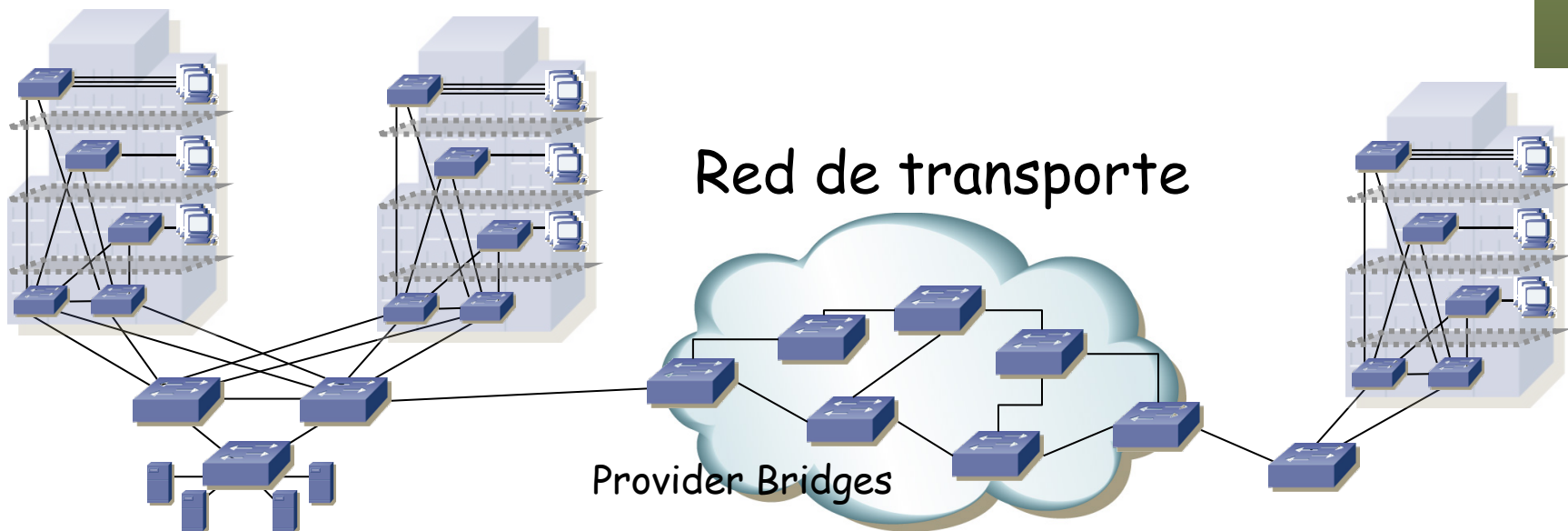


Provider (Backbone) Bridges

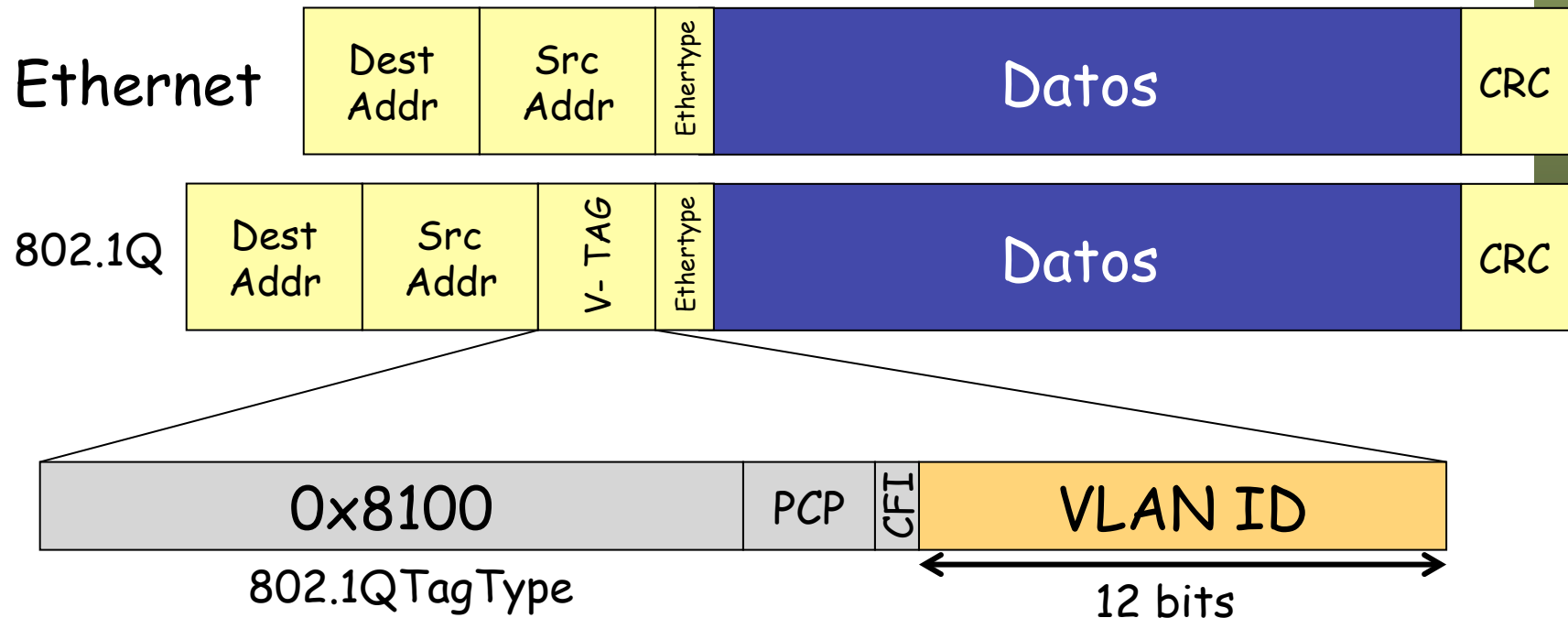
Provider Bridges

- El objetivo es que un proveedor pueda transportar tráfico Ethernet mediante una MAN/WAN Ethernet (...)
- (...)



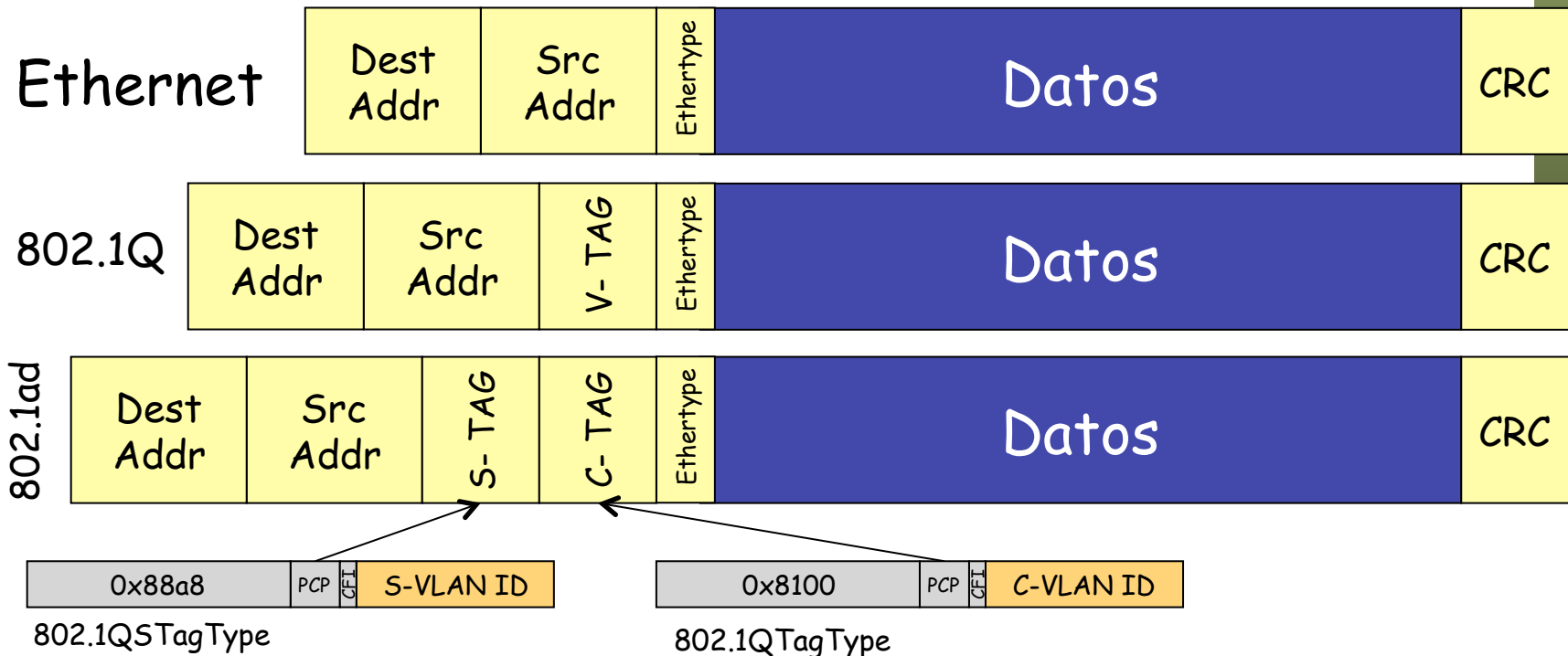
Provider Bridges

- El objetivo es que un proveedor pueda transportar tráfico Ethernet mediante una MAN/WAN Ethernet (...)
- Con 802.1Q el proveedor puede emplear tags de VLAN para diferenciar usuarios
- Pero esto impide transportar tráfico *tagged*
- 802.1ad (modificación a 802.1Q-2005) permite diferenciar entre las VLANs del cliente (C-VLAN) y las del servicio (S-VLAN) (...)



