

SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2021

R. Gandhi, Ed.
C. Filsfils
Cisco Systems, Inc.
N. Vaghamshi
Reliance
M. Nagarajah
Telstra
R. Foote
Nokia
July 11, 2020

**Enhanced Performance Delay and Liveness Monitoring in Segment Routing
Networks
draft-gandhi-spring-sr-enhanced-plm-02**

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 procedure for Enhanced Performance Delay and Liveness Monitoring (PDLM) in Segment Routing networks. The procedure uses the probe messages defined in [RFC 5357](#) (Two-Way Active Measurement Protocol (TWAMP) Light) and [RFC 8762](#) (Simple Two-Way Active Measurement Protocol (STAMP)) for end-to-end SR Paths including SR Policies with both SR-MPLS and SRv6 data planes.

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 <https://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."

This Internet-Draft will expire on January 12, 2021.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1.	Introduction	3
2.	Conventions Used in This Document	4
2.1.	Requirements Language	4
2.2.	Abbreviations	4
2.3.	Reference Topology	5
2.4.	Loopback Mode	5
3.	Probe Messages	5
3.1.	Example Provisioning Model	6
4.	Performance Delay and Liveness Monitoring	7
4.1.	Probe Message for SR-MPLS	7
4.2.	Probe Message for SRv6	8
5.	Enhanced Performance Delay and Liveness Monitoring	9
5.1.	Loopback Mode Enabled with Network Programming	9
5.2.	Probe Message with Network Programming for SR-MPLS	10
5.2.1.	Node Capability for Timestamp Label	11
5.2.2.	Timestamp Label Allocation	11
5.3.	Probe Message with Network Programming for SRv6	12
6.	ECMP Handling	13
7.	Failure Notification	13
8.	Security Considerations	14
9.	IANA Considerations	14
10.	References	15
10.1.	Normative References	15
10.2.	Informative References	15
	Acknowledgments	17
	Authors' Addresses	17

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 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 Liveness Monitoring for detecting faults as well as Performance Delay Measurement (DM) and Loss Measurement (LM) are essential requirements to provide Service Level Agreements (SLAs) in SR networks.

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 using probe messages. The TWAMP Light [Appendix I in [RFC5357](#)] and the Simple Two-way Active Measurement Protocol (STAMP) [[RFC8762](#)] provide simplified mechanisms for active performance measurement in IP networks, alleviating the need for control-channel signaling by using configuration data model to provision a test-channel.

[[I-D.gandhi-spring-twamp-srpm](#)] defines procedure for performance measurement using TWAMP Light messages with user-defined IP/UDP paths in SR networks. [[I-D.gandhi-spring-stamp-srpm](#)] defines similar procedure using STAMP messages in SR networks. The procedure for one-way and two-way modes defined for delay measurement can also be applied to liveness monitoring of SR Paths. However, it limits the scale for number of PM sessions and fault detection interval since the probe query messages need to be punted from the forwarding path (to slow path or control plane) and response messages need to be injected.

For Liveness Monitoring, Seamless Bidirectional Forwarding Detection (S-BFD) [[RFC7880](#)] can be used in Segment Routing networks. However, S-BFD requires protocol support on the reflector node to process the S-BFD packets as packets need to be punted from the forwarding path in order to send the reply thereby limiting the scale for number of PM sessions and fault detection interval. In addition, S-BFD protocol does not have the capability today to enable performance delay monitoring in SR networks. Enabling multiple protocols in SR networks, S-BFD for liveness monitoring and TWAMP Light or STAMP for performance delay monitoring increases the deployment and operational complexities in SR networks.

This document defines procedure for Enhanced Performance Delay and Liveness Monitoring (PDLM) in Segment Routing networks. The procedure uses the probe messages defined in [[RFC5357](#)] (TWAMP Light) and [[RFC8762](#)] (STAMP) for end-to-end SR Paths including SR Policies with both SR-MPLS and SRv6 data planes.

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

BFD: Bidirectional Forwarding Detection.

BSID: Binding Segment ID.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.

OWAMP: One-Way Active Measurement Protocol.

PDLM: Performance Delay 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.

STAMP: Simple Two-way Active Measurement Protocol.

TWAMP: Two-Way Active Measurement Protocol.

