Workgroup: MPLS Working Group
Internet-Draft:
draft-ietf-mpls-mna-usecases-02
Published: 13 March 2023
Intended Status: Informational
Expires: 14 September 2023
Authors: T. Saad, Cisco Systems, Inc.
K. Makhijani, Futurewei Technologies
H. Song, Futurewei Technologies
G. Mirsky, Ericsson

Use Cases for MPLS Network Action Indicators and MPLS Ancillary Data

Abstract

This document presents a number of use cases that have a common need for encoding network action indicators and associated ancillary data inside MPLS packets. There has been significant recent interest in extending the MPLS data plane to carry such indicators and ancillary data to address a number of use cases that are described in this document.

The use cases described in this document are not an exhaustive set, but rather the ones that are actively discussed by members of the IETF MPLS, PALS and DETNET working groups participating in the MPLS Open Design Team.

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 September 2023.

Copyright Notice

Copyright (c) 2023 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

This document describes important cases that require carrying additional ancillary data within the MPLS packets, as well as means to indicate the ancillary data is present, and a specific action needs to be performed on the packet.

These use cases have been identified by the MPLS Working Group Open Design Team working on defining MPLS Network Actions for the MPLS data plane. The MPLS Ancillary Data (AD) can be classified as:

* implicit, or "no-data" associated with a Network Action (NA) indicator,

* residing within the MPLS label stack and referred to as In Stack Data (ISD), and

* residing after the Bottom of MPLS label Stack (BoS) and referred to as Post Stack Data (PSD).

The use cases described in this document will be used to assist in identifying requirements and issues to be considered for future resolution by the working group.

1.1. Terminology

The following terminology is used in the document:

IETF Network Slice:
a well-defined composite of a set of endpoints, the connectivity requirements between subsets of these endpoints, and associated requirements; the term 'network slice' in this document refers to 'IETF network slice' as defined in [I-D.ietf-teas-ietf-network-slices].

Time-Bound Networking:
Networks that transport time-bounded traffic.

1.2. Acronyms and Abbreviations

ISD: In-stack data

PSD: Post-stack data

MNA: MPLS Network Action

NAI: Network Action Indicator

AD: Ancillary Data

2. Use Cases

2.1. No Further Fastreroute

MPLS Fast Reroute (FRR) [RFC4090], [RFC5286] and [RFC7490] is a useful and widely deployed tool for minimizing packet loss in the case of a link or node failure.

Several cases exist where, once FRR has taken place in an MPLS network and resulted in rerouting a packet away from the failure, a second FRR that impacts the same packet to rerouting is not helpful, and may even be disruptive.

For example, in such a case, the packet may continue to loop until its TTL expires. This can lead to link congestion and further packet loss. Thus, the attempt to prevent a packet from being dropped may instead affect many other packets. A proposal to address this is presented in [I-D.kompella-mpls-nffrr].

2.2. In-situ OAM

In-situ Operations, Administration, and Maintenance (IOAM) is used to collect operational and telemetry information while packets traverses a particular path in a network domain.

The term "in-situ" refers to the fact that the IOAM data fields are added to the data packets rather than being sent within the probe packets specifically dedicated to OAM or Performance Measurement (PM).

IOAM can run in two modes Edge-to-Edge (E2E) and Hop-by-Hop (HbH). In E2E mode, only the encapsulating and decapsulating nodes will process IOAM data fields. In HbH mode, the encapsulating and decapsulating nodes as well as intermediate IOAM-capable nodes process IOAM data fields. The IOAM data fields are defined in [I-D.ietf-ippm-ioam-data], and can be used for various OAM use-cases.

Several IOAM Options have been defined:

* Pre-allocated and Incremental

* Edge-to-Edge

* Proof-of-Transit

* Direct Export (see Section 2.2.1)

[I-D.gandhi-mpls-ioam-sr] defines how IOAM data fields are transported using the MPLS data plane encapsulations, including Segment Routing (SR) with MPLS data plane (SR-MPLS).

