EVPN Support for L3 Fast Convergence and Aliasing/Backup Path

draft-sajassi-bess-evpn-ip-aliasing-02

A. Sajassi (Cisco)
G. Badoni (Cisco)
P. Warade (Cisco)
S. Pasupula (Cisco)
J. Drake (Juniper)
J. Rabadan (Nokia)
Agenda

Introduction

Use-cases in rev 02

Conclusions and Next Steps
**Introduction**

**History**
- Version 0 was presented in IETF99 as an extension of the EVPN Aliasing procedures for EVPN Symmetric IRB model (MAC/IP routes)
- Version 1 added support for EVPN IP Prefix routes

**The current document**
Revision 2 clarifies all the use cases supported, including:
- Ethernet Segments for Host Routes in Symmetric IRB
- EVPN Layer-3 Interface-less IP-VRF-to-IP-VRF model
- Ethernet Segments for IP Prefix routes

Clarifies:
- how IP A-D per ES (Fast Convergence) routes are constructed
- How IP A-D per EVI (Aliasing and Backup) routes are constructed
Refreshing of the RFC7432 Aliasing procedures

**EVPN Routes**
- Leaf-1 advertises the MAC/IP route with ESI-1.
- Leaf-1 and Leaf-2 advertise an A-D per ES route for ES-1.
- Leaf-1 and Leaf-2 signal their attachment to ESI-1 in the BD via A-D per-EVI routes.

**Bridge Table**
- MAC-VRF BD 1
  - MAC M1
    - Destination: M1 / ESI-1 (Leaf-1, Leaf-2)
    - Type: EVPN
  - MAC M2
    - Destination: AC-5
    - Type: Learned

**Leaf-1**
- MAC-VRF BD 1
  - MAC M1
    - Type: Lag-1
  - MAC M2
    - Destination: Leaf-5
    - Type: EVPN

**Leaf-2**
- MAC-VRF BD 1
  - MAC M1
    - Type: Lag-1
  - MAC M2
    - Destination: Leaf-5
    - Type: EVPN

**Leaf-5**
- MAC DA = M1
- MAC SA = M2
- Payload

**CE2** (MAC M2)
- Leaf-5 installs M1 with two aliasing paths associated to ESI-1, and does per flow load balancing for the traffic destined to M1.

**Leaf-2** installs M1 with a destination of the ES attachment circuit.

**Leaf-1 receives MAC M1**

**EVPN Routes**
- Leaf-1 advertises the MAC/IP route with ESI-1.
- Leaf-1 and Leaf-2 advertise an A-D per ES route for ES-1.
- Leaf-1 and Leaf-2 signal their attachment to ESI-1 in the BD via A-D per-EVI routes.

**Bridge Table**
- MAC-VRF BD 1
  - MAC M1
    - Destination: M1 / ESI-1 (Leaf-1, Leaf-2)
    - Type: EVPN
  - MAC M2
    - Destination: AC-5
    - Type: Learned

**Leaf-1**
- MAC-VRF BD 1
  - MAC M1
    - Type: Lag-1
  - MAC M2
    - Destination: Leaf-5
    - Type: EVPN

**Leaf-2**
- MAC-VRF BD 1
  - MAC M1
    - Type: Lag-1
  - MAC M2
    - Destination: Leaf-5
    - Type: EVPN

**Leaf-5**
- MAC DA = M1
- MAC SA = M2
- Payload

**CE2** (MAC M2)
- Leaf-5 installs M1 with two aliasing paths associated to ESI-1, and does per flow load balancing for the traffic destined to M1.

