

# Configuración de STP para controlar la topología

Guillermo Pérez Trabado ©2016-2022

## Diseño de Infraestructuras de Redes

Dept. de Arquitectura de Computadores - Universidad de Málaga

### Escenario

En este documento vamos a aprender a configurar el Spanning Tree Protocol (STP) de los switches para controlar la forma en que se genera el **árbol de expansión mínima** del grafo formado por los switches de la red de un site.

Antes de empezar recordaremos qué es el Spanning Tree de una red de switches y cómo funciona el protocolo STP:

#### *Spanning Tree*

El enrutamiento usado por los switches Ethernet tiene dos propiedades que tenemos que destacar:

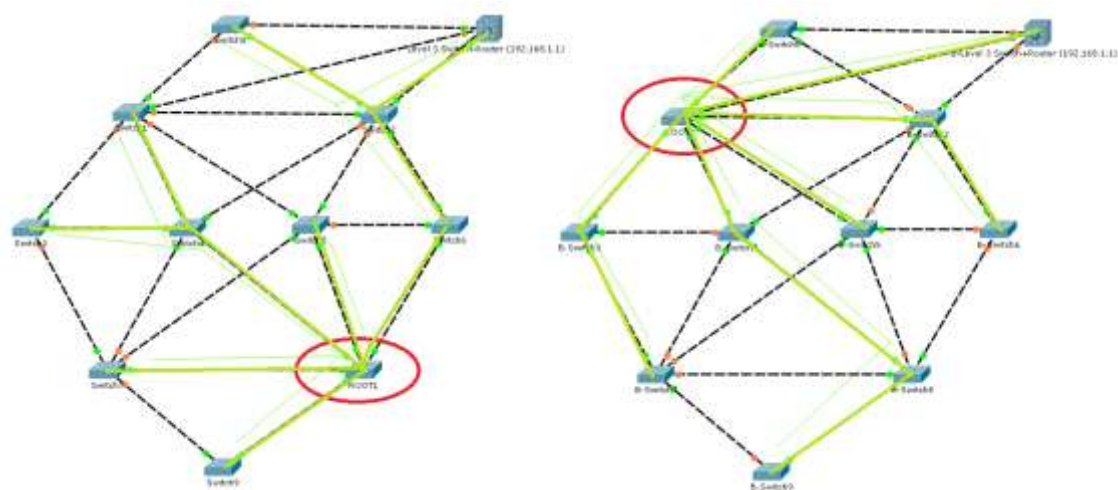
- Cuando un switch no conoce la ruta a un destino usa la inundación de la red emitiendo copias de un paquete por todos sus enlaces de salida (broadcast). Cada switch asume que el enlace por el que se recibe un paquete en cada momento es la ruta más corta a la MAC de origen de dicho paquete.
- Ethernet no incluyen ningún tipo de mecanismo de control de bucles, por lo que los switches asumen que la topología de la red **nunca debe contener un ciclo**.

Es decir, el encaminamiento por broadcast y el aprendizaje que la red es un árbol y nunca un grafo con ciclos. Por tanto, antes de permitir el enrutamiento, los switches deben garantizar que la red tiene topología de árbol desactivando los enlaces sobrantes que generarían ciclos.

#### *Minimum Spanning Tree*

El cálculo de los enlaces a desactivar no es trivial ya que se pretende que el árbol generado tenga sea óptimo desde el punto de vista del encaminamiento. El árbol buscado se denomina **minimum spanning tree (MST)**, que es el subgrafo acíclico con el menor coste global de caminos posible. Para un mismo grafo, es posible calcular más de un MST con el mismo coste global. En la figura siguiente puedes ver dos MST diferentes (en verde) calculados para el mismo grafo. Observa que:

- Cada uno de ellos tiene la raíz del MST en un nodo distinto del grafo original.
- El MST resultante es un árbol que en cada nivel prioriza la amplitud frente a la profundidad. Eso quiere decir que en caso de existir un enlace entre un nodo y la raíz, se prioriza elegir dicho camino directo en lugar de colgar de otro nodo ya que esto minimiza la distancia entre cualquier par de nodos medida en número de saltos (**hops**).



