Informazioni sui miglioramenti di Virtual Port Channel (vPC)

Sommario

Introduzione

Questo documento descrive i miglioramenti più comuni apportati al canale della porta virtuale (vPC) sugli switch Cisco Nexus in un dominio vPC.

Prerequisiti

Requisiti

Cisco consiglia di avere conoscenze base sugli scenari d'uso, la configurazione e l'implementazione del Virtual Port Channel (vPC). Per ulteriori informazioni sul vPC, consultare uno di questi documenti:

- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 10.1(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 9.3(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 9.2(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 7.x
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 7000, versione 8.x
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 7000, versione 7.x
- Guida alla progettazione e configurazione: best practice per i canali delle porte virtuali (vPC) sugli switch Cisco Nexus serie 7000

Componenti usati

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Premesse

Sin dalle prime versioni di Cisco NX-OS sugli switch per data center Cisco Nexus, la funzionalità Virtual Port Channel (vPC) è stata modificata numerose volte per migliorare l'affidabilità dei dispositivi connessi al vPC e per ottimizzare la modalità di inoltro di entrambi gli switch peer vPC. Comprendere lo scopo di ciascun miglioramento, gli effetti sul funzionamento del sistema e gli scenari di errore che può risolvere aiuta a capire perché e quando configurarli in un dominio vPC per soddisfare al meglio le esigenze e i requisiti aziendali.

Dispositivi interessati

La procedura descritta in questo documento è valida per tutti gli switch per data center Cisco Nexus con funzionalità vPC.

vPC Peer Switch

In questa sezione viene descritta la funzionalità migliorata vPC Peer Switch abilitata con il comando di configurazione **peer-switch** nel dominio vPC.

Panoramica

In molti ambienti, gli switch Nexus di un dominio vPC sono una coppia di switch di aggregazione o core che delimitano il confine tra i domini Ethernet di Layer 2 e i domini di routing di Layer 3. Entrambi gli switch sono configurati con più VLAN e sono responsabili del routing del traffico est-ovest tra VLAN e del traffico nord-sud. In questi ambienti, gli switch Nexus svolgono una funzione di root bridge per il protocollo STP (Spanning Tree Protocol).

In genere, un peer vPC viene configurato come bridge radice dello Spanning Tree impostando la relativa priorità Spanning Tree su un valore basso, ad esempio 0. L'altro peer vPC è configurato con una priorità Spanning Tree leggermente superiore, ad esempio 4096, che consente di assumere il ruolo di bridge radice all'interno dello Spanning Tree se il peer vPC che agisce come bridge radice non funziona. In questa configurazione, il peer vPC che agisce come root bridge genera unità Spanning Tree Bridge Protocol Data Unit (BPDU) con un ID bridge contenente l'indirizzo MAC del sistema.

Tuttavia, se il peer vPC che funge da bridge radice si guasta e causa l'assunzione del controllo da parte dell'altro peer vPC come bridge radice dello Spanning Tree, l'altro peer vPC genera BPDU Spanning Tree con un ID bridge contenente il proprio indirizzo MAC di sistema, che è diverso dall'indirizzo MAC di sistema del bridge radice originale. A seconda di come sono connessi i bridge a valle, l'impatto di questa modifica varia ed è descritto nelle sottosezioni seguenti.

Bridge non vPC con connessione ridondante

I bridge non connessi a vPC connessi a entrambi i peer vPC con collegamenti ridondanti (in modo che un collegamento sia in stato di blocco dalla prospettiva di uno Spanning Tree Protocol) che rilevano la modifica nella BPDU (e, di conseguenza, la modifica nel bridge radice) osservano una modifica nella porta radice. Altre interfacce di inoltro designato passano immediatamente a uno stato di blocco, quindi attraversano la macchina a stati finiti Spanning Tree Protocol (blocco, apprendimento e inoltro) con pause comprese tra l'equivalente del timer di ritardo di inoltro del protocollo Spanning Tree configurato (15 secondi per impostazione predefinita).

La modifica della porta root e il successivo attraversamento della macchina a stati finiti dello Spanning Tree Protocol possono causare interruzioni importanti nella rete. La funzionalità vPC Peer Switch è stata migliorata principalmente per evitare interruzioni della rete causate dalla disconnessione di uno dei peer vPC. Con il miglioramento di vPC Peer Switch, il bridge non connesso a vPC dispone ancora di un singolo collegamento ridondante in stato di blocco, ma passa immediatamente a uno stato di inoltro se la porta radice esistente diventa inattiva a causa di un errore del collegamento. Lo stesso processo si verifica quando il peer vPC offline torna online: l'interfaccia con il costo più basso per il bridge radice acquisisce il ruolo di porta radice e il collegamento ridondante passa immediatamente a uno stato di blocco. L'unico impatto del data

plane che viene osservato è l'inevitabile perdita di pacchetti in-flight che attraversavano il peer vPC mentre passava offline.

Bridge connessi al vPC

I bridge connessi a vPC nel dominio Spanning Tree rilevano la modifica nella BPDU (e, di conseguenza, la modifica nel bridge radice) e scaricano dinamicamente gli indirizzi MAC appresi dalle relative tabelle degli indirizzi MAC locali. Questo comportamento è inefficiente e non è necessario nelle topologie con dispositivi connessi a vPC che non dipendono dallo Spanning Tree Protocol per una topologia senza loop. I vPC sono visti come un'unica interfaccia logica dal punto di vista dello Spanning Tree Protocol proprio come i normali canali di porta, quindi la perdita di un peer vPC è simile alla perdita di un singolo collegamento all'interno di un membro del canale di porta. In entrambi gli scenari, il protocollo STP non cambia, quindi non è necessario eliminare dai bridge gli indirizzi MAC acquisiti dinamicamente nel dominio STP, perché lo scopo di questa eliminazione sarebbe permettere il comportamento flood-and-learn di Ethernet e la riacquisizione degli indirizzi MAC sulle nuove interfacce di inoltro.

Inoltre, l'eliminazione degli indirizzi MAC acquisiti dinamicamente potrebbe essere potenzialmente dannosa. Supponiamo di avere due host con traffico UDP unidirezionale, ad esempio un client TFTP che invia i dati a un server TFTP. In questo flusso, i dati vengono trasferiti principalmente dal client TFTP al server TFTP; raramente il server TFTP restituisce il pacchetto al client TFTP. Di conseguenza, dopo uno scaricamento degli indirizzi MAC appresi in modo dinamico nel dominio Spanning Tree, l'indirizzo MAC del server TFTP non viene appreso per qualche tempo. Ciò significa che i dati del client TFTP inviati al server TFTP vengono trasmessi su tutta la VLAN, in quanto il traffico è traffico unicast sconosciuto. Ciò può causare flussi di dati di grandi dimensioni che si spostano verso destinazioni indesiderate all'interno della rete e possono causare problemi di prestazioni se attraversano sezioni della rete con oversubscription.

Scopo della funzionalità vPC Peer Switch è proprio impedire questo comportamento inefficiente e superfluo in caso venga ricaricato o disattivato il peer vPC che svolge il ruolo di root bridge dello Spanning Tree per una o più VLAN.

Per attivare la funzionalità migliorata vPC Peer Switch, entrambi i peer vPC devono avere la stessa configurazione STP, compresa la stessa priorità su tutte le VLAN del vPC, e devono essere root bridge di almeno una VLAN. Una volta soddisfatti questi prerequisiti, usare il comando di configurazione **peer-switch** nel dominio vPC per attivare la funzionalità migliorata vPC Peer Switch.

> **Nota**: si sconsiglia di abilitare il miglioramento di vPC Peer Switch in un dominio vPC in cui nessuno degli switch peer vPC è lo Spanning Tree Protocol Root Bridge per una o più VLAN vPC. Attivare la funzionalità migliorata vPC Peer Switch solo se uno switch peer vPC, o entrambi, sono root bridge STP su una o più VLAN del vPC.

Una volta abilitati i miglioramenti allo switch peer vPC, entrambi i peer vPC iniziano a generare BPDU Spanning Tree identiche con un ID bridge contenente l'indirizzo MAC del sistema vPC condiviso da entrambi i peer vPC. Se un peer vPC viene ricaricato, la BPDU dello Spanning Tree originata dal peer vPC rimanente non cambia, quindi gli altri bridge nel dominio dello Spanning Tree non vedono alcuna modifica nel bridge radice e non reagiscono in modo subottimale alla modifica nella rete.

Avvertenze

Prima di configurare la funzionalità migliorata vPC Peer Switch nell'ambiente di produzione, tenere conto delle considerazioni seguenti.

I due peer vPC devono avere gli stessi valori di priorità nello Spanning Tree

Prima di attivare la funzionalità migliorata vPC Peer Switch, è necessario modificare la configurazione delle priorità Spanning Tree in tutte le VLAN del vPC in modo che siano identiche per entrambi i peer vPC.

Considerare la configurazione qui, in cui N9K-1 è configurato come il bridge radice dello Spanning Tree per le VLAN 1, 10 e 20 con priorità 0. N9K-2 è il ponte principale secondario dello Spanning Tree per le VLAN 1, 10 e 20 con priorità 4096.

```
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
```

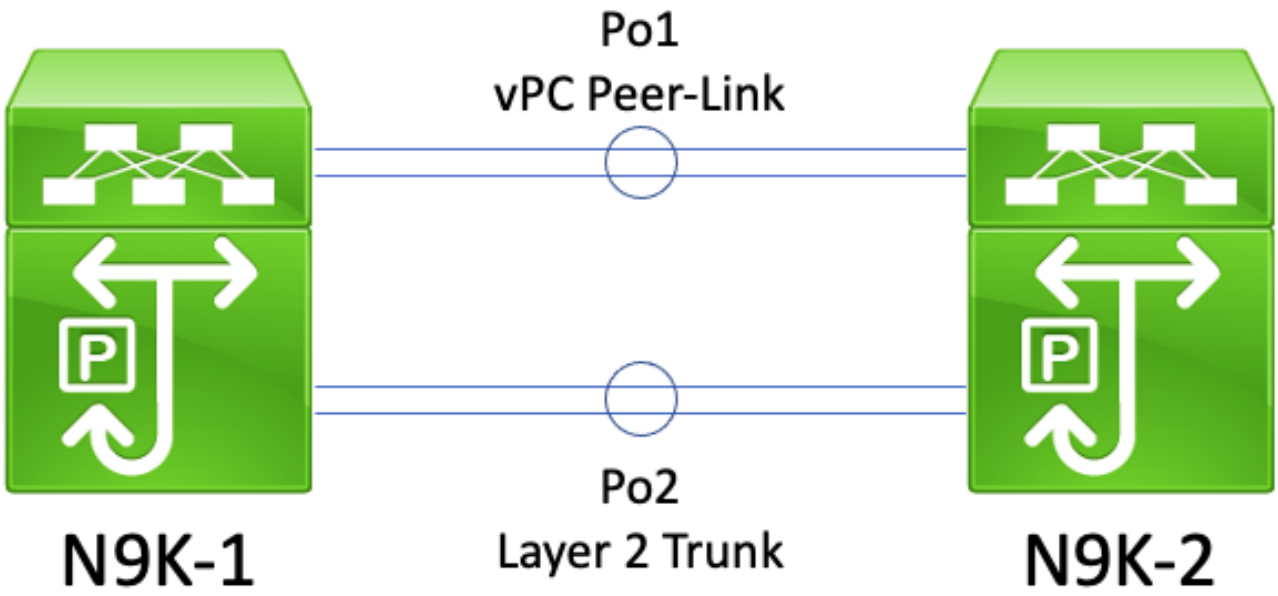
Prima di abilitare il miglioramento di vPC Peer Switch, è necessario modificare la configurazione della priorità dello Spanning Tree per le VLAN 1, 10 e 20 sul router N9K-2 in modo da farla corrispondere alla configurazione della priorità dello Spanning Tree per le stesse VLAN sul router N9K-1. Di seguito è riportato un esempio di questa modifica.

```
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
```

Conseguenze della funzionalità migliorata vPC Peer Switch sulle VLAN non vPC

Supponiamo di avere questa topologia:



In questa topologia, due peer vPC (N9K-1 e N9K-2) dispongono di due trunk di layer 2 tra loro: Po1 e Po2. Po1 è il vPC Peer-Link che trasporta le VLAN vPC, mentre Po2 è un trunk di layer 2 che trasporta tutte le VLAN non vPC. Se i valori di priorità dello Spanning Tree per le VLAN non vPC trasportate tramite Po2 sono identici sui router N9K-1 e N9K-2, ciascun peer vPC genera i frame Spanning Tree BPDU provenienti dall'indirizzo MAC del sistema vPC, identico su entrambi gli switch. Di conseguenza, N9K-1 sembra ricevere la propria BPDU Spanning Tree sulla Po2 per ciascuna VLAN non vPC, anche se N9K-2 è lo switch da cui proviene la BPDU Spanning Tree. Da una prospettiva dello Spanning Tree, N9K-1 blocca il Po2 su tutte le VLAN non vPC.

Si tratta di un comportamento normale. Per evitare questo comportamento o aggirare il problema, entrambi i peer vPC devono essere configurati con priorità Spanning Tree diverse su tutte le VLAN non vPC. In questo modo, un peer vPC diventa il bridge radice per la VLAN non vPC e passa il trunk di layer 2 tra i peer vPC a uno stato di inoltro designato. Analogamente, il peer vPC remoto esegue la transizione del trunk di layer 2 tra i peer vPC a uno stato radice designato. In questo modo, il traffico nelle VLAN non vPC può passare attraverso entrambi i peer vPC attraverso il trunk di layer 2.

Configurazione

Questo è un esempio di come configurare la funzionalità vPC Peer Switch.

Nell'esempio, N9K-1 è configurato come bridge radice dello Spanning Tree per le VLAN 1, 10 e 20 con priorità 0. N9K-2 è il ponte principale secondario dello Spanning Tree per le VLAN 1, 10 e 20 con priorità 4096.

```
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
```

Innanzitutto, la configurazione delle priorità dello Spanning Tree di N9K-2 deve essere modificata per essere identica a quella di N9K-1. Questo è un requisito obbligatorio affinché la funzionalità vPC Peer Switch funzioni come previsto. Se l'indirizzo MAC di sistema di N9K-2 è inferiore all'indirizzo MAC di sistema di N9K-1, N9K-2 usurpa il ruolo di bridge radice per il dominio Spanning Tree, il che fa sì che altri bridge nel dominio Spanning Tree scarichino le loro tabelle di indirizzi MAC locali per tutte le VLAN interessate. Ecco un esempio che spiega questa situazione.

```
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
```

Successivamente, possiamo abilitare la funzionalità vPC Peer Switch tramite il comando di configurazione **peer-switch** nel dominio vPC. In questo modo viene modificato l'ID bridge all'interno degli Spanning Tree BPDU originati da entrambi i peer vPC, il che fa sì che altri bridge nel dominio Spanning Tree scarichino le tabelle degli indirizzi MAC locali per tutte le VLAN interessate.

```
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#
```

Per verificare che la funzionalità vPC Peer Switch funzioni come previsto, è possibile usare il comando **show spanning-tree summary** per confermare che entrambi i peer vPC reclamino il ruolo di root bridge. L'output deve confermare anche che la funzionalità vPC Peer Switch è abilitata e operativa.

```
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
```

Utilizzare il comando **show spanning-tree vlan {x}** per visualizzare altre informazioni dettagliate su una VLAN specifica. Tutte le interfacce dello switch con ruolo vPC primario o operativo primario sono in uno stato di inoltro designato. Lo switch con il ruolo vPC secondario o operativo secondario ha tutte le sue interfacce in uno stato di inoltro designato, ad eccezione del vPC Peer-Link, che si trova nello stato di inoltro radice. Tenere presente che l'indirizzo MAC del sistema vPC visualizzato nell'output del comando **show vpc role** è identico per l'ID root bridge e l'ID bridge di ciascun peer vPC.

```
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
```

Infine, possiamo usare lo [strumento di acquisizione pacchetti del piano di controllo Ethanalyzer](#) su uno dei peer vPC per verificare che entrambi i peer vPC stiano generando le Spanning Tree BPDU con un ID bridge e un ID root bridge contenenti l'indirizzo MAC del sistema vPC condiviso da entrambi i peer vPC.

```
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  Cost = 0  Port = 0x9000

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  Cost = 0  Port = 0x9000
```

Conseguenze

L'impatto dell'abilitazione del miglioramento dello switch peer vPC varia a seconda che altri bridge nel dominio Spanning Tree siano connessi a entrambi i peer vPC tramite un vPC o che siano connessi in modo ridondante a entrambi i peer vPC senza un vPC.

