# 仮想ポートチャネル（vPC）の拡張機能の理解

## 内容

# はじめに

このドキュメントでは、vPCドメインのCisco Nexusスイッチで設定される一般的な仮想ポートチャネル(vPC)の拡張機能について説明します。

# 前提条件

## 要件

仮想ポートチャネル（vPC）の使用例、設定、および実装に関する基本情報を理解しておくことをお勧めします。この機能の詳細については、次の該当するドキュメントのいずれかを参照してください。

- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.3(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.2(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.1(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.3(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.2(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 7.x』
- 『Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide 8.x』
- 『Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide 7.x』
- 設計および設定ガイド：Cisco Nexus 7000シリーズスイッチの仮想ポートチャネル(vPC)のベストプラクティス

## 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、クリアな（デフォルト）設定で作業を開始しています。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。

# 背景説明

Cisco Nexus データセンター用スイッチでの Cisco NX-OS の登場以来、仮想ポートチャネル（vPC）機能は何度も拡張されています。それにより、障害シナリオでの vPC 接続デバイスの信

頼性が向上し、両方の vPC ピアスイッチの転送動作も最適化されています。各機能拡張の目的、機能拡張による動作の変更、および機能拡張によって解決される障害シナリオを理解することは、ビジネスニーズと要件を最大限に満たすために機能拡張を vPC ドメイン内で設定する必要がある理由とそのタイミングを理解するために役立ちます。

### 該当ハードウェア

このドキュメントで説明する手順は、vPC 対応のすべての Cisco Nexus データセンター用スイッチに適用できます。

## vPC ピア スイッチ

ここでは、vPC ドメイン設定コマンドの peer-switch で有効にする vPC ピアスイッチの機能拡張について説明します。

### 概要

多くの環境では、vPC ドメイン内の Nexus スイッチのペアは、レイヤ 2 スイッチド イーサネット ドメインとレイヤ 3 ルーテッドドメインの間の境界として機能するアグリゲーションスイッチまたはコアスイッチです。両方のスイッチは、複数の VLAN によって設定されており、VLAN 間の East-West トラフィックと North-South トラフィックのルーティングを担当します。これらの環境では、Nexus スイッチは、通常、スパニングツリープロトコルの観点からルートブリッジとしても機能します。

通常、1つのvPCピアは、スパニングツリーのプライオリティを0などの低い値に設定することによって、スパニングツリーのルートブリッジとして設定されます。もう一方のvPCピアは、4096などのやや高いスパニングツリープライオリティを使用して設定されており、ルートブリッジとして機能するvPCピアに障害が発生した場合に、スパニングツリー内でルートブリッジの役割を引き継ぐことができます。この構成では、ルートブリッジとして機能する vPC ピアは、システム MAC アドレスを含むブリッジ ID を持つスパニング ツリー ブリッジ プロトコル データ ユニット（BPDU）を発信します。

ただし、ルートブリッジとして機能するvPCピアに障害が発生し、他のvPCピアがスパニングツリールートブリッジとして引き継ぐ場合、他のvPCピアは、自身のシステムMACアドレスを含むブリッジIDでスパニングツリーBPDUを発信します。このブリッジIDは、元のルートブリッジのシステムMACアドレスとは異なります。この変更による影響は、ダウンストリームブリッジの接続方法によって異なります。詳細は次のサブセクションを参照してください。

#### 冗長接続の非 vPC ブリッジ

BPDUの変更（およびルートブリッジの変更）を検出する冗長リンク（スパニングツリープロトコルの観点から見て、1つのリンクがブロッキング状態にあるような）を使用して両方のvPCピアに接続されている非vPC接続ブリッジは、ルートポートの変更を監視します。他の指定フォワーディングインターフェイスはただちにブロッキングステートに移行し、設定されたスパニングツリープロトコル転送遅延タイマー（デフォルトでは15秒）と同等の間隔で、スパニングツリープロトコルの有限状態マシン（ブロッキング、ラーニング、フォワーディング）を通過します。

ルートポートの変更とそれに続くスパニングツリープロトコル有限ステートマシンの遷移は、ネットワーク内で大きな中断を引き起こす可能性があります。vPC ピアスイッチの機能拡張は、主に、vPC ピアの一方がオフラインになったときにこの問題によって引き起こされるネットワークの中断を防ぐために導入されました。vPCピアスイッチ拡張機能を使用すると、vPCに接続されていないブリッジは、ブロッキングステートにある単一の冗長リンクを保持しますが、リンク障害によって既存のルートポートがダウンした場合には、そのインターフェイスを即座にフォワーディングステートに移行します。オフラインvPCピアがオンラインに戻っても同じプロセスが発生します。ルートブリッジへのコストが最も低いインターフェイスがルートポートの役割を奪い、冗長リンクは即座にブロッキングステートに移行します。データプレーンに発生する唯一の影響は、オフラインになったvPCピアを通過する途中のパケットが、やむを得ず失われることです。

vPC 接続ブリッジ

スパニングツリードメイン内のvPC接続ブリッジは、BPDUの変更（およびルートブリッジの変更）を検出し、動的に学習されたMACアドレスをローカルMACアドレステーブルからフラッシュします。ループのないトポロジでは、スパニングツリープロトコルに依存しないvPC接続デバイスのトポロジでは、この動作は非効率的で不要です。vPCは、スパニングツリープロトコルの観点からは通常のポートチャネルと同様に単一の論理インターフェイスと見なされるため、vPCピアの損失はポートチャネルメンバー内の単一リンクの損失と似ています。どちらのシナリオでも、スパニングツリーは変更されないため、スパニングツリードメイン内のブリッジから動的に学習された MAC アドレスのフラッシュ（その目的は、イーサネットのフラッドアンドラーニング動作で、スパニングツリーの新たに転送するインターフェイスの MAC アドレスを再学習できるようにすること）は不要です。

さらに、動的に学習された MAC アドレスのフラッシュは、中断を引き起こす可能性があります。2 つのホストがほぼ一方向の UDP ベースのフローを持っているシナリオ（TFTP サーバーにデータを送信する TFTP クライアントなど）について考えてみます。このフローでは、大半のデータが、TFTP クライアントから TFTP サーバーに送信されます。TFTP サーバーが TFTP クライアントにパケットを送り返すことはほとんどありません。その結果、スパニングツリードメインで動的に学習されたMACアドレスがフラッシュされた後、TFTPサーバのMACがしばらくの間学習されません。つまり、TFTPサーバに送信されるTFTPクライアントのデータは、トラフィックが不明なユニキャストトラフィックであるため、VLAN全体にフラッディングされます。これにより、大量のデータフローがネットワーク内の意図しない場所に移動し、ネットワークのオーバーサブスクライブセクションを通過する場合にパフォーマンスの問題が発生する可能性があります。

vPC ピアスイッチの機能拡張は、1 つ以上の VLAN のスパニングツリールートブリッジとして機能する vPC ピアがリロードされたり、その電源がオフになった場合に、この非効率的で不要な動作が発生することを防ぐために導入されました。

vPCピアスイッチ機能拡張を有効にするには、両方のvPCピアが同一のスパニングツリープロトコル設定（すべてのvPC VLANのスパニングツリープライオリティ値を含む）を持ち、すべてのvPC VLANのルートブリッジになる必要があります。これらの前提条件が満たされた後に、vPCドメイン設定コマンドの peer-switch を設定して、vPC ピアスイッチの機能拡張を有効にする必要があります。

注:vPCピアスイッチ拡張機能は、すべてのVLANのルートを含むvPCドメインでのみサポートされます。

vPCピアスイッチ機能拡張を有効にすると、両方のvPCピアが、両方のvPCピアで共有されるvPCシステムのMACアドレスを含むブリッジIDを使用して、同一のスパニングツリーBPDUの生成を開始します。vPCピアがリロードされても、残りのvPCピアから発信されたスパニングツリーBPDUは変更されないため、スパニングツリードメイン内の他のブリッジはルートブリッジの変更を認識せず、ネットワークの変更に対して最適に反応しません。

警告

vPC ピアスイッチの機能拡張には、実稼働環境で設定する前に知っておく必要のあるいくつかの注意事項があります。

vPC ピア間でのスパニングツリー優先順位値の一致が必要

vPC ピアスイッチの機能拡張を有効にする前に、すべての vPC VLAN のスパニングツリー優先順位設定を変更して、両方の vPC ピア間で同一になるようにする必要があります。

ここで、N9K-1が、プライオリティ0のVLAN 1、10、および20のスパニングツリールートブリッジとして設定されている設定について考えます。N9K-2は、VLAN 1、10、および20のセカンダリスパニングツリールートブリッジで、プライオリティは4096です。

```
<#root>

N9K-1#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 0
interface port-channel1
  spanning-tree port type network

N9K-2#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 4096
interface port-channel1
  spanning-tree port type network
```

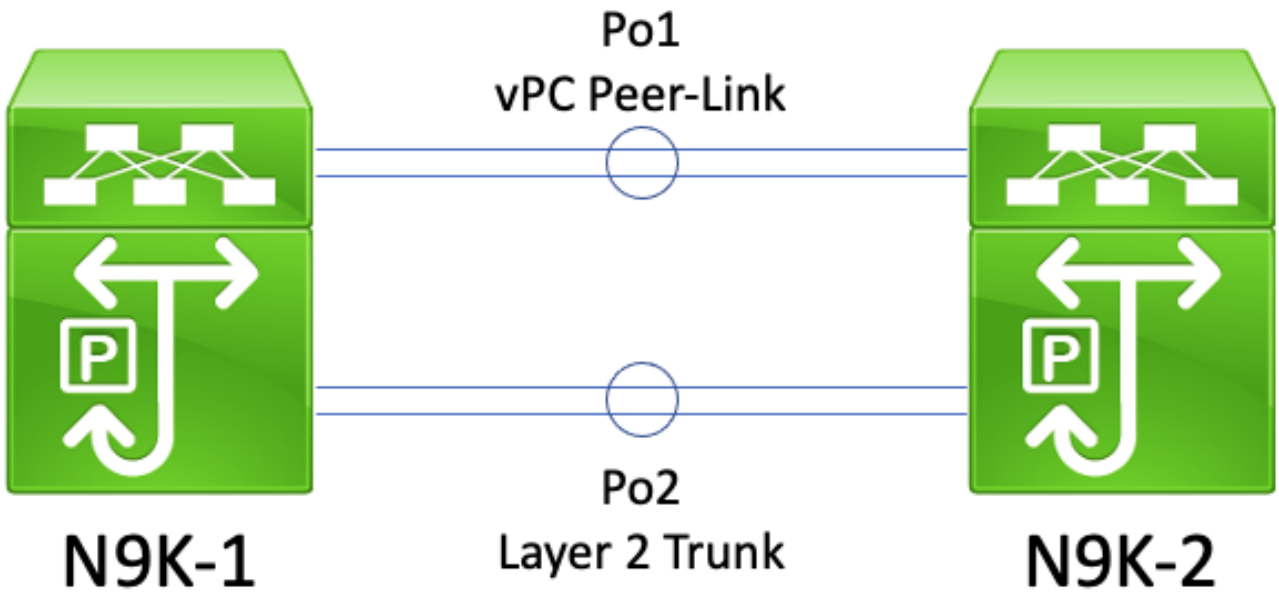
vPCピアスイッチ機能拡張を有効にする前に、N9K-2上のVLAN 1、10、および20のスパニングツリープライオリティ設定を、N9K-1上の同じVLANのスパニングツリープライオリティ設定と一致するように変更する必要があります。この変更の例を次に示します。

```
<#root>

N9K-2#

configure terminal
```

```
Enter configuration commands, one per line. End with CNTL/Z.
N9K-2(config)#

spanning-tree vlan 1,10,20 priority 0

N9K-2(config)#

end

N9K-2#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 0
interface port-channel1
  spanning-tree port type network

N9K-1#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 0
interface port-channel1
  spanning-tree port type network
```

非 vPC VLAN への vPC ピアスイッチの影響

次のトポロジについて考えてみます。




このトポロジでは、2つのvPCピア（N9K-1とN9K-2）の間に2つのレイヤ2トランク（Po1とPo2）があります。Po1はvPC VLANを伝送するvPCピアリンクで、Po2はすべての非vPC VLANを伝送するレイヤ2トランクです。Po2で伝送される非vPC VLANのスパニングツリーのプライオリティ値がN9K-1とN9K-2で同じ場合、各vPCピアが、両方のスイッチで同じvPCシステムMACアドレスからスパニングツリーBPDUフレームを発信します。その結果、N9K-2がスパニングツリーBPDUを発信したスイッチであっても、N9K-1は非vPC VLANごとにPo2で独自のスパニ

ングツリーBPDUを受信しているように見えます。スパニングツリーの観点からは、N9K-1はすべての非vPC VLANに対してPo2をブロッキングステートにします。

これは正常な動作です。この動作の発生を防ぐ、またはこの問題を回避するには、すべての非vPC VLAN で、両方の vPC ピアを異なるスパニングツリー優先順位値で設定する必要があります。これにより、1つのvPCピアが非vPC VLANのルートブリッジになり、vPCピア間のレイヤ2トランクをDesignated Forwarding状態に移行できます。同様に、リモートvPCピアは、vPCピア間のレイヤ2トランクを指定ルート状態に移行します。これにより、非vPC VLANのトラフィックは、レイヤ2トランクを介して両方のvPCピアに流れるようになります。

## コンフィギュレーション

ここでは、vPC ピアスイッチ機能を設定する方法の例を示します。

この例では、N9K-1がプライオリティ0のVLAN 1、10、および20のスパニングツリールートブリッジになるように設定されています。N9K-2は、VLAN 1、10、および20のセカンダリスパニングツリールートブリッジで、プライオリティは4096です。

<#root>

```
N9K-1#

show running-config vpc

<snip>
vpc domain 1
  role priority 150
  peer-keepalive destination 10.122.190.196

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

<snip>
vpc domain 1
  peer-keepalive destination 10.122.190.195

interface port-channel1
  vpc peer-link

N9K-1#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 0
interface port-channel1
  spanning-tree port type network

N9K-2#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 4096
interface port-channel1
  spanning-tree port type network
```

まず、N9K-2 のスパニングツリーの優先順位設定を、N9K-1 の設定と同じになるように変更する必要があります。これは、vPC ピアスイッチ機能が予期どおりに機能するための要件です。N9K-2のシステムMACアドレスがN9K-1のシステムMACアドレスよりも小さい場合、N9K-2はスパニングツリードメインのルートブリッジの役割を奪い、スパニングツリードメイン内の他のブリッジは該当するすべてのVLANのローカルMACアドレステーブルをフラッシュします。この現象の例を、次に示します。

<#root>

```
N9K-1#

show spanning-tree vlan 1


VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     689e.0baa.dea7
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    1      (priority 0 sys-id-ext 1)
             Address     689e.0baa.dea7
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC) P2p

N9K-2#

show spanning-tree vlan 1


VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     689e.0baa.dea7
             Cost        1
             Port        4096 (port-channel1)
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    4097   (priority 4096 sys-id-ext 1)
             Address     689e.0baa.de07
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Root FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC) P2p

N9K-2#

configure terminal

Enter configuration commands, one per line. End with CNTL/Z.
```

```
N9K-2(config)#
spanning-tree vlan 1,10,20 priority 0
N9K-2(config)#
end
N9K-2#
show spanning-tree vlan 1


VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     689e.0baa.de07
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    1      (priority 0 sys-id-ext 1)
             Address     689e.0baa.de07
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC) P2p
```

次に、vPC ドメイン設定コマンドの peer-switch を使用して vPC ピアスイッチ機能を有効にすることができます。これにより、両方のvPCピアから発信されたスパニングツリーBPDU内のブリッジIDが変更され、スパニングツリードメイン内の他のブリッジでは、該当するすべてのVLANのローカルMACアドレステーブルがフラッシュされます。

<#root>

```
N9K-1#
configure terminal
N9K-1(config)#
vpc domain 1
N9K-1(config-vpc-domain)#
peer-switch
N9K-1(config-vpc-domain)#
end
N9K-1#

N9K-2#
configure terminal
N9K-2(config)#
vpc domain 1
```

```
N9K-2(config-vpc-domain)#
peer-switch
N9K-2(config-vpc-domain)#
end
N9K-2#
```

show spanning-tree summary コマンドを使用して、両方の vPC ピアが vPC VLAN のルートブリッジであると主張していることを検証することにより、vPC ピアスイッチ機能が予期どおりに動作していることを確認できます。この出力には、vPC ピアスイッチ機能が有効になっており、動作可能であることも示されている必要があります。

<#root>

```
N9K-1#
show spanning-tree summary

Switch is in rapid-pvst mode
Root bridge for: VLAN0001, VLAN0010, VLAN0020
L2 Gateway STP                          is disabled
Port Type Default                       is disable
Edge Port [PortFast] BPDU Guard Default  is disabled
Edge Port [PortFast] BPDU Filter Default is disabled
Bridge Assurance                        is enabled
Loopguard Default                       is disabled
Pathcost method used                    is short
vPC peer-switch                         is enabled (operational)
STP-Lite                                is disabled

Name                   Blocking Listening Learning Forwarding STP Active
---------------------- -------- --------- -------- ---------- ----------
VLAN0001                     0         0        0          3          3
VLAN0010                     0         0        0          3          3
VLAN0020                     0         0        0          3          3
---------------------- -------- --------- -------- ---------- ----------
3 vlans                      0         0        0          9          9

N9K-2#
show spanning-tree summary

Switch is in rapid-pvst mode
Root bridge for: VLAN0001, VLAN0010, VLAN0020
L2 Gateway STP                          is disabled
Port Type Default                       is disable
Edge Port [PortFast] BPDU Guard Default  is disabled
Edge Port [PortFast] BPDU Filter Default is disabled
Bridge Assurance                        is enabled
Loopguard Default                       is disabled
Pathcost method used                    is short
vPC peer-switch                         is enabled (operational)
STP-Lite                                is disabled

Name                   Blocking Listening Learning Forwarding STP Active
---------------------- -------- --------- -------- ---------- ----------
```

```
VLAN0001                     0         0        0          3          3
VLAN0010                     0         0        0          3          3
VLAN0020                     0         0        0          3          3
---------------------- -------- --------- -------- ---------- ----------
3 vlans                      0         0        0          9          9
```

