

# ACI交換矩陣內轉發故障排除 — L3轉發：不同BD中的兩個端點

## 目錄

[簡介](#)

[背景資訊](#)

[L3轉發：不同BD中的兩個端點](#)

[高級故障排除工作流程](#)

[第一次檢查 — 驗證程式設計](#)

[第二次檢查 — 在枝葉節點上通過CLI驗證學習條目和表條目](#)

[第三種檢查 — 獲取資料包並分析轉發決策](#)

[已知端點的故障排除工作流](#)

[檢查BD的普及網關](#)

[檢查枝葉上的路由表](#)

[預設網關IP的ARP解析](#)

[入口枝葉源IP和MAC終端學習](#)

[入口枝葉目標IP查詢 — 已知遠端終結點](#)

[出口枝葉上的源IP學習](#)

[出口枝葉上的目標IP查詢](#)

[按資料路徑進行分類](#)

[未知終結點的疑難解答工作流](#)

[入口枝葉目標IP查詢](#)

[主幹上的COOP查詢 — 目標IP已知](#)

[主幹上的COOP查詢 — 目標IP未知](#)

[ACI轉發摘要](#)

## 簡介

本文檔介紹瞭解ACI L3轉發方案並對其進行故障排除的步驟。

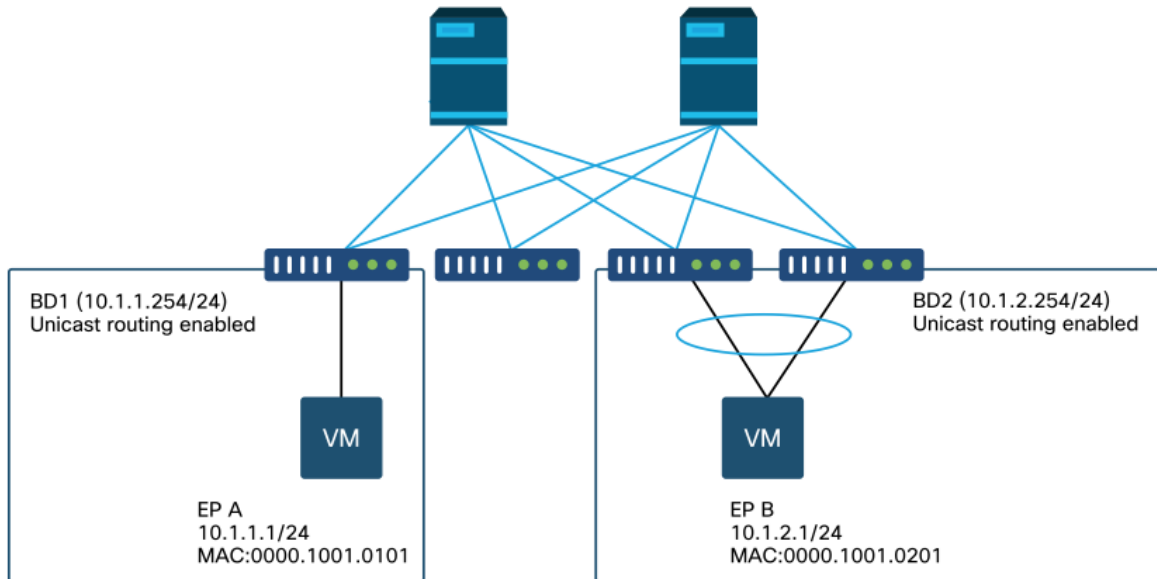
## 背景資訊

本文中的資料摘自 [思科以應用為中心的基礎設施第二版故障排除](#) 書，特別是 [交換矩陣內轉發 — L3轉發：不同BD中的兩個端點](#) 章節。

## L3轉發：不同BD中的兩個端點

本章介紹不同網橋域中的終端無法相互通訊的故障排除示例。這將是ACI交換矩陣路由的流。圖1說明了拓撲。

不同網橋域中的端點



## 高級故障排除工作流程

以下是典型的故障排除步驟和驗證命令：

### 第一次檢查 — 驗證程式設計

- 應該將BD沉浸式網關推送到枝葉節點。
- 到達目標BD子網的路由應推送到葉節點。
- 應解析主機的預設網關的ARP。

### 第二次檢查 — 在枝葉節點上通過CLI驗證學習條目和表條目

- 檢查源枝葉節點和目標枝葉節點瞭解端點以及它是否瞭解目標端點：終結點表 — 「顯示終結點」。TEP destination - 'show interface tunnel <x>'。在「show ip route <TEP address> vrf overlay-1」命令中查詢TEP目標。
- 檢查脊柱節點瞭解端點：'顯示合作公寓內部資訊'。

