Konfiguration und Fehlerbehebung für VXLAN vPC Fabric Peering für NX-OS

Inhalt

Einleitung

In diesem Dokument wird beschrieben, wie vPC Fabric-Peering für NXOS- und BUM-Datenverkehrsfluss konfiguriert und verifiziert wird.

Voraussetzungen

Anforderungen

Cisco empfiehlt, sich mit folgenden Themen vertraut zu machen:

- vPC (virtueller Port-Channel)
- Virtual Extensible LAN (VXLAN)

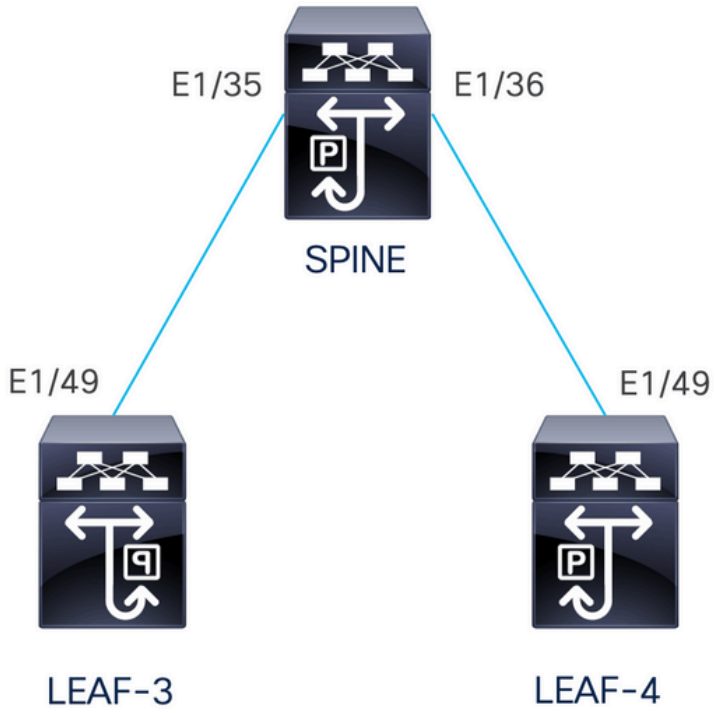
Verwendete Komponenten

Die Informationen in diesem Dokument basierend auf folgenden Software- und Hardware-Versionen:

- N9K-C93240YC-FX2 für Leaf-Switches Version: 10.3(3)
- N9K-C9336C-FX2 für Spine-Switch Version: 10.3(3)

Die Informationen in diesem Dokument beziehen sich auf Geräte in einer speziell eingerichteten Testumgebung. Alle Geräte, die in diesem Dokument benutzt wurden, begannen mit einer gelöschten (Nichterfüllungs) Konfiguration. Wenn Ihr Netzwerk in Betrieb ist, stellen Sie sicher, dass Sie die möglichen Auswirkungen aller Befehle kennen.

Netzwerkdiagramm



vPC Fabric Peering bietet eine erweiterte Dual-Homing-Zugriffslösung, ohne dass physische Ports für vPC Peer Link vergeudet werden müssen. Mit dieser Funktion bleiben alle Merkmale eines herkömmlichen vPC erhalten.

In dieser Bereitstellung sind Leaf-3 und Leaf-4 als vPC mit Fabric-Peering konfiguriert.

Konfiguration

TCAM-Konfiguration

Vor der Konfiguration wird der TCAM-Speicher überprüft:

```
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
```

vPC Fabric-Peering erfordert die Anwendung des TCAM-Carving für die Region "ing-flow-redirect". Für das Erstellen von TCAMs muss die Konfiguration gespeichert und der Switch neu geladen werden, bevor die Funktion verwendet werden kann.

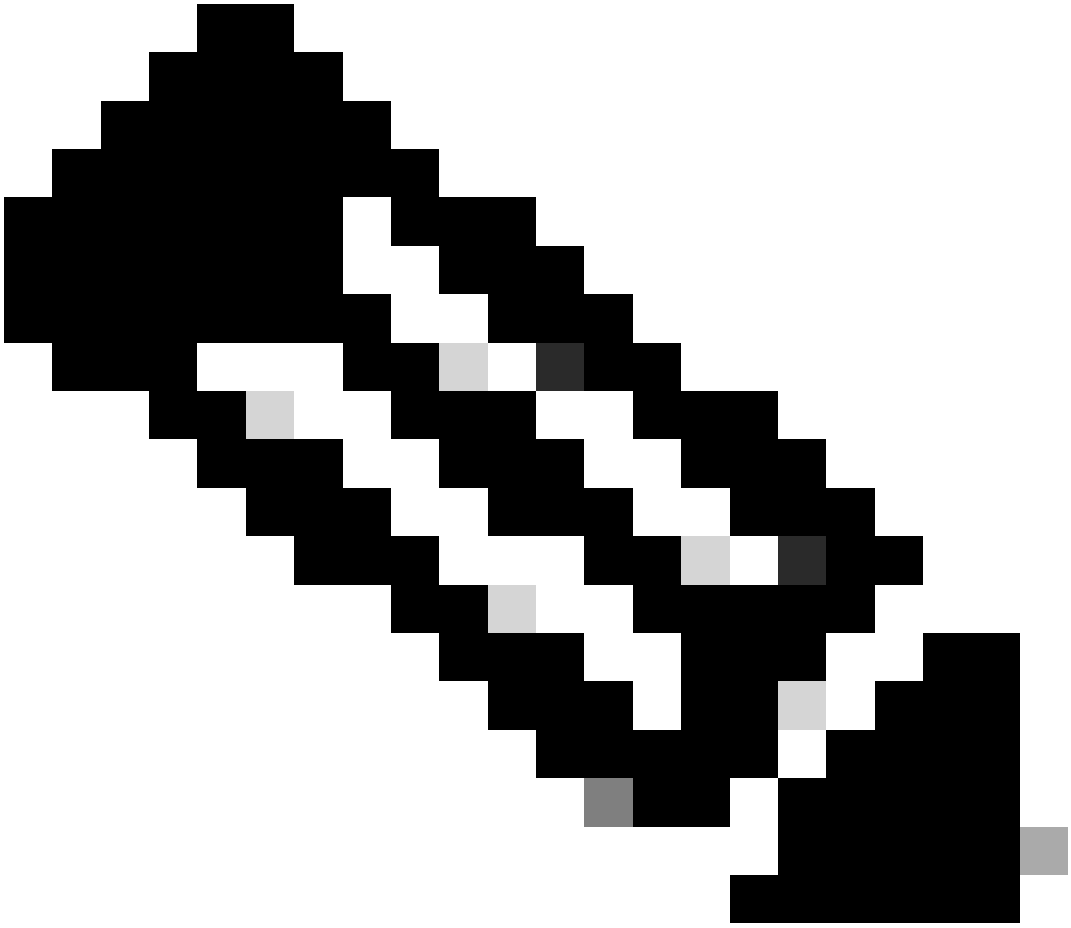
Der Abstand auf dem TCAM ist doppelt so groß, sodass mindestens 512 zugewiesen werden können.