Bridge non vPC con connessione ridondante

Se un bridge non vPC con connessione ridondante a entrambi i peer vPC, di cui una in stato di blocco dal punto di vista dell'STP, rileva un cambiamento nel root bridge annunciato dalle BPDU, la porta root del bridge tra le due interfacce ridondanti potrebbe cambiare. Di conseguenza, le altre interfacce di inoltro designato passano immediatamente allo stato di blocco, quindi attraversano la macchina a stati finiti del protocollo Spanning Tree (blocco, acquisizione e inoltro) a intervalli definiti dal timer di ritardo di inoltro configurato nell'STP, ossia per impostazione predefinita ogni 15 secondi. La modifica della porta root e il successivo attraversamento della macchina a stati finiti dello Spanning Tree Protocol possono causare interruzioni importanti nella rete.

Èopportuno ricordare che questo impatto si verifica ogni volta che il peer vPC che è attualmente il bridge radice per il dominio Spanning Tree non è in linea, ad esempio in caso di interruzione dell'alimentazione, guasto hardware o ricaricamento. Questo comportamento non è esclusivo della funzionalità migliorata vPC Peer Switch, ma è semplicemente causato da tale funzionalità abilitata quando un peer vPC diventa inattivo dal punto di vista del protocollo STP.

Bridge connessi al vPC

Se un bridge connesso con vPC rileva una modifica nel bridge radice Spanning Tree annunciato nelle BPDU Spanning Tree, il bridge scarica dinamicamente gli indirizzi MAC appresi dalla relativa tabella degli indirizzi MAC. Durante la configurazione della funzione vPC Peer Switch, è possibile osservare questo comportamento nei due scenari seguenti:

1. Quando entrambi i peer vPC hanno la stessa priorità configurata nello Spanning Tree, può verificarsi un cambio di root bridge se l'indirizzo MAC di sistema del peer vPC che precedentemente aveva assunto quel ruolo risulta più alto dell'indirizzo dell'altro peer vPC. Un esempio di questo scenario è mostrato nella sottosezione Configurazione di vPC Peer Switch in questo documento.
2. Quando la funzionalità vPC Peer Switch viene abilitata tramite il comando di configurazione del dominio vPC **peer-switch**, entrambi i peer vPC iniziano a funzionare come bridge radice del dominio Spanning Tree. Entrambi i peer vPC iniziano a generare BPDU Spanning Tree identici che si affermano come bridge radice del dominio Spanning Tree.

Nella maggior parte degli scenari e delle topologie, non viene osservato alcun impatto del piano dati come risultato di questi due scenari. Tuttavia, per un breve periodo di tempo, il traffico del data plane viene inondato all'interno di una VLAN a causa di un flooding unicast sconosciuto, in quanto l'indirizzo MAC di destinazione dei frame non viene imparato su alcuna porta switchport come risultato diretto dello scaricamento di indirizzi MAC appresi in modo dinamico. In alcune topologie, ciò può causare brevi periodi di calo di prestazioni o perdita di pacchetti se il traffico del

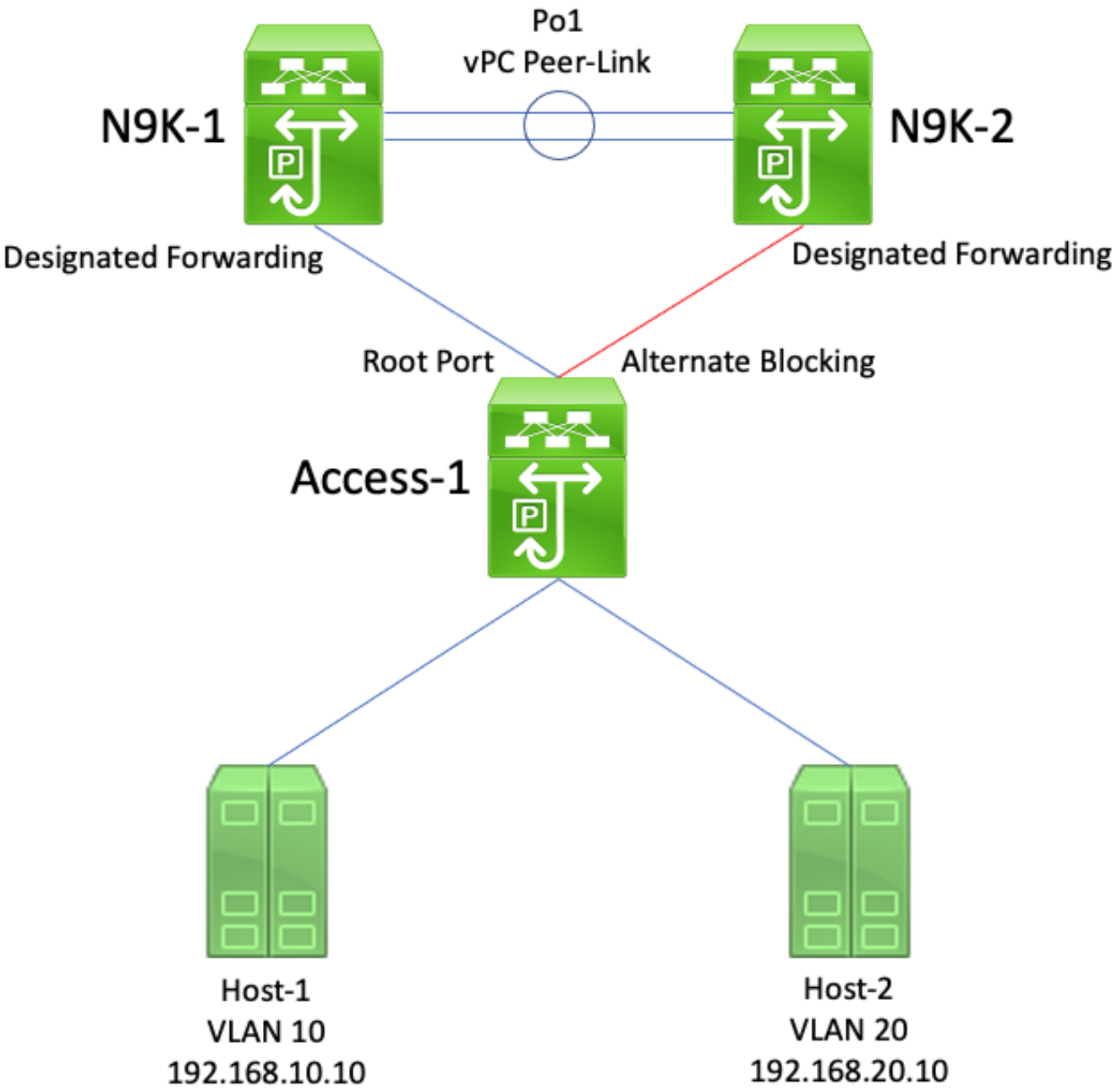
piano dati viene inoltrato a tutti i dispositivi di rete con oversubscription presenti nella VLAN. Ciò può anche causare problemi con i flussi di traffico unidirezionale ad alta intensità di larghezza di banda o con gli host inattivi (host che ricevono principalmente i pacchetti e inviano raramente i pacchetti), in quanto il traffico viene inondato all'interno della VLAN per un periodo di tempo esteso, anziché essere commutato direttamente sull'host di destinazione come al solito.

Èopportuno ricordare che questo impatto è correlato allo scaricamento degli indirizzi MAC appresi in modo dinamico dalla tabella degli indirizzi MAC dei bridge all'interno della VLAN interessata. Questo comportamento non è causato solo dalla funzionalità migliorata vPC Peer Switch o dal cambiamento del root bridge, ma può dipendere anche da una notifica di modifica alla topologia, o TCN (Topology Change Notification), generata da una porta non edge che diventa attiva nella VLAN.

Esempi di scenari di errore

I bridge non vPC con connessione ridondante riavviano la macchina a stati finiti

Supponiamo di avere questa topologia:



In questa topologia, N9K-1 e N9K-2 sono i peer vPC di un dominio vPC. N9K-1 è configurato con priorità Spanning Tree pari a 0 in tutte le VLAN ed è quindi il root bridge di tutte le VLAN. N9K-2 è configurato con priorità Spanning Tree pari a 4096 in tutte le VLAN ed è quindi il root bridge secondario di tutte le VLAN. Access-1 è uno switch connesso in modo ridondante a entrambi i peer N9K-1 e N9K-2 tramite porte di Layer 2. Queste switchport non sono associate a un port-channel, quindi il protocollo STP cambia lo stato del collegamento a N9K-1 in Designated Root e lo stato del collegamento a N9K-2 in Alternate Blocking (Blocco alternativo).

Supponiamo che N9K-1 diventi inattivo a causa di un guasto hardware, un'interruzione dell'alimentazione o il ricaricamento dello switch. N9K-2 si afferma come bridge radice per tutte le VLAN annunciando le BPDU dello Spanning Tree che usano il suo indirizzo MAC di sistema come ID bridge. In Access-1 viene rilevata una modifica nell'ID del bridge radice. Inoltre, la porta radice designata passa a uno stato di inattività/inattività, il che significa che la nuova porta radice designata è il collegamento che era in uno stato di blocco alternativo rivolto verso N9K-2.

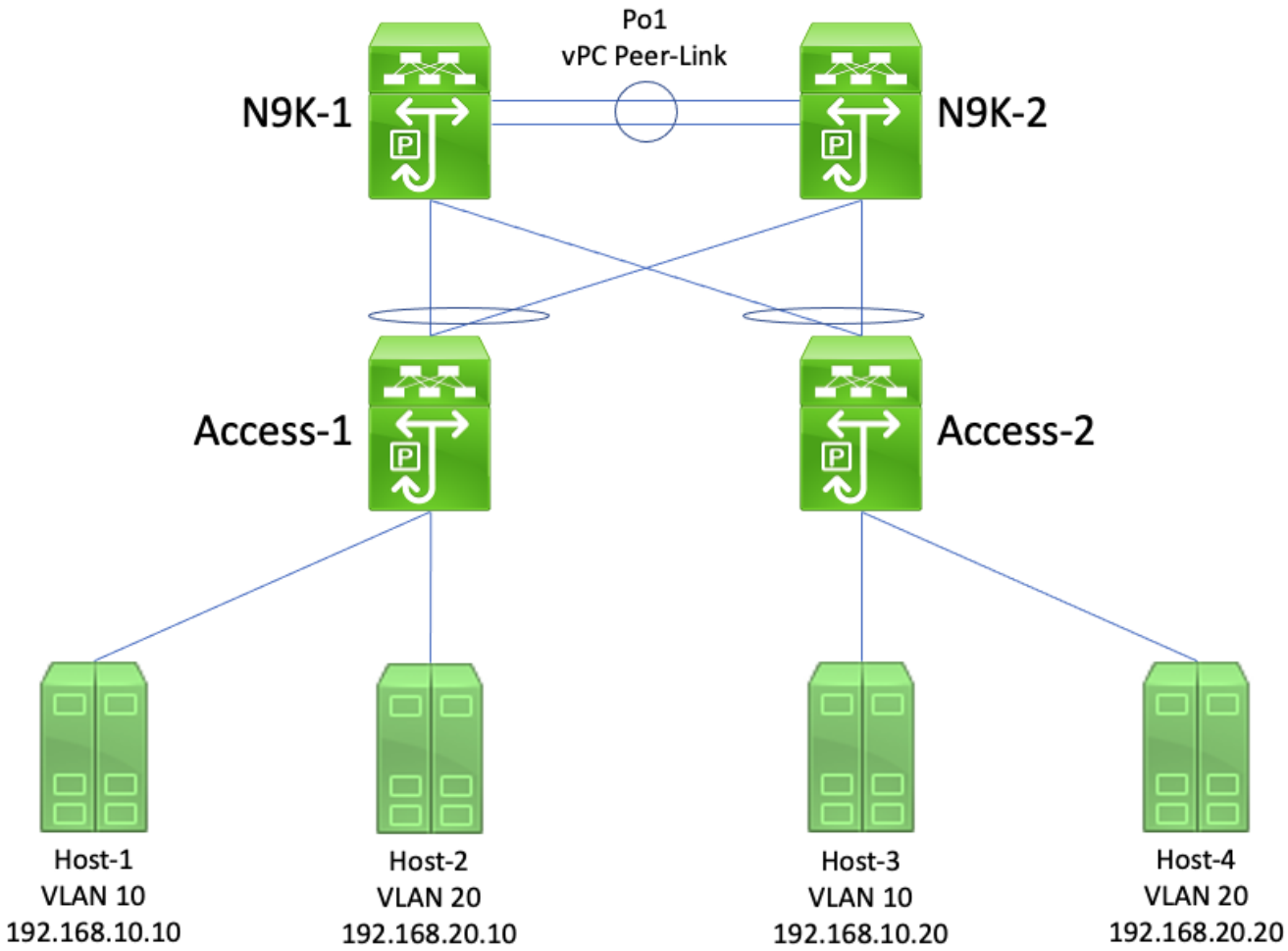
Questa modifica nelle porte radice designate determina l'esecuzione di tutte le porte Spanning Tree non edge nel computer a stato finito Spanning Tree Protocol (blocco, apprendimento e inoltro) con pause comprese tra quelle equivalenti al timer Forward Delay del protocollo Spanning

Tree configurato (15 secondi per impostazione predefinita). Questo processo può essere estremamente dannoso per la rete.

