SPRING Working Group R. Gandhi, Ed.
Internet-Draft C. Filsfils
Intended status: Standards Track Cisco Systems, Inc.
Expires: September 24, 2020 D. Voyer
Bell Canada
M. Chen
Huawei
B. Janssens
Colt

March 23, 2020

Performance Measurement Using STAMP for Segment Routing Networks
draft-gandhi-spring-stamp-srpm-00

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 specifies procedure for sending and processing probe query and response messages for Performance Measurement (PM) in Segment Routing networks. The procedure uses the mechanisms defined in RFC 8762 (Simple Two-Way Active Measurement Protocol (STAMP)) for Delay Measurement, and uses the mechanisms defined in this document for Loss Measurement. The procedure specified is applicable to SR-MPLS and SRv6 data planes and is used for both Links and end-to-end SR Policies.

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 September 24, 2020.

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. 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).

The Simple Two-way Active Measurement Protocol (STAMP) provides capabilities for the measurement of various performance metrics in IP networks using probe messages [RFC8762]. It eliminates the control-channel signaling by using configuration data model to provision a test-channel (e.g. UDP paths). [I-D.ietf-ippm-stamp-option-tlv] defines TLV extensions for STAMP messages.

The STAMP message with a TLV for "direct measurement" can be used for combined Delay + Loss measurement [I-D.ietf-ippm-stamp-option-tlv]. However, in order to use only for loss measurement purpose, it requires the node to support the delay measurement messages and support timestamp for these messages (which may also require clock synchronization). Furthermore, for hardware-based counter collection for direct-mode loss meaurement, the optional TLV based processing adds unnecessary overhead (as counters are not at well-known locations).

This document specifies procedures for sending and processing probe query and response messages for Performance Measurement in SR networks. The procedure uses the mechanisms defined in [RFC8762] (STAMP) (including the TLV extensions) for Delay Measurement (DM), and uses the mechanisms defined in this document for Loss Measurement. The procedure specified is applicable to SR-MPLS and SRv6 data planes and are used for both Links and end-to-end SR Policies. This document also defines mechanisms for handling ECMPs

of SR Policies for performance delay measurement. Unless otherwise specified, the mechanisms defined in [RFC8762] and [I-D.ietf-ippm-stamp-option-tlv] are not modified by this document.

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

BSID: Binding Segment ID.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

HMAC: Hashed Message Authentication Code.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.

NTP: Network Time Protocol.

OWAMP: One-Way Active Measurement Protocol.

PM: Performance Measurement.

PSID: Path Segment Identifier.

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.

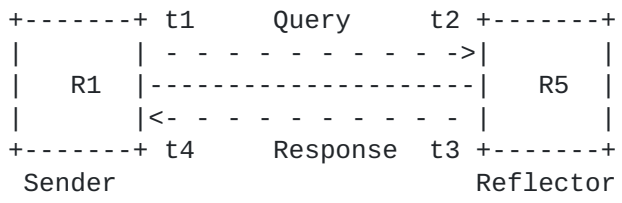
SRv6: Segment Routing with IPv6 data plane.

STAMP: Simple Two-way Active Measurement Protocol.

TC: Traffic Class.

2.3. Reference Topology

In the reference topology shown below, the sender node R1 initiates a probe query for performance measurement and the reflector node R5 sends a probe response for the query message received. The probe response is sent to the sender node R1. The nodes R1 and R5 may be directly connected via a Link or there exists a Point-to-Point (P2P) SR Policy [I-D.ietf-spring-segment-routing-policy] on node R1 with destination to node R5. In case of Point-to-Multipoint (P2MP), SR Policy originating from source node R1 may terminate on multiple destination leaf nodes [I-D.voyer-spring-sr-replication-segment].



Reference Topology

3. Overview

For one-way, two-way and round-trip delay measurements in Segment Routing networks, the probe messages defined in [RFC8762] are used. For direct-mode and inferred-mode loss measurements in Segment Routing networks, the messages defined in this document are used. Separate UDP destination port numbers are user-configured for delay and loss measurements from the range specified in [RFC8762]. The sender uses the UDP port number following the guidelines specified in Section 6 in [RFC6335]. The UDP destination port 862 can be used for delay measurement probes by default except for the case where the reflector needs to process STAMP TLVs. For both Links and end-to-end SR Policies, no PM session for delay or loss measurement is created on the reflector node R5 [RFC8762].

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

- For Delay Measurement, the probe messages are sent on the congruent path of the data traffic by the sender node, and are used to measure the delay experienced by the actual data traffic flowing on the Links and SR Policies.

- For Loss Measurement, the probe messages are sent on the congruent path of the data traffic by the sender 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 reflector node (incoming link or incoming SID needed since the reflector 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] and for SRv6 defined in [I-D.ali-spring-ioam-srv6] are used to carry PM information such as timestamp in-band as part of the data packets, and are outside the scope of this document.

3.1. Example Provisioning Model

An example of a provisioning model and typical measurement parameters for each user-configured destination UDP port for performance delay and loss measurements is shown in the following Figure 1:

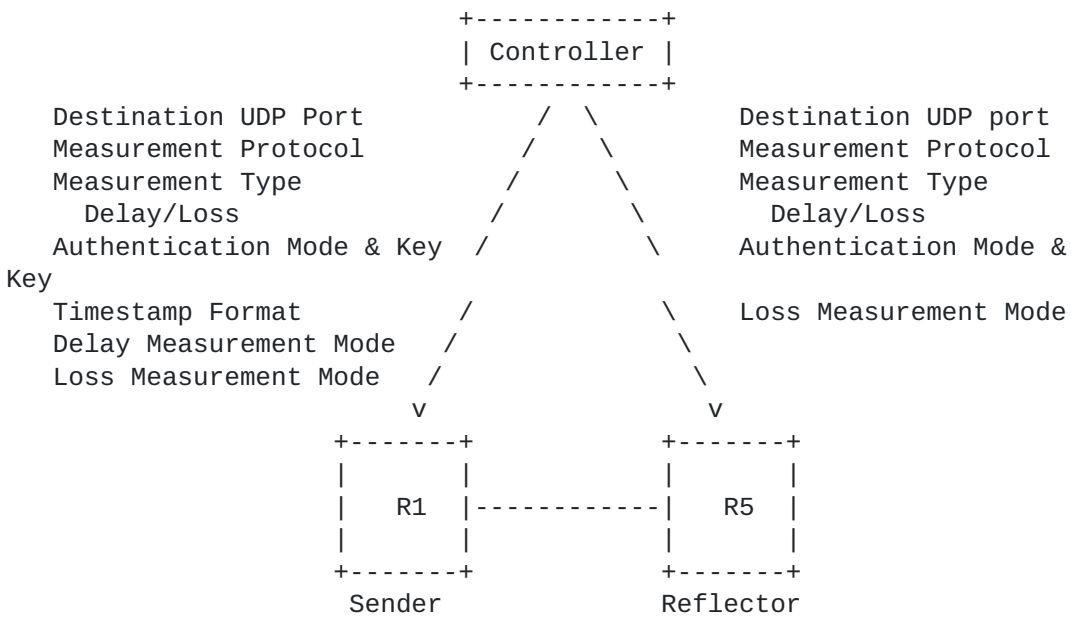


Figure 1: Example Provisioning Model

Example of Measurement Protocol is STAMP, the Timestamp Format is PTPv2 [IEEE1588] or NTP and the Loss Measurement mode is inferred-mode or direct-mode. The mechanisms to provision the sender and reflector nodes are outside the scope of this document.

