MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

V.N. Karazin Kharkiv National University

School of Mathematics and Computer Science

Department of Theoretical and Applied Informatics

Master's Thesis

Exploring Fault-Tolerant Network Protocols in Distributed Systems

Author:

Final year Master's Program student,

specialty - Computer Sciences and Information Technologies,
educational program: "Informatics"

Song Chong

Supervisor: Kyrylo Rukkas

Reviewer: Kyryl Korobchynskyi

Adviser: Oleksandr Barskyi

Kharkiv, 2024

ABSTRACT

Distributed systems typically require their networks to possess high reliability. High network reliability is generally reflected on three levels: systems, components, and links. A common method for achieving high network reliability is through redundancy mechanisms. Designers deploy redundant networks between nodes so that multiple physical networks can act as backups for one another. When all links are functioning normally, they share the network load evenly, improving network utilization. In the event of a link failure, the system can continue to use other functioning links to ensure normal operation, effectively realizing network fault tolerance and load balancing.

This study primarily explores the principles of network communication reliability, the design of redundant network reliability, and their engineering applications, with a particular focus on implementing reliable network communication backups and load balancing at the foundational level. The fault tolerance and load balancing techniques implemented differ from conventional methods. Without altering the existing configuration or operational methods of the physical redundant network hardware, a virtual network adapter is added alongside the network communication processing module developed for this project. This approach virtualizes the physical redundant networks into a single logical network. The goal is to design and implement a universal high-reliability redundant network communication solution that allows upper-layer applications to develop for a single virtual network, without worrying about underlying details such as redundancy and load balancing. This technique enables a local node to transparently communicate with other nodes in the redundant network with high reliability through the virtual network adapter. Based on current popular fault tolerance technologies and project requirements, this study focuses on three key fault tolerance technologies: link backup, equipment backup, and stacking. While enhancing network redundancy, the study also optimizes the network architecture to achieve true high availability and fault tolerance.

The document is divided into six chapters. Chapter One primarily discusses the background, research content, objectives, and significance of the project. Chapter Two provides a brief introduction to the basic concepts of fault tolerance technologies and fault handling. Chapter Three covers link backup technologies explored in the research, the involved technical implementations, and corresponding experimental verifications through a constructed network topology. Chapter Four focuses on equipment backup technologies, detailing their implementations and validation through network topology testing. Chapter Five introduces stacking technologies, along with their implementation details and experimental verifications. Chapter Six concludes the study and proposes future work to enhance the project further.

Keywords: Redundant network, fault tolerance, load balancing, virtual network adapter

PREFACE

1.1 Background of the Topic

In today's era of rapid advancements in information technology, the internet has seen swift development and widespread application. Computer networks have not only become essential communication infrastructure for society but also indispensable foundational components of various automated systems. Unlike conventional networks, distributed systems have made significant progress and have become the foundational architecture for many critical applications, such as industrial monitoring systems and commercial service systems. The application prospects of distributed systems in fields such as industry and military have attracted the attention of researchers both domestically and internationally. With economic development, the application of distributed systems in certain critical fields continues to expand, primarily in civilian and military domains.

In the civilian domain, distributed systems are often used for speed monitoring on automated highways, marine environment exploration, and information communication for controlling transportation systems, among other applications. For example, the automated power dispatching system must operate uninterruptedly 24/7; any operational disruption could lead to severe and incalculable consequences. Similarly, in the rapidly growing field of e-commerce, interruptions in service or the loss of critical data could result in significant financial losses.

In the military domain, distributed systems are frequently applied in areas such as reconnaissance by unmanned spacecraft, multi-target search by fighter jets, and the synchronization control of multiple satellite groups. As humans continue to expand their exploration of natural domains, the scale of engineering control systems and large-scale automated equipment has gradually increased, and the complexity of distributed systems has also grown. At the same time, the level of artificial intelligence in these systems is steadily improving. In high-tech fields such as aerospace, deep space exploration, nuclear power devices, and robotics, the analysis of distributed systems is often highly intricate.

These complex systems often exhibit characteristics such as randomness, suddenness, uncertainty, and diversity. Due to the increasingly intricate interconnections among subsystems, even minor faults within the system can lead to its complete failure, potentially resulting in catastrophic consequences. On one hand, this could threaten human life and safety, causing casualties, while on the other hand, it could lead to severe economic losses for society. This is particularly true in certain specialized industries, such as nuclear energy, satellite launches, and maritime navigation, where there are exceptionally high demands for system safety and reliability. The losses caused by failures in these specialized industries are often immeasurable. Therefore, ensuring the safe and efficient operation of these critical and even high-risk systems and improving their reliability is of great significance.

Fault-tolerance technology in distributed systems can enhance system safety. It ensures that when a component within a distributed system encounters issues or failures, the system can automatically recover and continue functioning without affecting the correct and stable execution of applications within the distributed system. This has become an urgent issue that needs to be addressed. To this end, it is essential to design systems

with high reliability. As the foundational components of computer network systems, distributed systems require highly reliable networks as one of the key bases for ensuring overall system reliability.

1.2 Research Content and Objectives

This study focuses on network reliability issues. By utilizing three key technologies—link backup, device backup, and stacking techniques—it aims to design a cross-platform universal module, analyze the reliability of links and devices, and propose a universal, high-reliability redundant network communication solution. High-reliability design is an integrated concept. Generally, when designing a highly reliable computer network system, the main areas of focus include:

- Designing a highly reliable network architecture
- Providing device-level and link-level redundancy and backup for critical links
- Reducing unplanned downtime
- Ensuring reliable network management
- Promptly detecting and locating faults

High network reliability refers primarily to the ability of the network to maintain uninterrupted service in the event of equipment or network failures. A network consists of numerous components, including routers, switches, servers, PCs, cables, wireless signals, software, and protocols. Ensuring network reliability necessitates ensuring the reliability of all these components. Since any network component can fail, achieving network reliability often requires a certain level of redundancy. The most critical components in a communication network are network devices and links. Therefore, reliability is primarily achieved through device redundancy and link redundancy. Providing one or more backup devices or links ensures a quick switchover to a standby device or link in the event of a failure in the primary device or link.

This study primarily explores the application of link backup, device backup, and stacking technologies in network design. Specifically, it emphasizes the inherent reliability of devices, enhances the reliability of individual links, and selects H3C network devices based on the Comware system. Utilizing communication protocols and software with strong redundancy path management capabilities—including Link Aggregation, RRPP, Smart Link, VRRP, and IRF—this research aims to improve network redundancy while optimizing the network architecture, ultimately achieving true high reliability in the network system.

1.3 Significance of the Research Topic

Distributed systems typically demand high reliability, and redundancy technology is one of the primary means to achieve this goal. Currently, there are various solutions for network redundancy, each with its advantages and disadvantages. However, there has yet to be seen a universal, cross-platform technological solution suitable for distributed systems that meets the following requirements:

1) Compatible with various mainstream operating system platforms.

2) Does not require support from specific hardware.

3) Completely transparent to upper-layer applications, with no need for dedicated interfaces.

4) Capable of load balancing during normal network operation and reliable failover during network failures.

5) Avoids introducing single points of failure on critical devices.

Based on these requirements, this research aims to develop a universal network redundancy solution applicable to distributed systems. This holds broad significance for improving the availability of distributed systems.

Chapter 2: Understanding Fault Tolerance Technology

Fault tolerance technology refers to the ability of a computer system to tolerate faults by utilizing redundant resources to complete processing and computation without compromising system performance, even when hardware or software failures occur. It is an important research area in the field of computer science and technology. In this chapter, we will first examine the basic concepts related to fault handling, followed by a discussion of fault models in computers.

2.1 Basic Concepts

2.1.1 The Concept of Fault Tolerance Characteristics

Fault tolerance plays a crucial role in distributed systems and is closely tied to the reliability of networked systems. Reliability in distributed systems encompasses the following characteristics:

Availability: Availability refers to the ability of a system to be immediately usable. It generally means that a system can correctly execute user commands at any given moment. In other words, a highly available system can respond to tasks promptly.

Reliability: Reliability is the ability of a system to perform its intended functions without failure within a specified time frame. It reflects the continuity of service provided by the network. The longer a system operates without failure, the higher its reliability. Compared to availability, reliability focuses on a system's performance over a period of time, whereas availability pertains to performance at a specific moment. For example, if a system experiences only 1 millisecond of downtime per hour, it achieves an availability of 99.9999%. However, it may not be considered highly reliable. Conversely, a system that operates continuously without crashing for two weeks but needs to shut down once can be deemed highly reliable, though its availability might only be 96%. Therefore, reliability and availability are distinct concepts.

Safety: Safety refers to the absence of catastrophic consequences when a system temporarily fails to function correctly. For instance, systems that control critical processes, such as nuclear power plants or spacecraft, require high safety standards. A temporary failure in such control systems could lead to disastrous outcomes. However, past experiments have shown that establishing a highly safe system is extremely challenging.

Maintainability: Maintainability measures how easily a failed system can be restored. A system with high maintainability often also has high availability, especially if it can automatically detect and repair faults.

2.2 Basic Fault Models

A failed system cannot fully deliver the services it was designed to provide. If we consider a distributed system as a cluster server communicating with clients, the inability to fully provide services means that the server, the communication channel, or both are not functioning as intended. The failure of a server may not necessarily be caused by issues within the server itself. If a server depends on other servers to deliver complete services, the failure could be due to issues in those other servers. Such interdependencies among modules in a distributed system are common. For instance, in the case of a highly available file server, if the hard drive fails, the file server may stop functioning properly. If this file server is a part of a distributed database, the entire database may also malfunction, potentially leaving only parts of the data accessible. Basic fault models are primarily categorized into crash faults, omission faults, timing faults, response faults, and arbitrary faults.

1.Crash Faults: These occur when a server crashes but was functioning properly before the crash. Typically, this happens when a server shuts down earlier than expected. A key characteristic of crash faults is that once a server fails, the client sending requests will no longer receive any messages from the server. If a machine experiences a crash fault, it cannot perform any tasks, and the only solution is to restart it.

2.Omission Faults: These occur when a server is running but fails to respond to a request. Omission faults can be divided into three scenarios:

（1）. Receive Omission Faults**: For example, while the connection between the client and server is intact, no thread on the server side is listening for the client's requests, resulting in no response to the client. In most cases, since the server is unaware of the client's request messages, receive omission faults do not affect the current state of the server.

（2）Send Omission Faults**: These occur when the server has processed a request message but fails to send the response message. For instance, this might happen if the send buffer overflows and the server does nothing to handle it. Compared to receive omission faults, the server is in a state where it has fully processed the client's request. However, due to the failure to send the response message, the server might reprocess the client's previous request.

（3）Non-Communication-Related Omission Faults**: These are associated with software errors and do not involve communication. For example, the server might enter an infinite loop or improperly manage memory, causing the program to remain "stuck" for a prolonged period.

3.Timing Faults: These are faults related to timing constraints. In synchronous systems, there are time limitations on clock drift rates, execution times, and message transmission times. Timing faults can occur when any of these constraints are violated, causing the server to fail to respond to the client within the specified time frame.

Failure Type	Impact	Description
Clock	Process	The local clock of the process exceeds the allowable range of deviation from the actual time.
Performance	Process	The process exceeds the allowable range for the interval between two processes.
Performance	Channel	Message transmission takes longer than the specified range.

*Time Sequence Fault Table

4.Response Failures: The server's response is incorrect. Generally, there are two scenarios: ① The server provides an erroneous response to a request. For example, a search engine automatically returns a webpage completely unrelated to the search keywords. ② State Transition Errors: When the server's reaction to a request does not meet expectations, a state transition error occurs. For instance, if a server receives an unrecognizable request and has no mechanism to handle such information, this kind of state transition error may arise.

5.Arbitrary Failures: These are the most difficult types of failures to resolve, often referred to as Byzantine failures. In practice, when an arbitrary failure occurs, the client must prepare for the worst-case scenario. The outputs generated by the server might be entirely unintended. Moreover, the errors in these outputs cannot be easily detected. An even more serious issue arises if the malfunctioning server is working collaboratively with other servers. The faulty server can influence the other servers to make incorrect decisions as well.

2.3 Fault Masking and Fault Handling

2.3.1 Fault Masking

Redundancy is a key technique for concealing faults. The main types of redundancy are divided into three categories: informational redundancy, physical redundancy, and temporal redundancy.

Informational Redundancy: This involves using extra bits to recover corrupted fragments. For example, the use of Hamming code technology primarily corrects errors that occur during data transmission.

Physical Redundancy: Adding extra devices allows a system to tolerate partial component failures or malfunctions. Physical redundancy can be either hardware-based or software-based. Examples of hardware redundancy include additional processors or I/O devices, while software redundancy involves additional versions of software modules.

Temporal Redundancy: This involves performing an operation once and, if necessary, repeating it. For instance, atomic transaction processing and atomic operations. If a fault occurs during atomic transaction processing or atomic operations, they will roll back to their pre-execution state, essentially as if they were never executed. This allows them to be re-executed, requiring only additional time. Temporal redundancy is particularly useful for temporary or intermittent faults.

2.3.2 Fault Handling

There are three main approaches to fault handling: active replication, passive replication, and semi-active replication.

Active Replication:

In the fault-tolerant active replication model, backup managers function as state machines. Each backup manager plays an equal role, forming a group. Clients send their requests simultaneously to the entire backup manager group, and then all backup managers independently process these requests using the same method. A failure in any individual backup manager does not disrupt the overall system performance because the remaining backup managers can still provide proper services. In this model, all replication modules work together and maintain tightly synchronized states.

Passive Replication:

In the fault-tolerant passive replication model, backup managers do not play equal roles; instead, they are divided into primary and secondary roles. There is one primary backup manager and one or more secondary backup managers. The essence of this model is that clients communicate with the primary backup manager to receive the requested services. The primary backup manager performs the corresponding operations and forwards updated copies to the secondary backup manager group. If the primary backup manager fails, one of the secondary backup managers is promoted to act as the new primary backup manager. In this model, only one module is active at any given time, and the state of the other modules is periodically updated via checkpoints from this active module.

Semi-Active Replication:

This is a hybrid approach combining aspects of both active and passive replication. Its main advantage is lower recovery overhead compared to either active or passive replication.

Among the three methods discussed above, active replication incorporates the concept of fault masking and primarily relies on physical redundancy. In contrast, passive replication detects errors primarily through mechanisms such as periodic checks, monitoring clocks, and self-testing loops.

2.4 Basic Knowledge of Dual-Machine Fault Tolerance

Dual-machine fault tolerance involves connecting two servers in a network using specific hardware and software technologies to enable mutual backup of data and applications. Since the two servers can monitor each other's operational status, if one server fails, the other server immediately takes over its tasks and continues to provide the same level of service, ensuring the system's continuous operation, availability, and reliability.

As described above, dual-machine fault tolerance primarily achieves two functions: monitoring and switching. Its design principle is based on the servers monitoring each other's applications or CPUs through a software system, while continuously exchanging signals. If one server encounters a failure and the other server does not receive the expected signal, it detects the issue and automatically activates the software's switching function. This function transfers the workload of the failed server to the designated server,

which restarts the services previously provided by the failed server, ensuring uninterrupted service delivery. The main difference between dual-machine fault tolerance and cluster fault tolerance lies in the number of servers involved: dual-machine fault tolerance specifically refers to two servers, while clusters involve multiple servers.

Dual-machine fault tolerance operates in two modes: hot standby and mutual backup.

Hot Standby: In this mode, one server is designated as the primary working server, while the other serves as the backup. Under normal conditions, the primary server is responsible for providing external services, and the backup server monitors the operational status of the primary server. Similarly, the primary server also monitors the backup server's functionality, as the backup server can also experience failures for various reasons. If the backup server encounters an issue, the primary server should promptly notify the administrator to resolve it, ensuring the backup server can function normally during the next switch. When the primary server fails and can no longer provide services, the backup server takes over and continues offering the same level of service. This guarantees uninterrupted service delivery to external users, achieving the goal of non-stop server operation. Once the primary server is repaired, it automatically resumes its original state of operation.

Mutual Backup: In this mode, both servers are designed as working servers, and under normal conditions, both provide services to external users while monitoring each other's operational status. If one server fails and cannot provide services, the other server continues its original tasks while taking over the workload of the failed server. This ensures uninterrupted service delivery to external users, achieving non-stop server operation. However, the workload on the functioning server inevitably increases, so the failed server must be repaired as soon as possible to reduce the duration of the increased load on the operational server. Once the failed server is repaired, it automatically resumes its original state of operation.

Each dual-machine fault tolerance mode has its advantages and disadvantages. The mutual backup mode has higher operational costs and is more complex to manage, while the main drawback of the hot standby mode is the longer switching time between the primary and backup servers, which typically takes at least five minutes and requires an additional restart.

The two servers are connected via a network, with core data mirrored between them. During normal operation, the primary server holds control, and data is mirrored in real-time to the backup server. Using Ethernet as the system's data link ensures that the primary server does not interfere with the backup server's work, allowing error detection and maintenance to occur in an independent environment. After the primary server recovers, control must be switched back to it, and data must be synchronized from the backup server to the primary server. This synchronization process is handled automatically in the background. Once data synchronization is complete, the dual-machine system resumes normal operation.

2.5 Summary of This Chapter

This chapter provides a brief explanation of the foundational concepts of fault tolerance. It introduces four main aspects: the basic concepts related to fault tolerance, fundamental

fault models, basic methods for fault masking and handling, and dual-system fault tolerance.

Chapter 3: Overview of Link Backup Technologies

3.1 Overview of Link Backup Technologies

Link backup technologies in distributed network systems primarily use three methods: link aggregation, RRPP, and Smart Link. Distributed aggregation technology further eliminates the issue of single-point failure in aggregation devices, enhancing the availability of aggregated links. Since aggregation members can reside on different devices within the system, even if an entire device hosting certain members fails, the aggregated link will not completely fail. Other functioning units will continue to manage and maintain the status of the remaining aggregated ports. This is particularly significant for core switching systems and network environments requiring high-quality services.

Metropolitan Area Networks (MANs) and enterprise networks often adopt ring networks to ensure high reliability. The technologies used in ring networks are generally RPR or Ethernet rings. RPR requires specialized hardware, making it relatively costly. In contrast, Ethernet ring technology has become increasingly mature and cost-effective, leading to a growing trend of using Ethernet rings in MANs and enterprise networks. Currently, the technologies available to address Layer 2 network loop issues include STP and RRPP. STP is relatively mature but has a convergence time measured in seconds. RRPP, on the other hand, is a link-layer protocol specifically designed for Ethernet rings. It offers faster convergence than STP, and its convergence time is independent of the number of nodes in the ring, making it suitable for networks with larger diameters.

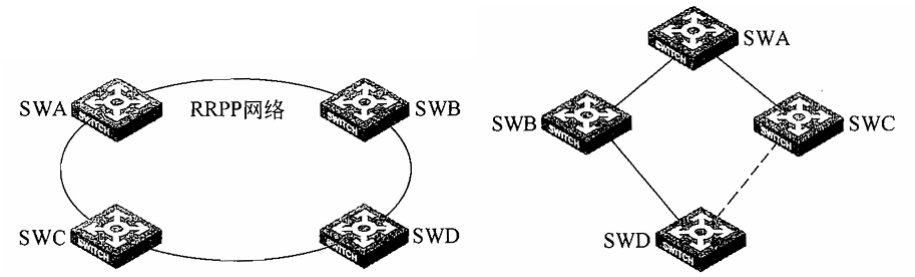
Smart Link is a solution designed for dual-uplink network configurations, enabling efficient and reliable link redundancy backup with rapid convergence in the event of a failure.

As shown in the diagram above, link aggregation combines multiple physical Ethernet ports into a single logical aggregation group. Through link aggregation, the upper-layer entity treats multiple physical links within the same aggregation group as a single logical link. Link aggregation allows data traffic to be distributed among the member ports of the aggregation group, increasing bandwidth. At the same time, member ports within the same aggregation group dynamically back each other up, enhancing connection reliability.

RRPP (Rapid Ring Protection Protocol) is a link-layer protocol specifically designed for Ethernet rings, as shown in the diagram below.When the Ethernet ring is intact, RRPP prevents broadcast storms caused by data loops. In the event that a link within the Ethernet ring disconnects, RRPP quickly restores communication paths between all nodes on the ring, offering fast convergence speeds.

To meet users' requirements for rapid link convergence while simplifying configuration, H3C has proposed the Smart Link solution for dual-uplink networking, as shown in the diagram. Smart Link provides redundant backup for primary and secondary links, ensuring that in the event of a failure on the primary link, traffic can quickly switch to the backup link. This ensures high convergence speeds. Smart Link technology is specifically

designed for dual-uplink networking, achieving convergence performance at the millisecond level. It is simple to configure and user-friendly.



3.1.1 Introduction to Link Aggregation
3.1.1.1 Background of Link Aggregation

As shown in the diagram below, link aggregation combines multiple physical Ethernet ports into a single logical aggregation group. The upper-layer entities using the link aggregation service treat multiple physical links within the same aggregation group as a single logical link.

Link aggregation enables the distribution of data traffic across multiple member ports within an aggregation group, thereby increasing bandwidth. At the same time, the member ports of the same aggregation group dynamically back each other up to enhance connection reliability.

As shown in the illustration below, the official standard for link aggregation technology is IEEE Standard 802.3ad, established by IEEE. Within the structure of IEEE 802.3, link aggregation is located between the MAC CLIENT and the MAC layers, functioning as an optional sublayer. The standard defines the objectives of link aggregation technology, the functions and operations of each module within the aggregation sublayer, as well as the principles of link aggregation control.

The stated objectives that the aggregation technology must achieve include enhancing link availability, linearly increasing bandwidth, balancing loads, enabling automatic configuration, ensuring rapid convergence, maintaining transmission quality, guaranteeing transparency to upper-layer users, and being backward compatible.

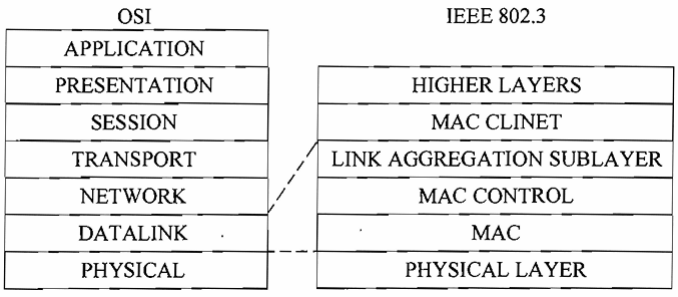


Figure: IEEE 802.3 Architecture

3.1.1.2 Concepts Related to Link Aggregation

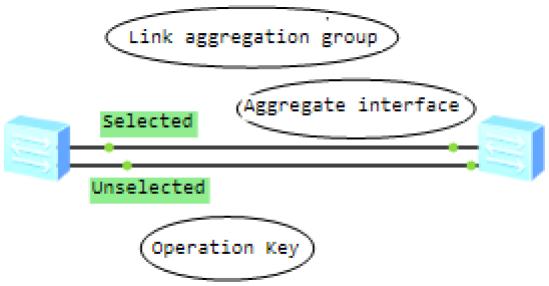


Diagram: Concepts Related to Link Aggregation

As shown in the figure above, the relevant concepts in link aggregation are as follows:

(1) Aggregated Interface: An aggregated interface is a logical interface that can be categorized into Layer 2 aggregated interfaces and Layer 3 aggregated interfaces.

(2) Aggregation Group: An aggregation group is a collection of Ethernet interfaces. The aggregation group is automatically created along with the aggregated interface, and its number corresponds to the aggregated interface's number. Based on the types of Ethernet interfaces that can be added to the aggregation group, it can be classified as either a Layer 2 aggregation group or a Layer 3 aggregation group.

