Workgroup: SPRING Working Group
Internet-Draft:
draft-gandhi-spring-sr-enhanced-plm-05
Published: 9 August 2021
Intended Status: Standards Track
Expires: 10 February 2022

Authors: R. Gandhi, Ed. C. Filsfils
Cisco Systems, Inc. Cisco Systems, Inc.
N. Vaghamshi M. Nagarajah R. Foote
Reliance Telstra Nokia

Enhanced Performance and Liveness Monitoring in Segment Routing Networks

Abstract

Segment Routing (SR) leverages the source routing paradigm. SR is applicable to both Multiprotocol Label Switching (SR-MPLS) and IPv6 (SRv6) data planes. This document defines procedures for Enhanced Performance and Liveness Monitoring of end-to-end SR paths including SR Policies for both SR-MPLS and SRv6 data planes, those reduce the deployment and operational complexities in a network.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

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 [RFC8402]. 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 Performance Measurement (PM) for delay and packet loss as well as Liveness Monitoring for Connectivity Verification

(CV) are essential requirements to provide Service Level Agreements (SLAs) in SR networks.

The Simple Two-Way Active Measurement Protocol (STAMP) provides capabilities for the measurement of various performance metrics in IP networks [RFC8762] without the use of a control channel to pre-signal session parameters. The STAMP can be used for performance measurement for delay and packet loss of SR paths. However, the STAMP requires protocol support on the Session-Reflector to process the STAMP test packets as packets are punted from the forwarding fast path on the Session-Reflector and STAMP reply test packets are generated. This limits the scale for number of STAMP test sessions and faster connectivity loss detection interval.

The Seamless Bidirectional Forwarding Detection (S-BFD) [RFC7880] can be used for connectivity verification of SR paths. However, S-BFD requires protocol support on the BFD-Reflector to process the S-BFD packets as packets need to be punted from the forwarding fast path and reply packets are generated. This limits the scale for number S-BFD sessions and faster connectivity loss detection interval.

Enabling multiple protocols, S-BFD for connectivity verification and STAMP for performance measurement increases the deployment and operational complexities in a network. Also, implementing multiple protocols in a hardware significantly increases the development cost.

This document defines procedures for Enhanced Performance and Liveness Monitoring of end-to-end SR paths including SR Policies for both SR-MPLS and SRv6 data planes, those reduce the deployment and operational complexities in a network. The procedures use the Performance and Liveness Monitoring (PLM) test packet formats defined in this document. The test packets have the transmit and receive timestamps at the same locations as the Session-Reflector STAMP test packets to leverage the existing hardware support for it.

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

S-BFD: Seamless Bidirectional Forwarding Detection.

BSID: Binding Segment ID.

ECMP: Equal Cost Multi-Path.

EB: Endpoint Behaviour.

HMAC: Hashed Message Authentication Code.

MBZ: Must be Zero.

MPLS: Multiprotocol Label Switching.

PLM: Performance and Liveness Monitoring.

PM: Performance Measurement.

PTP: Precision Time Protocol.

SID: Segment ID.

SL: Segment List.

SR: Segment Routing.

SRH: Segment Routing Header.

SR-MPLS: Segment Routing with MPLS data plane.

SRv6: Segment Routing with IPv6 data plane.

SSID: Sender Session Identifier.

STAMP: Simple Two-way Active Measurement Protocol.

TC: Traffic Class.

TTL: Time To Live.

2.3. Reference Topology

In the reference topology shown below, the Session-Sender S1 initiates a PLM test packet and the Session-Reflector R1 returns the PLM test packet. The PLM return test packet is transmitted back to the Session-Sender S1 on the same path (same set of links and nodes) or a different path in the reverse direction from the path taken towards the Session-Reflector R1.

