CCNA notebook

Huy Bui

February 20, 2018

Contents

Chapter 1

LAN Design

1.1 Hierarchical Design Model

A hierarchical LAN design includes the following three layers, as shown in Figure 1.1:

- **Access layer** provides endpoints and users direct access to the network
- **Distribution layer** aggregates access layers and provides connectivity to services.
- **Core layer** provides connectivity between distribution layers for large LAN environments.

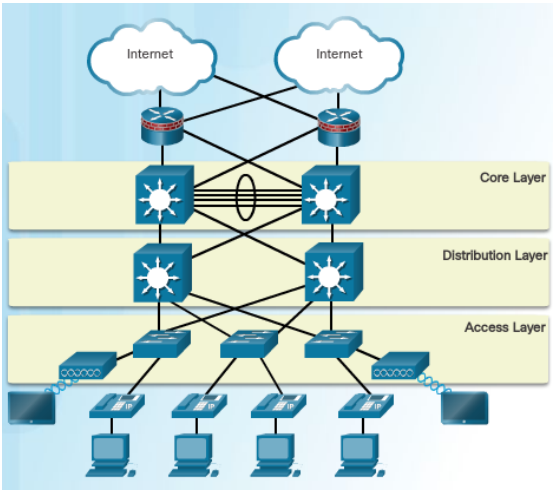

Figure 1.1: Three-layer hierarchical design model

Even though the hierarchical model has three layers, some smaller enterprise networks may implement a two-tier hierarchical design. In a two-tier hierarchical design, the core and distribution layers are collapsed into one layer, reducing cost and complexity, as shown in Figure 1.2.

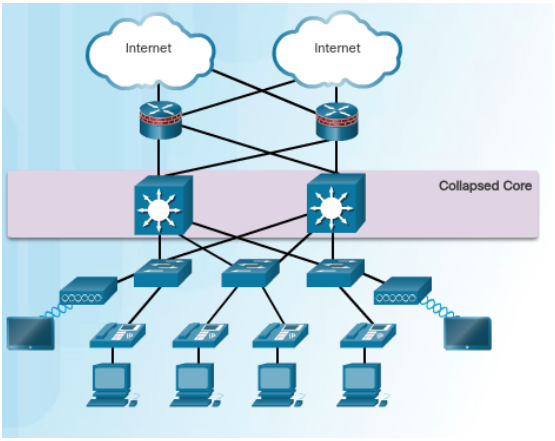

Figure 1.2: Collapsed Core

1.2 Expanding the network

To support a large, medium or small network, the network designer must develop a strategy to enable the network to be available and to scale effectively and easily. Included in a basic network design strategy are the following recommendations:

- Use expandable, modular equipment or clustered devices that can be easily upgraded to increase capabilities.
- Design a hierarchical network to include modules that can be added, upgraded, and modified, as necessary, without affecting the design of the other functional areas of the network.
- Create an IPv4 or IPv6 address strategy that is hierarchical.
- Choose routers or multilayer switches to limit broadcasts and filter other undesirable traffic from the network.

More advanced network design requirements will be described in the following sections.

1.2.1 Planning for redundancy

One method of implementing redundancy is by installing duplicate equipment and providing failover services for critical devices. Another method of implementing redundancy is redundant paths.

1.2.2 Failure domain

A failure domain is the area of a network that is impacted when a critical device or network service experiences problems. The use of redundant links and reliable enterprise-class equipment minimize the chance of disruption in a network. Smaller failure domains reduce the impact of a failure on company productivity.

In the hierarchical design model, it is easiest and usually least expensive to control the size of a failure domain in the distribution layer. In the distribution layer, network errors can be contained to a smaller area; thus, affecting fewer users.

Routers, or multilayer switches, are usually deployed in pairs, with access layer switches evenly divided between them. This configuration is referred to as a building, or departmental, switch block. Each switch block acts independently of the others. As a result, the failure does not affect a significant number of end users.

1.2.3 Increasing Bandwidth

Link aggregation allows an administrator to increase the amount of bandwidth between devices by creating one logical link made up of several physical links. EtherChannel is a form of link aggregation used in switched networks. The EtherChannel is seen as one logical link using an EtherChannel interface. Most configuration tasks are done on the EtherChannel interface, instead of on each individual port, ensuring configuration consistency throughout the links.

1.2.4 Expanding the Access Layer

To communicate wirelessly, end devices require a wireless NIC that incorporates a radio transmitter/receiver and the required software driver to make it operational. Additionally, a wireless router or a wireless access point (AP) is required for users to connect.

1.2.5 Fine-tuning Routing Protocols

Advanced routing protocols, such as OSPF and EIGRP are used in large networks. Link-state routing protocols such as Open Shortest Path First (OSPF) works well for larger hierarchical networks where fast convergence is important. Another popular routing protocol for larger networks is Enhanced Interior Gateway Routing Protocol (EIGRP). Cisco developed EIGRP as a proprietary distance vector routing protocol with enhanced capabilities.

1.3 Selecting network devices

1.3.1 Switch hardwares

There are five categories of switches for enterprise networks:

- **Campus LAN Switches** – To scale network performance in an enterprise LAN, there are core, distribution, access, and compact switches.
- **Cloud-Managed Switches** – The Cisco Meraki cloud-managed access switches enable virtual stacking of switches. They monitor and configure thousands of switch ports over the web, without the intervention of onsite IT staff.

- **Data Center Switches** – A data center should be built based on switches that promote infrastructure scalability, operational continuity, and transport flexibility. The data center switch platforms include the Cisco Nexus Series switches and the Cisco Catalyst 6500 Series switches.
- **Service Provider Switches** – Service provider switches fall under two categories: aggregation switches and Ethernet access switches. Aggregation switches are carrier-grade Ethernet switches that aggregate traffic at the edge of a network. Service provider Ethernet access switches feature application intelligence, unified services, virtualization, integrated security, and simplified management.
- **Virtual Networking** – Cisco Nexus virtual networking switch platforms provide secure multi-tenant services by adding virtualization intelligence technology to the data center network.

