Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 24, 2021

D. Voyer, Ed.
Bell Canada
C. Filsfils
D. Dukes, Ed.
Cisco Systems, Inc.
S. Matsushima
Softbank
J. Leddy
Individual Contributor
Z. Li
Huawei
J. Guichard
Futurewei

November 20, 2020

Deployments With Insertion of IPv6 Segment Routing Headers
draft-voyer-6man-extension-header-insertion-10

Abstract

SRv6 is deployed in multiple provider networks.

This document describes the usage of SRH insertion and deletion within the SR domain and how security and end-to-end integrity is guaranteed.

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

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 May 24, 2021.

Table of Contents

1. Introduction 2
2. Deployment Report 3
2.1. Deployments 3
2.2. Vendor and Open-Source Support 3
3. Deployment Experience With SRH Header Operarion 4
3.1. The SR Domain 5
4. Baseline Usecase 5
5. Choosing an SRv6 SID Block 6
6. Securing the SR Domain 7
7. MTU within the SR domain 7
8. VPN with SRv6 7
9. TILFA with SRv6 8
9.1. SRH Insertion Process 8
9.2. The Penultimate SID of the Inserted SRH is of PSP flavor 9
9.3. MTU-delta 9
10. Security Considerations 10
11. IANA Considerations 10
12. Contributors 10
13. References 10
13.1. Normative References 10
13.2. Informative References 11
Authors' Addresses 11

1. Introduction

[I-D.matsushima-spring-srv6-deployment-status] records multiple SRv6 deployments in multiple networks

In each deployment, traffic traversing an SR domain is encapsulated in an outer IPv6 header for its journey through the SR domain.

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

The SR domain always preserves the end-to-end integrity of traffic traversing it. No extension header is manipulated, inserted or removed from an inner transported packet. The packet leaving the SR domain is exactly the same (except for the hop-limit update) as the packet entering the SR domain.

The SR domain is designed with link MTU sufficiently greater than the MTU at the ingress edge of the SR domain.

2. Deployment Report

The following deployments are as of November 2019.

2.1. Deployments

Six operators have publicly reported SRv6 deployments with commercial traffic supported by linerate hardware. Each deployment follows the network design and SRH add/remove behavior described in this document.

Softbank

China Telecom

Iliad

China Unicom

CERNET2

MTN Uganda Ltd.

Further information can be found in [I-D.matsushima-spring-srv6-deployment-status]

2.2. Vendor and Open-Source Support

Eighteen unique implementations of SRv6 are available from multiple vendors and open source initiatives that support the SRH add/remove behavior described in this document:

Cisco ASR 9000

Cisco NCS 5500

Cisco NCS 560

Cisco NCS 540

Cisco ASR1000

Huawei ATN

Huawei CX600

Huawei NE40E

Hauwei ME60

Huawei NE5000E

Huawei NE9000

Huawei NG-OLT MA5800

Barefoot Tofino 1 NPU

Barefoot Tofino 2 NPU

Broadcom Jericho 1, 1+

Broadcom Jericho 2

Linux kernel

FD.io VPP

Marvell's Prestera Falcon CX 8500 family

Intel PAC N3000

Further information can be found in
[I-D.matsushima-spring-srv6-deployment-status]

3. Deployment Experience With SRH Header Operarion

3.1. The SR Domain

An SR Domain is defined in [RFC8402].

Section 5.2 of [I-D.ietf-6man-segment-routing-header] further describes the SR domain as a single system with delegation among components. It states:

> All intra SR Domain packets are of the SR Domain. The IPv6 header is originated by a node of the SR Domain, and is destined to a node of the SR Domain.
>
> All inter domain packets are encapsulated for the part of the packet journey that is within the SR Domain. The outer IPv6 header is originated by a node of the SR Domain, and is destined to a node of the SR Domain.

In other words, all packets within the SR domain have a source and destination address within the SR Domain.

The SR domain is secured as per Section 5.1 of [I-D.ietf-6man-segment-routing-header] and no external packet can enter the domain with a destination address equal to a segment of the domain.

In other words, no node outside the SR domain may send packets to, nor make direct use of, segments within the SR domain.

4. Baseline Usecase

The following abstract illustration shows the SR Domain, how traffic is encapsulated when traversing the SR domain, and (in subequent sections) how an SRH is inserted and processed for a packet traversing the SR domain. It is representative of all deployments in Section 2.1.

