

BGP EVPN: A view from the Field



Alex Mitev
SE, Cisco

3 September, 2026

L2VPN Intro

L2VPN Today: A brief History

- Virtual Private L2 connection over IP/MPLS enabled networks
 - VPWS / Martini (RFC 4447) – the original wire
 - VPLS for multipoint
 - RFC 4762 – manual + T-LDP
 - RFC 6074 – BGP auto-discovery + T-LDP (or L2TPv3)
 - RFC 4761 – BGP for both (label blocks)
 - Full mesh → H-VPLS (tames mesh, but concentrates MACs on hub PE) → PBB-VPLS (hides C-MACs behind B-MACs)
- Discovery got solved, learning and redundancy did NOT
 - Every option – active/standby
 - Flood-and-learn
 - L2-only

Why EVPN

REQUIREMENTS

- All active multihoming (per-flow LB, sub-second failover)
- CP MAC learning (mobility, no unknown-unicast flooding)
- Auto-discovery and unified control-plane across all services
- Integrated L2/L3 + DCI at scale
- MAC scale – hide host MACs from the core

Why EVPN

REQUIREMENTS

- All active multihoming (per-flow LB, sub-second failover)
- CP MAC learning (mobility, no unknown-unicast flooding)
- Auto-discovery and unified control-plane across all services
- Integrated L2/L3 + DCI at scale
- MAC scale – hide host MACs from the core

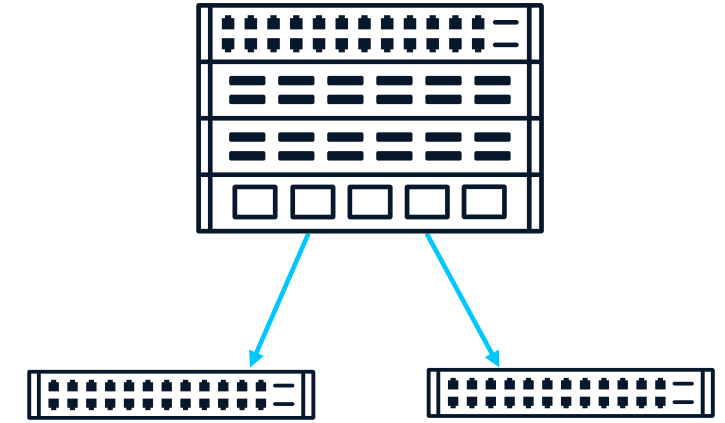
EVPN

- A/A via ESI + aliasing + mass withdrawal (RFC7432)
- MP-BGP AFI 25/SAFI 70, Type-2 routes/sequence-number
- RT-based Auto-discovery; (RFC7432/RFC8214)
- EVPN-IRB (RFC9135); rides MPLS/SRv6/VXLAN
- PBB-EVPN, control-plane B-MAC distribution (RFC7623)

Why EVPN...

Fewer Protocols, Independent Hardware

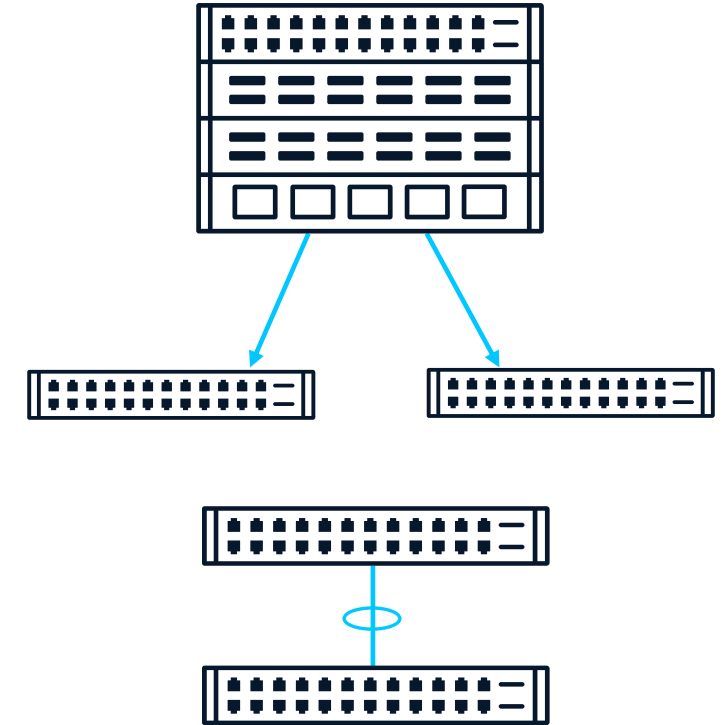
- Opportunity to erase the old model
 - Stop engineering around single points of failure – remove them
 - No more hitless-upgrade gamble -> drain, reboot, repeat
 - Independent boxes – move coordination into BGP



Why EVPN...

Fewer Protocols, Independent Hardware

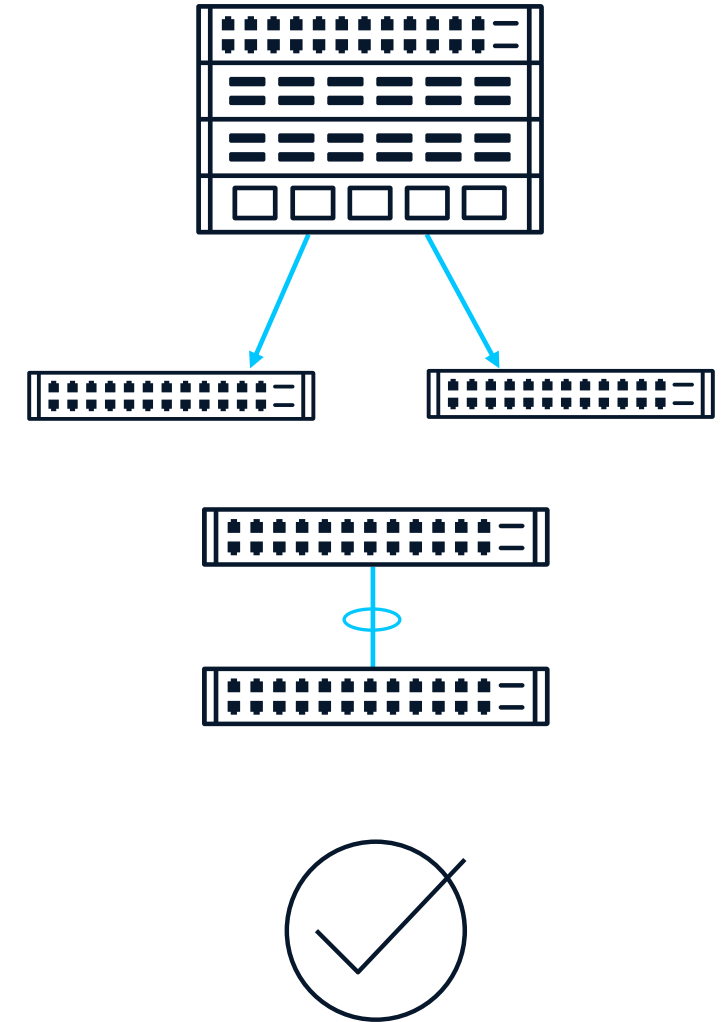
- Opportunity to erase the old model
 - Stop engineering around single points of failure – remove them
 - No more hitless-upgrade gamble -> drain, reboot, repeat
 - Independent boxes – move coordination into BGP
- Delete the peer-link and the sync protocol
 - No more split brain
 - N-way, multi-vendor multihoming



Why EVPN...

Fewer Protocols, Independent Hardware

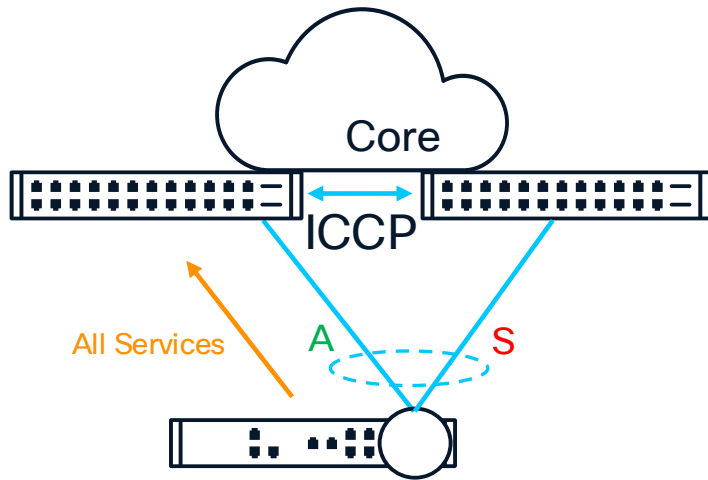
- Opportunity to erase the old model
 - Stop engineering around single points of failure – remove them
 - No more hitless-upgrade gamble -> drain, reboot, repeat
 - Independent boxes – move coordination into BGP
- Delete the peer-link and the sync protocol
 - No more split brain
 - N-way, multi-vendor multihoming
- Risk distributes – complexity changes shape
 - More failure events, smaller blast radius – each one survivable
 - Distributed control plane vs opaque proprietary in-box mechanism
 - Only as good as the multihoming design



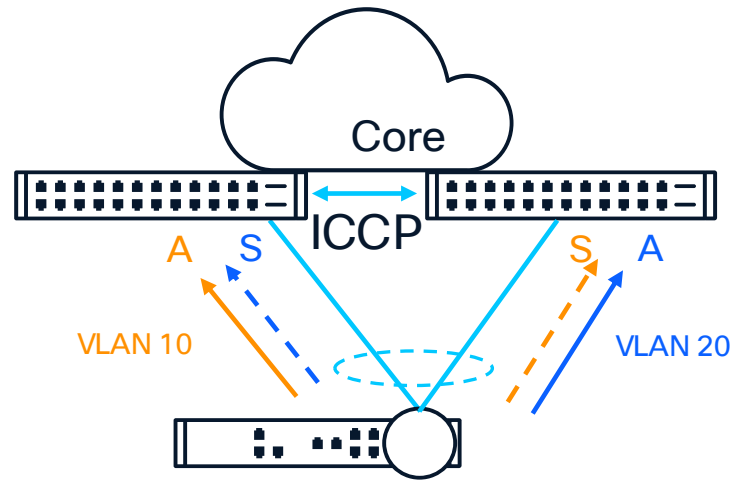
EVPN and Legacy MH Services

Legacy Designs

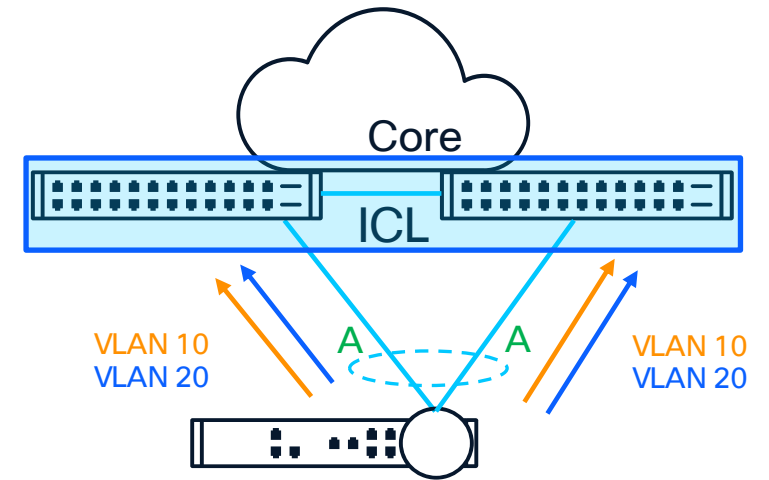
- MC-LAG
- ICCP-SM
- nV Edge



MC-LAG



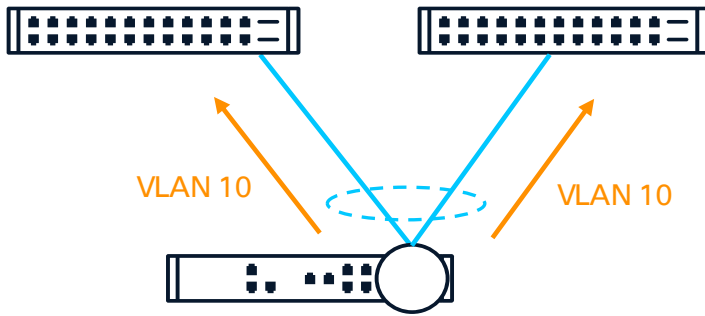
Pseudo MC-LAG
ICCP-SM



nV Edge

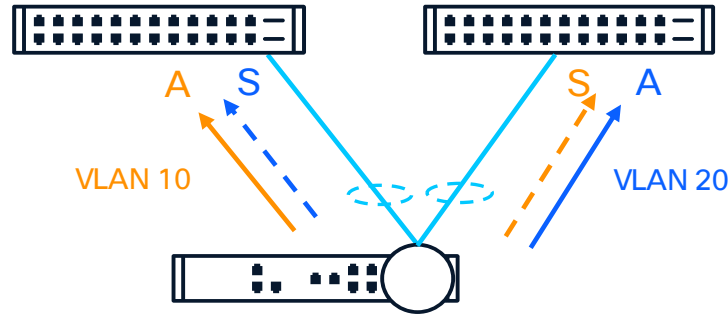
EVPN Design Option

All-Active
(per flow)



