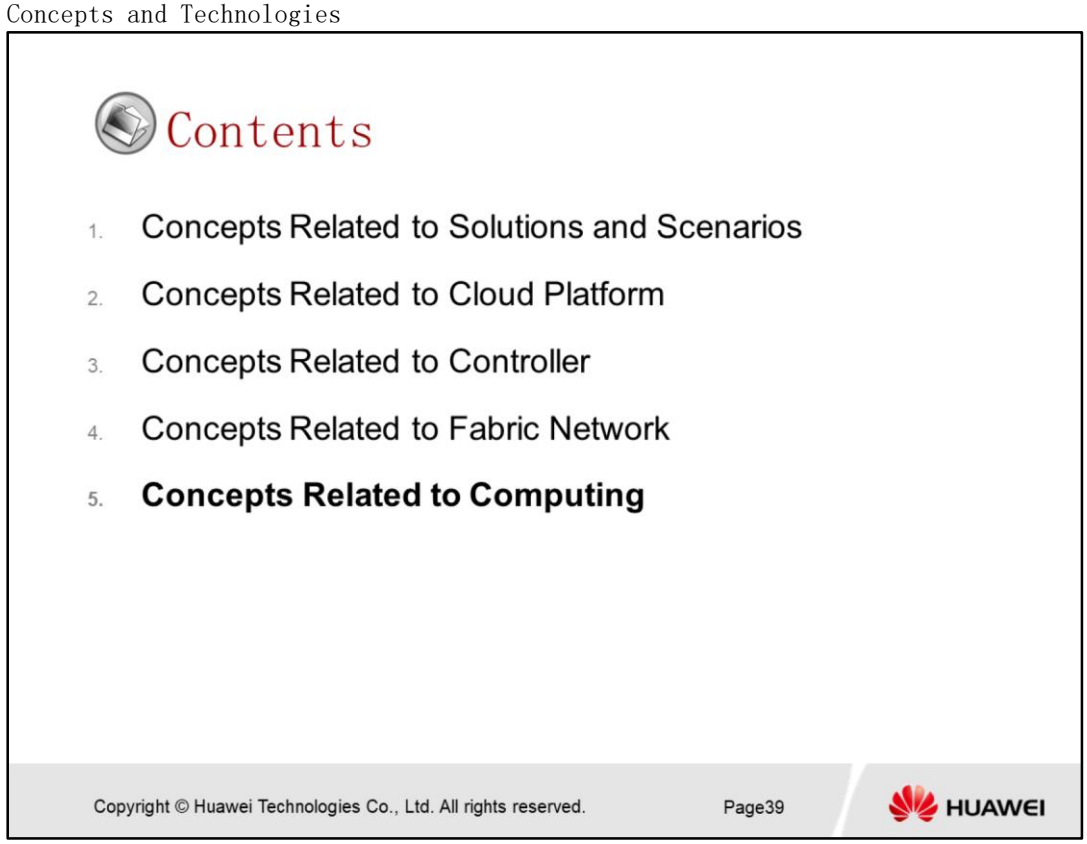


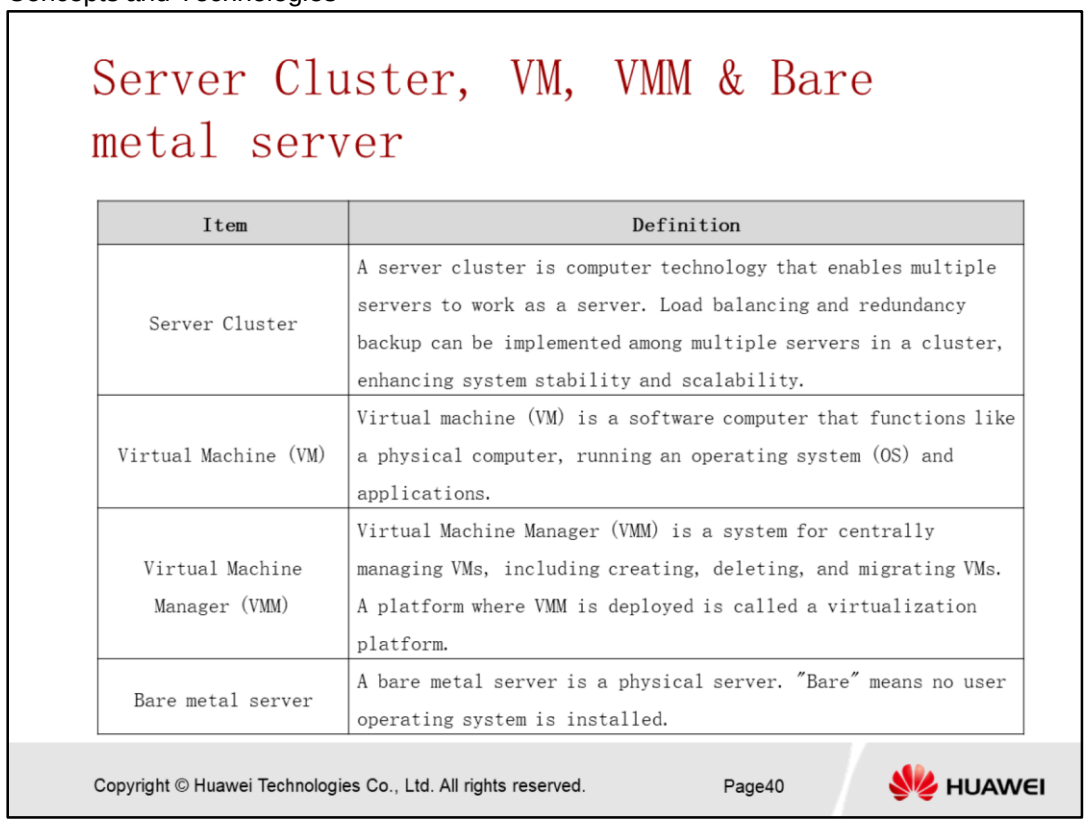
- In a large Layer 2 data center network, there are many physical and virtual machines. A large number of ARP broadcast packets will pose huge pressure on gateways and tenants. Gateways may be affected greatly because they need to learn and manage ARP entries. In addition, a massive number of ARP packets will occupy high bandwidth.

- Huawei CloudFabric Solution offers an ARP broadcast suppression solution. Diagram on the slide above shows the implementation of ARP broadcast suppression.

1. The spine gateways dynamically learn tenants' ARP entries and generate host information (including IP addresses, MAC addresses, VETP addresses, and VNI IDs) based on the entries.
2. The leaf switches synchronize host information generated by the spine gateways through BGP-EVPN.
3. When the leaf switches receive local ARP requests, they convert broadcast packets into unicast packets based on host information before forwarding packets.

- You can also configure ARP broadcast suppression on the AC-DCN. After the configuration, gateways convert tenants' ARP broadcast packets into ARP unicast packets based on routing information with known destination IP addresses, reducing ARP packet flooding.





