Nexus VPCループ回避について

内容

はじめに

このドキュメントでは、vPCループ回避がNexusベースのレイヤ3ネットワーク設計のトラフィック転送に影響を与える可能性があるシナリオについて説明します。

前提条件

要件

次の項目に関する知識があることが推奨されます。

- Nexus オペレーティング システムの CLI
- vPCの概念

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

- ソフトウェア10.4(4)

- ハードウェアN9K-C9364C-GX

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

今日のデータセンター環境では、冗長性とロードバランシングを実現するために、Cisco Nexus Virtual Port Channel(vPC)テクノロジーが不可欠です。vPCは、2つの異なるNexusスイッチへの接続を1つの論理ポートチャネルとして機能させることにより、ネットワークアーキテクチャを簡素化し、ダウンストリームデバイスの信頼性を向上させます。ただし、設定の詳細によっては、運用が複雑になる場合があります。

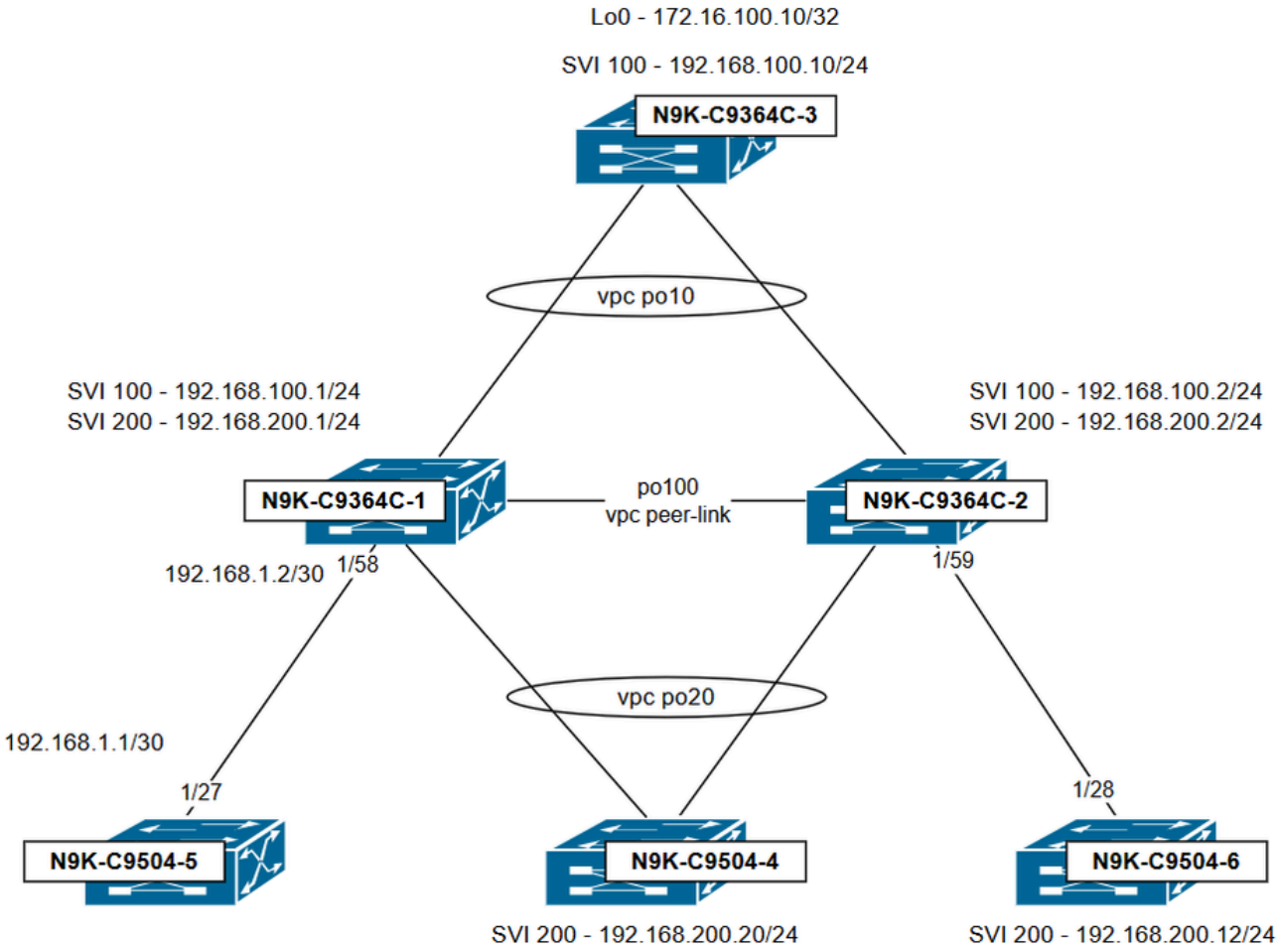
このドキュメントでは、vPCループ回避が重要になるシナリオについて説明し、トラフィック転送に対するその影響を検証します。このメカニズムを明確に理解することは、Nexusベースのインフラストラクチャで堅牢で効率的なレイヤ3接続を設計および維持し、トラフィックの中断を防ぎ、最適なネットワークパフォーマンスを維持することを目指しているネットワークエンジニアにとって非常に重要です。

問題

vPCを使用するCisco Nexus環境で、ネットワークオペレータは、vPCループ回避ルールによって発生する予期しないトラフィック転送動作を確認できます。トラフィックがvPCピアリンクを介して1つのvPCピアから別のvPCピアに移動する場合、トラフィックは両方のスイッチでアクティブなvPCポートチャネルを介して送信できません。その結果、すべての物理リンクがアップしているように見えても、接続のためにこのパスに依存しているデバイスでは、パケットのドロップや接続の喪失が発生する可能性があります。

復元力のあるネットワークトポロジの設計とトラブルシューティングでは、この動作を見過ごすと予期しないサービスの中断が発生し、ネットワークの問題の診断が困難になるため、vPCループ回避ルールを理解して説明することが不可欠です。

