SPRING Working Group
Internet-Draft
Intended Status: Informational

R. Gandhi, Ed.
C. Filsfils
S. Soni
P. Khordoc
Z. Ali
Cisco Systems, Inc.
D. Voyer
D. Bernier
Bell Canada
S. Salsano
Universita di Roma "Tor Vergata"
P. Ventre
CNIT

February 14, 2018

Performance Measurement in Segment Routing Networks with MPLS Data Plane
draft-gandhi-spring-sr-mpls-pm-00.txt

Abstract

RFC 6374 specifies protocol mechanisms to enable the efficient and accurate measurement of packet loss, one-way and two-way delay, as well as related metrics such as delay variation and channel throughput in MPLS networks. This document reviews how these mechanisms can be used for Performance Measurements in Segment Routing with MPLS data plane (SR-MPLS) networks.

Table of Contents

1. Introduction

Service provider's ability to satisfy Service Level Agreements (SLAs) depend on the ability to measure and monitor performance metrics for packet loss and one-way and two-way delay, as well as related metrics such as delay variation and channel throughput. The ability to monitor these performance metrics also provides operators with greater visibility into the performance characteristics of their networks, thereby facilitating planning, troubleshooting, and network performance evaluation.

[RFC6374] specifies protocol mechanisms to enable the efficient and accurate measurement of these performance metrics in MPLS 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. However, mechanisms in [RFC6374] are more suitable for Segment Routing when using MPLS data plane. This document reviews how these mechanisms can be used for performance measurements (PM) in Segment Routing with the MPLS data plane (SR-MPLS) networks.

2. Conventions Used in This Document

2.1. Abbreviations

ACH: Associated Channel Header.

DM: Delay Measurement.

ECMP: Equal Cost Multi-Path.

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

GAL: Generic Associated Channel (G-ACh) Label

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.

PM: Performance Measurement.

PTP: Precision Time Protocol.

SID: Segment ID.

TC: Traffic Class.

UCMP: Unequal Cost Multi-Path.

2.2. Terminology and Reference Topology

In this document, the following simple topology is used for illustration.

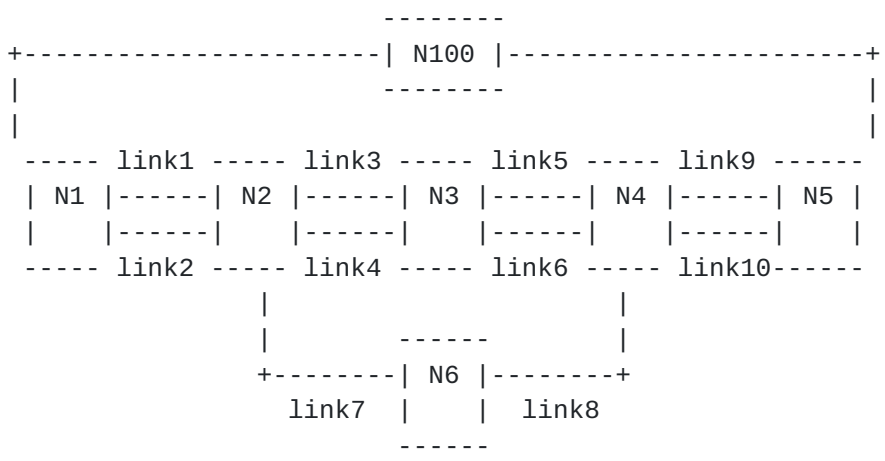


Figure 1: Reference Topology

In the reference topology in Figure 1:

Nodes N1, N2, N3, N4, N5 and N6 are SR-MPLS capable nodes.

N100 is a controller.

<L1, L2, ..., Ln> represents a MPLS Label stack for an SR policy where L1 is the first Label and Ln is the last Label as defined in Section 4 of [I-D.spring-segment-routing-policy].