## 1. Spanning Tree Protocol

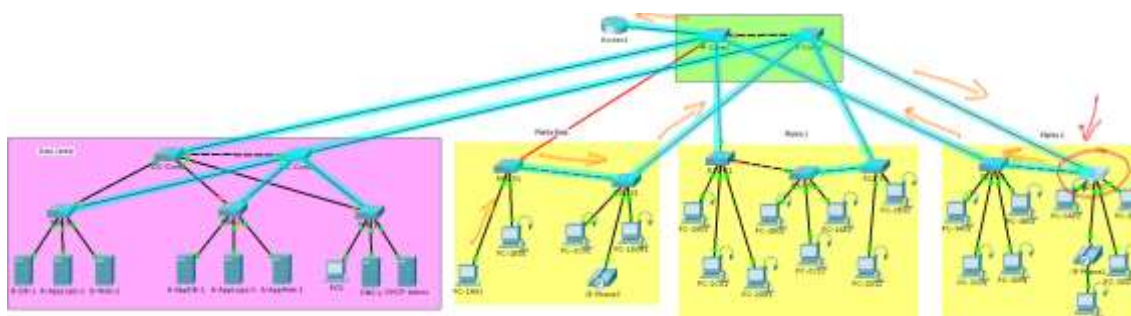
El MST de la red es construido entre todos los switches del grado mediante el protocolo STP. En este documento no vamos a tratar cómo funciona, pero sí cómo se puede controlar su funcionamiento de forma práctica para optimizar nuestro diseño de red.

Uno de los principales detalles a saber de STP es que los switches tienen que comenzar eligiendo qué switch va a ser la raíz del MST. El STP define una prioridad para cada switch de forma determinista. El switch con el valor más reducido es elegido como **root** de un grupo de switches. STP usa en cada switch dos campos para componer un valor numérico de la prioridad y conseguir que sea único:

- Un campo de prioridad de 16 bits que por defecto es la misma en todos los switches.
- La MAC del switch.

Eso quiere decir que si no se configura STP, el switch con la MAC más antigua será siempre elegido como root. Por un lado, eso nos proporciona una negociación automática de STP que funciona sin configuración alguna. Pero por otro, puede depararnos algunas sorpresas desagradables en nuestras configuraciones.

El caso del grafo del apartado anterior no es un ejemplo significativo de la estructura de red de un edificio con dos niveles de repartidores y un data center. En un edificio típico, la red es un árbol con dos raíces y dos caminos redundantes entre los switches de cada repartidor secundario y los core switches. El siguiente esquema muestra el MST generado para dicho árbol con doble raíz. Sin embargo, el administrador no ha configurado la prioridad de los switches y por tanto el **STP root** elegido es aquel con la MAC más baja (marcado por un círculo rojo en la figura). El MST de la red construido por STP puede verse dibujado en azul:



Se pueden observar que los enlaces con los LED en naranja han sido desactivados por STP para eliminar los ciclos del grafo original formado por los enlaces redundantes. El administrador, de forma intuitiva considera que la raíz de la red debería ser el Core-1 ya que ahí está conectado el router LAN que interconecta todas las VLANs entre sí. Todos los terminales deberían estar a la mínima distancia posible de dicho router.

Sin embargo, nos puede sorprender lo ineficiente de la topología negociada (en azul), ya que el MST calculado se ha optimizado para reducir la distancia de todos los nodos a switch raíz (que no es Core-1 como podríamos creer al ver el mapa de la red). La ruta marcada por las flechas requiere **7 hops** (saltos Ethernet) para llegar al router cuando podría llegarse en solo **3 hops**.

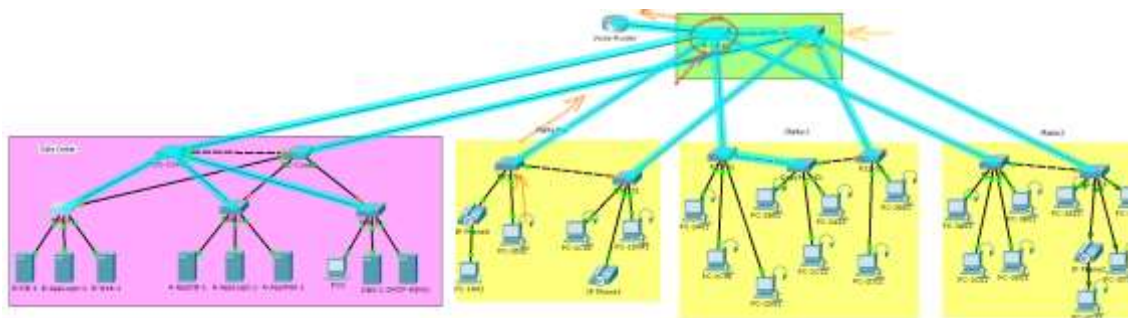
Por tanto, uno de los aspectos más importantes a configurar de una red Ethernet es la prioridad de los switches core, para forzar su elección como raíces del MST. Para ello usaremos el comando de configuración siguiente **solo en el switch que deseamos como primario**:

```
Core1(config)#spanning-tree vlan 1-1023 root primary
```

En caso de avería del core primario tenemos que garantizar que le reemplaza el core secundario. Por tanto, es necesario configurar este segundo switch con una prioridad ligeramente más baja para que se convierta en el nuevo root. Este comando **solo se ha de ejecutar en el secundario**:

```
Core2(config)#spanning-tree vlan 1-1023 root secondary
```

En la figura siguiente puede verse que ahora el STP root es el Core1 y que el MST generado está optimizado para que todos los terminales estén a la mínima distancia de dicho nodo, y por tanto también del router LAN. Ahora vemos que el camino para llegar desde la mayor parte de terminales hasta el router es de 3 hops y el camino más largo es de 4 hops. Una ventaja de elegir bien el root es que el RTT es uniforme dentro de la red.



