

ACI Layer 4-7 Policy-Based Redirect (PBR) Deep Dive and Tips

cisco Live !

Minako Higuchi
Technical Marketing Engineer

Cisco Webex App

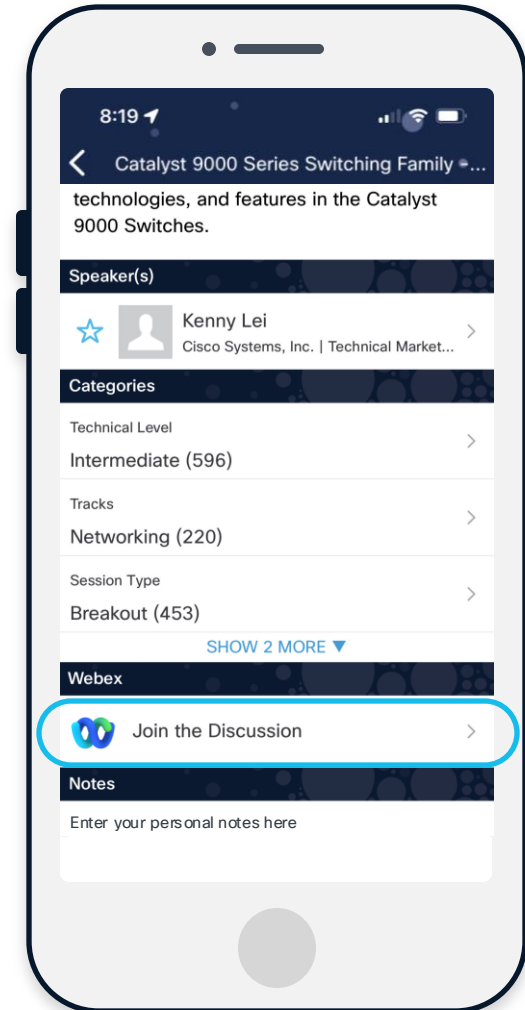
Questions?

Use Cisco Webex App to chat with the speaker after the session

How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 13, 2025.



Session Objectives

- At the end of the session, the participants should be able to:
 - Understand ACI PBR use cases.
 - Understand how ACI PBR works.
 - Understand design considerations.
 - Understand ACI PBR for Multi-Site (New configuration workflow)
- Initial assumption:
 - The audience already has a good knowledge of ACI main concepts: VRF, BD, EPG, ESG, L3Out, Contract, Multi-Pod, Multi-Site, Remote Leaf etc.
- Note: This session uses ESGs mainly, but the PBR features are applicable to EPGs and uSeg EPGs.

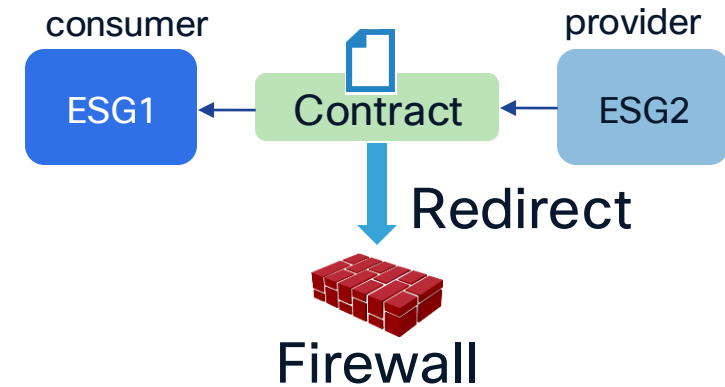
Agenda

- 01 ACI PBR Use cases**
- 02 PBR Forwarding and zoning-rules**
- 03 FAQs**
- 04 Multi-location DC design**

ACI PBR Use Cases

PBR (redirect) is one of the contract actions!

- Permit
- Deny
- **Redirect**
- Copy

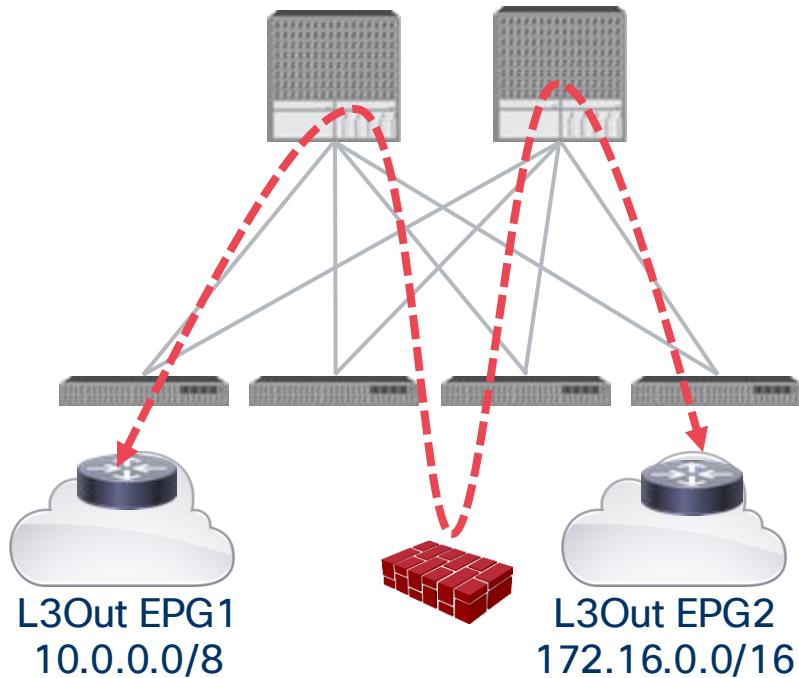


Where can we use PBR?

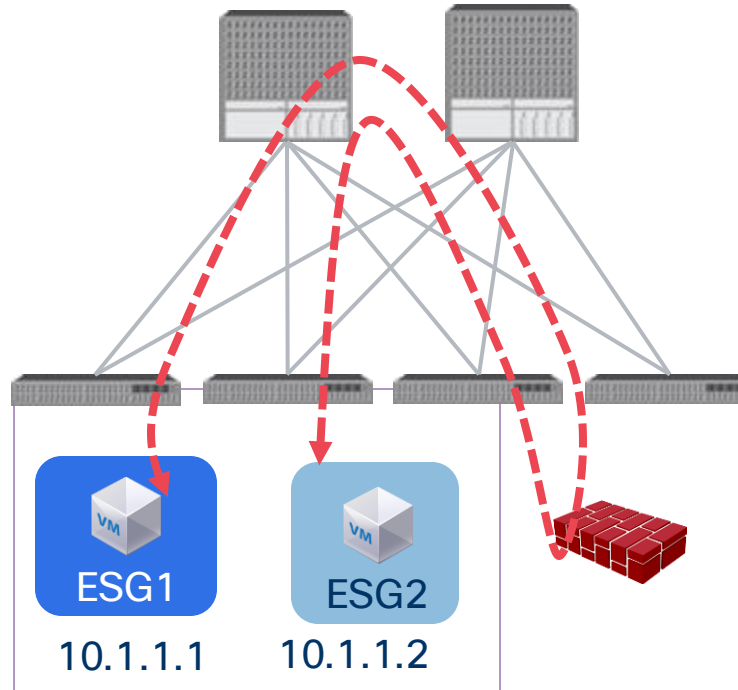


PBR is a contract action. It's based on source, destination EPG/ESG and filter matching.

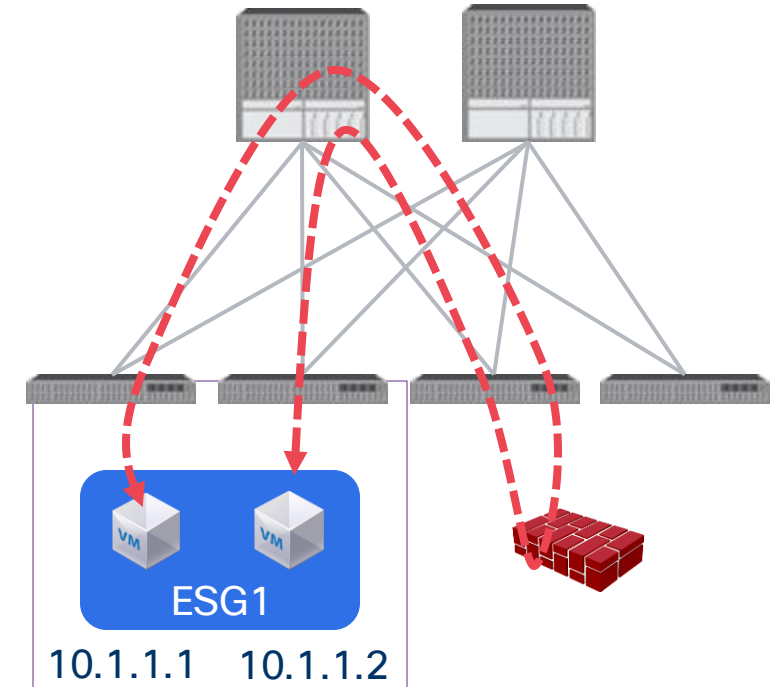
- Between EPGs or ESGs.
- Between L3Out EPGs.



- Between EPGs or ESGs in the same subnet.



- Between endpoints in the same EPG or ESG.



PBR use cases

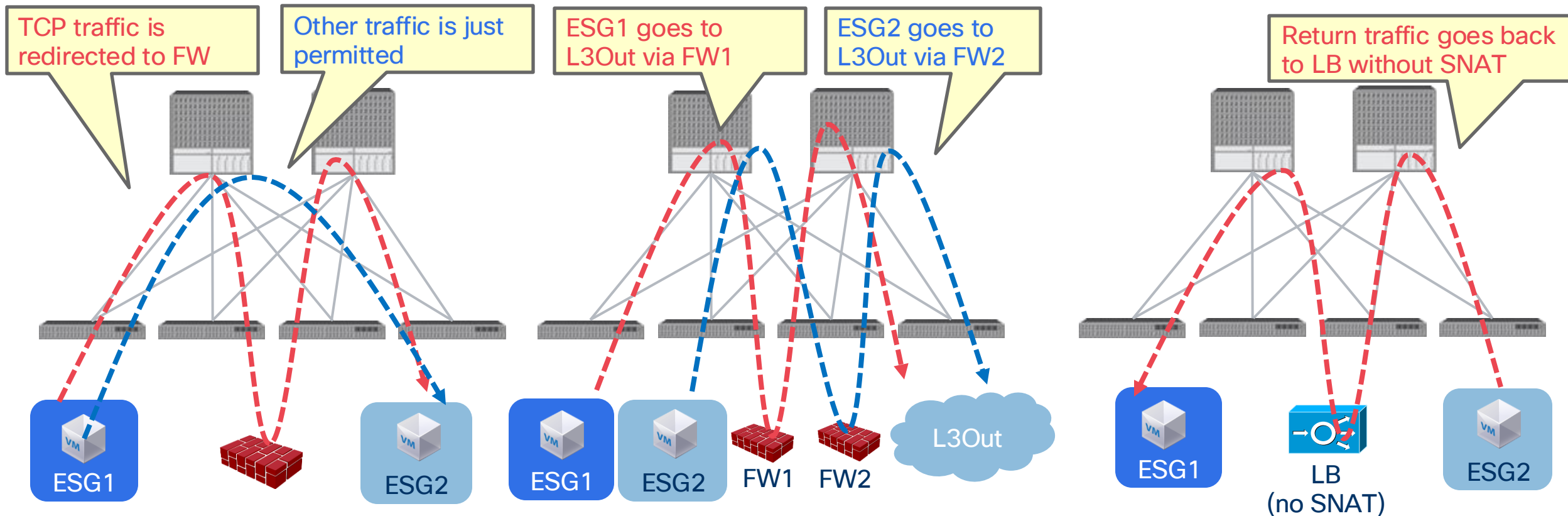


PBR can be applied to each direction

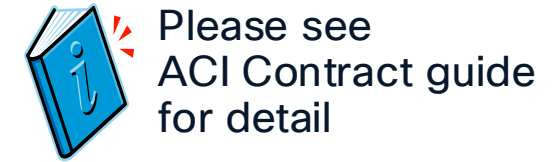
- Inspect specific traffic

- Use different Firewall

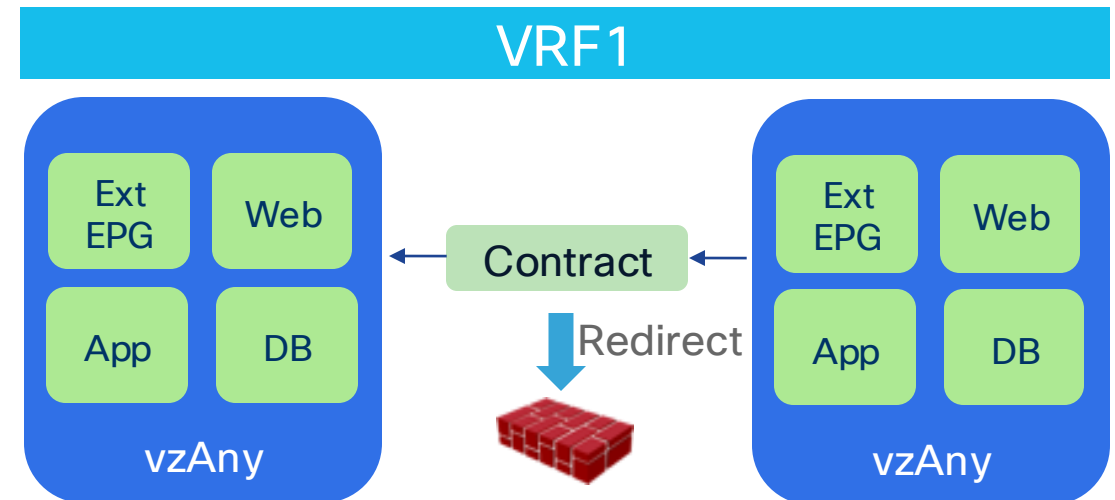
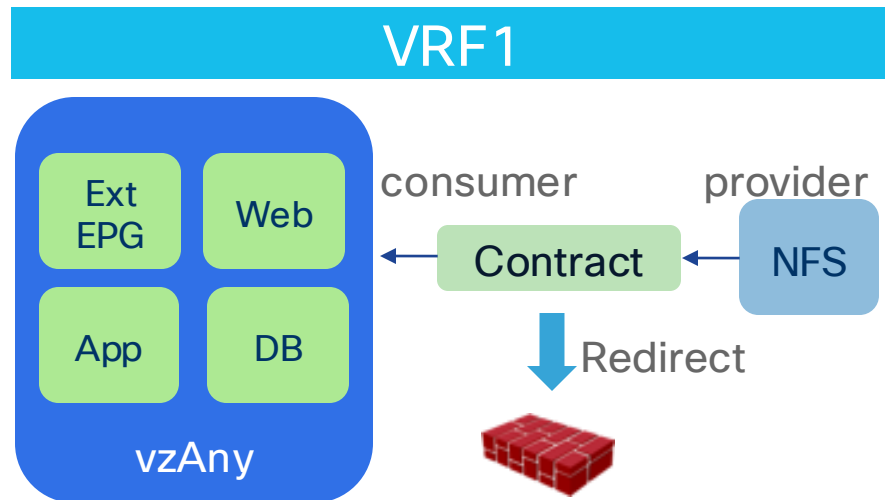
- LB without SNAT (uni-directional PBR)



Important note



- ACI must be Layer 3. (L2Out EPG is not supported)
- VRF must be in enforced mode. (PBR cannot be used in a VRF with unenforced mode)
 - If you want common permit or redirect rules in the VRF, you can use vzAny (All EPGs and ESGs in a VRF)
 - If you don't need contract enforcement for specific EPGs/ESGs in the VRF, you can still use Preferred Group.



PBR

Forwarding and zoning-rules

Zoning-rules (1-node Service Graph)

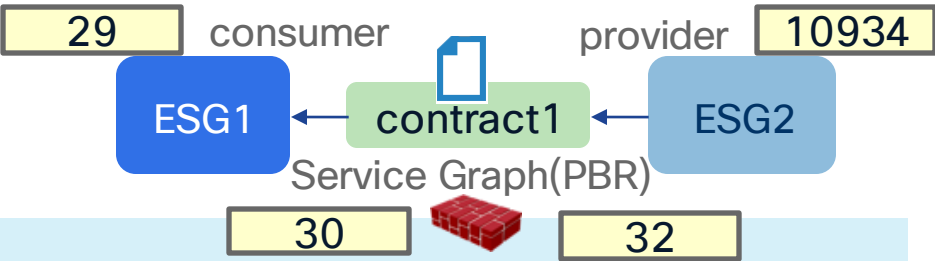
- Without PBR (permit action)



```
Pod1-Leaf1# show zoning-rule scope 2195459
```

Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
<snip>									
4157	29	10934	14	bi-dir	enabled	2195459	tenant1:contract1	permit	fully_qual(7)
4144	10934	29	14	uni-dir-ignore	enabled	2195459	tenant1:contract1	permit	fully_qual(7)

- With PBR (Service Graph)



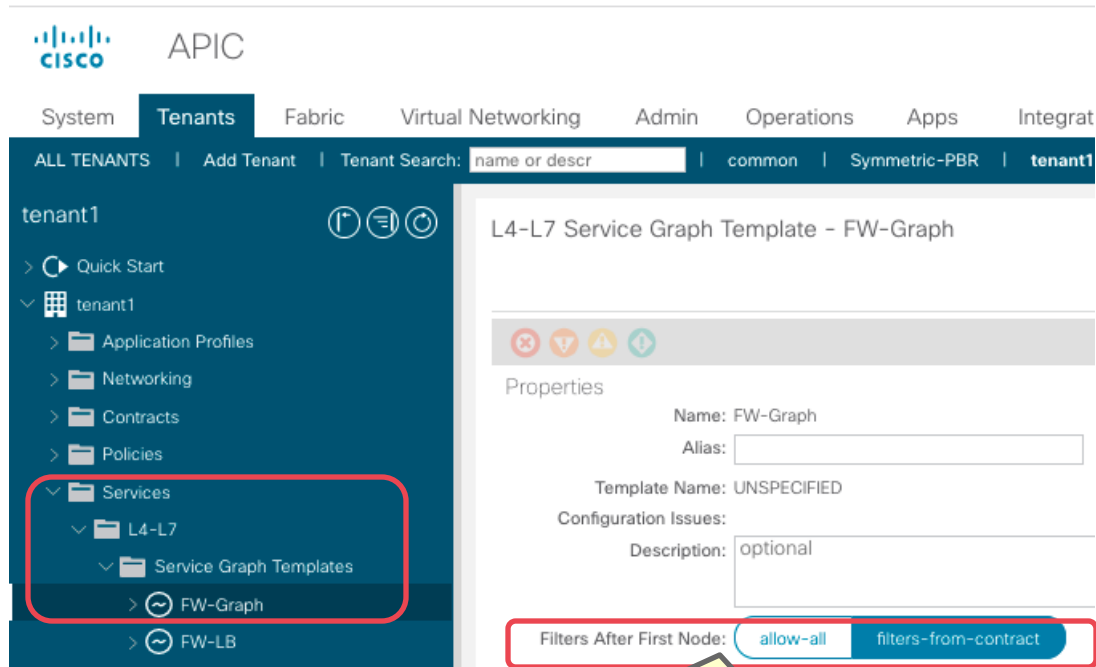
```
Pod1-Leaf1# show zoning-rule scope 2195459
```

Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
<snip>									
4144	29	10934	14	bi-dir	enabled	2195459		redir(destgrp-11)	fully_qual(7)
4157	10934	29	14	uni-dir-ignore	enabled	2195459		redir(destgrp-12)	fully_qual(7)
4140	32	10934	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4136	30	29	14	uni-dir	enabled	2195459		permit	fully_qual(7)

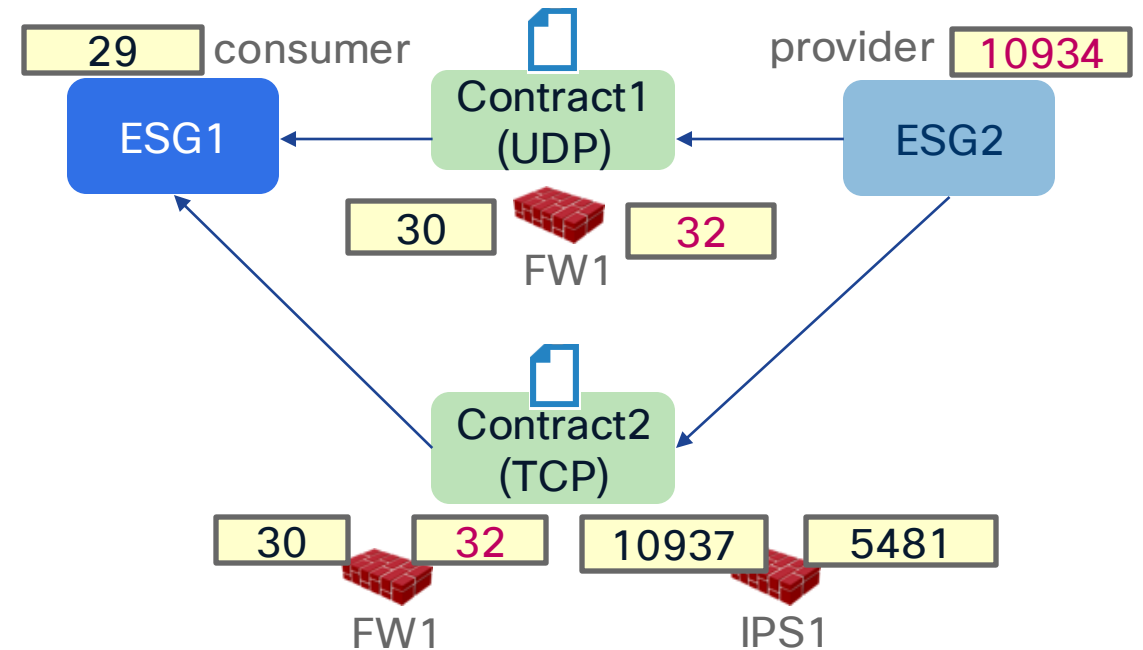
By default, unspecified default filter (any) is used for a zoning-rule entry without the consumer EPG.

Filter-from-contract

- To use the specific filter in the contract, “filters-from-contract” needs to be checked.
- Use case: use a different forwarding action based on the filter.



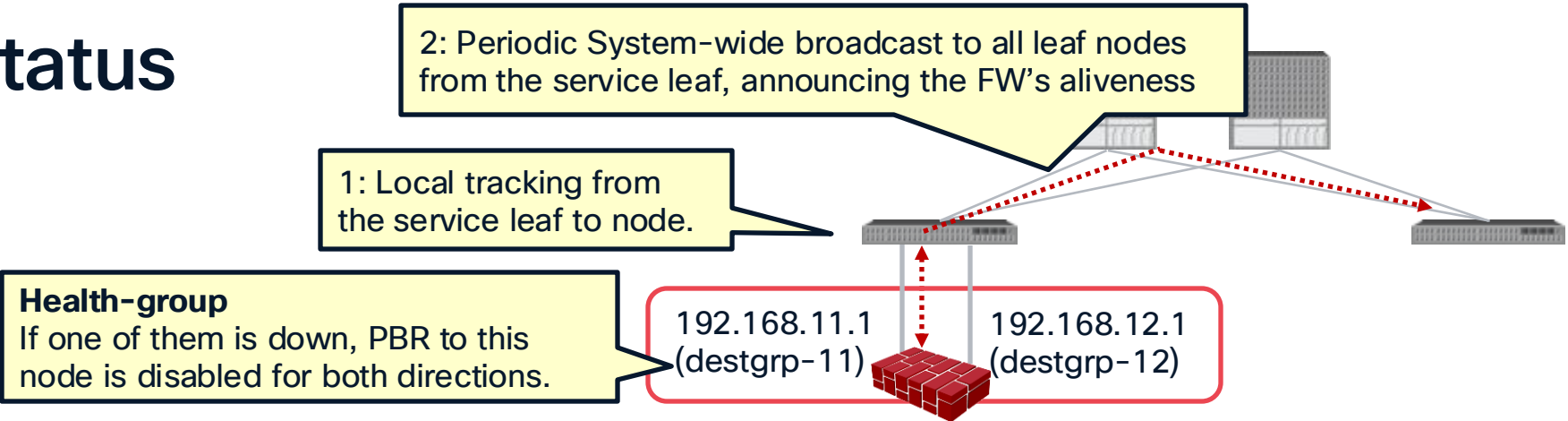
Default is “allow-all”



By default, forwarding actions are duplicated.

