# EVPN

Ethernet VPN är en L2VPN teknik som ska lösa de brister som finns i VPLS. Information om MAC-adresser transporteras via en egen BGP AFI.

EVPN inkluderar Ethernet over MPLS och Ethernet over VXLAN. Den sistnämnda är populär i nutida datacenterimplementationer.

En variant av EVPN som finns är PBB-EVPN (provider backbone).

# EVPN med VXLAN

BGP är kontrollplan och VXLAN är dataplan. RFC 7348. MAC eller MAC och IP enkapsuleras i UDP 4789. Trafik enkapsuleras och de-enkapsuleras av VTEP:ar.

## Begrepp

- **BUM** - Broadcast, multicast eller unknown unicast.
- **EVI** - Ethernet VPN Identifier.
- **ES** - Ethernet Segment.
- **ESI** - Ethernet Segment Identifier. Varje ES identifieras via en ESI. 10 bytes stort.
- **I-SID** - Instance Service Identifier. En identifierar som informerar en router hur L2-paketet från kunden ska mappas. 24 bytes stort värde.
- **VNI** - Virtual Network Instance. En logiskt nätverksinstans som hanterar L2 eller L3 tjänster och definierar en L2 broadcastdomän
- **VNID** - Virtual Network Identifier. 24 bitar stort segment ID, tillåter upp till 16 miljoner segment. VLAN mappas till VNID
- **VTEP** - Virtual tunnel endpoint. Routrar/switchar. Är i verkligheten i en loopbackadress.
- **NVE** - Network Virtualization Edge. Enhet som implementerar L2 och/eller L3 virtualisering

## BGP rötter / EVPN Route-Types

Ett antal typer av BGP rötter används i EVPN AFI.

| Route typ | Beskrivning |
| --- | --- |

| | |
|---|---|
| 0. Reserverad | |
| 1. Ethernet active discovery (AD) route | Används för att signalera tillbakadragande av MAC-adresser, aliasing och split-horizon etiketter. |
| 2. MAC/IP advertisement route | Innehåller MAC-adresser med EVPN ESI:er och MPLS etiketter. Måste innehålla en MAC-adress men kan innehålla IP-adress. |
| 3. Inclusive multicast route | Innehåller attributer för att representera ingress replication. |
| 4. Ethernet Segment (ES) route | |
| 5. IP Prefix Route | En route för ett helt nät. |
| 6. Selective Multicast Ethernet Tag Route | |
| 7. IGMP Join Synch Route | |
| 8. IGMP Leave Synch Route | |
| 9. Per-Region I-PMSI A-D route | |
| 10. S-PMSI A-D route | |
| 11. Leaf A-D route | |
| 12-255. Unassigned | |

# Typ 2

En typ 2 innehåller information tillhörande en host. En route typ 2 måste innehålla:

- MAC Address (/48)
- MPLS Label1 (L2VNI)
- Route Target for MAC-VRF

Den kan innehålla:

- IP Address (/32 eller /128)
- MPLS LAbel2 (3VNI*)
- Route Target för IP-VRF
- Router MAC

NVE lär sig IP-attribut genom ARP eller ND. Typ 2 importeras i en viss VRF/MAC-VRF.

# Typ 5

En route typ 5 måste innehålla:

- IP Prefix
- MPLS Label (L3VNI)
- Route Target för IP VRF
- Router MAC

Den kan även innehålla Gateway-IP. NVE lär sig IP-attribut genom redistribuering eller routing-protokol.

# Symmetric/asymmetric IRB

När ett paket transporteras mellan två VXLAN/EVPN routrar så märks paketet med ett VNID (Virtual Network Identifier). Beroende på vilken leverantör (och kanske konfiguration?) så används antingen symmetrisk eller asymmetrisk IRB.

Med asymmetrisk IRB så används taggas det routade paketet med VNID:t som används för det lokala segmentet på routern som paketet routas till. För att det ska fungera måste VNID:t finnas definierat på både den lokalswitchen och fjärrswitchen. I stora nät där det kan finnas enormt många VNI:er så skalar det här dåligt.

Med symmetrisk IRB så används samma, utpekade, VNI för att tagga paketen. I nve-interface konfigurationen pekas VNI ut. Kallas för transit L3 VNI. En transit L3 VNI används per tenant (VRF).

# Ingress replication

Används ingress replication för att transportera BUM-trafik (istället för multicast) så används typ 3 rötter för att dessa. Grannar går att se med `show nve vni ingress-replication`.

Ingress replication-grannar konfigureras statiskt. Se exempel (NXOS):

```
interface nve1
  no shutdown
  source-interface loopback1
  member vni 1000
    ingress-replication protocol static
      peer-ip 203.0.113.2
      peer-ip 203.0.113.3
      peer-ip 203.0.113.4
```

# ARP Suppression

ARP Suppression används för att minska ARP-trafiken inom en fabric. Närmaste VTEP till sluthosten blir proxy för samtliga ARP-frågor. Aktiveras per L2VNI. Verifiering med `show ip arp suppression-cache { detail }`.

Exempelkonfiguration (NXOS):

```
interface nve1
 no shutdown
 source-interface loopback1
 host-reachability protocol bgp
 member vni 202020 associate-vrf
 member vni 301010
  mcast-group 239.0.0.1
  suppress-arp
```

Switchen kommer då att snoopa efter ARP-frames och bygga Typ 2-rötter som populeras till grannar. Om en ARP-fråga kommer till en switch skickas den inte vidare till hosten utan switchen svarar direkt på frågan med sluthostens MAC-address. Alltså fungerar det inte som proxy-arp där switchen skulle svara med sin egen mac-address.

---

Revision #1
Created 26 April 2025 12:03:45 by Jehrlander
Updated 26 April 2025 12:03:48 by Jehrlander