ネットワーク図



このトポロジでは、vPCドメインはN9K-C9364C-1およびN9K-C9364C-2によって作成されます。両方のスイッチがvPC VLANとしてVLAN 100および200で設定され、SVIがVLANごとに設定されます。vPCドメインは、これらのVLAN間のVLAN間ルーティングを行います。 特に指定がない限り、vPCピアスイッチ間で共有されるHSRP仮想IP(VIP)は、トポロジ内の他のスイッチによってデフォルトルートのネクストホップとして使用されます。

- N9K-C9364C-1 SVI設定

```
interface Vlan100
no shutdown
no ip redirects
ip address 192.168.100.1/24
no ipv6 redirects
hsrp 100
ip 192.168.100.254

interface Vlan200
no shutdown
no ip redirects
ip address 192.168.200.1/24
no ipv6 redirects
hsrp 200
```

ip 192.168.200.254

- N9K-C9364C-2 SVI設定

interface Vlan100
no shutdown
no ip redirects
ip address 192.168.100.2/24
no ipv6 redirects
hsrp 100
ip 192.168.100.254

interface Vlan200
no ip redirects
ip address 192.168.200.2/24
no ipv6 redirects
hsrp 200
ip 192.168.200.254

シナリオ

シナリオ1:vPCピアでvPC VLANのSVIが管理上シャットダウンされている

a) vPCからvPCへのルーテッドトラフィックが影響を受ける

動作シナリオでは、N9K-C9504-4(VLAN 200)からN9K-C9364C-3(VLAN 100)へのpingが成功します。 tracerouteは、接続パスが192.168.200.2を通過し、それがN9K-C9364C-2に割り当てられていることを示しています。

<#root>

N9K-C9504-4#

ping 192.168.100.10

```
PING 192.168.100.10 (192.168.100.10): 56 data bytes
64 bytes from 192.168.100.10: icmp_seq=0 ttl=253 time=8.48 ms
64 bytes from 192.168.100.10: icmp_seq=1 ttl=253 time=0.618 ms
64 bytes from 192.168.100.10: icmp_seq=2 ttl=253 time=0.582 ms
64 bytes from 192.168.100.10: icmp_seq=3 ttl=253 time=0.567 ms
64 bytes from 192.168.100.10: icmp_seq=4 ttl=253 time=0.55 ms

--- 192.168.100.10 ping statistics ---
```

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

```
round-trip min/avg/max = 0.55/2.159/8.48 ms
N9K-C9504-4#
```

<#root>

N9K-C9504-4#

traceroute 192.168.100.10

traceroute to 192.168.100.10 (192.168.100.10), 30 hops max, 40 byte packets

1 192.168.200.2 (192.168.200.2) 1.129 ms 0.602 ms 0.724 ms   <<<---- SVI 200 on N9K-C9364C-2

2 192.168.100.10 (192.168.100.10) 1.001 ms 0.657 ms 0.588 ms

この時点で、トラフィックフローは次のように動作しています。

- N9K-C9364C-2は、vPCドメイン内の共有HSRP仮想MAC(VMAC)に設定された宛先MACアドレスを使用して、192.168.100.10宛ての192.168.200.20からのトラフィックを受信します。
- HSRPはvPC上のデータプレーンの観点からはアクティブ – アクティブモードで動作するため、N9K-C9364C-2ではVLAN 200からVLAN 100へのトラフィックがルーティングされ、トラフィックはvPC 10経由で転送されます。

SVI 200がN9K-C9364C-2でシャットダウンされ、N9K-C9364C-1ではアクティブのままであるシナリオを考えます。

<#root>

N9K-C9364C-1#

show ip interface brief

IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status
Vlan100 192.168.100.1 protocol-up/link-up/admin-up

Vlan200 192.168.200.1 protocol-up/link-up/admin-up <<<---- SVI 200 is up

N9K-C9364C-1#

<#root>

N9K-C9364C-2#

show ip interface brief

IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status
Vlan100 192.168.100.2 protocol-up/link-up/admin-up

Vlan200 192.168.200.2 protocol-down/link-down/admin-down <<<---- SVI 200 is down

```
N9K-C9364C-2#
```

vPCピア間でSVIの動作ステータスが異なるため、vPCドメイン内でタイプ2の不整合が検出されます。

<#root>

```
N9K-C9364C-1#
```

show vPC

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

vPC domain id : 100
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 : failed
```

Type-2 inconsistency reason : SVI type-2 configuration incompatible

```
vPC role : primary
Number of vPCs configured : 2
Peer Gateway : Enabled
Dual-active excluded VLANs : -
Graceful Consistency Check : Enabled
Auto-recovery status : Disabled
Delay-restore status : Timer is off.(timeout = 30s)
Delay-restore SVI status : Timer is off.(timeout = 10s)
Delay-restore Orphan-port status : Timer is off.(timeout = 0s)
Operational Layer3 Peer-router : Disabled
Virtual-peerlink mode : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------
1 Po100 up 1,100,200

vPC status
----------------------------------------------------------------------------
Id Port Status Consistency Reason Active vlans
-- ------------ ------ ----------- ------ ---------------
10 Po10 up success success 1,100,200
20 Po20 up success success 1,100,200
N9K-C9364C-1#
```

<#root>

```
N9K-C9364C-2#
```

show vPC

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

vPC domain id : 100
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 : failed
```

Type-2 inconsistency reason : SVI type-2 configuration incompatible

```
vPC role : secondary
Number of vPCs configured : 2
Peer Gateway : Enabled
Dual-active excluded VLANs : -
Graceful Consistency Check : Enabled
Auto-recovery status : Disabled
Delay-restore status : Timer is off.(timeout = 30s)
Delay-restore SVI status : Timer is off.(timeout = 10s)
Operational Layer3 Peer-router : Disabled
Virtual-peerlink mode : Disabled

vPC Peer-link status
---------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------
1 Po100 up 1,100,200

vPC status
----------------------------------------------------------------------------
Id Port Status Consistency Reason Active vlans
-- ------------ ------ ----------- ------ ---------------
10 Po10 up success success 1,100,200
20 Po20 up success success 1,100,200

N9K-C9364C-2#
```