- 32-to-10934: permit (contract1 with UDP)
- 32-to-10934: redirect to IPS1 (contract2 with TCP)

PBR destination status



```
Pod1-Leaf1# show service redir info
=====
LEGEND
TL: Threshold(Low)   |  TH: Threshold(High) |  HP: HashProfile   |  HG: HealthGrp   |  BAC: Backup-Dest |  TRA: Tracking   |  RES: Resiliency
=====

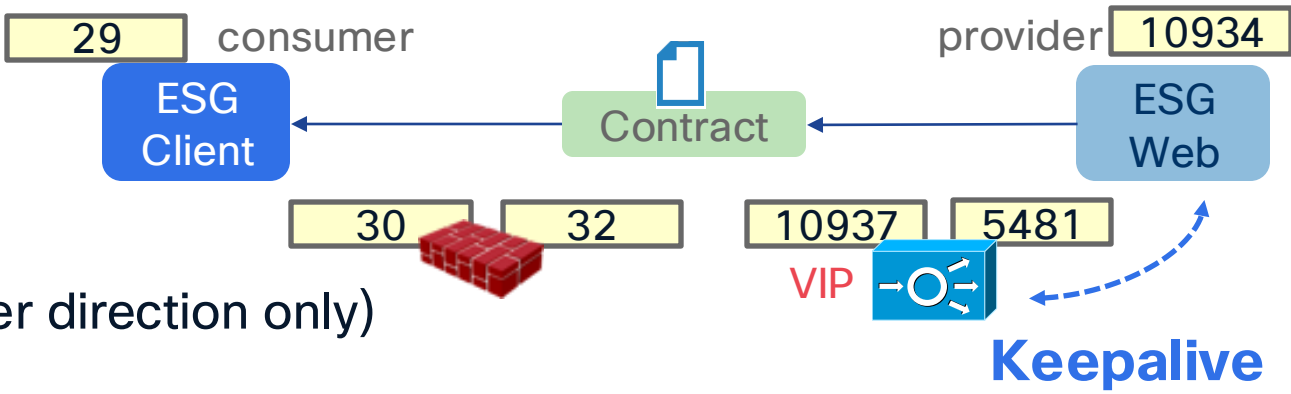
List of Dest Groups
GrpID Name          destination                                     HG-name          BAC  operSt  operStQual  TL  TH  HP  TRAC  RES
=====
11  destgrp-11      dest-[192.168.11.1]-[vxlan-2195459]  tenant1::HG1     N    enabled no-oper-grp  0   0   sym yes  no
12  destgrp-12      dest-[192.168.12.1]-[vxlan-2195459]  tenant1::HG1     N    enabled no-oper-grp  0   0   sym yes  no

List of destinations
Name                bdVnid          vMac          vrf          operSt  operStQual  HG-name
=====
dest-[192.168.11.1]-[vxlan-2195459]  vxlan-16678782  00:50:56:AF:6C:16  tenant1:VRF1  enabled  no-oper-dest  tenant1::HG1
dest-[192.168.12.1]-[vxlan-2195459]  vxlan-16121790  00:50:56:AF:DF:55  tenant1:VRF1  enabled  no-oper-dest  tenant1::HG1

List of Health Groups
HG-Name            HG-OperSt  HG-Dest                                     HG-Dest-OperSt
=====
tenant1::HG1       enabled    dest-[192.168.11.1]-[vxlan-2195459]]      up
                  enabled    dest-[192.168.12.1]-[vxlan-2195459]]      up
```


Zoning-rules (2-nodes Service Graph)

- With Service Graph (PBR)
 - First node: FW (PBR for both directions)
 - Second node: LB (PBR for provider to consumer direction only)



- Consumer to provider direction
- Provider to consumer direction

Pod1-Leaf1# show zoning-rule scope 2195459

Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
4195	29	10937	14	bi-dir	enabled	2195459		redir(destgrp-11)	fully_qual(7)
4196	32	10937	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4193	5481	10934	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4198	10934	29	14	uni-dir	enabled	2195459		redir(destgrp-17)	fully_qual(7)
4181	10937	29	14	uni-dir-ignore	enabled	2195459		redir(destgrp-12)	fully_qual(7)
4194	30	29	14	uni-dir	enabled	2195459		permit	fully_qual(7)

To permit traffic from the provider EPG to the LB (10934 to 5481), Direct Connect option must be enabled.

Direct Connect (False by default)



Direct Connect must be “True” for communication between the consumer/provider endpoint and the PBR destination.

- Tenant > Services > L4-L7 > Service Graph templates > Service Graph_NAME > Policy

APIC

System | **Tenants** | Fabric | Virtual Networking | Admin | Operations | Apps | Integrations

ALL TENANTS | Add Tenant | Tenant Search: name or descr | common | Symmetric-PBR | **tenant1** | PBR | floating

tenant1

- Quick Start
- tenant1
 - Application Profiles
 - Networking
 - Contracts
 - Policies
 - Services**
 - L4-L7**
 - Service Graph Templates**
 - FW-Graph**
 - FW-LB**
 - Router configurations
 - Devices
 - Imported Devices
 - Devices Selection Policies
 - Deployed Graph Instances
 - DNS Server Groups (Beta)
 - Identity Server Groups (Beta)
 - Security

L4-L7 Service Graph Template - FW-LB

Topology | **Policy**

Properties

Description: optional

Filters After First Node: allow-all | filters-from-contract

Function Nodes:

Name	Function Name	Function Type	Description
N1		GoTo	
N2		GoTo	

Terminal Nodes:

Name	Provider/Consumer	Description
T1	Consumer	
T2	Provider	

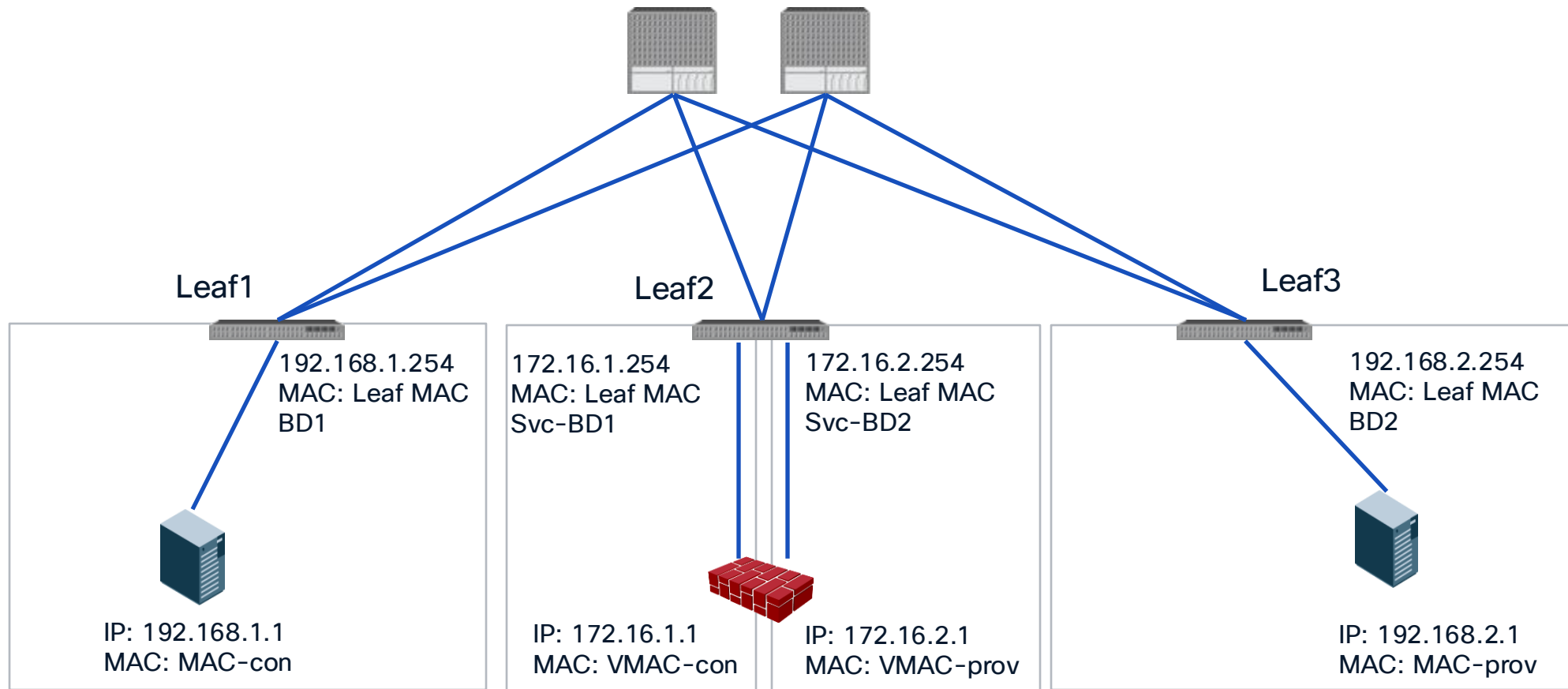
Connections:

Name	Connected Nodes	Direct Connect	Unicast Route	Adjacency Type	Description
C1	N1, T1	False	True	L3	
C2	N1, N2	False	True	L3	
C3	N2, T2	False	True	L3	

Default is “False”

How forwarding works

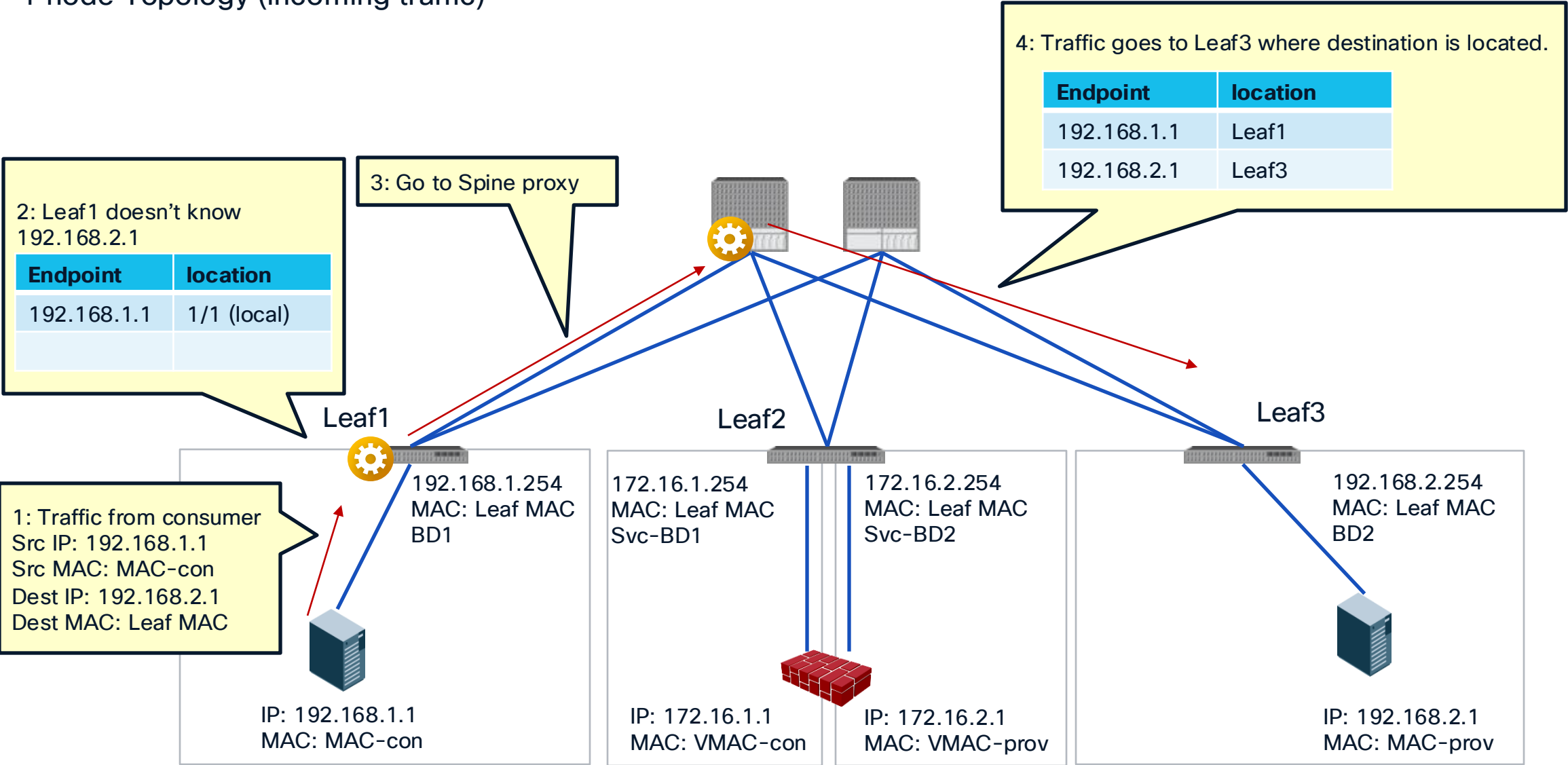
1 node Topology



How forwarding works

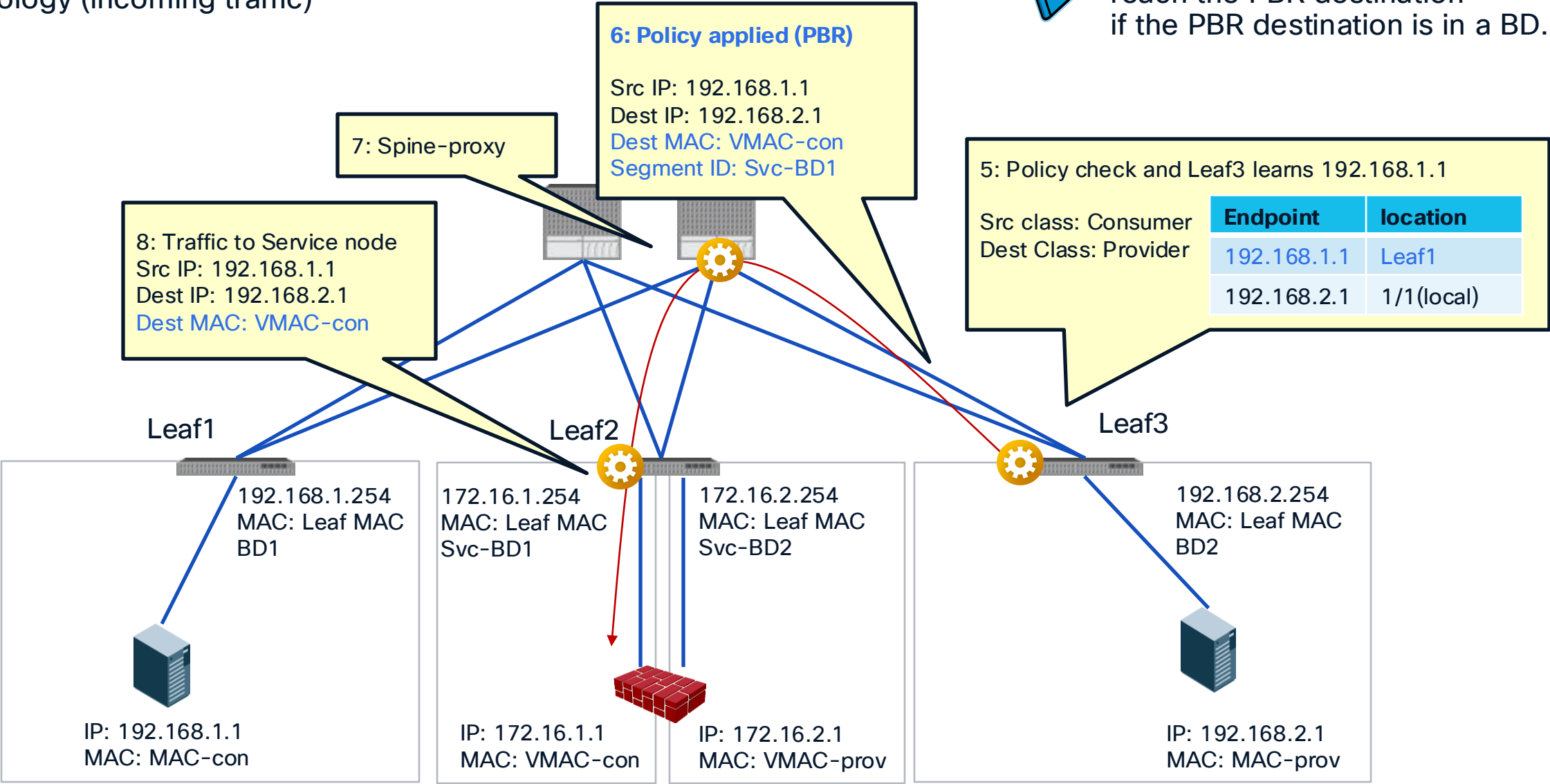
 = VXLAN Encap/Decap

1 node Topology (incoming traffic)



How forwarding works

1 node Topology (incoming traffic)



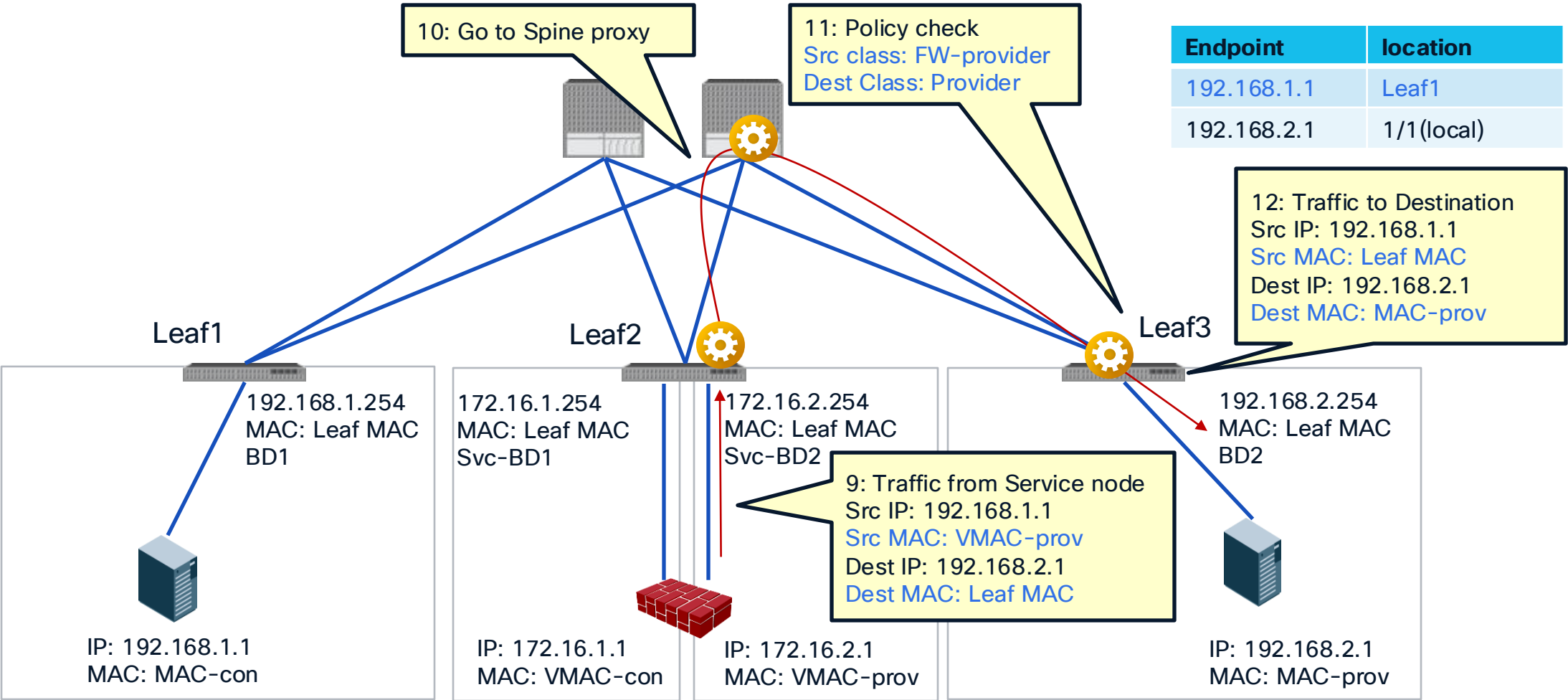
How forwarding works

1 node Topology (incoming traffic)



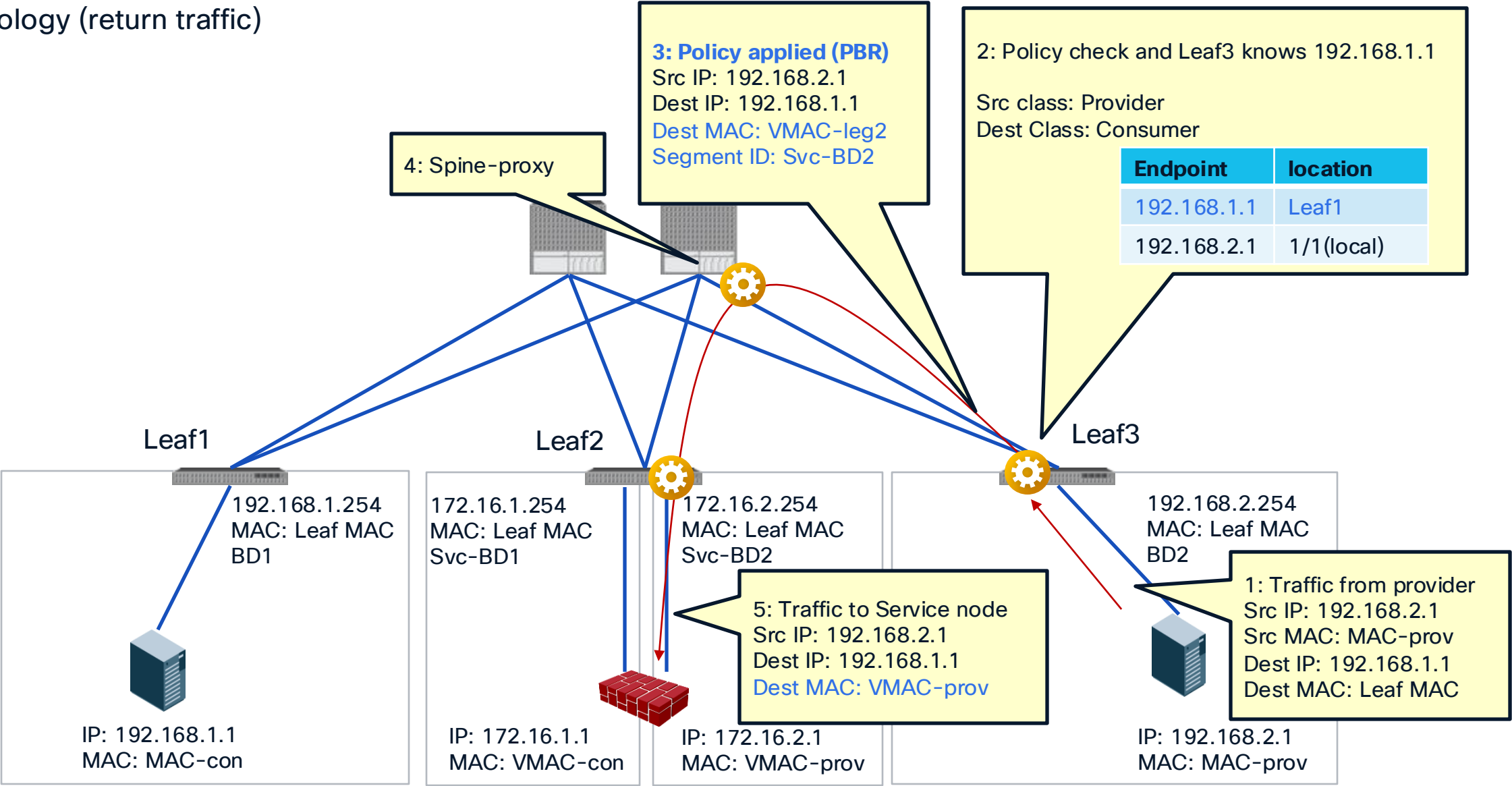
Dataplane IP learning is automatically disabled for the service EPG. (starting from 3.1)

Leaf3 doesn't re-learn 192.168.1.1 here
Because of disable dataplane IP learning



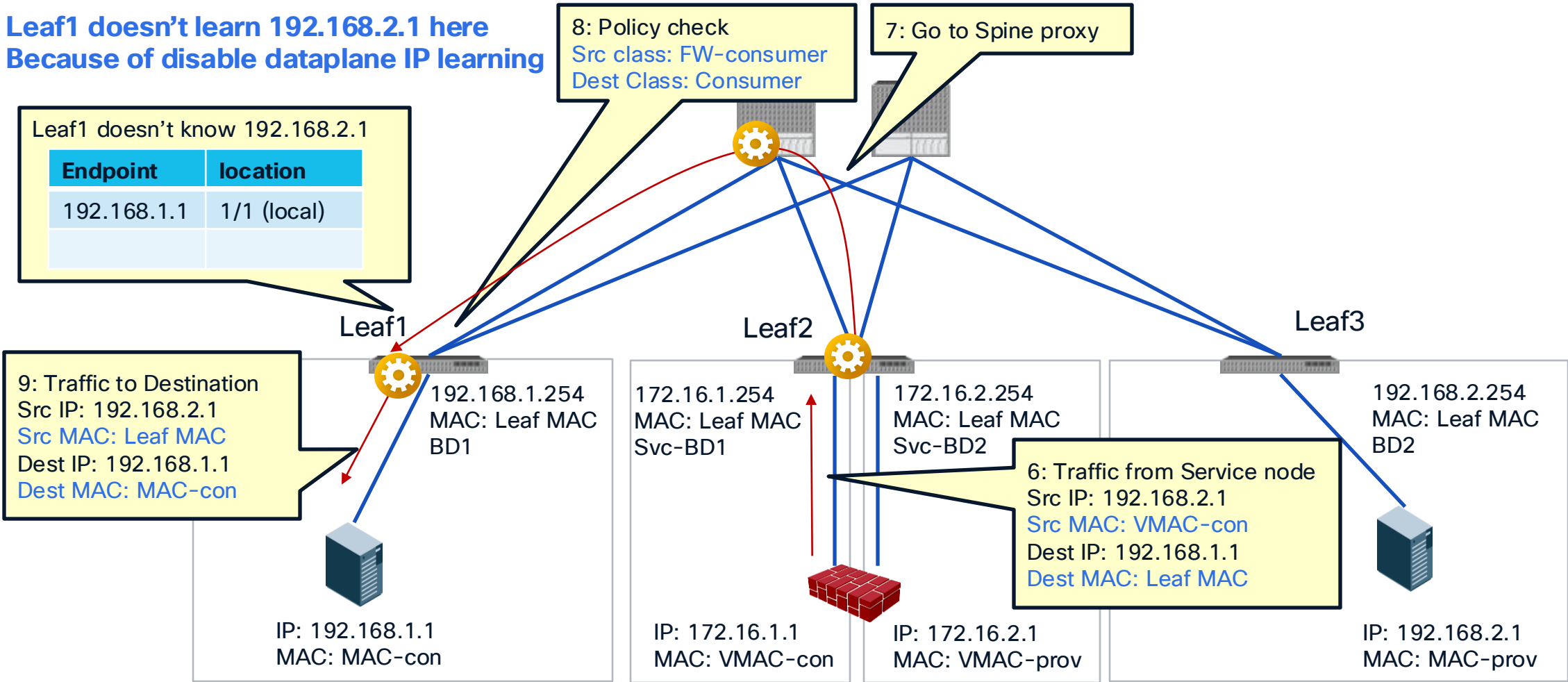
How forwarding works

1 node Topology (return traffic)



How forwarding works

1 node Topology (return traffic)



Where is the policy applied?



