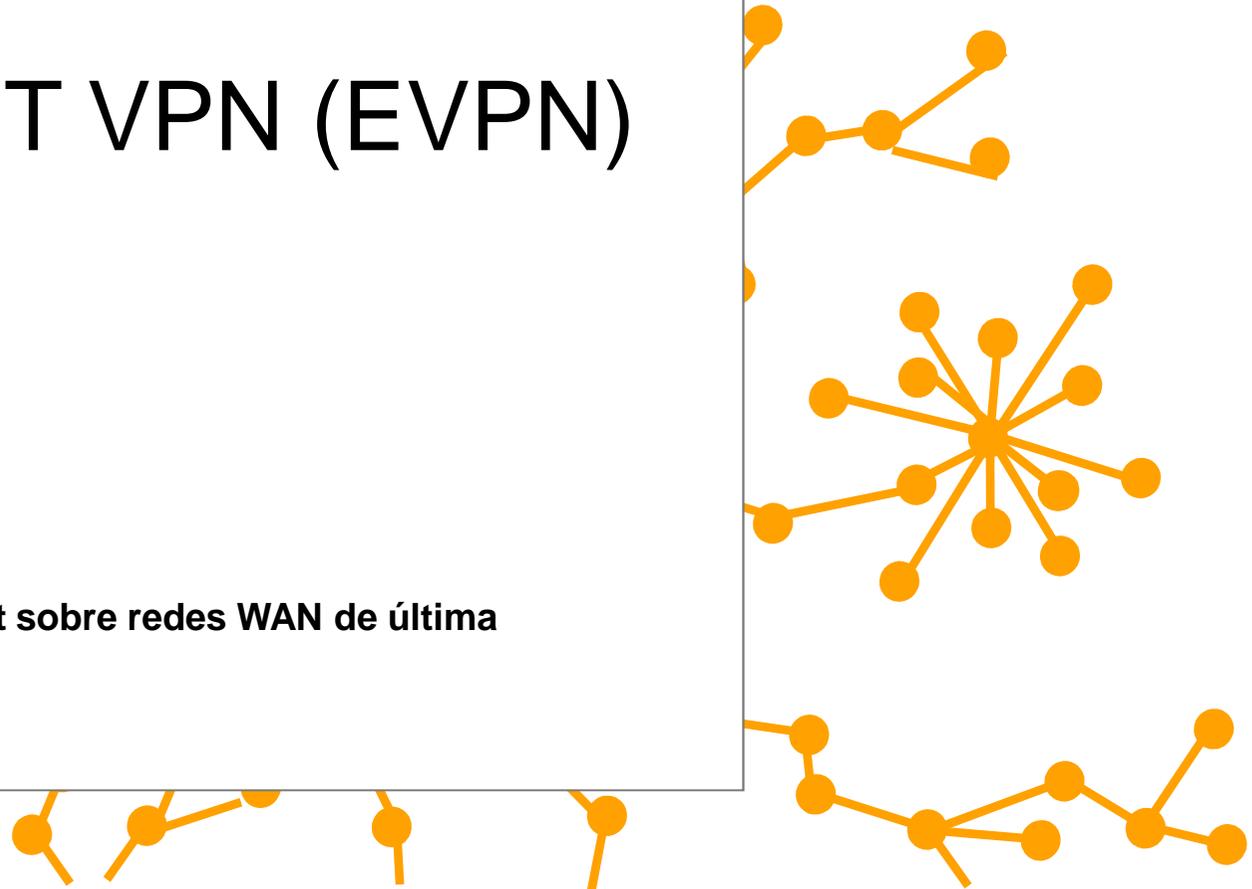


**indra**

# ETHERNET VPN (EVPN)

**Interconexión L2-Ethernet sobre redes WAN de última generación: E-VPN**



# AGENDA

1. Introducción
2. Limitaciones resueltas por EVPN
3. Funcionamiento de EVPN
4. Integración de L2 y L3: caso DCI (Juniper)
5. Referencias

# Introducción-I

## ¿Qué es EVPN?

- Es un nuevo protocolo para interconectar dominios de L2 (Ethernet) a través de redes IP/MPLS.
- En proceso de estandarización (drafts, RFCs) a través de varios grupos de trabajo de “IETF Internet L2 Working group”.
- Algunos de los autores de las especificaciones de EVPN son:
  - Fabricantes: Juniper, Cisco, Alcatel-Lucent
  - Service Providers: AT&T, Verizon, Bloomberg
- Actualmente ya existen implementaciones comerciales basadas en los drafts: Juniper (EVPN), Cisco (PBB-EVPN), Alcatel-Lucent (EVPN, PBB-EVPN)

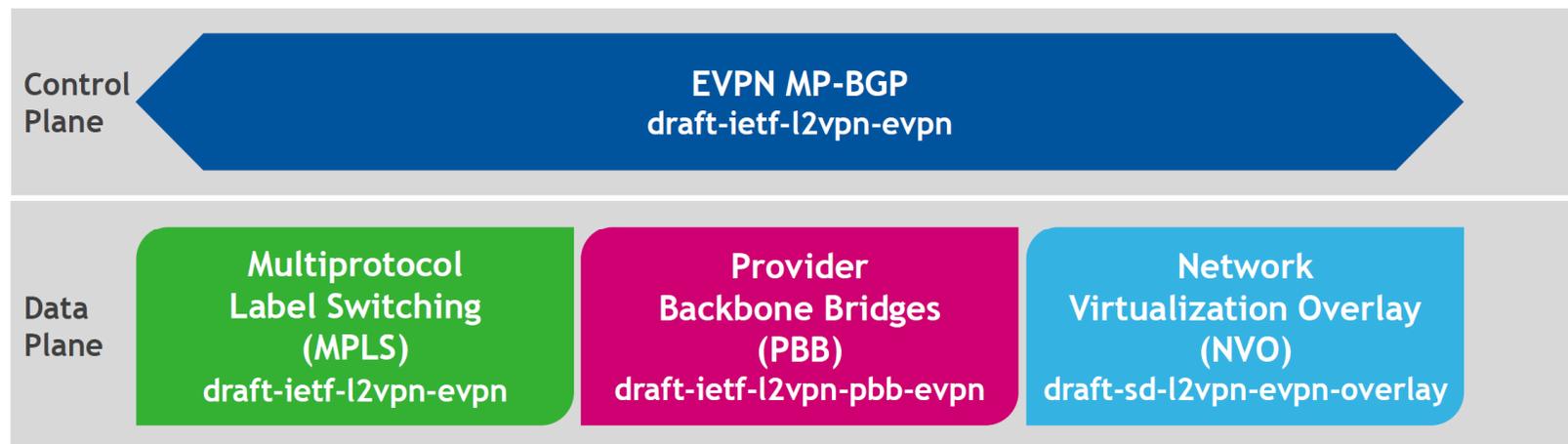
## ¿Por qué EVPN?

