
The Notebook of CCNA

HUY BUI



THE PUBLISHER

Contents

1 LAN Design	4
1.1 Hierarchical Design Model	4
1.2 Expanding the network	5
1.2.1 Planning for redundancy	5
1.2.2 Failure domain	5
1.2.3 Increasing Bandwidth	6
1.2.4 Expanding the Access Layer	6
1.2.5 Fine-tuning Routing Protocols	6
1.3 Selecting network devices	6
1.3.1 Switch hardwares	6
2 VTP	8
2.1 Overviews	8
2.2 Operations	8
2.3 VTP Caveats	9
3 Layer 3 Switching	11
4 STP	12
4.1 Issues with redundancy	12
4.1.1 Multiple-frame transmission	12
4.1.2 MAC Database Instability	12
4.1.3 Broadcast storm	12
4.2 Operation	13
4.2.1 Port Roles	13
4.2.2 Root bridge	13
4.2.3 Bridge ID (BID)	14
4.2.4 Port role decision	14
4.2.5 Root path cost	15
4.2.6 BPDU frame	15
4.3 Types of STP	15
4.3.1 Port states and PVST+ Operation	15
4.3.2 Rapid PVST+	15
5 EtherChannel	17
5.1 Introduction	17
5.1.1 Advantages	17
5.1.2 Implementation restrictions	17
5.2 Port Aggregation Protocol (PagP)	17
5.3 Link Aggregation Control Protocol (LACP)	18
5.4 Configuration	18

6	HSRP	19
6.1	Operations	19
6.1.1	Priority	20
6.1.2	Preemption	20
6.1.3	States and timers	20
6.2	Configuration	20
7	EIGRP	21
7.1	Basic features	21
7.1.1	Protocol Dependent Modules (PDM)	21
7.1.2	Reliable Transport Protocol (RTP)	21
7.2	Packet types	21
7.2.1	Hello packets	21
7.2.2	Update packets	21
7.2.3	Acknowledgment packets	22
7.2.4	Query and reply packets	22
7.3	Encapsulating EIGRP Messages	22
7.3.1	TLV fields	22
7.3.2	Packet header	22
7.4	Operation	22
7.4.1	Neighbor adjacency	22
7.4.2	Topology table	24
7.4.3	Metric	24
7.4.4	Delay	24
7.4.5	DUAL algorithm	25
7.5	Tune EIGRP	26
7.5.1	Automatic summarization	26
7.5.2	Hello and hold timers	27
8	OSPF	28
8.1	Introduction	28
8.1.1	Link-state operation	28
8.1.2	OSPF network types	28
8.1.3	OSPF cost	30
8.2	Protocol components	30
8.2.1	Data structure	30
8.2.2	Messages	30
8.2.3	Algorithm	31
8.3	DR election	31
8.4	Multiarea OSPF	33
8.4.1	Single-area vs Multi-area OSPF	33
8.4.2	Introduction to multiarea OSPF	34
9	ACL	35
9.1	ACL Operation Overview	35
9.1.1	ACEs Logic Operations	35
9.1.2	Inbound and Outbound ACL Logic	35
9.1.3	Numbered and Named ACLs	36
9.2	Standard ACL	36
9.2.1	Overview	36
9.2.2	Standard ACL placement	36
9.3	Extended ACLs	36
9.3.1	Overview	36
9.3.2	Extended ACL Placement	37
9.4	IPv6 ACLs	37
9.5	Configurations	38
9.6	Troubleshoot	40

10 Network Security and Monitoring	42
10.1 Security attacks	42
10.1.1 CDP Reconnaissance Attack	42
10.1.2 Telnet Attacks	42
10.1.3 MAC Address Table Flooding Attack	42
10.1.4 VLAN Attacks	42
10.1.5 DHCP Attacks	43
10.1.6 Cisco solution	44
10.1.7 The AAA framework	44
10.1.8 802.1X	44
10.2 SNMP	45
10.2.1 Introduction to SNMP	45
10.2.2 SNMP operation	45
10.3 SNMP configuration	46
10.3.1 SNMPv2	46
10.3.2 SNMPv3	47
10.4 SPAN	48
10.4.1 Introduction	48
10.4.2 Local SPAN	48
10.4.3 RSPAN	48
11 Quality of Service	49
11.1 Network transmission	49
11.1.1 Network quality	49
11.1.2 Traffic characteristics	49
11.2 Queueing algorithms	50
11.2.1 WFQ	50
11.2.2 CBWFQ	50
11.2.3 LLQ	51
11.3 QoS models	51
11.3.1 Best effort	51
11.3.2 Integrated services	51
11.3.3 Differentiated services	52
11.4 QoS implementation	53
11.4.1 Classification and Marking	53
11.4.2 Congestion Avoidance	54
11.4.3 Shaping and Policing	55

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.

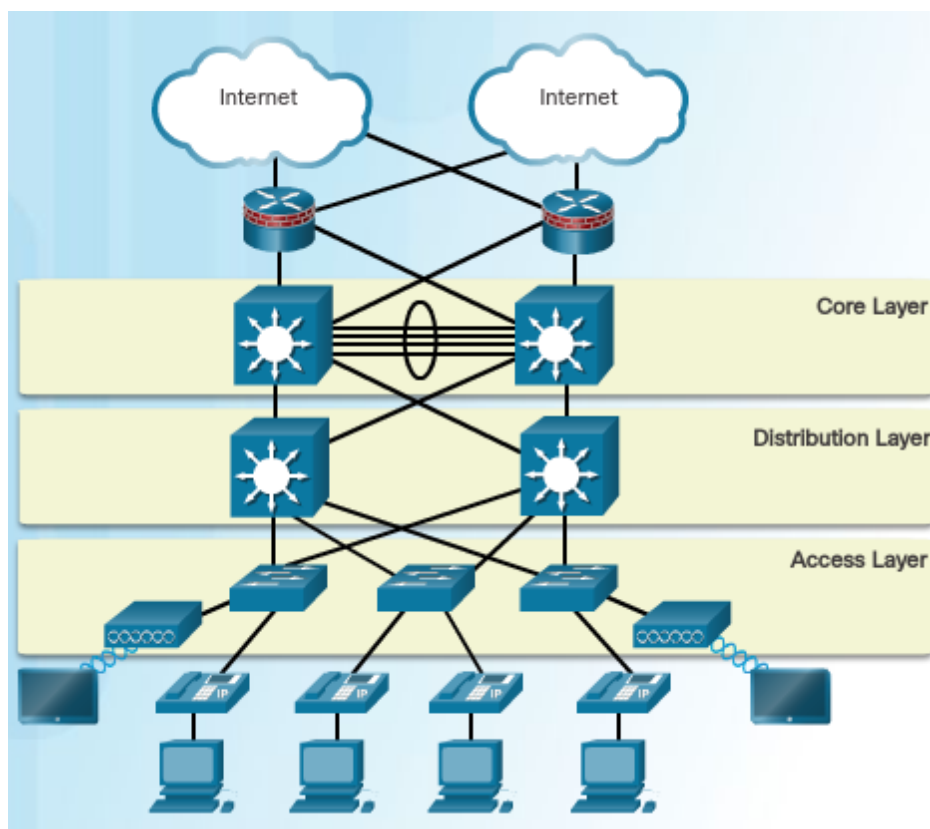


Figure 1.1: Three-layer hierarchical design model

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, as shown in Figure 1.2.

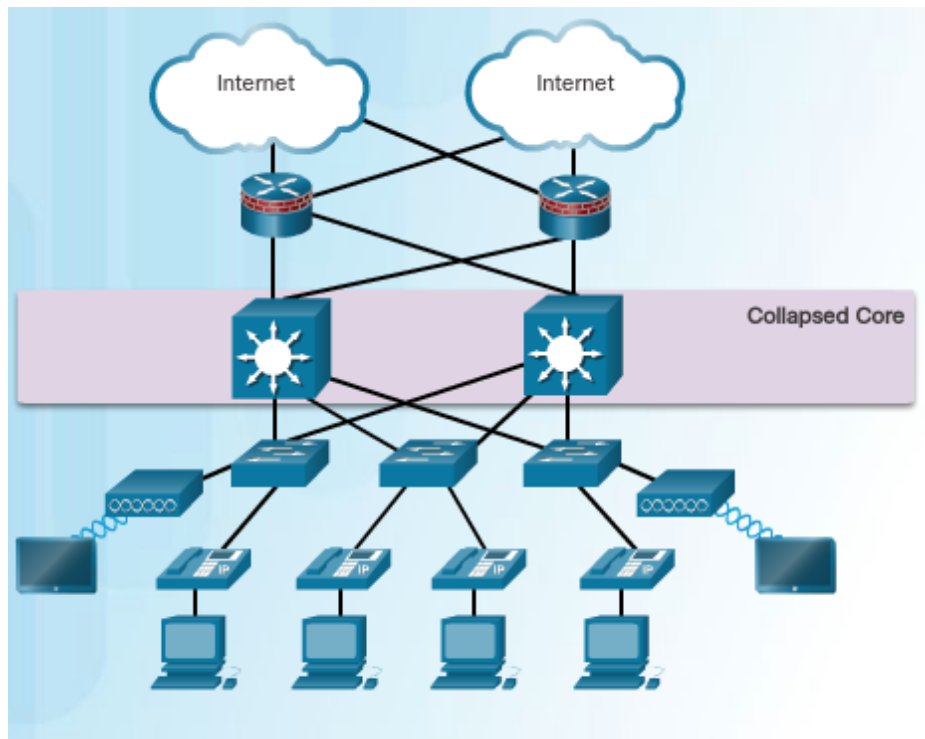


Figure 1.2: Collapsed Core

1.2 Expanding the 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.
- Design a hierarchical network to include modules that can be upgraded 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 hardware

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.

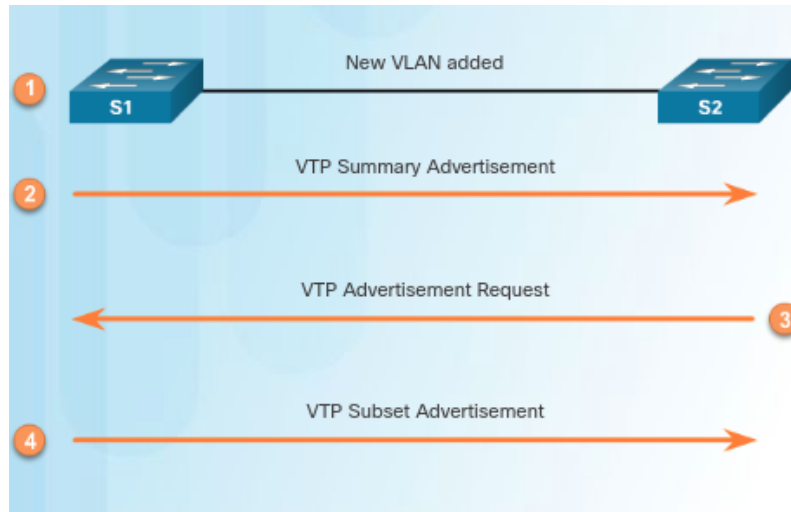


Figure 2.1: VTP operation

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 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).

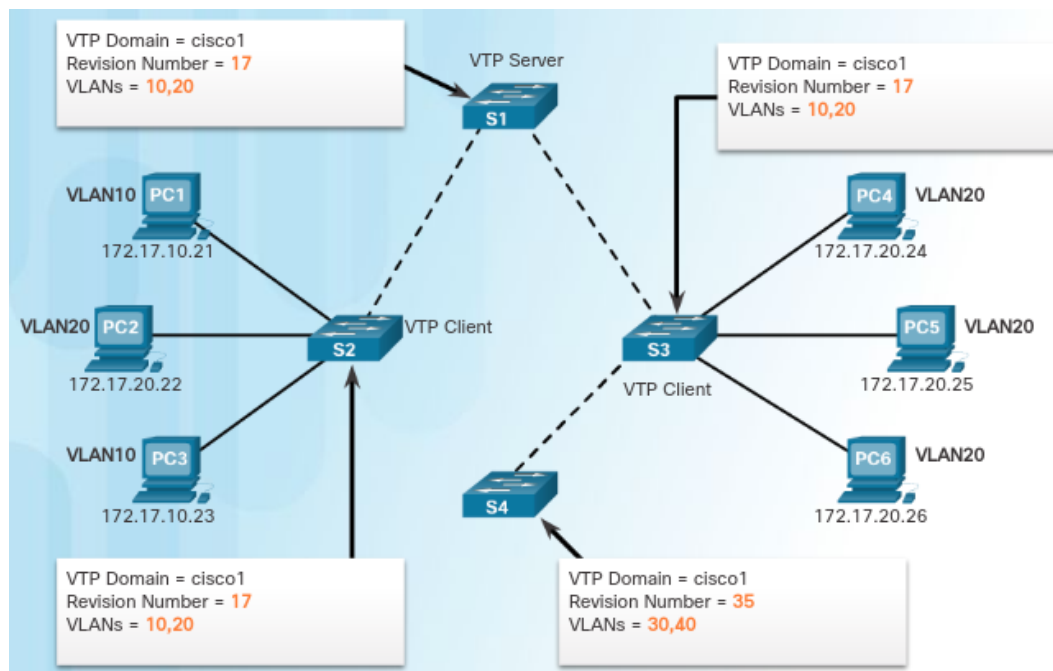


Figure 2.2: Incorrect VTP configuration revision number scenario

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

Copies of the same frame are received on different ports of the switch

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).

