基于Carmel ASIC的交换机(Nexus 5548/5596)中vPC对等链路上基于L2MP的转发

目录

简介

在vPC拓扑中，仅孤立端口流量或泛洪流量（未知单播、广播、组播）的对等链路上会看到用户流量。 对于此泛洪流量，交换机要求确保在vPC的一条支路上接收的泛洪流量不会发回另一个vPC支路上，以便数据包不会发回源或复制到其他vPC。

在基于Carmel的交换机(Nexus 55xx)中，vPC环路避免实施与基于Gatos(Nexus 5010/5020)的实施不同，后者使用单独的内部MCT VLAN在对等链路上泛洪流量。

由于基于Carmel的交换机支持L2MP或fabricpath，工程部门决定在对等链路上使用基于L2MP的转发。使用此型号时，vPC主交换机的交换机ID为2748(0xabc)，而vPC辅助交换机的交换机ID为2749(0xabd)。 模拟交换机ID 2750(0xabe)将用作传入vPC但通过对等链路发送的帧的源交换机ID。vPC主交换机上的所有端口将是FTAG 256的成员，而vPC辅助交换机上的所有端口将是FTAG 257的成员。在vPC主交换机中，只有孤立端口将是FTAG 257的成员，而在vPC辅助交换机中，孤立端口将是FTAG 256的成员。

