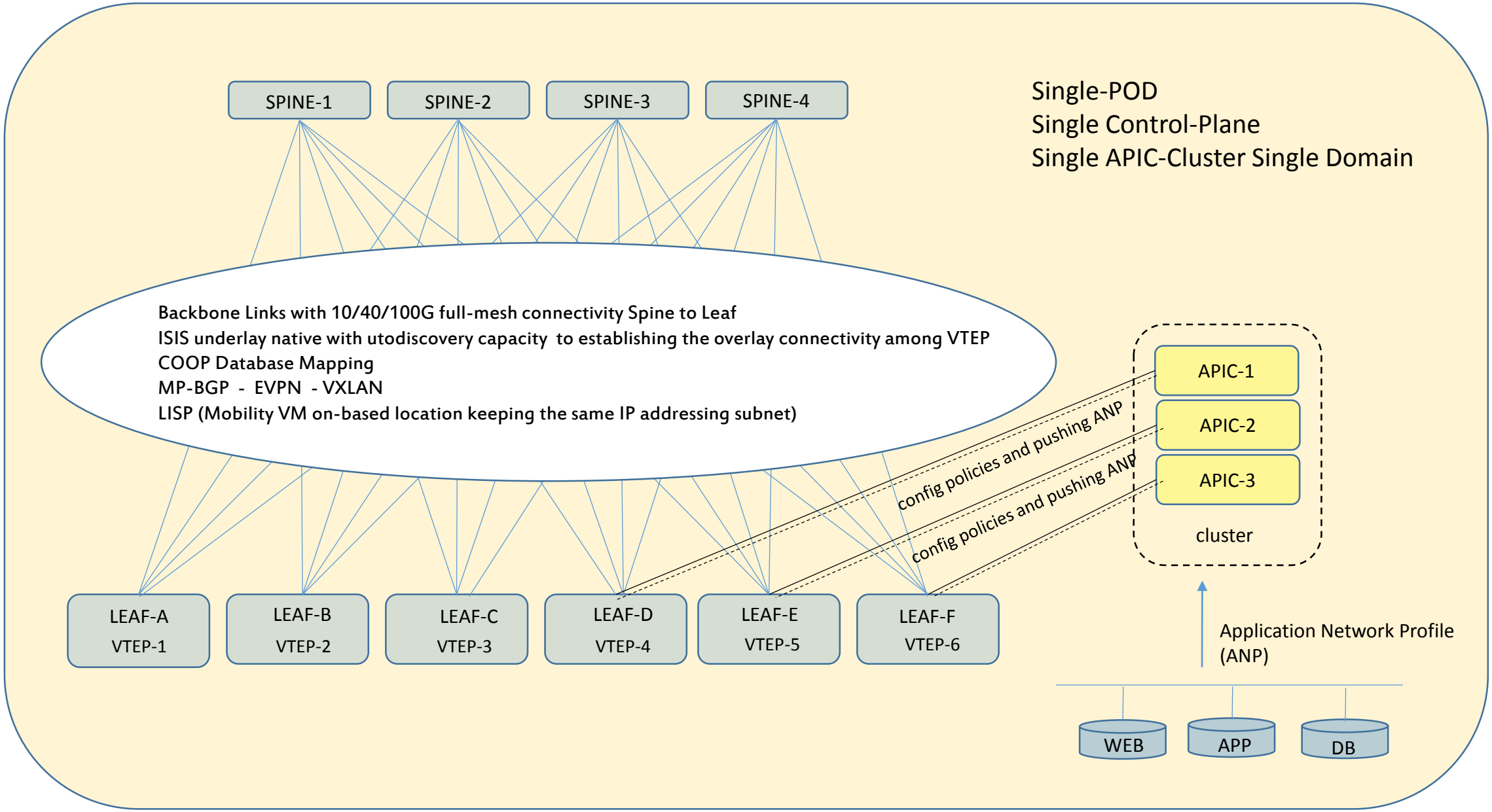
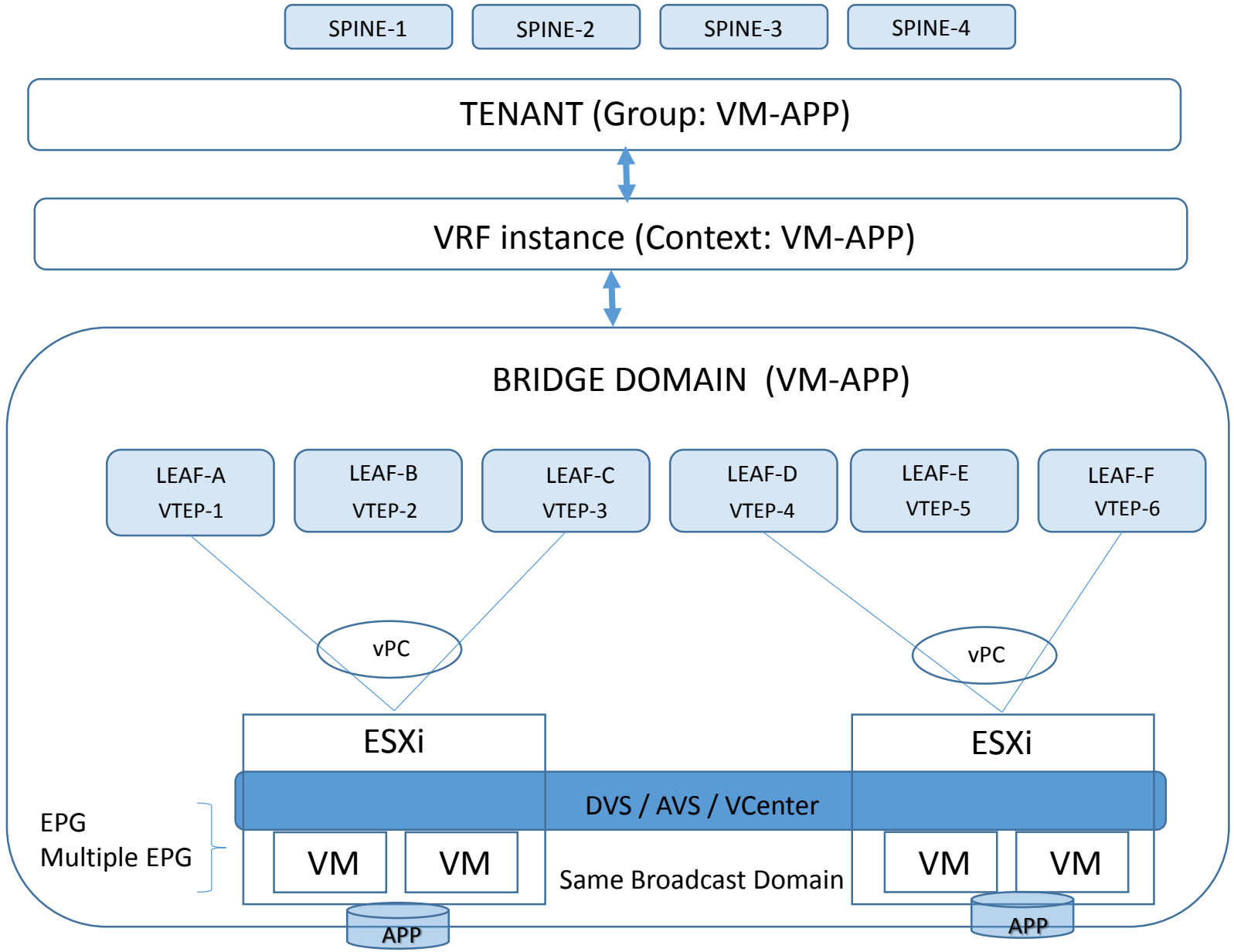
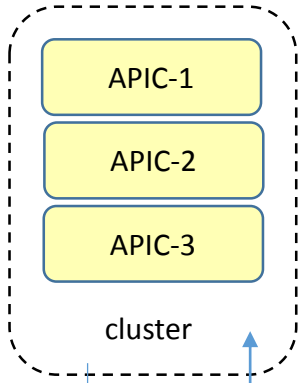
COSA SIGNIFICA ACI Fabric

Massimiliano Sbaraglia


SPINE-1
SPINE-2
SPINE-3
SPINE-4
Single-POD
Single Control-Plane
Single APIC-Cluster Single Domain
Backbone Links with 10/40/100G full-mesh connectivity Spine to Leaf
ISIS underlay native with utodiscovery capacity to establishing the overlay connectivity among VTEP
COOP Database Mapping
MP-BGP - EVPN - VXLAN
LISP (Mobility VM on-based location keeping the same IP addressing subnet)
APIC-1
APIC-2
APIC-3
cluster
config policies and pushing ANP
config policies and pushing ANP
LEAF-A
VTEP-1
LEAF-B
VTEP-2
LEAF-C
VTEP-3
LEAF-D
VTEP-4
LEAF-E
VTEP-5
LEAF-F
VTEP-6
Application Network Profile
(ANP)
WEB
APP
DB





Pushing Policies

Config Policies
Config ANP

Access Policies:

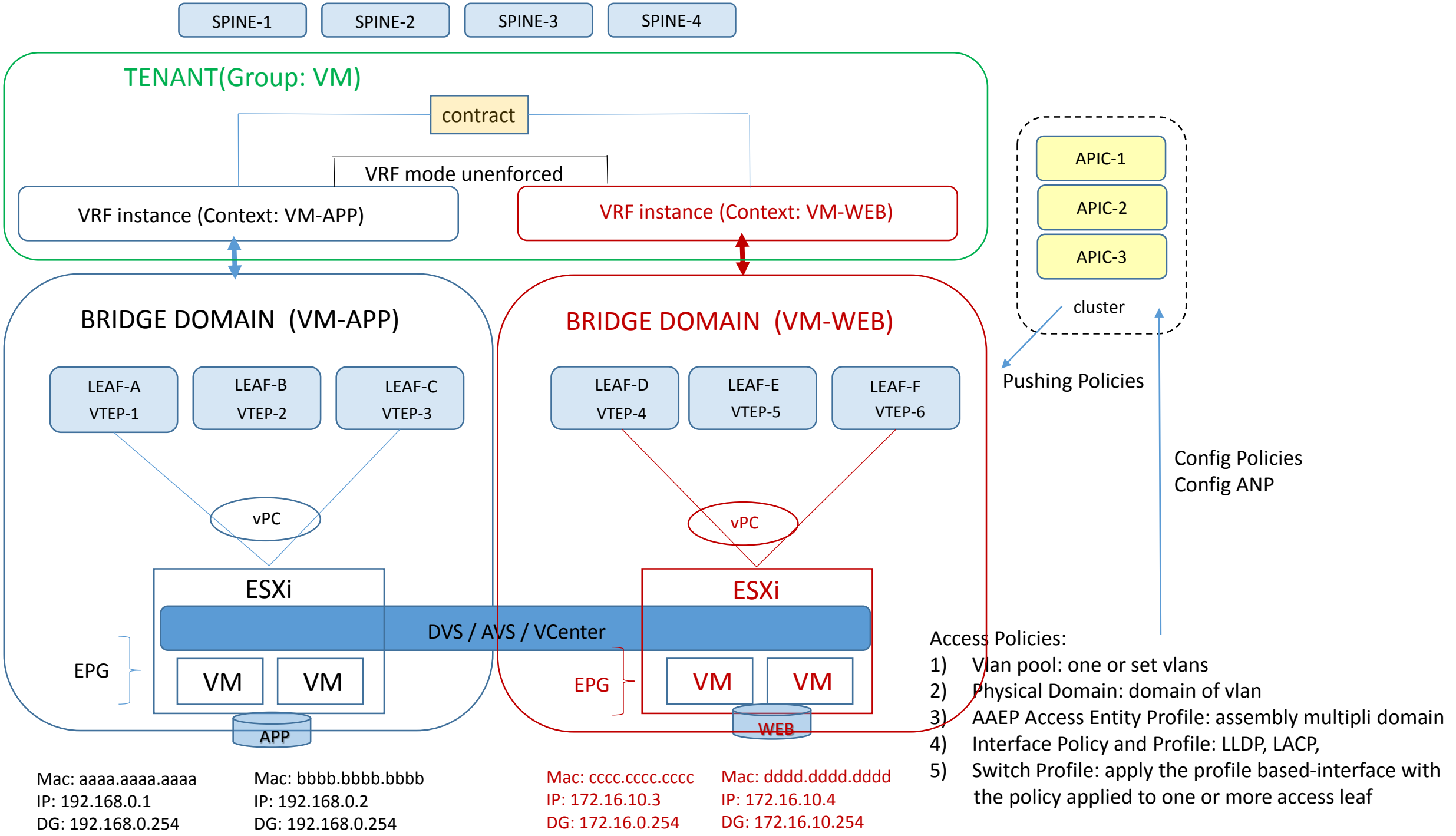
1) Vlan pool: one or set vlans
2) Physical Domain: domain of vlan
3) AAEP Access Entity Profile: assembly multipli domain
4) Interface Policy and Profile: LLDP, LACP,
5) Switch Profile: apply the profile based-interface with the pilicy applied to one or more access leaf

Mac: aaaa.aaaa.aaaa
IP: 192.168.0.1
DG: 192.168.0.254

Mac: bbbb.bbbb.bbbb
IP: 192.168.0.2
DG: 192.168.0.254

Mac: cccc.cccc.cccc
IP: 192.168.0.3
DG: 192.168.0.254

Mac: dddd.dddd.dddd
IP: 192.168.0.4
DG: 192.168.0.254

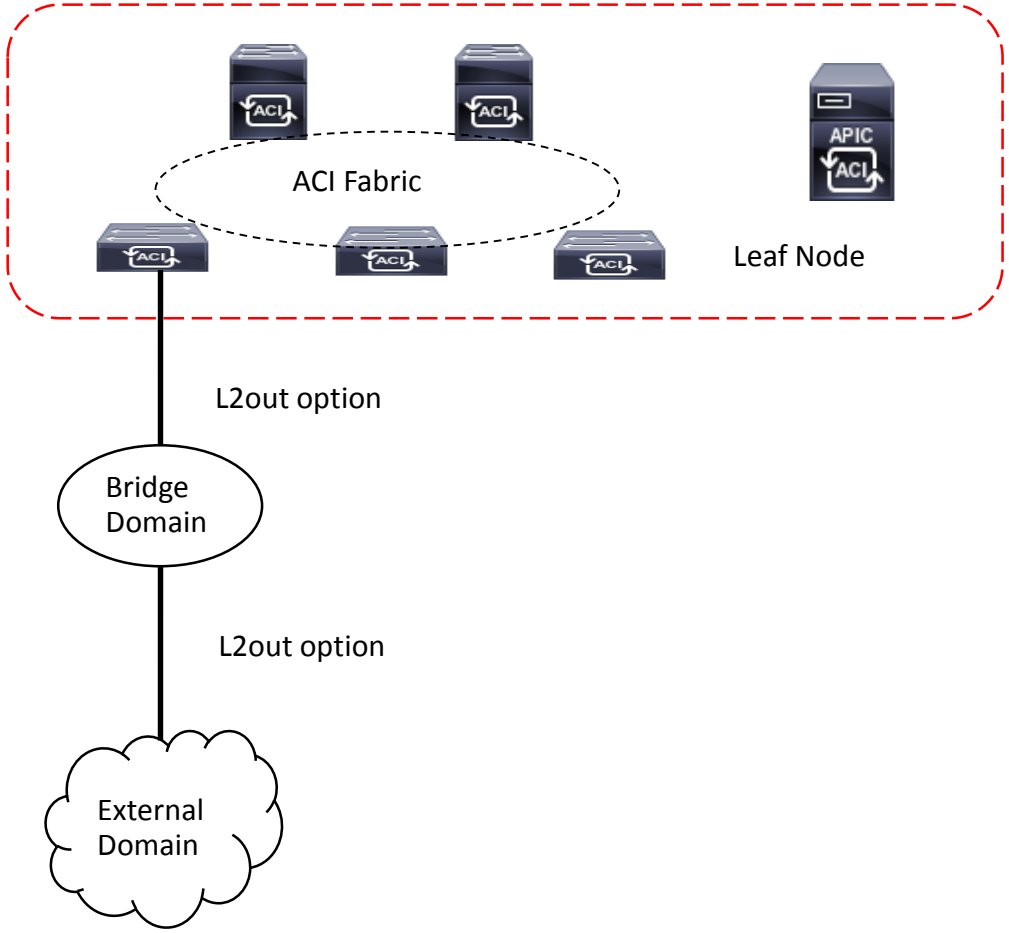

SPINE-1
SPINE-2
SPINE-3
SPINE-4
TENANT(Group: VM)
contract
VRF mode unenforced
VRF instance (Context: VM-APP)
VRF instance (Context: VM-WEB)
APIC-1
APIC-2
APIC-3
cluster
Pushing Policies
Config Policies
Config ANP
BRIDGE DOMAIN (VM-APP)
LEAF-A
VTEP-1
LEAF-B
VTEP-2
LEAF-C
VTEP-3
vPC
ESXi
DVS / AVS / VCenter
EPG
VM
VM
APP
BRIDGE DOMAIN (VM-WEB)
LEAF-D
VTEP-4
LEAF-E
VTEP-5
LEAF-F
VTEP-6
vPC
ESXi
EPG
VM
VM
WEB
Mac: aaaa.aaaa.aaaa
IP: 192.168.0.1
DG: 192.168.0.254
Mac: bbbb.bbbb.bbbb
IP: 192.168.0.2
DG: 192.168.0.254
Mac: cccc.cccc.cccc
IP: 172.16.10.3
DG: 172.16.0.254
Mac: dddd.dddd.dddd
IP: 172.16.10.4
DG: 172.16.10.254
Access Policies:
1) Vlan pool: one or set vlans
2) Physical Domain: domain of vlan
3) AAEP Access Entity Profile: assembly multipli domain
4) Interface Policy and Profile: LLDP, LACP,
5) Switch Profile: apply the profile based-interface with the policy applied to one or more access leaf