The Session-Sender S1 and Session-Reflector R1 are connected via an SR path [RFC8402]. The SR path may be an SR Policy [I-D.ietf-spring-segment-routing-policy] on node S1 (called head-end) with destination to node R1 (called tail-end).

```
                       T1
                      /
             +-------+    PLM Test Packet    +-------+
             |       | - - - - - - - - - - - |       |
             |   S1  |======================||   R1  |
             |       |<- - - - - - - - - - - |       |
             +-------+   Return Test Packet  +-------+
                      \
                       T4

            Session-Sender              Session-Reflector
                                        (Simply Forward)

                        Reference Topology
```

3. Overview

3.1. Loopback Mode

In loopback mode, the Session-Sender S1 initiates PLM test packets and the Session-Reflector R1 forwards them just like data packets for the regular traffic back to the Session-Sender S1. In this mode, the received PLM test packets are not punted out of the fast path in forwarding at the Session-Reflector. In other words, the Session-Reflector does not process them and generate Session-Reflector test packets. At the Session-Reflector, the loopback function simply makes the necessary changes to the encapsulation including IP and UDP headers to return the test packet to the Session-Sender. The Session-Reflector MUST NOT drop the loopback PLM test packets, for example, due to a local policy provisioned. No PLM test session is created on the Session-Reflector.

The Session-Sender MUST set the Destination UDP port to the UDP port it is configured to receive the PLM reply test packets. The UDP ports in the PLM test packets MUST follow the procedure defined in [RFC8762].

The IPv4 Time To Live (TTL) or IPv6 Hop Limit (HL) in the PLM test packet is set to 255.

3.2. Loopback Mode Enabled with Network Programming Function

In loopback mode enabled with network programming function, both transmit (T1) and receive (T2) timestamps in data plane are collected by the PLM test packets transmitted in loopback mode as shown in Figure 1. The network programming function optimizes the "operations of punt and generate the PLM test packet" on the Session-Reflector as timestamping is implemented in forwarding fast path in hardware. This helps to achieve higher PLM test session scale and faster connectivity loss detection.

```
                         T1                  T2
                        /                     \
               +-------+    PLM Test Packet    +-------+
               |       | - - - - - - - - - - - |       |
               |   S1  |======================||  R1   |
               |       |<- - - - - - - - - - - |       |
               +-------+   Return Test Packet  +-------+
                        \
                         T4

             Session-Sender                 Session-Reflector
                                             (Timestamp,
                                              Pop and Forward)
```

Figure 1: Loopback Mode Enabled with Network Programming Function

The Session-Sender adds transmit timestamp (T1) in the payload of the PLM test packet and clears the receive (T2) timestamp. The Session-Reflector adds the receive timestamp (T2) in the payload of the received PLM test packet in forwarding fast path in hardware without punting the test packet (e.g. to slow path or control-plane). The network programming function enables Session-Reflector to add the receive timestamp (T2) at a specific offset in the payload which is locally provisioned consistently in the network. The payload of the PLM test packet is not modified by the intermediate nodes.

The Session-Reflector only adds the receive timestamp if the source IP address (in case of SR-MPLS) or destination IP address (in case of SRv6) in the PLM test packet matches the local node address to ensure that the PLM test packet reaches the intended Session-Reflector and the receive timestamp is returned by the intended Session-Reflector.

3.3. Example Provisioning Model

An example provisioning model and typical measurement parameters are shown in Figure 2:

```
                               +------------+
                               | Controller |
                               +------------+
 PLM Mode                          /    \     Timestamp Label/SRV6 EB
   Loopback or Enhanced Mode      /      \      Timestamp Offset
 Timestamp Label/SRv6 EB         /        \     Timestamp Format
   Timestamp Format             /          \
 Missed Packet Count (N)       /            \
 Delay Threshold/Count (T/M)  /              \
 Packet Loss Threshold (XofY)/                \
                            v                  v
                        +-------+          +-------+
                        |       |          |       |
                        |   S1  |==========|   R1  |
                        |       |          |       |
                        +-------+          +-------+

                     Session-Sender    Session-Reflector
```

Figure 2: Example Provisioning Model

Example of a PLM mode is loopback mode and enhanced loopback mode. The values for Timestamp Label and SRv6 Endpoint Behaviour may be provisioned as described in Section 6. Example of Timestamp Format is 64-bit PTPv2 [IEEE1588]. Example of Timestamp Offset is 16 and 32 bytes for the PLM test packet formats defined in this document. Example threshold values configured for generating notifications are: Missed Packet Count (N), Delay Exceeded Threshold and Packet Count (T/M) and Packet Loss Threshold (XofY), as described in this document.