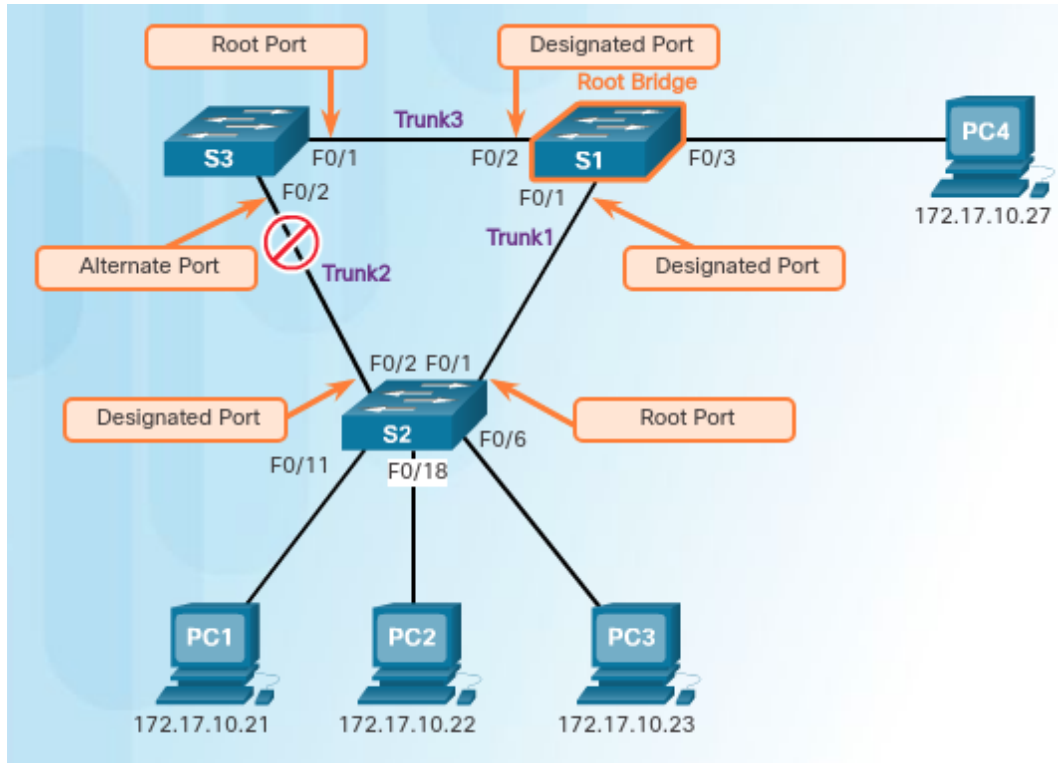


Figure 4.1: The STA designates a single switch as the root bridge and uses it as the reference point for all path calculations.

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 a customizable value that can be used to influence which switch becomes the

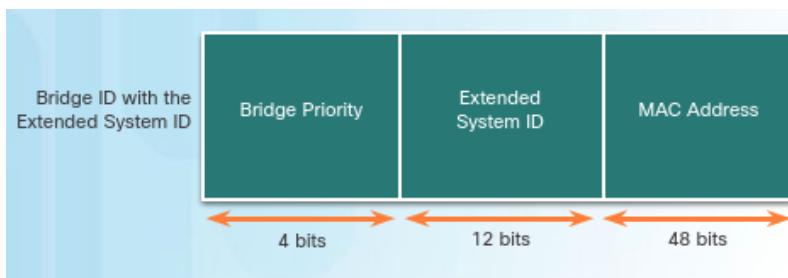


Figure 4.2: BID fields

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 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.

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

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.

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

5.4 Configuration

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.

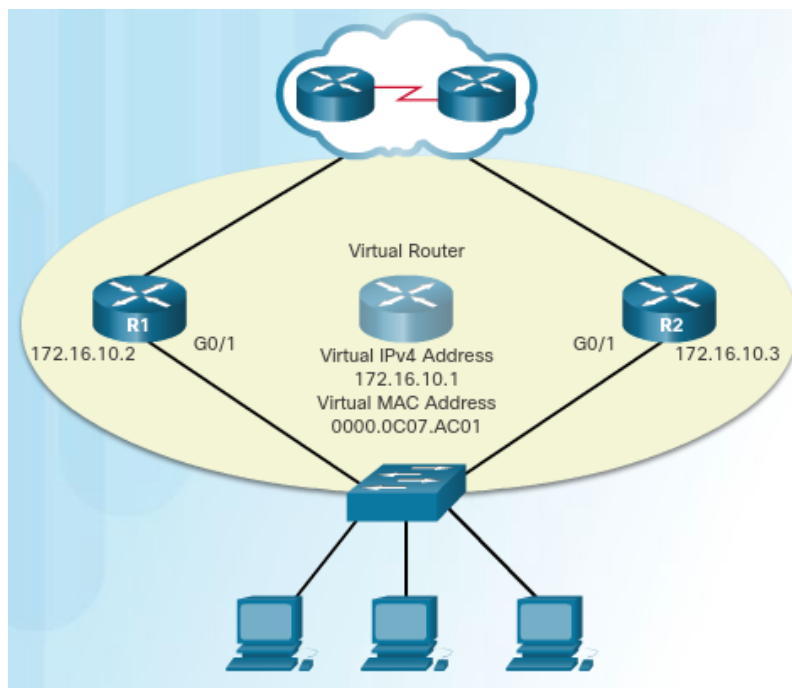


Figure 6.1: HSRP topology

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 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 *types* 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.

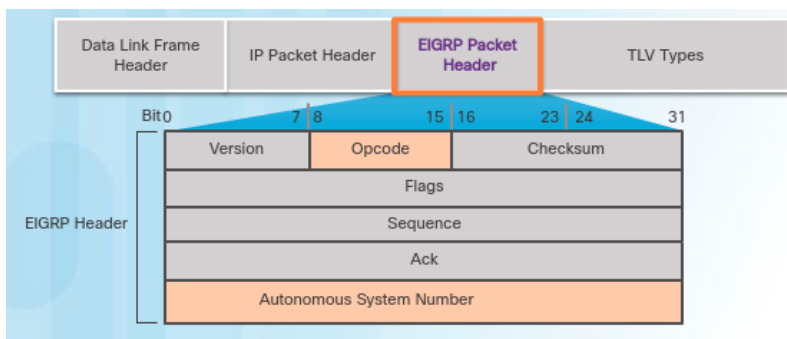


Figure 7.1: EIGRP packet header

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.



Figure 7.2: EIGRP paramters TLV fields

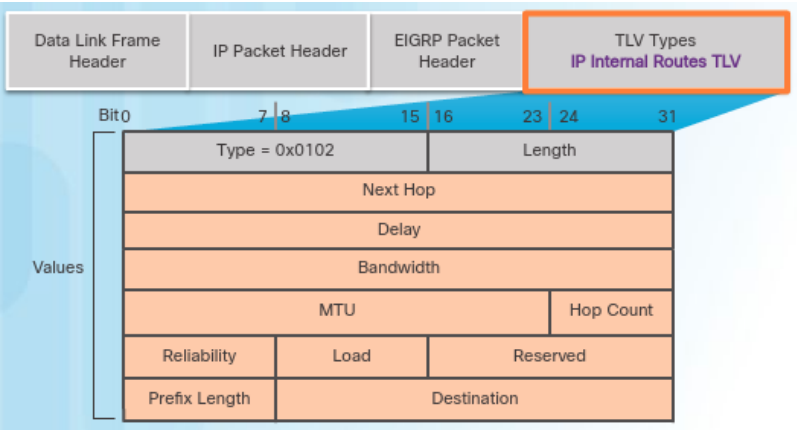


Figure 7.3: EIGRP internal routes TLV fields

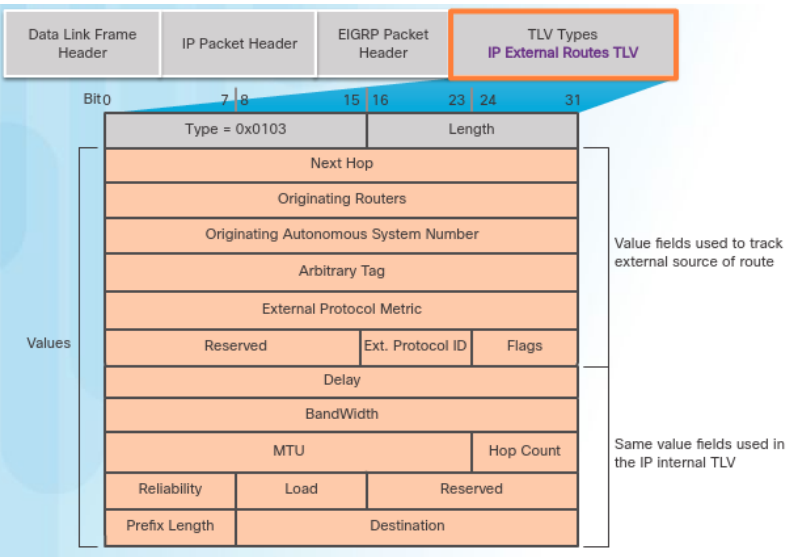


Figure 7.4: EIGRP external route TLV fields

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:

- $K1 (\text{bandwidth}) = 1$
- $K2 (\text{load}) = 0$
- $K3 (\text{delay}) = 0$
- $K4 (\text{reliability}) = 0$
- $K5 (\text{reliability}) = 0$

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 7.1 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
Ethernet	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 7.5).
- **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 7.5). 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.

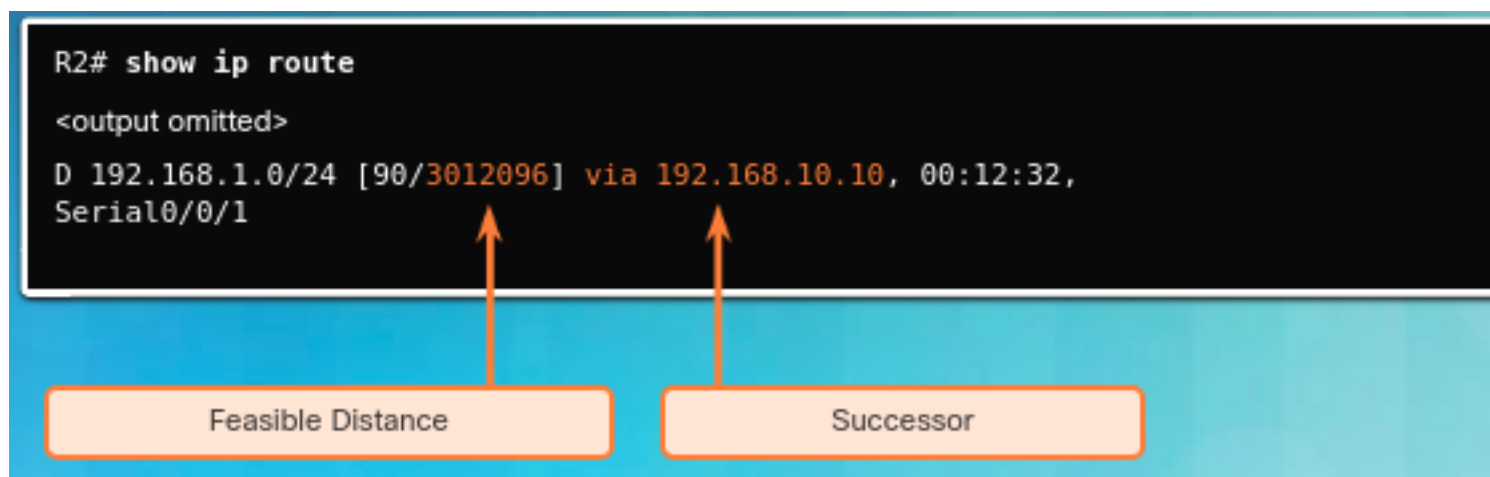


Figure 7.5: Successor and Feasible Distance

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.

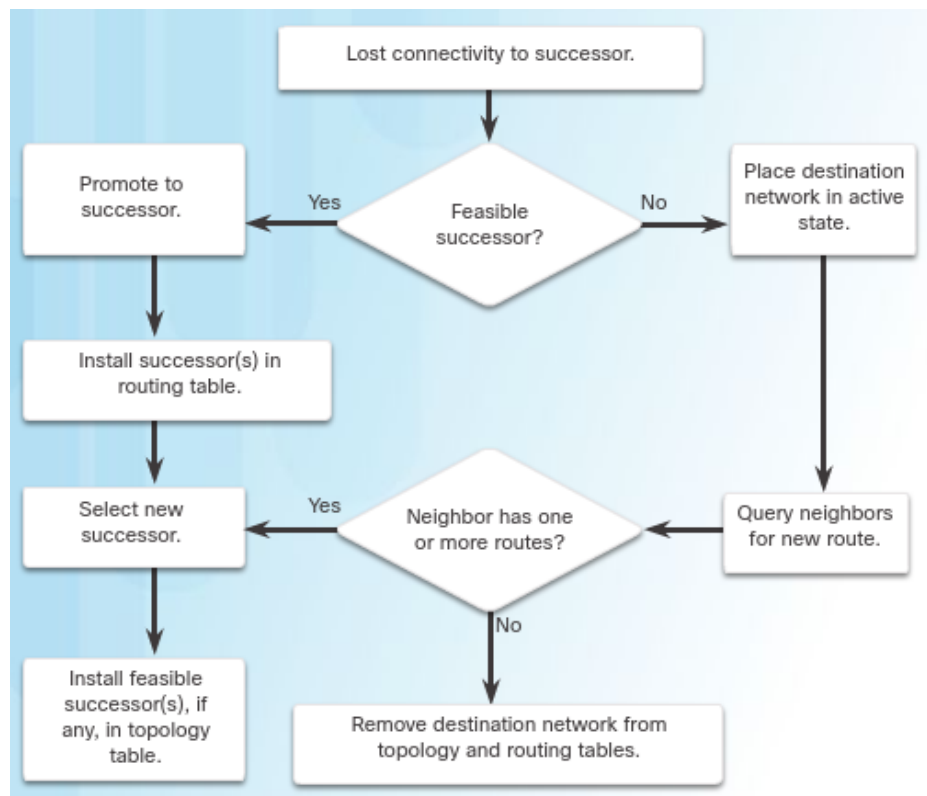


Figure 7.6: DUAL Finite State machine

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 7.6.