ACI (Application Centric Infrastructure) Cisco layer 2 steps di configurazione

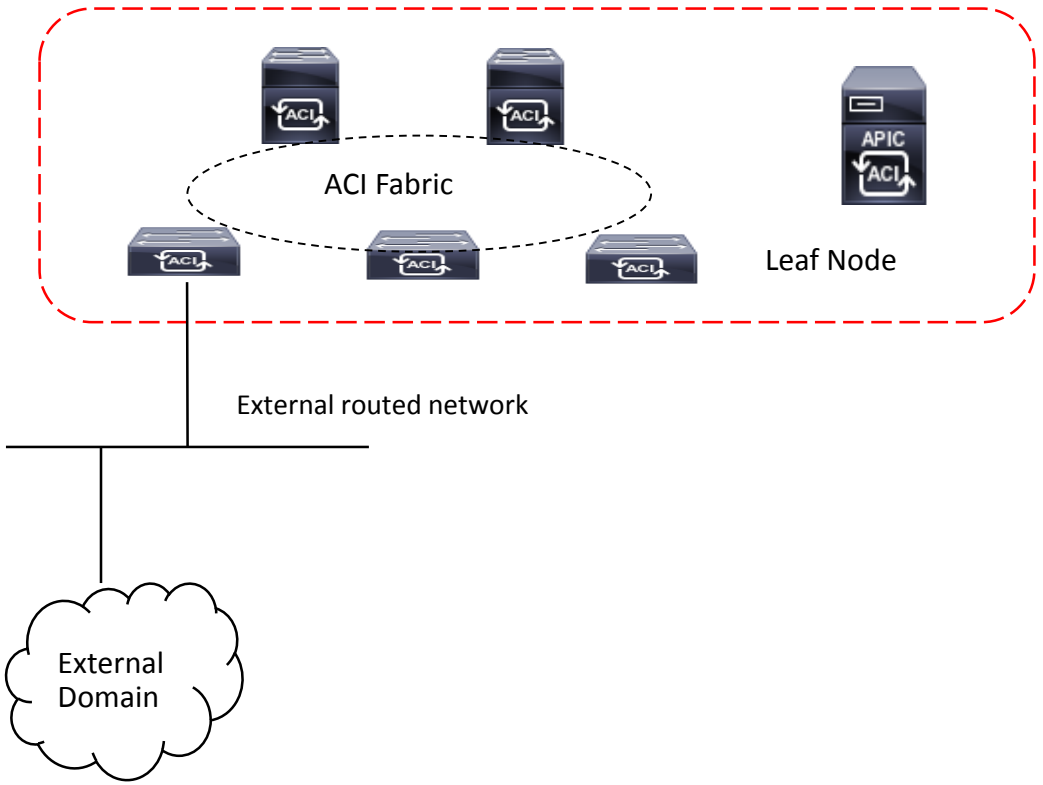
- VRF instances
- BD (Bridge Domain) associato alla VRF instance (senza abilitare nessun layer 3 IP SVIs subnet)
- Configurazione del Bridge Domain per ottimizzare la funzionalità di switching (hardware-proxy-mode) usando il mapping database oppure il tradizionale flood-and-learn
- EPG (End Point Group) relazionandoli ai bridge domain di riferimento; possiamo avere multipli EPG associati allo stesso bridge domain
- Creare policy Contracts tra EPG come necessario; possiamo anche considerare una comunicazione tra diversi EPG senza ausilio di filtri, settando la VRF instance in modalità < unenforced >
- Creare access policies switch e port profiles assegnando i parametri richiesti, associate al nodo Leaf di pertinenza

ACI (Application Centric Infrastructure) Cisco layer 2 option extending to external domain

- Enable flooding of layer 2 unknown unicast
- Enable ARP flooding
- Disable unicast routing (può essere abilitato successivamente ad una fase di migrazione ad esempio se gli end-point usano come IP gateway il sistema ACI Fabric
- L2Out option provvede ad una L2 extension da ACI Fabric ad un External domain bridged network

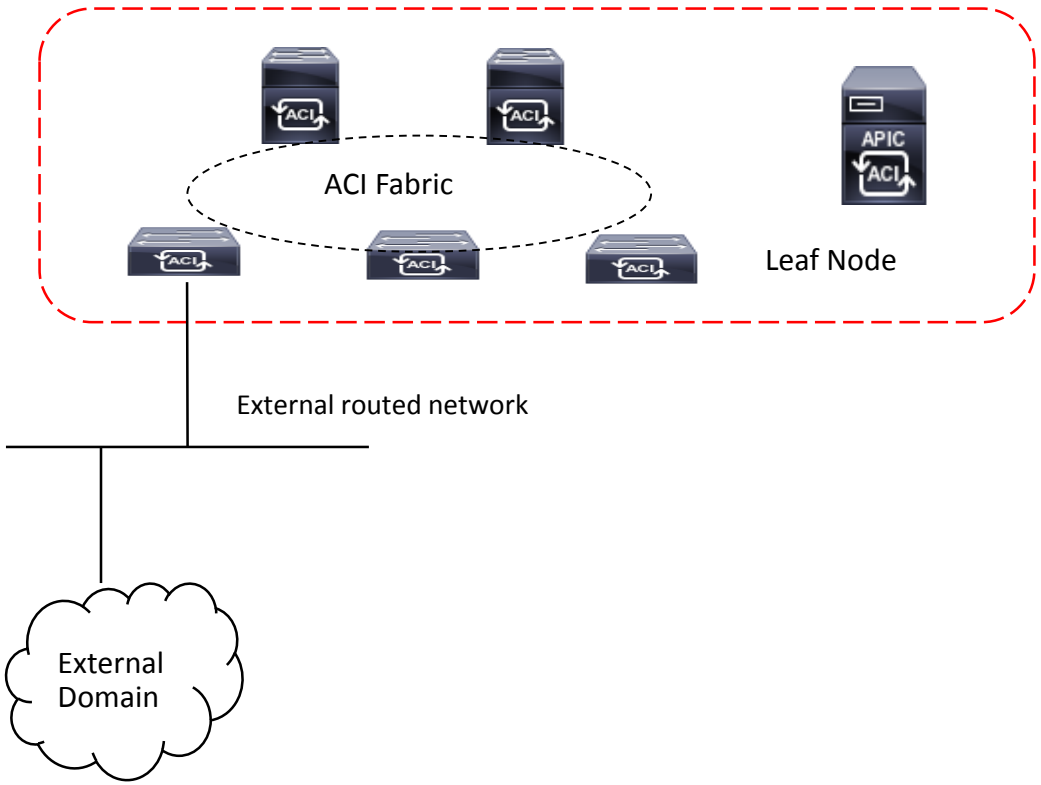


Aci fabric external network parameters



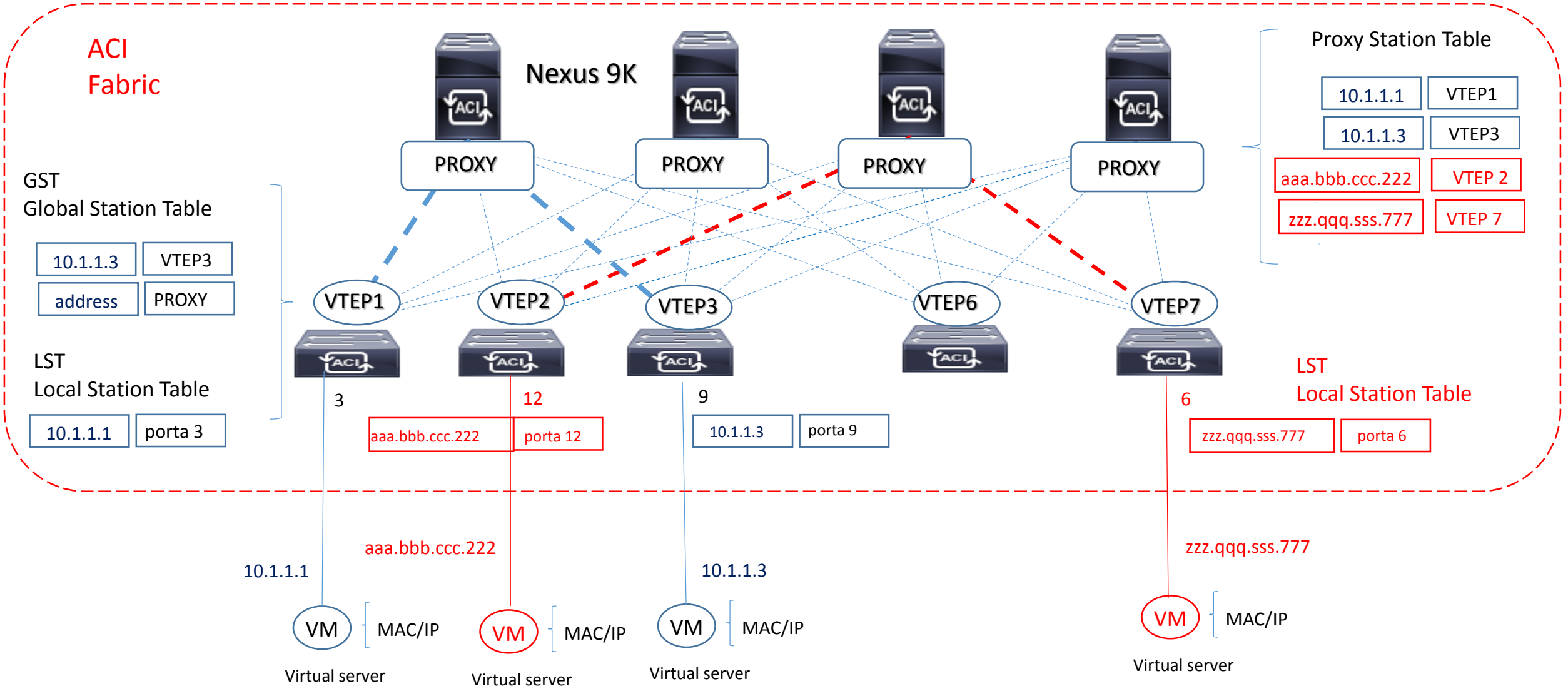
- Layer 3 interface routed: usata quando si connette un determinato external devices per tenant /VRF
- Subinterface with 802.1q tagging: usata quando vi è una connessione condivisa ad un determinato external devices attraverso tenants/ VRF-lite
- Switched Virtual Interface (SVI): usata quando entrambi i layer L2 ed L3 di connessione sono richiesti sulla stessa interfaccia
- La propagazione di external network all'interno di un dominio ACI Fabric utilizza il MP-BGP (Multi Protocol BGP) tra Spine e Leaf (si può avere anche la funzionalità di Route Reflector abilitato a livello Spine) all'interno di un unico AS

Aci fabric external network L3-out option



- Create un external routed network
- Set a layer 3 border leaf node for the L3 outside connection
- Set a layer 3 interface profile for the L3 outside connection
- Repeat step 2 and 3 if you need to add additional leaf nodes/interface
- Configure an external EPG (ACI Fabric maps the external L3 router to the external EPG by using the IP prefix and mask
- Configure a contract policies between the external and internal EPG (without this all connectivity to the outside will be blocked)

ACI (Application Centric Infrastructure) Cisco Control-Plane with mapping database

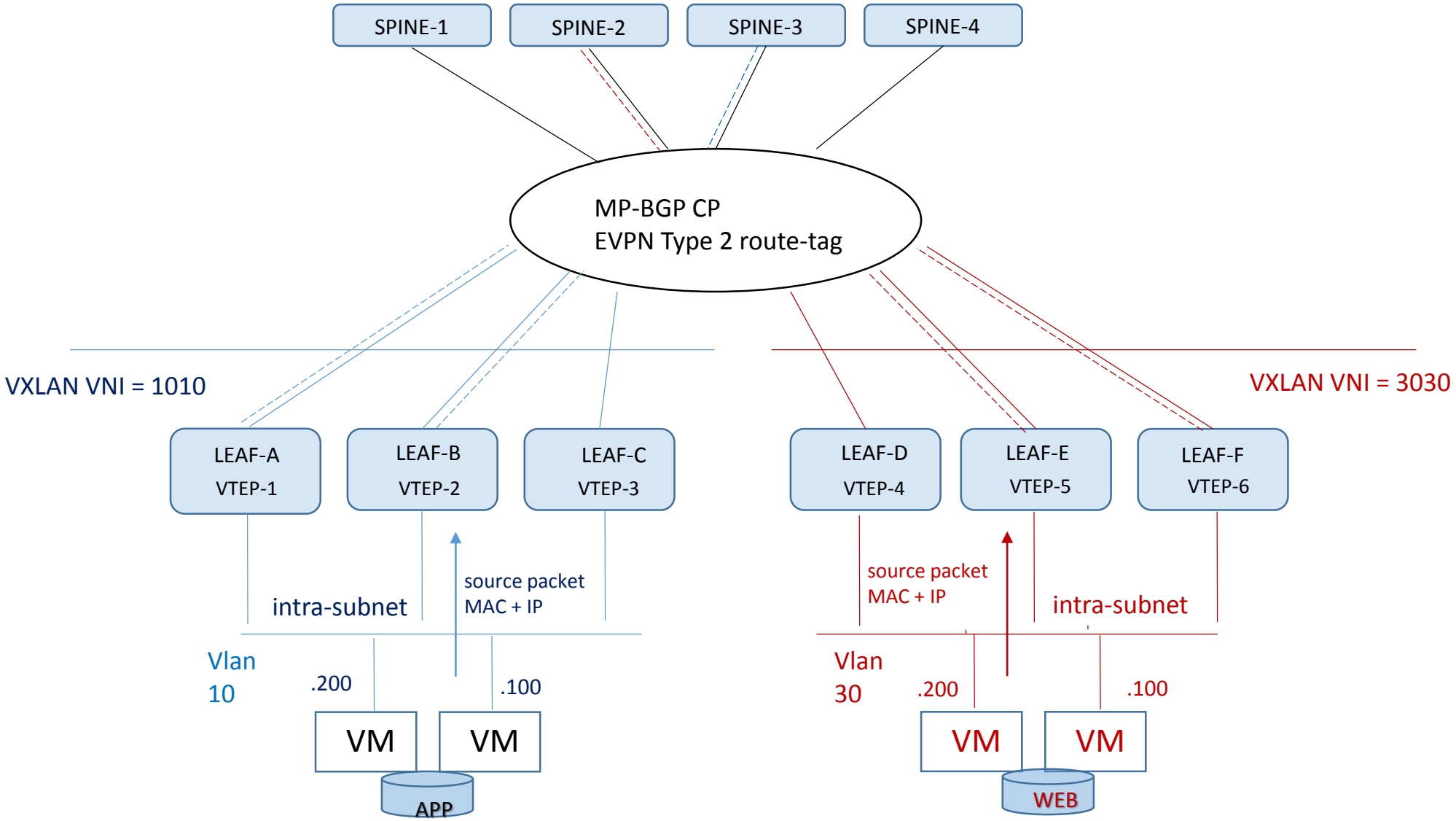


Learning Process End-Point information

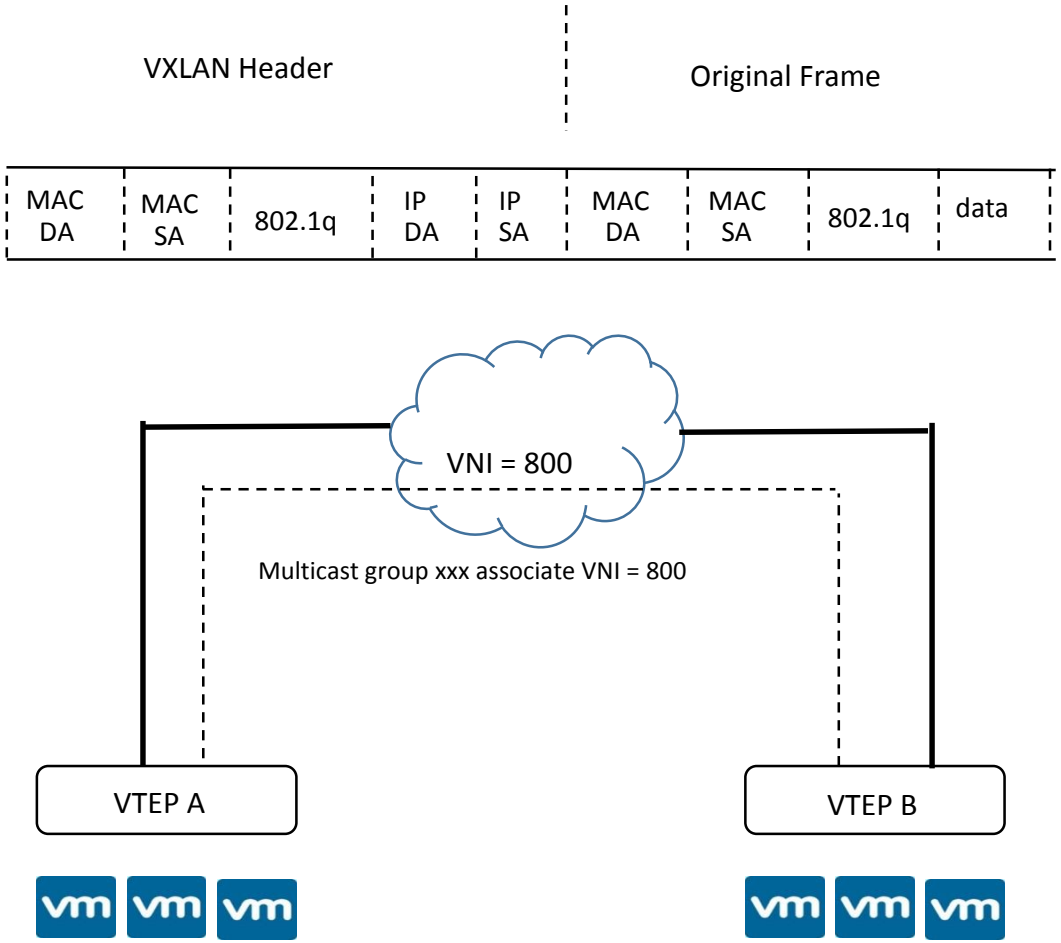
Il processo di learning Endpoint avviene a livello Edge Switch Leaf Node di una VXLAN EVPN Fabric, dove l'endpoint è direttamente connesso; le informazioni MAC address a livello locale sono calcolate attraverso la tabella di forwarding locale (data-plane table) mentre l'IP address è imparato attraverso meccanismi di ARP, GARP (Gratitous ARP) oppure IPv6 neighbor discovery message.

