Workgroup: MPLS Working Group
Internet-Draft:
draft-andersson-mpls-mna-operation-
architecture-00
Published: 26 September 2022
Intended Status: Informational
Expires: 30 March 2023
Authors: L. Andersson — Bronze Dragon Consulting
J. Guichard — Futurewei Technologies
H. Song — Futurewei Technologies
S. Bryant — University of Surrey

MPLS MNA Operational Architecture

Abstract

MPLS Network Action (MNA) allows MPLS packet to carry instruction and data for in-network services and functions in an MPLS network. This document describes the network operations to support MNAs and what actions an MNA capable Label Switching Router (LSR) takes when MNA is present or absent in an packet.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Multi-protocol Label Switching (MPLS) is a widely deployed forwarding technology. It uses label stack entries prepended to the payload. The label stack entries are used to identify the forwarding actions by each LSR. Actions may include pushing, swapping or popping the labels, and using the labels to determine the next hop for forwarding the packet. Labels may also be used to establish the context under which the packet is forwarded.

MPLS Network Actions (MNA) is used to support actions for Label Switched Paths (LSPs) and/or MPLS packets in addition to the normal forwarding. [I-D.ietf-mpls-mna-fwk] provides the architectural framework for MNA and [I-D.ietf-mpls-mna-requirements] provides the design requirements for MNA. MNA can support actions encoded within or below the label stack. The presence of MNA is indicated by a bSPL in the label stack.

This document specifies the architecture for the extension of MPLS to include MNA. MNAs carry information on in-network services and functions in an MPLS network. This document describes an architecture for MNAs and what actions an MNA capable Label Switching Router (LSR) takes when MNA is present or absent in an packet.

The MNA encoded below the label stack is supported by MPLS Extension Header (EH), which is described in [I-D.song-mpls-extension-header].

Below some example use cases for MPLS EH are listed. More use cases for MNA in general can be found in [I-D.ietf-mpls-mna-usecases]

* In-situ OAM: In-situ OAM (IOAM) records flow OAM information within user packets while the packets traverse a network.

* Network Telemetry and Measurement: A network telemetry and instruction header can be carried as an extension header to instruct a node what type of network measurements should be performed on the packets.

* Network Security: Security related functions may require user packets to carry some metadata.

* Segment Routing and Network Programming: MPLS extension header could support MPLS-based segment routing. The details will be described in a separate draft.

It is possible to distinguish between two types of MPLS MNAs, "Hop by Hop" (HBH) and "End to end" (E2E).

An HBH MNA is processed by every MNA capable node along an LSP, HBH MNAs MAY be inserted by an ingress LER or a transit LSR. A HBH MNA MUST be removed by an LSR along the LSP or by the egress LER. An LSR along the LSP may be configured to ignore HBH MNAs.

An E2E MNA will be inserted by an upstream LSR and, processed and MUST be removed by a downstream LSR, no other LSR in between will process the E2E MNA.

Note: This document separates the concepts of LSP and MNA path, and allows an MNA to be applied on any section of an LSP. Another extreme is to strictly limit that MNAs can only initiate and terminate at LERs. This is simpler yet inflexible. A decision needs to be made after examining all the potential use cases.

Only MNA capable LSRs will process MNAs if they are configured to do so. LSR that are MNA non-capable will ignore the MNA and forward the packet as if the information was not there.

This document describes the interaction between MNA capable neighbor LSRs, and between MNA capable LSRs and a neighbor that is MNA non-capable.

1.1. Requirement 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. Specification

This document specifies the use of MNA with MPLS.

2.1. MPLS Network Action Overview

Applications carried over an MPLS network may require that specific instructions and/or metadata are added to user packets. One such example is In-situ OAM (IOAM) [RFC9197]. It is likely that new applications will emerge over time.

One or more MNAs may be added by an ingress node to an MNA Path (NAP) and be removed by one or more MNA capable nodes along the NAP. Such ingress and egress nodes may be nodes at the head end and tail end of a Label Switched Path (LSP), or any other intermediate node of the LSP that is MNA capable. For more details on NAPs see Figure 1.

2.2. MNA Operation Terminology

This section lists the abbreviations and concepts that are used throughout this document in the context of MNA.

