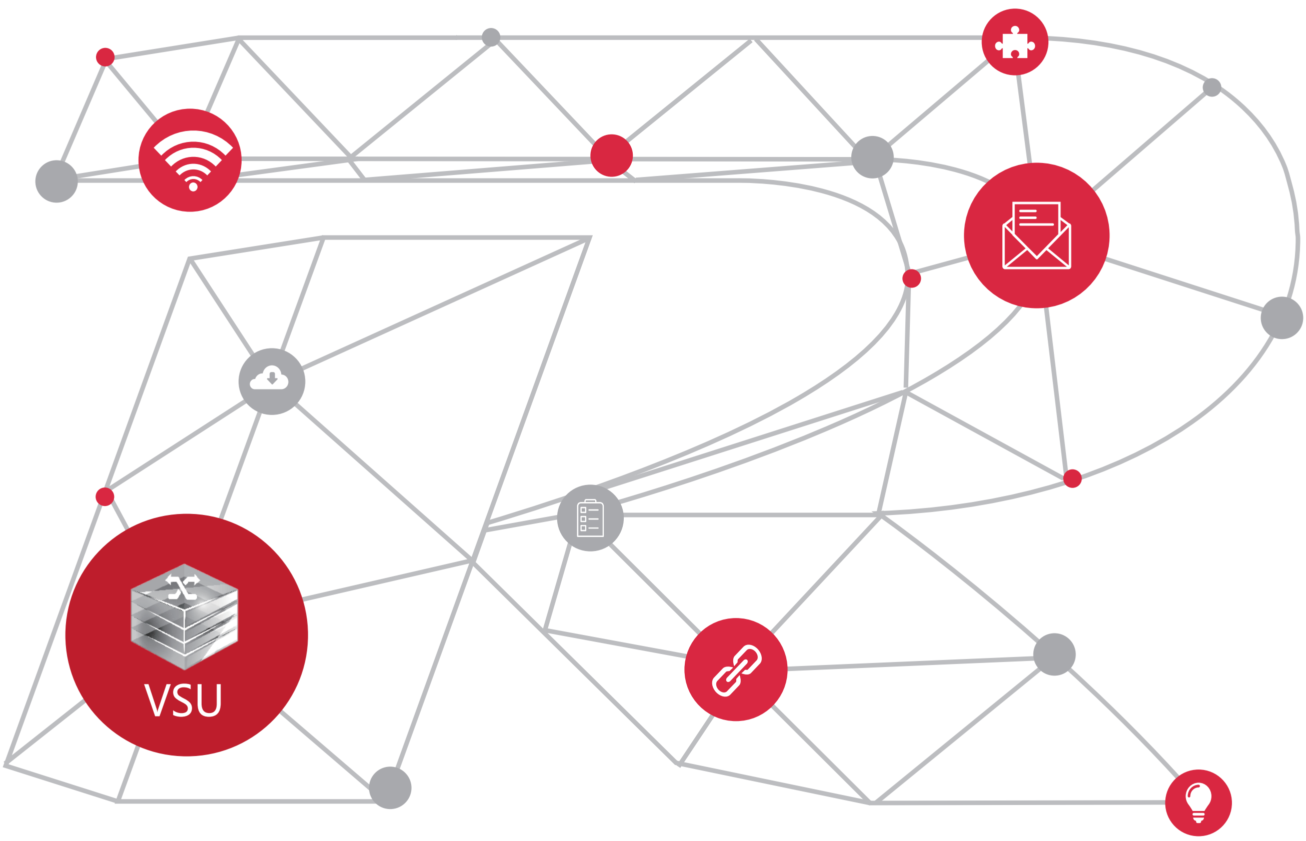
Ruijie Virtual Switching Unit

White Paper



Contents

Introduction

This document details the technologies, functionalities and performance indexes of the Virtual Switching Unit (VSU) from hardware components to application scenarios, from standalone chassis components to multi-chassis VSU.

VSU enables the merging of two or more physical chassis together into a single, logical entity. This technology enables enhancements in all areas of network design, including high availability, scalability, management, and maintenance.

Other similar technologies include Cisco Virtual Switching System (VSS), Huawei Cluster Switch System (CSS), and H3C Intelligent Resilient Framework (IRF). Those technologies are similar, and each has its unique characteristics and related patents, yet there is no unified standard.

In the network structure, high availability usually results in complex configuration and deployment, and network scalability complicates network management. It is difficult to enhance network high availability and scalability, and simplify network management at the same time. However, VSU solves the two bottlenecks and has the following advantages:

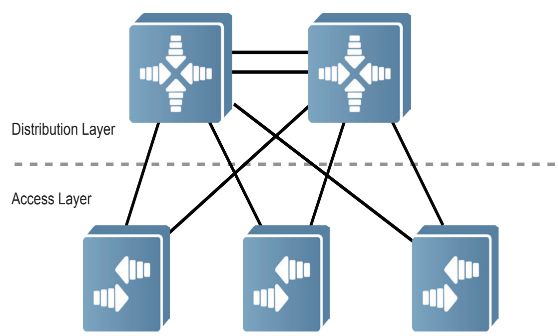
* **High availability and simplified network management;**
* **Simplified network scalability.**

• High Availability and Simplified Network Management

A VSU system enhances network high availability and simplifies management.

