Performance Measurement Using UDP Path for Segment Routing Networks
draft-gandhi-spring-rfc6374-srpm-udp-01

Abstract

Segment Routing (SR) is applicable to both Multiprotocol Label Switching (SR-MPLS) and IPv6 (SRv6) data planes. This document specifies procedures for using UDP path for sending and processing synthetic probe query and response messages for Performance Measurement (PM). The procedure uses the RFC 6374 defined mechanisms for Performance Delay and Loss Measurement. The procedure specified is applicable to SR-MPLS and SRv6 data planes for both links and end-to-end measurement for SR Policies. In addition, this document defines Return Path TLV for two-way performance measurement and Block Number TLV for loss measurement.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."
# 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
3. Probe Messages ................................................ 7
   3.1. Probe Query Message ................................. 7
   3.1.1. Delay Measurement Probe Query Message ............... 7
   3.1.2. Loss Measurement Probe Query Message ............... 7
   3.1.2.1. Block Number TLV ............................... 8
   3.1.3. Probe Query for SR Links .......................... 9
   3.1.4. Probe Query for End-to-end Measurement for SR Policy . 9
   3.1.4.1. Probe Query Message for SR-MPLS Policy .......... 9
   3.1.4.2. Probe Query Message for SRv6 Policy .......... 10
3.2. Probe Response Message ................................... 10
   3.2.1. One-way Measurement Mode .......................... 11
   3.2.1.1. SR Links and End-to-end Measurement for SR Policy .... 11
   3.2.1.2. Probe Response Message to Controller ............ 12
   3.2.2. Two-way Measurement Mode .......................... 12
   3.2.2.1. SR Links ........................................ 12
   3.2.2.2. End-to-end Measurement for SR Policy ............ 12
   3.2.2.3. Return Path TLV ................................ 12
   3.2.2.4. Probe Response Message for SR-MPLS Policy ........ 14
   3.2.2.5. Probe Response Message for SRv6 Policy .......... 14
   3.2.3. Loopback Measurement Mode ........................ 15
3.3. Checksum Complement .................................... 15
4. Performance Measurement for P2MP SR Policies .................. 15
5. ECMP Support for SR Policies ................................ 16
6. Sequence Numbers ............................................. 16
1. Introduction

Segment Routing (SR) technology 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, transit and destination nodes. SR Policies as defined in [I-D.spring-segment-routing-policy] are used to steer traffic through a specific, user-defined path using a stack of Segments. Built-in SR Performance Measurement (PM) is one of the essential requirements to provide Service Level Agreements (SLAs).

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. These protocols rely on control channel signaling to establish a test channel over an UDP path. These protocols lack support for IEEE 1588 timestamp [IEEE1588] format and direct-mode Loss Measurement (LM), which are required in SR networks [RFC6374]. The Simple Two-way Active Measurement Protocol (STAMP) [I-D.ippm-stamp] alleviates the control channel signaling by using configuration data model to provision test channels. In addition, the STAMP supports IEEE 1588 timestamp format for Delay Measurement (DM). The TWAMP Light from broadband forum [BBF.TR-390] provides simplified mechanisms for active performance measurement in Customer Edge IP networks. [Y1731] specifies the mechanisms to carry OAM messages specifically for Ethernet networks that include Ethernet Frame Delay and Loss measurements.

[RFC6374] specifies protocol mechanisms to enable the efficient and accurate measurement of performance metrics and can be used in SR networks with MPLS data plane [I-D.spring-rfc6374-srpm-mpls]. [RFC6374] addresses the limitations of the IP based performance measurement protocols as specified in Section 1 of [RFC6374]. The [RFC6374] requires data plane to support MPLS Generic Associated
Channel Label (GAL) and Generic Associated Channel (G-Ach), which may not be supported on all nodes in the network.

[RFC7876] specifies the procedures to be used when sending and processing out-of-band performance measurement probe response messages over an UDP return path for RFC 6374 based probe queries. [RFC7876] can be used to send out-of-band PM probe responses in both SR-MPLS and SRv6 networks for one-way performance measurement.

For SR Policies, there are ECMPs between the source and transit nodes, between transit nodes and between transit and destination nodes. Existing PM protocols (e.g. RFC 6374) do not define handling for ECMP forwarding paths in SR networks.

