Spanning-Tree Protocol (STP) - II

STP States

To participate in STP, each port of a switch must progress through several states. A port begins in a Disabled state moving through several passive states and finally into an active state if allowed to forward traffic. The STP port states are: Disabled, Blocking, Listening, Learning, and Forwarding.

. Ports that are administratively shut down by the network administrator or by the system due to a fault condition are in the Disabled state. This state is special and is not part of the normal STP progression for a port.

. After a port initializes, it begins in the Blocking state so that no bridging loops can form. In the Blocking state, a port cannot receive or transmit data and cannot add MAC addresses to its address table. Instead, a port is only allowed to receive BPDUs. Also, ports that are put into standby mode to remove a bridging loop enter the Blocking state.

. The port will be moved from the Blocking state to the Listening state if the switch thinks that the port can be selected as a Root Port or Designated Port. In the Listening state, the port still cannot send or receive data frames. However, the port is allowed to receive and send BPDUs so that it can actively participate in the Spanning-Tree topology process. Here the port is finally allowed to become a Root Port or Designated Port because the switch can advertise the port by sending BPDUs to other switches. Should the port lose its Root Port or Designated Port status, it is returned to the Blocking state.

. After a period of time called the Forward Delay in the Listening state, the port is allowed to move into the Learning state. The port still sends and receives BPDUs as before. In addition, the switch can now learn new MAC addresses to add into its address table.

. After another Forward Delay period in the Learning state, the port is allowed to move into the Forwarding state. The port can now send and receive data frames, collect MAC addresses into its address table, and send and receive BPDUs. The port is now a fully functioning switch port within the Spanning-Tree topology.

STP Timers

STP operates as switches send BPDUs to each other in an effort to form a loop-free topology. The BPDUs take a finite amount of time to travel from switch to switch. In addition, news of a topology change such as a link or Root Bridge failure can suffer from propagation delays as the announcement travels from one side of a network to the other. Because of the possibility of these delays, preventing the Spanning-Tree topology from converging until all switches have had time to receive accurate information is important. STP uses three timers for this purpose. There are three timers: Hello Time, Forward Delay, and Max Age.

. Hello Time is the time interval between Configuration BPDUs sent by the Root Bridge. The Hello Time value configured in the Root Bridge switch will determine the Hello Time for all non-root switches. However, all switches have a locally configured Hello Time that is used to time Topology Change Notification (TCN) BPDUs when they are retransmitted. The IEEE 802.1D standard specifies a default Hello Time value of two seconds.

. Forward Delay is the time interval that a switch port spends in both the Listening and Learning states. The default value is 15 seconds.

. Max Age is the time interval that a switch stores a BPDU before discarding it. While executing the STP, each switch port keeps a copy of the "best" BPDU that it has heard. If the source of the BPDU loses contact with the switch port, the switch will notice that a topology change has occurred after the Max Age time elapses and the BPDU is aged out. The default Max Age value is 20 seconds.

To announce a change in the active network topology, switches send a Topology Change Notification (TCN) BPDU. This occurs when a switch either moves a port into the Forwarding state or moves a port from Forwarding or Learning into the Blocking state. The switch sends a TCN BPDU out its Designated Port. The TCN BPDU carries no data about the change, but only informs recipients that a change has occurred. However, the switch will not send TCN BPDUs if the port has been configured with PortFast enabled. The switch will continue sending TCN BPDUs every Hello Time interval until it gets an acknowledgement from an

upstream neighbor. As the upstream neighbors receive the TCN BPDU, they will propagate it on toward the Root Bridge. When the Root Bridge receives the BPDU, the Root Bridge sends out an acknowledgement. The Root Bridge also sends out the Topology Change flag in a Configuration BPDU so that all other bridges will shorten their bridge table aging times down from the default 300 seconds to the Forward Delay value. This condition causes the learned locations of MAC addresses to be flushed out sooner than they normally would, easing the bridge table corruption that might occur due to the change in topology. However, any stations that are actively communicating during this time will be kept in the bridge table. This condition lasts for the sum of the Forward Delay and the Max Age.

The three STP timers can be adjusted. These timers need only be modified on the Root Bridge and any secondary or backup Root Bridges because the Root Bridge propagates all three timer values throughout the network in the Configuration BPDU.

Optional STP Features

Cisco has added several proprietary enhancements to STP and to the logic used by its switches. Also, the IEEE, which owns the STP specifications, has made other enhancements, some similar to Cisco's proprietary enhancements.

EtherChannel

EtherChannel combines from two to eight parallel Ethernet trunks between the same pair of switches, bundled into an EtherChannel. STP treats an EtherChannel as a single link, so if at least one of the links is up, STP convergence does not have to occur. With each pair of Ethernet links configured as an EtherChannel, STP treats each EtherChannel as a single link. Thus, both links to the same switch must fail for a switch to need to cause STP convergence. Without EtherChannel, if you have multiple parallel links between two switches, STP blocks all the links except one. With EtherChannel, all the parallel links can be up and working at the same time, while reducing the number of times STP must converge, which in turn makes the network more available.

EtherChannel also provides more network bandwidth. All trunks in an EtherChannel are either forwarding or blocking, because STP treats all the trunks in the same EtherChannel as one trunk. When an EtherChannel is in forwarding state, the switches forward traffic over all the trunks, providing more bandwidth.

PortFast

PortFast allows a switch to place a port in forwarding state immediately when the port becomes physically active. However, the only ports on which you can safely enable PortFast are ports on which you know that no bridges, switches, or other STP devices are connected. Thus, PortFast is most appropriate for connections to end-user devices. If you turn on PortFast for end-user devices, when an end-user PC boots, as soon as the Ethernet card is active, the switch port can forward traffic. Without PortFast, each port must wait MaxAge plus twice Forwarding Delay, which is 50 seconds with the default MaxAge and Forward Delay settings.

Rapid Spanning Tree (IEEE 802.1w)

The IEEE has improved the 802.1d protocol, which defines STP, with the definition of Rapid Spanning Tree Protocol (RSTP), as defined in standard 802.1w. RSTP is similar to STP in that it elects the root switch using the same parameters and tiebreakers; elects the root port on nonroot switches with the same rules; elects designated ports on each LAN segment with the same rules; and places each port in either a forwarding state or a blocking state, with the latter being called the discarding state instead of the blocking state.

RSTP can be deployed alongside traditional STP bridges and switches, with RSTP features working in switches that support it, and STP features working in the switches that support only STP. The advantage RSTP has over STP is improved network convergence when network topology changes occur. STP convergence has essentially wait periods: a switch must first cease to receive root BPDUs for MaxAge seconds before it can begin to transition any interfaces from blocking to forwarding. For any interfaces that need to transition from blocking to forwarding, the interface must endure Forward Delay seconds in listening state and Forward Delay more seconds in learning state before being placed in forwarding state. By default, these three wait periods of are 20, 15, and 15 seconds.

RSTP convergence times typically take less than 10 seconds. In some cases, they can be as low as 1 to 2 seconds.