Dépannage de PBR dans l'ACI

Table des matières

Introduction

Ce document décrit comment dépanner les environnements ACI (infrastructure axée sur les applications) avec la redirection basée sur les politiques (PBR) dans une structure de pod unique.

Conditions préalables

Exigences

Pour cet article, nous vous recommandons de disposer de connaissances générales sur les sujets suivants :

- Concepts ACI : Politiques d'accès, apprentissage des terminaux, contrats et L3out

Composants utilisés

Cet exercice de dépannage a été effectué sur la version 6.0(8f) de l'ACI à l'aide des commutateurs Nexus de deuxième génération N9K-C93180YC-EX et N9K-C93240YC-FX2.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l’incidence possible des commandes.

Abréviations

- BD : Domaine Bridge
- EPG : Groupe de terminaux
- ID de classe : Balise qui identifie un EPG
- Noeud PBR : Périphérique L4-L7 utilisé pour une destination PBR
- Connecteur client : interface de noeud PBR côté client
- Connecteur du fournisseur : Interface de noeud PBR côté fournisseur

Historique PBR

Version	Principales fonctionnalités
2,0(1 m)	• Les graphiques de service fournissent la fonctionnalité PBR.
3.x et versions antérieures	• Prise en charge de la redirection PBR (Policy-Based Redirect) multinoeud • Hachage résilient PBR
3,2(x)	• Le pare-feu PBR à un noeud est pris en charge dans les environnements multisites.
4,0(x)	• Le pare-feu PBR à deux noeuds est pris en charge dans les environnements multisites.
4.2(1)	• Une politique de sauvegarde pour la création de noeuds PBR de secours est désormais prise en charge.
4.2(3)	• L'option Filter-from-contract est disponible dans le modèle Service Graph à l'aide de l'interface utilisateur graphique.
5.0(1)	• Les sorties L3 sont prises en charge sur tous les côtés fournisseur des noeuds de service. • Les chemins de déploiement actif-actif/ECMP sont pris en charge pour les périphériques PBR de couche 1/couche 2.
5.2(1)	• Une destination PBR peut maintenant être dans une L3Out. • Vous pouvez suivre les noeuds de service à l'aide de l'URI HTTP. • Les modes Service Graph géré et hybride ne sont plus pris en charge. • L'adresse MAC de noeud PBR dynamique est prise en charge.
6.0(1)	• Redirection basée sur la politique symétrique basée sur le poids (PBR)

Considérations conceptuelles

- PBR fonctionne avec les appareils physiques et virtuels
- PBR peut être utilisé entre des EPG L3Out et des EPG, entre des EPG et entre des EPG L3Out. PBR n'est pas pris en charge si L2Out EPG fait partie du contrat
- Le PBR est pris en charge dans les environnements Cisco ACI multipod, multisite et leaf distants.
- Le fabric Cisco ACI doit être la passerelle des serveurs et du noeud PBR
- Le périphérique L4-L7 doit être déployé en mode d'accès (mode routé)
- Noeud PBR non pris en charge sur les commutateurs FEX
- Un domaine Bridge dédié est nécessaire pour le noeud PBR
- L'adresse MAC dynamique du noeud PBR est désormais prise en charge à l'aide des groupes d'intégrité.
- Il est recommandé d'activer la détection GARP sur le domaine de pont de noeud PBR
- Le filtre par défaut courant qui inclut le protocole ARP, le trafic Ethernet et tout autre trafic non IP ne doit pas être utilisé pour PBR
- Le noeud PBR peut se trouver entre des instances VRF ou dans l'une des instances VRF. Pour cette topologie, le noeud PBR n'est pas pris en charge pour être sur un troisième VRF, doit être configuré dans le VRF du consommateur ou du fournisseur

Informations générales

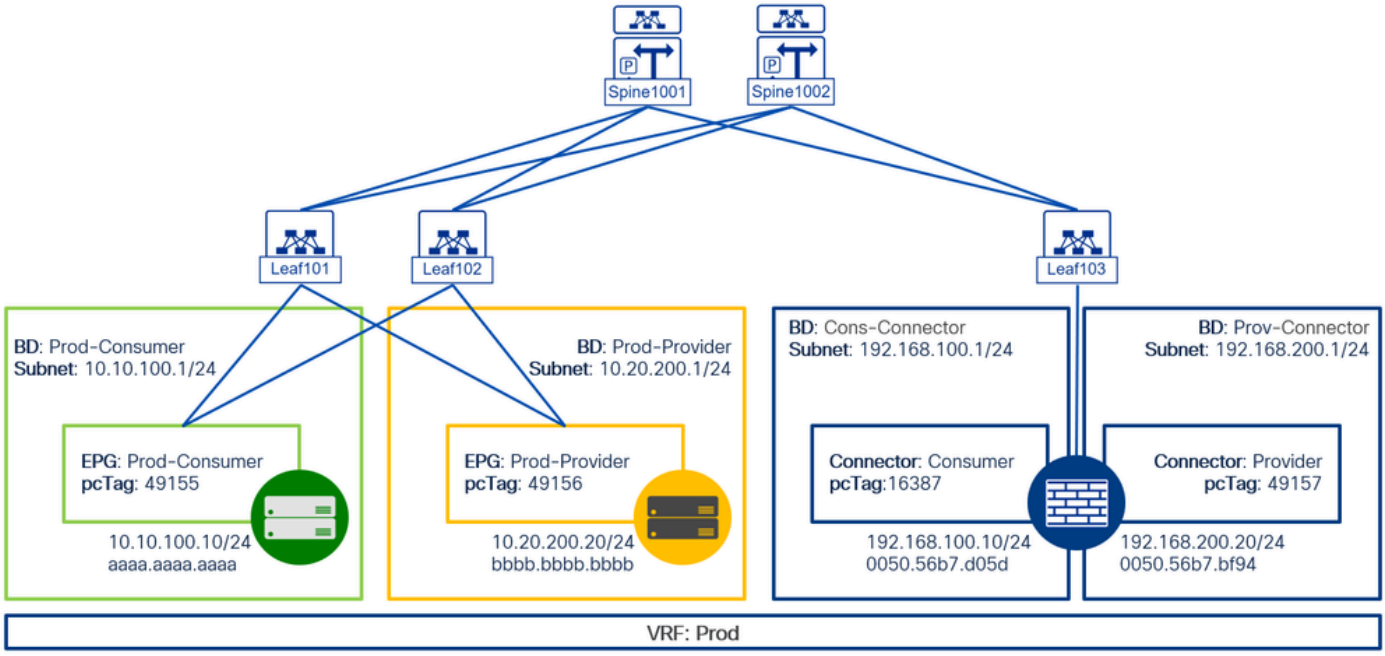
Vous trouverez des explications plus détaillées sur ELAM et Ftriage dans la bibliothèque CiscoLive On-Demand de la session BRKDCN-3900b.

En outre, toutes les directives de configuration sont disponibles dans le livre blanc Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper.

Scénario: VRF unique dans un fabric à pod unique

Diagramme du réseau

Topologie physique :



Dépannage

Étape 1 : Défauts

L'ACI génère une erreur en cas de problème au niveau de la configuration ou des interactions de stratégie. Des défaillances spécifiques ont été identifiées pour le processus de rendu PBR en cas de défaillance :

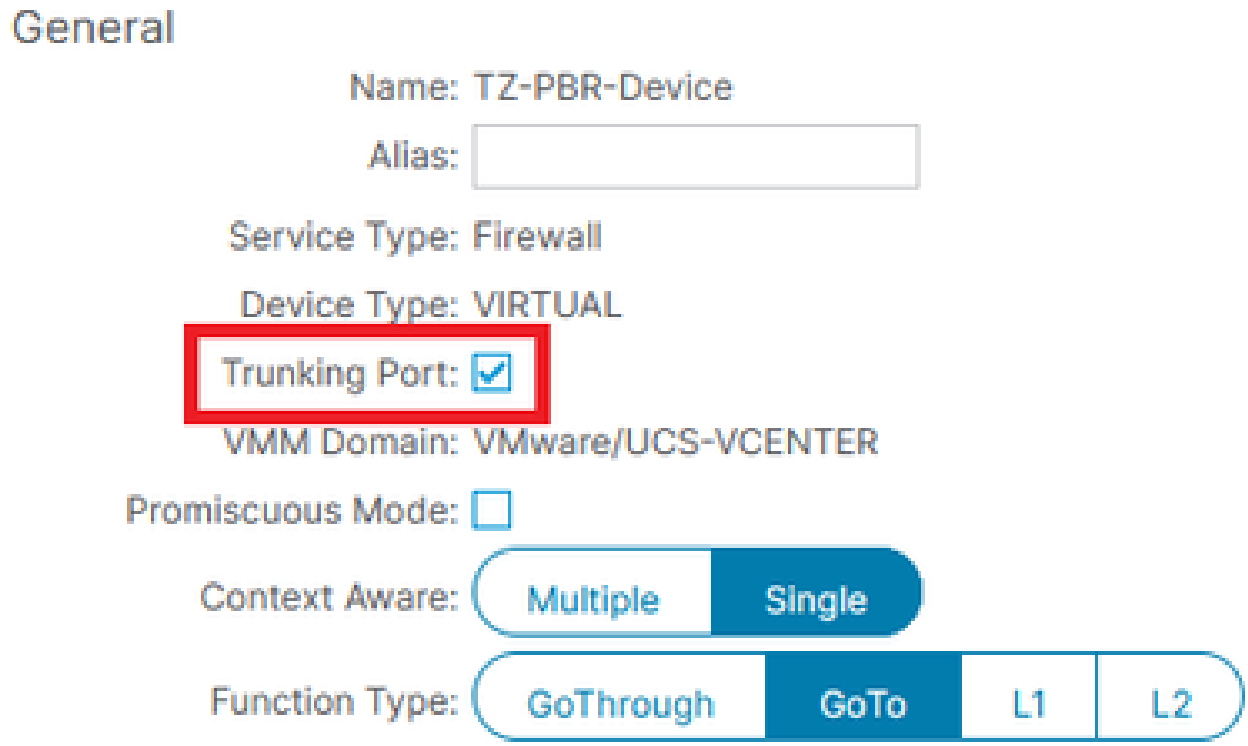
F1690 : La configuration n'est pas valide pour :

- échec d'attribution ID

  Cette erreur signifie que le VLAN encapsulé pour le noeud de service n'est pas disponible. Par exemple, aucun VLAN dynamique n'est disponible dans le pool de VLAN associé au domaine Virtual Machine Manager (VMM) utilisé par le périphérique logique.

  Résolution : Vérifiez le pool de VLAN dans le domaine utilisé par le périphérique logique. Si l'interface du périphérique logique se trouve dans un domaine physique, inspectez également la configuration du VLAN encapsulé. Ces paramètres sont disponibles sous Tenant > Services > L4-L7 > Devices and Fabric > Access Policies > Pools > VLAN.

