Encaminhamento baseado em L2MP através do link de peer vPC em switches baseados em Carmel ASIC (Nexus 5548/5596)

Contents

Introduction
Prerequisites
Requirements
Componentes Utilizados
Conventions
Prevenção de loop
Informações Relacionadas

Introduction

Nas topologias vPC, o tráfego do usuário será visto no peer-link somente para tráfego de porta órfã ou tráfego inundado (unicast, broadcast, multicast desconhecido). Para esse tráfego de inundação, há um requisito que os switches assegurem que o tráfego de inundação recebido em uma perna do vPC não seja enviado de volta para a outra perna do vPC, de modo que os pacotes não sejam enviados de volta para a origem ou duplicados para outros vPCs.

Nos switches baseados em Carmel (Nexus 55xx), a implementação para evitar loops de vPC é diferente em comparação com a implementação baseada em Gatos (Nexus 5010/5020) que usa uma VLAN MCT interna separada para o tráfego inundado através do link de mesmo nível.

Como os switches baseados em Carmel suportam L2MP ou fabricpath, a engenharia decidiu usar o encaminhamento baseado em L2MP através do peer-link. Com esse modelo, o switch primário vPC terá um switch-id 2748(0xabc) enquanto o secundário do vPC terá um switch-id 2749(0xabd). O Emulated switch-id 2750(0xabe) será usado como ID do switch de origem para quadros que ingressam em um vPC, mas são enviados pelo link de peer. Todas as portas no vPC principal serão membros do FTAG 256, enquanto que no vPC secundário serão membros do FTAG 257. No switch primário vPC, somente portas órfãs serão membros do FTAG 257 enquanto no switch secundário vPC, portas órfãs serão membros do FTAG 256.

Prerequisites

Requirements

Não existem requisitos específicos para este documento.

Componentes Utilizados

Este documento não se restringe a versões de software e hardware específicas.

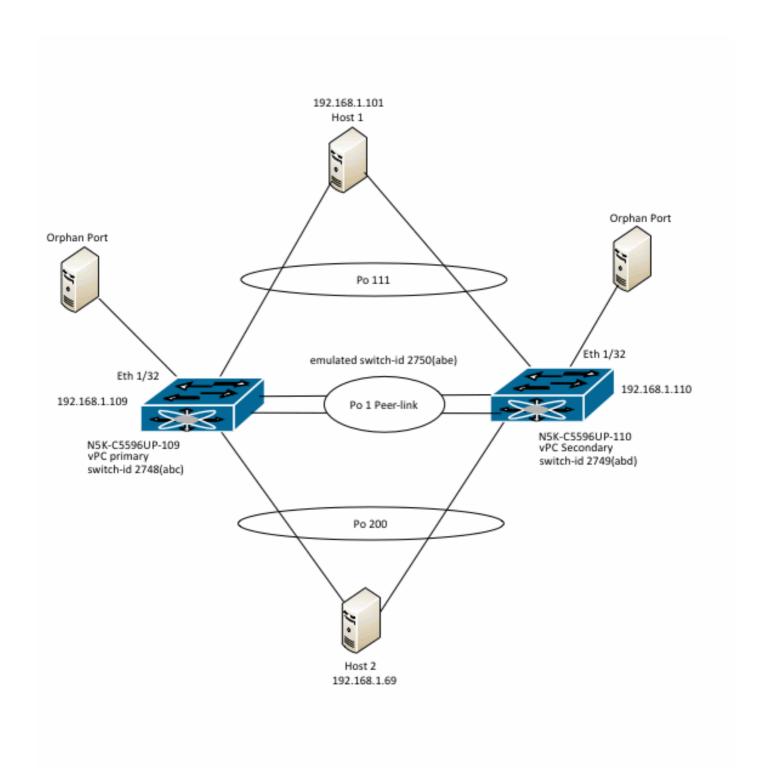
Conventions

Consulte as <u>Convenções de Dicas Técnicas da Cisco para obter mais informações sobre convenções de documentos.</u>

