# cisco live!







# ACI L4-L7 Policy-Based Redirect (PBR) Deep Dive and tips

Minako Higuchi
Technical Marketing Engineer, Cloud Networking Business Group
BRKDCN-3610



#### 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 17, 2022.



https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-3610

#### 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.
- What are not covered in this session.
  - Multi-Pod, Multi-Site, Remote Leaf and cloud ACI. We are going to focus on on-prem single pod.
- Initial assumption:
  - The audience already has a good knowledge of ACI main concepts: VRF, BD, EPG, ESG, L3Out, Contract etc

Note: This session uses ESGs mainly, but the PBR features should be applicable to EPGs and uSeg EPGs.





## Agenda

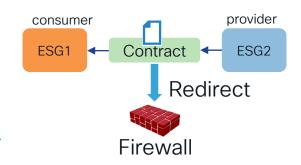
- ACI PBR Use cases
- PBR Forwarding and zoning-rules
- FAQ and Advanced use cases

## ACI PBR Use cases



# PBR (redirect) is one of the contract actions!

Permit, Deny, Redirect and Copy





#### Where can we use PBR?

Wherever contracts can be applied!

- Between EPGs or ESGs.
- Between L3Out EPGs.

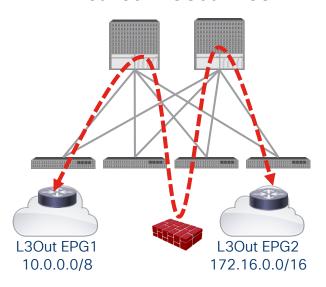
 Between EPGs or ESGs in the same subnet.

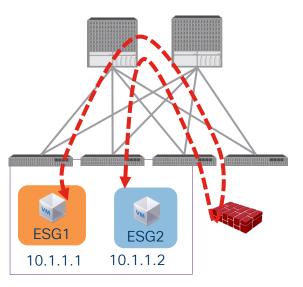


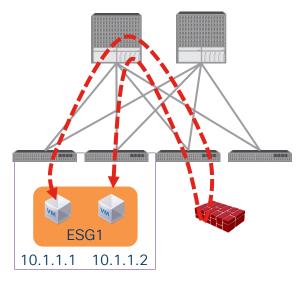
PBR is a contract action. It's based on source,

destination EPG/ESG and

filter matching.

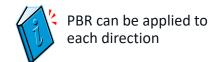








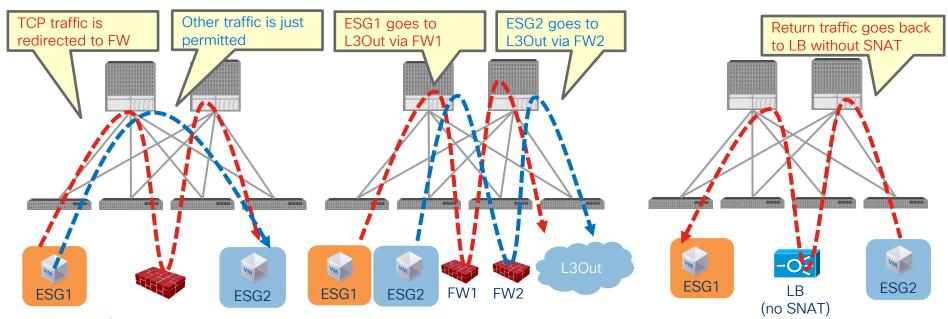
#### PBR use cases



Inspect specific traffic

Use different Firewall

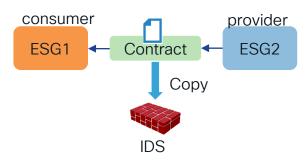
LB without SNAT (uni-directional PBR)

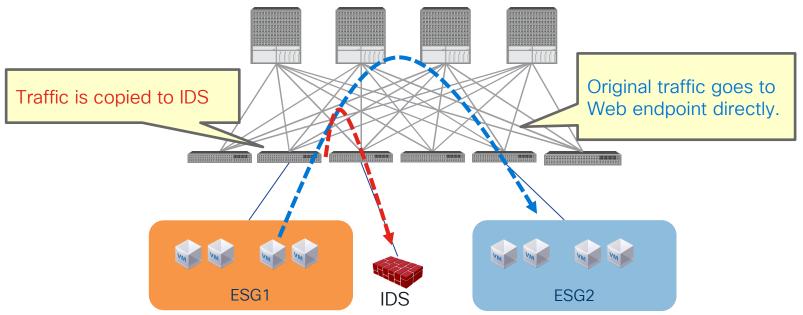




#### ACI Copy service

Copy specific traffic instead of redirect.