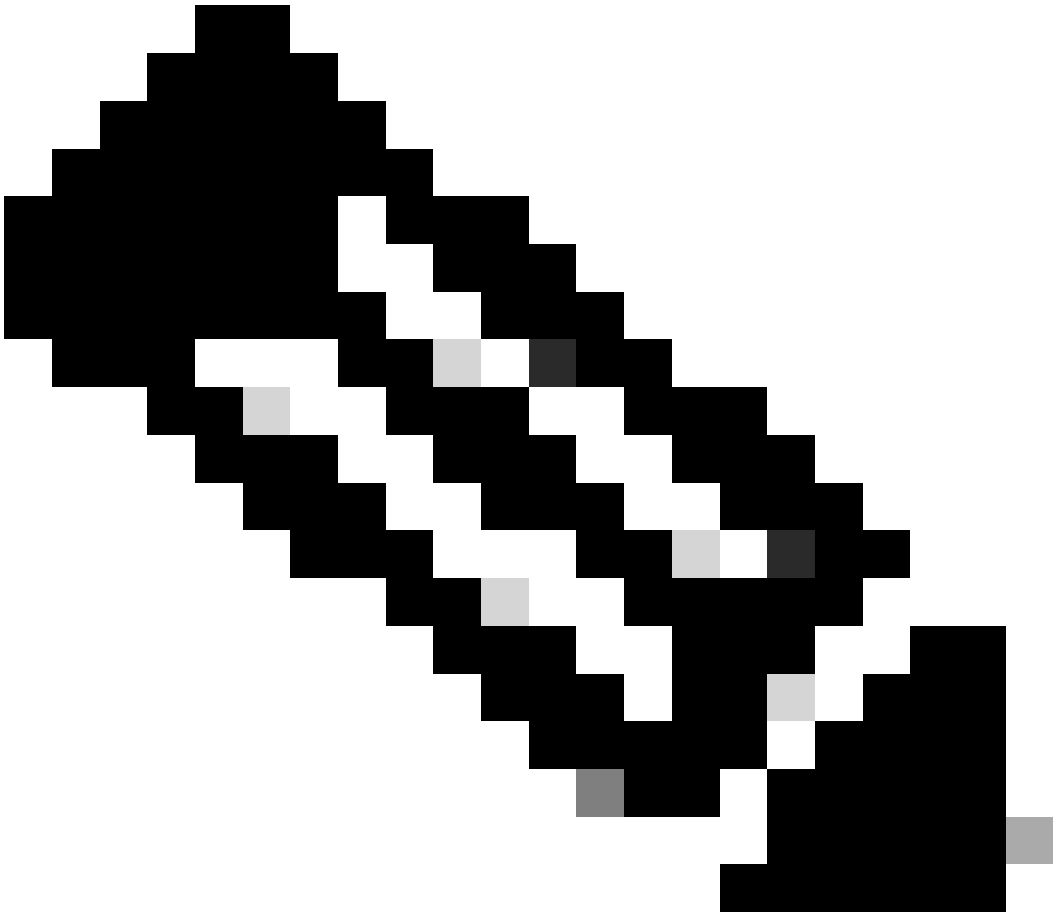
  Inversement, si l'interface de périphérique logique réside dans un domaine virtuel et se connecte à vos hôtes ESXi via une interface d'agrégation, assurez-vous que l'option de port d'agrégation est activée.



- Aucun contexte de périphérique trouvé pour LDev

Cette erreur indique que le périphérique logique n'est pas localisable pour le rendu Service Graph. Par exemple, aucune stratégie de sélection de périphérique ne peut correspondre au contrat associé au graphique de services.

Résolution : Vérifiez qu'une stratégie de sélection de périphérique est définie. La politique de sélection de périphérique spécifie les critères de sélection pour un périphérique de service et ses connecteurs, en fonction du nom du contrat, du nom du graphique de service et du nom du noeud dans le graphique de service. Vous pouvez le trouver sous Tenant > Services > L4-L7 > Device Selection Policy.

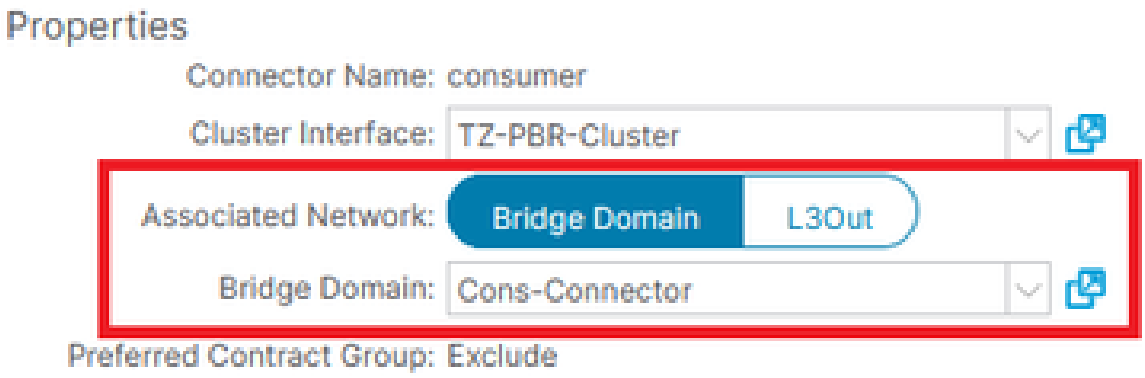


Remarque : Lorsque vous déployez un modèle Service Graph, l'ACI présélectionne le domaine de pont pour l'EPG source. Vous devez modifier ce domaine de pont pour le connecteur client PBR. Il en va de même pour le connecteur fournisseur.

- Aucun BD trouvé

  Cette erreur indique que le domaine de pont (BD) du noeud de service est introuvable. Par exemple, le BD n'est pas spécifié dans la politique de sélection de périphérique.

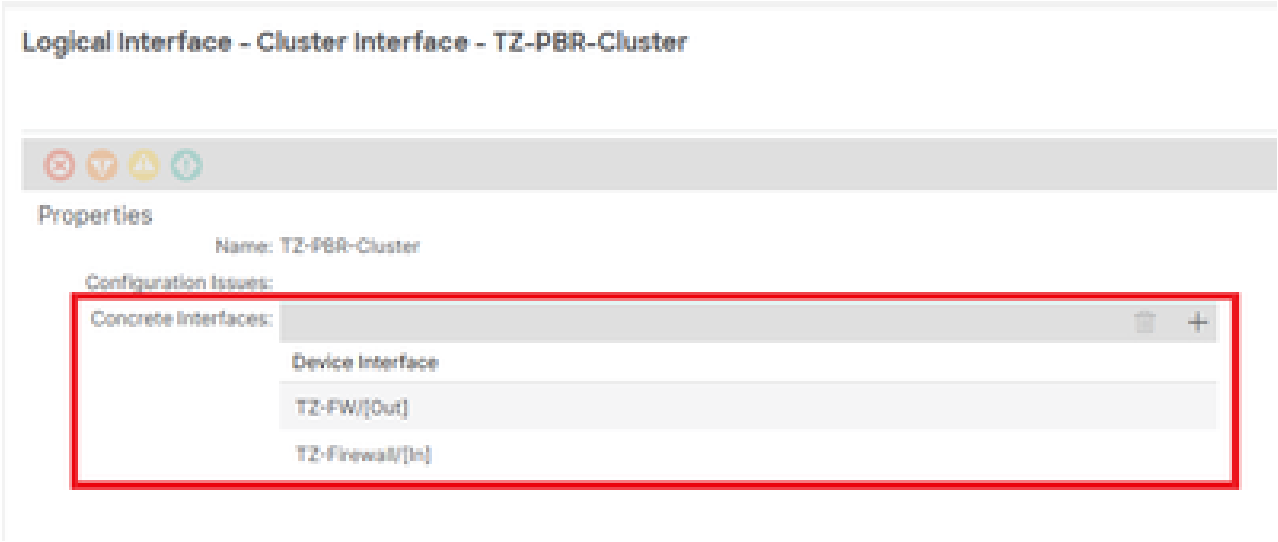
  Résolution : Assurez-vous que le BD est spécifié dans la stratégie de sélection de périphérique et que le nom du connecteur est correct. Cette configuration se trouve sous Locant > Services > L4-L7 > Devices Selection Policy > [Contract + SG] > [Consumer | Fournisseur].



- LIf n'a aucune relation avec CIf et LIf a un CIf non valide
- Aucune interface de cluster trouvée

  Ces défauts indiquent que le périphérique n'a aucune relation avec les interfaces de cluster.

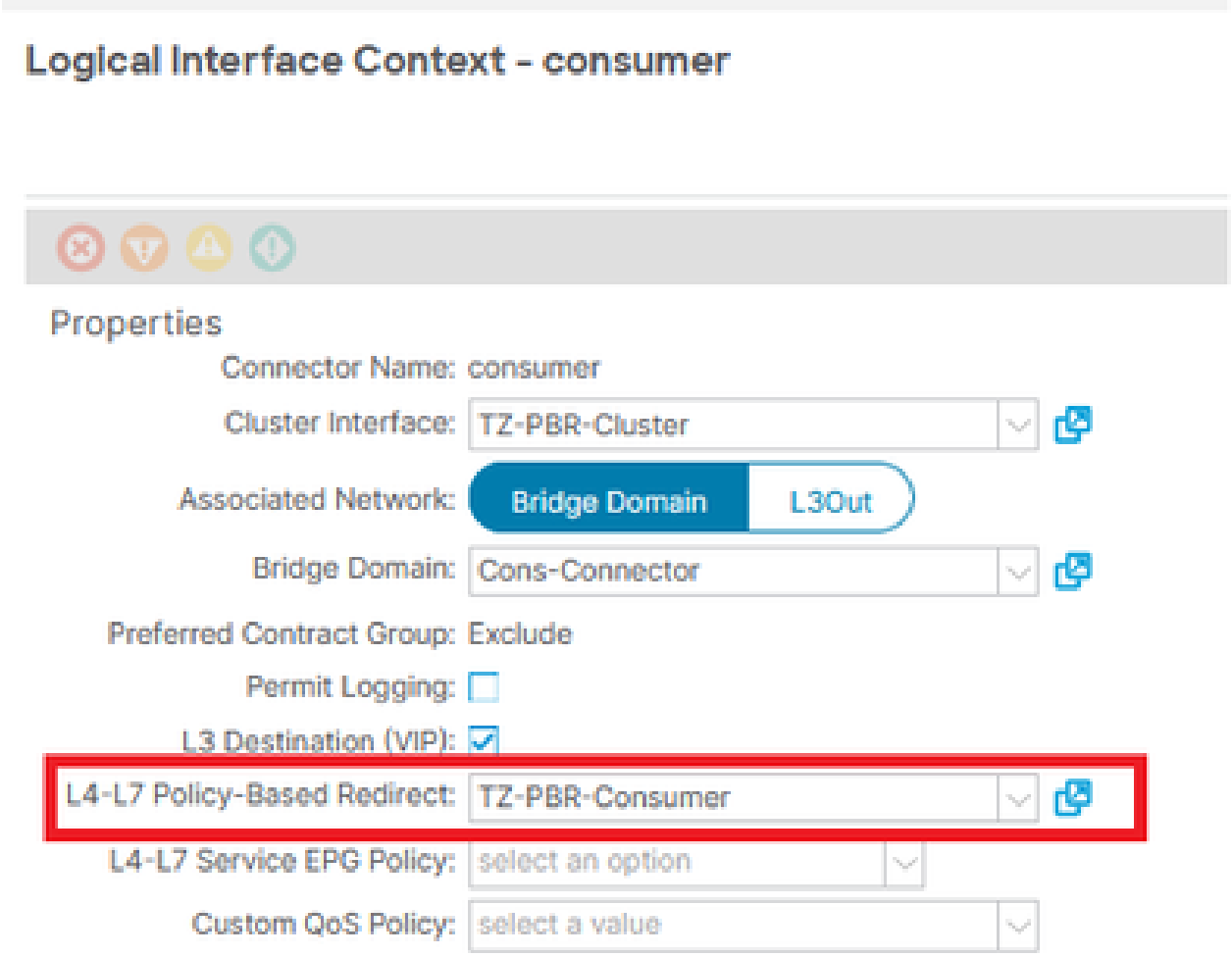
  Résolution : Assurez-vous que la configuration du périphérique de couche 4 à couche 7 (L4-L7) inclut un sélecteur d'interface concret spécifié. Cette configuration se trouve sous Tenant > Services > L4-L7 > Devices > [Device] > Cluster Interfaces.



- Stratégie de redirection de service non valide

  Cette erreur signifie que la stratégie PBR n'a pas été appliquée malgré l'activation de la redirection sur la fonction de service dans le graphique des services.

