

# LAN switching: Spanning Tree Protocol

Autor: Ing. Nicolás Matsunaga

Facultad de Ingeniería

1er cuatrimestre 2014

# Resumen

- 1 Spanning Tree Protocol
- 2 Variantes de STP
- 3 Preguntas

# El problema de los links redundantes

Ante un trama broadcast un switch debe

# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

Ante un trama con destino desconocido un switch debe

# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

Ante un trama con destino desconocido un switch debe copiar ésta en todos sus puertos (flooding). (puede haber CAM table corruption)

# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

Ante un trama con destino desconocido un switch debe copiar ésta en todos sus puertos (flooding). (puede haber CAM table corruption)

En capa 2 de Ethernet no existe mecanismo para mitigar *loops* (sí en IP,

# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

Ante un trama con destino desconocido un switch debe copiar ésta en todos sus puertos (flooding). (puede haber CAM table corruption)

En capa 2 de Ethernet no existe mecanismo para mitigar *loops* (sí en IP, TTL)



# El problema de los links redundantes

Ante un trama broadcast un switch debe copiar ésta en todos sus puertos

Ante un trama con destino desconocido un switch debe copiar ésta en todos sus puertos (flooding). (puede haber CAM table corruption)

En capa 2 de Ethernet no existe mecanismo para mitigar *loops* (sí en IP, TTL)

Como consecuencia, si existen loops se pueden producir **broadcast storms** o tormentas de paquetes

# Solución: Spanning Tree Protocol

Es un protocolo abierto definido por la IEEE y es la norma IEEE 802.1d (download gratis)

Su función es la de *evitar* loops y proveer caminos *redundantes*

Para ello convierte una topología mallada en un árbol (span a tree). Único camino entre cualesquiera dos nodos

Como consecuencia algunos puertos entran en *blocking state* y no participan del delivery de tramas. Los que sí participan están en *forwarding state*

STP es *transparente* para las estaciones de trabajo

# STP (1) - Cómo funciona

Cada switch debe ejecutar el *Spanning Tree Algorithm*

Los switches intercambian información a través de tramas llamadas **BPDU** (Bridge Protocol Data Unit)

Con esta información intercambiada deciden qué puertos poner en *blocking* o *forwarding*

## STP (2) - Spanning Tree Algorithm

Requisitos del STA:

- 1 configurar una topología *activa* de una *Bridged LAN* arbitraria en un **solo árbol**
- 2 proveer **redundancia**, reconfiguración automática del árbol ( $\neq$  loops temporales)
- 3 la topología activa deberá **estabilizarse** en una *Bridged LAN* de cualquier tamaño (alta probabilidad, plazo corto, conocido y acotado)
- 4 la topología activa será **predecible** y **reproducibile** y permitirá ser **elegida** por los administradores a través de parámetros
- 5 operación **transparente** a los terminales
- 6 BW para establecer y mantener el árbol (pequeño porcentaje e independiente del número de Bridges)

## STP (2) - Spanning Tree Algorithm

Requisitos del STA (cont.):

- 7 **requisitos de memoria** asociados a cada port deben ser **independientes** de la cantidad de Bridges o LANs (red de medio compartido  $\neq$  Bridged LAN)
- 8 los Bridges **no deben tener que ser configurados individualmente** antes de agregarse a la Bridged LAN, excepto la configuración de la MAC Address si así lo requiriese

## STP (3) - Spanning Tree Algorithm - Funcionamiento

El protocolo STP podría resumirse en las siguientes reglas:

- Ante falta de info, cada bridge se asignará a si mismo como Root Bridge

## STP (3) - Spanning Tree Algorithm - Funcionamiento

El protocolo STP podría resumirse en las siguientes reglas:

- Ante falta de info, cada bridge se asignará a si mismo como Root Bridge
- El Root Bridge producirá BPDUs cada *Hello Time* segundos

## STP (3) - Spanning Tree Algorithm - Funcionamiento

El protocolo STP podría resumirse en las siguientes reglas:

- Ante falta de info, cada bridge se asignará a si mismo como Root Bridge
- El Root Bridge producirá BPDUs cada *Hello Time* segundos
- Las BPDUs competirán para determinar su superioridad



