Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2019

D. Voyer, Ed.
C. Hassen
K. Gillis
Bell Canada
C. Filsfils
R. Parekh
Cisco Systems, Inc.
H. Bidgoli
Nokia

October 22, 2018

SR Replication Policy for P2MP Service Delivery
draft-voyer-spring-sr-p2mp-policy-01

Abstract

This document describes the SR policy architecture for P2MP service delivery.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document defines a variant of the SR Policy [I-D. ietf-spring-segment-routing-policy] for constructing a P2MP segment to support Point-to-Multipoint service delivery. We call it an SR Replication Policy.

A Point-to-Multipoint (P2MP) segment connects a Root node to a set of Leaf nodes in a Segment Routing Domain. We define two types of a P2MP segment: Spray and TreeSID.

Spray P2MP segment enables a Root node to directly replicate a packet using a SR path to each Leaf node.

For a TreeSID P2MP segment, a controller computes a tree from a Root node to a set of Leaf nodes via a set of Replication nodes. A packet is replicated at the Root node and on Replication nodes towards each Leaf node.

2. SR Replication Policy

The SR Replication policy is a variant of an SR policy [I-D.ietf-spring-segment-routing-policy]. This section is similar to section 2 of SR Policy draft [I-D.ietf-spring-segment-routing-policy], and applies equally to the Spray and TreeSID P2MP segments unless explicitly specified. A SR replication policy can be provisioned either locally or setup via controller.

A SR replication Policy is defined by following elements:

- o Root node: This is the headend of the P2MP segment.
- o Leaf nodes: A set of nodes that terminate the P2MP segment.
- o Constraints: Optional set of topological constraints to be satisfied by the P2MP segment.

A SR Replication Policy is identified through the tuple <Root node, color>.

Like any SR policy, a SR Replication Policy has a BSID [I-D.ietf-spring-segment-routing-policy] instantiated into the forwarding plane. For P2MP segments, the BSID is applicable only at the Root node.

For a TreeSID P2MP segment, the SR Replication policy also has an associated identifier, a TreeSID. The TreeSID is instantiated into the forwarding plane at Replication nodes and Leaf nodes of a P2MP segment. A packet is steered towards the set of Leaf nodes when the active SID of the packet is a TreeSID.

A SR Replication may comprise of multiple candidate paths. A candidate path is valid when all its SID-Lists are valid. The active candidate path is selected based on the tie breaking rules amongst the valid candidate-paths.

In the context of a SR Replication Policy, the selected path MAY have more than one SID-List. The weights of the SID-Lists is not applicable for a SR Replication Policy. They MUST be set to 1.

Any traffic steered into a SR Replication Policy is replicated along the SID-Lists of its selected path. Each SID-List takes a packet to either a Replication node or a Leaf node of a P2MP segment.

3. Steering

Traffic is steered into a SR Replication Policy in two ways

- o Based on a local policy-based routing at the Root node.

- o Based on remote classification and steering via the BSID of the SR Replication Policy at the Root node.

4. Spray P2MP segment

In a Spray P2MP segment, packet replication occurs only at the Root node. A SR Replication policy for a Spray P2MP segment is instantiated only at the Root node. There are no Replication nodes in these segments.

A packet, using this approach, is replicated directly to each Leaf node via a SR path from the Root to a given Leaf node.

5. TreeSID P2MP segment

In a TreeSID P2MP segment, packet replication occurs at the Root node and on Replication nodes towards the Leaf node.

A SR Replication policy instantiated on the Root node takes a packet from the Root node to Replication nodes towards the Leaf node. A Replication node MAY also be a Leaf node. The SR Replication policy instantiated at the Replication nodes take the packet down further to other Replication nodes or Leaf nodes.

5.1. Using Controller to build a P2MP Segment

```
                         North Bound                South Bound
                         Programming          ..... Programming
                         Interface                  Interface
                              |
                              |
                              v
                           +-----+ ...........................
              .............| PCE | .............             .
              .            +-----+             .             .
              .               .                .             .
              .               .                .             .
              .               .                .             .
              .               .                V             .
              .               .              +----+          .
              .               .              | N3 |          .
              .               .              +----+          .
              .               .                 | Leaf (L2)   .
              .               .                 |            .
              .               .                 |            .
              V               V                 |            V
            +----+          +-----+ -------------          +----+
            | N1 |----------| N2 |-------------------------| N4 |
            +----+          +-----+                        +----+
           Root (R)         Replication Node (M)           Leaf (L1)
```

Figure 1: Centralized Control Plane Model

5.1.1. SR Replication Policy Creation

A SR Replication policy can be instantiated and maintained in a centralized fashion using a Path Computation Element (PCE). This section outlines a high-level architecture for such an approach.

