

使用非vPC L2中继检验Nexus 9000系列ARP & MAC表同步行为

目录

[简介](#)

[背景信息](#)

[要求](#)

[使用的组件](#)

[拓扑](#)

[概述](#)

[相关信息](#)

简介

本文档介绍在共享非vPC第2层中继的Nexus 9000设备之间可能出现的ARP和MAC表行为。

背景信息

仅当SVI不使用用户定义的MAC地址，并且在vPC域下配置了vPC对等网关功能时，才会出现此行为。此外，只有当ARP表保持填充状态，而MAC地址表没有指定主机的MAC条目时，才可能看到此消息。

本文档中描述的行为是第一代Nexus交换机的ASIC限制，不影响Nexus 9300云扩展 (EX/FX/GX/C)交换机及更高版本，并且已记录为Cisco Bug ID [CSCuh94866的一部分](#)。

要求

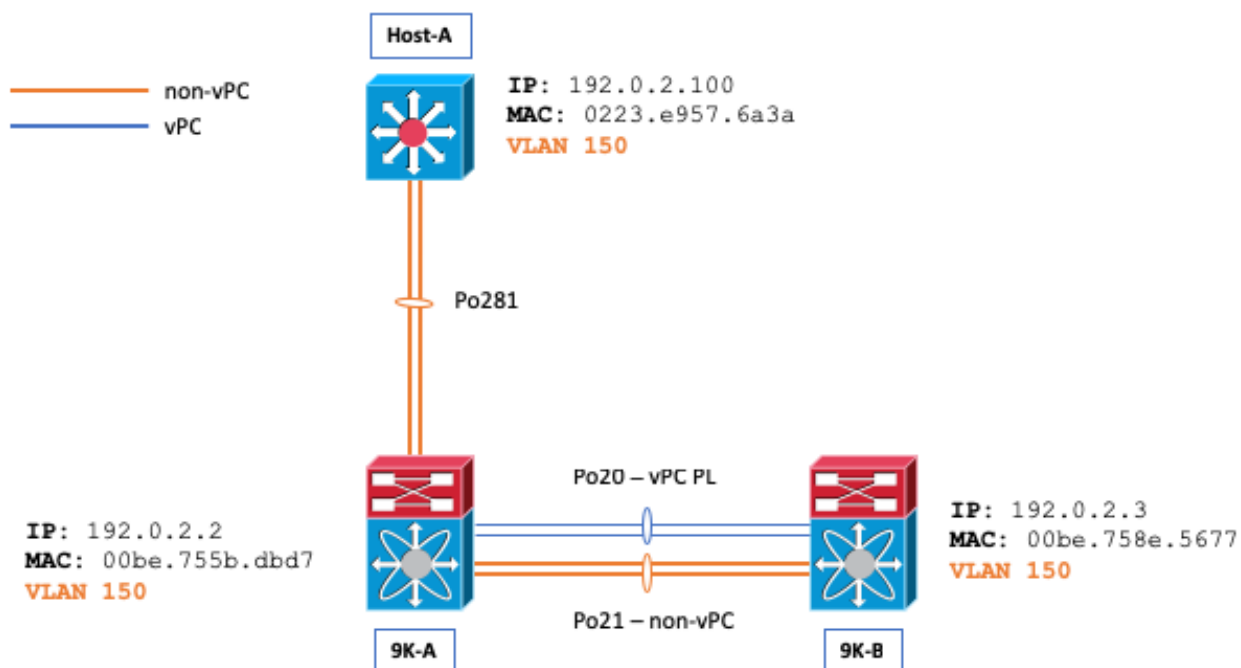
虚拟端口通道(vPC)、NXOS虚拟端口通道对等网关功能和Nexus操作系统(NXOS)的一般知识。

使用的组件

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

- Nexus 3000s/Nexus 9000s (仅限第一代)
- 虚拟端口通道功能(vPC)
- vPC对等网关功能
- 非vPC第2层(L2)中继
- 非vPC SVI
- NX-OS 7.0(3)I7(5)