- Aumenta de forma significativa la eficiencia y el control del tráfico sobre las actuales soluciones de L2VPN Eth (VPLS).
- Resuelve de forma óptima los requerimientos de “Virtual Machine mobility” entre Data Centers interconectados a través de redes WAN.

# Introducción-II

## Variantes de EVPN

- Existen tres variantes de la tecnología dependiendo de la solución adoptada en el plano de datos: EVPN, PBB-EVPN y EVPN-VXLAN.
- Todas ellas gestionan el aprendizaje de rutas entre PEs (plano de control) mediante MP-BGP.
- EVPN usa MPLS/IP en el plano de datos. Es la solución original de EVPN en la especificación base de la tecnología.
- PBB-EVPN combina IEEE 802.1ah PBB con EVPN para aumentar la escalabilidad de MACs, agregando C-MACs sobre B-MACs. EVPN-VXLAN usa VXLAN en el plano de datos.
- Esta presentación versa sobre la primera de ellas: EVPN.



# Limitaciones resueltas por EVPN - I

## Integración de servicios

- Permite ofrecer servicios de L2 y L3 sobre el mismo interfaz de acceso, VLAN y VPN.
- Operación y provisión similares a L3VPN, heredando sus características en cuanto a escalabilidad y control.

## Eficiencia

- Permite balanceo de carga por flujo en entornos “all-active multihoming” (con todos los links de acceso al servicio activos).
- Optimiza la gestión del tráfico Broadcast, Unknown Unicast y Multicast (BUM).

## Flexibilidad

- Diferentes opciones de encapsulación en el plano de datos: MPLS, IP, VXLAN.
- Soporta los servicios actuales E-LAN, E-Line y E-TREE.
- Simplifica la provisión y operación de los servicios usando una sola tecnología VPN.

# Limitaciones resueltas por EVPN-II

## Redundancia

- Facilita la convergencia rápida de la red ante fallos (link / nodo) o ante “Virtual Machine mobility”.

## Control

- El proceso de aprendizaje de MAC/IP entre PEs se realiza en el plano de control (MP-BGP) y no en el plano de datos.
- Implementa Proxy-ARP/ND en los PEs para responder a las solicitudes de MAC/IP de ubicaciones remotas en local.

## Escalabilidad

- Facilita la provisión mediante funcionalidades de auto-descubrimiento de PEs remotos o de PEs locales (multihoming).
- MP-BGP facilita la escalabilidad en plano de control (RR).
- Elimina la necesidad de mallado completo de PWs (T-LDP sessions) entre PEs participantes de una L2VPN.
  - Tráfico Unicast → MP2P LSPs (similar a L3VPNs).
  - Tráfico Multicast → MP2MP, MP2P ó P2P LSPs mediante “ingress replication”.

# Limitaciones resueltas por EVPN-III

## Comparativa VPLS vs EVPN

- EVPN resuelve las carencias de VPLS respecto a topologías multihoming “all-active”, control de tráfico y escalabilidad.

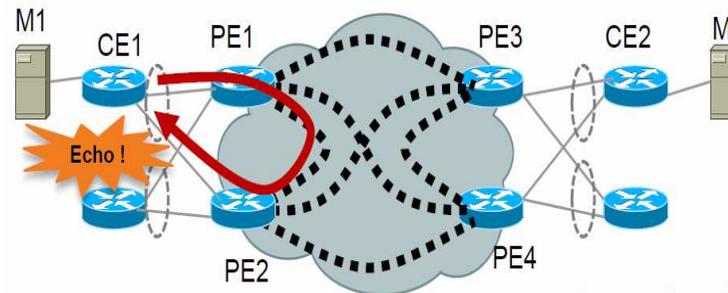
Requirement	VPLS	PBB-EVPN	EVPN
PE Auto-discovery	✓	✓	✓
Access multi-homing (Single Active)	✓	✓	✓
All-active multi-homing (flow based load balancing)		✓	✓
Fast convergence on failure	✓	✓	✓
MAC mobility	✓	✓	✓
Control plane scaling of large number of MACs	(No CP)	✓	✓
Policy up to MAC/IP address level granularity (enable flexible topologies)			✓
Avoid flooding of unknown unicast MAC and ARP			✓
Faster convergence based on local repair during failures			✓
Optimized VM (MAC+IP) mobility & inter-subnet routing			✓
Ease of configuration, operation and manageability			✓
Per service COS			✓

# Limitaciones resueltas por EVPN-IV

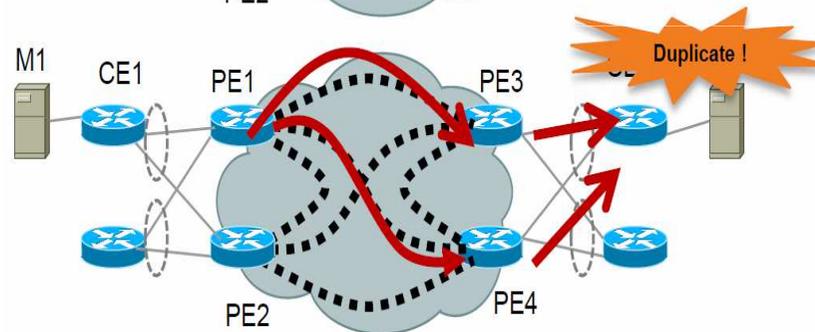
## Requerimientos no resueltos por VPLS

- Se evidencian en topologías “all-active multihoming” a la hora de evaluar redundancia y eficiencia en base a flujo de datos.

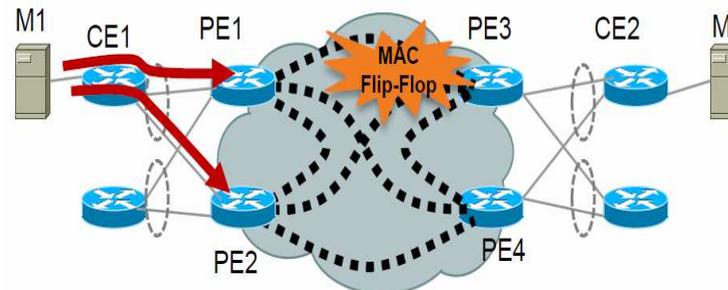
- Retorno de tráfico BUM al mismo segmento de red que lo origina.



- Duplicidad de tráfico BUM en sentido CORE→LAN (PE→CE).