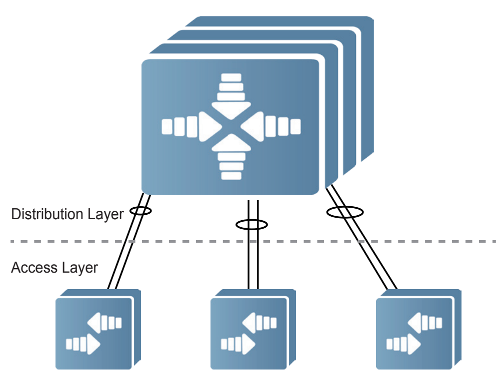
The typical network topology is designed as in Figure 1. To improve network availability, each access switch connects to two distribution switches. To enhance network availability, the topology is getting more and more complex and causes a loop. To eliminate loops and support gateway backup, Multiple Spanning Tree Protocol (MSTP), Virtual Router Redundancy Protocol (VRRP) and other protocols are required, which result in complicated configuration. To load-balance traffic, you have to enable Ethernet Ring Protection Switching (ERPS) and other multi-ring network technologies.

Figure 1 Typical Network Topology



In a VSU system, MSTP, VRRP, or ERPS are not required. With normal link aggregation, you can load-balance traffic, eliminate loops, support redundancy backup, and simplify network topology and management, as shown in Figure 2.

Figure 2 Ruijie VSU Network Topology



• Simplified Network Scalability

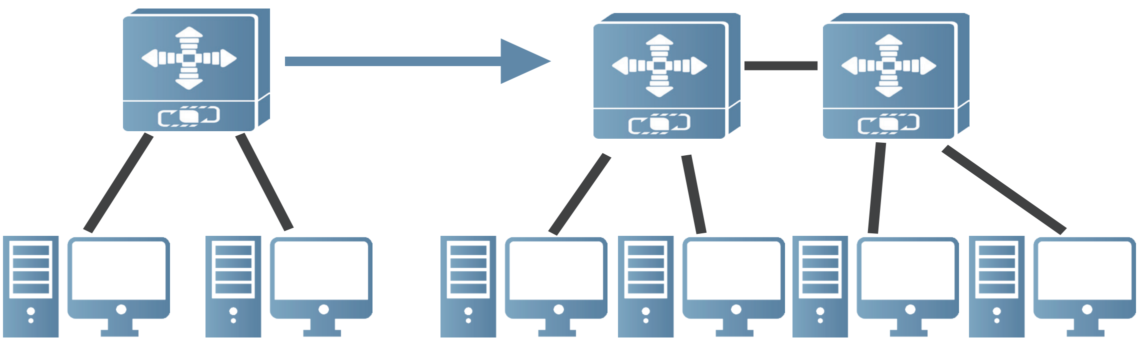
Except simplified network management, a VSU system simplifies network scalability.

Generally, network scalability complicates network management. When ports of a standalone switch are insufficient, the option is to replace it with high-port-density devices or add devices.

Replacement of devices results in a waste of investment, service interruption, and configuration reload.

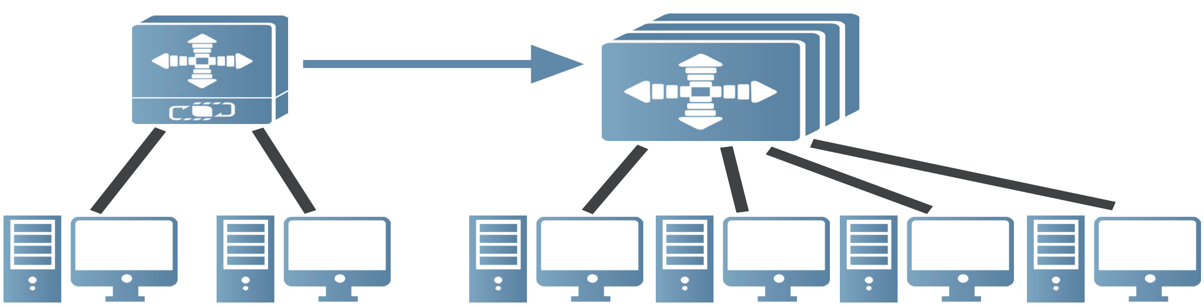
Deploying additional devices requires you to adjust configurations (for example, configuring VLAN) of the existing and additional devices because of the topology change (as shown in Figure 3). The more devices you add, the more repeated logins and switchovers you perform.

