Verificar o comportamento de sincronização da tabela MAC do Nexus 9000 Series ARP & com tronco L2 não vPC

Contents

Introduction

Este documento descreve o comportamento do ARP e da tabela MAC que pode ocorrer entre os dispositivos Nexus 9000 que compartilham um tronco de Camada 2 não-vPC.

Informações de Apoio

Esse comportamento ocorre somente quando os SVIs não usam endereços MAC definidos pelo usuário e o recurso de gateway par do vPC é configurado no domínio do vPC. Além disso, ele só pode ser visto quando a tabela ARP permanece preenchida, enquanto a tabela de endereços MAC não tem uma entrada MAC para um determinado host.

O comportamento descrito neste documento é uma limitação ASIC dos switches Nexus de primeira geração e não afeta os switches Nexus 9300 Cloud Scale (EX/FX/GX/C) e posteriores e foi documentado como parte da ID de bug da Cisco CSCuh94866.

Requirements

Conhecimento geral do Virtual Port Channel (vPC), do recurso de gateway par do Virtual Port Channel do NXOS e do Nexus Operating System (NXOS).

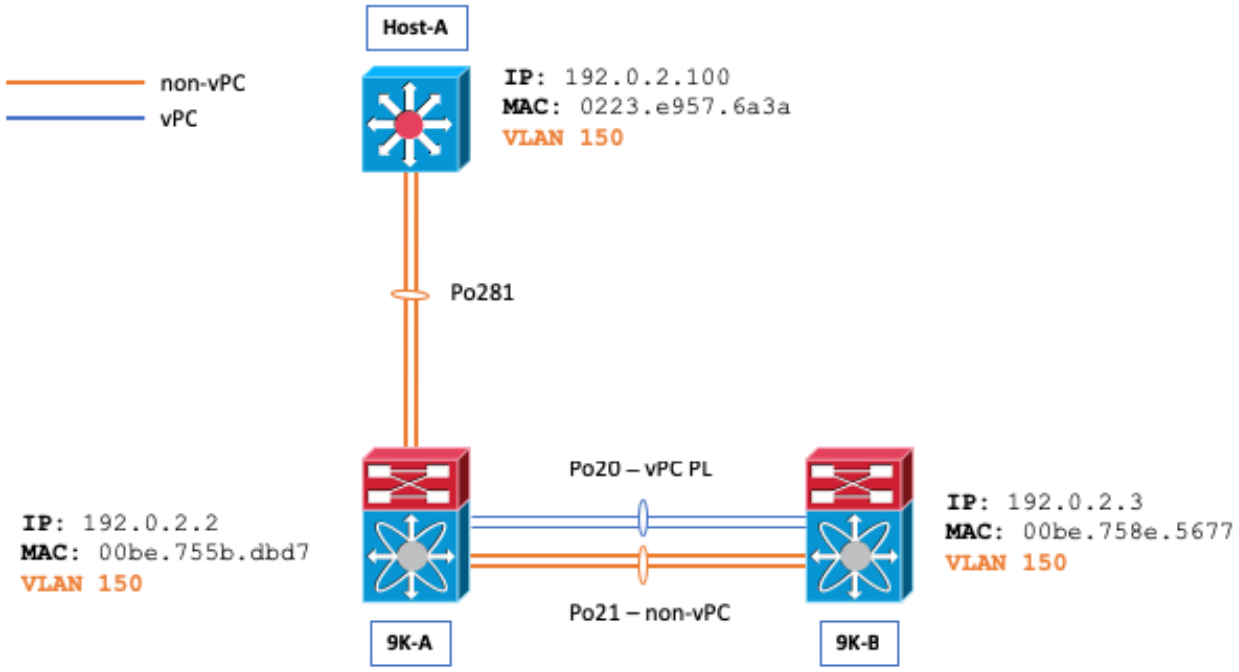
Componentes Utilizados

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

- Nexus 3000s/Nexus 9000s (somente primeira geração)
- Recurso de canal de porta virtual (vPC)
- Recurso de gateway de mesmo nível vPC

