Nexus VPC 루프 방지 이해

목차

소개

이 문서에서는 Nexus 기반 레이어 3 네트워크 설계에서 vPC 루프 회피가 트래픽 포워딩에 영향을 줄 수 있는 시나리오에 대해 설명합니다.

사전 요구 사항

요구 사항

다음 주제에 대한 지식을 보유하고 있으면 유용합니다.

- Nexus 운영 체제 CLI
- vPC 개념

사용되는 구성 요소

이 문서의 정보는 다음 소프트웨어 및 하드웨어 버전을 기반으로 합니다.

- 소프트웨어 10.4(4)

- 하드웨어 N9K-C9364C-GX

이 문서의 정보는 특정 랩 환경의 디바이스를 토대로 작성되었습니다. 이 문서에 사용된 모든 디바이스는 초기화된(기본) 컨피그레이션으로 시작되었습니다. 현재 네트워크가 작동 중인 경우 모든 명령의 잠재적인 영향을 미리 숙지하시기 바랍니다.

배경 정보

오늘날의 데이터 센터 환경에서 Cisco Nexus vPC(Virtual Port Channel) 기술은 이중화 및 로드 밸런싱을 지원하는 데 필수적입니다. 두 개의 개별 Nexus 스위치에 대한 연결이 단일 논리적 포트 채널 기능을 하도록 허용함으로써 vPC는 네트워크 아키텍처를 간소화하고 다운스트림 디바이스의 안정성을 향상시킵니다. 그러나 특정 컨피그레이션 세부사항으로 인해 운영이 복잡해질 수 있습니다.

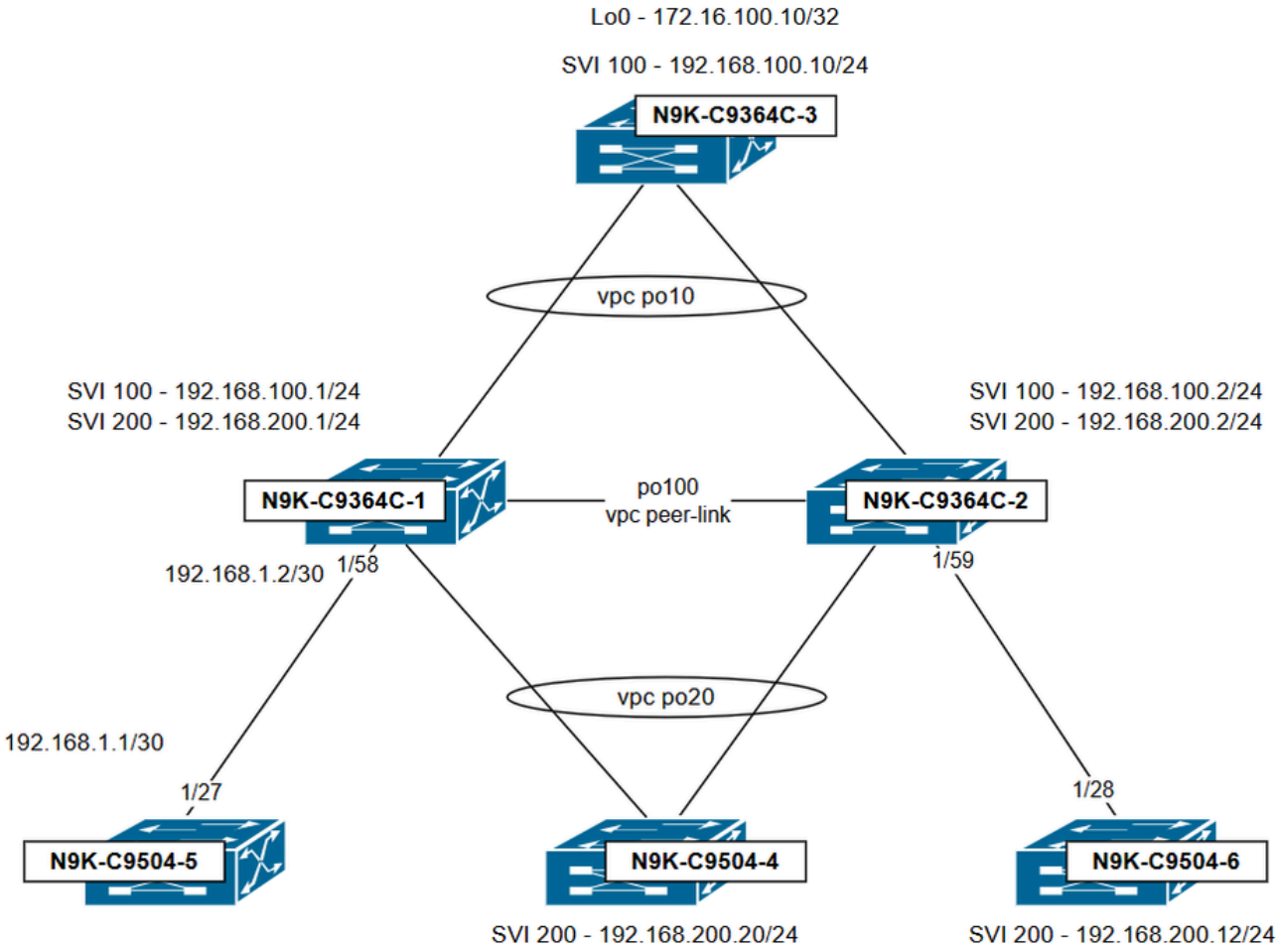
이 문서에서는 vPC 루프 회피가 중요해지는 시나리오를 살펴보고 트래픽 포워딩에 미치는 영향을 살펴봅니다. Nexus 기반 인프라에서 강력하고 효율적인 레이어 3 연결을 설계 및 유지하려는 네트워크 엔지니어에게는 이 메커니즘을 명확하게 이해하는 것이 매우 중요하며, 이를 통해 트래픽 중단을 방지하고 최적의 네트워크 성능을 유지할 수 있습니다.

