排除ACI中的PBR故障

目录

简介

本文档介绍如何在单个Pod交换矩阵中使用基于策略的重定向(PBR)对以应用为中心的基础设施(ACI)环境进行故障排除。

先决条件

要求

对于本文而言，建议您了解以下主题的一般知识：

- ACI概念：访问策略、终端学习、合同和L3out

使用的组件

此故障排除练习是在ACI版本6.0(8f)上进行的，使用第二代Nexus交换机N9K-C93180YC-EX和N9K-C93240YC-FX2。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

缩写

- BD:网桥域
- EPG:终端组

- 类 ID:标识EPG的标记
- PBR节点：用于PBR目标的L4-L7设备
- 消费者连接器：面向用户端的PBR节点接口
- 提供程序连接器：面向提供商端的PBR节点接口

PBR历史记录

version	主要功能
2.0（1米）	• 服务图提供PBR功能。
3.x及更低版本	• 多节点基于策略的重定向(PBR)支持 • PBR弹性散列
3.2(x)	• 多站点环境中支持单节点防火墙PBR。
4.0(x)	• 多站点环境中支持双节点防火墙PBR。
4.2(1)	• 现在支持用于创建备用PBR节点的备份策略。
4.2(3)	• 使用GUI的Service Graph模板中提供Filter-from-contract选项。
5.0(1)	• 服务节点的所有提供商端都支持L3Outs。 • 第1层/第2层PBR设备支持主动 — 主动部署/ECMP路径。
5.2(1)	• 现在，PBR目标可以位于L3Out中。 • 您可以使用HTTP URI跟踪服务节点。 • 不再支持服务图托管和混合模式。 • 支持动态PBR节点mac地址。
6.0(1)	• 基于权重的对称策略重定向(PBR)

设计注意事项

- PBR可同时与物理和虚拟设备配合使用
- PBR可以在L3Out EPG和EPG之间、EPG之间以及L3Out EPG之间使用。如果L2Out EPG是合同的一部分，则不支持PBR
- 思科ACI多Pod、多站点和远程枝叶环境支持PBR。
- 思科ACI交换矩阵必须是服务器和PBR节点的网关
- L4-L7设备必须在转至模式（路由模式）下部署

- FEX交换机不支持PBR节点
- PBR节点需要专用网桥域
- 现在，使用健康组支持PBR节点的动态MAC地址。
- 建议在PBR节点网桥域上启用基于GARP的检测
- 包括ARP、以太网流量和其他非IP流量的常见默认过滤器不得用于PBR
- PBR节点可以在VRF实例之间或在一个VRF实例内。对于此拓扑，不支持在第三个VRF上使用PBR节点，必须在消费者或提供商VRF中配置

背景信息

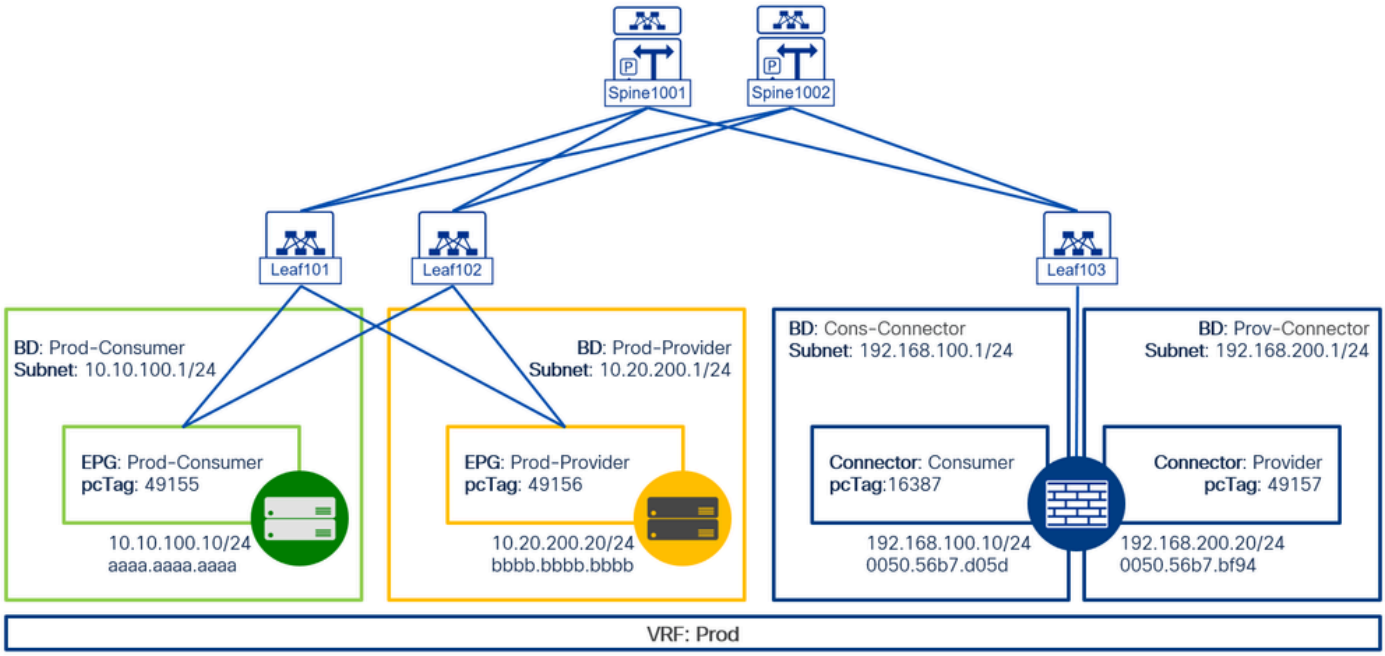
有关ELAM和Ftriage的更多说明，请参阅会话BRKDCN-3900b中的CiscoLive按需库。

此外，所有配置指南都可以在Cisco Application Centric Infrastructure Policy-Based Redirect Service Graph Design White Paper中找到。

场景:单个Pod交换矩阵中的单个VRF

网络图

物理拓扑:



故障排除

步骤 1：故障

当配置或策略交互存在问题时，ACI会生成故障。在出现故障时，已为PBR呈现过程识别出特定故障：

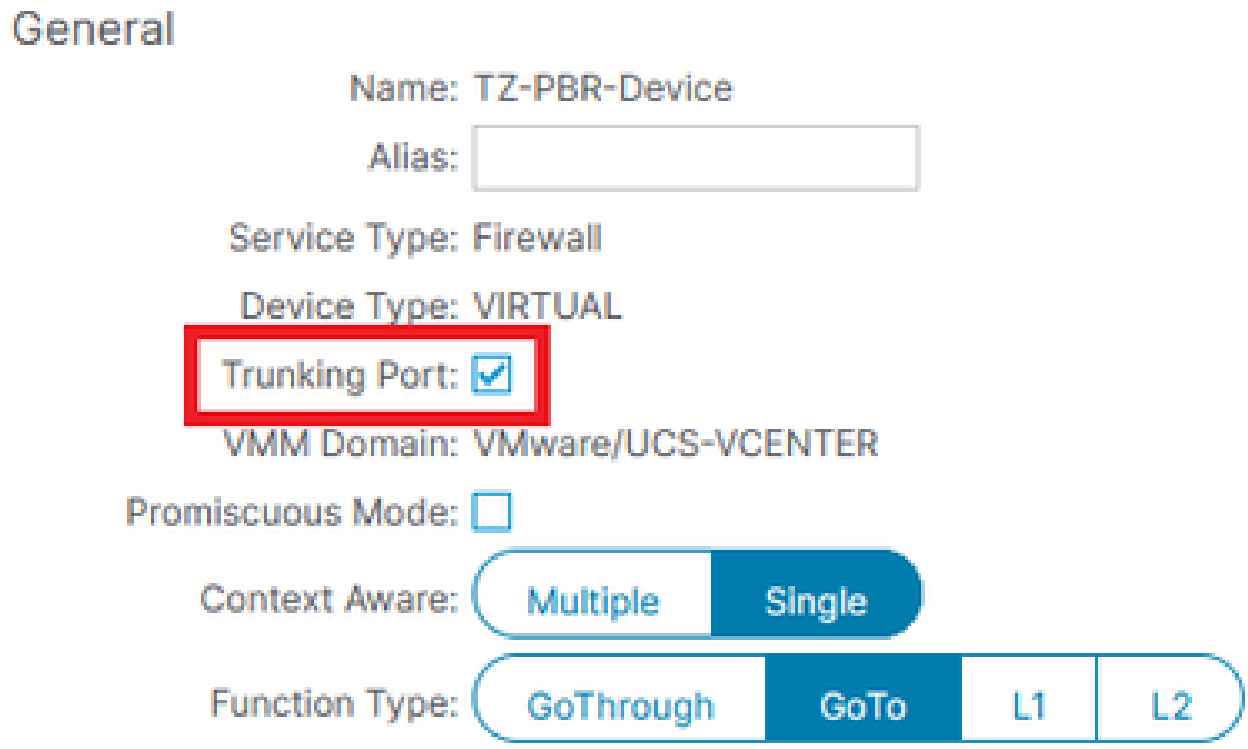
F1690:配置无效，原因如下：

- ID分配失败

此故障表示服务节点的封装VLAN不可用。例如，在与逻辑设备使用的Virtual Machine Manager(VMM)域关联的VLAN池中可能没有可用的动态VLAN。

分辨率：验证逻辑设备使用的域内的VLAN池。如果逻辑设备接口在物理域内，还要检查封装的VLAN配置。这些设置可在Tenant > Services > L4-L7 > Devices and Fabric > Access Policies > Pools > VLAN下找到。

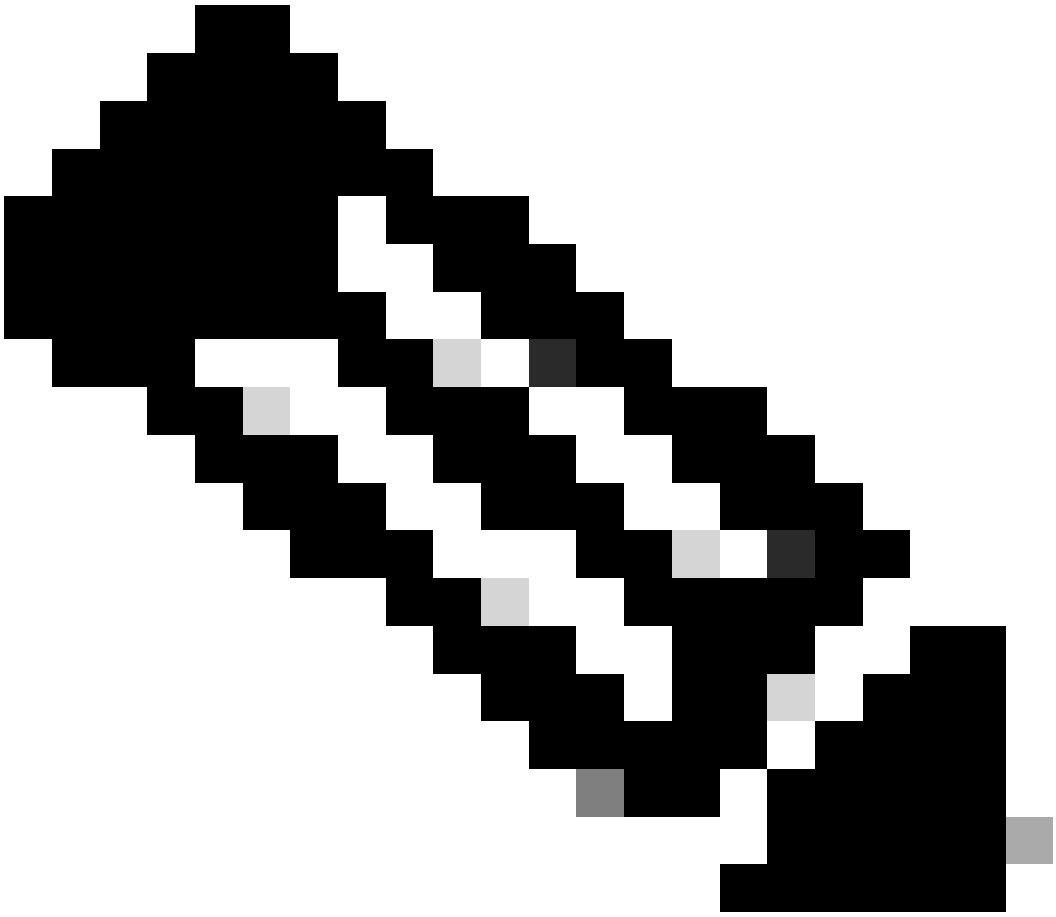
相反，如果逻辑设备接口位于虚拟域内并通过中继接口连接到您的ESXi主机，请确保启用中继端口选项。



- 找不到LDev的设备上下文

此故障表示逻辑设备不可定位用于服务图呈现。例如，可能没有与与Service Graph关联的合同相匹配的设备选择策略。

分辨率：验证设备选择策略已定义。设备选择策略根据合同名称、服务图名称和服务图中的节点名称指定服务设备及其连接器的选择条件。可在Tenant > Services > L4-L7 > Device Selection Policy下找到此项。

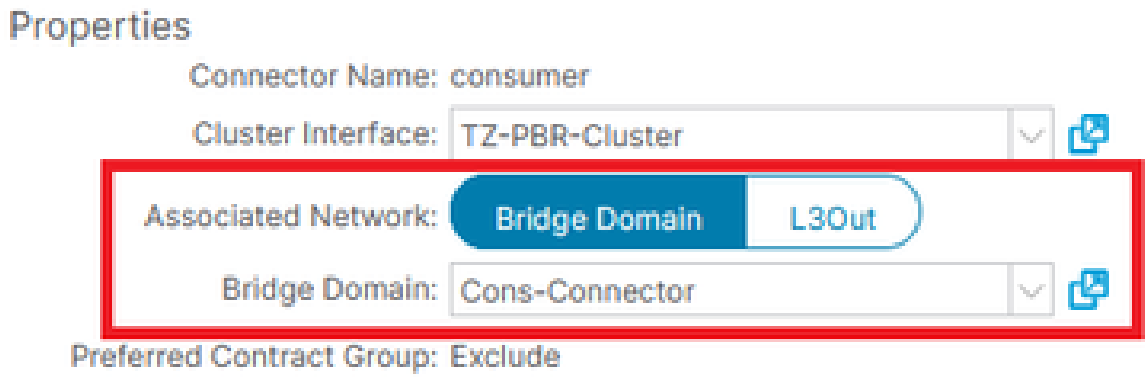


注意：部署服务图模板时，ACI会预先为源EPG选择网桥域。您必须为PBR使用者连接器更改此网桥域。这也适用于提供商连接器。

- 未找到BD

此故障表示找不到服务节点的桥接域(BD)。例如，设备选择策略中未指定BD。

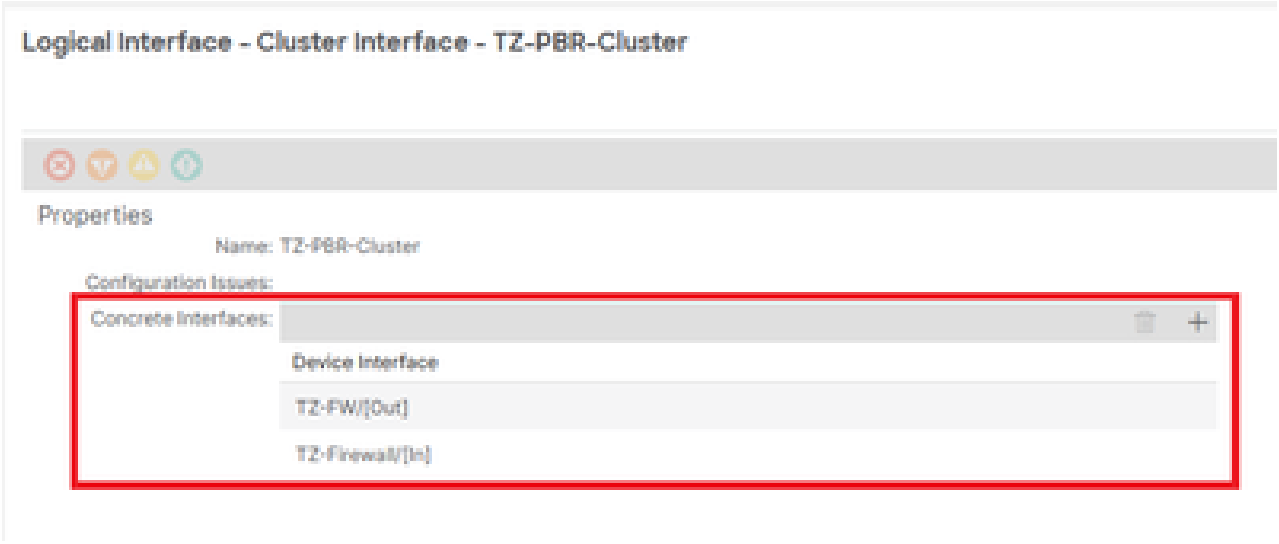
分辨率：确保已在设备选择策略中指定BD，并且连接器名称正确。此配置位于Tenant > Services > L4-L7 > Devices Selection Policy > [Contract + SG] > [Consumer |提供商].