There are some terminologies that an administrator to be able to choose the right switch platform:

- **Port density** is the number of ports available on a single switch.
- **Forwarding rates** define the processing capabilities of a switch by rating how much data the switch can process per second.
- **Wire speed** is the data rate that each Ethernet port on the switch is capable of attaining. Data rates can be 100 Mb/s, 1 Gb/s, 10 Gb/s, or 100 Gb/s.
- **PoE** (Power over Ethernet) allows the switch to deliver power to a device over the existing Ethernet cabling. This feature can be used by IP phones and some wireless access points.
- **Multilayer switches**, so called Layer-3 switches, are typically deployed in the core and distribution layers of an organization's switched network

Router hardware

There are three categories of routers:

- **Branch Routers** – Branch routers optimize branch services on a single platform while delivering an optimal application experience across branch and WAN infrastructures.
- **Network Edge Routers** – Network edge routers enable the network edge to deliver high-performance, highly secure, and reliable services that unite campus, data center, and branch networks.
- **Service Provider Routers** – Service provider routers differentiate the service portfolio and increase revenues by delivering end-to-end scalable solutions and subscriber-aware services.

Chapter 2

VTP

2.1 Overviews

VLAN trunking protocol (VTP) allows a network administrator to manage VLANs on a switch configured as a VTP server. The VTP server distributes and synchronizes VLAN information over trunk links to VTP-enabled switches throughout the switched network. The following provides a brief description of important components of VTP:

VTP domain A VTP domain consists of all interconnected switches. All switches in a domain share VLAN configuration details. Switches resides in different domains do not exchange VTP messages. The boundary of a VTP domain is a router or a layer-3 switch.

Revision number The revision number is a 32-bit number that indicates the level of revision for a VTP advertisements. Each VTP device tracks the VTP configuration revision number that is assigned to it. Each time that you make a VLAN change in a VTP device, the configuration revision is incremented by one. Therefore, this number is used to determine whether the received information is more recent than the current version.

password Switches in the same domain are configured with the same password for security reason.

VTP modes A switch can be configured as a VTP server, client, or transparent.

VTP server stores the VLAN information in NVRAM (vlan.dat), then advertises it to other switches. VLAN configuration is allowed, and affects the entire VTP domain.

VTP client stores the VLAN information in RAM, therefore, a switch reset deletes all VLAN information. VLAN configuration is not allowed.

VTP transparent Switches in this mode do not participate in VTP except to forward VTP advertisements to VTP clients and VTP server. VLANs that are created, renamed, or deleted on transparent switches are local to that switch only.

Each switch in a VTP domain sends periodic VTP advertisements so that its neighbors can update VLAN configuration. VTP includes three types of advertisements:

- **Summary advertisements** – These inform adjacent switches of VTP domain name and configuration revision number. By default, Cisco switches issue summary advertisements every five minutes.
- **Advertisement request** – These are in response to a summary advertisement message when the summary advertisement contains a higher configuration revision number than the current value.
- **Subset advertisements** – These contain VLAN information including any changes.

2.2 Operations

When the switch receives a summary advertisement packet, the switch compares the VTP domain name to its own VTP domain name. If the name is different, the switch simply ignores the packet. If the name is the same, the switch then compares the configuration revision to its own revision. If its own configuration revision number is higher or equal to the packet's configuration revision number, the packet is ignored. If its own configuration revision number is lower, an advertisement request is sent asking for the subset advertisement message.

The subset advertisement message contains the VLAN information with any changes. When you add, delete, or change a VLAN on the VTP server, the VTP server increments the configuration revision and issues a summary advertisement. One or several subset advertisements follow the summary advertisement containing the VLAN information including any changes. This process is shown in the figure 2.1.

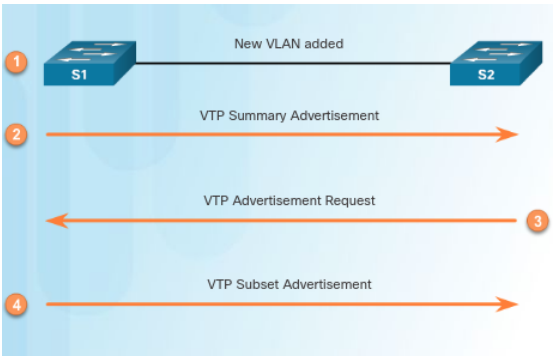

Figure 2.1: VTP operation

2.3 VTP Caveats

Adding a VTP-enabled switch to an existing VTP domain will wipe out the existing VLAN configurations in the domain if the new switch is configured with

different VLANs and has a higher configuration revision number than the existing VTP server (see figure 2.2). Therefore, when a switch is added to a network, ensure that it has a default VTP configuration. The VTP configuration revision

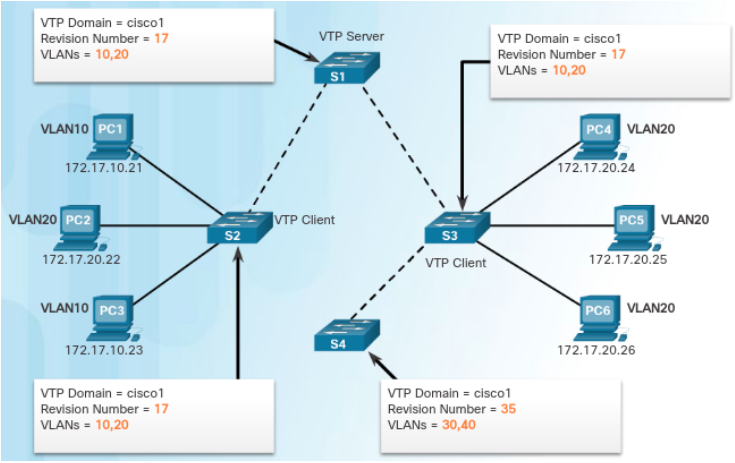

Figure 2.2: Incorrect VTP configuration revision number scenario

number is stored in NVRAM and is not reset if you erase switch configuration and reload it. To reset VTP configuration revision number to zero you have two options:

- Change the switch's VTP domain to a nonexistent VTP domain and then change the domain back to the original name.
- Change the switch's VTP mode to transparent and then back to previous VTP mode (Recommended).

