BMWG
Internet-Draft
Intended status: Informational
Expires: September 8, 2022

G. Fioccola
E. Vasilenko
P. Volpato
Huawei Technologies
L. Contreras
Telefonica

March 7, 2022

Benchmarking Methodology for IPv6 Segment Routing
draft-vfv-bmwg-srv6-bench-meth-01

Abstract

This document defines a methodology for benchmarking Segment Routing (SR) performance for Segment Routing over IPv6 (SRv6). It builds upon [RFC2544], [RFC5180] and [RFC8402].

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 September 8, 2022.

Copyright Notice

Table of Contents

1. Introduction 2
2. SRv6 Forwarding 3
3. Test Methodology 5
3.1. Test Setup 5
3.2. IGP and BGP Support 6
3.3. Frame Formats and Sizes 6
3.4. Protocol Addresses 6
3.5. Traffic with SRH 6
4. Reporting Format 7
5. SRv6 Forwarding Benchmarking Tests 8
5.1. Throughput 9
5.1.1. Throughput of a Source Node 9
5.1.2. Throughput of a Segment Endpoint Node 9
5.1.3. Throughput of a Transit Node 9
5.2. Latency 10
5.3. Frame Loss 10
5.4. System Recovery 10
5.5. Reset 10
6. SR Policy: protection performance 11
7. Security Considerations 11
8. IANA Considerations 11
9. Acknowledgements 11
10. References 11
10.1. Normative References 11
10.2. Informative References 12
Authors' Addresses 13

1. Introduction

Segment Routing (SR), defined in [RFC8402], leverages the source routing paradigm. The headend node steers a packet through an SR Policy [I-D.ietf-spring-segment-routing-policy], instantiated as an ordered list of segments. A segment, referred to by its Segment Identifier (SID), can have a semantic local to an SR node or global within an SR domain. SR supports per-flow explicit routing while maintaining per-flow state only at the ingress nodes to the SR domain.

However, there is no standard method defined to compare and contrast the foundational SR packet forwarding capabilities of network devices. This document aims to extend the efforts of [RFC1242] and [RFC2544] to SR network.

The SR architecture can be instantiated on two data-plane: SR over MPLS (SR-MPLS) and SR over IPv6 (SRv6). This document is limited to SRv6.

SR can be applied to the IPv6 architecture with a new type of routing header called the SR Header (SRH) [RFC8754]. An instruction is associated with a segment and encoded as an IPv6 address. An SRv6 segment is also called an SRv6 SID. An SR Policy is instantiated as an ordered list of SRv6 SIDs in the routing header. The active segment is indicated by the Destination Address (DA) of the packet.

For Segment Routing, PUSH, NEXT, and CONTINUE are operations applied by the forwarding plane.

PUSH consists of the insertion of a segment at the top of the segment list. In SRv6, the top of the segment list is represented by the first segment in the SRH.

NEXT consists of the inspection of the next segment. The active segment is completed and the next segment becomes active. In SRv6, NEXT is implemented as the copy of the next segment from the SRH to the destination address of the IPv6 header.

CONTINUE happens when the active segment is not completed; hence, it remains active. In SRv6, the CONTINUE operation is the plain IPv6 forwarding action of a regular IPv6 packet according to its destination address.

[RFC5180] provides benchmarking methodology recommendations that address IPv6-specific aspects, such as evaluating the forwarding performance of traffic containing extension headers.

The purpose of this document is to describe a methodology specific to the benchmarking of Segment Routing. The methodology described is a complement for [RFC5180].

2. SRv6 Forwarding

In IPv6, a Prefix-SID is allocated in the form of an IPv6 address. For the IPv6 data plane, a new type of IPv6 Routing Extension Header, called Segment Routing Header (SRH) has been defined. The SRH contains the Segment List as an ordered list of IPv6 addresses: each address in the list is a SID. A dedicated field, referred to as

Segments Left, is used to maintain the pointer to the active SID of the Segment List.

There are three different categories of nodes that may be involved in segment routing networks.

The SR source node is the headend node and steers a packet into an SR Policy. It can be a host originating an IPv6 packet or an SR domain ingress router encapsulating a received packet into an outer IPv6 packet and inserts the SRH in the outer IPv6 header. It sets the first SID of the SR Policy as IPv6 Destination Address of the packet.

The SR transit node forwards packets destined to a remote segment as a normal IPv6 packet on the basis of the IPv6 destination address, because the IPv6 destination address does not locally match with a segment. Indeed, according to [RFC8200] the only node allowed to inspect the Routing Extension Header (and therefore the SRH) is the node corresponding to the destination address of the packet.