The mechanisms to provision the Session-Sender and Session-Reflector are outside the scope of this document.

4. PLM Test Packet Formats

The PLM test packet formats for unauthenticated and authenticated modes are defined in this document as shown in Figure 3. They have the transmit and receive timestamps at the same locations as the Session-Reflector STAMP test packets to leverage the existing hardware support for it.

```
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                     Sequence Number                           |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                     Transmit Timestamp (T1)                   |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Transmit Error Estimate      |  SSID                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                     Receive Timestamp (T2)                    |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      |                     MBZ (12 Octets)                           |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Receive Error Estimate       |  MBZ                          |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                     MBZ (4 Octets)                            |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

              PLM Test Packet Format in Unauthenticated Mode


      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
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                        Sequence Number                        |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                        MBZ (12 octets)                        |
      |                                                               |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                        Transmit Timestamp (T1)                |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Transmit Error Estimate      |  SSID                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                        MBZ (4 octets)                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                        Receive Timestamp (T2)                 |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      |                        MBZ (32 octets)                        |
      |                                                               |
      |                                                               |
      |                                                               |
      |                                                               |
      |                                                               |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  Receive Error Estimate       |                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               +
      |                        MBZ (6 octets)                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      |                        MBZ (16 octets)                        |
      |                                                               |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      |                        HMAC (16 octets)                       |
      |                                                               |
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

PLM Test Packet Format in Authenticated Mode

Figure 3: PLM Test Packet Formats

Sequence Number is the sequence number of the PLM test packet according to its transmit order. It starts with zero and is incremented by one for each subsequent PLM test packet.

SSID (16-bits): PLM Sender Session Identifier. Uses the procedure for SSID defined in [RFC8762].

Transmit Timestamp and Transmit Error Estimate are the Session-Sender's transmit timestamp and error estimate for the PLM test packet, respectively.

Receive Timestamp and Receive Error Estimate are the Session-Reflector's receive timestamp and error estimate, respectively.

The timestamp and error estimate fields follow the definition and formats defined in Section 4.1.2 in [RFC8762]. The timestamp format used by default is 64-bit PTPv2 [IEEE1588].

HMAC: The use of the HMAC field is described in Section 4.4 of [RFC8762].

MBZ: Must be Zero. It MUST be all zeroed on the transmission and MUST be ignored on receipt.

5. PLM Procedure

For performance and liveness monitoring of an end-to-end SR path including SR Policy, PLM test packets are transmitted in loopback mode.

For SR Policy, the PLM test packets are transmitted using the Segment List (SL) of the Candidate-Path [I-D.ietf-spring-segment-routing-policy]. When a Candidate-Path has more than one Segment Lists, multiple PLM test packets MUST be transmitted, one using each Segment List.

5.1. PLM for SR-MPLS Policies

The PLM test packets MUST be transmitted using the MPLS header for each Label Stack of the SR-MPLS Policy Candidate-Path(s) as shown in Figure 4.

In case of SR-MPLS paths, the SR-MPLS header can contain the MPLS label stack of the forward path or both forward and the reverse paths. In the first case, the PLM return test packets are received by the Session-Sender via IP/UDP [RFC0768] return path and the MPLS header is removed by the Session-Reflector.

In the second case, the Segment List of the reverse SR path is added in the PLM test packet header to receive the return test packet on a specific path, either using the Binding SID [I-D.ietf-pce-binding-label-sid] or Segment List of the Reverse SR Policy [I-D.ietf-pce-sr-bidir-path]. In this case, the MPLS header is not removed by the Session-Reflector.

In both cases, the Session-Sender MUST set the Destination Address equal to the Session-Sender address and the Source Address equal to the Session-Reflector address in the IP header of the test packets.