Una volta avvenuto il processo di apprendimento MAC + IP a livello locale, queste informazioni vengono annunciate dai rispettivi VTEP attraverso il MP-BGP EVPN control-plane utilizzando le EVPN route-type 2 advertisement trasmette a tutti i VTEP Edge devices che appartengono alla stessa VXLAN EVPN Fabric.

Di conseguenza, tutti gli edge devices imparano le informazioni end-point che appartengono ai rispettivi VNI (VXLAN segment Network Identifier) ed essere importate all'interno della propria forwarding table.

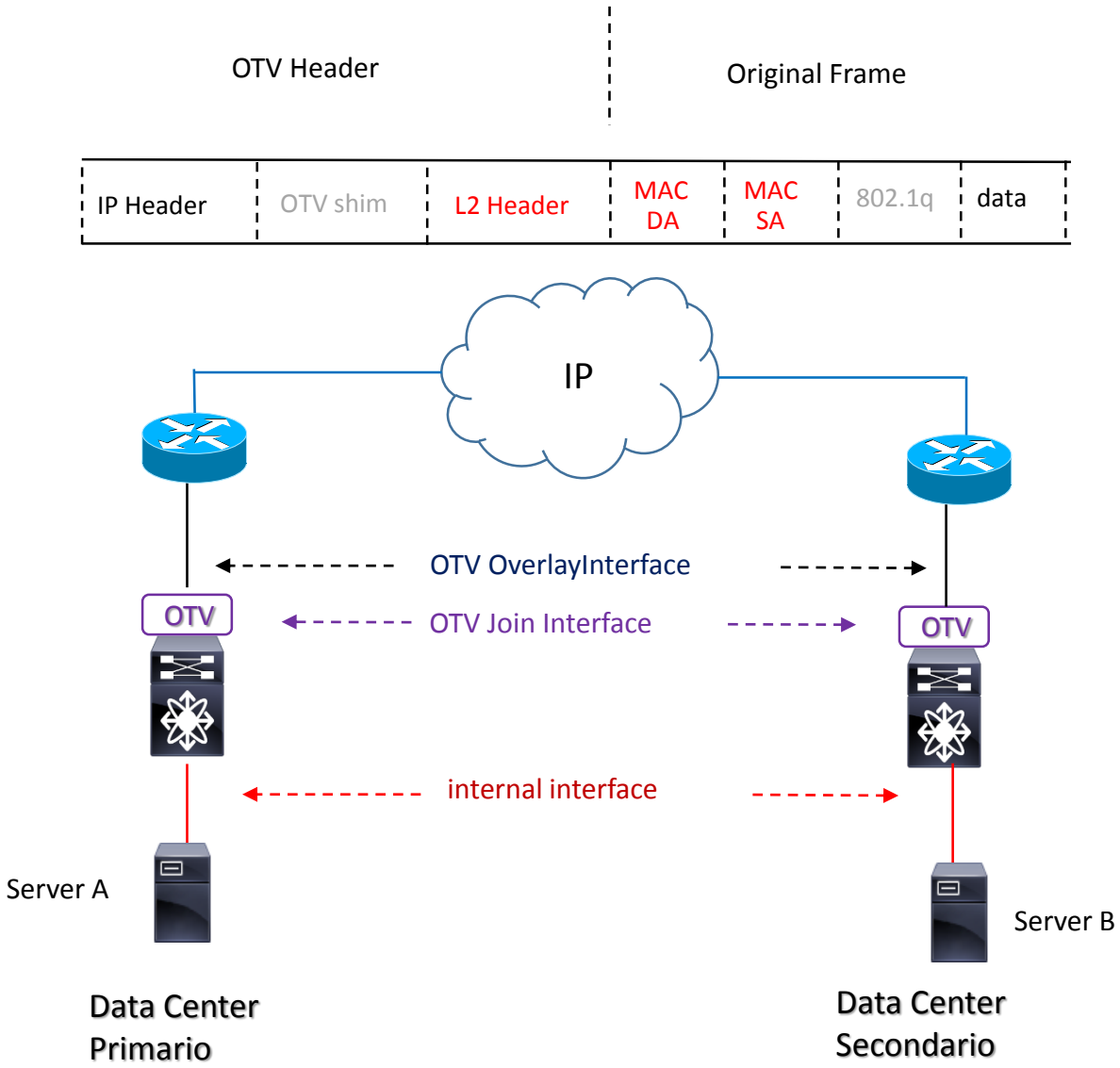

SPINE-1
SPINE-2
SPINE-3
SPINE-4
MP-BGP CP
EVPN Type 2 route-tag
VXLAN VNI = 1010
VXLAN VNI = 3030
LEAF-A
VTEP-1
LEAF-B
VTEP-2
LEAF-C
VTEP-3
LEAF-D
VTEP-4
LEAF-E
VTEP-5
LEAF-F
VTEP-6
intra-subnet
source packet
MAC + IP
source packet
MAC + IP
intra-subnet
Vlan
10
.200
.100
Vlan
30
.200
.100
VM
VM
VM
VM
APP
WEB

DCI vxlan (virtual extensible lan)



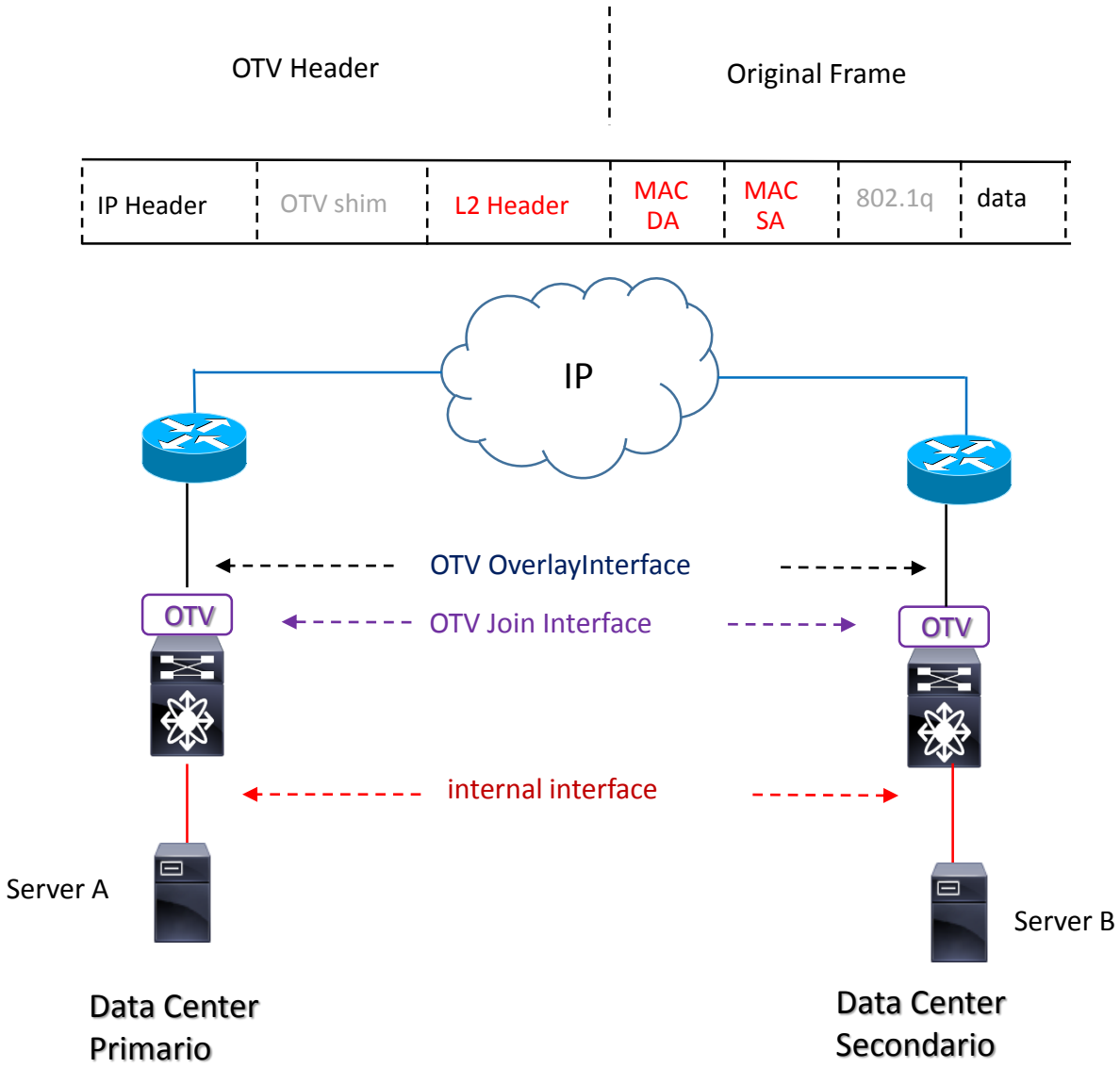
- VXLAN è un meccanismo che permette di aggregare e tunnellizzare (VTEP) multipli layer 2 subnetwork attraverso una infrastruttura layer 3 IP network
- Ogni VXLAN segment è associato con un unico 24 bit VXLAN Network Identifier differente chiamato VNI
- Questo 24 bit VNI permette di scalare da il classico 4096 vlans con 802.1q a più di 16 milioni di possibili virtual networks
- Le VMs servers all'interno di un dominio layer 2 utilizzano la stessa subnet IP e sono mappati con lo stesso valore VNI
- VXLAN mantiene l'indentità di ciascuna VMs mappando il valore di MAC address della VM con il valore VNI (possiamo avere duplicate MAC address all'interno di un datacenters domain ma con il limite che non possono essere mappati con lo stesso VNI)
- Il gateway VTEP deve essere configurato associando il dominio L2 or L3 al VNI network value e quest'ultimo ad un gruppo IP multicast; quest'ultima configurazione permette ai VTEP la costruzione di una forwarding table attraverso l'infrastruttura di rete

DCI OTV CISCO (overlay transport virtualization)



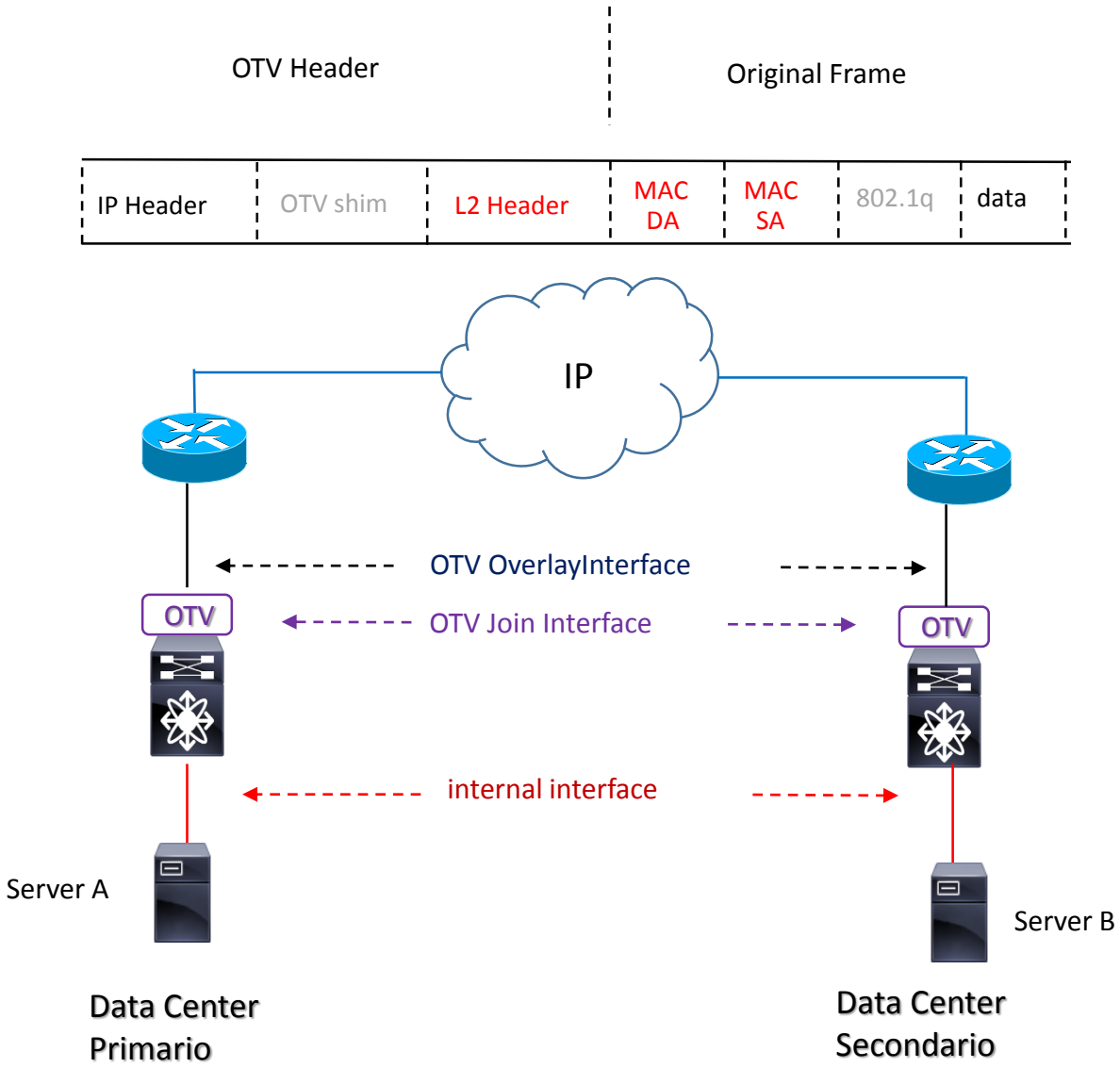
- OTV è una infrastruttura inter-datacenters e provvede a L2 extensions presenrvando fault-isolation, resilienza e load-balancing
- Il requisito è che deve esserci connettività IP tra i due datacenters
- OTV introduce il concetto di Layer 2 MAC routing (MAC in IP) che abilita il piano di controllo (control-plane) di annunciare la raggiungibilità MAC addressess; con il piano di controllo MAC address learning, OTV non trasmette (flood) unknown unicast traffic e il traffico ARP è trasmesso solo in modo controllato
- OTV non propaga BPDU STP attraverso l'infrastruttura di trasporto overlay
- OTV utilizza Nexus Cisco con VDC (Virtual Context Domain) ed è mandatorio avere vlans extended con layer 3 SVI (switched virtual interface) per una data vlan
- La funzionalità site-vlan è utilizzata per la scoperta di edge devices remoti in una topologia multi-homed: in aggiunta al site-vlan, l'edge devices mantiene una seconda OTV adiacenza con gli altri edge devices appartenenti allo stesso datacenter

DCI OTV CISCO (overlay transport virtualization)



- **OTV Edge Device:** performa le funzionalità e le operazioni OTV; riceve le frame ethernet traffic per tutte le vlans soggette ad L2-extensions tra data centers OTV peers e dinamicamente le incapsula dentro IP packets che sono trasmessi via overlay transport infrastructure
- **OTV internal interface:** sono le interfacce di un edge device che connette il datacenter locale con una configurazione generalmente in trunk trasportando multiple vlans. Non prevedono nessuna configurazione OTV compliant
- **OTV join interface:** sono le interfacce uplink di un edge device che si affacciano alla rete core overlay IP; questo tipo di interfacce sono point-to-point layer 3 routed, subinterface, port-channel oppure port-channel subinterface (No loopback) ed hanno lo scopo di essere le sorgenti di traffico OTV incapsulato e trasmesso verso l'infrastrutura overlay
- **OTV overlay interface:** sono interfacce logiche virtuali dove risiede tutta la configurazione OTV; incapsula le frame layer 2 in IP unicast o multicast packets che sono trasmesse verso altri datacenters. Questo pemette agli edge device di performare un dinamico encapsulations.

DCI OTV CISCO (overlay transport virtualization)



- **OTV site vlan:** è una funzionalità utilizzata per scoprire altri Edge Devices in una topologia multi-homed
- **OTV site ID:** sappiamo che le adiancenze OTV sono costruite via le join interface attraverso la rete IP overlay; ogni edge device all'interno dello stesso site hanno lo stesso site-id configurato; dalla release NX-OS 5.2.1 una seconda OTV adiancenza è mantenuta con lo scopo di protezione in caso di partizionamento di site-vlan tra edge devices all'interno dello stesso site.
- **AED authoritatve edge device:** è responsabile della trasmissione di layer 2 traffic incluso unicast, multicast e broadcast; è responsabile di annunciare la raggiungibilità dei mac-addresses verso i datacenters remoti;

EVPN MP-BGP

EVPN (Ethernet Virtual Private Network) collega un gruppo di users sites usando un virtual bridge layer 2;

Tratta indirizzi MAC come address ruotabili e distribuisce queste informazioni via MP-BGP;

Utilizzato in ambienti Data Centers multi-tenancy con end-point virtualizzati; supporta encapsulamento VXLAN e lo scambio di indirizzi IP host e IP-Prefix.