Prevenção de loop

Para quadros unicast/multicast de broadcast/desconhecidos que entram no switch primário vPC, eles serão enviados com um FTAG de 256 pelo link de peer. Quando o switch secundário vPC obtém esse quadro através do link par do vPC, ele inspeciona o FTAG e, desde seus 256, o switch secundário do vPC só o envia aos membros do FTAG 256, que serão apenas portas órfãs. Para o tráfego de inundação do vPC secundário, ele será enviado com FTAG de 257 e quando o switch primário vPC receber esse quadro, ele enviará o quadro de inundação recebido somente aos membros do FTAG 257, que serão apenas portas órfãs. É assim que os switches baseados em Carmel implementam a prevenção de loop de vPC.

Para mergulhar profundamente o encaminhamento baseado em L2MP/FTAG de quadros de inundação pelo link de mesmo nível, essa topologia é usada:



O N5K-C5596UP-109 e o N5K-C5596UP-100 são um par vPC de switches Nexus 5596 executando o NX-OS 5.2(1)N1(2a). O N5K-C5596UP-109 é o switch principal do vPC e o N5K-C5596UP-110 é o switch secundário do vPC. O canal de porta 1 é o peer-link do vPC. Os endereços IP mostrados pertencem à interface VLAN 1 dos switches. Os hosts 1 e 2 são switches Cisco conectados via vPC na VLAN 1. Eles são chamados de host 1 e host 2 neste documento. Há uma porta órfã na VLAN 1 conectada a Eth1/32 em ambos os switches.

Aqui estão algumas saídas de comandos dos switches:

```
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
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
  Po1 up
vPC status
_____
            Status Consistency Reason
_____
111 Po111
            up
                 success
                          success
            up success
200 Po200
                          success
N5K-C5596UP-109# show platform fwm info 12mp 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
----::Going through port database. Please be patient.::-----
          Orphan Ports
          Eth1/32
N5K-C5596UP-110# show vpc
```

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

Legend:

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

l Pol up 1

vPC status

id Port Status Consistency Reason Active vlans

111 Poll1 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.

 ${\tt N5K-C5596UP-109\#}$ show platform fwm info l2mp ftag all

L2MP FTAG

ftag[0x9565blc] 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

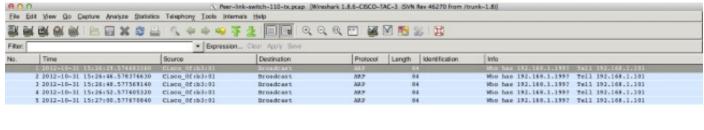
```
ifmap[0x88400fc]
ifmap idx 6: ref 1, lu_mcq_alloced 0, lu_mcq 15 (orig 15) 'not pruned'
ifmap idx 6: prune_ifmap 0, prune ref count 0, prune_unvisited 0
ifmap_idx 6: oifls_macg_ref_cnt 0, num_oifls 0
ifmap idx 6: ifs - sup-eth1 sup-eth2 Po200 Po1 Po111 Eth1/32 Po127
rpf: (0x0)
alternate: 0
intf:
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:
Pol (0x1600000)
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 12mp ftag all
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:
Pol (0x16000000)
ftag_ucast_index: 1
ftag_flood_index: -1
ftag_mcast_index: 32
ftag_alt_mcast_index: 48
ftag[0x956acbc] id: 257 (0x101)
```

