MPLS Working Group
Internet-Draft
Intended status: Informational
Expires: April 9, 2022

L. Andersson
Bronze Dragon Consulting
J. Guichard
H. Song
Futurewei Technologies
S. Bryant
University of Surrey

October 6, 2021

MPLS Extension Header Architecture
draft-andersson-mpls-eh-architecture-02

Abstract

Extension Headers (EH) carry information on in-network services and functions in an MPLS network. This document describes an architecture for EHs and what actions an EH capable Label Switching Router (LSR) takes when finding or not finding an EH in the packet.

Multiprotocol Label Switching (MPLS) is a widely deployed forwarding technology. It uses label stack entries that are pre-pended to either the EH or the ACH which in turn is pre-pended 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.

The extension headers are carried after the MPLS Label Stack, and the presence of EHs are indicated in the label stack by an Extension Header Indicator (EHI).

This Internet-Draft will expire on April 9, 2022.

Copyright Notice

Table of Contents

1. Introduction

This document specifies the architecture for the extension of MPLS to include Extension Headers (EH). EHs carry information on in-network services and functions in an MPLS network. This document describes an architecture for EHs and what actions an EH capable Label Switching Router (LSR) takes when finding or not finding an EH in the packet,

The extension headers are carried after the MPLS Label Stack, and the presence of EHs are indicated in the label stack by an Extension Header Indicator (EHI).

Below some exmple use cases are listed. More details will be found in [I-D.song-mpls-extension-header]

- 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 EHs, "hop-by hop" (HBH) and "End to end" (E2E).

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

An E2E EH will be inserted by the ingress LSR and, processed and MUST be removed by the egress LSR, no other LSR along the LSP will process the E2E EH.

Only EH capable LSRs will process EHs, LSR that are EH non-capable will ignore the EH and forward the packet as if the information was not there.

This document describes the interaction between EH capable neighbour LSRs, and between EH capable LSRs and a neighbour that is EH 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 Extension Headers (EH) with MPLS. Further information on EH processing and formats will be found in [I-D.song-mpls-extension-header].

2.1. Extension Header Overview

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

One or more EHs may be added by an ingress node to an Extension Header Path (EHP) and be removed by one or more EH capable nodes along the EHP. 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 EH capable. For more details on EHPs see Figure 1.

2.2. Extension Header Terminology

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

- EH - Extension Header
- EHI - Extension Header Indicator
- 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:

- EH capable node - an LSR that can process Extension Headers and announce its EH capability
- EH capable LSR - this may be used interchangeably with EH capable node.
- EH non-capable node - an LSR that is unaware of and unable to process Extension Headers.
- EH path - an EH path starts at the node adding an EH and ends at the node that removes it.

3. Extension Header Basics

3.1. General Principles

Any EH capable node along an LSP may add an EH as long as it can be verified that there is another EH capable LSR downstream that can remove it. Any EH capable node downstream may remove an EH. An EH path starts when one or more EHs are added and ends where the last EH is removed. If there is no node downstream capable to remove the EH, 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 MPLS EH architecture an EH capable node assumes that all user packets on the default LSP carry EHs. As an optimization a second parallel LSP may be instantiated using a Forwarding Equivalence Class (FEC) that does not permit EHs, thus indicating to the LSR that there are no EHs in the packet.

3.2. LSPs in a EH capable Network

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

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

3.3. EH capable nodes

EH capable nodes may process Extension Headers, i.e. add, augment, remove or do required processing at a transit node.

An EH capable node may not add an extension header 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 EH capable node which will remove the EH.

3.4. EH path and LSP

EH capable nodes may process Extension Headers, i.e. add, remove or do required processing at a transit node.

Figure 1 will be used for illustration.