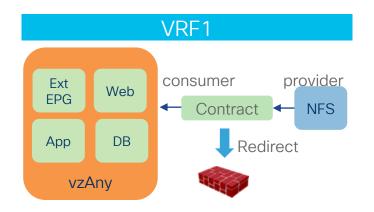


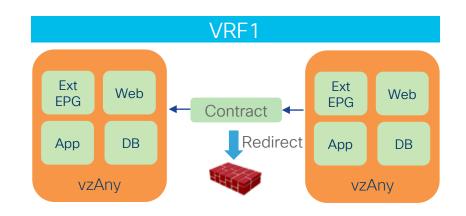


#### Important note



- ACI must be L3. (L2Out EPG is not supported)
- VRF must be 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)



		_	scope 21954		ı	ı	1	ı	+	
Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority	ĺ
<snip></snip>	,	T			T	,		,		
4157	29	10934	14	bi-dir	enabled	2195459	tenant1:contract1	permit	fully_qual(7)	
4144	10934	29	14				tenant1:contract1		fully_qual(7) +	- 1

With PBR (Service Graph)

29 consumer provider 10934

ESG1 contract1 ESG2

Service Graph(PBR)

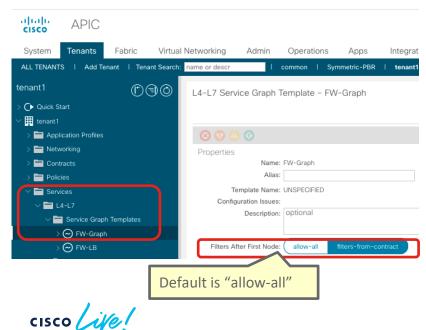
		-	scope 2195	459 +	+	+	+	30	32
Rule ID	SrcEPG	DstEPG	FilterID	Dir	operSt	Scope	Name	Action	Priority
snip>		+	+	+	+	+	+		+
4144	29	10934	14	bi-dir	enabled	2195459		redir(destgrp-11)	fully qual(7)
4157	10934	29	14	uni-dir-ignore	enabled	2195459		redir(destgrp-12)	fully qual(7)
4140	32	10934	default	uni-dir	enabled	2195459		permit	src dst any(9)
I 4136 I	30	1 29	14	uni-dir	enabled	2195459		permit	fully qual(7)

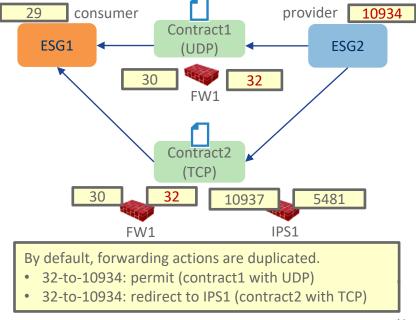
By default, unspecified default filter (any) is used for a zoning-rule entry without the consumer EPG.



#### 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.





#### PBR destination status

2: Periodic System-wide broadcast to all leaf nodes from the service leaf, announcing the FW's aliveness

1: Local tracking from the service leaf to node.

#### Health-group

If one of them is down, PBR to this node is disabled for both direction.

192.168.11.1 (destgrp-11)

192.168.12.1 (destgrp-12)

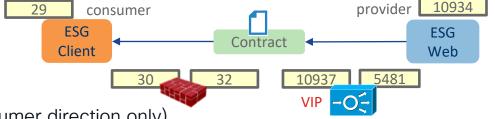
HILIANIAN IN THE STREET

LEGEND		h)   UD. Hashburfile		L DIC. Darkur D			
TL: Threshold(Low)	TH: Threshold(Hig	h)		BAC: Backup-D ======			: Resiliency
List of Dest Groups GrpID Name =====	destination		HG-name	BAC operSt	operStQual	TL TH	HP TRAC RES
11 destgrp-11 12 destgrp-12	dest-[192.168.11.1 dest-[192.168.12.1		tenant1::HG1 tenant1::HG1	N enabled N enabled	no-oper-grp no-oper-grp		sym yes no sym yes no
List of destination Name ====	s	bdVnid	vMac	vrf ====	operSt	operStQual	HG-name
dest-[192.168.11.1] dest-[192.168.12.1]	•	vxlan-16678782 vxlan-16121790	00:50:56:AF:6C:16 00:50:56:AF:DF:55	tenant1:VRF1 tenant1:VRF1	enabled enabled	no-oper-dest no-oper-dest	tenant1::HG1 tenant1::HG1
ist of Health Grou MG-Name	ps	HG-OperSt HG-Dest					HG-Dest-OperS
====== tenant1::HG1			= 192.168.11.1]-[vxlan 192.168.12.1]-[vxlan				up 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)



od1-Leaf1#	show zon	ning-rule	scope 219	Provider	er to provide to consum				
+ Rule ID	SrcEPG	•	+   FilterID		   operSt	· •	+   Name	•	+   Priority
+ snip>		+	+	-+	+	·	+	+	+
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)
	10937	29	14	uni-dir-ignore	enabled	2195459		redir(destgrp-12)	fully qual(7)
4181				uni-dir	enabled	2195459		permit	fully qual(7)

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



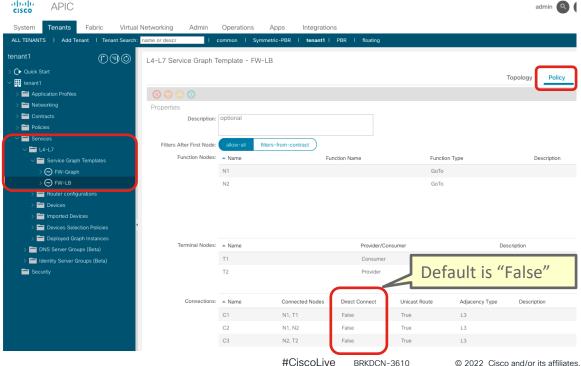


Direct Connect must be "True" for communication between the consumer/provider and the PBR destination.