문제

vPC를 사용하는 Cisco Nexus 환경에서 네트워크 운영자는 vPC 루프 회피 규칙으로 인해 발생하는 예기치 않은 트래픽 포워딩 동작을 관찰할 수 있습니다. 트래픽이 vPC 피어 링크를 통해 한 vPC 피어에서 다른 피어로 이동하면 두 스위치에서 모두 활성화된 vPC 포트 채널을 통해 나갈 수 없습니다. 그 결과, 이 연결 경로에 따라 디바이스는 모든 물리적 링크가 가동 중인 것으로 보이는 경우에도 패킷이 삭제되거나 연결이 끊길 수 있습니다.

이러한 동작을 간과하면 예기치 않은 서비스 중단이 발생하고 네트워크 문제를 진단하기가 더욱 어려워질 수 있으므로, vPC 루프 회피 규칙을 이해하고 이를 설명하는 것은 복원력 있는 네트워크 토폴로지를 설계하고 트러블슈팅하는 데 필수적입니다.

네트워크 다이어그램



이 토폴로지에서는 vPC 도메인이 N9K-C9364C-1 및 N9K-C9364C-2로 구성됩니다. 두 스위치 모두 VLAN 100 및 200을 vPC VLAN으로 구성하고 각 VLAN에 대해 SVI를 설정합니다. vPC 도메인은 이러한 VLAN 간의 VLAN 간 라우팅을 담당합니다. 달리 지정되지 않는 한 vPC 피어 스위치 간에 공유되는 HSRP VIP(가상 IP)는 토폴로지의 다른 스위치에 의해 기본 경로에 대한 다음 홉으로 사용됩니다.

- N9K-C9364C-1 SVI 구성

인터페이스 Vlan100
종료 안 함
no ip redirects
ip 주소 192.168.100.1/24
ipv6 리디렉션 없음
hsrp 100
ip 192.168.100.254

인터페이스 Vlan200
종료 안 함
no ip redirects
ip 주소 192.168.200.1/24
ipv6 리디렉션 없음
hsrp 200

ip 192.168.200.254

- N9K-C9364C-2 SVI 구성

인터페이스 Vlan100
종료 안 함
no ip redirects
ip 주소 192.168.100.2/24
ipv6 리디렉션 없음
hsrp 100
ip 192.168.100.254

인터페이스 Vlan200
no ip redirects
ip 주소 192.168.200.2/24
ipv6 리디렉션 없음
hsrp 200
ip 192.168.200.254

시나리오

시나리오 1: vPC VLAN용 SVI가 관리상 vPC 피어에서 종료됨

a) vPC에서 vPC로 라우팅된 트래픽이 영향을 받음

작업 시나리오에서 N9K-C9504-4(VLAN 200)는 N9K-C9364C-3(VLAN 100)을 성공적으로 ping할 수 있습니다. Traceroute는 연결 경로가 N9K-C9364C-2에 할당된 192.168.200.2를 통과함을 나타냅니다.

<#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
```

현재 트래픽 흐름은 다음과 같이 작동합니다.

- N9K-C9364C-2는 192.168.100.10으로 향하는 192.168.200.20에서 트래픽을 수신하며, 대상 MAC 주소는 vPC 도메인 내에서 공유 HSRP VMAC(Virtual MAC)로 설정됩니다.
- HSRP는 vPC의 데이터 플레인 관점에서 액티브-액티브 모드에서 작동하므로 N9K-C9364C-2는 트래픽을 VLAN 200에서 VLAN 100으로 라우팅한 다음 vPC 10을 통해 전달합니다.

SVI 200이 N9K-C9364C-2에서는 종료되지만 N9K-C9364C-1에서는 활성 상태로 유지되는 시나리오를 고려해 보십시오.