Please see
ACI Contract guide
for detail

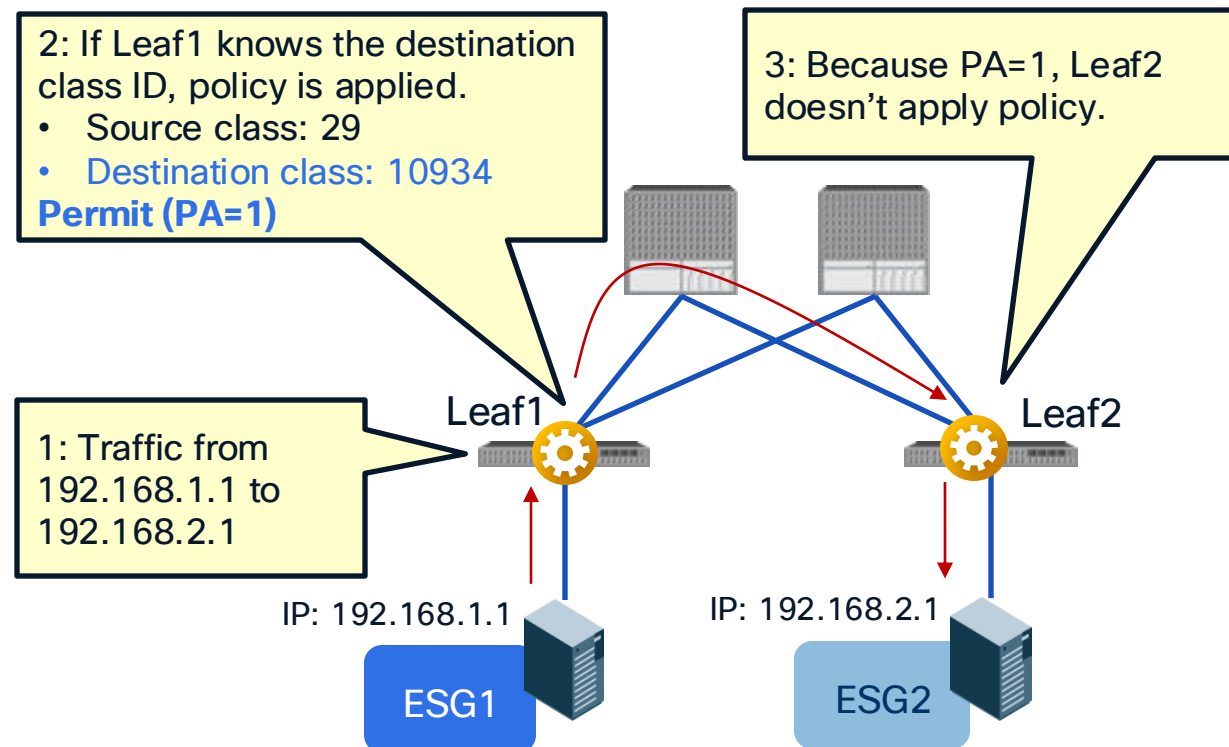
Scenario	VRF enforcement mode	Consumer	Provider	Policy enforced on
Intra-VRF	Ingress/egress	EPG	EPG	<ul style="list-style-type: none"> If destination endpoint is learned: ingress leaf If destination endpoint is not learned: egress leaf
	ingress	EPG	L3Out EPG	Consumer leaf (non-border leaf)
	ingress	L3Out EPG	EPG	Provider leaf (non-border leaf)
	egress	EPG	L3Out EPG	Border leaf -> non-border leaf traffic
	egress	L3Out EPG	EPG	<ul style="list-style-type: none"> If destination endpoint is learned: border leaf If destination endpoint is not learned: non-border leaf Non-border leaf-> border leaf traffic
	Ingress/egress	L3Out EPG	L3Out EPG	Border leaf
Inter-VRF	Ingress/egress	EPG	EPG	Consumer leaf
	Ingress/egress	EPG	L3Out EPG	Consumer leaf (non-border leaf)
	Ingress/egress	L3Out EPG	EPG	Ingress leaf
	Ingress/egress	L3Out EPG	L3Out EPG	Ingress leaf

How ingress/egress leaf enforcement works?

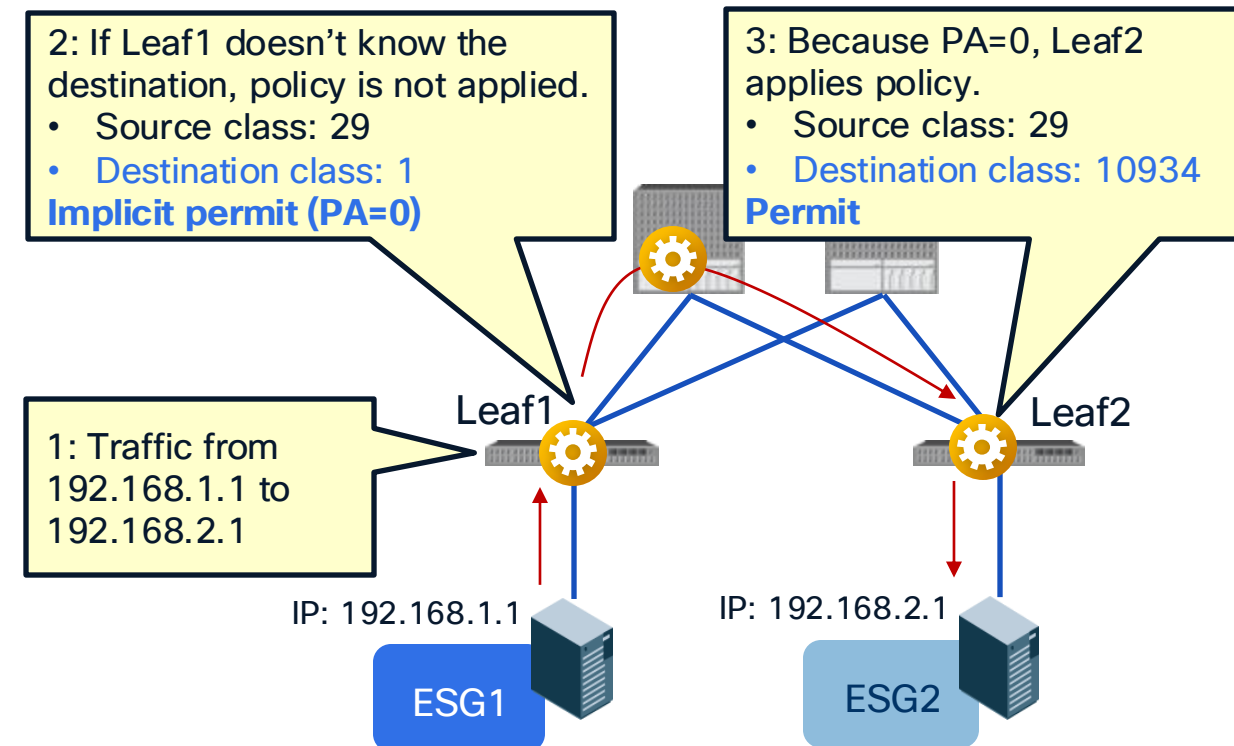
Policy Applied (PA) bit



• Intra-VRF ESG-to-ESG egress leaf enforcement



• Intra-VRF ESG-to-ESG ingress leaf enforcement



Contract Priority

Look at your zoning-rule priority and then filter priority!

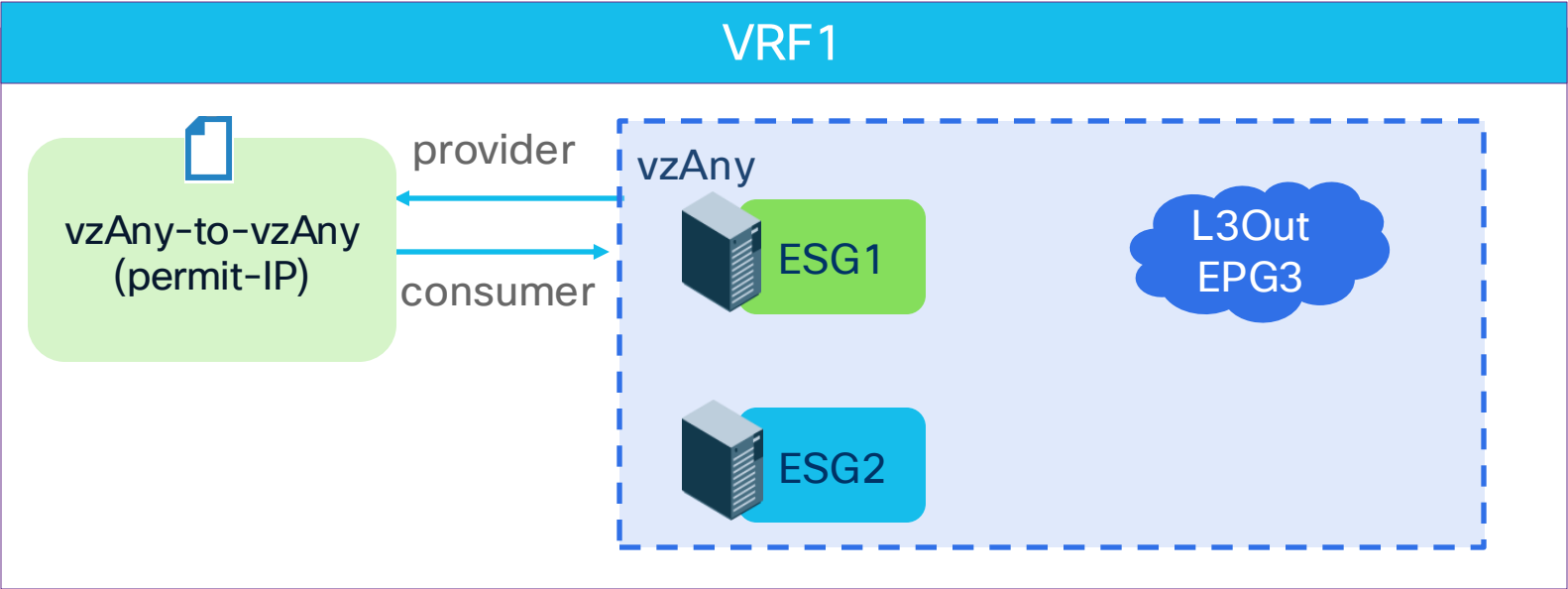


Please see
ACI Contract guide
for detail

- More specific EPGs win over vzAny and preferred groups.
 - EPG-to-EPG wins over EPG-to-vzAny/vzAny-to-EPG that wins over vzAny-to-vzAny.
 - Specific source wins over specific destination. (EPG-to-vzAny wins over vzAny-to-EPG)
- Deny actions win. Specific protocol wins.
 - If the zoning-rule priority is the same, deny wins over redirect or permit action.
 - Between redirect and permit, a more specific protocol and a specific L4 protocol wins.
- More specific L4 rules win.
 - Specific filter wins over “any” filter.
 - Specific destination wins over specific source (“s-any to d-80” wins over “s-80 to d-any”)

Example 1

What's the forwarding action?



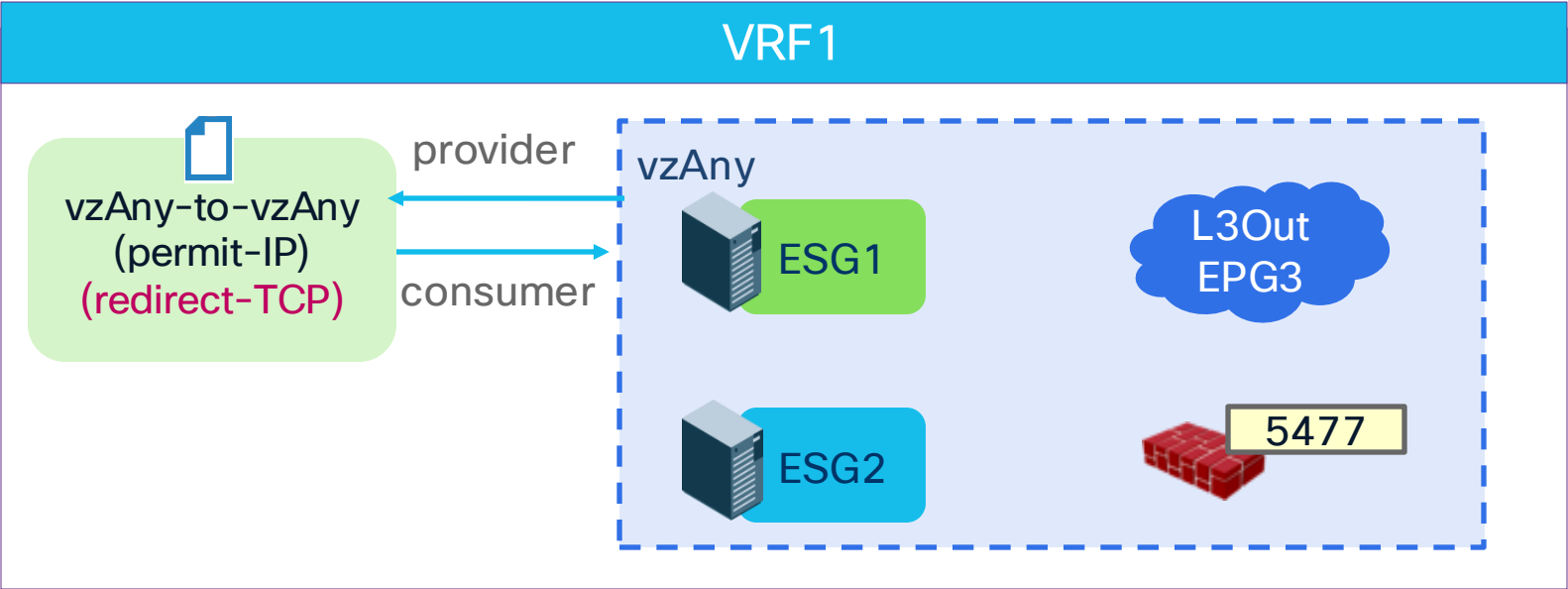
- ESG1-to-ESG2 (IP)
 - Permit
- ESG1-to-L3OutEPG3 (IP)
 - Permit
- ESG2-to-L3OutEPG3 (IP)
 - Permit

```
Pod1-Leaf1# show zoning-rule scope 2195459
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Rule ID | SrcEPG | DstEPG | FilterID | Dir | operSt | Scope | Name | Action | Priority |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
<snip>
| 4194 | 0 | 0 | 74 | uni-dir | enabled | 2195459 | tenant1:vzAny-to-vzAny | permit | any_any_filter(17) |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```


Example 2

What's the forwarding action?



- ESG1-to-ESG2 (TCP)
- Redirect
- ESG1-to-ESG2 (UDP)
- Permit

More specific L4 rules win though the zoning-rule priority is the same.

