VPNs de capa 2 basadas en MPLS

Ingeniería de tráfico - Curso 2014-15

Área de Ingeniería Telemática - Departamento de Automática

Universidad de Alcalá

José Manuel Giménez

VPNs de capa 2 basadas en MPLS

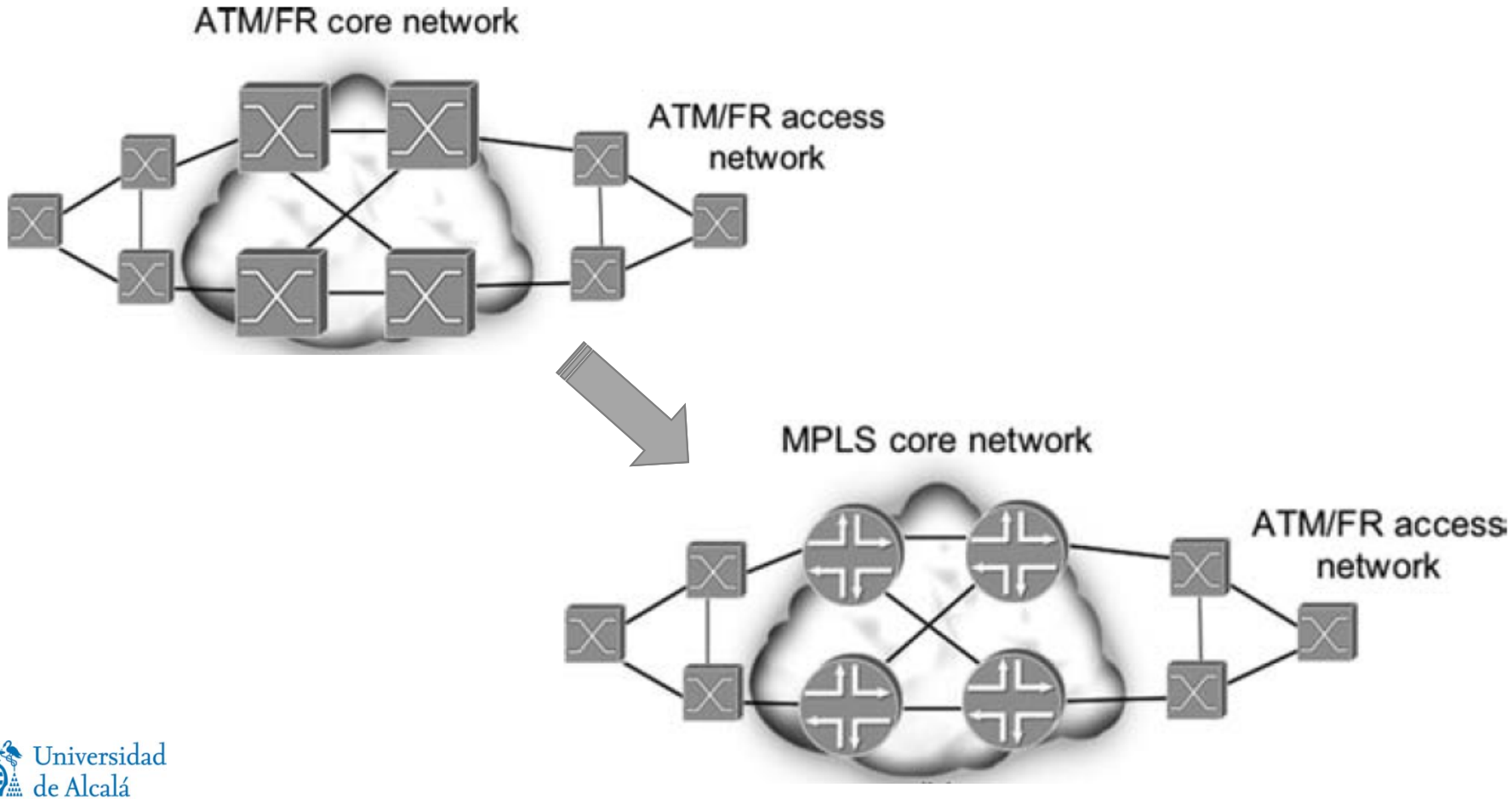
- Introducción a las L2VPNs.
- L2VPNs vs. L3VPNs.
- Transporte de capa 2 sobre MPLS.
- Plano de reenvío.
- Plano de control.
- VPNs de capa 2.5.

Introducción a las L2VPNs

- Una fuente importante de ingresos de los proveedores procede de la interconexión de redes LAN corporativas.
  - Se interconectan en capa 2: el proveedor transporta celdas ATM o tramas FR con una determinada tasa para cada circuito virtual.
    - Ese tráfico puede ser vídeo o incluso tráfico de voz entre centralitas → Requisitos de QoS.
  - Típicamente siguiendo una arquitectura "*hub and spoke*".
- Las L2VPN permiten ofrecer a los clientes que tienen contratado un servicio ATM o FR un servicio similar, migrando la red a MPLS.
  - El proveedor puede mantener una única red (que es MPLS) en lugar de múltiples: ATM, FR, Ethernet…
  - Ahorro en CAPEX y OPEX (CAPital EXpenditures /OPerating EXpenses) para el proveedor puesto que puede emplear una única red para dar conectividad de capas 2 y 3.

Introducción a las L2VPNs

- Si las redes de acceso se basan en FR o ATM se puede migrar únicamente el núcleo de la red:



Introducción a las L2VPNs

- Las L2VPNs basadas en MPLS se les llama en ocasiones AToM: *Any Transport over MPLS*.
- Posibilidad de ofrecer servicios Ethernet sobre la red MPLS:
  - El proveedor transporta tramas Ethernet entre sedes del cliente sobre la red MPLS del proveedor.
  - Muchas redes corporativas actuales se basan o están migrando a Ethernet.
    - Emplear Ethernet para el transporte entre sedes es una extensión natural.
    - Equipos Ethernet son menos costosos que los ATM o FR.
  - Además de ofrecer Ethernet punto a punto se pueden ofrecer servicios Ethernet multipunto: VPLS (*Virtual Private LAN Service*).

L2VPNs vs. L3VPNs

	L2VPNs	L3VPNs
Modelo de conectividad	Overlay	Peer
Interacción para routing entre cliente y proveedor	NO	Si CE y PE pueden intercambiar información de routing
Protocolo de capa 3 entre las sedes del cliente	Cualquier protocolo de capa 3. La red del proveedor solo se ocupa de transportar PDUs de capa 2, no es consciente del protocolo de capa 3	Se requiere IP. Excluye L3VPNs para transportar protocolos de capa 3 que no sea IP (p.e. IPX, Apple Talk)
Nº de conexiones entre CE y PE	Múltiples conexiones (físicas o lógicas) entre cada CE y su correspondiente PE local, una por cada CE remoto con el que se pretenda conexión	Solo una conexión entre cada CE y su correspondiente PE local. Cada PE es responsable del routing del tráfico hacia el CE de egreso apropiado

L2VPNs vs. L3VPNs

- ¿Qué opción escoger?
  - Depende de los protocolos que deseen transportarse.
  - Depende del grado de implicación que desee el cliente en cuanto al encaminamiento.
- El proveedor puede ofrecer L2VPNs y L3VPNs sobre una misma infraestructura MPLS.