拓扑



概述

假设主机A和N9K-B之间的ARP和MAC地址表为空，并且从主机A ping N9K-B。

```
Host-A# ping 192.0.2.3
PING 192.0.2.3 (192.0.2.3): 56 data bytes
36 bytes from 192.0.2.100: Destination Host Unreachable
Request 0 timed out
64 bytes from 192.0.2.3: icmp_seq=1 ttl=254 time=1.011 ms
64 bytes from 192.0.2.3: icmp_seq=2 ttl=254 time=0.763 ms
64 bytes from 192.0.2.3: icmp_seq=3 ttl=254 time=0.698 ms
64 bytes from 192.0.2.3: icmp_seq=4 ttl=254 time=0.711 ms

--- 192.0.2.3 ping statistics ---
5 packets transmitted, 4 packets received, 20.00% packet loss
round-trip min/avg/max = 0.698/0.795/1.011 ms
```

从主机A执行ping操作会导致主机A发送ARP请求9K-B。ARP请求从Po21的N9K-A（在VLAN上泛洪）进入，同时从Po20（通过思科交换矩阵服务[CFS]进行隧道）退出。因此，9K-B上的MAC地址表被正确填充，并且ARP条目被插入N9K-B的ARP表中，该表指向主机A的MAC地址0223.e957.6a3a的Po21（非vPC L2中继）。

```
N9K-B# show ip arp 192.0.2.100
```

```
Flags: * - Adjacencies learnt on non-active FHRP router
+ - Adjacencies synced via CFSOE
# - Adjacencies Throttled for Glean
CP - Added via L2RIB, Control plane Adjacencies
PS - Added via L2RIB, Peer Sync
RO - Re-Originated Peer Sync Entry
D - Static Adjacencies attached to down interface
```

```
IP ARP Table
Total number of entries: 1
```

```
Address      Age      MAC Address  Interface  Flags
192.0.2.100  00:01:07  0223.e957.6a3a  Vlan150
```

```
N9K-B# show mac address-table address | i i 6a3a
* 150      0223.e957.6a3a  dynamic 0      F      F      Po21
```

```
N9K-B# show ip arp detail | i 3a
192.0.2.100  00:03:22  0223.e957.6a3a  Vlan150      port-channel21  <<<< Expected port-
channel
```

当从N9K-B的MAC地址表中删除主机A的MAC地址时，可以看到此问题。可以出于各种原因删除MAC地址，例如MAC地址老化、生成树协议(STP)拓扑更改通知(TCN)、通过命令行界面运行clear mac address-table dynamic命令等。

```
N9K-B# show ip arp 192.0.2.100
```

```
Flags: * - Adjacencies learnt on non-active FHRP router
+ - Adjacencies synced via CFSOE
# - Adjacencies Throttled for Glean
CP - Added via L2RIB, Control plane Adjacencies
PS - Added via L2RIB, Peer Sync
RO - Re-Originated Peer Sync Entry
D - Static Adjacencies attached to down interface
```

```
IP ARP Table
```

```
Total number of entries: 1
```

```
Address      Age      MAC Address  Interface  Flags
192.0.2.100  00:00:29  0223.e957.6a3a  Vlan150      <<< ARP remains populated
```

```
N9K-B# show mac address-table address 0223.e957.6a3a
```

```
Legend:
```

```
* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen,+ - primary entry using vPC Peer-Link,
(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
```

```
VLAN      MAC Address      Type      age      Secure NTFY Ports
```

```
-----+-----+-----+-----+-----+-----+-----
```

```
N9K-B# ping 192.0.2.100
```

```
PING 192.0.2.100 (192.0.2.100): 56 data bytes
```

```
64 bytes from 192.0.2.100: icmp_seq=0 ttl=253 time=1.112 ms
```

```
64 bytes from 192.0.2.100: icmp_seq=1 ttl=253 time=0.647 ms
```

```
64 bytes from 192.0.2.100: icmp_seq=2 ttl=253 time=0.659 ms
```

```
64 bytes from 192.0.2.100: icmp_seq=3 ttl=253 time=0.634 ms
```

```
64 bytes from 192.0.2.100: icmp_seq=4 ttl=253 time=0.644 ms
```

```
--- 192.0.2.100 ping statistics ---
```

```
5 packets transmitted, 5 packets received, 0.00% packet loss
```

```
round-trip min/avg/max = 0.634/0.739/1.112 ms
```

