Entender a prevenção de loop do Nexus VPC

Contents

Introdução

Este documento descreve cenários em que a Prevenção de Loop vPC pode afetar o encaminhamento de tráfego em projetos de rede de Camada 3 baseados em Nexus.

Pré-requisitos

Requisitos

A Cisco recomenda que você tenha conhecimento destes tópicos:

- CLI do sistema operacional Nexus
- Conceitos de vPC

Componentes Utilizados

As informações neste documento são baseadas nestas versões de software e hardware:

- Software 10.4(4)

- Hardware N9K-C9364C-GX

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Informações de Apoio

Nos ambientes de data center atuais, a tecnologia Cisco Nexus Virtual Port Channel (vPC) é essencial para permitir redundância e balanceamento de carga. Ao permitir que conexões com dois switches Nexus separados funcionem como um único canal de porta lógica, o vPC simplifica a arquitetura de rede e melhora a confiabilidade para dispositivos downstream. No entanto, alguns detalhes de configuração podem apresentar complexidades operacionais.

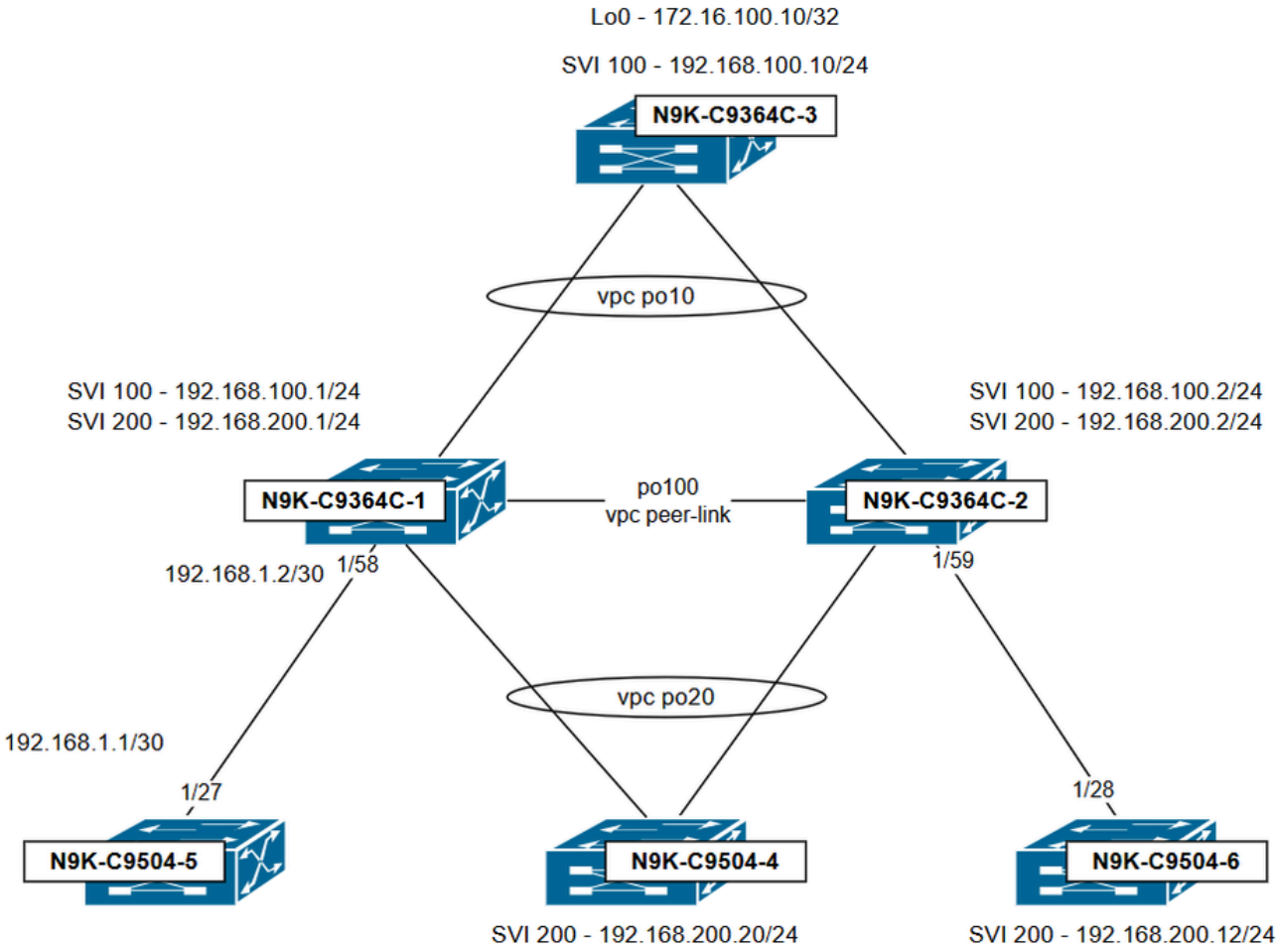
Este documento explora cenários onde a Prevenção de Loop vPC se torna significativa e examina seu impacto no encaminhamento de tráfego. Uma compreensão clara desse mecanismo é crucial para engenheiros de rede que buscam projetar e manter uma conectividade eficiente e robusta da camada 3 na infraestrutura baseada no Nexus, ajudando a evitar interrupções de tráfego e a manter o desempenho ideal da rede.