Transporte de capa 2 sobre MPLS

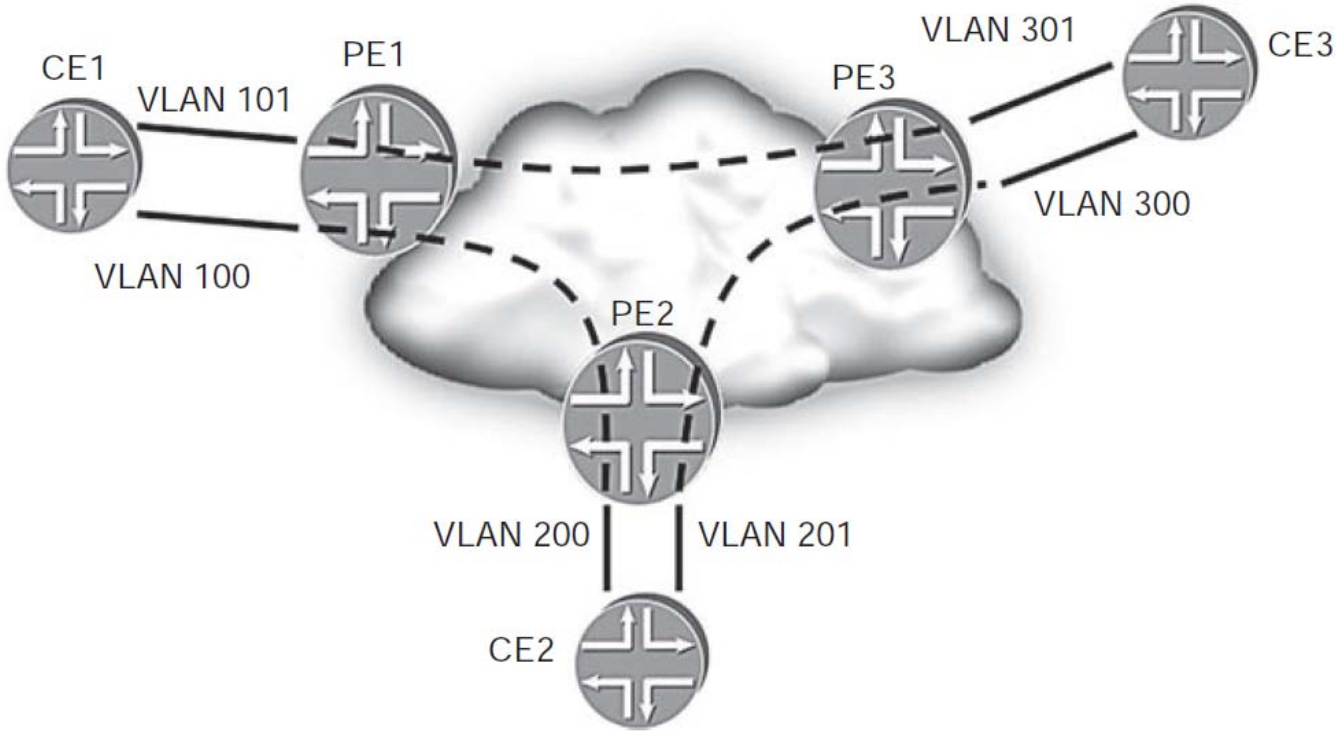
- Existen dos alternativas en función del plano de control:
  - Señalización basada en LDP.
  - Señalización basada en BGP.
- En el plano de reenvío ambas propuestas son similares:
  - Encapsulan las tramas de capa 2 del mismo modo para ser transportadas por la red MPLS.
- Una conexión de capa 2 punto a punto sobre una red MPLS se suele denominar pseudocable (*pseudowire*).
  - Se desea que la red MPLS sea invisible al cliente, de manera que este puede pensar que los dos CEs interconectados por el pseudocable están directamente conectados.
- Una L2VPN basada en MPLS se compone de un conjunto de pseudocables que conectan CEs siguiendo una topología escogida por el usuario.

Transporte de capa 2 sobre MPLS

- Ejemplos de protocolos de capa 2 que pueden transportarse sobre una red MPLS:
  - ATM. Existen dos modos, principalmente:
    1. Se transportan PDUs AAL5 sobre el pseudocable.
    2. Se transportan celdas ATM sobre el pseudocable.
  - Ethernet. El mapeo de tráfico en pseudocables puede hacerse:
    1. Por VLAN.
    2. Por puerto.
  - Frame Relay. El mapeo de tráfico en pseudocables puede hacerse:
    1. Por DLCI.
    2. Por puerto.

Transporte de capa 2 sobre MPLS

- Ejemplo:



- AC: *Attachment Circuit*, nombre que se le da a cada circuito de acceso.

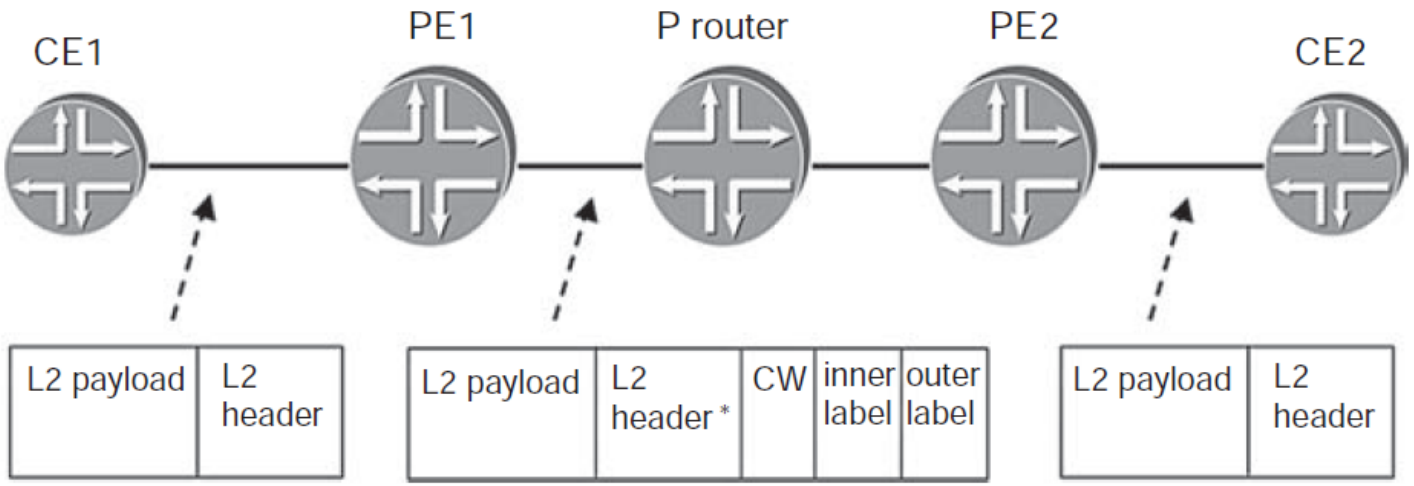
Plano de reenvío

- RFC4905 (junio 2007):
  - "*Encapsulation Methods for Transport of Layer 2 Frames over MPLS Networks*".
  - "Categoría: Histórica".
  - Describe como se encapsulan las tramas de capa 2 para su transporte sobre una red MPLS.
  - Reemplazada por diversas RFCs, una para cada protocolo de transporte, del grupo de trabajo de IETF denominado PWE3 (*PseudoWire Emulation Edge-to-Edge*).
    - Sin embargo, se parecen bastante a RFC4905.

Plano de reenvío

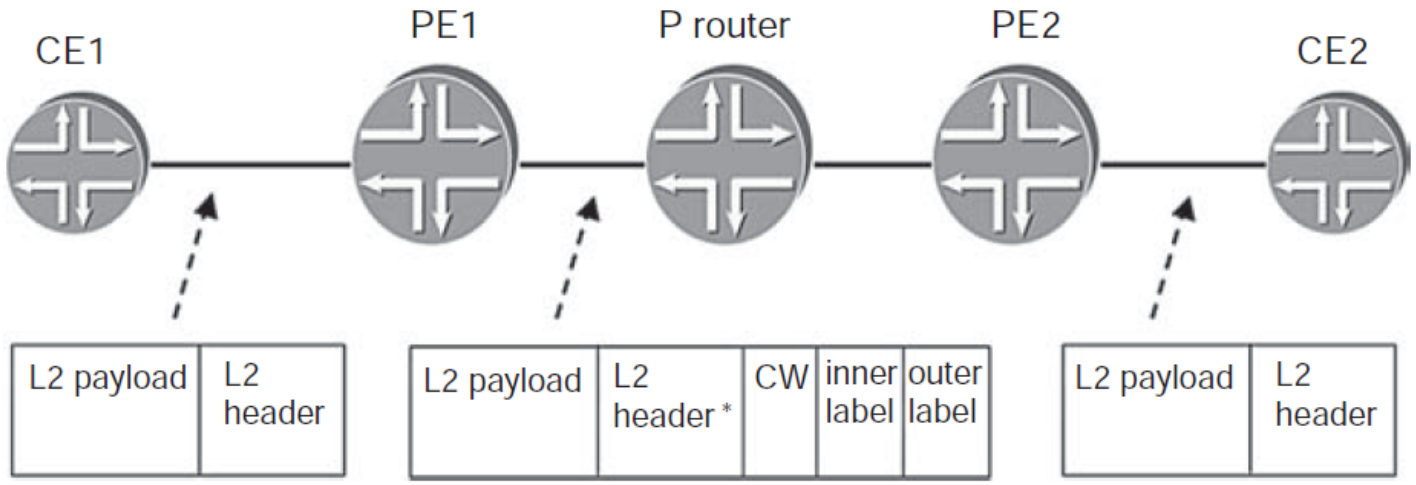
- Ejemplo: cuando llega un trama de capa 2 a PE1:

1. Se eliminan las partes de la trama de capa 2 que no se van a transportar.
   - Por ejemplo, en Ethernet se elimina el FCS.
2. En ocasiones se le añade una palabra de control (CW – Control Word) de 4 bytes.
   - CW puede incluir un número de secuencia para que el PE de salida detecte tramas desordenadas en recepción.
   - Puede contener *flags* correspondientes a campos de la cabecera de capa 2 para evitar el envío de la cabecera de capa 2 completa.



