

排除ACI交换矩阵内转发 — 第2层转发故障

目录

[简介](#)

[背景信息](#)

[概述](#)

[拓扑](#)

[GUI检查](#)

[已知第2层单播流量的故障排除工作流程](#)

[入口枝叶源EP MAC学习](#)

[入口枝叶目标MAC终端查找](#)

[入口枝叶交换机发送到主干交换机](#)

[主干转发](#)

[出口枝叶远程EP MAC学习](#)

[出口枝叶目标MAC查找](#)

[验证是否在主干交换机的COOP EP回拨中正确获知了两个终端](#)

[使用ELAM Assistant的ELAM输出](#)

[使用CLI的入口枝叶ELAM](#)

[使用fTriage跟踪流](#)

[未知第2层单播流量的故障排除工作流程 — BD处于泛洪模式](#)

[查找BD GIPo](#)

[ELAM — 入口枝叶 — 泛洪流量](#)

[绘制FTAG拓扑](#)

[ELAM — 出口枝叶 — 泛洪流量](#)

[未知第2层单播流量的故障排除工作流程 — 硬件代理中的BD](#)

[第2层转发摘要](#)

[ACI交换矩阵第2层转发行为](#)

简介

本文档介绍了解第2层转发并对其进行故障排除的步骤

背景信息

本文档中的内容摘自 [思科以应用为中心的基础设施故障排除 \(第二版\)](#) 书，特别是 [交换矩阵内转发 — L2转发：同一BD中的两个终端 — 无单播路由](#) 第章。

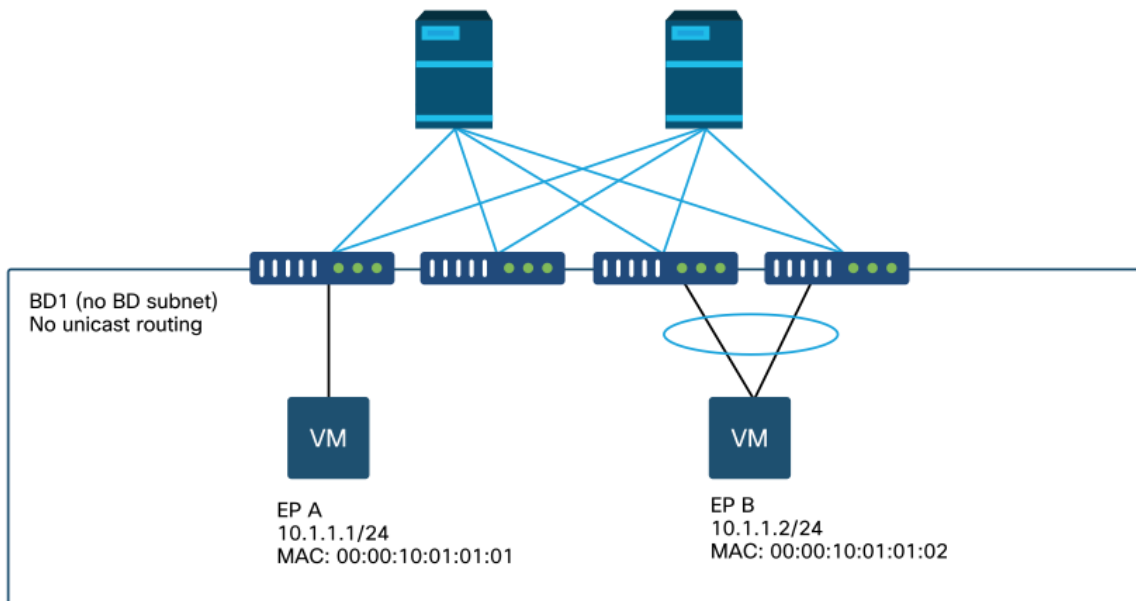
概述

本节介绍同一网桥域和子网中的终端无法相互通信的故障排除示例。下图说明了BD没有任何子网且单播路由被禁用的拓扑。

通常，对具有终端连接的流量进行故障排除时，建议开始识别一对终端。请参阅以下带有EP A和

EP B的拓扑。这两个IP地址分别为10.1.1.1/24和10.1.1.2/24。MAC地址分别为00:00:10:01:01:01和00:00:10:01:01:02。

拓扑



本节提供三种方案：

1. 已知的第2层单播流。
2. 泛洪模式下具有BD的未知第2层单播流。
3. 硬件代理模式下的BD第2层未知单播流。

接下来的故障排除流程可按以下方案总结：

- 1级检查：GUI验证所了解的配置、故障和终端。
- 2级检查：枝叶交换机上的CLI: 检查源枝叶交换机和目标枝叶交换机是否了解终端。检查主干节点是否学习COOP中的终端。
- 3级检查：数据包捕获: ELAM (ELAM Assistant或CLI) 验证帧是否存在。分类以跟踪流。

GUI检查

第一级故障排除是从GUI验证终端MAC已正确获知。这可以通过终端所在的EPG的operational选项卡完成。

'EPG Operational选项卡>客户端终端'

								Summary	Policy	Operational	Stats	Health
								Client End-Points	Configured Access Policies	Contracts	Controller End-Points	Deployed Leaves
MAC	IP	Learning Source	Hosting Server	Reporting Controller Name	Interface	Multicast Address	Encap Address					
00:00:10:01:01:01	---	learned	---	---	Pod-1/Node-101/eth1/3 (learned)	---	vlan-2501					
00:00:10:01:01:02	---	learned	---	---	Pod-1/Node-103-104/N3k-3-VPC3-4 (learned)	---	vlan-2501					

Objects Per Page: 15

在这种情况下，终端A和B都显示在GUI中。GUI显示它们的MAC地址、它们连接到交换矩阵的接口以及封装 — 在本例中，两者都处于封装VLAN 2501中。

预计不会从ACI交换矩阵获取IP地址，因为单播路由已在BD级别禁用。

请参阅以上屏幕截图中的学习源列。如果它表示“已获取”，则ACI枝叶交换机至少收到一个来自终端的数据包。

由于在本例中终端是从ACI交换矩阵学习的，因此请继续查看已知第2层单播流量的下一个故障排除案例。

已知第2层单播流量的故障排除工作流程

入口枝叶源EP MAC学习

对于同一BD中的第2层转发，ACI将仅学习源MAC并根据目标MAC进行转发。MAC地址是在BD范围内学习的。

首先，检查终端是否已获知：

```
leaf1# show endpoint mac 0000.1001.0101
Legend:
s - arp                H - vtep                V - vpc-attached        p - peer-aged
R - peer-attached-rl  B - bounce              S - static               M - span
D - bounce-to-proxy   O - peer-attached       a - local-aged          m - svc-mgr
L - local              E - shared-service

+-----+-----+-----+-----+-----+
---+
VLAN/          Encap          MAC Address          MAC Info/          Interface
Domain         VLAN           IP Address           IP Info
+-----+-----+-----+-----+-----+
---+
4/Prod:VRF1                vlan-2501    0000.1001.0101 L
```

eth1/3

以上输出提供以下信息：

- MAC地址0000.1001.0101通过vrf Prod:VRF1中的encapsulation vlan-2501在端口ethernet 1/3上本地获取（本地标志为L）。
- 请参阅上述输出中的“VLAN/域”列。此处列出的VLAN ID是内部VLAN。

