在Catalyst 9000系列交換機上實施BGP EVPN保護覆蓋分段

目錄

簡介

本檔案介紹如何在Catalyst 9000系列交換器上實作BGP EVPN VXLAN保護的重疊分段。

必要條件

需求

思科建議您瞭解以下主題：

- BGP EVPN VxLAN概念
- BGP EVPN單播故障排除
- BGP EVPN VxLAN路由策略

採用元件

本文中的資訊係根據以下軟體和硬體版本：

- Catalyst 9300
- Catalyst 9400
- Catalyst 9500
- Catalyst 9600
- Cisco IOS® XE 17.12.1及更高版本

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

背景資訊

高級功能說明

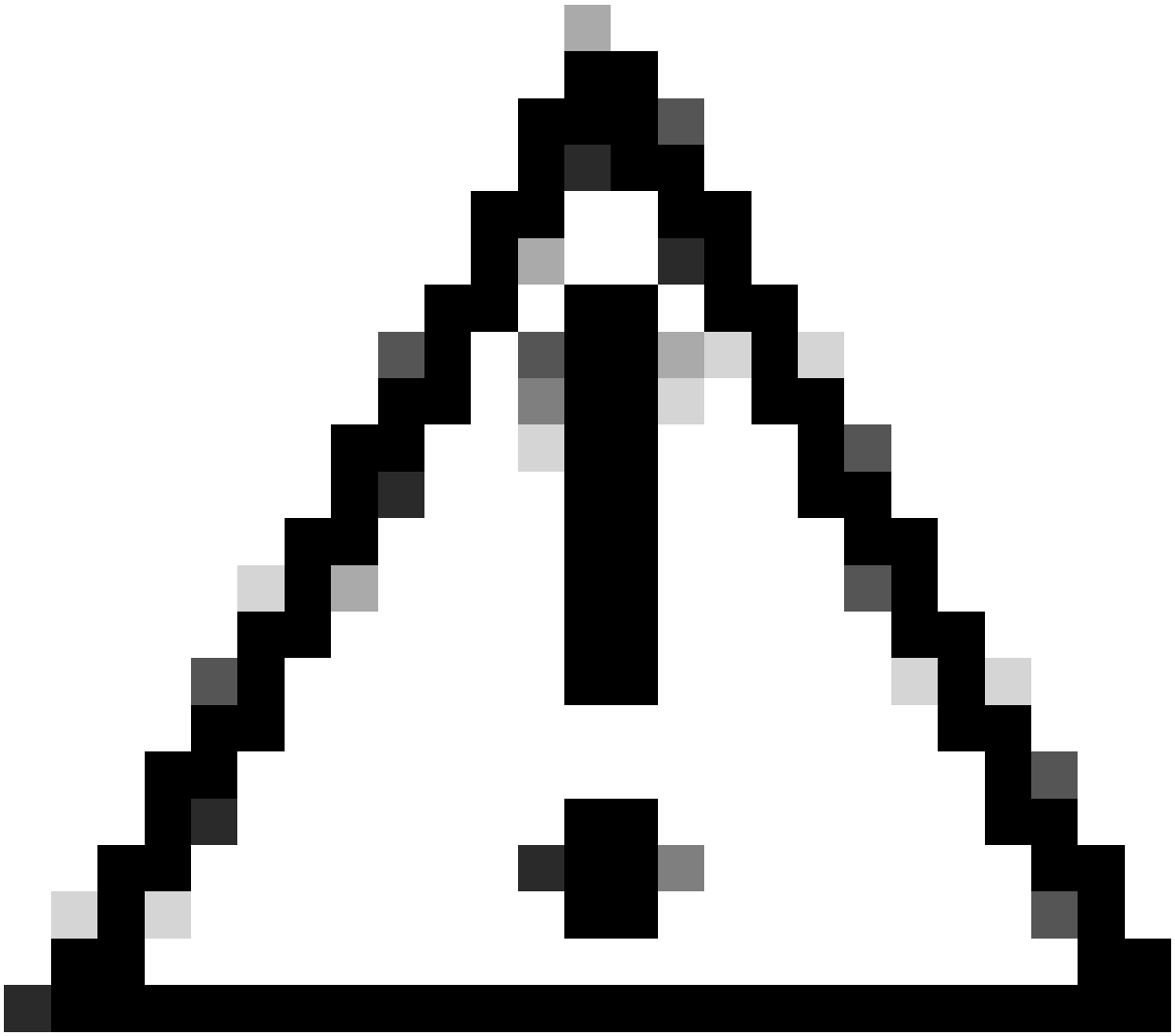
受保護分段功能是一種安全措施，可防止埠相互轉發流量，即使它們位於同一個VLAN和同一交換機上也是如此

- 此功能類似於「switchport protected」或專用Vlan，但適用於EVPN交換矩陣。
- 此設計強制所有流量傳送到CGW，在傳送到最終目的地之前，防火牆可以檢查這些流量。
- 流量是受控制的、確定性的，並且易於使用集中式安全裝置進行檢查。

檔案詳細資訊

本檔案是第2部分或第3部分相互關聯的檔案：

- 檔案1：在Catalyst 9000系列交換器上實作BGP EVPN路由原則說明如何控制重疊中的BGP BUM流量，必須首先設定
- 檔案2：本檔案。本文基於文檔1的覆蓋設計和策略，描述了「protected」關鍵字的實現
- 文檔3：在Catalyst 9000系列交換機上實施BGP EVPN DHCP第2層中繼介紹DHCP中繼在僅L2 VTEP上的工作方式



注意：在實施受保護分段配置之前，您必須實施文檔1中的配置。

受保護的區段型別

完全隔離

- 僅允許北到南通訊，

- 網關透過「預設網關通告」CLI通告到交換矩陣中

大部分隔離

- 允許北向南通訊（在此使用案例中，根據防火牆流量策略允許東/西流量流）
- 允許從東向西通訊（基於防火牆流量策略）
- 網關位於交換矩陣外部，並且不會使用「預設網關通告」CLI通告SVI

交換機行為

- 即使主機連線到同一台交換機，它們也不能直接相互通訊(當主機位於同一VRF/Vlan/網段時，ARP請求不會傳送到同一台交換機上的其他埠)。
- L2 VTEP之間沒有BUM流量(使用路由原則設定過濾的IMET字首)
- 來自主機的所有資料包都將中繼到邊界枝葉以轉發。（這意味著主機1要與同一枝葉上的主機2通訊，流量被固定到CGW）。

路由型別2處理

- 接入枝葉會通告本地RT2，並設定E-Tree擴展社群和枝葉標誌。
- 接入枝葉不會安裝資料平面中設定了E-Tree Extended Community和Leaf標誌的任何遠端RT2。
- 接入枝葉不會將彼此RT2安裝在資料平面中。
- 接入枝葉和邊界枝葉(CGW)相互將RT2安裝在資料平面中。
- 無需在Access Leaf或Border Leaf上更改配置。

設計摘要

- 對於廣播(BUM)，RT3拓撲是中心輻射型，以強制廣播流量（例如ARP）到達GCW。
- 為了考慮主機移動性，RT2在BGP控制平面處為全網狀（當主機從一個VTEP移動到另一個VTEP時，RT2中的序列號會遞增）
- 資料平面有選擇地安裝MAC地址。
    - 枝葉僅安裝包含DEF GW屬性的本地MAC和RT2
    - CGW沒有受保護的KW，並且在其資料平面中安裝了所有本地MAC和遠端RT2。

技術

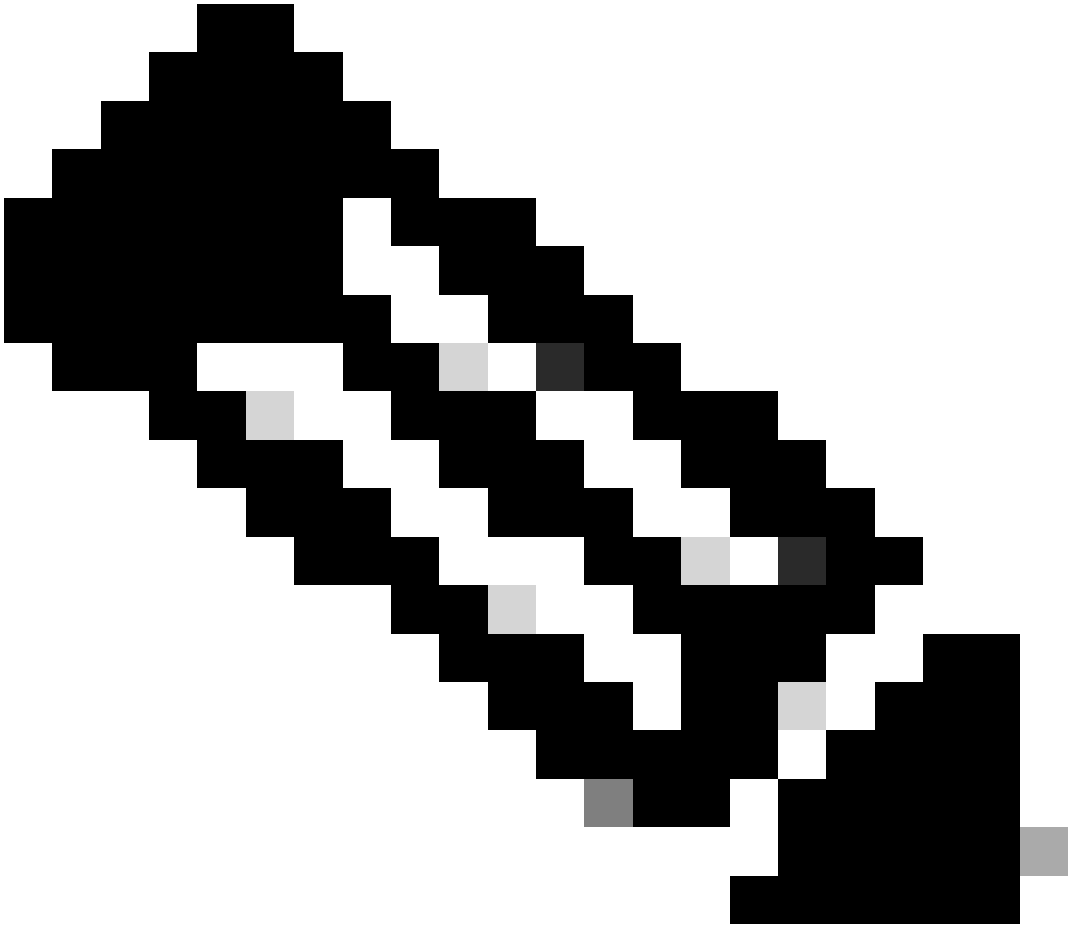
VRF	虛擬路由轉送	定義與其他VRF和全局IPv4/IPv6路由域分開的第3層路由域
AF	地址系列	定義BGP處理的型別字首和路由資訊
AS	自治系統	屬於某個網路或網路集合的一組網際網路可路由IP字首，它們全部由單個實體或組織管理、控制和監督

EVPN	乙太網路虛擬私人網路	允許BGP傳輸第2層MAC和第3層IP資訊的擴展是EVPN，它使用多協定邊界網關協定(MP-BGP)作為協定，以分發屬於VXLAN重疊網路的可達性資訊。
VXLAN	虛擬可擴充LAN（區域網路）	VXLAN的用途是克服VLAN和STP的固有限制。建議的IETF標準[RFC 7348]可提供與VLAN相同的乙太網第2層網路服務，但具有更高的靈活性。從功能上講，它是UDP內MAC封裝協定，在第3層底層網路上作為虛擬重疊運行。
CGW	集中網關	以及網關SVI不在每個枝葉上的EVPN的實施。相反，所有路由都由使用不對稱IRB（整合路由和橋接）的特定枝葉完成
DEF網關	預設閘道	在「l2vpn evpn」配置部分下，透過「default-gateway advertise enable」命令增加到MAC/IP字首的BGP擴展社群屬性。
IMET (RT3)	內含組播乙太網標籤（路由）	也稱為BGP型別3路由。此路由型別用於EVPN中在VTEP之間傳送BUM（廣播/未知單點傳播/多點傳送）流量。
RT2	路由型別2	代表主機MAC或閘道MAC-IP的BGP MAC或MAC/IP首碼
EVPN管理器	EVPN管理員	各種其他元件的中央管理元件（例如：從SISF獲知並向L2RIB傳送訊號）
SISF	交換機整合安全功能	EVPN使用的唯一主機跟蹤表，用於瞭解枝葉上存在哪些本地主機
L2RIB	第2層路由資訊庫	在用於管理BGP、EVPN管理器、L2FIB之間的互動的中間元件中
FED	轉發引擎驅動程式	對ASIC（硬體）層進行程式設計
MATM	Mac位址表管理員	IOS MATM：僅安裝本地地址和 FED MATM：安裝從控制平面獲知的本地和遠端地址的硬體表，屬於硬體轉發平面的一部分