For two-way measurements for SR Policies, there is a need to specify a return path in the form of a Segment List in PM probe query messages without requiring any SR Policy state on the destination node. Existing protocols do not have such mechanisms to specify return path in the PM probe query messages.

This document specifies a procedure for using UDP path for sending and processing synthetic probe query and response messages for Performance Measurement that does not require to bootstrap PM sessions. The procedure uses RFC 6374 defined mechanisms for Performance Delay and Loss Measurement and unless otherwise specified, the procedures from RFC 6374 are not modified. The procedure specified is applicable to both SR-MPLS and SRv6 data planes. The procedure can be used for both SR links and end-to-end performance measurement for SR Policies. This document also defines mechanisms for handling Equal Cost Multi-Paths (ECMPs) of SR Policies for performance delay measurement. In addition, this document defines Return Path TLV for two-way performance measurement, Block Number TLV for loss measurement and Sequence Number TLV.

2. Conventions Used in This Document

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Abbreviations

ACH: Associated Channel Header.
BSID: Binding Segment ID.

DFlag: Data Format Flag.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

G-ACh: Generic Associated Channel (G-ACh).

GAL: Generic Associated Channel (G-ACh) Label.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.


PM: Performance Measurement.

PSID: Path Segment Identifier.

PTP: Precision Time Protocol.

SID: Segment ID.

SL: Segment List.

SR: Segment Routing.

SR-MPLS: Segment Routing with MPLS data plane.

SRv6: Segment Routing with IPv6 data plane.


TC: Traffic Class.


URO: UDP Return Object.

2.3. Reference Topology

In the reference topology, the querier node R1 initiates a probe
query for performance measurement and the responder node R5 sends a probe response for the query message received. The probe response may be sent to the querier node R1 or to a controller node R100. The nodes R1 and R5 may be directly connected via a link enabled with Segment Routing or there exists a Point-to-Point (P2P) SR Policy [I-D.spring-segment-routing-policy] on node R1 with destination to node R5. In case of Point-to-Multipoint (P2MP), SR Policy originating from source node R1 may terminate on multiple destination leaf nodes [I-D.spring-sr-p2mp-policy].

```
------
| R100 |
------
 ^
------
 | Response |
 +--------+  Query  +--------+
 | R1  |---------------------| R5  |
 |<--- |---------------------|<--- |
 +------+  Response  +------+
```

Reference Topology

For delay and loss measurements, for both links and end-to-end SR Policies, no PM session is created on the responder node R5. One-way delay and two-way delay measurements are defined in Section 2.4 of [RFC6374]. Transmit and Receive packet loss measurements are defined in Section 2.2 and Section 2.6 of [RFC6374]. One-way loss measurement provides receive packet loss whereas two-way loss measurement provides both transmit and receive packet loss.

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

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

- For Loss Measurement, the probe messages are sent on the congruent path of the data traffic by the querier node, and are used to collect the receive traffic counters for the incoming link or incoming SID where the probe query messages are received at the responder node (incoming link or incoming SID used as the responder node has no PM session state present).

The In-Situ Operations, Administration, and Maintenance (IOAM)
mechanisms for SR-MPLS defined in [I-D.spring-ioam-sr-mpls] and for SRv6 defined in [I-D.spring-srv6-oam] are used to carry PM information in-band as part of the data traffic, and are outside the scope of this document.

3. Probe Messages

3.1. Probe Query Message

In this document, UDP path is used for Delay and Loss measurements for SR links and end-to-end SR Policies. The user-configured UDP ports are used for identifying PM probe packets and to avoid signaling to bootstrap PM sessions. This approach is similar to the one defined in STAMP protocol [I-D.ippm-stamp]. The IPv4 TTL or IPv6 Hop Limit field of the IP header MUST be set to 255.

3.1.1. Delay Measurement Probe Query Message

The message content for Delay Measurement for probe query message using UDP header [RFC768] is shown in Figure 1. The DM probe query message is sent with user-configured Destination UDP port number for DM. The Destination UDP port can also be used as Source port for two-way delay measurement, since the message has a flag to distinguish between query and response. The DM probe query message contains the payload for delay measurement defined in Section 3.2 of [RFC6374].