(3) Status of Aggregated Member Ports: The member ports in an aggregation group have two states. The Selected state indicates that the port can participate in forwarding user data, while the Unselected state indicates that the port cannot forward user data. The speed and duplex mode of the aggregated port are determined by its Selected member ports. The speed of the aggregated port is the sum of the speeds of the Selected member ports, and the duplex mode is consistent with that of the Selected member ports.

(4) Operation Key: The Operation Key is a configuration combination automatically generated during link aggregation by the aggregation control, based on certain configurations of member ports. It includes configurations such as port speed, duplex mode, and link status. Within an aggregation group, member ports in the Selected state share the same Operation Key.

(5) First-Class Configuration: This type of configuration can be applied to both aggregation ports and member ports, but it does not factor into the calculation of the Operation Key. Examples include GVRP and MSTP.

(6) Second-Class Configuration: The contents of this type of configuration are shown in the table below. In the same aggregation group, if the Second-Class Configuration of a member port differs from that of the aggregation port, the member port cannot become a Selected port.

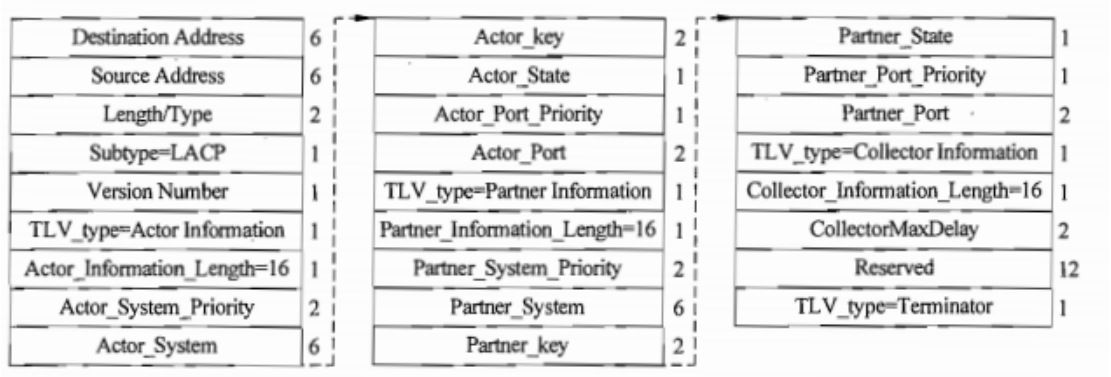
Category	Configuration Content
Port Isolation	Is the port part of the isolation group
QinQ Configuration	Port QinQ functionality enable/disable status, VLAN Tag TPID value, added outer VLAN Tag, mapping between inner and outer VLAN priorities, strategies for adding outer VLAN Tags to different inner VLAN IDs, and inner VLAN ID replacement relationships.
VLAN Configuration	VLANs allowed on the port, the default VLAN ID of the port, the link type of the port (Trunk, Hybrid, or Access), VLAN configuration based on IP subnets, VLAN configuration based on protocols, and whether VLAN packets are configured to carry Tags.
MAC Address Learning Configuration	Does it support MAC address learning? Is there a limit on the maximum number of MAC addresses that can be learned per port? If the MAC address table is full, will it continue to forward traffic?

3.1.1.3 LACP Protocol

The LACP (Link Aggregation Control Protocol) is based on the IEEE802.3ad standard and is a protocol designed to enable the dynamic aggregation and de-aggregation of links. LACP uses LACPDU (Link Aggregation Control Protocol Data Unit) to exchange information with the peer device.

In the LACP protocol, the two ends of a link are referred to as the Actor and the Partner. Both sides exchange LACPDU packets to inform the other party of their system priority, system MAC address, port priority, port number, and operational Key. The peer device, upon receiving this information, compares it with data stored for other ports and selects the ports that can be aggregated. Both parties agree on the addition or removal of ports from a specific dynamic aggregation group, thereby determining which links can be included in the same aggregation group and when a particular link may join the group.

As shown in the figure below, the fields are explained as follows:



(1) Destination Address: The destination address is a multicast address.

(2) Source Address: MAC address of the sending port.

(3) Length/Type: 0x8099.

(4) Subtype = LACP: LACP protocol.

(5) Version Number: LACP version number (0x01).

(6) TLV_type = Actor Information: Information about the Actor side.

(7) Actor_Information_Length = 16: Length of Actor side information.

(8) Actor_System_Priority: System priority of the Actor side.

(9) Actor_System: System information of the Actor side.

(10) Actor_key: Key for the Actor side.

(11) Actor_State: State of the Actor side.

(12) Actor_Port_Priority: Port priority of the Actor side.

(13) Actor_Port: Port information of the Actor side.

(14) TLV_type = Partner Information: Information about the Partner side.

(15) Partner_Information_Length = 16: Length of Partner side information.

(16) Partner_System_Priority: System priority of the Partner side.

(17) Partner_System: System information of the Partner side.

(18) Partner_key: Key for the Partner side.

(19) Partner_State: State of the Partner side.

(20) Partner_Port_Priority: Port priority of the Partner side.

(21) Partner_Port: Port information of the Partner side.

(22) TLV_type = Collector Information: Information about the Collector.

(23) Collector_Information_Length = 16: Length of Collector information.

(24) CollectorMaxDelay: Maximum delay time for the Collector.

(25) Reserved: Reserved section.

(26) TLV_type = Terminator: Terminator information.

3.1.2 Link Aggregation Modes

Based on the method of aggregation, link aggregation can be divided into two modes:

(1) Static Aggregation Mode: In this mode, ports are prohibited from enabling LACP, and devices do not exchange information with peer devices. The selection of the reference port is based on information from the local device.

(2) Dynamic Aggregation Mode: In this mode, the LACP protocol of the ports is automatically enabled, allowing the exchange of LACP packets with peer devices. The selection of the reference port is determined based on the information exchanged between the local and peer devices.

The static aggregation process is illustrated below. In static aggregation mode, the LACP protocol for member ports is disabled. The system sets the selected state of member ports according to the following principles:

(1) When there are ports in the aggregation group in an UP state, the system selects the port with the highest priority in the following order: full-duplex/high-speed, full-duplex/low-speed, half-duplex/high-speed, half-duplex/low-speed. The selected port must be in the UP state and have the same secondary configuration as the corresponding aggregated port in the group. If multiple ports share the highest priority, the port with the smallest port number is chosen as the reference port.

(2) Ports that match the reference port in port attributes and secondary configuration and are in the UP state are eligible to become candidate ports in a Selected state. Other ports are placed in an Unselected state.

(3) The number of ports in the Selected state within an aggregation group is limited. If the number of candidate ports does not exceed the limit, all candidate ports are in a Selected state, while other ports remain in an Unselected state. When the number of candidate ports exceeds the limit, the system selects some candidate ports to stay in the Selected state based on ascending port numbers, while ports with larger port numbers transition to the Unselected state.

(4) If all members in the aggregation group are in the DOWN state, all group members will be in the Unselected state.

(5) Ports that cannot aggregate with the reference port due to hardware constraints will remain in the Unselected state.

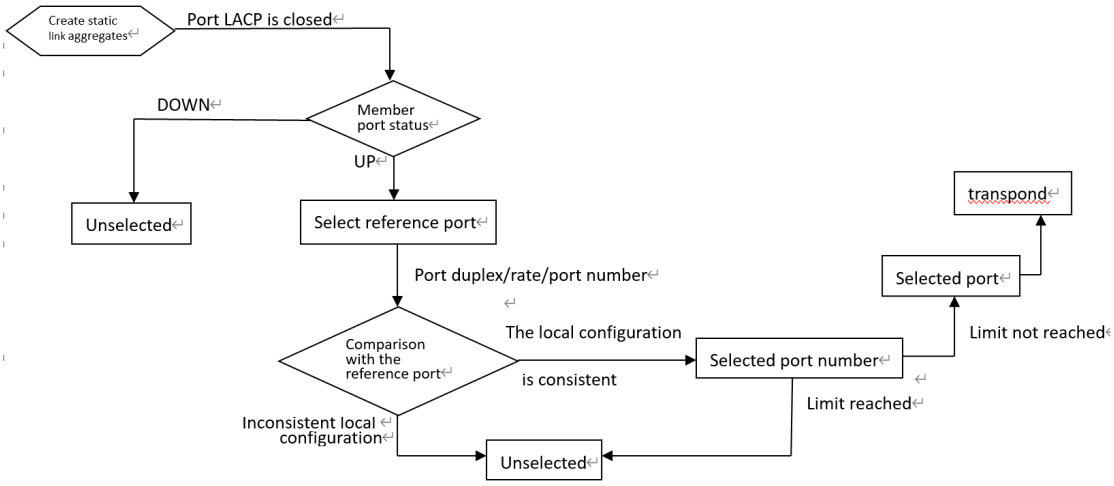


Diagram: Static Aggregation Process

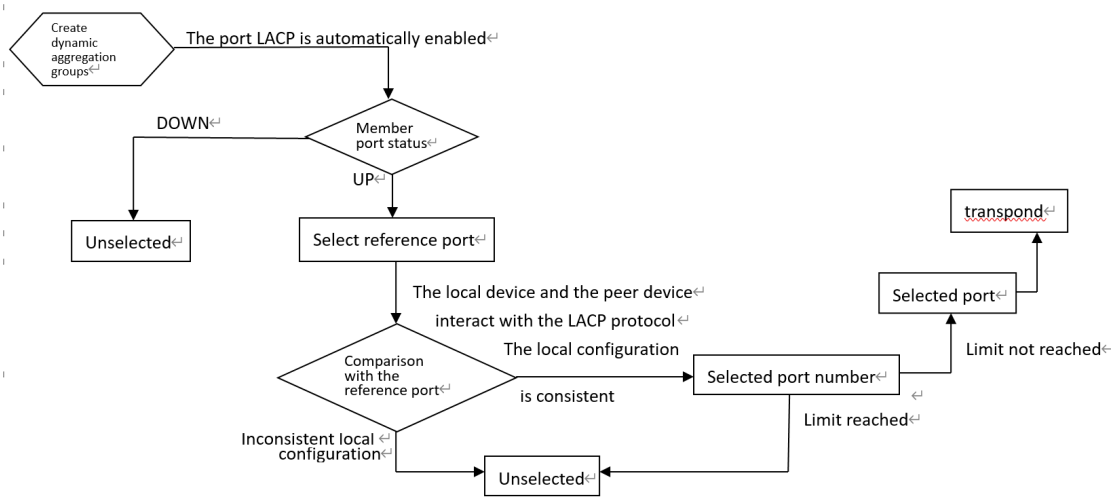


Figure: Dynamic Aggregation Process

The dynamic aggregation process is illustrated in the figure above. When an aggregation group is configured in dynamic aggregation mode, the LACP protocol for member ports in the group is automatically enabled. In dynamic aggregation mode, ports in the Selected state can send and receive LACP protocol packets. Unselected ports in an UP state, if their configurations match those of the corresponding aggregation ports, can also send and receive LACP protocol packets.

In dynamic aggregation mode, the local system and the peer system negotiate to determine the status of ports on both sides based on the port ID of the system with the better device ID. The detailed negotiation steps are as follows:

1. Compare the device IDs of the two systems (Device ID = LACP protocol priority of the system + MAC address of the system). First, compare the LACP protocol priority of the systems. If the priorities are the same, then compare the MAC addresses. The system with the smaller device ID is considered better (lower LACP protocol priority and MAC address indicate a smaller device ID).

2. Compare the port IDs of the better system (Port ID = LACP protocol priority of the port + port number). For each port of the better system, compare the LACP protocol priority first, and if the priorities are the same, compare the port numbers. The port with the smallest port ID is chosen as the reference port (lower LACP protocol priority and port number result in a smaller port ID).

3. A port is considered a candidate port for the Selected state only if it matches the configuration attributes of the reference port, matches the secondary configuration, is in the UP state, and its peer port's attributes also match those of the peer reference port. Otherwise, the port remains in the Unselected state.

4. The number of ports in the Selected state within an aggregation group is limited. If the number of candidate ports does not reach the limit, all candidate ports will be in the Selected state, and the remaining ports will be in the Unselected state. If the number of candidate ports exceeds the limit, the system selects ports to remain in the Selected state following the ascending order of port IDs, while ports with larger port IDs become Unselected. Additionally, the peer device detects this state change and adjusts the status of the corresponding ports accordingly.

5. Ports that cannot be aggregated with the reference port due to hardware limitations will remain in the Unselected state.

3.1.3 Configuration of Link Aggregation

3.1.3.1 Commands Related to Link Aggregation

The basic steps for configuring a static link aggregation group are as follows:

(1) In the system view, create a Layer 2 aggregation port and enter the Layer 2 aggregation port view. The configuration command is:

`interface bridge-aggregation interface-number`

(2) Exit to the system view and then enter the Ethernet port view. The configuration command is:

`interface interface-type interface-number`

(3) In the Ethernet port view, add the Ethernet port to the static aggregation group. The configuration command is:

`port link-aggregation group number`

When a static aggregation port is deleted, the system will automatically remove the corresponding aggregation group, and all member ports in the group will be automatically removed from the aggregation group.

For a static aggregation mode, users must ensure that the Selected status of the ports on both ends of the same link, which belong to two different devices, remains consistent. Otherwise, the aggregation functionality will not work correctly.

The basic steps for configuring a dynamic link aggregation group are as follows:

(1) In the system view, create a Layer 2 aggregation port and enter the Layer 2 aggregation port view. The configuration command is:

`interface bridge-aggregation interface-number`

(2) In the aggregation port view, configure the aggregation group to operate in dynamic aggregation mode. The configuration command is:

`link-aggregation mode dynamic`

(3) Exit to the system view and enter the Ethernet port view. The configuration command is:

`interface interface-type interface-number`

(4) In the Ethernet port view, add the Ethernet port to the dynamic aggregation group. The configuration command is:

`port link-aggregation group number`

When a dynamic aggregation port is deleted, the system will automatically remove the corresponding aggregation group, and all member ports in the group will be automatically removed from the aggregation group.

For dynamic aggregation mode, both ends of the system will automatically negotiate the Selected status of the ports in their respective aggregation groups for the same link. Users only need to ensure that ports aggregated together in one system are also

aggregated together at the remote end; the aggregation functionality will then work properly.

(5) In the system view, configure the system LACP protocol priority. The configuration command is:

`lacp system-priority system-priority`

(6) In the Ethernet port view, configure the LACP protocol priority of the port. The configuration command is:

`lacp port-priority port-priority`

(7) In the system view, configure the load-sharing mode for the aggregation group. The configuration command is:

`link-aggregation load-sharing mode { destination-ip | destination-mac | destination-port | ingress-port | source-ip | source-mac | source-port }`

Both the system's LACP protocol priority and the port's LACP protocol priority are set to 32768 by default. Changing the system's LACP protocol priority will affect the Selected and Unselected status of dynamic aggregation group members.

For load-sharing aggregation groups, the system uses a Hash algorithm to achieve load balancing. This algorithm can utilize different Hash keys for calculations (i.e., different load-sharing modes). Information found in packet headers, such as MPLS labels, IP addresses, MAC addresses, ingress port numbers, and their combinations, can all be used as Hash keys. By altering the load-sharing mode, traffic load balancing for the aggregation group can be flexibly achieved.

3.1.3.2 Link Aggregation Experiment

Figure: Static Link Aggregation Configuration Experiment

As shown in the figure above, SWA and SWB establish a static link aggregation. The switches are connected using trunk ports, and the default VLAN for the ports is VLAN1.

(1) Configure SWA.

```
[SWA]interface GigabitEthernet1/0/1
[SWA-GigabitEthernet1/0/1]port link-type trunk
[SWA-GigabitEthernèt1/0/1]port trunk permit vlan 1 10
[SWA]interface GigabitEthernet1/0/2
```

```
[SWA-GigabitEthernet1/0/2]port link-type trunk
[SWA-GigabitEthernetl/0/2]port trunk permit vlan 1 10
[SWA]interface GigabitEthernet1/0/3
[SWA-GigabitEthernet1/0/3]port link-type trunk
[SWA-GigabitEthernet1/0/3]port trunk permit vlan 1 10
[SWA]interface bridge-aggregation 1
[SWA-Bridge-Aggregationl]port link-type trunk
[SWA-Bridge-Aggregationl]port trunk permit vlan 1 10
[SWA]interface gigabitethernet 1/0/1
[SWA-GigabitEthernetl/0/1]port link-aggregation group 1
[SWA]interface gigabitethernet 1/0/2
[SWA-GigabitEthernet1/0/2]port link-aggregation group 1
[SWA] interface gigabitethernet 1/0/3
[SWA-GigabitEthernet1/0/3]port link-aggregation group 1
```

(2) Configure SWB.

```
[SWB]interface GigabitEthernet1/0/1
[SWB-GigabitEthernet1/0/1]port link-type trunk
[SWB-GigabitEthernet1/0/1]port trunk permit vlan 1 10
[SWB]interface GigabitEthernet1/0/2
[SWB-GigabitEthernet1/0/2]port link-type trunk
[SWB-GigabitEthernetl/0/2]port trunk permit vlan 1 10
[SWB]interface GigabitEthernet1/0/3
[SWB-GigabitEthernet1/0/3]port link-type trunk
[SWB-GigabitEthernet1/0/3]port trunk permit vlan 1 10
[SWB]interface bridge-aggregation 1
[SWB-Bridge-Aggregationl]port link-type trunk
[SWB-Bridge-Aggregation1]port trunk permit vlan 1 10
[SWB] interface gigabitethernet 1/0/1
```

```
[SWB-GigabitEthernet1/0/1] port link-aggregation group 1
[SWB] interface gigabitethernet 1/0/2
[SWB-GigabitEthernet1/0/2]port link-aggregation group 1
[SWB] interface gigabitethernet 1/0/3
[SWB-GigabitEthernet1/0/3]port link-aggregation group 1
```

(3) After the configuration is complete, establish a static link aggregation between SWA and SWB. Check the Layer 2 aggregation port entries on SWA.

```
[SWA]display interface Bridge-Aggregation 1
Bridge-Aggregation1 current state: UP
IP Packet Frame Type: PKTFMT_ETHNT_2, Hardware Address: 000f-e207-f2e0
Description: Bridge-Aggregation1 Interface
3Gbps-speed mode, full-duplex mode
Link speed type is autonegotiation, link duplex type is autonegotiation
PVID: 1
Port link-type: trunk
  VLAN passing   : 1(default vlan), 10
  VLAN permitted: 1(default vlan), 10
  Trunk port encapsulation: IEEE 802.1q
```

From the displayed information, it can be seen that the Layer 2 aggregate port is already UP, with a port speed of 3 Gbps. The detailed link aggregation information for the port is as follows:

```
[SWA]display link-aggregation verbose
Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing
Port Status: S -- Selected, U -- Unselected
Flags:   A -- LACP_Activity, B -- LACP_Timeout, C -- Aggregation,
         D -- Synchronization, E -- Collecting, F -- Distributing,
         G -- Defaulted, H -- Expired

Aggregation Interface: Bridge-Aggregation1
Aggregation Mode: Static
```

```
Loadsharing Type: Shar
  Port              Status     Oper-Key
--------------------------------------------------------------------------------
  GE1/0/1           S          3
  GE1/0/2           S          3
  GE1/0/3           S          3
```

Based on the displayed information, it can be seen that the aggregation group mode is set to Static, and ports GigabitEthernet1/0/1 through GigabitEthernet1/0/3 are designated as Selected ports.

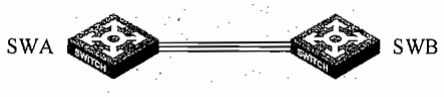


Figure: Dynamic Link Aggregation Experiment

As shown in the figure above, SWA and SWB establish dynamic link aggregation. The switches are connected using trunk ports, with the default VLAN for the ports being VLAN1.

(1) Configure SWA.

```
[SWA]interface GigabitEthernet1/0/1
[SWA-GigabitEthernet1/0/1]port link-type trunk
[SWA-GigabitEthernetl/0/1]port trunk permit vlan 1 10
[SWA]interface GigabitEthernetl/0/2
[SWA-GigabitEthernetl/0/2]port link-type trunk
[SWA-GigabitEthernetl/0/2]port trunk permit vlan 1 10
[SWA]interface GigabitEthernetl/0/3
[SWA-GigabitEthernet1/0/3]port link-type trunk
[SWA-GigabitEthernetl/0/3]port trunk permit vlan 1 10
[SWA]interface bridge-aggregation 1
[SWA-BridgeAggregationl]link-aggregation mode dynamic
[SWA-Bridge-Aggregationl]port link-type trunk
[SWA-Bridge-Aggregation1]port trunk permit vlan 1 10
[SWA]interface gigabitethernet 1/0/1
[SWA-GigabitEthernet1/0/1]port link-aggregation group 1
[SWA] interface gigabitethernet 1/0/2
[SWA-GigabitEthernetl/0/2]port link-aggregation group 1
[SWA] interface gigabitethernet 1/0/3
[SWA-GigabitEthernet1/0/3]port link-aggregation group 1
```

(2) Configure SWB.

```
[SWB]interface GigabitEtheret1/0/1
[SWB-GigabitEthernet1/0/1]port link-type trunk
[SWB-GigabitEthernet1/0/1]port trunk permit vlan 1 10
[SWB]interface GigabitEthernet1/0/2
```

```
[SWB-GigabitEthernet1/0/2]port link-type trunk

[SWB-GigabitEthernet1/0/2]port trunk permit vlan 1 10

[SWB]interface GigabitEthernet1/0/3

[SWB-GigabitEthernet1/0/3]port link-type trunk

[SWB-GigabitEthernet1/0/3]port trunk permit vlan 1 10

[SWB]interface bridge-aggregation 1

[SWB-Bridge Aggregationl] link-aggregation mode dynamic

[SWB-Bridge-Aggregationl]port link-type trunk

[SWB-Bridge-Aggregationl]port trunk permit vlan 1 10

[SWB] interface gigabitethernet 1/0/1

[SWB-GigabitEthernet1/0/1]port link-aggregation group 1

[SWB] interface gigabitethernet 1/0/2

[SWB-GigabitEthernet1/0/2]port link-aggregation group 1

[SWB] interface gigabitethernet 1/0/3

[SWB-GigabitEthernet1/0/3]port link-aggregation group 1
```

(3) After the configuration is complete, SWA and SWB establish a dynamic link aggregation. On SWA, check the Layer 2 aggregation port entries as follows:

```
[SWA]display interface Bridge-Aggregation 1
Bridge-Aggregation1 current state: UP
IP Packet Frame Type: PKTFMT_ETHNT_2, Hardware Address: 000f-e207-f2e0
Description: Bridge-Aggregation1 Interface
3Gbps-speed mode, full-duplex mode
Link speed type is autonegotiation, link duplex type is autonegotiation
PVID: 1
Port link-type: trunk
  VLAN passing   : 1(default vlan), 10
  VLAN permitted: 1(default vlan), 10
  Trunk port encapsulation: IEEE 802.1q
```

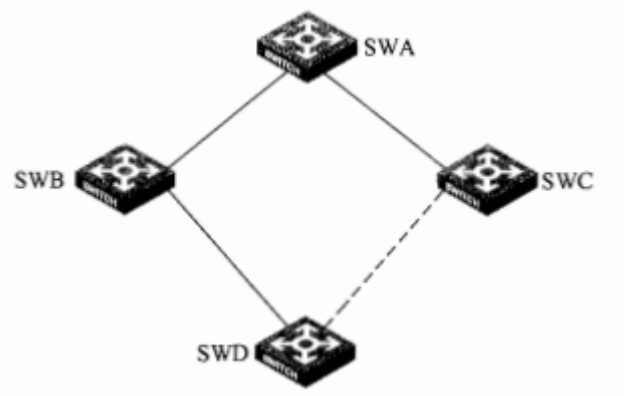
From the displayed information, it can be seen that the Layer 2 aggregated port is already UP, with a port speed of 3Gbps. The detailed link aggregation information for the port is as follows:

```
[SWA]display link-aggregation verbose

Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing
Port Status: S -- Selected, U -- Unselected
Flags:  A -- LACP_Activity, B -- LACP_Timeout, C -- Aggregation,
        D -- Synchronization, E -- Collecting, F -- Distributing,
        G -- Defaulted, H -- Expired

Aggregation Interface: Bridge-Aggregation1
Aggregation Mode: Dynamic
Loadsharing Type: Shar
System ID: 0x8000, 000f-e245-6bc0
Local:
  Port            Status  Priority  Oper-Key  Flag
--------------------------------------------------------------------------------
  GE1/0/1         S       32768     2         {ACDEF}
  GE1/0/2         S       32768     2         {ACDEF}
  GE1/0/3         S       32768     2         {ACDEF}
Remote:
  Actor           Partner Priority  Oper-Key  SystemID                Flag
--------------------------------------------------------------------------------
  GE1/0/1         25      32768     2         0x8000, 000f-e23d-59f0  {ACDEF}
  GE1/0/2         26      32768     2         0x8000, 000f-e23d-59f0  {ACDEF}
  GE1/0/3         27      32768     2         0x8000, 000f-e23d-59f0  {ACDEF}
```

Based on the displayed information, it can be seen that the aggregation group mode is set to Dynamic, and ports GigabitEthernet1/0/1 through GigabitEthernet1/0/3 have been designated as Selected ports.