Chapter 3

Layer 3 Switching

Inter-VLAN routing using the router-on-a-stick method was simple to implement because routers were usually available in every network. However, most modern enterprise networks use multilayer switches to achieve high-packet processing rates using hardware-based switching. Layer 3 switches usually have packet-switching throughputs in the millions of packets per second (pps), whereas traditional routers provide packet switching in the range of 100,000 pps to more than 1 million pps.

Many users are in separate VLANs, and each VLAN is usually a separate subnet. Therefore, it is logical to configure the distribution switches as Layer 3 gateways for the users of each access switch VLAN. This implies that each distribution switch must have IP addresses matching each access switch VLAN. This can be achieved by using Switch Virtual Interfaces (SVIs) and routed ports.

- **Routed port** – A pure Layer 3 interface similar to a physical interface on a Cisco IOS router.
- **Switch virtual interface (SVI)** – A virtual VLAN interface for inter-VLAN routing. In other words, SVIs are the virtual-routed VLAN interfaces.

A routed port is a physical port that acts similarly to an interface on a router. Unlike an access port, a routed port is not associated with a particular VLAN. A routed port behaves like a regular router interface. Unlike Cisco IOS routers, routed ports on a Cisco IOS switch do not support subinterfaces. Routed ports are used for point-to-point links. In a switched network, routed ports are mostly configured between switches in the core and distribution layer.

An SVI is a virtual interface that is configured for each VLAN that exists on the switch. It is considered to be virtual because there is no physical port dedicated to the interface. It can perform the same functions for the VLAN as a router interface would, and can be configured in much the same way as a router interface (i.e., IP address, inbound/outbound ACLs, etc.).

Chapter 4

STP

4.1 Issues with redundancy

Multiple cabled paths between switches provide physical redundancy in a switched network. This improves the reliability and availability of the network. However, when there is no spanning tree implementation on the switches, a Layer 2 loop occurs. A Layer 2 loop can result in the three primary issues:

- Duplicate unicast frames
- MAC database instability
- Broadcast storm

4.1.1 Multiple-frame transmission

Broadcast frames are not the only type of frames that are affected by loops. Unknown unicast frames sent onto a looped network can result in duplicate frames arriving at the destination device. An unknown unicast frame is when the switch does not have the destination MAC address in its MAC address table and must forward the frame out all ports, except the ingress port.

4.1.2 MAC Database Instability

Ethernet frames do not have a time to live (TTL) attribute. As a result, if there is no mechanism enabled to block continued propagation of these frames on a switched network, they continue to propagate between switches endlessly.

Broadcast frames are forwarded out all switch ports, except the original ingress port. If there is more than one path to the destination, the frames may be forwarded back to the original switch, and create an endless loop. When a loop occurs, the MAC address table on a switch to constantly change with the updates from the broadcast frames, which results in MAC database instability.

See this link for more explanation.

4.1.3 Broadcast storm

A broadcast storm occurs when there are so many broadcast frames caught in a Layer 2 loop that all available bandwidth is consumed. Consequently, no

bandwidth is available for legitimate traffic and the network becomes unavailable for data communication. This is an effective denial of service (DoS).

There are other consequences of broadcast storms. Because broadcast traffic is forwarded out every port on a switch, all connected devices have to process all the broadcast traffic that is being flooded endlessly around the looped network. This can cause the end device to malfunction because of the processing requirements needed to sustain such a high traffic load on the NIC.

4.2 Operation

4.2.1 Port Roles

IEEE 802.1D STP and RSTP use the Spanning Tree Algorithm (STA) to determine which switch ports on a network must be put in blocking state to prevent loops from occurring. The STA designates a single switch as the root bridge and uses it as the reference point for all path calculations (see figure 4.1).

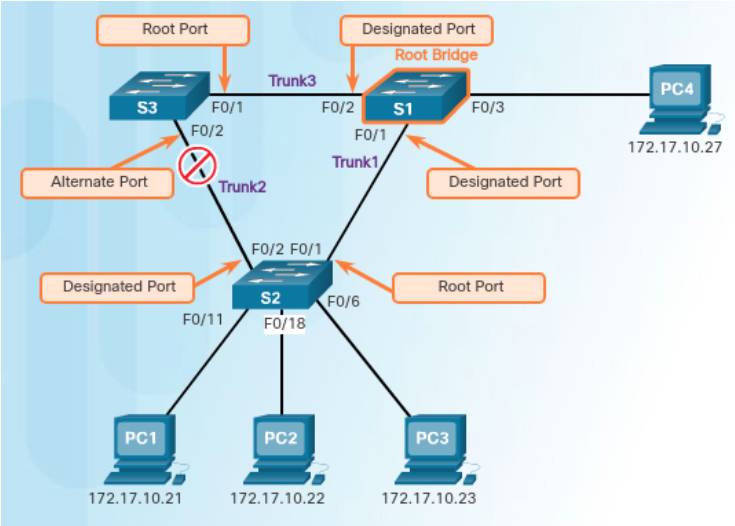

Figure 4.1: The STA designates a single switch as the root bridge and uses it as the reference point for all path calculations.

After the root bridge has been determined, the STA calculates the shortest path to the root bridge, while each switch uses the STA to determine which ports to block. When the STA has determined which paths are most desirable relative to each switch, it assigns port roles to the participating switch ports:

- **Root port** – Switch ports closest to the root bridge in terms of overall cost to the root bridge.
- **Designated port** – All non-root ports that are still permitted to forward traffic on the network.
- **Alternate port** – Alternate ports and backup ports are in discarding or blocking state to prevent loops.

Sometimes, the administrator wants to determine port roles without calculating port cost. He/She should keep in mind the following tips:

- There can only be one root port per non-root switch
- If one end of a segment (the link between two switches) is a root port, then the other end is a designated port.
- All ports on the root bridge are designated ports.
- Alternate ports are selected only on segments where neither end is a root port.