The IOAM data may be added after the bottom of the MPLS label stack. The IOAM data fields can be of fixed or incremental size as defined in [I-D.ietf-ippm-ioam-data]. [I-D.gandhi-mpls-ioam] describes the applicability of IOAM to MPLS dataplane. The encapsulating MPLS node needs to know if the decapsulating MPLS node can process the IOAM data before adding it in the packet. In HbH IOAM mode, nodes that are capable of processing IOAM will intercept and process the IOAM data accordingly. The presence of IOAM header and optional IOAM data will betransparent to nodes that do not support or do not participate in the IOAM process.

2.2.1. In-situ OAM Direct Export

IOAM Direct Export (DEX) [I-D.ietf-ippm-ioam-direct-export] is an IOAM Option-Type in which the operational state and telemetry information is collected according to the specified profile and exported in a manner and format defined by a local policy.

In IOAM DEX, the user data packet is only used to trigger the IOAM data to be directly exported or locally aggregated without being pushed into in-flight data packets.

2.3. Network Slicing

An IETF Network Slice service provides connectivity coupled with a set of network resource commitments and is expressed in terms of one or more connectivity constructs. A slice-flow aggregate

[I-D.bestbar-teas-ns-packet] refers to the set of traffic streams from one or more connectivity constructs belonging to one or more IETF Network Slices that are mapped to a set of network resources and provided the same forwarding treatment. The packets associated with a slice-flow aggregate may carry a marking in the packet's network layer header to identify this association and this marking is referred to as Flow-Aggregate Selector (FAS). The FAS is used to map the packet to the associated set of network resources and provide the corresponding forwarding treatment to the packet.

A router that requires forwarding of a packet that belongs to a slice-flow aggregate may have to decide on the forwarding action to take based on selected next-hop(s), and the forwarding treatment (e.g., scheduling and drop policy) to enforce based on the associated per-hop behavior.

In this case, the routers that forward traffic over resources that are shared by multiple slice-flow aggregates need to identify the slice aggregate packets in order to enforce the associated forwarding action and treatment.

MNA can be used to indicate the action and carry ancillary data for packets traversing Label Switched Paths (LSPs). An MNA network action can be used to carry the FAS in MPLS packets.

2.3.1. Dedicated Identifier as Flow-Aggregate Selector

A dedicated Identifier that is independent of forwarding can be carried in the packet as a Flow-Aggregate Selector (FAS). This can be encoded in the MPLS packet as defined in [I-D.kompella-mpls-mspl4fa], [I-D.li-mpls-enhanced-vpn-vtn-id], and [I-D.decraene-mpls-slid-encoded-entropy-label-id]. The FAS is used to associate the packets belonging to Slice-Flow Aggregate to the underlying Network Resource Partition (NRP) as described in [I-D.bestbar-teas-ns-packet].

When MPLS packets carry a dedicated FAS identifier, the MPLS LSRs use the forwarding label to select the forwarding next-hop(s), and use the FAS in the MPLS packet to infer the specific forwarding treatment that needs to be applied on the packet.

The FAS can be encoded within an MPLS label carried in the packet's MPLS label stack. All MPLS packets that belong to the same flow aggregate MAY carry the same FAS identifier.

2.3.2. Forwarding Label as a Flow-Aggregate Selector

[RFC3031] states in Section 2.1 that: 'Some routers analyze a packet's network layer header not merely to choose the packet's next hop, but also to determine a packet's "precedence" or "class of service"'.

It is possible by assigning a unique MPLS forwarding label to each flow aggregate (FEC) to distinguish the packets forwarded to the same destination. from other flow aggregates. In this case, LSRs can use the top forwarding label to infer both the forwarding action and the forwarding treatment to be invoked on the packets.

2.4. Generic Delivery Functions

The Generic Delivery Functions (GDF), defined in [I-D.zzhang-intarea-generic-delivery-functions], provide a new mechanism to support functions analogous to those supported through the IPv6 Extension Headers mechanism. For example, GDF can support fragmentation/reassembly functionality in the MPLS network by using the Generic Fragmentation Header. MNA can support GDF by placing a GDF header in an MPLS packet within the Post-Stack Data block [I-D.ietf-mpls-mna-fwk]. Multiple GDF headers can also be present in the same MPLS packet organized as a list of headers.