3.1.2 Overview of Smart Link

3.1.2.1 Background of Smart Link

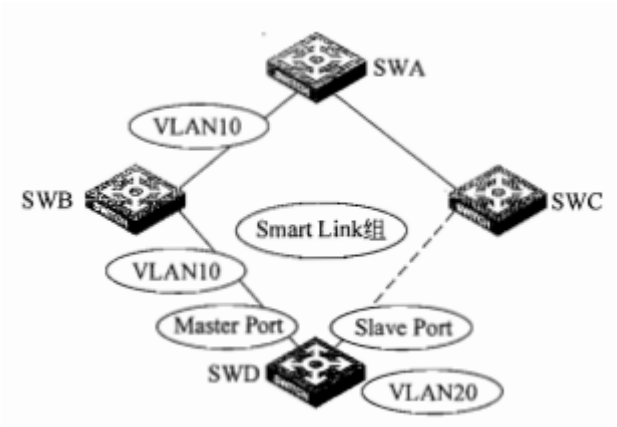


Smart Link Generation Diagram

Dual-uplink networking can enhance network reliability but introduces the issue of loops. Typically, loops can be eliminated using STP (Spanning Tree Protocol) or RRPP (Rapid Ring Protection Protocol). However, STP only achieves convergence speeds at the second-level, making it unsuitable for users with high demands for fast convergence times. While RRPP meets the requirements for convergence speed, it involves a higher level of configuration complexity and is mainly suited for more complex ring network setups. To meet user demands for rapid link convergence while simplifying configuration, H3C has proposed the Smart Link solution specifically for dual-uplink networking, as shown in the diagram above. Smart Link enables redundant backup between primary and secondary links, allowing traffic to quickly switch to the backup link in case of a failure on

the primary link. This ensures high convergence speed. Smart Link technology is designed exclusively for dual-uplink networking, achieving millisecond-level convergence performance with simple configuration, making it user-friendly and easy to operate.

3.1.2.2 Smart Link Related Concepts



As shown in the diagram above, the related concepts in Smart Link are as follows:

1. Smart Link Group: A Smart Link group, also referred to as a Flexible Link Group, contains only two ports: one designated as the primary port and the other as the secondary port. Under normal circumstances, only one port is in a forwarding state while the other is blocked and remains on standby.

2. Primary Port: The primary (Master) port is a role in a Smart Link group. When both ports in the Smart Link group are in the UP state, the primary port is prioritized to enter the forwarding state. However, the primary port does not always maintain the forwarding state; when its link fails, the secondary port, which is in standby, switches to the forwarding state. Without preemption configured, the primary port remains in standby even if its link recovers, until the next link switchover occurs.

3. Secondary Port: The secondary (Slave) port is the other role in a Smart Link group. When both ports in the Smart Link group are in the UP state, the secondary port remains in standby. However, it transitions to the forwarding state when the primary port encounters a link failure.

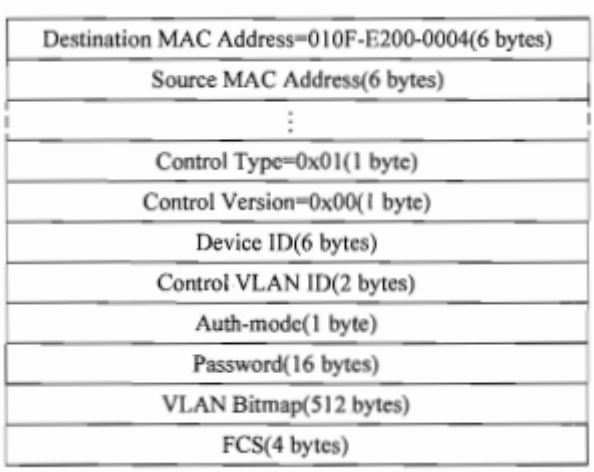
4. Flush Packet: When a Smart Link group undergoes a link switchover, the existing forwarding table entries become invalid in the new network topology, requiring all devices in the network to update their MAC address forwarding table entries and ARP/ND table entries. The Smart Link group sends Flush packets to notify other devices to refresh these MAC address forwarding table entries and ARP/ND table entries.

5. Transmit Control VLAN: A transmit control VLAN is a VLAN used to send Flush packets. When a link switchover occurs, the device broadcasts Flush packets within the transmit control VLAN.

6. Receive Control VLAN: A receive control VLAN is used to receive and process Flush packets. During a link switchover, the device receives and processes Flush packets belonging to the receive control VLAN to refresh MAC address entries and ARP/ND table entries.

7. Protected VLAN: A protected VLAN is a user data VLAN controlled by the Smart Link group to determine its forwarding state. Different Smart Link groups on the same port protect different VLANs. A port's forwarding state on a protected VLAN is determined by the port's role and state within its respective Smart Link group.

The Flush packet uses IEEE 802.3 encapsulation, which includes fields such as Destination MAC, Source MAC, Control VLAN ID, and VLAN Bitmap. As shown in the figure below, the fields are explained as follows:



1. The Destination MAC refers to an unknown multicast address, which can be distinguished as a Flush packet by checking if the address is 0x010F-E200-0004.

2. The Source MAC represents the source MAC address of the Flush packet.

3. The Control Type indicates the control type. Currently, there is only one type, which deletes MAC address forwarding table entries and ARP table entries (type 0x01).

4. The Control Version represents the version number. The current version is 0x00, which allows for future extensions.

5. The Device ID represents the bridge MAC address of the device sending the Flush packet.

6. The Control VLAN ID indicates the ID of the VLAN used for sending the control message.

7. Auth-mode specifies the authentication mode, used together with the Password for future security enhancements.

8. The VLAN Bitmap represents a VLAN bitmap, which carries the list of VLANs for which the address table needs to be refreshed.

9. FCS (Frame Check Sequence) is used for verifying the validity of the packet.

When a link switchover occurs, the VLAN Bitmap field in the Flush packet is populated with all VLAN IDs associated with ports in the forwarding state within the group before the switchover. The Control VLAN ID field is populated with the control VLAN ID

configured for the Smart Link group. Once the Flush packet is constructed, it is broadcast on the control VLAN through the new link (the link in the forwarding state after the switchover).

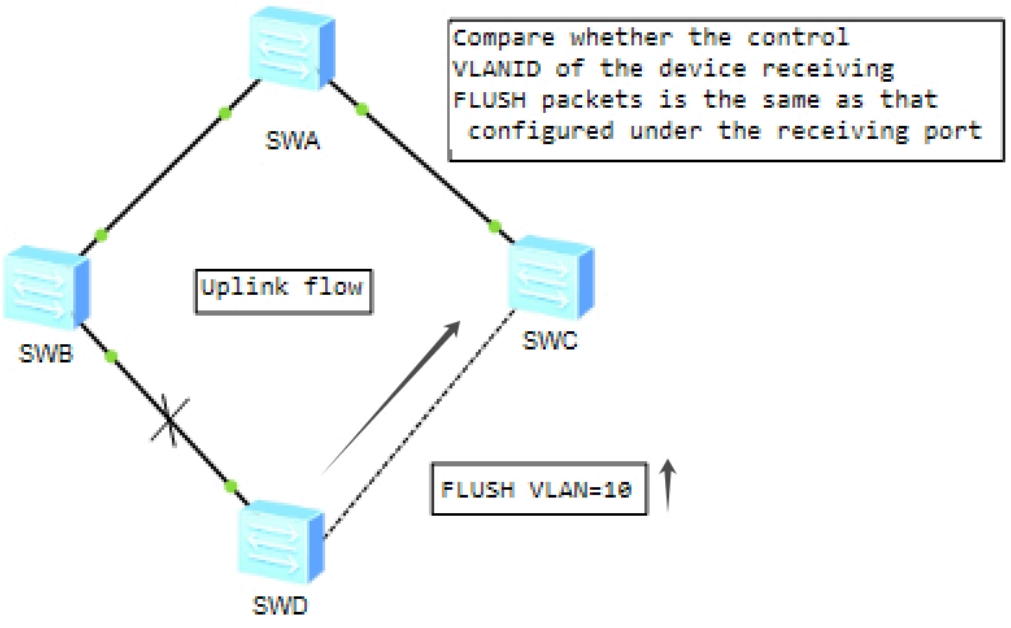
When a device receives a Flush packet, it checks whether the Control VLAN ID in the Flush packet matches the Control VLAN ID configured for the port that received the packet. If the two Control VLAN IDs do not match, the device will not process the Flush packet and will simply forward it. If the two Control VLAN IDs do match, the device extracts the VLAN Bitmap data from the Flush packet and deletes all MAC address forwarding table entries and ARP table entries learned within these VLANs.

To ensure that the Flush packet is correctly transmitted within the control VLAN, all ports on the link between the master and slave ports in the Smart Link group must belong to the control VLAN. If any port does not belong to the control VLAN, the Flush packet will fail to be sent or forwarded. Users send Flush packets with the VLAN tag retained. If the tag is to be removed, the default VLAN of the opposite port must match the control VLAN; otherwise, the Flush packet will not be transmitted within the control VLAN.

If upstream devices are not configured with the receiving control VLAN for handling Flush packets, or if the configured receiving control VLAN does not match the Control VLAN in the Flush packet, the device will not process the Flush packet and will simply forward it.

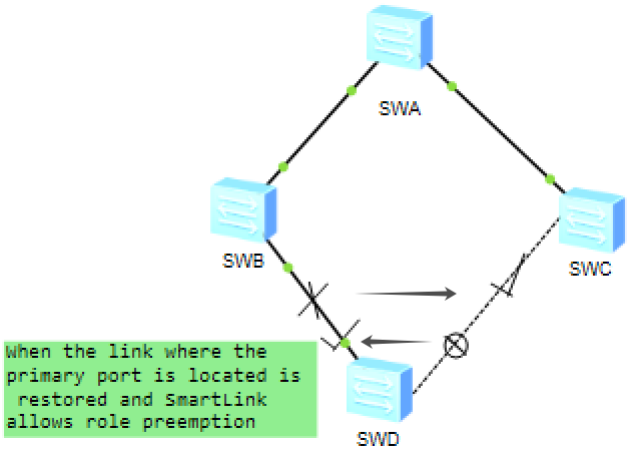
3.1.2.3 Smart Link Operating Mechanism

When the primary link fails, triggering a link switchover, the MAC address table entries and ARP table entries on network devices may become incorrect. To address this, a mechanism is required to update both MAC and ARP information. When devices that do not support the Smart Link feature interact with devices that do, the devices automatically refresh the MAC address table entries and ARP table entries through incoming traffic. This method of updating the MAC address table and ARP table relies on upstream traffic as the trigger. During the switchover process, traffic will be interrupted.



As shown in the figure above, when interfacing with a device that supports the Smart Link function, the Smart Link group sends Flush messages over the new link to refresh the MAC address forwarding table entries and ARP table entries. When the upstream device receives the Flush message, it deletes the MAC table entries and ARP table entries learned from VLANs within the VLAN Bitmap. If any ARP entries are deleted, the device will automatically trigger an update of the ARP entries. The entire link switching process is completed within milliseconds, with virtually no traffic loss.

The Smart Link link backup mechanism refers to the automatic process where, if a port in the forwarding state experiences a link failure, the Smart Link group will block that port and switch the previously blocked standby port to the forwarding state. When a port switches to the forwarding state, the system outputs log information to notify the user.

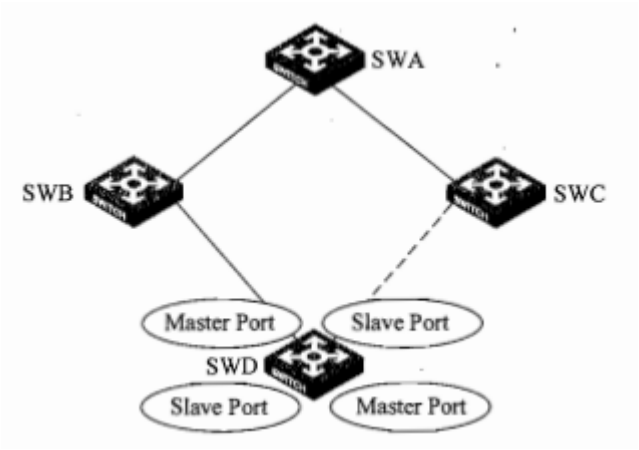


As shown in the figure above, the Smart Link role preemption mechanism refers to the preemption process that occurs after the primary port's link is restored. The link connected to the primary port serves as the active link, while the link connected to the secondary port acts as the standby link. When a failure occurs on the primary port's link, the primary port is automatically blocked and switched to standby mode, while the secondary port switches to forwarding mode. Once the primary port's link is restored, if the Smart Link group configuration allows role preemption, the secondary port will be automatically blocked and switched to standby mode, while the primary port resumes forwarding mode.

The protection VLAN of a Smart Link group is implemented by referencing the MSTP instance.

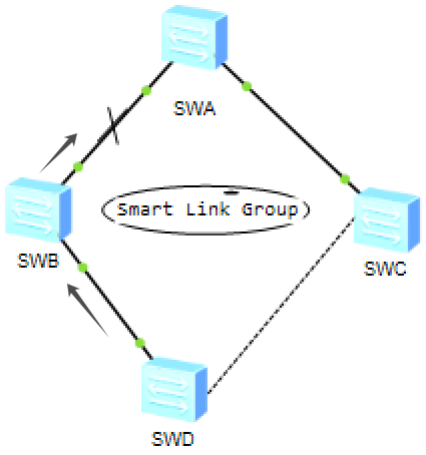
In the same ring network, multiple VLAN data flows may coexist. Smart Link can achieve traffic load balancing by forwarding the data traffic of different VLANs along the paths determined by different Smart Link groups.

By configuring a single port as a member port of multiple Smart Link groups (each Smart Link group having a different protection VLAN) and ensuring that the port's forwarding status differs across groups, it is possible to enable data traffic for different VLANs to be forwarded along different paths. This achieves the goal of load balancing, as illustrated below.



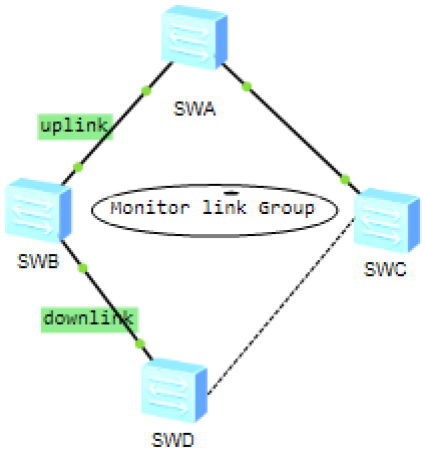
3.1.3 Overview of Monitor Link

3.1.3.1 Background of Monitor Link



As shown in the figure above, Monitor Link is a type of port linkage solution primarily used for network applications in conjunction with the Smart Link protocol to monitor the uplink on devices. It triggers changes in the downlink's UP/DOWN status based on the uplink's UP/DOWN status, thereby enabling the switchover of backup links controlled by the Smart Link protocol on downstream devices.

3.1.3.2 Monitor Link Related Concepts



As shown in the figure above, the related concepts in Monitor Link are as follows:

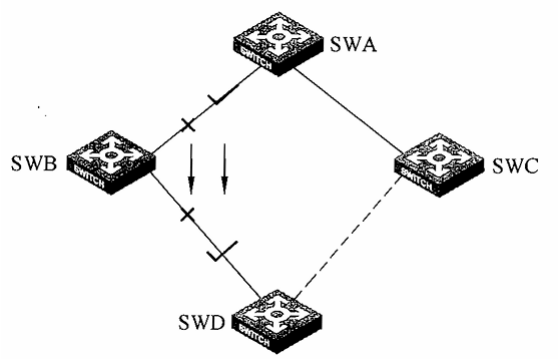
(1) Monitor Link Group: Also known as a monitoring link group, each group consists of an uplink and a downlink, with member roles determined by user configuration. Both the uplinks and downlinks can include multiple member ports, but each member can only belong to one Monitor Link Group. Member ports can be Layer 2 Ethernet ports or Layer 2 aggregation ports.

(2) Uplink: The uplink is the monitored link within the Monitor Link Group. When there are no uplink members in the Monitor Link Group, or all uplink member ports are in the DOWN state, the Monitor Link Group is considered to be in the DOWN state. If at least one uplink member in the Monitor Link Group is in the UP state, the Monitor Link Group is considered to be in the UP state.

(3) Downlink: The downlink is the passive link within the Monitor Link Group. When the UP/DOWN status of the Monitor Link Group changes, the Monitor Link adjusts the status of the downlink member ports accordingly to ensure alignment with the status of the Monitor Link Group.

3.1.3.3 Monitor Link Operating Mechanism

Each Monitor Link group independently handles uplink monitoring and downlink coordination, ensuring that the status of the downlink ports changes in response to the status of the uplink ports.



As shown in the figure above, when there are no uplink member ports in the Monitor Link group or all uplink member ports are in the DOWN state, the Monitor Link group itself will enter the DOWN state. This will force all its downlink member ports into the DOWN state as well. Once any uplink member port transitions from the DOWN state to the UP state, the Monitor Link group will recover to the UP state and will restore all its downlink member ports to the UP state.

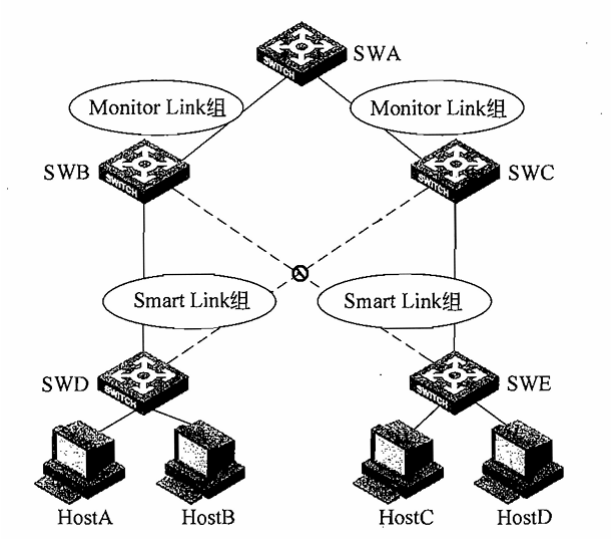
3.1.4 Specific Network Examples of Smart Link Technology & Monitor Link Technology

As shown in the figure below, in this network environment, Smart Link is configured on SWD and SWE. One of the dual-uplink links is blocked, while the other remains in the normal forwarding state. When a failure occurs on the forwarding link, the Smart Link group quickly detects it and switches the link.

As shown in the figure above, when there are no uplink member ports in the Monitor Link group or when all uplink member ports are in the DOWN state, the Monitor Link group will enter the DOWN state and will forcibly set its downlink member ports to the DOWN state as well. Once any uplink member port changes from the DOWN state to the UP state, the Monitor Link group will return to the UP state and restore all its downlink member ports to the UP state.

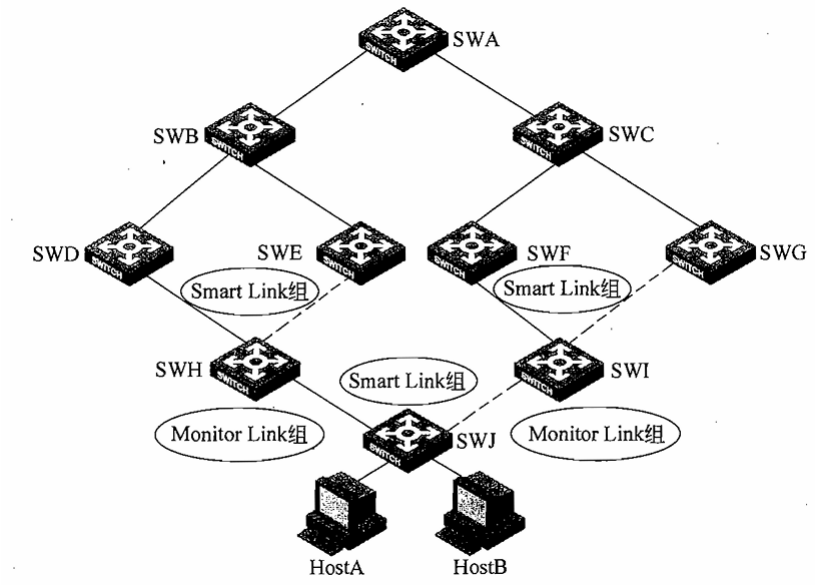
3.1.4 Smart Link Technology & Monitor Link Technology Specific Networking Experiment

As shown in the figure below, in this networking environment, Smart Link is configured on SWD and SWE. One of the dual uplinks is blocked, while the other remains in normal forwarding mode. When a failure occurs on the forwarding link, the Smart Link group quickly detects the issue and switches the link.



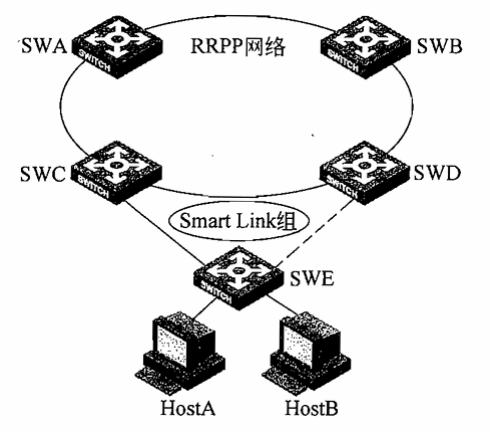
In this network environment, the link failure between SWA and SWB (or SWC) cannot be directly detected by SWD (or SWE). It is necessary to configure a Monitor Link group on SWB (or SWC). Once the Monitor Link group detects a failure in the uplink port or the link where the uplink port is located, it will forcibly shut down the downlink port, thereby triggering a link switchover within the Smart Link group on SWD and SWE. When the uplink port or link failure is resolved, the downlink port will automatically reopen, enabling SWD (or SWE) to quickly detect changes in the link status between SWA and SWB (or SWC).

As shown in the diagram below, it is a network application where the Smart Link is cascaded with the Monitor Link.



The Monitor Link group supports Smart Link as one of its uplink member ports. By combining Smart Link and Monitor Link technologies, cascading of backup links can be achieved. The implementation method is as follows: a Smart Link group functions as an

uplink port for a Monitor Link group, while the downstream ports of this Monitor Link group connect to the primary or secondary ports of another Smart Link group.