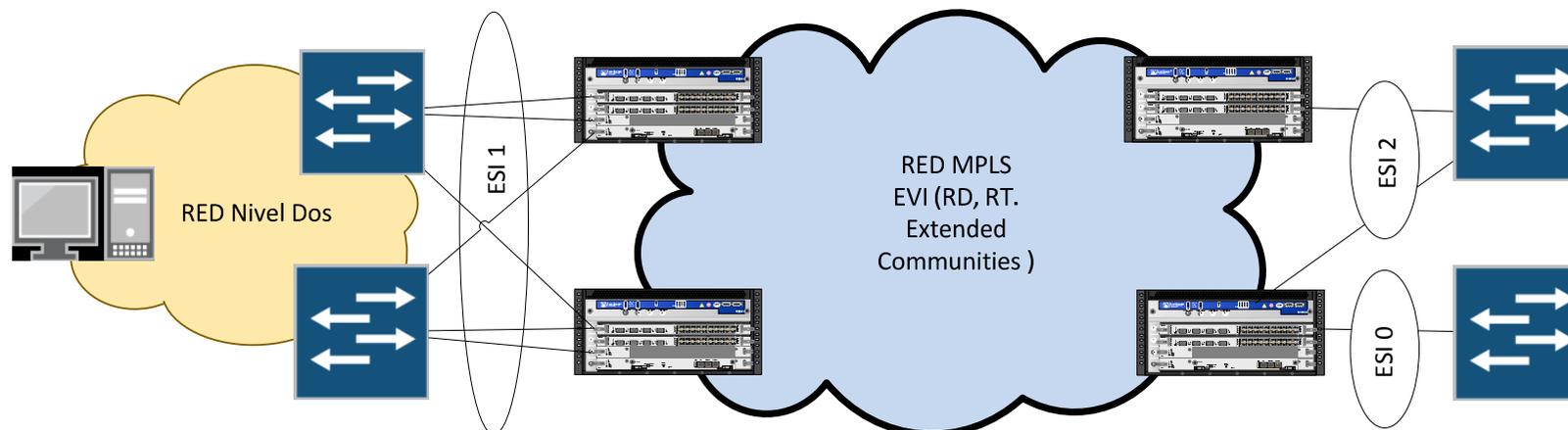
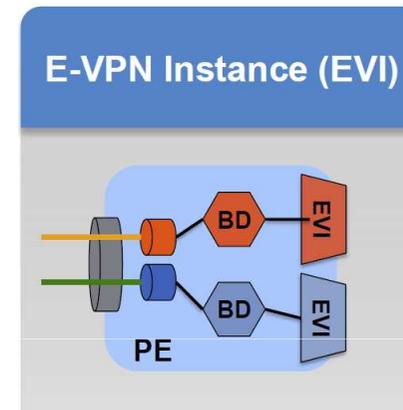
- “MAC flip-flopping” debido a balanceo de carga en enlaces Multi-Chassis LAG.



# Funcionamiento de EVPN-I

## Nomenclatura

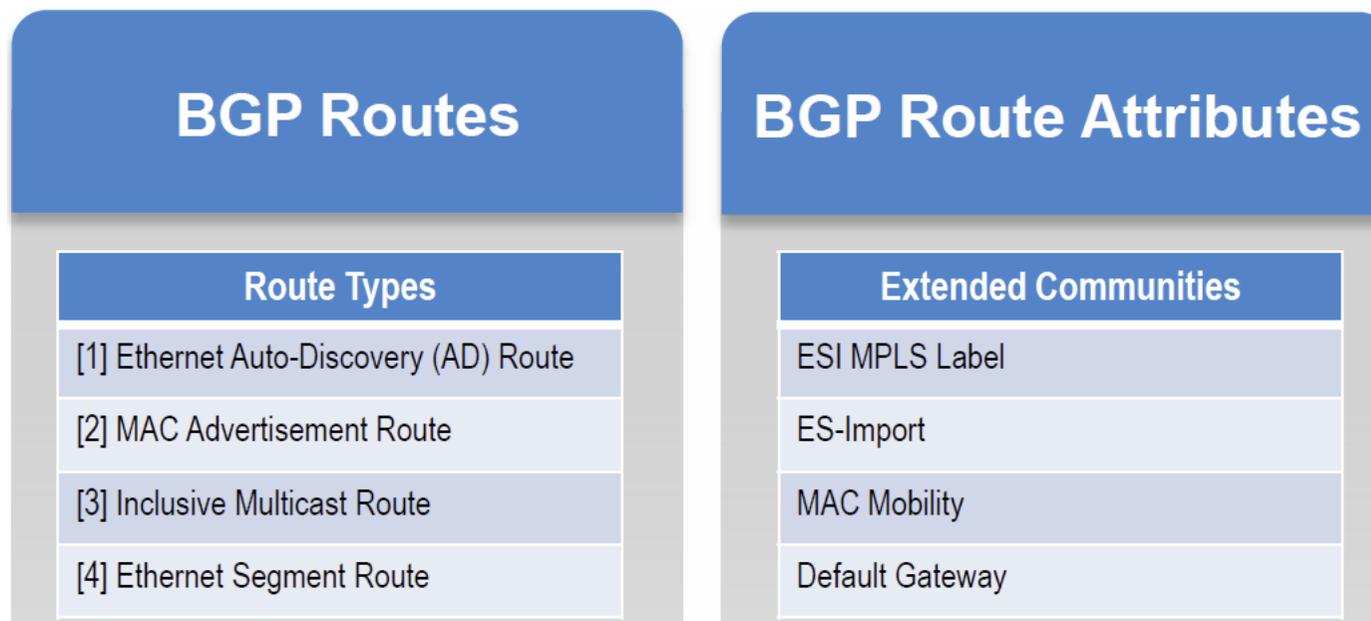
- EVI → Instancia tipo EVPN que se extiende a los PEs que participan en ella.
- MAC-VRF → tabla de rutas de una instancia EVPN.
- ESI → segmento de red conectado a los PEs que implementan la instancia EVPN
  - Single-Homed Device (SHD) → ESI = 0
  - Multi-Homed Device (MHD)
  - Single-Homed Network (SHN) → ESI = 0
  - Multi-Homed Network (MHN)
- Ethernet TAG → Identifica un dominio de broadcast en particular (por ejemplo una VLAN, un BD).
- Modo S-A → redundancia “single-active”.
- Modo A-A → redundancia “all-active”.



