

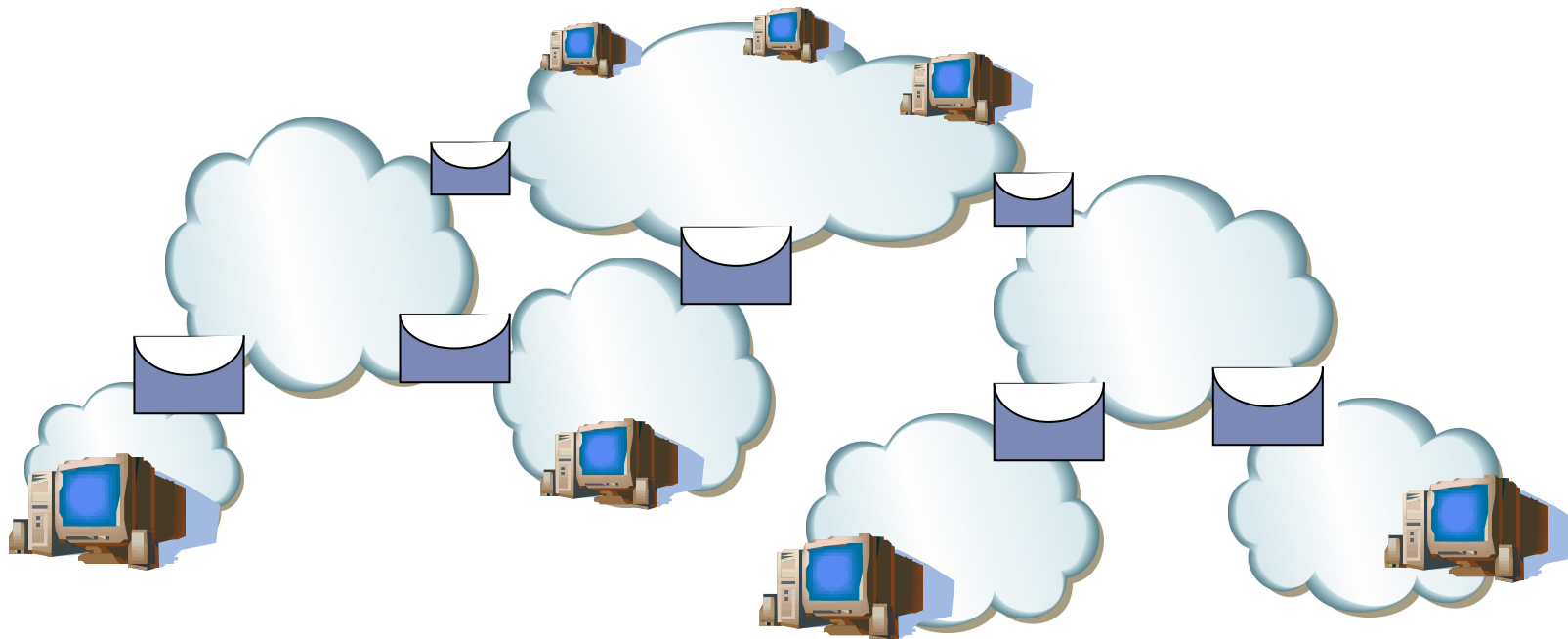
# Spanning Tree Protocol

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

Redes de Banda Ancha  
5º Ingeniería de Telecomunicación

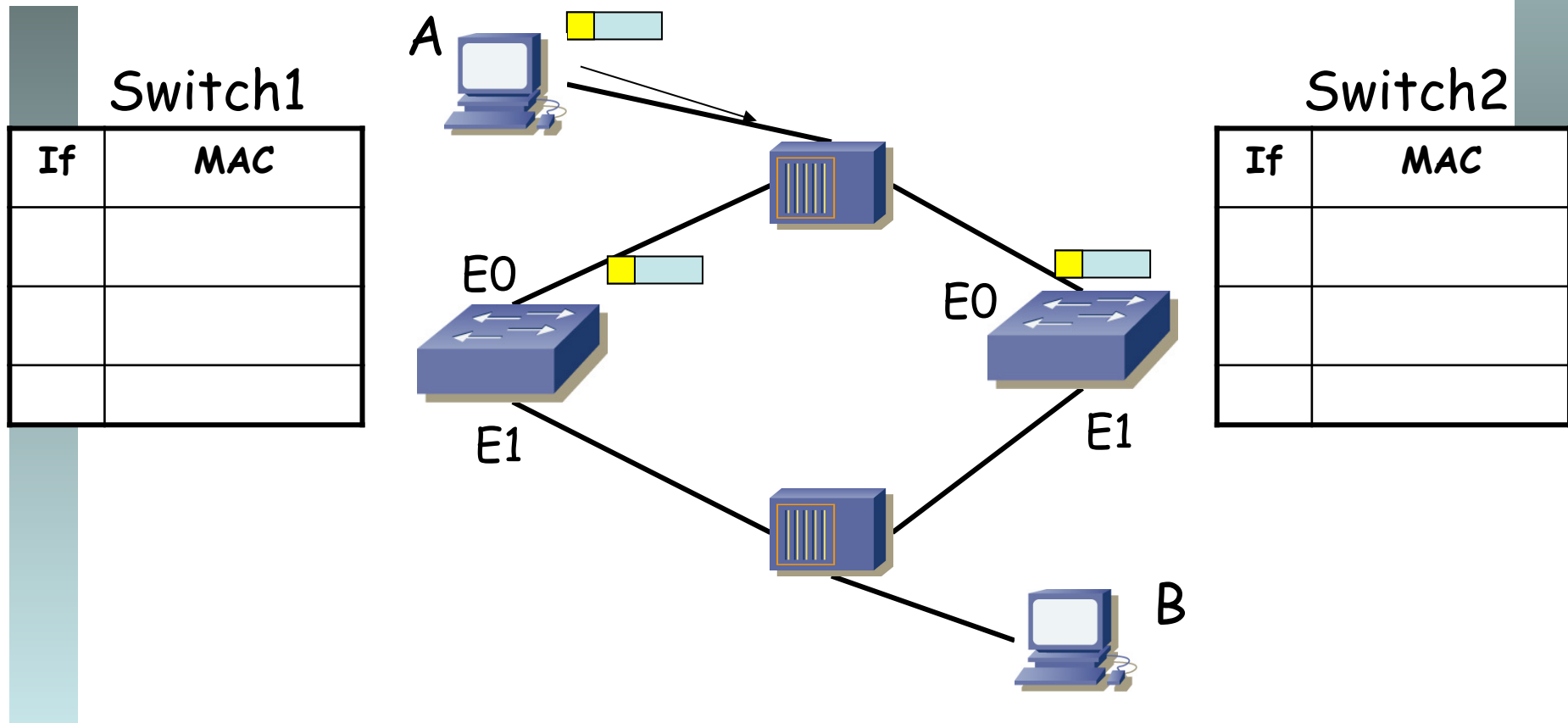
# Caminos alternativos

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



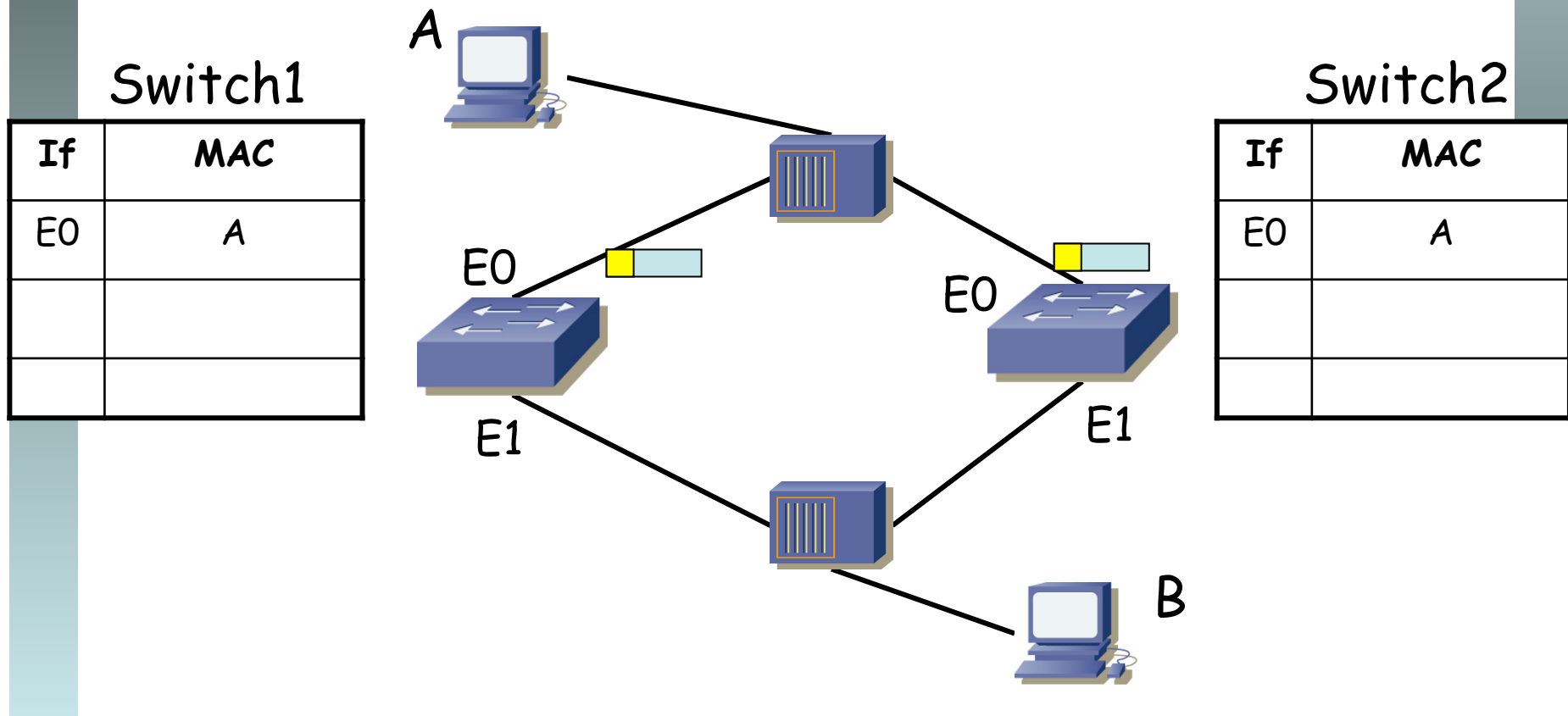
# Caminos alternativos

- El host A envía una trama al host B



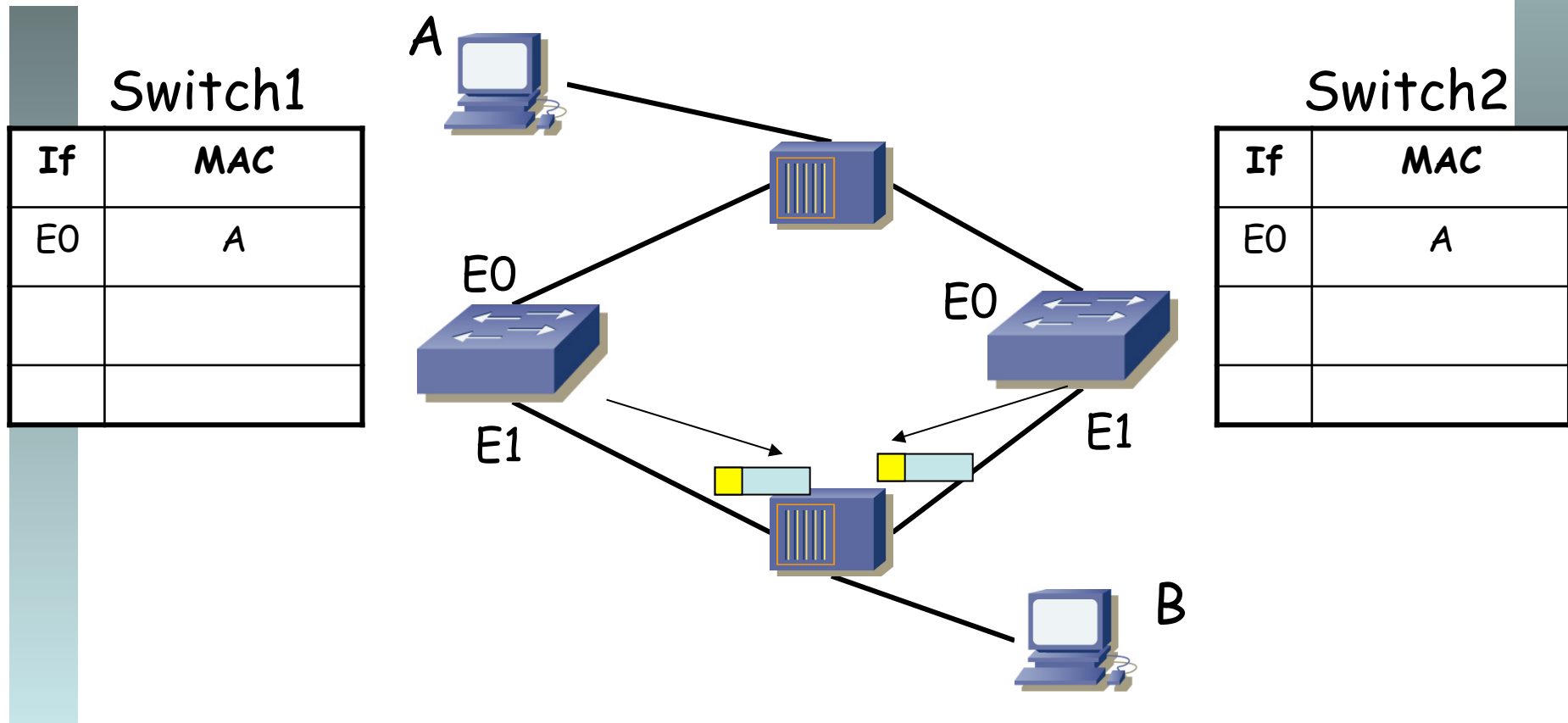
# Caminos alternativos

- Switch1 y Switch2 aprenden la localización del host A



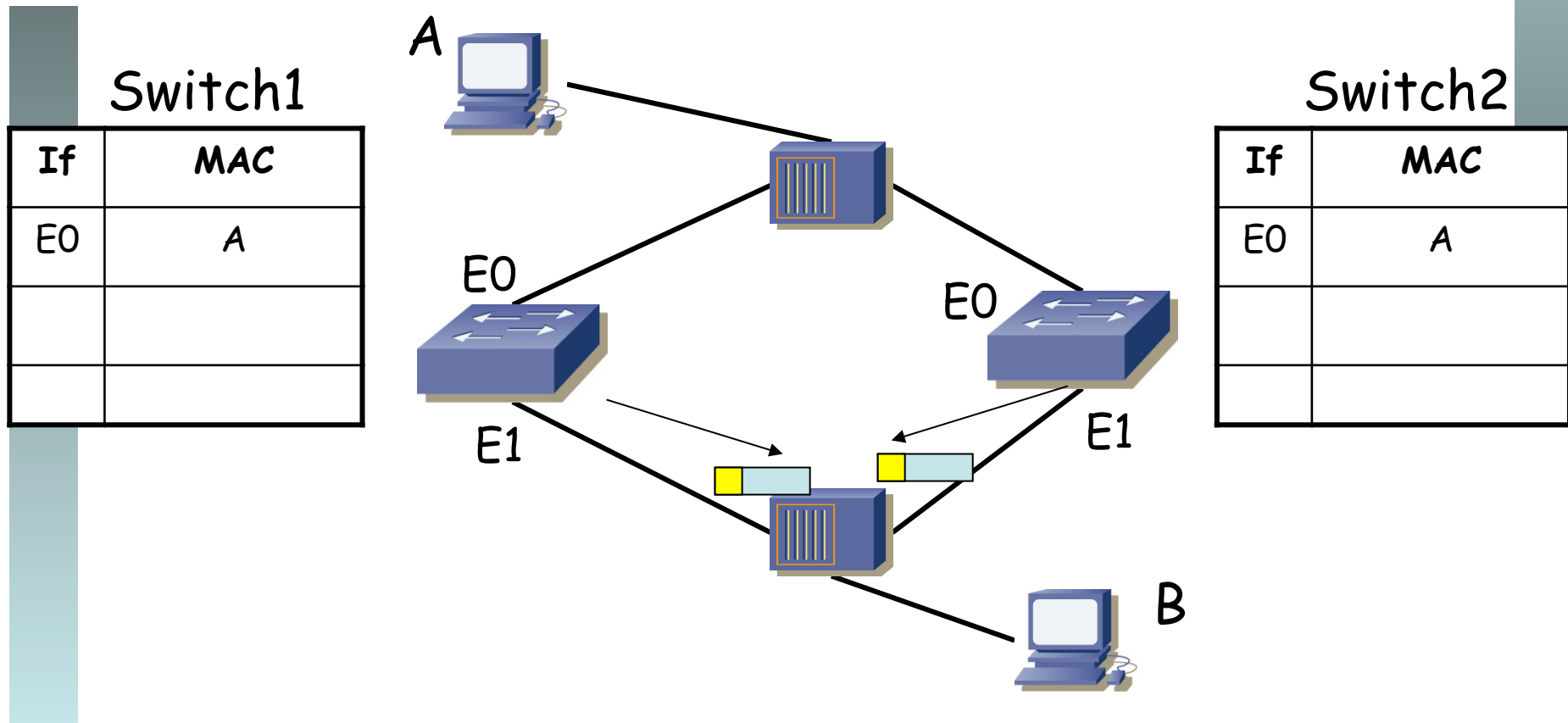
# Caminos alternativos

- Los conmutadores no conocen al destino