* MNA - MPLS Network Action

* MNAI - MPLS Network Action Indicator, a bSPL in the label stack.

* LDP DoD - LDP Downstream on Demand

* LDP DU - LDP Downstream Unsolicited

* LSP - Label Switched Path

* LSR - Label Switching Router

The following concepts new for MPLS are defined:

* MNA capable node - an LSR that can process MNAs and announce its MNA capability
* MNA capable LSR - this may be used interchangeably with MNA capable node.
* non-MNA-capable node - an LSR that is unaware of and unable to process MNAs.
* NAP - A network action path for a specific network action, which is sub-path of an LSP. An NAP starts at the node adding an MNA and ends at the node that removes it.

3. MNA Basics

3.1. General Principles

Any MNA capable node along an LSP may add an MNA as long as it can be verified that there is another MNA capable LSR downstream that can remove it. Any MNA capable node downstream can be configured to remove an MNA. An NAP starts when an MNA is added and ends where it is removed. If there is no node downstream capable to remove the MNA, it MUST NOT be added. It is assumed that a control plane will make this determination, the specification of which is outside the scope of this document.

In the context of the MNA, an MNA capable node assumes that all user packets on the default LSP carry MNAs. As an optimization a second parallel LSP may be instantiated using a Forwarding Equivalence Class (FEC) that does not permit MNAs, thus indicating to the LSR that there are no MNAs in the packet.

3.2. LSPs in an MNA capable Network

For an MNA capable LSP between two MNA capable LSRs there are two label mappings:

* first, a label mapping for the FEC that indicates that the packet carries IP
* second, a label mapping for a new FEC indicating that there are no MNAs in the packet

3.3. MNA capable nodes

MNA capable nodes may process MNAs, i.e. add, augment, remove or do required processing.

An MNA capable node may not add an MNA to a packet if unless it is sure that there is a downstream node that can remove it.

If an LSP forks due to ECMP, the node that does the forking MUST be sure that all LSP branches (which may be re-merged) eventually terminate at an MNA capable node which will remove the MNA.

3.4. NAP and LSP

MNA capable nodes may process MNAs, i.e. add, remove or do required processing.

Figure 1 is used for illustration of NAPs.

```
<------------------LSP-------------------->

A------b------c------D------E------F------G

<------------------NAP1------------------->
<------------NAP2----------->
                     <--------NAP3-------->
                     <-----NAP4---->
                                   <-NAP5->
```

A, D, E, F and G are MNA capable nodes

b and c are non-MNA-capable nodes.

Figure 1: NAP vs. LSP

LSP - the LSP originates at ingress LSR A and terminates at egress LSR G, packets flow from A to G.

NAP1 - NAP1 originates with the MNA capable node A adding an MNA to the packet and terminates when the MNA capable node G removes the MNA.

NAP2 - NAP2 originates with the MNA capable node A adding an MNA to the packet and terminates when the MNA capable node E removes the MNA. i.e. the NAP2 is shorter than the LSP.

NAP3 - NAP3 originates with the MNA capable node D adding an MNA to the packet and terminates when the MNA capable node G removes the MNA.

NAP4 - NAP4 originates with the MNA capable node D adding an MNA to the packet and terminates when the MNA capable node F removes the MNA, i.e. it is not necessary that an NAP originates or terminate on an MPLS LER.

NAP5 - NAP5 originates with the MNA capable node F adding an MNA to the packet and terminates when the MNA capable node G removes the MNA.

Further discussion on the information needed in the packet to identify and process the post stack MNAs can be found in [I-D.song-mpls-extension-header].

3.5. Announcement of MNA Capability

A node that is MNA capable MUST have a way to announce this capability to other nodes in the same domain. Additions to the IGPs should be a baseline for such capabilities.

3.6. LSP establishment with LDP Downstream on Demand (DoD) in an MNA capable network

LSPs for MNA handling and processing in an MPLS network may be set up by LDP [RFC5036], a centralized controller and/or MPLS-SR. To enable this small extensions to the protocols are required.

In the examples in Section 3.6 and Section 3.7 we for simplicity assume that the payload of the packet is IP. It is of course possible that the payload will be a Pseudo-Wire (PW) or a Virtual Private Network (VPN). This will be described in a later version of the document.