## STP (3) - Spanning Tree Algorithm - Funcionamiento

El protocolo STP podría resumirse en las siguientes reglas:

- Ante falta de info, cada bridge se asignará a si mismo como Root Bridge
- El Root Bridge producirá BPDUs cada *Hello Time* segundos
- Las BPDUs competirán para determinar su superioridad
- Una  $BPDU_A$  es superior a  $BPDU_B$  :

si el Root Bridge  $ID_A < \text{Root Bridge } ID_B$

si  $\Rightarrow$  si Root Path Cost $_A < \text{Root Path Cost}_B$

si  $\Rightarrow$  si Bridge ID $_A < \text{Bridge ID}_B$

si  $\Rightarrow$  si Port ID $_A < \text{Port ID}_B$

## STP (3) - Spanning Tree Algorithm - Funcionamiento

El protocolo STP podría resumirse en las siguientes reglas:

- Ante falta de info, cada bridge se asignará a si mismo como Root Bridge
- El Root Bridge producirá BPDUs cada *Hello Time* segundos
- Las BPDUs competirán para determinar su superioridad
- Una  $BPDU_A$  es superior a  $BPDU_B$  :

si el Root Bridge  $ID_A < \text{Root Bridge } ID_B$

si  $\Rightarrow$  si Root Path Cost $_A < \text{Root Path Cost}_B$

si  $\Rightarrow$  si Bridge  $ID_A < \text{Bridge } ID_B$

si  $\Rightarrow$  si Port  $ID_A < \text{Port } ID_B$

- Aquellos bridges que no sean Root deberán reaccionar ante la recepción de una BPDU

## STP (4) - Spanning Tree Algorithm - BPDU

La BPDU se encapsula en 802.2 LLC (SSAP y DSAP 0x42, ETH dst 01:80:C2:00:00:00) y hay de 2 tipos:

- Topology Change Notification
- Configuration, que contiene los campos:
  - **Root ID** 8 Bytes = 2B prio + 6 MAC Address
  - **Root Path Cost** 4 Bytes
  - **Bridge ID** 8 Bytes = 2B prio + 6 MAC Address (propia)
  - **Port ID** 2B = 1B prio + 1 Port number (o 4b prio + 12b Port number 802.1t)
  - Message Age: antigüedad del mensaje
  - Max Age: (conf) máxima antigüedad (def: 20, 6-40)
  - Hello Time: (conf) tiempo entre BPDUs (def: 2, 1-10)
  - Forward Delay: (conf) tiempo en estados transitorios - *Learning, Listening* (def: 15, 4-30)

# STP (5) - Spanning Tree Algorithm - BPDU 2

## Path Cost

Bandwidth	Revised STP Cost
4 Mbps	250
10 Mbps	100
16 Mbps	62
45 Mbps	39
100 Mbps	19
155 Mbps	14
622 Mbps	6
1 Gbps	4
10 Gbps	2

# STP (6) - Spanning Tree Algorithm - Port states

Estado de los ports (IEEE 802.1d)

STP State	STP Activity	Usr.Fr.Fwd'
Disabled	Port is not active; it is not participating in any STP activity	No
<b>Listening</b>	Port is sending and receiving BPDUs.	No
<b>Learning</b>	Building loop-free bridging table.	No
<b>Forwarding</b>	Sending and receiving user data.	<b>Yes</b>
<b>Blocking</b>	Not permitting user traffic out the port.	No

Notas:

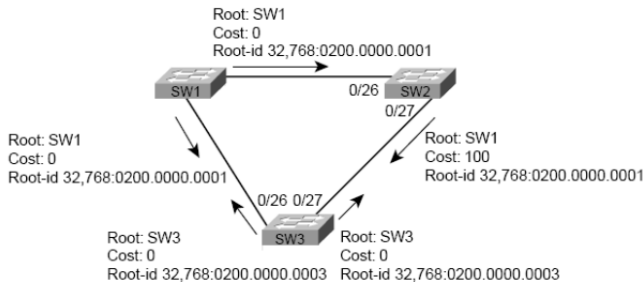
- Un port en Blocking sólo escucha BPDUs
- en Listening, Learning y Forwarding se envían y reciben BPDUs