# Funcionamiento de EVPN-II

## Plano de control / datos en EVPN

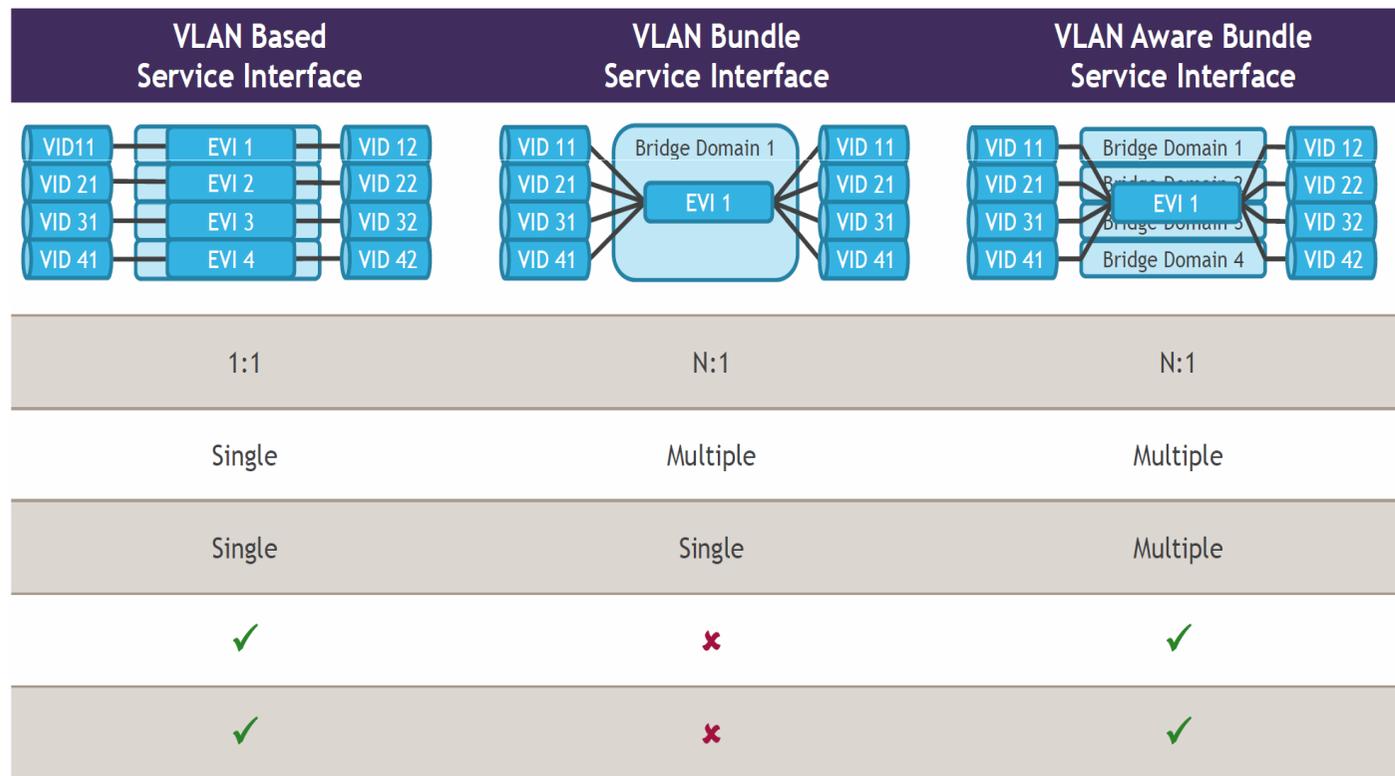
- Entre PEs se introduce el concepto de “routing” de direcciones MAC (o mejor dicho, duplas MAC/IP) mediante MP-BGP.
- Entre PE-CE, se mantiene el aprendizaje de MACs por “flood-and-learn” (en plano de datos).
- El uso de MP-BGP entre PEs permite aplicar políticas de routing y forwarding de MACs, criterios de TE etc.
- Anuncios MP-BGP AFI = 25 (VPLS) SAFI = 70 (EVPN), además de Route Distinguisher, Route Targets, BGP communities etc.



# Funcionamiento de EVPN-III

## Interfaces de servicio

- Se definen tres tipos de interfaces de servicio: “VLAN based”, “VLAN bundle” y “VLAN aware”.
- Existe un cuarto tipo que definiría el servicio más básico, denominado “Port based”.



# Funcionamiento de EVPN-IV

## Tipos de rutas y communities BGP

- TIPO 1 → tiene diferentes usos para entornos multihoming. Usa la community “ESI Label”.
- TIPO 2 → ruta principal de EVPN. Anuncia las duplas MAC/IP a otros PEs de la red, así como el ESI donde se ubica. Usa las communities “MAC Mobility” y “Default Gateway”.
- TIPO 3 → ruta usada para descubrimiento de PEs multihoming y elección de DF, usa la community “ES-Import”.

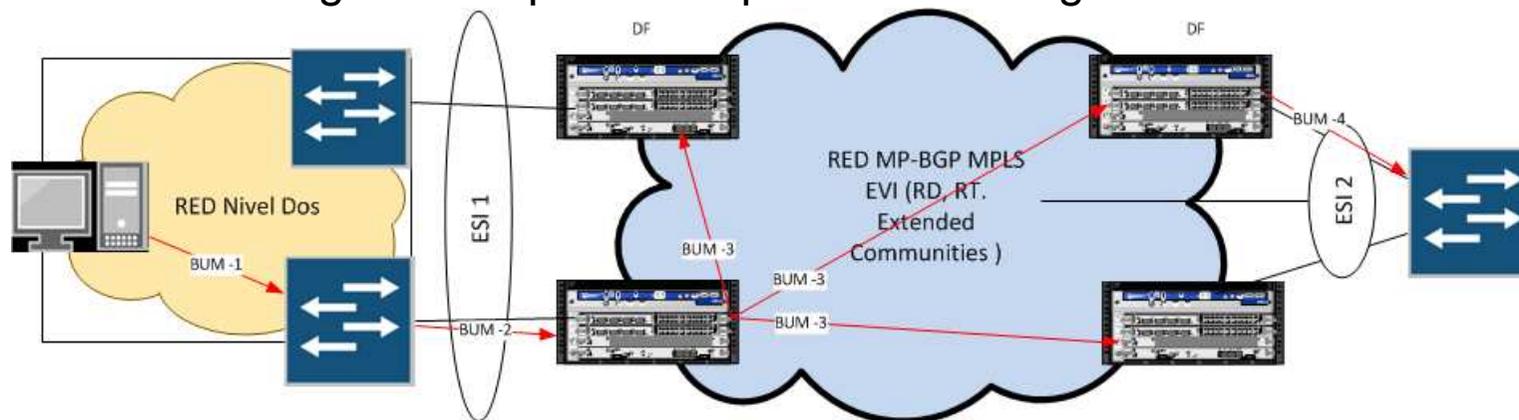
Route Type	Route Description	Route Usage	Reference
1	Ethernet Auto-Discovery (A-D) Route	Endpoint Discovery, Aliasing, Mass-Withdraw	draft-ietf-l2vpn-evpn
2	MAC Advertisement Route	MAC/IP Advertisement	draft-ietf-l2vpn-evpn
3	Inclusive Multicast Route	BUM Flooding Tree	draft-ietf-l2vpn-evpn
4	Ethernet Segment Route	Ethernet Segment Discovery, DF Election	draft-ietf-l2vpn-evpn
5	IP Prefix Route	IP Route Advertisement	draft-rabadan-l2vpn-evpn-prefix-advertisement