It is anticipated that the difference in establishment procedures for IP, PW and VPN will be minor.

It is possible to use the simplified physical topology show in Figure 2 which uses LDP Downstream on Demand (DoD) to illustrate how LSP setup work in a network with a mix of MNA capable and non-MNA-capable nodes. In LDP DoD the action to set up an LSP is taken by the node at the head-end of the potential LSP.

```
    +---+      +---+      +---+      +---+      +---+
    |   |      |   |      |   |      |   |      |   |
    | A +------+ b +------+ D +------+ E +------+ G +
    |   |      |   |      |   |      |   |      |   |
    +---+      +---+      +---+      +---+      +---+
A, D, E, and G are MNA capable nodes

b is a non-MNA-capable node.
```

Figure 2: MNA topology I

The following steps would be taken assuming that node A wants to set up connectivity with node G to support MNA handling and processing:

* A sends an LDP Label Request message to b, indicating that an MNA capable LSP should be set up to G. A keeps track of the outstanding request.

* b is not MNA capable and treats the Label Request as a normal request, however, the information indicating that an MNA capable LSP is requested is transitive and sent to D.

* D receives the Label Request, forwards it to E, and keeps track of the outstanding request.

* E treats the label request the same way as D, and forward it to G.

* G receives the label request, finds out that it is the egress node for this LSP. G allocates two labels one for the IP FEC and one for the new "no MNA present" FEC. G sends a label mapping to E with both labels, and asks E to PHP both LSPs.

* E receives the label mapping and installs PHP for both the IP FEC and for the new "no MNA present"-FEC. E allocates two labels one for the IP FEC (label value 201) and one for the new FEC (label value 301). E sends a label mapping message to D, with the two labels.

* D receives the label mapping message and installs label 201 for the IP FEC and label value 301 for the new FEC. Since D know that b is not MNA capable it will only allocate one label (202 for the IP FEC) and send a label mapping message to with that label.

* b receives the label mapping messages and installs label 202 for the IP FEC. Since b is not MNA capable it will only allocate one

label (203 for the IP FEC). b sends a label mapping message to A with that label.

* A receives the label mapping and installs label value 203 for the IP FEC.

This will result in installed labels like this.

```
+---+         +---+         +---+         +---+         +---+
|   |...203...|   |...202...|   |...201...|   |...php...|   |
| A +---------+ b +---------+ D +---------+ E +---------+ G +
|   |         |   |         |   |...301...|   |...php...|   |
+---+         +---+         +---+         +---+         +---+
```

A, D, E and G are MNA capable nodes.

b is a non-MNA-capable node.

Figure 3: MNA topology II

3.7. LSP establishment with LDP Downstream Unsolicited (DU) in an MNA capable network

In LDP Downstream Unsolicited (DU) the initiative to establish a LSP is taken by the egress router. The egress will establish an LSP to every prefix it learns of from the IGP. With the exception from how the set up of the LSP(s) are triggered the label mappings are similar to how it is done with LDP DoD.

The same topology as in the LDP DoD example Figure 2 will be used for LDP DU.

* G learns that an MNA capable LSP to egress LSR A is needed. G allocates two labels one for the IP FEC and one for the new "no MNA present" FEC. G sends a label mapping to E with both labels, and asks E to PHP both LSPs.

* E receives the label mapping and installs PHP for both the IP FEC and for the new "no MNA present"-FEC. E allocates two labels one for the IP FEC (label value 201) and one for the new FEC (label value 301). E sends a label mapping message to D, with the two labels.

* D receives the label mapping message and installs label 201 for the IP FEC and label value 301 for the new FEC. Since D know that

b is not MNA capable it will only allocate one label (202 for the IP FEC) and send a label mapping message to with that label.

* b receives the label mapping messages and installs label 202 for the IP FEC. Since b is not MNA capable it will only allocate one label (203 for the IP FEC). b sends a label mapping message to A with that label.

* A receives the label mapping and installs label value 203 for the IP FEC.

* This will result in the exact the same label mappings as in the DoD Example, see Figure 3.

3.8. Forwarding Behavior of MNA Capable Nodes

An MNA capable node will always search the label stack for MNAs, with the exception of when a packet is received on the new "no MNA present" FEC.