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.

Chapter 8

OSPF

8.1 Introduction

8.1.1 Link-state operation

1. **Establish neighbor adjacencies** – An OSPF-enabled router sends Hello packets out all OSPF-enabled interfaces to determine if neighbors are present on those links.
2. **Exchange Link-State Advertisements** – Routers flood their LSAs (which contain the state and cost of each directly connected link) to adjacent neighbors. Those neighbors then immediately flood the LSAs to other directly connected neighbors, until all routers in the area have all LSAs.
3. **Build the Topology Table** – Routers build the topology table (LSDB) based on the received LSAs.
4. **Execute the SPF Algorithm** – Routers use the SPF algorithm to create the SPF tree.
5. **Insert best path to routing table** – From the SPF tree, the best paths are offered to the IP routing table. The route will be inserted into the routing table unless there is a route source to the same network with a lower administrative distance, such as a static route. Routing decisions are made based on the entries in the routing table, not LSDB.

8.1.2 OSPF network types

OSPF defines five network types:

- **Point-to-point** – Two routers interconnected over a common link. No other routers are on the link. (Figure 8.1(a))
- **Multiaccess** – also called Broadcast multiaccess, Multiple routers interconnected over an Ethernet network. (Figure 8.1(b))
- **Nonbroadcast multiaccess (NBMA)** – Multiple routers interconnected in a network that does not allow broadcasts, such as Frame Relay. (Figure 8.1(c))
- **Point-to-multipoint** – Multiple routers interconnected in a hub-and-spoke topology over an NBMA network (Figure 8.1(d)).
- **Virtual links** – Special OSPF network used to interconnect distant OSPF areas to the backbone area (Figure 8.1(e)). In this scenario, area 51 cannot connect directly to area 0. A special OSPF area must be configured to connect area 51 to area 0. The R1 and R2 in area 1 must be configured as a virtual link.

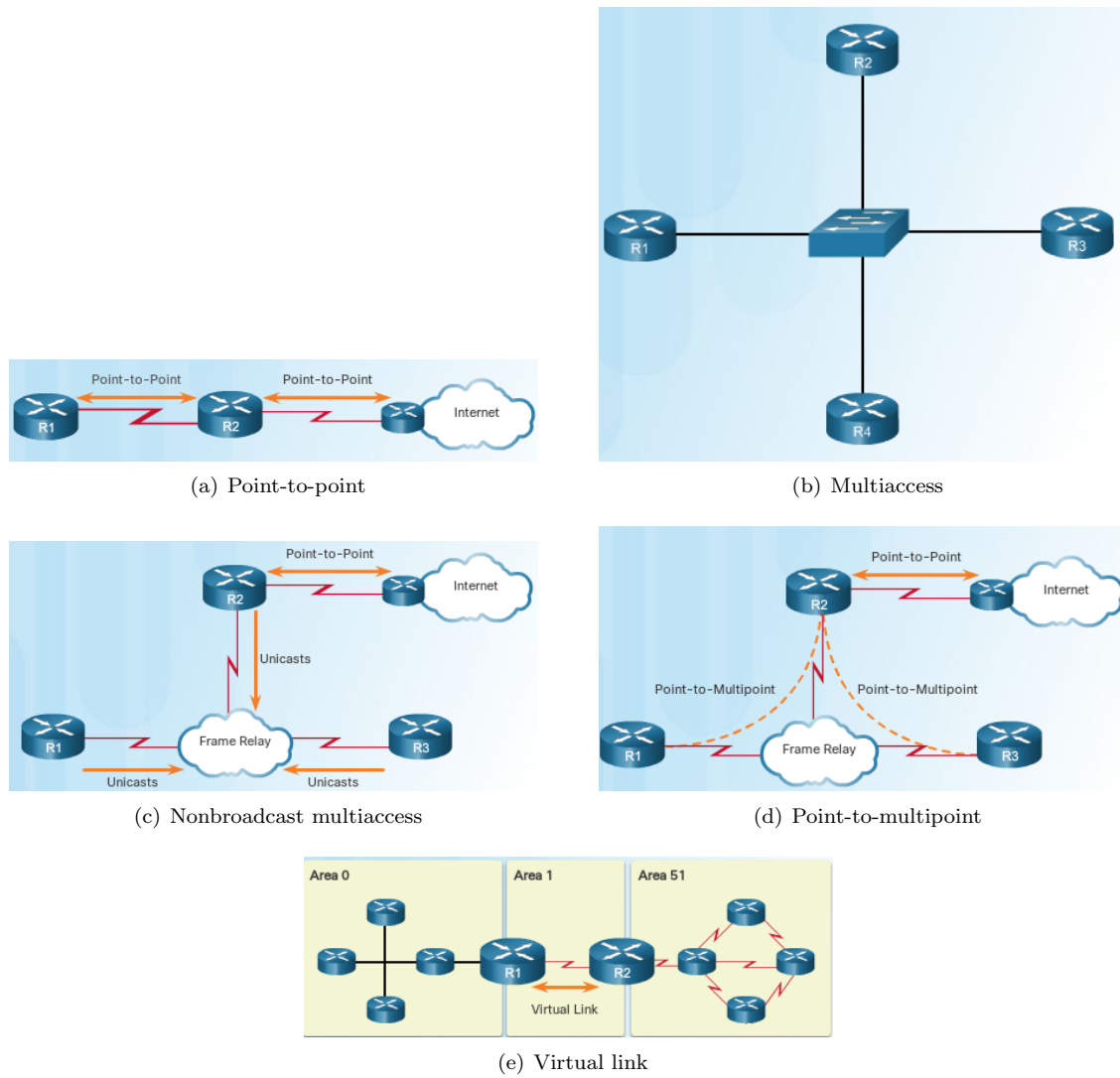


Figure 8.1: OSPF network types

8.1.3 OSPF cost

OSPF uses cost as a metric, where lower cost indicates a better path. The cost of an interface is inversely proportional to the bandwidth of the interface:

$$\text{cost} = \frac{\text{Reference bandwidth}}{\text{Interface bandwidth}}$$

The default reference bandwidth is 100 Mb/s, therefore the formula is:

$$\text{cost} = \frac{10^8}{\text{Interface bandwidth in bps}}$$

Notice that any interfaces faster than 100 Mb/s share the same cost 1, because the OSPF cost value must be an integer. To avoid this, changing reference bandwidth to a higher value than 100 Mb/s is required. *Note!* Changing the reference bandwidth does not actually affect the physical bandwidth of the device; rather, it simply affects the calculation used to determine the metric.

8.2 Protocol components

The OSPF routing protocol has three main components:

- Data structures
- Routing protocol messages
- Algorithm

8.2.1 Data structure

Data structures are the tables or databases that OSPF builds in order to operate. OSPF creates and maintains three databases. These databases are kept and maintained in RAM.

- **Adjacency database** (neighbor table) is a list of all neighbor routers, and unique for each router. It can be viewed using `show ip ospf neighbor` command.
- **Link-state database or LSDB** (topology table) shows the network topology and is identical for all routers in one area. It can be viewed using `show ip ospf database` command.
- **Forwarding database** (Routing table) is a list of best routes to reach networks.

8.2.2 Messages

OSPF messages are transmitted to multicast address 01-00-5E-00-00-05 and 01-00-5E-00-00-06 in MAC address, 224.0.0.5 and 224.0.0.4 in IPv4, or FF02::5 and FF02::6 in IPv6. The protocol field in IP packet header is set to 89 for OSPF protocol. OSPF uses five types of packets to convey routing information:

- **Hello packets** establish neighbor adjacency, and facilitate the DR, BDR election in multiaccess network.
- **Database description (DBD) packets** contain an abbreviated LSDB.
- **Link-State Request (LSR) packets** request additional information about network.
- **Link-State Update (LSU) packets** are sent only to neighbors every 30 minutes, or as a response to LSRs, or when a change is perceived.
- **Link-State Acknowledgment (LSAck) packets** are used to confirm receipt of the LSU.

Hello and dead intervals

The frequency at which the router sends hello packets is specified by hello interval in packet header. The default hello interval on point-to-point and multiaccess network is 10 seconds, on NBMA is 30 seconds. Another important timer is dead interval, which is the period that the router waits to receive a hello packet before declaring the neighbor down. By default, dead interval is four times hello interval. When configuring these two timers on a router, please note that dead interval must be larger than hello interval, and they must be identical on peer routers.

Link-State Advertisement

The link-state advertisement (LSA) is a basic communication means of the OSPF routing protocol. It describes a building block of the OSPF LSDB. Individually, they act as database records and provide specific OSPF network details.

LSAs are not Link-State Packets, they are actually packaged inside LSU and DBD packets to convey different kinds of routing information. The use of terms LSU and LSA can sometimes confusing because these terms are often used interchangeably. However, they are different: *An LSU contains one or more LSAs.*

LSA has its own header, which includes link-state type, link's cost, sequence number, the address of advertising router, and link-ID. The *link ID* field identifies the piece of the routing domain based on LSA type (see table 8.1).

Table 8.1: Link-State Advertisement types

LSA type	Sending router	Description	Flooding area	Link ID
1	all routers	Introduce directly connected networks to its neighbors	one area	router ID of the originating router
2	DR	Give other routers info about multiaccess network	one area	IP interface address of DR
3	ABR	Propagates info of each area to all other routers	between areas	network address
4	ABR	Identify ASBR and provide the route to it	entire routing domain	router ID of ASBR
5	ASBR, ABR	Advertise external network	entire routing domain	external network address

8.2.3 Algorithm

When an OSPF router is initially connected to a network, it goes through the following states in order:

1. **Down state** (send hello packets but not able to receive them)
2. **Init state** (hello packets are received)
3. **Two-way state** (elect a DR and a BDR)
4. **ExStart state** (decide which router will send the DBD packets first, the router with higher router ID will be the first one to send DBD packets)
5. **Exchange state** (exchange DBD packets)
6. **Loading state** (if the information in DBD packets is different from the LSDB, a router will transition to this state to gain additional route information, using LSRs)
7. **Full state** (reach convergence)

8.3 DR election

Multiaccess network Multiaccess networks can create challenges the flooding of LSAs. Ethernet network interconnects routers over a common link, therefore each router considers all counterparts in multiaccess network are its neighbors and establish adjacency to each of them (see figure 8.2). This could lead to extensive flooding of LSPs when OSPF is initialized or when topology changes occur.

Designated Router (DR) The solution to managing the number of adjacencies and the flooding of LSAs on a multiaccess network is the Designated Router (DR). On multiaccess networks, OSPF elects a DR to be the collection and distribution point for LSAs sent and received. The router ID of DR can be viewed using `show ip ospf interface` command on any routers within the multiaccess network.

Backup Designated Router (BDR) A BDR is also elected in case the DR fails. The BDR listens passively to this exchange and maintains a relationship with all the routers. If the DR stops producing Hello packets, the BDR promotes itself and assumes the role of DR.

DROTHER All other routers become DROTHER (a router that is neither the DR nor the BDR). DROTHERs only form full adjacencies with the DR and BDR in the network (see figure 8.3). Instead of flooding LSAs to all routers in the network, DROTHERs only send their LSAs to the DR and BDR using the multicast address 224.0.0.6 (all DR routers).

To “keep in touch” with neighbor routers, DROTHER still form 2-way adjacencies with any DROTHERs that join the multiaccess network. This means that all DROTHER routers in the multiaccess network still receive Hello packets from all other DROTHER routers. In this way, they are aware of all routers in the network.

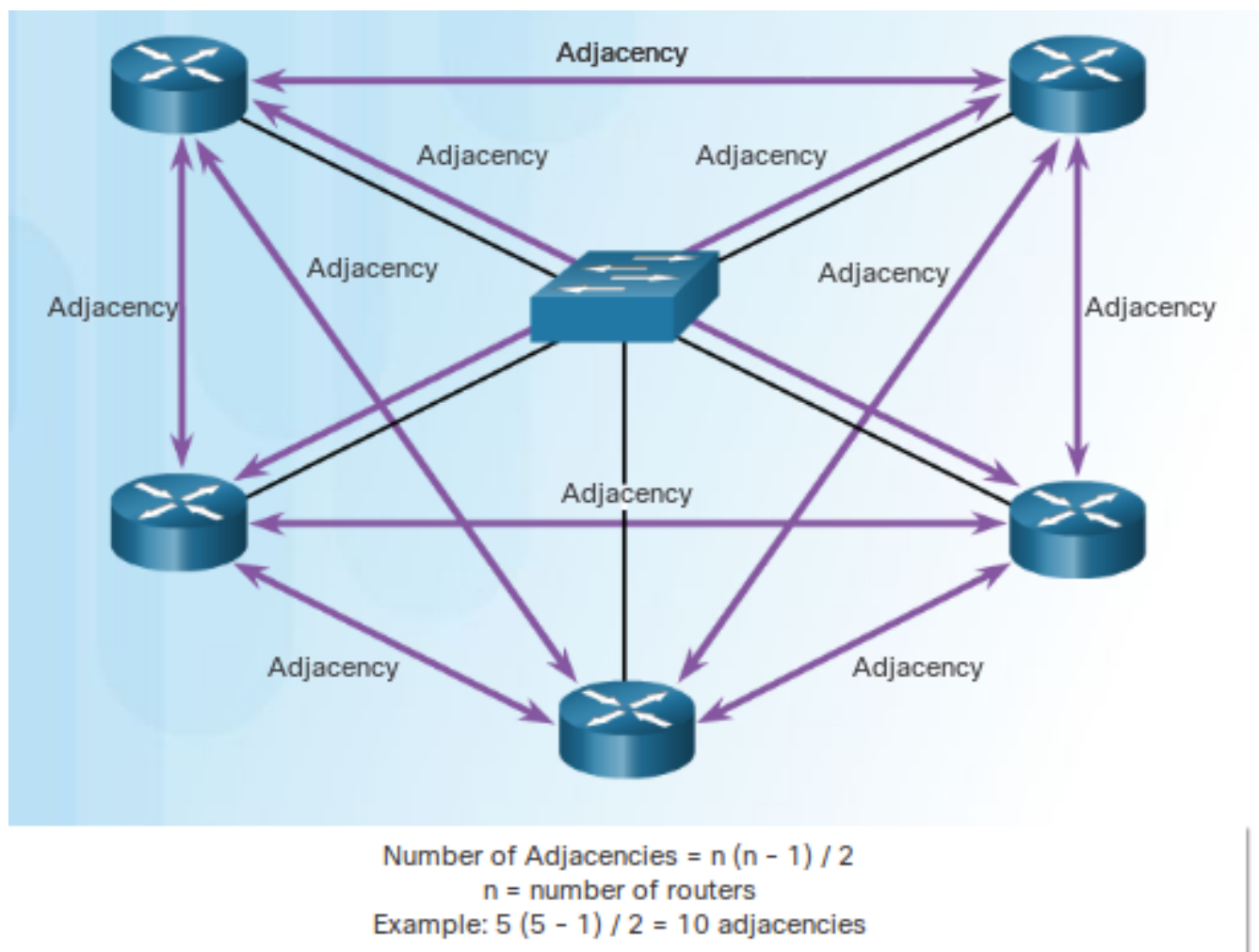


Figure 8.2: Creating adjacencies with every neighbors in multiaccess network

The OSPF DR and BDR election decision is based on the following criteria, in sequential order:

1. The routers in the network elect the router with the highest interface priority as the DR. The router with the second highest interface priority is elected as the BDR.
2. If the interface priorities are equal, then the router with the highest router ID is elected the DR. The router with the second highest router ID is the BDR.