特定の VLAN に関する詳細情報を確認するには、show spanning-tree vlan{x} コマンドを使用します。プライマリまたは動作可能なプライマリvPCロールを持つスイッチは、そのすべてのインターフェイスがDesignated Forwarding状態になります。セカンダリvPCロールまたは動作可能セカンダリvPCロールを保持しているスイッチは、vPCピアリンクを除くすべてのインターフェイスをDesignated Forwarding状態にします。ただし、vPCピアリンクはRoot Forwarding状態です。show vpc role の出力に示される vPC システム MAC アドレスが、各 vPC ピアのルートブリッジ ID およびブリッジ ID と同じであることに注意してください。

<#root>

```
N9K-1#

show vpc role


vPC Role status
----------------------------------------------------
vPC role                        : primary
Dual Active Detection Status    : 0
vPC system-mac                  : 00:23:04:ee:be:01
vPC system-priority             : 32667
vPC local system-mac            : 68:9e:0b:aa:de:a7
vPC local role-priority         : 150
vPC local config role-priority  : 150
vPC peer system-mac             : 68:9e:0b:aa:de:07
vPC peer role-priority          : 32667
vPC peer config role-priority   : 32667

N9K-1#

show spanning-tree vlan 1


VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     0023.04ee.be01
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    1      (priority 0 sys-id-ext 1)
             Address     0023.04ee.be01
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Desg FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC) P2p

N9K-2#
```

```
show vpc role


vPC Role status
----------------------------------------------------
vPC role                        : secondary
Dual Active Detection Status    : 0
vPC system-mac                  : 00:23:04:ee:be:01
vPC system-priority             : 32667
vPC local system-mac            : 68:9e:0b:aa:de:07
vPC local role-priority         : 32667
vPC local config role-priority  : 32667
vPC peer system-mac             : 68:9e:0b:aa:de:a7
vPC peer role-priority          : 150
vPC peer config role-priority   : 150

N9K-2#

show spanning-tree vlan 1


VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
             Address     0023.04ee.be01
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    1      (priority 0 sys-id-ext 1)
             Address     0023.04ee.be01
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Po1              Root FWD 1         128.4096 (vPC peer-link) Network P2p
Po10             Desg FWD 1         128.4105 (vPC) P2p
Po20             Desg FWD 1         128.4115 (vPC) P2p
```

最後に、いずれかの vPC ピアで Ethanalyzer コントロール プレーン パケット キャプチャ ユーティリティを使用すると、両方の vPC ピアが両方の vPC ピア間で共有される vPC システム MAC アドレスを含むブリッジ ID およびルートブリッジ ID を持つスパニングツリー BPDU を発信していることを確認できます。

<#root>

```
N9K-1#

ethanalyzer local interface inband display-filter stp limit-captured-frames 0

<snip>
Capturing on inband
2021-05-13 01:59:51.664206 68:9e:0b:aa:de:d4 -> 01:80:c2:00:00:00 STP RST. Root = 0/1/00:23:04:ee:be:01

N9K-2#

ethanalyzer local interface inband display-filter stp limit-captured-frames 0

<snip>
Capturing on inband
```

```
2021-05-13 01:59:51.777034 68:9e:0b:aa:de:34 -> 01:80:c2:00:00:00 STP RST. Root = 0/1/00:23:04:ee:be:01
```

## 影響

vPCピアスイッチ拡張機能を有効にした場合の影響は、スパニングツリードメイン内の他のブリッジがvPC経由で両方のvPCピアに接続されているか、vPCを使用せずに両方のvPCピアに冗長接続されているかによって異なります。

### 冗長接続の非 vPC ブリッジ

両方の vPC ピアへの冗長リンク（スパニングツリープロトコルの観点からは一方のリンクが Blocking 状態）を持つ非 vPC 接続ブリッジが、スパニングツリー BPDU でアドバタイズされたスパニング ツリー ルート ブリッジの変更を検出した場合、ブリッジのルートポートが 2 つの冗長インターフェイス間で変更される可能性があります。次に、これにより、その他の指定転送インターフェイスが、すぐに Blocking 状態に移行し、設定されたスパニングツリープロトコル転送遅延タイマー（デフォルトでは 15 秒）に相当する間隔でスパニングツリープロトコル有限ステートマシン（ブロッキング、ラーニング、および転送）を遷移します。ルートポートの変更とそれに続くスパニングツリープロトコル有限ステートマシンの遷移は、ネットワーク内で大きな中断を引き起こす可能性があります。

この影響は、現在スパニングツリードメインのルートブリッジになっているvPCピアがオフライン（停電、ハードウェア障害、リロードなど）になると発生することに注意してください。この動作は、vPC ピアスイッチの機能拡張に固有のものではありません。vPC ピアスイッチの機能拡張を有効にすると、スパニングツリーの観点から vPC ピアがオフラインになるのと同様の動作が発生します。

### vPC 接続ブリッジ

vPC接続ブリッジは、スパニングツリーBPDUでアドバタイズされたスパニングツリールートブリッジの変更を検出すると、動的に学習されたMACアドレスをMACアドレステーブルからフラッシュします。vPCピアスイッチ機能を設定する際に、次の2つのシナリオでこの動作を確認できます。

1. スパニングツリー優先順位値が両方の vPC ピアの間で一致するように設定されている場合、以前はルートブリッジではなかった vPC ピアのシステム MAC アドレスが以前にルートブリッジであった vPC ピアよりも低いと、スパニング ツリー ルート ブリッジが一方の vPC ピアからもう一方の vPC ピアに変わる可能性があります。このシナリオの例は、このドキュメントの「vPC ピアスイッチ」にある「設定」セクションに示されています。
2. peer-switch vPCドメイン設定コマンドでvPCピアスイッチ機能を有効にすると、両方のvPCピアがスパニングツリードメインのルートブリッジとして動作を開始します。両方のvPCピアが、スパニングツリードメインのルートブリッジとして自身をアサートする、同一のスパニングツリーBPDUの生成を開始します。

ほとんどのシナリオとトポロジでは、これら2つのシナリオのいずれでもデータプレーンへの影響は見られません。ただし、短時間の間は、フレームの宛先MACアドレスがダイナミックに学習されたMACアドレスのフラッシュの直接的な結果としてどのスイッチポートでも学習されないため

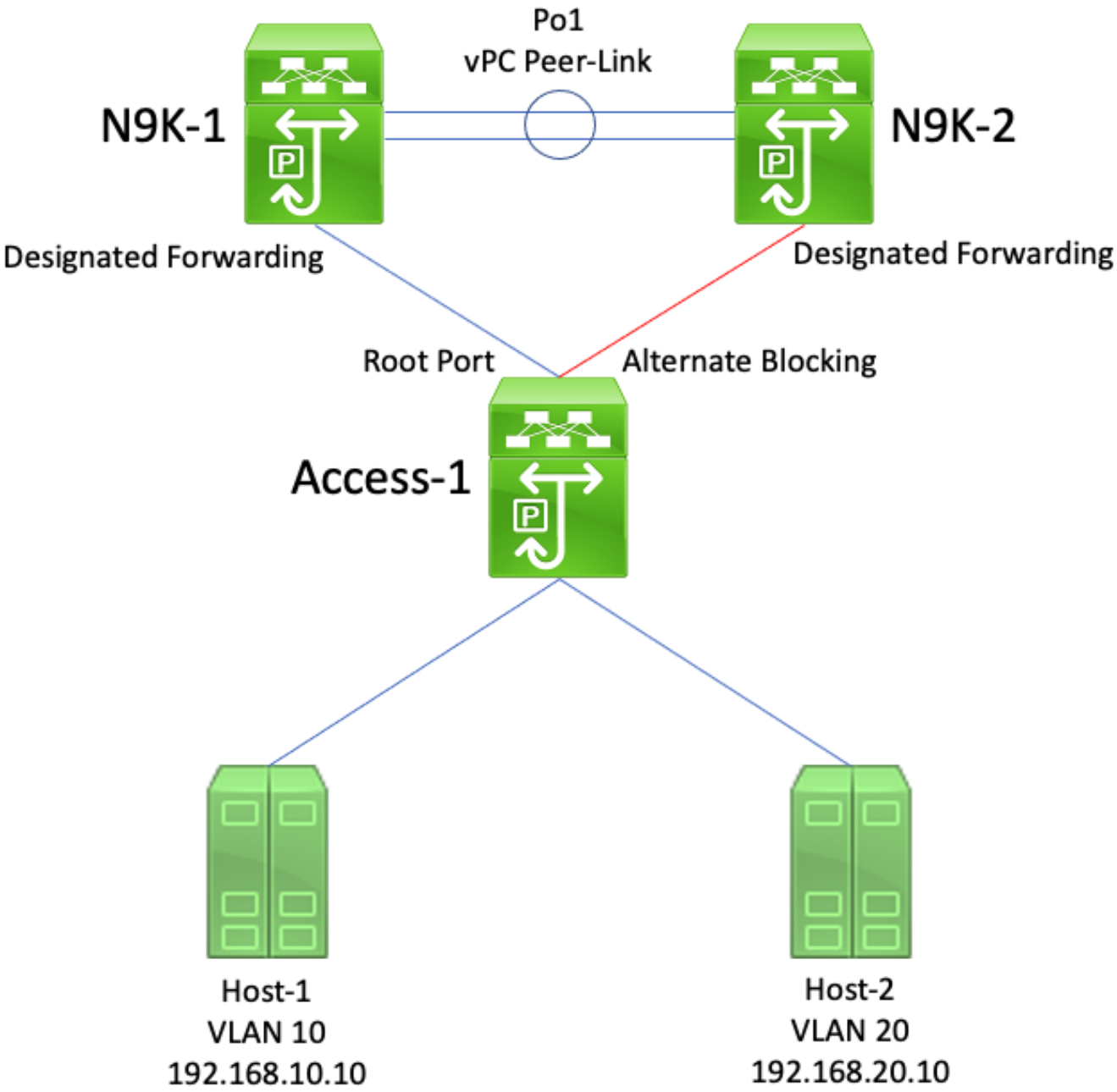
、データプレーントラフィックは未知のユニキャストフラッディングが原因でVLAN内でフラッディングされます。これにより、一部のトポロジでは、VLAN 内のオーバーサブスクライブされたネットワークデバイスへのデータプレーントラフィックのフラッディングが発生する場合に、短期間のパフォーマンスの問題またはパケット損失が発生する可能性があります。また、通常のように宛先ホストに直接スイッチングされるのではなく、長時間VLAN内でトラフィックがフラッディングされるため、帯域幅を大量に消費する単方向トラフィックフローやサイレントホスト（主にパケットを受信し、ほとんどパケットを送信しないホスト）で問題が発生する可能性があります。

この影響は、影響を受けるVLAN内のブリッジのMACアドレステーブルから動的に学習されたMACアドレスのフラッシュに関連していることに注意してください。この動作は、vPC ピアスイッチの機能拡張やルートブリッジの変更に固有のものではなく、VLAN 内で非エッジポートが起動することで生成されるトポロジ変更通知が原因である場合もあります。

## 障害シナリオの例

### 有限ステートマシンを再起動させる冗長接続の非 vPC ブリッジ

次のトポロジについて考えてみます。




このトポロジでは、N9K-1 と N9K-2 が vPC ドメイン内 vPC ピアです。N9K-1 は、すべての VLAN のスパニングツリー優先順位値が 0 に設定されているため、すべての VLAN のルートブリッジになります。N9K-2 は、すべての VLAN のスパニングツリー優先順位値が 4096 に設定されているため、すべての VLAN のセカンダリルートブリッジになります。Access-1 は、レイヤ 2 スイッチポートを介して N9K-1 と N9K-2 の両方に冗長接続されているスイッチです。これらのスイッチポートはポートチャネルにバンドルされていないため、スパニングツリープロトコルにより、N9K-1 に接続されているリンクは Designated Root 状態になり、N9K-2 に接続されているリンクは Alternate Blocking 状態にします。

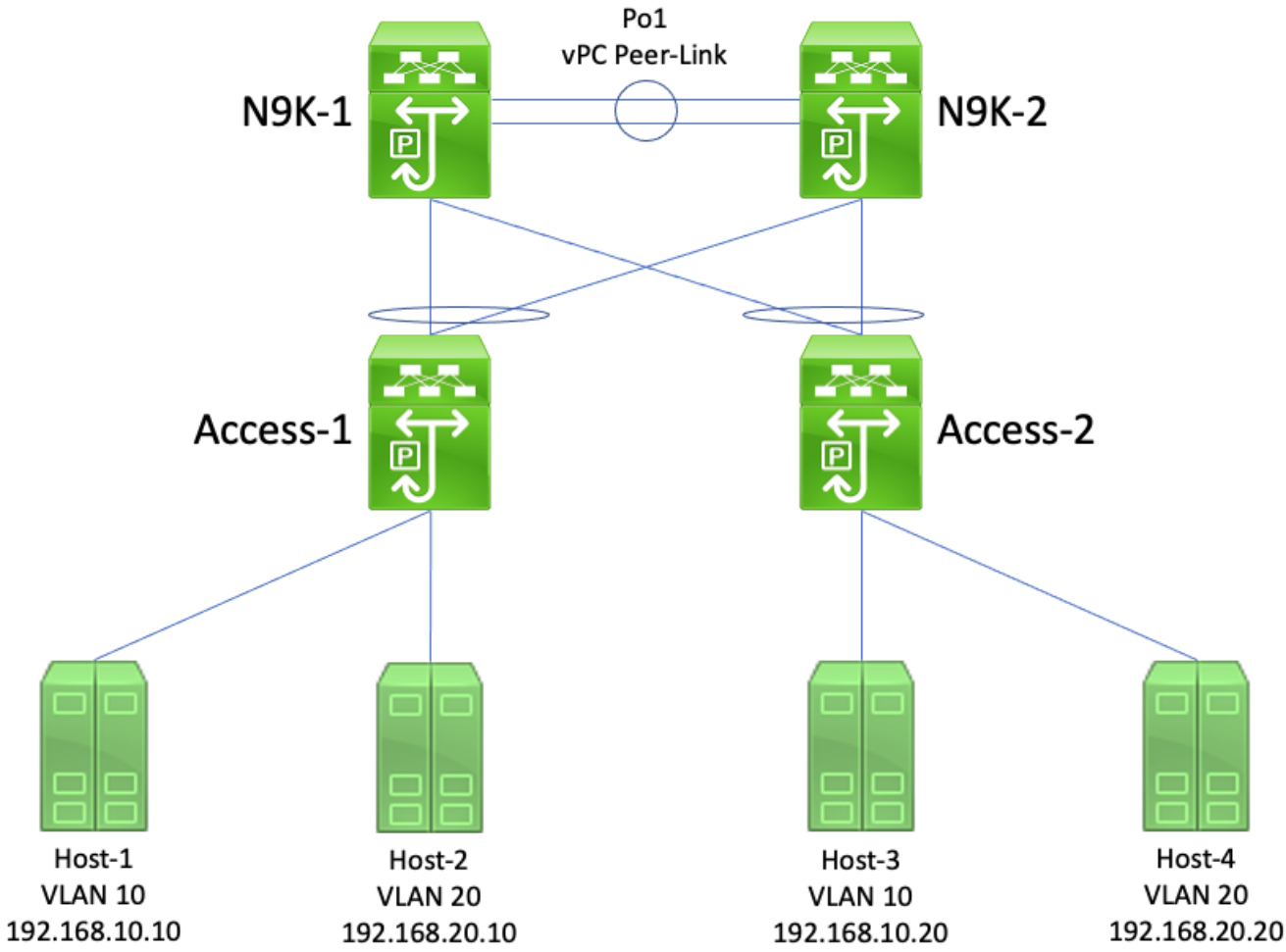
ハードウェア障害、電源障害、またはスイッチのリロードが原因で N9K-1 がオフラインになる障害シナリオについて考えてみます。N9K-2は、自身のシステムMACアドレスをブリッジIDとして使用してスパニングツリーBPDUをアドバタイズすることにより、自身がすべてのVLANのルートブリッジであることを示します。Access-1はルートブリッジのIDの変更を認識します。さらに、その指定ルートポートはdown/down状態に移行します。つまり、新しい指定ルートポートは、N9K-2に面した代替ブロッキング状態にあったリンクです。

この指定ルートポートの変更により、すべての非エッジスパニングツリーポートは、設定されたスパニングツリープロトコル転送遅延タイマー（デフォルトでは15秒）と同等の間で一時停止しながら、スパニングツリープロトコルの有限状態マシン（ブロッキング、ラーニング、フォワーディング）を通過します。このプロセスは、ネットワークに大きな中断を引き起こす可能性があります。

vPCピアスイッチ機能拡張を有効にした同じ障害シナリオでは、N9K-1とN9K-2の両方が、共有vPCシステムのMACアドレスをブリッジIDとして使用して、同一のスパニングツリーBPDUを送信します。N9k-1に障害が発生すると、N9K-2はこの同じスパニングツリーBPDUの送信を続行します。その結果、Access-1はただちにに N9K-2への代替ブロッキングリンクを指定ルート状態に移行し、リンクを介したトラフィックの転送を開始します。さらに、スパニング ツリー ルート ブリッジ ID が変更されないという事実により、非エッジポートのスパニングツリープロトコル有限ステートマシンの遷移が防止され、ネットワークで観測される中断が減少します。

動的に学習された MAC アドレスをフラッシュする vPC 接続ブリッジ

次のトポロジについて考えてみます。




このトポロジでは、N9K-1とN9K-2はvPCドメイン内のvPCピアであり、VLAN 10とVLAN 20の間でVLAN間ルーティングを実行します。N9K-1は、VLAN 10とVLAN 20に対してスパニングツリーのプライオリティ値を0に設定し、N9K-1を両方のVLANのルートブリッジにします。N9K-2 は、VLAN 10 と VLAN 20 のスパニングツリー優先順位値が 4096 に設定されているため、両方の VLAN のセカンダリルートブリッジになります。Host-1、Host-2、Host-3、および Host-4 はすべ

て、相互に継続的に通信しています。

ハードウェア障害、電源障害、またはスイッチのリロードが原因で N9K-1 がオフラインになる障害シナリオについて考えてみます。N9K-2は、自身のシステムMACアドレスをブリッジIDとして使用してスパニングツリーBPDUをアドバタイズすることで、自身をVLAN 10とVLAN 20のルートブリッジとして主張します。Access-1とAccess-2はルートブリッジのIDの変更を認識し、スパニングツリーは同じままですが（つまり、N9K-1とN9K-2に面するvPCは指定ルートポートのままになります）、Access-1とAccess-2は、VLAN 10とVLAN 20で動的に学習されたすべてのMACアドレスのMACアドレスをフラッシュします。

