

Nexus 9000 : VXLAN Xconnectの設定と確認

内容

[概要](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[概要](#)

[トポロジ](#)

[設定](#)

[確認](#)

[トラブルシューティング](#)

[警告](#)

[パケットキャプチャ](#)

概要

このドキュメントでは、Nexus 9000スイッチでVXLAN Xconnectを設定および確認する方法のクイックリファレンスについて説明します。

前提条件

要件

VXLAN EVPNに関する知識があることが推奨されます。

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

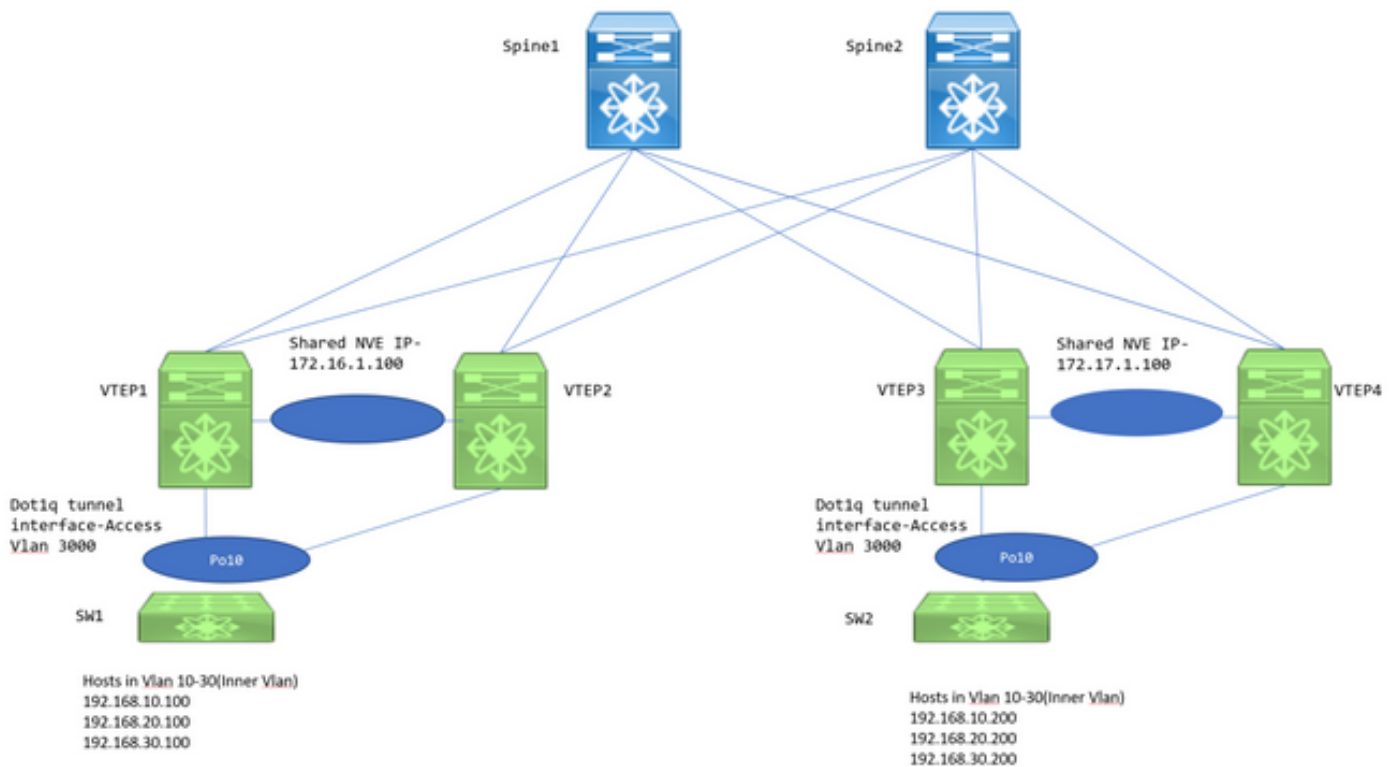
- N9K-C93180YC-EX
- NXOS 9.2(1)

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期（デフォルト）設定の状態から起動しています。対象のネットワークが実稼働中である場合には、どのようなコマンドについても、その潜在的な影響について確実に理解しておく必要があります。

概要

VXLAN Xconnectは、あるリーフから別のリーフへのデータおよび制御パケット用のポイントツーポイントトンネルのメカニズムです。内部Dot1qタグは保持され、Xconnect VNIDとして指定された外部VNID内でVXLANがカプセル化されます。Link Layer Discovery Protocol(LLDP)、Cisco Discovery Protocol(CDP)、スパニングツリープロトコル(STP)などのレイヤ2制御フレームはVXLANカプセル化され、トンネルの他の端に送信されます。

トポロジ



VTEP1、VTEP2、VTEP3、およびVTEP4は、ダウンストリームスイッチからの内部dot1qタグが保持されるように設定された2つのvPC VTEPペアで、VXLANカプセル化の際に外部VLAN IDのVXLAN VNID(VTEP)を... すべてのVTEPはN9K-C93180YC-EXです。