先决条件

要求

本文档没有任何特定的要求。

使用的组件

本文档不限于特定的软件和硬件版本。

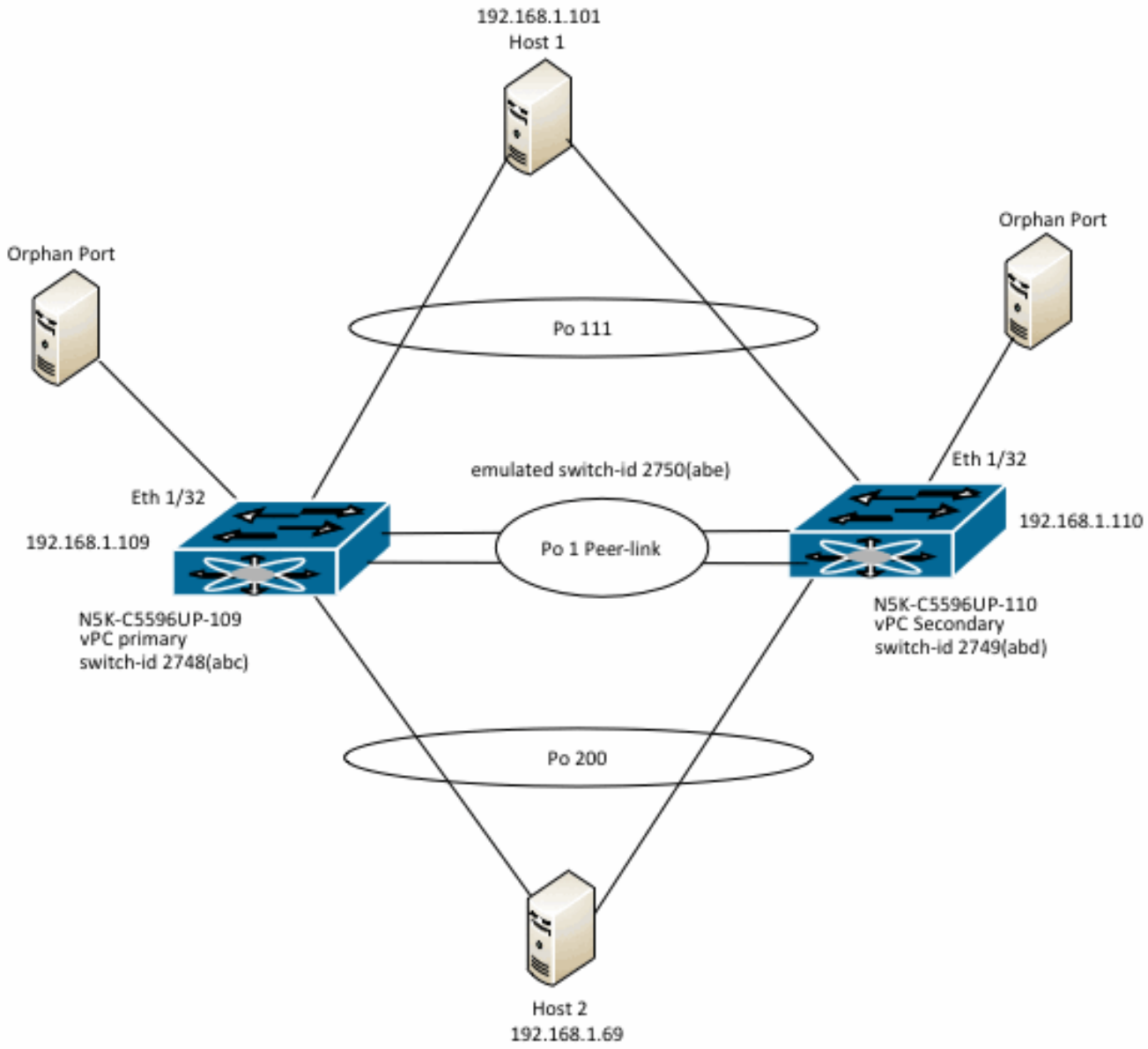
规则

有关文档规则的详细信息，请参阅 Cisco 技术提示规则。

环路避免

对于进入vPC主交换机的广播/未知单播/组播帧，它们将通过对等链路以FTAG 256发送出去。当vPC辅助交换机通过vPC对等链路获取此帧时，它会检查FTAG，自其256以来，vPC辅助交换机只会将其发送到仅是孤立端口的FTAG 256成员。对于来自vPC辅助的泛洪流量，它将以FTAG 257发送，当vPC主交换机收到此帧时，它仅将收到的泛洪帧发送给仅作为孤立端口的FTAG 257的成员。这是基于Carmel的交换机如何实现vPC环路避免的方法。

为了深入探讨基于L2MP/FTAG的泛洪帧在对等链路间的转发，请使用此拓扑：



N5K-C5596UP-109和N5K-C5596UP-100是运行NX-OS 5.2(1)N1(2a)的Nexus 5596交换机的vPC对。N5K-C5596UP-109是vPC主交换机，N5K-C5596UP-110是vPC辅助交换机。Port-channel 1是vPC对等链路。显示的IP地址属于交换机的接口VLAN 1。主机1和主机2是通过VLAN 1中的vPC连接的思科交换机。本文档中称为主机1和主机2。VLAN 1中有孤立端口连接到两台交换机上的Eth1/32。

以下是交换机的一些命令输出：

```
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
```

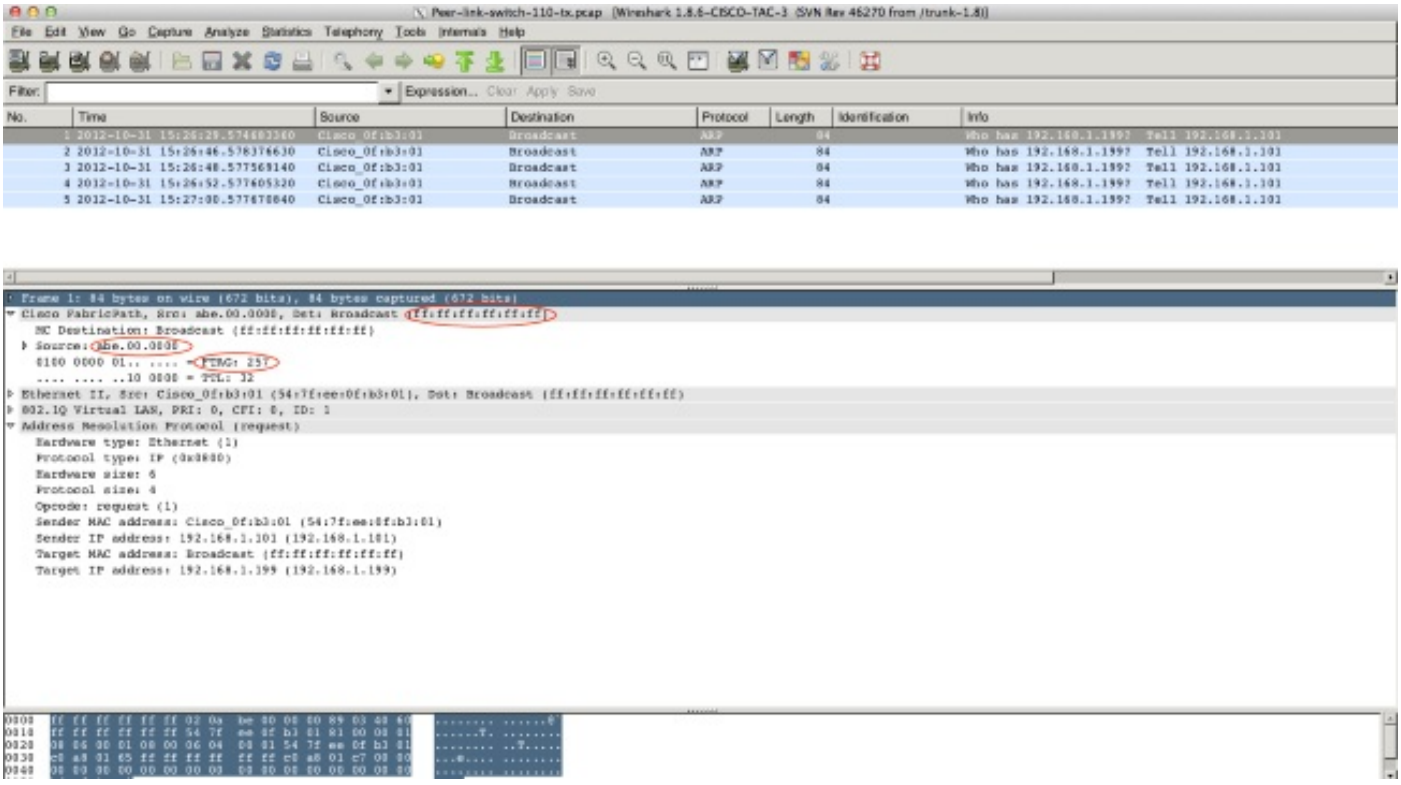
```
ifmap[0x88400fc]
ifmap idx 6: ref 1, lu_mcq_alloced 0, lu_mcq 15 (orig 15) 'not pruned'
ifmap idx 6: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 6: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 6: ifs - sup-eth1 sup-eth2 Po200 Po1 Po111 Eth1/32 Po127
rpf: (0x0)
alternate: 0
intf:
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
------------------------------------------------------------------------
```