2.3. Reference Topology

In the reference topology shown below, the nodes R1 and R5 are connected via Point-to-Point (P2P) SR Path such as SR Policy [[I-D.ietf-spring-segment-routing-policy](#)] originating on node R1 with endpoint on node R5.

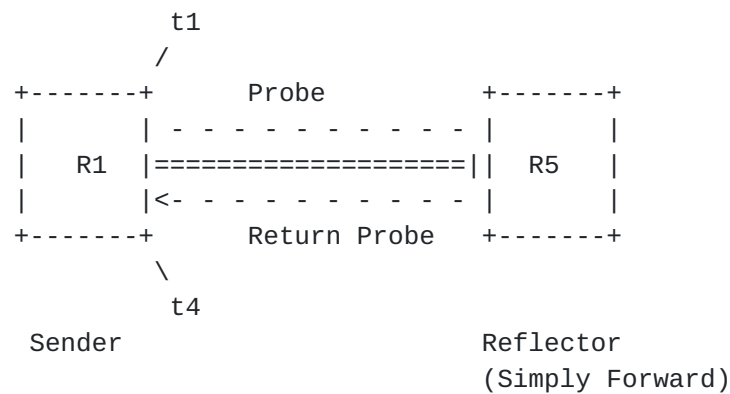


Figure 1: Reference Topology

2.4. Loopback Mode

In loopback mode, the sender node R1 initiates probe messages and the reflector node R5 forwards them back to the sender node R1 just like data packets for the normal traffic. The probe messages are not punted at the reflector node and it does not process them and generate response messages. The reflector node must not drop the loopback probe messages, for example, due to a local policy provisioned on the node.

3. Probe Messages

The TWAMP Light probe messages for delay measurement as defined in [[RFC5357](#)] or STAMP probe messages as defined in [[RFC8762](#)] are sent by the sender node R1 towards the reflector node R5 in loopback mode as shown in Figure 1. The probe messages are sent by the sender node on the congruent path of the data traffic flowing on the SR Path.

Both Source and Destination UDP ports in the probe messages are allocated dynamically or user-configured from the range specified in [[RFC8762](#)] and are different than the ports used for TWAMP Light and STAMP sessions. The Source and Destination IP addresses in the probe messages are set to the reflector and the sender node addresses,

respectively (representing the reverse path). The IPv4 Time To Live (TTL) and IPv6 Hop Limit (HL) are set to 255.

No PM session is created on the reflector node R5. As the probe message is not punted on the reflector node for processing, the Sender copies the 'Sequence Number' in 'Session-Sender Sequence Number' field directly. Also, the Sender Timestamp, Sender Error Estimate and Sender TTL fields [RFC5357] [RFC8762] in the probe message are not used. The rest of the fields are set as defined in [RFC5357] [RFC8762]

Timestamp format preferred is 64-bit PTPv2 [IEEE1588] as specified in [RFC8186], implemented in hardware. The NTP timestamp format MUST be supported [RFC5357], however, since PTPv2 is widely used, it SHOULD also be supported. In addition to adding the timestamp in the message, the "Error Estimate" field in the payload of the message can be updated using the procedure defined in [RFC4656].