## STP (7) - Spanning Tree Algorithm - Rol del Port

- **Designated ports** son los que se “conducen hacia afuera” del Root Bridge en un designated Bridge (elegido en una LAN para llegar al root).
  - En el root bridge, todos sus ports son designated ports
  - Solo hay un designated port por segmento LAN
  - Designated ports están en forwarding
- **Root ports** es el port que me lleva hacia el Root Bridge.
  - Hay un solo root port por bridge (que no sea el Root Bridge)
  - Root ports están en forwarding
- **Nondesignated ports** cualquier port que no sea elegido root port o designated port y esté activo.
  - Nondesignated están en blocking

# STP (8) - Root Election Process (1)

Figure 2-3 Root Election Process



**Default Priority = 32768 = 0x8000**

## STP (8) - Root Election Process (2)

Figure 2-4 SW1 Wins Election

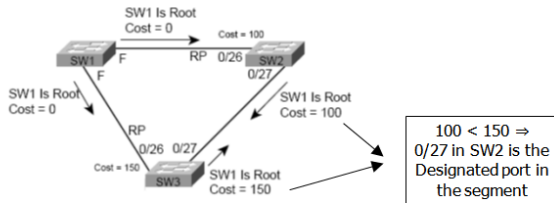


Table 2-3 State of Each Interface

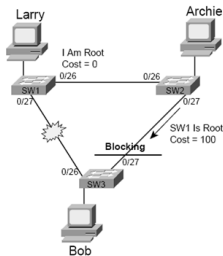
Bridge Interface	State	Reason That the Interface Is in Forwarding State
SW1, E 0/26	Forwarding	The interface is on the root bridge
SW1, E 0/27	Forwarding	The interface is on the root bridge
SW2, E 0/26	Forwarding	The root port
SW2, E 0/27	Forwarding	The designated port on the LAN segment to SW3
SW3, E 0/26	Forwarding	The root port
SW3, E 0/27	Blocking	Not the root bridge, not the root port, and not designated port



## STP (9) - Link Failure Reaction

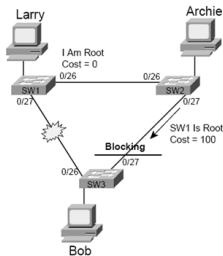
- SW3 ceases to receive the hello message in its root port, interface 0/26.

*Reacting to Link Failure Between SW1 and SW3*



## STP (9) - Link Failure Reaction

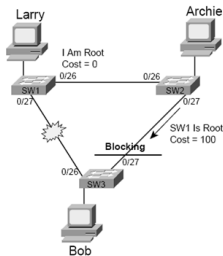
*Reacting to Link Failure Between SW1 and SW3*



- SW3 ceases to receive the hello message in its root port, interface 0/26.
- SW3 waits MaxAge seconds before deciding that it is no longer receiving the same root BPDU in its root port (20 seconds is the default).

## STP (9) - Link Failure Reaction

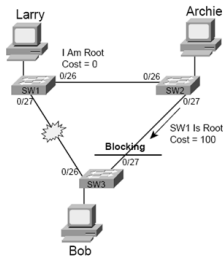
*Reacting to Link Failure Between SW1 and SW3*



- SW3 ceases to receive the hello message in its root port, interface 0/26.
- SW3 waits MaxAge seconds before deciding that it is no longer receiving the same root BPDU in its root port (20 seconds is the default).
- At that point, SW3 places port fastethernet 0/27 in listening state for Forward Delay seconds (15 seconds is the default).

## STP (9) - Link Failure Reaction

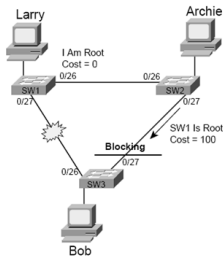
*Reacting to Link Failure Between SW1 and SW3*



- SW3 ceases to receive the hello message in its root port, interface 0/26.
- SW3 waits MaxAge seconds before deciding that it is no longer receiving the same root BPDU in its root port (20 seconds is the default).
- At that point, SW3 places port fastethernet 0/27 in listening state for Forward Delay seconds (15 seconds is the default).
- After that, SW3 places fastethernet 0/27 in learning state for another Forward Delay seconds before transitioning the port to forwarding state.

