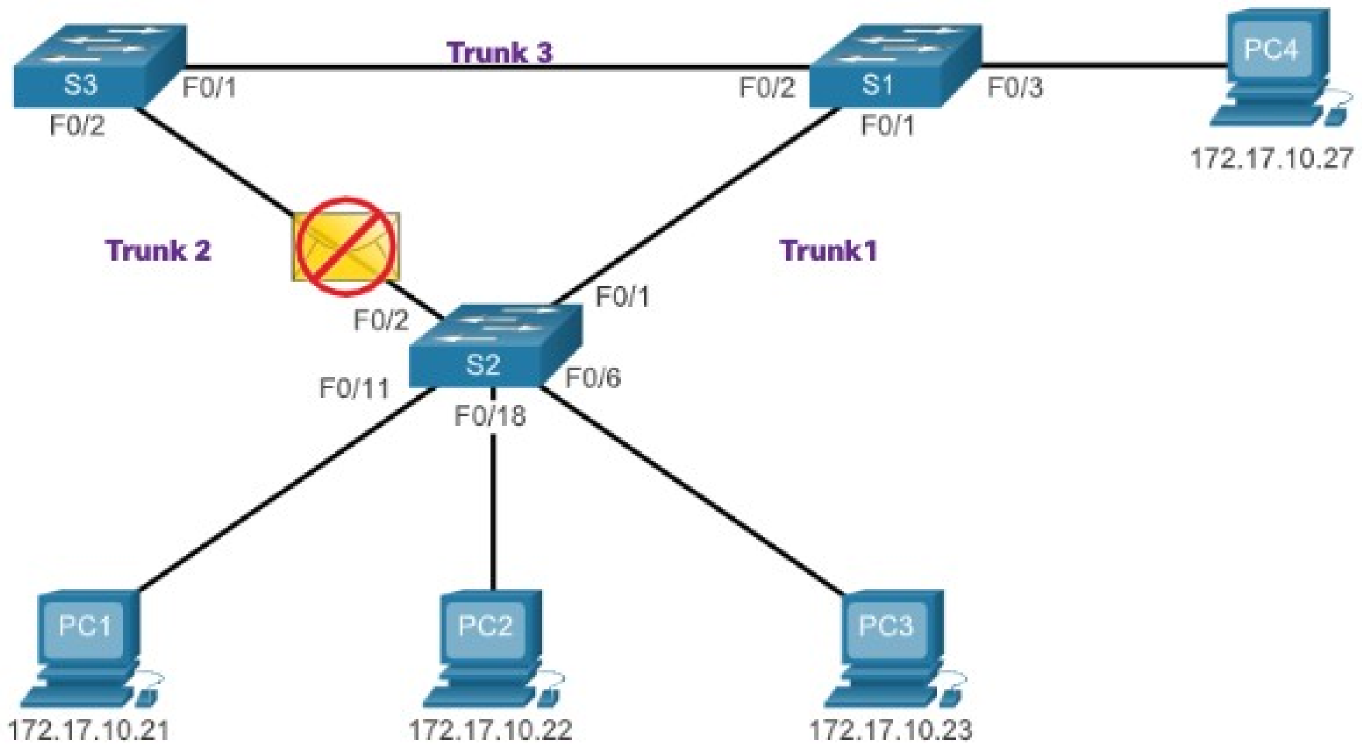
Réseaux fixes
II.1 REDONDANCE
Spanning-Tree Protocol
EtherChannels

Luc Deneire

EII-5, Option Réseaux et Objets Connectés (ROC)

Purpose of STP
Spanning Tree Protocol

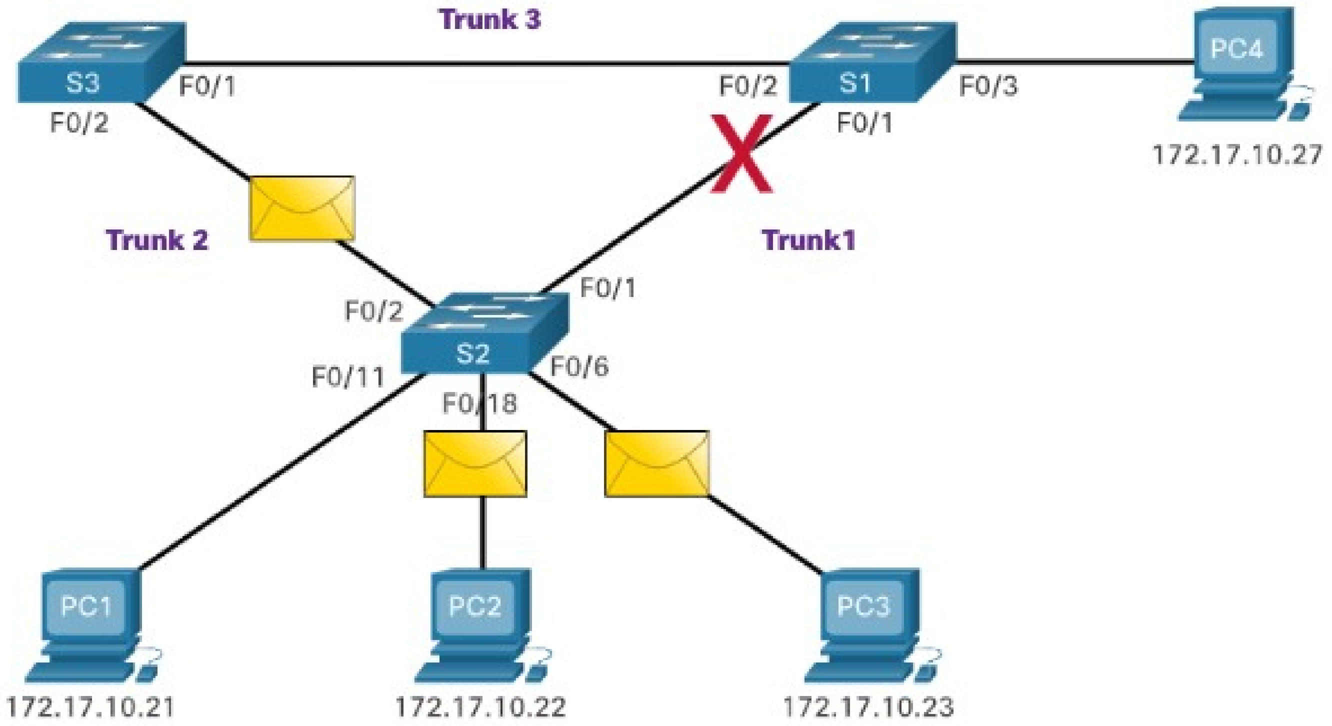
- Spanning Tree Protocol (STP) is a loop-prevention network protocol that allows for redundancy while creating a loop-free Layer 2 topology.
- STP logically blocks physical loops in a Layer 2 network, preventing frames from circling the network forever.



S2 drops the frame because it received it on a blocked port.

STP Recalculation

STP compensates for a failure in the network by recalculating and opening up previously blocked ports.



Issues with Redundant Switch Links

- Path redundancy : eliminating the possibility of a single point of failure.
- Layer 2 loop can occurs.
- A Layer 2 loop can result in the network becoming unusable due to
  - MAC address table instability,
  - link saturation, and high
  - CPU utilization on switches and end-devices,.
- Layer 2 Ethernet : no loop prevention.
- Both IPv4 and IPv6 include a mechanism that limits the number of times a Layer 3 networking device can retransmit a packet. A router will decrement the TTL (Time to Live) in every IPv4 packet, and the Hop Limit field in every IPv6 packet. When these fields are decremented to 0, a router will drop the packet.
- STP was developed specifically as a loop prevention mechanism for Layer 2 Ethernet.

Layer 2 Loops

- Without STP enabled, Layer 2 loops can form, causing *broadcast*, *multicast* and *unknown unicast* frames to loop endlessly. This can bring down a network quickly.
- When a loop occurs, the *MAC address table* on a switch will constantly *change* with the updates from the broadcast frames, which results in MAC database instability. This can cause high CPU utilization, which makes the switch unable to forward frames.
- An unknown unicast frame is when the switch does not have the destination MAC address in its MAC address table and must forward the frame out all ports, except the ingress port.

Broadcast Storm

- A broadcast storm .
- Broadcast storms can disable a network within seconds by overwhelming switches and end devices. can be caused by
  - a hardware problem such as a faulty NIC
  - from a Layer 2 loop in the network.
- Layer 2 broadcasts in a network, such as ARP Requests are very common. Layer 2 multicasts are typically forwarded the same way as a broadcast by the switch. IPv6 packets are never forwarded as a Layer 2 broadcast, ICMPv6 Neighbor Discovery uses Layer 2 multicasts.
- A host caught in a Layer 2 loop is not accessible to other hosts on the network. Additionally, due to the constant changes in its MAC address table, the switch does not know out of which port to forward unicast frames.
- 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.

The Spanning Tree Algorithm

- 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*." Her spanning tree algorithm (STA) creates a loop-free topology by selecting a single root bridge where all other switches determine a single least-cost path.
- 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.

The Spanning Tree Algorithm (Cont.)

How does the STA create a loop-free topology?

- Selecting a Root Bridge: This bridge (switch) is the reference point for the entire network to build a spanning tree around.
- Block Redundant Paths: 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. When a port is blocked, user data is prevented from entering or leaving that port.
- Create a Loop-Free Topology: A blocked port has the effect of making that link a non-forwarding link between the two switches. This creates a topology where each switch has only a single path to the root bridge, similar to branches on a tree that connect to the root of the tree.
- Recalculate in case of Link Failure: 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. STP recalculations can also occur any time a new switch or new inter-switch link is added to the network.