Problema

Em um ambiente Cisco Nexus usando vPC, os operadores de rede podem observar um comportamento inesperado de encaminhamento de tráfego causado pela regra de prevenção de loop do vPC. Quando o tráfego trafega de um peer vPC para outro pelo link peer vPC, ele não pode sair por nenhum canal de porta vPC que esteja ativo em ambos os switches. Como resultado, os dispositivos que dependem desse caminho para a conectividade podem experimentar pacotes descartados ou perda de conectividade, mesmo que todos os links físicos pareçam estar ativos.

Entender e contabilizar a regra de prevenção de loop do vPC é essencial para projetar e solucionar problemas de topologias de rede resilientes, já que ignorar esse comportamento pode levar a interrupções inesperadas do serviço e tornar os problemas de rede mais difíceis de diagnosticar.

Diagrama de Rede



Nessa topologia, o domínio vPC é feito por N9K-C9364C-1 e N9K-C9364C-2. Ambos os switches são configurados com as VLANs 100 e 200 como VLANs vPC, e as SVIs são configuradas para cada VLAN. O domínio vPC é responsável pelo roteamento entre VLANs entre essas VLANs. A menos que especificado de outra forma, o IP virtual (VIP) do HSRP compartilhado entre os switches pares do vPC é usado como o próximo salto para a rota padrão pelos outros switches na topologia.

- Configuração de SVI N9K-C9364C-1

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

interface Vlan200
no shutdown
no ip redirects
endereço ip 192.168.200.1/24
no ipv6 redirects
```

hsrp 200
ip 192.168.200.254

- Configuração de SVI N9K-C9364C-2

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

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

Cenários

Cenário 1: O SVI para VLAN vPC é desligado administrativamente no peer vPC

a) O tráfego roteado de vPC para vPC é afetado

Em um cenário de trabalho, o N9K-C9504-4 (VLAN 200) pode fazer ping com êxito no N9K-C9364C-3 (VLAN 100). Traceroute indica que o caminho de conexão passa por 192.168.200.2, que é atribuído a N9K-C9364C-2.

<#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
```

Neste ponto, o fluxo de tráfego está funcionando desta maneira:

- O N9K-C9364C-2 recebe o tráfego de 192.168.200.20 destinado a 192.168.100.10, com o endereço MAC de destino definido como o MAC virtual HSRP (VMAC) compartilhado dentro do domínio vPC.
- Como o HSRP opera no modo Ativo-Ativo a partir de uma perspectiva do plano de dados no vPC, o N9K-C9364C-2 roteia o tráfego da VLAN 200 para a VLAN 100 e o encaminha pelo vPC 10.

Considere um cenário em que o SVI 200 é desligado no N9K-C9364C-2, mas permanece ativo no N9K-C9364C-1:

```
<#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#
```

Devido à diferença no status operacional dos SVIs entre os pares do vPC, uma inconsistência de tipo 2 é detectada no domínio do vPC:

<#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#
```

Neste estágio, o tráfego de 192.168.200.20 a 192.168.100.10 não é mais bem-sucedido:

<#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#
```

Um ping colorido (um ping com um tamanho de MTU especificado) é usado para rastrear o caminho seguido por este tráfego:

<#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#
```

De acordo com os contadores de interface no N9K-C9364C-2, esse tráfego é recebido no canal de porta 20 e encaminhado para o canal de porta 100 (o link par do vPC):

<#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#
```

Esse comportamento ocorre porque o SVI 200 é desligado no N9K-C9364C-2, impedindo o roteamento local do tráfego para a VLAN 200. Nesse cenário, o tráfego é ligado através do link peer do vPC para o N9K-C9364C-1, para que o dispositivo execute o roteamento entre VLANs.

