Performance Measurement for
Segment Routing Networks with MPLS Data Plane
draft-gandhi-mpls-rfc6374-sr-01

Abstract

Segment Routing (SR) leverages the source routing paradigm. RFC 6374 specifies protocol mechanisms to enable the efficient and accurate measurement of packet loss, one-way and two-way delay, as well as related metrics such as delay variation in MPLS networks using probe messages. This document utilizes these mechanisms for Performance Delay and Loss Measurements in Segment Routing (SR) networks with MPLS data plane (SR-MPLS), for both SR links and end-to-end SR Policies. In addition, this document defines Return Path TLV for two-way performance measurement and Block Number TLV for loss measurement.

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 http://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."

Copyright Notice
Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents
1. Introduction .................................................. 3
2. Conventions Used in This Document ......................... 4
   2.1. Requirements Language ................................ 4
   2.2. Abbreviations ......................................... 4
   2.3. Reference Topology .................................... 5
3. Overview ....................................................... 5
4. Probe Query and Response Packets ............................ 6
   4.1. Probe Packet Header for SR-MPLS Policies ............ 6
   4.2. Probe Packet Header for SR-MPLS Links ............... 7
   4.3. Probe Response Message for SR-MPLS Links and Policies 7
      4.3.1. One-way Measurement Mode ....................... 7
      4.3.2. Two-way Measurement Mode ....................... 7
      4.3.3. Loopback Measurement Mode ..................... 8
   4.4. Return Path TLV ....................................... 8
5. Performance Delay Measurement ............................... 9
   5.1. Delay Measurement Message Format .................... 9
   5.2. Timestamps ........................................... 10
6. Performance Loss Measurement ............................... 10
   6.1. Loss Measurement Message Format ..................... 11
   6.2. Block Number TLV .................................... 11
7. Performance Measurement for P2MP SR Policies ................ 11
8. ECMP for SR-MPLS Policies .................................. 12
9. SR Link Extended TE Metrics Advertisements ................ 13
10. Security Considerations ................................... 13
11. IANA Considerations ...................................... 13
12. References ................................................ 14
   12.1. Normative References ................................ 14
   12.2. Informative References ............................. 14
Acknowledgments ............................................... 17
Contributors .................................................. 17
Authors’ Addresses ........................................... 17
1. Introduction

Service provider’s ability to satisfy Service Level Agreements (SLAs) depend on the ability to measure and monitor performance metrics for packet loss and one-way and two-way delay, as well as related metrics such as delay variation. The ability to monitor these performance metrics also provides operators with greater visibility into the performance characteristics of their networks, thereby facilitating planning, troubleshooting, and network performance evaluation.

Segment Routing (SR) leverages the source routing paradigm and greatly simplifies network operations for Software Defined Networks (SDNs). SR is applicable to both Multiprotocol Label Switching (SR-MPLS) and IPv6 (SRv6) data planes. SR takes advantage of the Equal-Cost Multipaths (ECMPs) between source and transit nodes, between transit nodes and between transit and destination nodes. SR Policies as defined in [I-D.spring-segment-routing-policy] are used to steer traffic through a specific, user-defined paths using a stack of Segments. Built-in SR Performance Measurement (PM) is one of the essential requirements to provide Service Level Agreements (SLAs).

[RFC6374] specifies protocol mechanisms to enable the efficient and accurate measurement of performance metrics in MPLS networks using probe messages. The One-Way Active Measurement Protocol (OWAMP) defined in [RFC4656] and Two-Way Active Measurement Protocol (TWAMP) defined in [RFC5357] provide capabilities for the measurement of various performance metrics in IP networks. However, mechanisms defined in [RFC6374] are more suitable for Segment Routing (SR) when using MPLS data plane (SR-MPLS). [RFC6374] also supports IEEE 1588 timestamps [IEEE1588] and "direct mode" Loss Measurement (LM), which are required in SR networks.

[RFC7876] specifies the procedures to be used when sending and processing out-of-band performance measurement probe replies over an UDP return path when receiving RFC 6374 based probe queries. These procedures can be used to send out-of-band PM replies for both SR-MPLS links and Policies [I-D.spring-segment-routing-policy] for one-way measurement.