Extended Community Type	Extended Community Description	Extended Community Usage	Reference
0x06/0x01	ESI Label Extended Community	Split Horizon Label	draft-ietf-l2vpn-evpn
0x06/0x02	ES-Import Route Target	Redundancy Group Discovery	draft-ietf-l2vpn-evpn
0x06/0x00	MAC Mobility Extended Community	MAC Mobility	draft-ietf-l2vpn-evpn
0x03/0x030d	Default Gateway Extended Community	Default Gateway	draft-ietf-l2vpn-evpn, bgp-extended-communities

# Funcionamiento de EVPN-V

## “All-active multihoming” y elección de DF

- Los PEs conectados al mismo ESI se autodescubren mediante la ruta “Ethernet Segment”.
- Cada PE habilita de forma automática filtros para importar rutas tipo 4 con la communitie “ES-Import”.
- Uno de los PEs conectados al mismo ESI se elige como DF (Designated Forwarder).
- Se encarga de enviar el tráfico BUM recibido del core hacia los CEs conectados al ESI (el resto de PEs filtran ese tráfico → CE).
- La elección de DF en un site se realiza de forma flexible:
  - Pueden existir múltiples DFs para balanceo de carga.
  - Se elige un DF por ESI o por Ethernet tag.

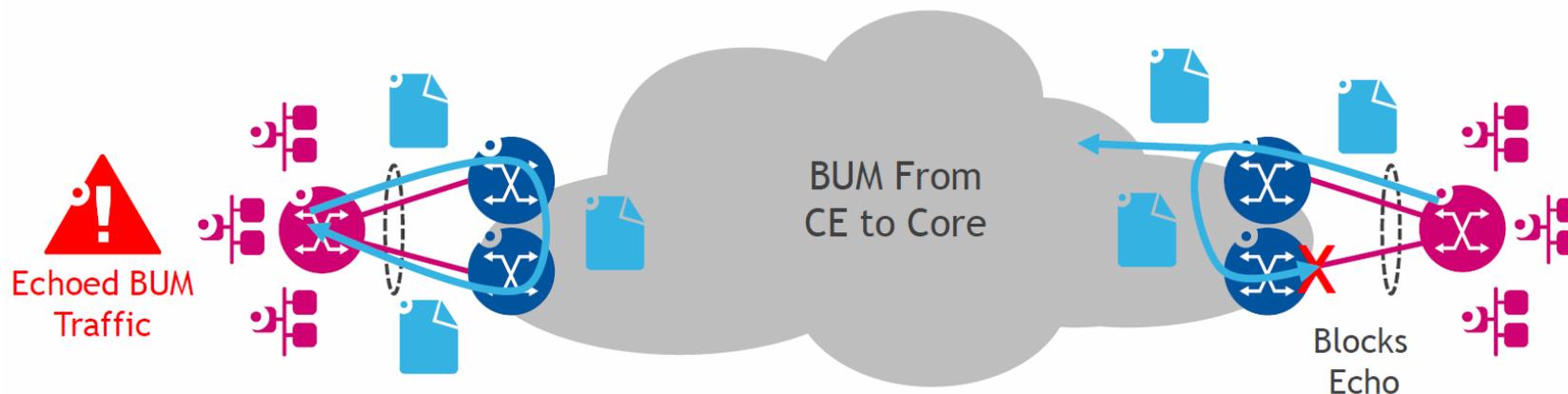


—BUM-3— Trafico BUM marcado con ESI label AD Route

# Funcionamiento de EVPN-VI

## “All-active multihoming” y “split horizon”

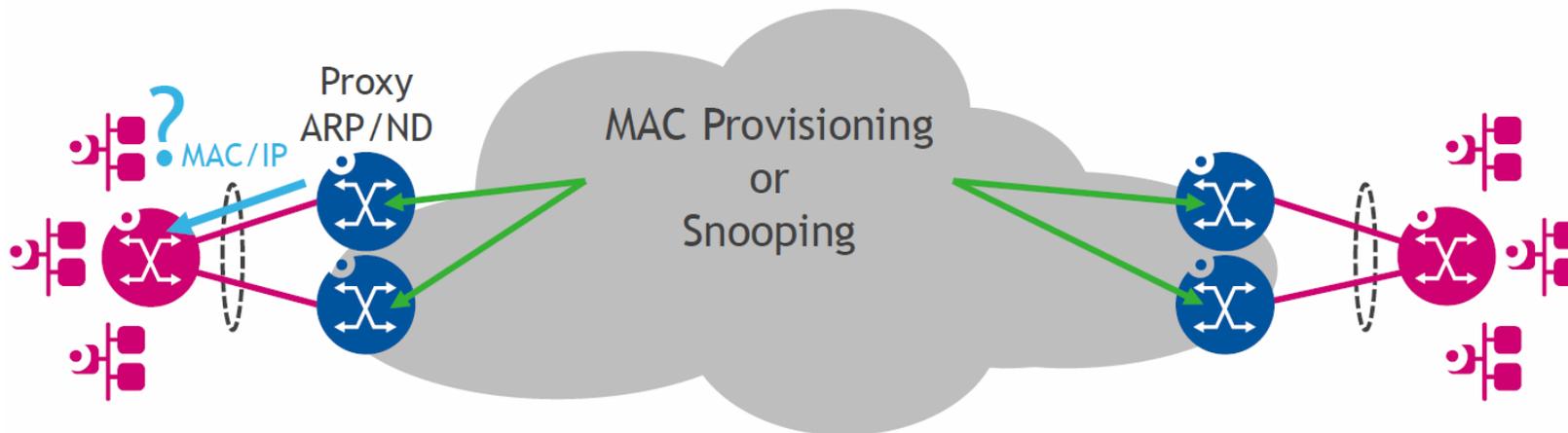
- Los PEs anuncian una etiqueta MPLS (“ESI MPLS Label”) asociada a cada segmento de red con topología multihoming.
- Esa etiqueta “split horizon” se envía asociada a las rutas tipo 1 (“Ethernet A-D”) en forma de communitie BGP.
- Mecanismo anti-bucles “Split horizon” → no reenviar el tráfico por el interfaz a través del que se recibe.
- Cuando un PE envía tráfico BUM hacia el core, replica la trama hacia todos los PEs que participan en la EVI, y añade la etiqueta “ESI MPLS Label” para identificar el ESI donde se originó.
- Los PEs conectados al mismo segmento de red identifican la etiqueta y filtran el tráfico BUM hacia los CEs conectados al ESI.