Examinando os contadores de interface em N9K-C9364C-1, confirma-se que os pacotes alcançam esse dispositivo pelo link par do vPC, no entanto, não há pacotes de saída observados no canal de porta 10 do vPC, que se conecta a 192.168.100.10.

<#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#
```

Mesmo que o tráfego chegue a N9K-C9364C-1 por meio do link de peer do vPC, ele não é encaminhado para o canal de porta 10 do vPC. Isso ocorre porque o bit egress_vsl_drop é definido como 1 para esse vPC, o que acontece quando o mesmo canal de porta do vPC está operacional no switch de peer (nesse caso, 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:
```

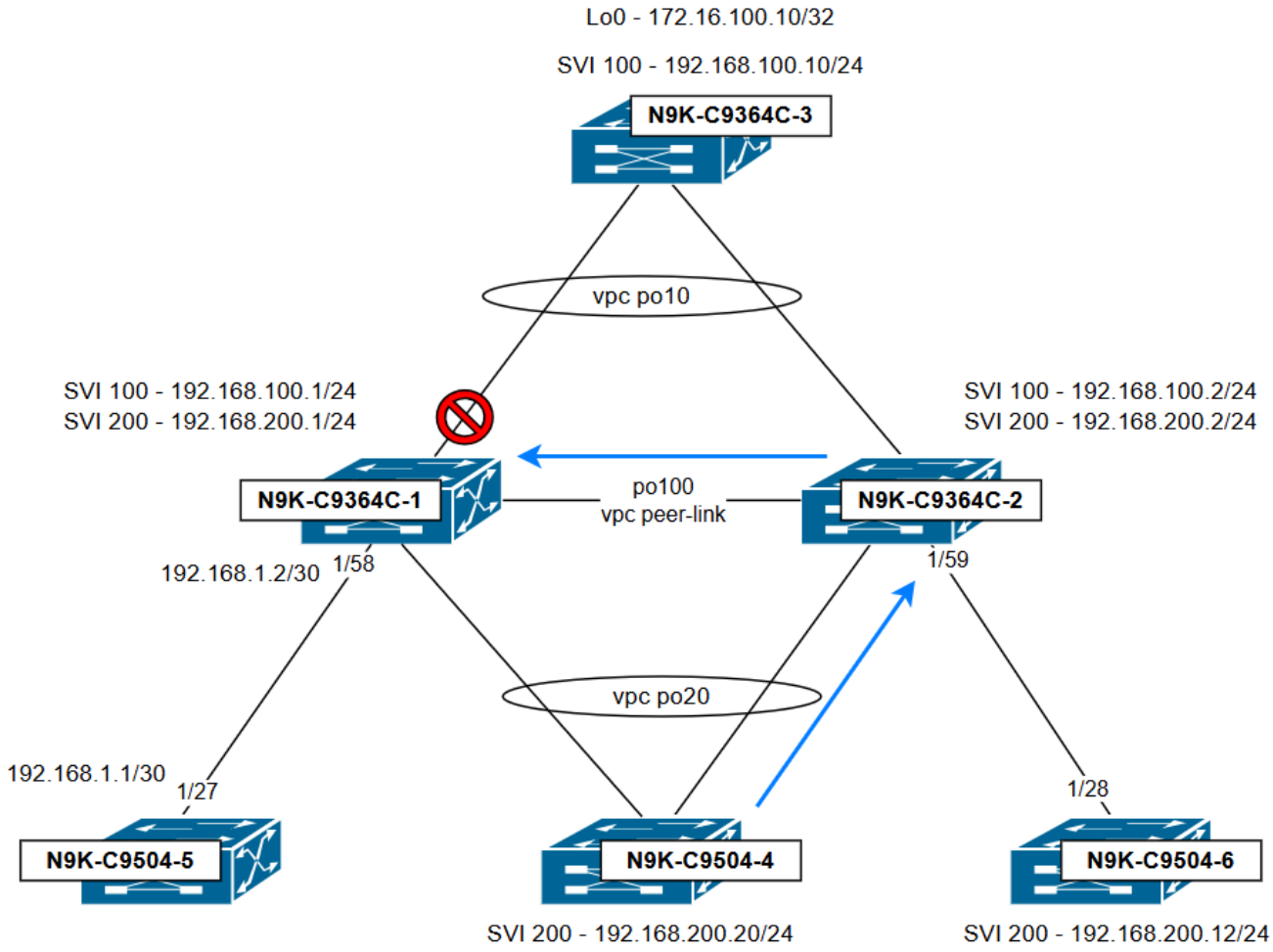
vPC state: Up Old Compat Status: Pass

```
vPC Peer Information:
```

