Configuration et dépannage de l'appairage de fabric VXLAN vPC pour NXOS

Table des matières

Introduction

Ce document décrit comment configurer et vérifier l'appairage de fabric vPC pour le flux de trafic NXOS et BUM.

Conditions préalables

Exigences

Cisco recommande de connaître les sujets suivants :

- vPC (virtual port-channel)
- Réseau local virtuel extensible (VXLAN)

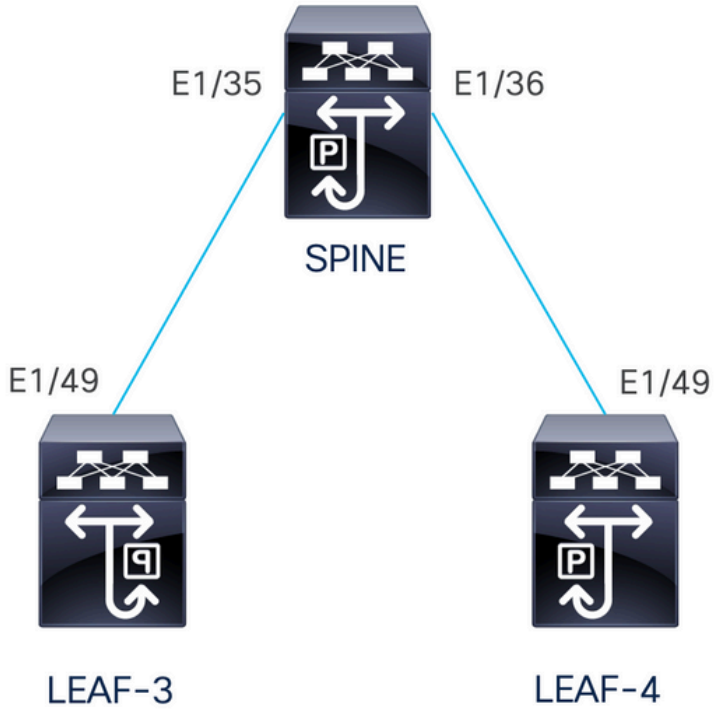
Composants utilisés

Les informations contenues dans ce document sont basées sur les versions de matériel et de logiciel suivantes :

- N9K-C93240YC-FX2 pour commutateurs Leaf Version : 10.3(3)
- N9K-C9336C-FX2 pour commutateur Spine Version : 10.3(3)

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.

Diagramme du réseau



L'appairage de fabric vPC fournit une solution d'accès à double hébergement améliorée sans gaspiller les ports physiques pour la liaison d'homologue vPC. Cette fonctionnalité préserve toutes les caractéristiques d'un vPC traditionnel.

Dans ce déploiement, nous avons Leaf-3 et Leaf-4 configurés comme vPC avec l'appairage de fabric.

Configuration

Configuration TCAM

Avant la configuration, il y a une vérification de la mémoire TCAM :

```
LEAF-4(config-if)# sh hardware access-list tcam region
                                NAT ACL[nat] size = 0
                     Ingress PACL [ing-ifacl] size = 0
                                  VACL [vacl] size = 0
                      Ingress RACL [ing-racl] size = 2304
                   Ingress L2 QOS [ing-l2-qos] size = 256
       Ingress L3/VLAN QOS [ing-l3-vlan-qos] size = 512
                        Ingress SUP [ing-sup] size = 512
 Ingress L2 SPAN filter [ing-l2-span-filter] size = 256
 Ingress L3 SPAN filter [ing-l3-span-filter] size = 256
                    Ingress FSTAT [ing-fstat] size = 0
                                  span [span] size = 512
                       Egress RACL [egr-racl] size = 1792
                         Egress SUP [egr-sup] size = 256
             Ingress Redirect [ing-redirect] size = 0
                  Egress L2 QOS [egr-l2-qos] size = 0
        Egress L3/VLAN QOS [egr-l3-vlan-qos] size = 0
     Ingress Netflow/Analytics [ing-netflow] size = 512
                        Ingress NBM [ing-nbm] size = 0
                         TCP NAT ACL[tcp-nat] size = 0
          Egress sup control plane[egr-copp] size = 0
   Ingress Flow Redirect [ing-flow-redirect] size = 0 <<<<<<<<
Ingress PACL IPv4 Lite [ing-ifacl-ipv4-lite] size = 0
Ingress PACL IPv6 Lite [ing-ifacl-ipv6-lite] size = 0
                  Ingress CNTACL [ing-cntacl] size = 0
                   Egress CNTACL [egr-cntacl] size = 0
                     MCAST NAT ACL[mcast-nat] size = 0
                      Ingress DACL [ing-dacl] size = 0
     Ingress PACL Super Bridge [ing-pacl-sb] size = 0
   Ingress Storm Control [ing-storm-control] size = 0
         Ingress VACL redirect [ing-vacl-nh] size = 0
                      Egress PACL [egr-ifacl] size = 0
                Egress Netflow [egr-netflow] size = 0
```

L'appairage de fabric vPC nécessite l'application de la découpe TCAM de la région ing-flow-redirect. Le découpage TCAM nécessite d'enregistrer la configuration et de recharger le commutateur avant d'utiliser la fonction.