Plano de reenvío

3. PE1 busca el valor de la etiqueta VPN que espera PE2 para la trama y añade dicha etiqueta MPLS a la trama.
4. PE1 determina la manera de alcanzar PE2.
   - El proveedor tiene distintas maneras para hacer el túnel por el núcleo de su red:
     - De manera similar a lo que ocurre con las L3VPN.
     - Puede usar, por ejemplo, LDP, RSVP-TE, con lo que se tendría un túnel MPLS.
     - O usar, túneles GRE o IPSec, por lo que no se trataría de un túnel MPLS.
     - Se puede compartir el mismo túnel entre L3VPN, tráfico de capa 2 e incluso VPLS.
   - Si se emplea LDP o RSVP-TE:
     - PE1 determina el valor de la etiqueta MPLS necesaria para alcanzar PE2 y apila dicha nueva etiqueta.
   - Si no se emplea MPLS (GRE o IPSec), PE1 determina el túnel adecuado para alcanzar PE2.



n MPLS

Plano de reenvío

5. PE2, al recibir el paquete, examina el valor de la etiqueta VPN antes de eliminarla.

Si el paquete llega por un túnel MPLS y no se emplea PHP, PE2 debe primero eliminar la etiqueta de transporte para mostrar la etiqueta VPN.

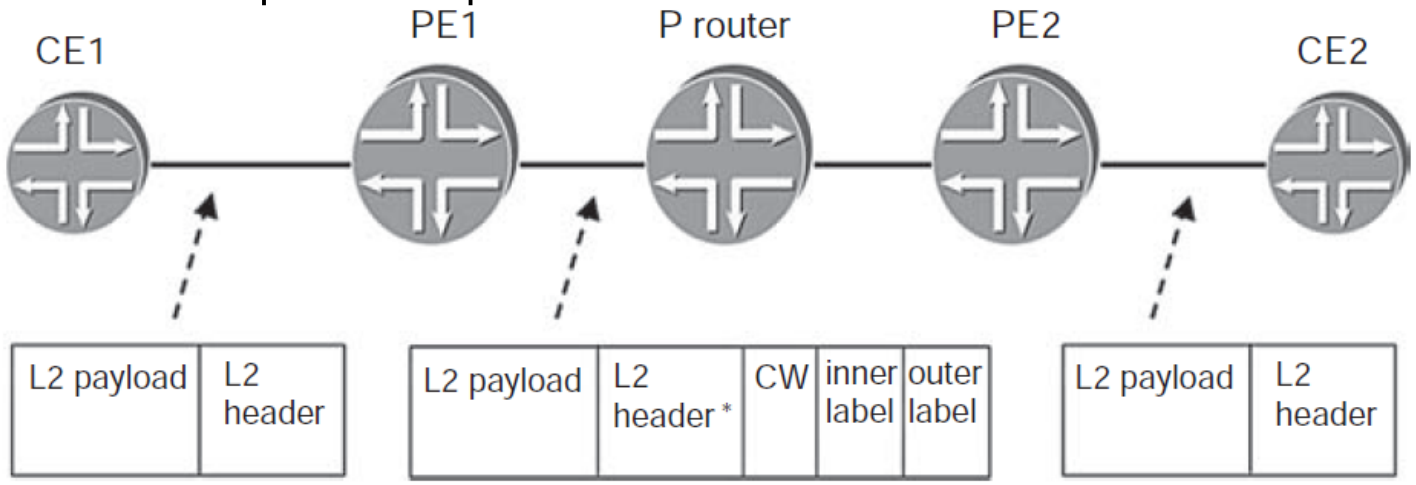
A partir de la etiqueta VPN decide que la trama debe enviarse por la VLAN200 de CE2.

Si CW está presente, PE2 puede comprobar el número de secuencia.

PE2 regenera la trama:

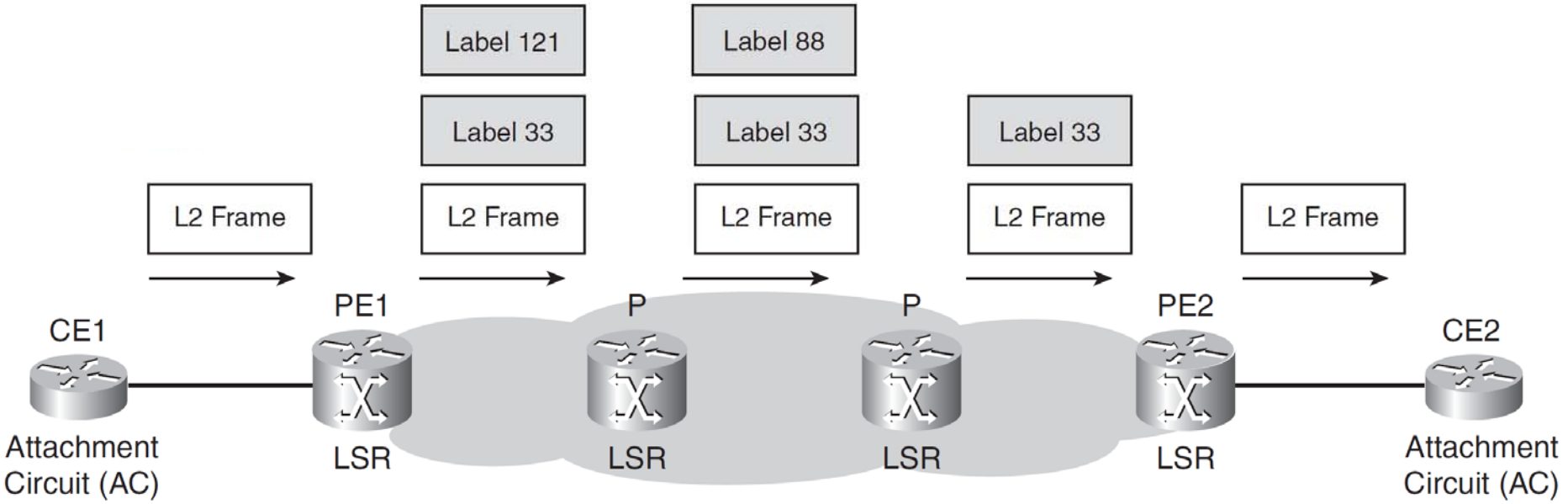
- Puede incluir la determinación de los bits de control de la cabecera de la trama a partir de la CW.
- En el ejemplo, la trama Ethernet llegó con VLAN100, pero como CE2 espera VLAN200, PE2 escribe el valor correcto.

PE2 envía la trama por el AC que conduce a CE2.



