Comprendre les améliorations du canal de port virtuel (vPC)

Table des matières

Introduction

Ce document décrit les améliorations courantes du Virtual Port Channel (vPC) configurées sur les commutateurs Cisco Nexus dans un domaine vPC.

Conditions préalables

Exigences

Cisco vous conseille de comprendre les informations de base relatives au scénario, à la configuration et à la mise en œuvre du canal de port virtuel (vPC). Pour plus d'informations sur cette fonctionnalité, consultez l'un des documents applicables suivants :

- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.3(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.2(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.1(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 9.3(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 9.2(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 7.x
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 7000 8.x
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 7000 7.x
- Guide de conception et de configuration : Bonnes pratiques pour les canaux de port virtuel (vPC) sur les commutateurs Cisco Nexus 7000

Composants utilisés

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l'incidence possible des commandes.

Informations générales

Depuis l'intégration de Cisco NX-OS sur les commutateurs de centre de données Cisco Nexus, la

fonctionnalité de canaux de ports virtuels (vPC) a bénéficié de nombreuses améliorations qui améliorent la fiabilité des périphériques connectés au vPC pendant les scénarios de défaillance et optimisent le comportement de transfert des deux commutateurs homologues vPC. Le fait de comprendre l'objectif de chaque amélioration, le changement de comportement introduit par l'amélioration et les scénarios de défaillance que l'amélioration résout peut vous aider à comprendre pourquoi et quand une amélioration doit être configurée dans un domaine vPC pour mieux répondre aux besoins et aux exigences de l'entreprise.

Matériel applicable

La procédure décrite dans ce document s'applique à tous les commutateurs de centre de données Cisco Nexus compatibles avec vPC.

Commutateur homologue vPC

Cette section décrit l'amélioration du commutateur homologue vPC, qui est activée avec la commande de configuration de domaine vPC peer-switch.

Aperçu

Dans de nombreux environnements, une paire de commutateurs Nexus dans un domaine vPC sont des commutateurs d'agrégation ou principaux servant de frontière entre les domaines Ethernet commutés de couche 2 et les domaines de routage de couche 3. Les deux commutateurs sont configurés avec plusieurs VLAN et sont responsables du routage du trafic est-ouest entre les VLAN ainsi que du trafic nord-sud. Dans ces environnements, les commutateurs Nexus agissent également comme des ponts racine du point de vue du protocole Spanning Tree.

Normalement, un homologue vPC est configuré comme pont racine du Spanning Tree en réglant sa priorité Spanning Tree à une valeur faible, telle que 0. L'autre homologue vPC est configuré avec une priorité Spanning Tree légèrement plus élevée, telle que 4096, ce qui lui permet de jouer le rôle de pont racine dans le Spanning Tree si l'homologue vPC agissant en tant que pont racine tombe en panne. Avec cette configuration, l'homologue vPC agissant en tant que pont racine crée des BPDU (Bridge Protocol Data Unit) Spanning Tree avec un ID de pont contenant son adresse MAC système.

Cependant, si l'homologue vPC agissant comme pont racine échoue et fait en sorte que l'autre homologue vPC prenne le relais en tant que pont racine Spanning Tree, l'autre homologue vPC émet des BPDU Spanning Tree avec un ID de pont contenant son adresse MAC système, qui est différente de l'adresse MAC système du pont racine d'origine. Selon le mode de connexion des ponts en aval, l'impact de ce changement varie et est décrit dans les sous-sections suivantes.

Ponts non vPC connectés de manière redondante

Les ponts non connectés au vPC qui sont connectés aux deux homologues vPC avec des liaisons redondantes (de sorte qu'une liaison est dans un état de blocage du point de vue du protocole Spanning Tree) qui détectent la modification de la trame BPDU (et, par conséquent, la

modification du pont racine) observent une modification du port racine. D'autres interfaces de transfert désigné passent immédiatement à l'état Blocage, puis traversent la machine à états finis du protocole Spanning Tree (Blocage, Apprentissage et Transfert) avec des pauses entre elles équivalentes au temporisateur de délai de transfert du protocole Spanning Tree configuré (15 secondes par défaut).

Le changement de port racine et la traversée ultérieure de l'automate fini du protocole Spanning Tree peuvent provoquer une perturbation importante au sein du réseau. L'amélioration du commutateur homologue vPC a été introduite principalement pour éviter les perturbations du réseau causées par ce problème si l'un des homologues vPC venait à se déconnecter. Avec l'amélioration du commutateur homologue vPC, le pont non connecté au vPC possède toujours une liaison redondante unique qui est dans un état de blocage, mais passe immédiatement cette interface à un état de transfert si le port racine existant tombe en panne en raison d'une défaillance de liaison. Le même processus se produit lorsque l'homologue vPC hors connexion se reconnecte : l'interface présentant le coût le plus faible pour le pont racine prend le rôle de port racine et la liaison redondante passe immédiatement à l'état de blocage. Le seul impact observé sur le plan de données est la perte inévitable de paquets en vol qui traversaient l'homologue vPC lors de sa mise hors ligne.

Ponts connectés au vPC

Les ponts connectés au vPC dans le domaine Spanning Tree détectent la modification de la trame BPDU (et, par conséquent, la modification du pont racine) et vident les adresses MAC apprises dynamiquement de leurs tables d'adresses MAC locales. Ce comportement est inefficace et inutile dans les topologies avec des périphériques connectés au protocole Spanning Tree qui ne dépendent pas du protocole Spanning Tree pour une topologie sans boucle. Les vPC sont considérés comme une interface logique unique du point de vue du protocole Spanning Tree, tout comme les canaux de port normaux, de sorte que la perte d'un homologue vPC est similaire à la perte d'une liaison unique au sein d'un membre de canal de port. Dans les deux cas, le spanning tree ne change pas, de sorte que la purge des adresses MAC apprises dynamiquement à partir des ponts dans le domaine du spanning tree (dont le but est de permettre au comportement d'inondation et de détection d'Ethernet de réapprendre les adresses MAC sur les interfaces du spanning tree qui commencent à transférer) n'est pas nécessaire.

En outre, la purge des adresses MAC apprises dynamiquement pourrait potentiellement être perturbatrice. Voici un scénario dans lequel deux hôtes ont un flux de type UDP en grande partie unidirectionnel (comme un client TFTP envoyant des données à un serveur TFTP). Dans ce flux, les données circulent principalement du client TFTP vers le serveur TFTP. Le serveur TFTP envoie rarement un paquet de retour vers le client TFTP. Par conséquent, après un vidage d'adresses MAC apprises dynamiquement dans le domaine Spanning Tree, l'adresse MAC du serveur TFTP n'est pas apprise pendant un certain temps. Cela signifie que les données du client TFTP envoyées vers le serveur TFTP sont diffusées à travers le VLAN, car le trafic est du trafic unicast inconnu. Cela peut entraîner le déplacement de flux de données volumineux vers des emplacements inattendus au sein du réseau et peut entraîner des problèmes de performances s'ils traversent des sections surabonnées du réseau.

L'amélioration du commutateur homologue vPC a été introduite pour éviter que ce comportement

inefficace et inutile ne se produise au cas où l'homologue vPC agissant en tant que pont racine Spanning Tree pour un ou plusieurs VLAN est rechargé ou mis hors tension.

Pour activer l'amélioration du commutateur d'homologue vPC, les deux homologues vPC doivent avoir la même configuration de protocole Spanning Tree (y compris les valeurs de priorité Spanning Tree pour tous les VLAN vPC) et être le pont racine pour tous les VLAN vPC. Une fois que ces conditions préalables sont remplies, la commande de configuration de domaine vPC peer-switch doit être configurée pour activer l’amélioration du commutateur homologue vPC.

Remarque : l'amélioration du commutateur homologue vPC est uniquement prise en charge sur un domaine vPC qui contient la racine pour tous les VLAN.

Une fois l'amélioration du commutateur homologue vPC activée, les deux homologues vPC commencent à émettre des BPDU Spanning Tree identiques avec un ID de pont contenant l'adresse MAC système vPC partagée par les deux homologues vPC. Si un homologue vPC est rechargé, la trame BPDU Spanning Tree qui est émise par l'homologue vPC restant ne change pas, de sorte que les autres ponts dans le domaine Spanning Tree ne voient aucune modification dans le pont racine et ne réagissent pas de manière sous-optimale à la modification dans le réseau.

Mises en garde

L’amélioration du commutateur homologue vPC comporte certaines mises en garde que vous devez connaître avant de la configurer dans un environnement de production.

Les valeurs de priorité du Spanning Tree doivent correspondre entre les homologues vPC

Avant d’activer l’amélioration du commutateur homologue vPC, la configuration de priorité Spanning Tree pour tous les VLAN vPC doit être modifiée afin qu’elle soit identique entre les deux homologues vPC.

Examinez la configuration suivante, où N9K-1 est configuré comme étant le pont racine Spanning Tree pour les VLAN 1, 10 et 20 avec une priorité de 0. N9K-2 est le pont racine Spanning Tree secondaire pour les VLAN 1, 10 et 20 avec une priorité de 4096.

<#root>

```
N9K-1#

show running-config spanning-tree

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

N9K-2#

show running-config spanning-tree

spanning-tree vlan 1,10,20 priority 4096
```

```
interface port-channel1
  spanning-tree port type network
```

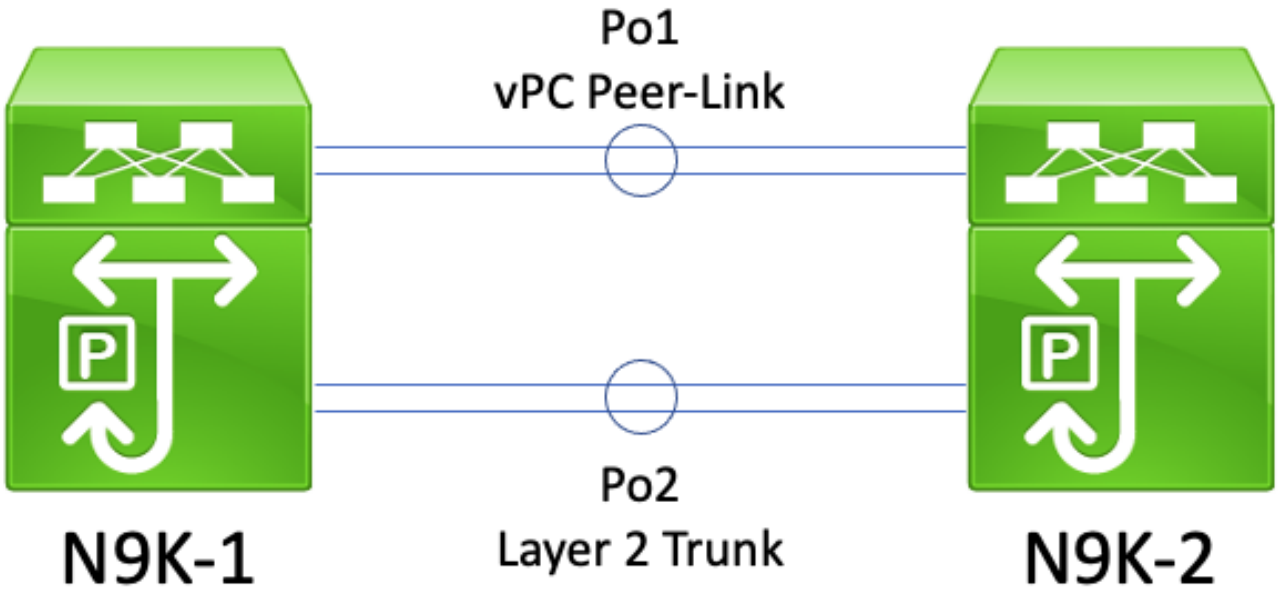
Avant d'activer l'amélioration du commutateur homologue vPC, vous devez modifier la configuration de priorité Spanning Tree pour les VLAN 1, 10 et 20 sur N9K-2 afin qu'elle corresponde à la configuration de priorité Spanning Tree pour les mêmes VLAN sur N9K-1. Un exemple de cette modification est présenté ici.

<#root>

```
N9K-2#

configure terminal

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

spanning-tree vlan 1,10,20 priority 0

N9K-2(config)#

end

N9K-2#

show running-config spanning-tree

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

N9K-1#

show running-config spanning-tree

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

Effet du commutateur homologue vPC sur les VLAN non vPC

Examinez la topologie suivante :



Dans cette topologie, deux homologues vPC (N9K-1 et N9K-2) sont séparés par deux lignes principales de couche 2 : Po1 et Po2. Po1 est le lien homologue vPC transportant des VLAN vPC, tandis que Po2 est une ligne principale de couche 2 transportant tous les VLAN non vPC. Si les valeurs de priorité Spanning Tree pour les VLAN non-vPC transportés sur Po2 sont identiques sur N9K-1 et N9K-2, alors chaque homologue vPC émet des trames BPDU Spanning Tree provenant de l'adresse MAC système vPC, qui est identique sur les deux commutateurs. Par conséquent, N9K-1 semble recevoir sa propre BPDU Spanning Tree sur Po2 pour chaque VLAN non vPC, même si N9K-2 est le commutateur à l'origine de la BPDU Spanning Tree. Du point de vue du Spanning Tree, N9K-1 place Po2 dans un état de blocage pour tous les VLAN non-vPC.

C’est un comportement attendu. Pour éviter ce comportement ou pour contourner ce problème, les deux homologues vPC doivent être configurés avec des valeurs de priorité Spanning Tree différentes sur tous les VLAN non vPC. Cela permet à un homologue vPC de devenir le pont racine pour le VLAN non-vPC et de faire passer l'agrégation de couche 2 entre les homologues vPC à un état de transfert désigné. De même, l'homologue vPC distant fait passer l'agrégation de couche 2 entre les homologues vPC à un état Racine désignée. Cela permet au trafic des VLAN non vPC de circuler entre les deux homologues vPC via l'agrégation de couche 2.

Configuration

Vous trouverez ici un exemple de configuration de la fonctionnalité de commutateur homologue vPC.

Dans cet exemple, N9K-1 est configuré pour être le pont racine Spanning Tree pour les VLAN 1, 10 et 20 avec une priorité de 0. N9K-2 est le pont racine Spanning Tree secondaire pour les VLAN 1, 10 et 20 avec une priorité de 4096.

<#root>

N9K-1#

```
show running-config vpc

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

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

<snip>
vpc domain 1
  peer-keepalive destination 10.122.190.195

interface port-channel1
  vpc peer-link

N9K-1#

show running-config spanning-tree

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

N9K-2#

show running-config spanning-tree

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

Tout d'abord, la configuration de priorité Spanning Tree de N9K-2 doit être modifiée pour être identique à celle de N9K-1. Cela est obligatoire pour que la fonctionnalité de commutateur homologue vPC fonctionne comme prévu. Si l'adresse MAC système de N9K-2 est inférieure à l'adresse MAC système de N9K-1, alors N9K-2 usurpe le rôle de pont racine pour le domaine Spanning Tree, ce qui amène les autres ponts du domaine Spanning Tree à vider leurs tables d'adresses MAC locales pour tous les VLAN affectés. Un exemple de ce phénomène est présenté ici.

<#root>

```
N9K-1#

show spanning-tree vlan 1


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

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

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

N9K-2#
```

show spanning-tree vlan 1

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

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

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

N9K-2#
```

configure terminal

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

spanning-tree vlan 1,10,20 priority 0

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

end

```
N9K-2#
```

show spanning-tree vlan 1

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

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

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

```
Po10            Desg FWD 1         128.4105 (vPC) P2p
Po20            Desg FWD 1         128.4115 (vPC) P2p
```

Ensuite, nous pouvons activer la fonctionnalité de commutateur homologue vPC à l'aide de la commande de configuration de domaine vPC peer-switch. Cela modifie l'ID de pont dans les unités BPDU Spanning Tree émises par les deux homologues vPC, ce qui entraîne le vidage des tables d'adresses MAC locales de tous les VLAN affectés par les autres ponts du domaine Spanning Tree.

```
<#root>

N9K-1#

configure terminal

N9K-1(config)#

vpc domain 1

N9K-1(config-vpc-domain)#

peer-switch

N9K-1(config-vpc-domain)#

end

N9K-1#

N9K-2#

configure terminal

N9K-2(config)#

vpc domain 1

N9K-2(config-vpc-domain)#

peer-switch

N9K-2(config-vpc-domain)#

end

N9K-2#
```

Vous pouvez vérifier que la fonctionnalité de commutateur homologue vPC fonctionne comme prévu en validant les revendications des deux homologues vPC comme étant le pont racine pour les VLAN vPC à l'aide de la commande show spanning-tree summary. Cette sortie devrait également indiquer que la fonctionnalité de commutateur homologue vPC est activée et opérationnelle.

```
<#root>

N9K-1#
```

```
show spanning-tree summary

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

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

N9K-2#
show spanning-tree summary

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

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

Utilisez la commande show spanning-tree vlan{x} pour afficher des informations plus détaillées sur un VLAN spécifique. Toutes les interfaces du commutateur doté du rôle vPC principal ou principal opérationnel sont à l'état de transfert désigné. Le commutateur qui détient le rôle secondaire ou secondaire opérationnel du vPC a toutes ses interfaces dans un état de transfert désigné, à l'exception de la liaison entre homologues vPC, qui est dans un état de transfert racine. Notez que l’adresse MAC du système vPC affichée dans la sortie de show vpc role est identique à l’ID de pont racine et à l’ID de pont de chaque homologue vPC.

