LSoE Based PE-CE Control Plane for EVPN
draft-malhotra-bess-evpn-lsoe-00

N. Malhotra, K. Patel,
Arrcus

J. Rabadan
Nokia

IETF 104, March 2019
Prague
Motivation – EVPN-IRB

No PE-CE Control Plane

- Data-plane source MAC gleaning for CE MAC learning.
- ARP snooping for CE IPv4-MAC learning.
- ND snooping for CE IPv6-MAC learning.
- Local policy configuration for CE Prefix learning.

Optional PE-CE Control Plane

- PE-CE learning decoupled from non-deterministic periodicity of CE data traffic, in extreme case, from a silent CE host.
- PE-CE learning decoupled from non-deterministic CE behavior of unsolicited ARP and NAs on CE boot-up and CE moves.
- PE-CE learning decoupled from non-deterministic hashing of data, ARP, and ND packets over all-active multi-homed PE-CE LAG interfaces.
- PE-CE learning decoupled from data packet driven ARP/ND gleaning latency on a PE.
- PE-CE prefix learning when CE does not run a dynamic routing protocol.
PE-CE Control Plane for EVPN - Overview

• Based on Link State over Ethernet (LSoE) – link layer auto-discovery protocol
• Procedures and TLVs for MAC learning via PE-CE control plane
• Procedures and TLVs for IPv4-MAC learning via PE-CE control plane
• Procedures and TLVs for IPv6-MAC learning via PE-CE control plane
• Procedures and TLVs for IPv4 and IPv6 Prefix learning via PE-CE control plane
• Procedures for learning with EVPN all-active multi-homing LAG
• Use Cases
PE-CE Control Plane for EVPN – Single Homed CE

- PE-CE LSoE session on PE AC interface via HELLO, OPEN
- Overlay Host TLVs for MAC and IP-MAC learning
- Overlay Prefix TLVs for prefix learning
- Remote Overlay Host TLVs for ARP / ND suppression

- PE – CE Session
- MAC, MAC-IP TLVs
- Prefix TLVs
PE-CE Control Plane – EVPN All-active Multi-Homing

- Separate PE-CE LSoE session to each PE
- Learning TLVs exchanged over each session
- PE-CE failure handling
Use Cases – Simplified All-active Multi-Homing

- Automatic MAC, ARP, ND table sync
- Only handle solicited ARP/ND response
- MAC ECMP without aliasing
- IP ECMP without RT-2 re-origination
Use Cases – Simplified Workload Mobility

- Move detection decoupled from CE data traffic
- Move detection decoupled from CE unsolicited ARP / NA
- No black-holing, flooding to silent hosts
- No probing on MAC, MAC-IP local/remote mismatch
- No probing on MAC age-outs
- Deterministic convergence on host move
Current Workload Mobility

1. Data Frame
2. Punt/MAC Event to CPU
3. Probe
4. Reply
5. RT-2
6. Probe
7. Probe Failure
5. MAC-IP Sync

PE-1 — BGP-EVPN — PE-x — PE-y

Workload Move
BD stretch

CE1

Current Workload Mobility
Use Cases – Overlay ARP/ND Learning

- Pre-emptive remote CE to CE ARP/ND learning to eliminate all flooding
- No ARP/ND gleaning latency on PE for inter-subnet routing
Use Cases – Prefix Learning

- Prefix learning when CE is not running dynamic routing
- Avoids local policy configuration on PE for advertising prefixes behind a CE
PE-CE Control Plane for EVPN - Summary

- Deterministic convergence on boot-up, moves, failures
- Decoupled from non-deterministic CE behavior – data flows, unsolicited ARP/NA
- Decoupled from non-deterministic hashing of data frames on all-active MH LAG
- Prefix learning in absence of dynamic routing on CE
- Extensible for additional use cases
Thank You

Neeraj Malhotra, Keyur Patel (Arrcus)
Jorge Rabadan (Nokia)