

Nexus 9000: Konfigurieren und Überprüfen von VXLAN Xconnect

Inhalt

[Einleitung](#)

[Voraussetzungen](#)

[Anforderungen](#)

[Verwendete Komponenten](#)

[Überblick](#)

[Topologie](#)

[Konfigurieren](#)

[Überprüfung](#)

[Hinweise](#)

[Paketerfassung](#)

Einleitung

Das Dokument beschreibt eine Kurzreferenz zur Konfiguration und Verifizierung von VXLAN Xconnect für Nexus 9000-Switches.

Voraussetzungen

Anforderungen

Cisco empfiehlt, dass Sie VXLAN EVPN kennen.

Verwendete Komponenten

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

- N9K-C93180YC-EX
- NXOS 9.2(1)

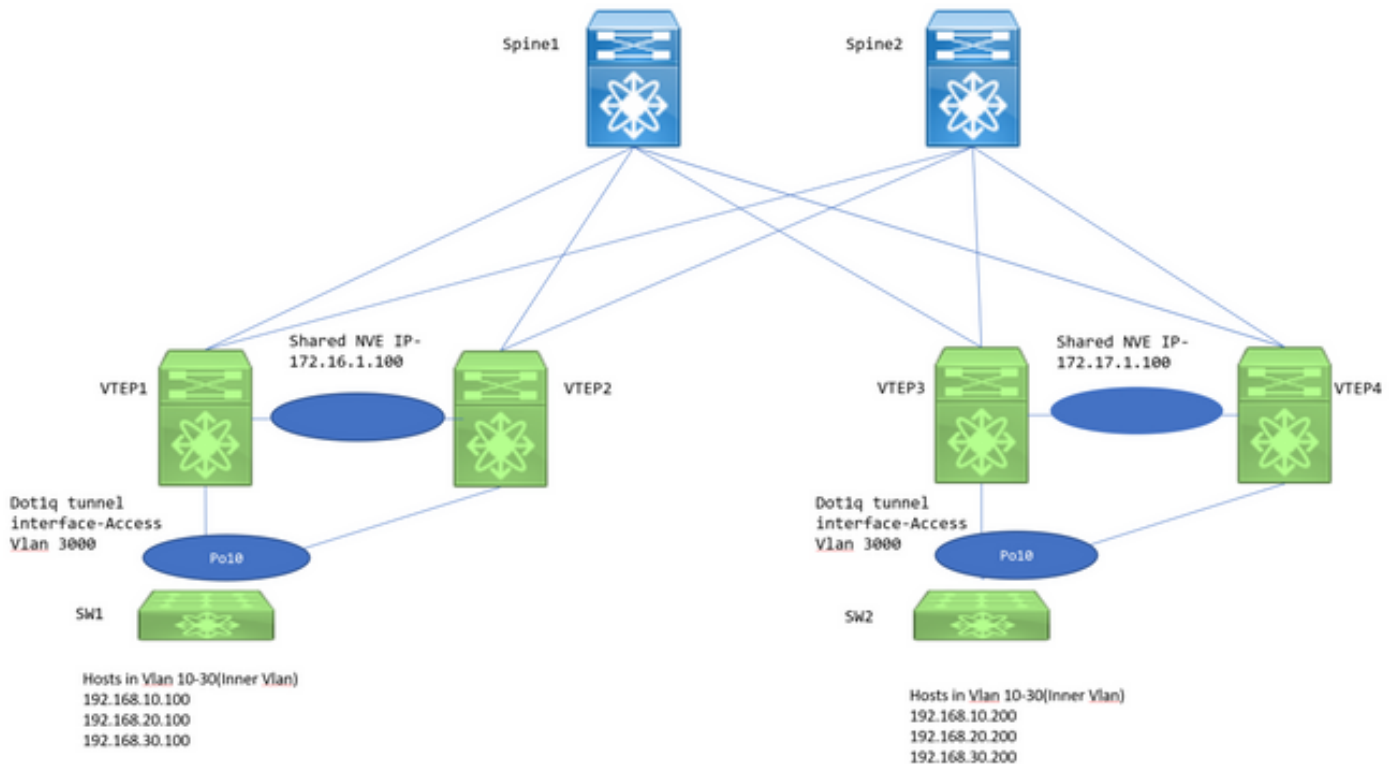
Die Informationen in diesem Dokument beziehen sich auf Geräte in einer speziell eingerichteten Testumgebung. Alle Geräte, die in diesem Dokument benutzt wurden, begannen mit einer gelöschten (Nichterfüllungs) Konfiguration. Wenn Ihr Netz Live ist, überprüfen Sie, ob Sie die mögliche Auswirkung jedes möglichen Befehls verstehen.

