CHAPTER 2

LAN Redundancy

Objectives

Upon completion of this chapter, you will be able to answer the following questions:

- What are the issues that you should be concerned with when implementing a redundant network?
- How does IEEE 802.1D STP operate?
- What are the different varieties of spanning tree?
- How does PVST+ operate in a switched LAN environment?
- How does Rapid PVST+ operate in a switched LAN environment?
- What are the commands to configure PVST+ in a switched LAN environment?
- What are the commands to configure Rapid PVST+ in a switched LAN environment?
- What are the common STP configuration issues?
- What are the purpose and operation of First Hop Redundancy Protocols?
- What are the different varieties of First Hop Redundancy Protocols?
- What are the commands to verify HSRP and GLBP implementations?

Key Terms

This chapter uses the following key terms. You can find the definitions in the Glossary.

First Hop Redundancy Protocols (FHRP) *page 51*

broadcast storm *page 54*

time to live (TTL) *page 54*

root bridge *page 59*

bridge protocol data unit (BPDU) *page 59*

blocking state *page 60*

Rapid Spanning Tree Protocol (RSTP) *page 61*

Multiple Spanning Tree Protocol (MSTP) *page 61*

IEEE-802.1D-2004 *page 61*

bridge ID (BID) *page 61*

extended system ID *page 62*

root port *page 62*

designated port *page 63*

alternate and backup port *page 63*

disabled port *page 63*

default port cost *page 64*

bridge priority *page 74*

Common Spanning Tree (CST) *page 78*

Introduction (2.0.1.1)

Network redundancy is a key to maintaining network reliability. Multiple physical links between devices provide redundant paths. The network can then continue to operate when a single link or port has failed. Redundant links can also share the traffic load and increase capacity.

Multiple paths need to be managed so that Layer 2 loops are not created. The best paths are chosen, and an alternate path is immediately available should a primary path fail. The Spanning Tree Protocols are used to manage Layer 2 redundancy.

Redundant devices, such as multilayer switches or routers, provide the capability for a client to use an alternate default gateway should the primary default gateway fail. A client can now have multiple paths to more than one possible default gateway. *First Hop Redundancy Protocols* are used to manage how a client is assigned a default gateway, and to be able to use an alternate default gateway should the primary default gateway fail.

This chapter focuses on the protocols used to manage these forms of redundancy. It also covers some of the potential redundancy problems and their symptoms.

Class Activity 2.0.1.2: Stormy Traffic

It is your first day on the job as a network administrator for a small- to medium-sized business. The previous network administrator left suddenly after a network upgrade took place for the business.

During the upgrade, a new switch was added. Since the upgrade, many employees complain that they are having trouble accessing the Internet and servers on your network. In fact, most of them cannot access the network at all. Your corporate manager asks you to immediately research what could be causing these connectivity problems and delays.

So you take a look at the equipment operating on your network at your main distribution facility in the building. You notice that the network topology seems to be visually correct and that cables have been connected correctly, routers and switches are powered on and operational, and switches are connected together to provide backup or redundancy.

However, one thing you do notice is that all of your switches' status lights are constantly blinking at a very fast pace to the point that they almost appear solid. You think you have found the problem with the connectivity issues your employees are experiencing.

Use the Internet to research STP. As you research, take notes and describe

- Broadcast storm
- Switching loops
- The purpose of STP
- Variations of STP

Complete the reflection questions that accompany the PDF file for this activity. Save your work and be prepared to share your answers with the class.

Spanning Tree Concepts (2.1)

This section focuses on the purpose and operation of the Spanning Tree Protocol.

Purpose of Spanning Tree (2.1.1)

STP provides the mechanism to have redundant links at Layer 2 while avoiding the potential for loops and MAC address database instability.

Redundancy at OSI Layers 1 and 2 (2.1.1.1)

The three-tier hierarchical network design that uses core, distribution, and access layers with redundancy attempts to eliminate a single point of failure on the network. Multiple cabled paths between switches provide physical redundancy in a switched network. This improves the reliability and availability of the network. Having alternate physical paths for data to traverse the network makes it possible for users to access network resources, despite path disruption.

The following steps explain how redundancy works in the topology shown in Figure 2-1.

1. PC1 is communicating with PC4 over a redundant network topology.
2. When the network link between S1 and S2 is disrupted, the path between PC1 and PC4 is automatically adjusted to compensate for the disruption (shown in Figure 2-1).
3. When the network connection between S1 and S2 is restored, the path is then readjusted to route traffic directly from S2 to S1 to get to PC4.

Note

To view an animation of these steps, refer to the online course.

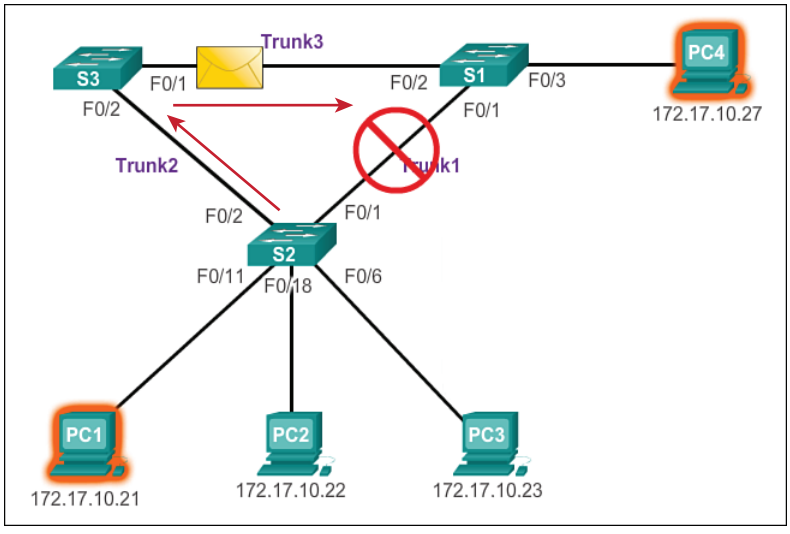


Figure 2-1 Redundancy in a Hierarchical Network

For many organizations, the availability of the network is essential to supporting business needs; therefore, the network infrastructure design is a critical business element. Path redundancy is a solution for providing the necessary availability of multiple network services by eliminating the possibility of a single point of failure.

Note

The OSI Layer 1 redundancy is illustrated using multiple links and devices, but more than just physical planning is required to complete the network setup. For the redundancy to work in a systematic way, the use of OSI Layer 2 protocols such as STP is also required.

Redundancy is an important part of hierarchical design for preventing disruption of network services to users. Redundant networks require adding physical paths, but logical redundancy must also be part of the design. However, redundant paths in a switched Ethernet network can cause both physical and logical Layer 2 loops.

Logical Layer 2 loops can occur because of the natural operation of switches, specifically, the learning and forwarding process. When multiple paths exist between two devices on a network, and there is no spanning tree implementation on the switches, a Layer 2 loop occurs. A Layer 2 loop can result in three primary issues:

- **MAC database instability:** Instability in the content of the MAC address table results from copies of the same frame being received on different ports of the switch. Data forwarding can be impaired when the switch consumes the resources that are coping with instability in the MAC address table.

- **Broadcast storms:** Without some loop-avoidance process, each switch can flood broadcasts endlessly. This situation is commonly called a *broadcast storm*.
- **Multiple frame transmission:** Multiple copies of unicast frames can be delivered to destination stations. Many protocols expect to receive only a single copy of each transmission. Multiple copies of the same frame can cause unrecoverable errors.

Issues with Layer 1 Redundancy: MAC Database Instability (2.1.1.2)

Ethernet frames do not have a *time to live* (TTL) attribute, like IP packets. As a result, if there is no mechanism enabled to block continued propagation of these frames on a switched network, they continue to propagate between switches endlessly, or until a link is disrupted and breaks the loop. This continued propagation between switches can result in MAC database instability. This can occur because of broadcast frames forwarding.

Broadcast frames are forwarded out all switch ports, except the original ingress port. This ensures that all devices in a broadcast domain are able to receive the frame. If there is more than one path for the frame to be forwarded out, an endless loop can result. When a loop occurs, it is possible for the MAC address table on a switch to constantly change with the updates from the broadcast frames, resulting in MAC database instability.

The following steps demonstrate the MAC database instability issue. Figure 2-2 shows a snapshot during Step 4.

1. PC1 sends out a broadcast frame to S2. S2 receives the broadcast frame on F0/11. When S2 receives the broadcast frame, it updates its MAC address table to record that PC1 is available on port F0/11.
2. Because it is a broadcast frame, S2 forwards the frame out all ports, including Trunk1 and Trunk2. When the broadcast frame arrives at S3 and S1, they update their MAC address tables to indicate that PC1 is available out port F0/1 on S1 and out port F0/2 on S3.
3. Because it is a broadcast frame, S3 and S1 forward the frame out all ports, except the ingress port. S3 sends the broadcast frame from PC1 to S1. S1 sends the broadcast frame from PC1 to S3. Each switch updates its MAC address table with the incorrect port for PC1.
4. Each switch again forwards the broadcast frame out all of its ports, except the ingress port, resulting in both switches forwarding the frame to S2 (shown in Figure 2-2).

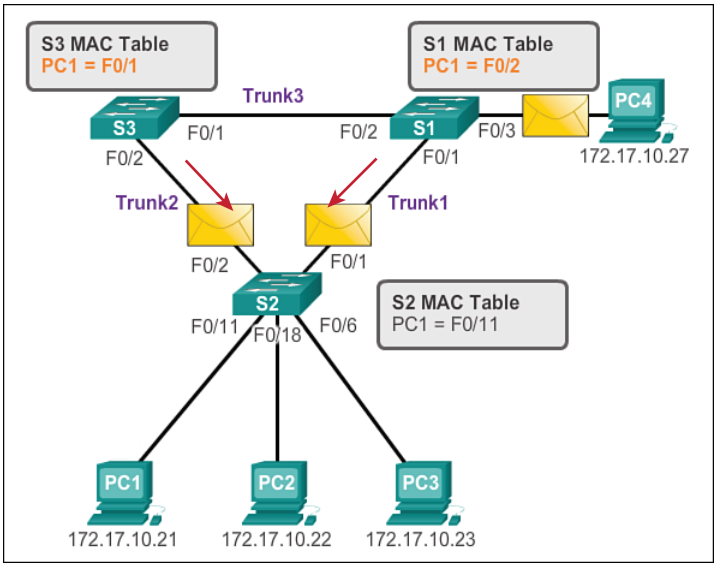


Figure 2-2 MAC Database Instability Example

5. When S2 receives the broadcast frames from S3 and S1, the MAC address table is updated again, this time with the last entry received from the other two switches.

Note

To view an animation of these steps, refer to the online course.

This process repeats over and over again until the loop is broken by physically disconnecting the connections causing the loop or powering down one of the switches in the loop. This creates a high CPU load on all switches caught in the loop. Because the same frames are constantly being forwarded back and forth between all switches in the loop, the CPU of the switch must process a lot of data. This slows down performance on the switch when legitimate traffic arrives.

A host caught in a network loop is not accessible to other hosts on the network. Additionally, because of the constant changes in the MAC address table, the switch does not know out of which port to forward unicast frames. In the previous example, the switches will have the incorrect ports listed for PC1. Any unicast frame destined for PC1 loops around the network, just as the broadcast frames do. More and more frames looping around the network eventually create a broadcast storm.

Issues with Layer 1 Redundancy: Broadcast Storms (2.1.1.3)

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.

A broadcast storm is inevitable on a looped network. As more devices send broadcasts over the network, more traffic is caught within the loop, consuming resources. This eventually creates a broadcast storm that causes the network to fail.

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 broadcast traffic that is being flooded endlessly around the looped network. This can cause the end device to malfunction because of the high processing requirements for sustaining such a high traffic load on the NIC.

The following steps demonstrate the broadcast storm issue. Figure 2-3 shows the final result during Step 6.

1. PC1 sends a broadcast frame out onto the looped network.
2. The broadcast frame loops between all the interconnected switches on the network.
3. PC4 also sends a broadcast frame out on to the looped network.
4. The PC4 broadcast frame also gets caught in the loop between all the interconnected switches, just like the PC1 broadcast frame.
5. As more devices send broadcasts over the network, more traffic is caught within the loop, consuming resources. This eventually creates a broadcast storm that causes the network to fail.
6. When the network is fully saturated with broadcast traffic that is looping between the switches, new traffic is discarded by the switch because it is unable to process it. In Figure 2-3, S2 is now discarding additional frames.

Note

To view an animation of these steps, refer to the online course.

Because devices connected to a network are regularly sending out broadcast frames, such as ARP requests, a broadcast storm can develop in seconds. As a result, when a loop is created, the switched network is quickly brought down.

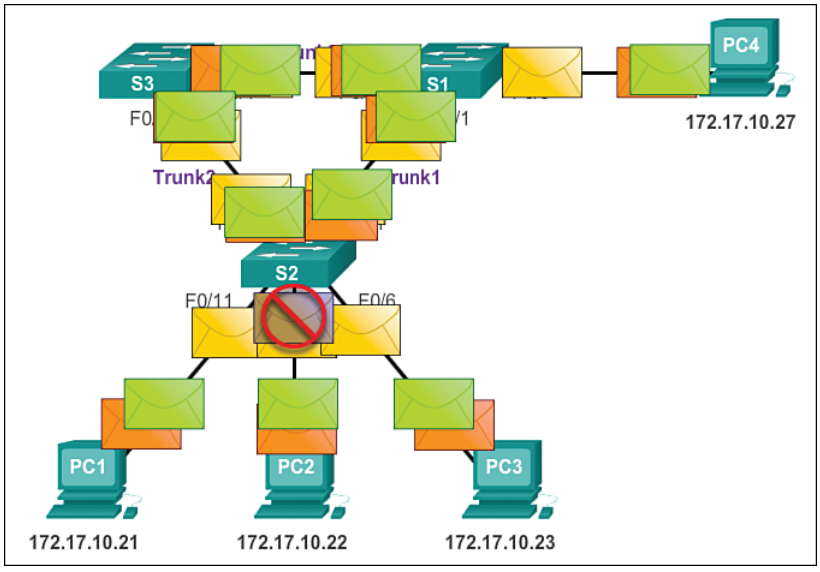


Figure 2-3 Broadcast Storms

Issues with Layer 1 Redundancy: Duplicate Unicast Frames (2.1.1.4)

Broadcast frames are not the only type of frames that are affected by loops. Unicast frames sent onto a looped network can result in duplicate frames arriving at the destination device.

The following steps demonstrate the duplicate unicast frames issue. Figure 2-4 shows a snapshot during Step 5 and Step 6.

1. PC1 sends a unicast frame destined for PC4.
2. S2 does not have an entry for PC4 in its MAC table, so it floods the unicast frame out all switch ports in an attempt to find PC4.
3. The frame arrives at switches S1 and S3.
4. S1 does have a MAC address entry for PC4, so it forwards the frame out to PC4.
5. S3 also has an entry in its MAC address table for PC4, so it forwards the unicast frame out Trunk3 to S1.
6. S1 receives the duplicate frame and forwards the frame out to PC4.
7. PC4 has now received the same frame twice.

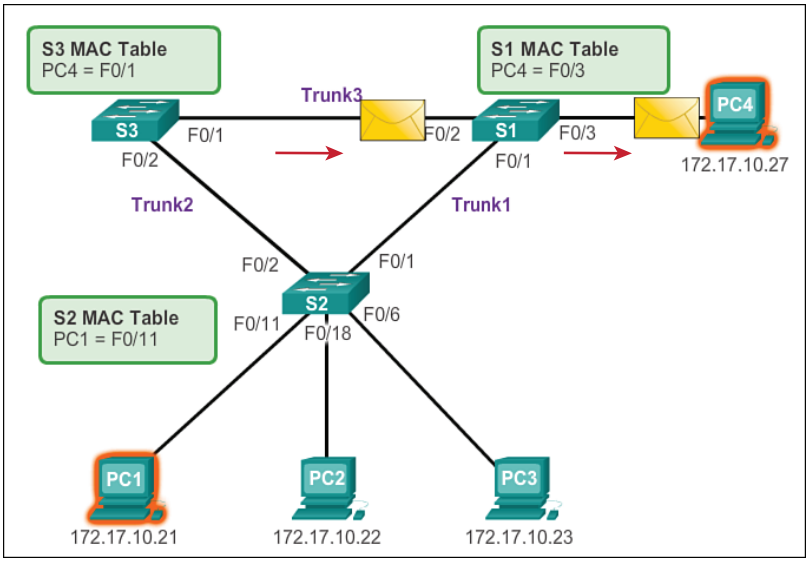


Figure 2-4 S1 and S3 Send Duplicate Frame to PC4

Note

To view an animation of these steps, refer to the online course.

Most upper-layer protocols are not designed to recognize, or cope with, duplicate transmissions. In general, protocols that make use of a sequence-numbering mechanism assume that the transmission has failed and that the sequence number has recycled for another communication session. Other protocols attempt to hand the duplicate transmission to the appropriate upper-layer protocol to be processed and possibly discarded.

Layer 2 LAN protocols, such as Ethernet, lack a mechanism to recognize and eliminate endlessly looping frames. Some Layer 3 protocols implement a TTL mechanism that limits the number of times a Layer 3 networking device can retransmit a packet. Lacking such a mechanism, Layer 2 devices continue to retransmit looping traffic indefinitely. A Layer 2 loop-avoidance mechanism, STP, was developed to address these problems.

To prevent these issues from occurring in a redundant network, some type of spanning tree must be enabled on the switches. Spanning tree is enabled, by default, on Cisco switches to prevent Layer 2 loops from occurring.

Packet Tracer Activity

Packet Tracer Activity 2.1.1.5: Examining a Redundant Design

In this activity, you will observe how STP operates, by default, and how it reacts when faults occur. Switches have been added to the network "out of the box." Cisco

switches can be connected to a network without any additional action required by the network administrator. For the purpose of this activity, the bridge priority was modified.

STP Operation (2.1.2)

STP uses the concepts of a *root bridge*, port roles, and path costs to calculate which links to use in a redundant topology.

