基於Carmel ASIC的交換機(Nexus 5548/5596)中的vPC對等鏈路上的L2MP轉發

目錄

簡介

在vPC拓撲中，只有孤立埠流量或泛洪流量（未知的單播、廣播、組播）在對等鏈路上可見使用者流量。 對於這種泛洪流量，交換機需要確保在vPC的一個分支上接收的泛洪流量不會在另一個vPC分支上返回，以便資料包不會發回源或複製到其他vPC。

在基於Carmel的交換機(Nexus 55xx)中，vPC環路避免實施與基於Gatos(Nexus 5010/5020)的實施不同，後者針對對等鏈路上的泛洪流量使用單獨的內部MCT VLAN。

由於基於Carmel的交換機支援L2MP或fabricpath，因此工程部門決定使用基於L2MP的對等鏈路轉發。在此型號中，vPC主交換機的交換機ID為2748(0xabc)，而vPC輔助交換機的交換機ID為2749(0xabd)。 2750(0xabe)的模擬交換機ID將用作輸入vPC但通過對等鏈路傳送的幀的源交換機ID。vPC主交換機上的所有埠將是FTAG 256的成員，vPC輔助交換機上的所有埠將是FTAG 257的成員。在vPC主交換機中，只有孤立埠將是FTAG 257的成員，而在vPC輔助交換機中，孤立埠將是FTAG 256的成員。