+---------------------------------------------------------------+
| IP Header                                                     |
| . Source IP Address = Querier IPv4 or IPv6 Address            |
| . Destination IP Address = Responder IPv4 or IPv6 Address      |
| . Protocol = UDP                                              |
| . Router Alert Option Not Set                                  |
| +---------------------------------------------------------------+  
| UDP Header                                                   |
| . Source Port = As chosen by Querier                           |
| . Destination Port = User-configured Port for Delay Measurement.
| .                                                                |
| +---------------------------------------------------------------+  
| Payload = Message as specified in Section 3.2 of RFC 6374     |
| .                                                                |
+---------------------------------------------------------------+

Figure 1: DM Probe Query Message

3.1.2. Loss Measurement Probe Query Message
The message content for Loss measurement probe query message using UDP header [RFC768] is shown in Figure 2. As shown, the LM probe query message is sent with user-configured Destination UDP port number for LM. Different Destination UDP ports are used for direct-mode and inferred-mode loss measurements. The Destination UDP port can also be used as Source port for two-way loss measurement, since the message has a flag to distinguish between query and response. The LM probe query message contains the payload for loss measurement defined in Section 3.1 of [RFC6374].

+---------------------------------------------------------------+
| IP Header                                                     |
| . Source IP Address = Querier IPv4 or IPv6 Address             |
| . Destination IP Address = Responder IPv4 or IPv6 Address       |
| . Protocol = UDP                                               |
| . Router Alert Option Not Set                                  |
| +---------------------------------------------------------------+
| UDP Header                                                    |
| . Source Port = As chosen by Querier                           |
| . Destination Port = User-configured Port for Loss Measurement  |
| +---------------------------------------------------------------+
| Payload = Message as specified in Section 3.1 of RFC 6374      |
| +---------------------------------------------------------------+

Figure 2: LM Probe Query Message

3.1.2.1. Block Number TLV

The Loss Measurement using Alternate-Marking method defined in [RFC8321] requires to identify the Block Number (or color) of the traffic counters carried by the probe query and response messages. Probe query and response messages specified in [RFC6374] for Loss Measurement do not define any means to carry the Block Number.

