Workgroup: MPLS Working Group
Internet-Draft: draft-jags-mpls-ps-mna-hdr-01
Published: 20 October 2023
Intended Status: Standards Track
Expires: 22 April 2024
Authors: J. Rajamanickam, Ed.  R. Gandhi, Ed.
Cisco Systems, Inc.  Cisco Systems, Inc.
R. Zigler  T. Li  J. Dong
Broadcom  Juniper Networks  Huawei Technologies

# Post-Stack MPLS Network Action (MNA) Solution

## Abstract

This document defines the Post-Stack MPLS Network Action (MNA) solution for carrying Network Actions and Ancillary Data after the MPLS label stack based on In-Stack MNA solution defined in draft-ietf-mpls-mna-hdr. MPLS Network Actions can be used to influence packet forwarding decisions, carry additional OAM information in the MPLS packet or perform user-defined operations. This document addresses the MNA requirements specified in draft-ietf-mpls-mna-requirements. This document follows the MNA framework specified in draft-ietf-mpls-mna-fwk.

## 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 22 April 2024.

## Copyright Notice

**Table of Contents**

## 1. Introduction

[RFC3032] defines the encoding of the MPLS label stack, the basic structure used to define a forwarding path. Forthcoming applications require MPLS packets to perform special network actions and carry optional Ancillary Data (AD) that can affect the packet forwarding decision or trigger OAM logging, for example. Ancillary Data can be used to carry additional information, such as a IOAM, Path tracing etc. Several MNA applications are described in [I-D.ietf-mpls-mna-usecases]. User-defined network actions allow new, local actions to be defined. In some cases, more Ancillary Data may required to be carried in the MPLS header, so these kind of Network Actions and its Ancillary data are encoded after the MPLS Stack. These are called as Post-Stack Data.

This document defines the syntax and semantics of Post-Stack Network Actions and their corresponding Ancillary Data based on the In-Stack MNA solution defined in [I-D.ietf-mpls-mna-hdr]. This document addresses the requirements specified in [I-D.ietf-mpls-mna-requirements]. This document follows the framework specified in [I-D.ietf-mpls-mna-fwk].

## 2. Conventions Used in This Document

### 2.1. Requirements Language

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

### 2.2. Abbreviations

The terminology defined in [I-D.ietf-mpls-mna-fwk] and [I-D.ietf-mpls-mna-requirements] are used in this document.

| Abbreviation | Meaning | Reference |
|---|---|---|
| AD | Ancillary Data | [I-D.ietf-mpls-mna-requirements] |
| bSPL | Base Special Purpose Label | [RFC9017] |
| BOS | Bottom Of Stack | [RFC3032] |
| HBH | Hop-By-Hop Scope | [I-D.ietf-mpls-mna-fwk] |
| I2E | Ingress-To-Egress Scope | [I-D.ietf-mpls-mna-fwk] |
| IHS | I2E, HBH, or Select Scope | [I-D.ietf-mpls-mna-hdr] |
| ISD | In-Stack Data | [I-D.ietf-mpls-mna-requirements] |

| Abbreviation | Meaning | Reference |
|---|---|---|
| LSE | Label Stack Entry | [RFC3032] |
| MNA | MPLS Network Actions | [I-D.ietf-mpls-mna-fwk] |
| NAI | Network Action Indicator | [I-D.ietf-mpls-mna-requirements] |
| NAL | Network Action Length | [I-D.ietf-mpls-mna-hdr] |
| NAS | Network Action Sub-Stack | [I-D.ietf-mpls-mna-fwk] |
| NASL | Network Action Sub-Stack Length | [I-D.ietf-mpls-mna-hdr] |
| OAM | Operations And Management | [RFC4377] |
| P | Post-Stack Network Action Indicator Bit | This document |
| PSD | Post-Stack Data | [I-D.ietf-mpls-mna-requirements] and [I-D.ietf-mpls-mna-fwk] |
| PSMNA | Post-Stack MPLS Network Action | This document |
| PS-MNA-OP | Post-Stack MPLS Network Action Opcode | This document |
| TC | Traffic Class | [RFC5462] |
| TTL | Time To Live | [RFC3032] |

Table 1: Abbreviations

## 3. Overview

A Flag in the In-Stack NAS header [I-D.ietf-mpls-mna-hdr] indicates the presence of the Post-Stack MNA. The Post-Stack MNA's are encoded after the MPLS Label Stack (BoS).