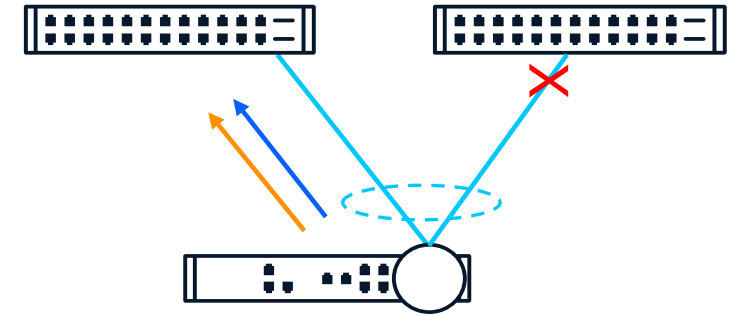
- Single LAG at CE
- VLAN goes to both PE
- Traffic hashed per-flow
- NDF blocks BUM

Single-Active
(per VLAN)



- Multiple LAG at CE
- VLAN active on one PE
- Traffic hashed per-VLAN
- NDF blocks all traffic in both directions per VLAN

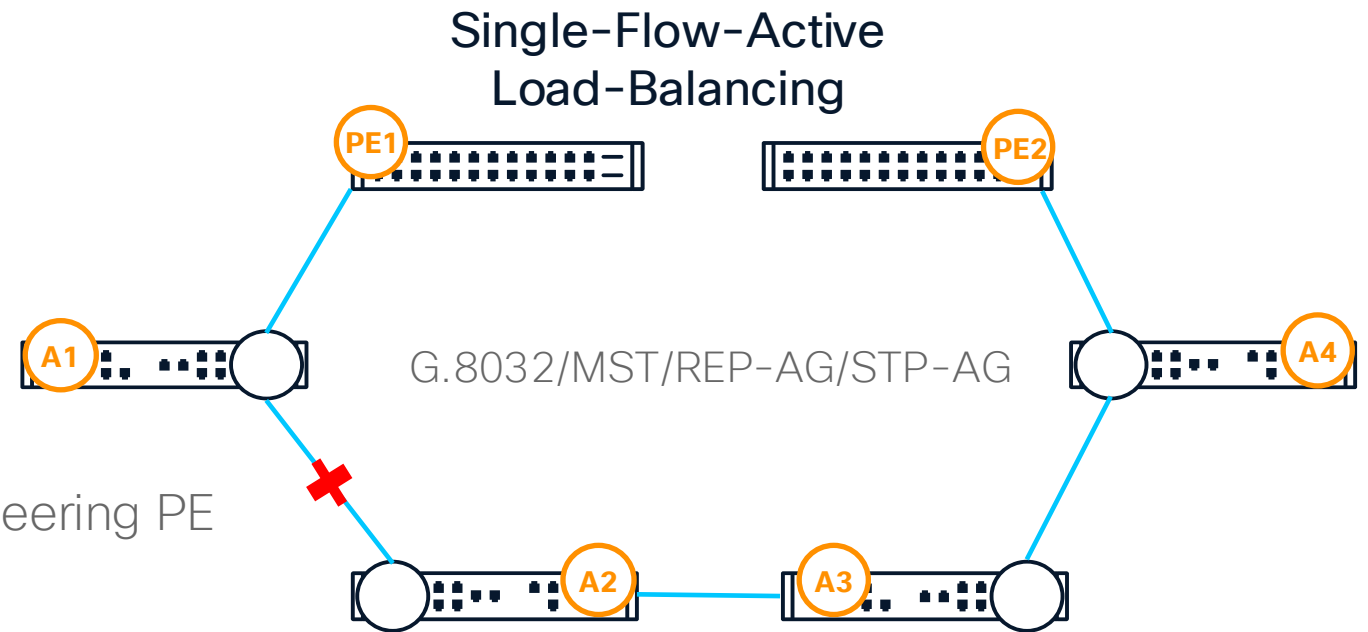
Port-Active
(per port)



- Single LAG at CE
- Port active on one PE
- Traffic hashed per-port
- Port on NDF is blocked

EVPN Design Option

- Flows may arrive at PE1 or PE2: no DF Election
- Use case is L2-access protocols
 - STP-AG, REP-AG ... who break loops
 - Protocol must break any L2 Loop
- Mobility within the ESI – predictable
 - Speculation (mitigate ARP probing)
 - MAC and MAC-IP synchronisation
 - Pre-program ADJ but FIB Nexthop is peering PE
 - PE1: A1 ADJ
 - PE2: A1 via PE1, A3 pre-programmed ADJ



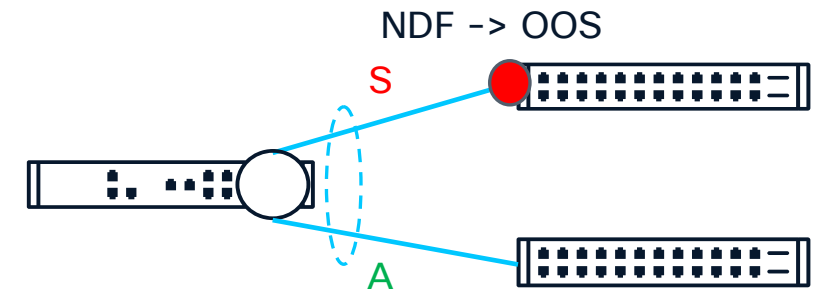
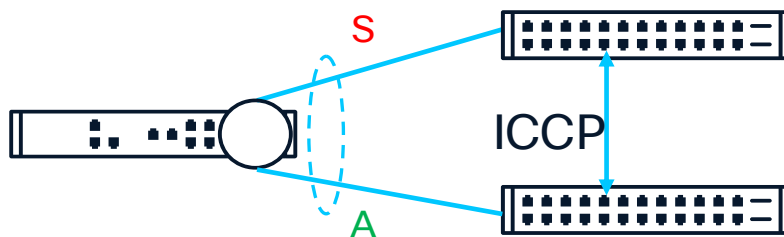
EVPN Migration.

Or what we see on the field

MC-LAG to EVPN Port-Active

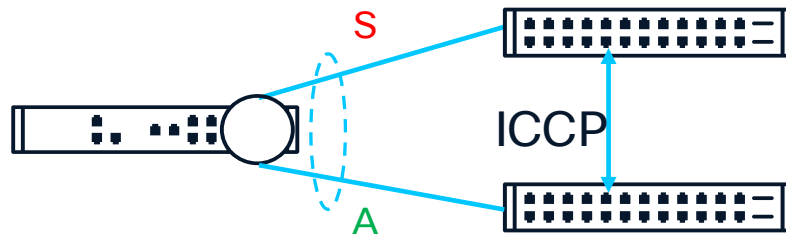
- No config change on the CE
- Active/Backup is preserved
- QoS is accurate and preserved
- Standby PE sets the LACP state of OOS
- Support L2 and L3 services

```
evpn
  evi 123
    advertise-mac
    !
    !
  interface Bundle-Ether123
    ethernet-segment
      identifier type 0 12.31.23.12.31.23.12.31.01
      load-balancing-mode port-active
    !
  !
!
l2vpn
  bridge group VPLS
  bridge-domain 123
  interface Bundle-Ether123.123
  !
  vfi 123
    neighbor 10.0.0.1 pw-id 123
    !
    neighbor 10.0.0.3 pw-id 123
    !
    neighbor 10.0.0.4 pw-id 123
    !
  evi 123
  !
```

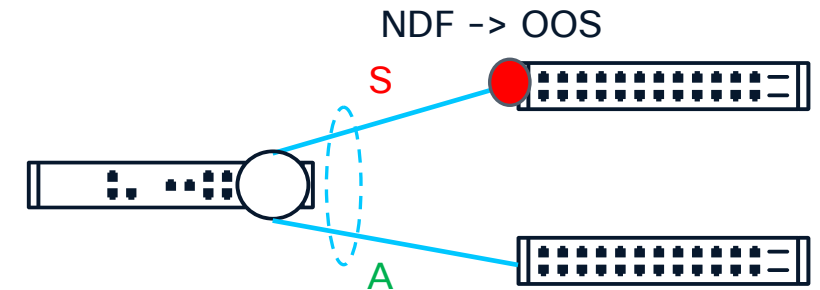


MC-LAG to EVPN Port-Active

- After that, we **CLEAN**



```
redundancy
no iccp*
!
interface Bundle-Ether123
no mlacp iccp-group 123
no mlacp port-priority 1
lacp system mac 1234.1234.1234
lacp system priority 1
```