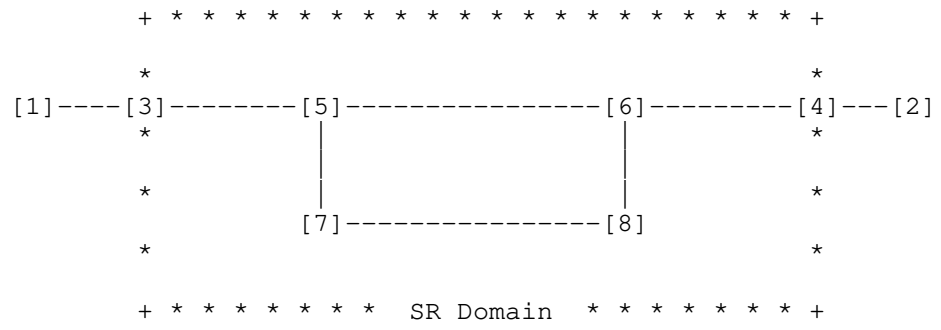

Figure 1

- 3 and 4 are SR Domain edge routers
- 5, 6, 7, and 8 are all SR Domain routers
- 1 and 2 are hosts outside the SR Domain

Since all inter domain packets are encapsulated for the part of the packet journey that is within the SR Domain, a packet sent from 1 and destined to 2 is encapsulated in an outer IP v6 header between nodes 3 and 4.

5. Choosing an SRv6 SID Block

Without revealing the specifics of each deployment, the following well-known technique can be used:

Obtain a GUA block from the respective registry (e.g. PPPP:PPP0::/28)

Sub-allocate a block for SID allocation (e.g. PPPP:PPPB:BB00::/40)

Allocate a /64 SID locator to each node in the domain that needs to provide network instruction (e.g. node 4 gets PPPP:PPPB:BB00:0004::/64 as a SID locator)

Vendors and operators have automated the process of locator selection, the details of which are outside the scope of this document.

6. Securing the SR Domain

The security measures defined in [I-D.ietf-6man-segment-routing-header] Section 5.1 are applied.

Protection level 1: filter external traffic entering the SR domain. For example, node 4 (on its interface from node 2) applies an ingress ACL that drops any packet with DA within the PPPP:PPPB:BB00::/40 block.

Protection level 2: filter internal traffic. For example, node 4 (on its interface from node 6) applies an ingress ACL that drops any packet with DA in PPPP:PPPB:BB00:0004::/64 block if SA is not within the block PPPP:PPP0::/28

Vendors and operators have automated the application of these protection levels, the details of which are outside the scope of this document.

7. MTU within the SR domain

The deployments, Section 2.1, leverage the extensively used practice of ensuring an MTU within the domain is bigger than the MTU on the external links of the domain.

More specifically, the MTU difference (MTU-Delta) is designed to be larger than the maximum encapsulation overhead deemed required by the deployment.

The exact number is operator specific and is outside the scope of this document. Some indications on how to plan this are provided in the following sections.

Any packet exceeding the MTU of a link generates an IPv6 ICMP error message "packet too big" back to the source of the packet.

8. VPN with SRv6

The deployments involve the creation of commercial SRv6-based VPN traffic as described in [I-D.ietf-bess-srv6-services].

The salient point to note is that no SRH needs to be inserted to realize an SRv6 VPN service.

The ingress PE encapsulates the inner packet in an outer header and sets the outer DA to the END.DT/DX SID signaled by the egress PE.

MTU-Delta must be >= 40 bytes to allow for the outer IPv6 encapsulation without fragmentation.

9. TILFA with SRv6

The deployments involve the delivery of sub-50msec TILFA protection to the commercial SRv6-based VPN traffic transported by the operator network [I-D.ietf-rtgwg-segment-routing-ti-lfa].

In these deployments, when a failure is detected, the Point of Local Repair (PLR) inserts an SRH implementing the precomputed TILFA backup path.

The following salient points are discussed:

SRH insertion process

The penultimate SID of the inserted SRH is of PSP flavor

MTU-delta planning

9.1. SRH Insertion Process

When an SRH is inserted by an intermediate node it walks the IPv6 header chain to the first header after the IPv6 header and inserts the SRH prior to that header.