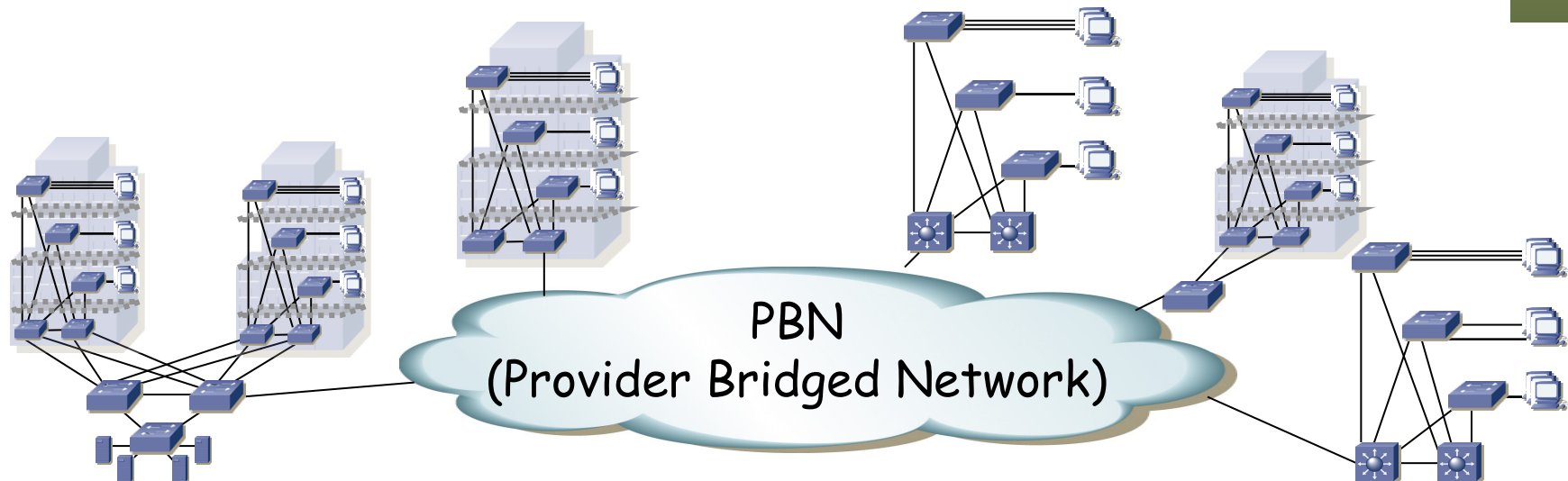
Provider Bridges

- El objetivo es que un proveedor pueda transportar tráfico Ethernet mediante una MAN/WAN Ethernet (...)
- Con 802.1Q el proveedor puede emplear tags de VLAN para diferenciar usuarios
- Pero esto impide transportar tráfico *tagged*
- 802.1ad (modificación a 802.1Q-2005) permite diferenciar entre las VLANs del cliente (C-VLAN) y las del servicio (S-VLAN) (...)



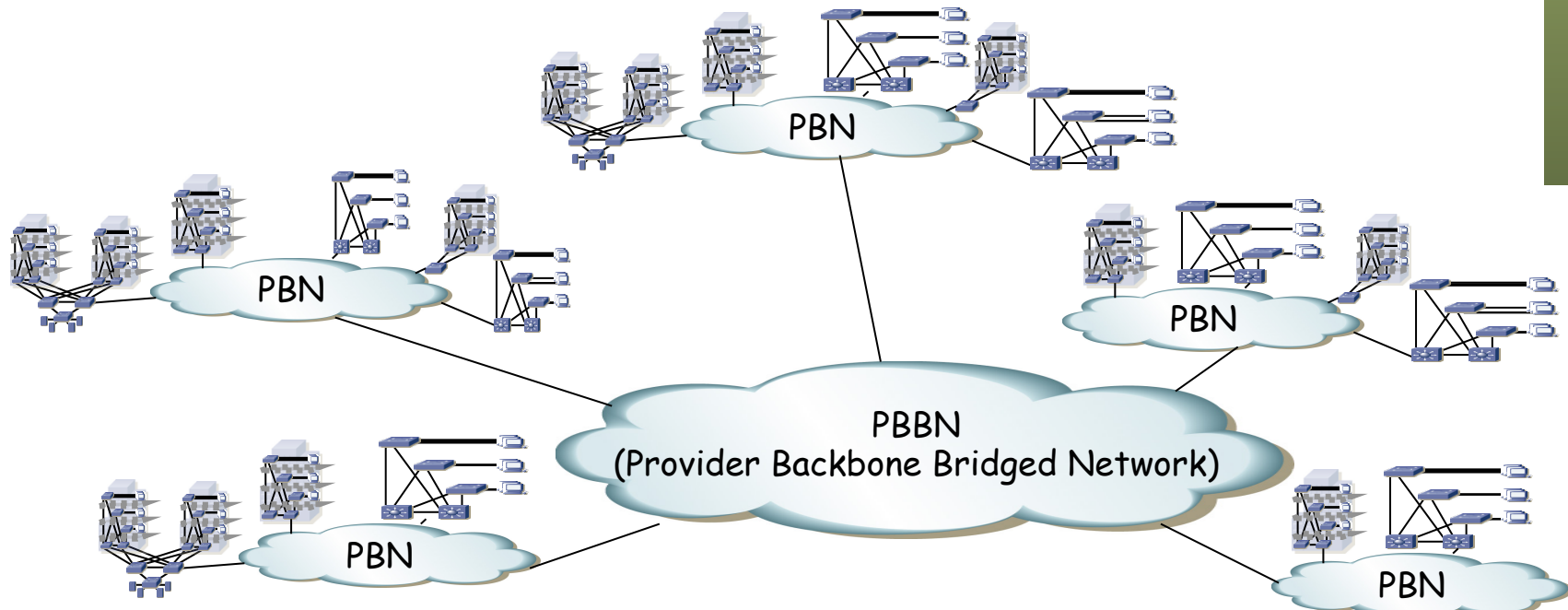
Problemas y evolución

- 802.1ad implica que los puentes del proveedor ven gran número de direcciones MAC
- Solo permite 4094 clientes simultáneos
- 802.1ah (modifica 802.1Q-2005): “*Provider Backbone Bridges*”
- Posibilita conectar PBNs (*Provider Bridged Networks*) a través de una PBBN (*Provider Backbone Bridged Network*) (...)



Problemas y evolución

- 802.1ad implica que los puentes del proveedor ven gran número de direcciones MAC
- Solo permite 4094 clientes simultáneos
- 802.1ah (modifica 802.1Q-2005): “*Provider Backbone Bridges*”
- Posibilita conectar PBNs (*Provider Bridged Networks*) a través de una PBBN (*Provider Backbone Bridged Network*) (...)



PBBNs

- Se define el Backbone Service Instance Tag (I-TAG) (... ..)

Ethernet



802.1Q



802.1ad



PBBNs

- Se define el Backbone Service Instance Tag (I-TAG) (... ..)

Ethernet

802.1Q

Client Dest Addr	Client Src Addr	Ethertype	Datos	CRC
------------------------	-----------------------	-----------	-------	-----

Client Dest Addr	Client Src Addr	C- TAG	Ethertype	Datos	CRC
------------------------	-----------------------	--------	-----------	-------	-----

Client Dest Addr	Client Src Addr	S- TAG	C- TAG	Ethertype	Datos	CRC
------------------------	-----------------------	--------	--------	-----------	-------	-----

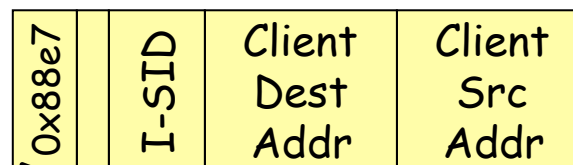
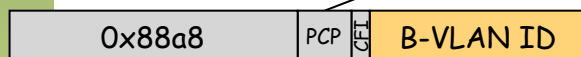
802.1ah 802.1ad

Dest Addr	Src Addr	B- TAG	I-TAG	S- TAG	C- TAG	Ethertype	Datos	CRC
--------------	-------------	--------	-------	--------	--------	-----------	-------	-----

PBBNs

- Se define el Backbone Service Instance Tag (I-TAG) (... ..)
- Se encapsula la trama Ethernet 802.1ad dentro de otra:
 - B-TAG: *Backbone VLAN tag* (idéntico a un S-TAG)
 - I-SID: *Backbone Service Identifier* (24 bits)
 - C-DA: *Encapsulated Customer Destination Address*. La dirección MAC destino de la trama encapsulada (Client Dest Addr en la figura)
 - C-SA: *Encapsulated Customer Source Address*. La dirección MAC origen de la trama encapsulada (Client Src Addr en la figura)
- Direcciones MAC origen y destino son de los equipos frontera de la PBBN
- Los conmutadores de la PBBN NO ven las direcciones MAC de los equipos de cliente (van encapsuladas)