The reflector node R5 uses the parameters for the timestamp format and delay measurement mode (i.e. one-way, two-way or loopback mode) from the received probe query message.

4. Probe Messages

4.1. Probe Query Message

The probe messages defined in [RFC8762] are used for Delay Measurement for Links and end-to-end SR Policies. For Loss Measurement, the probe messages defined in this document are used.

The Sender IPv4 or IPv6 address is used as the source address. When known, the reflector IPv4 or IPv6 address is used as the destination address. If not known, the address in the range of 127/8 for IPv4 or 0:0:0:0:0:FFFF:7F00/104 for IPv6 is used as destination address. This is the case for example, when using SR Policy with IPv4 endpoint of 0.0.0.0 or IPv6 endpoint of ::0 [I-D.ietf-spring-segment-routing-policy].

4.1.1. Delay Measurement Query Message

The message content for Delay Measurement probe query message using UDP header [RFC0768] is shown in Figure 2. The DM probe query message is sent with user-configured Destination UDP port number for DM. The Destination UDP port cannot be used as Source port, since the message does not have any indication to distinguish between the query and response message. The payload of the DM probe query message contains the delay measuremen message defined in [RFC8762].

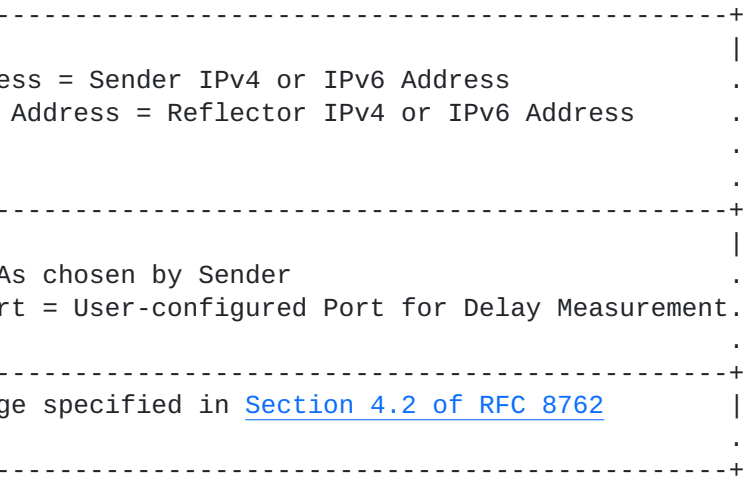


Figure 2: DM Probe Query Message

Timestamp field is eight bytes and use the format defined in Section 4.2.1 of [RFC8762]. It is recommended to use the IEEE 1588v2 Precision Time Protocol (PTP) truncated 64-bit timestamp format

[IEEE1588] as specified in [RFC8186], with hardware support in Segment Routing networks.

4.1.1.1. Delay Measurement Authentication Mode

When using the authenticated mode for delay measurement, the matching authentication type (e.g. HMAC-SHA-256) and key are user-configured on both the sender and reflector nodes. A separate user-configured destination UDP port is used for the delay measurement in authentication mode due to the different probe message format.

4.1.2. Loss Measurement Query Message

The message content for Loss Measurement probe query message using UDP header [RFC0768] is shown in Figure 3. The LM probe query message is sent with user-configured Destination UDP port number for LM, which is a different Destination UDP port number than DM. Separate Destination UDP ports are used for direct-mode and inferred-mode loss measurements. The Destination UDP port cannot be used as Source port, since the message does not have any indication to distinguish between the query and response message. The LM probe query message contains the payload for loss measurement as defined in Figure 7 and Figure 8.

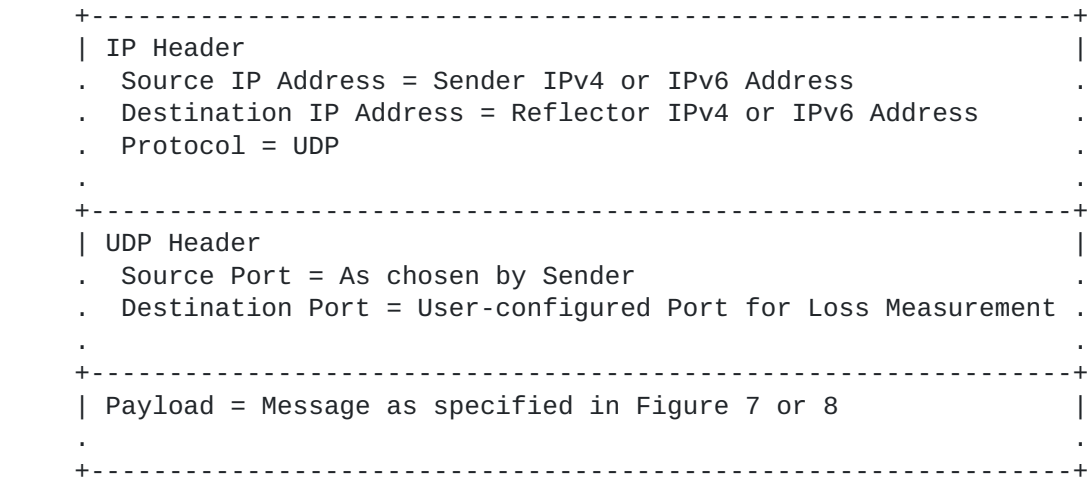


Figure 3: LM Probe Query Message

4.1.2.1. Loss Measurement Authentication Mode

When using the authenticated mode for loss measurement, the matching authentication type (e.g. HMAC-SHA-256) and key are user-configured on both the sender and reflector nodes. A separate user-configured

destination UDP port is used for the loss measurement in authentication mode due to the different message format.

4.1.3. Probe Query for Links

The probe query message as defined in Figure 2 for delay measurement and Figure 3 for loss measurement is sent on the congruent path of the data traffic. The probe messages are routed over the Link for both delay and loss measurement.

4.1.4. Probe Query for End-to-end Measurement for SR Policy

The performance delay and loss measurement for segment routing is applicable to both SR-MPLS and SRv6 Policies.

4.1.4.1. Probe Query Message for SR-MPLS Policy

The probe query messages for end-to-end performance measurement of an SR-MPLS Policy is sent using its SR-MPLS header containing the MPLS segment list as shown in Figure 4.