この段階で、192.168.200.20から192.168.100.10へのトラフィックは成功しなくなります。

<#root>

```
N9K-C9504-4#
```

ping 192.168.100.10

```
PING 192.168.100.10 (192.168.100.10): 56 data bytes
Request 0 timed out
Request 1 timed out
Request 2 timed out
Request 3 timed out
Request 4 timed out

--- 192.168.100.10 ping statistics ---
5 packets transmitted, 0 packets received, 100.00% packet loss
N9K-C9504-4#
```

色付きのping（MTUサイズが指定されたping）を使用して、このトラフィックが通過するパスをトレースします。

<#root>

N9K-C9504-4#

ping 192.168.100.10 count 100 timeout 0 packet-size 1030

PING 192.168.100.10 (192.168.100.10): 1030 data bytes
Request 0 timed out
Request 1 timed out
---- snip -----
Request 98 timed out
Request 99 timed out

--- 192.168.100.10 ping statistics ---

100 packets transmitted, 0 packets received, 100.00% packet loss

N9K-C9504-4# ^C
N9K-C9504-4#

N9K-C9364C-2のインターフェイスカウンタに従って、このトラフィックはポートチャネル20で受信され、ポートチャネル100（vPCピアリンク）に転送されます。

<#root>

N9K-C9364C-2#

show interface port-channel 20 counters detailed all | i "1024 to|po" ; sh int port-channel 10 counters

port-channel20

52.

Rx Packets from 1024 to 1518 bytes: = 100   <<<----- Ingress vPC po20

60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0

port-channel100

52. Rx Packets from 1024 to 1518 bytes: = 0
60.

Tx Packets from 1024 to 1518 bytes: = 100 <<<----- Egress po100 (vPC peer-link)

```
N9K-C9364C-2#
```

この動作は、SVI 200がN9K-C9364C-2上でシャットダウンされ、VLAN 200のトラフィックのローカルルーティングを妨げるためです。このシナリオでは、トラフィックはvPCピアリンク経由でN9K-C9364C-1にブリッジされるため、デバイスはVLAN間ルーティングを実行します。

N9K-C9364C-1のインターフェイスカウンタを見ると、vPCピアリンク経由でこのデバイスにパケットが到達していることが確認できますが、192.168.100.10に接続されているvPCポートチャネル10では発信パケットが観察されません。

<#root>

```
N9K-C9364C-1#
```

show interface port-channel 20 counters detailed all | i "1024 to|po" ; sh int port-channel 10 counters

```
port-channel20
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60.
```

Tx Packets from 1024 to 1518 bytes: = 0 <<<----- Expected egress vPC po10. No packets!!!

```
port-channel100
52.
```

Rx Packets from 1024 to 1518 bytes: = 100 <<<----- Ingress po100 (vPC peer-link)

```
60. Tx Packets from 1024 to 1518 bytes: = 0
N9K-C9364C-1#
```

トラフィックはvPCピアリンクを介してN9K-C9364C-1に到達しますが、vPCポートチャネル10には転送されません。これは、このvPCのegress_vsl_dropビットが1に設定されているためです。これは、ピアスイッチで同じvPCポートチャネルが動作可能な場合に発生します（この場合はN9K-C9364C-2）。

<#root>

```
N9K-C9364C-1#
```

show system internal eltm info interface Po10 | i i vsl

egress_vsl_drop = 1

```
N9K-C9364C-1#
```

<#root>

N9K-C9364C-1#

show system internal vPCm info interface Po10 | i "Peer stat|Inform|vPC sta"

IF Elem Information:
MCECM DB Information:

vPC state: Up Old Compat Status: Pass

vPC Peer Information:

Peer state: Up <<<------------- vPC 10 up on peer

PSS Information:

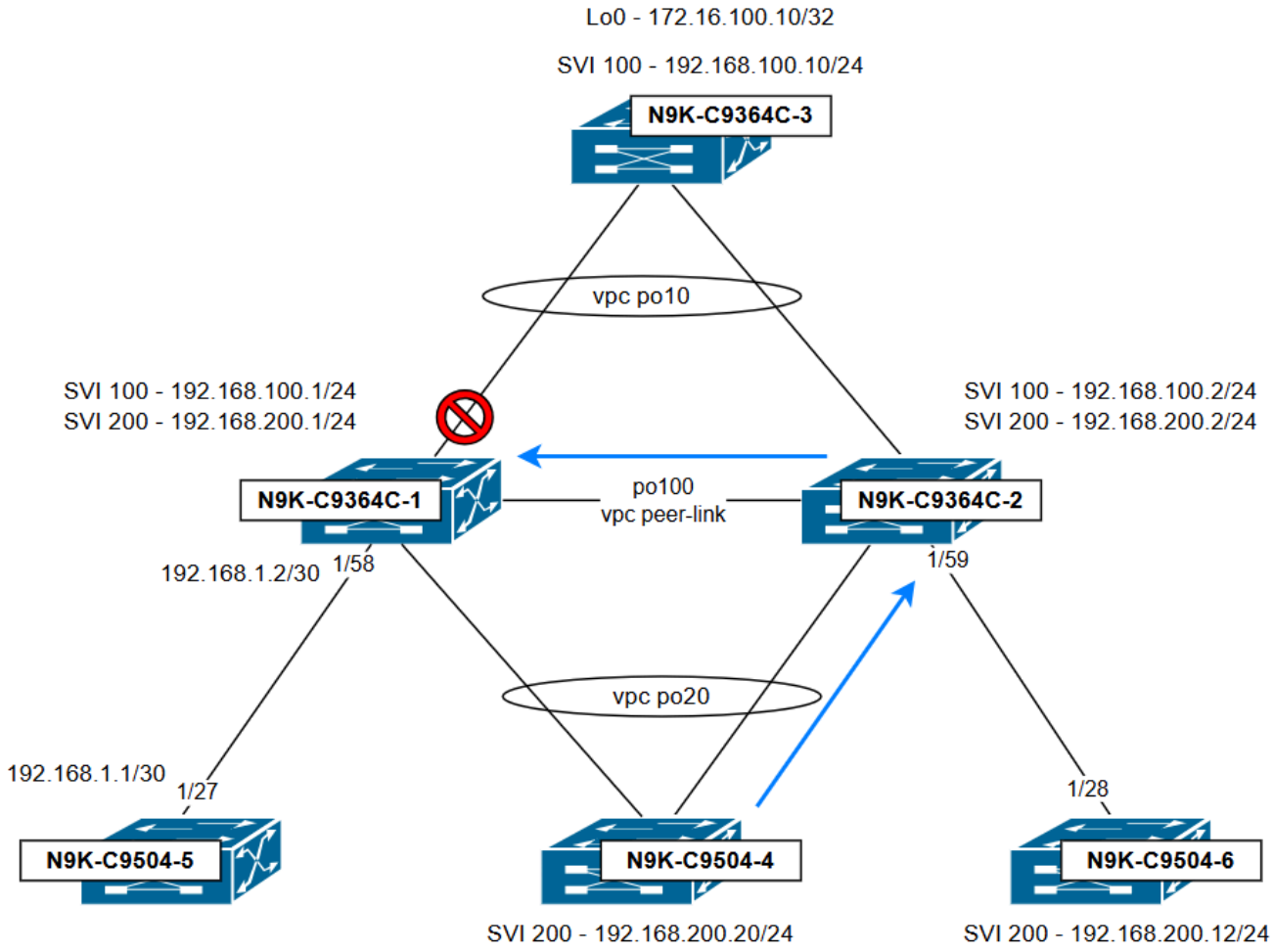
vPC state: Up Old Compat Status: Pass

vPC Peer Information:

Peer state: Up <<<------------- vPC 10 up on peer

Shared Database Information:
Application database Information:
Lock Information:
N9K-C9364C-1#

トラフィックフローと、トラフィックフローがドロップされるポイントを示すトポロジ:



結論：

N9K-C9364C-1は、vPCループ回避ルールが原因でトラフィックをドロップします。vPCピアリンク経由で受信したトラフィックは、両方のスイッチでアクティブなvPCポートチャネルから転送できません。」この問題を回避するには、SVIの管理ステータスが両方のスイッチで一致し、設定が対称的であることを確認します。

b)孤立ホストからvPCホストへのルーテッドトラフィックが影響を受ける

SVI 200がN9K-C9364C-2でシャットダウンされ、N9K-C9364C-1でアクティブのままであるという同じシナリオを検討します。N9K-C9504-6(VLAN 200)からN9K-C9364C-3(VLAN 100)へのpingが失敗します。

<#root>

```
N9K-C9504-6#

ping 192.168.100.10 packet-size 1030 count 100 timeout 0


PING 192.168.100.10 (192.168.100.10): 1030 data bytes
Request 0 timed out
Request 1 timed out
```

```
Request 2 timed out
---- snip -----
Request 97 timed out
Request 98 timed out
Request 99 timed out

--- 192.168.100.10 ping statistics ---

100 packets transmitted, 0 packets received, 100.00% packet loss


N9K-C9504-6#
```

色付きのping（MTUサイズが指定されたping）を使用して、このトラフィックが通過するパスをトレースします。

<#root>

```
N9K-C9364C-2#

show interface eth1/59 counters detailed all | i "1024 to|Eth" ; sh int port-channel 10 counters detaile


Ethernet1/59

52. Rx Packets from 1024 to 1518 bytes: = 100 <<<----- Ingress port to N9K-C9504-6


60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel100
52. Rx Packets from 1024 to 1518 bytes: = 0

60. Tx Packets from 1024 to 1518 bytes: = 100 <<<----- Egress po100 (vPC peer-link)


N9K-C9364C-2#
```

<#root>

```
N9K-C9364C-1#

show interface port-channel 10 counters detailed all | i "1024 to|po" ; sh int port-channel 100 counters


port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0

60. Tx Packets from 1024 to 1518 bytes: = 0 <<<----- Expected egress vPC po10. No packets!!!


port-channel100

52. Rx Packets from 1024 to 1518 bytes: = 100 <<<----- Ingress po100 (vPC peer-link)


60. Tx Packets from 1024 to 1518 bytes: = 0
```

```
N9K-C9364C-1#
```

トラフィックはvPCピアリンクを介してN9K-C9364C-1に到達しますが、vPCポートチャネル10には転送されません。これは、このvPCのegress_vsl_dropビットが1に設定されているためです。これは、ピアスイッチで同じvPCポートチャネルが動作可能な場合に発生します（この場合はN9K-C9364C-2）。

