IP Prefix Advertisement in EVPN

draft-rabadan-l2vpn-evpn-prefix-advertisement-00

Jorge Rabadan

Wim Hendericks

Florin Balus
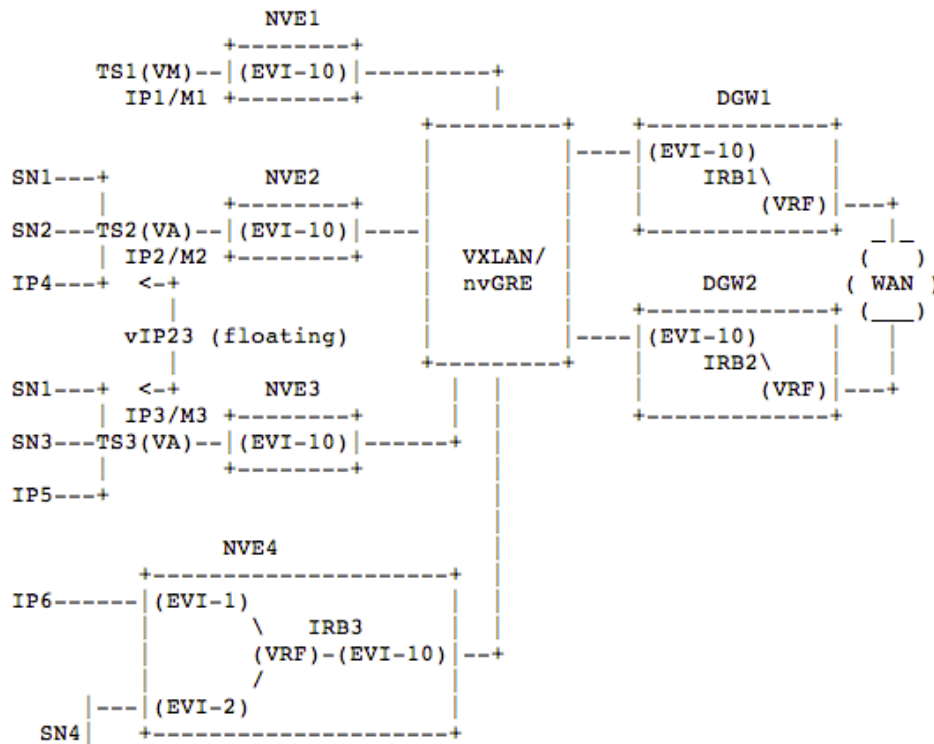
Senad Palislamovic

Aldrin Isaac

Problem statement: flexible inter-subnet connectivity is required within the DC

```
                      NVE1
                    +--------+
     TS1(VM)--|(EVI-10)|----------+              DGW1
        IP1/M1 +--------+         |        +-------------+
                                  |   +----|(EVI-10)     |
SN1---+           NVE2            |   |    | IRB1\       |
      |         +--------+        |   |    |    (VRF)|---+
SN2---TS2(VA)--|(EVI-10)|----     |   |    +-------------+   _|_
      | IP2/M2 +--------+    |  VXLAN/  |                   (   )
IP4---+  <-+                 |  nvGRE   |    DGW2          ( WAN )
         |                   |        +-------------+      (___)
      vIP23 (floating)       |   +----|(EVI-10)     |
         |                   |   |    | IRB2\       |
SN1---+  <-+      NVE3       | | |    |    (VRF)|---+
      | IP3/M3 +--------+    | | |    +-------------+
SN3---TS3(VA)--|(EVI-10)|------+ |
      |         +--------+      |
IP5---+                         |
                                |
             NVE4               |
        +---------------------+ |
IP6------|(EVI-1)            | |
        |        \    IRB3   | |
        |        (VRF)-(EVI-10)|--+
        |        /            |
        |---|(EVI-2)          |
     SN4|   +---------------------+
```

TS = Tenant System

VM = Virtual Machine

VA = Virtual Appliances

SN = Subnet

EVI = EVPN instance

IRB = Integrated Routing and Bridging interface
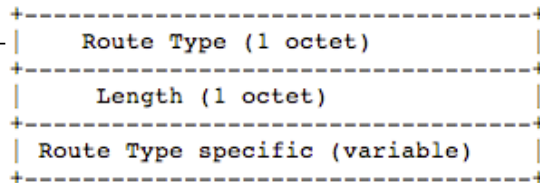
NVE = Network Virtualization Edge

- Subnets can be located:
  - Behind IRB interfaces
  - Behind VA (Virtual Appliances) such as Firewalls, Load Balancers, NAT routers

- Virtual Appliance characteristics:
  - They do not run dynamic routing protocols
  - They support basic redundancy mechanisms based on a floating IP
  - They can move

- EVPN NVE/CP requirements:
  - Need to advertise IP Prefixes on behalf of the VAs (but not route)
  - IP Prefix next-hop is hosted in the NVE and can be an IRB, an overlay floating IP, a MAC or an ESI

NVEs must advertise prefixes from different overlay next-hops

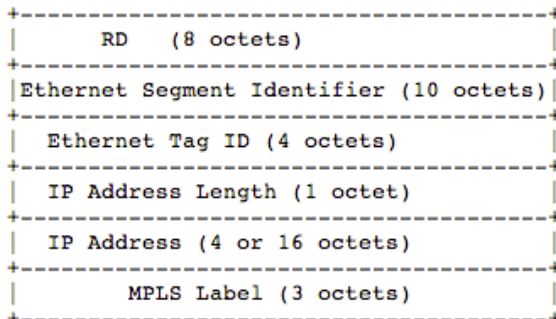# Solution: The prefix-advertisement route (route-type 5)

```
+-----------------------------------+
|         Route Type (1 octet)      |
+-----------------------------------+
|          Length (1 octet)         |
+-----------------------------------+
|     Route Type specific (variable)|
+-----------------------------------+
```

Route-types:

1- Ethernet A-D route

2- MAC-advertisement route

3- Inclusive Multicast route

4- Ethernet Segment route

5- IP Prefix route (OPTIONAL)

Route-type 5 specific content

```
+-----------------------------------+
|         RD    (8 octets)          |
+-----------------------------------+
|Ethernet Segment Identifier (10 octets)|
+-----------------------------------+
|       Ethernet Tag ID (4 octets)  |
+-----------------------------------+
|     IP Address Length (1 octet)   |
+-----------------------------------+
|     IP Address (4 or 16 octets)   |
+-----------------------------------+
|         MPLS Label (3 octets)     |
+-----------------------------------+
```

- RT 2 (MAC-advertisement) used for advertising MACs and MAC-IP for ARP resolution)

  – Also used for MAC Mobility

- RT 5 (IP-prefix) used to advertise Prefixes independently of the mac-advertisement routes

- Possible Prefix next-hops in the overlay topology

  – ESI, MAC address, IRB IP address, Floating IP address

The prefix-advertisement route benefits

- Clean identification of a prefix

- NVEs running EVPN but not supporting the OPTIONAL prefix-advertisement route can easily identify it an ignore it without processing the route

- No MAC information is compared by BGP when selecting routes

- Flexible overlay next-hop (IRB, floating IP, MAC, ESI) addressing different end-point types

- Decouples prefix-advertisement from mac mobility procedures

- Supports VA resiliency procedures

Next steps

- Feedback from WG

- Merge or keep it as separate draft?