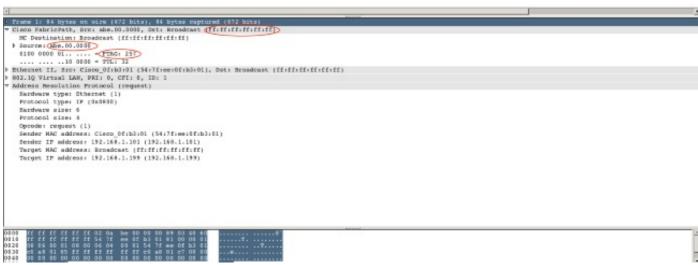
```
Topology ID: 0x111
Ftag flags: 0 (invalid ftag-flags)
Is stale: FALSE
ftag_mask[0x97359bc]
ifindex array:
0x160000c7 0x16000066 0x1600006e
0x1a01f000 0x15010000 0x15020000
0x1600007e 0x16000000
ifmap[0x95c624c]
ifmap idx 7: ref 1, lu_mcq_alloced 0, lu_mcq 16 (orig 16) 'not pruned'
ifmap idx 7: prune_ifmap 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:
Pol (0x16000000)
ftag_ucast_index: 0
ftag_flood_index: 1
ftag_mcast_index: 32
ftag_alt_mcast_index: 48
```

Teste 1: Tráfego ARP de broadcast que entra no vPC secundário

Um IP 192.168.1.199 inexistente é enviado por ping do host 1 (192.168.1.101). Devido a isso, o host 1 continua enviando uma solicitação ARP de broadcast perguntando "quem é 192.168.1.199". O host 1 acontece ao hash desse tráfego de broadcast para o switch secundário vPC N5K-C5596UP-110, que, por sua vez, o inunda para todas as portas na VLAN 1 incluindo Po1, que é o link par do vPC.

Um SPAN TX do canal de porta 1 é capturado para examinar os cabeçalhos do caminho de estrutura desse broadcast ARP, que é um quadro de vários destinos na terminologia FP. Examine o cabeçalho do caminho da estrutura desse quadro multidestino.





Como o quadro entra por meio de um vPC(vPC 111), o ID do switch de origem é

abe.00.0000.

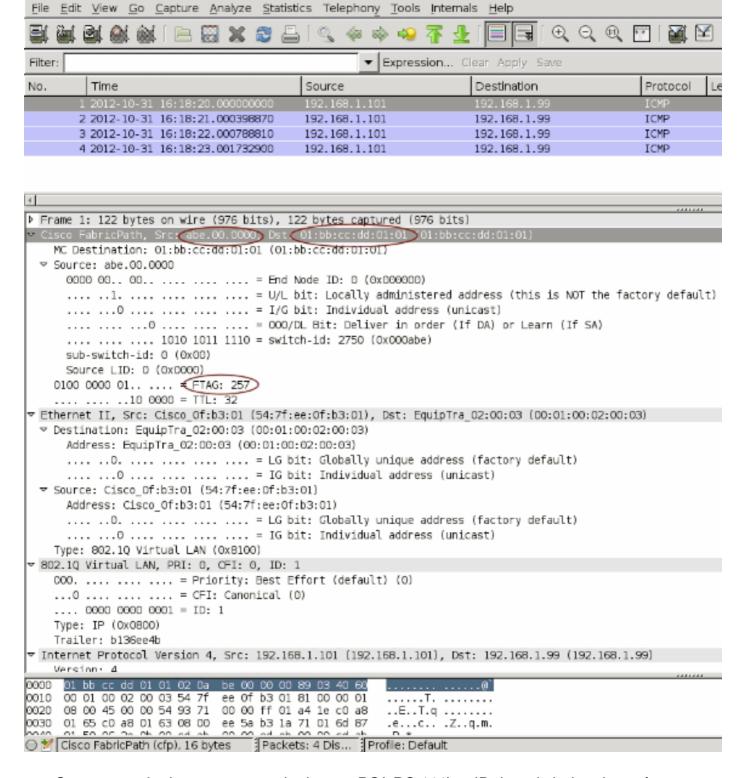
- O destino é um MAC de broadcast FF:FF:FF:FF:FF
- FTAG é 257.

Quando esse quadro entra no switch principal do vPC, ele inspeciona o FTAG 257. Como somente portas órfãs são membros do FTAG 257, esse quadro ARP de broadcast será enviado somente para Eth 1/32.