As shown in the figure above, this is an application of Smart Link in conjunction with RRPP for network deployment.

In this networking environment, RRPP is enabled on SWA, SWB, SWC, and SWD to provide link redundancy and backup. Since the two ports connected between SWC and SWD have already activated the RRPP function, STP cannot be enabled on them. Therefore, link backup on SWE can only be achieved by configuring a Smart Link group.

3.1.5 Smart Link & Monitor Link Configuration

3.1.5.1 Smart Link & Monitor Link Relevant Commands

To avoid creating loops that could lead to broadcast storms, ports must be manually shut down before being configured as member ports (main port or secondary port) of a Smart Link group. Once the Smart Link group configuration is complete, these ports can then be re-enabled. During this process, the administrator must disable both the STP and RRPP functions on the relevant ports and ensure that the ports are not part of any link aggregation group or service loopback group.

The `protected-vlan` command is used to configure the list of VLANs protected by the Smart Link group via an indirect reference to the MSTP instance.

The `flush enable` command requires each Smart Link group to be configured with a different control VLAN.

Member ports of a Smart Link group can be configured in either the Smart Link group view or the port view, with the configuration outcomes being identical in both cases. The basic steps to configure a Smart Link group are as follows:

(1) In the system view, create a Smart Link group and enter the Smart Link group view. The configuration command is:
smart-link group group-id

(2) Configure the protection VLAN of the Smart Link group in the Smart Link group view. The configuration command is:
protected-vlan reference-instance instance-id-list

(3) Enable the function of sending Flush messages in the Smart Link group view. The configuration command is:

flush enable [control-vlanvlan-id]

(4) Configure the member ports of the Smart Link group in the Smart Link group view. The configuration command is:

port interface-type interface-number { master | slave }

Alternatively, configure the member ports in the port view. The configuration command is:

port smart-link group group-id { master | slave }

(5) Configure the preemption mode to role preemption mode in the Smart Link group view. The configuration command is:

preemption mode role

During the configuration process, to avoid unnecessary DOWN/UP changes in the downlink, a port can belong to only one Monitor Link group. The uplink members of the Monitor Link group must be configured first.

The basic steps for configuring a Monitor Link group are as follows:

(1) Create a Monitor Link group in the system view and enter the Monitor Link group view. The configuration command is:

monitor-link group group-id

(2) Configure the uplink members in the Monitor Link group view. The configuration command is:

port interface-type interface-number uplink

Alternatively, configure the uplink members in the port view. The configuration command is:

port monitor-link group group-id uplink

(3) Configure the downlink members in the Monitor Link group view. The configuration command is:

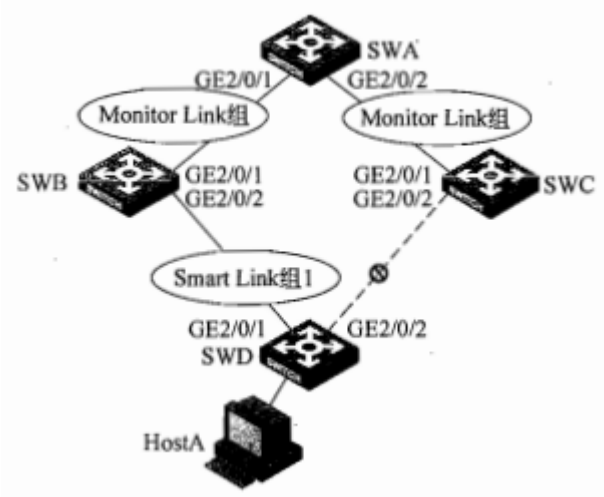
port interface-type interface-number downlink

Alternatively, configure the downlink members in the port view. The configuration command is:

port monitor-link group group-id downlink

3.1.5.2 Smart Link & Monitor Link Configuration Experiment

As shown in the diagram below, switches SWA, SWB, SWC, and SWD are interconnected through their respective Ethernet ports, and HostA connects to the Ethernet port of SWD via an Ethernet cable to access the network.



SWD has dual uplinks to SWA, providing flexible backup for the dual uplink connections. It sends and receives Flush messages within VLAN1 to protect all VLANs. On SWD, GigabitEthernet2/0/1 serves as the primary port, while GigabitEthernet2/0/2 is the secondary port.

SWB and SWC are capable of receiving Flush messages.

(1) Configure SWA.

```
[SWA]interface GigabitEthernet2/0/1
[SWA-GigabitEthernet2/0/1] undo stp
[SWA-GigabitEthernet2/0/1]port link-type trunk
[SWA-GigabitEthernet2/0/1]port trunk permit vlan all
[SWA-GigabitEthernet2/0/1]smart-link flush enable control-vlan 1
[SWA]interface GigabitEthernet2/0/2
[SWA-GigabitEthernet2/0/2]undostp
[SWA-GigabitEthernet2/0/2]port link-type trunk
[SWA-GigabitEthernet2/0/2]port trunk permit vlan all
[SWA-GigabitEthernet2/0/2]smart-link flush enable control-vlan 1
```

(2) Configure SWB.

```
[SWB] interface GigabitEthernet2/0/1
[SWB-GigabitEthernet2/0/1]undostp
[SWB-GigabitEthernet2/0/1]port link-type trunk
[SWB-GigabitEthernet2/0/1]port trunk permit vlan all
[SWB-GigabitEthernet2/0/1]smart-link flush enable control-vlan 1
[SWB]interface GigabitEthernet2/0/2
[SWB-GigabitEthernet2/0/2]undostp
[SWB-GigabitEthernet2/0/2]port link-type trunk
```

[SWB-GigabitEthernet2/0/2]port trunk permit vlan all

[SWB-GigabitEthernet2/0/2]smart-link flush enable control-vlan 1

(3) Configure SWC.

[SWC]interface GigabitEthernet2/0/1

[SWC-GigabitEthernet2/0/1]undostp

[SWC-GigabitEthernet2/0/1]port link-type trunk

[SWC-GigabitEthernet2/0/1port trunk permit vlan all

[SWC-GigabitEthernet2/0/1]smart-link flush enable control-vlan 1

[SWC] interface GigabitEthernet2/0/2

[SWC-GigabitEthernet2/0/2]undostp

[SWC-GigabitEthernet2/0/2]port link-type trunk

[SWC-GigabitEthernet2/0/2]port trunk permit vlan all

[SWC-GigabitEthernet2/0/2]smart-link flush enable control-vlan 1

(4) Configure SWD.

[SWD]interface GigabitEthernet2/0/1

[SWD-GigabitEthernet2/0/1]undostp

[SWD-GigabitEthernet2/0/1]port link-type trunk

[SWD-GigabitEthernet2/0/1]port trunk permit vlan all

[SWD]interface GigabitEthernet2/0/2

[SWD-GigabitEthernet2/0/2] undo stp

[SWD-GigabitEthernet2/0/2]port link-type trunk

[SWD-GigabitEthernet2/0/2]port trunk permit vlan all

[SWD]smart-link group 1

[SWD-smlk-groupl]protected-vlan reference-instance 0 to 32

[SWD-smlk-groupl]port GigabitEthernet2/0/1master

[SWD-smlk-groupl]port GigabitEthernet2/0/2 slave

[SWD-smlk-groupl]smart-link flush enable control-vlan 1

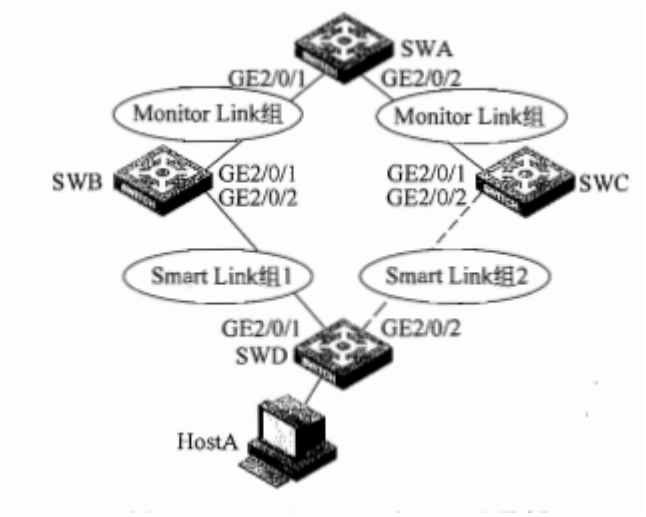
(5) After completing the configuration, check the status of the Smart Link group on the SWD as shown below.

```
[SWD]display smart-link group all
Smart link group 1 information:
  Device ID: 000f-e269-42d1
  Preemption mode: NONE
  Preemption delay: 1(s)
  Control VLAN: 1
  Protected VLAN: Reference Instance 0 to 32
```

Member	Role	State	Flush-count	Last-flush-time
GigabitEthernet2/0/1	MASTER	ACTIVE	1	16:07:11 2000/04/26
GigabitEthernet2/0/2	SLAVE	STANDBY	2	15:53:12 2000/04/26

Based on the information above, SWD created Smart Link Group 1, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/2.

As shown in the figure below, switches SWA, SWB, SWC, and SWD are interconnected through their respective Ethernet ports, and HostA connects to SWD's Ethernet port via an Ethernet link to access the network.



On SWD, dual uplink flexible backup is implemented. The traffic of Smart Link Group 1's reference instance 0 (bound to VLANs 1–100) flows through the link connected to SWB towards SWA, while the traffic of Smart Link Group 2's reference instance 2 (bound to VLANs 101–200) flows through the link connected to SWC towards SWA. Smart Link Group 1 and Group 2 respectively transmit and receive Flush messages within VLAN 10 and VLAN 101.

（1） Configure SWA.

[SWA]interface GigabitEthernet2/0/1

[SWA-GigabitEthernet2/0/1]undostp

[SWA-GigabitEthernet2/0/1]port link-type trunk

[SWA-GigabitEthernet2/0/1]port trunk permit vlan all

[SWA-GigabitEthernet2/0/1]smart-link flush enable control-vlan 10 101

[SWA]interface GigabitEthernet2/0/2

[SWA-GigabitEthernet2/0/2]undostp

[SWA-GigabitEthernet2/0/2]port link-type trunk

[SWA-GigabitEthernet2/0/2]port trunk permit vlan all

[SWA-GigabitEthernet2/0/2]smart-link flush enable control-vlan 10 101

（2） Configure SWB.

[SWB] interface GigabitEthernet2/0/1

[SWB-GigabitEthernet2/0/1]undostp

[SWB-GigabitEthernet2/0/1]port link-type trunk

[SWB-GigabitEthernet2/0/1]port trunk permit vlan all

[SWB-GigabitEthernet2/0/1]smart-link flush enable control-vlan 10 101

[SWB] interface GigabitEthernet2/0/2

[SWB-GigabitEthernet2/0/2]undostp

[SWB-GigabitEthernet2/0/2]port link-type trunk

[SWB-GigabitEthernet2/0/2]port trunk permit vlan all

[SWB-GigabitEthernet2/0/2]smart-link flush enable control-vlan 10 101

（3） Configure SWC.

[SWC]interface GigabitEthernet2/0/1

[SWC-GigabitEthernet2/0/1]undostp

[SWC-GigabitEthernet2/0/1]port link-type trunk

[SWC-GigabitEthernet2/0/1]port trunk permit vlan all

[SWC-GigabitEthernet2/0/1]smart-link flush enable control-vlan 10 101

[SWC]interface GigabitEthernet2/0/2

[SWC-GigabitEthernet2/0/2]undostp

[SWC-GigabitEthernet2/0/2]port link-type trunk

[SWC-GigabitEthernet2/0/2]port trunk permit vlan all

[SWC-GigabitEthernet2/0/2]smart-link flush enable control-vlan 10 101

（4） Configure SWD.

[SWD]vlan 1 to 200

[SWD]stp region-configuration

[SWD-mst-region] instance 0 vlan 1 to 100

[SWD-mst-region]instance 2 vlan 101 to 200

[SWD-mst-region]active region-configuration

[SWD]interface GigabitEthernet2/0/1

[SWD-GigabitEthernet2/0/1]undostp

[SWD-GigabitEthernet2/0/1]port link-type trunk

[SWD-GigabitEthernet2/0/1]port trunk permit vlan all

[SWD]interface GigabitEthernet2/0/2

[SWD-GigabitEthernet2/0/2]undostp

[SWD-GigabitEthernet2/0/2]port link-type trunk

[SWD-GigabitEthernet2/0/2]port trunk permit vlan all

[SWD]smart-link group 1

[SWD-smlk-groupl]protected-vlan reference-instance 0

[SWD-smlk-groupl]port gigabitethernet2/0/1 master

[SWD-smlk-group1]port gigabitethernet2/0/2 slave

[SWD-smlk-groupl]preemption mode role

[SWD-smlk-groupl]flush enable control-vlan 10

[SWD] smart-link group 2

[SWD-smlk-group2]protected-vlanreference-instance 2

[SWD-smlk-group2]port gigabitethernet2/0/2 master

[SWD-smlk-group2]port gigabitethernet2/0/1 slave

[sWD-smlk-group2]preemption mode role

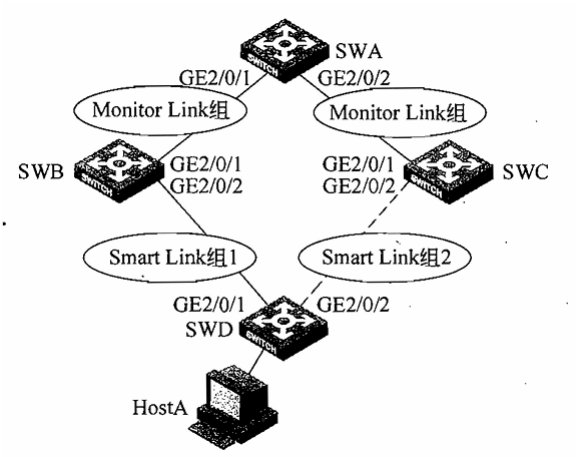
[SWD-smlk-group2]flush enable control-vlan 101

(5) After completing the configuration, check the status of the Smart Link group on the SWD as shown below.

```
[SWD] display smart-link group all
Smart link group 1 information:
Device ID: 000f-e269-42d1
Preemption mode: ROLE
Control VLAN: 10
Protected VLAN: Reference Instance 0
Member                  Role      State     Flush-count  Last-flush-time
--------------------------------------------------------------------------------
GigabitEthernet2/0/1    MASTER    ACTVIE    5            16:37:20 2000/04/26
GigabitEthernet2/0/2    SLAVE     STANDBY   1            17:45:20 2000/04/26

Smart link group 2 information:
Device ID: 000f-e269-42d1

  Preemption mode: ROLE
  Control VLAN: 101
  Protected VLAN: Reference Instance 2
  Member                  Role      State     Flush-count  Last-flush-time
  --------------------------------------------------------------------------------
  GigabitEthernet2/0/2    MASTER    ACTVIE    5            16:37:20 2000/04/26
  GigabitEthernet2/0/1    SLAVE     STANDBY   1            17:45:20 2000/04/26
```

The experiment shows that SWD has created Smart Link Group 1 and Group 2. The primary port for Smart Link Group 1 is GigabitEthernet2/0/1, with the secondary port being GigabitEthernet2/0/2; the primary port for Smart Link Group 2 is GigabitEthernet20/2, with the secondary port being GigabitEthernet2/0/1.

As shown in the diagram below, switches SWA, SWB, SWC, and SWD are interconnected through their respective Ethernet ports. Host A connects to the network via an Ethernet line linked to SWD's Ethernet port.



SWB and SWC can receive Flush messages and are configured with a Monitor Link group. When the ports GigabitEthernet2/0/1 or GigabitEthernet2/0/2 on device SWA fail and go DOWN, the access device SWD can detect the link failure and complete the uplink switchover for the dual-uplink backup in the Smart Link group.

(1) 配置 SWB.

[SWB]monitor-link group 1

[SWB-mtlk-group1]portgigabitethernet2/0/1 uplink

[SWB-mtlk-group1]port gigabitethernet2/0/2 downlink

(2) 配置 SWC.

[SWC]monitor-link group 1

[SWC-mtlk-group1]portgigabitethernet2/0/1 uplink

[SWC-mtlk-groupl]port gigabitethernet2/0/2 downlink

(3) After completing the configuration, check the status of the Monitor Link group on the SWB as shown below.

```
<SWB>display monitor-link group 1
Monitor link group 1 information:
Group status: UP
Last-up-time: 16:52:40 2000/04/26
Last-down-time: -
Member                    Role          Status
--------------------------------------------------------------------------------
GigabitEthernet2/0/1      UPLINK        UP

GigabitEthernet2/0/2      DOWNLINK      UP
```

The experiment shows that a Monitor Link Group 1 was created on the SWB, with the uplink port being GigabitEthernet20/1 and the downlink port being GigabitEthernet2/0/2.

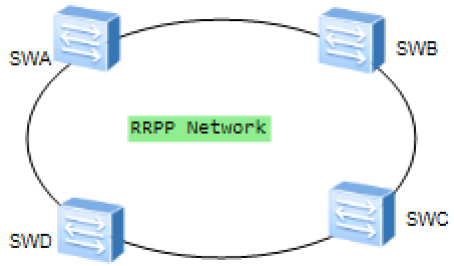
(4) After the configuration is completed, check the status of the Monitor Link group on the SWC as shown below..

```
<SWC>display monitor-link group 1
Monitor link group 1 information:
Group status: UP
Last-up-time: 16:56:35 2000/04/26
Last-down-time: -
Member                    Role          Status
--------------------------------------------------------------------------------
GigabitEthernet2/0/1      UPLINK        UP
GigabitEthernet2/0/2      DOWNLINK      UP
```

From the above experimental information, it can be seen that a Monitor Link Group 1 was created on the SWC, with the uplink port being GigabitEthernet2/0/1 and the downlink port being GigabitEthernet2/0/2.

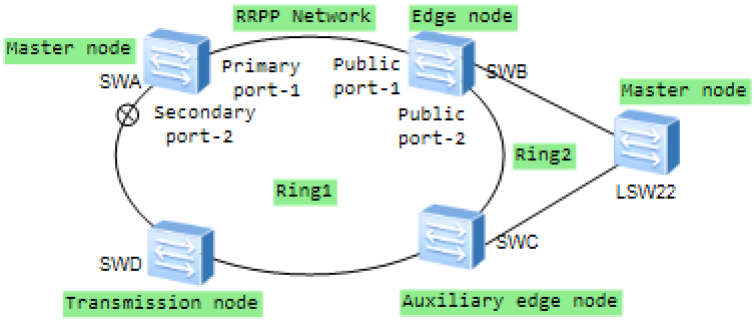
3.1.4 RRPP Technology

3.1.4.1 Functions of RRPP



As shown in the figure above, RRPP is a link-layer protocol specifically designed for Ethernet rings. It prevents broadcast storms caused by data loops when the Ethernet ring is intact and can quickly restore communication paths between nodes on the ring when a link in the Ethernet ring breaks. RRPP offers faster convergence compared to STP, and its convergence time is independent of the number of nodes on the ring, making it suitable for networks with larger diameters.

3.1.4.2 Basic Concepts of RRPP



As shown in the network topology diagram above, commonly used concepts and terms in RRPP ring networks include the following:

1. RRPP Domain: Devices that share the same domain ID and control VLAN and are interconnected form an RRPP domain. An RRPP domain consists of elements such as the RRPP major ring, sub-ring, control VLAN, master node, transit nodes, master port and secondary port, common ports, and edge ports.

2. RRPP Ring: An RRPP ring is an Ethernet network topology connected in a circular structure. RRPP rings are classified as either major rings or sub-rings. The roles of these rings can be determined by specifying the RRPP ring level, with the major ring designated as level 0 and sub-rings as level 1. An RRPP domain can contain one or more RRPP rings, but it can only have one major ring, with all others being sub-rings. The status of an RRPP ring includes the "Complete" state, where all physical links in the ring network are intact, and the "Failed" state, where a physical link in the ring is broken.

3. Node: Every device on an RRPP ring is referred to as a node. The role of a node is determined by user configuration.

4. Master Node: Each ring has one and only one master node. The master node initiates the active detection mechanism for the ring's status and makes decisions when changes in the network topology occur. The master node operates in two states: "Complete State" and "Failed State."

5. Transit Node: All nodes on the major ring other than the master node, and all nodes on sub-rings except the master node and the nodes where the sub-ring intersects the major ring, are considered transit nodes. Transit nodes monitor the status of their directly connected RRPP links and report link changes to the master node, which decides how to handle them. Transit nodes can be in one of the following three states: "Link-Up State," "Link-Down State," and "Preforwarding State" (temporary blocked state).

6. Edge Node: A node that exists on both the major ring and a sub-ring is a special type of transit node. It functions as a transit node on the major ring but is an edge node on the sub-ring. Edge nodes have the same three states as transit nodes. The state transitions of edge nodes are similar to those of transit nodes, except that when a link status change triggers a state transition on an edge node, it only affects the edge port's state.

7. Auxiliary Edge Node: Like edge nodes, auxiliary edge nodes are located on both the major ring and a sub-ring and are a special type of transit node. They function as transit nodes on the major ring and as auxiliary edge nodes on the sub-ring. Auxiliary edge nodes are used in pairs with edge nodes to detect the integrity of the major ring and prevent loops. They also share the same three states as transit nodes and have similar state transitions, with the difference being that state transitions triggered by link status changes only affect the edge port status.

8. Control VLAN: Control VLANs are used to carry RRPP protocol messages. All ports connected to an RRPP ring must be included in the control VLAN, and only these ports can be added to the VLAN. Each RRPP domain has two control VLANs: the primary control VLAN and the sub-control VLAN. The control VLAN for the major ring is called the primary control VLAN, and the one for a sub-ring is called the sub-control VLAN. When configuring, only the primary control VLAN needs to be specified; the system automatically assigns the VLAN with an ID value that is one higher than the primary control VLAN as the sub-control VLAN. All sub-rings in the same RRPP domain share the same sub-control VLAN. Additionally, no IP addresses can be configured on interfaces associated with either the primary or sub-control VLAN.

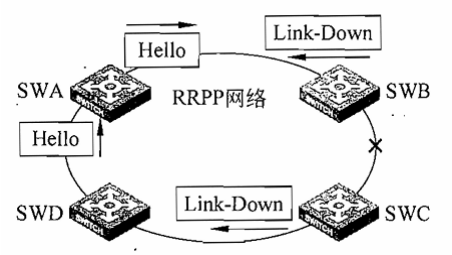
9. Data VLAN: In contrast to control VLANs, data VLANs are used to transmit data packets. A data VLAN can include both RRPP and non-RRPP ports.

10. Master Port and Secondary Port: Both master and transit nodes have two ports connected to an RRPP ring: one is the master port, and the other is the secondary port. The role of each port is determined by user configuration. For master nodes, the master port and secondary port have distinct functions: the master port is used to send loop-detection messages, while the secondary port receives these messages. When the RRPP ring is in a healthy state, the secondary port of the master node logically blocks the data VLAN but permits control VLAN messages to pass through. If the RRPP ring enters a failed state, the secondary port unblocks the data VLAN and forwards its packets. For transit nodes, there is no functional difference between the master port and the secondary port; both are used for transmitting protocol and data packets on the RRPP ring.

11. Common Ports and Edge Ports: Common ports are the ports on edge nodes and auxiliary edge nodes that connect to the major ring. Each edge node or auxiliary edge node has two such common ports configured on the major ring. Edge ports, on the other hand, are the ports on edge nodes and auxiliary edge nodes that connect only to sub-rings.

3.1.4.3 RRPP Operating Mechanism

The operating mechanism of RRPP primarily includes the Polling mechanism and the Link State Change Notification mechanism. The Polling mechanism is a method where the master node of the RRPP ring actively monitors the health status of the ring network. The Link State Change Notification mechanism provides a faster way to detect topology changes in the ring network compared to the Polling mechanism.



