Workgroup: MPLS Working Group
Internet-Draft:
draft-ietf-mpls-inband-pm-encapsulation-03
Published: 1 July 2022
Intended Status: Standards Track
Expires: 2 January 2023
Authors: W. Cheng (China Mobile)
X. Min, Ed. (ZTE Corp.)
T. Zhou (Huawei)
X. Dong (FiberHome)
Y. Peleg (Broadcom)

Encapsulation For MPLS Performance Measurement with Alternate Marking Method

Abstract

This document defines the encapsulation for MPLS performance measurement with alternate marking method, which performs flow-based packet loss, delay, and jitter measurements on live traffic.

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 2 January 2023.

Table of Contents

1. Introduction

[I-D.ietf-ippm-rfc8321bis] describes a performance measurement method, which can be used to measure packet loss, delay, and jitter on live traffic. Since this method is based on marking consecutive batches of packets, the method is often referred to as Alternate Marking Method. [RFC8372] discusses the desired capabilities for MPLS flow identification, in order to perform a better in-band performance monitoring of user data packets.

This document defines the encapsulation for MPLS performance measurement with alternate marking method, which performs flow-based packet loss, delay, and jitter measurements on live traffic. The encapsulation defined in this document supports monitoring at intermediate nodes, as well as flow identification at both transport and service layers.

This document employs a method, other than Synonymous Flow Label (SFL), to accomplish MPLS flow identification. The method described in this document is complementary to the SFL method [RFC8957] [I-D.ietf-mpls-sfl-control], the former mainly aims at hop-by-hop performance measurement, and the latter mainly aims at edge-to-edge performance measurement. Different sets of flows may use different methods.

The method described in this document is also complementary to the In-situ OAM method [RFC9197] [I-D.ietf-ippm-ioam-direct-export], the former doesn't introduce any new header whereas the latter introduces a new In-situ OAM header, furthermore, the former requests the network nodes to report the data used for performance measurement, and the latter requests the network nodes to report the data used for operational and telemetry information collection. One set of flows may use both of the two methods concurrently.

1.1. Conventions Used in This Document

1.1.1. Abbreviations

ACL: Access Control List

bSPL: Base Special Purpose Label

ECMP: Equal-Cost Multipath

ELC: Entropy Label Capability

ERLD: Entropy Readable Label Depth

FLC: Flow-ID Label Capability

FLI: Flow-ID Label Indicator

FRLD: Flow-ID Readable Label Depth

LSP: Label Switched Path

MPLS: Multi-Protocol Label Switching

NMS: Network Management System

PHP: Penultimate Hop Popping

PM: Performance Measurement

PW: PseudoWire

SFL: Synonymous Flow Label

SID: Segment ID

SR: Segment Routing

TC: Traffic Class

TTL: Time to Live

VC: Virtual Channel

VPN: Virtual Private Network

1.1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Flow-based PM Encapsulation in MPLS

Flow-based MPLS performance measurement encapsulation with alternate marking method has the following format:

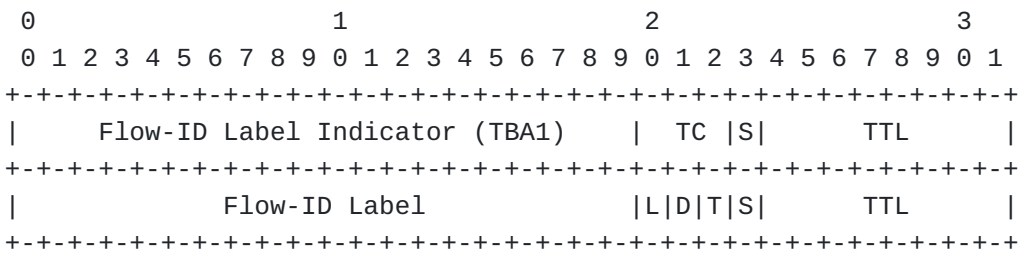

Figure 1: Flow-based PM Encapsulation in MPLS

Flow-ID Label Indicator (FLI) is a Base Special Purpose Label (bSPL) as defined in [RFC9017]. The FLI is defined in this document as value TBA1.

Traffic Class (TC) and Time to Live (TTL) for the FLI SHOULD follow the same field values of that label immediately preceding the FLI.