# Funcionamiento de EVPN-VII

## Gestión de flooding “Unknown Unicast” y ARP/ND

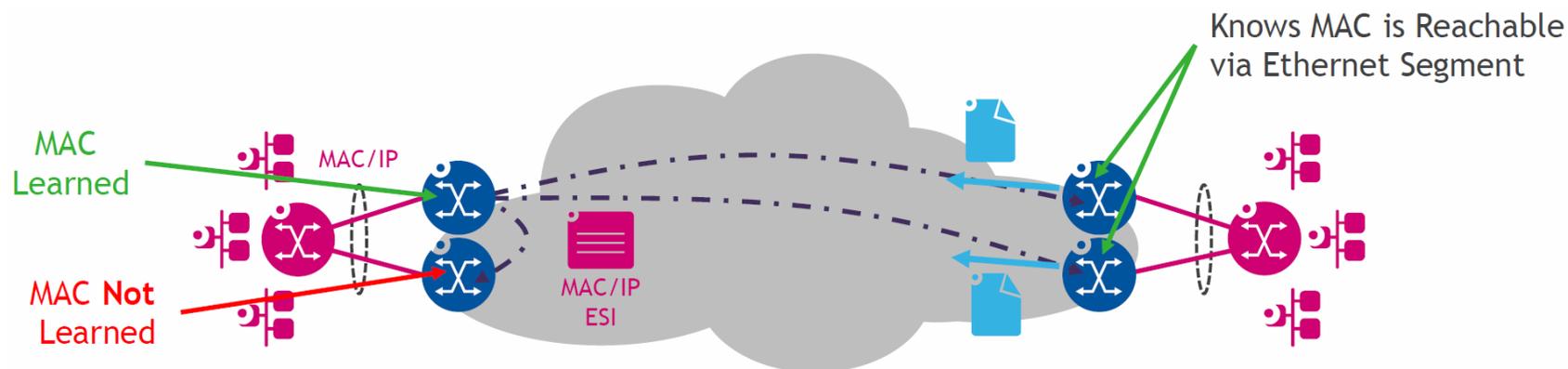
- Los PEs participantes de la EVI realizan las funciones de Proxy-ARP/ND en base a las tablas de rutas EVPN de cada instancia.
- “ARP/ND/DHCP snooping” optimiza y reduce el tráfico de tipo “unknown unicast” entre sites de una EVI.
- Provisionando de forma estática las direcciones MAC de un ESI se puede reducir aún más (o incluso eliminar) el tráfico “unknown unicast”.



# Funcionamiento de EVPN-VIII

## “All-active multihoming” y “aliasing”

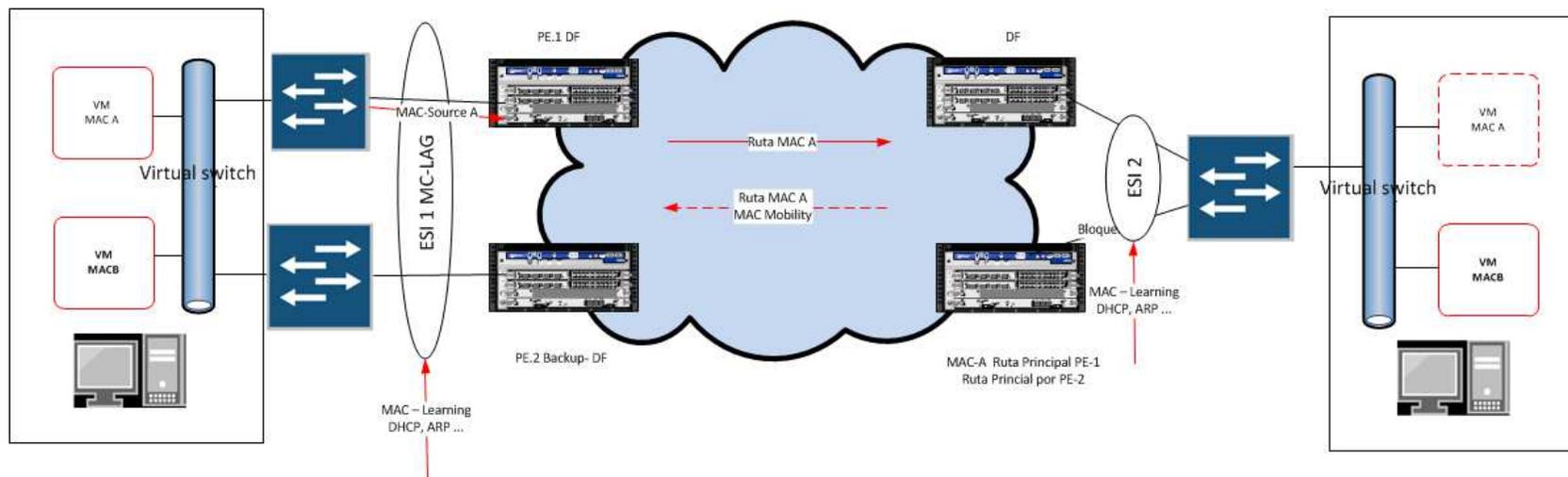
- Las rutas tipo 2 “MAC Advertisement” incluyen información sobre el ESI de la EVI donde reside la MAC de cada host.
- Las rutas tipo 1 “Ethernet A-D” incluyen información que indica los PEs conectados a cada ESI.
- Combinando esta información, los PEs remotos puede identificar las direcciones MAC que se encuentran en cada ESI, independientemente del PE que anunció esas MAC.
- Multihoming modo S-A → permite el uso de paths de backup, mejorando la convergencia de la red ante fallos de link /nodo.
- Multihoming modo A-A → permite el balanceo de carga por flujo hacia servidores conectados a PEs remotos.