SR policy is defined in Section 3 of [I-D.spring-segment-routing-policy].

3. Probe Packets

3.1. Probe Packets for SR-MPLS Policies

As described in [RFC6374], Section 2.9.1, MPLS PM probe messages flow over the MPLS Generic Associated Channel (G-ACh). A probe packet for an SR policy contains SR-MPLS label stack [I-D.spring-segment-routing-policy], with the G-ACh Label (GAL) at the bottom of the stack. The GAL is followed by an Associated Channel Header (ACH), which identifies the message type, and the

message body following the ACH as shown in Figure 2.

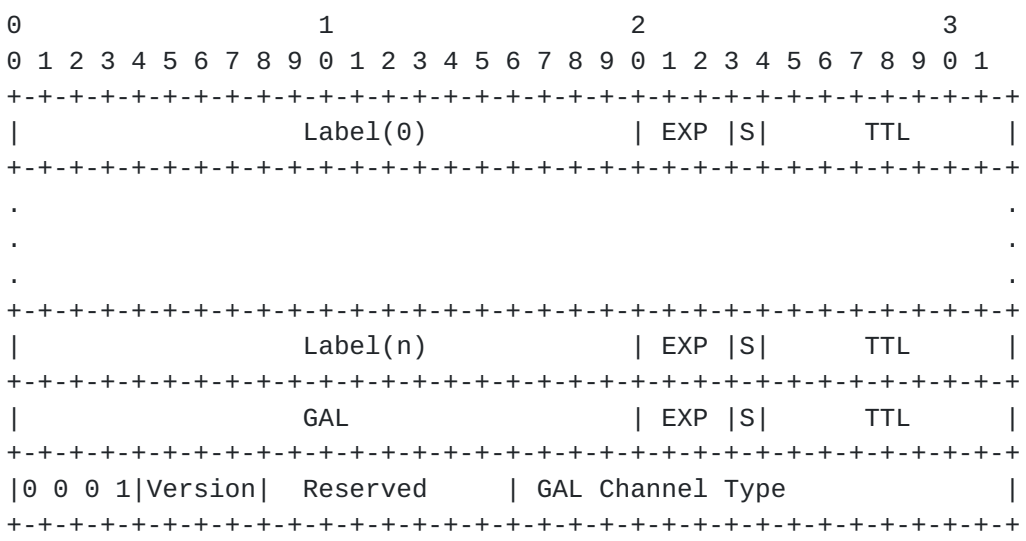


Figure 2: Probe Packet Header for an SR-MPLS Policy

3.2. Probe Packets for SR-MPLS Links

As described in [RFC6374], Section 2.9.1, MPLS PM probe messages flow over the MPLS Generic Associated Channel (G-ACh). A probe packet for SR-MPLS links contains G-ACh Label (GAL). The GAL is followed by an Associated Channel Header (ACH), which identifies the message type, and the message body following the ACH as shown in Figure 3.

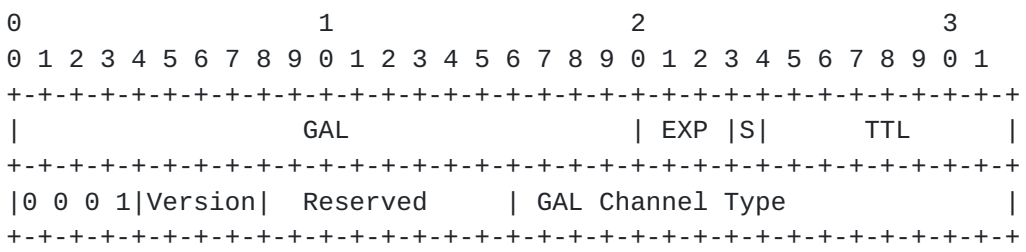


Figure 3: Probe Packet Header for an SR-MPLS Link

3.3. Probe Reply Message

3.3.1. One-way Measurement Probe Reply

For one-way performance measurement [RFC7679], the PM querier node can receive "out-of-bands" probe replies by properly setting the UDP Return Object (URO) TLV in the probe message. The URO TLV (Type=131)

is defined in [RFC7876] and includes the UDP-Destination-Port and IP Address. In particular, if the querier sets its own IP address in the URO TLV the probe response is sent back by the responder node to the querier node.