```
<#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 피어 간 SVI의 작동 상태 차이로 인해 vPC 도메인 내에서 Type-2 불일치가 감지됩니다.

<#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#
```

이 단계에서는 192.168.200.20에서 192.168.100.10으로의 트래픽이 더 이상 성공하지 않습니다.

<#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#
```

컬러 ping(지정된 MTU 크기의 ping)은 이 트래픽에서 사용하는 경로를 추적하는 데 사용됩니다.

<#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-2의 인터페이스 카운터에 따르면, 이 트래픽은 포트 채널 20에서 수신되어 포트 채널 100(vPC 피어 링크)으로 전달됩니다.

<#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#

이 동작은 SVI 200이 N9K-C9364C-2에서 종료되어 VLAN 200에 대한 트래픽의 로컬 라우팅을 방

해하기 때문에 발생합니다. 이 시나리오에서 트래픽은 vPC 피어 링크를 통해 N9K-C9364C-1에 브리지되므로 디바이스에서 VLAN 간 라우팅을 수행합니다.

N9K-C9364C-1의 인터페이스 카운터를 보면 패킷이 vPC 피어 링크를 통해 이 디바이스에 도달하는 것이 확인되지만 192.168.100.10에 연결하는 vPC 포트 채널 10에서는 발신 패킷이 관찰되지 않습니다.

```
<#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#
```

트래픽이 vPC 피어 링크를 통해 N9K-C9364C-1에 도착하더라도 vPC 포트 채널 10으로 전달되지 않습니다. 이는 이 vPC에 대해 egress_vsl_drop 비트가 1로 설정되었기 때문입니다. 이는 동일한 vPC 포트 채널이 피어 스위치(이 경우에는 N9K-C9364C-2)에서 작동할 때 발생합니다.

```
<#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:
```

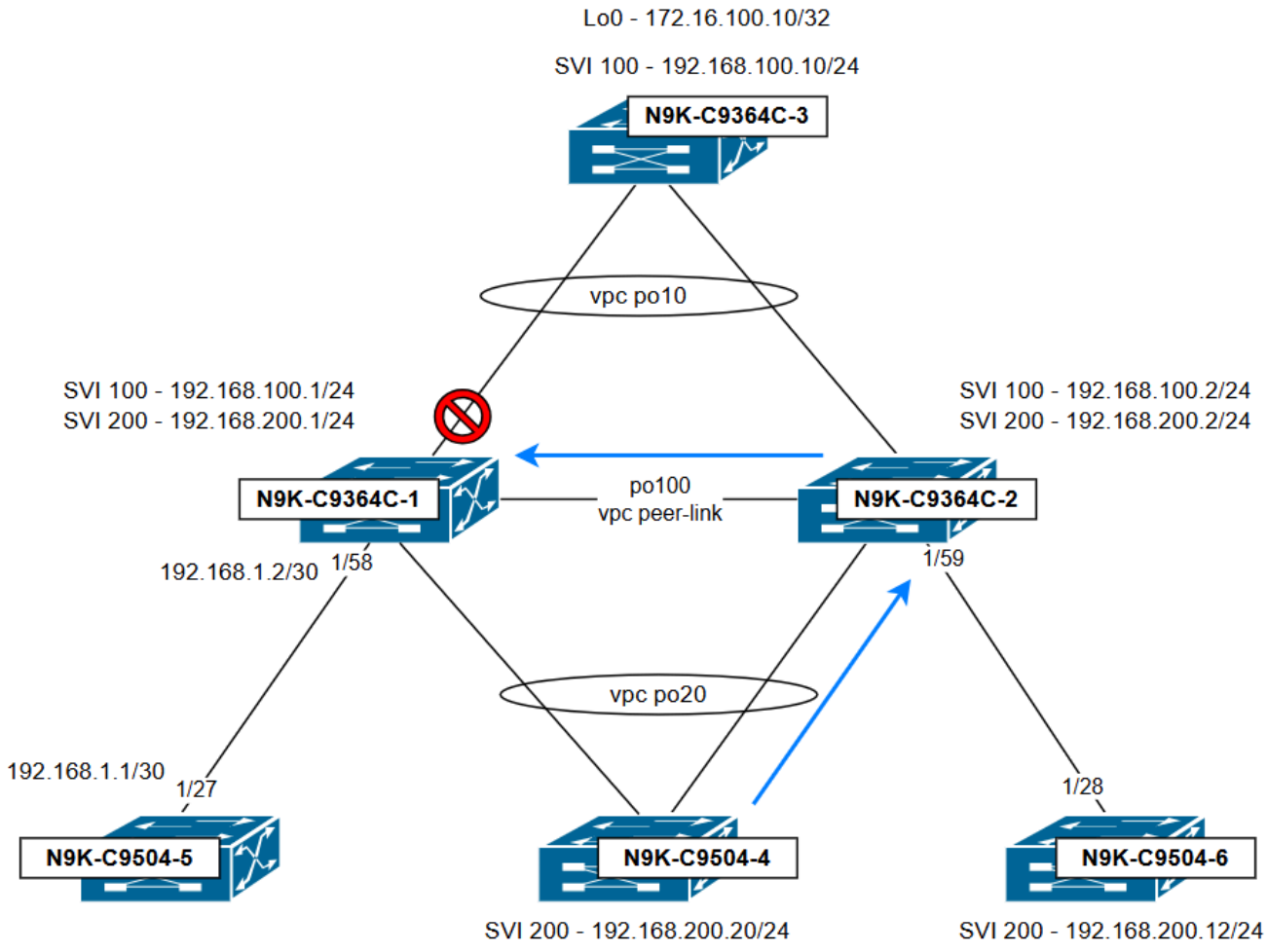
vPC state: Up Old Compat Status: Pass

```
vPC Peer Information:
```