TCAM-Carving

In diesem Szenario hat ing-racl genügend Speicherplatz, um 512 zu verwenden und diese 512 zu ing-flow-redirect zuzuweisen.

```
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
```

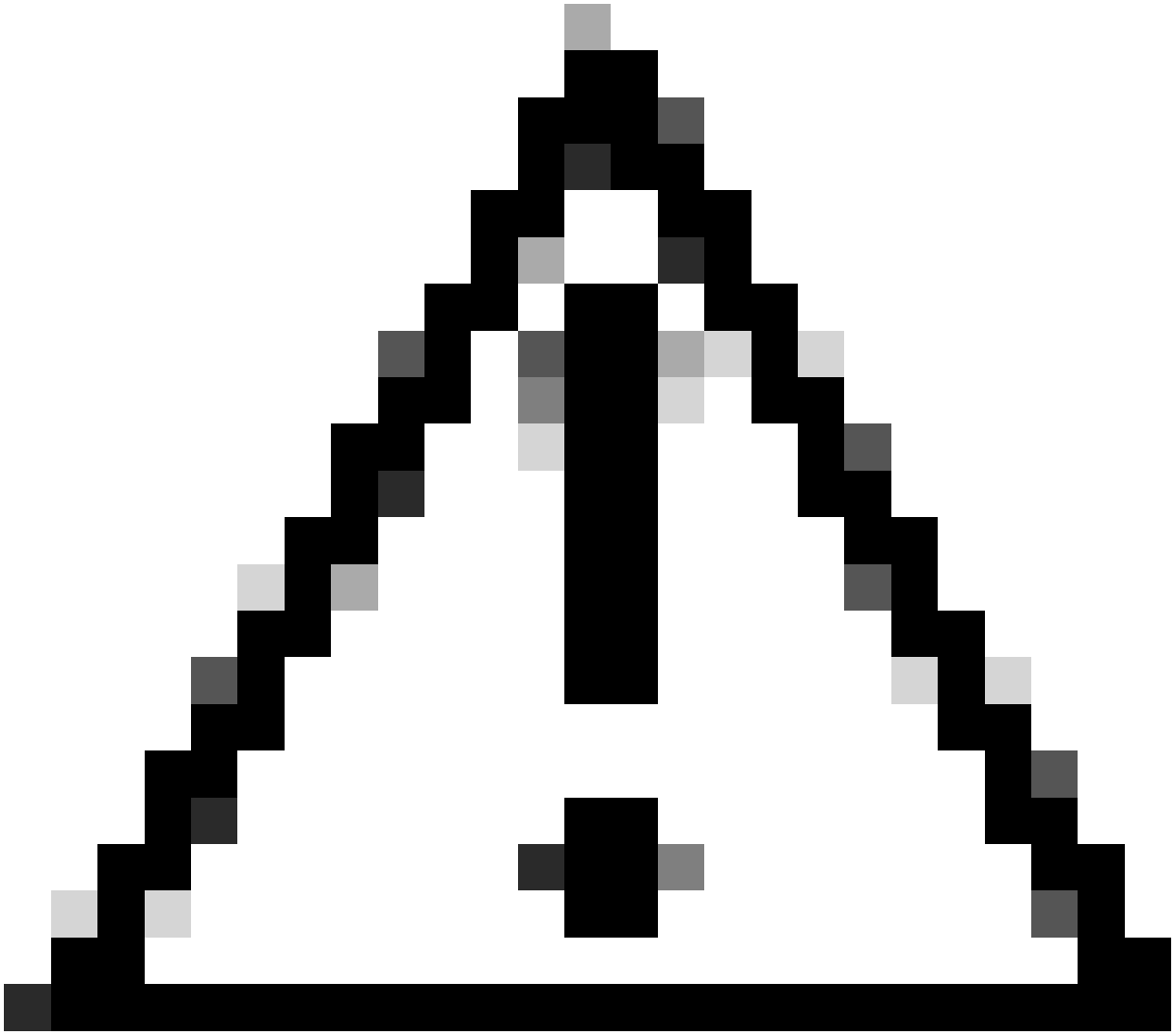


Hinweis: Bei der Konfiguration von vPC-Fabric-Peering über DCNM erfolgt die TCAM-Partitionierung, sie muss jedoch neu geladen werden, damit sie in Kraft tritt.

