Spanning Tree Procolol

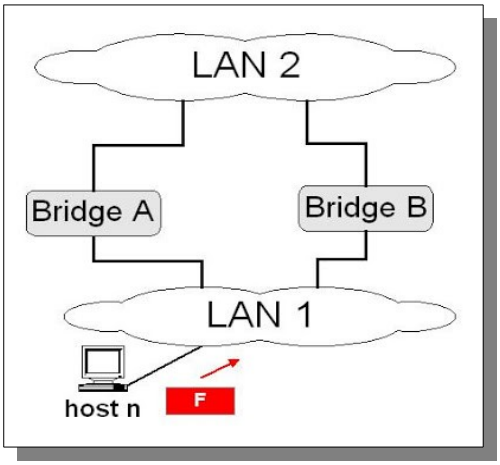
El STP (Spanning Tree Protocol) asegura que exista sólo una ruta lógica entre todos los destinos de la red, al realizar un bloqueo de forma intencional a aquellas rutas redundantes que puedan ocasionar un bucle.

El protocolo usado por STP es el IEEE 802.1D

Un puerto se considera bloqueado cuando el tráfico de la red no puede ingresar ni salir del puerto. Esto no incluye las tramas de unidad de datos del protocolo comúnmente llamadas (BPDU) utilizadas por STP para evitar bucles. Las rutas físicas aún existen para proporcionar la redundancia, pero las mismas se deshabilitan para evitar que se generen bucles. Si alguna vez la ruta es necesaria para compensar la falla de un cable de red o de un switch, STP vuelve a calcular las rutas y desbloquea los puertos necesarios para permitir que la ruta redundante se active.

La función del STP es la de gestionar la presencia de bucles en topologías de red debido a la existencia de enlaces redundantes. El protocolo permite a los dispositivos de interconexión activar o desactivar automáticamente los enlaces de conexión, de forma que se garantice que la topología esté libre de bucles.

Considere dos LANs que estén conectadas por dos bridges, asuma que el *host n* está transmitiendo un f*rame (Trama) F*, con un destino desconocido.



¿Que pasaría?, estaríamos en presencia de lo siguiente:

- El Bridge A envía este Frame a la LAN 2.
- El Bridge B ve este Frame F en la LAN 2 (with unknown destination, con destino desconocido), y envía este Frame a la LAN 1.
- El Bridge A hace lo mismo y el reenvío continua.

¿Dónde está el problema?, ¿cuál es la solución?

El problema está en que ocurre un bucle, también conocido como *loop* o *lazo*. Cuando hay bucles en la topología de red, los dispositivos de interconexión de nivel de enlace (como un puente de red o un conmutador de paquetes) reenvían indefinidamente las tramas Broadcast y Multicast, al no existir ningún campo TTL (Time To Live, Tiempo de Vida) en la Capa 2, tal y como ocurre en la Capa 3. Se consume entonces una gran cantidad de ancho de banda, y en muchos casos la red queda inutilizada. La solución consiste en permitir la existencia de enlaces físicos redundantes, pero creando una topología lógica libre de bucles. STP permite solamente una

trayectoria activa a la vez entre dos dispositivos de la red (esto previene los bucles) pero mantiene los caminos redundantes como reserva, para activarlos en caso de que el camino inicial falle.

El árbol de expansión (Spanning Tree) permanece vigente hasta que ocurre un cambio en la topología, situación que el protocolo es capaz de detectar de forma automática. El máximo tiempo de duración del árbol de expansión es de cinco minutos. Cuando ocurre uno de estos cambios, el puente raíz actual redefine la topología del árbol de expansión o se elige un nuevo puente raíz.

Funcionamiento.

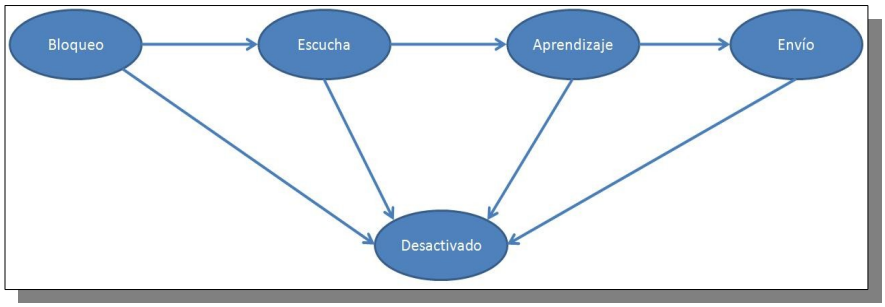
Este algoritmo cambia una red física con forma de malla, en la que existen bucles, por una red lógica en forma de árbol en la que no existe ningún bucle. Los puentes se comunican mediante mensajes de configuración llamados Bridge Protocol Data Unit (BPDU).

