

MPLS
Internet-Draft
Intended status: Standards Track
Expires: 14 February 2026

H. Song
Futurewei Technologies
G. Fioccola
Huawei Technologies
R. Gandhi
Cisco Systems
13 August 2025

MPLS On-Path Telemetry Network Action Flag for OAM
draft-song-mpls-on-path-telemetry-flag-00

Abstract

This document describes the postcard-based on-path telemetry with packet marking (PBT-M) using an MPLS Network Actions (MNA) flag to support OAM in MPLS networks. The scheme uses a single bit in the MNA header opcode dedicated for the flag-based actions. The document provides the solutions to address the requirements for applying PBT-M in MPLS networks.

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 14 February 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components

extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
1.1. Requirements Language	3
2. PBT-M: Direct Export for On-path Telemetry based on Packet Marking	3
3. New Requirements	5
4. Design Considerations	6
4.1. Packet Marking	6
4.2. Flow Path Discovery	6
4.3. Packet Identity for Export Data Correlation	7
4.4. Load Control	7
5. Implementation Recommendation	8
5.1. Configuration	8
5.2. Data Export	8
5.3. Use Cases	8
6. Security Considerations	8
7. IANA Considerations	9
8. Acknowledgments	9
9. References	9
9.1. Normative References	9
9.2. Informative References	9
Authors' Addresses	10

1. Introduction

To gain detailed data plane visibility to support effective network OAM, it is essential to be able to examine the trace of user packets along their forwarding paths. Such on-path flow data reflect the state and status of each user packet's real-time experience and provide valuable information for network monitoring, measurement, and diagnosis.

The telemetry data include but not limited to the detailed forwarding path, the timestamp/latency at each network node, and, in case of packet drop, the drop location as well as the reason. The emerging programmable data plane devices allow user-defined data collection or conditional data collection based on trigger events. Such on-path flow data are from and about the live user traffic, which complements the data acquired through other passive and active OAM mechanisms such as IPFIX [RFC7011] and ICMP [RFC4560].

On-path telemetry was developed to cater to the need of collecting on-path flow data. There are two basic modes for on-path telemetry: the passport mode (e.g., IOAM trace option [RFC9197]) and the postcard mode (e.g., IOAM direct export option (DEX) [RFC9326]).

In MPLS networks, MPLS Network Action (MNA) [I-D.ietf-mpls-mna-fwk] extends the MPLS label stack by supporting extra in-stack network actions and ancillary data encoded in stack, the in-stack MNA header is described in [I-D.ietf-mpls-mna-hdr]. MNA also extends the MPLS payload by supporting extra post-stack network actions and ancillary data encoded post-stack, the post-stack MNA header is described in [I-D.jags-mpls-ps-mna-hdr].

This document describes the method to apply a new variation of the postcard mode on-path telemetry, PBT-M, to MPLS network using an MNA flag only. PBT-M does not require a telemetry instruction header but a single trigger bit in MNA flags. The similar mechanism has been adopted by SRv6 for SRv6 OAM [RFC9259], which uses the O-bit in SRH flags as the marking bit to trigger the on-path telemetry. The key benefits of PBT-M are its low overhead and high flexibility. However, PBT-M introduces some unique requirements that need to be considered. This document discusses the requirements and their solutions in MPLS networks.

1.1. 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. PBT-M: Direct Export for On-path Telemetry based on Packet Marking

As the name suggests, PBT-M only needs a marking-bit in the existing headers of user packets to trigger the telemetry data collection and export. The sketch of PBT-M is as follows. If on-path data need to be collected, the user packet is marked at the path head node. At each PBT-M-aware node, if the mark is detected and data collection is enabled, a postcard packet (i.e., the dedicated OAM packet triggered by a marked user packet) is generated and sent to a collector. The postcard contains the data requested by the management plane. The requested data are configured by the management plane. Once the collector receives all the postcards for a single user packet, it can infer the packet's forwarding path and analyze the data set. The path end node is configured to un-mark the packets to its original format if necessary.

The overall architecture of PBT-M is depicted in Figure 1.