Figure 3 Typical Network Topology Change



However, in Ruijie VSU network (Ruijie VSU supports up to nine member devices), you can reduce the number of devices used for network maintenance to 1/9. You can combine nine switches that originally formed a ring into a VSU, that is, a logical entity. At the initial phase of network design, only one switch is required. With the increasing terminal devices, more switches are involved. With VSU, enhancing network scalability is as simple as inserting a line card into a chassis (as shown in Figure 4).

Figure 4 Ruijie VSU Network Topology Change



Concepts

To better understand switches and VSU systems, we will first introduce hardware components of general switches and VSU, as well as VSU systems.

• General Switches’ Hardware Components

The standalone switch components are shown in Figure 5 and Figure 6. A box switch consists of a general CPU, and one or more application-specific integrated circuits (ASICs). Some ASIC integrates general CPU, and the logic is the same. If the output ports of one ASIC is insufficient, more chips are needed. In this situation, a box switch requires only one CPU, and the CPU manages multiple ASICs (usually through PCIE interface). The ASICs transmit internal messages of chips through chip-specific hardware interfaces. The chip-specific hardware interfaces are not compatible between different suppliers, or even between different product series of the same supplier.

Figure 5 Standalone Switch Components(a)

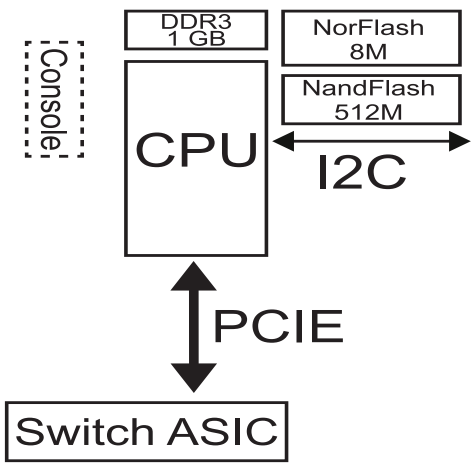
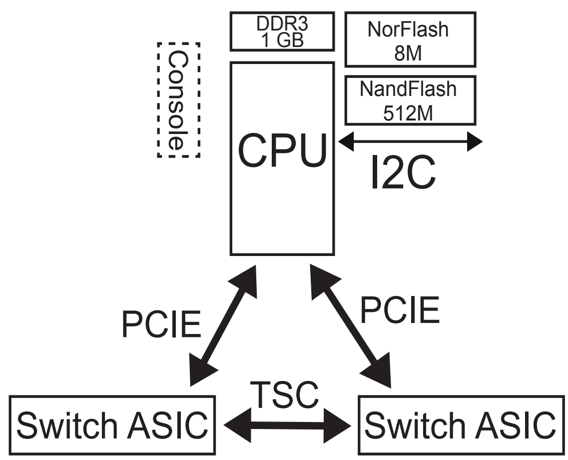


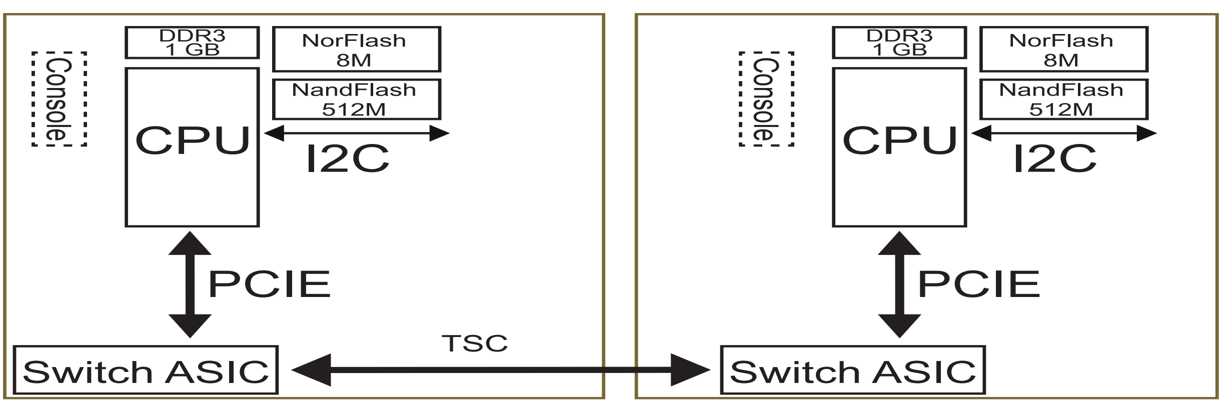
Figure 6 Standalone Switch Components(b)



The vital difference between multiple ASICs switch and a single ASIC switch is that multiple ASICs have to synchronize some entries to all ASICs, such as MAC address table. When a user adds a filter address, the CPU writes the address to all ASICs, to ensure that no matter which ASIC ports the packet (to be filtered) is input, the entry is retrieved directly. For example, link aggregation ports support their member ports to be distributed on different ASICs, which require the CPU to configure each ASIC link aggregation respectively.