BPDU	**:** Bridge Protocol Data Unit, se envían cada 2 segundos.
Bridge ID	**:** Es igual a la Prioridad + Mac Adrress , La prioridad tiene un valor de 32.768
Root Bridge	**:** Es el switch principal en una red Spanning Tree (el que tenga el menor *Bridge ID*). Es el punto de referencia para los demas switches y desde donde se realizara el calculo de la topologia
Non Root Bridge	**:** Son los switches que no son *Root Bridge*
Root Port	**:** Es el puerto que se conecta directamente o el que tiene el menor costo hacia un *Root Bridge*.
Designated Port	**:** Es el *Root Port*
Nodesignated Port	**:** Son los puertos que no se conectan directamente al *Root Bridge*.
Forwarding Port	**:** Cuando es *Root Port*, *Designated Port*, se inicia el envió de tramas, cuando esta en *Forwarding Port*
Block Port	**:** En un puerto por donde no se envían ni se reciben tramas

Estado de los puertos.

Los estados en los que puede estar un puerto son los siguientes:

Bloqueo	En este estado sólo se pueden recibir BPDU's. Las tramas de datos se descartan y no se actualizan las tablas de direcciones MAC (mac-address-table).
Escucha	A este estado se llega desde Bloqueo. En este estado, los switches determinan si existe alguna otra ruta hacia el puente raíz. En el caso que la nueva ruta tenga un coste mayor, se vuelve al estado de Bloqueo. Las tramas de datos se descartan y no se actualizan las tablas ARP. Se procesan las BPDU.
Aprendizaje	A este estado se llega desde Escucha. Las tramas de datos se descartan pero ya se actualizan las tablas de direcciones MAC (aquí es donde se aprende por primera vez). Se procesan las BPDU.
Envío	A este estado se llega desde aprendizaje. Las tramas de datos se envían y se actualizan las tablas de direcciones MAC (mac-address-table). Se procesan las BPDU.
Desactivado	A este estado se llega desde cualquier otro. Se produce cuando un administrador deshabilita el puerto o éste falla. No se procesan las BPDU.



STP en la teoría.

El algoritmo consta de tres partes y requiere que cada conmutador tenga un ID y pueda saber el estado de cada puerto conectado a él. Las partes son:

- Se elige el bridge con el valor ID bajo (pequeño, menor) como el bridge root.
- Cada bridge calcula el camino más corto hacia el root y marca el puerto correspondiente como "root port".
- Para cada LAN todos los bridges conectados a él deben acordar cuál de ellos será el designado. Para hacerlo intercambian paquetes llamados BPDUs. El bridge designado será, en orden de preferencia:
  - El más cercano al root.
  - El más cercano y con menor ID en caso de que sea necesario desempatar.

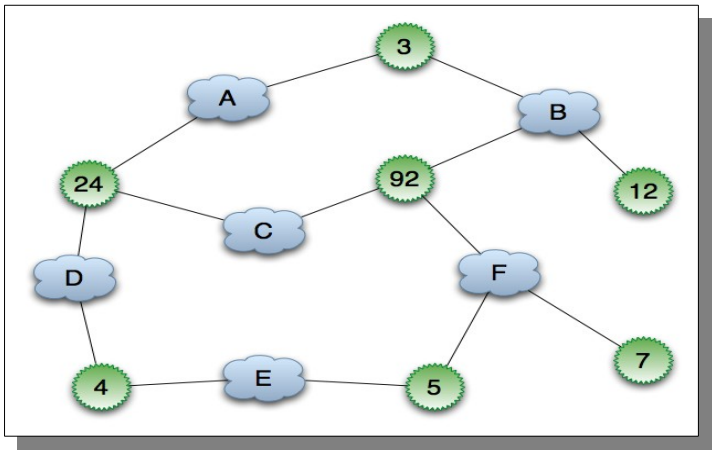
STP en la práctica.

Los bridges se sincronizan enviándose paquetes llamados BPDUs que contienen:

- ID del emisor.
- ID del que el emisor piensa que es el root.
- Distancia del emisor al root.

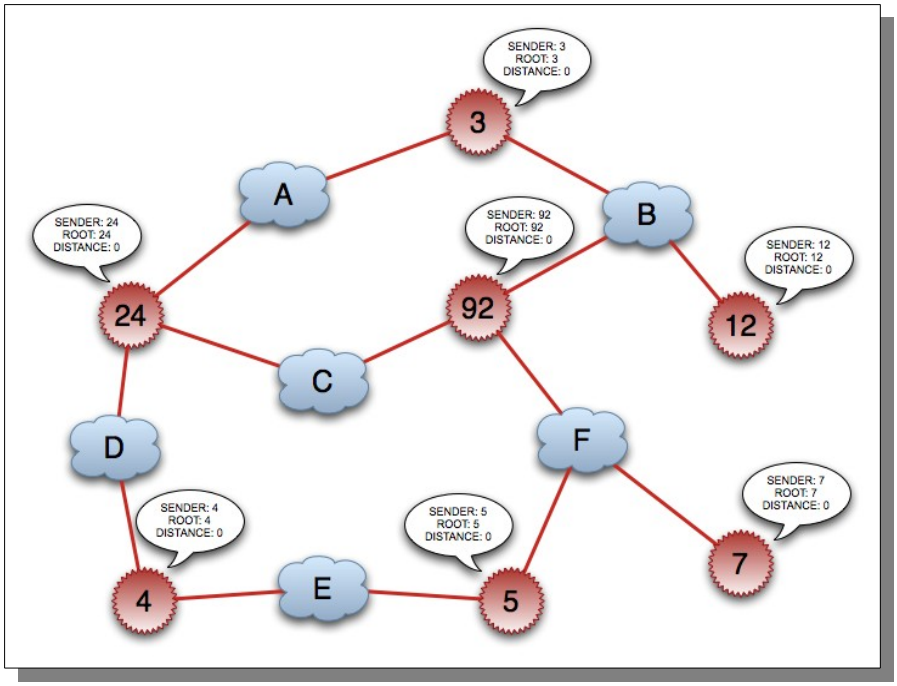
Ejemplo.