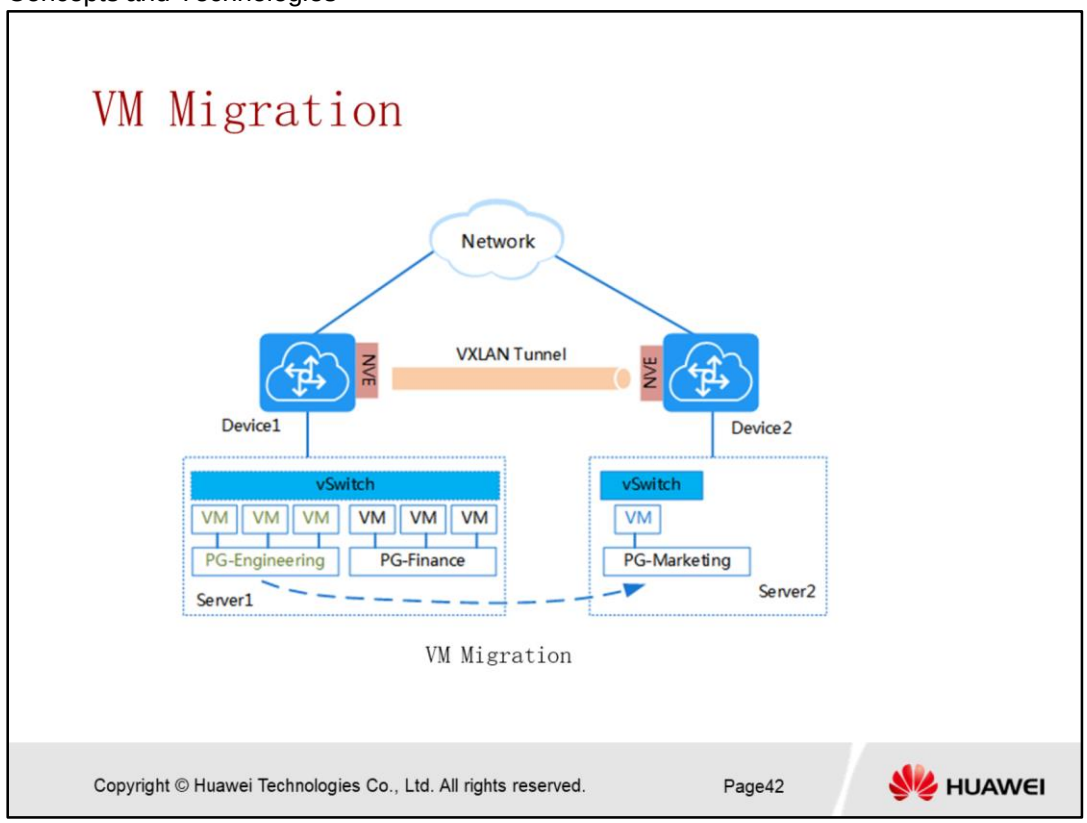
- **Server Cluster**
  - A server cluster is computer technology that enables multiple servers to work as a server.
  - Load balancing and redundancy backup can be implemented among multiple servers in a cluster, enhancing system stability and scalability.
  - In most cases, the AC-DCN is deployed in cluster mode. In the cluster networking, all servers where the AC-DCN is installed are cluster nodes. Deploy at least three AC-DCN cluster nodes.

- **Virtual machine (VM)**
  - A virtual machine (VM) is a software computer that functions like a physical computer, running an operating system (OS) and applications.
  - Multiple VMs can run on the same physical server. These VMs share the same underlying hardware, and function as one physical server for applications. Each VM has independent virtual resources, including an OS, CPU, memory, disk space, and I/O devices (such as NICs).
  - VMs provide the following functions:
    - Partitioning: You can run multiple VMs on one physical server.
    - Isolation: You can isolate different VMs on one physical server.
    - Encapsulation: You can save a VM in files, and migrate and copy this VM by moving and copying the files.

- **Virtual Machine Manager (VMM)**
  - Virtual Machine Manager (VMM) is a system for centrally managing VMs, including creating, deleting, and migrating VMs.
  - A platform where VMM is deployed is called a virtualization platform. Mainstream virtualization platforms include the open-source KVM and Xen and the private VMware vSphere and Microsoft Hyper-V.
  - In Huawei CloudFabric Solution, VMware, Microsoft, and Kernel-based Virtual Machine (KVM) are selected as VMM systems.
  - KVM performance is improved by hardware-assisted virtualization. With the continuous and rapid optimization of Linux system versions, KVM will surpass Xen in the near future. Therefore, in the open-source virtualization scenario, KVM is preferred.
  - Among all open-source KVM systems, Red Hat's Linux KVM and Huawei's FusionSphere UVP are the first choice. (UVP refers to a unified virtualization platform.)Red Hat is an industry leader, and Huawei's FusionSphere UVP is the fastest-growing system in China.

