MPLS Working Group R. Gandhi, Ed.
Internet-Draft C. Filsfils
Intended status: Standards Track Cisco Systems, Inc.
Expires: January 27, 2020 D. Voyer
Bell Canada
S. Salsano
Universita di Roma "Tor Vergata"
M. Chen
Huawei
July 26, 2020

Performance Measurement Using RFC 6374 for Segment Routing Networks with MPLS Data Plane

draft-ietf-mpls-rfc6374-sr-00

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 networks with MPLS data plane (SR-MPLS), for both SR-MPLS Links and end-to-end Paths including SR-MPLS 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

Copyright Notice

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.

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.ietf-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 when using MPLS data plane (SR-MPLS). [RFC6374] also supports "direct mode" Loss Measurement (LM), which is required in SR networks.

[RFC7876] specifies the procedures to be used when sending and processing out-of-band performance measurement probe responses over an UDP return path when receiving RFC 6374 based probe queries. These procedures can be used to send out-of-band probe responses for both SR-MPLS Links and Policies 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-MPLS Links and end-to-end Paths including SR-MPLS 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-MPLS 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 defined in [RFC7471], [RFC8570], and [RFC8571].

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.

NTP: Network Time Protocol.

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 querier node R1 initiates a performance measurement probe query and the responder node R5 sends a probe response message for the query message received. The probe response message is typically sent back to the querier node R1.

SR is enabled with MPLS data plane on nodes R1 and R5. The nodes R1 and R5 may be directly connected via a Link enabled with MPLS or there exists a Point-to-Point (P2P) SR-MPLS Path e.g. Policy [I-D.ietf-spring-segment-routing-policy] on node R1 (called head-end) with destination to node R5 (called tail-end).

```
                  t1                t2
                 /                   \
         +-------+       Query        +-------+
         |       | - - - - - - - - - ->|       |
         |   R1  |=====================|   R5  |
         |       |<- - - - - - - - - - |       |
         +-------+       Response     +-------+
                 \                   /
                  t4                t3
          Querier                      Responder
```

Figure 1: Reference Topology

3. Overview

For one-way, two-way and round-trip delay measurements, the procedures defined in Section 2.4 and Section 2.6 of [RFC6374] are used. For transmit and receive packet loss measurements, the procedures defined in Section 2.2 and Section 2.6 of [RFC6374] are used. For both SR-MPLS Links and end-to-end 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:

o 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 SR-MPLS Links and Policies.

o 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 needed since the responder node does not have PM session state present).

The In-Situ Operations, Administration, and Maintenance (IOAM) mechanisms for SR-MPLS defined in [I-D.gandhi-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 Messages

4.1. Probe Message for SR-MPLS Links

As described in Section 2.9.1 of [RFC6374], probe query and response messages flow over the MPLS Generic Associated Channel (G-ACh). A probe message 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 2. The probe messages are routed over the Links for both delay and loss measurement using the same procedure described in [RFC6374]. For SR-MPLS Links, the TTL value is set to 1 in the SR-MPLS header for one-way and two-way measurement modes.