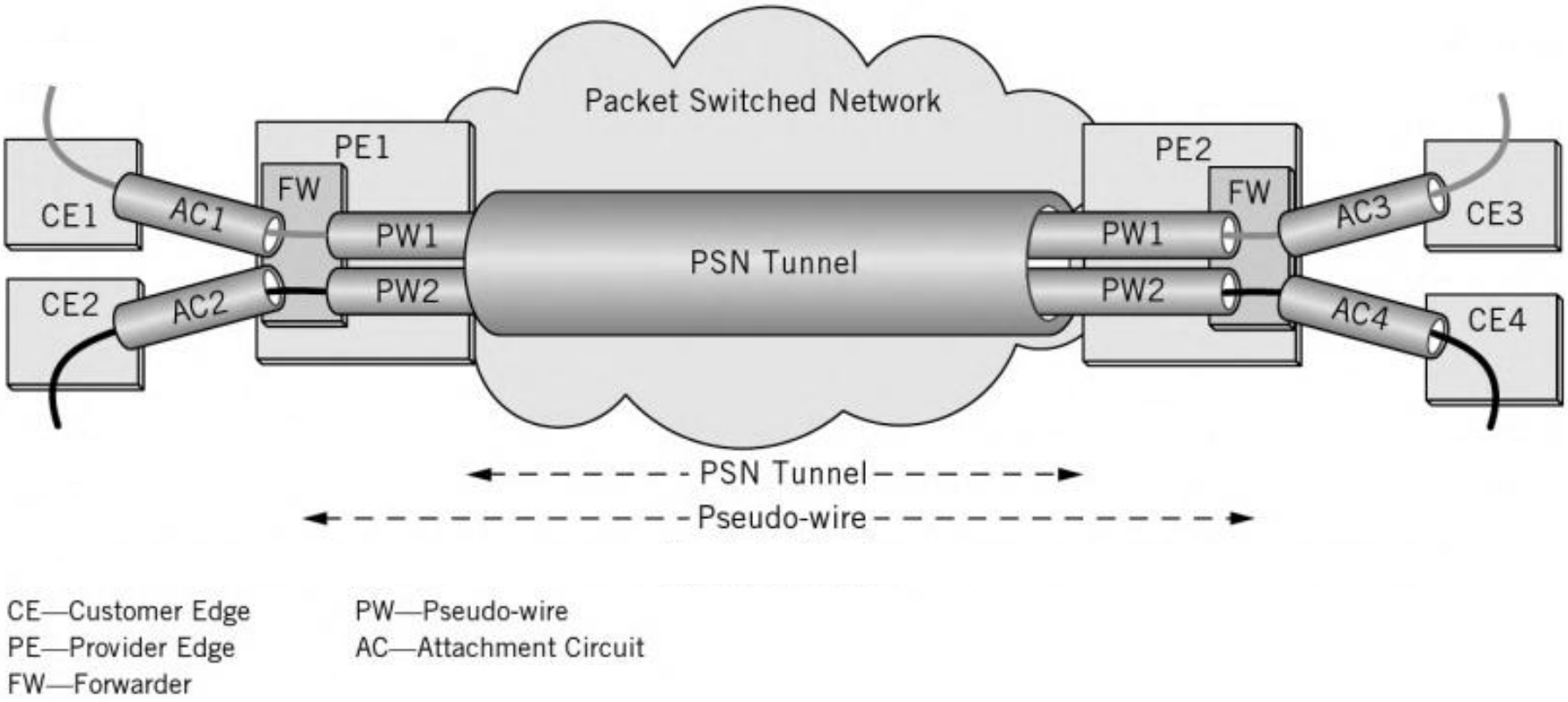
n MPLS

Plano de reenvío



- Como en las L3VPN, se puede usar MPLS para las etiquetas VPN pero el túnel de transporte puede hacerse mediante MPLS (con LDP o RSVP-TE), GRE o IPSec.

Plano de reenvío



Plano de reenvío

- Aplicación a tecnologías concretas:
  - Celdas ATM.
  - AAL5 de ATM.
  - Frame Relay.
  - Ethernet.

Plano de reenvío: celdas ATM

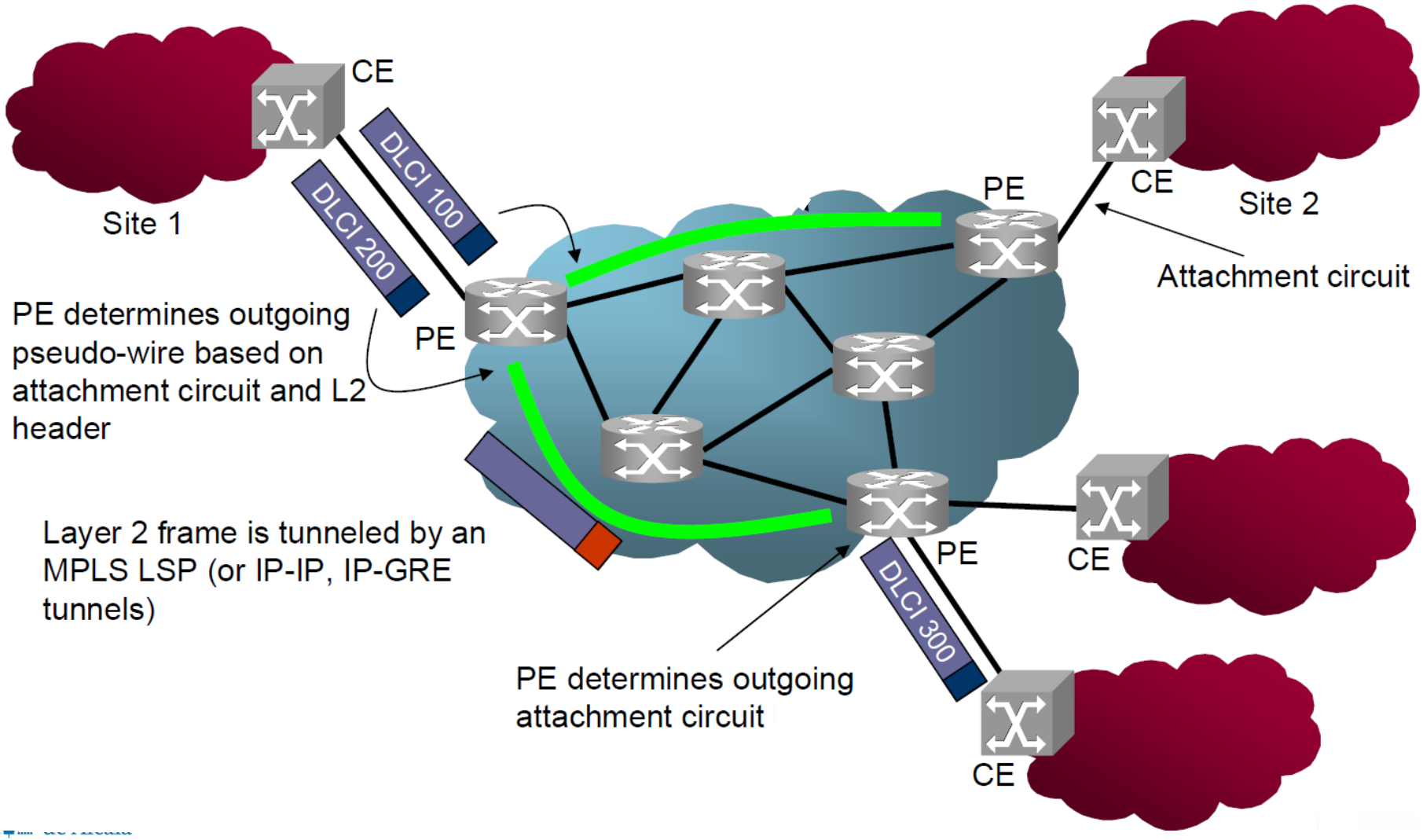
- Dos modos:
  - N a 1 (obligatorio):
    - Las celdas de una o más conexiones ATM se mapean en un único pseudocable.
    - Se mantienen los campos VPI/VCI al transportarse por el núcleo de la red:
      - Para que el PE de salida sepa a qué VPI/VCI pertenece cada celda.
    - Puede usarse, por ejemplo, para que todas las conexiones de un puerto se transporten a otro puerto remoto de la red.
  - 1 a 1 (opcional):
    - Las celdas de cada conexión se mapean en un único pseudocable.
    - No se envía el valor de VPI/VCI, puesto que puede regenerarse por el PE de salida.
- En ninguno se envía el HEC (*Header Error Check*) → Lo regenera el PE de salida.
- Se pueden enviar múltiples celdas ATM en un único paquete MPLS.

Plano de reenvío: AAL5 de ATM y FR

- AAL5 de ATM:
  - Modo 1 a 1: cada VC se mapea en un pseudocable.
  - Es más eficiente transportar paquetes AAL5 que celdas ATM.
- *Frame Relay* (FR):
  - Modo 1 a 1:
    - Cada DLCI se mapea en un único pseudocable.
    - No se transporta la cabecera FR ni el FCS.
    - Se puede usar la CW para transportar ciertos campos de la cabecera FR.
  - Modo puerto (opcional):
    - Todos los DLCIs de un puerto particular se transportan sobre un pseudocable.
    - Debe transportarse la dirección FR.

Plano de reenvío: FR

- Ejemplo:



Plano de reenvío: Ethernet

- Dos modos:
  - Por VLAN:
    - Cada VLAN se mapea en un único pseudocable.
  - Por puerto:
    - Cada puerto Ethernet se mapea en un único pseudocable.
- El PE de entrada elimina FCS y lo reconstruye el PE de salida.
- La CW, si se usa, solo tiene como campo útil el número de secuencia (16b).

Plano de control

- Dos opciones básicas:
  - LDP.
  - BGP.
- Ambas tienen en común:
  - Cuando se recibe tráfico de un CE local, permite a un PE saber el valor de la etiqueta VPN que espera un PE remoto que se conecta al CE remoto.
  - Señalizar las características del pseudocable.
  - Asumen que el pseudocable es bidireccional.
  - Proveen medios para que un PE pueda indicar a un PE remoto que hay un problema de conectividad.