  Résolution : Assurez-vous que la stratégie PBR est configurée dans les paramètres de la stratégie de sélection de périphérique. Cette configuration se trouve sous Locant > Services > L4-L7 > Devices Selection Policy > [Contract + SG] > [Consumer | Fournisseur].



F0759 : graph-render-failure - "Le graphique de service pour le locataire < > n'a pas pu être instancié. La configuration du noeud de fonction < noeud > n'est pas valide."

Impossible d'instancier le graphique de service pour le service partagé spécifié en raison d'une configuration non valide du nom de noeud de fonction.

Cette erreur suggère qu'il y a des problèmes de configuration liés aux conditions susmentionnées.

En outre, lors des déploiements initiaux, cette défaillance peut se produire temporairement, puis être rapidement résolue. Cela se produit en raison du processus de rendu que l'ACI doit suivre pour déployer toutes les politiques.

Résolution : Étudiez toute erreur supplémentaire signalée et résolvez-la en conséquence.

F0764 : configuration-failed - "La configuration des périphériques C4-C7 < périphérique > pour le client < locataire > n'est pas valide."

Impossible d'instancier le graphique de service pour le locataire spécifié en raison d'une configuration non valide de la stratégie de périphérique PBR.

Cette erreur suggère qu'il y a des problèmes de configuration liés aux conditions susmentionnées.

Résolution : Étudiez toute erreur supplémentaire signalée et résolvez-la en conséquence.

F0772 : configuration-failed - "La configuration LIf < cluster > pour les unités L4-L7 < device > pour le locataire < tenant > n'est pas valide."

Impossible d'instancier le graphique de service pour le locataire spécifié en raison d'une configuration non valide de la sélection de l'interface de cluster de périphériques PBR.

Cette erreur suggère qu'il y a des problèmes de configuration liés aux conditions susmentionnées.

Résolution : Étudiez toute erreur supplémentaire signalée et résolvez-la en conséquence.

Étape 2 : Apprentissage des terminaux source et de destination

Assurez-vous que vos points d'extrémité source et de destination sont reconnus dans le fabric, ce qui nécessite une configuration de base :

- Objet VRF (Virtual Routing and Forwarding).
- Objet de domaine de pont (BD) avec routage de monodiffusion activé. Bien que l'activation de l'apprentissage du plan de données IP soit considérée comme une meilleure pratique, elle n'est pas obligatoire.
- Objet Endpoint Group (EPG), applicable à la fois aux domaines virtuels et physiques.
- Politiques d'accès : Vérifiez que la chaîne de configuration de votre stratégie d'accès est terminée et qu'aucune erreur n'est signalée sur l'EPG.

Pour confirmer l'apprentissage du point de terminaison sur l'EPG et l'interface corrects, exécutez cette commande sur le noeud leaf où le point de terminaison est appris, également connu sous le nom de noeud leaf de calcul :

<#root>

show system internal epm endpoint [ip | mac] [x.x.x.x | eeee.eeee.eeee]

<#root>

Leaf101#

show system internal epm endpoint ip 10.10.100.10

MAC :

aaaa.aaaa.aaaa

::: Num IPs : 1

IP# 0 : 10.10.100.10

```
 ::: IP# 0 flags : ::: l3-sw-hit: No
Vlan id : 57 ::: Vlan vnid : 10865 :::
```

VRF name : TZ:Prod

```
BD vnid : 16056291 ::: VRF vnid : 2162692
Phy If : 0x16000008 ::: Tunnel If : 0
Interface :
```

port-channel9

```
Flags : 0x80004c05 :::
```

sclass : 49155

```
 ::: Ref count : 5
EP Create Timestamp : 02/18/2025 15:00:18.767228
EP Update Timestamp : 02/18/2025 15:04:57.908343
EP Flags : local|VPC|IP|MAC|sclass|timer|

::::


Leaf101#
```

Cette commande vous permet d'identifier le pcTag (sclass) associé à l'EPG où le point d'extrémité est classé, ainsi que de récupérer l'interface, l'étendue VRF et les informations d'adresse MAC.

Si vous ne connaissez pas l'emplacement de votre point d'extrémité source ou de destination, vous pouvez toujours vous aider avec cette commande sur le contrôleur APIC :

<#root>

show endpoint [ip | mac] [x.x.x.x | eeee.eeee.eeee]

<#root>

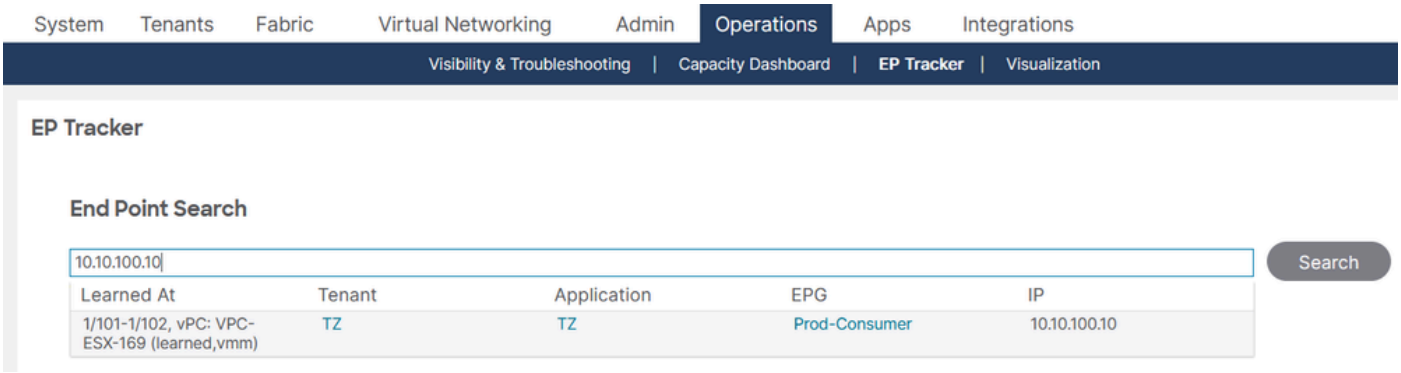
```
APIC#
```

show endpoint ip 10.10.100.10

```
Legends:
(P):Primary VLAN
(S):Secondary VLAN


Dynamic Endpoints:
Tenant : TZ
Application : TZ
AEPg : Prod-Consumer

End Point MAC     IP Address                                Source       Node       Interface
----------------- ----------------------------------------- ------------ ---------- --------------------
```

```
AA:AA:AA:AA:AA:AA 10.10.100.10

                              learned,vmm

101 102    vpc VPC-ESX-169

                vlan-2673       not-applicable    2025-02-18T15:16:40.

Total Dynamic Endpoints: 1
Total Static Endpoints: 0
APIC#
```

Dans l'interface utilisateur graphique, la fonction EP Tracker est accessible en accédant à Operations > EP Tracker pour la surveillance et la gestion des terminaux.



Grâce aux informations collectées à partir des points d'extrémité source et de destination, vous pouvez désormais vous concentrer sur le déploiement des stratégies PBR.

Étape 3 : Rediriger le contrat

PBR est intégré au cadre Service Graph. Par conséquent, un modèle Service Graph doit être déployé et configuré sur les commutateurs source et de destination conformément à un contrat. En utilisant les informations pcTags collectées à l'étape précédente, vous pouvez vérifier si un groupe de terminaux (EPG) est redirigé vers un groupe Service Graph en exécutant cette commande.

<#root>

```
show zoning-rule scope [ vrf_scope ]
```

Dans les règles de zonage, ces règles doivent être prises en compte :

1. EPG source vers EPG de destination avec un groupe de redirection associé
2. EPG fantôme de noeud PBR source vers EPG source
3. EPG de destination vers EPG source avec un groupe de redirection associé. Ce groupe peut être identique ou différent de la configuration précédente.
4. EPG fantôme de noeud PBR de destination vers EPG de destination

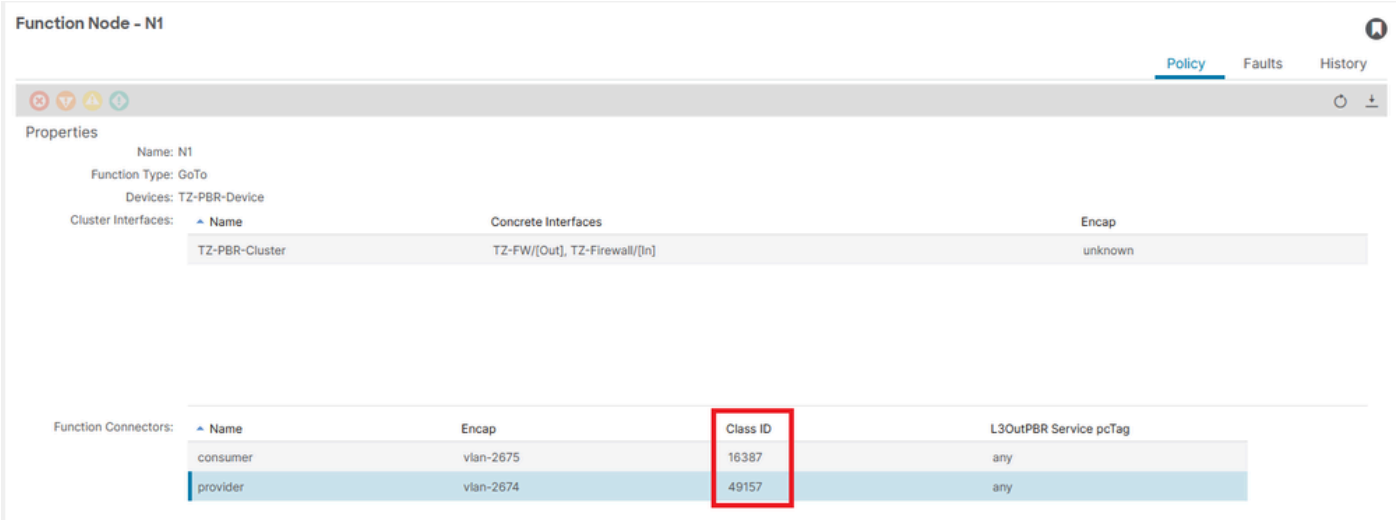
<#root>

Leaf101#

show zoning-rule scope 2162692

```
+---------+--------+--------+----------+----------------+---------+---------+------+------------------+
| Rule ID | SrcEPG | DstEPG | FilterID |       Dir      |  operSt |  Scope  | Name |      Action      |
+---------+--------+--------+----------+----------------+---------+---------+------+------------------+
|   4565  |  49155 |  49156 |  default |     bi-dir     | enabled | 2162692 |      | redir(destgrp-8) |
|   4565  |  49156 |  49155 |  default | uni-dir-ignore | enabled | 2162692 |      | redir(destgrp-9) |
|   4973  |  16387 |  49155 |  default |     uni-dir    | enabled | 2162692 |      |      permit      |
|   4564  |  49157 |  49156 |  default |     uni-dir    | enabled | 2162692 |      |      permit      |
+---------+--------+--------+----------+----------------+---------+---------+------+------------------+
Leaf101#
```

Pour vérifier le pcTag des groupes de terminaux fantômes créés lors du processus de déploiement du routage PBR (Policy-Based Routing), accédez à Tenants > [TENANT_NAME] > Services > L4-L7 > Deployed Graph Instance > [SG_NAME] > Function Node - N1.



- Analyseur de contrat

Le script met en corrélation les règles de zonage, les filtres, les statistiques et les noms de groupes de terminaux. Vous pouvez exécuter ce script en toute sécurité directement sur un leaf ACI ou un APIC. Lorsqu'il est exécuté sur le contrôleur APIC, il collecte des objets concrets sur tous les commutateurs Leaf, ce qui peut prendre plusieurs minutes pour les déploiements de politiques de grande envergure.

À partir de la version 3.2 de l'ACI, le Contract_parser est intégré à l'image et disponible sur le leaf. Il vous suffit de saisir contract_parser.py à partir du shell iBash.