Para entender mejor como se dividen las partes del algoritmo y como termina convergiendo aplicaremos el algoritmo sobre la siguiente topología y veremos la reacción de los bridges.



Al principio todos los bridges se creen root, por lo tanto preparan **BPDU's** que dicen eso y lo envían por todos sus puertos para avisarle a los bridges adyacentes.

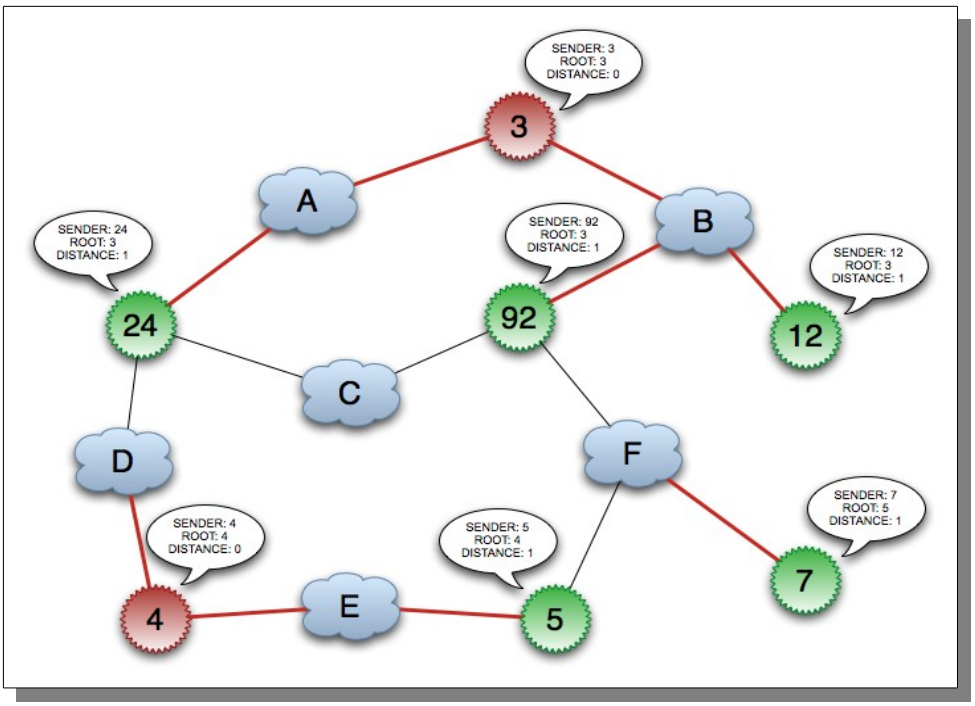
Los bridges root están marcados de color rojo y los puertos del root también.



En la "primera oleada" los paquetes de los bridges adyacentes llegan. Ahora, si un bridge considera que un paquete que llegó es mejor al suyo cambia su estado.

Un paquete es considerado mejor cuando:

- El root ID recibido es menor.
- El root ID es igual pero la distancia al root es menor.
- El root ID y la distancia son iguales pero el ID del emisor es menor.