Sobald die Änderung vorgenommen wurde, wird sie im folgenden Befehl angezeigt:

```
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
```



Vorsicht: Stellen Sie sicher, dass das Gerät nach den Änderungen am TCAM neu geladen wird. Andernfalls wird die VPC aufgrund von Änderungen, die nicht am TCAM

---

vorgenommen wurden, nicht angezeigt.

Konfiguration für vPC

VPC-Domäne

Für LEAF-3 und LEAF-4 in der vPC-Domäne muss die Konfiguration die IP-Adressen für den Keep-Alive und den virtuellen Peer-Link festlegen.

```
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
```

Verbindung aufrecht halten

Jede direkte Layer-3-Verbindung zwischen vPC-Peers darf nur für den Peer-Keep-Alive verwendet werden. Es muss sich in einer separaten VRF-Instanz befinden, die nur für den Keep-Alive-Modus vorgesehen ist. In diesem Szenario wird das Schnittstellenmanagement des Switches verwendet.

```
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
```

Layer-3-Schnittstelle für die virtuelle Peer-Verbindung

Die für den virtuellen Peer-Link verwendete Layer-3-Schnittstelle darf nicht mit der für den Keep-Alive verwendeten Schnittstelle identisch sein. Sie können entweder den gleichen Loopback verwenden, der für das Underlay verwendet wird, oder es kann sich um einen dedizierten Loopback auf dem Nexus handeln.

Hierbei ist "loopback0" für das Underlay und "loopback2" ein dediziertes Loopback für den virtuellen Peer-Link, während "loopback1" die Schnittstelle ist, die unserer Schnittstelle NVE zugeordnet ist.

```
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
```

VPC-Peer-Link

Dem Peer-Link muss ein Port-Channel zugewiesen werden, auch wenn dem Port-Channel keine physische Schnittstelle zugewiesen wird.

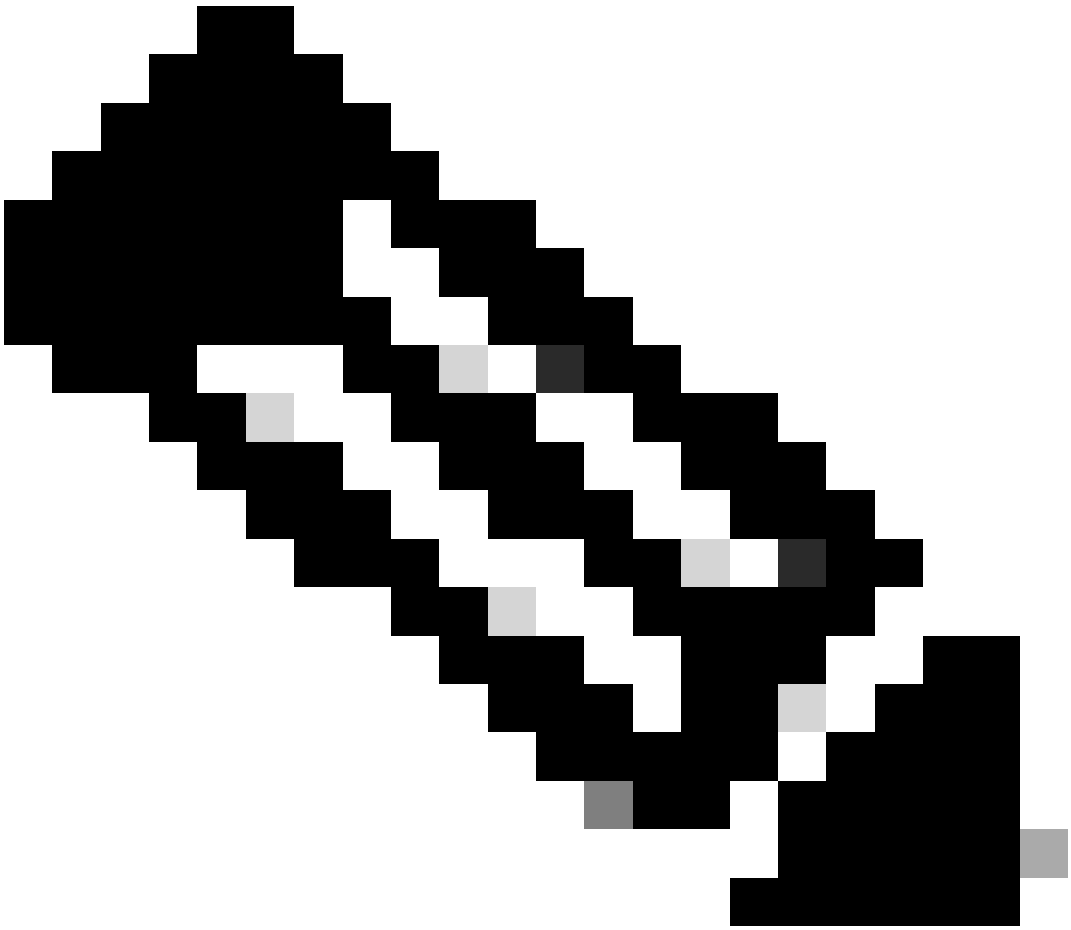
```
LEAF-3(config-if)# sh run interface port-channel 1 membership

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

Up-Links

Der letzte Teil der Konfiguration besteht in der Konfiguration der Verbindungen auf beiden Leafs zum SPINE mit der Kommando-Port-Typ-Struktur.