## Direct Connect (False by default)

Tenant > Services > L4-L7 > Service Graph templates > Service Graph\_NAME

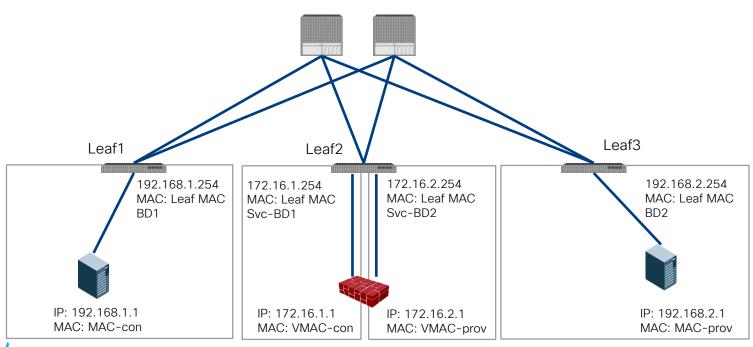
> Policy





#### How forwarding works

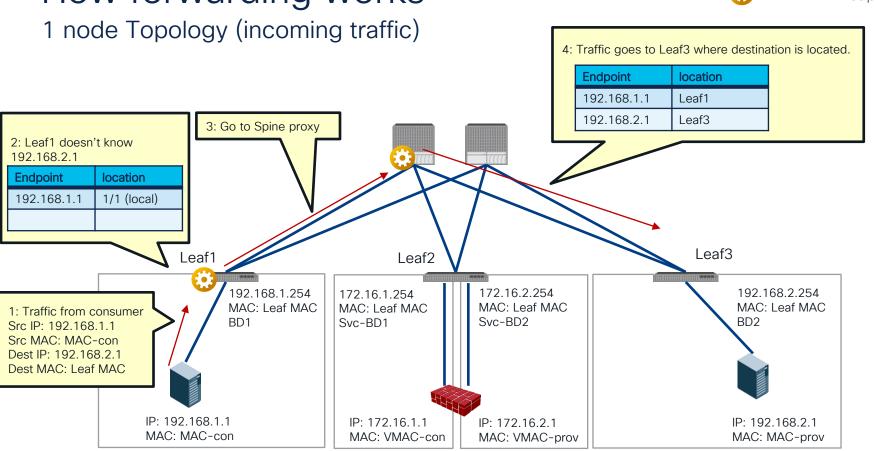
1 node Topology









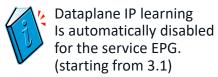


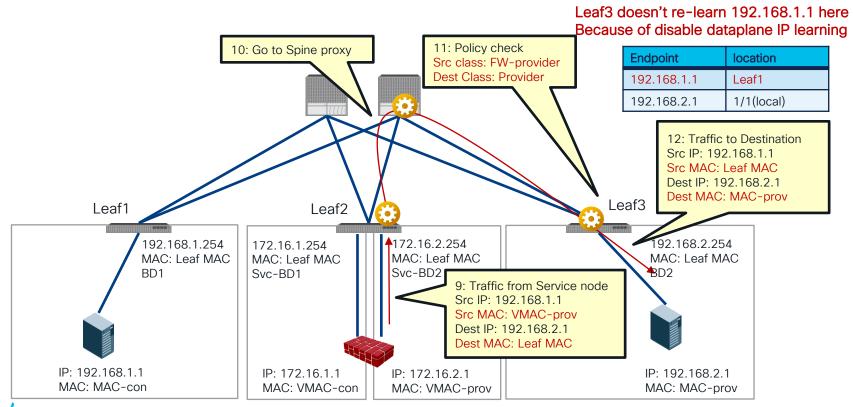
#### How forwarding works Leaf applied policy. It's always spine-proxy to reach the PBR destination 1 node Topology (incoming traffic) if the PBR destination is in a BD. 6: Policy applied (PBR) Src IP: 192.168.1.1 Dest IP: 192.168.2.1 Dest MAC: VMAC-con 7: Spine-proxy Segment ID: Svc-BD1 5: Policy check and Leaf3 learns 192.168.1.1 **Endpoint** location Src class: Consumer 8: Traffic to Service node Dest Class: Provider 192.168.1.1 Leaf1 Src IP: 192.168.1.1 Dest IP: 192.168.2.1 192.168.2.1 1/1(local) Dest MAC: VMAC-con Leaf3 Leaf1 Leaf2 172.16.2.254 192.168.1.254 172.16.1.254 192.168.2.254 MAC: Leaf MAC MAC: Leaf MAC MAC: Leaf MAC MAC: Leaf MAC Svc-BD2 BD2 BD1 Svc-BD1 IP: 192.168.1.1 IP: 172.16.1.1 IP: 172.16.2.1 IP: 192.168.2.1 MAC: MAC-con MAC: VMAC-con MAC: VMAC-prov MAC: MAC-prov