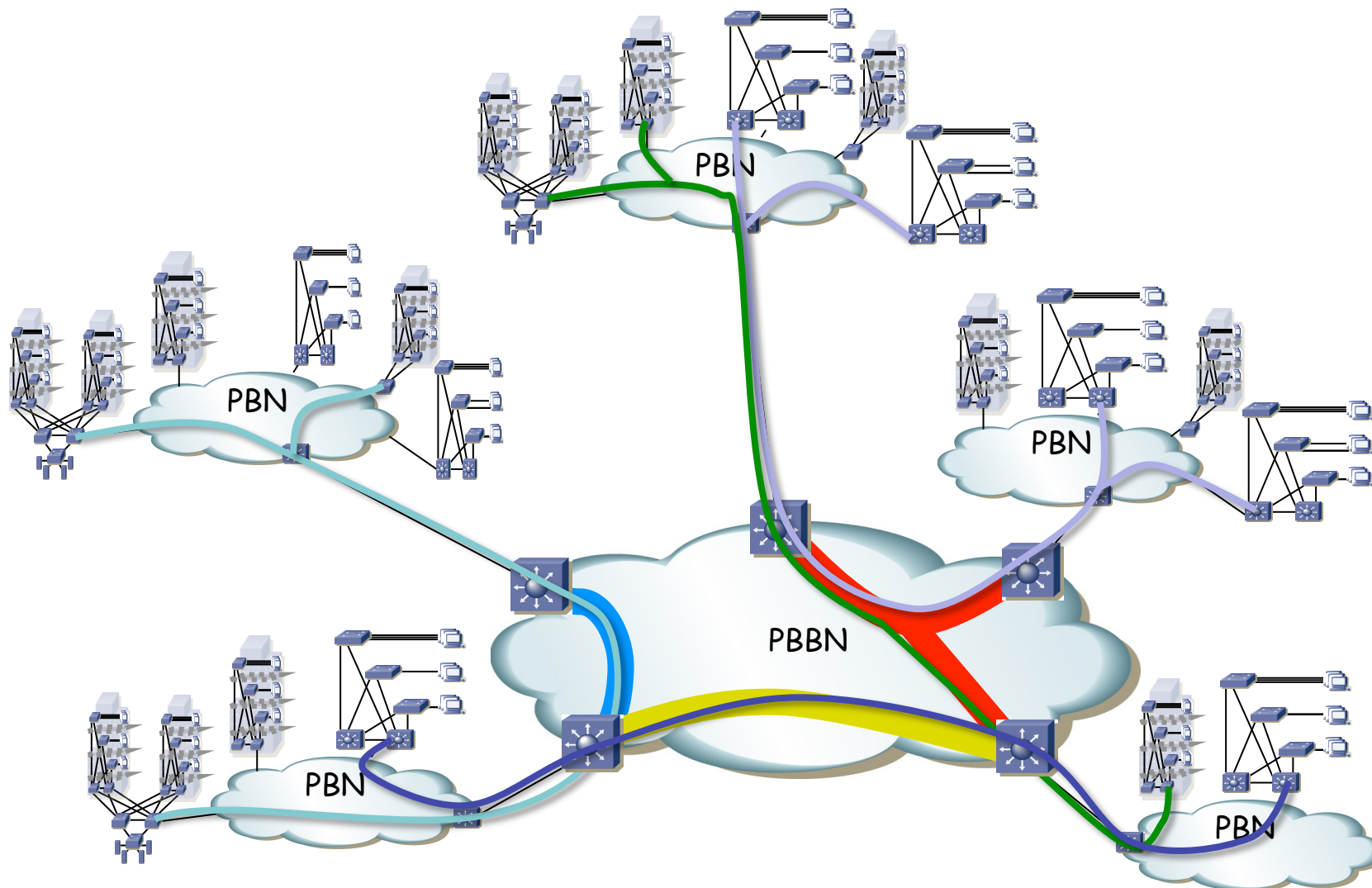
802.1ah



802.1QITagType

← 16 bytes →

PBBNs



Traffic Engineering

- 802.1Qay-2009 “Provider Backbone Bridge Traffic Engineering”, amendment a 802.1Q-2005
- Define funcionamiento de Ethernet *orientado a conexión* usando trama PBB (PBB-TE)
- Para ello se crearían *Ethernet Switched Paths* (ESPs) desde el plano de gestión (un agente externo configura switches)
- Un ESP es como un LSP, es también unidireccional
- El camino viene identificado por las direcciones origen y destino del ESP y el identificador de VLAN del B-TAG (serían la etiqueta)
- La etiqueta NO cambia en el camino, manteniendo el funcionamiento del plano de datos de Ethernet
- El ESP puede ser punto-a-punto o punto-a-multipunto (entonces la dirección destino es una dirección MAC de grupo)
- Desactiva el *learning* para esos VIDs y descarta tramas en esa VLAN con destino desconocido
- Posibilita servicios tipo *Carrier Ethernet* (<http://metroethernetforum.org>)
- 802.1Q-2014 tiene ya 1832 páginas

SPB

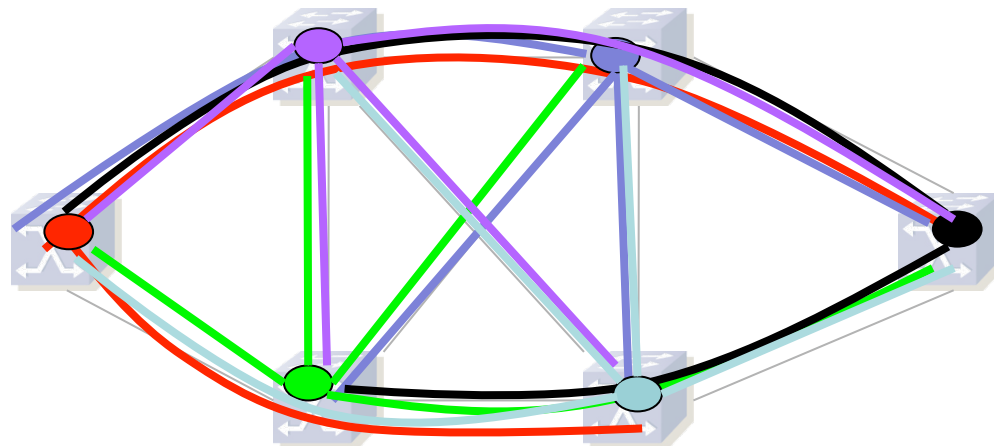
SPB: Objetivos

- *Shortest Path Bridging* (802.1aq, ya en 802.1Q-2014)
- Busca escalar la capa 2 al orden del millar de puentes
- Busca emplear múltiples caminos de igual coste
- Evitar aprendizaje de direcciones MAC donde sea posible
- Tiempos de recuperación en los centenares de milisegundos
- Reutiliza ASICs existentes
- Mantener orden y simetría en los caminos (para una VLAN)
- Compatibilidad con el resto de 802.1 (DCB, OA&M, etc)



SPB

- Sustituye {R|M}STP, pudiendo interoperar con ellos
- Para una VLAN ahora tenemos (al menos) 3 posibilidades:
 - Emplear el Internal Spanning Tree (IST)
 - Emplear una Multiple Spanning Tree Instance (MSTI)
 - Emplear un “SPT set” (set of Shortest Path Trees)
- Cada SPT del SPT set tiene como raíz del árbol a un puente diferente del dominio (...)

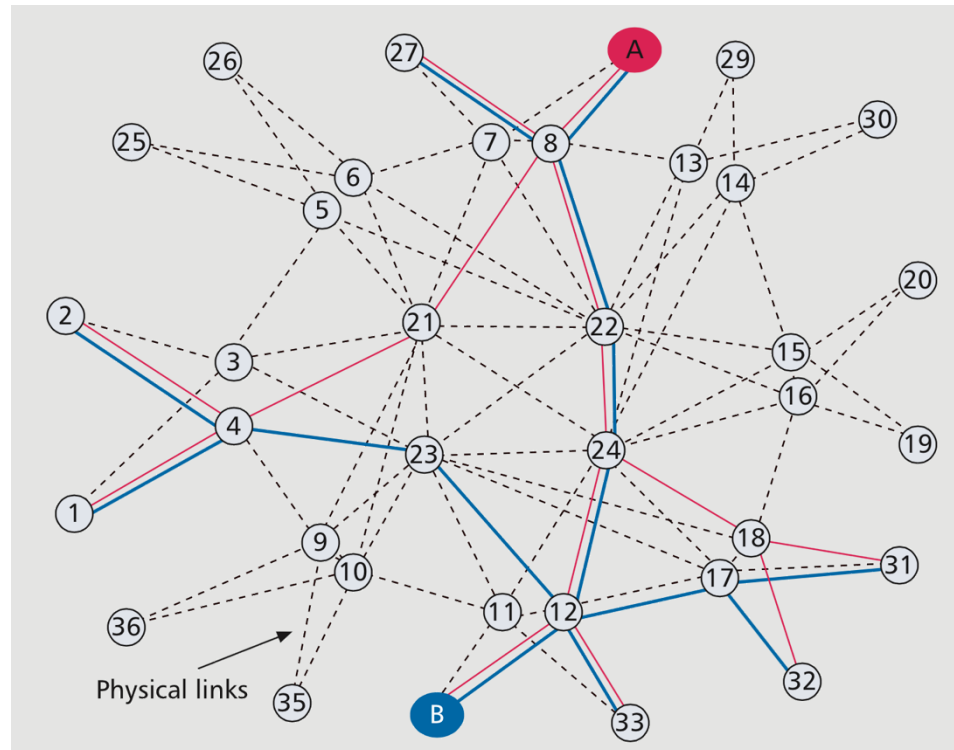


SPB: Control plane

- En el plano de control emplea IS-IS con algunas extensiones (ISIS-SPB)
- ISIS-SPB mantiene al menos un SPT (Shortest Path Tree) para cada puente con él como raíz
- Un puente envía tramas solo por uno de esos árboles