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

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

     <------------------EHP1------------------->
     <------------EHP2----------->
                          <--------EHP3-------->
                          <-----EHP4---->
                                        <-EHP5->
```

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

b and c are non-EH capable nodes.

Figure 1: EH path vs. LSP

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

EHP1 - EHP1 originates with the EH capable node A adding an extension header to the packet and terminates when the EH capable node G removes the EH

EHP2 - EHP2 originates with the EH capable node A adding an extension header to the packet and terminates when the EH capable node E removes the EH. i.e. the EH path is shorter than the LSP

EHP3 - EHP3 originates with the EH capable node D adding an extension header to the packet and terminates when the EH capable node G removes the EH.

EHP4 - EHP4 originates with the EH capable node D adding an extension header to the packet and terminates when the EH capable node F removes the EH, i.e. it is not necessary that an EH path originates or terminate on an MPLS LER.

EHP5 - EHP5 originates with the EH capable node F adding an extension header to the packet and terminates when the EH capable node G removes the EH

Further discussion on the information needed in the packet to identify and process EHs are found in [I-D.song-mpls-extension-header].

3.5. Announcement of EH Capability

A node that is EH 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 EH capable network

LSPs for EH 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 Pseudowire (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 EH capable and EH non-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 EH capable nodes

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

Figure 2: EH topology

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

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

- o b is not EH capable and treat the Label Request as a normal request, however, the information indicating that an EH capable LSP is requested is transitive and sent to D.

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

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

- o 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 EH present" FEC. G sends a label mapping to E with both labels, and asks E to PHP both LSPs.

- o E receives the label mapping and installs PHP for both the IP FEC and for the new "no EH 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.

- o 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 EH capable it will only allocate one label (202 for the IP FEC) and send a label mapping message to with that label.

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

- o 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 EH capable nodes.

b is a non-EH capable node.

Figure 3: EH topology

3.7. LSP establishment with LDP Downstream Unsolicited (DU) in an EH 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 topoplogy as in the LDP DoD example Figure 2 will be used for LDP DU.

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

- o E receives the label mapping and installs PHP for both the IP FEC and for the new "no EH 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.

o 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 EH capable it will only allocate one label (202 for the IP FEC) and send a label mapping message to with that label.

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

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

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

3.8. Forwarding Behavior of EH Capable Nodes

A EH capable node will always search the label stack for EHs, with the exception of when a packet is received on the new FEC (no EH present).

Non-EH capable nodes will never search the label stack for EHs.

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

If Node A sends a packet with an extension header folling the label stack:

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 EH capable it will search the stack to find an EH-indicator. Since there is EH present, D will decide whether it should process the extension header 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 "EH may present" and will search the packet for an EH. When the EH is found by E it will, if required, process the EH, 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 EHI. It will find the EH and since G is the egress node it will

do necessary processing and as a last step remove the EH. G will forward the packet based on the IP address.

If Node A sends a packet without an extension header:

1. A sends a packet with label 203 without an EH 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 EH capable it will search the stack to find an EH. Since there is no EH present, D will swap the label to 301 and send it to E (FEC indicates no EH present).

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

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

3.9. EH for RSVP-TE tunnels

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

3.10. Ways to indicate an EH after the Label Stack

There are several ways to indicate the presence of EHs after the label stack. This will be discussed in a separate document.

4. EH in VPNs

TBA

5. EH and MPLS-SR

TBA

6. Extension Header Applications

TBA

7. EH distribution and EH capability announcement

TBA

8. Security Considerations

TBA

9. IANA Considerations

MPLS extension headers 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 EH 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	Reference	Note/Reg. Date

			Discipline			
TBD	TBD	No EH present	DoD or DU	This Document	TBA	

Table 1: No EH present

10. Acknowledgements

-

-

11. References

11.1. Normative References

[I-D.song-mpls-extension-header]
Song, H., Li, Z., Zhou, T., Andersson, L., and Z. Zhang, "MPLS Extension Header", draft-song-mpls-extension-header-05 (work in progress), July 2021.

[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

[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>.

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