Nexus 9000: Configureren en controleren VXLAN Xconnect

Inhoud

Inleiding

Het document beschrijft een snelle referentie hoe u VXLAN Xconnect op Nexus 9000 Switches kunt configureren en controleren.

Voorwaarden

Vereisten

Cisco raadt u aan kennis te hebben over VXLAN EVPN.

Gebruikte componenten

De informatie in dit document is gebaseerd op de volgende software- en hardware-versies:

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

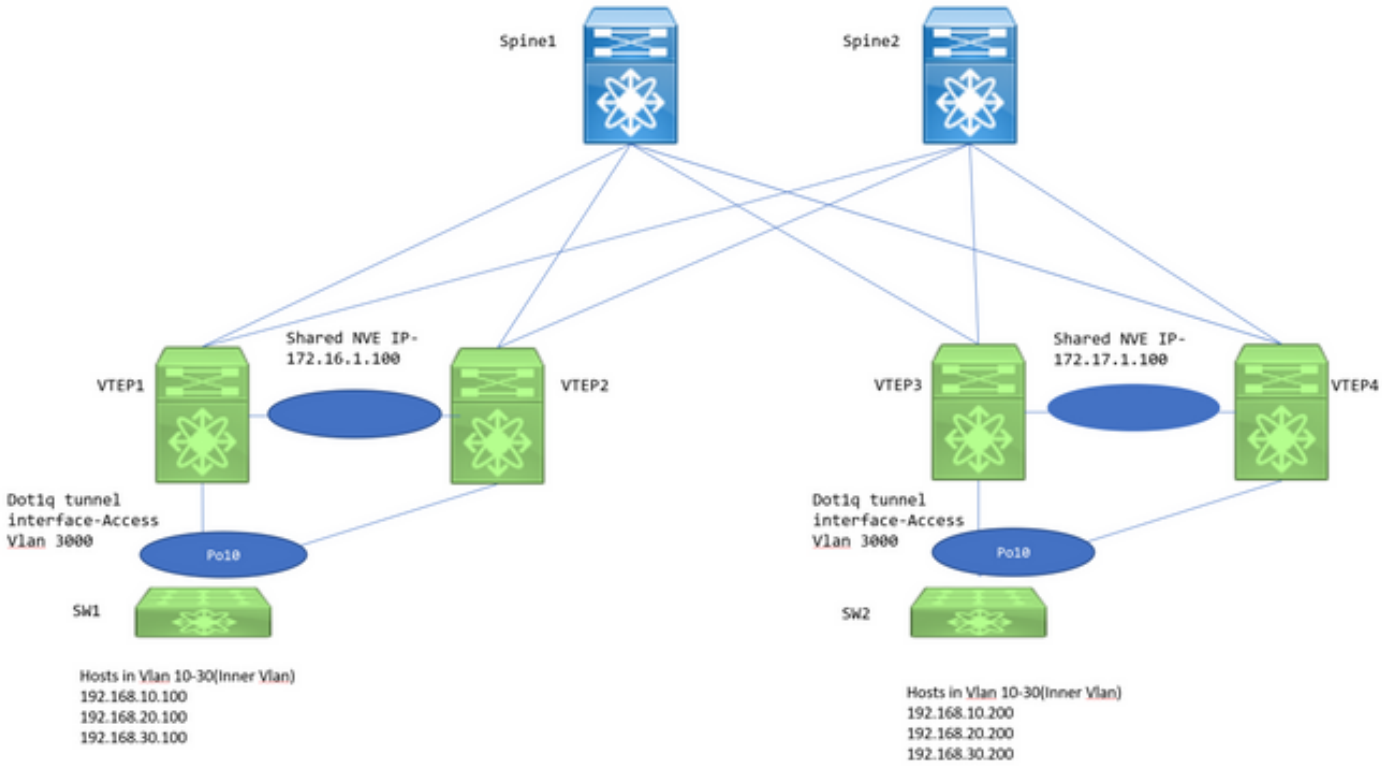
De informatie in dit document is gebaseerd op de apparaten in een specifieke laboratoriumomgeving. Alle apparaten die in dit document worden beschreven, hadden een opgeschoonde (standaard)configuratie. Als uw netwerk live is, moet u de potentiële impact van elke opdracht begrijpen.

Overzicht

VXLAN Xconnect is een mechanisme voor een point-to-point tunnel voor gegevens en besturingspakketten van de ene Leaf naar de andere. Inner Dot1q Tags zijn bewaard en VXLAN is ingekapseld in de buitenste VPN die wordt gespecificeerd als de Xconnect VPN. Layer 2 Control

Frames zoals Link Layer Discovery Protocol (LLDP), Cisco Discovery Protocol (CDP), Spanning Tree Protocol (STP) zijn VXLAN ingekapseld en verzonden naar andere uiteinden van de Tunnel.

Topologie



VTEP1, VTEP2, VTEP3 en VTEP4 zijn twee vPC VTEP-paren die zodanig zijn geconfigureerd dat de inwendige punt1q tags van downstreamswitches bewaard blijven en dat, wanneer VXLAN ingekapseld is, u VXLAN VPN van VLAN-ID gebruikt om naar de afgelegen VTEP te verzenden. Alle VTEP's zijn N9K-C93180YC-EX.

Downstream switches zijn Nexus 3ks, geconfigureerd met Switch Virtual Interface (SVIs) in respectievelijke VLAN's om de hosts na te bootsen.