3.3.1.1. Probe Reply Message to Controller

As shown in Figure 1, if the querier node N1 requires the probe reply to be sent to the controller N100, it sets the IP address of N100 in the Address field of the URO TLV of the PM probe query message.

3.3.2. Two-way Measurement Probe Reply

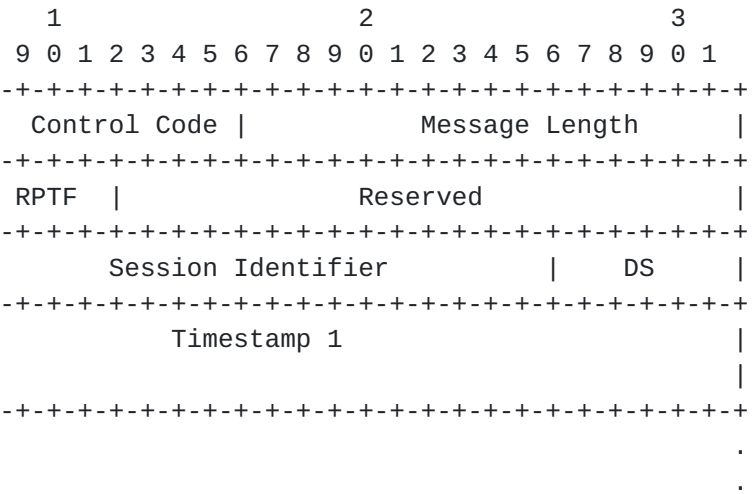
For two-way performance measurement [RFC6374], when using a bidirectional channel, the probe reply message is sent back to the querier node using a message similar to the probe query message as an SR-MPLS packet. In this case, the "control code" in the probe message is set to "in-band response requested".

4. Performance Delay Measurement

4.1. Delay Measurement Message Format

As defined in [RFC6374], MPLS DM probe messages use Associated Channel Header (ACH) (value 0x000C for delay measurement) [RFC6374], which identifies the message type, and the message body following the ACH. For both SR-MPLS policies and links, the same MPLS DM ACH value is used.

The DM message payload as defined in [RFC6374] is used for SR-MPLS delay measurement, for both SR-MPLS policies and SR-MPLS links. The DM message payload format is defined as following in [RFC6374]:



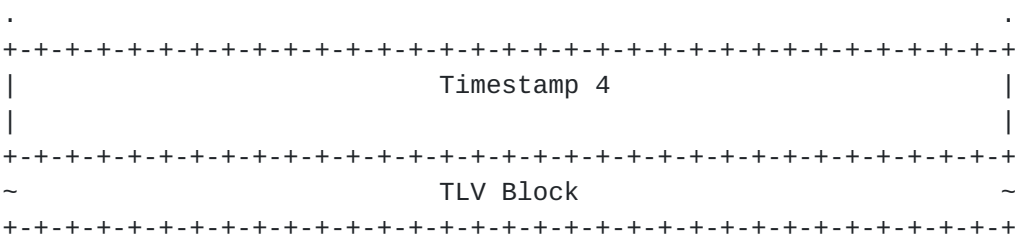


Figure 4: Delay Measurement Message Format

The meanings of the fields are summarized in the following table, see [RFC6374] for details.