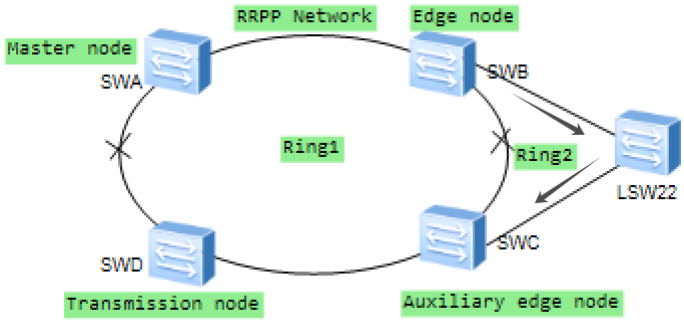
As shown in the figure above, the Polling mechanism enables the master node to periodically send Hello packets from its primary port, which propagate through each transit node along the ring. If the ring is intact, the master node's secondary port will receive the Hello packet before the timer expires, and the master node will maintain the blocking state of the secondary port. If the ring is broken, the master node's secondary port will not receive the Hello packet before the timer expires. In this case, the master node will unblock the data VLAN on the secondary port and send a Common-Flush-FDB packet to notify all transit nodes to update their respective MAC table entries and ARP/ND table entries.

In the link state change notification mechanism, the notifier is the transit node. When a transit node, edge node, or assistant edge node detects that any port belonging to the RRPP domain has gone DOWN, it immediately sends a Link-Down packet to the master node. Upon receiving the Link-Down packet, the master node immediately unblocks the data VLAN on its secondary port and sends a Common-Flush-FDB packet to notify all transit nodes, edge nodes, and assistant edge nodes to update their respective MAC

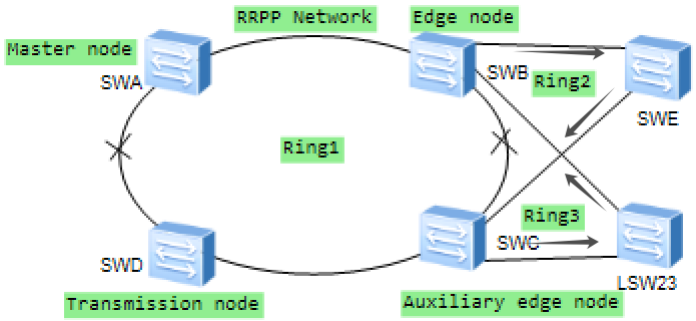
table entries and ARP/ND table entries. Once all nodes have updated their entries, the data flow switches to the normal link.

When a port belonging to the RRPP domain on a transit node, edge node, or assistant edge node comes back UP, the master node may take some time to detect the ring's restoration. During this period, a temporary loop may form in the network for the data VLAN, potentially causing a broadcast storm. To prevent temporary loops, non-master nodes immediately block their ports connecting to the ring (allowing only control VLAN packets to pass) upon detecting that these ports have come back UP. The ports are unblocked only after confirming that no loop will occur.

Protocol packets for sub-rings are transmitted between the edge ports of edge nodes and assistant edge nodes via channels provided by the main ring. The main ring functions as if it were a single node on the sub-ring, and the sub-ring protocol packets are treated as data packets within the main ring.

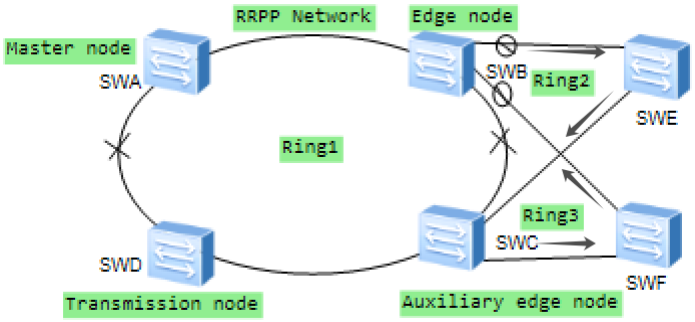


As shown in the figure above, when a fault occurs on the primary ring link and the channel for sub-ring protocol messages between the edge node and the auxiliary edge node is interrupted (a failure on the common link between the primary ring and the sub-ring, along with at least one additional failure on the non-common link), the sub-ring master node will not receive the Hello message it sent out. Consequently, the Fail timer times out, and the sub-ring master node transitions to the Failed state, releasing the secondary port to ensure the network path remains unobstructed.



As shown in the diagram above, the two sub-rings, Ring2 and Ring3, which are dual-homed in the network, are interconnected via the edge nodes SWB and SWC. When a failure occurs on the primary ring, Ring1, both main ring links between the edge node and the auxiliary edge node go into a DOWN state. Due to deficiencies in the link status notification mechanism, the root nodes of sub-rings Ring2 and Ring3 release their respective secondary ports, creating loops between devices and resulting in a broadcast storm.

To eliminate the broadcast storm issues caused by dual-homing in RRPP (Rapid Ring Protection Protocol) networking, a sub-ring protocol message channel status detection mechanism is introduced within the primary ring. This mechanism requires coordination between the edge node and the auxiliary edge node. Its purpose is to block the edge port of the edge node before the secondary ports of the sub-rings' root nodes are released, thereby preventing the formation of data loops between the sub-rings. The edge node serves as the initiator and decision-maker for detection, while the auxiliary edge node acts as a channel status listener, responsible for promptly notifying the edge node of any changes in channel status.

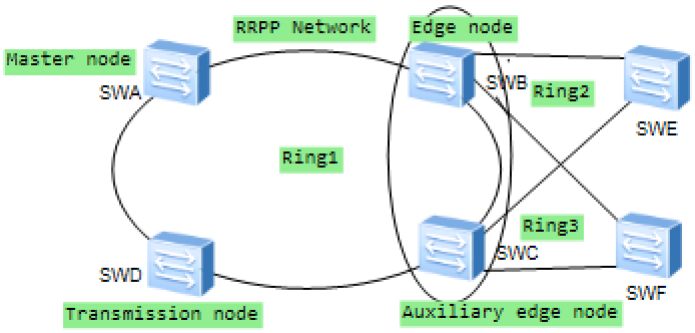


As described above, the edge nodes of the RRPP sub-ring periodically send Edge-Hello packets into the main ring through the shared ports of the main ring. These packets are sequentially forwarded by each node on the main ring to the auxiliary edge node. If the auxiliary edge node receives the Edge-Hello packet within the specified time, it indicates that the communication channel is functioning properly. Conversely, if the packet is not received, it signifies a channel interruption.

When the auxiliary edge node detects a disruption in the sub-ring protocol communication channel, it immediately sends a Major-Fault packet from its edge port to the edge node through the sub-ring link. Upon receiving the Major-Fault packet, the edge node blocks its edge port to prevent a loop.

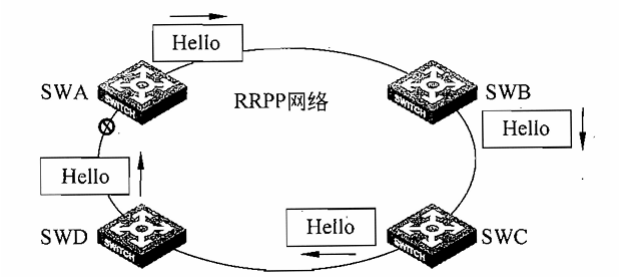
When the main ring link recovers, restoring communication between the edge node and the auxiliary edge node, the sub-ring protocol communication channel also resumes normal operation. The primary node of the sub-ring then receives its own Hello packet through the secondary port, switching to the Complete state and blocking the secondary port. The primary node of the sub-ring sends a Complete-Flush-FDB packet from its primary port, and upon receiving this packet, the edge node unblocks its edge port.

Due to the activation of the sub-ring protocol communication channel status checking mechanism, certain fault conditions may cause the RRPP ring to break, rendering backup functionality unavailable.

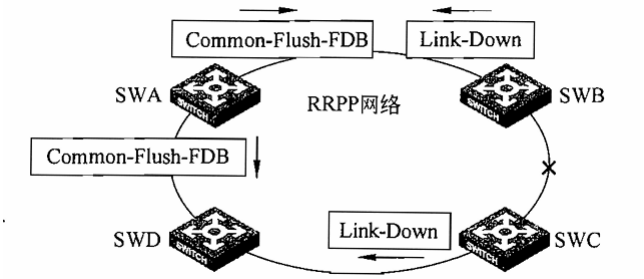


As shown in the figure above, within the RRPP ring group configured at the edge node, only the active sub-ring with the smallest domain ID and ring ID sends Edge-Hello packets. In the auxiliary edge node ring group, any active sub-ring that receives an Edge-Hello packet will notify other active sub-rings. By configuring RRPP ring groups on the edge node and auxiliary edge node respectively, only one sub-ring sends or receives Edge-Hello packets, thereby reducing the impact on the device's CPU.

3.1.4.4 RRPP Ring Topology Change Process



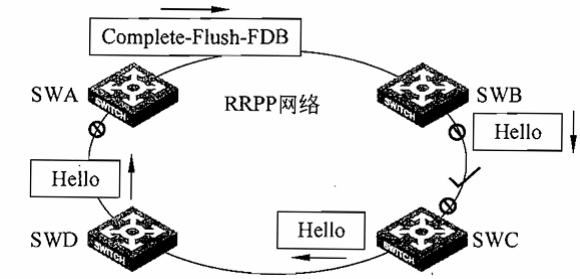
As shown in the figure above, when all links in the entire ring network are in the UP state, the RRPP ring is in the Complete (healthy) state, and the status of the master node reflects the Complete state of the entire ring network. When the ring network is in the Complete state, to prevent broadcast loops caused by data packets, the master node blocks its secondary port. The master node periodically sends Hello packets from its primary port, which pass sequentially through each transit node before finally returning to the master node from its secondary port.



As shown in the figure above, when a link in the entire ring network is in a DOWN state, the RRPP ring enters the Failed state, and the status of the master node reflects the Failed state of the entire ring network. When an RRPP port on a transmission node switch experiences a link DOWN event, the node will send a Link-Down message from the RRPP port paired with the failed port, which is in the UP state, to notify the master node. Upon receiving the Link-Down message, the master node immediately switches its status to the Failed state and unblocks the secondary port. Since the network topology has changed, to prevent misdirected traffic, the master node must also refresh the FDB

table and send a Common-Flush-FDB message from the primary port to notify all transmission nodes to refresh their FDB tables.

The fault report is initiated by the transmission node. If the Link-Down message is unfortunately lost during transmission, the master node's ring network status detection mechanism comes into play. If the master node's secondary port does not receive a Hello message sent by the master node within the specified time, it also determines that a fault has occurred in the ring network. The fault handling process in this case is the same as when the transmission node proactively reports the fault.



As shown in the figure above, when the RRPP ring topology state and the master node state are in the Complete state, the corresponding state of the transmission nodes should be in the Link-Up state. When the RRPP ring topology state and the master node state are in the Failed state, some of the transmission nodes should be in the Link-Down state. When the RRPP port on the transmission node's switch recovers, the state of the transmission node does not immediately transition from Link-Down to Link-Up. Instead, it first transitions to the Preforwarding state. In this state, the recovered port is blocked to prevent the formation of temporary loops.

The recovery of the data channel in the ring network is actively initiated by the master node. The master node periodically sends Hello packets from the primary port. Once all the faulty links in the ring network are restored, the master node will receive its own Hello packet from the secondary port. Upon receiving its own Hello packet, the master node first transitions its state back to the Complete state, blocks the secondary port, and then sends a Complete-Flush-FDB packet from the primary port. Transmission nodes in the Preforwarding state, upon receiving the Complete-Flush-FDB packet, transition back to the Link-Up state, unblock the temporarily blocked ports, and refresh the FDB table. If the Complete-Flush-FDB packet is lost during transmission, a backup mechanism is in place to recover the temporarily blocked ports of the transmission nodes. Specifically, if a transmission node in the Preforwarding state does not receive a Complete-Flush-FDB packet from the master node within a specified time, the transmission node will independently unblock the temporarily blocked ports and resume data communication.

3.1.4.5 RRPP Protocol Messages

RRPP protocol messages include six types: Health (Hello), Link-Down, Common-Flush-FDB, Complete-Flush-FDB, Edge-Hello, and Major-Fault.

Health (Hello): A health check message initiated by the master node to verify the integrity of the network loop.

Link-Down: A link-down message sent by a transmission node, edge node, or auxiliary edge node when a direct link status changes to DOWN. This notifies the master node of a link-down event on the loop, indicating the physical loop has disappeared.

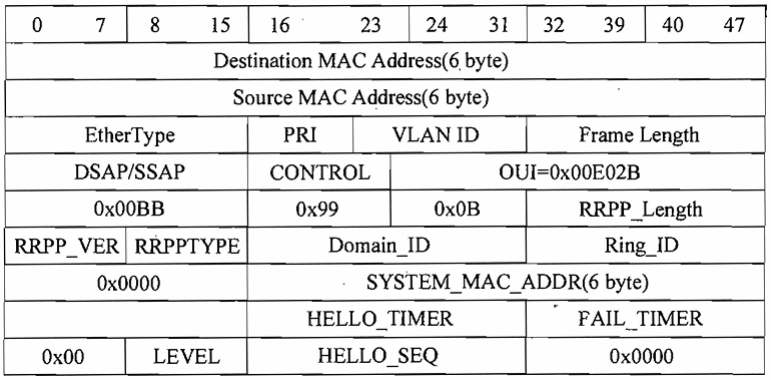
Common-Flush-FDB: A flush FDB (Forwarding Database) message initiated by the master node to notify transmission nodes, edge nodes, or auxiliary edge nodes to update their respective MAC address forwarding tables.

Complete-Flush-FDB: A recovery flush FDB message initiated by the master node to notify transmission nodes, edge nodes, or auxiliary edge nodes to update their respective MAC address forwarding tables. It also instructs transmission nodes to unblock temporarily blocked ports.

Edge-Hello:A primary loop integrity check message initiated by the edge node of a sub-ring and received by the auxiliary edge node of the same sub-ring. This message is used by the sub-ring to verify the loop integrity of the primary ring in its domain.

Major-Fault: A major fault notification message. If the auxiliary edge node of a sub-ring does not receive the Edge-Hello message from the edge node within a specified time, it reports a fault in the primary ring of its domain to the edge node.

The fields of the RRPP message are explained as shown in the figure below:



1) Destination MAC Address: 48 bits, the destination MAC address of the protocol packet.

2) Source MAC Address: 48 bits, the source MAC address of the protocol packet, which is 0x000FE203FD75.

3) EtherType: 16 bits, the encapsulation type field of the packet, which is 0x8100, indicating Tagged encapsulation.

4) PRI: 4 bits, the Class of Service (COS) priority, which is 0xE0.

5) VLAN ID: 12 bits, the VLAN ID of the packet.

6) Frame Length: 16 bits, the length of the Ethernet frame, which is 0x48.

7) DSAP/SSAP: 16 bits, Destination Service Access Point/Source Service Access Point, which is 0xAAAA.

8) CONTROL: 8 bits, which is 0x03.

9) OUI: 24 bits, which is 0x00E02B.

10) RRPP Length: 16 bits, the length of the RRPP data unit, which is 0x40.

11) RRPP VER: 16 bits, the version information of RRPP, which is 0x0001.

12) Domain_ID: 16 bits, the ID of the RRPP domain to which the packet belongs.

13) Ring ID: 16 bits, the ID of the RRPP ring to which the packet belongs.

14) SYSTEM MAC ADDR: 48 bits, the bridge MAC address of the node sending the packet.

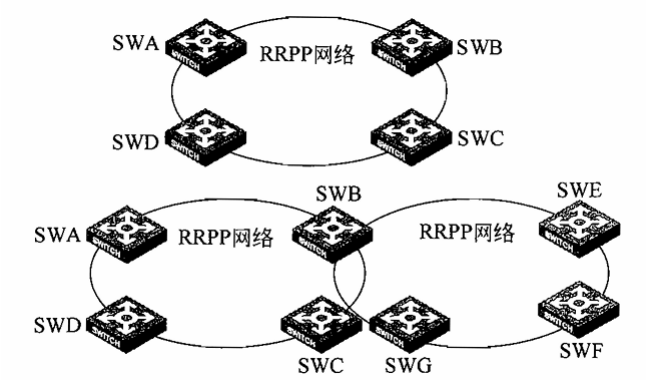
15) HELLO_TIMER: 16 bits, the timeout value of the Hello timer used by the node sending the packet, measured in seconds (s).

16) FAIL TIMER: 16 bits, the timeout value of the Fail timer used by the node sending the packet, measured in seconds (s).

17) LEVEL: 8 bits, the level of the RRPP ring to which the packet belongs.

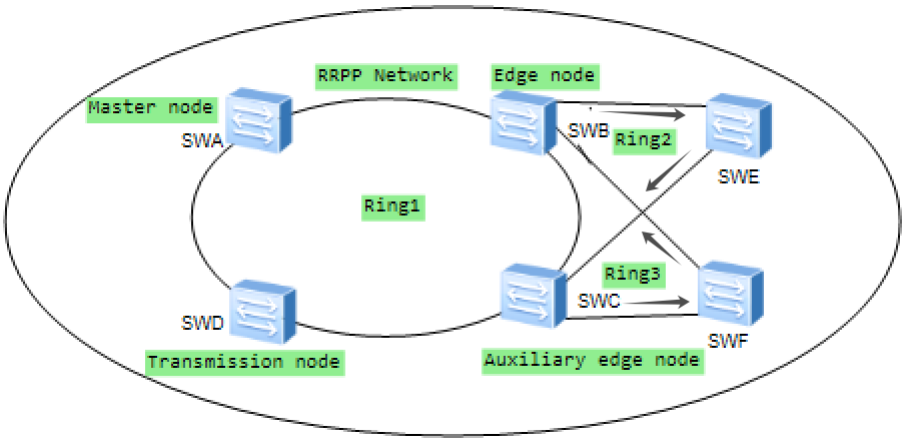
18) HELLO SEQ: 16 bits, the sequence number of the Hello packet.

3.1.4.6 RRPP Network Configuration Experiment



As shown in the figure above, a single-ring RRPP refers to a network topology with only one ring. In this case, it is only necessary to define one RRPP domain and one RRPP ring. This type of network configuration is characterized by fast response to topology changes and short convergence times, making it suitable for applications where there is only one ring in the network.

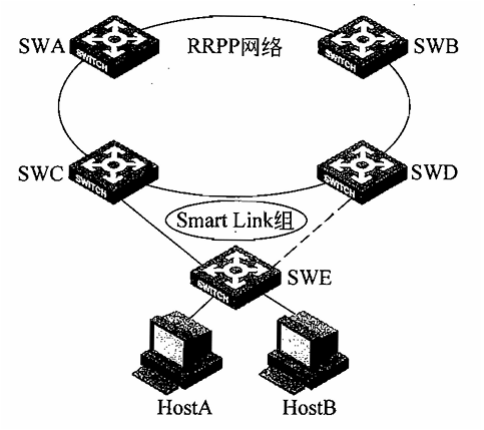
Tangent rings in RRPP refer to a network topology with two or more rings, where each pair of rings shares only one common node. In this scenario, each ring must belong to a different RRPP domain. This configuration is commonly used in large-scale networks or when peer-level networks require regional management.



As shown in the figure above, an RRPP intersecting ring refers to a network topology where there are two or more rings with two shared nodes between them. In this case,

only one RRPP domain needs to be defined, with one ring designated as the primary ring and the others as sub-rings.

A typical application of an RRPP intersecting ring is dual-homing networking: the primary node of the sub-ring can achieve dual-homing uplink through two edge nodes, providing uplink link redundancy.



As shown in the diagram above, this is an example of a networking application where Smart Link and RRPP are used together.

In this network setup, RRPP is enabled on devices SWA, SWB, SWC, and SWD to provide link redundancy. Since the two ports connected between SWC and SWD already have RRPP enabled and cannot enable STP, link redundancy on SWE can only be achieved by configuring a Smart Link group.

3.1.4.7 RRPP Configuration Commands

Before configuring RRPP, you need to establish an Ethernet ring network topology.

Since RRPP does not have an automatic election mechanism, detection and protection of the ring network can only be achieved when configurations on all nodes within the network are correct. Thus, ensuring configuration accuracy is critical.

The steps for configuring RRPP are as follows:

1. Create an RRPP Domain in System View

Command:

rrpp domain domain-id

When creating an RRPP domain, you must specify a domain ID. The domain ID uniquely identifies one RRPP domain, and the same domain ID must be configured on all nodes within the same RRPP domain.

2. Configure the Control VLAN in RRPP Domain View

Command:

control-vlanvlan-id

Before configuring an RRPP ring, you must first set up the control VLAN. The same control VLAN must be configured on all nodes in the same RRPP domain. Note that QinQ and VLAN mapping functions cannot run within the control VLAN; otherwise, RRPP packets cannot be transmitted properly.

3. Configure the Protected VLAN in RRPP Domain View

Command:

protected-vlan reference-instance instance-id-list

Before configuring an RRPP ring, you must also set up the protected VLAN. All VLANs that the RRPP ports permit must be protected within the RRPP domain. The same protected VLANs must be configured across all nodes in the same RRPP domain.

4. Configure the Current Device as the Master Node in RRPP Domain View

Command:

ring ring-id **node-mode master** [**primary-port** interface -type interface-number][**secondary-port**interface-type interface-number] **level** level-value

5. **Configure the Current Device as a Transit Node in RRPP Domain View**

Command:

ring ring-id **node-mode transit** [**primary-port** interface-type interface-number] [**secondary-port**interface-type interface-number]**level** level-value

When configuring an RRPP ring, first properly configure the RRPP ports on each node that is to join the ring. Following this, configure all nodes within the RRPP ring. RRPP ports must be Layer 2 Ethernet ports, Layer 2 GE ports, Layer 2 XGE ports, or Layer 2 aggregated ports, but they cannot belong to aggregation groups, service loopback groups, or Smart Link groups.

6. Set the Current Device as the Edge Node for a Subring in RRPP Domain View

Command:

ring ring-id **node-mode edge**[**edge-port** interface-type interface-number]

When configuring an edge node, the primary ring must be configured first, followed by the subring.

7. Set the Current Device as an Assistant Edge Node for a Subring in RRPP Domain View

Command:

ring ring-id **node-mode** assistant-edge [edge-port interface-type inter face-number]

When configuring an assistant edge node, configure the primary ring first, followed by the subring. For both edge nodes and assistant edge nodes, the primary ring must be enabled before enabling subrings, and subrings within the RRPP domain must be disabled before disabling the primary ring.

8. Enable RRPP in System View

Command:

rrpp enable

9. Enable the RRPP Ring in RRPP Domain View

Command:

ring ring-id **enable**

After enabling both RRPP and the RRPP ring, the RRPP functionality on the current device is activated. By grouping subrings sharing the same edge/assistant-edge configurations, the number of Edge-Hello packets sent and received can be reduced. Ring groups should be configured on both the edge and assistant-edge nodes. Note, however, that a subring can only belong to one ring group, and subrings configured in the edge and assistant-edge nodes must be identical; otherwise, the ring group will not function properly.