Configureren

1. Buiten VLAN dat in deze Xconnect-topologie wordt gebruikt, is 3000. Dit is de VPN- en Xconnect-configuratie.

```
VTEP1# sh run vlan 3000

vlan 3000
  vn-segment 1003000
  xconnect
```

2. Feature NGO moet worden ingeschakeld en heeft deze configuratie nodig.

```
VTEP1# sh run ngoam
```

```
feature ngoam

ngoam install acl
ngoam xconnect hb-interval 5000
```

3. Dot1q-tunnelconfiguratie naar de stroomafwaarts gelegen 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
```

De vPC-configuraties zijn alleen vereist wanneer VTEPs worden ingezet als vPC. Anders slaat u de vPC-configuraties over die in dit document worden genoemd. VXLAN Xconnect is ook instelbaar op een standalone VTEP.

4. Multicastgroep moet worden gedefinieerd onder de NVE-interface om de verzending te regelen. Opmerking om ip pim sparse-mode op relevante uplinks in te schakelen en PIM RP te definiëren zodat multicast routing en PIM berichten correct worden uitgewisseld. PIM wordt doorgaans gedefinieerd op de spinelaag.

```
VTEP1# sh run int nve1

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

5. Infrarisch VLAN moet worden gespecificeerd en toegestaan als inheems VLAN binnen de peer link. Deze stap is nodig voor vPC VTEP's.

```
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-configuratie: L2VPN-EVPN-buurten zijn nodig tussen rif en wervelkolom om de Type 3-routes te ruilen die nodig zijn om de VXLAN Xconnect in te stellen.

- Hier zijn de IP-adressen - 192.168.100.1 en 192.168.100.2 de Spines in de topologie. Normaal gesproken worden de L2VPN-buurten gevormd naar de Spines. Spines vormen alle Leaf switches in een iBGP Scenario als Routeswitchclients.
- Het wordt aanbevolen om afzonderlijke Loopbacks voor BGP/OSPF- en NVE-doeleinden te

gebruiken.

```
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
```

Opmerking: STP moet worden uitgeschakeld in het Xconnect-VLAN. MAC learning zal niet plaatsvinden binnen Xconnect VLAN wat betekent dat er in wezen geen Type 2 bgp l2vpn VPN-updates is voor MAC-adressen. Dit betekent dat verkeer van één VLAN ingekapseld zal worden met het IP-adres van de buitenbestemming dat is ingesteld op de Mcast-groep (239.30.30.30) gedefinieerd voor Xconnect VLAN.

Verifiëren

Gebruik dit gedeelte om te bevestigen dat de configuratie correct werkt.

1. BGP-wijk.

```
 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. Ontvang type 3-prefixes.

```
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. NU 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. ngo-controles.

```
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
```

Als de NGO-sessie eenmaal is begonnen, zouden de N3k's elkaar in de CDP zien. STP-BPDU's worden ook getunneld zodat de switches het ook eens worden over de plaatsing van de root-brug.

5. Verificaties bij de downstreamswitches.

```
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
```

Problemen oplossen

Er is momenteel geen specifieke troubleshooting-informatie beschikbaar voor deze configuratie.

Caveats

1. De dot1q tunnelinterfaces blijven steken in toestand met **fouten** in een Xconnect VXLAN-instelling als de configuraties binnen vPC-switches niet consistent zijn. Hieronder staan enkele gevallen waarin de interface foutloos wordt gemaakt;

- Als het VLAN to VN-segment niet op beide vPC-switches is gedefinieerd.
- Als de NVE to multicast groep niet op beide vPC-switches is gedefinieerd.
- Als de NGO's niet worden ontvangen (gebruik ethanalyzer met filter=**cfm** om de verstoppingspakketten van de NGO's te vangen).
- Zelfs als de dot1q tunnelinterface in een vPC-instelling is aangesloten, is het nog steeds vereist om de multicast groep onder de NVE-interface te configureren voor het VN-segment dat deel uitmaakt van Xconnect op beide switches.
- De kern van de NGO AM wordt verwerkt/verstuurd door de vPC Primaire switch. De hartslag berichten die op vPC secundair landen zullen worden gesynchroon aan de primaire

2. Wanneer Xconnect in een VLAN is geconfigureerd, wordt het verkeer van de ene locatie naar de andere ingekapseld met het buitenste doeladres=multicast adres dat onder de NVE-interface voor dat specifieke vn-segment is gedefinieerd. Het wordt aanbevolen om een unieke multicast groep voor Xconnect VLAN's te gebruiken. Multicast in de kern/wervelkolom moet functioneel zijn.

3. Multicastverkeer kan beide vPC-boxen op de afgelegen kant van Xconnect raken; De winnaar van Decap (het vakje dat de BUM kan decapsuleren) zal echter slechts één switch in een vPC-paar zijn. Dit kan worden geverifieerd met de opdracht-**show** het **verzenden van multicast routegroep <groepsadres> bron <SRC IP>**. Als de vlag die hier wordt afgebeeld een kleine case **v** is, betekent dit dat het vakje loser is en als het een Upper case **V** is, dan is de doos de decap winnaar, zodat het multicast verkeer kan decapsuleren en verder naar beneden kan sturen.

4. Op 93180YC-gebaseerde platforms, wanneer de host wordt aangesloten op 9k1 en er geen OIL voor S, G op 9k1 is, wordt een kopie van het multicast-pakket naar de vPC-peer verzonden met behulp van een speciale insluiting van IP-bron > 127.0.0.1 en IP-doelmap gedeeld door NVE IP-nummer en indien de 9k2 is gebruikt IL voor S, G inzending, dan zal de verkeersgeleiding door de 9k2 naar de afgelegen locaties worden verzorgd.

Packet Capture

Hier is een screenshot van een pakketvastlegging die tijdens de ruggengraat switch is genomen:

```
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
```

- Inner punt1q header=10 is bewaard gebleven
- Gebruikte VNI is 1003000 (wat de VPN-id van het buitenste VLAN is)
- Het IP-adres van de bestemming zou de multicast groep zijn die onder de NF-interface is gedefinieerd