Spanning Tree Algorithm: Introduction (2.1.2.1)

Redundancy increases the availability of the network topology by protecting the network from a single point of failure, such as a failed network cable or switch. When physical redundancy is introduced into a design, loops and duplicate frames occur. Loops and duplicate frames have severe consequences for a switched network. The Spanning Tree Protocol (STP) was developed to address these issues.

STP ensures that there is only one logical path between all destinations on the network by intentionally blocking redundant paths that could cause a loop. A port is considered blocked when user data is prevented from entering or leaving that port. This does not include *bridge protocol data unit* (BPDU) frames that are used by STP to prevent loops. Blocking the redundant paths is critical to preventing loops on the network. The physical paths still exist to provide redundancy, but these paths are disabled to prevent the loops from occurring. If the path is ever needed to compensate for a network cable or switch failure, STP recalculates the paths and unblocks the necessary ports to allow the redundant path to become active.

In Figure 2-5, all switches have STP enabled:

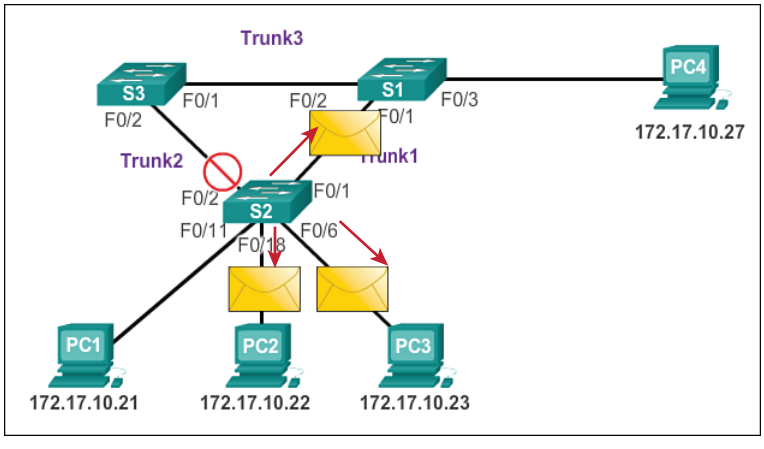


Figure 2-5 Normal STP Operation

1. PC1 sends a broadcast out onto the network.
2. S2 is configured with STP and has set the port for Trunk2 to a *blocking state*, as shown in Figure 2-5. The blocking state prevents ports from being used to forward user data, thus preventing a loop from occurring. S2 forwards a broadcast frame out all switch ports, except the originating port from PC1 and the port for Trunk2.
3. S1 receives the broadcast frame and forwards it out all of its switch ports, where it reaches PC4 and S3. S3 forwards the frame out the port for Trunk2 and S2 drops the frame. The Layer 2 loop is prevented.

Note

To view an animation of these steps, refer to the online course.

In Figure 2-6, STP recalculates the path when a failure occurs.

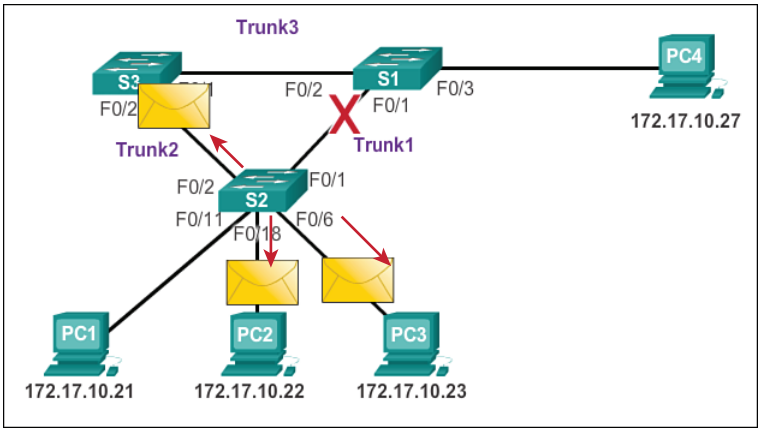


Figure 2-6 STP Compensates for Network Failure

1. PC1 sends a broadcast out onto the network.
2. The broadcast is then forwarded around the network, just as in the previous animation.
3. The trunk link between S2 and S1 fails, resulting in the previous path being disrupted.
4. S2 unblocks the previously blocked port for Trunk2 and allows the broadcast traffic to traverse the alternate path around the network, permitting communication to continue. If this link comes back up, STP reconverges and the port on S2 is again blocked.

Note

To view an animation of these steps, refer to the online course.

STP prevents loops from occurring by configuring a loop-free path through the network using strategically placed "blocking-state" ports. The switches running STP are able to compensate for failures by dynamically unblocking the previously blocked ports and permitting traffic to traverse the alternate paths.

Up to now, we have used the term Spanning Tree Protocol and the acronym STP. The usage of the Spanning Tree Protocol term and the STP acronym can be misleading. Many professionals generically use these to refer to various implementations of spanning tree, such as ***Rapid Spanning Tree Protocol (RSTP)*** and ***Multiple Spanning Tree Protocol (MSTP)***. In order to communicate spanning tree concepts correctly, it is important to refer to the particular implementation or standard in context. The latest IEEE documentation on spanning tree, IEEE-802-1D-2004, says "STP has now been superseded by the Rapid Spanning Tree Protocol (RSTP)." So one sees that the IEEE uses "STP" to refer to the original implementation of spanning tree and "RSTP" to describe the version of spanning tree specified in ***IEEE-802.1D-2004***. In this book, when the original Spanning Tree Protocol is the context of a discussion, the phrase "original 802.1D spanning tree" is used to avoid confusion.

Note

STP is based on an algorithm invented by Radia Perlman while working for Digital Equipment Corporation, and published in the 1985 paper "An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN."

Spanning Tree Algorithm: Port Roles (2.1.2.2)

IEEE 802.1D STP uses 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. In Figure 2-7, the root bridge (switch S1) is chosen through an election process. All switches participating in STP exchange BPDU frames to determine which switch has the lowest ***bridge ID*** (BID) on the network. The switch with the lowest BID automatically becomes the root bridge for the STA calculations.

Note

For simplicity, assume until otherwise indicated that all ports on all switches are assigned to VLAN 1. Each switch has a unique MAC address associated with VLAN 1.

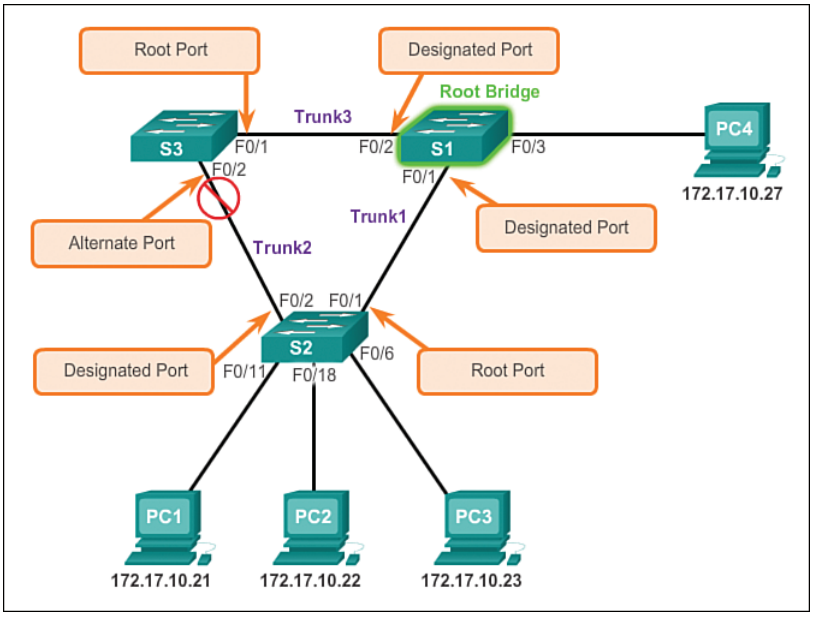


Figure 2-7 STP Algorithm

A BPDU is a messaging frame exchanged by switches for STP. Each BPDU contains a BID that identifies the switch that sent the BPDU. The BID contains a priority value, the MAC address of the sending switch, and an optional *extended system ID*. The lowest BID value is determined by the combination of these three fields.

After the root bridge has been determined, the STA calculates the shortest path to it. Each switch uses the STA to determine which ports to block. While the STA determines the best paths to the root bridge for all switch ports in the broadcast domain, traffic is prevented from being forwarded through the network. The STA considers both path and port costs when determining which ports to block. The path costs are calculated using port cost values associated with port speeds for each switch port along a given path. The sum of the port cost values determines the overall path cost to the root bridge. If there is more than one path to choose from, STA chooses the path with the lowest path cost.

When the STA has determined which paths are most desirable relative to each switch, it assigns port roles to the participating switch ports. The port roles describe their relation in the network to the root bridge and whether they are allowed to forward traffic:

- *Root ports*: Switch ports closest to the root bridge. In Figure 2-7, the root port on S2 is F0/1 configured for the trunk link between S2 and S1. The root port on S3 is F0/1, configured for the trunk link between S3 and S1. Root ports are selected on a per-switch basis.

- *Designated ports*: All nonroot ports that are still permitted to forward traffic on the network. In Figure 2-7, switch ports (F0/1 and F0/2) on S1 are designated ports. S2 also has its port F0/2 configured as a designated port. Designated ports are selected on a per-trunk basis. If one end of a trunk is a root port, the other end is a designated port. All ports on the root bridge are designated ports.
- *Alternate and backup ports*: Alternate ports and backup ports are configured to be in a blocking state to prevent loops. In the figure, the STA configured port F0/2 on S3 in the alternate role. Port F0/2 on S3 is in the blocking state. Alternate ports are selected only on trunk links where neither end is a root port. Notice in Figure 2-7 that only one end of the trunk is blocked. This allows for faster transition to a forwarding state, when necessary. (Blocking ports only come into play when two ports on the same switch are connected to each other through a hub or single cable.)
- *Disabled ports*: A disabled port is a switch port that is shut down.

Spanning Tree Algorithm: Root Bridge (2.1.2.3)

As shown in Figure 2-8, every spanning tree instance (switched LAN or broadcast domain) has a switch designated as the root bridge. The root bridge serves as a reference point for all spanning tree calculations to determine which redundant paths to block.

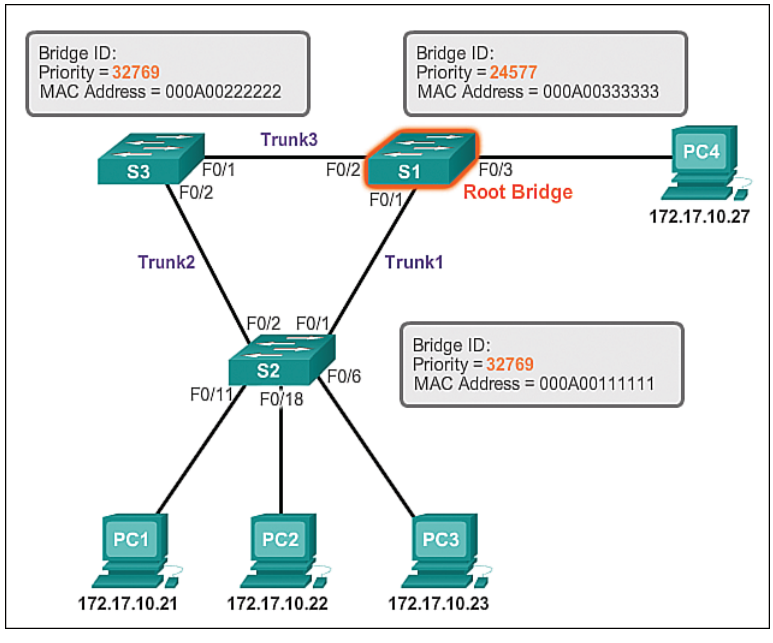


Figure 2-8 Root Bridge

An election process determines which switch becomes the root bridge.

Figure 2-9 shows the BID fields. The BID is made up of a priority value, an extended system ID, and the MAC address of the switch.

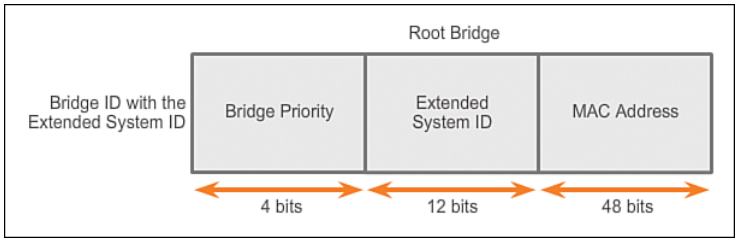


Figure 2-9 BID Fields

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.

As the switches forward their BPDU frames, adjacent switches in the broadcast domain read the root ID information from the BPDU frames. If the root ID from a BPDU received is lower than the root ID on the receiving switch, the receiving switch updates its root ID, identifying the adjacent switch as the root bridge. Actually, it might not be an adjacent switch, but could be any other switch in the broadcast domain. The switch then forwards new BPDU frames with the lower root ID to the other adjacent switches. Eventually, the switch with the lowest BID ends up being identified as the root bridge for the spanning tree instance.

There is a root bridge elected for each spanning tree instance. It is possible to have multiple distinct root bridges. If all ports on all switches are members of VLAN 1, there is only one spanning tree instance. The extended system ID plays a role in how spanning tree instances are determined.

Spanning Tree Algorithm: Path Cost (2.1.2.4)

When the root bridge has been elected for the spanning tree instance, the STA starts the process of determining the best paths to the root bridge from all destinations in the broadcast domain. The path information is determined by summing up the individual port costs along the path from the destination to the root bridge. Each "destination" is actually a switch port.

The *default port costs* are defined by the speed at which the port operates. As shown in Table 2-1, 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.

Table 2-1 Best Paths to the Root Bridge

Link Speed	Cost (Revised IEEE Specification)	Cost (Previous IEEE Specification)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100

Note

As newer, faster Ethernet technologies enter the marketplace, the path cost values can change to accommodate the different speeds available. The nonlinear numbers in Table 2-1 accommodate some improvements to the older Ethernet standard. The values have already been changed to accommodate the 10-Gb/s Ethernet standard. To illustrate the continued change associated with high-speed networking, Catalyst 4500 and 6500 switches support a longer path cost method. For example, 10 Gb/s has a 2000 path cost, 100 Gb/s has a 200 path cost, and 1 Tb/s has a 20 path cost.

Although switch ports have a default port cost associated with them, the port cost is configurable. The ability to configure individual port costs gives the administrator the flexibility to manually control the spanning tree paths to the root bridge.

To configure the port cost of an interface, enter the **spanning-tree cost** *value* command in interface configuration mode. The value can be between 1 and 200,000,000.

In Example 2-1, switch port F0/1 has been configured with a port cost of 25 using the **spanning-tree cost 25** interface configuration mode command on the F0/1 interface.

Example 2-1 Configure Port Cost

```
S2# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
S2(config)# interface f0/1
S2(config-if)# spanning-tree cost 25
S2(config-if)# end
S2#
```

To restore the port cost to the default value of 19, enter the **no spanning-tree cost** interface configuration mode command.

The path cost is equal to the sum of all the port costs along the path to the root bridge, as shown in Figure 2-10.

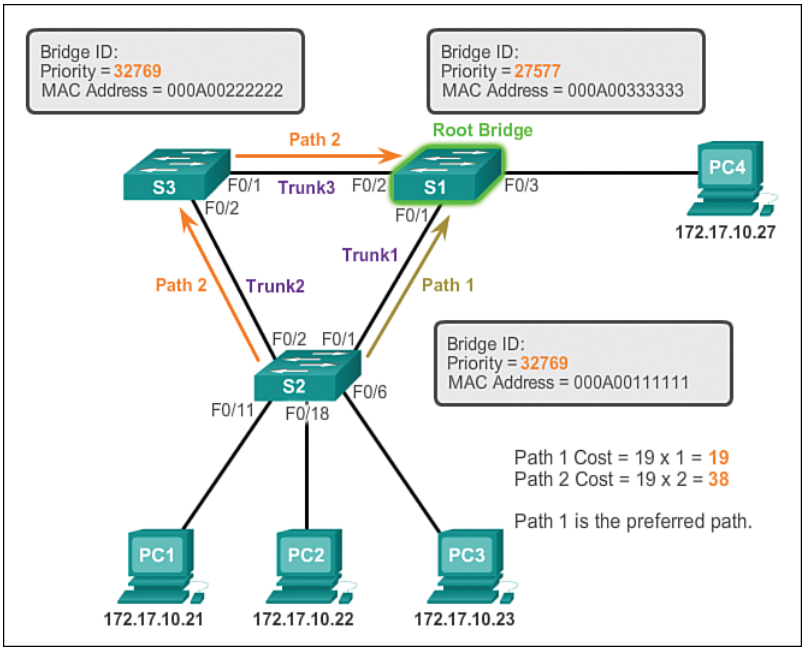


Figure 2-10 Path Cost

Paths with the lowest cost become preferred, and all other redundant paths are blocked. In the example, the path cost from S2 to the root bridge S1, over path 1 is 19 (based on the IEEE-specified individual port cost), while the path cost over path 2 is 38. Because path 1 has a lower overall path cost to the root bridge, it is the preferred path. STP then configures the redundant path to be blocked, preventing a loop from occurring.

To verify the port and path cost to the root bridge, enter the **show spanning-tree** command, as shown in Example 2-2.

Example 2-2 **show spanning-tree** Command

```
S1# show spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    27577
              Address     000A.0033.0033
              Cost        19
              Port        1
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    32769  (priority 32768 sys-id-ext 1)
              Address     000A.0011.1111
```

```
            Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
            Aging Time  15  sec
Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------------
Fa0/1               Root FWD 19        128.1    Edge P2p
Fa0/2               Desg FWD 19        128.2    Edge P2p
```

The Cost field near the top of the output is the total path cost to the root bridge. This value changes depending on how many switch ports must be traversed to get to the root bridge. In the output, each interface is also identified with an individual port cost of 19.

802.1D BPDU Frame Format (2.1.2.5)

The spanning tree algorithm depends on the exchange of BPDUs to determine a root bridge. As shown in Table 2-2, a BPDU frame contains 12 distinct fields that convey path and priority information used to determine the root bridge and paths to the root bridge.

