非vPC L2トランクでのNexus 9000シリーズARP&MACテーブルの同期動作の確認

内容

概要

このドキュメントでは、非vPCレイヤ2トランクを共有するNexus 9000デバイス間で発生する可能性があるARPテーブルとMACテーブルの動作について説明します。

背景説明

この動作は、SVIがユーザ定義のMACアドレスを使用せず、vPCピアゲートウェイ機能がvPCドメインで設定されている場合にのみ発生します。また、ARPテーブルにデータが入っている間だけ、MACアドレステーブルに特定のホストのMACエントリがない場合にも表示されます。

このドキュメントで説明されている動作は、第1世代のNexusスイッチのASIC制限であり、Nexus 9300 Cloud Scale(EX/FX/GX/C)スイッチ以降には影響しません。Cisco Bug ID CSCuh94866の一部として文書化されています。

要件

仮想ポートチャネル(vPC)、NXOS仮想ポートチャネルピアゲートウェイ機能、およびNexusオペレーティングシステム(NXOS)に関する一般知識。

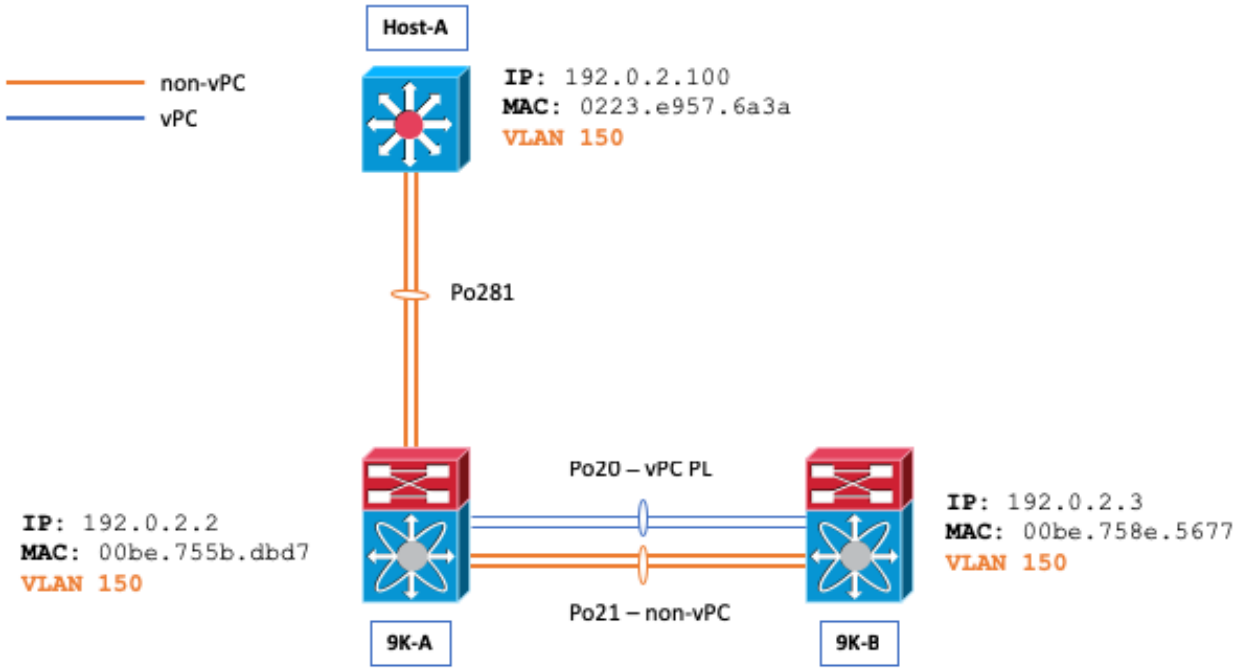
使用するコンポーネント

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期（デフォルト）設定の状態から起動しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

- Nexus 3000/Nexus 9000（第1世代のみ）
- 仮想ポートチャネル機能(vPC)
- vPCピアゲートウェイ機能
- 非vPCレイヤ2(L2)トランク

- 非vPC SVI
- NX-OS 7.0(3)I7(5)

トポロジ



概要

Host-AとN9K-Bの間のARPおよびMACアドレステーブルが空で、Host-AからN9K-Bへのpingが開始されるシナリオを考えます。

```
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
```

ホストAからのpingにより、ホストAは9K-Bに対するARP要求を送信します。ARP要求はN9K-A（VLAN上でフラッディング）のPo21から出力されますが、Po20(Cisco Fabric Services(CFS)経由でトンネリング)でも出力されます。その結果、9K-BのMACアドレステーブルは正しく入力され、N9K-BのARPテーブルにARPエントリが挿入されます。このARPエントリは、ホストAのMACアドレス0223.e957.6a3aのPo21（非vPC L2トランク）を指しています。

```
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
```

この問題は、ホストAのMACアドレスがN9K-BのMACアドレステーブルから削除されると発生する可能性があります。 MACアドレスが削除される理由は、MACアドレスのエージング、スパニングツリープロトコル(STP)トポロジ変更通知(TCN)、コマンドラインインターフェイスによる**clear mac address-table dynamic**コマンドの実行など、さまざまです。

```
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
```

pingは引き続き成功しますが、ARPエントリはPo21(vPC PL)ではなくPo20(vPC PL)を指すようになりました。VLAN 150は非VPC VLANであるため、これは予期されるポートチャネルではありません。

```
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.
```

両方のNexus 9000スイッチでshow ip arp internal event-history eventコマンドを使用して、パケットがCisco Fabric Services(CFS)経由でトンネリングされることを示すことができます。

```
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
```

また、9K-Bで一連のdebugコマンドdebug ip arpを使用して、この動作を詳細に説明することもできます。

```
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応答はホストAから9K-Aに入り、次に9K-Bにトンネリングされます。peer-gateway vPCドメイン拡張が有効になっているため、9K-AがARP応答をコントロールプレーンにパントしていることに注目してください。これにより、非vPC VLANであっても、9K-AがN9K-Bの代わりにパケットをルーティングします。

```
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
```

NX-OSのEthanalyzerコントロールプレーンパケットキャプチャ機能を使用して、9K-Bのコントロールプレーンがこのの応答をネイティブに認識しないことを示すことができます。

```
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**：イベントや状況のシーケンスによっては、N9K-BからHost-Aへのパケット損失が発生する可能性があります

```
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
```

この動作は、SVIユーザ定義MACアドレスがvPC上の隣接関係のルーティングに使用されていない場合でも、非vPC SVIで設定されていない場合に発生します。この動作は、第1世代のNexus 9000スイッチにのみ適用されます。

この動作を回避するには、影響を受けるSVIのMACアドレスを変更します。

```
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**：ハードウェアの制限により、デバイスごとに一度に設定できるユーザ定義MACアドレスは16個だけです。これは『Cisco Nexus 9000シリーズNX-OSインターフェイスコンフィギュレーションガイド』に記載されています。

回避策を適用した後、NX-OSのEthanalyzerコントロールプレーンパケットキャプチャ機能を使用して、9K-AがARP応答をコントロールプレーンにパントしない方法を示すことができます。

```
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
```

関連情報

レイヤ2の非vPCトランク、ルーティングの隣接関係、およびSVIユーザ定義MAC要件の詳細については、『仮想ポートチャネルを介したルーティングのトポロジの作成』ドキュメントを参照してください。

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。