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

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

Topologia que ilustra o fluxo de tráfego e o ponto em que ele é descartado:



Conclusão:

N9K-C9364C-1 descarta o tráfego devido à regra de prevenção de loop vPC: O tráfego recebido pelo link par do vPC não pode ser encaminhado por nenhum canal de porta do vPC que esteja ativo em ambos os switches."Para evitar esse problema, verifique se o status administrativo das SVIs é consistente em ambos os switches e se suas configurações são simétricas.

b) O tráfego roteado do host órfão para o vPC é afetado

Considerando o mesmo cenário em que o SVI 200 é desligado em N9K-C9364C-2, mas permanece ativo em N9K-C9364C-1. Um ping de N9K-C9504-6 (VLAN 200) para N9K-C9364C-3 (VLAN 100) não tem êxito.

<#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#
```

Um ping colorido (um ping com um tamanho de MTU especificado) é usado para rastrear o caminho seguido por este tráfego:

<#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
```

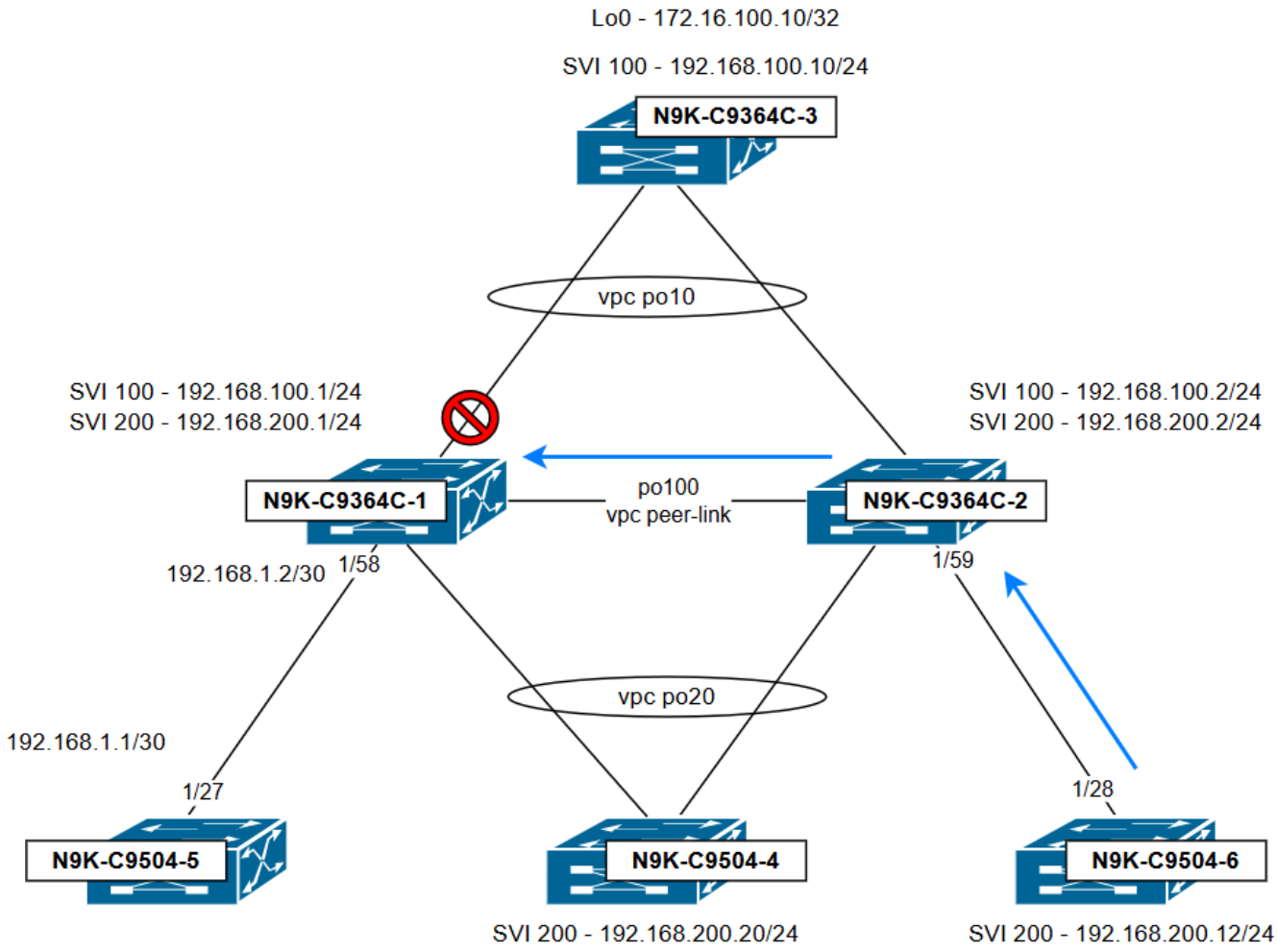
```
N9K-C9364C-1#
```

Mesmo que o tráfego chegue a N9K-C9364C-1 por meio do link de peer do vPC, ele não é encaminhado para o canal de porta 10 do vPC. Isso ocorre porque o bit egress_vsl_drop é definido como 1 para esse vPC, o que acontece quando o mesmo canal de porta do vPC está operacional no switch de peer (nesse caso, 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#
```