Nello stesso scenario di errore in cui il miglioramento dello switch peer vPC è abilitato, sia N9K-1 che N9K-2 trasmettono BPDU Spanning Tree identiche utilizzando l'indirizzo MAC di sistema vPC condiviso come ID bridge. Se N9k-1 non funziona, N9K-2 continua a trasmettere la stessa BPDU dello Spanning Tree. Di conseguenza, Access-1 esegue immediatamente la transizione del collegamento Blocco alternativo verso N9K-2 a uno stato Radice designato e inizia l'inoltro del traffico attraverso il collegamento. Inoltre, poiché l'ID root bridge dello Spanning Tree non cambia, non si rischia che le porte non edge passino attraverso la macchina a stati finiti dello Spanning Tree Protocol, riducendo così le possibilità di interruzioni nella rete.

I bridge connessi al vPC eliminano gli indirizzi MAC acquisiti in modo dinamico

Supponiamo di avere questa topologia:



In questa topologia, N9K-1 e N9K-2 sono peer vPC in un dominio vPC che eseguono il routing tra VLAN tra la VLAN 10 e la VLAN 20. N9K-1 è configurato con un valore di priorità dello Spanning Tree pari a 0 per la VLAN 10 e la VLAN 20, rendendo N9K-1 il bridge radice per entrambe le VLAN. N9K-2 è configurato con priorità Spanning Tree di 4096 per la VLAN 10 e la VLAN 20 ed è quindi il root bridge secondario di entrambe le VLAN. Host-1, Host-2, Host-3 e Host-4 comunicano continuamente tra loro.

Supponiamo che N9K-1 diventi inattivo a causa di un guasto hardware, un'interruzione dell'alimentazione o il ricaricamento dello switch. N9K-2 si afferma come root bridge per le VLAN 10 e VLAN 20 annunciando le Spanning Tree BPDU che usano il suo indirizzo MAC di sistema

come ID bridge. Access-1 e Access-2 rilevano una modifica nell'ID del bridge radice e, sebbene lo spanning tree resti lo stesso (ovvero, il vPC rivolto verso N9K-1 e N9K-2 rimane una porta radice designata), sia Access-1 che Access-2 scaricano il proprio indirizzo MAC di tutti gli indirizzi MAC appresi in modo dinamico nella VLAN 10 e nella VLAN 20.

Nella maggior parte degli ambienti, l'eliminazione degli indirizzi MAC acquisiti in modo dinamico causa conseguenze minime. Nessun pacchetto viene perso (a parte i pacchetti persi perché trasmessi a N9K-1 mentre era disconnesso), ma il traffico viene temporaneamente inviato a tutto il dominio di broadcast come traffico unknown unicast (modalità flooding) e tutti gli switch del dominio di broadcast acquisiscono di nuovo gli indirizzi MAC in modo dinamico.

Nello stesso scenario di errore, ma con la funzionalità migliorata vPC Peer Switch abilitata, i due peer N9K-1 e N9K-2 trasmetterebbero Spanning Tree BPDU identiche usando l'indirizzo MAC di sistema vPC condiviso come ID bridge. Se N9k-1 non funziona, N9K-2 continua a trasmettere la stessa BPDU dello Spanning Tree. Di conseguenza, Access-1 e Access-2 non sono a conoscenza di alcuna modifica nella topologia dello Spanning Tree: dal loro punto di vista, le BPDU dello Spanning Tree del bridge radice sono identiche, quindi non è necessario scaricare gli indirizzi MAC appresi in modo dinamico dalle VLAN rilevanti. In questo scenario di errore, il traffico unknown unicast non viene inviato a tutti i domini di broadcast.

Gateway dei peer vPC

In questa sezione viene descritta la funzionalità migliorata vPC Peer Gateway (Gateway dei peer vPC), abilitata con il comando di configurazione **peer-gateway** nel dominio vPC.

Panoramica

Gli switch Nexus configurati in un dominio vPC eseguono un inoltro First Hop Redundancy Protocol (FHRP) a doppia azione per impostazione predefinita. Ciò significa che se un peer vPC riceve un pacchetto con un indirizzo MAC di destinazione appartenente a un gruppo HSRP (Hot Standby Router Protocol) o VRRP (Virtual Router Redundancy Protocol) configurato sullo switch, lo switch instrada il pacchetto in base alla relativa tabella di routing locale, indipendentemente dallo stato del control plane HSRP o VRRP. In altre parole, ci si aspetta che un peer vPC in stato HSRP Standby o VRRP Backup indirizzi i pacchetti destinati all'indirizzo MAC virtuale di HSRP o VRRP.

Quando un peer vPC instrada un pacchetto destinato a un indirizzo MAC virtuale FHRP, riscrive il pacchetto con un nuovo indirizzo MAC di origine e destinazione. L'indirizzo MAC di origine è l'indirizzo MAC dell'interfaccia virtuale commutata (SVI) del peer vPC all'interno della VLAN a cui il pacchetto è indirizzato. L'indirizzo MAC di destinazione è l'indirizzo MAC associato all'indirizzo IP dell'hop successivo per l'indirizzo IP di destinazione del pacchetto in base alla tabella di routing locale del peer vPC. Negli scenari di routing tra VLAN, l'indirizzo MAC di destinazione del pacchetto dopo la riscrittura è l'indirizzo MAC dell'host a cui il pacchetto è destinato.

In alcuni host, per ottimizzare i risultati, non si seguono le regole di inoltro standard. Con questo comportamento, quando deve rispondere a un pacchetto in arrivo, l'host non cerca nella tabella di routing e/o nella cache ARP. Al contrario, inverte gli indirizzi MAC di origine e di destinazione del pacchetto in arrivo per il pacchetto di risposta. In altre parole, l'indirizzo MAC di origine del pacchetto in arrivo diventa l'indirizzo MAC di destinazione del pacchetto di risposta e l'indirizzo MAC di destinazione del pacchetto in arrivo diventa l'indirizzo MAC di origine del pacchetto di risposta. Nel comportamento di inoltro standard invece, l'host cerca nella tabella di routing locale

e/o nella cache ARP e imposta l'indirizzo MAC di destinazione del pacchetto di risposta sull'indirizzo MAC virtuale FHRP.

Il diverso comportamento dell'host può violare la regola vPC Loop Avoidance (Prevenzione dei loop vPC), in caso il pacchetto di risposta generato dall'host sia indirizzato a un peer vPC, ma venga trasmesso dal vPC all'altro peer vPC. L'altro peer vPC riceve il pacchetto destinato a un indirizzo MAC di proprietà del peer vPC e inoltra il pacchetto dal collegamento peer vPC verso il peer vPC proprietario dell'indirizzo MAC presente nel campo dell'indirizzo MAC di destinazione del pacchetto. Il peer vPC proprietario dell'indirizzo MAC tenta di instradare il pacchetto localmente. Se il pacchetto deve uscire da un vPC, il peer vPC lo scarta per aver violato la regola di prevenzione del loop vPC. È possibile quindi che si verifichino problemi di connettività o perdita di pacchetti in alcuni flussi provenienti da o destinati a un host con questa modalità di funzionamento non standard.

La funzionalità migliorata vPC Peer Gateway è stata introdotta proprio per evitare che un tale comportamento non standard causi la perdita di pacchetti. Viene quindi consentito a un peer vPC di indirizzare localmente i pacchetti destinati all'indirizzo MAC dell'altro peer vPC in modo che i pacchetti destinati al peer vPC remoto non debbano passare per il collegamento vPC Peer-Link. In altre parole, la funzionalità migliorata vPC Peer Gateway permette a un peer vPC di indirizzare i pacchetti "per conto del" peer vPC remoto. La funzionalità migliorata vPC Peer Gateway può essere abilitata con il comando di configurazione **peer-gateway** nel dominio vPC.

Avvertenze

Instabilità delle adiacenze del protocollo di routing unicast sui vPC o sulle VLAN del vPC

Se si formano adiacenze del protocollo di routing unicast dinamico tra due peer vPC e un router connesso al vPC o un router connesso tramite una porta non vPC, si potrebbe osservare un'instabilità delle adiacenze che cambiano incessantemente stato dopo aver abilitato la funzionalità migliorata vPC Peer Gateway se non è stata configurata immediatamente anche la funzionalità Routing/Layer 3 over vPC. Questi scenari di errore sono descritti in dettaglio negli esempi Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway e Adiacenze del protocollo di routing unicast su una VLAN del vPC con vPC Peer Gateway in questo documento.

Per risolvere il problema, abilitare la funzionalità migliorata Routing/Layer 3 over vPC con il comando di configurazione **layer3 peer-router** nel dominio vPC subito dopo aver abilitato la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC.

Disattivazione automatica dei reindirizzamenti ICMP e ICMPv6

Quando il miglioramento di vPC Peer Gateway è abilitato, la generazione di pacchetti di reindirizzamento ICMP e ICMPv6 viene disabilitata automaticamente su tutte le SVI VLAN vPC (ossia, qualsiasi SVI associato a una VLAN che sia trunking attraverso il collegamento peer vPC). Lo switch imposta quindi i comandi **no ip redirects** e **no ipv6 redirects** sulle SVI di tutte le VLAN del vPC. Ciò impedisce allo switch di generare pacchetti di reindirizzamento ICMP in risposta ai pacchetti che entrano nello switch ma hanno indirizzi IP e MAC di destinazione del peer vPC.

Se sono necessari pacchetti di reindirizzamento ICMP o ICMPv6 nell'ambiente in uso all'interno di una VLAN specifica, è necessario escludere questa VLAN dallo sfruttamento del miglioramento di vPC Peer Gateway usando il comando di configurazione del dominio vPC **peer-gateway exclude-**

vlan<vlan-id>.

> **Nota**: il comando **peer-gateway exclude-vlan <vlan-id>** vPC domain configuration non è supportato sugli switch Nexus serie 9000.

Configurazione

Riportiamo qui un esempio di configurazione della funzionalità vPC Peer Gateway.

Nell'esempio, N9K-1 e N9K-2 sono i peer vPC di un dominio vPC. Entrambi i peer vPC dispongono di un gruppo HSRP configurato per la VLAN 10. N9K-1 è il router attivo HSRP con priorità 150, mentre N9K-2 è il router di standby HSRP con priorità predefinita di 100.

```
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)
```

Su N9K-1 la SVI della VLAN 10 ha un indirizzo MAC pari a 00ee.ab67.db47, su N9K-2 l'indirizzo MAC della SVI della VLAN 10 è 00ee.abd8.747f. L'indirizzo MAC virtuale HSRP della VLAN 10 è 0000.0c07.ac0a. In questo stato, l'indirizzo MAC della SVI della VLAN 10 di ciascuno switch e l'indirizzo MAC virtuale HSRP sono presenti nella tabella degli indirizzi MAC di ogni switch. L'indirizzo MAC VLAN 10 SVI e l'indirizzo MAC virtuale HSRP di ciascuno switch hanno il flag Gateway (G) presente, che indica che lo switch instrada localmente i pacchetti destinati a questo indirizzo MAC.

Tenere presente che nella tabella degli indirizzi MAC di N9K-1, l'indirizzo MAC della SVI della VLAN 10 di N9K-2 non è contrassegnato con il flag Gateway. Analogamente, nella tabella degli indirizzi MAC di N9K-2 anche l'indirizzo MAC della SVI della VLAN 10 di N9K-1 non è contrassegnato con il flag Gateway.

```
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)
```

La funzionalità vPC Peer Gateway può essere abilitata con il comando di configurazione **peer-gateway** nel dominio vPC. In questo modo lo switch può indirizzare localmente i pacchetti ricevuti con un indirizzo MAC di destinazione appartenente all'indirizzo MAC del peer vPC appreso sul collegamento peer vPC. A tale scopo, impostare il flag Gateway per l'indirizzo MAC del peer vPC nella tabella degli indirizzi MAC dello switch.

```
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#
```

Per verificare che vPC Peer Gateway funzioni come previsto, verificare che l'indirizzo MAC del vPC peer sia contrassegnato con il flag Gateway nella tabella degli indirizzi MAC.

```
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)
```

Conseguenze

L'impatto dell'abilitazione del miglioramento di vPC Peer Gateway può variare a seconda della topologia circostante e del comportamento degli host connessi, come descritto nelle seguenti sottosezioni. Se nessuna delle sottosezioni seguenti si applica all'ambiente, l'abilitazione del miglioramento di vPC Peer Gateway non comporta interruzioni e non ha alcun impatto sull'ambiente.

Instabilità delle adiacenze del protocollo di routing unicast sui vPC o sulle VLAN del vPC

Se si formano adiacenze del protocollo di routing unicast dinamico tra due peer vPC e un router connesso al vPC o un router connesso tramite una porta non vPC, si potrebbe osservare un'instabilità delle adiacenze che cambiano incessantemente stato dopo aver abilitato la funzionalità migliorata vPC Peer Gateway se non è stata configurata immediatamente anche la funzionalità Routing/Layer 3 over vPC. Questi scenari di errore sono descritti in dettaglio negli esempi Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway e Adiacenze del protocollo di routing unicast su una VLAN del vPC con vPC Peer Gateway in questo documento.

Per risolvere il problema, abilitare la funzionalità migliorata Routing/Layer 3 over vPC con il comando di configurazione **layer3 peer-router** nel dominio vPC subito dopo aver abilitato la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC.

Disattivazione automatica dei reindirizzamenti ICMP e ICMPv6

Quando il miglioramento di vPC Peer Gateway è abilitato, la generazione di pacchetti di reindirizzamento ICMP e ICMPv6 viene disabilitata automaticamente su tutte le SVI VLAN vPC (ossia, qualsiasi SVI associato a una VLAN che sia trunking attraverso il collegamento peer vPC).

Lo switch imposta quindi i comandi **no ip redirects** e **no ipv6 redirects** sulle SVI di tutte le VLAN del vPC. Ciò impedisce allo switch di generare pacchetti di reindirizzamento ICMP in risposta ai pacchetti che entrano nello switch ma hanno indirizzi IP e MAC di destinazione del peer vPC.