- LIf与CIf无关，并且LIf具有无效的CIf
- 未找到集群接口

这些故障表明设备与集群接口没有关系。

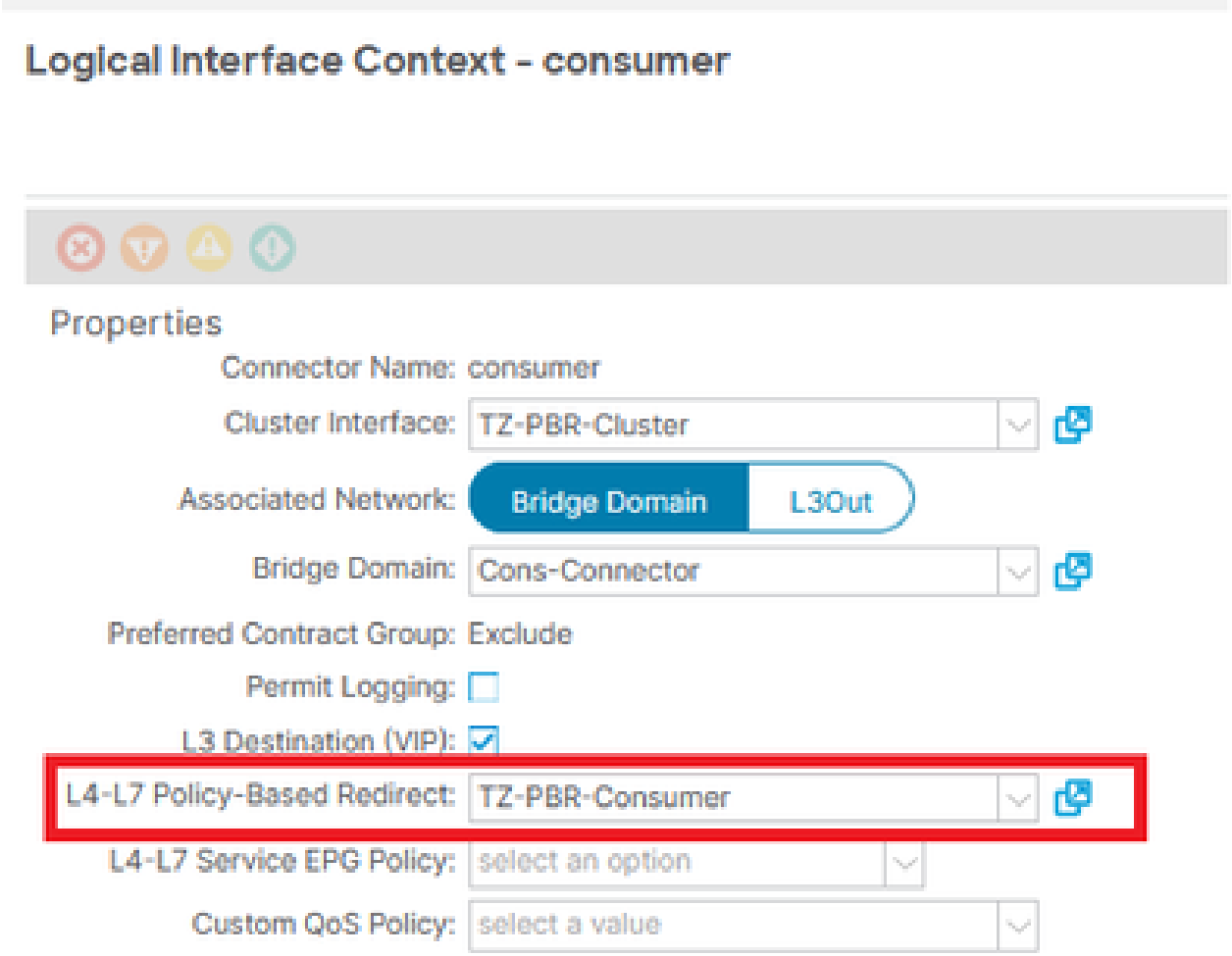
分辨率：确保第4层到第7层(L4-L7)设备配置包含指定的具体接口选择器。此配置位于Tenant > Services > L4-L7 > Devices > [Device] > Cluster Interfaces。



- 服务重定向策略无效

此故障表示尽管已在服务图中的服务功能上激活重定向，但尚未应用PBR策略。

分辨率：确保PBR策略在设备选择策略设置中配置。此配置位于Tenant > Services > L4-L7 > Devices Selection Policy > [Contract + SG] > [Consumer |提供商].



F0759:graph-rendering-failure — “无法实例化租户< tenant >的服务图。Function node < node > configuration is invalid."

由于功能节点名称的配置无效，无法实例化指定租户的服务图。

此错误表明存在与上述条件相关的配置问题。

此外，在初始部署期间，此故障可能暂时出现，然后得到及时解决。这是由于ACI部署所有策略所经历的呈现过程所致。

分辨率：调查已报告的任何其他故障并相应地解决它们。

F0764:configuration-failed — “L4-L7 Devices configuration < device > for tenant < tenant > is invalid.”

由于PBR设备策略的配置无效，无法实例化指定租户的服务图。

此错误表明存在与上述条件相关的配置问题。

分辨率：调查已报告的任何其他故障并相应地解决它们。

F0772:configuration-failed - "LIf configuration < cluster > for L4-L7 Devices < device > for tenant < tenant > is invalid."

由于PBR设备集群接口选择的配置无效，无法实例化指定租户的服务图。

此错误表明存在与上述条件相关的配置问题。

分辨率：调查已报告的任何其他故障并相应地解决它们。

步骤 2：源和目标终端学习

确保在交换矩阵中识别您的源终端和目标终端，这就需要进行基本配置：

- 虚拟路由和转发(VRF)对象。
- 启用单播路由的网桥域(BD)对象。虽然启用IP数据平面学习被视为最佳实践，但并非强制性的。
- 终端组(EPG)对象，适用于虚拟和物理域。
- 访问策略：验证访问策略配置链是否完整，以及EPG上是否未报告任何故障。

要在正确的EPG和接口上确认终端学习，请在获取终端（也称为计算枝叶）后的枝叶上执行此命令：

<#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#
```

此命令允许您标识与终端分类的EPG关联的pcTag(class)，以及检索接口、VRF范围和MAC地址信息。

如果您不知道源或目标端点的位置，您可以随时在APIC上执行此命令：

<#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
----------------- ----------------------------------------- ------------ ---------- --------------------
```