- **Bare metal server**
  - A bare metal server is a physical server. "Bare" means no user operating system is installed.
  - Bare metal servers offer excellent computing performance, and satisfy service requirements on high performance, stability, and data security and monitoring. Therefore, they are also called dedicated user servers.
  - Similar to other physical servers, resources of bare metal servers are distributed and

released by FusionSphere or other cloud platforms

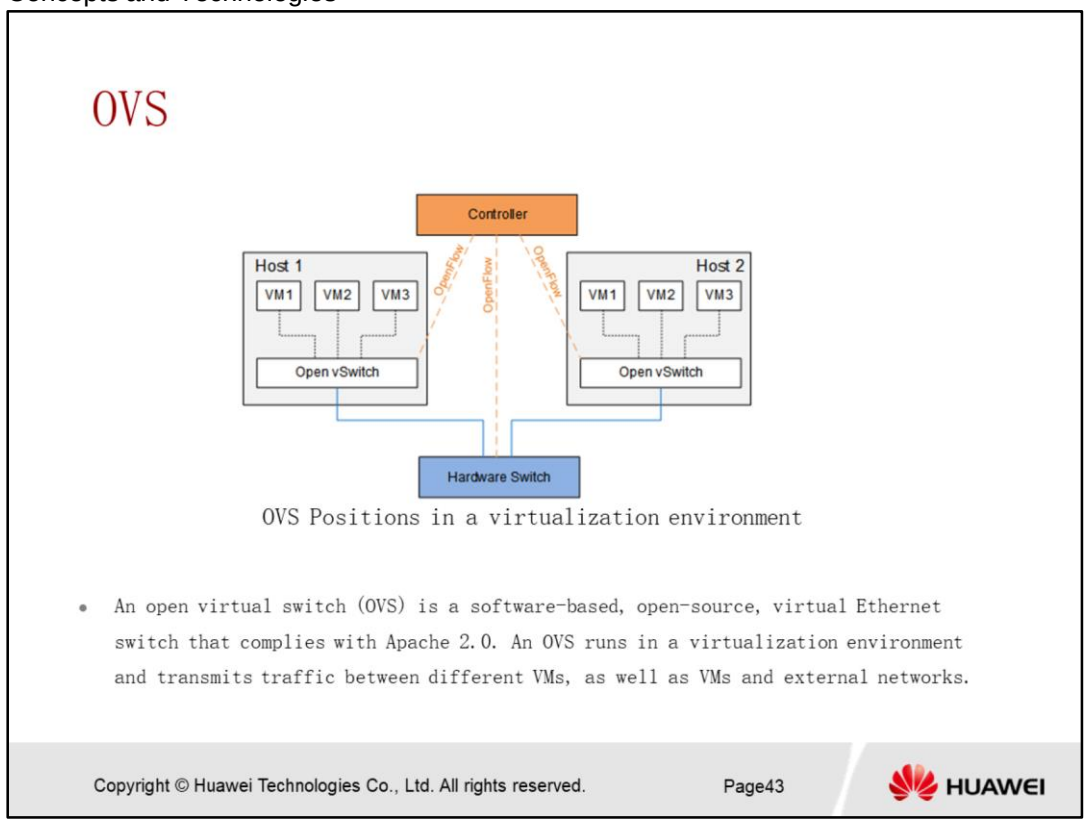


- During VM migration, a VM migrates from one server to another server.

- As shown in the figure above , an enterprise has two servers in its data center. Server1 serves the engineering and finance departments, and Server2 serves the marketing department. The computing resources on Server1 are inadequate, whereas the computing resources on Server2 are not fully utilized. The network administrator wants to migrate the engineering department to Server2 without affecting services.

- Available VM migrations are as follows:

1. **Live migration**
   - Live migration is also called online migration. VMs are migrated from one server to another server when they are running. During live migration, services are not interrupted.
2. **Cold migration**
   - Cold migration is also called offline migration. VMs are migrated from one server to another server when they stop running. During cold migration, services are interrupted.