Steps to a Loop-Free Topology

Using the STA, STP builds a loop-free topology in a four-step process:

1. Elect the root bridge.
2. Elect the root ports.
3. Elect designated ports.
4. Elect alternate (blocked) ports.

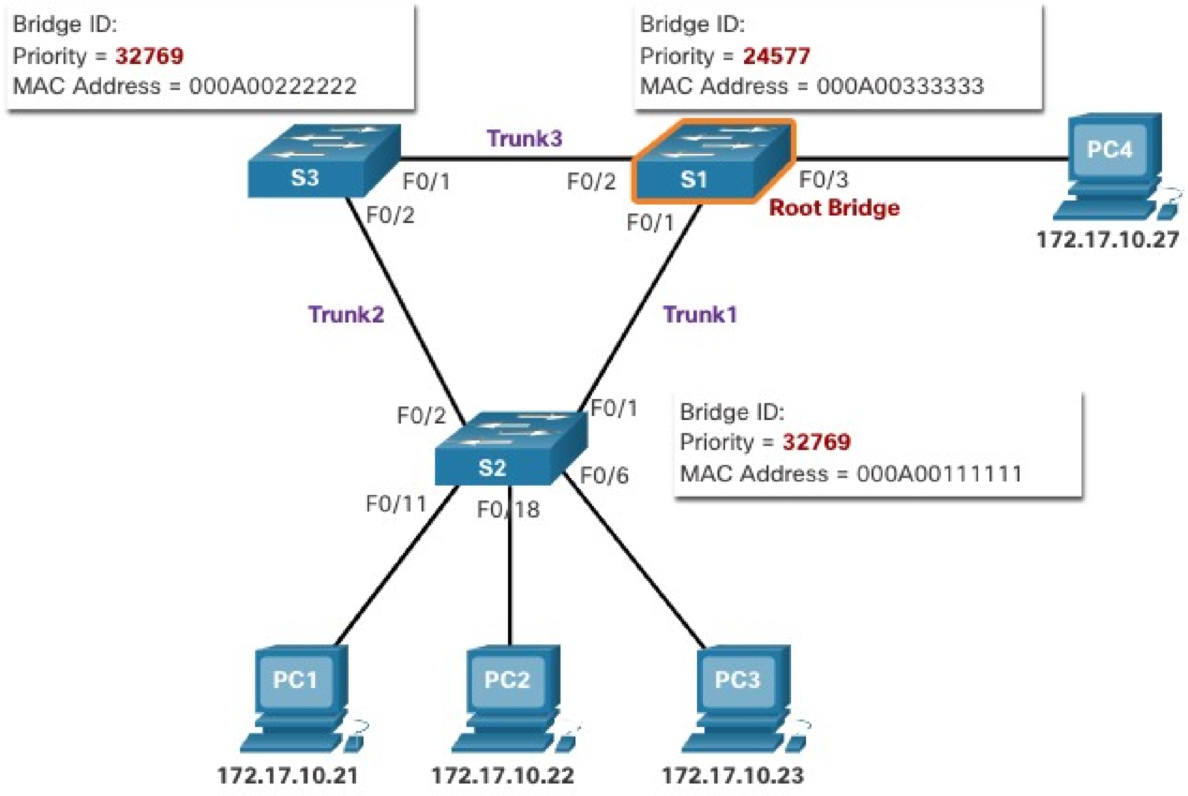
- During STA and STP functions, switches use Bridge Protocol Data Units (BPDUs) to share information about themselves and their connections. BPDUs are used to elect the root bridge, root ports, designated ports, and alternate ports.
- Each BPDU contains a bridge ID (BID) that identifies which switch sent the BPDU. The BID is involved in making many of the STA decisions including root bridge and port roles.
- The BID contains a priority value, the MAC address of the switch, and an extended system ID. The lowest BID value is determined by the combination of these three fields.

Steps to a Loop-Free Topology (Cont.)

- **Bridge Priority:** The default priority value for all Cisco switches is the decimal value 32768. The range is 0 to 61440 in increments of 4096. A lower bridge priority is preferable. A bridge priority of 0 takes precedence over all other bridge priorities.
- **Extended System ID:** The extended system ID value is a decimal value added to the bridge priority value in the BID to identify the VLAN for this BPDU.
- **MAC address:** When two switches are configured with the same priority and have the same extended system ID, the switch having the MAC address with the lowest value, expressed in hexadecimal, will have the lower BID.

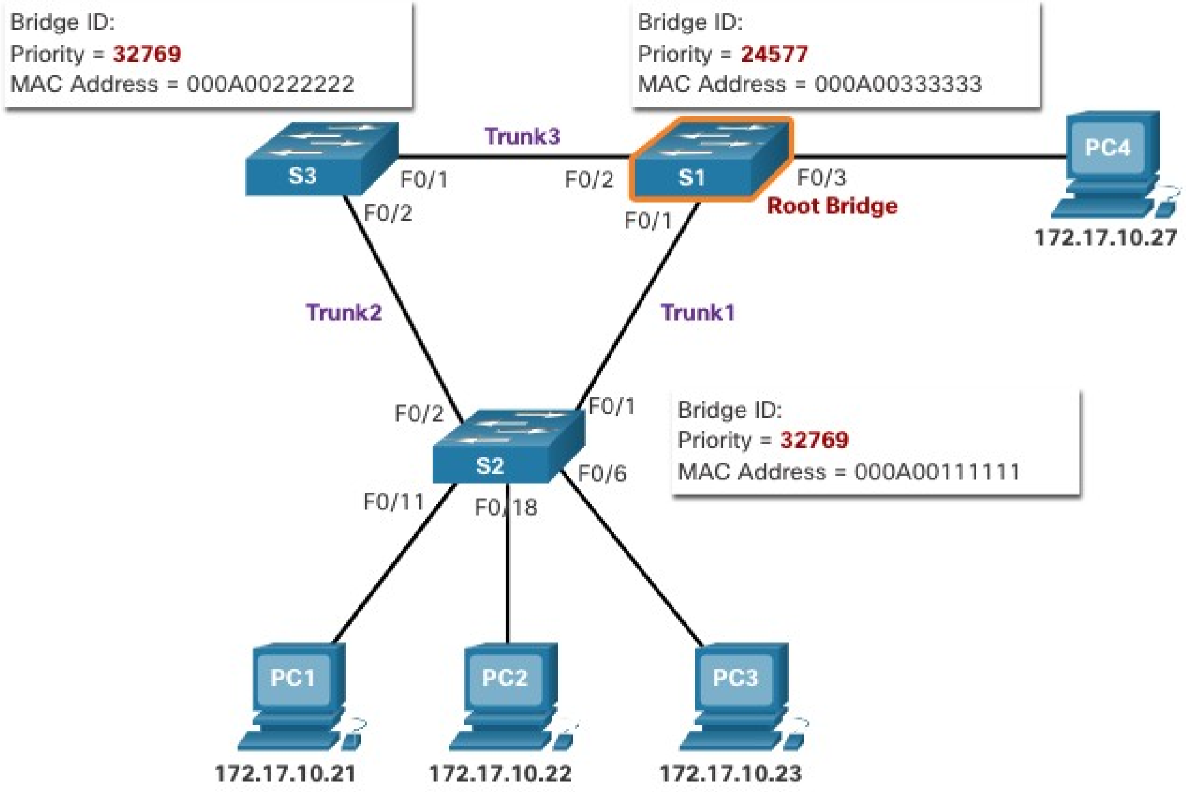
1. Elect the Root Bridge

- The STA designates a single switch as the root bridge and uses it as the reference point for all path calculations. Switches exchange BPDUs to build the loop-free topology beginning with selecting the root bridge.
- 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 BPDU frames contain the BID of the sending switch and the BID of the root bridge, known as the Root ID.
- The switch with the lowest BID will become the root bridge. At first, all switches declare themselves as the root bridge with their own BID set as the Root ID. Eventually, the switches learn through the exchange of BPDUs which switch has the lowest BID and will agree on one root bridge.



Impact of Default BIDs

- Because the default BID is 32768, it is possible for two or more switches to have the same priority. In this scenario, where the priorities are the same, the switch with the lowest MAC address will become the root bridge. The administrator should configure the desired root bridge switch with a lower priority.
- In the figure, IF all switches were configured with the same priority of 32769. Here the MAC address becomes the deciding factor as to which switch becomes the root bridge. The switch with the lowest hexadecimal MAC address value is the preferred root bridge. In this example, S2 has the lowest value for its MAC address and is elected as the root bridge for that spanning tree instance.
- **Note**: The priority of all the switches is 32769. The value is based on the 32768 default bridge priority and the extended system ID (VLAN 1 assignment) associated with each switch (32768+1).