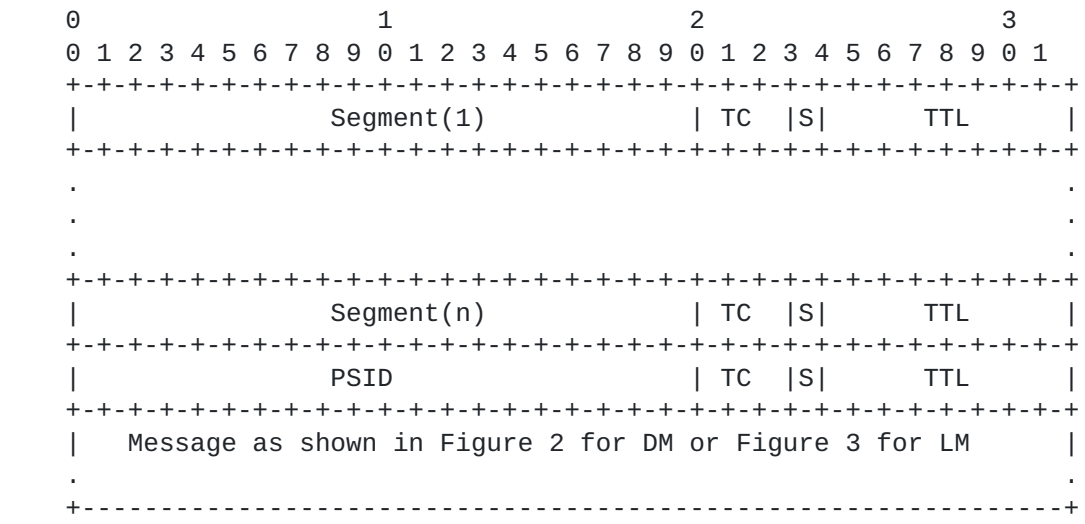


Figure 4: Probe Query Message for SR-MPLS Policy

The Segment List (SL) can be empty to indicate Implicit NULL label case for a single-hop SR Policy.

The Path Segment Identifier (PSID) [I-D.ietf-spring-mpls-path-segment] of the SR-MPLS Policy is used for accounting received traffic on the egress node for loss measurement.

4.1.4.2. Probe Query Message for SRv6 Policy

An SRv6 Policy setup using the SRv6 Segment Routing Header (SRH) and a Segment List as defined in [RFC8754]. For SRv6, network programming is defined in [I-D.ietf-spring-srv6-network-programming].

The probe query messages for end-to-end performance measurement of an SRv6 Policy is sent using its SRH with Segment List as shown in Figure 5.

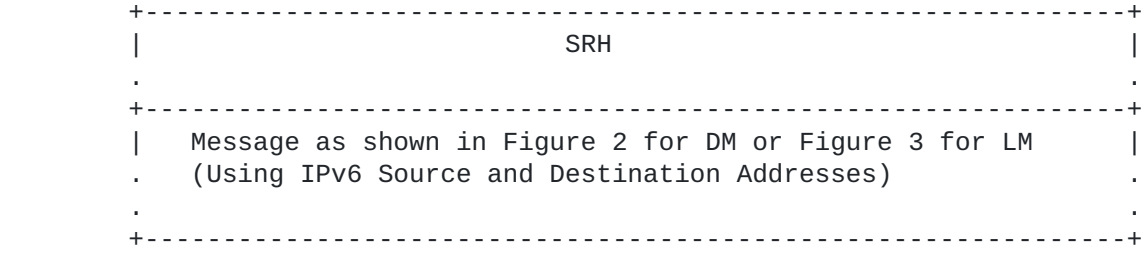


Figure 5: Probe Query Message for SRv6 Policy

For delay measurement of SRv6 Policy using SRH, END function END.OTP [I-D.ietf-6man-spring-srv6-oam] is used with the target SRv6 SID to punt probe messages on the target node, as shown in Figure 5. Similarly, for loss measurement of SRv6 Policy, END function END.OP [I-D.ietf-6man-spring-srv6-oam] is used with target SRv6 SID to punt probe messages on the target node.

4.1.5. Control Code Field for STAMP Messages

The Control Code field is defined for delay and loss measurement probe query and response messages for STAMP protocol in unauthenticated and authenticated modes. The modified delay measurement probe query and response message format for STAMP is shown in Figure 6. This message format is backwards compatible with the message format defined in STAMP [RFC8762] as its reflector MUST ignore the received field (previously identified as MBZ). The usage of the Control Code is not limited to the SR networks and can be used for various bidirectional paths in a network.

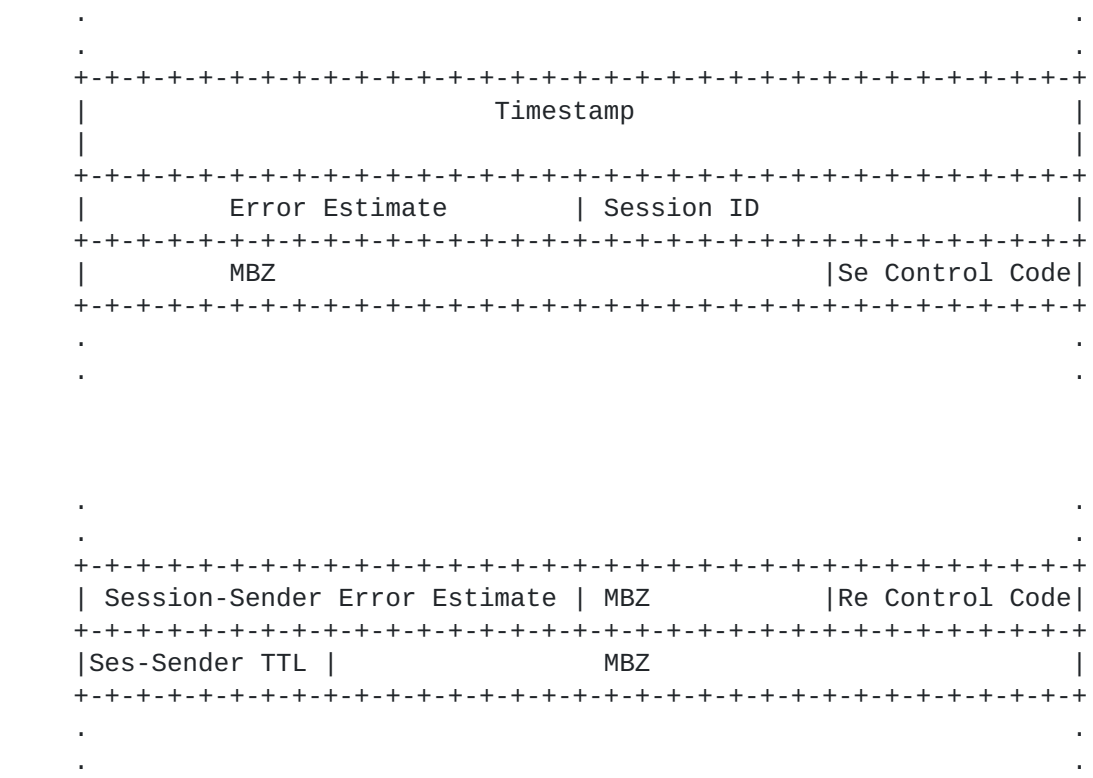


Figure 6: Sender and Reflector Control Code in STAMP DM Message

Sender Control Code: Set as follows in STAMP probe query message.

For a Query:

0x0: Out-of-band Response Requested. Indicates that the probe response is not required over the same path in the reverse direction. This is also the default behavior.

0x1: In-band Response Requested. Indicates that this query has been sent over a bidirectional path and the probe response is required over the same path in the reverse direction. The bidirectional path does not have to be an SR path.

Reflector Control Code: Set as follows in STAMP probe response message.

For a Response:

0x1: Error - Invalid Message. Indicates that the operation failed because the received query message could not be processed.

Additional Error Codes to be defined in future.

4.1.6. Loss Measurement Query Message Formats for STAMP

In this document, STAMP probe query message formats are defined for loss measurement as shown in Figure 7 and Figure 8. The message formats are hardware efficient due to the well-known locations of the counters. They are similar to the delay measurement message formats (e.g. location of the Counter and Timestamp) and do not require any backwards compatibility or support for the existing DM message formats from [RFC8762] as different user-configured destination UDP port is used for loss measurement.