<#root>

Leaf101#

contract_parser.py --sepg 49155

```
Key:
[prio:RuleId] [vrf:{str}] action protocol src-epg [src-l4] dst-epg [dst-l4] [flags][contract:{str}] [hi

[7:4999] [vrf:TZ:Prod] log,

redir

 ip tn-TZ/ap-TZ/epg-

Prod-Consumer(49155)

 tn-TZ/ap-TZ/epg-

Prod-Provideer(49156)

 [contract:uni/tn-TZ/brc-

TZ-PBR-Contract

] [

hit=81

]

destgrp-8

 vrf:TZ:Prod ip:

192.168.100.10

 mac:

00:50:56:B7:D0:5D

 bd:uni/tn-TZ/

BD-Cons-Connector

Leaf101#
```

Cette commande fournit des détails tels que l'action du contrat, les EPG source et de destination, le nom du contrat en cours d'utilisation et le nombre d'occurrences.

Étape 4 : Groupe de redirection

Après avoir identifié le groupe de redirection en fonction du contrat appliqué aux règles de zonage, l'étape suivante consiste à déterminer les adresses IP et MAC du ou des périphériques ciblés pour la redirection. Pour y remédier, exécutez la commande suivante :

<#root>

```
show service redir info group [ destgrp_ID ]
```

<#root>

```
Leaf101#
```

show service redir info group 8

```
================================================================================================================
LEGEND
TL: Threshold(Low) | TH: Threshold(High) | HP: HashProfile | HG: HealthGrp | BAC: Backup-Dest | TRA: Tra
================================================================================================================
GrpID Name      destination                              HG-name        BAC W   operSt  operStQual   TL  TH
===== ====      ===========                              ============== === === ======= ============ === ===
8     destgrp-8 dest-[
```

192.168.100.10

```
]-[vxlan-
```

2162692

```
] Not attached   N   1
```

enabled

```
 no-oper-grp  0   0   sym no  no
Leaf101#
Leaf101# show service redir info group 9
================================================================================================================
LEGEND
TL: Threshold(Low) | TH: Threshold(High) | HP: HashProfile | HG: HealthGrp | BAC: Backup-Dest | TRA: Tra
================================================================================================================
GrpID Name      destination                              HG-name        BAC W   operSt  operStQual   TL  TH
===== ====      ===========                              ============== === === ======= ============ === ===
9     destgrp-9 dest-[
```

192.168.200.20

```
]-[vxlan-
```

2162692

```
] Not attached   N   1
```

enabled

```
 no-oper-grp  0   0   sym no  no
Leaf101#
```

Cette commande nous permet de déterminer l'état opérationnel (OperSet) de notre groupe de redirection, l'adresse IP configurée dans la section PBR de couches 4 à 7 et le VNID du VRF associé aux domaines de pont de noeud PBR. Vous devez maintenant déterminer l'adresse MAC configurée :

<#root>

show service redir info destinations ip [PBR-node IP] vnid [VRF_VNID]

<#root>

```
Leaf101#
```

show service redir info destination ip 192.168.100.10 vnid 2162692

```
================================================================================================================
LEGEND
TL: Threshold(Low) | TH: Threshold(High) | HP: HashProfile | HG: HealthGrp | BAC: Backup-Dest | TRA: Tra
================================================================================================================
Name                                    bdVnid         vMac              vrf     operSt  operStQual   HG-n
====                                    ======         ====              ====    =====   =========    ====
dest-[
```

192.168.100.10

```
]-[vxlan-
```

2162692

```
] vxlan-
```

15826939

00:50:56:B7:D0:5D

```
 TZ:Prod
```

enabled

```
 no-oper-dest Not attached
Leaf101#
Leaf101# show service redir info destination ip 192.168.200.20 vnid 2162692
================================================================================================================
LEGEND
TL: Threshold(Low) | TH: Threshold(High) | HP: HashProfile | HG: HealthGrp | BAC: Backup-Dest | TRA: Tra
================================================================================================================
Name                                    bdVnid         vMac              vrf     operSt  operStQual   HG-n
====                                    ======         ====              ====    =====   =========    ====
dest-[
```

192.168.200.20

```
]-[vxlan-
```

2162692

```
] vxlan-
```

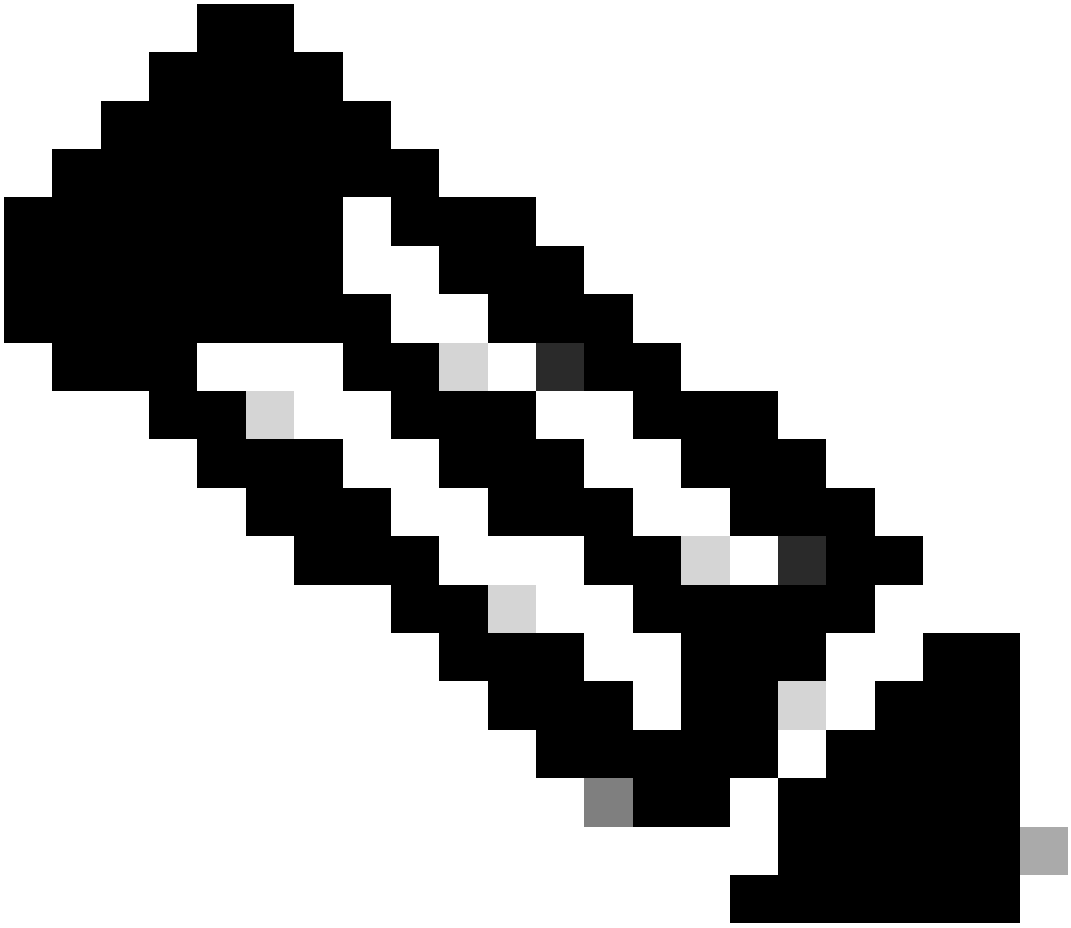
16646036 00:50:56:B7:BF:94

```
 TZ:Prod
```

enabled

```
 no-oper-dest Not attached
Leaf101#
```

En plus des détails mentionnés précédemment, cette commande fournit des informations précieuses, notamment le nom VRF, le VNID BD et l'adresse MAC configurée du noeud PBR.



Remarque : Il est important de noter que les adresses IP et MAC sont toutes deux configurées par l'utilisateur à ce stade, ce qui signifie que des erreurs typographiques peuvent se produire lors de la définition du routage basé sur des politiques de couches 4 à 7.

Étape 5 : noeud PBR ne recevant aucun trafic.

Un problème courant rencontré avec le transfert PBR est l'absence de trafic atteignant le noeud PBR. Une cause fréquente de ce problème est une adresse MAC incorrectement spécifiée dans la configuration du routage basé sur les stratégies L4-L7.

Pour vérifier l’exactitude de l’adresse MAC configurée dans le routage basé sur des politiques de couches 4 à 7, exécutez la commande précédemment utilisée à l’étape 2. Cette commande peut être exécutée sur le commutateur leaf désigné comme noeud terminal de service, où le noeud est censé être appris.

<#root>

show system internal epm endpoint [ip | mac] [x.x.x.x | eeee.eeee.eeee]

<#root>

Leaf103#

show system internal epm endpoint ip 192.168.100.10

MAC :

0050.56b7.d05d

::: Num IPs : 1
IP# 0 :

192.168.100.10

::: IP# 0 flags : ::: l3-sw-hit: Yes ::: flags2 :

dp-lrn-dis

Vlan id : 71 ::: Vlan vnid : 10867 ::: VRF name : TZ:Prod
BD vnid : 15826939 ::: VRF vnid :

2162692

Phy If : 0x16000008 ::: Tunnel If : 0
Interface : port-channel19
Flags : 0x80004c25 ::: sclass : 16387 ::: Ref count : 5
EP Create Timestamp : 02/19/2025 12:07:44.065032
EP Update Timestamp : 02/19/2025 15:27:03.400086
EP Flags : local|vPC|peer-aged|IP|MAC|sclass|timer|

::::

Leaf103#

.....

Leaf103#

show system internal epm endpoint ip 192.168.200.20

MAC :

0050.56b7.bf94

::: Num IPs : 1
IP# 0 :

192.168.200.20

::: IP# 0 flags : ::: l3-sw-hit: Yes ::: flags2 :

dp-lrn-dis

```
Vlan id : 60 ::: Vlan vnid : 10866 ::: VRF name : TZ:Prod
BD vnid : 16646036 ::: VRF vnid :
```

2162692

```
Phy If : 0x16000008 ::: Tunnel If : 0
Interface : port-channel19
Flags : 0x80004c25 ::: sclass : 49157 ::: Ref count : 5
EP Create Timestamp : 02/19/2025 13:51:03.377942
EP Update Timestamp : 02/19/2025 15:28:34.151877
EP Flags : local|vPC|peer-aged|IP|MAC|sclass|timer|

::::


Leaf103#
```

Vérifiez que l'adresse MAC enregistrée dans la table EPM correspond à celle configurée dans le groupe de redirection de service. Même les erreurs typographiques mineures doivent être corrigées pour garantir un routage correct du trafic vers la destination du noeud PBR.

Étape 6 : Flux de trafic.

- TRIAGE

Un outil CLI pour le contrôleur APIC conçu pour automatiser la configuration et l'interprétation des processus ELAM de bout en bout. L'outil permet aux utilisateurs de spécifier un flux particulier et le commutateur leaf d'où provient le flux. Il exécute séquentiellement des ELAM sur chaque noeud pour analyser le chemin de transmission du flux. Cet outil est particulièrement utile dans les topologies complexes où le chemin des paquets n’est pas facilement discernable.