Plano de control

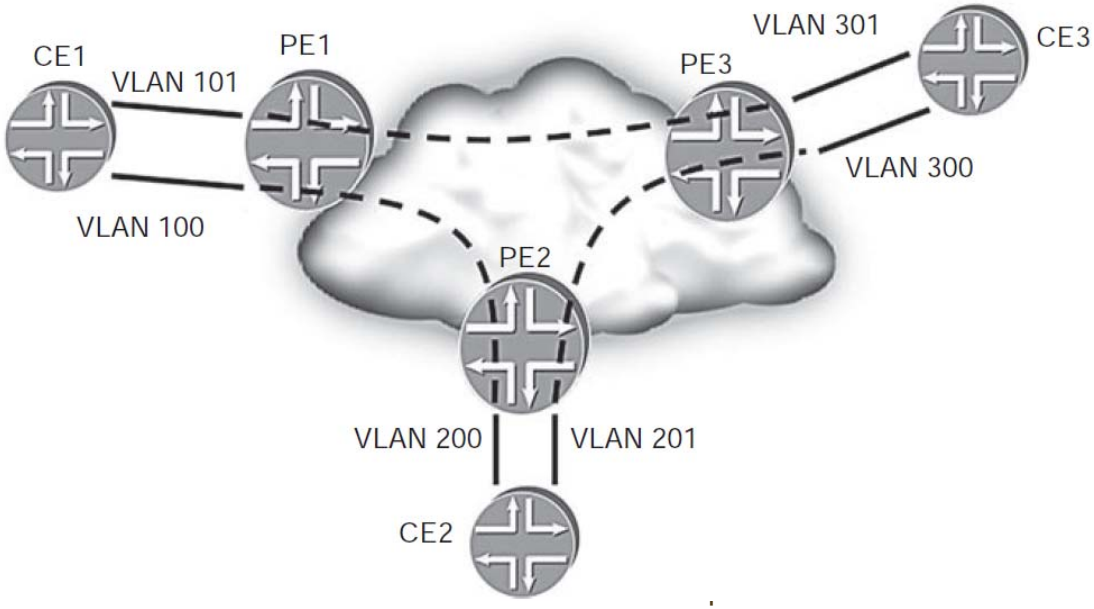
- Ambas propuestas (LDP y BGP) difieren especialmente en la manera en la que un PE sabe con qué PE remoto necesita construir un pseudocable.
  - En la propuesta LDP original, debía configurarse manualmente.
  - La propuesta BGP tiene capacidades de autodescubrimiento → No requiere la configuración manual.
  - La propuesta LDP original se amplió para permitir el autodescubrimiento usando BGP.

Plano de control: propuesta LDP original

- En la propuesta LDP original no existe el concepto de L2VPN, sino que trata de la señalización de pseudocables individuales.
- Se establece una sesión LDP (*targeted*) entre cada par de PEs de la red.
  - Para conocer el PE remoto se usa configuración manual.
  - La sesión LDP entre PE-PE se emplea para comunicar la etiqueta PW que empleará cada pseudocable.
  - Pueden haber múltiples pseudocables entre dos PEs concretos, usando la misma sesión LDP.
    - Se requiere un identificador para cada uno de ellos: CV ID o PW ID.
    - Ambos extremos del pseudocable se configuran con el mismo PW ID.

Plano de control: propuesta LDP original

- Ejemplo:
  - Mediante configuración, en PE1 se crea una asociación entre VLAN100, un PW ID y una dirección IP de PE2 (que se usa para la sesión *targeted* de LDP).
  - En PE2 se crea una asociación equivalente entre VLAN200, el mismo PW ID que en PE1 y una dirección IP de PE1.
  - PE1 emplea LDP para comunicar el valor de la etiqueta PW que PE2 debe emplear al enviar paquetes a PE1.
    - Ídem entre PE2 y PE1.



Plano de control: propuesta LDP original

- Se emplea el modo *Unsolicited Downstream* (UD) y modo de retención de etiqueta liberal.
- Un mensaje de mapeo de etiqueta LDP contiene dos TLVs obligatorios:
  - TLV etiqueta.
  - TLV FEC.
- Se define un nuevo tipo de FEC (llamado "*FEC 128*" o "*PWid FEC*":
  - Sirve para identificar el pseudocable al cual se asocia una etiqueta.
  - A diferencia de los FEC que contienen información de prefijos de red IP, ahora cada FEC TLV solo puede contener un elemento FEC.

Plano de control: propuesta LDP original

- El FEC 128 puede emplearse cuando ambos extremos del pseudocable han sido provistos del mismo PW ID.
- El FEC 128 contiene:
  - *PW Type*: tipo de circuito virtual.
  - *PW ID*: junto con *PW Type* sirve para identificar a un PW.
  - Bit CW (*Control Word*): indica si se usa palabra de control en este pseudocable (CW=1) o no (CW=0).
- Se debe emplear el mismo *PW ID* y *PW Type* en ambos extremos del pseudocable.
- Usando FEC 128 ambos extremos del pseudocable inician, de manera independiente, el establecimiento de un LSP unidireccional.
  - LSPs del sentido de ida y de vuelta con el mismo PW Type y PW ID → Forman un único pseudocable.

Plano de control: propuesta LDP original

- No hay concepto de VPN en este esquema:
  - Los pseudocables se crean de manera independiente.
  - Si se requiere una estructura *full mesh*, cada vez que se añada un nuevo CE hay que crear pseudocables manualmente con los CEs ya existentes de esa VPN.
    - Problema de escalabilidad, mitigado mediante el uso de la propuesta BGP y de la propuesta LDP original extendida con BGP.

VPNs de capa 2.5

- Los esquemas propuestos para L2VPN requieren que ambos extremos empleen el mismo protocolo de capa 2 en ambos extremos → Relajar esa restricción. Útil, por ejemplo cuando:
  - Diferentes tecnologías en diferentes sedes debido a una nueva adquisición o a estar en proceso de migración.
- Interconexión de tecnologías de capa 2 → VPNs de capa 2.5.
- Funcionamiento:
  - Cuando un PE recibe una trama se le quita el encapsulamiento de capa 2, dejándose el paquete de capa 3.
    - Como no se puede indicar qué protocolo de capa 3 se emplea → Solo funciona con IP.
  - Al paquete IP se le añaden las etiquetas MPLS y PW para transportarlo por la red.
  - El PE de salida extrae el paquete IP y le añade la cabecera apropiada a la tecnología de capa 2 empleada en la red de salida.

VPLS

Ingeniería de tráfico - Curso 2013-14

Área de Ingeniería Telemática - Departamento de Automática

Universidad de Alcalá

José Manuel Giménez

VPLS

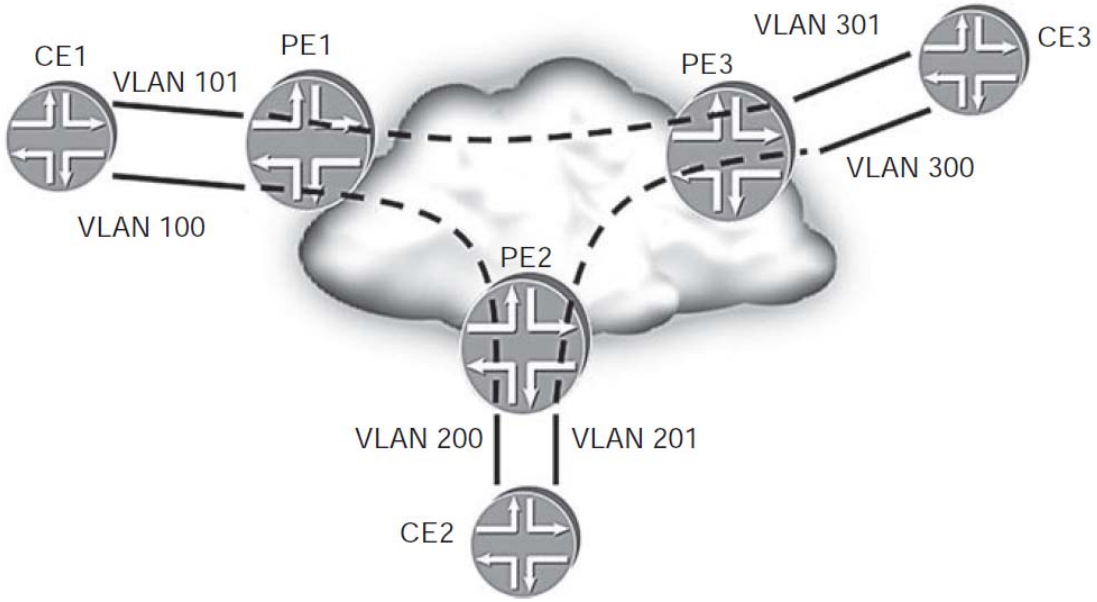
- Introducción.
- Funcionamiento.
- Plano de reenvío.