Überblick

VXLAN Xconnect ist ein Mechanismus für einen Point-to-Point-Tunnel für Daten- und Steuerungs Pakete von einem Leaf zum anderen. Innere dot1q-Tags bleiben erhalten, und VXLAN wird in die äußere VNID eingekapselt, die als Xconnect-VNID angegeben ist. Layer-2-

Kontrollrahmen wie Link Layer Discovery Protocol (LLDP), Cisco Discovery Protocol (CDP) und Spanning Tree Protocol (STP) sind VXLAN-gekapselt und werden an andere Enden des Tunnels übertragen.

Topologie



Bei VTEP1, VTEP2, VTEP3 und VTEP4 handelt es sich um zwei vPC-VTEP-Paare, die so konfiguriert sind, dass die inneren dot1q-Tags der Downstream-Switches erhalten bleiben. Wenn das VXLAN eingekapselt ist, verwenden Sie die VXLAN-VNID der äußeren VLAN-ID, um das VTEP an das Remote-Netzwerk weiterzuleiten. Alle VTEPs sind N9K-C93180YC-EX.

Downstream-Switches sind Nexus 3ks, die in den jeweiligen VLANs mit Switch Virtual Interface (SVIs) konfiguriert sind, um die Hosts nachzuziehen.

Konfigurieren

1. Das in dieser Xconnect-Topologie verwendete äußere VLAN lautet 3000. Dies ist bei der VNID- und Xconnect-Konfiguration der Fall.

```
VTEP1# sh run vlan 3000  
  
vlan 3000  
  vn-segment 1003000  
  xconnect
```

2. Die Funktion NGOAM muss aktiviert sein und muss entsprechend konfiguriert werden.

```
VTEP1# sh run ngoam
```

```
feature ngoam
```

```
ngoam install acl
```

```
ngoam xconnect hb-interval 5000
```

3. Dot1q-Tunnelkonfiguration zum Downstream-Switch

```
VTEP1# sh run int po10
```

```
interface port-channel10
  switchport
  switchport mode dot1q-tunnel
  switchport access vlan 3000
  speed 40000
  no negotiate auto
  vpc 10
```

Die vPC-Konfigurationen sind nur erforderlich, wenn VTEPs als vPC bereitgestellt werden. Andernfalls überspringen Sie die in diesem Dokument erwähnten vPC-Konfigurationen. VXLAN Xconnect kann auch auf einem eigenständigen VTEP konfiguriert werden.

4. Die Multicast-Gruppe muss unter der NVE-Schnittstelle definiert werden, um die Weiterleitung zu übernehmen. Hinweis zur Aktivierung des **ip pim sparse-mode** auf relevanten Uplinks und zur Definition des PIM-RP, sodass Multicast-Routing und PIM-Nachrichten entsprechend ausgetauscht werden. In der Regel wird PIM RP auf der Spine-Ebene definiert.

```
VTEP1# sh run int nve1
```

```
no shutdown
host-reachability protocol bgp
source-interface loopback1
member vni 1003000 mcast-group 239.30.30.30
```

5. Das Infra-VLAN muss als natives VLAN innerhalb der Peer-Verbindung angegeben und zugelassen werden. Dieser Schritt ist für vPC-VTEPs erforderlich.

```
VTEP1# sh run span|infra
no spanning-tree vlan 3000
system nve infra-vlans 999
```

```
VTEP1# sh run int po1
```

```
interface port-channel1
  switchport
  switchport mode trunk
  switchport trunk native vlan 999
  spanning-tree port type network
  vpc peer-link
```

6. BGP/EVPN-Konfiguration: L2VPN-EVPN-Nachbarschaften werden zwischen Leaf/Spine benötigt, um die Type 3-Routen auszutauschen, die zum Herstellen der VXLAN Xconnect erforderlich sind.