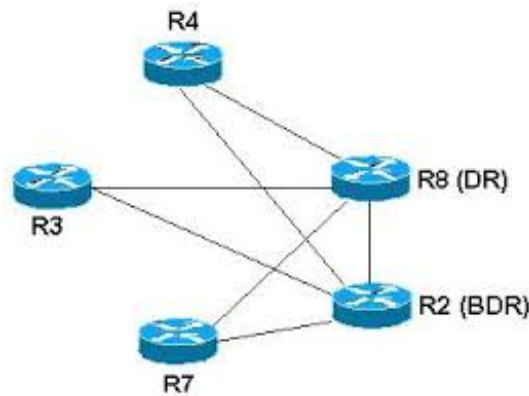


Figure 8.3: DROTHERs only form full adjacencies with the DR and BDR in the network

OSPF DR and BDR elections are not pre-emptive. If a new router with a higher priority or higher router ID is added to the network after the DR and BDR election, the newly added router does not take over the DR or the BDR role. This is because those roles have already been assigned. The addition of a new router does not initiate a new election process.

If the DR fails, the BDR is automatically promoted to DR. This is the case even if another DROTHER with a higher priority or router ID is added to the network after the initial DR/BDR election. However, after a BDR is promoted to DR, a new BDR election occurs and the DROTHER with the higher priority or router ID is elected as the new BDR. After the DR is elected, it remains the DR until one of the following events occurs:

- The DR fails
- The OSPF process on the DR fails or is stopped
- The multiaccess interface on the DR fails or is shutdown

8.4 Multiarea OSPF

8.4.1 Single-area vs Multi-area OSPF

An OSPF area is a group of routers that share the same link-state information in their LSDBs. Single-area OSPF is useful in smaller networks, however, if an area becomes too big, the following issues must be addressed:

- Large routing table (no summarization by default)
- Large link-state database (contains the entire network topology)
- Frequent SPF algorithm calculations (even small changes in a specific area of the network makes the routers recalculate the SPF algorithm and update the LSDB for the entire routing domain.)

Multiarea OSPF have these advantages:

- Smaller routing table (There are fewer routing table entries as network addresses can be summarized between areas.)
- Reduced link-state update overhead (Fewer routers exchanging LSAs because LSA flooding stops at the area boundary.)
- Reduced frequency of SPF calculations (Routing still occurs between the areas, however, the CPU intensive routing operation of recalculating the SPF algorithm is done only for routes within an area.)

8.4.2 Introduction to multiarea OSPF

Two-layer area hierarchy

To make OSPF more efficient and scalable, Multiarea OSPF is implemented in a two-layer area hierarchy:

- **Backbone (Transit) area** – A backbone area directly connected with all other areas. All traffic moving from one area to another area must traverse the backbone area. Generally, end users are not found within a backbone area. The backbone area is also called OSPF area 0.
- **Regular (Non-backbone) area** – Connects users and resources. By default, a regular area does not allow traffic from another area to use its links to reach other areas. All traffic from other areas must cross a transit area.

Types of routers

There are four different types of OSPF routers:

- Internal router (have all interfaces in the same area)
- Backbone router (reside in backbone area)
- Area border router or ABR (have interfaces attached to multiple areas)
- Autonomous System Boundary Router or ASBR (have at least one interface attached to an external network)

Multiarea configuration

The optimal number of routers per area varies based on factors such as network stability, but Cisco recommends the following guidelines:

- An area should have no more than 50 routers.
- A router should not be in more than three areas.
- Any single router should not have more than 60 neighbors.

Also keep the following notes in mind:

- Propagating type 3 and 5 LSAs can cause significant flooding problems. For this reason, it is strongly recommended that manual route summarization be configured on the ABRs and ASBR.
- Receiving a type 3 LSA does not cause a router to run the SPF algorithm, but the routes being advertised in the type 3 LSAs are appropriately updated to the routing table.
- In the routing table, OSPF intra-area routes start with O, inter-area routes start with O IA, external routes start with O E1 (or O E2).
- SPF algorithm calculates the best paths in order: calculate intra-area routes → calculate inter-area routes → calculate external routes.

Chapter 9

ACL

9.1 ACL Operation Overview

An ACL contains a sequential list of permit or deny statements, known as access control entries (ACEs). ACEs are also commonly called ACL statements.

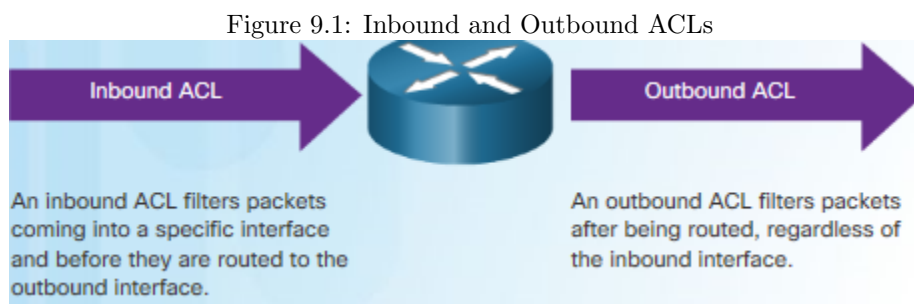
9.1.1 ACEs Logic Operations

ACLs are processed in a top down manner. When an ACL is inspected, if the information in a packet header and an ACL statement match, the remaining statements are not examined, and the packet is either denied or permitted through as specified by the ACL. If a packet header does not match an ACL statement, the packet is tested against the next statement in the list. This matching process continues until the end of the list is reached. If no conditions match, the address is rejected.

In a nut shell, ACL always stops testing conditions after the first match, therefore, the order of the ACEs is critical. At the end of every ACL is a statement is an implicit deny any statement and because of this statement, an ACL should have at least one permit statement in it; otherwise, the ACL blocks all traffic.

9.1.2 Inbound and Outbound ACL Logic

The Figure 9.1 shows the logic of routing and ACL processes. When a packet arrives at a router interface, the router checks for an ACL on the inbound interface. If an ACL exists, the packet is tested against the statements in the list.



If the packet matches a statement, the packet is either permitted or denied. If the packet is accepted, it is then checked against routing table entries to determine the destination interface. If a routing table entry exists for the destination, the packet is then switched to the outgoing interface, otherwise the packet is dropped.

Next, the router checks whether the outgoing interface has an ACL. If an ACL exists, the packet is tested against the statements in the list. If the packet matches a statement, it is either permitted or denied.

If there is no ACL or the packet is permitted, the packet is encapsulated in the new Layer 2 protocol and forwarded out the interface to the next device.

9.1.3 Numbered and Named ACLs

Standard and extended ACLs can be created using either a number or a name to identify the ACL and its list of statements.

Numbered ACL Assign a number based on the following rules:

- (1 to 99) and (1300 to 1999): Standard ACL
- (100 to 199) and (2000 to 2699): Extended ACL

Named ACL Assign a name based on the following rules:

- Cannot contain spaces or punctuation
- Names are case-sensitive
- Can contain alphanumeric characters
- It is suggested that the name be written in CAPITAL LETTER

9.2 Standard ACL

9.2.1 Overview

A standard IPv4 ACL can filter traffic based on source IP addresses only. Unlike an extended ACL, it cannot filter traffic based on Layer 4 ports.

Because standard ACLs do not specify destination addresses, place them as close to the destination as possible. If a standard ACL was placed at the source of the traffic, the "permit" or "deny" will occur based on the given source address no matter where the traffic is destined.

9.2.2 Standard ACL placement

In the figure 9.2, the administrator wants to prevent traffic originating in the 192.168.10.0/24 network from reaching the 192.168.30.0/24 network.

Following the basic placement guidelines of placing the standard ACL close to the destination, the figure shows two possible interfaces on R3 to apply the standard ACL:

- R3 S0/0/1 interface - Applying a standard ACL to prevent traffic from 192.168.10.0/24 from entering the S0/0/1 interface will prevent this traffic from reaching 192.168.30.0/24 and all other networks that are reachable by R3. This includes the 192.168.31.0/24 network. Because the intent of the ACL is to filter traffic destined only for 192.168.30.0/24, a standard ACL should not be applied to this interface.
- R3 G0/0 interface - Applying the standard ACL to traffic exiting the G0/0 interface will filter packets from 192.168.10.0/24 to 192.168.30.0/24. This will not affect other networks that are reachable by R3. Packets from 192.168.10.0/24 will still be able to reach 192.168.31.0/24.

9.3 Extended ACLs

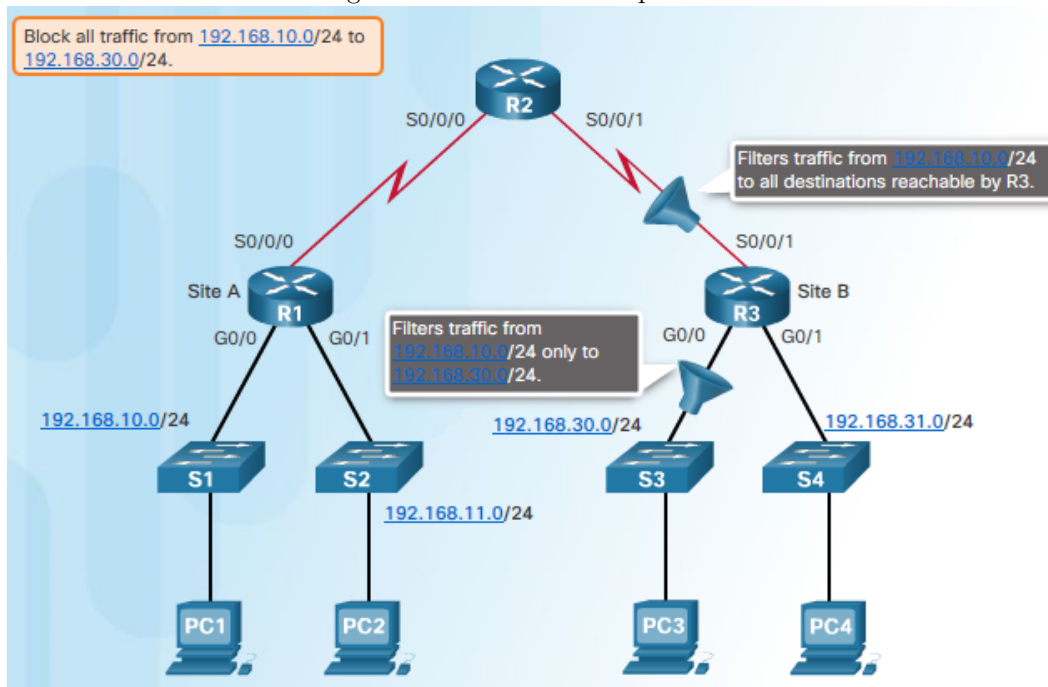
9.3.1 Overview

Extended ACLs filter packets based on:

- Protocol type (e.g. IP, ICMP, UDP, TCP)
- Source and destination IP addresses
- Source and destination TCP and UDP ports (HTTP port 80, SSH port 22, etc.)

Extended ACLs are used more often than standard ACLs because they provide a greater degree of control. We usually locate extended ACLs as close as possible to the source of the traffic to be filtered. This way, undesirable traffic is denied close to the source network without crossing the network infrastructure.

Figure 9.2: Standard ACL placement



9.3.2 Extended ACL Placement

In figure 9.3, the administrator wants to deny Telnet and FTP traffic from the .11 network to Company B's 192.168.30.0/24 (.30, in this example) network. At the same time, all other traffic from the .11 network must be permitted to leave Company A without restriction.

A better solution is to place an extended ACL on R1. There are two possible interfaces on R1 to apply the extended ACL:

- R1 S0/0/0 interface (outbound) - One possibility is to apply an extended ACL outbound on the S0/0/0 interface. Because the extended ACL can examine both source and destination addresses, only FTP and Telnet packets from 192.168.11.0/24 will be denied. Other traffic from 192.168.11.0/24 and other networks will be forwarded by R1. The disadvantage of placing the extended ACL on this interface is that all traffic exiting S0/0/0 must be processed by the ACL including packets from 192.168.10.0/24.
- R1 G0/1 interface (inbound) - Applying an extended ACL to traffic entering the G0/1 interface means that only packets from the 192.168.11.0/24 network are subject to ACL processing on R1. Because the filter is to be limited to only those packets leaving the 192.168.11.0/24 network, applying the extended ACL to G0/1 is the best solution.