测试 1：广播ARP流量进入vPC辅助

从主机1(192.168.1.101)对不存在的IP 192.168.1.199执行ping操作。 因此，主机1不断发出广播ARP请求，询问“谁是192.168.1.199”。 主机1会将此广播流量散列到vPC辅助交换机N5K-C5596UP-110，然后vPC辅助交换机N5K-C5596UP-110，从而将其泛洪到VLAN 1中的所有端口，包括vPC对等链路Po1。

捕获端口通道1的TX SPAN，以查看此ARP广播的交换矩阵路径报头，该ARP广播是FP术语中的多目标帧。查看此多目标帧的交换矩阵路径报头。



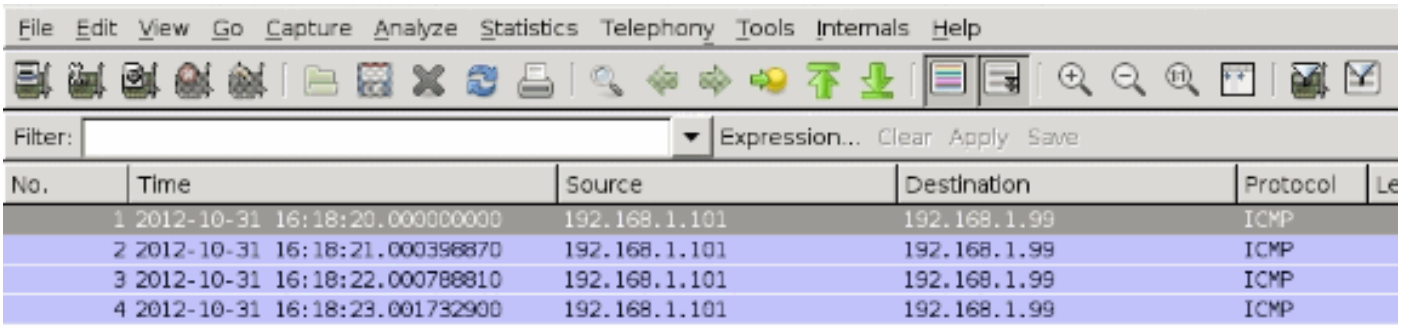
- 由于帧通过vPC(vPC 111)进入，因此源交换机ID为abe.00.0000。
- 目的地是广播MAC FF:FF:FF:FF:FF:FF
- FTAG是257。