<#root>

```
APIC #
```

ftriage -user admin route -sip 10.10.100.10 -dip 10.20.200.20 -ii LEAF:101,102

```
Starting ftriage
```

Log file name for the current run is: ftlog_2025-02-25-10-26-05-108.txt

```
2025-02-25 10:26:05,116 INFO /controller/bin/ftriage -user admin route -sip 10.10.100.10 -dip 10.20.200
Request password info for username: admin
Password:
2025-02-25 10:26:31,759 INFO ftriage: main:2505 Invoking ftriage with username: admin
2025-02-25 10:26:34,188 INFO ftriage: main:1546 Enable Async parellel ELAM with 2 nodes
2025-02-25 10:26:57,927 INFO ftriage: fcls:2510
```

LEAF101

```
: Valid ELAM for asic:0 slice:0 srcid:64 pktid:1913
2025-02-25 10:26:59,120 INFO ftriage: fcls:2863
```

LEAF101

: Signal ELAM found for Async lookup
2025-02-25 10:27:00,620 INFO ftriage: main:1317 L3 packet

Seen on LEAF101

Ingress:

Eth1/45 (Po9)

Egress: Eth1/52 Vnid: 2673
2025-02-25 10:27:00,632 INFO ftriage: main:1372 LEAF101: Incoming Packet captured with [

SIP:10.10.100.10, DIP:10.20.200.20

]
...
2025-02-25 10:27:08,665 INFO ftriage: main:480 Ingress

BD(s) TZ:Prod-Consumer

2025-02-25 10:27:08,666 INFO ftriage: main:491 Ingress

Ctx: TZ:Prod Vnid: 2162692

...
2025-02-25 10:27:45,337 INFO ftriage: pktrec:367 LEAF101:

traffic is redirected

...
2025-02-25 10:28:10,701 INFO ftriage: unicast:1550 LEAF101:

traffic is redirected

to vnid:15826939

mac:00:50:56:B7:D0:5D

via tenant:TZ

graph:TZ-PBR-SG

contract:

TZ-PBR-Contract

...
2025-02-25 10:28:20,339 INFO ftriage: main:975

Found peer-node SPINE1001

and IF: Eth1/1 in candidate list
...
2025-02-25 10:28:39,471 INFO ftriage: main:1366

SPINE1001

: Incoming Packet captured with Outer [SIP:10.2.200.64, DIP:10.2.64.97] .... Inner [

SIP:10.10.100.10, DIP:10.20.200.20

]
2025-02-25 10:28:39,472 INFO ftriage: main:1408

SPINE1001

: Outgoing packet's Vnid:

15826939

2025-02-25 10:28:58,469 INFO ftriage: fib:524

SPINE1001: Proxy in spine

...
2025-02-25 10:29:07,898 INFO ftriage: main:975

Found peer-node LEAF103

. and IF: Eth1/50 in candidate list
...
2025-02-25 10:29:35,331 INFO ftriage: main:1366

LEAF103

: Incoming Packet captured with Outer [SIP:10.2.200.64, DIP:10.2.200.64] .... Inner [

SIP:10.10.100.10, DIP:10.20.200.20

]
...
2025-02-25 10:29:50,277 INFO ftriage: ep:128 LEAF103: pbr traffic with dmac:

00:50:56:B7:D0:5D

2025-02-25 10:30:07,374 INFO ftriage: main:800 Computed egress encap string

vlan-2676

2025-02-25 10:30:13,326 INFO ftriage: main:535 Egress

Ctx TZ:Prod

2025-02-25 10:30:13,326 INFO ftriage: main:536 Egress BD(s):

TZ:Cons-Connector

...
2025-02-25 10:30:18,812 INFO ftriage: misc:908 LEAF103: caller unicast:581

EP if(Po19)

 same as egr if(Po19)
2025-02-25 10:30:18,812 INFO ftriage: misc:910 LEAF103: L3 packet caller unicast:668 getting

bridged

 in SUG
2025-02-25 10:30:18,813 INFO ftriage: main:1822 dbg_sub_nexthop function returned values on node LEAF10
2025-02-25 10:30:19,378 INFO ftriage: acigraph:794 : Ftriage Completed with hunch: matching service dev
APIC #

- ELAM :

Le module ELAM (Embedded Logic Analyzer Module) est un outil de diagnostic qui permet aux utilisateurs d'établir des conditions spécifiques dans le matériel pour capturer le paquet ou la trame initiale répondant à ces critères. Lorsqu'une capture est réussie, l'état ELAM est indiqué comme déclenché. Lors du déclenchement, l'ELAM est désactivé, ce qui permet de collecter un vidage de données, ce qui facilite l'analyse des nombreuses décisions de transfert exécutées par l'ASIC du commutateur pour ce paquet ou cette trame. ELAM fonctionne au niveau de l'ASIC, en s'assurant qu'il n'affecte pas le CPU ou d'autres ressources du commutateur.

Structure de la syntaxe de commande. Cette structure a été collectée dans le livre Troubleshoot ACI Intra-Fabric Forwarding Tools

```
vsh_lc                                        [This command enters the line card shell where ELAMs are r
debug platform internal <asic> elam asic 0    [refer to the ASICs table]
```

Définir les conditions à déclencher

```
trigger reset                                                 [ensures no existing triggers are running]
trigger init in-select <number> out-select <number>           [determines what information about a packet is
set outer/inner                                               [sets conditions]
start                                                         [starts the trigger]
status                                                        [checks if a packet is captured]
```

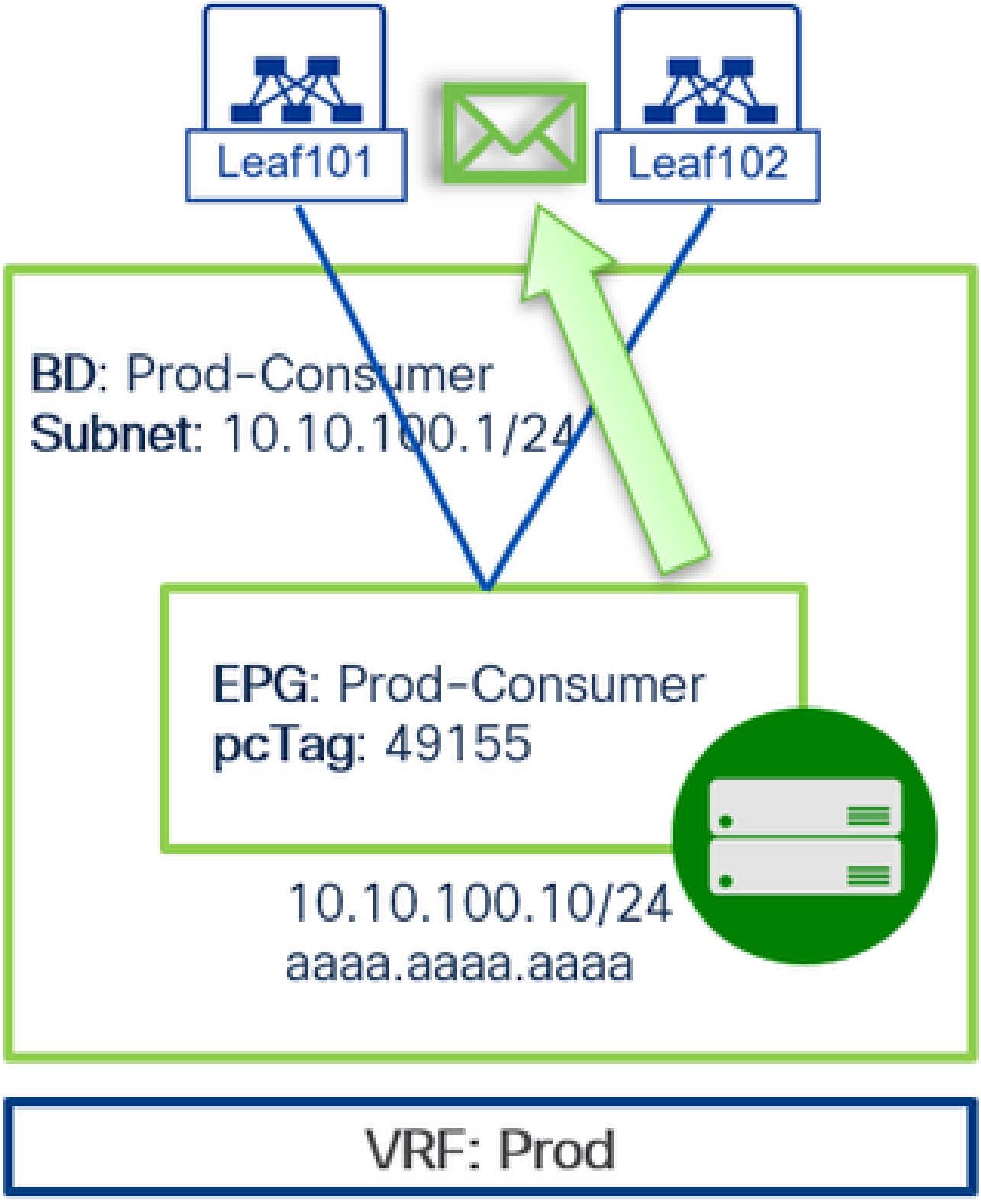
Générez le vidage contenant l'analyse des paquets.

```
ereport             [display detailed forwarding decision for the packet]
```

- Flux de trafic

Il est essentiel de comprendre le flux de trafic sur tous les périphériques en question. L'outil Ftriage fournit un excellent résumé de ce flux. Cependant, pour une validation détaillée étape par étape et pour obtenir des informations plus détaillées sur le processus de réception des paquets, vous pouvez exécuter le module ELAM (Embedded Logic Analyzer Module) à chaque point de la topologie du réseau.

1. Le trafic entrant se produit au niveau du noeud leaf de calcul où le serveur source est appris. Dans ce scénario spécifique, comme la source est positionnée derrière une interface vPC, ELAM doit être configuré sur les homologues vPC. Cela est nécessaire car l'interface physique sélectionnée par l'algorithme de hachage est indéterminée.



<#root>

LEAF101#

vsh_lc

module-1#

debug platform internal tah elam asic 0

```
module-1(DBG-elam)#
trigger reset

module-1(DBG-elam)#
trigger init in-select 6 out-select 1

module-1(DBG-elam-insel6)#
reset

module-1(DBG-elam-insel6)#
set outer ipv4 src_ip 10.10.100.10 dst_ip 10.20.200.20

module-1(DBG-elam-insel6)#
start

module-1(DBG-elam-insel6)#
status

ELAM STATUS
===========
Asic 0 Slice 0 Status
Triggered

Asic 0 Slice 1 Status Armed

module-1(DBG-elam-insel6)#
 ereport

================================================================================================================
Trigger/Basic Information
================================================================================================================
...
Incoming Interface : 0x40( 0x40 )


>>> Eth1/45


...

----------------------------------------------------------------------------------------------------------------
Outer L2 Header
----------------------------------------------------------------------------------------------------------------
Destination MAC : 0022.BDF8.19FF >>> Bridge-domain MAC address


Source MAC : AAAA.AAAA.AAAA
```

802.1Q tag is valid : yes(0x1)
CoS : 0(0x0)

Access Encap VLAN : 2673

(0xA71)

Outer L3 Header

L3 Type : IPv4
IP Version : 4
DSCP : 0
IP Packet Length : 84 (= IP header(28 bytes) + IP payload)
Don't Fragment Bit : set
TTL : 64
IP Protocol Number : ICMP
IP CheckSum : 6465(0x1941)

Destination IP : 10.20.200.20

Source IP : 10.10.100.10

Contract Lookup Key

IP Protocol : ICMP(0x1)
L4 Src Port : 2048(0x800)
L4 Dst Port : 7345(0x1CB1)

sclass (src pcTag) : 49155(0xC003) >>> Prod-Consumer

EPG

dclass (dst pcTag) : 49156(0xC004)

>>> Prod-Provider EPG

src pcTag is from local table : yes

>>> EPGs are known locally

derived from a local table on this node by the lookup of src IP or MAC
Unknown Unicast / Flood Packet : no
If yes, Contract is not applied here because it is flooded

Sideband Information

---

ovector : 176(0xB0) >>> Eth1/52

Ovec in "show plat int hal l2 port gpd"

Opcode : OPCODE_UC

---

sug_luc_latch_results_vec.luc3_0.