[RFC6374] defines probe query and response messages that can include one or more optional TLVs. New TLV Type (value TBA2) is defined in this document to carry Block Number (16-bit) for the traffic counters in the probe query and response messages for loss measurement. The format of the Block Number TLV is shown in Figure 11:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
```

Figure 11: Block Number TLV

The Block Number TLV is optional. The PM querier node SHOULD only insert one Block Number TLV in the probe query message and the responder node in the probe response message SHOULD return the first Block Number TLV from the probe query messages and ignore other Block Number TLVs if present. In both probe query and response messages, the counters MUST belong to the same Block Number.

3.1.3. Probe Query for SR Links

The probe query message as defined in Figure 1 is sent on the congruent path of the data traffic for performance Delay measurement. Similarly, the probe query message as defined in Figure 2 is sent on the congruent path of the data traffic for performance Loss measurement.

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

3.1.4.1. Probe Query Message for SR-MPLS Policy

The message content for the probe query message using UDP header for end-to-end performance measurement of SR-MPLS Policy is shown in Figure 3.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Segment List(1)        | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Segment List(n)        | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                PSID                   | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Message as shown in Figure 1 for DM or Figure 2 for LM      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
The Segment List (SL) can be empty to indicate Implicit NULL label case.

The Path Segment Identifier (PSID) [I-D.spring-mpls-path-segment] of the SR-MPLS Policy is used for an egress node for loss measurement. The PSID is not required for delay measurement.

3.1.4.2. Probe Query Message for SRv6 Policy

An SRv6 Policy is setup using the SRv6 Segment Routing Header (SRH) and a Segment List as defined in [I-D.6man-segment-routing-header]. The probe query messages using UDP header for end-to-end performance measurement of an SRv6 Policy is sent using its SRv6 Segment Routing Header (SRH) and Segment List as shown in Figure 4.

Figure 4: Probe Query Message for SRv6 Policy

For delay measurement of SRv6 Policy using SRH, END function END.OTP [I-D.spring-srv6-oam] is used with the target SRv6 SID to punt probe messages on the target node, as shown in Figure 4. Similarly, for loss measurement of SRv6 Policy, END function END.OP [I-D.spring-srv6-oam] is used with target SRv6 SID to punt probe messages on the target node.

3.2. Probe Response Message

When the received probe query message does not contain any UDP Return Object (URO) TLV [RFC7876], the probe response message is sent using the IP/UDP information from the probe query message. The content of the probe response message is shown in Figure 5.
. Source IP Address =Responder IPv4 or IPv6 Address
. Destination IP Address = Source IP Address from Query
. Protocol = UDP
. Router Alert Option Not Set

+---------------------------------------------------------------+
| UDP Header                                                    |
. Source Port = As chosen by Responder                         |
. Destination Port = Source Port from Query                     |

+---------------------------------------------------------------+
| Message as specified in Section 3.2 of RFC 6374 for DM, or    |
. Message as specified in Section 3.1 of RFC 6374 for LM       |

+---------------------------------------------------------------+

Figure 5: Probe Response Message

When the received probe query message contains UDP Return Object (URO) TLV [RFC7876], the probe response message uses the IP/UDP information from the URO in the probe query message. The content of the probe response message is shown in Figure 6.

+---------------------------------------------------------------+
| IP Header                                                     |
. Source IP Address = Responder IPv4 or IPv6 Address            |
. Destination IP Address = URO.Address                         |
. Protocol = UDP                                               |
. Router Alert Option Not Set                                  |

+---------------------------------------------------------------+
| UDP Header                                                    |
. Source Port = As chosen by Responder                         |
. Destination Port = URO.UDP-Destination-Port                  |

+---------------------------------------------------------------+
| Message as specified in Section 3.2 of RFC 6374 for DM, or    |
. Message as specified in Section 3.1 of RFC 6374 for LM       |

+---------------------------------------------------------------+

Figure 6: Probe Response Message Using URO from Probe Query

3.2.1. One-way Measurement Mode

3.2.1.1. SR Links and End-to-end Measurement for SR Policy

In one-way performance measurement mode, the probe response message
as defined in Figure 5 or Figure 6 is sent out-of-band for both SR links and SR Policies.

The PM querier node can receive probe response message back by setting its own IP address as Source Address of the header or by adding URO TLV in the probe query message and setting its own IP address in the IP Address in the URO TLV (Type=131) [RFC7876]. The "control code" in the probe query message is set to "out-of-band response requested". The "Source Address" TLV (Type 130), and "Return Address" TLV (Type 1), if present in the probe query message, are not used to send probe response message.

3.2.1.2. Probe Response Message to Controller

As shown in the Reference Topology, if the querier node requires the probe response message to be sent to the controller R100, it adds URO TLV in the probe query message and sets the IP address of R100 in the IP Address field and user-configured UDP port for DM and for LM in the UDP-Destination-Port field of the URO TLV (Type=131) [RFC7876].

3.2.2. Two-way Measurement Mode

3.2.2.1. SR Links

In two-way performance measurement mode, when using a bidirectional link, the probe response message as defined in Figure 5 or Figure 6 is sent back on the congruent path of the data traffic to the querier node for SR links. In this case, the "control code" in the probe query message is set to "in-band response requested" [RFC6374].

3.2.2.2. End-to-end Measurement for SR Policy

In two-way performance measurement mode, when using a bidirectional path, the probe response message is sent back on the congruent path of the data traffic to the querier node for end-to-end measurement of SR Policies. In this case, the "control code" in the probe query message is set to "in-band response requested" [RFC6374].

3.2.2.3. Return Path TLV

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

[RFC6374] defines DM and LM probe query messages that can include one
or more optional TLVs. New TLV Type (TBA1) is defined in this
document for Return Path to carry reverse SR path for probe response
messages (in the payload of the message). The format of the Return
Path TLV is shown in Figure 7A and 7B:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type = TBA1   |    Length     |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Return Path Sub-TLVs                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Figure 7A: Return Path TLV

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type      |    Length     |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Segment List(1)                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Figure 7B: Segment List Sub-TLV in Return Path TLV
```

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

- Type (value 1): SR-MPLS Label Stack of the Reverse SR Policy
- Type (value 2): SR-MPLS Binding SID [I-D.pce-binding-label-sid] of
  the Reverse SR Policy
- Type (value 3): SRv6 Segment List of the Reverse SR Policy
- Type (value 4): SRv6 Binding SID [I-D.pce-binding-label-sid] of
  the Reverse SR Policy
With sub-TLV Type 1, the Segment List(1) can be used by the responder node to compute the next-hop IP address and outgoing interface to send the probe response messages.