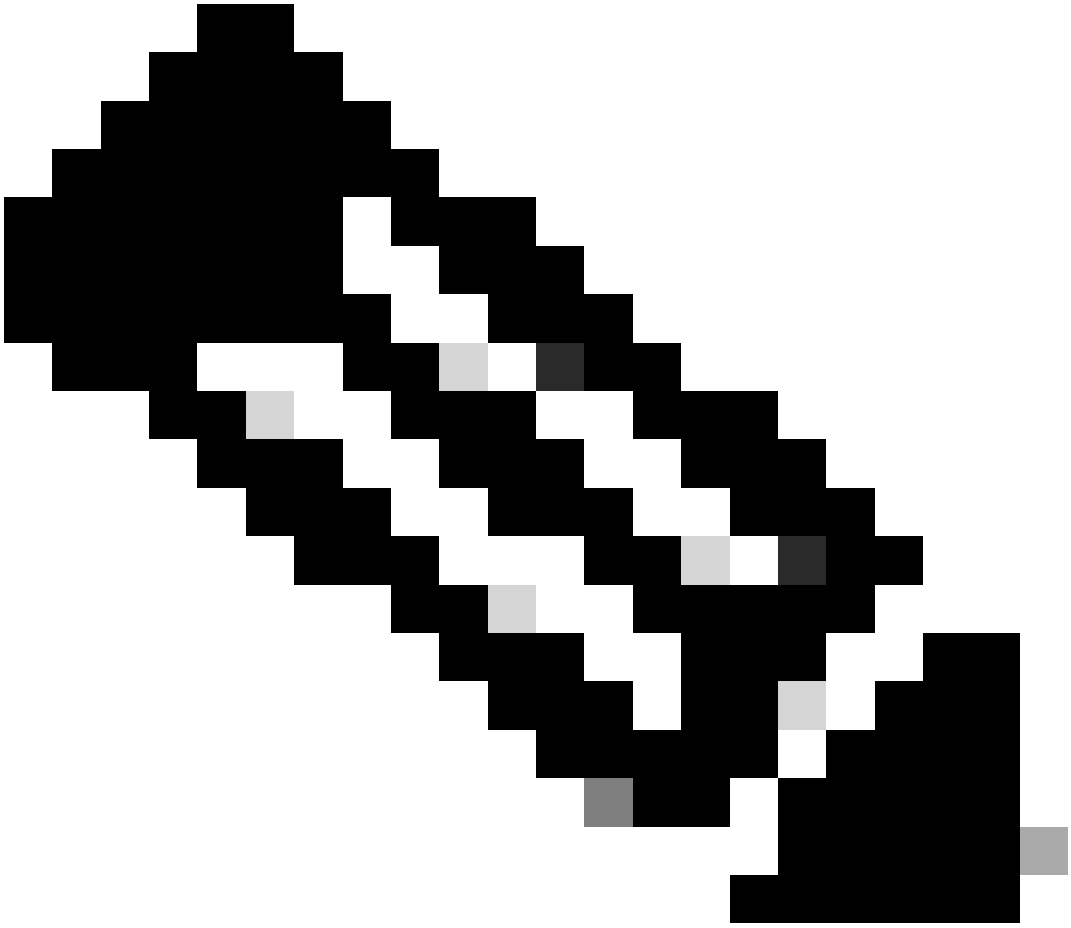
Cet espace sur la TCAM est double largeur, donc le minimum que nous pouvons attribuer est 512.

Sculpture TCAM

Dans ce scénario, ing-racl a assez d'espace pour prendre 512 et assigner ces 512 à ing-flow-redirect.

```
LEAF-4(config-if)# hardware access-list tcam region ing-racl 1792
Please save config and reload the system for the configuration to take effect

LEAF-4(config)# hardware access-list tcam region ing-flow-redirect 512
Please save config and reload the system for the configuration to take effect
```

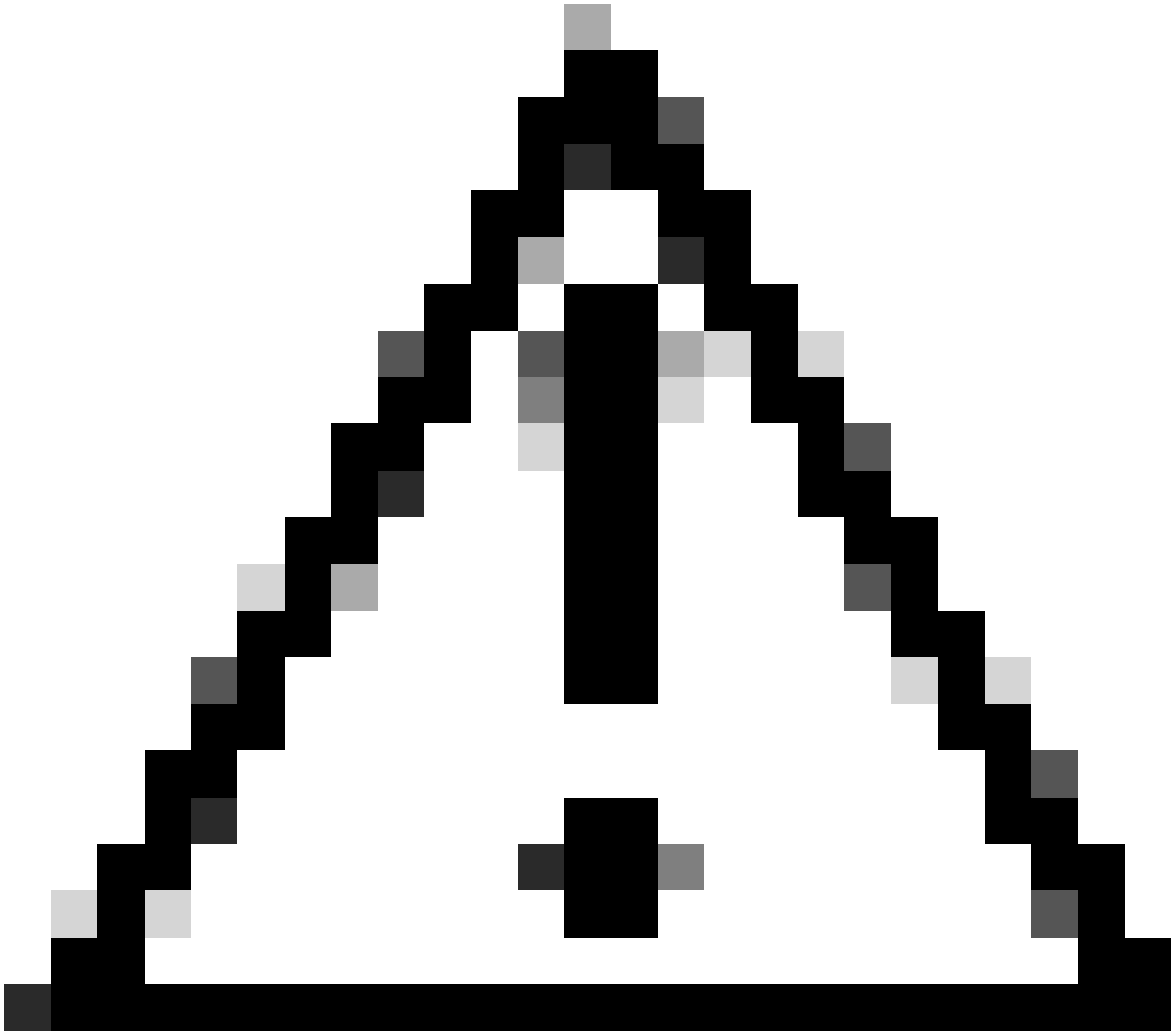


Remarque : lors de la configuration de l'appairage de fabric vPC via DCNM, le découpage TCAM va être effectué, mais il nécessite un rechargement pour prendre effet