ほとんどの環境では、動的に学習された MAC アドレスのフラッシュは、最小限の影響しか引き起こしません。パケットの損失はありません（障害発生時に N9K-1 に送信されたパケットの損失を除く）が、ブロードキャストドメイン内のすべてのスイッチが動的 MAC アドレスを再学習している間、不明なユニキャストトラフィックとして各ブロードキャストドメイン内でトラフィックの一時的なフラッディングが発生します。

同じ障害シナリオで、vPC ピアスイッチの機能拡張が有効になっている場合は、N9K-1 と N9K-2 の両方が共有 vPC システム MAC アドレスをブリッジ ID として使用して、同一のスパニングツリー BPDU を送信します。N9k-1に障害が発生すると、N9K-2はこの同じスパニングツリーBPDUの送信を続行します。その結果、Access-1とAccess-2はスパニングツリートポロジの変更が発生したことを認識しません。Access-1とAccess-2の観点から見ると、ルートブリッジのスパニングツリーBPDUは同一であるため、関連するVLANから動的に学習されたMACアドレスをフラッシュする必要はありません。これにより、この障害シナリオで、各ブロードキャストドメインにおける不明なユニキャストトラフィックのフラッディングが防止されます。

# vPC ピアゲートウェイ

ここでは、vPC ドメイン設定コマンドの peer-gateway で有効にする vPC ピアゲートウェイの機能拡張について説明します。

## 概要

vPC ドメインで設定された Nexus スイッチは、デフォルトでデュアルアクティブ First Hop Redundancy Protocol（FHRP）転送を実行します。つまり、いずれかのvPCピアが、スイッチで設定されているホットスタンバイルータプロトコル(HSRP)または仮想ルータ冗長プロトコル(VRRP)グループに属する宛先MACアドレスを持つパケットを受信した場合、スイッチは、HSRPまたはVRRPコントロールプレーンの状態に関係なく、ローカルルーティングテーブルに従ってパケットをルーティングします。言い換えると、HSRP Standby または VRRP Backup 状態の vPC ピアの期待される動作は、HSRP または VRRP 仮想 MAC アドレス宛てのパケットのルーティングです。

vPCピアがFHRP仮想MACアドレス宛てのパケットをルーティングする際に、パケットは新しい送信元および宛先MACアドレスで書き換えられます。送信元MACアドレスは、パケットがルーティングされるVLAN内のvPCピアのスイッチ仮想インターフェイス(SVI)のMACアドレスです。宛先MACアドレスは、vPCピアのローカルルーティングテーブルに従って、パケットの宛先IPアドレスのネクストホップIPアドレスに関連付けられたMACアドレスです。VLAN間ルーティングのシナリオでは、パケットが書き換えられた後のパケットの宛先MACアドレスは、パケットの最終

的な宛先であるホストのMACアドレスです。

一部のホストは、最適化機能のために標準の転送動作に従いません。このように動作するホストは、着信パケットに応答するときに、ルーティングテーブルや ARP キャッシュのルックアップを実行しません。代わりに、ホストは、応答パケット用に着信パケットの送信元 MAC アドレスと宛先 MAC アドレスを反転させます。つまり、着信パケットの送信元 MAC アドレスが応答パケットの宛先 MAC アドレスになり、着信パケットの宛先 MAC アドレスが応答パケットの送信元 MAC アドレスになります。この動作は、ローカルルーティングテーブルや ARP キャッシュのルックアップを実行し、応答パケットの宛先 MAC アドレスを FHRP 仮想 MAC アドレスに設定する標準の転送動作に従うホストとは異なります。

この非標準のホスト動作では、ホストによって生成された応答パケットが、一方の vPC ピア宛てであるものの、その vPC を出て、もう一方の vPC ピアに向かう場合、vPC ループ回避ルールに違反する可能性があります。もう一方のvPCピアは、自身のvPCピアが所有するMACアドレス宛てのパケットを受信し、パケットの宛先MACアドレスフィールドに存在するMACアドレスを所有するvPCピアに向けてvPCピアリンクからパケットを転送します。MACアドレスを所有するvPCピアは、ローカルでのパケットのルーティングを試みます。パケットがvPCから出力される必要がある場合、vPCピアはvPCループ回避ルールに違反してこのパケットをドロップします。その結果、この非標準の動作を利用してホストと送受信される一部のフローについて、接続の問題またはパケット損失が発生する可能性があります。

この非標準の動作を利用するホストによって発生するパケット損失を排除するために、vPC ピアゲートウェイの機能拡張が導入されました。これは、一方の vPC ピアがもう一方の vPC ピアの MAC アドレス宛てのパケットをローカルにルーティングできるようにすることで実現されます。これにより、リモート vPC ピア宛てのパケットは、ルーティングされるために vPC ピアリンクから出る必要がなくなります。言い換えると、この vPC ピアゲートウェイの機能拡張により、一方の vPC ピアがリモート vPC ピアの「代わり」にパケットをルーティングできるようになります。この vPC ピアゲートウェイの機能拡張は、vPC ドメイン設定コマンドの peer-gateway で有効にすることができます。

## 警告

vPC または vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係のフラッピング

2 つの vPC ピアと vPC 接続のルータまたは vPC 孤立ポートを介して接続されたルータの間に動的ユニキャストルーティングプロトコル隣接関係が形成されている場合、vPC ピアゲートウェイの機能拡張を有効すると、その後に vPC を介したルーティング/レイヤ 3 の機能拡張を設定しないときには、ルーティングプロトコル隣接関係の継続的なフラッピングが始まる可能性があります。これらの障害シナリオについては、このドキュメントの「vPC ピアゲートウェイのある、vPC を介したユニキャスト ルーティング プロトコル隣接関係」セクションにある障害シナリオの例と「vPC ピアゲートウェイのある、vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係」セクションで詳しく説明しています。

この問題を解決するには、vPC ドメイン設定コマンドの peer-gateway で vPC ピアゲートウェイの機能拡張を有効にした直後に、vPC ドメイン設定コマンドの layer3 peer-router で vPC を介したルーティング/レイヤ 3 の機能拡張を有効にてください。

ICMP および ICMPv6 リダイレクトの自動無効化

vPCピアゲートウェイ機能拡張を有効にすると、すべてのvPC VLAN SVI（つまり、vPCピアリンク経由でトランクされるVLANに関連付けられたすべてのSVI）で、ICMPおよびICMPv6リダイレクトパケットの生成が自動的に無効になります。スイッチは、すべての vPC VLAN SVI で no ip redirects および no ipv6 redirects を設定することにより、これを実行します。これにより、スイッチは、スイッチに入るパケットのうち、そのスイッチの vPC ピアの宛先 MAC アドレスおよび IP アドレス持つものに応答して ICMP リダイレクトパケットを生成しなくなります。

特定のVLAN内の環境でICMPまたはICMPv6リダイレクトパケットが必要な場合は、peer-gateway exclude-vlan <vlan-id> vPCドメイン設定コマンドを使用して、このVLANをvPCピアゲートウェイ拡張機能の利用から除外する必要があります。

---

注:peer-gateway exclude-vlan <vlan-id> vPCドメイン設定コマンドは、Nexus 9000シリーズスイッチではサポートされていません。

---

## コンフィギュレーション

ここでは、vPC ピアゲートウェイ機能を設定する方法の例を示します。

この例では、N9K-1 と N9K-2 が vPC ドメイン内 vPC ピアです。両方のvPCピアに、VLAN 10用に設定されたHSRPグループがあります。N9K-1はプライオリティ150のHSRPアクティブルータで、N9K-2はデフォルトのプライオリティ100のHSRPスタンバイルータです。

<#root>

```
N9K-1#

show running-config vpc

<snip>
vpc domain 1
  role priority 150
  peer-keepalive destination 10.82.140.43

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

<snip>
vpc domain 1
  peer-keepalive destination 10.82.140.42

interface port-channel1
  vpc peer-link

N9K-1#

show running-config interface vlan 10

<snip>
```

```
interface Vlan10
  no shutdown
  ip address 192.168.10.2/24
  hsrp 10
    preempt
    priority 150
    ip 192.168.10.1

N9K-2#

show running-config interface vlan 10

<snip>
interface Vlan10
  no shutdown
  ip address 192.168.10.3/24
  hsrp 10
    ip 192.168.10.1

N9K-1#

show hsrp interface vlan 10 brief

*:IPv6 group   #:group belongs to a bundle
                      P indicates configured to preempt.
                      |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   150  P Active   local            192.168.10.3     192.168.10.1    (conf)

N9K-2#

show hsrp interface vlan 10 brief

*:IPv6 group   #:group belongs to a bundle
                      P indicates configured to preempt.
                      |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   100    Standby  192.168.10.2     local            192.168.10.1    (conf)
```

N9K-1 の VLAN 10 SVI の MAC アドレスは 00ee.ab67.db47 であり、N9K-2 の VLAN 10 SVI の MAC アドレスは 00ee.abd8.747f です。VLAN 10 の HSRP 仮想 MAC アドレスは 0000.0c07.ac0a です。この状態では、各スイッチの VLAN 10 SVI の MAC アドレスと HSRP 仮想 MAC アドレスが各スイッチの MAC アドレステーブルに存在します。各スイッチのVLAN 10 SVI MACアドレスとHSRP仮想MACアドレスには、ゲートウェイ(G)フラグが付いています。これは、このMACアドレス宛てのパケットがスイッチによってローカルにルーティングされることを示しています。

N9K-1 の MAC アドレステーブルには、N9K-2 の VLAN 10 SVI の MAC アドレスに存在するゲートウェイフラグがないことに注意してください。同様に、N9K-2 の MAC アドレステーブルには、N9K-1 の VLAN 10 SVI の MAC アドレスに存在するゲートウェイフラグがありません。

<#root>

```
N9K-1#

show mac address-table vlan 10
```

```
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
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    sup-eth1(R)
G   10     00ee.ab67.db47   static   -         F      F    sup-eth1(R)
*   10     00ee.abd8.747f   static   -         F      F    vPC Peer-Link(R)

N9K-2#
```

**show mac address-table vlan 10**

```
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
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    vPC Peer-Link(R)
*   10     00ee.ab67.db47   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.abd8.747f   static   -         F      F    sup-eth1(R)
```

vPC ピアゲートウェイの機能拡張は、vPC ドメイン設定コマンドの peer-gateway を使用して有効にすることができます。これにより、vPCピアリンクで学習されたvPCピアのMACアドレスに属する宛先MACアドレスを持つ受信パケットを、スイッチがローカルにルーティングできるようになります。これは、スイッチの MAC アドレステーブルに含まれる vPC ピアの MAC アドレスにゲートウェイフラグを設定することで実現されます。

<#root>

```
N9K-1#
```

**configure terminal**

```
Enter configuration commands, one per line. End with CNTL/Z.
N9K-1(config)#
```

**vpc domain 1**

```
N9K-1(config-vpc-domain)#
```

**peer-gateway**

```
N9K-1(config-vpc-domain)#
```

**end**

```
N9K-1#

N9K-2#
```

**configure terminal**

```
Enter configuration commands, one per line. End with CNTL/Z.
N9K-2(config)#
```

**vpc domain 1**

```
N9K-2(config-vpc-domain)#
```

```
peer-gateway

N9K-2(config-vpc-domain)#

end

N9K-2#
```

vPC ピアの MAC に関して MAC アドレステーブルにゲートウェイフラグが存在することを検証することにより、vPC ピアゲートウェイの機能拡張が予期どおりに動作していることを確認できます。

<#root>

```
N9K-1#

show mac address-table vlan 10

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
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    sup-eth1(R)
G   10     00ee.ab67.db47   static   -         F      F    sup-eth1(R)
G   10     00ee.abd8.747f   static   -         F      F    vPC Peer-Link(R)

N9K-2#

show mac address-table vlan 10

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
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.ab67.db47   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.abd8.747f   static   -         F      F    sup-eth1(R)
```

## 影響

vPCピアゲートウェイ拡張機能を有効にした場合の影響は、周辺のトポロジや、次のサブセクションで説明する接続ホストの動作によって異なります。次のサブセクションのいずれも環境に適用されない場合、vPCピアゲートウェイ拡張機能を有効にしても中断は発生せず、環境への影響もありません。

### vPC または vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係のフラッピング

2 つの vPC ピアと vPC 接続のルータまたは vPC 孤立ポートを介して接続されたルータの間に動的ユニキャストルーティングプロトコル隣接関係が形成されている場合、vPC ピアゲートウェイ

の機能拡張を有効すると、その後に vPC を介したルーティング/レイヤ 3 の機能拡張を設定しないときには、ルーティングプロトコル隣接関係の継続的なフラッピングが始まる可能性があります。これらの障害シナリオについては、このドキュメントの「vPC ピアゲートウェイのある、vPC を介したユニキャスト ルーティング プロトコル隣接関係」セクションにある障害シナリオの例と「vPC ピアゲートウェイのある、vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係」セクションで詳しく説明しています。

この問題を解決するには、vPC ドメイン設定コマンドの peer-gateway で vPC ピアゲートウェイの機能拡張を有効にした直後に、vPC ドメイン設定コマンドの layer3 peer-router で vPC を介したルーティング/レイヤ 3 の機能拡張を有効にてください。

ICMP および ICMPv6 リダイレクトの自動無効化

vPCピアゲートウェイ機能拡張を有効にすると、すべてのvPC VLAN SVI（つまり、vPCピアリンク経由でトランクされるVLANに関連付けられたすべてのSVI）で、ICMPおよびICMPv6リダイレクトパケットの生成が自動的に無効になります。スイッチは、すべての vPC VLAN SVI で no ip redirects および no ipv6 redirects を設定することにより、これを実行します。これにより、スイッチは、スイッチに入るパケットのうち、そのスイッチの vPC ピアの宛先 MAC アドレスおよび IP アドレス持つものに応答して ICMP リダイレクトパケットを生成しなくなります。

特定のVLAN内の環境でICMPまたはICMPv6リダイレクトパケットが必要な場合は、peer-gateway exclude-vlan <vlan-id> vPCドメイン設定コマンドを使用して、このVLANをvPCピアゲートウェイ拡張機能の利用から除外する必要があります。

---

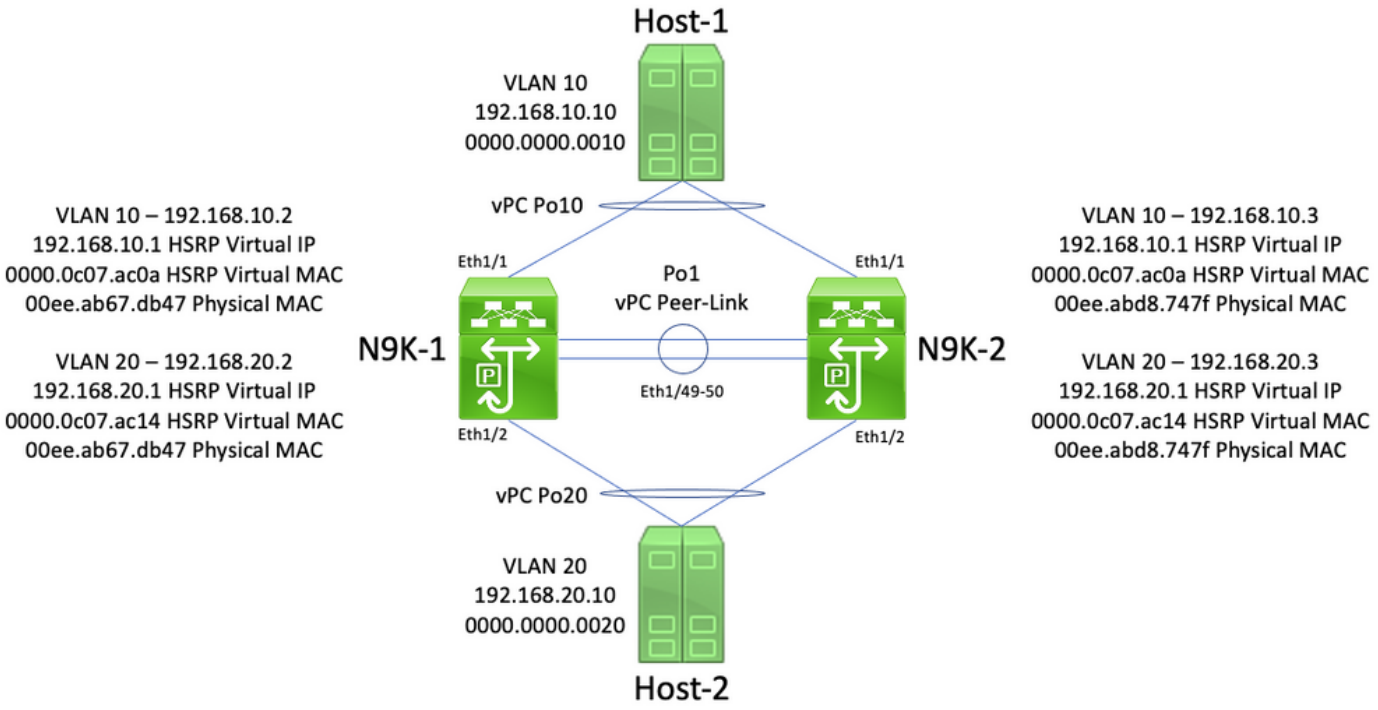
注:peer-gateway exclude-vlan <vlan-id> vPCドメイン設定コマンドは、Nexus 9000シリーズスイッチではサポートされていません。