9.4 IPv6 ACLs

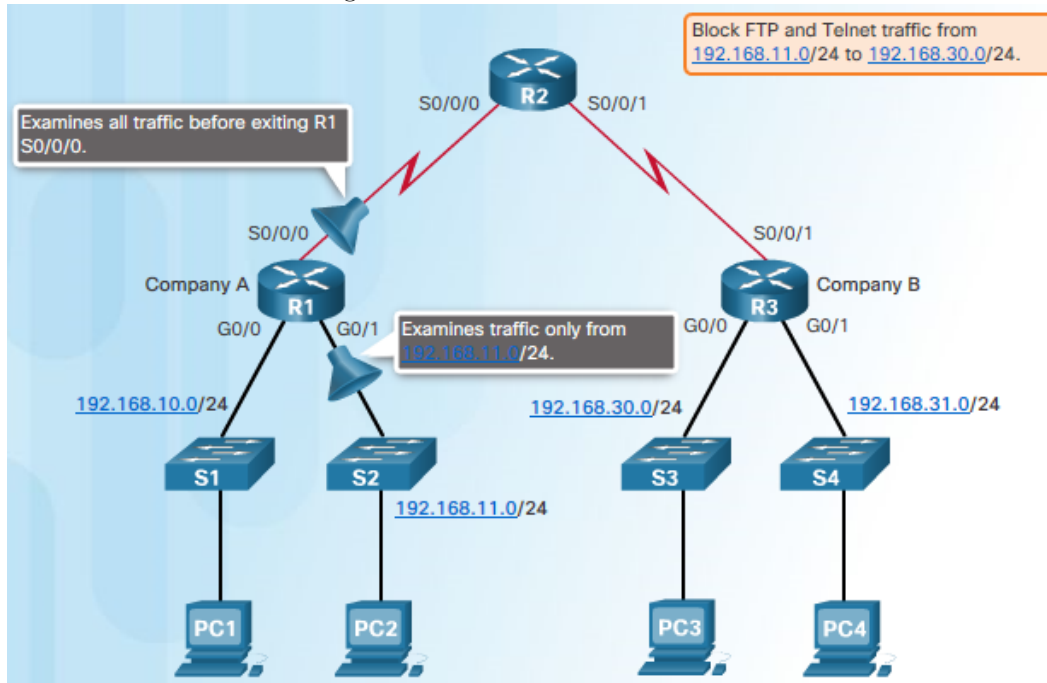
In IPv4 there are two types of ACLs, standard and extended and both types of ACLs can be either numbered or named ACLs. With IPv6, there is only one type of ACL, which is equivalent to an IPv4 extended named ACL and there are no numbered ACLs in IPv6.

Because IPv6 ACLs must be configured with both a source and a destination, they should be applied closest to the source of the traffic.

An IPv4 ACL and an IPv6 ACL cannot share the same name. There are three significant differences between IPv4 and IPv6 ACLs:

- The command used to apply an IPv6 ACL to an interface is `ipv6 traffic-filter` command.
- IPv6 ACLs do not use wildcard masks but instead specifies the prefix-length

Figure 9.3: Extended ACL Placement



- Besides `deny ipv6 any any`, An IPv6 ACL adds two implicit permit statements at the end of each IPv6 access list: `permit icmp any any nd-na` and `permit icmp any any nd-ns`

9.5 Configurations

Example 9.1. The figure shows an example of an ACL designed to permit a single network. Only traffic from the 192.168.10.0/24 network will be permitted out the Serial 0/0/0 interface.

```
R1(config)# access-list 1 permit 192.168.10.0 0.0.0.255
R1(config)# interface s0/0/0
R1(config-if)# ip access-group 1 out
```

Example 9.2. Figure below shows the commands used to configure a standard named ACL on router R1, interface G0/0, which denies host 192.168.11.10 access to the 192.168.10.0 network.

```
R1(config)# ip access-list standard NO_ACCESS
R1(config-std-nacl)# deny host 192.168.11.10
R1(config-std-nacl)# permit any
R1(config-std-nacl)# exit
R1(config)# interface g0/0
R1(config-if)# ip access-group NO_ACCESS out
```

Example 9.3. Design an IPv4 named access list HQServer to prevent any computers attached to the g0/0 interface of the Branch router from accessing HQServer.pka (172.16.0.1). All other traffic is permitted. Configure the access list on the appropriate router, apply it to the appropriate interface and in the appropriate direction.

```
Branch(config)#ip access-list extended HQServer
Branch(config-ext-nacl)#deny ip any host 172.16.0.1
Branch(config-ext-nacl)#permit ip any any
Branch(config-ext-nacl)#exit
Branch(config)#int g0/0
Branch(config-if)#ip access-group HQServer in
```

Example 9.4. Design an IPv4 named access list BranchServer to prevent any computers attached to the Gigabit Ethernet 0/0 interface of the HQ router from accessing the HTTP and HTTPS service of the Branch server

(172.16.128.1/20). All other traffic is permitted. Configure the access list on the appropriate router, apply it to the appropriate interface and in the appropriate direction.

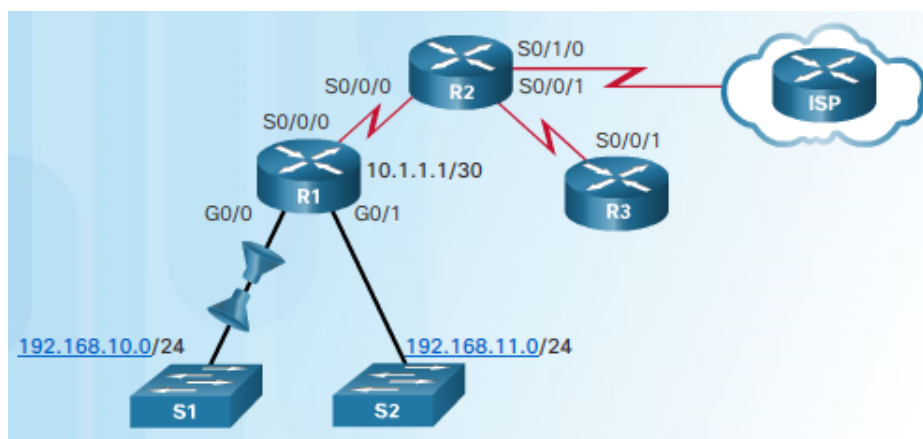
```
HQ(config)#ip access-list extended BranchServer
HQ(config-ext-nacl)#deny tcp any host 172.16.128.1 eq 80
HQ(config-ext-nacl)#deny tcp any host 172.16.128.1 eq 443
HQ(config-ext-nacl)#permit ip any any
HQ(config-ext-nacl)#exit
HQ(config)#int g0/0
HQ(config-if)#ip access-group HQServer in
```

Example 9.5. Design an IPv6 access-list named NO-B1 to prevent any IPv6 traffic originating on B1 (2001:DB8:ACAD:B1::2/64) to reach the BranchServer.pka (2001:DB8:ACAD:B2::3/64). No traffic should be permitted from B1 to BranchServer.pka. Apply the IPv6 access to the most appropriated location (interface and direction).

```
Branch(config)#ipv6 access-list NO-B1
Branch(config-ipv6-acl)#deny ipv6 host 2001:DB8:ACAD:B1::2 host 2001:DB8:ACAD:B2::3
Branch(config-ipv6-acl)#permit ipv6 any any
Branch(config-ipv6-acl)#exit
Branch(config)#int g0/1
Branch(config-if)#ipv6 traffic-filter NO-B1 out
```

Example 9.6. The network administrator configured an ACL to allow users from the 192.168.10.0/24 network to browse both insecure and secure websites. In this topology (figure 9.4) the interface closest to the source of the target traffic is the G0/0 interface of R1. Web request traffic from users on the 192.168.10.0/24 LAN is inbound to

Figure 9.4: Extended ACL example



(a)

```
R1(config)# access-list 103 permit tcp 192.168.10.0 0.0.0.255 any eq 80
R1(config)# access-list 103 permit tcp 192.168.10.0 0.0.0.255 any eq 443
R1(config)# access-list 104 permit tcp any 192.168.10.0 0.0.0.255 established
R1(config)# interface g0/0
R1(config-if)# ip access-group 103 in
R1(config-if)# ip access-group 104 out
```

(b)

the G0/0 interface. Return traffic from established connections to users on the LAN is outbound from the G0/0 interface. The example applies the ACL to the G0/0 interface in both directions. The inbound ACL, 103, checks for the type of traffic. The outbound ACL, 104, checks for return traffic from established connections. This will restrict 192.168.10.0 Internet access to allow only website browsing.

Example 9.7. Configure an extended IPv4 ACL named INTOHQ such that:

- Allow any hosts from the Internet to access the County DNS Svr. There should be two ACEs, one for TCP and the other UDP. Both use port 53.
- Allow any hosts from the Internet to access the County Web Svr. Only port 80 is needed.
- Allow return TCP traffic from the Internet that was initiated from the hosts in the Central networks to pass (with the established keyword).
- Apply the ACL to the Central S0/0/0 interface.

```
ip access-list extended INTOHQ
permit tcp any host 172.16.10.5 eq 53
permit udp any host 172.16.10.5 eq 53
permit tcp any host 172.16.10.10 eq 80
permit tcp any any established
exit
interface s0/0/0
ip access-group INTOHQ IN
exit
```

Example 9.8. Configure an extended ACL named SNMPACCESS such that

- The SNMP operation runs UDP on port 161.
- Allow only the County-Admin-PC to access the Central router for the SNMP connection.
- SNMP connections from other hosts on the Central LAN should fail.
- Allow all other IP traffic.
- Apply this ACL on the Central router, G0/0 interface.

```
ip access-list extended SNMPACCESS
permit udp host 192.168.10.5 host 192.168.10.1 eq 161
deny udp any host 192.168.10.1 eq 161
permit ip any any
exit
interface g0/0
ip access-group SNMPACCESS in
exit
```

9.6 Troubleshoot

Using the `show access-lists` command to reveal most of the common ACL errors. The most common errors are entering ACEs in the wrong order and not applying adequate criteria to the ACL rules. Following these steps to troubleshoot ACL:

1. Check the criteria of ACL rules
2. Check the order of ACEs
3. Check the direction of ACL (inbound, outbound)
4. Check the location of ACL (which router, which interface). Remember that extended ACLs are placed as close as possible to the source and standard ACLs are placed as close as possible to the destination.

Example 9.9. The 192.168.10.0/24 network cannot use TFTP to connect to the 192.168.30.0/24 network.

```
R3# show access-lists 120
Extended IP access list 120
 10 deny tcp 192.168.10.0 0.0.0.255 any eq telnet
 20 deny tcp 192.168.10.0 0.0.0.255 host 192.168.31.12 eq smtp
 30 permit tcp any any
```


Solution: The 192.168.10.0/24 network cannot use TFTP to connect to the 192.168.30.0/24 network because TFTP uses the transport protocol UDP. Statement 30 in access list 120 allows all other TCP traffic. However, because TFTP uses UDP instead of TCP, it is implicitly denied. Recall that the implied deny any statement does not appear in show access-lists output and therefore matches are not shown. Statement 30 should be **permit ip any any**.

Example 9.10. The 192.168.11.0/24 network can use Telnet to connect to 192.168.30.0/24, but according to company policy, this connection should not be allowed. The results of the show access-lists 130 command indicate that the permit statement has been matched.

```
R1# show access-lists 130
Extended IP access list 130
 10 deny tcp any eq telnet any
 20 deny tcp 192.168.11.0 0.0.0.255 host 192.168.31.12 eq smtp
 30 permit tcp any any (12 match(es))
```

The 192.168.11.0/24 network can use Telnet to connect to the 192.168.30.0/24 network because the Telnet port number in statement 10 of access list 130 is listed in the wrong position in the ACL statement. Statement 10 currently denies any source packet with a port number that is equal to Telnet. To deny Telnet traffic inbound on G0/1, deny the destination port number that is equal to Telnet, for example, 10 deny tcp 192.168.11.0 0.0.0.255 192.168.30.0 0.0.0.255 eq telnet.

Chapter 10

Network Security and Monitoring

10.1 Security attacks

10.1.1 CDP Reconnaissance Attack

The Cisco Discovery Protocol (CDP) is a proprietary Layer 2 link discovery protocol. It is enabled on all Cisco devices by default. CDP broadcasts are sent unencrypted and unauthenticated. Therefore, an attacker could interfere with the network infrastructure by sending crafted CDP frames containing bogus device information to directly-connected Cisco devices.

To mitigate the exploitation of CDP, limit the use of CDP on devices or ports. For example, disable CDP on edge ports that connect to untrusted devices.

10.1.2 Telnet Attacks

There are two types of Telnet attacks:

- Brute Force Password Attack: The attacker tries to discover the administrative password.
- Telnet DoS Attack: The attacker continuously requests Telnet connections in an attempt to render the Telnet service unavailable and preventing an administrator from remotely accessing a switch.

10.1.3 MAC Address Table Flooding Attack

MAC address tables are limited in size. MAC flooding attacks exploit this limitation with fake source MAC addresses until the switch MAC address table is full. When the MAC address table becomes full of fake MAC addresses, the switch enters into what is known as fail-open mode. In this mode, the switch broadcasts all frames to all machines on the network. As a result, the attacker can capture all of the frames, even frames that are not addressed to its MAC address table.

Configure **port security** on the switch to mitigate MAC address table overflow attacks.