SPB: Caminos

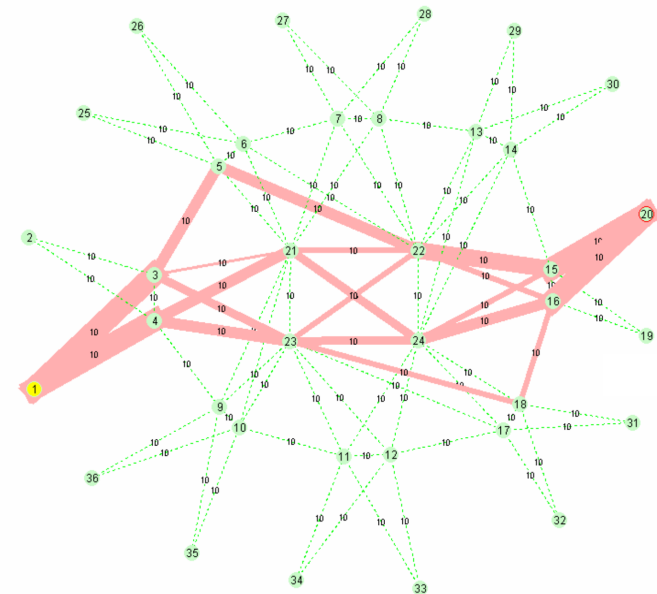
- Calcula caminos deterministas y simétricos
- Si hay varios caminos de coste mínimo entre 2 nodos solo se emplea uno de ellos en un SPT set
- Todos los nodos emplean la misma técnica de desempate para lograr caminos simétricos



D. Allan, et al. "Shortest Path Bridging:
Efficient Control of Larger Ethernet
Networks. IEEE Comm.Magazine

SPB: Load Balancing

- Al calcular SPTs se está distribuyendo la carga más que con un solo árbol de expansión
- Además puede calcular múltiples Equal Cost Trees (ECTs)
- Cada ECT resulta de un algoritmo de desempate
- Define 16 algoritmos de desempate, lo cual da hasta 16 ECTs
- SPTs resultado de un tipo de desempate forman un “SPT set”
- El balanceo de carga se hace distribuyendo VLANs por ellos
- No hace balanceo a nivel de paquete ni de flujo



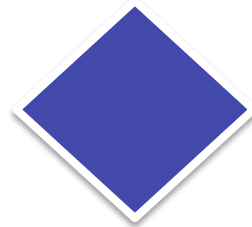
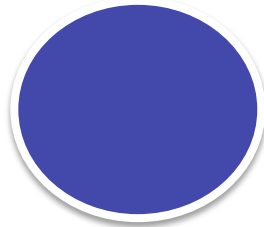
SPB: Loops

- Loops pueden producirse temporalmente
- Cuando las bases de datos de IS-IS no están sincronizadas
- Emplea RPF (Reverse Path Forwarding) para intentar evitarlos
- También emplea técnicas para evitar tener un loop mientras converge IS-IS
- Pero sigue sin haber TTL



SPB: Modos

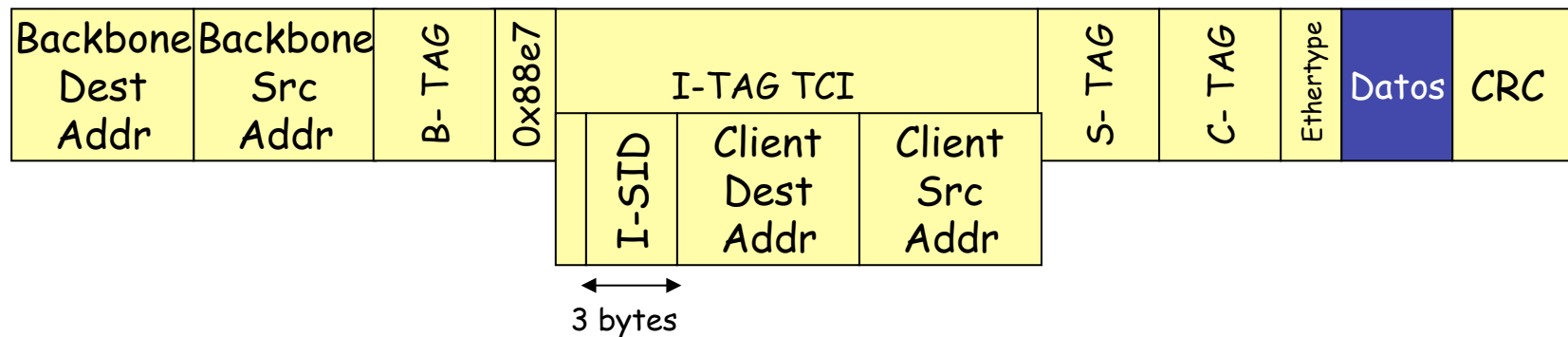
- En la asignación de tramas de una VLAN a un SPT se definen dos modos posibles:
 - SPBM: *Shortest Path Bridging MAC*
 - SPBV: *Shortest Path Bridging VID*
- Todos los puentes emplean el mismo modo para una VLAN concreta



SPBM

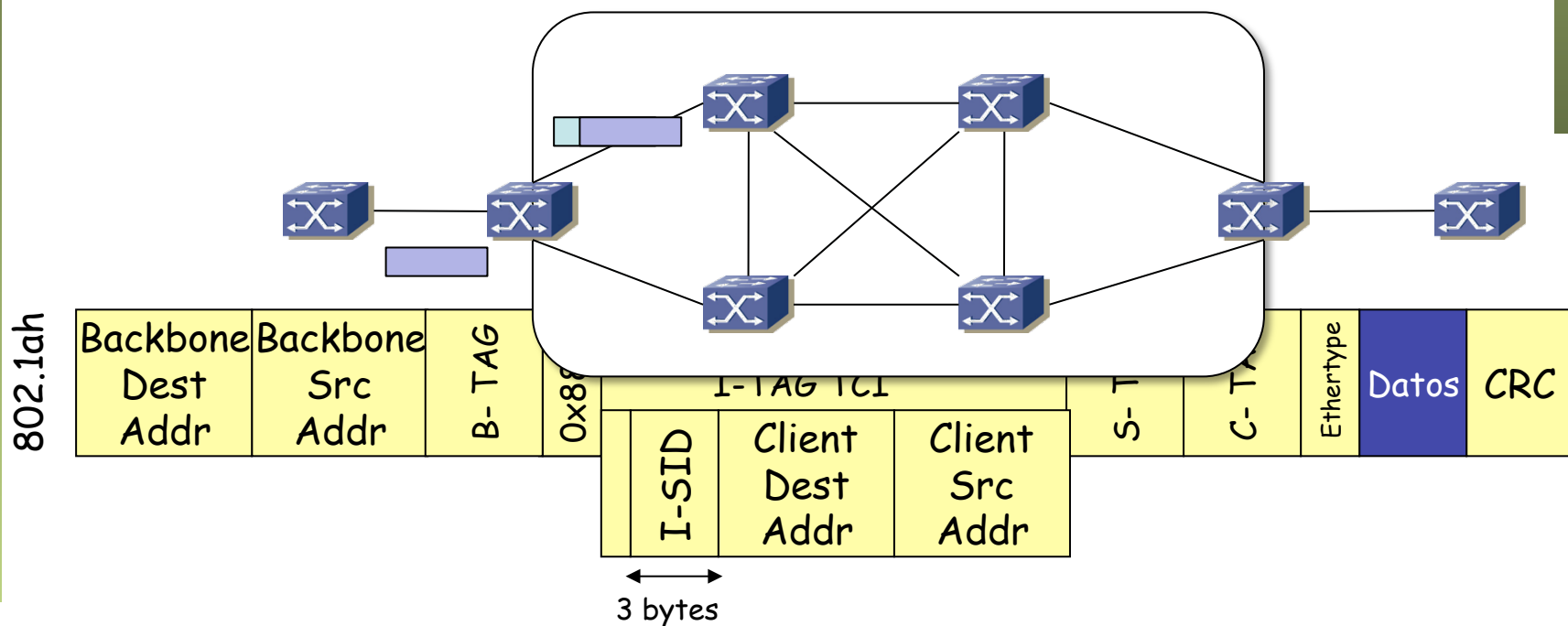
- *Shortest Path Bridging MAC Mode*
- Emplea encapsulado 802.1ah (MAC-in-MAC)
- Transporta transparentemente las VLANs de usuario
- El I-SID permite distinguir una gran cantidad de servicios
- Las tramas se transportan solo entre puertos que mapeen al mismo I-SID
- Cada I-SID se mapea a un B-VID (VLAN del backbone)
- La pareja de B-VID y dirección origen del backbone identifican al SPT

802.1ah



SPBM

- Solo los puentes frontera aprenden MACs de usuarios
- Si no tienen aprendida una dirección de usuario harán multicast
- Los puentes del backbone solo ven las MACs frontera
- Aprenden cómo llegar a esas direcciones mediante ISIS-SPB
- Es decir, el aprendizaje típico de puente se puede desactivar
- Hace una búsqueda de la dirección MAC destino para averiguar a qué puente frontera del Backbone mandar la trama



SPBM: Multicast

- Para las tramas BUM: tramas de usuario con destino desconocido en el nodo frontera o tramas con destino multicast/broadcast
- Dos alternativas seleccionables por usuario
 - Replicación en el nodo frontera y unicast
 - Construcción de árboles multicast con raíz en cada nodo
- Unicast y multicast sigue el mismo camino
- En caso de árboles para el multicast emplea una dirección MAC origen específica para el grupo
 - SPSourceID identifica al puente
 - I-SID