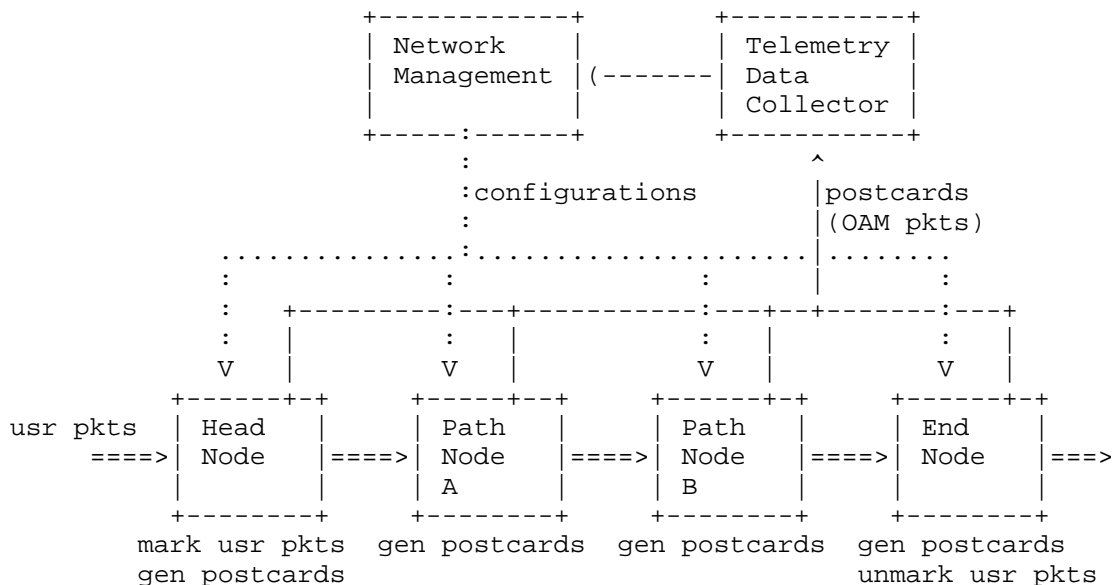


Figure 1: Architecture of PBT-M

The advantages of PBT-M are summarized as follows.

- * 1: PBT-M avoids augmenting user packets with new headers and the signaling for telemetry data collection remains in the data plane.
- * 2: PBT-M is extensible for collecting arbitrary new data to support possible future use cases. The data set to be collected can be configured through the management plane or control plane.
- * 3: PBT-M can avoid interfering with the normal forwarding. The collected data are free to be transported independently through in-band or out-of-band channels. The data collecting, processing, assembly, encapsulation, and transport are, therefore, decoupled from the forwarding of the corresponding user packets and can be performed in data-plane slow-path if necessary.
- * 4: For PBT-M, the types of data collected from each node can vary depending on application requirements and node capability.

- * 5: PBT-M makes it easy to secure the collected data without exposing it to unnecessary entities. For example, both the configuration and the telemetry data can be encrypted and/or authenticated before being transported, so passive eavesdropping and a man-in-the-middle attack can both be deterred.
- * 6: Even if a user packet under inspection is dropped at some node in the network, the postcards collected from the preceding nodes are still valid and can be used to diagnose the packet drop location and reason.
- * 7: Raw data can be processed or aggregated in data plane to reduce the exporting traffic load.

3. New Requirements

Although PBT-M has some unique features, it also introduces a few new requirements.

- * Req. 1 (Packet Marking Bit): A user packet needs to be marked to trigger the path-associated data collection. Since PBT-M aims to avoid the need to augment user packets with new headers, it needs to reserve or reuse a single bit from the existing header fields.
- * Req. 2 (Configuration): Since the packet header will not carry telemetry instructions anymore, the data plane devices need to be configured to know what data to collect. However, in general, the forwarding path of a flow packet (due to ECMP or dynamic routing) is unknown beforehand (note that there are some notable exceptions, such as segment routing). If the per-flow customized data collection is required, configuring the data set for each flow at all data plane devices might be expensive in terms of configuration load and data plane resources.
- * Req. 3 (Data Correlation): Due to the variable transport latency, the dedicated postcard packets for a single packet may arrive at the collector out of order or be dropped in networks for some reason. In order to infer the packet forwarding path, the collector needs some information from the postcard packets to identify the user packet affiliation and the order of path node traversal.
- * Req. 4 (Overhead and Security): Since each postcard packet has its header, the overall network bandwidth overhead of PBT-M can be high. A large number of postcards could add processing pressure on data collecting servers. That can be used as an attack vector for DoS.