Une fois la modification effectuée, elle sera reflétée dans la commande :

```
513E-B-11-N9K-C93240YC-FX2-4# sh hardware access-list tcam region
                                       NAT ACL[nat] size = 0
                             Ingress PACL [ing-ifacl] size = 0
                                       VACL [vacl] size = 0
                               Ingress RACL [ing-racl] size = 2304
                          Ingress L2 QOS [ing-l2-qos] size = 256
                Ingress L3/VLAN QOS [ing-l3-vlan-qos] size = 512
                               Ingress SUP [ing-sup] size = 512
          Ingress L2 SPAN filter [ing-l2-span-filter] size = 256
          Ingress L3 SPAN filter [ing-l3-span-filter] size = 256
                         Ingress FSTAT [ing-fstat] size = 0
                                       span [span] size = 512
                            Egress RACL [egr-racl] size = 1792
                              Egress SUP [egr-sup] size = 256
```

```
            Ingress Redirect [ing-redirect] size = 0
                 Egress L2 QOS [egr-l2-qos] size = 0
       Egress L3/VLAN QOS [egr-l3-vlan-qos] size = 0
     Ingress Netflow/Analytics [ing-netflow] size = 512 <<<<<
                    Ingress NBM [ing-nbm] size = 0
                     TCP NAT ACL[tcp-nat] size = 0
         Egress sup control plane[egr-copp] size = 0
   Ingress Flow Redirect [ing-flow-redirect] size = 0
Ingress PACL IPv4 Lite [ing-ifacl-ipv4-lite] size = 0
Ingress PACL IPv6 Lite [ing-ifacl-ipv6-lite] size = 0
                Ingress CNTACL [ing-cntacl] size = 0
                 Egress CNTACL [egr-cntacl] size = 0
                   MCAST NAT ACL[mcast-nat] size = 0
                    Ingress DACL [ing-dacl] size = 0
     Ingress PACL Super Bridge [ing-pacl-sb] size = 0
   Ingress Storm Control [ing-storm-control] size = 0
         Ingress VACL redirect [ing-vacl-nh] size = 0
                    Egress PACL [egr-ifacl] size = 0
```



Attention : assurez-vous que le périphérique est rechargé après les modifications sur la TCAM, sinon le VPC ne va pas apparaître en raison de modifications non appliquées sur

---

la TCAM.

Configuration pour vPC

Domaine VPC

Sur les LEAF-3 et LEAF-4 dans le domaine VPC, la configuration consiste à spécifier les adresses IP pour le keepalive et la liaison d'homologue virtuel

```
vpc domain 1
  peer-keepalive destination 192.168.1.1 source 192.168.1.2 vrf management
  virtual peer-link destination 10.10.10.2 source 10.10.10.1 dscp 56

interface port-channel1
  vpc peer-link
```

Maintien en vie

Toute liaison directe de couche 3 entre des homologues vPC ne doit être utilisée que pour le maintien de la connexion entre homologues. Il doit se trouver dans un VRF distinct dédié au maintien de la connexion uniquement. Dans ce scénario, nous utilisons la gestion d'interface du commutateur.

```
LEAF-3
interface mgmt0
  vrf member management
  ip address 192.168.1.1/24

LEAF-4
interface mgmt0
  vrf member management
  ip address 192.168.1.2/24
```

Interface de couche 3 pour la liaison entre homologues virtuels

L'interface de couche 3 utilisée pour la liaison homologue virtuelle ne doit pas être la même que celle utilisée pour le test d'activité, vous pouvez utiliser le même bouclage utilisé pour le sous-réseau ou il peut s'agir d'un bouclage dédié sur le Nexus

Ici, le loopback0 est pour le sous-réseau et le loopback2 est un loopback dédié pour la liaison homologue virtuelle, tandis que le loopback1 est l'interface associée à notre interface NVE.

```
LEAF-3
interface loopback0
  ip address 10.1.1.1/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  ip address 172.16.1.2/32
  ip address 172.16.1.1/32 secondary
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback2
  ip address 10.10.10.2/32
  ip router ospf 1 area 0.0.0.0


LEAF-4
interface loopback0
  ip address 10.1.1.2/32
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback1
  ip address 172.16.1.3/32
  ip address 172.16.1.1/32 secondary
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode

interface loopback2
  ip address 10.10.10.1/32
  ip router ospf 1 area 0.0.0.0
```

Peer-link VPC

Un port-channel doit être attribué à la liaison homologue même si nous n'allons pas attribuer une interface physique au port-channel.

