SR Multicast: Non-source-routed Options

draft-zzhang-mboned-non-source-routed-sr-mcast-00

Jeffrey Zhang, Ice Wijnands, Hooman Bidgoli, Yisong Liu

Mboned WG, IETF120

A Few Clarifications

- An informational draft
  - For discussions and considerations
  - A replacement of draft-zzhang-pim-sr-multicast
  - Rehomed from PIM WG – was presented in PIM WG in IETF119
- Source-routed options not considered in this document
  - Not that they’re not good – they’re just not in the scope of this document
  - BIER is not source-routed – the BitString does not encode the (sub-)tree

SR Principles

- #1 SR Principle – no per-flow/tunnel states in the network
  - Loose/strict path encoded in packets themselves
- #2 SR Principle/Characteristics – optional use of controllers
  - Ingress routers may be instructed by omnipotent/omniscient god-box controllers on how packets should be steered

Multicast Options per #1 Principle

- Ingress Replication (IR)
  - Simple, mature, but inefficient replication
  - Applicable for certain scenarios – e.g., low-rate/fanout sporadic flows
  - Can use both SR or non-SR paths – IP, MPLS or SRv6
- Bit Index Explicit Replication (BIER)
  - Efficient replication w/o per-tree state - perfect solution for SR Multicast
  - Works for all situations
    - Scenarios where convergence under scale is critical benefit the most
- Interestingly, both IR and BIER are developed independently of SR

BIER Status

- Good protocol progress in IETF
  - Base architecture, encapsulation, and signaling RFCs
  - Extensions and brownfield deployment solutions
- Limited deployment due to the chicken-and-egg dilemma
  - New encapsulation/forwarding requires new/programmable ASIC
  - Operators are very interested in BIER, but had to back off due to platform limitations
  - Vendors back off due to unclear demand
- Prime time is coming – pioneering vendors/operators will break the dilemma
  - ASIC capability from several major vendors across edge/access/core platforms
  - Interop testing among major vendors

Multicast Options per #2 Principle

- Controller-calculated trees
  - The trees could still be set up using traditional signaling
- Controller-calculated and controller-signaled trees
  - SR-P2MP: draft-ietf-pim-sr-p2mp-policy
  - BGP-signaled mLDP/IP multicast: draft-ietf-bess-bgp-multicast-controller
    - The only relevance to mLDP is the use of mLDP FEC as a tree identifier
      - This is a good option for transitioning from traditional mLDP-based MVPN

SR-P2MP Considerations

- With the MPLS data plane, SR-P2MP is similar to mLDP/RSVP P2MP
  - Tree nodes have identical per-tunnel forwarding state: label in label out
  - The difference is in the control plane: different tree identifiers and signaling
- SR-P2MP avoids tree-state on non-replication nodes
  - This is good for sparse replication situations
  - PIM/mLDP supports that as well
    - PIM adjacency over tunnels or mLDP signaling over targeted sessions
      - In the case of SR-P2MP, this is avoided by the use of controller

SRv6-P2MP

- With the SRv6 data plane, the tree node state is like the MPLS case
- The incoming/outgoing tree-identifying label corresponds to the FUNCT bits in an SRv6 DST address
- The Locator in an SRv6 DST address gets a packet to a downstream tree node
  - Whether it is directly connected or not; this is like the “base” label
  - SRH can be used to explicitly steer the packet to a remote downstream node
  - Each tree node needs to update the DST address
- New/programmable ASIC is needed to support SRv6-P2MP
  - Though those supporting SRv6 unicast may already support SRv6-P2MP
    - Older SR-incapable ASIC is out of luck – IPv6 multicast can be used in this case
  - Since IPv6 Multicast and SRv6-P2MP all use an IPv6 header, IPv6 Multicast is a very good alternative to SRv6-p2mp
    - Especially when you use controller-signaled IP multicast

Traditional Solutions

- Traditional IR/PIM/P2MP solutions can still be used
  - If they address the use cases w/o pain points
  - Plain old IPv6 multicast may be a better choice than SRv6 in many situations
- While SR removes the need for LDP/RSVP for unicast label distribution, they can still be used for multicast purposes only
  - Unless it is important to use new signaling from controllers
    - To completely remove LDP/RSVP
  - The key is the operator's preference or choice of taste

Summary: Order of Considerations

1. BIER
   - if efficient replication w/o per-tree state is important, and,
   - most replication nodes support BIER
2. Traditional Multicast solutions (IR/PIM/P2MP)
   - If they address the needs w/o pain points, especially if:
     - No need/desire to use controllers, and,
     - Willing to run PIM/mLDP/RSVP for multicast purposes
       - While not using LDP/RSVP for unicast
3. Controller-signaled Multicast
   - SR-P2MP – just keep in mind that they still have per-tree state on tree nodes
   - BGP-signaled mLDP or IP Multicast
     - PIM/controller-signaled IPv6 Multicast is a very good choice for non-MPLS IPv6 networks

Next Steps

- Discussions, please!
- WG adoption?