## STP (9) - Link Failure Reaction

### Reacting to Link Failure Between SW1 and SW3



- SW3 ceases to receive the hello message in its root port, interface 0/26.
- SW3 waits MaxAge seconds before deciding that it is no longer receiving the same root BPDU in its root port (20 seconds is the default).
- At that point, SW3 places port fastethernet 0/27 in listening state for Forward Delay seconds (15 seconds is the default).
- After that, SW3 places fastethernet 0/27 in learning state for another Forward Delay seconds before transitioning the port to forwarding state.
- So, a total of MaxAge plus twice Forward Delay, or 50 seconds, is required.

## STP (10) - Consideraciones

- Cisco inventó 2 violaciones a STP
  - PortFast (entra directamente a forwarding, BPDU guard disabling) sino  $\text{Max Age} + 2 \times \text{FDT} = 50$
  - UplinkFast (alternativo pasa a forwarding ante link down)
- Link aggregation evita que ante la caída de un link haya reconvergencia

# Rapid Spanning Tree Protocol (RSTP)

- Estándar IEEE 802.1w
- Tiempos de convergencia menores (menores a 10, a veces 1-2)
- Deployment simultaneo con switches que solo soportan STP
  - RSTP features en switches que lo soportan
  - STP features en switches que soportan sólo 802.1d
- utiliza las mismas reglas para elegir Root Bridge, Root Ports, y Designated Ports

# STP vs RSTP

- STP: blocking a forwarding
  - debe dejar de recibir BPDUs del root port por Max Age segundos (20s)
  - para que una interfaz pase de blocking a forwarding (Listening, Learning  $2 \times \text{FDT} = 30\text{s}$ )
- RSTP mejora este tiempo cuando hay cambios de topología (típicamente menos de 10s)

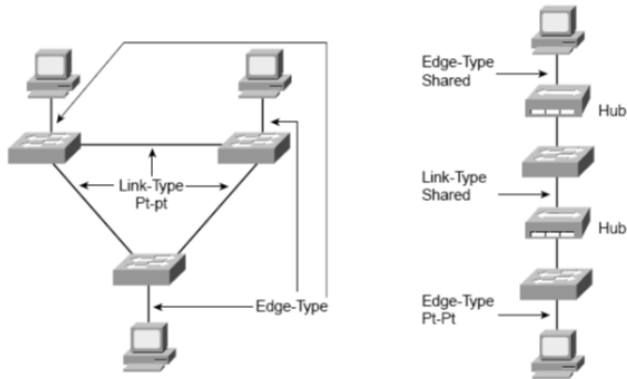


## RSTP Physical connectivity types

- RSTP distingue las conexiones entre link-type (entre switches) y edge-type (end-user)
- Hay 2 clases de conexiones link-type
  - point-to-point
  - shared
- RSTP reduce los tiempos de convergencia para conexiones point-to-point y edge-type.
- No mejora la convergencia para conexiones compartidas (hubs).

# RSTP Physical connectivity types

Figure 2-7 RSTP Link and Edge Types



# Aún más STP

- Cisco PVST Per-VLAN Spanning Tree
- MSTP Multiple Spanning Tree Protocol (IEEE 802.1s)  
(regiones e instancias)

# Preguntas

Which of the following interface states could a switch interface settle into after STP has completed building a spanning tree?

- Listening
- Blocking
- Forwarding
- Learning

# Preguntas

Which of the following are transitory port states used only during the process of STP convergence?

- Blocking
- Forwarding
- Listening
- Learning
- Discarding
- Disabled

# Preguntas

Which of the following bridge IDs would win election as root, assuming that the switches with these bridge IDs were in the same network?

- 32768:0200.1111.1111
- 32768:0200.2222.2222
- 200:0200.1111.1111
- 200:0200.2222.2222
- 40,000:0200.1111.1111
- 40,000:0200.2222.2222

# Preguntas

Which of the following facts determines how often a root bridge or switch sends a BPDU message?

- The hello interval configured on a nonroot switch
- The hello interval configured on the root switch
- It is always every 2 seconds.
- The switch reacts to BPDUs received from the root switch by sending another BPDU 2 seconds after receiving the root BPDU.