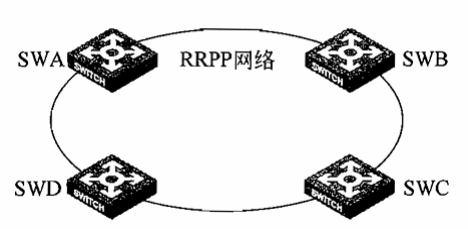
To create an RRPP ring group, use the following command in System View:

rrpp ring-group ring-group-id

To add a subring to an RRPP ring group, use the following command in System View:

domain domain-id **ring** ring-id-list

3.1.4.8 RRPP Network Experiment (Part 1)



As shown in the figure above, SWA, SWB, SWC, and SWD form RRPP Domain 1. The control VLAN for this domain is VLAN 4092, which protects all VLANs. SWA serves as the primary node of the main ring, with GigabitEthernet2/0/1 as the primary port and GigabitEthernet2/0/2 as the secondary port. SWB, SWC, and SWD function as transit nodes in the main ring, with GigabitEthernet2/0/1 as the primary port and GigabitEthernet2/0/2 as the secondary port. The specific configuration steps are as follows:

(1) Configure SWA.

[SWA]interface GigabitEthernet2/0/1

[SWA-GigabitEthernet2/0/1]undostp

[SWA-GigabitEthernet2/0/1]port link-type trunk

[SWA-GigabitEthernet2/0/1]port trunk permit vlan all

[SWA] interface GigabitEthernet2/0/2

[SWA-GigabitEthernet2/0/2]undostp

[SWA-GigabitEthernet2/0/2]port link-type trunk

[SWA-GigabitEthernet2/0/2]port trunk permit vlan all

[SWA]rrpp domain 1

[SWA-rrpp-domain1]control-vlan4092

[SWA-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWA-rrpp-domain1]ring 1 node-mode master primary-port gigabitethernet 2/0/1 secondary-portgigabitethernet 2/0/2 level 0

[SWA-rrpp-domain1]ring1 enable

[SWA]rrpp enable

(2) Configure SWB.

[SWB]interface GigabitEthernet2/0/1

[SWB-GigabitEthernet2/0/1]undostp

[SWB-GigabitEthernet2/0/1]port link-type trunk

[SWB-GigabitEthernet2/0/1]port trunk permit vlan all

[SWB]interface GigabitEthernet2/0/2

[SWB-GigabitEthernet2/0/2]undostp

[SWB-GigabitEthernet2/0/2]port link-type trunk

[SWB-GigabitEthernet2/0/2]port trunk permit vlan all

[SWB]rrpp domain 1

[SWB-rrpp-domain1]control-vlan 4092

[SWB-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWB-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 2/0/1secondary-portgigabitethernet 2/0/2 level 0

[SWB-rrpp-domainl]ring 1 enable

[SWB]rrpp enable

(3) Configure SWC.

[SWC]interface GigabitEthernet2/0/1

[SWC-GigabitEthernet2/0/1]undostp

[SWC-GigabitEthernet2/0/1]port link-type trunk

[SWC-GigabitEthernet2/0/1]port trunk permit vlan all

[SWC]interface GigabitEthernet2/0/2

[SWC-GigabitEthernet2/0/2]undostp

[SWC-GigabitEthernet2/0/2]port link-type trunk

[SWC-GigabitEthernet2/0/2]port trunk permit vlan all

[SWC]rrpp domain 1

[SWC-rrpp-domainl]control-vlan 4092

[SWC-rrpp-domain1] protected-vlan reference-instance 0 to 32

[SWC-rrpp-domain1]ring 1 node-mode transit primary-portgigabitethernet2/0/1secondary-portgigabitethernet 2/0/2 level 0

[SWC-rrpp-domainl]ring 1 enable

[SWC]rrpp enable

(4) Configure SWD.

[SWD]interface GigabitEthernet2/0/1

[SWD-GigabitEthernet2/0/1]undostp

[SWD-GigabitEthernet2/0/1]port link-type trunk

[SWD-GigabitEthernet2/0/1]port trunk permit vlan all

[SWD]interface GigabitEthernet2/0/2

[SWD-GigabitEthernet2/0/2]undostp

[SWD-GigabitEthernet2/0/2]port link-type trunk

[SWD-GigabitEthernet2/0/2]port trunk permit vlan all

[SWD]rrpp domain 1

[SWD-rrpp-domain1]control-vlan 4092

[SWD-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWD-rrpp-domainl]ring 1 node-mode transit primary-port gigabitethernet 2/0/1secondary-portgigabitethernet 2/0/2 level 0

[SWD-rrpp-domainl]ring 1 enable

[SWD]rrpp enable

(5) After completing the configuration, check the RRPP status on the SWA as shown below.

```
[SWA]dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Master
Ring State      : Complete
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: BLOCKED
```

Based on the above information, it can be seen that SWB is a transmission node of RRPP Ring1, with the main port as GigabitEthernet2/0/1 and the auxiliary port as GigabitEthernet2/0/1.

(6) After the configuration is completed, check the RRPP status on the SWB, as shown below.

```
<SWB>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: UP
```

From the above information, it can be seen that SWB is the transmission node of RRPP Ring1, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/1.

(7) After the configuration is completed, check the RRPP status on the SWC, as shown below.

```
<SWC>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -

Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: UP
```

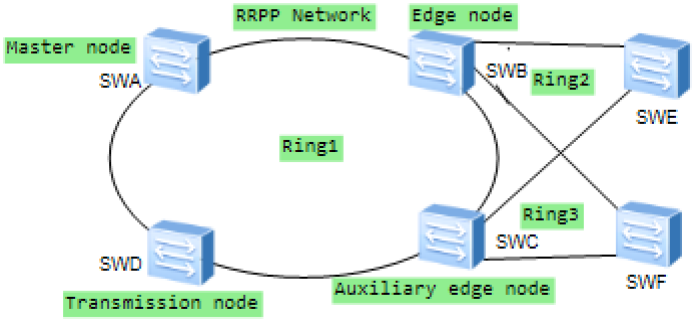
Based on the above information, it can be observed that SWC serves as the RRPP Ring1 transmission node, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/2.

(8) After the configuration is completed, check the RRPP status on the SWD, as shown below.

```
<SWD>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: UP
```

Based on the above information, it can be seen that SWD is the transmission node for RRPP Ring1, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/2.

3.1.4.9 RRPP Network Experiment (Part 2)



As shown in the figure above, SWA, SWB, SWC, SWD, SWE, and SWF form RRPP Domain 1, with the control VLAN for this domain being VLAN 4092, protecting all VLANs. SWA, SWB, SWC, and SWD form Major Ring 1, while SWB, SWC, and SWE form Sub-Ring 2, and SWB, SWC, and SWF form Sub-Ring 3.

SWA serves as the master node of the major ring, with GigabitEthernet2/0/1 as the master port and GigabitEthernet2/0/2 as the secondary port. SWE functions as the master node of Sub-Ring 2, with GigabitEthernet2/0/1 as the master port and GigabitEthernet2/0/2 as the secondary port. SWF is the master node of Sub-Ring 3, with GigabitEthernet2/0/1 as the master port and GigabitEthernet2/0/2 as the secondary port. SWB acts as both a transit node in the major ring and an edge node in the sub-rings, with GigabitEthernet2/0/3 and GigabitEthernet2/0/4 serving as edge ports. SWC is both a transit node in the major ring and an auxiliary edge node in the sub-rings, with GigabitEthernet2/0/3 and GigabitEthernet2/0/4 serving as edge ports. SWD is a transit node in the major ring, with GigabitEthernet2/0/1 as the master port and GigabitEthernet2/0/2 as the secondary port. The detailed configuration steps are as follows:

(1) Configure SWA.

[SWA]interface GigabitEthernet2/0/1

[SWA-GigabitEthernet2/0/1]undostp

[SWA-GigabitEthernet2/0/1]port link-type trunk

[SWA-GigabitEthernet2/0/1]port trunk permit vlan all

[SWA]interface GigabitEthernet2/0/2

[SWA-GigabitEthernet2/0/2]undostp

[SWA-GigabitEthernet2/0/2]port link-type trunk

[SWA-GigabitEthernet2/0/2]port trunk permit vlan all

[SWA]rrpp domain 1

[SWA-rrpp-domain1]control-vlan 4092

[SWA-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWA-rrpp-domain1]ring1 node-mode master primary-port gigabitethernet 2/0/1 secondary-portgigabitethernet 2/0/2 level 0

[SWA-rrpp-domain1]ring1 enable

[SWA]rrpp enable

(2) Configure SWB.

[SWB]interface GigabitEthernet2/0/1

[SWB-GigabitEthernet2/0/1]undostp

[SWB-GigabitEthernet2/0/1]port link-type trunk

[SWB-GigabitEthernet2/0/1]port trunk permit vlan all

[SWB]interface GigabitEthernet2/0/2

[SWB-GigabitEthernet2/0/2]undostp

[SWB-GigabitEthernet2/0/2]port link-type trunk

[SWB-GigabitEthernet2/0/2]port trunk permit vlan all

[SWB]interface GigabitEthernet2/0/3

[SWB-GigabitEthernet2/0/3]undostp

[SWB-GigabitEthernet2/0/3]port link-type trunk

[SWB-GigabitEthernet2/0/3] port trunk permit vlan all

[SWB]interface GigabitEthernet2/0/4

[SWB-GigabitEthernet2/0/4]undostp

[SWB-GigabitEthernet2/0/4]port link-type trunk

[SWB-GigabitEthernet2/0/4]port trunk permit vlan all

[SWB]rrpp domain 1

[SWB-rrpp-domain1]control-vlan 4092

[SWB-rrpp-domain1]protected-ylan reference-instance 0 to 32

[SWB-rrpp-domain1]ring 1 node-mode transit primary-port gigabitethernet 2/0/1 secondary-portgigabitethernet 2/0/2 level 0

[SWB-rrpp-domain1]ring l enable

[SWB-rrpp-domain1]ring 2 node mode edge edge-port gigabitethernet 2/0/3

[SWB-rrpp-domainl]ring 2 enable

[SWB-rrpp-domainl]ring 3 nodemode edge edge-port gigabitethernet 2/0/4

[SWB-rrpp-domainl]ring 3 enable

[SWB]rrpp enable

(3) Configure SWC

[SWC]interface GigabitEthernet2/0/1

[SWC-GigabitEthernet2/0/1]undostp

[SWC-GigabitEthernet2/0/1]port link-type trunk

[SWC-GigabitEthernet2/0/1]port trunk permit vlan all

[SWC]interface GigabitEthernet2/0/2

[SWC-GigabitEthernet2/0/2]undostp

[SWC-GigabitEthernet2/0/2]port link-type trunk

[SWC-GigabitEthernet2/0/2]port trunk permit vlan all

[SWC]interface GigabitEthernet2/0/3

[SWC-GigabitEthernet2/0/3]undostp

[SWC-GigabitEthernet2/0/3]port link-type trunk

[SWC-GigabitEthernet2/0/3]port trunk permit vlan all

[SWC]interface GigabitEthernet2/0/4

[SWC-GigabitEthernet2/0/4]undostp

[SWC-GigabitEthernet2/0/4]port link-type trunk

[SWC-GigabitEthernet2/0/4]port trunk permit vlan all

[SWC]rrpp domain 1

[SWC-rrpp-domain1]control-vlan 4092

[SWC-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWC-rrpp-domain1]ring 1 node-mode transit primary-port gigabitethernet 2/0/1 secondary-portgigabitethernet 2/0/2 level 0

[SWC-rrpp-domain1]ring 1 enable

[SWC-rrpp-domain1]ring 2 node-mode edge assistant-edge gigabitethernet 2/0/3

[SWC-rrpp-domain1]ring 2 enable

[SWC-rrpp-domain1]ring 3 node-mode edge assistant-edge gigabitethernet 2/0/4

[SWC-rrpp-domain1]ring 3 enable

ISWC]rrpp enable

(4) Configure SWD

[SWD]interface GigabitEthernet2/0/1

[SWD-GigabitEthernet2/0/1]undostp

[SWD-GigabitEthernet2/0/1]port link-type trunk

[SWD-GigabitEthernet2/0/1]port trunk permit vlan all

[SWD]interface GigabitEthernet2/0/2

[SWD-GigabitEthernet2/0/2]undostp

[SWD-GigabitEthernet2/0/2]port link-type trunk

[SWD-GigabitEthernet2/0/2]port trunk permit vlan all

[SWD]rrpp domain 1

[SWD-rrpp-domain1]control-vlan 4092

[SWD-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWD-rrpp-domain1]ring l node-mode transit primary-portgigabitethernet 2/0/1secondary-portgigabitethernet 2/0/2 level 0

[SWD-rrpp-domain1]ring 1 enable

[SWD]rrpp enable

(5) Configure SWE

[SWE]interface GigabitEthernet2/0/1

[SWE-GigabitEthernet2/0/1]undostp

[SWE-GigabitEthernet2/0/1]port link-type trunk

[SWE-GigabitEthernet2/0/1]port trunk permit vlan all

[SWE]interface GigabitEthernet2/0/2

[SWE-GigabitEthernet2/0/2]undostp

[SWE-GigabitEthernet2/0/2]port link-type trunk

[SWE-GigabitEthernet2/0/2]port trunk permit vlan all

[SWE]rrpp domain 1

[SWE-rrpp-domain1]control-vlan 4092

[SWE-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWE-rrpp-domain1]ring 2 node-mode masterprimary-port gigabitethernet 2/0/1 secondary-portgigabitethernet 2/0/2 level 1

[SWE-rrpp-domain1]ring 2 enable

[SWE]rrpp enable

(6) Configure SWF

[SWF]interface GigabitEthernet2/0/1

[SWF-GigabitEtheret2/0/1]undostp

[SWF-GigabitEthernet2/0/1]port link-type trunk

[SWF-GigabitEthernet2/0/1]port trunk permit vlan all

[SWF] interface GigabitEthernet2/0/2

[SWF-GigabitEthernet2/0/2]undostp

[SWF-GigabitEthernet2/0/2]port link-type trunk

[SWF-GigabitEthernet2/0/2]port trunk permit vlan all

[SWF] rrpp domain 1

[SWF-rrpp-domainl]control-vlan 4092

[SWF-rrpp-domain1]protected-vlan reference-instance 0 to 32

[SWF-rrpp-domain1] ring 3 node-mode master primary-port gigabitethernet2/0/1 secondary-port gigabitethernet 2/0/2 level 1

[SWF-rrpp-domain1]ring 3 enable

[SWF]rrpp enable

(7) After the configuration is complete, check the RRPP status on the SWA as shown below.

```
[SWA]dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092     Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Master
Ring State      : Complete
Enable Status   : Yes     Active Status: Yes
Primary port    : GigabitEthernet2/0/1       Port status: UP
Secondary port  : GigabitEthernet2/0/2       Port status: BLOCKED
```

Based on the information above, SWA is the primary node of RRPP Ring1, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/2.

(8) After completing the configuration, check the RRPP status on SWB, as shown below.

```
<SWB>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1          Port status: UP
Secondary port  : GigabitEthernet2/0/2          Port status: UP

Ring ID         : 2
Ring Level      : 1
Node Mode       : Edge
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Common port     : GigabitEthernet2/0/1          Port status: UP
                  GigabitEthernet2/0/2          Port status: UP
Edge port       : GigabitEthernet2/0/3          Port status: UP

Ring ID         : 3
Ring Level      : 1
Node Mode       : Edge
Ring State      : -

  Enable Status   : Yes      Active Status: Yes
  Common port     : GigabitEthernet2/0/1          Port status: UP
                    GigabitEthernet2/0/2          Port status: UP
  Edge port       : GigabitEthernet2/0/4          Port status: UP
```

From the information above, it can be seen that SWB serves as both the transmission node of the main ring and the edge node of the sub-ring, with GigabitEthernet2/0/3 and GigabitEthernet2/0/4 functioning as edge ports.

(9) After the configuration is complete, check the RRPP status on SWC as shown below.

```
<SWC>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: UP

Ring ID         : 2
Ring Level      : 1
Node Mode       : Assistant-edge
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Common port     : GigabitEthernet2/0/1        Port status: UP
                  GigabitEthernet2/0/2        Port status: UP
Edge port       : GigabitEthernet2/0/3        Port status: UP

Ring ID         : 3
Ring Level      : 1
Node Mode       : Assistant-edge
Ring State      : -
Enable Status   : Yes      Active Status: Yes
Common port     : GigabitEthernet2/0/1        Port status: UP
                  GigabitEthernet2/0/2        Port status: UP
Edge port       : GigabitEthernet2/0/4        Port status: UP
```

Based on the information above, it can be seen that SWC serves as the transmission node for the main ring and as an auxiliary edge node for the sub-ring. GigabitEthernet2/0/3 and GigabitEthernet2/0/4 are edge ports.

(10) After completing the configuration, check the RRPP status on SWD as shown below.

```
<SWD>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 32
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 1
Ring Level      : 0
Node Mode       : Transit
Ring State      : -
Enable Status   : Yes      Active Status: Yes


Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: UP
```

From the above information, it can be seen that SWD is the transmission node for RRP PRing1, with the primary port being GigabitEthernet2/0/1 and the secondary port being GigabitEthernet2/0/2.

(11) After completing the configuration, check the RRPP status on SWE, as shown below.

```
<SWE>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 16
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 2
Ring Level      : 1
Node Mode       : Master
Ring State      : Complete
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: BLOCKED
```

From the information above, it can be seen that SWE is the primary node of Subring 2, with GigabitEthernet2/0/1 as the primary port and GigabitEthernet2/0/2 as the secondary port.

(12) After the configuration is complete, check the RRPP status on SWF as shown below.

```
<SWF>dis rrpp verbose domain 1
Domain ID       : 1
Control VLAN    : Major 4092      Sub 4093
Protected VLAN  : Reference Instance 0 to 16
Hello Timer     : 1   sec   Fail Timer : 3   sec

Ring ID         : 3
Ring Level      : 1
Node Mode       : Master
Ring State      : Complete
Enable Status   : Yes      Active Status: Yes
Primary port    : GigabitEthernet2/0/1        Port status: UP
Secondary port  : GigabitEthernet2/0/2        Port status: BLOCKED
```

Based on the information above, SWF serves as the main node for Sub-Ring 3, with GigabitEthernet2/0/1 as the primary port and GigabitEthernet2/0/2 as the secondary port.

Chapter Four: Backup Technology for Equipment

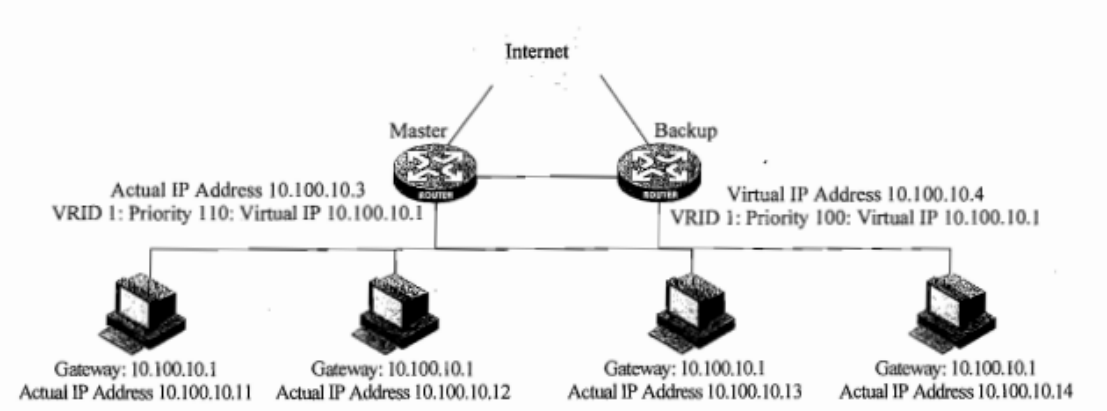
4.1 Overview of Equipment Backup Technology

Distributed network systems encounter various types of failures, and risks are unavoidable. Equipment failure is a common issue within these systems. The simplest way to mitigate equipment failures is through redundant design. By providing backups for the equipment itself and between devices, the impact of failures on user operations can be minimized. Equipment self-backup technology primarily refers to the redundant design of the equipment itself.

High-end switches from H3C support dual main control board failover technology. The two main control boards serve as a primary board and a standby board, with the primary board handling normal operations while the standby board remains in a hot standby state. If the primary board fails and cannot function properly, the standby board can switch states in a very short time, ensuring that operations are disrupted as little as possible. This primary-standby backup is applied to the main control boards of distributed network products to enhance the reliability of network equipment.

Key components of H3C high-end switches, such as the main control board, switching network board, and power system, support redundant hot backup. The AC/DC power supply employs N+1 redundancy to ensure the system operates normally, while the fan system features a 1:1 hot backup and provides automatic speed adjustment based on temperature.

VRRP (Virtual Router Redundancy Protocol) is a fault-tolerant protocol that ensures when the next-hop device of a host fails, another device can promptly take over, thereby maintaining the continuity and reliability of communication.



VRRP allows routers capable of serving as gateways to be added to a VRRP group, forming a virtual router. A VRRP group consists of one Master device and several Backup devices, with the Master performing the actual forwarding function. If the Master fails, one of the Backup devices takes over as the new Master, assuming its role.

4.2 Introduction to VRRP

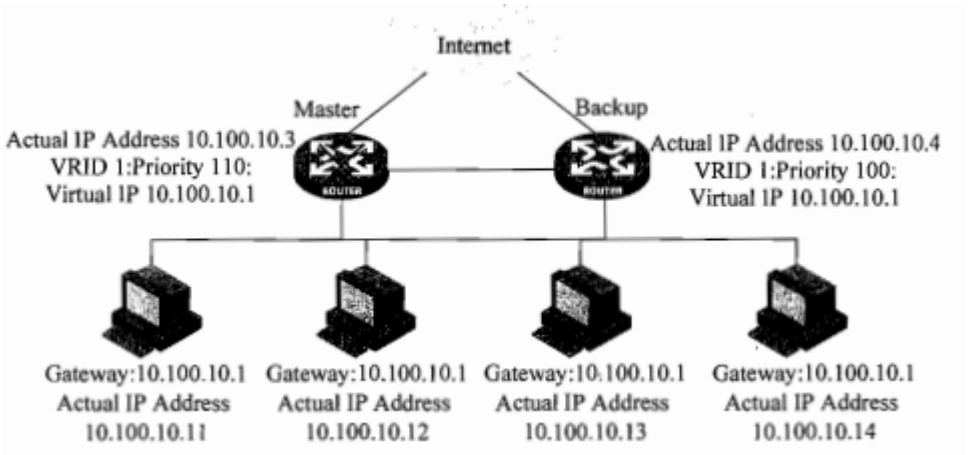
Generally speaking, all hosts within the same subnet are configured with a default route that points to the gateway as the next hop. When the gateway fails, all hosts in this subnet that have the gateway as their default route will be unable to communicate with external networks. Virtual Router Redundancy Protocol (VRRP) helps to prevent network interruptions caused by a single point of failure at the local area network gateway.

4.2.1 Background of VRRP

Typically, all hosts in the same subnet set a common default route pointing to the gateway. Packets directed to other subnets are sent through the default route to the gateway, which then forwards them to enable communication with external networks. If the gateway fails, all hosts relying on this gateway as their default route will lose their ability to communicate with external networks.

