Spanning Tree Protocol STP Tutorial

June 1st, 2011 Go to comments

To provide for fault tolerance, many networks implement redundant paths between devices using multiple switches. However, providing redundant paths between segments causes packets to be passed between the redundant paths endlessly. This condition is known as a bridging loop.

(Note: the terms bridge, switch are used interchangeably when discussing STP)

To prevent bridging loops, the IEEE 802.1d committee defined a standard called the spanning tree algorithm (STA), or spanning tree protocol (STP). Spanning-Tree Protocol is a link management protocol that provides path redundancy while preventing undesirable loops in the network. For an Ethernet network to function properly, only one active path can exist between two stations.

Let's see a situation when there is no loop-avoidance process in operation. Suppose you have two switches connected with redundant links. One switch connected to PC A and the other switch connected to PC B.

Now PC A wants to talk to PC B. It then sends a broadcast, say an Address Resolution Protocol (ARP) to find out where the location of PC B, the green arrow shows a broadcast frame sent by PC A.

When the switch A receives a broadcast frame, it forwards that frame to all ports except the port where it receives the request -> SwA forwards that ARP frame out of fa0/0 and fa0/1 ports.

Suppose SwB receives the broadcast frame from fa0/0 first then it will forward that frame to the two other links (fa0/1 and fa0/5 of SwB).

The other broadcast frame from SwA comes to fa0/1 of SwB so SwB forwards it to fa0/0 and fa0/5.

As you can see, SwA has sent 2 broadcast frames out of its fa0/0 and fa0/1, SwB receives each of them, creates 2 copies and sends one of them back to SwA (the other is sent to PC B).

When SwA receives these broadcast frames it continues broadcasting them again to its other interfaces, this will keep going on forever until you shutdown the network. This phenomenon is called a **broadcast storm.**

Broadcast storm consumes entire bandwidth and denies bandwidth for normal network traffic. Broadcast storm is a serious network problem and can shut down entire network in seconds.

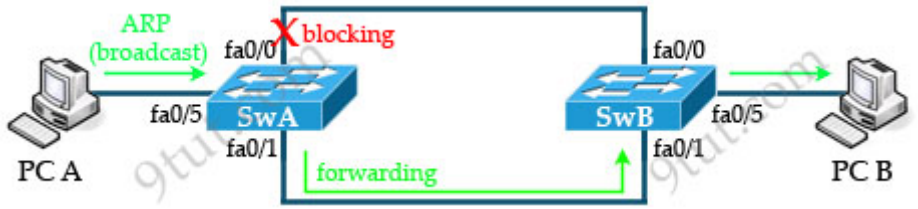
Other problems:

Multiple frame transmission: Multiple copies of unicast frames may 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. In the above example, if the first frame is not a ARP broadcast but a unicast and SwA and SwB haven't learned about the destination in that frame yet then they flood the frame on all ports except the originating port. The same phenomenon occurs and PC B will receive more than one copy of that frame.

MAC Database Instability: MAC database instability results when multiple copies of a frame arrive on different ports of a switch. We can see it in the above example too when the two ports on SwB (fa0/0 and fa0/1) receive the same frame.

Now you learned about problems when there is no looping-avoidance mechanism running on the network. All of these problems can be solved with the Spanning Tree Protocol (STP)

STP prevents loop by blocking one of switch's port. For example, by blocking port fa0/0 of SwA, no data traffic is sent on this link and the loop in the network is eliminated.



But how STP decides which port should be blocked. The whole process is more complex than what is shown above. We will learn it in the next part.