```
Pod1-Leaf1# show zoning-rule scope 2195459
```

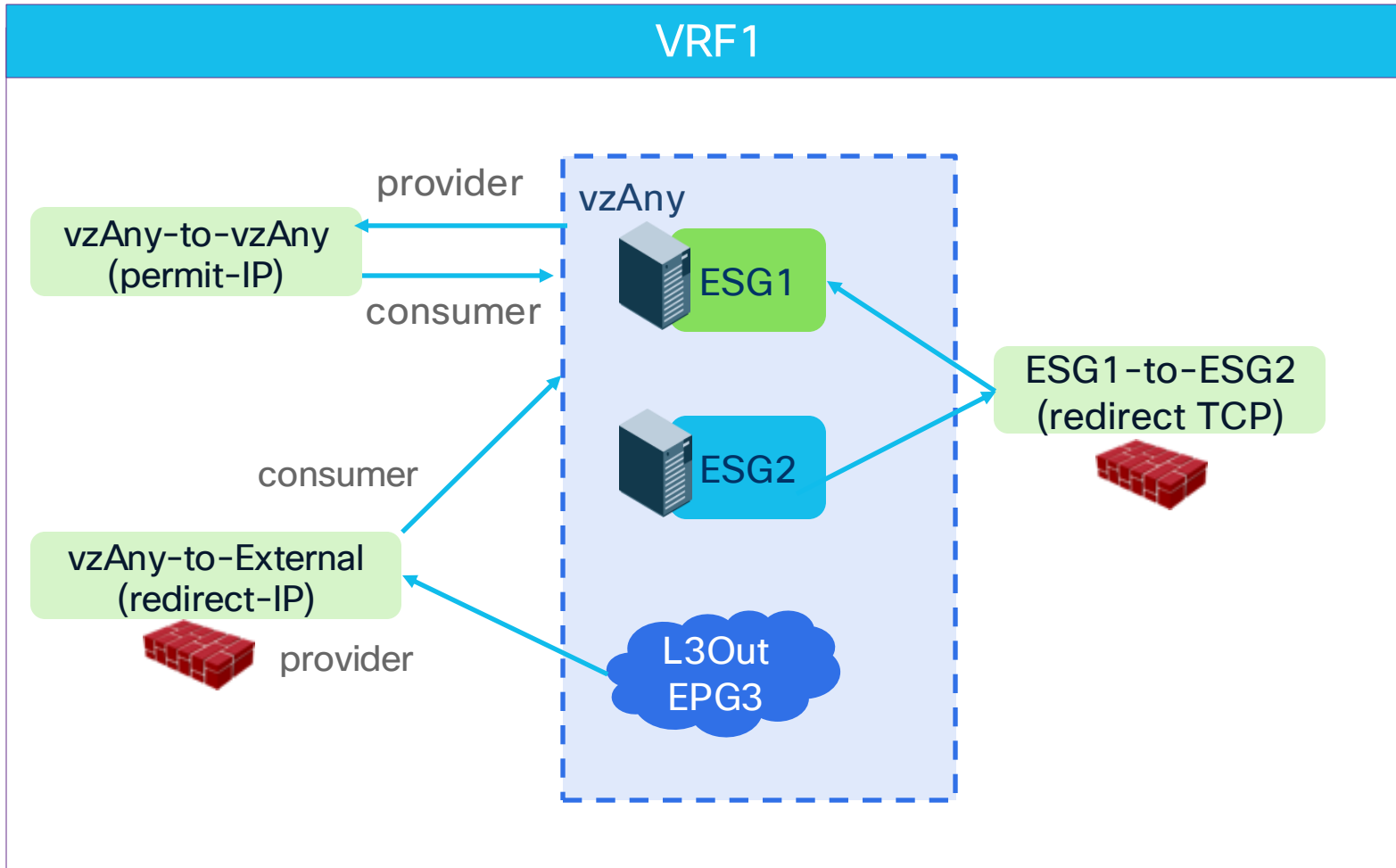
Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
4194	0	0	74	uni-dir	enabled	2195459	tenant1:vzAny-to-vzAny	permit	any_any_filter(17)
4248	0	0	14	uni-dir	enabled	2195459		redir(destgrp-20)	any_any_filter(17)
4186	5477	0	14	uni-dir	enabled	2195459		permit	shsrc_any_filt_perm(10)
4193	5477	0	default	uni-dir	enabled	2195459		permit	shsrc_any_any_perm(11)

In this example:

- Filter ID 74: Permit-IP all
- Filter ID 14: Permit-TCP all

Example 3

What's the forwarding action?



- ESG1-to-ESG2 (TCP)
 - Redirect
- ESG1-to-L3OutEPG3 (IP)
 - Redirect
- ESG1-to-ESG2 (UDP)
 - Permit

Example 3

Why?



- **ESG-to-ESG (priority 7)** wins over External-to-vzAny/vzAny-to-External (priority 13 or 14) that wins over **vzAny-to-vzAny (priority 17)**.

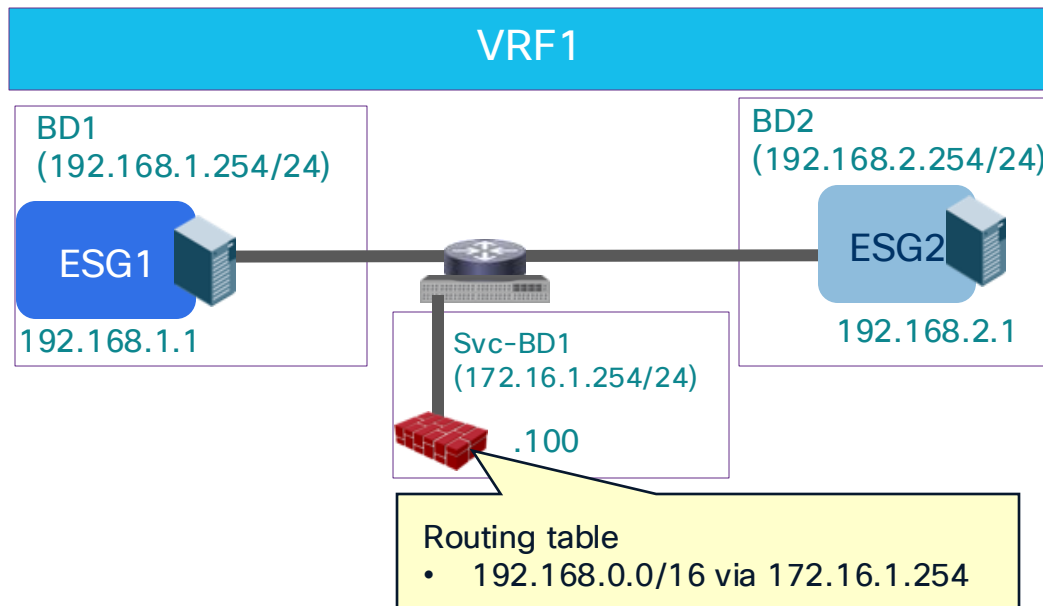
Pod1-Leaf1# show zoning-rule scope 2195459

Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
4194	0	0	74	uni-dir	enabled	2195459	tenant1:vzAny-to-vzAny	permit	any_any_filter(17)
4172	0	32782	74	uni-dir	enabled	2195459		redir(destgrp-1)	any_dest_filter(14)
4196	5477	32782	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4201	32782	0	74	uni-dir	enabled	2195459		redir(destgrp-1)	src_any_filter(13)
4242	5477	0	74	uni-dir	enabled	2195459		permit	shsrc_any_filt_perm(10)
4186	24	10936	14	bi-dir	enabled	2195459		redir(destgrp-1)	fully_qual(7)
4193	5477	10936	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4209	5477	24	14	uni-dir	enabled	2195459		permit	fully_qual(7)
4248	10936	24	14	uni-dir-ignore	enabled	2195459		redir(destgrp-1)	fully_qual(7)

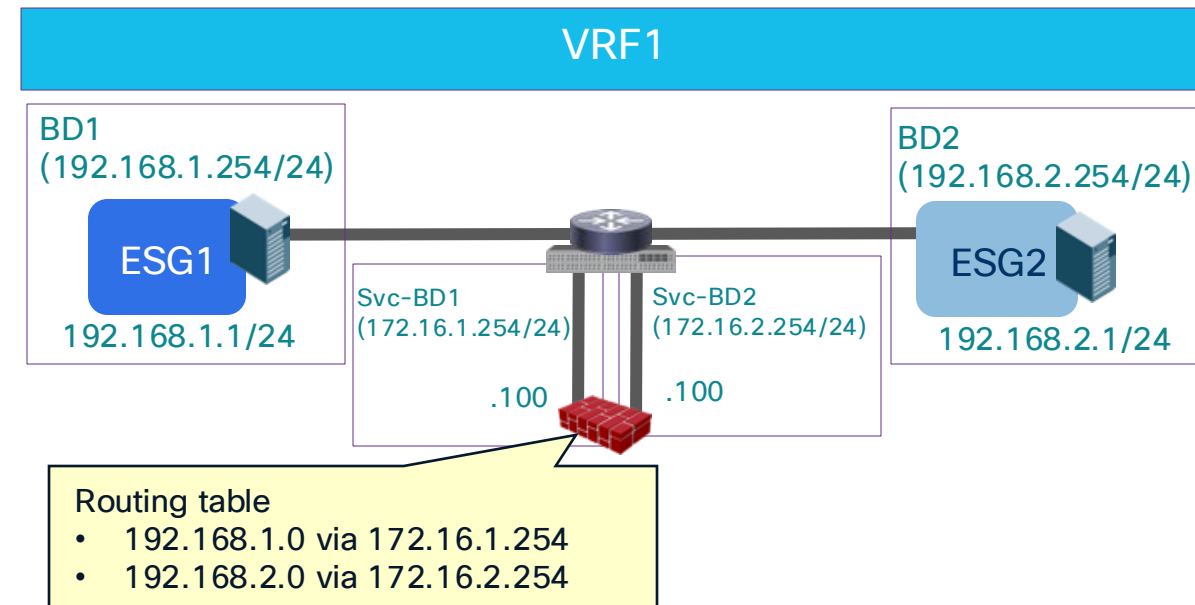
FAQs and advanced use cases

One-arm vs Two-arm?

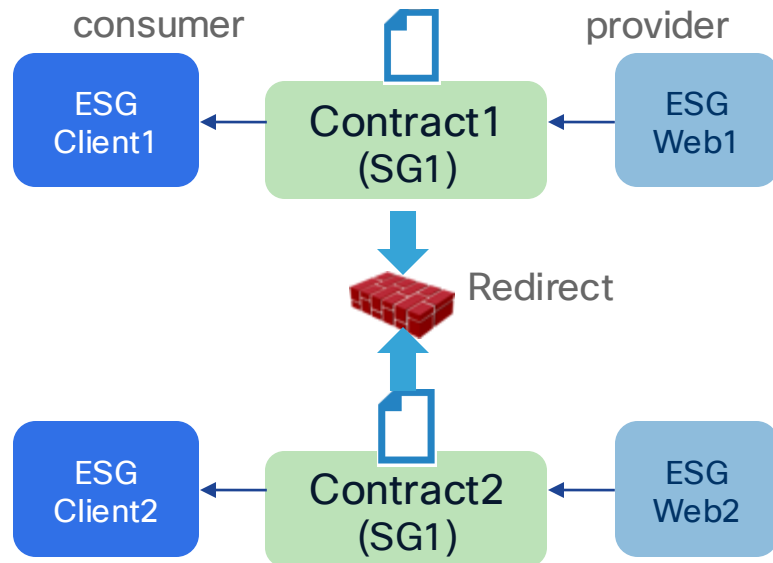
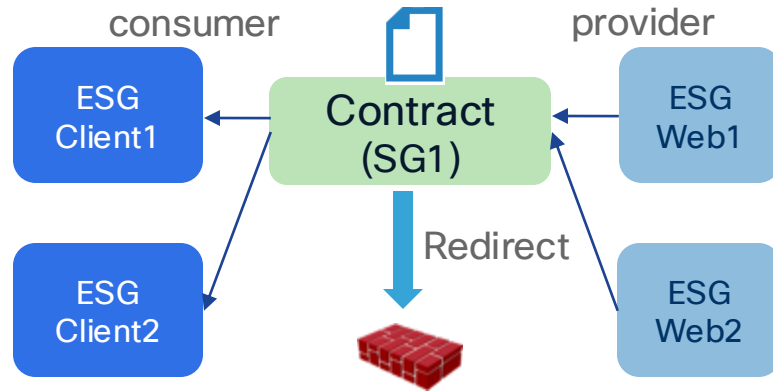
- One-arm
 - Simple routing design on service node.
 - One-arm must be used for intra-subnet or intra-EPG/ESG contract.
 - Some firewall doesn't allow intra-interface traffic by default.



- Two-arm
 - Need to manage routing design on service node.
 - Different security level on each interface.



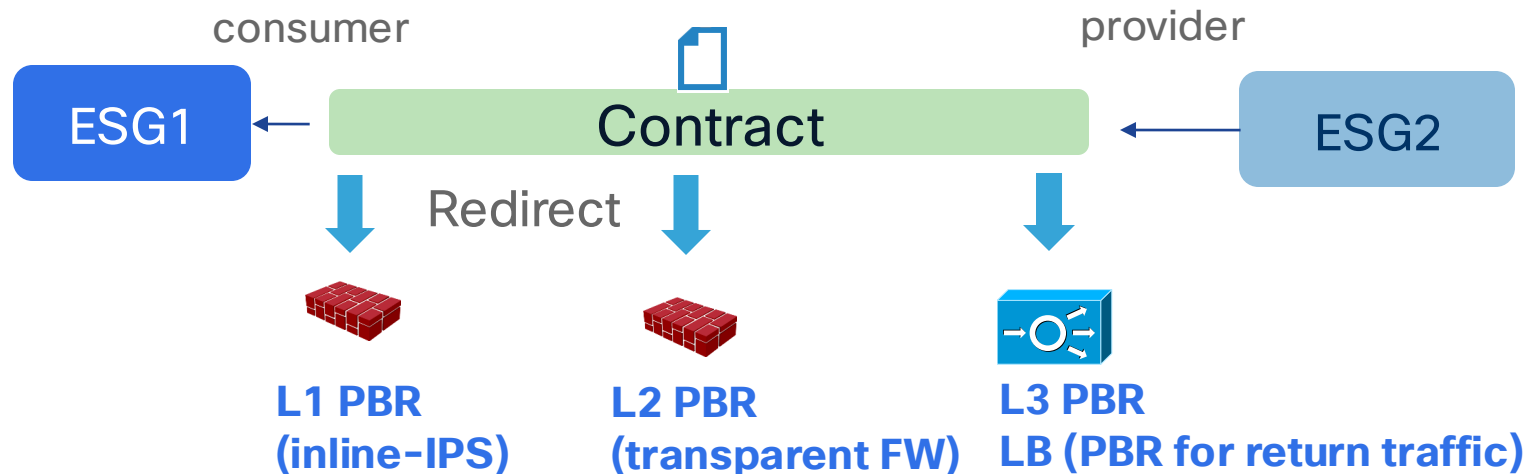
Can we reuse same PBR destination multiple times?



- Multiple consumer/provider ESGs/EPGs
- Multiple contracts can use the same PBR destination and Service Graph.
- Note
 - It could consume more TCAM resources if many EPGs consume and provide the same contract. The use of vzAny or ESG might be more efficient.
 - Depending on routing design, one-arm mode deployment may be required.

What types of devices can be PBR destinations?

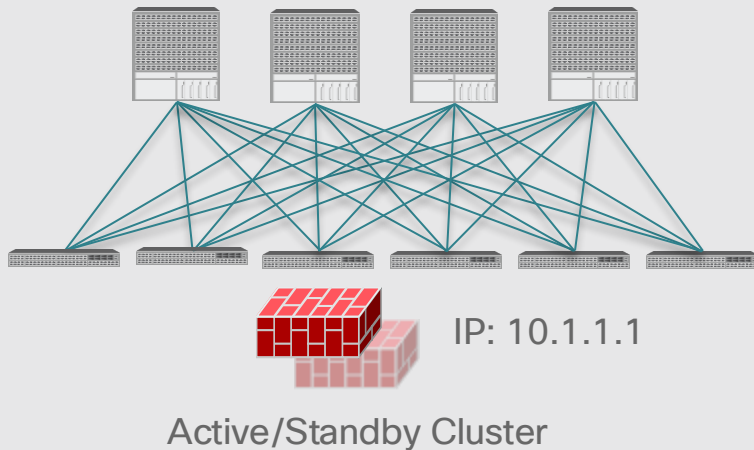
- Prior to ACI Release 5.0, a PBR destination must be an L3 routed device (L3 PBR).
- Starting from ACI Release 5.0, L1/L2 PBR is supported to insert L1/L2 devices.
 - Insert firewall without relying on BD/VLAN stitching.
 - L1/L2 service device BD must be dedicated BD that cannot be shared with other endpoints.
 - L1/L2/L3 PBR can be mixed in a service graph.



What are HA options?

One PBR destination IP
One Logical device with two concrete devices

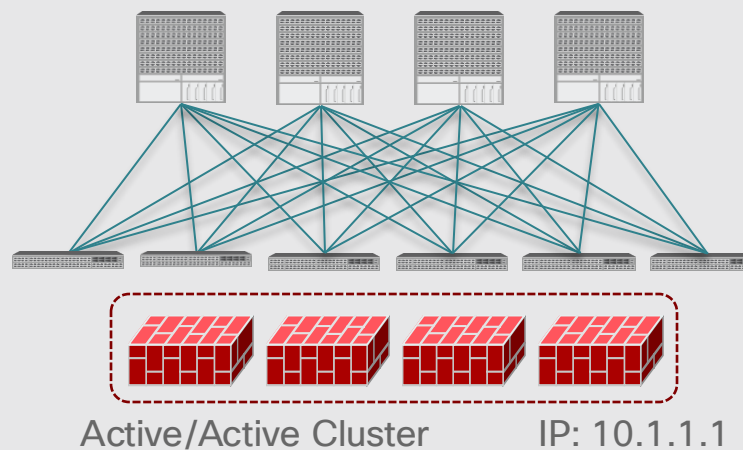
Active/Standby Cluster



- PBR is not mandatory
- The Active/Standby pair represents a single MAC/IP entry.

One PBR destination IP
One Logical device with one concrete device

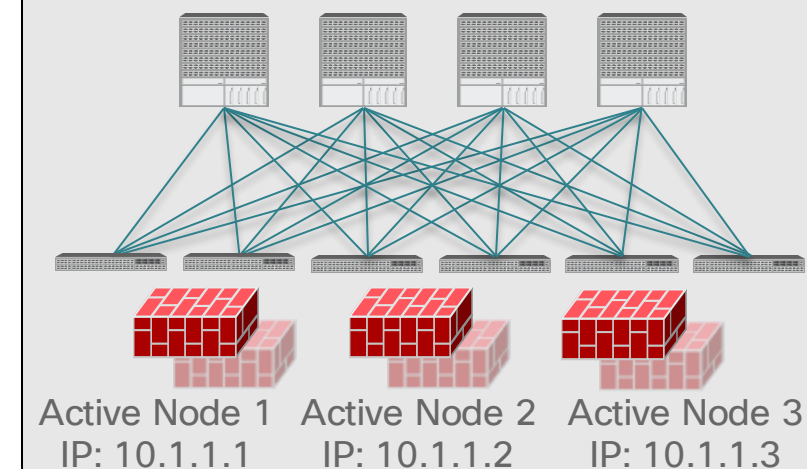
Active/Active Cluster (‘Scale-Up’ Model)



- PBR is required if the cluster is stretched across pods.
- The Active/Active cluster represents a single MAC/IP entry.
- Spanned Ether-Channel Mode supported with Cisco ASA/FTD platforms