必要條件

需求

本文件沒有特定需求。

採用元件

本文件所述內容不限於特定軟體和硬體版本。

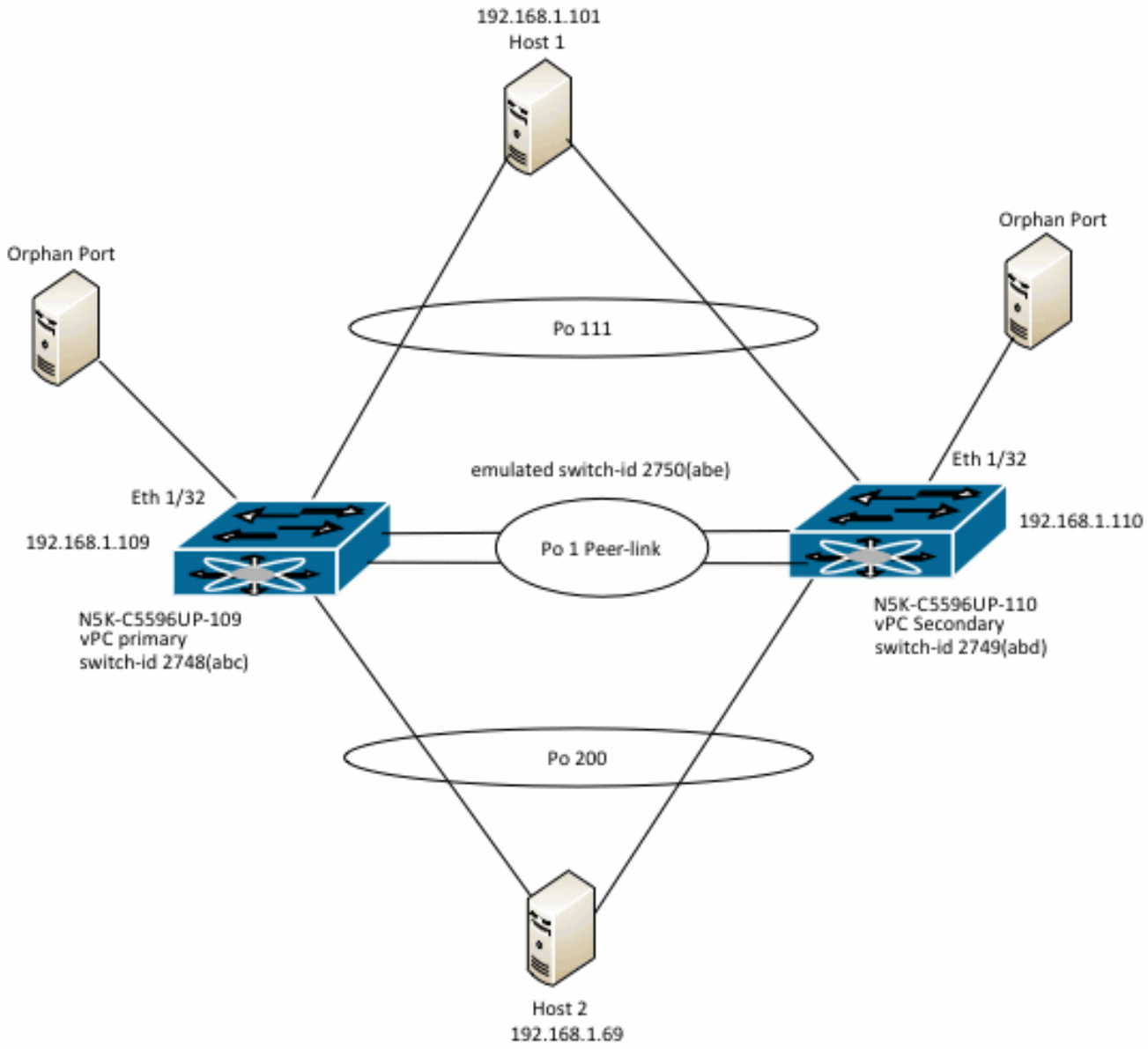
慣例

如需文件慣例的詳細資訊，請參閱思科技術提示慣例。

環路避免

對於進入vPC主交換機的廣播/未知單播/組播幀，將通過對等鏈路以256的FTAG傳送出去。當vPC輔助交換機通過vPC對等鏈路獲得此幀時，它會檢查FTAG，並且自其256以來，vPC輔助交換機只將其傳送到FTAG 256成員，該成員將僅是孤立埠。對於來自vPC輔助裝置的泛洪流量，將使用FTAG 257傳送該泛洪流量；當vPC主交換機收到此幀時，僅將收到的泛洪幀傳送到FTAG 257的成員，該成員將僅作為孤立埠。基於Carmel的交換機就是這樣實現vPC環路避免的。

為了深入研究基於L2MP/FTAG的泛洪幀在對等鏈路上的轉發，使用以下拓撲：



N5K-C5596UP-109和N5K-C5596UP-100是運行NX-OS 5.2(1)N1(2a)的Nexus 5596交換機的vPC對。N5K-C5596UP-109是vPC主交換機，N5K-C5596UP-110是vPC輔助交換機。埠通道1是vPC對等鏈路。圖中所示的IP地址屬於交換機的介面VLAN 1。主機1和主機2是通過VLAN 1中的vPC連線的Cisco交換機。在本文檔中，這些交換機稱為主機1和主機2。VLAN 1中有孤立埠連線到兩台交換機的Eth1/32。

以下是交換器的一些指令輸出：

```
N5K-C5596UP-109# show vpc
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 2
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
Peer gateway excluded VLANs       : -
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status  Active vlans
-- ---- ------ --------------------------------------------------
1    Po1    up      1

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
111    Po111       up     success     success                    1
200    Po200       up     success     success                    1



N5K-C5596UP-109# show platform fwm info l2mp myswid

switch id
---------------------------------------------------------------------

switch id manager
  -------------------------------------------------------------------
    vpc role: 0
    my primary switch id: 2748 (0xabc)
    emu switch id: 2750 (0xabe)
    peer switch id: 2749 (0xabd)
N5K-C5596UP-109# show vpc orphan-ports
Note:
--------::Going through port database. Please be patient.::--------

VLAN            Orphan Ports
-------         -------------------------
1               Eth1/32




N5K-C5596UP-110# show vpc

Legend:
                (*) - local vPC is down, forwarding via vPC peer-link
```

```
vPC domain id                     : 2
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
Per-vlan consistency status       : success
Type-2 consistency status         : success
vPC role                          : secondary
Number of vPCs configured         : 2
Peer Gateway                      : Enabled
Peer gateway excluded VLANs       : -
Dual-active excluded VLANs        : -
Graceful Consistency Check        : Enabled
Auto-recovery status              : Disabled
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status  Active vlans
-- ---- ------ --------------------------------------------------
1    Po1    up      1

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                     Active vlans
------ ----------- ------ ----------- -------------------------- -----------
111    Po111       up     success     success                    1
200    Po200       up     success     success                    1



N5K-C5596UP-110# show platform fwm info l2mp myswid

switch id
---------------------------------------------------------------------
  switch id manager
  -------------------------------------------------------------------
    vpc role: 1
    my primary switch id: 2749 (0xabd)
    emu switch id: 2750 (0xabe)
    peer switch id: 2748 (0xabc)
N5K-C5596UP-110# show vpc orphan-ports
Note:
--------::Going through port database. Please be patient.::--------

VLAN            Orphan Ports
-------         -------------------------
1               Eth1/32



Now lets check on default FTAGs used and its members.

N5K-C5596UP-109# show platform fwm info l2mp ftag all
L2MP FTAG
---------------------------------------------------------------------

ftag[0x9565b1c] id: 256 (0x100)
Topology ID: 0x111
Ftag flags: 0 (invalid ftag-flags)
Is stale: FALSE
ftag_mask[0x973eca4]
ifindex array:
0x160000c7 0x1600006e 0x1a01f000
0x15010000 0x15020000 0x1600007e
0x16000000
```