• VSU Hardware Components

Figure 7 VSU Hardware Components



As shown in Figure 7, two switches form a VSU. Functionally, this is equivalent to one switch configured with two CPUs, with one CPU managing one ASIC respectively.

The two switches are connected through VSU-specific or ordinary cables. The two ports, which form the VSU, work in the same mode as how multiple ASICs are connected in a standalone chassis. The purpose is to transmit internal messages between ASICs, with multiple chips coordinated as one chip.

In a VSU system, when users add a static filter address, the address is written to all ASICs of the VSU system, to ensure that the packet is filtered correctly whichever the VSU member devices the packet is input. Member ports of a link aggregation distribute on one of the VSU member devices. The difference is that though a VSU system has multiple CPUs, only one CPU receives input at one time, and this CPU is located on the active switch. The active switch deals with user requests and distributes the requests to CPUs of all VSU member devices, and then the member devices write the requests to ASICs.

A standalone chassis manages one CPU. A VSU system manages multiple CPUs, and the CPUs communicate with in-band mode. There is no independent communication path between CPUs. The benefit is that you can use ordinary twisted-pair cable or fiber to form a VSU.

• VSU System

A VSU system has multiple standalone switches. And for users, there is only one entity. A set of protocols is required for multiple switches, to assign user data and configuration requests to each member device.

Each member device is assigned one of the three roles: active, standby, or candidate. A VSU system enables one active switch, one standby switch at most (some solutions enables multiple standby switches, and this document introduces the single standby switch technology), and one or more candidate switches (if any).

The active switch deals with user requests (such as commands issued on the Console, Telnet requests, Simple Network Management Protocol (SNMP) requests, and so on), and various packets of management protocols on switches (such as BPDU of STP, LSA of OSPF, and so on) with the highest priority.

The standby switch takes over the responsibilities of the active switch within 50 ms when power failure or breakdown occurs to the active switch, and nominates a candidate switch (if any) to be the standby. In this process, the standby switch receives backup data from the active switch continuously.

Candidate switches receive commands from the active switch, complete ASIC configuration, and forward data as required by users.

The assignment of the three roles is as follows:

Principle for determining the active switch

1. The switch with the highest priority specified by users is the active switch;
2. If not specified, the switch with the numerically smallest MAC address is the active switch.

Principle for determining the standby switch

1. The switch connecting directly to the active switch is the standby switch; or
2. The switch with the highest priority specified by users is the standby switch;
3. If not specified, the one with the numerically smallest MAC address is the standby switch.

In a VSU system, if a switch is neither active nor standby, it is candidate.

VSU Core Technologies

A robust VSU system supports inter-device link aggregation, inter-device packet mirroring, non-blocking communication between VSU member devices, coexistence of multiple VSU systems within a network, VSU ring, VSU system split and merge, dual-active detection, fast switchover, and in-service software upgrade. You can form a VSU within a long distance, start up a VSU system in a short time, and manage VSU through console and MGMT ports of any member devices. Ruijie VSU supports up to nine member devices.

• Inter-device Link Aggregation

VSU supports inter-device link aggregation. As previously introduced, whether a switch is composed of one or more ASICs, ports on different ASICs can form a link aggregation port. The functionalities and performance of the inter-device link aggregation is the same with that of member ports on the same ASIC.

In a VSU system, ports of different member devices also form a link aggregation port. Only in this way can switches at the distribution layer in Figure 1 form a VSU. When one switch breaks down, the network still works, thus providing redundancy backup and availability. When both switches work, they can load-balance traffic and extend life of the switches.

• Inter-device Packet Mirroring

VSU supports inter-device packet mirroring. The mirroring function is mainly used to troubleshoot and monitor networks.

Similar to inter-device link aggregation, the source port and the destination port of the mirror can be distributed on different ASICs. In a VSU system, the source port and the destination port can also be distributed on different VSU member devices.

• Fast to Form VSU System

Ruijie VSU technology supports fast setup and startup of VSU systems.

To support a VSU, the switch software becomes dramatically complicated. For most similar technologies, there is a sharp rise in startup time after enabling a VSU.

However, after optimizing design, the startup time of a Ruijie VSU system is close to the startup time of a single switch. Different systems and different versions have different optimization solutions.

• Non-blocking Communication Between VSU Member Devices

VSU supports non-blocking communication between member devices. For example, two switches with 24 10/100BASE-T ports form a VSU. To achieve non-blocking communication across devices, the bandwidth of the connection ports between switches must be greater than 2.4G in one direction, and support full-duplex mode. Therefore, you can deduce the bandwidth of the VSU created by switches of other ports.