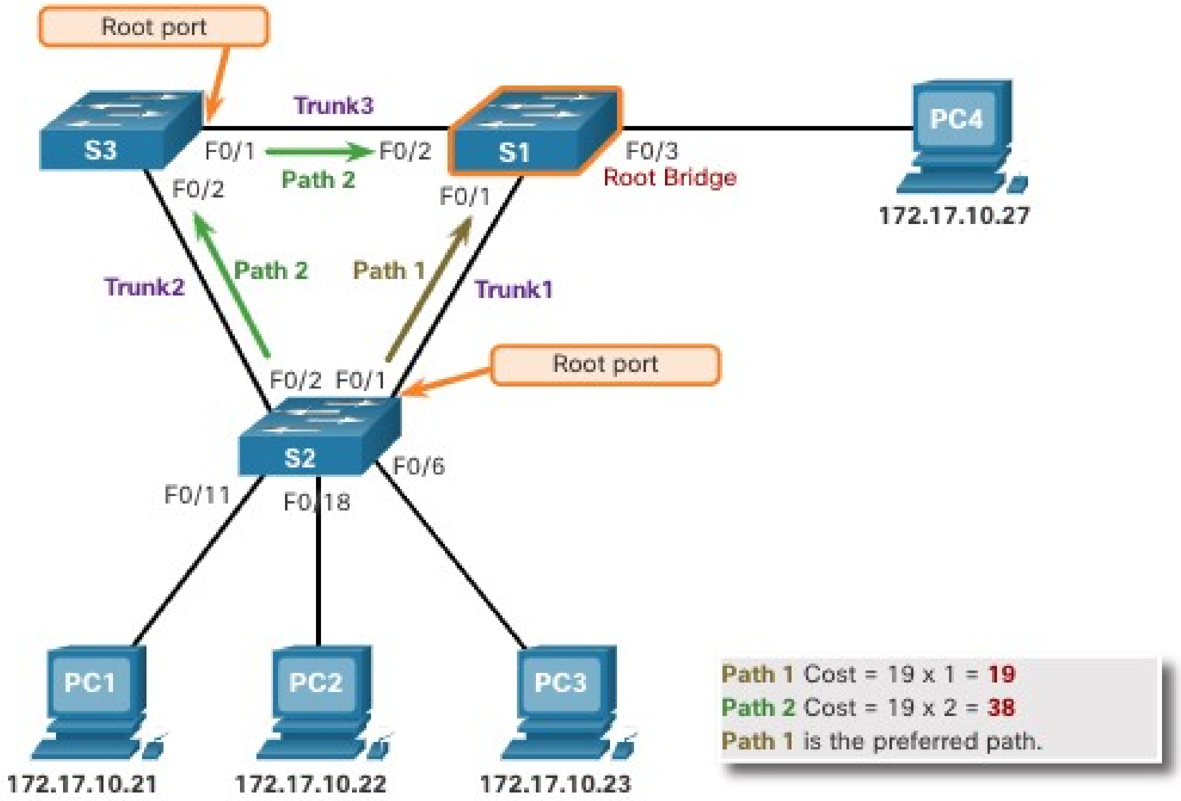
Determine the Root Path Cost

- When the root bridge has been elected for a given spanning tree instance, the STA starts determining the best paths to the root bridge from all destinations in the broadcast domain. The path information, known as the internal root path cost, is determined by the sum of all the individual port costs along the path from the switch to the root bridge.
- When a switch receives the BPDU, it adds the ingress port cost of the segment to determine its internal root path cost.
- The default port costs are defined by the speed at which the port operates. The table shows the default port costs suggested by IEEE. Cisco switches by default use the values as defined by the IEEE 802.1D standard, also known as the short path cost, for both STP and RSTP.
- 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.

Link Speed	STP Cost: IEEE 802.1D-1998	RSTP Cost: IEEE 802.1w-2004
10 Gbps	2	2,000
1 Gbps	4	20,000
100 Mbps	19	200,000
10 Mbps	100	2,000,000

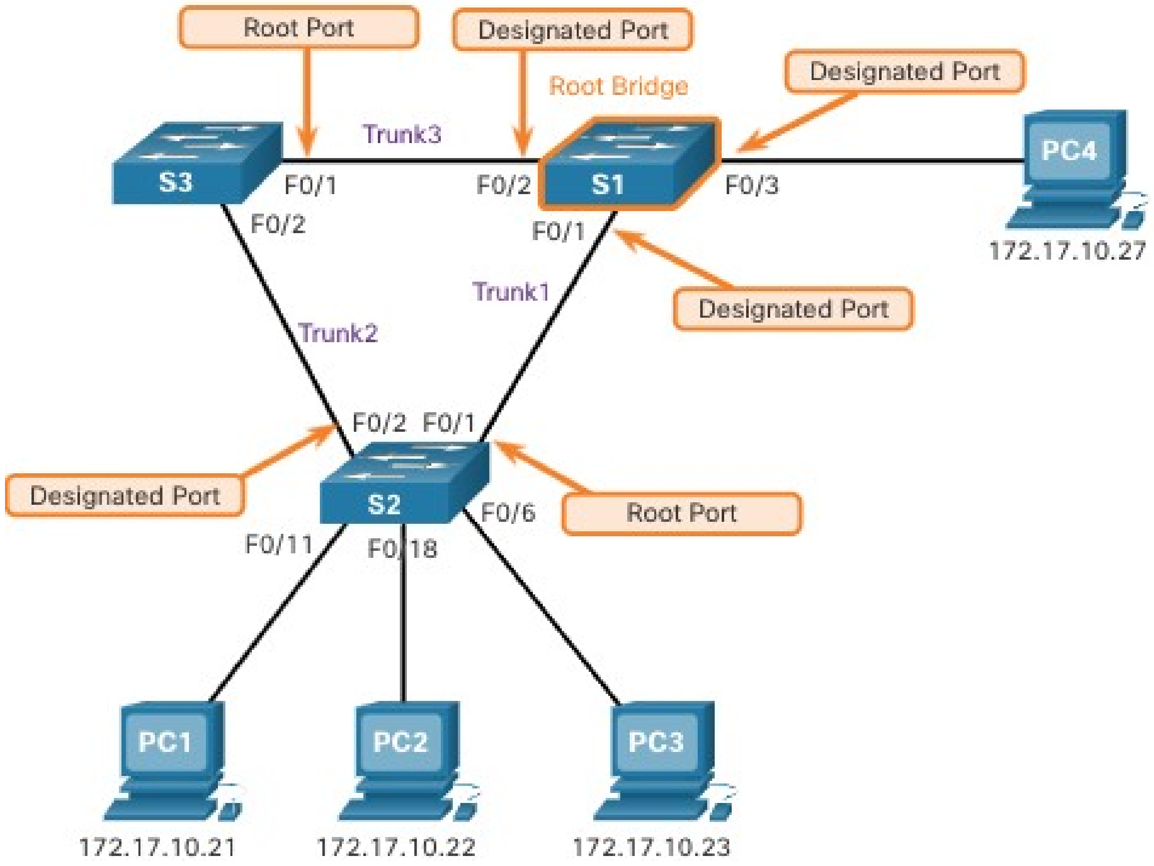
2. Elect the Root Ports

- After the root bridge has been determined, the STA algorithm is used to select the root port. Every non-root switch will select one root port. The root port is the port closest to the root bridge in terms of overall cost to the root bridge. This overall cost is known as the internal root path cost.
- The internal root path cost is equal to the sum of all the port costs along the path to the root bridge, as shown in the figure. Paths with the lowest cost become preferred, and all other redundant paths are blocked. In the example, the internal root path cost from S2 to the root bridge S1 over path 1 is 19 while the internal root 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 and F0/1 becomes the root port on S2.



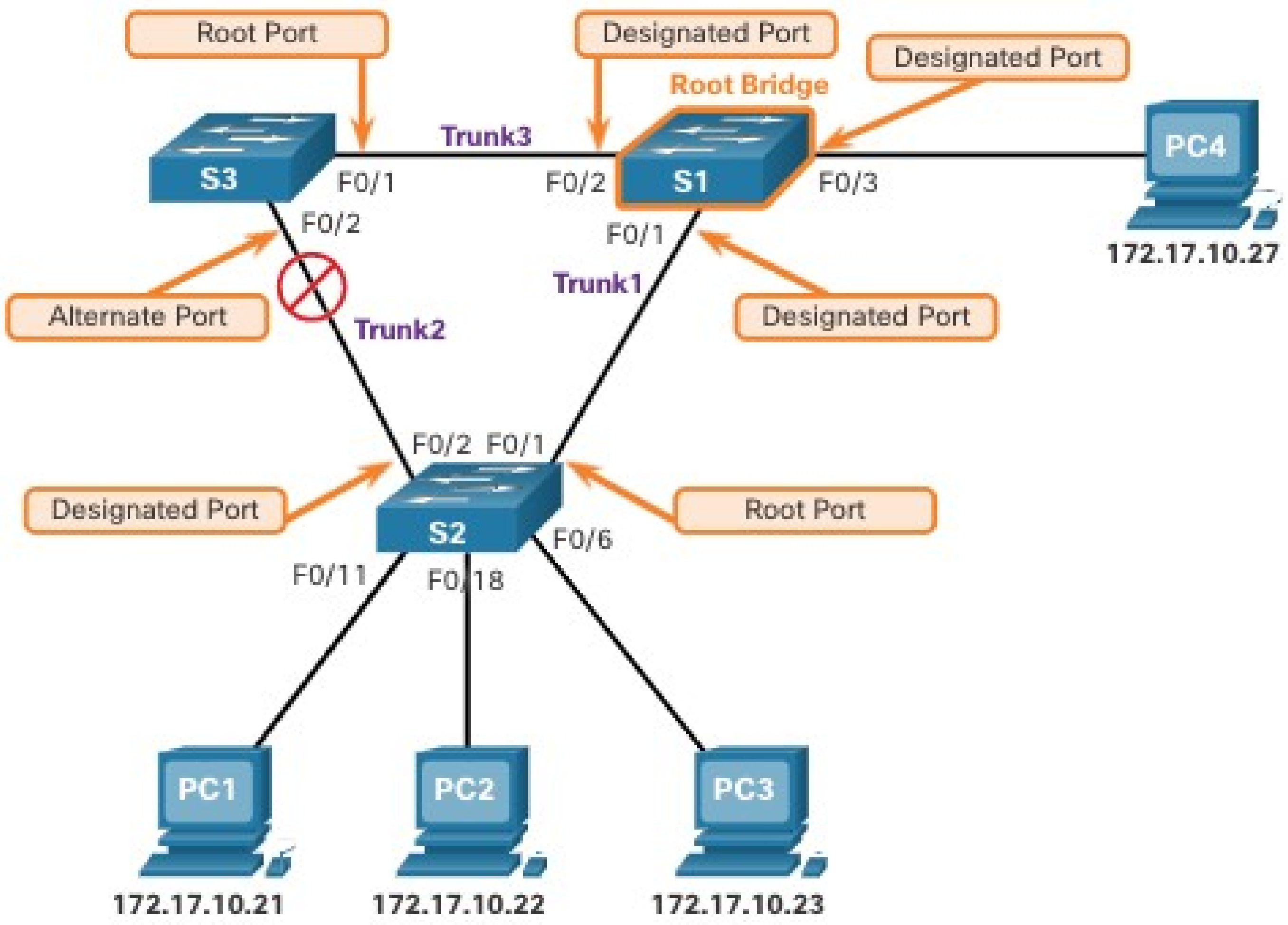
3. Elect Designated Ports

- Every segment between two switches will have one designated port. The designated port is a port on the segment that has the internal root path cost to the root bridge. In other words, the designated port has the best path to receive traffic leading to the root bridge.
- What is not a root port or a designated port becomes an alternate or blocked port.
- All ports on the root bridge are designated ports.
- If one end of a segment is a root port, the other end is a designated port.
- All ports attached to end devices are designated ports.
- On segments between two switches where neither of the switches is the root bridge, the port on the switch with the least-cost path to the root bridge is a designated port.