- Hier sind die IP-Adressen 192.168.100.1 und 192.168.100.2 die Spines in der Topologie. In der Regel werden die L2VPN EVPN-Nachbarschaften zu den Spines gebildet. Spines

konfigurieren alle Leaf-Switches als Routen-Reflektor-Clients in einem iBGP-Szenario.

- Es wird empfohlen, separate Loopbacks für BGP/OSPF- und NVE-Zwecke zu verwenden.

```
feature bgp

router bgp 65000
  router-id 192.168.100.3
  neighbor 192.168.100.1
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
      send-community
      send-community extended
  neighbor 192.168.100.2
    remote-as 65000
    update-source loopback0
    address-family l2vpn evpn
send-community
send-community extended evpn vni 1003000 l2 rd auto route-target import auto route-target export auto
```

Anmerkung: STP muss im Xconnect VLAN deaktiviert werden. Das MAC-Lernen findet innerhalb des Xconnect-VLAN nicht statt. Dies bedeutet im Wesentlichen, dass es keine Typ-2-bgp l2vpn-Ereignis-Updates für MAC-Adressen gibt. Daher wird der Datenverkehr von einem VTP mit der äußeren Ziel-IP-Adresse gekapselt, die auf die Mcast-Gruppe (239.30.30.30) für das Xconnect-VLAN festgelegt ist.

Überprüfung

1. BGP-Nachbarschaft.

```
VTEP1# sh bgp l2vpn evpn sum
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 192.168.100.3, local AS number 65000
BGP table version is 14, L2VPN EVPN config peers 2, capable peers 1
4 network entries and 5 paths using 756 bytes of memory
BGP attribute entries [3/492], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [2/8]

Neighbor      V    AS MsgRcvd MsgSent  TblVer  InQ  OutQ Up/Down  State/PfxRcd
192.168.100.1  4 65000    92     90     14    0    0 01:21:41  2
```

2. Präfixe vom Typ 3 empfangen.

```
VTEP1# sh bgp l2vpn evpn
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 14, Local Router ID is 192.168.100.3
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

Network          Next Hop          Metric      LocPrf      Weight Path
Route Distinguisher: 192.168.100.3:35767 (L2VNI 1003000)
*>l[3]:[0]:[32]:[172.16.1.100]/88
                172.16.1.100          100         32768 i
* i[3]:[0]:[32]:[172.17.1.100]/88<<< bgp type 3
```

```

172.17.1.100          100          0 i
*>i          172.17.1.100          100          0 i

```

Route Distinguisher: 192.168.100.5:35767

```

*>i[3]:[0]:[32]:[172.17.1.100]/88
172.17.1.100          100          0 i

```

Route Distinguisher: 192.168.100.6:35767

```

*>i[3]:[0]:[32]:[172.17.1.100]/88
172.17.1.100          100          0 i

```

3. NVE-Peering.

```

VTEP1# sh nve peer
Interface Peer-IP          State LearnType Uptime  Router-Mac
-----
nve1      172.17.1.100            Up    CP         00:58:06 n/a

```

```

VTEP1# show nve vni
Codes: CP - Control Plane      DP - Data Plane
       UC - Unconfigured       SA - Suppress ARP
       SU - Suppress Unknown Unicast

```

```

Interface VNI      Multicast-group  State Mode Type [BD/VRF]  Flags
-----
nve1      1003000  239.30.30.30    Up    CP    L2 [3000]          Xconn <<<

```

4. NGOAM-Kontrollen.

```
VTEP1# show ngoam xconnect sess all
```

```

States: LD = Local interface down, RD = Remote interface Down
        HB = Heartbeat lost, DB = Database/Routes not present
        * - Showing Vpc-peer interface info