入口枝叶目标MAC终端查找

假设目标MAC已知（已知单播）。

```
leaf1# show endpoint mac 0000.1001.0102
```

Legend:

```
s - arp          H - vtep          V - vpc-attached    p - peer-aged
R - peer-attached-r1 B - bounce        S - static          M - span
D - bounce-to-proxy O - peer-attached  a - local-aged     m - svc-mgr
L - local        E - shared-service
```

```
-----+-----+-----+-----+-----+
----+
      VLAN/          Encap          MAC Address          MAC Info/          Interface
      Domain          VLAN          IP Address          IP Info
-----+-----+-----+-----+-----+
----+
7/Prod:VRF1          vxlan-16351141    0000.1001.0102
tunnel4
```

以上输出提供以下信息：

- MAC地址0000.1001.0102不在本地学习。
- 从接口tunnel 4获知。
- 它通过封装VXLAN-16351141获取，该封装对应于网桥域的BD_VNID（VXLAN网络ID）。

接下来，使用“show interface tunnel <x>”命令检查隧道接口的目标

```
leaf1# show interface tunnel 4
```

```
Tunnel4 is up
  MTU 9000 bytes, BW 0 Kbit
  Transport protocol is in VRF "overlay-1"
  Tunnel protocol/transport is vxlan
  Tunnel source 10.0.88.95/32 (lo0)
  Tunnel destination 10.0.96.66
  Last clearing of "show interface" counters never
  Tx
  0 packets output, 1 minute output rate 0 packets/sec
  Rx
  0 packets input, 1 minute input rate 0 packets/sec
```

因此，数据包将封装到使用源TEP IP 10.0.88.95（分配给loopback0）的VXLAN中，并发送到目标TEP IP 10.0.96.66。

确认源IP:

```
leaf1# show ip interface loopback 0 vrf overlay-1
```

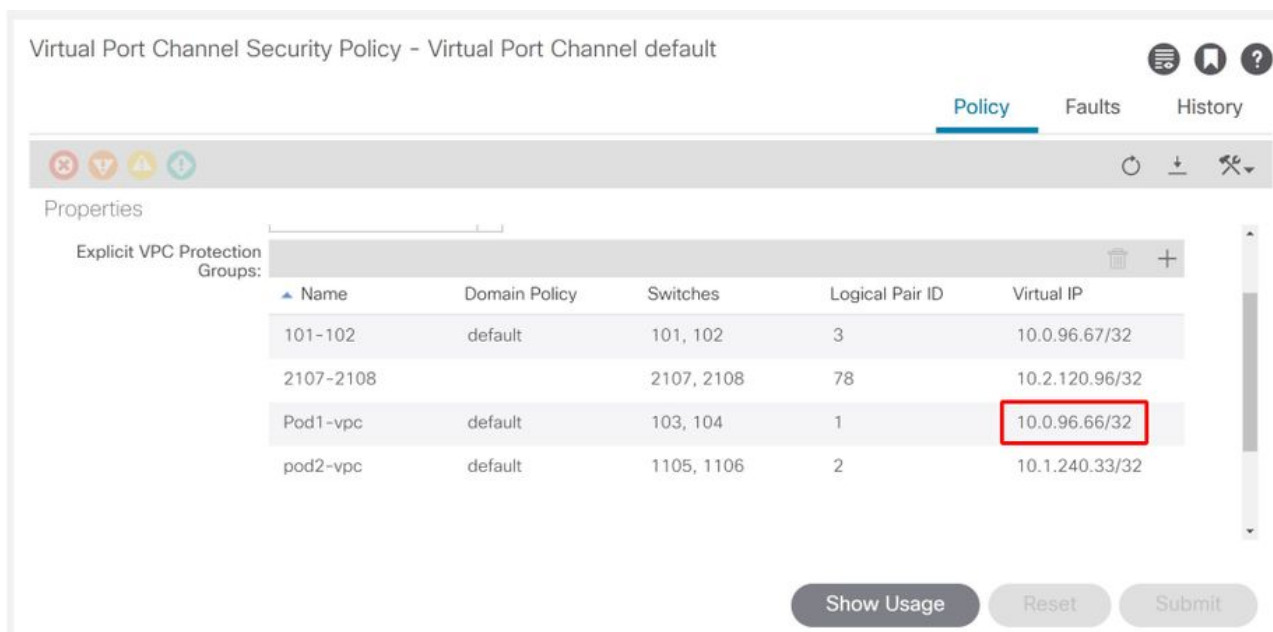
```
IP Interface Status for VRF "overlay-1"
lo0, Interface status: protocol-up/link-up/admin-up, iod: 4, mode: ptep
IP address: 10.0.88.95, IP subnet: 10.0.88.95/32
IP broadcast address: 255.255.255.255
```

IP primary address route-preference: 0, tag: 0

目的TEP IP 10.0.96.66可以是以下其中一项：

- 另一个枝叶的PTEP地址（可使用aciddiag fmvread检查）
- VPC VIP(可在“GUI”>“交换矩阵”>“访问策略”>“策略”>“交换机”>“虚拟端口通道默认值”中查看（请参阅下面的截图）
- 主干交换机上的环回IP。在主干交换机上使用“show ip interface vrf overlay-1”命令对此进行验证。

显式VPC保护组



入口枝叶交换机发送到主干交换机

入口枝叶现在将帧封装到VXLAN中，并将外部目标IP设置为10.0.96.66，这是前面的“show interface tunnel 4”命令中列出的隧道目标IP。它将使用网桥域的VNID(vxlan-16351141)将其封装到VXLAN中，如前面的“show endpoint mac 0000.1001.0102”命令输出所示。

根据VRF重叠-1中的IS-IS路由确定发送位置：

```
leaf1# show ip route 10.0.96.66 vrf overlay-1
IP Route Table for VRF "overlay-1"
'*' denotes best ucast next-hop
'***' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
```

```
10.0.96.66/32, ubest/mbest: 4/0
```

```
*via 10.0.88.65, Eth1/49.10, [115/3], 2w5d, isis-isis_infra, isis-l1-int
```

```
*via 10.0.88.94, Eth1/50.128, [115/3], 2w5d, isis-isis_infra, isis-l1-int
```