How Spanning Tree Protocol (STP) works

SPT must performs three steps to provide a loop-free network topology:

1. Elects one root bridge

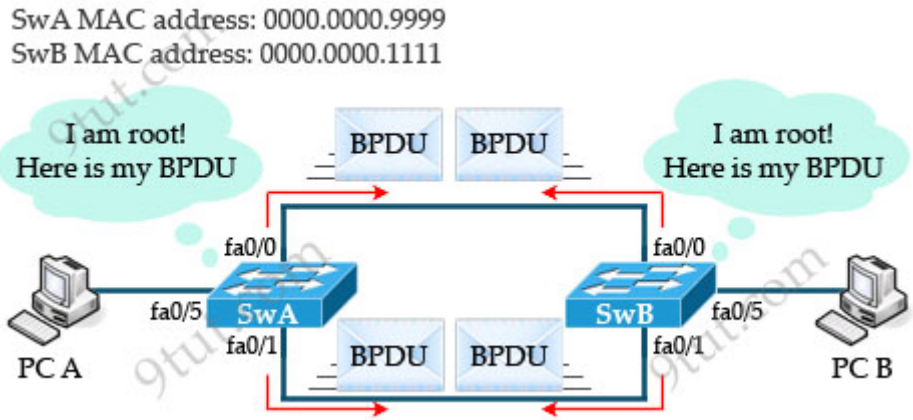
2. Select one root port per nonroot bridge

3. Select one designated port on each network segment

Now let's have a closer look from the beginning, when you have just turned on the switches…

1. Elects one root bridge

A fun thing is that when turned on, each switch claims itself as the root bridge immediately and starts sending out multicast frames called Bridge Protocol Data Units (BPDUs), which are used to exchange STP information between switches.



A BPDU contains many fields but there are 4 most important fields for STP to operate correctly:

*** The Bridge IDs of the Root Bridge and the Bridge ID of the Transmitting Bridge:**

In the initial stage, each switch claims itself as a root bridge so the bridge ID of the root bridge and the bridge ID of the transmitting bridge are the same.

The Bridge ID is composed of the **bridge priority** value (0-65535, 2 bytes) and the **bridge MAC address** (6 bytes).

Bridge ID = Bridge Priority + MAC Address

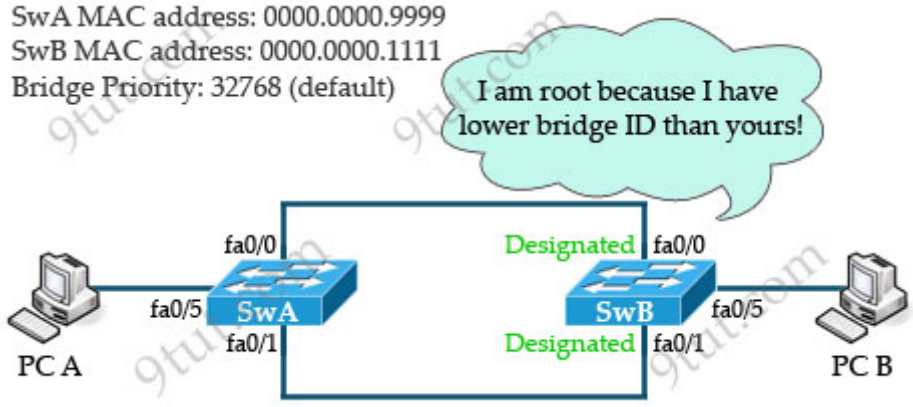
For example:

+ The bridge priority of SwA is 32768 and its MAC address is 0000.0000.9999 -> the bridge ID of SwA is 32768:0000.0000.9999

+ The bridge priority of SwB is 32768 and its MAC address is 0000.0000.1111 -> the bridge ID of SwB is 32768:0000.0000.1111

The root bridge is the bridge with the lowest bridge ID.

To compare two bridge IDs, the priority is compared first. If two bridges have equal priority, then the MAC addresses are compared. In the above example, both SwA and SwB have the same bridge ID (32768) so they will compare their MAC addresses. Because SwB has lower MAC address it will become root bridge.



On the root bridge, all ports are designated ports. Designated ports are in the forwarding state and can send and receive traffic.

Note: The default bridge priority value is 32768. An administrator can decide which bridge will become the root bridge by lowering the priority value (thus lowering Bridge ID). For example, we can lower SwA's bridge priority to 28672(smaller than 32768) to make it root bridge. But notice that the bridge priority number can be incremented only in step of 4096.

In conclusion, STP decides which switch will become root bridge by comparing the Bridge ID in the BPDUs. The bridge priorities are compare first; if they are equal then the MAC addresses will be used. Because each switch has a unique MAC address so surely one root bridge will be elected.