The Return Path TLV is optional. The PM querier node MUST only insert one Return Path TLV in the probe query message and the responder node MUST only process the first Return Path TLV in the probe query message and ignore other Return Path TLVs if present. The responder node MUST send probe response message back on the reverse path specified in the Return Path TLV and MUST NOT add Return Path TLV in the probe response message.

3.2.2.4. Probe Response Message for SR-MPLS Policy

The message content for sending probe response message on the congruent path of the data traffic using UDP header for two-way end-to-end performance measurement of an SR-MPLS Policy is shown in Figure 8. The SR-MPLS label stack in the packet header is built using the Segment List received in the Return Path TLV in the probe query message.

```
+---------------+---------------+---------------+---------------+
|                | Segment List(1)| Segment List(n)|                |
+---------------+---------------+---------------+---------------+
|                | TC |S|      TTL    |                |
|                |               |               |                |
|                | Message as shown in Figure 5 or 6 |                |
+---------------+---------------+---------------+---------------+
```

Figure 8: Probe Response Message for SR-MPLS Policy

The Path Segment Identifier (PSID) [I-D.spring-mpls-path-segment] of the forward SR-MPLS Policy can be used to find the reverse SR-MPLS Policy to send the probe response message for two-way measurement in the absence of Return Path TLV defined in the following Section.

3.2.2.5. Probe Response Message for SRv6 Policy

The message content for sending probe response message on the congruent path of the data traffic using UDP header for two-way end-to-end performance measurement of an SRv6 Policy is shown in
Figure 9. For SRv6 Policy using SRH, the SRv6 SID list in the SRH of the probe response message is built using the SRv6 Segment List received in the Return Path TLV in the probe query message.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          SRH                                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.   END.OTP (DM) or END.OP (LM) with Target SRv6 SID            .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Message as shown in Figure 5 or 6 (with IPv6 Addresses)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: Probe Response Message for SRv6 Policy

3.2.3. Loopback Measurement Mode

The Loopback measurement mode defined in Section 2.8 of [RFC6374] can be used to measure round-trip delay of a bidirectional Path. The probe query messages in this case either carry the reverse Path information as part of the SR header or set the querier address in the destination address in the IP header. The responder node does not process the PM probe messages and generate response messages.

3.3. Checksum Complement

For both delay and loss measurement, when the probe packets are updated with timestamp or counter, UDP Checksum field also need updating since these packets are transported over UDP. As an alternative, the Checksum Complement field (2 Bytes) can be optionally updated using the procedure defined in [RFC7820]. The Checksum Complement field can be any unused field in the probe message and is a local behavior.

4. 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.spring-segment-routing-policy] are also equally applicable to the Point-to-Multipoint (P2MP) SR Policies [I-D.spring-sr-p2mp-policy] as following:

- The querier root node sends probe query messages using the either Spray P2MP segment or TreeSID P2MP segment defined in
5. 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. Following mechanisms can be used in PM probe messages to take advantage of the hashing function in forwarding plane to influence the path taken by them.

- The mechanisms described in [RFC8029] and [RFC5884] for handling ECMPs are also applicable to the performance measurement. In the IP/UDP header of the PM probe messages, 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 a particular ECMP path. As specified in [RFC6437], 3-tuple of Flow Label, Source Address and Destination Address fields in the IPv6 header can also be used.

- For SR-MPLS Policy, entropy label [RFC6790] can be used in the PM probe messages.

- For SRv6 Policy using SRH, Flow Label in the SRH [I-D.6man-segment-routing-header] of the PM probe messages can be used.

6. Sequence Numbers

The message formats for DM and LM [RFC6374] can carry either timestamp or sequence number but not both. There are cases where both timestamp and sequence number are desired for both DM and LM. Sequence numbers can be useful when some probe query messages are lost or they arrive out of order. In addition, the sequence numbers
can be useful for detecting denial-of-service (DoS) attacks on UDP ports.