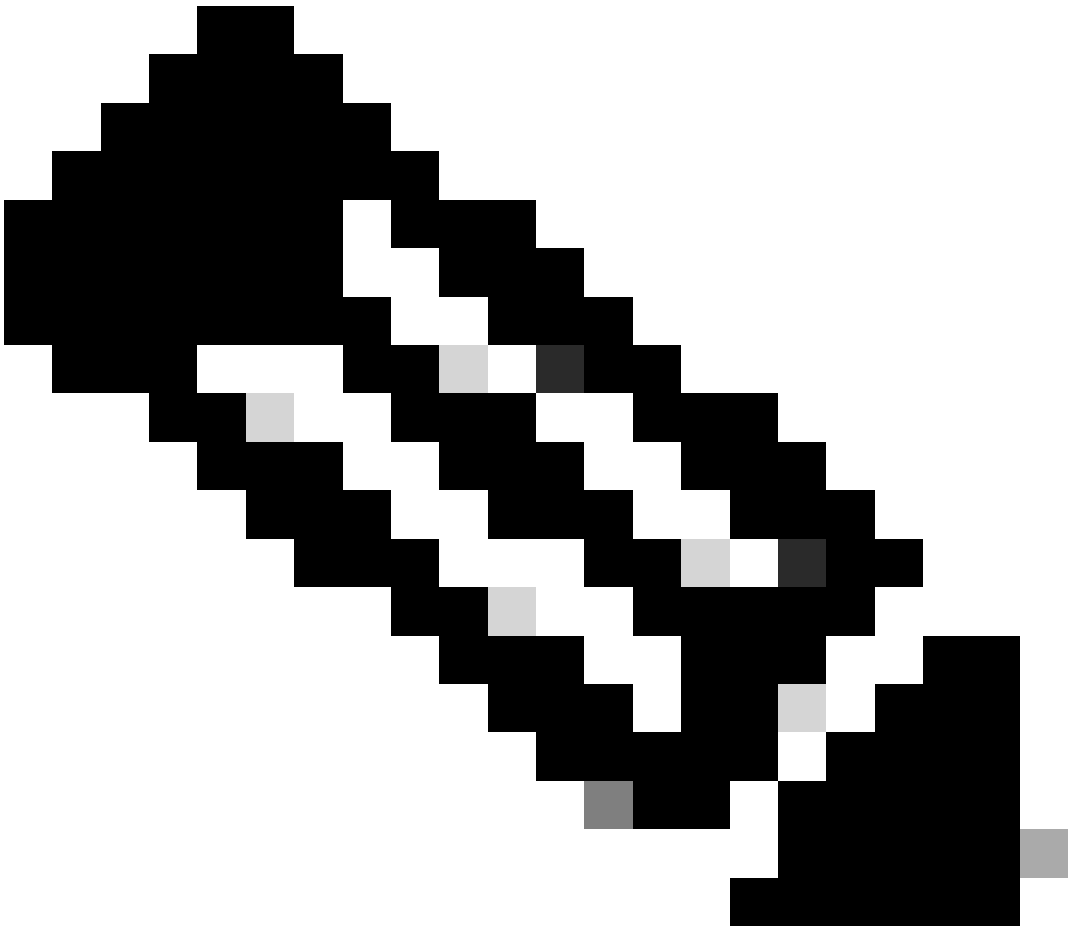
```
LEAF-3(config-if)# sh run interface port-channel 1 membership

interface port-channel1
  switchport
  switchport mode trunk
  spanning-tree port type network
  vpc peer-link
```

Liaisons ascendantes

La dernière partie de la configuration consiste à configurer les liaisons sur les deux leafs vers le SPINE avec la commande port-type fabric.

```
interface Ethernet1/49
  port-type fabric <<<<<<<<
  medium p2p
  ip unnumbered loopback0
  ip router ospf 1 area 0.0.0.0
  ip pim sparse-mode
  no shutdown
```



Remarque : si vous ne configurez pas le fabric de type de port, vous ne pouvez pas voir le keepalive généré par le Nexus

---

Configuration SPINES

Sur les spines, il est recommandé de définir la QoS pour qu'elle corresponde à la valeur DSCP configurée sur le domaine VPC puisque la liaison homologue d'appairage de fabric vPC est établie sur le réseau de transport.

Les messages CFS d'informations de plan de contrôle utilisés pour synchroniser les informations d'état de port, les informations VLAN, le mappage VLAN à VNI, les adresses MAC d'hôte et les groupes de surveillance IGMP sont transmis sur le fabric. Les messages CFS sont marqués avec la valeur DSCP appropriée, qui doit être protégée dans le réseau de transport.

```
class-map type qos match-all CFS
  match dscp 56

policy-map type qos CFS
  class CFS
    Set qos-group 7 <<< Depending on the platform it can be 4

interface Ethernet 1/35-36
  service-policy type qos input CFS
```

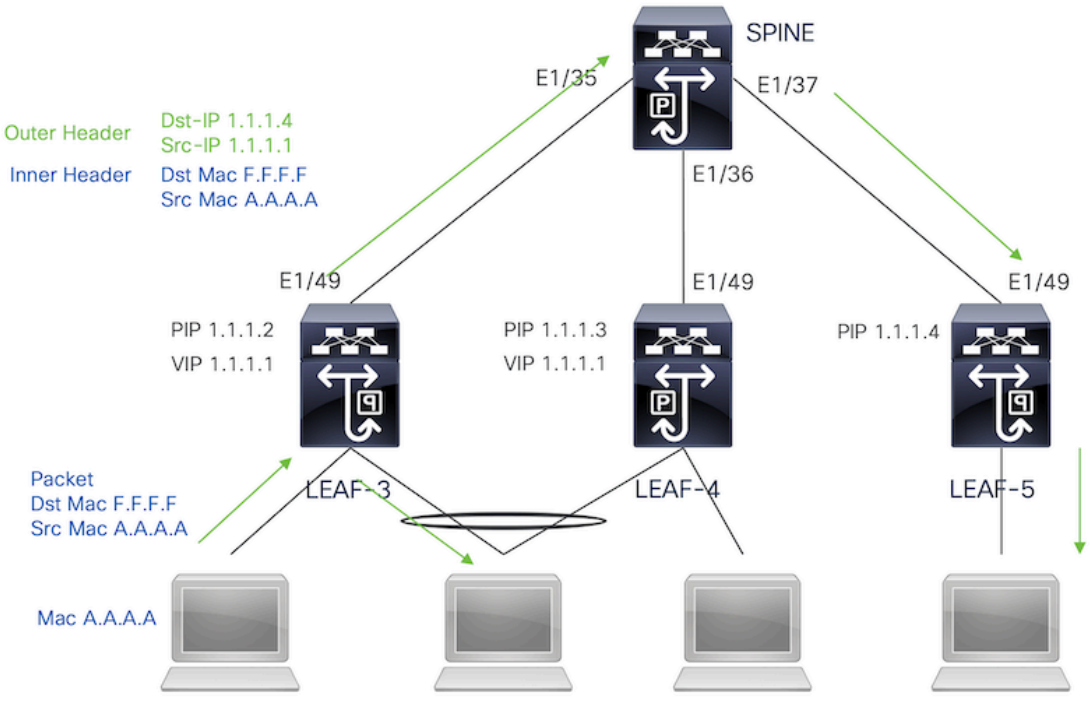
Trafic de diffusion, de monodiffusion inconnue et de multidiffusion avec encapsulation de réplication en entrée

Lorsque le nexus reçoit un paquet qui doit être diffusé, il génère 2 copies du paquet.