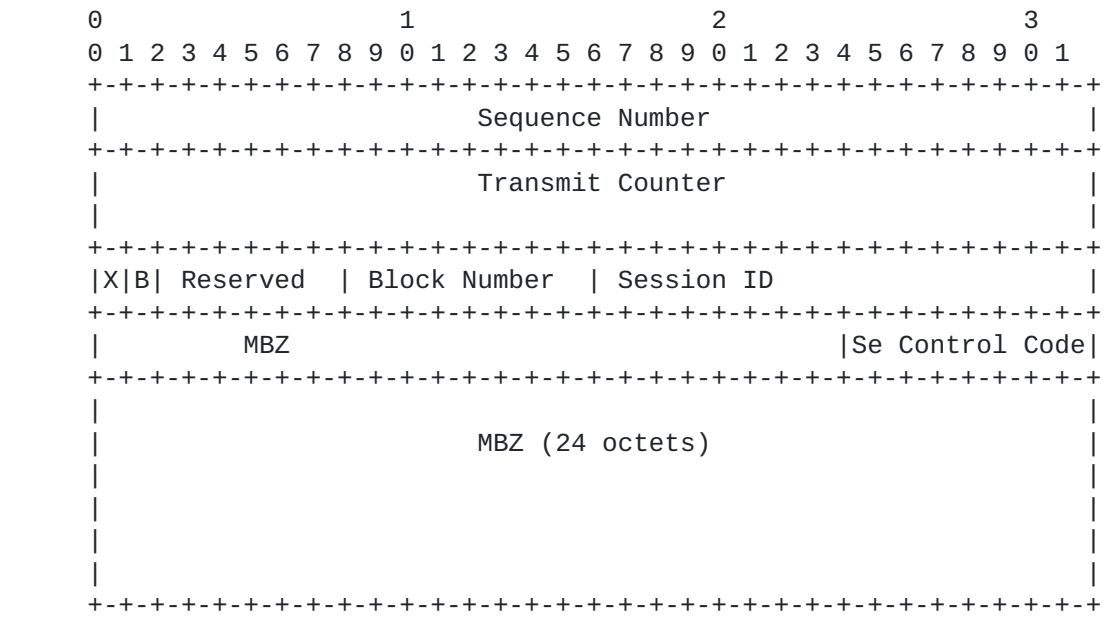


Figure 7: STAMP LM Probe Query Message - Unauthenticated Mode

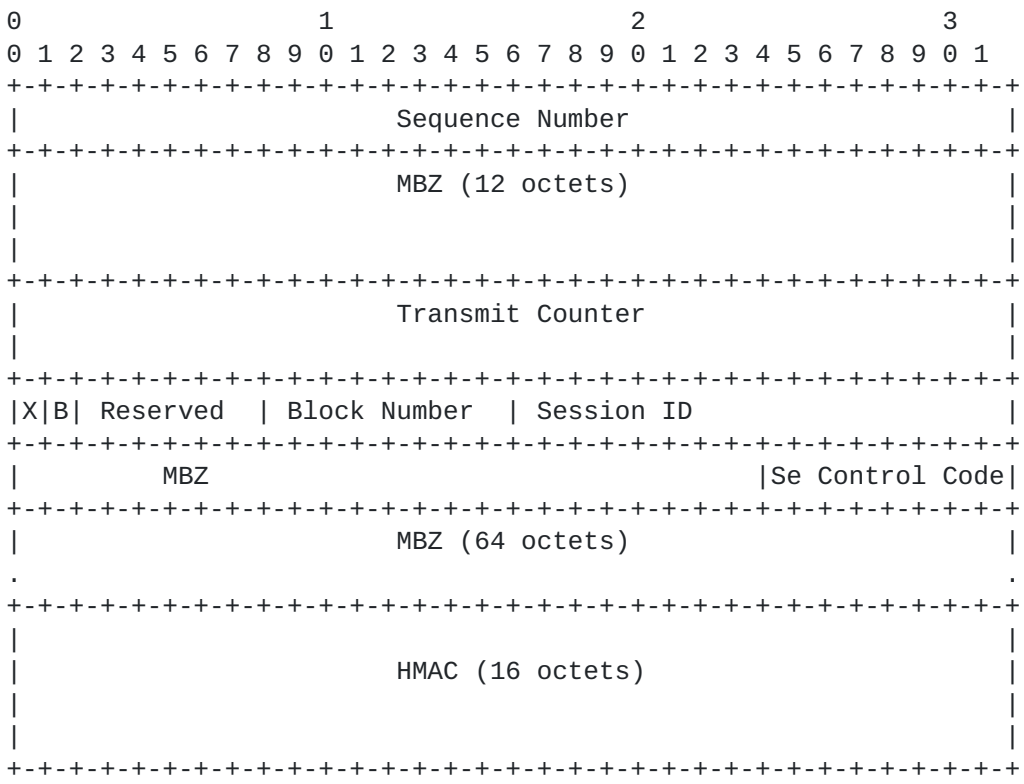


Figure 8: STAMP LM Probe Query Message - Authenticated Mode

Sequence Number (32-bit): As defined in [RFC8762].

Transmit Counter (64-bit): The number of packets or octets sent by the sender node in the query message and by the reflector node in the response message. The counter is always written at the well-known location in the probe query and response messages.

Receive Counter (64-bit): The number of packets or octets received at the reflector node. It is written by the reflector node in the probe response message.

Sender Counter (64-bit): This is the exact copy of the transmit counter from the received query message. It is written by the reflector node in the probe response message.

Sender Sequence Number (32-bit): As defined in [RFC8762].

Sender TTL: As defined in Section 7.1.

LM Flags: The meanings of the Flag bits are:

X: Extended counter format indicator. Indicates the use of extended (64-bit) counter values. Initialized to 1 upon creation (and prior to transmission) of an LM Query and copied from an LM Query to an LM response. Set to 0 when the LM message is transmitted or received over an interface that writes 32-bit counter values.

B: Octet (byte) count. When set to 1, indicates that the Counter 1-4 fields represent octet counts. The octet count applies to all packets within the LM scope, and the octet count of a packet sent or received includes the total length of that packet (but excludes headers, labels, or framing of the channel itself). When set to 0, indicates that the Counter fields represent packet counts.

Block Number (8-bit): The Loss Measurement using Alternate-Marking method defined in [RFC8321] requires to color the data traffic. To be able to compare 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.

HMAC: The PM probe message in authenticated mode includes a key Hashed Message Authentication Code (HMAC) ([RFC2104]) hash. Each probe query and response messages are authenticated by adding Sequence Number with Hashed Message Authentication Code (HMAC) TLV. It can use HMAC-SHA-256 truncated to 128 bits (similarly to the use of it in IPSec defined in [RFC4868]); hence the length of the HMAC field is 16 octets.

HMAC uses its own key and the mechanism to distribute the HMAC key is outside the scope of this document.

In authenticated mode, only the sequence number is encrypted, and the other payload fields are sent in clear text. The probe message MAY include Comp.MBZ (Must Be Zero) variable length field to align the packet on 16 octets boundary.

4.2. Probe Response Message

The probe response message is sent using the IP/UDP information from the received probe query message. The content of the probe response message is shown in Figure 9.