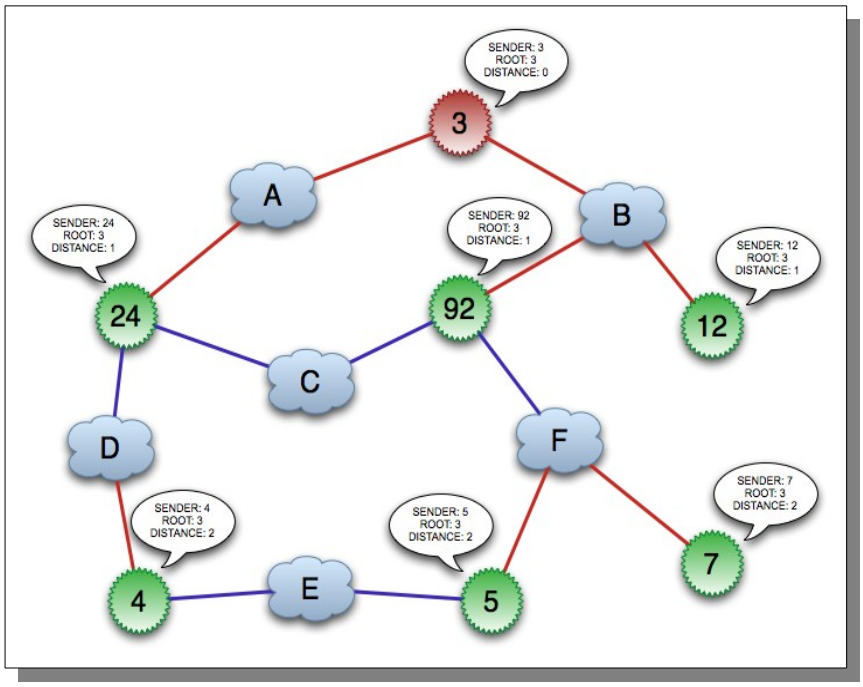


Ahora vemos como quedaría ahora el estado de cada bridge. Están marcados en rojo los "root ports" que son los puertos de cada bridge que apuntan al root. Para entender mejor los casos anteriormente mencionados vamos a ver dónde se dieron en nuestra red.

Los bridges 24, 92 y 12 recibieron un paquete del bridge número 3, vieron que existe un bridge con número menor estricto que ellos por lo tanto cambian su estado y forward los recibido por los puertos que no recibieron el paquete. Más adelante veremos como reaccionan los bridges de más abajo.

Los bridges de la segunda línea (los de más abajo) todavía no recibieron el paquete del bridge número 3, por lo tanto sólo mejoraron un poco lo conocido dentro de su alcance. El bridge 7 sabe que el 5 es mejor que él (e ignoró el paquete del 92 por ser mucho peor que él). El bridge número 4 todavía no se vió superado y se cree el rey del mundo.

En la próxima iteración los paquetes del bridge 3 llegan a la línea de abajo y se estabiliza la sección del root.



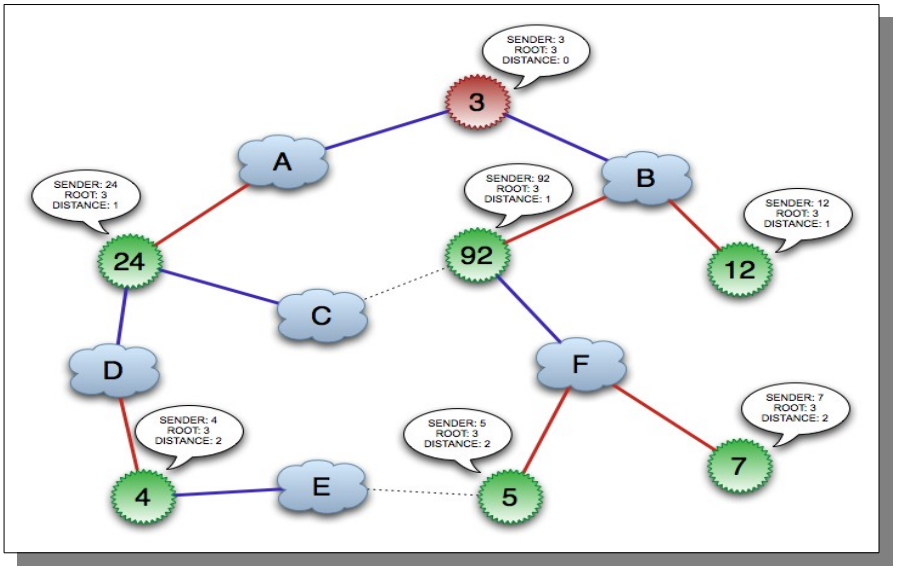
Es importante ver que durante esta segunda iteración no produjeron paquetes nuevos los bridges que cambiaron su estado durante la primera iteración (de la primera línea) sino que forwardearon el paquete proveniente del bridge número 3. Con esto me refiero a que, por ejemplo, el bridge número 12 no generó un paquete y lo envió por su root port. Si lo hubiera hecho habríamos tenido problemas graves de convergencia en cuanto a los puertos designados que es lo que vamos a ver ahora.

Los puertos designados son los que cada bridge cree que tiene que forwardear. El próximo paso es bloquear algunos de estos puertos para convertir el grafo en un árbol.

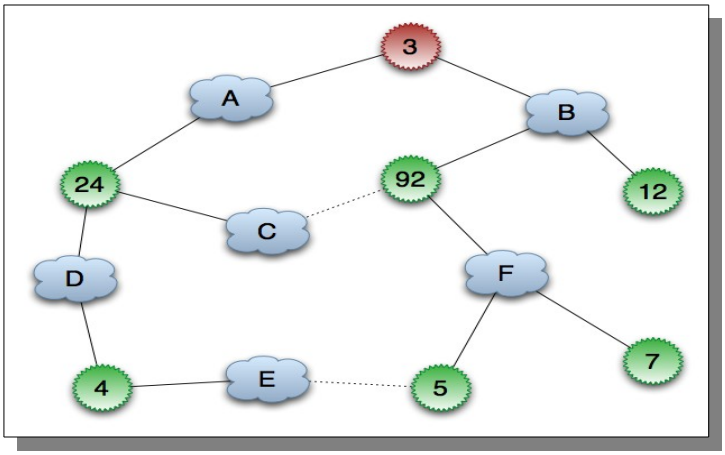
Los bridges envían paquetes hacia las LANs (no por los root ports sino por los designados) y escuchan a los demás bridges.