VPLS: introducción

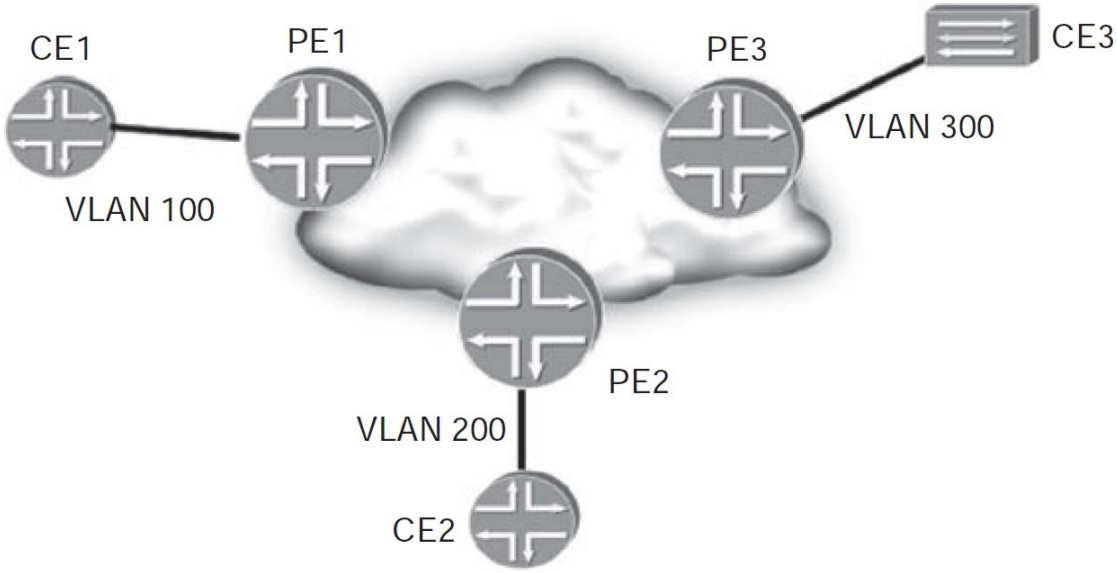
- VPLS: *Virtual Private LAN Service*.
- Cada cliente ve el resto de sus sedes como si estuvieran en la misma LAN, pese a que en realidad están conectadas mediante la red del proveedor.
- En L3VPN el cliente debe encargarse de configurar el encaminamiento entre CE y PE.
- En L2VPN el cliente debe controlar la creación de los pseudocables.
- ¿Y si una compañía no tiene expertos para sus comunicaciones?
  - VPLS permite la interconexión de equipos sobre la red de un proveedor como si estuvieran en la misma LAN.
  - Requiere menos conocimientos por parte del cliente que L3VPN y L2VPN.

VPLS: introducción

- En L2VPN se requieren múltiples interfaces lógicas (ej. VLAN) entre cliente y PE.

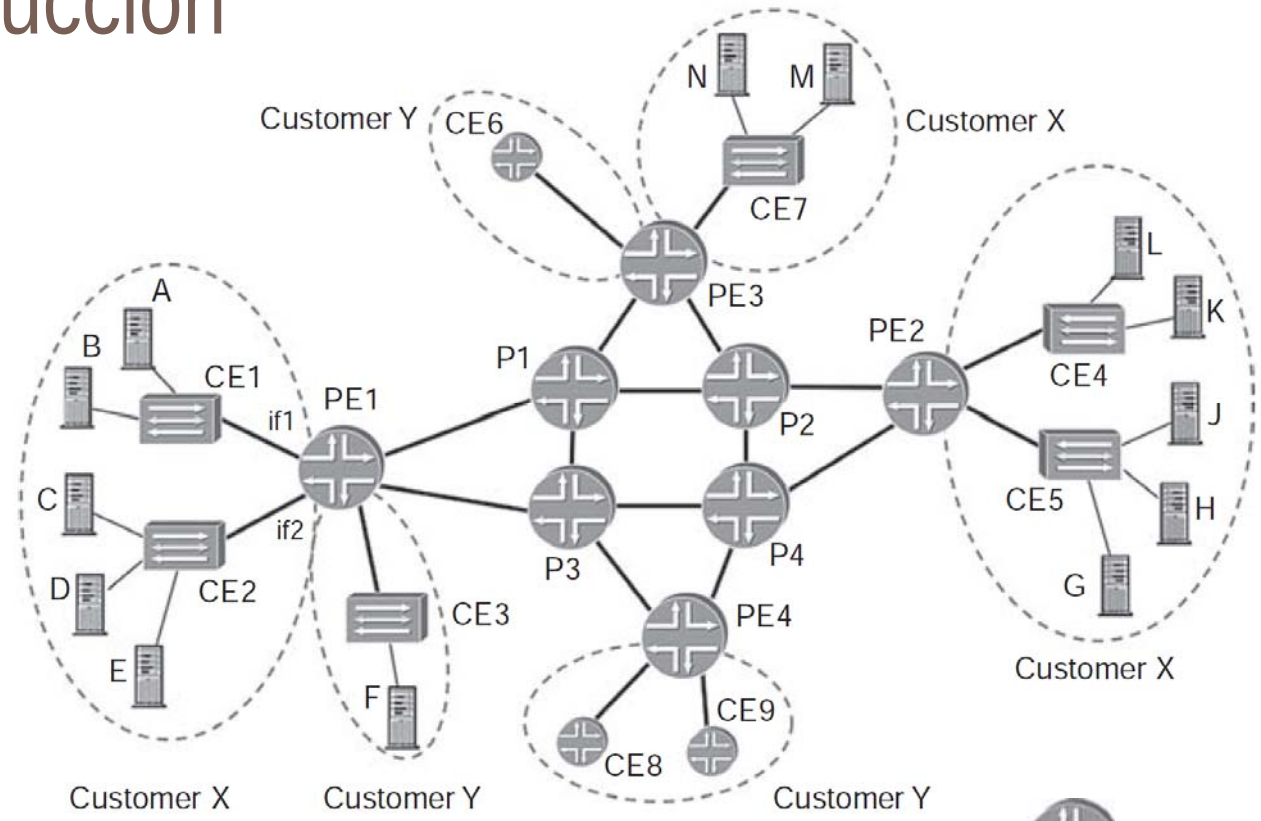


- En VPLS una única interfaz lógica es suficiente.
  - Se debe a que VPLS es multipunto.
  - Cada PE sabe a qué sedes debe enviar cada trama.

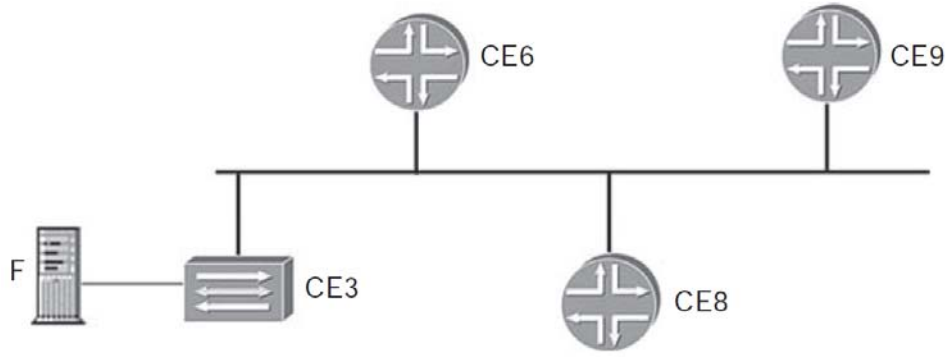


VPLS: introducción

- Red real:



- Red que ve un cliente de Y:



VPLS: funcionamiento

- Para cada VPLS, los routers PE están unidos mediante pseudocables de manera *full mesh*.
  - Cuando un PE recibe una trama de otro PE sabe a qué VPLS pertenece la trama en función de su etiqueta de pseudocable.
  - Además de la etiqueta de pseudocable se requiere un túnel de transporte: MPLS, GRE o IPSec.
- Problema: en el caso de Ethernet, cada PE debe saber donde enviar cada trama en función de su dirección MAC destino.
  - Direcciones MAC no siguen ninguna jerarquía en las sedes de la VPN.
  - Los PEs tienen un mecanismo de aprendizaje de direcciones MAC aprendiendo de la dirección MAC origen cada vez que reciben una trama.