<#root>

```
N9K-1#
```

show vpc role

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

N9K-1#
```

show spanning-tree vlan 1

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

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

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

N9K-2#
```

show vpc role

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

N9K-2#
```

show spanning-tree vlan 1

```
VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    1
```

```
             Address     0023.04ee.be01
             This bridge is the root
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

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

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

Enfin, nous pouvons utiliser [l'utilitaire de saisie de paquets du plan de commande Ethanalyzer] sur l'un ou l'autre des homologues vPC pour confirmer que les deux homologues vPC sont à l'origine des BPDU Spanning Tree avec un ID de pont et un ID de pont racine contenant l'adresse MAC du système vPC partagé entre les deux homologues vPC.

<#root>

```
N9K-1#

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

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

N9K-2#

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

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

Incidence

L'impact de l'activation de l'amélioration du commutateur homologue vPC varie selon que les autres ponts du domaine Spanning Tree sont connectés aux deux homologues vPC via un vPC ou s'ils sont connectés de manière redondante aux deux homologues vPC sans vPC.

Ponts non vPC connectés de manière redondante

Si un pont non connecté au vPC avec des liens redondants vers les deux homologues vPC (de sorte qu'un lien est dans un état bloquant du point de vue du protocole Spanning Tree) détecte un changement dans le pont racine Spanning Tree annoncé dans les BPDU Spanning Tree, le port racine du pont peut changer entre les deux interfaces redondantes. Cela peut alors amener d'autres interfaces de transfert désigné à passer immédiatement à l'état bloquant, puis à traverser l'automate fini du protocole Spanning Tree (blocage, apprentissage et transfert) avec des pauses

équivalentes à la minuterie de délai de transfert du protocole Spanning Tree configurée (15 secondes par défaut). Le changement de port racine et la traversée ultérieure de l'automate fini du protocole Spanning Tree peuvent provoquer une perturbation importante au sein du réseau.

Il est utile de mentionner que cet impact se produit lorsque l'homologue vPC qui est actuellement le pont racine pour le domaine Spanning Tree passe hors ligne (par exemple en cas de panne d'alimentation, de panne matérielle ou de rechargement). Ce comportement n'est pas spécifique à l'amélioration du commutateur homologue vPC : l'activation de l'amélioration du commutateur homologue vPC entraîne simplement un comportement similaire à celui d'un homologue vPC qui se déconnecte du point de vue du spanning tree.

Ponts connectés au vPC

Si un pont connecté à un vPC détecte une modification dans le pont racine Spanning Tree annoncé dans les unités BPDU Spanning Tree, le pont vide les adresses MAC apprises dynamiquement de sa table d'adresses MAC. Lors de la configuration de la fonctionnalité de commutateur homologue vPC, vous pouvez observer ce comportement dans les deux scénarios suivants :

1. Lorsque les valeurs de priorité du Spanning Tree sont configurées pour correspondre aux deux homologues vPC, le pont racine du Spanning Tree peut passer d'un homologue vPC à un autre si l'homologue vPC qui n'était pas le pont racine auparavant a une adresse MAC système inférieure à celle de l'homologue vPC qui était auparavant le pont racine. Un exemple de ce scénario est présenté dans la section Configuration du commutateur homologue vPC de ce document.
2. Lorsque la fonctionnalité Commutateur d'homologue vPC est activée via la commande de configuration de domaine vPC peer-switch, les deux homologues vPC commencent à fonctionner en tant que ponts racine du domaine Spanning Tree. Les deux homologues vPC commencent à émettre des BPDU Spanning Tree identiques s'affirmant comme le pont racine du domaine Spanning Tree.

Dans la plupart des scénarios et des topologies, aucun impact sur le plan de données n'est observé suite à l'un ou l'autre de ces deux scénarios. Cependant, pendant une courte période, le trafic du plan de données est inondé au sein d'un VLAN en raison d'une inondation de monodiffusion inconnue, car l'adresse MAC de destination des trames n'est pas apprise sur un port de commutateur en conséquence directe du vidage des adresses MAC apprises dynamiquement. Dans certaines topologies, cela peut provoquer des problèmes de performances ou des pertes de paquets pendant de courtes périodes si le trafic du plan de données est inondé vers des périphériques réseau surabonnés dans le VLAN. Cela peut également entraîner des problèmes avec les flux de trafic unidirectionnel à forte consommation de bande passante ou les hôtes silencieux (hôtes qui reçoivent principalement des paquets et envoient rarement des paquets), car ce trafic est diffusé dans le VLAN pendant une période prolongée au lieu d'être commuté directement vers l'hôte de destination comme d'habitude.

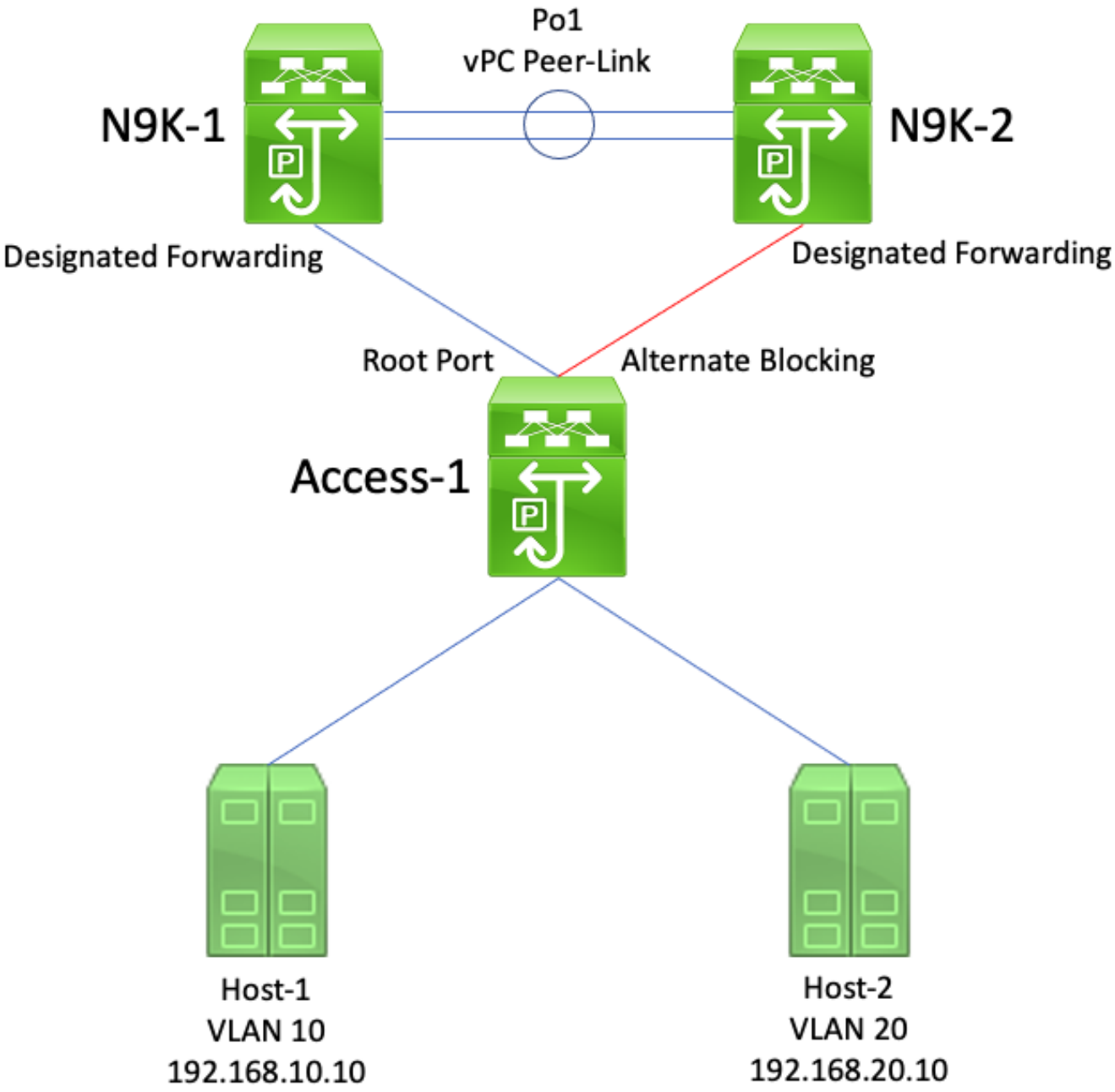
Il est utile de mentionner que cet impact est lié au vidage des adresses MAC apprises dynamiquement à partir de la table d'adresses MAC des ponts dans le VLAN affecté. Ce comportement n'est pas spécifique à l'amélioration du commutateur homologue vPC ou à une

modification du pont racine. Il peut également être provoqué par une notification de changement de topologie générée en raison d'un port non périphérique qui se manifeste dans le VLAN.

Exemples de scénarios de défaillance

Ponts non vPC connectés de manière redondante qui redémarrent l'automate fini

Examinez la topologie suivante :



Dans cette topologie, N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC. N9K-1 est configuré avec une valeur de priorité Spanning Tree de 0 pour tous les VLAN, faisant de N9K-1 le pont racine pour tous les VLAN. N9K-2 est configuré avec une valeur de priorité Spanning Tree de 4096 pour tous les VLAN, faisant de N9K-2 le pont racine secondaire pour tous les VLAN. Accès-1 est un commutateur connecté de manière redondante à N9K-1 et N9K-2 grâce à des

ports de commutateur de couche 2. Ces ports de commutateur ne sont pas regroupés dans un canal de port, de sorte que le protocole Spanning Tree place le lien connecté à N9K-1 dans un état racine désigné et le lien connecté à N9K-2 dans un état bloquant secondaire.

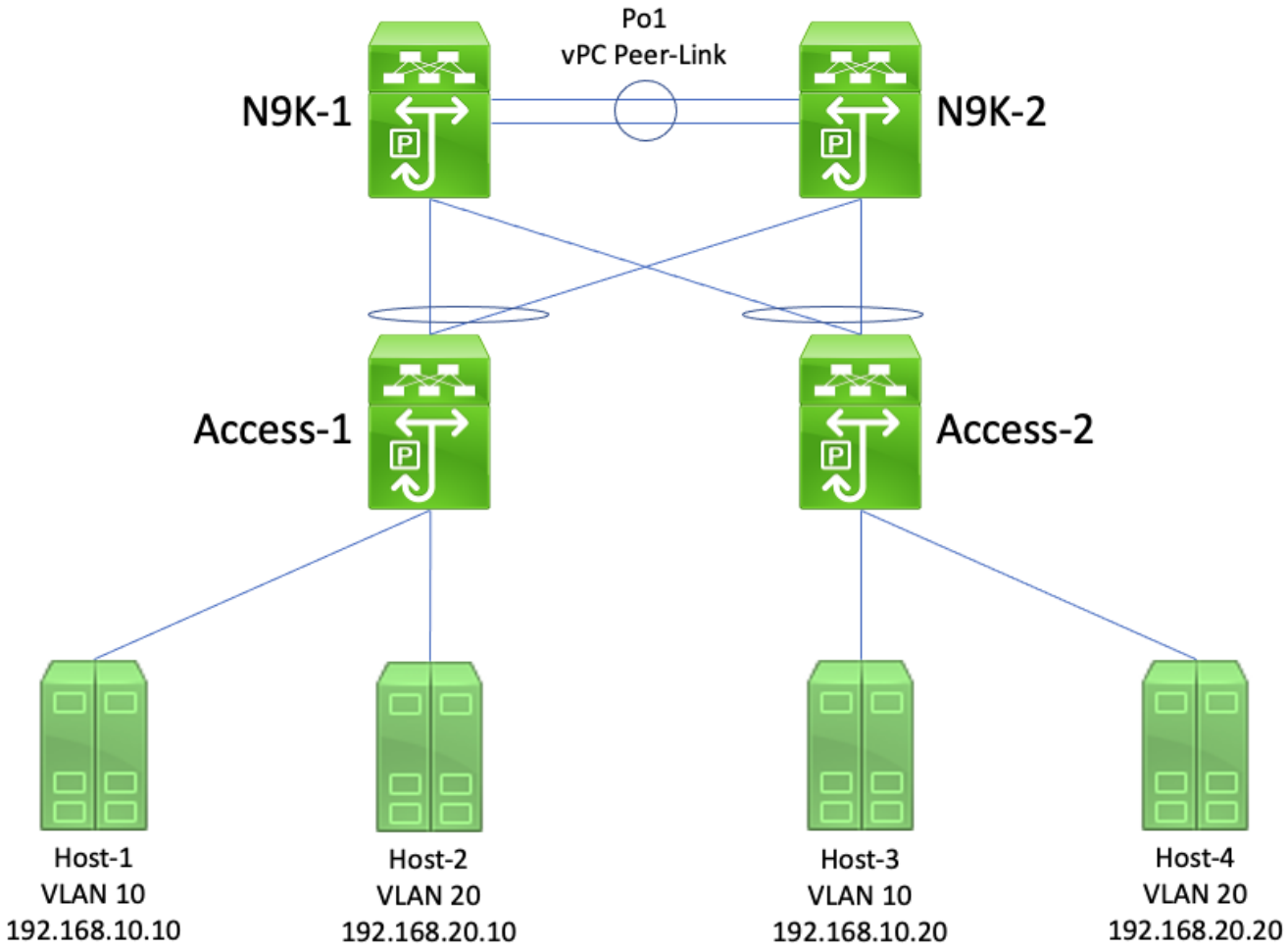
Voici un scénario de défaillance dans lequel N9K-1 passe hors ligne en raison d'une défaillance matérielle, d'une panne de courant ou d'un rechargement du commutateur. N9K-2 s'affirme comme le pont racine pour tous les VLAN en annonçant les BPDU Spanning Tree en utilisant son adresse MAC système comme ID de pont. Access-1 constate une modification de l'ID du pont racine. En outre, son port racine désigné passe à l'état down/down, ce qui signifie que le nouveau port racine désigné est la liaison qui était dans un état de blocage alternatif face à N9K-2.

Ce changement dans les ports racine désignés entraîne tous les ports Spanning Tree non périphériques à traverser la machine à états finis du protocole Spanning Tree (Blocage, Apprentissage et Transfert) avec des pauses entre les deux équivalentes au temporisateur de délai de transfert du protocole Spanning Tree configuré (15 secondes par défaut). Ce processus peut être extrêmement perturbateur pour le réseau.

Dans le même scénario de défaillance avec l'amélioration du commutateur homologue vPC activée, N9K-1 et N9K-2 transmettent des unités BPDU Spanning Tree identiques en utilisant l'adresse MAC système vPC partagée comme ID de pont. Si N9k-1 échoue, N9K-2 continue de transmettre cette même unité BPDU Spanning Tree. Par conséquent, Access-1 fait immédiatement passer la liaison de blocage alternatif vers N9K-2 à un état de racine désignée et commence à transférer le trafic sur la liaison. De plus, le fait que l'ID du pont racine Spanning Tree ne change pas empêche les ports non périphériques de passer par l'automate fini du protocole Spanning Tree, ce qui réduit la quantité de perturbations observées dans le réseau.

Les ponts connectés au vPC purgent les adresses MAC apprises dynamiquement

Examinez la topologie suivante :



Dans cette topologie, N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC qui effectuent le routage inter-VLAN entre le VLAN 10 et le VLAN 20. N9K-1 est configuré avec une valeur de priorité Spanning Tree de 0 pour le VLAN 10 et le VLAN 20, faisant de N9K-1 le pont racine pour les deux VLAN. N9K-2 est configuré avec une valeur de priorité Spanning Tree de 4096 pour le VLAN 10 et le VLAN 20, faisant de N9K-2 le pont racine secondaire pour les deux VLAN. Hôte-1, Hôte-2, Hôte-3 et Hôte-4 communiquent tous en permanence entre eux.

Voici un scénario de défaillance dans lequel N9K-1 passe hors ligne en raison d'une défaillance matérielle, d'une panne de courant ou d'un rechargement du commutateur. N9K-2 s'affirme comme le pont racine pour VLAN 10 et VLAN 20 en annonçant les BPDU Spanning Tree en utilisant son adresse MAC système comme ID de pont. Access-1 et Access-2 voient un changement dans l'ID du pont racine, et bien que le Spanning Tree reste le même (ce qui signifie que le vPC faisant face à N9K-1 et N9K-2 reste un port racine désigné) Access-1 et Access-2 vident leur adresse MAC de toutes les adresses MAC apprises dynamiquement dans VLAN 10 et VLAN 20.

Dans la plupart des environnements, la purge des adresses MAC apprises dynamiquement a un impact minimal. Aucun paquet n'est perdu (à l'exception de ceux qui ont été perdus alors qu'ils étaient transmis à N9K-1 lors de sa défaillance), mais le trafic est temporairement inondé dans chaque domaine de diffusion en tant que trafic de monodiffusion inconnu pendant que tous les commutateurs du domaine de diffusion réapprennent les adresses MAC dynamiques.

Dans le même scénario de défaillance où l'amélioration du commutateur homologue vPC est

activée, N9K-1 et N9K-2 transmettraient des BPDU Spanning Tree identiques en utilisant l’adresse MAC du système vPC partagé comme ID de pont. Si N9k-1 échoue, N9K-2 continue de transmettre cette même unité BPDU Spanning Tree. Par conséquent, Access-1 et Access-2 ignorent qu'une modification de la topologie Spanning Tree a eu lieu. De leur point de vue, les unités BPDU Spanning Tree du pont racine sont identiques, il n'est donc pas nécessaire de vider les adresses MAC apprises dynamiquement à partir des VLAN concernés. Cela empêche l’inondation de trafic de monodiffusion inconnu dans chaque domaine de diffusion dans ce scénario de défaillance.

Passerelle homologue vPC

Cette section décrit l’amélioration de la passerelle homologue vPC, qui est activée avec la commande de configuration de domaine vPC peer-gateway.

Aperçu

Les commutateurs Nexus configurés dans un domaine vPC effectuent par défaut un protocole FHRP (First Hop Redundancy Protocol) double actif. Cela signifie que si l'un des homologues vPC reçoit un paquet avec une adresse MAC de destination appartenant à un groupe HSRP (Hot Standby Router Protocol) ou VRRP (Virtual Router Redundancy Protocol) configuré sur le commutateur, ce dernier achemine le paquet selon sa table de routage locale, quel que soit son état de plan de contrôle HSRP ou VRRP. En d’autres termes, on s’attend à ce qu’un homologue vPC dans un état de veille HSRP ou de secours VRRP achemine les paquets destinés à l’adresse MAC virtuelle HSRP ou VRRP.

Lorsqu'un homologue vPC achemine un paquet destiné à une adresse MAC virtuelle FHRP, il réécrit le paquet avec une nouvelle adresse MAC source et de destination. L'adresse MAC source est l'adresse MAC de l'interface virtuelle commutée (SVI) de l'homologue vPC au sein du VLAN vers lequel le paquet est acheminé. L'adresse MAC de destination est l'adresse MAC associée à l'adresse IP du tronçon suivant pour l'adresse IP de destination du paquet, conformément à la table de routage locale de l'homologue vPC. Dans les scénarios de routage inter-VLAN, l'adresse MAC de destination du paquet après sa réécriture est l'adresse MAC de l'hôte auquel le paquet est finalement destiné.

Certains hôtes ne respectent pas le comportement de transfert standard en tant que fonctionnalité d’optimisation. Avec ce comportement, l’hôte n’effectue pas de recherche dans le tableau de routage ni dans la mémoire cache ARP lorsqu’il répond à un paquet entrant. L’hôte inverse plutôt les adresses MAC source et de destination du paquet entrant pour le paquet de réponse. En d’autres termes, l’adresse MAC source du paquet entrant devient l’adresse MAC de destination du paquet de réponse, et l’adresse MAC de destination du paquet entrant devient l’adresse MAC source du paquet de réponse. Ce comportement diffère de celui d’un hôte qui respecte le comportement de transfert standard, car il effectuerait plutôt une recherche dans le tableau de routage local et/ou dans la mémoire cache ARP et indiquerait l’adresse MAC virtuelle FHRP comme adresse MAC de destination du paquet de réponse.

Ce comportement d’hôte non standard peut enfreindre la règle de prévention de boucle vPC si le

paquet de réponse généré par l'hôte est adressé à un homologue vPC, mais sort du vPC vers l'autre homologue vPC. L'autre homologue vPC reçoit le paquet destiné à une adresse MAC appartenant à son homologue vPC et transfère le paquet de l'homologue Peer-Link vPC vers l'homologue vPC qui possède l'adresse MAC présente dans le champ d'adresse MAC de destination du paquet. L'homologue vPC propriétaire de l'adresse MAC tente d'acheminer le paquet localement. Si le paquet doit sortir d'un vPC, l'homologue vPC abandonne ce paquet pour violation de la règle d'évitement de boucle vPC. Par conséquent, vous pourriez subir des problèmes de connectivité ou des pertes de paquets pour certains flux provenant ou destinés à un hôte utilisant ce comportement non standard.

L'amélioration de la passerelle homologue vPC a été introduite pour éliminer la perte de paquets causée par les hôtes utilisant ce comportement non standard. Pour ce faire, il permet à un homologue vPC d'acheminer localement les paquets destinés à l'adresse MAC de l'autre homologue vPC, de sorte que les paquets destinés à l'homologue vPC distant n'aient pas besoin de sortir du lien homologue vPC pour être acheminés. En d'autres termes, l'amélioration de la passerelle homologue vPC permet à un homologue vPC d'acheminer les paquets « au nom » de l'homologue vPC distant. L'amélioration de la passerelle homologue vPC peut être activée avec la commande de configuration du domaine vPC peer-gateway.

Mises en garde

Oscillation des contiguïtés de protocole de routage de monodiffusion sur les vPC ou les VLAN vPC

Si des contiguïtés de protocole de routage de monodiffusion dynamique sont formées entre deux homologues vPC et un routeur connecté au vPC ou un routeur connecté par un port orphelin vPC, une oscillation continue des contiguïtés de protocole de routage peut se produire après l'activation de l'amélioration de la passerelle d'homologue vPC si le routage ou la couche 3 sur vPC n'est pas configurée immédiatement après. Ces scénarios de défaillance sont décrits en détail dans les sections Contiguïtés de protocole de routage de monodiffusion sur un vPC avec passerelle homologue vPC et Contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC avec passerelle homologue vPC de ce document.

Pour résoudre ce problème, activez l'amélioration du routage/couche 3 sur vPC avec la commande de configuration du domaine vPC layer3 peer-router immédiatement après avoir activé l'amélioration de la passerelle homologue vPC avec la commande de configuration du domaine vPC peer-gateway.

Désactivation automatique des redirections ICMP et ICMPv6

Lorsque l'amélioration de la passerelle d'homologue vPC est activée, la génération de paquets de redirection ICMP et ICMPv6 est automatiquement désactivée sur toutes les interfaces SVI de VLAN vPC (c'est-à-dire, toute interface SVI associée à un VLAN qui est agrégé sur la liaison d'homologue vPC). Pour ce faire, le commutateur configure no ip redirects et no ipv6 redirects sur tous les SVI de VLAN vPC. Cela empêche un commutateur de générer des paquets de redirection ICMP en réponse aux paquets qui entrent dans le commutateur, mais qui ont l'adresse MAC de

destination et l'adresse IP de l’homologue vPC du commutateur.

Si des paquets de redirection ICMP ou ICMPv6 sont nécessaires dans votre environnement au sein d'un VLAN spécifique, vous devez exclure ce VLAN de l'utilisation de l'amélioration de la passerelle d'homologue vPC à l'aide de la commande de configuration de domaine peer-gateway exclude-vlan <vlan-id> vPC.

Remarque : la commande de configuration de domaine vPC peer-gateway exclude-vlan <vlan-id> n'est pas prise en charge sur les commutateurs de la gamme Nexus 9000.

Configuration

Vous trouverez ici un exemple de la configuration de la fonctionnalité de passerelle homologue vPC.

Dans cet exemple, N9K-1 et N9K-2 sont des homologues vPC dans un domaine de vPC. Les deux homologues vPC ont un groupe HSRP configuré pour le VLAN 10. N9K-1 est le routeur HSRP actif avec une priorité de 150, tandis que N9K-2 est le routeur en attente HSRP avec une priorité par défaut de 100.

<#root>