Multiple PBR destination IPs (Symmetric PBR)
One Logical device with multiple concrete devices

Independent Active Nodes (‘Scale-Out’ Model)

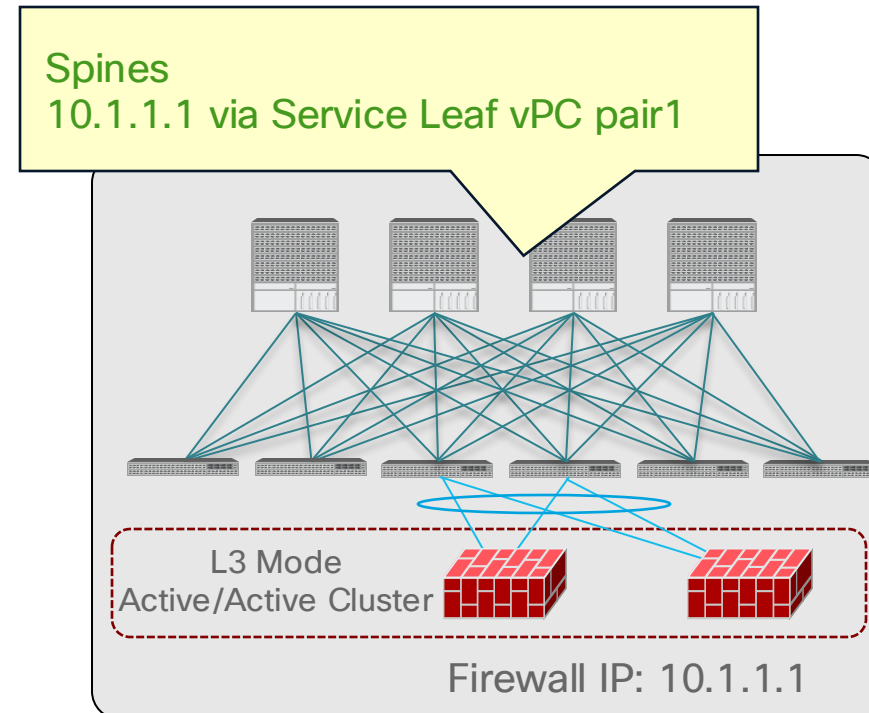
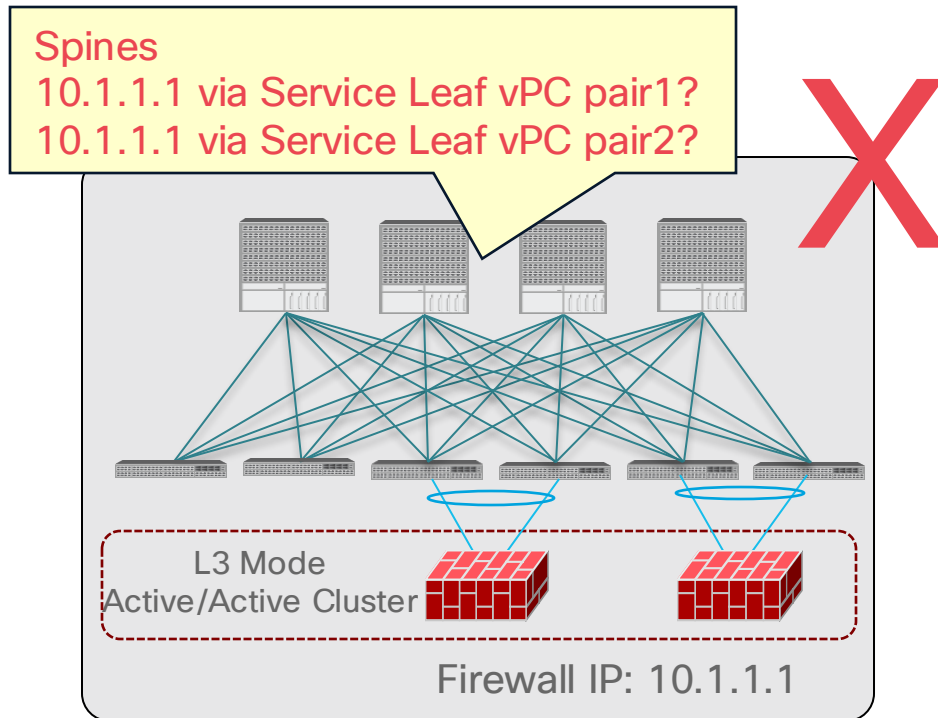


- PBR is required.
- Each Active node represent a unique MAC/IP entry.
- Use of Symmetric PBR to ensure each flow is handled by the same Active node in both directions

Active/Active cluster

One PC/vPC to all devices in the cluster

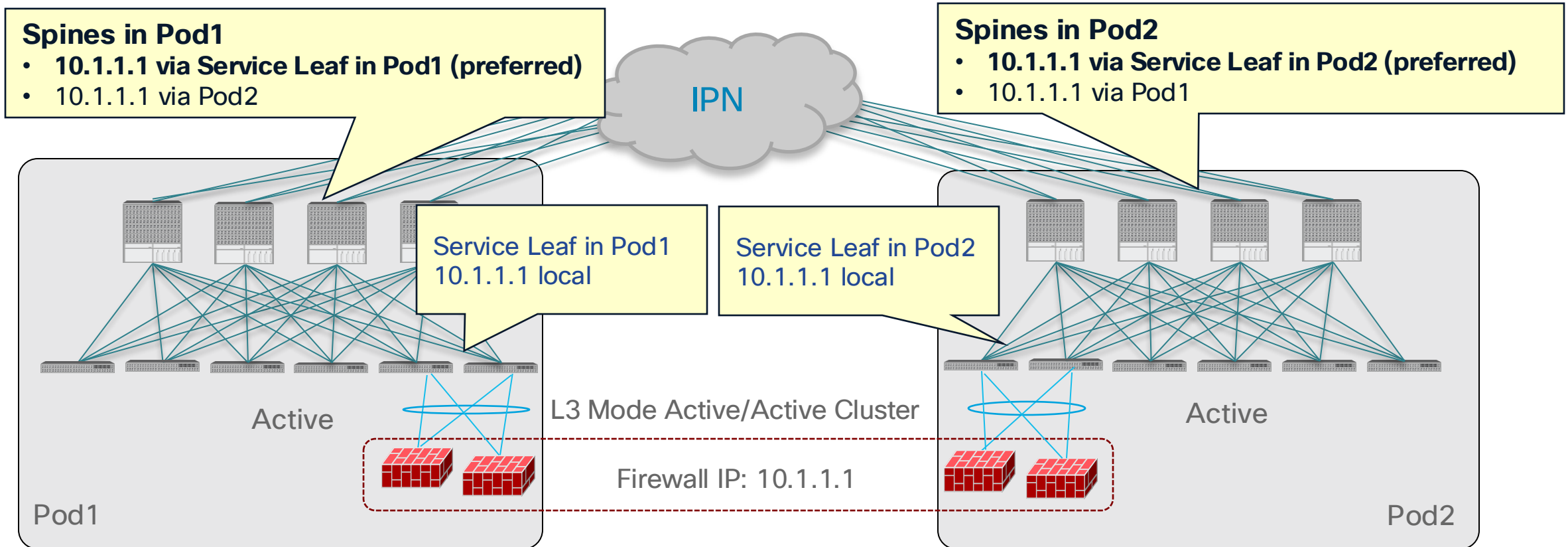
- Firewalls in the same cluster must be connected via the same PC/vPC in each pod. Otherwise, the same endpoint will be learned via different locations, which results in endpoint flapping.



Active/Active cluster across pods

Anycast service

- For Multi-pod, Anycast service feature must be enabled.



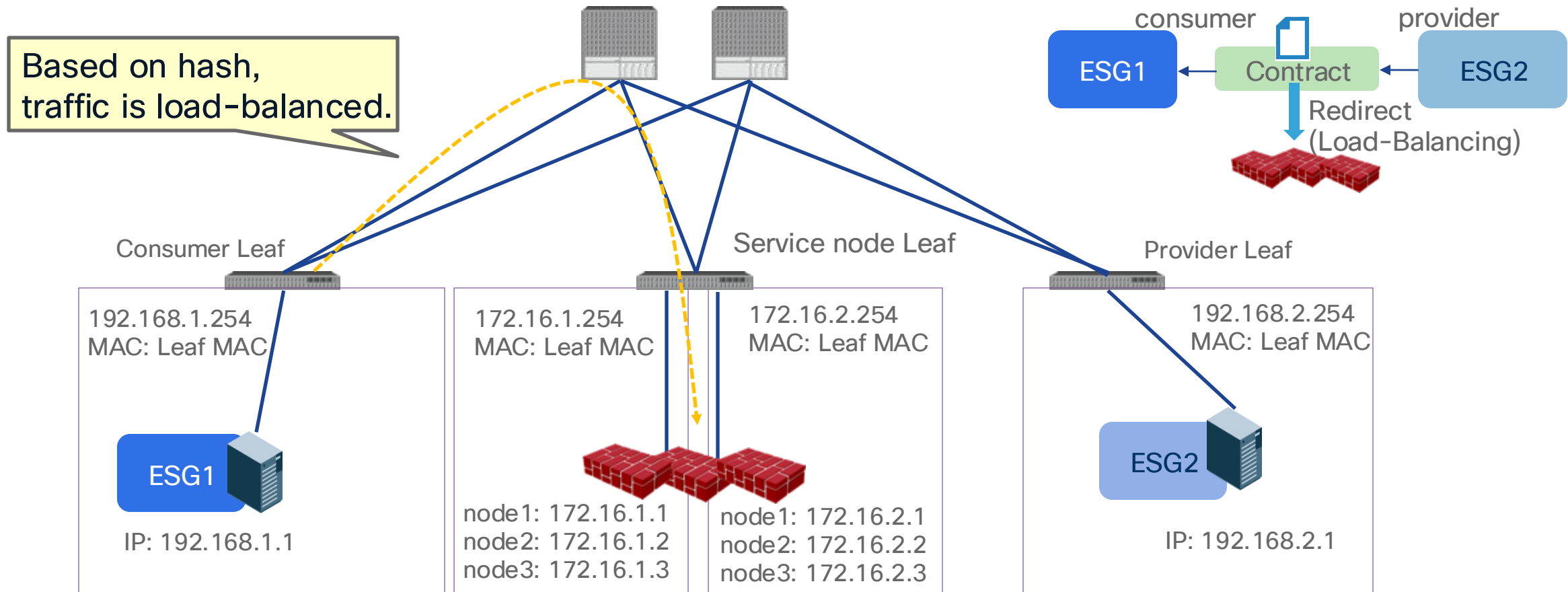
Independent Active Nodes

Symmetric PBR: Scale Firewall Easily



PBR destinations can be distributed across multiple leaf nodes.

- Ensure incoming and return traffic go to the same firewall



Independent Active Nodes

Symmetric PBR: Hash algorithm option

- Source IP, Destination IP and Protocol number (default)
- Source IP only
- Destination IP only

Create L4-L7 Policy-Based Redirect

Name: FW-external

Description: optional

Destination Type: L1 L2 L3

Rewrite source MAC: ☐

IP SLA Monitoring Policy: select an option

Enable Pod ID Aware Redirection: ☐

Hashing Algorithm: Destination IP Source IP Source IP, Destination IP and Protocol number

Enable Anycast: ☐

Resilient Hashing Enabled: ☐

L3 Destinations:

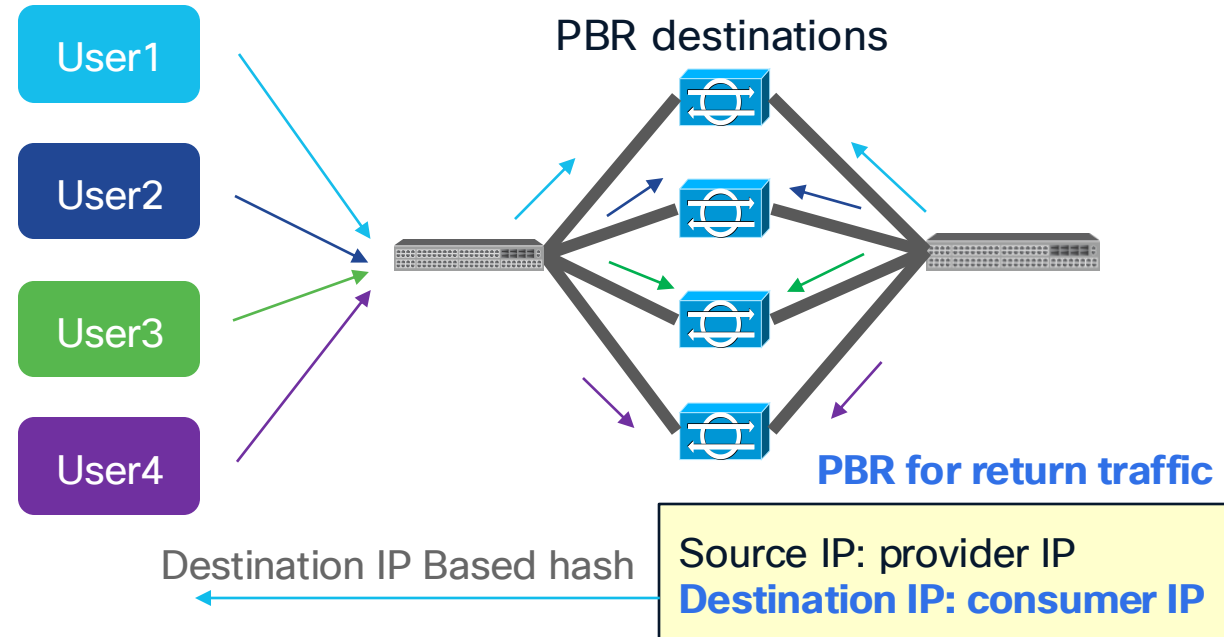
IP	Destination MAC Name	Redirect Health Group	Additional IPv4/IPv6	Description	Oper Status
----	----------------------	-----------------------	----------------------	-------------	-------------

Example: same user (IP) will go through the same device

PBR for incoming traffic

Source IP: consumer IP
Destination IP: provider IP

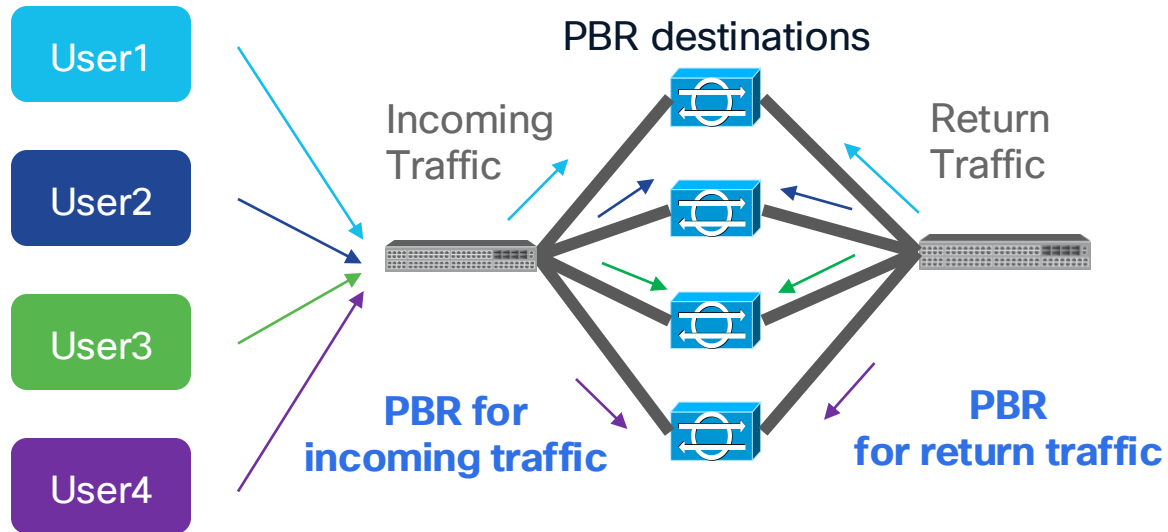
Source IP Based hash



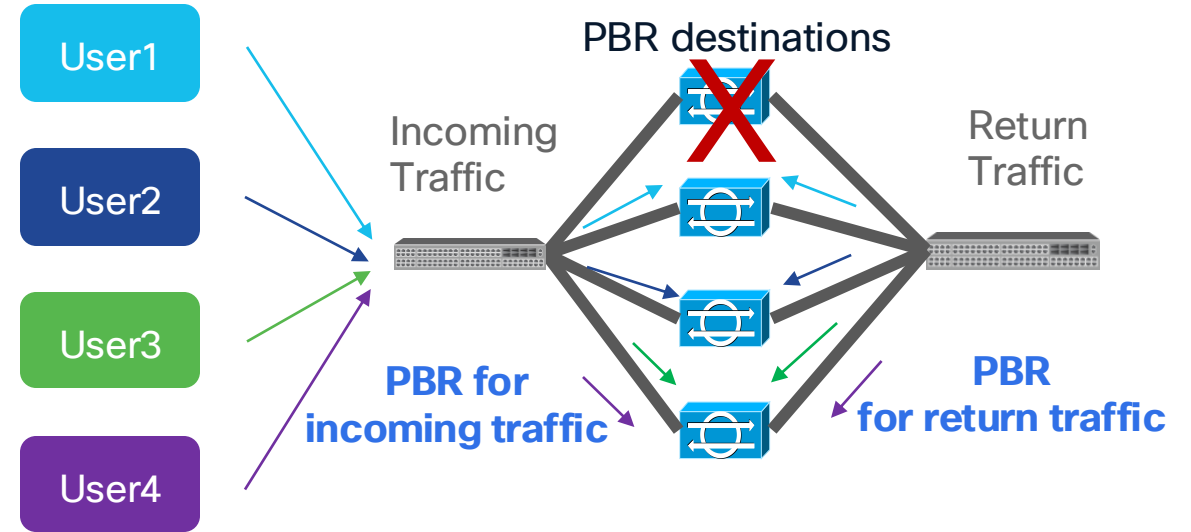
What happens if an L4-L7 device is down?

- If one of the PBR nodes goes down, existing traffic flows will be rehashed. This could lead to the connection being reset.

Thanks to Symmetric PBR, incoming and return traffic go to same PBR node.



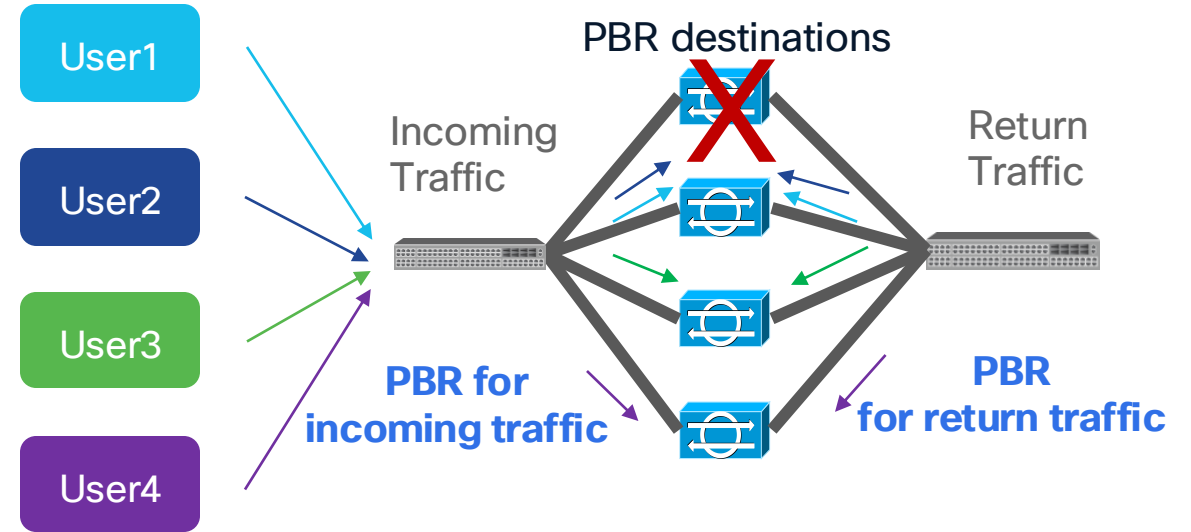
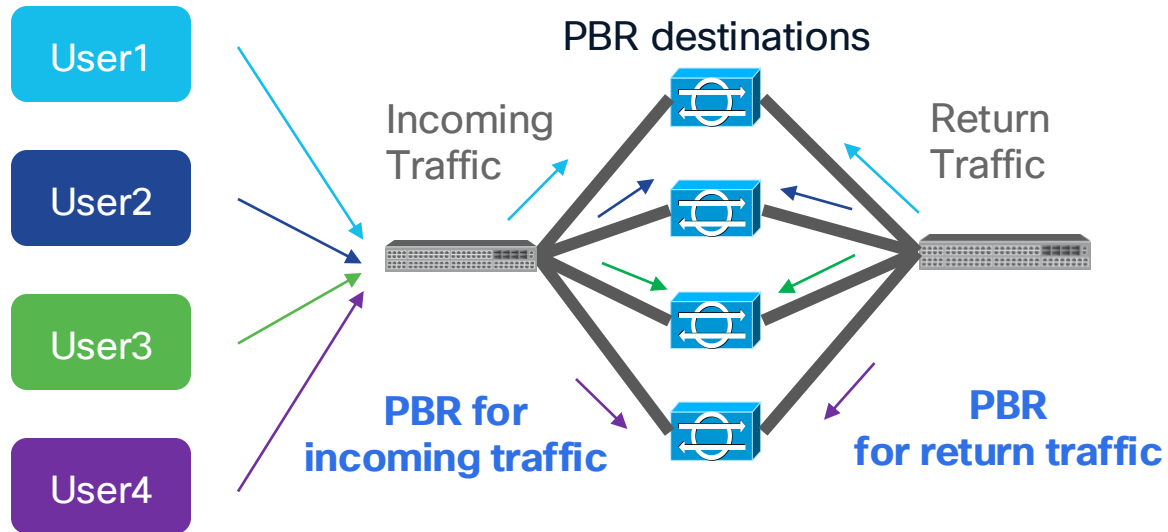
Some traffic could be load-balanced to different PBR nodes that don't have existing connection info.



I want to minimize impact on the existing flow!

With Resilient Hash

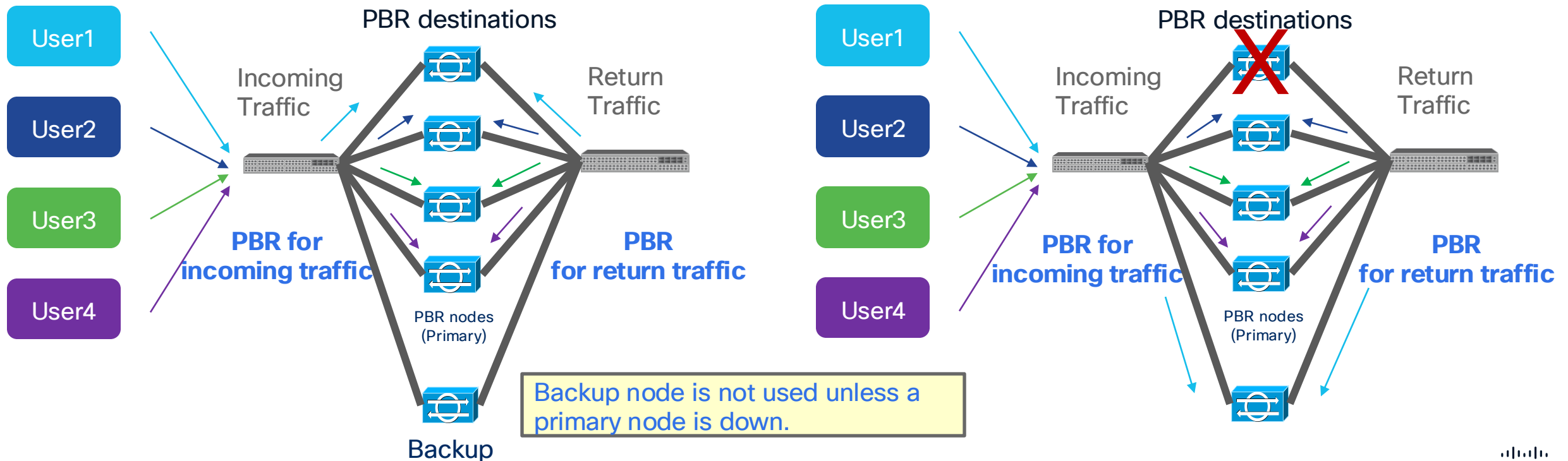
- With Resilient Hash PBR, only the traffics that went through failed node will be rerouted to one of the available nodes.



Can we use standby PBR destination?

Resilient Hash PBR with N+M backup

- As all the traffic that went through the failed node will go to one of the available nodes, capacity of the node is a concern. (The node would have doubled amount of traffic compared with usual)
- Instead of using one of the available primary nodes, a backup node in the group will be used. (N+M)

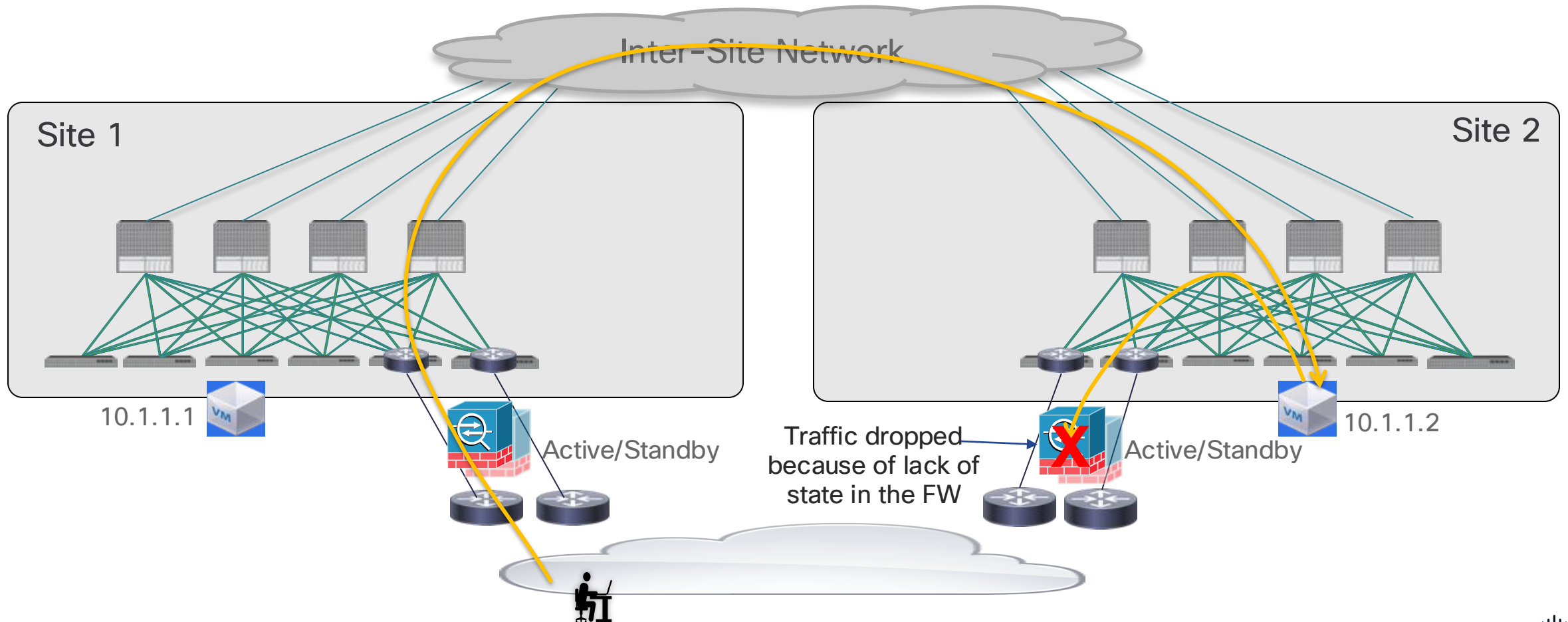


Multi-location Data Centers

Service insertion in multiple DC locations

What is the challenge of service insertion in multiple DC locations?

- Traffic Symmetry is important

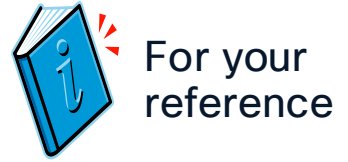


Multi-location Data Centers

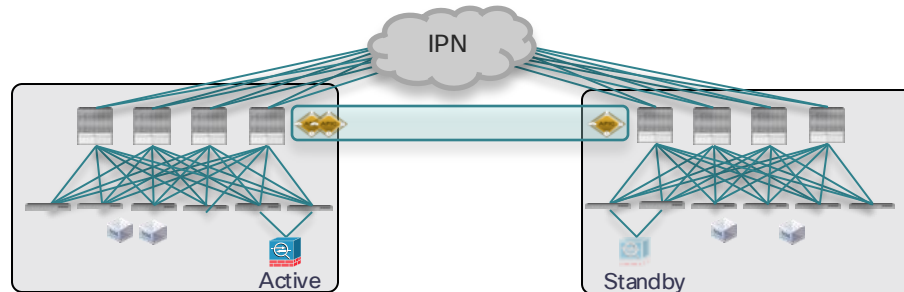
- Multi-Pod
- Multi-Site

ACI Multi-Pod

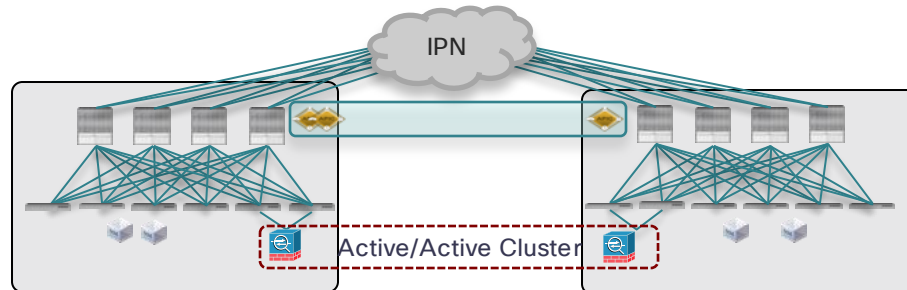
Design options



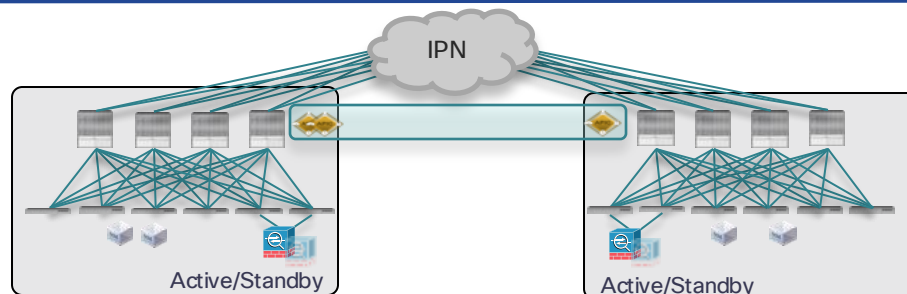