BRKDCN-3610

#### How forwarding works

1 node Topology (incoming traffic)





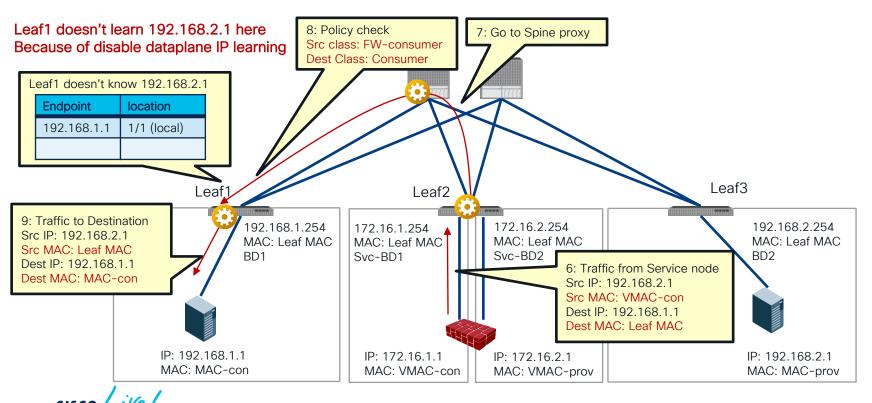
BRKDCN-3610

How forwarding works 1 node Topology (return traffic) 2: Policy check and Leaf3 knows 192.168.1.1 3: Policy applied (PBR) Src IP: 192.168.2.1 Src class: Provider Dest IP: 192.168.1.1 Dest Class: Consumer Dest MAC: VMAC-leg2 Segment ID: Svc-BD2 4: Spine-proxy location **Endpoint** 192.168.1.1 Leaf1 192.168.2.1 1/1(local) Leaf3 Leaf2 Leaf1 172.16.2.254 192.168.2.254 192.168.1.254 172.16.1.254 MAC: Leaf MAC MAC: Leaf MAC MAC: Leaf MAC MAC: Leaf MAC Svc-BD2 BD2 BD1 Svc-BD1 1: Traffic from provider 5: Traffic to Service node Src IP: 192.168.2.1 Src IP: 192.168.2.1 Src MAC: MAC-prov Dest IP: 192.168.1.1 Dest IP: 192.168.1.1 Dest MAC: VMAC-prov Dest MAC: Leaf MAC IP: 192.168.1.1 IP: 172.16.2.1 IP: 172.16.1.1 IP: 192.168.2.1 MAC: MAC-con MAC: VMAC-con MAC: VMAC-prov MAC: MAC-prov

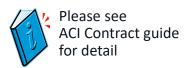
BRKDCN-3610

#### How forwarding works

1 node Topology (return traffic)



### Where is the policy applied?



Scenario	VRF enforcement mode	Consumer	Provider	Policy enforced on
Intra-VRF	Ingress/egress	EPG	EPG	<ul> <li>If destination endpoint is learned: ingress leaf*</li> <li>If destination endpoint is not learned: egress leaf</li> </ul>
	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> <li>If destination endpoint is learned: border leaf</li> <li>If destination endpoint is not learned: non-border leaf</li> <li>Non-border leaf-&gt; border leaf traffic</li> <li>Border leaf</li> </ul>
	Ingress/egress	L3Out EPG	L3Out EPG	Ingress 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



## Please see ACI Contract guide for detail

#### **Contract Priority**

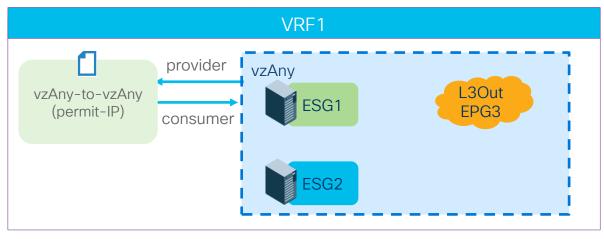
Look at your zoning-rule priority and then filter priority!

- 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 zoning-rule priority is 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?

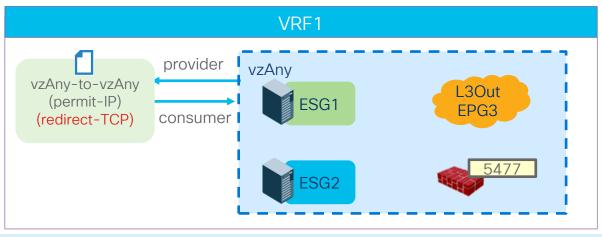


- ESG1-to-ESG2 (IP)
  Permit
- ESG1-to-L3OutEPG3 (IP)
  Permit
- ESG2-to-L3OutEPG3 (IP)
  Permit



### Example 2

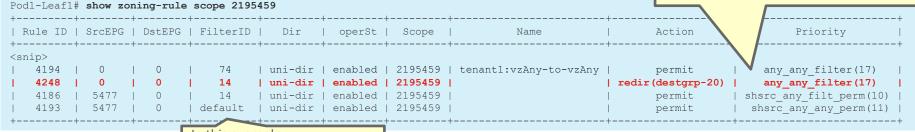
What's the forwarding action?