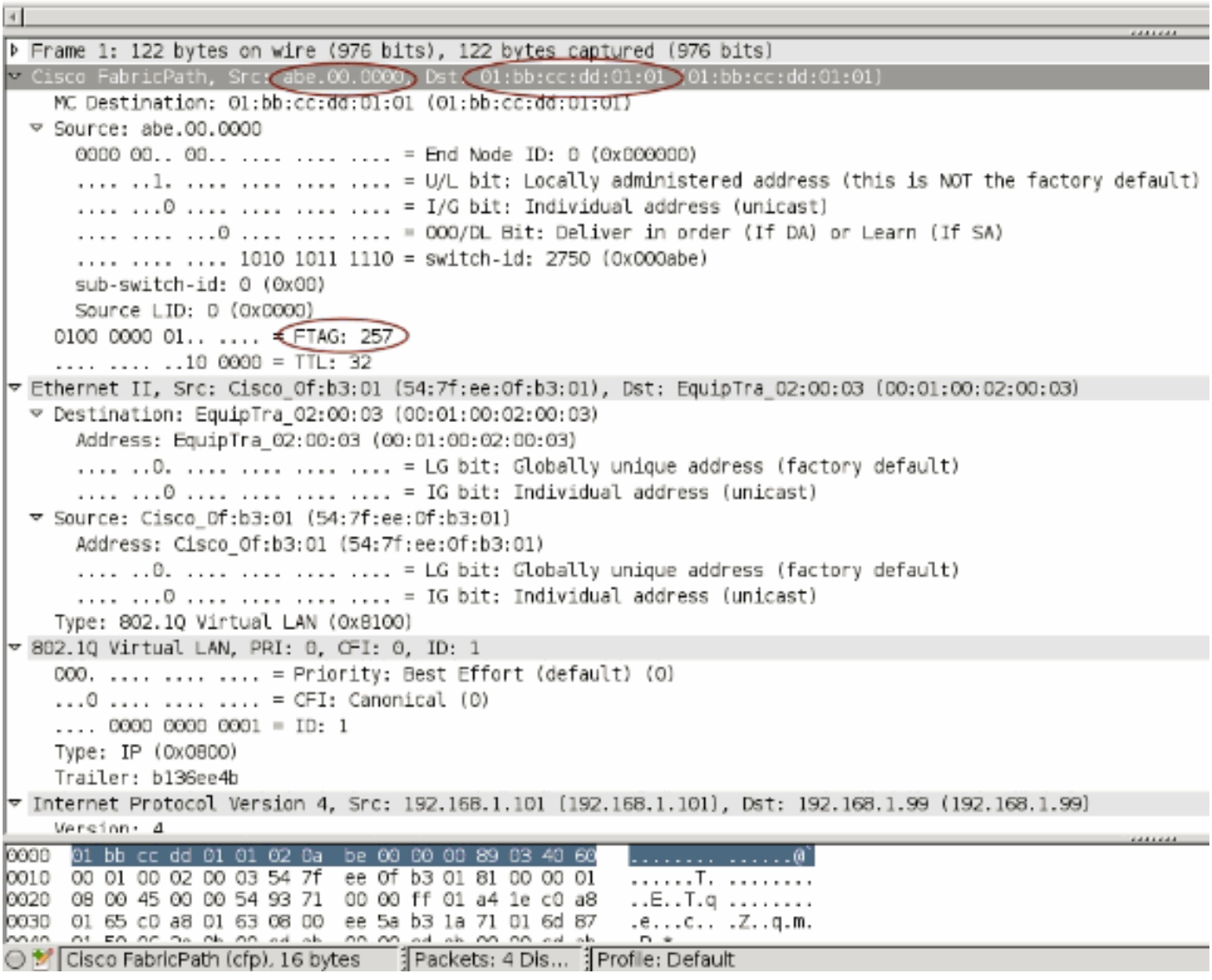
当此帧进入vPC主交换机时，它将检查FTAG 257。由于只有孤立端口是FTAG 257的成员，因此此广播ARP帧将仅发送到Eth 1/32。

测试 2：进入vPC辅助的未知单播帧

为了引入未知单播流量，在主机1上，我为192.168.1.99设置了静态ARP，静态MAC为0001.0002.0003，并对192.168.1.99执行ping操作。ICMP回应请求到达N5K-C5596UP-110，由于不知道MAC 0001.0002.0003在何处，因此它将该帧泛洪到VLAN中，包括对等链路。