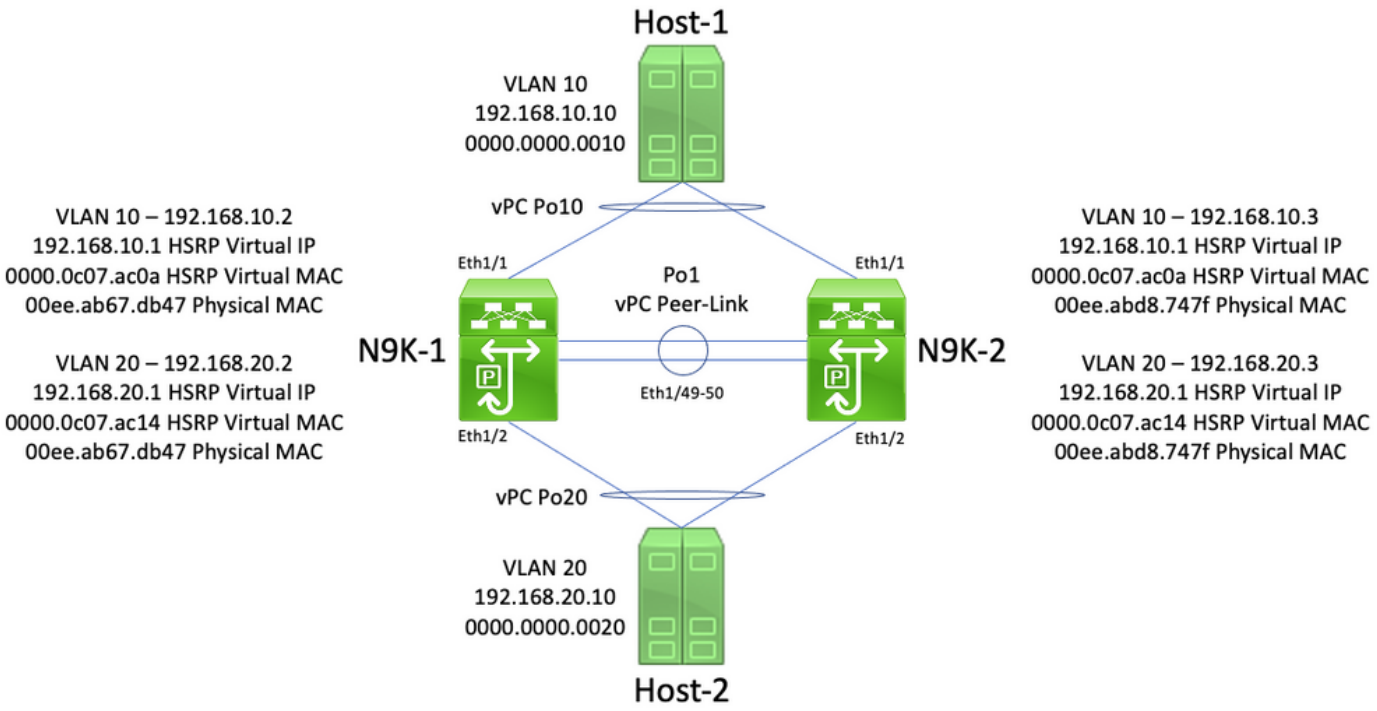
Se sono necessari pacchetti di reindirizzamento ICMP o ICMPv6 nell'ambiente in uso all'interno di una VLAN specifica, è necessario escludere questa VLAN dallo sfruttamento del miglioramento di vPC Peer Gateway usando il comando di configurazione del dominio vPC **peer-gateway exclude-vlan<vlan-id>**.

> **Nota**: il comando **peer-gateway exclude-vlan <vlan-id>** vPC domain configuration non è supportato sugli switch Nexus serie 9000.

Esempi di scenari di errore

Host connessi al vPC con inoltro non standard

Supponiamo di avere questa topologia:

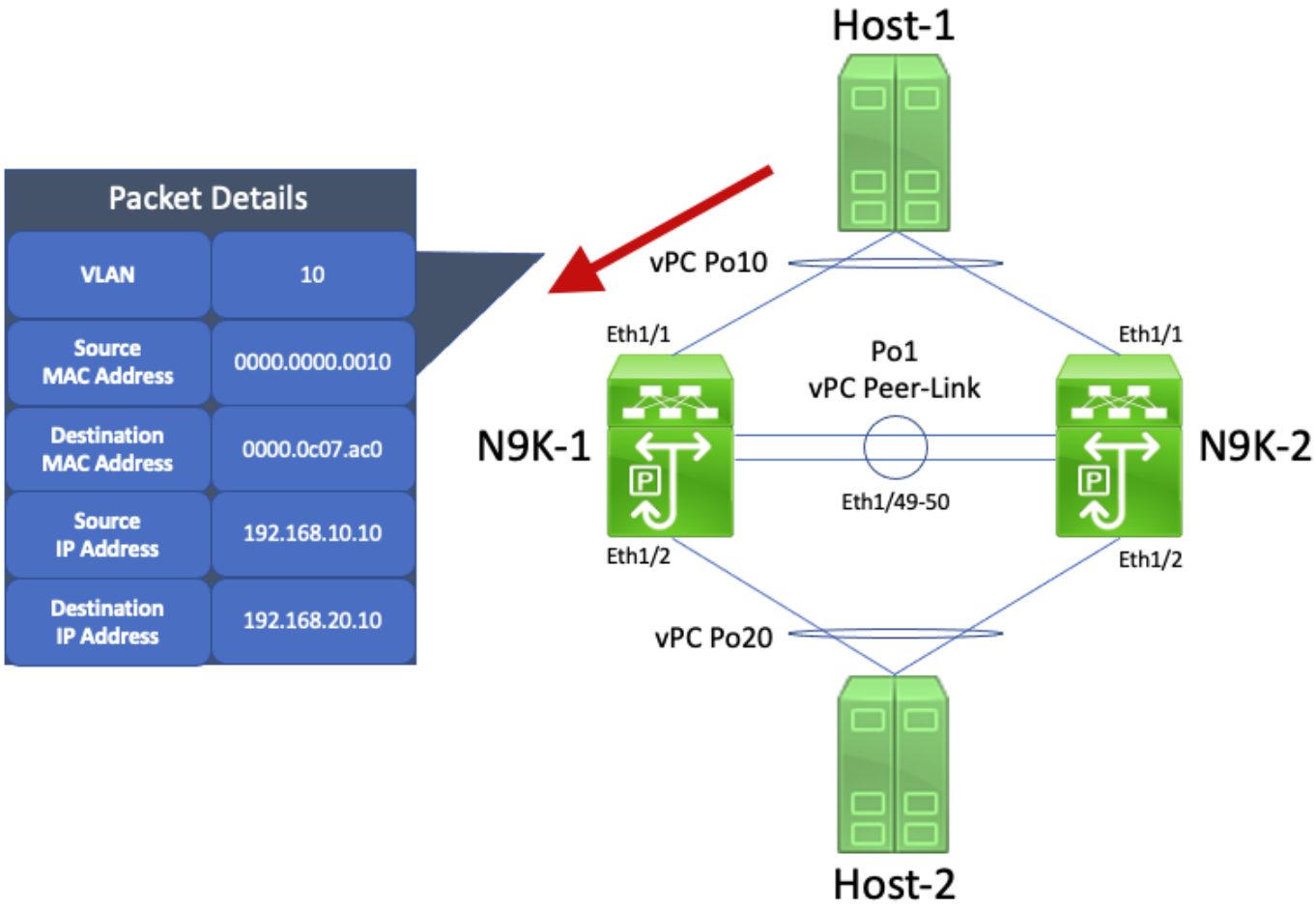


In questa topologia, N9K-1 e N9K-2 sono peer vPC in un dominio vPC che eseguono il routing tra VLAN tra la VLAN 10 e la VLAN 20. Po1 è l'interfaccia del collegamento tra i peer vPC, o vPC Peer-Link. Un host denominato Host-1 è connesso tramite vPC Po10 a N9K-1 e N9K-2 nella VLAN 10. L'host 1 possiede un indirizzo IP di 192.168.10.10 con un indirizzo MAC di 0000.0000.0010. Un host denominato Host-2 è connesso tramite vPC Po20 a N9K-1 e N9K-2 nella VLAN 20. L'host 2 possiede un indirizzo IP di 192.168.20.10 con un indirizzo MAC di 0000.0000.0020.

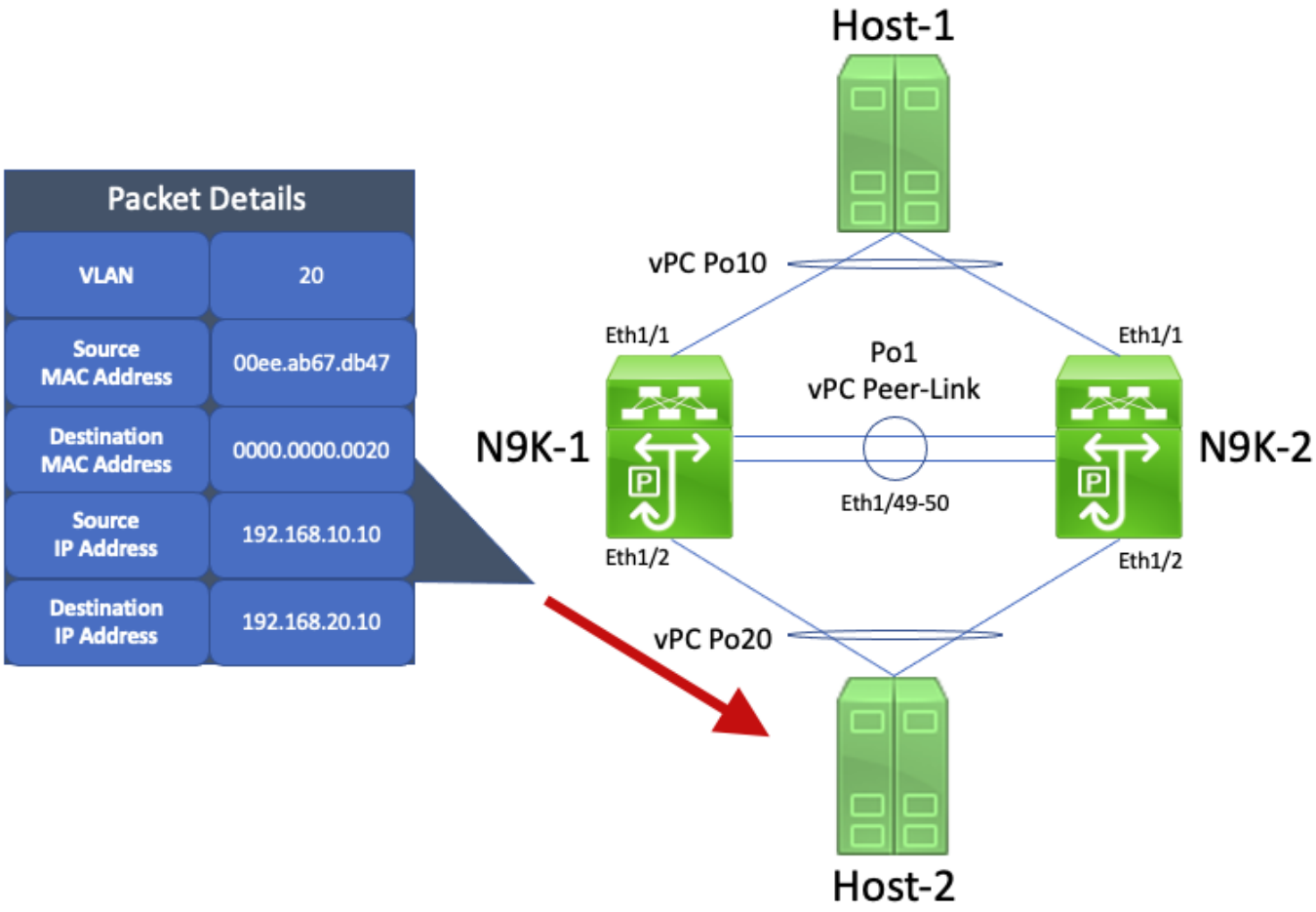
N9K-1 e N9K-2 hanno entrambi le SVI nella VLAN 10 e nella VLAN 20 con protocollo HSRP attivato per ciascuna SVI. L'interfaccia VLAN 10 di N9K-1 ha un indirizzo IP di 192.168.10.2 e l'interfaccia VLAN 20 di N9K-1 ha un indirizzo IP di 192.168.20.2. Entrambe le SVI di N9K-1 hanno un indirizzo MAC fisico di 00ee.ab67.db47. L'interfaccia VLAN 10 di N9K-2 ha un indirizzo IP di 1921.68.10.3, mentre l'interfaccia VLAN 20 di N9K-2 ha un indirizzo IP di 192.168.20.3. Entrambe

le SVI di N9K-2 hanno un indirizzo MAC fisico di 00ee.abd8.747f. L'indirizzo IP virtuale HSRP della VLAN 10 è 192.168.10.1, l'indirizzo MAC virtuale HSRP è 0000.0c07.ac0a. L'indirizzo IP virtuale HSRP della VLAN 20 è 192.168.20.1, l'indirizzo MAC virtuale HSRP è 0000.0c07.ac14.

Si prenda in considerazione uno scenario in cui l'host 1 invia un pacchetto di richiesta echo ICMP all'host 2. Dopo che l'host 1 ha risolto l'ARP per il gateway predefinito (indirizzo IP virtuale HSRP), l'host 1 segue il comportamento di inoltro standard e genera un pacchetto di richiesta echo ICMP con indirizzo IP di origine 192.168.10.10, indirizzo IP di destinazione 192.168.20.10, indirizzo MAC di origine 000.000.0010 e indirizzo MAC di destinazione 000.0c07.ac0a. Questo pacchetto inizia verso N9K-1. Ecco un diagramma grafico esplicativo.

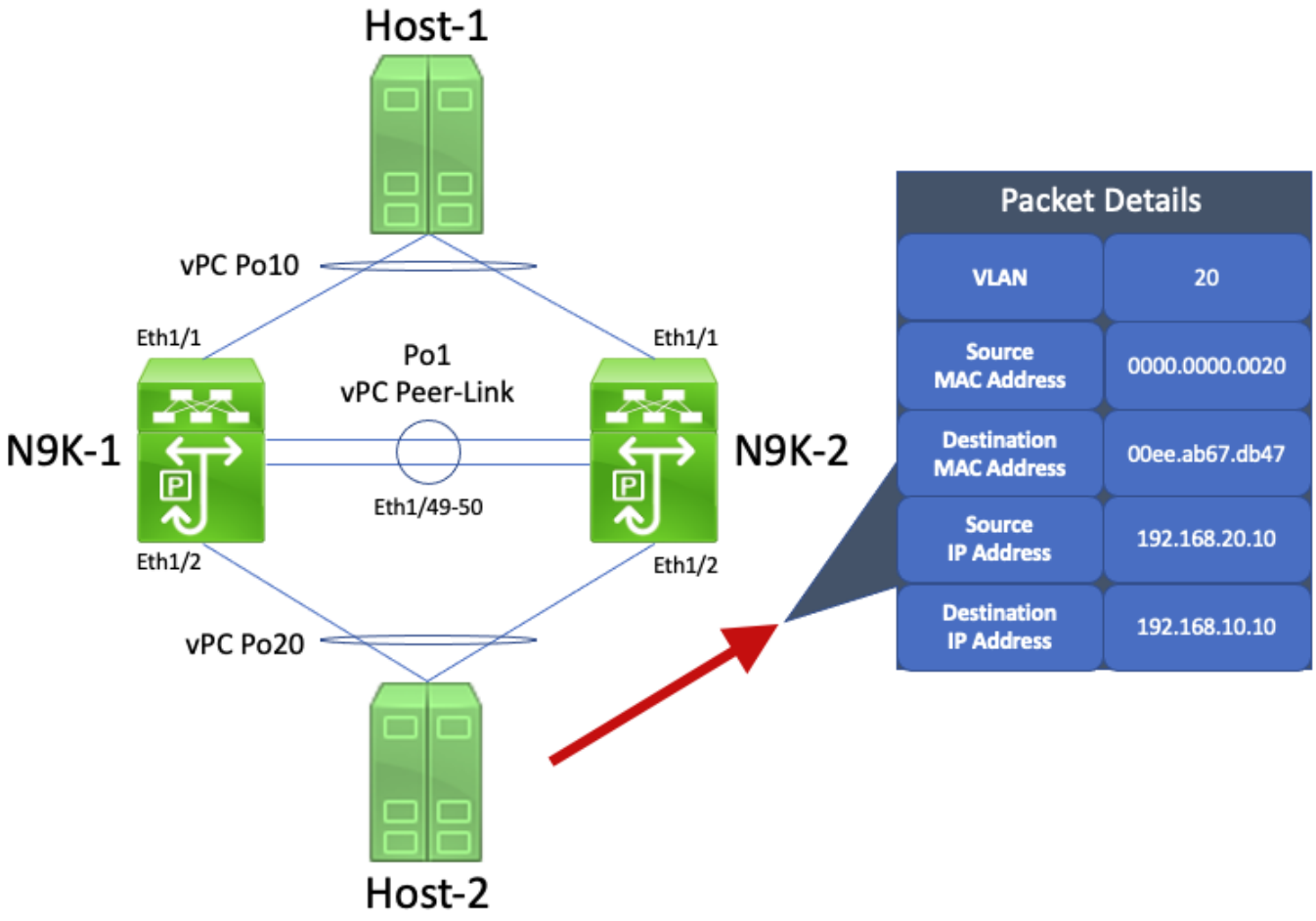


N9K-1 riceve il pacchetto. Poiché il pacchetto è destinato all'indirizzo MAC virtuale HSRP, N9K-1 può indirizzarlo in base alla tabella di routing locale a prescindere dallo stato del piano di controllo HSRP. Questo pacchetto viene indirizzato dalla VLAN 10 alla VLAN 20. Come parte del routing del pacchetto, N9K-1 esegue la riscrittura del pacchetto riindirizzando i campi degli indirizzi MAC di origine e destinazione del pacchetto. Il nuovo indirizzo MAC di origine del pacchetto è l'indirizzo MAC fisico associato alla VLAN 20 SVI (00ee.ab67.db47) della N9K-1 e il nuovo indirizzo MAC di destinazione è l'indirizzo MAC associato all'host-2 (0000.000.0020). Ecco un diagramma grafico esplicativo.

