BGP AS AN MVPN PE-CE Protocol

draft-keyupate-l3vpn-mvpn-pe-ce-00

Keyur Patel, Yakov Rekhter, Eric Rosen

IETF 87, July 2013, Berlin, Germany

Motivation

- RFC6513 and RFC6514 describes procedures and protocol extensions to handle VPN multicast traffic using BGP
- RFC6513 and RFC6514 assumes PIM and mLDP as a PE-CE protocol
  - As the number of PE-CE connections increase, the overhead on PE due to (frequent) periodic PIM state refresh can be an issue
- Natural to extend BGP as a PE-CE protocol for MVPNs

BGP MVPN PE-CE Extensions

- New SAFI known as a C-MCAST SAFI to exchange routes
  - C-MCAST SAFI is the only multicast SAFI used on the PE-CE session
- New Route Types
- Procedures for generating and processing the Source Tree Join Route and Shared Tree Join Routes
- Procedures for pruning sources from the Shared Tree entries

BGP C-MCAST SAFI

- New SAFI to carry customer multicast routes
- New Capability to signal exchange of a new SAFI
- C-MCAST NLRI is defined as:
  - Route Type
  - Length
  - Route Type Specific Information
- Route Types are defined as:
  - Shared Tree Join Routes
  - Source Tree Join Routes
  - Source Prune A-D Routes

C-MCAST Join and Prune Routes

- C-MCAST Join Routes are defined as:
  - Multicast Source Length
  - Multicast Source
  - Multicast Group Length
  - Multicast Group
- AFI Value dictates length check for Join Route Types
  - Specific to IPv4 and IPv6 AFIs only
- Shared Tree Join routes have RP as a Source address
- NLRI format of C-MCAST Source Prune Routes is same as C-MCAST Join Route

Procedures for Originating BGP Source/ Shared Tree Join Routes

- Originated at the CE router by Join/Prunes received from PIM
  - Upstream neighbor selection done using procedures defined in RFC4601
- Originated at the PE router by Join/Prunes received by BGP MVPN SAFI
  - Upstream neighbor selection done using procedures defined in RFC6513
  - Route Type Field is appropriately converted
- Join Routes SHOULD carry IPv4 or IPv6 Address specific RT with the upstream neighbor information to facilitate the filtering of received Joins by the BGP receivers
- Implementations COULD use local filtering methods on the sending side to ensure that the Join messages are directed to the appropriate BGP neighbor avoid attaching IPv4 or IPv6 Address specific RT
- Withdrawals are generated when Source Tree or a Shared Tree entries are deleted

Procedures for Processing Sourced/ Shared Tree Join Routes

- Inbound Filtering MUST be done if the received Join has IPv4 or IPv6 Address specific RT
- Received Join message is used to create/modify or delete the corresponding multicast entry
  - The multicast entry is created when the Join is received for the first time
  - The PE-CE interface is either added or removed from the corresponding multicast entry's OIF list when a withdrawal is received
  - Multicast entry is removed when the OIF list becomes empty

Procedures for Pruning Sources off the Shared Tree

- Source Prune routes SHOULD carry IPv4 or IPv6 Address specific RT with the upstream neighbor information to facilitate the filtering of received Source A-D route by the BGP receivers
- Implementations COULD use local filtering methods on the sending side to ensure that the Source Prune routes are directed to the appropriate BGP neighbor to avoid attaching IP or IPv6 Address specific RT
- Upon received a Source Prune route for a given (S, G) entry a particular Source address is pruned off the corresponding (*, G) state
- Upon receiving a withdrawal for a Source Prune route for a given (S, G) entry a pruned source address is added back to the (*, G) entry

Questions?