# Funcionamiento de EVPN-IX

## MAC mobility / duplicación

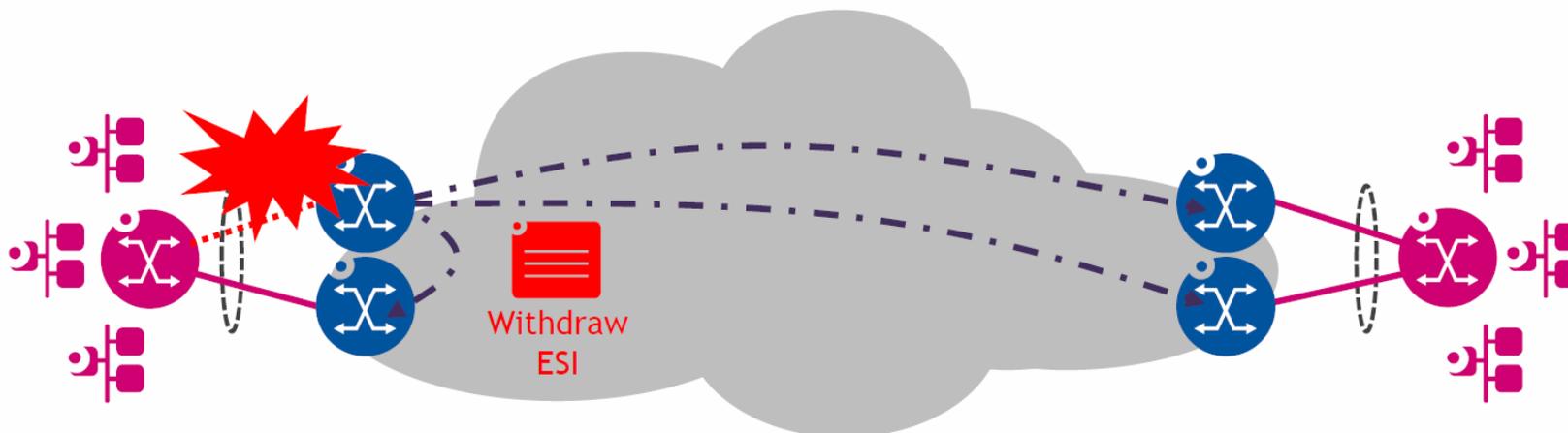
- En entornos virtualizados es bastante frecuente el movimiento de VM entre CPDs.
- Esto implica que una dirección MAC que en un instante dado se anuncia desde un PE/ESI (ruta antigua, incorrecta) de red cambia de ubicación y pasa e anunciarse desde otro PE/ESI de la EVI (ruta nueva, correcta).
- El método usado para identificar la ruta válida de entre las existentes en la red para cada MAC se basa en un contador asociado a la communitie “MAC mobility”, que se añade a cada ruta tipo 2 “MAC Advertisement”.
- La ruta con el mayor número de secuencia es la válida.



# Funcionamiento de EVPN-X

## Optimización de convergencia ante fallos

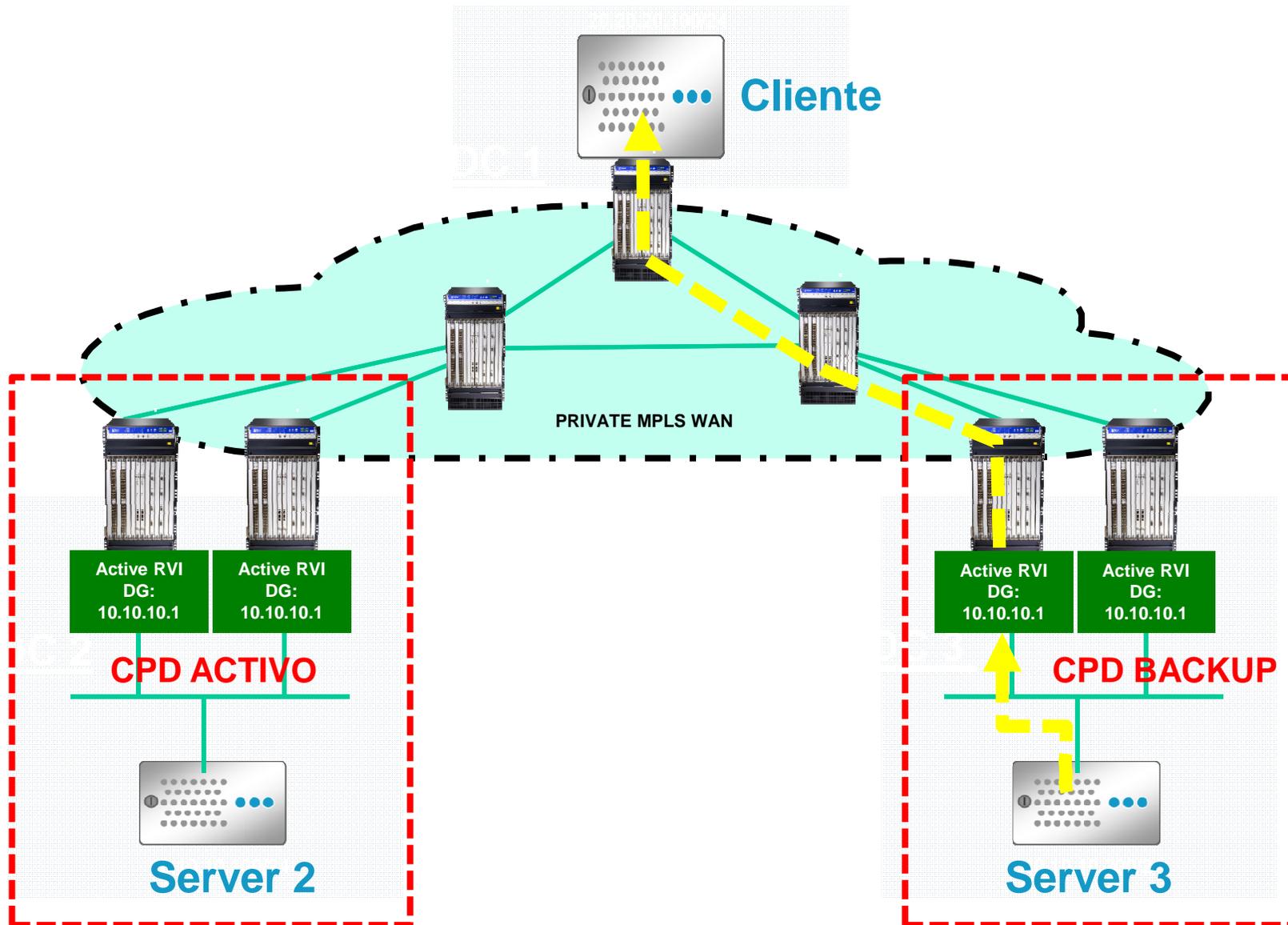
- Los PEs aprenden mediante BGP las MACs que se encuentran en cada ESI (ruta tipo 2), así como los PEs conectados a cada ESI (ruta tipo 1).
- Cuando un PE detecta un fallo en el link que lo conecta con un ESI, todas las MACs que aprende por ese path dejan de ser accesibles para él.
- Para indicarlo a los PEs remotos, el PE local simplemente tiene que eliminar la ruta “Ethernet A-D” para dicho ESI. Los PEs remotos eliminarán de forma masiva todas las rutas MAC/ESI alcanzadas a través del PE que la origina.