```
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
```



Hinweis: Wenn Sie die Port-Typ-Fabric nicht konfigurieren, können Sie nicht erkennen, dass der Keepalive vom Nexus generiert wird.

SPINES-Konfiguration

Für die Spines wird empfohlen, die QoS so einzustellen, dass sie mit dem in der vPC-Domäne konfigurierten DSCP-Wert übereinstimmt, da die vPC-Fabric-Peering-Peer-Verbindung über das Transportnetzwerk hergestellt wird.

CFS-Nachrichten auf Kontrollebene, die zum Synchronisieren von Port-Statusinformationen, VLAN-Informationen, VLAN-zu-VNI-Zuordnung, Host-MAC-Adressen und IGMP-Snooping-Gruppen verwendet werden, werden über die Fabric übertragen. CFS-Nachrichten werden mit dem entsprechenden DSCP-Wert markiert, der im Transportnetzwerk geschützt werden muss.

```
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
```

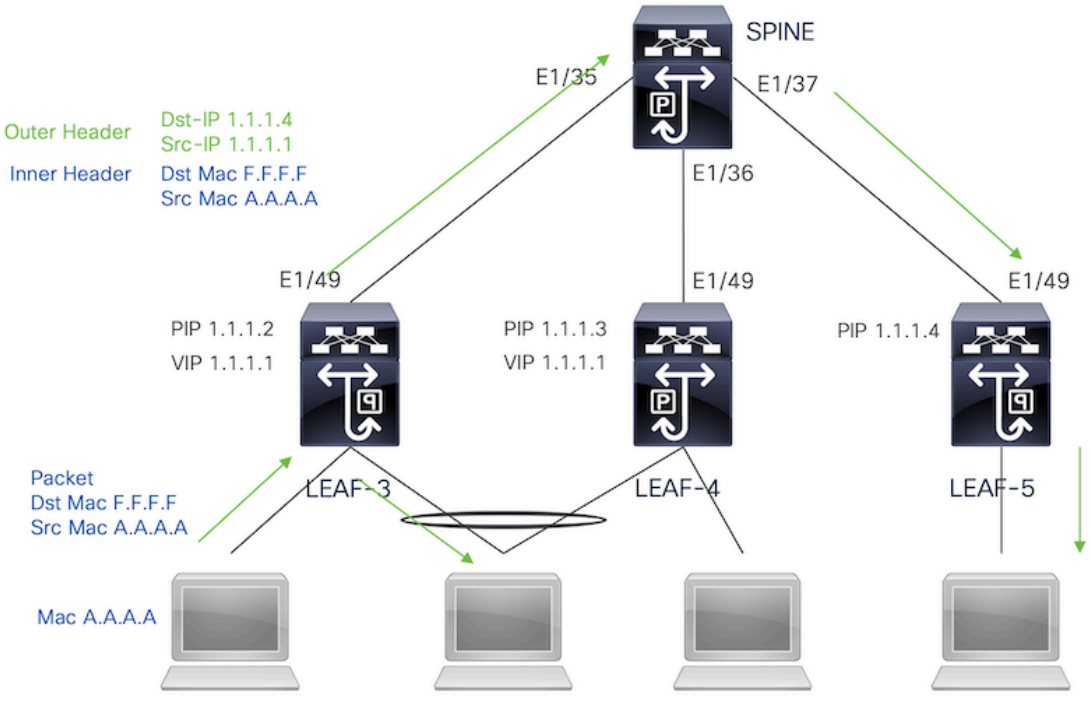