4.2.2 Root bridge

Every switch has its own BID and a root ID:

4.2.3 Bridge ID (BID)

This field is used to uniquely identify the switch in the election process. It includes the priority, extended system ID, and MAC address (see figure 4.2). The bridge priority value is automatically assigned, but can be modified by the administrator. The extended system ID is used to specify a VLAN ID or a multiple spanning tree protocol (MSTP) instance ID. The bridge priority is

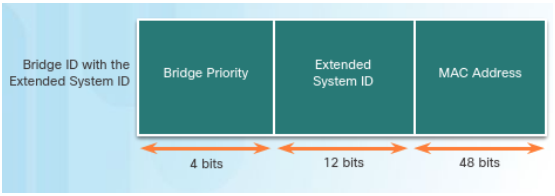

Figure 4.2: BID fields

a customizable value that can be used to influence which switch becomes the root bridge. The switch with the lowest priority, which implies the lowest BID, becomes the root bridge because a lower priority value takes precedence. The default priority value for all Cisco switches is the decimal value 32768. The range is 0 to 61440 in increments of 4096.

The extended system ID reserves the rightmost 12 bits for the VLAN ID and the far left 4 bits for the bridge priority. This explains why the bridge priority value can only be configured in multiples of 4096, or 2^12.

When two switches are configured with the same priority and have the same extended system ID, the switch having the MAC address with the lowest value, expressed in hexadecimal, will have the lower BID. Initially, all switches are configured with the same default priority value. The MAC address is then the deciding factor as to which switch is going to become the root bridge.

root ID

This field indicates the BID of the root bridge. When a switch first boots, the root ID is the same as the bridge ID (BID). However, as the election occurs, the lowest BID replaces the local root ID to identify the root bridge.

Election process

All switches in the broadcast domain participate in the election process. After a switch boots, it begins to send out BPDU frames every two seconds. These BPDUs contain the switch BID and the root ID.

Assuming that switch A,B,C,D resides in the same STP domain. As the switch A forward its BPDU frames, adjacent switch B reads the root ID information from the BPDU frames. If the root ID from the BPDU received is lower than the root ID on B, then the B updates its root ID, identifying the A as the root bridge. Switch B then forwards new BPDU frames with the lower root ID to switch C.

The same process repeats, switch C compares its current root ID with the root ID identified in the frames, and then updates its current root ID if needed. Eventually, the switch with the lowest BID ends up being identified as the root bridge for the spanning tree instance.

4.2.4 Port role decision

After the root bridge is elected, the STA determines port roles on interconnecting links.

The root bridge automatically configures all of its switch ports in the designated role. Non-root switches transition ports with the lowest root path cost to root ports, and the other to either designated or alternate port (because there can only be one root port per non-root switch).

Next step is to decide which port to configure as a designated or alternative port. On the segment, where root ports have already been defined, two switches exchange BPDU frames, which contain the BID. Generally, the switch with the lower BID has its port configured as a designated port while the other has its port configured as an alternate port. However, keep in mind that the first priority is the lowest path cost to the root bridge and that the sender's BID is used only if the port costs are equal.

4.2.5 Root path cost

The STA considers both path and port costs when determining which ports to block.The path information, known as the internal root path cost, is determined by summing up the individual port costs along the path from the switch to the root bridge. The default port costs are defined by the speed at which the port operates. 10 Gb/s Ethernet ports have a port cost of 2, 1 Gb/s Ethernet ports have a port cost of 4, 100 Mb/s Fast Ethernet ports have a port cost of 19, and 10 Mb/s Ethernet ports have a port cost of 100.

Switches send BPDUs, which include the root path cost. This is the cost of the path from the sending switch to the root bridge. When a switch receives the BPDU, it adds the ingress port cost of the segment to determine its internal root path cost.

4.2.6 BPDU frame

A Bridge Protocol Data Units (BPDU) is a frame exchanged by switches for STP. The spanning tree algorithm depends on the exchange of BPDUs to determine a root bridge. BPDUs have a destination MAC address of 01:80:C2:00:00:00,

which is a multicast address for the spanning tree group. A BPDU frame contains 12 distinct fields:

- The first four fields identify the protocol, version, message type, and status flags.
- The next four fields are *root ID*, *bridge ID (BID)*, root path cost. They are used to identify the root bridge and the root path cost to the root bridge.
- The last four fields are all timer fields that determine how frequently BPDU messages are sent and how long the information received through the BPDU process is retained.

4.3 Types of STP

- **STP** – This is the original IEEE 802.1D version. The protocol assumes one spanning tree instance for the entire bridged network.
- **PVST+** – This is a Cisco enhancement of STP that provides a separate 802.1D spanning tree instance for each VLAN configured in the network.
- **802.1D-2004** – This is an updated version of the STP standard, incorporating IEEE 802.1w.
- **RSTP or IEEE 802.1w** – This is an evolution of STP.
- **Rapid PVST+** – This is a Cisco enhancement of RSTP that uses PVST+.
- **MSTP** – Multiple Spanning Tree Protocol. This is an IEEE standard inspired by the earlier Cisco proprietary Multiple Instance STP (MISTP) implementation. MSTP maps multiple VLANs into the same spanning tree instance. The Cisco implementation of MSTP is MST, which provides up to 16 instances of RSTP

4.3.1 Port states and PVST+ Operation

To facilitate the learning of the logical spanning tree, each switch port transitions through five possible port states and three BPDU timers:

4.3.2 Rapid PVST+

Rapid PVST+ is the Cisco implementation of RSTP on a per-VLAN basis. An independent instance of RSTP runs for each VLAN.

RSTP does not have a blocking port state. Port states are defined as discarding, learning, or forwarding. RSTP also speeds the recalculation of the spanning tree when the Layer 2 network topology changes. If a port is configured to be an alternate port or a backup port, it can immediately change to a forwarding state without waiting for the network to converge.