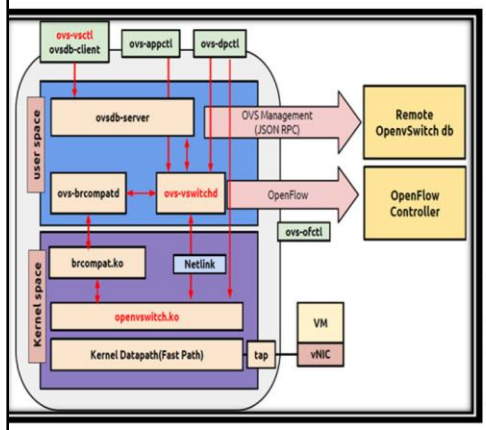


- An open virtual switch (OVS) is a software-based, open-source, virtual Ethernet switch that complies with Apache 2.0.
- The objective of OVS is to manage and configure VM networks in a better way. OVS supports various management interfaces, including NetFlow, sFlow, RSPAN, ERSPAN, and CLI.
- Compared with physical switches, OVS has the following advantages:
  - Flexible configuration
    - OVS is based on software. Therefore, hundreds of OVSs can be deployed on one physical server, and the port quantity can be adjusted as required.
  - Lower costs
    - A switching speed of 10 Gbit/s can be achieved through software configuration.
- An OVS runs in a virtualization environment and transmits traffic between different VMs, as well as VMs and external networks.
- The working principle of OVSs is similar to that or physical switches. An OVS connects one physical NIC to multiple virtual NICs. An OVS maintains a mapping table to find links to VMs and forward traffic to them. A mapping table is maintained in an OVS so that data can be forwarded through VM links according to MAC

addresses.

- Additionally, OVSs support the OpenFlow protocol for communicating with the AC-DCN.

OVSDB and OVS



OVS Architecture

No	Module	Explanation
1	ovs-vswitchd	Daemon to achieve switching function, and is compatible with Linux kernel module, based on current exchange flow-based switching
2	ovsdb-server	Lightweight database service, mainly to save the entire configuration information for the OVS, including interfaces, exchange content, VLAN, and so on. Ovs-vswitchd will work according to the configuration information in the database.
3	ovs-vsctl	Query or change the configuration information of ovs-vswitchd. This tool operation will update the database in ovsdb-server.
4	ovs-brcompatd	Replace Linux original bridge, listen to bridge ioctl commands
5	ovsdb-tool	Manage OVS database
6	ovsdb-client	OVS database JSON-RPC client
7	ovs-dpctl	Manages the datapath kernel module, can control the forwarding rules
8	ovs-appctl	Send commands to OVS daemon (e.g. log settings command)
9	ovs-ofctl	Manage the OpenFlow switch
10	ovs-pki	Manage the OpenFlow public key infrastructure