4. Elect Alternate (Blocked) Ports

If a port is not a root port or a designated port, then it becomes an alternate (or backup) port. Alternate ports are in discarding or blocking state to prevent loops. In the figure, the STA has configured port F0/2 on S3 in the alternate role. Port F0/2 on S3 is in the blocking state and will not forward Ethernet frames. All other inter-switch ports are in forwarding state. This is the loop-prevention part of STP.



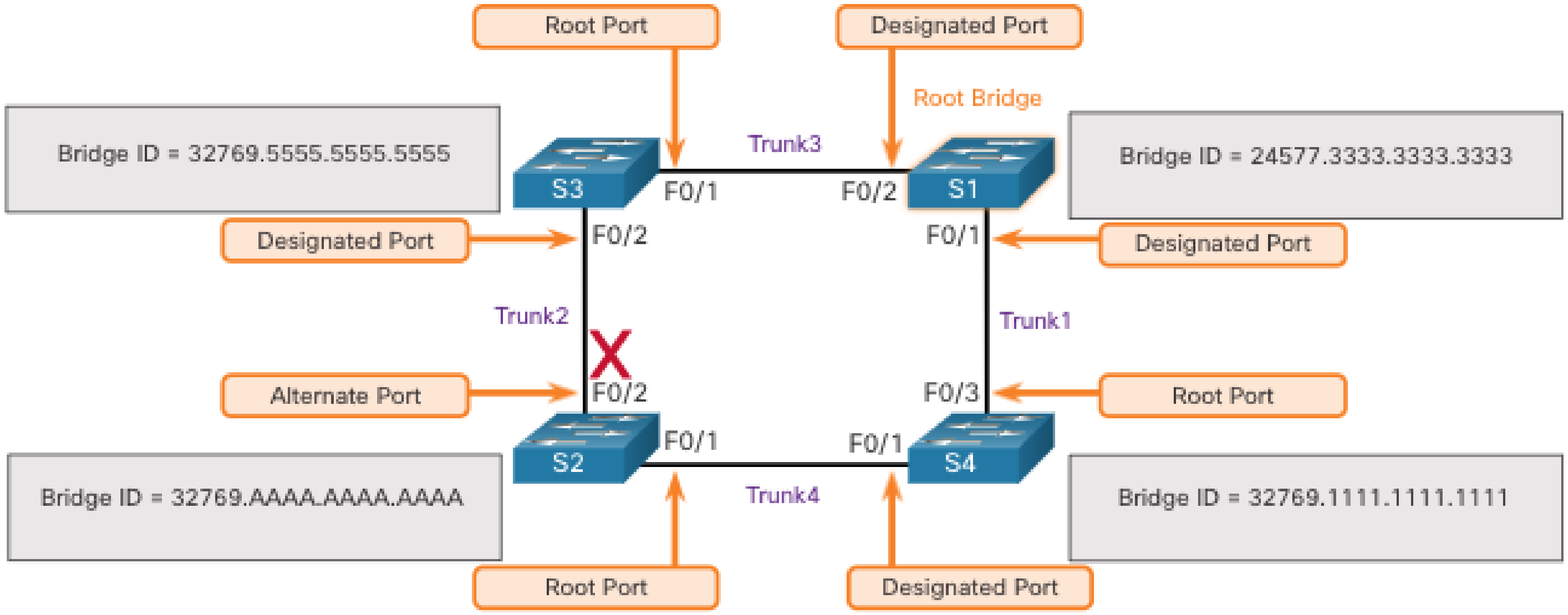
Elect a Root Port from Multiple Equal-Cost Paths

When a switch has multiple equal-cost paths to the root bridge, the switch will determine a port using the following criteria:

- Lowest sender BID
- Lowest sender port priority
- Lowest sender port ID

Elect a Root Port from Multiple Equal-Cost Paths (Cont.)

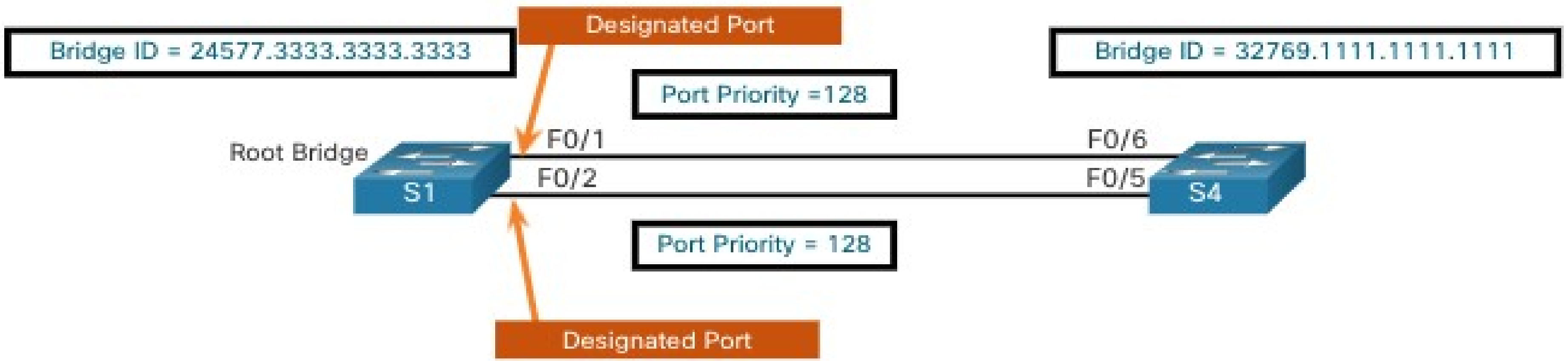
Lowest Sender BID: This topology has four switches with switch S1 as the root bridge. Port F0/1 on switch S3 and port F0/3 on switch S4 have been selected as root ports because they have the root path cost to the root bridge for their respective switches. S2 has two ports, F0/1 and F0/2 with equal cost paths to the root bridge. The bridge IDs of S3 and S4, will be used to break the tie. This is known as the sender's BID. S3 has a BID of 32769.5555.5555.5555 and S4 has a BID of 32769.1111.1111.1111. Because S4 has a lower BID, the F0/1 port of S2, which is the port connected to S4, will be the root port.



Elect a Root Port from Multiple Equal-Cost Paths (Cont.)

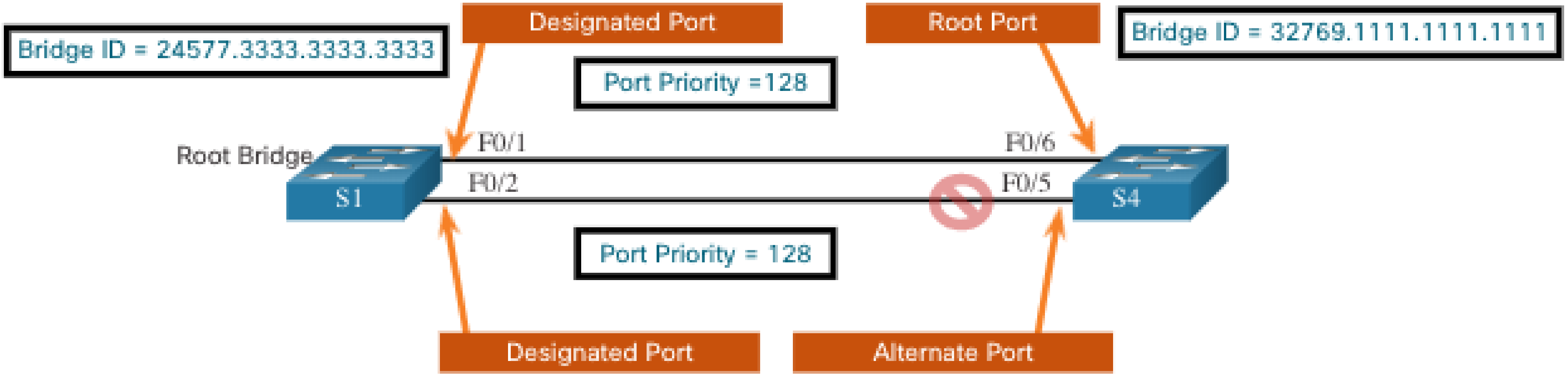
Lowest Sender Port Priority: This topology has two switches which are connected with two equal-cost paths between them. S1 is the root bridge, so both of its ports are designated ports.

- S4 has two ports with equal-cost paths to the root bridge. Because both ports are connected to the same switch, the sender's BID (S1) is equal. So the first step is a tie.
- Next, is the sender's (S1) port priority. The default port priority is 128, so both ports on S1 have the same port priority. This is also a tie. However, if either port on S1 was configured with a lower port priority, S4 would put its adjacent port in forwarding state. The other port on S4 would be a blocking state.



Elect a Root Port from Multiple Equal-Cost Paths (Cont.)