10.1.4 VLAN Attacks

The attacker attempts to gain VLAN access by configuring a host to spoof a switch and use the 802.1Q trunking protocol and the Cisco-proprietary Dynamic Trunking Protocol (DTP) feature to trunk with the connecting switch. If successful and the switch establishes a trunk link with the host and the attacker can then access all the VLANs on the switch and hop (i.e., send and receive) traffic on all the VLANs. The best way to prevent basic VLAN attacks:

- Disable DTP (auto trunking) negotiations on non-trunking ports and non-trunking ports using the **switchport nonegotiate** interface configuration command. Additionally explicitly force the access ports by using the **switchport mode access** interface configuration command..
- Manually enable the trunk link on a trunking port using the **switchport mode trunk** interface configuration command.

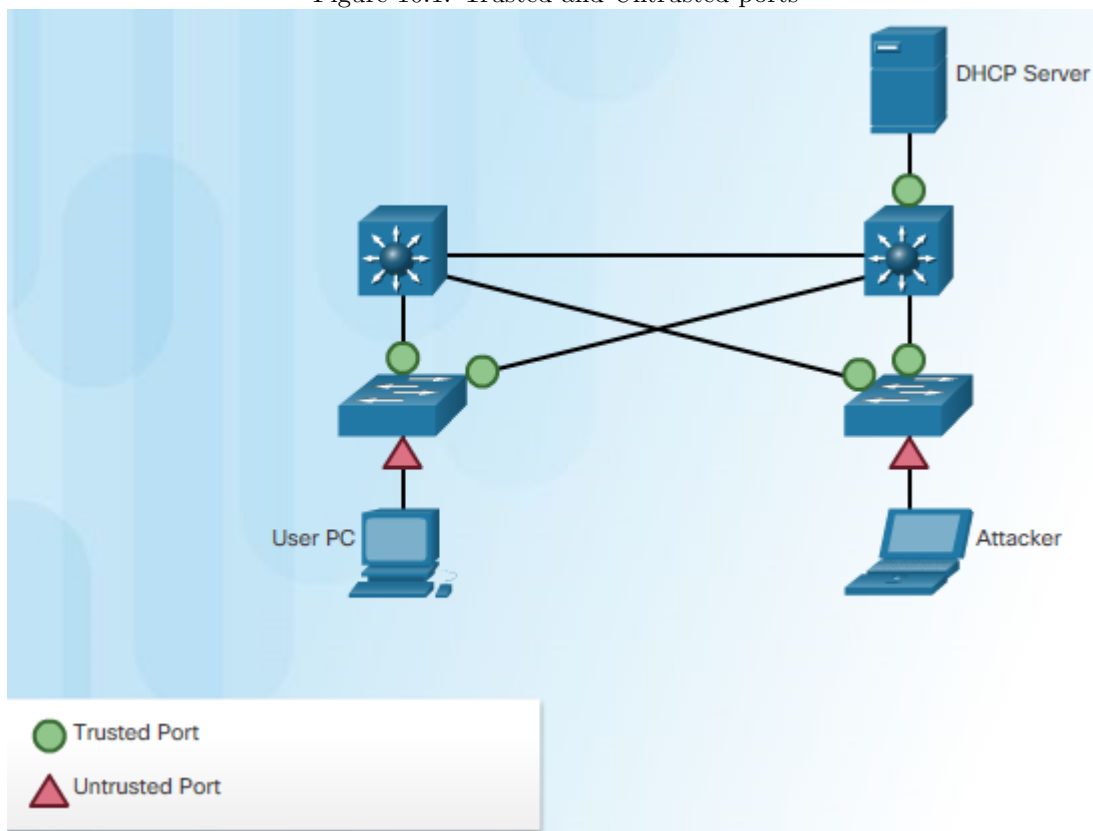
- Disable DTP (auto trunking) negotiations on trunking
- Set the native VLAN to be something other than VLAN 1.
- Disable unused ports and assign them to an unused VLAN.

10.1.5 DHCP Attacks

There are two types of DHCP attacks which can be performed against a switched network:

- DHCP spoofing attack: A DHCP spoofing attack occurs when a rogue DHCP server is connected to the network and provides false IP configuration parameters to legitimate clients. Security best practices recommend using DHCP snooping to mitigate DHCP spoofing attacks.
- DHCP starvation attack: An attacker floods the DHCP server with bogus DHCP requests and eventually leases all of the available IP addresses in the DHCP server pool. After these IP addresses are issued, the server cannot issue any more addresses, and this situation produces a DoS¹ attack as new clients cannot obtain network access.

Figure 10.1: Trusted and Untrusted ports



DHCP snooping recognizes two types of ports (see figure 10.1):

- Trusted DHCP ports: Only ports connecting to upstream DHCP servers should be trusted. These ports should lead to legitimate DHCP servers replying with DHCP Offer and DHCP Ack messages. Trusted ports must be explicitly identified in the configuration.
- Untrusted ports: These ports connect to hosts that should not be providing DHCP server messages. By default, all switch ports are untrusted.

¹A DoS attack is any attack that is used to overload specific devices and network services with illegitimate traffic, thereby preventing legitimate traffic from reaching those resources.

10.1.6 Cisco solution

There are four Cisco switch security solutions to help mitigate Layer 2 attacks:

- Port security prevents MAC address flooding, DHCP starvation
- DHCP snooping prevents DHCP spoofing and DHCP starvation
- DAI (Dynamic ARP inspection) prevents ARP spoofing and ARP poisoning
- IPSG (IP Source Guard) prevents MAC and IP address spoofing

10.1.7 The AAA framework

The Authentication, Authorization, and Accounting (AAA) framework is used to help secure device access.

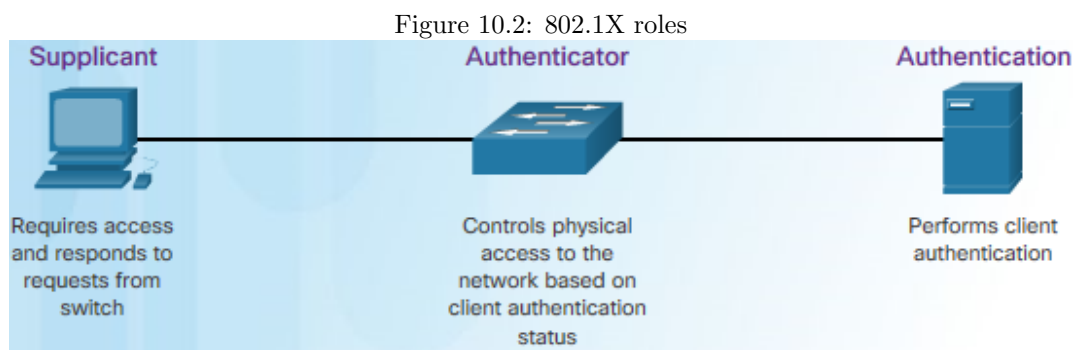
An AAA-enabled router uses either the Terminal Access Controller Access Control System (TACACS+) protocol or the Remote Authentication Dial-In User Service (RADIUS) protocol to communicate with the AAA server. While both protocols can be used to communicate between a router and AAA servers, TACACS+ is considered the more secure protocol. This is because all TACACS+ protocol exchanges are encrypted, while RADIUS only encrypts the user's password. RADIUS does not encrypt user names, accounting information, or any other information carried in the RADIUS message.

Cisco provides two common methods of implementing AAA services:

- Local AAA Authentication - Local AAA uses a local database for authentication. This method is sometimes known as self-contained authentication. This method stores usernames and passwords locally in the Cisco router, and users authenticate against the local database. Local AAA is ideal for small networks.
- Server-Based AAA Authentication - Server-based AAA authentication is a much more scalable solution. With server-based method, the router accesses a central AAA server. The AAA server contains the usernames and password for all users and serves as a central authentication system for all infrastructure devices.

10.1.8 802.1X

The IEEE 802.1X standard defines a port-based access control and authentication protocol. IEEE 802.1X restricts unauthorized workstations from connecting to a LAN through publicly accessible switch ports. With 802.1X port-



based authentication, the devices in the network have specific roles, as shown in the figure 10.2:

- Client (Supplicant): The device is a PC running 802.1X-compliant client software.
- Switch (Authenticator): This controls physical access to the network based on the authentication status of the client. The switch requests identifying information from the client, verifies that information with the authentication server, and relays a response to the client.
- Authentication server: validates the identity of the client and notifies the switch or other authenticator such as a wireless access point whether the client is authorized to access the LAN and switch services.

10.2 SNMP

10.2.1 Introduction to SNMP

Simple Network Management Protocol (SNMP) was developed to allow administrators to manage nodes such as servers, workstations, routers, switches, and security appliances, on an IP network. The SNMP system consists of three elements:

- **SNMP manager:** is part of a network management system (NMS), run SNMP management software.
- **SNMP agents** (managed node): responsible for providing access to the local MIB. The SNMP agent and MIB reside on SNMP client devices.
- **MIB** (Management Information Base): store data about the device and operational statistics

10.2.2 SNMP operation

SNMP requests

The SNMP manager uses the get and set actions to perform the operations described in the table in Figure 10.3.

Figure 10.3: SNMP operations

Operation	Description
get-request	Retrieves a value from a specific variable.
get-next-request	Retrieves a value from a variable within a table; the SNMP manager does not need to know the exact variable name. A sequential search is performed to find the needed variable from within a table.
get-bulk-request	Retrieves large blocks of data, such as multiple rows in a table, that would otherwise require the transmission of many small blocks of data. (Only works with SNMPv2 or later.)
get-response	Replies to a get-request , get-next-request , and set-request sent by an NMS.
set-request	Stores a value in a specific variable.

SNMP Agent Traps

An NMS periodically polls the SNMP agents residing on managed devices, by querying the device for data using the get request. Using this process, a network management application can collect information to monitor traffic loads and to verify device configurations of managed devices.

Periodic SNMP polling does have disadvantages. First, there is a delay between the time that an event occurs and the time that it is noticed (via polling) by the NMS. Second, there is a trade-off between polling frequency and bandwidth usage.

To mitigate these disadvantages, it is possible for SNMP agents to generate and send traps to inform the NMS immediately of certain events. Traps are unsolicited messages alerting the SNMP manager to a condition or event on the network.

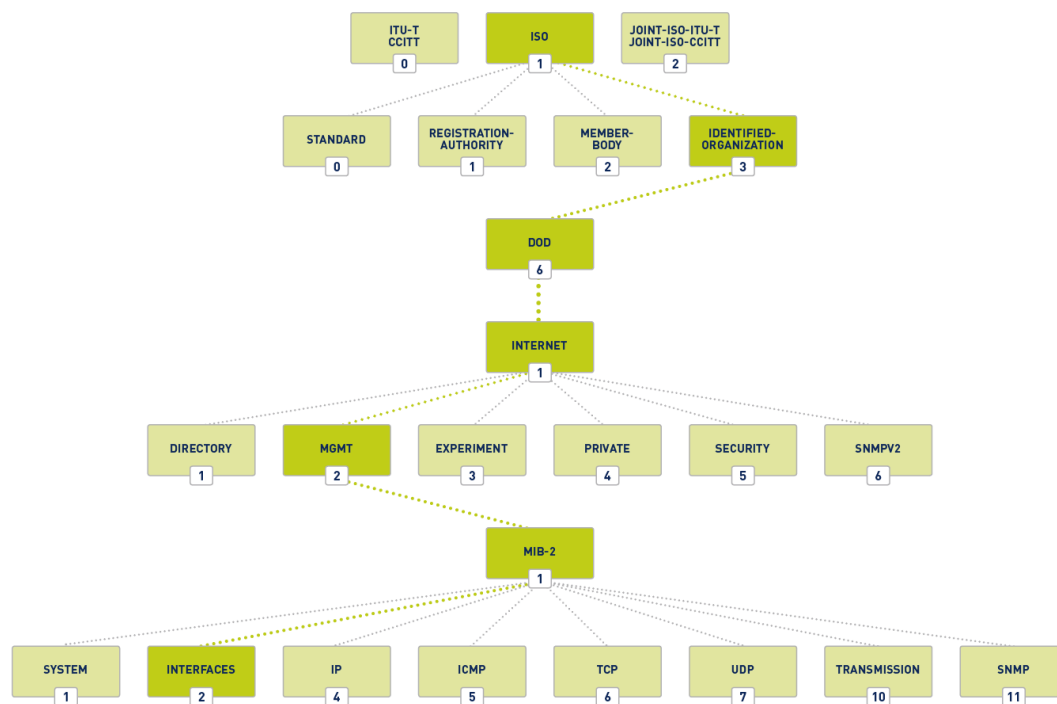
Community string

SNMPv1 and SNMPv2c use community strings that control access to the MIB. Community strings are plaintext passwords. There are two types of community strings: Read-only (**ro**) and Read-write (**rw**).

object ID

MIB saves data in variables and organizes them hierarchically. Formally, the MIB defines each variable as an object ID (OID). OIDs uniquely identify managed objects in the MIB hierarchy (figure 10.4). For example, OIDs belonging to Cisco, are numbered as follows: .iso (1).org (3).dod (6).internet (1).private (4).enterprises (1).cisco (9). Therefore the OID is 1.3.6.1.4.1.9. The data is retrieved via the `snmpget` utility, issued on the NMS. Using

Figure 10.4: OID tree



the snmpget utility, one can manually retrieve real-time data or have the NMS run a report which would give you a period of time that you could use the data to get the average.