```

```

Vlan      Peer-ip/vni      XC-State      Local-if/State      Rmt-if/State
=====
3000  172.17.1.100 / 1003000      Active              Po10 / UP              Po10 / UP

```

```

VTEP1# show ngoam xconnect sess 3000
Vlan ID: 3000
Peer IP: 172.17.1.100 VNI : 1003000
State: Active <<< State should be active
Last state update: 12/10/2018 17:13:45.337
Local interface: Po10 State: UP
Local vpc interface Po10 State: UP
Remote interface: Po10 State: UP
Remote vpc interface: Po10 State: UP

```

Sobald die NGOAM-Sitzung beendet ist, sehen die N3Ks einander in CDP. STP-BPDUs werden ebenfalls getunnelt, sodass die Switches auch die Root-Bridge-Platzierung vereinbaren.

5. Überprüfungen an den Downstream-Switches.

```
SW1(config)# sh span vl 10
```

```

VLAN0010
Spanning tree enabled protocol rstp
Root ID      Priority      32778
Address      7079.b348.6cb7

```

```
This bridge is the root
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
```

```
Bridge ID Priority 32778 (priority 32768 sys-id-ext 10)
Address 7079.b348.6cb7
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
```

Interface	Role	Sts	Cost	Prio.	Nbr	Type
Po10	Desg	FWD	1	128.	4105	P2p

```
SW2(config)# sh span vl 10
```

```
VLAN0010
Spanning tree enabled protocol rstp
Root ID Priority 32778
Address 7079.b348.6cb7
Cost 1
Port 4105 (port-channel10)
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID Priority 32778 (priority 32768 sys-id-ext 10)
Address 707d.b964.9441
Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
```

Interface	Role	Sts	Cost	Prio.	Nbr	Type
Po10	Root	FWD	1	128.	4105	P2p

```
SW1(config)# show ip int b
IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status
Vlan10 192.168.10.100 protocol-up/link-up/admin-up
Vlan20 192.168.20.100 protocol-up/link-up/admin-up
Vlan30 192.168.30.100 protocol-up/link-up/admin-up
```

```
SW2(config)# show ip int b
IP Interface Status for VRF "default"(1)
Interface IP Address Interface Status
Vlan10 192.168.10.200 protocol-up/link-up/admin-up
Vlan20 192.168.20.200 protocol-up/link-up/admin-up
Vlan30 192.168.30.200 protocol-up/link-up/admin-up
```

```
SW1(config)# ping 192.168.10.200
PING 192.168.10.200 (192.168.10.200): 56 data bytes
64 bytes from 192.168.10.200: icmp_seq=0 ttl=254 time=0.826 ms
64 bytes from 192.168.10.200: icmp_seq=1 ttl=254 time=0.531 ms
64 bytes from 192.168.10.200: icmp_seq=2 ttl=254 time=0.54 ms
64 bytes from 192.168.10.200: icmp_seq=3 ttl=254 time=0.522 ms
64 bytes from 192.168.10.200: icmp_seq=4 ttl=254 time=0.571 ms
```

Hinweise

1. Die dot1q-Tunnelschnittstellen bleiben in einer Xconnect VXLAN-Konfiguration im Fehlermodus "disabled" (Deaktiviert), wenn die Konfigurationen innerhalb der vPC-Switches nicht konsistent sind. Im Folgenden sind einige Fälle aufgeführt, in denen die Schnittstelle fehlerhaft deaktiviert wird.

- Wenn das VLAN-zu-VN-Segment nicht auf beiden vPC-Switches definiert ist.
- Wenn die NVE to Multicast-Gruppe nicht auf beiden vPC-Switches definiert ist.
- Wenn die NGOAM-Heartbeats nicht empfangen werden (verwenden Sie den

Ethernet II, Src: Cisco_2a:89:a7 (70:79:b3:2a:89:a7), Dst: IPv4mcast_1e:1e:1e (01:00:5e:1e:1e:1e)

Ethernet II, Src: Cisco_64:94:41 (70:7d:b9:64:94:41), Dst: Cisco_48:6c:b7 (70:79:b3:48:6c:b7)

802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 10

000. = Priority: Best Effort (default) (0)