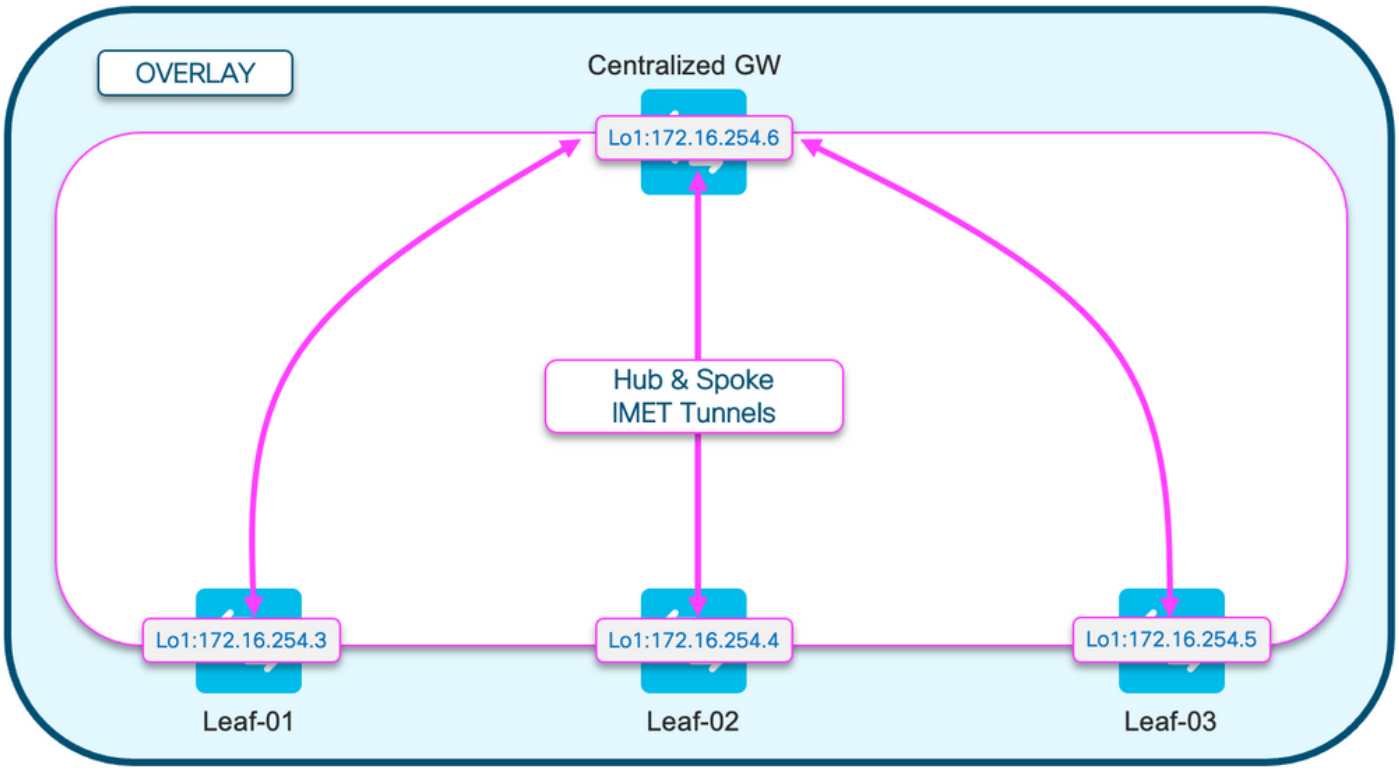
流程圖

路由型別2 (RT2)圖

此圖顯示了第2類MAC/MAC-IP主機字首的完整網狀設計。

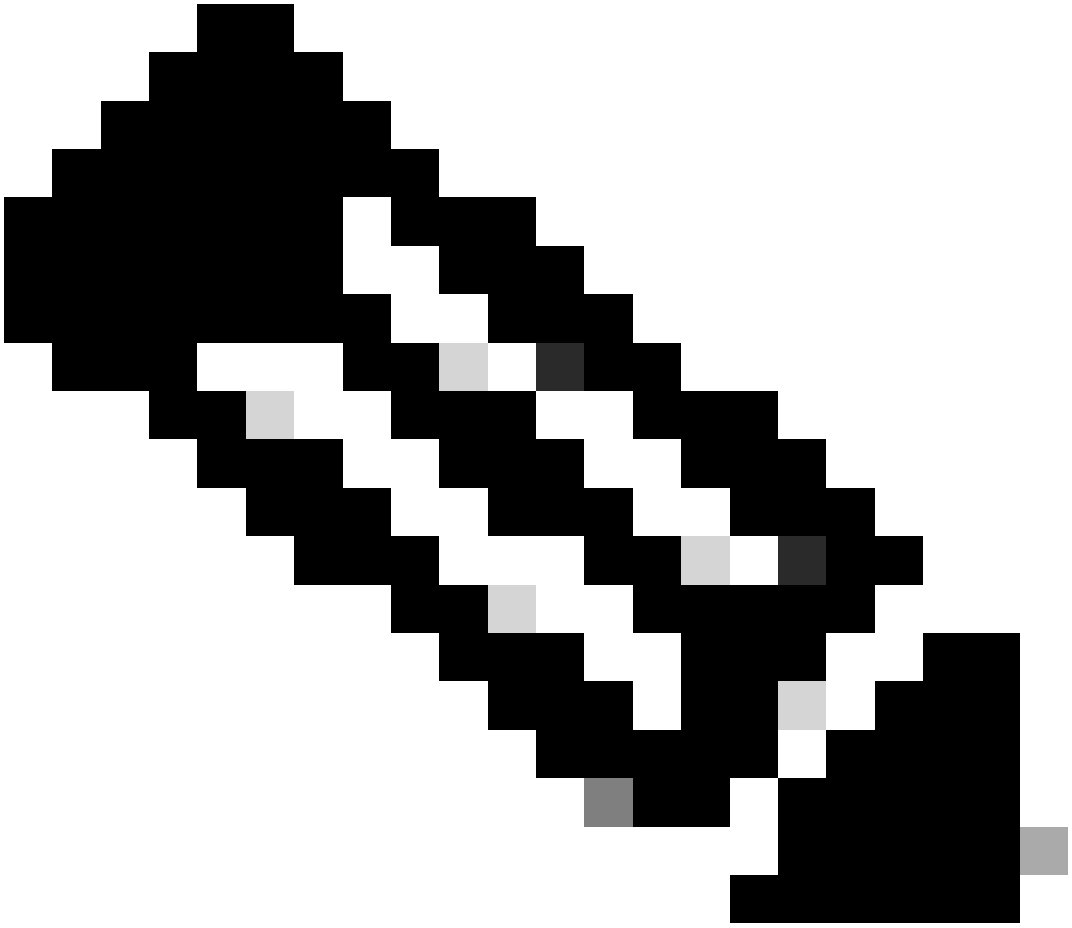


註：需要全網狀網路才能支援移動性和漫遊

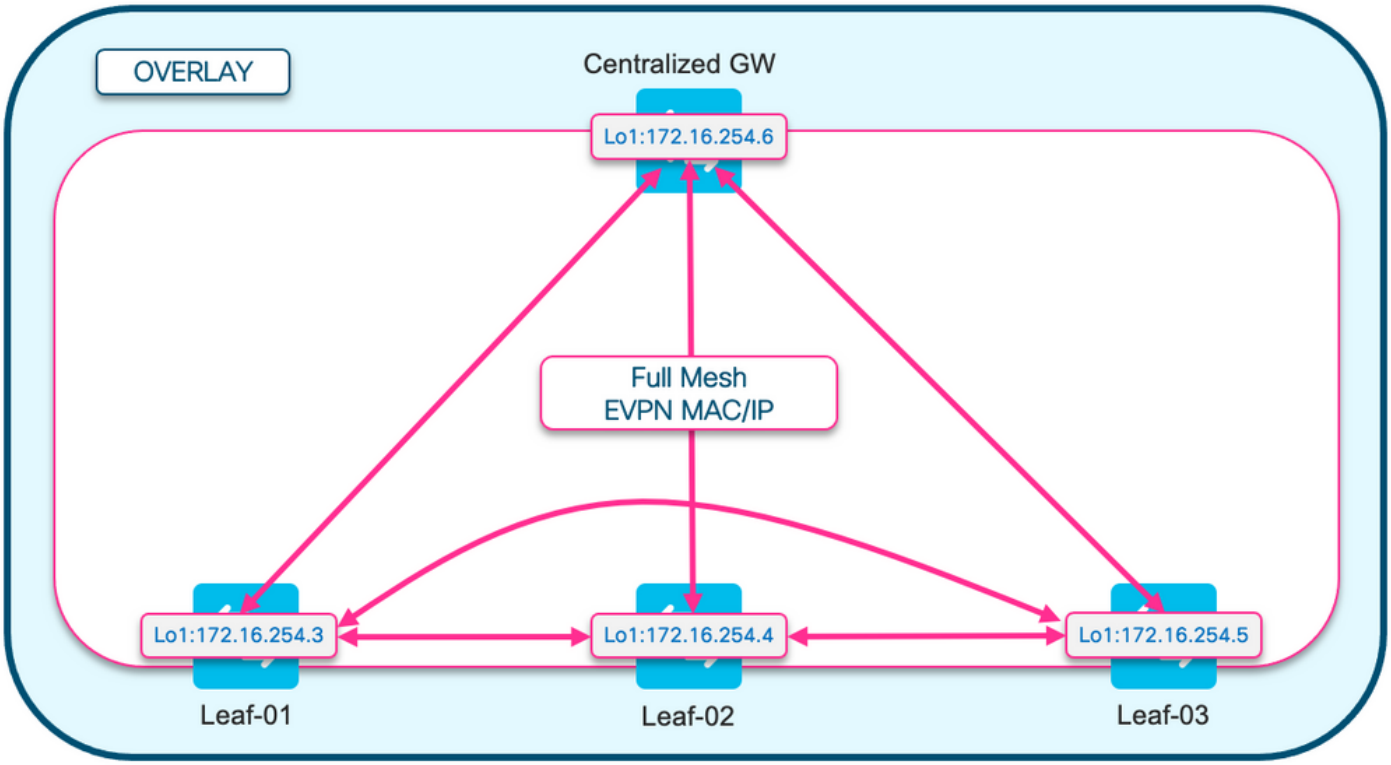


路由型別3 (RT3)圖

此圖顯示廣播IMET (RT3)隧道的星型設計

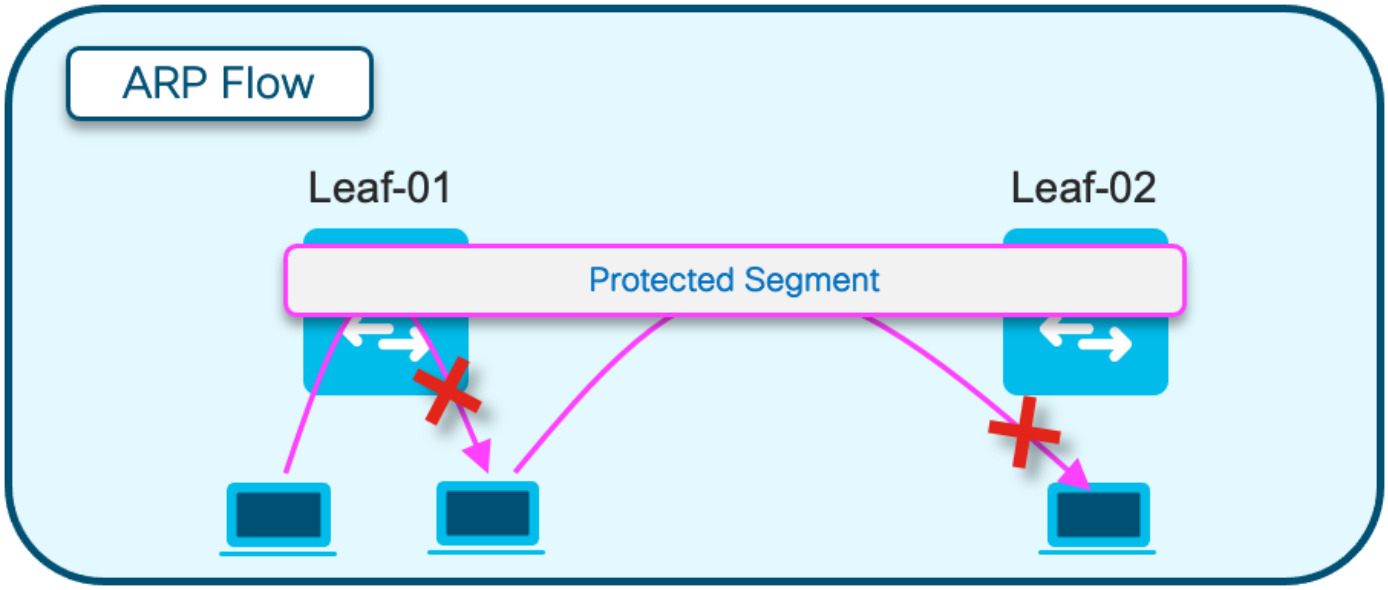


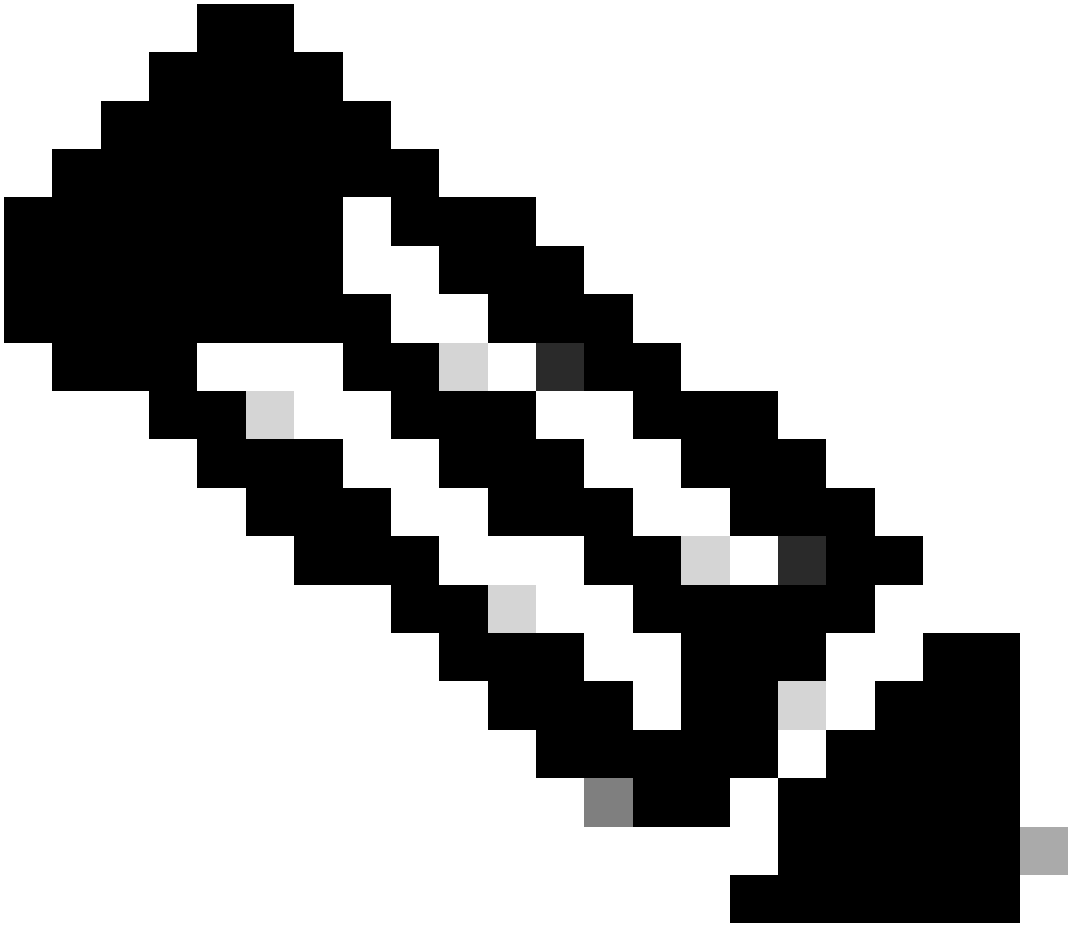
注意：中心輻射型廣播是必需的，以防止具有同一網段的枝葉直接互相傳送廣播。



位址解析(ARP)圖表

此圖顯示ARP不能到達同一EPVN網段中的任何主機。當另一台主機的主機ARP時，只有CGW會獲得此ARP並回覆



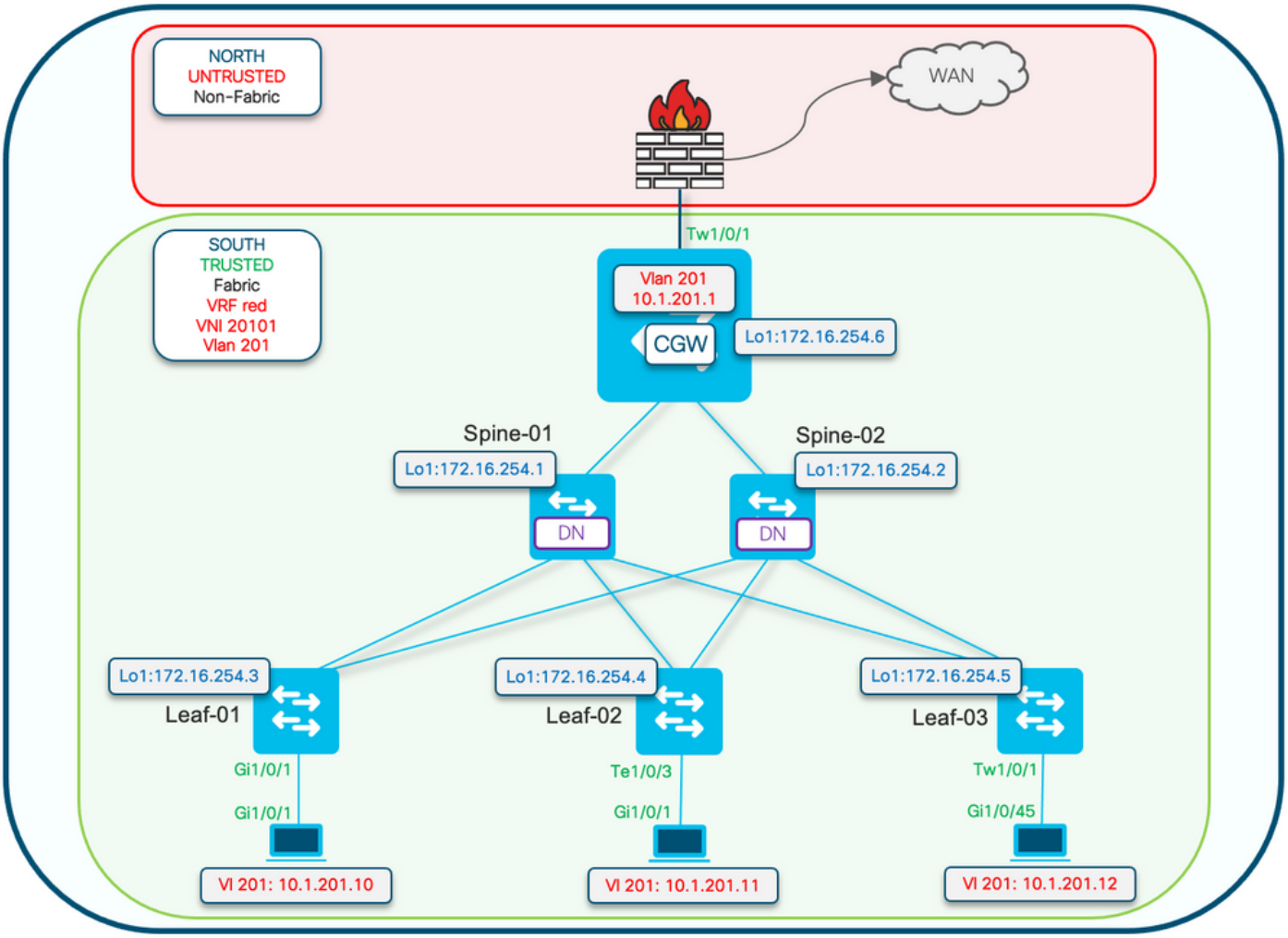


注意：此ARP行為更改透過使用「protected」關鍵字進行例項化。

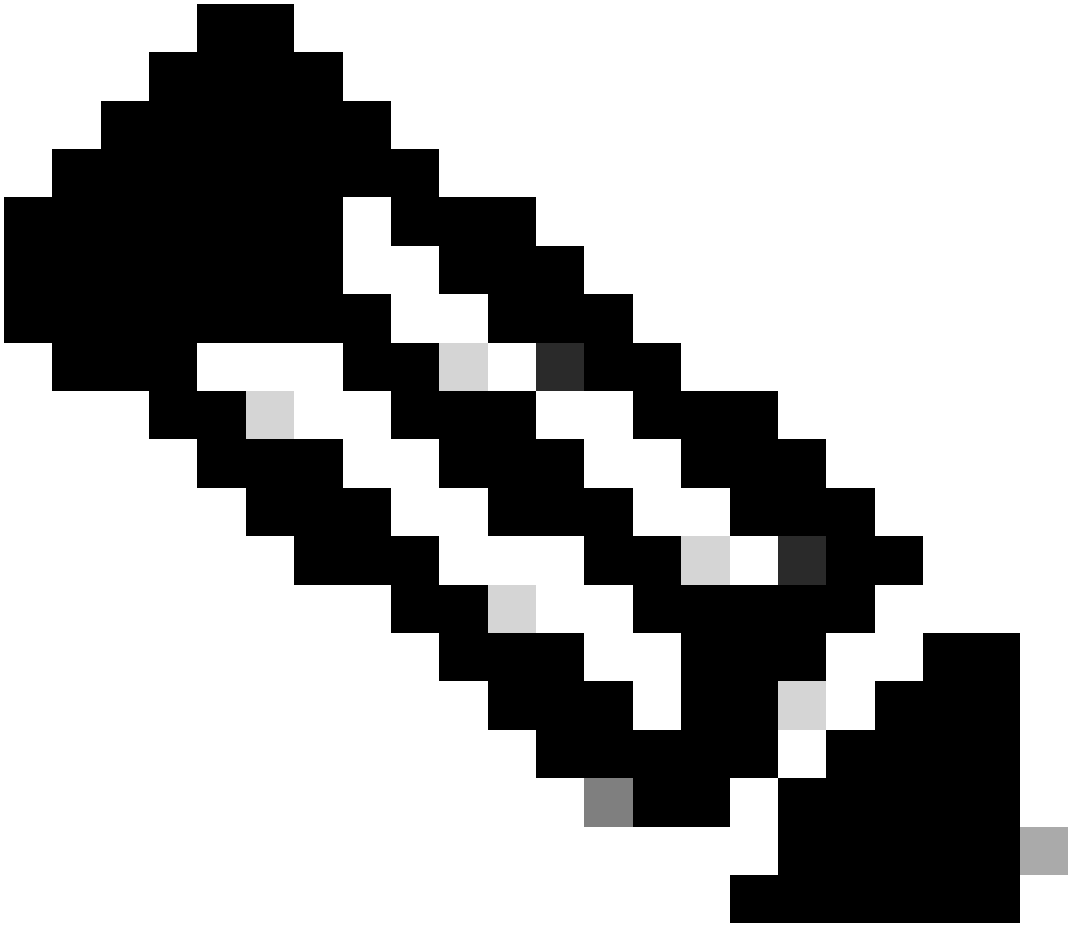
範例：成員evpn-instance 202 vni 20201 protected

配置（完全隔離）

網路圖表



受保護的配置關鍵字應用於枝葉交換機。CGW是一個混合裝置，會安裝所有mac地址。



注意：[在Catalyst 9000系列交換機上實施BGP EVPN路由策略](#)中顯示了控制導入/導出IMET字首的路由策略社群清單和路由對映配置。 本文檔中僅顯示受保護的段差異。

枝葉01（基本EVPN配置）

<#root>

```
Leaf-01#

show run | sec l2vpn
l2vpn evpn

 replication-type static

 flooding-suppression address-resolution disable  <-- Disables ARP caching so ARP is always sent up to t

 router-id Loopback1
l2vpn evpn

instance 201
```

```
 vlan-based
 encapsulation vxlan

replication-type ingress                          <-- Sets segment to use Unicast replication of BUM traff

 multicast advertise enable
```

<#root>

```
Leaf01#

show run | sec vlan config

vlan configuration 201
 member evpn-instance 201 vni 20101

protected  <-- protected keyword added
```

CGW（基本配置）

<#root>