Take S2910C-24GT2XS-HP-E switches for example, an S2910C-24GT2XS-HP-E switch supports up to four 10G-ports (two of them are expansion modules) to form a VSU, and the bandwidth is 40G in total. A standalone chassis has 24 Gigabit Ethernet ports, and requires 24G bandwidth in total, thus supporting three S2910C-24GT2XS-HP-E switches to create a non-blocking VSU. Three S2910C-24GT2XS-HP-E switches form a ring topology, and the bandwidth between each two switches is 20G. To achieve non-blocking communication, the VSU is required to provide 144Gbps (bidirectional) switching capacity, and the backplane bandwidth of the full mesh structure formed by three S2910C-24GT2XS-HP-E switches is 240Gbps (bidirectional), so the bandwidth is sufficient for 144Gbps required by VSU.

The following traffic distribution cases will help you better understand how traffic is distributed for non-blocking communication.

1. Each switch possesses 24G traffic input, and distributes 12G traffic output to the two adjacent switches respectively. And the 24G traffic is input from the adjacent two switches. The inter-device link carries 12G traffic in each direction, and 8G bandwidth is free. The 8G free bandwidth is used to cope with HASH collisions of the aggregation link.
2. Considering the aggregation link HASH collisions, unequal traffic distribution might occur.
3. If 24G traffic inputs from one switch, and outputs from the other switch, then the third switch only forwards traffic within devices. Hence the bandwidth is significant enough.

Figure 8

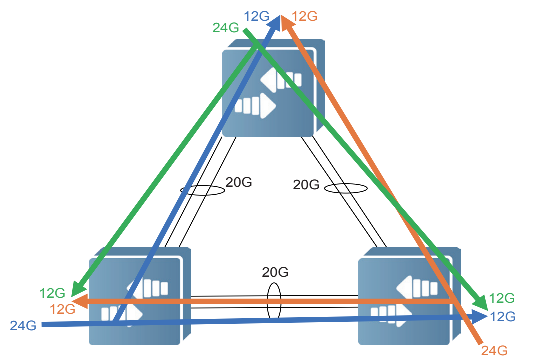


Figure 9

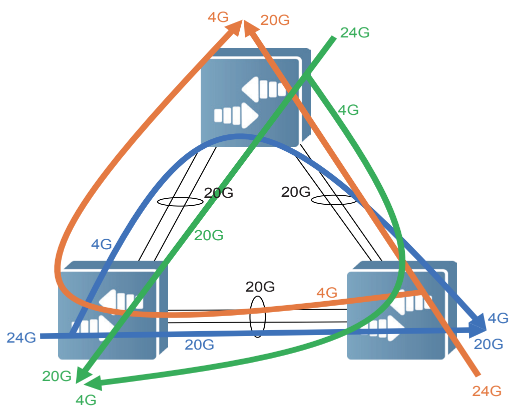
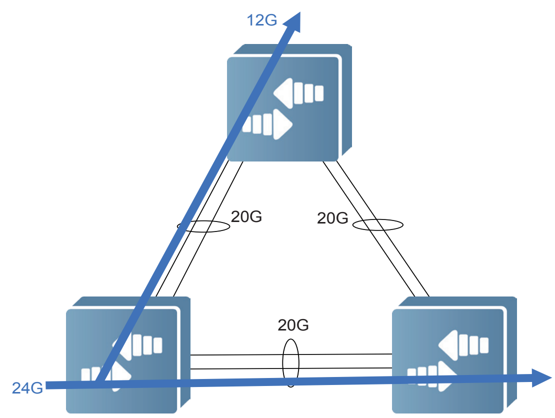


Figure 10



With the listed cases, you can deduce other traffic distribution situations. When there are more than three switches, non-blocking communication cannot be achieved because it is not the full mesh connection.

• VSU Management Through Console and MGMT Ports of Any Member Devices

To simplify network management, Ruijie VSU supports out-band management through the Console or MGMT port of any switch, and in-band management through Layer 3 interfaces of any switch.

Each VSU device has three roles (active, standby, and candidate), and each role has respective responsibilities. To manage a VSU, you had better connect to the console port of the active switch or other service ports, which is the shortest path for physical data transmission.

If a VSU does not support the aforementioned management, it is difficult for users to identify which switch is currently active, because switches usually do not have display devices.

• Maximum Number of VSU Member Devices

The maximum number of Ruijie VSU member devices varies because different series have different limitations. It is nine for the access layer and the distribution layer, four for the core layer.

The number of member devices supported by a VSU is restricted by two factors:

a.The number of ASIC chips that support distributed collaborative work. In general, ASIC supports distributed coordination of tens or hundreds of chips, so in theory, VSU member devices range from 10 to 100.

b.CPU capacity and memory. More devices mean more user configurations and messages sent to the CPU, thus consuming more memories.

The CPU capacity and memory of a switch is usually not high, so the above factor b is the bottleneck of increasing VSU member devices.

• Coexistence of Multiple VSU Systems Within a Network

When multiple VSU systems coexist within a network, they distinguish each other with unique domain identifiers (domain IDs). Only switches with the same domain ID can form a VSU. The domain ID ranges from 1 to 255, and the default value is 100.