- Reenvían por todos los puertos menos por donde recibieron



# Caminos alternativos

- Host B recibe la trama
- Switch2 recibe la trama que envió Switch1
- Switch1 recibe la trama que envió Switch2

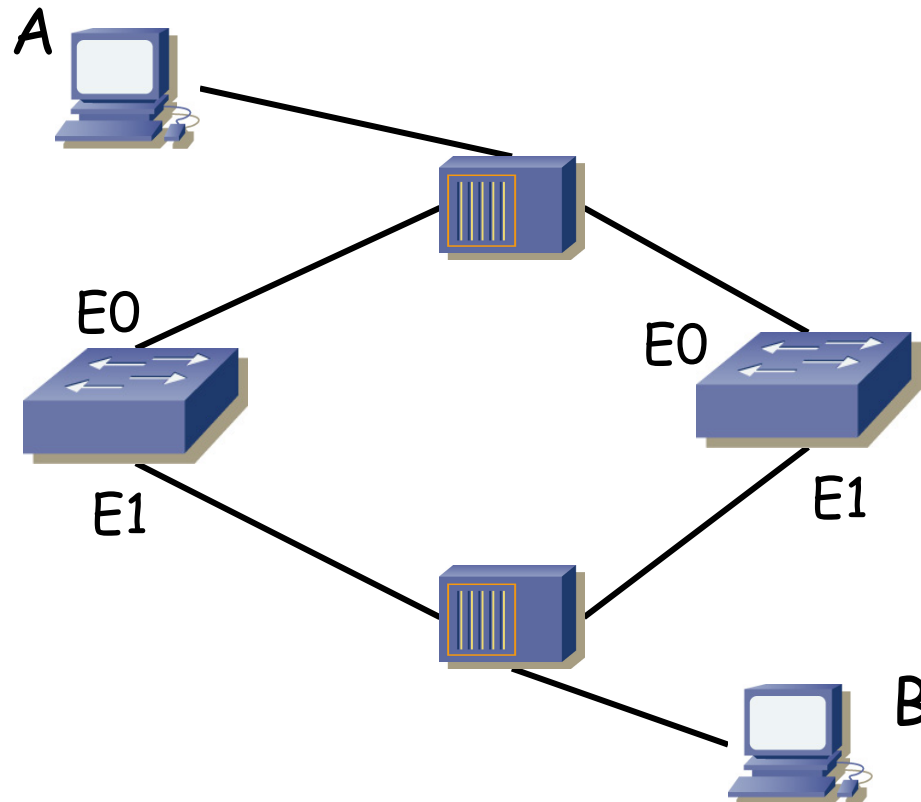


# Caminos alternativos

- Aprenden una nueva ubicación del host A

Switch1

If	MAC
E1	A

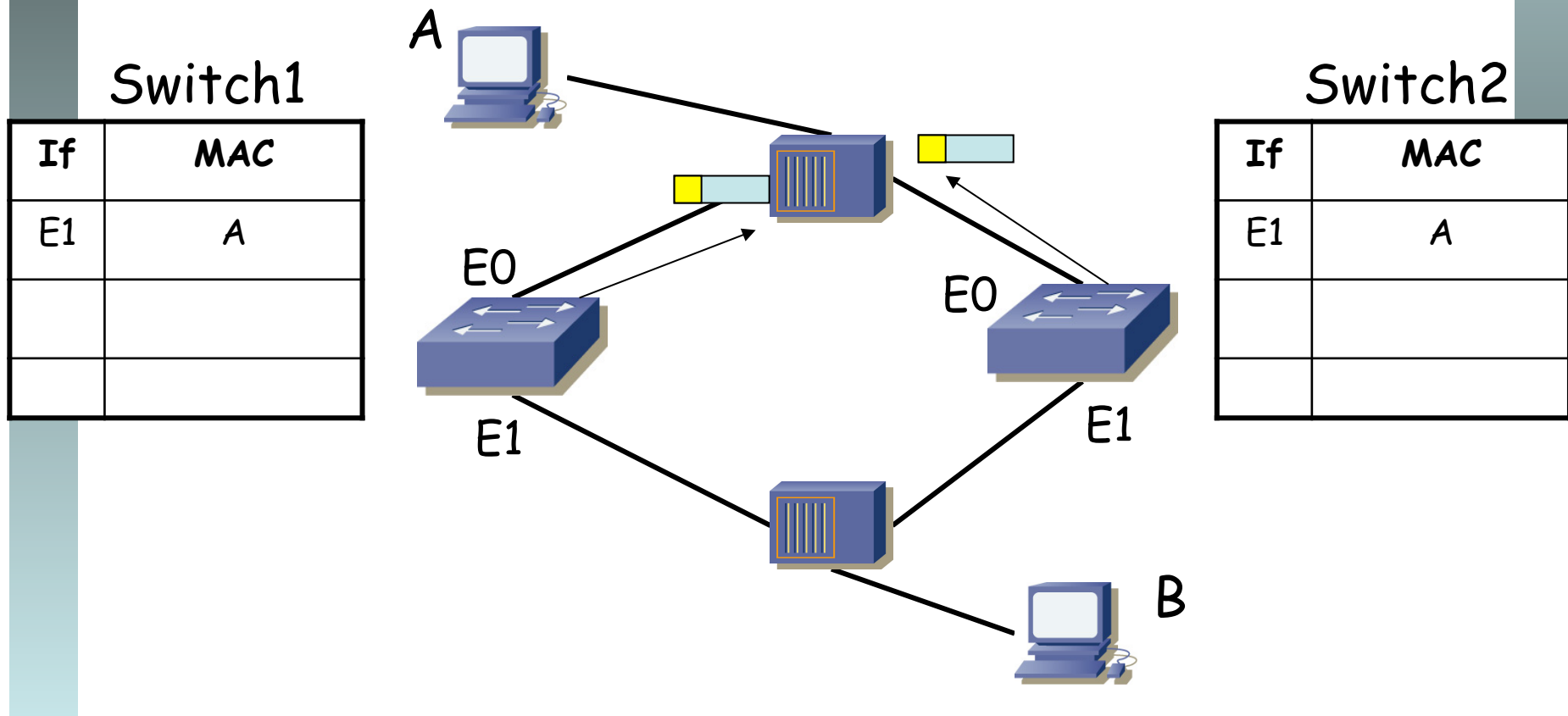


Switch2

If	MAC
E1	A

# Caminos alternativos

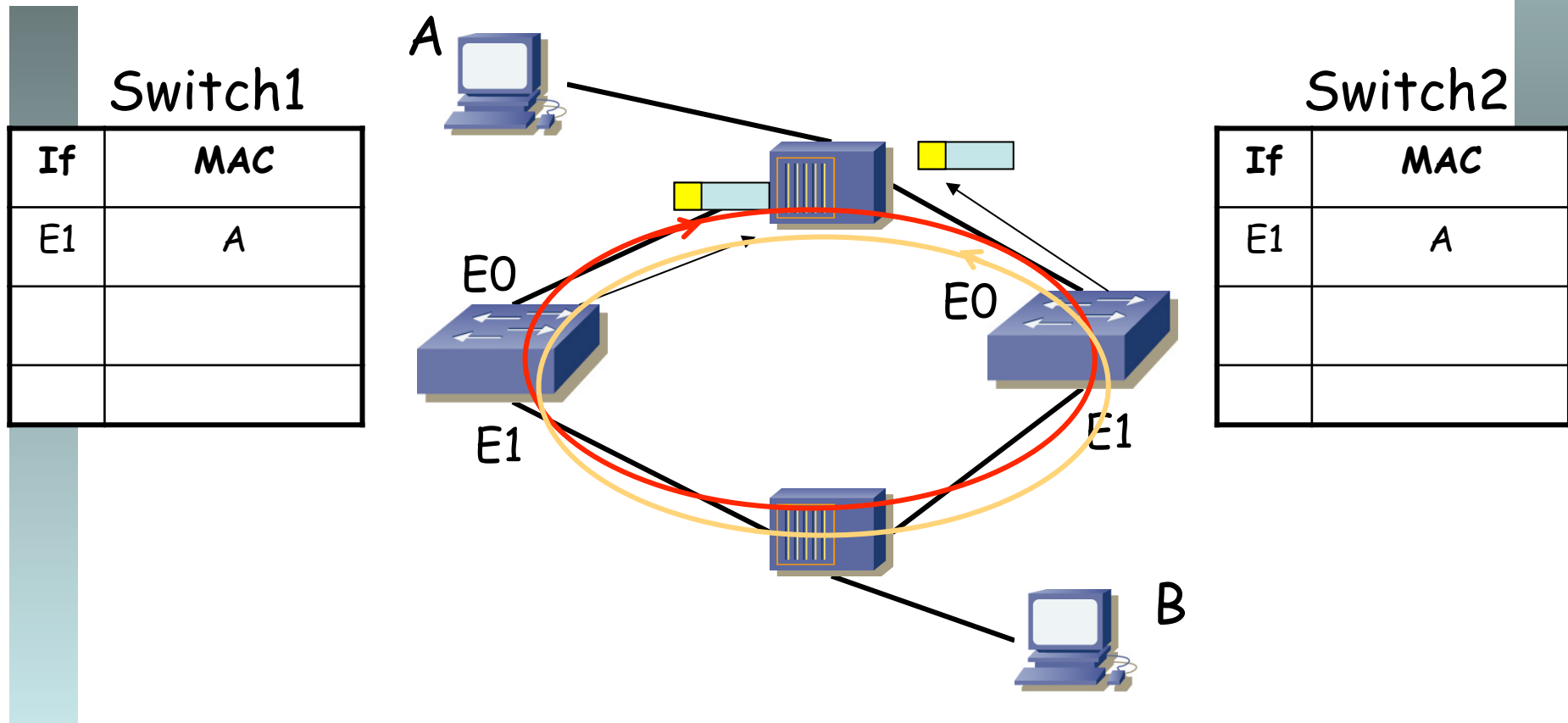
- Aprenden una nueva ubicación del host A
- Y reenvían por todos los puertos menos por donde recibieron la trama