Teste 2: Quadro unicast desconhecido entrando no vPC secundário

Para introduzir tráfego unicast desconhecido, no host 1, configurei um ARP estático para 192.168.1.99 com um MAC estático de 0001.0002.0003 e fiz um ping para 192.168.1.99. A solicitação de eco ICMP chega em N5K-C5596UP-110 e, como não sabe onde está o MAC 0001.0002.0003, inunda esse quadro na VLAN incluindo link de peer.

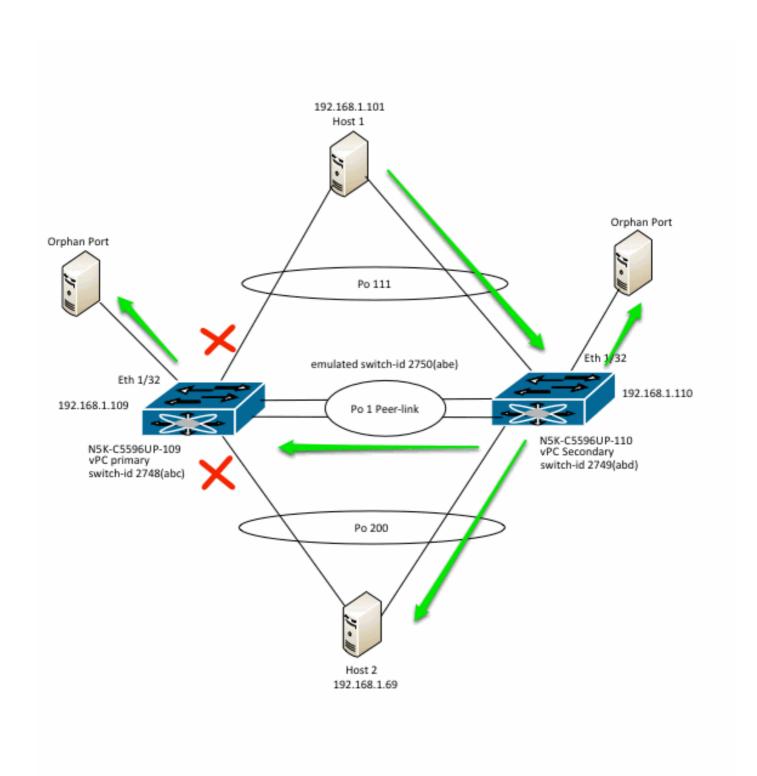
Um SPAN TX do canal de porta 1 é capturado para examinar os cabeçalhos do caminho de estrutura desse quadro de inundação unicast desconhecido, que é um quadro de vários destinos na terminologia FP. Examine o cabeçalho do caminho da estrutura desse quadro multidestino.



- Como o quadro ingressa por meio de um vPC(vPC 111), o ID do switch de origem é abe.00.000
- O destino é um MAC multicast 01:bb:cc:dd:01:01
- FTAG é 257.

Quando esse quadro entra no switch principal do vPC, ele inspeciona o FTAG 257. Como somente portas órfãs são membros do FTAG 257, este vPC primário inundará esse quadro somente para a porta órfã Eth 1/32.

Devido ao mecanismo acima, o fluxo para o tráfego inundado que chega ao switch secundário do vPC é o seguinte.