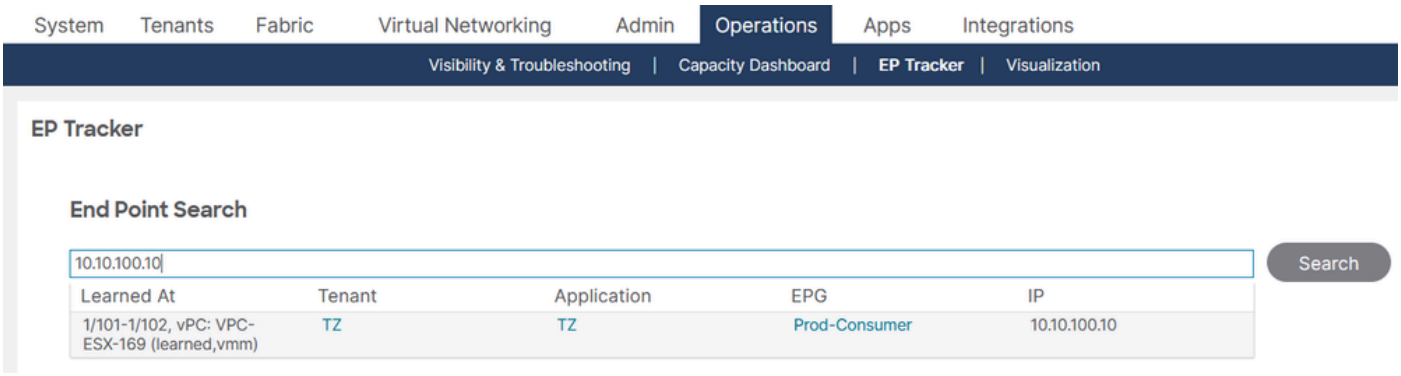
AA:AA:AA:AA:AA:AA 10.10.100.10

```
                            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#
```

在GUI中，可以导航到操作(Operations)> EP跟踪器访问EP跟踪器功能，以进行终端监控和管理。



利用从源端点和目标端点收集的信息，您现在可以集中精力部署PBR策略。

步骤 3：重定向合同

PBR集成在服务图框架中。因此，必须根据合同在源交换机和目标交换机上部署和配置Service Graph模板。利用上一步收集的pcTags信息，您可以执行此命令来确定是否已将终端组(EPG)重定向到服务图组。

<#root>

show zoning-rule scope [vrf_scope]

在分区规则中，必须考虑以下规则：

1. 源EPG到目标EPG与关联的重定向组
2. 源PBR节点影子EPG到源EPG
3. 目标EPG到具有关联重定向组的源EPG。此组可以与之前的配置相同或不同。
4. 目标PBR节点影子EPG到目标EPG

<#root>

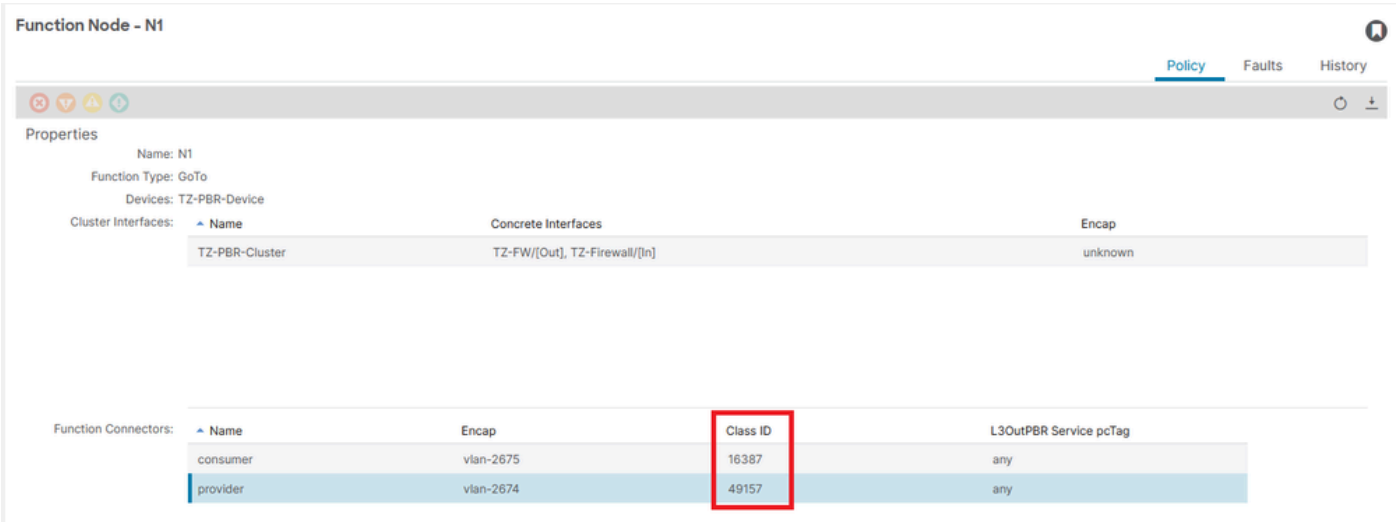
Leaf101#

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#
```

要验证在基于策略的路由(PBR)部署过程中创建的影子EPG的pcTag，请导航到Tenants > [TENANT_NAME] > Services > L4-L7 > Deployed Graph Instance > [SG_NAME] > Function

Node - N1。



- 合同解析器

该脚本将分区规则、过滤器、统计信息和EPG名称相关联。您可以安全地直接在ACI枝叶或APIC上执行此脚本。在APIC上运行时，它会收集所有枝叶交换机上的具体对象，对于大型策略部署，这可能需要几分钟的时间。

从ACI版本3.2开始，contract_parser捆绑在映像中并在枝叶上可用。只需从iBash shell输入contract_parser.py。

<#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#

此命令提供合同操作、源和目标EPG、使用的合同名称和命中计数等详细信息。

步骤 4：重定向组

根据应用于分区规则的合同识别重定向组后，下一步是确定重定向目标设备的IP和MAC地址。为此，请执行以下命令：

<#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#
```

此命令使我们能够确定重定向组的运行状态(OperSt)、L4-L7 PBR部分中配置的IP地址，以及与PBR节点网桥域关联的VRF的VNID。您现在需要确定配置的MAC地址：

<#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-
====                                ======         ====              ====    =====   =========    ===
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-
====                                    ======         ====              ====    =====   =========    ====
dest-[
```

192.168.200.20

```
]-[vxlan-
```

2162692

```
] vxlan-
```

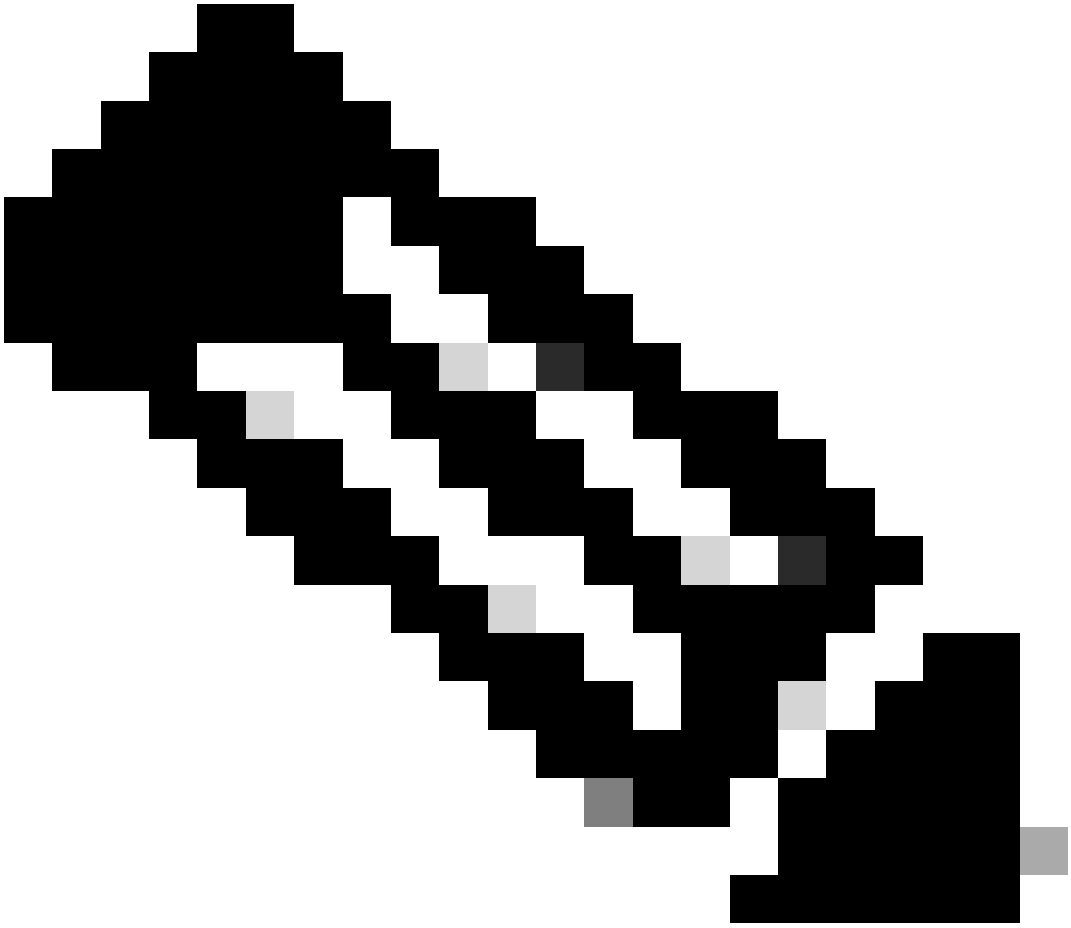
16646036 00:50:56:B7:BF:94