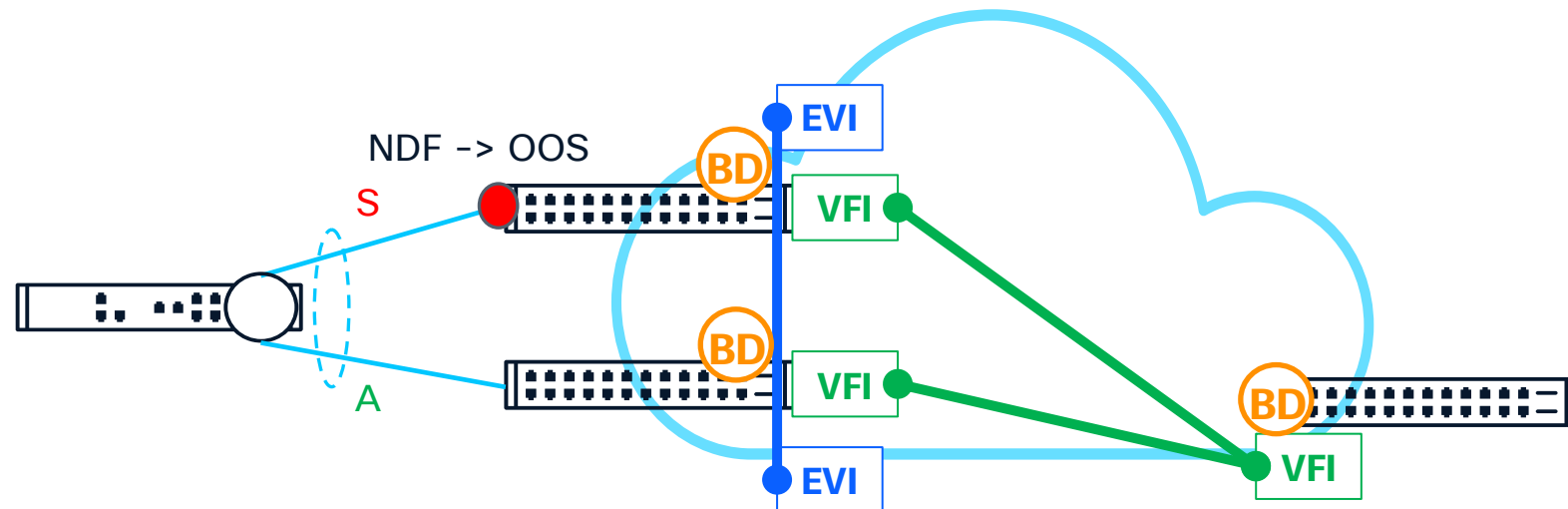


* Make sure no other services are using ICCP

VPLS to EVPN

- Just need to configure the EVI on the new PE

```
!2vpn
  bridge group VPLS
  bridge-domain 123
  interface Bundle-Ether34.123
  !
  vfi 113
    neighbor 10.0.0.1 pw-id 123
    !
    neighbor 10.0.0.3 pw-id 123
    !
    neighbor 10.0.0.4 pw-id 123
    !
  evi 123
  !
```



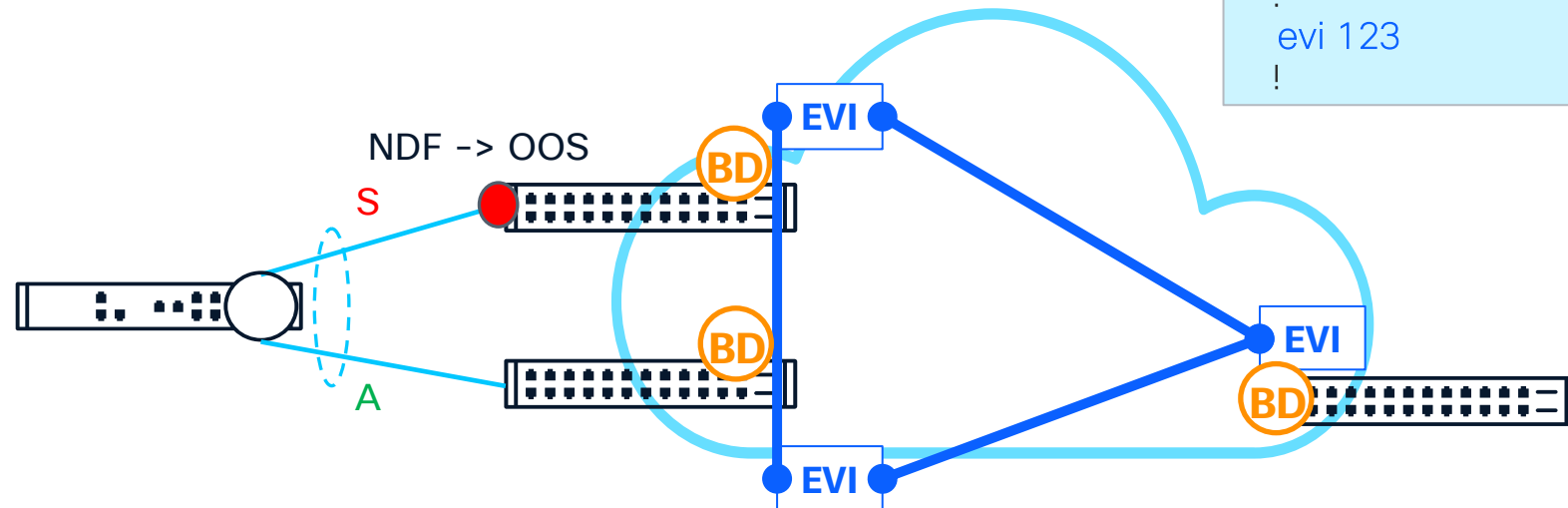
VPLS to EVPN

- After all PEs discover each other
 - Pseudowires are shut down
 - EVPN replaces VPLS

```

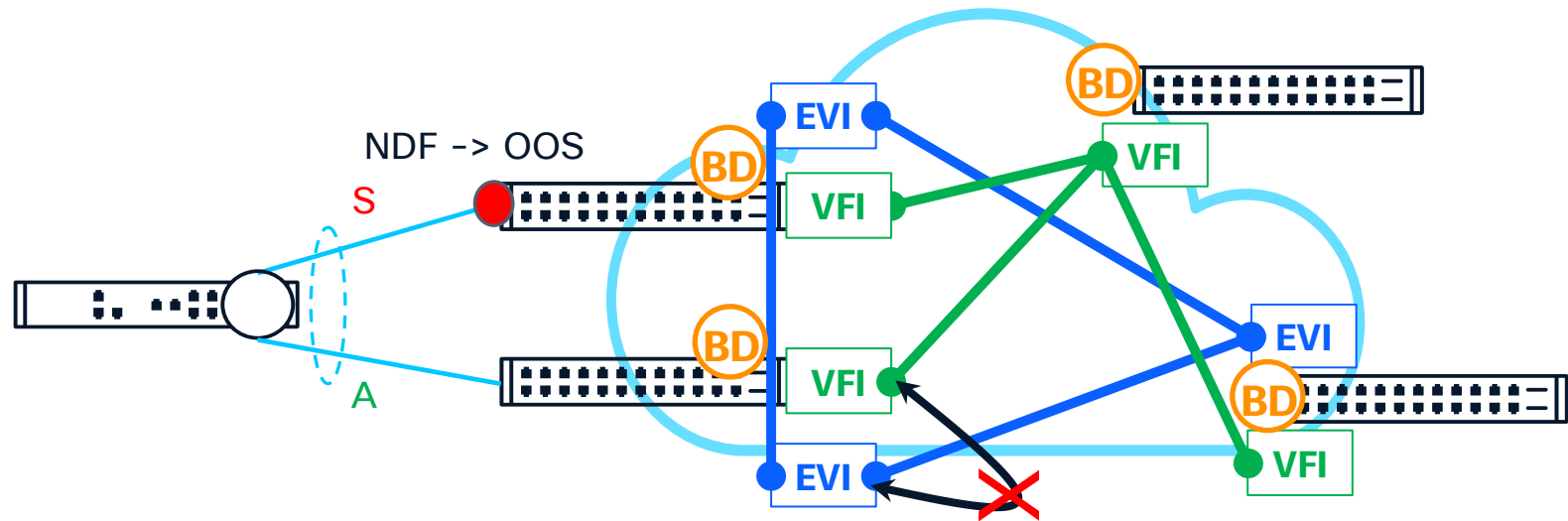
evpn
  evi 123
    advertise-mac
    !
    !
    !
  !
  l2vpn
    bridge group VPLS
    bridge-domain 123
    interface Bundle-Ether123.123
    !
    vfi 113
      neighbor 10.0.0.1 pw-id 123
      !
      neighbor 10.0.0.3 pw-id 123
      !
      neighbor 10.0.0.4 pw-id 123
      !
    evi 123
    !
  !

```



VPLS to EVPN

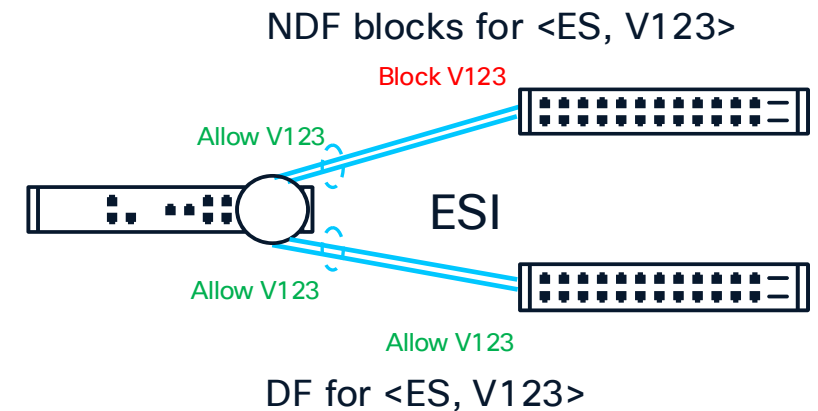
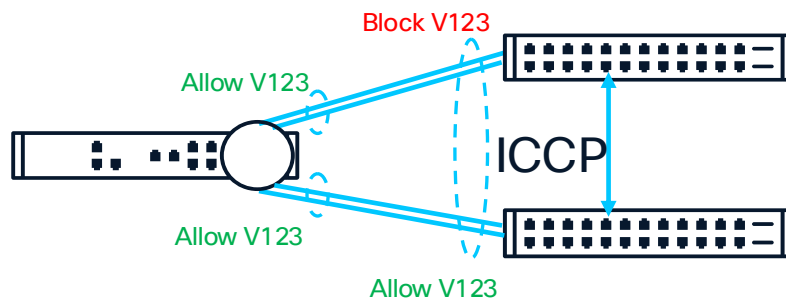
- VFI by default is in Split Horizon Group 1
 - Split Horizon Group 1 protects from loops in MPLS core
- EVI is also by default in Split Horizon Group 1
 - No PE will forward data between EVI and VFI



ICCP-SM to EVPN Single Active

- No config change on the CE
- Single/Active and per-VLAN LB is preserved
- QoS is accurate and preserved
- Support L2 and L3 services
- NDF blocks per <ESI, VLAN> – interface stays UP, CE cannot see who is DF
- MAC flush toward CE on DF change: STP TCN or MVRP

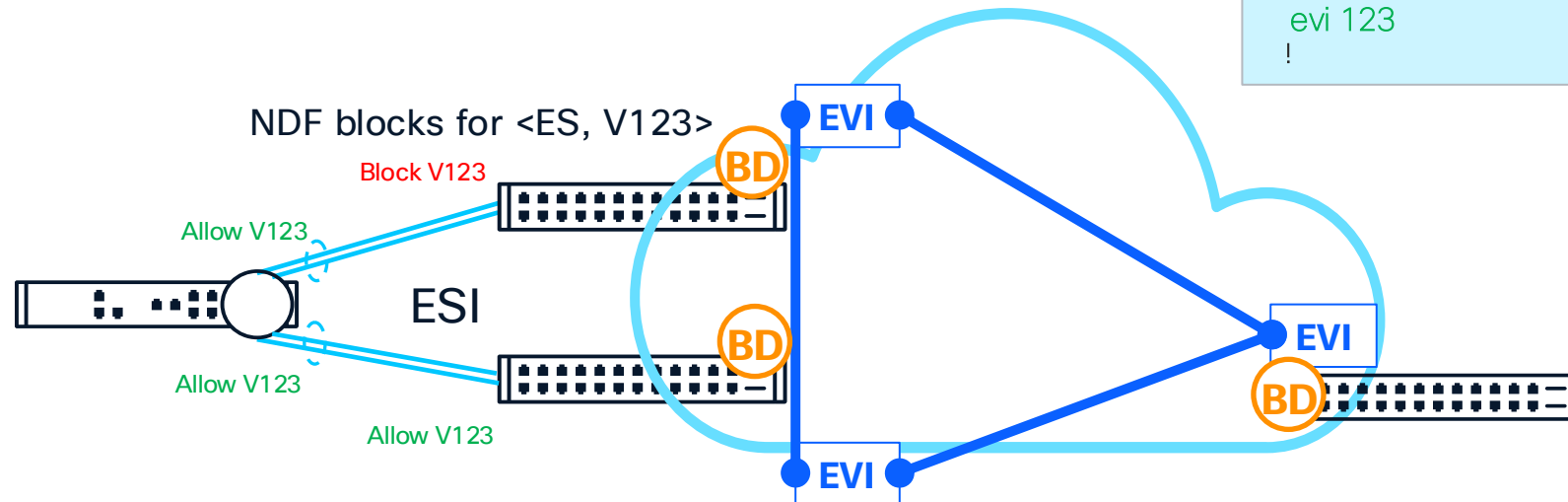
```
evpn
...
!
interface Bundle-Ether123
  ethernet-segment
  identifier type 0 12.31.23.12.31.23.12.31.01
  load-balancing-mode single-active
!
```