3.1. Example Provisioning Model

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

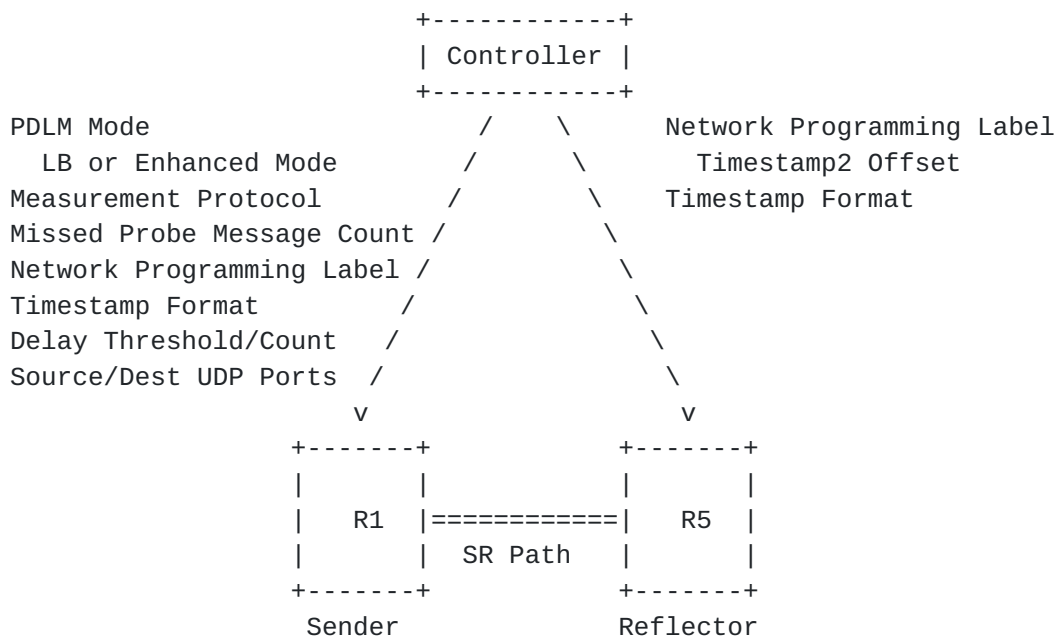


Figure 2: Example Provisioning Model

Example of Measurement Protocol is TWAMP Light and STAMP, example of Timestamp Format is 64-bit PTPv2 [IEEE1588] and NTP, etc.

The mechanisms to provision the sender and reflector nodes are outside the scope of this document.

4. Performance Delay and Liveness Monitoring

For performance delay and liveness monitoring of an end-to-end SR Path including SR Policy, PM probes in loopback mode is used. The PM probe messages are sent by the sender (head-end) node R1 to the reflector (endpoint) node R5 of the SR Policy as shown in Figure 1.

The probe messages are sent using the Segment List (SL) of the Candidate-paths of the SR Policy [[I-D.ietf-spring-segment-routing-policy](#)]. When a Candidate-path has more than one Segment Lists, multiple probe messages are sent, one using each Segment List. The return probe messages are received by the sender node via IP/UDP [[RFC0768](#)] return path by default. The Segment List of the return SR path can be added in the probe message header to receive the return probe message on a specific path using the mechanisms defined in [[I-D.ietf-pce-binding-label-sid](#)] and [[I-D.ietf-pce-sr-bidir-path](#)].

4.1. Probe Message for SR-MPLS

The TWAMP Light or STAMP probe messages for SR-MPLS data plane are sent using the MPLS header containing the label stack of the SR Policy as shown in Figure 3. In case of IP/UDP return path, the MPLS header is removed by the reflector node. The label stack can contain a reverse SR-MPLS path to receive the return probe message on a specific path. In this case, the MPLS header will not be removed by the reflector node.