```
 TZ:Prod
```

enabled

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

除了上述详细信息，此命令还提供有价值的见解，包括VRF名称、BD VNID和PBR节点的已配置MAC地址。



注意：需要注意的是，IP和MAC地址在此阶段都是用户配置的，这意味着在第4-7层基于策略的路由定义期间可能出现排版错误。

步骤 5：PBR节点不接收任何流量。

PBR转发遇到的普遍问题是没有流量到达PBR节点。此问题的一个常见原因是L4-L7基于策略的路由配置中的MAC地址指定不正确。

要验证在L4-L7基于策略的路由中配置的MAC地址的准确性，请执行步骤2中以前使用的命令。此命令可在指定为服务枝叶的枝叶交换机上执行，预期节点将在该枝叶交换机上获知。

<#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#
```

验证EPM表中记录的MAC地址是否与服务重定向组中配置的MAC地址匹配。即使是轻微的排版错误也必须纠正，以确保适当的流量路由到PBR节点目标。

步骤 6：通信流.

- FTRIAGE

用于APIC的CLI工具旨在自动配置和解释ELAM端到端流程。该工具允许用户指定特定流以及流源所在的枝叶交换机。它按顺序在每个节点上执行ELAM，以分析流的转发路径。在数据包路径不易识别的复杂拓扑中，此工具尤其有用。

```
<#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:

嵌入式逻辑分析器模块(ELAM)是一个诊断工具，使用户能够在硬件中建立特定条件，以捕获满足这些条件的初始数据包或帧。捕获成功后，ELAM状态显示为已触发。触发时，ELAM被禁用，允许收集数据转储，这有助于分析交换机ASIC为该数据包或帧执行的许多转发决策。ELAM在ASIC级别运行，确保它不会影响交换机的CPU或其他资源。

命令语法的结构。此结构摘自ACI交换矩阵内转发工具故障排除一书

