

Spanning Tree Protocol

Area de Ingeniería Telemática
<http://www.tlm.unavarra.es>

Programación de Redes
Grado en Ingeniería Informática, 3º

Temas de teoría

1. Introducción
2. Campus LAN
 - Ethernet conmutada para LANs
 - VLANs
 - **Protección en LANs Ethernet**
 - WLANs
3. Encaminamiento
4. Tecnologías de acceso y WAN

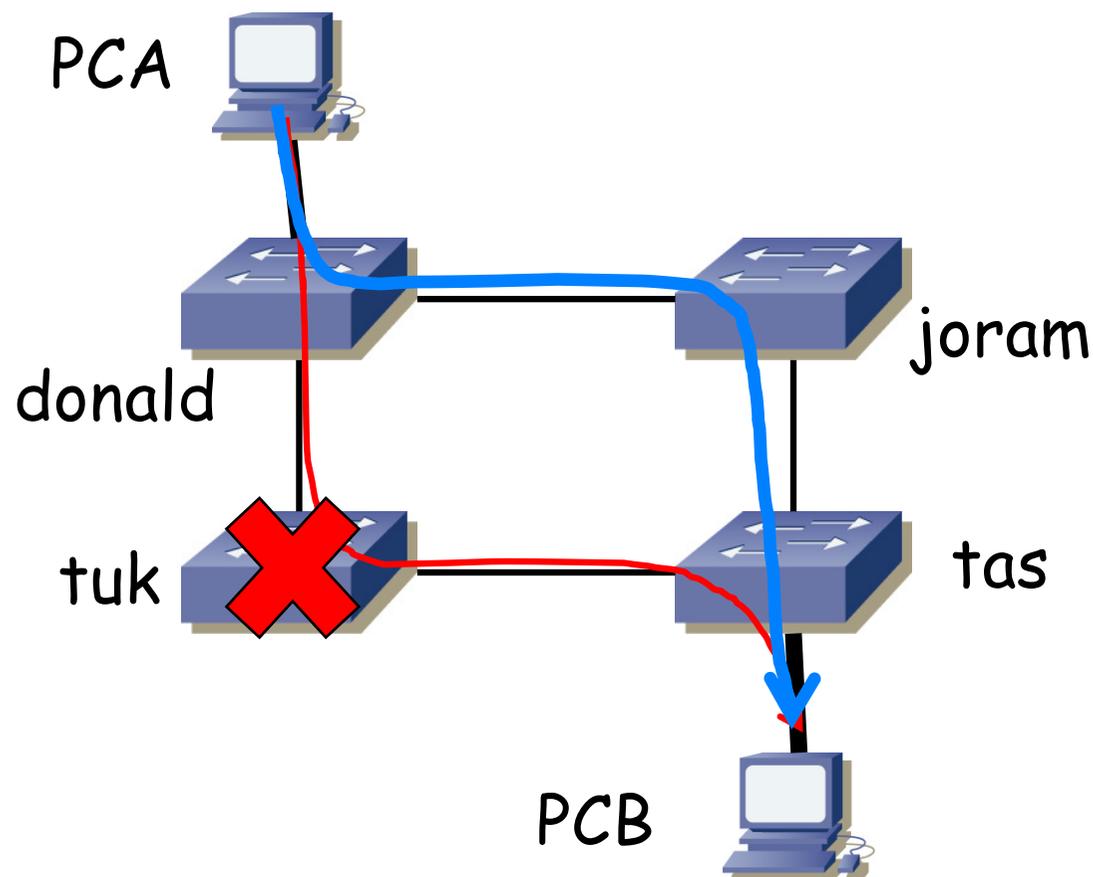
Objetivos

- Comprender qué necesidades cubre STP
- Conocer a grandes rasgos el funcionamiento de STP
- Conocer soluciones para la interacción entre STP y VLANs

Bucles en LANs con puentes

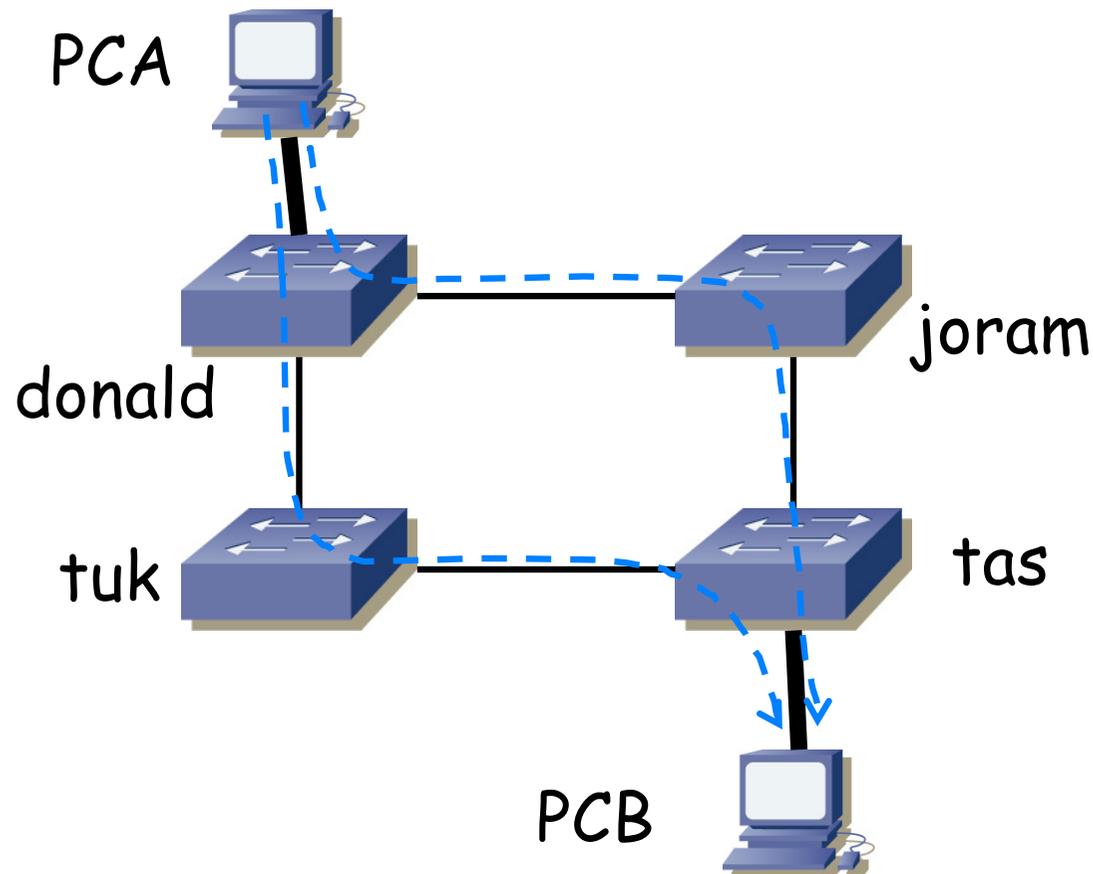
Caminos alternativos

- Ofrecerían la posibilidad de:
 - Reconfiguración ante fallos (...)
 - (...)