The Post-Stack MNA encoding contains two main parts:

*Post-Stack Network Action Indicator

*Post-Stack Network Action Encoding

## 4. Post-Stack Network Action Indicator

A reserved bit (21st bit from left in LSE Format B) in the In-Stack MNA header described in [I-D.ietf-mpls-mna-hdr] is used to indicate the presence of the Post-Stack Network Action.

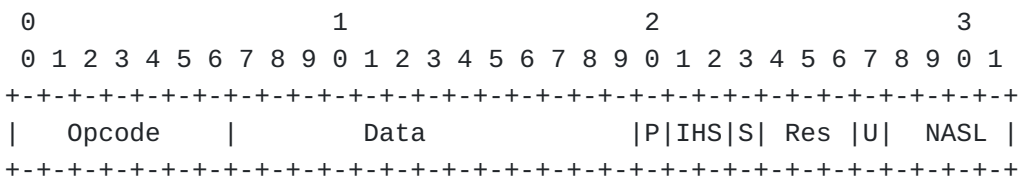


Figure 1

The below are the flags applicable to Post-Stack MNA encoding purposes defined in [I-D.ietf-mpls-mna-hdr].

- *P (1 Bit) : Indicates the presence of the Post-Stack MNA

- *IHS (2 Bit) : Indicates the combined scope of the In-Stack and the Post-Stack Network Actions. Each scope with P bit set will have its corresponding Post-Stack MNA sub-stack.

- *U (1 Bit) : Indicates the combined Unknown Action Handling of the In-Stack and the Post-Stack Network Actions

## 5. Post-Stack Network Action Encoding

The Post-Stack Network Action and its Ancillary Data are encoded after the MPLS Label Stack (BoS). The Post-Stack Network Action may carry multiple Post-Stack Network Actions and its corresponding Ancillary Data.

This consist of two main parts:

- *Post-Stack Network Action Top Header

- *Post-Stack Network Action Header

### 5.1. Post-Stack Network Action Top Header

This header is overall for all the Post-Stack Network Actions that are encoded.

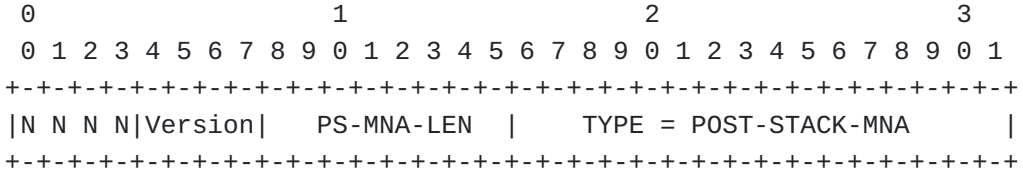

Figure 2

*NNNN (4 bits): This first nibble identifies the start of the Post-Stack Network Actions. A new value can be assigned by IANA (value TBA1). Generic Associated Channel (0001b) can be used instead.

*Version (4 bits): This is Post-Stack MNA version. The initial version will be 0.

*PS-MNA-LEN (8 bits): Post-Stack MNA Total Length in words. This excludes the Post-Stack Top header.

*TYPE (16 bits): Type is set to POST-STACK-MNA. The type value is an IANA allocated value.

### 5.2. Post-Stack Network Action Header

This header encodes a single Post-Stack Network Action. Using this scheme, multiple Post-Stack Network Action and its corresponding Ancillary data can be encoded.

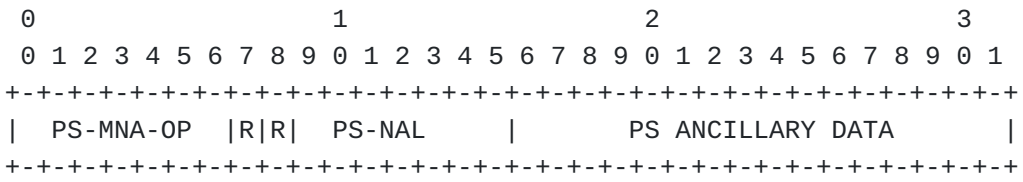


Figure 3

*PS-MNA-OP (7 bits): Post-Stack Network Action Opcode. Opcode "0" is reserved and other opcodes will be assigned by IANA accordingly.

*R (2 bits): Reserved bits

*PS-NAL (7 bits): Post-Stack Network Action Length for the respective Network Action. This value is in the order of words excluding current word.