```
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]
```

设置触发条件

```
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]
```

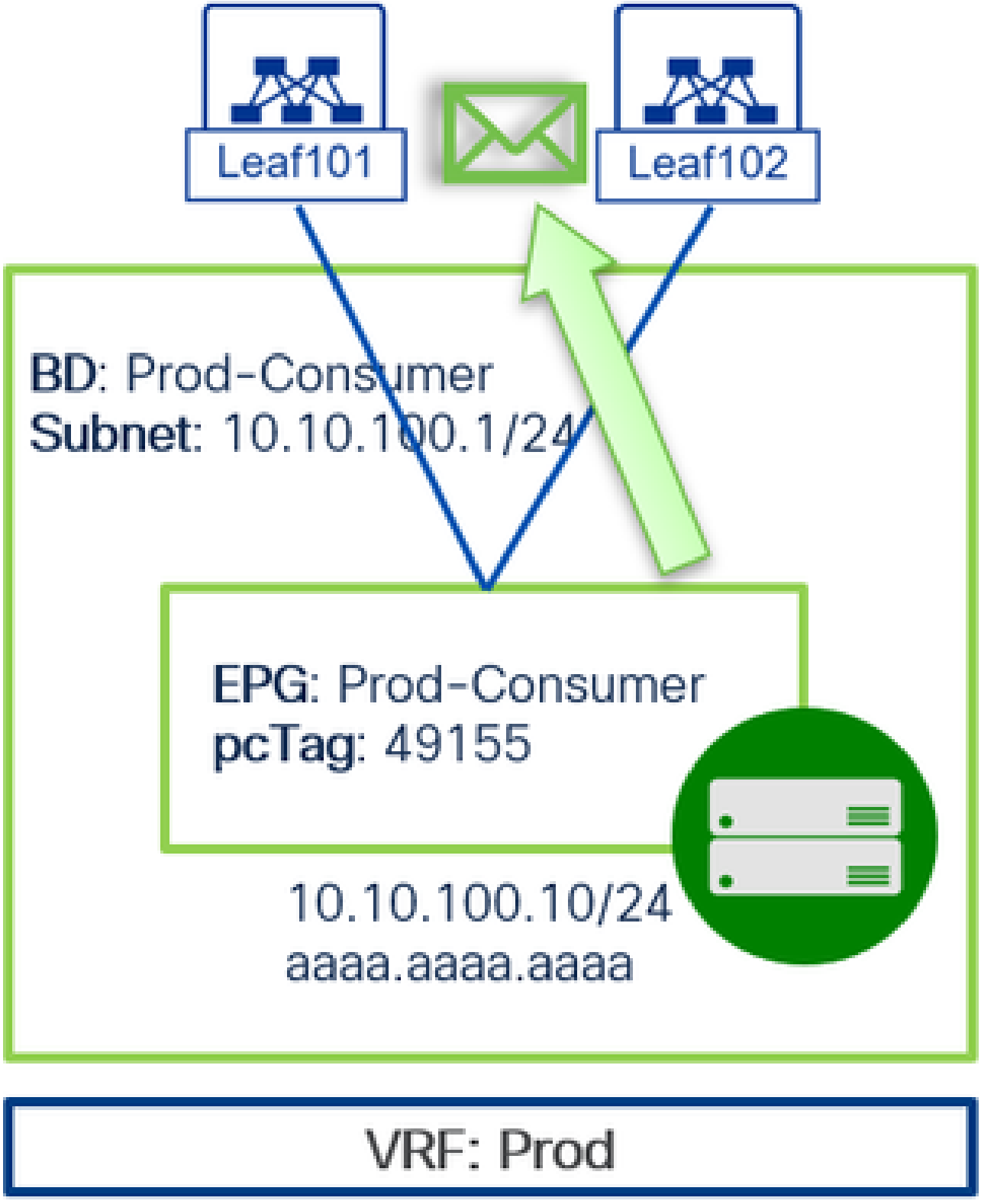
生成包含数据包分析的转储。

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

- 流量传输

理解流经所有有问题的设备的流量至关重要。Ftriage工具提供了此流程的精彩摘要。但是，为了进行详细的分步验证，并深入了解数据包接收过程，您可以在网络拓扑中的每一点执行嵌入式逻辑分析器模块(ELAM)。

1.入口流量发生在获取源服务器的计算枝叶上。在此特定场景中，由于源位于vPC接口之后，必须在vPC对等体上配置ELAM。这是必要的，因为散列算法选择的物理接口是不确定的。



<#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

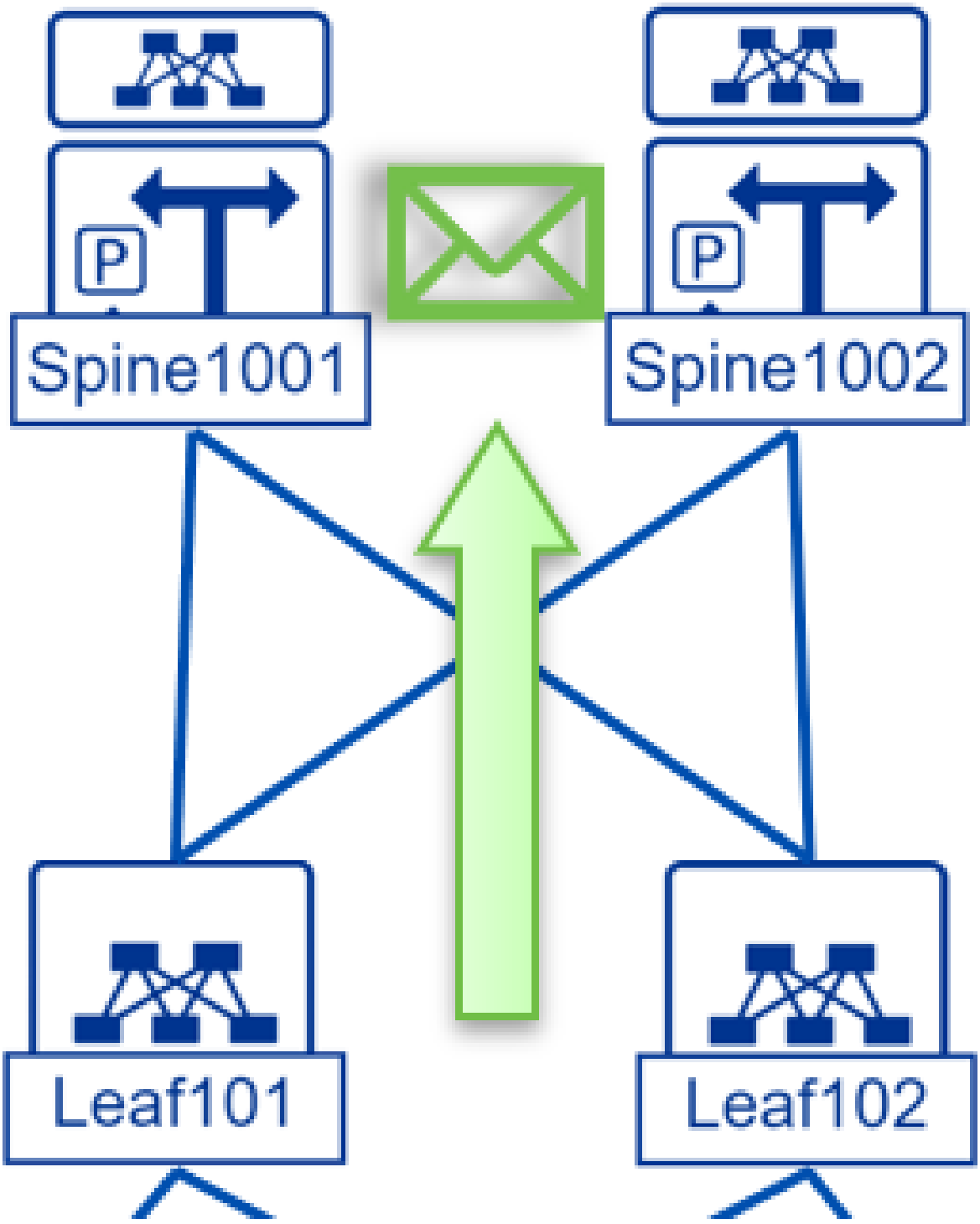
根据提供的信息，显然数据包通过基于策略的路由(PBR)重定向，因为service_redir选项已启用。此外，还可以检索class和dclass值。在此特定场景中，交换机可识别dclass。但是，如果EPM表中不存在目标端点，则dclass值默认为1。

此外，入口接口由SRCID确定，出口接口由向量值确定。通过在vsh_lc级别执行此命令，可以将这些值转换为前端端口：

<#root>

show platform internal hal l2 port gpd

2.流程中的后续步骤涉及到达主干交换机，以将目的MAC地址映射到PBR节点。由于流量封装在VXLAN报头中，在脊柱或远程枝叶上执行ELAM需要使用选择14来正确解码封装。



<#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.

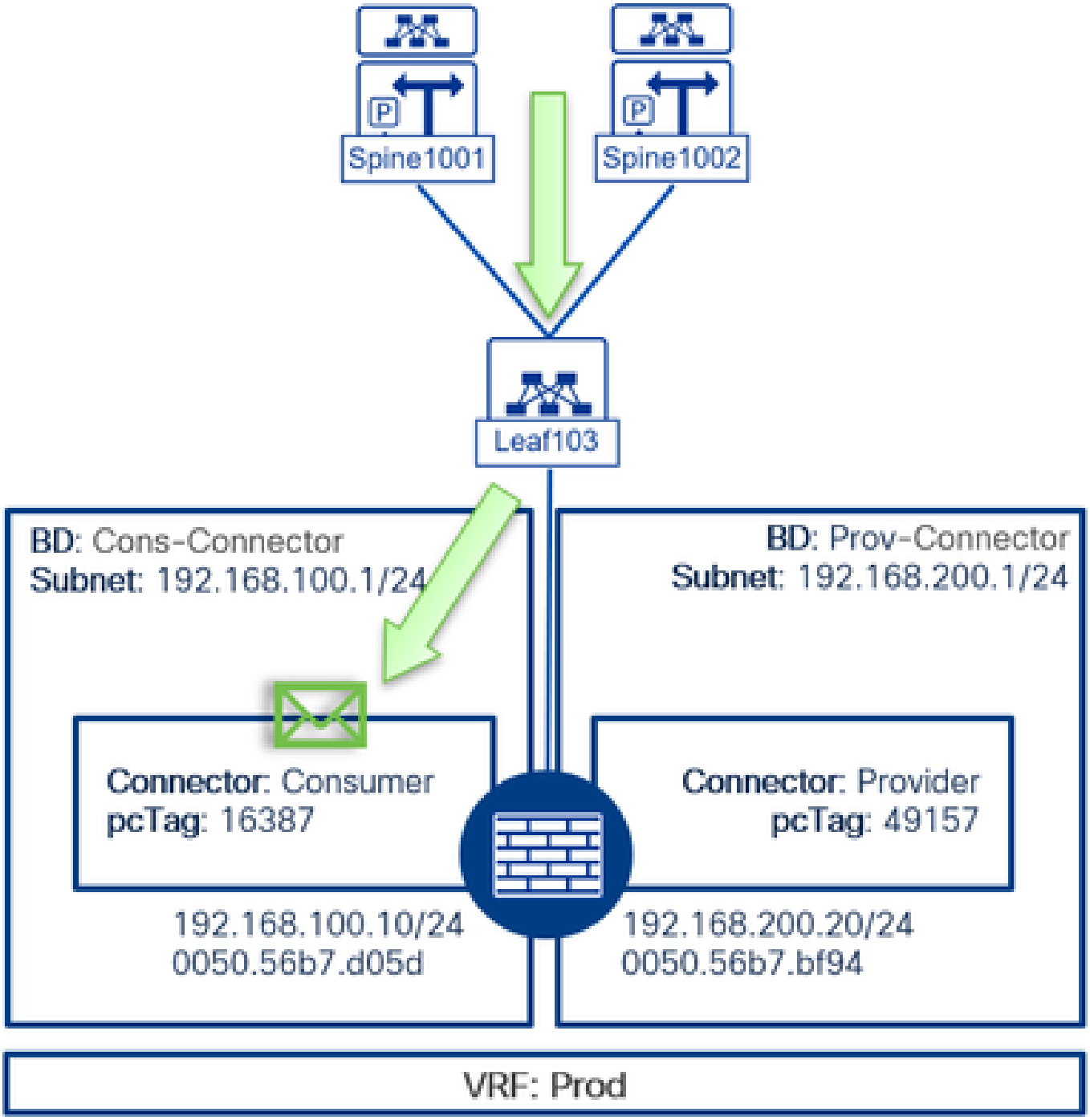
ovector_idx: 0x1F0 >>> Eth1/10

从之前的输出可以明显看出，目的MAC地址已重写为防火墙的MAC地址。随后，执行COOP查找以标识MAC的目标发布方，然后将数据包转发到交换机的相应接口。

您可以通过执行以下命令，利用网桥域VNID和防火墙的MAC地址模拟主干上的此查找：