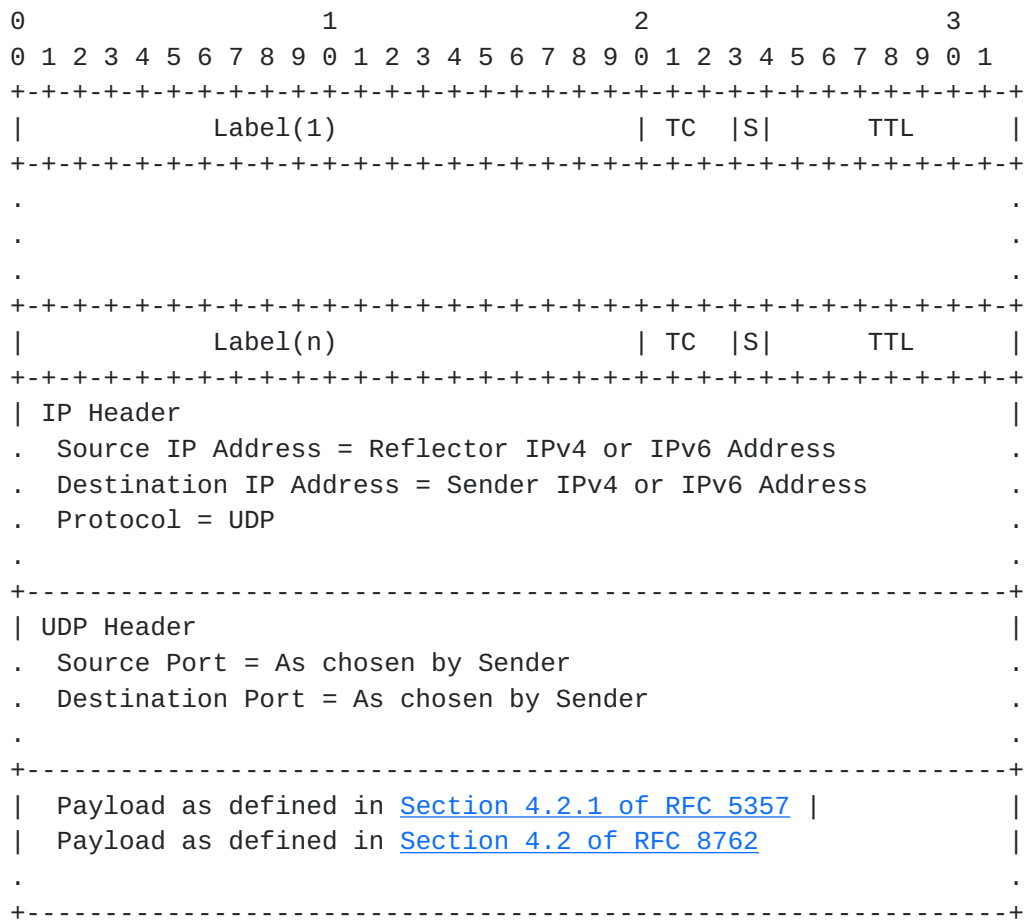


Figure 3: Example Probe Message for SR-MPLS

4.2. Probe Message for SRv6

The TWAMP Light or STAMP probe messages for SRv6 data plane are sent using the Segment Routing Header (SRH) [[RFC8754](#)] containing the Segment List of the SR Policy as shown in Figure 4. In case of IP/UDP return path, the SRH is removed by the reflector node. The Segment List can contain a reverse SRv6 path to receive the return probe message on a specific path. In this case, the SRH will not be removed by the reflector node.



Figure 4: Example Probe Message for SRv6

5. Enhanced Performance Delay and Liveness Monitoring

The enhanced performance delay and liveness monitoring of an end-to-end SR Path including SR Policy is defined using the PM probes in "loopback mode enabled with network programming".

5.1. Loopback Mode Enabled with Network Programming

In "loopback mode enabled with network programming", both transmit (t1) and receive (t2) timestamps in data plane are collected by the probe messages sent in loopback mode as shown in Figure 5. The network programming function optimizes the "operations of punt, add receive timestamp and inject the probe packet" on the reflector node and it is implemented in hardware. The payload of the probe message is not modified by any intermediate nodes.

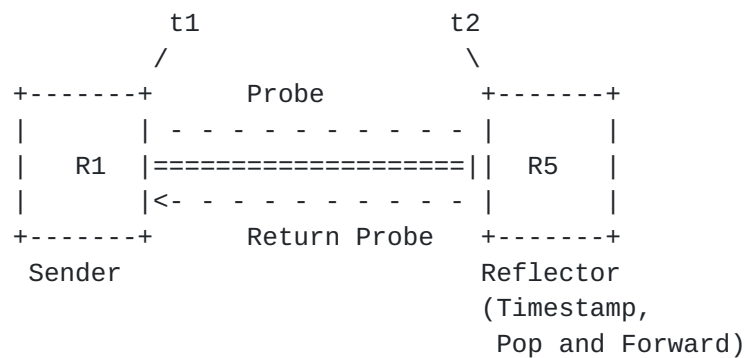


Figure 5: Loopback Mode Enabled with Network Programming

The sender node adds transmit (t1) timestamp in the payload of the TWAMP Light or STAMP probe message and clears the receive (t2) timestamp. The reflector node adds the receive timestamp in the payload of the received probe message without punting the message to slow-path (or control-plane). The reflector node only adds the receive timestamp if the source or destination address in the probe message matches the local node address to ensure that the receive timestamp is returned by the intended reflector node.

The network programming function enables the node to add receive timestamp in the payload of the probe message at a specific offset which is locally provisioned consistently in the network. In TWAMP Light message defined in [Section 4.2.1 of \[RFC5357\]](#) or STAMP message defined in [\[RFC8762\]](#) for delay measurement, the 64-bit receive timestamp is added at byte-offset 16 which is from the start of the payload.