因此，存在使用eth1/49和1/50（主干交换机的交换矩阵上行链路）到目标的ECMP（等价多路径）路由。

主干转发

主干上的VRF重叠-1路由表显示，主机路由10.0.96.66可通过枝叶3或枝叶4到达。预期这是10.0.96.66是枝叶交换机103和104的VPC VIP:

```
spinel# show ip route 10.0.96.66 vrf overlay-1
IP Route Table for VRF "overlay-1"
'*' denotes best ucast next-hop
'***' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.0.96.66/32, ubest/mbest: 2/0
  *via 10.0.88.91, eth1/3.35, [115/2], 02w05d, isis-isis_infra, isis-l1-int
  *via 10.0.88.90, eth1/4.39, [115/2], 02w05d, isis-isis_infra, isis-l1-int
```

```
spinel# show lldp neighbors | egrep "1\3 |1\4 "
leaf3          Eth1/3          120          BR          Eth1/49
leaf4          Eth1/4          120          BR          Eth1/49
```

出口枝叶远程EP MAC学习

在这种情况下，目的TEP是VPC对，因此数据包将到达leaf3或leaf4。请参阅下面的命令输出。Leaf4应显示类似的输出。假设它们是同一VPC对的一部分，所有终端都在两个枝叶交换机之间同步。

出口枝叶上第2层流量的终端学习基于源MAC地址，该源MAC地址在与接收数据包中的VNID对应的BD中获取。这可以在终端表中进行验证。

源MAC地址位于VXLAN-S中的隧道2616351141。

隧道26进入TEP IP 10.0.88.95，即leaf1:

```
leaf3# show endpoint mac 0000.1001.0101
Legend:
s - arp          H - vtep          V - vpc-attached  p - peer-aged
R - peer-attached-rl B - bounce        S - static        M - span
D - bounce-to-proxy O - peer-attached a - local-aged    m - svc-mgr
L - local        E - shared-service

-----+-----+-----+-----+-----+
----+
      VLAN/          Encap          MAC Address          MAC Info/          Interface
      Domain          VLAN          IP Address          IP Info
-----+-----+-----+-----+-----+
----+
136/Prod:VRF1          vxlan-16351141    0000.1001.0101
tunnel26
```

```
leaf3# show interface tunnel 26
Tunnel26 is up
  MTU 9000 bytes, BW 0 Kbit
  Transport protocol is in VRF "overlay-1"
  Tunnel protocol/transport is ipvlan
  Tunnel source 10.0.88.91/32 (lo0)
  Tunnel destination 10.0.88.95
  Last clearing of "show interface" counters never
  Tx
  0 packets output, 1 minute output rate 0 packets/sec
  Rx
  0 packets input, 1 minute input rate 0 packets/sec
```

```
leaf3# acidiag fmvread | egrep "10.0.88.95"
      101      1      leaf1      FDO20160TPA      10.0.88.95/32      leaf
active 0
```

出口枝叶目标MAC查找

“show endpoint”命令确认目标MAC是在port-channel 1后面学习的，并且使用封装VLAN-2501

```
leaf3# show endpoint mac 0000.1001.0102
Legend:
s - arp          H - vtep          V - vpc-attached    p - peer-aged
R - peer-attached-rl B - bounce        S - static          M - span
D - bounce-to-proxy O - peer-attached  a - local-aged     m - svc-mgr
L - local        E - shared-service

+-----+-----+-----+-----+-----+
----+
      VLAN/          Encap          MAC Address          MAC Info/          Interface
      Domain          VLAN          IP Address          IP Info
+-----+-----+-----+-----+-----+
----+
135/Prod:VRF1          vlan-2501    0000.1001.0102 LpV
pol
```

这表示该帧将离开leaf3接口port-channel 1上的ACI交换矩阵，并封装VLAN ID 2501。您可以在GUI的Tenant Operational选项卡下找到BD VNID。

验证是否在主干交换机的COOP EP回拨中正确获知了两个终端

应在所有主干节点上同步COOP EP回程。可以使用BD VNID作为密钥并输入EP MAC地址来检查COOP EP回购。

此流量的源MAC地址是从隧道下一跳10.0.88.95获知的，该下一跳是leaf1的TEP IP。此外，命令输出显示与正确的网桥域对应的VNID 16351141。

```
spinel# show coop internal info repo ep key 16351141 00:00:10:01:01:01
```

```
Repo Hdr Checksum : 24197
Repo Hdr record timestamp : 10 01 2019 10:16:50 278195866
Repo Hdr last pub timestamp : 10 01 2019 10:16:50 283699467
Repo Hdr last dampen timestamp : 01 01 1970 00:00:00 0
Repo Hdr dampen penalty : 0
Repo Hdr flags : IN_OBJ EXPORT ACTIVE
EP bd vnid : 16351141
EP mac : 00:00:10:01:01:01
flags : 0x80
repo flags : 0x122
Vrf vnid : 2097154
Epg vnid : 0
EVPN Seq no : 0
Remote publish timestamp: 01 01 1970 00:00:00 0
Snapshot timestamp: 10 01 2019 10:16:50 278195866
Tunnel nh : 10.0.88.95
MAC Tunnel : 10.0.88.95
IPv4 Tunnel : 10.0.88.95
IPv6 Tunnel : 10.0.88.95
ETEP Tunnel : 0.0.0.0
```

根据leaf3和leaf4的VPC VIP 10.0.96.66获知此流量的目标MAC。同时列出了EP BD VNID

16351141，它与正确的BD相对应。

```
spinel# show coop internal info repo ep key 15302583 00:00:10:01:01:02
```

```
Repo Hdr Checksum : 16897
Repo Hdr record timestamp : 10 01 2019 11:05:46 351360334
Repo Hdr last pub timestamp : 10 01 2019 11:05:46 352019546
Repo Hdr last dampen timestamp : 01 01 1970 00:00:00 0
Repo Hdr dampen penalty : 0
Repo Hdr flags : IN_OBJ EXPORT ACTIVE
EP bd vnid : 16351141
  EP mac : 00:00:10:01:01:02
flags : 0x90
repo flags : 0x122
Vrf vnid : 2097154
Epg vnid : 0
EVPN Seq no : 0
Remote publish timestamp: 01 01 1970 00:00:00 0
Snapshot timestamp: 10 01 2019 11:05:46 351360334
Tunnel nh : 10.0.96.66
MAC Tunnel : 10.0.96.66
IPv4 Tunnel : 10.0.96.66
IPv6 Tunnel : 10.0.96.66
ETEP Tunnel : 0.0.0.0
```