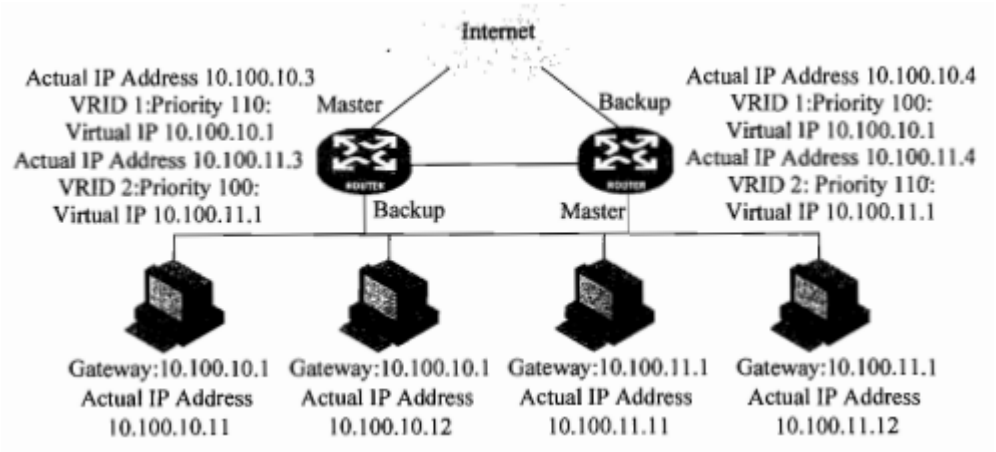
4.2.2 Application of VRRP

VRRP allows multiple routers to be grouped into a backup group, creating a virtual router. In the VRRP master-backup configuration, only the Master router acts as the gateway. If the Master router fails, other Backup routers will elect one to take over the Master's role, as illustrated below. As long as there is at least one functioning router in the backup group, the virtual router remains operational, thus avoiding network interruptions caused by the gateway's single point of failure.



The primary backup method requires only one backup group, where different routers have varying priorities within that group, with the router holding the highest priority becoming the Master router.

The VRRP load balancing method involves multiple routers sharing the traffic simultaneously, thus requiring two or more backup groups. Each backup group contains one Master router and several Backup routers, with different Master routers for each group. A single router can be part of multiple VRRP backup groups, holding different priorities in each group.



As shown in the diagram above, to achieve load balancing of business traffic between routers, the default gateway for hosts within the local area network must be set to different virtual routers. When configuring priorities, it is important to ensure that the VRRP priorities of the routers in the backup group are cross-correlated.

4.2.3 How VRRP Works

4.2.3.1 VRRP Standard Protocol

As a fault-tolerant protocol, VRRP simplifies the configuration of hosts while enhancing reliability. VRRP messages are sent using a fixed multicast address of 224.0.0.18. In local area networks (such as Ethernet) with multicast or broadcast capabilities, VRRP ensures a reliable default link even when a specific router fails, effectively preventing network interruptions caused by a single link failure without requiring modifications to dynamic routing protocols or routing discovery configurations.

Key terms related to VRRP include:

(1) VRRP Backup Group: A group of VRRP-enabled routers within a local area network is designated as a backup group, functioning as a virtual router. Backup groups can be classified as single or multiple backup groups.

(2) Virtual Router Identifier (VRID): Ranging from 1 to 255, this identifier is user-configurable to distinguish different backup groups. Routers with the same VRID form a VRRP backup group.

(3) Master and Backup Routers: The Master router is elected from all routers in the backup group based on priority, serving as the gateway. Other routers act as Backup routers.

(4) IP Address Owner: A router whose interface IP address matches the virtual IP address is referred to as the IP Address Owner.

(5) Virtual MAC Address: A virtual router is assigned a virtual MAC address. As specified by RFC2338, the format of the virtual MAC address is 00-00-5E-00-01-{VRID}. When the virtual router responds to an ARP request, it responds with the virtual MAC address instead of the interface's actual MAC address.

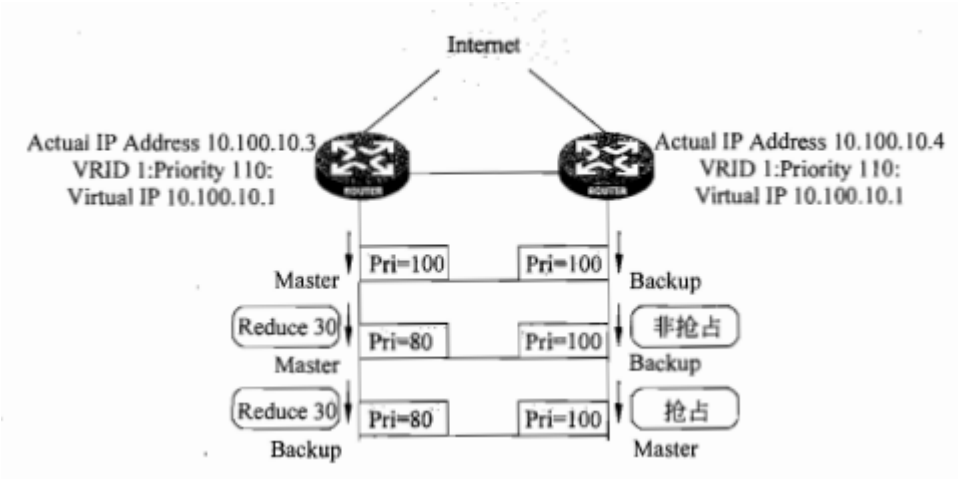
(6) Priority: In VRRP, the status of each router in the backup group is determined by its priority. The router with the highest priority in the group becomes the Master router. When router priorities are equal, the primary IP addresses of the interfaces are compared; the one with the higher IP address has a higher priority. Priority values range from 0 to 255, with higher values indicating higher priority. The default priority is 100, but it can be configured to range from 1 to 254. A priority of 0 is reserved for system use, while 255 is reserved for the IP Address Owner.

(7) Preemption Mode: If a router in the backup group operates in preemption mode, it will send a VRRP advertisement when it detects that its priority is higher than that of the current Master router, prompting a reelection of the Master router. Consequently, the original Master router will become a Backup router.

(8) Non-Preemption Mode: If a router operates in non-preemption mode, it will not become the Master router even if a Backup router is configured with a higher priority, as long as the Master router is functioning.

(9) Authentication Type: VRRP defines three types of authentication: No Authentication, Simple Clear Text Passwords, and MD5 Authentication.

4.2.3.2 VRRP Working Procedure



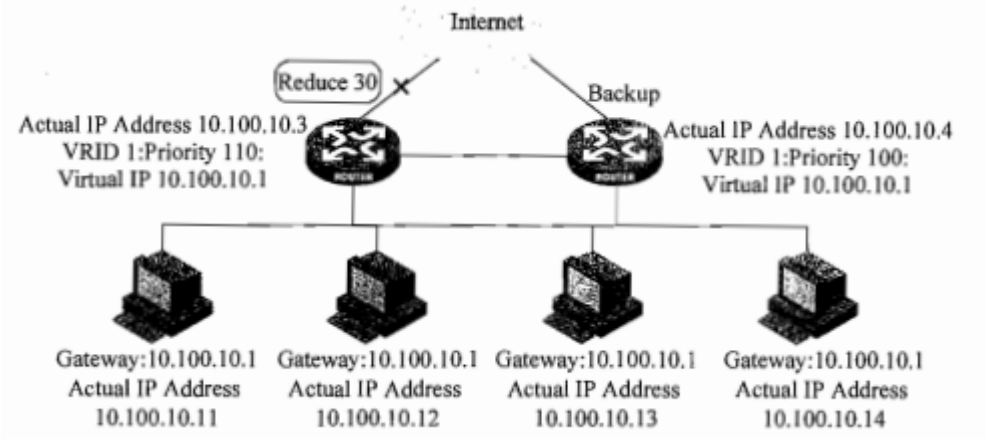
As shown in the diagram above, after the router enables the VRRP function, it determines its role within the backup group based on its priority. The router with the highest priority becomes the Master router, while those with lower priorities become Backup routers. The Master router periodically sends VRRP advertisement messages to inform the other routers in the backup group that it is functioning properly; the Backup routers then start a timer and wait for the arrival of these advertisement messages.

In non-preemption mode, as long as the Master router is functioning, the routers in the backup group will always remain in either the Master or Backup state. Even if a Backup router is later configured with a higher priority, it will not become the Master router. In preemption mode, when a Backup router receives a VRRP advertisement message, it compares its own priority with the priority specified in the advertisement. If its priority is higher, the router will become the Master; otherwise, it will remain a Backup.

If the timer of a Backup router times out and it has still not received VRRP advertisement messages from the Master router, it assumes that the Master router is no longer functioning properly. At this point, the Backup router considers itself to be the Master and sends out VRRP advertisement messages. The routers in the backup group will then elect a new Master router based on priority, taking over the message forwarding function.

4.2.3.3 VRRP Monitoring Interface Function

The VRRP monitoring interface function enhances the backup capabilities significantly.

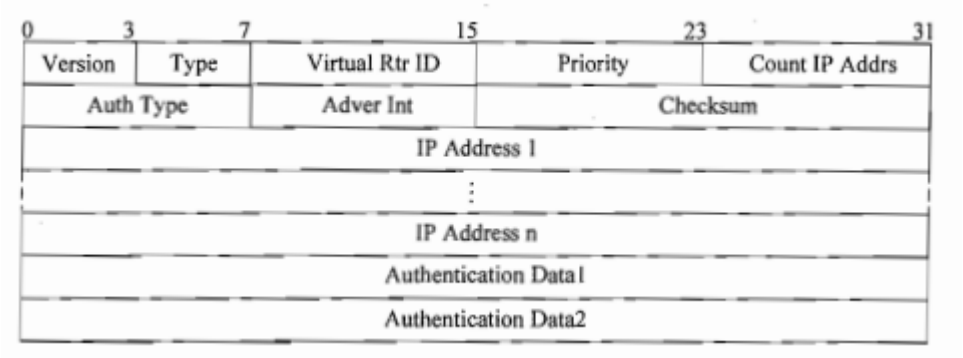


As shown in the figure above, the VRRP backup group is unable to detect the failure of the upstream link. When the router connected to the upstream link experiences a failure and is in the Master state, it will result in hosts within the local area network being unable to access the external network or accessing it through a non-optimal path.

If the router is configured to monitor a specified interface, when the interface connected to the upstream link is in a DOWN or Removed state, the router will proactively lower its priority. This will allow other routers in the backup group to have a higher priority than this router, enabling the router with the highest priority to become the Master and take on the forwarding task.

4.2.4 VRRP Messages and State Machine

1. VRRP Messages



As shown in the diagram above, the fields of the VRRP message are explained as follows:

(1) Version: Protocol version number; the version number for VRRPv2 is 2.

(2) Type: The type of VRRP message. There is only one type of VRRPv2 message, which is the VRRP Advertisement message, and this field has a value of 1.

(3) Virtual Router ID (VRID): The virtual router number (i.e., backup group number) with a value range of 1 to 255.

(4) Priority: The priority of the router within the backup group, with a value range of 0 to 255; a higher value indicates a higher priority.

(5) Count IP Addrs: The number of virtual IP addresses in the backup group. One backup group can correspond to multiple virtual IP addresses.

(6) Auth Type: Authentication type. A value of 0 indicates no authentication, 1 indicates simple character authentication, and 2 indicates MD5 authentication.

(7) Adver Int: The interval for sending advertisement messages, measured in seconds, with a default of 1 second.

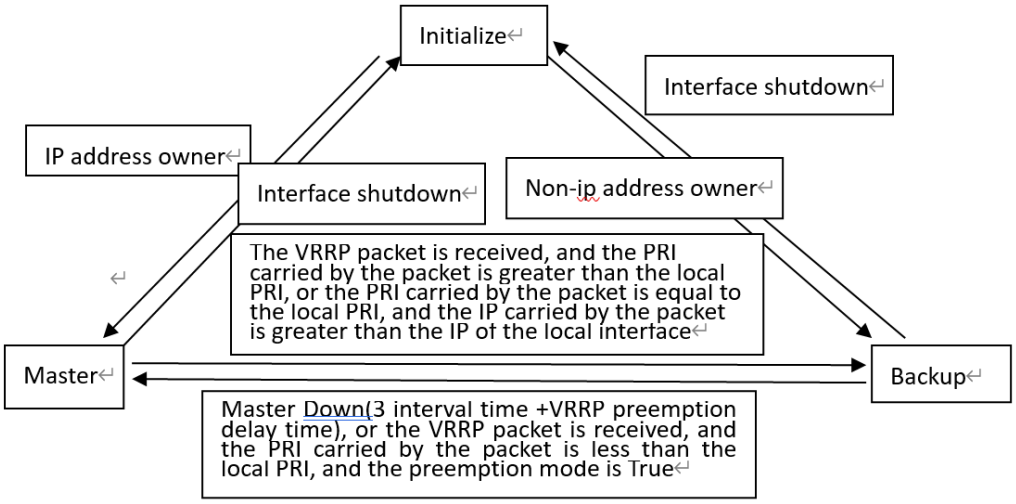
(8) Checksum: A 16-bit checksum used to detect data corruption in the VRRP message.

(9) IP Address: The entry for the backup group's virtual IP address. The number of addresses included is defined in the Count IP Addrs field.

(10) Authentication Data: The authentication data, currently only used for simple character authentication; fill with 0 for other authentication methods.

2. VRRP States

There are three states in VRRP: Initialize, Master, and Backup.



As shown in the figure above, the router enters the Initialize state after booting up. Upon receiving a Startup message from the interface, the router will transition to the Backup or Master state (where the priority is 255). While in the Initialize state, the router will not process VRRP packets.

When the router is in the Master state, it regularly sends VRRP broadcast packets, responds to ARP requests for the virtual IP address using the virtual MAC address, rather than the physical MAC address of the interface. The destination MAC address in the IP packets will be the virtual MAC address. If it is the owner of the virtual IP address, it will accept the IP packets destined for that address; otherwise, it will discard them. In the Master state, the router will only switch to Backup if it receives a VRRP packet with a priority higher than its own or a VRRP packet with a priority equal to its local priority while the packet's interface IP is greater than its local interface IP. When the router receives a Shutdown event from the interface, it will revert to the Initialize state.

When the router is in the Backup state, it receives the VRRP broadcast packets sent by the Master, does not respond to ARP requests for the virtual IP address, and discards IP packets with a destination MAC address that is the virtual MAC address. It will also discard IP packets with a destination IP address that is the virtual IP address. The Backup will only transition to Master when the Master_Down timer expires. If the router receives a VRRP packet with a priority lower than its own, it will discard that packet without resetting the timer. After several such instances, when the Master_Down timer expires, the router will change to Master state. Upon receiving a Shutdown event from the interface, it will revert to the Initialize state.

4.2.5 Configuring VRRP

4.2.5.1 VRRP Configuration Commands

The steps to configure VRRP are as follows:

(1) In the system view, configure the correspondence between the virtual IP address and the MAC address. This step is optional. The configuration command is:

vrrp method { real-mac |virtual-mac}

After configuring the correspondence between the backup group's virtual IP address and MAC address, the Master router will use the defined MAC address as the source MAC address for outgoing packets. This allows hosts on the internal network to learn the mapping between the IP address and MAC address, enabling them to correctly forward packets destined for other subnets to the Master router.

There are two types of correspondence for the virtual IP address and MAC address:

1. The virtual IP address corresponds to a virtual MAC address. By default, when a backup group is created, the router automatically generates a corresponding virtual MAC address, pairing it with the virtual IP address. If this mapping is used, hosts on the internal network do not need to update their IP address and MAC address bindings when the Master router changes.

2. The virtual IP address corresponds to the actual MAC address of an interface. When there is an IP address owner in the backup group, configuring the virtual IP address to correspond with a virtual MAC address can lead to a scenario where one IP address is associated with two MAC addresses. Therefore, users can configure the backup group's virtual IP address to correspond to the actual MAC address, allowing packets sent by hosts to be forwarded to the IP address owner using the actual MAC address.

(2) In the interface view, create a backup group and configure the virtual IP address for the group. The configuration command is:

vrrpvrid virtual-router-id **virtual-ip** virtual-address

When creating a VRRP backup group, users need to configure the virtual IP address for the backup group. If the interface is connected to multiple subnets, it is possible to configure multiple virtual IP addresses for a single backup group, allowing for redundancy among routers in different subnets. When specifying the first virtual IP address for the backup group, the VRRP backup group is automatically generated. If users later assign additional virtual IP addresses to this backup group, those addresses will simply be added to the list of virtual IP addresses in the backup group.

（3）In the interface view, configure the priority of the router within the backup group. The configuration command is:

vrrpvridvirtual-router-id **priority** priority-talue

（4）In the interface view, configure the routers in the backup group to operate in a preemptive mode and set the preemption delay time. This step is optional. The configuration command is:

vrrpvrid virtual-router-id **preempt-mode** [**timer delay** delay-value]

The priority of the IP address owner is always set to 255 and does not require user configuration. The IP address owner operates in a preemptive manner and does not

allow the configuration of monitoring for specified interfaces or Track items. To configure VRRP monitoring of specified interfaces, use the command in the interface view.

vrrpvrid virtual-router-id **trackinterface** interface-type interface-number [**reduced** priority-reduced]

When the status of the monitored interface changes from DOWN or Removed to UP, the priority of the corresponding router will automatically revert. Different backup groups on a single interface can be set with different authentication methods and passwords; however, members within the same backup group must use the same authentication method and password. To configure the authentication for sending and receiving VRRP messages in the backup group, use the command in the interface view:

vrrpvrid*virtual-router-id***authentication-mode**{md5 | **simple**} key

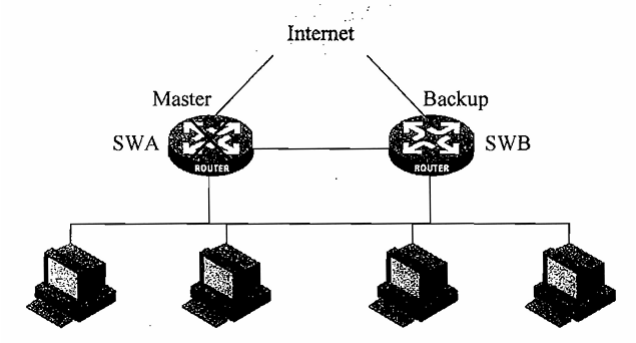
Excessive network traffic or discrepancies in timers across different routers can cause the backup router's timer to time out abnormally, leading to a state transition. To address this issue, you can extend the interval for sending VRRP advertisement packets. To configure the interval at which the Master router in the backup group sends these VRRP advertisement packets, use the following command in the interface view.

vrrpvrid*virtual-router-id***timer advertise** adver-interual

To display the status information of the VRRP backup group, use the following command：

display vrrp[**verbose**][**interface***interface-type interface-number* [**vrid** virtual-router-id]]

4.2.5.2 VRRP Network Experiment-1



As shown in the diagram above, SWA and SWB are interconnected via their respective Ethernet ports. SWA and SWB belong to backup group 1 with a virtual IP address of 192.168.0.254/24. When SWA is functioning normally, local area network traffic is forwarded through SWA; when SWA fails, local area network traffic is forwarded through SWB. The specific configuration steps are as follows:

(1) Configure SWA.

```
[SWA]vlan 10
[SWA]interface Vlan-interface 10
[SWA-Vlan-interface10]ip add 192.168.0.252 255.255.255.0
[SWA-Vlan-interface10]vrrpvrid 1 virtual-ip 192.168.0.254
[SWA-Vlan-interface10]vrrpvrid 1 priority 120
[SWA-Vlan-interface10]vrrpvrid 1 preempt-mode
```

(2) Configure SWB.

```
[SWB]vlan 10
[SWB]interface Vlan-interface 10
[SWB-Vlan-interface10]ip add 192.168.0.253 255.255.255.0
[SWB-Vlan-interface10]vrrpvrid 1 virtual-ip 192.168.0.254
[SWB-Vlan-interface10]vrrpvrid 1 preempt-mode
```

(3) After completing the configuration, check the VRRP status on the SWA as shown below:

```
<SWA>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 1
  Interface     : Vlan-interface10
  VRID          : 1                 Adver. Timer  : 1
  Admin Status  : UP                State         : Master
  Config Pri    : 120               Run Pri       : 120

  Preempt Mode  : YES               Delay Time    : 5
  Auth Type     : NONE
  Virtual IP    : 192.168.0.254
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 192.168.0.252
```

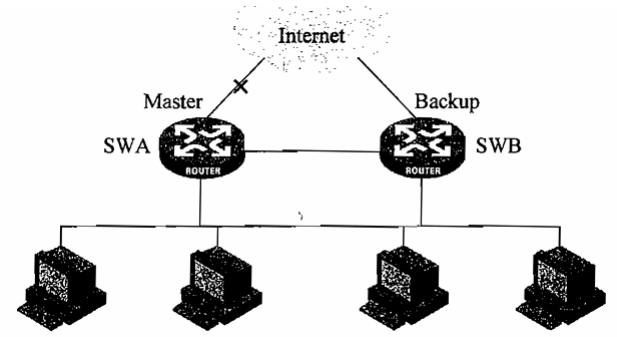
From the information above, it can be seen that SWA is the Master device for VRRP Backup Group 1, and the virtual IP address for Backup Group 1 is 192.168.0.254/24.

(4) After configuration is completed, check the VRRP status on SWB, as shown below.

```
<SWB>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 1
  Interface     : Vlan-interface10
  VRID          : 1                 Adver. Timer  : 1
  Admin Status  : UP                State         : Backup
  Config Pri    : 100               Run Pri       : 100
  Preempt Mode  : YES               Delay Time    : 5
  Auth Type     : NONE
  Virtual IP    : 192.168.0.254
  Master IP     : 192.168.0.252
```

From the above information, it can be seen that SWB is the backup device for VRRP backup group 1, and the virtual IP address for backup group 1 is 192.168.0.254/24.

4.2.5.3 VRRP Network Experiment-2



As shown in the figure above, SWA and SWB are connected to each other through their respective Ethernet ports. SWA and SWB belong to Backup Group 1, which has a virtual IP address of 192.168.0.254/24. When SWA is functioning normally, local area network (LAN) traffic is forwarded through SWA; if the upstream link of SWA becomes unavailable, the LAN traffic is forwarded through SWB. The specific steps for configuring the experiment are as follows:

(1) Configure SWA.

```
[SWA]vlan 100
[SWA-vlan100]portGigabitEthernetl/0/1
[SWA]int vlan 100
[SWA-Vlan-interfacel00]ip add 192.168.255.1 255.255.255.252
[SWA]vlan 10
[SWA]interface Vlan-interface 10
[SWA-Vlan-interfacel0]ip add 192.168.0.252 255.255.255.0
[SWA-Vlan-interfacel0]vrrpvrid l virtual-ip 192.168.0.254
[SWA-Vlan-interfacel0]vrrpvrid l track interface Vlan-
interface 100 reduced 30
[SWA-Vlan-interfacel0]vrrpvrid l priority 120
[SWA-Vlan-interfacel0]vrrpvrid 1 preempt-mode
```

(2) Configure SWB.

```
[SWB]vlan 101
[SWB-vlan101]port GigabitEthernet1/0/1
[SWB]interface vlan 101
[SWB-Vlan-interface101]ip add 192.168.255.5 255.255.255.252
[SWB]vlan 10
[SWB]interface Vlan-interface 10
[SWB-Vlan-interfacel0]ip add 192.168.0.253 255.255.255.0
[SWB-Vlan-interfacel0]vrrpvrid 1 virtual-ip 192.168.0.254
[SWB-Vlan-interfacel0]vrrpvrid 1 preempt-mode
```

(3) After the configuration is complete, disable the upstream link on the SWA.

```
[SWA]interface GigabitEthernet1/0/1
[SWA-GigabitEthernet1/0/1]shutdown
```

(4) Check the VRRP status on the SWA, as shown below.

```
<SWA>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 1
  Interface     : Vlan-interface10
  VRID          : 1                 Adver. Timer  : 1
  Admin Status  : UP                State         : Backup
  Config Pri    : 120               Run Pri       : 90
  Preempt Mode  : YES               Delay Time    : 5
  Auth Type     : NONE
  Virtual IP    : 192.168.0.254
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 192.168.0.253
```

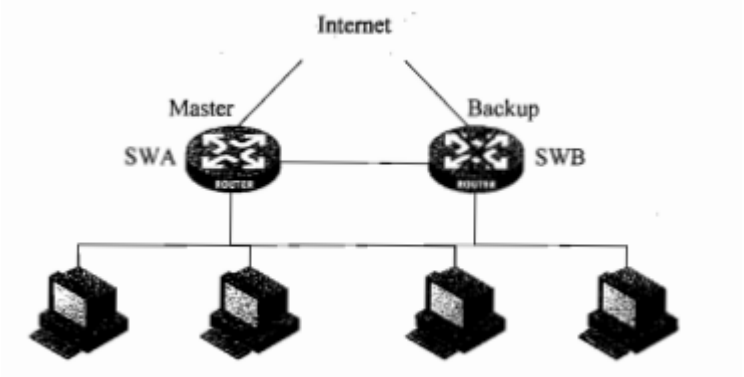
It can be seen that SWA serves as the Backup device for VRRP Backup Group 1, with the virtual IP address of Backup Group 1 set to 192.168.0.254/24.