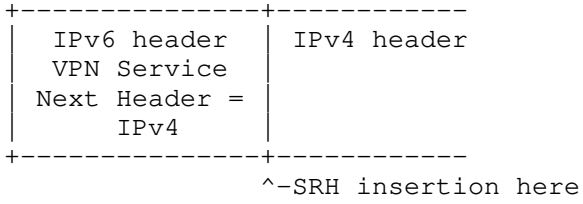

Figure 2

An SR Policy headend within the SR domain inserts an SRH as follows:

1. Determine where to insert the SRH.

2. Copy the destination address from the IPv6 header to Segment List[0] of the SRH to be inserted. This ensures the original destination address is restored upon execution of the final segment in the inserted SRH.

3. Increase the IPv6 header payload length field by the length in bytes of the inserted SRH.

If the resulting payload length exceeds 2^16 bytes generate an ICMP "Packet To Big" error message to the source with an MTU of 2^16 minus the length in bytes of the SRH and discard the packet. Note: this does not occur in reported deployments given the MTU design constraint.

4. Set the SRH next header field to the value in the next header field of the header that will precede the SRH.

5. Set the next header field of the header that will precede the SRH to the routing extension header (43)

6. Set the IPv6 destination address to the first segment in the segment list of the SRH to be inserted. This segment may or may not be present in the SRH depending on the use of a reduced SRH, see section 4.1.1 of [I-D.ietf-6man-segment-routing-header].

7. Insert the SRH into the packet at the location it should be inserted and resubmit the packet to the IPv6 module for transmission to the new destination.

9.2. The Penultimate SID of the Inserted SRH is of PSP flavor

The TILFA protection service is essentially a transparent service: it seeks to make the loss of a link, node or SRLG invisible to the transport service. It is also a very transient service as it only lasts a few hundreds of msec while the IGP converges.

Consistent with this transparent service definition, the deployments leverage a TILFA computation that ensures that the penultimate SID of the inserted SRH is of PSP flavor.

9.3. MTU-delta

The vendors reporting the listed deployments have collectively deployed TILFA in tens of SR-MPLS networks, in 6 SRv6 networks and have simulated their SRv6 algorithm in tens of collected real topologies. The inferred experience is that the probability that a TILFA backup path requires more than 2 SRv6 SIDs is very rare.

MTU-Delta must be >= 80 bytes.

40 bytes (VPN service)

+ 8 (SRH) (TILFA)

+ 2 * 16 (2 TILFA SID’s)

The maximum encapsulation size of any node within the SR domain is limited to a specific value, this maximum is used to calculate the maximum link MTU of interfaces ingress to the SR domain.

10. Security Considerations

Section 6 describes the method of securing the SR domain in the deployments listed.

All security considerations discussed in [I-D.ietf-6man-segment-routing-header] are equally applicable when an SRH insertion is performed.

11. IANA Considerations

This document doesn't introduce any IANA request.

12. Contributors

The authors would like to thank the following for their contributions: Robert Raszuk, Stefano Previdi, Stefano Salsano, Antonio Cianfrani, David Lebrun, Olivier Bonaventure, Prem Jonnalagadda, Milad Sharif, Hani Elmalky, Ahmed Abdelsalam, Arthi Ayyangar, Dirk Steinberg, Wim Henderickx.

13. References

13.1. Normative References

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-26 (work in progress), October 2019.

[I-D.ietf-bess-srv6-services]
Dawra, G., Filsfils, C., Talaulikar, K., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "SRv6 BGP based Overlay services", draft-ietf-bess-srv6-services-05 (work in progress), November 2020.

[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>.

[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>.

13.2. Informative References

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
Litkowski, S., Bashandy, A., Filsfils, C., Decraene, B., Francois, P., Voyer, D., Clad, F., and P. Camarillo, "Topology Independent Fast Reroute using Segment Routing", draft-ietf-rtgwg-segment-routing-ti-lfa-04 (work in progress), August 2020.

[I-D.matsushima-spring-srv6-deployment-status]
Matsushima, S., Filsfils, C., Ali, Z., Li, Z., and K. Rajaraman, "SRv6 Implementation and Deployment Status", draft-matsushima-spring-srv6-deployment-status-09 (work in progress), November 2020.

Authors' Addresses

Daniel Voyer (editor)
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Clarence Filsfils
Cisco Systems, Inc.
Belgium

Email: cfilsfil@cisco.com

Darren Dukes (editor)
Cisco Systems, Inc.
Ottawa
Canada

Email: ddukes@cisco.com

Satoru Matsushima
Softbank
Japan

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

John Leddy
Individual Contributor
USA

Email: john@leddy.net

Zhenbin Li
Huawei
China

Email: lizhenbin@huawei.com

James Guichard
Futurewei
USA

Email: james.n.guichard@futurewei.com