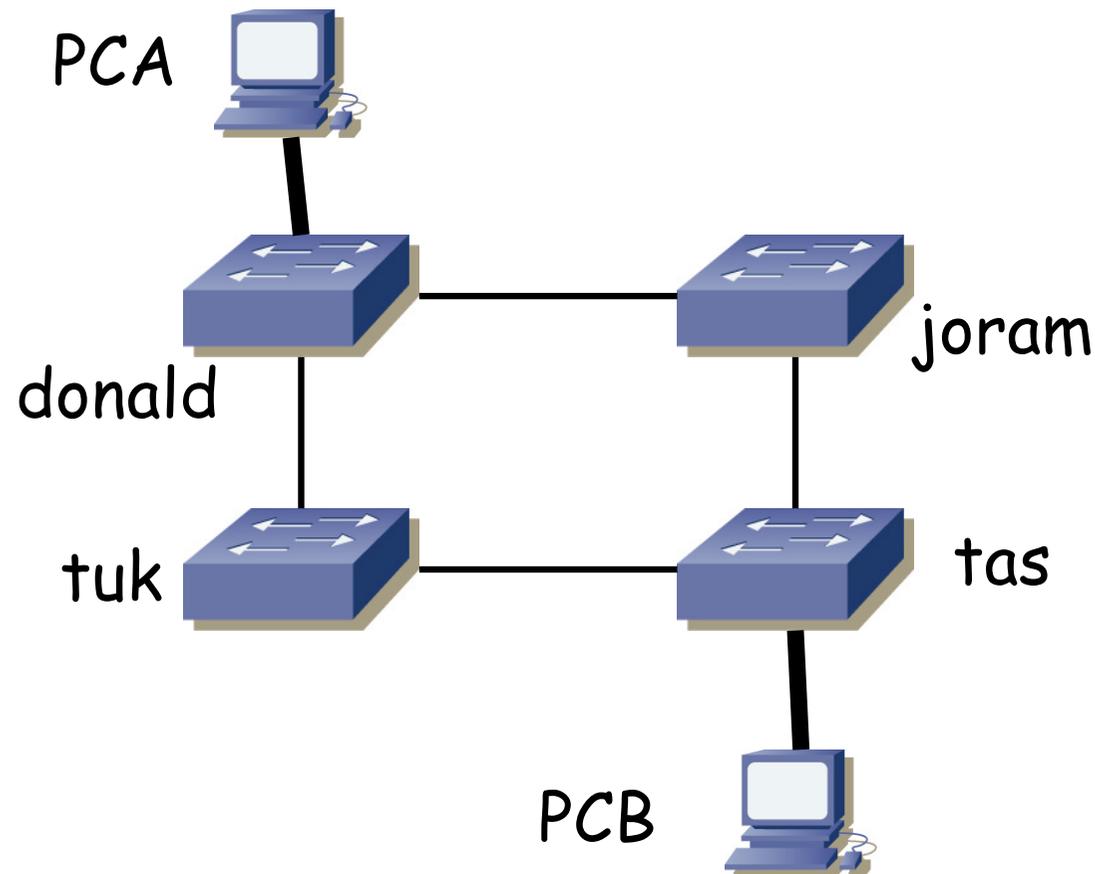
Caminos alternativos

- Ofrecerían la posibilidad de:
 - Reconfiguración ante fallos
 - Balanceo de carga
- Requiere tomar decisiones de encaminamiento



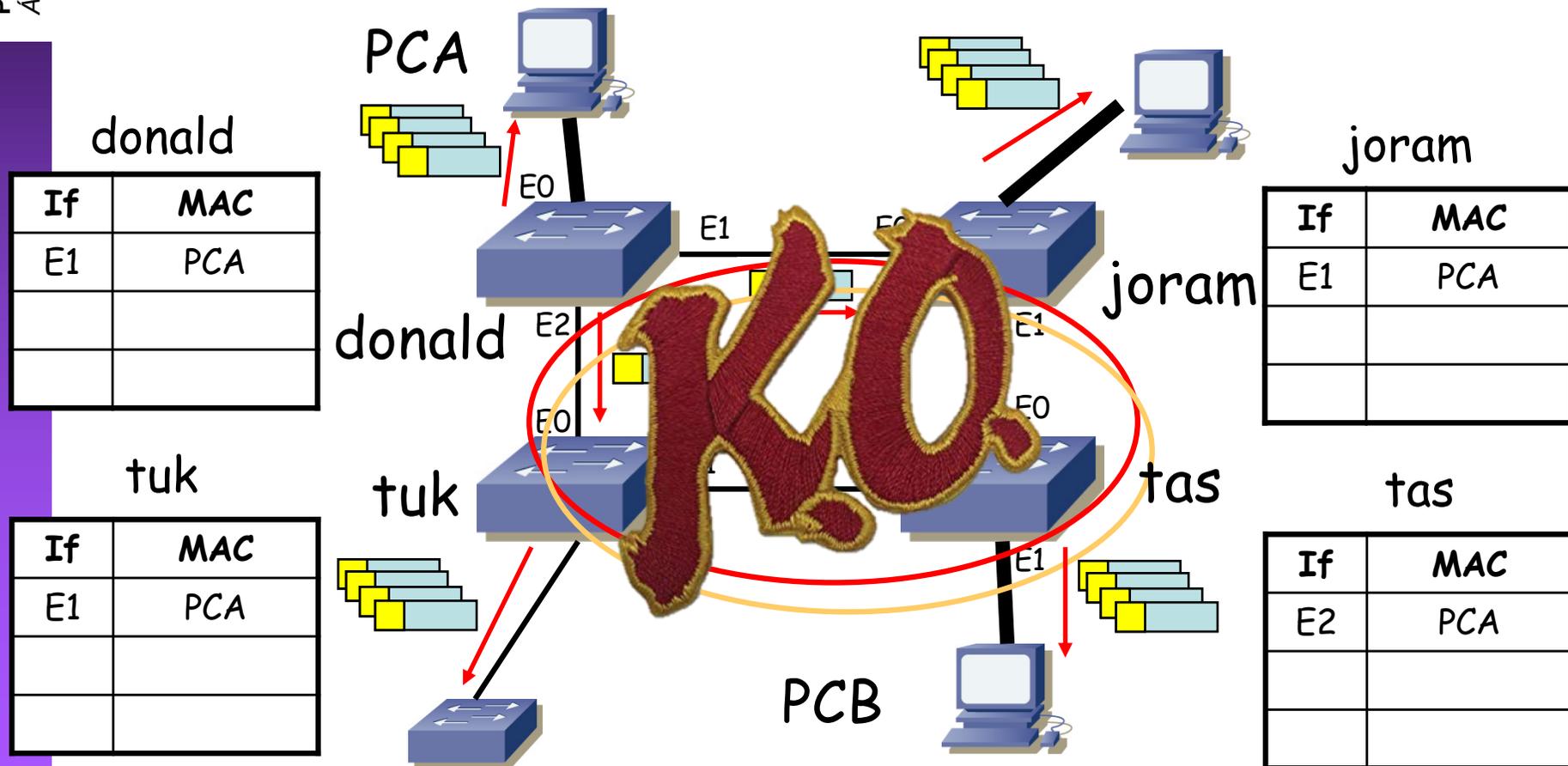
Pero...

- Ejemplo:
 - PCA envía una trama al PCB y los switches no han aprendido la MAC de PCB
 - O PCA envía una trama a broadcast
 - En ambos casos habrá inundación



Pero...

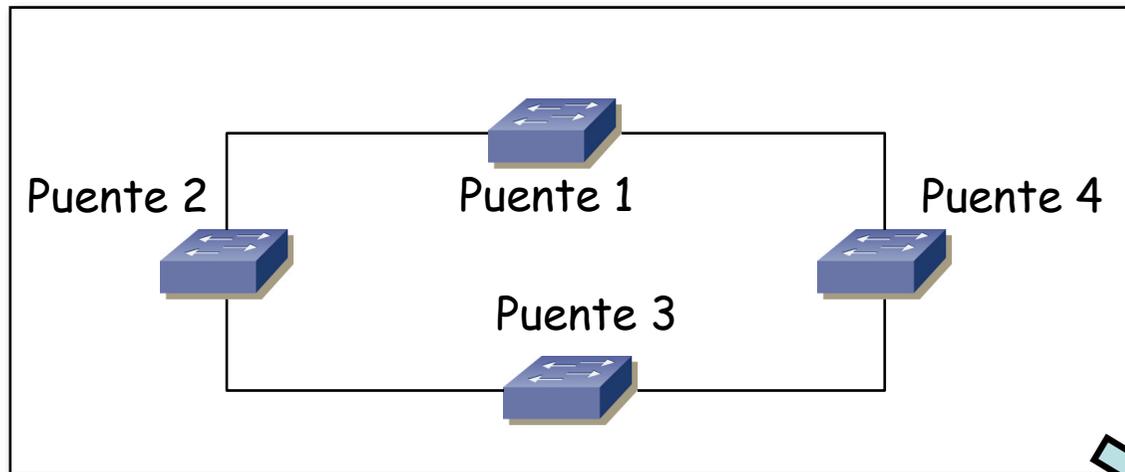
- Ejemplo:
 - Seguimos con los dos paquetes, uno en cada sentido del anillo
 - Esto no tiene fin
 - Además los paquetes salen por todos los puertos hacia hosts y otros switches



Spanning-Tree Protocol

Spanning-Tree Protocol (STP)

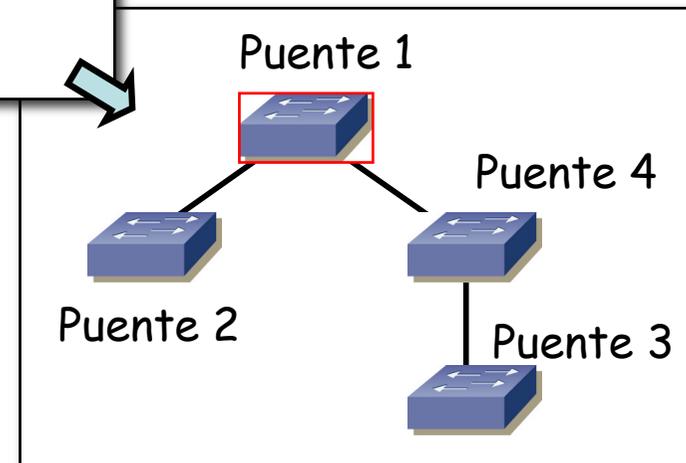
- Calcula una topología libre de ciclos
- A partir del grafo de la topología crea un árbol
- ¿Cómo? Seleccionan un puente como “raíz” y desde él calculan un árbol
- Bloquean puertos de enlaces que no están en el árbol