<#root>

```
N9K-C9364C-1#
```

show system internal eltm info interface Po10 | i i vsl

egress_vsl_drop = 1

```
N9K-C9364C-1#
```

<#root>

```
N9K-C9364C-1#
```

show system internal vpcm info interface Po10 | i "Peer stat|Inform|vPC sta"

```
IF Elem Information:
MCECM DB Information:
```

vPC state: Up Old Compat Status: Pass

```
vPC Peer Information:
```

Peer state: Up <<<------------ vPC 10 up on peer

```
PSS Information:
```

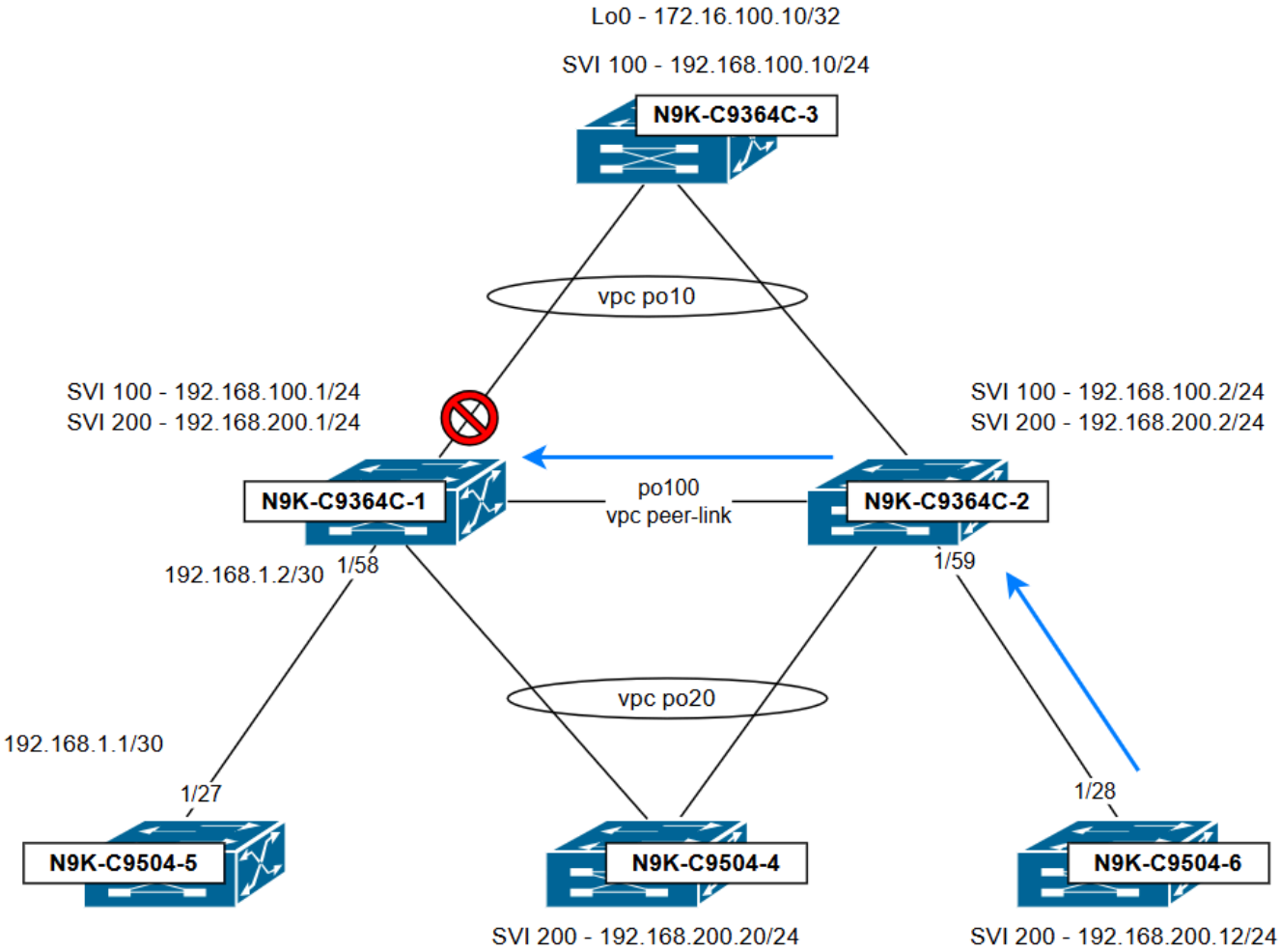
vPC state: Up Old Compat Status: Pass

```
vPC Peer Information:
```

Peer state: Up <<<------------ vPC 10 up on peer

```
Shared Database Information:
Application database Information:
Lock Information:
N9K-C9364C-1#
```

トラフィックフローと、トラフィックフローがドロップされるポイントを示すトポロジ：



結論：

トラフィックがN9K-C9364C-2に接続された孤立ホストから送信された場合でも、vPCループ回避ルールにより、トラフィックはN9K-C9364C-1によってドロップされます。vPCピアリンク経由で受信されたトラフィックは、両方のスイッチでアクティブなvPCポートチャネルからは転送できません。ピアスイッチ上の入力ポートがvPCであるか孤立ポートであるかは無関係です。重要なのは、トラフィックがvPCピアリンク経由で入力され、両方のスイッチ上でアクティブなvPC宛てに送信されることです。 この問題を回避するには、SVIの管理ステータスが両方のスイッチで一貫しており、設定が対称的であることを確認します。

シナリオ2:すべてのvPCとSVIが稼働している – ネクストホップがvPCピアを指している

このシナリオでは、vPCドメイン内のすべてのSVIおよびvPCポートチャネルがアップしています。ただし、レイヤ3インターフェイス経由でN9K-C9364C-1に接続されているN9K-C9504-5から、N9K-C9364C-3上のLoopback 0にpingを実行できません。

N9K-C9504-5からのtracerouteは、パケットが最初に192.168.1.2の即時ネクストホップに到達してから、N9K-C9364C-2に関連付けられた192.168.100.2に進むことを示します。

<#root>

```
N9K-C9504-5#
```

traceroute 172.16.100.10

```
traceroute to 172.16.100.10 (172.16.100.10), 30 hops max, 40 byte packets
1 192.168.1.2
```

(192.168.1.2)

```
 1.338 ms 0.912 ms 0.707 ms
2 192.168.100.2
```

(192.168.100.2)

```
 0.948 ms 0.751 ms 0.731 ms
3 * * *
4 * * *
N9K-C9504-5#
```

N9K-C9364C-1からのネクストホップ検証（このトラフィックの初期ホップ）は、宛先が192.168.100.2を介して到達可能であることを示しています。これはN9K-C9364C-2のSVI 100に対応します。