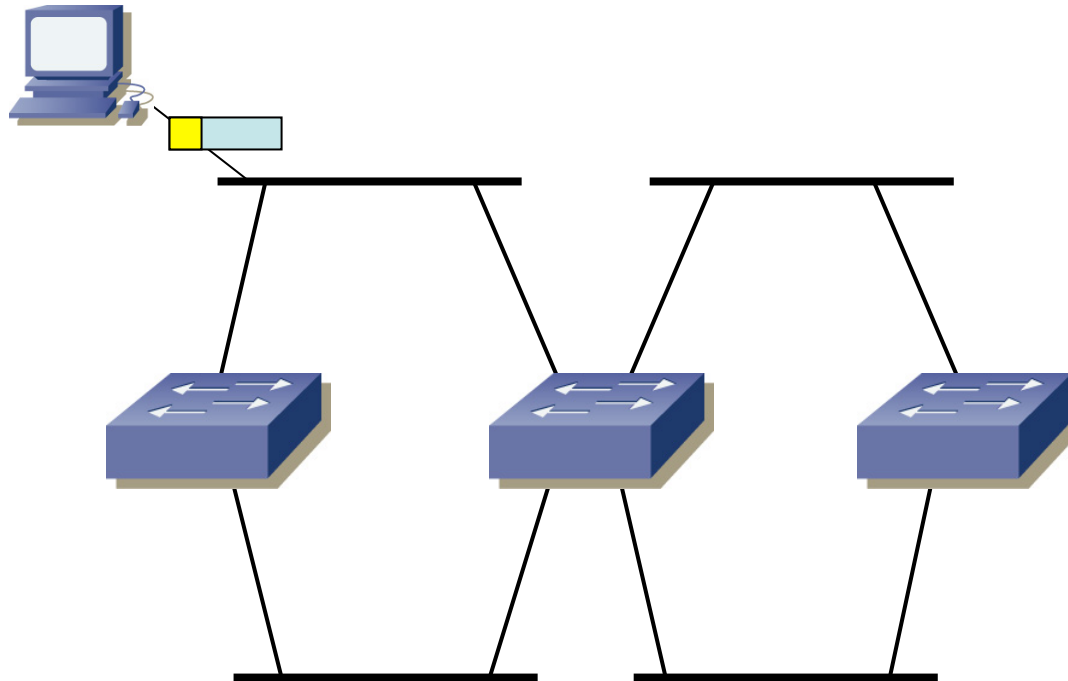
# Caminos alternativos

- Y se repite...
- No hay TTL en la trama Ethernet
- Además todos los hosts la deberían procesar



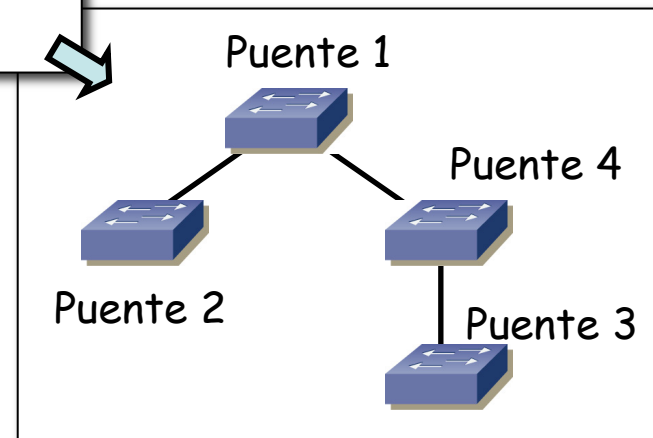
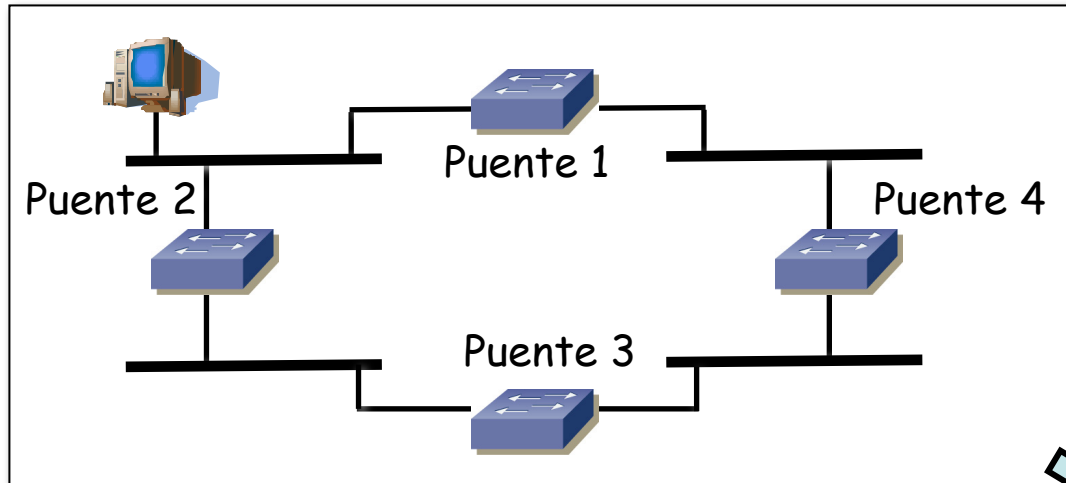
# Ejercicio

- PC envía trama de broadcast



# Spanning-Tree Protocol (STP)

- Calcula una topología libre de ciclos
- A partir del grafo de la topología crea un árbol
- Desactiva los enlaces sobrantes
- IEEE 802.1D

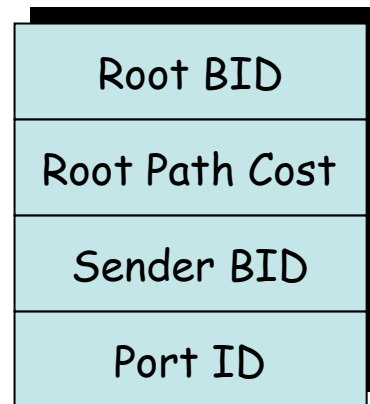
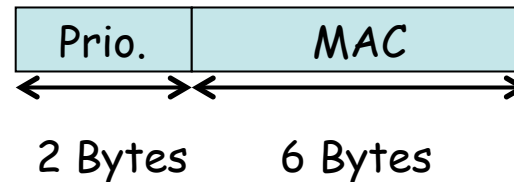


Radia Perlman (1983)

# Spanning-Tree Protocol (STP)

## BPDUs

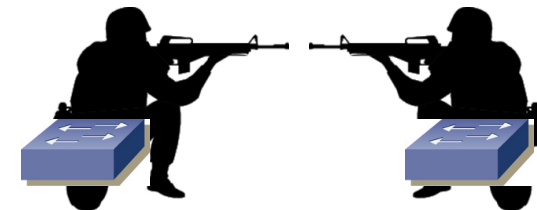
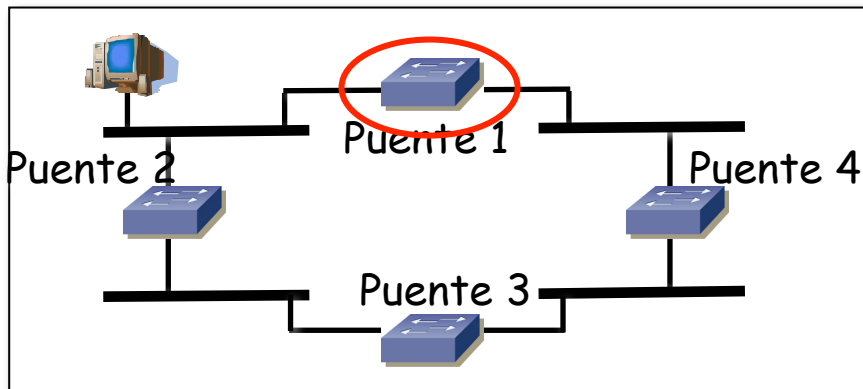
- Bridge Protocol Data Units
- Enviadas periódicamente por los puentes
- Destino 01:80:C2:00:00:00 (Bridge Group Address)
- No son reenviadas
- BID = Bridge ID
- Información importante:



# Spanning-Tree Protocol (STP)

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

- Raíz para el árbol
- 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



# Spanning-Tree Protocol (STP)

## Path Cost

- Asociado a cada LAN
- Según la velocidad
- Originalmente  $1000 / Velocidad(Mbps)$
- 802.1D-2004 :

Table 17-3—Port Path Cost values

Link Speed	Recommended value	Recommended range	Range
<=100 Kb/s	200 000 000 <sup>*</sup>	20 000 000–200 000 000	1–200 000 000
1 Mb/s	20 000 000 <sup>a</sup>	2 000 000–200 000 000	1–200 000 000
10 Mb/s	2 000 000 <sup>a</sup>	200 000–20 000 000	1–200 000 000
100 Mb/s	200 000 <sup>a</sup>	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

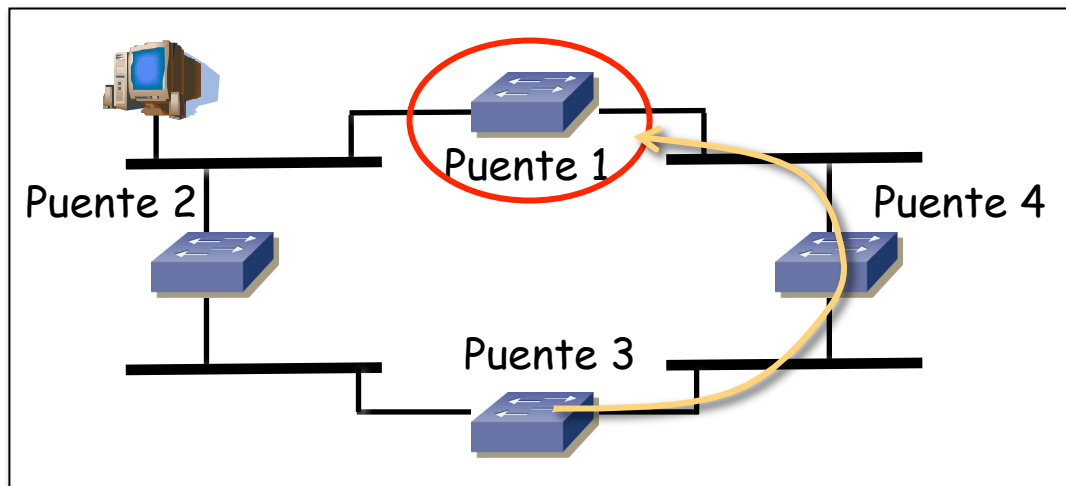
<sup>\*</sup>Bridges conformant to IEEE Std 802.1D, 1998 Edition, i.e., that support only 16-bit values for Path Cost, should use 65 535 as the Path Cost for these link speeds when used in conjunction with Bridges that support 32-bit Path Cost values.