*PS ANCILLARY DATA (16 bits): Post-Stack Ancillary Data associated with the Network Action

## 6. In-Stack Special Opcode Allocation

Some of the In-Stack MNA Opcodes are allocated to support Post-Stack Network Action. They are as follows.

### 6.1. Post-Stack Network Action Offset

Opcode: TBA2

Purpose: This opcode carries the start offset of the Post-Stack Network Action Top Header.

LSE Format: B or C (defined in [I-D.ietf-mpls-mna-hdr])

Data: The data value of the LSE contains the offset from the MPLS BOS in units of 4 octets. This allows the Generic Control Word (0000b) [RFC4385] and G-ACh (0001b) [RFC5586] fields to be placed immediately after the BOS. In the absence of this opcode, the Post-Stack Network Action is encoded immediately after the MPLS BOS. A data value of 1 indicates that the PAH starts 4 octets after the BOS.

Scope: This opcode can be used with any scope.

### 6.2. PS-IS-NA Ordering

Opcode: TBA3

Purpose: In cases where the ordering of network action is significant and where some of the network actions reside in Post-Stack Network Action, this opcode can be used to insert Post-Stack network actions into the order of execution. The 'P' bit and 'O' bit MUST be set in the NAS's Format B LSE if this opcode is used.

LSE Format: B, C, D (defined in [I-D.ietf-mpls-mna-hdr])

Data: The data field contains one or more 7-bit Post-Stack MNA Opcode. When used with LSE Format B, only one PS MNA Opcode is carried. Two PS MNA opcodes can be carried in a Format C LSE, and if Format D LSEs are used, each may carry up to three PS MNA opcodes. The PS MNA opcodes are the stored concatenated in the most significant bits of the data field. If multiple indicators are carried, the most significant PS MNA opcode is evaluated to the least significant. PS MNA opcodes do not span LSEs. If some PS MNA opcode positions are not to be used, then the opcode should be set to value 0.

Scope: This opcode can be used with any scope.

## 7. Node Capability Signaling

The ingress node which is adding a Post-Stack MNA MUST make sure that the egress node is capable of MNA and removes the Post-Stack MNA.

- *Each participating node MUST signal the network actions that it supports.

- *Each participating node MUST signal its Maximum Post-Stack MNA Length that could encoded.

The above capability signaling will be added in appropriate protocols. Signaling details are outside the scope of this document.

## 8. Processing the Network Action Sub-Stack

This section defines the specific responsibilities for nodes along a MPLS path.

### 8.1. Encapsulating Node Responsibilities

The encapsulating node MAY add Post-Stack MNA to the packet in accordance with its policies, the placement restrictions, and the limitations.

The encapsulating node MUST NOT add a Post-Stack MNA to the packet if the decapsulation node does not support Post-Stack MNA.

If the encapsulating node is also a transit node, then it MUST also respect transit node responsibilities.

### 8.2. Transit Node Responsibilities

A transit node MAY change the Ancillary Data in the Post-Stack MNA.

A transit node MUST respect the Unknown Action Handling value encoded in the NAS.

A node that removes the last copy of a NAS that has the P bit set MUST remove all Post-Stack Network Actions.

### 8.3. Penultimate Node Responsibilities

In addition to the transit node responsibilities above, the penultimate node MUST NOT remove the last copy of a HBH or I2E NAS when it is exposed after removing the forwarding (transport) label. This allows the egress node to process the NAS.

### 8.4. Decapsulating Node Responsibilities

The decapsulating node MUST remove any Post-Stack MNA it receives.

## 9. Security Considerations

The security considerations in [RFC3032] also apply to this document.

In addition, MNA creates a new dimension in security concerns:

* The actions of an encapsulating node can affect any or all of the nodes along the path. In the most common and benign situations, such as a syntactically incorrect packet, this could result in packet loss or corruption.

* The semantics of a network action are unbounded and may be insecure. A network action could be defined that made arbitrary changes to the memory of the forwarding router, which could then be used by the encapsulating node to compromise every MNA capable router in the network. The IETF needs to ensure that only secure network actions are defined.

* The MNA architecture supports locally defined network actions. For such actions, there will be limited oversight to ensure that the semantics do not create security issues. Implementors and network operators will need to ensure that locally defined network actions do not compromise the security of the network.