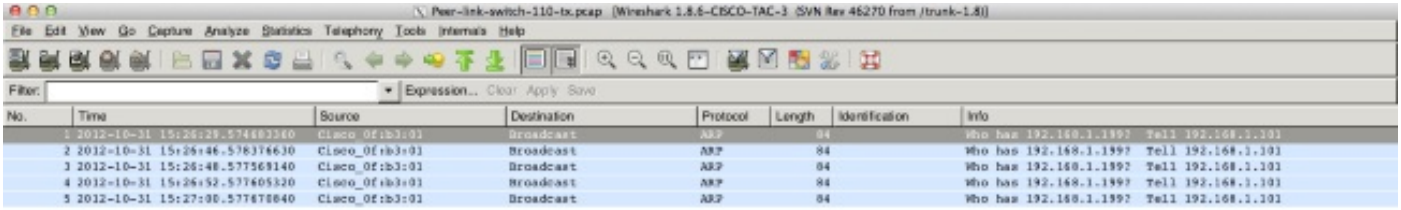
```
ifmap[0x88400fc]
ifmap idx 6: ref 1, lu_mcq_alloced 0, lu_mcq 15 (orig 15) 'not pruned'
ifmap idx 6: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 6: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 6: ifs - sup-eth1 sup-eth2 Po200 Po1 Po111 Eth1/32 Po127
rpf: (0x0)
alternate: 0
intf:
Po1 (0x16000000)
ftag_ucast_index: 1
ftag_flood_index: 1
ftag_mcast_index: 32
ftag_alt_mcast_index: 48
------------------------------------------------------------------
ftag[0x9565e3c] id: 257 (0x101)
Topology ID: 0x111
Ftag flags: 0 (invalid ftag-flags)
Is stale: FALSE
ftag_mask[0x95612b4]
ifindex array:
0x1a01f000 0x15010000 0x15020000
0x16000000
ifmap[0x883b81c]
ifmap idx 11: ref 1, lu_mcq_alloced 0, lu_mcq 14 (orig 14) 'not pruned'
ifmap idx 11: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 11: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 11: ifs - sup-eth1 sup-eth2 Po1 Eth1/32
rpf: (0x0)
alternate: 1
intf:
Po1 (0x16000000)
ftag_ucast_index: 0
ftag_flood_index: -1
ftag_mcast_index: 0
ftag_alt_mcast_index: 0
------------------------------------------------------------------
N5K-C5596UP-109#

N5K-C5596UP-110# show platform fwm info l2mp ftag all
L2MP FTAG
------------------------------------------------------------------
ftag[0x956a99c] id: 256 (0x100)
Topology ID: 0x111
Ftag flags: 0 (invalid ftag-flags)
Is stale: FALSE
ftag_mask[0x98b4764]
ifindex array:
0x16000066 0x1a01f000 0x15010000
0x15020000 0x16000000
ifmap[0x9635adc]
ifmap idx 4: ref 1, lu_mcq_alloced 0, lu_mcq 15 (orig 15) 'not pruned'
ifmap idx 4: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 4: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 4: ifs - sup-eth1 sup-eth2 Po103 Po1 Eth1/32
rpf: (0x0)
alternate: 1
intf:
Po1 (0x16000000)
ftag_ucast_index: 1
ftag_flood_index: -1
ftag_mcast_index: 32
ftag_alt_mcast_index: 48
------------------------------------------------------------------
ftag[0x956acbc] id: 257 (0x101)
```

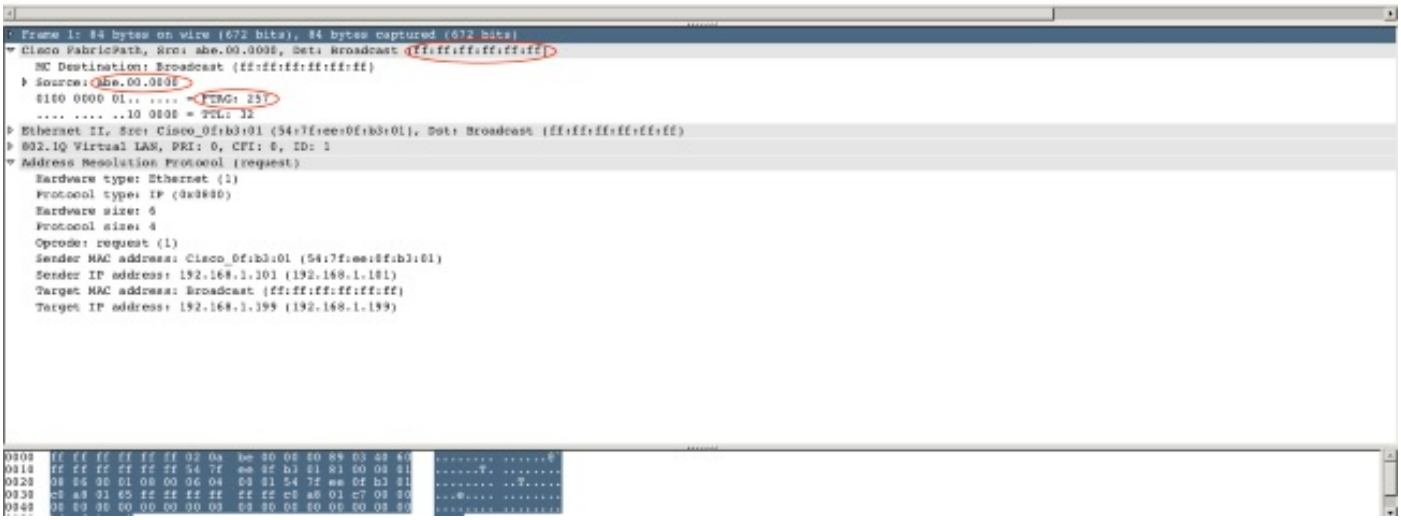
```
Topology ID: 0x111
Ftag flags: 0 (invalid ftag-flags)
Is stale: FALSE
ftag_mask[0x97359bc]
ifindex array:
0x160000c7 0x16000066 0x1600006e
0x1a01f000 0x15010000 0x15020000
0x1600007e 0x16000000
ifmap[0x95c624c]
ifmap idx 7: ref 1, lu_mcq_alloced 0, lu_mcq 16 (orig 16) 'not pruned'
ifmap idx 7: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 7: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 7: ifs - sup-eth1 sup-eth2 Po200 Po103 Po1 Po111 Eth1/32 Po127
rpf: (0x0)
alternate: 0
intf:
Po1 (0x16000000)
ftag_ucast_index: 0
ftag_flood_index: 1
ftag_mcast_index: 32
ftag_alt_mcast_index: 48
------------------------------------------------------------------------
```