2.5. Delay Budgets for Time-Bound Applications

The routers in a network can perform two distinct functions on incoming packets, namely forwarding (where the packet should be sent) and scheduling (when the packet should be sent). IEEE-802.1 Time Sensitive Networking (TSN) and Deterministic Networking provide several mechanisms for scheduling under the assumption that routers are time-synchronized. The most effective mechanisms for delay minimization involve per-flow resource allocation.

Segment Routing (SR) is a forwarding paradigm that allows encoding forwarding instructions in the packet in a stack data structure, rather than being programmed into the routers. The SR instructions are contained within a packet in the form of a First-in First-out stack dictating the forwarding decisions of successive routers. Segment routing may be used to choose a path sufficiently short to be capable of providing a bounded end-to-end latency but does not influence the queueing of individual packets in each router along that path.

When carried over the MPLS data plane, a solution is required to enable the delivery of such packets that can be delivered to their final destination by a given time budget. One approach to address this usecase in SR-MPLS was described in [I-D.stein-srtsn].

2.5.1. Stack Based Methods for Latency Control

One efficient data structure for inserting local deadlines into the headers is a "stack", similar to that used in Segment Routing to carry forwarding instructions. The number of deadline values in the stack equals the number of routers the packet needs to traverse in the network, and each deadline value corresponds to a specific router. The Top-of-Stack (ToS) corresponds to the first router's deadline while the Bottom-of-Stack (BoS) refers to the last's. All local deadlines in the stack are later or equal to the current time (upon which all routers agree), and times closer to the ToS are always earlier or equal to times closer to the BoS.

The ingress router inserts the deadline stack into the packet headers; no other router needs to be aware of the requirements of the time-bound flows. Hence admitting a new flow only requires updating the information base of the ingress router.

MPLS LSRs that expose the Top of Stack (ToS) label can also inspect the associated "deadline" carried in the packet (either in MPLS stack as ISD or after BoS as PSD).

2.6. NSH-based Service Function Chaining

[RFC8595] describes how Service Function Chaining (SFC) can be realized in an MPLS network by emulating the NSH by using only MPLS label stack elements.

The approach in [RFC8595] introduces some limitations that are discussed in [I-D.lm-mpls-sfc-path-verification]. This approach, however, can benefit from the framework introduced with MNA in [I-D.ietf-mpls-mna-fwk].

For example, it may be possible to extend NSH emulation using MPLS labels [RFC8595] to support the functionality of NSH Context Headers, whether fixed or variable-length. One of the use cases could support Flow ID [I-D.ietf-sfc-nsh-tlv] that may be used for load-balancing among Service Function Forwarders (SFFs) and/or the Service Function (SF) within the same SFP.

2.7. Network Programming

In SR, an ingress node steers a packet through an ordered list of instructions, called "segments". Each one of these instructions represents a function to be called at a specific location in the network. A function is locally defined on the node where it is executed and may range from simply moving forward in the segment list to any complex user-defined behavior.

Network Programming combines Segment Routing (SR) functions to achieve a networking objective that goes beyond mere packet routing.

It may be desirable to encode a pointer to function and its arguments within an MPLS packet transport header. For example, in MPLS we can encode the FUNC::ARGs within the label stack or after the Bottom of Stack to support the equivalent of FUNC::ARG in SRv6 as described in [RFC8986].

3. Existing MPLS Use cases

There are serveral services that can be transported over MPLS networks today. These include providing Layer-3 (L3) connectivity (e.g. for unicast and multicast L3 services), and Layer-2 (L2) connectivity (e.g. for unicast Pseudo-Wires (PWs), multicast E-Tree, and broadcast E-LAN L2 services). In those cases, the user service traffic is encapsulated as the payload in MPLS packets.