Hay que detallar que en realidad los switches de Cisco usan por defecto una versión mejorada de STP llamada PVSTP (per VLAN STP), que negocia un árbol MST distinto (y un root distinto) por cada VLAN. Normalmente el core switch principal debería ser la raíz del MST de todas las VLANs, por lo que el comando especifica la prioridad para todas las VLANs posibles a la vez. Sin embargo, en algunas situaciones puede ser interesante balancear el tráfico de distintas VLANs entre distintos enlaces troncales. Esto se consigue definiendo un root switch diferente para cada grupo de VLANs. Por ejemplo:

```
Core1(config)#spanning-tree vlan 10-15,100,102-199 root primary
Core1(config)#spanning-tree vlan 16-29,101,200-210 root secondary
y
```

```
Core2(config)#spanning-tree vlan 10-15,100,102-199 root secondary
Core2(config)#spanning-tree vlan 16-29,101,200-210 root primary
```

Con esta configuración, por ejemplo, el switch intermedio del segundo repartido enviaría el tráfico de las VLANs 100 y 102 por el trunk hacia el switch de su izquierda y el tráfico de las VLANs 101 y 200 por el enlace hacia el switch de su derecha, aprovechando la capacidad de ambos enlaces.

## 2. Rapid per VLAN Spanning Tree Protocol (RPVST)

STP implementa un algoritmo distribuido en el que un grupo de switches resuelven un problema de teoría de grafos obteniendo información de sus vecinos y enviando la información que han averiguado. De esa forma, el conocimiento sobre la topología del grafo se va distribuyendo a todos los nodos hasta que todos conocen la misma información (convergencia de la red). El problema del STP es que el intercambio de información tiene lugar solo cada vez que un temporizador lo indica cada 15s. El número de intercambios necesario para que la información se propague depende del diámetro del grafo (la distancia máxima entre dos nodos). El resultado es que la negociación es bastante lenta.

Para acelerar el tiempo de convergencia de STP se desarrolló una variante conocida como Rapid STP (y también Rapid PVSTP) que mejora la implementación del algoritmo convergiendo mucho más rápido. Para usarlo, en **todos los switches** hemos de cambiar el modo de funcionamiento de STP con el comando:

```
(config)#spanning-tree mode rapid-pvst
```

Otra aceleración posible de la convergencia es evitar que un switch de acceso intente negociar STP o RSTP en los puertos donde se conectan terminales. Por defecto los switches lo hacen por precaución en prevención de que puedan encontrar otro switch conectado a un puerto de acceso. El siguiente comando deshabilita por defecto la negociación STP en aquellos puertos que están configurados en modo **access**:

```
(config)#spanning-tree portfast default
```

Este último comando solo debe usarse en los **switches de acceso** que tienen puertos con terminales.

## 3. Monitorización de STP

Aunque son raros los problemas de configuración de STP, no sería extraño que queramos examinar el switch raíz de cada VLAN y otros detalles en caso de sospechar que hay problemas de rendimiento que podrían ser debidos a la topología MST.

Los comandos de configuración muestran el estado desglosado para cada VLAN definida en el switch. Algunos de los datos más interesantes son las direcciones MAC del Root ID (root switch de una VLAN) y del Bridge ID (descripción del switch local). Si son idénticas, como en este ejemplo, estamos viendo la configuración del **root** para una VLAN. Además, el propio comando indica en su salida que estamos en el root switch:

```
P-Core1#sh spanning-tree
VLAN0001
  Spanning tree enabled protocol rstp
    Root ID    Priority      24577
              Address       0030.F2C4.20D9
              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 0030.F2C4.20D9
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
-----
Gi1/1 Desg FWD 4 128.2 Shr
Gi8/1 Desg FWD 4 128.9 P2p
Gi9/1 Desg FWD 4 128.10 P2p
Gi0/1 Desg FWD 4 128.1 Shr
Gi2/1 Desg FWD 4 128.3 Shr

```

Si son diferentes, como en este otro ejemplo, estamos en un nodo distinto al root:

```

R1-S1#sh spanning-tree
VLAN0001
Spanning tree enabled protocol rstp
Root ID Priority 24577
Address 0030.F2C4.20D9
Cost 4
Port 1(GigabitEthernet0/1)
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID Priority 32769 (priority 32768 sys-id-ext 1)
Address 0002.1687.E375
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
-----
Gi0/1 Root FWD 4 128.1 Shr
Gi2/1 Desg FWD 19 128.3 P2p
Gi3/1 Desg FWD 19 128.4 P2p
Gi9/1 Desg FWD 4 128.10 P2p

```