Broadcast-, unbekannter Unicast- und Multicast-Datenverkehr mit Kapselung der Eingangsreplikation

Wenn der Nexus ein Paket empfängt, das ausgestrahlt werden muss, generiert er 2 Kopien des Pakets.

1. An alle Remote-VTEPS in der Flood-Liste für den VNI inkl. lokaler Access Ports

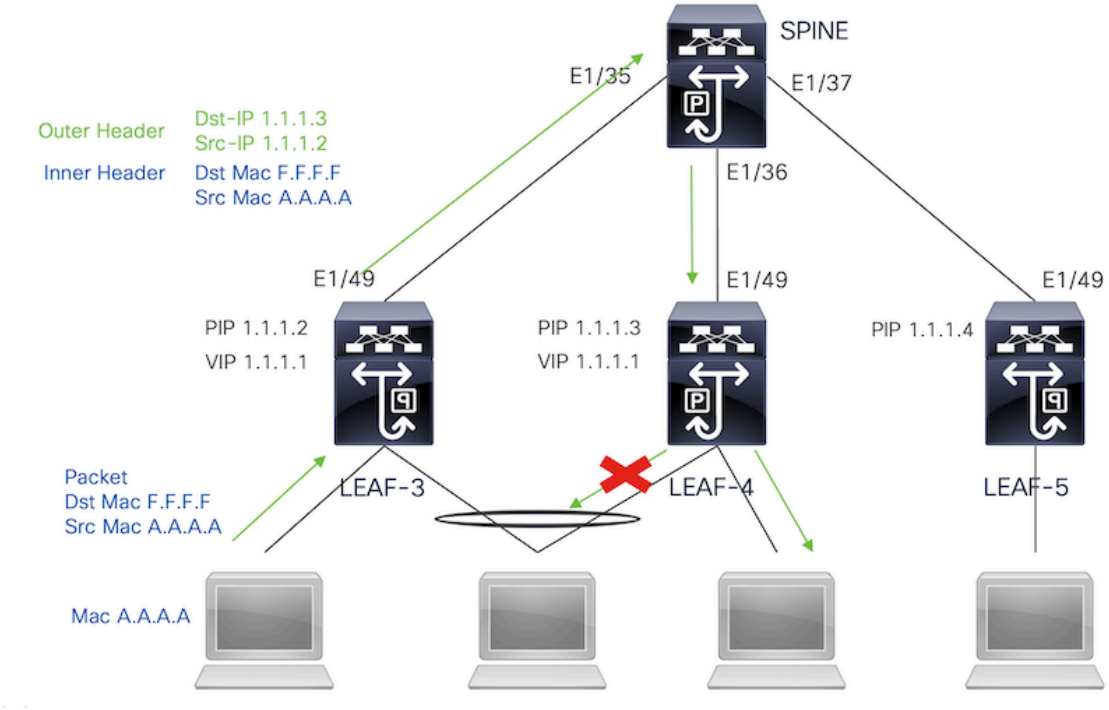
2. Mit dem Remote-vPC-Peer

Für die erste Kopie kapselte der Nexus den Datenverkehr mithilfe der Quell-IP der sekundären IP-Adresse und der Ziel-IP der Remote-VTEP sowie zu den lokalen Access-Ports.



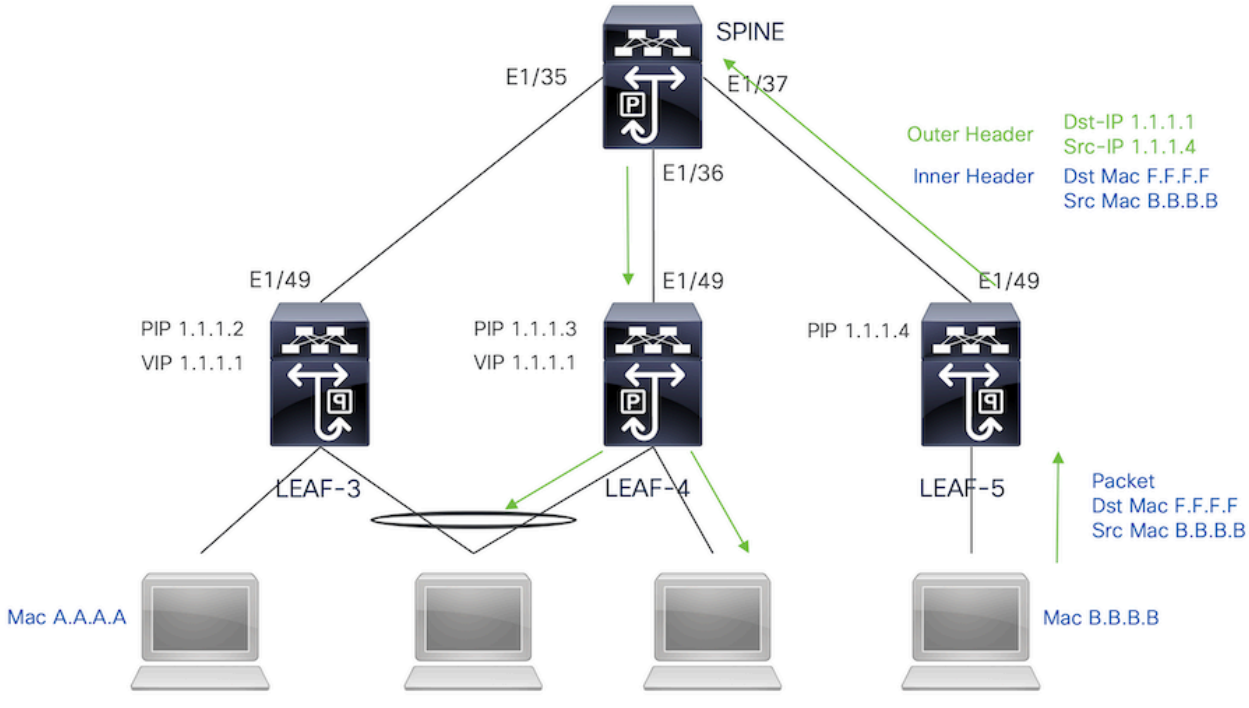
Für die zweite Kopie wird sie an den Remote-vPC-Peer gesendet. Dabei ist die Quell-IP die primäre IP des Loopbacks, und die Ziel-IP ist die PIP des Remote-vPC-Peers.

Sobald das Paket vom Spine empfangen wurde, leitet das Remote-VTEP das Paket nur an die verwaisten Ports weiter.



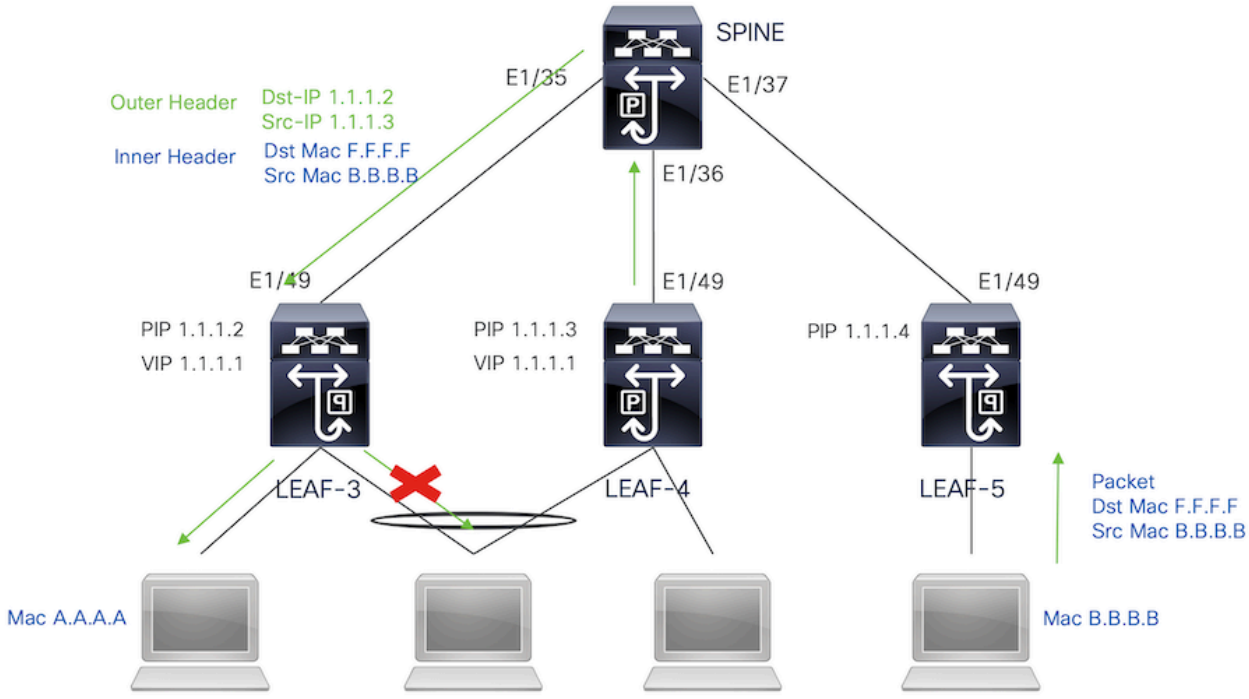
Broadcast-, unbekannter Unicast- und Multicast-Datenverkehr mit Dekapsulation der eingehenden Replikation