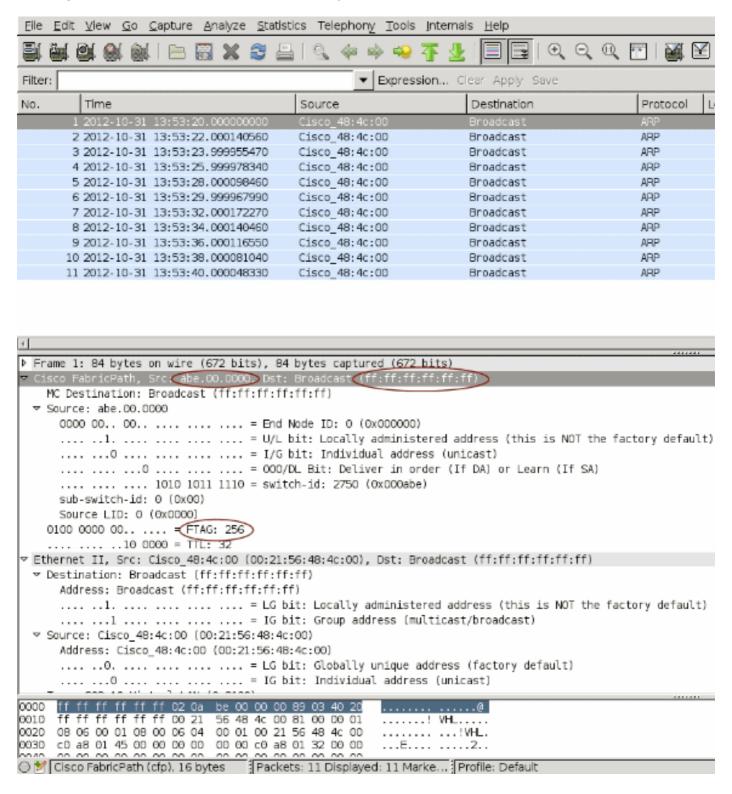


Teste 3: Tráfego ARP de broadcast que entra no vPC Primary

Um IP 192.168.1.200 inexistente é enviado por ping do host 2 (192.168.1.69). Devido a isso, o host 2 continua enviando uma solicitação ARP de broadcast perguntando "quem é 192.168.1.200". O host 2 acontece de hash desse tráfego de broadcast para o switch primário vPC N5K-C5596UP-109, que, por sua vez, o inunda para todas as portas na VLAN 1, incluindo

Po1, que é o peer-link do vPC.

Um SPAN TX do canal de porta 1 é capturado para examinar os cabeçalhos do caminho de estrutura desse broadcast ARP, que é um quadro de vários destinos na terminologia FP. Examine o cabeçalho do caminho da estrutura desse quadro multidestino.



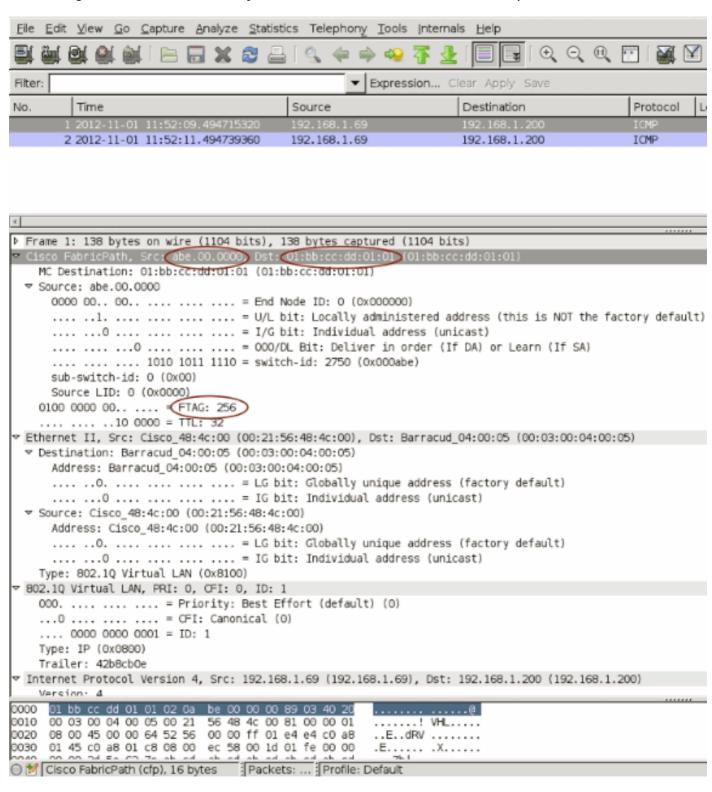
- Como o quadro ingressa por meio de um vPC(vPC 200), o ID do switch de origem é abe.00.000
- O destino é um MAC de broadcast FF:FF:FF:FF:FF
- FTAG é 256.