5.2. Probe Message with Network Programming for SR-MPLS

In this document, new Timestamp Label (value TBD1) is defined for SR-MPLS data plane to enable network programming function for "timestamp, pop and forward" the received packet.

In the probe message for SR-MPLS, Timestamp Label is added in the MPLS header as shown in Figure 6, to collect "Receive Timestamp" field in the payload of the TWAMP Light [\[RFC5357\]](#) or STAMP probe message. The label stack for the reverse SR-MPLS path can be added after the Timestamp Label to receive the return probe message on a specific path. When a node receives a message with Timestamp Label, after timestamping the message at a specific offset, the node pops the Timestamp Label and forwards the message using the next label or IP header in the message (just like the data packets for the normal traffic).

- o Labels assigned by IANA with value TBA1 from the Extended Special-Purpose MPLS Values [[I-D.ietf-mpls-spl-terminology](#)].
- o Labels allocated by a Controller from the global table of the egress node. The Controller provisions the label on both ingress and egress nodes.
- o Labels allocated by the egress node. The signaling or IGP flooding extension for this is outside the scope of this document.

5.3. Probe Message with Network Programming for SRv6

In this document, new Endpoint function "Timestamp and Forward (TSF)" (value TBD2) is defined for Segment Routing Header (SRH) [[RFC8754](#)] for SRv6 data plane to enable network programming function for "timestamp and forward" the received message.

In the probe message for SRv6, END.TSF function is added for the Endpoint Segment Identifier (SID) in SRH [[RFC8754](#)] as shown in Figure 7, to collect "Receive Timestamp" field in the payload of the TWAMP Light [[RFC5357](#)] or STAMP probe message. When a node receives a packet with END.TSF function for the target SID which is local, after timestamping the packet at a specific offset, the node forwards the packet using the next SID or IP header in the packet (just like the packets for the normal traffic).