Da die Ziel-IP für BUM-Datenverkehr, der von einem anderen VTEP empfangen wird, die VIP ist, die der Datenverkehr zu einem der VPC-Geräte hasht, entkapselt sie das Paket und sendet es an die Access-Ports.



Damit der Datenverkehr zu den verwaisten Ports gelangt, die mit dem Remote-VPC-Peer verbunden sind, generiert der Nexus eine Kopie des Pakets und sendet dieses nur an den Remote-VPC. Dabei wird die primäre IP-Adresse als Quell-/Ziel-IP verwendet.

Sobald der Nexus den Datenverkehr auf dem Remote-vPC-Peer empfangen hat, entkapselt er ihn und leitet ihn nur an verwaiste Ports weiter.



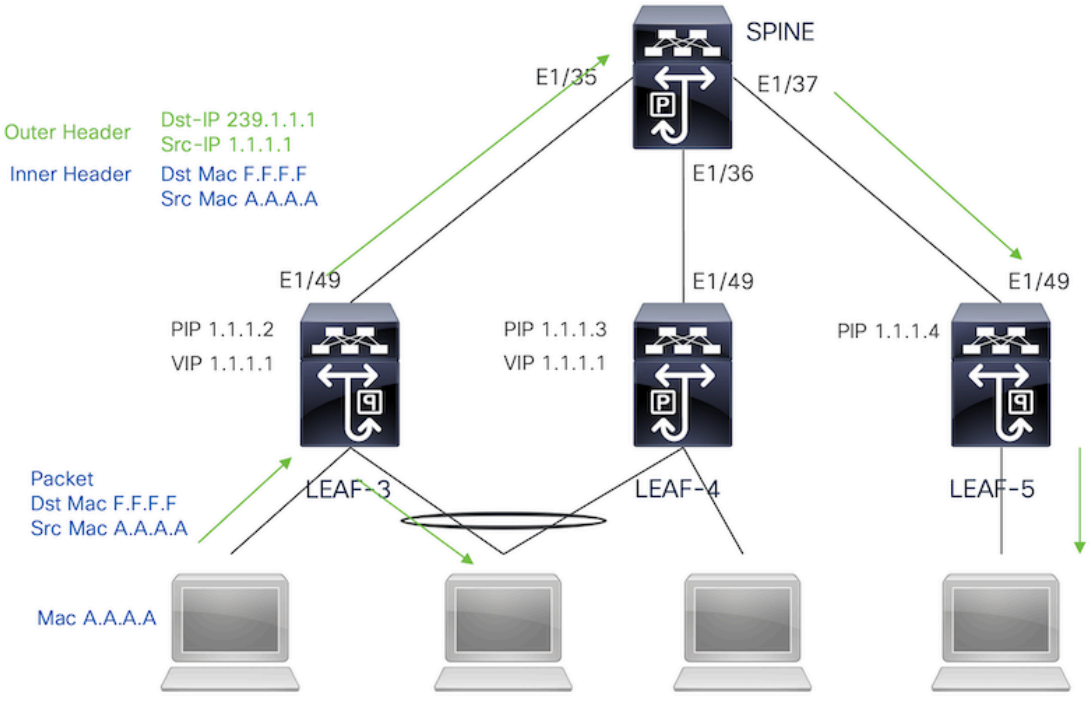
Broadcast-, Unicast- und Multicast-Datenverkehr mit Multicast-Kapselung

Wenn der Nexus ein Paket empfängt, das ausgestrahlt werden muss, generiert er 2 Kopien des Pakets.

1. Das Paket wird an alle OIFs im Multicast-S,G-Eintrag einschließlich lokaler Zugriffsports gesendet.