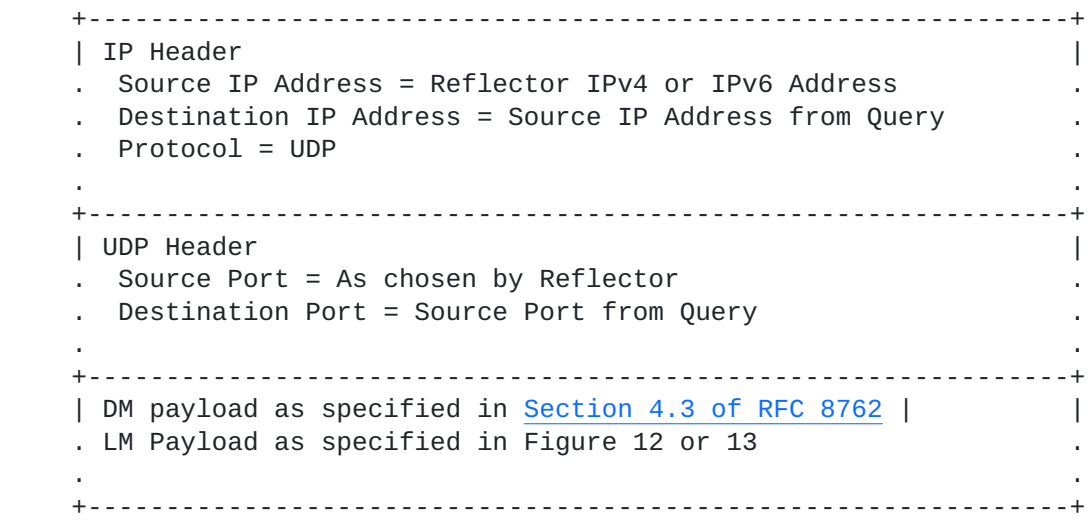


Figure 9: Probe Response Message

4.2.1. One-way Measurement Mode

In one-way performance measurement mode, the probe response message as defined in Figure 9 is sent back out-of-band to the sender node, for both Links and SR Policies. The Sender Control Code is set to "Out-of-band Response Requested". In this delay measurement mode, as per Reference Topology, all timestamps t1, t2, t3, and t4 are collected by the probes. However, only timestamps t1 and t2 are needed to measure one-way delay.

4.2.2. Two-way Measurement Mode

In two-way performance measurement mode, when using a bidirectional path, the probe response message as defined in Figure 9 is sent back to the sender node on the congruent path of the data traffic on the same reverse direction Link or associated reverse SR Policy [I-D.ietf-pce-sr-bidir-path]. The Sender Control Code 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 needed to measure two-way delay.

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.

4.2.2.1. Probe Response Message for SR-MPLS Policy

The message content for sending probe response message for two-way end-to-end performance measurement of an SR-MPLS Policy is shown in Figure 10.

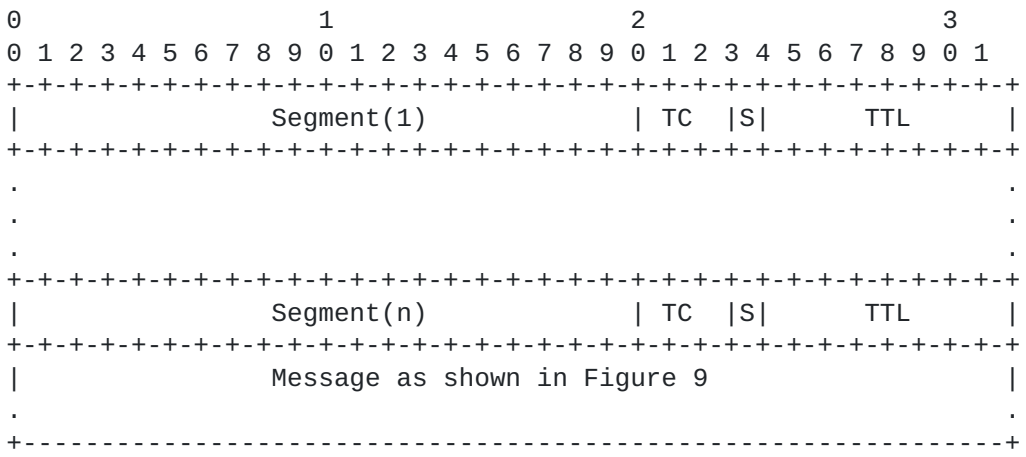


Figure 10: Probe Response Message for SR-MPLS Policy

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

4.2.2.2. Probe Response Message for SRv6 Policy

The message content for sending probe response message on the congruent path of the data traffic for two-way end-to-end performance measurement of an SRv6 Policy with SRH is shown in Figure 11.

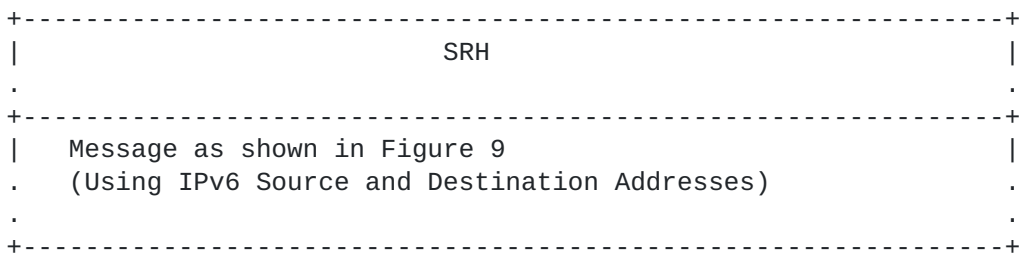


Figure 11: Probe Response Message for SRv6 Policy

4.2.3. Loopback Measurement Mode

The Loopback measurement mode can be used to measure round-trip delay for a bidirectional SR Path. The IP header of the probe query message contains the destination address equals to the sender address and the source address equals to the reflector address. Optionally, the probe query message can carry the reverse path information (e.g. reverse path label stack for SR-MPLS) as part of the SR header. The probe messages are not punted at the reflector node and it does not process them and generate response messages. The Sender Control Code is set to the default value of 0. In this mode, as the probe packet is not punted on the reflector node for processing, the querier copies the 'Sequence Number' in 'Session-Sender Sequence Number' directly. In this delay measurement mode, as per Reference Topology, the timestamps t1 and t4 are collected by the probes. Both these timestamps are needed to measure round-trip delay.

4.2.4. Loss Measurement Response Message Formats for STAMP

In this document, STAMP probe response message formats are defined for loss measurement as shown in Figure 12 and Figure 13.