Typical options for an Active/Active DC use case



- Active and Standby pair deployed across Pods
- No issues with asymmetric flows



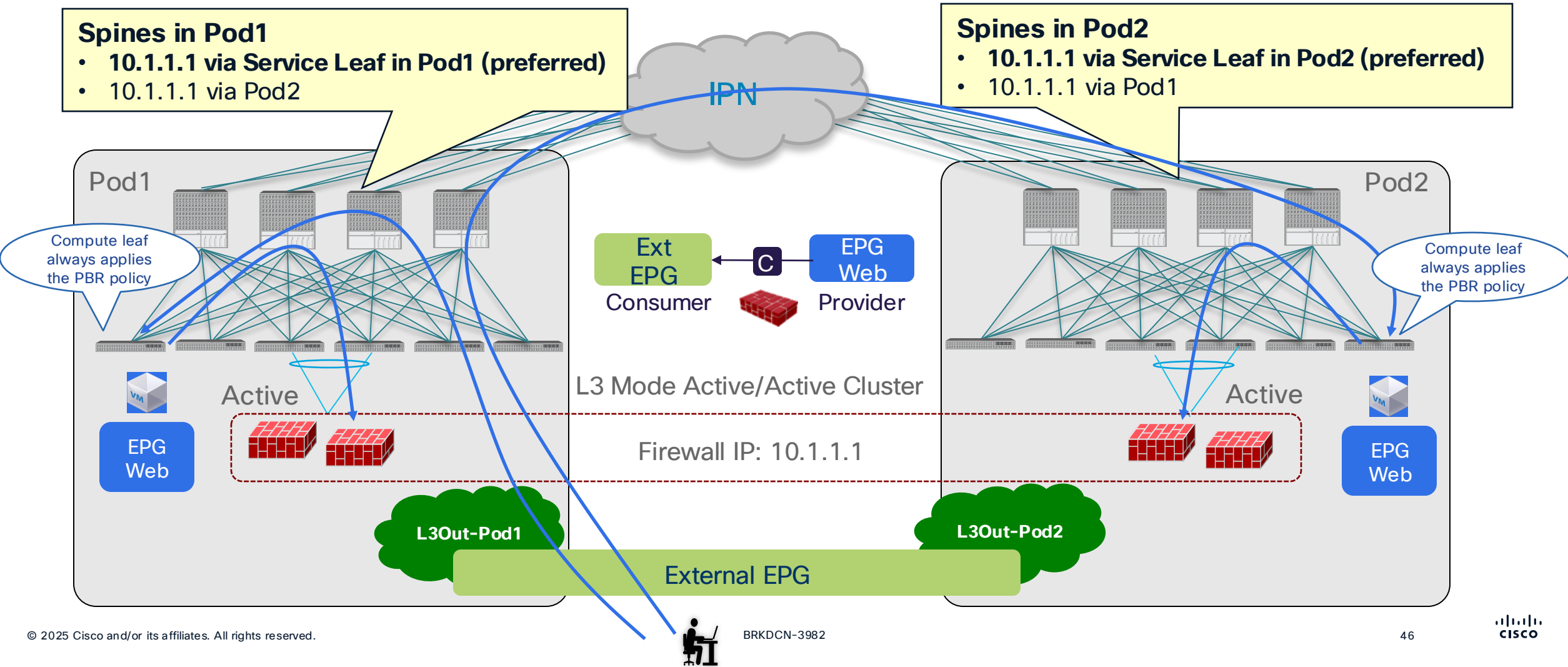
- Active/Active FW cluster nodes stretched across Sites (single logical FW)
- Requires the ability of discovering the same MAC/IP info in separate pods at the same time
- Supported from ACI release 3.2(4d) with the use of Service-Graph with PBR



- Independent Active/Standby pairs deployed in separate Pods
- Use of Symmetric PBR to avoid the creation of asymmetric paths crossing different active FW nodes

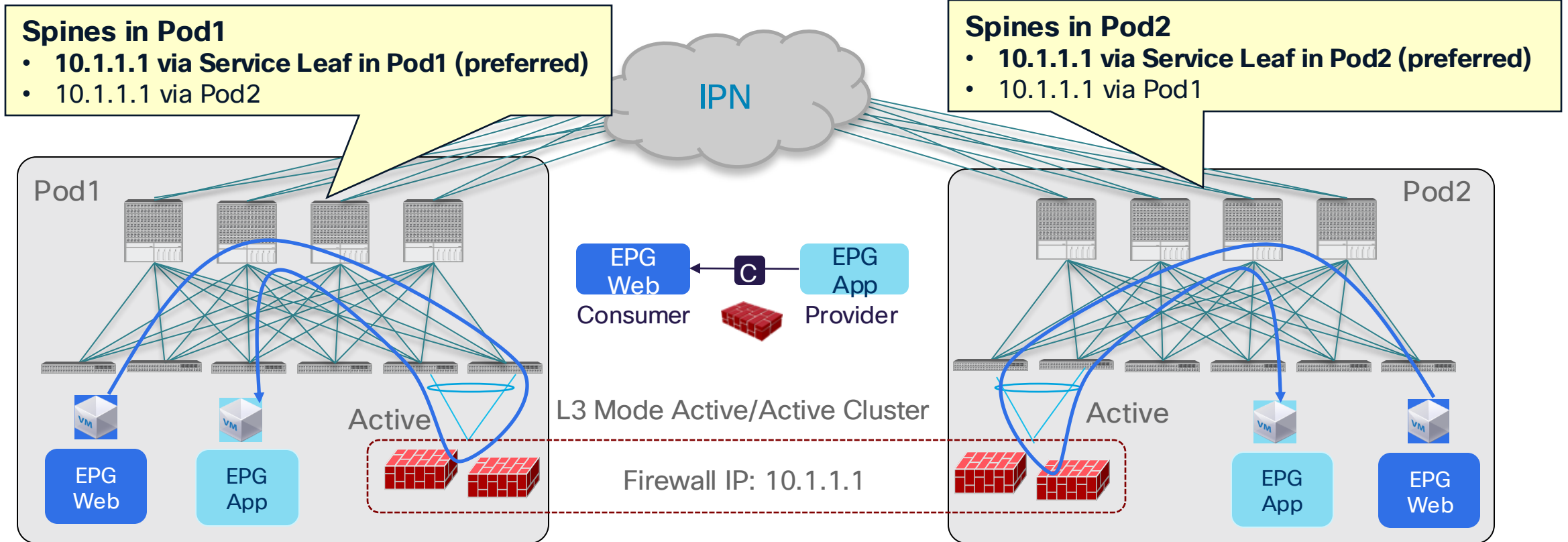
ACI Multi-Pod: Active/Active cluster across pods

North-South Traffic Flow



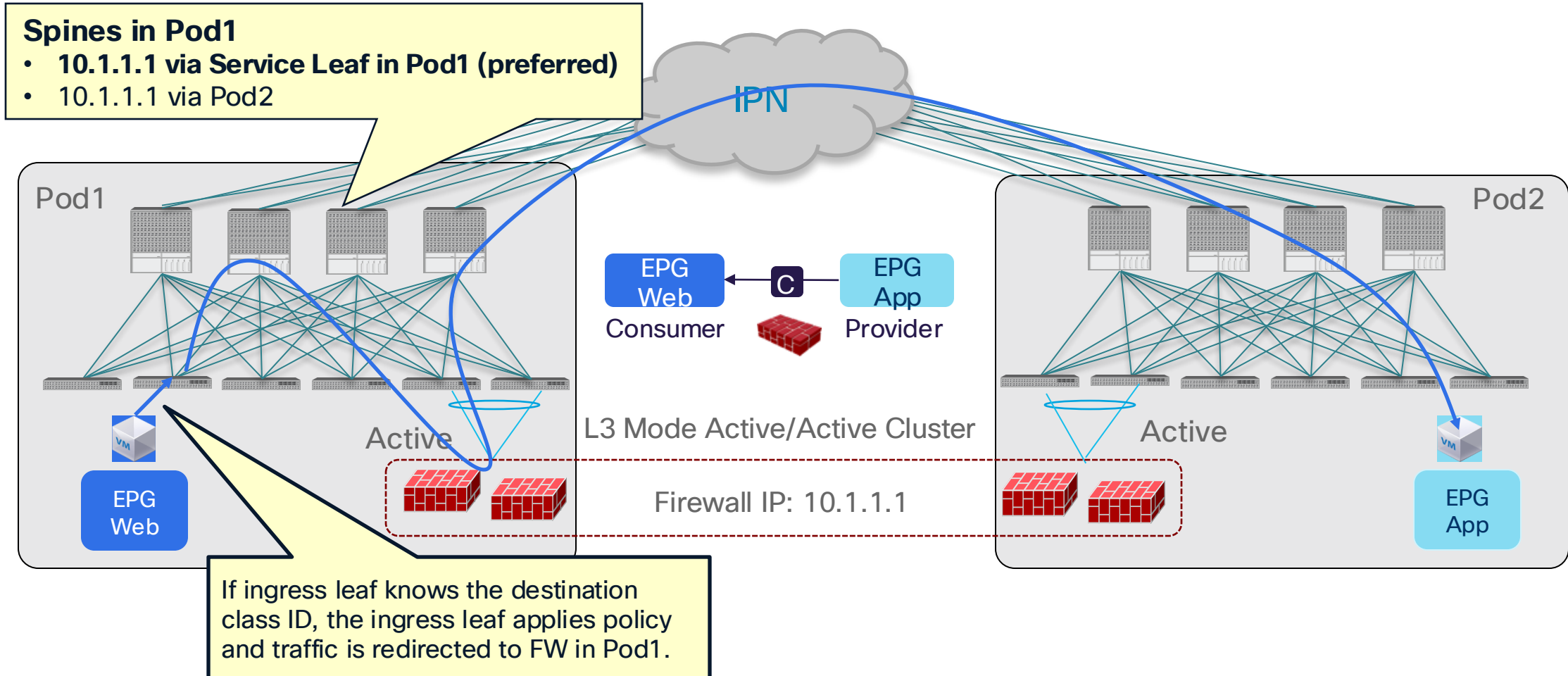
ACI Multi-Pod: Active/Active cluster across pods

East-West Traffic Flow (Intra-Pod)



ACI Multi-Pod: Active/Active cluster across pods

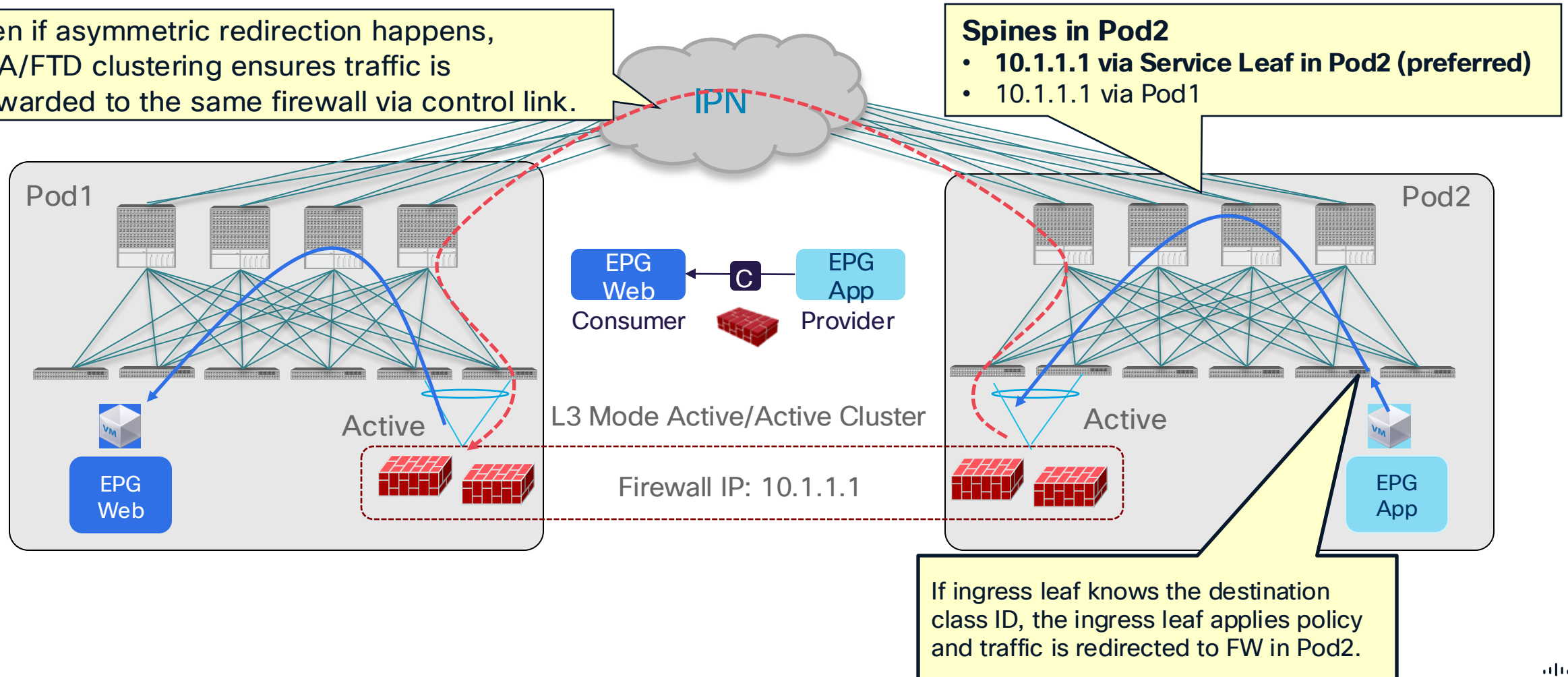
East-West Traffic Flow (Inter-Pod) incoming traffic



ACI Multi-Pod: Active/Active cluster across pods

East-West Traffic Flow (Inter-Pod) return traffic

Even if asymmetric redirection happens, ASA/FTD clustering ensures traffic is forwarded to the same firewall via control link.

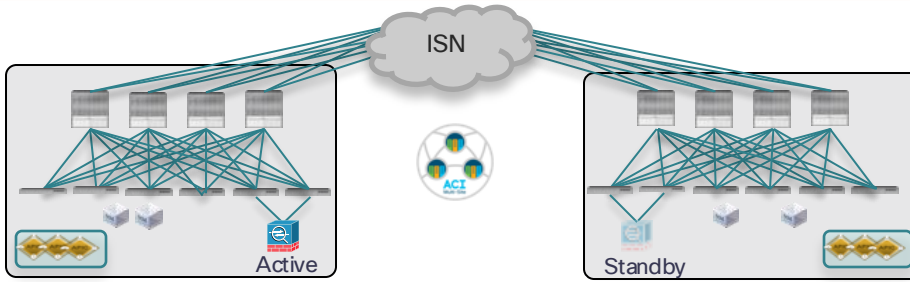


ACI Multi-Site

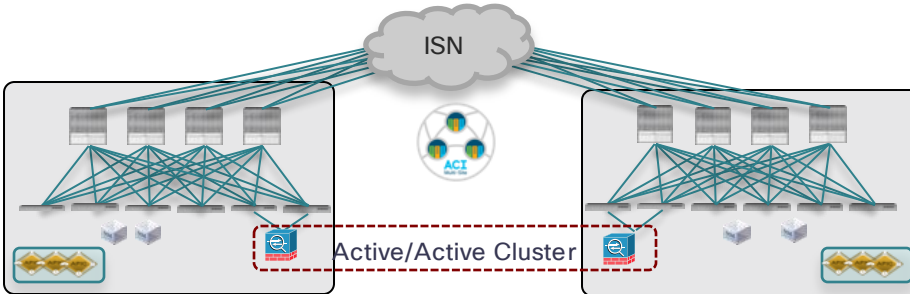
Design options



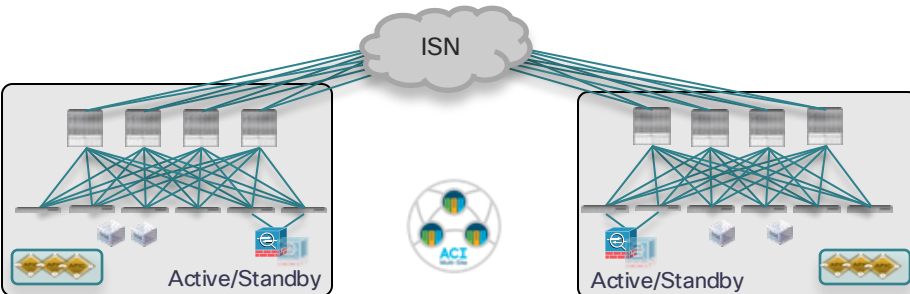
Deployment options fully supported with ACI Multi-Pod



- Active and Standby pair deployed across Pods
- **Limited supported options**



- Active/Active FW cluster nodes stretched across Sites (single logical FW)
- **Not supported**



- **Recommended deployment model for ACI Multi-Site**
- Supported from 3.2 release with the use of Service Graph with Policy Based Redirection (PBR)

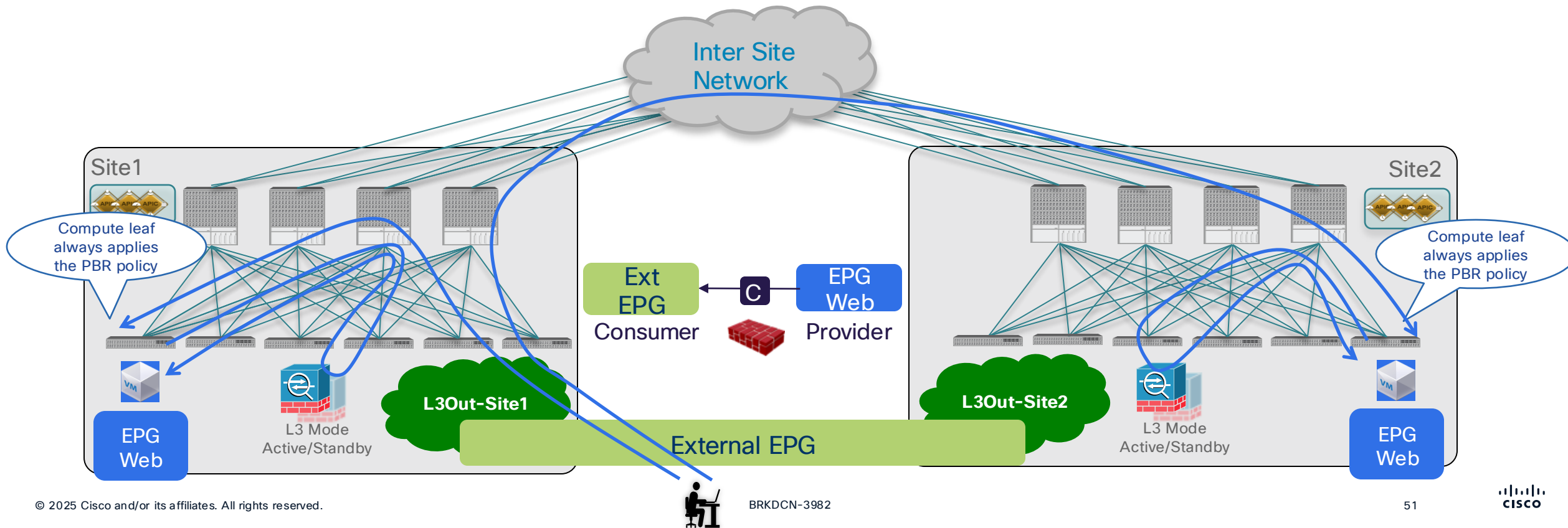
ACI Multi-Site: service nodes in each site

North-South Traffic Flow: compute leaf enforcement



Policy is always applied on the compute leaf

- North-South (L3Out-to-EPG) intra-VRF and inter-VRF contract with PBR
 - For inter-VRF contract, L3Out must be the provider.



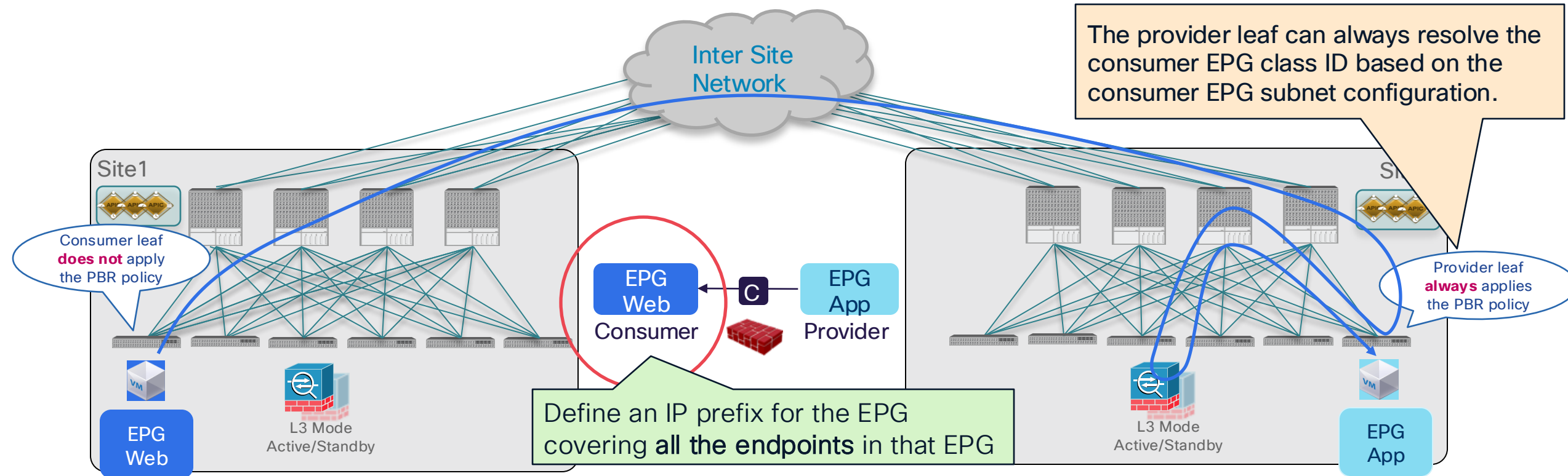
ACI Multi-Site: service nodes in each site

East-West Traffic Flow: provider leaf enforcement



Policy is always applied on the provider leaf

- East-West (EPG-to-EPG) intra-VRF and inter-VRF contract with PBR
 - **The consumer EPG subnet must be configured**, which means the design must be 1 BD subnet = 1 EPG (network centric).



How to ensure the provider leaf enforcement?

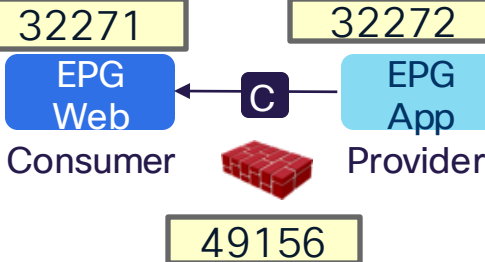
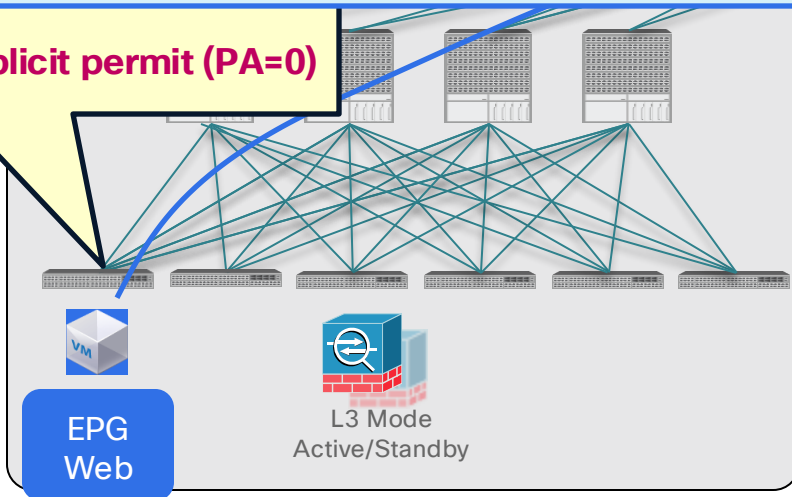
Special rule for consumer-to-provider traffic

- **redir_override**: If the destination is NOT a local endpoint, the leaf doesn't apply policy (PA=0)

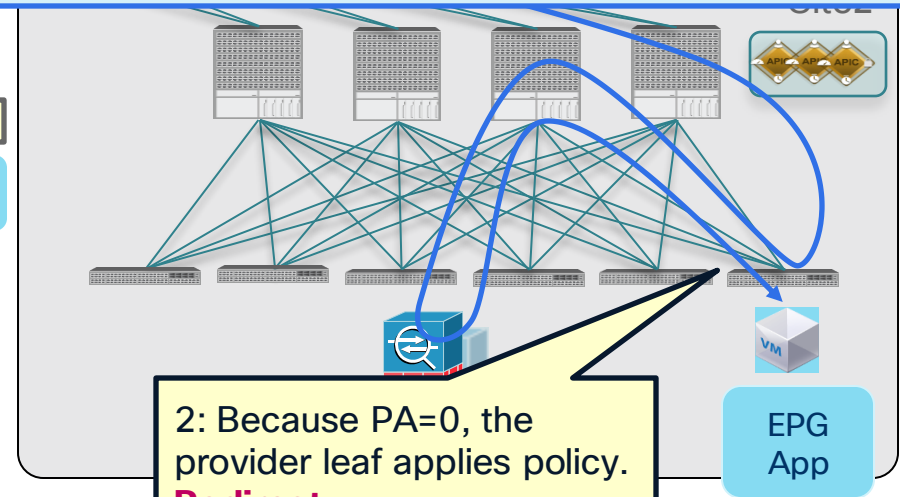
```
Pod1-Leaf1# show zoning-rule scope 2195459
```

Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
<snip>									
4144	32271	32272	14	bi-dir	enabled	2195459		redir(destgrp-1),redir_override	fully_qual(7)
4157	32272	32271	14	uni-dir-ignore	enabled	2195459		redir(destgrp-1)	fully_qual(7)
4140	49156	32272	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4136	49156	32271	14	uni-dir	enabled	2195459		permit	fully_qual(7)

1: Implicit permit (PA=0)



2: Because PA=0, the provider leaf applies policy.
Redirect



How to ensure the provider leaf enforcement?

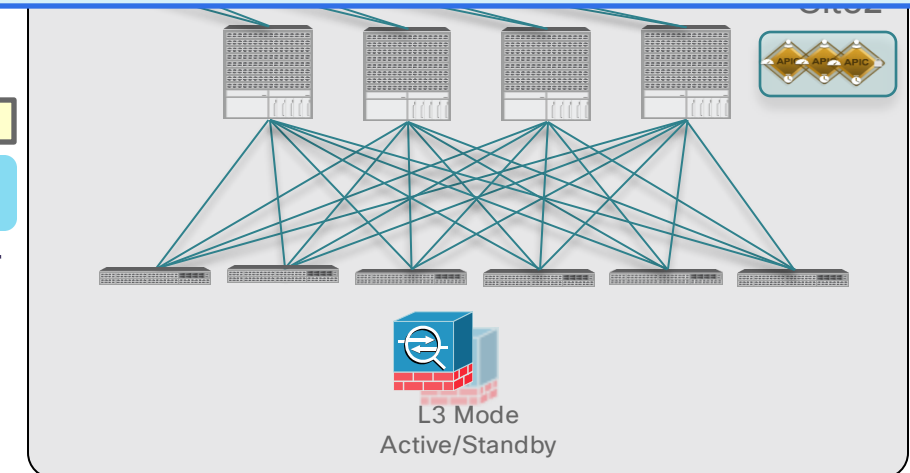
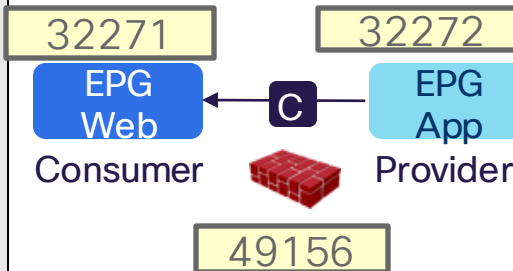
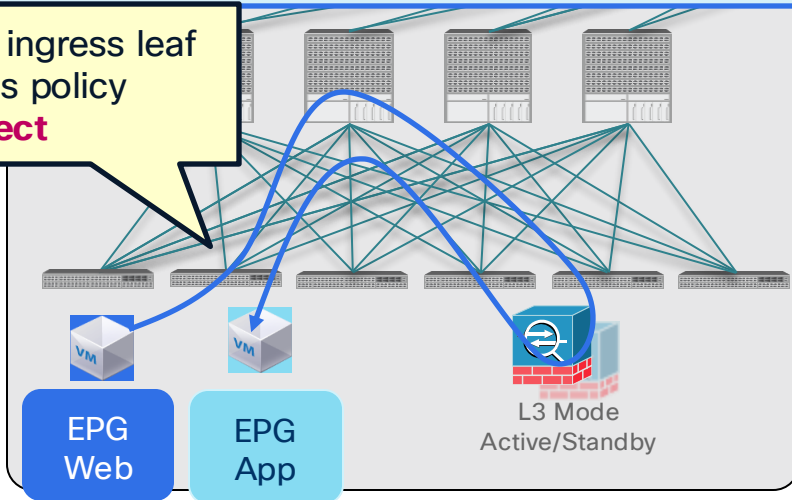
Special rule for consumer-to-provider traffic

- If the destination is under the same leaf, the leaf applies policy.