Si reciben un root ID igual pero una distancia menor al root bloquean este puerto y dejan que el otro bridge se encargue de mandar los mensajes desde y hacia esa LAN. En este paso ya no debería haber paquetes circulando con root ID menor al mío (porque dijimos que ya habíamos acordado un root). Sin embargo si ese fuera el caso simplemente se procede como antes y en la próxima iteración quedaremos listos para sincronizar los puertos designados.
Si reciben un root ID y una misma distancia se desempata con el menor ID del emisor.

En la siguiente figura podemos ver como quedan asignados los puertos:



Finalmente al converger el algoritmo se tiene un árbol y todos los bridges están sincronizados. Cada un intervalo de tiempo se generan paquetes con información para mantener la asignación y estar atentos a los cambios.



¿Cuándo utilizar el STP?

Este protocolo se utiliza cuando nos encontramos en redes con topologías redundantes. Es en este tipo de redes que se pueden producir bucles (loops o lazos), los cuales pueden producir los siguientes problemas:

- Tormentas o Inundaciones de Broadcast o Multicast. Los Broadcast en la red son enviados una y otra vez circulando sin fin en la misma, dado que en Ethernet no existe como en IP un campo TTL. Lógicamente al no eliminarse la situación se agrava con cada nuevo Broadcast.
- Copias múltiples de Tramas. Con la redundancia es muy probable que un host reciba una trama repetida, dado que la trama podría llegar por dos enlaces diferentes.
- Inestabilidad o Inconsistencia de Tablas de Direcciones MAC (Media Access Control, Control de Acceso al Medio). Una trama que proviene de una dirección MAC en particular podría llegar desde enlaces diferentes.

Como efecto de estos problemas se reduce el tráfico de usuario y la red parecerá estar inactiva o extremadamente lenta.

Entonces alguien dirá, si este tipo de topologías puede acarrear tantos problemas, ¿por qué utilizarlas?, ¿por qué mejor no utilizar otra topología de red y así evitar estos problemas sin necesidad de utilizar el STP?; todo tiene su razón de ser y estas preguntas tienen sus respuestas.

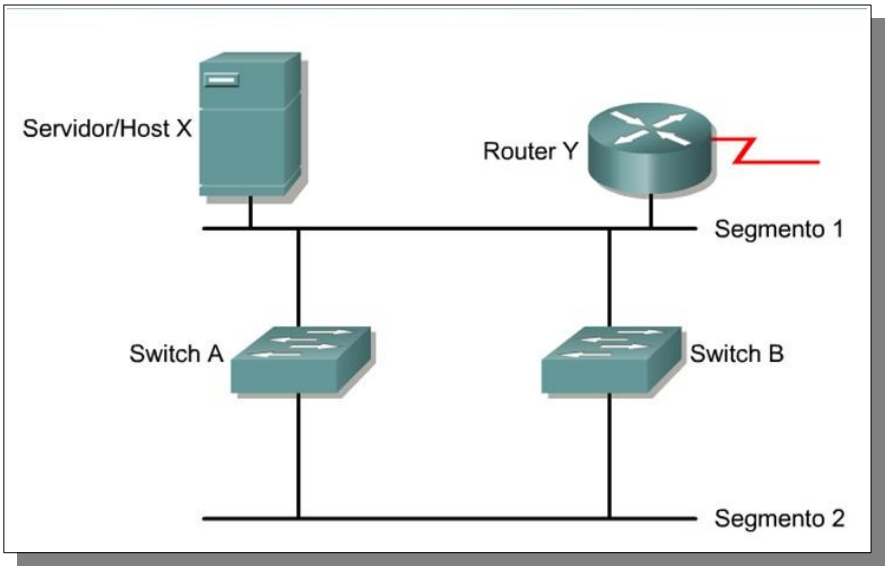
Las empresas y organizaciones requieren cada vez más de redes de computadoras para realizar su trabajo, pero además necesitan disponibilidad de la red con un tiempo de actividad continuo. 100 % de tiempo de actividad es imposible, pero se puede llevar a lo más alto; un 99,999 % por ejemplo, que equivale a 1 hora inactiva cada 4000 días o aproximadamente 5,25 minutos de inactividad por año.

Debido a esto los administradores de redes ven la necesidad del uso de estas topologías. Ya mencionamos las desventajas o problemas que acarrean las mismas, ahora resumamos algunas ventajas:

Se tiene una mayor disponibilidad de la red, es decir permiten más tiempo de actividad de la red.
Protección contra el tiempo de inactividad o no disponibilidad. El tiempo de inactividad puede deberse a la falla de un solo enlace, puerto o dispositivo.