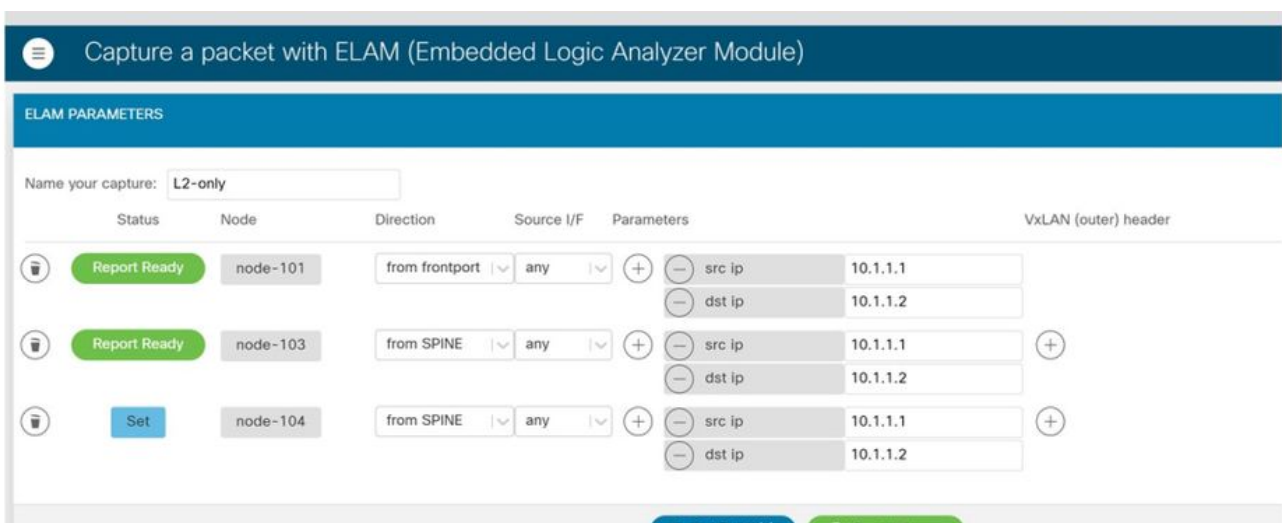
使用ELAM Assistant的ELAM输出

ELAM Assistant是一款功能强大的ACI应用，可以简化ACI交换矩阵上ELAM捕获的执行。

可以在多个枝叶节点上同时启动ELAM Assistant触发器。因此，可以在枝叶1、枝叶3和枝叶4中并行检查特定数据包。

配置的ELAM捕获将显示如下。如观察，数据包在leaf1（节点-101）和leaf3（节点-103）上可见。

ELAM Assistant — 参数



leaf1(node-101)的报告显示以下内容：

- Captured Packet Information输出确认数据包进入eth1/3且具有正确的MAC和IP信息。
- 数据包转发信息显示它已在eth1/49上转发到TEP IP 10.0.96.66。

ELAM Assistant — leaf1(node-101) — 捕获的数据包信息

Basic Information	
Device Type	LEAF
Packet Direction	ingress (front panel port -> leaf)
Incoming I/F	eth1/3

L2 Header	
Destination MAC	0000.1001.0102
Source MAC	0000.1001.0101
Access Encap VLAN	2501
CoS	0

L3 Header	
L3 Type	IPv4
Destination IP	10.1.1.2
Source IP	10.1.1.1
IP Protocol	0x1 (ICMP)
DSCP	0
TTL	255

No Vx

ELAM Assistant — leaf1(node-101) — 数据包转发信息

Packet Forwarding Information	
Forward Result	
Destination Type	To another ACI node (or AVS/AVE)
Destination TEP	10.0.96.66 (vPC (103_104))
Destination Physical Port	eth1/49
Sent to SUP/CPU instead	no
SUP Redirect Reason (SUP code)	NONE
Contract	
Destination EPG pcTag (dclass)	32770 (Prod:App:EPG1)
Source EPG pcTag (sclass)	32770 (Prod:App:EPG1)
Contract was applied	1 (Contract was applied on this node)
Drop	

在出口枝叶上的枝叶3 (节点-103) 上，观察到以下内容：

在leaf3的捕获数据包信息中，它从eth1/49进入。外部IP地址确认以下内容：

- 源TEP:10.0.88.95
- 目标TEP:10.0.96.66
- VNID:16351141(BD VNID)

ELAM Assistant — leaf3(node-103) — 捕获的数据包信息

Captured Packet Information	
Basic Information	
Device Type	LEAF
Packet Direction	egress (spine LC -> leaf)
Incoming I/F	eth1/49

L3 Header (Outer VxLAN)	
L3 Type	IPv4
Destination IP	10.0.96.66 (vPC (103_104))
Source IP	10.0.88.95 (bdsol-aci32-leaf1)
IP Protocol	0x11 (UDP)
DSCP	0
TTL	31
Don't Fragment Bit	0x0 (0x0)

L4 Header (Outer VxLAN)	
L4 Type	ivxLAN
DL (Don't Learn) Bit	0 (not set)
Src Policy Applied Bit	1 (Contract was applied on the previous node)
Dst Policy Applied Bit	1 (Contract was applied on the previous node)
Source EPG (sclass / src pcTag)	0x8002 / 32770 (Prod:App:EPG1)
VRF/BD VNID	15302583 (Prod:BD1)

数据包转发信息显示流量在端口通道1上转发，特别是以太网1/12。

Packet Forwarding Information

Forward Result	
Destination Type	To a local port
Destination Logical Port	Po1
Destination Physical Port	eth1/12
Sent to SUP/CPU instead	no
SUP Redirect Reason (SUP code)	NONE

Contract	
Destination EPG pcTag (dclass)	32770 (Prod:App:EPG1)
Source EPG pcTag (sclass)	32770 (Prod:App:EPG1)
Contract was applied	1 (Contract was applied on this node)

Drop	
Drop Code	no drop

使用CLI的入口枝叶ELAM

建议使用ELAM Assistant，因为它简化了运行ELAM捕获的操作。但是，也可以在ACI交换机上使用CLI命令生成ELAM报告。下面是如何执行此操作的示例。

使用所示的触发序列捕获入口枝叶上的数据包。有关ELAM选项的详细信息，请参阅“工具”部分。

- 在本示例中，ASIC是“tah”**作为**枝叶（部件号结尾“—EX”）。
- “in-select 6”用于捕获来自没有VXLAN封装的下行链路端口的数据包。
- “out-select 1”确保丢弃矢量也显示（在丢包的情况下）。
- 需要“reset”命令以确保之前的所有触发器都已清除。
- 即使这是桥接流，ELAM仍能查看IP报头。因此，“ipv4 src_ip”和“dst_ip”可用于设置触发器。

```
module-1# debug platform internal tah elam asic 0
module-1(DBG-elam)# trigger init in-select ?
10 Outer14-inner14-ieth
13 Outer(12|13|14)-inner(12|13|14)-noieth
14 Outer(12(vntag)|13|14)-inner(12|13|14)-ieth
15 Outer(12|13|14)-inner(12|13|14)-ieth
6 Outer12-outer13-outer14
7 Inner12-inner13-inner14
8 Outer12-inner12-ieth
9 Outer13-inner13

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 10.1.1.1 dst_ip 10.1.1.2
module-1(DBG-elam-insel6)# start
```