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

Figure 2: Probe Message Header for an SR-MPLS Link

4.2. Probe Message for SR-MPLS Policies

As described in Section 2.9.1 of [RFC6374], probe query and response messages flow over the MPLS Generic Associated Channel (G-ACh). A probe message for an end-to-end SR-MPLS Policy measurement contains SR-MPLS label stack [I-D.ietf-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 3. For SR-MPLS Policies, the TTL value is set to 255 in the SR-MPLS header.

```
 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(1)             | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Label(n)             | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  GAL (value 13)       | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version| Reserved      | GAL Channel Type              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Example Probe Message Header for an End-to-end SR-MPLS Policy

The SR-MPLS label stack can be empty (as shown in Figure 2) to indicate Implicit NULL label case.

For SR-MPLS Policy performance measurement, in order to ensure that the probe query message is processed by the intended responder node, Destination Address TLV (Type 129) [RFC6374] MAY be sent in the probe query message. The responder node only returns Success in Control Code if it is the intended destination for the probe query. Otherwise, it MUST return 0x15: Error - Invalid Destination Node Identifier [RFC6374].

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 querier node can receive "out-of-band" probe responses 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 node 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". In this delay measurement mode, as per Reference Topology, timestamps t1 and t2 are collected by the probes to measure one-way delay. The one-way mode is applicable to both SR-MPLS Links and Policies.

4.3.2. Two-way Measurement Mode

In two-way performance measurement mode [RFC6374], when using a bidirectional SR path [I-D.ietf-pce-sr-bidir-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-MPLS Link or associated SR-MPLS Policy [I-D.ietf-pce-sr-bidir-path] 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". In this delay measurement mode, as per Reference Topology, all timestamps t1, t2, t3, and t4 are collected by the probes. All four timestamps are used to measure two-way delay. The two-way mode is applicable to both SR-MPLS Links and Policies.

Specifically, the probe response message is sent back on the incoming physical interface where the probe query message is received. This is useful for example, in case of two-way measurement mode for Link delay.

The Path Segment Identifier (PSID) [I-D.ietf-spring-mpls-path-segment] of the forward SR-MPLS Policy in the probe query can be used to find the associated reverse SR-MPLS Policy [I-D.ietf-pce-sr-bidir-path] to send the probe response message for two-way measurement of SR-MPLS Policy unless when using the Return Path TLV.

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-MPLS Path [I-D.ietf-pce-sr-bidir-path]. The probe query messages in this case carries the reverse 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 probe messages and generate response messages, and hence Loopback Request object (Type 3) is not required for SR. In this delay measurement mode, as per Reference Topology, the timestamps t1 and t4 are collected by the probes. Both these timestamps are used to measure round-trip delay.

4.4. Return Path TLV Extensions

For two-way performance measurement, the responder node needs to send the probe response message on a specific reverse path. The querier 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-MPLS Policy state.

For one-way performance measurement, the querier node address may not be reachable via IP route from the responder node. The querier 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 information in the probe query messages (in the payload). The format of the Return Path TLV is shown in Figure 4 and Figure 5:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type = TBA1  |     Length    |       Reserved                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Return Path Sub-TLVs                      |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length    |       Reserved                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Label(1)                                  |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Label(n)                                  |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Segment List Sub-TLV in Return Path TLV

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

o Type (value 1): SR-MPLS Label Stack of the Reverse Path

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

The Return Path TLV is a Mandatory TLV Type. The querier 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.

5. 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-MPLS Links and end-to-end Policies measurements, 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-MPLS Links and end-to-end 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], with hardware support in Segment Routing networks.

6. 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 (PSID) [I-D.ietf-spring-mpls-path-segment] is used for accounting received traffic on the egress node of the SR-MPLS Policy as shown in Figure 6. 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.

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

Figure 6: Example 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-MPLS Links and end-to-end Policies measurements, 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-MPLS Links and end-to-end Policies.

6.2. Block Number TLV Extensions

The loss measurement using Alternate-Marking method defined in [RFC8321] requires to color the data traffic. To be able to correlate 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. The probe query and response messages currently specified in [RFC6374] for loss measurement do not identify the Block Number of the counters. The Block Number can also be used to aggregate performance metrics collected.

[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 7:

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

Figure 7: Block Number TLV

The Block Number TLV is a Mandatory TLV Type. The querier 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.

6.3. Combined Loss/Delay Measurement Message Format

As defined in [RFC6374], Combined DM+LM probe query and response messages use Associated Channel Header (ACH) (value 0x000D for direct loss and delay measurement or value 0x000E for inferred loss and delay measurement), which identifies the message type, and the message payload following the ACH. For both SR-MPLS Links and end-to-end Policies measurements, the same MPLS ACH value is used.

The message payload as defined in Section 3.3 of [RFC6374] is used for SR-MPLS combined delay and loss measurement, for both SR-MPLS Links and end-to-end Policies.

7. Performance Measurement for P2MP SR-MPLS Policies

The Point-to-Multipoint (P2MP) SR-MPLS Path that originates from a root node terminates on multiple destinations called leaf nodes (e.g. P2MP SR-MPLS Policy [I-D.voyer-pim-sr-p2mp-policy] or P2MP Transport [I-D.shen-spring-p2mp-transport-chain]).

The procedures for delay and loss measurement described in this document for P2P SR-MPLS Policies are also equally applicable to the P2MP SR-MPLS Policies. The procedure for one-way measurement is defined as following:

- o The querier root node sends probe query messages using the Tree-SID defined in [I-D.voyer-pim-sr-p2mp-policy] for the P2MP SR-MPLS Policy as shown in Figure 8.

o The probe query messages can contain the replication SID as defined in [I-D.voyer-spring-sr-replication-segment].

o 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 querier root node to identify the responder leaf nodes of the P2MP SR-MPLS Policy.

o The P2MP root node measures the delay and loss performance for each P2MP leaf node of the end-to-end P2MP SR-MPLS Policy.