ICCP-SM to EVPN Single Active

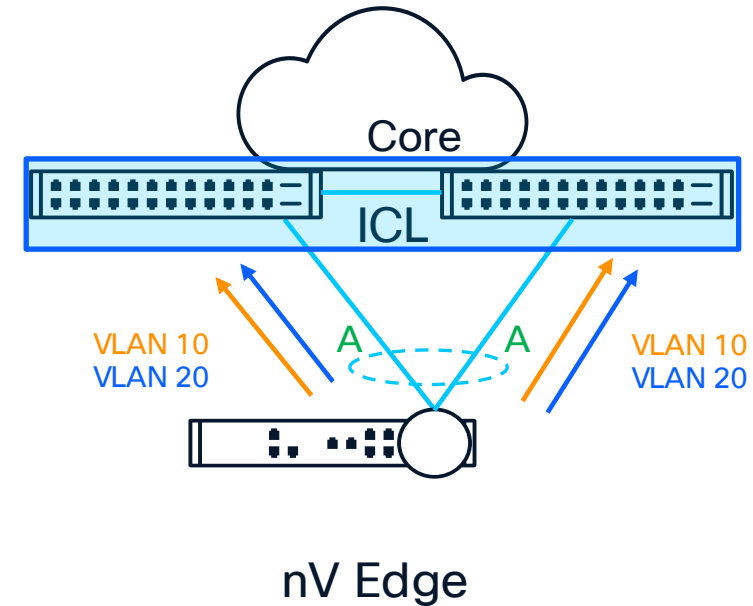
- Nothing changes in case of VPLS to EVPN
- Once all PE discover each other through EVPN
 - VPLS PWs are shut down

```
evpn
  evi 123
    advertise-mac
    !
    !
    interface Bundle-Ether123
      ethernet-segment
        identifier type 0 12.31.23.12.31.23.12.31.01
        load-balancing-mode single-active
    !
  !
  l2vpn
    bridge group VPLS
    bridge-domain 123
    interface Bundle-Ether123.123
    !
    vfi 123
      neighbor 10.0.0.1 pw-id 123
    !
      neighbor 10.0.0.3 pw-id 123
    !
      neighbor 10.0.0.4 pw-id 123
    !
    evi 123
  !
!
```



nV Edge to EVPN

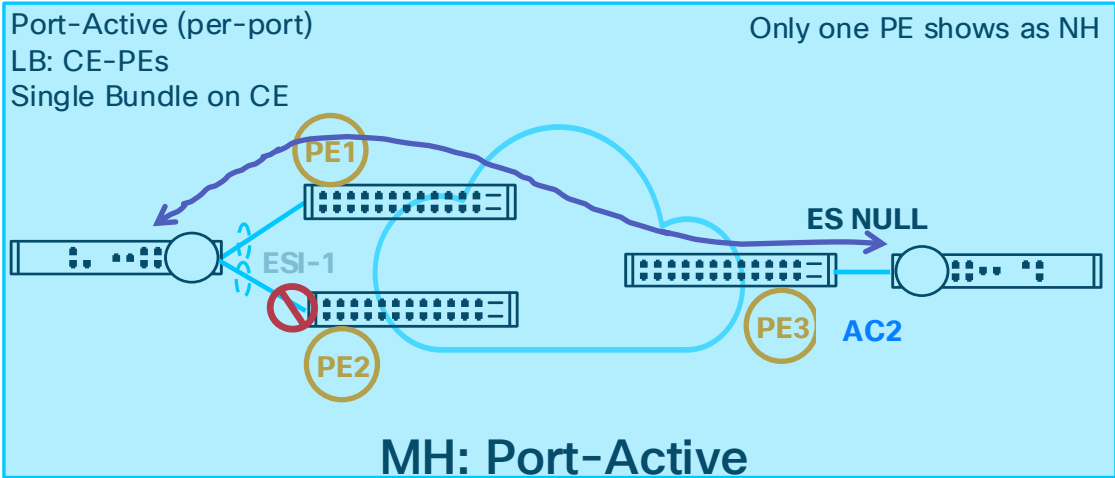
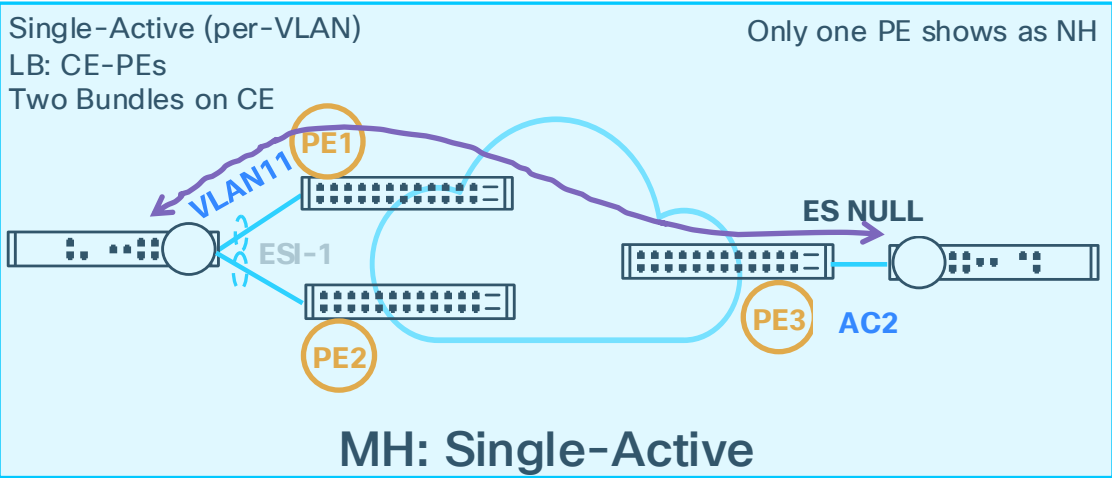
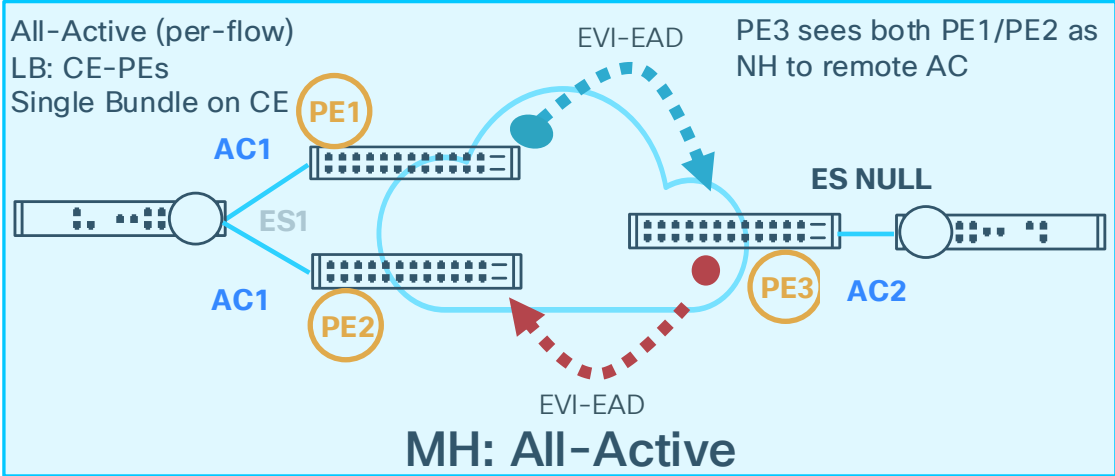
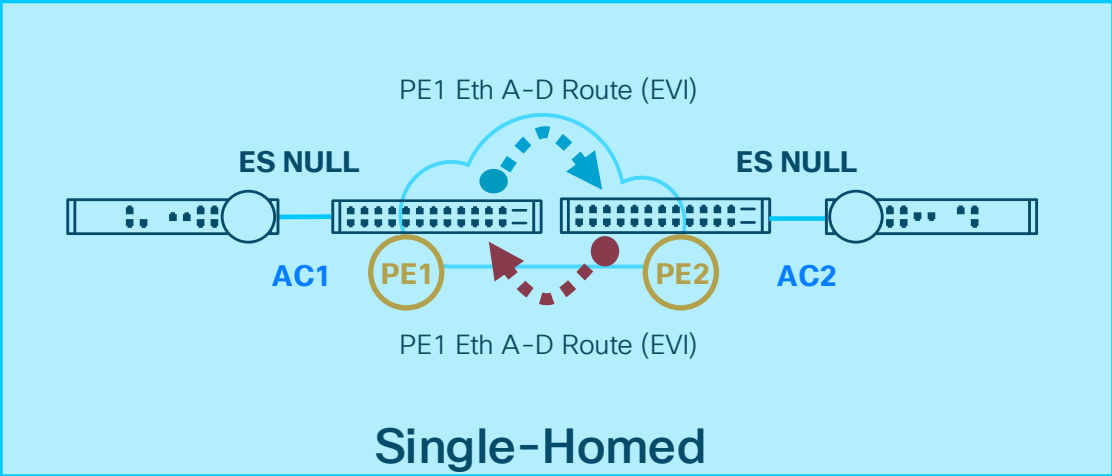
- nV Edge is the one migration that can't be done in service
- Two options:
 - EVPN All-Active
 - No config change on the CE
 - Keep the all-active design and per-flow LB
 - vES feature is required for VPLS
 - EVPN Port-Active
 - No config change on the CE
 - QoS policy doesn't change
- nV Edge and similar clustering solutions will require de-clustering*



* Downtime required

EVPN VPWS

EVPN VPWS Design Option



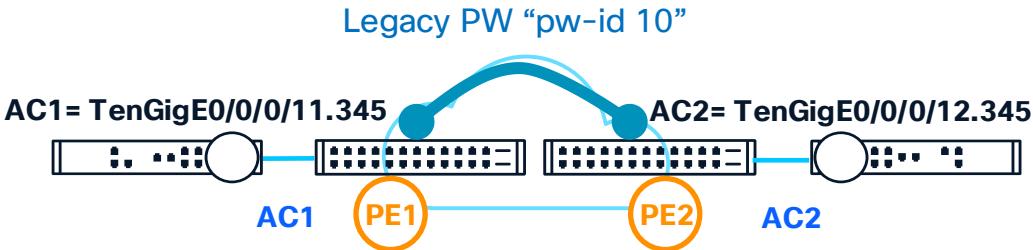
Seamless Migration to EVPN VPWS

PE1

```
l2vpn
xconnect group P2P
p2p LEGACY
interface TenGigE0/0/0/11.345
neighbor ipv4 10.0.0.2 pw-id 10
```

PE2

```
l2vpn
xconnect group P2P
p2p LEGACY
interface TenGigE0/0/0/12.345
neighbor ipv4 10.0.0.1 pw-id 10
```