Field	Meaning
Version	Protocol version
Flags	Message control flags
Control Code	Code identifying the query or response type
QTF	Querier timestamp format (see Section 3.4 in [RFC6374])
RTF	Responder timestamp format (see Section 3.4 in [RFC6374])
RPTF	Responder's preferred timestamp format
Reserved	Reserved for future specification
Session Identifier	Set arbitrarily by the querier
Differentiated Services (DS) Field	Differentiated Services Code Point (DSCP) being measured
Timestamp 1-4	64-bit timestamp values (see Section 3.4 in [RFC6374])
TLV Block	Optional block of Type-Length-Value fields

4.2. Timestamping

[RFC6374], Section 3.4 defines timestamp format that can be used for delay measurement. The IEEE 1588 Precision Time Protocol (PTP) timestamp format [IEEE1588] is used by default as described in Appendix A of [RFC6374], but it may require hardware support. As an

alternative, Network Time Protocol (NTP) timestamp format is also supported in [RFC6374].

Note that for one-way delay measurement, Clock synchronization between the querier and responder nodes using methods detailed in [RFC6374] is required. Two-way delay measurement does not require clock to be synchronized between the querier and responder nodes.

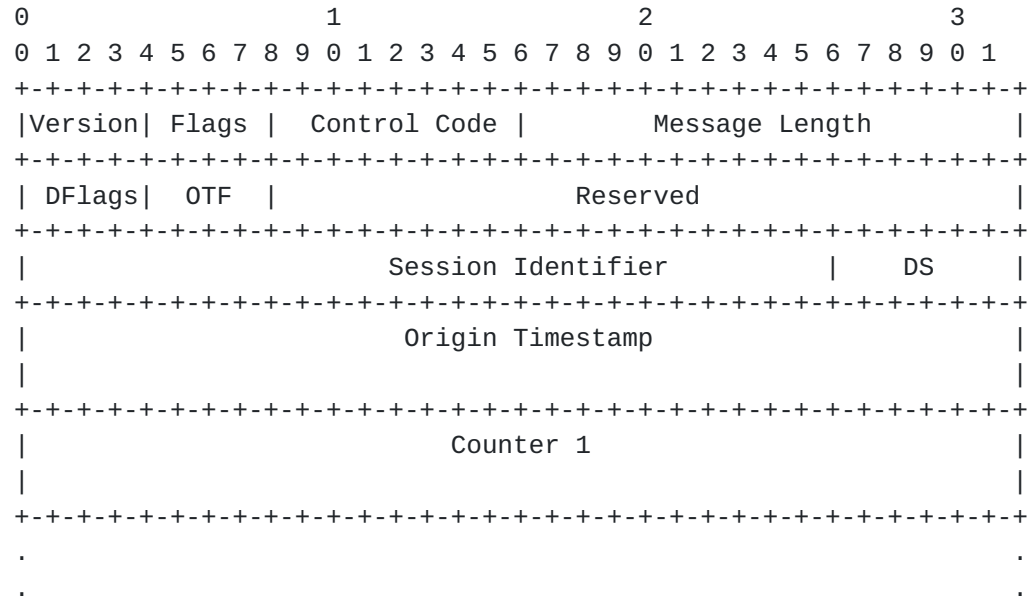
5. Performance Loss Measurement

The performance LM protocol can perform two distinct kinds of loss measurement as described in [RFC6374], Section 2.9.8. In inferred mode, LM will measure the loss of specially generated test messages in order to infer the approximate data plane loss level. Inferred loss measurement provides only approximate loss accounting. In direct mode, LM will directly measure data plane packet loss. Direct loss measurement provides perfect loss accounting, but may require hardware support.

5.1. Loss Measurement Message Format

As defined in [RFC6374], MPLS LM probe messages use Associated Channel Header (ACH) (value 0x000A for direct loss measurement or value 0x000B for inferred loss measurement), which identifies the message type, and the message body following the ACH. For both SR-MPLS policies and SR-MPLS links, the same MPLS LM ACH value is used.

