

Nexus 9000: VXLAN Xconnect 구성 및 확인

목차

[소개](#)

[사전 요구 사항](#)

[요구 사항](#)

[사용되는 구성 요소](#)

[개요](#)

[토폴로지](#)

[구성](#)

[다음을 확인합니다.](#)

[문제 해결](#)

[주의 사항](#)

[패킷 캡처](#)

소개

이 문서에서는 Nexus 9000 스위치에서 VXLAN Xconnect를 구성하고 확인하는 방법에 대한 빠른 참조를 설명합니다.

사전 요구 사항

요구 사항

Cisco에서는 VXLAN EVPN에 대해 알고 있는 것이 좋습니다.

사용되는 구성 요소

이 문서의 정보는 다음 소프트웨어 및 하드웨어 버전을 기반으로 합니다.

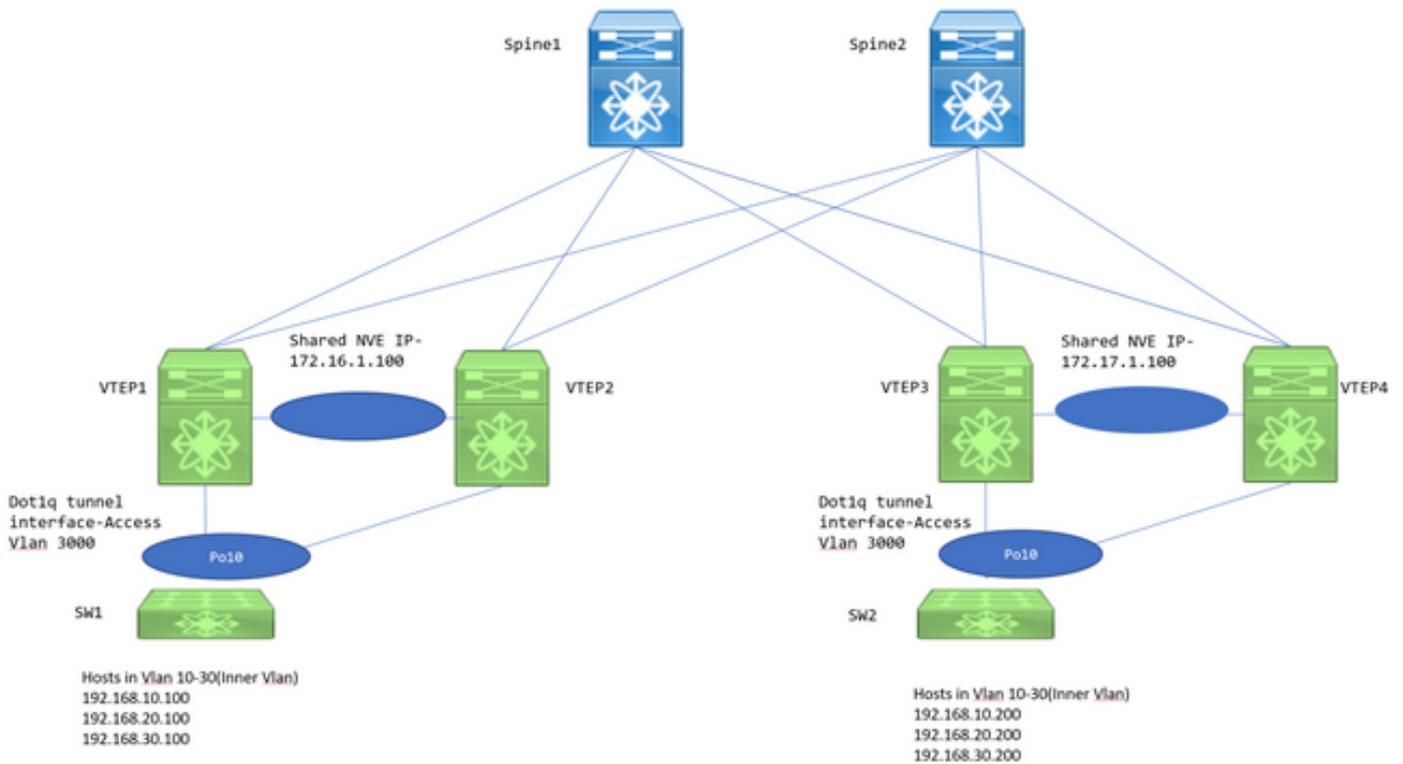
- N9K-C93180YC-EX
- NXOS 9.2(1)