- Se va agregando en un camino creando el *Root Path Cost*

# Spanning-Tree Protocol (STP)

## BPDUs

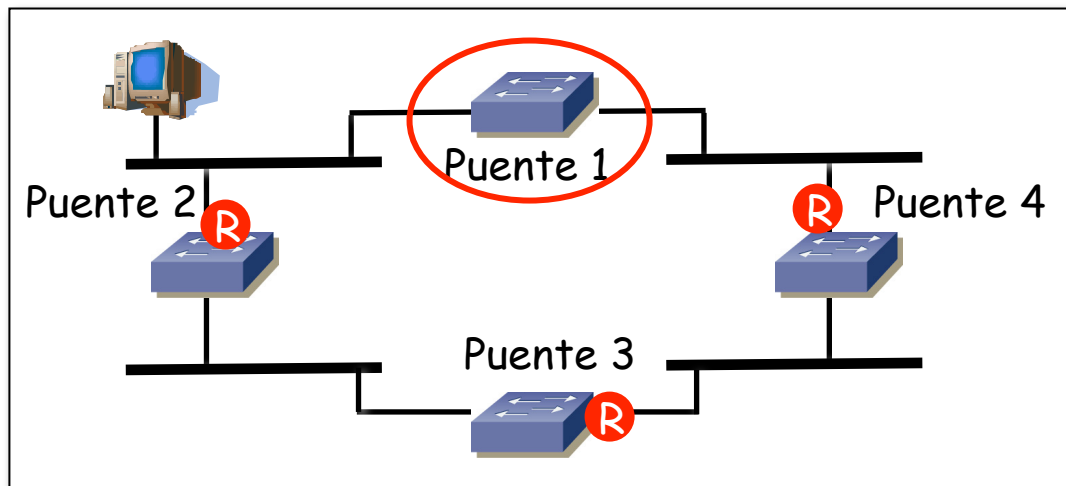
- Se pueden “comparar” entre si y decidir si una BPDU recibida por un puerto es “mejor” que otra
- “Mejor” en el sentido de “mejor” camino a la raíz
- Relacionado con el “coste” hasta la raíz y ocasionalmente con el puerto por el que se recibió
- Incluye dependencia con la velocidad de los tramos



# Spanning-Tree Protocol (STP)

## *Root Port*

- Puerto con menor *Root Path Cost*
- Puente raíz es el único sin un puerto raíz

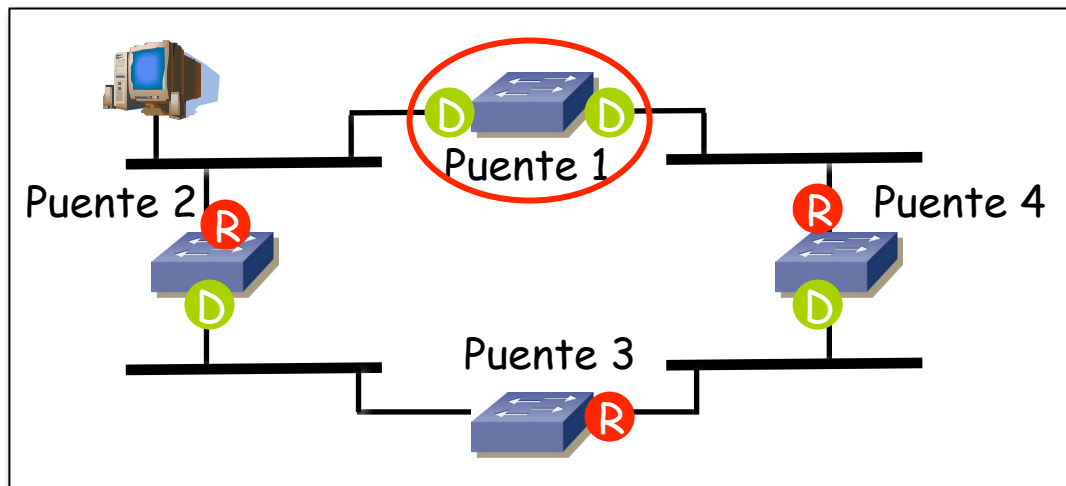




# Spanning-Tree Protocol (STP)

## *Designated Port*

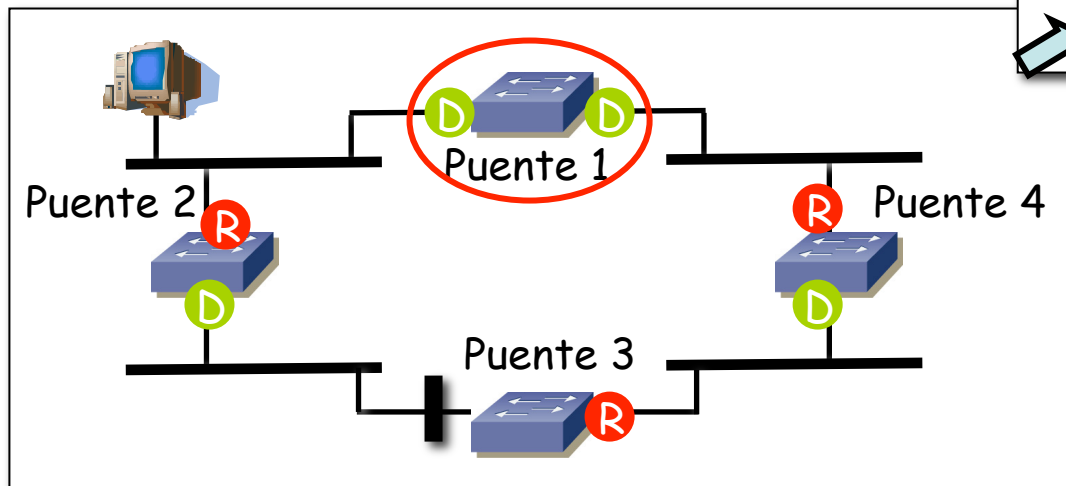
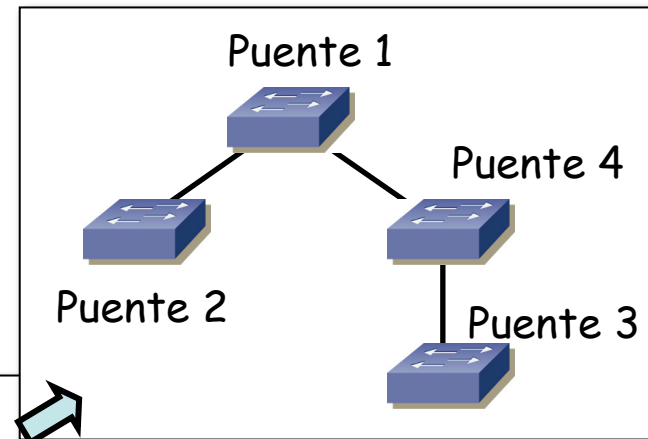
- Del puente conectado a una LAN con mejor camino hasta la raíz
- Uno por segmento



# Spanning-Tree Protocol (STP)

## ***Blocked Port***

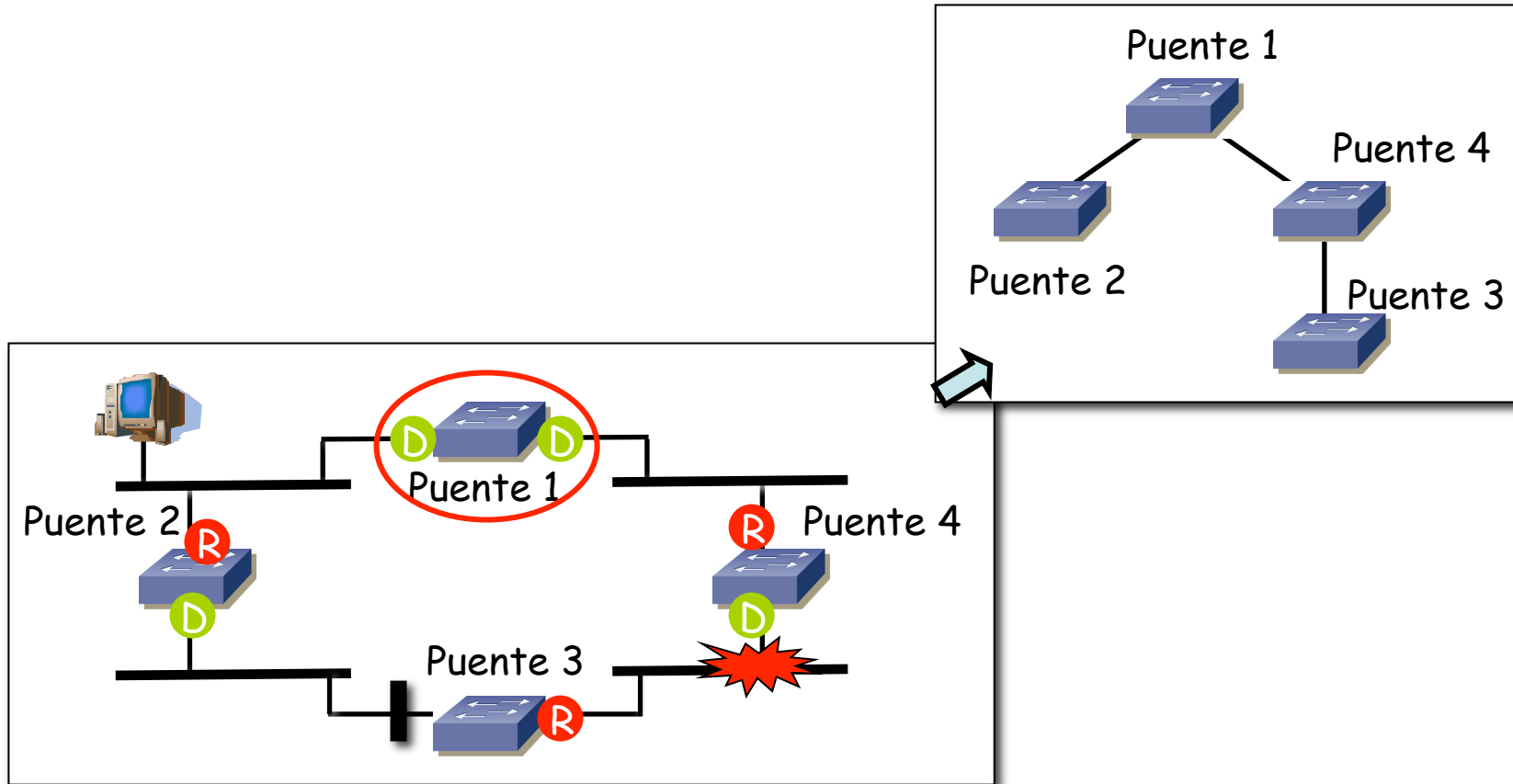
- No aprenden MACs ni reenvían tramas
- Se aceptan BPDUs
- Es un puerto *alternativo* o de *backup*
- Todos aquellos que ni son *Root* ni *Designated*



# Spanning-Tree Protocol (STP)

## Cambios en la topología

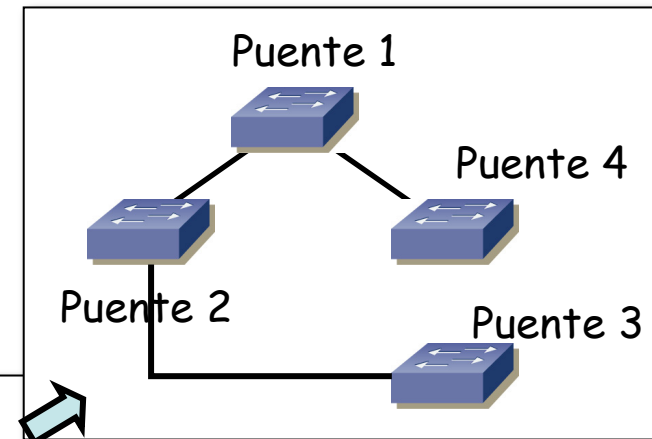
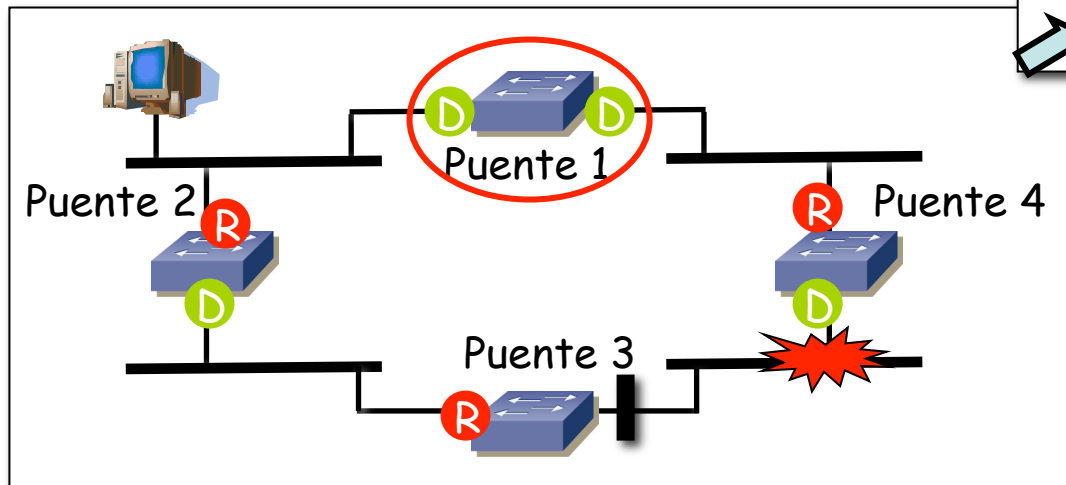
- Ante un fallo (...)
- Recalcular árbol (...)