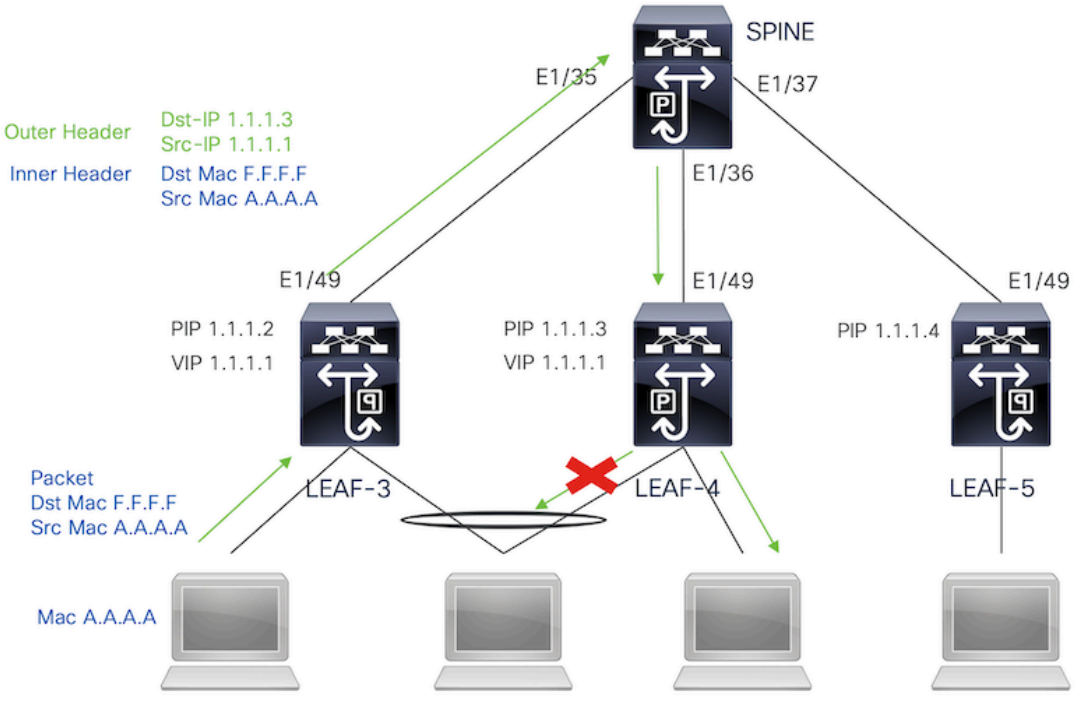
2. Mit dem Remote-vPC-Peer

Für die erste Kopie kapselte der Nexus den Datenverkehr mithilfe der Quell-IP der sekundären IP-Adresse und der Ziel-IP-Adresse der konfigurierten Multicast-Gruppe.



Für die zweite Kopie wird sie an den Remote-VPC-Peer gesendet. Die Quell-IP ist dabei die sekundäre IP des Loopbacks und die Ziel-IP die PIP des Remote-VPC-Peers.

Sobald das Paket vom Spine empfangen wurde, leitet das Remote-VTEP das Paket nur noch an die verwaisten Ports weiter.

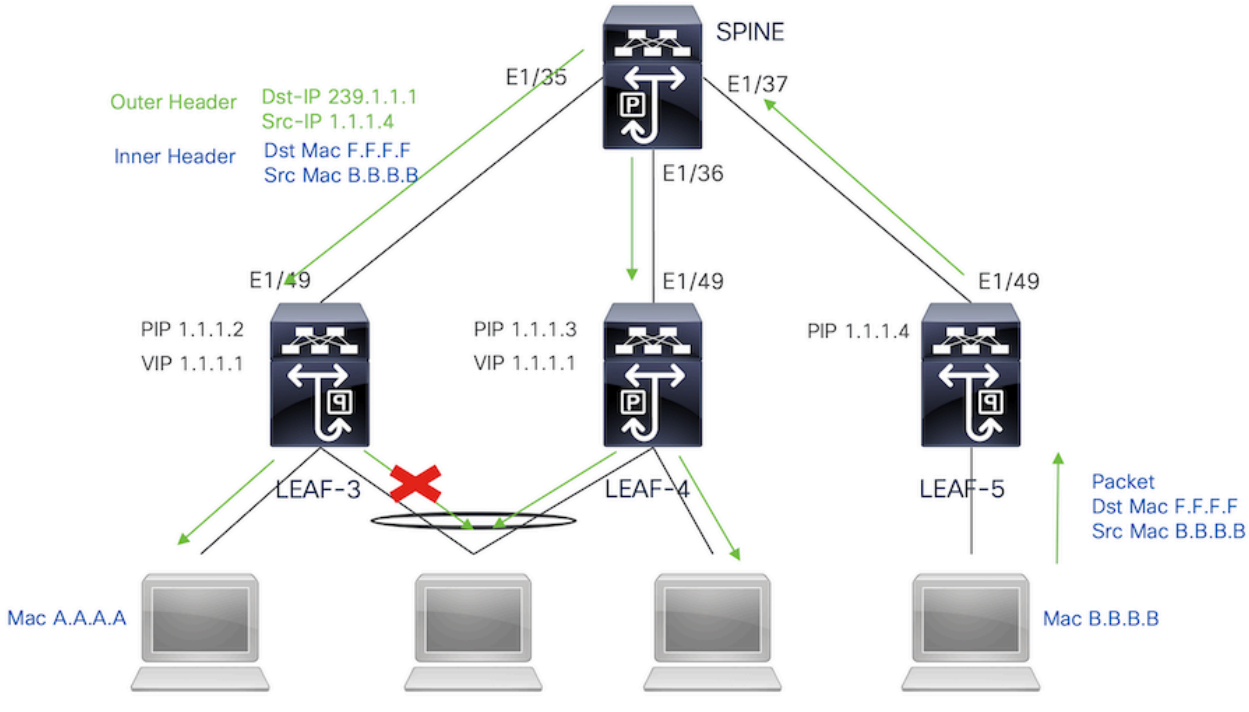


Broadcast-, unbekannter Unicast- und Multicast-Datenverkehr mit Multicast-Entkapselung

Für den Entkapselungsprozess wird das Paket an beide vPC-Peers gesendet. Nur ein VPC-Gerät leitet den Datenverkehr über die VPC-Port-Channels weiter. Dies wird vom im Befehl angezeigten Forwarder entschieden.

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

VPC DF: FORWARDER
```