---

## 障害シナリオの例

非標準転送動作を持つ vPC 接続ホスト

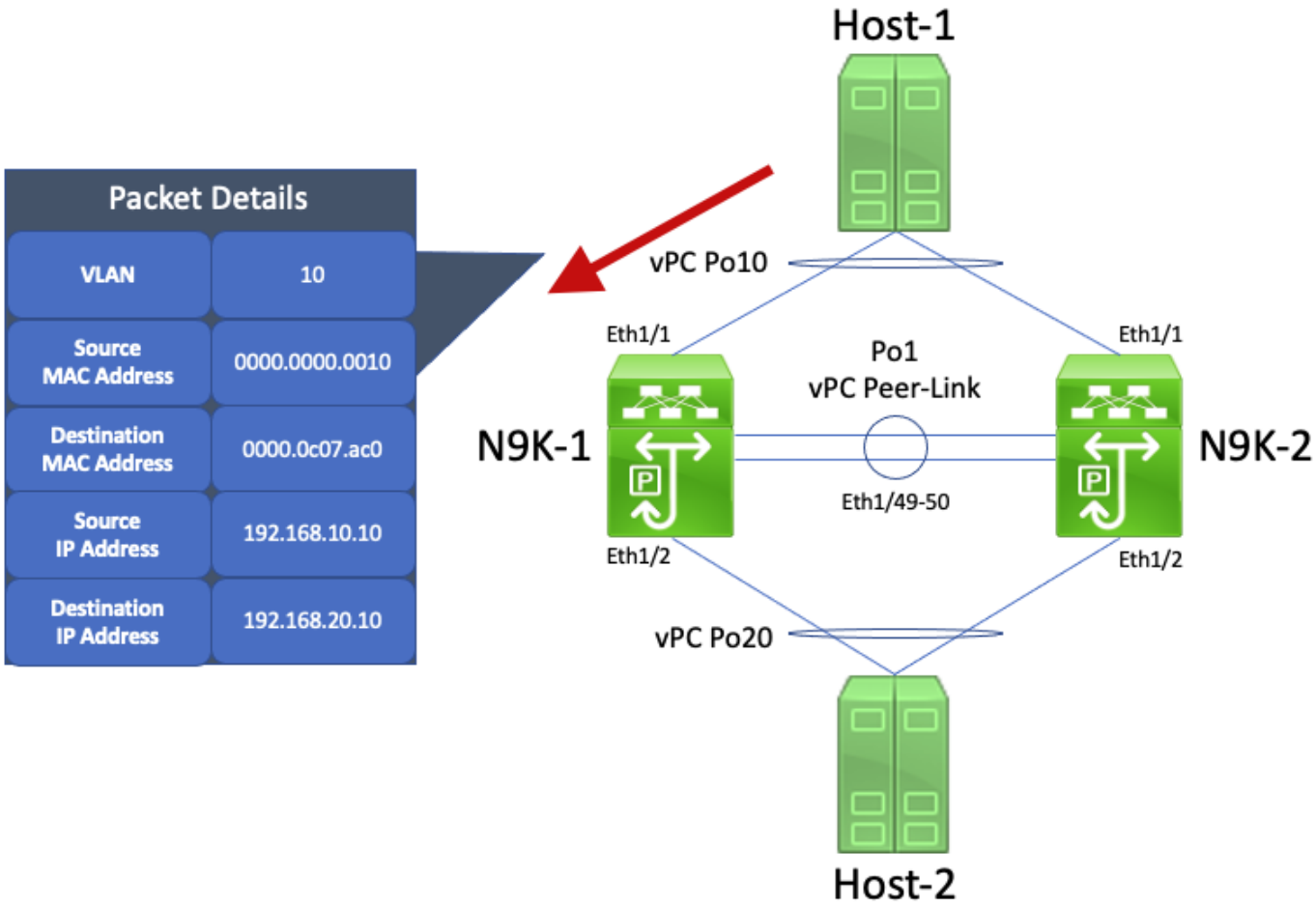
次のトポロジについて考えてみます。



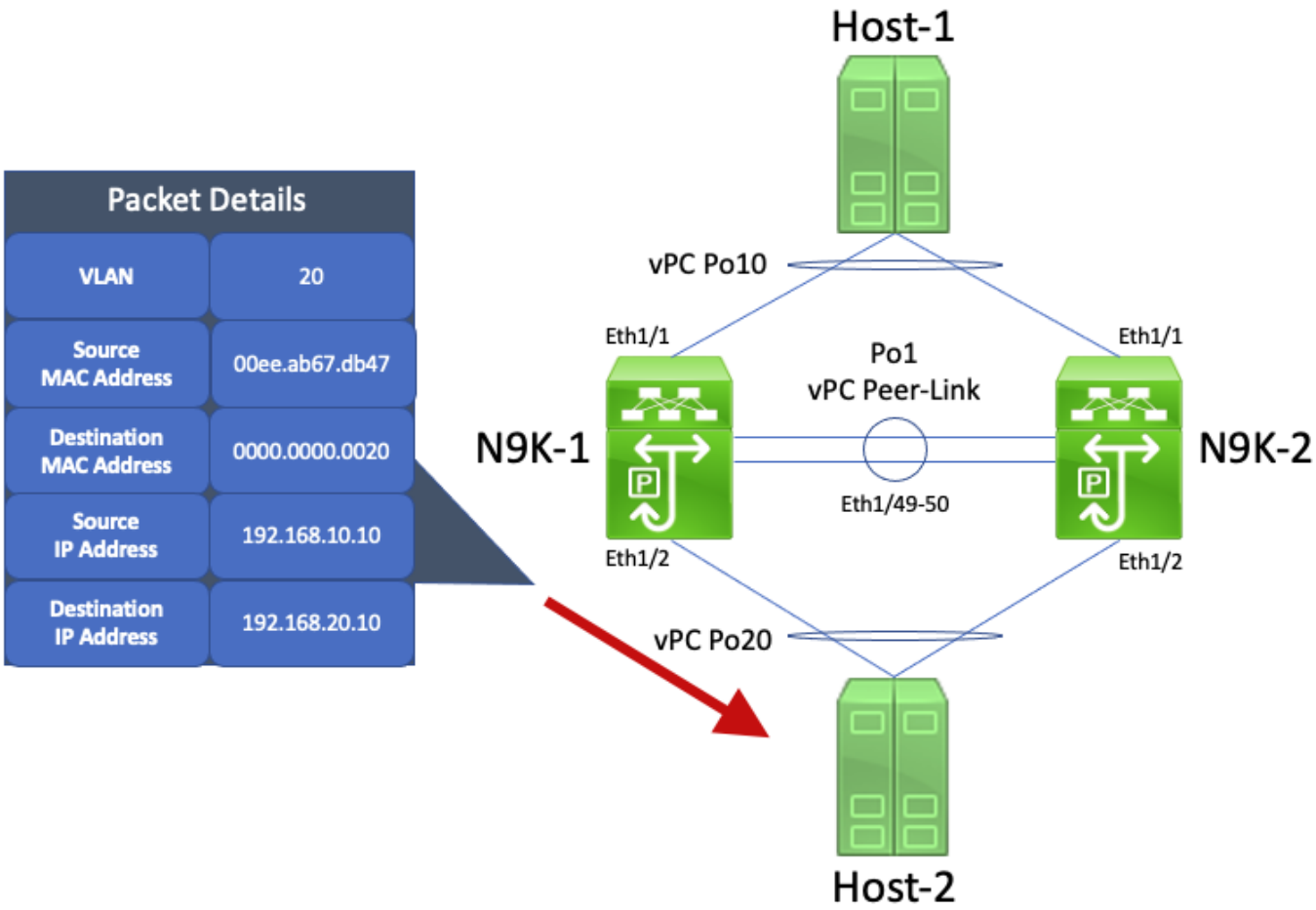

このトポロジでは、N9K-1とN9K-2はvPCドメイン内のvPCピアであり、VLAN 10とVLAN 20の間でVLAN間ルーティングを実行します。Po1 インターフェイスは vPC ピアリンクです。Host-1という名前のホストは、vPC Po10を介してVLAN 10のN9K-1およびN9K-2に接続されています。Host-1は、IPアドレス192.168.10.10、MACアドレス0000.0000.0010を所有しています。Host-2という名前のホストは、vPC Po20を介してVLAN 20のN9K-1およびN9K-2に接続されています。Host-2は、IPアドレス192.168.20.10とMACアドレス0000.0000.0020を所有しています。

N9K-1 と N9K-2 はどちらも、VLAN 10 と VLAN 20 に SVI があり、各 SVI で HSRP がアクティブになっています。N9K-1のVLAN 10インターフェイスのIPアドレスは192.168.10.2で、N9K-1のVLAN 20インターフェイスのIPアドレスは192.168.20.2です。N9K-1のSVIの物理MACアドレスは両方とも00ee.ab67.db47です。N9K-2のVLAN 10インターフェイスのIPアドレスは1921.68.10.3で、N9K-2のVLAN 20インターフェイスのIPアドレスは192.168.20.3です。N9K-2のSVIの物理MACアドレスは両方とも00ee.abd8.747fです。VLAN 10 の HSRP 仮想 IP アドレスは 192.168.10.1、HSRP 仮想 MAC アドレスは 0000.0c07.ac0a です。VLAN 20 の HSRP 仮想 IP アドレスは 192.168.20.1、HSRP 仮想 MAC アドレスは 0000.0c07.ac14 です。

Host-1がICMPエコー要求パケットをHost-2に送信するシナリオを考えます。Host-1がデフォルトゲートウェイ（HSRP仮想IPアドレス）のARPを解決すると、Host-1は標準の転送動作に従って、送信元IPアドレス192.168.10.10、宛先IPアドレス192.168.20.10、送信元MACアドレス0000.0000.0010、宛先MACアドレス0000010c0を0としてICMPIPパケットをを生成します07.AC0Aです。このパケットはN9K-1に向けて出力されます。これの視覚的な例を、次に示します。

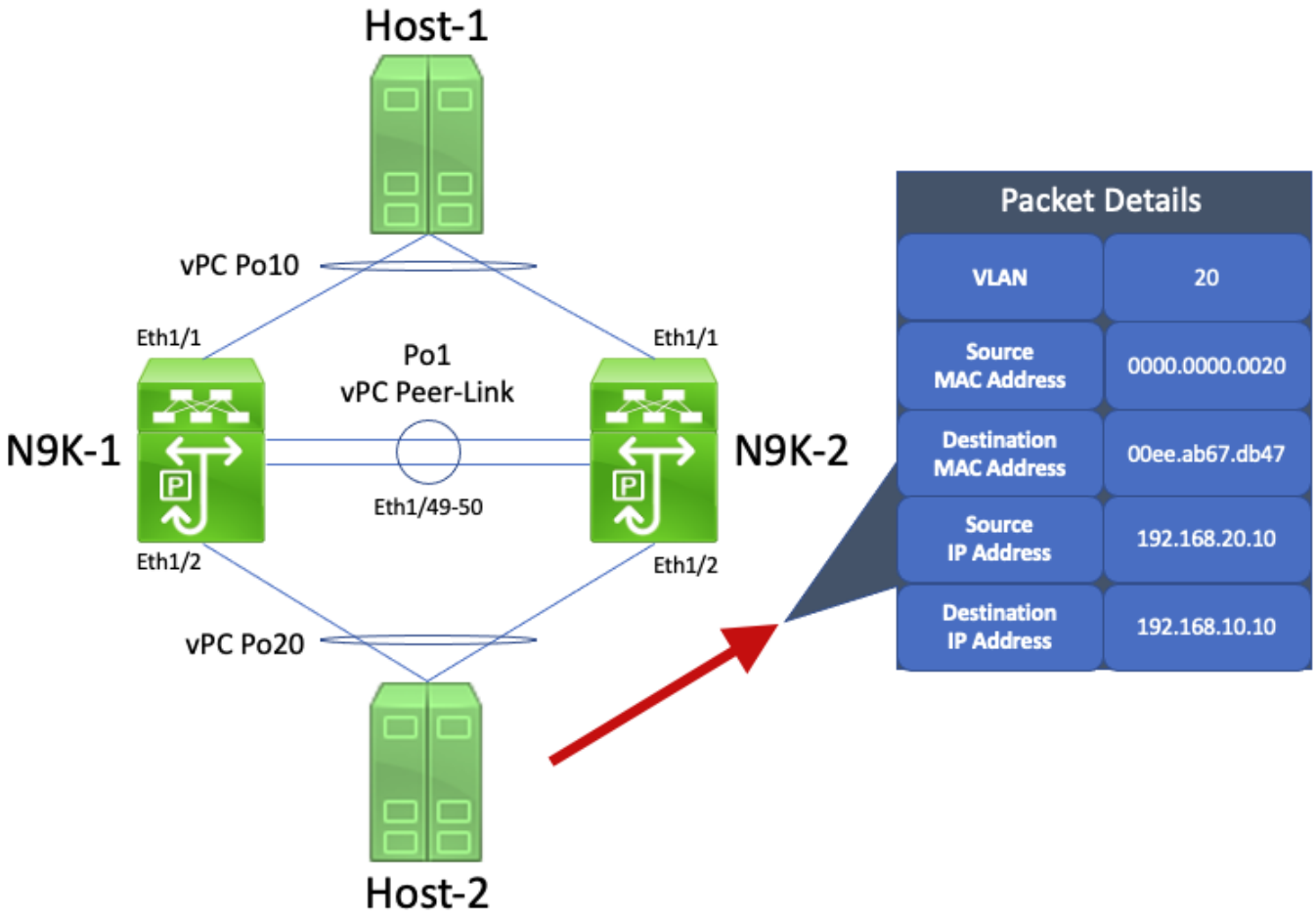



N9K-1 は、このパケットを受信します。このパケットは HSRP 仮想 MAC アドレス宛てであるため、N9K-1 は、HSRP コントロールプレーンの状態に関係なくローカル ルーティング テーブルに従ってこのパケットをルーティングすることができます。このパケットはVLAN 10からVLAN 20にルーティングされます。パケットのルーティングの一部として、N9K-1は、パケットの送信元および宛先MACアドレスフィールドを再アドレッシングすることで、パケットの書き換えを実行します。パケットの新しい送信元MACアドレスは、N9K-1のVLAN 20 SVIに関連付けられた物理MACアドレス(00ee.ab67.db47)で、新しい宛先MACアドレスは、ホスト2(0000.0000.0020)に関連付けられたMACアドレスです。これの視覚的な例を、次に示します。

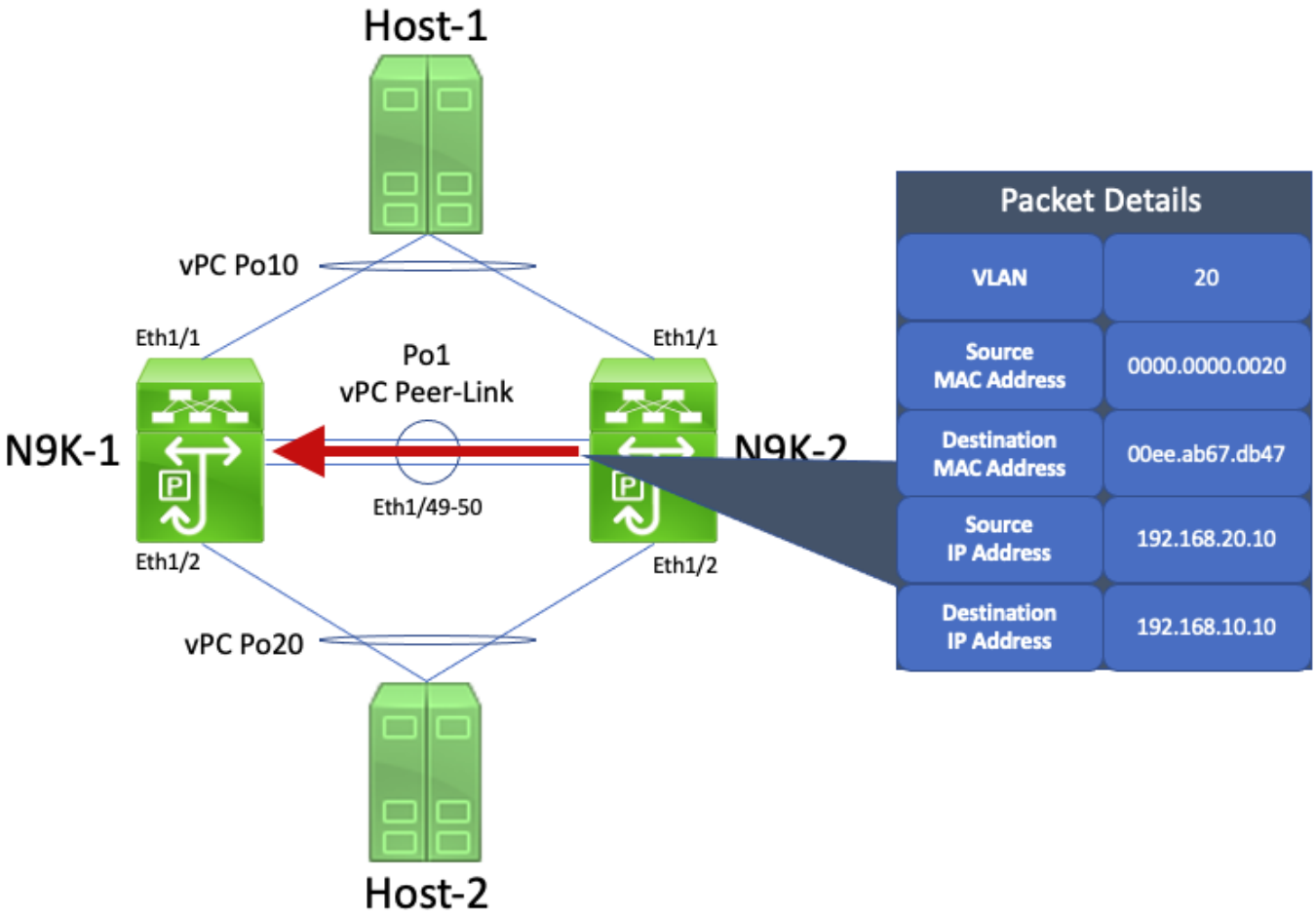



Host-2 は、このパケットを受信し、Host-1 の ICMP エコー要求パケットに応答して ICMP エコー応答パケットを生成します。ただし、Host-2 が標準の転送動作に従わない場合があります。転送を最適化するために、Host-2 は、Host-1 の IP アドレス ( 192.168.10.10 ) のルーティングテーブルまたは ARP キャッシュのルックアップを実行しません。代わりに、Host-2 が最初に受信した ICMP エコー要求パケットの送信元 MAC アドレスと宛先 MAC アドレスのフィールドを反転させます。その結果、Host-2によって生成されるICMPエコー応答パケットの送信元IPアドレスは192.168.20.10、宛先IPアドレスは192.168.10.10、送信元MACアドレスは0000.0000.0020、宛先MACアドレスは00ee.ab67.db47になります。