# Spanning-Tree Protocol (STP)

## Cambios en la topología

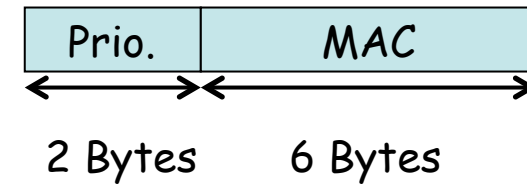
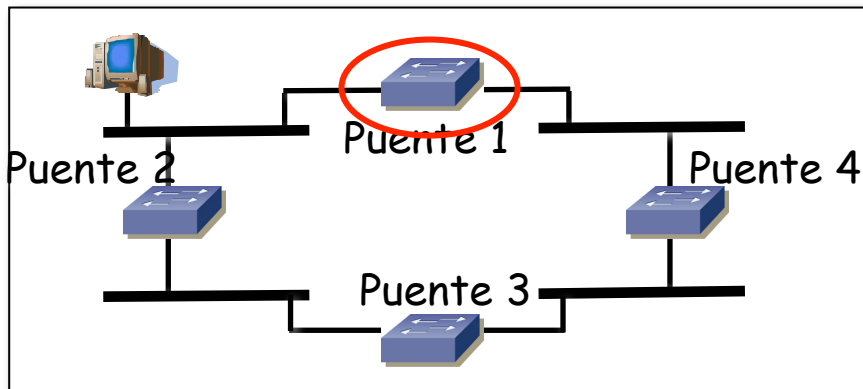
- Ante un fallo (...)
- Recalcular árbol (...)
- Tiempo de convergencia:  
30-60 segs



# Spanning-Tree Protocol (STP)

## Selección del puente raíz

- Por defecto todos 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



# RSTP

## Rapid Spanning-Tree Protocol

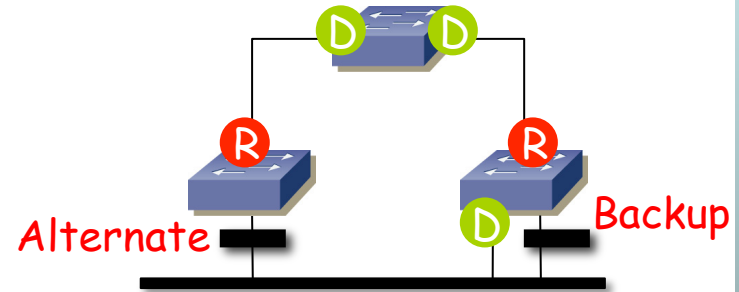
- IEEE 802.1w
- STP obsoleto
- RSTP en 802.1D-2004
- Tiempos de convergencia de 2-3 segs
- Tres estados posibles para un puerto: Discarding, Learning y Forwarding

STP Port State	RSTP Port State	Included in active topology?	Learning addresses?	MAC addresses?
Disabled	Discarding	No	No	No
Blocking	Discarding	No	No	No
Listening	Discarding	Yes	No	No
Learning	Learning	Yes	Yes	Yes
Forwarding	Forwarding	Yes	Yes	Yes

# RSTP

## Port Roles:

- *Root y Designated* (sin cambios)
- *Alternate y Backup*:
  - Corresponden a lo que antes eran *blocked port*
  - *Backup* es todo puerto que no es ni *Root* ni *Designated* y el puente es *Designated* para esa LAN (si no, es *Alternate*)
  - Un *Alternate port* da un camino alternativo hacia el root frente al puerto que se tiene como *Root*
  - *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
  - *Alternate* está bloqueado porque se han recibido BPDUs mejores (menor coste) de otro switch en el mismo segmento
  - *Backup* está bloqueado porque se han recibido BPDUs mejores **del mismo switch** en el mismo segmento



# RSTP

- Evita loops temporales cuando se producen fallos o retirada de equipos
- No protege ante loops temporales formados mediante repetidores
- Se pueden configurar puertos como *edge* para que pasen inmediatamente al estado *forwarding*
- Pueden coexistir en la LAN puentes que implementen STP y RSTP



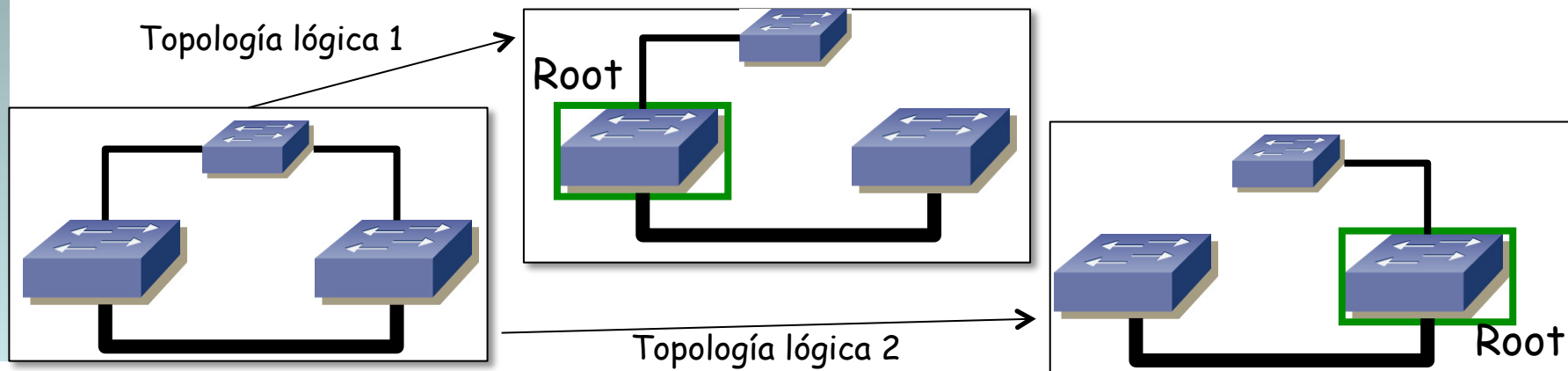
# VLANs y Spanning Trees

## Solución básica:

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

## MSTP

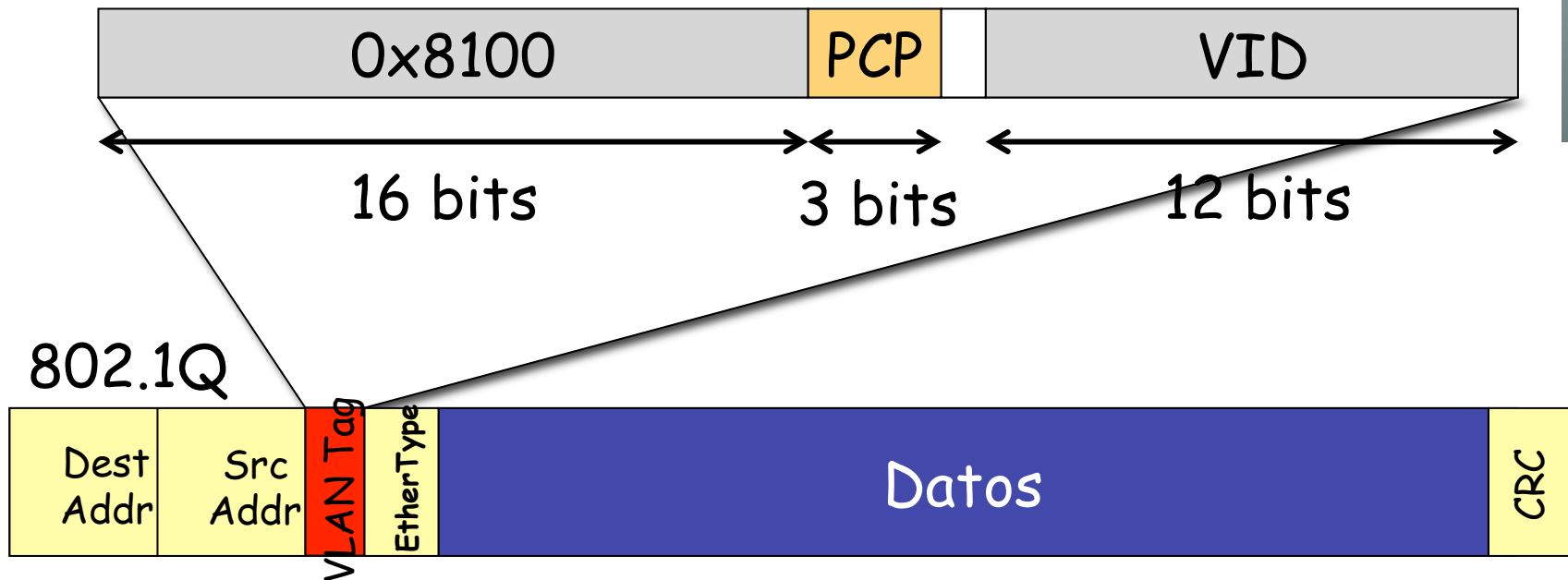
- MSTP = Multiple Spanning Tree Protocol (modificación 802.1s a 802.1Q)
- Un ST por *grupo* de VLANs (que puede ser de una)
- Una topología lógica por VLAN o por grupo de VLANs
- Para cada grupo se pueden cambiar parámetros de ST, por ejemplo la prioridad para cambiar el Root Bridge
- Ejemplo: topología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



# Otros estándares y modificaciones de Ethernet

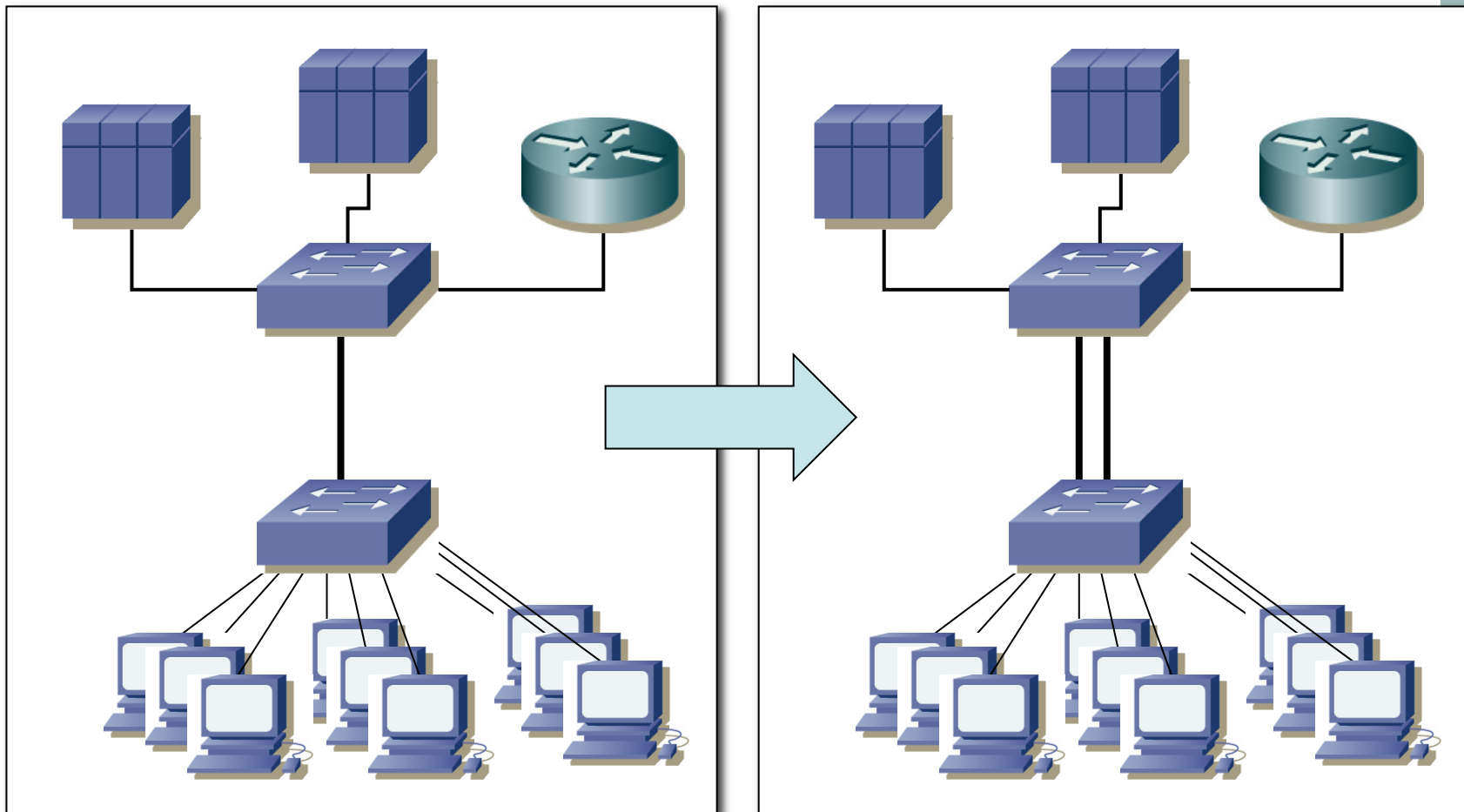
# Frame Priority

- IEEE 802.1p (ahora parte de 802.1D)
- Classes of Service (CoS)
- PCP = Priority Code Point
- Permite aplicar técnicas de planificación
- Si VID=0 solo se indica la prioridad