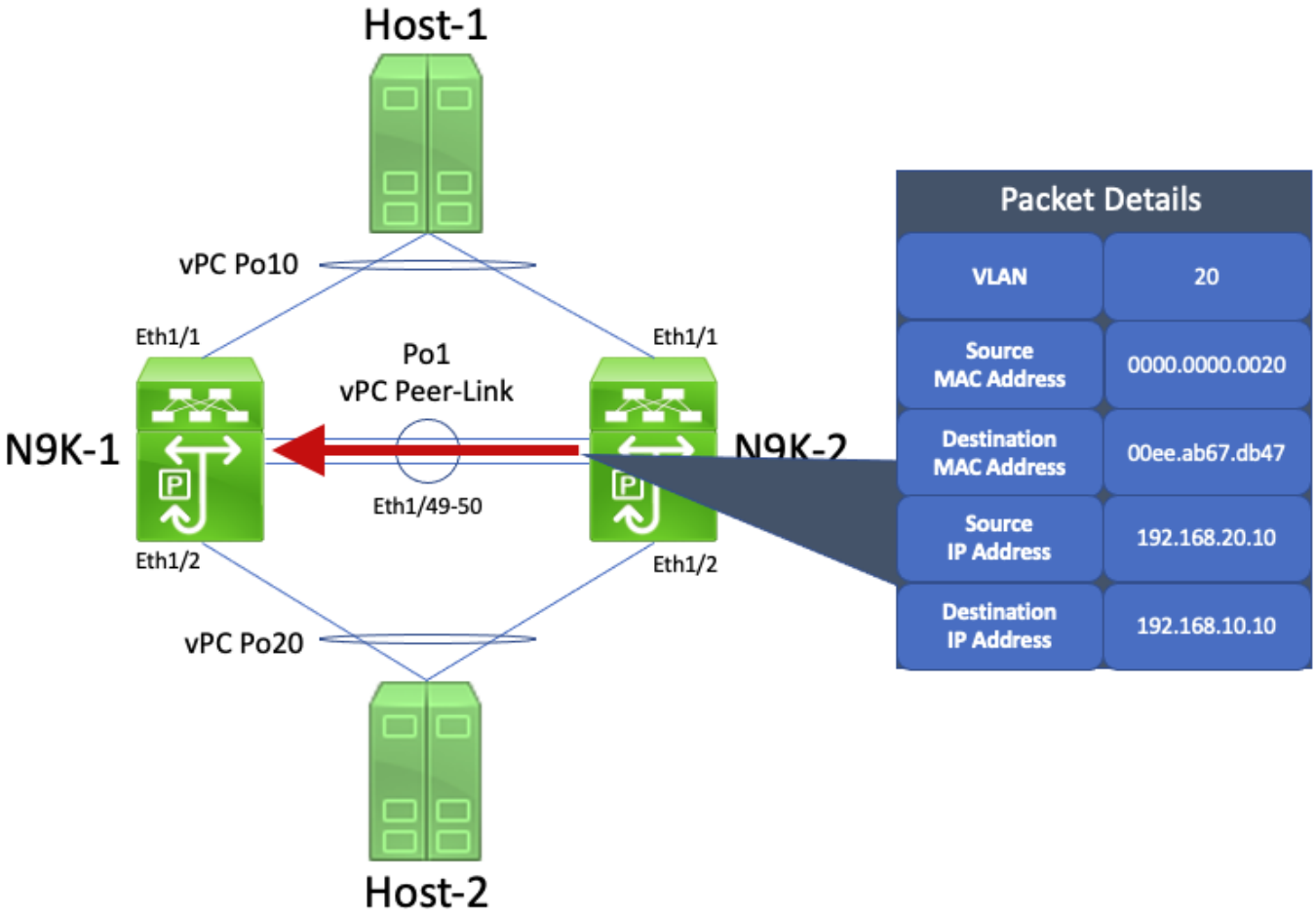


L'Host-2 riceve il pacchetto e genera un pacchetto di risposta echo ICMP in risposta al pacchetto di richiesta echo ICMP dell'Host-1. Tuttavia, l'Host-2 non applica in questo caso una politica di inoltro standard. Per ottimizzare l'inoltro, l'Host-2 non cerca l'indirizzo IP dell'Host-1 (192.168.10.10) nella tabella di routing o nella cache ARP, ma inverte i campi degli indirizzi MAC di origine e di destinazione del pacchetto di richiesta echo ICMP originariamente ricevuto dall'Host-2. Di conseguenza, il pacchetto di risposta echo ICMP generato dall'host 2 ha un indirizzo IP di origine di 192.168.20.10, un indirizzo IP di destinazione di 192.168.10.10, un indirizzo MAC di origine di 0000.000.0020 e un indirizzo MAC di destinazione di 00ee.ab67.db47.

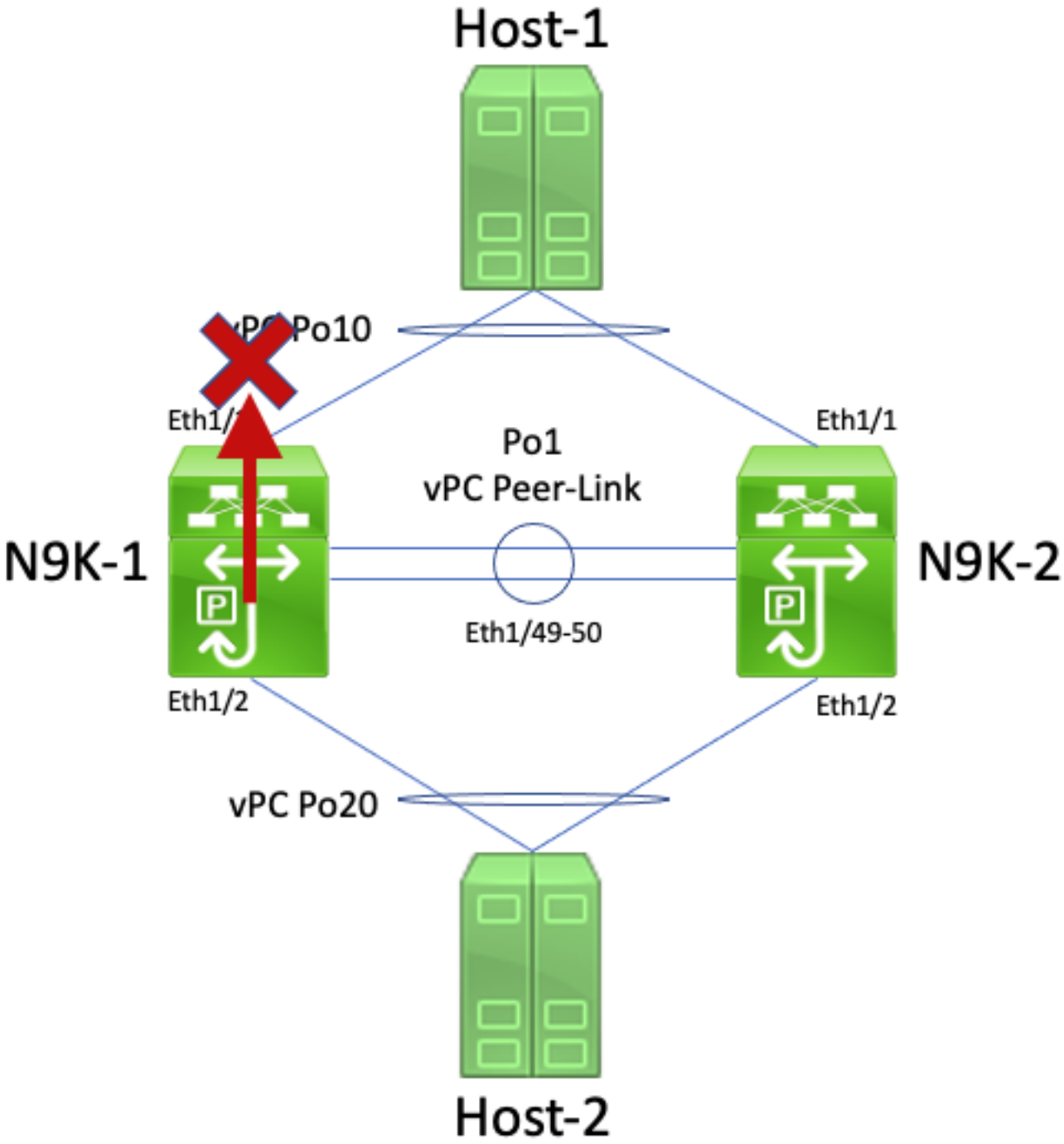
Se il pacchetto di risposta echo ICMP inizia a N9K-1, il pacchetto viene inoltrato all'host-1 senza problemi. Tuttavia, supponiamo di avere un pacchetto di risposta echo ICMP inoltrato a N9K-2, come mostrato qui.



N9K-2 riceve il pacchetto. Poiché questo pacchetto è destinato all'indirizzo MAC fisico della VLAN 20 SVI della N9K-1, N9K-2 inoltra il pacchetto attraverso vPC Peer-Link verso N9K-1, in quanto N9K-2 non può inoltrare il pacchetto per conto di N9K-1. Ecco un diagramma grafico esplicativo.



N9K-1 riceve il pacchetto. Poiché il pacchetto è destinato all'indirizzo MAC virtuale della SVI della VLAN 20 di N9K-1, N9K-1 può indirizzarlo in base alla tabella di routing locale a prescindere dallo stato del piano di controllo HSRP. Questo pacchetto viene indirizzato dalla VLAN 20 alla VLAN 10. Tuttavia, l'interfaccia in uscita per questa route si risolve in vPC Po10, attivo su N9K-2. Questa è una violazione della regola vPC Loop Avoidance: se N9K-1 riceve un pacchetto tramite vPC Peer-Link, N9K-1 non può inoltrarlo da un'interfaccia vPC se la stessa interfaccia vPC è attiva su N9K-2. L'utente N9K-1 rifiuta questo pacchetto come risultato di questa violazione. Ecco un diagramma grafico esplicativo.



Il problema può essere risolto abilitando la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC. Ciò consente al N9K-2 di indirizzare il pacchetto di risposta echo ICMP (e altri pacchetti indirizzati in modo simile) per conto del N9K-1, anche se l'indirizzo MAC di destinazione del pacchetto è di proprietà del N9K-1 e non del N9K-2. Di conseguenza, N9K-2 può inoltrare questo pacchetto dall'interfaccia vPC Po10 invece di inoltrarlo attraverso vPC Peer-Link.

Routing/Layer 3 su vPC (layer3 peer-router)

In questa sezione viene descritta la funzionalità migliorata Routing/Layer 3 over vPC (Routing/Layer 3 su vPC), abilitata con il comando di configurazione **layer3 peer-router** nel dominio vPC.

Nota: la creazione di adiacenze del protocollo di routing multicast (ovvero adiacenze PIM (Protocol Independent Multicast) su un vPC non è supportata con il miglioramento Routing/Layer 3 su vPC abilitato.

Panoramica

In alcuni ambienti, i clienti potrebbero chiedere di collegare una coppia di switch Nexus tramite vPC e di formare adiacenze del protocollo di routing unicast sul vPC con entrambi i peer vPC. Oppure, potrebbero voler collegare un router a un unico peer vPC tramite una VLAN del vPC e formare adiacenze del protocollo di routing unicast con entrambi i peer vPC sulla VLAN. Il router connesso al vPC avrebbe così un percorso ECMP (Equal-Cost Multi-Path) per i prefissi annunciati da entrambi gli switch Nexus. Questa situazione potrebbe essere preferibile ad avere dei collegamenti di routing dedicati tra il router connesso al vPC e i due peer vPC per ridurre gli indirizzi IP, che sarebbero 3 anziché 4, o per avere una configurazione più semplice, con interfacce indirizzate sulle SVI, in particolare negli ambienti VRF-Lite che potrebbero richiedere l'uso di interfacce secondarie.

Le adiacenze del protocollo di routing unicast su un vPC non erano supportate in passato sulle piattaforme Cisco Nexus. Tuttavia, i clienti potrebbero aver implementato una topologia che ne permette il formarsi corretto su un vPC. Dopo aver apportato delle modifiche alla rete, ad esempio aver aggiornato il software sul router connesso al vPC o sui peer vPC o dopo un failover del firewall, le adiacenze del protocollo di routing unicast su un vPC non funzionano più con conseguente perdita di pacchetti sul traffico del piano dati o mancata formazione delle adiacenze su uno o entrambi i peer vPC. I dettagli tecnici di questi scenari sono discussi nella sezione Esempi di scenari di errore in questo documento.

La funzionalità Routing/Layer 3 over vPC è stata introdotta per supportare la formazione di adiacenze del protocollo di routing unicast su un vPC. Tale scopo viene raggiunto autorizzando l'inoltro dei pacchetti del protocollo di routing unicast con TTL pari a 1 sul vPC Peer-Link senza di diminuire il valore TTL del pacchetto. Le adiacenze del protocollo di routing unicast potranno così formarsi sul vPC o sulla VLAN del vPC senza problemi. La funzionalità Routing/Layer 3 over vPC può essere abilitata con il comando di configurazione **layer3 peer-router** nel dominio vPC dopo aver abilitato la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC.

Le versioni software NX-OS che supportano la funzionalità Routing/Layer 3 over vPC su ciascuna piattaforma Cisco Nexus sono riportate nella Tabella 2 ("Supporto delle adiacenze dei protocolli di routing sulle VLAN del vPC") del documento Topologie supportate per il routing su vPC (Virtual Port Channel) sulle piattaforme Nexus.