*** The cost to reach the root from this bridge (Root Path Cost):** This value is set to 0 at the beginning of STP root bridge election process since all bridges claim to be the root. The cost range is 0-65535.

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

The root path cost is used to elect root port and we will discuss in the next part.

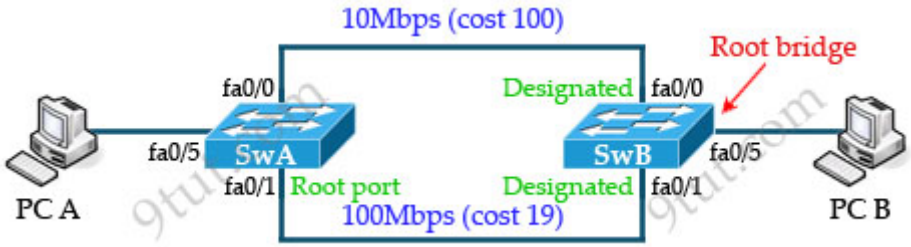
*** The Port ID**: The transmitting switch port ID, will be discussed later.

2. Select one root port per nonroot bridge

Root port is the port that is closest to the root bridge, which means it is the port that receiving the lowest-cost BPDU from the root.

Every non-root bridge must have a root port. All root ports are placed in forwarding state.

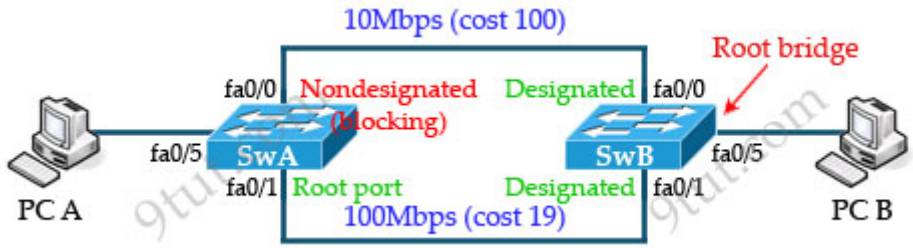
In the above example, if we suppose the upper link (between two fa0/0 interfaces) are 10Mbps and the lower link (between two fa0/1 interfaces) is 100Mbps link then fa0/1 of SwA will become root port as it has lower cost than fa0/0 (cost 19 < cost 100).



The root port election is much more complex when there are many switches so we wrote a separate tutorial. If you want to find out please read our STP Root Port Election Tutorial.

3. Select one designated port on each network segment

STP selects one designated port per segment to forward traffic. Other switch ports on the segment typically become nondesignated ports and are blocked. Therefore interface fa0/0 of SwA will become nondesignated port (blocking state). In blocking state, although switches cannot send data traffic but can still receive BPDUs.



Now the network reaches a state called **convergence**. Convergence in STP occurs when all ports on bridges and switches have transitioned to either forwarding or blocking states. No data is forwarded until convergence is complete so the time for convergence when network topology changes is very important. Fast convergence is very desirable in large networks. The normal convergence time is 50 seconds for 802.1D STP (which is rather slow) but the timers can be adjusted.

STP switch port states

When STP is enabled, every switch in the network goes through the blocking state and the transitory states of listening and learning. The ports then stabilize to the forwarding or blocking state.

* Blocking – no user data is sent or received but it may go into forwarding mode if the other links in use fail and the spanning tree algorithm determines the port may transition to the forwarding state. BPDU data is still received in blocking state but discards frames, does not learn MAC address.

* Listening – The switch processes BPDUs and awaits possible new information that would cause it to return to the blocking state, discards frames and MAC address.

* Learning – receives and transmits BPDUs and learns MAC addresses but does not yet forward frames.

* Forwarding – receives and sends data, normal operation, learns MAC address, receives and transmits BPDUs.