RSTP introduces new types of port: edge port. An RSTP edge port is a switch port that is never intended to be connected to another switch. It immediately transitions to the forwarding state when enabled. The RSTP edge port

Table 4.1: Five port states in PVST+

	Port states				
Operation allowed	Blocking	Listening	Learning	Forwarding	Disabled
Learn MAC addresses	YES	YES	YES	YES	NO
Forward data frames received on interface	NO	NO	YES	YES	NO
Forward data frames switched from another interface	NO	NO	NO	YES	NO
Receive and process BPDUs	NO	NO	NO	YES	NO

concept corresponds to the PVST+ PortFast feature. An edge port is directly connected to an end station and assumes that no switch device is connected to it.

Chapter 5

EtherChannel

5.1 Introduction

5.1.1 Advantages

EtherChannel technology has many advantages:

- Most configuration tasks can be done on the EtherChannel interface instead of on each individual port, ensuring configuration consistency throughout the links.
- EtherChannel creates an aggregation that is seen as one logical link.
- EtherChannel provides redundancy.
- Load balancing takes place between links that are part of the same EtherChannel.
- EtherChannel relies on existing switch ports. There is no need to upgrade the link to a faster and more expensive connection to have more bandwidth.

5.1.2 Implementation restrictions

The EtherChannel provides full-duplex bandwidth between one switch and another switch or host. Currently each EtherChannel can consist of up to eight compatibly-configured Ethernet ports. However, interface types cannot be mixed. For example, Fast Ethernet and Gigabit Ethernet cannot be mixed within a single EtherChannel.

EtherChannel creates a one-to-one relationship; that is, one EtherChannel link connects only two devices. The individual EtherChannel group member port configuration must be consistent on both devices. It is mandatory that all ports have the same speed, duplex setting, and VLAN information. Any port modification after the creation of the channel also changes all other channel ports. If the physical ports of one side are configured as trunks, the physical ports of the other side must also be configured as trunks within the same native VLAN. Additionally, all ports in each EtherChannel link must be configured as Layer 2 ports.

5.2 Port Aggregation Protocol (PagP)

PAgP (pronounced "Pag – P") is a Cisco-proprietary protocol that aids in the Passivematic creation of EtherChannel links. PAgP helps create the EtherChannel link by detecting the configuration of each side and ensuring that links are compatible so that the EtherChannel link can be enabled when needed. PAgP can be configured in one of three models:

- **On** – This mode forces the interface to channel without PAgP. Interfaces configured in the on mode do not exchange PAgP packets.
- **PAgP Active** – This PAgP mode places an interface in an active negotiating state in which the interface initiates negotiations with other interfaces by sending PAgP packets.
- **PAgP Passive** – This PAgP mode places an interface in a passive negotiating state in which the interface responds to the PAgP packets that it receives, but does not initiate PAgP negotiation.

The modes must be compatible on each side as shown in table 5.1.

Table 5.1: PAgP Establishment

S1	S2	EtherChannel establishment
Active	Passive/Active	Yes
On	On	Yes
Passive	Passive/On	No
Not configured	Passive/Active/On	No
Active	on	No

5.3 Link Aggregation Control Protocol (LACP)

LACP is part of an IEEE specification (802.3ad) that allows several physical ports to be bundled to form a single logical channel. Because LACP is an IEEE standard, it can be used to facilitate EtherChannels in multivendor environments, including Cisco devices. LACP allows for eight active links, and also eight standby links. A standby link will become active should one of the current active links fail. PAgP can be configured in one of three models:

- **On** – This mode forces the interface to channel without LACP. Interfaces configured in the on mode do not exchange LACP packets.
- **LACP active** – Similar to PAgP Active mode.
- **LACP passive** – Similar to PAgP Passive mode. negotiation.

The modes must be compatible on each side as shown in table 5.1.

5.4 Configuration

Table 5.2: LACP Establishment

S1	S2	EtherChannel establishment
Active	Passive/Active	Yes
On	On	Yes
Passive	Passive/On	No
Not configured	Passive/Active/On	No
Active	on	No

Chapter 6

HSRP

6.1 Operations

One way to prevent a single point of failure at the default gateway, is to implement a virtual router. To implement this type of router redundancy, multiple routers are configured by First Hop Redundancy Protocol (FHRP) to work together as an illusion of a single router to the hosts on the LAN. One the most popular options for FHRP is Hot Standby Router Protocol (HSRP). HSRP was designed by Cisco to allow for gateway redundancy without any additional configuration on end devices.

One of the routers is selected by HSRP to be the active router. The active router will act as the default gateway for end devices. The other router will become the standby router. The default gateway address is a virtual IPv4 address along with a virtual MAC address that is shared amongst both HSRP routers. End devices use this virtual IPv4 address as their default gateway address.(See figure 6.1)

The HSRP virtual IPv4 address is configured by the network administrator. The virtual MAC address is created automatically.

6.1.1 Priority

HSRP priority can be used to determine the active router. The router with the highest HSRP priority will become the active router. By default, the HSRP priority is 100. If the priorities are equal, the router with the numerically highest IPv4 address is elected as the active router.

6.1.2 Preemption

By default, after a router becomes the active router, it will remain the active router even if another router comes online with a higher HSRP priority. This means that the router which boots up first will become the active router if there are no other routers online during the election process. To force a new HSRP election process, preemption must be enabled. Preemption is the ability of an HSRP router to trigger the re-election process.

With preemption enabled, a router that comes online with a higher HSRP priority will assume the role of the active router. Preemption only allows a

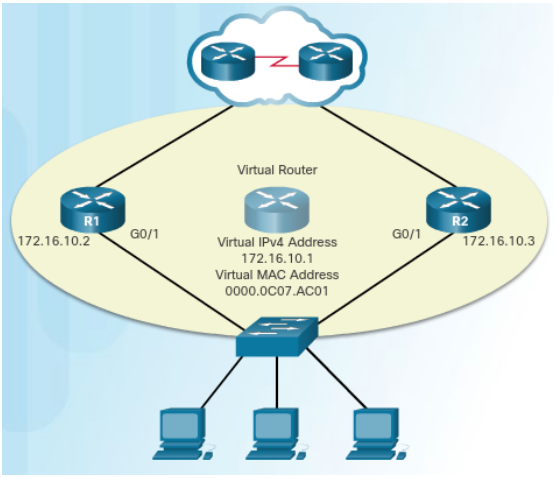

Figure 6.1: HSRP topology

router to become the active router if it has a higher priority. A router enabled for preemption, with equal priority but a higher IPv4 address will not preempt an active router.