Avvertenze

Generazione reiterata di syslog VPC-2-L3_VPC_UNEQUAL_WEIGHT

Dopo l'abilitazione del miglioramento di Routing/Layer 3 su vPC, entrambi i peer vPC iniziano a generare syslog simili a quelli riportati di seguito ogni ora:

```
2021 May 26 19:13:47.079 switch %VPC-2-L3_VPC_UNEQUAL_WEIGHT: Layer3 peer-router is enabled.
Please make sure both vPC peers have the same L3 routing configuration.
2021 May 26 19:13:47.351 switch %VPC-2-L3_VPC_UNEQUAL_WEIGHT: Unequal weight routing is not
supported in L3 over vPC. Please make sure both vPC peers have equal link cost configuration
```

Nessuno di questi syslog è indice di un problema nello switch. I syslog avvisano semplicemente l'amministratore che configurazione, costo e peso del routing devono essere identici su entrambi i peer vPC quando la funzionalità Routing/Layer 3 over vPC è abilitata in modo che i due peer vPC possano indirizzare il traffico allo stesso modo. Non indicano necessariamente differenze nella configurazione, nel costo o nel peso del routing dei due peer vPC.

I syslog possono essere disabilitati con questa configurazione.

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

Questa configurazione deve essere eseguita su entrambi i peer vPC per disabilitare il syslog su entrambi i peer vPC.

Traffico Data Plane con TTL di 1 software inoltrato a causa dell'ID bug Cisco CSCvs82183 e ID bug Cisco CSCvw16965

Quando il miglioramento di Routing/Layer 3 su vPC è abilitato sugli switch Nexus serie 9000 dotati di un ASIC Cloud Scale che esegue una versione software NX-OS precedente alla versione software NX-OS 9.3(6), il traffico del piano dati non associato a un protocollo di routing unicast con TTL pari a 1 viene indirizzato al supervisore e inoltrato nel software anziché nell'hardware. A seconda che lo switch Nexus sia uno switch a chassis fisso (detto anche "Top of Rack") o modulare (detto anche "End of Row") o una versione software NX-OS corrente dello switch, la causa principale di questo problema può essere attribuita al difetto software Cisco ID bug CSCvs82183 o guasto software Cisco ID bug CSCvw16965 . Entrambi i problemi software riguardano solo gli switch Nexus serie 9000 dotati di un ASIC su scala cloud - nessun'altra piattaforma hardware Cisco Nexus è interessata da entrambi i problemi. Per ulteriori dettagli, consultare le informazioni di ciascun difetto software.

Per evitare questi difetti software, Cisco consiglia di aggiornare il software NX-OS alla versione 9.3(6) o successive. In generale, Cisco consiglia di aggiornare regolarmente le versioni del software NX-OS sugli switch Nexus serie 9000 menzionate nel documento Versioni Cisco NX-OS consigliate per gli switch Cisco Nexus serie 9000.

Configurazione

Questo è un esempio di come configurare la funzionalità Routing/Layer 3 over vPC.

Nell'esempio, N9K-1 e N9K-2 sono i peer vPC di un dominio vPC. Su entrambi i peer vPC la funzionalità vPC Peer Gateway, necessaria per abilitare Routing/Layer 3 over vPC, è già stata abilitata. Entrambi i peer vPC dispongono di una SVI nella VLAN 10, abilitata nel processo OSPF 1. N9K-1 e N9K-3 sono bloccati in uno stato OSPF EXSTART/EXCHANGE con un router OSPF connesso al vPC con un indirizzo IP e ID router adiacente di 192.168.10.3.

```
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
```

Possiamo abilitare la funzionalità Routing/Layer 3 over vPC usando il comando di configurazione **layer3 peer-router** nel dominio vPC. In questo modo si impedisce a un peer vPC di ridurre il valore TTL dei pacchetti del protocollo di routing unicast instradati in seguito all'abilitazione del miglioramento di vPC Peer Gateway.

```
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#
```

Èpossibile verificare che Routing/Layer 3 over vPC funzioni come previsto confermando che l'adiacenza OSPF con il dispositivo vicino connesso al vPC passi allo stato FULL subito dopo aver abilitato Routing/Layer 3 over vPC.

```
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
```

Conseguenze

La funzionalità Routing/Layer 3 over vPC non ha conseguenze intrinseche sul dominio vPC. Ciò significa che quando si abilita il miglioramento di Routing/Layer 3 su vPC, né il peer vPC sospende i vPC, né il traffico del piano dati viene influenzato intrinsecamente dall'abilitazione di questo miglioramento.

Tuttavia, se le adiacenze del protocollo di routing dinamico, precedentemente inattive, diventano attive come conseguenza dell'aver abilitato la funzionalità Routing/Layer 3 over vPC, a seconda del ruolo delle adiacenze interessate, i prefissi annunciati tramite tali adiacenze e lo stato corrente della tabella di routing unicast, si potrebbero osservare alcune interruzioni della rete.

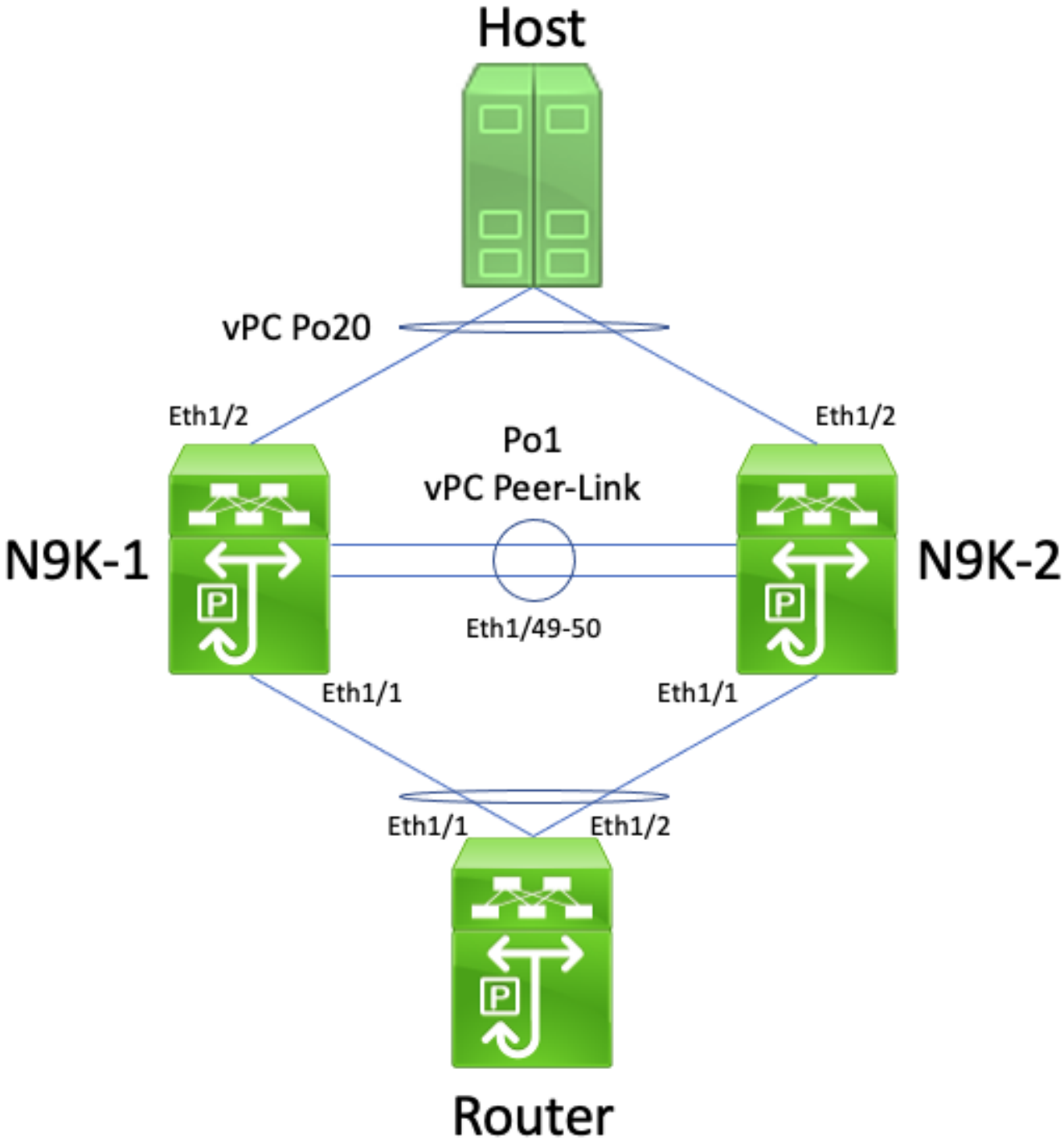
Per questo motivo, Cisco consiglia ai clienti di abilitare questo miglioramento durante un intervento di manutenzione in modo che si possa verificare un'interruzione del control plane o del data plane, a meno che i clienti non siano estremamente certi che le adiacenze del protocollo di routing interessate non influiscano in modo significativo sul funzionamento della rete.

Cisco consiglia anche di riesaminare attentamente la sezione Avvertenze di questo documento per verificare che non vi siano difetti nella versione del software NX-OS in uso che possano causare l'elaborazione del traffico del piano dati con TTL 1 nel software anziché nell'hardware.

Esempi di scenari di errore

Adiacenze del protocollo di routing unicast su un vPC senza vPC Peer Gateway

Supponiamo di avere questa topologia:



In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway non è abilitata. Po1 è l'interfaccia del collegamento tra i peer vPC, o vPC Peer-Link. Un router con nome host Router è connesso tramite vPC Po10 a N9K-1 e N9K-

2. Un host è collegato a N9K-1 e N9K-2 tramite vPC Po20. L'interfaccia Po10 del router è un canale di porta routing attivato in un protocollo di routing unicast. N9K-1 e N9K-2 hanno interfacce SVI attivate con lo stesso protocollo di routing unicast e si trovano nello stesso dominio di broadcast del Router.

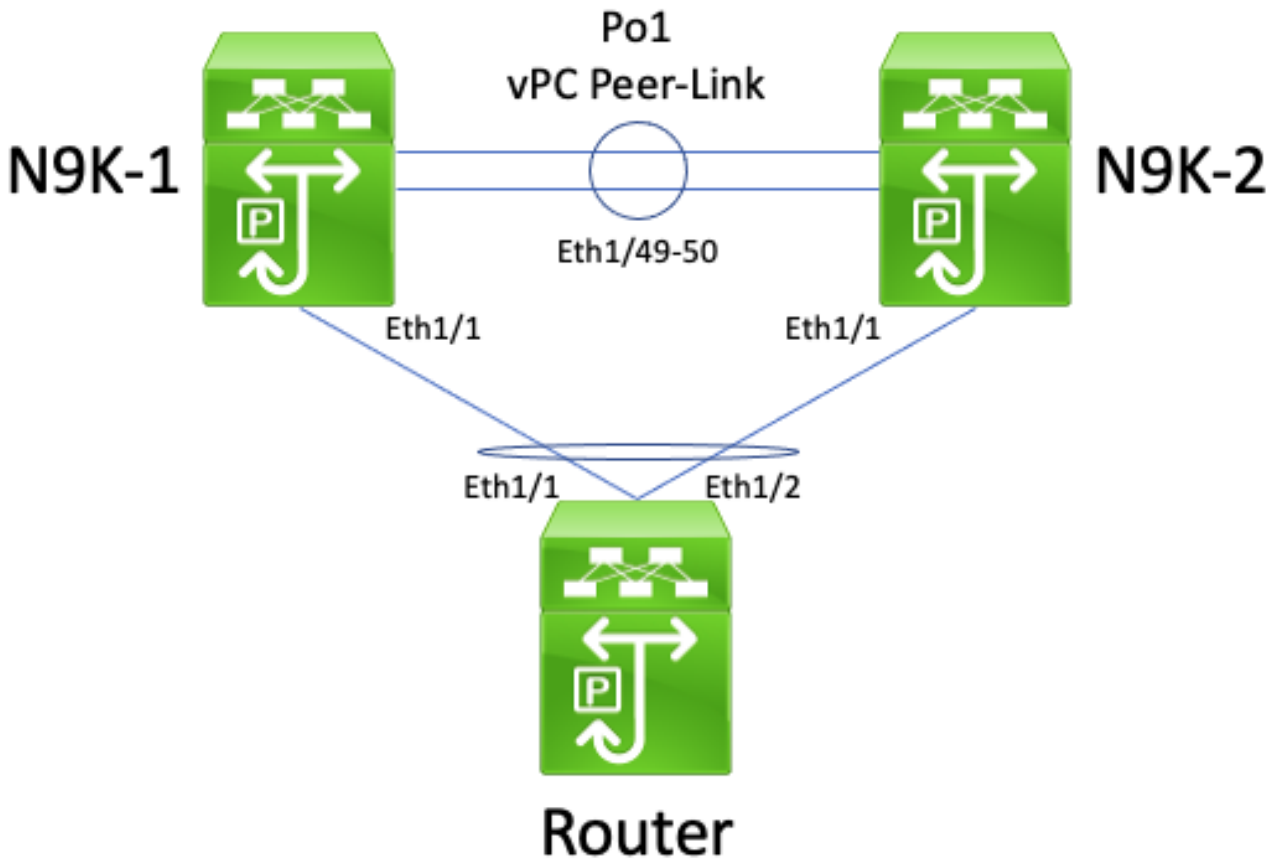
Le adiacenze del protocollo di routing unicast su un vPC senza la funzionalità vPC Peer Gateway abilitata non sono supportate perché la decisione di hashing ECMP del router connesso al vPC e la decisione di hashing del port-channel di Layer 2 potrebbero essere diverse. In questa topologia, le adiacenze tra i protocolli di routing si formerebbero correttamente tra i router N9K-1 e N9K-2. Prendere in considerazione il flusso del traffico tra il router e l'host. Il traffico del piano dati indirizzato all'Host attraverso il Router può essere riscritto con un indirizzo MAC di destinazione che appartiene all'indirizzo MAC della SVI di N9K-1 (a causa della decisione di hashing ECMP presa dal router), ma che viene trasmesso sull'interfaccia Ethernet1/2 (a causa della decisione di hashing del port-channel di Layer 2 presa dal router).

N9K-2 riceve questo pacchetto e lo inoltra attraverso il vPC Peer-Link, poiché l'indirizzo MAC di destinazione appartiene a N9K-1 e il miglioramento di vPC Peer Gateway (che consente a N9K-2 di instradare il pacchetto per conto di N9K-1) non è abilitato. N9K-1 riceve questo pacchetto sul vPC Peer-Link e riconosce che dovrebbe inoltrarlo fuori dalla sua rete Ethernet1/2 in vPC Po20. In questo modo si viola la regola vPC Loop Avoidance, quindi N9K-1 scarta il pacchetto nell'hardware. Ne derivano problemi di connettività o perdita di pacchetti su alcuni flussi che attraversano il dominio vPC.

