SPRING
Internet-Draft
Intended status: Standards Track
Expires: December 22, 2017

C. Filsfils
S. Previdi
P. Psenak
L. Ginsberg
Cisco Systems, Inc.
June 20, 2017

Segment Routing Recursive Information
draft-filsfils-spring-sr-recursing-info-05

Abstract

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF).

There are use cases where it is desirable to utilize a SID associated with a given node in order to transport traffic destined to different local services supported by such node. This document defines the mechanism to do so and illustrates it.

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].

Table of Contents

1. Introduction

Segment Routing (SR) as defined in [I-D.ietf-spring-segment-routing] utilizes forwarding instructions called "segments" to direct packets through the network. When an MPLS dataplane is in use Segment Identifiers (SIDs) are assigned to prefixes and are associated with a specific algorithm. SIDs may be Local or Global in scope. When a SID has global scope the SID is mapped via the Segment Routing Global Block (SRGB) to a node specific label which acts as the forwarding instruction.

There are use cases where it is desirable to utilize a SID associated with a node N to transport traffic destined to different local services supported by N. This document defines the mechanism to do so and illustrates it.

2. Use Cases

In some deployments, multiple loopback addresses are configured in a single router. Each of these loopback addresses can serve as the address of the node. Specific addresses within this set of node addresses may be used as the endpoint for a particular service or capability. If the number of labeled entries installed in the forwarding plane is a concern, the use of a single label for the set of node addresses (or for some subset of the set of node addresses) can be used in order to reduce the number of forwarding entries required to reach any of the node addresses. This, in turn, would require sharing of a SID among multiple prefixes.

Some deployments attach different services to an edge router in a network via unique interfaces. Rather than assigning a unique SID for the address associated with each service the desired behavior is to use a Node-SID to reach the edge router and then utilize a service specific Local SID to direct the packet to the correct service.

The first use case is a sub-case of the second one where the local SID is not present (e.g. encoded as implicit-null). Hence in the remainder of this document, we will focus on the more generic use case, i.e., utilizing a single Node-SID in combination with an optional Local SID to transport traffic up to the node and then have the node apply the correct service based on the local SID (if present).

3. Segment Routing Recursing Information (SRRI)

This document introduces and defines the concept of a "Segment Routing Recursing Information (SRRI) that needs to be carried into IGPs (ISIS and OSPF), e.g., as a TLV (or SubTLV) attached to the prefix advertisement.

The description in this document is protocol agnostic and can be applied to both IGPs (IS-IS and OSPF). The protocol specific format of this advertisement will be defined in protocol specific specifications.

Advertisement of a prefix P with SRRI (R, Alg, SID-L) indicates that a remote node M MUST use a segment list {SID(R), SID-L) to transport traffic to P; where SID(R) is the prefix SID of R for the specified algorithm. The generic advertisement format is then:

IPv4 SRRI

```
        +-----------------------+
        | Flags                 | 1 byte
        +-----------------------+
        | Algorithm             | 1 byte
        +-----------------------+
        | Recursing SID Address | 4 bytes
        +-----------------------+
        | Local SID             | 4 bytes (optional)
        +-----------------------+
```

IPv6 SRRI

```
        +-----------------------+
        | Flags                 | 1 byte
        +-----------------------+
        | Algorithm             | 1 byte
        +-----------------------+
        | Recursing SID Address | 16 bytes
        +-----------------------+
        | Local SID             | 4 bytes (optional)
        +-----------------------+
```

where:

- "Flags" is one byte field of flags. Only one flag is currently defined: the V-flag. If set, the receiver of the SRRI MUST verify that the originator of the prefix the SRRI is attached to and the prefix covering the Recursing SID address are originated by the same node ("same origin node").

- "Algorithm" defines the algorithm related to the prefix reachability as defined in [I-D.ietf-spring-segment-routing].

- "Recursing SID Address" contains an IPv4 or IPv6 address whose Node-SID is to be used in order to reach the prefix with which the SRRI information is associated.

- "Local SID" is the local SID allocated to the prefix the SRRI is attached to.