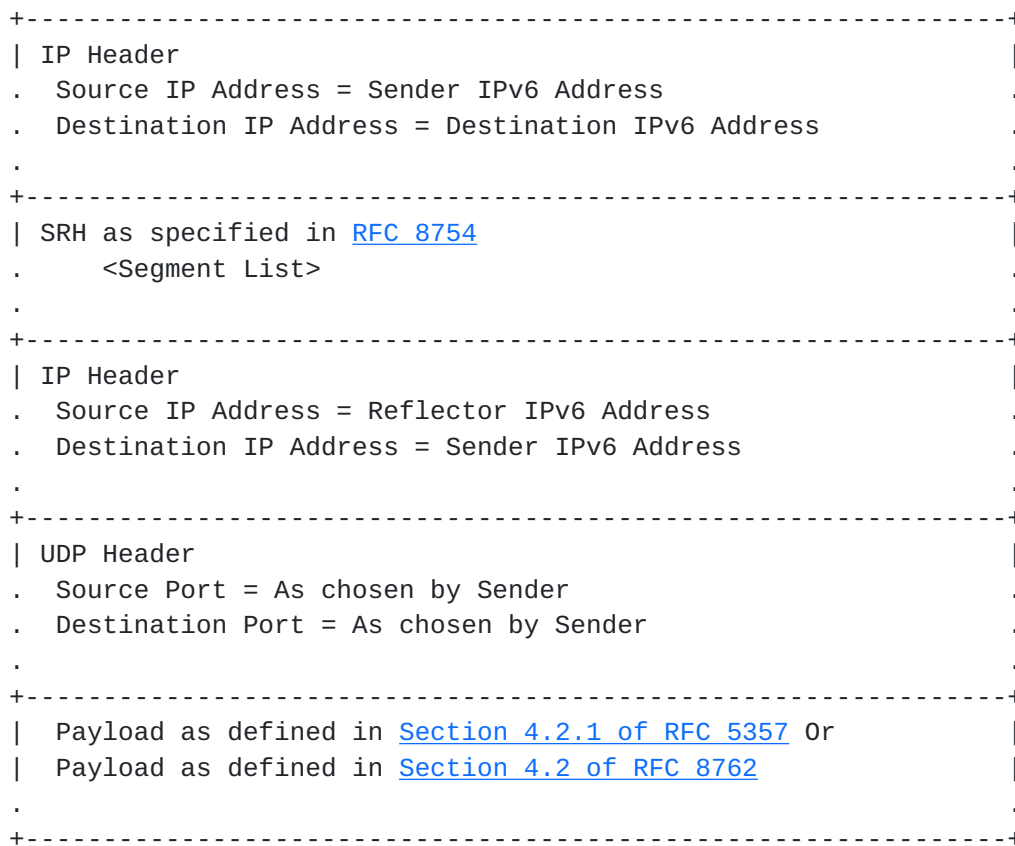


Figure 7: Example Probe Message with Endpoint Function for SRv6

6. ECMP Handling

An SR Policy can have ECMPs between the source and transit nodes, between transit nodes and between transit and destination nodes. The PM probe messages need to be sent to traverse different ECMP paths to monitor the liveness for an end-to-end SR Policy.

Forwarding plane has various hashing functions available to forward packets on specific ECMP paths. In IPv4 header of the PM probe messages, sweeping of Destination Address in 127/8 range can be used to exercise different ECMP paths in the loopback mode as long as the return path is also SR-MPLS. The Flow Label field in the outer IPv6 header can also be used for sweeping to exercise different ECMP paths.

7. Failure Notification

Liveness failure for SR Path is notified when consecutive N number of return probe messages are not received at the sender node, where N (Missed Probe Message Count) is locally provisioned value. Similarly, delay metrics are notified when consecutive M number of

probe messages have measured delay values exceed user-configured thresholds (absolute and percentage), where M is also locally provisioned value.

In loopback mode, the timestamps t1 and t4 are used to measure round-trip delay. In loopback mode enabled with network programming, the timestamps t1 and t2 are used to measure one-way delay.

8. Security Considerations

The Performance Delay and Liveness Monitoring is intended for deployment in the well-managed private and service provider networks. As such, it assumes that a node involved in a monitoring operation has previously verified the integrity of the path and the identity of the reflector node. If desired, attacks can be mitigated by performing basic validation and sanity checks, at the sender, of the timestamp fields in received probe messages. The minimal state associated with these protocols also limits the extent of disruption that can be caused by a corrupt or invalid message to a single probe cycle. Use of HMAC-SHA-256 in the authenticated mode protects the data integrity of the probe messages. Cryptographic measures may be enhanced by the correct configuration of access-control lists and firewalls.

9. 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	This document

IANA is requested to allocate, within the "SRv6 Endpoint Behaviors Registry" sub-registry belonging to the top-level "Segment-routing with IPv6 data plane (SRv6) Parameters" registry [[I-D.ietf-spring-srv6-network-programming](#)], the following allocation:

Value	Endpoint Behavior	Reference
TBA2	END.TSF (Timestamp and Forward)	This document

10. References

10.1. 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>>.
- [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>>.
- [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>>.

10.2. Informative References

- [IEEE1588] IEEE, "1588-2008 IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", March 2008.
- [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>>.
- [RFC8186] Mirsky, G. and I. Meilik, "Support of the IEEE 1588 Timestamp Format in a Two-Way Active Measurement Protocol (TWAMP)", [RFC 8186](#), DOI 10.17487/RFC8186, June 2017, <<https://www.rfc-editor.org/info/rfc8186>>.

- [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>>.
- [I-D.gandhi-spring-twamp-srpm]
Gandhi, R., Filsfils, C., Voyer, D., Chen, M., and B. Janssens, "Performance Measurement Using TWAMP Light for Segment Routing Networks", [draft-gandhi-spring-twamp-srpm-09](#) (work in progress), June 2020.
- [I-D.gandhi-spring-stamp-srpm]
Gandhi, R., Filsfils, C., Voyer, D., Chen, M., and B. Janssens, "Performance Measurement Using STAMP for Segment Routing Networks", [draft-gandhi-spring-stamp-srpm-01](#) (work in progress), June 2020.
- [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.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", [draft-ietf-spring-srv6-network-programming-16](#) (work in progress), June 2020.
- [I-D.ietf-mpls-spl-terminology]
Andersson, L., Kompella, K., and A. Farrel, "Special Purpose Label terminology", [draft-ietf-mpls-spl-terminology-02](#) (work in progress), May 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-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

TBD

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