Table 2-2 BPDU Fields

Field Number	Bytes	Field
1–4	2	Protocol ID
	1	Version
	1	Message type
	1	Flags
5–8	8	Root ID
	4	Cost of path
	8	Bridge ID
	2	Port ID
9–12	2	Message age
	2	Max age
	2	Hello time
	2	Forward delay

The first four fields identify the protocol, version, message type, and status flags.

- The next four fields are used to identify the root bridge and the cost of the path 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 (next topic) is retained.

Figure 2-11 shows a BPDU frame that was captured using Wireshark.

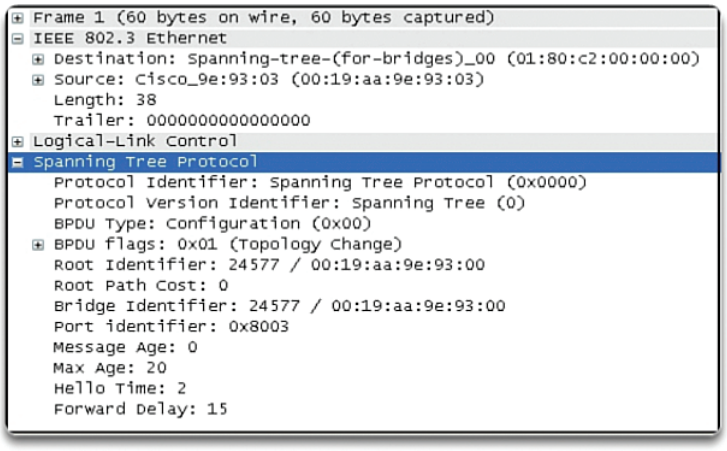


Figure 2-11 Captured BPDU Frame

In the capture, the BPDU frame contains more fields than previously described. The BPDU message is encapsulated in an Ethernet frame when it is transmitted across the network. The 802.3 header indicates the source and destination addresses of the BPDU frame. This frame has a destination MAC address of 01:80:C2:00:00:00, which is a multicast address for the spanning tree group. When a frame is addressed with this MAC address, each switch that is configured for spanning tree accepts and reads the information from the frame; all other devices on the network disregard the frame.

Also note in the capture, the root ID and the BID are the same in the captured BPDU frame. This indicates that the frame was captured from a root bridge. The timers are all set to the default values.

BPDU Propagation and Process (2.1.2.6)

Each switch in the broadcast domain initially assumes that it is the root bridge for a spanning tree instance, so the BPDU frames sent contain the BID of the local switch as the root ID. By default, BPDU frames are sent every two seconds after a switch is booted; that is, the default value of the Hello timer specified in the BPDU frame

is two seconds. Each switch maintains local information about its own BID, the root ID, and the path cost to the root.

When adjacent switches receive a BPDU frame, they compare the root ID from the BPDU frame with the local root ID. If the root ID in the BPDU is lower than the local root ID, the switch updates the local root ID and the ID in its BPDU messages. These messages indicate the new root bridge on the network. The distance to the root bridge is also indicated by the path cost update. For example, if the BPDU was received on a Fast Ethernet switch port, the path cost would increment by 19. If the local root ID is lower than the root ID received in the BPDU frame, the BPDU frame is discarded.

After a root ID has been updated to identify a new root bridge, all subsequent BPDU frames sent from that switch contain the new root ID and updated path cost. That way, all other adjacent switches are able to see the lowest root ID identified at all times. As the BPDU frames pass between other adjacent switches, the path cost is continually updated to indicate the total path cost to the root bridge. Each switch in the spanning tree uses its path costs to identify the best possible path to the root bridge.

The following summarizes the BPDU process:

Note

Priority is the initial deciding factor when electing a root bridge. If the priorities of all the switches are the same, the device with the lowest MAC address becomes the root bridge.

1. Initially, each switch identifies itself as the root bridge. S2 forwards BPDU frames out all switch ports. (See Figure 2-12.)

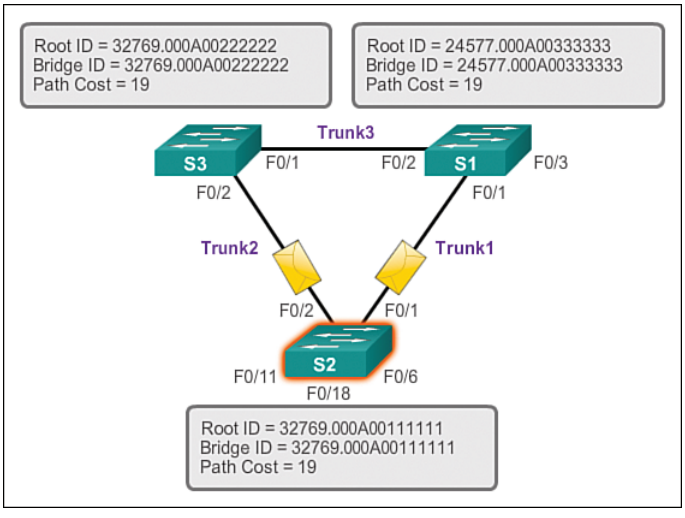


Figure 2-12 BPDU Process: Step 1

2. When S3 receives a BPDU from switch S2, S3 compares its root ID with the BPDU frame it received. The priorities are equal, so the switch is forced to examine the MAC address portion to determine which MAC address has a lower value. Because S2 has a lower MAC address value, S3 updates its root ID with the S2 root ID. At that point, S3 considers S2 as the root bridge. (See Figure 2-13.)

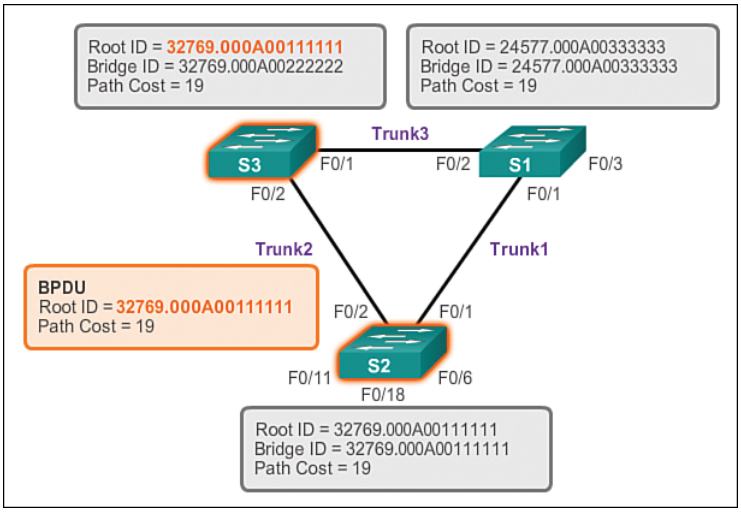


Figure 2-13 BPDU Process: Step 2

3. When S1 compares its root ID with the one in the received BPDU frame, it identifies its local root ID as the lower value and discards the BPDU from S2. (See Figure 2-14.)

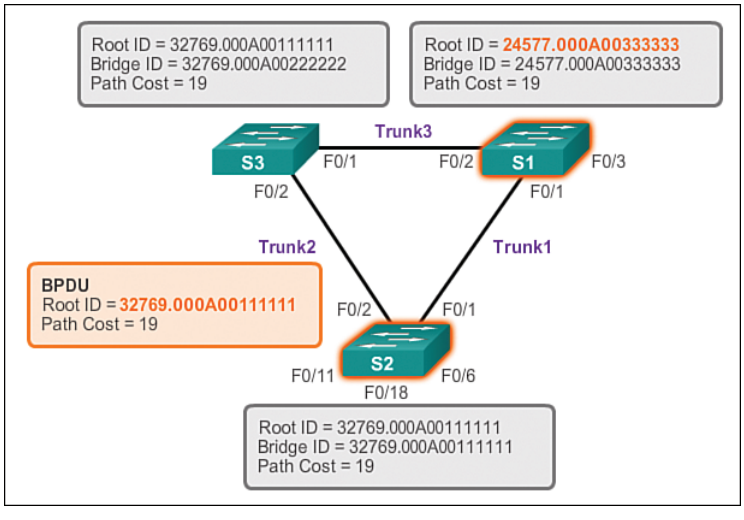


Figure 2-14 BPDU Process: Step 3

4. When S3 sends out its BPDU frames, the root ID contained in the BPDU frame is that of S2. (See Figure 2-15.)

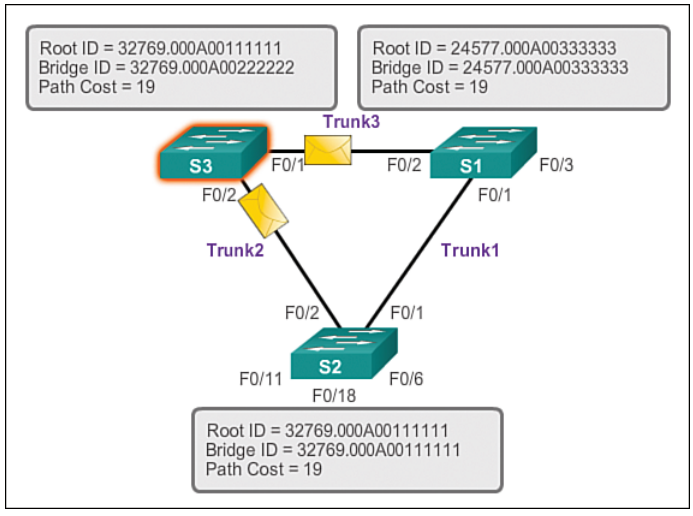


Figure 2-15 BPDU Process: Step 4

5. When S2 receives the BPDU frame, it discards it after verifying that the root ID in the BPDU matched its local root ID. (See Figure 2-16.)

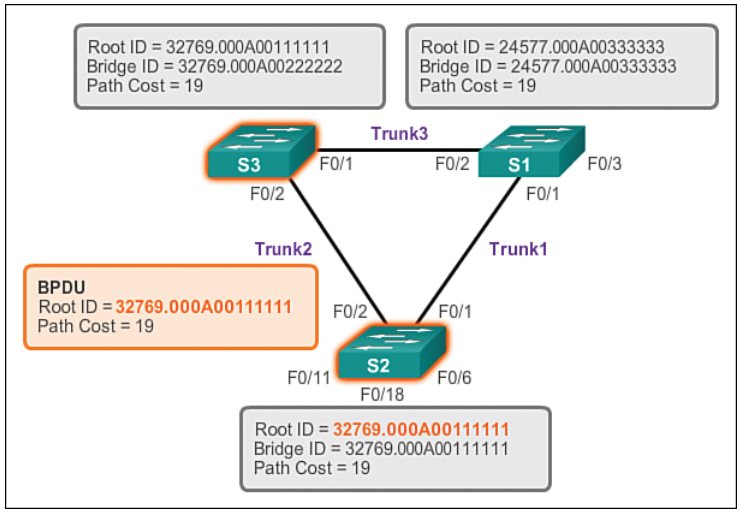


Figure 2-16 BPDU Process: Step 5

6. Because S1 has a lower priority value in its root ID, it discards the BPDU frame received from S3. (See Figure 2-17.)

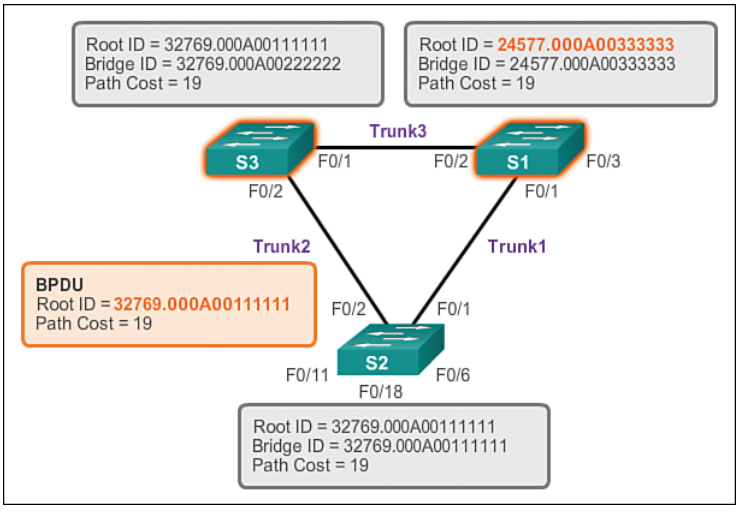


Figure 2-17 BPDU Process: Step 6

7. S1 sends out its BPDU frames. (See Figure 2-18.)

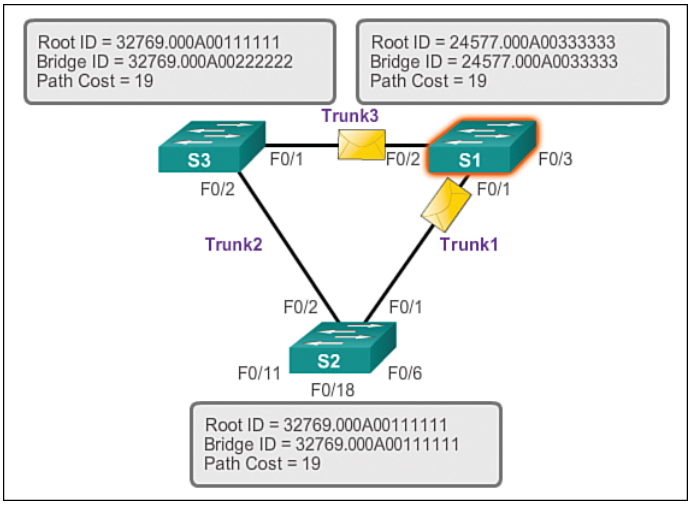


Figure 2-18 BPDU Process: Step 7

8. S3 identifies the root ID in the BPDU frame as having a lower value and, therefore, updates its root ID values to indicate that S1 is now the root bridge. (See Figure 2-19.)

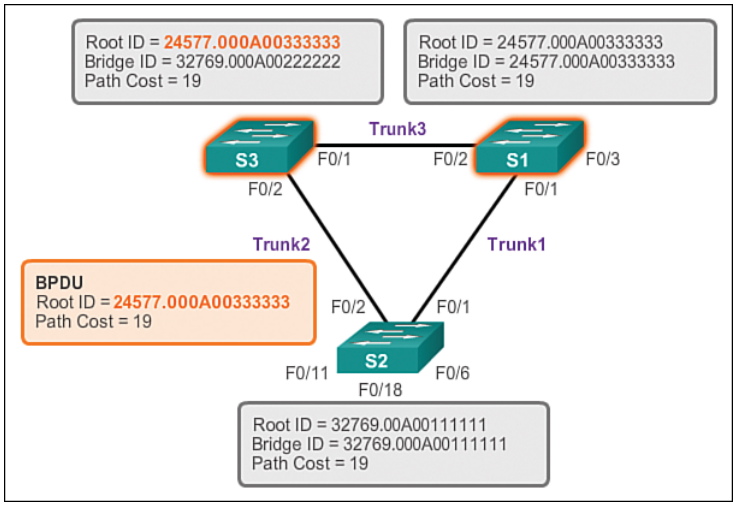


Figure 2-19 BPDU Process: Step 8

9. S2 identifies the root ID in the BPDU frame as having a lower value and, therefore, updates its root ID values to indicate that S1 is now the root bridge. (See Figure 2-20.)

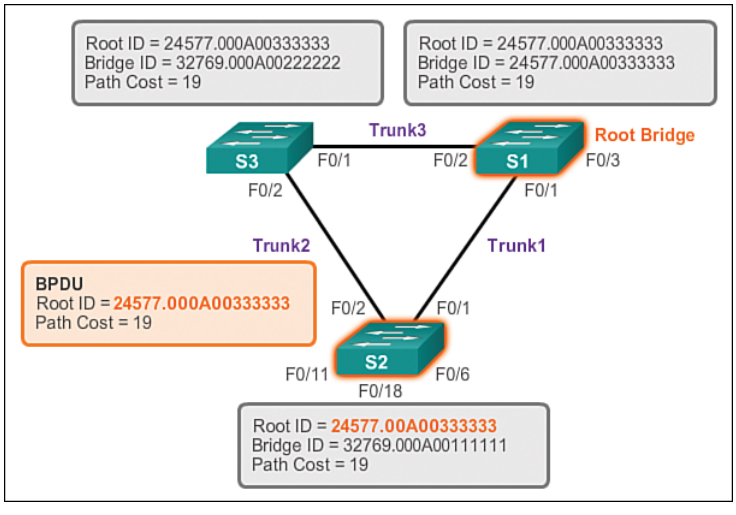


Figure 2-20 BPDU Process: Step 9

Extended System ID (2.1.2.7)

The bridge ID (BID) is used to determine the root bridge on a network. The BID field of a BPDU frame contains three separate fields:

- Bridge priority
- Extended system ID
- MAC address

Each field is used during the root bridge election.

Bridge Priority

The ***bridge priority*** is a configurable value that can be used to influence which switch becomes the root bridge. The switch with the lowest priority, which implies the lowest BID, becomes the root bridge because a lower-priority value takes precedence. For example, to ensure that a specific switch is always the root bridge, set the priority to a lower value than the rest of the switches on the network. The default priority value for all Cisco switches is 32768. The range is 0 to 61440 in increments of 4096. Valid priority values are 0, 4096, 8192, 12288, 16384, 20480, 24576, 28672, 32768, 36864, 40960, 45056, 49152, 53248, 57344, and 61440. All other values are rejected. A bridge priority of 0 takes precedence over all other bridge priorities.

Extended System ID

Early implementations of IEEE 802.1D were designed for networks that did not use VLANs. There was a single common spanning tree across all switches. For this reason, in older Cisco switches, the extended system ID could be omitted in BPDU frames. As VLANs became common for network infrastructure segmentation, 802.1D was enhanced to include support for VLANs, requiring the VLAN ID to be included in the BPDU frame. VLAN information is included in the BPDU frame through the use of the extended system ID. All newer switches include the use of the extended system ID by default.

As shown in Figure 2-21, the bridge priority field is 2 bytes or 16 bits in length; 4 bits are used for the bridge priority and 12 bits for the extended system ID, which identifies the VLAN participating in this particular STP process.

Using these 12 bits for the extended system ID reduces the bridge priority to 4 bits. This process 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}. If the far left bits are 0001, the bridge priority

is 4096; if the far left bits are 1111, the bridge priority is 61440 (= 15 x 4096). The Catalyst 2960 and 3560 Series switches do not allow the configuration of a bridge priority of 65536 (= 16 x 4096) because it assumes the use of a fifth bit that is unavailable because of the use of the extended system ID.