## 10. IANA Considerations

### 10.1. Post-Stack MNA Nibble

This document requests that IANA allocate a value (TBA1) for the Post-Stack MNA Nibble (NNNN) to indicate the start of the Post-Stack Network Actions. The reference should be this document.

### 10.2. In-Stack Network Action Opcodes

The In-Stack Network Action Opcodes for In-Stack Network Action Opcode registry (to be created by in [[I-D.ietf-mpls-mna-hdr]])are defined in the document as follows

| Opcode | Description | Reference |
|---|---|---|
| TBA2 | Offset of start of Post-Stack Network Action Header | This document |
| TBA3 | PS-IS-NA Ordering | This document |

Table 2: In-Stack Network Action Flags With Ancillary Data Registry

### 10.3. Top Header Types Registry

This document requests that IANA create a new registry with the name "Top Header Types". The registration procedure for this registry is "IETF Review". The fields are "Type" (integer), "Description" (string), and "Reference" (string). Type is an integer 0-65535.

The initial assignments for this registry are:

| Type | Description | Reference |
|---|---|---|
| 0 | Reserved | This document |
| 1 | POST-STACK-MNA-TYPE | This document |
| 2-65520 | IETF Review | This document |
| 65521-65524 | Experimental Use | This document |
| 65525-65535 | Private Use | This document |

Table 3: Top Header Types Registry

### 10.4. Post-Stack Network Action Opcodes

This document requests that IANA create a new registry with the name "Post-Stack Network Action Opcodes". The registration procedure for this registry is "IETF Review". The fields are "Opcode" (integer), "Description" (string), and "Reference" (string). Opcode is an integer 0-127.

The initial assignments for this registry are:

| Opcode | Description | Reference |
|---|---|---|
| 0 | Reserved | This document |
| 1-110 | IETF Review | This document |
| 111-114 | Experimental Use | This document |
| 115-126 | Private Use | This document |

Table 4: Post-Stack Network Action Opcodes Registry

## 11. Appendix A: Examples

### 11.1. Post-Stack Network Action Encoding

#### 11.1.1. NAS that only Indicates Post-Stack NAs

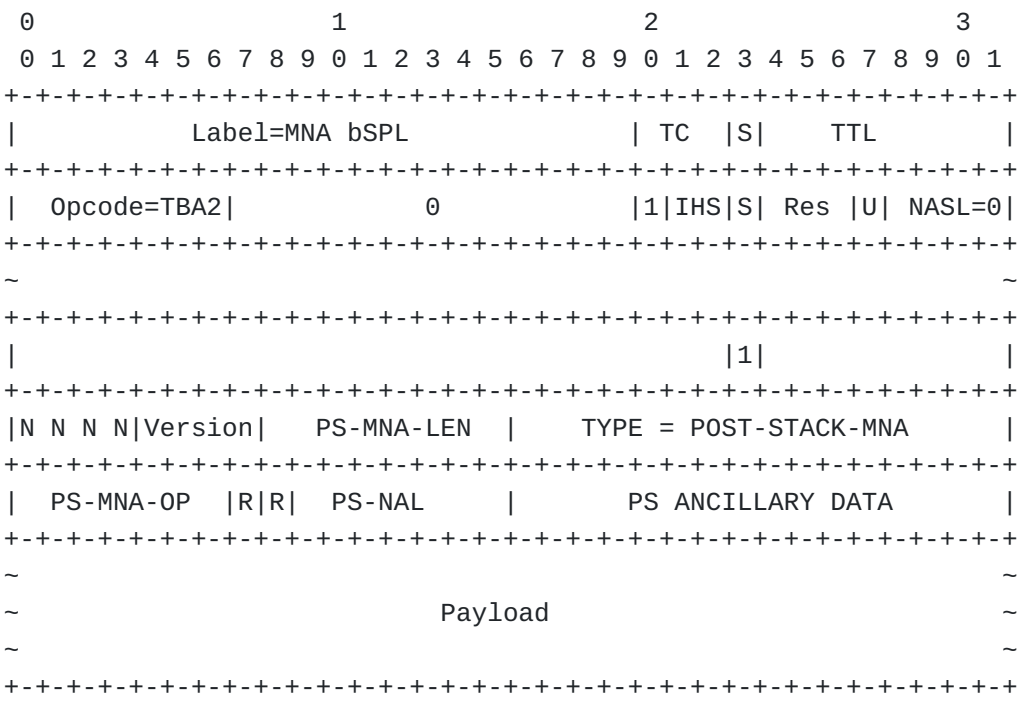


Figure 4: NAS encoding only Post-Stack NAs

In some cases, the NAS may encode only the presence of Post-Stack NAs. In this case, the P-Bit is set. The IHS field indicates the scope of the Post-Stack NAs (I2E, HBH, Select).