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

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

트래픽 흐름 및 트래픽이 중단된 지점을 보여주는 토폴로지:



결론:

N9K-C9364C-1은 vPC 루프 회피 규칙으로 인해 트래픽을 삭제합니다. vPC 피어 링크를 통해 수신된 트래픽은 두 스위치에서 모두 활성 상태인 vPC 포트 채널 밖으로 전달될 수 없습니다."이 문제를 방지하려면 두 스위치에서 SVI의 관리 상태가 일치하고 컨피그레이션이 대칭인지 확인하십시오.

b) 분리된 호스트에서 vPC 호스트로 라우팅된 트래픽이 영향을 받음

SVI 200이 N9K-C9364C-2에서 종료되었지만 N9K-C9364C-1에서 활성 상태로 유지되는 동일한 시나리오를 고려해 보십시오. N9K-C9504-6(VLAN 200)에서 N9K-C9364C-3(VLAN 100)으로의 ping은 성공하지 못합니다.

<#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#
```

컬러 ping(지정된 MTU 크기의 ping)은 이 트래픽에서 사용하는 경로를 추적하는 데 사용됩니다.

<#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#

트래픽이 vPC 피어 링크를 통해 N9K-C9364C-1에 도착하더라도 vPC 포트 채널 10으로 전달되지 않습니다. 이는 이 vPC에 대해 egress_vsl_drop 비트가 1로 설정되었기 때문입니다. 이는 동일한 vPC 포트 채널이 피어 스위치(이 경우에는 N9K-C9364C-2)에서 작동할 때 발생합니다.

<#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:

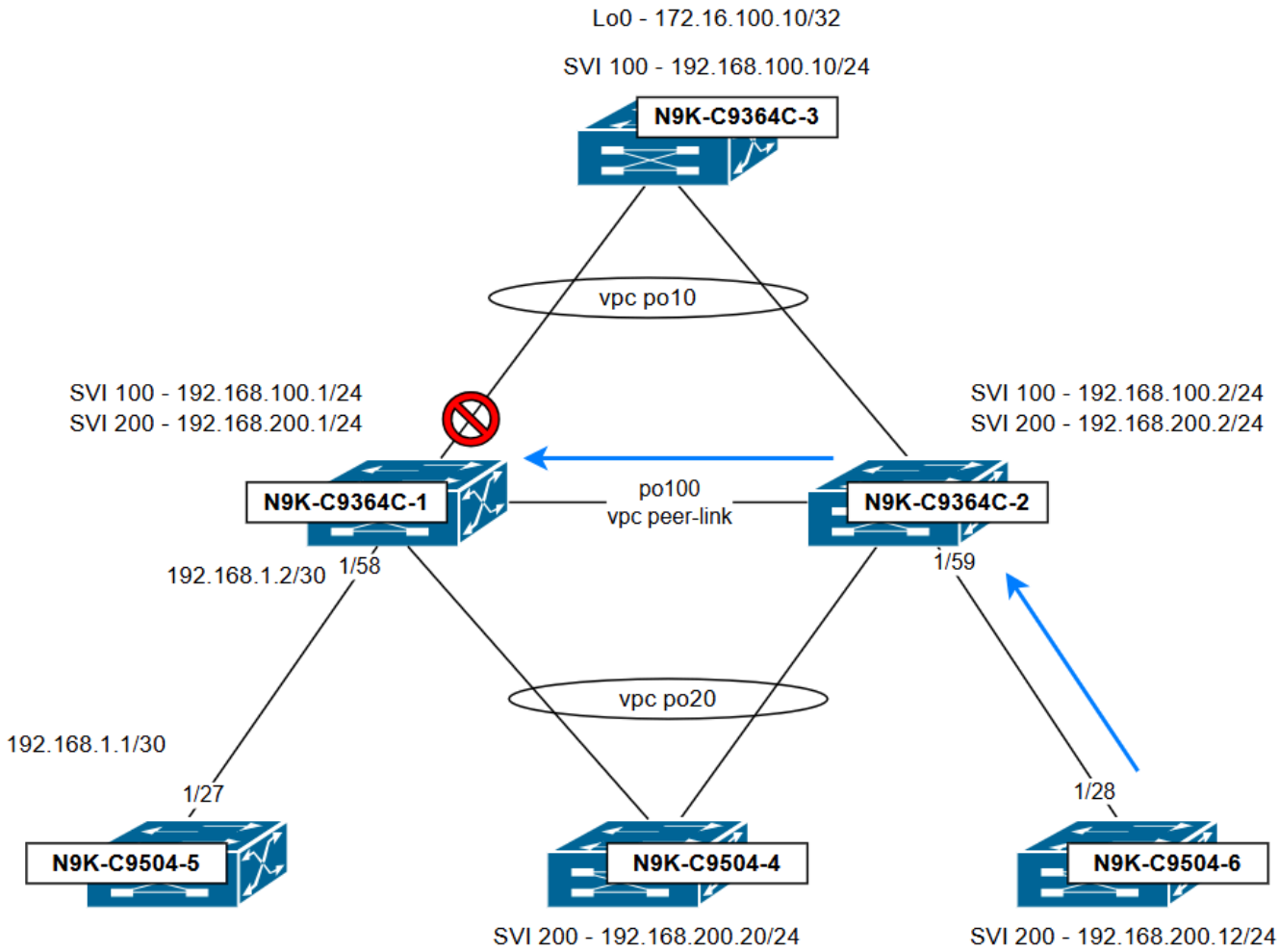
vPC state: Up Old Compat Status: Pass

vPC Peer Information:

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

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

트래픽 흐름 및 트래픽 흐름 중단 지점을 보여주는 토폴로지:



결론:

트래픽은 N9K-C9364C-2에 연결된 고아 호스트에서 시작되지만 vPC 루프 회피 규칙 때문에 N9K-C9364C-1에 의해 삭제됩니다. vPC 피어 링크를 통해 수신된 트래픽은 두 스위치에서 모두 활성 상

태인 vPC 포트 채널 밖으로 전달될 수 없습니다. 피어 스위치의 인그레스 포트가 vPC인지 고아 포트인지 여부는 관련이 없습니다. 중요한 것은 트래픽이 vPC 피어 링크를 통해 들어오고 두 스위치에서 모두 활성 상태인 vPC를 목적지로 한다는 것입니다. 이 문제를 방지하려면 SVI의 관리 상태가 두 스위치에서 일치하고 컨피그레이션이 대칭인지 확인하십시오.

시나리오 2: 모든 vPC 및 SVI가 가동 - vPC 피어를 가리키는 다음 홉

이 시나리오에서는 vPC 도메인 내의 모든 SVI 및 vPC 포트 채널이 가동됩니다. 그러나 레이어 3 인터페이스를 통해 N9K-C9364C-1에 연결된 N9K-C9504-5는 N9K-C9364C-3에서 루프백 0을 ping할 수 없습니다.

N9K-C9504-5의 traceroute는 패킷이 먼저 192.168.1.2에서 바로 다음 홉에 도달한 다음 N9K-C9364C-2와 연결된 192.168.100.2로 진행됨을 나타냅니다.

```
<#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-1의 다음 홉 확인(이 트래픽에 대한 초기 홉)은 대상이 192.168.100.2를 통해 연결할 수 있음을 보여줍니다. 이는 N9K-C9364C-2의 SVI 100에 해당합니다.

```
<#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#
```

컬러 ping(지정된 MTU 크기의 ping)은 이 트래픽에서 사용하는 경로를 추적하는 데 사용됩니다.