EVPN MP-BGP control plane

- informazioni layer 2 (MAC address) e layer 3 (host IP address) imparate localmente da ogni VTEP sono propagate ad altri VTEP permettendo funzionalità di switching e routing all'interno della stessa fabbrica;
- le routes sono annunciate tra VTEP attraverso route-target policy;
- utilizzo di VRF e route-distinguisher per routes/subnet;
- Le informazioni layer 2 sono distribuite tra VTEP con la funzionalità di ARP cache per minimizzare il flooding;
- le sessioni L2VPN EVPN tra VTEP possono essere autenticate via MD5 per mitigare problematiche di sicurezza (Rogue VTEP)

In genere un data centers IaaS costruito su una architettura Spine-Leaf utilizza per migliorare le sue performance di raggiungibilità layer 2 e 3 un processo ECMP (Equal Cost Multi Path) via IGP.

In caso di crescita della Fabric con la separazione multi-tenant, si può pensare a meccanismi di scalabilità come il protocollo BGP e scegliere se utilizzare Internal-BGP oppure external-BGP in considerazione anche di meccanismi ECMP molto utili in ambienti datacenters

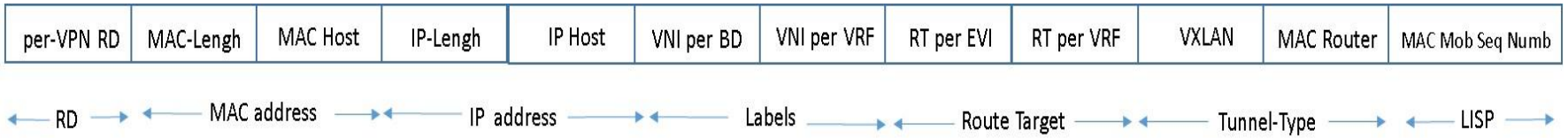
IBGP richiede sessioni tra tutti i PE VTEP e l'impiego di Router Reflector aiuta molto in termini di scalabilità delle sessioni configurati a livello Spine; questo tipo standard di soluzione, in ogni caso, riflette solo il best-single-prefix verso i loro client ed nella soluzione di utilizzare ECMP bisogna configurare un BGP add-path feature per aggiungere ECMP all'interno degli annuncia da parte dei RRs

EBGP, invece, supporta ECMP senza add-path ed è semplice nella sua tradizionale configurazione; con EBGP ogni devices della Fabric utilizza un proprio AS (Autonomous System)

EVPN MP-BGP route-type

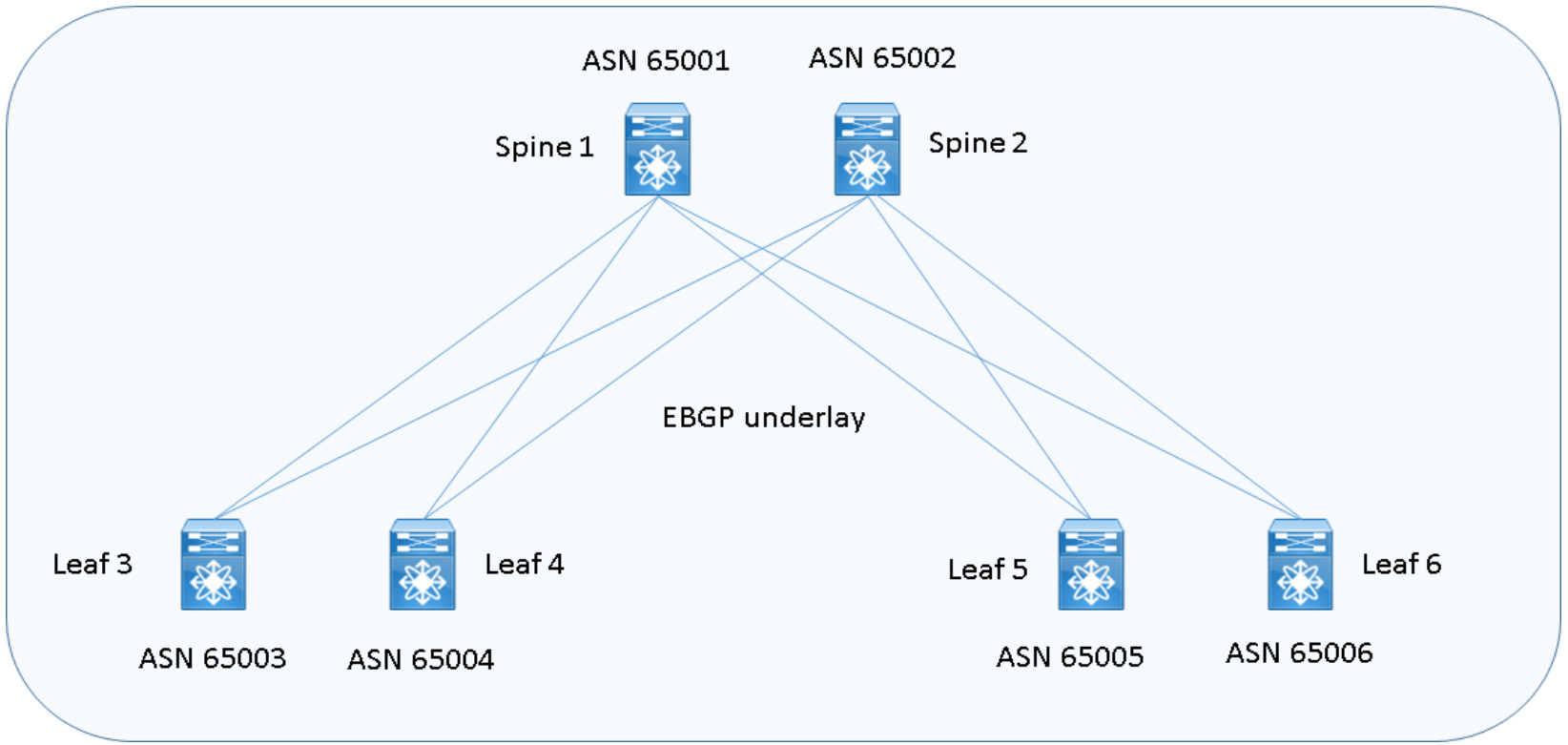
MP-BGP EVPN utilizza due routing advertisement:

- ✓ **Route type 2:** usato per annunciare host MAC ed IP address information per gli endpoint direttamente collegati alla VXLAN EVPN Fabric, ed anche trasportare extended community attribute, come route-target, router MAC address e sequence number

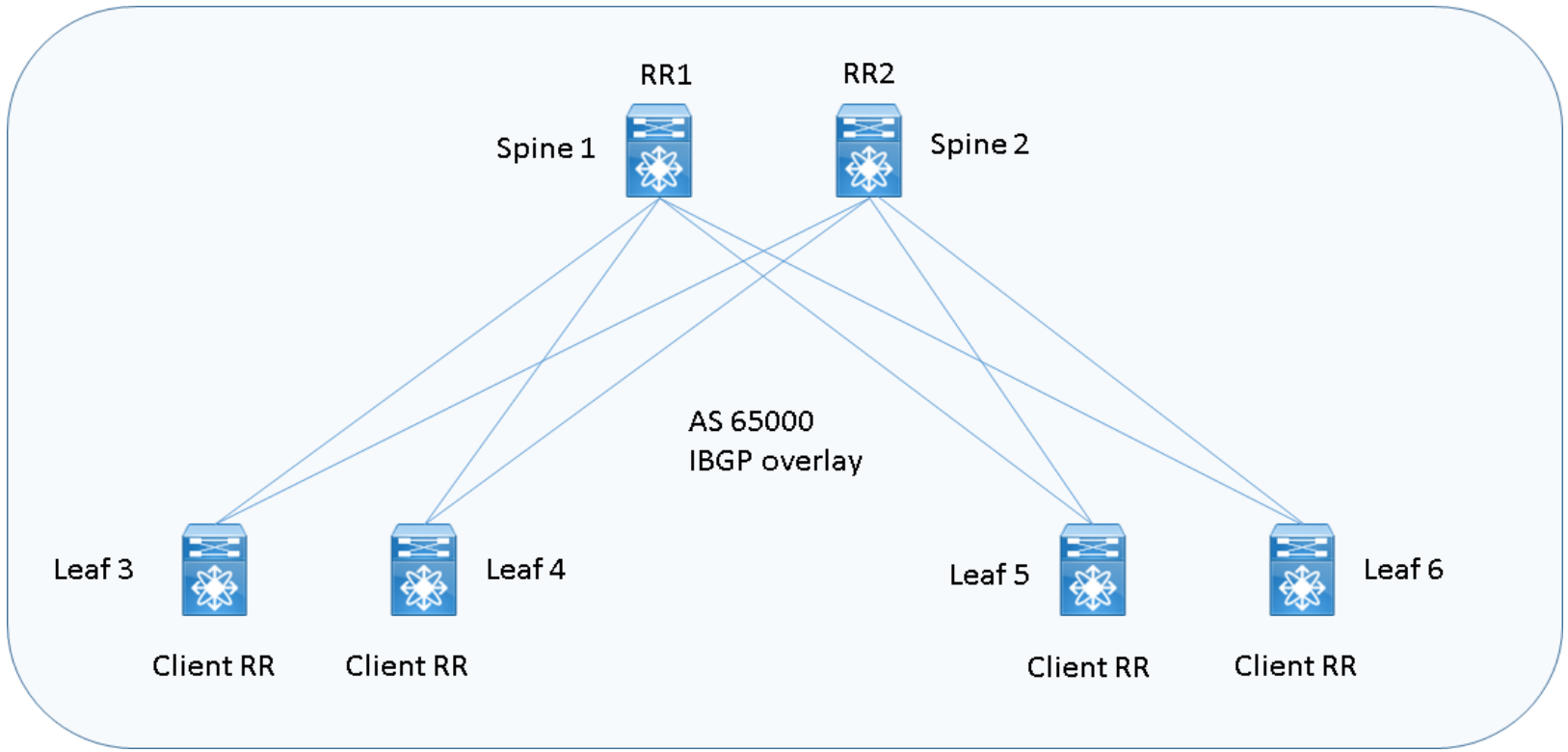


- ✓ **Route type 5:** annuncio di IP Prefix oppure host routes (loopback interface) ed anche trasporto di extended community attribute, come route-target, router MAC address e sequence number

EVPN E-BGP and ASN underlay design



EVPN I-BGP and ASN underlay design



Distributed Anycast Protocol Gateway

Protocolli FHRP quali HSRP, VRRP e GLBP hanno funzionalità di alta affidabilità layer 3 attraverso meccanismi active-standby routers e VIP address gateway condiviso.
Distributed Anycast Protocol, supera la limitazione di avere solo due routers peers HSRP/VRRP in ambienti Data Centers, costruendo una VXLAN EVPN VTEP Fabric con una architettura di tipo Spine-Leaf.

Distributed Anycast Proocol offre i seguenti vantaggi:

- ✓ stesso IP address gateway per tutti gli Edge Switch; ogni endpoint ha come gateway il proprio local VTEP il quale ruota poi il traffico esternamente ad altri VTEP attraverso una rete IP core (questo vale sia per VXLAN EVPN costruito come Fabric locale che geograficamente distribuito);

- ✓ la funzionalità di ARP suppression permette di ridurre il flooding all'interno del proprio dominio di switching (Leaf to Edge Switch);

- ✓ permette il moving di host/server continuando a mantenere lo stesso IP address gateway configurato nel local VTEP, all'interno di ciascuna VXLAN EVPN Fabric locale o geograficamente distribuita;

- ✓ No FHRP Filtering tra VXLAN EVPN Fabrics

Intra-Subnet and Inter-Subnet communication via EVPN Fabric

La comunicazione tra due end-point intra-subnet (stessa subnet IP) ubicati su EVPN Fabric differenti è stabilito attraverso la combinazione di creare un bridge domain L2 VXLAN (all'interno di ogni Fabric) e un L2 extension segment di rete IP address tra Fabrics;

La comunicazione tra due end-point inter-subnet (differente subnet IP) avviene sempre tra due endpoint EVPN ubicati in differenti Fabrics, ma con due differenti subnets IP default gateway.

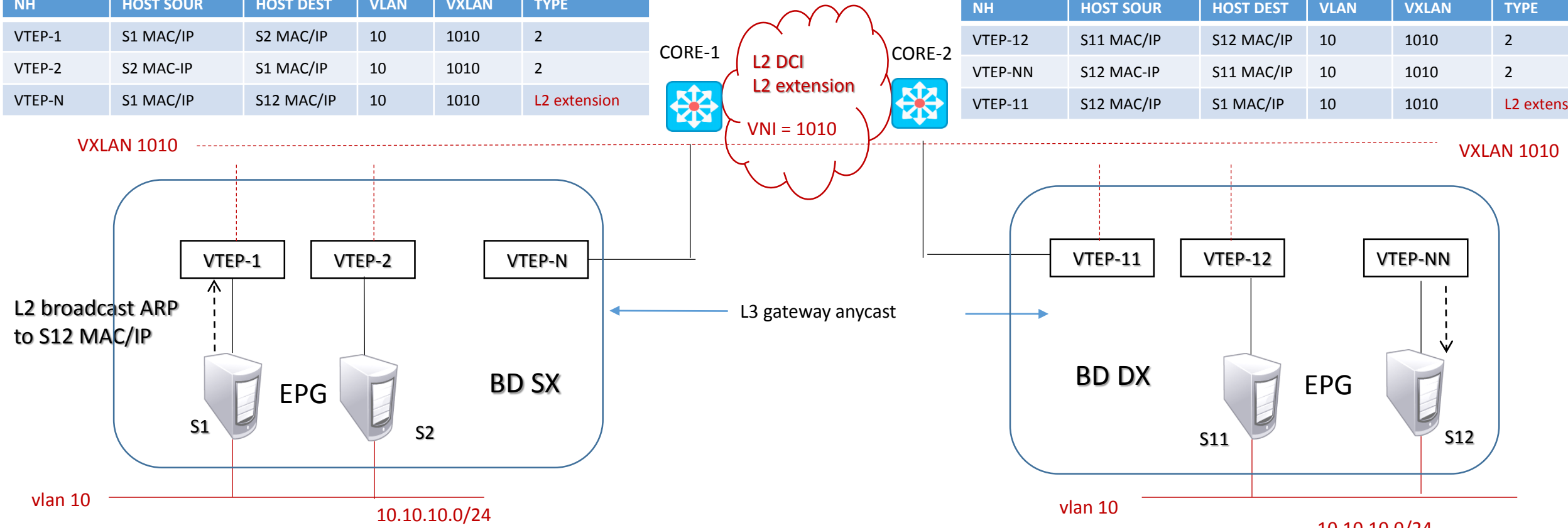
Intra-Subnet design communication via EVPN Fabric

SPINE-1 SPINE-2

NH	HOST SOUR	HOST DEST	VLAN	VXLAN	TYPE
VTEP-1	S1 MAC/IP	S2 MAC/IP	10	1010	2
VTEP-2	S2 MAC-IP	S1 MAC/IP	10	1010	2
VTEP-N	S1 MAC/IP	S12 MAC/IP	10	1010	L2 extension

SPINE-1 SPINE-2

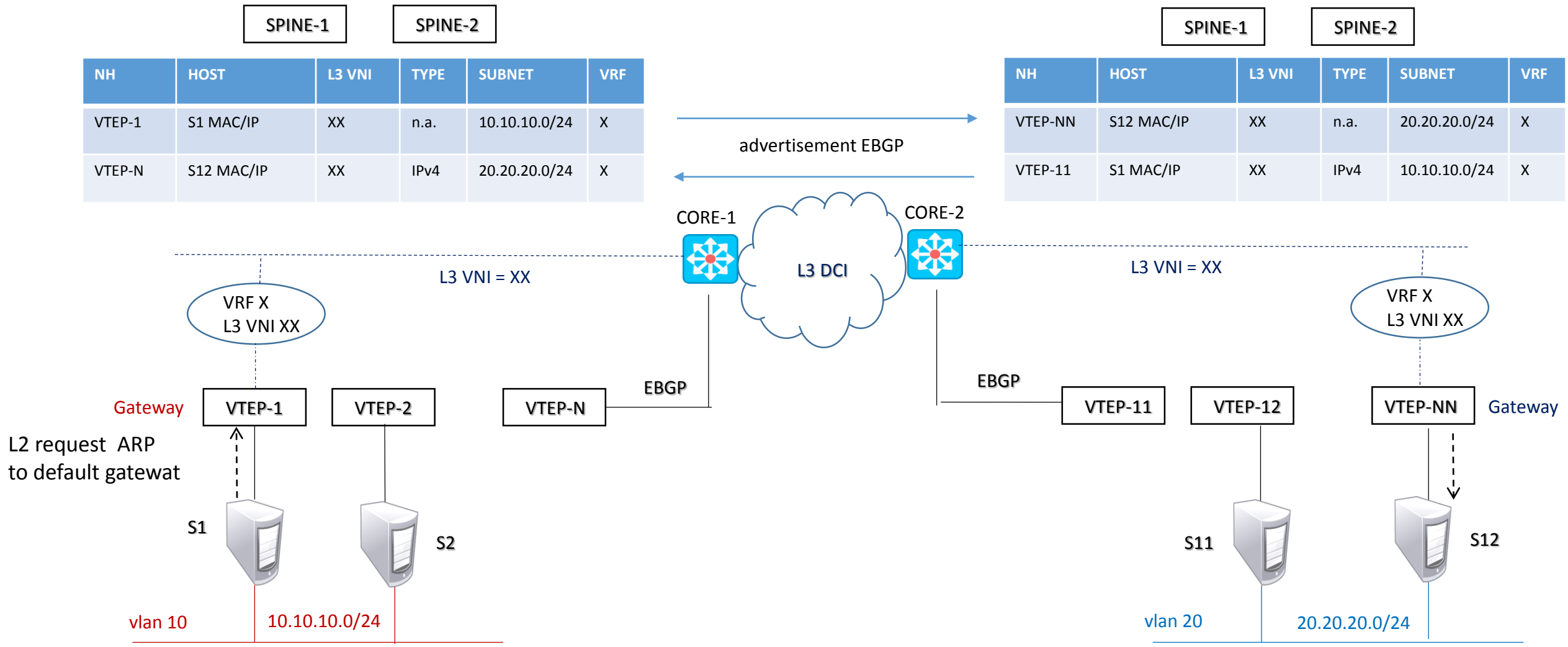
NH	HOST SOUR	HOST DEST	VLAN	VXLAN	TYPE
VTEP-12	S11 MAC/IP	S12 MAC/IP	10	1010	2
VTEP-NN	S12 MAC-IP	S11 MAC/IP	10	1010	2
VTEP-11	S12 MAC/IP	S1 MAC/IP	10	1010	L2 extension



