PFM-SD extension for EVPN multihoming

draft-mankamana-pim-pfm-sd-extension-evpn-mh

Mankamana Mishra, IJsbrand Wijnands, Ryan, Hooman Bidgoli, Zhaohui (Jeffrey) Zhang

Nov 7th 2023, IETF 118
Problem Statement

ECMP path to source 10.1.1.3, pick one of the Next hop to send PIM join.

Source 10.1.1.3 lookup to send PIM join.

From CE source 10.1.1.3 is active on Leaf-4, but join came to Leaf-3.

SRC lookup 10.1.1.3 prefix: 10.1.1.0/24 NH Leaf-3 Leaf-4.

BVI prefix: 10.1.1.0/24.
How we handle source with EVPN multi-homing

Traffic need to be sent to peer with over EVPN tunnel
Deployment challenge with current approach

• Spine to leaf traffic is flowing two times using same link. It requires to provision extra bandwidth

• Need to provision device which can process double the traffic, sending over EVPN tunnels and routing at same time
PFM-SD Quick recap

- RFC 8364 defines flood mechanism in PIM network for source discovery
- As soon as new source is learnt, PFM-SD protocol extension takes care of flooding the source and group information in PIM domain
Optimal solution, what is needed?

Information needed about where (10.1.1.3, 232.1.1.1) is active?

SRC lookup 10.1.1.3 prefix: 10.1.1.0/24
NH Leaf-3 Leaf-4

BVI prefix: 10.1.1.0/24

Source 10.1.1.3 lookup to send PIM join

BVI prefix: 10.1.1.0/24

PIM

(10.1.1.3, 232.1.1.1)

Information needed about where (10.1.1.3, 232.1.1.1) is active?

SRC lookup 10.1.1.3 prefix: 10.1.1.0/24 NH Leaf-3 Leaf-4

BVI prefix: 10.1.1.0/24
Achieving Optimal solution First Step

1. Multicast source becomes active
2. PFM-SD floods in whole domain
1. Membership request for \( (S,G) \)
2. PIM join initiated based on information flooded
3. Spine knows exactly which next hop traffic is active

Look at PIM DB to find where the group is active

- Source 10.1.1.3 lookup to send PIM join
- PFM-SD update
- Source : 10.1.1.3
- Group : 232.1.1.1
- NH Leaf-4
- BVI prefix: 10.1.1.0/24
- (10.1.1.3, 232.1.1.1)
Next step optimization

• Using flood mechanism to whole network unnecessary produce huge amount of data across big PIM doing

• We need further extension to make sure flood are not going beyond the fork point where final ECMP decision is to be made.

• This extension would come in next revision.