MVPN/EVPN Controller
for
Provider Tunnel Discovery

draft-zzhang-bess-mvpn-evpn-controller

Jeffrey Zhang (Juniper), Rishabh Parekh (Cisco)

Sandy Zhang (ZTE), Hooman Bidgoli (Nokia)

BESS, IETF112

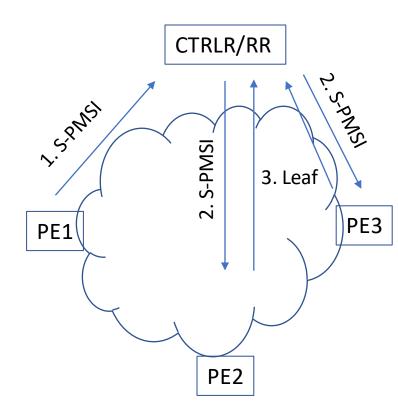# Controller Usage in MVPN/EVPN

- Calculation and Signaling of Provider Tunnels
  - draft-ietf-pim-sr-p2mp-policy
  - draft-ietf-bess-mvpn-evpn-sr-p2mp
    - Ingress PEs collect tree information (root, leaves, constraints) and pass to controller
- Setting up overlay forwarding state on PEs
  - draft-ietf-bess-bgp-multicast-controller
    - The procedures can be used to set up not only provider tunnels but also (C-S,C-G) state in VRFs directly
- *Participating in MVPN/EVPN signaling to collect/signal provider tunnel information at overlay*
  - draft-zzhang-bess-mvpn-evpn-controller
  - What this presentation is about

Background for Tree/Leaf Discovery

- PMSI – Provider Multicast Service Interface
  - A logical interface used by a PE to send C-multicast data to all/some PEs
- Inclusive/Selective-PMSI A-D routes announce tunnel used to instantiate the I/S-PMSI
- Ingress PE advertises I/S-PMSI A-D route and may track leaves explicitly
  - For I-PMSI, tunnel leaves include all those who advertise I-PMSI A-D route
  - For S-PMSI, egress PEs send Leaf A-D routes in response
- Ingress PE triggers tunnel setup
  - Local RSVP-TE module signals the tree based on local or controller calculation, ***or,***
  - Requests a controller to calculate and signal SR-P2MP tree
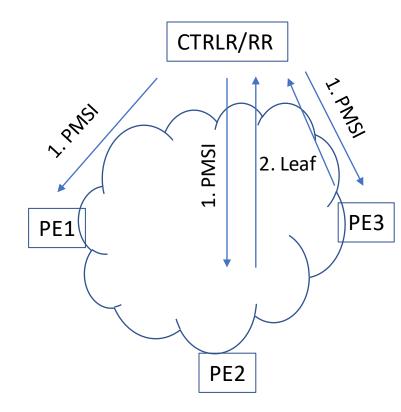   - draft-ietf-bess-mvpn-evpn-sr-p2mp

Tree/Leaf Discovery by Controller w/o Involving Ingress PEs

- The controller participates in MVPN/EVPN signaling to discover tree/leaf information
  - It may be running BGP protocol already, e.g., BGP-LS for topology information
- It receives I/S-PMSI A-D routes to learn tree identification and inclusive tunnel leaves
- It receives Leaf A-D routes to learn selective tunnel leaves
- To direct Leaf A-D routes to the controller, an ingress PE attaches one or more Controller Extended Community in I/S-PMSI A-D routes
  - The Controller EC encodes a controller's address, and tunnel leaves use it to construct a Route Target to direct Leaf A-D routes to the controller
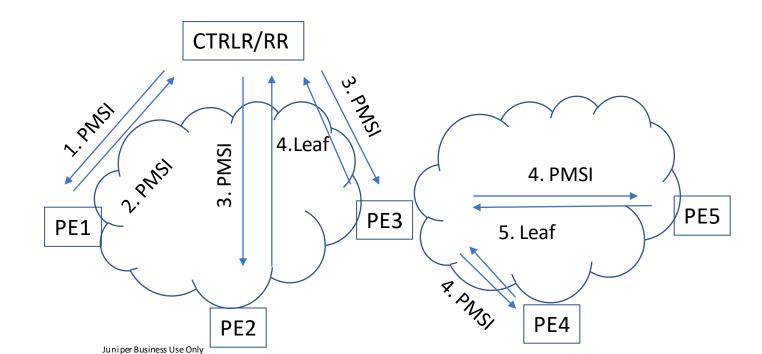
Controller Advertising A-D Routes

- Instead of ingress PEs advertising the I/S-PMSI A-D routes, a controller can advertise on their behalf
  - If the ingress PEs advertises the routes, it would be based on certain configurations, which must be from a central planning entity anyway
  - This facilitates PCE Initiated P2MP trees

- This also makes Domain-wide Common Block (DCB) label allocation easier
  - DCB labels are like "global labels" – e.g., all PEs use the same label value for a VPN (when tunnel aggregation is used – one underlay tunnel for many VPNs)

- Ingress PE would treat an I/S-PMSI A-D route advertised on its behalf as if it advertised route itself and set up forwarding state accordingly

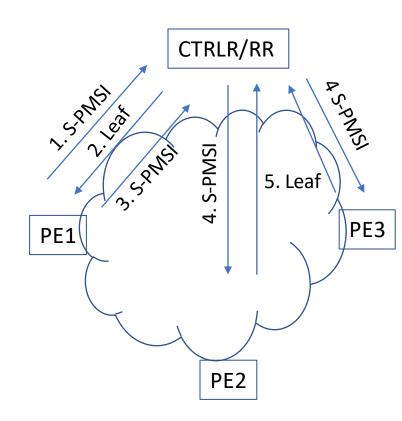- Some special considerations for tunnel segmentation in next slides

CTRLR/RR

1. PMSI

1. PMSI

1. PMSI

2. Leaf

1. PMSI

PE1

PE3

PE2

Tunnel Segmentation Consideration

- With tunnel segmentation, I/S-PMSI A-D routes advertised from a controller on behalf of an ingress PE need to be targeted at the ingress PE, who will then re-advertise
  - Otherwise, the I/S-PMSI A-D routes may not be propagated along desired path of segmentation points
- Although this is a detour, it relieves the ingress PEs from the burden of advertising the routes based on configuration
  - And makes DCB label allocation easy as previously mentioned

Ingress triggered I/S-PMSI A-D Routes

- **There may still be cases for ingress PE triggered I/S-PMSI A-D routes**
  - E.g., selective tunnels only when certain flows' data rate exceeds a threshold
- **Yet we may need a controller to allocate DCB labels for them**
  - Instead of giving each ingress PE its own DCB (that are accepted by all PEs)
- **For that the ingress PE first targets the routes at the controller**
  - Controller allocates the DCB label and send it back in Leaf A-D route
  - Ingress PE then re-advertise the routes normally
- **This extra round-trip (step #1/#2 in diagram) is basically a DCB label allocation request/response**

CTRLR/RR

1. S-PMSI
2. Leaf
3. S-PMSI
4. S-PMSI
4. S-PMSI
5. Leaf
4. S-PMSI

PE1

PE3

PE2

Summary

- Controller participating in MVPN/EVPN signaling to discover/advertise provider tunnel information
  - Controller learning tunnel ID and leaf information
    - PCE delegation w/o relying on ingress PEs
  - Controller advertising I/S-PMSI A-D routes on behalf of PEs
    - PCE initiation
  - Controller allocating DCB labels dynamically for ingress PE's I/S-PMSI routes
    - To be added: ESI labels from DCB
- Comments appreciated!