```
CGW#

show running-config | beg l2vpn evpn instance 201

l2vpn evpn instance 201 vlan-based
 encapsulation vxlan
 replication-type ingress

 default-gateway advertise enable    <-- adds the BGP attribute EVPN DEF GW:0:0 to the MAC/IP prefix

 multicast advertise enable
```

<#root>

```
CGW#

show running-config | sec vlan config

vlan configuration 201
 member evpn-instance 201 vni 20101
```

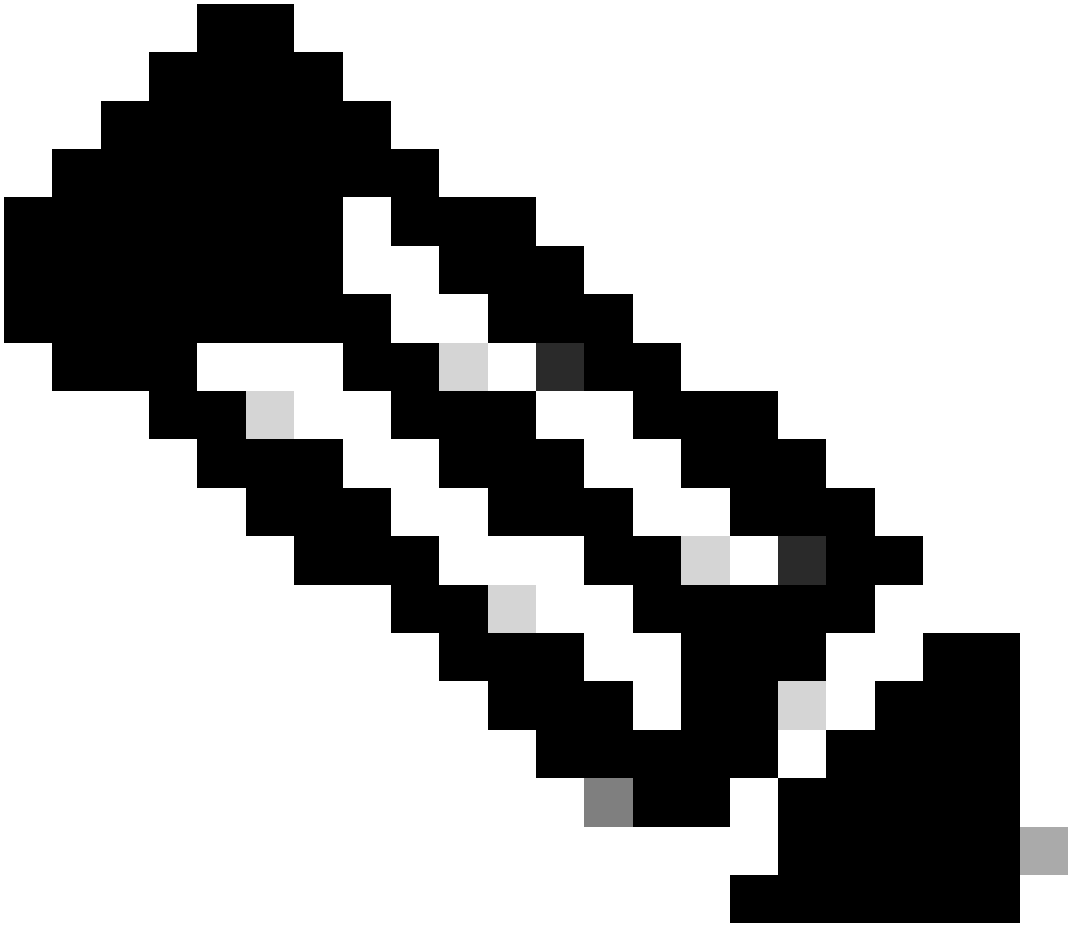
<#root>

```
CGW#

show run int nve 1

Building configuration...

Current configuration : 313 bytes
!
```

```
interface nve1
 no ip address
 source-interface Loopback1
 host-reachability protocol bgp
 member vni 20101 ingress-replication local-routing  <-- 'ingress-replication' (Unicast all BUM traffic)
```

<#root>

```
CGW#
show run interface vlan 201

Building configuration...

Current configuration : 231 bytes
!
interface Vlan201

mac-address 0000.beef.cafe          <-- MAC is static in this example for viewing simplicity. This is no
 vrf forwarding red                  <-- SVI is in VRF red
 ip address 10.1.201.1 255.255.255.0
 no ip redirects
 ip local-proxy-arp                 <-- Sets CGW to Proxy reply even for local subnet ARP requests
 ip pim sparse-mode
 ip route-cache same-interface      <-- This is auto added when local-proxy-arp is configured. However,
 ip igmp version 3
 no autostate
```