The SR segment endpoint node receives packets whose IPv6 destination address is locally configured as a segment. It creates Forwarding Information Base (FIB) entries for its local SIDs. For each SR packet, it inspects the SRH and replaces the IPv6 destination address with the new active segment.

The operations applied by the SRv6 packet processing are different at the SR source, Transit and SR segment endpoint nodes.

The processing of the SR source node corresponds to the sequence of the insertion of the SRH, composed of SIDs stored in reverse order, and setting of the IPv6 Destination Address as first SID of the SR Policy. It can be performed by encapsulating a packet into an outer IPv6 packet with an SRH. Another possibility is to perform the insertion of an SRH as a new header between the IPv6 header and the Next Header (e.g. the Transport Layer Header, TCP or UDP). This option only applies to IPv6 packets and it is especially suited in case the source host is acting as headend node.

The processing of the SR segment endpoint node corresponds to the detection of the new active segment, which is the next segment in the Segment List and the related modification of the IPv6 destination address of the outer IPv6 header. Then packets are forwarded on the basis of the IPv6 forwarding table.

The processing of the SR transit node corresponds to normal forwarding of the packets containing the SR header. In SRv6 the transit nodes do not need to be SRv6 aware, as every IPv6 router can act as an SRv6 transit node since any IPv6 node will maintain a plain

IPv6 FIB entry for any prefix, no matter if the prefix represents a segment or not.

[I-D.ietf-spring-segment-routing-policy] specifies the concepts of SR Policy and steering into an SR Policy. The header of a packet steered in an SR Policy is augmented with the ordered list of segments associated with that SR Policy. SR Policy state is instantiated only on the headend node, that steers a flow into an SR Policy. Indeed intermediate and endpoint nodes do not require any state to be maintained. SR Policies can be instantiated on the headend dynamically and on demand basis. Moreover, signaling can be used in the case of a controller based deployment. For all these reasons, SR Policies scale better than traditional TE mechanisms.

In addition to the basic SRv6 packet processing, the SRv6 Network Programming model [RFC8986] describes a set of functions that can be associated to segments and executed in a given SRv6 node.

Examples of such functions are described in [RFC8986], but, in practice, any behavior and function can be associated to a local SID in a node, in order to apply any special processing on the packet. Obviously, the definition of a standardized set of segment routing functions facilitates the deployment of SR domains with interoperable equipment from multiple vendors.

According to [RFC8986], 128 bit SID can be logically split into three fields and interpreted as LOCATOR:FUNCTION:ARGS (in short LOC:FUNCT:ARG) where LOC includes the L most significant bits, FUNCT the following F bits and ARG the remaining A bits, where 128=L+F+A. The LOC corresponds to an IPv6 prefix (for example with a length of 48, 56 or 64 bits) that can be distributed by the routing protocols and provides the reachability of a node that hosts a number of functions. All the different functions residing in a node have a different FUNCT code, so that their SIDs will be different. The ARG bits are used to provide information (arguments) to a function. From the routing point of view, the solution is scalable, as a single prefix is distributed for a node, which implements a potentially large number of functions and related arguments.

3. Test Methodology

3.1. Test Setup

The Device Under Test (DUT) is connected to the test ports on the test tool according to [RFC2544].

The test topology recommended for the SRv6 performance evaluation are the same as IPv6 and are described in [RFC5180] and [RFC2544], in

both single-port and multi-port scenarios. Single-port testing measures per-interface forwarding performance, while multi-port testing measures the scalability of forwarding performance across the entire platform.

3.2. IGP and BGP Support

It is RECOMMENDED that all of the ports on the DUT and test tool support a Segment Routing extensions for dynamic Interior Gateway Protocol (IGP) for routing such as IS-IS [I-D.ietf-lsr-isis-srv6-extensions] and OSPF [I-D.ietf-lsr-ospfv3-srv6-extensions] as well as Border Gateway Protocol (BGP) [I-D.ietf-bess-srv6-services].

As specified in [RFC8402], in the context of an IGP-based distributed control plane, two topological segments are defined: the IGP-Adjacency segment and the IGP-Prefix segment; while, in the context of a BGP-based distributed control plane, two topological segments are defined: the BGP peering segment and the BGP-Prefix segment.

The distribution method that is used (e.g. OSPF, IS-IS, BGP) MUST be reported.

3.3. Frame Formats and Sizes

The tests for SRv6 will use the Frame characteristics as described in [RFC5180].