ダウンストリームスイッチはNexus 3ksで、各VLANにスイッチ仮想インターフェイス(SVI)を設定してホストを模倣します。

設定

1.このXconnectトポロジで使用される外部VLANは3000です。これは、VNIDとXconnectの設定を持つものです。

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

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

2.機能NGAMを有効にする必要があり、この設定が必要です。

```
VTEP1# sh run ngoam
```

```
feature ngoam
```

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

3.ダウンストリームスイッチへのDot1qトンネル設定。

```
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の設定は、VTEPがvPCとして導入されている場合にのみ必要です。それ以外の場合は、このドキュメントで説明されているvPC設定をスキップします。VXLAN Xconnectは、スタンドアロンVTEPでも設定できます。

4.転送を処理するには、NVEインターフェイスでマルチキャストグループを定義する必要があります。関連するアップリンクで**ip pim sparse-mode**を有効にすると、マルチキャストルーティングとPIMメッセージが適切に交換されるようにPIM RPも定義することに注意してください。通常、PIM RPはスパインレイヤで定義されます。

```
VTEP1# sh run int nve1

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

5.インフラストラクチャVLANを指定し、ピアリンク内のネイティブVLANとして許可する必要があります。この手順は、vPC VTEPに必要です。

```
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設定：L2VPN EVPNネイバーシップは、VXLAN Xconnectの確立に必要なタイプ3ルートを交換するために、リーフ/スパイン間で必要です。

- ここでは、IPアドレス192.168.100.1と192.168.100.2がトポロジ内のスパインです。通常、L2VPN EVPNネイバーシップはスパインに対して形成されます。スパインは、すべてのリーフスイッチをiBGPシナリオのルートリフレクタクライアントとして設定します。
- BGP/OSPFおよびNVEの目的には、別々のループバックを使用することを推奨します。

```
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

```

注：Xconnect VLAN内でSTPを無効にする必要があります。MACラーニングはXconnect VLAN内では発生しません。これは基本的に、MACアドレスに対するType 2 bgp l2vpn evpnアップデートがないことを意味します。そのため、1つのvtepからのトラフィックは、Xconnect VLANに定義されたMcast-group(239.30.30.30)に設定されます。

確認

ここでは、設定が正常に機能しているかどうかを確認します。

1. BGPネイバーシップ。

```

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. タイプ3プレフィクスの受信

```

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					

3. NVEピアリング。

```
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チェック。

```
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
```

NGOAMセッションが起動すると、N3kはCDPで互いを認識します。STP BPDUもトンネリングされるため、スイッチはルートブリッジの配置についても同意します。

5. ダウンストリームスイッチでの検証。

```
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
```

トラブルシューティング

現在、この設定に関する特定のトラブルシューティング情報はありません。

警告

1. dot1qトンネルインターフェイスが、vPCスイッチ内の設定に一貫性がない場合は、Xconnect VXLANセットアップで**error disabled**状態のままになります。インターフェイスがエラーディセーブル状態になる場合を次に示します。

- VNセグメントへのVLANが両方のvPCスイッチで定義されていない場合。
- NVEからマルチキャストグループが両方のvPCスイッチで定義されていない場合。
- NGOAMハートビートが受信されない場合(NGOAMハートビートパケットを捕捉するには、ethanalyzerにfilter=cfmを使用します)。

- dot1qトンネルインターフェイスがvPCセットアップで孤立して接続されている場合でも、両方のスイッチでXconnectの一部であるVNセグメントのNVEインターフェイスでマルチキャストグループを設定する必要があります。
- NGOAMハートビートは、vPCプライマリスイッチによって処理/送信されます。vPCセカンダリに到着したハートビートメッセージがプライマリに同期される

2. XconnectがVLANに設定されている場合、あるサイトから別のサイトへのトラフィックは、その特定のvnセグメントのNVEインターフェイスで定義された外部宛先アドレス=マルチキャストアドレスでカプセル化されます。Xconnect VLANには一意のマルチキャストグループを使用することを推奨します。コア/スパインのマルチキャストは機能している必要があります。

3. マルチキャストトラフィックがXconnectのリモート側の両方のvPCボックスにヒットする可能性があります。ただし、Decapの勝者 (BUMをカプセル化解除できるボックス) は、vPCペアの1つのスイッチのみになります。これは、**show forwarding multicast route group <Group address> source <SRC IP>** コマンドを使用して確認できます。ここで示すフラグが小文字のvの場合は、ボックスがdecap loserであることを意味し、大文字のVの場合は、ボックスがdecapの勝者であるため、マルチキャストトラフィックをカプセル化してさらに転送できます。

4. 93180YCベースのプラットフォームでは、ホストが9k1に孤立接続され、9k1にS、G用のOILがない場合、マルチキャストパケットのコピーがvPCピアに送信されます。これは、送信元IP-> 127.0.0.1および宛先IP共有NVE k2にはS、Gエントリ用のOILがあり、その後、9k2によってリモートサイトへのトラフィック転送が処理されます。

パケット キャプチャ

スパインスイッチで取得したパケットキャプチャのスクリーンショットを次に示します。

```

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

```

- 内部dot1qヘッダー=10は保持されます
- 使用されるVNIは1003000 (外部VLANのVNID)
- 宛先IPアドレスは、nveインターフェイスで定義されたマルチキャストグループです