 IEEE
802.1D

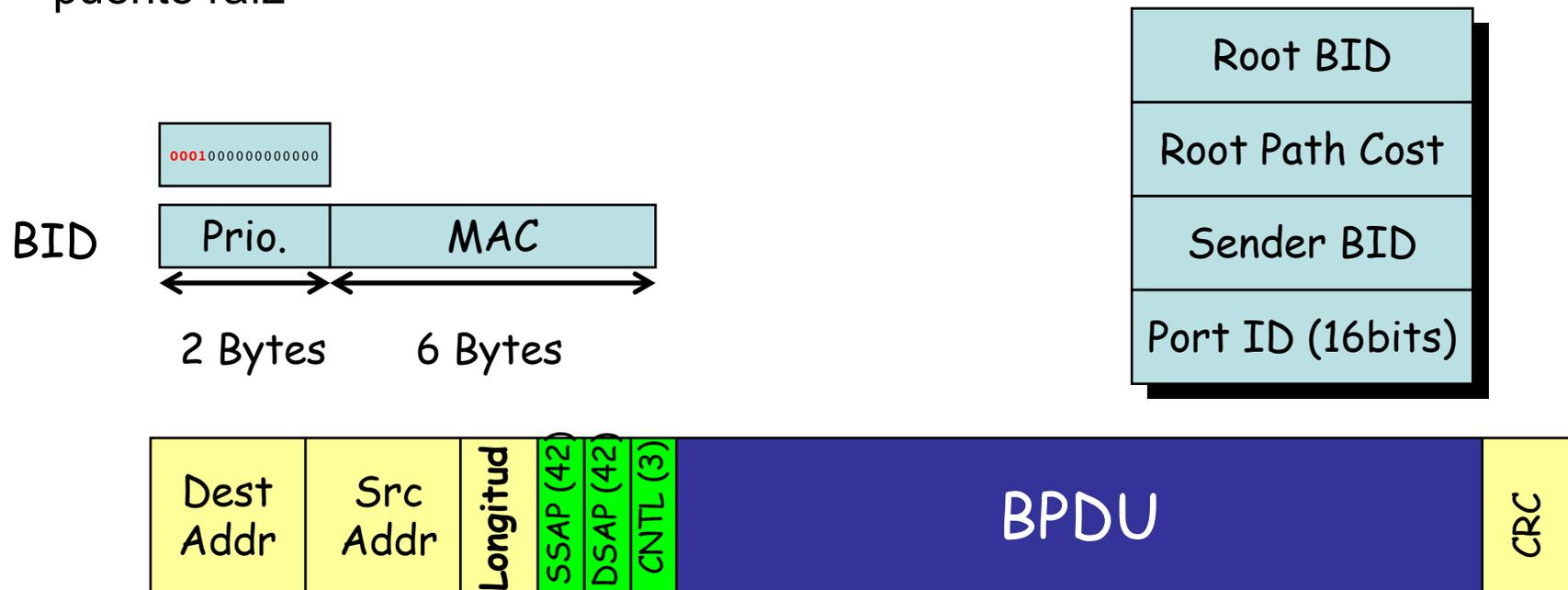


Radia Perlman



STP: BPDUs

- Bridge Protocol Data Units
- Enviadas periódicamente por los puentes, por todos sus puertos
- Destino 01:80:C2:00:00:00 (Bridge Group Address)
- No son reenviadas
- BID = Bridge ID = { Prioridad, MAC }
- Prioridad (default: 32768) en múltiplos de 4096
- En la BPDUs se anuncia el coste que este switch cree tener hasta el puente raíz



STP: BPDUs

No.	Time	Source	Destination	Total L	Source Port	Request Method	Status Code	Info
1	1193234155...	Cisco_87:85:...	Spanning-tree...					Conf. Roo
2	1193234157...	Cisco_87:85:...	Spanning-tree...					Conf. Roo

> Frame 1: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
> IEEE 802.3 Ethernet
> Logical-Link Control

▼ Spanning Tree Protocol

- Protocol Identifier: Spanning Tree Protocol (0x0000)
- Protocol Version Identifier: Spanning Tree (0)
- BPDU Type: Configuration (0x00)
- > BPDU flags: 0x00
- > Root Identifier: 32768 / 100 / 00:1c:0e:87:78:00
- Root Path Cost: 4
- > Bridge Identifier: 32768 / 100 / 00:1c:0e:87:85:00
- Port identifier: 0x8004
- Message Age: 1
- Max Age: 20
- Hello Time: 2
- Forward Delay: 15

Root BID

Root Path Cost

Sender BID

Port ID (16bits)

```

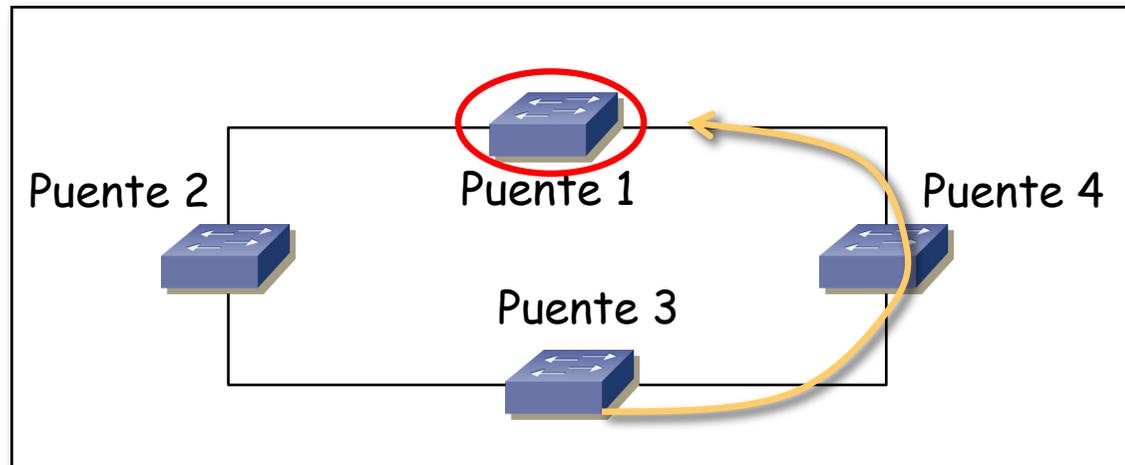
0000  01 80 c2 00 00 00 00 1c 0e 87 85 04 00 26 42 42  .....&BB
0010  03 00 00 00 00 00 80 64 00 1c 0e 87 78 00 00 00  .....d...x...
0020  00 04 80 64 00 1c 0e 87 85 00 80 04 01 00 14 00  ...d.....
0030  02 00 0f 00 00 00 00 00 00 00 00 00  .....
    
```

Spanning Tree Protocol (stp), 35 bytes Packets: 96 · Displayed: 96 (100.0%) Profile: Default



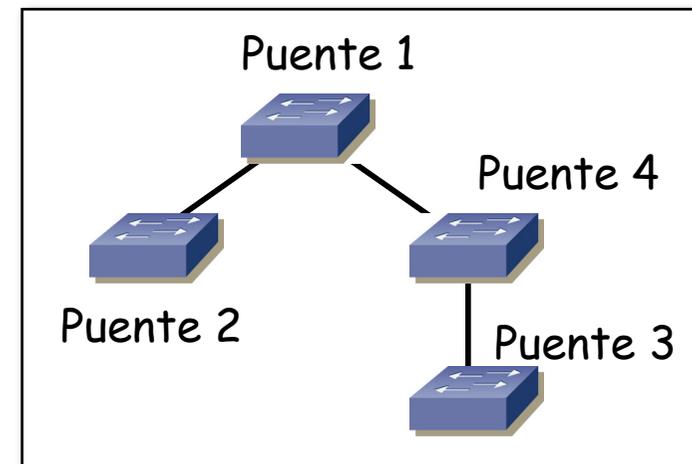
STP: BPDUs

- Se pueden “comparar” entre sí y decidir si una BPDU recibida por un puerto es “mejor” que otra
- “Mejor” en el sentido de “mejor” camino a la raíz
- Para ello cada enlace tendrá un coste y se suman
- El coste de un enlace se configura por defecto en función de su velocidad pero **debe ser configurable**



Dos detalles importantes

- Los conmutadores siguen enviando BPDUs tras calcular árbol
 - Un conmutador no tiene forma de saber que se ha alcanzado una topología estable
 - Seguir enviando y aceptando BPDUs permite recalculer el árbol ante fallos
- El plano de datos del conmutador no cambia
 - Es decir, los conmutadores siguen reenviando los paquetes de usuario en función de sus tablas
 - Y siguen aprendiendo igual
 - Simplemente algunos puertos descartan las tramas que reciben y no se reenvían tramas de usuario por ellos
 - Para decidir por dónde va una trama necesitamos saber qué puertos están bloqueados pero por lo demás sigue la misma lógica de siempre

