

BGP Multicast Signaling Update

draft-ietf-bess-bgp-multicast-controller-07

Jeffrey Zhang, Robert Raszuk, Dante Pacella, Arkadiy Gulko

BESS WG, IETF111

Main Changes

- *Replication State* route type replaces Leaf A-D route type
 - Used to signal replication state from controllers to tree nodes
 - Format is consistent with Leaf A-D route for maximum code reuse
 - Name change is to reduce confusion in relationship to BGP-MVPN
 - No BGP-MVPN knowledge is needed
- Clarifications on BGP-MVPN replacement use case
 - Service controller and tunnel controller could be different
- SR-P2MP signaling
 - New “segment-list” tunnel instead of “P2P Policy” tunnel to reference an existing tunnel (p2p/p2mp) as a replication branch
 - Discussions on relationship with another signaling option
 - draft-hb-idr-sr-p2mp-policy

Two Options for SR-P2MP Signaling

- BGP-MCAST: draft-ietf-bess-bgp-multicast-controller
- SRTE-P2MP: draft-hb-idr-sr-p2mp-policy
 - Based on draft-ietf-idr-segment-routing-te-policy (SRTE-P2P)
 - A new SAFI is still needed, can't reuse "SR-Policy" SAFI (dedicated to unicast)
- Common grounds
 - New NLRI route type in a new SAFI for replication segment
 - With TEA attached for forwarding information

BGP-MCAST

- Covers IP multicast, mLDP, SR-P2MP and label-identified trees
 - Support MP2MP
 - Support SR, yet not SR-specific
 - Support both overlay and underlay signaling
 - Not tied to BGP-MVPN but can replace it
- Tunnel Encapsulation Attribute (TEA) encodes forwarding information

SRTE-P2MP

- Specifically, only for SR-P2MP
 - Do not support MP2MP (yet)
- Also uses a new SAFI
 - Can't reuse "SR-Policy" SAFI from SRTE-unicast
- Separate NLRI route types for policy and replication segments
 - Policy – for roots, including tree identification, leaves, Candidate Paths, etc.
 - No forwarding information
 - Replication segments – forwarding info for roots, leaves and transit nodes
 - SRTE-P2P has only one route for both policy and forwarding information – all encoded in one TEA tunnel
- Two new TEA tunnels for policy and replication segments respectively

Differences wrt SR-P2MP signaling

- No separate route for “SR-P2MP policy” in BGP-MCAST
 - Policy info carried in attribute of route for replication segment on root
- Major difference in TEA
 - SRTE-P2MP is similar to SRTE-P2P:
 - Policy info encoded in TEA
 - All replication/ECMP branches encoded in a single TEA tunnel
 - BGP-MCAST is more aligned with “traditional” TEA
 - Does not carry “policy” info in TEA
 - Replication/ECMP branches are TEA tunnels

TEA Review and Use in SRTE-P2P

- Initial unicast use case
 - An egress router advertises a BGP route to an ingress router, with a TEA encoding a bunch of tunnels
 - The ingress router resolve the route using the TEA tunnels instead of protocol next hop
 - ECMP or Load Balance (LB) out of all listed tunnels
- Use in SRTE-P2P
 - An SRTE P2P Policy candidate path to a destination can ECMP/LB traffic out of several “segment lists”
 - All those segment lists are encoded as sub-TLVs in a single “SR Policy” tunnel in the TEA
 - This is done to encode policy info (each TLV in a TEA is a “tunnel” so policy info has to be part of a tunnel)

SRTE-P2MP Follows P2P Model

SR P2MP Policy SAFI NLRI:

<route-type p2mp-policy>

Attributes:

Tunnel Encaps Attribute (23)

Tunnel Type: (TBD, P2MP-Policy)

Preference

Policy Name

leaf-list (optional)

remote-end point

remote-end point

...

path-instance

active-instance-id

instance-id

instance-id

...

None of the blue info really belong to TEA

replication segment SAFI NLRI:

<route-type: tree replication-segment>

Attributes:

Tunnel Encaps Attribute (23)

Tunnel Type: (TBD Replication-Segment)

replication-sid

downstream-nodes

segment-list

segment

segment

...

segment-list

segment

segment

...

...

TEA for BGP-MCAST

- Each tunnel in the TEA is a replication branch
 - Instead of having all branches encoded in a single tunnel
 - Policy information encoded in a separate attribute
 - No route for policy; just route for replication segment
 - This is more aligned with initial unicast TEA case
- Incoming/Binding SID also encoded as a tunnel
 - Enables/simplifies MP2MP

Tunnel Encap Attribute

Incoming (RPF) 10.1.1.1 Tree Label 100	tunnel1
10.2.1.1 Tree Label 100	tunnel2
10.3.1.1 Tree Label 100	tunnel3
10.4.1.1 Tree Label 200	tunnel4

Possible Ways Forward – Ongoing Discussions

- Common SAFI for both BGP-MCAST and SRTE-P2MP
- For SR-P2MP:
 - Maybe a route to encode policy information as in draft-hb
 - Common route type for replication state on tree nodes
 - Different styles of TEAs to encode forwarding information
 - BGP-MCAST: TEA with tunnels as replication branches
 - Works for IP multicast and mLDP/SR-P2MP tunnels
 - SRTE-P2MP: TEA with a single tunnel that encodes multiple downstream nodes
 - Each downstream node with multiple segment lists to reach it
 - Similar to P2P policy

Considerations

- The SRTE-P2MP style TEA is not applicable to IP multicast or mLDP tunnels
 - BGP-MCAST style is needed anyway and can cover all
- Do we want/need the SRTE-P2MP style
 - OK for implementations that only care about SR-P2MP and already have SRTE-P2P baseline
 - Duplicate effort in standards and implementation as long as controller signaled IP multicast or mLDP tunnel is needed
- Should we fold draft-hb into BGP-MCAST or keep it separate?
 - Just the TEA style is different

Next Steps

- Seek comments
- Add specification details
- BGP-MCAST can and should proceed on its own
 - With considerations for sharing SAFI and route type with SRTE-P2MP
 - With considerations for potential merging from SRTE-P2MP
 - Or keep SRTE-P2MP separate