### 第三種檢查 — 獲取資料包並分析轉發決策

- 使用ELAM ( ELAM Assistant或CLI ) 驗證幀是否存在。
- 或使用fTriage跟蹤流量。

## 已知端點的故障排除 workflow

### 檢查BD的普及網關

在本示例中，將使用以下源端點和目標端點：

- leaf1下的EP A 10.1.1.1。

- VPC對leaf3和leaf4下的EP B 10.1.2.1。

應看到以下沉浸式網關：

- 10.1.1.254/24，用於leaf1上的BD1網關。
- 10.1.2.254/24，用於leaf3和leaf4上的BD2網關。

可以使用以下命令檢查此問題：枝葉節點上的「show ip interface vrf <vrf name>」。

leaf1:

```
leaf1# show ip interface vrf Prod:VRF1
IP Interface Status for VRF "Prod:VRF1"
vlan7, Interface status: protocol-up/link-up/admin-up, iod: 106, mode: pervasive
  IP address: 10.1.1.254, IP subnet: 10.1.1.0/24
  IP broadcast address: 255.255.255.255
  IP primary address route-preference: 0, tag: 0
```

leaf3和4:

```
leaf3# show ip interface vrf Prod:VRF1
IP Interface Status for VRF "Prod:VRF1"
vlan1, Interface status: protocol-up/link-up/admin-up, iod: 159, mode: pervasive
  IP address: 10.1.2.254, IP subnet: 10.1.2.0/24
  IP broadcast address: 255.255.255.255
  IP primary address route-preference: 0, tag: 0
```

```
leaf4# show ip interface vrf Prod:VRF1
IP Interface Status for VRF "Prod:VRF1"
vlan132, Interface status: protocol-up/link-up/admin-up, iod: 159, mode: pervasive
  IP address: 10.1.2.254, IP subnet: 10.1.2.0/24
  IP broadcast address: 255.255.255.255
  IP primary address route-preference: 0, tag: 0
```

請注意，leaf3和leaf4具有相同的通用網關地址，但可能會看到SVI的不同VLAN封裝。

- leaf3使用VLAN 1。
- leaf4使用VLAN 132。

這是因為VLAN 1或VLAN 132是枝葉上的本地VLAN。

如果未將沉浸式網關IP地址推送到枝葉，請在APIC GUI中驗證不存在阻止部署VLAN的故障。

## 檢查枝葉上的路由表

Leaf1在子網10.1.2.0/24中沒有任何端點，但是它必須具有到該子網的路由才能到達該子網：

```
leaf1# show ip route 10.1.2.0/24 vrf Prod:VRF1
IP Route Table for VRF "Prod:VRF1"
 '*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

10.1.2.0/24, ubest/mbest: 1/0, attached, direct, pervasive
  *via 10.0.8.65%overlay-1, [1/0], 00:22:37, static, tag 4294967294
    recursive next hop: 10.0.8.65/32%overlay-1
```

請注意，標籤為「沉浸式」和「直接」的路由的下一跳為10.0.8.65。這是所有主幹上存在的任播

v4環回地址。

```
leaf1# show isis dteps vrf overlay-1 | egrep 10.0.8.65
10.0.8.65          SPINE    N/A          PHYSICAL,PROXY-ACAST-V4
```

同樣，leaf3和leaf4應具有10.1.1.0/24的路由。

```
leaf3# show ip route 10.1.1.1 vrf Prod:VRF1
IP Route Table for VRF "Prod:VRF1"
'*' denotes best ucast next-hop
 '**' denotes best mcast next-hop
 '[x/y]' denotes [preference/metric]
 '%<string>' in via output denotes VRF <string>

10.1.1.0/24, ubest/mbest: 1/0, attached, direct, pervasive
  *via 10.0.8.65%overlay-1, [1/0], 00:30:25, static, tag 4294967294
    recursive next hop: 10.0.8.65/32%overlay-1
```

如果缺少這些路由，很可能因為BD1中的EPG和BD2中的EPG之間沒有合約。如果BD1中的枝葉下沒有本地終結點，則BD1普及網關不會推送到枝葉。如果EPG中有本地終端與BD1中的另一個EPG有合約，則會在枝葉上獲取BD1子網。