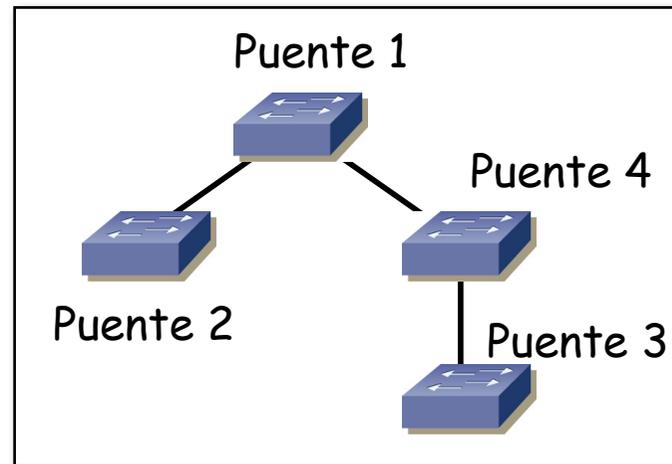


STP/RSTP: Mecanismos

Root bridge

Mecanismos

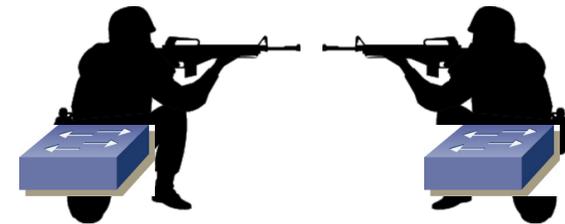
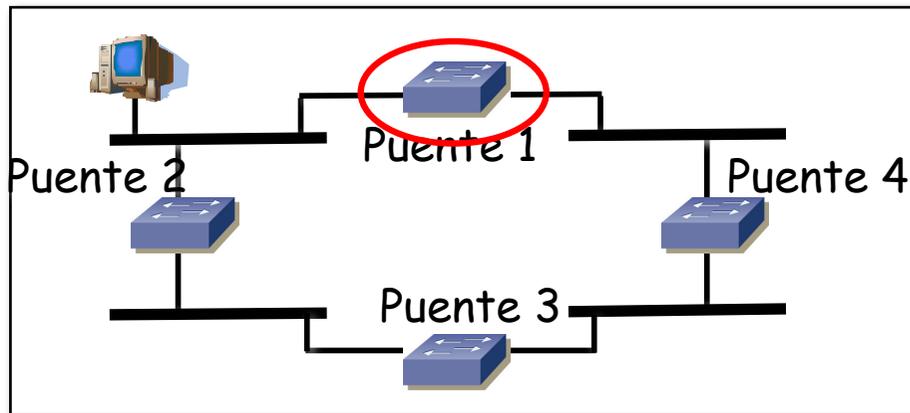
- Vamos a intentar entender cómo se calcula el árbol
- No nos interesa el “transitorio” sino solamente el estado final
- El mecanismo es básicamente *distance-vector* (Bellman-Ford distribuido)



Spanning-Tree Protocol (STP)

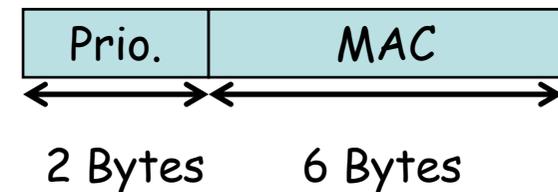
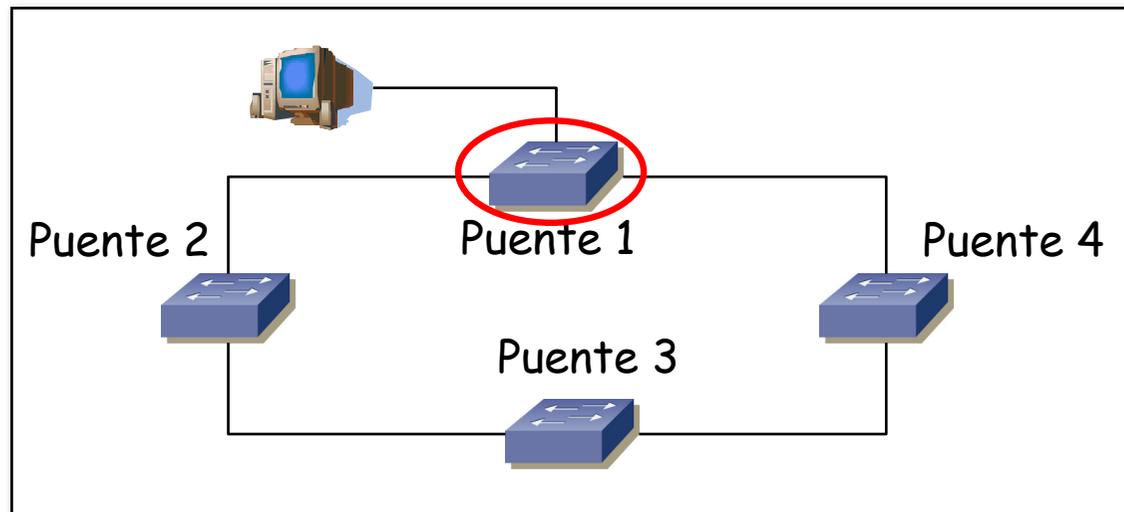
Selección de un *Root Bridge* (Root War !!!)

- Raíz para el árbol
- No es un primer paso sino que para cualquier BPDU que se recibe se decide si anuncia mejor root
- A partir de un valor de prioridad y una MAC del puente
 - Vienen en las BPDU
 - Puente de prioridad más baja (def. 0x8000)
 - MAC más baja en caso de empate



STP: Root Bridge

- Por defecto todos los puentes la misma prioridad
- Gana el de dirección MAC más baja
- Primeros 3 bytes de la MAC son el OUI
- ¡ Luego el ganador depende del fabricante !
- Cuidado pues puede ser el conmutador más lento
- Selección manual con el campo de prioridad



Path cost

STP: Path Cost

- Asociado a cada LAN; según su velocidad o **administrativamente**
- Se va agregando en un camino creando el *Root Path Cost*
- En un switch se calcula a partir de los anuncios recibidos en cada puerto, sumando el coste del puerto por el que han llegado
- 802.1D-1998:

Table 8-5—Path Cost Parameter Values

Parameter	Link Speed	Recommended value	Recommended range	Range
Path Cost	4 Mb/s	250	100–1000	1–65 535
Path Cost	10 Mb/s	100	50–600	1–65 535
Path Cost	16 Mb/s	62	40–400	1–65 535
Path Cost	100 Mb/s	19	10–60	1–65 535
Path Cost	1 Gb/s	4	3–10	1–65 535
Path Cost	10 Gb/s	2	1–5	1–65 535

- ¿Y si tuviéramos un enlace de 2x10Gbps? ¿Coste 1?
- ¿Y si tuviéramos un enlace de 4x10Gbps?

STP: Path Cost

- Estos valores recomendados cambian en 802.1t
- 802.1D-2004 (desde 802.1t-2001):
 - Recomendado $2 \times 10^7 / \text{Link_Speed_en_Mbps}$
 - Podemos acumular al menos 20 saltos del peor coste sin exceder el máximo de un contador de 32bits

Table 17-3—Port Path Cost values

Link Speed	Recommended value	Recommended range	Range
<=100 Kb/s	200 000 000 [*]	20 000 000–200 000 000	1–200 000 000
1 Mb/s	20 000 000 ^a	2 000 000–200 000 000	1–200 000 000
10 Mb/s	2 000 000 ^a	200 000–20 000 000	1–200 000 000
100 Mb/s	200 000 ^a	20 000–2 000 000	1–200 000 000
1 Gb/s	20 000	2 000–200 000	1–200 000 000
10 Gb/s	2 000	200–20 000	1–200 000 000
100 Gb/s	200	20–2 000	1–200 000 000
1 Tb/s	20	2–200	1–200 000 000
10 Tb/s	2	1–20	1–200 000 000

^{*}Bridges conformant to IEEE Std 802.1D, 1998 Edition, i.e., that support only 16-bit values for Path Cost, should use 65 535 as the Path Cost for these link speeds when used in conjunction with Bridges that support 32-bit Path Cost values.