10.3 SNMP configuration

10.3.1 SNMPv2

```

R1(config)# snmp-server community batonaug ro SNMP_ACL
R1(config)# snmp-server location NOC_SNMP_MANAGER
R1(config)# snmp-server contact Wayne World
R1(config)# snmp-server host 192.168.1.3 version 2c batonaug
R1(config)# snmp-server enable traps
R1(config)# ip access-list standard SNMP_ACL
R1(config-std-nacl)# permit 192.168.1.3
  
```

1. (Required) Configure the community string and access level (read-only or read-write) with the snmp-server community string ro — rw command.
2. (Optional) Document the location of the device using the snmp-server location text command.
3. (Optional) Document the system contact using the snmp-server contact text command.
4. (Optional) Restrict SNMP access to NMS hosts (SNMP managers) that are permitted by an ACL: define the ACL and then reference the ACL with the snmp-server community string access-list-number-or-name command. This command can be used both to specify a community string and to restrict SNMP access via ACLs. Step 1 and Step 4 can be combined into one step, if desired; the Cisco networking device combines the two commands into one if they are entered separately.
5. (Optional) Specify the recipient of the SNMP trap operations with the snmp-server hosthost-id [version 1—2c — 3 [auth — noauth — priv]] community-string command. By default, no trap manager is defined.

6. (Optional) Enable traps on an SNMP agent with the `snmp-server enable traps notification-types` command. If no trap notification types are specified in this command, then all trap types are sent. Repeated use of this command is required if a particular subset of trap types is desired.

Note! To verify SNMP configuration, use any of the variations of the `show snmp` command.

Note! Only the first step is required, the rest are optional.

Note! By default, SNMP does not have any traps set. Without this command, SNMP managers must poll for all relevant information.

10.3.2 SNMPv3

SNMPv3 provides three security features: Message integrity and authentication, Encryption, Access control. SNMPv3 can be secured with the four steps.

Step 1: Configure an ACL to permit access to the protected management network.

```
Router(config)# ip access-list standard acl-name
Router(config-std-nacl)# permit source_net
```

Step 2: Configure an SNMP view.

```
Router(config)# snmp-server view view-name oid-tree
```

Step 3: Configure an SNMP group.

```
Router(config)# snmp-server group group-name v3 priv read view-name access [acl-
number | acl-name]
```

Step 4: Configure a user as a member of the SNMP group.

```
Router(config)# snmp-server user username group-name v3 auth {md5 | sha} auth-
password priv {des | 3des | aes {128 | 192 | 256}} privpassword
```

```
1 R1(config)# ip access-list standard PERMIT-ADMIN
2 R1(config-std-nacl)# permit 192.168.1.0 0.0.0.255
3 R1(config-std-nacl)# exit
4 R1(config)# snmp-server view SNMP-RO iso included
5 R1(config)# snmp-server group ADMIN v3 priv read SNMP-RO access PERMIT-ADMIN
6 R1(config)# snmp-server user BOB ADMIN v3 auth sha cisco12345 priv aes 128 cisco54321
7 R1(config)# end
```

line 1,2 The above example configures a standard ACL named PERMIT-ADMIN.

line 4 An SNMP view is named SNMP-RO and is configured to include the entire ISO tree from the MIB.

line 5 An SNMP group is configured with the name ADMIN. SNMP is set to version 3 with authentication and encryption required. The group is allowed read-only access to the view SNMP-RO. Access for the group is limited by the PERMIT-ADMIN ACL.

line 6 An SNMP user, BOB, is configured as a member of the group ADMIN. Authentication is set to use SHA, and an authentication password is configured. Although R1 supports up to AES 256 encryption, the SNMP management software only supports AES 128. Therefore, the encryption is set to AES 128 and an encryption password is configured.

10.4 SPAN

10.4.1 Introduction

A packet analyzer is typically software that captures packets entering and exiting a network interface card (NIC). However, the basic operation of a modern switched network disables the packet analyzer ability to capture traffic from other sources. For instance, a user running Wireshark can only capture traffic going to their NIC.

The solution to this dilemma is to enable port mirroring. The port mirroring feature allows a switch to copy and send Ethernet frames from specific ports to the destination port connected to a packet analyzer. The Switched Port Analyzer (SPAN) feature on Cisco switches is a type of port mirroring that sends copies of the frame entering a port, out another port on the same switch.

SPAN is commonly implemented to deliver traffic to specialized devices including: Packet analyzers (such as Wireshark) and IPSs (Intrusion Prevention Systems)

10.4.2 Local SPAN

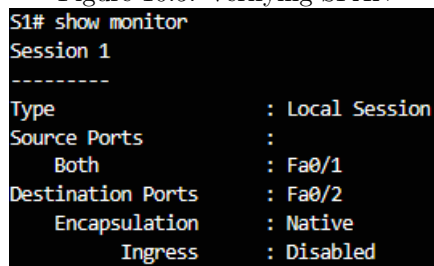
Local SPAN is when traffic on a switch is mirrored to another port on that switch. A SPAN session is the association between source ports (or VLANs) and a destination port. Traffic entering or leaving the source port (or VLAN) is replicated by the switch on the destination port. There are three important things to consider when configuring SPAN:

- The destination port cannot be a source port, and the source port cannot be a destination port.
- The number of destination ports is platform-dependent. Some platforms allow for more than one destination port.
- The destination port is no longer a normal switch port. Only monitored traffic passes through that port.

```
S1(config)# monitor session 1 source interface f0/1
S1(config)# monitor session 1 destination interface f0/2
S1(config)# end
S1# show monitor
```

The above command is used to associate a source port and a destination port with a SPAN session. A separate monitor session command is used for each session. A VLAN can be specified instead of a physical port. Use `show monitor` to verify SPAN configuration.

Figure 10.5: Verifying SPAN



```
S1# show monitor
Session 1
-----
Type                : Local Session
Source Ports        :
  Both              : Fa0/1
Destination Ports    : Fa0/2
Encapsulation        : Native
Ingress              : Disabled
```

In the output of show command shown in Figure 10.5, the session number is 1, the source port for both traffic directions (receive and transmit) is F0/1, and the destination port is F0/2. The ingress SPAN is disabled on the destination port, so only traffic that leaves the destination port is copied to that port.

10.4.3 RSPAN

Remote SPAN (RSPAN) allows source and destination ports to be in different switches. RSPAN uses two sessions. One session is used as the source and one session is used to copy or receive the traffic from a VLAN. The traffic for each RSPAN session is carried over trunk links in a user-specified RSPAN VLAN that is dedicated (for that RSPAN session) in all participating switches.

Chapter 11

Quality of Service

11.1 Network transmission

11.1.1 Network quality

When the volume of traffic is greater than what can be transported across the network, devices queue, or hold, the packets in memory until resources become available to transmit them. If the number of packets to be queued continues to increase, the memory within the device fills up and packets are dropped.

Network congestion causes delay. Two types of delays are fixed and variable. A fixed delay is a specific amount of time a specific process takes, such as how long it takes to place a bit on the transmission media. A variable delay take an unspecified amount of time and is affected by factors such as how much traffic is being processed. Jitter is the variation in the delay of received packets.

A device implements QoS only when it is experiencing some type of congestion. Without any QoS mechanisms in place, packets are processed in the order in which they are received. When congestion occurs, network devices such as routers and switches can drop packets.

The *playout delay buffer* must buffer these packets and then play them out in a steady stream. The *digital signal processor (DSP)* interpolates what it thinks the audio should be and no problem is audible to the user.

11.1.2 Traffic characteristics

Voice

Voice traffic is predictable and smooth, as shown in the figure. However, voice is very sensitive to delays and dropped packets; there is no reason to re-transmit voice if packets are lost. Therefore, voice packets must receive a higher priority than other types of traffic.

Latency should be no more than 150 milliseconds (ms). Jitter should be no more than 30 ms, and voice packet loss should be no more than 1%. Voice traffic requires at least 30 Kbps of bandwidth.

Video

Video traffic tends to be unpredictable, inconsistent, and bursty compared to voice traffic. Compared to voice, video is less resilient to loss and has a higher volume of data per packet.

Latency should be no more than 400 milliseconds (ms). Jitter should be no more than 50 ms, and video packet loss should be no more than 1%. Video traffic requires at least 384 Kbps of bandwidth.

Data

Data traffic is relatively insensitive to drops and delays compared to voice and video. The two main factors a network administrator needs to ask about the flow of data traffic are the following: Does the data come from an

interactive application? Is the data mission critical?

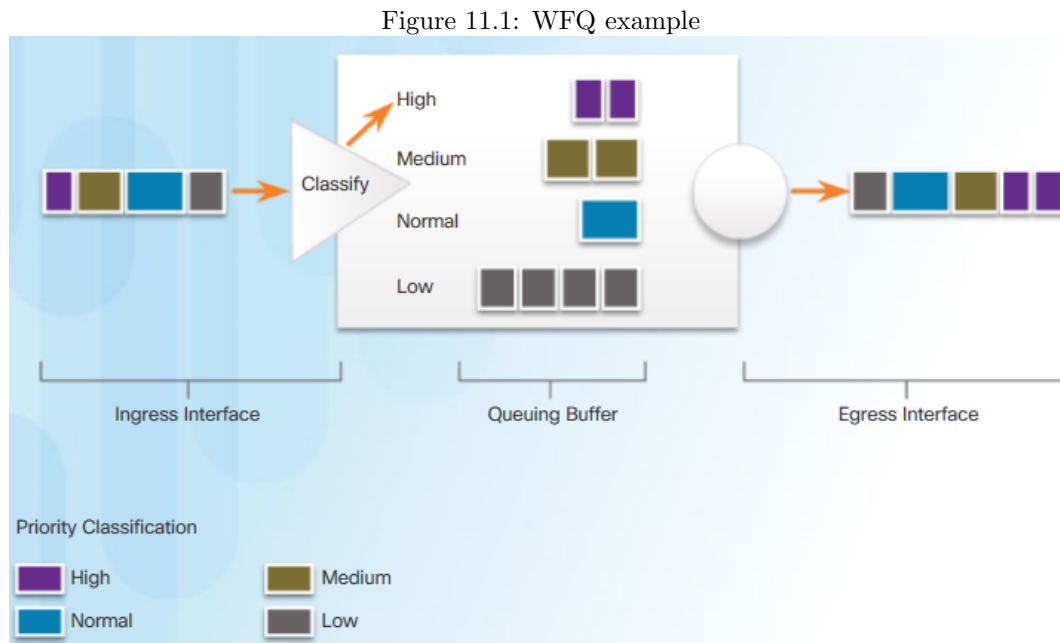
11.2 Queueing algorithms

Queueing is a congestion management tool that can buffer, prioritize, and, if required, reorder packets before being transmitted to the destination.

11.2.1 WFQ

WFQ (Weighted Fair Queuing) is an automated scheduling method that provides fair bandwidth allocation to all network traffic.

WFQ applies priority, or weights, to identified traffic and classifies it into conversations or flows, as shown in the figure 11.1. WFQ then determines how much bandwidth each flow is allowed relative to other flows. WFQ classifies traffic into different flows based on packet header addressing (IP addresses, MAC addresses, port numbers, protocol, and ToS value).



WFQ has limitations. It is not supported with tunneling and encryption. It does not offer the degree of precision control over bandwidth allocation that CBWFQ offers.

11.2.2 CBWFQ

Class-Based Weighted Fair Queuing (CBWFQ) extends the standard WFQ functionality to provide support for user-defined traffic classes. For CBWFQ, you define traffic classes based on match criteria including protocols, access control lists (ACLs), and input interfaces. A FIFO queue is reserved for each class, and traffic belonging to a class is directed to the queue for that class, as shown in the figure.

To characterize a class, you assign it bandwidth, weight, and queue limit. After a queue has reached its configured queue limit, adding more packets to the class causes tail drop. Tail drop means a router simply discards any packet that arrives at the tail end of a queue that has completely used up its packet-holding resources.

11.2.3 LLQ

The Low Latency Queuing (LLQ) feature brings strict priority queuing (PQ) to CBWFQ. Strict PQ allows delay-sensitive data such as voice to be sent before packets in other queues.

Without LLQ, CBWFQ services fairly based on weight; no class of packets may be granted strict priority. This scheme poses problems for voice traffic that is largely intolerant of delay, especially variation in delay. With LLQ, delay-sensitive data such as voice is sent first (before packets in other queues).

11.3 QoS models

The three models for implementing QoS are: Best-effort model, Integrated services (IntServ), Differentiated services (DiffServ). QoS is really implemented in a network using either IntServ or DiffServ.

11.3.1 Best effort

The best-effort model (meaning no QoS) treats all network packets in the same way, so an emergency voice message is treated the same way a digital photograph attached to an email is treated. The table 11.1 in the figure lists the benefits and drawbacks of the best effort model.

Table 11.1: Pros and Cons of Best-effort	
Benefits	Drawbacks
Most scalable	No guarantees of delivery
Scalability is limited by bandwidth	Packets can arrive in any order
No special QoS mechanism required	No packets have preferential treatment
Easy to deploy	Critical data is treated the same as casual one

11.3.2 Integrated services

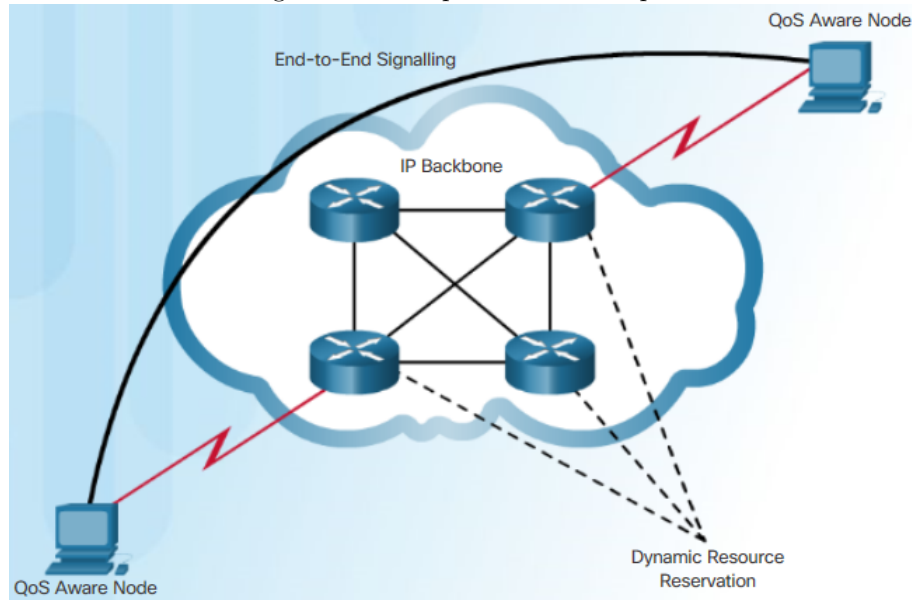
Integrated Services (IntServ) is a multiple-service model that can accommodate multiple QoS requirements.