Quando esse quadro entra no switch secundário vPC, ele inspeciona o FTAG 256. Como somente portas órfãs são membros do FTAG 256, esse quadro ARP de broadcast será enviado

somente para Eth 1/32.

Teste 4: Quadro unicast desconhecido entrando no vPC Primary

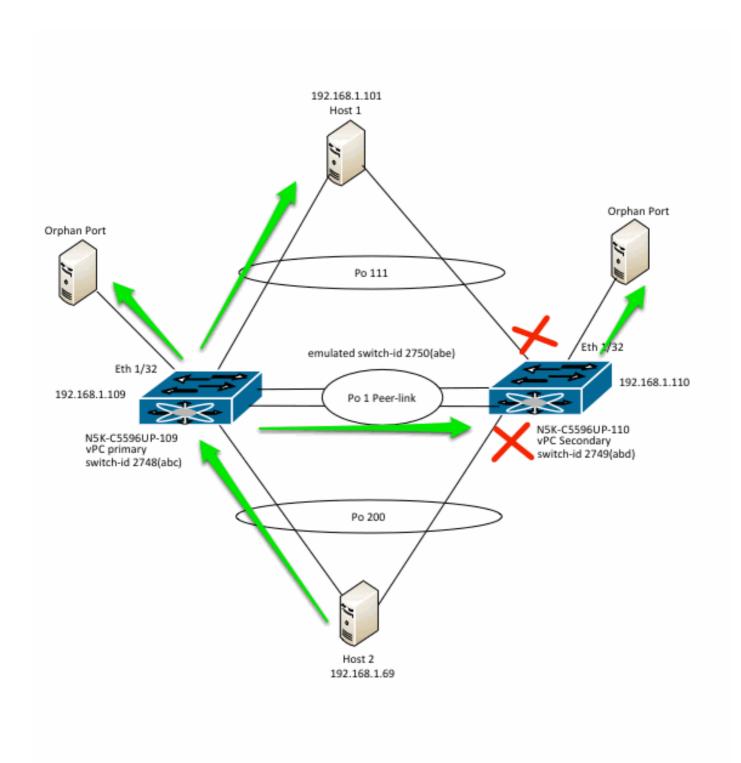
Para introduzir tráfego unicast desconhecido, no host 2, um ARP estático para 192.168.1.200 é configurado com um MAC estático de 0003.0004.0005 e 192.168.1.200 é executado ping. A solicitação de eco ICMP tem hashes para o N5K-C5596UP-109 principal do vPC e, como ele não sabe onde está o MAC 0003.0004.0005, inunda esse quadro na VLAN incluindo link de peer. Um SPAN TX do canal de porta 1 é capturado para examinar os cabeçalhos do caminho de estrutura desse quadro de inundação unicast desconhecido, que é um quadro de vários destinos na terminologia FP. Examine o cabeçalho do caminho da estrutura desse quadro multidestino.



- Como o quadro ingressa por meio de um vPC(vPC 200), o ID do switch de origem é abe.00.000
- O destino é um MAC multicast 01:bb:cc:dd:01:01 usado para inundação unicast desconhecida
- FTAG é 256.

Quando esse quadro entra no switch secundário vPC, ele inspeciona o FTAG 257. Como somente portas órfãs são membros do FTAG 256, este vPC primário inundará esse quadro somente para a porta órfã Eth 1/32.

Devido ao mecanismo acima, o fluxo para o tráfego inundado que entra no switch principal do vPC é o seguinte.



Informações Relacionadas

• Suporte Técnico e Documentação - Cisco Systems