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

PROGRAMACIÓN DE REDES
Área de Ingeniería Telemática

Estados y RSTP

STP: Port States

- STP definía 5 estados posibles para un puerto: *disabled*, *listening*, *learning*, *blocking*, *forwarding*
- Estos estados mezclaban por un lado si el puerto reenviaba o no tramas y por otro el papel que jugaba el puerto en el árbol
- RSTP separa *port states* de *port roles*
 - Los *estados* definen si se reenvían las tramas y si se aprenden direcciones MAC
 - Los *roles* definen el papel que juega el puerto en el árbol y con él se deduce el estado (hay estados transitorios)

RSTP

Rapid Spanning-Tree Protocol

- STP obsoleto y retirado del estándar
- RSTP es IEEE 802.1w
- RSTP es el STP que aparece en 802.1D-2004
- Tiempos de convergencia de 2-3 segs (aunque según la topología puede llegar a 30s y cuentas a infinito)
- Tres **estados** posibles para un puerto:
 - *Discarding*: ni envía ni acepta paquetes de usuario
 - *Learning*: no envía ni acepta paquetes de usuario pero aprende MACs
 - *Forwarding*: funcionamiento normal
- No vamos a detallar el diagrama de estados con sus transiciones ni cómo se adapta a cambios
- Sí vamos a detallar el significado de los *roles* pues ayudan a entender cómo se calcula el árbol

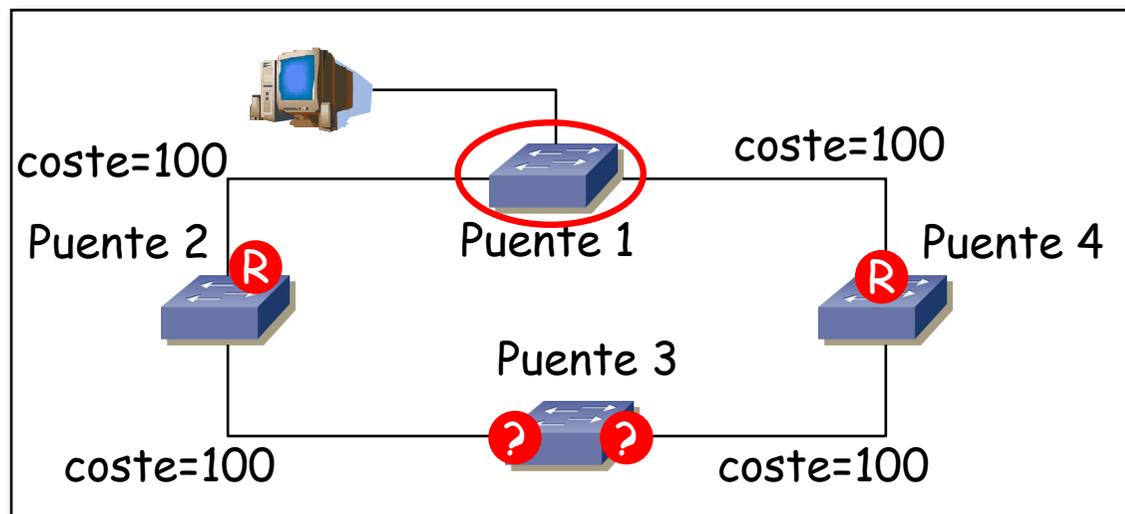
RSTP: Port Roles

Root port

RSTP: Port Roles

Root Port (puerto raíz)

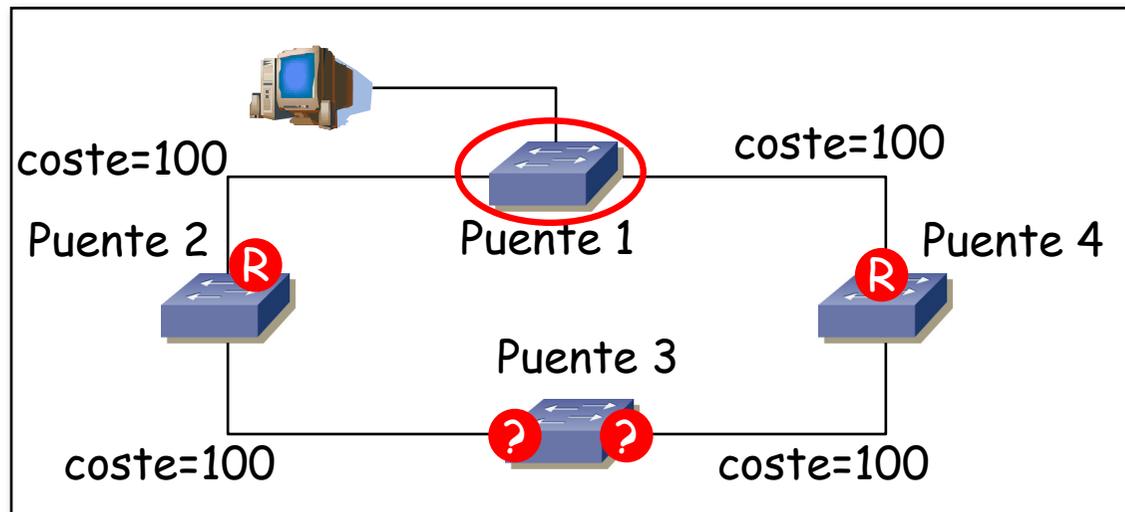
- Uno en cada puente salvo en el puente raíz
- El puente raíz es el único sin un puerto raíz
- No confundáis “Puente” raíz (*Root Bridge*) con “Puerto” raíz (*Root Port*)
- Es el puerto de un conmutador que tiene el menor *Root Path Cost anunciado*+*Port Cost* (menor coste hasta la raíz)
- En este ejemplo supongamos que todos los puertos tienen el mismo coste
- Está claro cuál es *root port* en los puentes 2 y 4 pero ¿y en 3?



RSTP: Port Roles

¿Empates?

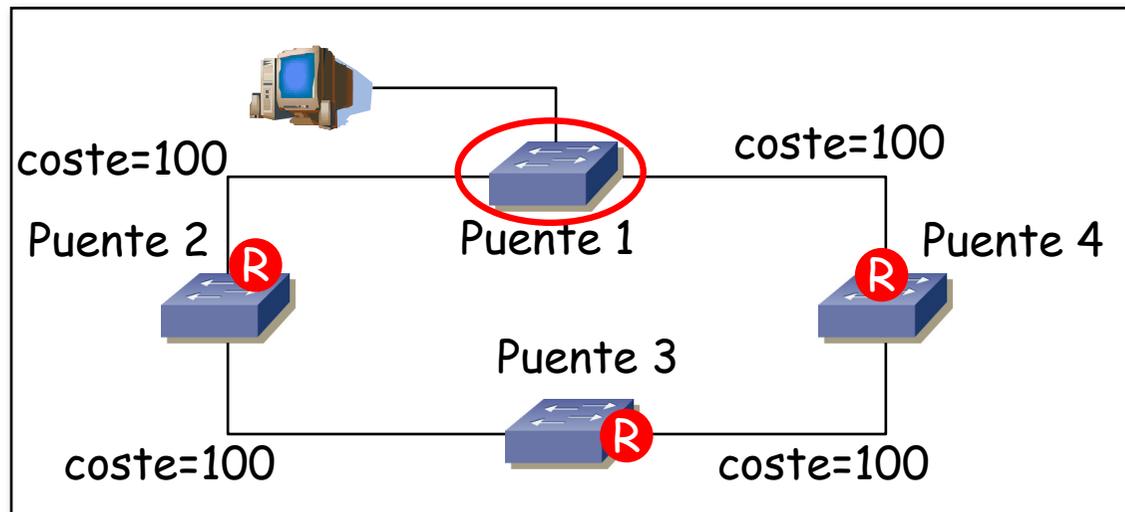
- En el puente 3 los costes que recibe de puente 2 y de puente 4 hasta la raíz son el mismo = 100
- A cada uno le suma el coste del puerto por el que lo ha recibido y empatan (le sale 200 en ambos)
- Entonces se comparan los BID de los puentes que hacen el anuncio
- El anuncio que venga del BID menor gana (...)



RSTP: Port Roles

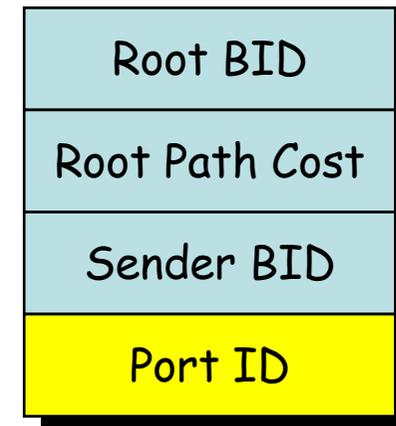
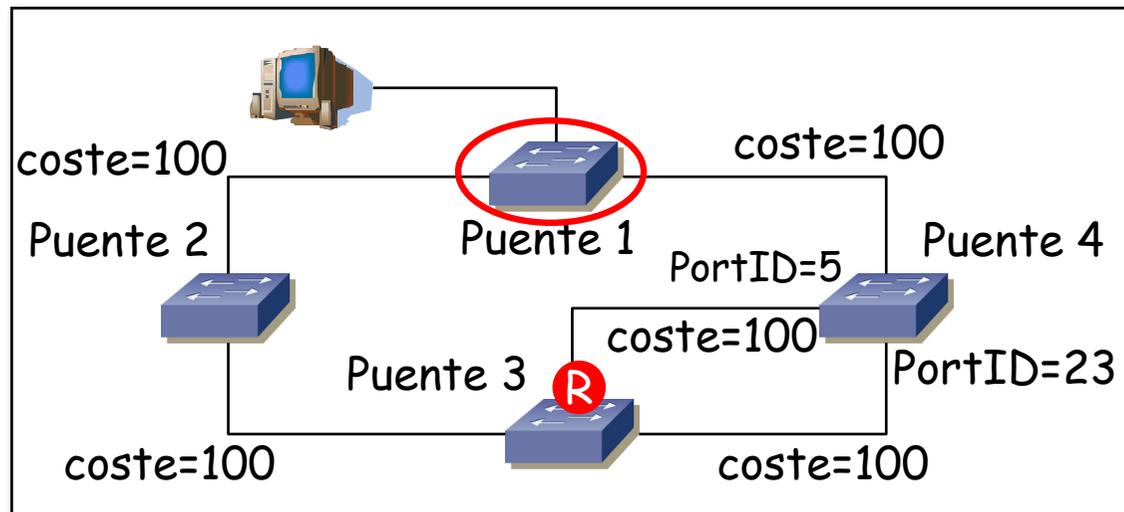
Desempate