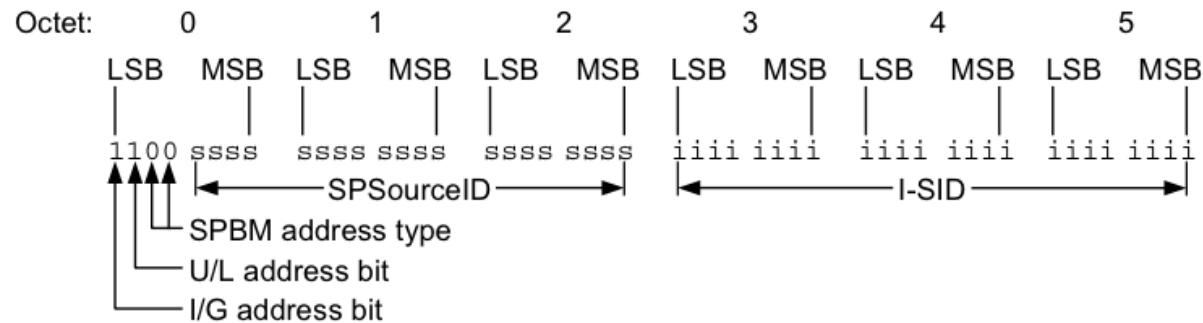
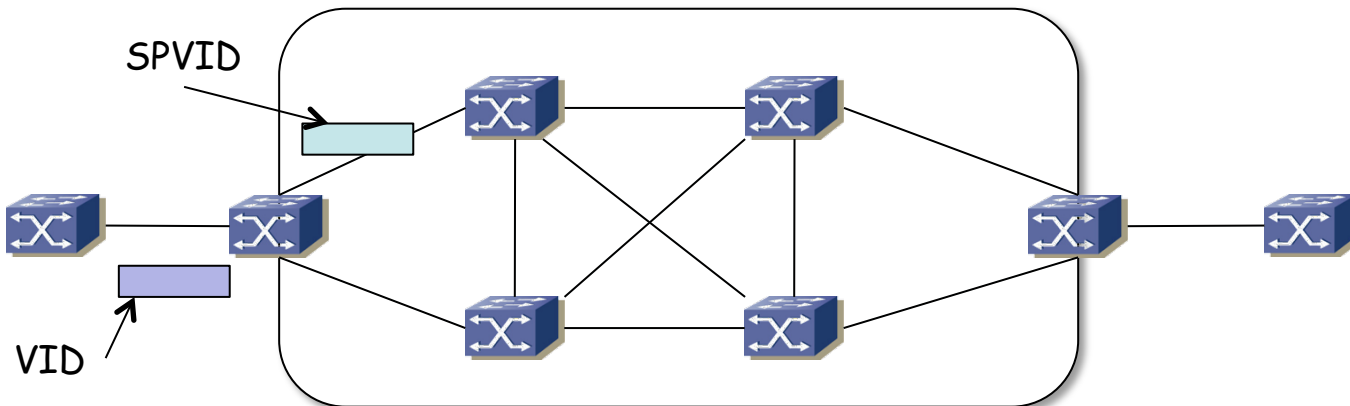


Figure 27-3—SPBM Group MAC Addresses used by this standard

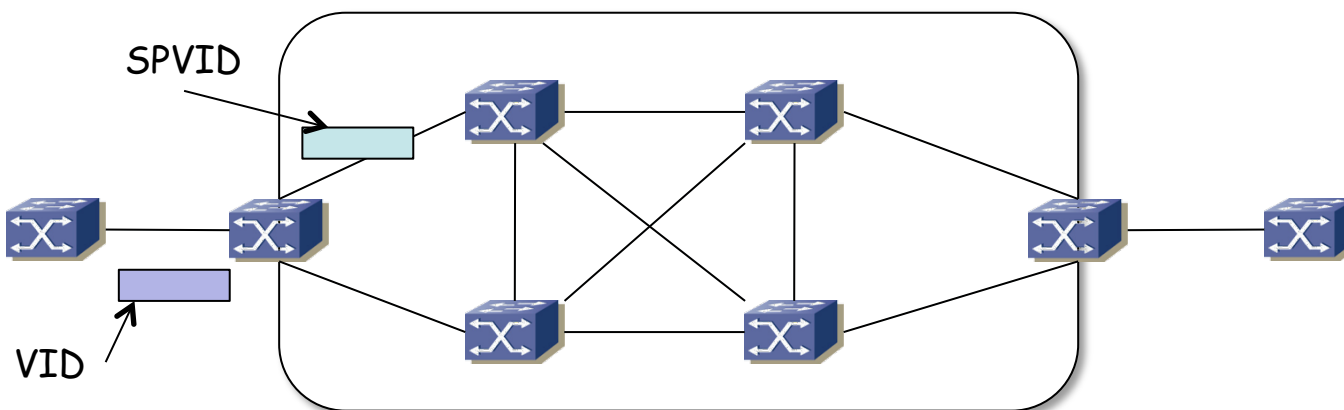
SPBV

- *Shortest Path Bridging VID*
- Requiere menos configuración que SPBM
- Un nodo frontera recibe tramas de una VLAN
- Calcula un SPVID (*Shortest Path VLAN Identifier*) en base al VID y el nodo de entrada (un SPVID para cada pareja)
- Se envía la trama con el SPVID como VID
- El cambio se deshace en el nodo de salida
- El mapeo se sincroniza entre los nodos empleando ISIS-SPB



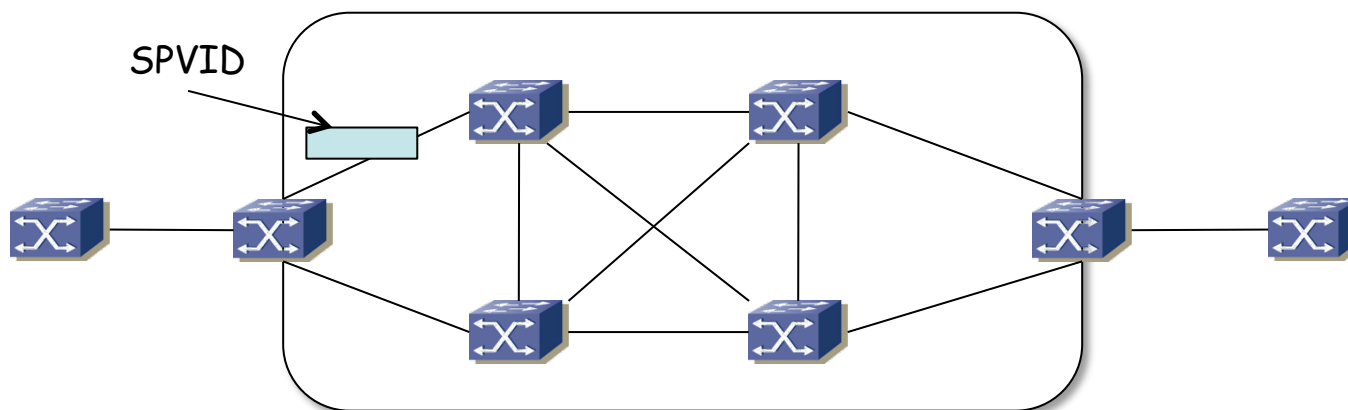
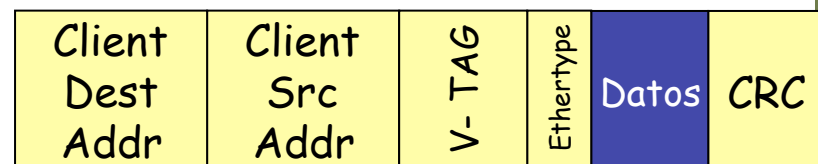
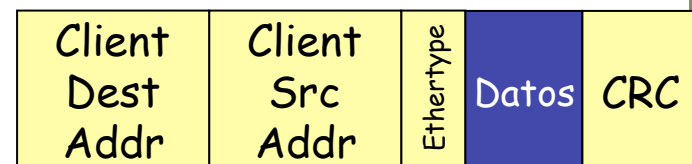
SPBV

- Es decir, en el backbone se emplea una VLAN para cada SPT
- Eso limita a $(n^{\circ}\text{VLANs_en_backbone} \times n^{\circ}\text{Nodos}) < 4095$
- El aprendizaje de bases de datos de filtrado en el backbone es compartido entre todos los SPVID de la misma VLAN de usuario
- Si no puede asignar un SPVID no cambia el VID y emplea el IST



SPBV

- Si la trama original no tiene encapsulado de VLAN se añade el SPVID con 802.1Q
- Se retira a la salida
- Si tiene encapsulado 802.1Q se modifica el V-TAG
- No separa el direccionamiento de espacio del usuario del empleado en el backbone

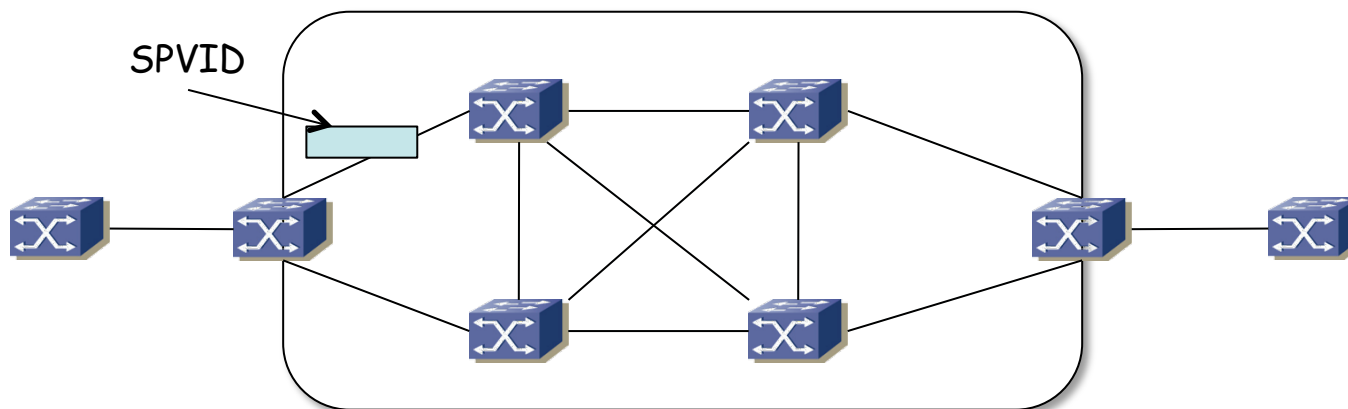


SPBV

- Si la trama original tiene encapsulado 802.1ad (Q-in-Q) se modifica el S-TAG con el SPVID