#### 11.1.2. NAS with both In-Stack and Post-Stack NAs

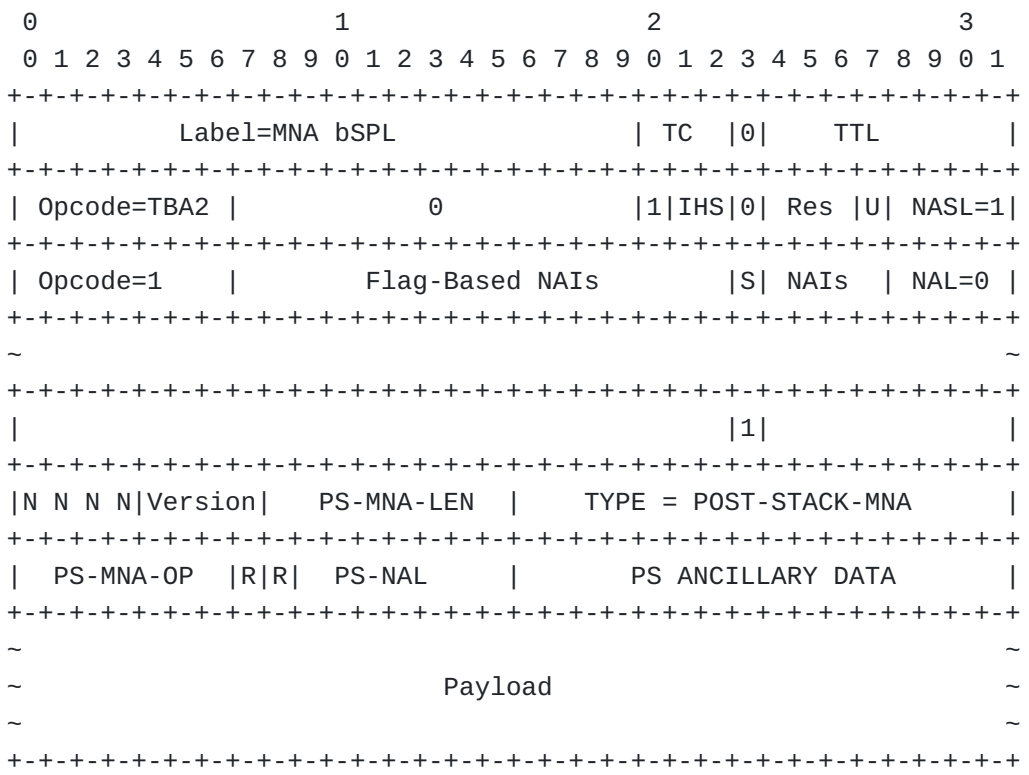


Figure 5: NAS with In-Stack and Post-Stack NAs

In some cases, the NAS may encode In-Stack NAs and indicate the presence of Post-Stack NAs. In this case, P-Bit is set. The NASL is set to "1", indicating the presence of one additional LSE. The IHS field indicates the scope of both the In-Stack and Post-Stack NAs.