请注意，ping仍然成功；但是，我们的ARP条目现在指向Po20(vPC PL)而不是Po21，后者不是预期的端口通道，因为VLAN 150是非VPC VLAN:

```
N9K-B# show ip arp detail | i i 6a3a
```

```
Flags: * - Adjacencies learnt on non-active FHRP router
+ - Adjacencies synced via CFSOE
# - Adjacencies Throttled for Glean
```

CP - Added via L2RIB, Control plane Adjacencies
PS - Added via L2RIB, Peer Sync
RO - Re-Originated Peer Sync Entry

IP ARP Table for context default

Total number of entries: 2

| Address | Age | MAC Address | Interface | Physical Interface | Flags |
|-------------|----------|----------------|-----------|--------------------|---|
| 192.0.2.100 | 00:15:54 | 0223.e957.6a3a | Vlan150 | port-channel20 | <<< Not Po21 once the issue is triggered. |

您可以在两台Nexus 9000交换机上使用**show ip arp internal event-history event**命令来演示数据包通过思科交换矩阵服务(CFS)进行隧道传输：

```
N9K-B# show ip arp internal event-history event | i i tunnel
[116] [27772]: Tunnel Packets came with: vlan: 150, L2-SMAC :0223.e957.6a3a, L2-DMAC:
00be.758e.5677
```

```
[116] [27772]: Received tunneled packet on iod: Vlan150, physical iod: port-channel20
```

```
N9K-A# show ip arp internal event-history event | i i tunnel
[116] [28142]: Tunnel Packets sent with: vlan: 150, L2-SMAC :0223.e957.6a3a, L2-DMAC:
00be.758e.5677
```

```
[116] [28142]: Tunnel it to peer destined to remote SVI's Gateway MAC. Peer Gateway Enabled
```

您还可以在9K-B上使用**debug ip arp**系列调试命令来详细描述此行为：

```
N9K-B# debug logfile TAC_ARP
```

```
N9K-B# debug ip arp packet
```

```
N9K-B# debug ip arp event
```

```
N9K-B# debug ip arp error
```

```
N9K-B# show debug logfile TAC_ARP | beg "15:31:23"
```

```
2018 Oct 11 15:31:23.954433 arp: arp_send_request_internal: Our own address 192.0.2.3 on
interface Vlan150, sender_pid =27661
```

```
2018 Oct 11 15:31:23.955221 arp: arp_process_receive_packet_msg: Received tunneled packet on
iod: Vlan150, physical iod: port-channel20
```

```
2018 Oct 11 15:31:23.955253 arp: arp_process_receive_packet_msg: Tunnel Packets came with: vlan:
150, L2-SMAC :0223.e957.6a3a, L2-DMAC: 00be.758e.5677
```

```
2018 Oct 11 15:31:23.955275 arp: (context 1) Receiving packet from Vlan150, logical interface
Vlan150 physical interface port-channel20, (prty 6) Hrd type 1 Prot type 800 Hrd len 6 Prot len
4 OP 2, Pkt size 46
```

```
2018 Oct 11 15:31:23.955293 arp: Src 0223.e957.6a3a/192.0.2.100 Dst 00be.758e.5677/192.0.2.3
```

```
2018 Oct 11 15:31:23.955443 arp: arp_add_adj: arp_add_adj: Updating MAC on interface Vlan150,
phy-interface port-channel20, flags:0x1
```

```
2018 Oct 11 15:31:23.955478 arp: arp_adj_update_state_get_action_on_add: Different
MAC(0223.e957.6a3a) Successful action on add Previous State:0x10, Current State:0x10 Received
event:Data Plane Add, entry: 192.0.2.100, 0000.0000.0000, Vlan150, action to be taken
send_to_am:TRUE, arp_aging:TRUE
```

```
2018 Oct 11 15:31:23.955576 arp: arp_add_adj: Entry added for 192.0.2.100, 0223.e957.6a3a, state
2 on interface Vlan150, physical interface port-channel20, ismct 0. flags:0x10, Rearp (interval:
0, count: 0), TTL: 1500 seconds update_shm:TRUE
```

```
2018 Oct 11 15:31:23.955601 arp: arp_add_adj: Adj info: iod: 77, phy-iod: 91, ip: 192.0.2.100,
mac: 0223.e957.6a3a, type: 0, sync: FALSE, suppress-mode: ARP Suppression Disabled flags:0x10
```