Topologia que ilustra o fluxo de tráfego e o ponto em que ele é descartado:



Conclusão:

Mesmo que o tráfego se origine de um host órfão conectado a N9K-C9364C-2, ele é descartado por N9K-C9364C-1 devido à regra de prevenção de loop de vPC: O tráfego recebido pelo link par do vPC não pode ser encaminhado por nenhum canal de porta do vPC que esteja ativo em ambos os switches. Se a porta de ingresso no peer switch é uma porta vPC ou órfã é irrelevante; o que importa é que o tráfego entre por meio do link par do vPC e seja destinado a um vPC que esteja ativo em ambos os switches. Para evitar esse problema, verifique se o status administrativo dos SVIs é consistente em ambos os switches e se suas configurações são simétricas.

Cenário 2: Todos os vPCs e SVIs estão ativos - Pontos do próximo salto para o peer do vPC

Neste cenário, todos os SVIs e canais de porta vPC dentro do domínio vPC estão ativos. No entanto, o N9K-C9504-5, que está conectado ao N9K-C9364C-1 por meio de uma interface de Camada 3, não consegue fazer ping no Loopback 0 no N9K-C9364C-3.

Um traceroute de N9K-C9504-5 indica que o pacote primeiro alcança seu próximo salto imediato em 192.168.1.2 e, em seguida, prossegue para 192.168.100.2, que está associado a N9K-C9364C-2.

<#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#

A verificação do próximo salto de N9K-C9364C-1 (o salto inicial para esse tráfego) mostra que o destino é alcançável por meio de 192.168.100.2, que corresponde a SVI 100 em N9K-C9364C-2.

