SPRING Working Group
Internet-Draft
Intended Status: Informational
Expires: March 18, 2019

R. Gandhi, Ed.
C. Filsfils
Cisco Systems, Inc.
D. Voyer
Bell Canada
S. Salsano
Universita di Roma "Tor Vergata"
P. L. Ventre
CNIT
M. Chen
Huawei

September 14, 2018

Performance Measurement in Segment Routing Networks with MPLS Data Plane

draft-gandhi-spring-sr-mpls-pm-03

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 probe messages. This document reviews how these mechanisms can be used for Delay and Loss Performance Measurements (PM) in Segment Routing (SR) networks with MPLS data plane (SR-MPLS), for both SR links and end-to-end SR Policies. The performance measurements for SR links are used to compute extended Traffic Engineering (TE) metrics for delay and loss and are advertised in the network using 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) 2018 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

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). The [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 links and SR Policies [I-D.spring-segment-routing-policy] for one-way measurement.

This document reviews how probe based mechanisms defined in [RFC6374] can be used for Delay and Loss Performance Measurements (PM) in SR networks with MPLS data plane, for both SR links and end-to-end SR Policies. The performance measurements for SR links are used to compute extended Traffic Engineering (TE) metrics for delay and loss and are advertised in the network using routing protocol extensions.

2. Conventions Used in This Document

2.1. Abbreviations

ACH: Associated Channel Header.

DFLag: Data Format Flag.

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.

NTP: Network Time Protocol.

PM: Performance Measurement.

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.

2.2. 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 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].

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

Figure 1: Reference Topology

Both delay and loss performance measurement is performed in-band for the traffic traversing between node R1 and 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.

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. 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.

```
 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Label(0)             | EXP |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Label(n)             | EXP |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  GAL                  | EXP |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version|  Reserved     | GAL Channel Type              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

The SR-MPLS label stack can be empty 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). 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  GAL                  | EXP |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version|  Reserved     | GAL Channel Type              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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 Probe Response Message

For one-way performance measurement [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 Probe Response Message

For two-way performance measurement [RFC6374], when using a bidirectional channel, the probe response message is sent back to the querier node in-band 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 [I-D.spring-mpls-path-segment] [I-D.pce-sr-path-segment] of the forward SR Policy can be used to find the reverse SR Policy to send the probe response message.

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 Policies, the same MPLS DM ACH value is used.

The DM message payload as defined in [RFC6374] is used for SR-MPLS delay measurement, for both SR links and end-to-end SR Policies. The DM message payload format is defined as following in [RFC6374]:

```
 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version| Flags |  Control Code |           Message Length      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  QTF  |  RTF  | RPTF  |              Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Session Identifier         |    DS     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Timestamp 1                        |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Timestamp 4                        |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                            TLV Block                          ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Delay Measurement Message Payload Format

The meanings of the fields are summarized in the following table, see [RFC6374] for details.

Field	Meaning
Version	Protocol version
Flags	Message control flags
Control Code	Code identifying the query or response type
QTF	Querier timestamp format

	(see Section 3.4 of [RFC6374])
RTF	Responder timestamp format (see Section 3.4 of [RFC6374])
RPTF	Responder’s preferred timestamp format
Reserved	Reserved for future specification
Session Identifier	Set arbitrarily by the querier
Differentiated Services (DS) Field	Differentiated Services Code Point (DSCP) being measured
Timestamp 1-4	64-bit timestamp values (see Section 3.4 of [RFC6374])
TLV Block	Optional block of Type-Length-Value fields

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], but it may require hardware support. As an alternative, Network Time Protocol (NTP) timestamp format can also be used [RFC6374].

Note that for one-way delay measurement, clock synchronization between the querier and responder nodes using the methods detailed in [RFC6374] is required. The two-way delay measurement does 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].

- o 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.

- o 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 [I-D.spring-mpls-path-segment] [I-D.pce-sr-path-segment] is required for accounting received traffic on the egress node of the SR-MPLS 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 Policies, the same MPLS LM ACH value is used.