```
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      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IP Header                                                     |
.  Source IP Address = Session-Reflector IPv4 or IPv6 Address   .
.  Destination IP Address = Session-Sender IPv4 or IPv6 Address .
.  Protocol = UDP                                               .
.                                                               .
+---------------------------------------------------------------+
| UDP Header                                                    |
.  Source Port = As chosen by Session-Sender                    .
.  Destination Port = As chosen by Session-Sender               .
.                                                               .
+---------------------------------------------------------------+
|  Payload as defined in Figure 3                               |
.                                                               .
+---------------------------------------------------------------+
```

Figure 4: Example PLM Test Packet for SR-MPLS

5.2. PLM for SRv6 Policies

The PLM test packets for SRv6 data plane MUST be transmitted using the Segment Routing Header (SRH) [RFC8754] for each Segment List of the SRv6 Policy Candidate-Path(s) as shown in Figure 5.

In case of SRv6 paths, the SRH can contain the Segment List of the forward path or both forward and the reverse paths. In the first case, an inner IPv6 header (after SRH and before UDP header) MUST be added that contains the Destination Address equal to the Session-Sender address and the Source Address equal to the Session-Reflector address as shown in Figure 5. In this case, the SRH is removed by the Session-Reflector and IP/UDP return path is used.

In the second case, the Segment List of the reverse SR path is added in the SRH to receive the return test packet on a specific path, either using the Binding SID [I-D.ietf-pce-binding-label-sid] or Segment List of the Reverse SR Policy [I-D.ietf-pce-sr-bidir-path]. In this case, the SRH is not removed by the Session-Reflector and an inner IPv6 header is not required. When the PLM return test packet contains an SRH at the Session-Sender, the procedure defined for upper-layer header processing for SRv6 SIDs in [RFC8986] MUST be used to process the UDP header in the received PLM test packets.

```
+-----------------------------------------------------------------+
| IP Header                                                       |
.  Source IP Address = Session-Sender IPv6 Address                .
.  Destination IP Address = Destination IPv6 Address              .
.                                                                 .
+-----------------------------------------------------------------+
| SRH as specified in RFC 8754                                    |
.     <Segment List>                                              .
.                                                                 .
+-----------------------------------------------------------------+
| IP Header                                                       |
.  Source IP Address = Session-Reflector IPv6 Address             .
.  Destination IP Address = Session-Sender IPv6 Address           .
.                                                                 .
+-----------------------------------------------------------------+
| UDP Header                                                      |
.  Source Port = As chosen by Session-Sender                      .
.  Destination Port = As chosen by Session-Sender                 .
.                                                                 .
+-----------------------------------------------------------------+
|  Payload as defined in Figure 3                                 |
.                                                                 .
+-----------------------------------------------------------------+
```

Figure 5: Example PLM Test Packet for SRv6

6. Enhanced PLM Procedure

For enhanced performance and liveness monitoring of an end-to-end SR path including SR Policy, PLM test packets are transmitted in loopback mode enabled with network programming function to timestamp and forward the packet.

6.1. Enhanced PLM with Timestamp Label for SR-MPLS Policies

In this document, two new Timestamp Labels are defined for SR-MPLS data plane to enable network programming function for "timestamp, pop and forward" the received test packet, one for unauthenticated mode and one for authenticated mode.

In the PLM test packets for SR-MPLS Policies, a Timestamp Label is added in the MPLS header as shown in Figure 6, to collect "Receive Timestamp" field in the payload of the PLM test packet. The Label Stack for the reverse SR-MPLS path can be added after the Timestamp Label to receive the PLM return test packet on a specific path. When a Session-Reflector receives a packet with Timestamp Label, after timestamping the packet at a specific offset, the Session-Reflector pops the Timestamp Label and forwards the packet using the next label or IP header in the packet (just like the data packets for the regular traffic).