VTEP	NEXT HOP	HOST	TYPE
1	S1 MAC/IP	LOCAL	
1	S2 MAC/IP	VTEP-2	2
1	S12 MAC/IP	VTEP-N	2

VTEP	NEXT HOP	HOST	TYPE
VTEP-NN	S12 MAC/IP	LOCAL	
VTEP-NN	S11 MAC/IP	VTEP-12	2
VTEP-NN	S1 MAC/IP	VTEP-11	2

Inter-Subnet design communication via EVPN Fabric



NH	HOST	L3 VNI	TYPE	SUBNET	VRF
VTEP-1	S1 MAC/IP	XX	n.a.	10.10.10.0/24	X
VTEP-N	S12 MAC/IP	XX	IPv4	20.20.20.0/24	X

NH	HOST	L3 VNI	TYPE	SUBNET	VRF
VTEP-NN	S12 MAC/IP	XX	n.a.	20.20.20.0/24	X
VTEP-11	S1 MAC/IP	XX	IPv4	10.10.10.0/24	X

VTEP	NEXT HOP	HOST	TYPE
1	S1 MAC/IP	LOCAL	
1	S2 MAC/IP	VTEP-2	2
1	S12 MAC/IP	Request ARP	5

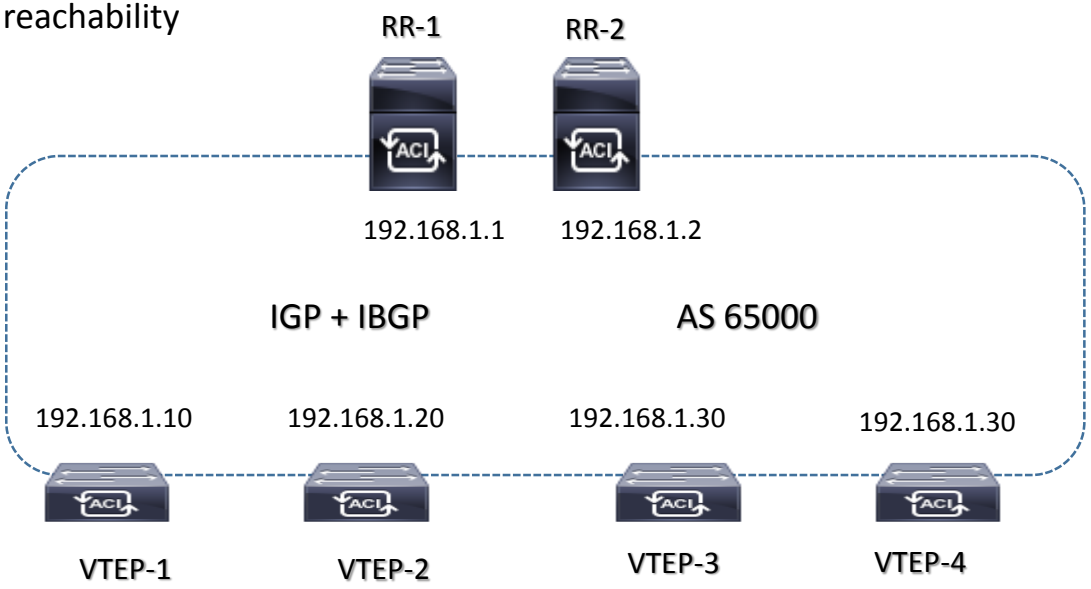
VTEP	NEXT HOP	HOST	TYPE
VTEP-NN	S12 MAC/IP	LOCAL	
VTEP-NN	S11 MAC/IP	VTEP-12	2
VTEP-NN	S1 MAC/IP	Request ARP	5

EVPN I-BGP Configurations VTEP (VXLAN Tunnel End-Point)

feature bgp
feature nv overlay → enable VTEP (required on Leaf or Border)
fature nv overlay evpn → enable EVPN control-plane in BGP

@ only on LEAF

interface nve1 → enable interface VTEP
source-interface loopback0 → enable source interface with loopback
host-reachability protocol bgp → enable BGP for host reachability



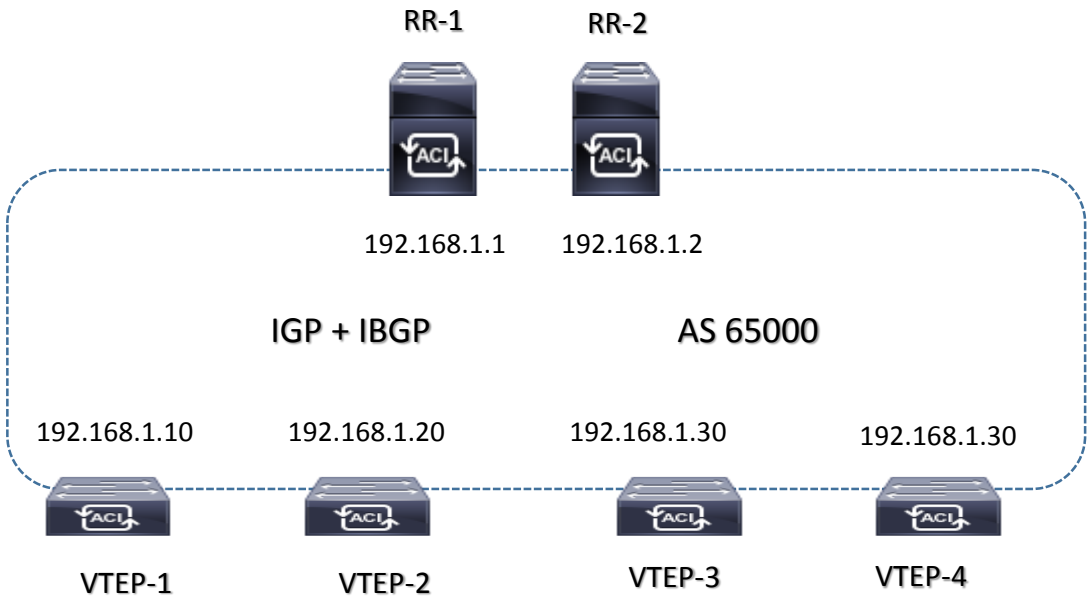
EVPN I-BGP Configurations Overlay Control Plane

 # SPINE RR1

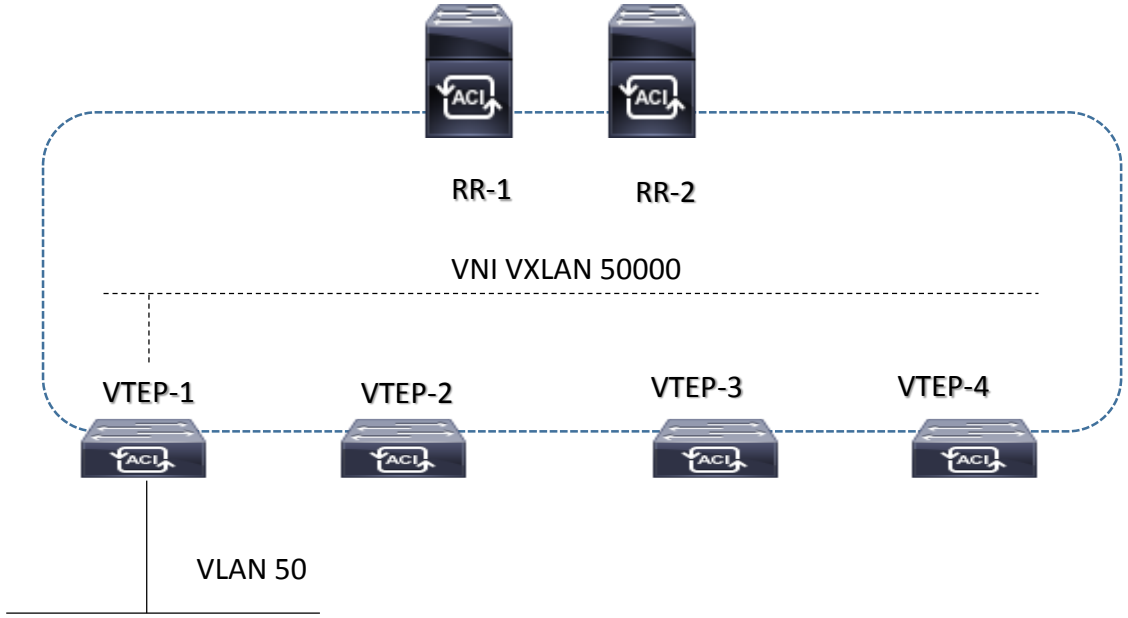
```
router bgp 65000
router-id 192.168.1.1
address-family ipv4 unicast
neighbor 192.168.1.10  remote-as 65000
  update-source loopback0
  address-family l2vpn evpn
  send-community both
  route-reflector client
```

 # LEAF VTEP-1

```
router bgp 65000
router-id 192.168.1.10
address-family ipv4 unicast
neighbor 192.168.1.1 remote-as 65000
  update-source loopback0
  address-family l2vpn evpn
  send-community both
```



EVPN I-BGP Configurations VLAN to VXLAN



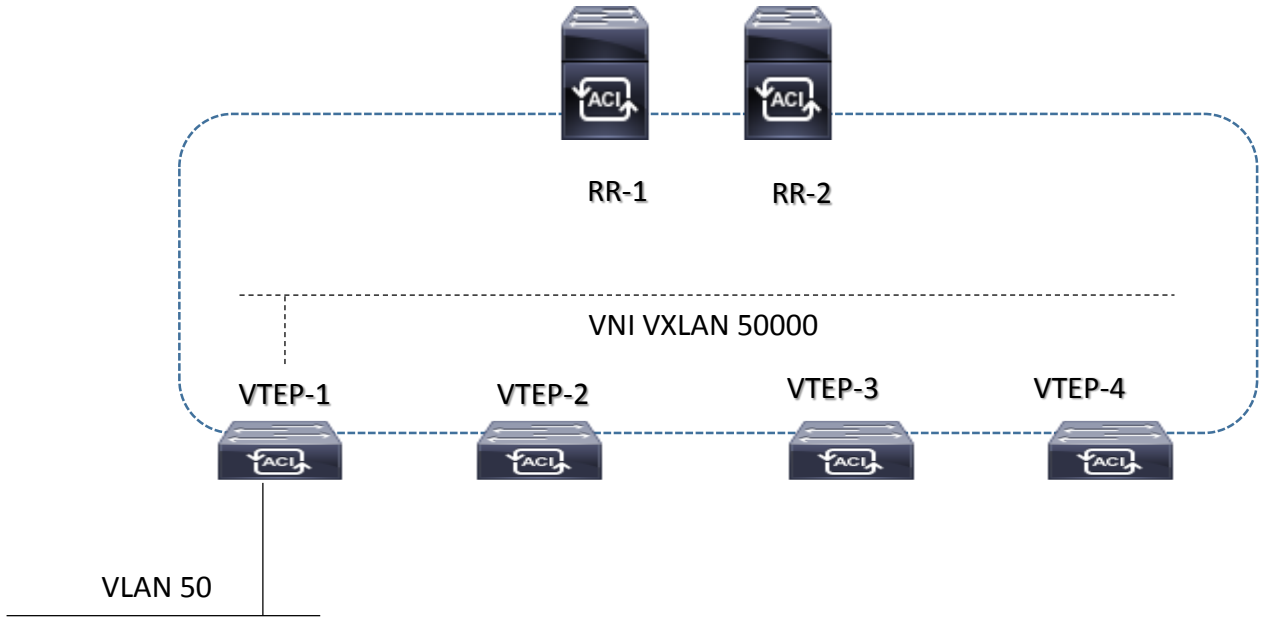
 # Mapping IEEE 802.1q vlan-id TO VXLAN VNI

```
feature  vn-segment-vlan-based
!
vlan 50
  vn-segment 50000
!
evpn
  vni 50000 l2
  rd auto                          → # RD is default calculated as VNI:BGP Router ID
  route-target import auto         → # RT is default calculated as BGP AS:VNI
  route-target export auto
!
interface  nve1
  source-interface loopback0
  host-reachability protocol bgp
  member vni 50000
    mcast-group 239.239.239.10
    suppress-arp
```

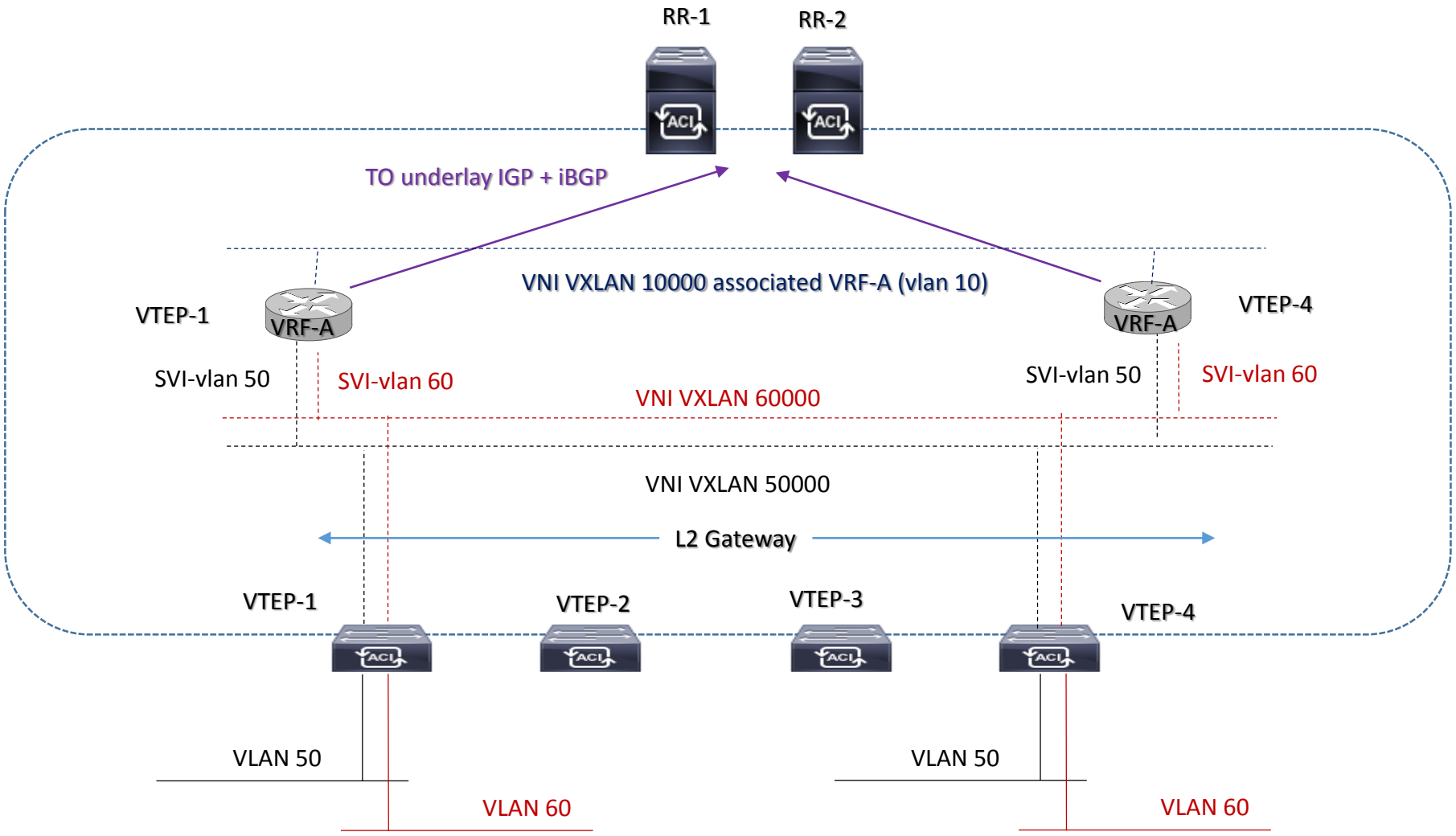
EVPN I-BGP Configurations Routing Resource on VXLAN

