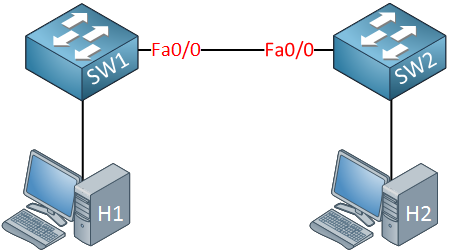
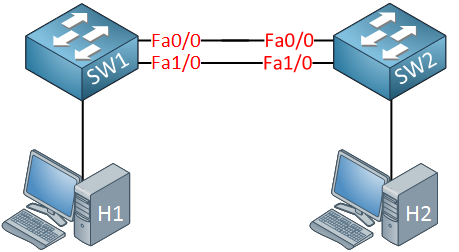
Why do we need spanning tree?

What is a loop, and how do we get one? Let me show you an example:



In the picture above, we have two switches. These switches are connected with a single cable, so there is a **single point of failure**. To get rid of this single point of failure, we will add another cable:



With the extra cable, we now have **redundancy**. Unfortunately for us, redundancy also brings **loops**. Why do we have a loop in the scenario above? Let me describe it to you:

1. H1 sends an ARP request because it’s looking for the MAC address of H2. An ARP request is a **broadcast** frame.
2. SW1 will forward this broadcast frame on all it interfaces, except the interface where it received the frame on.
3. SW2 will receive both broadcast frames.

Now, what does SW2 do with those broadcast frames?

1. It will forward it from every interface except the interface where it received the frame.
2. This means that the frame that was received on interface Fa0/0 will be forwarded on Interface Fa1/0.
3. The frame that was received on Interface Fa1/0 will be forwarded on Interface Fa0/0.

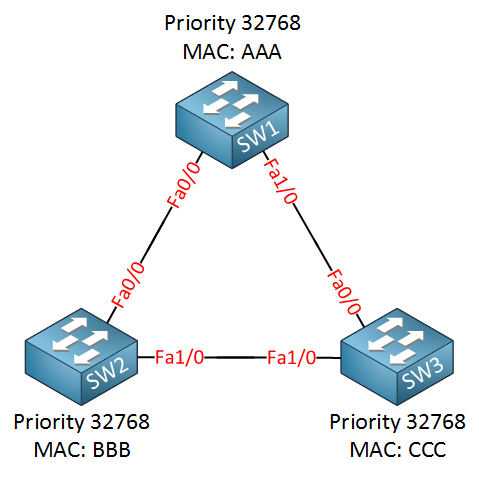
Do you see where this is going? We have a loop! Both switches will keep forwarding over and over again until the following happens:

* You fix the loop by disconnecting one of the cables.
* One of your switches will crash because they are overburdened with traffic.

Ethernet frames**don’t have a TTL** (Time to Live) value, so they will loop around forever. Besides ARP requests, many frames are broadcasted. For example, whenever the switch doesn’t know about a destination MAC address, it will be flooded.

How spanning tree solves loops

Spanning tree will help us to create a**loop-free topology** by blocking certain interfaces. Let’s take a look at how spanning tree work! Here’s an example:

******

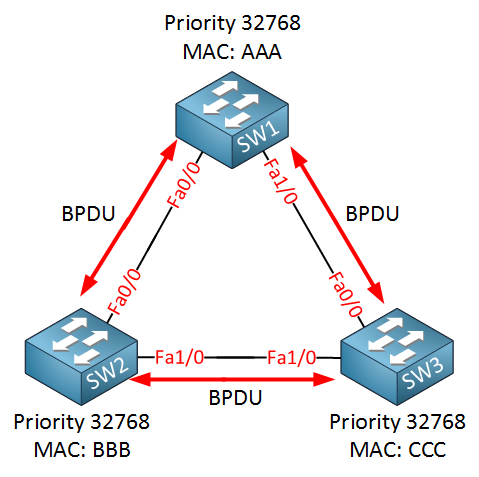
We have three switches, and as you can see, we have added redundancy by connecting the switches in a triangle, this also means we have a loop here. I have added the MAC addresses but simplified them for this example:

* SW1: MAC AAA
* SW2: MAC BBB
* SW3: MAC CCC

Since spanning tree is enabled, all our switches will send a special frame to each other called a **BPDU (Bridge Protocol Data Unit)**. In this BPDU, there are two pieces of information that spanning tree requires:

* **MAC address**
* **Priority**

The **MAC address** and the **priority** together make up the **bridge ID**. The BPDU is sent between switches as shown in the following picture:

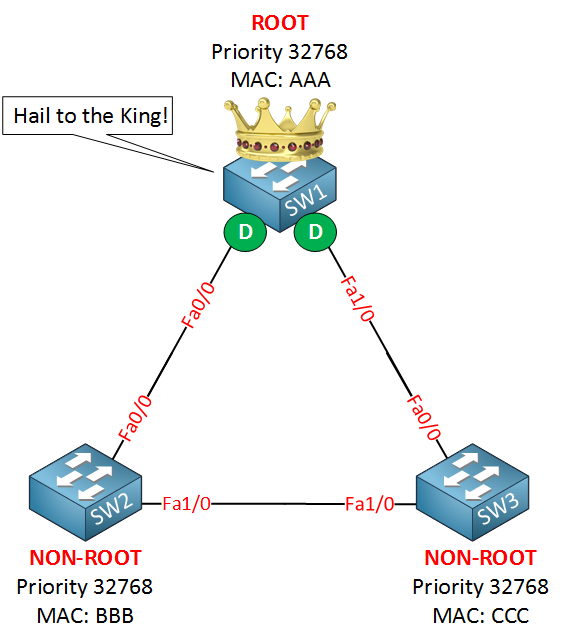


Spanning tree requires the bridge ID for its calculation. Let me explain how it works:

* First of all, spanning tree will **elect a root bridge**; this root bridge will be the one that has the best “bridge ID”.
* The switch with the**lowest bridge ID** is the best one.
* By default, the priority is **32768,** but we can change this value if we want.

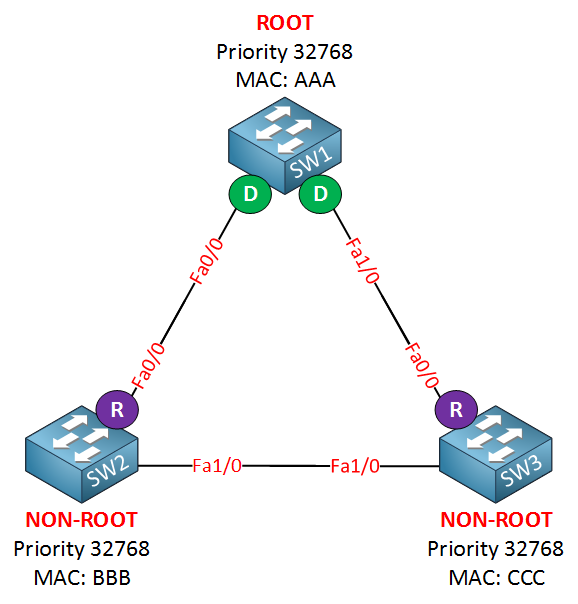
So who will become the root bridge? In our example, SW1 will become the root bridge! Priority and MAC address make up the bridge ID. Since the priority is the same on all switches, it will be the MAC address that is the tiebreaker. SW1 has the lowest MAC address thus the best bridge ID and will become the root bridge.

The ports on our root bridge are always **designated,** which means they are in a **forwarding** state. Take a look at the following picture:



Above, you see that SW1 has been elected as the root bridge and the “D” on the interfaces stands for designated.

Now we have agreed on the root bridge, our next step for all our **“non-root” bridges** (so that’s every switch that is not the root) will have to find the**shortest path to our root bridge**! The shortest path to the root bridge is called the**“root port”**. Take a look at my example:



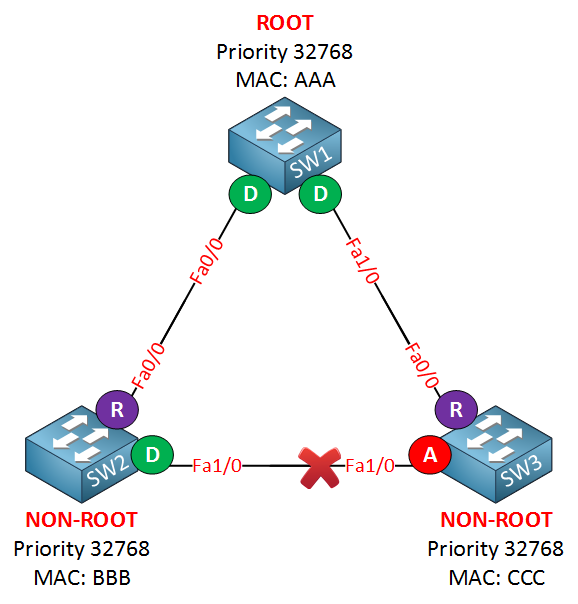
I’ve put an **“R” for “root port”** on SW2 and SW3. Their Fa0/0 interface is the shortest path to get to the root bridge. In my example, I’ve kept things simple, but “shortest path” in spanning tree means it will actually look at the **speed of the interface.**Each interface has a certain cost, and the path with the lowest cost will be used. Here’s an overview of the interfaces and their cost:

* 10 Mbit = Cost 100
* 100 Mbit = Cost 19
* 1000 Mbit = Cost 4

Excellent!…we have designated ports on our root bridge and root ports on our non-root bridges, we still have a loop, however, so we need to shut down a port between SW2 and SW3 to break that loop. So which port are we going to shut down? The one on SW2 or the one on SW3? We’ll look again at the best bridge ID:

* Bridge ID = Priority + MAC address.

Lower is better, both switches have the same priority, but the MAC address of SW2 is lower, which means that SW2 will “win this battle”. SW3 is our loser here which means it will have to block its port, effectively breaking our loop! Take a look at my example:



## What is STP and how does it work?

Spanning Tree Protocol (STP) is a Layer 2 network protocol used to prevent looping within a network topology. STP was created to avoid the problems that arise when computers exchange data on a local area network ([LAN](https://www.techtarget.com/searchnetworking/definition/local-area-network-LAN)) that contains redundant paths. If the flow of traffic is not carefully monitored and controlled, the data can be caught in a loop that circles around network segments, affecting performance and bringing traffic to a near halt.

Networks are often configured with redundant paths when connecting network segments. Although [redundancy](https://www.techtarget.com/whatis/definition/redundancy) can help protect against disaster, it can also lead to [bridge](https://www.techtarget.com/searchsecurity/definition/bridge) or switch looping. Looping occurs when data travels from a source to a destination along redundant paths and the data begins to circle around the same paths, becoming amplified and resulting in a broadcast storm.

STP can help prevent bridge looping on LANs that include redundant links. Without STP, it would be difficult to implement that redundancy and still avoid network looping. STP monitors all network links, identifies redundant connections and disables the ports that can lead to looping.

LANs are often divided into multiple network segments, and they use bridges to connect the individual segment pairs. Each message, called a frame, goes through the bridge before being sent to the intended destination. The bridge determines whether the message is for a destination within the same segment as the sender's or for another segment and then forwards the message accordingly. When used in the context of STP, the term *bridge* can also refer to a [network switch](https://www.techtarget.com/searchnetworking/definition/switch).

A bridge looks at the destination address and, based on its understanding of which computers are on which segments, forwards the data on the right path via the correct outgoing [port](https://www.techtarget.com/searchnetworking/definition/port). Network segmentation and bridging can reduce the amount of competition for a network path by half -- assuming each segment has the same number of computers. As a result, the network is much less likely to come to a halt.

A segmented LAN is often designed with redundant bridges and paths to ensure that communications can continue in the event that a network link becomes unavailable. However, this makes the network more susceptible to looping, so a system must be put into place to prevent this possibility, which is where STP comes in.

When STP is enabled, each bridge learns which computers are on which segment by sending a first-time message to network segments. Through this process, the bridge discovers the computers' locations and records the details in a table. When subsequent messages are sent, the bridge uses the table to determine which segment to forward them to. Enabling the bridge to learn about the network on its own is known as *transparent bridging*, a process that eliminates the need for an administrator to set up bridging manually.

In a network that contains redundant paths, bridges need to continually understand the [topology of the network](https://www.techtarget.com/searchnetworking/definition/network-topology) to control the flow of traffic and prevent looping. To do this, they exchange bridge protocol data units (BPDUs) via an extended LAN that uses a spanning tree protocol. BPDUs are data messages that provide the bridges with network information that's used to carry out STP operations.

At the heart of STP is the spanning tree algorithm that runs on each STP-enabled bridge. The [algorithm](https://www.techtarget.com/whatis/definition/algorithm) was specifically designed to avoid bridge loops when redundant paths exist. It uses the BPDUs to identify redundant links and select the best data path for forwarding messages. The algorithm also controls [packet](https://www.techtarget.com/searchnetworking/definition/packet) forwarding by setting the port state.