```
  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      |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |            Extension Label (15)       | TC  |S|      TTL      |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |       Timestamp Label (TBA1 or TBA2)  | TC  |S|      TTL      |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | IP Header                                                     |
  .  Source IP Address = Session-Reflector IPv4 or IPv6 Address   .
  .  Destination IP Address = Session-Sender IPv4 or IPv6 Address .
  .  Protocol = UDP                                               .
  .                                                               .
  +---------------------------------------------------------------+
  | UDP Header                                                    |
  .  Source Port = As chosen by Session-Sender                    .
  .  Destination Port = As chosen by Session-Sender               .
  .                                                               .
  +---------------------------------------------------------------+
  |  Payload as defined in Figure 3                               |
  .                                                               .
  +---------------------------------------------------------------+
```

Figure 6: Example PLM Test Packet with Timestamp Label for SR-MPLS

6.1.1. Timestamp Label Allocation

The timestamp Labels for unauthenticated and authenticated modes can be allocated using one of the following methods:

- Labels (values TBA1 and TBA2) assigned by IANA from the "Extended Special-Purpose MPLS Values" [RFC9017]. For Label (value TBA1), the timestamp offset is fixed at byte-offset 16 from the start of the payload for the unauthenticated mode, and Label (value TBA2) at byte-offset 32 from the start of the payload for the authenticated mode, both using the timestamp format 64-bit PTPv2.

- Labels allocated by a Controller from the global table of the Session-Reflector. The Controller provisions the labels on both Session-Sender and Session-Reflector, as well as timestamp offsets and timestamp formats.

- Labels allocated by the Session-Reflector. The signaling and IGP flooding extension for the labels (including timestamp offsets and timestamp formats) are outside the scope of this document.

6.1.2. Node Capability for Timestamp Label

The PLM Session-Sender needs to know if the Session-Reflector can process the Timestamp Label to avoid dropping PLM test packets. The signaling extension for this capability exchange is outside the scope of this document.

6.2. Enhanced PLM with Timestamp Endpoint Function for SRv6 Policies

The [RFC8986] defines SRv6 Endpoint Behaviours (EB) for SRv6 nodes. In this document, two new Timestamp Endpoint Behaviours are defined for Segment Routing Header (SRH) [RFC8754] to enable "Timestamp and Forward (TSF)" function for the received test packets, one for unauthenticated mode and one for authenticated mode.

In the PLM test packets for SRv6 Policies, Timestamp Endpoint Function (End.TSF) is carried with the target Segment Identifier (SID) in SRH [RFC8754] as shown in Figure 7, to collect "Receive Timestamp" field in the payload of the PLM test packet. The Segment List for the reverse path can be added after the target SID to receive the PLM return test packet on a specific path. When a Session-Reflector receives a packet with Timestamp Endpoint (End.TSF) for the target SID which is local, after timestamping the packet at a specific offset, the Session-Reflector forwards the packet using the next SID in the SRH or inner IPv6 header in the packet (just like the data packets for the regular traffic).