(5) After completing the configuration, check the VRRP status on SWB, as shown below.

```
<SWB>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 1
  Interface     : Vlan-interface10
  VRID          : 1                 Adver. Timer  : 1
  Admin Status  : UP                State         : Master
  Config Pri    : 100               Run Pri       : 100
  Preempt Mode  : YES               Delay Time    : 0
  Auth Type     : NONE
  Virtual IP    : 192.168.0.254
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 192.168.0.253
```

It is clear that SWB is the Master device for VRRP backup group 1, which utilizes a virtual IP address of 192.168.0.254/24.

4.2.5.4 VRRP Network Experiment-3



As shown in the figure above, SWA and SWB are interconnected through their respective Ethernet ports. SWA and SWB belong to Backup Group 1 with a virtual IP address of 192.168.0.254/24, and Backup Group 2 with a virtual IP address of 192.168.1.254/24. When both SWA and SWB are functioning normally, traffic from VLAN10 in the local area

network is forwarded through SWA, while traffic from VLAN20 is forwarded through SWB. If SWB fails, both the VLAN10 and VLAN20 traffic in the local area network will be forwarded through SWA. The specific configuration steps are as follows:

(1) Configure SWA.

```
[SWA]vlan 10
[SWA]interface Vlan-interface 10
[SWA-Vlan-interface10]ip add 192.168.0.252 255.255.255.0
[SWA-Vlan-interface10]vrrpvrid 1 virtual-ip 192.168.0.254
[SWA-Vlan-interface10]vrrpvrid 1 priority 120
[SWA-Vlan-interface10]vrrpvrid 1 preempt-mode
[SWA]vlan 20
[SWA]interface Vlan-interface 20
[SWA-Vlan-interface20]ip add 192.168. 1.252 255.255.255.0
[SWA-Vlan-interface20]vrrpvrid 2 virtual-ip 192.168.1.254
[SWA-Vlan-interface20]vrrpvrid 2 priority 100
[SWA-Vlan-interface20]vrrpvrid 2 preempt-mode
```

(2) Configure SWB.

```
[SWB]vlan 10
[SWB]interface Vlan-interface 10
[SWB-Vlan-interface10]ip add 192.168.0.253 255.255.255.0
[SWB-Vlan-interface10]vrrpvrid 1 virtual-ip 192.168.0.254
[SWB-Vlan-interface10]vrrpvrid 1 priority 100
[SWB-Vlan-interface10]vrrpvrid 1 preempt-mode
[SWB]vlan 20
[SWB]interface Vlan-interface 20
[SWB-Vlan-interface20]ip add 192.168. 1.253 255.255.255.0
[SWB-Vlan-interface20]vrrpvrid 2 virtual-ip 192.168.1.254
[SWB-Vlan-interface20]vrrpvrid 2 priority 120
[SWB-Vlan-interface20]vrrpvrid 2 preempt-mode
```

(3) Check the VRRP status on SWA, as shown below.

```
<SWA>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 2
  Interface     : Vlan-interface10
  VRID          : 1                 Adver. Timer   : 1
  Admin Status  : UP                State          : Master
```

```
Config Pri      : 120                   Run Pri         : 120
Preempt Mode    : YES                   Delay Time      : 0
Auth Type       : NONE
Track IF        : Vlan100               Pri Reduced     : 30
Virtual IP      : 192.168.0.254
Virtual MAC     : 0000-5e00-0101
Master IP       : 192.168.0.252

Interface       : Vlan-interface20
VRID            : 2                     Adver. Timer    : 1
Admin Status    : UP                    State           : Backup
Config Pri      : 100                   Run Pri         : 100
Preempt Mode    : YES                   Delay Time      : 0
Auth Type       : NONE
Virtual IP      : 192.168.1.254
Master IP       : 192.168.1.253     IPv4 Standby Information:
Run Method      : VIRTUAL-MAC
```

It can be seen that SWA is the Master device for VRRP Backup Group 1, with a virtual IP address of 192.168.0.254/24. SWA is also the Backup device for VRRP Backup Group 2, which has a virtual IP address of 192.168.1.254/24.

(4) After the configuration is complete, check the VRRP status on SWB, as shown below.

```
<SWB>dis vrrp verbose
  IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
  Total number of virtual routers: 2
  Interface     : Vlan-interface10
  VRID          : 1                     Adver. Timer    : 1
  Admin Status  : UP                    State           : Backup
  Config Pri    : 100                   Run Pri         : 100
  Preempt Mode  : YES                   Delay Time      : 0
  Auth Type     : NONE
  Virtual IP    : 192.168.0.254
  Master IP     : 192.168.0.252

  Interface     : Vlan-interface20
  VRID          : 2                     Adver. Timer    : 1
  Admin Status  : UP                    State           : Master
  Config Pri    : 120                   Run Pri         : 120
  Preempt Mode  : YES                   Delay Time      : 0
  Auth Type     : NONE
  Virtual IP    : 192.168.1.254
  Virtual MAC   : 0000-5e00-0102
  Master IP     : 192.168.1.253         IPv4 Standby Information:
  Run Method    : VIRTUAL-MAC
```

SWB serves as the Backup device for VRRP backup group 1 and as the Master device for VRRP backup group 2. The virtual IP address for backup group 1 is 192.168.0.254/24, while the virtual IP address for backup group 2 is 192.168.1.254/24.

Chapter Five: Stacking Technology

5.1 Overview of Stacking Technology

IRF (Intelligent Resilient Framework) is an enhanced stacking technology that introduces innovations and enhancements in areas such as high reliability and redundant backups.

IRF stacking allows for global cross-device link aggregation, providing comprehensive link-level protection. Additionally, it implements three-layer routing redundancy across devices, supporting distributed processing for multiple unicast and multicast routing protocols, as well as hot backup technologies for various routing protocols.

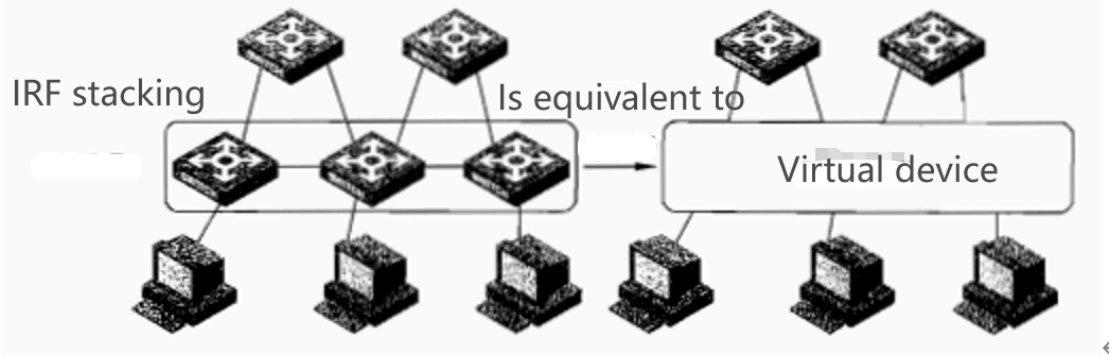
IRF stacking achieves distributed operation of Layer 2 protocols within the Fabric, improving the utilization and reliability of units within the stack while reducing inter-device protocol dependencies.

All individual devices within an IRF stack are referred to as member devices. The physical stacking ports between member devices support aggregation, and the physical connections between the stacking system and upstream/downstream devices also support aggregation, thereby enhancing the reliability of the stacking system through multi-link backups.

IRF employs 1:N redundancy, where the Master handles the traffic, while the Slave acts as its backup, remaining in sync with the Master at all times. If the Master encounters a fault, IRF will select one of the Slaves to become the new Master. Due to strict configuration and data synchronization protocols maintained during the operation of the stacking system, the new Master can seamlessly manage and operate the IRF stack without impacting existing network functions and services. Moreover, the presence of multiple Slave devices further enhances system reliability.

When the IRF member devices are modular distributed devices, they possess multiple master control boards and multiple interface boards. For stacking modular distributed devices, IRF does not forego the redundancy protection provided by the primary and backup master control boards of each modular distributed member device, despite IRF technology's backup capabilities. Instead, it unifies the management of primary and backup master control boards from each member device as a shared resource, further improving system reliability.

5.1.1 Functions of IRF Stacking



As shown in the figure above, all individual devices in an IRF stack are referred to as member devices. IRF connects multiple member devices through stacking ports to form a virtual "logical device." This integration allows member devices to function as a single entity in both management and usage. IRF offers the following features:

(1) Simplified management. Once the IRF stack is formed, it can be managed as a unified logical device, allowing administrators to log in to manage the entire IRF stack and all its member devices without needing to connect to each member device individually for configuration and management.

(2) Enhanced performance. The various control protocols running on the logical device created by IRF operate as a single entity. For instance, the IRF stack runs routing protocols and computes routing tables as if it were a single device, eliminating the need for member devices to run spanning tree protocols. This reduces the volume of protocol messages exchanged between devices and shortens convergence times.

(3) Flexible scalability. IRF allows for elastic scaling based on demand, ensuring user investment is protected. New devices can be added to or removed from the IRF architecture with "hot-swappable" capabilities, allowing for uninterrupted operation of other devices.

(4) High reliability. The high reliability of IRF is reflected in three aspects: links, devices, and protocols. The physical ports interconnecting member devices support aggregation, and the physical connections between the IRF stack and upstream/downstream devices also support aggregation. This multi-link backup enhances link reliability. The IRF stack can quickly detect failures in internal member devices and respond promptly, ensuring uninterrupted operation of the entire stack.

5.1.2 Basic Concepts of IRF Stacking

Commonly used concepts in IRF stacking technology include:

(1) Master: A type of member device elected by a role to manage the entire stack. Only one member device can be designated as the Master at any given time in a stack.

(2) Slave: A type of member device also elected by a role, subordinate to the Master device, and operates as a backup. All devices in the stack, except the Master, are Slave devices. Multiple Slave devices may exist within a stack.

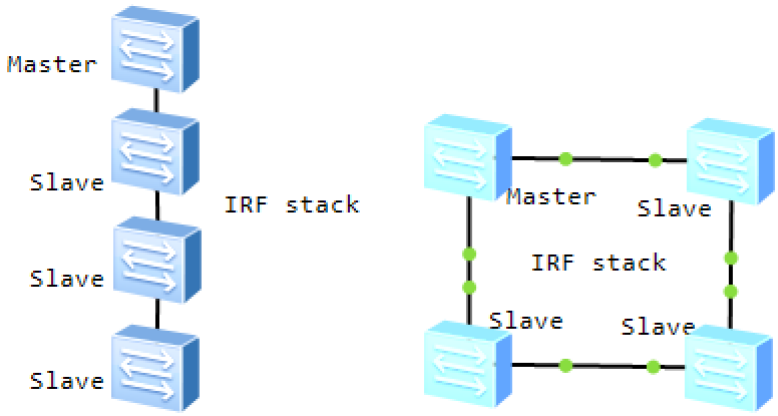
(3) Physical Stacking Ports: For IRF to function properly, member devices must be physically connected. The physical ports used for stacking connectivity on the devices are referred to as physical stacking ports.

(4) Stacking Ports: Physical stacking ports need to be bound with logical stacking ports, which are simply referred to as stacking ports.

(5) Aggregated Stacking Ports: A stacking port can either be bound to a physical stacking port or be formed by aggregating multiple physical stacking ports. A stacking port formed from multiple physical stacking ports is called an aggregated stacking port.

5.1.3 IRF Stacking Physical Topology

There are two types of physical topologies for IRF stacking:



(1) Chain topology. Connect one device's left stacking port (right stacking port) to another device's right stacking port (left stacking port) using stacking cables, and repeat this for the next devices. The right stacking port (left stacking port) of the first device is not connected to the left stacking port (right stacking port) of the last device. This type of connection is also known as chain connection.

(3) Ring topology. Connect the right stacking port (left stacking port) of the first device in the chain topology to the left stacking port (right stacking port) of the last device. This type of connection is also referred to as ring connection.

5.1.4 Formation of IRF Stacking

The operation of IRF stacking is divided into three stages: topology collection, role election, and stack maintenance. Upon device startup, topology collection is conducted first, followed by participation in role election. Only after successful processing can the stacking system form and operate normally.

Each device in the stack collects the overall topology by interacting with directly adjacent member devices through Hello messages. These Hello messages carry topology information, including the stacking port connection relationships, member device IDs, member device priorities, and the MAC addresses of the member bridges.

Each member device locally records its known topology information. Initially, member devices only record their own topology information. Once the stacking port status changes to UP, member devices periodically send out their known topology information from the stacking port. When a member device receives topology information from a directly neighboring device, it updates its locally recorded topology information. After a period of collection, all devices will gather complete topology information, a process referred to as topology convergence. Following topology convergence, the system immediately enters the role election stage.

The stacking system consists of multiple member devices, each having a defined role—Master or Slave. The process of determining the roles of member devices is called role election. Role elections occur when there are changes in the topology, such as stack establishment, the addition of new devices, stack splits, or the merging of two stacks.

The rules for role election are as follows (starting from the first rule, if there are multiple optimal candidates, the next rule is evaluated until a unique optimal member is found, at which point the election stops. This optimal member becomes the Master device of the stack, while the others are designated as Slave devices):

(1) The current Master is preferred over non-Master members.

(2) When all member devices are chassis-based distributed devices, the local active control board is preferred over the local standby control board.

(3) When all member devices are chassis-based distributed devices, the standby control board of the original Master is preferred over the control boards of non-Master members.

(4) Members with higher priorities are preferred.

(5) Members with longer operational times are preferred.

(6) Members with smaller MAC addresses are preferred.

During the role election phase, the Master is also responsible for handling member ID conflicts, software version loading, and stack merging management. After the role election is completed, stack management will proceed to the stack maintenance phase.

5.1.5 IRF Stack Maintenance

The main function of stack maintenance is to monitor the joining and leaving of member devices and to continuously gather new topology information to ensure the normal operation of the stack. During the stack maintenance process, the topology collection continues, and when new member devices join, two different actions are taken based on the status of the newly added device.

(1) If the newly joined device has not formed a stack, it will be designated as a Slave.

(2) If the joining device has already formed a stack, it effectively results in a merge of two stacks. In this case, both stacks will conduct a stack election, and the losing party's member devices will need to reboot and join the winning party as Slave devices.

During the stack maintenance process, there are two ways to determine if a member device has left:

(1) Normally, directly adjacent member devices will periodically exchange Hello messages (typically every 200ms). If no Hello messages are received from direct neighbors for multiple cycles (usually 10 cycles), the member device is considered to have left the stack system, and it will be isolated from the topology.

(2) If a stack port is found to be DOWN, the member device with that port will urgently broadcast a notification to other members in the stack to immediately recalculate the current topology, without waiting for the Hello message timeout to handle it.

When a member device leaves, if it is a Slave device, the system essentially loses a backup control board and the physical resources associated with that board. If it is the Master device that leaves, the stack system will conduct a new election, and the newly elected Master will take over all functions of the original Master.

When a single device leaves the stack, it returns to standalone operation. If multiple interconnected devices leave, they will form two independent stacks, a situation referred to as stack splitting.

5.1.6 IRF Configuration and Maintenance

5.1.6.1 IRF Stacking-Related Commands

In the system view, bind the device's logical stacking ports to the physical stacking ports, while also enabling the stacking feature on the current device. The configuration command is:

```
irf member member-id irf-portirf-port-id port port-list
```

In a stack, members are marked by member numbers, and device numbers are also used in the stack configuration file to differentiate port configurations on various member devices. In the system view, configure the IRF member number. The configuration command is:

```
irf member member-id renumber new-member-id
```

In system view, configure the priority of the specified member device in the stack. The configuration command is:

```
irf member member-id priority-priority
```

If the automatic stacking loading feature is not enabled, when the software version of a device participating in the stack does not match that of the Master device, newly added or lower-priority devices may fail to start correctly. In this case, the user must manually upgrade the device version before adding it to the stack. Once the automatic loading feature is enabled, when member devices join the stack, their software version numbers will be compared with the Master device's. If they do not match, the devices will automatically download the boot file from the Master device, then restart using the new system boot file to rejoin the stack. If the filename of the newly downloaded boot file matches an existing file on the device, the original boot file will be overwritten.

Configuration commands for stack ports and member priority must be followed by a device restart to take effect. After the IRF stack is formed, users can log into the stack system console through the AUX or Console port of any member device. By configuring an IP address on the VLAN interfaces of any member device and ensuring routing is reachable, users can remotely access the stack system via Telnet, Web, or SNMP.

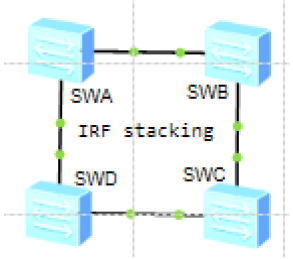
When a user logs into the stack, they are actually logging into the Master device of the stack. The Master is the configuration and control center of the stack system. After configuring on the Master, these settings will be synchronized to the Slave devices. At this point, a redirection to the Slave device is required to log in to it.

In system view, the command to configure redirection to the Slave device is:

```
irf switch-to member-id
```

Only the following commands are permitted on Slave devices: display, quit, return, system-view, debugging terminal debugging, terminal trapping, terminal logging.

5.1.6.2 IRF Stacked Network Experiment



As shown in the diagram above, SWA, SWB, SWC, and SWD form an IRF stack, configured with member numbers 1, 2, 3, and 4, using 10-gigabit modules as the stacking modules. The Port1 of the stacking module is connected to Port2 of another stacking module. The specific configuration steps are as follows:

(1) Configure SWA.

```
[SWA]irf member 1 irf-port 1 port 1
[SWA]irf member 1 irf-port 2 port 2
[SWA]irf member 1 renumber 1
```

(2) Configure SWB.

```
[SWB]irf member 1 irf-port 1 port 1
[SWB]irf member 1 irf-port 2 port 2
[SWB]irf member 1 renumber 2
```

(3) Configure SWC.

```
[SWC]irf member 1 irf-port 1 port 1
[SWC]irf member 1 irf-port 2 port 2
[SWC]irf member 1 renumber 3
```

(4) Configure SWD.

```
[SWD]irf member 1 irf-port 1 port 1
[SWD]irf member 1 irf-port 2 port 2
[SWD]irf member 1 renumber 4
```

Power off the four devices and connect them using the stacking cables as per the network diagram. Then, power them all on to form a stack. SWA, SWB, SWC, and SWD stack together to create a new device labeled as SWA_NEW.

(5) Check the IRF stacking status on SWA_NEW, as shown below.

```
<SWA_NEW>display irf configuration
MemberID    NewID    IRF-Port1    IRF-Port2
 *1          1        1            2
  2          2        1            2
  3          3        1            2
  4          4        1            2

--------------------------------------------------------------------------------
 * indicates the device is the master.
 + indicates the device through which the user logs in.
```

From the above information, it can be seen that SWA is the Master node in the IRF stack, while SWB, SWC, and SWD are the Slave nodes in the IRF stack.

Summary and Outlook

Fault tolerance technology is a critical means to ensure the stability, reliability, and continuity of systems. With the expansion of network scale and the complexity of business demands, its technological development is evolving

towards diversification and intelligence. This article focuses on three major fault tolerance solutions: link backup technology, device backup technology, and stacking technology, in conjunction with current mainstream fault tolerance technologies and actual project requirements. An analysis of their technical characteristics, application status, and future challenges highlights their significant roles and developmental trends in the field of fault tolerance.

Link backup technology is one of the foundational technologies in fault tolerance mechanisms. By configuring a primary path and a backup path, it can quickly switch to the backup link when the primary link fails, ensuring the continuity and stability of communication. Link backups can be either static or dynamic; the former is simple to deploy and suitable for smaller networks, but lacks flexibility. The latter relies on dynamic routing protocols (such as OSPF, BGP, etc.) for automatic path switching, which enhances network resilience and reduces management complexity. As networks become virtualized and scaled up, the advantages of dynamic link backup become increasingly evident, though challenges such as switch delays and insufficient link resource allocation also emerge. In the context of rapid advancements in cloud computing and edge computing, link backup will explore intelligent directions by introducing artificial intelligence and machine learning technologies to construct backup systems capable of predicting link failures and making adaptive adjustments. Additionally, multipath transmission technologies (like MP-TCP) will achieve link load balancing, further enhancing backup efficiency. In the future, link backup will advance towards greater efficiency, real-time capabilities, and intelligence to meet the increasingly demanding fault tolerance requirements in complex networks.

Device backup technology focuses on fault tolerance at the hardware device level, preventing device failures from affecting the overall operation of the network through the deployment of redundant devices. In traditional models, the primary-backup device model and the clustered device model are the main approaches for device backup. The primary-backup device model centers on a "one primary, one backup" structure, activating the backup device when the primary device fails. This is commonly seen in the deployment of critical devices such as routers, firewalls, and switches. Conversely, the clustered device model shares data traffic through multiple devices working in cooperation, achieving redundancy among devices so that when one device fails, others can take over its tasks. Currently, emerging virtualization and containerization technologies offer more flexibility and efficiency for device backup, such as achieving logical backup fault tolerance through virtual machine live migration, thereby reducing reliance on physical devices. Additionally, distributed collaborative backup models have become a research focus, distributing device tasks across distributed nodes and enhancing collaboration capabilities through Software Defined Networking (SDN) and Network Functions Virtualization (NFV) technologies. In the future, device backup technology will develop towards distribution, intelligence, and green economics, particularly by employing low-energy backup strategies to

minimize resource consumption, thus improving the efficiency and sustainability of device-level fault tolerance.

Stacking technology provides solutions for fault tolerance from a system architecture perspective, centering on logically stacking multiple devices into a unified virtual device to achieve resource sharing and task load transfer. Stacking technology boasts easy deployment and cost-effectiveness, making it widely applicable in enterprise campus networks and small to medium-sized business networks. However, as the scale of device stacking increases, the stacked structure inevitably faces management complexity and bottleneck issues in communication links. Particularly, when stacking links fail, it can lead to decreased system performance or even failure propagation. Therefore, future optimizations of stacking technology will focus on large-scale stacking deployment and enhancing stacking efficiency. Through distributed stacking architectures and multi-layer stacking methods, the burden on stacking links can be reduced, improving system resilience. Additionally, the integration with virtualization technology is an important direction for the development of stacking technology. By leveraging SDN technology, physical device stacking logic can be transformed into virtual stacking, simplifying stacking management and enhancing scalability. In the future, stacking technology will advance towards greater flexibility, collaboration, and performance, continuing to play an irreplaceable role in fault tolerance design.

In summary, link backup technology, device backup technology, and stacking technology construct a fault tolerance defense line for modern network systems from different levels, ensuring stability and high availability in complex environments. In future developments, fault tolerance technology will benefit from further advancements in artificial intelligence, big data, SDN, NFV, among others, gradually evolving towards intelligence, multidimensional collaboration, and sustainability. Researching multi-level interlinked fault tolerance mechanisms based on links, devices, and architectures will pave new paths for enhancing fault tolerance capabilities and support the development of more efficient, stable, and reliable networks and computing systems, providing a solid technical foundation for the rapid advancement of information technology.