For L2 service traffic, it is possible to use A Control Word (CW) [RFC4385] and [RFC5085] immediately after the MPLS header to disambiguate the type of MPLS payload, prevent possible packet misordering, and allow for fragmentation. In this case, the first nibble the data that immediately follows after MPLS bottom of stack is set to 0000b to identify the presence of PW CW.

In addition to providing connectivity to user traffic, MPLS may also transport OAM data (e.g. over MPLS G-AChs [RFC5586]). In this case, the first nibble of the data that immediately follows after MPLS bottom of stack is set to 0001b, it indicates the presence of a control channel associated witha PW, LSP, or Section.

Bit Index Explicit Replication (BIER) [RFC8296] traffic can also be encapsulated over MPLS. In this case, BIER has defined 0101b as the value for the first nibble in the data that immediately appears after the bottom of the label stack for any BIER encapsulated packet over MPLS.

For pseudowires, the G-ACh uses the first four bits of the PW control word to provide the initial discrimination between data packets and packets belonging to the associated channel, as described in [RFC4385].

It is expected that new use cases described in this document and within the MNA framework [I-D.ietf-mpls-mna-fwk] will allow for the co-existance and backward compatibility with all such existing MPLS services.

4. Co-existence of Usecases

Two or more of the aforementioned use cases MAY co-exist in the same packet. This may require the presence of multiple ancilary data (whether In-stack or Post-stack ancillary data) to be present in the same MPLS packet.

For example, IOAM may provide key functions along with network slicing to help ensure that critical network slice SLOs are being met by the network provider. In this case, IOAM is able to collect key performance measurement parameters of network slice traffic flows as it traverses the transport network.

5. IANA Considerations

This document has no IANA actions.

6. Security Considerations

This document introduces no new security considerations.

7. Acknowledgement

The authors gratefully acknowledge the input of the members of the MPLS Open Design Team.

8. Contributors

The following individuals contributed to this document:

Loa Andersson
Bronze Dragon Consulting
Email: loa@pi.nu

9. Informative References

[I-D.bestbar-teas-ns-packet]
Saad, T., Beeram, V. P., Dong, J., Wen, B., Ceccarelli, D., Halpern, J. M., Peng, S., Chen, R., Liu, X., Contreras, L. M., Rokui, R., and L. Jalil, "Realizing Network Slices in IP/MPLS Networks", Work in Progress, Internet-Draft, draft-bestbar-teas-ns-packet-10, 5 May 2022, <https://datatracker.ietf.org/doc/html/draft-bestbar-teas-ns-packet-10>.

[I-D.decraene-mpls-slid-encoded-entropy-label-id]
Decraene, B., Filsfils, C., Henderickx, W., Saad, T., Beeram, V. P., and L. Jalil, "Using Entropy Label for Network Slice Identification in MPLS networks.", Work in Progress, Internet-Draft, draft-decraene-mpls-slid-encoded-entropy-label-id-05, 12 December 2022, <https://datatracker.ietf.org/doc/html/draft-decraene-mpls-slid-encoded-entropy-label-id-05>.

[I-D.gandhi-mpls-ioam] Gandhi, R., Brockners, F., Wen, B., Decraene, B., and H. Song, "MPLS Data Plane Encapsulation for In Situ OAM Data", Work in Progress, Internet-Draft, draft-gandhi-mpls-ioam-10, 10 March 2023, <https://datatracker.ietf.org/doc/html/draft-gandhi-mpls-ioam-10>.

[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", Work in Progress, Internet-Draft, draft-gandhi-mpls-ioam-sr-06, 18 February 2021, <https://datatracker.ietf.org/doc/html/draft-gandhi-mpls-ioam-sr-06>.

[I-D.ietf-ippm-ioam-data] Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-data-17, 13 December 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-ippm-ioam-data-17>.

[I-D.ietf-ippm-ioam-direct-export] Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In Situ Operations, Administration, and Maintenance (IOAM) Direct Exporting", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-direct-export-11, 23 September 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-ippm-ioam-direct-export-11>.

[I-D.ietf-mpls-mna-fwk] Andersson, L., Bryant, S., Bocci, M., and T. Li, "MPLS Network Actions Framework", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-fwk-03, 11 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-fwk-03>.