Below is a quick summary of STP states:

State	Can forward data?	Learn MAC?	Timer	Transitory or Stable State?
Blocking	No	No	Max Age (20 sec)	Stable
Listening	No	No	Forward Delay (15 sec)	Transitory
Learning	No	Yes	Forward Delay	Transitory
Forwarding	Yes	Yes		Stable

* **MaxAge** – How long any bridge should wait, after beginning to not hear hellos, before trying to change the STP topology. Usually this is a multiple of the hello time; the default is 20 seconds.

* **Forward Delay** – Delay that affects the time involved when an interface changes from blocking state to forwarding state. A port stays in listening state and then learning state for the number of seconds defined by the forward delay. This timer is covered in more depth shortly.

The spanning tree algorithm provides the following benefits:

* Eliminates bridging loops

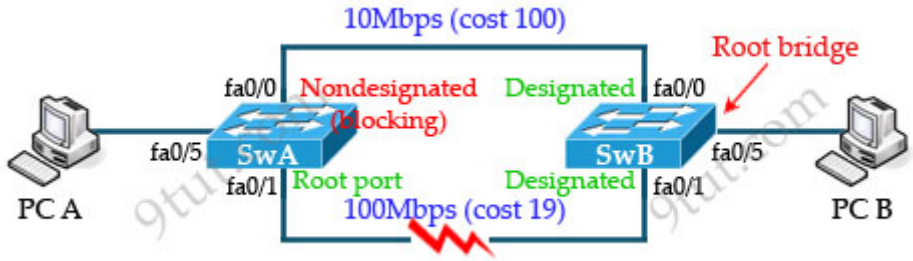
* Provides redundant paths between devices

* Enables dynamic role configuration

* Recovers automatically from a topology change or device failure

* Identifies the optimal path between any two network devices

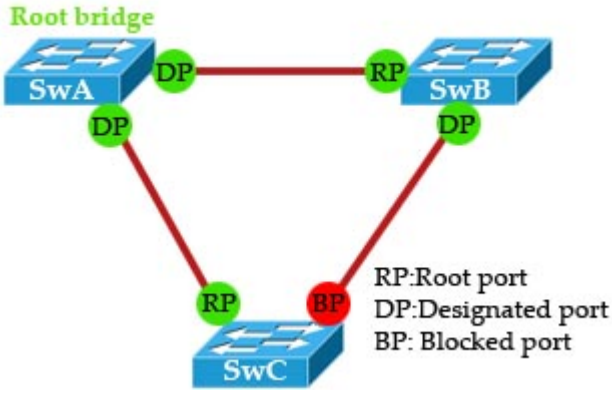
Now let's take an example using the same network as above but we suppose that the bottom 100Mbps connection is broken.



When the lower link is broken, SwA must wait for Max Age seconds before it begins to transition fa0/0 interface from blocking to listening state. In listening state it must wait for the Forward Delay seconds to move to the Learning state. Next it continues waiting for more Forward Delay seconds. If no BPDU is received, it is then placed in forwarding state. These three waiting periods of (by default) 20, 15, and 15 seconds create STP's relatively slow convergence.

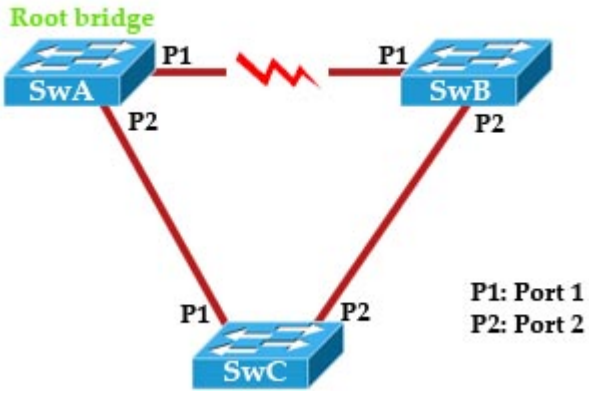
How STP performs when a link fails

Suppose we have a topology with three switches as shown below:



In which SwA is elected the root bridge, the link between SwB and SwC is being blocked. When STP is converged, the port roles are shown above.