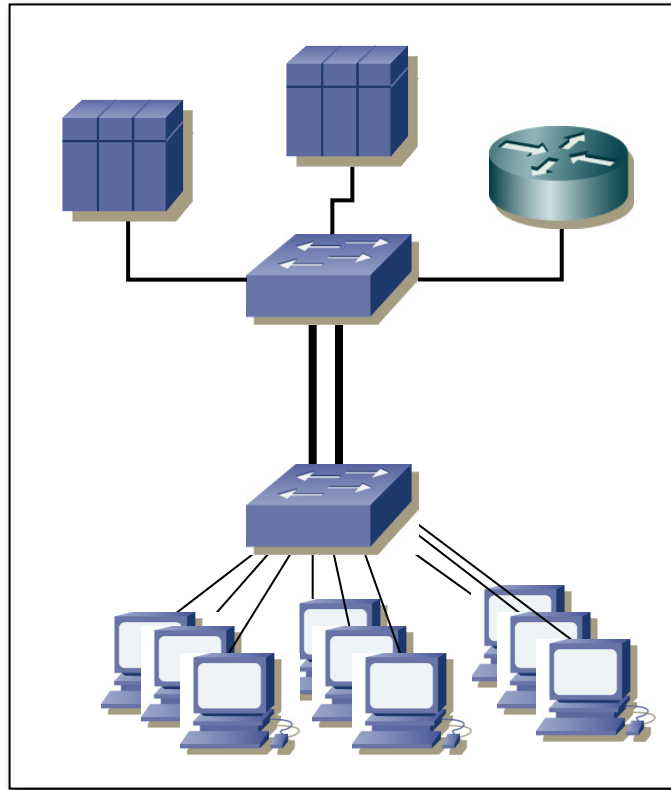
# Link Aggregation

- IEEE 802.3ad
- Ahora 802.1AX



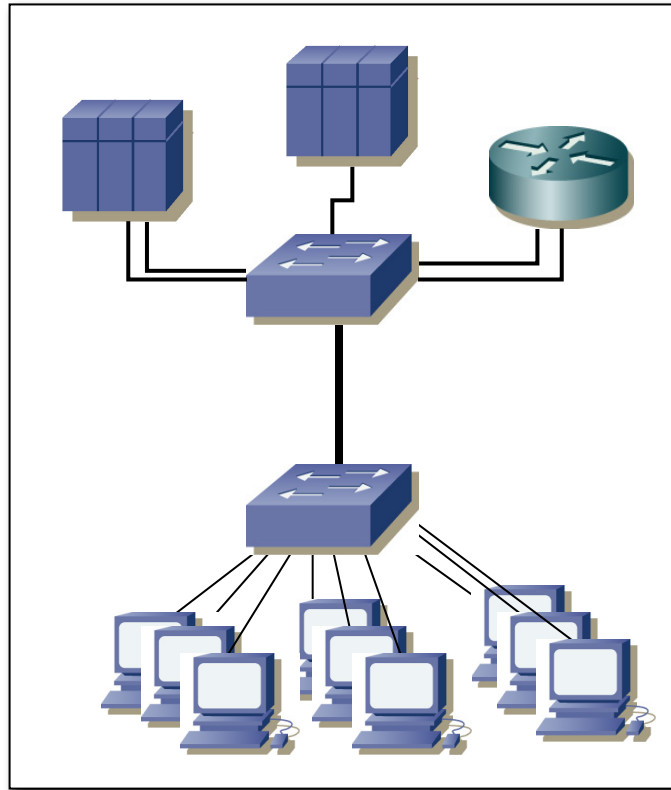
# Link Aggregation

- Tipos de agregación:
  - *Switch-to-switch*



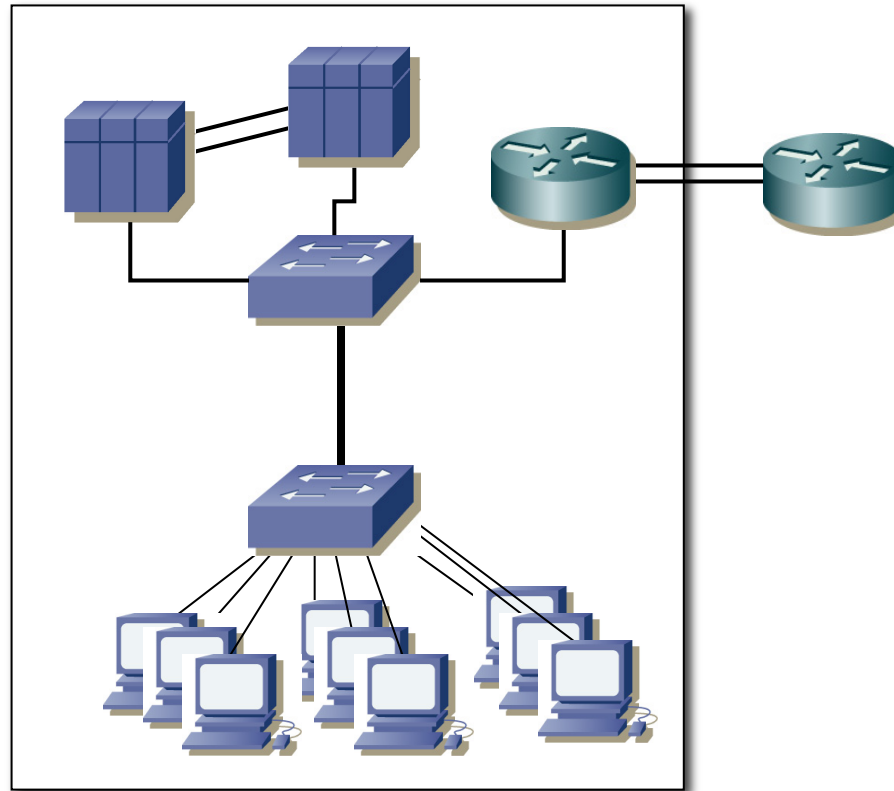
# Link Aggregation

- Tipos de agregación:
  - *Switch-to-station*



# Link Aggregation

- Tipos de agregación:
  - *Station-to-station*



# Link Aggregation

- Implementado entre el subnivel MAC y el LLC
- Los enlaces se agregan en Grupos
- El agregado: como un solo interfaz
- Conversación: tramas de la misma MAC → MAC y prioridad
- Mantiene el orden de las tramas de la misma conversación
- ¿Cómo? (...)

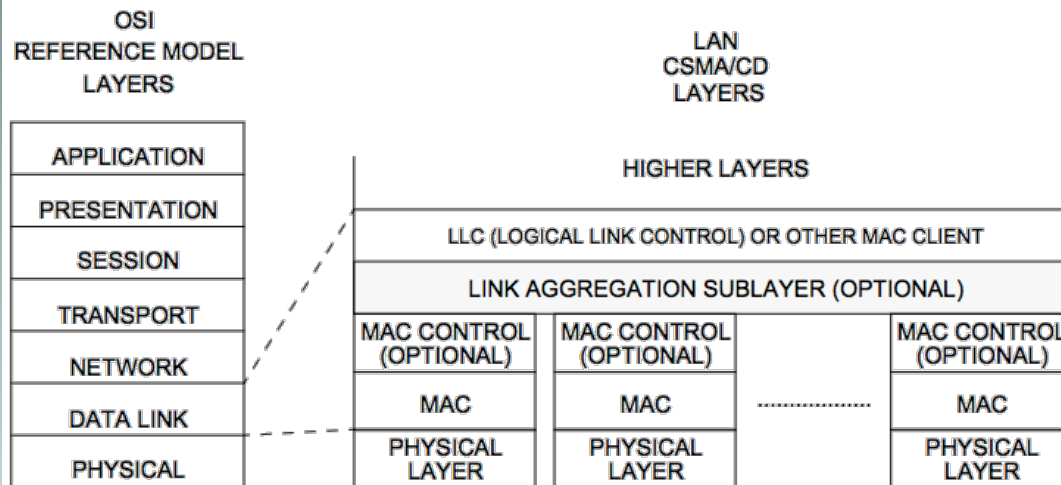
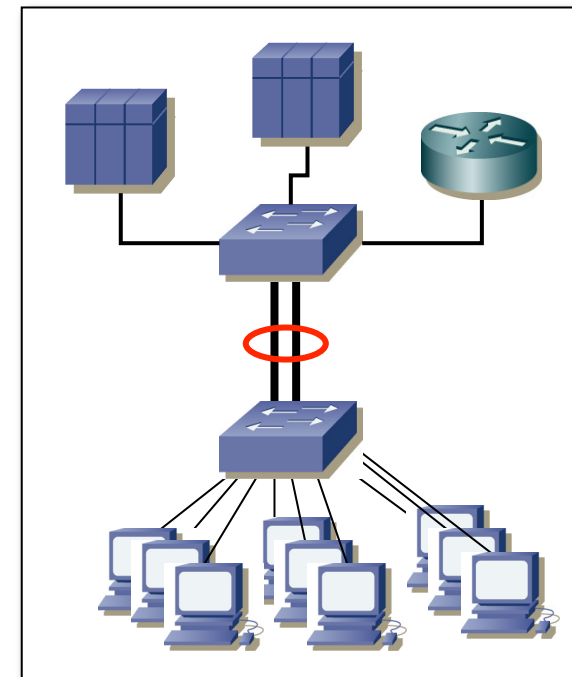


Figure 5-2—Architectural positioning of Link Aggregation sublayer

IEEE 802.1AX





# Link Aggregation

- Implementado entre el subnivel MAC y el LLC
- Los enlaces se agregan en Grupos
- El agregado: como un solo interfaz
- Conversación: tramas de la misma MAC → MAC y prioridad
- Mantiene el orden de las tramas de la misma conversación
- ¿Cómo? Mandándolas siempre por el mismo enlace del grupo

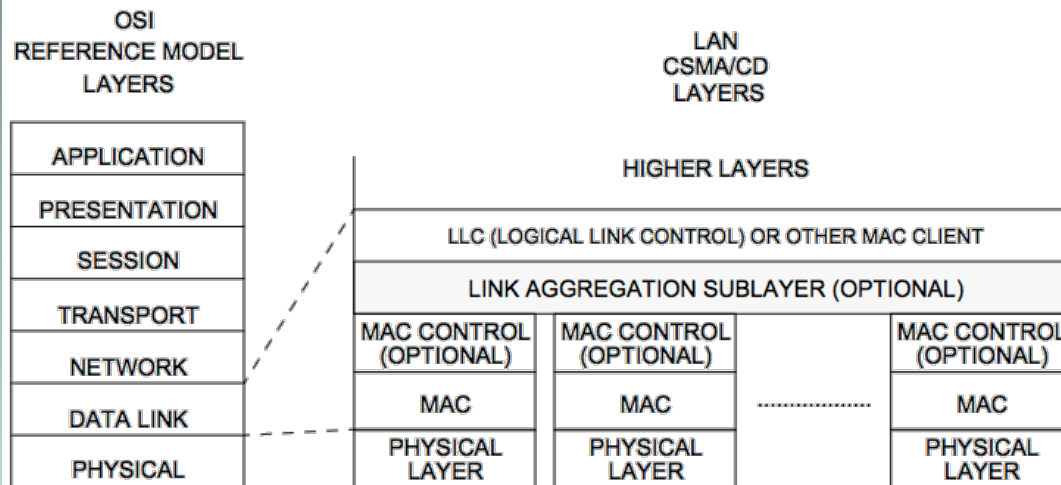
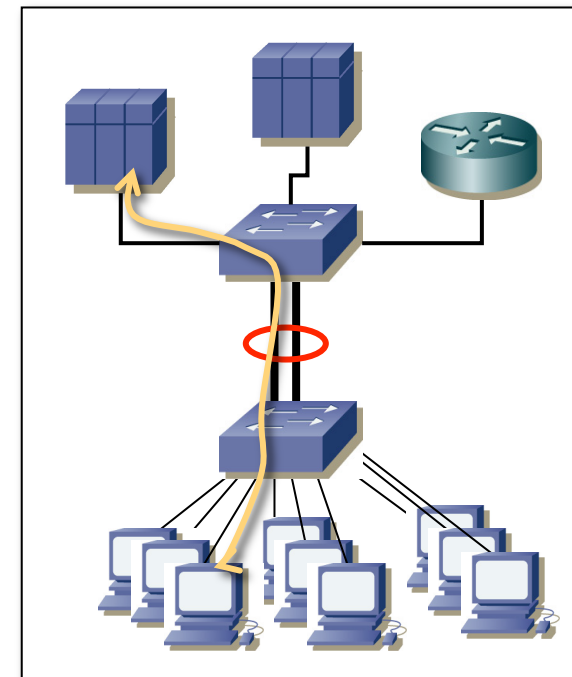


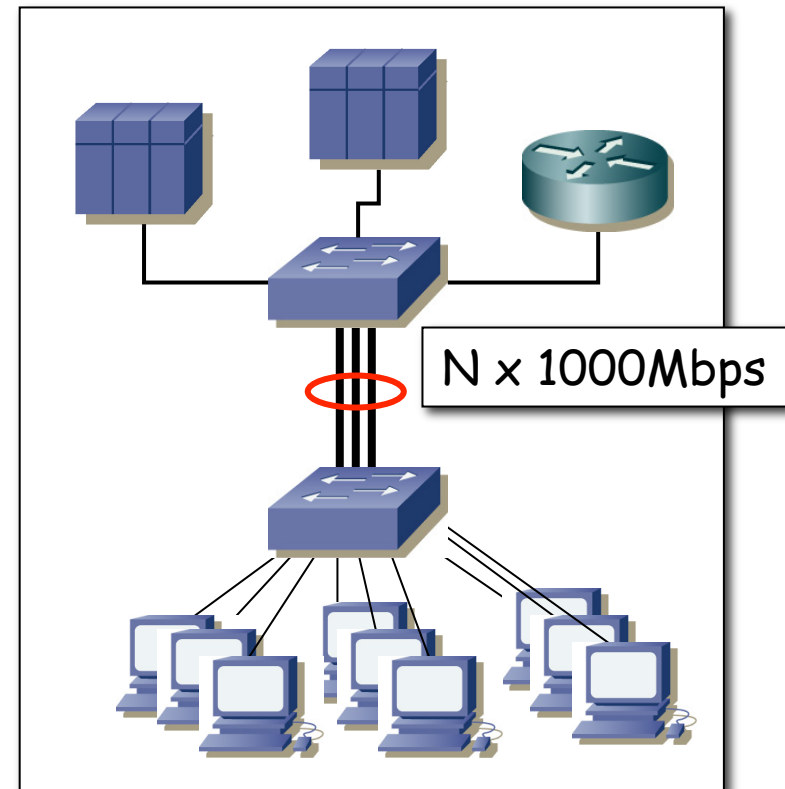
Figure 5-2—Architectural positioning of Link Aggregation sublayer