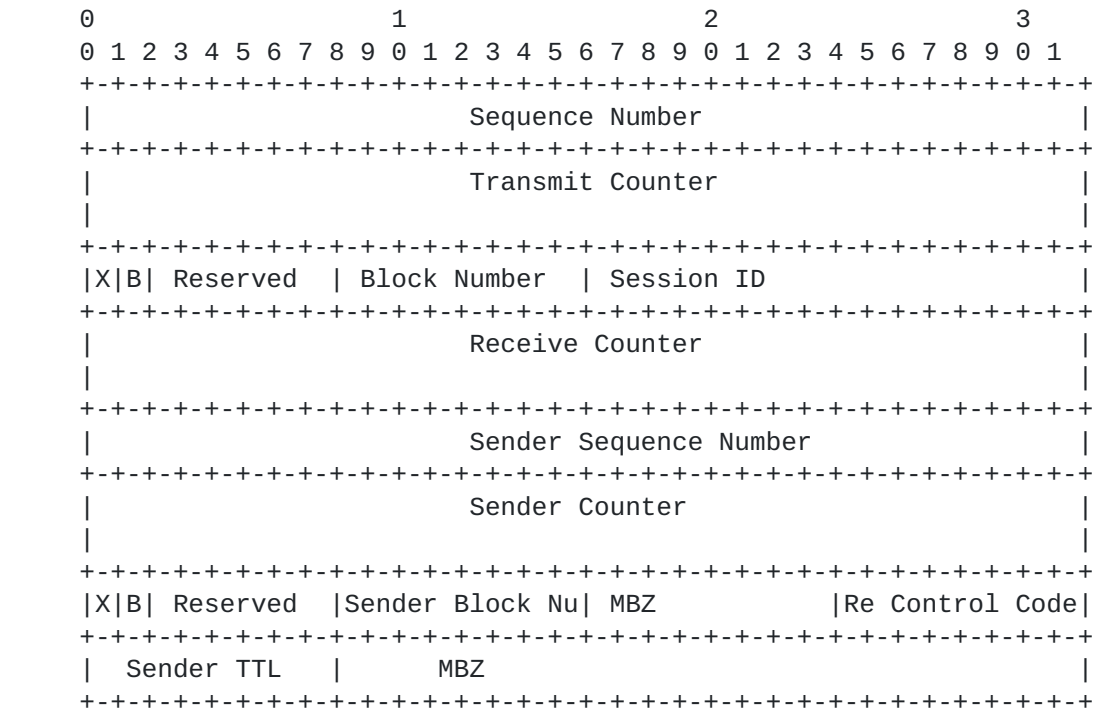


Figure 12: STAMP LM Probe Response Message - Unauthenticated Mode

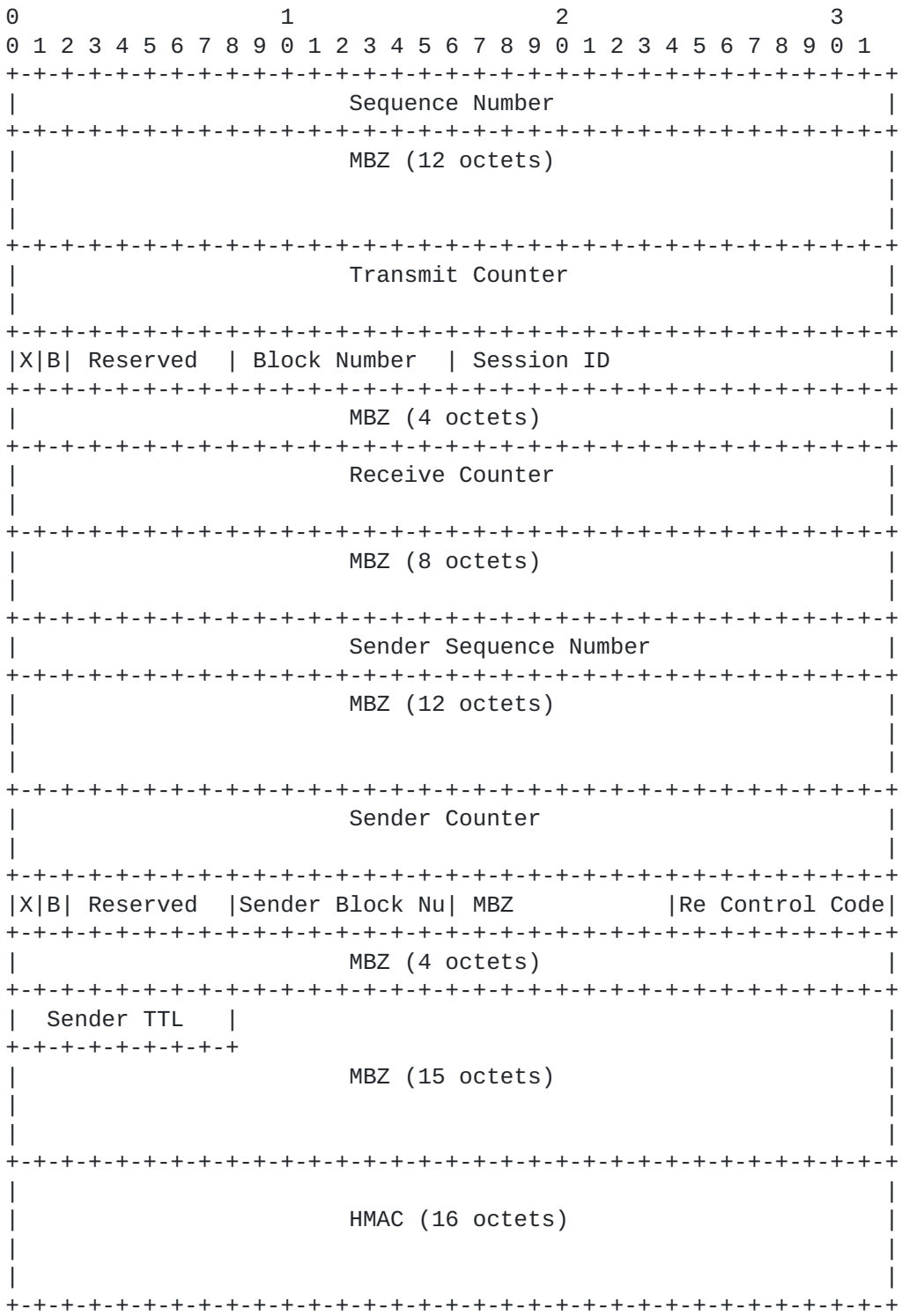


Figure 13: STAMP LM Probe Response Message - Authenticated Mode

4.3. Node Address TLV for STAMP Message

In this document, Node Address TLV is defined for STAMP message [I-D.ietf-ippm-stamp-option-tlv] and has the following format shown in Figure 14:

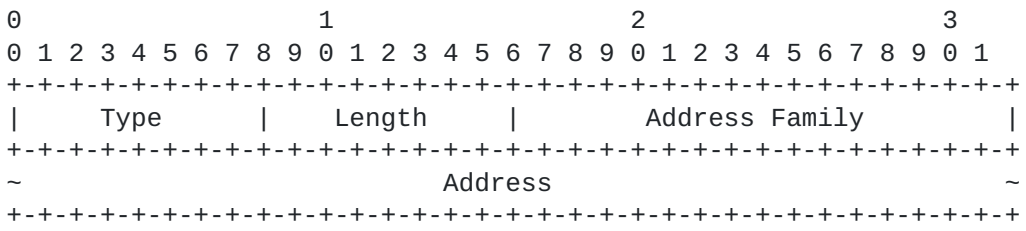


Figure 14: Node Address TLV Format

The Address Family field indicates the type of the address, and it SHALL be set to one of the assigned values in the "IANA Address Family Numbers" registry.

The following Type is defined and it contains Node Address TLV:

Destination Node Address (value TBA1):

The Destination Node Address TLV is optional. The Destination Node Address TLV indicates the address of the intended recipient node of the probe message. The reflector node SHOULD NOT send response if it is not the intended destination node of the probe query message. This check is useful for example, for performance measurement of SR Policy when using the destination address in 127/8 range for IPv4 or in 0:0:0:0:0:FFFF:7F00/104 range for IPv6.