Flow-ID label is used as MPLS flow identification [RFC8372], its value MUST be unique within the administrative domain. Flow-ID values can be allocated by an external NMS/controller, based on measurement object instance such as LSP or PW. There is a one-to-one mapping between Flow-ID and flow. The specific method on how to allocate the Flow-ID values is described in Section 4.

Flow-ID label can be placed at either the bottom or the middle of the MPLS label stack, and the Flow-ID label MAY appear multiple times in a label stack. Section 2.1 of this document provides several examples to illustrate how to apply Flow-ID label in a label stack. TTL for the Flow-ID label MUST be zero to ensure that it is not used inadvertently for forwarding. S bit for the Flow-ID Label depends on whether the Flow-ID label is placed at the bottom of the MPLS label stack.

Besides flow identification, a color-marking field is also necessary for alternate marking method. To achieve the purpose of coloring the MPLS traffic, as well as the distinction between hop-by-hop measurement and edge-to-edge measurement, TC for the Flow-ID label is defined as follows:

*L(oss) bit is used for coloring the MPLS packets for loss measurement.

*D(elay) bit is used for coloring the MPLS packets for delay/ jitter measurement.

*T(ype) bit is used to indicate the measurement type. When T bit is set to 1, it means edge-to-edge performance measurement. When T bit is set to 0, it means hop-by-hop performance measurement.

2.1. Examples for Applying Flow-ID Label in a label stack

Three examples on different layout of Flow-ID label (4 octets) are illustrated as follows:

(1) Layout of Flow-ID label when applied to MPLS transport.

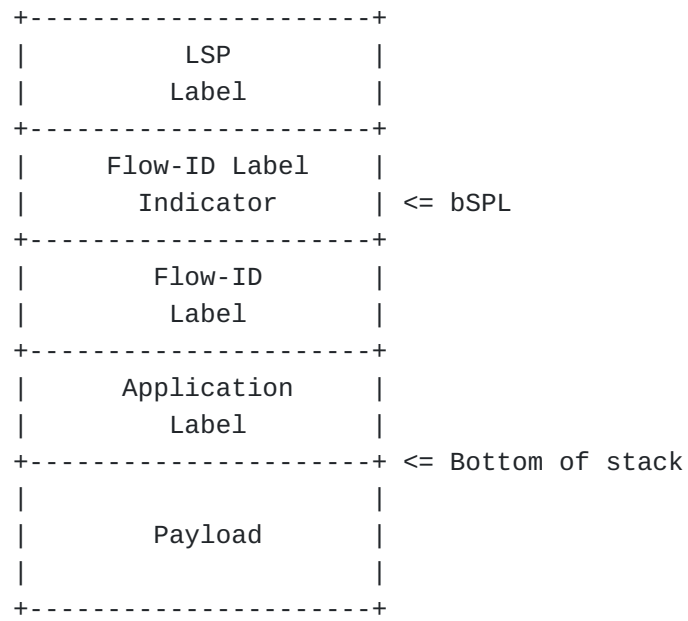

Figure 2: Applying Flow-ID to MPLS transport

Note that here if penultimate hop popping (PHP) is in use, the PHP LSR that recognizes the bSPL MAY choose not to pop the bSPL and the following Flow-ID label, otherwise the egress LSR would be excluded from the performance measurement.

Also note that in other examples of applying Flow-ID to MPLS transport, one LSP label can be substituted by multiple SID labels in the case of using SR Policy, and the combination of bSPL and Flow-ID label can be placed between SID labels, as specified in Section 5.

(2) Layout of Flow-ID label when applied to MPLS service.

