

PIM
Internet-Draft
Intended status: Informational
Expires: February 16, 2019

Z. Zhang
Juniper Networks
IJ. Wijnands
Cisco Systems
A. Dolganow
Nokia
L. Geng
China Mobile
August 15, 2018

Multicast in SR Networks
draft-zzhang-pim-sr-multicast-00

Abstract

With Segment Routing (SR) architecture [[rfc8402](#)], a unicast flow can be routed through an SR network following an explicit path specified in the packet, w/o the need for per-flow state in the network. As a result, the otherwise needed protocols to signal the per-flow unicast state can also be removed from the network. This document discusses options for multicast in an SR network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 16, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of

Internet-Draft

SR-multicast

August 2018

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Traditional Multicast Technologies	2
3.	Bit Index Explicit Replication	3
4.	Multicast Trees Set up by Other Means	4
5.	SR Specific Solutions	4
6.	Summary	5
7.	Security Considerations	5
8.	Acknowledgements	5
9.	Informative References	5
	Authors' Addresses	7

[1.](#) Introduction

With Segment Routing (SR) architecture [[RFC8402](#)], a unicast flow can be routed through an SR network following an explicit path specified in the packet, w/o the need for per-flow state in the network. As a result, the otherwise needed protocols to signal the per-flow unicast state can also be removed from the network.

This document summarizes options for multicast in an SR network, including traditional multicast technologies, BIER, and various SR specific proposals. The pros and cons of each solution are listed for considerations by operators and vendors.

[2.](#) Traditional Multicast Technologies

Traditional multicast technologies include PIM [[RFC7761](#)], RSVP-TE P2MP [[RFC4875](#)], and mLDP [[RFC6388](#)]. They all require per-tree state on nodes on the tree, and the corresponding protocols to signal and maintain the state. An incoming packet's IP header or MPLS label is looked up, and the packet is forwarded according to the matched state.

While SR allows simplification of state, protocols and centralized

SDN-control, the traditional methods of delivering multicast traffic run contrary to those SR goals.

An alternative is Ingress Replication (IR) - an upstream node of a multicast packet tunnels individual copies to some downstream nodes

across some intermediate nodes. In an SR network, the unicast tunnels used for IR could be traditional IP/MPLS tunnels or could be SR tunnels.

While with IR there is no per-tree state on those intermediate nodes, multiple copies of the same packet may be sent over the same link. If efficient replication is required, PIM/mLDP/RSVP-TE P2MP can still be used for multicast even in an SR network. Deploying SR in the network does not mandate changing the solution that is in place for Multicast.

While mLDP uses the same LDP Session (and packet encoding) as is used by unicast and there is code sharing between LDP and mLDP with regards to neighbor discovery, session setup and Label allocation management, the protocol logic for setting up mLDP tunnel is completely different. It is understood that there is a desire to remove LDP from the network when SR is deployed, but there is really no problem to continue to run mLDP for Multicast. [RFC7473](#) documents a solution where LDP can be run in mLDP-Only mode by simply not exchanging any Unicast bindings.

Similarly, RSVP-TE P2MP tunnel setup is orthogonal from P2P tunnel setup as well. It can be used for multicast even in an SR network if bandwidth reservation and/or per-tunnel explicit path are required.

[3.](#) Bit Index Explicit Replication

Bit Index Explicit Replication (BIER) [[RFC8279](#)] is a new multicast technology that achieves efficient replication as with PIM tree or mLDP/RSVP-TE P2MP tunnels but without requiring per-tree/tunnel state in the transit nodes as with IR.

BIER is a perfect fit for an SR network. A BIER domain can be used as provider/underlay tunnels for MVPN/EVPN-BUM (just as traditional PIM tree or mLDP/RSVP P2MP tunnels), or can be part of end-to-end PIM trees [[I-D.ietf-bier-pim-signaling](#)] (similar to mLDP inband

signaling [[RFC6826](#)]).

However, BIER uses a new forwarding plane that cannot be implemented on many deployed routers. On the other hand, the entry point barrier is decreasing and BIER is becoming more realistic with the following developments:

- o Several major router vendors have wide deployment of platforms capable of BIER (with a software upgrade)
- o Several ASIC vendors have BIER support on their latest/upcoming ASIC offering

- o There have been very good discussions in BIER WG for methods of brown field deployments ([[I-D.przygienda-bier-migration-options](#)] [[I-D.zzhang-bier-php](#)])
- o More and more customers now have concrete plans and requirements for BIER deployment in their networks

[4.](#) Multicast Trees Set up by Other Means

Some operators may accept that they need to maintain per-tree state in their SR network for efficient multicast, and their applications do not require too much per-tree multicast state. However, since unicast no longer needs LDP or RSVP-TE protocol in an SR network, they really want to remove PIM/LDP/RSVP-TE protocol from their networks, and make use of controllers. In that case, the multicast trees could be set up statically from management plane (e.g., netconf), or dynamically via BGP [[I-D.zzhang-bess-bgp-multicast](#)]. [[I-D.zzhang-bess-bgp-multicast-controller](#)].

[5.](#) SR Specific Solutions

There are two multicast solutions that are specifically tied to SR technologies - "TreeSID" and "Spray" [[I-D.voyer-spring-sr-p2mp-policy](#)]. There is also a proposal [[I-D.allan-pim-sr-mpls-multicast-framework](#)] that uses calculated multicast trees based on flooded membership information.

1. TreeSID essentially refers to P2MP tunnels setup not by mLDP or RSVP-TE. It could be statically configured, or instantiated from

a controller.

