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 MPLS Segment Routing
draft-vfv-bmwg-srmpls-bench-meth-01

Abstract

This document defines a methodology for benchmarking Segment Routing (SR) performance for Segment Routing over MPLS (SR-MPLS). It builds upon [RFC2544], [RFC5695] 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 will expire on September 8, 2022.

Copyright Notice

Table of Contents

1. Introduction 2
2. SR-MPLS Forwarding 3
3. Test Methodology 5
  3.1. Test Setup 5
  3.2. IGP and BGP Support 5
  3.3. Frame Formats and Sizes 5
4. Reporting Format 6
5. SR-MPLS Forwarding Benchmarking Tests 6
  5.1. Throughput 6
    5.1.1. Throughput for SR-MPLS PUSH 6
    5.1.2. Throughput for SR-MPLS NEXT 6
    5.1.3. Throughput for SR-MPLS CONTINUE 7
  5.2. Latency 7
  5.3. Frame Loss 7
  5.4. System Recovery 7
  5.5. Reset 8
6. SR Policy: protection performance 8
7. Security Considerations 8
8. IANA Considerations 9
9. Acknowledgements 9
10. References 9
  10.1. Normative References 9
  10.2. Informative References 10
Authors' Addresses 11

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 SR-MPLS.

SR can be directly applied to the Multiprotocol Label Switching (MPLS) architecture with no change to the forwarding plane [RFC8660]. A segment is encoded as an MPLS label. An SR Policy is instantiated as a stack of labels.

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 SR-MPLS, the top of the segment list is the outer label of the label stack.

NEXT consists of the inspection of the next segment. The active segment is completed and the next segment becomes active. In SR-MPLS, NEXT is implemented as a POP of the top label.

CONTINUE happens when the active segment is not completed; hence, it remains active. In SR-MPLS, the CONTINUE operation is implemented as a SWAP of the top label.

[RFC5695] describes a methodology specific to the benchmarking of MPLS forwarding devices, by considering the most common MPLS packet forwarding scenarios and corresponding performance measurements.

The purpose of this document is to describe a methodology specific to the benchmarking of Segment Routing. The methodology described is a complement for [RFC5695].

2. SR-MPLS Forwarding

In MPLS, a Prefix-SID is allocated in the form of an MPLS label. For SR-MPLS, Segment Routing does not require any change to the MPLS forwarding plane. An SR Policy is instantiated through the MPLS Label Stack: the Segment IDs (SIDs) of a Segment List are inserted as MPLS Labels. The classical forwarding functions available for MPLS networks allow implementing the SR operations.

The operations applied by the SR-MPLS forwarding plane are PUSH, NEXT, and CONTINUE.

The PUSH operation corresponds to the Label Push function, according to the MPLS label pushing rules specified in [RFC3032]. It consists of pushing one or more MPLS labels on top of an incoming packet then sending it out of a particular physical interface or virtual interface towards a particular next hop.

The NEXT operation corresponds to the Label Pop function, that consists of removing the topmost label. The action before and/or after the popping depends on the instruction associated with the active SID on the received packet prior to the popping. It is equivalent to Penultimate Hop Popping (PHP).

The CONTINUE operation corresponds to the Label Swap function, according to the MPLS label-swapping rules in [RFC3031]. It consists of associating an incoming label with an outgoing interface and outgoing label and forwarding the packet on the outgoing interface. It is equivalent to Ultimate Hop Popping (UHP).

The encapsulation of an IP packet into an SR-MPLS packet is performed at the edge of an SR-MPLS domain, reusing the MPLS Forwarding Equivalent Class (FEC) concept. A Forwarding Equivalent Class (FEC) can be associated with an SR Policy ([RFC8660]). When pushing labels onto a packet's label stack, the Time-to-Live (TTL) field and the Traffic Class (TC) field of each label stack entry must also be set.

All SR nodes in the SR domain use an IGP signaling extension to advertise their own prefix SIDs. After receiving advertised prefix SIDs, each SR node calculates the prefix SIDs to the advertisers. The prefix SID advertisement can be an absolute value advertisement or an index value advertisement. In this regard, the mapping of Segments to MPLS Labels (SIDs) is an important process in the SR-MPLS data plane. Each router can advertise its own available label space to be used for Global Segments called Segment Routing Global Block (SRGB) and an identical range of labels (SRGB) should be used in all routers in order to simplify services and operations. In the SR domain Global Segments can be identified by an index, which has to be re-mapped into a label, or by an absolute value. This is relevant for the nodes that perform the NEXT operation to the segments, because the label for the next segments needs to be crafted accordingly.