Per risolvere il problema, abilitare la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC, quindi abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer3 peer-router** nel dominio vPC. Per ridurre al minimo i problemi di rete, abilitare entrambe le funzionalità vPC in rapida successione in modo da evitare lo scenario di errore descritto in Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway.

Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway

Supponiamo di avere questa topologia:



In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway è abilitata. Po1 è l'interfaccia del collegamento tra i peer vPC, o vPC Peer-Link. Un router con nome host Router è connesso tramite vPC Po10 a N9K-1 e N9K-2. L'interfaccia Po10 del router è un canale di porta routing attivato in un protocollo di routing unicast. N9K-1 e N9K-2 hanno interfacce SVI attivate con lo stesso protocollo di routing unicast e si trovano nello stesso dominio di broadcast del Router.

Le adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway abilitata non sono supportate perché la funzionalità vPC Peer Gateway potrebbe impedire la formazione di adiacenze del protocollo di routing unicast tra il router connesso a vPC e entrambi i peer vPC. In questa topologia, un protocollo di routing adiacente tra il router e l'host N9K-1 o N9K-2 potrebbe non raggiungere il risultato previsto, a seconda di come i pacchetti del protocollo di routing unicast sono stati originati dal router nell'hash N9K-1 o N9K-2 in vPC Po10.

Tutti i router sono in grado di inviare e ricevere senza problemi i pacchetti del protocollo di routing multicast link-local, chiamati comunemente pacchetti "Hello", in quanto questi pacchetti vengono inoltrati a tutte le porte della VLAN del vPC (modalità flooding). Tuttavia, supponiamo di avere uno scenario in cui un pacchetto del protocollo di routing unicast originato dal Router e destinato a N9K-1 viene trasmesso sull'interfaccia Ethernet1/2 indirizzato a N9K-2 in seguito alla decisione di hashing del port-channel di Layer 2 del Router. Questo pacchetto è destinato all'indirizzo MAC SVI di N9K-1, ma è in entrata nell'interfaccia Ethernet1/1 di N9K-2. N9K-2 vede che il pacchetto è destinato all'indirizzo MAC SVI di N9K-1, che viene installato nella tabella degli indirizzi MAC di N9K-2 con il flag "G", o "Gateway", a causa del miglioramento di vPC Peer Gateway abilitato. Di conseguenza, N9K-2 tenta di instradare localmente il pacchetto del protocollo di routing unicast per conto di N9K-1.

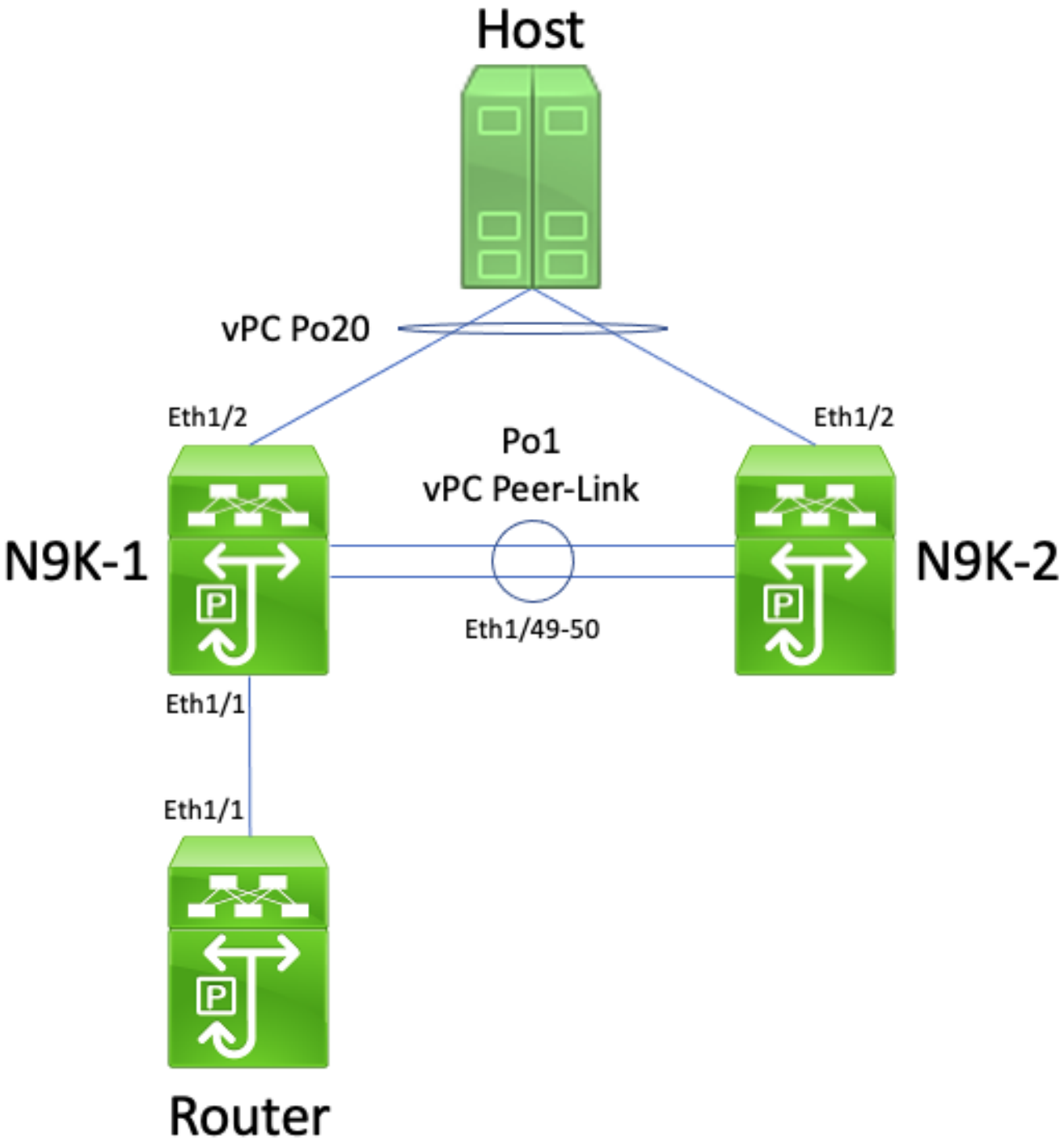
Tuttavia, instradando il pacchetto, il valore TTL (Time to Live) del pacchetto viene diminuito e il valore TTL della maggior parte dei pacchetti del protocollo di routing unicast è 1. Di conseguenza,

il valore TTL del pacchetto viene diminuito a 0 e scartato da N9K-2. Dal punto di vista di N9K-1, N9K-1 riceve pacchetti del protocollo di routing multicast locale rispetto al collegamento dal router ed è in grado di inviare pacchetti del protocollo di routing unicast al router, ma non riceve pacchetti del protocollo di routing unicast dal router. Di conseguenza, N9K-1 elimina l'adiacenza del protocollo di routing con il router e riavvia la macchina a stati finiti locale per il protocollo di routing. Analogamente, il router riavvia la macchina a stati finiti locale per il protocollo di routing.

Per risolvere il problema, è possibile abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer 3 peer-router** nel dominio vPC. Ciò permette di inoltrare i pacchetti unicast con TTL pari a 1 sul collegamento vPC Peer-Link senza diminuire il TTL del pacchetto. Le adiacenze del protocollo di routing unicast potranno così formarsi sul vPC o sulla VLAN del vPC senza problemi.

Adiacenze del protocollo di routing unicast su una VLAN del vPC senza vPC Peer Gateway

Supponiamo di avere questa topologia:



In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway non è abilitata. Po1 è l'interfaccia del collegamento tra i peer vPC, o vPC Peer-Link. Un router con un nome host di Router è collegato tramite Ethernet1/1 a Ethernet1/1 di N9K-1. L'interfaccia Ethernet1/1 del router è un'interfaccia di routing attivata tramite un protocollo di routing unicast. N9K-1 e N9K-2 hanno interfacce SVI attivate con lo stesso protocollo di routing unicast e si trovano nello stesso dominio di broadcast del Router.

Le adiacenze del protocollo di routing unicast su una VLAN del vPC senza funzionalità vPC Peer Gateway abilitata non sono supportate perché la decisione di hashing ECMP del router connesso alla VLAN del vPC può indurre N9K-2 a eliminare il traffico del piano dati per violazione della regola vPC Loop Avoidance. In questa topologia, le adiacenze tra i protocolli di routing si formerebbero correttamente tra i router N9K-1 e N9K-2. Prendere in considerazione il flusso del traffico tra il router e l'host. Il traffico del piano dati indirizzato all'Host attraverso il Router può essere riscritto con un indirizzo MAC di destinazione che appartiene all'indirizzo MAC della SVI di

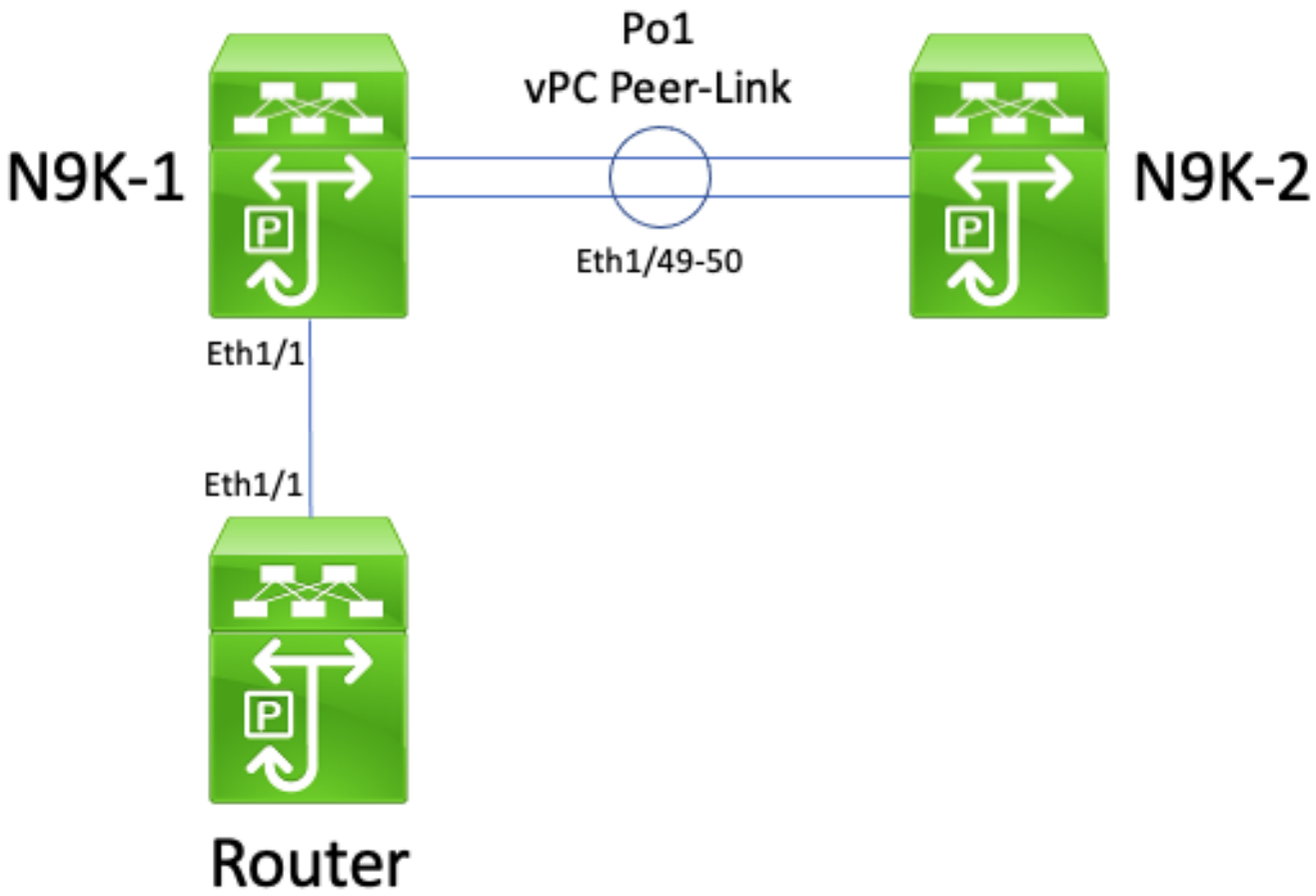
N9K-2 (a causa della decisione di hashing ECMP presa dal router) ma che viene trasmesso sull'interfaccia Ethernet1/1 diretto a N9K-1.

N9K-1 riceve questo pacchetto e lo inoltra attraverso il vPC Peer-Link, poiché l'indirizzo MAC di destinazione appartiene a N9K-2 e il miglioramento di vPC Peer Gateway (che consente a N9K-1 di instradare il pacchetto per conto di N9K-2) non è abilitato. N9K-2 riceve questo pacchetto sul vPC Peer-Link e riconosce che avrebbe dovuto inoltrarlo fuori dalla sua rete Ethernet1/2 in vPC Po20. In questo modo si viola la regola vPC Loop Avoidance, quindi N9K-2 scarta il pacchetto nell'hardware. Ne derivano problemi di connettività o perdita di pacchetti su alcuni flussi che attraversano il dominio vPC.

Per risolvere il problema, abilitare la funzionalità vPC Peer Gateway con il comando di configurazione **peer-gateway** nel dominio vPC, quindi abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer3 peer-router** nel dominio vPC. Per ridurre al minimo i problemi di rete, abilitare entrambe le funzionalità vPC in rapida successione in modo da evitare lo scenario di errore descritto in Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway.

Adiacenze del protocollo di routing unicast su una VLAN del vPC con vPC Peer Gateway

Supponiamo di avere questa topologia:



In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway è abilitata. Po1 è l'interfaccia del collegamento tra i peer vPC, o vPC Peer-Link. Un router con un nome host di Router è collegato tramite Ethernet1/1 a Ethernet1/1 di N9K-1. L'interfaccia Ethernet1/1 del router è un'interfaccia di routing attivata tramite un protocollo di routing unicast. N9K-1 e N9K-2 hanno interfacce SVI attivate con lo stesso

protocollo di routing unicast e si trovano nello stesso dominio di broadcast del Router.

Le adiacenze del protocollo di routing unicast su una VLAN vPC con il miglioramento di vPC Peer Gateway abilitato non sono supportate perché il miglioramento di vPC Peer Gateway impedisce la formazione di adiacenze del protocollo di routing unicast tra il router connesso alla VLAN vPC e il peer vPC a cui il router connesso alla VLAN vPC non è connesso direttamente. In questa topologia, una adiacenza del protocollo di routing tra il router e l'N9K-2 non riesce a raggiungere il risultato previsto, in quanto i pacchetti del protocollo di routing unicast di routing N9K-1 destinati all'indirizzo MAC SVI dell'N9K-2 sono stati abilitati per il miglioramento di vPC Peer Gateway. Poiché i pacchetti vengono indirizzati, il valore Time To Live (TTL) viene ridotto. In genere il TTL dei pacchetti del protocollo di routing unicast è 1 e un TTL sceso a 0 induce il router a eliminarlo.