測試1:進入vPC輔助裝置的廣播ARP流量

從主機1(192.168.1.101)對不存在的IP 192.168.1.199執行ping操作。 因此，主機1不斷發出廣播ARP請求，詢問「誰是192.168.1.199」。 主機1碰巧將該廣播流量雜湊到vPC輔助交換機N5K-C5596UP-110，後者又將其泛洪到VLAN 1中的所有埠，包括vPC對等鏈路Po1。

擷取連線埠通道1的TX SPAN以檢視此ARP廣播（FP術語中的多目標訊框）的光纖路徑標頭。檢視此多目標幀的交換矩陣路徑報頭。





- 由於幀通過vPC(vPC 111)進入，因此源交換機ID是abe.00.0000。
- 目的地是廣播MAC FF:FF:FF:FF:FF:FF
- FTAG是257。

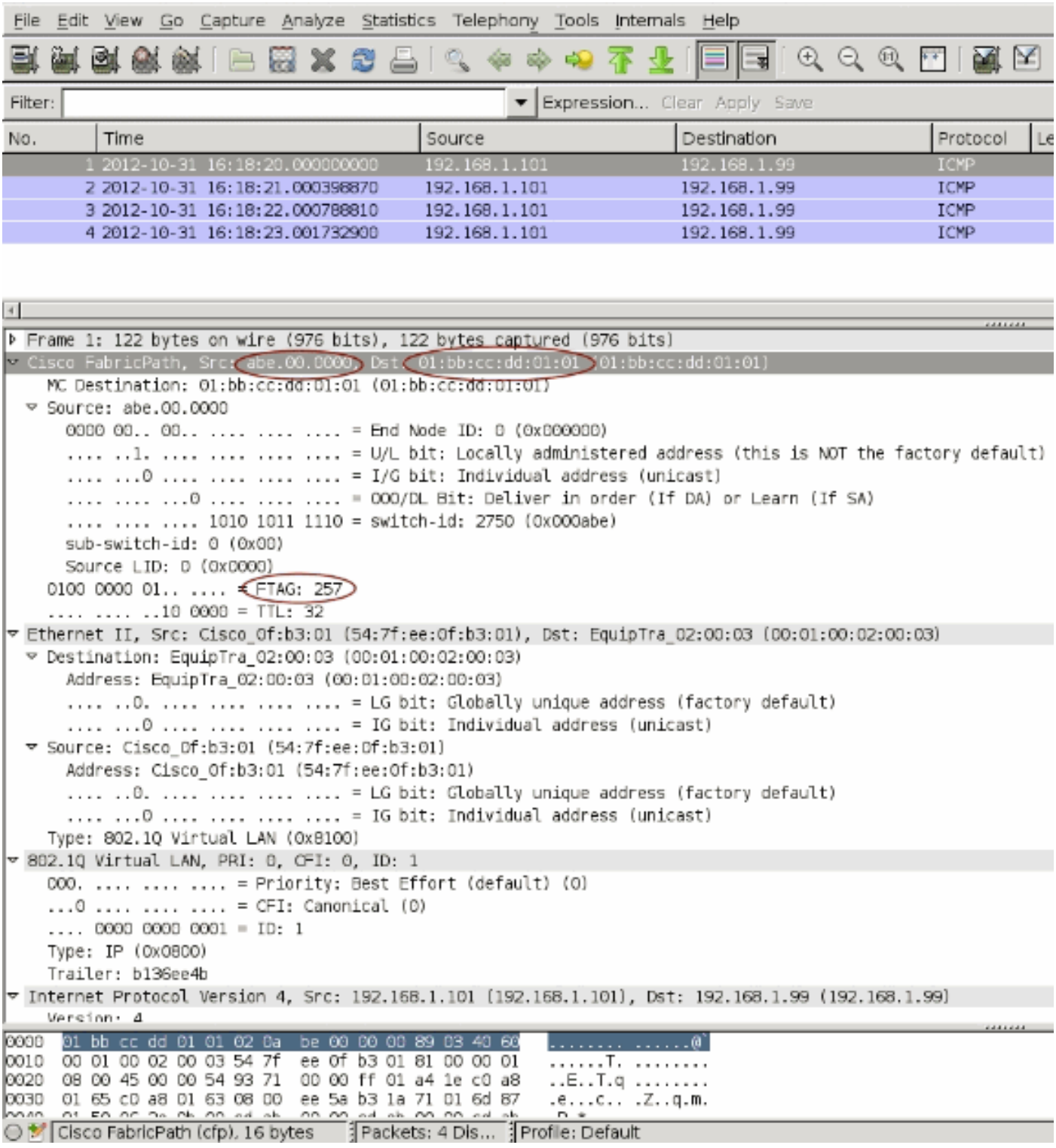
此幀進入vPC主交換機時，將檢查FTAG 257。因為只有孤立埠是FTAG 257的成員，所以此廣播

ARP幀將只傳送到Eth 1/32。

測試2:進入vPC輔助節點的未知單播幀

為了引入未知的單點傳播流量，在主機1上，我為192.168.1.99設定靜態ARP，靜態MAC為0001.0002.0003，並對192.168.1.99執行ping。ICMP回應請求到達N5K-C5596UP-110，因為它不知道MAC 0001.002.003的位置會將該幀泛洪到VLAN（包括對等鏈路）。