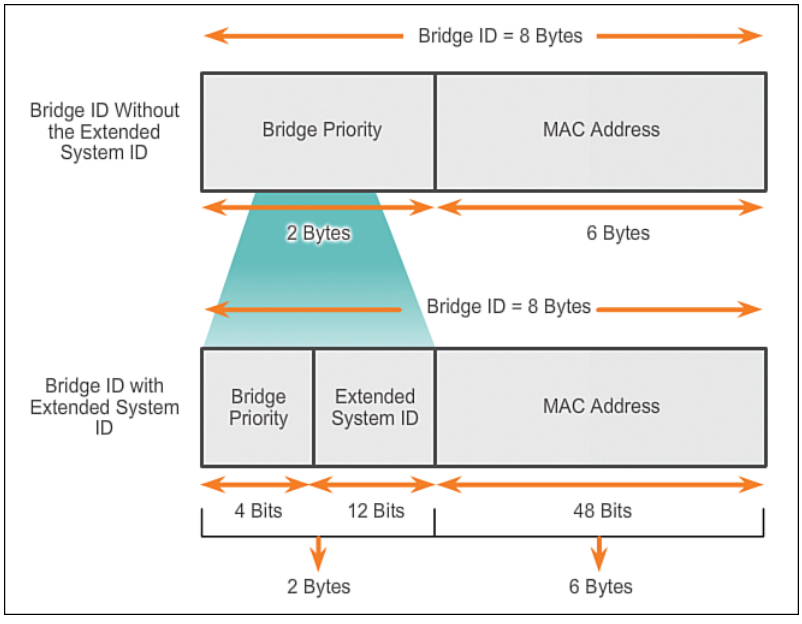


Figure 2-21 BID Fields

The extended system ID value is added to the bridge priority value in the BID to identify the priority and VLAN of the BPDU frame.

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 hexadecimal value will have the lower BID. Initially, all switches are configured with the same default priority value. The MAC address is then the deciding factor on which switch is going to become the root bridge. To ensure that the root bridge decision best meets network requirements, it is recommended that the administrator configure the desired root bridge switch with a lower priority. This also ensures that the addition of new switches to the network does not trigger a new spanning tree election, which can disrupt network communication while a new root bridge is being selected.

In Figure 2-22, S1 has a lower priority than the other switches; therefore, it is preferred as the root bridge for that spanning tree instance.

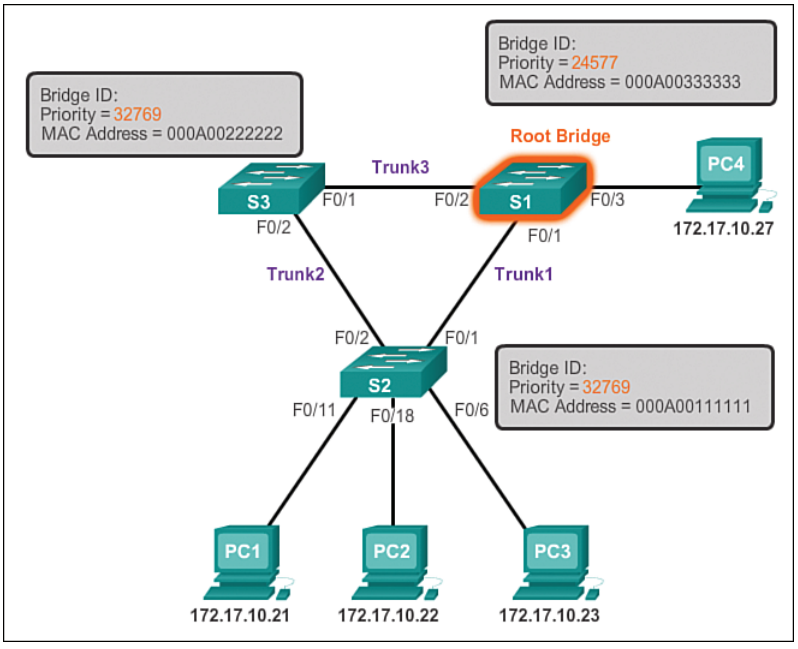


Figure 2-22 Priority-Based Decision

When all switches are configured with the same priority, as is the case with all switches kept in the default configuration with a priority of 32768, the MAC address becomes the deciding factor for which switch becomes the root bridge, as shown in Figure 2-23.

Note

In the example, the priority of all the switches is 32769. The value is based on the 32768 default priority and the VLAN 1 assignment associated with each switch (32768+1).

The MAC address with the lowest hexadecimal value is considered to be the preferred root bridge. In the example, S2 has the lowest value for its MAC address and is, therefore, designated as the root bridge for that spanning tree instance.

Interactive Graphic

Activity 2.1.2.8: Identify 802.1D Port Roles

Go to the course online to perform this practice activity.

Video

Video Demonstration 2.1.2.9: Observing Spanning Tree Protocol Operation

View the video in the online course for an understanding of STP operation.

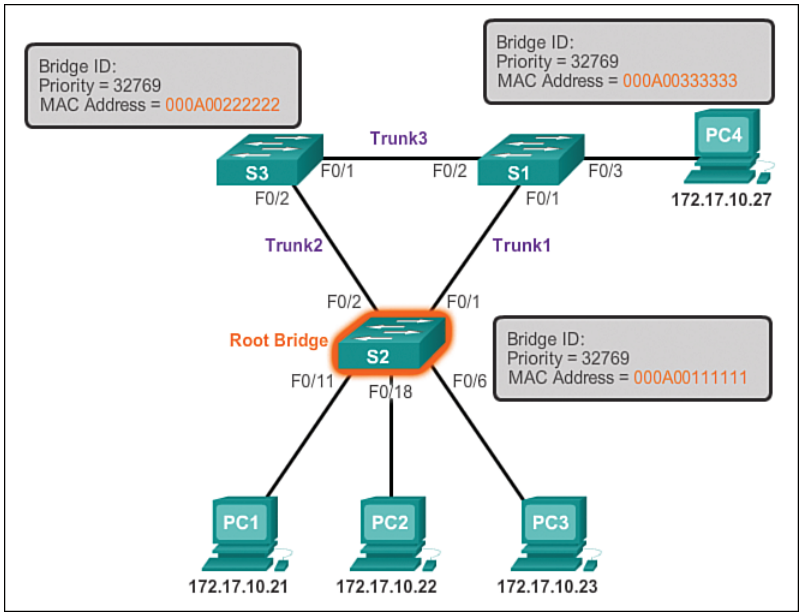


Figure 2-23 MAC-Based Decision

Lab 2.1.2.10: Building a Switched Network with Redundant Links

In this lab, you will complete the following objectives:

- Part 1: Build the Network and Configure Basic Device Settings
- Part 2: Determine the Root Bridge
- Part 3: Observe STP Port Selection Based on Port Cost
- Part 4: Observe STP Port Selection Based on Port Priority

Varieties of Spanning Tree Protocols (2.2)

STP has evolved into several different versions since the original specification. Some versions are IEEE standards, while others are proprietary. This section reviews the features unique to each of the more popular STP versions.

Overview (2.2.1)

To begin to understand the scope of STP versions available, let's briefly look at a list of all of them.

List of Spanning Tree Protocols (2.2.1.1)

Several varieties of spanning tree protocols have emerged since the original IEEE 802.1D.

The varieties of spanning tree protocols include

- **STP:** This is the original IEEE 802.1D version (802.1D-1998 and earlier) that provides a loop-free topology in a network with redundant links. *Common Spanning Tree (CST)* assumes one spanning tree instance for the entire bridged network, regardless of the number of VLANs.
- *PVST+*: This is a Cisco enhancement of STP that provides a separate 802.1D spanning tree instance for each VLAN configured in the network. The separate instance supports *PortFast*, UplinkFast, BackboneFast, *BPDU guard*, BPDU filter, root guard, and loop guard.
- **802.1D-2004:** This is an updated version of the STP standard, incorporating *IEEE 802.1w*.
- **Rapid Spanning Tree Protocol (RSTP) or IEEE 802.1w:** This is an evolution of STP that provides faster convergence than STP.
- *Rapid PVST+*: This is a Cisco enhancement of RSTP that uses PVST+. Rapid PVST+ provides a separate instance of 802.1w per VLAN. The separate instance supports PortFast, BPDU guard, BPDU filter, root guard, and loop guard.
- **Multiple Spanning Tree Protocol (MSTP):** 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 and combines many VLANs with the same physical and logical topology into a common RSTP instance. Each instance supports PortFast, BPDU guard, BPDU filter, root guard, and loop guard.

A network professional, whose duties include switch administration, might be required to decide which type of spanning tree protocol to implement.

Note

The legacy Cisco-proprietary features UplinkFast and BackboneFast are not described in this course. These features are superseded by the implementation of Rapid PVST+, which incorporates these features as part of the implementation of the RSTP standard.

Characteristics of the Spanning Tree Protocols (2.2.1.2)

These are characteristics of the various spanning tree protocols:

- **STP:** Assumes one *IEEE 802.1D* spanning tree instance for the entire bridged network, regardless of the number of VLANs. Because there is only one instance, the CPU and memory requirements for this version are lower than for the other protocols. However, because there is only one instance, there is only one root bridge and one tree. Traffic for all VLANs flows over the same path, which can lead to suboptimal traffic flows. Because of the limitations of 802.1D, this version is slow to converge.
- **PVST+:** A Cisco enhancement of STP that provides a separate instance of the Cisco implementation of 802.1D for each VLAN that is configured in the network. The separate instance supports PortFast, UplinkFast, BackboneFast, BPDU guard, BPDU filter, root guard, and loop guard. Creating an instance for each VLAN increases the CPU and memory requirements, but allows for per-VLAN root bridges. This design allows the spanning tree to be optimized for the traffic of each VLAN. Convergence of this version is similar to the convergence of 802.1D. However, convergence is per-VLAN.
- **RSTP** (or *IEEE 802.1w*)**:** An evolution of spanning tree that provides faster convergence than the original 802.1D implementation. This version addresses many convergence issues, but because it still provides a single instance of STP, it does not address the suboptimal traffic flow issues. To support that faster convergence, the CPU usage and memory requirements of this version are slightly higher than those of CST, but less than those of RSTP+.
- **Rapid PVST+:** A Cisco enhancement of RSTP that uses PVST+. It provides a separate instance of 802.1w per VLAN. The separate instance supports PortFast, BPDU guard, BPDU filter, root guard, and loop guard. This version addresses both the convergence issues and the suboptimal traffic flow issues. However, this version has the largest CPU and memory requirements.
- **MSTP:** The *IEEE 802.1s* standard, inspired by the earlier Cisco-proprietary MISTP implementation. To reduce the number of required STP instances, MSTP maps multiple VLANs that have the same traffic flow requirements into the same spanning tree instance.
- **MST:** The Cisco implementation of MSTP, which provides up to 16 instances of RSTP (802.1w) and combines many VLANs with the same physical and logical topology into a common RSTP instance. Each instance supports PortFast, BPDU guard, BPDU filter, root guard, and loop guard. The CPU and memory requirements of this version are less than those of Rapid PVST+, but more than those of RSTP.

Table 2-3 summarizes these STP characteristics.

Table 2-3 Spanning Tree Protocol Characteristics

Protocol	Standard	Resources Needed	Convergence	Tree Calculation
STP	802.1D	Low	Slow	All VLANs
PVST+	Cisco	High	Slow	Per VLAN
RSTP	802.1w	Medium	Fast	All VLANs
Rapid PVST+	Cisco	Very high	Fast	Per VLAN
MSTP	802.1s, Cisco	Medium or high	Fast	Per Instance

The default spanning tree mode for Cisco Catalyst switches is PVST+, which is enabled on all ports. PVST+ has much slower convergence after a topology change than Rapid PVST+.

Note

It is important to distinguish between the legacy IEEE 802.1D-1998 (and earlier) standard and the IEEE 802.1D-2004 standard. IEEE 802.1D-2004 incorporates RSTP functionality, while IEEE 802.1D-1998 refers to the original implementation of the spanning tree algorithm. Newer Cisco switches running newer versions of the IOS, such as Catalyst 2960 switches with IOS 15.0, run PVST+ by default, but incorporate many of the specifications of IEEE 802.1D-1998 in this mode (such as alternate ports in place of the former nondesignated ports). But to run rapid spanning tree on such a switch, it still must be explicitly configured for rapid spanning tree mode.

Interactive Graphic

Activity 2.2.1.3: Identify Types of Spanning Tree Protocols

Go to the course online to perform this practice activity.

PVST+ (2.2.2)

PVST+ is a Cisco implementation of STP and is the default STP mode on Cisco Catalyst switches.

Overview of PVST+ (2.2.2.1)

The original IEEE 802.1D standard defines a Common Spanning Tree (CST) that assumes only one spanning tree instance for the entire switched network, regardless of the number of VLANs. A network running CST has these characteristics:

- No load sharing is possible. One uplink must block for all VLANs.
- The CPU is spared. Only one instance of spanning tree must be computed.

Cisco developed PVST+ so that a network can run an independent instance of the Cisco implementation of IEEE 802.1D for each VLAN in the network. With PVST+, it is possible for one trunk port on a switch to be blocking for a VLAN while not blocking for other VLANs. PVST+ can be used to implement Layer 2 load balancing. Because each VLAN runs a separate instance of STP, the switches in a PVST+ environment require greater CPU process and BPDU bandwidth consumption than a traditional CST implementation of STP.

In a PVST+ environment, spanning tree parameters can be tuned so that half of the VLANs forward on each uplink trunk. In Figure 2-24, port F0/3 on S2 is the forwarding port for VLAN 20, and F0/2 on S2 is the forwarding port for VLAN 10.

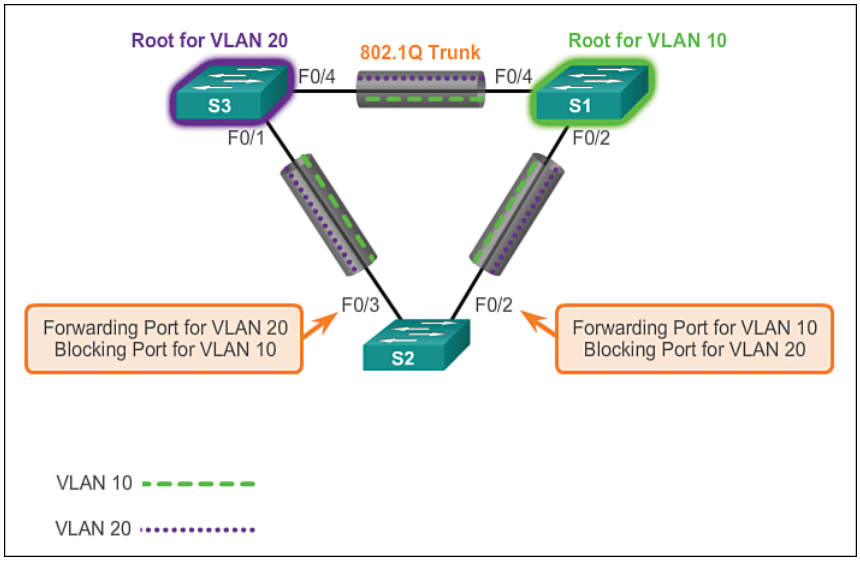


Figure 2-24 PVST+ Example

This is accomplished by configuring one switch to be elected the root bridge for half of the VLANs in the network, and a second switch to be elected the root bridge for the other half of the VLANs. In the figure, S3 is the root bridge for VLAN 20, and S1 is the root bridge for VLAN 10. Multiple STP root bridges per VLAN increase redundancy in the network.

Networks running PVST+ have these characteristics:

- Optimum load balancing can result.
- One spanning tree instance for each VLAN maintained can mean a considerable waste of CPU cycles for all the switches in the network (in addition to the bandwidth that is used for each instance to send its own BPDU). This would only be problematic if a large number of VLANs are configured.

Port States and PVST+ Operation (2.2.2.2)

STP facilitates the logical loop-free path throughout the broadcast domain. The spanning tree is determined through the information learned by the exchange of the BPDU frames between the interconnected switches. To facilitate the learning of the logical spanning tree, each switch port transitions through five possible port states and three BPDU timers.

The spanning tree is determined immediately after a switch is finished booting up. If a switch port transitions directly from the blocking to the forwarding state without information about the full topology during the transition, the port can temporarily create a data loop. For this reason, STP introduces the following five port states that ensure that no loops are created during the creation of the logical spanning tree:

- **Blocking:** The port is an alternate port and does not participate in frame forwarding. The port receives BPDU frames to determine the location and root ID of the root bridge switch and what port roles each switch port should assume in the final active STP topology.
- ***Listening***: Listens for the path to the root. STP has determined that the port can participate in frame forwarding according to the BPDU frames that the switch has received thus far. At this point, the switch port not only receives BPDU frames, but it also transmits its own BPDU frames and informs adjacent switches that the switch port is preparing to participate in the active topology.
- ***Learning***: Learns the MAC addresses. The port prepares to participate in frame forwarding and begins to populate the MAC address table.
- ***Forwarding***: The port is considered part of the active topology. It forwards data frames and sends and receives BPDU frames.
- ***Disabled***: The Layer 2 port does not participate in spanning tree and does not forward frames. The disabled state is set when the switch port is administratively disabled.

Table 2-4 summarizes the operations that are allowed during each port state.

Table 2-4 Port States

Operation Allowed	Port State				
	Blocking	Listening	Learning	Forwarding	Disabled
Can receive and process BPDUs	yes	yes	yes	no	—
Can forward data frames received on interface	no	no	no	yes	no

Operation Allowed	Port State				
	Blocking	Listening	Learning	Forwarding	Disabled
Can forward data frames switched from another interface	no	no	no	yes	no
Can learn MAC addresses	no	no	yes	yes	no

Note that the number of ports in each of the various states (blocking, listening, learning, or forwarding) can be displayed with the **show spanning-tree summary** command.