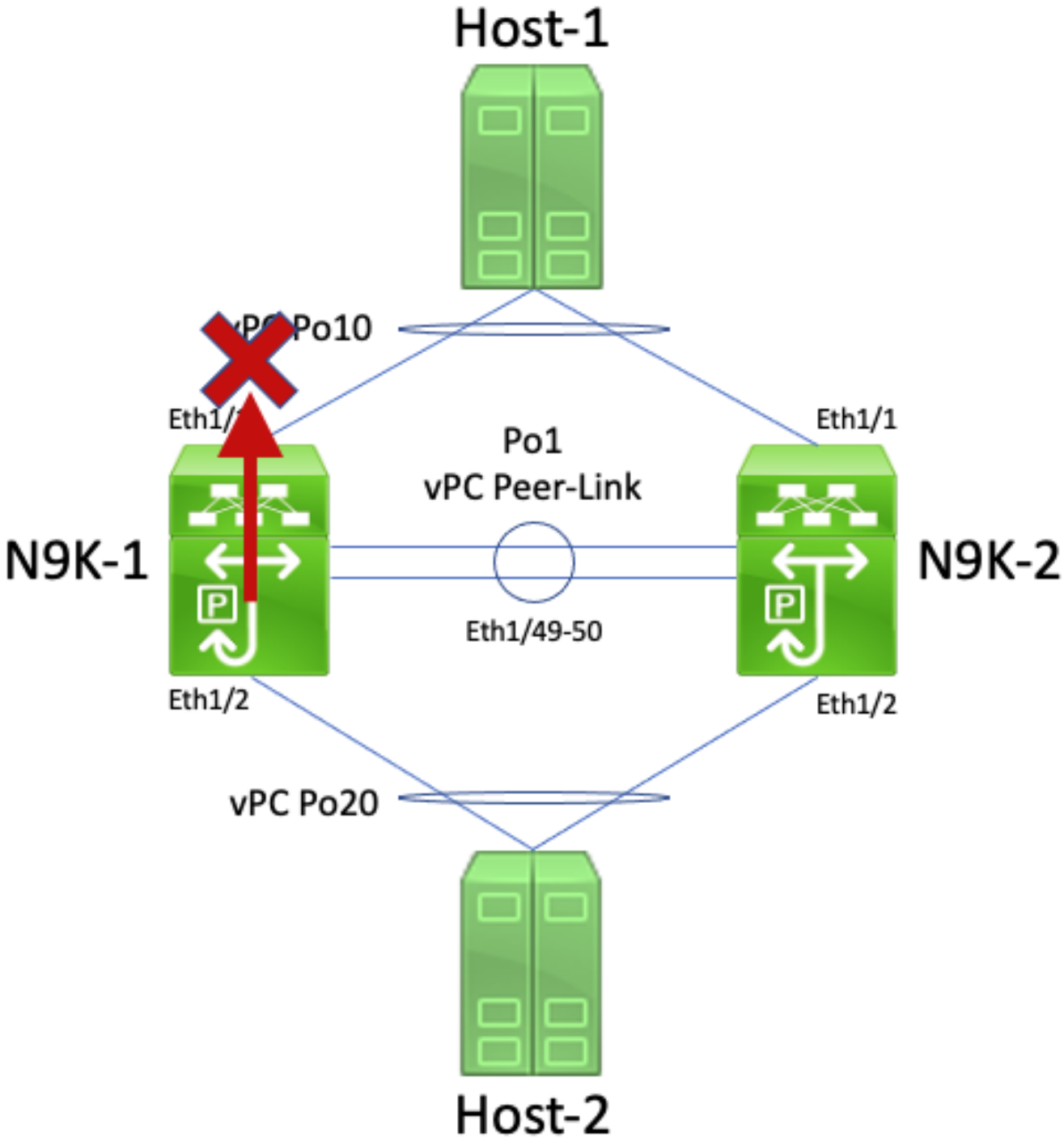
このICMPエコー応答パケットがN9K-1に向けて出力される場合、このパケットは問題なくHost-1に転送されます。ただし、次に示すように、この ICMP エコー応答パケットが N9K-2 に向かうシナリオについて考えてみます。




N9K-2 は、このパケットを受信します。このパケットはN9K-1のVLAN 20 SVIの物理MACアドレス宛てであるため、N9K-2はN9K-1の代わりにこのパケットをルーティングできないため、N9K-2はこのパケットをvPCピアリンク経由でN9K-1に転送します。これの視覚的な例を、次に示します。




N9K-1 は、このパケットを受信します。このパケットは N9K-1 の VLAN 20 SVI の物理 MAC アドレス宛てであるため、N9K-1 は、HSRP コントロールプレーンの状態に関係なくローカル ルーティング テーブルに従ってこのパケットをルーティングすることができます。このパケットはVLAN 20からVLAN 10にルーティングされます。ただし、このルートの出力インターフェイスはvPC Po10に解決され、vPC Po10はN9K-2でアップ状態です。これはvPCループ回避ルールの違反です。N9K-1がvPCピアリンク経由でパケットを受信した場合、同じvPCインターフェイスがN9K-2でアップ状態であれば、N9K-1はそのパケットををvインターフェイスから転送できません。この違反の結果、N9K-1はこのパケットをドロップします。これの視覚的な例を、次に示します。




PC ドメイン設定コマンドの peer-gateway を使用して vPC ピアゲートウェイの機能拡張を有効にすることにより、この問題を解決することができます。これにより、パケットの宛先MACアドレスがN9K-2ではなくN9K-1によって所有されている場合でも、N9K-2はN9K-1に代わってICMPエコー応答パケット（および同様にアドレス指定されたその他のパケット）をルーティングできます。その結果、N9K-2はこのパケットをvPC Po10インターフェイスから転送でき、vPCピアリンク経由で転送する必要はありません。

## vPC を介したルーティング/レイヤ 3（レイヤ 3 ピアルータ）

ここでは、vPC ドメイン設定コマンドの layer3 peer-router で有効にする vPC を介したルーティング/レイヤ 3 の機能拡張について説明します。

---

注:vPC上でのマルチキャストルーティングプロトコルの隣接関係(つまり、Protocol Independent Multicast(PIM)の隣接関係)の形成は、Routing/Layer 3 over vPC機能拡張が有効になっている場合はサポートされません。

---

## 概要

一部の環境では、vPC を介してルータを Nexus スイッチのペアに接続し、両方の vPC ピアを使用して vPC を介したユニキャスト ルーティング プロトコル隣接関係を形成する必要があります。また、vPC VLAN を介してルータを単一の vPC ピアに接続し、両方の vPC ピアを使用して vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係を形成する必要がある場合もあります。その結果、vPC に接続されたルータは、両方の Nexus スイッチによってアドバタイズされるプレフィックスに関して等コストマルチパス ( ECMP ) を持ちます。これは、vPC 接続ルータと両方の vPC ピアの間で専用ルーティングリンクを使用して、IP アドレスの使用を節約 ( 必要な IP アドレスが 4 つではなく 3 つ ) したり、構成の複雑さ ( 特にサブインターフェイスを必要とする VRF-Lite 環境で SVI とともに使用されるルーテッドインターフェイス ) を軽減するよりも、望ましい可能性があります。

これまで、vPC を介したユニキャスト ルーティング プロトコル隣接関係の形成は、Cisco Nexus プラットフォームではサポートされていませんでした。ただし、サポートされていない場合でも、ユニキャスト ルーティング プロトコル隣接関係が vPC を介して問題なく形成されるトポロジが実装されている場合があります。ネットワークに何らかの変更 ( vPC 接続ルータまたは vPC ピア自体のソフトウェアのアップグレード、ファイアウォールのフェールオーバーなど ) が加えられると、vPC を介したユニキャスト ルーティング プロトコル隣接関係が機能しなくなり、データプレーントラフィックでパケット損失が発生するか、一方または両方の vPC ピアを使用してユニキャスト ルーティング プロトコル隣接関係を稼働させることができなくなります。これらのシナリオが失敗する理由およびサポートされない理由の背後にある技術的な詳細については、このドキュメントの「障害シナリオの例」セクションで説明します。

vPC を介したルーティング/レイヤ 3 の機能拡張は、vPC を介したユニキャスト ルーティング プロトコル隣接関係の形成のサポートを追加するために導入されました。これは、TTL が 1 のユニキャスト ルーティング プロトコル パケットを、パケットの TTL を減らすことなく vPC ピアリンクを介して転送できるようにすることで実現されます。その結果、ユニキャスト ルーティング プロトコル隣接関係を、vPC または vPC VLAN を介して問題なく形成できます。vPC を介したルーティング/レイヤ 3 の機能拡張は、vPC ドメイン設定コマンドの peer-gateway で vPC ピアゲートウェイの機能拡張を有効にした後に、vPC ドメイン設定コマンドの layer3 peer-router で有効にすることができます。

各 Cisco Nexus プラットフォーム向けに vPC を介したルーティング/レイヤ 3 の機能拡張のサポートが導入された NX-OS ソフトウェアリリースは、『Supported Topologies for Routing over Virtual Port Channel on Nexus Platforms』内の表 2 (「Routing Protocols Adjacencies Support over vPC VLANs」) に示されています。

## 警告

時折発生する VPC-2-L3_VPC_UNEQUAL_WEIGHT Syslog

Routing/Layer 3 over vPC機能拡張が有効になると、両方のvPCピアで次のようなsyslogが1時間に1回生成され始めます。

```
2021 May 26 19:13:47.079 switch %VPC-2-L3_VPC_UNEQUAL_WEIGHT: Layer3 peer-router is enabled. Please mak
2021 May 26 19:13:47.351 switch %VPC-2-L3_VPC_UNEQUAL_WEIGHT: Unequal weight routing is not supported i
```

これらの Syslog はどちらも、スイッチの問題を示すものではありません。これらの Syslog は、vPC を介したルーティング/レイヤ 3 の機能拡張が有効になっている場合に、両方の vPC ピアがトラフィックを同じようにルーティングできる状態を確保するために、ルーティングの設定、コスト、および重みが両方の vPC ピアで同一である必要があることを管理者に警告します。これは、必ずしも、一致しないルーティングの設定、コスト、または重みがいずれかの vPC ピアに存在することを示しているわけではありません。

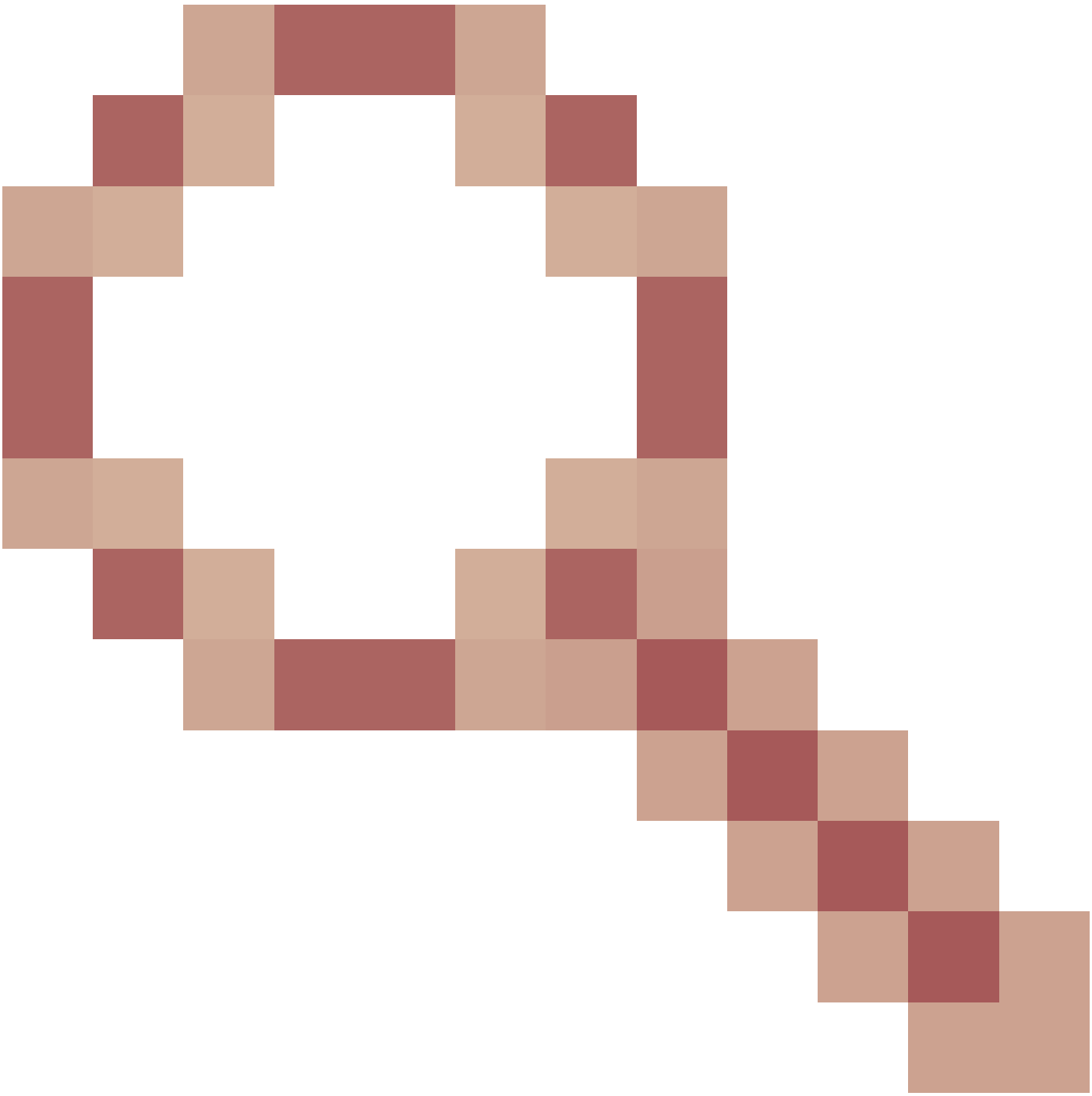
これらの Syslog は、ここに示す設定によって無効にすることができます。

<#root>

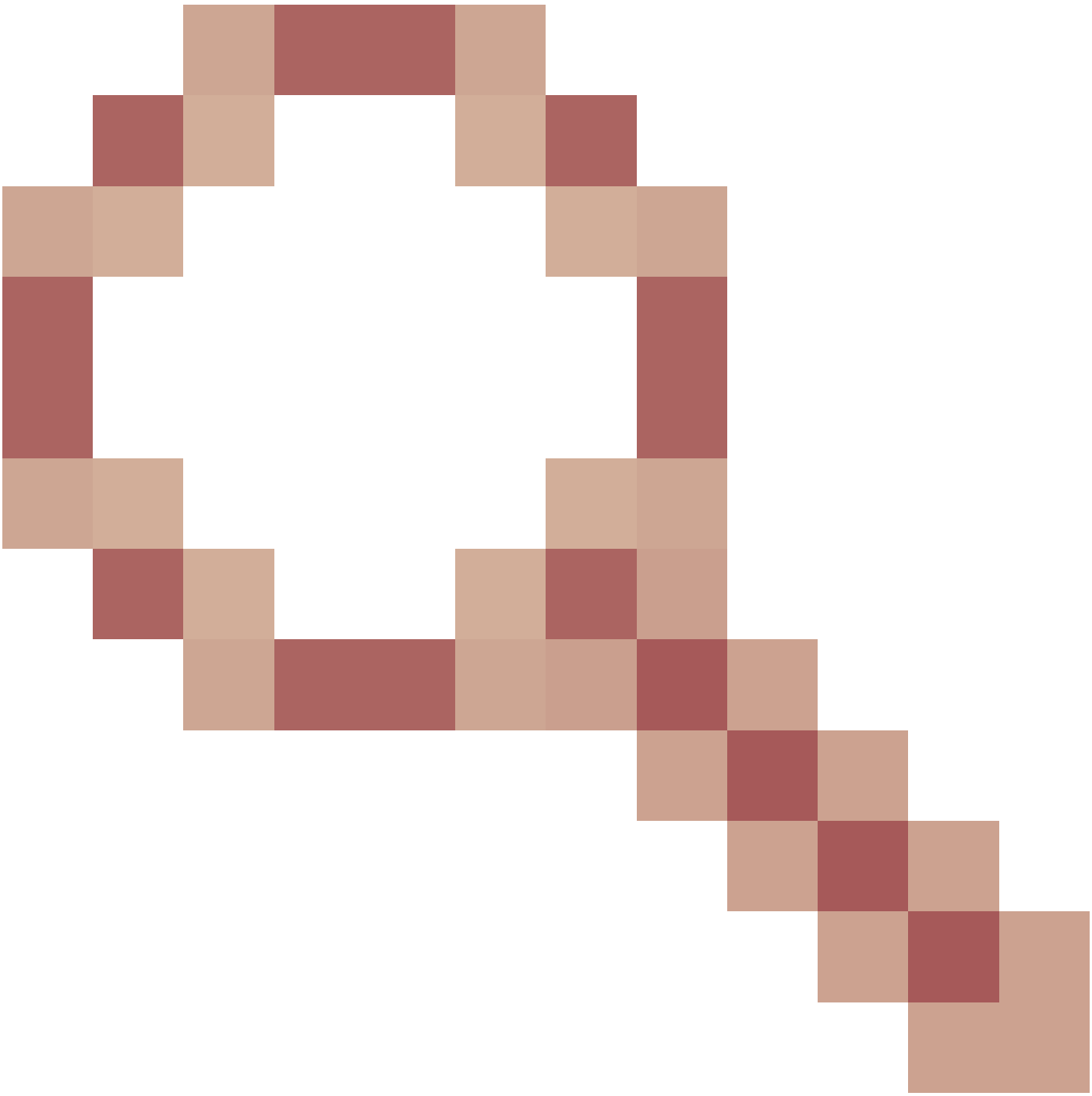
```
switch#
configure terminal
switch(config)#
vpc domain 1
switch(config-vpc-domain)#
no layer3 peer-router syslog
switch(config-vpc-domain)#
end
switch#
```