1. À tous les VTEP distants dans la liste de diffusion pour le VNI, y compris les ports d'accès locaux

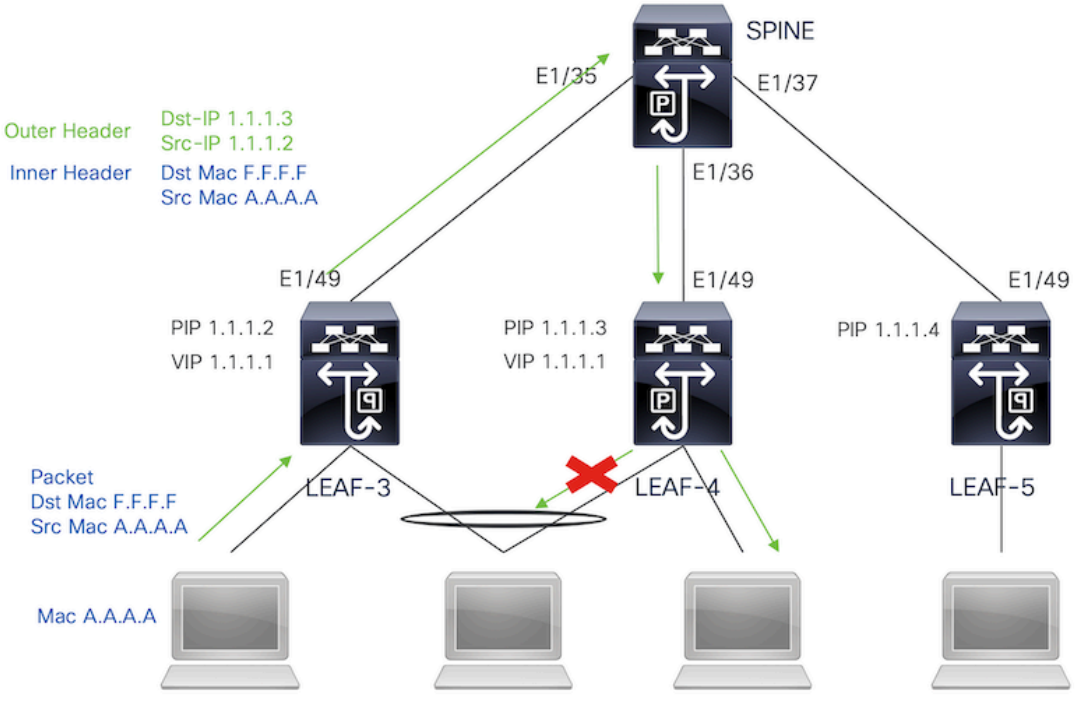
2. À l'homologue VPC distant

Pour la première copie, le Nexus a encapsulé le trafic en utilisant l'adresse IP source de l'adresse IP secondaire et l'adresse IP de destination du VTEP distant, ainsi que vers les ports d'accès locaux.



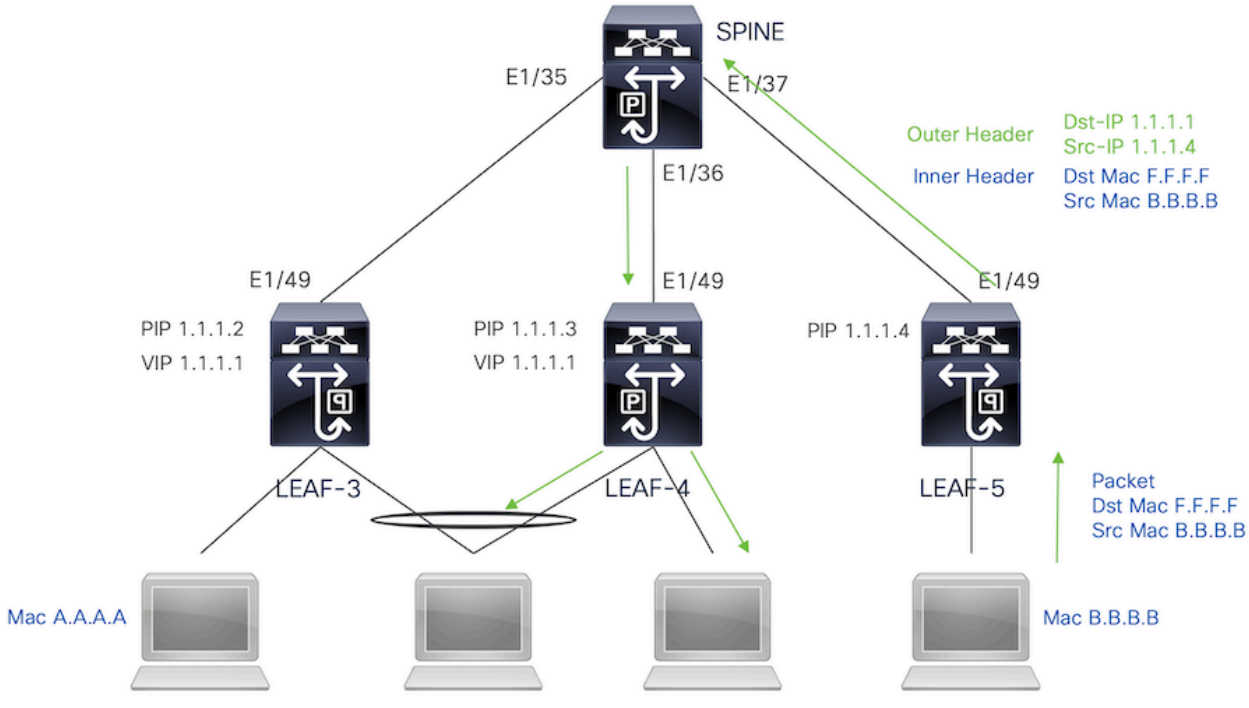
Pour la deuxième copie, elle sera envoyée à l'homologue VPC distant, l'IP source sera la principale adresse de bouclage et l'IP de destination sera celle de l'homologue VPC distant.

Une fois reçu le paquet de la colonne vertébrale, le VTEP distant ne va transférer le paquet qu'aux ports orphelins.