要查看是否收到数据包，请检查ELAM状态。如果存在触发器，则意味着捕获到匹配条件的数据包

o

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

```
ELAM STATUS
```

```
=====
```

```
Asic 0 Slice 0 Status Triggered
```

```
Asic 0 Slice 1 Status Armed
```

下一个输出显示使用“ereport”命令显示报表。输出非常长，因此此处仅粘贴开头。但请注意，完整的报告会保存在枝叶文件系统中的某个位置供以后分析使用。文件名还包含获取ELAM的时间戳。

```
leaf1# ls -al /var/log/dme/log/elam_2019-09-30-03m-23h-14s.txt
```

```
-rw-rw-rw- 1 root root 699106 Sep 30 23:03 /var/log/dme/log/elam_2019-09-30-03m-23h-14s.txt
```

“ereport”用于验证数据包是否已被接收，以及信息是否如预期一样（源和目标MAC、源和目标IP等）

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

```
Python available. Continue ELAM decode with LC Pkg
```

```
ELAM REPORT
```

```
=====
```

```
Trigger/Basic Information
```

```
=====
```

```
ELAM Report File           : /tmp/logs/elam_2019-09-30-03m-23h-14s.txt
In-Select Trigger          : Outerl2-outerl3-outerl4( 6 )
Out-Select Trigger         : Pktrw-sideband-drpvec( 1 )
ELAM Captured Device       : LEAF
Packet Direction           : ingress
Triggered ASIC type        : Sugarbowl
Triggered ASIC instance    : 0
Triggered Slice            : 0
Incoming Interface         : 0x24( 0x24 )
( Slice Source ID(Ss) in "show plat int hal l2 port gpd" )
```

```
=====
```

```
Captured Packet
```

```
-----
```

```
Outer Packet Attributes
```

```
-----
```

```
Outer Packet Attributes    : l2uc ipv4 ip ipuc ipv4uc
Opcode                     : OPCODE_UC
```

```
-----
```

```
Outer L2 Header
```

```
-----
```

```
Destination MAC           : 0000.1001.0102
Source MAC                 : 0000.1001.0101
802.1Q tag is valid       : yes( 0x1 )
CoS                        : 0( 0x0 )
Access Encap VLAN         : 2501( 0x9C5 )
```

```
-----
```

```
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     : not set
TTL                    : 255
IP Protocol Number     : ICMP
IP CheckSum             : 51097( 0xC799 )
Destination IP         : 10.1.1.2
Source IP              : 10.1.1.1

```

```

=====
=====
                                         Forwarding Lookup ( FPB )
=====
=====
-----
-----

```

```

Destination MAC (Lookup Key)
-----
-----

```

```

Dst MAC Lookup was performed      : yes
Dst MAC Lookup BD                 : 522( 0x20A )
( Hw BDID in "show plat int hal l2 bd pi" )
Dst MAC Address                   : 0000.1001.0102

```

```

-----
-----
Destination MAC (Lookup Result)
-----
-----

```

```

Dst MAC is Hit                   : yes
Dst MAC is Hit Index             : 6443( 0x192B )
( phy_id in "show plat int hal objects ep l2 mac (MAC) extensions" )
or ( HIT IDX in "show plat int hal l3 nexthops" for L3OUT/L3 EP)
.....

```

使用fTriage跟踪流

分类从APIC CLI运行，可用于通过ACI交换矩阵执行完整路径。至少指定入口枝叶（节点-101）、源IP和目标IP。在此特定情况下，它是桥接（第2层）流，因此使用fTriage网桥选项。

请注意，fTriage在当前目录中生成日志文件。此日志文件将包含收集的所有日志和ELAM报告。这样便可以在每一跳捕获数据包。输出的简短版本如下：

```

apic1# ftrriage bridge -ii LEAF:101 -sip 10.1.1.1 -dip 10.1.1.2
fTriage Status: {"dbgFtrriage": {"attributes": {"operState": "InProgress", "pid": "12181",
"apicId": "1", "id": "0"}}}
Starting ftrriage
Log file name for the current run is: ftlog_2019-10-01-18-53-24-125.txt
2019-10-01 18:53:24,129 INFO      /controller/bin/ftrriage bridge -ii LEAF:101 -sip 10.1.1.1 -dip
10.1.1.2
2019-10-01 18:53:49,280 INFO      ftrriage:      main:1165 Invoking ftrriage with default password
and default username: apic#fallback\\admin
2019-10-01 18:54:10,204 INFO      ftrriage:      main:839 L2 frame Seen on leaf1 Ingress: Eth1/3
Egress: Eth1/49 Vnid: 15302583
2019-10-01 18:54:10,422 INFO      ftrriage:      main:242 ingress encaps string vlan-2501
2019-10-01 18:54:10,427 INFO      ftrriage:      main:271 Building ingress BD(s), Ctx
2019-10-01 18:54:12,288 INFO      ftrriage:      main:294 Ingress BD(s) Prod:BD1