4. Design Considerations

To address the above requirements, we propose several design details for applying PBT-M in MPLS networks.

4.1. Packet Marking

To trigger the path-associated data collection, usually, a single bit from some header field is sufficient. The proposed action encoding is shown in Figure 2 changed from [I-D.ietf-mpls-mna-hdr].

```

    0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+-----+-----+-----+-----+-----+-----+-----+-----+
|          NASI=bSPL          | TC | S |          TTL          |
+-----+-----+-----+-----+-----+-----+-----+-----+
| NAI-Opcode=1 | P |          | R | IHS | S | U | NASL=0 | NAL=0 |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 2: Action Encoding

In the figure, NAI-Opcode is Network Action Indicator Opcode. If the Opcode has the value of 'one', then the 13 bits following the Opcode carries NAI-flags. P-flag, the first flag bit in the flag-based network action field, is used as the PBT indicator. If the bit is set to '1', a node is triggered to collect and export the telemetry data as configured by the control plane.

Note that the in-stack MNA encoding may take different form as the header encoding draft evolves, and these flag-based on-path telemetry use cases would adapt to the change.

4.2. Flow Path Discovery

In case the path that a flow traverses is unknown in advance, all PBT-M-aware nodes should be configured to react to the marked packets by exporting some basic data, such as node ID and TTL before a data set template for that flow is configured. This way, the management plane can learn the flow path dynamically.

If the management plane wants to collect the on-path data for some flow, it configures the head node with a probability or time interval for the flow packet marking. When the first marked packet is forwarded in the network, the PBT-M-aware nodes will export the basic data set to the collector. Hence, the flow path is identified. If other data types need to be collected, the management plane can further configure the data set's template to the target nodes on the

flow's path. The PBT-M-aware nodes collect and export data accordingly if the packet is marked and a data set template is present.

If the flow path is changed for any reason, the new path can be quickly learned by the collector. Consequently, the management plane controller can be directed to configure the nodes on the new path. The outdated configuration can be automatically timed out or explicitly revoked by the management plane controller.

4.3. Packet Identity for Export Data Correlation

The collector needs to correlate all the postcard packets for a single user packet. Once this is done, the TTL (or the timestamp, if the network time is synchronized) can be used to infer the flow forwarding path. The key issue here is to correlate all the postcards for the same user packet.

The first possible approach includes the flow ID in the OAM packets. In case of MPLS, the MPLS label stack can serve as the flow ID. If the packet marking interval is large enough, the flow ID is enough to identify a user packet. As a result, it can be assumed that all the exported postcard packets for the same flow during a short time interval belong to the same user packet.

Alternatively, if the network is synchronized, then the flow ID plus the timestamp at each node can also infer the postcard affiliation. However, some errors may occur under some circumstances. For example, two consecutive user packets from the same flows are marked, but one exported postcard from a node is lost. It is difficult for the collector to decide to which user packet the remaining postcard is related. In many cases, such a rare error has no catastrophic consequence. Therefore it is tolerable.

4.4. Load Control

PBT-M should not be applied to all the packets all the time. It is better to be used in an interactive environment where the network telemetry applications dynamically decide which subset of traffic is under scrutiny. The network devices can limit the packet marking rate through sampling and metering. The postcard packets can be distributed to different servers to balance the processing load.

Because PBT-M sends telemetry data by dedicated postcard packets, it allows data aggregation and compression. Each node can process the generated raw data according to the configured local data-export policies. Such policies may specify how raw data is used to calculate performance metrics, e.g., max, min, mean, percentile, etc.

5. Implementation Recommendation

5.1. Configuration

Access lists with an optional sampler, [RFC5476], should be configured and attached at the ingress of the PBT-M encapsulation node's to select the intended flows for PTB-M.

