

Nexus 9000 : Configuration et vérification de VXLAN Xconnect

Contenu

[Introduction](#)

[Conditions préalables](#)

[Conditions requises](#)

[Components Used](#)

[Aperçu](#)

[Topologie](#)

[Configuration](#)

[Vérification](#)

[Dépannage](#)

[Cavates](#)

[Capture de paquets](#)

Introduction

Le document décrit une référence rapide sur la configuration et la vérification de VXLAN Xconnect sur les commutateurs Nexus 9000.

Conditions préalables

Conditions requises

Cisco vous recommande d'avoir connaissance de VXLAN EVPN.

Components Used

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

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

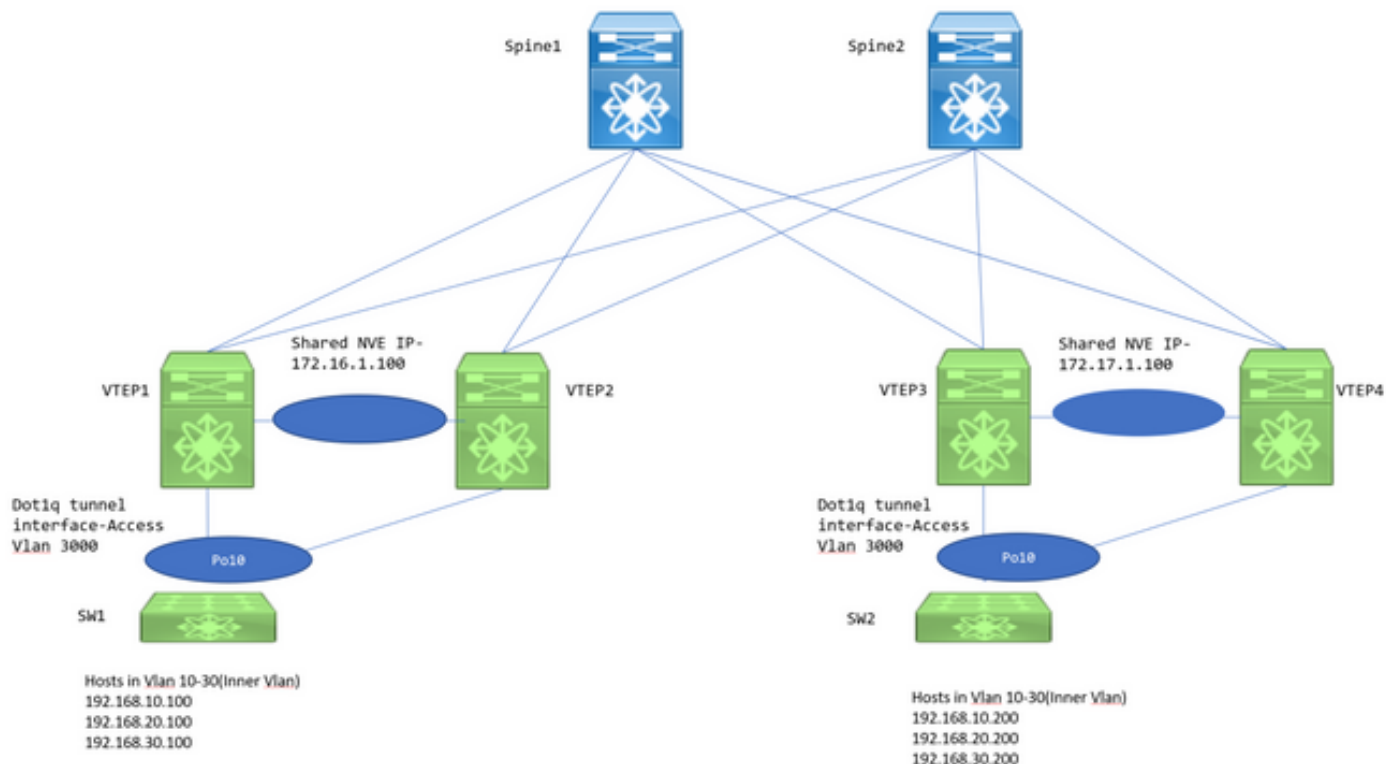
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

Aperçu

VXLAN Xconnect est un mécanisme pour un tunnel point à point pour les paquets de données et de contrôle d'un noeud à l'autre. Les balises Dot1q internes sont conservées et VXLAN encapsulé dans le VNID externe, spécifié sous le nom Xconnect VNID. Les trames de contrôle de couche 2,

telles que le protocole LLDP (Link Layer Discovery Protocol), le protocole CDP (Cisco Discovery Protocol) et le protocole STP (Spanning Tree Protocol), sont encapsulées par VXLAN et envoyées à d'autres extrémités du tunnel.

Topologie



VTEP1, VTEP2, VTEP3 et VTEP4 sont deux paires VTEP vPC configurées de telle sorte que les balises dot1q internes des commutateurs en aval soient préservées et lorsque VXLAN est encapsulé, utilisez le VNID VXLAN de l'ID de VLAN externe afin de l'envoyer au VTEP distant. Tous les VTEP sont N9K-C93180YC-EX.

Les commutateurs en aval sont des Nexus 3ks configurés avec des interfaces SVI (Switch Virtual Interface) dans les VLAN respectifs pour simuler les hôtes.

Configuration

1. Le VLAN externe utilisé dans cette topologie Xconnect est 3000. Il s'agit de celui avec la configuration VNID et Xconnect.