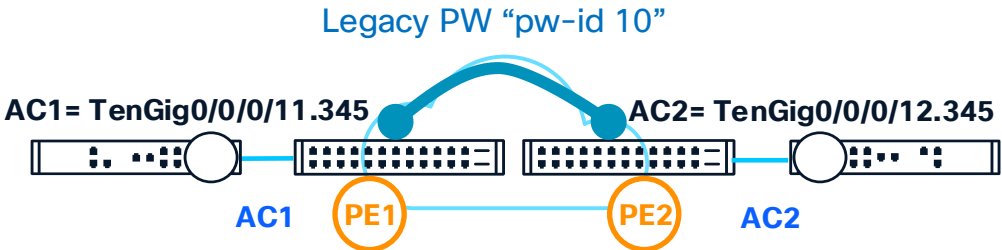
RP/0/RP0/CPU0:PE1-am-lab#show l2vpn xconnect

XConnect		Segment 1		Segment 2			
Group	Name	ST	Description	ST	Description	ST	
P2P	LEGACY	UP	Te0/0/0/11.345	UP	10.0.0.1	10	UP

RP/0/RP0/CPU0:PE1-am-lab#

Seamless Migration to EVPN VPWS

```
l2vpn
xconnect group P2P
p2p LEGACY
vpws-seamless-integration
interface TenGigE0/0/0/11.345
neighbor ipv4 10.0.0.2 pw-id 10
!
p2p EVPN_VPWS
interface TenGigE0/0/0/11.345
neighbor evpn evi 10 service 10
```



```
RP/0/RP0/CPU0:PE1-am-lab# sho l2vpn xconnect
```

XConnect Group	Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
P2P	EVPN_VPWS	DN	Te0/0/0/11.345	UP	EVPN 10,10,10.0.0.1	DN
P2P	LEGACY	UP	Te0/0/0/11.345	UP	10.0.0.1 10	UP

```
RP/0/RP0/CPU0:PE1-am-lab# #show bgp l2vpn evpn rd 10.0.0.1:10
Network Next Hop Metric LocPrf Weight Path
Route Distinguisher: 10.0.0.1:10 (default for vrf VPWS:10)
*> [1][0000.0000.0000.0000.0000][10]/120 0.0.0.0 0 i
```

Seamless Migration to EVPN VPWS

I2vpn

PE1

xconnect group P2P

p2p LEGACY

vpws-seamless-integration

interface TenGigE0/0/0/11.345

neighbor ipv4 10.0.0.2 pw-id 10

!

p2p EVPN_VPWS

interface TenGigE0/0/0/11.345

neighbor evpn evi 10 service 10

I2vpn

PE2

xconnect group P2P

p2p LEGACY

vpws-seamless-integration

interface TenGigE0/0/0/12.345

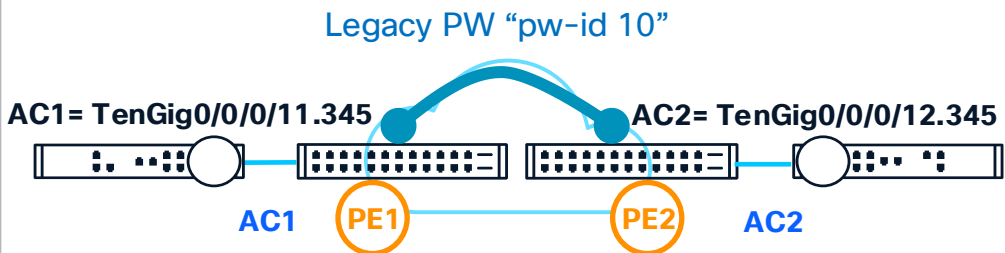
neighbor ipv4 10.0.0.1 pw-id 10

!

p2p EVPN_VPWS

interface TenGigE0/0/0/12.345

neighbor evpn evi 10 service 10



RP/0/RP0/CPU0:PE1-am-lab# sh I2vpn xconnect

XConnect Group	Name	ST	Segment 1 Description	ST	Segment 2 Description	ST
P2P	EVPN_VPWS	UP	Te0/0/0/11.345	UP	EVPN 10,10,10.0.0.2	UP
P2P	LEGACY	DN	Te0/0/0/11.345	SB(SI)	10.0.0.1 10	UP

RP/0/RP0/CPU0:PE1-am-lab# # show bgp I2vpn evpn rd 10.0.0.1:10

Network

Next Hop

Metric LocPrf Weight Path

Route Distinguisher: 10.0.0.1:10 (default for vrf VPWS:10)

*> [1I0000.0000.0000.0000.0000I10]/120

0.0.0.0

0 i

* i

10.0.0.2

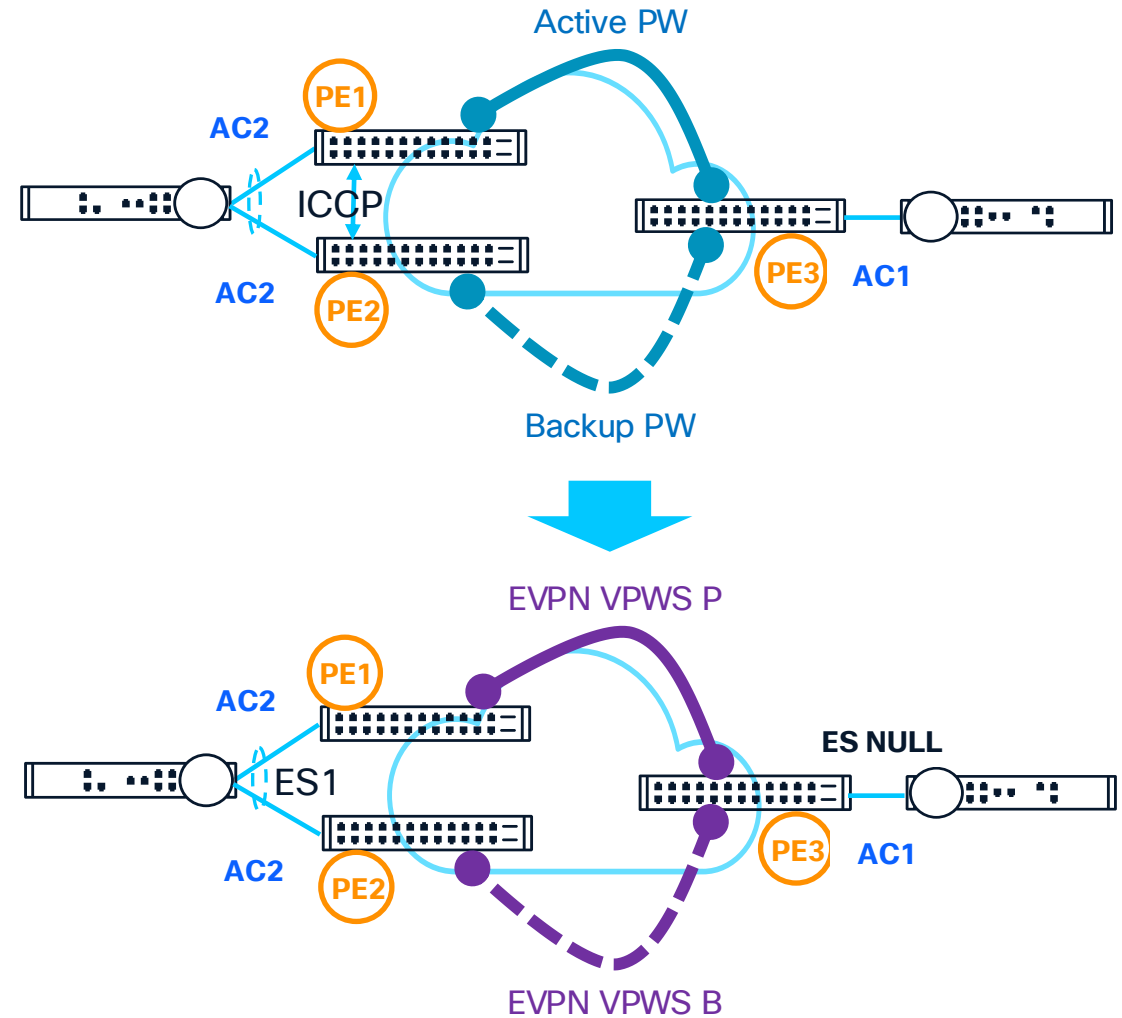
100

0 i 0 i

When both Legacy PW and EVPN VPWS exist for the same AC, EVPN VPWS gets preference over the Legacy PW

MH Migration to EVPN VPWS (MC-LAG)

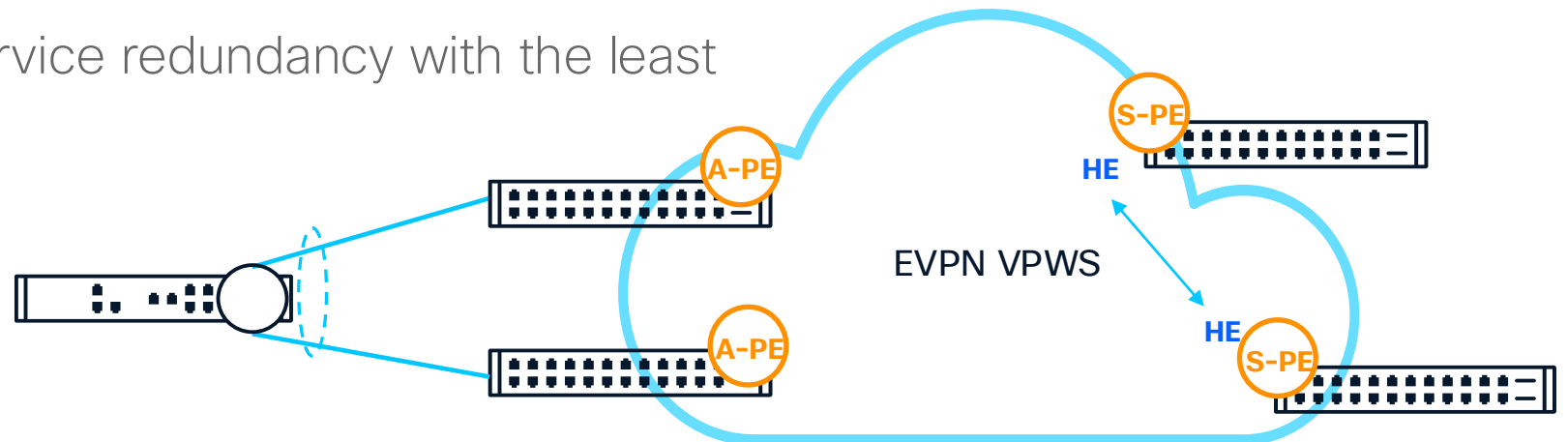
1. PE3 – Seamless Migration configuration
2. PE2
 1. Bundle shutdown
 2. EVPN VPWS PA Configuration
 3. Remove mLACP/ICCP config
3. PE1
 1. EVPN VPWS PA Configuration
 2. Remove mLACP/ICCP config
4. PE2: Bring up Bundle Interface