Based on the requirements and node capability, the flow data could be exported at each transit node and at the end edge node with IPFIX [RFC7011].

5.2. Data Export

The data decomposition can be achieved on the PBT-M-aware node exporting the data or on the IPFIX data collection. [I-D.spiegel-ippm-ioam-rawexport] describes how data is being exported when decomposed at IPFIX data collection. When being decomposed on the PBT-M-aware node the data can be aggregated according to section 5 of [RFC7015].

5.3. Use Cases

In MPLS networks, MLD is a great concern which limits the MNA size and in turn the OAM capability. Moreover, for SR-MPLS, Maximum SID Depth(MSD) as well as PMTU in SR Policy are critical issues for SR path instantiation by a controller. PBT-M can become a critical solution to ensure that flexible on-path telemetry can be viable for operators by eliminating telemetry data from being carried in-situ in the SR-TE policy path.

This draft provides a critical optimization that fills the gaps with IOAM DEX related to packet marking triggers using existing mechanisms as well as flow path discovery mechanisms to avoid configuration of on path data plane node complexity and helps mitigate SR MSD and PMTU issues.

6. Security Considerations

Only the ingress node is allowed to set these flag bits. The other on-path nodes can only react to the bit values. The tampering of these flag-based actions would result in DoS attack or unreliable measurements. Therefore, security measures must be taken to ensure the proper functioning of these actions.

7. IANA Considerations

This document requires IANA to assign a bit for PBT-M action (TBA1) from the MPLS "In-Stack MPLS Network Action Indicator Flags" registry created in [I-D.ietf-mpls-mna-hdr].

8. Acknowledgments

9. References

9.1. Normative References

[I-D.ietf-mpls-mna-hdr]

Rajamanickam, J., Gandhi, R., Zigler, R., Song, H., and K. Kompella, "MPLS Network Action (MNA) Sub-Stack Solution", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-hdr-13, 20 July 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-hdr-13>>.

[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>>.

[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>>.

[RFC7015] Trammell, B., Wagner, A., and B. Claise, "Flow Aggregation for the IP Flow Information Export (IPFIX) Protocol", RFC 7015, DOI 10.17487/RFC7015, September 2013, <<https://www.rfc-editor.org/info/rfc7015>>.

[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>>.

9.2. Informative References

[I-D.ietf-mpls-mna-fwk]

Andersson, L., Bryant, S., Bocci, M., and T. Li, "MPLS Network Actions (MNA) Framework", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-fwk-15, 27 December 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-fwk-15>>.

[I-D.jags-mpls-ps-mna-hdr]

Rajamanickam, J., Gandhi, R., Zigler, R., Li, T., and J. Dong, "Post-Stack MPLS Network Action (MNA) Solution", Work in Progress, Internet-Draft, draft-jags-mpls-ps-mna-hdr-05, 20 March 2025, <<https://datatracker.ietf.org/doc/html/draft-jags-mpls-ps-mna-hdr-05>>.

[I-D.spiegel-ippm-ioam-rawexport]

Spiegel, M., Brockners, F., Bhandari, S., and R. Sivakolundu, "In-situ OAM raw data export with IPFIX", Work in Progress, Internet-Draft, draft-spiegel-ippm-ioam-rawexport-07, 12 February 2024, <<https://datatracker.ietf.org/doc/html/draft-spiegel-ippm-ioam-rawexport-07>>.

[RFC4560] Quittek, J., Ed. and K. White, Ed., "Definitions of Managed Objects for Remote Ping, Traceroute, and Lookup Operations", RFC 4560, DOI 10.17487/RFC4560, June 2006, <<https://www.rfc-editor.org/info/rfc4560>>.

[RFC5476] Claise, B., Ed., Johnson, A., and J. Quittek, "Packet Sampling (PSAMP) Protocol Specifications", RFC 5476, DOI 10.17487/RFC5476, March 2009, <<https://www.rfc-editor.org/info/rfc5476>>.

[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>>.

[RFC9259] Ali, Z., Filmsils, C., Matsushima, S., Voyer, D., and M. Chen, "Operations, Administration, and Maintenance (OAM) in Segment Routing over IPv6 (SRv6)", RFC 9259, DOI 10.17487/RFC9259, June 2022, <<https://www.rfc-editor.org/info/rfc9259>>.

[RFC9326] Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In Situ Operations, Administration, and Maintenance (IOAM) Direct Exporting", RFC 9326, DOI 10.17487/RFC9326, November 2022, <<https://www.rfc-editor.org/info/rfc9326>>.

Authors' Addresses

Haoyu Song
Futurewei Technologies
United States of America

Email: haoyu.song@futurewei.com

Giuseppe Fioccola
Huawei Technologies
Germany
Email: giuseppe.fioccola@huawei.com

Rakesh Gandhi
Cisco Systems
Canada
Email: rgandhi@cisco.com