 # Define VLAN for VRF routing instances

```
vlan 50
  vn-segment 50000
!
interface  vlan 50
 no shutdown
 mtu 9216
 vrf member VRF-A
 ip forward
!
vrf context VRF-A
 vni 50000
 rd auto
   address-family ipv4unicast
   route-target both auto
   route-target both auto evpn
```



EVPN I-BGP Design Distributed IP Anycast Gateway



Vlan-ID ha significato solo locale al VTEP

EVPN I-BGP Configurations Distributed IP Anycast Gateway

Define VLAN 50 and 60

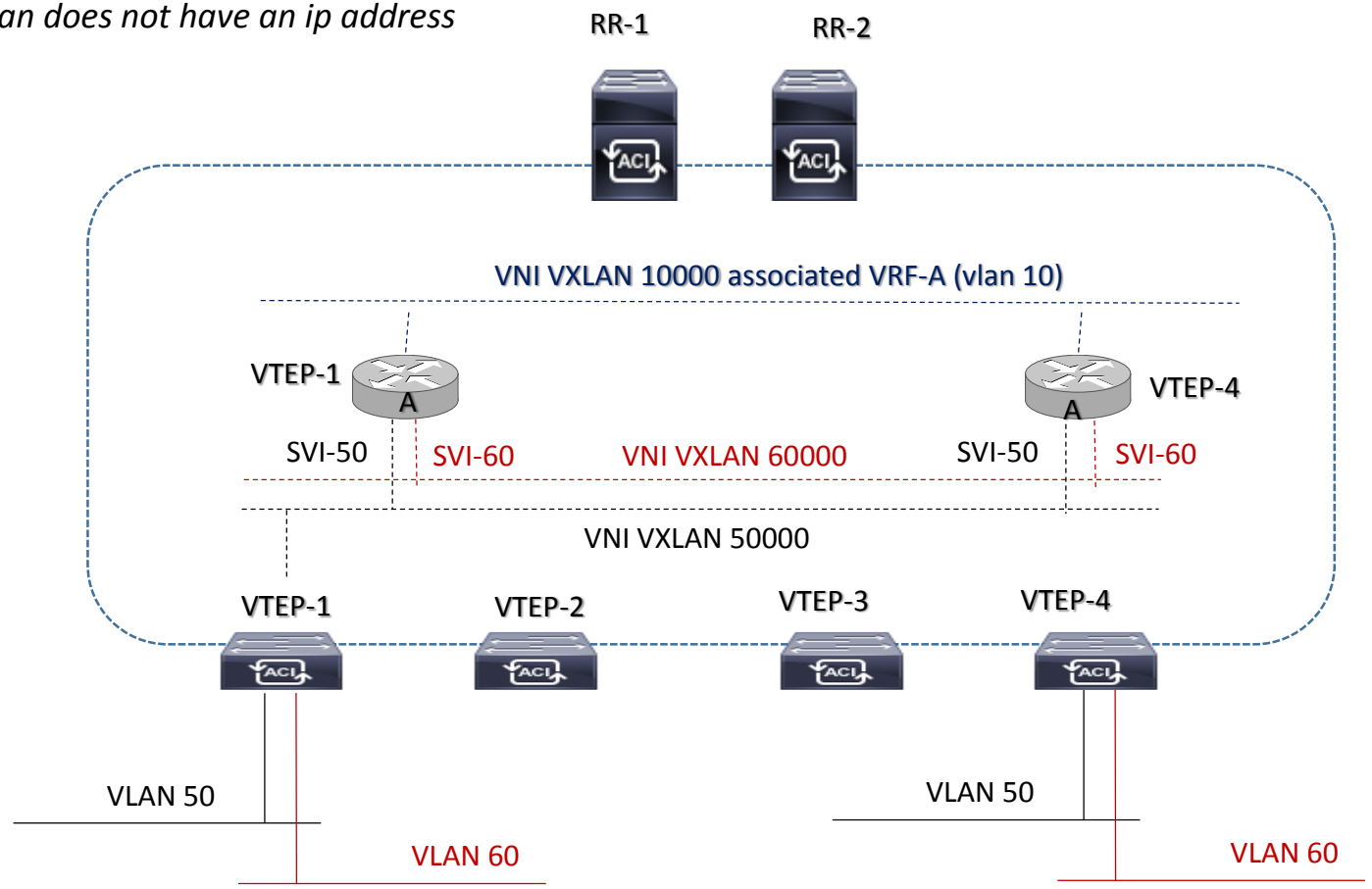
```
features interface-vlan
fabric-forwarding anycast-gateway-mac  < mac-address >
!
vlan 50
  vn-segment 50000
!
vlan 60
  vn-segment 60000
!
interface  vlan 50
 no shutdown
 mtu 9216
 vrf member VRF-A
 ip address 50.50.50.1/24 tag 123
fabric forwarding mode anycast-gateway
!
interface  vlan 60
 no shutdown
 mtu 9216
 vrf member VRF-A
 ip address 60.60.60.1/24 tag 123
fabric forwarding mode anycast-gateway
```

→ *un MAC address per VTEP; tutti i VTEP dovrebbero avere lo stesso MAC Address*

EVPN I-BGP Configurations Routing on VXLAN (1/1)

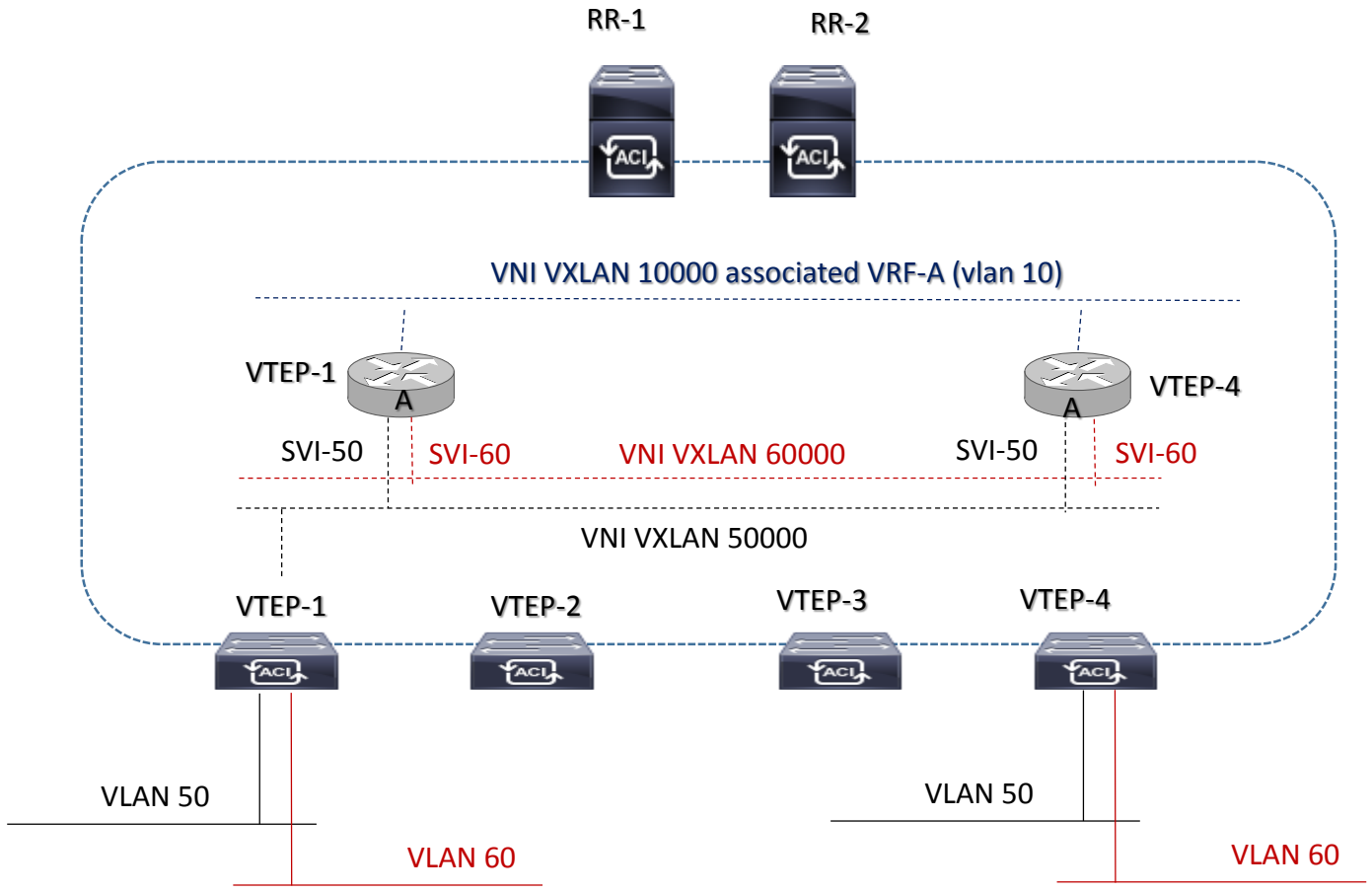
```
vlan 10          → # vlan 10 is used as Layer 3 VNI to route inter-vlan routing
vn-segment 10000
!
interface vlan 10 → # Layer 3 VNI associated interface vlan does not have an ip address
vrf member VRF-A
no shutdown
!
interface  nve1
  source-interface loopback0
  host-reachability protocol bgp
  member vni 50000
    mcast-group 239.239.239.10
    suppress-arp
member vni 10000 associate-vrf
!
member vni 60000
    mcast-group 239.239.239.11
    suppress-arp
member vni 10000 associate-vrf
!
```

segue ./.



EVPN I-BGP Configurations Routing on VXLAN (1/2)

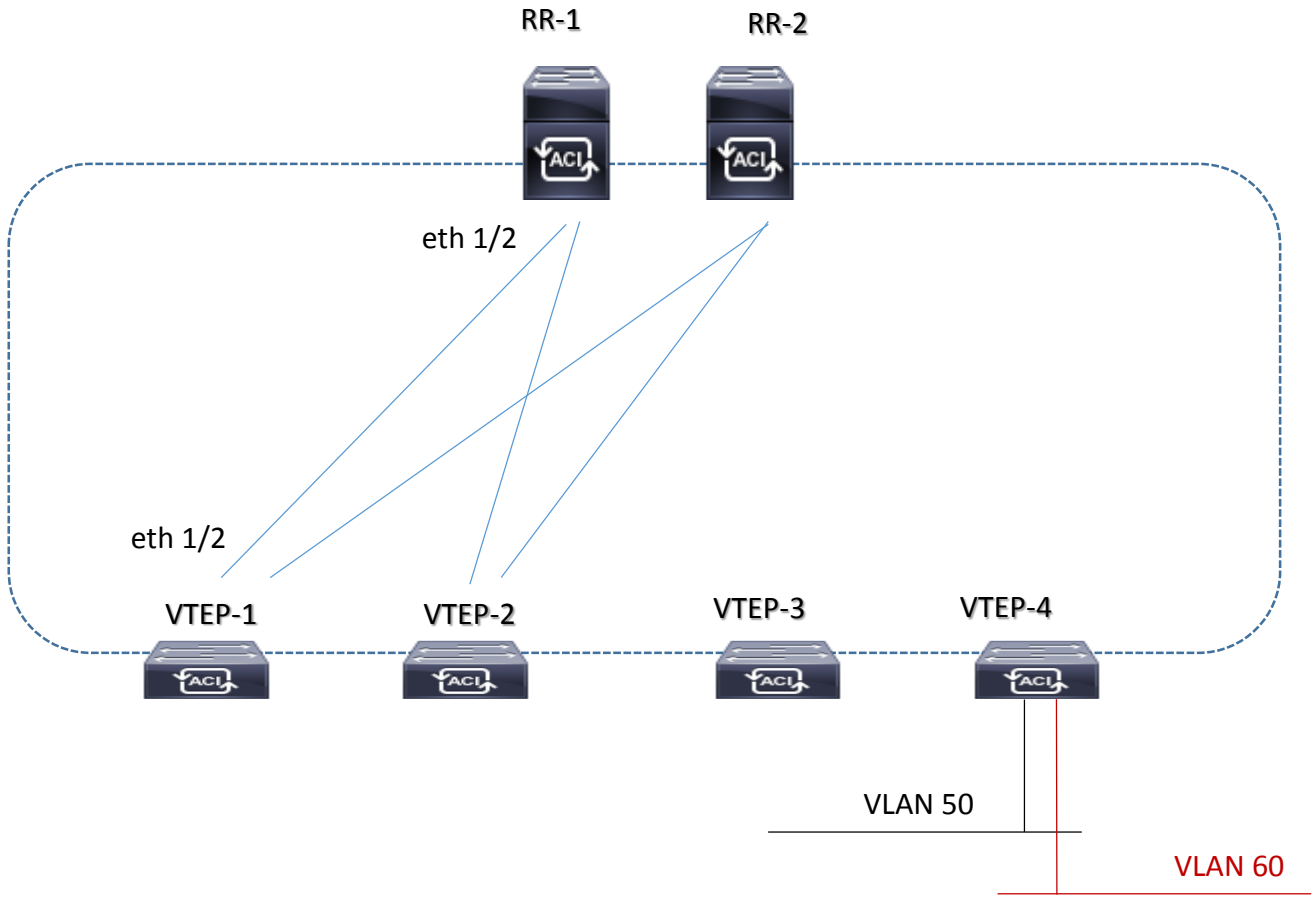
```
route-map RED-SUBNET permit 10
match 123
!
router bgp 65000
 vrf VRF-A
 advertise l2vpn evpn
 redistribuite direct route-map RED-SUBNET
 maximum-path ibgp 2
```