```

```

2019-10-01 18:54:12,288 INFO      ftriage:      main:301  Ingress Ctx: Prod:VRF1
2019-10-01 18:54:12,397 INFO      ftriage:      pktrec:490 leaf1: Collecting transient losses
snapshot for LC module: 1
2019-10-01 18:54:30,079 INFO      ftriage:      main:933  SMAC 00:00:10:01:01:01 DMAC
00:00:10:01:01:02
2019-10-01 18:54:30,080 INFO      ftriage:      unicast:973 leaf1: <- is ingress node
2019-10-01 18:54:30,320 INFO      ftriage:      unicast:1215 leaf1: Dst EP is remote
2019-10-01 18:54:31,155 INFO      ftriage:      misc:659  leaf1: L2 frame getting bridged in SUG
2019-10-01 18:54:31,380 INFO      ftriage:      misc:657  leaf1: Dst MAC is present in SUG L2 tbl
2019-10-01 18:54:31,826 INFO      ftriage:      misc:657  leaf1: RwdMAC DIPO(10.0.96.66) is one of
dst TEPs ['10.0.96.66']
2019-10-01 18:56:16,249 INFO      ftriage:      main:622  Found peer-node spine1 and IF: Eth1/1 in
candidate list
2019-10-01 18:56:21,346 INFO      ftriage:      node:643  spine1: Extracted Internal-port GPD Info
for lc: 1
2019-10-01 18:56:21,348 INFO      ftriage:      fcls:4414 spine1: LC trigger ELAM with IFS: Eth1/1
Asic :0 Slice: 0 Srcid: 32
2019-10-01 18:56:54,424 INFO      ftriage:      main:839  L2 frame Seen on spine1 Ingress: Eth1/1
Egress: LC-1/0 FC-24/0 Port-0 Vnid: 15302583
2019-10-01 18:56:54,424 INFO      ftriage:      pktrec:490 spine1: Collecting transient losses
snapshot for LC module: 1
2019-10-01 18:57:15,093 INFO      ftriage:      fib:332  spine1: Transit in spine
2019-10-01 18:57:21,394 INFO      ftriage:      unicast:1252 spine1: Enter dbg_sub_nexthop with
Transit inst: ig infra: False glbs.dipo: 10.0.96.66
2019-10-01 18:57:21,508 INFO      ftriage:      unicast:1417 spine1: EP is known in COOP (DIPO =
10.0.96.66)
2019-10-01 18:57:25,537 INFO      ftriage:      unicast:1458 spine1: Infra route 10.0.96.66 present
in RIB
2019-10-01 18:57:25,537 INFO      ftriage:      node:1331 spine1: Mapped LC interface: LC-1/0 FC-
24/0 Port-0 to FC interface: FC-24/0 LC-1/0 Port-0
2019-10-01 18:57:30,616 INFO      ftriage:      node:460  spine1: Extracted GPD Info for fc: 24
2019-10-01 18:57:30,617 INFO      ftriage:      fcls:5748 spine1: FC trigger ELAM with IFS: FC-
24/0 LC-1/0 Port-0 Asic :0 Slice: 2 Srcid: 0
2019-10-01 18:57:49,611 INFO      ftriage:      unicast:1774 L2 frame Seen on FC of node: spine1
with Ingress: FC-24/0 LC-1/0 Port-0 Egress: FC-24/0 LC-1/0 Port-0 Vnid: 15302583
2019-10-01 18:57:49,611 INFO      ftriage:      pktrec:487 spine1: Collecting transient losses
snapshot for FC module: 24
2019-10-01 18:57:53,110 INFO      ftriage:      node:1339 spine1: Mapped FC interface: FC-24/0 LC-
1/0 Port-0 to LC interface: LC-1/0 FC-24/0 Port-0
2019-10-01 18:57:53,111 INFO      ftriage:      unicast:1474 spine1: Capturing Spine Transit pkt-
type L2 frame on egress LC on Node: spine1 IFS: LC-1/0 FC-24/0 Port-0
2019-10-01 18:57:53,530 INFO      ftriage:      fcls:4414 spine1: LC trigger ELAM with IFS: LC-1/0
FC-24/0 Port-0 Asic :0 Slice: 0 Srcid: 64
2019-10-01 18:58:26,497 INFO      ftriage:      unicast:1510 spine1: L2 frame Spine egress Transit
pkt Seen on spine1 Ingress: LC-1/0 FC-24/0 Port-0 Egress: Eth1/3 Vnid: 15302583
2019-10-01 18:58:26,498 INFO      ftriage:      pktrec:490 spine1: Collecting transient losses
snapshot for LC module: 1
2019-10-01 18:59:28,634 INFO      ftriage:      main:622  Found peer-node leaf3 and IF: Eth1/49 in
candidate list
2019-10-01 18:59:39,235 INFO      ftriage:      main:839  L2 frame Seen on leaf3 Ingress: Eth1/49
Egress: Eth1/12 (Po1) Vnid: 11364
2019-10-01 18:59:39,350 INFO      ftriage:      pktrec:490 leaf3: Collecting transient losses
snapshot for LC module: 1
2019-10-01 18:59:54,373 INFO      ftriage:      main:522  Computed egress encaps string vlan-2501
2019-10-01 18:59:54,379 INFO      ftriage:      main:313  Building egress BD(s), Ctx
2019-10-01 18:59:57,152 INFO      ftriage:      main:331  Egress Ctx Prod:VRF1
2019-10-01 18:59:57,153 INFO      ftriage:      main:332  Egress BD(s): Prod:BD1
2019-10-01 18:59:59,230 INFO      ftriage:      unicast:1252 leaf3: Enter dbg_sub_nexthop with Local
inst: eg infra: False glbs.dipo: 10.0.96.66
2019-10-01 18:59:59,231 INFO      ftriage:      unicast:1257 leaf3: dbg_sub_nexthop invokes
dbg_sub_eg for vip
2019-10-01 18:59:59,231 INFO      ftriage:      unicast:1784 leaf3: <- is egress node
2019-10-01 18:59:59,377 INFO      ftriage:      unicast:1833 leaf3: Dst EP is local
2019-10-01 18:59:59,378 INFO      ftriage:      misc:657  leaf3: EP if(Po1) same as egr if(Po1)

```

```

2019-10-01 18:59:59,378 INFO      ftrriage:      misc:659  leaf3: L2 frame getting bridged in SUG
2019-10-01 18:59:59,613 INFO      ftrriage:      misc:657  leaf3: Dst MAC is present in SUG L2 tbl
2019-10-01 19:00:06,122 INFO      ftrriage:      main:961  Packet is Exiting fabric with peer-
device: n3k-3 and peer-port: Ethernet1/16

```

未知第2层单播流量的故障排除工作流程 — BD处于泛洪模式

在本示例中，目的MAC未知。入口枝叶上的目标MAC查找未显示任何输出。

```

leaf1# show endpoint mac 0000.1001.0102
Legend:
s - arp          H - vtep          V - vpc-attached  p - peer-aged
R - peer-attached-rl B - bounce       S - static        M - span
D - bounce-to-proxy O - peer-attached a - local-aged    m - svc-mgr
L - local        E - shared-service

-----+-----+-----+-----+-----+
---+
      VLAN/                               Encap          MAC Address      MAC Info/        Interface
      Domain                             VLAN            IP Address       IP Info
-----+-----+-----+-----+-----+
---+