- **Lowest Sender Port ID:** The last tie-breaker is the lowest sender's port ID. Switch S4 has received BPDUs from port F0/1 and port F0/2 on S1. The decision is based on the sender's port ID, not the receiver's port ID. Because the port ID of F0/1 on S1 is lower than port F0/2, the port F0/6 on switch S4 will be the root port. This is the port on S4 that is connected to the F0/1 port on S1.
- Port F0/5 on S4 will become an alternate port and placed in the blocking state.



STP Timers and Port States

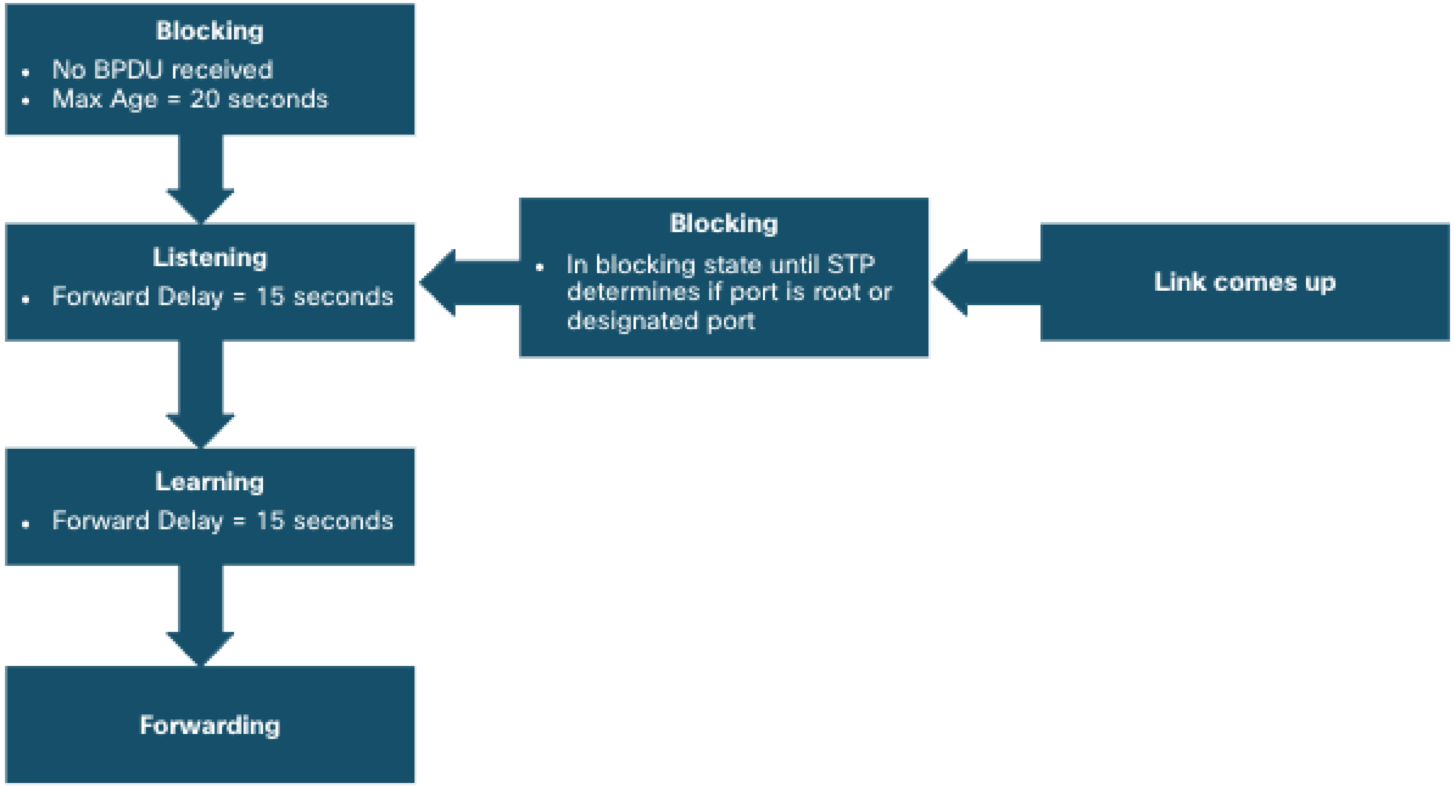
STP convergence requires three timers, as follows:

- **Hello Timer** -The hello time is the interval between BPDUs. The default is 2 seconds but can be modified to between 1 and 10 seconds.
- **Forward Delay Timer** -The forward delay is the time that is spent in the listening and learning state. The default is 15 seconds but can be modified to between 4 and 30 seconds.
- **Max Age Timer** -The max age is the maximum length of time that a switch waits before attempting to change the STP topology. The default is 20 seconds but can be modified to between 6 and 40 seconds.

Note: The default times can be changed on the root bridge, which dictates the value of these timers for the STP domain.

STP Timers and Port States (Cont.)

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. If a switch port transitions directly from the blocking state 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 has five ports states, four of which are operational port states as shown in the figure. The disabled state is considered non-operational.



Operational Details of Each Port State

The table summarizes the operational details of each port state

Port State	BPDU	MAC Address Table	Forwarding Data Frames
Blocking	Receive only	No update	No
Listening	Receive and send	No update	No
Learning	Receive and send	Updating table	No
Forwarding	Receive and send	Updating table	Yes
Disabled	None sent or received	No update	No

Per-VLAN Spanning Tree

STP can be configured to operate in an environment with multiple VLANs. In Per-VLAN Spanning Tree (PVST) versions of STP, there is a root bridge elected for each spanning tree instance. This makes it possible to have different root bridges for different sets of VLANs. STP operates a separate instance of STP for each individual VLAN. If all ports on all switches are members of VLAN 1, then there is only one spanning tree instance.

Different Versions of STP

- Many professionals generically use spanning tree and STP to refer to the 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 implementation or standard of spanning tree 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)."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.
- Because the two protocols share much of the same terminology and methods for the loop-free path, the primary focus will be on the current standard and the Cisco proprietary implementations of STP and RSTP.
- Cisco switches running IOS 15.0 or later, run PVST+ by default. This version incorporates many of the specifications of IEEE 802.1D-2004, such as alternate ports in place of the former non-designated ports. Switches must be explicitly configured for rapid spanning tree mode in order to run the rapid spanning tree protocol.

Different Versions of STP (Cont.)

STP Variety	Description
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. Also called Common Spanning Tree (CST), it assumes one spanning tree instance for the entire bridged network, regardless of the number of VLANs.
PVST+	Per-VLAN Spanning Tree (PVST+) is a Cisco enhancement of STP that provides a separate 802.1D spanning tree instance for each VLAN configured in the network. PVST+ 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.
RSTP	Rapid Spanning Tree Protocol (RSTP) or IEEE 802.1w is an evolution of STP that provides faster convergence than STP.
Rapid PVST+	This is a Cisco enhancement of RSTP that uses PVST+ and provides a separate instance of 802.1w per VLAN. Each separate instance supports PortFast, BPDU guard, BPDU filter, root guard, and loop guard.
MSTP	Multiple Spanning Tree Protocol (MSTP) 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.
MST	Multiple Spanning Tree (MST) is the Cisco implementation of MSTP, 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.

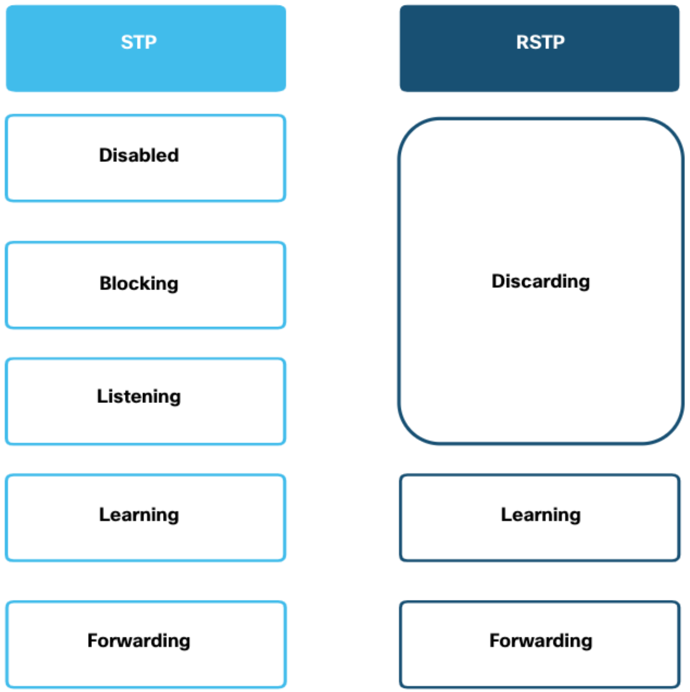
RSTP Concepts

- RSTP (IEEE 802.1w) supersedes the original 802.1D while retaining backward compatibility. The 802.1w STP terminology remains primarily the same as the original IEEE 802.1D STP terminology. Most parameters have been left unchanged. Users that are familiar with the original STP standard can easily configure RSTP. The same spanning tree algorithm is used for both STP and RSTP to determine port roles and topology.
- RSTP increases the speed of 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. If a port is configured to be an alternate port it can immediately change to a forwarding state without waiting for the network to converge.