この設定は、両方のvPCピアでsyslogを無効にするために、両方のvPCピアで実行する必要があります。

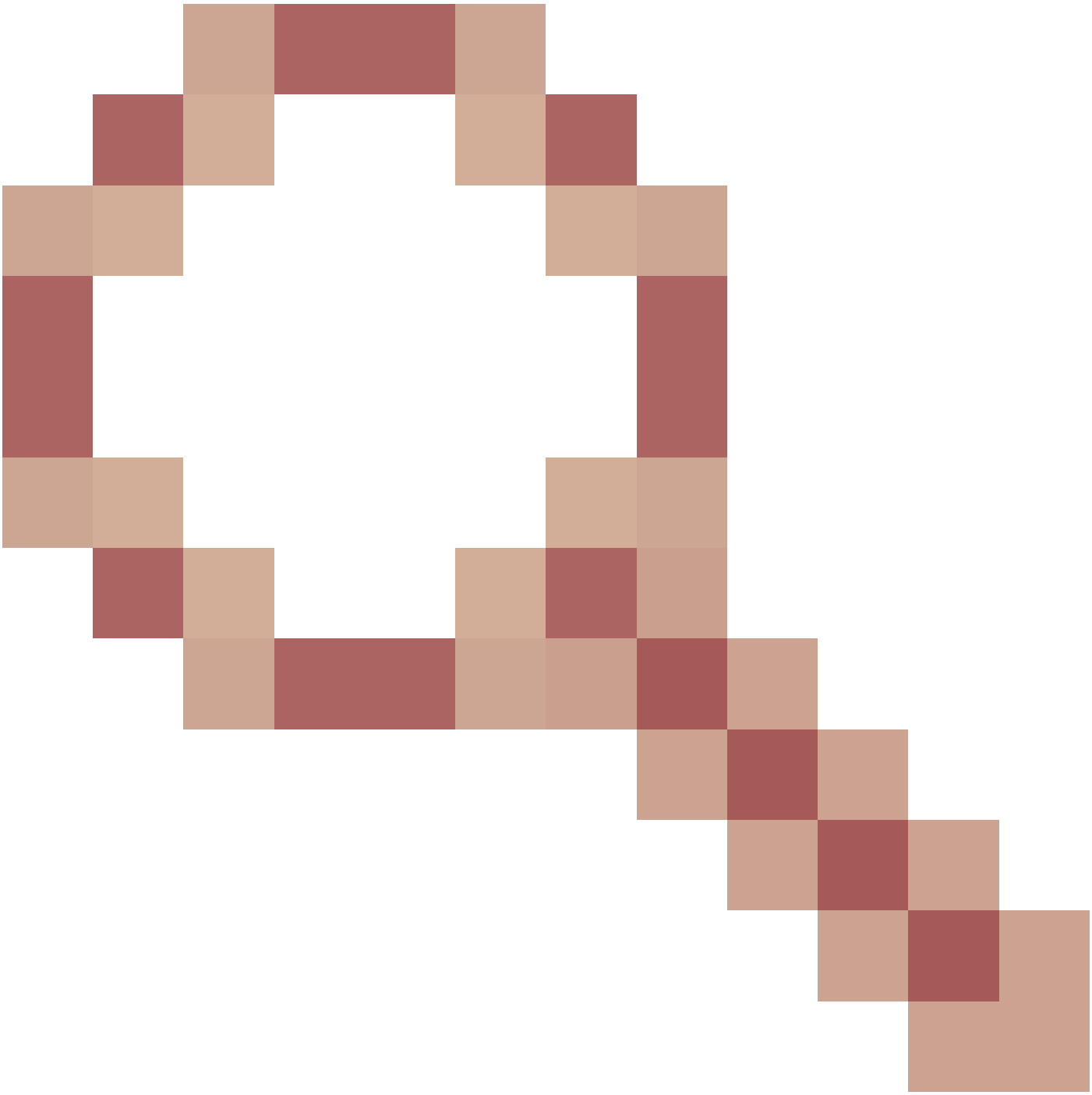
Cisco Bug ID CSCvs82183により、TTLが1のソフトウェアで転送されたデータプレーントラフィック



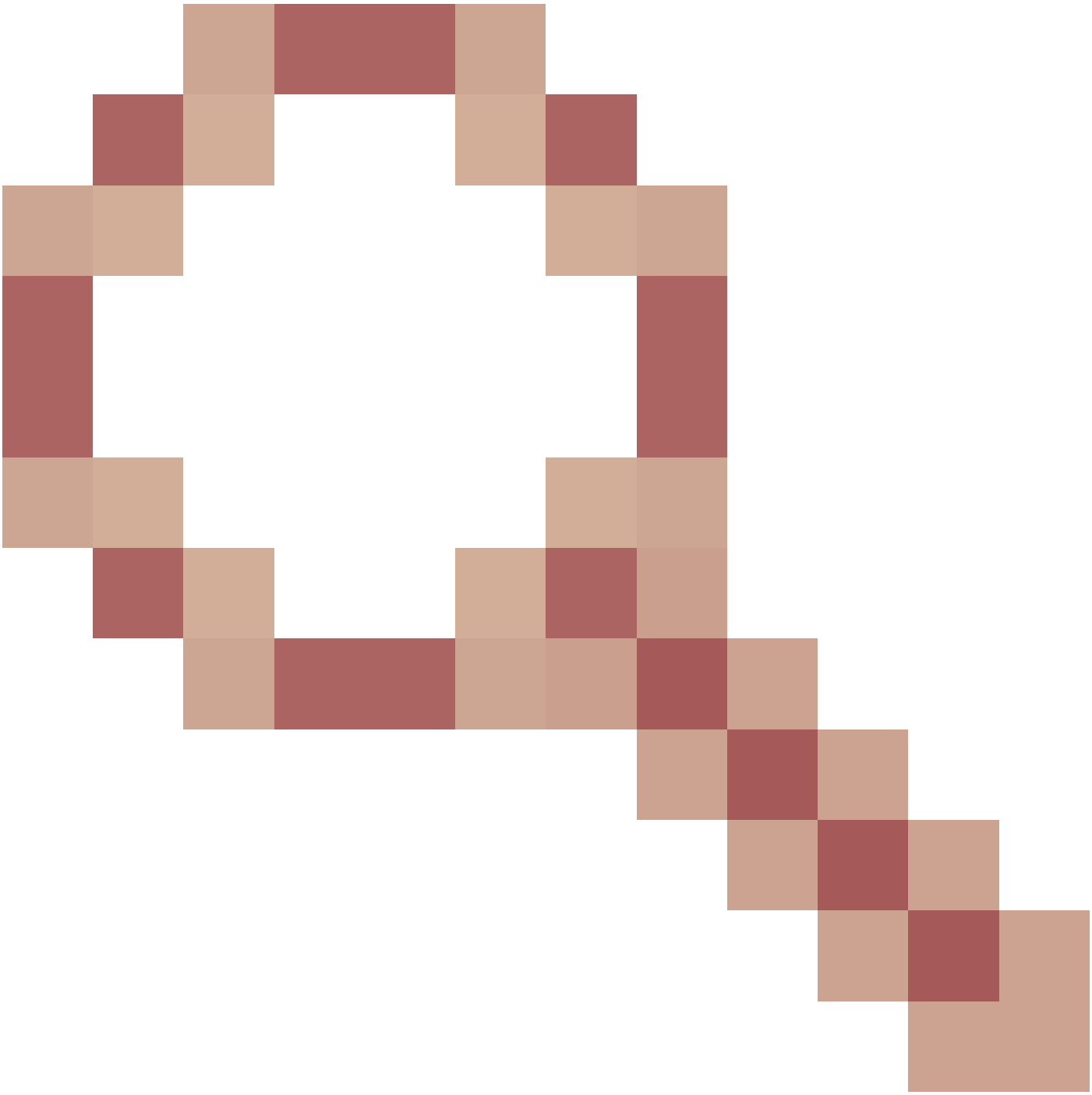
およびCisco Bug ID CSCvw16965



NX-OSソフトウェアリリース9.3(6)より前のNX-OSソフトウェアリリースを実行するクラウドスケールASICを搭載したNexus 9000シリーズスイッチでRouting/Layer 3 over vPC機能拡張を有効にすると、TTLが1のユニキャストルーティングプロトコルに関連付けられていないデータプレーントラフィックがスーパーバイザにパントされ、ハードウェアではなくソフトウェアで転送されます。Nexusスイッチが固定シャーシ（「トップオブラック」とも呼ばれる）スイッチか、モジュラシャーシ（「エンドオブロー」とも呼ばれる）スイッチか、およびスイッチの現在のNX-OSソフトウェアリリースかによって、この問題の根本的な原因はソフトウェア不具合Cisco Bug ID CSCvs82183



またはソフトウェア不具合Cisco Bug ID CSCvw16965



を参照。どちらのソフトウェア不具合も、Cloud Scale ASICを搭載したNexus 9000シリーズスイッチにのみ影響します。他のCisco Nexusハードウェアプラットフォームはいずれの問題にも該当しません。詳細については、各ソフトウェア不具合の情報を参照してください。

これらのソフトウェア不具合を回避するために、NX-OS ソフトウェアリリース 9.3(6) 以降にアップグレードすることをお勧めします。また、一般的な推奨事項として、『Recommended Cisco NX-OS Releases for Cisco Nexus 9000 Series Switches』で参照されている Nexus 9000 シリーズ スイッチの現在推奨されている NX-OS ソフトウェアリリースに定期的にアップグレードすることをお勧めします。

## コンフィギュレーション

ここでは、vPC を介したルーティング/レイヤ 3 の機能拡張を設定する方法の例を示します。

この例では、N9K-1 と N9K-2 が vPC ドメイン内 vPC ピアです。両方の vPC ピアで vPC ピアゲートウェイの機能拡張がすでに有効になっています。これは、vPC を介したルーティング/レイヤ 3 の機能拡張を有効にするために必要です。両方のvPCピアのVLAN 10にSVIがあり、OSPFプロセス1で有効になっています。N9K-1とN9K-3は、IPアドレスとネイバーIDが192.168.10.3のvPC接続のOSPFルータとのOSPF EXSTART/EXCHANGE状態でスタックしています。

<#root>

```
N9K-1#

show running-config vpc

<snip>
vpc domain 1
  role priority 150
  peer-keepalive destination 10.122.190.196
  peer-gateway

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

<snip>
vpc domain 1
  peer-keepalive destination 10.122.190.195
  peer-gateway

interface port-channel1
  vpc peer-link

N9K-1#

show running-config interface Vlan10


interface Vlan10
  no shutdown
  no ip redirects
  ip address 192.168.10.1/24
  no ipv6 redirects
  ip router ospf 1 area 0.0.0.0

N9K-2#

show running-config interface Vlan10


interface Vlan10
  no shutdown
  no ip redirects
  ip address 192.168.10.2/24
  no ipv6 redirects
  ip router ospf 1 area 0.0.0.0

N9K-1#

show running-config ospf
```

```
feature ospf

router ospf 1

interface Vlan10
  ip router ospf 1 area 0.0.0.0
```

```
N9K-2#
```

**show running-config ospf**

```
feature ospf

router ospf 1

interface Vlan10
  ip router ospf 1 area 0.0.0.0
```

```
N9K-1#
```

**show ip ospf neighbors**

```
 OSPF Process ID 1 VRF default
 Total number of neighbors: 3
 Neighbor ID     Pri State            Up Time  Address         Interface
 192.168.10.2      1 TWOWAY/DROTHER   00:08:10 192.168.10.2    Vlan10
 192.168.10.3      1 EXCHANGE/BDR     00:07:43 192.168.10.3    Vlan10
```

```
N9K-2#
```

**show ip ospf neighbors**

```
 OSPF Process ID 1 VRF default
 Total number of neighbors: 3
 Neighbor ID     Pri State            Up Time  Address         Interface
 192.168.10.1      1 TWOWAY/DROTHER   00:08:21 192.168.10.1    Vlan10
 192.168.10.3      1 EXSTART/BDR      00:07:48 192.168.10.3    Vlan10
```

vPC ドメイン設定コマンドの layer3 peer-router を使用して、vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることができます。これにより、vPCピアゲートウェイ機能拡張が有効になった結果、vPCピアがユニキャストルーティングプロトコルパケットのTTLを減らすことがなくなります。

<#root>

```
N9K-1#
```

**configure terminal**

```
Enter configuration commands, one per line. End with CNTL/Z.
N9K-1(config)#
```

**vpc domain 1**

```
N9K-1(config-vpc-domain)#
layer3 peer-router
N9K-1(config-vpc-domain)#
end
N9K-1#
N9K-2#
configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
N9K-2(config)#
vpc domain 1
N9K-2(config-vpc-domain)#
layer3 peer-router
N9K-2(config-vpc-domain)#
end
N9K-2#
```

vPC を介したルーティング/レイヤ 3 の機能拡張を有効にした直後に、vPC 接続 OSPF ネイバーとの OSPF 隣接関係が FULL 状態に移行することを検証することにより、vPC を介したルーティング/レイヤ 3 の機能拡張が予期どおりに動作していることを確認できます。

<#root>

```
N9K-1#
show ip ospf neighbors

 OSPF Process ID 1 VRF default
 Total number of neighbors: 3
 Neighbor ID     Pri State            Up Time  Address         Interface
 192.168.10.2      1 TWOWAY/DROTHER   00:12:17 192.168.10.2    Vlan10
 192.168.10.3      1 FULL/BDR         00:00:29 192.168.10.3    Vlan10

N9K-2#
show ip ospf neighbors

 OSPF Process ID 1 VRF default
 Total number of neighbors: 3
 Neighbor ID     Pri State            Up Time  Address         Interface
 192.168.10.1      1 TWOWAY/DROTHER   00:12:27 192.168.10.1    Vlan10
 192.168.10.3      1 FULL/BDR         00:00:19 192.168.10.3    Vlan10
```

## 影響

vPC を介したルーティング/レイヤ 3 の機能拡張を有効にしても、本質的に、vPC ドメインが影響を受けることはありません。つまり、Routing/Layer 3 over vPC拡張機能を有効にしても、vPCピアはvPCを一時停止せず、データプレーントラフィックもこの拡張機能を有効にしても本質的に影響を受けません。

ただし、vPC を介したルーティング/レイヤ 3 の機能拡張を有効にしていなかったために以前にダウンした動的ルーティングプロトコル隣接が、この機能拡張を有効にした結果として突然稼働すると、影響を受けるルーティングプロトコル隣接関係の役割、これらの隣接関係を通じてアドバタイズされる特定のプレフィックス、およびユニキャスト ルーティング テーブルの現在の状態に応じて、vPC を介したルーティング/レイヤ 3 の機能拡張を有効にしたときに何らかの中断が観測される可能性があります。

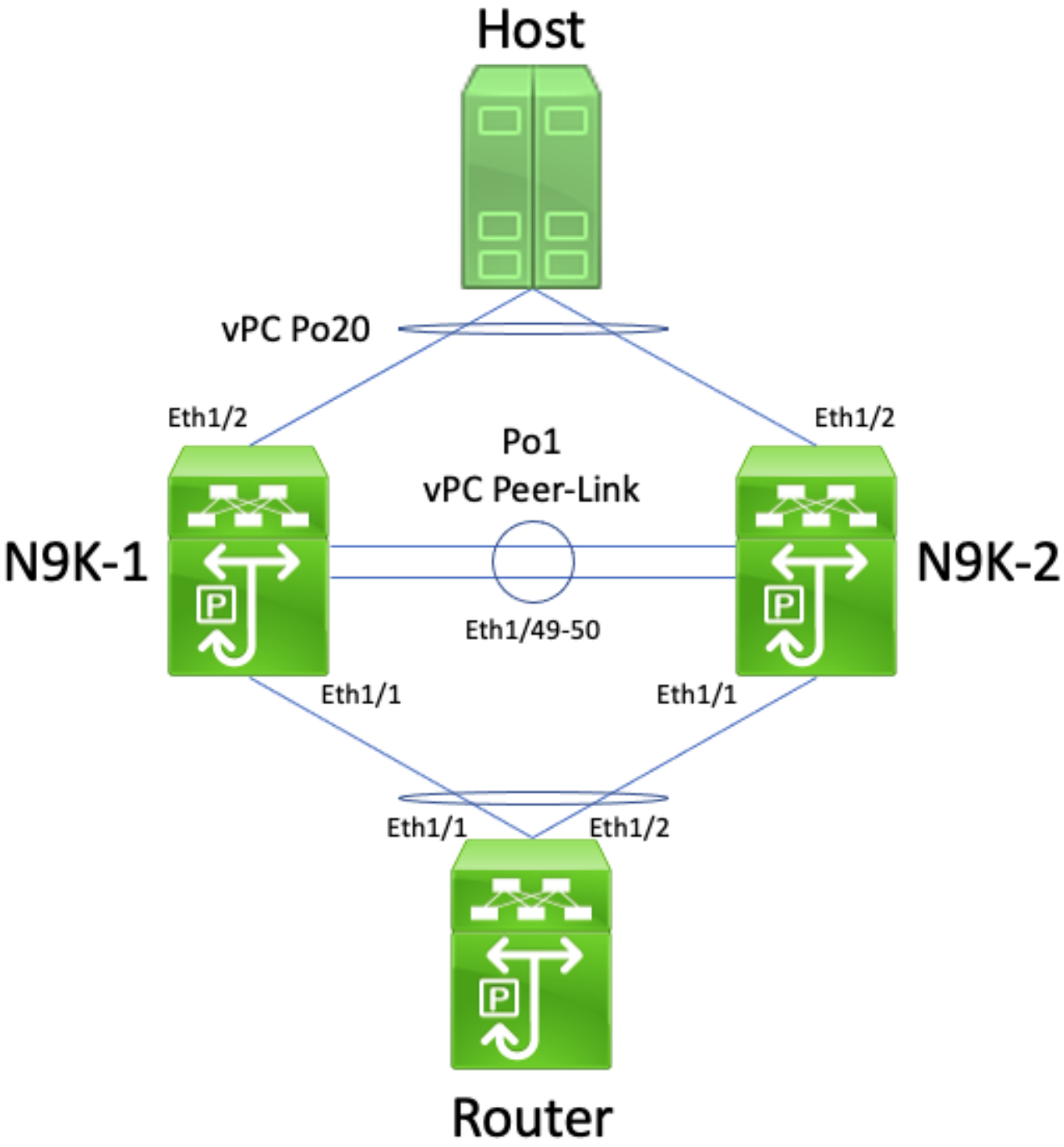
このため、シスコでは、影響を受けるルーティングプロトコルの隣接関係がネットワークの運用に大きな影響を与えるものではないと確信できる場合を除き、コントロールプレーンとデータプレーンの中断が発生する可能性を想定して、メンテナンスの時間帯にこの拡張機能を有効にすることを推奨しています。