**What are STP port states?**

When STP is enabled on a network bridge, each port is set to one of five states to control frame forwarding:

1. **Disabled.** The port does not participate in frame forwarding or STP operations.
2. **Blocking.** The port does not participate in frame forwarding and discards frames received from the attached network segment. However, the port continues to listen for and process BPDUs.
3. **Listening.** From the blocking state, the port transitions to the listening state. The port discards frames from the attached network segment or forwarded from another port. However, it receives BPDUs and redirects them to the switch module for processing.
4. **Learning.** The port moves from the listening state to the learning state. It listens for and processes BPDUs but discards frames from the attached network segment or forwarded from another port. It also starts updating the address table with the information it's learned. In addition, it processes user frames but does not forward those frames.
5. **Forwarding.** The port moves from the learning state to the forwarding state and starts forwarding frames across the network segments. This includes frames from the attached network segment and those forwarded from another port. The port also continues to receive and process BPDUs, and the address table continues to be updated.

STP moves from the blocking state through the forwarding state in relatively short order, usually between 15 to 20 seconds for each state. Every port starts in the blocking state. If it's been disabled, the port enters directly into the blocking state upon being enabled. STP balances the states across ports to avoid bridge looping, while still making redundancy possible.

**What are STP modes?**

To understand STP modes, it helps to go back to STP's beginnings. The original spanning tree protocol and algorithm were invented in 1985 by Radia Perlman when she was working at Digital Equipment Corporation. Spanning tree protocols were later standardized by the Institute of Electrical and Electronics Engineers ([IEEE](https://www.techtarget.com/whatis/definition/IEEE-Institute-of-Electrical-and-Electronics-Engineers)). Since then, the protocol has evolved in a number of ways, and new variations have been introduced.

The following table provides an overview of the most common spanning tree protocols. However, not all bridges and switches support every one of these protocols, and there are other spanning tree-inspired protocols not listed here.

### STP Port Roles

* **Root -**Ports on non-root switches with the best cost path to root bridge. These ports forward data to the root bridge.
* **Designated**- Ports on root and designated switches. All ports on the root bridge will be designated.
* **Blocked**- All other ports to bridges or switches are in a blocked state. Access ports going to workstations or PCs are not affected.

### STP Election Process

When switches are first turned on, they will send configuration BPDUs containing their BIDs, with each switch initially believing themselves to be the root bridge. However, when a switch receives a BPDU with a superior (lower value) BID, that switch will stop originating configuration BPDUs and will instead relay these superior BPDUs to its neighboring switches.

Once a root bridge has finally been announced, a second election process begins to determine the “root port” selection process (the port on a switch that will forward frames to the root bridge). This process will follow the steps below until a root port is elected:

1. A switch port receives superior BPDUs from another switch and identifies that switch as the root bridge.
2. The port with the lowest root path is selected as the root port, if possible.
3. If the path cost is the same, the switch will select the port with the lowest sender BID as the selected root port.
4. If the sender BID is the same (usually the same switch), the port with the lowest physical port number on the sending switch will be selected as the root bridge (as the final tie-breaker).

Spanning Tree Priority: Root Primary and Root Secondary

Spanning Tree Protocol is a Layer 2 loop prevention mechanism that will block one port on the network switch if it detects a loop of broadcast messages within its architecture. By default, spanning trees are enabled on most interconnected Cisco switches. Switches send out Bridge Protocol Data Unit (BPDU) on all active interfaces. BPDU contains STP information needed to elect a root switch and detect loops.

STP Root Bridge Election

The switch assigns a root bridge within the interconnected switches. A root bridge is the central point of all switches and will be responsible for forwarding the traffic. The switch selects a root bridge by using the switch priority and the MAC address. Each switch has its own bridge ID and has a default priority value of 32768. The root bridge is taking precedence over the MAC address. If a switch has the lowest bridge priority value among the switches within the LAN, then it will be elected as the spanning tree root bridge.

If all the spanning tree bridge priority has the same priority value on all the switches, then the MAC address will be the tiebreaker. The lowest MAC address will be elected as the Root Bridge. Most of the older switches have a lower value of MAC address and have lower bandwidth and limited CPU/memory as compared to newer switches. Electing an older switch as the root bridge will cause a suboptimal operation on your network.