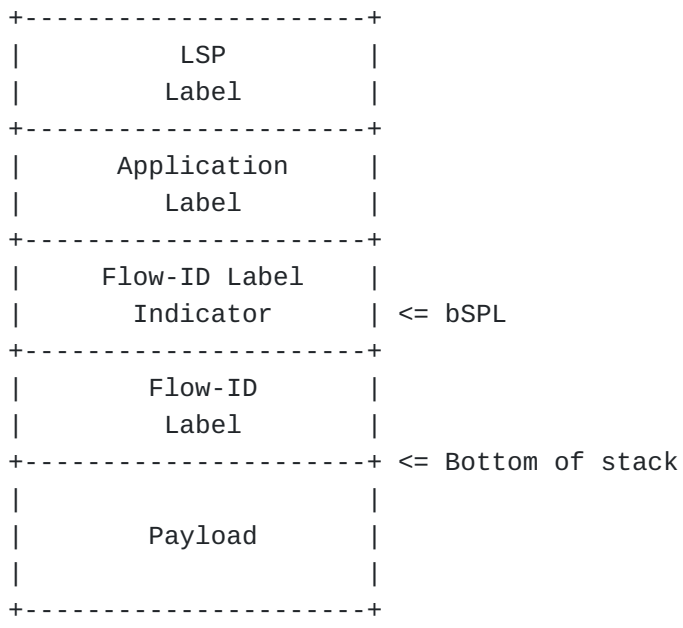

Figure 3: Applying Flow-ID to MPLS service

Note that here application label can be MPLS PW label, MPLS Ethernet VPN label or MPLS IP VPN label, and it's also called VC label as defined in [RFC4026].

(3) Layout of Flow-ID label when applied to both MPLS transport and MPLS service.

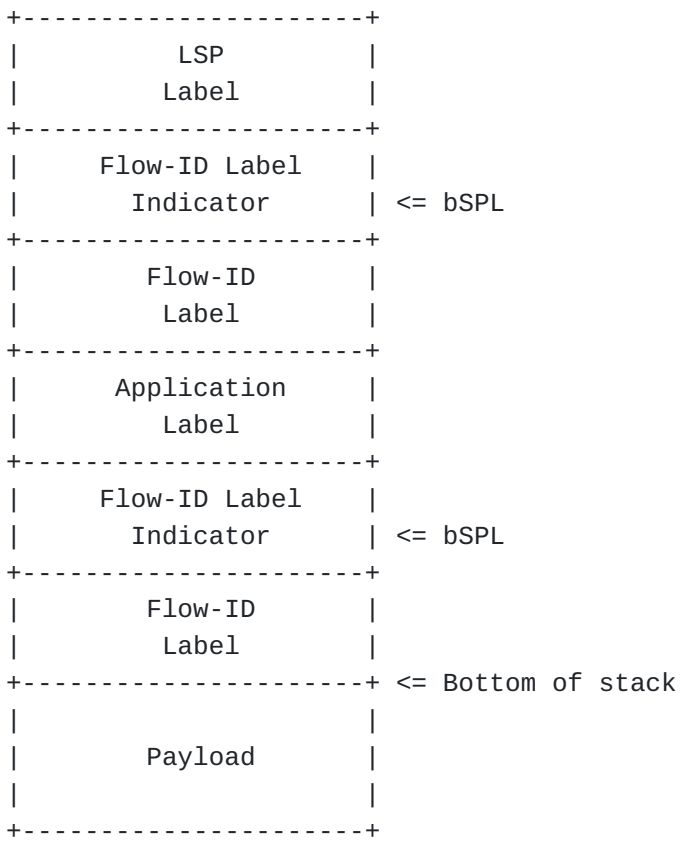

Figure 4: Applying Flow-ID to both MPLS transport and MPLS service

Note that for this example the two Flow-ID values appearing in a label stack MUST be different, that is to say, Flow-ID label applied to MPLS transport and Flow-ID label applied to MPLS service share the same value space. Also note that the two Flow-ID label values are independent from each other, e.g., two packets can belong to the same VPN flow but different LSP flows, or two packets can belong to two different VPN flows but the same LSP flow.

3. Procedures of Encapsulation, Look-up and Decapsulation

The procedures for Flow-ID label encapsulation, look-up and decapsulation are summarized as follows:

*The ingress node inserts the Flow-ID Label Indicator and the Flow-ID label into the MPLS label stack. At the same time, the ingress node sets the Flow-ID value, two color-marking bits and the T bit, as defined in this document.

*If the hop-by-hop measurement is applied, i.e., the T bit is set to 0, then whether the transit node or the egress node is the processing node. If the edge-to-edge measurement is applied, i.e., the T bit is set to 1, then only the egress node is the processing node. The processing node looks up the Flow-ID label

with the help of the Flow-ID Label Indicator, and exports the collected data, such as the Flow-ID, block counters and timestamps, to an external NMS/controller, referring to the alternate marking method. Note that while looking up the Flow-ID label, the transit node needs to perform some deep packet inspection beyond the label (at the top of the label stack) used to take forwarding decisions.