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

2. La fonctionnalité NGOAM doit être activée et nécessite cette configuration.

```
VTEP1# sh run ngoam
```

```
feature ngoam
```

```
ngoam install acl
```

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

3. Configuration du tunnel Dot1q vers le commutateur en aval.

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

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

Les configurations vPC ne sont requises que lorsque les VTEP sont déployés en tant que vPC. Sinon, ignorez les configurations vPC mentionnées dans ce document. VXLAN Xconnect est également configurable sur un VTEP autonome.

4. Le groupe de multidiffusion doit être défini sous l'interface NVE pour prendre en charge le transfert. Remarque pour activer **ip pim sparse-mode** sur les liaisons ascendantes pertinentes et définir le RP PIM, de sorte que le routage de multidiffusion et les messages PIM soient échangés de manière appropriée. En règle générale, le RP PIM est défini au niveau de la couche dorsale.

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

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

5. Le VLAN infrarouge doit être spécifié et autorisé en tant que VLAN natif au sein de la liaison homologue. Cette étape est nécessaire pour les VTEP vPC.

```
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. Configuration BGP/EVPN : Les voisins EVPN L2VPN sont nécessaires entre leaf/spine pour échanger les routes de type 3 nécessaires à l'établissement de VXLAN Xconnect.

- Ici, les adresses IP - 192.168.100.1 et 192.168.100.2 sont les Spines de la topologie. Généralement, les voisins EVPN L2VPN sont formés sur les Spines. Les spines configurent tous les commutateurs Leaf en tant que clients de réflecteur de route dans un scénario iBGP.

- Il est recommandé d'utiliser des bouclages distincts pour les fonctions BGP/OSPF et NVE.

```
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
```

Note: STP doit être désactivé dans le VLAN Xconnect. L'apprentissage MAC ne se produira pas dans le VLAN Xconnect, ce qui signifie essentiellement qu'il n'y a pas de mises à jour evpn bgp l2vpn de type 2 pour les adresses MAC. Par conséquent, le trafic d'un VLAN sera encapsulé avec l'adresse IP de destination externe définie sur le Mcast-group(239.30.30.30) défini pour le VLAN Xconnect.

Vérification

Référez-vous à cette section pour vous assurer du bon fonctionnement de votre configuration.

1. Quartier BGP.

```
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. Recevoir des préfixes de type 3.

```
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)
*>1[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. Contrôles NGOAM.

```

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

```

Une fois la session NGOAM terminée, les N3k se voient dans CDP. Les BPDU STP sont également tunnelisés, de sorte que les commutateurs conviennent également du placement du pont racine.

5. Vérifications sur les commutateurs en aval.

```

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
```

Dépannage

Il n'existe actuellement aucune information de dépannage spécifique pour cette configuration.

Cavates

1. Les interfaces de tunnel dot1q seront bloquées en **état de désactivation des erreurs** dans une configuration Xconnect VXLAN si les configurations des commutateurs vPC ne sont pas cohérentes. Ci-dessous figurent certains cas où l'interface sera désactivée par erreur ;

- Si le VLAN vers le segment VN n'est pas défini sur les deux commutateurs vPC.
- Si le groupe NVE à multidiffusion n'est pas défini sur les deux commutateurs vPC.
- Si les pulsations NGOAM ne sont pas reçues (utilisez ethanalyzer avec filter=**cfm** pour attraper les paquets de pulsations NGOAM).
- Même si l'interface de tunnel dot1q est orpheline connectée dans une configuration vPC, il est toujours nécessaire de configurer le groupe de multidiffusion sous l'interface NVE pour le segment VN qui fait partie de Xconnect sur les deux commutateurs.
- Les pulsations NGOAM sont traitées/envoyées par le commutateur principal vPC. Les messages de pulsation qui atterrissent sur vPC secondaire seront synchronisés sur le principal

2. Lorsque Xconnect est configuré dans un VLAN, le trafic d'un site à l'autre est encapsulé avec l'adresse de destination externe=adresse de multidiffusion définie sous l'interface NVE pour ce segment particulier de VPN. Il est recommandé d'utiliser un groupe de multidiffusion unique pour les VLAN Xconnect. La multidiffusion dans le coeur/la colonne vertébrale doit être fonctionnelle.

3. Le trafic multidiffusion peut atteindre les deux boîtes vPC du côté distant de Xconnect ; Cependant, le gagnant du Decap (la boîte qui peut décapsuler le BUM) ne sera qu'un seul commutateur dans une paire vPC. Ceci peut être vérifié à l'aide de la commande **show forwarding multicast route group <adresse de groupe> source <adresse IP SRC>**. Si l'indicateur représenté ici est un **v** minuscule, cela signifie que la case est un décodeur perdu et s'il s'agit d'un **V majuscule**, la case est le décodeur gagnant et peut donc décapsuler le trafic de multidiffusion et le transmettre plus bas.

4. Sur les plates-formes basées sur 93180YC, lorsque l'hôte est orphelin connecté à 9k1 et qu'il n'y a pas d'OIL pour S, G sur 9k1, une copie du paquet multicast est envoyée à l'homologue vPC à l'aide d'une encapsulation spéciale de Source IP-> 127.0.0.1 et de Destination IP-> partagée NVE et si 9k2 a OIL pour S, G, puis le transfert du trafic sera pris en charge par le 9k2 vers les sites distants.

Capture de paquets

Voici une capture d'écran d'une capture de paquets effectuée au niveau du commutateur dorsal :

```

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

```

- En-tête dot1q interne=10 préservé
- VNI utilisé est 1003000 (qui est le VNID du VLAN externe)
- L'adresse IP de destination est le groupe de multidiffusion défini sous l'interface nve