擷取連線埠通道1的TX SPAN來檢視此未知的單點傳播泛洪訊框（FP術語中的多目的地訊框）的光纖路徑標頭。檢視此多目標幀的交換矩陣路徑報頭。

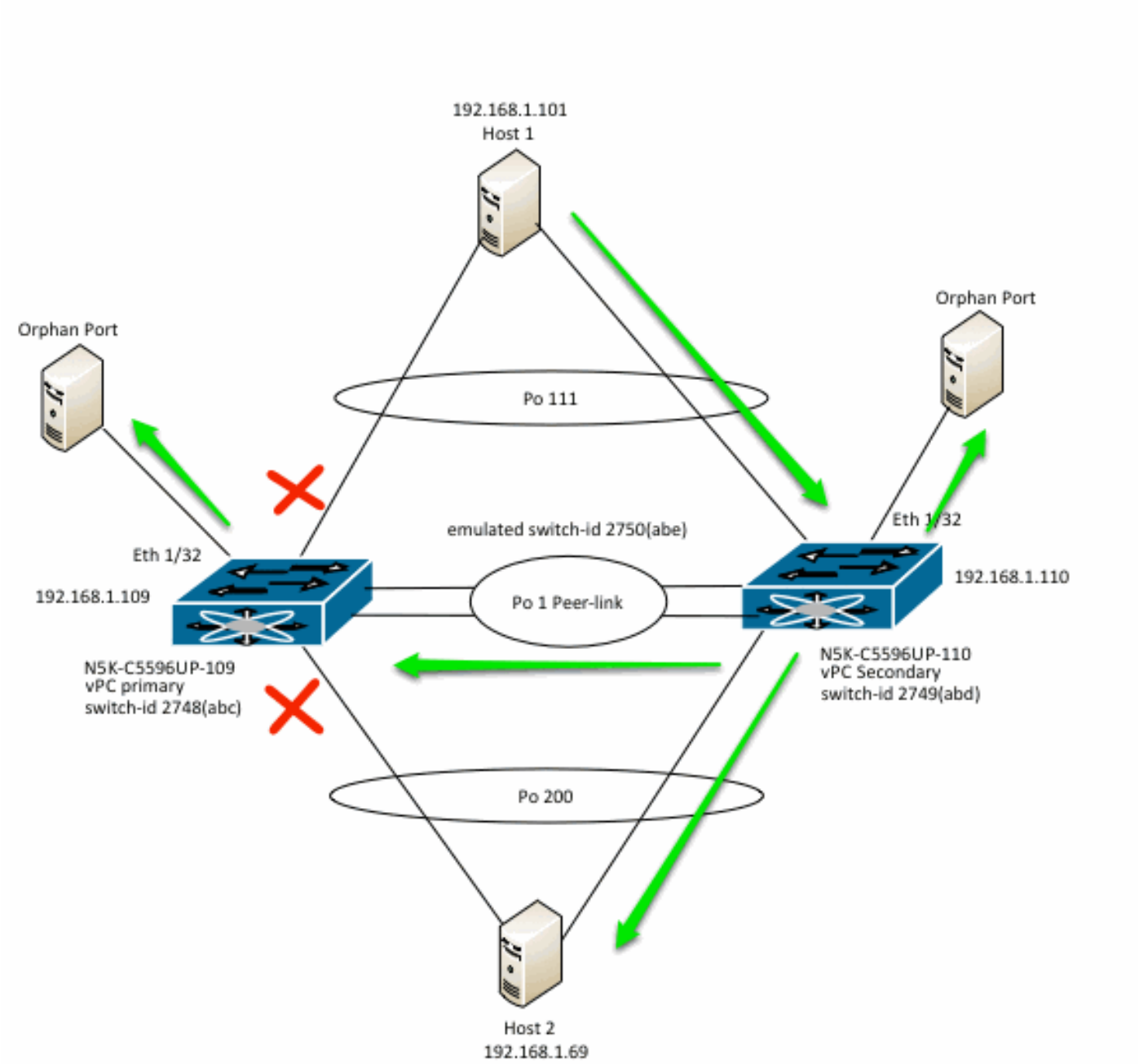


- 由於幀通過vPC(vPC 111)進入，因此源交換機ID是abe.00.0000
- 目的地是組播MAC 01:bb:cc:dd:01:01

- FTAG是257。

當此幀進入vPC主交換機時，它將檢查FTAG 257。因為只有孤立埠是FTAG 257的成員，所以此vPC主交換機將僅將該幀泛洪到孤立埠Eth 1/32。