捕获端口通道1的TX SPAN，以查看此未知单播泛洪帧的交换矩阵路径报头，该帧是FP术语中的多目标帧。查看此多目标帧的交换矩阵路径报头。



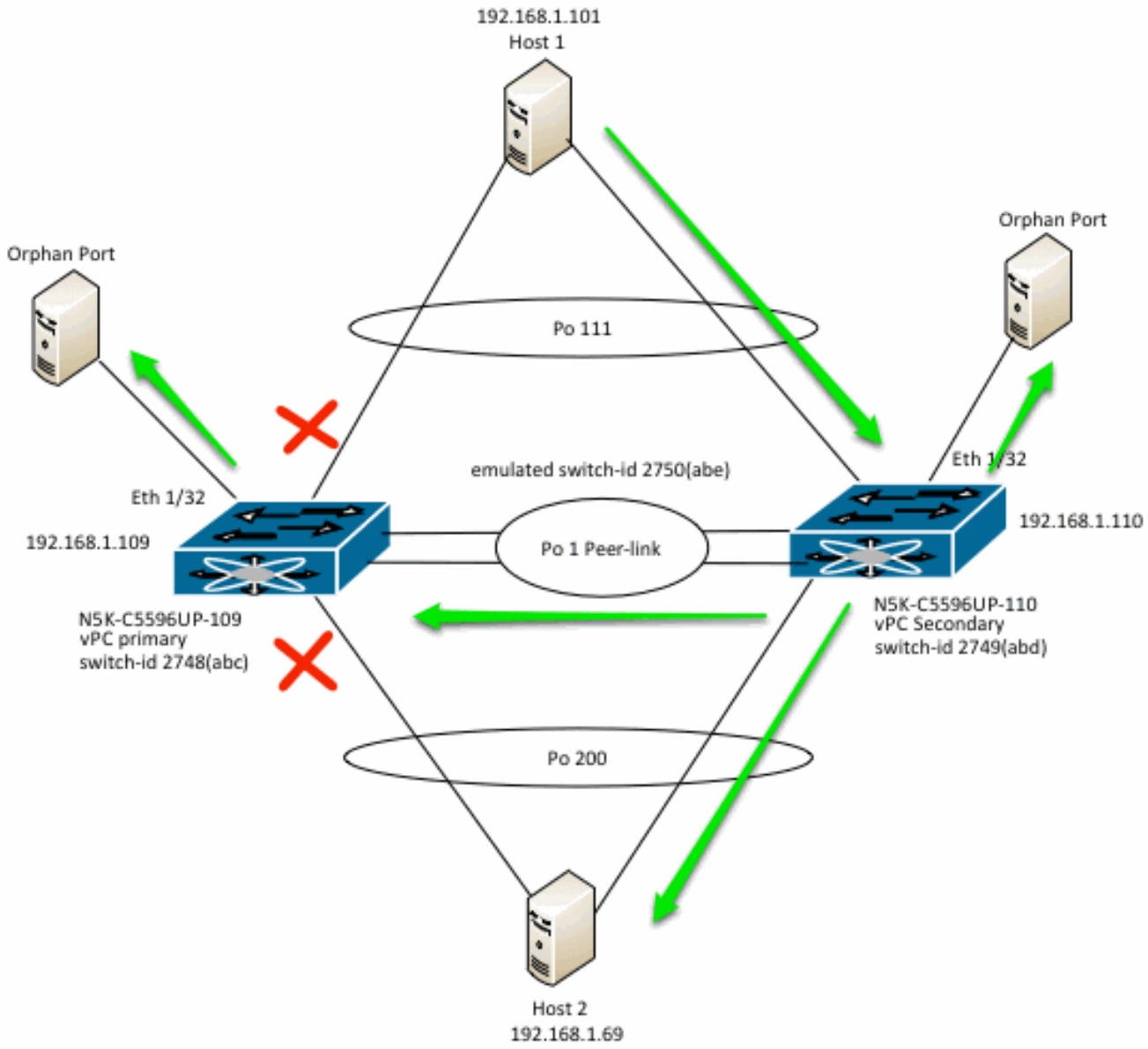


- 由于帧通过vPC(vPC 111)进入，因此源交换机ID为abe.00.0000

- 目标是组播MAC 01:bb:cc:dd:01:01
- FTAG是257。

当此帧进入vPC主交换机时，它将检查FTAG 257。由于只有孤立端口是FTAG 257的成员，因此此vPC主交换机将仅将此帧泛洪到孤立端口Eth 1/32。