[I-D.ietf-sfc-nsh-tlv] Wei, Y., Elzur, U., Majee, S., Pignataro, C., and D. E. Eastlake, "Network Service Header (NSH) Metadata Type 2 Variable-Length Context Headers", Work in Progress, Internet-Draft, draft-ietf-sfc-nsh-tlv-15, 20 April 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-sfc-nsh-tlv-15>.

[I-D.ietf-teas-ietf-network-slices] Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for IETF Network Slices", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-19, 21 January 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-19>.

[I-D.kompella-mpls-mspl4fa] Kompella, K., Beeram, V. P., Saad, T., and I. Meilik, "Multi-purpose Special Purpose Label for Forwarding Actions", Work in Progress, Internet-Draft, draft-kompella-mpls-mspl4fa-03, 10 July 2022, <https://datatracker.ietf.org/doc/html/draft-kompella-mpls-mspl4fa-03>.

[I-D.kompella-mpls-nffrr] Kompella, K. and W. Lin, "No Further Fast Reroute", Work in Progress, Internet-Draft, draft-kompella-mpls-nffrr-03, 8 July 2022, <https://datatracker.ietf.org/doc/html/draft-kompella-mpls-nffrr-03>.

[I-D.li-mpls-enhanced-vpn-vtn-id] Li, Z. and J. Dong, "Carrying Virtual Transport Network (VTN) Information in MPLS Packet", Work in Progress, Internet-Draft, draft-li-mpls-enhanced-vpn-vtn-id-03, 16 October 2022, <https://datatracker.ietf.org/doc/html/draft-li-mpls-enhanced-vpn-vtn-id-03>.

[I-D.lm-mpls-sfc-path-verification] Yao, L. and G. Mirsky, "MPLS-based Service Function Path(SFP) Consistency Verification", Work in Progress, Internet-Draft, draft-lm-mpls-sfc-path-verification-03, 11 June 2022, <https://

datatracker.ietf.org/doc/html/draft-lm-mpls-sfc-path-verification-03>.

[I-D.stein-srtsn] Stein, Y. J., "Segment Routed Time Sensitive Networking", Work in Progress, Internet-Draft, draft-stein-srtsn-01, 29 August 2021, <https://datatracker.ietf.org/doc/html/draft-stein-srtsn-01>.

[I-D.zzhang-intarea-generic-delivery-functions]
Zhang, Z. J., Bonica, R., Kompella, K., and G. Mirsky, "Generic Delivery Functions", Work in Progress, Internet-Draft, draft-zzhang-intarea-generic-delivery-functions-03, 11 July 2022, <https://datatracker.ietf.org/doc/html/draft-zzhang-intarea-generic-delivery-functions-03>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <https://www.rfc-editor.org/info/rfc4090>.

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, DOI 10.17487/RFC4385, February 2006, <https://www.rfc-editor.org/info/rfc4385>.

[RFC5085] Nadeau, T., Ed. and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, DOI 10.17487/RFC5085, December 2007, <https://www.rfc-editor.org/info/rfc5085>.

[RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, DOI 10.17487/RFC5286, September 2008, <https://www.rfc-editor.org/info/rfc5286>.

[RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <https://www.rfc-editor.org/info/rfc5586>.

[RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <https://www.rfc-editor.org/info/rfc7490>.

[RFC8296]
Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <https://www.rfc-editor.org/info/rfc8296>.

[RFC8595] Farrel, A., Bryant, S., and J. Drake, "An MPLS-Based Forwarding Plane for Service Function Chaining", RFC 8595, DOI 10.17487/RFC8595, June 2019, <https://www.rfc-editor.org/info/rfc8595>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

Authors' Addresses

Tarek Saad
Cisco Systems, Inc.

Email: tsaad.net@gmail.com

Kiran Makhijani
Futurewei Technologies

Email: kiranm@futurewei.com

Haoyu Song
Futurewei Technologies

Email: haoyu.song@futurewei.com

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com