As specified in [RFC5180], for Ethernet, the following frame sizes SHOULD be used for benchmarking over this media type: 64, 128, 256, 512, 1024, 1280, and 1518 bytes. Note that the recommended 1518-byte frame size represents the maximum size of an untagged Ethernet frame. A frame size commonly used in operational environments is 1522 bytes, the max length for a VLAN-tagged frame.

3.4. Protocol Addresses

IANA reserved an IPv6 address block for use with IPv6 benchmark testing (see [RFC5180]). IPv6 source and destination addresses for the test streams SHOULD belong to the IPv6 range assigned by IANA.

3.5. Traffic with SRH

The extension header chain recommended in [RFC5180] for testing is: Routing header (24-32 bytes), Destination options header (8 bytes), Fragment header (8 bytes). This was considered a real-life extension-header chain but it does not fit well for SRv6.

The length of the SRH is (n x 16 + 8) bytes, where n is the number of segments. So, for most of the SRv6 application the recommendation of [RFC5180] is not enough. In addition, it is worth mentioning that the length of SRv6 packets is increased in Topology Independent Loop-Free Alternate (TI-LFA) Fast Reroute (FRR), binding SID, and microloop avoidance scenarios.

For SRv6, the extension header chain characteristics and length that are used MUST be reported and the DUT MUST traverse the chain of extension headers, so the impact on performance can be observed.

4. Reporting Format

There are new parameters that MUST be added to the parameters specified in [RFC5180] and [RFC2544]:

- SRv6 types of nodes.
- Number of Segments considered in the SRH.
- Extension header chain (including SRH) characteristics and length.
- Global SIDs or Local SID forwarding behavior.
- SR Headend or Endpoint Behaviors eventually associated with a SID, as specified in [RFC8986].

For the sake of completeness, the following Figure 1 reports all the SR Headend or Endpoint Behaviors, as defined in [RFC8986]. But, in most cases, it may not be necessary to test all the services and it is possible to select a subset.

H.Encaps	SR Headend with Encapsulation in an SR Policy
H.Encaps.Red	H.Encaps with Reduced Encapsulation
H.Encaps.L2	H.Encaps Applied to Received L2 Frames
H.Encaps.L2.Red	H.Encaps.Red Applied to Received L2 Frames
End	Endpoint
End.X	Endpoint with L3 cross-connect
End.T	Endpoint with specific IPv6 table lookup
End.DX6	Endpoint with decapsulation and IPv6 cross-

	connect
End.DX4	Endpoint with decapsulation and IPv4 cross-connect
End.DT6	Endpoint with decapsulation and specific IPv6 table lookup
End.DT4	Endpoint with decapsulation and specific IPv4 table lookup
End.DT46	Endpoint with decapsulation and specific IP table lookup
End.DX2	Endpoint with decapsulation and L2 cross-connect
End.DX2V	Endpoint with decapsulation and VLAN L2 table lookup
End.DT2U	Endpoint with decapsulation and unicast MAC L2 table lookup
End.DT2M	Endpoint with decapsulation and L2 table flooding
End.B6.Encaps	Endpoint bound to an SRv6 Policy with encapsulation
End.B6.Encaps.Red	End.B6.Encaps with reduced SRH
End.BM	Endpoint bound to an SR-MPLS Policy

Figure 1: SR Policy Headend and Endpoint Behaviors

5. SRv6 Forwarding Benchmarking Tests

This document recommends the same benchmarking tests described in [RFC2544] and [RFC5180] while observing the DUT setup and the traffic setup considerations described above. Indeed, the specificities of SRv6, for example the SRH processing, require additional benchmarking steps.

5.1. Throughput

This section contains the description of the tests that are related to the characterization of a DUT's SRv6 traffic forwarding throughput.

The list of segments for SRv6 is represented as a list of IPv6 addresses, included in the SRH. There are three distinct types of nodes that are involved in segment routing networks.

5.1.1. Throughput of a Source Node

Objective: To obtain the DUT's Throughput during the packet processing of a Source Node. It is when the Source SR node, which corresponds to the headend node, encapsulates a received packet into an outer IPv6 packet and inserts the SR Header (SRH) as a Routing Extension Header in the outer IPv6 header. The Segment List in the SRH is composed of SIDs and the Source SR node sets the first SID of the SR Policy as IPv6 Destination Address of the packet.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.1.2. Throughput of a Segment Endpoint Node

Objective: To obtain the DUT's Throughput during the packet processing of a Segment Endpoint Node. It is when the SR Segment Endpoint node receives packets whose IPv6 destination address is locally configured as a segment. The SR Segment Endpoint node inspects the SR header: it detects the new active segment, i.e. the next segment in the Segment List, modifies the IPv6 destination address of the outer IPv6 header and forwards the packet on the basis of the IPv6 forwarding table.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.1.3. Throughput of a Transit Node