#### 11.1.3. NASes with Multiple Scopes

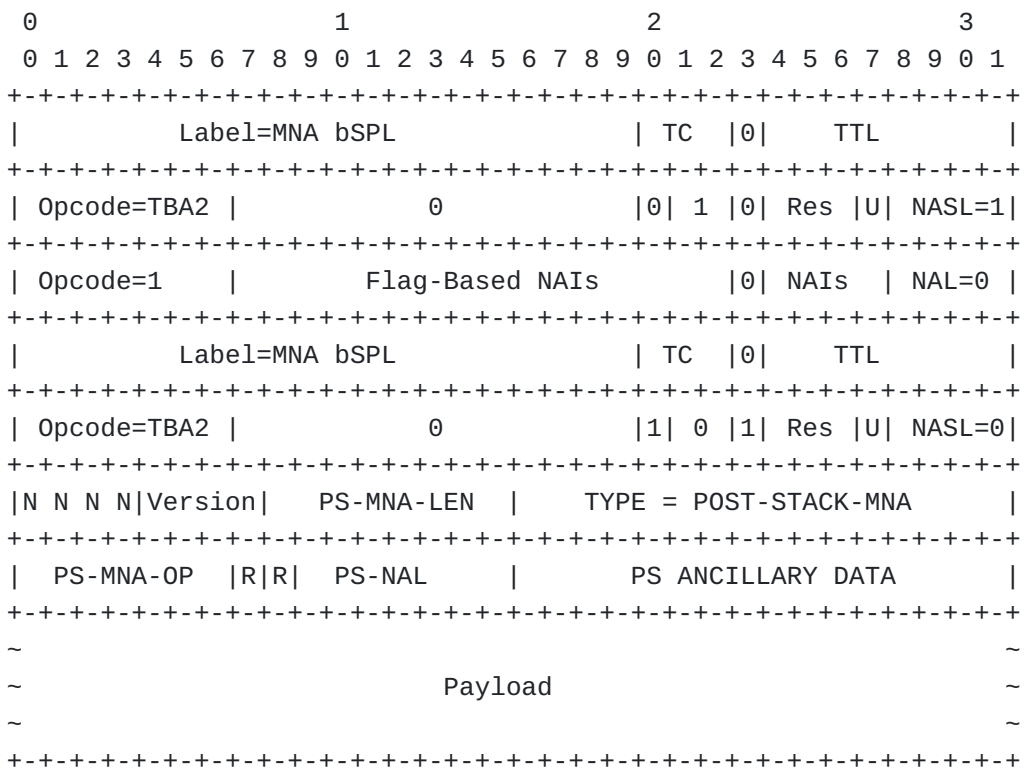


Figure 6: NASes with multiple scopes

In some cases the label stack may need to carry In-Stack NAs with Hop-By-Hop scope and Post-Stack NAs with I2E scope. In this case, there will be two NASes in the label stack. In this case, the first NAS will encode the In-Stack NA with the Hop-By-Hop scope and the second NAS will encode the presence of I2E scoped Post-Stack NAs.

### 11.2. Post-Stack Network Action with two Opcodes

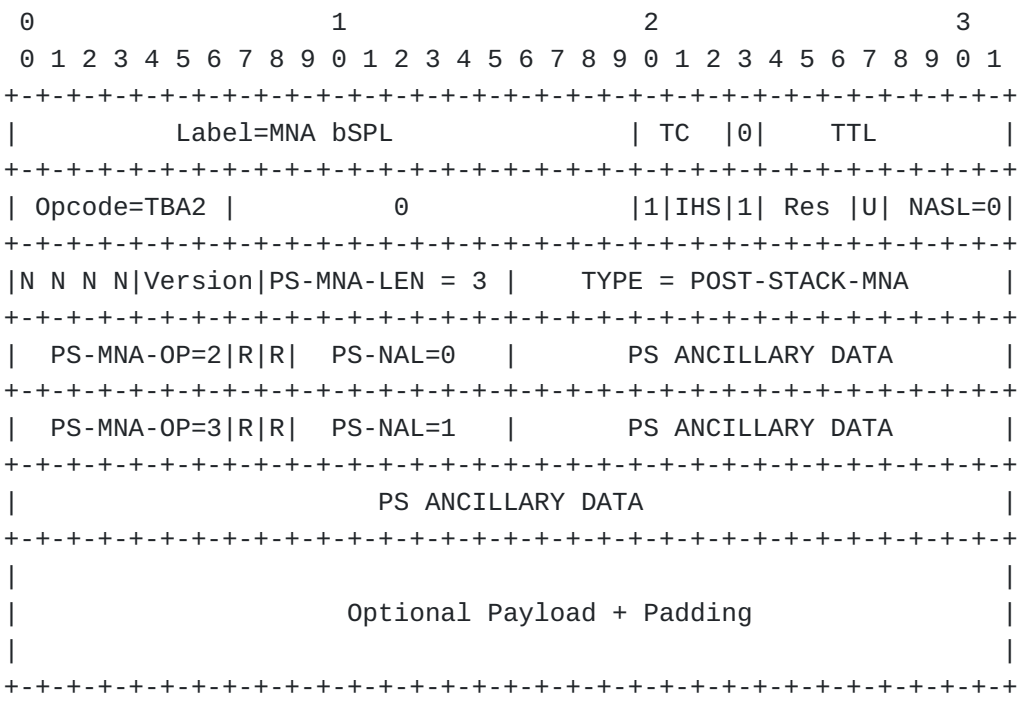


Figure 7: Post-Stack NA Example

This is an example of Post-Stack MNA encoding, that encode two different Post-Stack Network Actions.

