draft-sajassi-bess-evpn-mvpn-seamless-interop-00.txt

A. Sajassi (Cisco), S. Thoria (Cisco), N. Fazlollahi (Cisco), A. Gupta (Avi Networks)

IETF 99, July 2017
Prague
Main Reasons for Seamless Interop

- Lower cost by not needing GW devices
- Optimum forwarding within a CO among EVPN and MVPN PEs
- Less provisioning
Requirements

- Optimum forwarding
- Optimum replications
- Support for all-active and single-active multi-homing
- Inter-AS support
- Support for all EVPN service interfaces
- Distribute anycast gateway (host gateway)
- Selective & aggregate selective tunnels
- Optimization of host (*,G) and (S,G) state storage
EVPN PE Model
MVPN Network
All-Active Multi-Homing & Split-Horizon Filtering
All-Active Multi-Homing & Split-Horizon Filtering

- Existing SH filtering per RFC 7432 can NOT be used - i.e., MAC-VRF context is lost and SH filtering is valid only in context of a MAC-VRF/BD

- What to do?
  - Use Local-bias mechanism of evpn-overlay with following changes:
  - Adapt it for MPLS overlay (instead of VxLAN)
  - Apply the local bias to all BDs of an IP-VRF
All-Active Multi-Homing & Split-Horizon Filtering using Local Bias
Intra-DC Solution

1. EVPN-IRB PEs modeled as MVPN PEs using IP-VRF facing the core

2. Customer MAC-VRFs connected to the IP-VRF using IRB interface, modeled as MVPN ACs.

3. One-to-one or many-to-one mapping between BDs and MAC-VRFs

4. Incoming traffic on ingress leaf is routed/bridged conventionally for local receivers

5. Incoming traffic on ingress leaf for all kinds of remote receivers (L2/L3) is delivered to the IP-VRF via the IRB interface

6. Within the fabric, both L2 and L3 traffic to remote leaves is encapsulated with the (I-PMSI or S-PMSI) tunnel encap associated with the IP-VRF

7. Each egress leaf will then locally replicate traffic from IP-VRF to its local MAC-VRFs attached via IRB-interfaces which have interested receivers

8. Tenant multicast signaling terminated at IP-VRF (IGMP reports/IGMP Queries/Mcast Data packets received on server facing interfaces are not sent to the core)

9. Receiver interest is carried using BGP MVPN control plane
Solution – Cont.

8. Selective mdt is supported: ingress leaf will originate a \((Cs, Cg) - (Ds, Dg)\) mapping for the given VRF where:

- \((Cs, Cg)\) represents overlay stream
- \((Ds, Dg)\) represents underlay source, group

9. Only leaf nodes which have interest in a given \((Cs, Cg)\) will join the respective \((Ds, Dg)\)
Inclusive PMSI (I-PMSI)

EVPN cloud

PE1
IP-VRF1
MAC-VRF1
MAC-VRF2

PE2
IP-VRF1
MAC-VRF1

PE3
IP-VRF1
MAC-VRF2
MAC-VRF3

S
R1
R2
R3
Selective PMSI (S-PMSI)
Data Center Interconnect (DCI) solution
Data Center Interconnect (DCI) solution
Source Active Discovery in Data Center Interconnect (DCI) solution

- Define a new optional non-transitive attribute and carry with SA-AD when translating received SA-AD on GW
- Attribute discarded by legacy MVPN PEs
- GW processes the attribute on received SA-AD and discards the SA-AD