```
N9K-1#

show running-config vpc

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

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

<snip>
vpc domain 1
  peer-keepalive destination 10.82.140.42

interface port-channel1
  vpc peer-link

N9K-1#

show running-config interface vlan 10

<snip>
interface Vlan10
  no shutdown
  ip address 192.168.10.2/24
  hsrp 10
    preempt
```

```
    priority 150
    ip 192.168.10.1

N9K-2#

show running-config interface vlan 10

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

N9K-1#

show hsrp interface vlan 10 brief

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

N9K-2#

show hsrp interface vlan 10 brief

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

L'adresse MAC VLAN 10 SVI de K9K-1 est 00ee.ab67.db47, et l'adresse MAC VLAN 10 SVI de N9K-2 est 00ee.abd8.747 L'adresse MAC virtuelle HSRP pour VLAN 10 est 0000.0c07.ac0a. Dans cet état, l'adresse MAC VLAN 10 SVI de chaque commutateur et l'adresse MAC virtuelle HSRP sont indiquées dans le tableau d'adresses MAC de chaque commutateur. L'adresse MAC SVI du VLAN 10 de chaque commutateur et l'adresse MAC virtuelle HSRP ont l'indicateur Gateway (G) présent, qui indique que le commutateur achemine localement les paquets destinés à cette adresse MAC.

Notez que le tableau d'adresses MAC de N9K-1 ne contient pas d'indicateur de passerelle pour l'adresse MAC du VLAN 10 SVI de N9K-2. De même, le tableau d'adresses MAC de N9K-2 ne contient pas d'indicateur de passerelle pour l'adresse MAC du VLAN 10 SVI de N9K-1.

<#root>

```
N9K-1#

show mac address-table vlan 10

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

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

N9K-2#
```

show mac address-table vlan 10

```
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    vPC Peer-Link(R)
*   10     00ee.ab67.db47   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.abd8.747f   static   -         F      F    sup-eth1(R)
```

Nous pouvons activer l’amélioration de la passerelle homologue vPC grâce à la commande de configuration du domaine vPC peer-gateway. Cela permet au commutateur d'acheminer localement les paquets reçus avec une adresse MAC de destination appartenant à l'adresse MAC de leur homologue vPC apprise sur le Peer-Link vPC. Cela se fait en réglant l’adresse MAC de l’homologue vPC comme indicateur de passerelle dans le tableau d’adresses MAC du commutateur.

<#root>

```
N9K-1#
```

configure terminal

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

vpc domain 1

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

peer-gateway

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

end

```
N9K-1#

N9K-2#
```

configure terminal

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

vpc domain 1

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

peer-gateway

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

```
end

N9K-2#
```

Vous pouvez vérifier que l'amélioration de la passerelle homologue vPC fonctionne comme prévu en validant la présence de l'indicateur de passerelle dans le tableau d'adresses MAC pour l'adresse MAC de l'homologue vPC.

```
<#root>

N9K-1#

show mac address-table vlan 10

Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    sup-eth1(R)
G   10     00ee.ab67.db47   static   -         F      F    sup-eth1(R)
G   10     00ee.abd8.747f   static   -         F      F    vPC Peer-Link(R)

N9K-2#

show mac address-table vlan 10

Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link,
        (T) - True, (F) - False, C - ControlPlane MAC, ~ - vsan
   VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
G   10     0000.0c07.ac0a   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.ab67.db47   static   -         F      F    vPC Peer-Link(R)
G   10     00ee.abd8.747f   static   -         F      F    sup-eth1(R)
```

Incidence

L'impact de l'activation de l'amélioration de la passerelle d'homologue vPC peut varier en fonction de la topologie environnante et du comportement des hôtes connectés, comme décrit dans les sous-sections suivantes. Si aucune des sous-sections suivantes ne s'applique à votre environnement, l'activation de l'amélioration de la passerelle d'homologue vPC n'est pas gênante et n'a pas d'impact sur votre environnement.

Oscillation des contiguïtés de protocole de routage de monodiffusion sur les vPC ou les VLAN vPC

Si des contiguïtés de protocole de routage de monodiffusion dynamique sont formées entre deux homologues vPC et un routeur connecté au vPC ou un routeur connecté par un port orphelin vPC, une oscillation continue des contiguïtés de protocole de routage peut se produire après l'activation

de l’amélioration de la passerelle d’homologue vPC si le routage ou la couche 3 sur vPC n’est pas configurée immédiatement après. Ces scénarios de défaillance sont décrits en détail dans les sections Contiguïtés de protocole de routage de monodiffusion sur un vPC avec passerelle homologue vPC et Contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC avec passerelle homologue vPC de ce document.

Pour résoudre ce problème, activez l’amélioration du routage/couche 3 sur vPC avec la commande de configuration du domaine vPC layer3 peer-router immédiatement après avoir activé l’amélioration de la passerelle homologue vPC avec la commande de configuration du domaine vPC peer-gateway.

Désactivation automatique des redirections ICMP et ICMPv6

Lorsque l'amélioration de la passerelle d'homologue vPC est activée, la génération de paquets de redirection ICMP et ICMPv6 est automatiquement désactivée sur toutes les interfaces SVI de VLAN vPC (c'est-à-dire, toute interface SVI associée à un VLAN qui est agrégé sur la liaison d'homologue vPC). Pour ce faire, le commutateur configure no ip redirects et no ipv6 redirects sur tous les SVI de VLAN vPC. Cela empêche un commutateur de générer des paquets de redirection ICMP en réponse aux paquets qui entrent dans le commutateur, mais qui ont l'adresse MAC de destination et l'adresse IP de l’homologue vPC du commutateur.

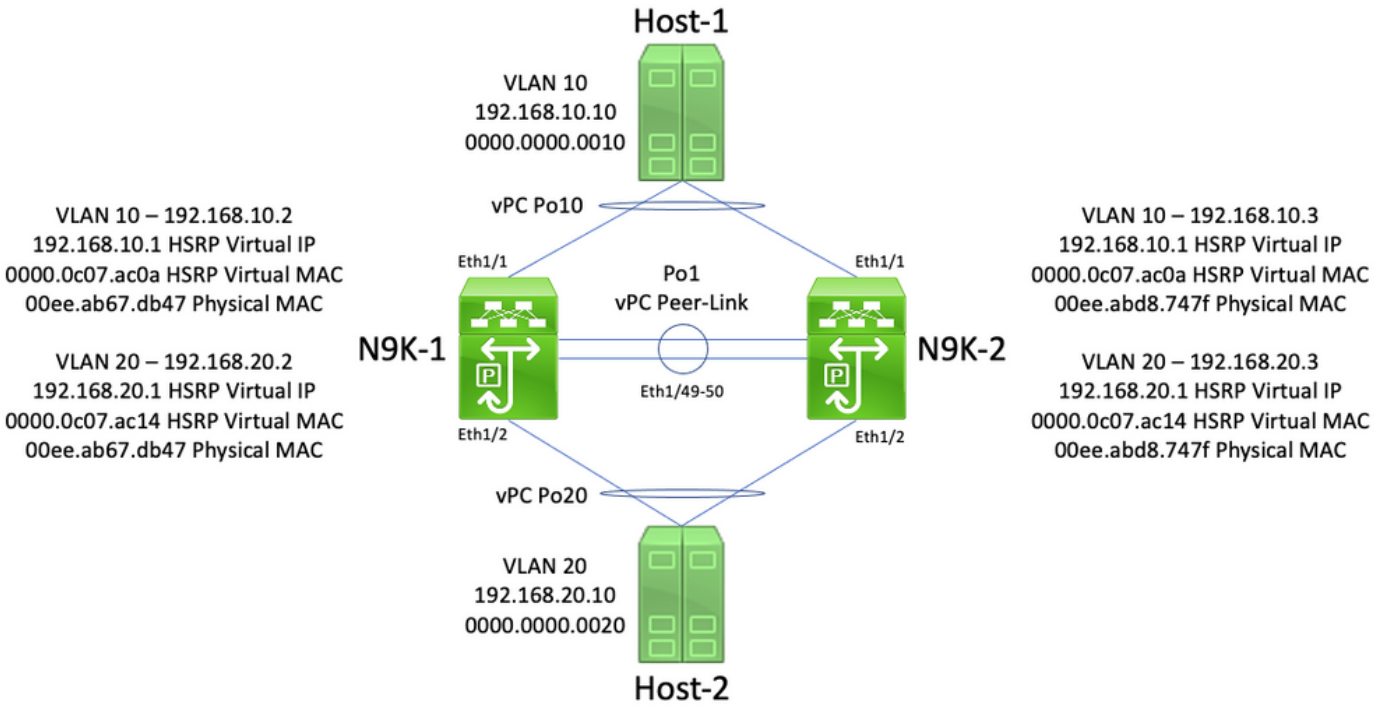
Si des paquets de redirection ICMP ou ICMPv6 sont nécessaires dans votre environnement au sein d'un VLAN spécifique, vous devez exclure ce VLAN de l'utilisation de l'amélioration de la passerelle d'homologue vPC à l'aide de la commande de configuration de domaine peer-gateway exclude-vlan <vlan-id> vPC.

Remarque : la commande de configuration de domaine vPC peer-gateway exclude-vlan <vlan-id> n'est pas prise en charge sur les commutateurs de la gamme Nexus 9000.