```
Pod1-Leaf1# show zoning-rule scope 2195459
```

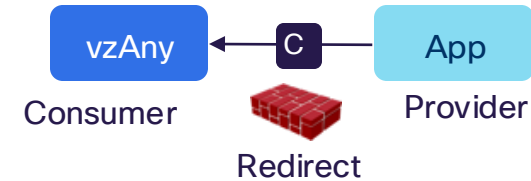
Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
<snip>									
4144	32271	32272	14	bi-dir	enabled	2195459		redir(destgrp-1),redir_override	fully_qual(7)
4157	32272	32271	14	uni-dir-ignore	enabled	2195459		redir(destgrp-1)	fully_qual(7)
4140	49156	32272	default	uni-dir	enabled	2195459		permit	src_dst_any(9)
4136	49156	32271	14	uni-dir	enabled	2195459		permit	fully_qual(7)

1: the ingress leaf applies policy
Redirect



Multi-Site PBR Update

- **vzAny-to-EPG**
- vzAny-to-vzAny
- Configuration workflow

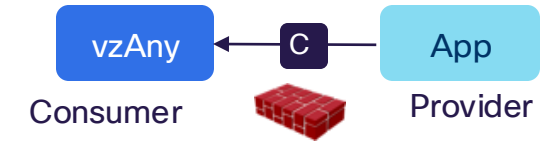


ACI Multi-Site vzAny-to-EPG PBR

NDO 4.2(3)/ACI 6.0(4)

Challenges

- How to keep traffic symmetric
 - Provider leaf enforcement
- How to ensure the provider leaf nodes can resolve destination class ID without EPG subnet.
 - Conversational learning

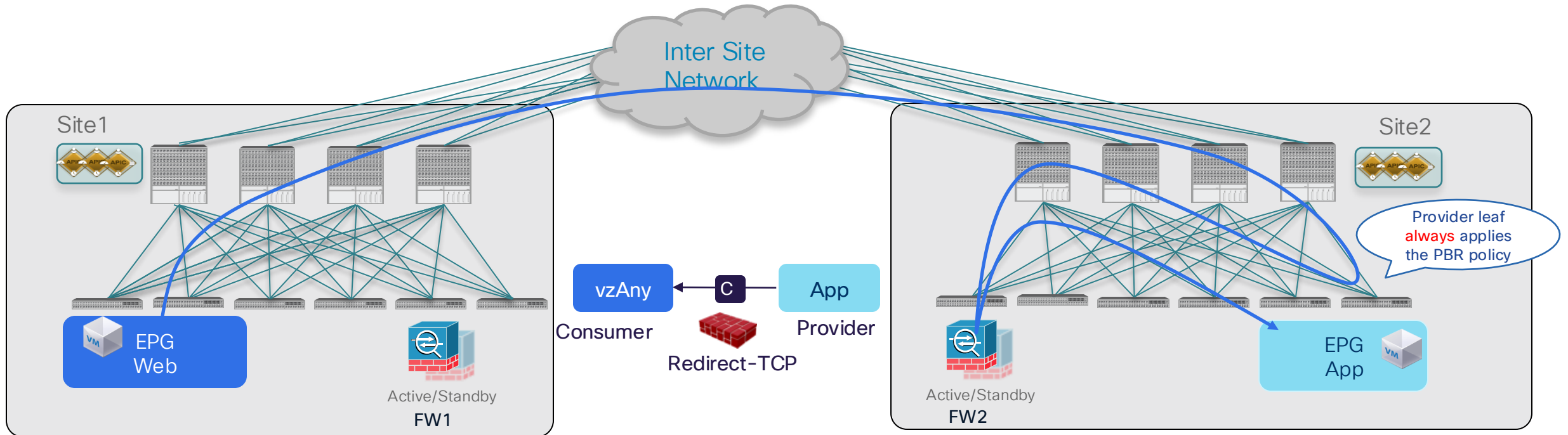


ACI Multi-Site vzAny-EPG PBR

NDO 4.2(3)/ACI 6.0(4)

Consumer to provider direction

- Provider leaf enforcement to keep traffic symmetric.



- Provider leaf enforcement to keep traffic symmetric.

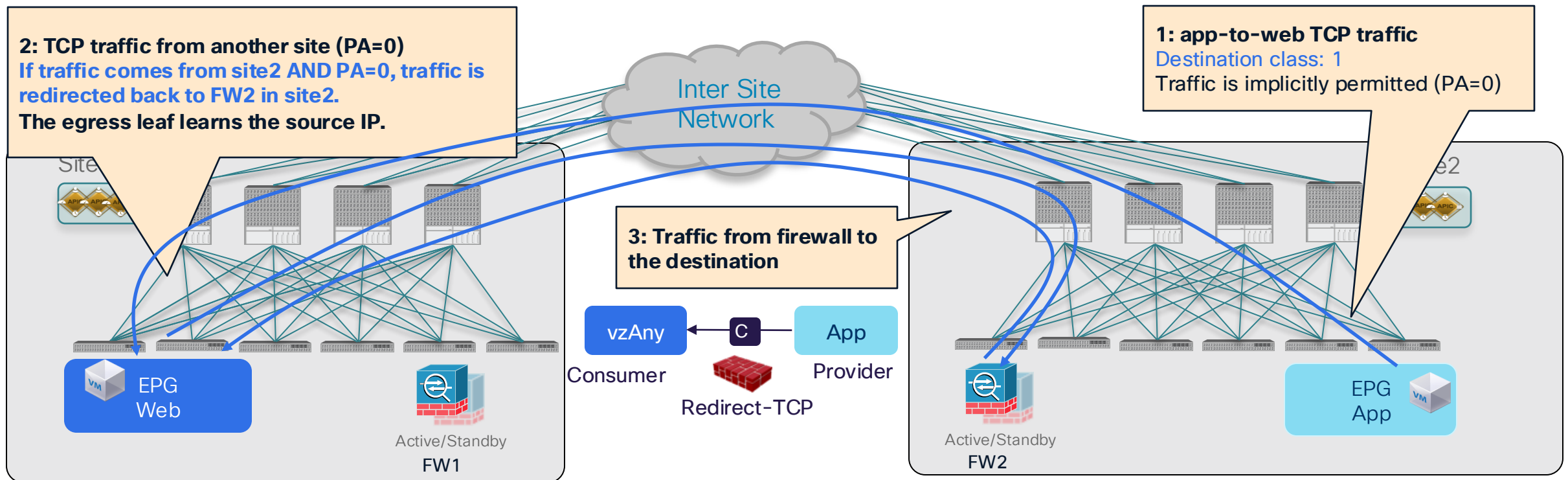


ACI Multi-Site vzAny-EPG PBR

NDO 4.2(3)/ACI 6.0(4)

What if the provider leaf doesn't know the consumer endpoint? (1/2)

- Force traffic inspected by the service device in the provider site

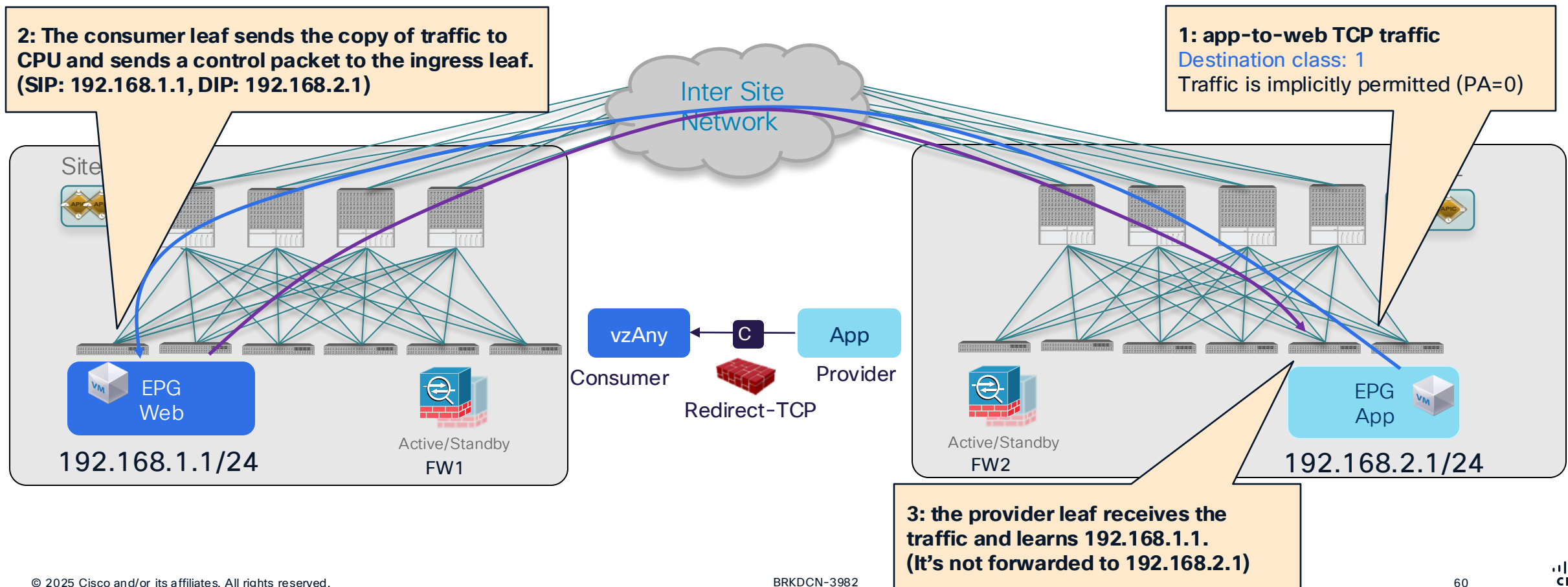


ACI Multi-Site vzAny-EPG PBR

NDO 4.2(3)/ACI 6.0(4)

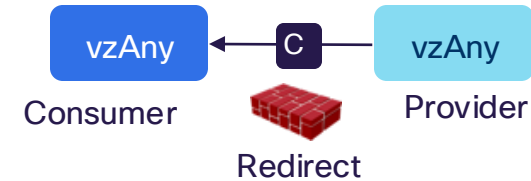
What if the provider leaf doesn't know the consumer endpoint? (2/2)

- Conversational Learning to get the ingress leaf learn the destination EP.



Multi-Site PBR Update

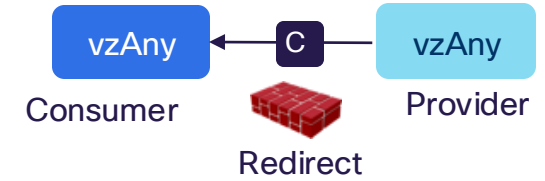
- vzAny-to-EPG
- **vzAny-to-vzAny**
- Configuration workflow



ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

Challenges



- How to keep traffic symmetric

→ redirect “inter-site” traffic in both source and destination sites.

Note: If it’s intra-site traffic, redirect doesn’t happen twice.

- How to ensure the ingress leaf nodes can resolve the destination class ID without the EPG subnet.

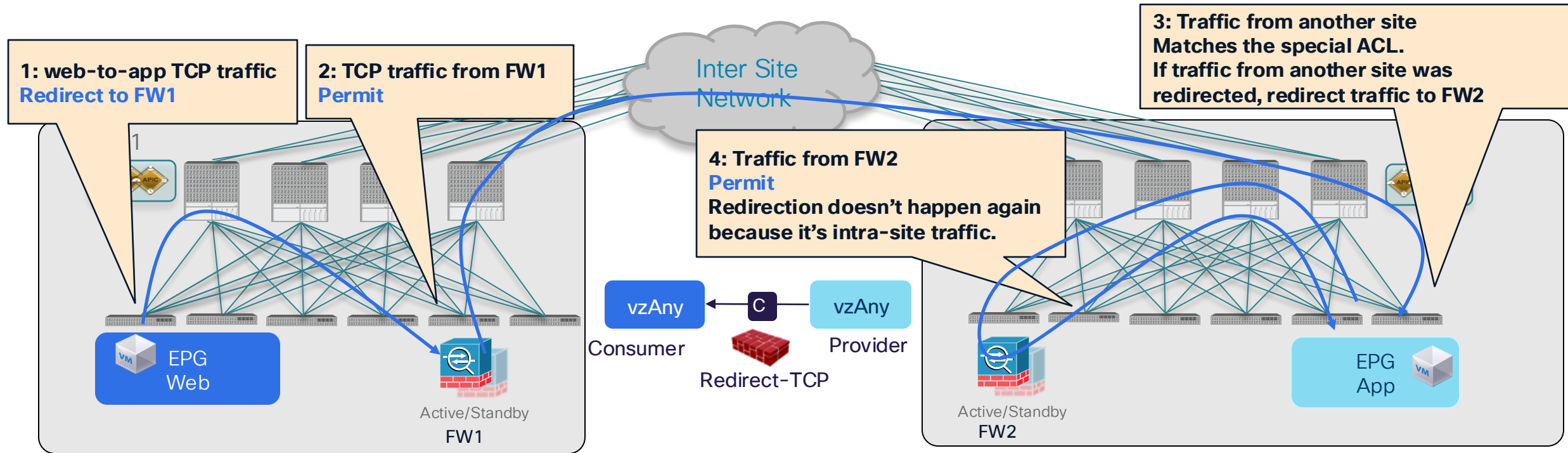
→ Conversational learning

ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

Consumer to provider direction

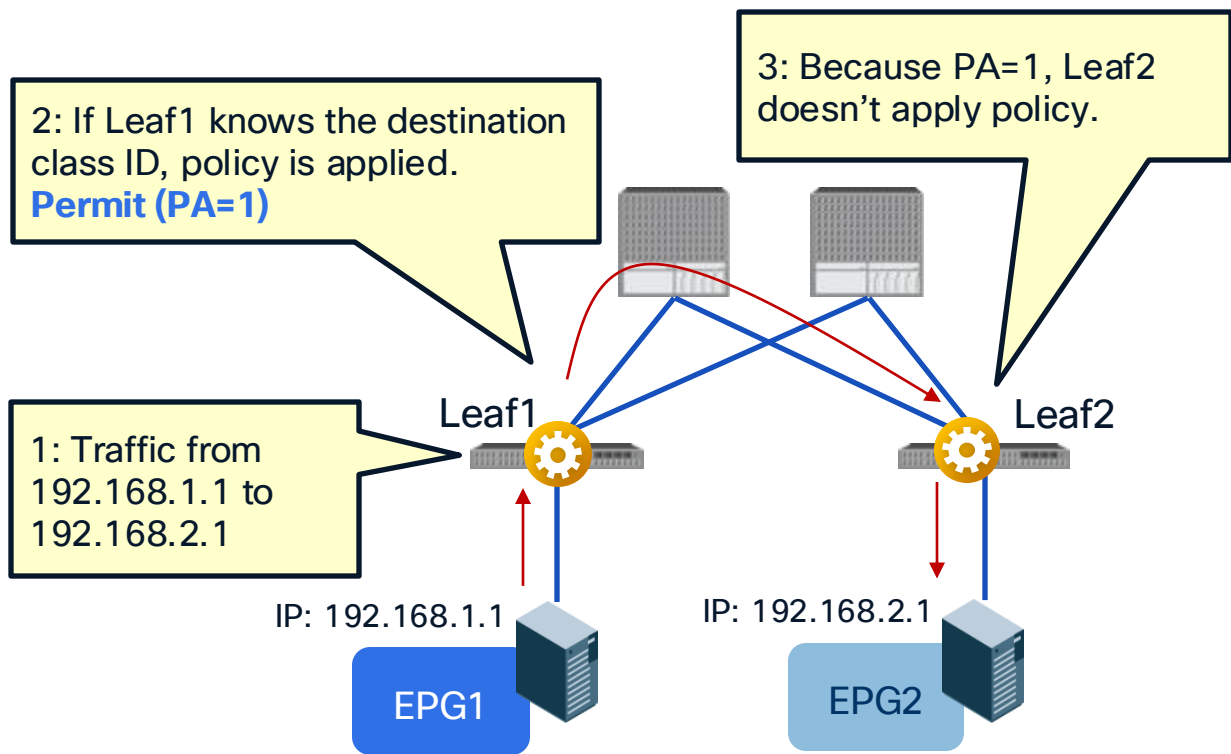
- Redirect “inter-site” traffic in both source and destination sites.



How to identify traffic was redirected?

Policy Applied (PA) bit

- PA bit (2 bit): Source Policy (SP) bit and Destination Policy (DP) bit



	SP	DP	Behavior
PA=1	1	1	The egress leaf doesn't apply policy because policy was applied.
PA=0	0	0	The egress leaf should apply policy because policy is not applied yet.



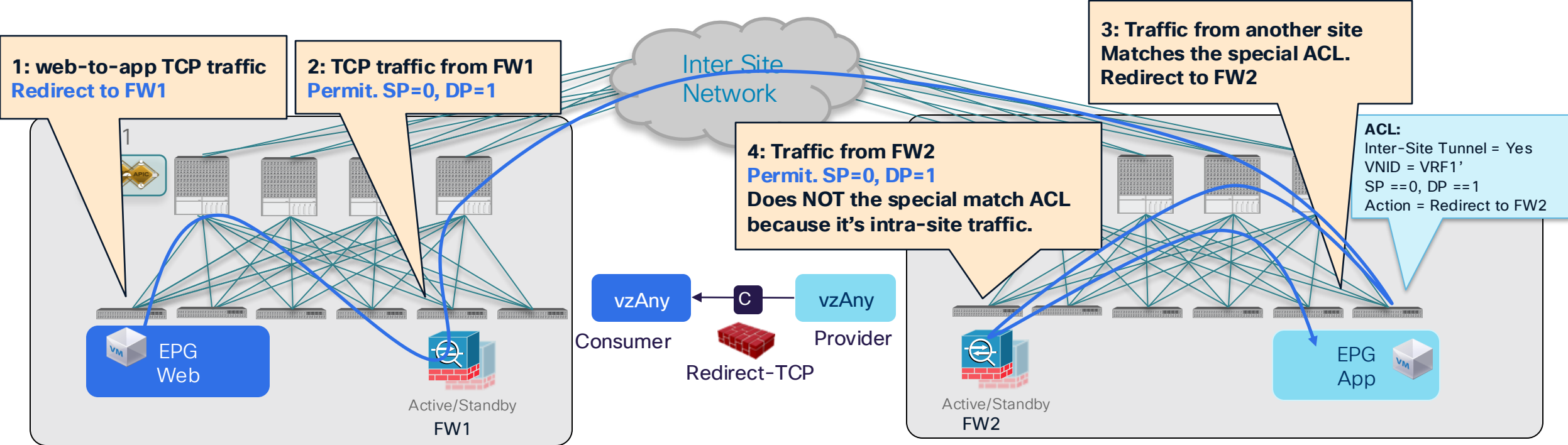
“SP=0, DP=1” is used for traffic from service EPG to indicate traffic needs to be redirected again

ACI Multi-Site vzAny-to-vzAny PBR

Consumer to provider direction



SP=0, DP=1
for traffic from
the service EPG

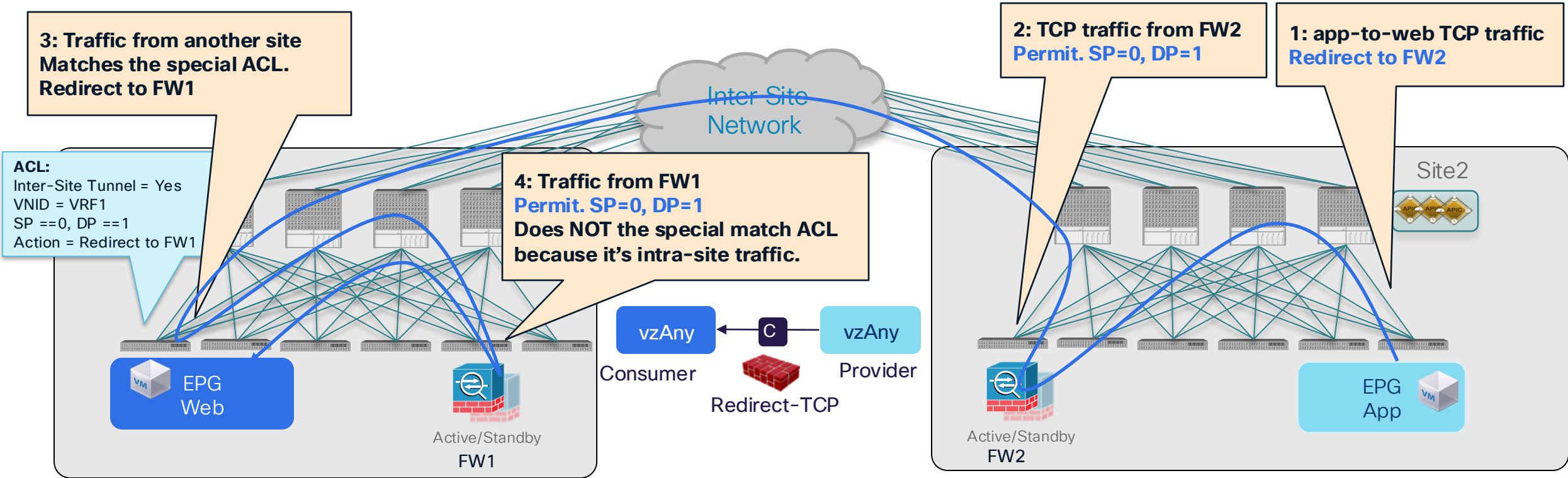


ACI Multi-Site vzAny-to-vzAny PBR

Provider to consumer direction



SP=0, DP=1
for traffic from
the service EPG

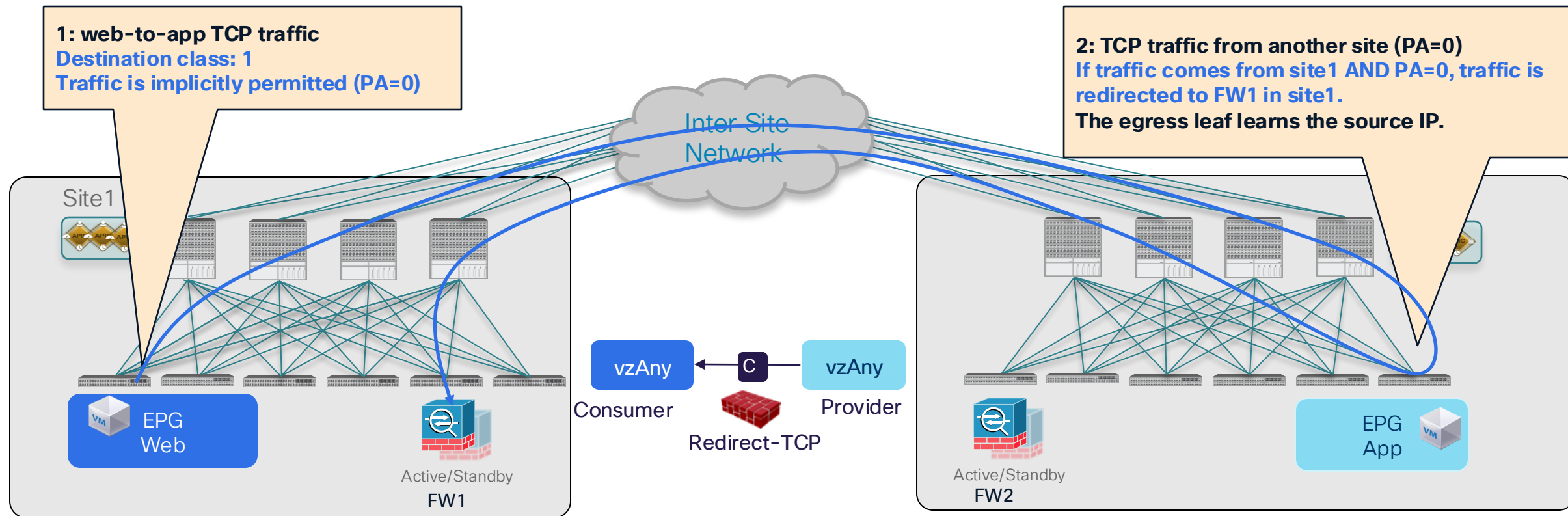


ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

What if the ingress leaf doesn't know the destination class ID (1/3)

- Force traffic inspected by the service device in the source site.

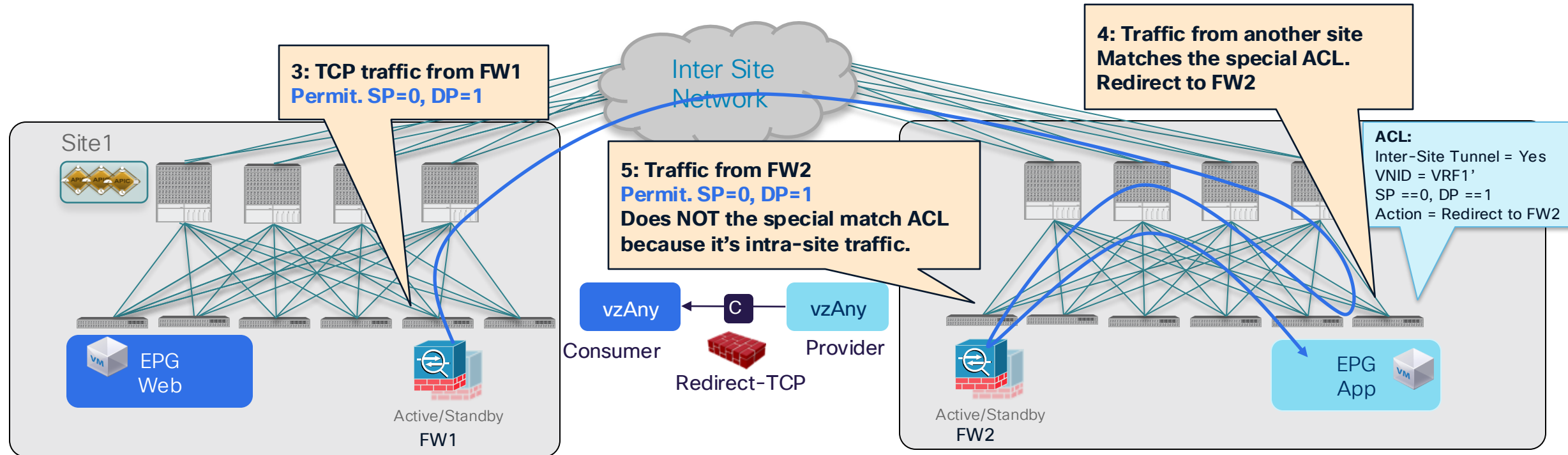


ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

What if the ingress leaf doesn't know the destination class ID (2/3)

- Force traffic inspected by the service device in the destination site

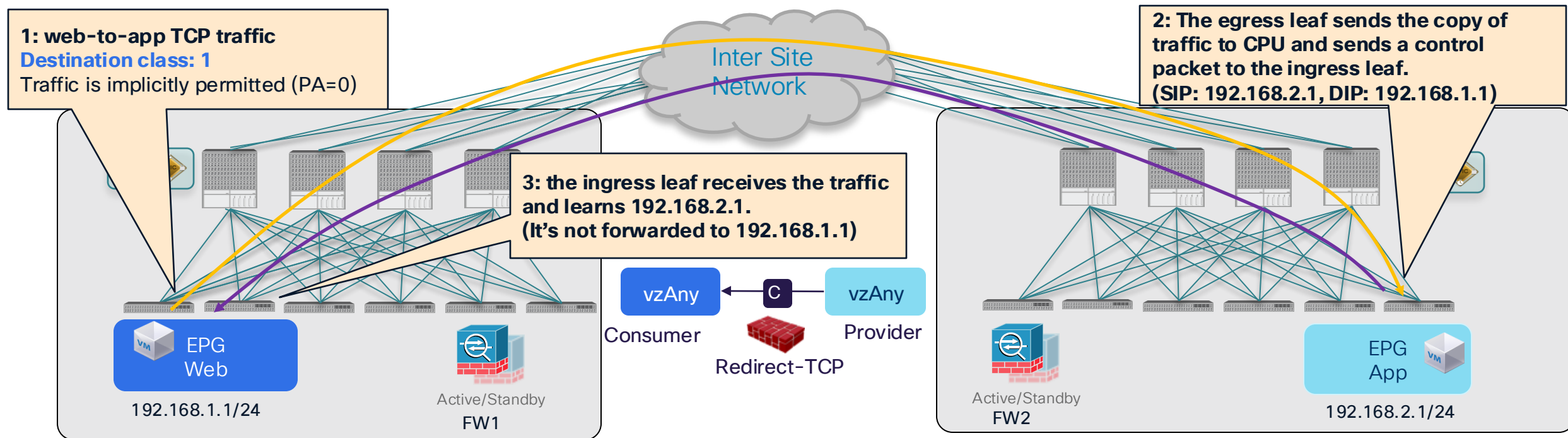


ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

What if the ingress leaf doesn't know the destination class ID (3/3)

- Conversational Learning to get the ingress leaf learn the destination EP.

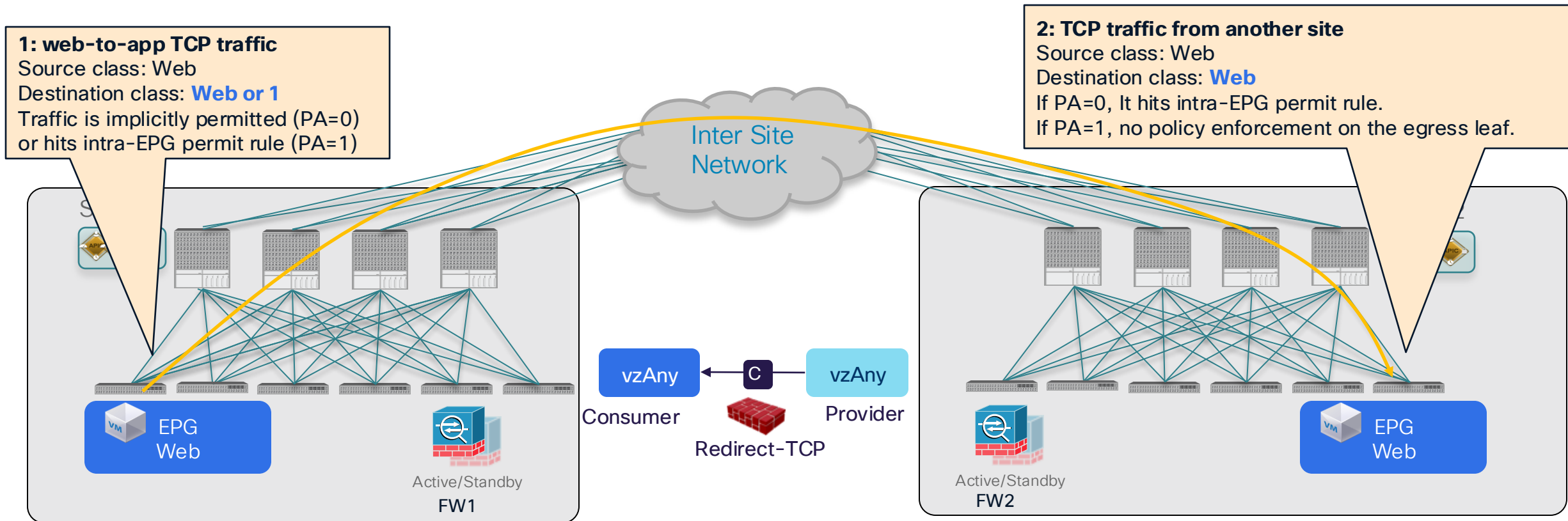


ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

Intra-EPG traffic

- Intra-EPG permit rule (priority 3) wins over vzAny-to-vzAny rule (priority 17).

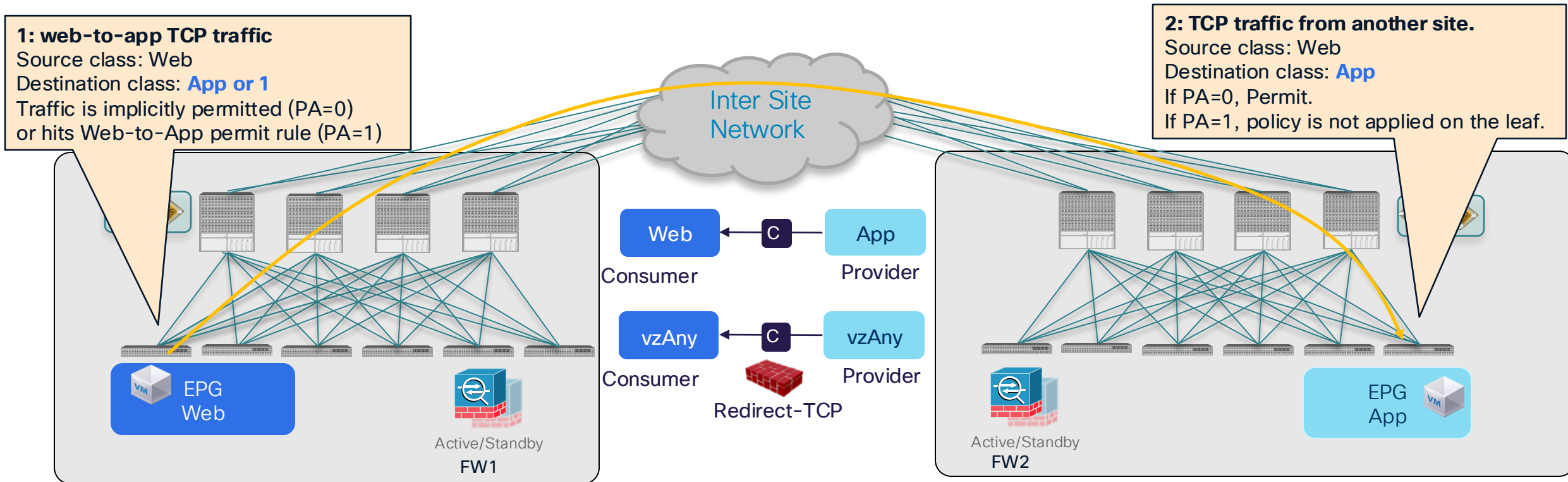


ACI Multi-Site vzAny-to-vzAny PBR

NDO 4.2(3)/ACI 6.0(4)

Bypass firewall for specific EPG-to-EPG traffic

- EPG-to-EPG permit rule (priority 7 or 9) wins over vzAny-to-vzAny rule (priority 17).



ACI Multi-Site

vzAny PBR and L3Out-to-L3Out PBR

	vzAny-to-vzAny	vzAny-to-EPG	vzAny-to-L3Out	L3Out-to-L3Out
Redirection	Both sites	Site for the specific EPG	Both sites	Both sites
Service node	One-node One-arm	One-node One-arm	One-node One-arm	One-node One-arm
		Starting from 6.1(3) Two-node Two-arm		
VRF	Intra-VRF	Intra-VRF	Intra-VRF	Intra-VRF and Inter-VRF

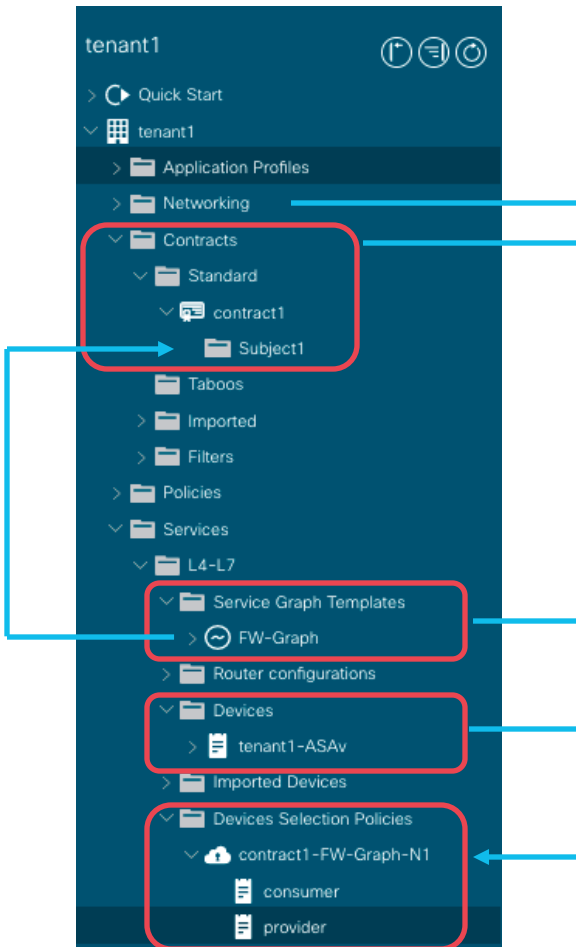
- No need to configure EPG subnets.
- By configuring specific EPG-to-EPG permit contract, firewall can be bypassed.
- Each site needs to have PBR destination with decent high availability within the site.
- ESG with PBR is not supported in Multi-Site (6.1(4) roadmap)

Multi-Site PBR Update

- vzAny-to-EPG
- vzAny-to-vzAny
- Configuration workflow

NEW

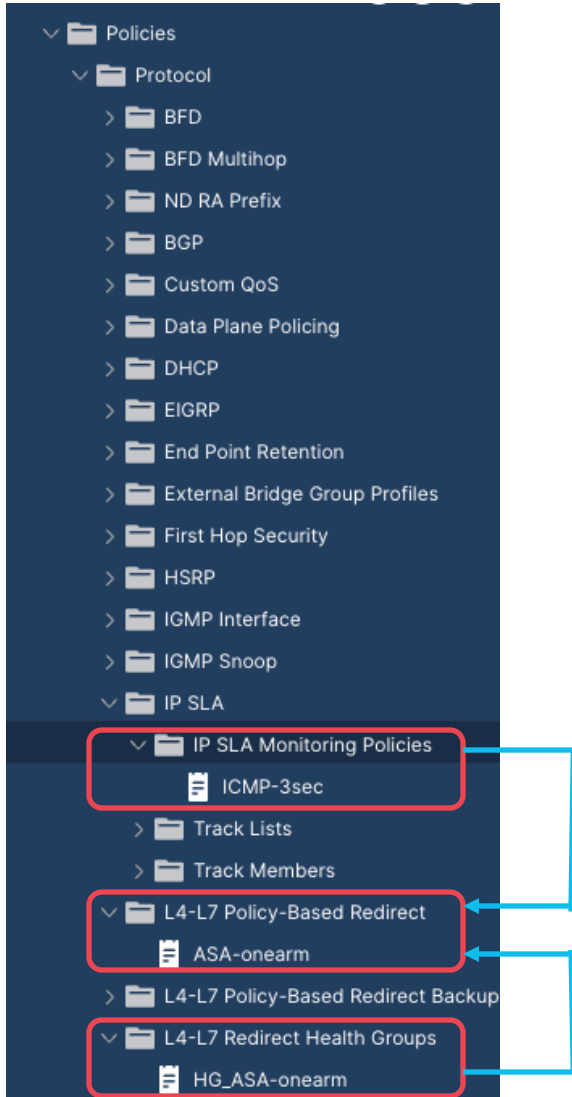
Recap: Configuration for Service Graph



- Contract
- Service Graph template
 - Service Graph template is attached to a contract subject
- L4-L7 Device
 - Physical domain (static path) or VMM domain (VM name and interfaces)
 - Cluster interfaces
- Device Selection Policy
 - It's based on
 - Contract name
 - Service Graph template name
 - Node name in the Service Graph

Recap: Configuration for PBR

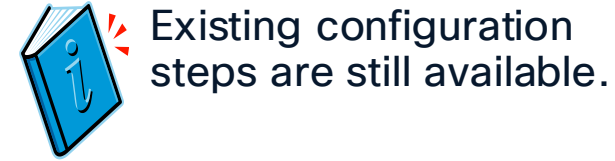
PBR requires additional configurations



- L4-L7 PBR Policy
 - PBR destination IP (and MAC)
 - PBR related options (hash option, resilient hash etc)
 - Tracking configuration (IP-SLA, Health-group)
- IP-SLA policy (optional)
 - Protocol: ICMP/TCP/HTTP/L2Ping
 - Interval etc
- L4-L7 Redirect Health-group (optional)

Multi-Site L4-L7 configuration

Previous and new L4-L7 Configuration steps



- Previous L4-L7 configuration steps: APIC local config + NDO config

APIC config (each site)

APIC Admin



1. Configure Tracking options (optional)
 - IP-SLA policy
 - Health-Group
2. Create a PBR policy
3. Create a L4-L7 Device

NDO config

NDO Admin



1. Create a Service Graph template
2. Attach the Service Graph template to a contract
3. Select the cluster interface, BDs and PBR policies required for Device Selection Policy on APIC

- New L4-L7 configuration steps: NDO config ONLY

NDO config

NDO Admin



1. Configure an IP-SLA policy (optional)
2. Configure a Service Device template
3. Insert the Service Device to a contract

vzAny-to-vzAny PBR configuration Example (Video)

- Overview
- Operate
- Configure**
- Admin

Configure / Tenant Templates

Tenant Template

Refresh Audit Logs

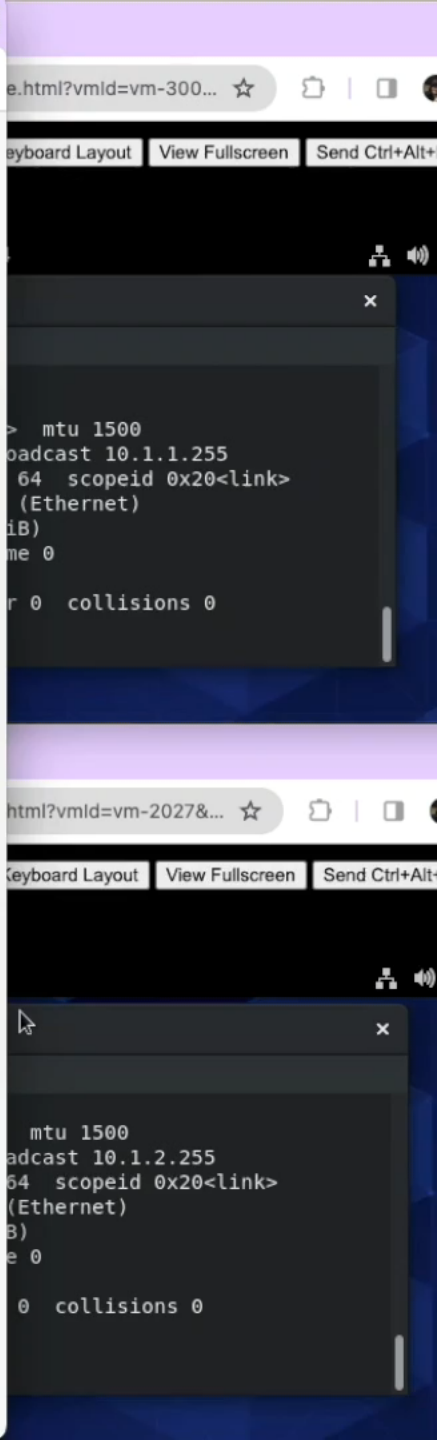
Applications L3Out Monitoring Policies Service Device Tenant Policies

Filter by attributes Add Schema

Name	Templates	Tenants	Policies	
Max-Schema	4  4	1	18	...
ServiceChaining	1  1	1	13	...

10 Rows

Page 1 of 1 << < 1-2 of 2 >> >>



Conclusion

Summary

- How ACI PBR works, use cases and design tips
- Flexible traffic redirection
 - Redirect specific traffic based on contract.
 - Intra-subnet and intra-EPG/ESG redirection
 - Any-to-Any, Any-to-EPG/ESG redirection
- Scale easily
 - Symmetric PBR with tracking and resilient hash
 - PBR destinations can be L1/L2/L3 devices anywhere in the fabric
- Multi-Location Data Centers
 - Multi-Site vzAny PBR is available!
 - New L4-L7 configuration workflow on NDO
 - Multi-Site ESG PBR will be available in 6.1(4)
- For more information, please check ACI PBR white paper!

Complete your session evaluations



Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to win 1 of 5 full conference passes to Cisco Live 2026.



Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.



Level up and earn exclusive prizes!



Complete your surveys in the Cisco Live mobile app.

Continue your education



Visit the Cisco Showcase for related demos.



Book your one-on-one Meet the Engineer meeting.



Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs.



Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Thank you

CISCO Live !





Appendix:

- Useful Links**
- NDO Configuration UI**

Useful Links

- Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper
 - <https://www.cisco.com/c/en/us/solutions/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739971.html>
- Cisco ACI Contract Guide
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743951.html>
- Service Graph Design with Cisco ACI (Updated to Cisco APIC Release 5.2) White Paper
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-2491213.html>
- ACI Fabric Endpoint Learning White Paper
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html>

Useful Links

- Cisco ACI and F5 BIG-IP Design Guide White Paper
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743890.html>
- Cisco ACI Multi-Pod and Service Node Integration White Paper
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739571.html>
- Cisco ACI Multi-Site and Service Node Integration White Paper
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-743107.html>
- ACI Multi-Site/Multi-Pod and F5 BIG-IP Design Guide
 - <https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/aci-multi-site-pod-f5-ip-design-guide.html>

Multi-Site L4-L7 configuration

1: Configure an IP-SLA policy (optional)

- Tenant Policy Template
 - NDO -> Configure -> Tenant Template -> Tenant Policies
 - -> Create Object -> Create an IPSLA Monitoring Policy

The screenshot displays the Cisco Nexus Dashboard Orchestrator interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', and 'Orchestrator'. The left sidebar shows navigation options: Overview, Operate, Configure (highlighted with a red box), and Admin. The main content area is titled 'Tenant Policies' and includes a breadcrumb path: 'Configure / Tenant Templates [Tenant Policies] / PBR-Tenant-Policies' (highlighted with a red box). Below this, there are tabs for 'Template Properties', 'Site1', and 'Site2'. The 'Template Properties' tab is active, showing a 'Template Summary' table with columns: Type, Tenant, Template Status, Associated Sites, and Last Action. The table contains one row: 'Tenant Policy Template', 'PBR', 'Out Of Sync', and '2' (highlighted with a red circle). The 'Associated Sites' column shows 'In Sync 0' and 'Out of Sync 2'. The 'Last Action' column shows 'Updated'. Below the table is a 'Filter' input field. To the right of the table are buttons for 'Edit Template', 'Deploy Template', and 'Actions'. At the bottom right of the main content area is a 'Create Object' button (highlighted with a red box). On the right side of the interface, there is a list of policy types: 'Route Map Policy for Multicast', 'Route Map Policy for Route Control', 'Custom QoS Policy', 'DHCP Relay Policy', 'DHCP Option Policy', 'IGMP Interface Policy', 'IGMP Snooping Policy', 'MLD Snooping Policy', 'L3Out Node Routing Policy', 'L3Out Interface Routing Policy', 'BGP Peer Prefix Policy', 'IPSLA Track List', and 'IPSLA Monitoring Policy' (highlighted with a red box). To the right of this list is a form for 'ICMP-3-sec'. The form includes fields for 'Name' (ICMP-3-sec), 'SLA Type' (ICMP), 'SLA Frequency (sec)' (1), 'Detect Multiplier' (3), 'Req Data Size (bytes)' (28), 'Type of Service' (0), 'Operation Timeout (milliseconds)' (900), 'Threshold (milliseconds)' (900), and 'IPv6 Traffic Class' (0). The 'SLA Type' field is highlighted with a red box.