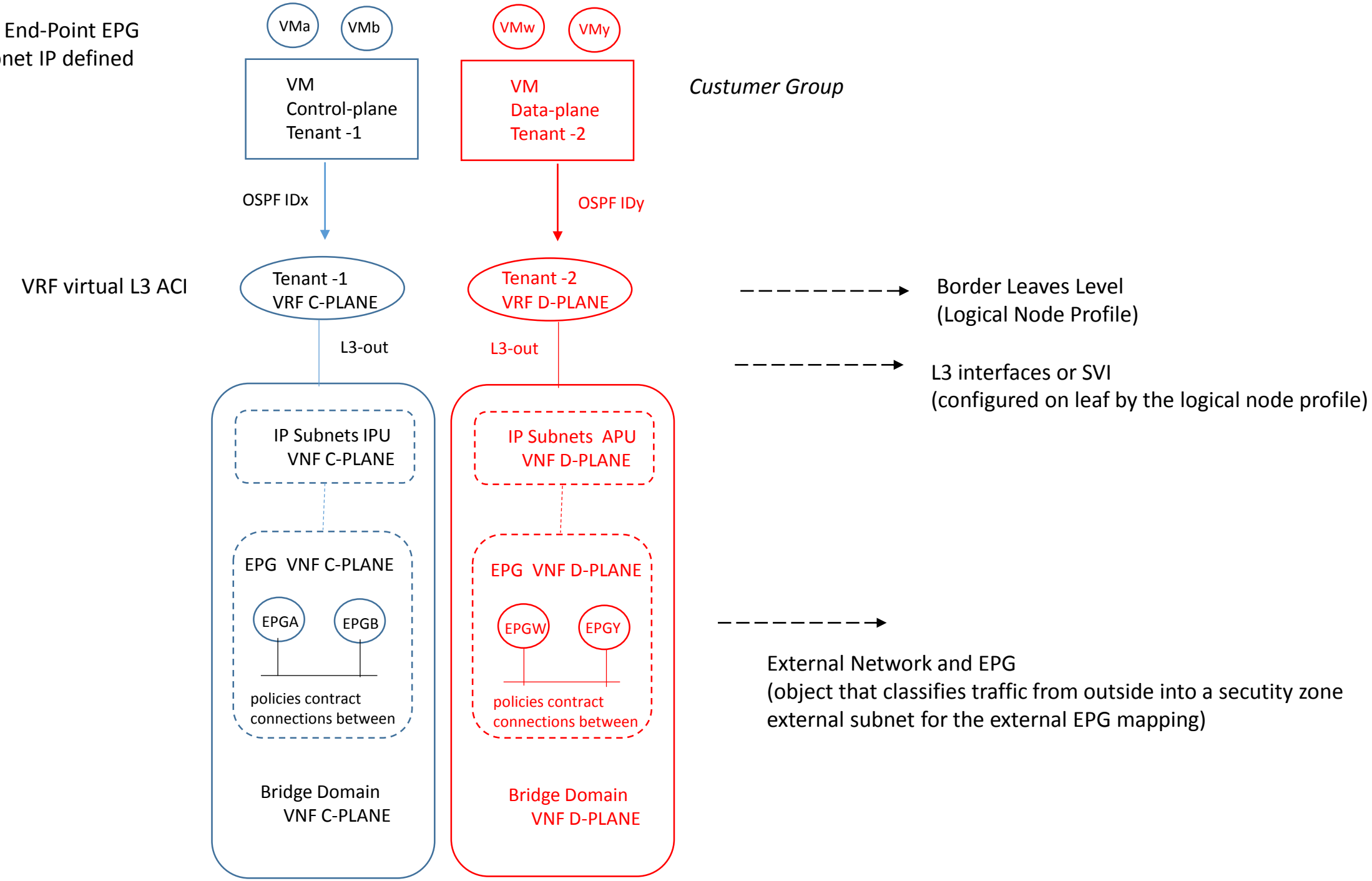
EVPN I-BGP Configurations IGP with OSPF

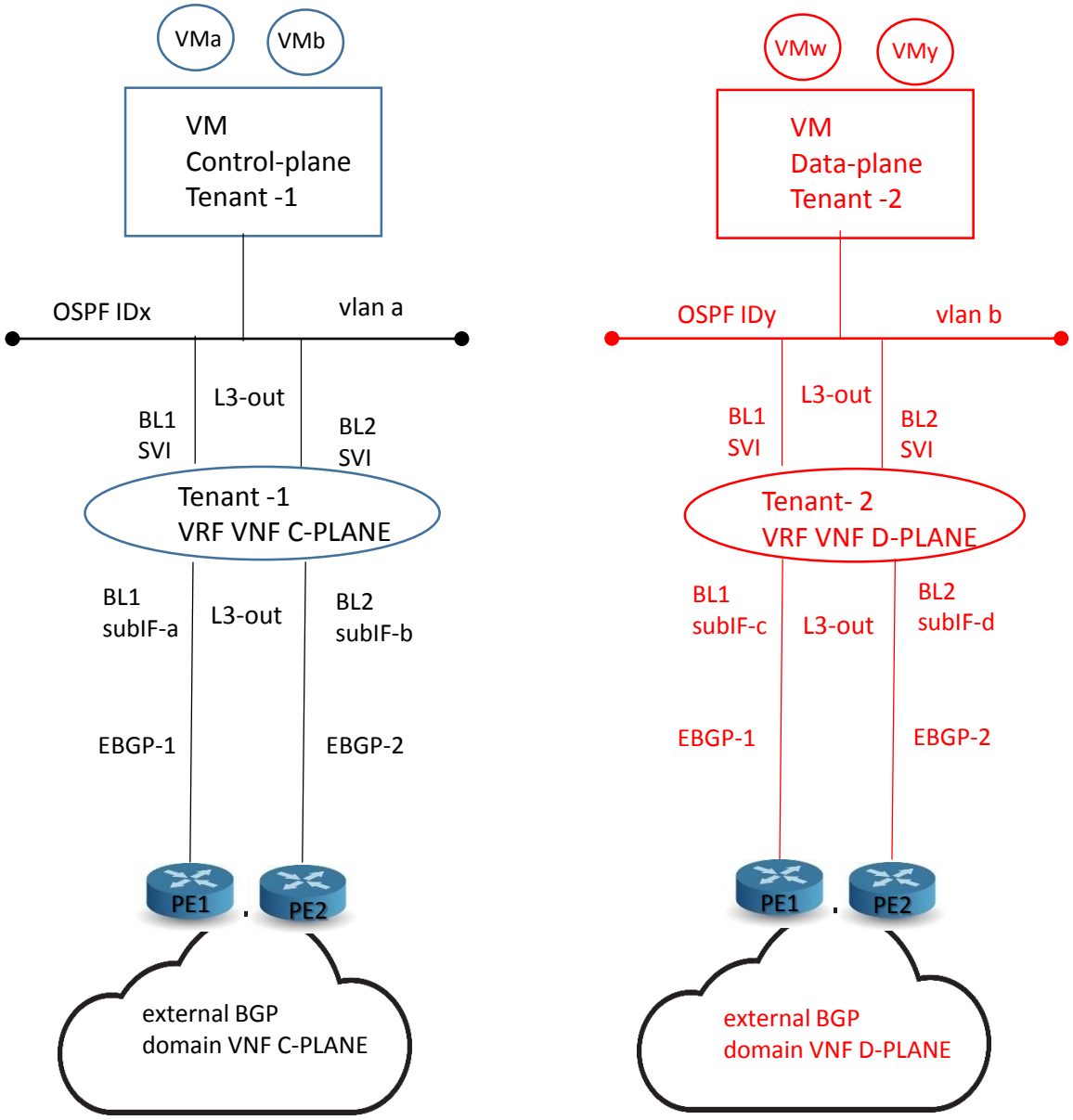
VTEP1:

```
feature ospf
feature pim
!
ip pim rp-address 192.168.1.1 group-list 224.0.0.0/4
ip pim ssm range 232.0.0.0/8
!
interface ethernet 1/2
desscription to-SPINE
no switchport
ip address 10.1.1.2/30
ip route ospf UNDERLAY area  0.0.0.0
ip pim sparse-mode
no shutdown
!
interface loopback 0
description «loopback for BGP»
ip address 192.168.1.10/32
ip route ospf UNDERLAY area  0.0.0.0
ip pim sparse-mode
no shutdown
!
router ospf UNDERLAY
```



External End-Point EPG
with subnet IP defined
in ACI





1) Configure a VRF for each customer Tenant

2) Configure L3 out policy associated with a VRF

2a) define the logical node profile → Border Leaves

2b) logical interface profile → SVI interface on the BL defined by the logical node profile

2c) external network and EPG: object that classifier traffic from the outside into the fabric (security zone)

3) L3-out must be referred by the Bridge Domain where subnet need to be advertised to the outside

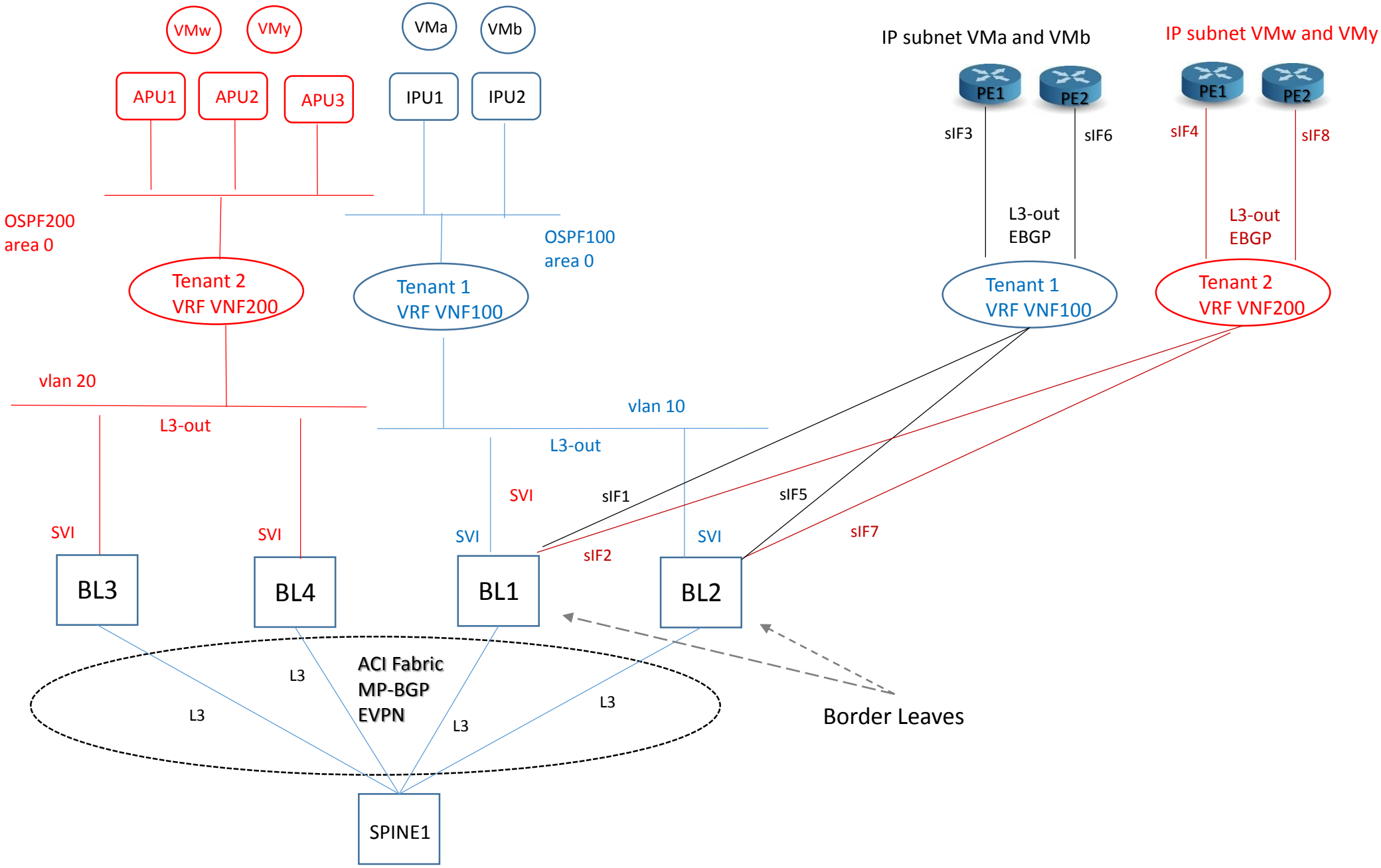
4) L3 out policies provide IP connectivity between a VRF and an external IP netw each L3-out is associated with one VRF instances only.

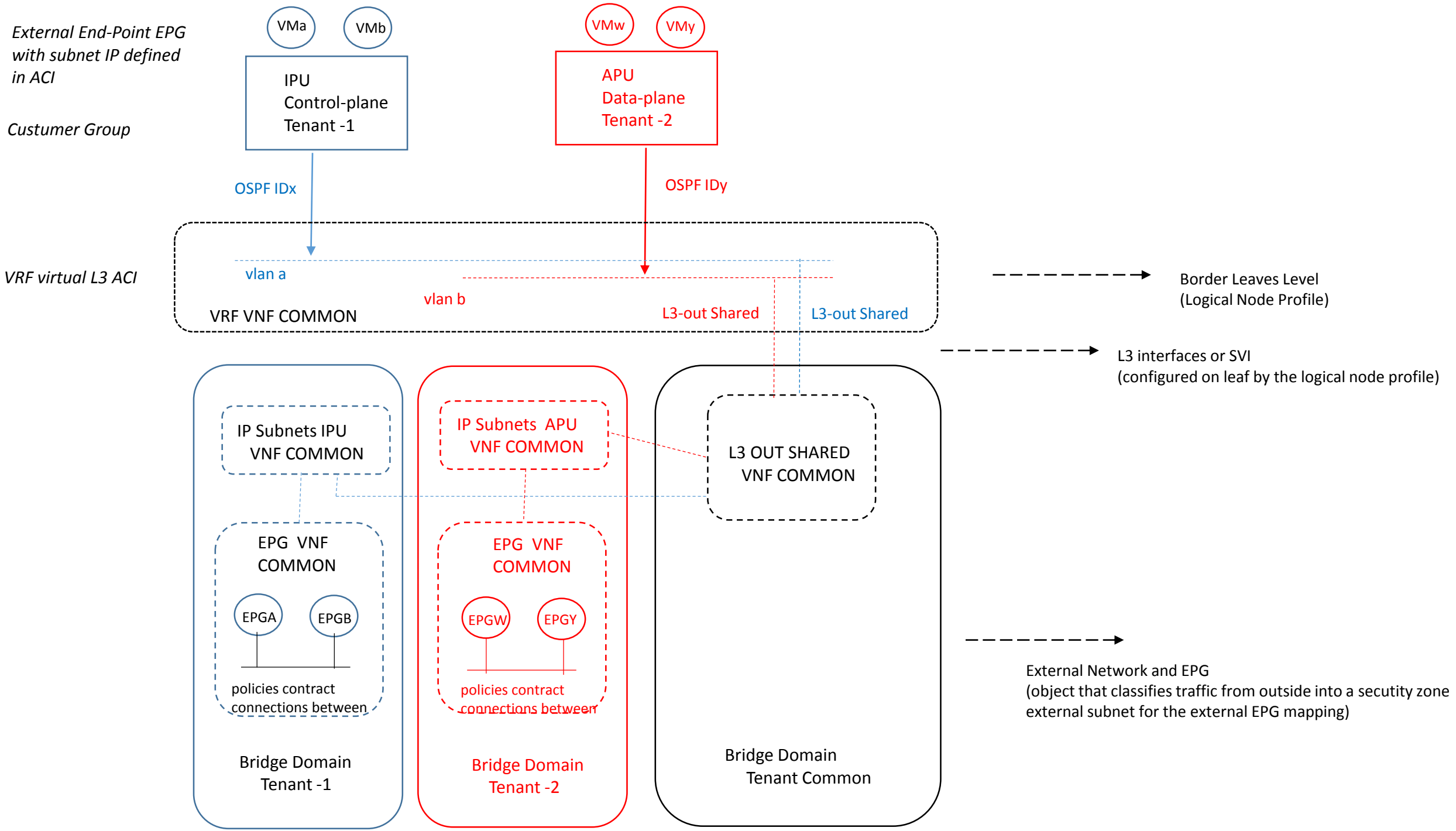
5) For subnet defined in the BD to be announced to the outside router, follow:

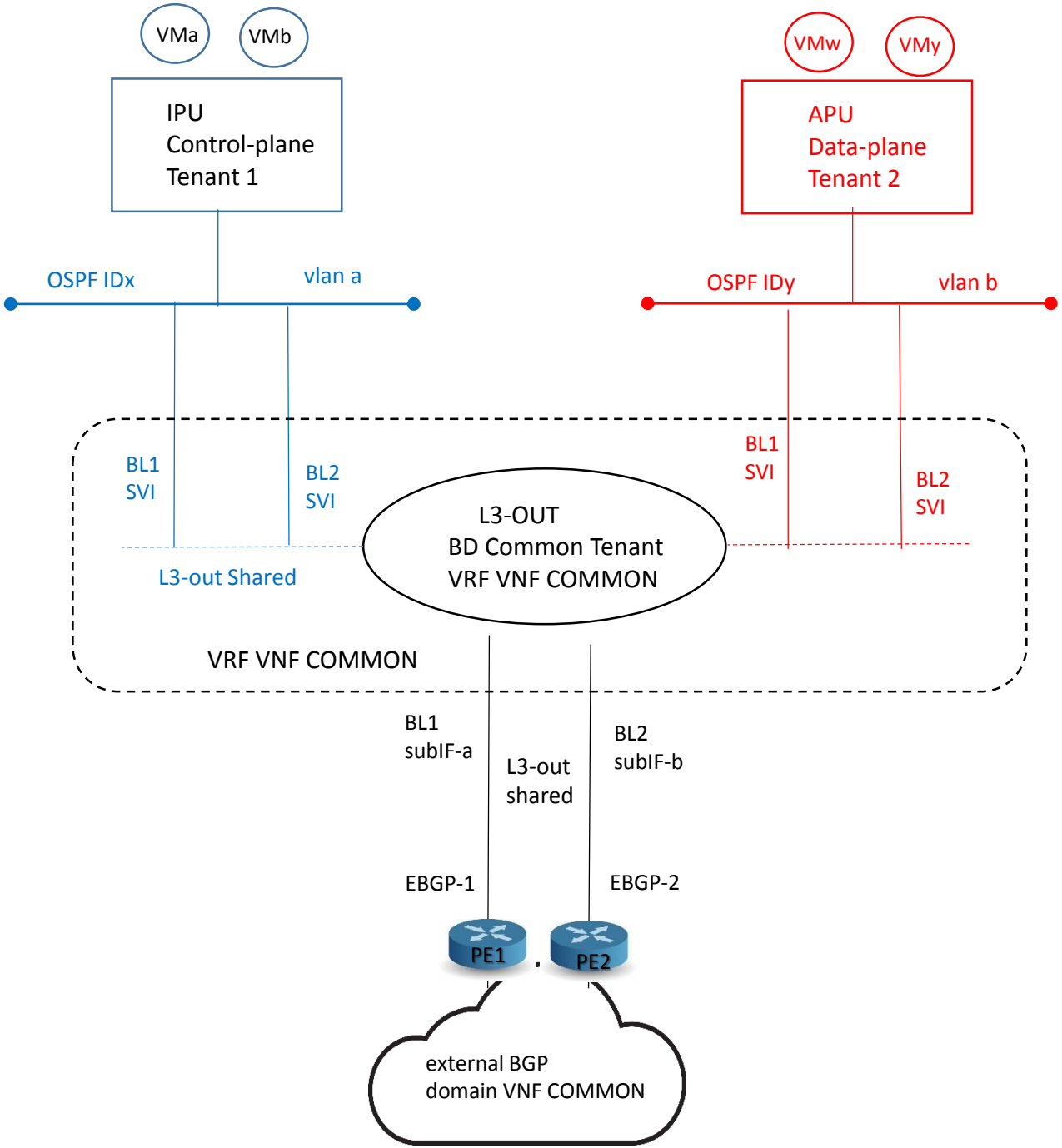
5a) the sunbet need to be defined as advertised externally

5b) the BD must have a relationship with the L3 out connection

5c) a contract must exist between layer 3 external EPG and the EPG associated with BD; if this contract is not in place, the advertisment of the subnet cannot occur.