If multiple VSU systems cannot coexist within a network, VSU member devices can only be connected in point-to-point mode (by using special cables). It is difficult to lay out cables across floors or buildings

• VSU Ring

VSU topology supports two structures: the linear topology (Figure 11) and the ring topology (Figure 12).

The ring topology is more reliable than the linear topology. In the linear topology, when a Virtual Switching Link (VSL) breaks down (for example, a VSU member device shuts down), the VSU splits. However, in the ring topology, when a VSL fails, the ring topology becomes the linear topology, and the VSU services are not affected.

Figure 11 Linear Topology

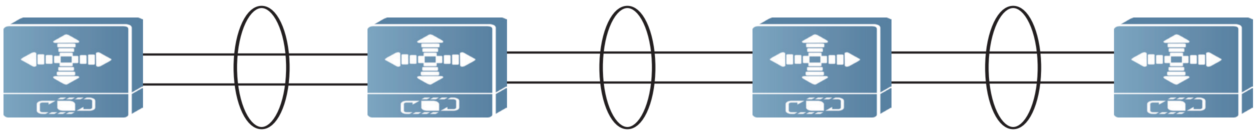
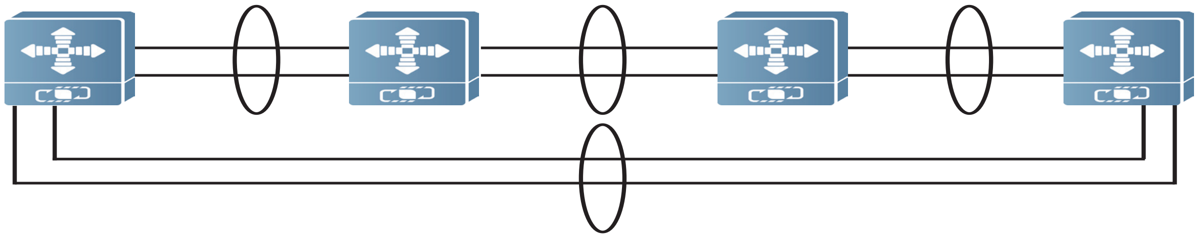


Figure 12 Ring Topology

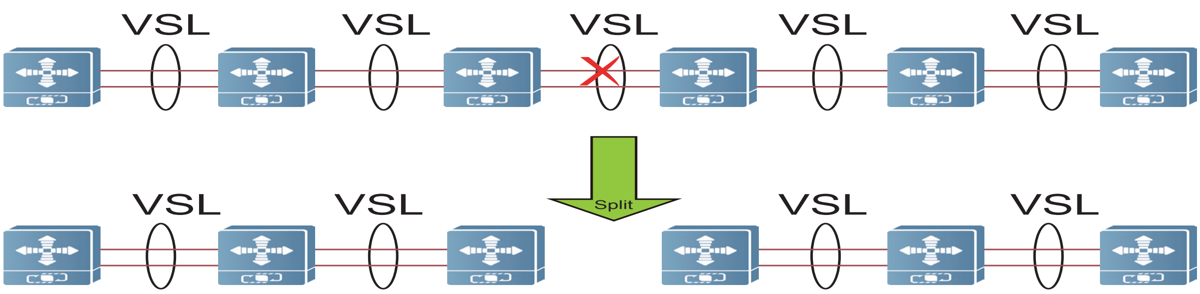


• VSU System Split

In the linear topology, when a VSL breaks down, a VSU system splits into two. The advantage of VSU split is that even if a VSL breaks down, only partial system is under influence, and access devices of the rest VSUs still work well.

After a VSU split, two switches duplicate the IP address and MAC address in a network, causing IP address conflicts and network chaos. Thus, dual-active detection is required to prevent network failures caused by VSU split.

Figure 13 VSU System Split



• Dual-active Detection

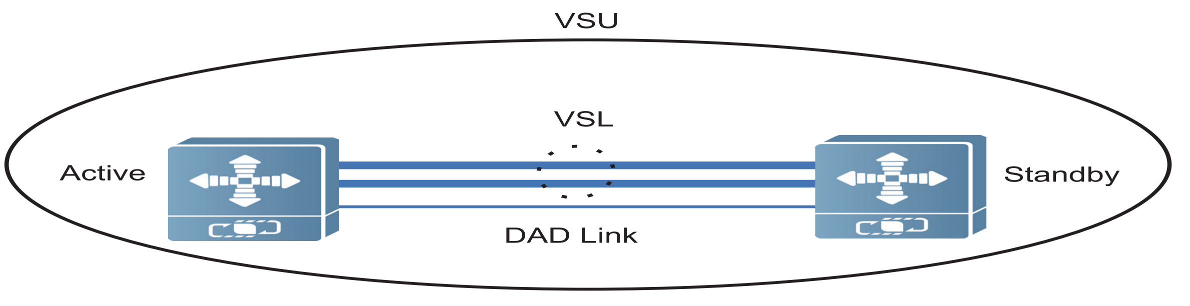
Ruijie supports bidirectional forwarding detection (BFD) and aggregation port to detect dual-active chassis.