注意：在CGW未應用BGP策略。允許CGW接收和傳送所有字首型別(RT2、RT5 / RT3)。

驗證（完全隔離）

EVI詳細資料

<#root>

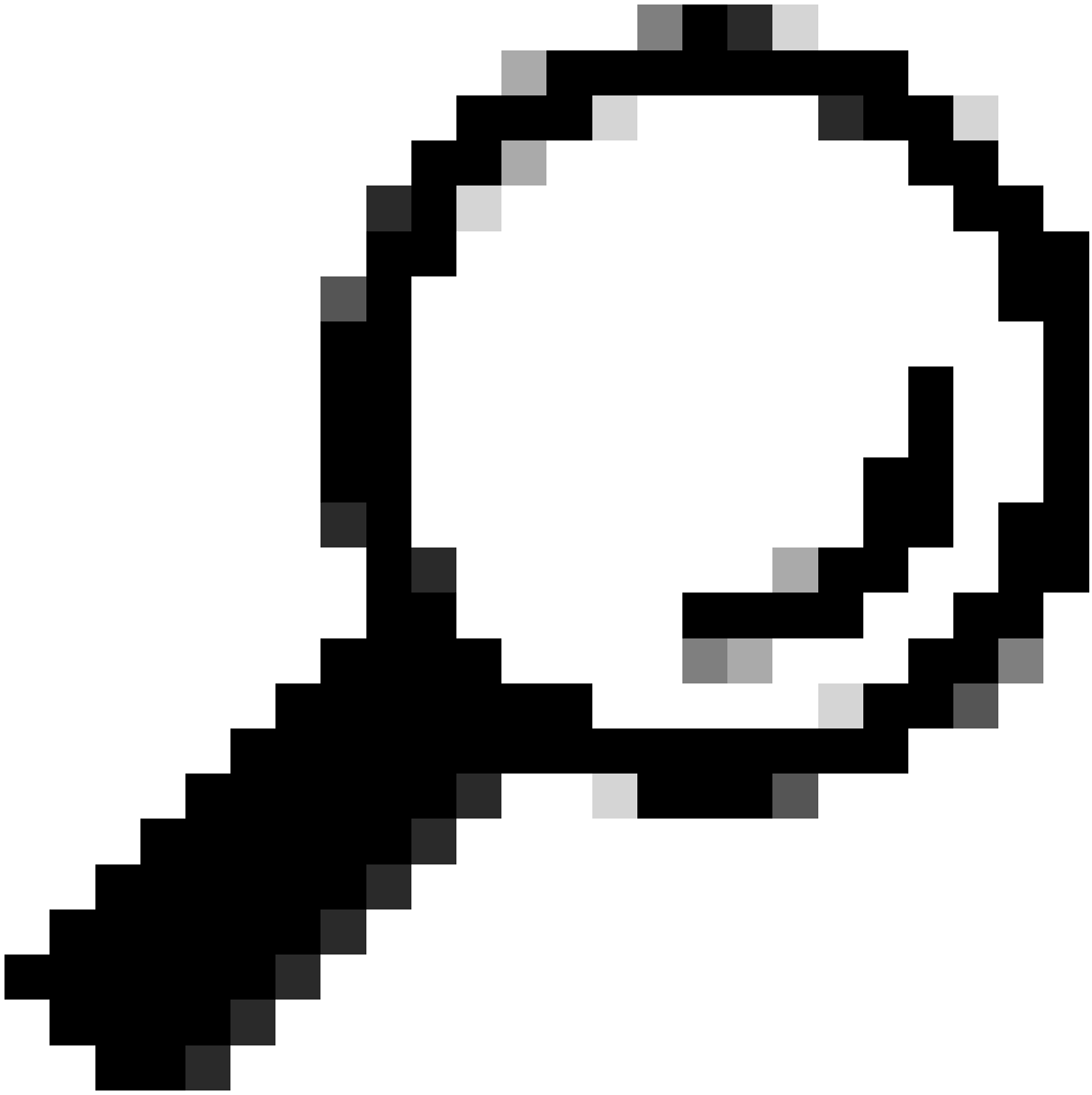
```
Leaf01#

sh l2vpn evpn evi 201 detail

EVPN instance:       201 (VLAN Based)
  RD:                172.16.254.3:201 (auto)
  Import-RTs:        65001:201
  Export-RTs:        65001:201
  Per-EVI Label:     none
  State:             Established
  Replication Type:  Ingress
```

```
  Encapsulation:     vxlan
  IP Local Learn:    Enabled (global)
  Adv. Def. Gateway: Disabled (global)
  Re-originate RT5:  Disabled
  Adv. Multicast:    Enabled
  AR Flood Suppress: Disabled (global)


Vlan:              201
    Protected:       True (local access p2p blocked)  <-- Vlan 201 is in protected mode

<...snip...>
```

本地RT2生成（本地主機到RT2）

驗證從本地主機學習到RT2生成的元件依賴關係鏈：

- SISF（當枝葉沒有SVI時，SISF仍透過主機的ARP幀收集主機資訊）
- EVPN管理器
- L2RIB
- BGP

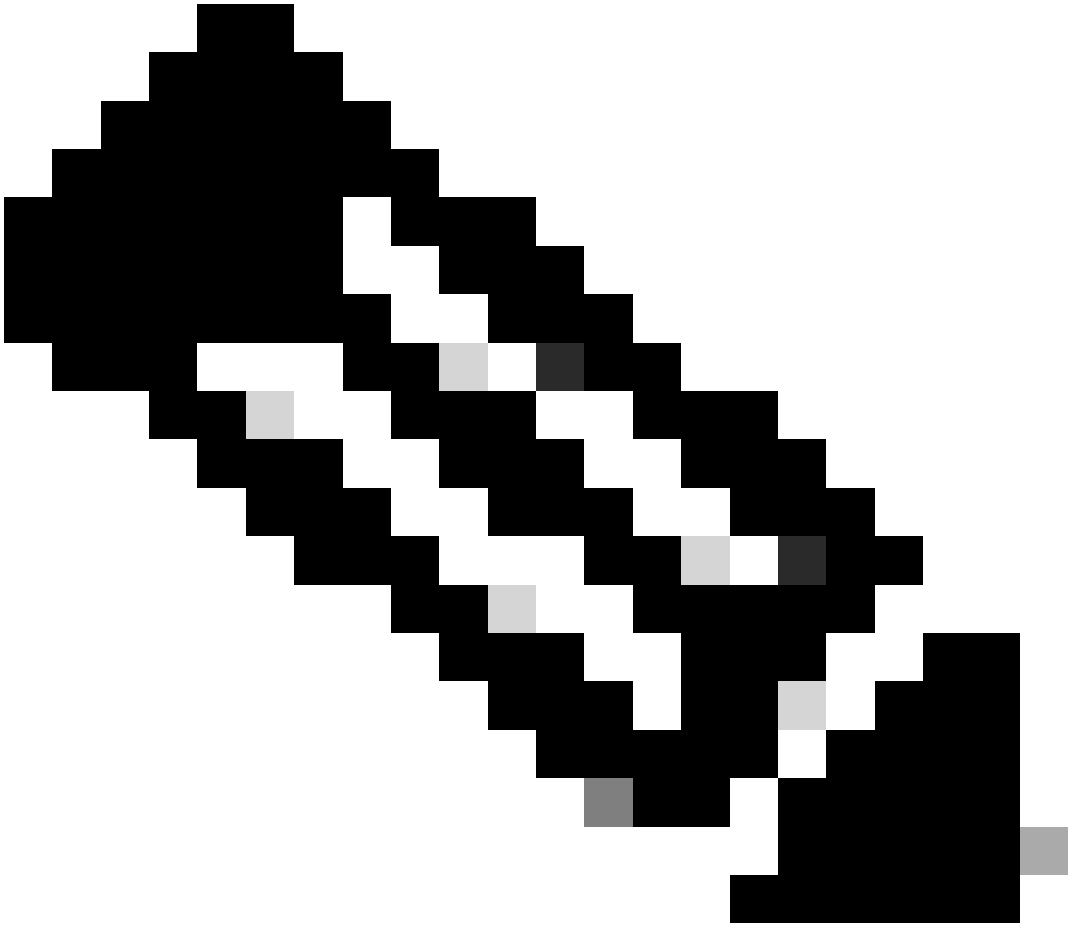


提示：如果先前的元件未正確設定，整個相依性鏈會中斷（例如：SISF沒有en專案，則BGP無法建立RT2）。

SISF

驗證SISF已在DB中獲知主機（從DHCP或ARP獲知主機資訊）

- SISF從IOS-MATM learning獲取MAC條目，然後向上傳送至EVPN Mgr（必須使用策略「evpn-sisf-policy」進行MAC可訪問）。
- SISF在本地VTEP上收集IP/MAC繫結，並使用EVPN管理器將資訊程式設計為透過BGP到其他枝葉的/32路由。



注意：在此場景中，主機有一個靜態IP，因此SISF使用ARP來收集主機詳細資訊。在「Modly Isolated」（大部分隔離）部分中顯示DHCP和DHCP監聽。

<#root>

```
Leaf01#

show device-tracking database vlanid 201

vlanDB has 1 entries for vlan 201, 1 dynamic
Codes: L - Local, S - Static, ND - Neighbor Discovery, ARP - Address Resolution Protocol, DH4 - IPv4 DHC
Preflevel flags (prlvl):
0001:MAC and LLA match     0002:Orig trunk            0004:Orig access
0008:Orig trusted trunk    0010:Orig trusted access   0020:DHCP assigned
0040:Cga authenticated     0080:Cert authenticated    0100:Statically assigned


    Network Layer Address                  Link Layer Address     Interface  vlan       prlvl      age

ARP

 10.1.201.10
```

```
0006.f601.cd43


Gi1/0/1

    201        0005       3mn        REACHABLE  86 s

<-- Gleaned from local host ARP Request
```

EVPN管理員

EVPN Mgr瞭解本地MAC並安裝到L2RIB中。EVPN Mgr也從L2RIB獲取遠端MAC，但條目僅用於處理MAC移動性

確認EVPN管理器已使用SISF條目更新

<#root>

```
Leaf01#

show l2vpn evpn mac evi 201

MAC Address    EVI   VLAN  ESI                       Ether Tag  Next Hop(s)
-------------- ----- ----- ------------------------- ---------- ---------------

0006.f601.cd43 201   201

0000.0000.0000.0000.0000 0

         Gi1/0/1:201    <-- MAC in VLan 201 local interface Gi1/0/1:service instance 201

<...snip...>
```

L2RIB

- L2RIB從EVPN管理器學習本地MAC並傳送到BGP和L2FIB。
- L2RIB還負責從BGP學習遠端MAC以更新EVPN管理器和L2FIB。
- L2RIB需要「本地」和「遠端」，其他元件才能正確更新。
- L2RIB元件位於本地和遠端MAC學習之間，具體取決於需要更新的方向/元件

驗證從EVPN管理器使用本地MAC更新了L2RIB

<#root>

```
Leaf01#

show l2route evpn mac topology 201      <-- View the overall topology for this segment

  EVI       ETag

Prod
```

```
    Mac Address                                          Next Hop(s) Seq Number
----- ---------- ----- -------------- ------------------------------------------------- ----------
  201          0
BGP
 0000.beef.cafe                                  V:20101 172.16.254.6          0
<-- produced by BGP who updated L2RIB (remote learn)
  201          0
L2VPN
 0006.f601.cd43                                          Gi1/0/1:201          0
<-- produced by EVPN Mgr who updated L2RIB (local learn)


Leaf01#
show l2route evpn mac mac-address 0006.f601.cd43 detail
EVPN Instance:            201
Ethernet Tag:             0
Producer Name:            L2VPN                <-- Produced by local
MAC Address:              0006.f601.cd43       <-- Host MAC Address
Num of MAC IP Route(s):   1
Sequence Number:          0
ESI:                      0000.0000.0000.0000.0000
Flags:                    B()
Next Hop(s):              Gi1/0/1:201 (E-LEAF) <-- Port:Instance and info about the Role (Leaf)
```

BGP

驗證BGP是否已由L2RIB更新

<#root>

```
Leaf01#
show bgp l2vpn evpn route-type 2 0 0006.f601.cd43 *
BGP routing table entry for [2][172.16.254.3:201][0][48][0006F601CD43][0][*]/20, version 268232
Paths: (1 available, best #1,
table evi_201
)
<-- In the totally isolated evi context
  Advertised to update-groups:
     2
  Refresh Epoch 1
  Local
```

```
 0.0.0.0 (via default) from 0.0.0.0

 (172.16.255.3)

<-- from 0.0.0.0 indicates local


      Origin incomplete, localpref 100, weight 32768, valid, sourced,

local

, best

<-- also indicates local


      EVPN ESI: 00000000000000000000, Label1 20101
      Extended Community: RT:65001:201 ENCAP:8

EVPN E-Tree:flag:1

,label:0

<-- EVPN e-Tree attribute with Leaf flag = 1 (added to indicate this is a host address)

      Local irb vxlan vtep:
        vrf:not found, l3-vni:0
        local router mac:0000.0000.0000
        core-irb interface:(not found)


 vtep-ip:172.16.254.3                                                    <-- Local VTEP Loopback

      rx pathid: 0, tx pathid: 0x0
      Updated on Sep 14 2023 20:16:17 UTC
```

遠端RT2學習（預設網關RT2）

BGP

驗證BGP已獲取CGW RT2字首

<#root>

```
Leaf01#

show bgp l2vpn evpn route-type 2 0 0000.beef.cafe 10.1.201.1

BGP routing table entry for [2][172.16.254.3:201][0][48][0000BEEFCAFE][32][10.1.201.1]/24, version 1141
Paths: (1 available, best #1,

table evi_201

)

<-- EVI context is 201

  Flag: 0x100
  Not advertised to any peer
  Refresh Epoch 2
  Local, imported path from [2][172.16.254.6:201][0][48][0000BEEFCAFE][32][10.1.201.1]/24 (global)
    172.16.254.6 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
```

```
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      EVPN ESI: 00000000000000000000,

Label1 20101                <-- Correct segment identifier

      Extended Community: RT:65001:201 ENCAP:8

EVPN DEF GW:0:0    <-- Default gateway attribute is added via the 'default gateway advertise CLI'

      Originator: 172.16.255.6, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Sep 1 2023 15:27:45 UTC
```