- Suponiendo que BID del puente 4 < BID del puente 2 sería puerto raíz de puente 3 el que va al puente 4



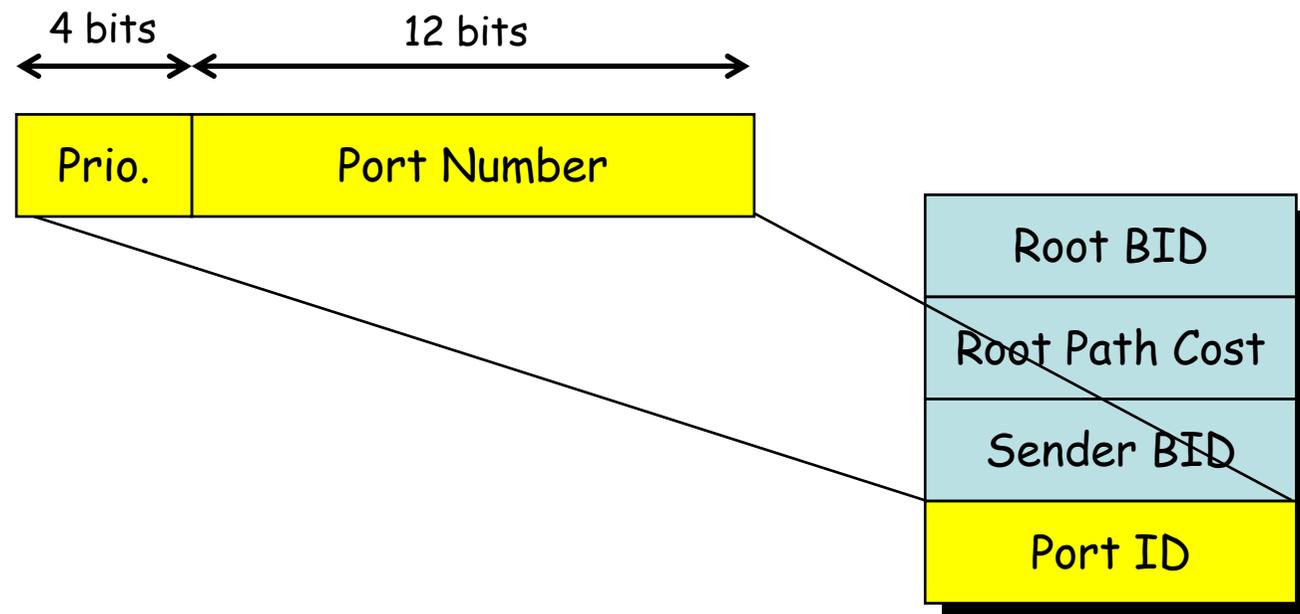
RSTP: Port Roles

- Hemos dicho que cuando hay que elegir un camino a la raíz se toma el menor coste agregado
- Si hay empate el menor BID
- ¿Hay más posibilidades de empate?
- En este caso empatarían los puertos a puente 4 (suponiendo que el BID del puente 4 es menor que el del puente 2)
- ¿Desempate? Menor identificador de **puerto**
- Es decir, en realidad es un solo coste compuesto:
 - (RootBID, RPC, SenderBID, PortBID, PortBIDrecv), el menor gana



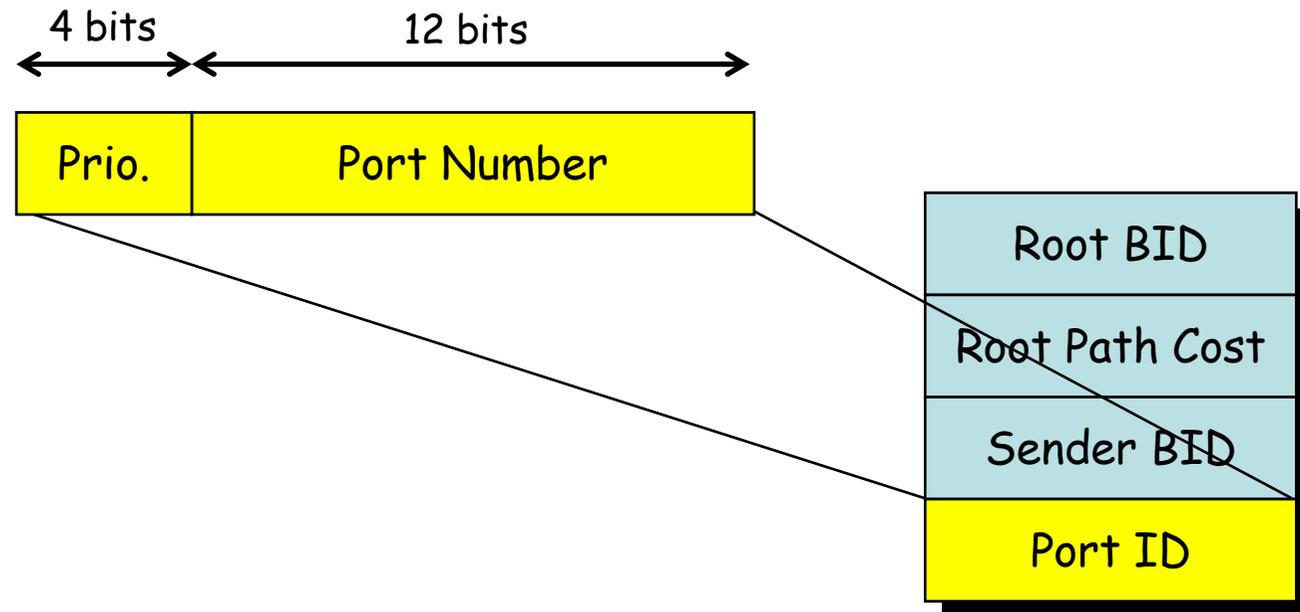
Port ID

- Es un valor de 16bits
- Tiene una componente de prioridad y un número de puerto
- Es decir, el valor de prioridad, si lo entendemos como el primer byte ignorando los últimos 4 bits, va en múltiplos de 16



Port ID

- > Frame 1: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
- > IEEE 802.3 Ethernet
- > Logical-Link Control
- ▼ Spanning Tree Protocol
 - Protocol Identifier: Spanning Tree Protocol (0x0000)
 - Protocol Version Identifier: Spanning Tree (0)
 - BPDU Type: Configuration (0x00)
 - > BPDU flags: 0x00
 - > Root Identifier: 32768 / 100 / 00:1c:0e:87:78:00
 - Root Path Cost: 4
 - > Bridge Identifier: 32768 / 100 / 00:1c:0e:87:85:00
 - Port identifier: 0x8004
 - Message Age: 1
 - Max Age: 20
 - Hello Time: 2
 - Forward Delay: 15

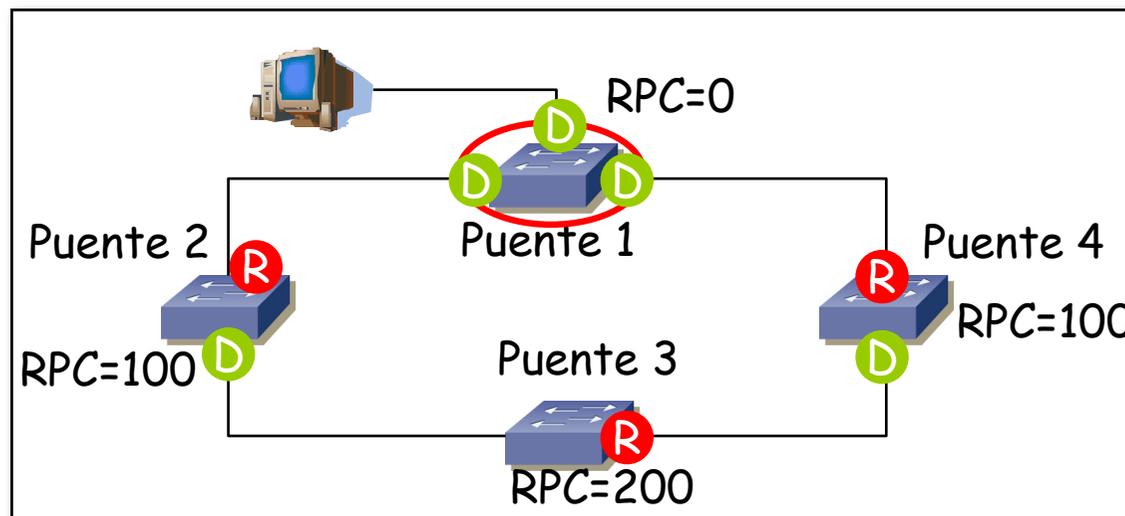


Designated port

RSTP: Port Roles

Designated Port (puerto designado)