When a VSL breaks down, the active switch and the standby switch are assigned to different VSUs, which result in VSU split. For the standby switch, the active switch fails, so the standby switch takes over as the active switch. From the active switch perspective, the standby switch breaks down, and a candidate switch takes over the standby role. After the VSU splits, two VSUs with the same configuration exist in the network. At Layer 3, any virtual interfaces (VLAN interface, loopback interface, and so on) duplicate configuration on both switches, causing IP address conflict.

To enable BFD and aggregation port to detect dual-active chassis, you have to set up a dual-active detection link (DAD link) between the two switches. When a VSL is disconnected, the two switches send detection packets through DAD link. If one switch receives the detection packets from its peer, it indicates that the peer switch is still running normally, and two chassis exist.

*** BFD detection (Figure 14): The BFD dual-active detection (DAD) port must be a routing port. Layer 3 aggregation ports, or switch virtual interfaces (SVI) cannot act as BFD detection ports. When the detection port is switched from Layer 3 routing interface to other types of ports, the BFD DAD configuration is automatically cleared and a prompt is displayed. BFD detection adopts scalable BFD, and you cannot configure the DAD ports with the existing BFD configuration and commands.**

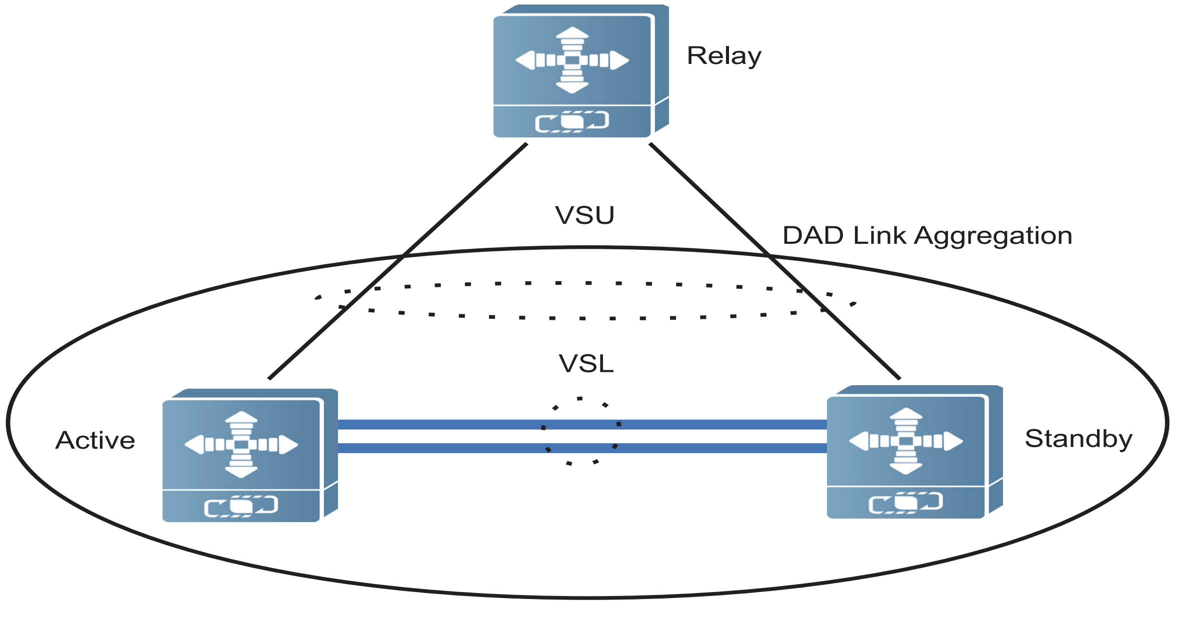
Figure 14 BFD Detection



*** Aggregation port detection (Figure 15): The mechanism of detecting dual-active chassis based on the aggregation port is similar to BFD. When a VSL breaks down and generates two active chassis, the two chassis send each other private messages specific to aggregated ports to detect multiple chassis.**

The main difference between BFD detection and aggregation port detection is that the aggregation port detection is configured on the cross-switch service aggregation port (not the VSL aggregation port), and requires neighboring switches to forward private detection packets.