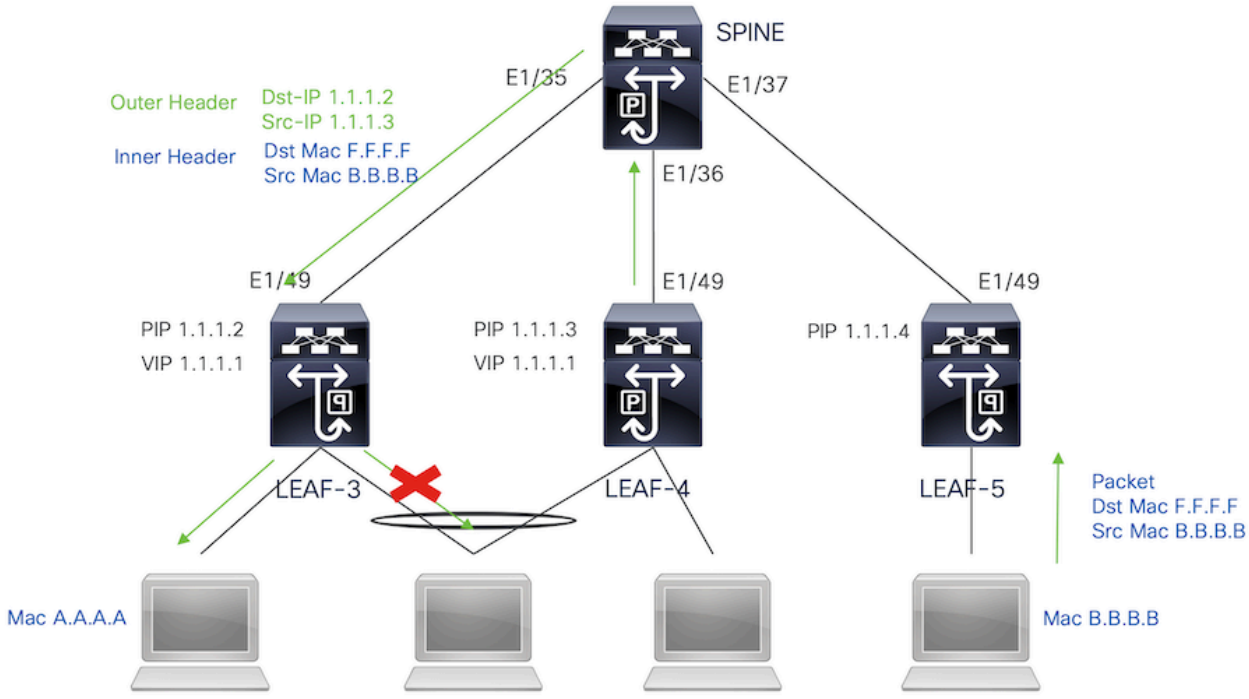
Trafic de diffusion, de monodiffusion inconnue et de multidiffusion avec décapsulation de réplication en entrée

Puisque l'IP de destination pour le trafic BUM reçu d'un autre VTEP est le VIP que le trafic hache vers l'un des périphériques VPC, il décapsule le paquet et l'envoie aux ports d'accès.



Pour que le trafic atteigne les ports orphelins connectés sur l'homologue VPC distant, le nexus génère une copie du paquet et va l'envoyer uniquement au VPC distant en utilisant l'adresse IP principale comme adresse IP source/de destination.

Une fois reçu sur l'homologue vpc distant, le nexus décapsule le trafic et le transfère uniquement aux ports orphelins.



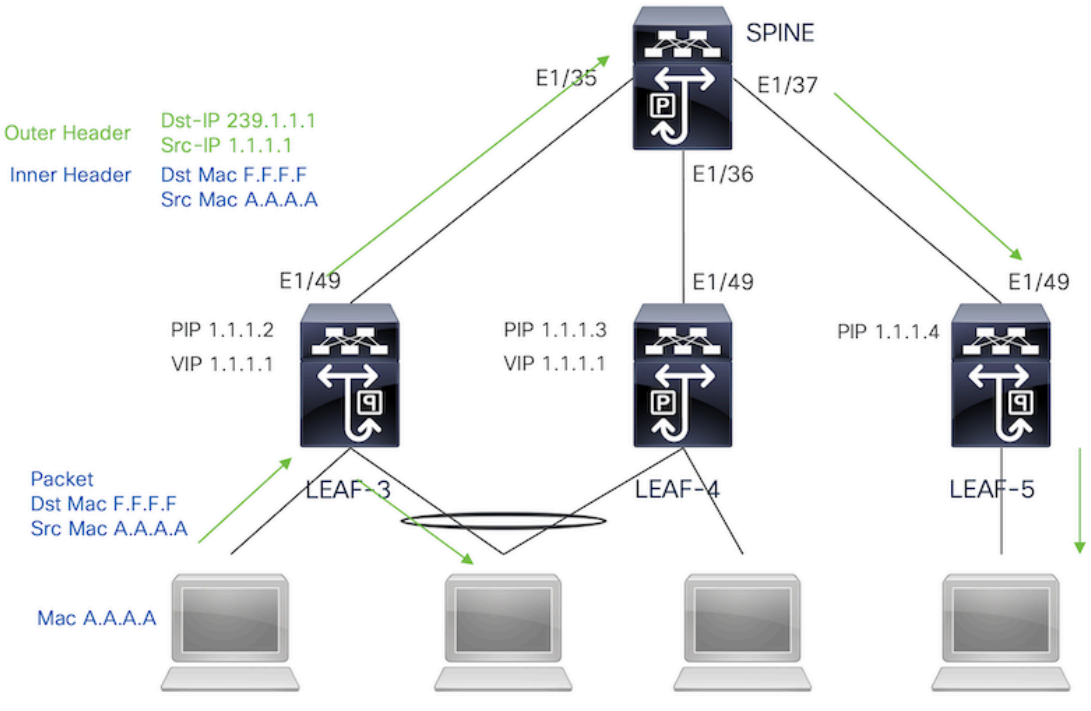
Trafic de diffusion, de monodiffusion inconnue et de multidiffusion avec encapsulation multidiffusion

Lorsque le nexus reçoit un paquet qui doit être diffusé, il génère 2 copies du paquet.