- En un segmento de LAN puede haber varios puertos de conmutador
- El puerto de conmutador en un segmento de LAN con menor *Root Path Cost+Port Cost* es el puerto designado para la LAN
- Uno por segmento de LAN, el del switch con menor RPC al puente raíz
- Si hay empate por costes se desempata por el BID
- Ante nuevo empate desempataría por Port ID

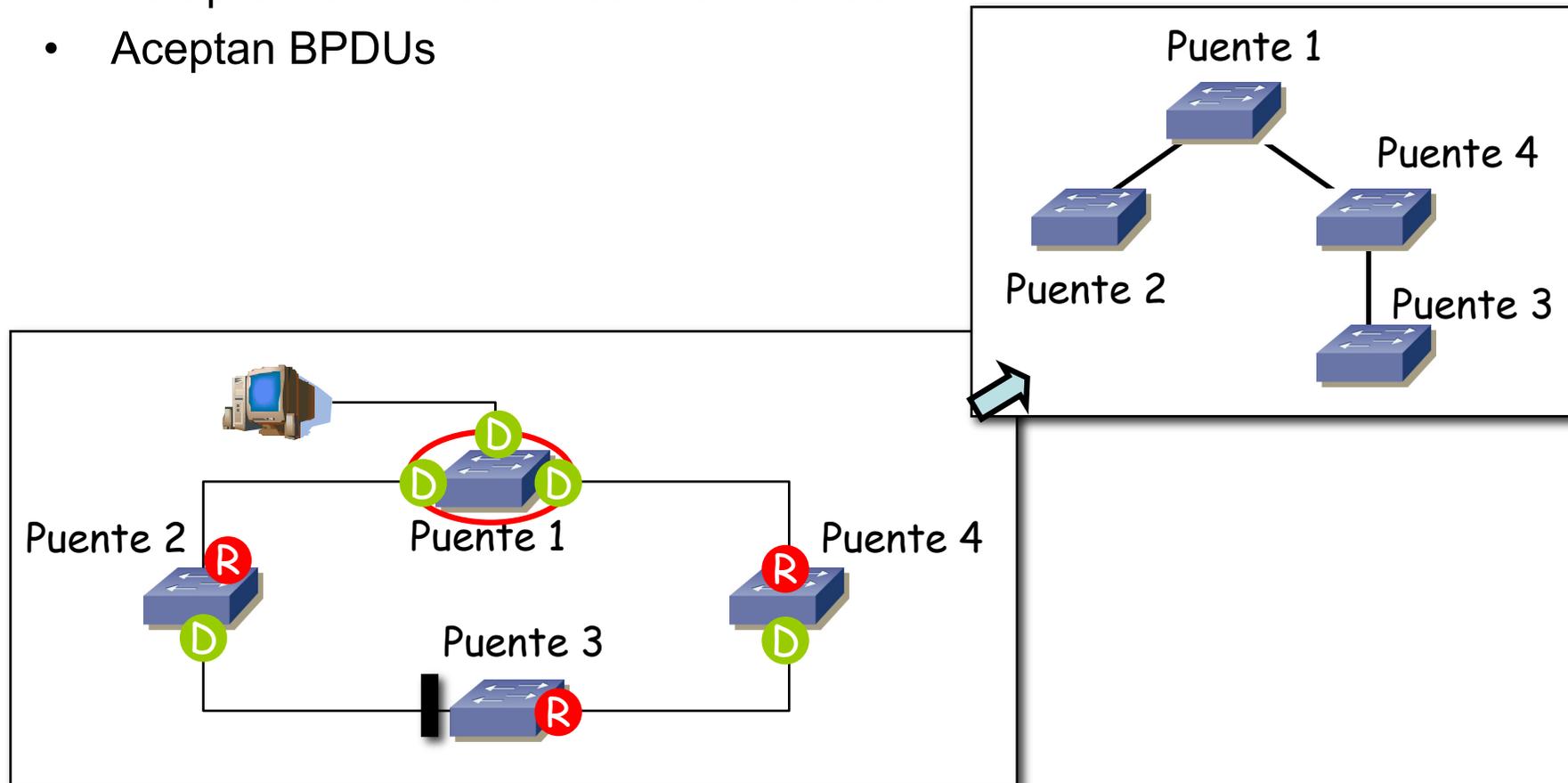


Puertos bloqueados

RSTP: Port Roles

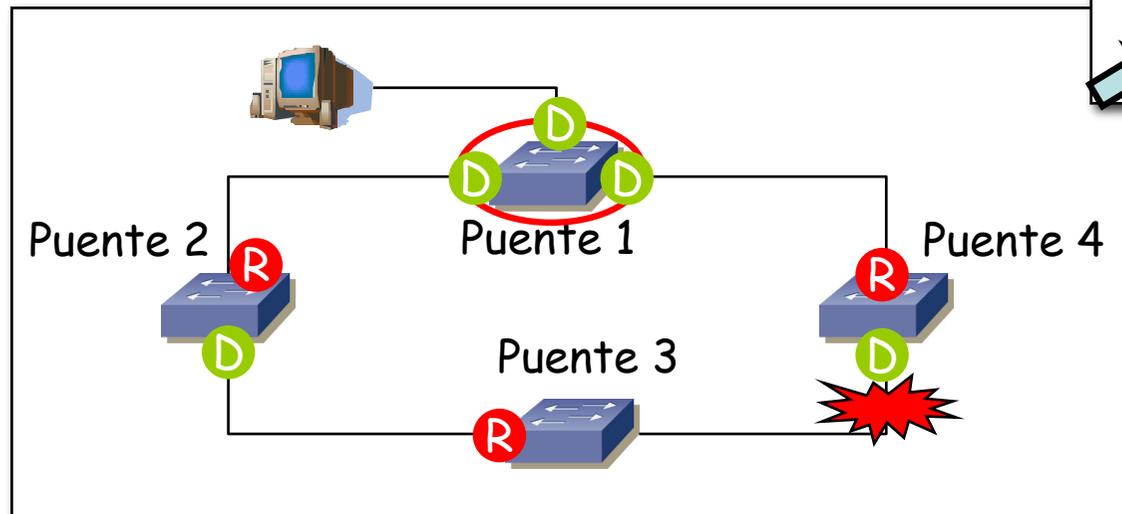
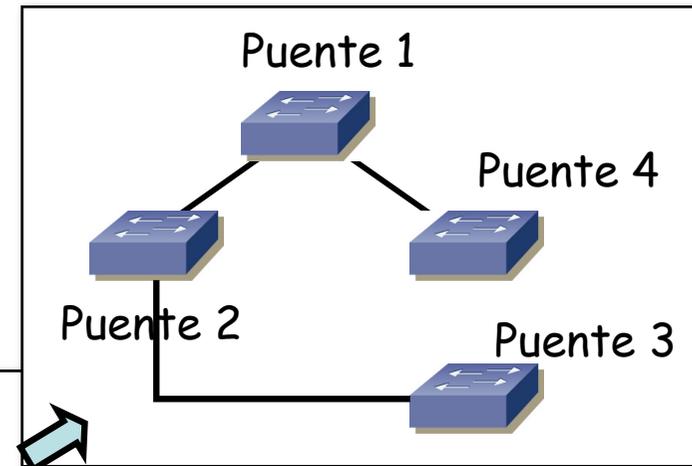
Blocked Port (no es un rol en RSTP)

- Se suele llamar así a los puertos con rol *Alternate* o *Backup*, que son los puertos que rompen los ciclos en la topología (herencia de STP)
- Todos los que no son ni *Root port* ni *Designated port*
- No aprenden MACs ni reenvían tramas
- Aceptan BPDUs



RSTP: cambios en la topología

- Ante un fallo (switch, puerto, enlace...) se recalcula el árbol
- ¿Cómo?
 - Se dejan de recibir BPDUs del adyacente donde se produce el fallo
 - Caduca la información que daban esas BPDUs
 - Otro camino pasa a ser mejor
 - Se mandan nueva BPDUs
- Tiempo de convergencia:
 - STP clásico 30-60 segs
 - RSTP 2-3 segs (mecanismos extra)



upna

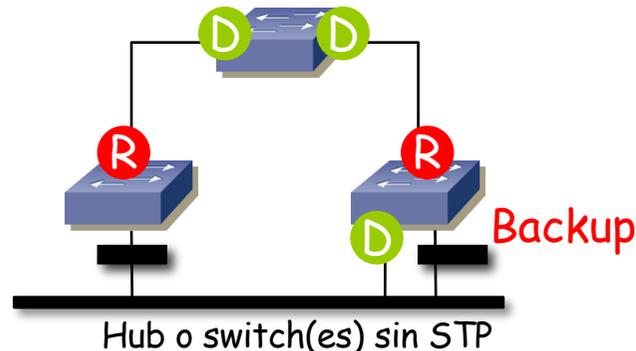
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

