Partitioned MDT (PMSI)

71st IETF Philadelphia

Mboned WG

IJsbrand Wijnands,Ice@cisco.com

What is a partitioned MDT

- draft-rosen-l3vpn-mvpn-profiles-00, section 3
- A partitioned MDT is a combination between an emulated-LAN and a P2MP network.
- The tree is a MP2MP LSP
- You setup a partitioned MDT per ingress PE!
  - Only setup the tree if you want to receive data…
- The root of the MP2MP is the ingress PE.

Partitioned MDT setup

- The root of the MP2MP is the ingress PE.
  - Automatically selected based on the next-hop for a customer multicast source or RP.
- Setup of a partitioned MDT is based on receiving joins from customers.
- MDT is setup only if there is data to forward.

Partitioned MDT setup

1. Customer PIM join for (S,G) or (*,G).

2. On the egress PE, Lookup S or RP in RT and find next-hop of the ingress PE.

3. If not already exists, join MP2MP LSP with ingress PE as root.
   - Note, FEC opaque value can either be a BGP signaled ID or statically configured VPN ID.

4. Once tree is setup, send PIM join over the Tunnel to the ingress PE.

5. Ingress PE will receive the join and forward data on the MDT.

MP2MP LSP

- Provides information about the direction a packet is traveling (up/down).

- The upstream path installation on the egress PE is an indication the root is active.

- Saves on state and labels compared to a full mesh of P2MP LSPs.

- Simple procedures to setup compared to PIM bidir because no DF election is necessary.

PIM signaling

- PIM can run on the partitioned MDT without modifications because it's a bidirectional LSP.
- Supports PIM-SM, PIM-SSM and PIM-Bidir.

PIM signaling (cont)

- No asserts will be triggered, single forwarder on the tree.
  - Allows anycast sources to be supported.
  - Downstream PE's RPF against the ingress PE MDT.

- No PIM DF election necessary
  - The root is the DF.
  - A MP2MP LSP gives the ability to know which packets are traveling upstream and which downstream. No need to use a second upstream assigned label to implement DF procedure.

- PIM over Reliable Transport (PORT) possible alternative.

PIM adjacencies

- No need to send PIM Hello's over the MDT.
- The upstream path is an implicit indication the MDT root is active and ready to receive.
- The root of the tree may still send PIM hello's if needed, the leafs don't.
- Can also use BGP auto discovery routes to maintain some sort of PIM adjacency.

RP Mapping (Auto-RP/BSR)

- RP mappings need to be announced ahead of time before PIM can build the tree.

- There are 3 different solutions:
  - Unicast Auto-RP/BSR messages between PE's.
  - Announce Auto-RP/BSR senders in BGP and let interested PE's setup a MDT to it.
  - Build a static default MDT per VPN for Auto-RP/BSR.

Deployment

- If sources/RP's are co-located in few sites, provides optimal re-use of MDT and uses less resources compared to static MDT's.

- First user joining to 'a' source/RP behind a ingress PE may experience initial MDT setup delay.

- Customers with sources in every site are better of with a static MDT (rosen-model).

Deployment (cont)

- PIM LAN procedures are run with a sub-set of the PE's within a VPN.
    - Only the PE's interested in the same ingress PE for the same VPN participate in the PIM procedures.