For each VLAN in a switched network, PVST+ performs four steps to provide a loop-free logical network topology:

1. **Elects one root bridge:** Only one switch can act as the root bridge (for a given VLAN). The root bridge is the switch with the lowest bridge ID. On the root bridge, all ports are designated ports (in particular, no root ports).
2. **Selects the root port on each nonroot bridge:** STP establishes one root port on each nonroot bridge. The root port is the lowest-cost path from the nonroot bridge to the root bridge, indicating the direction of the best path to the root bridge. Root ports are normally in the forwarding state.
3. **Selects the designated port on each segment:** On each link, STP establishes one designated port. The designated port is selected on the switch that has the lowest-cost path to the root bridge. Designated ports are normally in the forwarding state, forwarding traffic for the segment.
4. **The remaining ports in the switched network are alternate ports:** Alternate ports normally remain in the blocking state, to logically break the loop topology. When a port is in the blocking state, it does not forward traffic, but can still process received BPDU messages.

Extended System ID and PVST+ Operation (2.2.2.3)

In a PVST+ environment, the extended switch ID, shown in Figure 2-25, ensures that each switch has a unique BID for each VLAN.

For example, the VLAN 2 default BID would be 32770 (priority 32768, plus the extended system ID of 2). If no priority has been configured, every switch has the same default priority, and the election of the root for each VLAN is based on the MAC address. This method is a random means of selecting the root bridge.

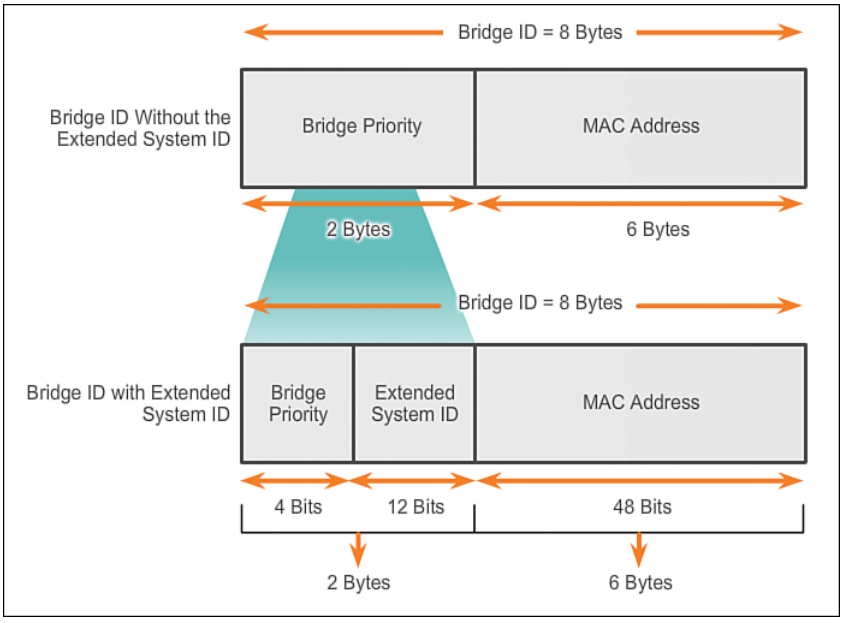


Figure 2-25 PVST+ and the Extended System ID

There are situations where the administrator might want a specific switch to be selected as the root bridge. This can be for a variety of reasons, including the switch is more centrally located within the LAN design, the switch has higher processing power, or the switch is simply easier to access and manage remotely. To manipulate the root bridge election, simply assign a lower priority to the switch that should be selected as the root bridge.

Interactive Graphic

Activity 2.2.2.4: Identifying PVST+ Operation

Go to the course online to perform this practice activity.

Rapid PVST+ (2.2.3)

Rapid PVST+ is the Cisco-proprietary implementation of RSTP.

Overview of Rapid PVST+ (2.2.3.1)

RSTP (IEEE 802.1w) is an evolution of the original 802.1D standard and is incorporated into the IEEE 802.1D-2004 standard. The 802.1w STP terminology remains primarily the same as the original IEEE 802.1D STP terminology. Most parameters have been left unchanged, so users familiar with STP can easily configure the new

protocol. Rapid PVST+ is simply the Cisco implementation of RSTP on a per-VLAN basis. With Rapid PVST+, an independent instance of RSTP runs for each VLAN.

Figure 2-26 shows a network running RSTP.

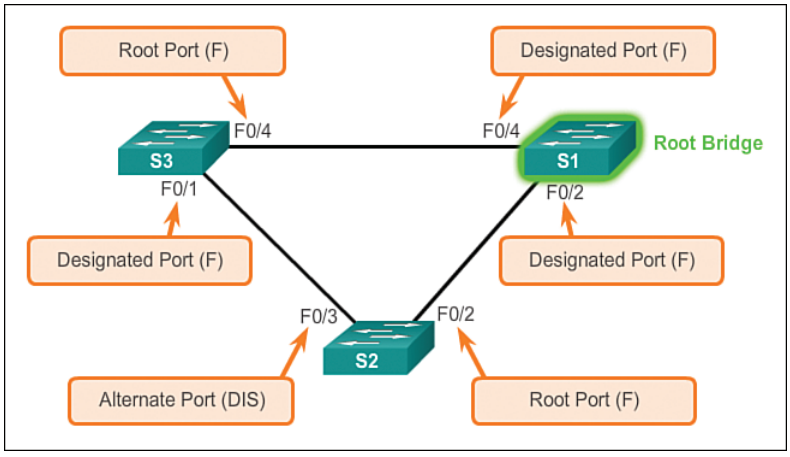


Figure 2-26 RSTP Port Roles

S1 is the root bridge with two designated ports in a forwarding state. RSTP supports a new port type: Port F0/3 on S2 is an alternate port in discarding state. Notice that there are no blocking ports. RSTP does not have a blocking port state. RSTP defines port states as discarding, learning, or forwarding.

RSTP speeds the recalculation of the spanning tree when the Layer 2 network topology changes. RSTP can achieve much faster convergence in a properly configured network, sometimes in as little as a few hundred milliseconds. RSTP redefines the type of ports and their state. If a port is configured to be an alternate port or a backup port, it can immediately change to forwarding state without waiting for the network to converge. The following briefly describes RSTP characteristics:

- RSTP is the preferred protocol for preventing Layer 2 loops in a switched network environment. Many of the differences were established by Cisco-proprietary enhancements to the original 802.1D. These enhancements, such as BPDUs carrying and sending information about port roles only to neighboring switches, require no additional configuration and generally perform better than the earlier Cisco-proprietary versions. They are now transparent and integrated in the protocol's operation.
- Cisco-proprietary enhancements to the original 802.1D, such as UplinkFast and BackboneFast, are not compatible with RSTP.
- RSTP (802.1w) supersedes the original 802.1D while retaining backward compatibility. Much of the original 802.1D terminology remains and most parameters

are unchanged. In addition, 802.1w is capable of reverting to legacy 802.1D to interoperate with legacy switches on a per-port basis. For example, the RSTP spanning tree algorithm elects a root bridge in exactly the same way as the original 802.1D.

- RSTP keeps the same BPDU format as the original IEEE 802.1D, except that the version field is set to 2 to indicate RSTP, and the flags field uses all 8 bits.
- RSTP is able to actively confirm that a port can safely transition to the forwarding state without having to rely on any timer configuration.

RSTP BPDU (2.2.3.2)

RSTP uses type 2, version 2 BPDUs. The original 802.1D STP uses type 0, version 0 BPDUs. However, a switch running RSTP can communicate directly with a switch running the original 802.1D STP. RSTP sends BPDUs and populates the flag byte in a slightly different manner than in the original 802.1D:

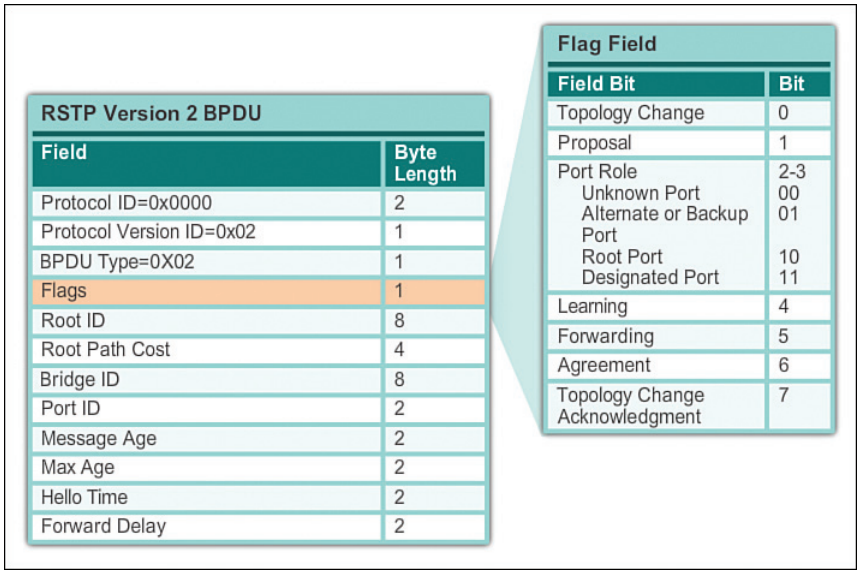
- Protocol information can be immediately aged on a port if Hello packets are not received for three consecutive Hello times, six seconds by default, or if the max age timer expires.
- Because BPDUs are used as a keepalive mechanism, three consecutively missed BPDUs indicate lost connectivity between a bridge and its neighboring root or designated bridge. The fast aging of the information allows failures to be detected quickly.

Note

Like STP, an RSTP switch sends a BPDU with its current information every Hello time period (two seconds, by default), even if the RSTP bridge does not receive any BPDUs from the root bridge.

As shown in Figure 2-27, RSTP uses the flag byte of version 2 BPDU:

- Bits 0 and 7 are used for topology change and acknowledgment as they are in the original 802.1D.
- Bits 1 and 6 are used for the Proposal Agreement process (used for rapid convergence).
- Bits from 2 to 5 encode the role and state of the port.
- Bits 4 and 5 are used to encode the port role using a 2-bit code.



RSTP Version 2 BPDU

Field	Byte Length
Protocol ID=0x0000	2
Protocol Version ID=0x02	1
BPDU Type=0X02	1
Flags	1
Root ID	8
Root Path Cost	4
Bridge ID	8
Port ID	2
Message Age	2
Max Age	2
Hello Time	2
Forward Delay	2

Flag Field

Field Bit	Bit
Topology Change	0
Proposal	1
Port Role Unknown Port Alternate or Backup Port Root Port Designated Port	2-3 00 01 10 11
Learning	4
Forwarding	5
Agreement	6
Topology Change Acknowledgment	7

Figure 2-27 RSTP BPDU

Edge Ports (2.2.3.3)

An RSTP *edge port* is a switch port that is never intended to be connected to another switch device. 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. RSTP edge ports should immediately transition to the forwarding state, thereby skipping the time-consuming original 802.1D listening and learning port states.

The Cisco RSTP implementation, Rapid PVST+, maintains the PortFast keyword, using the **spanning-tree portfast** command for edge port configuration. This makes the transition from STP to RSTP seamless.

Figure 2-28 shows examples of ports that can be configured as edge ports.

Figure 2-29 shows examples of ports that are nonedge ports.

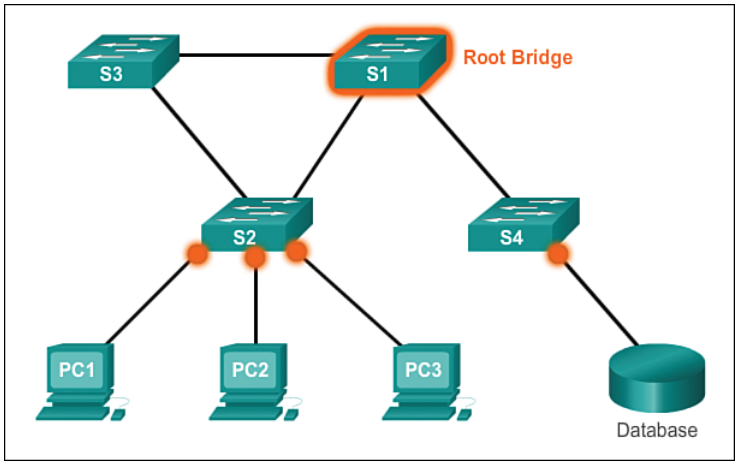


Figure 2-28 Edge Ports

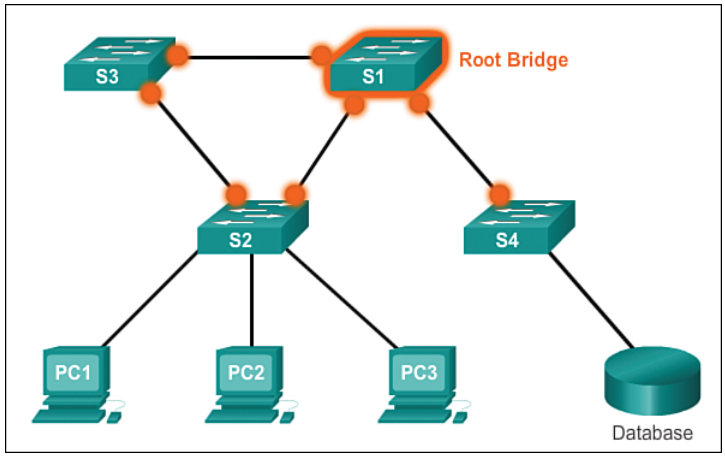


Figure 2-29 Nonedge Ports

Note

Configuring an edge port to be attached to another switch is not recommended. This can have negative implications for RSTP because a temporary loop can result, possibly delaying the convergence of RSTP.

Link Types (2.2.3.4)

The link type provides a categorization for each port participating in RSTP by using the duplex mode on the port. Depending on what is attached to each port, two different link types can be identified, as shown in Figure 2-30:

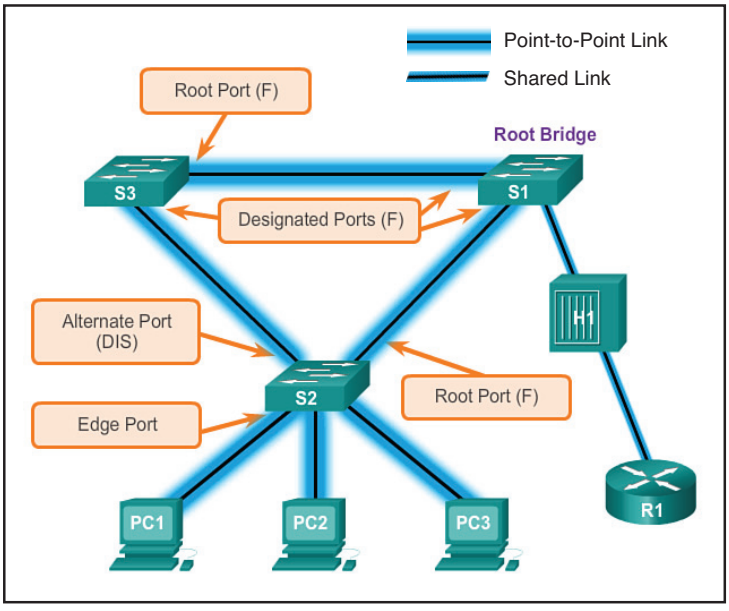


Figure 2-30 RSTP Link Types

- *Point-to-Point Link*: A port operating in full-duplex mode typically connects a switch to a switch and is a candidate for rapid transition to a forwarding state.
- *Shared Link*: A port operating in half-duplex mode connects a switch to a hub that attaches multiple devices.

The link type can determine whether the port can immediately transition to a forwarding state, assuming that certain conditions are met. These conditions are different for edge ports and nonedge ports. Nonedge ports are categorized into two link types, point-to-point and shared. The link type is automatically determined, but can be overridden with an explicit port configuration using the **spanning-tree link-type** *parameter* command.

Edge port connections and point-to-point connections are candidates for rapid transition to the forwarding state. However, before the link-type parameter is considered, RSTP must determine the port role. Characteristics of port roles with regard to link types include the following:

- Root ports do not use the link-type parameter. Root ports are able to make a rapid transition to the forwarding state as soon as the port is in sync.
- Alternate and backup ports do not use the link-type parameter in most cases.
- Designated ports make the most use of the link-type parameter. Rapid transition to the forwarding state for the designated port occurs only if the link-type parameter is set to *point-to-point*.

Interactive Graphic

Activity 2.2.3.5: Identify Port Roles in Rapid PVST+

Go to the course online to perform this practice activity.

Interactive Graphic

Activity 2.2.3.6: Compare PVST+ and Rapid PVST+

Go to the course online to perform this practice activity.

Spanning Tree Configuration (2.3)

Although STP runs by default, there are some configurations that allow the network administrator to modify the version and behavior of STP, including root bridge election, speeding up convergence, and load balancing.

PVST+ Configuration (2.3.1)

In this topic, we review the commands to modify the default PVST+ configuration.

Catalyst 2960 Default Configuration (2.3.1.1)

Table 2-5 shows the default spanning tree configuration for a Cisco Catalyst 2960 Series switch. Notice that the default spanning tree mode is PVST+.

Table 2-5 Default Switch Configuration

Feature	Default Setting
Enable state	Enabled on VLAN 1
Spanning-tree mode	PVST+
Switch priority	32768
Spanning-tree port priority (configurable on a per-interface basis)	128
Spanning-tree port cost (configurable on a per-interface basis)	1000 Mbps: 4 100 Mbps: 19 10 Mbps: 100
Spanning-tree VLAN port priority (configurable on a per-VLAN basis)	128

Feature	Default Setting
Spanning-tree VLAN port cost (configurable on a per-VLAN basis)	1000 Mbps: 4
	100 Mbps: 19
	10 Mbps: 100
Spanning-tree timers	Hello time: 2 seconds
	Forward-delay time: 15 seconds
	Maximum-aging time: 20 seconds
	Transmit hold count: 6 BPDUs

Configuring and Verifying the Bridge ID (2.3.1.2)

When an administrator wants a specific switch to become a root bridge, the bridge priority value must be adjusted to ensure that it is lower than the bridge priority values of all the other switches on the network. There are two different methods to configure the bridge priority value on a Cisco Catalyst switch, as shown in Figure 2-31.