## 預設網關IP的ARP解析

由於本地終端所在的枝葉應該具有沉浸式網關，因此對沉浸式網關的ARP請求應始終由本地枝葉解析。可以使用以下命令在本地枝葉上檢查此問題：

```
leaf1# show ip arp internal event-history event | egrep 10.1.1.1
[116] TID 26571:arp_handle_arp_request:6135: log_collect_arp_pkt; sip = 10.1.1.1; dip =
10.1.1.254;interface = Vlan7; phy_inteface = Ethernet1/3; flood = 0; Info = Sent ARP response.
[116] TID 26571:arp_process_receive_packet_msg:8384: log_collect_arp_pkt; sip = 10.1.1.1; dip
= 10.1.1.254;interface = Vlan7; phy_interface = Ethernet1/3;Info = Received arp request
```

## 入口枝葉源IP和MAC終端學習

對於第3層轉發，ACI將執行第3層源IP學習和目標IP查詢。學習的IP地址範圍確定為VRF。

您可以在EPG的「操作」頁籤中的GUI上檢查此項。請注意，此處會同時得知IP和MAC。

## EPG運行端點



End Point	MAC	IP	Learning Source	Hosting Server	Reportin: Controllr Name	Interface	Multicast Encap Address
EP-00:00:10:01:01:01	00:00:10:01:01:01	10.1.1.1	learned	---	---	Pod-1/Node-101/eth1/0 (learned)	vlan-2501
EP-00:00:10:01:01:02	00:00:10:01:01:02	10.1.1.2...	learned	---	---	Pod-1/Node-103-104/N3k-3-VRPC3-4 (learned)	vlan-2501

## EPG運營端點 — 詳細資訊

End Point	MAC	IP	Learning Source	Hosting Server	Reporting Controller Name	Interface	Multicast Address	Encap
EP-00:00:10:01:01:01	00:00:10:01:01:01	10.1.1.1	learned	---	---	Pod-1/Node-101/eth1/3 (learned)	---	vlan-2501
EP-00:00:10:01:01:02	00:00:10:01:01:02	10.1.1.2...	learned	---	---	Pod-1/Node-103-104/N3k-3-VPC3-4 (learned)	---	vlan-2501

檢查是否已在本地枝葉上獲取本地終結點。在leaf1上檢查是否已獲知IP 10.1.1.1:

```
leaf1# show endpoint ip 10.1.1.1
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/
      Domain
+-----+-----+-----+-----+
46          vlan-2501    0000.1001.0101 L
eth1/3
Prod:VRF1   vlan-2501    10.1.1.1 L
                        eth1/3
```

如上所述，端點內容為：

- BD ( BD的內部VLAN為46 )，採用EPG(vlan-2501)的VLAN封裝，並在eth1/3上獲知MAC地址
- 採用IP 10.1.1.1的VRF(Prod:VRF1)

這可以理解為相當於傳統網路中的ARP條目。ACI不會在終端的ARP表中儲存ARP資訊。終結點僅在終結點表中可見。

枝葉上的ARP表僅用於L3Out下一跳。

```
leaf1# show ip arp vrf Prod:VRF1
Flags: * - Adjacencies learnt on non-active FHRP router
      + - Adjacencies synced via CFSOE
      # - Adjacencies Throttled for Glean
      D - Static Adjacencies attached to down interface      IP ARP Table for context Prod:VRF1
Total number of entries: 0
Address      Age      MAC Address      Interface
<NO ENTRY >
```

### 入口枝葉目標IP查詢 — 已知遠端終結點

假設目標IP為已知 ( 已知單播 )，下面是目標IP 10.1.2.1的「show endpoint」輸出。這是遠端學習，因為它不駐留在枝葉1上，具體指向本地學習它的隧道介面 ( 隧道4 )。

遠端終端僅包含IP或MAC，決不能同時包含這兩個條目。僅當在本地獲取終端時，才會發生同一終端中的MAC地址和IP地址。

```
leaf1# show endpoint ip 10.1.2.1
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
+-----+-----+-----+-----+-----+
---+
Prod:VRF1          10.1.2.1 p
tunnel4

```

```

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 ipvlan
  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

```

目標TEP是枝葉3和4 VPC對的任播TEP，並通過到主幹的上行鏈路獲知。

```

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], 02w06d, isis-isis_infra, isis-l1-int
  *via 10.0.128.64, eth1/51.8, [115/3], 02w06d, isis-isis_infra, isis-l1-int
  *via 10.0.88.64, eth1/52.126, [115/3], 02w06d, isis-isis_infra, isis-l1-int
  *via 10.0.88.94, eth1/50.128, [115/3], 02w06d, isis-isis_infra, isis-l1-int