6.1.3 States and timers

When an interface is configured with HSRP or is first activated with an existing HSRP configuration, the router sends and receives HSRP hello packets to begin the process of determining which state it will assume in the HSRP group. The active and standby HSRP routers send hello packets to the HSRP group multicast address every 3 seconds, by default. The standby router will become active if it does not receive a hello message from the active router after 10 seconds. However, to avoid increased CPU usage and unnecessary standby state changes, do not set the hello timer below 1 second or the hold timer below 4 seconds.

6.2 Configuration

Complete the following steps to configure HSRP (see figure **??** for example):

1. Configure HSRP version 2.
2. Configure the virtual IP address for the group.
3. Configure the priority for the desired active router to be greater than 100.
4. Configure the active router to preempt the standby router in cases where the active router comes online after the standby router.

```
R1(config)# interface g0/1
R1(config-if)# ip address 172.16.10.2 255.255.255.0
R1(config-if)# standby version 2
R1(config-if)# standby 1 ip 172.16.10.1
```

```
R1(config-if)# standby 1 priority 150
R1(config-if)# standby 1 preempt
R1(config-if)# no shutdown
```

Chapter 7

EIGRP

7.1 Basic features

7.1.1 Protocol Dependent Modules (PDM)

EIGRP has the capability for routing different protocols, including IPv4 and IPv6. EIGRP does so by using protocol-dependent modules (PDMs). PDMs are responsible for network layer protocol-specific tasks.

7.1.2 Reliable Transport Protocol (RTP)

EIGRP was designed as a network layer independent routing protocol. Because of this design, EIGRP cannot use the services of UDP or TCP. Instead, EIGRP uses the Reliable Transport Protocol (RTP) for the delivery and reception of EIGRP packets.

RTP includes both reliable delivery and unreliable delivery of EIGRP packets, similar to TCP and UDP, respectively. For example, an EIGRP update packet is sent reliably over RTP and requires an acknowledgment. An EIGRP Hello packet is also sent over RTP, but unreliably. This means that EIGRP Hello packets do not require an acknowledgment.

RTP can send EIGRP packets as unicast or multicast. Multicast EIGRP packets use the multicast address 224.0.0.10 for IPv4 and FF02::A for IPv6.

7.2 Packet types

7.2.1 Hello packets

EIGRP uses small Hello packets to discover other EIGRP-enabled routers on directly connected links. Hello packets are used by routers to form EIGRP neighbor adjacencies.

On most modern networks, EIGRP Hello packets are sent as multicast packets every five seconds. However, on multipoint, non-broadcast multiple access (NBMA) networks with access links of T1 (1.544 Mb/s) or slower, Hello packets are sent as unicast packets every 60 seconds.

EIGRP uses a Hold timer to determine the maximum time the router should wait to receive the next Hello before declaring that neighbor as unreachable. By

default, the hold time is three times the Hello interval, or 15 seconds on most networks and 180 seconds on low-speed NBMA networks. If the hold time expires, EIGRP declares the route as down and DUAL searches for a new path by sending out queries.

7.2.2 Update packets

EIGRP sends Update packets to propagate routing information. Update packets are sent only when necessary. EIGRP updates contain only the routing information needed and are sent only to those routers that require it.

EIGRP uses the terms *partial update* and *bounded update* when referring to its updates. A partial update means that the update only includes information about route changes. A bounded update refers to the sending of partial updates only to the routers that are affected by the changes. Bounded updates help EIGRP minimize the bandwidth that is required to send EIGRP updates.

7.2.3 Acknowledgment packets

EIGRP sends Acknowledgment (ACK) packets when reliable delivery is used. An EIGRP acknowledgment is an EIGRP Hello packet without any data. RTP uses reliable delivery for Update, Query, and Reply packets.

7.2.4 Query and reply packets

DUAL uses Query and Reply packets when searching for networks and other tasks. Queries and replies use reliable delivery. Queries can use multicast or unicast, whereas replies are always sent as unicast.

7.3 Encapsulating EIGRP Messages

7.3.1 TLV fields

The data portion of an EIGRP message is encapsulated in a packet (figure 7.1). This data field is called *type*, *length*, *value* (TLV). The *type*s of TLVs relevant to this course are EIGRP parameters (figure 7.2), IP internal routes (figure 7.3), and IP external routes (figure 7.4). The *length* field identifies the size (in bytes) of the *value* field. The *value* field contains data for EIGRP message.

The EIGRP parameters include the weights that EIGRP uses for its composite metric (K1 – K5). By default, only bandwidth and delay are weighted. Both are weighted equally; therefore, the K1 field for bandwidth and the K3 field for delay are both set to one (1). The other K values are set to zero (0).

Each IP internal routes and IP external routes contains one route entry and the metric information for that route. The Update packet parameters identify IP internal and external routes. The IP internal message is used to advertise EIGRP routes within an autonomous system, whereas the IP external message is used to import default static route, as well as routes outside the autonomous system, into the EIGRP routing process.

7.3.2 Packet header

In the packet header (figure 7.1), the protocol field is set to 88 to indicate EIGRP, and the destination address is set to the multicast 224.0.0.10 for IPv4, and FF02::A for IPv6. If the EIGRP packet is encapsulated in an Ethernet frame, the destination MAC address is also a multicast address, 01-00-5E-00-00-0A.

Another important field in the packet header is opcode field, which specifies EIGRP packet type. Specifically, it identifies the EIGRP messages as either type 1 = Update, type 3 = Query, type 4 = Reply, type 5 = Hello.