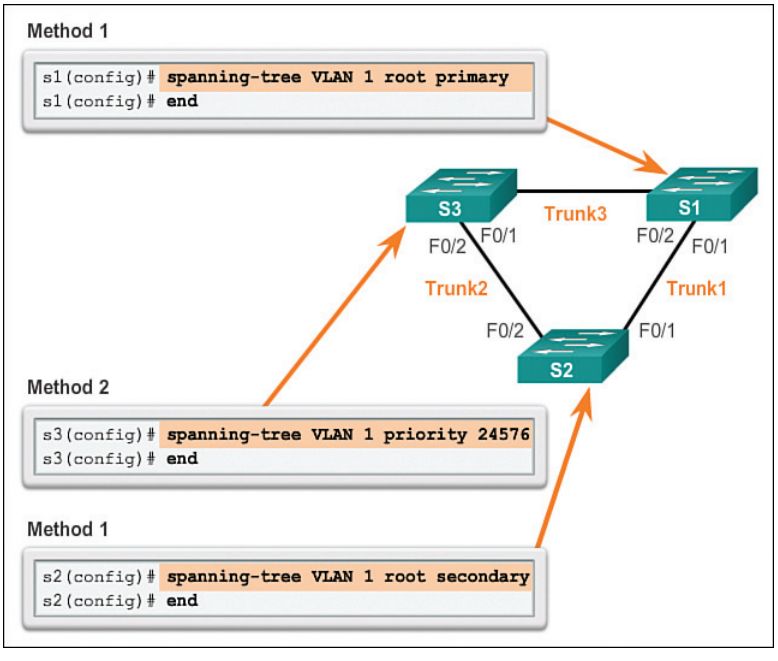


Figure 2-31 Methods for Configuring the BID

Method 1

To ensure that the switch has the lowest bridge priority value, use the **spanning-tree vlan** *vlan-id* **root primary** command in global configuration mode. The priority for the switch is set to the predefined value of 24,576 or to the highest multiple of 4,096, less than the lowest bridge priority detected on the network.

If an alternate root bridge is desired, use the **spanning-tree vlan** *vlan-id* **root secondary** global configuration mode command. This command sets the priority for the switch to the predefined value of 28,672. This ensures that the alternate switch becomes the root bridge if the primary root bridge fails. This assumes that the rest of the switches in the network have the default 32,768 priority value defined.

In Figure 2-31, S1 has been assigned as the primary root bridge using the **spanning-tree vlan 1 root primary** command, and S2 has been configured as the secondary root bridge using the **spanning-tree vlan 1 root secondary** command.

Method 2

Another method for configuring the bridge priority value is using the **spanning-tree vlan** *vlan-id* **priority** *value* global configuration mode command. This command gives more granular control over the bridge priority value. The priority value is configured in increments of 4,096 between 0 and 61,440.

In Figure 2-31, S3 has been assigned a bridge priority value of 24,576 using the **spanning-tree vlan 1 priority 24576** command.

To verify the bridge priority of a switch, use the **show spanning-tree** command. In Example 2-3, the priority of the switch has been set to 24,576. Also notice that the switch is designated as the root bridge for the spanning tree instance.

Example 2-3 Verifying That S3 Is the Root Bridge

```
S3# show spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    24577
              Address     000A.0033.0033
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    24577  (priority 24576 sys-id-ext 1)
              Address     000A.0033.3333
              Hello Time  2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  300
```

```
Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- ------------------------
Fa0/1               Desg FWD 4         128.1    P2p
Fa0/2               Desg FWD 4         128.2    P2p
```

PortFast and BPDU Guard (2.3.1.3)

PortFast is a Cisco feature for PVST+ environments. When a switch port is configured with PortFast, that port transitions from blocking to forwarding state immediately, bypassing the usual 802.1D STP transition states (the listening and learning states). You can use PortFast on access ports to allow these devices to connect to the network immediately, rather than waiting for IEEE 802.1D STP to converge on each VLAN. Access ports are ports that are connected to a single workstation or to a server, as shown in Figure 2-32.

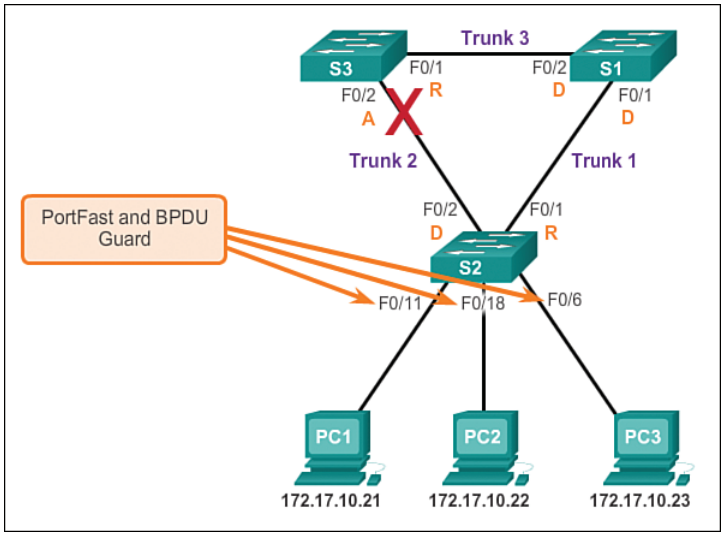


Figure 2-32 PortFast and BPDU Guard

In a valid PortFast configuration, BPDUs should never be received, because that would indicate that another bridge or switch is connected to the port, potentially causing a spanning tree loop. Cisco switches support a feature called BPDU guard. When it is enabled, BPDU guard puts the port in an *error-disabled* state on receipt of a BPDU. This will effectively shut down the port. The BPDU guard feature provides a secure response to invalid configurations because you must manually put the interface back into service.

Cisco PortFast technology is useful for DHCP. Without PortFast, a PC can send a DHCP request before the port is in forwarding state, denying the host from getting a usable IP address and other information. Because PortFast immediately changes the state to forwarding, the PC always gets a usable IP address.

Note

Because the purpose of PortFast is to minimize the time that access ports must wait for spanning tree to converge, it should only be used on access ports. If you enable PortFast on a port connecting to another switch, you risk creating a spanning tree loop.

To configure PortFast on a switch port, enter the **spanning-tree portfast** interface configuration mode command on each interface that PortFast is to be enabled. The **spanning-tree portfast default** global configuration mode command enables PortFast on all nontrunking interfaces.

To configure BPDU guard on a Layer 2 access port, use the **spanning-tree bpduguard enable** interface configuration mode command. The **spanning-tree portfast bpduguard default** global configuration command enables BPDU guard on all PortFast-enabled ports.

To verify that PortFast and BPDU guard have been enabled for a switch port, use the **show running-config** command. PortFast and BPDU guard are disabled, by default, on all interfaces.

In Example 2-4, the FastEthernet 0/11 interface is configured with PortFast and BPDU guard.

Example 2-4 Configuring and Verifying PortFast and BPDU Guard

```
S2(config)# interface FastEthernet 0/11
S2(config-if)# spanning-tree portfast
%Warning: portfast should only be enabled on ports connected to a single
 host. Connecting hubs, concentrators, switches, bridges, etc... to this
 interface  when portfast is enabled, can cause temporary bridging loops.
 Use with CAUTION

%Portfast has been configured on FastEthernet0/11 but will only
 have effect when the interface is in a non-trunking mode.

S2(config-if)# spanning-tree bpduguard enable
S2(config-if)# end
S2# show running-config interface f0/11
Building configuration...

Current configuration : 90 bytes
```

```
!
interface FastEthernet0/11
 spanning-tree portfast
 spanning-tree bpduguard enable
end
```

PVST+ Load Balancing (2.3.1.4)

The topology in Figure 2-33 shows three switches with 802.1Q trunks connecting them.

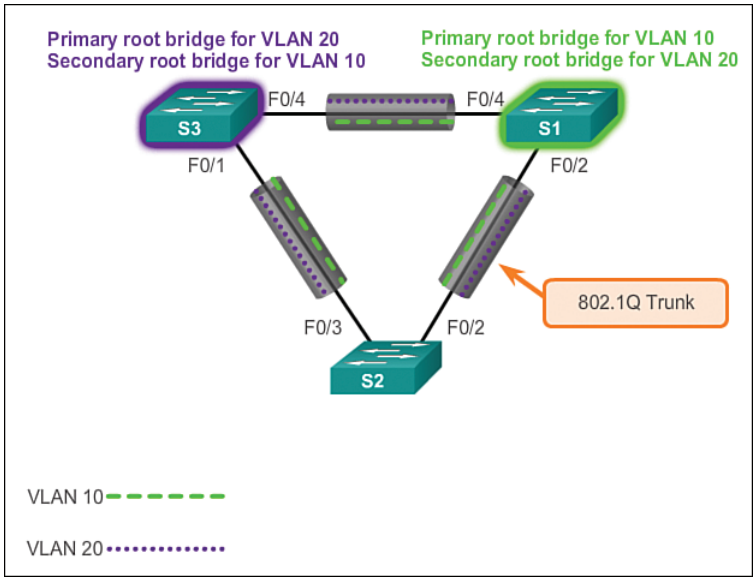


Figure 2-33 Configure PVST+

There are two VLANs, 10 and 20, that are being trunked across these links. The goal is to configure S3 as the root bridge for VLAN 20 and S1 as the root bridge for VLAN 10. Port F0/3 on S2 is the forwarding port for VLAN 20 and the blocking port for VLAN 10. Port F0/2 on S2 is the forwarding port for VLAN 10 and the blocking port for VLAN 20.

In addition to establishing a root bridge, it is also possible to establish a secondary root bridge. A secondary root bridge is a switch that can become the root bridge for a VLAN if the primary root bridge fails. Assuming that the other bridges in the VLAN retain their default STP priority, this switch becomes the root bridge if the primary root bridge fails.

The steps to configure PVST+ on this example topology are

Step 1. Select the switches that you want for the primary and secondary root bridges for each VLAN. For example, in Figure 2-33, S3 is the primary bridge for VLAN 20 and S1 is the secondary bridge for VLAN 20.

Step 2. Configure the switch to be a primary bridge for the VLAN by using the **spanning-tree vlan** *number* **root primary** command, as shown in Example 2-5.

Step 3. Configure the switch to be a secondary bridge for the VLAN by using the **spanning-tree vlan** *number* **root secondary** command.

Another way to specify the root bridge is to set the spanning tree priority on each switch to the lowest value so that the switch is selected as the primary bridge for its associated VLAN.

Notice that in Example 2-5, S3 is configured as the primary root bridge for VLAN 20 and S1 is configured as the primary root bridge for VLAN 10. S2 retained its default STP priority.

Example 2-5 Configuring Primary and Secondary Root Bridge for Each VLAN

```
S3(config)# spanning-tree vlan 20 root primary
S3(config)# spanning-tree vlan 10 root secondary
```

```
S1(config)# spanning-tree vlan 10 root primary
S1(config)# spanning-tree vlan 20 root secondary
```

Example 2-5 also shows that S3 is configured as the secondary root bridge for VLAN 10, and S1 is configured as the secondary root bridge for VLAN 20. This configuration enables spanning tree load balancing, with VLAN 10 traffic passing through S1 and VLAN 20 traffic passing through S3.

Another way to specify the root bridge is to set the spanning tree priority on each switch to the lowest value so that the switch is selected as the primary bridge for its associated VLAN, as shown in Example 2-6.

Example 2-6 Configuring the Lowest Possible Priority to Ensure That the Switch Is Root

```
S3(config)# spanning-tree vlan 20 priority 4096
```

```
S3(config)# spanning-tree vlan 20 priority 4096
```

The switch priority can be set for any spanning tree instance. This setting affects the likelihood that a switch is selected as the root bridge. A lower value increases the probability that the switch is selected. The range is 0 to 61,440 in increments of 4,096; all other values are rejected. For example, a valid priority value is 4,096 x 2 = 8,192.

As shown in Example 2-7, the **show spanning-tree active** command displays spanning tree configuration details for the active interfaces only.

Example 2-7 Verifying STP Active Interfaces

```
S1# show spanning-tree active
<output omitted>
VLAN0010
  Spanning tree enabled protocol ieee
  Root ID     Priority    4106
              Address     ec44.7631.3880
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    4106   (priority 4096 sys-id-ext 10)
              Address     ec44.7631.3880
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  300 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------------------
Fa0/3               Desg FWD 19        128.5    P2p
Fa0/4               Desg FWD 19        128.6    P2p
```

The output shown is for S1 configured with PVST+. There are a number of Cisco IOS command parameters associated with the **show spanning-tree** command.

In Example 2-8, the output shows that the priority for VLAN 10 is 4,096, the lowest of the three respective VLAN priorities.

Example 2-8 Verifying the S1 STP Configuration

```
S1# show running-config | include span
spanning-tree mode pvst
spanning-tree extend system-id
spanning-tree vlan 1 priority 24576
spanning-tree vlan 10 priority 4096
spanning-tree vlan 20 priority 28672
```

Packet Tracer Activity

Packet Tracer 2.3.1.5: Configuring PVST+

In this activity, you will configure VLANs and trunks, and examine and configure the Spanning Tree Protocol primary and secondary root bridges. You will also optimize the switched topology using PVST+, PortFast, and BPDU guard.

Rapid PVST+ Configuration (2.3.2)

Because PVST+ is the default STP mode, Rapid PVST+ must be explicitly configured.

Spanning Tree Mode (2.3.2.1)

Rapid PVST+ is the Cisco implementation of RSTP. It supports RSTP on a per-VLAN basis. The topology in Figure 2-34 has two VLANs: 10 and 20.

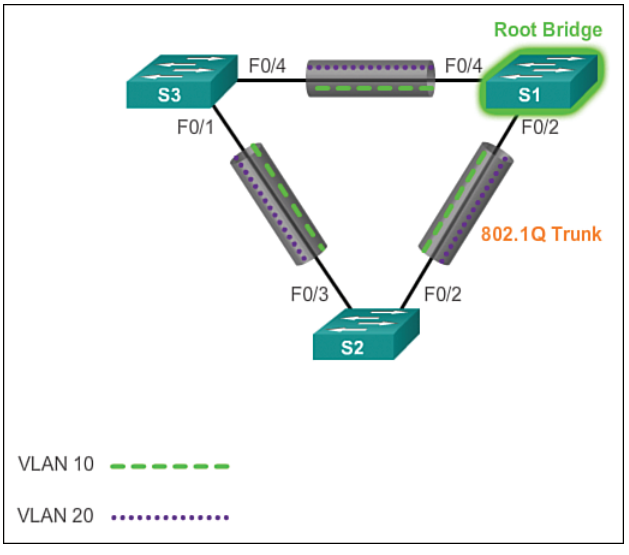


Figure 2-34 Configure Rapid PVST+

Note

The default spanning tree configuration on a Catalyst 2960 Series switch is PVST+. A Catalyst 2960 switch supports PVST+, Rapid PVST+, and MST, but only one version can be active for all VLANs at any time.

Rapid PVST+ commands control the configuration of VLAN spanning tree instances. A spanning tree instance is created when an interface is assigned to a VLAN and is

removed when the last interface is moved to another VLAN. As well, you can configure STP switch and port parameters before a spanning tree instance is created. These parameters are applied when a spanning tree instance is created.

Table 2-6 displays the Cisco IOS command syntax needed to configure Rapid PVST+ on a Cisco switch.

Table 2-6 Rapid PVST+ Configuration Commands

Description	Command Syntax
Enter global configuration mode.	**configure terminal**
Configure Rapid PVST+ spanning-tree mode.	**spanning-tree mode rapid-pvst**
Enter interface configuration mode.	**interface** *interface-id*
Specify that the link type for this port is point-to-point.	**spanning-tree link-type point-to-point**
Return to privileged EXEC mode.	**end**
Clear all detected STP.	**clear spanning-tree detected-protocols**

The **spanning-tree mode rapid-pvst** global configuration mode command is the one required command for the Rapid PVST+ configuration. When specifying an interface to configure, valid interfaces include physical ports, VLANs, and port channels. The VLAN ID range is 1 to 4094 when the enhanced software image (EI) is installed and 1 to 1005 when the standard software image (SI) is installed. The port-channel range is 1 to 6.

Example 2-9 shows Rapid PVST+ commands configured on S1.

Example 2-9 Configuring Rapid PVST+ on S1

```
S1# configure terminal
S1(config)# spanning-tree mode rapid-pvst
S1(config)# interface f0/2
S1(config-if)# spanning-tree link-type point-to-point
S1(config-if)# end
S1# clear spanning-tree detected-protocols
```

In Example 2-10, the **show spanning-tree vlan 10** command shows the spanning tree configuration for VLAN 10 on switch S1.

Example 2-10 Verifying That VLAN 10 Is Using RSTP

```
S1# show spanning-tree vlan 10

VLAN0010
  Spanning tree enabled protocol rstp
  Root ID     Priority    4106
              Address     ec44.7631.3880
              This bridge is the root
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    4106   (priority 4096 sys-id-ext 10)
              Address     ec44.7631.3880
              Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  300 sec

Interface           Role Sts Cost      Prio.Nbr Type
------------------- ---- --- --------- -------- --------------------------------
Fa0/3               Desg FWD 19        128.5    P2p Peer(STP)
Fa0/4               Desg FWD 19        128.6    P2p Peer(STP)
```

Notice that the BID priority is set to 4,096. In the output, the statement "Spanning tree enabled protocol rstp" indicates that S1 is running Rapid PVST+. Because S1 is the root bridge for VLAN 10, all of its interfaces are designated ports.

In Example 2-11, the **show running-config** command is used to verify the Rapid PVST+ configuration on S1.

Example 2-11 Verifying the Rapid PVST+ Configuration

```
S1# show running-config | include span
spanning-tree mode rapid-pvst
spanning-tree extend system-id
spanning-tree vlan 1 priority 24576
spanning-tree vlan 10 priority 4096
spanning-tree vlan 20 priority 28672
spanning-tree link-type point-to-point
```

Note

Generally, it is unnecessary to configure the point-to-point *link-type* parameter for Rapid PVST+, because it is unusual to have a shared *link-type*. In most cases, the only difference between configuring PVST+ and Rapid PVST+ is the **spanning-tree mode rapid-pvst** command.