Exemples de scénarios de défaillance

Hôtes connectés au vPC avec comportement de transfert non standard

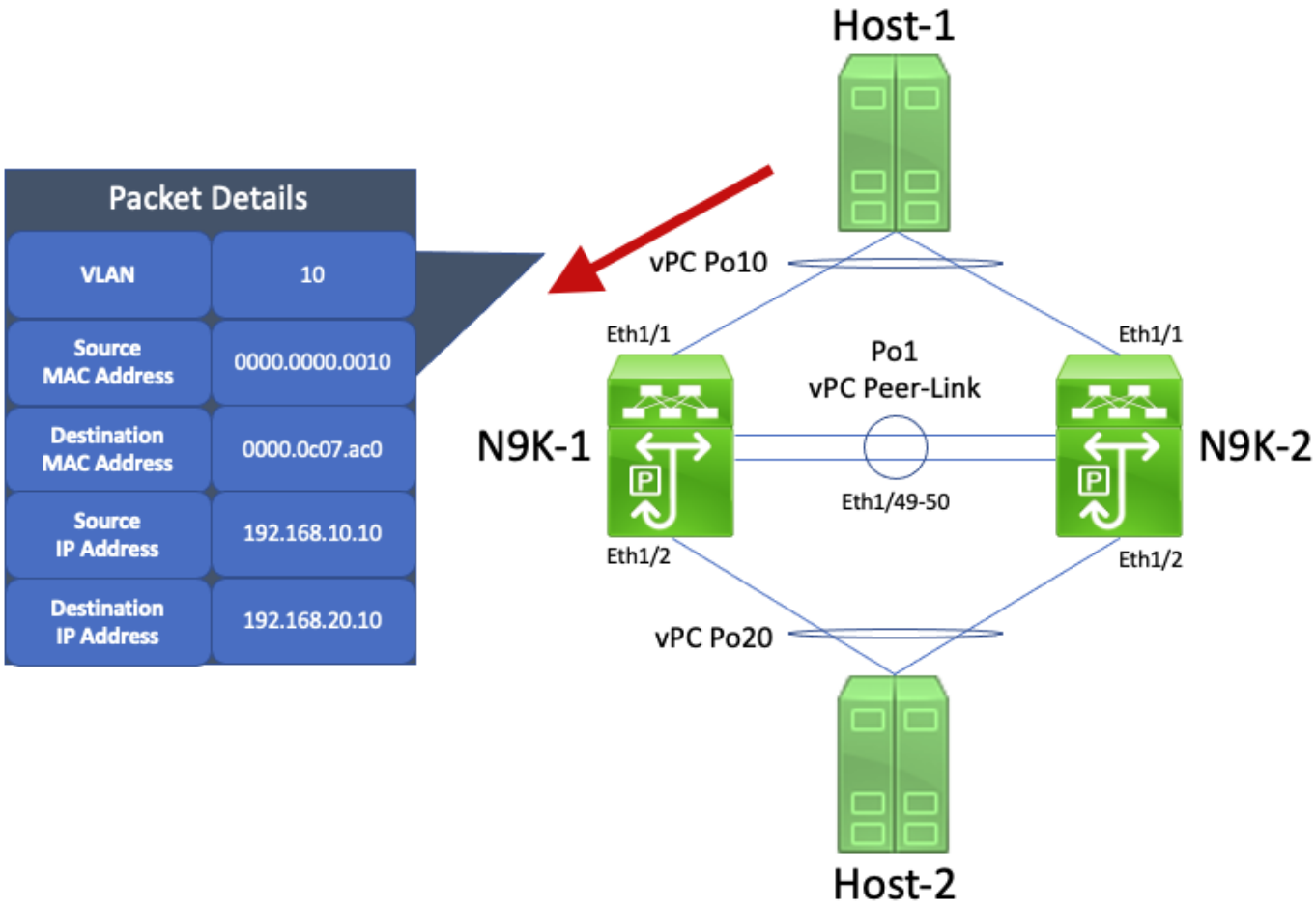
Examinez la topologie suivante :



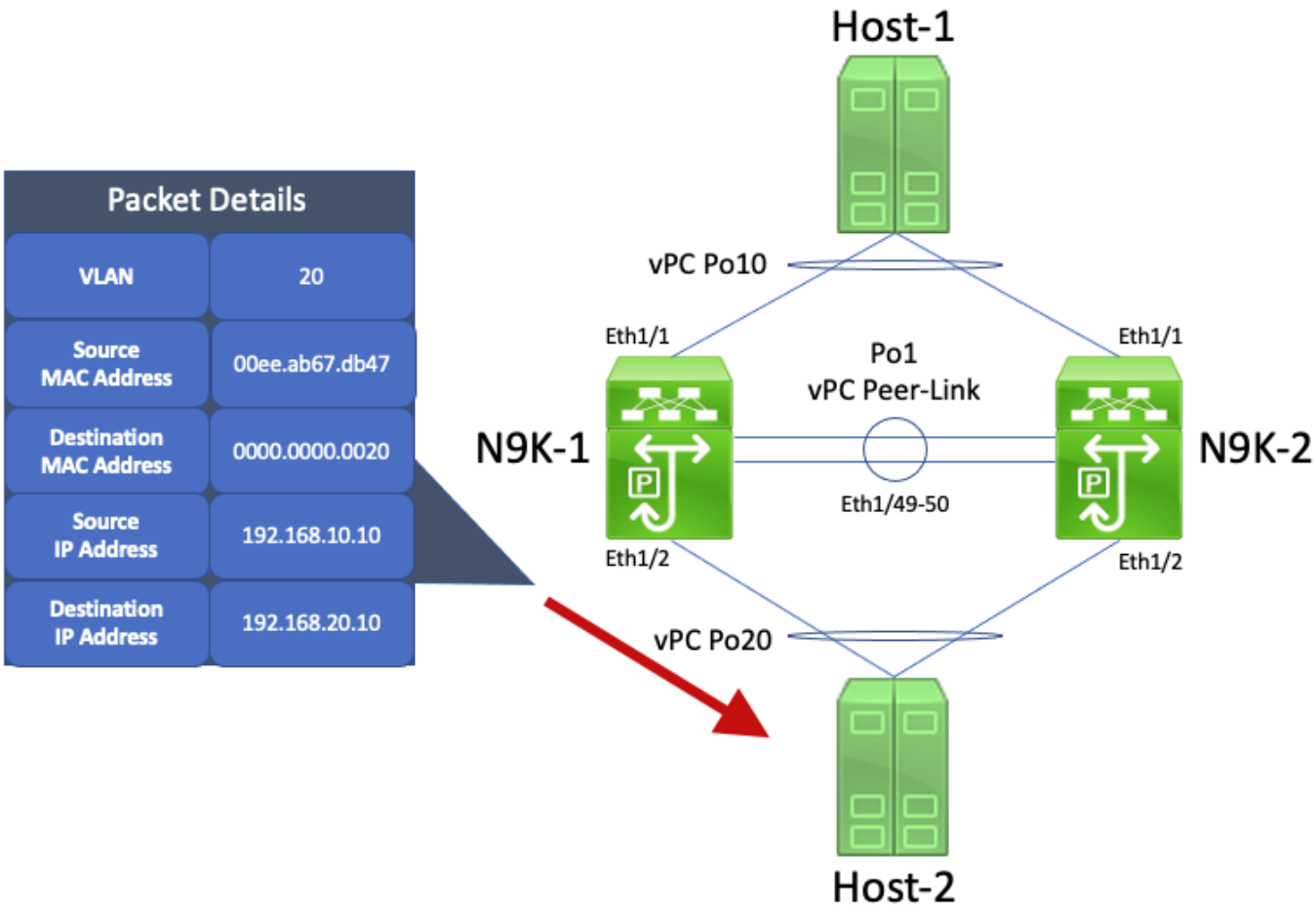
Dans cette topologie, N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC qui effectuent le routage inter-VLAN entre le VLAN 10 et le VLAN 20. L'interface Po1 est le lien homologue vPC. Un hôte nommé Hôte-1 est connecté avec le vPC Po10 à N9K-1 et N9K-2 dans le VLAN 10. Hôte-1 possède l'adresse IP 192.168.10.10 avec l'adresse MAC 0000.0000.0010. Un hôte nommé Hôte-2 est connecté avec le vPC Po20 à N9K-1 et N9K-2 dans le VLAN 20. Hôte-2 possède l'adresse IP 192.168.20.10 avec l'adresse MAC 0000.0000.0020.

N9K-1 et N9K-2 ont tous deux des SVI dans le VLAN 10 et le VLAN 20 et le HSRP est activé dans chaque SVI. L'interface VLAN 10 de N9K-1 possède l'adresse IP 192.168.10.2, et l'interface VLAN 20 de N9K-1 possède l'adresse IP 192.168.20.2. Les deux SVI de N9K-1 possèdent l'adresse MAC physique 00ee.ab67.db47. L'interface VLAN 10 de N9K-2 possède l'adresse IP 1921.68.10.3, et l'interface VLAN 20 de N9K-2 possède l'adresse IP 192.168.20.3. Les deux SVI de N9K-2 possèdent l'adresse MAC physique 00ee.abd8.747f. L'adresse IP virtuelle HSRP pour le VLAN 10 est 192.168.10.1 et l'adresse MAC virtuelle HSRP est 0000.0c07.ac0a. L'adresse IP virtuelle HSRP pour le VLAN 20 est 192.168.20.1 et l'adresse MAC virtuelle HSRP est 0000.0c07.ac14.

Voici un scénario dans lequel Hôte-1 envoie un paquet de demande Echo ICMP à Hôte-2. Une fois que l'Hôte-1 a résolu l'ARP pour sa passerelle par défaut (l'adresse IP virtuelle HSRP), l'Hôte-1 effectue le comportement de transfert standard et génère un paquet de demande Echo ICMP avec l'adresse IP source 192.168.10.10, l'adresse IP de destination 192.168.20.10 , l'adresse MAC source 0000 0000.0010 et l'adresse MAC de destination 0000.0c07.ac0a. Ce paquet sort vers N9K-1. Un exemple visuel de ce processus est montré ici.

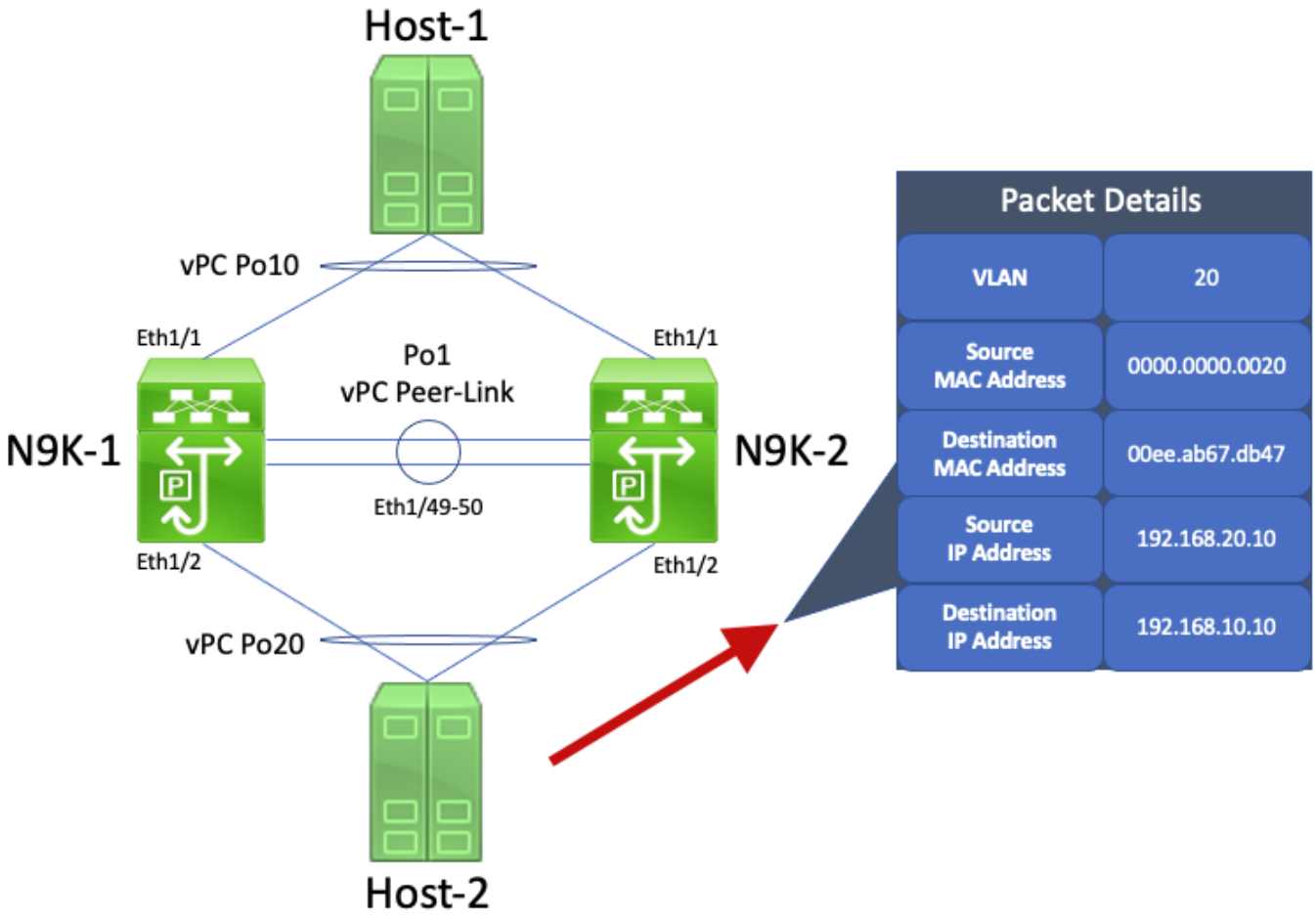


N9K-1 reçoit ce paquet. Comme ce paquet est destiné à l'adresse MAC virtuelle HSRP, N9K-1 est en mesure d'acheminer ce paquet en fonction de son tableau de routage local, quel que soit l'état de son plan de commande HSRP. Ce paquet est routé du VLAN 10 au VLAN 20. Dans le cadre du routage du paquet, N9K-1 réécrit le paquet en réadressant les champs d'adresse MAC source et de destination du paquet. La nouvelle adresse MAC source du paquet est l'adresse MAC physique associée à l'interface SVI VLAN 20 de N9K-1 (00ee.ab67.db47) et la nouvelle adresse MAC de destination est l'adresse MAC associée à l'hôte 2 (0000.000.0020). Un exemple visuel de ce processus est montré ici.

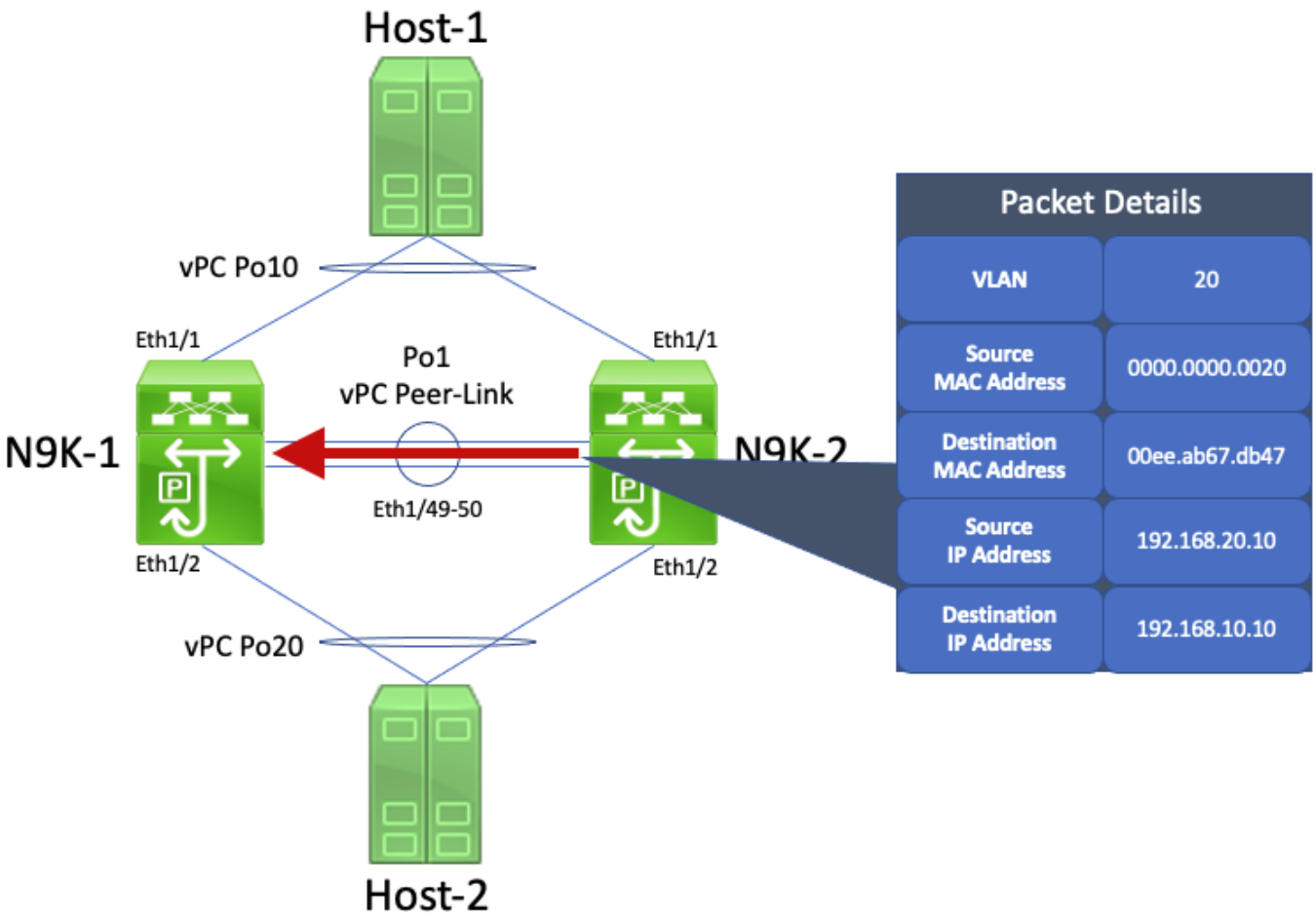


Hôte-2 reçoit ce paquet et génère un paquet de réponse Echo ICMP en réponse au paquet de demande Echo ICMP de l'Hôte-1. Toutefois, lorsque Hôte-2 ne respecte pas le comportement de transfert standard. Pour optimiser son transfert, Hôte-2 n'effectue pas de recherche dans le tableau de routage ou dans la mémoire cache ARP pour trouver l'adresse IP de l'Hôte-1 (192.168.10.10). Il inverse plutôt les champs d'adresse MAC source et de destination du paquet de demande Echo ICMP que l'Hôte-2 avait initialement reçu. Par conséquent, le paquet de réponse d'écho ICMP généré par l'hôte 2 possède l'adresse IP source 192.168.20.10, l'adresse IP de destination 192.168.10.10, l'adresse MAC source 000.000.0020 et l'adresse MAC de destination 00ee.ab67.db47.