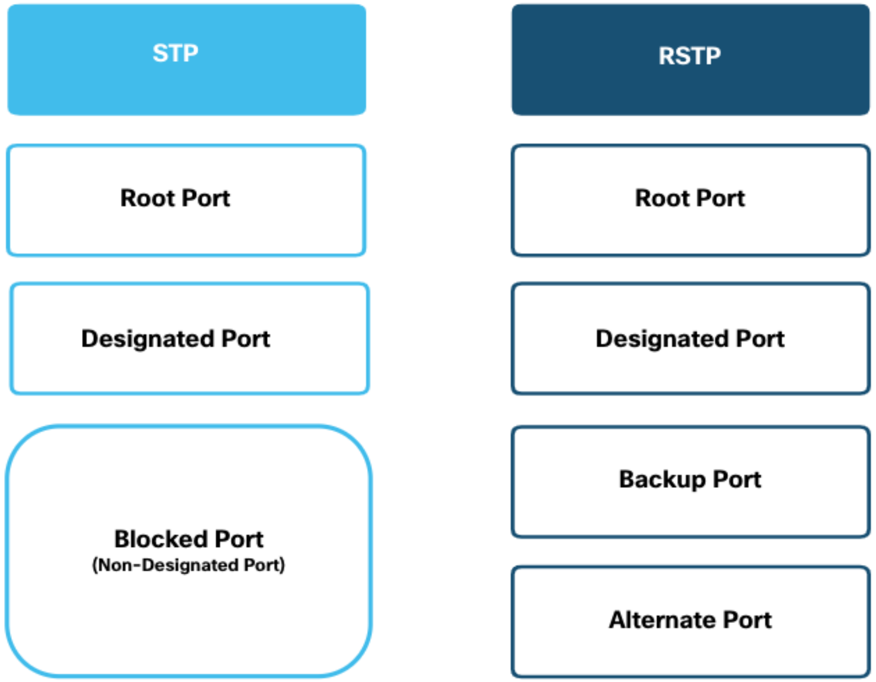
Note: Rapid PVST+ is the Cisco implementation of RSTP on a per-VLAN basis. With Rapid PVST+ an independent instance of RSTP runs for each VLAN.

RSTP Port States and Port Roles

There are only three port states in RSTP that correspond to the three possible operational states in STP. The 802.1D disabled, blocking, and listening states are merged into a unique 802.1w discarding state.

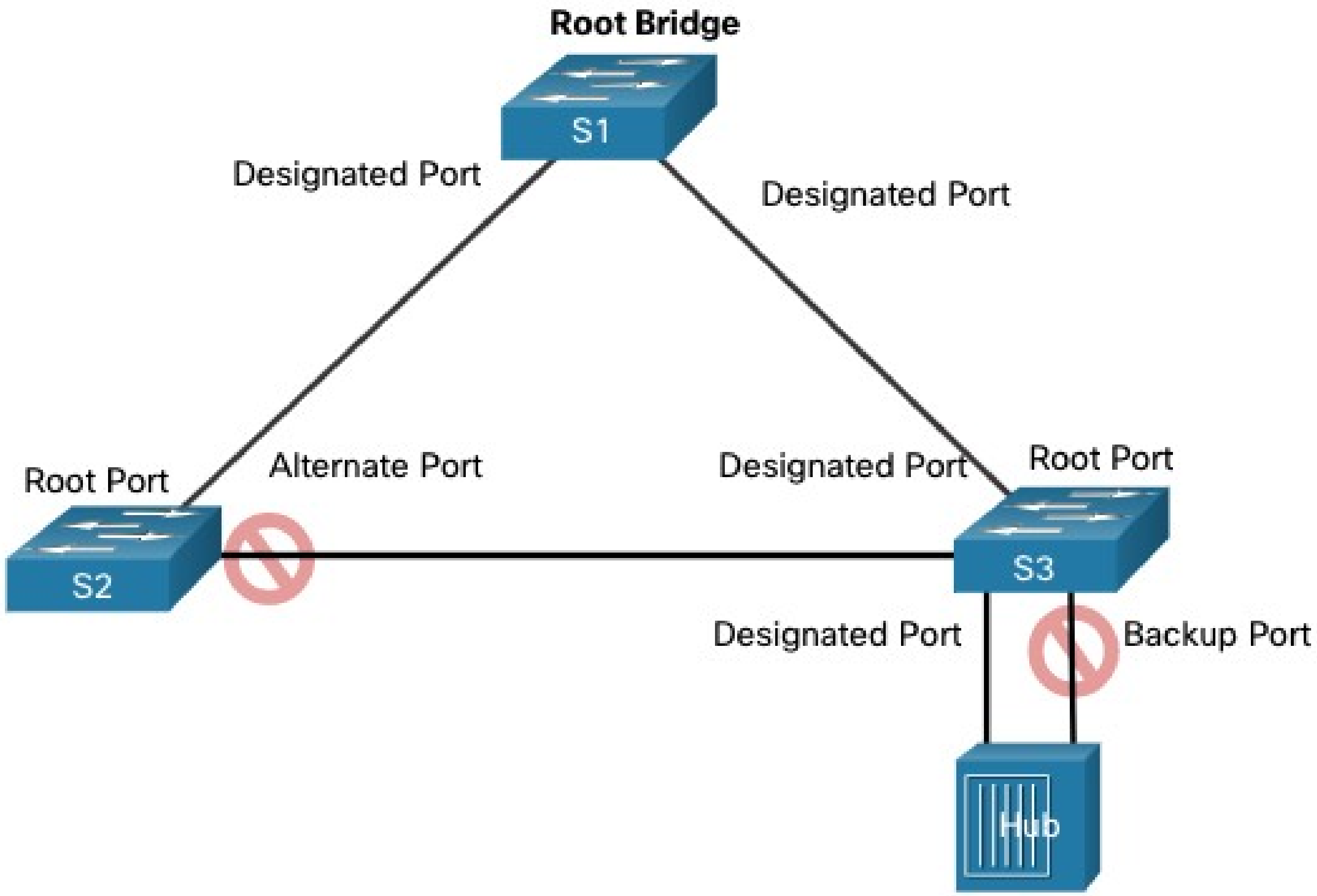


Root ports and designated ports are the same for both STP and RSTP. However, there are two RSTP port roles that correspond to the blocking state of STP. In STP, a blocked port is defined as not being the designated or root port. RSTP has two port roles for this purpose.



RSTP Port States and Port Roles (Cont.)

The alternate port has an alternate path to the root bridge. The backup port is a backup to a shared medium, such as a hub. A backup port is less common because hubs are now considered legacy devices.



PortFast and BPDU Guard

- When a device is connected to a switch port or when a switch powers up, the switch port goes through both the listening and learning states, each time waiting for the Forward Delay timer to expire. This delay is 15 seconds for each state for a total of 30 seconds. This can present a problem for DHCP clients trying to discover a DHCP server because the DHCP process may timeout. The result is that an IPv4 client will not receive a valid IPv4 address.
- When a switch port is configured with PortFast, that port transitions from blocking to forwarding state immediately, avoiding the 30 second delay. You can use PortFast on access ports to allow devices connected to these ports to access the network immediately. PortFast 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.
- A PortFast-enabled switch port should never receive BPDUs because that would indicate that switch is connected to the port, potentially causing a spanning tree loop. Cisco switches support a feature called BPDU guard. When enabled, it immediately puts the switch port in an errdisabled (error-disabled) state upon receipt of any BPDU. This protects against potential loops by effectively shutting down the port. The administrator must manually put the interface back into service.

Alternatives to STP

- Over the years, organizations required greater resiliency and availability in the LAN. Ethernet LANs went from a few interconnected switches connected to a single router, to a sophisticated hierarchical network design including access, distribution and core layer switches.
- Depending on the implementation, Layer 2 may include not only the access layer, but also the distribution or even the core layers. These designs may include hundreds of switches, with hundreds or even thousands of VLANs. STP has adapted to the added redundancy and complexity with enhancements, as part of RSTP and MSTP.
- An important aspect to network design is fast and predictable convergence when there is a failure or change in the topology. Spanning tree does not offer the same efficiencies and predictabilities provided by routing protocols at Layer 3.
- Layer 3 routing allows for redundant paths and loops in the topology, without blocking ports. For this reason, some environments are transitioning to Layer 3 everywhere except where devices connect to the access layer switch. In other words, the connections between access layer switches and distribution switches would be Layer 3 instead of Layer 2.