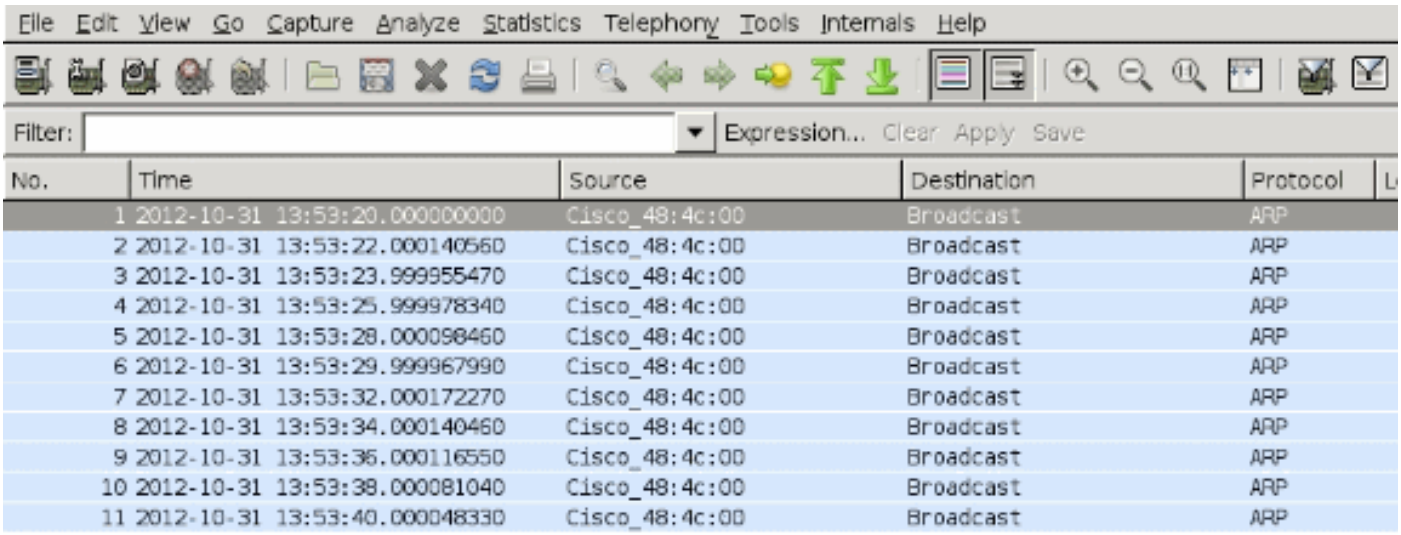
由於上述機制，以下是流入vPC輔助交換機的泛洪流量。

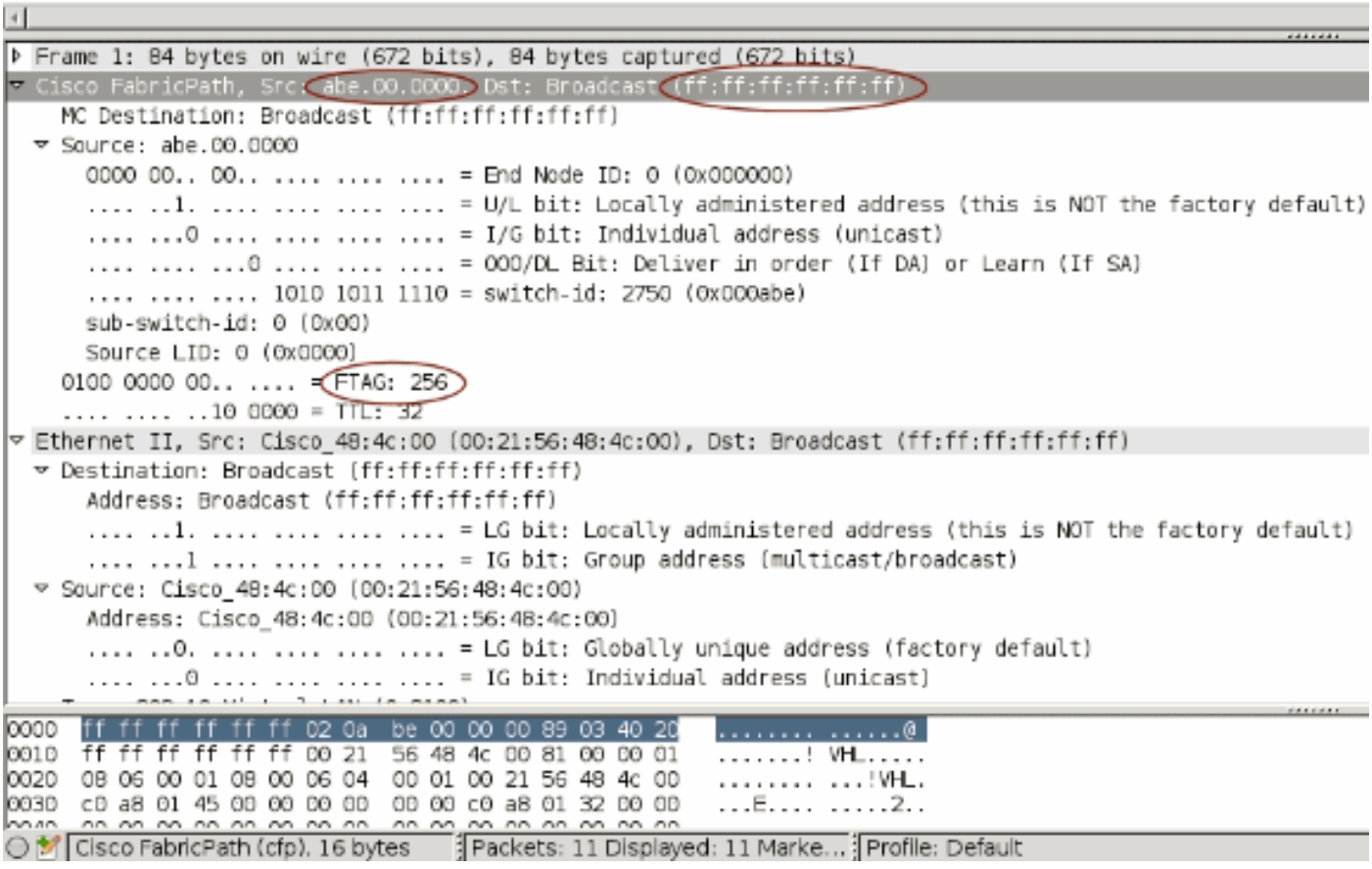


測試3:進入vPC主交換機的廣播ARP流量

從主機2(192.168.1.69)對不存在的IP 192.168.1.200執行ping操作。 因此，主機2不斷發出廣播ARP請求，詢問「誰是192.168.1.200」。 