[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.

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 recommended topology for SR-MPLS Forwarding Benchmarking should be the same as MPLS and it is described in [RFC5695] for both single-port and multi-port scenarios. Indeed, the number of ports is a parameter that MUST be reported.

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 [RFC8667] and OSPF [RFC8665] as well as Border Gateway Protocol (BGP) [RFC8669].

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 peer 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 SR-MPLS will use the Frame characteristics as described in [RFC5695].

Note that [RFC5695] requires exactly a single entry in the MPLS label stack in an MPLS packet. In other words, the depth of the label stack is set to one.

To ensure successful delivery of Layer 2 frames carrying SR-MPLS packets and realistic benchmarking, it is RECOMMENDED to set the media MTU value to the effective maximum frame payload size (payload of 1500 octets for Ethernet).

The number of entries in the label stack MUST be reported. In addition, it MUST be chosen taking into account this condition.

4. Reporting Format

There are new parameters that MUST be replaced or added to the parameters specified in [RFC5695]:

- o SR-MPLS Forwarding Operations (PUSH/ NEXT/ CONTINUE).
- o Number of Segments considered in the MPLS Label Stack.
- o Global SIDs or Local SID forwarding behavior.
- o SR Policy headend or endpoint behavior.

5. SR-MPLS Forwarding Benchmarking Tests

This document recommends the same benchmarking tests described in [RFC2544] and [RFC5695] while observing the DUT setup and the traffic setup considerations specific for SR-MPLS as described above. It may 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 SR-MPLS traffic forwarding throughput.

The list of segments for SR-MPLS is represented as a stack of MPLS labels. There are three distinct operations to be tested: PUSH, NEXT and CONTINUE. These correspond to the three forwarding operations of an MPLS packet: PUSH (or LSP Ingress), SWAP, or POP (or LSP Egress).

5.1.1. Throughput for SR-MPLS PUSH

Objective: To obtain the DUT's Throughput during PUSH forwarding operation. It is similar to label Push or LSP Ingress forwarding operation, as per [RFC5695]. Non-reserved MPLS label values MUST be used.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] but adding the additional parameters specified in Section 4.

5.1.2. Throughput for SR-MPLS NEXT

Objective: To obtain the DUT's Throughput during NEXT forwarding operation. It is equivalent to MPLS Label Pop or Penultimate Hop

Popping (PHP), as per [RFC5695]. Non-reserved MPLS label values MUST be used.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] but adding the additional parameters specified in Section 4.

5.1.3. Throughput for SR-MPLS CONTINUE

Objective: To obtain the DUT's Throughput during CONTINUE forwarding operation. It is equivalent to MPLS Label Swap or Ultimate Hop Popping (UHP), as per [RFC5695]. Non-reserved MPLS label values MUST be used.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] but adding the additional parameters specified in Section 4.

5.2. Latency

Objective: To determine the latency as defined in [RFC5695] for each of the SR-MPLS forwarding operations.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] but adding the additional parameters specified in Section 4.

5.3. Frame Loss

Objective: To determine the frame-loss rate (as defined in [RFC5695]) for each of the SR-MPLS forwarding operations of a DUT throughout the entire range of input data rates and frame sizes.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] 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 SR-MPLS forwarding operations.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695]. 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 SR-MPLS forwarding operations.

Procedure: Same as [RFC5695].

Reporting Format: Same as [RFC5695] 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. [RFC6894] provides detailed test cases with different topologies and scenarios that should be considered to effectively benchmark MPLS-FRR protection mechanisms and failover times on the data plane. The same approach can be considered also for Segment Routing 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 a PUSH operation 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>.

[RFC5695] Akhter, A., Asati, R., and C. Pignataro, "MPLS Forwarding Benchmarking Methodology for IP Flows", RFC 5695, DOI 10.17487/RFC5695, November 2009, <https://www.rfc-editor.org/info/rfc5695>.

[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>.

[RFC6894] Papneja, R., Vapiwala, S., Karthik, J., Poretsky, S., Rao, S., and JL. Le Roux, "Methodology for Benchmarking MPLS Traffic Engineered (MPLS-TE) Fast Reroute Protection", RFC 6894, DOI 10.17487/RFC6894, March 2013, <https://www.rfc-editor.org/info/rfc6894>.

[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>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

10.2. Informative References

[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.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019,

<https://www.rfc-editor.org/info/rfc8665>.

[RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <https://www.rfc-editor.org/info/rfc8667>.

[RFC8669] Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <https://www.rfc-editor.org/info/rfc8669>.

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