L2RIB

驗證BGP更新的L2RIB

- L2RIB從EVPN管理器學習本地MAC並傳送到BGP和L2FIB。L2RIB還負責從BGP學習遠端MAC以更新EVPN管理器和L2FIB。
- L2RIB需要「本地」和「遠端」，其他元件才能正確更新。
- L2RIB元件位於本地和遠端MAC學習之間，具體取決於需要更新的方向和元件。

<#root>

```
Leaf01#

show l2route evpn default-gateway host-ip 10.1.201.1

  EVI       ETag  Prod    Mac Address                                Host IP
----- ---------- ----- -------------- ------------------------------------ ------------------------

201

          0

BGP


0000.beef.cafe


10.1.201.1

                              V:20101 172.16.254.6

<-- L2RIB has the MAC-IP of the Gateway programmed
```

L2FIB

在L2FIB中驗證

- 負責將MAC的FED更新為在硬體中程式設計的元件。

- L2FIB安裝到FED-MATM中的遠端MAC條目不會傳送到IOS-MATM。（IOS-MATM僅顯示本地MAC，而FED-MATM同時顯示本地和遠端MAC）。
- L2FIB輸出僅顯示遠端MAC（它不負責對本地MAC進行程式設計）。

<#root>

```
Leaf01#

show l2fib bridge-domain 201 address unicast 0000.beef.cafe


MAC Address                  :

0000.beef.cafe                                        <-- CGW MAC

Reference Count              : 1
Epoch                        : 0

Producer                     : BGP                                                    <-- Learned fro

Flags                        : Static
Adjacency                    :

VXLAN_UC

  PL:2973(1) T:VXLAN_UC [MAC]20101:

172.16.254.6  <-- CGW Loopback IP

PD Adjacency                 : VXLAN_UC  PL:2973(1) T:VXLAN_UC [MAC]20101:172.16.254.6
Packets                      : 6979
Bytes                        : 0
```

FED

在FED MATM中驗證

- 在配置了「protected關鍵字」的枝葉的硬體級別，您應該只看到CGW預設網關MAC和本地主機MAC。
- 交換機檢視DEF GW屬性的RT2字首以確定哪個遠端MAC適合安裝。

<#root>

```
Leaf01#

show platform software fed switch active matm macTable vlan 201

VLAN   MAC

Type

  Seq#    EC_Bi  Flags  machandle           siHandle            riHandle            diHandle

Con

--------------------------------------------------------------------------------------------------

201    0000.beef.cafe
```

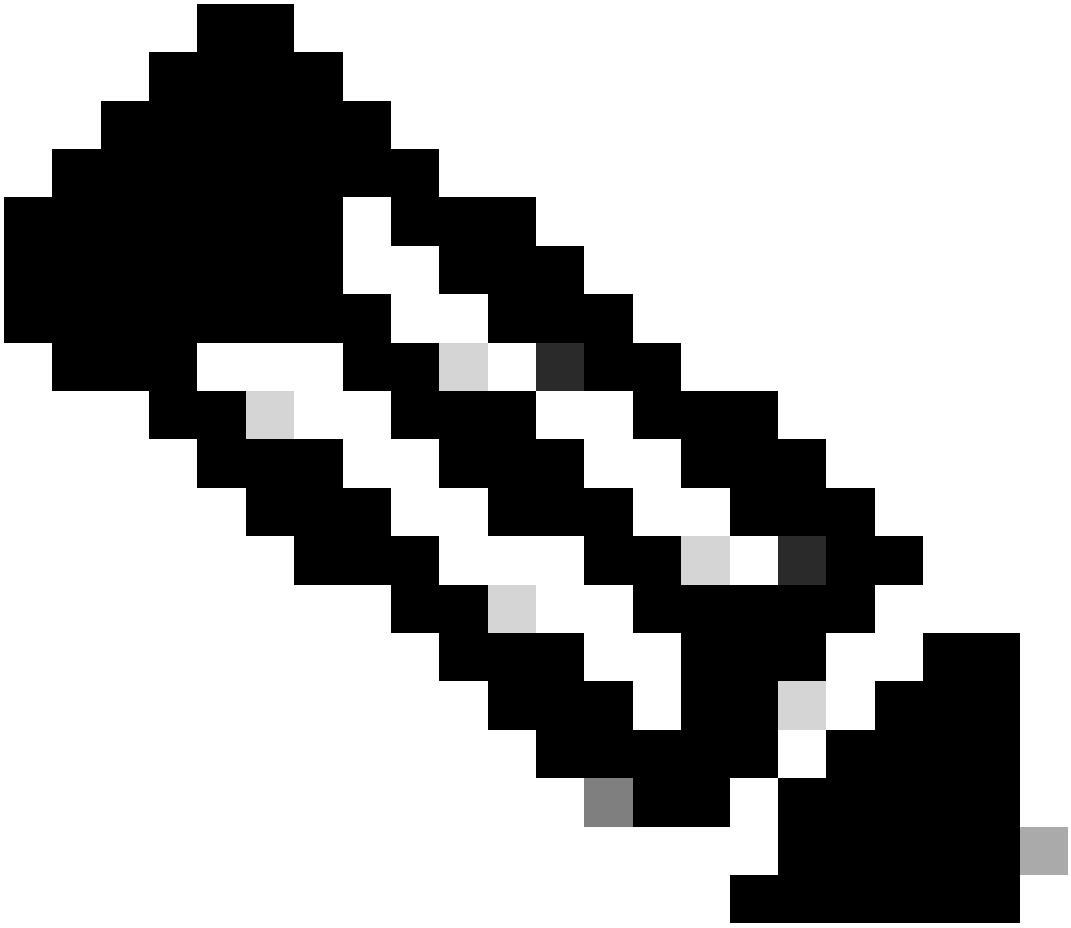
```
0x5000001

       0      0     64  0x7a199d182498      0x7a199d183578

0x71e059173e08

      0x0                          0       82

VTEP 172.16.254.6

 adj_id 9

No
<-- Only remote MAC installed in Fed is the Default Gateway (0x5000001 type) Conn = No (meaning not dire

201    0006.f601.cd01

0x1

    2458      0      0  0x7a199d1a2248      0x7a199d19eef8      0x0                 0x7a199c6f7cd8

201    0006.f601.cd43        0x1    8131      0      0  0x7a199d195a98      0x7a199d19eef8      0x0

<-- Two local MAC addresses (0x1 type) Conn = Yes (directly connected)



Total Mac number of addresses:: 5
Summary:
Total number of secure addresses:: 0
Total number of drop addresses:: 0
Total number of lisp local addresses:: 0
Total number of lisp remote addresses:: 3
*a_time=aging_time(secs)  *e_time=total_elapsed_time(secs)
Type:

MAT_DYNAMIC_ADDR           0x1

  MAT_STATIC_ADDR            0x2  MAT_CPU_ADDR               0x4  MAT_DISCARD_ADDR           0x8
MAT_ALL_VLANS             0x10  MAT_NO_FORWARD            0x20  MAT_IPMULT_ADDR           0x40  MAT_RESV
MAT_DO_NOT_AGE           0x100  MAT_SECURE_ADDR          0x200  MAT_NO_PORT              0x400  MAT_DROP
MAT_DUP_ADDR            0x1000  MAT_NULL_DESTINATION    0x2000  MAT_DOT1X_ADDR          0x4000  MAT_ROUT
MAT_WIRELESS_ADDR      0x10000  MAT_SECURE_CFG_ADDR    0x20000  MAT_OPQ_DATA_PRESENT   0x40000  MAT_WIRE
MAT_DLR_ADDR          0x100000  MAT_MRP_ADDR          0x200000  MAT_MSRP_ADDR         0x400000  MAT_LISP

MAT_LISP_REMOTE_ADDR 0x1000000

  MAT_VPLS_ADDR        0x2000000

MAT_LISP_GW_ADDR     0x4000000

<-- the addition of these values = 0x5000001

MAT_LISP_REMOTE_ADDR 0x1000000
MAT_LISP_GW_ADDR 0x4000000
MAT_DYNAMIC_ADDR 0x1
```

資料平面鄰接

確認FED條目後的最後一個步驟是解析重寫索引(RI)

<#root>

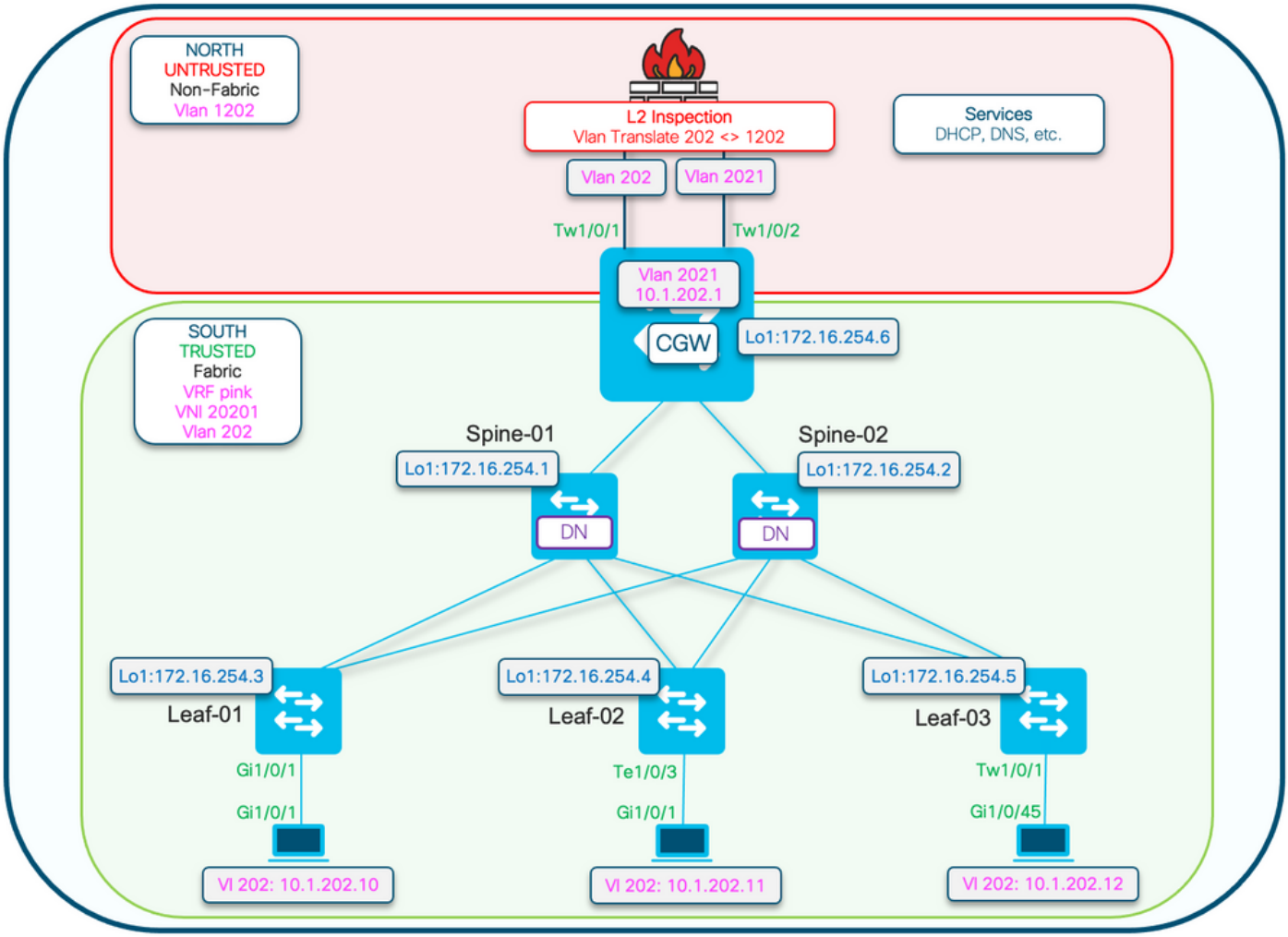
```
Leaf01#

sh platform hardware fed switch active fwd-asic abstraction print-resource-handle 0x71e059173e08 0
<-- 0x71e059173e08 is taken from previous FED command riHandle for the CGW MAC

Handle:0x71e059173e08 Res-Type:ASIC_RSC_RI Res-Switch-Num:255 Asic-Num:255 Feature-ID:AL_FID_L2_WIRELESS
priv_ri/priv_si Handle: 0x71e05917b8d8Hardware Indices/Handles: index0:0x38  mtu_index/l3u_ri_index0:0x0
Features sharing this resource:58 (1)]


Brief Resource Information (ASIC_INSTANCE# 0)
---------------------------------------
ASIC#:0 RI:56 Rewrite_type:AL_RRM_REWRITE_LVX_IPV4_L2_PAYLOAD_ENCAP_EPG(116) Mapped_rii:LVX_L3_ENCAP_L2_

 Src IP:     172.16.254.3      <-- source tunnel IP
 Dst IP:     172.16.254.6      <-- dest tunnel IP

 iVxlan dstMac:     0x9db:0x00:0x00
 iVxlan srcMac:     0x00:0x00:0x00
 IPv4 TTL:     0
 iid present:     0

 lisp iid:     20101            <-- Segment 20101

 lisp flags:     0

 dst Port:     4789             <-- VxLAN

 update only l3if:     0
 is Sgt:     0
 is TTL Prop:     0
 L3if LE:     53 (0)
 Port LE:     281 (0)
 Vlan LE:     8 (0)
```