Objective: To obtain the DUT's Throughput during the packet processing of a Transit Node. It is when a Transit node forwards the packet containing the SR header as a normal IPv6 packet because the IPv6 destination address does not locally match with a segment.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.2. Latency

Objective: To determine the latency as defined in [RFC5180] for each of the SRv6 forwarding operations.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.3. Frame Loss

Objective: To determine the frame-loss rate (as defined in [RFC5180]) for each of the SRv6 forwarding operations of a DUT throughout the entire range of input data rates and frame sizes.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.4. System Recovery

Objective: To characterize the speed at which a DUT recovers from an overload condition for each of the SRv6 forwarding operations.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

5.5. Reset

Objective: To characterize the speed at which a DUT recovers from a device or software reset for each of the SRv6 forwarding operations.

Procedure: Same as [RFC5180].

Reporting Format: Same as [RFC5180] but adding the additional parameters specified in Section 4.

6. SR Policy: protection performance

[RFC6414] provides common terminology and metrics for benchmarking the performance of protection mechanisms.

An SR Policy can be used for Traffic Engineering (TE), Operations, Administration, and Maintenance (OAM), or Fast Reroute (FRR) reasons. Protection allows that, in the event the interface associated with the Adj-SID is down, the packet can still be forwarded via an alternate path. The use of protection is clearly a policy-based decision that determines, for example, that the packet processing by the source node is done to forward a packet over a backup path calculated using TI-LFA. There are 2 different protection mechanisms for SR-TE: Segment protection specified in [I-D.ietf-spring-segment-protection-sr-te-paths] and Path protection introduced in [I-D.ietf-spring-segment-routing-policy].

7. Security Considerations

Benchmarking methodologies are limited to technology characterization in a laboratory environment, with dedicated address space and constraints. Special capabilities SHOULD NOT exist in the DUT/SUT specifically for benchmarking purposes. Any implications for network security arising from the DUT/SUT SHOULD be identical in the lab and in production networks. The benchmarking network topology is an independent test setup and MUST NOT be connected to devices that may forward the test traffic into a production network or misroute traffic to the test management network.

There are no specific security considerations within the scope of this document.

8. IANA Considerations

This document has no IANA actions.

9. Acknowledgements

TBD

10. References

10.1. Normative References

[RFC1242] Bradner, S., "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, DOI 10.17487/RFC1242, July 1991, <https://www.rfc-editor.org/info/rfc1242>.

[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>.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, DOI 10.17487/RFC2544, March 1999, <https://www.rfc-editor.org/info/rfc2544>.

[RFC5180] Popoviciu, C., Hamza, A., Van de Velde, G., and D. Dugatkin, "IPv6 Benchmarking Methodology for Network Interconnect Devices", RFC 5180, DOI 10.17487/RFC5180, May 2008, <https://www.rfc-editor.org/info/rfc5180>.

[RFC6414] Poretsky, S., Papneja, R., Karthik, J., and S. Vapiwala, "Benchmarking Terminology for Protection Performance", RFC 6414, DOI 10.17487/RFC6414, November 2011, <https://www.rfc-editor.org/info/rfc6414>.

[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>.

[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>.

10.2. Informative References

[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-12 (work in progress), March 2022.

[I-D.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Segment Routing over IPv6 Dataplane", draft-ietf-lsr-isis-srv6-extensions-18 (work in progress), October 2021.

[I-D.ietf-lsr-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-ietf-lsr-ospfv3-srv6-extensions-03 (work in progress), November 2021.

[I-D.ietf-spring-segment-protection-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu, "Segment Protection for SR-TE Paths", draft-ietf-spring-segment-protection-sr-te-paths-02 (work in progress), January 2022.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-19 (work in progress), March 2022.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

Authors' Addresses

Giuseppe Fioccola
Huawei Technologies
Riesstrasse, 25
Munich 80992
Germany

Email: giuseppe.fioccola@huawei.com

Eduard Vasilenko
Huawei Technologies
17/4 Krylatskaya str.
Moscow 121614
Russia

Email: vasilenko.eduard@huawei.com

Paolo Volpato
Huawei Technologies
Via Lorenteggio, 240
Milan 20147
Italy

Email: paolo.volpato@huawei.com

Luis Miguel Contreras Murillo
Telefonica
Spain

Email: luismiguel.contrerasmurillo@telefonica.com