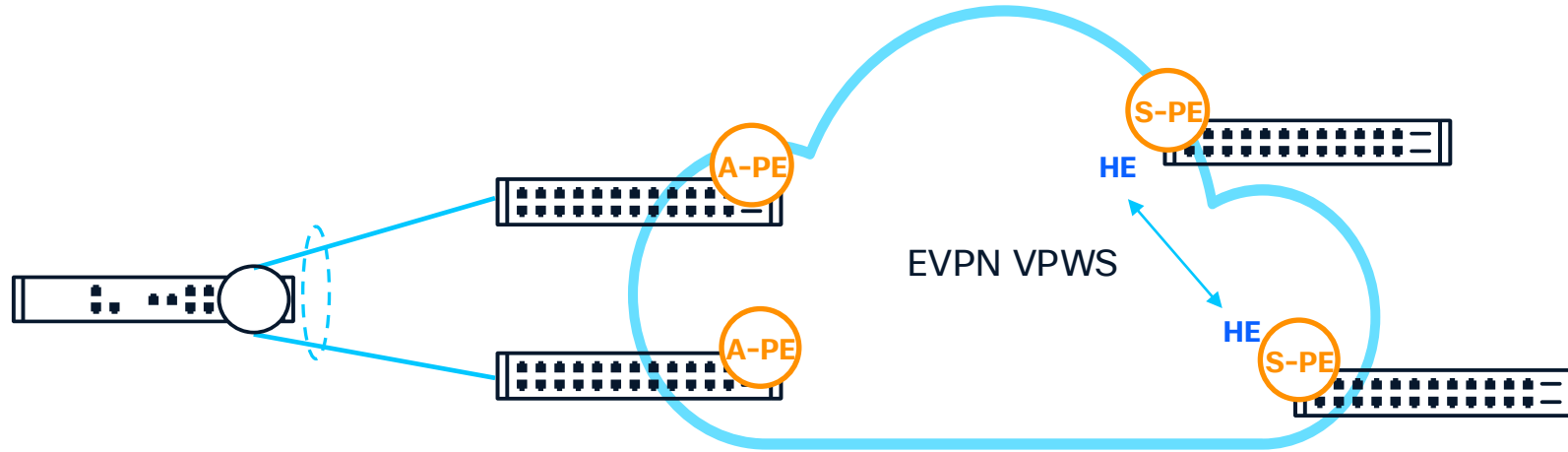
Other EVPN Services

EVPN Service/Feature Requests

- Access/agg/pre-agg network to be purely Layer 2. No L3 service on those devices. Instead, the service PE will be used to offer L3 services (DHCP, DNS, L2/L3 ACL, accurate QoS, routing)
- Efficiently backhaul the internal L3 services to L3 devices while still able to monitor and provide assurance on per-service granularity. Services must be resilient L3 services. Availability and Failover are the only requirements.
- Need a scalable L3 GW/service redundancy with the least states as possible



EVPN HE



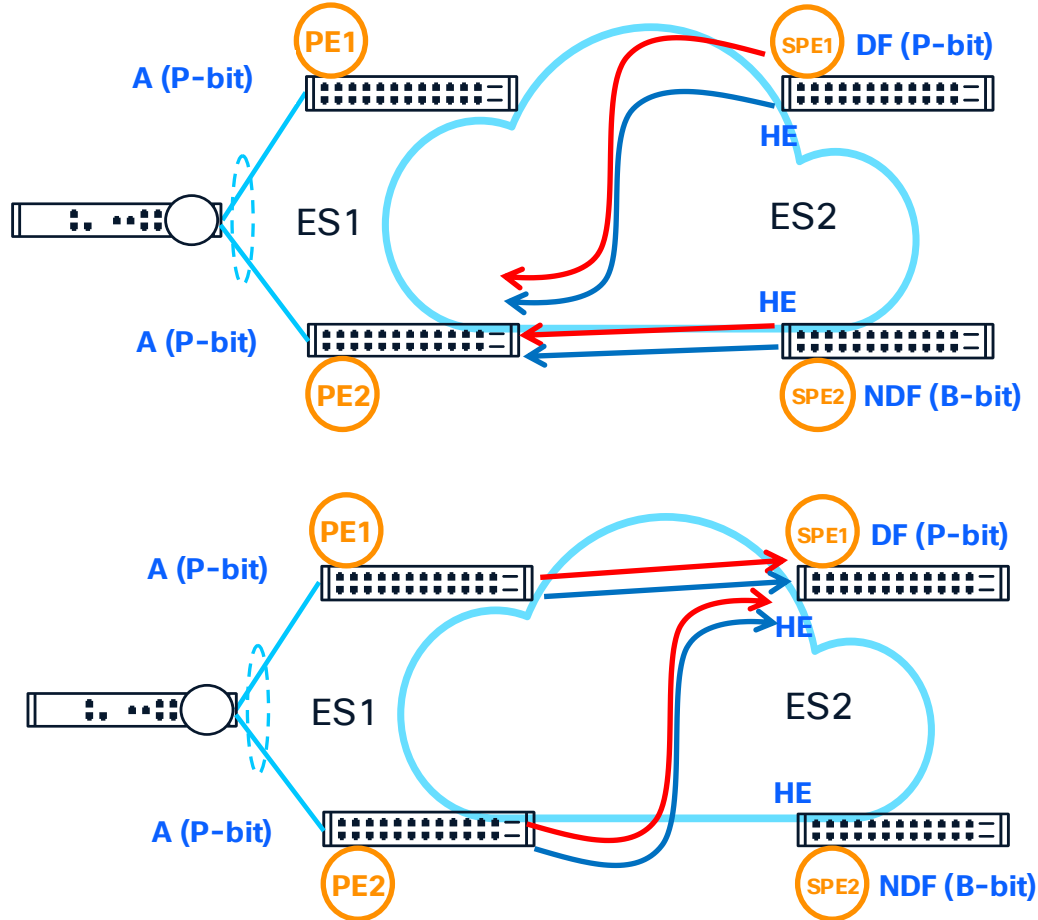
Access Modes A-PE

- All-Active EVPN-VPWS
- Port-Active EVPN-VPWS
- Single-Active EVPN-VPWS

Service Modes S-PE

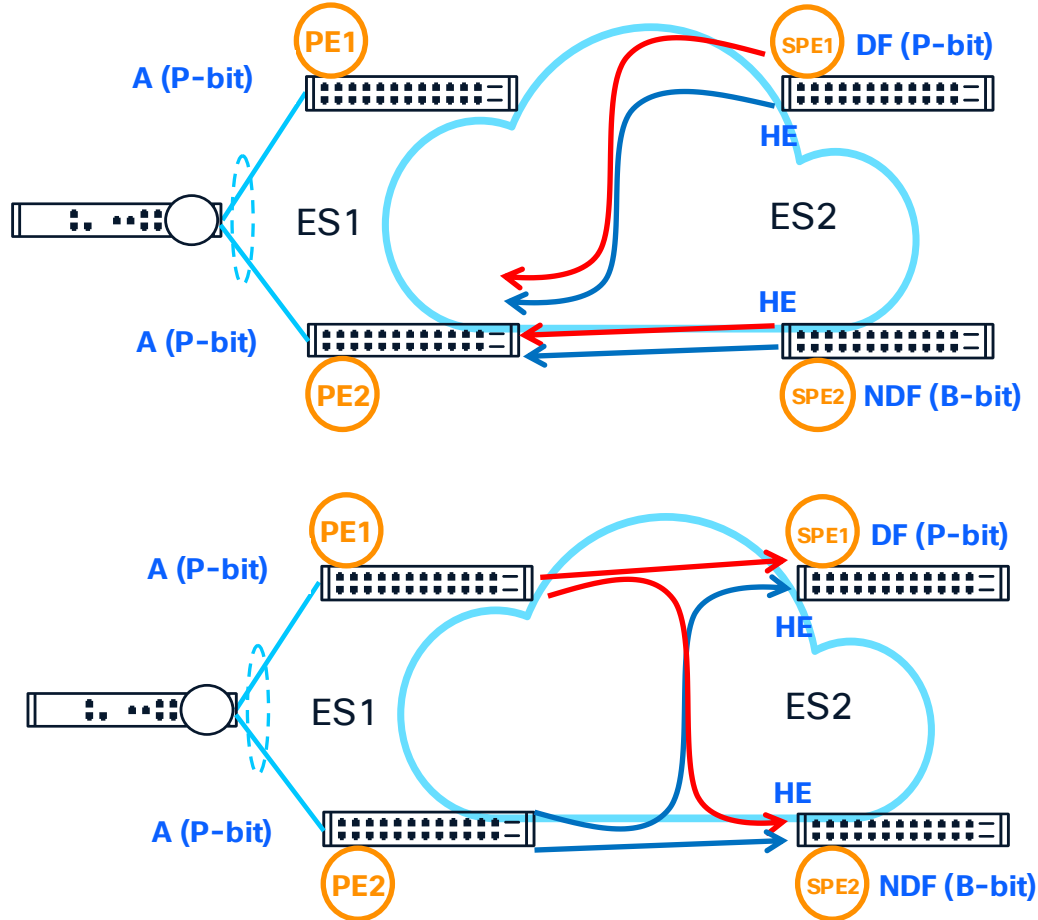
- All-Active EVPN-VPWS
- Port-Active EVPN-VPWS
- Single-Active EVPN-VPWS

EVPN HE - SA



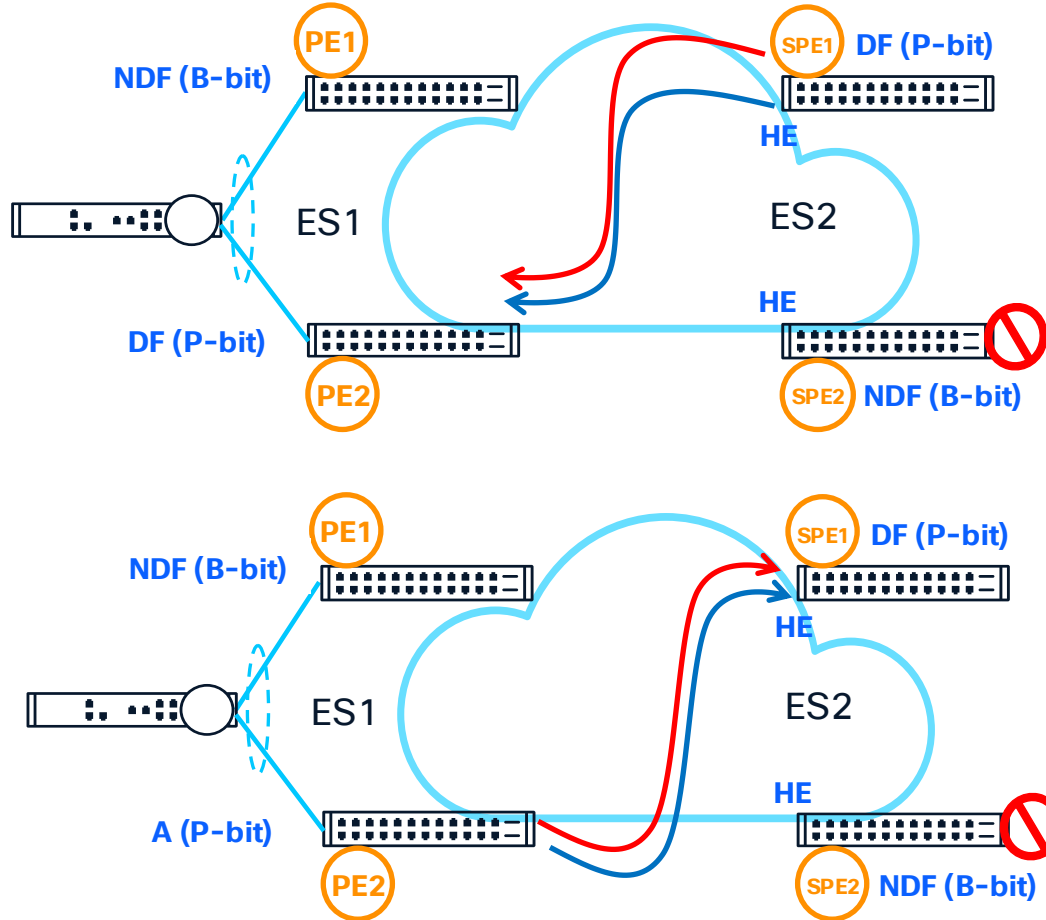
1. S-PE (Core) to Access
 - NDF and DF both bring PW-Ether and main intf. **UP**
 - NDF sends B-bit, DF sends P-bit to A-Pes
 - Both S-PE forward all VLANs to the same PE – DF A-PE
2. Core facing side is **UP/UP** and data-plane is **Forward/Forward**
3. Only HE DF node attracts traffic from Access and forwards to Core