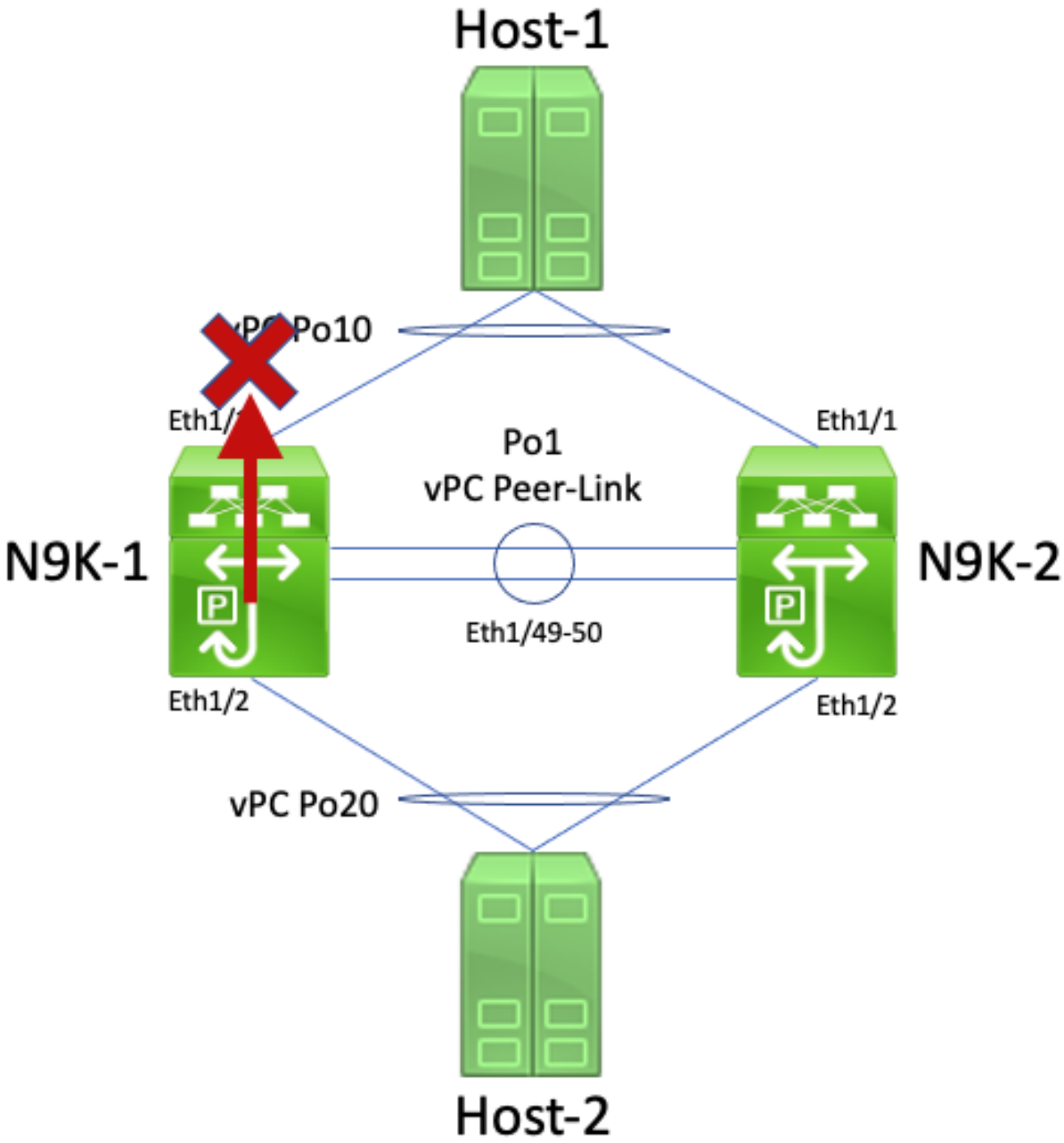
Si ce paquet de réponse d'écho ICMP sort vers N9K-1, il est transféré vers l'hôte 1 sans problème. Cependant, voici un scénario dans lequel ce paquet de réponse Echo ICMP sort vers N9K-2.



N9K-2 reçoit ce paquet. Comme ce paquet est destiné à l'adresse MAC physique de l'interface SVI VLAN 20 de N9K-1, N9K-2 transfère ce paquet via le Peer-Link vPC vers N9K-1, car N9K-2 ne peut pas acheminer ce paquet pour le compte de N9K-1. Un exemple visuel de ce processus est montré ici.



N9K-1 reçoit ce paquet. Puisque ce paquet est destiné à l'adresse MAC physique du VLAN 20 SVI de N9K-1, N9K-1 peut acheminer ce paquet en fonction de son tableau de routage local, quel que soit l'état de son plan de commande HSRP. Ce paquet est routé du VLAN 20 vers le VLAN 10. Cependant, l'interface de sortie pour cette route se résout en vPC Po10, qui est actif sur N9K-2. Il s'agit d'une violation de la règle d'évitement de boucle vPC : si N9K-1 reçoit un paquet via l'Peer-Link vPC, N9K-1 ne peut pas transférer ce paquet à partir d'une interface vPC si la même interface vPC est active sur N9K-2. N9K-1 abandonne ce paquet à la suite de cette violation. Un exemple visuel de ce processus est montré ici.



Vous pouvez résoudre ce problème en activant l'amélioration de la passerelle homologue vPC avec la commande de configuration du domaine vPC peer-gateway. Cela permet à N9K-2 d'acheminer le paquet de réponse Echo ICMP (et d'autres paquets adressés de manière similaire) au nom de N9K-1, même si l'adresse MAC de destination du paquet appartient à N9K-1 et non à N9K-2. Par conséquent, N9K-2 peut transférer ce paquet hors de son interface vPC Po10 au lieu de le transférer sur le lien homologue vPC.

Routage/couche 3 sur vPC (routeur homologue de couche 3)

Cette section décrit l'amélioration du routage/couche 3 sur vPC, qui est activée avec la commande de configuration du domaine vPC peer-router.

Remarque : la formation de contiguïtés de protocole de routage multidiffusion (à savoir, les contiguïtés PIM [Protocol Independent Multicast]) sur un vPC n'est pas prise en charge avec l'amélioration de routage/couche 3 sur vPC activée.

Aperçu

Dans certains environnements, les clients souhaitent connecter un routeur à une paire de commutateurs Nexus par vPC et créer des contiguïtés de protocole de routage de monodiffusion sur le vPC avec les deux homologues vPC. Les clients pourraient également vouloir connecter un routeur à un homologue vPC unique avec un VLAN vPC et former des contiguïtés de protocole de routage de monodiffusion avec les deux homologues vPC sur le VLAN vPC. Par conséquent, le routeur connecté au vPC aurait des chemins multiples à coûts égaux (ECMP) pour les préfixes annoncés par les deux commutateurs Nexus. Cela peut être préférable à l'utilisation de liens de routage dédiés entre le routeur connecté au vPC et les deux homologues vPC pour conserver l'utilisation des adresses IP (3 adresses IP nécessaires au lieu de 4 adresses IP) ou réduire la complexité de la configuration (interfaces routées en parallèle à des SVI, en particulier dans les environnements VRF-Lite qui nécessiteraient des sous-interfaces).

Historiquement, la formation de contiguïtés de protocole de routage de monodiffusion sur un vPC n'était pas prise en charge sur les plateformes Cisco Nexus. Cependant, les clients peuvent avoir mis en œuvre une topologie dans laquelle des contiguïtés de protocole de routage de monodiffusion se forment sur un vPC sans problème, même si elles ne sont pas prises en charge. Après certains changements dans le réseau (comme une mise à niveau logicielle du routeur connecté au vPC ou des homologues vPC eux-mêmes, un basculement de pare-feu, etc.), les contiguïtés de protocole de routage de monodiffusion sur un vPC cessent de fonctionner, ce qui entraîne une perte de paquets pour le trafic de plan de données ou les contiguïtés de protocole de routage de monodiffusion ne parviennent pas à trouver un ou deux homologues vPC. Les détails techniques expliquant pourquoi ces scénarios échouent et ne sont pas pris en charge sont décrits dans la section Exemples de scénarios de défaillance de ce document.

L'amélioration du routage/couche 3 sur vPC a été introduite pour ajouter la prise en charge de la formation de contiguïtés de protocole de routage de monodiffusion sur un vPC. Pour ce faire, les paquets de protocole de routage de monodiffusion avec une durée de vie (TTL) de 1 peuvent être transférés sur le lien homologue vPC sans décrémenter la TTL du paquet. Par conséquent, les contiguïtés de protocole de routage de monodiffusion peuvent être formées sur un vPC ou un VLAN vPC sans problème. L'amélioration du routage/couche 3 sur vPC peut être activée à l'aide de la commande de configuration du domaine vPC layer3 peer-router une fois que l'amélioration de la passerelle homologue vPC a été activée à l'aide de la commande de configuration du domaine vPC peer-gateway.

Les versions logicielles de NX-OS qui ont introduit la prise en charge de l'amélioration du routage/couche 3 sur vPC pour chaque plateforme Cisco Nexus sont documentées dans le tableau 2 (« Prise en charge des contiguïtés des protocoles de routage sur VLAN vPC ») dans le

[document Topologies prises en charge pour le routage sur canal de port virtuel sur les plateformes Nexus.]

Mises en garde

Journaux systèmes VPC-2-L3_VPC_UNEQUAL_WEIGHT occasionnels

Une fois l'amélioration du routage/de la couche 3 sur vPC activée, les deux homologues vPC commencent à générer des syslogs semblables à l'un des suivants une fois par heure :

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

Aucun de ces journaux système n'indique un problème avec le commutateur. Ces journaux système sont des avertissements pour l'administrateur que la configuration, le coût et la pondération du routage doivent être identiques sur les deux homologues vPC lorsque l'amélioration du routage/couche 3 sur le vPC est activée afin de s'assurer que les deux homologues vPC sont en mesure d'acheminer le trafic de manière identique. Cela n'indique pas nécessairement que la configuration de routage, le coût ou la pondération ne correspondent pas sur l'un ou l'autre des homologues VPC.

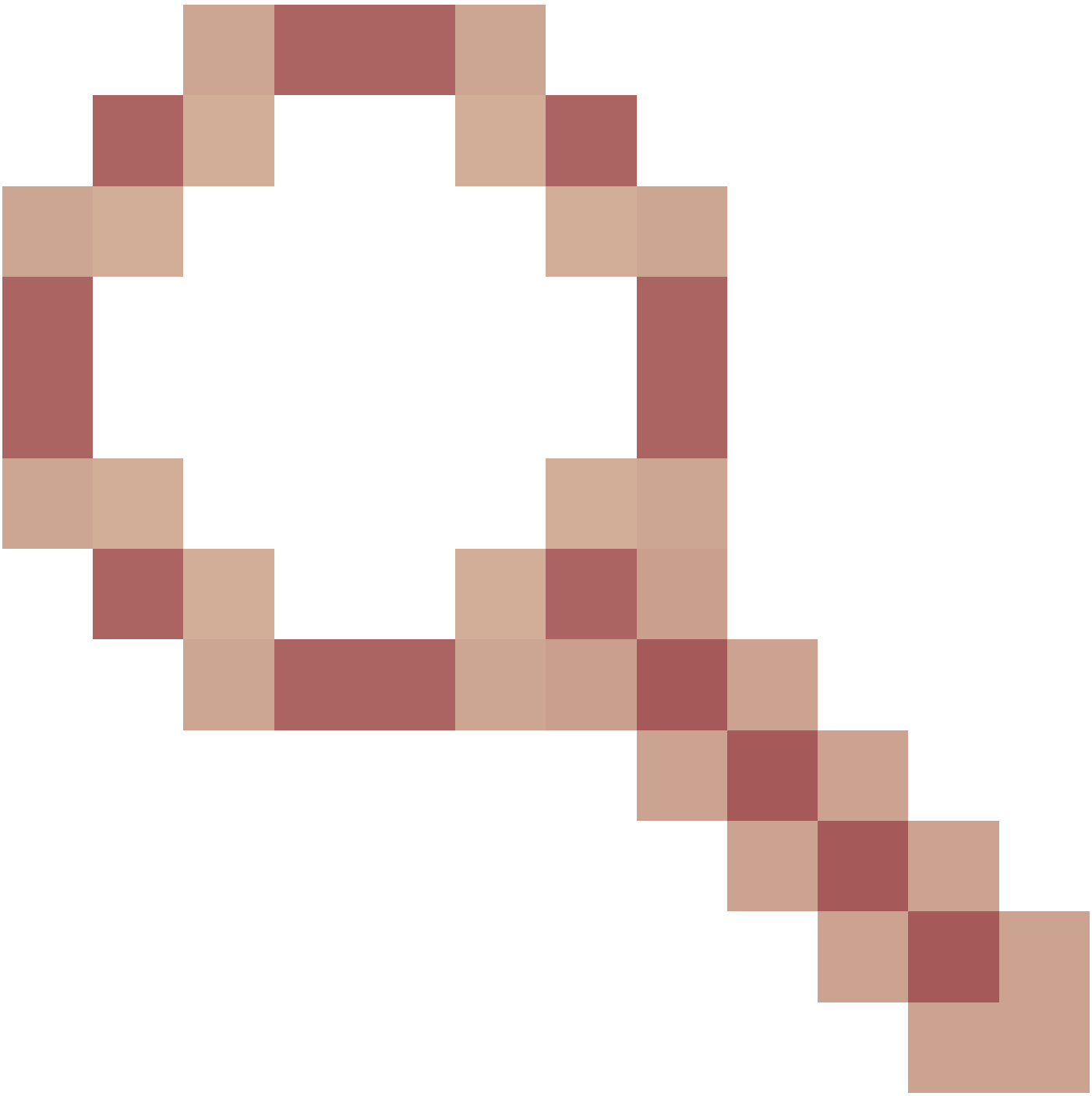
Ces journaux système peuvent être désactivés en effectuant la configuration indiquée ici.

<#root>

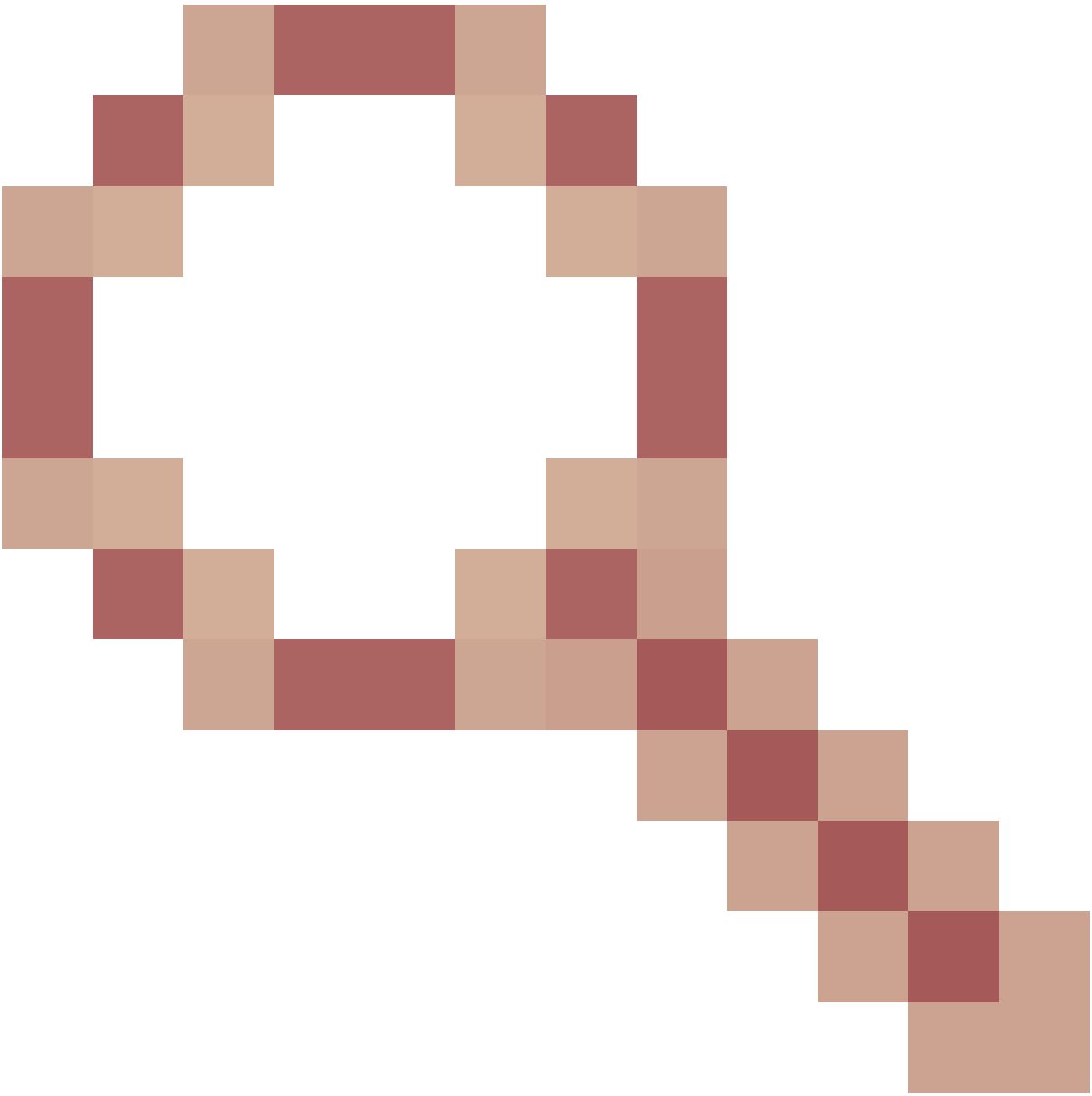
```
switch#
configure terminal
switch(config)#
vpc domain 1
switch(config-vpc-domain)#
no layer3 peer-router syslog
switch(config-vpc-domain)#
end
switch#
```

Cette configuration doit être effectuée sur les deux homologues vPC pour désactiver le journal système sur les deux homologues vPC.