IEEE 802.1AX



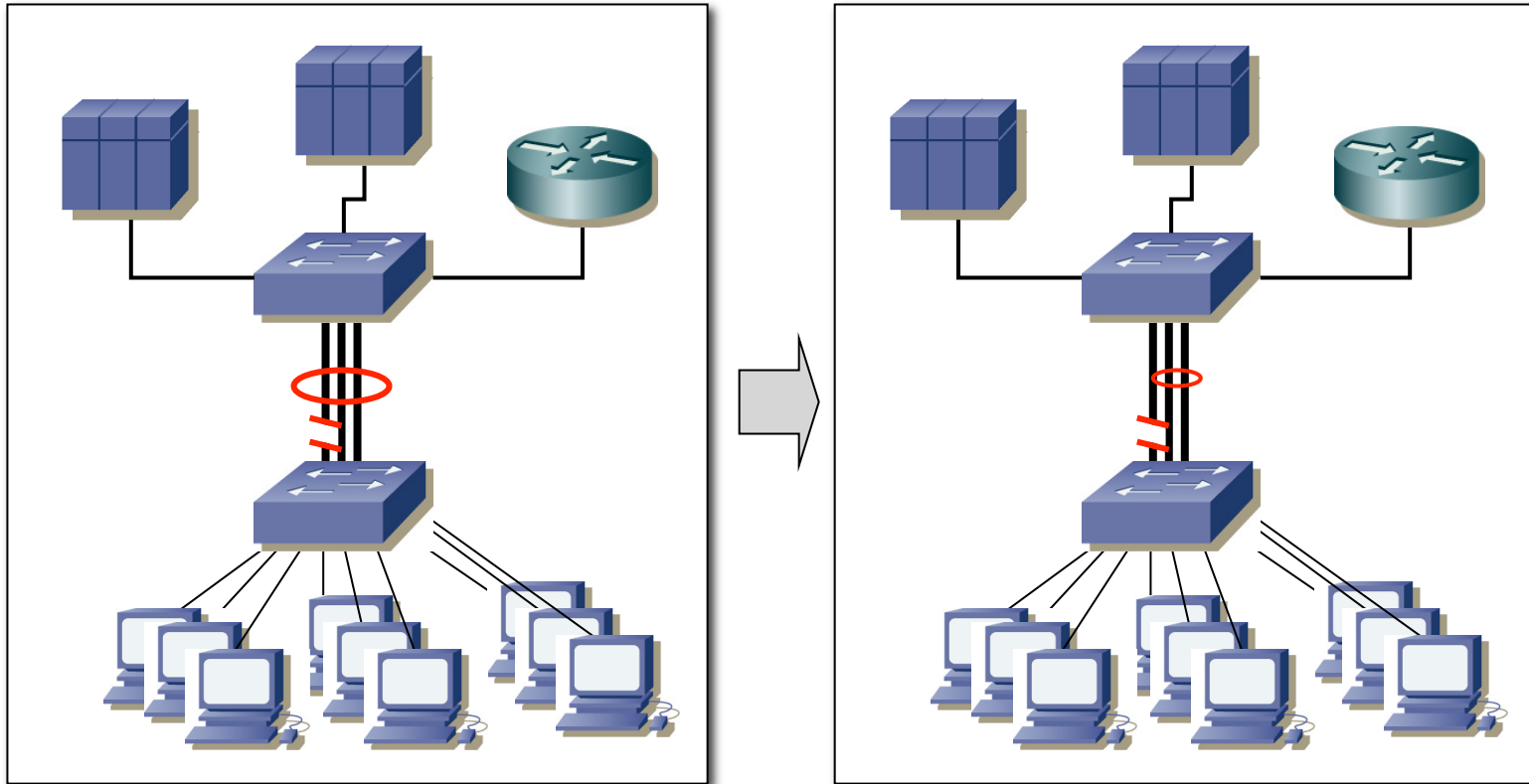
# Link Aggregation

Mayor ancho de banda



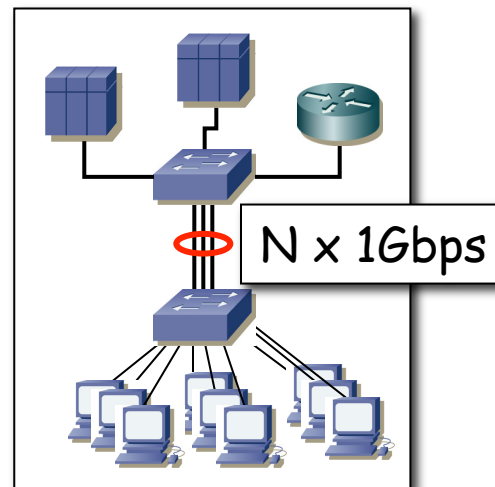
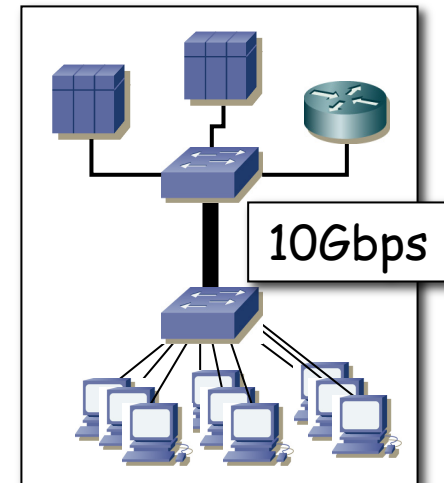
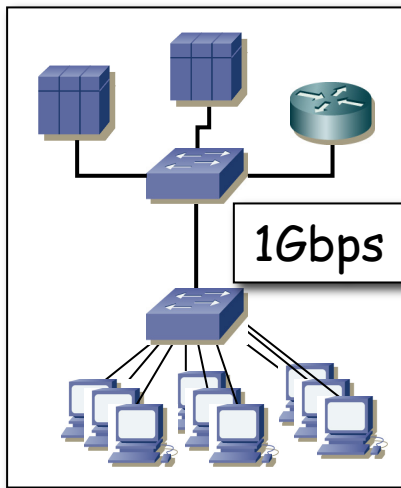
# Link Aggregation

## Mayor disponibilidad



# Link Aggregation

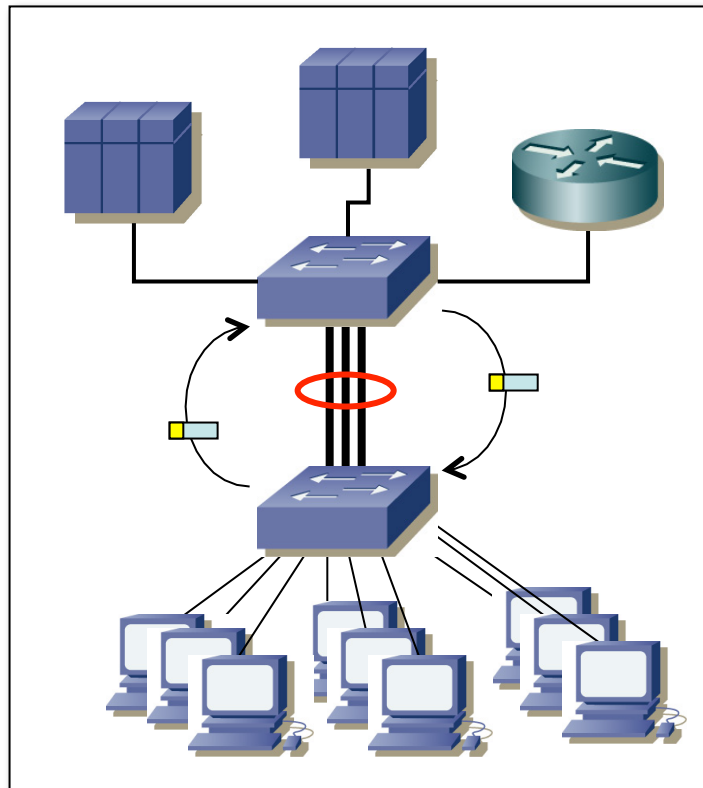
## Mayor granularidad



# Link Aggregation

## Configuración automática

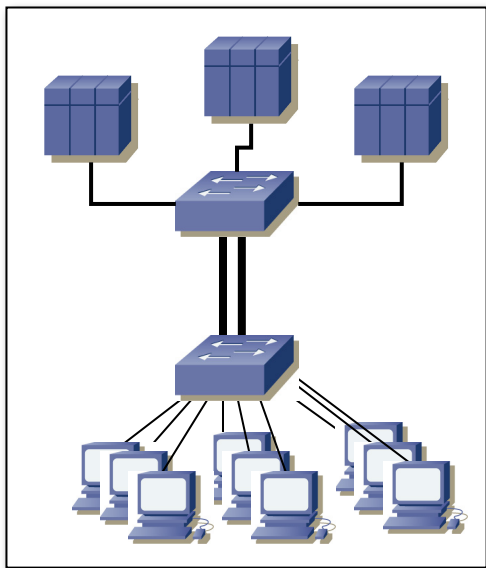
- Link Aggregation Control Protocol (LACP)
- Rápida reconfiguración (<1seg)



# Link Aggregation

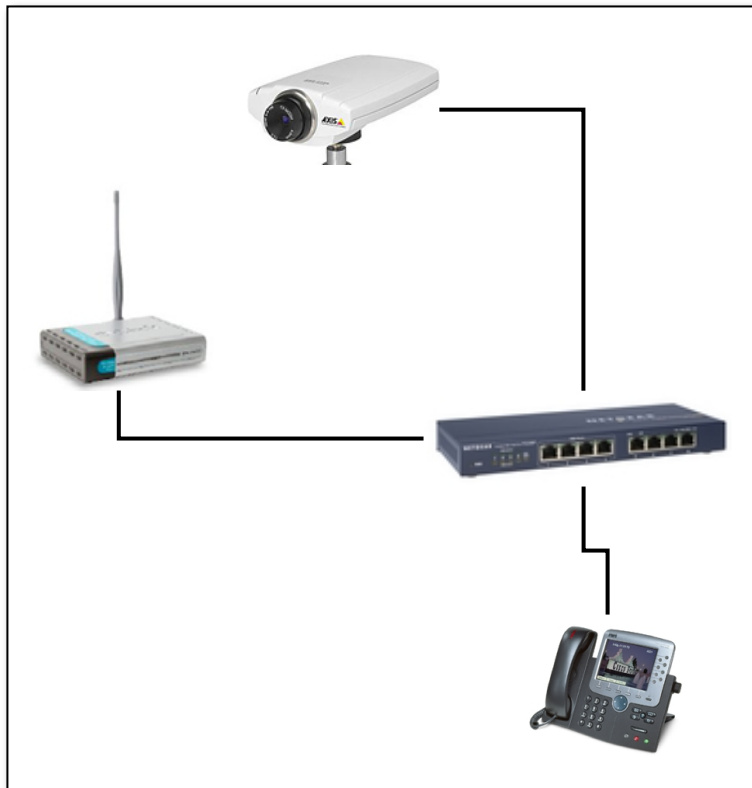
## Limitaciones:

- No más de 2 extremos
- Solo 802.3 (ej. no FDDI)
- No soporta enlaces half-duplex
- No puede agregar enlaces de diferentes velocidades
- Si la conversación es directamente  $A \rightarrow B$  no puede repartir el flujo en más de un enlace (salvo con información de niveles superiores)