```
 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              Tree-SID                 | TC  |S|      TTL      |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 .                                                               .
 .                                                               .
 .                                                               .
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              GAL (value 13)           | TC  |S|      TTL      |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |0 0 0 1|Version| Reserved      | GAL Channel Type              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: Example Probe Query with Tree-SID for SR-MPLS Policy

The probe query messages can also be sent using the scheme defined for P2MP Transport using Chain Replication that may contain Bud SID as defined in [I-D.shen-spring-p2mp-transport-chain].

The considerations for two-way mode for performance measurement for P2MP SR-MPLS Policy (e.g. for bidirectional SR-MPLS Path) are outside the scope of this document.

8. ECMP for SR-MPLS Policies

An SR-MPLS 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-MPLS Policy can result in ECMP paths via transit nodes part of that Anycast group. The probe messages need to be sent to traverse different ECMP paths to measure performance delay of each of the ECMP path of an SR-MPLS 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 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-MPLS Policy are outside the scope of this document.

9. SR-MPLS Link Extended TE Metrics Advertisements

The extended TE metrics for SR-MPLS 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 also 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.ketant-lsr-ospf-l2bundles] and ISIS [RFC8668] using the same mechanisms defined in [RFC7471] and [RFC8570], respectively.

10. Backwards Compatibility

The procedures defined in this document are backwards compatible with the procedures defined in [RFC6374] at both querier and responder nodes. If the responder does not support the new Mandatory TLV Types defined in this document, it MUST return Error 0x17: Unsupported Mandatory TLV Object as per [RFC6374].

11. Security Considerations

This document describes the procedures for performance delay and loss measurement for SR-MPLS networks, for both SR-MPLS Links and end-to-end 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].

12. IANA Considerations

IANA is requested to allocate a value for the following Mandatory Return Path TLV Type for [RFC6374] to be carried in probe query message from the "MPLS Loss/Delay Measurement TLV Object" registry contained within the "Generic Associated Channel (G-ACh) Parameters" registry set:

- Type TBA1: Return Path TLV

IANA is requested to create a sub-registry for "Return Path Sub-TLV Type" for the Return Path TLV. All code points in the range 1 through 32759 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC8126]. Code points in the range 32760 through 65279 in this registry shall be allocated according to the "First Come First Served" procedure as specified in [RFC8126]. Remaining code points are allocated according to Table 1:

Value	Description	Reference
0- 32767	Mandatory TLV, unassigned	IETF Review
32768 - 65279	Optional TLV, unassigned	First Come First Served
65280 - 65519	Experimental	This document
65520 - 65534	Private Use	This document
65535	Reserved	This document

Table 1: Return Path Sub-TLV Type Registry

IANA is requested to allocate the values for the following Sub-TLV Types from this registry.

- Type (value 1): SR-MPLS Label Stack of the Reverse Path

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

IANA is also requested to allocate a value for the following Mandatory Block Number TLV Type for RFC 6374 to be carried in the probe query and response messages for loss measurement from the "MPLS Loss/Delay Measurement TLV Object" registry contained within the "Generic Associated Channel (G-ACh) Parameters" registry set:

- Type TBA2: Block Number TLV

13. References

13.1. Normative References

[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>.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC7876] Bryant, S., Sivabalan, S., and S. Soni, "UDP Return Path for Packet Loss and Delay Measurement for MPLS Networks", RFC 7876, DOI 10.17487/RFC7876, July 2016, <https://www.rfc-editor.org/info/rfc7876>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

13.2. 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, DOI 10.17487/RFC4656, September 2006, <https://www.rfc-editor.org/info/rfc4656>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, DOI 10.17487/RFC5357, October 2008, <https://www.rfc-editor.org/info/rfc5357>.

