Nexus 9000: Configurazione e verifica della connessione XLAN della VXLAN

Sommario

Introduzione

In questo documento viene descritto un rapido riferimento su come configurare e verificare le VXLAN Xconnect sugli switch Nexus 9000.

Prerequisiti

Requisiti

Cisco raccomanda la conoscenza di VXLAN EVPN.

Componenti usati

Le informazioni fornite in questo documento si basano sulle seguenti versioni software e hardware:

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

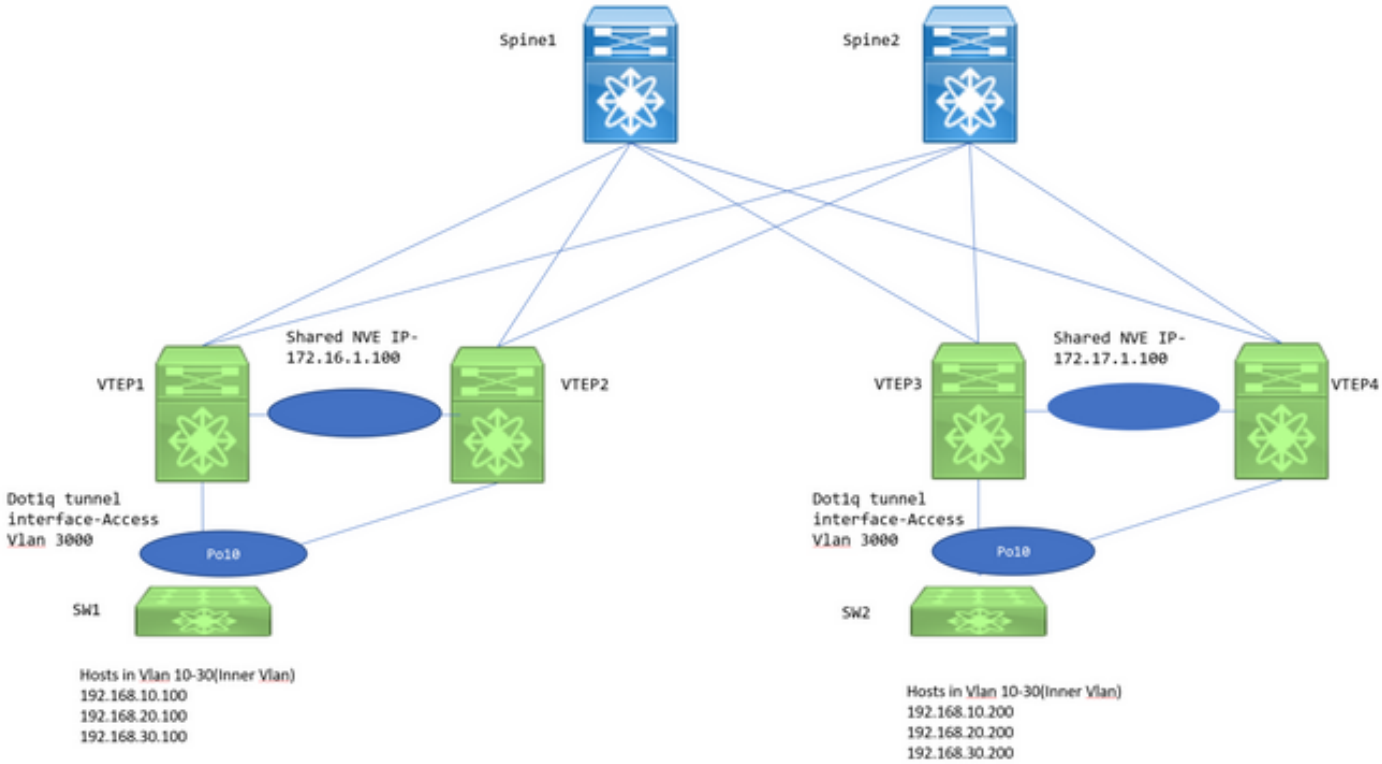
Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Panoramica

VXLAN Xconnect è un meccanismo per un tunnel point-to-point per pacchetti di dati e controllo da una foglia all'altra. I tag Inner Dot1q vengono conservati e la VXLAN incapsulata nel VNID esterno

specificato come VNID Xconnect. I frame di controllo di layer 2, come il protocollo LLDP (Link Layer Discovery Protocol), il protocollo CDP (Cisco Discovery Protocol), il protocollo STP (Spanning Tree Protocol) sono incapsulati tramite VXLAN e inviati ad altre estremità del tunnel.

Topologia



VTEP1, VTEP2, VTEP3 e VTEP4 sono due coppie di VTEP vPC configurate in modo che i tag dot1q interni degli switch in downstream vengano mantenuti e, quando la VXLAN è incapsulata, usare il VXLAN VNID dell'ID della VLAN esterna per inviarlo al VTEP remoto. Tutti i VTEP sono N9K-C93180YC-EX.

Gli switch in downstream sono Nexus 3k configurati con SVI (Switch Virtual Interface) nelle rispettive VLAN per simulare gli host.

Configurazione

1. La VLAN esterna utilizzata in questa topologia Xconnect è 3000. Si tratterebbe di quello con la configurazione VNID e Xconnect.

```
VTEP1# sh run vlan 3000

vlan 3000
  vn-segment 1003000
  xconnect
```

2. La funzionalità NGOAM deve essere abilitata e deve essere configurata.

```
VTEP1# sh run ngoam
```

```
feature ngoam

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

3. Configurazione del tunnel Dot1q per lo switch a valle.