service_redir: 0x1 >>> Service Redir 0x1 = PBR was applied

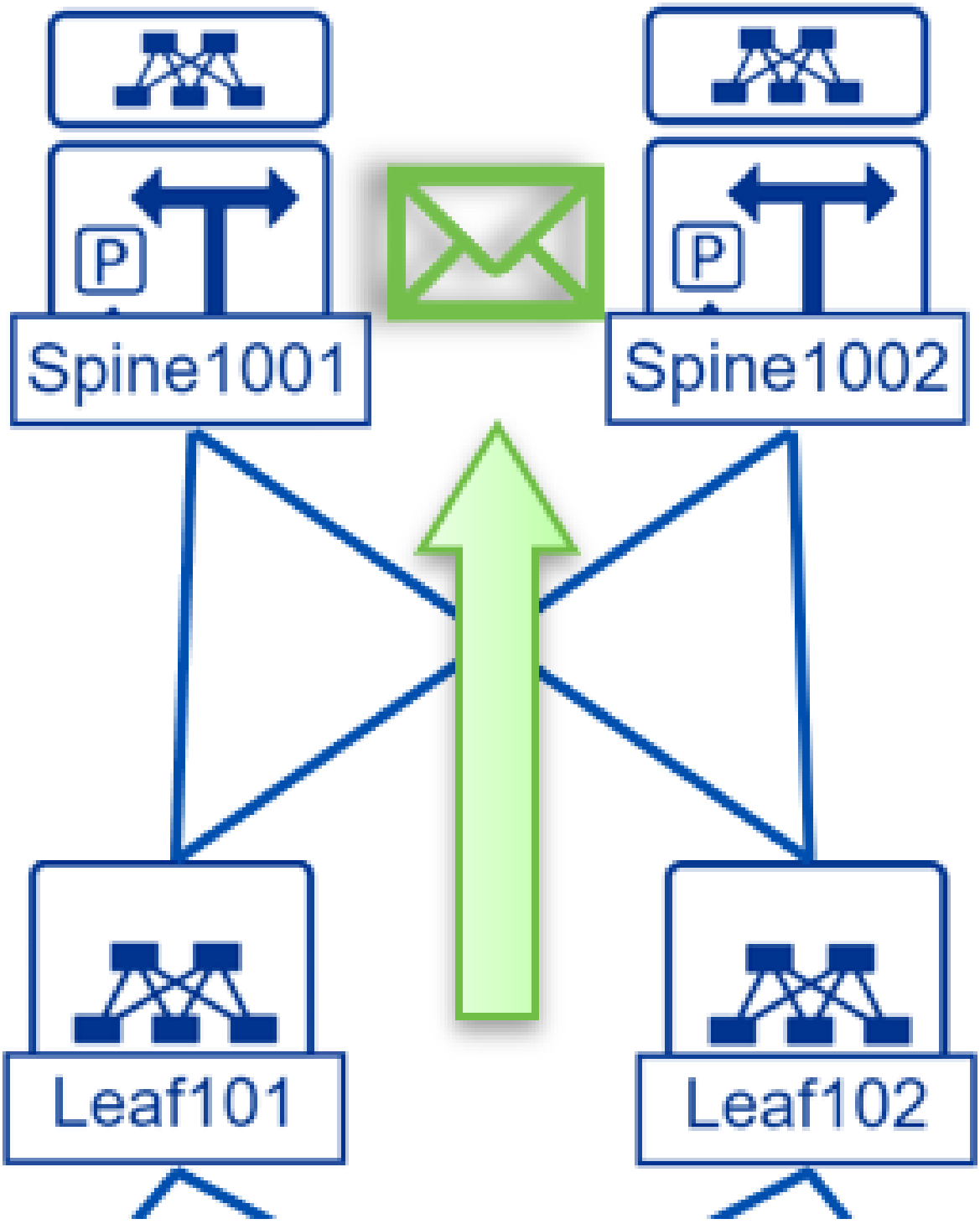
D’après les informations fournies, il est évident que le paquet est redirigé via le routage PBR (Policy-Based Routing), car l’option service_redir est activée. En outre, vous récupérez les valeurs sclass et dclass. Dans ce scénario particulier, le commutateur reconnaît la dclass. Toutefois, si le point de terminaison de destination n'est pas présent dans la table EPM, la valeur dclass est définie par défaut sur 1.

En outre, l'interface d'entrée est déterminée par le SRCID, et l'interface de sortie est identifiée par les valeurs de vecteur. Ces valeurs peuvent être traduites en port avant en exécutant cette commande au niveau vsh_lc :

<#root>

show platform internal hal l2 port gpd

2. L’étape suivante du flux consiste à atteindre le commutateur dorsal pour mapper l’adresse MAC de destination au noeud PBR. Puisque le trafic est encapsulé dans des en-têtes VXLAN, l'exécution d'ELAM sur un spine ou un leaf distant nécessite l'utilisation de la commande in-select 14 pour décoder correctement l'encapsulation.



<#root>

SPINE1001#

vsh_lc

module-1#

debug platform internal roc elam asic 0

```
module-1(DBG-elam)#
trigger reset

module-1(DBG-elam)#
trigger init in-select 14 out-selec 0

module-1(DBG-elam-insel14)#
reset

module-1(DBG-elam-insel14)#
set inner ipv4 src_ip 10.10.100.10 dst_ip 10.20.200.20

module-1(DBG-elam-insel14)#
start

module-1(DBG-elam-insel14)#
status

ELAM STATUS
===========
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Armed
Asic 0 Slice 2 Status Triggered

Asic 0 Slice 3 Status Armed

module-1(DBG-elam-insel14)#
ereport



================================================================================================================
Trigger/Basic Information
================================================================================================================
Incoming Interface :
0x48( 0x48 )  >>> Eth1/1

( Slice Source ID(Ss) in "show plat int hal l2 port gpd" )
Packet from vPC peer LEAF : yes
Packet from tunnel (remote leaf/avs) : yes

----------------------------------------------------------------------------------------------------------------
Outer L2 Header
----------------------------------------------------------------------------------------------------------------
Destination MAC : 000D.0D0D.0D0D
Source MAC : 000C.0C0C.0C0C

----------------------------------------------------------------------------------------------------------------
Inner L2 Header
```

Inner Destination MAC : 0050.56B7.D05D >>> Firewall MAC

Source MAC : AAAA.AAAA.AAAA

Outer L3 Header

L3 Type : IPv4
DSCP : 0
Don't Fragment Bit : 0x0
TTL : 32
IP Protocol Number : UDP
Destination IP : 10.2.64.97
Source IP : 10.2.200.64

Inner L3 Header

L3 Type : IPv4
DSCP : 0
Don't Fragment Bit : 0x1
TTL : 63
IP Protocol Number : ICMP

Destination IP : 10.20.200.20

Source IP : 10.10.100.10

Outer L4 Header

L4 Type : iVxLAN
Don't Learn Bit : 1
Src Policy Applied Bit : 1
Dst Policy Applied Bit : 1

sclass (src pcTag) : 0xc003 >>> pcTag 49155 (Prod-Consumer)

VRF or BD VNID : 15826939(0xF17FFB) >>> BD: Prod-Consumer

Sideband Information

Opcode : OPCODE_UC

---

bky_elam_out_sidebnd_no_spare_vec.

ovector_idx: 0x1F0 >>> Eth1/10

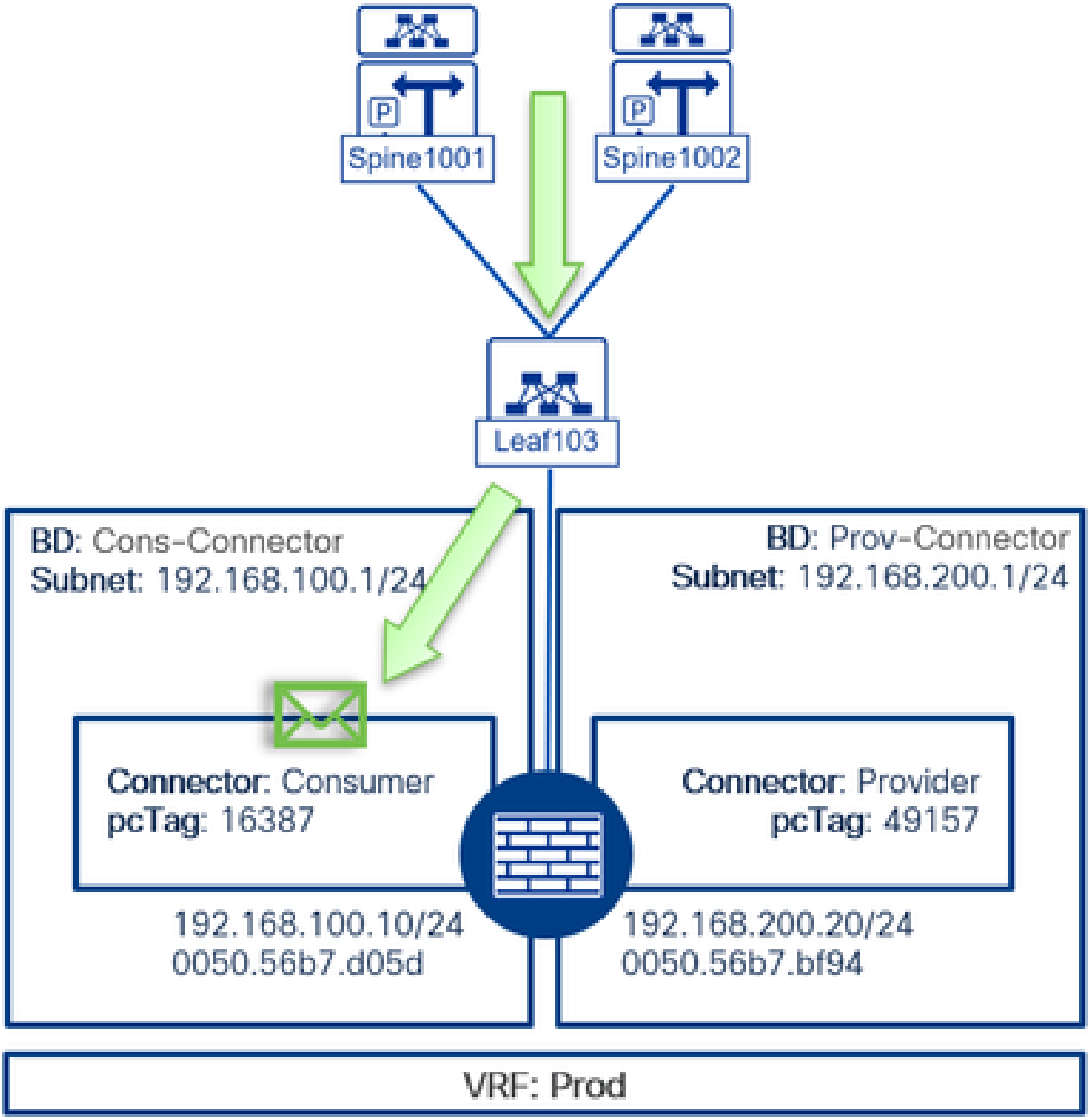
D'après le résultat précédent, il est évident que l'adresse MAC de destination est réécrite dans l'adresse MAC du pare-feu. Ensuite, une recherche COOP est effectuée pour identifier l'éditeur de destination de l'adresse MAC, et le paquet est ensuite transféré à l'interface correspondante du commutateur.