It uses resource reservation and admission-control mechanisms as building blocks to establish and maintain QoS. Each individual communication must explicitly specify its traffic descriptor and requested resources to the network (Figure ??). The edge router performs admission control to ensure that available resources are sufficient in the network.

IntServ uses the Resource Reservation Protocol (RSVP) to signal the QoS needs of an application’s traffic along devices in the end-to-end path through the network. The table 11.2 lists the benefits and drawbacks of the IntServ model.

Table 11.2: Pros and Cons of IntServ	
Benefits	Drawbacks
Explicit end-to-end resource admission control	Resource intensive
Per-request policy admission control	Not scalable
Signaling of dynamic port numbers	

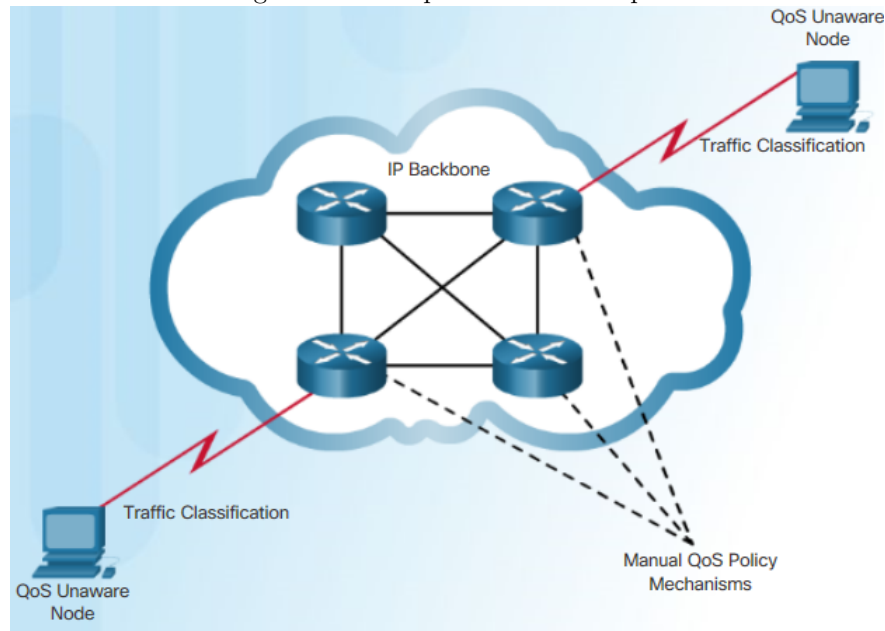
Figure 11.2: Simple IntServ example



11.3.3 Differentiated services

The DiffServ design overcomes the limitations of both the best-effort and IntServ models. Unlike IntServ and hard QoS in which the end-hosts signal their QoS needs to the network, DiffServ does not use signaling. It works on the provisioned-QoS model, where network elements are set up to service multiple classes of traffic each with varying QoS requirements (Figure 11.3).

Figure 11.3: Simple DiffServ example



As a host forwards traffic to a router, the router classifies the flows into aggregates (classes) and provides the appropriate QoS policy for the classes. DiffServ enforces and applies QoS mechanisms on a hop-by-hop basis, uniformly applying global meaning to each traffic class to provide both flexibility and scalability.

Table 11.3 lists the benefits and drawbacks of the DiffServ model.

Table 11.3: Pros and Cons of DiffServ

Benefits	Drawbacks
Highly scalable	No absolute guarantee of delivery
Many different levels of quality	Requires complex mechanisms

11.4 QoS implementation

11.4.1 Classification and Marking

Before a packet can have a QoS policy applied to it, the packet has to be classified. Classification and marking allows us to identify or mark types of packets. Marking means that we are adding a value to the packet header. Devices receiving the packet look at this field to see if it matches a defined policy.

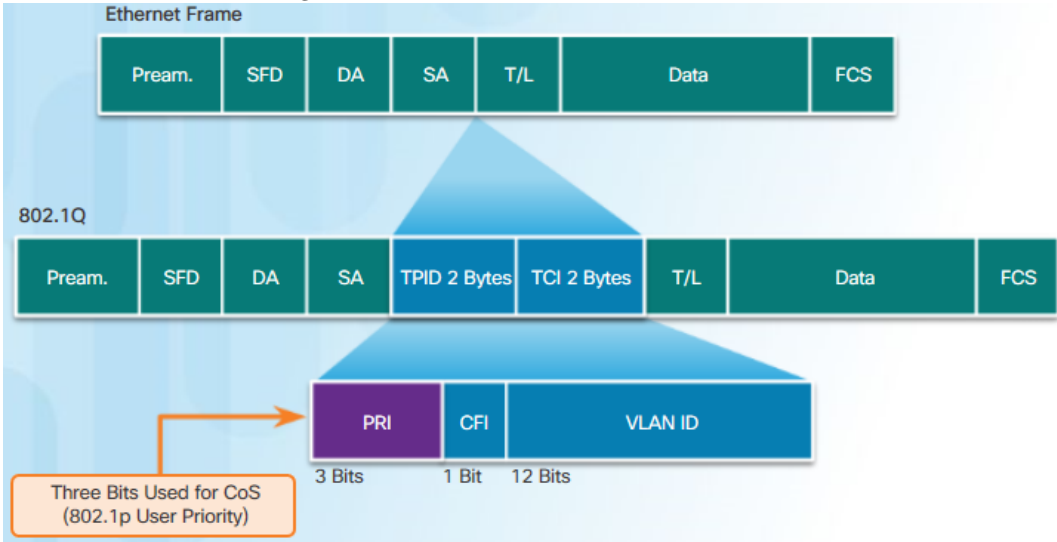
Traffic should be classified and marked as close to its source as technically and administratively feasible. This defines the trust boundary.

Trusted endpoints have the capabilities and intelligence to mark application traffic to the appropriate Layer 2 CoS and/or Layer 3 DSCP values. Examples of trusted endpoints include IP phones, wireless access points, videoconferencing gateways and systems, IP conferencing stations, and more.

Marking at Layer 2

802.1Q is the IEEE standard that supports VLAN tagging at layer 2 on Ethernet networks. The 802.1Q standard also includes the QoS prioritization scheme known as IEEE 802.1p. The 802.1p standard uses the first three bits in the Tag Control Information (TCI) field (Figure 11.4). Known as the Priority (PRI) field, this 3-bit field identifies the Class of Service (CoS) markings.

Figure 11.4: Ethernet Class of Service values

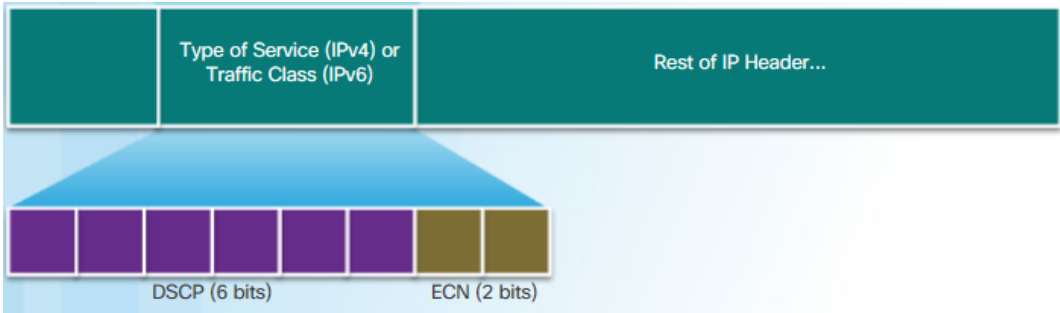


Marking at Layer 3

Both IPv4 and IPv6 support an 8-bit field for marking, the Type of Service (ToS) field for IPv4 and the Traffic Class field for IPv6. Figure 11.5 displays the contents of the 8-bit field. The field has 6-bits allocated for QoS, called the **DiffServ** Code Point (DSCP) field. The remaining two IP Extended Congestion Notification (ECN) bits can be used by ECN-aware routers to mark packets instead of dropping them. The ECN marking informs downstream

routers that there is congestion in the packet flow.

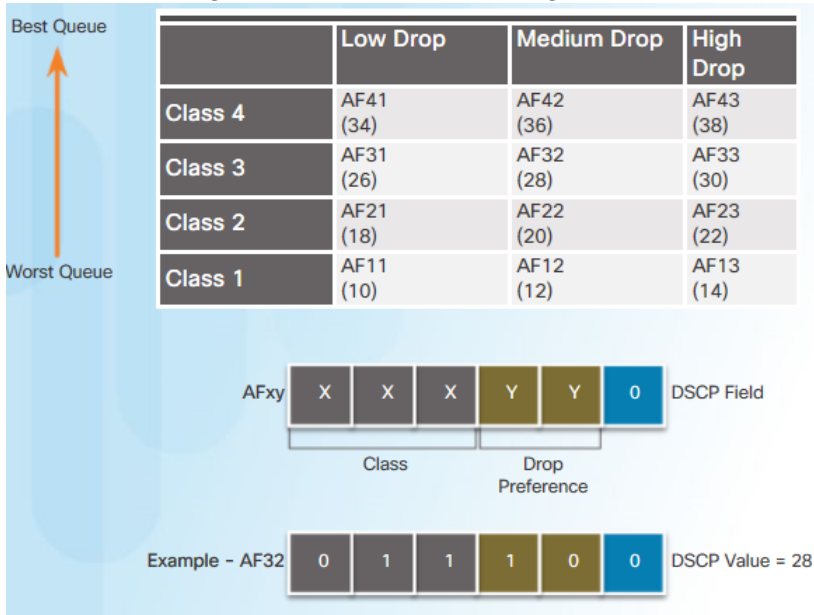
Figure 11.5: Type of Service/Traffic Class Field



The DSCP values are organized into three categories:

- Best-Effort (BE)** When a router experiences congestion, these packets will be dropped. No QoS plan is implemented.
- Expedited Forwarding (EF)** DSCP decimal value is 46 (binary 101110). At Layer 3, Cisco recommends that EF only be used to mark voice packets.
- Assured Forwarding (AF)** use the 5 most significant DSCP bits to indicate queues and drop preference. As shown in Figure 11.6, the first 3 most significant bits are used to designate the class. The 4th and 5th most significant bits are used to designate the drop preference. The 6th most significant bit is set to zero. The AF_{xy} formula shows how the AF values are calculated. For example, AF32 belongs to class 3 (binary 011) and has a medium drop preference (binary 10).

Figure 11.6: Assured forwarding values



11.4.2 Congestion Avoidance

Congestion avoidance tools are simpler than Congestion management. These tools can monitor the average depth of the queue, as represented in the figure. When the queue is below the minimum threshold, there are no drops. As the queue fills up to the maximum threshold, a small percentage of packets are dropped. When the maximum

threshold is passed, all packets are dropped.

Cisco IOS QoS includes weighted random early detection (WRED) as a possible congestion avoidance solution. Using WRED helps avoid tail drops and maximizes network use and TCP-based application performance. In case of UDP-based traffic, methods such as queuing and compression techniques help to reduce and even prevent UDP packet loss.

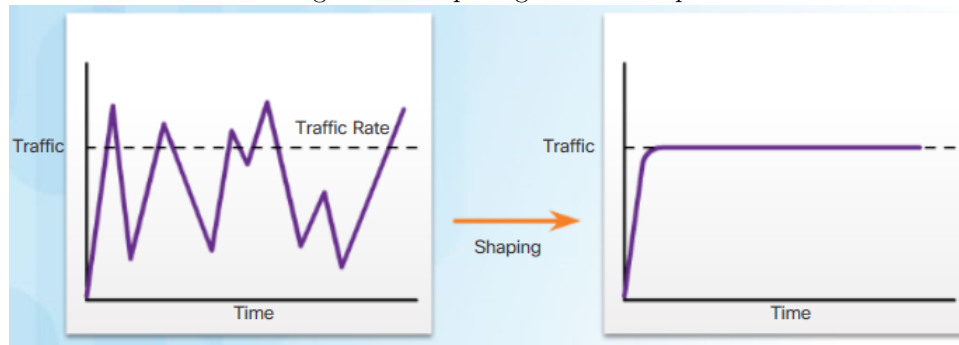
11.4.3 Shaping and Policing

Traffic shaping and traffic policing are two mechanisms provided by Cisco IOS QoS software to prevent congestion.

Traffic shaping

Traffic shaping retains excess packets in a queue and then schedules the excess for later transmission over increments of time. The result of traffic shaping is a smoothed packet output rate, as shown in Figure 11.7.

Figure 11.7: Spacing traffic example



Ensure that you have sufficient memory when enabling shaping. In addition, shaping requires a scheduling function (CBFWQ, LLQ) for later transmission of any delayed packets.

Traffic policing

Shaping is an outbound concept; packets going out an interface get queued and can be shaped. In contrast, policing is applied to inbound traffic on an interface. When the traffic rate reaches the configured maximum rate, excess traffic is dropped (or remarked). See also Figure 11.8.

Figure 11.8: Spacing traffic example



Ensure that you have sufficient memory when enabling shaping. In addition, shaping requires a scheduling function (CBFWQ, LLQ) for later transmission of any delayed packets.