*The processing node may also pop the Flow-ID Label Indicator and the Flow-ID label from the MPLS label stack. The egress node pops the whole MPLS label stack, and this document doesn't introduce any new process to the decapsulated packet.

4. Procedures of Flow-ID allocation

There are two ways of allocating Flow-ID, one way is to allocate Flow-ID by manual trigger from the network operator, and the other way is to allocate Flow-ID by automatic trigger from the ingress node, details are as follows:

*In the case of manual trigger, the network operator would manually input the characteristics (e.g. IP five tuples and IP DSCP) of the measured flow, then the NMS/controller would generate one or two Flow-IDs based on the input from the network operator, and provision the ingress node with the characteristics of the measured flow and the corresponding allocated Flow-ID(s).

*In the case of automatic trigger, the ingress node would identify the flow entering the measured path, export the characteristics of the identified flow to the NMS/controller by IPFIX [RFC7011], then the NMS/controller would generate one or two Flow-IDs based on the characteristics exported from the ingress node, and provision the ingress node with the characteristics of the identified flow and the corresponding allocated Flow-ID(s).

The policy pre-configured at the NMS/controller decides whether one Flow-ID or two Flow-IDs would be generated. If the performance measurement on MPLS service is enabled, then one Flow-ID applied to MPLS service would be generated; If the performance measurement on MPLS transport is enabled, then one Flow-ID applied to MPLS transport would be generated; If both of them are enabled, then two Flow-IDs respectively applied to MPLS service and MPLS transport would be generated, in this case the transit node needs to look up both of the two Flow-IDs by default, and that can be changed by configuration to, e.g., look up only the Flow-ID applied to MPLS transport.

Whether using manual trigger or automatic trigger, the NMS/ controller MUST guarantee every generated Flow-ID is unique within the administrative domain.

5. FLC and FRLD Considerations

Analogous to the Entropy Label Capability (ELC) defined in Section 5 of [RFC6790] and the Entropy Readable Label Depth (ERLD) defined in Section 4 of [RFC8662], the Flow-ID Label Capability (FLC) and the Flow-ID Readable Label Depth (FRLD) are defined in this document. Both FLC and FRLD have the similar semantics with ELC and ERLD to a router, except that the Flow-ID is used in its flow identification function while the Entropy is used in its load-balancing function.

The ingress node MUST insert each Flow-ID label at an appropriate depth, which ensures the node to which the Flow-ID label is exposed has the FLC. The ingress node SHOULD insert each Flow-ID label within an appropriate FRLD, which is the minimum FRLD of all on-path nodes that need to read and use the Flow-ID label in question. How the ingress node knows the FLC and FRLD of all on-path nodes is outside the scope of this document, whereas [I-D.xzc-lsr-mpls-flc-frld] provides a method to achieve that.

When SR paths are used as transport, the label stack grows as the number of on-path segments increases, if the number of on-path segments is high, that may become a challenge for the Flow-ID label to be placed within an appropriate FRLD. In order to overcome this potential challenge, an implementation MAY provide flexibility to the ingress node to place Flow-ID label between SID labels, i.e., multiple identical Flow-ID labels at different depths MAY be interleaved with SID labels, when that happens a sophisticated network planning may be needed and it's beyond the scope of this document.

6. Equal-Cost Multipath Considerations

Analogous to what's described in Section 5 of [RFC8957], under conditions of Equal-Cost Multipath (ECMP), the introduction of Flow-ID label may lead to the same problem as caused by SFL, and the two solutions proposed for SFL would also apply here.

7. Security Considerations

This document introduces the performance measurement domain that is the scope of a Flow-ID label. The Flow-ID Label Indicator and Flow-ID label MUST NOT be signaled and distributed outside one performance measurement domain. Improper configuration so that the Flow-ID label being passed from one domain to another would likely result in potential Flow-ID conflicts.

To prevent packets carrying Flow-ID label from leaking from one domain to another, the domain boundary nodes SHOULD deploy some policies (e.g., ACL) to filter out the packets. Specifically, in the sending edge, the domain boundary node SHOULD filter out the packets that carry the Flow-ID Label Indicator and are sent to other domain; in the receiving edge, the domain boundary node SHOULD drop the packets that carry the Flow-ID Label Indicator and are from other domains.