VMw
VMy
VMa
VMb
APU1
APU2
APU3
IPU1
IPU2
OSPF200
area 0
OSPF100
area 0
Tenant 2
VRF VNF200
Tenant 1
VRF VNF100
vlan 20
L3-out
vlan 10
L3-out
SVI
SVI
SVI
SVI
BL3
BL4
BL1
BL2
ACI Fabric
MP-BGP
EVPN
L3
L3
L3
L3
SPINE1
IP subnet VMa and VMb
IP subnet VMw and VMy
PE1
PE2
PE1
PE2
sIF3
sIF6
sIF4
sIF8
L3-out
EBGP
L3-out
EBGP
Tenant 1
VRF VNF100
Tenant 2
VRF VNF200
sIF1
sIF2
sIF5
sIF7
Border Leaves


External End-Point EPG
with subnet IP defined
in ACI
Custumer Group
VMa
VMb
IPU
Control-plane
Tenant -1
VMw
VMy
APU
Data-plane
Tenant -2
OSPF IDx
OSPF IDy
VRF virtual L3 ACI
vlan a
vlan b
VRF VNF COMMON
L3-out Shared
L3-out Shared
Border Leaves Level
(Logical Node Profile)
L3 interfaces or SVI
(configured on leaf by the logical node profile)
IP Subnets IPU
VNF COMMON
IP Subnets APU
VNF COMMON
L3 OUT SHARED
VNF COMMON
EPG VNF
COMMON
EPG VNF
COMMON
EPGA
EPGB
EPGW
EPGY
policies contract
connections between
policies contract
connections between
Bridge Domain
Tenant -1
Bridge Domain
Tenant -2
Bridge Domain
Tenant Common
External Network and EPG
(object that classifies traffic from outside into a secutity zone
external subnet for the external EPG mapping)



1) Configure a VRF under the common Tenant

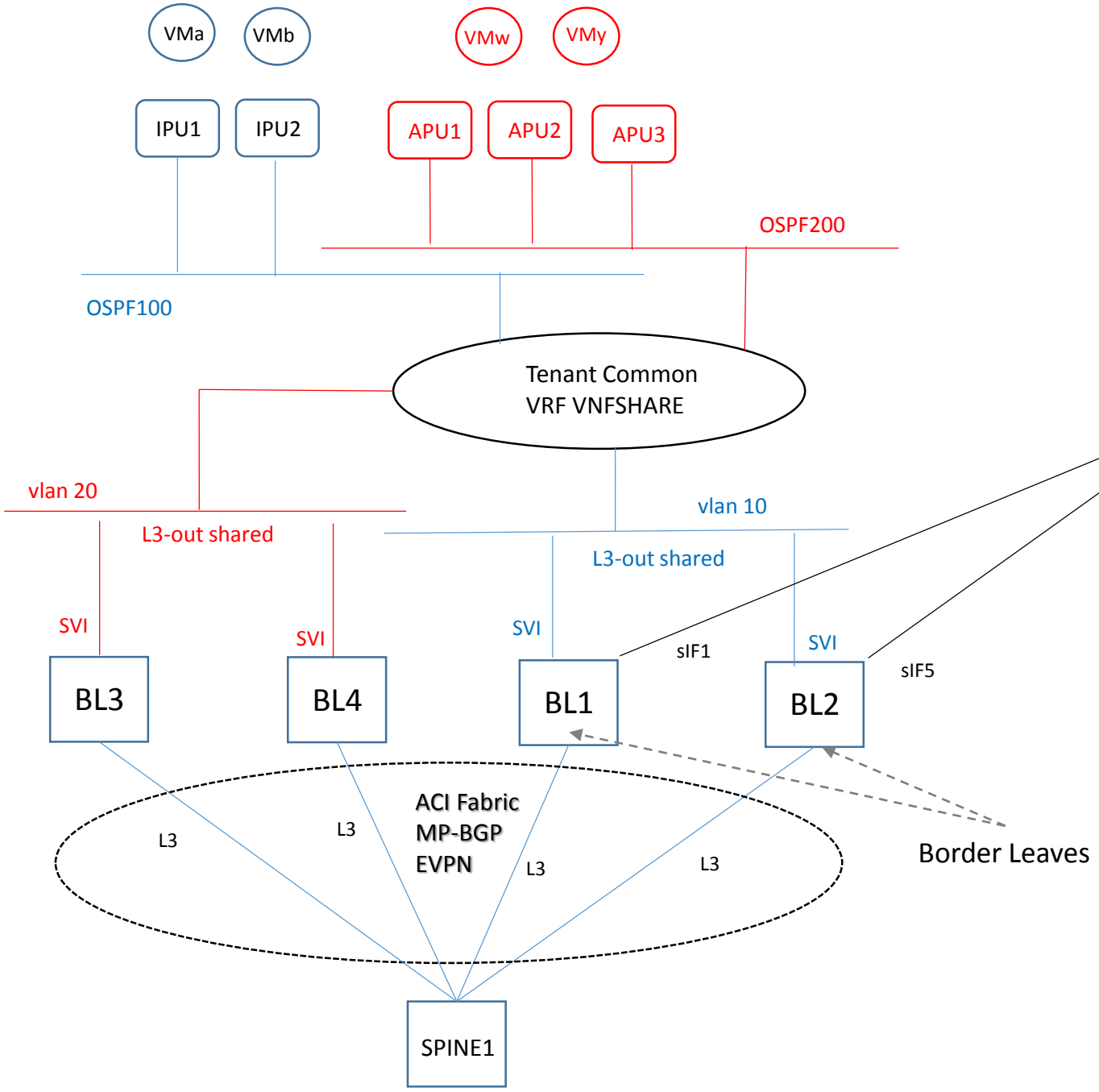
2) Configure L3 out connection under the common Tenant and associate it with the VRF instances

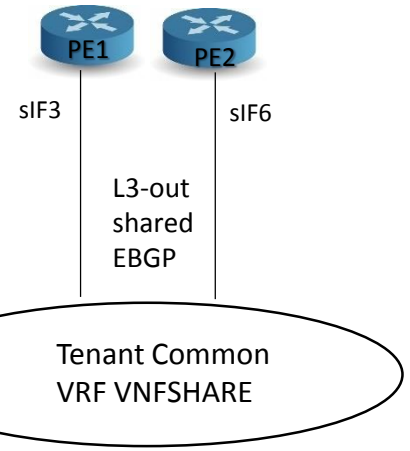
3) Configure a Bridge Domain and subnet under each customer Tenant

4) Associate the Bridge Domain with a VRF in the common Tenant and the L3-out connection

5) Under each Tenant configure EPG and associate the EPG with a BD in the Tenant itself

6) Configure contracts and application profiles under each Tenant





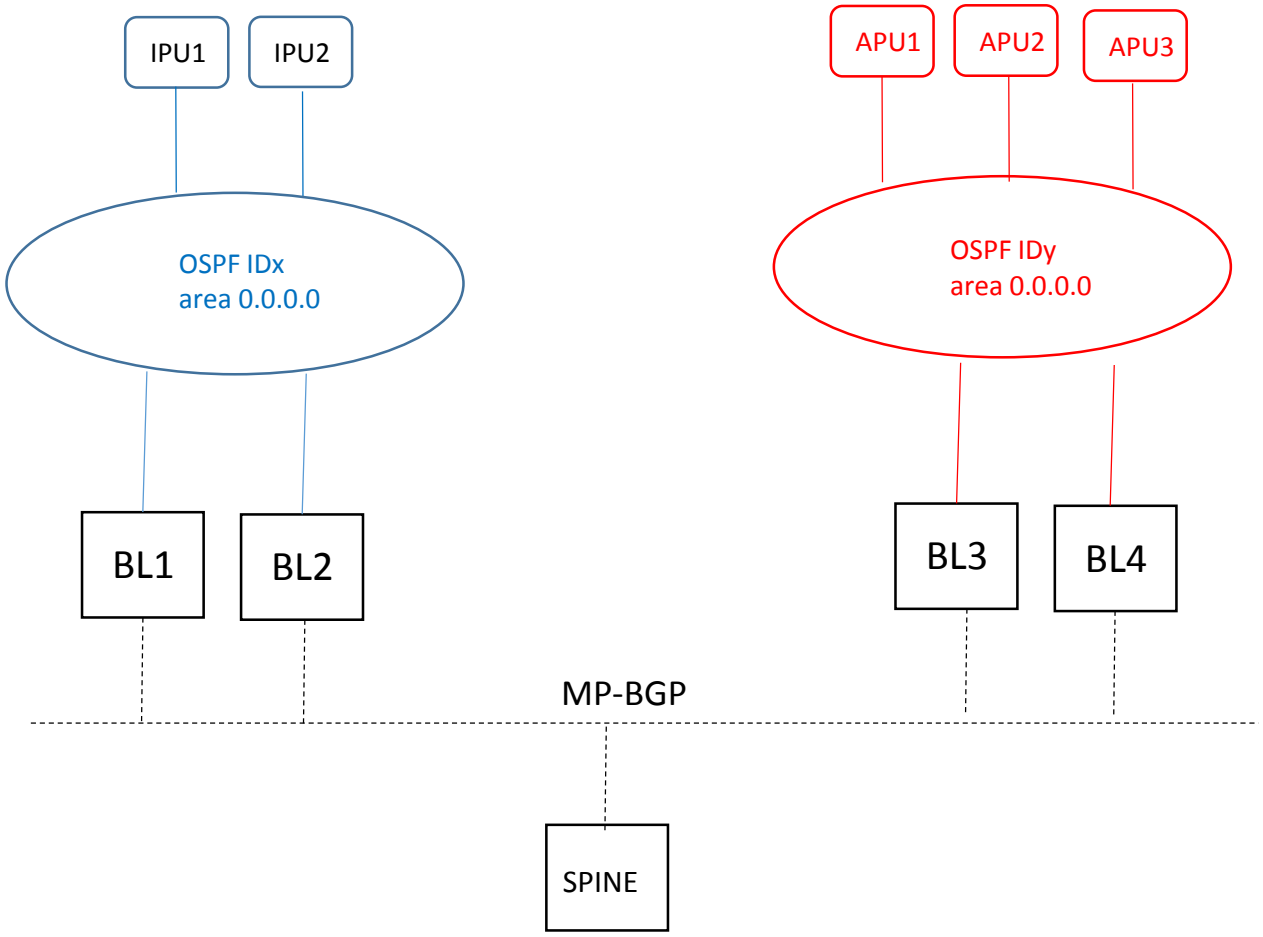
L3-out
connection usual

marked with:

Shared Route Control Subnet: mean that the network if learned from the outside through the VRF can be leaked to other VRF (via contract with the external EPG);

Shared Security Import Subnet: mean which subnet learned from a shared VRF belong to this external EPG for the purpose of contract filtering when establishing a cross-VRF Contract)

OSPF areas on different Border Leaf are different OSPF areas



ACI border leaf running OSPF are always AS boundary (ASBR)

all external routes learned in OSPF are redistribuite into MP-BGP

MP-BGP routes are redistribuite into OSPF as external type2

OSPF areas on different border leaf (pairs BL) are different OSPF areas, even if ID match

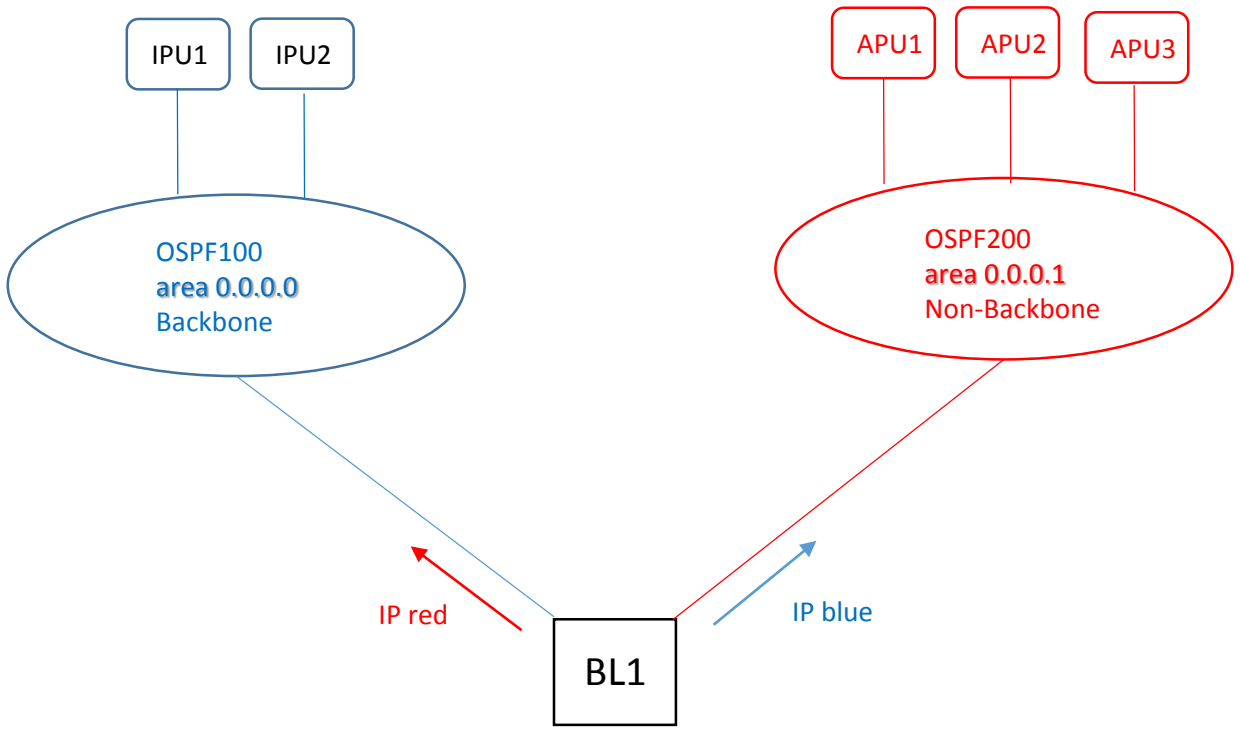
IPv4 and IPv6 support

ACI Border Leaf follow OSPF rules which as:

multiple areas but NO backbone (both areas)area the routes are not advertised between areas;

No backbone area and backbone area are advertised between them

OSPF areas on the same Border Leaf need different area type to be advertised



ACI border leaf running OSPF are always AS boundary (ASBR)

all external routes learned in OSPF are redistribuite into MP-BGP

MP-BGP routes are redistribuite into OSPF as external type2

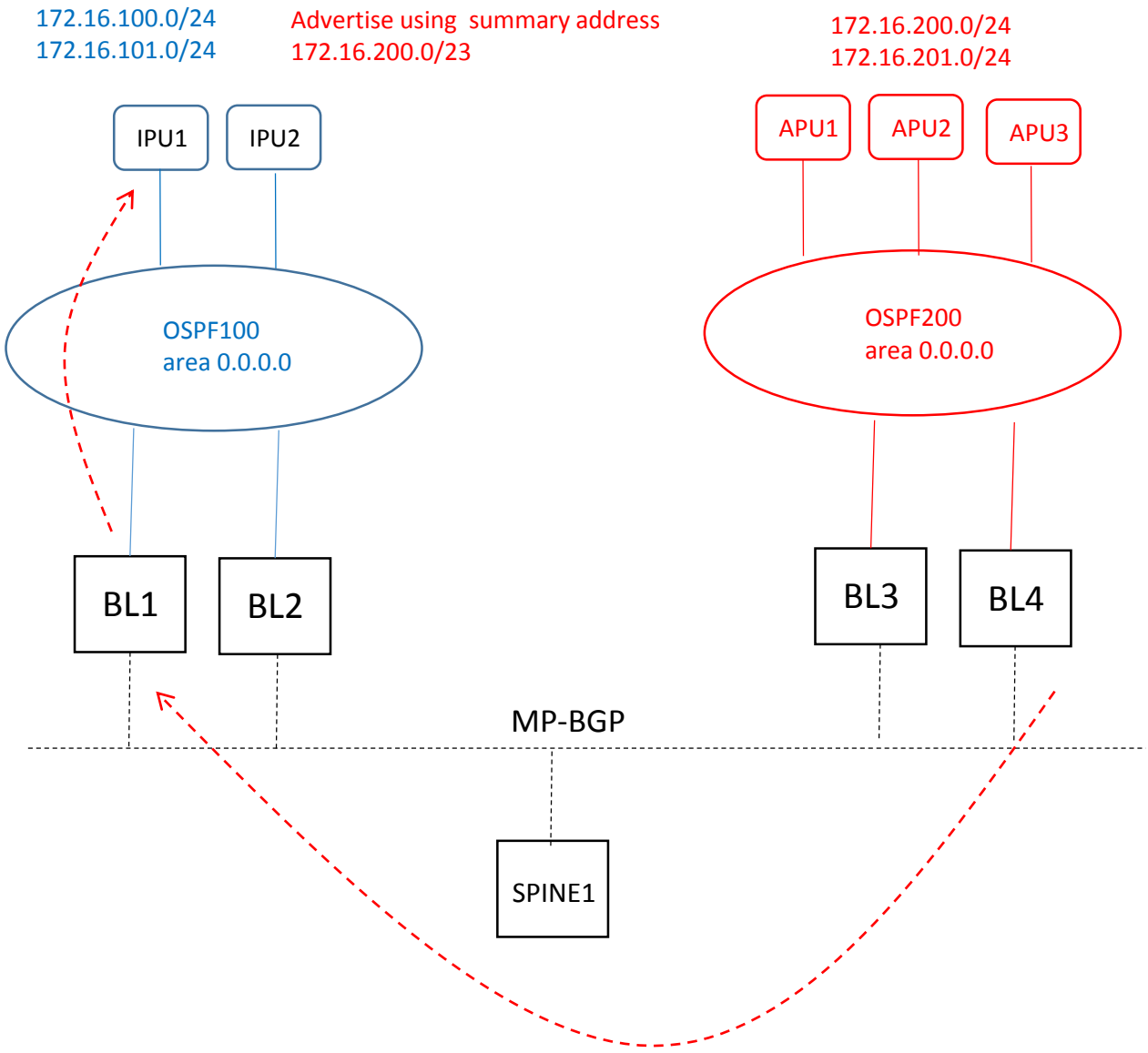
OSPF areas on different border leaf (pairs BL) are different OSPF areas, even if ID match

IPv4 and IPv6 support

ACI Border Leaf follow OSPF rules which as:

multiple areas but NO backbone (both areas)area the routes are not advertised between areas;
No backbone are and backbone area are advertised between them

OSPF summarization rules



Two otions are available with ACI:

External route summarization (equivalent to the summary address config)

Inter-area summarization (equivalent to the area range config)

When Tenant routes are injected into OSPF, ACI Leaf where L3-out connection resides is acting as an ASBR; in this case the summary address config (thats is external route sumarization) should be used.

For scenario where there are two L3-out connection and each using a different area and attached to the same border leaf switch, the area range config will be used to summarize.