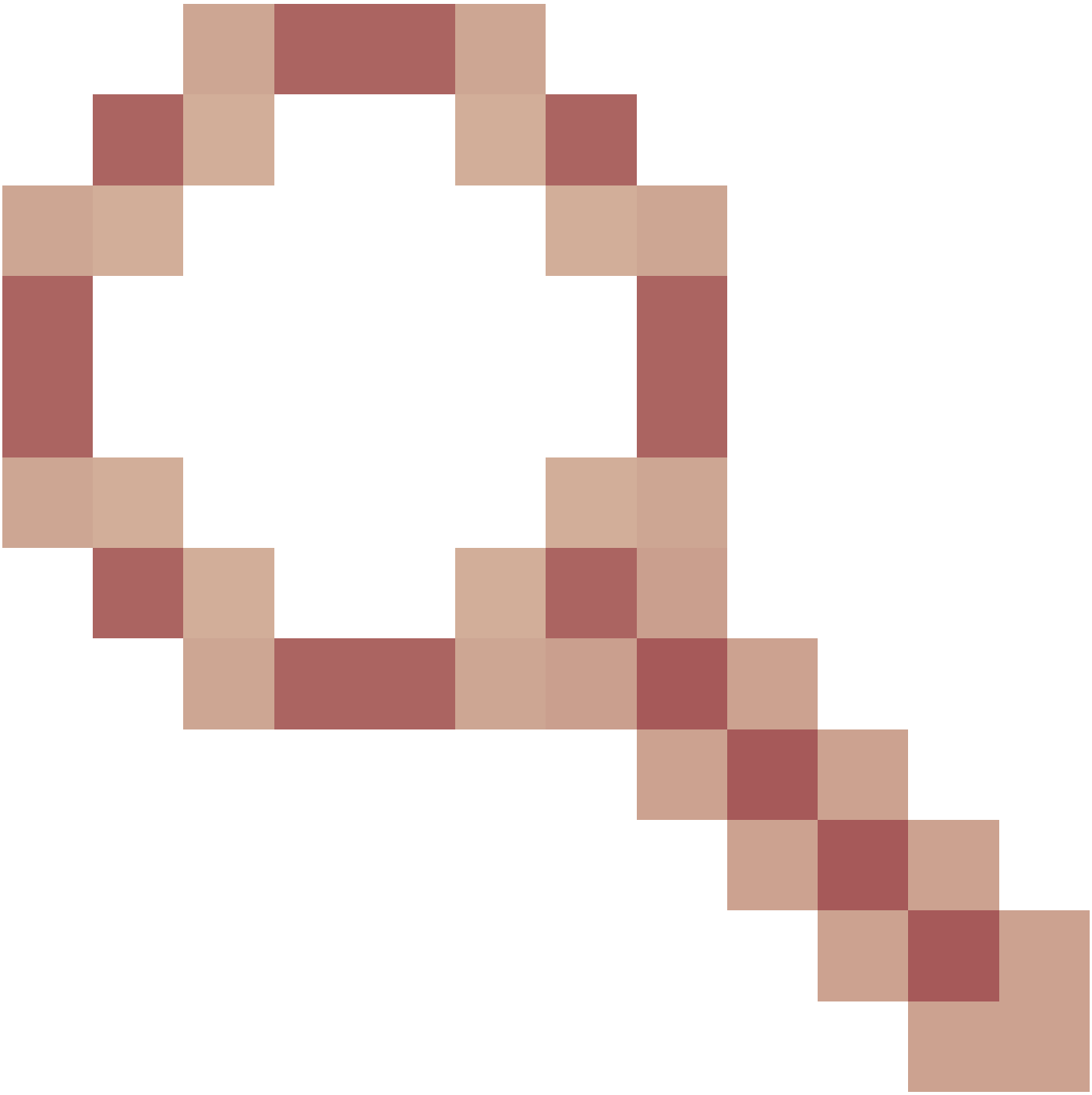
Trafic du plan de données avec TTL de 1 logiciel transféré en raison de l'ID de bogue Cisco [CSCvs82183]



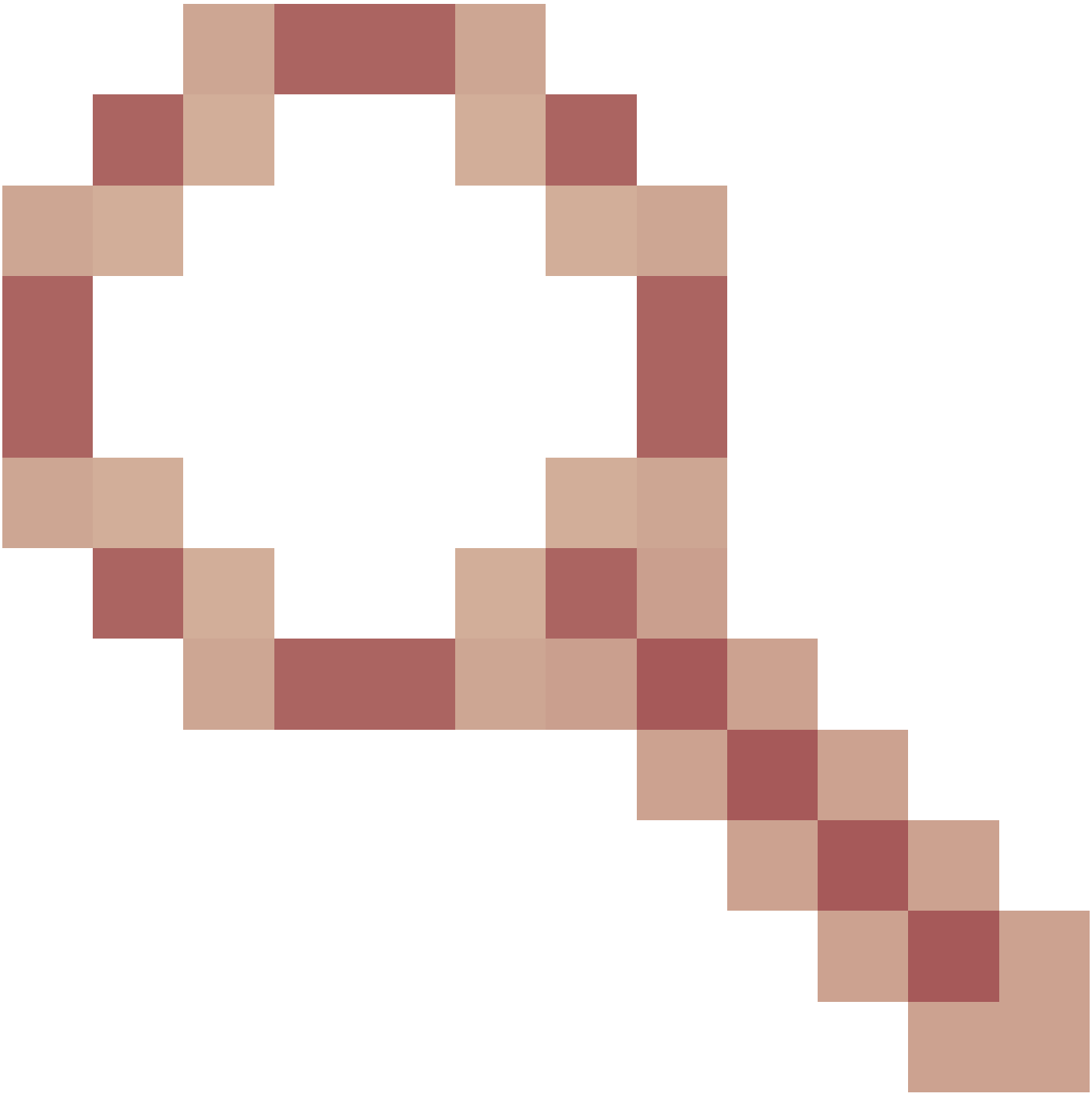
et ID de bogue Cisco CSCvw16965



Lorsque l'amélioration Routage/Couche 3 sur vPC est activée sur les commutateurs de la gamme Nexus 9000 équipés d'un ASIC Cloud Scale exécutant une version du logiciel NX-OS antérieure à la version 9.3(6) du logiciel NX-OS, le trafic du plan de données qui n'est pas associé à un protocole de routage monodiffusion ayant une durée de vie de 1 est envoyé au superviseur et transféré dans le logiciel plutôt que dans le matériel. Selon que le commutateur Nexus est un commutateur à châssis fixe (également appelé « Top of Rack ») ou un commutateur à châssis modulaire (également appelé « End of Row »), ainsi que la version actuelle du logiciel NX-OS du commutateur, la cause principale de ce problème peut être attribuée à l'un ou l'autre défaut logiciel ID de bogue Cisco CSCvs82183



ou défaut logiciel ID de bogue Cisco CSCvw16965



. Les deux défauts logiciels n’affectent que les commutateurs de la série Nexus 9000 équipés d’un Cloud Scale ASIC. Aucune autre plateforme matérielle Cisco Nexus n’est affectée par l’un ou l’autre de ces problèmes. Pour plus de détails, consultez les informations contenues dans chaque défaut logiciel.

Pour éviter ces défauts logiciels, Cisco conseille de mettre à niveau vers la version logicielle NX-OS 9.3(6) ou une version ultérieure. À titre de recommandation générale, Cisco conseille d’effectuer régulièrement une mise à niveau vers la version logicielle actuelle de NX-OS pour les commutateurs de la série Nexus 9000 référencés dans le document des versions recommandées de Cisco NX-OS pour les commutateurs de la série Cisco Nexus 9000.

Configuration

Vous trouverez ici les instructions pour configurer l’amélioration du routage/couche 3 sur vPC.

Dans cet exemple, N9K-1 et N9K-2 sont des homologues vPC dans un domaine de vPC. L'amélioration de la passerelle homologue vPC est déjà activée pour les deux homologues vPC, ce qui est nécessaire pour activer l'amélioration du routage/couche 3 sur vPC. Les deux homologues vPC ont un SVI dans le VLAN 10, qui est activé au cours du processus OSPF 1. N9K-1 et N9K-3 sont bloqués dans un état OSPF EXSTART/EXCHANGE avec un routeur OSPF connecté au vPC possédant l'adresse IP et l'ID de voisin 192.168.10.3.

<#root>

```
N9K-1#

show running-config vpc

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

interface port-channel1
  vpc peer-link

N9K-2#

show running-config vpc

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

interface port-channel1
  vpc peer-link

N9K-1#

show running-config interface Vlan10


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

N9K-2#

show running-config interface Vlan10


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

N9K-1#

show running-config ospf
```

```
feature ospf

router ospf 1

interface Vlan10
  ip router ospf 1 area 0.0.0.0

N9K-2#
```

show running-config ospf

```
feature ospf

router ospf 1

interface Vlan10
  ip router ospf 1 area 0.0.0.0

N9K-1#
```

show ip ospf neighbors

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

N9K-2#
```

show ip ospf neighbors

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

Nous pouvons activer l'amélioration du routage/couche 3 sur vPC grâce à la commande de configuration du domaine vPC layer3 peer-router. Cela empêche un homologue vPC de décrémenter la durée de vie des paquets de protocole de routage monodiffusion routés suite à l'activation de l'amélioration de la passerelle d'homologue vPC.

<#root>

```
N9K-1#
```

configure terminal

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

vpc domain 1

```
N9K-1(config-vpc-domain)#
layer3 peer-router
N9K-1(config-vpc-domain)#
end
N9K-1#

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

Vous pouvez vérifier que l'amélioration du routage/couche 3 sur vPC fonctionne comme prévu en validant que la contiguïté OSPF avec le voisin OSPF connecté au vPC passe à l'état FULL peu de temps après l'activation de l'amélioration du routage/couche 3 sur vPC.