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

```
 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version| Flags |  Control Code |        Message Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| DFlags|  OTF  |                  Reserved                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Session Identifier          |    DS     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Origin Timestamp                       |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Counter 1                           |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Counter 4                           |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                           TLV Block                           ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Loss Measurement Message Payload Format

The meanings of the fields are summarized in the following table, see

[RFC6374] for details.

Field	Meaning
Version	Protocol version
Flags	Message control flags
Control Code	Code identifying the query or response type
Message Length	Total length of this message in bytes
Data Format Flags (DFlags)	Flags specifying the format of message data
Origin Timestamp Format (OTF)	Format of the Origin Timestamp field
Reserved	Reserved for future specification
Session Identifier	Set arbitrarily by the querier
Differentiated Services (DS) Field	Differentiated Services Code Point (DSCP) being measured
Origin Timestamp	64-bit field for query message transmission timestamp
Counter 1-4	64-bit fields for LM counter values
TLV Block	Optional block of Type-Length-Value fields

6. 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 are also equally applicable to the Point-to-Multipoint (P2MP) SR Policies.

The responder node may add the "Source Address" TLV (Type 130) [RFC6374] in the probe response message. This TLV allows the querier node to identify the responder node for the SR Policy.

7. 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:

- o For OSPF, ISIS, and BGP-LS, protocol extensions defined in [RFC7471], [RFC7810] [I-D.lsr-isis-rfc7810bis], and [I-D.idr-te-pm-bgp] are used, respectively for advertising the extended TE link metrics in the network.

- o The extended TE link delay metrics advertised are minimum-delay, maximum-delay, average-delay, and delay-variance for one-way.

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

- o The one-way delay metrics can be computed using two-way measurement by dividing the measured delay values by 2.

- o The extended TE link loss metric advertised is one-way percentage packet loss.

8. 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]. This document does not introduce any additional security considerations other than those covered in [RFC6374], [RFC7471], [RFC7810], and [RFC7876].

9. IANA Considerations

This document does not require any IANA actions.

10. References

10.1. Normative References

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS networks’, RFC 6374, September 2011.

[RFC7876] Bryant, S., Sivabalan, S., and Soni, S., "UDP Return Path for Packet Loss and Delay Measurement for MPLS Networks", RFC 7876, July 2016.

10.2. Informative References

[IEEE1588] IEEE, "1588-2008 IEEE Standard for a Precision Clock

Synchronization Protocol for Networked Measurement and Control Systems", March 2008.

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, September 2006.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, October 2008.

[RFC5481] Morton, A. and B. Claise, "Packet Delay Variation Applicability Statement", RFC 5481, March 2009.

[RFC7679] Almes, G., et al., "A One-Way Delay Metric for IP Performance Metrics (IPPM)’, RFC 7679, January 2016.

[RFC7471] Giacalone, S., et al., "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, March 2015.

[RFC7810] Previdi, S., et al., "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, May 2016.

[I-D.lsr-isis-rfc7810bis] Ginsberg, L., et al., "IS-IS Traffic Engineering (TE) Metric Extensions", draft-ietf-lsr-isis-rfc7810bis, work in progress.

[I-D.idr-te-pm-bgp] Ginsberg, L. Ed., et al., "BGP-LS Advertisement of IGP Traffic Engineering Performance Metric Extensions", draft-ietf-idr-te-pm-bgp, work in progress.

[I-D.spring-segment-routing-policy] Filsfils, C., et al., "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy, work in progress.

[I-D.spring-sr-p2mp-policy] Voyer, D. Ed., et al., "SR Replication Policy for P2MP Service Delivery", draft-voyer-spring-sr-p2mp-policy, work in progress.

[I-D.spring-mpls-path-segment] Cheng, W., et al., "Path Segment in MPLS Based Segment Routing Network", draft-cheng-spring-mpls-path-segment, work in progress.

[I-D.pce-sr-path-segment] Li, C., et al., "Path Computation Element Communication Protocol (PCEP) Extension for Path Identification in Segment Routing (SR)", draft-li-pce-sr-path-segment, work in progress.

Acknowledgments

To be added.

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

Daniel Bernier
Bell Canada
Email: daniel.bernier@bell.ca

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

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

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