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

  Permit

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

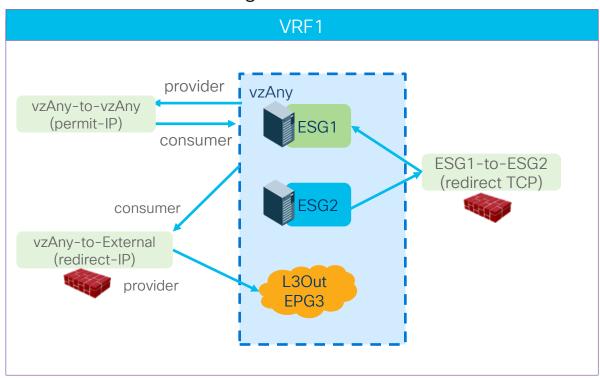




- Filter ID 74: Permit-IP all
- Filter ID 14: Permit-TCP all

### 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?









24

10936

32782

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).

ule ID			FilterID		operSt		Name	Action	Priority
ip>	+	+	+	-+	-+	+	†	+	T
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	I	permit	src dst any(9)
4201	32782	0	74	uni-dir	enabled	2195459	I	redir(destgrp-1)	src any filter(13)
4242	5477	0	74	uni-dir	enabled	2195459		permit	shsrc_any_filt_perm(1
4186	24	10936	14	bi-dir	enabled	2195459	l	redir(destgrp-1)	fully qual(7)
4193	5477	10936	default	uni-dir	enabled	2195459	l .	permit	src dst any(9)
4209	5477	24	14	uni-dir	enabled	2195459	l .	permit	fully_qual(7)
4248	10936	24	14	uni-dir-ignore	enabled	2195459	l .	redir(destgrp-1)	fully qual(7)



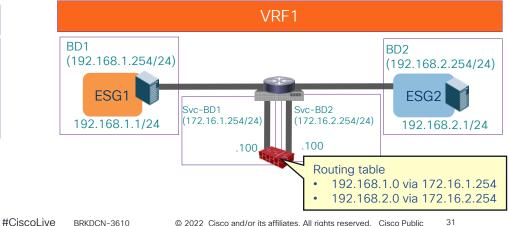
# FAQ and advanced use cases



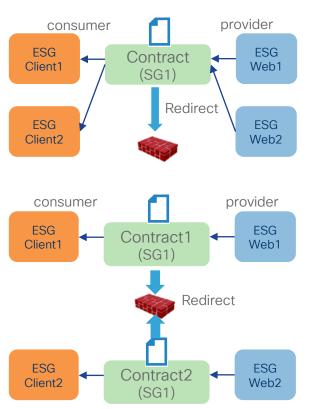
#### One-arm vs Two-arm?

- 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.
- VRF1 BD2 BD1 (192.168.2.254/24) (192.168.1.254/24) ESG2 ESG1 192.168.2.1 192.168.1.1 Svc-BD1 (172.16.1.254/24) .100 Routing table 192.168.0.0/16 via 172.16.1.254

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



#### Can we reuse same PBR destination multiple times?



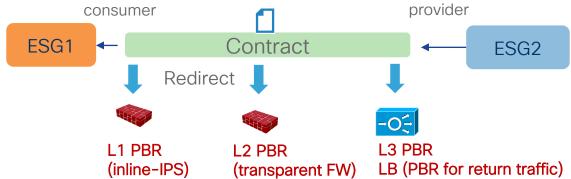
- Multiple consumer/provider ESGs/EPGs
- Multiple contracts can use the same PBR destination and Service Graph.

#### Note

- It could consume more TCAM resources if many EPGs consume and provide the same contract. The use of vzAny might be more efficient.
- Depending on routing design, one-arm mode deployment may be required.

#### What types of devices can be PBR destinations? L1/L2/L3 device

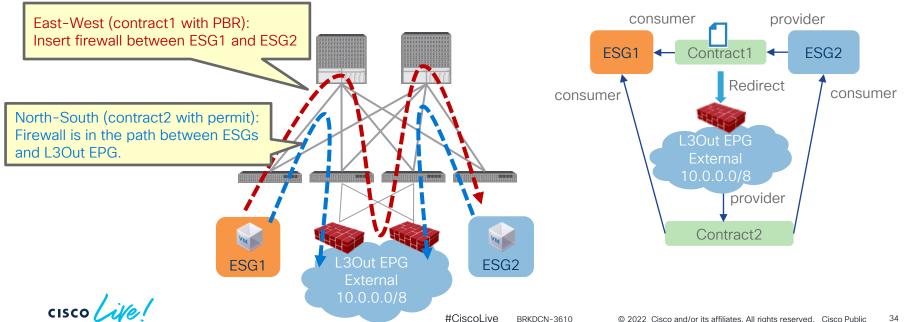
- Prior to ACI Release 5.0, Symmetric PBR destination must be L3 routed device (L3 PBR).
- Starting from ACI Release 5.0, L1/L2 Symmetric PBR is supported to insert L1/L2 devices.
  - Insert firewall without relying on BD/VLAN stitching.
  - L1/L2 service device BD must be dedicated BD that cannot be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





#### Can we use North-South firewall for East-West inspection? PBR destination in an L3Out

- Prior to ACI Release 5.2, PBR destination must be in a BD.
- Starting from ACI Release 5.2, PBR destination can be in an L3Out.