<#root>

```
N9K-C9364C-1#
```

show ip route 172.16.100.10

```
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

172.16.100.0/24, ubest/mbest: 1/0
*
```

via 192.168.100.2

```
, [1/0], 00:05:05, static
N9K-C9364C-1#
```

色付きのping（MTUサイズが指定されたping）を使用して、このトラフィックが通過するパスをトレースします。

<#root>

```
N9K-C9364C-1#
```

show interface e1/58 counters detailed all | i "1024 to|Eth" ; sh int port-channel 100 counters detailed

```
Ethernet1/58
52.
```

```
Rx Packets from 1024 to 1518 bytes: = 100 <<<----- Ingress Eth1/58


60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel100
52. Rx Packets from 1024 to 1518 bytes: = 0
60.

Tx Packets from 1024 to 1518 bytes: = 100 <<<----- Egress po100 (vPC peer-link)


port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
N9K-C9364C-1#
```

<#root>

```
N9K-C9364C-2# sh int port-channel 100 counters detailed all | i "1024 to|po" ; sh int port-channel 10 c
port-channel100
52.

Rx Packets from 1024 to 1518 bytes: = 100 <<<----- Ingress po100 (vPC peer-link)


60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60.

Tx Packets from 1024 to 1518 bytes: = 0 <<<----- Egress vPC po10, no packets!!!


N9K-C9364C-2#
```

トラフィックはvPCピアリンクを介してN9K-C9364C-2に到達しますが、vPCポートチャネル10には転送されません。これは、このvPCのegress_vsl_dropビットが1に設定されているためです。これは、同じvPCポートチャネルがピアスイッチで動作可能な場合に発生します（この場合はN9K-C9364C-1）。

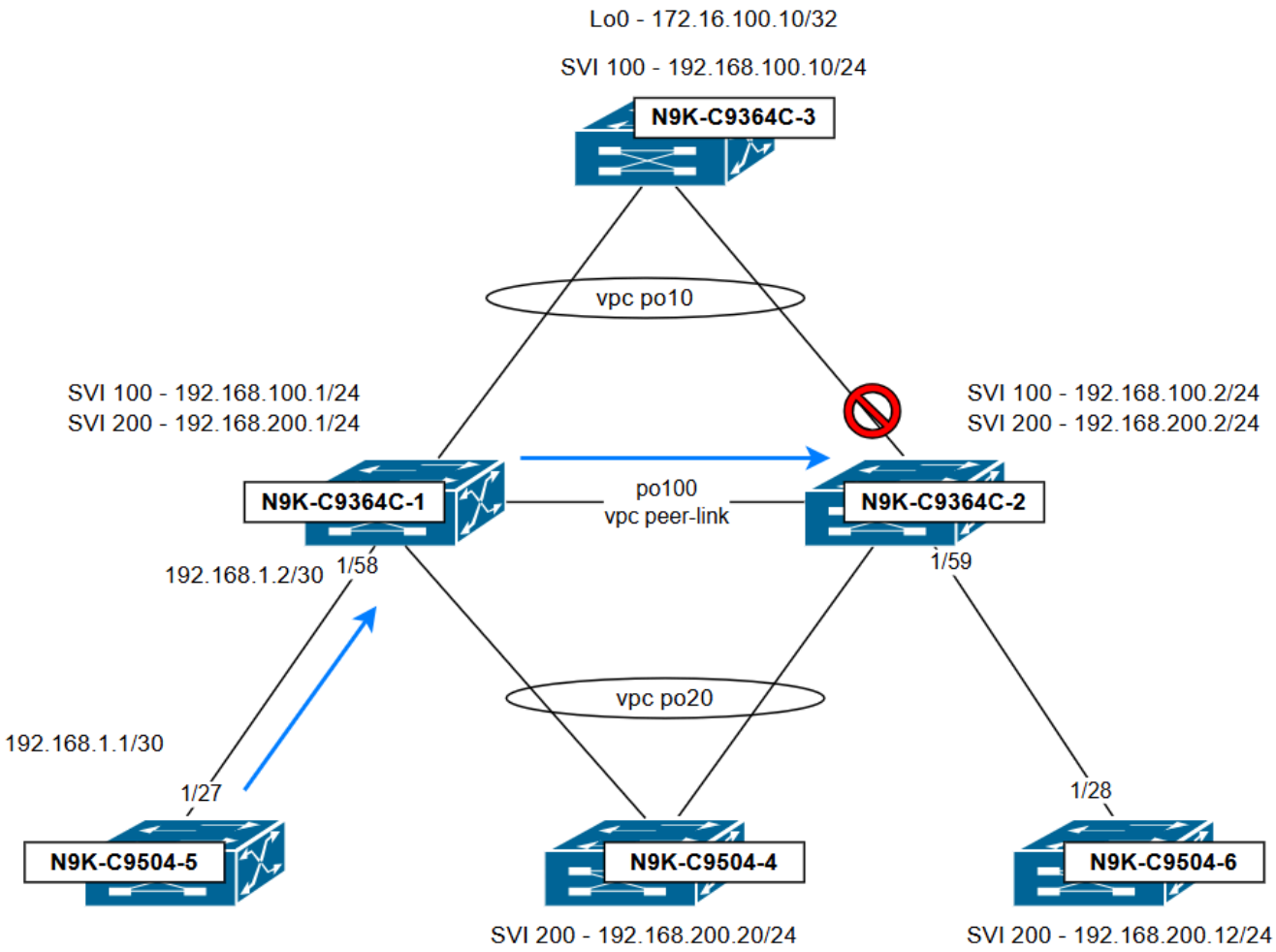
<#root>

```
N9K-C9364C-2#

show system internal eltm info interface Po10 | i i vsl


egress_vsl_drop = 1


N9K-C9364C-2#
```

