Weighted Multi-Path Procedures for EVPN All-Active Multi-Homing

draft-ietf-bess-evpn-unequal-lb-03

Neeraj Malhotra (Cisco)

Ali Sajassi (Cisco)

Jorge Rabadan (Nokia)

John Drake (Juniper)

Samir Thoria (Cisco)

Avinash Lingala (AT&T)

IETF 106, Nov 2019

Singapore

RECAP

Optimally handle scenarios with unequal PE-CE link bandwidth distribution within a multi-homed Ethernet Segment:

- Load-balance overlay unicast flows “unequally” in proportion to each PE’s link bandwidth share in a LAG
- Load-share DF role “unequally” in proportion to each PE’s link bandwidth share in a LAG

Both overlay unicast and BUM flows load-balanced in proportion to PE-CE link bandwidth share in EVPN all-active multi-homed LAG.

Recent Updates

- Section 4.4 – added handling for new DF election algorithm proposed in Weighted HRW draft (draft-mohanty-bess-weighted-hrw-01) - BW weighted score computation for each PE minimizes reassignment.

Status

- Ready for WGLC
- AI from last meeting to to check implementation status

BACKUP

Solution Summary

Unicast Traffic Load-Balancing

- Local PE
  - Advertises per-ESI link-band-width attribute as part of per-ESI EAD RT-1
- Remote PE
  - ESI Path-list is computed in proportion to received link-band-width attribute from each PE

DF Election

- New “BW” capability bit (28) in DF Election Extended-Community indicates desire to augment specified DF election algorithm to be “BW aware” as specified in section 4 of this draft
- Local PE
  - Advertises additional per-ES link-band-width attribute with per-ES RT-4
- Remote PE
  - Type 0 (service carving): Candidate PE list computed in proportion to bandwidth share
  - Type 1 and 4 (HRW): Candidate hash computations for each PE in proportion to it’s bandwidth share
  - Weighted HRW (Type TBD): BW weighted score computation for each PE
  - Type 2 (Preference): additional link-band-width tie-breaker based on PE’s bandwidth share