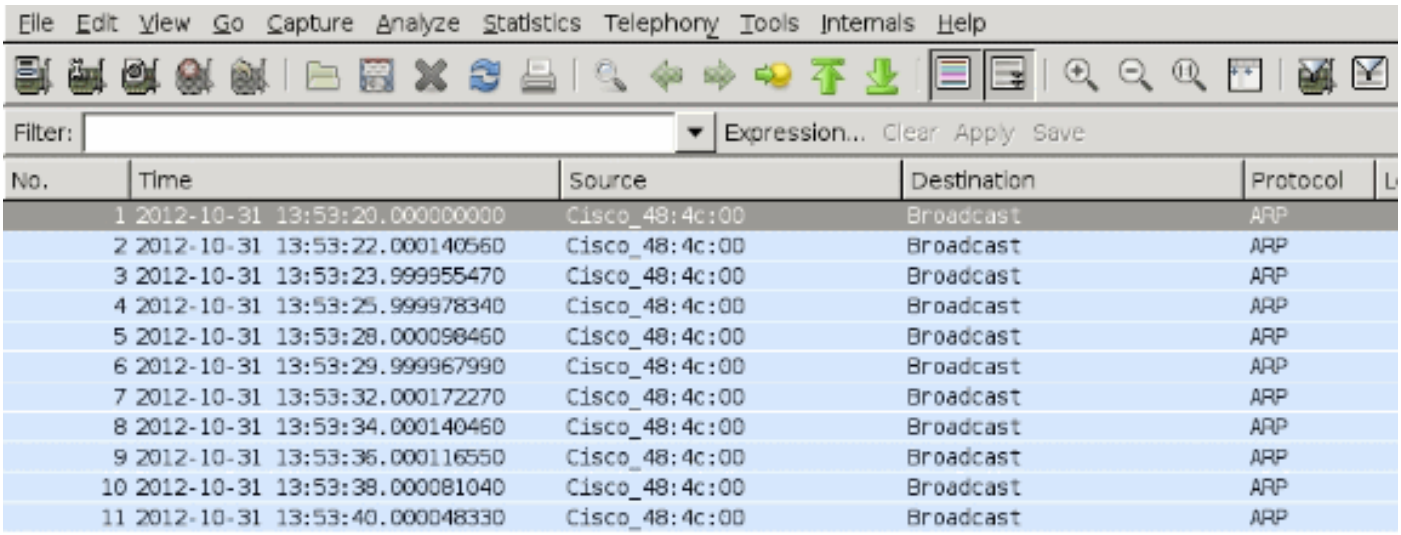
由于上述机制，以下是流入vPC辅助交换机的泛洪流量的流。

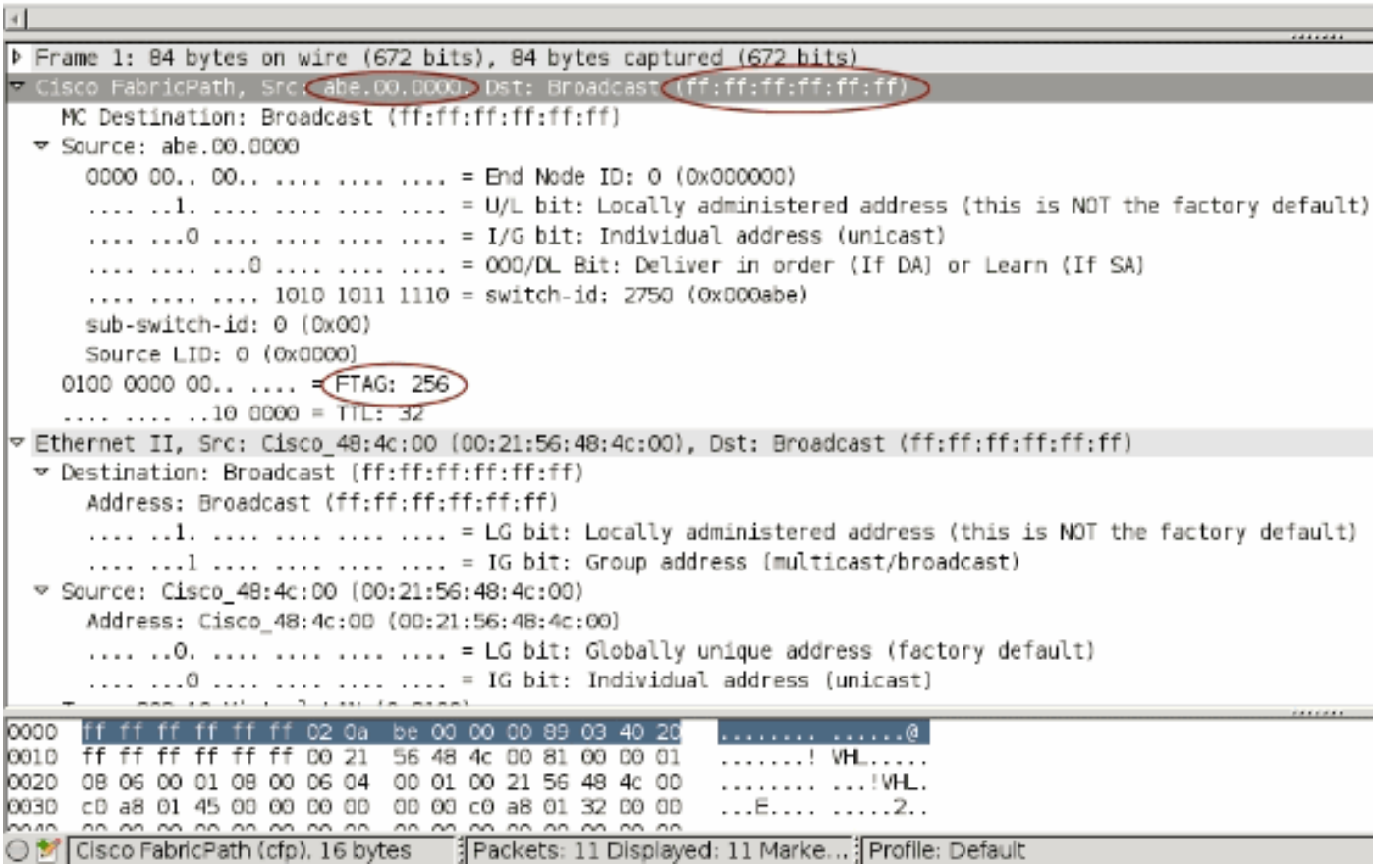


测试 3：广播ARP流量进入vPC主要

从主机2(192.168.1.69)对不存在的IP 192.168.1.200执行ping操作。 因此，主机2不断发出广播ARP请求，询问“谁是192.168.1.200”。 主机2会将此广播流量散列到vPC主交换机N5K-C5596UP-109，而vPC主交换机N5K-C5596UP-109又将其泛洪到VLAN 1中的所有端口，包括vPC对等链路Po1。