Non-MNA-capable nodes will never search the label stack for MNAs.

Given the configuration in Figure 3 packets will be forwarded as follows through the network.

If Node A sends a packet with a post-stack MNA:

1. A sends a packet with label 203 with an EH after the label stack to b

2. b receives the packet and swaps the label to 202 and forward it to D.

3. D receives the packet, and since D is MNA capable it will search the stack to find an MNAI. Since there is MNA present, D will decide whether it should process the MNA or not. When that decision is taken and potential processing is done, D will swap the label to 201 and send it to E.

4. E receives the packet on LSP with a FEC that indicates that "MNA may present" and will search the packet for an MNA. When the MNA is found by E it will, if required, process the MNA, after that the top label is popped and the packet is forwarded to G.

5. G receives the packet, it will search the label stack to find the MNAI. It will find the MNA and since G is the egress node it will do necessary processing and as a last step remove the MNA. G will forward the packet based on the IP address.

If Node A sends a packet without any MNA:

1. A sends the packet with label 203 to b

2. b receives the packet and swaps the label to 202 and forward it to D.

3. D receives the packet, and since D is MNA capable it will search the stack to find an MNA. Since there is no MNA present, D will swap the label to 301 and send it to E (FEC indicates no MNA present).

4. E receives the packet on FEC "no MNA present" and understand that it does not need to search the packet for an MNA. E pops the label and forward to G

5. G receives the packet on FEC "no MNA present" and understand that it does not need to search the packet for an MNA. G will forward it based on the IP address.

3.9. MNA for RSVP-TE tunnels

Extension Headers for RSVP-TE tunnels is for further study. Essentially it expected to be similar to the LDP case.

4. MNA in VPNs

TBA

5. MNA and MPLS-SR

TBA

6. MNA distribution and MNA capability announcement

TBA

7. Security Considerations

TBA

8. IANA Considerations

MPLS MNA will require code point allocations from more than one IANA registry. It is not yet decided which document that will make which allocation. However, tentatively the "No MNA present" FEC will be assigned from this document.

IANA is requested to allocate lowest free value from the "IETF Review" range as new FEC from the "Forwarding Equivalence Class

(FEC) Type Name Space" in the "Label Distribution Protocol (LDP) Parameters", like this:

Value	Hex	Name	Label Distribution Discipline	Reference	Note/Reg. Date
TBD	TBD	No MNA present	DoD or DU	This Document	TBA

Table 1: No MNA present

9. Acknowledgments

TBA

10. References

10.1. Normative References

[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-01, 8 September 2022, <https://www.ietf.org/archive/id/draft-ietf-mpls-mna-fwk-01.txt>.

[I-D.ietf-mpls-mna-requirements] Bocci, M., Bryant, S., and J. Drake, "Requirements for MPLS Network Action Indicators and MPLS Ancillary Data", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-requirements-03, 19 August 2022, <https://www.ietf.org/archive/id/draft-ietf-mpls-mna-requirements-03.txt>.

[I-D.ietf-mpls-mna-usecases] Saad, T., Makhijani, K., Song, H., and G. Mirsky, "Use Cases for MPLS Network Action Indicators and MPLS Ancillary Data", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-usecases-00, 19 May 2022, <https://www.ietf.org/archive/id/draft-ietf-mpls-mna-usecases-00.txt>.

[I-D.song-mpls-extension-header] Song, H., Li, Z., Zhou, T., Andersson, L., Zhang, Z., Gandhi, R., Rajamanickam, J., and J. Bhattacharya, "Support MPLS Network Actions using Post-Stack Extension Headers", Work in Progress, Internet-Draft, draft-song-mpls-extension-header-10, 1 September 2022, <https://www.ietf.org/archive/id/draft-song-mpls-extension-header-10.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>.

10.2. Informative References

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <https://www.rfc-editor.org/info/rfc5036>.

[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

Loa Andersson
Bronze Dragon Consulting

Email: loa@pi.nu

James N Guichard
Futurewei Technologies

Email: james.n.guichard@futurewei.com

Haoyu Song
Futurewei Technologies

Email: haoyu.song@futurewei.com

Stewart Bryant
University of Surrey

Email: stewart.bryant@gmail.com