Multicast VPN fast upstream failover

draft-ietf-bess-mvpn-fast-failover

Thomas Morin
Robert Kebler
Greg Mirsky

IETF-102 July 2018, Montreal

Proposed Solution

- local procedures allowing an egress PE (a PE connected to a receiver site) to take into account the status of P-tunnels to determine the Upstream Multicast Hop (UMH) for a given (C-S, C-G). This method does not provide a "fast failover" solution when used alone, but can be used with the following sections for a "fast failover" solution.
- protocol extensions that can speed up failover by not requiring any multicast VPN routing message exchange at recovery time.
- a "hot leaf standby" mechanism, that uses a combination of these two mechanisms.

Updates

Clarified mpBFD operation in sections:

- BFD Discriminator
  - Upstream PE Procedures
  - Downstream PE Procedures
- Per PE-CE link BFD Discriminator
  - Upstream PE Procedures
  - Downstream PE Procedures

Editorial and nits cleanup

Upstream PE Procedures: Bring It Up

When it is desired to track the P-tunnel status using p2mp BFD session, the Upstream PE:

- MUST initiate BFD session and set bfd.SessionType = MultipointHead as described in [I-D.ietf-bfd-multipoint];
- MUST use [Ed.note] address as destination IP address when transmitting BFD control packets;
- MUST use the IP address of the Upstream PE as source IP address when transmitting BFD control packets;
- MUST include the BGP-BFD Attribute in the x-PMSI A-D Route with BFD Discriminator value set to My Discriminator value.

If tracking of the P-tunnel by using a p2mp BFD session is to be enabled after the P-tunnel has been already signaled, then the procedure described above MUST be followed. Note that x-PMSI A-D Route MUST be re-sent with exactly the same attributes as before and the BGP-BFD Attribute included.

Upstream PE Procedures: Take It Down

If P-tunnel is already signaled, and P-tunnel status tracked using the p2mp BFD session and it is desired to stop tracking P-tunnel status using BFD, then:

- x-PMSI A-D Route MUST be re-sent with exactly the same attributes as before, but the BGP-BFD Attribute MUST be excluded;
- the p2mp BFD session SHOULD be deleted.

Downstream PE Procedures: Bring It Up

On receiving the BGP-BFD Attribute in the x-PMSI A-D Route, the Downstream PE:

- MUST associate the received BFD discriminator value with the P-tunnel originating from the Root PE;
- MUST create p2mp BFD session and set bfd.SessionType = MultipointTail as described in [I-D.ietf-bfd-multipoint];
- MUST use the source IP address of a BFD control packet, the value of BFD Discriminator from the BGP-BFD Attribute to properly demultiplex BFD sessions.

After the state of the p2mp BFD session is up, i.e. bfd.SessionState = Up, the session state will then be used to track the health of the P-tunnel.

Downstream PE Procedures: Take It Down

If the Downstream PE's P-tunnel is already up, its state being monitored by the p2mp BFD session, and the Downstream PE receives the new x-PMSI A-D Route without the BGP-BFD Attribute, the Downstream PE:

- MUST accept the x-PMSI A-D Route; o MUST stop receiving BFD control packets for this p2mp BFD session;
- SHOULD delete the p2mp BFD session associated with the P-tunnel;
- SHOULD NOT switch the traffic to the Standby Upstream PE.

When such a procedure is used, in the context where fast restoration mechanisms are used for the P-tunnels, leaf PEs should be configured to wait before updating the UMH, to let the P-tunnel restoration mechanism happen. A configurable timer MUST be provided for this purpose, and it is recommended to provide a reasonable default value for this timer.

Next steps

- Review and update, if necessary, after the IESG LC of mpBFD closes
- Your comments, suggestions, questions always welcome and greatly appreciated
- WG LC