Packet Tracer Activity

Packet Tracer 2.3.2.2: Configuring Rapid PVST+

In this activity, you will configure VLANs and trunks, and examine and configure the Spanning Tree primary and secondary root bridges. You will also optimize them by using rapid PVST+, PortFast, and BPDU guard.

Lab 2.3.2.3: Configuring Rapid PVST+, PortFast, and BPDU Guard

In this lab, you will complete the following objectives:

- Part 1: Build the Network and Configure Basic Device Settings
- Part 2: Configure VLANs, Native VLAN, and Trunks
- Part 3: Configure the Root Bridge and Examine PVST+ Convergence
- Part 4: Configure Rapid PVST+, PortFast, BPDU Guard, and Examine Convergence

STP Configuration Issues (2.3.3)

If STP configuration is left unchanged, the algorithm might not choose the best root bridge. So it is usually desirable to change the configuration. This topic reviews some of the common issues that can occur when the STP configuration is modified.

Analyzing the STP Topology (2.3.3.1)

To analyze the STP topology, follow these steps as shown in Figure 2-35:

Step 1. Discover the Layer 2 topology. Use network documentation if it exists or use the **show cdp neighbors** command to discover the Layer 2 topology.

Step 2. After discovering the Layer 2 topology, use STP knowledge to determine the expected Layer 2 path. It is necessary to know which switch is the root bridge.

Step 3. Use the **show spanning-tree vlan** command to determine which switch is the root bridge.

Step 4. Use the **show spanning-tree vlan** command on all switches to find out which ports are in the blocking or forwarding state and confirm your expected Layer 2 path.

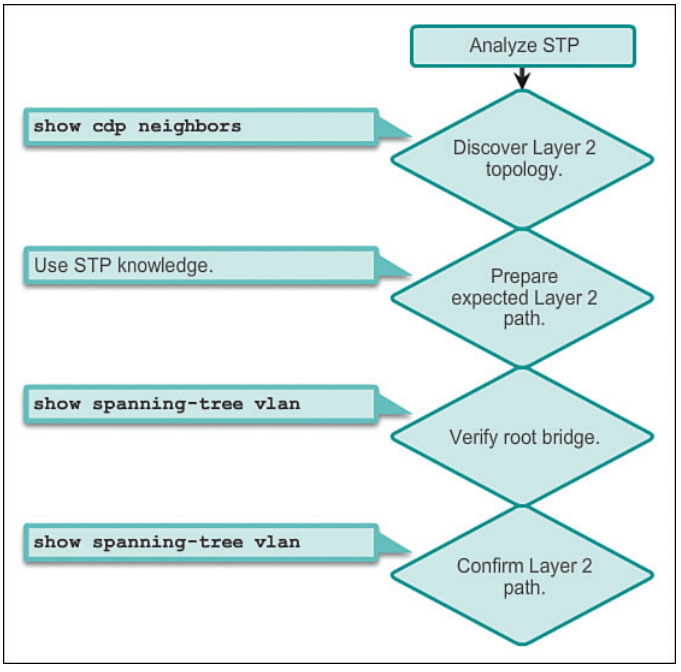


Figure 2-35 Analyzing the STP Topology

Expected Topology Versus Actual Topology (2.3.3.2)

In many networks, the optimal STP topology is determined as part of the network design and then implemented through manipulation of STP priority and cost values. Situations can occur where STP was not considered in the network design and implementation, or where it was considered or implemented before the network underwent significant growth and change. In such situations, it is important to know how to analyze the actual STP topology in the operational network.

A big part of troubleshooting consists of comparing the actual state of the network against the expected state of the network and spotting the differences to gather clues about the troubleshooting problem. A network professional should be able to examine the switches and determine the actual topology, and be able to understand what the underlying spanning tree topology should be.

Overview of Spanning Tree Status (2.3.3.3)

Using the **show spanning-tree** command without specifying any additional options provides a quick overview of the status of STP for all VLANs that are defined on a switch. If interested only in a particular VLAN, limit the scope of this command by specifying that VLAN as an option.

Use the **show spanning-tree vlan** *vlan_id* command to get STP information for a particular VLAN. Use this command to get information about the role and status of each port on the switch. In Figure 2-36, the example output on switch S1 shows all three ports in the forwarding (FWD) state and the role of the three ports as either designated ports or root ports. Any ports being blocked display the output status as "BLK."

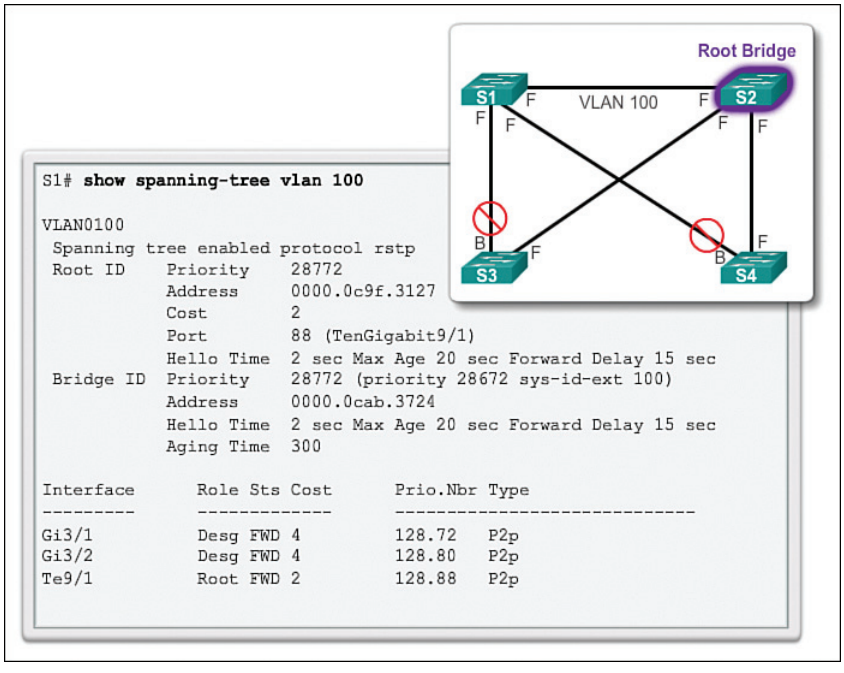


Figure 2-36 show spanning-tree vlan vlan_id Command

The output also gives information about the BID of the local switch and the root ID, which is the BID of the root bridge.

Spanning Tree Failure Consequences (2.3.3.4)

With many protocols, a malfunction means that you lose the functionality that the protocol was providing. For example, if OSPF malfunctions on a router, connectivity to networks that are reachable through that router might be lost. This would generally not affect the rest of the OSPF network. If connectivity to the router is still available, it is possible to troubleshoot to diagnose and fix the problem.

With STP, there are two types of failure. The first is similar to the OSPF problem; STP might erroneously block ports that should have gone into the forwarding state. Connectivity might be lost for traffic that would normally pass through this switch, but the rest of the network remains unaffected. The second type of failure is much

more disruptive. It happens when STP erroneously moves one or more ports into the forwarding state.

Remember that an Ethernet frame header does not include a TTL field, which means that any frame that enters a bridging loop continues to be forwarded by the switches indefinitely. The only exceptions are frames that have their destination address recorded in the MAC address table of the switches. These frames are simply forwarded to the port that is associated with the MAC address and do not enter a loop. However, any frame that is flooded by a switch enters the loop. This can include broadcasts, multicasts, and unicasts with a globally unknown destination MAC address.

Figure 2-37 graphically displays the consequences and corresponding symptoms of STP failure.

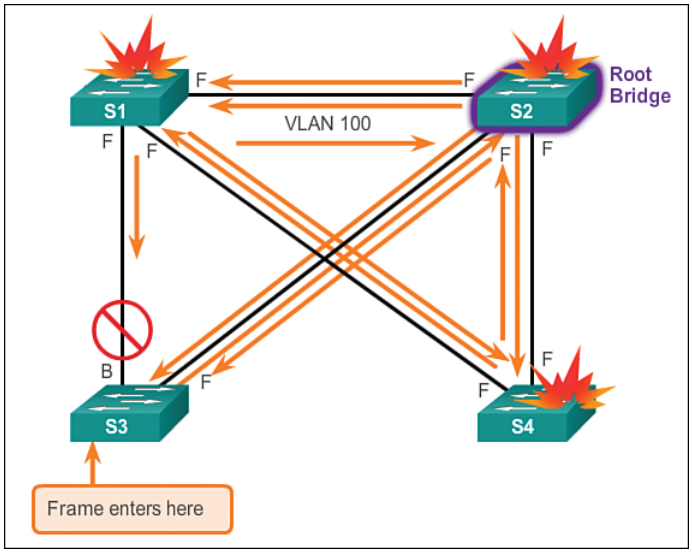


Figure 2-37 STP Failure

The load on all links in the switched LAN quickly starts increasing as more and more frames enter the loop. This problem is not limited to the links that form the loop, but also affects any other links in the switched domain because the frames are flooded on all links. When the spanning tree failure is limited to a single VLAN, only links in that VLAN are affected. Switches and trunks that do not carry that VLAN operate normally.

If the spanning tree failure has created a bridging loop, traffic increases exponentially. The switches will then flood the broadcasts out multiple ports. This creates copies of the frames every time the switches forward them.

When control plane traffic starts entering the loop (for example, OSPF Hellos or EIGRP Hellos), the devices that are running these protocols quickly start getting overloaded. Their CPUs approach 100 percent utilization while they are trying to process an ever-increasing load of control plane traffic. In many cases, the earliest indication of this broadcast storm in progress is that routers or Layer 3 switches are reporting control plane failures and that they are running at a high CPU load.

The switches experience frequent MAC address table changes. If a loop exists, a switch might see a frame with a certain source MAC address coming in on one port and then see another frame with the same source MAC address coming in on a different port a fraction of a second later. This will cause the switch to update the MAC address table twice for the same MAC address.

Because of the combination of very high load on all links and the switch CPUs running at maximum load, these devices typically become unreachable. This makes it very difficult to diagnose the problem while it is happening.

Repairing a Spanning Tree Problem (2.3.3.5)

One way to correct spanning tree failure is to manually remove redundant links in the switched network, either physically or through configuration, until all loops are eliminated from the topology. When the loops are broken, the traffic and CPU loads should quickly drop to normal levels, and connectivity to devices should be restored.

Although this intervention restores connectivity to the network, it is not the end of the troubleshooting process. All redundancy from the switched network has been removed, and now the redundant links must be restored.

If the underlying cause of the spanning tree failure has not been fixed, chances are that restoring the redundant links will trigger a new broadcast storm. Before restoring the redundant links, determine and correct the cause of the spanning tree failure. Carefully monitor the network to ensure that the problem is fixed.

Interactive Graphic

Activity 2.3.3.6: Troubleshoot STP Configuration Issues

Go to the course online to perform this practice activity.

First Hop Redundancy Protocols (2.4)

The term First Hop Redundancy Protocol (FHRP) refers to a collection of protocols that transparently provide end users with at least one redundant default gateway.

Concept of First Hop Redundancy Protocols (2.4.1)

With redundant routers and redundant links, it is possible to configure a redundant default gateway.

Default Gateway Limitations (2.4.1.1)

Spanning tree protocols enable physical redundancy in a switched network. However, a host at the access layer of a hierarchical network also benefits from alternate default gateways. If a router or router interface (that serves as a default gateway) fails, the hosts configured with that default gateway are isolated from outside networks. A mechanism is needed to provide alternate default gateways in switched networks where two or more routers are connected to the same VLANs.

Note

For the purposes of the discussion on router redundancy, there is no functional difference between a multilayer switch and a router at the distribution layer. In practice, it is common for a multilayer switch to act as the default gateway for each VLAN in a switched network. This discussion focuses on the functionality of *routing*, regardless of the physical device used.

In a switched network, each client receives only one default gateway. There is no way to configure a secondary gateway, even if a second path exists to carry packets off the local segment.

In Figure 2-38, R1 is responsible for routing packets from PC1.

If R1 becomes unavailable, the routing protocols can dynamically converge. R2 now routes packets from outside networks that would have gone through R1. However, traffic from the inside network associated with R1, including traffic from workstations, servers, and printers configured with R1 as their default gateway, is still sent to R1 and dropped.

End devices are typically configured with a single IP address for a default gateway. This address does not change when the network topology changes. If that default gateway IP address cannot be reached, the local device is unable to send packets off the local network segment, effectively disconnecting it from the rest of the network. Even if a redundant router exists that could serve as a default gateway for that segment, there is no dynamic method by which these devices can determine the address of a new default gateway.

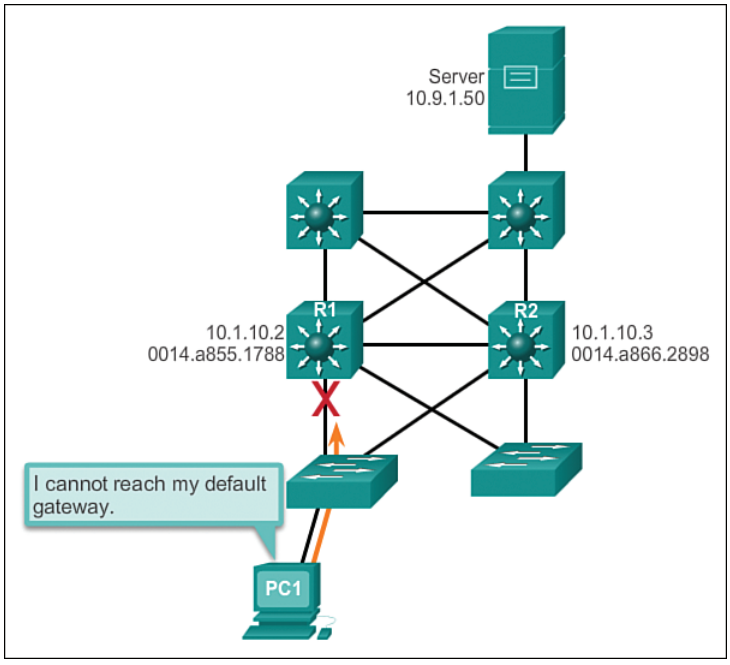


Figure 2-38 Default Gateway Limitations

Router Redundancy (2.4.1.2)

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 to work together to present the illusion of a single router to the hosts on the LAN, as shown in Figure 2-39. By sharing an IP address and a MAC address, two or more routers can act as a single virtual router.

The IP address of the virtual router is configured as the default gateway for the workstations on a specific IP segment. When frames are sent from host devices to the default gateway, the hosts use ARP to resolve the MAC address that is associated with the IP address of the default gateway. The ARP resolution returns the MAC address of the virtual router. Frames that are sent to the MAC address of the virtual router can then be physically processed by the currently active router within the virtual router group. A protocol is used to identify two or more routers as the devices that are responsible for processing frames that are sent to the MAC or IP address of a single virtual router. Host devices send traffic to the address of the virtual router. The physical router that forwards this traffic is transparent to the host devices.

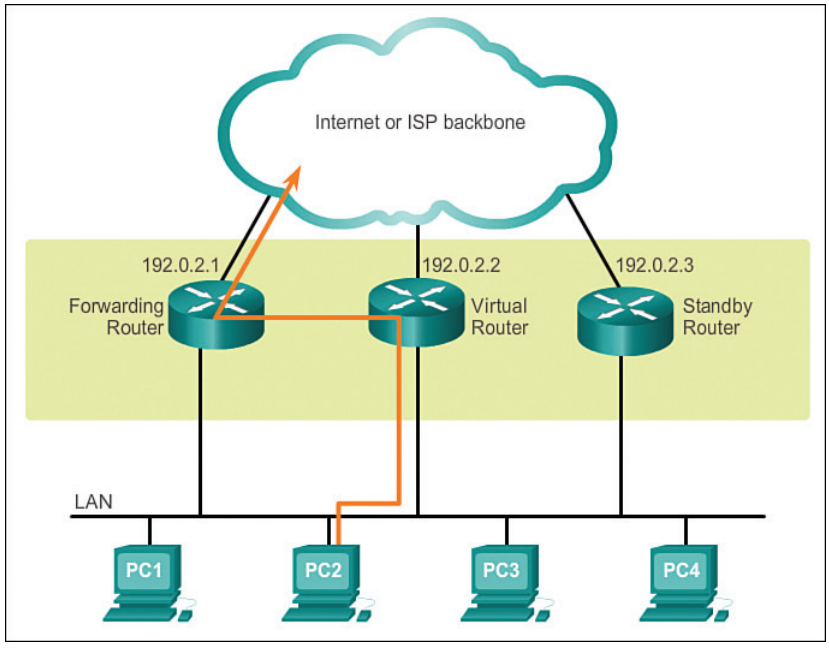


Figure 2-39 Router Redundancy

A redundancy protocol provides the mechanism for determining which router should take the active role in forwarding traffic. It also determines when the forwarding role must be taken over by a standby router. The transition from one forwarding router to another is transparent to the end devices.

The ability of a network to dynamically recover from the failure of a device acting as a default gateway is known as first hop redundancy.

Steps for Router Failover (2.4.1.3)

When the active router fails, the redundancy protocol transitions the standby router to the new active router role, as shown in Figure 2-40.

These are the steps that take place when the active router fails:

1. The standby router stops seeing Hello messages from the forwarding router.
2. The standby router assumes the role of the forwarding router.
3. Because the new forwarding router assumes both the IP and MAC addresses of the virtual router, the host devices see no disruption in service.

Interactive Graphic

Activity 2.4.1.4: Identify FHRP Terminology

Go to the course online to perform this practice activity.