```

可以使用「show system internal epm endpoint ip <ip>」命令收集IP 10.1.2.1的其他端點資訊。

```

leaf1# show system internal epm endpoint ip 10.1.2.1
MAC : 0000.0000.0000 ::: Num IPs : 1
IP# 0 : 10.1.2.1 ::: IP# 0 flags : ::: l3-sw-hit: No
Vlan id : 0 ::: Vlan vnid : 0 ::: VRF name : Prod:VRF1
BD vnid : 0 ::: VRF vnid : 2097154
Phy If : 0 ::: Tunnel If : 0x18010004
Interface : Tunnel4
Flags : 0x80004420 ::: sclass : 32771 ::: Ref count : 3
EP Create Timestamp : 10/01/2019 13:53:16.228319
EP Update Timestamp : 10/01/2019 14:04:40.757229
EP Flags : peer-aged|IP|sclass|timer|
:::

```

在輸出檢查中：

- VRF VNID已填充 — 這是用於將VXLAN中的幀封裝到交換矩陣的VNID。
- MAC地址是0000.0000.0000，因為從不會在遠端IP條目上填充MAC地址。
- BD VNID是未知的，對於路由幀，入口枝葉充當路由器並執行MAC重寫。這意味著遠端枝葉將不能檢視目標的BD，只能檢視VRF。

現在，該幀將封裝到發往遠端TEP 10.0.96.66的VXLAN幀中，其VXLAN ID為2097154，即VRF的VNID。它將在overlay-1路由表（IS-IS路由）中路由，並到達目的地TEP。此處它將到達leaf3或

leaf4 , 因為10.0.96.66是leaf3和leaf4 VPC對的任播TEP地址。

## 出口枝葉上的源IP學習

此處的輸出來自leaf3 , 但在leaf4上則類似。當資料包到達leaf3 ( 目標枝葉和TEP的所有者 ) 時 , leaf將學習VRF中資料包的源IP。

```
leaf3# show endpoint ip 10.1.1.1
```

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
+-----+-----+-----+-----+-----+
---+
Prod:VRF1                                10.1.1.1 p
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
```

目標TEP 10.0.88.95是leaf1的TEP地址 , 通過到脊柱的所有上行鏈路獲知。

## 出口枝葉上的目標IP查詢

最後一步是出口枝葉查詢目標IP。檢視10.1.2.1的終端表。

這將提供以下資訊 :

- 出口枝葉知道目的地10.1.2.1 ( 類似於路由表中的/32主機路由 ) , 並且路由是在正確的VRF中獲知的。
- 出口枝葉知道MAC 0000.1001.0201 ( 端點資訊 ) 。
- 輸出枝葉知道目的地為10.1.2.1的流量必須封裝在vlan-2502中 , 並在連線埠通道1(po1)上發出。

```
leaf3# show endpoint ip 10.1.2.1
```

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/ Domain	Encap VLAN	MAC Address IP Address	MAC Info/ IP Info	Interface
2	vlan-2502	0000.1001.0201	LpV	pol
Prod:VRF1	vlan-2502	10.1.2.1	LpV	pol

## 按資料路徑進行分類

使用APIC中的fTriage跟蹤資料路徑流。請記住，fTriage依賴於ELAM，因此它需要真正的資料流。這樣可確認完整的資料路徑，並確認資料包退出枝葉3埠1/16上的交換矩陣。