Details:

PS-MNA-LEN=3: This is the Total Length of Post-Stack MNAs.

PS-MNA-OP=2: Post-Stack MNA Opcode "2".

PS-NAL=0: Post-Stack Network Action does not contain any additional data.

PS-MNA-OP=3: Post-Stack MNA Opcode "3".

PS-NAL=1: Post-Stack Network Action contains 1 additional word to carry its Ancillary data.

### 11.3. Post-Stack Network Action with two different scopes

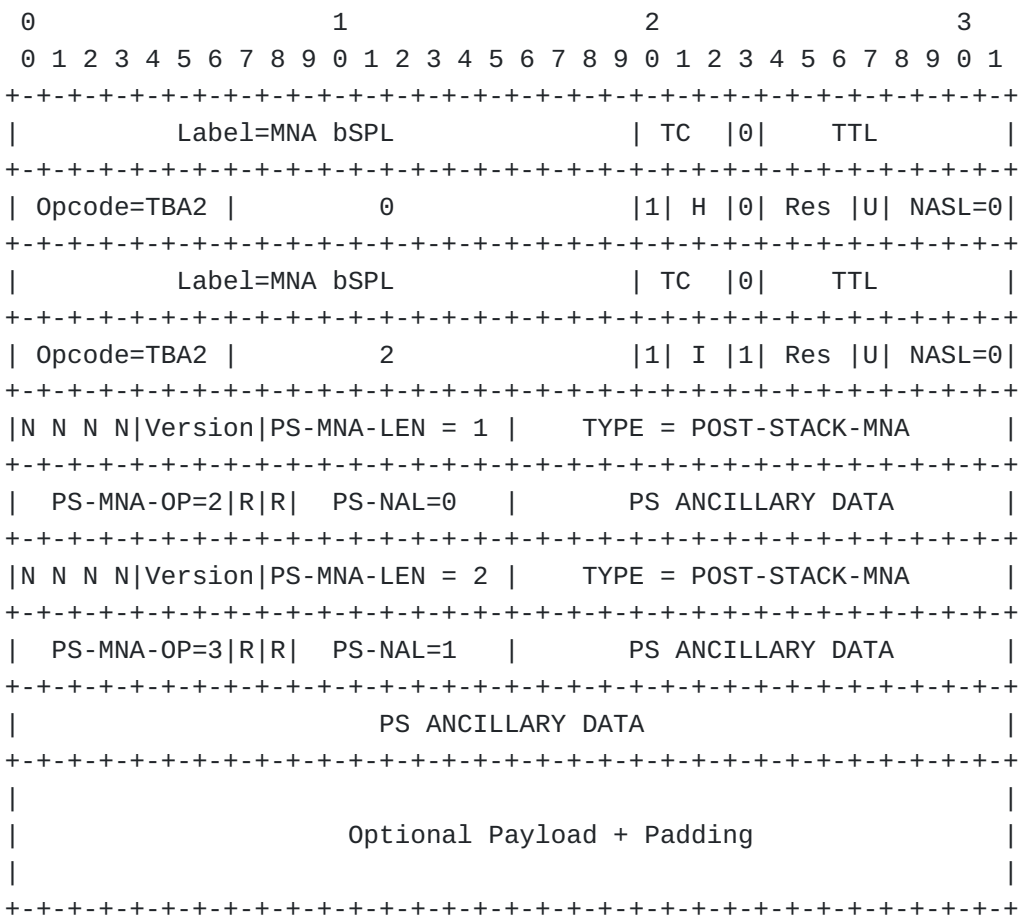


Figure 8: Post-Stack NA Example

This is an example of Post-Stack MNA encoding, that encode two different different scoped Post-Stack Network Actions. The first scope is Hop-By-Hop and the second scope is Ingress-To-Egress scoped PSD data.

Details:

Opcode:TBA2: This the offset of the Hop-By-Hop scoped PSD data. This value of this opcode is "0"

Opcode:TBA2: This the offset of the Ingress-To-Egress scoped PSD data. This value of this opcode is "2" (i.e) the PSD stack starts from second word after the MPLS Bottom Of Stack

### 11.4. Network Action Processing Order

The semantics of a network action can vary widely and the results of processing one network action may affect the processing of a subsequent network action.