This document utilizes the probe-based mechanisms defined in [RFC6374] for Performance Delay and Loss Measurements in SR networks with MPLS data plane, for both SR links and end-to-end SR Policies. In addition, this document defines Return Path TLV for two-way performance measurement and Block Number TLV for loss measurement. The Performance Measurements (PM) for SR links are used to compute extended Traffic Engineering (TE) metrics for delay and loss and can be advertised in the network using the routing protocol extensions.
2. Conventions Used in This Document

2.1. 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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Abbreviations

ACH: Associated Channel Header.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

G-ACh: Generic Associated Channel (G-ACh).

GAL: Generic Associated Channel (G-ACh) Label.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.


PM: Performance Measurement.

PSID: Path Segment Identifier.

PTP: Precision Time Protocol.

SID: Segment ID.

SL: Segment List.

SR: Segment Routing.

SR-MPLS: Segment Routing with MPLS data plane.

TC: Traffic Class.

TE: Traffic Engineering.

URO: UDP Return Object.
2.3. Reference Topology

In the reference topology shown in Figure 1, the sender node R1 initiates a performance measurement probe query and the responder node R5 sends a probe response for the query message received. The probe response is typically sent back to the sender node R1. The nodes R1 and R5 may be directly connected via a link enabled with Segment Routing or there exists a Point-to-Point (P2P) SR Policy [I-D.spring-segment-routing-policy] on node R1 with destination to node R5. In case of Point-to-Multipoint (P2MP), SR Policy originating from source node R1 may terminate on multiple destination leaf nodes [I-D.spring-sr-p2mp-policy].

```
+-------+        Query        +-------+
|       | - - - - - - - - - ->|       |
|   R1  |---------------------|   R5  |
|       |< - - - - - - - - - |       |
+-------+       Response      +-------+

Figure 1: Reference Topology
```

3. Overview

One-way delay and two-way delay measurement procedure defined in Section 2.4 of [RFC6374] are used. Transmit and Receive packet loss measurement procedures defined in Section 2.2 and Section 2.6 of [RFC6374] are used. One-way loss measurement provides receive packet loss whereas two-way loss measurement provides both transmit and receive packet loss. For both links and end-to-end SR Policies, no PM session for delay or loss measurement is created on the responder node R5 [RFC6374].

For Performance Measurement, probe query and response messages are sent as following:

- For Delay Measurement, the probe messages are sent on the congruent path of the data traffic by the sender node, and are used to measure the delay experienced by the actual data traffic flowing on the links and SR Policies.

- For Loss Measurement, the probe messages are sent on the congruent path of the data traffic by the sender node, and are used to collect the receive traffic counters for the incoming link or incoming SID where the probe query messages are received at the responder node (incoming link or incoming SID needed since the
For SR Policy performance measurement, in order to ensure that the probe query message is processed by the intended responder node, destination address TLV [RFC6374] can be sent in the probe query message. The responder node only replies if it is the intended destination for the probe query.

The In-Situ Operations, Administration, and Maintenance (IOAM) mechanisms for SR-MPLS defined in [I-D.mpls-ioam-sr] are used to carry PM information in-band as part of the data traffic packets, and are outside the scope of this document.

4. Probe Query and Response Packets

4.1. Probe Packet Header for SR-MPLS Policies

As described in Section 2.9.1 of [RFC6374], MPLS PM probe query and response messages flow over the MPLS Generic Associated Channel (G-ACh). A probe packet for an end-to-end measurement for SR Policy contains SR-MPLS label stack [I-D.spring-segment-routing-policy], with the G-ACh Label (GAL) at the bottom of the stack (with S=1). The GAL is followed by an Associated Channel Header (ACH), which identifies the message type, and the message payload following the ACH as shown in Figure 2.