The autonomous system number specifies the EIGRP routing process. Unlike RIP, multiple instances of EIGRP can run on a network. The autonomous system number is used to track each running EIGRP process.

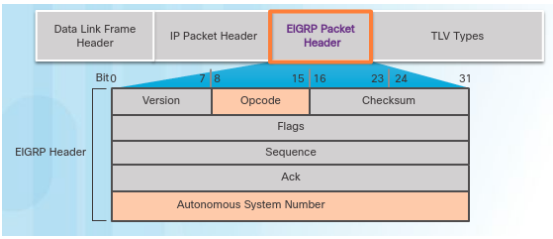

Figure 7.1: EIGRP packet header

Figure 7.2: EIGRP paramters TLV fields

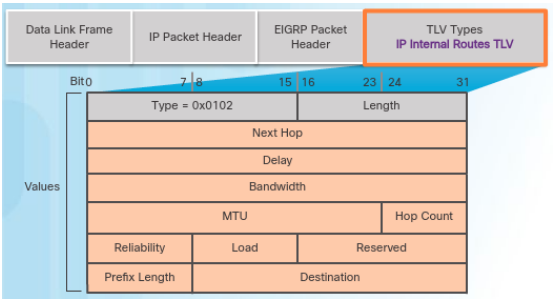

Figure 7.3: EIGRP internal routes TLV fields

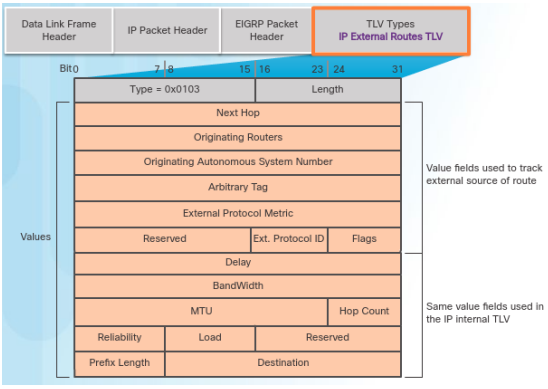

Figure 7.4: EIGRP external route TLV fields

7.4 Operation

7.4.1 Neighbor adjacency

EIGRP uses Hello packets to establish and maintain neighbor adjacencies. To accomplish this, two EIGRP routers must use the same EIGRP metric parameters (K values) and both must be configured using the same autonomous system number.

Each EIGRP router maintains a neighbor table, which contains a list of routers on shared links that have an EIGRP adjacency with this router. The neighbor table is used to track the status of these EIGRP neighbors.

When a new router A comes up on the link and it floods hello packets through all of its EIGRP-configured interfaces. The neighbor routers receive those packets, adds A to their neighbor tables, and sends hello as well as update packets to A. Router A, then add those routers to its neighbor table and updates its topology table with information from received packets.

7.4.2 Topology table

Each EIGRP router maintains a topology table for each routed protocol configured, such as IPv4 and IPv6. The topology table includes route entries for every destination that the router learns from its directly connected EIGRP neighbors.

When a router receives the EIGRP update from neighbor, it adds all update entries to its topology table. Because EIGRP update packets use reliable delivery, the router replies with an EIGRP acknowledgment packet informing its neighbors that it has received the update.

7.4.3 Metric

By default, EIGRP uses the following values in its composite metric to calculate the preferred path to a network:

- **Bandwidth** – The slowest bandwidth among all of the outgoing interfaces, along the path from source to destination.

- **Delay** – The cumulative (sum) of all interface delay along the path (in tens of microseconds).

Default composite formula:

$$\text{metric} = (K1 \times \text{bandwidth} + K1 \times \text{bandwidth}) \times 256$$

Complete composite formula:

$$\text{metric} = \left(K1 \times \text{bandwidth} + \frac{K2 \times \text{bandwidth}}{256 \times \text{load}} + K3 \times \text{delay}\right) \times \frac{K5}{K4 + \text{reliability}} \times 256$$

This is a conditional formula. If $K5 = 0$, the last term is replaced by 1. Default values for each parameter:

- K1 (bandwidth) = 1
- K2 (load) = 0
- K3 (delay) = 0
- K4 (reliability) = 0
- K5 (reliability) = 0

Bandwidth

EIGRP uses the slowest bandwidth along the path to the destination network. EIGRP divides a reference bandwidth value of 10^7 by the interface bandwidth value in kb/s. If the result is not a whole number, then the value is rounded down. For example, 10^7 divided by 1024 equals 9765.625. The .625 is dropped to yield 9765 for the bandwidth portion of the composite metric.

7.4.4 Delay

EIGRP uses the sum of all delays along the path to the destination. The sum of these delays is divided by 10. See also table **??** for default delay values

For example, along the path R1→R2→R3, the Serial 0/0/1 interface on R2 has a delay of 20,000 microseconds, the Gigabit 0/0 interface on R3 has a delay of 10 microseconds. The delay is $(20000 + 10) \div 10 = 2001$.

Table 7.1: Default delay values

Media	Delay
Ehternet	1000
Fast Ethernet	100
Gigabit Ethernet	10
Serial link	20000

7.4.5 DUAL algorithm

EIGRP uses the Diffusing Update Algorithm (DUAL) to provide the best loop-free path and loop-free backup paths. DUAL uses several terms, which are discussed in more detail throughout this section:

- Successor – A successor is a neighboring router that is used for packet forwarding and is the least-cost route to the destination network. The IP address of a successor is shown in a routing table entry right after the word via (see figure **??**).
- Feasible Distance (FD) – FD is the lowest calculated metric to reach the destination network. FD is the metric listed in the routing table entry as the second number inside the brackets (see figure **??**). As with other routing protocols, this is also known as the metric for the route.
- Feasible Successor (FS) – An FS is a neighbor that has a loop-free backup path to the same network as the successor, and it satisfies the Feasibility Condition (FC). FS is not represented in the routing table until the Successor is down.
- Reported Distance (RD) – The RD is simply an EIGRP neighbor's FD to the same destination network.
- Feasible Condition or Feasibility Condition (FC) – The FC is met when a neighbor's RD to a network is less than the local router's feasible distance to the same destination network. If the RD is less, it represents a loop-free path.