<#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#

Um ping colorido (um ping com um tamanho de MTU especificado) é usado para rastrear o caminho seguido por este tráfego:

<#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#

Mesmo que o tráfego chegue a N9K-C9364C-2 por meio do link de peer do vPC, ele não é encaminhado para o canal de porta 10 do vPC. Isso ocorre porque o bit egress_vsl_drop é definido como 1 para esse vPC, o que acontece quando o mesmo canal de porta do vPC está operacional no switch de peer (nesse caso, N9K-C9364C-1).

<#root>

N9K-C9364C-2#

show system internal eltm info interface Po10 | i i vsl

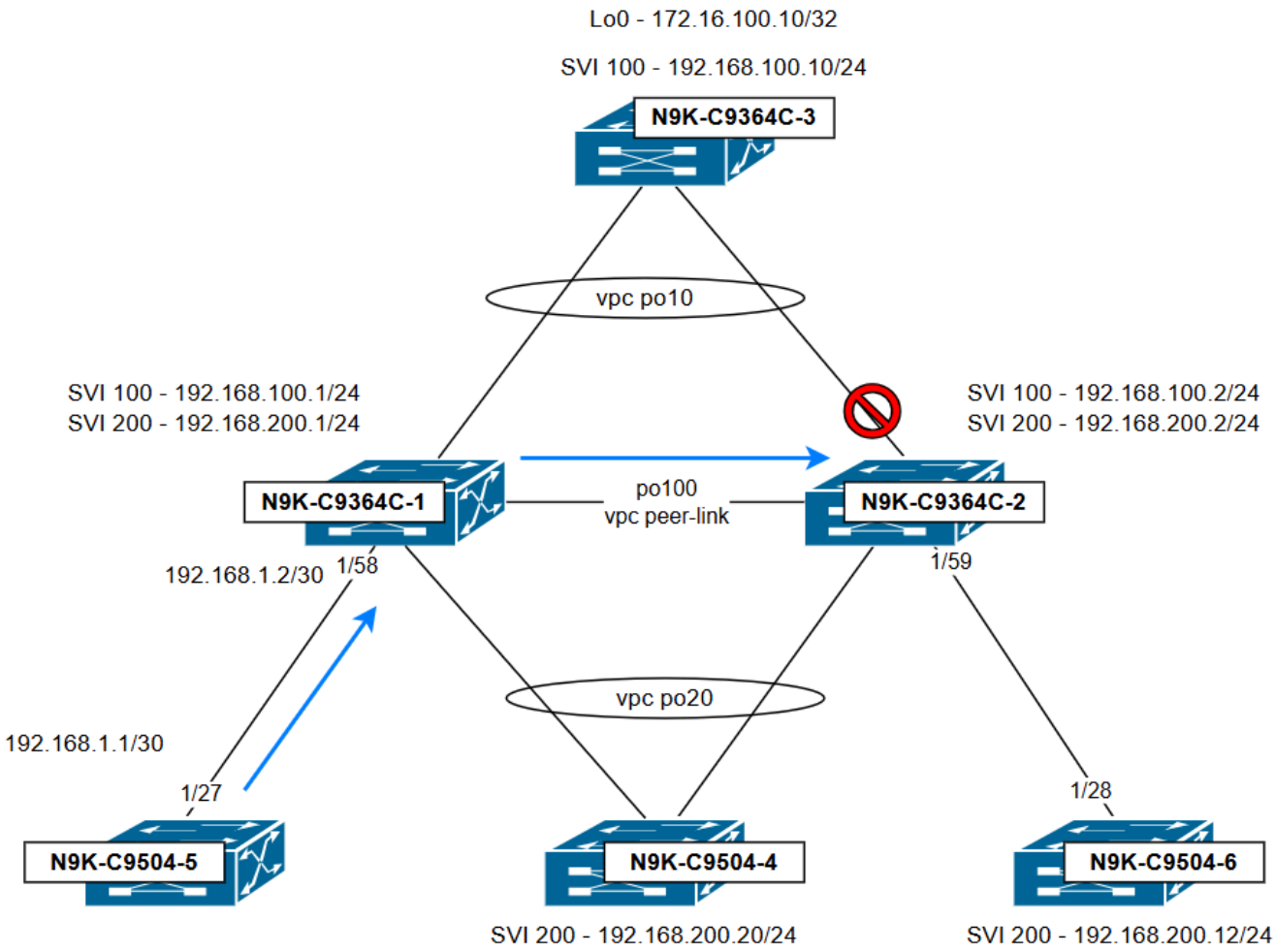
egress_vsl_drop = 1

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#
```

Topologia que ilustra o fluxo de tráfego e o ponto em que ele é descartado:



Conclusão:

O problema é observado porque o N9K-C9364C-1 usa o N9K-C9364C-2 como o próximo salto, enviando tráfego pelo link par do vPC antes de tentar sair pelo vPC 10. O tráfego é descartado devido à regra de prevenção de loop do vPC: O tráfego recebido pelo link par do vPC não pode ser encaminhado por nenhum canal de porta do vPC que esteja ativo em ambos os switches. Para evitar esse problema, verifique se as rotas (dinâmicas ou estáticas) com um próximo salto por meio de um canal de porta do vPC estão configuradas em ambos os switches pares do vPC, de modo que o tráfego não precise cruzar o link par do vPC e sair por um vPC.