主機2碰巧將該廣播流量雜湊到vPC主交換機N5K-C5596UP-109，後者又將其泛洪到VLAN 1中的所有埠，包括vPC對等鏈路Po1。

擷取連線埠通道1的TX SPAN以檢視此ARP廣播（FP術語中的多目標訊框）的光纖路徑標頭。檢視此多目標幀的交換矩陣路徑報頭。



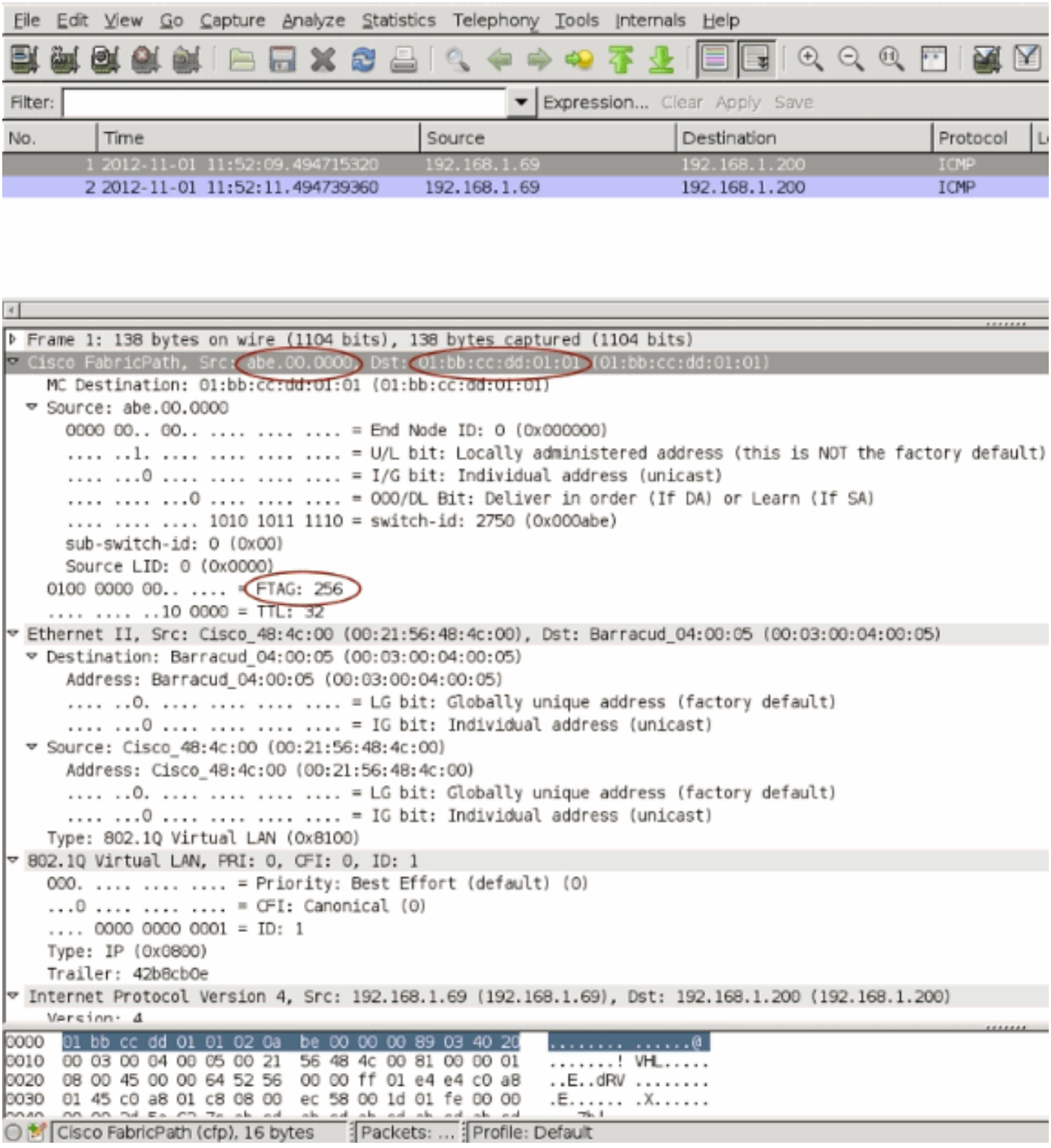


- 由於幀通過vPC(vPC 200)進入，因此源交換機ID是abe.00.0000
- 目的地是廣播MAC FF:FF:FF:FF:FF:FF
- FTAG是256。