<#root>

```
N9K-C9364C-2# show system internal vPCm info interface Po10 | i "Peer stat|Inform|vPC sta"
IF Elem Information:
MCECM DB Information:

vPC state: Up Old Compat Status: Pass


vPC Peer Information:

Peer state: Up <<<------------ vPC 10 up on peer


PSS Information:
vPC state: Up Old Compat Status: Pass
vPC Peer Information:

Peer state: Up <<<------------ vPC 10 up on peer


Shared Database Information:
Application database Information:
Lock Information:
N9K-C9364C-2#
```

トラフィックフローと、トラフィックフローがドロップされるポイントを示すトポロジ：



結論：

この問題は、N9K-C9364C-1がネクストホップとしてN9K-C9364C-2を使用し、vPC 10を経由して出ようとする前にvPCピアリンク経由でトラフィックを送信するために発生します。 トラフィックは、vPCループ回避ルールが原因でドロップされる:vPCピアリンク経由で受信したトラフィックは、両方のスイッチ上でアクティブなvPCポートチャネルから転送できません。 この問題を回避するには、トラフィックがvPCピアリンクを通過してvPC経由で出力される必要がないように、vPCポートチャネルを介したネクストホップを含むルート（ダイナミックまたはスタティック）が両方のvPCピアスイッチで設定されていることを確認します。

シナリオ3:すべてのvPCとSVIがアップの状態 – VPCピアゲートウェイ機能がオフ

このシナリオでは、すべてのSVIおよびvPCポートチャネルがvPCドメインでアップしていますが、vPCピアゲートウェイ機能はオフになっています。 この時点で、N9K-C9504-4(VLAN 200)はN9K-C9364C-3(VLAN 100)にpingを実行できません。

```
<#root>

N9K-C9504-4#

ping 192.168.100.10


PING 192.168.100.10 (192.168.100.10): 56 data bytes
Request 0 timed out
Request 1 timed out
Request 2 timed out
Request 3 timed out
Request 4 timed out

--- 192.168.100.10 ping statistics ---

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


N9K-C9504-4#
```

N9K-C9504-4からのネクストホップ検証は、宛先が192.168.200.2を介して到達可能であることを示しています。これは、N9K-C9364C-2上のSVI 200に対応し、vPCポートチャネル20を介して接続されています。

```
<#root>

N9K-C9504-4#

show ip route 192.168.100.10


IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
```

```
'%<string>' in via output denotes VRF <string>

0.0.0.0/0, ubest/mbest: 1/0
*via
```

192.168.200.2

```
, [1/0], 01:22:46, static
N9K-C9504-4#
```

<#root>

```
N9K-C9504-4#
```

show ip arp detail | i 192.168.200.2

192.168.200.2

```
 00:08:05
```

a478.06de.7edb

```
 Vlan200 port-channel20 default
```

色付きのping（MTUサイズが指定されたping）を使用して、このトラフィックが通過するパスをトレースします。ここで、インターフェイスカウンタは、N9K-C9364C-1がポートチャネル20を介して192.168.200.20から192.168.100.10へのトラフィックを受信し、それをvPCピアリンク（ポートチャネル100）に送信することを示します

<#root>

```
N9K-C9364C-1#
```

show interface port-channel 20 counters detailed all | i "1024 to|po" ; sh int port-channel 10 counters

```
port-channel20
52.
```

Rx Packets from 1024 to 1518 bytes: = 100   <<<----- Ingress vPC 20

```
60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel100
52. Rx Packets from 1024 to 1518 bytes: = 0
60.
```

Tx Packets from 1024 to 1518 bytes: = 100   <<<----- Egress po100 (vPC peer-link)

```
N9K-C9364C-1#
```

N9K-C9364C-2はvPCピアリンク（ポートチャネル100）経由でトラフィックを受信しますが、vPCポートチャネル10には転送しません。

<#root>

```
N9K-C9364C-2#
```

```
show int port-channel 20 counters detailed all | i "1024 to|po" ; sh int port-channel 10 counters detail
```

```
port-channel20
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
port-channel10
52. Rx Packets from 1024 to 1518 bytes: = 0
60. Tx Packets from 1024 to 1518 bytes: = 0
```

```
   <<<----- Egress vPC po10, no packets!!!
```

```
port-channel100
```

```
52. Rx Packets from 1024 to 1518 bytes: = 100  <<<----- Ingress po100 (vPC peer-link)
```

```
60. Tx Packets from 1024 to 1518 bytes: = 0
N9K-C9364C-2#
```

トラフィックはvPCピアリンクを介してN9K-C9364C-2に到達しますが、vPCポートチャネル10には転送されません。これは、このvPCのegress_vsl_dropビットが1に設定されているためです。これは、ピアスイッチで同じvPCポートチャネルが動作可能な場合に発生します（この場合はN9K-C9364C-1）。

ピアゲートウェイが無効になっているため、N9K-C9364C-1は自身のローカルMACアドレスにアドレス指定されたパケットだけをルーティングできます。その結果、a478.06de.7edb（N9K-C9364C-2からのMAC）宛てのパケットは、vPCピアリンク経由でN9K-C9364C-1によって転送されます。

<#root>

```
N9K-C9364C-1#
```

```
show mac address-table add a478.06de.7edb
```