```

apic1# ftrriage route -ii LEAF:101 -sip 10.1.1.1 -dip 10.1.2.1
fTriage Status: {"dbgFtrriage": {"attributes": {"operState": "InProgress", "pid": "6888",
"apicId": "1", "id": "0"}}}
Starting ftrriage
Log file name for the current run is: ftlog_2019-10-01-21-17-54-175.txt
2019-10-01 21:17:54,179 INFO      /controller/bin/ftrriage route -ii LEAF:101 -sip 10.1.1.1 -dip
10.1.2.1
2019-10-01 21:18:18,149 INFO      ftrriage:      main:1165 Invoking ftrriage with default password
and default username: apic#fallback\admin
2019-10-01 21:18:39,194 INFO      ftrriage:      main:839 L3 packet Seen on bdsol-aci32-leaf1
Ingress: Eth1/3 Egress: Eth1/51 Vnid: 2097154
2019-10-01 21:18:39,413 INFO      ftrriage:      main:242 ingress encap string vlan-2501
2019-10-01 21:18:39,419 INFO      ftrriage:      main:271 Building ingress BD(s), Ctx
2019-10-01 21:18:41,240 INFO      ftrriage:      main:294 Ingress BD(s) Prod:BD1
2019-10-01 21:18:41,240 INFO      ftrriage:      main:301 Ingress Ctx: Prod:VRF1
2019-10-01 21:18:41,349 INFO      ftrriage:      pktrec:490 bdsol-aci32-leaf1: Collecting transient
losses snapshot for LC module: 1
2019-10-01 21:19:05,747 INFO      ftrriage:      main:933 SIP 10.1.1.1 DIP 10.1.2.1
2019-10-01 21:19:05,749 INFO      ftrriage:      unicast:973 bdsol-aci32-leaf1: <- is ingress node
2019-10-01 21:19:08,459 INFO      ftrriage:      unicast:1215 bdsol-aci32-leaf1: Dst EP is remote
2019-10-01 21:19:09,984 INFO      ftrriage:      misc:657 bdsol-aci32-leaf1:
DMAC(00:22:BD:F8:19:FF) same as RMAC(00:22:BD:F8:19:FF)
2019-10-01 21:19:09,984 INFO      ftrriage:      misc:659 bdsol-aci32-leaf1: L3 packet getting
routed/bounced in SUG
2019-10-01 21:19:10,248 INFO      ftrriage:      misc:657 bdsol-aci32-leaf1: Dst IP is present in
SUG L3 tbl
2019-10-01 21:19:10,689 INFO      ftrriage:      misc:657 bdsol-aci32-leaf1: RwdMAC
DIPo(10.0.96.66) is one of dst TEPs ['10.0.96.66']
2019-10-01 21:20:56,148 INFO      ftrriage:      main:622 Found peer-node bdsol-aci32-spine3 and
IF: Eth2/1 in candidate list
2019-10-01 21:21:01,245 INFO      ftrriage:      node:643 bdsol-aci32-spine3: Extracted Internal-
port GPD Info for lc: 2
2019-10-01 21:21:01,245 INFO      ftrriage:      fcls:4414 bdsol-aci32-spine3: LC trigger ELAM with
IFS: Eth2/1 Asic :0 Slice: 0 Srcid: 32
2019-10-01 21:21:33,894 INFO      ftrriage:      main:839 L3 packet Seen on bdsol-aci32-spine3
Ingress: Eth2/1 Egress: LC-2/0 FC-22/0 Port-1 Vnid: 2097154
2019-10-01 21:21:33,895 INFO      ftrriage:      pktrec:490 bdsol-aci32-spine3: Collecting transient
losses snapshot for LC module: 2
2019-10-01 21:21:54,487 INFO      ftrriage:      fib:332 bdsol-aci32-spine3: Transit in spine
2019-10-01 21:22:01,568 INFO      ftrriage:      unicast:1252 bdsol-aci32-spine3: Enter
dbg_sub_nexthop with Transit inst: ig infra: False glbs.dipo: 10.0.96.66
2019-10-01 21:22:01,682 INFO      ftrriage:      unicast:1417 bdsol-aci32-spine3: EP is known in COOP
(DIPo = 10.0.96.66)
2019-10-01 21:22:05,713 INFO      ftrriage:      unicast:1458 bdsol-aci32-spine3: Infra route
10.0.96.66 present in RIB
2019-10-01 21:22:05,713 INFO      ftrriage:      node:1331 bdsol-aci32-spine3: Mapped LC interface:
LC-2/0 FC-22/0 Port-1 to FC interface: FC-22/0 LC-2/0 Port-1

```



