Authors: C. Filsfils   P. Camarillo, Ed.
         Cisco Systems, Inc.   Cisco Systems, Inc.
         J. Leddy   D. Voyer   S. Matsushima
         Individual Contributor   Bell Canada   SoftBank
         Z. Li
         Huawei Technologies

SRv6 NET-PGM extension: Insertion

Abstract

   Traffic traversing an SR domain is encapsulated in an outer IPv6
   header for its journey through the SR domain.

   To implement transport services strictly within the SR domain, the
   SR domain may require insertion or deletion of an SRH after the
   outer IPv6 header of the SR domain. Any segment within the SRH is
   strictly contained within the SR domain.

   This document extends SRv6 Network Programming [RFC8986] with new SR
   endpoint and transit behaviors to be performed only within the SR
   domain in any packet owned by the domain.

Requirements Language

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in
   BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
   capitals, as shown here.

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 17 February 2023.

Copyright Notice

Table of Contents

1. **Introduction**

Packets transiting an SR Domain may be steered into an SR Policy for
a variety of reasons. For example, a PLR router reroutes traffic on
a TI-LFA repair path [I-D.ietf-rtgwg-segment-routing-ti-lfa] or when
a Binding-SID is expanded [I-D.ietf-spring-segment-routing-policy].

This document extends the SRv6 Network Programming [RFC8986] model
with new endpoint and transit behaviors enabling the insertion of an
SRH after the outer IPv6 header of the SR domain. The operations

described in this document must take into account the considerations described in [I-D.voyer-6man-extension-header-insertion].

2. SRv6 endpoint behaviors

SRv6 Network Programming Section 4 defines a base set of SRv6 endpoint behaviors. This is extended with the behaviors described in this section.

2.1. End.B6.Insert: Endpoint bound to an SRv6 policy

The "Endpoint bound to an SRv6 Policy" is a variant of the End behavior.

One of its applications is to express scalable traffic-engineering policies across multiple domains. It is the one of the SRv6 instantiations of a Binding SID [RFC8402].

An End.B6.Insert SID is never the last segment in a SID list, and any SID instantiation must be associated with an SR Policy B[I-D.ietf-spring-segment-routing-policy].

When N receives a packet whose IPv6 DA is S and S is a local End.B6.Insert SID, does:

```
S01. When an SRH is processed {
S02.    If (Segments Left == 0) {
S03.       Send an ICMP Parameter Problem message to the Source Address
              Code TBD-SRH (SR Upper-layer Header Error),
              Pointer set to the offset of the upper-layer header,
              interrupt packet processing and discard the packet
S04.    }
S04.    If (IPv6 Hop Limit <= 1) {
S05.        Send an ICMP Time Exceeded message to the Source Address,
              Code 0 (Hop limit exceeded in transit),
              interrupt packet processing and discard the packet
S06.    }
S07.    max_LE = (Hdr Ext Len / 2) - 1
S08.    If ((Last Entry > max_LE) or (Segments Left > (Last Entry+1)){
S09.       Send an ICMP Parameter Problem to the Source Address,
              Code 0 (Erroneous header field encountered),
              Pointer set to the Segments Left field,
              interrupt packet processing and discard the packet
S11.    }
S12.    Decrement Hop Limit by 1
S13.    Insert a new SRH in between the IPv6 Header and the received
          SRH containing the list of segments of B
S14.    Set the IPv6 DA to the first segment of B
S15.    Resubmit the packet to the egress IPv6 FIB lookup and
            transmission to the new destination
S16. }
```

When processing the Upper-layer header of a packet matching a FIB
entry locally instantiated as an SRv6 End.B6.Insert SID, send an
ICMP parameter problem message to the Source Address and discard the
packet. Error code "SR Upper-layer Header Error", Pointer set to the
offset of the upper-layer header.

2.2. End.B6.Insert.Red: [...] with reduced SRH

This is an optimization of the End.B6.Insert behavior.

End.B6.Insert.Red reduces the size of the new SRH by one SID by
avoiding the insertion of the first SID in the pushed SRH. In this
way, the first SID is only written in the DA and the packet is
forwarded according to it.

The new SRH is created as described in Section 4.1.1 of [RFC8754].