當此幀進入vPC輔助交換機時，它將檢查FTAG 256。因為只有孤立埠是FTAG 256的成員，所以此廣播ARP幀將只傳送到Eth 1/32。

測試4:進入vPC主交換機的未知單播幀

為了引入未知的單點傳播流量，在主機2上，使用靜態MAC 0003.0004.0005設定192.168.1.200的靜態ARP，並且ping 192.168.1.200。ICMP回應請求雜湊到vPC主N5K-C5596UP-109，因為它不知道MAC 0003.0004.0005的位置，所以它會將此幀泛洪到VLAN（包括對等鏈路）。擷取連線埠通道1的TX SPAN以檢視此未知單點傳播泛洪訊框（FP術語中的多目標訊框）的光纖路徑標頭。檢視此多目標幀的交換矩陣路徑報頭。

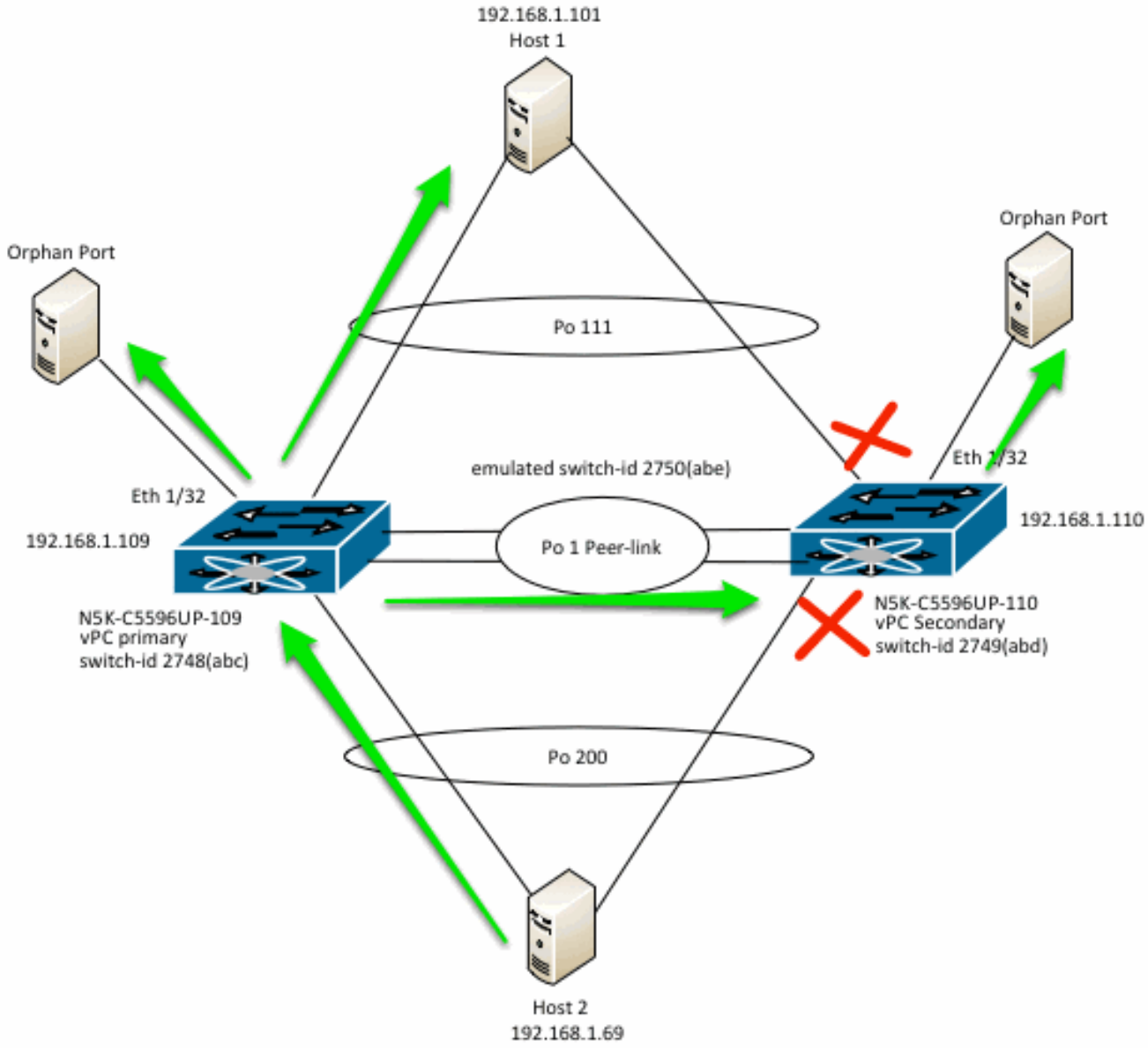


- 由於幀通過vPC(vPC 200)進入，因此源交換機ID是abe.00.0000

- 目的地是組播MAC 01:bb:cc:dd:01:01，用於未知的單點傳播泛洪
- FTAG是256。

當此幀進入vPC輔助交換機時，它將檢查FTAG 257。因為只有孤立埠是FTAG 256的成員，所以此vPC主交換機將僅將該幀泛洪到孤立埠Eth 1/32。

由於上述機制，以下為流入vPC主交換機的泛洪流量。



相關資訊

- 技術支援與文件 - Cisco Systems