<#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 co
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#
```

트래픽은 vPC 피어 링크를 통해 N9K-C9364C-2에 도착하지만 vPC 포트 채널 10으로 전달되지 않습니다. 이는 이 vPC에 대해 egress_vsl_drop 비트가 1로 설정되었기 때문입니다. 이는 동일한 vPC 포트 채널이 피어 스위치(이 경우 N9K-C9364C-1)에서 작동할 때 발생합니다.

<#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:

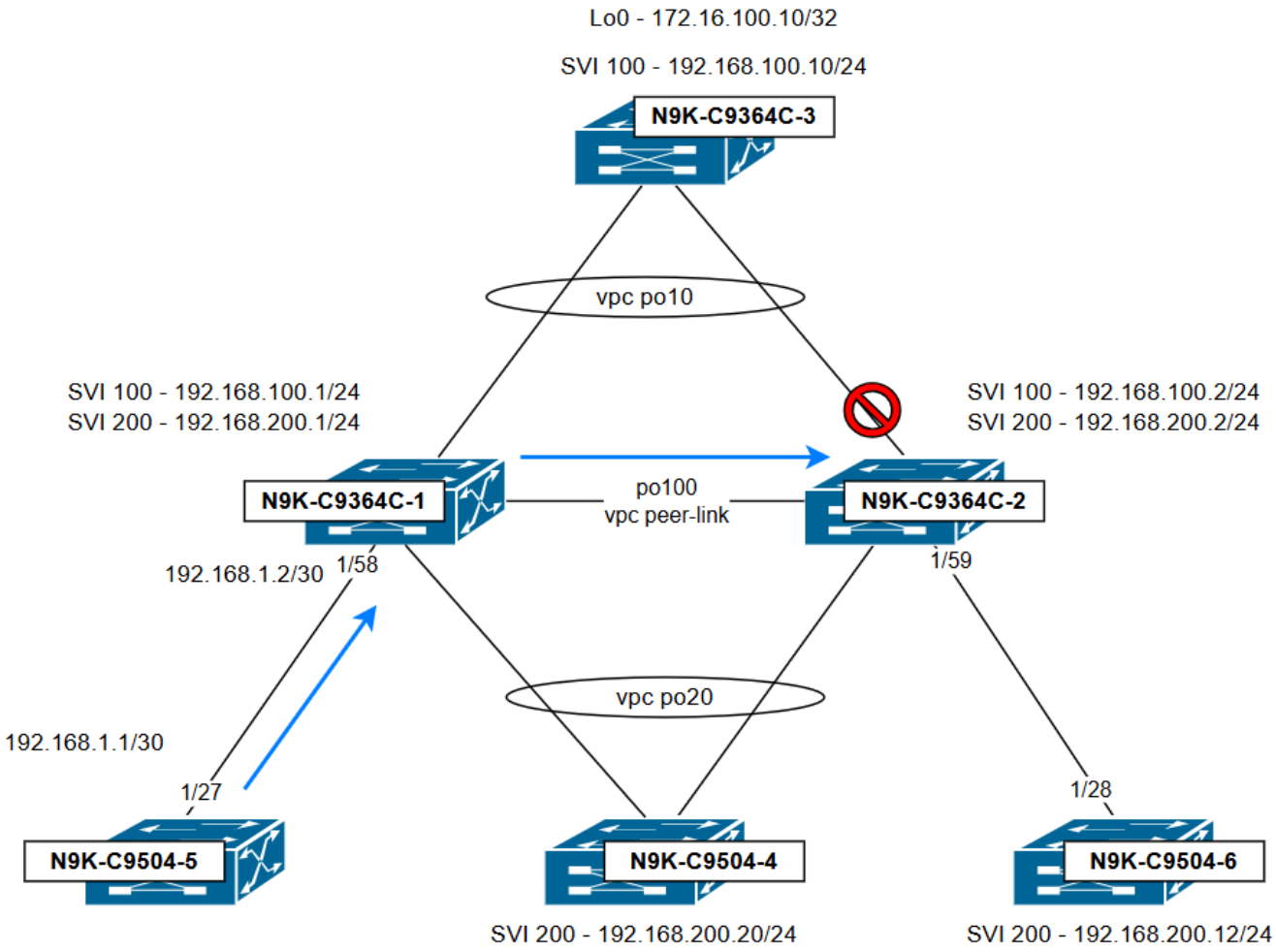
Peer state: Up <<<------------- vPC 10 up on peer

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#

트래픽 흐름 및 트래픽 흐름 중단 지점을 보여주는 토폴로지:



결론:

이 문제는 N9K-C9364C-1이 N9K-C9364C-2를 다음 홉으로 사용하여 vPC 10을 통해 나가려고 시도하기 전에 vPC 피어 링크를 통해 트래픽을 전송하기 때문에 발생합니다. vPC 루프 회피 규칙으로 인해 트래픽이 삭제됩니다. vPC 피어 링크를 통해 받은 트래픽은 두 스위치에서 모두 활성화된 vPC 포트 채널을 통해 전달할 수 없습니다. 이 문제를 방지하려면 트래픽이 vPC 피어 링크를 통해 vPC 피어 링크를 통해 나가고 vPC를 통해 이그레스(egress)될 필요가 없도록 vPC 포트 채널을 통해 다음 홉을 통한 경로(동적 또는 정적)가 두 vPC 피어 스위치에 구성되어 있는지 확인하십시오.

시나리오 3: 모든 vPC 및 SVI가 가동 - VPC 피어 게이트웨이 기능이 꺼짐

이 시나리오에서는 모든 SVI 및 vPC 포트 채널이 vPC 도메인에 있습니다. 그러나 vPC 피어 게이트웨이 기능이 꺼져 있습니다. 이 시점에서 N9K-C9504-4(VLAN 200)는 N9K-C9364C-3(VLAN 100)을 ping할 수 없습니다.