8. IANA Considerations

In the Special-Purpose MPLS Label Values registry, a new Base Special-Purpose MPLS Label Value for Flow-ID Label Indicator is requested from IANA as follows:

Base Special-Purpose MPLS Label Value	Description	Semantics Definition	Reference
TBA1 (12 is recommended)	Flow-ID Label Indicator	Section 2	This Document

Table 1: New Base Special-Purpose MPLS Label Value for Flow-ID Label Indicator

9. Acknowledgements

The authors would like to acknowledge Loa Andersson, Tarek Saad, Stewart Bryant, Rakesh Gandhi, Greg Mirsky, Aihua Liu, Shuangping Zhan and Ming Ke for their careful review and very helpful comments.

The authors would like to acknowledge Italo Busi and Chandrasekar Ramachandran for their insightful MPLS-RT review and very helpful comments.

10. Contributors

Minxue Wang
China Mobile
Email: wangminxue@chinamobile.com

Wen Ye
China Mobile
Email: yewen@chinamobile.com

11. References

11.1. Normative References

[I-D.ietf-ippm-rfc8321bis] Fioccola, G., Cociglio, M., Mirsky, G., Mizrahi, T., and T. Zhou, "Alternate-Marking Method", Work in Progress, Internet-Draft, draft-ietf-ippm-

rfc8321bis-02, 7 June 2022, <https://www.ietf.org/archive/id/draft-ietf-ippm-rfc8321bis-02.txt>.

[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>.

11.2. Informative References

[I-D.ietf-ippm-ioam-direct-export] Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In-situ OAM Direct Exporting", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-direct-export-09, 15 June 2022, <https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-direct-export-09.txt>.

[I-D.ietf-mpls-sfl-control] Bryant, S., Swallow, G., and S. Sivabalan, "A Simple Control Protocol for MPLS SFLs", Work in Progress, Internet-Draft, draft-ietf-mpls-sfl-control-02, 21 January 2022, <https://www.ietf.org/archive/id/draft-ietf-mpls-sfl-control-02.txt>.

[I-D.xzc-lsr-mpls-flc-frld] Min, X., Zhang, Z., and W. Cheng, "Signaling Flow-ID Label Capability and Flow-ID Readable Label Depth", Work in Progress, Internet-Draft, draft-xzc-lsr-mpls-flc-frld-00, 15 February 2022, <https://www.ietf.org/archive/id/draft-xzc-lsr-mpls-flc-frld-00.txt>.

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, DOI 10.17487/RFC4026, March 2005, <https://www.rfc-editor.org/info/rfc4026>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC8372] Bryant, S., Pignataro, C., Chen, M., Li, Z., and G. Mirsky, "MPLS Flow Identification Considerations", RFC 8372, DOI 10.17487/RFC8372, May 2018, <https://www.rfc-editor.org/info/rfc8372>.

[RFC8662] Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and J. Tantsura, "Entropy Label for Source Packet Routing in Networking (SPRING) Tunnels", RFC 8662, DOI 10.17487/RFC8662, December 2019, <https://www.rfc-editor.org/info/rfc8662>.

[RFC8957] Bryant, S., Chen, M., Swallow, G., Sivabalan, S., and G. Mirsky, "Synonymous Flow Label Framework", RFC 8957, DOI 10.17487/RFC8957, January 2021, <https://www.rfc-editor.org/info/rfc8957>.

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

[RFC9197] Brockners, F., Ed., Bhandari, S., Ed., and T. Mizrahi, Ed., "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", RFC 9197, DOI 10.17487/RFC9197, May 2022, <https://www.rfc-editor.org/info/rfc9197>.

Authors' Addresses

Weiqiang Cheng
China Mobile
Beijing
China

Email: chengweiqiang@chinamobile.com

Xiao Min (editor)
ZTE Corp.
Nanjing
China

Email: xiao.min2@zte.com.cn

Tianran Zhou
Huawei
Beijing
China

Email: zhoutianran@huawei.com

Ximing Dong
FiberHome
Wuhan
China

Email: dxm@fiberhome.com

Yoav Peleg
Broadcom
United States of America

Email: yoav.peleg@broadcom.com