Cenário 3: Todos os vPCs e SVIs estão ativos - o recurso de gateway par de VPC está desativado

Neste cenário, todos os SVIs e canais de porta vPC estão ativos no domínio vPC; no entanto, o recurso de gateway par de vPC está desativado. Nesse ponto, N9K-C9504-4 (VLAN 200) não pode fazer ping em N9K-C9364C-3 (VLAN 100).

```
<#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#
```

A verificação de próximo salto do N9K-C9504-4 mostra que o destino é alcançável por meio de 192.168.200.2, que corresponde ao SVI 200 no N9K-C9364C-2 e conectado por canal de porta 20 do vPC.

```
<#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
```

Um ping colorido (um ping com um tamanho de MTU especificado) é usado para rastrear o caminho seguido por esse tráfego. Aqui, os contadores de interface revelam que o N9K-C9364C-1 recebe o tráfego de 192.168.200.20 a 192.168.100.10 pelo canal de porta 20 e o envia para o link par do vPC (canal de porta 100)

<#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#
```

O N9K-C9364C-2 recebe o tráfego pelo link par do vPC (canal de porta 100), mas não o encaminha para o canal de porta 10 do vPC.

<#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#

Mesmo que o tráfego chegue a N9K-C9364C-2 por meio do link de peer do vPC, ele não é encaminhado para o canal de porta 10 do vPC. Isso ocorre porque o bit egress_vsl_drop é definido como 1 para esse vPC, o que acontece quando o mesmo canal de porta do vPC está operacional no switch de peer (nesse caso, N9K-C9364C-1).

Como o gateway par está desativado, o N9K-C9364C-1 pode rotear apenas pacotes endereçados a seu próprio endereço MAC local. Como resultado, os pacotes destinados a a a478.06de.7edb (MAC de N9K-C9364C-2) são encaminhados por N9K-C9364C-1 através do link peer do vPC.

<#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

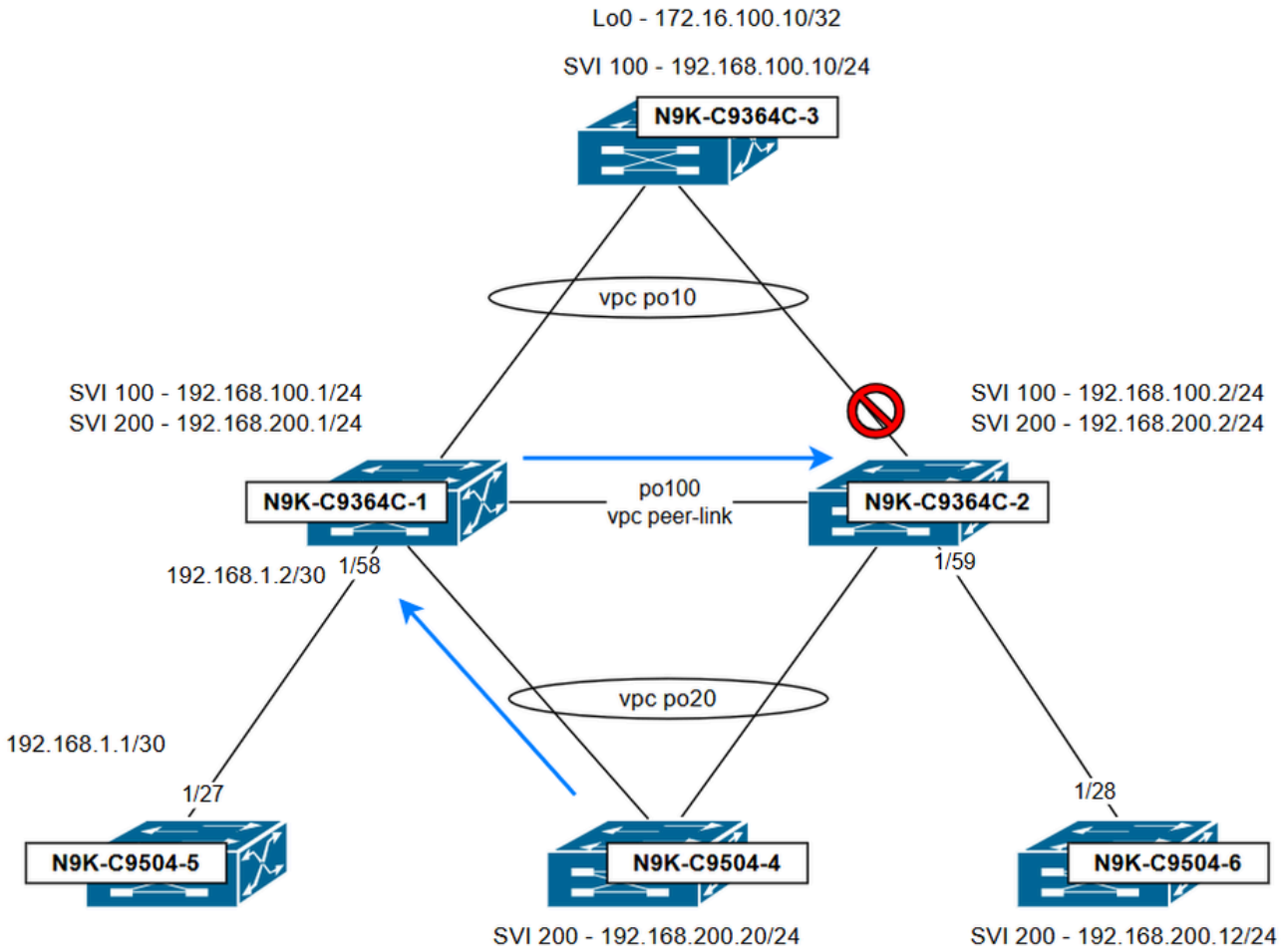
a478.06de.7edb

 static - F F

vPC Peer-Link