3. SR Policy Headend Behaviors

SRv6 Network Programming defines in Section 5 a set of SR Policy
Headend Behaviors. This is extended with the following behaviors
defined in this section.

3.1. H.Insert: SR Headend with insertion of an SRv6 Policy

Node N receives two packets P1=(A, B2) and P2=(A,B2)(B3, B2, B1; SL=1). B2 is neither a local address nor SID of N.

N steers the transit packets P1 and P2 into an SRv6 Policy with one SID list <S1, S2, S3>.

The "H.Insert" transit insertion behavior is defined as follows:

1.   insert the SRH (B2, S3, S2, S1; SL=3)   ;; Ref1, Ref1bis
2.   set the IPv6 DA = S1
3.   forward along the shortest path to S1

Ref1: The received IPv6 DA is placed as last SID of the inserted SRH.

Ref1bis: The SRH is inserted [I-D.voyer-6man-extension-header-insertion] before any other IPv6 Routing Extension Header.

After the H.Insert behavior, P1 and P2 respectively look like:

   1. (A, S1) (B2, S3, S2, S1; SL=3)

   2. (A, S1) (B2, S3, S2, S1; SL=3) (B3, B2, B1; SL=1)

3.2. H.Insert.Red: H.Insert with reduced insertion

The H.Insert.Red behavior is an optimization of the H.Insert behavior. It is defined as follows:

1.   insert the SRH (B2, S3, S2; SL=3)
2.   set the IPv6 DA = S1
3.   forward along the shortest path to S1

H.Insert.Red will reduce the size of the SRH by one segment by avoiding the insertion of the first SID in the pushed SRH. In this way, the first segment is only introduced in the DA and the packet is forwarded according to it.

After the H.Insert.Red behavior, P1 and P2 respectively look like:

   1. (A, S1) (B2, S3, S2; SL=3)

   2. (A, S1) (B2, S3, S2; SL=3) (B3, B2, B1; SL=1)

4. **Maximum H.Insert MSD Type**

   This document defines the MSD (Maximum SID Depth) for H.Insert
   behavior and requests the MSD type assignment from the IGP MSD-Types
   registry created by [RFC8491].

   The Maximum H.Insert MSD Type specifies the maximum number of SIDs
   that can be inserted as part of the "H.insert" behavior:

   1. Max H.insert Type: 43 (Suggested value - to be assigned by
      IANA)

   If the advertised value is zero or no value is advertised then the
   router is assumed not to support any variation of the "H.insert"
   behavior.

5. **IANA Considerations**

5.1. **SRv6 Endpoint Behaviors**

   This document requests IANA to allocate the following codepoints
   within the "SRv6 Endpoint Behaviors" sub-registry under the top-
   level "Segment Routing Parameters" registry.

Value	Hex	Endpoint behavior	Reference
13	0x000D	End.B6.Insert	[This.ID]
26	0x001A	End.B6.Insert.Red	[This.ID]

Table 1: IETF - SRv6 Endpoint Behaviors

5.2. **MSD Types**

   This document requests IANA to allocate the following codepoint
   within the "IGP MSD-Types" sub-registry under the top-level "IGP
   Parameters" registry.

Value	Hex	Endpoint behavior	Reference
43	0x2B	Max H.Insert	[This.ID]

Table 2: IETF - MSD Types

6. **Acknowledgements**

   The authors would like to acknowledge Stefano Previdi, Dave Barach,
   Mark Townsley, Peter Psenak, Thierry Couture, Kris Michielsen, Paul
   Wells, Robert Hanzl, Dan Ye, Gaurav Dawra, Faisal Iqbal, Jaganbabu
   Rajamanickam, David Toscano, Asif Islam, Jianda Liu, Yunpeng Zhang,
   Jiaoming Li, Narendra A.K, Mike Mc Gourty, Bhupendra Yadav, Sherif
   Toulan, Satish Damodaran, John Bettink, Kishore Nandyala Veera Venk,
   Jisu Bhattacharya and Saleem Hafeez.