<#root>

```
N9K-1#
show ip ospf neighbors

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

N9K-2#
show ip ospf neighbors

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

Incidence

L’activation de l’amélioration du routage/couche 3 sur vPC n’a intrinsèquement pas d’incidence sur le domaine vPC. Cela signifie que lorsque vous activez l'amélioration Routage/Couche 3 sur vPC, ni l'homologue vPC ne suspend de vPC, ni le trafic du plan de données n'est affecté de manière inhérente par l'activation de cette amélioration.

Toutefois, si des contiguïtés de protocole de routage dynamique qui étaient auparavant inactives parce que l’amélioration du routage/couche 3 sur vPC n'était pas activée se présentent soudainement à la suite de l’activation de cette amélioration, selon le rôle des contiguïtés de protocole de routage concernées, les préfixes spécifiques annoncés par le biais de ces contiguïtés et de l’état actuel du tableau de routage de monodiffusion, certaines perturbations peuvent être observées lors de l’activation de l’amélioration du routage/couche 3 sur vPC.

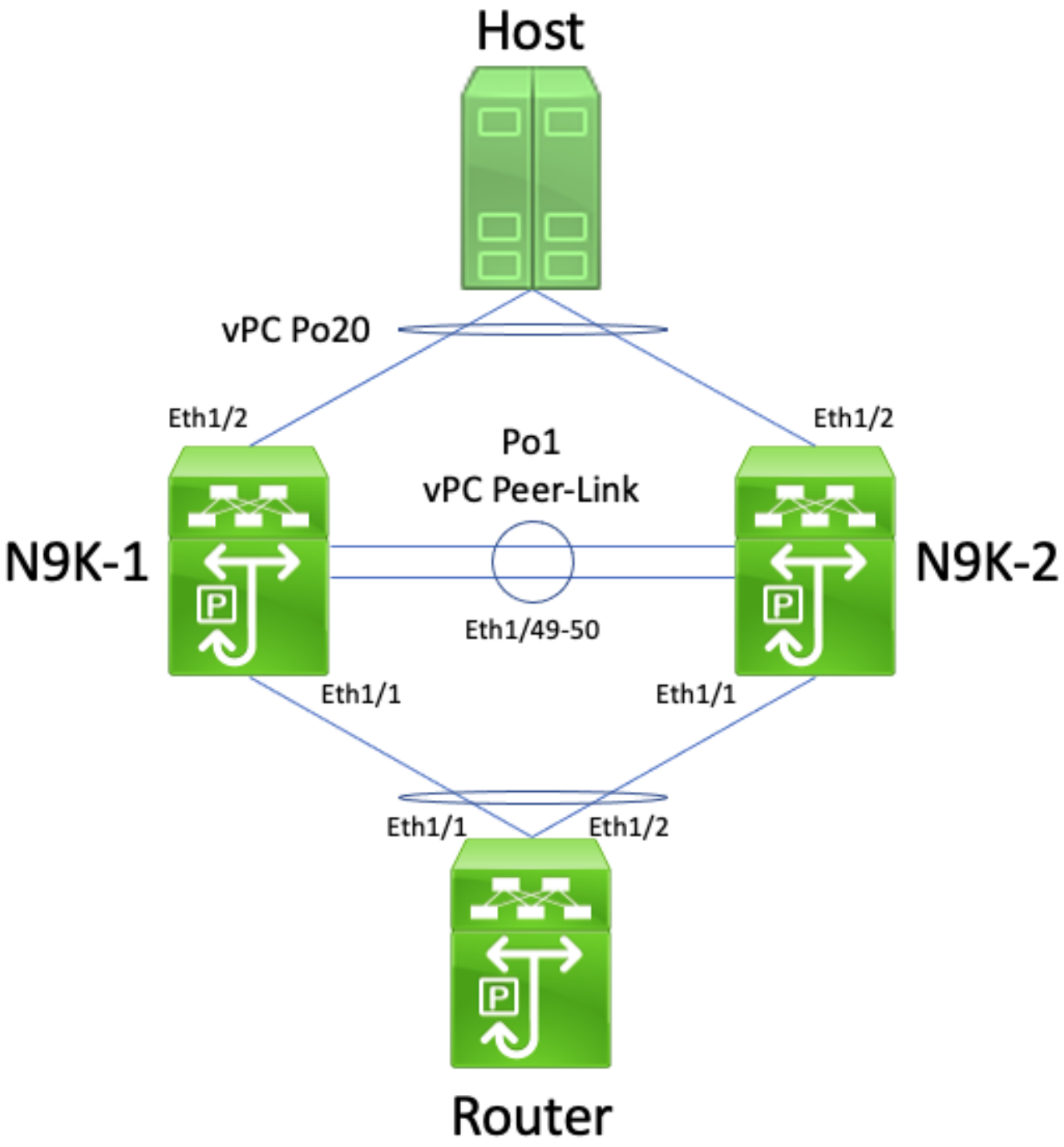
Pour cette raison, Cisco conseille aux clients d'activer cette amélioration au cours d'une fenêtre de maintenance en s'attendant à ce qu'il puisse y avoir une interruption du plan de contrôle et du plan de données, à moins qu'ils ne soient extrêmement confiants que les contiguïtés de protocole de routage affectées n'ont pas d'impact significatif sur le fonctionnement du réseau.

Cisco conseille également de passer attentivement en revue la section des mises en garde de ce document pour tout défaut logiciel affectant votre version logicielle NX-OS qui pourrait faire en sorte que le trafic de plan de données naturel avec une TTL de 1 soit traité dans le logiciel plutôt que dans le matériel.

Exemples de scénarios de défaillance

Contiguïtés de protocole de routage de monodiffusion sur un vPC sans passerelle homologue vPC

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues VPC dans un domaine VPC où l'amélioration de la passerelle homologue VPC n'est pas activée. L'interface Po1 est le lien homologue vPC. Un routeur avec le nom d'hôte Routeur est connecté avec vPC Po10 à N9K-1 et N9K-2. Un hôte est connecté à N9K-1 et N9K-2 avec vPC Po20. L'interface Po10 de Routeur est un canal de port routé qui est activé dans le cadre d'un protocole de routage de monodiffusion. N9K-1 et N9K-2 ont tous deux des interfaces SVI activées avec le même protocole de routage de monodiffusion et sont dans le même domaine de diffusion que Routeur.

Les contiguïtés de protocole de routage de monodiffusion sur un vPC sans l'amélioration de la passerelle homologue vPC ne sont pas prises en charge, car la décision de hachage ECMP du routeur connecté au vPC et sa décision de hachage de canal de port de couche 2 peuvent

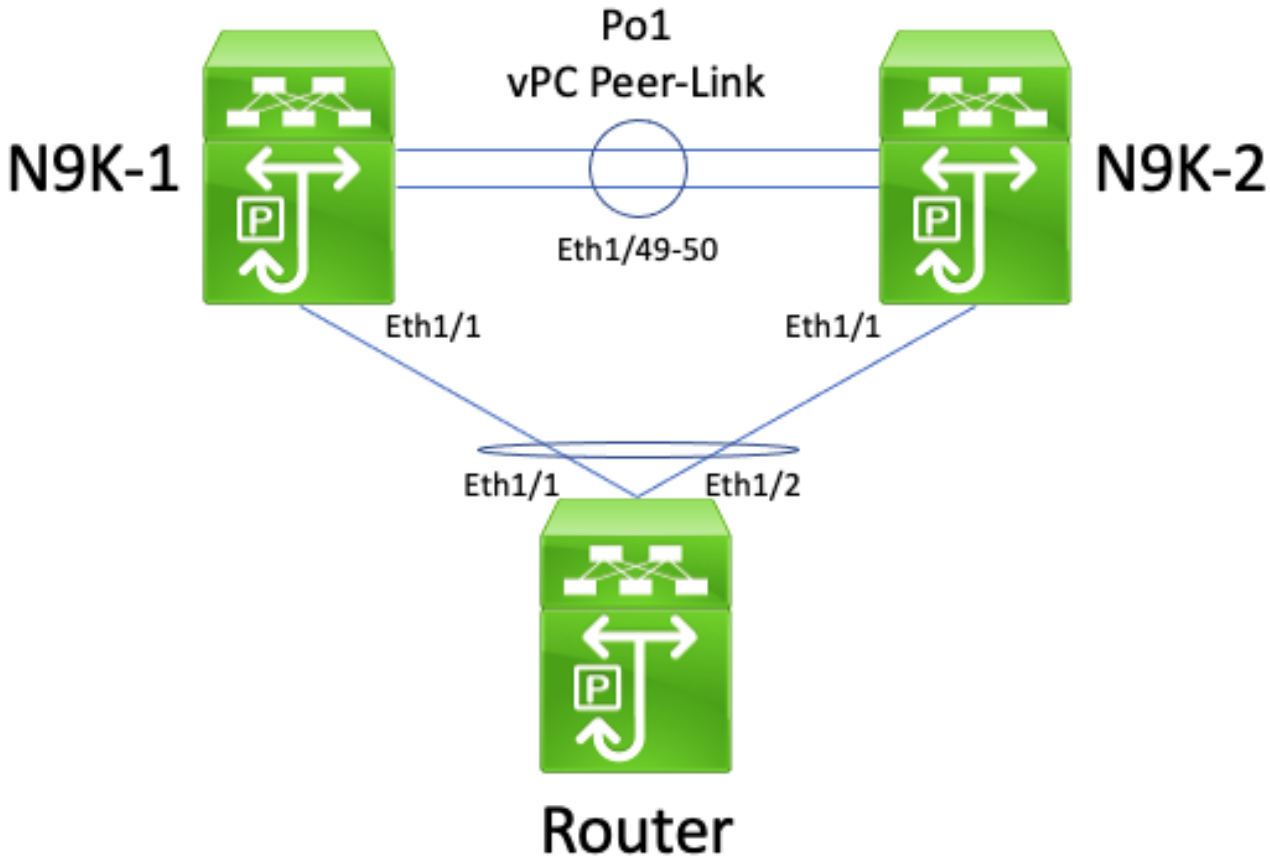
différer. Dans cette topologie, les contiguïtés de protocole de routage se formeraient correctement entre le routeur, N9K-1 et N9K-2. Examinez le flux de trafic entre Routeur et Hôte. Le trafic du plan de données traversant Routeur et destiné Hôte peut être réécrit avec une adresse MAC de destination appartenant à l'adresse MAC SVI de N9K-1 (en raison de la décision de hachage ECMP prise par le routeur), mais sort de l'interface Ethernet1/2 (en raison de la décision de hachage du canal de port de couche 2 prise par le routeur).

N9K-2 reçoit ce paquet et le transfère via la liaison d'homologue vPC, car l'adresse MAC de destination appartient à N9K-1 et l'amélioration de la passerelle d'homologue vPC (qui permet à N9K-2 d'acheminer le paquet pour le compte de N9K-1) n'est pas activée. N9K-1 reçoit ce paquet sur le lien homologue vPC et reconnaît qu'il devrait transférer le paquet hors de son Ethernet1/2 dans le vPC Po20. Cela enfreint la règle de prévention des boucles vPC, de sorte que N9K-1 abandonne le paquet dans le matériel. Par conséquent, vous pourriez observer des problèmes de connectivité ou des pertes de paquets pour certains flux qui traversent le domaine vPC dans cette topologie.

Vous pouvez résoudre ce problème en activant l'amélioration de la passerelle homologue vPC avec la commande de configuration du domaine vPC peer-gateway, puis en activant l'amélioration de routage/couche 3 sur vPC avec la commande de configuration du domaine vPC layer3 peer-router. Pour minimiser les perturbations, vous devez activer les deux améliorations de vPC rapidement afin que le scénario de défaillance décrit dans les contiguïtés de protocole de routage de monodiffusion sur vPC avec passerelle homologue vPC n'ait pas le temps de se produire.

Contiguïtés de protocole de routage de monodiffusion sur un vPC avec passerelle homologue vPC

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue VPC est activée. L'interface Po1 est le lien homologue vPC. Un routeur avec le nom d'hôte Routeur est connecté avec vPC Po10 à N9K-1 et N9K-2. L'interface Po10 de Routeur est un canal de port routé qui est activé dans le cadre d'un protocole de routage de monodiffusion. N9K-1 et N9K-2 ont tous deux des interfaces SVI activées avec le même protocole de routage de monodiffusion et sont dans le même domaine de diffusion que Routeur.

Les contiguïtés de protocole de routage de monodiffusion sur un vPC avec l'amélioration de la passerelle homologue vPC activée ne sont pas prises en charge, car l'amélioration de la passerelle homologue vPC pourrait empêcher la formation de contiguïtés de protocole de routage de monodiffusion entre le routeur connecté au vPC et les deux homologues vPC. Dans cette topologie, une contiguïté de protocole de routage entre le routeur et N9K-1 ou N9K-2 peut ne pas s'établir comme prévu, selon la manière dont les paquets de protocole de routage de monodiffusion provenant du routeur utilisent le hachage N9K-1 ou N9K-2 sur vPC Po10.

Tous les routeurs sont en mesure d'envoyer et de recevoir des paquets de protocole de routage de multidiffusion de lien local (généralement nommés paquets « Hello ») sans problème, car ces paquets sont acheminés avec succès vers le VLAN vPC. Cependant, voici un scénario dans lequel un paquet de protocole de routage de monodiffusion provenant de Routeur destiné à N9K-1 sort d'Ethernet1/2 vers N9K-2 en raison de la décision de hachage du canal de port de la couche 2 de Routeur. Ce paquet est destiné à l'adresse MAC SVI de N9K-1, mais entre dans l'interface Ethernet1/1 de N9K-2. N9K-2 constate que le paquet est destiné à l'adresse MAC SVI de N9K-1, qui est installée dans la table d'adresses MAC de N9K-2 avec l'indicateur « G », ou « Passerelle

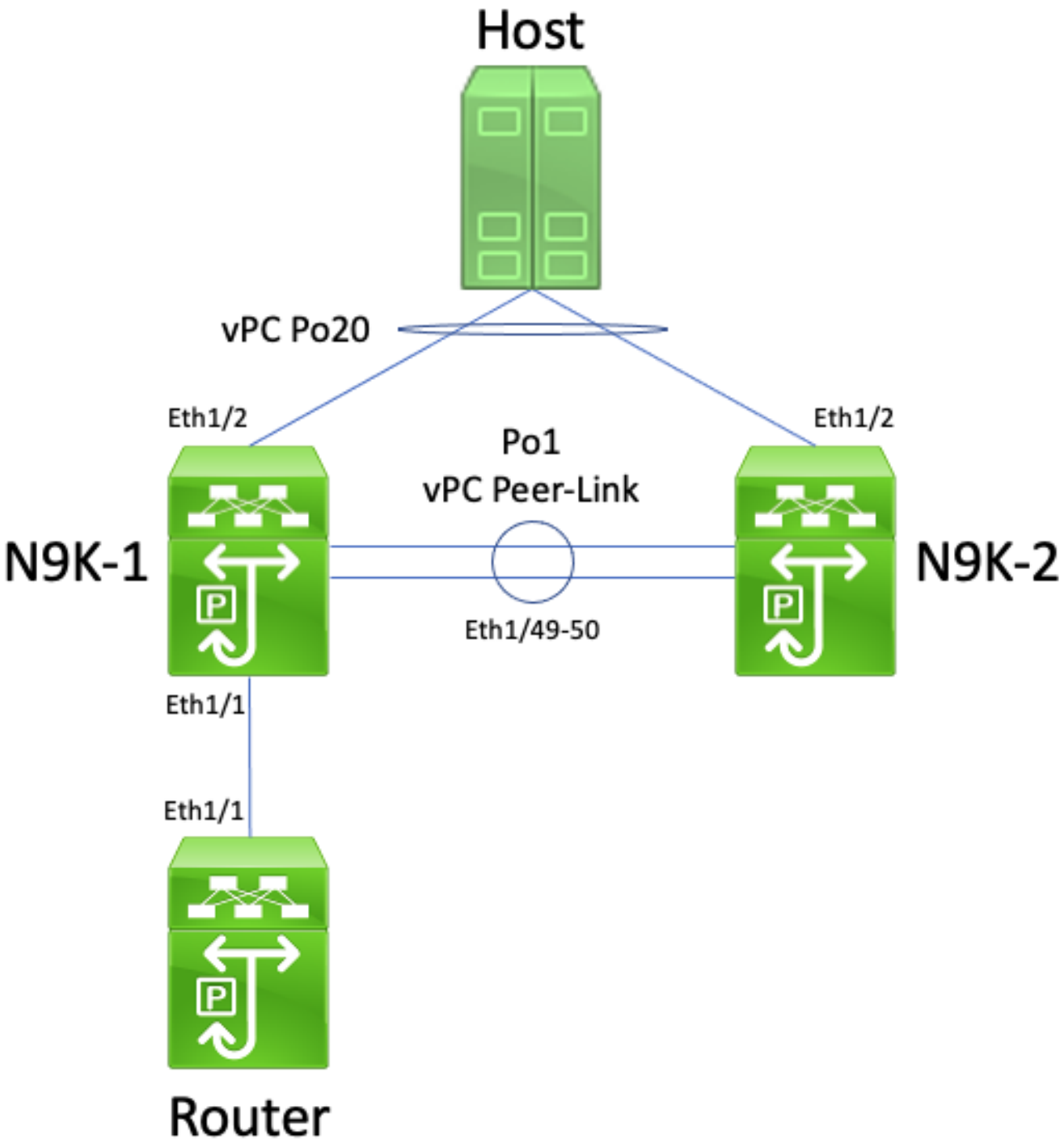
», en raison de l'activation de l'amélioration de la passerelle d'homologue vPC. Par conséquent, N9K-2 tente d'acheminer localement le paquet de protocole de routage de monodiffusion pour le compte de N9K-1.

Cependant, en acheminant le paquet, la durée de vie (TTL) du paquet est décrémentée et la durée de vie de la plupart des paquets de protocole de routage monodiffusion est de 1. Par conséquent, la durée de vie du paquet est décrémentée à 0 et abandonnée par N9K-2. Du point de vue de N9K-1, N9K-1 reçoit des paquets de protocole de routage de multidiffusion de lien local de Routeur et peut envoyer des paquets de protocole de routage de monodiffusion à Routeur, mais ne reçoit pas de paquets de protocole de routage de monodiffusion de Routeur. Par conséquent, N9K-1 supprime la contiguïté du protocole de routage avec le routeur et redémarre sa machine à état fini locale pour le protocole de routage. De même, le routeur redémarre sa machine à états finis locale pour le protocole de routage.

Vous pouvez résoudre ce problème en activant l'amélioration du routage/couche 3 sur vPC à l'aide de la commande de configuration du domaine vPC layer 3 peer-router. Cela permet aux paquets de protocole de routage de monodiffusion avec une TTL de 1 d'être transférés sur le lien homologue vPC sans décrémenter la TTL du paquet. Par conséquent, les contiguïtés de protocole de routage de monodiffusion peuvent être formées sur un vPC ou un VLAN vPC sans problème.

Contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC sans passerelle homologue vPC

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues VPC dans un domaine VPC où l'amélioration de la passerelle homologue VPC n'est pas activée. L'interface Po1 est le lien homologue vPC. Un routeur avec le nom d'hôte Routeur est connecté avec Ethernet1/1 au Ethernet1/1 de N9K-1. L'interface Ethernet1/1 de Routeur est une interface routée qui est activée avec un protocole de routage de monodiffusion. N9K-1 et N9K-2 ont tous deux des interfaces SVI activées avec le même protocole de routage de monodiffusion et sont dans le même domaine de diffusion que Routeur.

Les contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC sans que l'amélioration de la passerelle homologue vPC ne soit activée ne sont pas prises en charge, car la décision de hachage ECMP du routeur vPC connecté au VLAN peut amener N9K-2 à abandonner

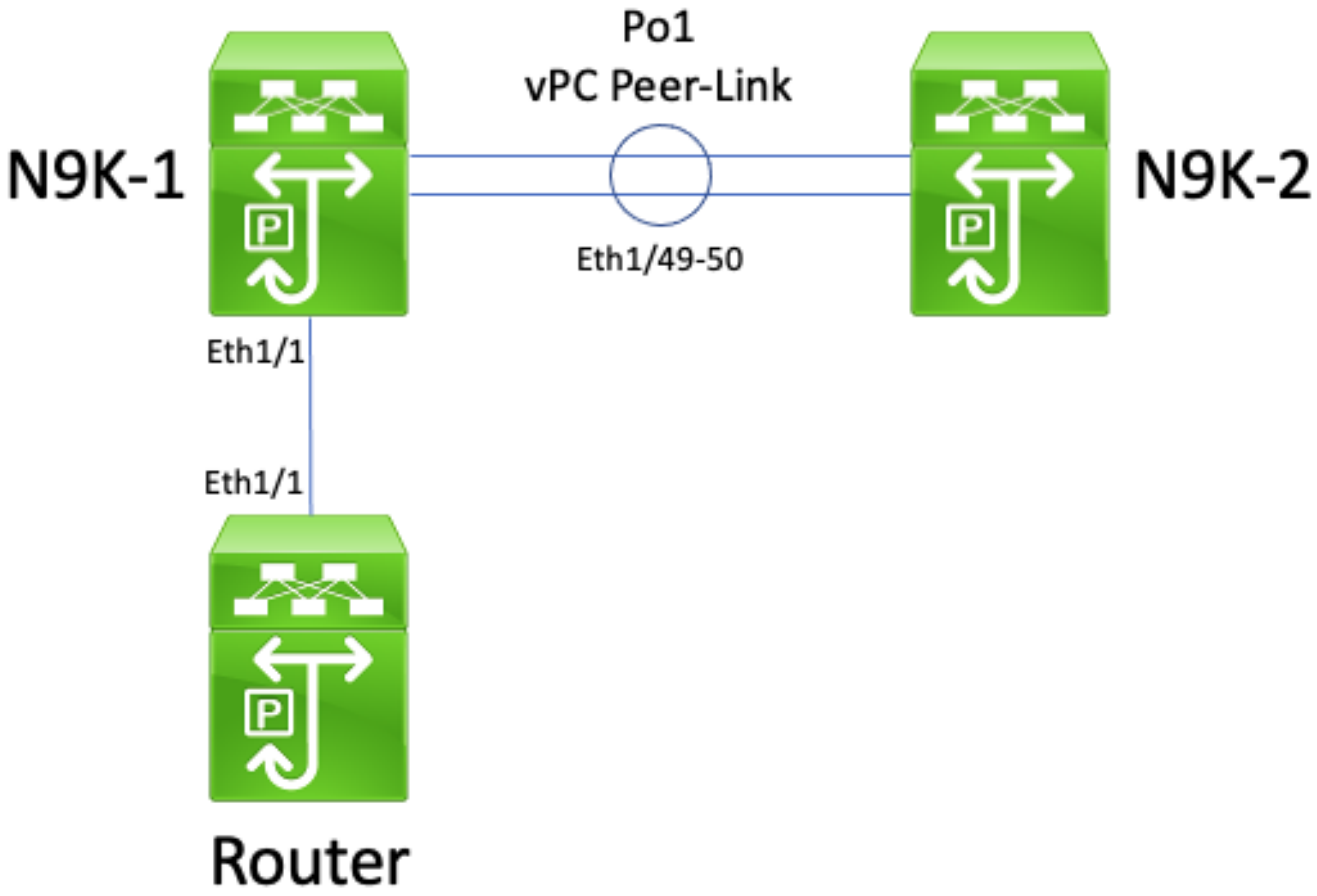
le trafic du plan de données pour violation de la règle de prévention des boucles vPC. Dans cette topologie, les contiguïtés de protocole de routage se formeraient correctement entre le routeur, N9K-1 et N9K-2. Examinez le flux de trafic entre Routeur et Hôte. Le trafic du plan de données traversant Routeur destiné à Hôte peut être réécrit avec une adresse MAC de destination appartenant à l'adresse MAC SVI de N9K-2 (en raison de la décision de hachage ECMP prise par le routeur) et sortir de l'interface Ethernet1/1 vers N9K-1.

N9K-1 reçoit ce paquet et le transfère via la liaison d'homologue vPC, car l'adresse MAC de destination appartient à N9K-2 et l'amélioration de la passerelle d'homologue vPC (qui permet à N9K-1 d'acheminer le paquet pour le compte de N9K-2) n'est pas activée. N9K-2 reçoit ce paquet sur le lien homologue vPC et reconnaît qu'il aurait besoin de transférer le paquet hors de son Ethernet1/2 dans le vPC Po20. Cela enfreint la règle de prévention des boucles vPC, de sorte que N9K-2 abandonne le paquet dans le matériel. Par conséquent, vous pourriez observer des problèmes de connectivité ou des pertes de paquets pour certains flux qui traversent le domaine vPC dans cette topologie.