Figure 15 Aggregation Port Detection

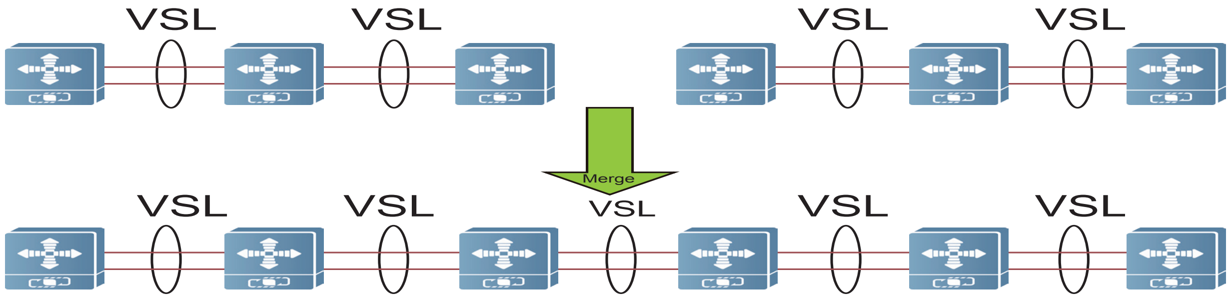


• VSU System Merge

The VSL breakdown might cause a VSU system to split into two. When the VSL breakdown restores, the two VSU systems merge into one.

As shown in Figure 16, two stable VSU systems work independently, and they have the same domain ID. By adding a VSL, the two VSU systems merge into one. This process is called VSU merge.

Figure 16 VSU System Merge



• Fast Switchover Between the Active and Standby Switches

The switchover between the active switch and the standby switch is fast, and the traffic interruption is less than 50 ms.

The switchover between the active switch and the standby switch is equivalent to the switchover between 2 control engines of the chassis.

If the active switch breaks down, the standby switch assumes the active role and provides network services continuously, and only terminal devices connecting to the original active switch are affected. The traffic interruption of terminal devices connecting to the rest VSUs is 50ms at most.

If you run the CLI command to force a switchover, a 50ms interruption (at most) occurs to all terminal devices.

• Long-distance VSU

VSU technology enables devices far from each other to form a VSU. To form a VSU within a distance of meters or hundreds of meters, or even tens of kilometers, you can choose ordinary twisted-pair cables (up to 100 meters), or fibers. The Gigabit fibers reach up to 100 km, the 10-Gigabit fibers reach up to 40 km, and the 40G fibers reach up to 10 km.

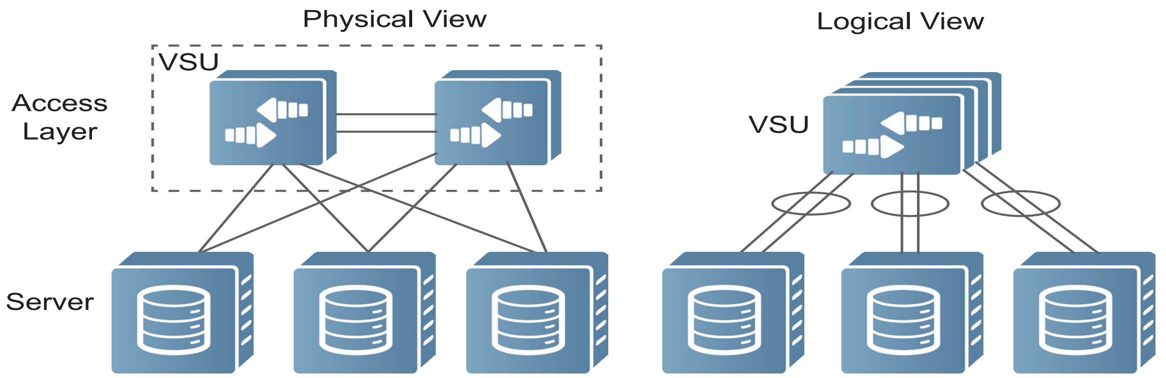
• In-service Software Upgrade

VSU supports uninterrupted services during software upgrade.

Some key networks require non-blocking operations, such as network operators, banks and other industries. With business development, new requirements are proposed for network equipment, which in turn requires switches to upgrade software system. Generally, the operating system upgrade takes effect after the device restarts. For example, upgrading the Windows system will cause service interruption.

A VSU system upgrades the standby switch first, and assigns the upgraded switch to take over as the active switch, and then upgrades the new standby switch. In this way, the entire system completes upgrade gradually. Thus in a topology that contains a backup link, the system is upgraded without service interruption. As shown in Figure 17, the VSU system upgrade has no effect on the services on servers.

Figure 17 VSU System Upgrade



FAQs

• Q: What are the differences between VSU management and cluster management?

A: Clusters use multiple identical devices to provide the same services at the same time, improving service capabilities and availability. VSU is a detailed technology to implement cluster management.

• Q: Will the capacity of routing information base (RIB) and MAC address table increase after multiple switches form a VSU?

A: Except for an increase in port number, other RIB and MAC address will not increase, and their capacity is the same as standalone chassis.

• Q: Can switches of different models of the same series form a VSU?

A: Yes. For example, RG-S5750C-28GT4XS-H switches and RG-S5750C-48GT4XS-H switches can form a VSU.

• Q: What makes VSU different from Cisco VSS and Huawei CSS?

A: VSU, VSS, and CSS implement switch cluster management respectively. Their functions are similar; yet each has its own characteristics.

Ruijie Networks Co.,Ltd

For further information,please visit our website **http://www.ruijienetworks.com**