PROGRAMACIÓN DE REDES
Área de Ingeniería Telemática

Alternate y Backup

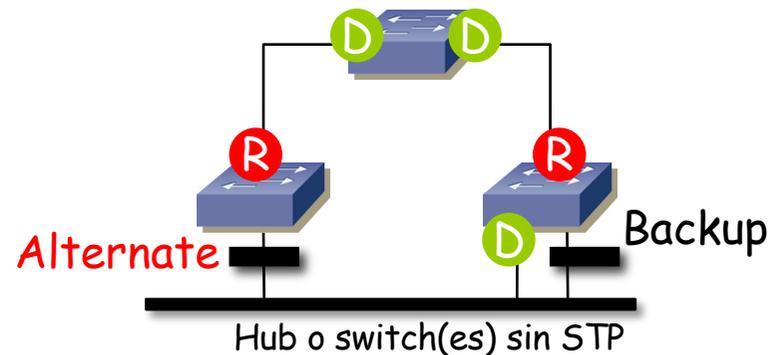
RSTP: *Port Roles*

- *Disabled*: puerto retirado mediante gestión
- *Alternate y Backup*:
 - Corresponden a lo que antes eran *blocked ports*
 - *Backup* es todo puerto que no es ni *Root* ni *Designated* y el puente es *Designated* para esa LAN
 - *Backup port* da un camino alternativo pero siguiendo el mismo camino que el *Root port*
 - *Backup port* solo existe donde haya 2+ enlaces de un puente a una LAN
 - *Backup* está bloqueado porque se han recibido BPDUs mejores **del mismo switch** en el mismo segmento
 - (...)



RSTP: *Port Roles*

- *Disabled*: puerto retirado mediante gestión
- *Alternate y Backup*:
 - Corresponden a lo que antes eran *blocked port*
 - Un *Alternate port* da un camino alternativo hacia el root frente al puerto que se tiene como *Root*
 - *Alternate* está bloqueado porque se han recibido BPDUs mejores (menor coste) de otro switch en el mismo segmento



upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

PROGRAMACIÓN DE REDES
Área de Ingeniería Telemática

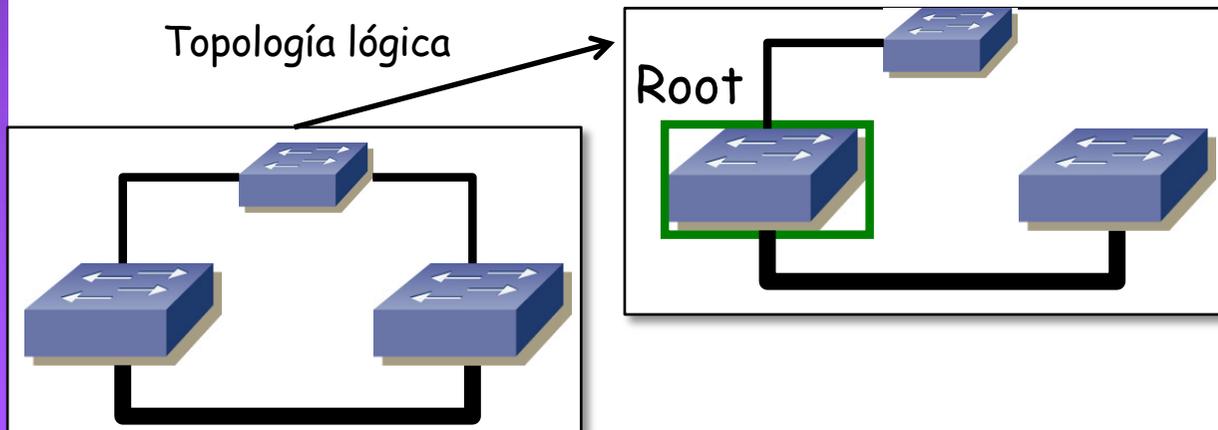
VLANs y STP

VLANs y Spanning Trees

Solución básica:

- Ignoramos las VLANs
- Un ST común a todas las VLANs (1 sola topología lógica, cómputo barato)
- CST = Common Spanning Tree

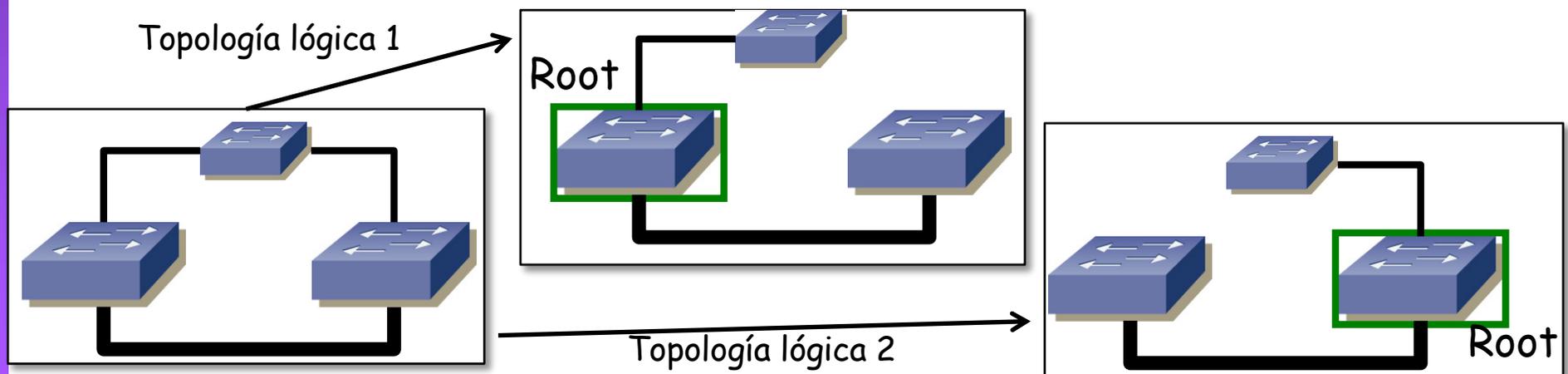
(...)



VLANs y Spanning Trees

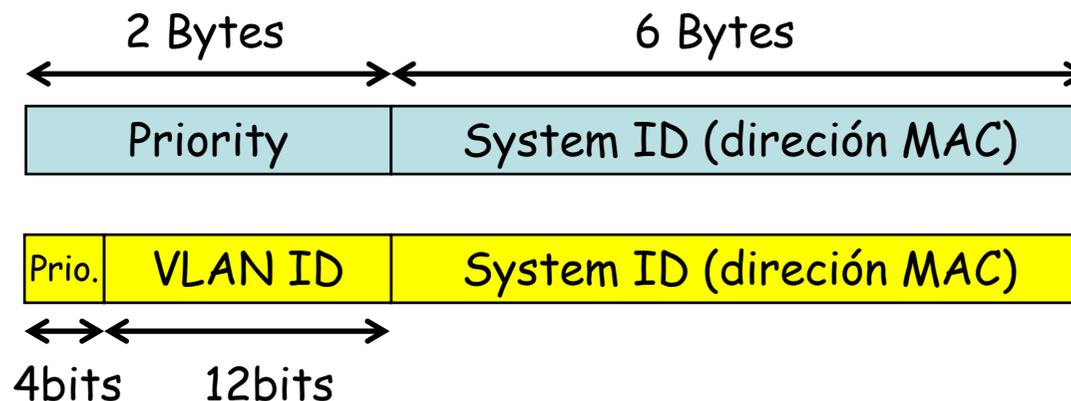
MSTP

- MSTP = Multiple Spanning Tree Protocol (modificación 802.1s a 802.1Q)
- Un ST por *grupo* de VLANs (que puede ser de una)
- Lo que se llama una MSTI (MST Instance)
- Una topología lógica por VLAN o por grupo de VLANs
- Para cada MSTI se pueden cambiar parámetros de ST, por ejemplo la prioridad para cambiar el Root Bridge o los costes de los enlaces
- Ejemplo: topología física con solo 2 posibles topologías lógicas, si se tienen N VLANs ($N > 2$) no es rentable calcular N STs



MSTP y BID

- El BID es de 8 bytes
- Haría falta uno diferente para el puente en cada VLAN
- Así como la dirección MAC se emplea como “system ID” se introduce (802.1t) un “extended system ID”
- Este “extended system ID” toma los 12 bits bajos del campo de prioridad
- Eso permite crear un BID para cada VLAN sin necesidad de más direcciones MAC
- Es lo que reduce los valores de prioridad a múltiplos de 4096 si entendemos como prioridad los 2 bytes



Resumen

- Topologías con redundancia ante fallos
- Proceso de elección de raíz
- Costes en los enlaces ajustables
- VLANs con árboles independientes o comunes
- Posibilidad de repartir carga por caminos redundantes con VLANs