...0 = DEI: Ineligible

.... 0000 0000 1010 = ID: 10

Type: IPv4 (0x0800)

Internet Protocol Version 4, Src: 192.168.10.200, Dst: 192.168.10.100

- Auch wenn die dot1q-Tunnelschnittstelle in einer vPC-Konfiguration verwaist verbunden ist, muss die Multicast-Gruppe unter der NVE-Schnittstelle für das VN-Segment, das Teil von Xconnect ist, auf beiden Switches konfiguriert werden.
- NGOAM-Heartbeats werden vom primären vPC-Switch verarbeitet/gesendet. Heartbeat-Nachrichten, die auf sekundärem vPC landen, werden mit dem primären synchronisiert

2. Wenn Xconnect in einem VLAN konfiguriert wird, wird der Datenverkehr zwischen den Standorten mit der äußeren Zieladresse=Multicast-Adresse gekapselt, die unter der NVE-Schnittstelle für das jeweilige VN-Segment definiert ist. Es wird empfohlen, für die Xconnect VLANs eine eindeutige Multicast-Gruppe zu verwenden. Multicast im Core/Spine muss funktionsfähig sein.

3. Multicast-Datenverkehr kann beide vPC-Boxen auf der Remote-Seite von Xconnect treffen; Der "Decap"-Gewinner (die Box, die das BUM entkapseln kann) ist jedoch nur ein Switch in einem vPC-Paar. Dies kann mithilfe der **Quelle** "Command-**show forward multicast route group <Group address> <SRC IP>** überprüft werden. Wenn die hier gezeigte Markierung ein **v** in Kleinbuchstaben ist, bedeutet dies, dass die Box ein Decap-Verlierer ist und wenn es sich um einen **V** in Großbuchstaben handelt, ist die Box der Gewinner des Decap-Falls und kann den Multicast-Datenverkehr entkapseln und weiter unten weiterleiten.

4. Wenn der Host mit 93180YC-basierten Plattformen verwaist ist und kein OIL für S, G für 9k1 vorhanden ist, wird eine Kopie des Multicast-Pakets mithilfe einer speziellen Kapselung von Quell-IP-> 127.0.0.1 und Ziel-IP-> gemeinsam genutzter NVE-IP an den vPC-Peer gesendet, und wenn der 9k2 OIL ist Bei S- und G-Einträgen wird die Weiterleitung des Datenverkehrs von 9k2 an die Remote-Standorte durchgeführt.

Paketerfassung

Hier ist ein Screenshot einer Paketerfassung, die am Spine-Switch durchgeführt wurde:

```
Frame 1: 152 bytes on wire (1216 bits), 152 bytes captured (1216 bits)
Ethernet II, Src: Cisco_2a:89:a7 (70:79:b3:2a:89:a7), Dst: IPv4mcast_1e:1e:1e (01:00:5e:1e:1e:1e)
Internet Protocol Version 4, Src: 172.17.1.100, Dst: 239.30.30.30
User Datagram Protocol, Src Port: 12860, Dst Port: 4789
Virtual eXtensible Local Area Network
  > Flags: 0x0800, VXLAN Network ID (VNI)
    Group Policy ID: 0
    VXLAN Network Identifier (VNI): 1003000
    Reserved: 0
Ethernet II, Src: Cisco_64:94:41 (70:7d:b9:64:94:41), Dst: Cisco_48:6c:b7 (70:79:b3:48:6c:b7)
802.1Q Virtual LAN, PRI: 0, DEI: 0, ID: 10
  000. .... = Priority: Best Effort (default) (0)
  ...0 .... = DEI: Ineligible
  .... 0000 0000 1010 = ID: 10
  Type: IPv4 (0x0800)
Internet Protocol Version 4, Src: 192.168.10.200, Dst: 192.168.10.100
```

- Der interne dot1q-Header=10 wird beibehalten.
- Verwendetes VNI ist 1003000 (das ist die VNID des äußeren VLANs).
- Die Ziel-IP-Adresse ist die Multicast-Gruppe, die unter der nve-Schnittstelle definiert wird.