また、TLS が 1 の自然なデータプレーントラフィックがハードウェアではなくソフトウェアで処理される原因となる可能性のある、NX-OS ソフトウェアリリースに影響を与えるソフトウェア不具合について、このドキュメントの「注意事項」セクションをよく確認することをお勧めします。

## 障害シナリオの例

vPC ピアゲートウェイのない、vPC を介したユニキャストルーティングプロトコル隣接関係

次に示すトポロジについて考えてみます。




このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっていない vPC ドメイン内の vPC ピアです。Po1 インターフェイスは vPC ピアリンクです。ルータのホスト名を持つルータが、vPC Po10を介してN9K-1およびN9K-2に接続されています。ホストは、vPC Po20を介してN9K-1およびN9K-2に接続されています。ルータのPo10インターフェイスは、ユニキャストルーティングプロトコルの下でアクティブになるルーテッドポートチャネルです。N9K-1 と N9K-2 はどちらも、同じユニキャスト ルーティング プロトコルでアクティブ化される SVI インターフェイスを備えており、Router と同じブロードキャストドメイン内にあります。

vPC ピアゲートウェイの機能拡張が有効になっていない場合、vPC を介したユニキャスト ルーティング プロトコル隣接関係は、vPC 接続ルータの ECMP ハッシュ決定とそのレイヤ 2 ポート

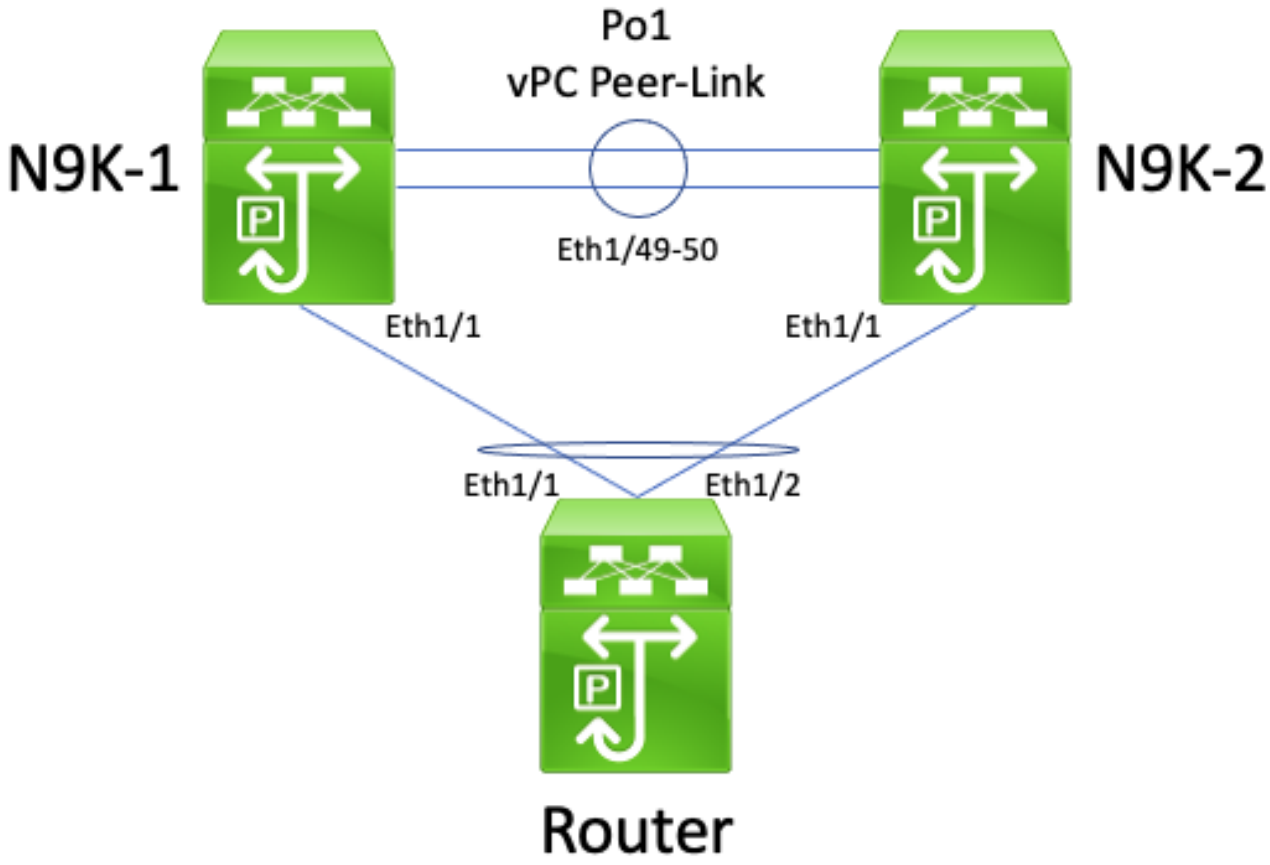
チャネルハッシュ決定が異なる可能性があるため、サポートされません。このトポロジでは、ルータ、N9K-1、およびN9K-2の間でルーティングプロトコルの隣接関係が正常に形成されます。ルータとホスト間のトラフィックフローを考慮します。Router を通過する、ホスト宛てのデータプレーントラフィックは、N9K-1 の SVI MAC アドレスに属している宛先 MAC アドレスで書き換えられる場合があります（ルータによる ECMP ハッシュ決定のため）が、Ethernet1/2 インターフェイスから送信されます（ルータによるレイヤ 2 ポートチャネルハッシュ決定のため）。

宛先MACアドレスがN9K-1に属し、vPCピアゲートウェイ機能拡張（N9K-2がN9K-1の代わりにパケットをルーティングできるようにする）が有効になっていないため、N9K-2はこのパケットを受信し、vPCピアリンク経由で転送します。N9K-1はこのパケットをvPCピアリンクで受信し、vPC Po20のEthernet1/2からパケットを転送する必要があることを認識します。これはvPCループ回避ルールに違反するため、N9K-1はハードウェアでパケットをドロップします。その結果、このトポロジで vPC ドメインを通過する一部のフローについて、接続の問題またはパケット損失が発生する可能性があります。

vPC ドメイン設定コマンドの peer-gateway を使用して vPC ピアゲートウェイの機能拡張を有効にしてから、vPC ドメイン設定コマンドの layer3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。中断を最小限に抑えるには、両方の vPC の機能拡張を短時間で連続して有効にすることで、「vPC ピアゲートウェイのない、vPC を介したユニキャスト ルーティング プロトコル隣接関係」で説明されている障害シナリオが発生する時間をなくす必要があります。

vPC ピアゲートウェイのある、vPC を介したユニキャスト ルーティング プロトコル隣接関係

次に示すトポロジについて考えてみます。




このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。Po1 インターフェイスは vPC ピアリンクです。ルータのホスト名を持つルータが、vPC Po10を介してN9K-1およびN9K-2に接続されています。ルータのPo10インターフェイスは、ユニキャストルーティングプロトコルの下でアクティブになるルーテッドポートチャネルです。N9K-1 と N9K-2 はどちらも、同じユニキャスト ルーティング プロトコルでアクティブ化される SVI インターフェイスを備えており、Router と同じブロードキャストドメイン内にあります。

vPC ピアゲートウェイの機能拡張が有効になっている場合、vPC を介したユニキャスト ルーティング プロトコル隣接関係は、vPC ピアゲートウェイの機能拡張のために vPC 接続ルータと両方の vPC ピアの間でユニキャスト ルーティング プロトコル隣接関係が形成されない可能性があるため、サポートされません。このトポロジでは、ルータによってN9K-1またはN9K-2のいずれかに発信されたユニキャストルーティングプロトコルパケットがvPC Po10を経由してどのようにハッシュされるかによっては、ルータとN9K-1またはN9K-2との間のルーティングプロトコル隣接関係が正常に確立されない場合があります。

リンクローカル マルチキャスト ルーティング プロトコル パケット ( 一般に「Hello」パケットと呼ばれます ) は vPC VLAN に正常にフラッディングされるため、すべてのルータが、これらのパケットを問題なく送受信できます。ただし、Router のレイヤ 2 ポートチャネルハッシュ決定のために、Router から送信される N9K-1 宛てのユニキャスト ルーティング プロトコル パケットが Ethernet1/2 を出て N9K-2 に向かうシナリオについて考えてみます。このパケットはN9K-1のSVI MACアドレス宛てですが、N9K-2のEthernet1/1インターフェイスに入ります。N9K-2は、パケットの宛先がN9K-1のSVI MACアドレスであることを認識します。このSVIアドレスは、vPCピアゲートウェイ機能拡張が有効になっているため、N9K-2のMACアドレステーブルに「G」フラグま

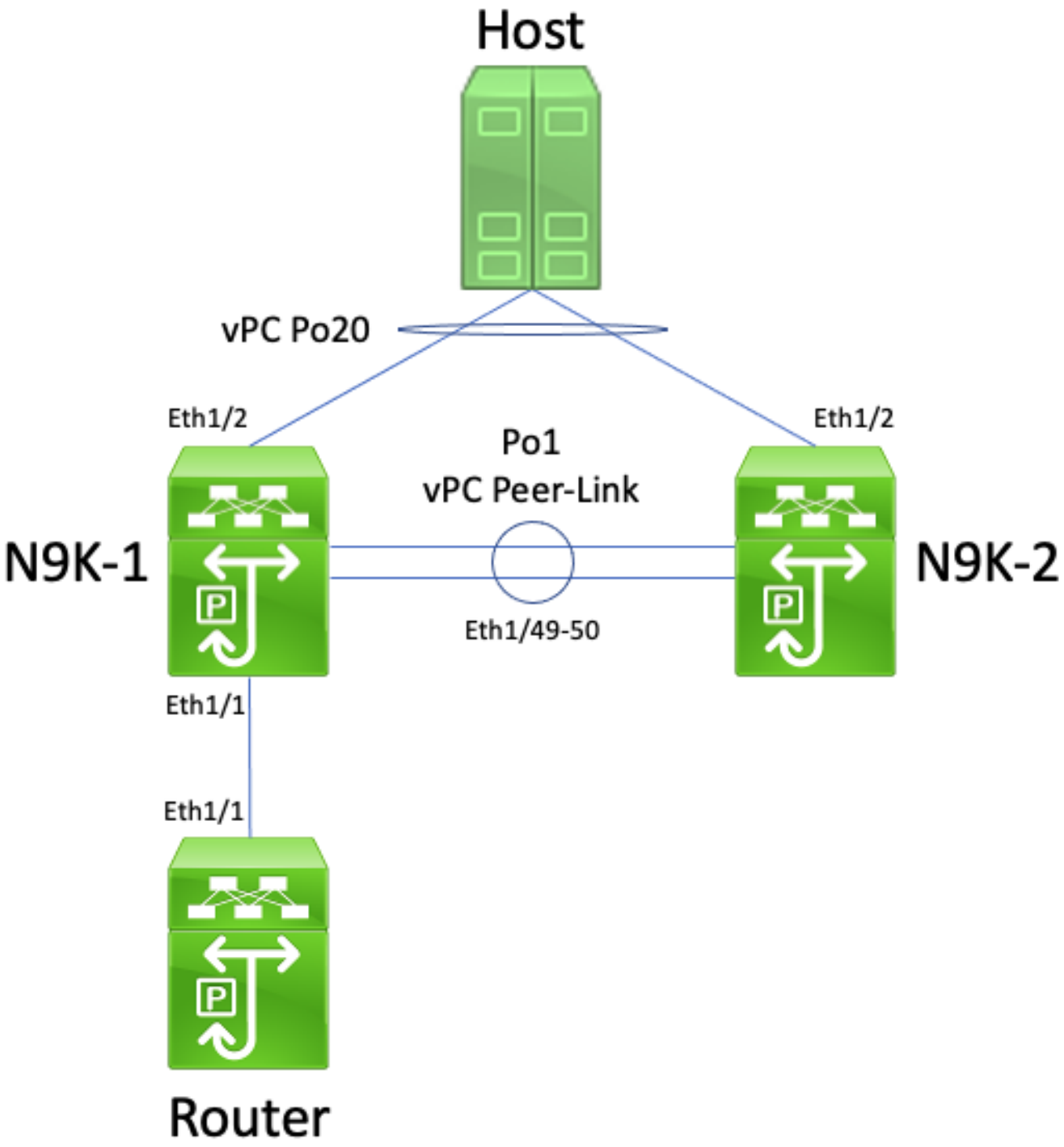
たは「Gateway」フラグとともにインストールされます。その結果、N9K-2はN9K-1の代わりにユニキャストルーティングプロトコルパケットをローカルでルーティングしようとします。

ただし、パケットをルーティングすることによって、パケットの存続可能時間(TTL)が減少し、ほとんどのユニキャストルーティングプロトコルパケットのTTLは1になります。その結果、パケットのTTLは0まで減少し、N9K-2だけ廃棄されます。N9K-1の観点からは、N9K-1はルータからリンクローカルマルチキャストルーティングプロトコルパケットを受信しており、ルータにユニキャストルーティングプロトコルパケットを送信できますが、ルータからユニキャストルーティングプロトコルパケットを受信していません。その結果、N9K-1はルータとのルーティングプロトコルの隣接関係を切断し、ルーティングプロトコルのローカル有限状態マシン(FSM)を再起動します。同様に、ルータはルーティングプロトコルに対してローカル有限状態マシンを再起動します。

PC ドメイン設定コマンドの layer 3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。これにより、TTL が 1 のユニキャスト ルーティング プロトコル パケットを、パケットの TTL を減らすことなく vPC ピアリンクを介して転送できるようになります。その結果、ユニキャスト ルーティング プロトコル隣接関係を、vPC または vPC VLAN を介して問題なく形成できます。

vPC ピアゲートウェイのない、vPC VLAN を介したユニキャスト ルーティングプロトコル隣接関係

次に示すトポロジについて考えてみます。




このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっていない vPC ドメイン内の vPC ピアです。Po1 インターフェイスは vPC ピアリンクです。ホスト名がルータであるルータは、Ethernet1/1を介してN9K-1のEthernet1/1に接続されています。ルータのEthernet1/1インターフェイスは、ユニキャストルーティングプロトコルでアクティブになっているルーテッドインターフェイスです。N9K-1 と N9K-2 はどちらも、同じユニキャスト ルーティング プロトコルでアクティブ化される SVI インターフェイスを備えており、Router と同じブロードキャストドメイン内にあります。

vPC ピアゲートウェイの機能拡張が有効になっていない場合、vPC VLAN を介したユニキャスト ルーティング プロトコル隣接関係は、vPC VLAN 接続ルータの ECMP ハッシュ決定のために N9K-2 が vPC ループ回避ルール違反としてデータプレーントラフィックをドロップする可能性が

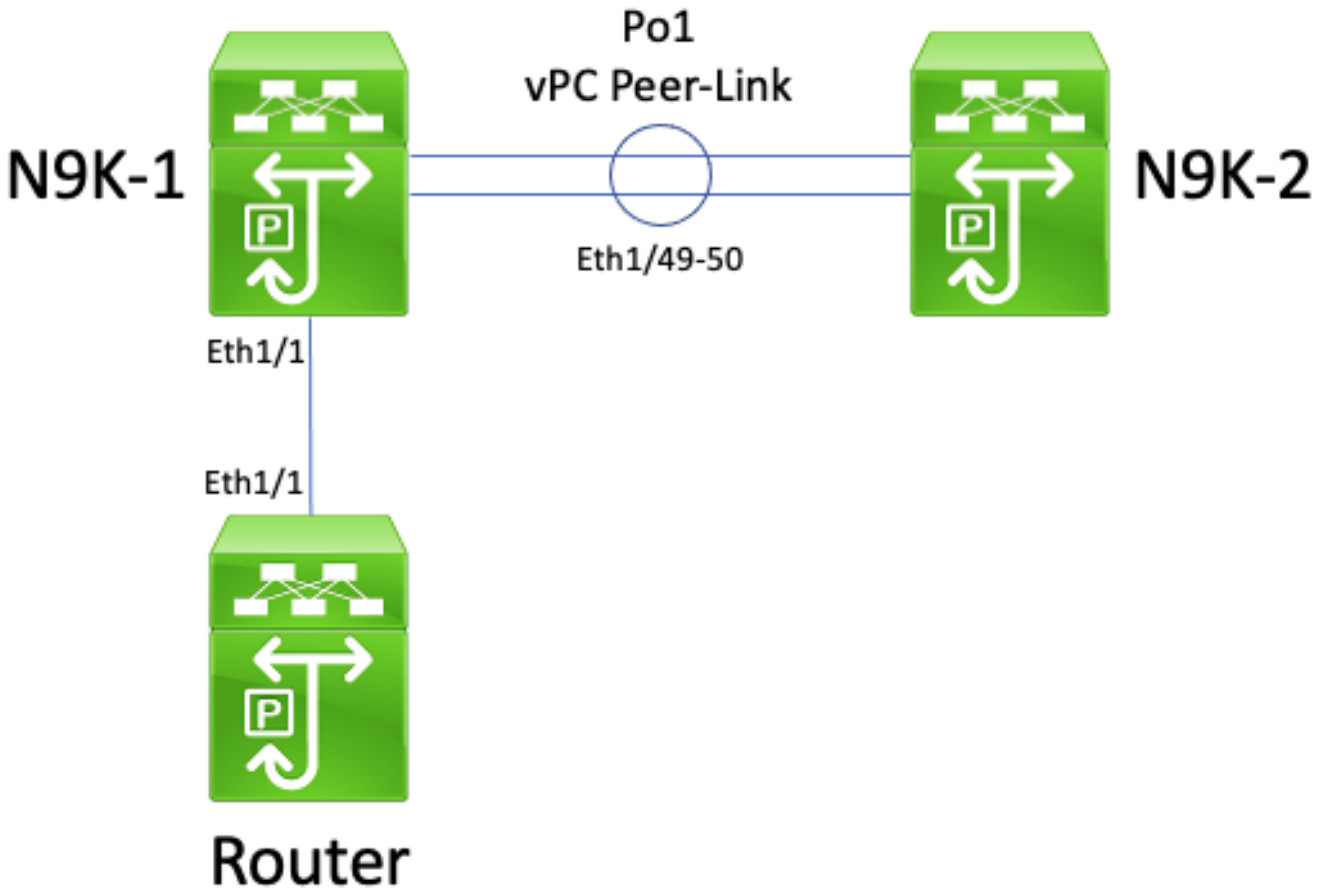
あるため、サポートされません。このトポロジでは、ルータ、N9K-1、およびN9K-2の間でルーティングプロトコルの隣接関係が正常に形成されます。ルータとホスト間のトラフィックフローを考慮します。Router を通過する、ホスト宛てのデータプレーントラフィックは、N9K-2 の SVI MAC アドレスに属している宛先 MAC アドレスで書き換えられる場合があります（ルータによる ECMP ハッシュ決定のため）が、Ethernet1/1 インターフェイスから N9K-1 に送信されます。