```
<#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.流量到达识别防火墙MAC地址的服务枝叶，然后转发到PBR节点。



<#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
```

---

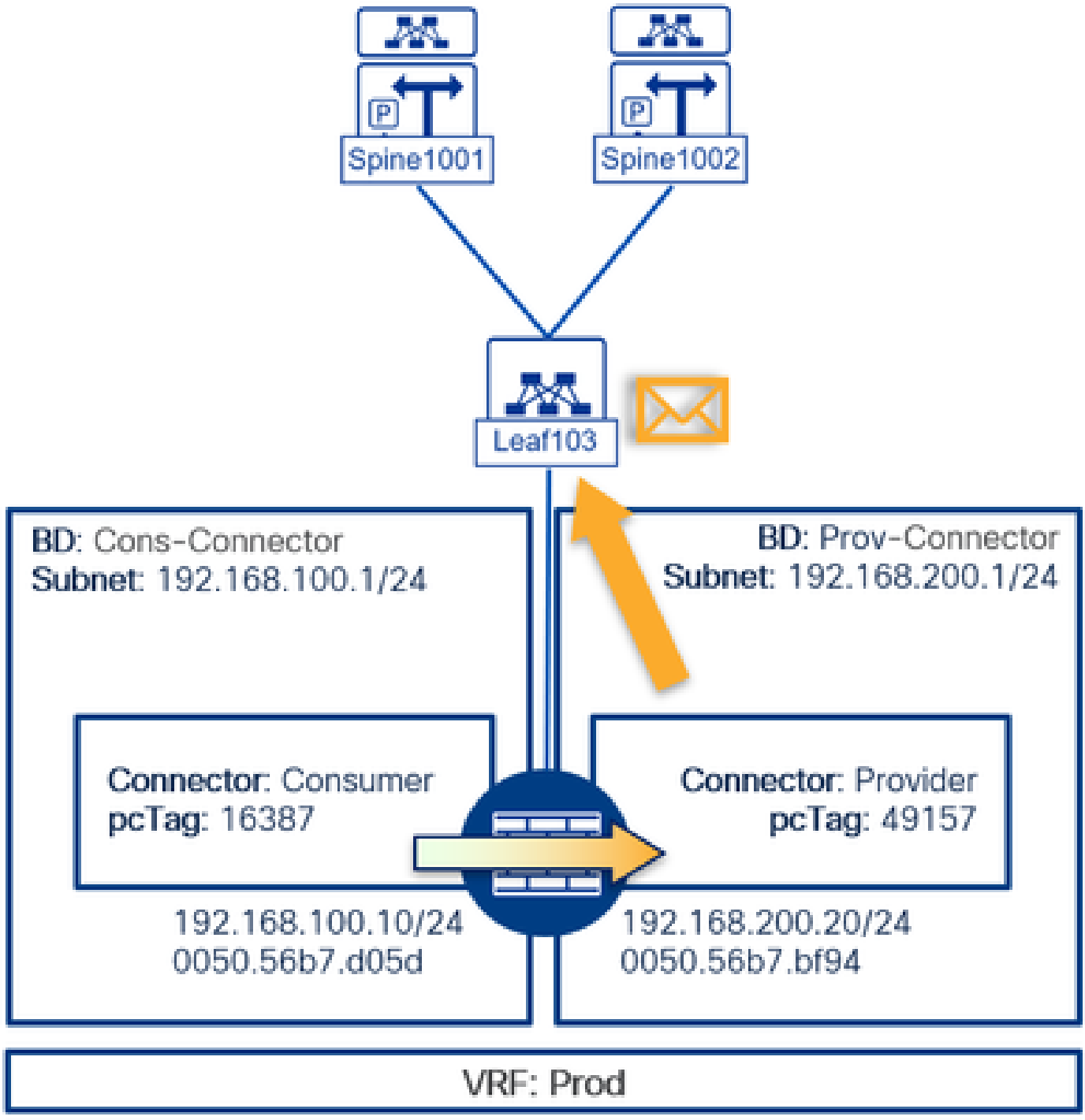
ovector

:

64(0x40) >>> Eth1/45

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

4.对于由PBR节点返回的打包程序，首先这个节点必须自己进行检查，并更改VRF、接口或VLAN。然后，数据包将转发回提供商连接器上的ACI:



<#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


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

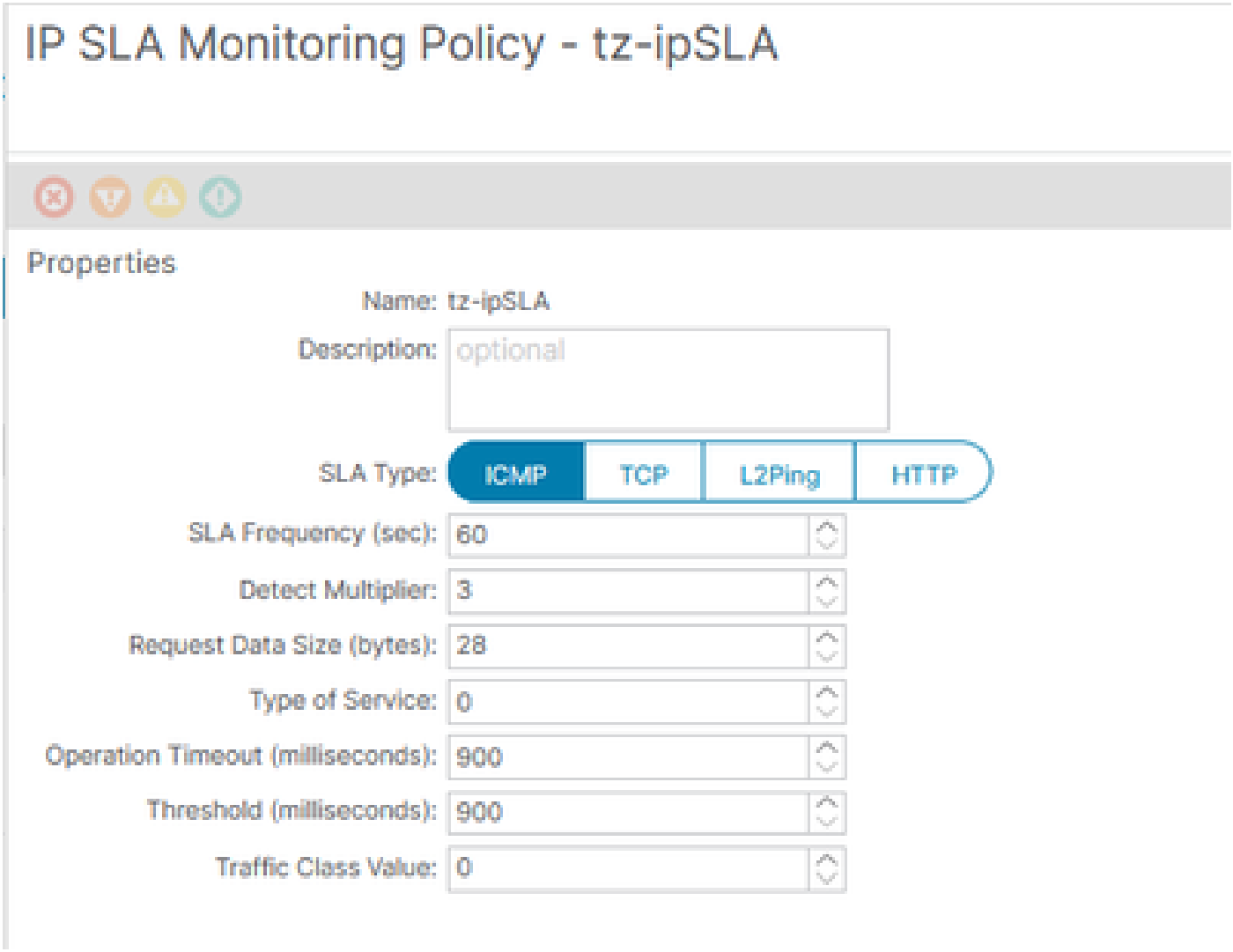
5. 剩余的网络流量遵循目前为止提到的既定步骤，即数据包返回主干交换机，以根据coop查找确定最终服务器目标。随后，将数据包定向到具有本地学习标志的计算枝叶交换机，并将此信息传播到主干上的COOP表。步骤2和步骤3的ELAM执行对于将数据包分发到其最终目的地是一致的。必须验证这些目的EPG pcTag和出口接口，以确保准确交付。

IP SLA

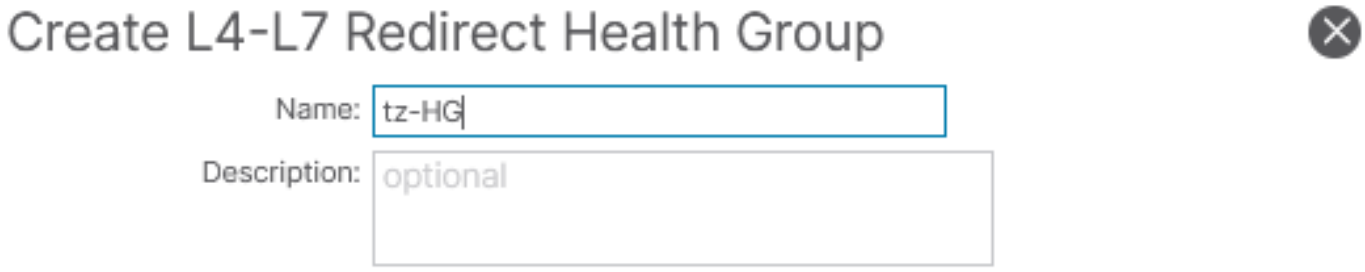
IP SLA用于评估网络路径的运行状态和性能。它有助于确保根据定义的策略基于实时网络条件有效地路由流量。在ACI中，PBR利用IP SLA做出明智的路由决策。如果IP SLA度量指示某个路径正在执行中，则PBR可以通过满足所需性能标准的备用路径重新路由流量。