2. Spray is Ingress Replication. In particular, in case of SRv6, instead of regular IP/MPLS tunneling, an SRv6 header encodes the "tunnel end" and the original destination multicast address. When the packet arrives at the tunnel end, the original multicast address is restored from the SRv6 header.
3. In [[I-D.allan-pim-sr-mpls-multicast-framework](#)], multicast membership information is flooded, and each node will calculate if it plays a role (as a root, leaf, or a branching point) for each multicast tree. If yes, corresponding forwarding state is installed to forward incoming multicast traffic accordingly.

The above three methods are really no different from existing technologies. 1 and 3 do not use existing tree signaling protocol but still require per-tree forwarding state on roots, leaves and branching points (for comparison, existing technologies does have per-tree state on all nodes on a tree - including those non-branching

points). 3 is similar to MOSPF [[RFC 1584](#)] and requires flooded receiver information throughout the domain. 2 is just IR with a different encapsulation (e.g. SRv6).

[6.](#) Summary

There is no one-for-all option when it comes to multicast in an SR network. Depending on specific deployment scenarios and requirements, the following could be considered in order:

- o If most of nodes can support it, BIER is the best choice, because it removes state from the network and simplifies the control-plane (like SR).
- o If efficient multicast replication is required, then run mLDP/RSVP-TE/PIM for multicast purpose if that is acceptable.
- o If there is a strong desire to create multicast forwarding state without relying on mLDP and/or RSVP-TE, one can use static configuration or controller signaling (e.g. SR specific TreeSID, generic BGP based signaling or netconf based provisioning).

- o Use Ingress Replication with SR unicast tunnels. This includes Spray Multicast if SRv6 is desired.

In case of MPLS based P2MP, a Binding SID could be optionally used to direct traffic to the root of the tunnel.

Notice that there are already many protocols defined (PIM, mLDP, RSVP-TE, BGP, MOSPF) that can be used to create multicast forwarding state. When designing a new protocol to do the same, we need to be sure the benefits are significant enough and out-weight the cost of developing it.

[7.](#) Security Considerations

This document does not seem to introduce new security risks.

[8.](#) Acknowledgements

The authors thank Eric Rosen, Tony Przygienda, and John Drake for their reviews, comments and suggestions.

[9.](#) Informative References

Zhang, et al. Expires February 16, 2019 [Page 5]

Internet-Draft SR-multicast August 2018

[I-D.allan-pim-sr-mpls-multicast-framework]

Allan, D., Tantsura, J., and I. Duncan, "A Framework for Computed Multicast Applied to SR-MPLS", [draft-allan-pim-sr-mpls-multicast-framework-00](#) (work in progress), June 2018.

[I-D.ietf-bier-pim-signaling]

Bidgoli, H., Dolganow, A., Kotalwar, J., Xu, F., mishra, m., and Z. Zhang, "PIM Signaling Through BIER Core", [draft-ietf-bier-pim-signaling-03](#) (work in progress), June 2018.

[I-D.przygienda-bier-migration-options]

Przygienda, T. and Z. Zhang, "BIER Migration Frameworks", [draft-przygienda-bier-migration-options-00](#) (work in

progress), June 2018.

- [I-D.voyer-spring-sr-p2mp-policy]
daniel.voyer@bell.ca, d., Hassen, C., Gillis, K.,
Filsfils, C., and A. Venkateswaran, "SR Replication Policy
for P2MP Service Delivery", [draft-voyer-spring-sr-p2mp-policy-00](#) (work in progress), June 2018.
- [I-D.zzhang-bess-bgp-multicast]
Zhang, Z., Patel, K., Wijnands, I., and a.
arkadiy.gulko@thomsonreuters.com, "BGP Based Multicast",
[draft-zzhang-bess-bgp-multicast-01](#) (work in progress),
March 2017.
- [I-D.zzhang-bess-bgp-multicast-controller]
Zhang, Z., Raszuk, R., Pacella, D., and A. Gulko,
"Controller Based BGP Multicast Signaling", [draft-zzhang-bess-bgp-multicast-controller-00](#) (work in progress),
September 2017.
- [I-D.zzhang-bier-php]
Zhang, Z., "BIER Penultimate Hop Popping", [draft-zzhang-bier-php-01](#) (work in progress), August 2018.
- [RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S.
Yasukawa, Ed., "Extensions to Resource Reservation
Protocol – Traffic Engineering (RSVP-TE) for Point-to-
Multipoint TE Label Switched Paths (LSPs)", [RFC 4875](#),
DOI 10.17487/RFC4875, May 2007,
<<https://www.rfc-editor.org/info/rfc4875>>.

- [RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B.
Thomas, "Label Distribution Protocol Extensions for Point-
to-Multipoint and Multipoint-to-Multipoint Label Switched
Paths", [RFC 6388](#), DOI 10.17487/RFC6388, November 2011,
<<https://www.rfc-editor.org/info/rfc6388>>.
- [RFC6826] Wijnands, IJ., Ed., Eckert, T., Leymann, N., and M.
Napierala, "Multipoint LDP In-Band Signaling for Point-to-

Multipoint and Multipoint-to-Multipoint Label Switched Paths", [RFC 6826](#), DOI 10.17487/RFC6826, January 2013, <<https://www.rfc-editor.org/info/rfc6826>>.

- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, [RFC 7761](#), DOI 10.17487/RFC7761, March 2016, <<https://www.rfc-editor.org/info/rfc7761>>.
- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", [RFC 8279](#), DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/info/rfc8279>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [RFC 8402](#), DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

IJsbrand Wijnands
Cisco Systems

EMail: ice@cisco.com

Andrew Dolganow
Nokia

EMail: andrew.dolganow@nokia.com

China Mobile

EMail: gengliang@chinamobile.com