- Son tolerantes a fallas, ya que eliminan únicos puntos de fallo
- Pueden asumir tareas ejecutas por rutas o dispositivos que fallan.

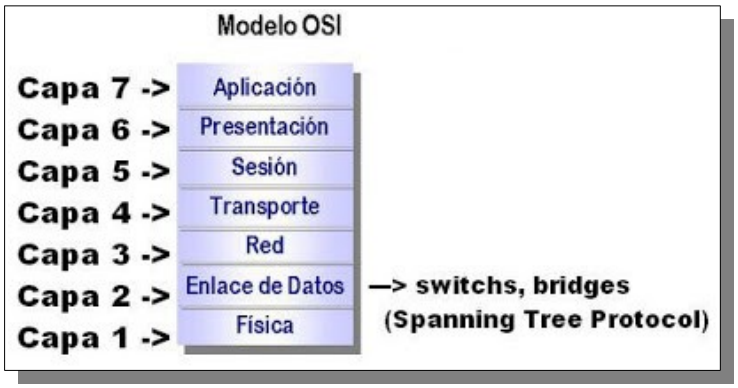
Para aclarar la situación tomemos un ejemplo más gráfico como la red de la siguiente figura. Si el Switch A falla, el tráfico puede continuar fluyendo desde el segmento 2 al segmento 1 y al Router a través del Swtich B.



Por lo que podríamos resumir que la confiabilidad se logra con equipos y diseños de red confiables, tolerantes a fallos (redes con topologías redundantes). Y podríamos utilizar en este contexto la palabra redundancia como sinónimo de confiabilidad.

¿Dónde se utiliza STP?

Como dice el dicho una imagen vale más que mil palabras:



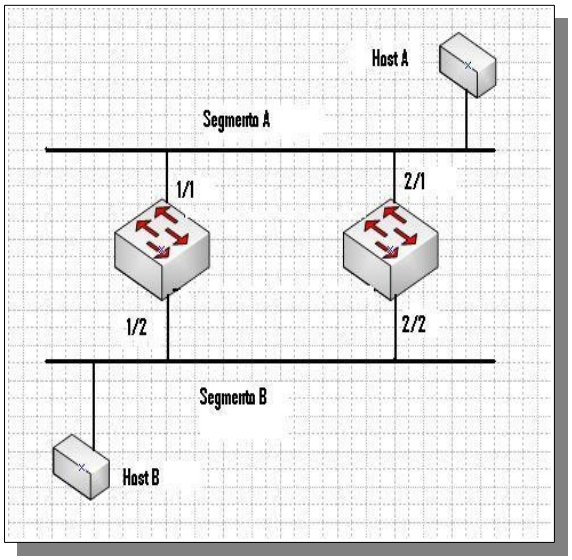
Este protocolo se utiliza solo en dispositivos de nivel 2 de la capa OSI (Open System Interconnection, Interconexión de Sistemas Abiertos), ósea en la capa de enlace de datos; es decir en los switchs y en los bridges (puentes). Debido a que en ésta capa no existe el campo TTL (Time To Live, Tiempo de Vida) que existe en la capa 3 (nivel de red) y que da un tiempo de vida a los paquetes para evitar este tipo de problemas.

¿Por qué se utiliza STP?

Se utiliza precisamente porque las redes redundantes producen bucles que acarrean los problemas mencionados anteriormente. Por lo que STP es el encargado de evitar y eliminar bucles en la red, dejandonos un path loop-free (camino sin bucles), de esta forma, STP asegura que sólo existe un camino a cada destino. En el caso en que un link (enlace) falle, el dispositivo pasa la interfaz del estado bloqueado, a estado forwarding (enviando), operativo.

Ahora tomemos como ejemplo la red de la siguiente figura. Nuestro Host A tiene dos posibles caminos hacia el

Host B; considerando un escenario en el cual, el Host A envía tráfico al Host B, pero ninguno de los switchs ha aprendido la MAC del Host B.



El Host A, transmite una trama al segmento A. Ambos switches reciben dicha trama tanto en el puerto 1/1 como en el 2/1, los switches actualizan sus tablas de direcciones situando al Host A en los puertos mencionados.

Ambos switches hacen un forward de la trama al Segmento B. No solo el Host B recibirá dicha trama dos veces (copias Múltiples de tramas), ambos switches recibirán la trama del otro switch también (escenario redundante).

Porque una de las características básicas del bridging transparente es la de escuchar la MAC origen para aprender el puerto correcto para usar con esa dirección, cada switch reaprenderá que el Host A reside en los puertos 1/2 y 2/2 respectivamente. Los switches asumen incorrectamente que todas las tramas para el Host A deben ser transmitadas al Segmento B (Inestabilidad en la Tabla de Direcciones MAC).

Los switches hacen un forward de la trama de nuevo al Segmento A, donde fue originada, por lo que aquí tenemos el bucle que debemos evitar. Si en vez de una simple trama nuestro Host A hubiera enviado un Broadcast, el problema sería aún mayor (Tormenta de Broadcast). Para colmo, el TTL no sirve para nada ya que estamos en capa 2, y TTL es de capa 3.