从版本5.2(1)开始，您可以为IP SLA启用动态MAC跟踪，这对于PBR节点进行故障转移并更改同一IP地址的MAC地址的情况非常有用，在静态部署中，每次更改PBR节点的MAC地址以继续发送流量时，都需要更改基于策略的重定向策略。使用IP SLA时，此策略中使用的MAC地址会中继到探测响应。在撰写本文时，可以使用不同的探针来确定PBR节点是否正常，这些探针包括：ICMP、TCP、L2Ping和HTTP。

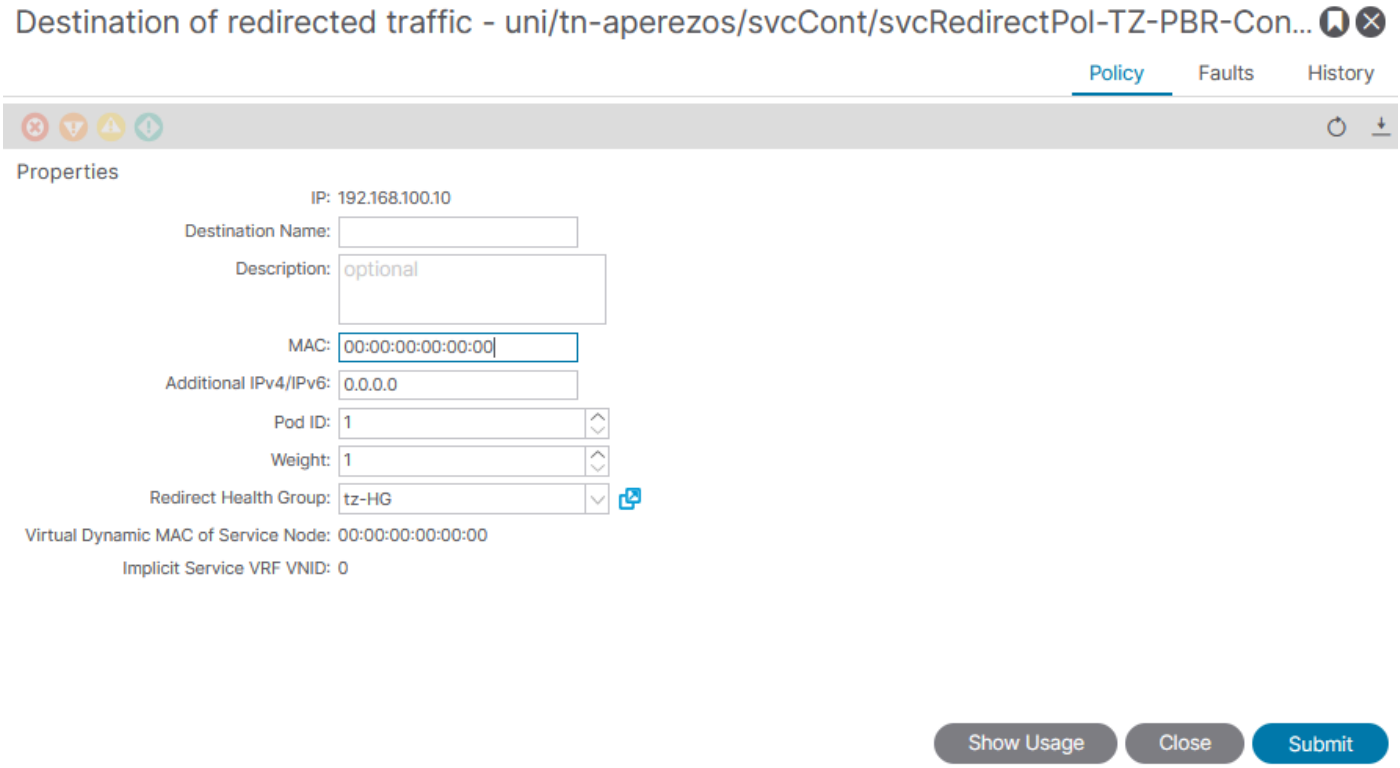
IP SLA策略的常规配置必须如下所示：



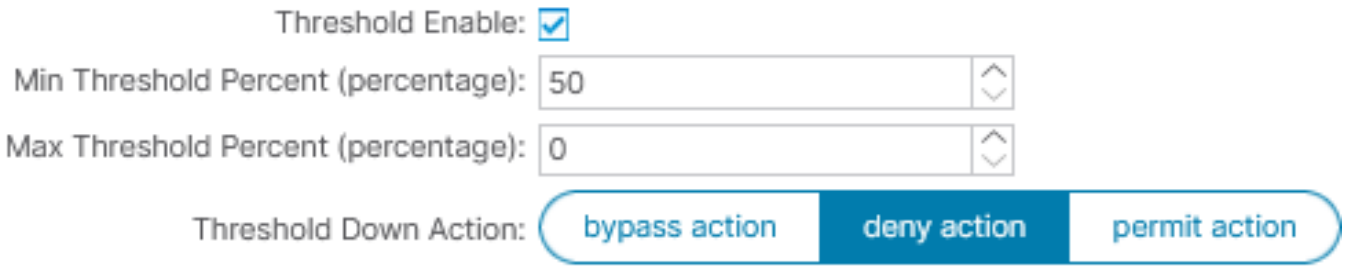
IP SLA策略必须通过健康组映射到PBR节点IP。



如果使用动态发现MAC地址的IP SLA，则此字段不能为空，但设置为所有0:



一旦运行状况组通过L4-L7基于策略的重定向策略中的L3目标与PBR节点IP关联，请设置阈值以定义不可达时的IP SLA行为。指定维护重定向所需的最小活动L3目标数。如果活动PBR节点的计数低于或超过阈值百分比，则整个组会受到所选Threshold Down Action的影响，从而停止重分发。此方法支持在故障排除期间绕过PBR节点，而不影响流量。



运行状况组将单个L4-L7策略库重定向策略或多个L4-L7策略库重定向策略的L3目标中定义的所有IP绑定在一起，只要使用相同的运行状况组策略即可。

```
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#
```

相关信息

- 思科以应用为中心的基础设施基于策略的重定向服务图设计白皮书
- ACI第4-7层基于策略的重定向(PBR)深入探讨和提示（Cisco Live演示）
- 排除Multipod PBR上的IP SLA故障

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。