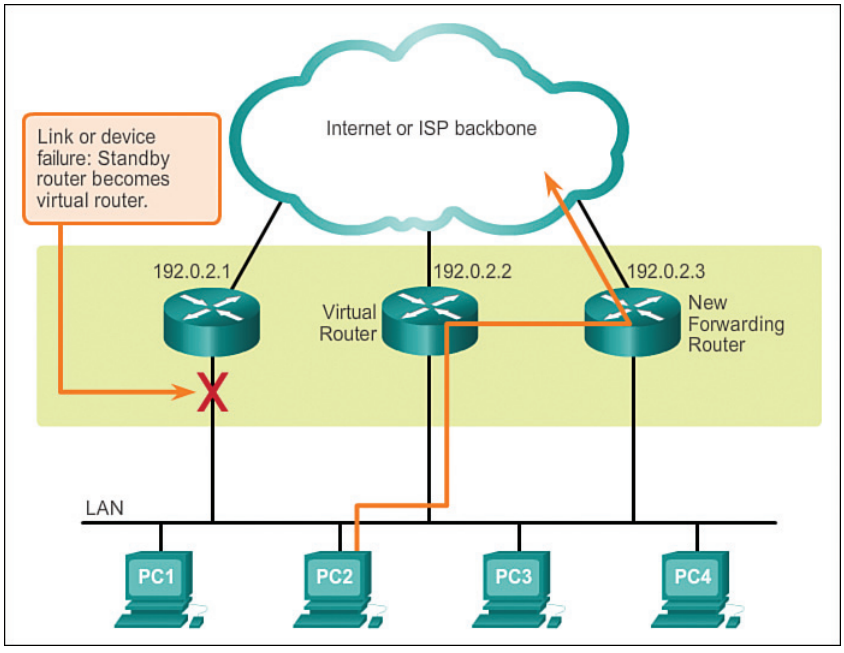


Figure 2-40 Router Failover Example

Varieties of First Hop Redundancy Protocols (2.4.2)

There are several options to choose from when configuring an FHRP.

First Hop Redundancy Protocols (2.4.2.1)

The following list defines the options available for First Hop Redundancy Protocols (FHRP).

- *Hot Standby Router Protocol (HSRP):* A Cisco-proprietary FHRP designed to allow for transparent failover of a first hop IPv4 device. HSRP provides high network availability by providing first-hop routing redundancy for IPv4 hosts on networks configured with an IPv4 default gateway address. HSRP is used in a group of routers for selecting an active device and a standby device. In a group of device interfaces, the active device is the device that is used for routing packets; the standby device is the device that takes over when the active device fails, or when preset conditions are met. The function of the HSRP standby router is to monitor the operational status of the HSRP group and to quickly assume packet-forwarding responsibility if the active router fails.
- **HSRP for IPv6:** A Cisco-proprietary FHRP providing the same functionality of HSRP, but in an IPv6 environment. An HSRP IPv6 group has a virtual MAC

address derived from the HSRP group number and a virtual IPv6 link-local address derived from the HSRP virtual MAC address. Periodic router advertisements (RA) are sent for the HSRP virtual IPv6 link-local address when the HSRP group is active. When the group becomes inactive, these RAs stop after a final RA is sent.

- *Virtual Router Redundancy Protocol* **version 2 (VRRPv2):** A nonproprietary election protocol that dynamically assigns responsibility for one or more virtual routers to the VRRP routers on an IPv4 LAN. This allows several routers on a multiaccess link to use the same virtual IPv4 address. A VRRP router is configured to run the VRRP protocol in conjunction with one or more other routers attached to a LAN. In a VRRP configuration, one router is elected as the virtual router master, with the other routers acting as backups, in case the virtual router master fails.
- **VRRPv3:** Provides the capability to support IPv4 and IPv6 addresses. VRRPv3 works in multivendor environments and is more scalable than VRRPv2.
- *Gateway Load Balancing Protocol* **(GLBP):** A Cisco-proprietary FHRP that protects data traffic from a failed router or circuit, like HSRP and VRRP, while also allowing load balancing (also called load sharing) between a group of redundant routers.
- **GLBP for IPv6:** A Cisco-proprietary FHRP providing the same functionality of GLBP, but in an IPv6 environment. GLBP for IPv6 provides automatic router backup for IPv6 hosts configured with a single default gateway on a LAN. Multiple first-hop routers on the LAN combine to offer a single virtual first-hop IPv6 router while sharing the IPv6 packet-forwarding load.
- *ICMP Router Discovery Protocol* **(IRDP):** Specified in RFC 1256, this is a legacy FHRP solution. IRDP allows IPv4 hosts to locate routers that provide IPv4 connectivity to other (nonlocal) IP networks.

Interactive Graphic

Activity 2.4.2.2: Identify the Type of FHRP

Go to the course online to perform this practice activity.

FHRP Verification (2.4.3)

This topic briefly reviews the tasks necessary to configure and verify HSRP and GLBP.

HSRP Verification (2.4.3.1)

Figure 2-41 shows an example topology for configuring either HSRP or GLBP.

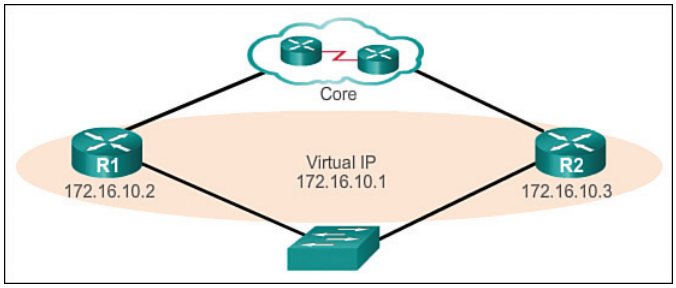


Figure 2-41 FHRP Configuration Topology

An HSRP active router has the following characteristics:

- Responds to default gateway's ARP requests with the virtual router's MAC.
- Assumes active forwarding of packets for the virtual router.
- Sends Hello messages.
- Knows the virtual router IP address.

An HSRP standby router has the following characteristics:

- Listens for periodic Hello messages.
- Assumes active forwarding of packets if it does not hear from the active router.

Use the **show standby** command to verify the HSRP state. In Example 2-12, the output shows that R1 is in the active state.

Example 2-12 Verifying That R1 Is the HSRP Active Router

```
R1# show standby
FastEthernet0/1 - Group 10
  State is Active
    2 state changes, last state change 00:04:01
  Virtual IP address is 172.16.10.1
  Active virtual MAC address is 0000.0c07.ac0a
    Local virtual MAC address is 0000.0c07.ac0a (v1 default)
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 0.528 secs
  Preemption disabled
  Active router is local
  Standby router is 172.16.10.3, priority 110 (expires in 10.576 sec)
  Priority 150 (configured 150)
  Group name is "hsrp-Fa0/1-10" (default)
```

GLBP Verification (2.4.3.2)

Although HSRP and VRRP provide gateway resiliency, for the standby members of the redundancy group, the upstream bandwidth is not used while the device is in standby mode.

Only the active router in HSRP and VRRP groups forwards traffic for the virtual MAC address. Resources that are associated with the standby router are not fully utilized. You can accomplish some load balancing with these protocols by creating multiple groups and assigning multiple default gateways, but this configuration creates an administrative burden.

GLBP is a Cisco-proprietary solution to allow automatic selection and simultaneous use of multiple available gateways in addition to automatic failover between those gateways. Multiple routers share the load of frames that, from a client perspective, are sent to a single default gateway address, as shown in Figure 2-42.

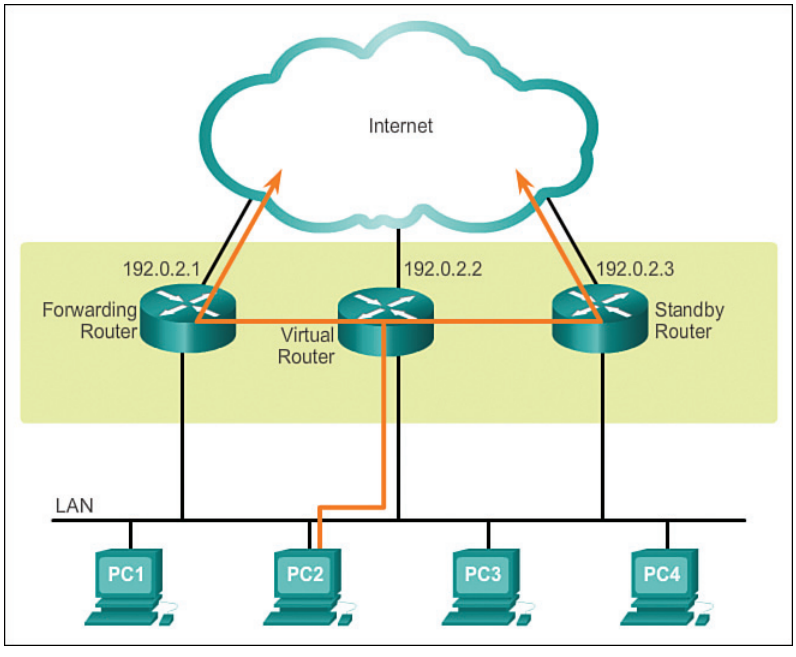


Figure 2-42 Gateway Load-Balancing Protocol

With GLBP, you can fully utilize resources without the administrative burden of configuring multiple groups and managing multiple default gateway configurations. GLBP has the following characteristics:

- Allows full use of resources on all devices without the administrative burden of creating multiple groups.
- Provides a single virtual IP address and multiple virtual MAC addresses.

- Routes traffic to single gateway distributed across routers.
- Provides automatic rerouting in the event of any failure.

Use the **show glbp** command to verify the GLBP status. Example 2-13 for R1 shows that GLBP group 10 is in the active state with virtual IP address 172.16.10.1. R1 is the active router for Forwarder 2.

Example 2-13 Verifying R1 GLBP Forwarding Roles

```
R1# show glbp
FastEthernet0/1 - Group 10
  State is Active
    2 state changes, last state change 00:02:50
  Virtual IP address is 172.16.10.1
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 1.408 secs
  Redirect time 600 sec, forwarder timeout 14400 sec
  Preemption disabled
  Active is local
  Standby is 172.16.10.3, priority 110 (expires in 7.776 sec)
  Priority 150 (configured)
  Weighting 100 (default 100), thresholds: lower 1, upper 100
  Load balancing: round-robin
  Group members:
    0016.c8ee.131a (172.16.10.3)
    001b.d4ef.5091 (172.16.10.2) local
  There are 2 forwarders (1 active)
  Forwarder 1
    State is Listen
      2 state changes, last state change 00:00:09
    MAC address is 0007.b400.0a01 (learnt)
    Owner ID is 0016.c8ee.131a
    Redirection enabled, 597.792 sec remaining (maximum 600 sec)
    Time to live: 14397.792 sec (maximum 14400 sec)
    Preemption enabled, min delay 30 sec
    Active is 172.16.10.3 (primary), weighting 100 (expires in 9.920 sec)
  Forwarder 2
    State is Active
      1 state change, last state change 00:05:57
    MAC address is 0007.b400.0a02 (default)
    Owner ID is 001b.d4ef.5091
    Redirection enabled
    Preemption enabled, min delay 30 sec
    Active is local, weighting 100
```

Syntax Checker — HSRP and GLBP (2.4.3.3)

Configuration of HSRP and GLBP is beyond the scope of this course. However, familiarity with the commands used to enable HSRP and GLBP aid in understanding the configuration output. For this reason, the syntax checker and subsequent lab are available as optional exercises.

Lab 2.4.3.4: Configuring HSRP and GLBP

In this lab, you will complete the following objectives:

- Part 1: Build the Network and Verify Connectivity
- Part 2: Configure First Hop Redundancy Using HSRP
- Part 3: Configure First Hop Redundancy Using GLBP

Summary (2.5)

Class Activity 2.5.1.1: Documentation Tree

The employees in your building are having difficulty accessing a web server on the network. You look for the network documentation that the previous network engineer used before he transitioned to a new job; however, you cannot find any network documentation whatsoever.

Therefore, you decide to create your own network record-keeping system. You decide to start at the access layer of your network hierarchy. This is where redundant switches are located, as well as the company servers, printers, and local hosts.

You create a matrix to record your documentation and include access layer switches on the list. You also decide to document switch names, ports in use, cabling connections, root ports, designated ports, and alternate ports.

Problems that can result from a redundant Layer 2 network include broadcast storms, MAC database instability, and duplicate unicast frames. STP is a Layer 2 protocol that ensures that there is only one logical path between all destinations on the network by intentionally blocking redundant paths that could cause a loop.

STP sends BPDU frames for communication between switches. One switch is elected as the root bridge for each instance of spanning tree. An administrator can control this election by changing the bridge priority. Root bridges can be configured to enable spanning tree load balancing by VLAN or by a group of VLANs, depending on the spanning tree protocol used. STP then assigns a port role to each participating port using a path cost. The path cost is equal to the sum of all the port costs along the path to the root bridge. A port cost is automatically assigned to each port; however, it can also be manually configured. Paths with the lowest cost become preferred, and all other redundant paths are blocked.

PVST+ is the default configuration of IEEE 802.1D on Cisco switches. It runs one instance of STP for each VLAN. A newer, faster-converging spanning tree protocol, RSTP, can be implemented on Cisco switches on a per-VLAN basis in the form of Rapid PVST+. Multiple Spanning Tree (MST) is the Cisco implementation of Multiple Spanning Tree Protocol (MSTP), where one instance of spanning tree runs for a defined group of VLANs. Features such as PortFast and BPDU guard ensure that hosts in the switched environment are provided immediate access to the network without interfering with spanning tree operation.

First Hop Redundancy Protocols, such as HSRP, VRRP, and GLBP, provide alternate default gateways for hosts in the redundant router or multilayer switched environment. Multiple routers share a virtual IP address and MAC address that is used as

the default gateway on a client. This ensures that hosts maintain connectivity in the event of the failure of one device serving as a default gateway for a VLAN or set of VLANs. When using HSRP or VRRP, one router is active or forwarding for a particular group while others are in standby mode. GLBP allows the simultaneous use of multiple gateways in addition to providing automatic failover.

Practice

The following activities provide practice with the topics introduced in this chapter. The Labs and Class Activities are available in the companion *Scaling Networks Lab Manual* (ISBN 978-1-58713-325-1). The Packet Tracer Activities PKA files are found in the online course.

Class Activities

- Class Activity 2.0.1.2: Stormy Traffic
- Class Activity 2.5.1.1: Documentation Tree

Labs

- Lab 2.1.2.10: Building a Switched Network with Redundant Links
- Lab 2.3.2.3: Configuring Rapid PVST+, PortFast, and BPDU Guard
- Lab 2.4.3.4: Configuring HSRP and GLBP

Packet Tracer Activities

- Packet Tracer Activity 2.1.1.5: Examining a Redundant Design
- Packet Tracer Activity 2.3.1.5: Configuring PVST+
- Packet Tracer Activity 2.3.2.2: Configuring Rapid PVST+

Check Your Understanding Questions

Complete all the review questions listed here to test your understanding of the topics and concepts in this chapter. The appendix "Answers to 'Check Your Understanding' Questions" lists the answers.

1. What is an accurate description of redundancy?

   A. Configuring a router with a complete MAC address database to ensure that all frames can be forwarded to the correct destination

   B. Designing a network to use multiple paths between switches to ensure that there is no single point of failure

   C. Designing a network to use multiple virtual devices to ensure that all traffic uses the best path through the internetwork

   D. Configuring a switch with proper security to ensure that all traffic forwarded through an interface is filtered

2. Which of the following issues are the result of a broadcast storm? (Choose two.)

   A. During a broadcast storm, constant changes to the MAC address table prevent a switch from accurately forwarding frames.

   B. In a network saturated with broadcast traffic, new traffic arriving at the switch will be forwarded into the broadcast domain, which further consumes available bandwidth.

   C. During a broadcast storm, switches with high-speed interfaces will forward traffic in half-duplex mode to conserve available bandwidth.

   D. Because of high processing demands during a broadcast storm, communication can fail between end stations in the broadcast domain.

   E. During a broadcast storm, a switch will forward a received broadcast out every port on the switch.

3. During the implementation of Spanning Tree Protocol, all switches are rebooted by the network administrator. What is the first step of the spanning-tree election process?

   A. Each switch determines the best path to forward traffic.

   B. Each switch determines what port to block to prevent a loop from occurring.

   C. Each switch with a lower root ID than its neighbor will not send BPDUs.

   D. All the switches send out BPDUs advertising themselves as the root bridge.

4. After the election of the root bridge has been completed, how will switches find the best paths to the root bridge?

   A. Each switch will analyze the sum of all port costs to reach the root and use the path with the lowest cost.
   B. Each switch will analyze the port states of all neighbors and use the designated ports to forward traffic to the root.
   C. Each switch will analyze the sum of the hops to reach the root and use the path with the fewest hops.
   D. Each switch will analyze the BID of all neighbors to reach the root and use the path through the lowest BID neighbors.

5. When PVST is running over a switched network, which port state can participate in BPDU frame forwarding based on BPDUs received, but does not forward data frames?

   A. Disabled
   B. Blocking
   C. Listening
   D. Forwarding

6. What are expectations of configuring PortFast on a switch port? (Choose two.)

   A. The switch port immediately transitions from the listening to the forwarding state.
   B. The switch port immediately processes any BPDUs before transitioning to the forwarding state.
   C. The switch port sends DHCP requests before transitioning to the forwarding state.
   D. The switch port should never receive BPDUs from end stations that are connected to the port.
   E. The switch port immediately transitions from the blocking to the forwarding state.

7. Which of the following port states are used by Rapid PVST+? (Choose three.)

   A. Learning
   B. Blocking
   C. Trunking
   D. Discarding
   E. Forwarding
   F. Listening

8. An administrator is troubleshooting a switch and wants to verify whether it is a root bridge. What command can be used to do this?

   A. **show vlan**
   B. **show spanning-tree**
   C. **show running-config**
   D. **show startup-config**

9. What is the initial approach that should be used to troubleshoot a broadcast storm in a switched network?

   A. Replace all instances of STP with RSTP.
   B. Insert redundant links to replace the failed STP links.
   C. Manually remove redundant links in the switched network.
   D. Replace the cables on failed STP links.

10. When first hop redundancy protocols are used, which of the following items will be shared by a set of routers that are presenting the illusion of being a single router? (Choose two.)

    A. Host name
    B. BID
    C. MAC address
    D. IP address
    E. Static route

11. A network administrator is overseeing the implementation of first hop redundancy protocols. Which of the following protocols will not be able to function with multivendor devices? (Choose two.)

    A. VRRP
    B. HSRP
    C. IRDP
    D. GLBP

12. Indicate the STP protocol the matches the description.

    _____ is a legacy standard that runs all VLANs in a single spanning tree instance.

    _____ is a Cisco enhancement of RSTP that provides a spanning tree instance for each VLAN.

    _____ allows multiple VLANs to run in a single spanning tree instance.

13. List the three steps that an FHRP initiates during a router failover process.