VPLS: funcionamiento

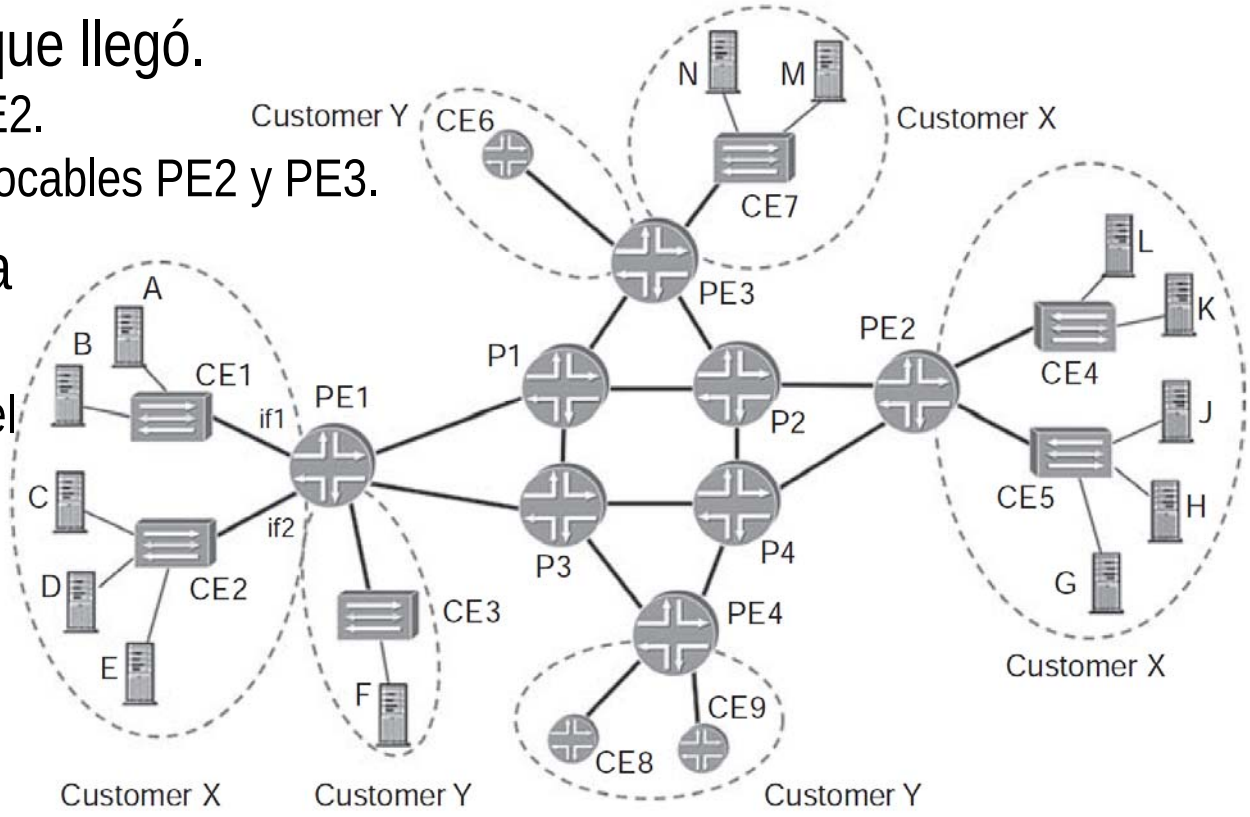
- Si el cliente usa como CE:
  - Conmutadores (*switches*): se deben aprender las direcciones MAC de todos los hosts que se conecten a dicho *switch*.
    - Problemas de escalabilidad porque las direcciones MAC no están jerarquizadas → Se suele limitar el tamaño de las tablas de direcciones MAC aprendidas.
    - Solo se recomienda para sedes pequeñas.
  - *Routers*: se aprende la dirección MAC de la interfaz Ethernet que se conecta del CE al PE.
    - Solución recomendada para sedes de clientes con muchos hosts.

VPLS: plano de reenvío

- El usuario configura qué puertos locales son miembros de cada VPLS en cada PE.
- Cada PE mantiene una tabla de reenvío separada para cada VPLS.
- Requisito: para cada VPLS, los *routers* PE deben estar totalmente conectados (topología *full mesh*).
  - Sin pasar por ningún PE intermedio (sí que se puede pasar por *routers* P).

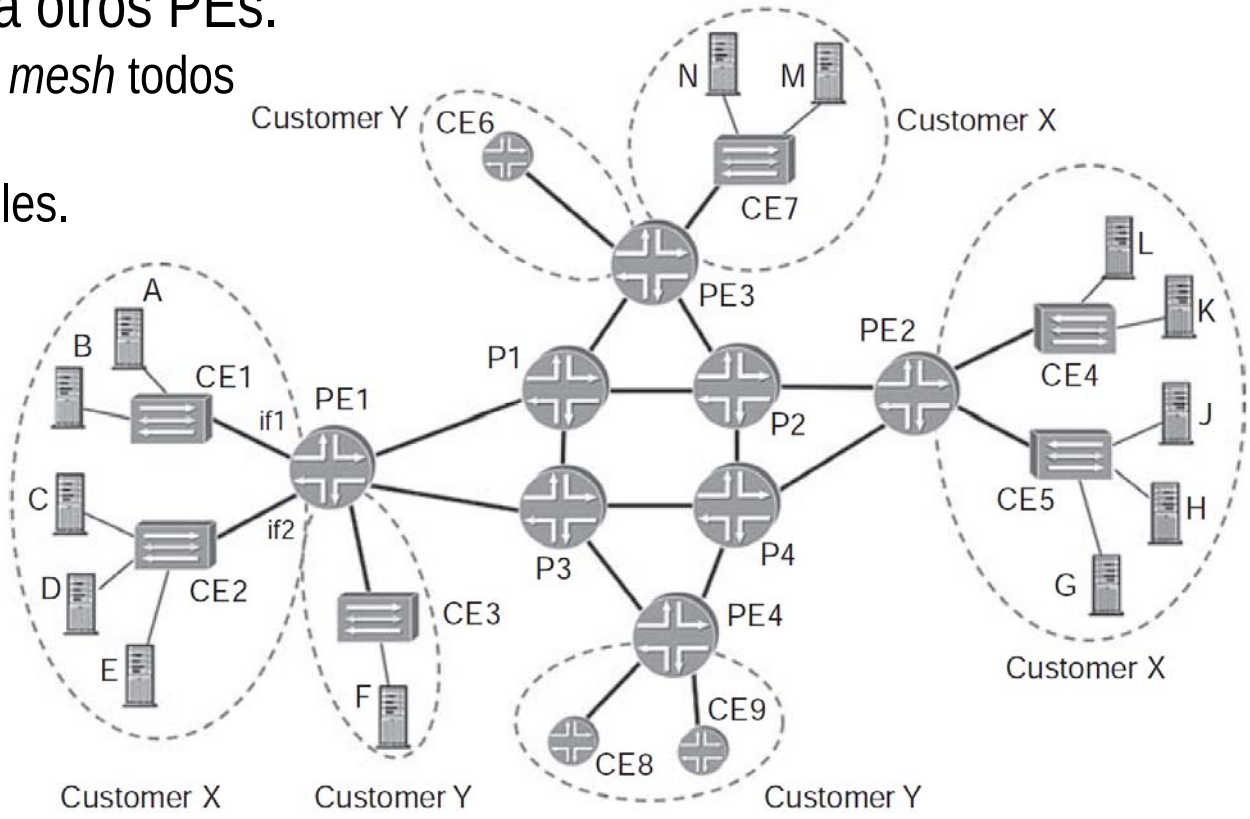
VPLS: plano de reenvío

- Ejemplo de reenvío de tramas *unicast* desde A (MAC A) hasta J (MAC J):
  - Supongamos que PE1 no sabe donde está MAC J.
- PE1 inunda la trama por todos los puertos excepto por el que llegó.
  - Se envía por puerto local CE2.
  - Se envían copias por pseudocables PE2 y PE3.
- PE2 y PE3 saben que la trama es de la VPLS X.
  - Debido al pseudocable por el que llega dicha trama.



VPLS: plano de reenvío

- Si ni PE2 ni PE3 conocen la MAC J:
  - PE2 inunda la trama a todos sus puertos locales (CE4 y CE5).
  - PE3 envía la trama a CE7 (a CE6 no puesto que no pertenece a VPLS X).
- Ni PE2 ni PE3 inundan a otros PEs.
  - Como todos los PEs son *full mesh* todos han recibido la trama.
  - Se evita la formación de bucles.