4.4. Return Path TLV for STAMP Message

For two-way performance measurement, the reflector node needs to send the probe response message on a specific reverse path. The sender node can request in the probe query message to the reflector node to send a response back on a given reverse path (e.g. co-routed bidirectional path). This way the destination node does not require any additional SR Policy state.

For one-way performance measurement, the sender node address may not be reachable via IP route from the reflector node. The sender node in this case needs to send its reachability path information to the reflector node.

[I-D.ietf-ippm-stamp-option-tlv] defines STAMP probe query messages that can include one or more optional TLVs. The TLV Type (value

TBA2) is defined in this document for Return Path that carries reverse path for STAMP probe response messages (in the payload of the message). The format of the Return Path TLV is shown in Figure 15:

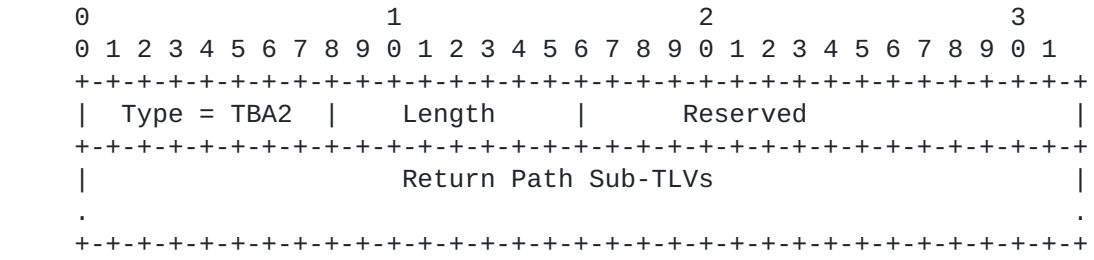


Figure 15: Return Path TLV

The following Type defined for the Return Path TLV contains the Node Address sub-TLV using the format shown in Figure 14:

- Type (value 0): Return Address. Target node address of the response different than the Source Address in the query

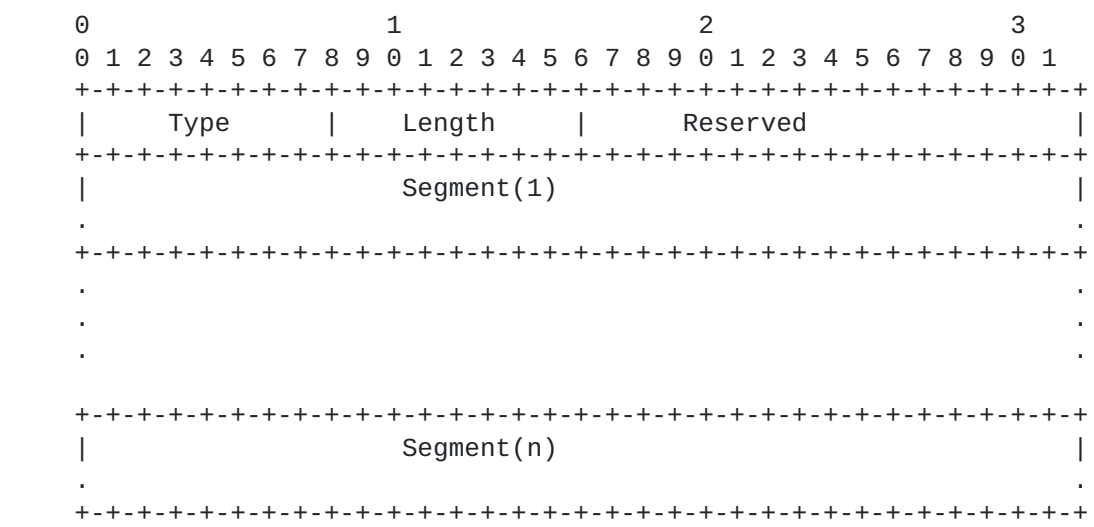


Figure 16: Segment List Sub-TLV in Return Path TLV

The Segment List Sub-TLV (shown in Figure 16) in the Return Path TLV can be one of the following Types:

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

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

- Type (value 3): SRv6 Segment List of the Reverse SR Path

- Type (value 4): SRv6 Binding SID [I-D.ietf-pce-binding-label-sid] of the Reverse SR Policy

The Return Path TLV is optional. The PM sender node MUST only insert one Return Path TLV in the probe query message and the reflector node MUST only process the first Return Path TLV in the probe query message and ignore other Return Path TLVs if present. The reflector 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. Performance Measurement for P2MP SR Policies

The procedures for delay and loss measurement described in this document for Point-to-Point (P2P) SR Policies [I-D.ietf-spring-segment-routing-policy] are also equally applicable to the Point-to-Multipoint (P2MP) SR Policies as following:

- The sender root node sends probe query messages using the Replication Segment defined in [I-D.voyer-spring-sr-replication-segment] for the P2MP SR Policy as shown in Figure 17.

- Each reflector leaf node sends its IP address in the Source Address of the probe response messages as shown in Figure 9. This allows the sender root node to identify the reflector leaf nodes of the P2MP SR Policy.

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

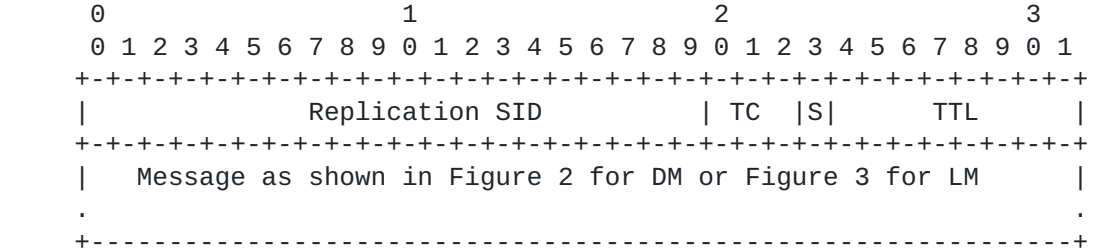


Figure 17: Query with Replication Segment for SR-MPLS Policy

6. ECMP Support for SR Policies

An SR Policy can have ECMPs between the source and transit nodes, between transit nodes and between transit and destination nodes. Usage of Anycast SID [RFC8402] by an SR Policy can result in ECMP paths via transit nodes part of that Anycast group. The PM probe messages need to be sent to traverse different ECMP paths to measure performance delay of an SR Policy.

Forwarding plane has various hashing functions available to forward packets on specific ECMP paths. The mechanisms described in [RFC8029] and [RFC5884] for handling ECMPs are also applicable to the performance measurement. In the IP header of the PM probe messages, sweeping of Destination Addresses in 127/8 range for IPv4 or 0:0:0:0:0:FFFF:7F00/104 range for IPv6 can be used to exercise particular ECMP paths. As specified in [RFC6437], Flow Label field in the outer IPv6 header can also be used for sweeping.

The considerations for performance loss measurement for different ECMP paths of an SR Policy are outside the scope of this document.

7. Additional Message Processing Rules

The processing rules defined in this section are applicable to the STAMP messages for delay and loss measurement for Links and end-to-end SR Policies.

7.1. TTL and Hop Limit

The TTL field in the IPv4 and MPLS headers of the probe query messages is set to 255 [RFC8762]. Similarly, the Hop Limit field in the IPv6 and SRH headers of the probe query messages is set to 255 [RFC8762].