**Leaf-2** installs M1 with a destination of the ES attachment circuit.

**Leaf-1 receives MAC M1**
Inter-Subnet Forwarding for host routes (RT2)
Aliasing for Symmetric IRB

1. Leaf-1 receives an ARP packet from CE1 and generates a MAC/IP route

2. Leaf-2 installs (synchronizes) 10.0.0.1 with a destination of the ES attachment circuit

3. Leaf-5 installs 10.0.0.1/32 with two aliasing paths associated to ESI-1, and does per flow load balancing for the traffic destined to 10.0.0.1 (Primary/Backup is also possible)

4. Leaf-5 installs 10.0.0.1/32 with two aliasing paths associated to ESI-1, and does per flow load balancing for the traffic destined to 10.0.0.1 (Primary/Backup is also possible)

EVPN Routes
- Leaf-1 advertises the MAC/IP route with ESI-1.
- Leaf-1 and Leaf-2 advertise IP A-D per ES routes for ESI-1.
- Leaf-1 and Leaf-2 signal their attachment to ESI-1 in the IP-VRF via IP A-D per-EVI routes.

EVPN MAC/IP route
- M1 / ESI-1
- L2/L3 label
- MAC/IP-VRF route-target
- NH Leaf-1

BGP EVPN

IP-VRF Route-Table
Prefix        Next-hop  Type
10.0.0.1/32   Leaf-5    IRB
(arp)
20.0.0.1/32   Leaf-5    EVPN
RT2

IP A-D per ES route
IP A-D per EVI route
ESI-1
MAC/IP-VRF route-target
NH Leaf-1

Leaf-2 installs (synchronizes) 10.0.0.1 with a destination of the ES attachment circuit

IP-VRF Route-Table
Prefix        Next-hop  Type
10.0.0.1/32   Leaf-5    EVPN
RT2
20.0.0.1/32   Leaf-5    EVPN
RT2

IP-VRF Route-Table
Prefix        Next-hop  Type
10.0.0.1/32   IRB       EVPN
RT2
20.0.0.1/32   Leaf-5    EVPN
RT2

IP-VRF Route-Table
Prefix        Next-hop  Type
10.0.0.1/32   ESI-1     evpn
(10.0.0.1)
Leaf-1

CE2
(IP 20.0.0.1)

IP DA = 10.0.0.1
IP SA = 20.0.0.1
Payload
Inter-Subnet Forwarding for prefix routes (RT5)
Aliasing for L2 and L3 Interface-less model

EVPN Routes
- Leaf-1 advertises the IP Prefix route with ESI-1.
- Leaf-1 and Leaf-2 advertise IP A-D per ES routes for ES-1
- Leaf-1 and Leaf-2 signal their attachment to ESI-1 in the IP-VRF via IP A-D per-EVI routes

BGP EVPN

EVPN IP Prefix route
10.0.0.1 / ESI-1
L3 label
NH Leaf-1
IP-VRF route-target

IP-VRF Route-Table
Prefix          Next-hop  Type
10.0.0.1/32     IRB      local
20.0.0.1/32     Leaf-5   EVPN

Leaf-1 receives an ARP packet from CE1 and generates a host route which is advertised in an IP Prefix route (RT5)

Leaf-1 advertises the IP Prefix route with ESI-1.
IP A-D per ES route
IP A-D per EVI route
ESI-1
IP-VRF route-target
NH Leaf-1

Leaf-2 installs (synchronizes) 10.0.0.1 with a destination of the ES attachment circuit

Ethernet Segment ESI-1

Leaf-5 installs 10.0.0.1/32 with two aliasing paths associated to ESI-1, and does per flow load balancing for the traffic destined to 10.0.0.1 (Primary/Backup is also possible)

This model supports IP Aliasing/Backup and fast convergence functions for Host routes and Prefix routes
Inter-Subnet Forwarding for prefix routes (RT5)
Aliasing for EVPN L3 Interface-less model

IP Aliasing allows ECMP for multi-rack VNFs that do not support BGP sessions to all the leaf nodes.

The Ethernet Segment is used exclusively as an L3 construct, associated to the reachability of the VNF loopback.
Conclusions and next steps

The authors ask for WG adoption
Thank you