#### Advanced use cases

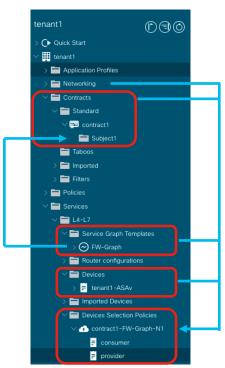
- Inter-VRF inter-tenant contract with PBR
  - The provider is in the common tenant. The consumer is in a user tenant.
  - The provider is in a user tenant. The consumer is in the common tenant.
  - The provider is in a user tenant. The consumer is in another user tenant.

- High Availability designs
  - Active/Standby
  - Active/Active
  - Independent Active nodes with Symmetric PBR



#### Inter-VRF, Inter-tenant contract with PBR

#### Configuration for PBR



- Contract in the provider or common tenant
- Service Graph template
  - · Service Graph template is attached to a subject
- · L4-L7 Device
- Device Selection Policy
  - It's based on
    - Contract name
    - Service Graph template name
    - Node name in the Service Graph
  - Then, select BD/L3Out etc, for the consumer and provider connector of the service node.

Note: vzAny cannot be a provider for an inter-VRF contract.

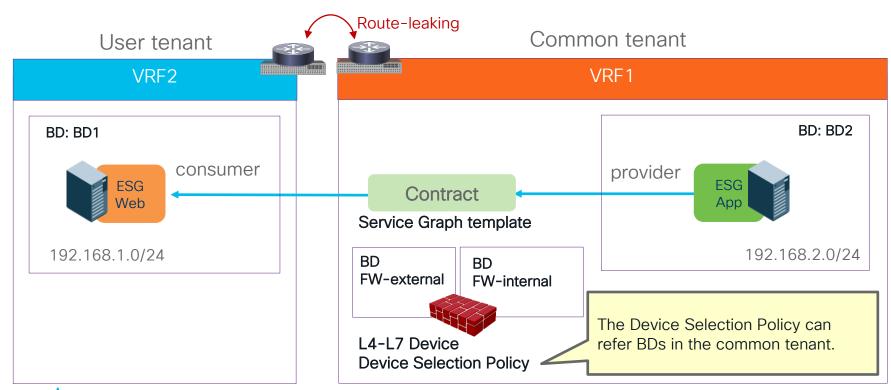
#### Important consideration

- Device Selection policy must be in the provider tenant.
- Device Selection policy must be able to refer:
  - L4-L7 Device
  - The BD/L3Out for the service device



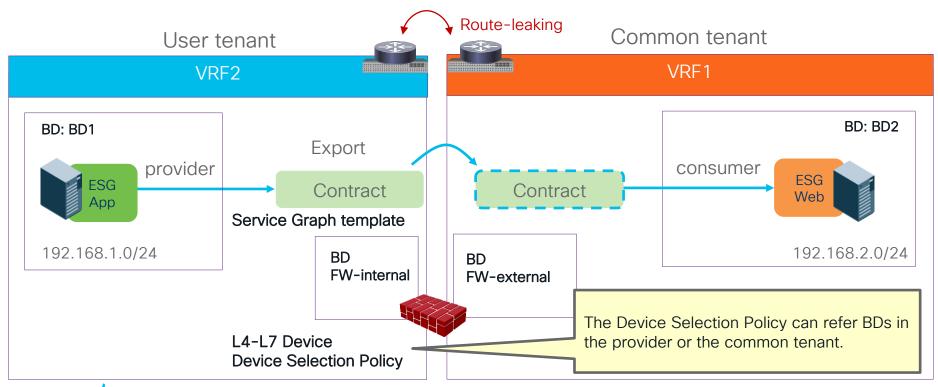
## Inter-VRF, Inter-tenant contract with PBR

Example 1: The provider is in the common tenant. (BDs for PBR destinations are in the provider tenant)



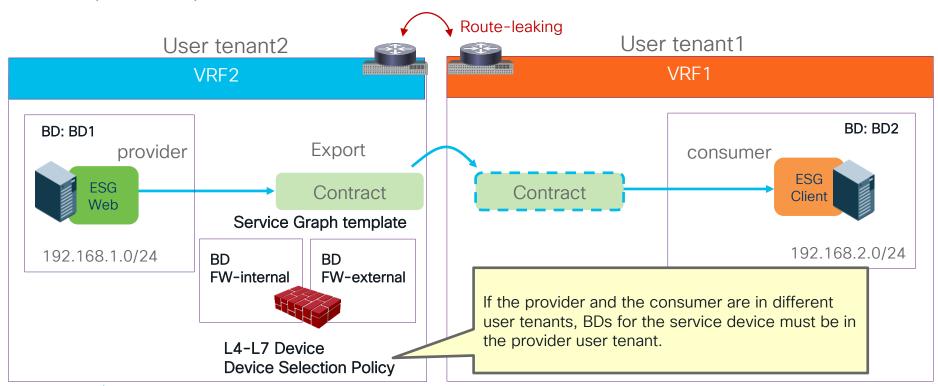
## Inter-VRF, Inter-tenant contract with PBR

Example 2: The provider is in a user tenant and the consumer is in the common tenant.