Now suppose the link between SwA and SwB goes down, let us see what and how STP will perform



1. First, P1 on SwB immediately goes down and SwB declares its link to SwA as down.
2. SwB considers its link to SwC (which is being blocked) as an alternate link to root port. SwB starts to transition P2 from the blocking state to listening state -> learning state -> forwarding state. Each of these stages lasts 15 seconds by default. Therefore port P2 on SwB will be hold blocking for 30 seconds before the network converges again. This downtime of the network is rather long (although we can tune the timers to 14 second downtime) and the users can feel it.

The noticeable downtime can be reduced significantly if we use Rapid Spanning Tree Protocol (RSTP). If you are interested in RSTP, please read my Rapid Spanning Tree Protocol Tutorial.

Comments (1) Comments

1. Anonymous
   November 16th, 2021

   Very helpful

Add a Comment
Name

Submit Comment
Subscribe to comments feed
9tut.net for ICND 1 & ICND 2 has been launched! Configure NAT – GNS3 Lab

Premium Member Zone

Welcome Gurjeet singh!

- Welcome Premium Member
- CCNA – New Questions Part 5
- CCNA – New Questions Part 6
- CCNA – New Questions Part 7
- CCNA – New Questions Part 8
- CCNA – New Questions Part 9
- Composite Quizzes

- Logout

CCNA 200-301

- Basic Questions
- Topology Architecture Questions
- Cloud & Virtualization Questions
- CDP & LLDP Questions
- Switch Questions
- VLAN & Trunking Questions
- VLAN & Trunking Questions 2
- STP & VTP Questions
- EtherChannel Questions
- TCP & UDP Questions
- IP Address & Subnetting Questions
- IP Routing Questions
- IP Routing Questions 2

- OSPF Questions
- OSPF Questions 2
- EIGRP Questions
- NAT Questions
- NTP Questions
- Syslog Questions
- HSRP Questions
- Access-list Questions
- AAA Questions
- Security Questions
- Security Questions 2
- DAI Questions
- IPv6 Questions
- DNS Questions
- QoS Questions
- Port Security Questions
- Wireless Questions
- Wireless Questions 2
- SDN Questions
- DNA Center Questions
- Drag Drop Questions
- Drag Drop Questions 2
- Drag Drop Questions 3
- VPN Questions
- DHCP Questions
- Automation Questions
- Miscellaneous Questions
- CCNA FAQs & Tips
- Share your new CCNA Experience

CCNA Self-Study

- Practice CCNA GNS3 Labs
- CCNA Knowledge
- CCNA Lab Challenges
- Puppet Tutorial
- Chef Tutorial
- Ansible Tutorial
- JSON Tutorial
- Layer 2 Threats and Security Features
- AAA TACACS+ and RADIUS Tutorial
- STP Root Port Election Tutorial
- GRE Tunnel Tutorial
- Basic MPLS Tutorial
- TCP and UDP Tutorial
- Border Gateway Protocol BGP Tutorial
- Point to Point Protocol (PPP) Tutorial
- WAN Tutorial
- DHCP Tutorial
- Simple Network Management Protocol SNMP Tutorial
- Syslog Tutorial
- Gateway Load Balancing Protocol GLBP Tutorial
- EtherChannel Tutorial
- Hot Standby Router Protocol HSRP Tutorial
- InterVLAN Routing Tutorial
- Cisco Command Line Interface CLI
- Cisco Router Boot Sequence Tutorial

- OSI Model Tutorial
- Subnetting Tutorial – Subnetting Made Easy
- Frame Relay Tutorial
- Wireless Tutorial
- Virtual Local Area Network VLAN Tutorial
- VLAN Trunking Protocol VTP Tutorial
- IPv6 Tutorial
- Rapid Spanning Tree Protocol RSTP Tutorial
- Spanning Tree Protocol STP Tutorial
- Network Address Translation NAT Tutorial
- Access List Tutorial
- RIP Tutorial
- EIGRP Tutorial
- OSPF Tutorial

Network Resources

- Free Router Simulators
  - CCNA Website
  - ENCOR Website
  - ENSDWI Website
  - ENARSI Website
  - DevNet Website
  - CCIE R&S Website
  - Security Website
  - Wireless Website
  - Design Website
  - Data Center Website
  - Service Provider Website
  - Collaboration Website

Top