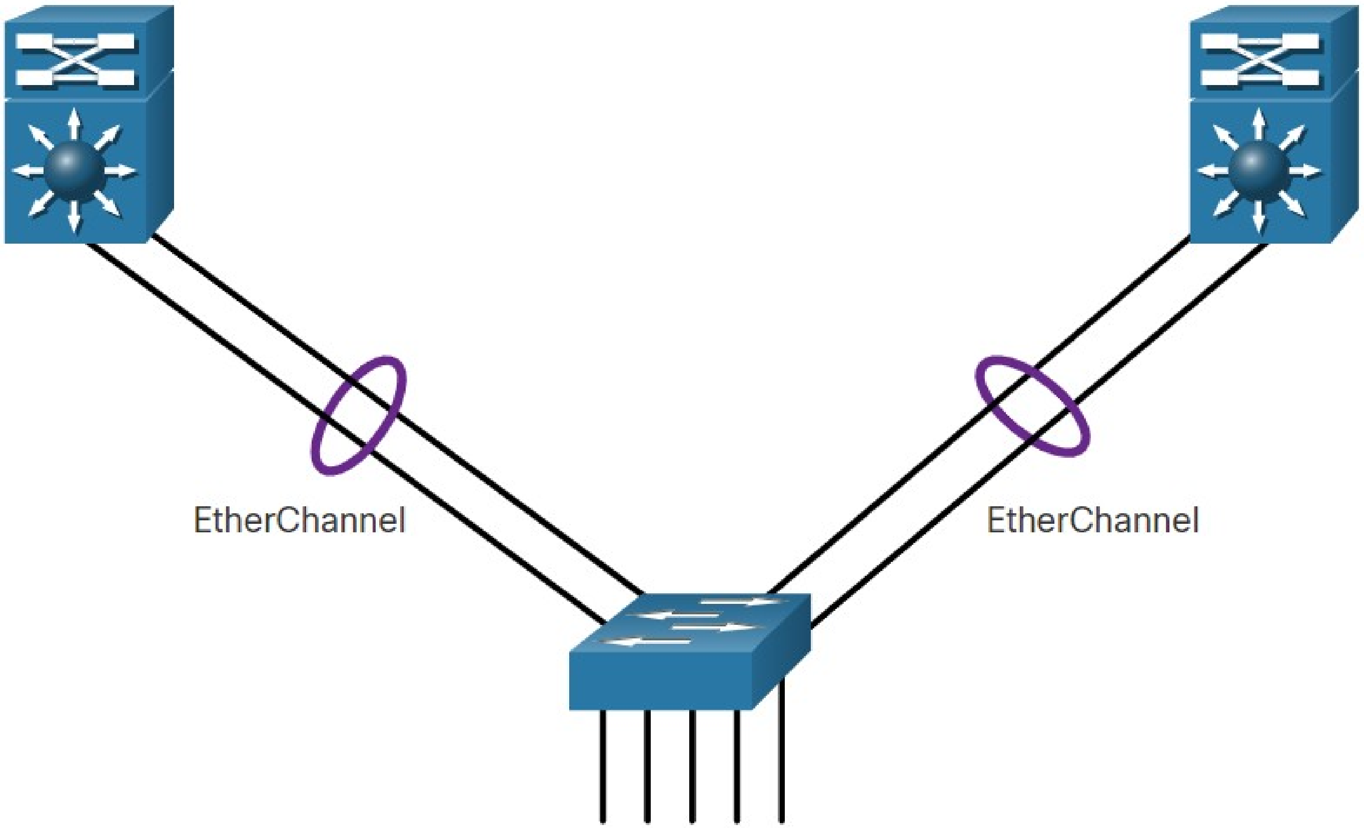
Link Aggregation

- There are scenarios in which more bandwidth or redundancy between devices is needed than what can be provided by a single link. Multiple links could be connected between devices to increase bandwidth. However, Spanning Tree Protocol (STP), which is enabled on Layer 2 devices like Cisco switches by default, will block redundant links to prevent switching loops.
- A link aggregation technology is needed that allows redundant links between devices that will not be blocked by STP. That technology is known as EtherChannel.
- EtherChannel is a link aggregation technology that groups multiple physical Ethernet links together into one single logical link. It is used to provide fault-tolerance, load sharing, increased bandwidth, and redundancy between switches, routers, and servers.
- EtherChannel technology makes it possible to combine the number of physical links between the switches to increase the overall speed of switch-to-switch communication.

EtherChannel

EtherChannel technology was originally developed by Cisco as a LAN switch-to-switch technique of grouping several Fast Ethernet or Gigabit Ethernet ports into one logical channel.

When an EtherChannel is configured, the resulting virtual interface is called a port channel. The physical interfaces are bundled together into a port channel interface, as shown in the figure.



Advantages of EtherChannel

EtherChannel technology has many advantages, including the following:

- Most configuration tasks can be done on the EtherChannel interface instead of on each individual port, ensuring configuration consistency throughout the links.
- EtherChannel relies on existing switch ports. There is no need to upgrade the link to a faster and more expensive connection to have more bandwidth.
- Load balancing takes place between links that are part of the same EtherChannel.
- EtherChannel creates an aggregation that is seen as one logical link. When several EtherChannel bundles exist between two switches, STP may block one of the bundles to prevent switching loops. When STP blocks one of the redundant links, it blocks the entire EtherChannel. This blocks all the ports belonging to that EtherChannel link. Where there is only one EtherChannel link, all physical links in the EtherChannel are active because STP sees only one (logical) link.
- EtherChannel provides redundancy because the overall link is seen as one logical connection. Additionally, the loss of one physical link within the channel does not create a change in the topology.

Implementation Restrictions

EtherChannel has certain implementation restrictions, including the following:

- Interface types cannot be mixed. For example, Fast Ethernet and Gigabit Ethernet cannot be mixed within a single EtherChannel.
- Currently each EtherChannel can consist of up to eight compatibly-configured Ethernet ports. EtherChannel provides full-duplex bandwidth up to 800 Mbps (Fast EtherChannel) or 8 Gbps (Gigabit EtherChannel) between one switch and another switch or host.
- The Cisco Catalyst 2960 Layer 2 switch currently supports up to six EtherChannels.
- The individual EtherChannel group member port configuration must be consistent on both devices. If the physical ports of one side are configured as trunks, the physical ports of the other side must also be configured as trunks within the same native VLAN. Additionally, all ports in each EtherChannel link must be configured as Layer 2 ports.
- Each EtherChannel has a logical port channel interface. A configuration applied to the port channel interface affects all physical interfaces that are assigned to that interface.

AutoNegotiation Protocols

EtherChannels can be formed through negotiation using one of two protocols, Port Aggregation Protocol (PAgP) or Link Aggregation Control Protocol (LACP). These protocols allow ports with similar characteristics to form a channel through dynamic negotiation with adjoining switches.

Note: It is also possible to configure a static or unconditional EtherChannel without PAgP or LACP.

PAgP Operation

PAgP (pronounced “Pag - P”) is a Cisco-proprietary protocol that aids in the automatic creation of EtherChannel links. When an EtherChannel link is configured using PAgP, PAgP packets are sent between EtherChannel-capable ports to negotiate the forming of a channel. When PAgP identifies matched Ethernet links, it groups the links into an EtherChannel. The EtherChannel is then added to the spanning tree as a single port.

When enabled, PAgP also manages the EtherChannel. PAgP packets are sent every 30 seconds. PAgP checks for configuration consistency and manages link additions and failures between two switches. It ensures that when an EtherChannel is created, all ports have the same type of configuration.

Note: In EtherChannel, it is mandatory that all ports have the same speed, duplex setting, and VLAN information. Any port modification after the creation of the channel also changes all other channel ports.

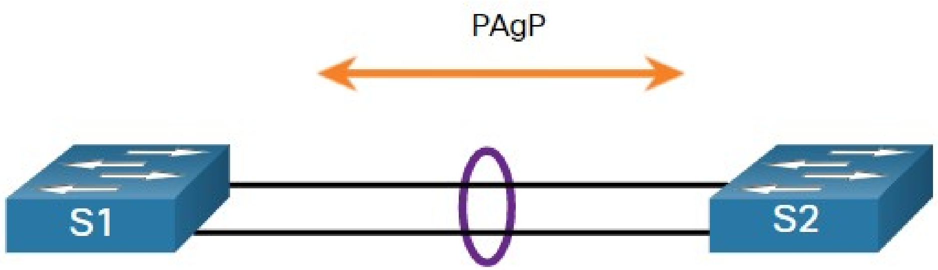
PAgP Operation (Cont.)

PAgP helps create the EtherChannel link by detecting the configuration of each side and ensuring that links are compatible so that the EtherChannel link can be enabled when needed. The modes for PAgP as follows:

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

The modes must be compatible on each side. If one side is configured to be in auto mode, it is placed in a passive state, waiting for the other side to initiate the EtherChannel negotiation. If the other side is also set to auto, the negotiation never starts and the EtherChannel does not form. If all modes are disabled by using the **no** command, or if no mode is configured, then the EtherChannel is disabled. The on mode manually places the interface in an EtherChannel, without any negotiation. It works only if the other side is also set to on. If the other side is set to negotiate parameters through PAgP, no EtherChannel forms, because the side that is set to on mode does not negotiate. No negotiation between the two switches means there is no checking to make sure that all the links in the EtherChannel are terminating on the other side, or that there is PAgP compatibility on the other switch.

PAgP Mode Settings Example



The table shows the various combination of PAgP modes on S1 and S2 and the resulting channel establishment outcome.

S1	S2	Channel Establishment
On	On	Yes
On	Desirable/Auto	No
Desirable	Desirable	Yes
Desirable	Auto	Yes
Auto	Desirable	Yes
Auto	Auto	No

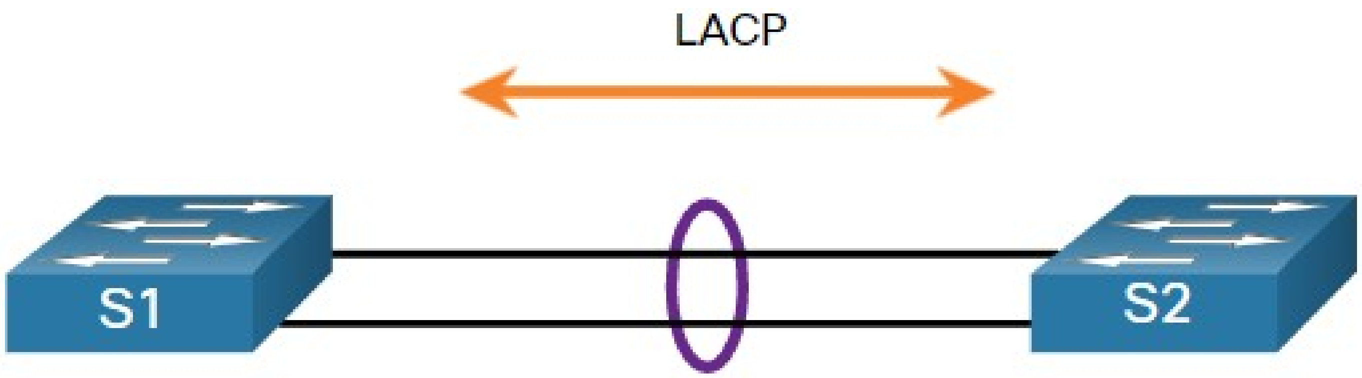
LACP Operation

LACP is part of an IEEE specification (802.3ad) that allows several physical ports to be bundled to form a single logical channel. LACP allows a switch to negotiate an automatic bundle by sending LACP packets to the other switch. It performs a function similar to PAgP with Cisco EtherChannel. Because LACP is an IEEE standard, it can be used to facilitate EtherChannels in multivendor environments. On Cisco devices, both protocols are supported.