802.1ad

Client Dest Addr	Client Src Addr	S- TAG	C- TAG	Ethertype	Datos	CRC
------------------------	-----------------------	--------	--------	-----------	-------	-----



ECMP 802.1Qbp

- 802.1Q-2014 tiene ya 1.832 páginas
- 802.1aq daba tres posibilidades:
 - Emplear el Internal Spanning Tree (IST)
 - Emplear una Multiple Spanning Tree Instance (MSTI)
 - Emplear un “SPT set” (set of Shortest Path Trees)
- 802.1Qbp añade dos posibilidades más
 - Emplear un conjunto de *Traffic Engineered Service Instances* (TESIs)
 - Emplear múltiples caminos de igual coste (hasta 16) que reparten el tráfico desde cada puente (ECMP)
- La solución ECMP solo se puede emplear en modo SPBM
- Permite repartir el tráfico de una misma VLAN
- Deja de garantizar que unicast y multicast sigan el mismo camino

ECMP 802.1Qbp

- Añade la posibilidad de un Flow-Tag (F-TAG)
- Ese campo contiene un identificador de flujo
- Puede hacer el balanceo en base a ese identificador (las tramas con el mismo valor siguen siempre el mismo árbol)
- De esa forma mantiene el orden en las tramas de un flujo
- El F-TAG contiene también un campo de TTL

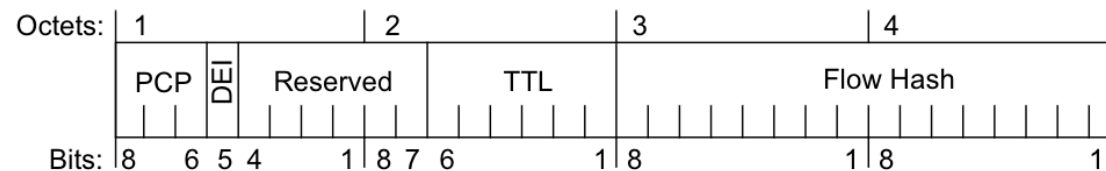
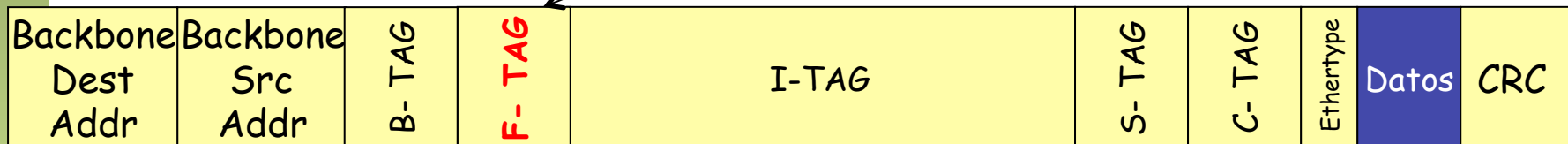


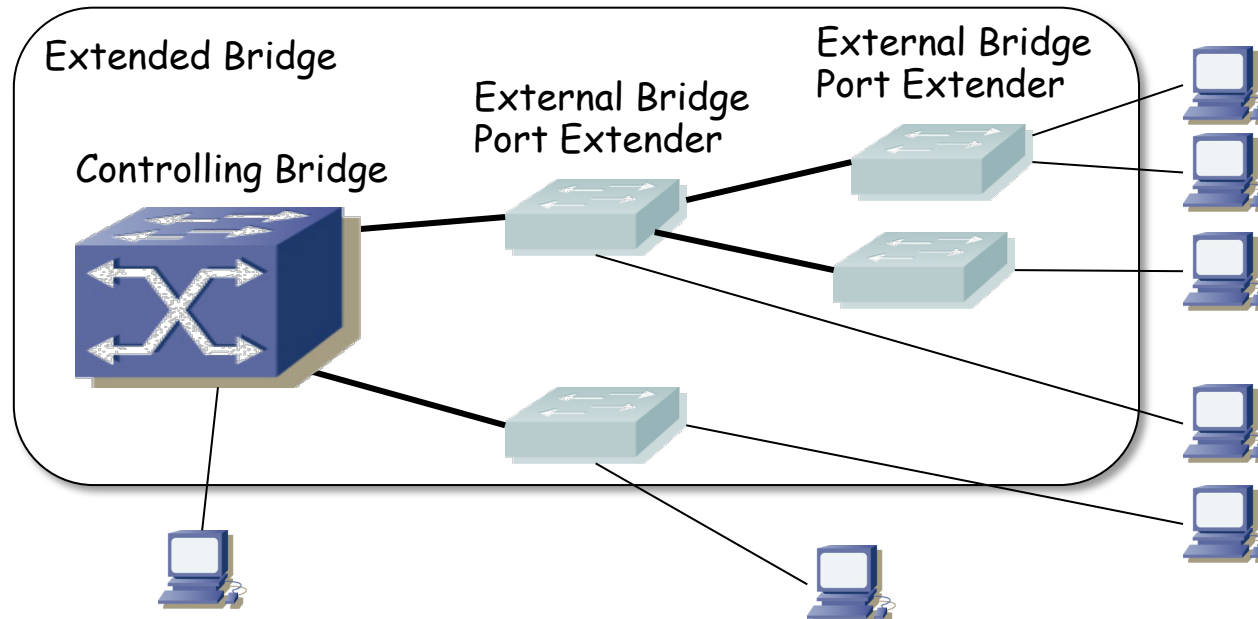
Figure 44-8—Flow Filtering TCI format



Extended Bridges y Virtual Bridging

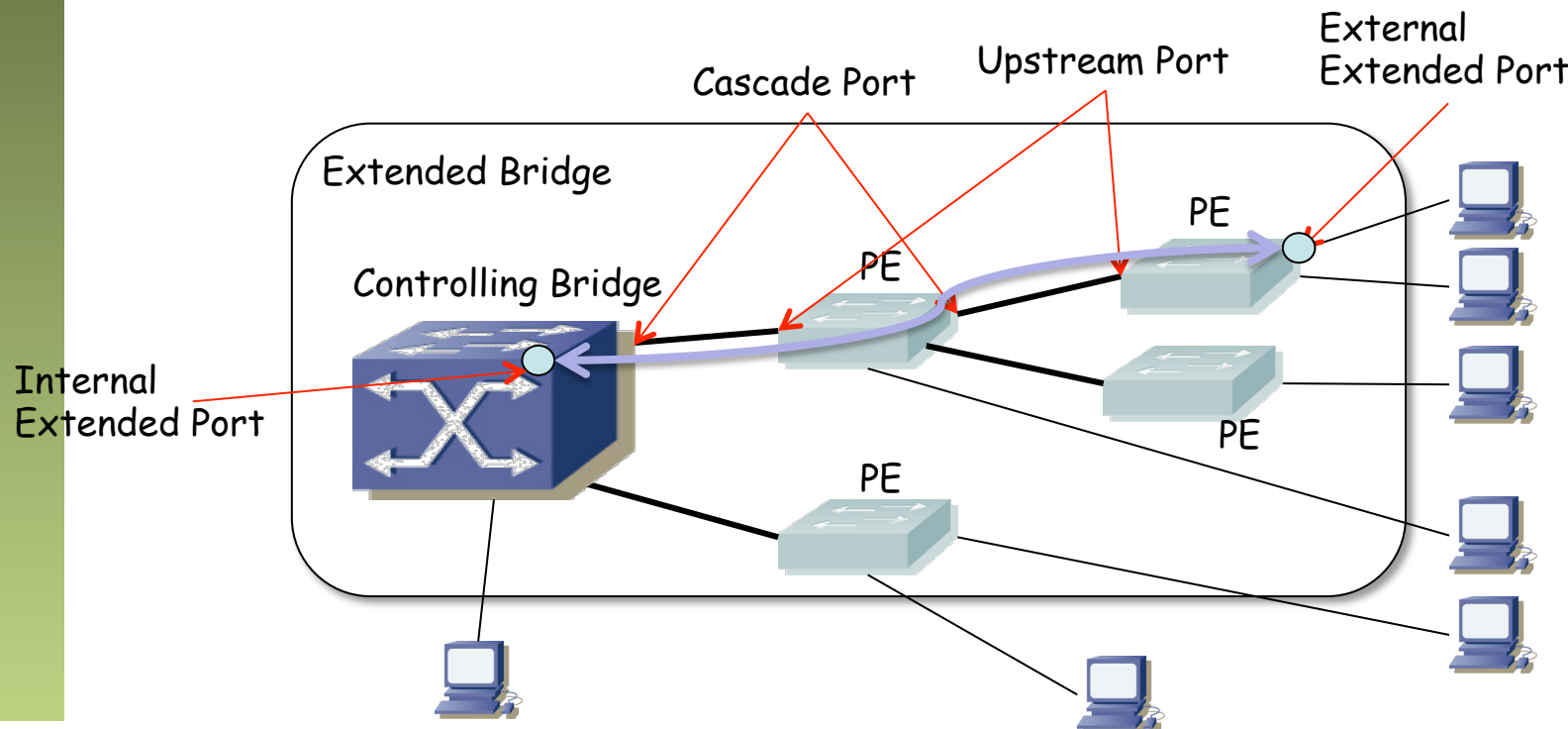
802.1BR Extended Bridge

- 802.1BR-2012 “Bridge Port Extension”
- Inicialmente 802.1Qbh (modificación a 802.1Q) pero lo cancelaron y lo movieron a un estándar independiente
- Introduce el *Extended Bridge*, que es simplemente un puente 802.1Q
- Está formado por *Controlling Bridge* y uno o más *Bridge Port Extenders*
- Forman un árbol con el Controlling Bridge como raíz (no STP interno)
- Todo ello forma un puente 802.1Q (el Extended Bridge) (...)
- Todo el extended bridge se gestiona como una unidad