6.1. Sequence Number TLV in Unauthenticated Mode

[RFC6374] defines DM and LM probe query and response messages that can include one or more optional TLVs. New TLV Type (value TBA3) is defined in this document to carry sequence number for probe query and response messages for delay and loss measurement. The format of the Sequence Number TLV is shown in Figure 10:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type TBA3   |    Length     |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Sequence Number                            |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: Sequence Number TLV - Unauthenticated Mode

- The sequence numbers start with 0 and are incremented by one for each subsequent probe query packet.
- The sequence number are independent for DM and LM messages.
- The sequence number can be of any length determined by the querier node.
- The Sequence Number TLV is optional.
- The PM querier node SHOULD only insert one Sequence Number TLV in the probe query message and the responder node in the probe response message SHOULD return the first Sequence Number TLV from the probe query message and ignore the other Sequence Number TLVs if present.
- When Sequence Number TLV is added, the DM and LM messages SHOULD NOT carry sequence number in the timestamp field of the message.

6.2. Sequence Number TLV in Authenticated Mode

The PM probe query and response packet format 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.

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

The OWAMP and TWAMP compute HMAC field using HMAC-SHA1 and can also be used with the procedure defined in this document.

HMAC uses own key and the definition of the mechanism to distribute the HMAC key is outside the scope of this document. Both the authentication type and key can be user-configured on both the querier and responder nodes.

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type TBA4   |    Length     |      Reserved                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Sequence Number                            |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                    Comp.MBZ                                   ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    HMAC (16 octets)                           |
|                                                               |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Figure 11: Sequence Number TLV - Authenticated Mode
```

- This TLV is mandatory in the authenticated mode.
- The node MUST discard the probe message if HMAC is invalid.
- The Sequence Number follows the same processing rule as defined in the unauthenticated mode.

7. 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 responder node. The security considerations described in Section 8 of [RFC6374] are applicable to this specification, and particular attention should be paid to the last three paragraphs.

Use of HMAC-SHA-256 in the authenticated mode defined in this document protects the data integrity of the probe messages. SRv6 has HMAC protection authentication defined for SRH [I-D.6man-segment-routing-header]. 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.

8. IANA Considerations

IANA is requested to allocate values for the following Return Path TLV Type for RFC 6374 to be carried in PM probe query messages:

- Type TBA1: Return Path TLV

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

- Type 1: SR-MPLS Label Stack of the Reverse SR Policy
- Type 2: SR-MPLS Binding SID of the Reverse SR Policy
- Type 3: SRv6 Segment List of the Reverse SR Policy
- Type 4: SRv6 Binding SID of the Reverse SR Policy

IANA is also requested to allocate a value for the following Block Number TLV Type for RFC 6374 to be carried in the PM probe query and response messages for loss measurement:

- Type TBA2: Block Number TLV

IANA is also requested to allocate a value for the following Sequence Number TLV Types for RFC 6374 to be carried in the PM probe query and response messages for delay and loss measurement:

- Type TBA3: Sequence Number TLV in Unauthenticated Mode
9. References

9.1. Normative References


9.2. Informative References


editor.org/info/rfc4868>.


Acknowledgments

The authors would like to thank Nagendra Kumar and Carlos Pignataro for the discussion on SRv6 Performance Measurement. The authors would like to thank Thierry Couture for various discussions on the use-cases for the performance measurement in segment routing networks. The authors would also like to thank Stewart Bryant for the discussion on UDP port allocation for Performance Measurement and Greg Mirsky for providing useful comments and suggestions.

Contributors

Sagar Soni
Cisco Systems, Inc.
Email: sagsoni@cisco.com

Patrick Khordoc
Cisco Systems, Inc.
Email: pkhordoc@cisco.com

Zafar Ali
Cisco Systems, Inc.
Email: zali@cisco.com

Authors’ Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada
Email: rgandhi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Email: cfilsfil@cisco.com

Daniel Voyer
Bell Canada
Email: daniel.voyer@bell.ca

Stefano Salsano
Universita di Roma "Tor Vergata"
Italy
Email: stefano.salsano@uniroma2.it

Pier LuigiVentre
CNIT
Italy
Email: pierluigi.ventre@cnit.it

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