捕获端口通道1的TX SPAN，以查看此ARP广播的交换矩阵路径报头，该ARP广播是FP术语中的多目标帧。查看此多目标帧的交换矩阵路径报头。





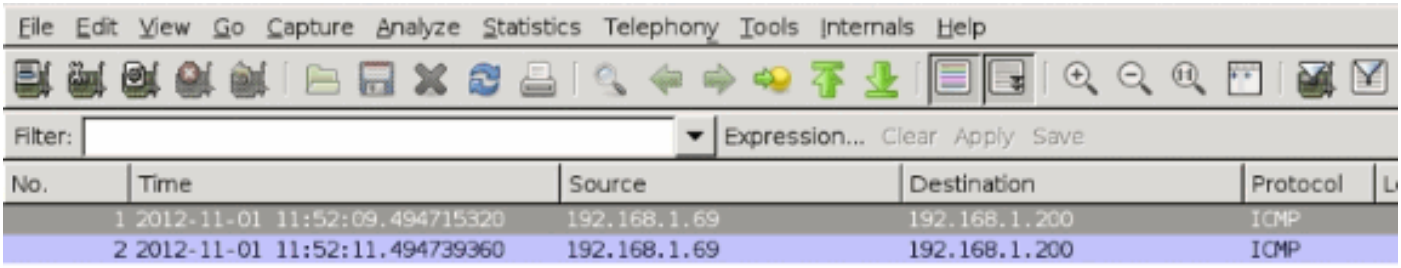
- 由于帧通过vPC(vPC 200)进入，因此源交换机ID为abe.00.0000
- 目的地是广播MAC FF:FF:FF:FF:FF:FF
- FTAG是256。

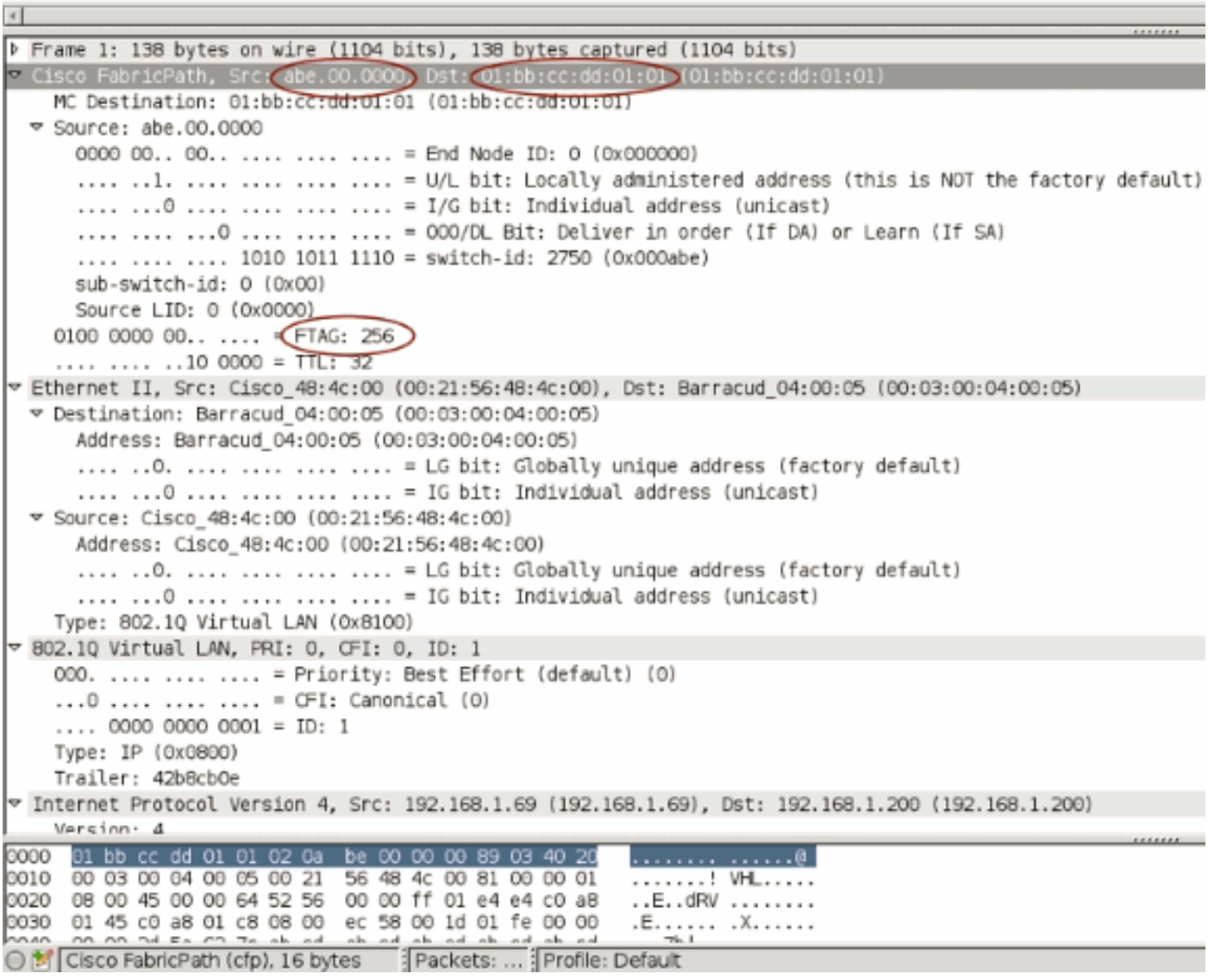
当此帧进入vPC辅助交换机时，它将检查FTAG 256。由于只有孤立端口是FTAG 256的成员，因此此广播ARP帧将仅发送到Eth 1/32。