```
+----------------------------------------------------------------+
| IP Header                                                      |
.  Source IP Address = Session-Sender IPv6 Address              .
.  Destination IP Address = Destination IPv6 Address            .
.                                                                .
+----------------------------------------------------------------+
| SRH as specified in RFC 8754                                   |
.     <Segment List>                                             .
.     SRv6 Endpoint End.TSF (value TBA3 or TBA4)                 .
.                                                                .
+----------------------------------------------------------------+
| IP Header                                                      |
.  Source IP Address = Session-Reflector IPv6 Address           .
.  Destination IP Address = Session-Sender IPv6 Address         .
.                                                                .
+----------------------------------------------------------------+
| UDP Header                                                     |
.  Source Port = As chosen by Session-Sender                     .
.  Destination Port = As chosen by Session-Sender                .
.                                                                .
+----------------------------------------------------------------+
|  Payload as defined in Figure 3                                |
.                                                                .
+----------------------------------------------------------------+
```

Figure 7: Example PLM Test Packet with Endpoint Function for SRv6

6.2.1. Timestamp Endpoint Function Assignment

The Timestamp Endpoint Functions for "Timestamp and Forward" can be signaled using one of the following methods:

- Timestamp Endpoint Functions (values TBA3 and TBA4) assigned by IANA from the "SRv6 Endpoint Behaviors Registry". For endpoint behaviour (value TBA3), the timestamp offset is fixed at byte-offset 16 from the start of the payload for the unauthenticated mode, and endpoint behaviour (value TBA4) at byte-offset 32 from the start of the payload for the authenticated mode, both using the timestamp format 64-bit PTPv2.

- Timestamp Endpoint Functions assigned by a Controller. The Controller provisions the values on both Session-Sender and Session-Reflector, as well as timestamp offsets and timestamp formats.

- Timestamp Endpoint Functions assigned by the Session-Reflector. The signaling and IGP flooding extension for the endpoint functions (including timestamp offsets and timestamp formats) are outside the scope of this document.

6.2.2. Node Capability for Timestamp Endpoint Function

The PLM Session-Sender needs to know if the Session-Reflector can process the Timestamp Endpoint Function to avoid dropping PLM test packets. The signaling extension for this capability exchange is outside the scope of this document.

7. ECMP Handling

An SR Policy can have ECMPs between the source and transit nodes, between transit nodes and between transit and destination nodes. The PLM test packets SHOULD be transmitted to traverse different ECMP paths to monitor an end-to-end SR Policy.

Forwarding plane has various hashing functions available to forward packets on specific ECMP paths.

For SR-MPLS Policy, sweeping of MPLS entropy label [RFC6790] values can be used in Session-Sender test packets and Session-Reflector test packets to take advantage of the hashing function in forwarding plane to influence the ECMP path taken by them.

In IPv4 header of the Session-Sender test packets, sweeping of Destination Address from the range 127/8 can be used to exercise ECMP paths. In this case, both the forward and the return paths MUST be SR-MPLS paths when using the loopback mode.

As specified in [RFC6437], Flow Label field in the outer IPv6 header can also be used for sweeping to exercise different IPv6 ECMP paths.

8. Example PLM Failure Notifications

In both loopback modes, the timestamps T1 and T4 are used to measure the round-trip delay. In loopback mode enabled with network programming function, the timestamps T1 and T2 are used to measure the one-way delay. The delay metrics for an end-to-end SR path are notified, for example, when consecutive M number of PLM test packets have measured delay values exceed user-configured threshold T, where M (Delay Exceeded Packet Count) and T (Absolute and Percentage Delay Exceeded Threshold) are also locally provisioned values.

The round-trip packet loss for an end-to-end SR path is calculated using the Sequence Number in the PLM test packets. The packet loss metric is notified when X number of PLM test packets were lost out of last Y number of PLM test packets transmitted by the Session-Sender, where Threshold XofY is locally provisioned value.