5.1.1.1. API

North-bound APIs on a PCE can be used to:

1. Create P2MP SR policy

2. Delete P2MP SR policy

3. Update P2MP SR policy

5.1.1.2. Invoking API

Operator shall interact with a PCE via REST, Netconf, gRPC, CLI. Yang model shall be be developed for this purpose as well.

5.1.2. TreeSID P2MP Segment Computation

Network operator passes the addresses of the root (R) and set of leaves {L} as well as Traffic Engineering (TE) attributes (e.g., constraints such as link color, optimization criteria such as latency) of the P2MP segment to PCE via a suitable North-Bound API. The PCE computes the tree, and if successful instantiates the P2MP segment on Root, Replication, and Leaf nodes.

Path constraints shall include link color affinity, bandwidth, disjointness (link, node, SRLG), delay bound, link loss, etc. Path shall be optimized based on IGP or TE metric or link latency.

Ideally, same P2MP SID SHOULD be used for forwarding entries at Root, Mid, and Leaf nodes. Different P2MP SIDs MAY be used at different node(s) if it is not feasible to use same P2MP SID. The P2MP SID is derived from SRLB of nodes. SIDs (BSID as well as P2MP SID) can also be assigned by operator.

A PCE can modify a P2MP segment following network element failure or in case a better path can be found based on the new network state. In this case, the PCE may want to setup the new tree and remove the old tree from the network in order to minimize traffic loss. As such, a separate P2MP SID can be used for the new tree.

A PCE shall be capable of computing paths across multiple IGP areas or levels as well as Autonomous Systems (ASs).

5.1.2.1. Topology Discovery

A PCE shall learn network topology, TE attributes of link/node as well as SIDs via dynamic routing protocols (IGP and/or BGP-LS). It may be possible for operators to pass topology information to PCE via north-bound API.

5.1.2.2. Capability and Attribute Discovery

It shall be possible for a node to advertise TreeSID capability via IGP and/or BGP-LS. Similarly, a PCE can also advertise its TreeSID capability via IGP and/or BGP-LS. Capability advertisement allows a network node to dynamically choose one or more PCE(s) to obtain services pertaining to SR Replication policies, as well a PCE to dynamically identify TreeSID capable nodes.

5.1.3. Instantiating TreeSID P2MP segment nodes

Once a PCE computes a tree for P2MP segment, it needs to instantiate the segment on the relevant network nodes. The PCE can use various protocols to program the forwarding entries, and these protocols are described below.

5.1.3.1. PCEP

PCE Protocol (PCEP)has been traditionally used:

1. For a head-end to obtain paths from a PCE.

2. A PCE to instantiate SR policies.

PCEP protocol can be stateful in that a PCE can have a stateful control of an SR policy on a head-end which has delegated the control of the SR policy to the PCE. PCEP shall be extended to provision and maintain forwarding entries in a stateful fashion.

5.1.3.2. BGP

BGP has been extended to instantiate and report SR policies. It shall be used to instantiate and maintain forwarding entries for SR Replication policies.

5.1.3.3. NetConf

TBD

5.1.4. Protection

5.1.4.1. Local Protection

A network link/node on the tree of a P2MP segment can be protected using SR policies computed by PCE. The backup SR policies shall be programmed in forwarding plane in order to minimize traffic loss when the protected link/node fails.

5.1.4.2. Path Protection

It is possible for PCE create a disjoint backup tree for providing end-to-end path protection.

6. Illustration

TBD

7. IANA Considerations

This document makes no request of IANA.

8. Security Considerations

There are no additional security risks introduced by this design.

9. Acknowledgements

The authors would like to acknowledge Siva Sivabalan.

10. Contributors

Arvind Venkateswaran
Cisco Systems, Inc.
San Jose
US

Email: arvvenka@cisco.com

Zafar Ali
Cisco Systems, Inc.
US

Email: zali@cisco.com

11. Normative References

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-02 (work in progress), October 2018.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

Authors' Addresses

Daniel Voyer (editor)
Bell Canada
Montreal
CA

Email: daniel.voyer@bell.ca

Clayton Hassen
Bell Canada
Vancouver
CA

Email: clayton.hassen@bell.ca

Kurtis Gillis
Bell Canada
Halifax
CA

Email: kurtis.gillis@bell.ca

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Rishabh Parekh
Cisco Systems, Inc.
San Jose
US

Email: riparekh@cisco.com

Hooman Bidgoli
Nokia
Ottawa
CA

Email: hooman.bidgoli@nokia.com