ARP应答从主机A进入9K-A，然后通过隧道传输到9K-B。请注意，由于对等网关vPC域增强已启用，9K-A将ARP应答传送到控制平面。这会导致9K-A代表N9K-B路由数据包，即使这是非vPC VLAN。

```
N9K-A# ethanalyzer local interface inband display-filter arp limit-c 0
```

```
Capturing on inband
```

```
2018-10-11 15:32:47.378648 00:be:75:8e:56:77 -> ff:ff:ff:ff:ff:ff ARP Who has 192.0.2.100? Tell
```

```
192.0.2.3 <<<<
```

```
2018-10-11 15:32:47.379262 02:23:e9:57:6a:3a -> 00:be:75:8e:56:77 ARP 192.0.2.100 is at  
02:23:e9:57:6a:3a
```

您可以使用NX-OS的Ethanalyzer控制平面数据包捕获功能来显示9K-B的控制平面不会在本地看到此ARP应答。

```
N9K-B# ethanalyzer local interface inband display-filter arp limit-c 0
```

```
Capturing on inband
```

```
2018-10-11 15:33:30.053239 00:be:75:8e:56:77 -> ff:ff:ff:ff:ff:ff ARP Who has 192.0.2.100? Tell  
192.0.2.3
```

```
2018-10-11 15:34:16.817309 00:be:75:8e:56:77 -> ff:ff:ff:ff:ff:ff ARP Who has 192.0.2.100? Tell  
192.0.2.3
```

```
2018-10-11 15:34:42.222965 00:be:75:8e:56:77 -> ff:ff:ff:ff:ff:ff ARP Who has 192.0.2.44? Tell  
192.0.2.43
```

```
<snip>
```

注意：根据事件顺序和情况，您可能会遇到从N9K-B到主机A的数据包丢失

```
N9K-B# ping 192.0.2.100
```

```
PING 192.0.2.100 (192.0.2.100): 56 data bytes
```

```
36 bytes from 192.0.2.3: Destination Host Unreachable
```

```
Request 0 timed out
```

```
Request 1 timed out
```

```
Request 2 timed out
```

```
Request 3 timed out
```

```
Request 4 timed out
```

```
--- 192.0.2.100 ping statistics ---
```

```
5 packets transmitted, 0 packets received, 100.00% packet loss
```

当未在非vPC SVI上配置SVI用户定义MAC地址时（即使这些地址不用于通过vPC路由邻接），也会发生此行为。此行为仅适用于第一代Nexus 9000交换机。

要解决此行为，请更改受影响的SVI的MAC地址。

```
N9K-A(config)# interface Vlan150
```

```
N9K-A(config-if)# mac-address 0000.aaaa.0030
```

```
N9K-A(config-if)# end
```

```
N9K-B(config)# interface Vlan150
```

```
N9K-B(config-if)# mac-address 0000.bbbb.0030
```

```
N9K-B(config-if)# end
```

注意：由于硬件限制，一次只能为每个设备配置16个用户定义的MAC地址。这在[Cisco Nexus 9000系列NX-OS接口配置指南](#)中进行了记录。

在应用此解决方法后，您可以使用NX-OS的Ethanalyzer控制平面数据包捕获功能来显示9K-A如何从不将ARP应答推送到其控制平面。

```
N9K-A# ethanalyzer local interface inband display-filter arp limit-c 0
```

```
Capturing on inband
```

2018-10-11 15:36:11.675108 00:00:bb:bb:00:30 -> ff:ff:ff:ff:ff:ff ARP Who has 192.0.2.100? Tell 192.0.2.3

相关信息

有关第2层非vPC中继、路由邻接和SVI用户定义MAC要求的详细信息，请参阅[创建虚拟端口通道路由拓扑](#)文档。

关于此翻译

思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。