宛先MACアドレスがN9K-2に属し、vPCピアゲートウェイ機能拡張（N9K-2の代わりにN9K-1がパケットをルーティングできるようにする）が有効になっていないため、N9K-1はこのパケットを受信し、vPCピアリンク経由で転送します。N9K-2はこのパケットをvPCピアリンクで受信し、vPC Po20のEthernet1/2からパケットを転送する必要があることを認識します。これはvPCループ回避ルールに違反するため、N9K-2はハードウェアでパケットをドロップします。その結果、このトポロジで vPC ドメインを通過する一部のフローについて、接続の問題またはパケット損失が発生する可能性があります。

vPC ドメイン設定コマンドの peer-gateway を使用して vPC ピアゲートウェイの機能拡張を有効にしてから、vPC ドメイン設定コマンドの layer3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。中断を最小限に抑えるには、両方の vPC の機能拡張を短時間で連続して有効にすることで、「vPC ピアゲートウェイのない、vPC を介したユニキャスト ルーティング プロトコル隣接関係」で説明されている障害シナリオが発生する時間をなくす必要があります。

vPC ピアゲートウェイのある、vPC VLAN を介したユニキャスト ルーティングプロトコル隣接関係

次に示すトポロジについて考えてみます。




このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。Po1 インターフェイスは vPC ピアリンクです。ホスト名がルータであるルータは、Ethernet1/1を介してN9K-1のEthernet1/1に接続されています。ルータのEthernet1/1インターフェイスは、ユニキャストルーティングプロトコルでアクティブになっているルーテッドインターフェイスです。N9K-1 と N9K-2 はどちらも、同じユニキャスト ルーティング プロトコルでアクティブ化される SVI インターフェイスを備えており、Router と同じブロードキャストドメイン内にあります。

vPCピアゲートウェイ機能拡張が有効になっているvPC VLAN上のユニキャストルーティングプロトコルの隣接関係はサポートされません。これは、vPCピアゲートウェイ機能拡張により、vPC VLAN接続ルータと、vPC VLAN接続ルータが直接接続されていないvPCピアとの間でユニキャストルーティングプロトコルの隣接関係が形成されないためです。このトポロジでは、vPCピアゲートウェイ機能拡張が有効になっているため、N9K-2のSVI MACアドレスを宛先とするN9K-1ルーティングユニキャストルーティングプロトコルパケットがN9K-2とN9K-2の間のルーティングプロトコル隣接関係を正常に確立できません。パケットはルーティングされているため、存続可能時間（TTL）を減らす必要があります。通常、ユニキャスト ルーティング プロトコル パケットの TTL は 1 であり、パケットの TTL を減らして 0 にするルータは、そのパケットをドロップする必要があります。

リンクローカル マルチキャスト ルーティング プロトコル パケット（一般に「Hello」パケットと呼ばれます）は vPC VLAN に正常にフラッディングされるため、すべてのルータが、これらのパケットを問題なく送受信できます。ただし、送信元がルータで宛先がN9K-2のユニキャストルーティングプロトコルパケットがEthernet1/1からN9K-1に向けて出力されるシナリオを考えます。

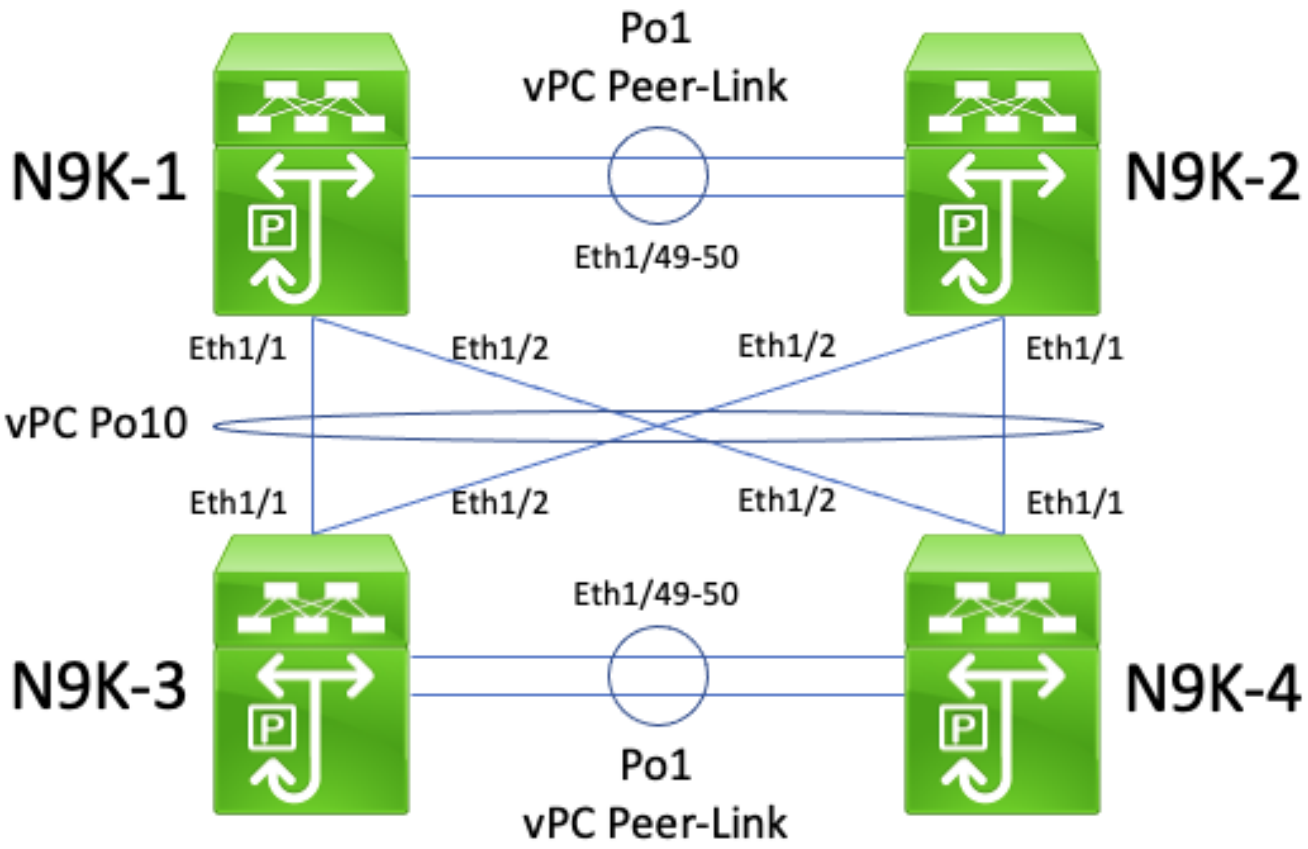
このパケットはN9K-2のSVI MACアドレス宛てですが、N9K-1のEthernet1/1インターフェイスに入ります。N9K-1は、パケットの宛先がN9K-2のSVI MACアドレスであることを認識します。このSVI MACアドレスは、vPCピアゲートウェイ機能拡張が有効になっているため、N9K-1のMACアドレステーブルに「G」フラグ ( 「ゲートウェイ」 ) とともにインストールされます。その結果、N9K-1はN9K-2の代わりにユニキャストルーティングプロトコルパケットをローカルでルーティングしようとします。

ただし、パケットをルーティングすることによって、パケットのTTLは減少し、ほとんどのユニキャストルーティングプロトコルパケットのTTLは1になります。その結果、パケットのTTLは0まで減少し、N9K-1だけドロップされます。N9K-2から見ると、N9K-2はルータからリンクローカルマルチキャストルーティングプロトコルパケットを受信しており、ルータにユニキャストルーティングプロトコルパケットを送信できますが、ルータからユニキャストルーティングプロトコルパケットを受信していません。その結果、N9K-2はルータとのルーティングプロトコルの隣接関係を切断し、ルーティングプロトコルのローカル有限状態マシン(FSM)を再起動します。同様に、ルータはルーティングプロトコルに対してローカル有限状態マシンを再起動します。

PC ドメイン設定コマンドの layer 3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。これにより、TTL が 1 のユニキャスト ルーティング プロトコル パケットを、パケットの TTL を減らすことなく vPC ピアリンクを介して転送できるようになります。その結果、ユニキャスト ルーティング プロトコル隣接関係を、vPC または vPC VLAN を介して問題なく形成できます。

vPC ピアゲートウェイのある、バックツーバック vPC を介したユニキャスト ルーティング プロトコル隣接関係

次に示すトポロジについて考えてみます。



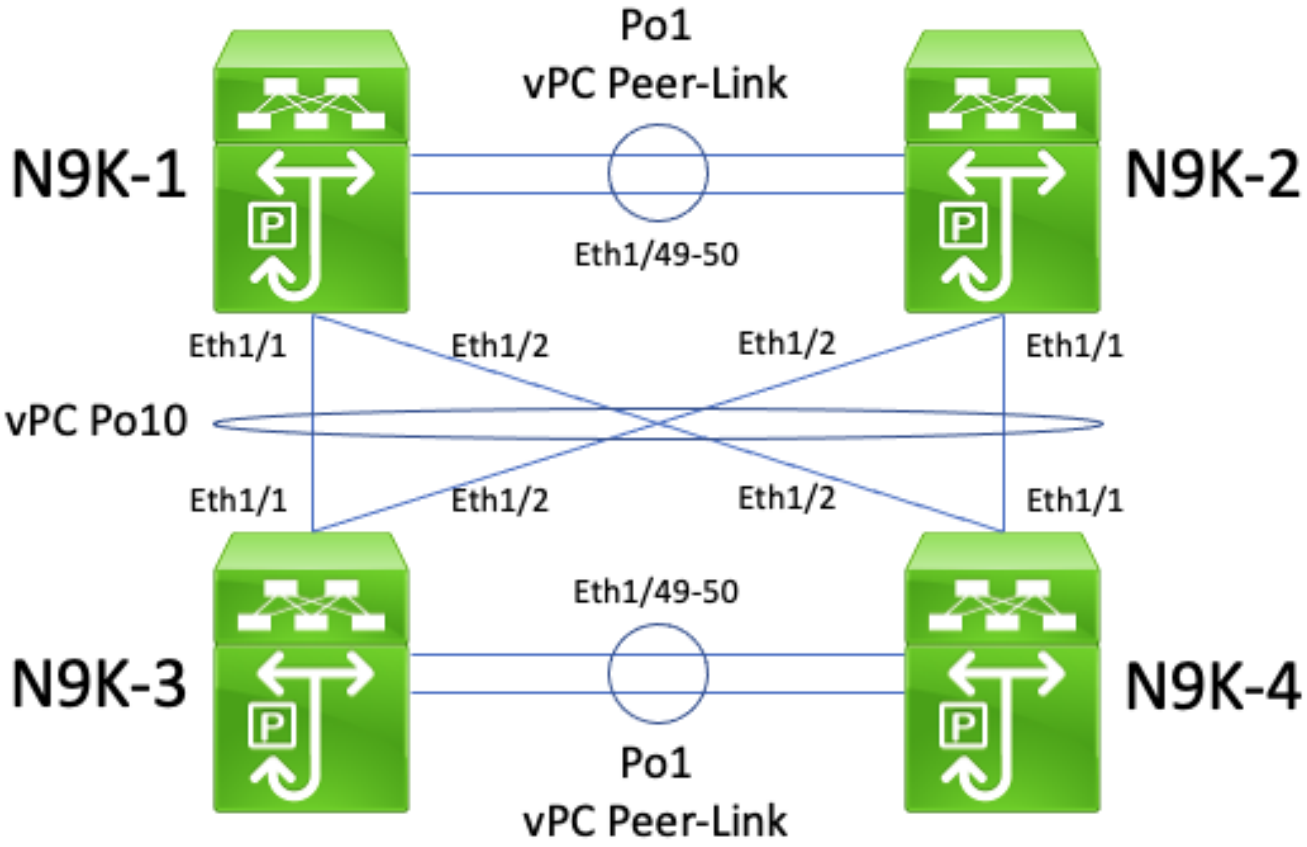

このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。Nexus スイッチの N9K-3 と N9K-4 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。両方のvPCドメインは、バックツーバックのvPC Po10を介して相互に接続されています。4つのスイッチはすべて、ユニキャストルーティングプロトコルでアクティブ化されたSVIインターフェイスを持ち、同じブロードキャストドメイン内にあります。

vPC ピアゲートウェイの機能拡張が有効になっている場合、バックツーバック vPC を介したユニキャスト ルーティング プロトコル隣接関係は、vPC ピアゲートウェイの機能拡張のために vPC ドメイン間でユニキャスト ルーティング プロトコル隣接関係が形成されない可能性があるため、サポートされません。このトポロジでは、N9K-1とN9K-3またはN9K-4（あるいはその両方）の間のルーティングプロトコルの隣接関係が、期待どおりに確立されない可能性があります。同様に、N9K-2 と N9K-3 または N9K-4 のいずれか（または両方）の間のルーティングプロトコル隣接関係が予期どおりに稼働しない場合があります。これは、ユニキャスト ルーティング プロトコル パケットが、あるルータ（たとえば、N9K-3）宛てに送信されても、発信元ルータのレイヤ 2 ポートチャネルハッシュ決定に基づいて別のルータ（たとえば、N9K-4）に転送される可能性があるためです。

この問題の根本原因は、このドキュメントの「vPC ピアゲートウェイのある、vPC を介したユニキャスト ルーティング プロトコル隣接関係」で説明されている根本原因と同じです。PC ドメイン設定コマンドの layer 3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。これにより、TTL が 1 のユニキャスト ルーティング プロトコル パケットを、パケットの TTL を減らすことなく vPC ピアリンクを介して転送できるようになります。その結果、ユニキャスト ルーティング プロトコル隣接関係は、バックツーバック vPC を介して問題なく形成できます。

プレフィックスが OSPF LSDB に存在するがルーティングテーブルには存在しない、vPC ピアゲートウェイのある vPC を介した OSPF 隣接関係

次に示すトポロジについて考えてみます。




このトポロジでは、Nexus スイッチの N9K-1 と N9K-2 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。Nexus スイッチの N9K-3 と N9K-4 が、vPC ピアゲートウェイの機能拡張が有効になっている vPC ドメイン内の vPC ピアです。両方の vPCドメインは、バックツーバックのvPC Po10を介して相互に接続されています。4つのスイッチはすべて、ユニキャストルーティングプロトコルでアクティブ化されたSVIインターフェイスを持ち、同じブロードキャストドメイン内にあります。N9K-4 はブロードキャストドメインの OSPF 代表ルータ（DR）であり、N9K-3 はブロードキャストドメインの OSPF バックアップ代表ルータ（BDR）です。

このシナリオでは、N9K-1 と N9K-3 の間の OSPF 隣接関係が、両方のスイッチの Ethernet1/1 から出るユニキャスト OSPF パケットのために FULL 状態に移行します。同様に、N9K-2 と N9K-3 の間の OSPF 隣接関係が、両方のスイッチの Ethernet1/2 から出るユニキャスト OSPF パケットのために FULL 状態に移行します。

ただし、このドキュメントの「vPC ピアゲートウェイのある、バックツーバック vPC を介したユニキャスト ルーティング プロトコル隣接関係」で説明されているように、ユニキャスト OSPF パケットが両方のスイッチの Ethernet1/1 から出て、N9K-2 と N9K-4 によってドロップされるため、N9K-1 と N9K-4 の間の OSPF 隣接関係は EXSTART または EXCHANGE 状態でスタックします。同様に、このドキュメントの「vPC ピアゲートウェイのある、バックツーバック vPC を介したユニキャスト ルーティング プロトコル隣接関係」で説明されているように、ユニキャスト

OSPF パケットが両方のスイッチの Ethernet1/2 から出て、N9K-1 と N9K-3 によってドロップされるため、N9K-2 と N9K-4 の間の OSPF 隣接関係は EXSTART または EXCHANGE 状態でスタックします。

その結果、N9K-1 および N9K-2 は、ブロードキャストドメインの BDR との FULL 状態になりますが、ブロードキャストドメインの DR との EXSTART または EXCHANGE 状態になります。ブロードキャストドメインの DR と BDR はどちらも、OSPF リンクステートデータベース（LSDB）の完全なコピーを保持しますが、DR または BDR のいずれかから OSPF を介して学習したプレフィックスをインストールするには、OSPF DROTHER ルータがブロードキャストドメインの DR との FULL 状態である必要があります。その結果、N9K-1とN9K-2の両方とも、N9K-3とN9K-4から学習されたプレフィックスがOSPF LSDBに存在しているように見えますが、これらのプレフィックスは、N9K-1とN9K-2がN9K-4（ブロードキャストドメインのDR）でFULL状態に移行するまで、ユニキャストルーティングテーブルにインストールされません。

PC ドメイン設定コマンドの layer 3 peer-router を使用して vPC を介したルーティング/レイヤ 3 の機能拡張を有効にすることにより、この問題を解決することができます。これにより、TTL が 1 のユニキャスト ルーティング プロトコル パケットを、パケットの TTL を減らすことなく vPC ピアリンクを介して転送できるようになります。その結果、ユニキャスト ルーティング プロトコル隣接関係は、バックツーバック vPC を介して問題なく形成できます。その結果、N9K-1（ブロードキャストドメインのDR）とN9K-2はN9K-4（ブロードキャストドメインのDR）でFULL状態に遷移し、OSPF経由でN9K-3とN9K-4から学習したプレフィックスをそれぞれのユニキャストルーティングテーブルに正常にインストールします。

## 関連情報

- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.3(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.2(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 10.1(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.3(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 9.2(x)』
- 『Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide, Release 7.x』
- 『Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide 8.x』
- 『Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide 7.x』
- 設計および設定ガイド：Cisco Nexus 7000シリーズスイッチの仮想ポートチャネル(vPC)のベストプラクティス
- Nexus プラットフォームにおける仮想ポート チャネルを介したルーティングでサポートされるトポロジ

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。