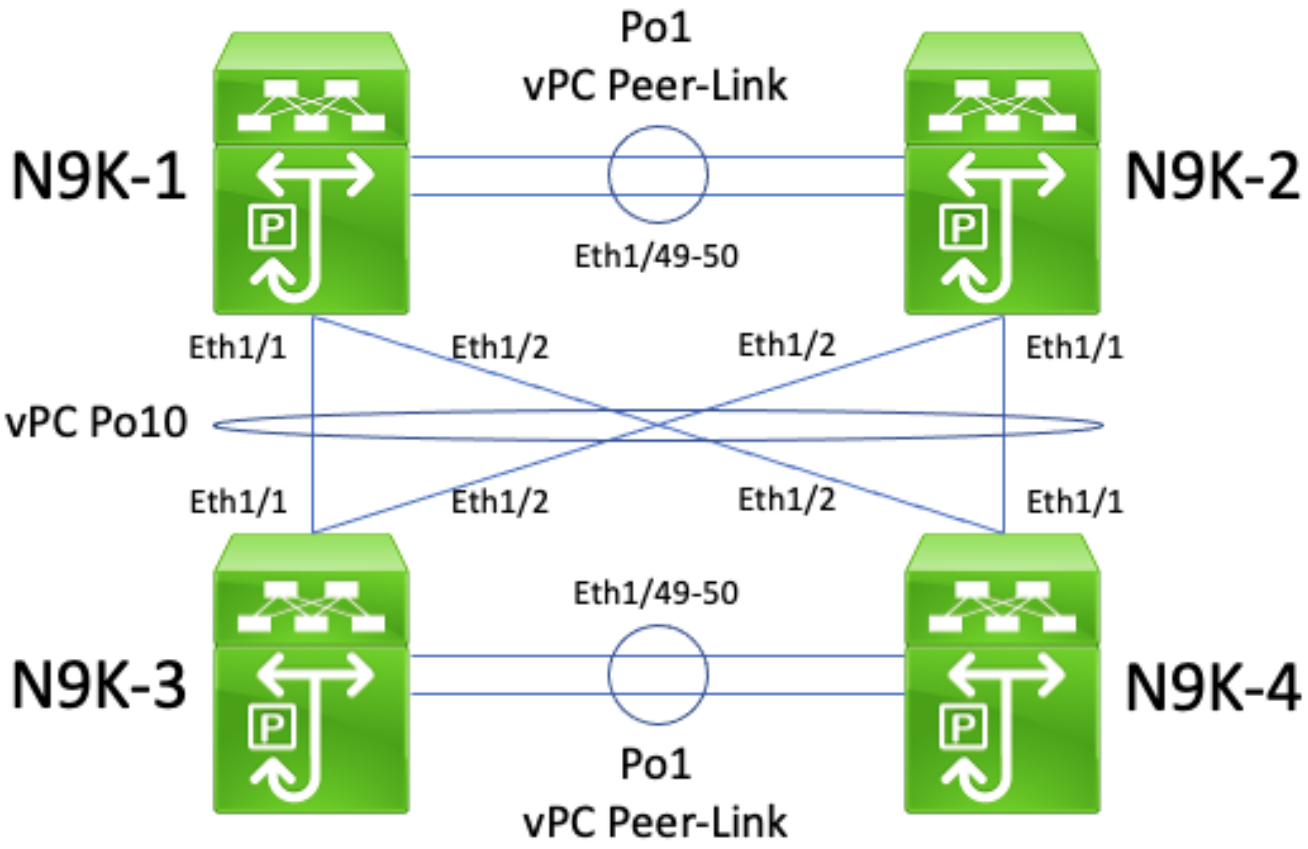
Tutti i router sono in grado di inviare e ricevere senza problemi i pacchetti del protocollo di routing multicast link-local, chiamati comunemente pacchetti "Hello", in quanto questi pacchetti vengono inoltrati a tutte le porte della VLAN del vPC (modalità flooding). Tuttavia, si consideri uno scenario in cui un pacchetto del protocollo di routing unicast inviato dal router destinato all'N9K-2 esce da Ethernet1/1 verso N9K-1. Questo pacchetto è destinato all'indirizzo MAC SVI di N9K-2, ma è in entrata nell'interfaccia Ethernet1/1 di N9K-1. N9K-1 vede che il pacchetto è destinato all'indirizzo MAC SVI di N9K-2, che viene installato nella tabella degli indirizzi MAC di N9K-1 con il flag "G", o "Gateway", a causa del miglioramento di vPC Peer Gateway abilitato. Di conseguenza, N9K-1 tenta di instradare localmente il pacchetto del protocollo di routing unicast per conto di N9K-2.

Tuttavia, instradando il pacchetto, il valore TTL del pacchetto viene diminuito e il valore TTL della maggior parte dei pacchetti del protocollo di routing unicast è 1. Di conseguenza, il valore TTL del pacchetto viene diminuito a 0 e scartato da N9K-1. Dal punto di vista di N9K-2, N9K-2 riceve pacchetti del protocollo di routing multicast locale rispetto al collegamento dal router ed è in grado di inviare pacchetti del protocollo di routing unicast al router, ma non riceve pacchetti del protocollo di routing unicast dal router. Di conseguenza, N9K-2 elimina l'adiacenza del protocollo di routing con il router e riavvia la macchina a stati finiti locale per il protocollo di routing. Analogamente, il router riavvia la macchina a stati finiti locale per il protocollo di routing.

Per risolvere il problema, è possibile abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer 3 peer-router** nel dominio vPC. Ciò permette di inoltrare i pacchetti unicast con TTL pari a 1 sul collegamento vPC Peer-Link senza diminuire il TTL del pacchetto. Le adiacenze del protocollo di routing unicast potranno così formarsi sul vPC o sulla VLAN del vPC senza problemi.

Adiacenze del protocollo di routing unicast su un vPC back-to-back con vPC Peer Gateway

Supponiamo di avere questa topologia:



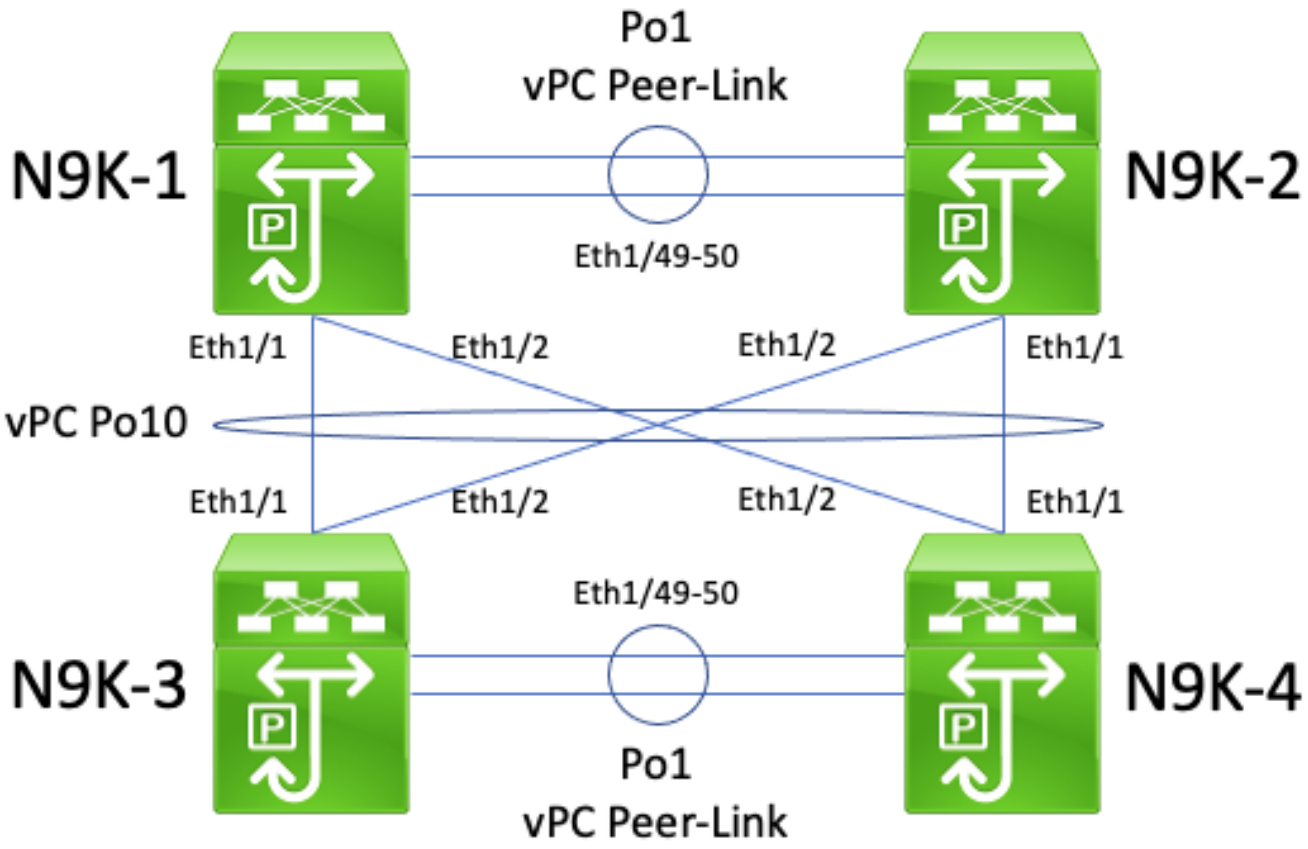
In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway è abilitata. Gli switch Nexus N9K-3 e N9K-4 sono i peer di un dominio vPC in cui la funzionalità migliorata vPC Peer Gateway è abilitata. Entrambi i domini vPC sono connessi tra loro tramite un vPC Po10 back-to-back. I quattro switch hanno interfacce SVI attivate con un protocollo di routing unicast e si trovano nello stesso dominio di broadcast.

Le adiacenze del protocollo di routing unicast su un vPC back-to-back con vPC Peer Gateway abilitata non sono supportate perché la funzionalità vPC Peer Gateway può impedire la formazione di adiacenze del protocollo di routing unicast tra un dominio vPC e l'altro. In questa topologia, un protocollo di routing adiacente tra N9K-1 e N9K-3 o N9K-4 (o entrambi) può non venire come previsto. Analogamente, è possibile che l'adiacenza del protocollo di routing tra N9K-2 e N9K-3 o N9K-4 (o entrambi) non si formi come previsto. Ciò perché i pacchetti del protocollo di routing unicast potrebbero essere destinati a un router (ad esempio, N9K-3) ma inoltrati a un router diverso (ad esempio, N9K-4) in base alla decisione di hashing del port-channel Layer 2 del router di origine.

La causa profonda di questo problema è già stata descritta nella sezione Adiacenze del protocollo di routing unicast su un vPC con vPC Peer Gateway in questo documento. Per risolvere il problema, è possibile abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer 3 peer-router** nel dominio vPC. Ciò permette di inoltrare i pacchetti unicast con TTL pari a 1 sul collegamento vPC Peer-Link senza diminuire il TTL del pacchetto. Le adiacenze del protocollo di routing unicast potranno così formarsi sul vPC back-to-back senza problemi.

Adiacenze OSPF sul vPC con vPC Peer Gateway abilitata e prefisso presente nel database OSPF LSDB ma non nella tabella di routing

Supponiamo di avere questa topologia:



In questa topologia, gli switch Nexus N9K-1 e N9K-2 sono i peer di un dominio vPC in cui la funzionalità vPC Peer Gateway è abilitata. Gli switch Nexus N9K-3 e N9K-4 sono i peer di un dominio vPC in cui la funzionalità migliorata vPC Peer Gateway è abilitata. Entrambi i domini vPC sono connessi tra loro tramite un vPC Po10 back-to-back. I quattro switch hanno interfacce SVI attivate con un protocollo di routing unicast e si trovano nello stesso dominio di broadcast. N9K-4 è il router OSPF designato, o DR (Designated Router), per il dominio di broadcast, mentre N9K-3 è il router OSPF designato di backup, o BDR (Backup Designated Router), del dominio di broadcast.

In questo scenario, un'adiacenza OSPF tra N9K-1 e N9K-3 passa allo stato FULL perché i pacchetti OSPF unicast vengono trasmessi sull'interfaccia Ethernet1/1 di entrambi gli switch. Analogamente, un'adiacenza OSPF tra N9K-2 e N9K-3 passa allo stato FULL perché i pacchetti OSPF unicast vengono trasmessi sull'interfaccia Ethernet1/2 di entrambi gli switch.

Tuttavia, un'adiacenza OSPF tra N9K-1 e N9K-4 è bloccata nello stato EXSTART o EXCHANGE in quanto i pacchetti OSPF unicast escono dall'interfaccia Ethernet1/1 di entrambi gli switch e vengono eliminati da N9K-2 e N9K-4 come descritto nella sezione Adiacenze del protocollo di routing unicast su vPC back-to-back con vPC Peer Gateway in questo documento. Analogamente, un'adiacenza OSPF tra N9K-2 e N9K-4 è bloccata nello stato EXSTART o EXCHANGE in quanto i pacchetti OSPF unicast escono dall'interfaccia Ethernet1/2 di entrambi gli switch e vengono eliminati da N9K-1 e N9K-3 come descritto nella sezione Adiacenze del protocollo di routing unicast su vPc back-to-back con vPC Peer Gateway in questo documento.

Di conseguenza, N9K-1 e N9K-2 sono nello stato FULL per il BDR del dominio di broadcast, ma sono nello stato EXSTART o EXCHANGE per il DR del dominio di broadcast. Sia il DR che il BDR di un dominio di broadcast conservano una copia completa del database OSPF LSDB (Link State Data Base), ma i router OSPF DROTHER devono avere lo stato FULL per il router designato del dominio di broadcast per installare i prefissi acquisiti tramite OSPF dal router designato o dal router designato di backup. Di conseguenza, sia N9K-1 che N9K-2 sembrano avere prefissi

appresi da N9K-3 e N9K-4 presenti in OSPF LSDB, ma tali prefissi non vengono installati nella tabella di routing unicast fino a quando N9K-1 e N9K-2 non passano a uno stato FULL con N9K-4 (il DR per il dominio broadcast).

Per risolvere il problema, è possibile abilitare la funzionalità Routing/Layer 3 over vPC con il comando di configurazione **layer 3 peer-router** nel dominio vPC. Ciò permette di inoltrare i pacchetti unicast con TTL pari a 1 sul collegamento vPC Peer-Link senza diminuire il TTL del pacchetto. Le adiacenze del protocollo di routing unicast potranno così formarsi sul vPC back-to-back senza problemi. Di conseguenza, N9K-1 e N9K-2 passano a uno stato FULL con N9K-4 (il DR per il dominio di trasmissione) e installano correttamente i prefissi appresi da N9K-3 e N9K-4 tramite OSPF nelle rispettive tabelle di routing unicast.

Informazioni correlate

- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 10.1(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 9.3(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 9.2(x)
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 9000, versione 7.x
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 7000, versione 8.x
- Guida alla configurazione delle interfacce NX-OS sui Cisco Nexus serie 7000, versione 7.x
- Guida alla progettazione e configurazione: best practice per i canali delle porte virtuali (vPC) sugli switch Cisco Nexus serie 7000
- Topologie supportate per il routing su vPC (Virtual Port Channel) sulle piattaforme Nexus

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l’accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).