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

Abstract

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 synthetic probe messages. This document reviews how these
mechanisms can be used 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. 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.

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 ............. 3
   2.1. Abbreviations .......................... 3
   2.2. Reference Topology ..................... 4
3. Probe Query and Response Packets ................. 5
   3.1. Probe Packet Header for SR-MPLS Policies ..... 5
   3.2. Probe Packet Header for SR-MPLS Links ........ 6
   3.3. Probe Response Message for SR-MPLS Links and Policies ... 6
       3.3.1. One-way Measurement Mode ............. 6
       3.3.2. Two-way Measurement Mode ............. 7
           3.3.2.1. Return Path TLV ................ 7
       3.3.3. Loopback Measurement Mode ............. 7
4. Performance Delay Measurement ................... 7
   4.1. Delay Measurement Message Format .......... 7
   4.2. Timestamps ................................ 8
5. Performance Loss Measurement .................... 8
   5.1. Loss Measurement Message Format .......... 9
       5.1.1. Block Number TLV .................... 9
6. Performance Measurement for P2MP SR Policies ...... 9
7. ECMP for SR-MPLS Policies ..................... 10
8. SR Link Extended TE Metrics Advertisements ....... 10
9. Security Considerations ....................... 11
10. IANA Considerations ........................ 11
11. References .................................. 11
    11.1. Normative References .................. 11
    11.2. Informative References ................ 11
Acknowledgments ................................ 14
Contributors .................................. 14
Authors’ Addresses ............................. 14
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.

[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 reviews how synthetic probe-based mechanisms defined in [RFC6374] can be used for Performance Delay and Loss Measurements in SR networks with MPLS data plane, for both SR links and end-to-end SR Policies. 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. Abbreviations

ACH: Associated Channel Header.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

G-ACh: Generic Associated Channel (G-ACh).
2.2. Reference Topology

In the reference topology shown in Figure 1, the querier 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 querier 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].
For delay and loss measurements, for both links and end-to-end SR Policies, no PM session is created on the responder node R5. One-way delay and two-way delay measurements are defined in Section 2.4 of [RFC6374]. Transmit and Receive packet loss measurements are defined in Section 2.2 and Section 2.6 of [RFC6374]. One-way loss measurement provides receive packet loss whereas two-way loss measurement provides both transmit and receive packet loss.

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

- For Delay Measurement, the probe messages are sent on the congruent path of the data traffic by the querier 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 querier 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 used as the responder node has no PM session state present).

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

3. Probe Query and Response Packets

3.1. 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.

```
+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+
|                  | Label(1)        | TC  |S|      TTL      |                  |
+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+
```

Figure 2: Probe Packet Header for an End-to-end SR-MPLS Policy

The SR-MPLS label stack can be empty (as shown in Figure 3) to indicate Implicit NULL label case.

### 3.2. 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.

```
+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+
|                  | Label(n)        | TC  |S|      TTL      |                  |
+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+-----------------+
```

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

### 3.3. Probe Response Message for SR-MPLS Links and Policies

#### 3.3.1. One-way Measurement Mode

In one-way performance measurement mode [RFC7679], the PM querier
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 querier sets its own IP address in the URO TLV, the probe response is sent back by the responder node to the querier node. In addition, the "control code" in the probe query message is set to "out-of-band response requested". The "Source Address" TLV (Type 130), and "Return Address" TLV (Type 1), if present in the probe query message, are not used to send probe response message.

3.3.2. 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 querier node on the congruent path of the data traffic on the reverse direction SR Link or SR Policy 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 reverse SR-MPLS Policy and to send back the probe response message for two-way measurement.

3.3.2.1. Return Path TLV

For two-way performance measurement, the querier node can request the responder node to send a response message back on a given reverse path (typically co-routed path for two-way measurement). Return Path TLV defined in [I-D.spring-rfc6374-srpm-udp] can be used to carry reverse SR path information as part of the payload of the probe query message.

3.3.3. 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. 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. Performance Delay Measurement

4.1. 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.

4.2. 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. As an alternative, Network Time Protocol (NTP) timestamp format can also be used [RFC6374].

Note that for one-way delay measurement mode, clock synchronization between the querier and responder nodes using the methods detailed in [RFC6374] is required. The two-way delay measurement mode and loopback measurement mode do not require clock synchronization between the querier and responder nodes.

5. 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 4. 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.
5.1. 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.

5.1.1. Block Number TLV

The Loss Measurement using Alternate-Marking method defined in [RFC8321] requires to identify the Block Number (or color) of the traffic counters carried by the probe query and response messages. Block Number TLV defined in [I-D.spring-rfc6374-srpm-udp] is used to carry Block Number for the traffic counters in the probe query and response messages for loss measurement.

6. Performance Measurement for P2MP SR Policies

The procedures for delay and loss measurement reviewed 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 querier root node sends probe query messages using the either Spray P2MP segment or TreeSID P2MP segment defined in [I-D.spring-sr-p2mp-policy] over the P2MP SR Policy as shown in Figure 5.
7. 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 an SR Policy.

Forwarding plane has various hashing functions available to forward packets on specific ECMP paths. For SR-MPLS Policy, entropy label [RFC6790] 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.

8. SR Link Extended TE Metrics Advertisements

The extended TE metrics for SR link delay and loss computed using the performance measurement procedures reviewed 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 extended TE link delay metrics advertised are minimum-delay,
maximum-delay, average-delay, and delay-variance for one-way.

- The delay-variance metric is computed as specified in Section 4.2 of [RFC5481].
- The 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 loss metric advertised is one-way percentage packet loss.

9. Security Considerations

This document reviews 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].

10. IANA Considerations

This document does not require any IANA actions.

11. References

11.1. Normative References


11.2. Informative References


Acknowledgments

The authors would like to thank Thierry Couture for various 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 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

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
Università di Roma "Tor Vergata"
Italy
Email: stefano.salsano@uniroma2.it

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

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