Vous pouvez simuler cette recherche sur la colonne vertébrale en exécutant cette commande, en utilisant le VNID de domaine de pont et l'adresse MAC du pare-feu :

<#root>

```
SPINE1001#

show coop internal info repo ep key 15826939 0050.56B7.D05D | egrep "Tunnel|EP" | head -n 3


EP bd vnid : 15826939
EP mac : 00:50:56:B7:D0:5D

Tunnel nh : 10.2.200.66


SPINE1001#
```

3. Le trafic atteint la feuille de service où l’adresse MAC du pare-feu est reconnue et ensuite transmise au noeud PBR.



<#root>

MXS2-LF101#

vsh_lc

module-1#

debug platform internal tah elam asic 0

module-1(DBG-elam)#

trigger reset

```
module-1(DBG-elam)#
trigger init in-select 14 out-select 1

module-1(DBG-elam-insel14)#
reset

module-1(DBG-elam-insel14)#
set inner ipv4 src_ip 10.10.100.10 dst_ip 10.20.200.20

module-1(DBG-elam-insel14)#
start

module-1(DBG-elam-insel14)#
status


ELAM STATUS
===========
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Triggered


module-1(DBG-elam-insel14)#
ereport


================================================================================================================
Trigger/Basic Information
================================================================================================================
Incoming Interface : 0x0( 0x0 ) >>> Eth1/17

( Slice Source ID(Ss) in "show plat int hal l2 port gpd" )
Packet from vPC peer LEAF : yes
Packet from tunnel (remote leaf/avs) : yes

----------------------------------------------------------------------------------------------------------------
Outer L2 Header
----------------------------------------------------------------------------------------------------------------
Destination MAC : 000D.0D0D.0D0D
Source MAC : 000C.0C0C.0C0C

----------------------------------------------------------------------------------------------------------------
Inner L2 Header
----------------------------------------------------------------------------------------------------------------
Inner
Destination MAC : 0050.56B7.D05D >>> Firewall MAC

Source MAC : AAAA.AAAA.AAAA
```

```
------------------------------------------------------------------------------------------------------------
Outer L3 Header
------------------------------------------------------------------------------------------------------------
L3 Type : IPv4
DSCP : 0
Don't Fragment Bit : 0x0
TTL : 32
IP Protocol Number : UDP
Destination IP : 10.2.200.66
Source IP : 10.2.200.64

------------------------------------------------------------------------------------------------------------
Inner L3 Header
------------------------------------------------------------------------------------------------------------
L3 Type : IPv4
DSCP : 0
Don't Fragment Bit : 0x1
TTL : 63
IP Protocol Number : ICMP
```

Destination IP : 10.20.200.20

Source IP : 10.10.100.10

```
------------------------------------------------------------------------------------------------------------
Outer L4 Header
------------------------------------------------------------------------------------------------------------
L4 Type : iVxLAN
Don't Learn Bit : 1
Src Policy Applied Bit : 1
Dst Policy Applied Bit : 1
```

sclass (src pcTag) : 0xc003 >>> pcTag 49155 (Prod-Consumer)

VRF or BD VNID : 15826939(0xF17FFB) >>> BD: Prod-Consumer

```
------------------------------------------------------------------------------------------------------------
Contract Lookup Key
------------------------------------------------------------------------------------------------------------
IP Protocol : ICMP( 0x1 )
L4 Src Port : 2048( 0x800 )
L4 Dst Port : 50664( 0xC5E8 )
```

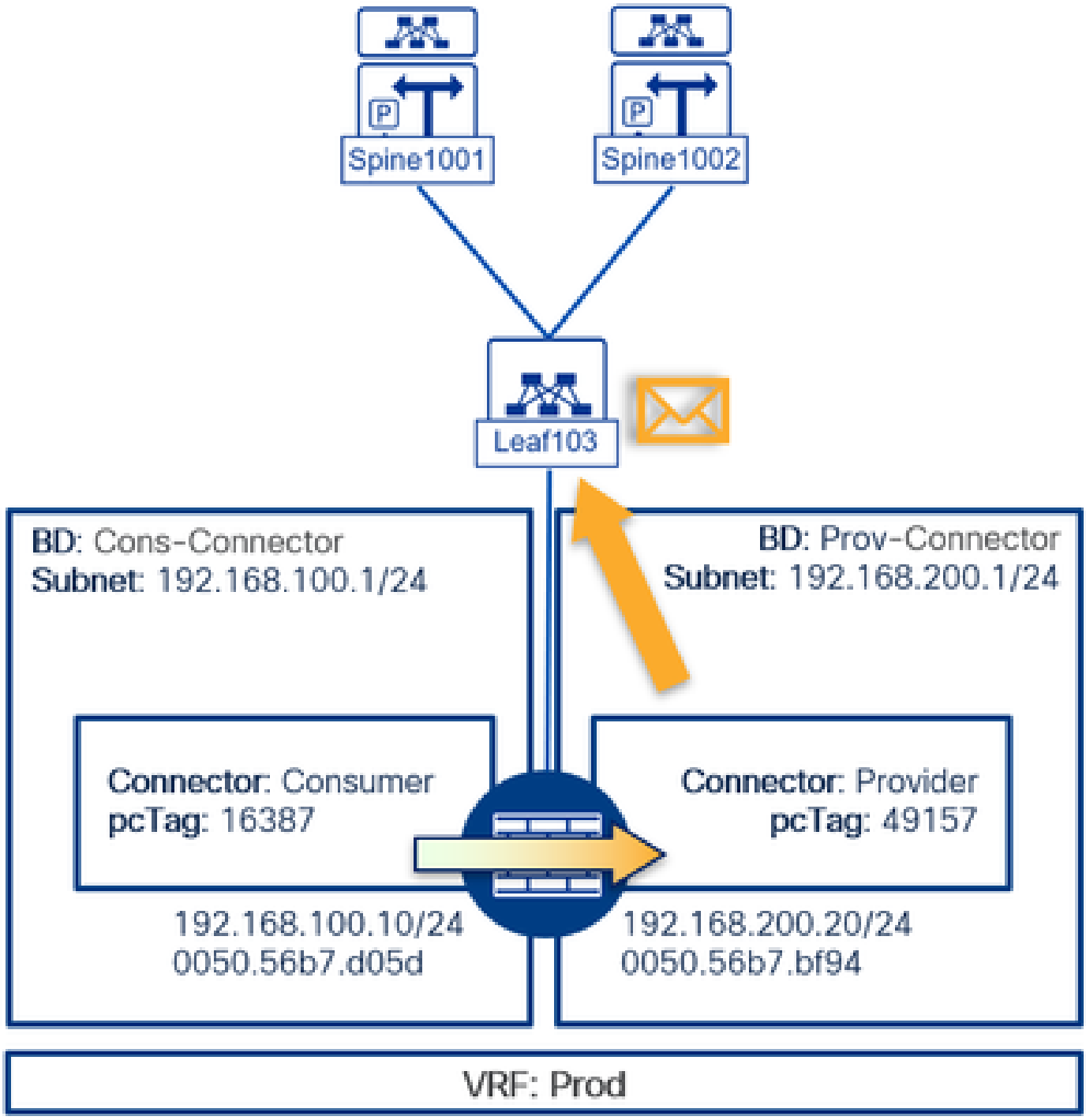
sclass (src pcTag) : 49155(0xC003) >>> Prod-Consumer EPG

dclass (dst pcTag) : 16387(0x4003) >>> Consumer connector EPG

```
src pcTag is from local table : no
derived from group-id in iVxLAN header of incoming packet
Unknown Unicast / Flood Packet : no
If yes, Contract is not applied here because it is flooded

------------------------------------------------------------------------------------------------------------
Sideband Information
```

```
----------------------------------------------------------------------------------------------------------------

ovector

 :

64( 0x40 ) >>> Eth1/45


Ovec in "show plat int hal l2 port gpd"
Opcode : OPCODE_UC
```

4. Pour le packer renvoyé par le noeud PBR, celui-ci doit d’abord effectuer sa propre vérification préalable et faire modifier le VRF, l’interface ou le VLAN. Le paquet est ensuite renvoyé à l'ACI sur le connecteur du fournisseur :



<#root>

LEAF103#

vsh_lc

module-1#

debug platform internal tah elam asic 0

module-1(DBG-elam)#

trigger reset

```
module-1(DBG-elam)#
trigger init in-select 6 out-select 1

module-1(DBG-elam-insel6)#
reset

module-1(DBG-elam-insel6)#
set outer ipv4 src_ip 10.10.100.10 dst_ip 10.20.200.20

module-1(DBG-elam-insel6)#
 set outer l2 src_mac 0050.56b7.bf94

module-1(DBG-elam-insel6)#
start

module-1(DBG-elam-insel6)#
stat

ELAM STATUS
===========
Asic 0 Slice 0 Status Triggered

Asic 0 Slice 1 Status Armed
module-1(DBG-elam-insel6)#
ereport


================================================================================================================
Trigger/Basic Information
================================================================================================================
...
Incoming Interface : 0x40( 0x40 )


>>> Eth1/45


...

----------------------------------------------------------------------------------------------------------------
Outer L2 Header
----------------------------------------------------------------------------------------------------------------
Destination MAC : 0022.BDF8.19FF
Source MAC : 0050.56B7.BF94

802.1Q tag is valid : yes( 0x1 )
```

CoS : 0(0x0)

Access Encap VLAN : 2006(0x7D6)

Outer L3 Header

L3 Type : IPv4
IP Version : 4
DSCP : 0
IP Packet Length : 84 (= IP header(28 bytes) + IP payload)
Don't Fragment Bit : set
TTL : 62
IP Protocol Number : ICMP
IP CheckSum : 46178(0xB462)

Destination IP : 10.20.200.20

Source IP : 10.10.100.10

Contract Lookup Key

IP Protocol : ICMP(0x1)
L4 Src Port : 2048(0x800)
L4 Dst Port : 37489(0x9271)

sclass (src pcTag) : 49157(0xC005) >>> Provider connector EPG

dclass (dst pcTag) : 49156(0xC004)

>>> Prod-Provider EPG

src pcTag is from local table : yes
derived from a local table on this node by the lookup of src IP or MAC
Unknown Unicast / Flood Packet : no
If yes, Contract is not applied here because it is flooded

Sideband Information

---

ovector : 176(0xB0) >>> Eth1/52

Ovec in "show plat int hal l2 port gpd"
Opcode : OPCODE_UC

---

sug_luc_latch_results_vec.luc3_0.