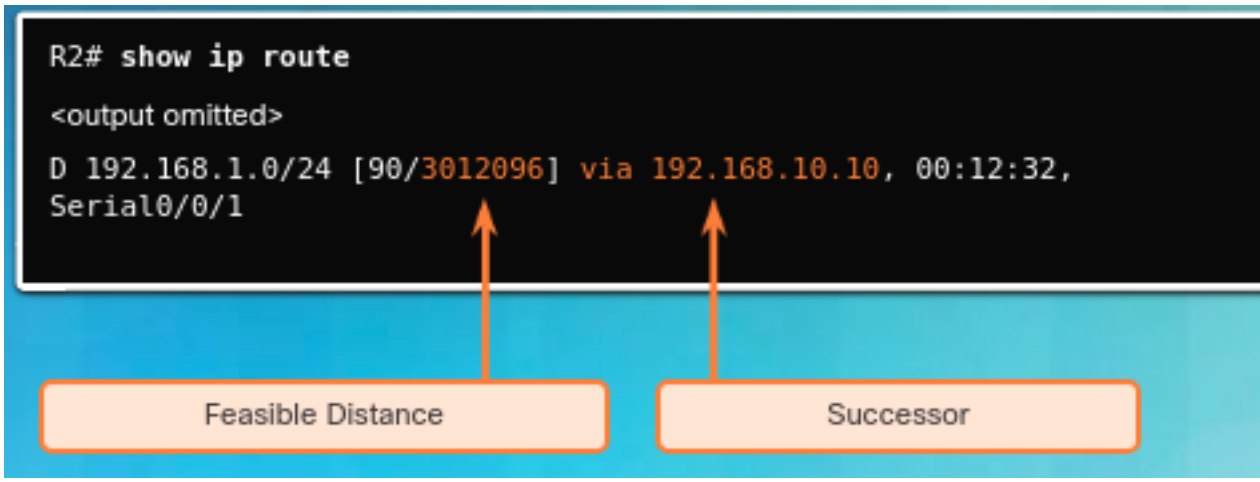

Figure 7.5: Successor and Feasible Distance

The decision process for all route computations is done by the DUAL Finite State Machine (FSM). The DUAL FSM tracks all routes and uses EIGRP metrics to select efficient, loop-free paths, and to identify the routes with the least-cost path to be inserted into the routing table.

Recomputation of the DUAL algorithm can be processor-intensive. EIGRP avoids recomputation whenever possible by maintaining a list of backup routes that DUAL has already determined to be loop-free. If the primary route in the routing table fails, the best backup route is immediately added to the routing table.

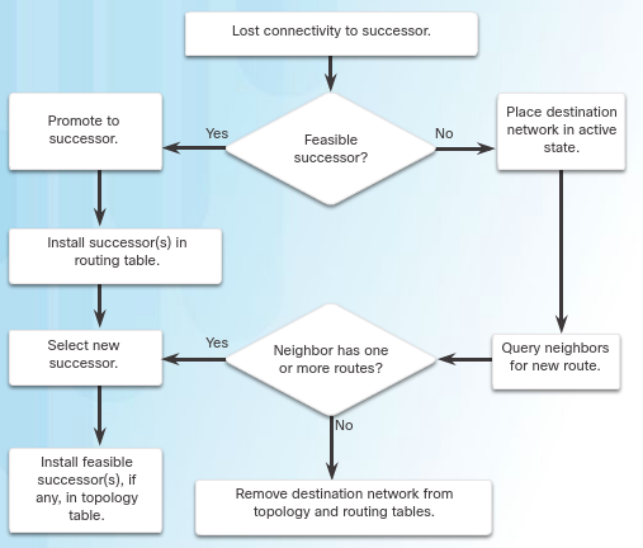

Figure 7.6: DUAL Finite State machine

When the successor is no longer available and there is no feasible successor, DUAL puts the route into an active state. DUAL sends EIGRP queries asking other routers for a path to the network. Other routers return EIGRP replies, letting the sender of the EIGRP query know whether or not they have a path to the requested network. If none of the EIGRP replies have a path to this network, the sender of the query does not have a route to this network. See also figure **??**.

7.5 Tune EIGRP

7.5.1 Automatic summarization

Route summarization allows a router to group networks together and advertises them as one large group using a single, summarized route. Summarization decreases the number of entries in routing updates and lowers the number of entries in local routing tables. It also reduces bandwidth utilization for routing updates and results in faster routing table lookups.

However, in classes IP network, the only way that all routers can find the best routes for each individual subnet is for neighbors to send subnet information. In this situation, automatic summarization should be disabled. When automatic summarization is disabled, updates include subnet information.

A problem associated with automatic route summarization is that a summary address also advertises networks that are not available on the advertising

router. For example, R1 is advertising the summary address 172.16.0.0/16, but it is only connected to 172.16.1.0/24. Therefore, R1 may receive incoming packets to destinations that do not exist (for example, 172.16.2.0/24). It then forwards a request to a destination network that does not exist, creating a routing loop.
EIGRP uses the Null0 interface to avoid this problem. The Null0 interface is a virtual IOS interface that is a route to nowhere. If R1 receives a packet destined for a network that is advertised by the classful mask but does not exist, it sends discards the packets by sending them to Null0.

Note! The Null0 summary route is removed when autosummarization is disabled using the no auto-summary router configuration mode command.

7.5.2 Hello and hold timers

The hold time tells the router the maximum time that the router should wait to receive the next Hello before declaring that neighbor as unreachable.Hello intervals and hold times are configurable on a per-interface basis and do not have to match with other EIGRP routers to establish or maintain adjacencies.

If the Hello interval is changed, ensure that the hold time value is not less than, the Hello interval. Otherwise, neighbor adjacency goes down after the hold time expires and before the next Hello interval.