1. Le paquet va être envoyé à tous les OIF de l'entrée S, G de multidiffusion, y compris les ports d'accès locaux

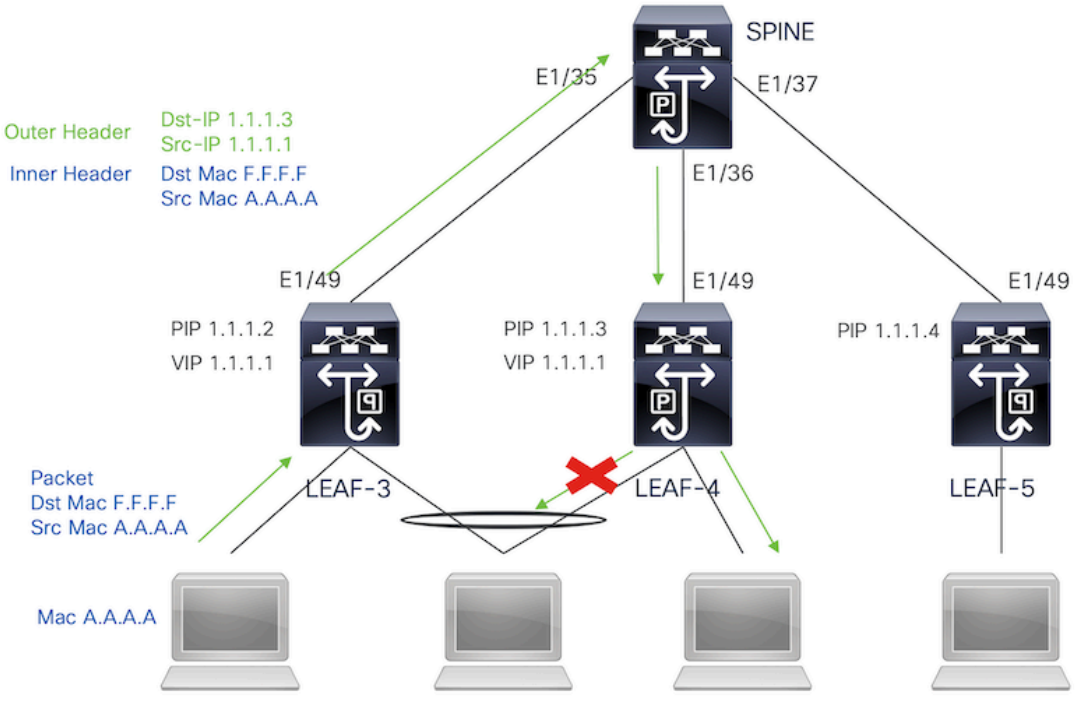
2. À l'homologue VPC distant

Pour la première copie, le Nexus a encapsulé le trafic en utilisant l'adresse IP source de l'adresse IP secondaire et l'adresse IP de destination du groupe de multidiffusion configuré.



Pour la deuxième copie, elle sera envoyée à l'homologue VPC distant, l'IP source sera le secondaire du bouclage et l'IP de destination sera le PIP de l'homologue VPC distant.

Une fois le paquet reçu de la colonne vertébrale, le VTEP distant ne transfère le paquet qu'aux ports orphelins.

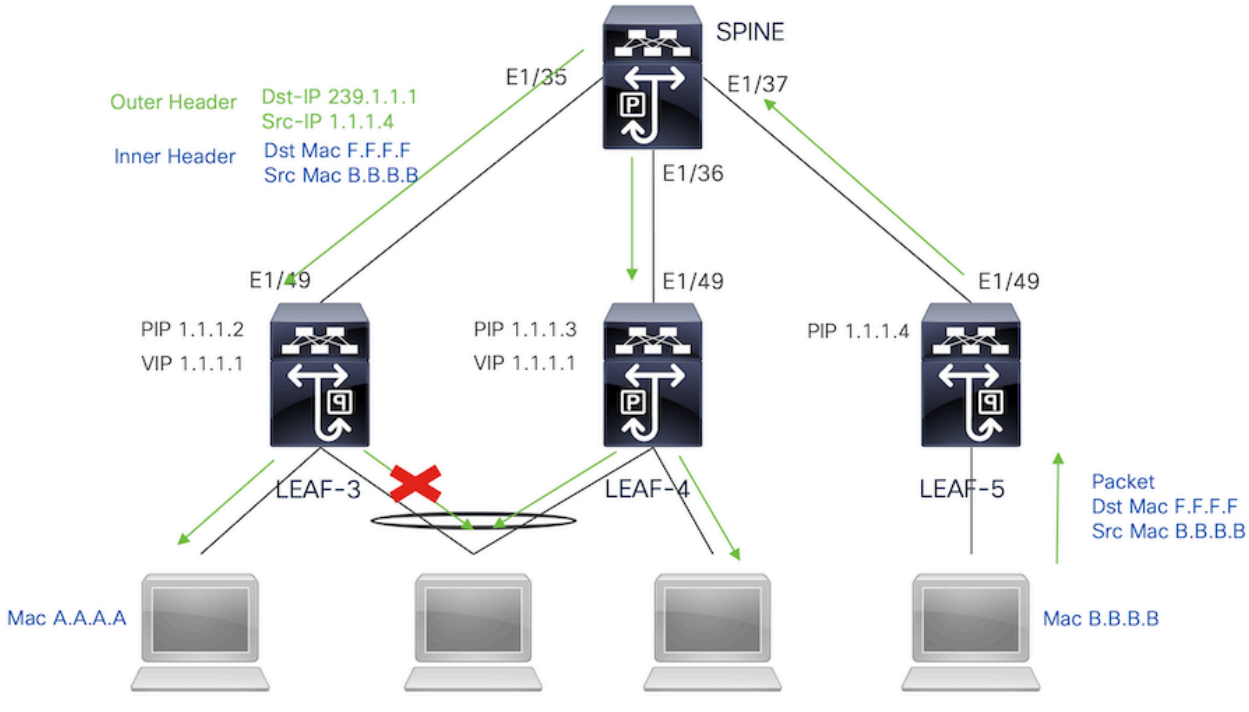


Trafic de diffusion, de monodiffusion inconnue et de multidiffusion avec décapsulation multidiffusion

Pour le processus de décapsulation, le paquet va arriver aux deux homologues VPC. Un seul périphérique VPC va transférer le trafic via les canaux de port VPC. Cette décision sera prise par le redirecteur affiché dans la commande.

```
module-1# show forwarding internal vpc-df-hash

VPC DF: FORWARDER
```



Vérifier

Pour vous assurer que le VPC est actif, exécutez les commandes suivantes :