Ante la necesidad de tener una red LAN (Local Area Network, Red de Área Local) redundante y dinámica libre de los problemas asociados a la redundancia resulta evidente que es imperioso un protocolo que sea capaz de resolver estas cuestiones. Es aquí donde entra en acción el Protocolo Spanning Tree (STP).

STP ejecuta un algoritmo llamado STA (Spanning Tree Algorithm, Algoritmo de Spanning Tree). Para encontrar links redundantes, STA escoge un punto de referencia en la red, y localiza los caminos redundantes hacia ese punto (Spanning Tree Root, Raíz del Árbol de Expansión). Si STA encuentra un camino redundante, selecciona un solo camino hacia el root, y bloquea el resto. El puerto bloqueado sigue recibiendo BPDU (Bridge Protocol Data Units, Unidad de Datos de Protocolo Puente), ya que en caso de que el otro puerto falle, este entra en funcionamiento.

Cost

By calculating and assigning the port cost of the switch ports, you can ensure that the shortest (lowest cost) distance to the root switch is used to transmit data. You can calculate and assign lower path cost values (port costs) to higher bandwidth ports by using either the short method (which is the default) or the long method. The short method uses a 16-bit format that yields values from 1-65535. The long method uses a 32-bit format that

yields values from 1-200,000,000. For more information on setting the default cost mode, see the "Configuring the PVST+ Default Port Cost Mode" section.

Calculating the Port Cost Using the Short Method

The IEEE 802.1D specification assigns 16-bit (short) default port cost values to each port that is based on bandwidth. You can also manually assign port costs between 1-65535. The 16-bit values are only used for ports that have not been specifically configured for port cost. Table 7-1 shows the default port cost values that are assigned by the switch for each type of port when you use the short method to calculate the port cost.

Velocidad del Puerto	Default Cost Value	Rango
10 Mbps	100	1 a 65535
100 Mbps	19	1 a 65535
1 Gbps	4	1 a 65535

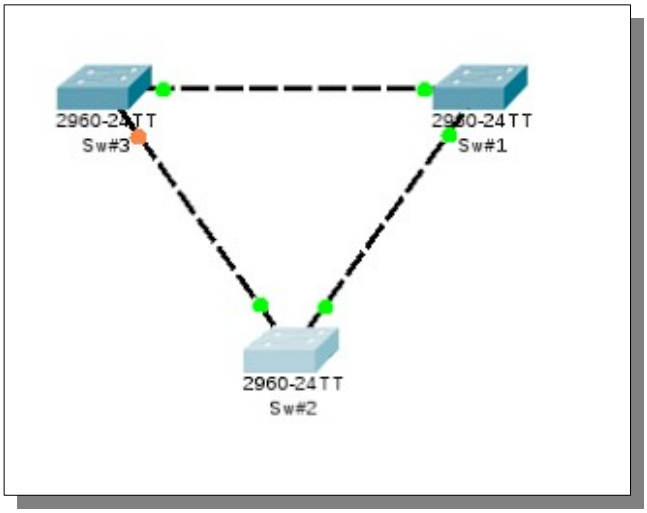
Configurando Spanning-tree

La primera forma consiste en asegurar la asignación de la prioridad mediante el uso del macro “root primary” es decir que el switch configura la Bridge ID Priority de manera automática, si en una topología en la que todos los equipos tengan por defecto el valor de 32768, el switch seleccionara el valor de 24576 dejando el valor siguiente de 28672 para el switch que se configure con la macro para definirlo como secundario. “root secondary”, esto garantiza que ambos switches contenga un Bridge ID menor que el resto de equipos.

Los comandos necesarios para realizar la configuración serían los siguientes:

sw1 enable
sw1 # show spanning-tree
sw1# configure terminal
sw1(config)# spanning-tree vlan 1 root primary

sw2 enable
sw2 # show spanning-tree
sw2# configure terminal
sw2(config)# spanning-tree vlan 1 root secondary



Para verificar la configuración realizada en ambos equipos basta con ejecutar el comando show spanning-tree. Principalmente porque acá es fácil ver lo que indica el gráficamente packet tracer.

sw#1#sh spanning-tree
VLAN0001
Spanning tree enabled protocol ieee
Root ID Priority 24577
Address 000B.BE09.C457
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 000B.BE09.C457
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Fa0/1 Desg FWD 19 128.1 P2p
Fa0/2 Desg FWD 19 128.2 P2p

Sw2#sh spanning-tree
VLAN0001
Spanning tree enabled protocol ieee
Root ID Priority 24577
Address 000B.BE09.C457
Cost 19
Port 1(FastEthernet0/1)
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID Priority 28673 (priority 28672 sys-id-ext 1)
Address 0009.7CDB.B952
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Fa0/1 Root FWD 19 128.1 P2p
Fa0/2 Desg LSN 19 128.2 P2p