Connectivity verification success for an end-to-end SR path is initially notified as soon as one or more PLM return test packets are received at the Session-Sender.

Connectivity verification failure for an end-to-end SR path is notified when consecutive N number of PLM return test packets are not received at the Session-Sender, where N (Missed PLM Packet Count) is a locally provisioned value.

In both loopback modes, a connectivity verification failure on the reverse direction path can cause the PLM return test packets to not reach the Session-Sender. This is also true in the case where the return test packets are generated by the stateless Session-Reflector. The stateful Session-Reflector can solve this issue by maintaining the forwarding direction state and notifying a connectivity verification failure to the Session-Sender.

9. Security Considerations

The PLM protocol is intended for deployment in limited domains [RFC8799]. As such, it assumes that a node involved in the PLM protocol operation has previously verified the integrity of the path and the identity of the far-end Session-Reflector.

If desired, attacks can be mitigated by performing basic validation and sanity checks, at the Session-Sender, of the timestamp fields in received PLM reply test packets. The minimal state associated with these protocols also limits the extent of disruption that can be caused by a corrupt or invalid packet to a single test cycle.

The security considerations specified in [RFC8762] also apply to the procedures defined in this document. Specifically, the message integrity protection using HMAC, as defined in Section 4.4 of [RFC8762] also apply to the procedure described in this document.

10. IANA Considerations

IANA maintains the "Special-Purpose Multiprotocol Label Switching (MPLS) Label Values" registry (see <https://www.iana.org/assignments/mpls-label-values/mpls-label-values.xml>). IANA is requested to allocate Timestamp Label value from the "Extended Special-Purpose MPLS Label Values" registry:

Value	Description	Reference
TBA1	Timestamp Label for offset 16 for Unauthenticated Mode	This document
TBA2	Timestamp Label for offset 32 for Authenticated Mode	This document

IANA is requested to allocate, within the "SRv6 Endpoint Behaviors Registry" sub-registry belonging to the top-level "Segment Routing Parameters" registry [RFC8986], the following allocation:

Value	Endpoint Behavior	Reference
TBA3	End.TSF (Timestamp and Forward) for offset 16 for Unauthenticated Mode	This document
TBA4	End.TSF (Timestamp and Forward) for offset 32 for Authenticated Mode	This document

11. References

11.1. Normative References

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, DOI 10.17487/RFC0768, August 1980, <https://www.rfc-editor.org/info/rfc768>.

[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>.

[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>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

11.2. Informative References

[IEEE1588] IEEE, "1588-2008 IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", March 2008.

[RFC6437]
Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/info/rfc6437>.

[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>.

[RFC7880] Pignataro, C., Ward, D., Akiya, N., Bhatia, M., and S. Pallagatti, "Seamless Bidirectional Forwarding Detection (S-BFD)", RFC 7880, DOI 10.17487/RFC7880, July 2016, <https://www.rfc-editor.org/info/rfc7880>.

[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>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

[RFC9017] Andersson, L., Kompella, K., and A. Farrel, "Special-Purpose Label Terminology", RFC 9017, DOI 10.17487/RFC9017, April 2021, <https://www.rfc-editor.org/info/rfc9017>.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-13, 28 May 2021, <https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-13.txt>.

[I-D.ietf-pce-binding-label-sid]
Sivabalan, S., Filsfils, C., Tantsura, J., Previdi, S., and C. L. (editor), "Carrying Binding Label/Segment Identifier in PCE-based Networks.", Work in Progress, Internet-Draft, draft-ietf-pce-binding-label-sid-10, 10 July 2021, <https://www.ietf.org/archive/id/draft-ietf-pce-binding-label-sid-10.txt>.

[I-D.ietf-pce-sr-bidir-path]
Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong, "Path Computation Element Communication Protocol (PCEP) Extensions for Associated Bidirectional Segment Routing (SR) Paths", Work in Progress, Internet-Draft, draft-ietf-pce-sr-bidir-path-07, 12 July 2021, <https://www.ietf.org/archive/id/draft-ietf-pce-sr-bidir-path-07.txt>.

Acknowledgments

The authors would like to thank Greg Mirsky, Mach Chen, Kireeti Kompella, and Adrian Farrel for providing the review comments.

Authors' Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.

Email: cfilsfil@cisco.com

Navin Vaghamshi
Reliance

Email: Navin.Vaghamshi@ril.com

Moses Nagarajah
Telstra

Email: Moses.Nagarajah@team.telstra.com

Richard Foote
Nokia

Email: footer.foote@nokia.com