Extended Bridge

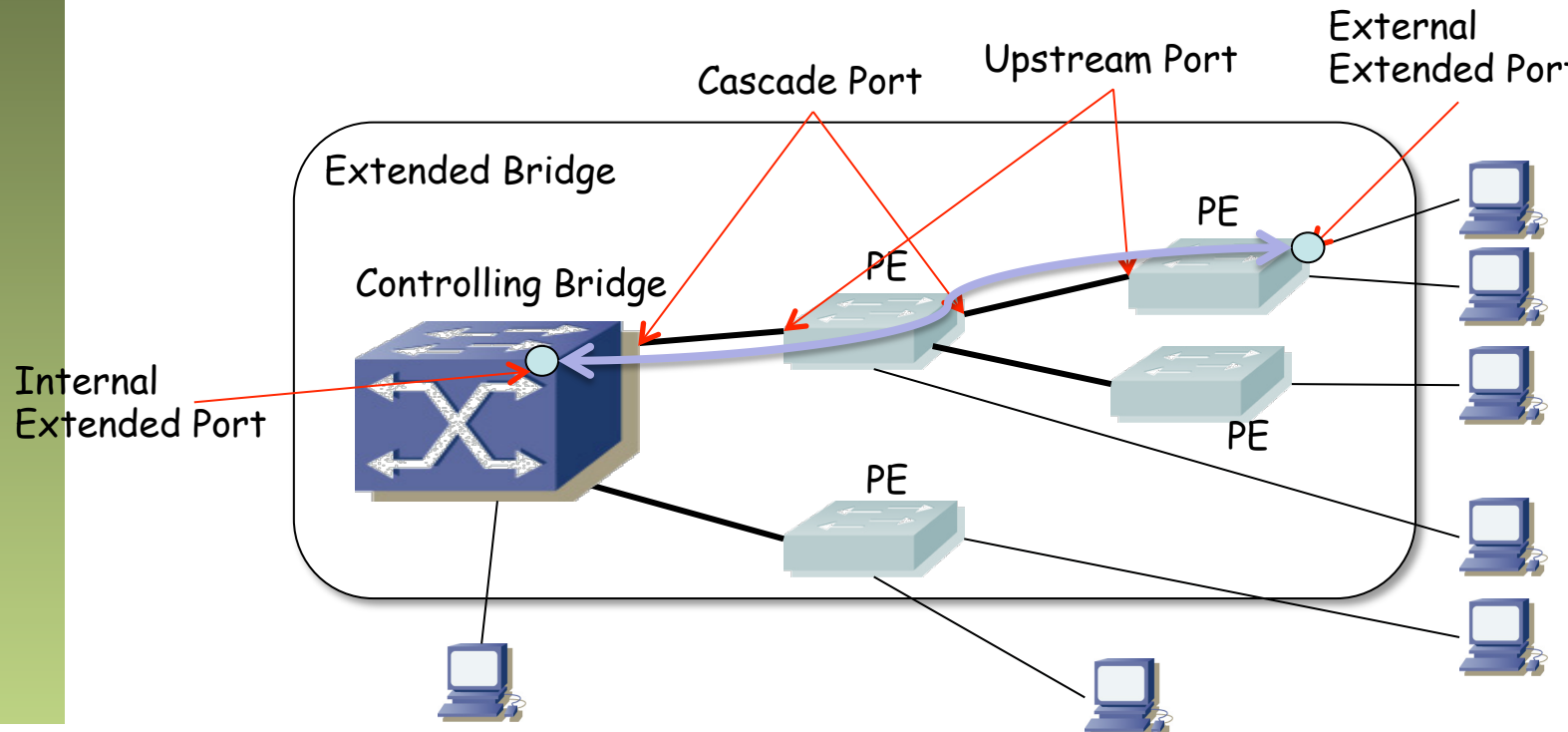
- Todos los puertos de los PE tienen su imagen en puertos internos del Controlling Bridge
- El camino bidireccional desde el puerto externo al interno en el Controlling Bridge se llama un *E-channel* (...)
- El tráfico se conmuta en el Controlling Bridge



Extended Bridge

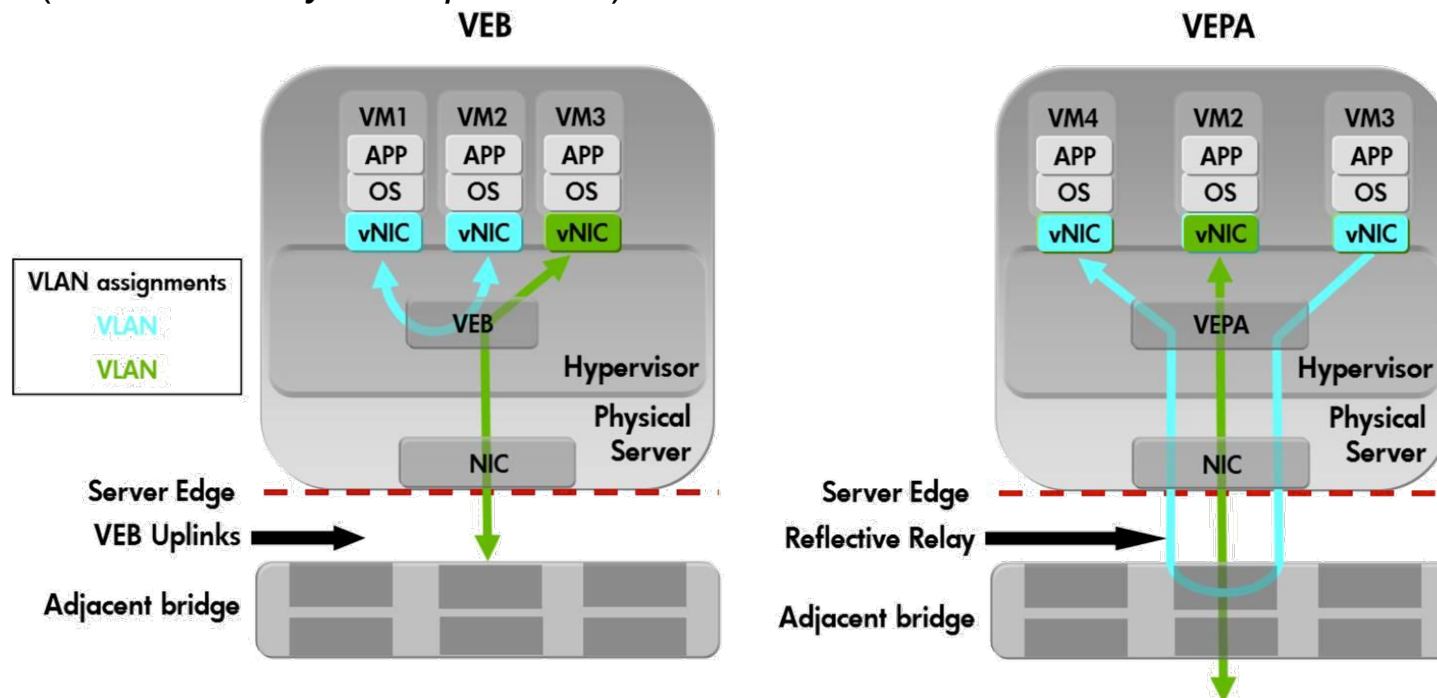
- Un E-channel se identifica por un E-channel Identifier (E-CID)
- A una trama que entra por un Extended Port se le añade un E-Tag
- Entre otros aspectos permite identificar al puerto por el que entró la trama (de hecho al E-channel)

Client Dest Addr	Client Src Addr	E- TAG	C- TAG	Ethertype	Datos	CRC
------------------------	-----------------------	--------	--------	-----------	-------	-----