```
(R)
* 200

a478.06de.7edb

 static - F F

vPC Peer-Link

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

Topologia que ilustra o fluxo de tráfego e o ponto em que ele é descartado:



Conclusão:

Se peer-gateway estiver habilitado, o tráfego roteado destinado ao endereço MAC do peer do vPC será processado localmente pela programação do MAC do peer como um gateway. Isso evita que o link par do vPC seja usado no caminho de tráfego e evita quedas causadas pela regra de prevenção de loop do vPC. Para evitar esses problemas, certifique-se de que o recurso de gateway par do vPC esteja habilitado no domínio do vPC.

Visão geral da solução

- Mantenha a configuração do SVI consistente nas VLANs do vPC.

As configurações de Interface Virtual Comutada Assimétrica (SVI) entre switches pares vPC podem levar a problemas críticos de encaminhamento de tráfego, incluindo blackholing de tráfego. Uma prática comum, mas sem suporte, que contribui para essa condição é testar o failover entre pares de vPC ao desligar SVIs em um lado. Esse método cria um estado de SVI assimétrico que a arquitetura do Nexus vPC não suporta, resultando em falhas de blackholing e encaminhamento de tráfego. Verifique se a configuração do SVI está sempre consistente em todas as VLANs do vPC para as quais o roteamento é necessário.

- Habilite o gateway par no domínio vPC.

O recurso de gateway de mesmo nível é um aprimoramento importante nas implantações do Cisco Nexus vPC. Quando habilitado no domínio vPC, ele permite que cada switch de peer vPC aceite e processe pacotes destinados ao endereço MAC virtual do peer vPC. Isso significa que qualquer par de vPC pode responder ao tráfego vinculado ao gateway, independentemente de qual switch recebeu originalmente o pacote. Sem a habilitação do gateway de mesmo nível, determinados tipos de tráfego, como pacotes enviados ao endereço MAC do gateway padrão, podem ser descartados se chegarem a um par e precisarem atravessar o link de mesmo nível e sair de uma porta membro do vPC. Verifique se o gateway par do vPC está configurado no domínio do vPC.

Informações Relacionadas

Entender os aprimoramentos do Virtual Port Channel (vPC)

Práticas recomendadas para Virtual Port Channels (vPC) no Nexus

Recurso de gateway de mesmo nível no Nexus 7000

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.