![Figure 2: Probe Packet Header for an End-to-end SR-MPLS Policy](image-url)

The SR-MPLS label stack can be empty (as shown in Figure 3) to indicate Implicit NULL label case.
4.2. Probe Packet Header for SR-MPLS Links

As described in Section 2.9.1 of [RFC6374], MPLS PM probe query and response messages flow over the MPLS Generic Associated Channel (G-ACh). A probe packet for SR-MPLS links contains G-ACh Label (GAL) (with S=1). The GAL is followed by an Associated Channel Header (ACH), which identifies the message type, and the message payload following the ACH as shown in Figure 3.

```
 0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             GAL (value 13)            | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version| Reserved      | GAL Channel Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Probe Packet Header for an SR-MPLS Link

4.3. Probe Response Message for SR-MPLS Links and Policies

4.3.1. One-way Measurement Mode

In one-way performance measurement mode [RFC7679], the PM sender node can receive "out-of-band" probe replies by properly setting the UDP Return Object (URO) TLV in the probe query message. The URO TLV (Type=131) is defined in [RFC7876] and includes the UDP-Destination-Port and IP Address. In particular, if the sender sets its own IP address in the URO TLV, the probe response is sent back by the responder node to the sender node. In addition, the "control code" in the probe query message is set to "out-of-band response requested".

4.3.2. Two-way Measurement Mode

In two-way performance measurement mode [RFC6374], when using a bidirectional path, the probe response message is sent back to the sender node on the congruent path of the data traffic on the reverse direction SR Link or associated SR Policy [I-D.bidir-sr] using a message with format similar to their probe query message. In this case, the "control code" in the probe query message is set to "in-band response requested".

A Path Segment Identifier (PSID) [I-D.spring-mpls-path-segment] of the forward SR-MPLS Policy can be used to find the associated reverse SR-MPLS Policy [I-D.bidir-sr] and to send back the probe response message for two-way measurement.
4.3.3. Loopback Measurement Mode

The Loopback measurement mode defined in Section 2.8 of [RFC6374] can be used to measure round-trip delay for a bidirectional SR Path [I-D.bidir-sr]. The probe query messages in this case carries the reverse SR Path label stack as part of the MPLS header. The GAL is still carried at the bottom of the label stack (with S=1). The responder node does not process the PM probe messages and generate response messages.

4.4. Return Path TLV

For two-way performance measurement, the responder node needs to send the probe response message on a specific reverse path. The sender node can request in the probe query message to the responder node to send a response message back on a given reverse path (e.g. co-routed path for two-way measurement). This way the destination node does not require any additional SR Policy state.

For one-way performance measurement, the sender node address may not be reachable via IP route from the responder node. The sender node in this case needs to send its reachability path information to the responder node.

[RFC6374] defines DM and LM probe query messages that can include one or more optional TLVs. New TLV Type (TBA1) is defined in this document for Return Path to carry reverse path for probe response messages (in the payload of the message). The format of the Return Path TLV is shown in Figure 4A and 4B:

```
0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type = TBA1 | Length | Reserved |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Return Path Sub-TLVs |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

**Figure 4A: Return Path TLV**

```
0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type | Length | Reserved |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Segment List Sub-TLV in the Return Path TLV can be one of the following Types:

- Type (value 1): SR-MPLS Label Stack of the Reverse SR Path
- Type (value 2): SR-MPLS Binding SID [I-D.pce-binding-label-sid] of the Reverse SR Policy

The Return Path TLV is optional. The PM sender node MUST only insert one Return Path TLV in the probe query message and the responder node MUST only process the first Return Path TLV in the probe query message and ignore other Return Path TLVs if present. The responder node MUST send probe response message back on the reverse path specified in the Return Path TLV and MUST NOT add Return Path TLV in the probe response message. In the absence of Return Path TLV, in two-way measurement mode, the probe response message is sent back on the incoming physical interface where the probe query message is received.

5. Performance Delay Measurement

5.1. Delay Measurement Message Format

As defined in [RFC6374], MPLS DM probe query and response messages use Associated Channel Header (ACH) (value 0x000C for delay measurement) [RFC6374], which identifies the message type, and the message payload following the ACH. For both SR links and end-to-end measurement for SR-MPLS Policies, the same MPLS DM ACH value is used.

The DM message payload as defined in Section 3.2 of [RFC6374] is used for SR-MPLS delay measurement, for both SR links and end-to-end SR Policies.
5.2. Timestamps

The Section 3.4 of [RFC6374] defines timestamp format that can be used for delay measurement. The IEEE 1588 Precision Time Protocol (PTP) timestamp format [IEEE1588] is used by default as described in Appendix A of [RFC6374], preferred with hardware support in Segment Routing networks.

Note that for one-way delay measurement mode, clock synchronization between the sender and responder nodes is required. The two-way delay measurement mode and loopback measurement mode do not require clock synchronization between the sender and responder nodes.

6. Performance Loss Measurement

The LM protocol can perform two distinct kinds of loss measurement as described in Section 2.9.8 of [RFC6374].

- In inferred mode, LM will measure the loss of specially generated test messages in order to infer the approximate data plane loss level. Inferred mode LM provides only approximate loss accounting.

- In direct mode, LM will directly measure data plane packet loss. Direct mode LM provides perfect loss accounting, but may require hardware support.

For both of these modes of LM, Path Segment Identifier (PSID) [I-D.spring-mpls-path-segment] is used for accounting received traffic on the egress node of the SR-MPLS Policy as shown in Figure 5. Different values of PSID can be used to measure packet loss per SR-MPLS Policy, per Candidate Path or per Segment List of the SR Policy.