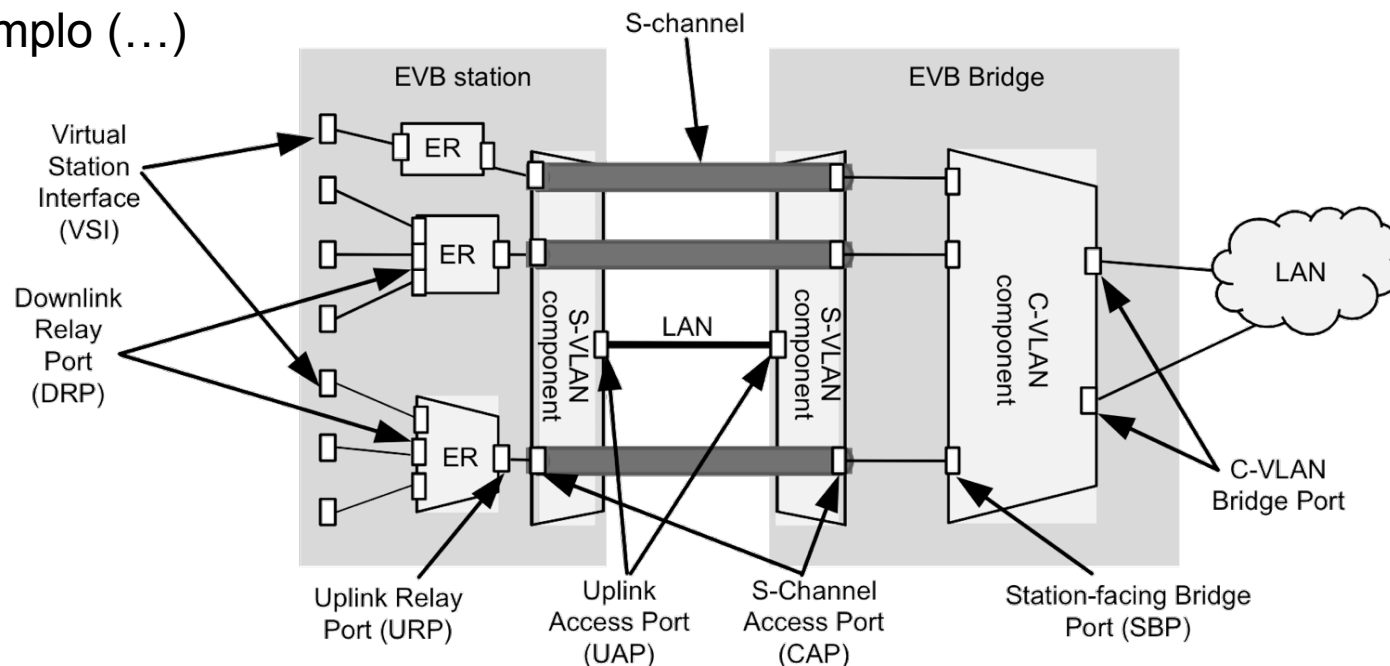
802.1Qbg EVB

- 802.1Qbg-2012 “Edge Virtual Bridging”, ya en 802.1Q-2014
- El hypervisor además de hacer de VEB (*Virtual Edge Bridge*) puede hacer de VEPA (*Virtual Edge Port Aggregator*)
- Se busca que el switch físico se encargue de conmutar el tráfico incluso entre las VMs del mismo host (resulta en mayor utilización de los enlaces)
- Es decir, el VEPA no conmuta internamente sino que manda al exterior
- El switch físico suele poder implementar más políticas que los switches software
- ¡ El switch debe reenviar tráfico por el mismo puerto por el cual le llegó ! (*reflective relay* o *hairpin turns*)



802.1Qbg *S-channels*

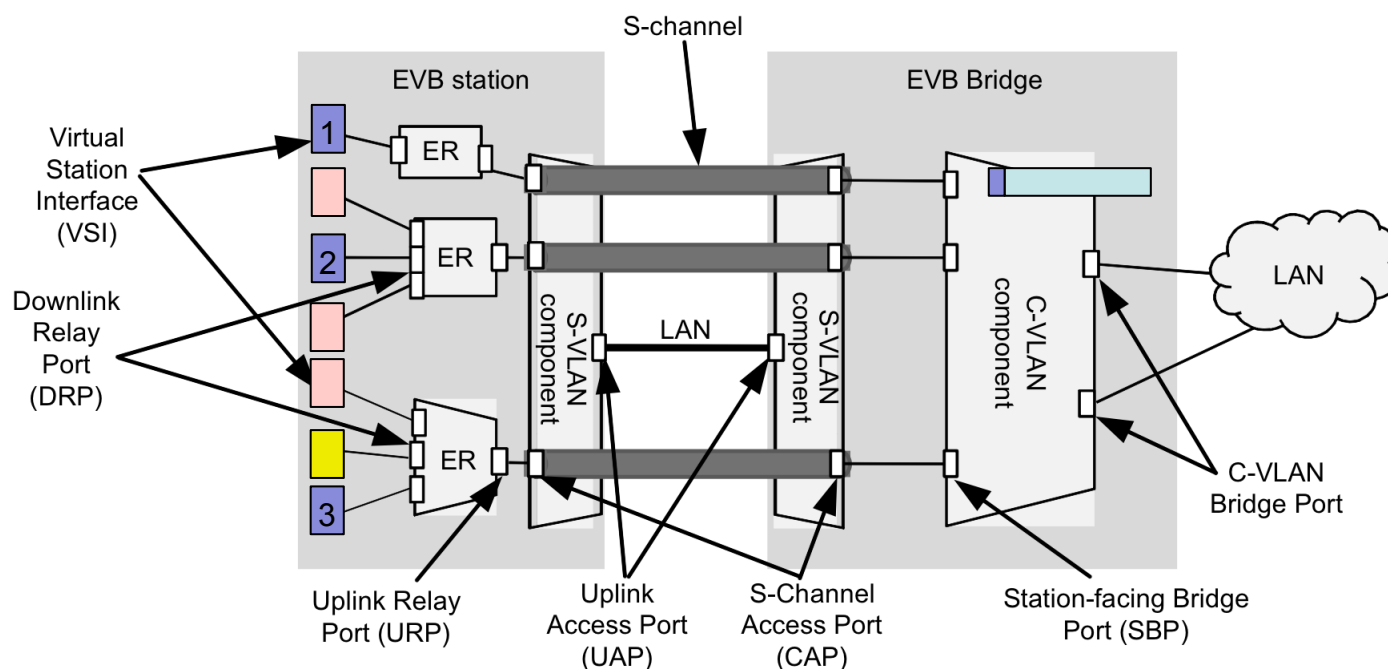
- Comunicación entre dos VMs en diferente vSwitch
- Multiplexar enlaces trunk lógicos sobre uno físico
- Se etiqueta la trama con un S-Tag (QinQ) para transportarla al switch físico e identificar al S-channel
- El S-channel identifica al vSwitch
- CDCP = “S-Channel Discovery and Configuration Protocol”
- CDCP emplea LLDP con nuevos TLVs
- Ejemplo (...)



802.1Qbg-2012 Figure 40-1—EVB architecture overview

802.1Qbg *S-channels*

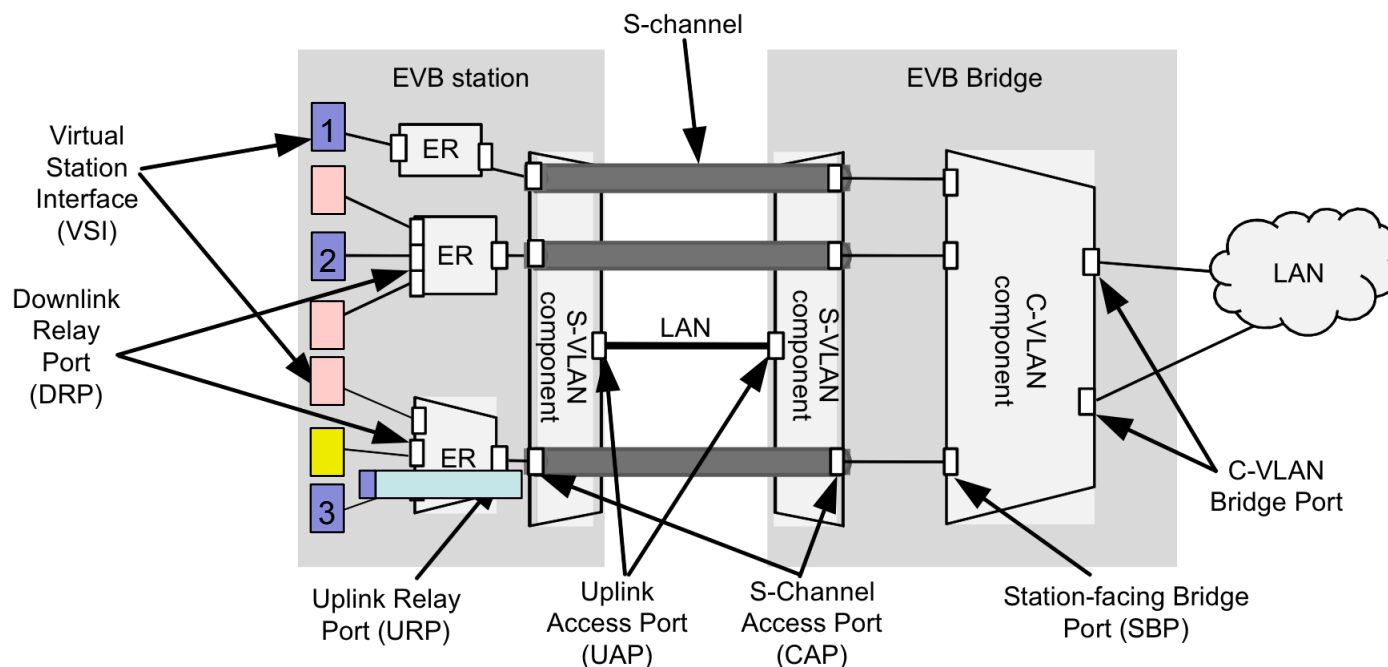
- Ejemplo
 - Trama de VSI 1 a VSI 3, en la VLAN azul
 - Se etiqueta con S-Tag para el S-channel
 - Se retira el S-Tag y se decide dónde está el destino
 - (...)



802.1Qbg-2012 Figure 40-1—EVB architecture overview

802.1Qbg *S-channels*

- Ejemplo
 - Trama de VSI 1 a VSI 3, en la VLAN azul
 - Se etiqueta con S-Tag para el S-channel
 - Se retira el S-Tag y se decide dónde está el destino
 - Se etiqueta con un nuevo S-Tag para el S-channel que lleva al destino
 - Se retira el S-Tag y el vSwitch entrega la trama al VSI
 - La trama no ha sido vista por el vSwitch donde está el VSI 2



802.1Qbg-2012 Figure 40-1—EVB architecture overview

802.1Qbg VDP

- VDP = “VSI discovery and configuration protocol”
- VSI = “Virtual Station Interface”
- Permite notificar al switch físico de la presencia de una VM
- Eso permite la configuración del puerto del switch con el perfil necesario para esa nueva VM
- Transportado como TLVs sobre ECP (*Edge Control Protocol*)
- ECP a su vez son tramas Ethernet con Ethertype 0x8940, incluye confirmaciones

Extended Bridges y Virtual Bridging

802.1Qca-2015

- *“Path Control and Reservation”, Amendment 24 a 802.1Q-2014*
- *“These extensions involve explicit path control, bandwidth reservation, and redundancy (protection, restoration) for data flows. Thus, this standard specifies bridging on explicit paths for unicast and multicast frames, specifying protocols to determine multiple active topologies.”*
- Extensiones a IS-IS: *ISIS-PCR (Path Control and Reservation)*
- Emplea el “Path Control Element” (PCE) definido por el IETF PCE WG
- El PCE es un puente o un host