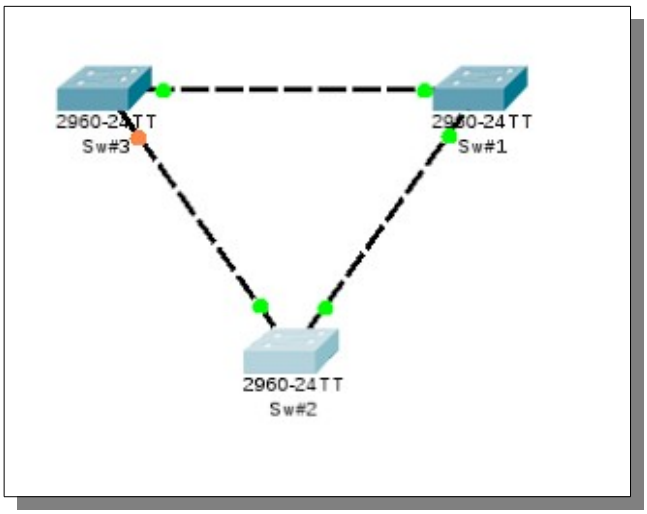
La segunda forma de definir el Root Bridge consiste en definir manualmente la prioridad del switch, mediante el comando “spanning-tree vlan id de la VLAN priority valor” para el ejemplo de la vlan 1 el comando seria el siguiente: “spanning-tree vlan 1 priority 4096” en este ejemplo utilizo el valor 4096, sin embargo bien puede usarse también el valor 0.

Con spanning-tree es posible definir rangos de vlan o bien vlans que no sean continuas, de la siguiente manera: 1,3-5,7,9-11 en este caso spanning-tree funciona para las vlan 1,3,4,5,6,7,9,10 y 11.

Los comandos necesarios para realizar la configuración serian los siguientes:

sw1# configure terminal
sw1(config)# spanning-tree vlan 1-2 priority 0

sw2# configure terminal
sw2(config)# spanning-tree vlan 1-2 priority 4096



Para verificar la configuración realizada en ambos equipos basta con ejecutar el comando show spanning-tree. En ambos comando es fácil identificar el valor del Root bridge y el valor local del switch, en el caso del sw#1 ambos valores serán iguales, pues este es el Root bridge, mientras que en el caso de los otros switches indicaran el valor del Root bridge y el valor local.

sw#1#sh spanning-tree
VLAN0001
Spanning tree enabled protocol ieee
Root ID Priority 1
Address 000B.BE09.C457
This bridge is the root
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID Priority 1 (priority 0 sys-id-ext 1)
Address 000B.BE09.C457
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
---------- ---- --- --------- -------- --------------------------------
Fa0/1 Desg FWD 19 128.1 P2p
Fa0/2 Desg FWD 19 128.2 P2p

Sw#2#sh spanning-tree
VLAN0001
Spanning tree enabled protocol ieee
Root ID Priority 1
Address 000B.BE09.C457
Cost 19
Port 1(FastEthernet0/1)
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID Priority 4097 (priority 4096 sys-id-ext 1)
Address 0009.7CDB.B952
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
---------- ---- --- --------- -------- --------------------------------
Fa0/1 Root FWD 19 128.1 P2p
Fa0/2 Desg FWD 19 128.2 P2p

Consultar sobre Spanning Tree rapid PVST (Per VLAN Spanning Tree)

Para usar Spanning Tree rapid VST debemos :

1. Activar Spanning Tree rapid VST dentro del modo configure terminal, usando la siguiente instrucción :
   Switch (config) # spanning-tree mode rapid-pvst
2. Compruébalo adicionando nuevo switch
3. Que pasa cuando adiciones un computador?

Basado en:
http://jedicerocool.blogspot.com.co/2010/04/cuando-donde-y-por-que-se-utiliza.html#.VkddjJf7tC1

Glosario

BPDU	**:** Bridge Protocol Data Unit, se envían cada 2 segundos.
Bridge ID	**:** Es igual a la Prioridad + Mac Adrress , La prioridad tiene un valor de 32.768
Root Bridge	**:** Es el switch principal en una red Spanning Tree (el que tenga el menor Bridge ID). Es el punto de referencia para todos los switches y desde donde se realizara el calculo de la topologia
Non Root Bridge	**:** Son los switches que no son *Root Bridge*
Root Port	**:** Es el puerto que se conecta directamente o el que tiene el menor costo hacia un *Root Bridge*.
Designated Port	**:** Es el Root Port
Nodesignated Port	**:** Son los puertos que no se conectan directamente al Root Bridge.
Forwarding Port	**:** Cuando es Root Port, Designated Port, se inicia el envió de tramas, cuando esta en Forwarding Port
Block Port	**:** En un puerto por donde no se envian ni se reciben tramas
Modo Bridge	**:** En modo bridge el router simplemente hace de pasarela y el trasiego de datos con Internet es tramitado directamente por el ordenador. Es decir, la función router queda anulada y el aparato no hace NAT (Network Address Translation = traducción de direcciones de red).

En modo bridge no hay que gestionar (abrir puertos) porque todo "pasa" sin restricciones. Por ese motivo se recomienda siempre que cuando el router está en este modo de operación se instale un firewall en el ordenador.