EVPN HE – A/A



1. S-PE (Core) to Access
 - NDF and DF both bring PW-Ether and main intf. **UP**
 - NDF sends B-bit, DF sends P-bit to A-Pes
 - Both S-PE forward all VLANs to the same PE – DF A-PE
 - In case A/A preferred next-hop CLI
2. Core facing side is **UP/UP** and data-plane is **Forward/Froward**
3. Access sends LB all traffic to both S-PE

EVPN HE – P/A

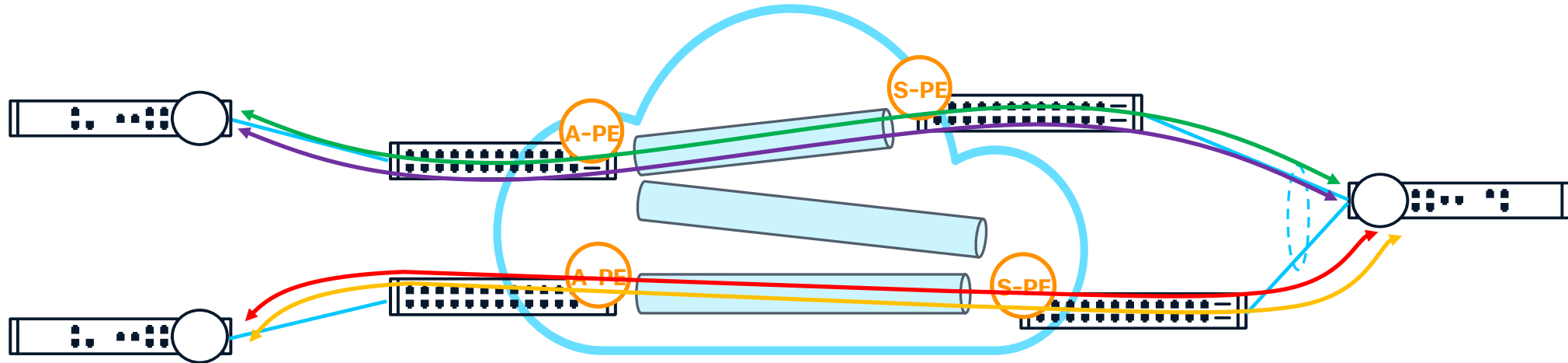


1. S-PE (Core) to Access
 - NDF node bring PW-Ether and main intf. **DOWN and** sends B-bit
 - DF sends P-bit to A-PEs
 - DF S-PE forward all VLANs to the same PE – DF A-PE
2. On DF S-PE intf. are **UP** and data-plane is **Forward**
3. Only DF S-PE forward to Core

Other notable EVPN use cases

Flexible Cross Connects

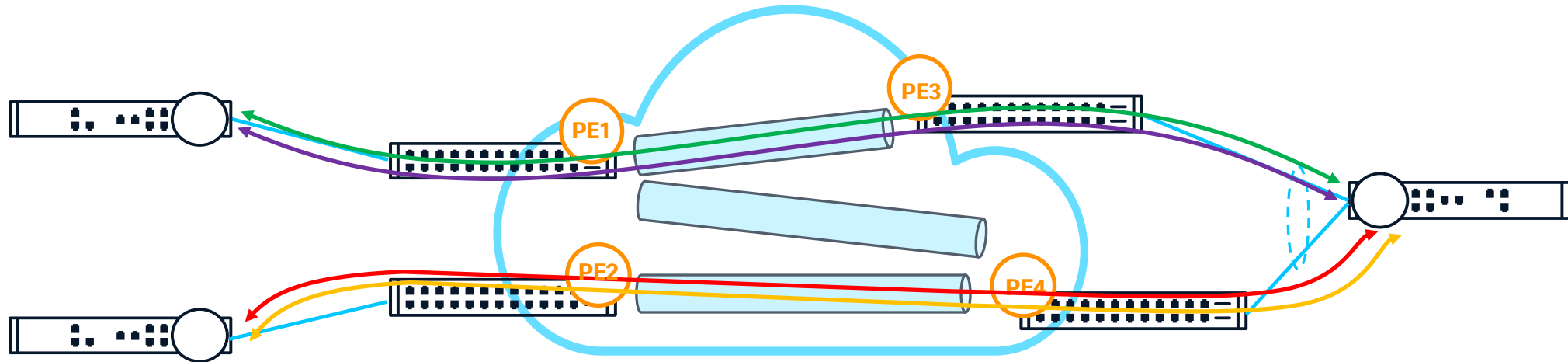
- Bundle 1000s of services into the least service tunnels in order to:
 - Reduce number of services labels associated with the tunnels and this associated OAM monitoring
 - Reduce BGP signalling



Other notable EVPN use cases

Flexible Cross Connects

- Single-Homed CE
 - Reduce both, numbers of PWs and BGP state
- Multi-homed CEs – VLAN Aware
 - Reduce the number of PWs but not BGP states
- Multi-homed CEs – VLAN unaware
 - Reduced the number of BGP states but we signal more PWs



Notable Use Cases

Australia

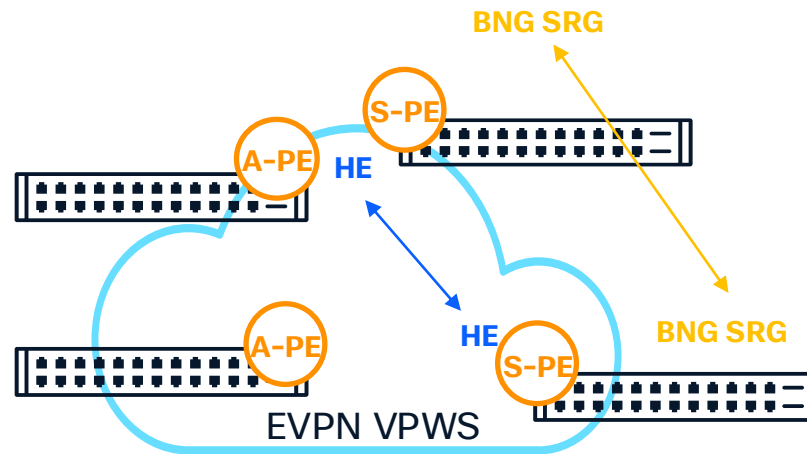
BNG and EVPN

Putting it all together

- Several Australian RSP, connected to the national backbone are building next gen service delivery for BNG application over EVPN
- Failover/Protection per S-vlan – **VLAN-aware**
- Other L3 services must be supported – L3VPN – **HE**
- Integration with Subscriber Redundancy – **?**
- Scalable delivery of 10,000s of services with access failure detection – **FXC**

```
l2vpn
router-id 9.9.9.5
!
flexible-xconnect-service vlan-aware evi 1000
interface Bundle-Ether100.1000
interface Bundle-Ether100.1001
!
```

```
interface Bundle-Ether100.1000 l2transport
description --- CVC_1000 ---
encapsulation dot1q 1000 second-dot1q any
mtu 1600
!
interface Bundle-Ether100.1001 l2transport
description --- CVC_1001 ---
encapsulation dot1q 1001 second-dot1q any
mtu 1600
!
interface Bundle-Ether100
ethernet-segment
identifier type 0 12.12.12.12.12.12.12.12
!
```

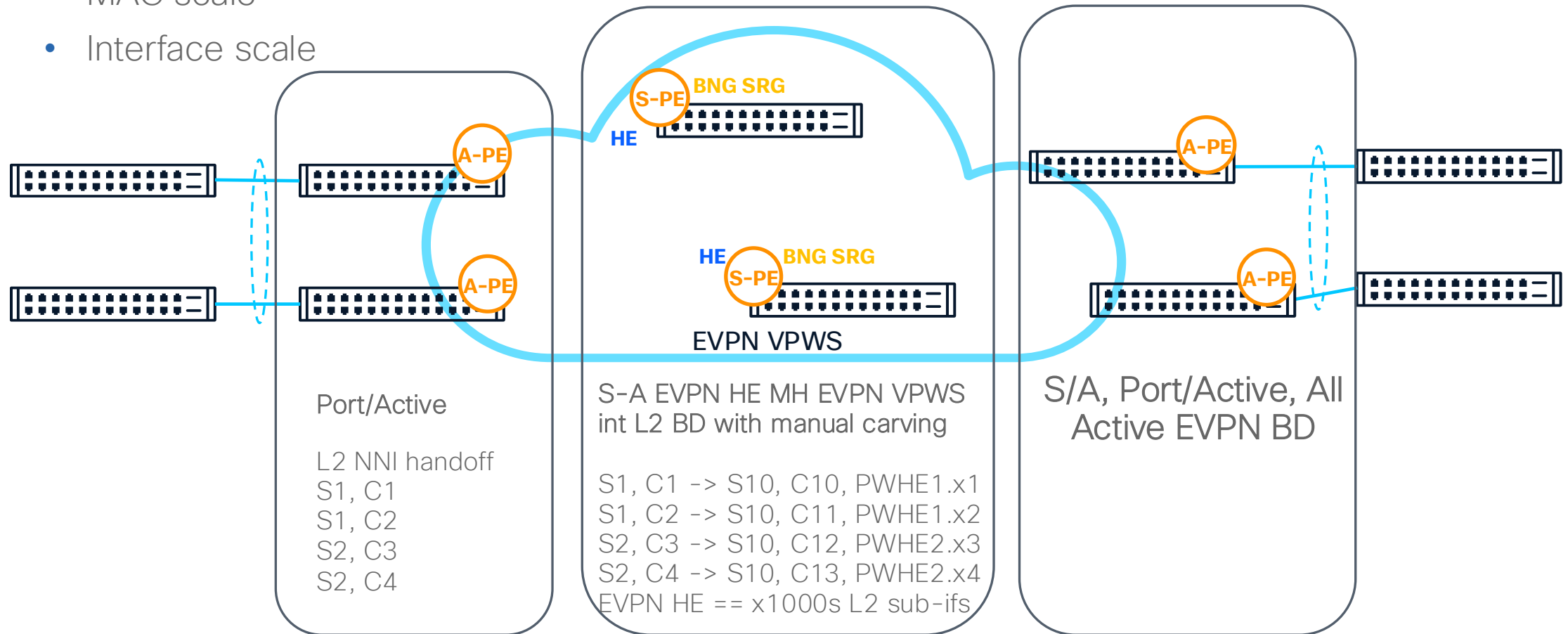


```
evpn
ethernet-segment
type 1 auto-generation-disable
!
evi 1000
preferred-nexthop lowest-ip
!
interface PW-Ether100
ethernet-segment
identifier type 0 10.12.12.12.12.12.12.10
service-carving preference-based
srg-driven
!
```