## Inter-VRF, Inter-tenant contract with PBR

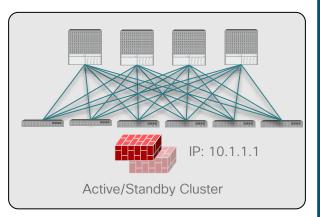
Example 3: The provider is in a user tenant and the consumer is in another user tenant.



## HA design options

One PBR destination IP
One Logical device with two concrete devices

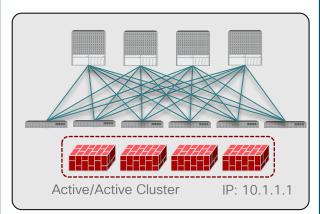
Active/Standby Cluster



- 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

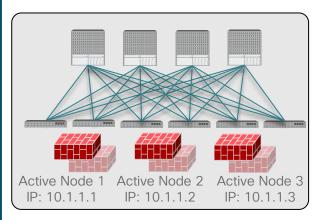
Active/Active Cluster ('Scale-Up' Model)



- 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)



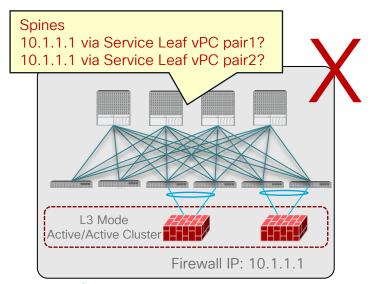
- PBR is required.
- Each Active node represent a unique MAC/IP entry.
- Use of Symmetric PBR to ensure each flow is handled by the same Active node in both directions

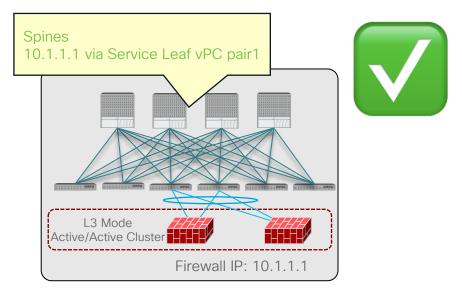


## Active/Active cluster

#### 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 is the flapping.



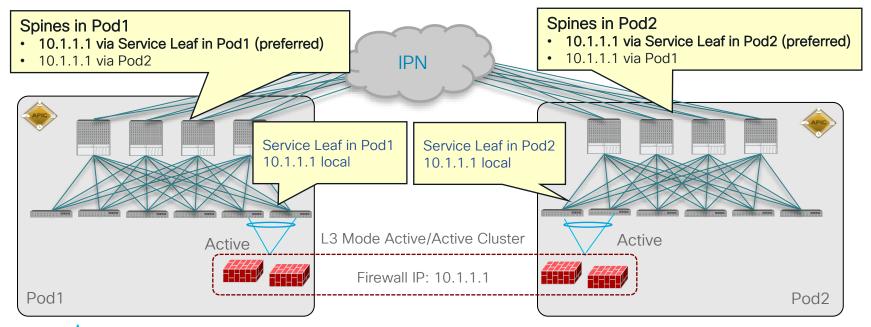




## Active/Active cluster across pods

#### Anycast service

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

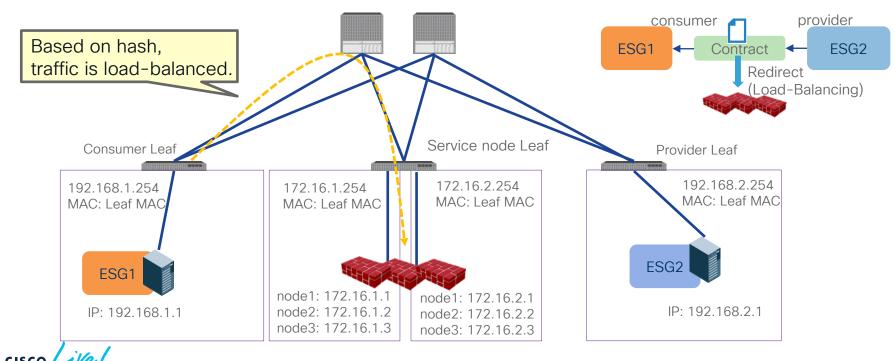


## PBR destinations can be distributed across multiple leaf nodes.

## Independent Active Nodes

Symmetric PBR: Scale Firewall Easily

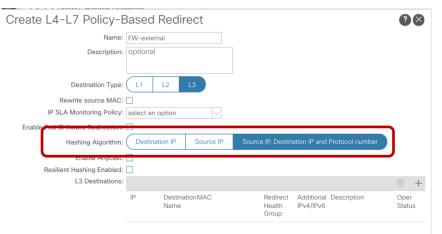
Ensure incoming and return traffic go to the same firewall



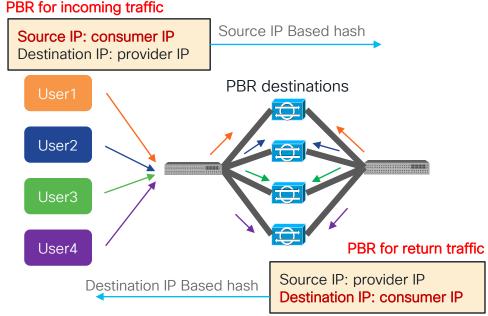
## Independent Active Nodes

#### Symmetric PBR: Hash algorithm option

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



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





## What happens if an L4-L7 device is down?

#### Without Resilient Hash (Default behavior)

• 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.