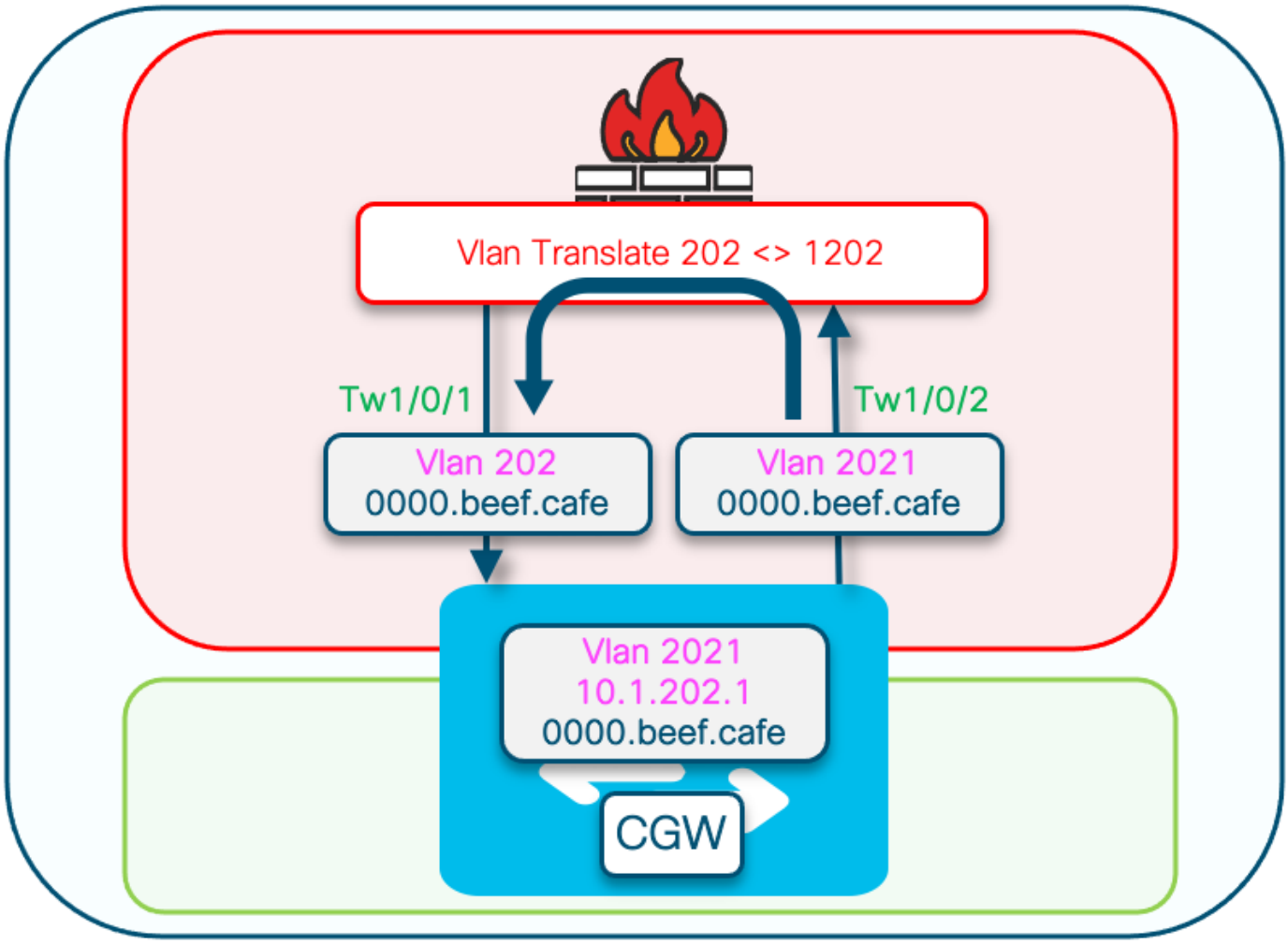


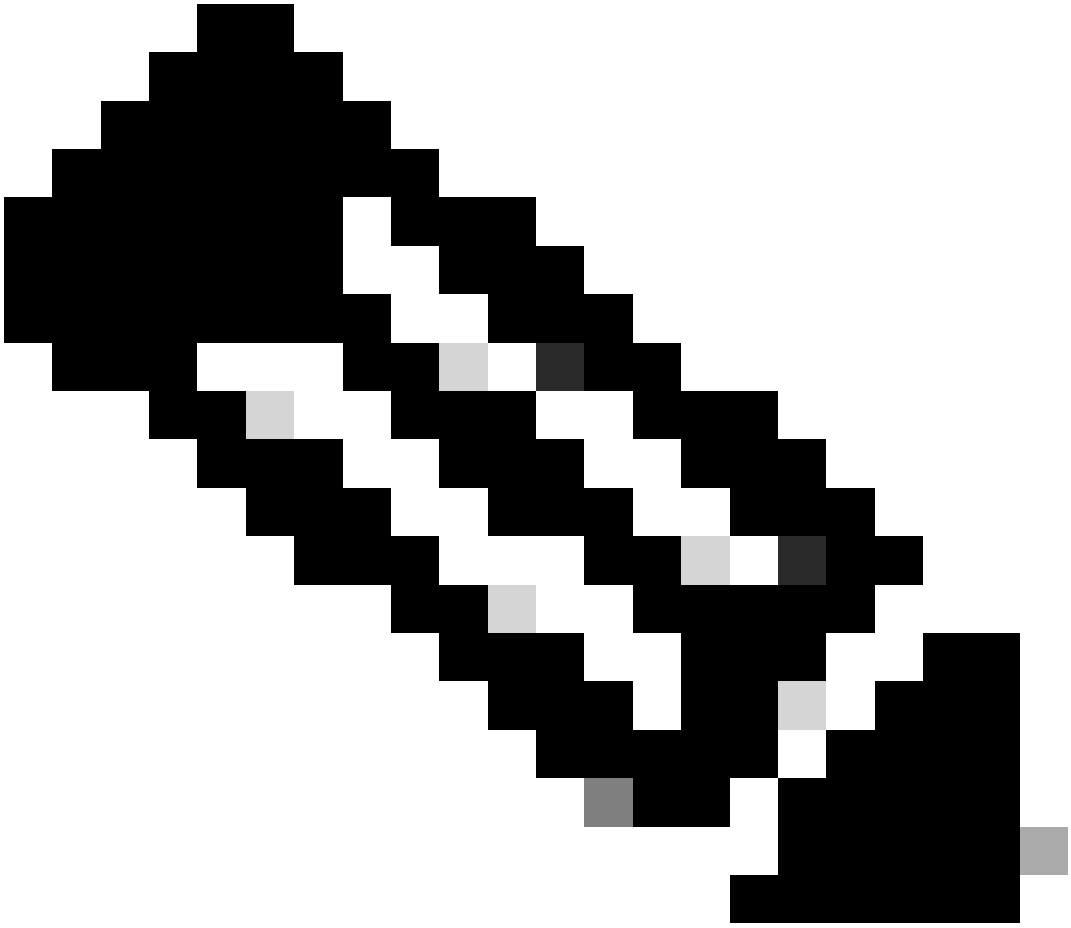
注意：您也可以使用「show platform software fed switch active matm macTable vlan 201 detail」，該命令使用FED命令將此命令連結到一個結果中

配置（部分隔離）

網路圖表


NORTH
UNTRUSTED
Non-Fabric
Vlan 1202
L2 Inspection
Vlan Translate 202 <> 1202
Services
DHCP, DNS, etc.
Vlan 202
Vlan 2021
Tw1/0/1
Tw1/0/2
Vlan 2021
10.1.202.1
CGW
Lo1:172.16.254.6
SOUTH
TRUSTED
Fabric
VRF pink
VNI 20201
Vlan 202
Spine-01
Lo1:172.16.254.1
DN
Spine-02
Lo1:172.16.254.2
DN
Lo1:172.16.254.3
Leaf-01
Lo1:172.16.254.4
Leaf-02
Lo1:172.16.254.5
Leaf-03
Gi1/0/1
Gi1/0/1
Te1/0/3
Gi1/0/1
Tw1/0/1
Gi1/0/45
Vl 202: 10.1.202.10
Vl 202: 10.1.202.11
Vl 202: 10.1.202.12


Vlan Translate 202 <> 1202
Tw1/0/1
Tw1/0/2
Vlan 202
0000.beef.cafe
Vlan 2021
0000.beef.cafe
Vlan 2021
10.1.202.1
0000.beef.cafe
CGW



注意：本部分僅介紹與完全隔離網段的區別。

- Routing-policy，用DEF GW屬性標籤GCW網關MAC IP
- 需要自定義裝置跟蹤策略以防止MAC抖動
- GW MAC IP的靜態裝置跟蹤繫結

枝葉01（基本EVPN配置）

<#root>

```
Leaf-01#

show run | sec l2vpn
l2vpn evpn

 replication-type static

 flooding-suppression address-resolution disable  <-- Disables ARP caching so ARP is always sent up to t

 router-id Loopback1
```

```
l2vpn evpn
instance 202
 vlan-based
 encapsulation vxlan

replication-type ingress
 multicast advertise enable
```

<#root>

```
Leaf01#
show run | sec vlan config

vlan configuration 202
 member evpn-instance 202 vni 20201
protected <-- protected keyword added
```

CGW（基本配置）

在nve下設定復制模式

<#root>

```
CGW#
show run int nve 1

Building configuration...

Current configuration : 313 bytes
!
interface nve1
 no ip address
 source-interface Loopback1
 host-reachability protocol bgp
 member vni 20201 ingress-replication local-routing <-- 'ingress-replication' (Unicast all BUM traffic)
end
```

配置外部網關SVI

<#root>

```
CGW#
show run interface vlan 2021
```

```
Building configuration...

Current configuration : 231 bytes
!
interface Vlan2021

mac-address 0000.beef.cafe          <-- MAC is static in this example for viewing simplicity. This is no

 vrf forwarding pink                  <-- SVI is in VRF pink

 ip address 10.1.202.1 255.255.255.0
 no ip redirects

 ip local-proxy-arp                 <-- Sets CGW to Proxy reply even for local subnet ARP requests

 ip pim sparse-mode

 ip route-cache same-interface      <-- This is auto added when local-proxy-arp is configured. However,

 ip igmp version 3
 no autostate
end
```

建立停用收集功能的策略

```
<#root>

device-tracking policy dt-no-glean

 <-- Configure device tracking policy to prevent MAC-IP flapping

 security-level glean
 no protocol ndp
 no protocol dhcp6
 no protocol arp
 no protocol dhcp4
```

連線到externalgatewayevi/vlan

```
<#root>

CGW#

show running-config | sec vlan config

vlan configuration 202
 member evpn-instance 202 vni 20201

 device-tracking attach-policy dt-no-glean <-- apply the new device tracking policy to the vlan configur
```

將靜態條目增加到externalgateway mac-ip的裝置跟蹤表中

```
<#root>
```

```
device-tracking binding vlan 202 10.1.202.1 interface TwentyFiveGigE1/0/1 0000.beef.cafe
```

<-- All static entries in device tracking table should be for external gateway mac-ip's.
If there is any other static entry in device tracking table, match ip/ipv6 configurations in route map

建立BGP路由對映以匹配RT2 MAC-IP字首並設定預設網關extendedcommunity

<#root>

```
route-map CGW_DEF_GW permit 10
```

match evpn route-type 2-mac-ip <-- match RT2 type MAC-IP

set extcommunity default-gw <-- Set Default-gateway (DEF GW 0:0) extended community

```
route-map CGW_DEF_GW permit 20
```

將路由對映應用到BGP路由反射器鄰居

<#root>

CGW#

sh run | s r bgp

```
address-family l2vpn evpn
 neighbor 172.16.255.1 activate
 neighbor 172.16.255.1 send-community both
 neighbor 172.16.255.1
```

route-map CGW_DEF_GW out <-- Sets the DEF GW Community when it advertises MAC-IP type RT2 to the RR

```
 neighbor 172.16.255.2 activate
 neighbor 172.16.255.2 send-community both
 neighbor 172.16.255.2
```

route-map CGW_DEF_GW out <-- Sets the DEF GW Community when it advertises MAC-IP type RT2 to the RR

驗證（部分隔離）

EVI詳細資料

<#root>

```
Leaf01#

show l2vpn evpn evi 202 detail

EVPN instance:       202 (VLAN Based)
  RD:                172.16.254.3:202 (auto)
  Import-RTs:        65001:202
  Export-RTs:        65001:202
  Per-EVI Label:     none
  State:             Established
  Replication Type:  Ingress
  Encapsulation:     vxlan
  IP Local Learn:    Enabled (global)
  Adv. Def. Gateway: Enabled (global)
  Re-originate RT5:  Disabled
  Adv. Multicast:    Enabled

  Vlan:              202
    Protected:       True (local access p2p blocked)  <-- Vlan 202 is in protected mode

<...snip...>
```

本地RT2生成（本地主機到RT2）

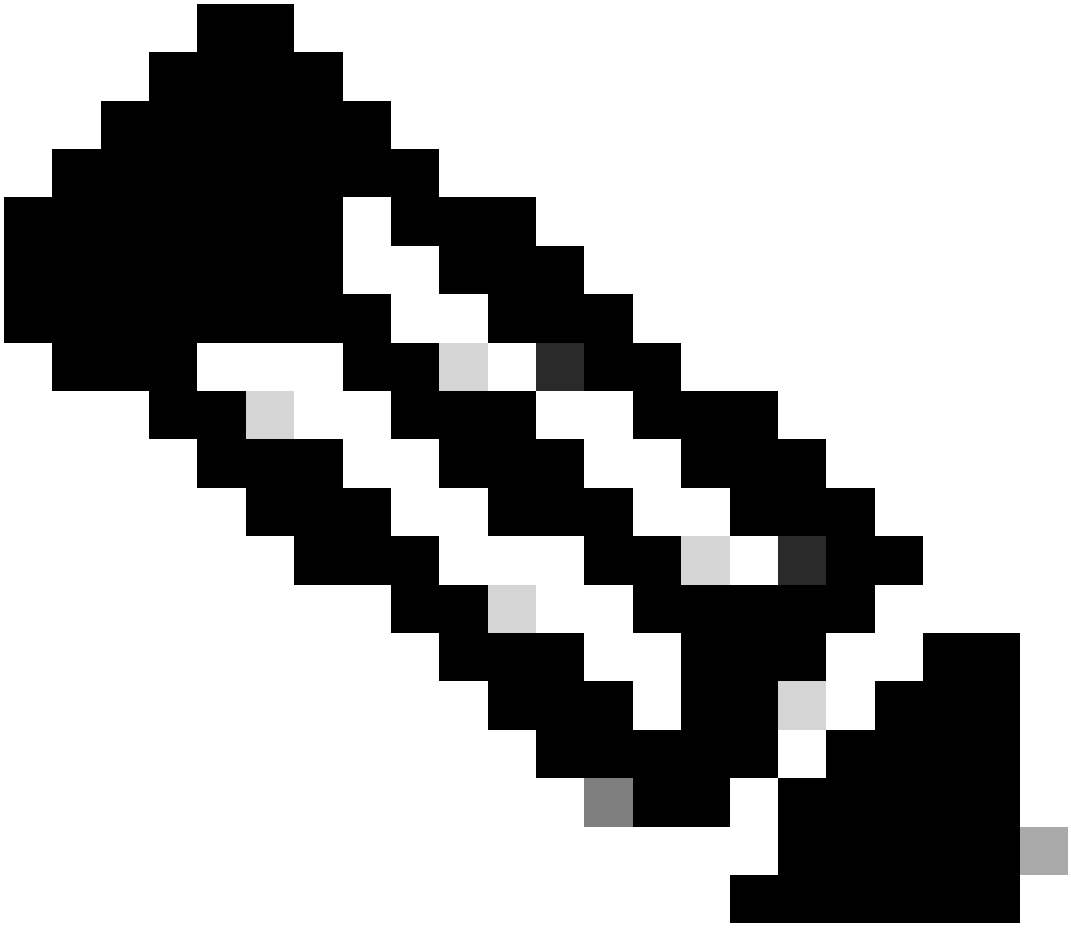
涵蓋在前一個完全隔離的示例中

遠端RT2學習（預設網關RT2）

涵蓋與「完全隔離」的不同之處

CGW預設網關字首（枝葉）

檢查字首是否具有適當的屬性，以便可以安裝到硬體中



注意：這對於DHCP L2中繼的正常運行至關重要

<#root>

Leaf01#

show bgp l2vpn evpn route-type 2 0 0000.beef.cafe 10.1.202.1

BGP routing table entry for [2][172.16.254.3:202][0][48][0000BEEFCAFE][32][10.1.202.1]/24, version 1846
Paths: (1 available, best #1,

table evi_202

)