Tenant Policies

Configure / Tenant Templates [Tenant Policies] / PBR-Tenant-Policies

Template Properties Site1 Site2

Template Summary

Type	Tenant	Template Status	Associated Sites	Last Action
Tenant Policy Template	PBR	Out Of Sync	2 In Sync 0 Out of Sync 2	Updated

Filter

IMPORT SELECT Create Object

Policy List:

- Route Map Policy for Multicast
- Route Map Policy for Route Control
- Custom QoS Policy
- DHCP Relay Policy
- DHCP Option Policy
- IGMP Interface Policy
- IGMP Snooping Policy
- MLD Snooping Policy
- L3Out Node Routing Policy
- L3Out Interface Routing Policy
- BGP Peer Prefix Policy
- IPSLA Track List
- IPSLA Monitoring Policy

ICMP-3-sec

Name * ICMP-3-sec

SLA Type ICMP

SLA Frequency (sec) 1

Detect Multiplier 3

Req Data Size (bytes) 28

Type of Service 0

Operation Timeout (milliseconds) 900

Threshold (milliseconds) 900

IPv6 Traffic Class 0

Multi-Site L4-L7 configuration

2: Configure a Service Device template (1/2)

- Service Device template: PBR policy + L4-L7 device network config at one
- NDO -> Configure -> Tenant Template -> Service Device -> Create Service Device Template

FW-OneArm

Common Properties

Name *

FW-OneArm

Device Location

ACI On-Prem Cloud

Device Type

Firewall Load Balancer Others

Device Mode

L3 L2 L1

Options that are not applicable are automatically grayed out. For example, if it's L1/L2 PBR, Device Type must be "Others"

Connectivity Mode

One Arm Two Arm Advanced

Interface Properties

Interface Name *

one-arm

Interface Type

BD L3Out

BD *

BD-Services X

Redirect

Yes No

IP SLA Monitoring Policy ⓘ

ICMP-3-sec X

Advanced Settings

Enabled Disabled

If it's Two Arm or Advanced, a table will show up and then each interface configuration can be done by clicking the pencil icon

Connectivity Mode

One Arm Two Arm Advanced

Interface Properties

Interface Name	Type	Redirect	IPSLA	
Internal	BD	No	-	
External	BD	No	-	

[Create Interface](#)

By default, other configuration options are hidden

Multi-Site L4-L7 configuration

2: Configure a Service Device template (2/2)

- New workflow hides Advanced configuration options unless it's required.

The diagram illustrates the configuration workflow for a Service Device template. It shows two main sections of the configuration interface:

General configuration options (highlighted in an orange box):

- Advanced Settings: Enabled / Disabled
- QoS Policy: Select QoS Policy >
- Preferred Group: ☐
- Load Balancing Hashing: Source IP, Destination IP and Protocol Number
- Pod Aware Redirection: ☐
- Anycast: ☐
- Rewrite Source MAC: ☐
- Advanced Tracking Options: Enabled / Disabled

PBR related configuration options (highlighted in an orange box):

- Advanced Tracking Options: Enabled / Disabled

Advanced Tracking Options (highlighted in an orange box):

- Advanced Tracking Options: Enabled / Disabled
- Static MAC configuration: ☐
- Tag Based Sorting: ☐
- Resilient Hash: Enabled / Disabled
- Threshold for Redirect Destinations: Enabled / Disabled

Resilient Hash (highlighted in an orange box):

- Resilient Hash: Enabled / Disabled
- Backup Redirects IP(s): ☐
- Threshold for Redirect Destinations: Enabled / Disabled
- Min Threshold for Redirect IP: 0
- Max Threshold for Redirect IP: 100
- Threshold Down Action: Permit

If Advanced Tracking option is enabled, more configuration options are shown

Multi-Site L4-L7 configuration

2: Configure a Service Device template for site level (1/3)

- Domain (physical or virtual domain) configuration is per site configuration.
- Select a site -> Select the Service Device Cluster

The screenshot displays the Cisco Nexus Dashboard Orchestrator interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', and 'Orchestrator'. The left sidebar shows navigation options: Overview, Operate, **Configure** (highlighted with a red box), and Admin. The main content area is titled 'Service Device Template' and shows the breadcrumb 'Configure / Tenant Templates [Service Device] / FW-Service' (highlighted with a red box). Below this, the 'Template Properties' section shows 'Site1' selected (highlighted with a red box) and 'Site2' available. The 'Template Summary' section displays the following information:

Type	Tenant	Template Status	Associated Sites	Last Action
Service Device Template	PBR	Out Of Sync	2 (circled in red) In Sync: 0 Out of Sync: 2	Deployment Successful Last Deployed: Aug 21, 2023 04:27 am

Below the summary, there is a 'Filter' input field and a 'Service Device Cluster' dropdown menu with 'FW-OneArm' selected (highlighted with a red box). The right sidebar contains buttons for 'Refresh', 'Audit Logs', 'Save', 'Open Site', 'Edit Template', and an 'Actions' dropdown.

Multi-Site L4-L7 configuration

2: Configure a Service Device template for site level (2/3)

- Physical domain

Service Device Cluster FW-OneArm on Site1

View Relationship

Common Properties

Interface Properties

Site Properties

Domain Type *

Physical VMM

Domain*

phys

Encap ranges: 56-56, 100-101, 102-102, 300-350, 351-400

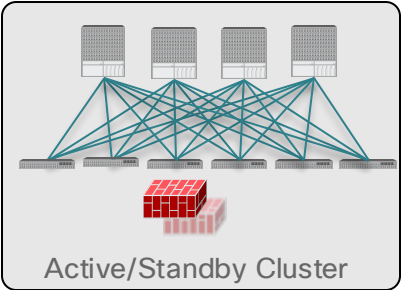
INTERFACE 1

Interface Name

one-arm

VLAN *

100



Specify the interfaces connected to FW nodes (Active/Standby mode)

For a physical domain:

- VLAN is mandatory (static allocation)

For a physical domain:

- Two Interfaces connected to the Active/Standby service devices (static path)

Specify the single IP address identifying the logical cluster

Fabric To Device Connectivity ⓘ			
Type *	Pod *	Node *	Path *
Virtual Port Channel	1	101,102	vPC-L101-L102-Port16
Virtual Port Channel	1	103,104	vPC-L103-L104-Port16
Add Fabric To Device Connectivity			
PBR Destinations			
IP Address *			
50.50.50.10			

Multi-Site L4-L7 configuration

2: Configure a Service Device template for site level (3/3)

- VMM domain

Service Device Cluster FW-Cluster on Site1 [View Relationship](#)

Common Properties

Interface Properties

Site Properties

Domain Type *

Physical VMM

Domain*

vDS-Site1 x

Encap ranges: 50-60, 100-110, 300-399, 480-480, 800-900

Trunking Port

☐ Enabled

Promiscuous Mode

☐ Enabled

INTERFACE 1

Interface Name

one-arm

VLAN

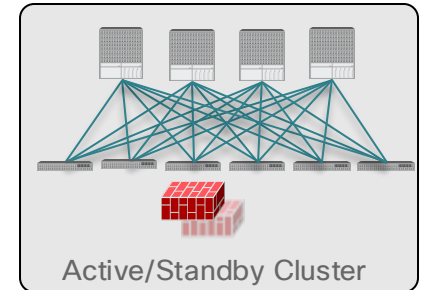
Enhanced LAG Option

LAG1

Select Domain Type and a domain

For a VMM domain:

- VLAN is not mandatory (dynamic allocation). If VLAN ID is specified, the VLAN ID must be part of a static VLAN range
- Enhanced LAG



Specify the FW VMs (Active/Standby mode)

Specify the single IP address identifying the logical cluster

VM Information* ⓘ

VM Name*	VNIC*
vCSA-7-Site1/ASAv-Pod1	Network adapter 2
vCSA-7-Site1/ASAv-Pod2	Network adapter 2

+ Add VM Information

PBR Destinations

IP Address *

50.50.50.10

For a VMM domain:

- VM Name and its interface
- PBR destination IP (If IP-SLA tracking is enabled, MAC configuration is not required)

Multi-Site L4-L7 configuration

Note:

- New workflow doesn't ask you some configuration options if they are not required. For example:
 - If tracking is enabled, NDO doesn't ask PBR destination MAC.
 - NDO doesn't ask Health-group configuration unless it's required.

The screenshot displays the Cisco ACI NDO configuration interface. On the left, the 'VM Information*' section shows two VMs: 'ACI-vDS-vcenter/Site1-ASAv-1' and 'ACI-vDS-vcenter/Site1-ASAv-2', both connected to 'Network adapter 5'. Below this, the 'PBR Destinations' section lists two IP addresses: '192.168.1.101' and '192.168.1.102'. A callout box points to these destinations, stating: 'If it's one-arm FW, NDO doesn't ask Health-group even though there are multiple PBR destinations.'

The central pane shows the 'NDO-L4-L7' policy tree. The 'L4-L7 Policy-Based Redirect' folder is expanded, showing sub-items like 'FW-onearm-one-arm', 'FW-two-arm-External', 'FW-two-arm-Internal', 'L4-L7 Policy-Based Redirect Backup', 'L4-L7 Redirect Health Groups' (highlighted with a red box), 'FW-onearm--ndo--implct--192.168.1.101', 'FW-onearm--ndo--implct--192.168.1.102', and 'FW-two-arm--ndo--implct'.

The right pane shows the 'L4-L7 Policy-Based Redirect - FW-onearm-one-arm' configuration. The 'Properties' section includes 'IP SLA Monitoring Policy: ICMP-3sec', 'Oper Status: Enabled', 'Threshold Enable: []', 'Enable Pod ID Aware Redirection: []', 'Hashing Algorithm: Destination IP', 'Source IP', and 'Source IP, Destination IP and Protocol num'. The 'L3 Destinations' table lists two destinations: '192.168.1.101' and '192.168.1.102', both with MAC address '00:00:00:00:00:00'. A callout box points to the 'Redirect Health Group' column, stating: 'NDO automatically configure Health-group'.

IP	Destination Name	MAC	Redirect Health Group
192.168.1.101		00:00:00:00:00:00	FW-onearm--ndo--implct--192.168.1.101
192.168.1.102		00:00:00:00:00:00	FW-onearm--ndo--implct--192.168.1.102

Multi-Site L4-L7 configuration

3: Insert the Service Device to a contract

- Just select which device you want to insert!
- NDO -> Configure -> Tenant Template -> Applications -> Select the Schema

The screenshot shows the Cisco Nexus Dashboard Orchestrator interface. The main panel displays the 'vzAny-to-vzAny' PBR Schema configuration. A red box highlights the 'Contracts' dropdown menu, which is set to 'vzAny-to-vzAny'. Another red box highlights the 'Service Chaining/Service Graph' section, showing a 'Service Chaining' button. A third red box highlights the 'Service Chaining' section, which contains a '+' icon. To the right, a 'Device Settings' modal is open, showing the 'Device Type' as 'Firewall', 'Device' as 'FW-OneArm', 'Consumer Interface' as 'one-arm', 'Consumer Connector Type' as 'Redirect', 'Provider Interface' as 'one-arm', and 'Provider Connector Type' as 'Redirect'. A red box highlights the 'Device Settings' modal. A yellow callout box at the bottom right states: 'If it's One-arm, the interface is automatically selected. Redirect can be enabled/disabled at each interface'.

Multi-Site L4-L7 configuration

Optional: required configuration for vzAny PBR

- Enable “L3 Multicast” and “Site-aware Policy Enforcement Mode” on the VRF

The screenshot shows the Cisco Nexus Dashboard interface for configuring VRF1. The left sidebar contains navigation tabs: Overview, Operate, Configure (selected), and Admin. The main content area is titled 'VRF1' and includes a 'PBR Schema' section. Under 'Template Properties', the 'On-Premises Properties' tab is active. It shows 'Policy Control Enforcement Preference' set to 'Enforced' and 'IP Data-Plane Learning' set to 'Enabled'. The 'L3 Multicast' checkbox is unchecked, and the 'Site-aware Policy Enforcement Mode' checkbox is also unchecked. A callout bubble points to these two settings with the text 'Both are disabled by default'.

RP is not required

The screenshot shows the 'Configure Rendezvous Points (RP)' section. The 'IP Address' field is empty. Below it, the 'vzAny' checkbox is unchecked. A callout bubble points to the 'RP is not required' text. The 'L3 Multicast' checkbox is checked, and the 'Site-aware Policy Enforcement Mode' checkbox is also checked.