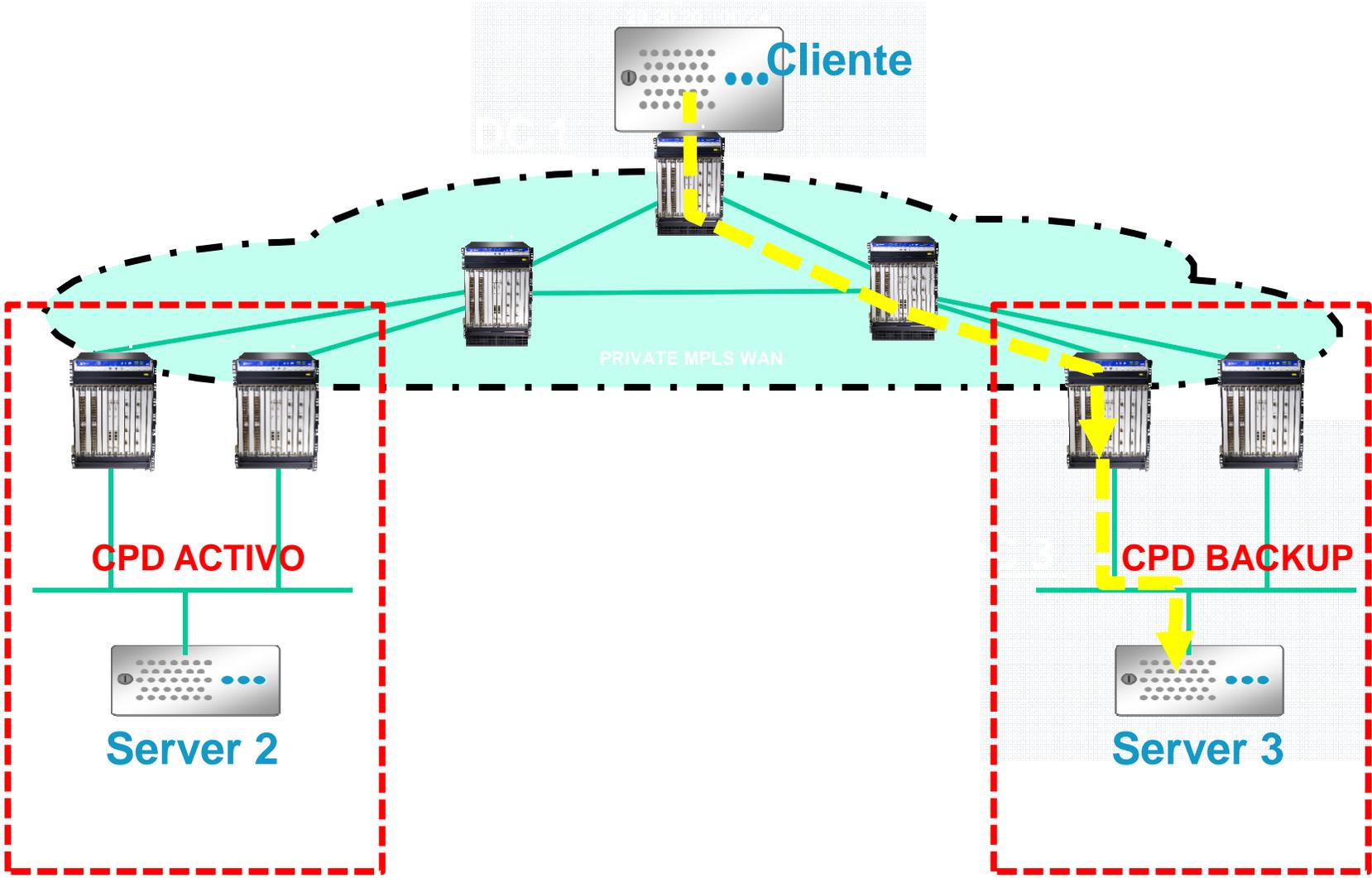
# Integración de L2 y L3: caso DCI (Juniper) - I

- Los interfaces L3 de la EVPN usan todos la misma IP (no se necesita HSRP/VRRP).
- La MAC de los interfaces L3 se anuncia a otros PEs con la community “Default Gateway”.
- El tráfico dirigido hacia esas MACs se enruta a L3 directamente por cualquier PE de la red. De esta manera se elimina el efecto “tromboning” en sentido CPD-clientes.
- La instancia VRF a la que se asocia el interfaz L3 de la EVPN crea direcciones /32 para las MACs presentes en cada CPD.
- De esta manera se puede exportar el rango de subred desde el CPD Principal y las direcciones /32 que en un instante concreto están presentes en el CPD de Backup.
- De esta manera se elimina el efecto “tromboning” en sentido clientes CPD.
- **VMTO: Virtual Machine Traffic Optimisation.**

# Integración de L2 y L3: caso DCI (Juniper) - II



# Integración de L2 y L3: caso DCI (Juniper) - III



Interconexión L2-Ethernet sobre rede; WAN de última generación: E-VPN

# Referencias

- <https://tools.ietf.org/html/draft-ietf-l2vpn-evpn-10>
- <https://tools.ietf.org/html/draft-ietf-l2vpn-pbb-evpn-07>
- <http://tools.ietf.org/html/draft-sd-l2vpn-evpn-overlay-03>
- RFC 7209: Requirements for Ethernet VPN (EVPN)
  
- EVPN & PBB-EVPN whitepaper (Cisco)  
[http://www.cisco.com/c/en/us/products/collateral/routers/asr-9000-series-aggregation-services-routers/whitepaper\\_c11-731864.html](http://www.cisco.com/c/en/us/products/collateral/routers/asr-9000-series-aggregation-services-routers/whitepaper_c11-731864.html)
  
- EVPN whitepaper (Juniper)  
<http://www.juniper.net/assets/us/en/local/pdf/whitepapers/2000553-en.pdf>
  
- EVPN overview (Alcatel-Lucent)  
<http://www2.alcatel-lucent.com/techzine/es/redes-vpn-ethernet-evpn-para-servicios-integrados-de-niveles-2-3/>



# indra

Alberto Aldave – [aaldave@indra.es](mailto:aaldave@indra.es)

José A. Díaz Linares – [jadiazl@indra.es](mailto:jadiazl@indra.es)

INDRA REDES  
C/ Anabel Segura, 7 – Planta Baja  
28108 - Alcobendas (Madrid) ,  
Madrid España

[www.indra.es](http://www.indra.es)