The SRRI is associated with a prefix reachability advertisement. The manner of this association is protocol specific.

The following apply to the SRRI:

- o The prefix reachability advertisement this SRRI is attached to MUST NOT have a Prefix-SID assigned for the algorithm specified in the SRRI.

- o Multiple "Recursing SID Address" and "Local SID" MAY be associated with the same parent prefix.

4. Illustration

4.1. Reference Diagram

The following reference topology diagram is used:

```
      A----B----D
       \    \  /
        \    \/
         \   /\
          \ /  \
           C----E

    |------|------|
      Area   Area
       0      1
```

A is in Area 0
B and C are ABRs
D and E are in Area 1

D has a loopback 1.0.0.4/32 and Node SID 16004
D has a local service 1.0.0.99/32 with Local SID 30004
E has a loopback 1.0.0.5/32 and Node SID 16005
E has a local service 1.0.0.99/32 with Local SID 30005

For simplicity, we abstracted the algorithm variable in the SRRI and process.

4.2. Description

D advertises prefix 1.0.0.99/32 with SRRI (1.0.0.4, 30004, V=1)

B receives the 1.0.0.99/32 prefix advertisement from D. Since the V-flag is set, B MUST confirm that D also originates the "Recursing SID address" 1.0.0.4/32 (i.e.: "same origin node").

If same origin node is not confirmed, then B does not install any SR RIB entry for prefix 1.0.0.99/32. If same origin node is confirmed, B installs an SR RIB entry for 1.0.0.99/32 which uses the segment list {16004, 30004} and the OIF to 16004.

Furthermore, B leaks the prefix in area 0 and advertises it with the SRRI (1.0.0.4, 30004, V=0). The V-flag unset indicates that area 0 nodes do not need to perform the "same origin node" check.

E advertises prefix 1.0.0.99/32 with the SRRI (1.0.0.5, 30005, V=1)

B does the same for the route from E as he did for the route from D.

Specifically, let us highlight two elements:

o First, B ends up with an ECMP SR-based RIB entry to 1.0.0.99/32:

- {16004, 30004} OIF to 16004
- {16005, 30005} OIF to 16005

o Second, B advertises 1.0.0.99/32 in area 0 with two SRRI's:

- (1.0.0.4, 30004, V=0)
- (1.0.0.5, 30005, V=0)

C does the same for the routes from D and E. Briefly, C advertises 1.0.0.99/32 in area 0 with two SRRI's:

- (1.0.0.4, 30004, V=0)
- (1.0.0.5, 30005, V=0)

A learns 1.0.0.99/32 from B and C and uses normal IGP process to select one, the other or both.

Let us assume A prefers the path via B.

A receives the 1.0.0.99/32 prefix advertisement from B. Since the V-flag is unset, no "same origin node" is verified. A installs an SR RIB entry for 1.0.0.99/32 with an ECMP set of path:

path 1 is {16004, 30004} and the OIF to 16004
path 2 is {16005, 30005} and the OIF to 16005

5. Benefits

The mechanism and SR Recursive Information defined in this document supports:

o single-area and multi-area deployments.

o single-homed and multi-homed prefixes.

o the presence of a Local-SID or not.

o Any combination of the above.

The Segment Routing Recursive Information minimize the number of global prefix SID's.

Finally, the Segment Routing Recursive Information is consistent with the MPLS FIB structure: different FEC's have different entries.

6. IANA Considerations

This document doesn't introduce any new code-point.

7. Security Considerations

TBD

8. Acknowledgements

We would like to thank Les Ginsberg and Ahmed Bashandy for their contributions.

9. Normative References

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-11 (work in progress), February 2017.

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

Authors' Addresses

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Stefano Previdi
Cisco Systems, Inc.
Italy

Email: stefano@previdi.net

Peter Psenak
Cisco Systems, Inc.
Apollo Business Center
Mlynske nivy 43
Bratislava 821 09
Slovakia

Email: ppsenak@cisco.com

Les Ginsberg
Cisco Systems, Inc.
US

Email: ginsberg@cisco.com