BNG, EVPN and Wholesale

Putting it all together

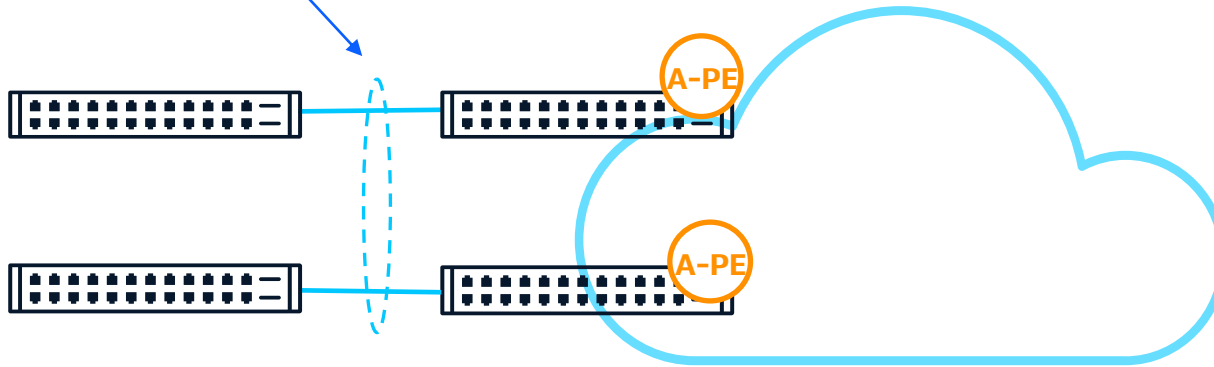
- Wholesale problem as a fan-in problem
 - MAC scale
 - Interface scale



BNG and EVPN

Access Driven DF Election

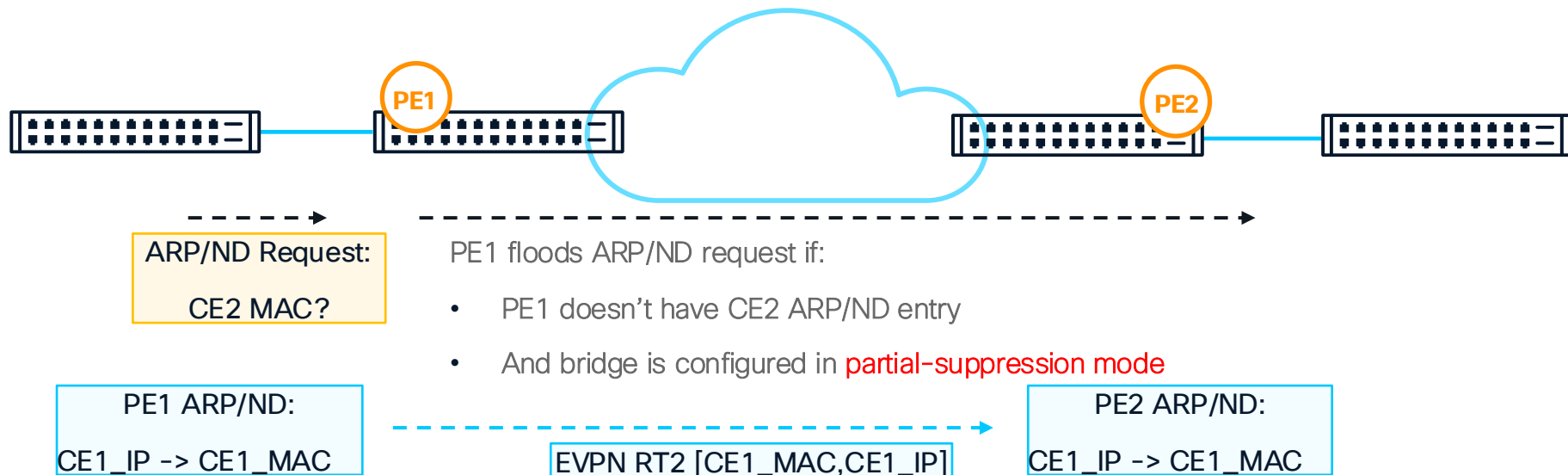
Single-homed Service



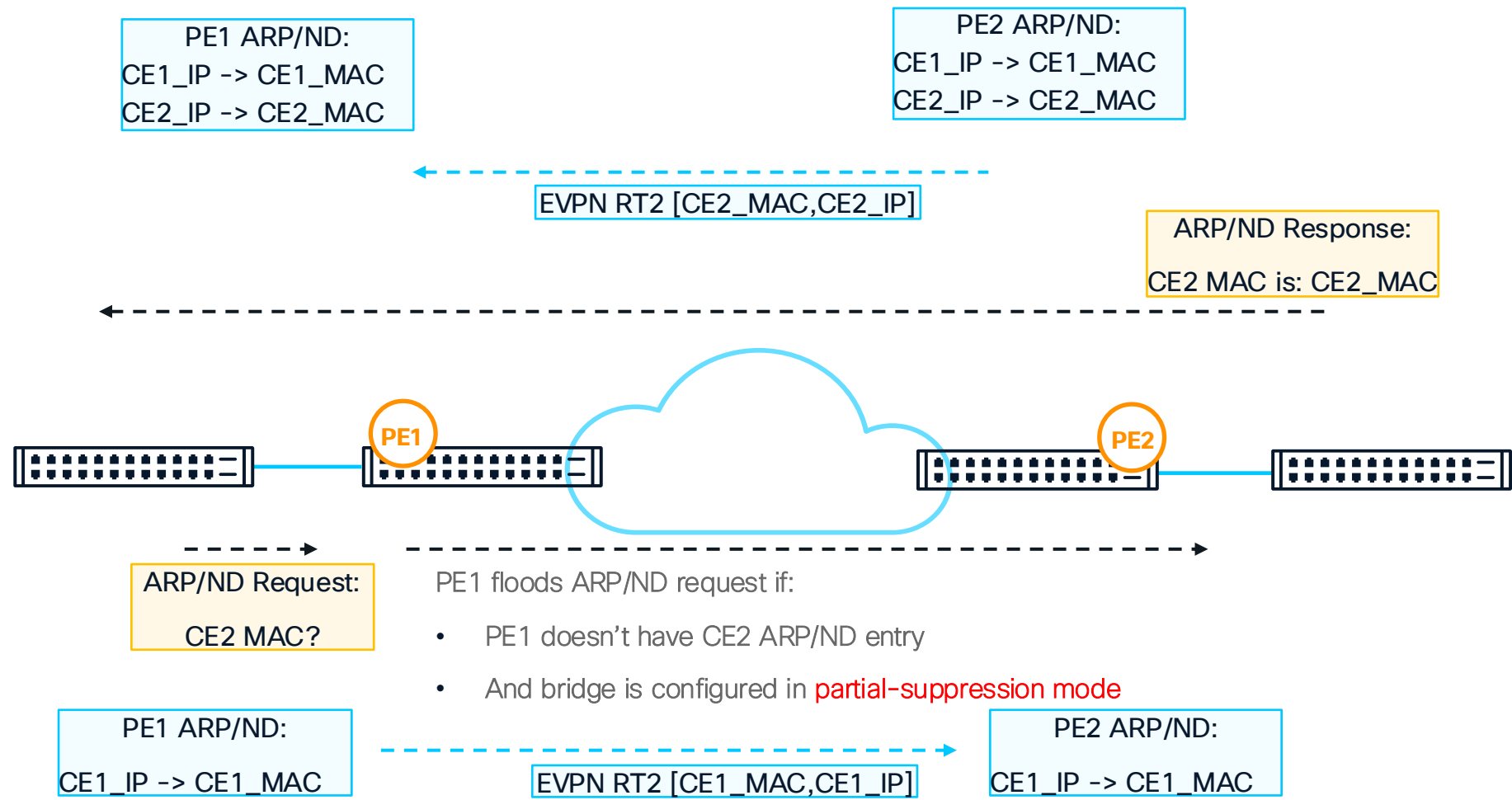
```
evpn
interface Bundle-Ether 1
  ethernet-segment
  identifier type 0 01.11.00.00.00.00.00.01
  load-balancing-mode port-active
  service-carving preference-based
  access-driven
!
```

ARP/ND Suppression

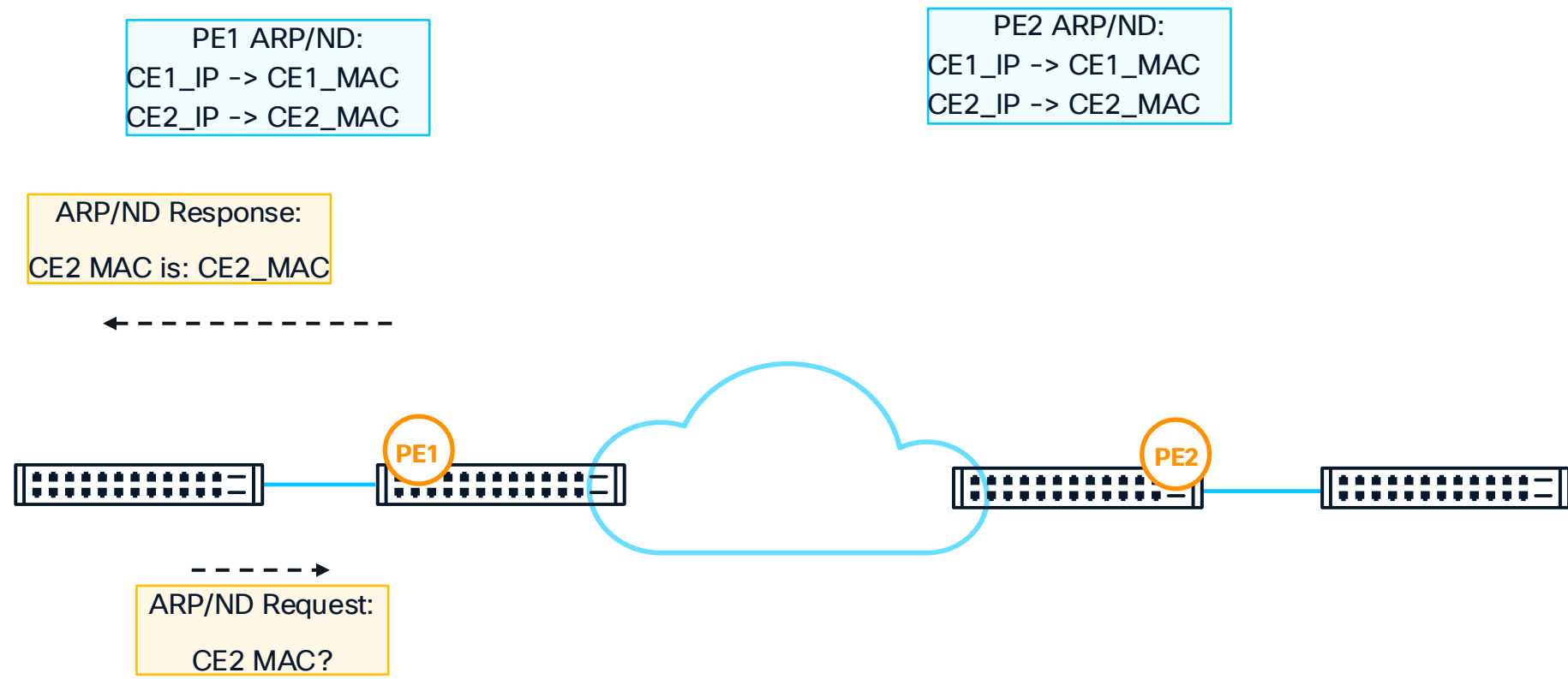
- Not the same as ARP/ND Proxy
 - ARP/ND suppression provides transparency to Host/CE
- Main goal is to eliminate ARP/ND
 - Bandwidth is not a concern (most of the time)
 - Host/CE procession is a concern in large L2 domains (IXP)
 - Host/CE ARP/ND stack implementation is not always "perfect"



ARP/ND Suppression



ARP/ND Suppression



Thank you