```

2019-10-01 21:22:10,799 INFO      ftriage:      node:460  bdsol-aci32-spine3: Extracted GPD Info
for fc: 22
2019-10-01 21:22:10,799 INFO      ftriage:      fcls:5748 bdsol-aci32-spine3: FC trigger ELAM with
IFS: FC-22/0 LC-2/0 Port-1 Asic :0 Slice: 2 Srcid: 24
2019-10-01 21:22:29,322 INFO      ftriage:      unicast:1774 L3 packet Seen on FC of node: bdsol-
aci32-spine3 with Ingress: FC-22/0 LC-2/0 Port-1 Egress: FC-22/0 LC-2/0 Port-1 Vnid: 2097154
2019-10-01 21:22:29,322 INFO      ftriage:      pktrec:487  bdsol-aci32-spine3: Collecting transient
losses snapshot for FC module: 22
2019-10-01 21:22:31,571 INFO      ftriage:      node:1339 bdsol-aci32-spine3: Mapped FC interface:
FC-22/0 LC-2/0 Port-1 to LC interface: LC-2/0 FC-22/0 Port-1
2019-10-01 21:22:31,572 INFO      ftriage:      unicast:1474 bdsol-aci32-spine3: Capturing Spine
Transit pkt-type L3 packet on egress LC on Node: bdsol-aci32-spine3 IFS: LC-2/0 FC-22/0 Port-1
2019-10-01 21:22:31,991 INFO      ftriage:      fcls:4414 bdsol-aci32-spine3: LC trigger ELAM with
IFS: LC-2/0 FC-22/0 Port-1 Asic :0 Slice: 1 Srcid: 0
2019-10-01 21:22:48,952 INFO      ftriage:      unicast:1510 bdsol-aci32-spine3: L3 packet Spine
egress Transit pkt Seen on bdsol-aci32-spine3 Ingress: LC-2/0 FC-22/0 Port-1 Egress: Eth2/3
Vnid: 2097154
2019-10-01 21:22:48,952 INFO      ftriage:      pktrec:490  bdsol-aci32-spine3: Collecting transient
losses snapshot for LC module: 2
2019-10-01 21:23:50,748 INFO      ftriage:      main:622  Found peer-node bdsol-aci32-leaf3 and
IF: Eth1/51 in candidate list
2019-10-01 21:24:05,313 INFO      ftriage:      main:839  L3 packet Seen on bdsol-aci32-leaf3
Ingress: Eth1/51 Egress: Eth1/12 (Pol) Vnid: 11365
2019-10-01 21:24:05,427 INFO      ftriage:      pktrec:490  bdsol-aci32-leaf3: Collecting transient
losses snapshot for LC module: 1
2019-10-01 21:24:24,369 INFO      ftriage:      nxos:1404 bdsol-aci32-leaf3: nxos matching rule
id:4326 scope:34 filter:65534
2019-10-01 21:24:25,698 INFO      ftriage:      main:522  Computed egress encaps string vlan-2502
2019-10-01 21:24:25,704 INFO      ftriage:      main:313  Building egress BD(s), Ctx
2019-10-01 21:24:27,510 INFO      ftriage:      main:331  Egress Ctx Prod:VRF1
2019-10-01 21:24:27,510 INFO      ftriage:      main:332  Egress BD(s): Prod:BD2
2019-10-01 21:24:30,536 INFO      ftriage:      unicast:1252 bdsol-aci32-leaf3: Enter dbg_sub_nexthop
with Local inst: eg infra: False glbs.dipo: 10.0.96.66
2019-10-01 21:24:30,537 INFO      ftriage:      unicast:1257 bdsol-aci32-leaf3: dbg_sub_nexthop
invokes dbg_sub_eg for vip
2019-10-01 21:24:30,537 INFO      ftriage:      unicast:1784 bdsol-aci32-leaf3: <- is egress node
2019-10-01 21:24:30,684 INFO      ftriage:      unicast:1833 bdsol-aci32-leaf3: Dst EP is local
2019-10-01 21:24:30,685 INFO      ftriage:      misc:657  bdsol-aci32-leaf3: EP if(Pol) same as
egr if(Pol)
2019-10-01 21:24:30,943 INFO      ftriage:      misc:657  bdsol-aci32-leaf3: Dst IP is present in
SUG L3 tbl
2019-10-01 21:24:31,242 INFO      ftriage:      misc:657  bdsol-aci32-leaf3: RW seg_id:11365 in
SUG same as EP segid:11365
2019-10-01 21:24:37,631 INFO      ftriage:      main:961  Packet is Exiting fabric with peer-
device: bdsol-aci32-n3k-3 and peer-port: Ethernet1/12

```

## 使用ELAM Assistant應用程式在出口枝葉上捕獲資料包

下面是使用leaf3上的ELAM Assistant應用捕獲的來自主幹的資料包。這顯示：

- 來自外部第4層資訊的VNID(VNID為2097154)。
- 外部L3標頭來源TEP和目的地TEP。

## ELAM助理 — L3流出口枝葉 (第1部分)

Device Type	LEAF
Packet Direction	egress (spine LC -> leaf)
Incoming I/F	eth1/51
<b>L2 Header</b>	
Destination MAC	000C.0C0C.0C0C
Source MAC	000C.0C0C.0C0C
Access Encap VLAN	No VLAN Tag
CoS	No VLAN Tag (= No CoS)
<b>L3 Header</b>	
L3 Type	IPv4
Destination IP	10.1.2.1
Source IP	10.1.1.1
IP Protocol	0x1 (ICMP)
DSCP	0
TTL	254
Don't Fragment Bit	0x0 (0x0)
IP Checksum	Unsupported for ELAM with VxLAN data
IP Packet Length	Unsupported for ELAM with VxLAN data

## ELAM助理 — L3流出口枝葉 ( 第2部分 )

L2 Header (Outer VxLAN)	
Destination MAC	000C.0C0C.0C0C
Source MAC	000D.0D0D.0D0D
Access Encap VLAN	2
CoS	0

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	IPvLAN
DL (Don't Learn) Bit	0 (not set)
Src Policy Applied Bit	0 (Contract has yet to be applied)
Dst Policy Applied Bit	0 (Contract has yet to be applied)
Source EPG (sclass / src pcTag)	0x8002 / 32770 (Prod:App:EPG1)
VRF/BD VNID	2097154 (Prod:VRF1)

「資料包轉發資訊」部分證明它是在埠通道1上發出的

## ELAM Assistant — L3出口枝葉 — 資料包轉發資訊

Packet Forwarding Information	
<b>Forward Result</b>	
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
<b>Contract</b>	
Destination EPG pcTag (dclass)	32771 (null)
Source EPG pcTag (sclass)	32770 (null)
Contract was applied	1 (Contract was applied on this node)
<b>Drop</b>	
Drop Code	no drop

## 未知終結點的疑難解答 workflow

本節介紹當入口枝葉不知道目標IP時有何不同。

### 入口枝葉目標IP查詢

第一步是檢查是否存在目標IP的終端學習。

