Multicast Scaling Considerations

draft-zzhang-pim-multicast-scaling-considerations

Jeffrey Zhang, Rishabh Parekh, Hooman Bidgoli, Sandy Zhang

IETF115, London

Scaling Dimensions & Solutions

- Number of receivers
  - IP multicast can scale to unlimited number of receivers
    - Group address is a logical/virtual representation of all receivers
- Number of flows
  - Many flows can be transported by fewer number of tunnels
    - Just like unicast case
  - No per-flow state needed where tunneling is needed
- Size of network
  - E2E multicast in a vast network can be tunneled by different means in different regions

Tunneling Technologies

- IP multicast, mLDP/RSVP-TE/SR P2MP
  - Per-tunnel state needed
- Ingress Replication
  - No per-tunnel state, but in-efficient
- BIER
  - Efficient replication w/o per-tunnel state
- BIER-TE
  - Per-tunnel TE w/o per-tunnel state
  - “Global” BPs (bit positions) encode replication branches
- BIER-RBS – enhanced BIER-TE
  - Local BPs in Recursive Bitstring Structure
  - draft-eckert-bier-cgm2-rbs

Tunnel Overlay Signaling

- Tunnel ingress needs to know which flows to be put onto which tunnel
  - For certain tunnels it also needs to know which nodes must be tunnel egresses
    - E.g., IR, RSVP-TE P2MP, BIER
- Tunnel egresses may need to join the tunnel
  - E.g., mLDP and IP multicast tunnel

mLDP/PIM/IGMP/MLD as Overlay

- Signaling protocol for tunneled multicast is used for overlay signaling
- mLDP signaling for mLDP traffic over other tunnels
  - mLDP over RSVP P2P/P2MP tunnels: RFC 7060
  - mLDP over BIER: draft-ietf-bier-mldp-signaling-over-bier
- PIM signaling for IP multicast traffic over other tunnels
  - PIM over PIM tunnels: Rosen/PIM MVPN
  - PIM over BIER tunnels: draft-ietf-bier-pim-signaling
- IGMP/MLD signaling for IP multicast traffic over BIER
  - ietf-bier-mld

BGP/LISP as Overlay

- Overlay signaling uses different protocol for the tunneled traffic
- BGP MVPN signaling for:
  - IP VPN Multicast
  - mLDP VPN multicast (mLDP being the PE-CE protocol for VPN)
  - Global Table Multicast (IP/mLDP traffic in the global/default routing instance)
- LISP signaling for IP multicast
  - RFC6831, RFC8378

Tunnel Segmentation in BGP MVPN

- PE-PE tunnels are referred to as PMSI tunnels
  - Identified by (overlay) PMSI routes
    - That binds overlay flows to underlay tunnels
  - Instantiated by an underlay tunnel whose type/instance is encoded in PMSI Tunnel Attribute (PTA) of the route
- A PMSI tunnel may have different instantiations in different regions
  - Each with a different type or instance
    - The PTA is updated when the PMSI route is re-advertised into next region
  - For technical/operational/administrative reasons
- Regional Border Routers (RBRs) are segmentation points
  - They need to maintain overlay state (PMSI routes)
  - They stitch upstream/downstream segments based on the overlay PMSI route

Scaling of Segmentation Points and Tunnel Ingress/Egress

- They need to maintain overlay state
  - E.g., IP multicast flow overlay state
- Scale up
  - Multicast forwarding state is not much different from unicast
    - A route points to forwarding instructions
      - ECMP branches for unicast or replication branches for multicast
    - Routes may share forwarding instructions
  - If you can scale up unicast, you can scale up multicast
    - You just need to be prepared to scale up, and to be able to control
    - Signaling protocol needs avoid soft state refreshes
- Scale out segmentation points
  - Use multiple segmentation points between two regions
    - Each responsible for different overlay flows

Summary

- There are existing and in-development solutions to scale multicast in all three dimensions
  - IPv4/IPv6 and MPLS/SRv6 agnostic
  - They can be used together for multi-dimension scaling
- Number of receivers: IP multicast
- Number of flows: tunneling w/ or w/o per-tunnel state
  - PIM/P2MP; IR/BIER/BIER-TE/RBS
- Vast network: tunnel segmentation
  - Scale up/out segmentation points