测试 4：传入vPC主要的未知单播帧

为引入未知单播流量，在主机2上，为192.168.1.200设置静态ARP，静态MAC为

0003.0004.0005,ping静态MAC为192.168.1.200。ICMP回应请求散列到vPC主N5K-C5596UP-109，由于它不知道MAC 0003.0004.0005在何处，因此它将该帧泛洪到VLAN中，包括对等链路。捕获端口通道1的TX SPAN，以查看此未知单播泛洪帧的交换矩阵路径报头，该帧是FP术语中的多目标帧。查看此多目标帧的交换矩阵路径报头。

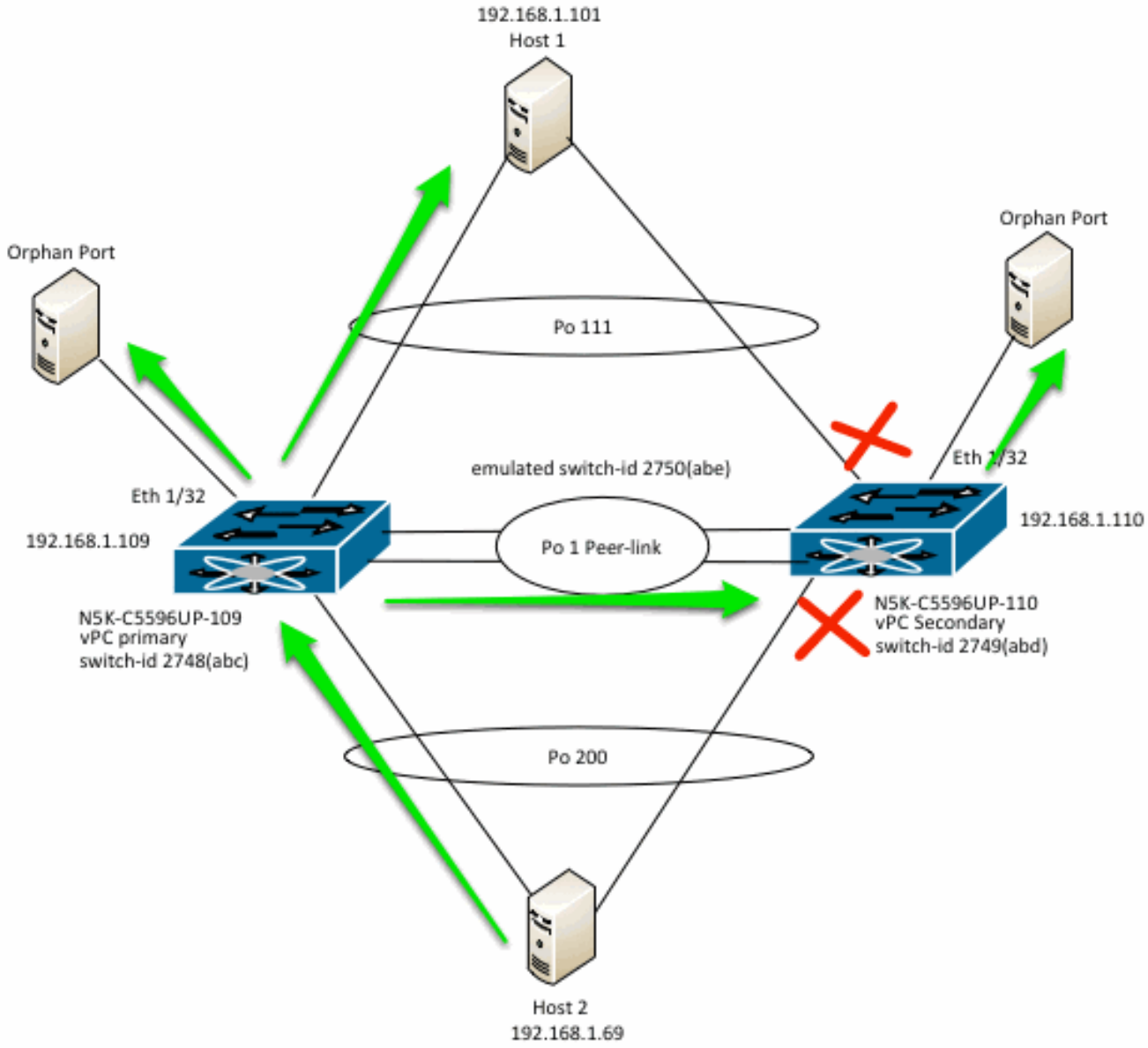




- 由于帧通过vPC(vPC 200)进入，因此源交换机ID为abe.00.0000
- 目标是用于未知单播泛洪的组播MAC 01:bb:cc:dd:01:01
- FTAG是256。

当此帧进入vPC辅助交换机时，它将检查FTAG 257。由于只有孤立端口是FTAG 256的成员，因此此vPC主要将此帧仅泛洪到孤立端口Eth 1/32。

由于上述机制，以下是流入vPC主交换机的泛洪流量的流量。



相关信息

- 技术支持和文档 - Cisco Systems