이 문서의 정보는 특정 랩 환경의 디바이스를 토대로 작성되었습니다. 이 문서에 사용된 모든 디바이스는 초기화된(기본) 컨피그레이션으로 시작되었습니다. 현재 네트워크가 작동 중인 경우, 모든 명령어의 잠재적인 영향을 미리 숙지하시기 바랍니다.

개요

VXLAN Xconnect는 데이터를 위한 포인트 투 포인트 터널의 메커니즘으로, 한 리프에서 다른 리프로 패킷을 제어합니다. 내부 Dot1q 태그는 유지되며 Xconnect VNID로 지정된 외부 VNID 내에서 VXLAN이 캡슐화됩니다. LLDP(Link Layer Discovery Protocol), CDP(Cisco Discovery Protocol), STP(Spanning Tree Protocol)와 같은 레이어 2 제어 프레임은 VXLAN으로 캡슐화되어 터널의 다른 끝으로 전송됩니다.

토폴로지



VTEP1, VTEP2, VTEP3 및 VTEP4는 다운스트림 스위치의 내부 dot1q 태그가 보존되고 VXLAN이 캡슐화된 경우 원격 VTEP로 전송하기 위해 외부 VLAN ID의 VXLAN VNID를 사용합니다. 모든 VTEP는 N9K-C93180YC-EX입니다.

다운스트림 스위치는 호스트를 모방하기 위해 각 VLAN에서 SVI(Switch Virtual Interface)로 구성된 Nexus 3k입니다.

구성

1. 이 Xconnect 토폴로지에서 사용되는 외부 VLAN은 3000입니다. 이는 VNID 및 Xconnect 컨피그레이션이 있는 것입니다.

```
VTEP1# sh run vlan 3000
```

```
vlan 3000  
  vn-segment 1003000  
  xconnect
```

2. NOGOAM 기능을 활성화해야 하며 이 구성이 필요합니다.

```
VTEP1# sh run ngoam
```

```
feature ngoam
```

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

3. 다운스트림 스위치에 대한 Dot1q 터널 구성

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

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

vPC 구성은 VTEP가 vPC로 구축된 경우에만 필요합니다. 그렇지 않으면 이 문서에 언급된 vPC 컨피그레이션을 건너뛴니다. VXLAN Xconnect는 독립형 VTEP에서도 구성할 수 있습니다.

4. 전송을 처리하려면 NVE 인터페이스에서 멀티캐스트 그룹을 정의해야 합니다. 관련 업링크에서 **ip pim sparse-mode**를 활성화하고 멀티캐스트 라우팅 및 PIM 메시지가 적절하게 교환되도록 PIM RP를 정의합니다. 일반적으로 PIM RP는 스파인 레이어에서 정의됩니다.

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

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

5. 피어 링크 내에서 인프라 VLAN을 지정하고 네이티브 VLAN으로 허용해야 합니다. 이 단계는 vPC VTEP에 필요합니다.

```
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 구성: VXLAN Xconnect를 설정하는 데 필요한 Type 3 경로를 교환하려면 리프/스파인 간에 L2VPN EVPN 인접 디바이스가 필요합니다.

- 여기서 IP 주소 192.168.100.1 및 192.168.100.2은 토폴로지의 Spine입니다. 일반적으로 L2VPN EVPN 인접 디바이스는 Spine에 형성됩니다. 스파인은 iBGP 시나리오에서 모든 리프 스위치를 경로 리플렉터 클라이언트로 구성합니다.
- BGP/OSPF 및 NVE를 위해 별도의 루프백을 사용하는 것이 좋습니다.