PBR destinations

Incoming Traffic

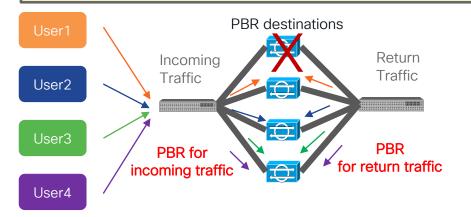
User2

User3

PBR for incoming traffic

PBR for return traffic

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

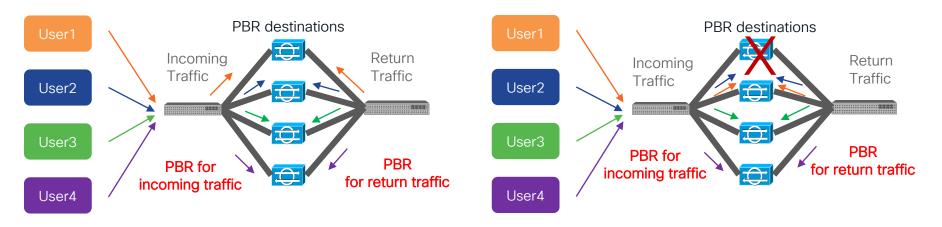




## I want to minimize impact on the existing flow!

#### With Resilient Hash

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

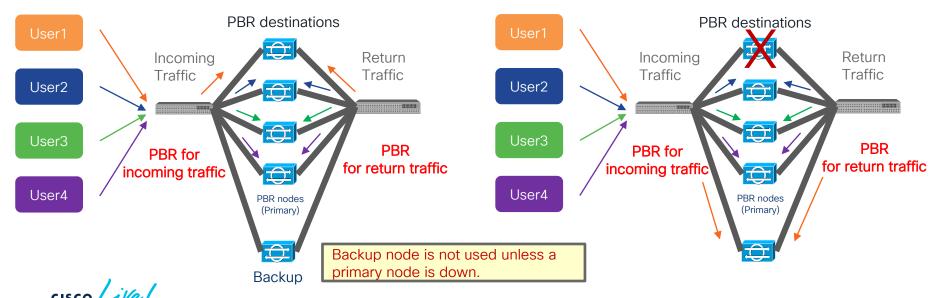




## Can we use standby PBR destination?

#### 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)



## Conclusions





## 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.
- For configuration steps, please check ACI PBR white paper!



## **Technical Session Surveys**

- Attendees who fill out a minimum of four session surveys and the overall event survey will get Cisco Live branded socks!
- Attendees will also earn 100 points in the Cisco Live Game for every survey completed.
- These points help you get on the leaderboard and increase your chances of winning daily and grand prizes.



## Cisco Learning and Certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. www.cisco.com/go/certs



(CLCs) are prepaid training vouchers redeemed directly with Cisco.



#### Learn



#### Train



#### Certify



#### Cisco U.

IT learning hub that guides teams and learners toward their goals

#### Cisco Digital Learning

Subscription-based product, technology, and certification training

#### Cisco Modeling Labs

Network simulation platform for design, testing, and troubleshooting

#### **Cisco Learning Network**

Resource community portal for certifications and learning



#### **Cisco Training Bootcamps**

Intensive team & individual automation and technology training programs

#### **Cisco Learning Partner Program**

Authorized training partners supporting Cisco technology and career certifications

## Cisco Instructor-led and Virtual Instructor-led training

Accelerated curriculum of product, technology, and certification courses



#### Cisco Certifications and Specialist Certifications

Award-winning certification program empowers students and IT Professionals to advance their technical careers

#### Cisco Guided Study Groups

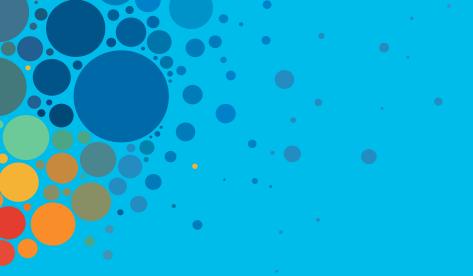
180-day certification prep program with learning and support

#### Cisco Continuing Education Program

Recertification training options for Cisco certified individuals

Here at the event? Visit us at The Learning and Certifications lounge at the World of Solutions





# 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 <u>www.CiscoLive.com/on-demand</u>



## Thank you



# cisco Live!





### 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/whitepaper-c11-739971.html

- Cisco ACI Contract Guide
  - https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centricinfrastructure/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-centricinfrastructure/white-paper-c11-2491213.html
- ACI Fabric Endpoint Learning White Paper

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centricinfrastructure/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