```
|                  PSID                 | TC  |S|      TTL      |
|                  GAL (value 13)       | TC  |S|      TTL      |
|0 0 0 1|Version| Reserved | GAL Channel Type |
```

Figure 5: With Path Segment Identifier for SR-MPLS Policy
6.1. Loss Measurement Message Format

As defined in [RFC6374], MPLS LM probe query and response messages use Associated Channel Header (ACH) (value 0x000A for direct loss measurement or value 0x000B for inferred loss measurement), which identifies the message type, and the message payload following the ACH. For both SR links and end-to-end measurement for SR-MPLS Policies, the same MPLS LM ACH value is used.

The LM message payload as defined in Section 3.1 of [RFC6374] is used for SR-MPLS loss measurement, for both SR links and end-to-end SR Policies.

6.2. Block Number TLV

The Loss Measurement using Alternate-Marking method defined in [RFC8321] requires to color the data traffic. To be able to compare the transmit and receive traffic counters of the matching color, the Block Number (or color) of the traffic counters is carried by the probe query and response messages for loss measurement. Probe query and response messages specified in [RFC6374] for Loss Measurement do not identify the Block Number of the counters.

[RFC6374] defines probe query and response messages that can include one or more optional TLVs. New TLV Type (value TBA2) is defined in this document to carry the Block Number (8-bit) of the traffic counters in the probe query and response messages for loss measurement. The format of the Block Number TLV is shown in Figure 6:

```
+------------------+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type TBA2   |    Length     | Reserved      | Block Number  |
+------------------+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: Block Number TLV

The Block Number TLV is optional. The PM sender node SHOULD only insert one Block Number TLV in the probe query message and the responder node in the probe response message SHOULD return the first Block Number TLV from the probe query messages and ignore other Block Number TLVs if present. In probe messages, the counters MUST belong to the same Block Number.

7. Performance Measurement for P2MP SR Policies
The procedures for delay and loss measurement described in this document for Point-to-Point (P2P) SR-MPLS Policies [I-D.spring-segment-routing-policy] are also equally applicable to the Point-to-Multipoint (P2MP) SR-MPLS Policies [I-D.spring-sr-p2mp-policy] as following:

- The sender root node sends probe query messages using the Replication Segment defined in [I-D.spring-sr-p2mp-policy] for the P2MP SR Policy as shown in Figure 7.
- Each responder leaf node adds the "Source Address" TLV (Type 130) [RFC6374] with its IP address in the probe response messages. This TLV allows the sender root node to identify the responder leaf nodes of the P2MP SR Policy.
- The P2MP root node measures the end-to-end delay and loss performance for each P2MP leaf node of the P2MP SR Policy.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Replication SID          | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              GAL (value 13)           | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version| Reserved      | GAL Channel Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: With Replication Segment for SR-MPLS Policy

8. ECMP for SR-MPLS Policies

An SR Policy can have ECMPs between the source and transit nodes, between transit nodes and between transit and destination nodes. Usage of Anycast SID [RFC8402] by an SR Policy can result in ECMP paths via transit nodes part of that Anycast group. The PM probe messages need to be sent to traverse different ECMP paths to measure performance delay of each of the ECMP path of an SR Policy.