Vérifiez l'accessibilité des adresses IP utilisées pour la liaison entre homologues virtuels.

```
LEAF-3# sh ip route 10.10.10.1
IP Route Table for VRF "default"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

10.10.10.1/32, ubest/mbest: 1/0
    *via 192.168.120.1, Eth1/49, [110/3], 01:15:01, ospf-1, intra

LEAF-3# ping 10.10.10.1
PING 10.10.10.1 (10.10.10.1): 56 data bytes
64 bytes from 10.10.10.1: icmp_seq=0 ttl=253 time=0.898 ms
64 bytes from 10.10.10.1: icmp_seq=1 ttl=253 time=0.505 ms
64 bytes from 10.10.10.1: icmp_seq=2 ttl=253 time=0.433 ms
64 bytes from 10.10.10.1: icmp_seq=3 ttl=253 time=0.465 ms
64 bytes from 10.10.10.1: icmp_seq=4 ttl=253 time=0.558 ms
```

```
LEAF-3(config-if)# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 1
Peer status                       : peer adjacency formed ok <<<<
```

```
vPC keep-alive status                 : peer is alive <<<<
Configuration consistency status      : success
Per-vlan consistency status           : success
Type-2 consistency status             : success
vPC role                              : secondary
Number of vPCs configured             : 0
Peer Gateway                          : Disabled
Dual-active excluded VLANs            : -
Graceful Consistency Check            : Enabled
Auto-recovery status                  : Disabled
Delay-restore status                  : Timer is off.(timeout = 30s)
Delay-restore SVI status              : Timer is off.(timeout = 10s)
Delay-restore Orphan-port status      : Timer is off.(timeout = 0s)
Operational Layer3 Peer-router        : Disabled
Virtual-peerlink mode                 : Enabled <<<<<<<

vPC Peer-link status
---------------------------------------------------------------------
id  Port  Status  Active vlans
--  ----  ------  --------------------------------------------------
1   Po1   up      1,10,50,600-604,608,610-611,614-618,638-639,
                  662-663,701-704
```

Pour vérifier les rôles du VPC, exécutez la commande suivante :

```
LEAF-3(config-if)# sh 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                  : d0:e0:42:e2:09:6f
vPC local role-priority               : 32667
vPC local config role-priority        : 32667
vPC peer system-mac                   : 2c:4f:52:3f:46:df
vPC peer role-priority                : 32667
vPC peer config role-priority         : 32667
```

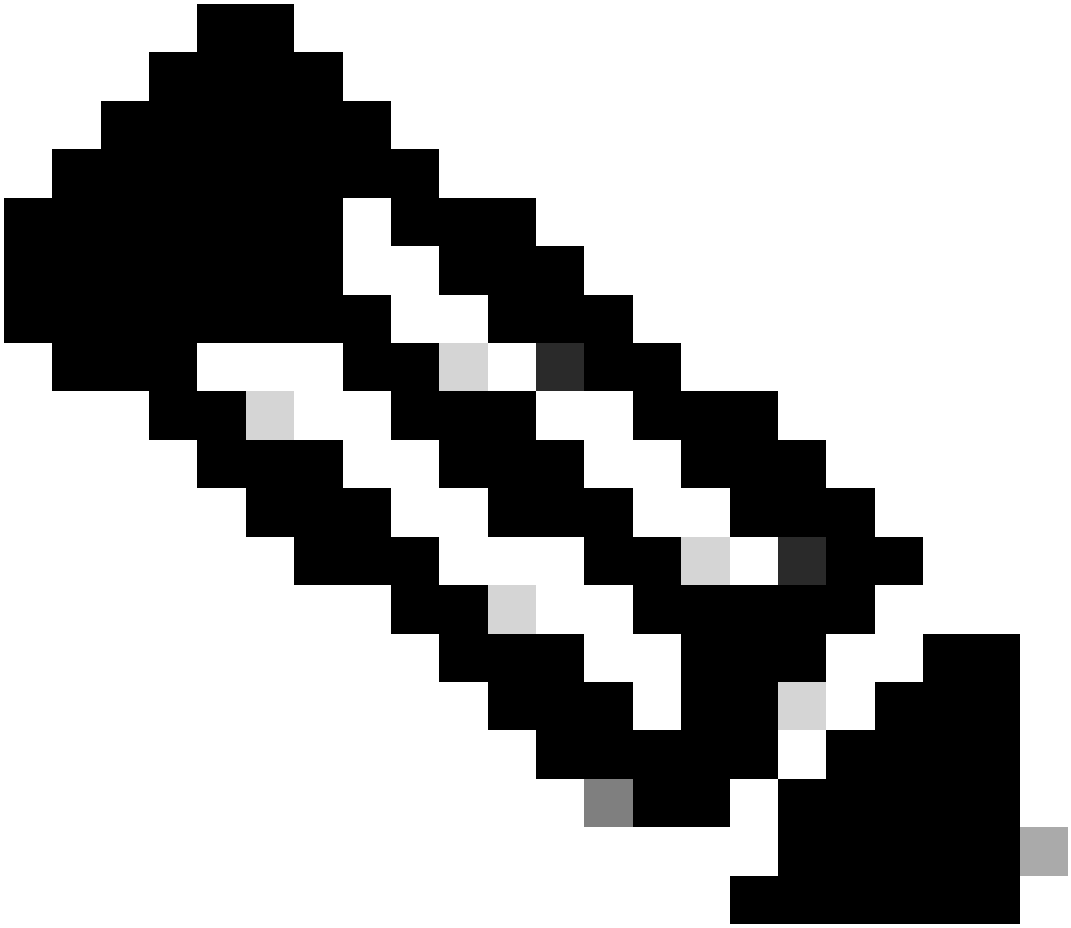
Tous les VLAN autorisés dans le port-channel de liaison homologue doivent être mappés à un VNI, au cas où ils ne seraient pas affichés comme incohérents

```
LEAF-3(config-if)# show vpc virtual-peerlink vlan consistency
Following vlans are inconsistent
1 608 610 611 614 615 616 617 618 638 639 701 702 703 704
```

Pour confirmer que la configuration des liaisons ascendantes est correctement programmée, exécutez la commande suivante :

```
LEAF-3(config-if)# show vpc fabric-ports
Number of Fabric port : 1
Number of Fabric port active : 1

Fabric   Ports      State
-------------------------------------
Ethernet 1/49       UP
```



Remarque : le NVE et l'interface de bouclage qui lui est associée vont s'afficher à moins que le VPC ne soit activé.

Informations connexes

- Assistance technique de Cisco et téléchargements

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.