The LM message payload as defined in [RFC6374] is used for SR-MPLS delay measurement, for both SR-MPLS policies and SR-MPLS links. The LM message payload format is defined as following in [RFC6374]:



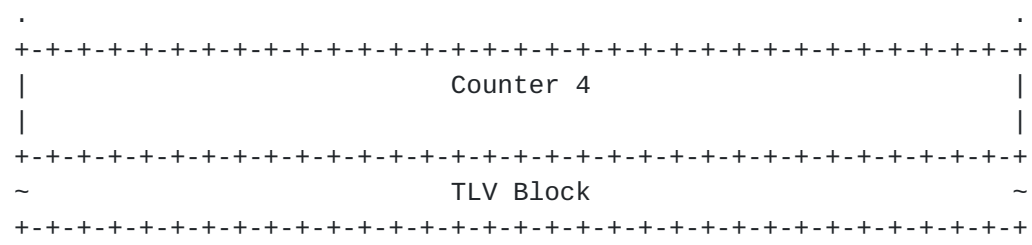


Figure 5: Loss Measurement Message Format

The meanings of the fields are summarized in the following table, see [RFC6374] for details.

Field	Meaning
Version	Protocol version
Flags	Message control flags
Control Code	Code identifying the query or response type
Message Length	Total length of this message in bytes
Data Format Flags (DFlags)	Flags specifying the format of message data
Origin Timestamp Format (OTF)	Format of the Origin Timestamp field
Reserved	Reserved for future specification
Session Identifier	Set arbitrarily by the querier
Differentiated Services (DS) Field	Differentiated Services Code Point (DSCP) being measured
Origin Timestamp	64-bit field for query message transmission timestamp
Counter 1-4	64-bit fields for LM counter values
TLV Block	Optional block of Type-Length-Value fields

6. SR-MPLS Link Metrics Advertisements

Performance metrics for link delay and packet loss calculated using the performance measurement procedures reviewed in this document can

be advertised in the routing domain. For OSPF, ISIS and BGP-LS, protocol extensions defined in [RFC7471], [RFC7810] and [I-D.idr-te-pm-bgp] are used, respectively for advertising the link delay metrics (minimum-delay, maximum-delay, average-delay, delay-variance) and loss metric in the network.

7. Security Considerations

This document reviews the procedures for performance measurement for SR-MPLS networks, for both SR-MPLS policies and SR-MPLS links, using the mechanisms defined in [RFC6374]. This document does not introduce any additional security considerations other than those covered in [RFC6374].

8. IANA Considerations

This document does not require any IANA actions.

9. References

9.1. Normative References

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

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS networks', RFC 6374, September 2011.

[RFC7876] Bryant, S., Sivabalan, S., and Soni, S., "UDP Return Path for Packet Loss and Delay Measurement for MPLS Networks", RFC 7876, July 2016.

9.2. Informative References

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protoco (OWAMP)", RFC 4656, September 2006.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, October 2008.

[RFC7679] Almes, G., et al., "A One-Way Delay Metric for IP Performance Metrics (IPPM)', RFC 7679, January 2016.

[RFC7471] Giacalone, S., et al., "OSPF Traffic Engineering (TE)

Metric Extensions", RFC 7471, March 2015.

[RFC7810] Previdi, S., et al., "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, May 2016.

[I-D.idr-te-pm-bgp] Ginsberg, L. Ed., et al., "BGP-LS Advertisement of IGP Traffic Engineering Performance Metric Extensions", draft-ietf-idr-te-pm-bgp, work in progress.

[I-D.spring-segment-routing-policy] Filsfils, C., et al., "Segment Routing Policy for Traffic Engineering", draft-filsfils-spring-segment-routing-policy, work in progress.

Acknowledgments

To be added.

Contributors

To be added.

Authors' Addresses

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

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

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

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

Daniel Bernier
Bell Canada
Email: daniel.bernier@bell.ca

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

Pier Luigi Ventre
CNIT
Italy
Email: pierluigi.ventre@cnit.it