Vous pouvez résoudre ce problème en activant l'amélioration de la passerelle homologue vPC avec la commande de configuration du domaine vPC peer-gateway, puis en activant l'amélioration de routage/couche 3 sur vPC avec la commande de configuration du domaine vPC layer3 peer-router. Pour minimiser les perturbations, vous devez activer les deux améliorations de vPC rapidement afin que le scénario de défaillance décrit dans les contiguïtés de protocole de routage de monodiffusion sur vPC avec passerelle homologue vPC n'ait pas le temps de se produire.

Contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC avec passerelle homologue vPC

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue VPC est activée. L'interface Po1 est le lien homologue vPC. Un routeur avec le nom d'hôte Routeur est connecté avec Ethernet1/1 au Ethernet1/1 de N9K-1. L'interface Ethernet1/1 de Routeur est une interface routée qui est activée avec un protocole de routage de monodiffusion. N9K-1 et N9K-2 ont tous deux des interfaces SVI activées avec le même protocole de routage de monodiffusion et sont dans le même domaine de diffusion que Routeur.

Les contiguïtés de protocole de routage de monodiffusion sur un VLAN vPC avec l'amélioration de la passerelle d'homologue vPC activée ne sont pas prises en charge car l'amélioration de la passerelle d'homologue vPC empêche la formation de contiguïtés de protocole de routage de monodiffusion entre le routeur connecté au VLAN vPC et l'homologue vPC auquel le routeur connecté au VLAN vPC n'est pas directement connecté. Dans cette topologie, une contiguïté de protocole de routage entre le routeur et N9K-2 ne s'établit pas comme prévu suite au routage de paquets de protocole de routage monodiffusion N9K-1 destinés à l'adresse MAC SVI de N9K-2 en raison de l'activation de l'amélioration de la passerelle d'homologue vPC. Comme les paquets sont acheminés, leur durée de vie (TTL) doit être décrémentée. Les paquets de protocole de routage de monodiffusion ont généralement une TTL de 1, et un routeur qui décrémente la TTL d'un paquet à 0 doit abandonner ce paquet.

Tous les routeurs sont en mesure d'envoyer et de recevoir des paquets de protocole de routage de multidiffusion de lien local (généralement nommés paquets « Hello ») sans problème, car ces paquets sont acheminés avec succès vers le VLAN vPC. Cependant, voici un scénario dans

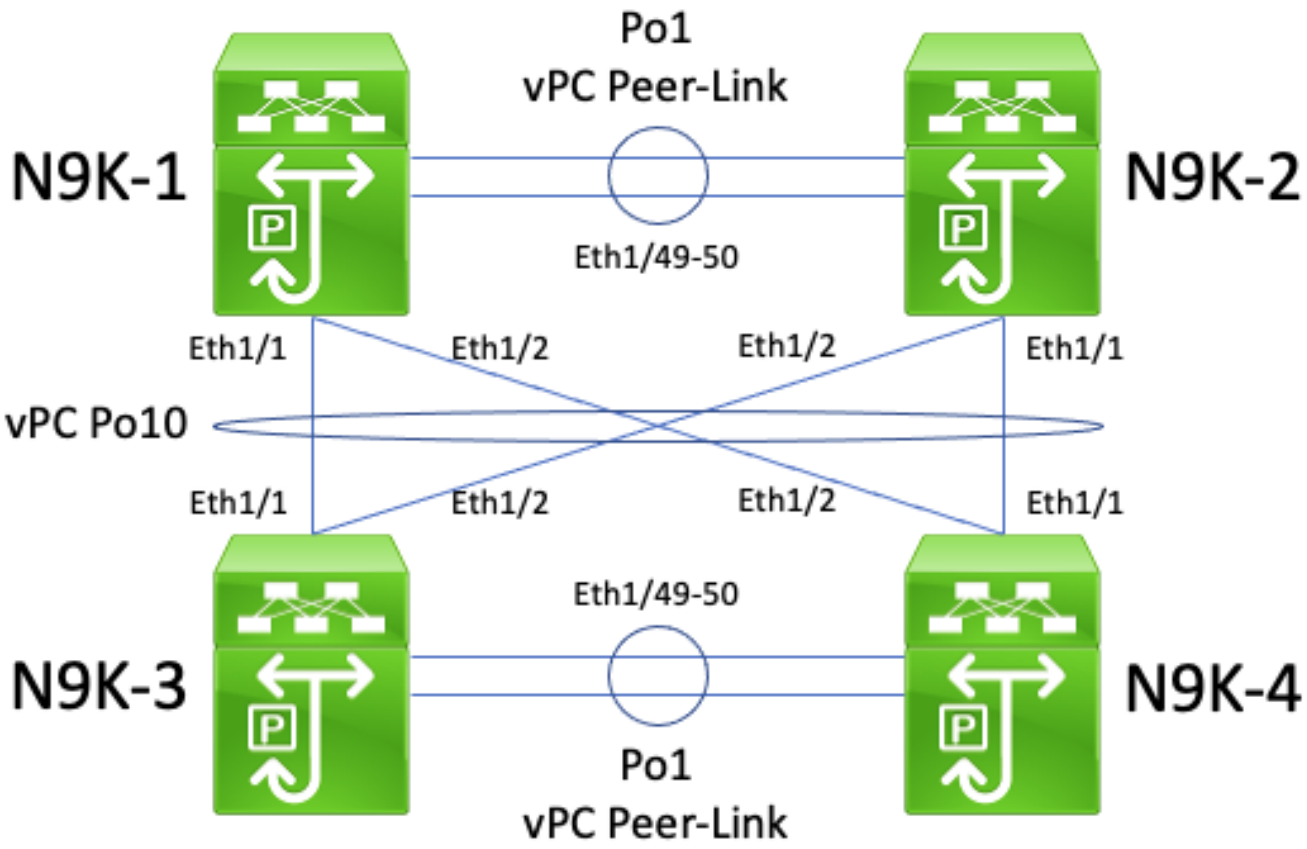
lequel un paquet de protocole de routage de monodiffusion provenant de Routeur destiné à N9K-2 sort d'Ethernet1/1 vers N9K-1. Ce paquet est destiné à l'adresse MAC SVI de N9K-2, mais entre dans l'interface Ethernet1/1 de N9K-1. N9K-1 constate que le paquet est destiné à l'adresse MAC SVI de N9K-2, qui est installée dans la table d'adresses MAC de N9K-1 avec l'indicateur « G », ou « Passerelle », en raison de l'activation de l'amélioration de la passerelle d'homologue vPC. Par conséquent, N9K-1 tente d'acheminer localement le paquet de protocole de routage de monodiffusion pour le compte de N9K-2.

Cependant, en acheminant le paquet, le TTL du paquet est décrémenté et le TTL de la plupart des paquets de protocole de routage monodiffusion est 1. Par conséquent, la durée de vie du paquet est décrémentée à 0 et abandonnée par N9K-1. Du point de vue de N9K-2, N9K-2 reçoit des paquets de protocole de routage de multidiffusion de lien local de Routeur et peut envoyer des paquets de protocole de routage de monodiffusion à Routeur, mais ne reçoit pas de paquets de protocole de routage de monodiffusion de Routeur. Par conséquent, N9K-2 supprime la contiguïté du protocole de routage avec le routeur et redémarre sa machine à état fini locale pour le protocole de routage. De même, le routeur redémarre sa machine à états finis locale pour le protocole de routage.

Vous pouvez résoudre ce problème en activant l'amélioration du routage/couche 3 sur vPC à l'aide de la commande de configuration du domaine vPC layer 3 peer-router. Cela permet aux paquets de protocole de routage de monodiffusion avec une TTL de 1 d'être transférés sur le lien homologue vPC sans décrémenter la TTL du paquet. Par conséquent, les contiguïtés de protocole de routage de monodiffusion peuvent être formées sur un vPC ou un VLAN vPC sans problème.

Contiguïtés de protocole de routage de monodiffusion sur vPC dos à dos avec passerelle homologue vPC

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue VPC est activée. Les commutateurs Nexus N9K-3 et N9K-4 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue vPC est activée. Les deux domaines vPC sont connectés l'un à l'autre par un vPC Po10 dos à dos. Les quatre commutateurs ont des interfaces SVI activées avec un protocole de routage de monodiffusion et se trouvent dans le même domaine de diffusion.

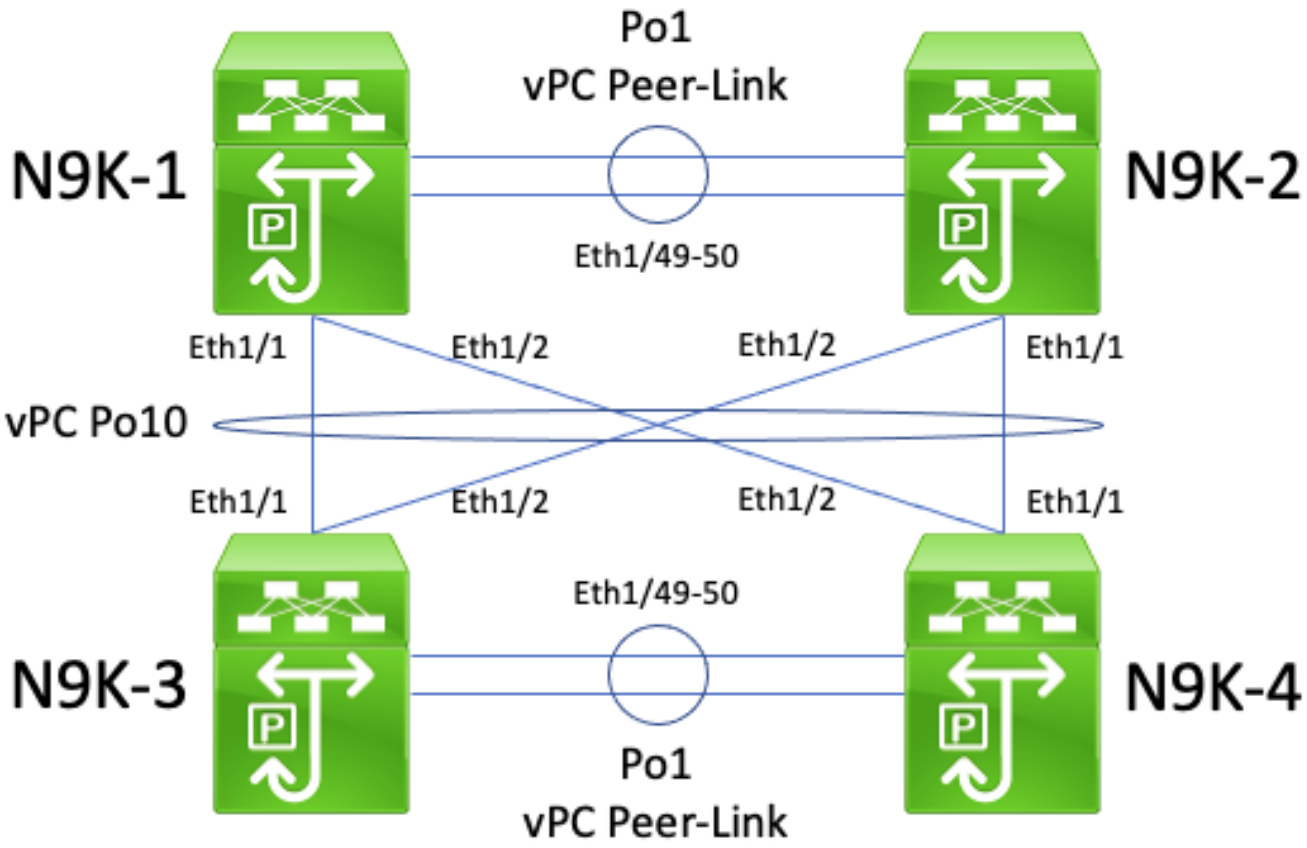
Les contiguïtés de protocole de routage de monodiffusion sur des vPC dos à dos avec l'amélioration de la passerelle homologue vPC activée ne sont pas prises en charge, car l'amélioration de la passerelle homologue vPC peut empêcher la formation de contiguïtés de protocole de routage de monodiffusion entre un domaine vPC et un autre domaine vPC. Dans cette topologie, une contiguïté de protocole de routage entre N9K-1 et N9K-3 ou N9K-4 (ou les deux) peut ne pas s'établir comme prévu. De même, une contiguïté de protocole de routage entre N9K-2 et N9K-3 ou N9K-4 (ou les deux) pourrait ne pas se produire comme prévu. Cela est causé par le fait que les paquets de protocole de routage de monodiffusion pourraient être destinés à un routeur (par exemple, N9K-3) mais être transférés à un routeur différent (par exemple, N9K-4) en fonction de la décision de hachage de canal de port de couche 2 du routeur d'origine.

La cause première de ce problème est identique à la cause première décrite dans la section Contiguïtés de protocole de routage de monodiffusion sur un vPC avec passerelle homologue vPC de ce document. Vous pouvez résoudre ce problème en activant l'amélioration du routage/couche 3 sur vPC à l'aide de la commande de configuration du domaine vPC layer 3 peer-router. Cela permet aux paquets de protocole de routage de monodiffusion avec une TTL de 1 d'être transférés sur le lien homologue vPC sans décrémenter la TTL du paquet. Par

conséquent, les contiguïtés de protocole de routage de monodiffusion peuvent être formées sur un vPC dos à dos sans problème.

Contiguïtés OSPF sur vPC avec passerelle homologue vPC où le préfixe est présent dans OSPF LSDB, mais pas dans le tableau de routage

Examinez la topologie suivante :



Dans cette topologie, les commutateurs Nexus N9K-1 et N9K-2 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue VPC est activée. Les commutateurs Nexus N9K-3 et N9K-4 sont des homologues vPC dans un domaine vPC où l'amélioration de la passerelle homologue vPC est activée. Les deux domaines vPC sont connectés l'un à l'autre par un vPC Po10 dos à dos. Les quatre commutateurs ont des interfaces SVI activées avec un protocole de routage de monodiffusion et se trouvent dans le même domaine de diffusion. N9K-4 est le routeur désigné OSPF (DR) pour le domaine de diffusion, tandis que N9K-3 est le routeur désigné de secours OSPF pour le domaine de diffusion.

Dans ce scénario, une contiguïté OSPF entre N9K-1 et N9K-3 passe à l'état FULL en raison des paquets OSPF de monodiffusion sortant de l'Ethernet1/1 des deux commutateurs. De même, une contiguïté OSPF entre N9K-2 et N9K-3 passe à l'état FULL en raison des paquets OSPF de monodiffusion sortant de l'Ethernet1/2 des deux commutateurs.

Cependant, une contiguïté OSPF entre N9K-1 et N9K-4 est bloquée dans un état EXSTART ou EXCHANGE en raison des paquets OSPF de monodiffusion sortant de l'Ethernet1/1 des deux commutateurs et abandonnés par N9K-2 et N9K-4 comme décrit dans la section Contiguïtés de

protocole de routage de monodiffusion sur vPC dos à dos avec passerelle homologue vPC de ce document. De même, une contiguïté OSPF entre N9K-2 et N9K-4 est bloquée dans un état EXSTART ou EXCHANGE en raison des paquets OSPF de monodiffusion sortant de l'Ethernet1/2 des deux commutateurs et abandonnés par N9K-1 et N9K-3 comme décrit dans la section Contiguïtés de protocole de routage de monodiffusion sur vPC dos à dos avec passerelle homologue vPC de ce document.

Par conséquent, N9K-1 et N9K-2 sont dans un état FULL avec le BDR pour le domaine de diffusion, mais sont dans un état EXSTART ou EXCHANGE avec le DR pour le domaine de diffusion. Le DR et le BDR d'un domaine de diffusion conservent une copie complète de la base de données d'état du lien (LSDB) OSPF, mais les routeurs OSPF DROTHER doivent être dans un état FULL avec le DR du domaine de diffusion afin d'installer les préfixes appris avec OSPF à partir du DR ou du BDR. Par conséquent, N9K-1 et N9K-2 semblent avoir des préfixes appris de N9K-3 et N9K-4 présents dans la LSDB OSPF, mais ces préfixes ne sont pas installés dans la table de routage de monodiffusion tant que N9K-1 et N9K-2 ne passent pas à l'état FULL avec N9K-4 (le DR pour le domaine de diffusion).

Vous pouvez résoudre ce problème en activant l'amélioration du routage/couche 3 sur vPC à l'aide de la commande de configuration du domaine vPC layer 3 peer-router. Cela permet aux paquets de protocole de routage de monodiffusion avec une TTL de 1 d'être transférés sur le lien homologue vPC sans décrémenter la TTL du paquet. Par conséquent, les contiguïtés de protocole de routage de monodiffusion peuvent être formées sur un vPC dos à dos sans problème. Par conséquent, N9K-1 et N9K-2 passent à l'état FULL avec N9K-4 (le DR pour le domaine de diffusion) et installent avec succès les préfixes appris de N9K-3 et N9K-4 via OSPF dans leurs tables de routage de monodiffusion respectives.

Informations connexes

- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.3(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.2(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 10.1(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 9.3(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 9.2(x)
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 9000, version 7.x
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 7000 8.x
- Guide de configuration des interfaces NX-OS de Cisco Nexus série 7000 7.x
- Guide de conception et de configuration : Bonnes pratiques pour les canaux de port virtuel (vPC) sur les commutateurs Cisco Nexus 7000
- Topologies prises en charge pour le routage sur canal de port virtuel sur les plateformes Nexus

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.