<-- the EVI context of 202 which matches the Vlan/EVI we are concerned about

  Not advertised to any peer
  Refresh Epoch 2
  Local, imported path from [2][172.16.254.6:202][0][48][0000BEEFCAFE][32][10.1.202.1]/24 (global)
    172.16.254.6 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, best

```
      EVPN ESI: 00000000000000000000,

Label1 20201              <-- Correct Segment ID

      Extended Community: RT:65001:202 ENCAP:8

EVPN DEF GW:0:0  <-- prefix has the Default GW attribute added

      Originator: 172.16.255.6, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Sep 7 2023 19:56:43 UTC
```

FED MATM（分葉）

<#root>

```
F241.03.23-9300-Leaf01#

show platform software fed active matm macTable vlan 202 mac 0000.beef.cafe

VLAN   MAC                  Type  Seq#    EC_Bi  Flags  machandle           siHandle            riHandl
-------------------------------------------------------------------------------------------------------

202    0000.beef.cafe

  0x5000001       0      0     64  0x71e058da7858      0x71e05916c0d8      0x71e059171678      0x0

VTEP 172.16.254.6

 adj_id 651

No
<-- MAC of Default GW is installed in FED
```

SISF (CGW)

<#root>

```
CGW#

sh device-tracking database vlanid 202

vlanDB has 1 entries for vlan 202, 0 dynamic
Codes: L - Local, S - Static, ND - Neighbor Discovery, ARP - Address Resolution Protocol, DH4 - IPv4 DH
Preflevel flags (prlvl):
0001:MAC and LLA match     0002:Orig trunk            0004:Orig access
0008:Orig trusted trunk    0010:Orig trusted access   0020:DHCP assigned
0040:Cga authenticated     0080:Cert authenticated    0100:Statically assigned


    Network Layer Address                  Link Layer Address     Interface  vlan       prlvl      ag
S   10.1.202.1                             0000.beef.cafe         Twe1/0/1   202        0100       13
```

IOS MATM (CGW)

<#root>

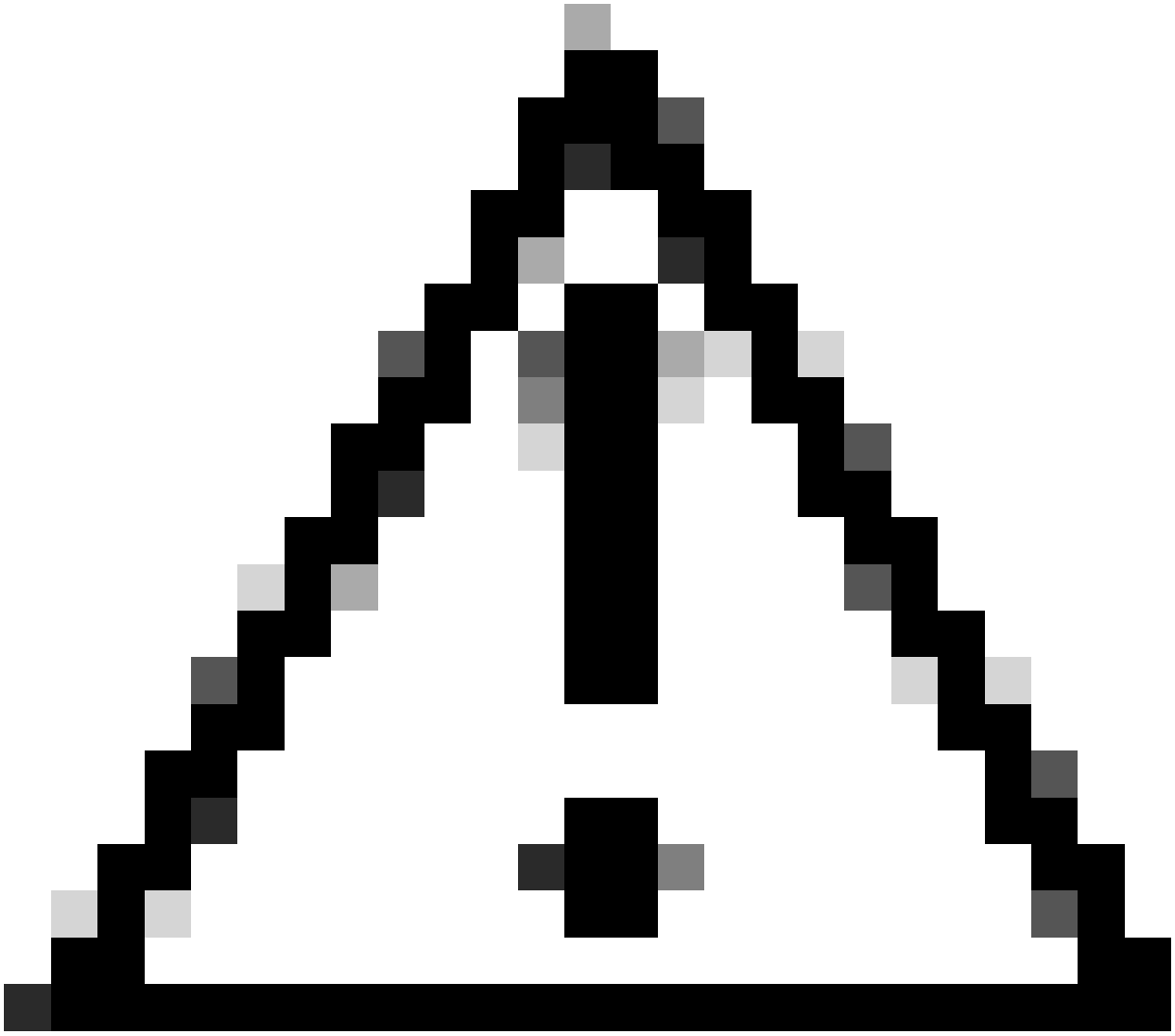
```
CGW#

show mac address-table address 0000.beef.cafe

          Mac Address Table
-------------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       --------    -----
 201    0000.beef.cafe    STATIC      Vl201

2021    0000.beef.cafe    STATIC      Vl2021   <-- The Vlan 2021 SVI MAC advertised out Tw1/0/1

 202    0000.beef.cafe    DYNAMIC     Twe1/0/1 <-- The Vlan 2021 SVI MAC learned dynamically after passi
```

疑難排解

位址解析(ARP)

隔離ARP問題的一般步驟

- 確認IMET通道已就緒
- 在CGW上行鏈路上捕獲以驗證從枝葉封裝的ARP接收
- 如果沒有ARP到達上行鏈路上的封裝
  - 驗證枝葉和CGW上的IMET隧道是否已就緒。
  - 在枝葉上行鏈路上捕獲，確認ARP已封裝並傳送。
  - 排除中繼路徑故障
- 如果ARP到達邊界IMET隧道捕獲，但未在VRF ARP表中程式設計。
  - 排除CPU/CoPP傳送路徑故障以確認傳送到CPU的ARP
  - 確認IP地址/客戶端資訊正確。
  - 調試VRF中的ARP以檢視可能影響ARP進程的因素
- 驗證主機上的CGW MAC是否安裝為下一跳/目標mac
- 確認CGW具有實際主機MAC的兩個ARP條目
- 驗證防火牆策略是否允許此類流量



注意：啟用調試時請務必小心！

確保您已停用泛洪抑制

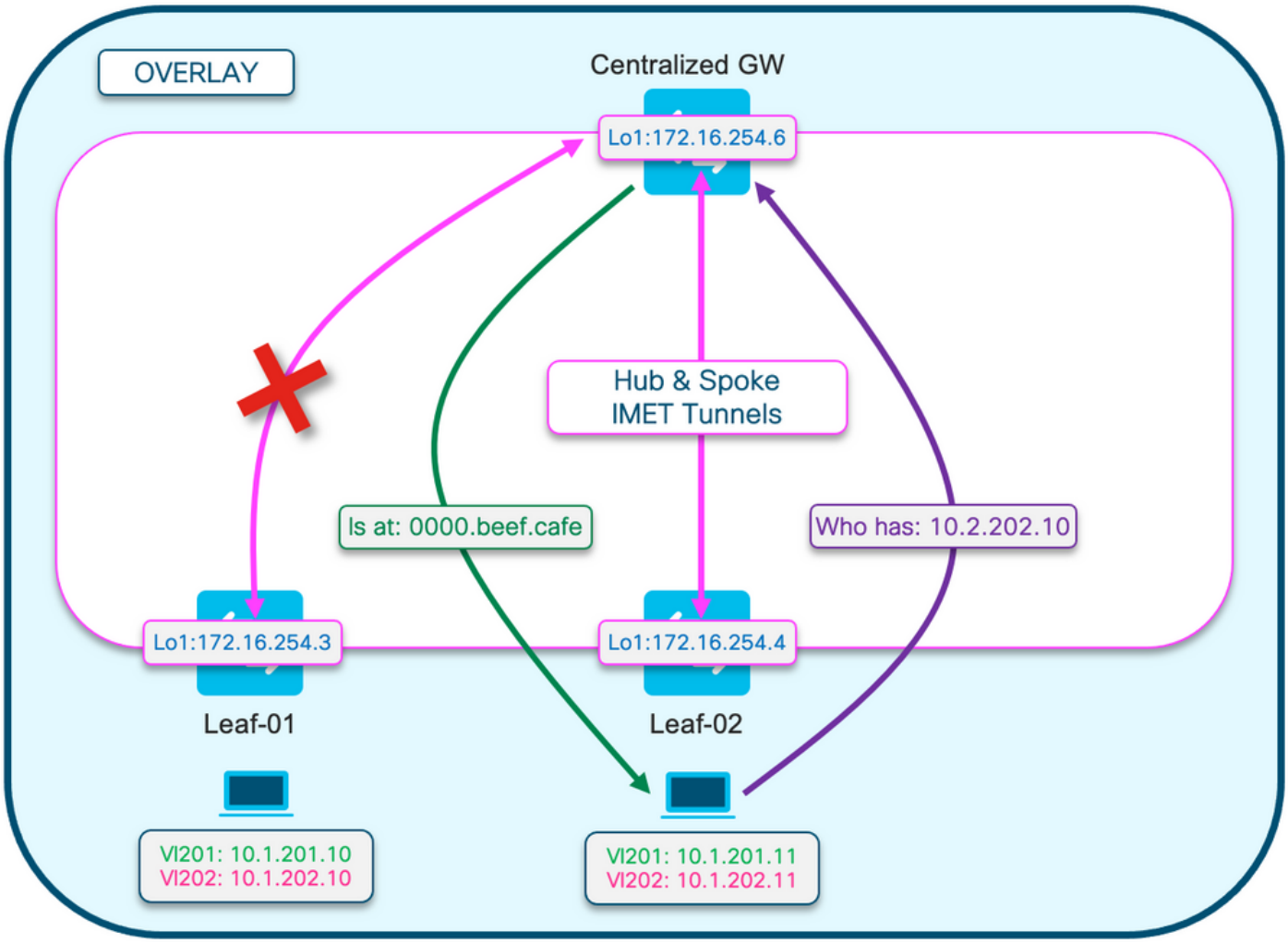
<#root>

```
Leaf-01#

show run | sec l2vpn
l2vpn evpn

 replication-type static

 flooding-suppression address-resolution disable <-- This CLI prevents a VTEP from trying to unicast oth
```

當枝葉-02上的主機解析枝葉-01上的主機的ARP時，ARP請求不會直接廣播到枝葉-01

- 相反，ARP會向上傳遞在Leaf-02上程式設計的唯一BUM隧道以指向CGW

- CGW不會將此消息轉發到Leaf-01，而是使用自己的MAC進行應答
- 這會導致所有通訊向上傳遞到CGW，然後路由到主機之間
- CGW路由資料包，即使它們位於同一本地子網中