```
leaf1# show endpoint ip 10.1.2.1
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
+-----+-----+-----+-----+
--+
<NO ENTRY>
```

終結點表中沒有與目標對應的路由，因此下一步是檢查路由表，查詢與目標匹配的最長字首路由：

```
leaf1# show ip route 10.1.2.1 vrf Prod:VRF1
IP Route Table for VRF "Prod:VRF1"
'*' denotes best ucast next-hop
'***' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.1.2.0/24, ubest/mbest: 1/0, attached, direct, pervasive
  *via 10.0.8.65%overlay-1, [1/0], 01:40:18, static, tag 4294967294
    recursive next hop: 10.0.8.65/32%overlay-1
```

落在/24 BD子網10.1.2.0/24上意味著枝葉將幀封裝在VXLAN中，目標TEP為10.0.8.65 (主幹上的任播v4)。幀將使用VXLAN ID，即VRF VNID。

### 主幹上的COOP查詢 — 目標IP已知

資料包將到達IP資料庫中執行COOP查詢的其中一個主幹。必須驗證源，並且需要從COOP資料庫正確學習目標IP。

要在COOP資料庫中查詢IP，關鍵是VRF VNID(2097154示例中為VRF VNID)

從下面的輸出中，可以確認COOP資料庫具有正確來自TEP 10.0.88.95(leaf1)的源IP條目。

```
spine1# show coop internal info ip-db key 2097154 10.1.1.1
IP address : 10.1.1.1
Vrf : 2097154
Flags : 0
EP bd vnid : 15302583
EP mac : 00:00:10:01:01:01
Publisher Id : 10.0.88.95
Record timestamp : 10 01 2019 14:16:50 522482647
```

```
Publish timestamp : 10 01 2019 14:16:50 532239332
Seq No: 0
Remote publish timestamp: 01 01 1970 00:00:00 0
URIB Tunnel Info
Num tunnels : 1
    Tunnel address : 10.0.88.95
    Tunnel ref count : 1
```

以下輸出顯示，COOP資料庫具有來自TEP 10.0.96.66 ( leaf3和4 VPC對的任播TEP ) 的目標IP條目正確