```
<#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-C9504-4의 다음 홉 확인에서는 192.168.200.2를 통해 대상에 연결할 수 있음을 보여 줍니다. 이는 N9K-C9364C-2의 SVI 200에 해당하며 vPC 포트 채널 20을 통해 연결됩니다.

```
<#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
```

컬러 ping(지정된 MTU 크기의 ping)은 이 트래픽에서 사용하는 경로를 추적하는 데 사용됩니다. 여기서 인터페이스 카운터는 N9K-C9364C-1이 포트 채널 20을 통해 192.168.200.20에서

192.168.100.10으로 트래픽을 수신하고 이를 vPC 피어 링크(포트 채널 100)로 전송함을 나타냅니다

<#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-2는 vPC 피어 링크(port-channel100)를 통해 트래픽을 수신하지만 vPC 포트 채널 10으로 전달하지는 않습니다.

<#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)

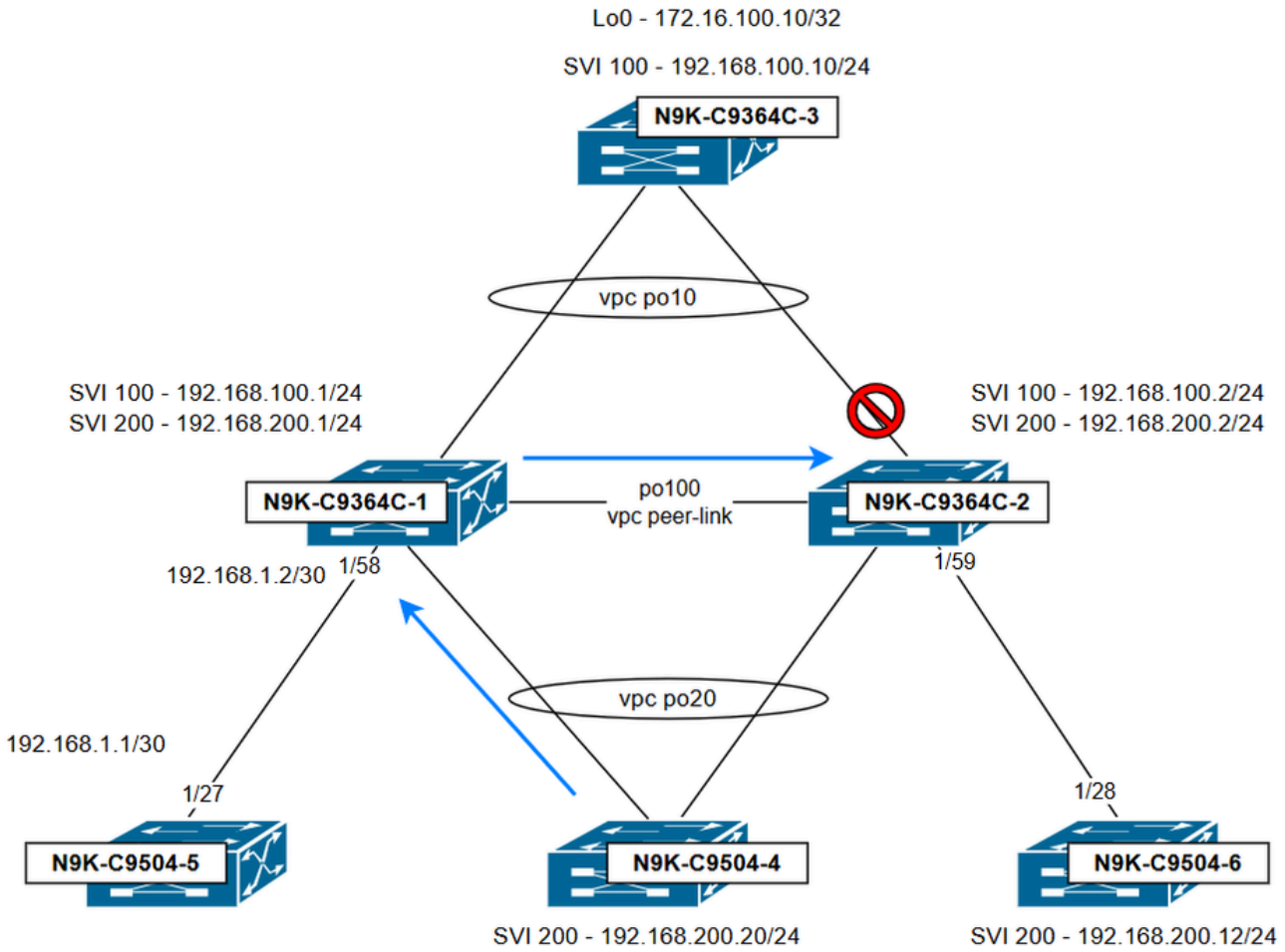
```
60. Tx Packets from 1024 to 1518 bytes: = 0
N9K-C9364C-2#
```