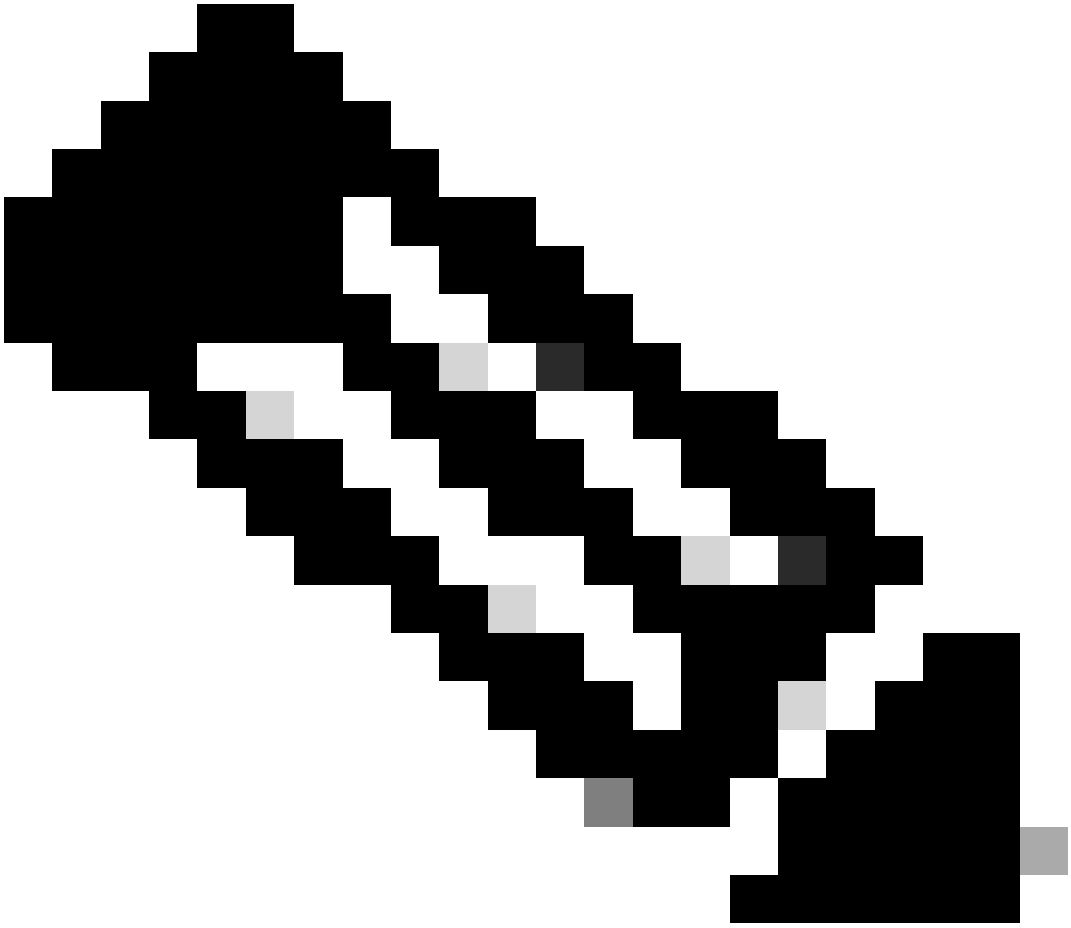
此圖有助於直觀顯示本部分中介紹的ARP解析過程流程。

ARP請求顯示為紫色

- 此ARP請求用於解析主機10.1.202.10 off Leaf-01的MAC地址
- 請注意，紫線在CGW處終止，並且未到達Leaf-01

ARP應答顯示為綠色

- 回覆包含用於Vlan 202的CGW SVI的MAC
- 請注意，綠線來自CGW，而不是來自實際主機



注意：紅色X表明此通訊不涉及向Leaf-01傳送資料流。

觀察每台相應主機上的ARP條目

<#root>

Leaf02-HOST#

sh ip arp 10.1.202.10

Protocol Address   Age (min) Hardware Addr   Type   Interface
Internet 10.1.202.10   1

0000.beef.cafe

  ARPA   Vlan202

<-- MAC address for Leaf01 host is CGW MAC

Leaf01-HOST#

sh ip arp 10.1.202.11

```
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  10.1.202.11             7
```

0000.beef.cafe

```
  ARPA   Vlan202
```

<-- MAC address for Leaf02 host is CGW MAC

觀察在CGW上如何獲取RT2字首。這是CGW路由資料包所必需的

<#root>

CGW#

sh bgp l2vpn evpn route-type 2 0 0006.f617.eec4 * <-- Leaf02 actual MAC

BGP routing table entry for [2][172.16.254.6:202][0][48][0006F617EEC4][0][*]/20, version 235458
Paths: (1 available, best #1,

table evi_202

)

```
  Not advertised to any peer
  Refresh Epoch 2
  Local, imported path from [2][172.16.254.4:202][0][48][0006F617EEC4][0][*]/20 (global)
    172.16.254.4 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      EVPN ESI: 00000000000000000000,
```

Label1 20201 <-- correct segment identifier

```
      Extended Community: RT:65001:202 ENCAP:8
```

EVPN E-Tree:flag:1

,label:0

<-- prefix contains the Leaf flag indicating this is a normal host

```
      Originator: 172.16.255.4, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Apr 9 2025 17:11:22 UTC
```

CGW#

sh bgp l2vpn evpn route-type 2 0 0006.f601.cd44 * <-- Leaf01 actual MAC

BGP routing table entry for [2][172.16.254.6:202][0][48][0006F601CD44][0][*]/20, version 235521
Paths: (1 available, best #1,

table evi_202)

```
  Not advertised to any peer
  Refresh Epoch 2
  Local, imported path from [2][172.16.254.3:202][0][48][0006F601CD44][0][*]/20 (global)
    172.16.254.3 (metric 3) (via default) from 172.16.255.1 (172.16.255.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      EVPN ESI: 00000000000000000000,
```

```
Label1 20201                          <-- correct segment identifier

      Extended Community: RT:65001:202 ENCAP:8

EVPN E-Tree:flag:1

,label:0

<-- prefix contains the Leaf flag indicating this is a normal host


      Originator: 172.16.255.3, Cluster list: 172.16.255.1
      rx pathid: 0, tx pathid: 0x0
      Updated on Apr 9 2025 17:17:06 UTC
```

捕獲上行鏈路上的ARP交換以確認雙向通訊

- 您可以在交換矩陣上行鏈路上使用嵌入式資料包捕獲(EPC)
- 此場景顯示Leaf01上行鏈路上的EPC。如有必要，在CGW上重複此相同過程

配置EPC

```
<#root>

Leaf01#

monitor capture 1 interface range te 1/1/2 , te 1/1/4 both match any buffer size 100


<-- both Uplinks toward fabric included
```

開始捕獲

```
<#root>

Leaf01#

monitor capture 1 start
```

啟動ping以觸發ARP請求（在本例中，ping是從Leaf01主機10.1.201.10到Leaf02主機10.1.201.11）

```
<#root>

Leaf01-HOST#

ping vrf red 10.1.201.11


Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.1.201.11, timeout is 2 seconds:
```

```
...!!
Success rate is 40 percent (2/5), round-trip min/avg/max = 1/1/1 ms
```

停止捕獲並檢查ARP幀

<#root>

```
Leaf01#
mon cap 1 stop

F241.03.23-9300-Leaf01#
show mon cap 1 buff br | i ARP

11
 8.153510 00:06:f6:01:cd:42 -> ff:ff:ff:ff:ff:ff ARP 110
Who has 10.1.201.11? Tell 10.1.201.10  <-- .10 requests .11 MAC (this is Frame 11)

12 8.154030 00:00:be:ef:ca:fe -> 00:06:f6:01:cd:42 ARP 110 10.1.201.11
is at 00:00:be:ef:ca:fe    <-- CGW replies with its MAC
```

檢視捕獲資料包的詳細資訊。如果要檢視有關資料包的更多資訊，請使用EPC的detail選項

- 請注意，為了簡潔起見，此輸出被裁剪到不同的位置

<#root>

```
Leaf01#
show mon cap 1 buffer detailed | beg Frame 11   <-- begin detail result from Frame 11 (ARP Request)
Frame 11: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface /tmp/epc_ws/wif_to_t
Ethernet II, Src: 00:00:00:00:00:00 (00:00:00:00:00:00), Dst: 00:00:00:00:00:00 (00:00:00:00:00:00)
    Destination: 00:00:00:00:00:00 (00:00:00:00:00:00)
        Address: 00:00:00:00:00:00 (00:00:00:00:00:00)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
    Source: 00:00:00:00:00:00 (00:00:00:00:00:00)
        Address: 00:00:00:00:00:00 (00:00:00:00:00:00)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
    Type: IPv4 (0x0800)
Internet Protocol Version 4, Src: 172.16.254.3, Dst: 172.16.254.6     <--- Outer tunnel IP header
```

```
    Source: 172.16.254.3
    Destination: 172.16.254.6
User Datagram Protocol, Src Port: 65483,
Dst Port: 4789  <-- VXLAN Dest port


Virtual eXtensible Local Area Network
    VXLAN Network Identifier
(VNI): 20101               <-- Verify the VNI for the segment you are investigating

    Reserved: 0
Ethernet II, Src: 00:06:f6:01:cd:42 (00:06:f6:01:cd:42), Dst: ff:ff:ff:ff:ff:ff (ff:ff:ff:ff:ff:ff)  <--



  Type: ARP (0x0806)

    Trailer: 00000000000000000000000000000000000000
Address Resolution Protocol (
request
)
 <-- is an ARP request


    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: request (1)


Sender MAC address: 00:06:f6:01:cd:42 (00:06:f6:01:cd:42)   <-- Sending host
    Sender IP address: 10.1.201.10
    Target MAC address: 00:00:00:00:00:00 (00:00:00:00:00:00)   <-- Trying to resolve MAC for host
    Target IP address: 10.1.201.11



Frame 12:
 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface /tmp/epc_ws/wif_to_ts_pipe, i
<-- ARP reply


Ethernet II,
Src: dc:77:4c:8a:6d:7f
 (dc:77:4c:8a:6d:7f),
Dst: 68:2c:7b:f8:87:48
 (68:2c:7b:f8:87:48)
<-- Underlay MACs



Internet Protocol Version 4, Src: 172.16.254.6, Dst: 172.16.254.3
User Datagram Protocol, Src Port: 65410, Dst Port: 4789
Virtual eXtensible Local Area Network
```

```
    VXLAN Network Identifier (VNI): 20101
    Reserved: 0
Ethernet II,
Src: 00:00:be:ef:ca:fe
 (00:00:be:ef:ca:fe),
Dst: 00:06:f6:01:cd:42
 (00:06:f6:01:cd:42)
 <-- Start of payload



 Type: ARP
 (0x0806)
    Trailer: 00000000000000000000000000000000000000
Address Resolution Protocol (
reply
)
<-- is an ARP reply


    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: reply (2)


 Sender MAC address: 00:00:be:ef:ca:fe (00:00:be:ef:ca:fe)   <-- Reply is that of the CGW MAC due to loc
    Sender IP address: 10.1.201.11
    Target MAC address: 00:06:f6:01:cd:42 (00:06:f6:01:cd:42)
    Target IP address: 10.1.201.10
```

CGW RT2網關字首

缺少網關字首

如前面關於部分隔離網段的部分中所述，需要在交換矩陣Vlan中獲取MAC

- 如果沒有超過MAC老化計時器的流量發往網關，則可能出現此問題。
- 如果缺少CGW網關字首，您需要確認存在MAC

```
<#root>
CGW#
show bgp l2vpn evpn route-type 2 0 0000.beef.cafe 10.1.202.1
% Network not in table <-- RT2 not generated on CGW
```

```
CGW#

show mac address-table address 0000.beef.cafe

          Mac Address Table
----------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       --------    -----
 201    0000.beef.cafe    STATIC      Vl201
2021    0000.beef.cafe    STATIC      Vl2021

<-- MAC is not learned in Fabric Vlan 202

Total Mac Addresses for this criterion: 2
```

網關字首缺失補救

在大多數生產網路中，隨時都可能存在一些流量。但是，如果您遇到此問題，可以使用以下選項之一來修正此問題：

- 增加靜態MAC條目，例如「mac address-table static 0000.beef.cafe vlan 202 interface TwentyFiveGigE1/0/1」
- 使用「mac address-table aging-time <seconds>」增加MAC老化計時器。（請記住，這會增加所有MAC地址的老化時間，因此首選靜態MAC選項）

缺少DEF GW屬性

對於部分隔離網段，有一些額外的配置可增加此屬性。

缺少DEF GW屬性修正

確認以下詳細資訊：

- 您運行的是17.12.1或更高版本
- 配置中存在SISF（裝置跟蹤） CLI
- 配置了route-map match & set命令，並將路由對映應用到BGP鄰居
- 您已刷新BGP通告(必須清除BGP才能使用新屬性重新通告字首

無線漫遊

頻繁漫遊可能導致BGP更新過於頻繁，在交換機宣告它擁有MAC並傳送RT2更新之前，應增加每個時間間隔的漫遊

- 當主機在不同交換機上的兩個AP之間移動時，就會發生這種情況。
- 漫遊的預設限制是每180秒5次

```
<#root>

Leaf01#

sh run | sec l2vpn
```

```
l2vpn evpn
 replication-type static
 flooding-suppression address-resolution disable
 ip duplication limit 10 time 180          <--- You can adjust this default in the global l2vpn section
 mac duplication limit 10 time 180


Leaf01#
sh l2vpn evpn summary

L2VPN EVPN
  EVPN Instances (excluding point-to-point): 4
    VLAN Based:   4
  Vlans: 4
  BGP: ASN 65001, address-family l2vpn evpn configured
  Router ID: 172.16.254.3
  Global Replication Type: Static
  ARP/ND Flooding Suppression: Disabled
  Connectivity to Core: UP
  MAC Duplication: seconds 180 limit 10

  MAC Addresses: 13
    Local:     6
    Remote:    7
    Duplicate: 0
  IP Duplication: seconds 180 limit 10

  IP Addresses: 7
    Local:     4
    Remote:    3
    Duplicate: 0

<...snip...>
```

要為TAC收集的命令

如果本指南未能解決您的問題，請收集顯示的命令清單，並將其附加到TAC服務請求中。

要收集的最少資訊

（在重新載入/復原動作之前收集資料的時間有限）。

- Show tech evpn
- Show tech
- Show tech sisf

要收集的詳細資訊

(如果有時間收集更完整的資料，則這是首選)

- show tech
- show tech evpn
- show tech platform evpn_vxlan switch <number>
- show tech platform
- show tech resource
- show tech sisf
- show tech isis
- show tech bgp
- show monitor event-trace evpn event all
- show monitor event-trace evpn error all
- 請求平台軟體跟蹤存檔

相關資訊

- 在Catalyst 9000系列交換器上實作BGP EVPN路由原則
- DHCP第2層中繼（即將推出）

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）。