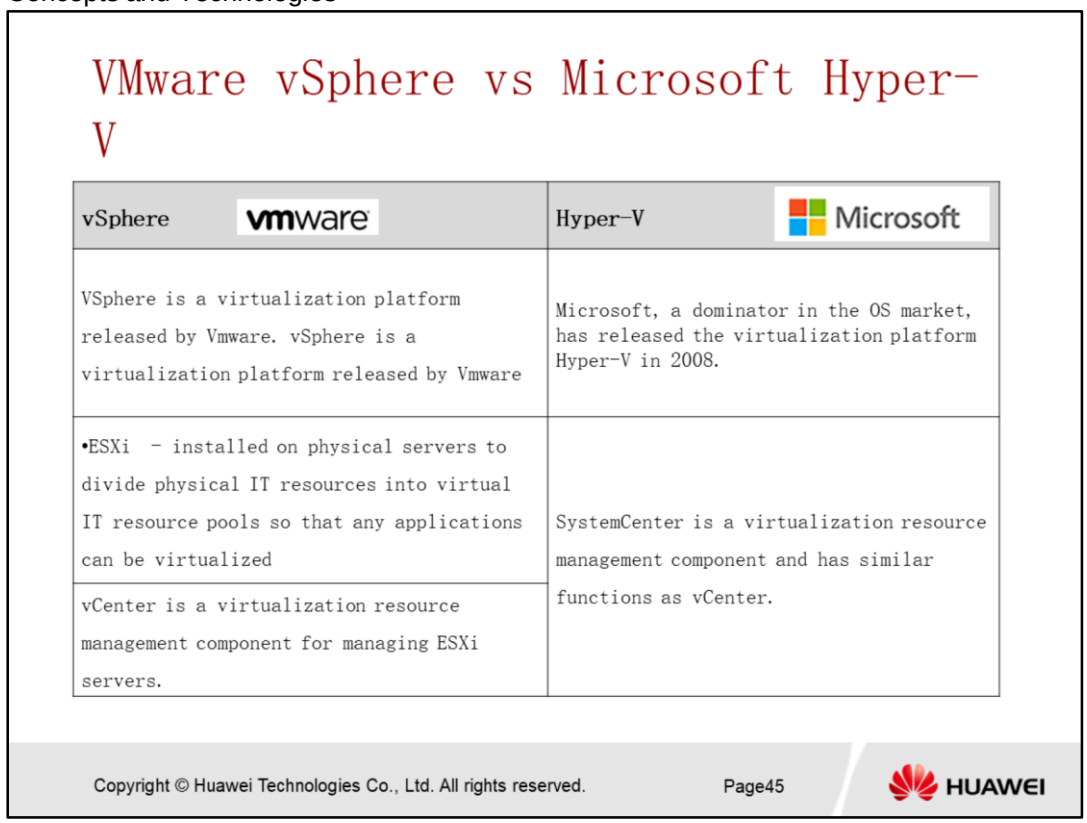
HUAWEI

- **Reference:** http://openvswitch.org/support/dist-docs/

- OVSDB stands for Open vSwitc Database Management Protocol; it is used to build communication between AC-DCN and vSwitch. An example of the application of connection establishment between AC-DCN and CE1800V (Huawei vSwitch) are as below:

1. CE1800V automatically connects to AC-DCN cluster south bound floating IP
2. After AC-DCN receives CE1800V connection request, it will record equipment information of CE1800V.
3. AC-DCN confirm internally on the connection node with CE1800V
4. Connection is established between AC-DCN and CE1800V.

- OVS stands for Open Virtual Switch is a open source virtual switch designed for the ease of management and configure network for virtual network; thus a OVS can support multiple types of management interface protocols such as Netflow, sFlow, RSPAN, ERSPAN, CLI etc.
- The main functions of OVS is to forward traffic between VMs, and to allow VMs to communicate with external network.



- **Vmware vSphere**
- VMware is founded in 1998 and claims to be an important promoter of virtualization technology.
- vSphere is a virtualization platform released by VMware. This platform is similar to Microsoft Hyper-V.
- vSphere is an ideal fundamental platform in the cloud computing environment for its following features:
  - High scalability, performance, and availability, allowing users to virtualize any applications
  - Powerful and simple tools for management on the creation, resource sharing, deployment, and migration of VMs.

- **Microsoft Hyper-V**
- Microsoft, a dominator in the OS market, has released the virtualization platform Hyper-V in 2008.
- Hyper-V uses hypervisor (a system management program) technology, which is similar to that in VMware. Hyper-V is designed to compete with VMware. Hyper-V manages and schedules VM creation and running, and enables virtualization of hardware resources.
- SystemCenter is a virtualization resource management component and has similar functions as vCenter. The external interfaces provided SystemCenter need to be optimized by installing an AC-DCN plug-in. This plug-in optimizes VM login and logout message detection and native

vSwitch management and control.

- In Huawei CloudFabric Solution, SystemCenter is connected to a cloud platform in the cloud-network integration scenario, or is connected to the AC-DCN in the network virtualization scenario.