```
VTEP1# sh run int po10

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

Le configurazioni vPC sono necessarie solo quando i VTEP vengono installati come vPC. In caso contrario, ignorare le configurazioni vPC menzionate in questo documento. VXLAN Xconnect è configurabile anche su un VTEP standalone.

4. Per gestire l'inoltro, il gruppo multicast deve essere definito nell'interfaccia NVE. Nota: abilitare la **modalità sparse ip pim** sugli uplink rilevanti e definire anche PIM RP in modo che il routing multicast e i messaggi PIM vengano scambiati correttamente. In genere, PIM RP viene definito sul livello della curva guida.

```
VTEP1# sh run int nve1

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

5. È necessario specificare la VLAN a infrarossi e consentirla come VLAN nativa all'interno del collegamento peer. Questo passaggio è necessario per i 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. Configurazione BGP/EVPN: Per scambiare le route di tipo 3 necessarie per stabilire la connessione VXLAN Xconnect, è necessario che tra le connessioni VPN L2VPN siano presenti router secondari con dorso.

- Gli indirizzi IP 192.168.100.1 e 192.168.100.2 sono gli aculei della topologia. In genere le aree adiacenti di L2VPN EVPN sono formate per gli Spines. Gli spine configurano tutti gli switch

foglia come client di riflessione della route in uno scenario iBGP.
- Si consiglia di utilizzare loopback separati per scopi BGP/OSPF e 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
```

> **Nota:** Il protocollo STP deve essere disabilitato nella VLAN Xconnect. L'apprendimento degli indirizzi MAC non verrà eseguito nella VLAN Xconnect, il che significa essenzialmente che non sono presenti aggiornamenti evpn bgp l2vpn di tipo 2 per gli indirizzi MAC. Per questo motivo, il traffico proveniente da un vtep verrà incapsulato con l'indirizzo IP di destinazione esterna impostato sul gruppo Mcast (239.30.30.30) definito per la VLAN Xconnect.

Verifica

Per verificare che la configurazione funzioni correttamente, consultare questa sezione.

1. Vicinato 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. Ricevi prefissi di tipo 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)
*>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. Controlli 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
```

Una volta terminata la sessione della NGOAM, i N3k si vedrebbero nel CDP. Anche le BPDU STP sono tunneling, quindi gli switch concordano sul posizionamento del bridge radice.

5. Verifiche sugli switch a valle.

```
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
```

Risoluzione dei problemi

Al momento non sono disponibili informazioni specifiche per la risoluzione dei problemi di questa configurazione.

Avvertenze

1. Le interfacce del tunnel dot1q saranno bloccate in **stato err-disabled** in una configurazione VXLAN Xconnect se le configurazioni negli switch vPC non sono coerenti. Di seguito sono riportati alcuni dei casi in cui l'interfaccia verrà disabilitata a causa di un errore;

- Se il segmento da VLAN a VN non è definito su entrambi gli switch vPC.
- Se il gruppo da NVE a multicast non è definito su entrambi gli switch vPC.
- Se gli heartbeat NGOAM non vengono ricevuti (usare ethanalyzer con filter=**cfm** per catturare i pacchetti heartbeat NGOAM).
- Anche se l'interfaccia del tunnel dot1q è orfana e connessa in una configurazione vPC, è comunque necessario configurare il gruppo multicast nell'interfaccia NVE per il segmento VN che fa parte di Xconnect su entrambi gli switch.
- Gli heartbeat NGOAM vengono elaborati/inviati dallo switch primario vPC. I messaggi heartbeat che atterrano sul vPC secondario verranno sincronizzati sul server primario

2. Quando Xconnect è configurato in una VLAN, il traffico tra due siti è incapsulato con l'indirizzo di destinazione esterno=indirizzo multicast definito nell'interfaccia NVE per quel particolare segmento VLAN. Si consiglia di utilizzare un gruppo multicast univoco per le VLAN Xconnect. Il multicast nel core/dorso deve essere funzionale.

3. Il traffico multicast potrebbe colpire entrambi i vPC box sul lato remoto di Xconnect; Tuttavia, il vincitore Decap (la scatola che può decapsulare il BUM) sarà solo uno switch in una coppia di vPC. È possibile verificare questa condizione tramite il comando **show forwarding multicast route group <indirizzo gruppo> source <IP SRC>**. Se il Flag mostrato qui è un **v** minuscolo, significa che la scatola è decaffeinato e se è un **V** maiuscolo, la scatola è il vincitore decap e quindi può decapsulare il traffico multicast e inoltrarlo ulteriormente verso il basso.

4. Sulle piattaforme basate su 93180YC, quando l'host è orfano e collegato a 9k1 e se non c'è OIL per S, G su 9k1, una copia del pacchetto multicast viene inviata al peer vPC utilizzando uno speciale incapsulamento di Source IP-> 127.0.0.1 e Destination IP-> shared NVE IP e se il 9k2 ha OIL per S, G entry, allora il traffico di inoltro sarà gestito dai 9k2 verso i siti remoti.

Acquisizione pacchetti

Di seguito è riportata la schermata di un'acquisizione presa dal commutatore sul dorso:

```
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 dot1q header=10 viene mantenuto
- Il VNI utilizzato è 1003000 (ossia il VNID della VLAN esterna)
- L'indirizzo IP di destinazione sarà il gruppo multicast definito nell'interfaccia nve