```
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

```

참고: Xconnect VLAN 내에서 STP를 비활성화해야 합니다. MAC 학습은 Xconnect VLAN 내에서 수행되지 않습니다. 이는 기본적으로 MAC 주소에 대한 Type 2 bgp l2vpn 이벤트 업데이트가 없음을 의미합니다. 이로 인해 한 vtep의 트래픽은 Xconnect VLAN에 대해 정의된 Mcast-group(239.30.30.30)으로 외부 대상 IP 주소로 캡슐화됩니다.

다음을 확인합니다.

이 섹션을 사용하여 컨피그레이션이 제대로 작동하는지 확인합니다.

1. BGP 인접 디바이스

```

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. 수신 유형 3 접두어.

```

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. 제10조 피어링

```
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. 응고지 확인

```
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
```

NOGOAM 세션이 시작되면 N3k는 CDP에서 서로를 볼 수 있습니다. STP BPDU도 터널링되므로 스위치도 루트 브리지 배치에 동의합니다.

5. 다운스트림 스위치의 확인

```
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
```

문제 해결

현재 이 컨피그레이션에 사용할 수 있는 특정 문제 해결 정보가 없습니다.

주의 사항

1. vPC 스위치 내의 컨피그레이션이 일치하지 않을 경우 Xconnect VXLAN 설정에서 dot1q 터널 인터페이스가 오류 비활성화 상태로 유지됩니다. 다음은 인터페이스가 error disable된 경우입니다.

- 두 vPC 스위치 모두에서 VLAN-to-VN 세그먼트가 정의되지 않은 경우.
- 멀티캐스트 그룹에 대한 NVE가 두 vPC 스위치 모두에서 정의되지 않은 경우
- NGOAAM 하트비트가 수신되지 않은 경우(NGOOAM 하트비트 패킷을 포착하려면 ethanalyzer with filter=**cfm**을 사용합니다).
- dot1q 터널 인터페이스가 vPC 설정에 연결된 고아 인터페이스일지라도 두 스위치에서 모두 Xconnect의 일부인 VN 세그먼트의 NVE Interface에서 멀티캐스트 그룹을 구성해야 합니다.

- vPC 기본 스위치에서 NGOAAM 하트비트를 처리/전송합니다. vPC 보조 서버에 있는 하트비트 메시지는 기본 메시지와 동기화됩니다.

2. Xconnect가 VLAN에 구성된 경우 한 사이트에서 다른 사이트로의 트래픽은 해당 특정 vn 세그먼트에 대해 NVE 인터페이스 아래에 정의된 외부 목적지 주소=멀티캐스트 주소로 캡슐화됩니다. Xconnect VLAN에 고유한 멀티캐스트 그룹을 사용하는 것이 좋습니다. 코어/스파인의 멀티캐스트가 작동해야 합니다.

3. 멀티캐스트 트래픽이 Xconnect의 원격지에 있는 vPC상자 모두에 도달할 수 있을 것 그러나 Decap 상자(BUM을 역캡슐화할 수 있는 상자)는 vPC 쌍에서 하나의 스위치만 됩니다. 이는 명령 - show forwarding multicast route group <Group address> source <SRC IP>를 사용하여 확인할 수 있습니다. 여기에 표시된 Flag가 소문자 v인 경우, 상자가 패소되어 있고 Upper case V인 경우, 디맵이 선택되므로 멀티캐스트 트래픽을 역캡슐화하고 더 아래로 전달할 수 있습니다.

4. 93180YC 기반 플랫폼에서 호스트가 9k1에 연결되어 있고 S, G, 9k1에 OIL이 없을 경우, 멀티캐스트 패킷의 복사본이 소스 IP-> 127.0.0.1 및 대상 IP-> 공유 IP를 사용하여 vPC 피어로 전송되고 9k2에 S, G, G에 대한 OIL이 있을 경우 포워딩에 주의하여 트래픽을 처리합니다. 원격 사이트로 9k2를 연결합니다.

패킷 캡처

다음은 스파인 스위치에서 촬영한 패킷 캡처의 스크린샷입니다.

```

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

```

- 내부 dot1q 헤더=100이 보존됩니다.
- 사용된 VNI는 1003000입니다(외부 VLAN의 VNID).
- 대상 IP 주소는 nve 인터페이스 아래에 정의된 멀티캐스트 그룹이 됩니다.