트래픽은 vPC 피어 링크를 통해 N9K-C9364C-2에 도착하지만 vPC 포트 채널 10으로 전달되지 않습니다. 이는 이 vPC에 대해 egress_vsl_drop 비트가 1로 설정되었기 때문입니다. 이는 동일한 vPC 포트 채널이 피어 스위치(이 경우에는 N9K-C9364C-1)에서 작동할 때 발생합니다.

피어 게이트웨이가 비활성화되었으므로 N9K-C9364C-1은 자체 로컬 MAC 주소로 주소가 지정된 패킷만 라우팅할 수 있습니다. 따라서 a478.06de.7edb(N9K-C9364C-2의 MAC)로 향하는 패킷은 N9K-C9364C-1에서 vPC 피어 링크를 통해 전달됩니다.

```
<#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#
```

트래픽 흐름 및 트래픽 흐름 중단 지점을 보여주는 토폴로지:



결론:

피어 게이트웨이가 활성화된 경우, vPC 피어의 MAC 주소로 향하는 라우팅된 트래픽은 피어 MAC을 게이트웨이로 프로그래밍하여 로컬에서 처리됩니다. 이렇게 하면 vPC 피어 링크가 트래픽 경로에서 사용되지 않으며 vPC 루프 회피 규칙으로 인한 삭제를 방지할 수 있습니다. 이러한 문제를 방지하려면 vPC 도메인에서 vPC 피어 게이트웨이 기능이 활성화되어 있는지 확인합니다.

솔루션 개요

- vPC VLAN에서 SVI 컨피그레이션을 일관되게 유지합니다.

vPC 피어 스위치 간 SVI(Asymmetric Switched Virtual Interface) 컨피그레이션은 트래픽 블랙홀링을 비롯한 중요한 트래픽 포워딩 문제로 이어질 수 있습니다. 이 상태에 기여하는 일반적이지만 지원되지 않는 방법은 한 쪽에서 SVI를 종료하여 vPC 피어 간의 장애 조치를 테스트하는 것입니다. 이 방법을 사용하면 Nexus vPC 아키텍처가 지원하지 않는 비대칭 SVI 상태가 생성되므로 트래픽 블랙홀링 및 포워딩 오류가 발생합니다. 라우팅이 필요한 모든 vPC VLAN에서 SVI 컨피그레이션이 항상 일치하는지 확인합니다.

- vPC 도메인에서 피어 게이트웨이를 활성화합니다.

피어 게이트웨이 기능은 Cisco Nexus vPC 구축에서 매우 향상되었습니다. vPC 도메인에서 활성화

되면 각 vPC 피어 스위치가 vPC 피어의 가상 MAC 주소로 향하는 패킷을 수락하고 처리할 수 있습니다. 즉, vPC 피어 중 어느 스위치가 패킷을 처음 수신했는지에 관계없이 게이트웨이 바운드 트래픽에 응답할 수 있습니다. 피어 게이트웨이를 활성화하지 않으면 특정 유형의 트래픽(예: 기본 게이트웨이 MAC 주소로 전송된 패킷)이 한 피어에 도착하고 그렇지 않으면 피어 링크를 통과하여 vPC 멤버 포트를 종료해야 하는 경우 이러한 트래픽은 삭제될 수 있습니다. vPC 피어 게이트웨이가 vPC 도메인에 구성되어 있는지 확인합니다.

관련 정보

vPC(Virtual Port Channel) 개선 사항 이해

Nexus의 vPC(Virtual Port Channel) 모범 사례

Nexus 7000의 피어 게이트웨이 기능

이 번역에 관하여

Cisco는 전 세계 사용자에게 다양한 언어로 지원 콘텐츠를 제공하기 위해 기계 번역 기술과 수작업 번역을 병행하여 이 문서를 번역했습니다. 아무리 품질이 높은 기계 번역이라도 전문 번역가의 번역 결과물만큼 정확하지는 않습니다. Cisco Systems, Inc.는 이 같은 번역에 대해 어떠한 책임도 지지 않으며 항상 원본 영문 문서(링크 제공됨)를 참조할 것을 권장합니다.