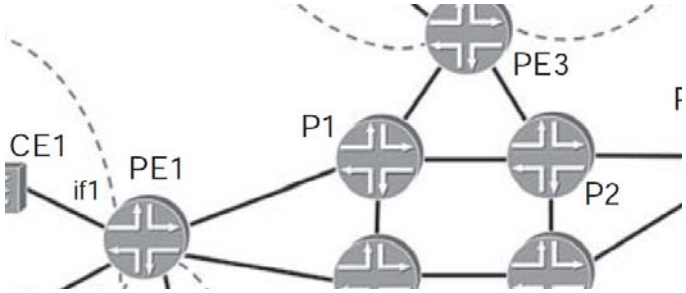
VPLS: plano de reenvío

- La recepción de tramas con dirección MAC origen A permite aprender la localización de A en términos del puerto o pseudocable por el que llegó la trama.
  - PE1 incluye en su tabla de reenvío:

Dirección MAC	Siguiente salto
A	if1

  - PE2 y PE3 incluyen en sus tablas de reenvío una asociación entre MAC A y sus respectivos pseudocables hacia PE1:

Dirección MAC	Siguiente salto
A	PW_PE1

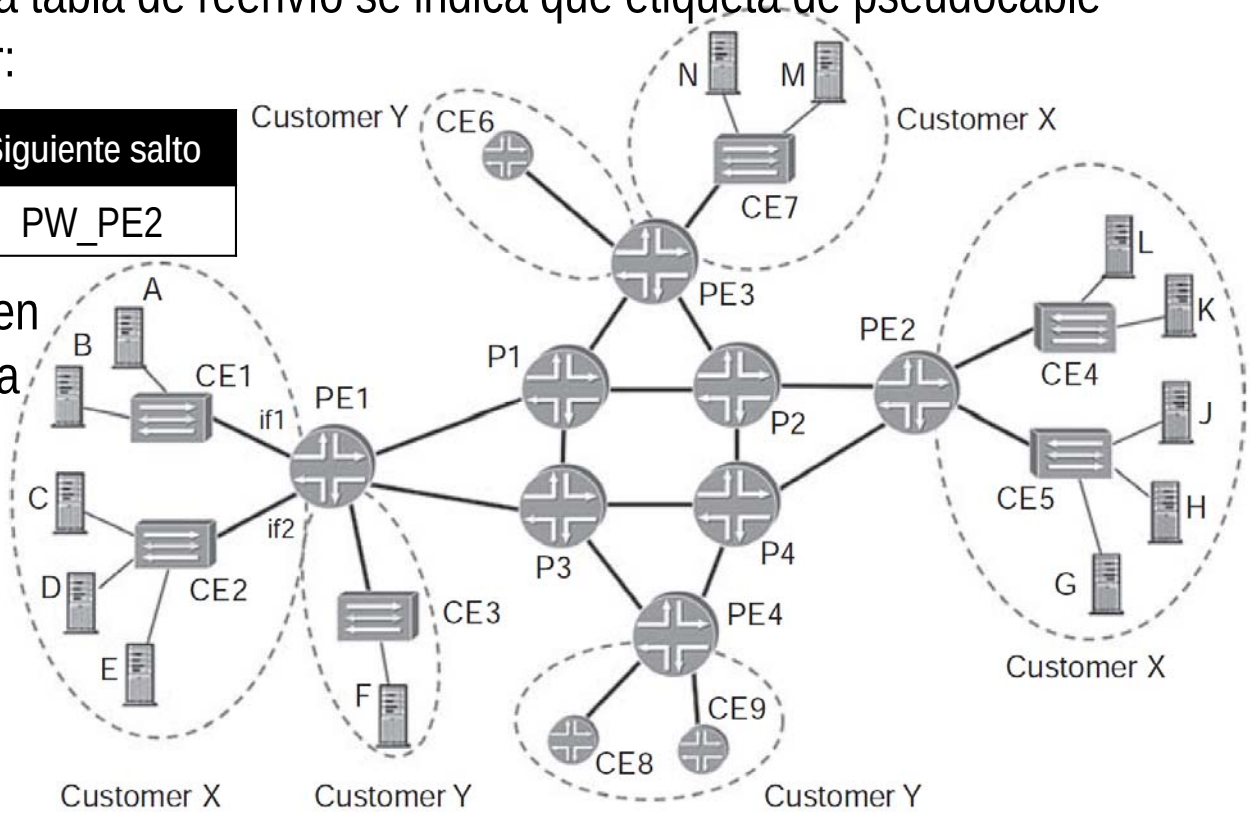


VPLS: plano de reenvío

- Si J empieza a enviar tramas a A:
  - PE2 tiene una entrada en su tabla de reenvío que indica que hay que mandarla por el pseudocable hacia PE1 (no necesita inundar).
  - Cuando PE1 recibe la trama aprende que para llegar a J hay que usar el pseudocable hacia PE2.
    - En realidad, en la tabla de reenvío se indica qué etiqueta de pseudocable hay que emplear:

Dirección MAC	Siguiente salto
J	PW_PE2

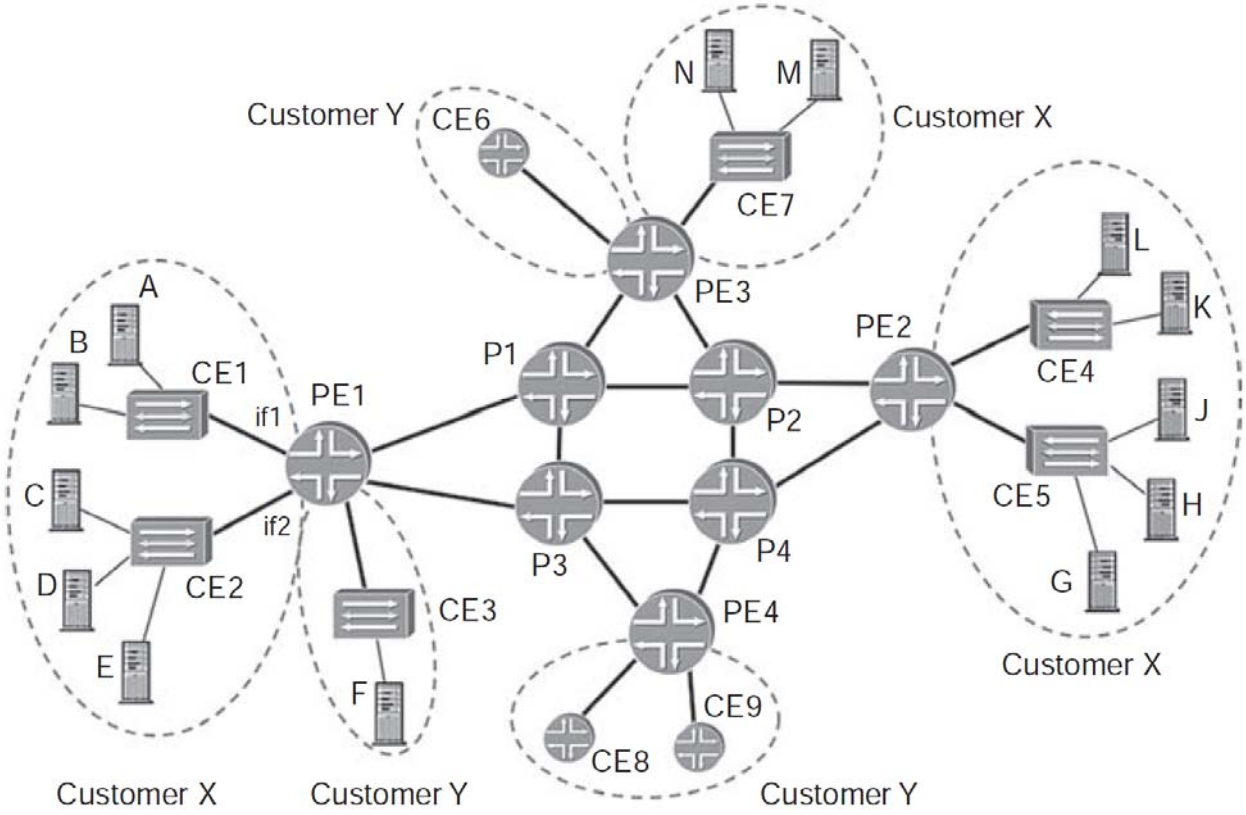
  - Como PE1 ya tiene en su tabla de reenvío la dirección MAC A la envía por if1.



VPLS: plano de reenvío

- Se pueden emplear túneles construidos mediante MPLS, GRE o IPSec.
- Un túnel de transporte se puede compartir entre varios pseudocables.
- Aprendizaje de PE1 (PW: etiqueta pseudocable, T: etiqueta transporte):

Dir MAC	Siguiente salto
A	if1
B	if1
C	if2
D	if2
E	if2
G	PW: 200, T: 410
H	PW: 200, T: 410
J	PW: 200, T: 410
K	PW: 200, T: 410
L	PW: 200, T: 410
M	PW: 300, T: 235
N	PW: 300, T: 235



VPLS: plano de reenvío

- Las tramas *broadcast* se envían mediante inundación.
- Resumen:
  - No se emplea el anuncio de direcciones MAC usando el plano de control, sino que se realiza el aprendizaje de dichas direcciones.
  - Se suelen eliminar las direcciones MAC aprendidas una vez se cumple cierto tiempo sin usarlas.
  - Se suele limitar el tamaño máximo de las tablas de reenvío eliminando las direcciones MAC que llevan más tiempo sin usarse.

Bibliografía

- I. Minei y J. Lucek, "*MPLS-Enabled Applications*", John Wiley & Sons, 3rd Ed, 2011.
- L. De Ghein, "*MPLS Fundamentals*", Cisco Press, 2007.
- V. Alwayn, "*Advanced MPLS Design and Implementation*", Cisco Press, 2001.