```

假设L2未知单播的BD设置为“泛洪”，下面是高层次的情况：

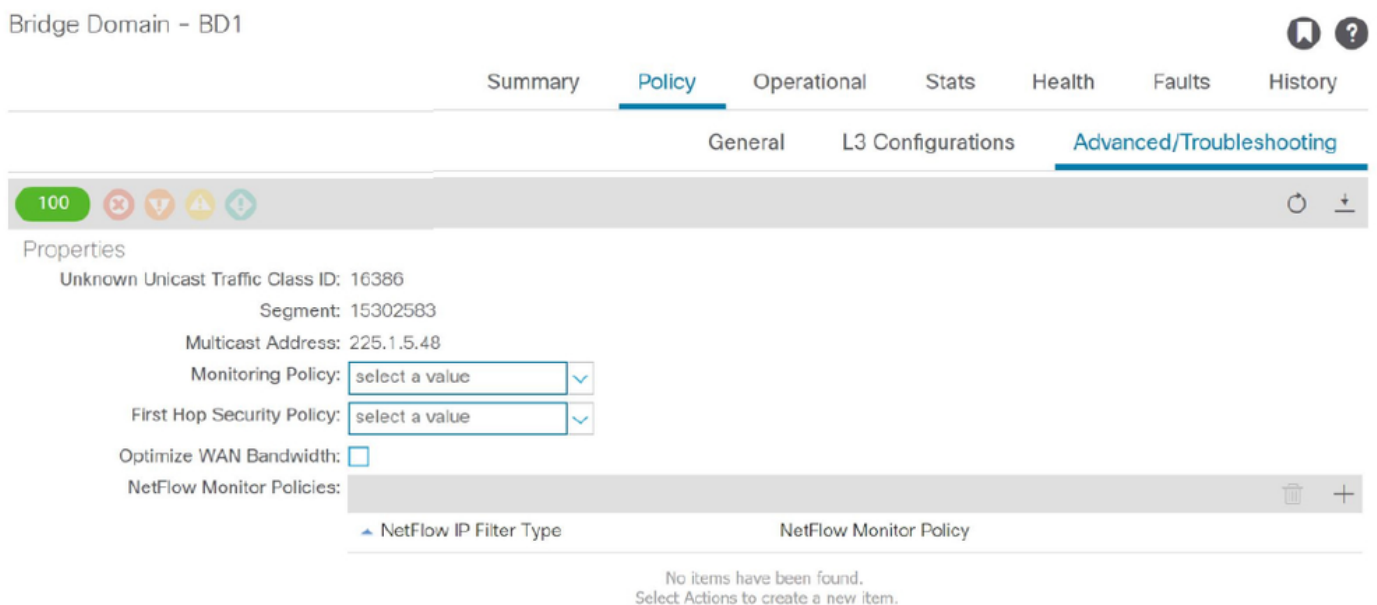
1. 入口枝叶将对数据包报头进行散列以将其分配到一个FTAG (从0到15)。
2. 入口枝叶将使用BD VNID将帧封装在VXLAN数据包中。外部目标IP将是BD GIPo + FTAG。
3. 它将按照树拓扑在交换矩阵中泛洪，并应到达已部署BD的每个枝叶节点。

此部分将突出显示可以检查的内容。

查找BD GIPo

GUI标识BD用于多目标流量的组播组225.1.5.48。

BD GIPo



使用ELAM Assistant，检查入口枝叶上的ELAM报告。这显示帧在BD中泛洪，并在所有交换矩阵上行链路（此处eth1/49、1/50、1/51和1/52）上退出。

ELAM助理 — 入口枝叶 — 数据包转发信息

Packet Forwarding Information	
Forward Result	
Destination Type	Flood in BD
Destination Ports	eth1/51, eth1/50, eth1/52, eth1/49 (overlay (Fabric uplink))
vPC Designated Forwarder (DF)	yes
Sent to SUP/CPU as well	no
SUP Redirect Reason (SUP code)	NONE
Contract	
Destination EPG pcTag (dclass)	16386 (null)
Source EPG pcTag (sclass)	32770 (null)
Contract was applied	0 (Contract was not applied on this node)
Drop	
Drop Code	no drop

要查找入口枝叶选择的FTAG值，请转到ELAM Assistant的原始报告。

```
sug_lu2ba_sb_info.mc_info.mc_info_nopad.ftag: 0xC
```

将十六进制值0xC转换为十进制时，结果为FTAG 12。

绘制FTAG拓扑

FTAG拓扑由IS-IS计算。为每个FTAG值创建一个树拓扑，带有允许最佳负载分布拓扑的根接口和输出接口列表。

使用以下命令显示本地FTAG拓扑。在下面的示例中，我们在spine1上使用FTAG ID 12拓扑。

```
spine1# show isis internal mcast routes ftag
IS-IS process: isis_infra
  VRF : default
FTAG Routes
=====
```



```
FTAG ID: 12 [Enabled] Cost:( 2/ 11/ 0)
```

```
-----  
Root port: Ethernet1/4.39  
OIF List:  
  Ethernet1/11.11  
  Ethernet1/12.12
```

在大型ACI交换矩阵中绘制完整的FTAG拓扑是一项漫长而复杂的任务。“aci-ftag-viewer” Python脚本(<https://github.com/agccie/aci-ftag-viewer>)可以复制到APIC上。它一次生成交换矩阵的完整FTAG拓扑。

以下输出显示了多Pod交换矩阵的Pod1中的FTAG 12树，并包括IPN设备上的FTAG拓扑。

这表明，如果流量从枝叶101进入ACI交换矩阵，它将经过以下脚本输出中列出的以下路径。

```
admin@apic1:tmp> python aci_ftag_viewer.py --ftag 12 --pod 1  
#####  
# Pod 1 FTAG 12  
# Root spine-204  
# active nodes: 8, inactive nodes: 1  
#####  
spine-204  
+- 1/1 ----- 1/52 leaf-101  
+- 1/2 ----- 1/52 leaf-102  
+- 1/3 ----- 1/52 leaf-103  
+- 1/4 ----- 1/52 leaf-104  
      +- 1/49 ----- 1/4 spine-201  
      |                                     +- 1/11 ..... (EXT) Eth2/13 n7706-01-Multipod-A1  
      |                                     +- 1/12 ..... (EXT) Eth2/9 n7706-01-Multipod-A2  
      |  
      +- 1/50 ----- 1/4 spine-202  
      |                                     +- 1/11 ..... (EXT) Eth2/14 n7706-01-Multipod-A1  
      |                                     +- 1/12 ..... (EXT) Eth2/10 n7706-01-Multipod-A2  
      |  
      +- 1/51 ----- 2/4 spine-203  
      |                                     +- 2/11 ..... (EXT) Eth2/15 n7706-01-Multipod-A1  
      |                                     +- 2/12 ..... (EXT) Eth2/11 n7706-01-Multipod-A2  
+- 1/11 ..... (EXT) Eth2/16 n7706-01-Multipod-A1  
+- 1/12 ..... (EXT) Eth2/12 n7706-01-Multipod-A2
```

ELAM — 出口枝叶 — 泛洪流量

在这种情况下，泛洪流量会到达ACI交换矩阵中的每个枝叶。因此，它将到达leaf3和leaf4，它们都是VPC对。这两个枝叶节点都具有通往目的地的VPC。为避免重复的数据包，VPC对仅选择一个枝叶将泛洪流量转发到目标。选择的枝叶称为VPC DF枝叶（VPC指定的转发器枝叶）。