```
spinel# show coop internal info ip-db key 2097154 10.1.2.1
IP address : 10.1.2.1
Vrf : 2097154
Flags : 0
EP bd vnid : 15957974
EP mac : 00:00:10:01:02:01
Publisher Id : 10.0.88.90
Record timestamp : 10 01 2019 14:52:52 558812544
Publish timestamp : 10 01 2019 14:52:52 559479076
Seq No: 0
Remote publish timestamp: 01 01 1970 00:00:00 0
URIB Tunnel Info
Num tunnels : 1
    Tunnel address : 10.0.96.66
    Tunnel ref count : 1
```

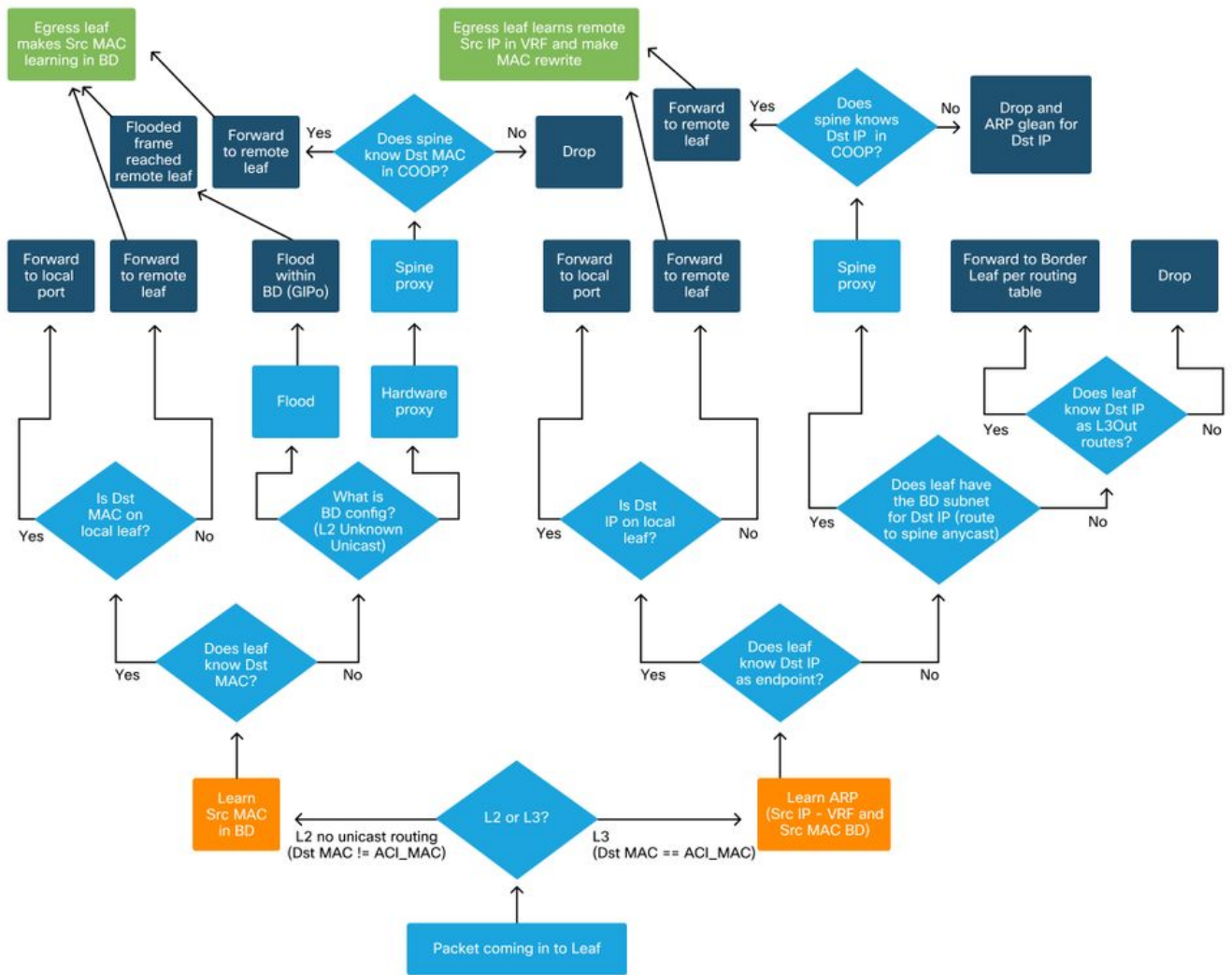
在此案例中，COOP知道目的地IP，因此會將VXLAN封包中外部IP標頭的目的地IP重新寫入10.0.96.66，然後傳送到leaf3或leaf4 ( 取決於ECMP雜湊 )。請注意，VXLAN幀的源IP未更改，因此它仍然是leaf1 PTEP。

## 主幹上的COOP查詢 — 目標IP未知

如果目標IP的COOP條目未填充 ( 無提示終端或過期 )，則主幹將生成一個ARP收集來解析它。有關詳細資訊，請參閱「多Pod轉發」部分。

## ACI轉發摘要

下圖總結了第2層和第3層的ACI轉發使用案例。



## 關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。