### 11.4.1. Post-Stack NA Processing Order

By default, Post-Stack NAs follow the ordering of the encoding. However, the PS-IS-NA ordering opcode can be used to override the default ordering and interleave Post-Stack network actions with In-Stack network actions.

In some cases, Post-Stack NAs needs to be processed before In-Stack NAs. This section shows how to prioritize the Post-Stack NAs over In-Stack NAs.

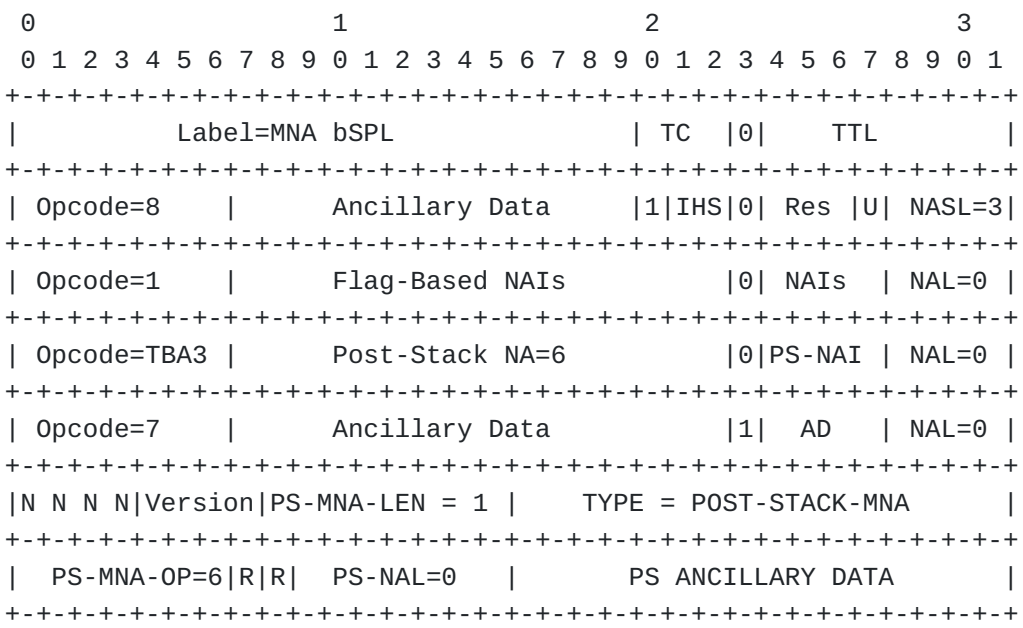


Figure 9: Post-Stack and In-Stack NA processing order

In the above example, opcode 8 is processed first, then the Flag-Based NAIs, followed by Post-Stack NA Opcode 6, and finally opcode 7.

## 12. References

### 12.1. 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-05, 19 October

2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-fwk-05>.

**[I-D.ietf-mpls-mna-requirements]** Bocci, M., Bryant, S., and J. Drake, "Requirements for MPLS Network Actions", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-requirements-07, 18 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-requirements-07>.

**[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-03, 6 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-hdr-03>.

**[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>.

**[RFC3032]** Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

**[RFC4377]** Nadeau, T., Morrow, M., Swallow, G., Allan, D., and S. Matsushima, "Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks", RFC 4377, DOI 10.17487/RFC4377, February 2006, <https://www.rfc-editor.org/info/rfc4377>.

**[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>.

**[RFC5462]** Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

**[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>.

**[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>.

**[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>.

### 12.2. Informative References

**[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-03, 15 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-usecases-03>.

## Acknowledgments

The authors would like to thank the authors and contributors of the draft-ietf-mpls-mna-hdr as this document borrows some text from the earlier version of that document.

## Contributors

The following people have substantially contributed to this document:

Jisu Bhattacharya
Cisco Systems, Inc.
Email: jisu@cisco.com

John Drake
Juniper Networks
United States
Email: jdrake@juniper.net

Figure 10

## Authors' Addresses

Jaganbabu Rajamanickam (editor)
Cisco Systems, Inc.
Canada

Email: jrajaman@cisco.com

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Royi Zigler
Broadcom

Email: royi.zigler@broadcom.com

Tony Li
Juniper Networks

Email: tony.li@tony.li

Jie Dong
Huawei Technologies
Beijing
China

Email: jie.dong@huawei.com