When using the Destination IPv4 Address from the 127/8 range, the TTL in the IPv4 header is set to 1 [RFC8029]. Similarly, when using the Destination IPv6 Address from the 0:0:0:0:0:FFFF:7F00/104 range, the Hop Limit field in the inner IPv6 header is set to 1 whereas in the outer IPv6 header is set to 255.

For Link performance delay and loss measurements, the TTL and Hop Limit field in the probe message is set to 1 in both one-way and two-way measurement modes.

7.2. Router Alert Option

The Router Alert IP option is not set when using the routable Destination IP Address in the probe messages.

When using the Destination IPv4 Address from the 127/8 range, to be able to punt probe packets on the reflector node, the Router Alert IP Option of value 0x0 [RFC2113] for IPv4 MAY be added [RFC8029]. Similarly, when using the Destination IPv6 Address from the 0:0:0:0:0:FFFF:7F00/104 range, the Router Alert IP Option of value 69 [RFC7506] for IPv6 MAY be added in the destination option header, Section 4.6 of [RFC8200]. For SRv6 Policy using SRH, it is added in the inner IPv6 header.

7.3. UDP Checksum

The UDP Checksum Complement for delay and loss measurement messages follows the procedure defined in [RFC7820] and can be optionally used with the procedures defined in this document.

For IPv4 and IPv6 probe messages, where the hardware is not capable of re-computing the UDP checksum or adding checksum complement [RFC7820], the sender node sets the UDP checksum to 0 [RFC6936] [RFC8085]. The receiving node bypasses the checksum validation and accepts the packets with UDP checksum value 0 for the UDP port being used for PM delay and loss measurements.

8. Security Considerations

The performance measurement is intended for deployment in well-managed private and service provider networks. As such, it assumes that a node involved in a measurement operation has previously verified the integrity of the path and the identity of the far-end reflector node.

If desired, attacks can be mitigated by performing basic validation and sanity checks, at the sender, of the counter or timestamp fields in received measurement response messages. The minimal state associated with these protocols also limits the extent of measurement disruption that can be caused by a corrupt or invalid message to a single query/response cycle.

Use of HMAC-SHA-256 in the authenticated mode protects the data integrity of the probe messages. SRv6 has HMAC protection authentication defined for SRH [RFC8754]. Hence, PM probe messages for SRv6 may not need authentication mode. Cryptographic measures may be enhanced by the correct configuration of access-control lists and firewalls.

9. IANA Considerations

IANA is requested to allocate a value for the following optional Destination Address TLV Type for [I-D.ietf-ippm-stamp-option-tlv] to be carried in PM probe messages:

- Type TBA1: Destination Node Address TLV

IANA is also requested to allocate a value for the following optional Return Path TLV Type for [I-D.ietf-ippm-stamp-option-tlv] to be carried in PM probe query messages:

- Type TBA2: Return Path TLV

IANA is also requested to allocate the values for the following Sub-TLV Types for the Return Path TLV.

- Type (value 0): Return Address
- Type (value 1): SR-MPLS Label Stack of the Reverse SR Path
- Type (value 2): SR-MPLS Binding SID [I-D.ietf-pce-binding-label-sid] of the Reverse SR Policy
- Type (value 3): SRv6 Segment List of the Reverse SR Path
- Type (value 4): SRv6 Binding SID [I-D.ietf-pce-binding-label-sid] of the Reverse SR Policy

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>.

[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>.

[I-D.ietf-6man-spring-srv6-oam] Ali, Z., Filsfils, C., Matsushima, S., Voyer, D., and M. Chen, "Operations, Administration, and Maintenance (OAM) in Segment Routing Networks with IPv6 Data plane (SRv6)", draft-ietf-6man-spring-srv6-oam-03 (work in progress), December 2019.

[I-D.ietf-ippm-stamp-option-tlv] Mirsky, G., Xiao, M., Nydell, H., Foote, R., Masputra, A., and E. Ruffini, "Simple Two-way Active Measurement Protocol Optional Extensions", draft-ietf-ippm-stamp-option-tlv-03 (work in progress), February 2020.

10.2. Informative References

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

[RFC2104] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, DOI 10.17487/RFC2104, February 1997, <https://www.rfc-editor.org/info/rfc2104>.

[RFC2113] Katz, D., "IP Router Alert Option", RFC 2113, DOI 10.17487/RFC2113, February 1997, <https://www.rfc-editor.org/info/rfc2113>.

[RFC4868] Kelly, S. and S. Frankel, "Using HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512 with IPsec", RFC 4868, DOI 10.17487/RFC4868, May 2007, <https://www.rfc-editor.org/info/rfc4868>.

[RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, DOI 10.17487/RFC5884, June 2010, <https://www.rfc-editor.org/info/rfc5884>.

[RFC6335] Cotton, M., Eggert, L., Touch, J., Westerlund, M., and S. Cheshire, "Internet Assigned Numbers Authority (IANA) Procedures for the Management of the Service Name and Transport Protocol Port Number Registry", BCP 165, RFC 6335, DOI 10.17487/RFC6335, August 2011, <https://www.rfc-editor.org/info/rfc6335>.

[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>.

[RFC6936] Fairhurst, G. and M. Westerlund, "Applicability Statement for the Use of IPv6 UDP Datagrams with Zero Checksums", RFC 6936, DOI 10.17487/RFC6936, April 2013, <https://www.rfc-editor.org/info/rfc6936>.

[RFC7506] Raza, K., Akiya, N., and C. Pignataro, "IPv6 Router Alert Option for MPLS Operations, Administration, and Maintenance (OAM)", RFC 7506, DOI 10.17487/RFC7506, April 2015, <https://www.rfc-editor.org/info/rfc7506>.

[RFC7820] Mizrahi, T., "UDP Checksum Complement in the One-Way Active Measurement Protocol (OWAMP) and Two-Way Active Measurement Protocol (TWAMP)", RFC 7820, DOI 10.17487/RFC7820, March 2016, <https://www.rfc-editor.org/info/rfc7820>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <https://www.rfc-editor.org/info/rfc8085>.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[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>.

[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.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-06 (work in progress), December 2019.

[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-02 (work in progress), November 2019.

[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.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-14 (work in progress), March 2020.

[I-D.ietf-pce-binding-label-sid] Sivabalan, S., Filsfils, C., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-ietf-pce-binding-label-sid-02 (work in progress), March 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.ali-spring-ioam-srv6]
Ali, Z., Gandhi, R., Filsfils, C., Brockners, F., Kumar, N., Pignataro, C., Li, C., Chen, M., and G. Dawra, "Segment Routing Header encapsulation for In-situ OAM Data", draft-ali-spring-ioam-srv6-02 (work in progress), November 2019.

[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-01 (work in progress), February 2020.

Acknowledgments

The authors would like to thank Thierry Couture for the discussions on the use-cases for Performance Measurement in Segment Routing. The authors would also like to thank Greg Mirsky for reviewing this document and providing useful comments and suggestions. Patrick Khordoc and Radu Valceanu, both from Cisco Systems have helped significantly improve the mechanisms defined in this document. The authors would like to acknowledge the earlier work on the loss measurement using TWAMP described in draft-xiao-ippm-twamp-ext-direct-loss. The authors would also like to thank Sam Aldrin for the discussions to check for broken path.

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

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Bart Janssens
Colt

Email: Bart.Janssens@colt.net