LACP provides the same negotiation benefits as PAgP. LACP helps create the EtherChannel link by detecting the configuration of each side and making sure that they are compatible so that the EtherChannel link can be enabled when needed. The modes for LACP are as follows:

- **On** - This mode forces the interface to channel without LACP. Interfaces configured in the on mode do not exchange LACP packets.
- **LACP active** - This LACP mode places a port in an active negotiating state. In this state, the port initiates negotiations with other ports by sending LACP packets.
- **LACP passive** - This LACP mode places a port in a passive negotiating state. In this state, the port responds to the LACP packets that it receives but does not initiate LACP packet negotiation.

LACP Mode Settings Example



The table shows the various combination of LACP modes on S1 and S2 and the resulting channel establishment outcome.

S1	S2	Channel Establishment
On	On	Yes
On	Active/Passive	No
Active	Active	Yes
Active	Passive	Yes
Passive	Active	Yes
Passive	Passive	No

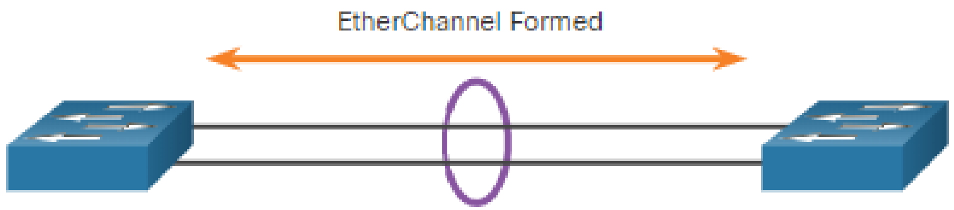
Configuration Guidelines

The following guidelines and restrictions are useful for configuring EtherChannel:

- **EtherChannel support** - All Ethernet interfaces must support EtherChannel with no requirement that interfaces be physically contiguous.
- **Speed and duplex** - Configure all interfaces in an EtherChannel to operate at the same speed and in the same duplex mode.
- **VLAN match** - All interfaces in the EtherChannel bundle must be assigned to the same VLAN or be configured as a trunk (shown in the figure).
- **Range of VLANs** - An EtherChannel supports the same allowed range of VLANs on all the interfaces in a trunking EtherChannel. If the allowed range of VLANs is not the same, the interfaces do not form an EtherChannel, even when they are set to **auto** or **desirable** mode.

Configuration Guidelines (Cont.)

- The figure shows a configuration that would allow an EtherChannel to form between S1 and S2.
- If these settings must be changed, configure them in port channel interface configuration mode. Any configuration that is applied to the port channel interface also affects individual interfaces. However, configurations that are applied to the individual interfaces do not affect the port channel interface. Therefore, making configuration changes to an interface that is part of an EtherChannel link may cause interface compatibility issues.
- The port channel can be configured in access mode, trunk mode (most common), or on a routed port.



S1 Port Configurations

Speed	1 Gbps
Duplex	Full
VLAN	10

S2 Port Configurations

Speed	1 Gbps
Duplex	Full
VLAN	10

LACP Configuration Example

Configuring EtherChannel with LACP requires the following three steps:

- **Step 1.** Specify the interfaces that compose the EtherChannel group using the **interface range** *interface* global configuration mode command. The **range** keyword allows you to select several interfaces and configure them all together.
- **Step 2.** Create the port channel interface with the **channel-group** *identifier* **mode active** command in interface range configuration mode. The identifier specifies a channel group number. The **mode active** keywords identify this as an LACP EtherChannel configuration.
- **Step3.** To change Layer 2 settings on the port channel interface, enter port channel interface configuration mode using the **interface port-channel** command, followed by the interface identifier. In the example, S1 is configured with an LACP EtherChannel. The port channel is configured as a trunk interface with the allowed VLANs specified.

```
S1(config)# interface range FastEthernet 0/1 - 2
S1(config-if-range)# channel-group 1 mode active
Creating a port-channel interface Port-channel 1
S1(config-if-range)# exit
S1(config-if)# interface port-channel 1
S1(config-if)# switchport mode trunk
S1(config-if)# switchport trunk allowed vlan 1,2,20
```

Verify EtherChannel

As always, when you configure devices in your network, you must verify your configuration. If there are problems, you will also need to be able to troubleshoot and fix them. There are a number of commands to verify an EtherChannel configuration:

- The **show interfaces port-channel** command displays the general status of the port channel interface.
- The **show etherchannel summary** command displays one line of information per port channel.
- The **show etherchannel port-channel** command displays information about a specific port channel interface.
- The **show interfaces etherchannel** command can provide information about the role of a physical member interface of the EtherChannel.

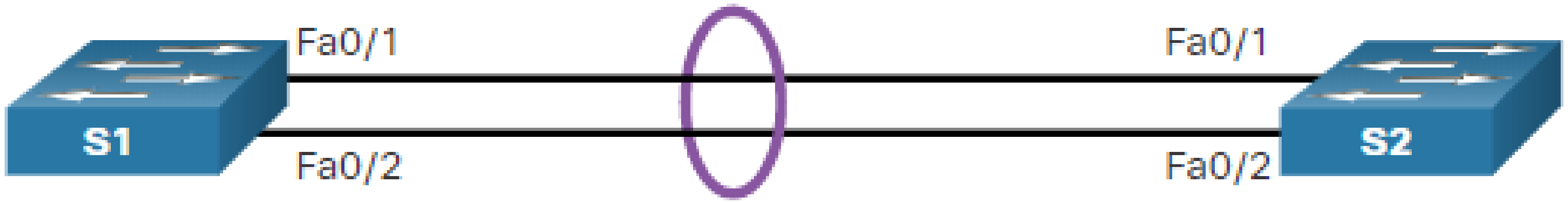
Common Issues with EtherChannel Configurations

All interfaces within an EtherChannel must have the same configuration of speed and duplex mode, native and allowed VLANs on trunks, and access VLAN on access ports. Ensuring these configurations will significantly reduce network problems related to EtherChannel. Common EtherChannel issues include the following:

- Assigned ports in the EtherChannel are not part of the same VLAN, or not configured as trunks. Ports with different native VLANs cannot form an EtherChannel.
- Trunking was configured on some of the ports that make up the EtherChannel, but not all of them. It is not recommended that you configure trunking mode on individual ports that make up the EtherChannel. When configuring a trunk on an EtherChannel, verify the trunking mode on the EtherChannel.
- If the allowed range of VLANs is not the same, the ports do not form an EtherChannel even when PAgP is set to the **auto** or **desirable** mode.
- The dynamic negotiation options for PAgP and LACP are not compatibly configured on both ends of the EtherChannel.

Troubleshoot EtherChannel Example

In the figure, interfaces F0/1 and F0/2 on switches S1 and S2 are connected with an EtherChannel. However, the EtherChannel is not operational.



Troubleshoot EtherChannel Example (Cont.)

Step 1. View the EtherChannel Summary Information: The output of the **show etherchannel summary** command indicates that the EtherChannel is down.

```
S1# show etherchannel summary
Flags:  D - down        P - bundled in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      N - not in use, no aggregation
        f - failed to allocate aggregator
        M - not in use, minimum links not met
        m - not in use, port not aggregated due to minimum links not met
        u - unsuitable for bundling
        w - waiting to be aggregated
        d - default port
        A - formed by Auto LAG
Number of channel-groups in use: 1
Number of aggregators:           1
Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------------------
1      Po1(SD)         -        Fa0/1(D)    Fa0/2(D)
```

Troubleshoot EtherChannel Example (Cont.)

Step 2. View Port Channel Configuration: In the **show run | begin interface port-channel** output, more detailed output indicates that there are incompatible PAgP modes configured on S1 and S2.

```
S1# show run | begin interface port-channel
interface Port-channel1
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
!
interface FastEthernet0/1
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
 channel-group 1 mode on
!
interface FastEthernet0/2
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
 channel-group 1 mode on
!======================================
S2# show run | begin interface port-channel
interface Port-channel1
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
!
interface FastEthernet0/1
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
 channel-group 1 mode desirable
!
interface FastEthernet0/2
 switchport trunk allowed vlan 1,2,20
 switchport mode trunk
 channel-group 1 mode desirable
```

Troubleshoot EtherChannel Example (Cont.)

Step 3: Correct the Misconfiguration: To correct the issue, the PAgP mode on the EtherChannel is changed to desirable.

Note: EtherChannel and STP must interoperate. For this reason, the order in which EtherChannel-related commands are entered is important, which is why you see interface Port-Channel 1 removed and then re-added with the **channel-group** command, as opposed to directly changed. If one tries to change the configuration directly, STP errors cause the associated ports to go into blocking or errdisabled state.

```
S1(config)# no interface port-channel 1
S1(config)# interface range fa0/1 - 2
S1(config-if-range)# channel-group 1 mode desirable
Creating a port-channel interface Port-channel 1
S1(config-if-range)# no shutdown
S1(config-if-range)# exit
S1(config)# interface range fa0/1 - 2
S1(config-if-range)# channel-group 1 mode desirable
S1(config-if-range)# no shutdown
S1(config-if-range)# interface port-channel 1
S1(config-if)# switchport mode trunk
S1(config-if)# end
S1#
```

Troubleshoot EtherChannel Example (Cont.)

Step 4. Verify EtherChannel is Operational: The EtherChannel is now active as verified by the output of the **show etherchannel summary** command.

```
S1# show etherchannel summary
Flags:  D - down        P - bundled in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      N - not in use, no aggregation
        f - failed to allocate aggregator
        M - not in use, minimum links not met
        m - not in use, port not aggregated due to minimum links not met
        u - unsuitable for bundling
        w - waiting to be aggregated
        d - default port
        A - formed by Auto LAG
Number of channel-groups in use: 1
Number of aggregators:           1
Group  Port-channel  Protocol    Ports
------+-------------+-----------+-------------------------------------------
1      Po1(SU)         PAgP      Fa0/1(P)    Fa0/2(P)
```