[RFC5481] Morton, A. and B. Claise, "Packet Delay Variation Applicability Statement", RFC 5481, DOI 10.17487/RFC5481, March 2009, <https://www.rfc-editor.org/info/rfc5481>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7679] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Delay Metric for IP Performance Metrics (IPPM)", STD 81, RFC 7679, DOI 10.17487/RFC7679, January 2016, <https://www.rfc-editor.org/info/rfc7679>.

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <https://www.rfc-editor.org/info/rfc7471>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

[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>.

[RFC8570] Ginsberg, L., Ed., Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 8570, DOI 10.17487/RFC8570, March 2019, <https://www.rfc-editor.org/info/rfc8570>.

[RFC8571] Ginsberg, L., Ed., Previdi, S., Wu, Q., Tantsura, J., and C. Filsfils, "BGP - Link State (BGP-LS) Advertisement of IGP Traffic Engineering Performance Metric Extensions", RFC 8571, DOI 10.17487/RFC8571, March 2019, <https://www.rfc-editor.org/info/rfc8571>.

[RFC8668] Ginsberg, L., Ed., Bashandy, A., Filsfils, C., Nanduri, M., and E. Aries, "Advertising Layer 2 Bundle Member Link Attributes in IS-IS", RFC 8668, DOI 10.17487/RFC8668, December 2019, <https://www.rfc-editor.org/info/rfc8668>.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-07 (work in progress), May 2020.

[I-D.voyer-pim-sr-p2mp-policy]
Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. Zhang, "Segment Routing Point-to-Multipoint Policy", draft-voyer-pim-sr-p2mp-policy-01 (work in progress), April 2020.

[I-D.voyer-spring-sr-replication-segment]
Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. Zhang, "SR Replication Segment for Multi-point Service Delivery", draft-voyer-spring-sr-replication-segment-03 (work in progress), June 2020.

[I-D.shen-spring-p2mp-transport-chain]
Shen, Y., Zhang, Z., Parekh, R., Bidgoli, H., and Y. Kamite, "Point-to-Multipoint Transport Using Chain Replication in Segment Routing", draft-shen-spring-p2mp-transport-chain-02 (work in progress), April 2020.

[I-D.ietf-pce-binding-label-sid]
Filsfils, C., Sivabalan, S., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-ietf-pce-binding-label-sid-03 (work in progress), June 2020.

[I-D.ietf-spring-mpls-path-segment]
Cheng, W., Li, H., Chen, M., Gandhi, R., and R. Zigler, "Path Segment in MPLS Based Segment Routing Network", draft-ietf-spring-mpls-path-segment-02 (work in progress), February 2020.

[I-D.gandhi-mpls-ioam-sr]
Gandhi, R., Ali, Z., Filsfils, C., Brockners, F., Wen, B., and V. Kozak, "MPLS Data Plane Encapsulation for In-situ OAM Data", draft-gandhi-mpls-ioam-sr-02 (work in progress), March 2020.

[I-D.ketant-lsr-ospf-l2bundles]
Talaulikar, K. and P. Psenak, "Advertising L2 Bundle Member Link Attributes in OSPF", draft-ketant-lsr-ospf-l2bundles-01 (work in progress), January 2020.

[I-D.ietf-pce-sr-bidir-path]
Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong, "PCEP Extensions for Associated Bidirectional Segment Routing (SR) Paths", draft-ietf-pce-sr-bidir-path-02 (work in progress), March 2020.

Acknowledgments

The authors would like to thank Thierry Couture for the discussions on the use-cases for the performance measurement in segment routing networks. Authors would like to thank Patrick Khordoc for implementing the mechanisms defined in this document. 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. Thanks to Huaimo Chen, Yimin Shen, and Xufeng Liu for MPLS-RT expert review.

Contributors

Sagar Soni
Cisco Systems, Inc.
Email: sagsoni@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