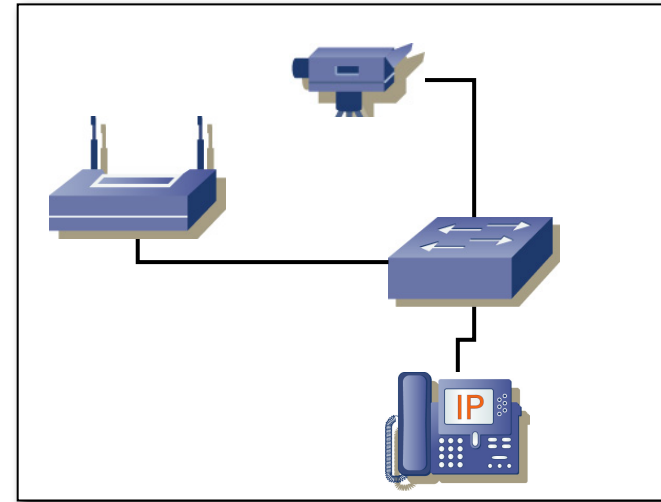


# Power over Ethernet (PoE)

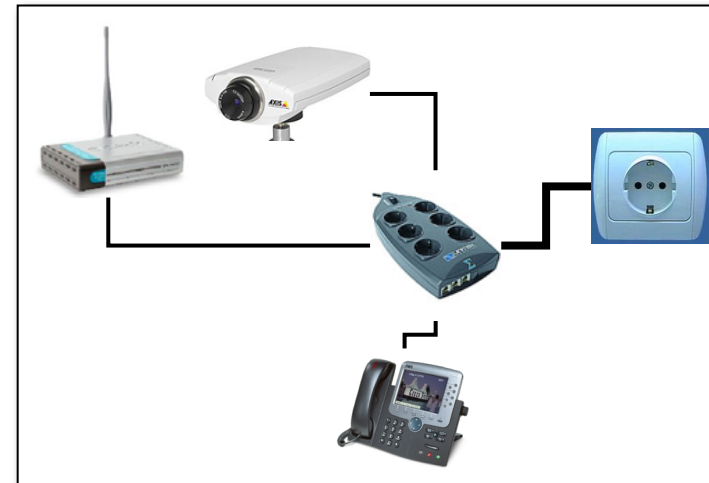
- IEEE 802.3af (cláusula 33 de 802.3-2005)
- Además de datos, corriente sobre cable Cat.3 ó 5 (15.4 W)
- PoE+ (802.3at) (30W)



=



+



# 802.3x

- Flow-control en el nivel de enlace
- Para enlaces full-duplex
- Cuando receptor cerca de saturación de buffer
- Envía trama Ethernet PAUSE que indica al otra extremo que detenga toda transmisión
- Son tramas de control MAC (Ethertype 0x8808)
- Emplea un campo de 2 bytes para indicar el tiempo de pausa
- Una unidad de pausa es el tiempo de transmisión de 512 bits
- Tramas enviadas a MAC multicast reservada (01:80:C2:00:00:01) que no son reenviadas

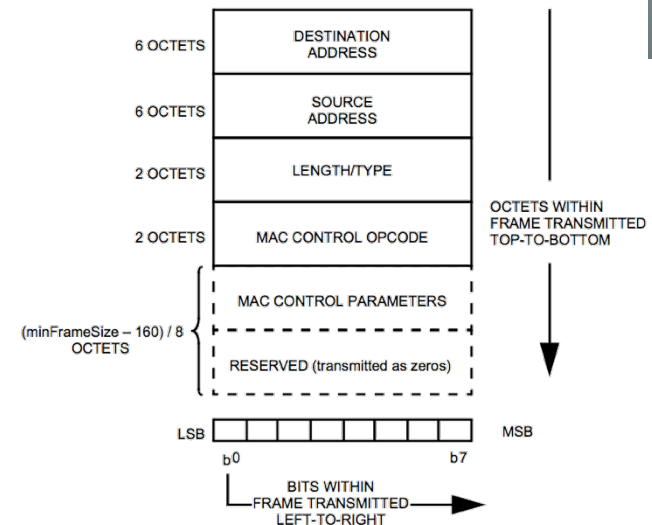
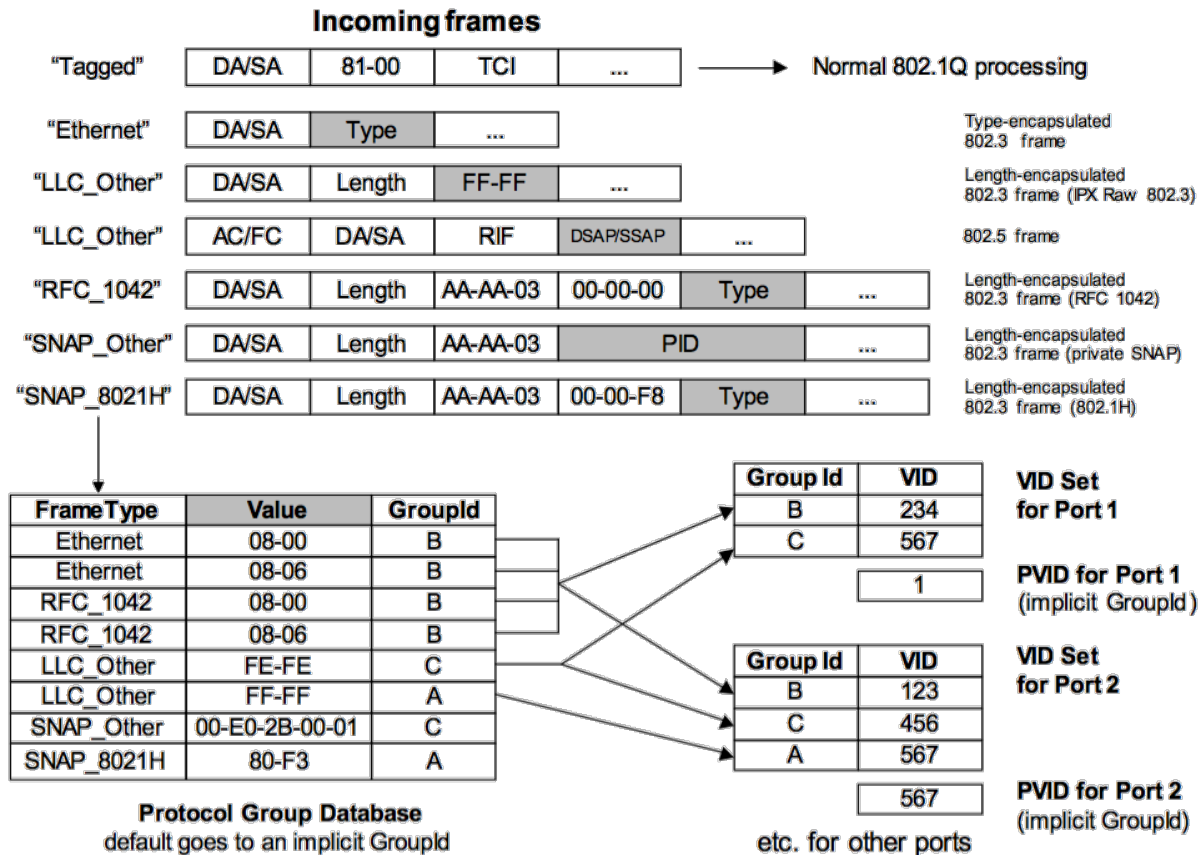


Figure 31-3—MAC Control frame format



# 802.1v

- Modificación a 802.1Q (ya es parte de él)
- Se permite etiquetar tramas entrantes en base a Puerto y “Protocolo”
  - “Protocolo” = tipo de encapsulado (Ethernet, RFC\_1042, SNAP\_8021H, etc)
  - Un segmento puede tener varias VLANs empleando encapsulado nativo (no 802.1Q o solo para prioridad con VID=0) distinguidas en base al “protocolo”



# GARP (802.1D-2004)

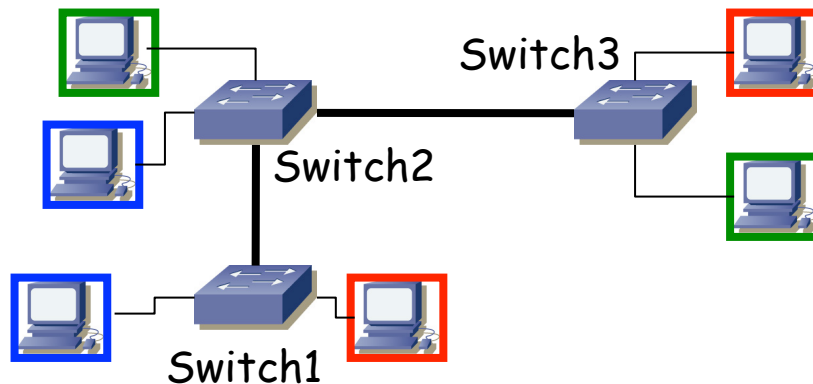
- Generic Attribute Registration Protocol
- Permite a los participantes registrar atributos en otros participantes en la Bridged LAN
- Los puentes distribuyen estos registros
- Los participantes pueden comprobar el estado de registros en un dispositivo
- Emplea LLC Type 1
- Se han reservado direcciones MAC 01:80:C2:00:00:20 hasta 01:80:C2:00:00:2F para diferentes aplicaciones GARP que se vayan definiendo
- El puente que implemente esas aplicaciones no reenviará directamente esas tramas sino que las procesará
- El puente que no las implemente las reenviará como cualquier otra trama multicast
- GMRP usa 01:80:C2:00:00:20
- GVRP usa 01:80:C2:00:00:21

# GMRP

- GMRP = GARP Multicast Registration Protocol
- GMRP ofrece mecanismos para que los hosts y puentes registren que desean recibir un flujo multicast
- Empleando estos registros y filtrado se limita el multicast a los segmentos que llevan a máquinas que han mostrado interés
- Los grupos son abiertos = cualquiera puede enviar a él
- Hosts pueden acceder a esta info de registro y si son fuente del grupo decidir no enviar si no hay destinatarios declarados (*source pruning*)

# GVRP

- GARP VLAN Registration Protocol
- Permite registrar de forma dinámica que por un puerto se puede alcanzar a hosts de una VLAN
- Dispositivos que soportan GVRP pueden conocer las VLANs que tienen miembros y a través de qué puertos se les alcanza
- Permite que un switch conozca la existencia de una VLAN y conmute tramas de la misma sin necesidad de crear esa VLAN en el switch ni asignar puertos manualmente
- Ejemplo 1: Switch 2 aprende mediante GVRP la existencia de la VLAN roja y sus puertos a Switch 1 y 3 reenviarán sus tramas
- Ejemplo 2: Switch 1 no tiene hosts de VLAN verde. Switch 2 no necesita reenviarle *floodings* a esa VLAN (broadcast o búsqueda de host)



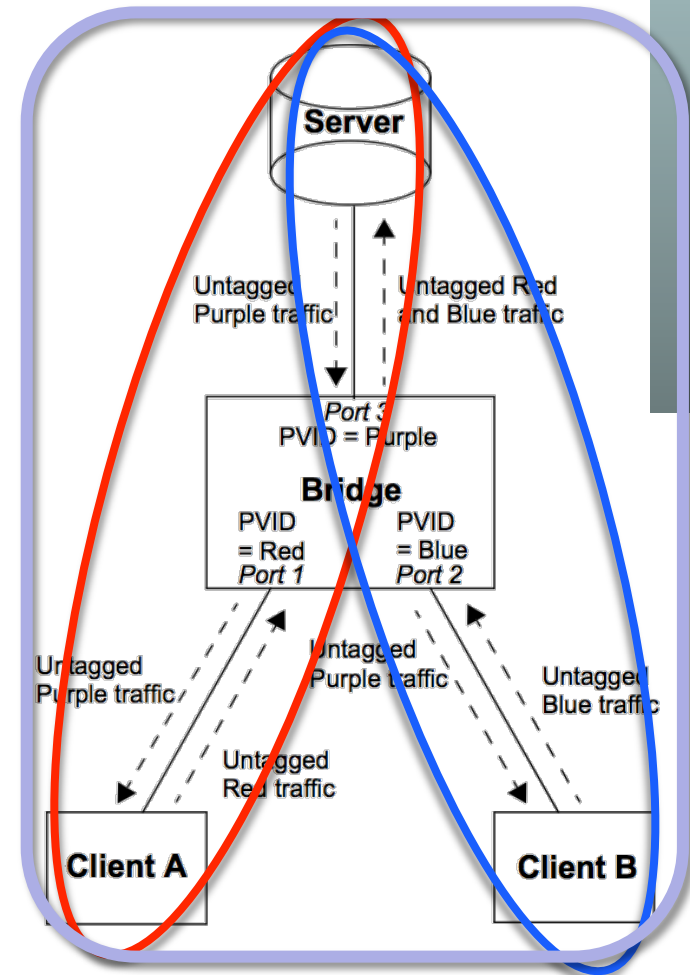
# 802.1ak-2007

- Modifica a 802.1Q-2005
- Para escenarios MAN/WAN
- Reemplaza GARP y sus aplicaciones
- En redes grandes GVRP y GMRP tardan demasiado en recuperación ante fallos
- MRP (Multiple Registration Protocol) sustituye a GARP
- MVRP (Multiple VLAN Registration Protocol) sustituye a GVRP
- MMRP (Multiple MAC Registration Protocol) sustituye a GMRP

# Asymmetric VLAN

## Ejemplo (801.Q-2005 Anexo B.1.3)

- Cliente A transmite al servidor por Puerto 1 que aprende su MAC en VLAN Roja
- Servidor transmite respuesta a A por Puerto 3 que la clasifica como VLAN Púrpura
- VLANs Roja y Púrpura comparten información (SVL = Shared VLAN Learning)
- Trama de la VLAN Púrpura sale por Puerto 1 hacia el Cliente A
- Todo tramas sin etiqueta de VLAN
- Es como si la VLAN Púrpura existiera en los 3 puertos (...)
- Es decir, varias VLAN se utilizan en el mismo puerto sin etiquetado 802.1Q



# Resumen

- (R)STP
- RSTP y VLANs
- Prioridades (802.1p)
- Agregación de enlaces (802.1ad)
- PoE (802.3af)
- 802.3x
- GARP/GMRP/GVRP
- MRP/MMRP/MVRP