7. Contributors

Daniel Bernier

Bell Canada

Canada

Email: daniel.bernier@bell.ca

Dirk Steinberg

Lapishills Consulting Limited

Cyprus

Email: dirk@lapishills.com

Robert Raszuk

Bloomberg LP

United States of America

Email: robert@raszuk.net

Bart Peirens

Proximus

Belgium

Email: bart.peirens@proximus.com

Hani Elmalky

Ericsson

United States of America

Email: hani.elmalky@gmail.com

Prem Jonnalagadda

Barefoot Networks

United States of America

Email: prem@barefootnetworks.com

Milad Sharif

Barefoot Networks

United States of America

Email: msharif@barefootnetworks.com

David Lebrun

Google

Belgium

Email: dlebrun@google.com

Stefano Salsano

Universita di Roma "Tor Vergata"

Italy

Email: stefano.salsano@uniroma2.it

Ahmed AbdelSalam

Gran Sasso Science Institute

Italy

Email: ahmed.abdelsalam@gssi.it

Gaurav Naik

Drexel University

United States of America

Email: gn@drexel.edu

Arthi Ayyangar

Arista

United States of America

Email: arthi@arista.com

Satish Mynam

Innovium Inc.

United States of America

      Email: smynam@innovium.com

      Wim Henderickx

      Nokia

      Belgium

      Email: wim.henderickx@nokia.com

      Shaowen Ma

      Juniper

      Singapore

      Email: mashao@juniper.net

      Ahmed Bashandy

      Individual

      United States of America

      Email: abashandy.ietf@gmail.com

      Francois Clad

      Cisco Systems, Inc.

      France

      Email: fclad@cisco.com

      Kamran Raza

      Cisco Systems, Inc.

      Canada

      Email: skraza@cisco.com

      Darren Dukes

      Cisco Systems, Inc.

      Canada

      Email: ddukes@cisco.com

      Patrice Brissete

Cisco Systems, Inc.

Canada

Email: pbrisset@cisco.com

Zafar Ali

Cisco Systems, Inc.

United States of America

Email: zali@cisco.com

8. References

8.1. Normative References

[I-D.voyer-6man-extension-header-insertion]
          Voyer, D., Filsfils, C., Dukes, D., Matsushima, S.,
          Leddy, J., Li, Z., and J. Guichard, "Deployments With
          Insertion of IPv6 Segment Routing Headers", Work in
          Progress, Internet-Draft, draft-voyer-6man-extension-
          header-insertion-10, 20 November 2020, <https://
          www.ietf.org/archive/id/draft-voyer-6man-extension-
          header-insertion-10.txt>.

[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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
          2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
          May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[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>.

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg,
          "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC
          8491, DOI 10.17487/RFC8491, November 2018, <https://
          www.rfc-editor.org/info/rfc8491>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy,
          J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing
          Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March
          2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8986]
          Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer,
          D., Matsushima, S., and Z. Li, "Segment Routing over IPv6
          (SRv6) Network Programming", RFC 8986, DOI 10.17487/
          RFC8986, February 2021, <https://www.rfc-editor.org/info/
          rfc8986>.

8.2. Informative References

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
          Litkowski, S., Bashandy, A., Filsfils, C., Francois, P.,
          Decraene, B., and D. Voyer, "Topology Independent Fast
          Reroute using Segment Routing", Work in Progress,
          Internet-Draft, draft-ietf-rtgwg-segment-routing-ti-
          lfa-08, 21 January 2022, <https://www.ietf.org/archive/
          id/draft-ietf-rtgwg-segment-routing-ti-lfa-08.txt>.

[I-D.ietf-spring-segment-routing-policy]
          Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A.,
          and P. Mattes, "Segment Routing Policy Architecture",
          Work in Progress, Internet-Draft, draft-ietf-spring-
          segment-routing-policy-18, 17 February 2022, <https://
          www.ietf.org/archive/id/draft-ietf-spring-segment-
          routing-policy-18.txt>.

Authors' Addresses

Clarence Filsfils
Cisco Systems, Inc.
Belgium

Email: cf@cisco.com

Pablo Camarillo Garvia (editor)
Cisco Systems, Inc.
Spain

Email: pcamaril@cisco.com

John Leddy
Individual Contributor
United States of America

Email: john@leddy.net

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Satoru Matsushima
SoftBank
1-9-1,Higashi-Shimbashi,Minato-Ku, Tokyo 105-7322
Japan

Email: satoru.matsushima@g.softbank.co.jp

Zhenbin Li
Huawei Technologies
China

Email: lizhenbin@huawei.com