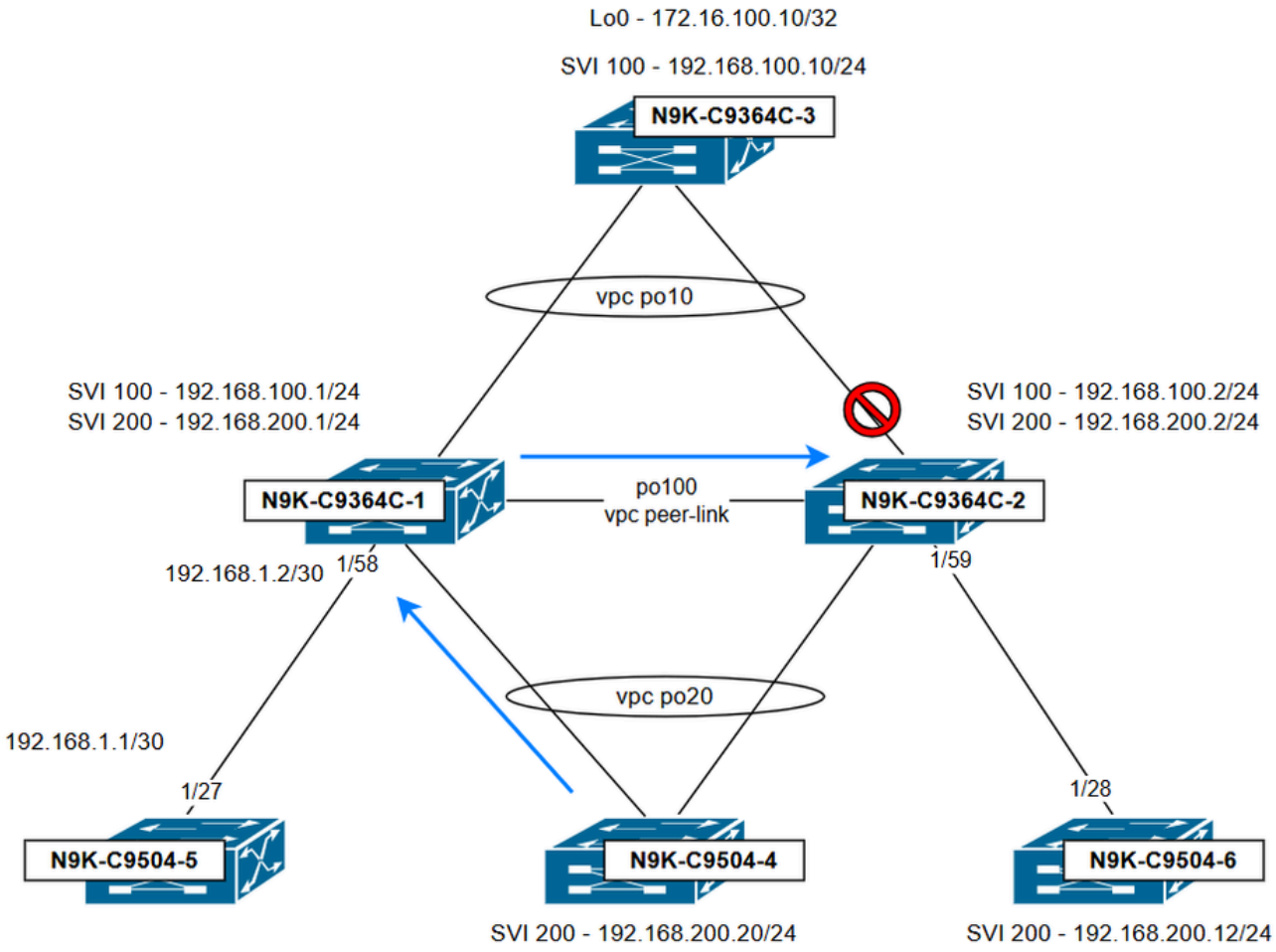
```
Legend:
* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen,+ - primary entry using vPC Peer-Link,
(T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
VLAN MAC Address Type age Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
* 100
```

```
a478.06de.7edb
```

```
 static - F F
```

```
vPC Peer-Link

(R)
* 200

a478.06de.7edb

 static - F F

vPC Peer-Link

(R)
N9K-C9364C-1#
```

トラフィックフローと、トラフィックフローがドロップされるポイントを示すトポロジ：



結論：

ピアゲートウェイが有効な場合、vPCピアのMACアドレス宛てのルーティングされたトラフィックは、ピアのMACをゲートウェイとしてプログラミングすることにより、ローカルで処理されます。これにより、vPCピアリンクがトラフィックパスで使用されなくなり、vPCループ回避ルールによるドロップが回避されます。このような問題を回避するには、vPCドメインでvPCピアゲートウェイ機能が有効になっていることを確認します。

ソリューションの概要

- vPC VLAN上でSVI設定の一貫性を維持する

vPCピアスイッチ間の非対称スイッチ仮想インターフェイス(SVI)設定は、トラフィックのブラックホールなどの重要なトラフィック転送問題の原因となる可能性があります。この状況に寄与する一般的ではあるがサポートされていない手法は、一方の側でSVIをシャットダウンしてvPCピア間のフェールオーバーをテストすることです。この方式では、Nexus vPCアーキテクチャでサポートされていない非対称のSVI状態が作成され、その結果、トラフィックのブラックホールと転送障害が発生します。ルーティングが必要なすべてのvPC VLANでSVI設定が常に一貫していることを確認します。

- vPCドメインでピアゲートウェイを有効にします。

ピアゲートウェイ機能は、Cisco Nexus vPCの導入における重要な拡張機能です。vPCドメインで有効にすると、各vPCピアスイッチがvPCピアの仮想MACアドレス宛てのパケットを受け入れて処理できるようになります。つまり、パケットを最初に受信したスイッチに関係なく、いずれかのvPCピアがゲートウェイに送られるトラフィックに応答できます。 ピアゲートウェイが有効になっていない場合、特定のタイプのトラフィック（デフォルトゲートウェイのMACアドレスに送信されるパケットなど）が1つのピアに到着した場合にドロップされる可能性があり、それ以外の場合は、ピアリンクを通過してvPCメンバーポートから発信される必要があります。vPCドメインでvPCピアゲートウェイが設定されていることを確認します。

関連情報

[仮想ポートチャネル（vPC）の拡張機能の理解]

[Nexus上の仮想ポートチャネル(vPC)のベストプラクティス]

[Nexus 7000のピアゲートウェイ機能]

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。