service_redir: 0x0

---

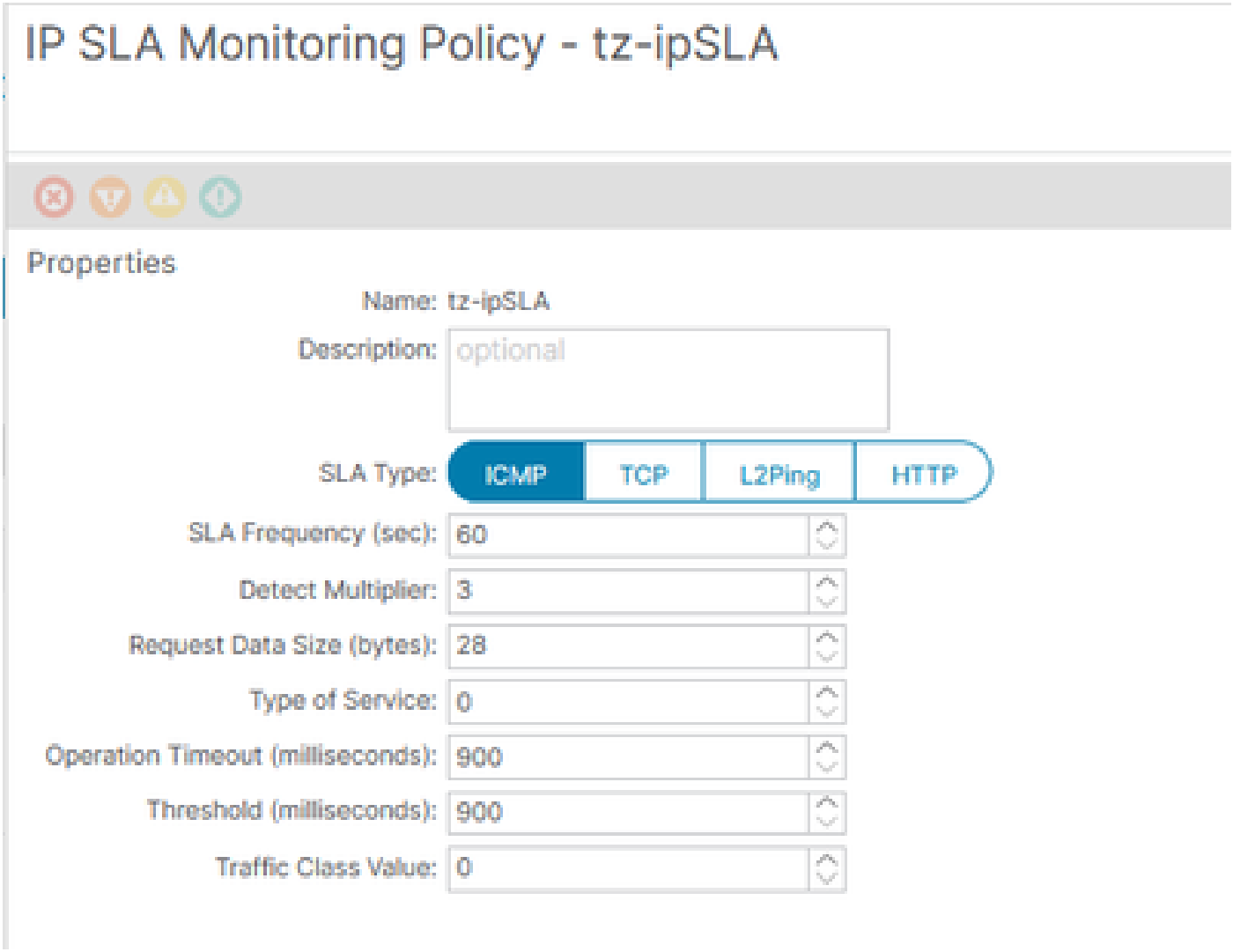
5. Le trafic réseau restant respecte les étapes établies mentionnées jusqu'à présent, dans lesquelles les paquets retournent aux commutateurs spine pour déterminer la destination finale du serveur, sur la base de la recherche de coopération. Ensuite, les paquets sont dirigés vers le commutateur de feuille de calcul qui possède l'indicateur d'apprentissage local et diffusent ces informations à la table COOP sur les spines. L'exécution d'ELAM pour les étapes 2 et 3 est cohérente pour la distribution du paquet vers sa destination finale. Il est essentiel de valider ces informations sur l'EPG de destination pcTag et l'interface de sortie pour garantir une livraison précise.

IP SLA

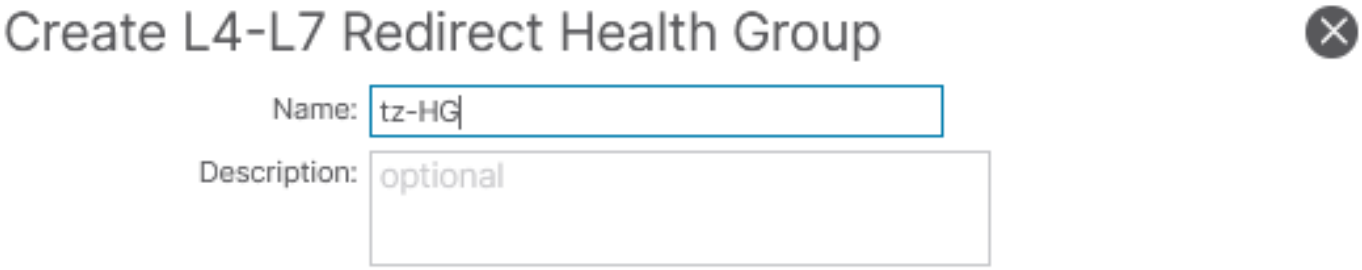
IP SLA est utilisé pour évaluer l’état opérationnel et les performances des chemins réseau. Il permet de garantir que le trafic est acheminé efficacement conformément aux politiques définies, en fonction des conditions du réseau en temps réel. Dans l'ACI, PBR exploite IP SLA pour prendre des décisions de routage éclairées. Si les métriques IP SLA indiquent qu'un chemin est en cours d'exécution, PBR peut réacheminer le trafic par d'autres chemins qui répondent aux critères de performance requis.

À partir de la version 5.2(1), vous pouvez activer le suivi MAC dynamique pour IP SLA, ce qui est utile dans les scénarios où un noeud PBR bascule et modifie l'adresse MAC pour la même adresse IP. Dans les déploiements statiques, une modification de la stratégie de redirection basée sur la stratégie doit être effectuée chaque fois que le noeud PBR modifie son adresse MAC pour continuer à envoyer du trafic. Avec IP SLA, l'adresse MAC utilisée dans cette stratégie est transmise à la réponse de la sonde. Différentes sondes peuvent être utilisées pour déterminer si le noeud PBR est sain ou non, au moment de cette écriture, celles-ci incluent : ICMP, TCP, L2Ping et HTTP.

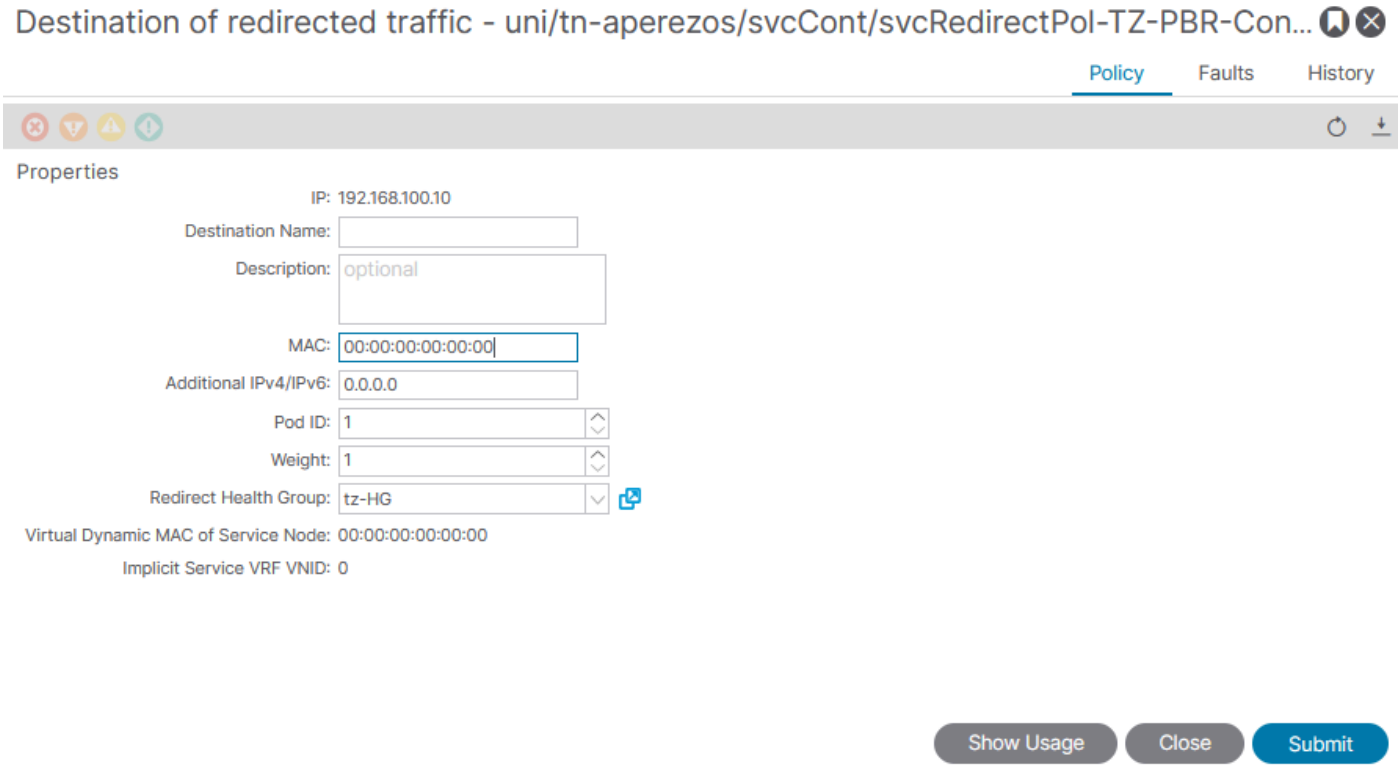
La configuration générale d’une stratégie IP SLA doit ressembler à ceci :



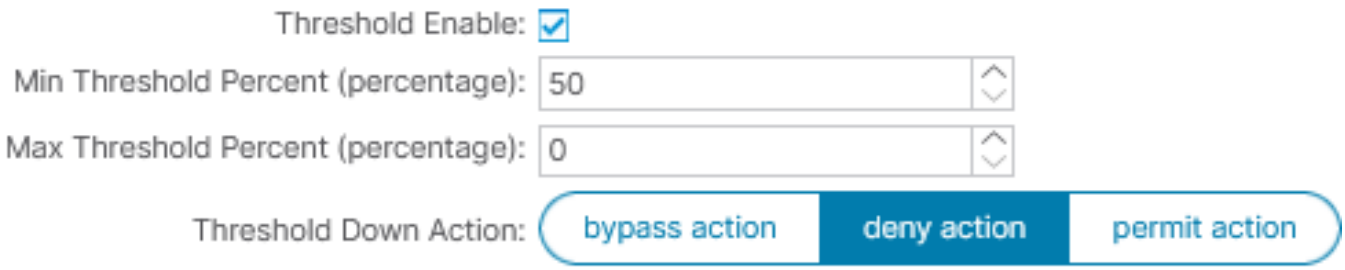
La stratégie IP SLA doit être mappée à l'adresse IP du noeud PBR au moyen d'un groupe d'intégrité.



Si IP SLA est utilisé pour détecter dynamiquement l'adresse MAC, ce champ ne peut pas être vide, mais uniquement défini sur des 0 :



Une fois qu'un groupe d'intégrité s'associe à l'adresse IP du noeud PBR via la destination de couche 3 dans la stratégie de redirection basée sur les stratégies de couches 4 à 7, définissez des seuils pour définir le comportement IP SLA en cas d'inaccessibilité. Spécifiez le nombre minimal de destinations L3 actives requises pour gérer la redirection. Si le nombre de noeuds PBR actifs est inférieur ou supérieur au pourcentage de seuil, l'ensemble du groupe est affecté par l'action Seuil inférieur sélectionnée, ce qui arrête la redistribution. Cette approche permet de contourner le noeud PBR pendant le dépannage sans affecter le flux de trafic.



Les groupes d'intégrité relient toutes les adresses IP définies dans les destinations de couche 3 d'une seule stratégie de redirection de base de stratégie de couche 4 à 7 ou de plusieurs stratégies de redirection de base de stratégie de couche 4 à 7, tant que la même stratégie de groupe d'intégrité est utilisée.

```
Leaf101# show service redir info health-group aperezos::tz-HG
==========================================================================================================
LEGEND
TL: Threshold(Low) | TH: Threshold(High) | HP: HashProfile | HG: HealthGrp | BAC: Backup-Dest | TRA: Tr
==========================================================================================================
HG-Name HG-OperSt HG-Dest HG-Dest-OperSt
======= ========= ======= ==============
aperezos::tz-HG enabled dest-[192.168.100.10]-[vxlan-2162692]] up
```

```
                    dest-[192.168.200.20]-[vxlan-2162692]] up
Leaf1011#
```

Informations connexes

- Livre blanc sur la conception du graphique de services de redirection basés sur les politiques d'infrastructure axée sur les applications Cisco
- ACI Couche 4-7 Redirection basée sur les politiques (PBR) - Explications et conseils approfondis (présentation Cisco Live)
- Dépannage du SLA IP sur le PBR multipod

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.