- Tronco de Camada 2 (L2) não vPC
- SVIs não vPC
- NX-OS 7.0(3)I7(5)

Topologia



Overview

Considere um cenário em que as tabelas ARP e Endereço MAC estejam vazias entre o Host-A e o N9K-B, e um ping seja iniciado do Host-A para o 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
```

O ping do Host-A faz com que o Host-A envie uma Solicitação ARP para 9K-B. A Solicitação ARP sai do Po21 no N9K-A (inundado na VLAN) enquanto também sai do Po20 (encapsulado via Cisco Fabric Services [CFS]). Como resultado, a tabela de endereços MAC em 9K-B é preenchida corretamente e uma entrada ARP é inserida na tabela ARP de N9K-B, que aponta para Po21 (o tronco L2 não-vPC) para o endereço MAC do Host-A de 0223.e957.6a3a.

```
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
```

O problema pode ser observado quando o endereço MAC para o Host-A é removido da tabela de endereços MAC de N9K-B. O endereço MAC pode ser removido por várias razões, como o envelhecimento do endereço MAC, as Topology Change Notifications (TCNs) do Spanning Tree Protocol (STP), a execução do comando **clear mac address-table dynamic** através da interface de linha de comando e assim por diante.

```
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
```

Observe que os pings ainda são bem-sucedidos; no entanto, nossa entrada ARP agora aponta para Po20 (o vPC PL) em vez de Po21, que não é o canal de porta esperado, pois a VLAN 150 é uma VLAN não VPC:

```
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.
```

Você pode usar o comando **show ip arp internal event-history event** em ambos os switches Nexus 9000 para demonstrar que os pacotes são encapsulados via Cisco Fabric Services (CFS):

```
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
```

Você também pode usar a série **debug ip arp** de comandos debug em 9K-B para detalhar esse comportamento também:

```
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
```

A Resposta ARP entra em 9K-A do Host-A e é então encapsulada em 9K-B. Observe que 9K-A envia a Resposta ARP para o plano de controle, pois o aprimoramento do domínio **peer-gateway** vPC foi habilitado. Isso faz com que 9K-A roteie o pacote em nome de N9K-B, mesmo que seja uma VLAN não-vPC.

```
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
```

Você pode usar o recurso de captura de pacotes do plano de controle do Ethanalyzer do NX-OS para mostrar que o plano de controle de 9K-B nunca vê essa Resposta ARP nativamente.

```
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>
```

> **Cuidado**: Dependendo da sequência de eventos e circunstâncias, você poderá experimentar a perda de pacotes de N9K-B para Host-A

```
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
```

Esse comportamento ocorre quando os endereços MAC definidos pelo usuário do SVI não são configurados em SVIs não vPC, mesmo quando não são usados para adjacências de roteamento sobre vPC. Esse comportamento aplica-se apenas aos switches Nexus 9000 de primeira geração.

Para contornar esse comportamento, altere o endereço MAC dos SVIs afetados.

```
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
```

> **Observação**: devido a uma limitação de hardware, você pode ter apenas 16 endereços MAC definidos pelo usuário configurados por dispositivo de cada vez. Isso está documentado no Guia de configuração de interfaces do NX-OS do Cisco Nexus 9000 Series.

Depois que a solução for aplicada, você poderá usar o recurso de captura de pacote do plano de controle do Ethanalyzer do NX-OS para mostrar como o 9K-A nunca direciona a Resposta ARP para o plano de controle.

```
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
```

Informações Relacionadas

Consulte o documento Create Topologies for Routing over Virtual Port Channel para obter mais informações sobre troncos não vPC de Camada 2, adjacências de roteamento e requisitos MAC definidos pelo usuário do SVI.

Sobre esta tradução
A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês (link fornecido) seja sempre consultado.