Überprüfung

Führen Sie die folgenden Befehle aus, um sicherzustellen, dass der vPC betriebsbereit ist:

Überprüfen Sie die Erreichbarkeit der für die virtuelle Peer-Verbindung verwendeten IP-Adressen.

```
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
```

Führen Sie den folgenden Befehl aus, um die Rollen für den vPC zu überprüfen:

```
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
```

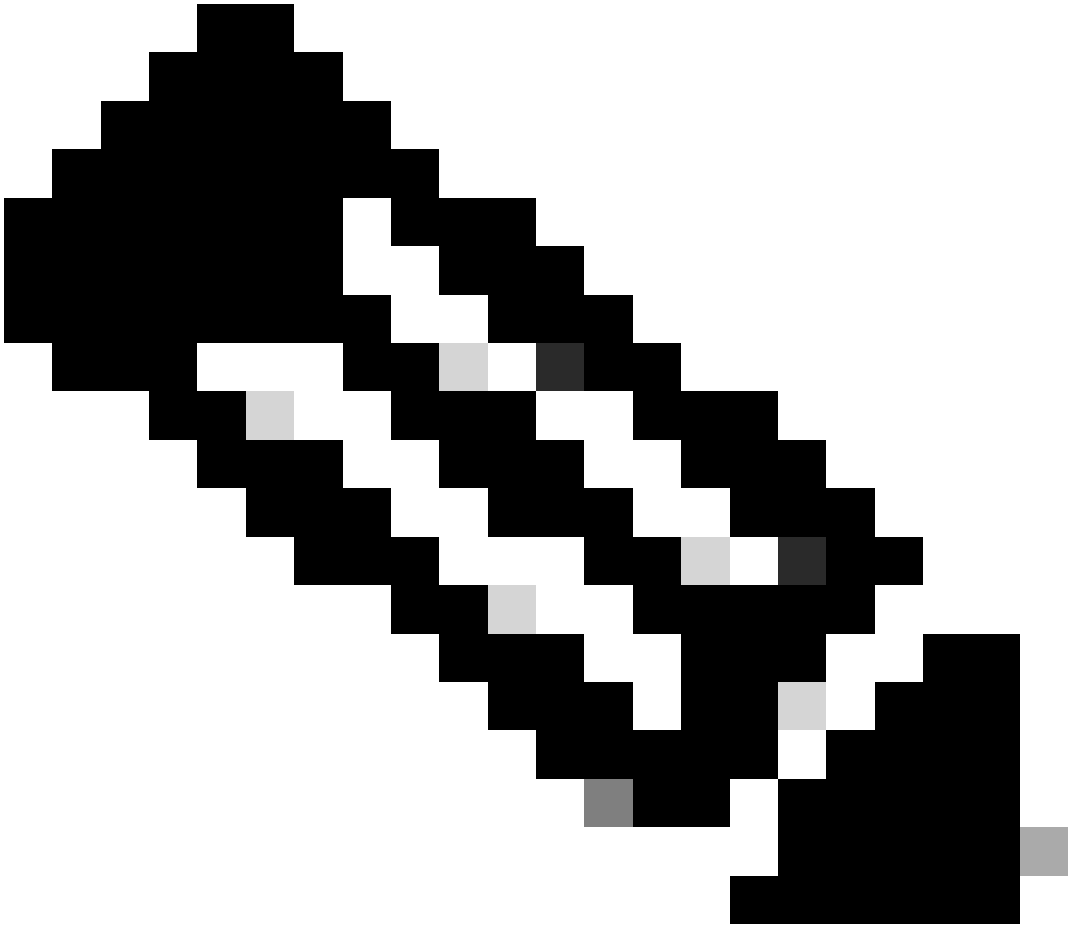
Alle im Peer-Link-Port-Channel zulässigen VLANs müssen einem VNI zugeordnet werden. Andernfalls werden sie als inkonsistent angezeigt.

```
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
```

Führen Sie den folgenden Befehl aus, um zu bestätigen, dass die Konfiguration auf den Up-Links richtig programmiert ist:

```
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
```



Hinweis: Die NVE und die zugehörige Loopback-Schnittstelle werden angezeigt, es sei denn, die VPC ist aktiv.

Zugehörige Informationen

- Technischer Support und Downloads von Cisco

Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.