Forwarding plane has various hashing functions available to forward packets on specific ECMP paths. For SR-MPLS Policy, sweeping of entropy label [RFC6790] values can be used in PM probe messages to take advantage of the hashing function in forwarding plane to influence the ECMP path taken by them.

The considerations for performance loss measurement for different...
ECMP paths of an SR Policy are outside the scope of this document.

9. SR Link Extended TE Metrics Advertisements

The extended TE metrics for SR link delay and loss computed using the performance measurement procedures described in this document can be advertised in the routing domain as follows:

- For OSPF, ISIS, and BGP-LS, protocol extensions defined in [RFC7471], [RFC8570], and [RFC8571] are used, respectively for advertising the extended TE link metrics in the network.

- The advertised delay-variance metric is computed as specified in Section 4.2 of [RFC5481].

- The extended TE link one-way delay metrics can be computed using two-way delay measurement or round-trip delay measurement from loopback mode by dividing the measured delay values by 2.

- The extended TE link delay and loss metrics are advertised for Layer 2 bundle members in OSPF [I-D.lsr-ospf-l2bundles] and ISIS [I-D.isis-l2bundles] using the same mechanisms defined in [RFC7471] and [RFC8570], respectively.

10. Security Considerations

This document describes the procedures for performance delay and loss measurement for SR-MPLS networks, for both links and end-to-end SR Policies using the mechanisms defined in [RFC6374] and [RFC7876]. This document does not introduce any additional security considerations other than those covered in [RFC6374], [RFC7471], [RFC8570], [RFC8571], and [RFC7876].

11. IANA Considerations

IANA is requested to allocate a value for the following optional Return Path TLV Type for RFC 6374 to be carried in PM probe query messages:

- Type TBA1: Return Path TLV

IANA is requested to allocate the values for the following Sub-TLV Types for the Return Path TLV for RFC 6374.
o Type (value 1): SR-MPLS Label Stack of the Reverse SR Path

o Type (value 2): SR-MPLS Binding SID [I-D.pce-binding-label-sid] of the Reverse SR Policy

IANA is also requested to allocate a value for the following optional Block Number TLV Type for RFC 6374 to be carried in the PM probe query and response messages for loss measurement:

o Type TBA2: Block Number TLV

12. References

12.1. Normative References


12.2. Informative References


[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and


[I-D.lsr-ospf-l2bundles] Talaulikar, K., et al., "Advertising L2 Bundle Member Link Attributes in OSPF"
draft-ketant-lsr-ospf-l2bundles, work in progress.

draft-ietf-isis-l2bundles, work in progress.

[I-D.bidir-sr] Li, C., et al., "PCEP Extensions for Associated Bidirectional Segment Routing (SR) Paths",
draft-li-pce-sr-bidir-path, work in progress.
Acknowledgments

The authors would like to thank Thierry Couture for the discussions on the use-cases for the performance measurement in segment routing networks. The authors would like to thank Greg Mirsky for providing many useful comments and suggestions. The authors would also like to thank Stewart Bryant, Sam Aldrin, Tarek Saad, and Rajiv Asati for their review comments.

Contributors

Sagar Soni
Cisco Systems, Inc.
Email: sagsoni@cisco.com

Patrick Khordoc
Cisco Systems, Inc.
Email: pkhordoc@cisco.com

Zafar Ali
Cisco Systems, Inc.
Email: zali@cisco.com

Pier Luigi Ventre
CNIT
Italy
Email: pierluigi.ventre@cnit.it

Authors’ Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada
Email: rgandhi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Email: cfilsfil@cisco.com

Daniel Voyer
Bell Canada
Email: daniel.voyer@bell.ca

Stefano Salsano  
Universita di Roma "Tor Vergata"  
Italy  
Email: stefano.salsano@uniroma2.it

Mach(Guoyi) Chen  
Huawei  
Email: mach.chen@huawei.com