这可以在两个枝叶节点上使用以下触发器在ELAM中检查。

```
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)# set inner ipv4 src_ip 10.1.1.1 dst_ip 10.1.1.2  
module-1(DBG-elam-insel14)# start
```

leaf3输出：

```
module-1(DBG-elam-insel14)# ereport | egrep vpc.*df
```

```
sug_lub_latch_results_vec.lub4_1.vpc_df: 0x1
```

leaf4输出：

```
module-1(DBG-elam-insell14)# ereport | egrep vpc.*df
sug_lub_latch_results_vec.lub4_1.vpc_df: 0x0
```

在上述输出中，leaf3为“vpc_df”字段设置了值“0x1”，而leaf4为“vpc_df”字段设置了“0x0”。因此，指定的转发器将是leaf3。leaf3会将其VPC链路上的泛洪数据包转发到目标EP。

未知第2层单播流量的故障排除工作流程 — 硬件代理中的BD

列出的当前场景是BD处于硬件代理模式的第2层未知单播流量。在这种情况下，如果入口枝叶不知道目的MAC地址，则会将数据包转发到主干任播代理MAC地址。主干将对目标MAC执行COOP查找。

如果查找成功（如下所示），主干会将外部目标IP重写为隧道目标（此处10.0.96.66），并将其发送到leaf3-leaf4 VPC对。

```
spinel# show coop internal info repo ep key 15302583 00:00:10:01:01:02
```

```
Repo Hdr Checksum : 16897
Repo Hdr record timestamp : 10 01 2019 11:05:46 351360334
Repo Hdr last pub timestamp : 10 01 2019 11:05:46 352019546
Repo Hdr last dampen timestamp : 01 01 1970 00:00:00 0
Repo Hdr dampen penalty : 0
Repo Hdr flags : IN_OBJ EXPORT ACTIVE
EP bd vnid : 16351141
  EP mac : 00:00:10:01:01:02
  flags : 0x90
  repo flags : 0x122
  Vrf vnid : 2097154
  Epg vnid : 0
  EVPN Seq no : 0
Remote publish timestamp: 01 01 1970 00:00:00 0
Snapshot timestamp: 10 01 2019 11:05:46 351360334
Tunnel nh : 10.0.96.66
MAC Tunnel : 10.0.96.66
IPv4 Tunnel : 10.0.96.66
IPv6 Tunnel : 10.0.96.66
ETEP Tunnel : 0.0.0.0
```

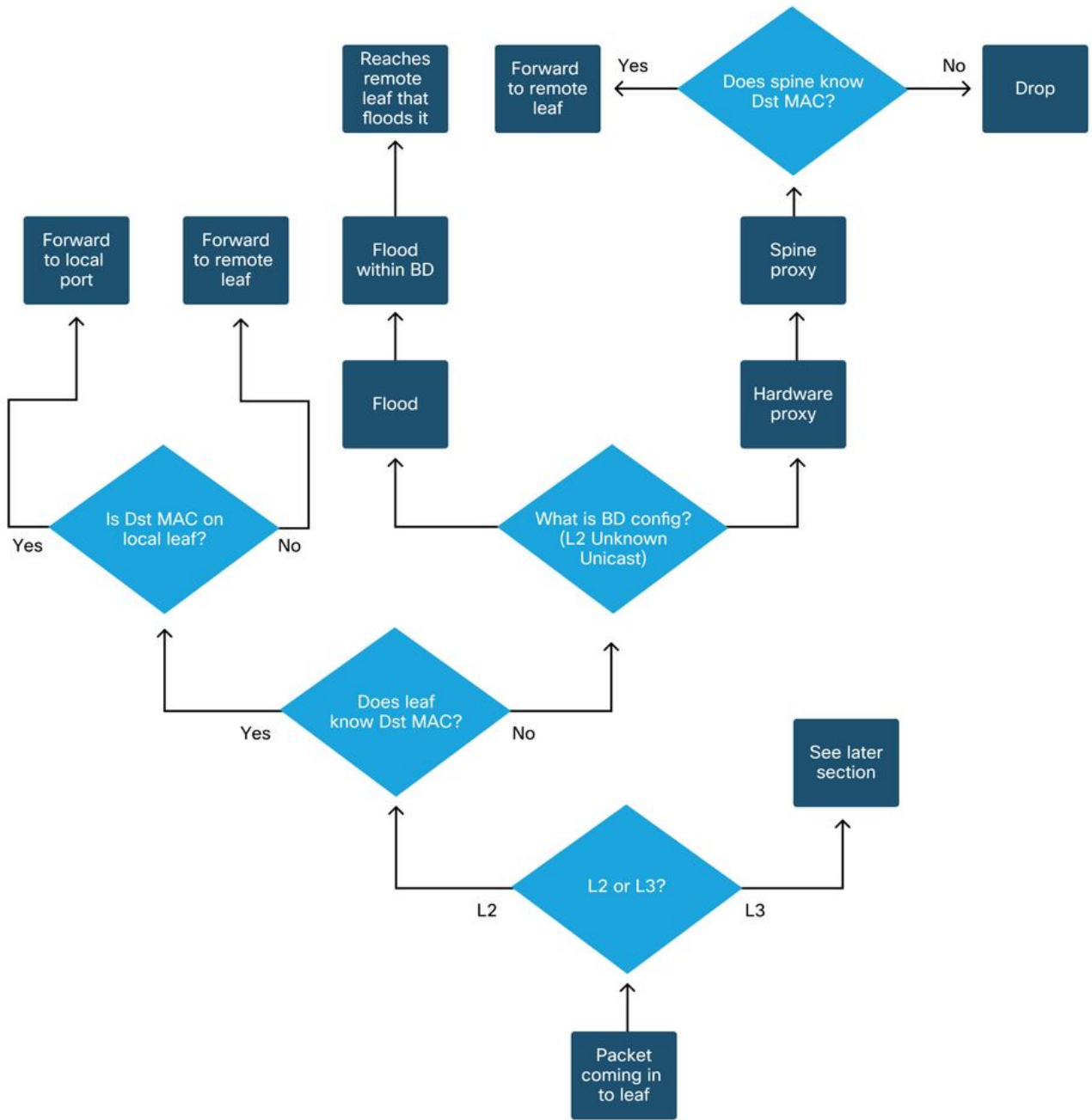
如果查找失败（终端在ACI交换矩阵中未知），主干将丢弃未知单播。

```
spinel# show coop internal info repo ep key 15302583 00:00:10:01:01:02
Key not found in repo
```

第2层转发摘要

下图总结了ACI交换矩阵中第2层流量可能的转发行为。

ACI交换矩阵第2层转发行为



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）。