

MPLS Working Group
Internet-Draft
Intended status: Informational
Expires: September 22, 2019

A. Malis
S. Bryant
Huawei Technologies
J. Halpern
Ericsson
W. Henderickx
Nokia
March 21, 2019

**MPLS Transport Encapsulation For The SFC NSH
draft-ietf-mpls-sfc-encapsulation-04**

Abstract

This document describes how to use a Service Function Forwarder (SFF) Label (similar to a pseudowire label or VPN label) to indicate the presence of a Service Function Chaining (SFC) Network Service Header (NSH) between an MPLS label stack and the packet original packet/frame. This allows SFC packets using the NSH to be forwarded between SFFs over an MPLS network, and to select one of multiple SFFs in the destination MPLS node.

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 September 22, 2019.

Copyright Notice

Copyright (c) 2019 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- [1. Introduction](#) [2](#)
- [1.1. Terminology](#) [3](#)
- [2. MPLS Encapsulation Using an SFF Label](#) [3](#)
- [2.1. MPLS Label Stack Construction at the Sending Node](#) [4](#)
- [2.2. SFF Label Processing at the Destination Node](#) [5](#)
- [3. Equal Cost Multipath \(ECMP\) Considerations](#) [5](#)
- [4. Operations, Administration, and Maintenance \(OAM\) Considerations](#) [6](#)
- [5. IANA Considerations](#) [6](#)
- [6. Security Considerations](#) [6](#)
- [7. Acknowledgements](#) [6](#)
- [8. References](#) [6](#)
- [8.1. Normative References](#) [6](#)
- [8.2. Informative References](#) [7](#)
- Authors' Addresses [8](#)

1. Introduction

As discussed in [[RFC8300](#)], a number of transport encapsulations for the Service Function Chaining (SFC) Network Service Header (NSH) already exist, such as Ethernet, UDP, GRE, and others.

This document describes an MPLS transport encapsulation for the NSH and how to use a Service Function Forwarder (SFF) [[RFC7665](#)] Label to indicate the presence of the NSH in the MPLS packet payload. This allows SFC packets using the NSH to be forwarded between SFFs in an MPLS transport network, where MPLS is used to interconnect the network nodes that contain one or more SFFs. The label is also used to select between multiple SFFs in the destination MPLS node.

This encapsulation is equivalent from an SFC perspective to other transport encapsulations of packets using the NSH. This can be illustrated by adding an additional line to the example of a next-hop SPI/SI-to-network overlay network locator mapping in Table 1 of [[RFC8300](#)]:


```

+-----+-----+-----+-----+-----+
| SPI | SI | Next Hop(s) | Transport Encapsulation |
+-----+-----+-----+-----+-----+
| 25 | 220 | Label 5467 | MPLS |
+-----+-----+-----+-----+-----+
    
```

Table 1: Extension to [RFC 8300](#) Table 1

SFF Labels are similar to other service labels at the bottom of an MPLS label stack that denote the contents of the MPLS payload being other than a normally routed IP packet, such as a layer 2 pseudowire, an IP packet that is routed in a VPN context with a private address, or an Ethernet virtual private wire service.

This informational document follows well-established MPLS procedures and does not require any actions by IANA or any new protocol extensions.

Note that using the MPLS label stack as a replacement for the SFC NSH, covering use cases that do not require per-packet metadata, is described elsewhere [[I-D.ietf-mpls-sfc](#)].

1.1. Terminology

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. MPLS Encapsulation Using an SFF Label

The encapsulation is a standard MPLS label stack [[RFC3032](#)] with an SFF Label at the bottom of the stack, followed by a NSH as defined by [[RFC8300](#)] and the NSH original packet/frame.

Much like a pseudowire label, an SFF Label MUST be allocated by the downstream receiver of the NSH from its per-platform label space, since the meaning of the label is identical independent of which incoming interface it is received [[RFC3031](#)].

If a receiving node supports more than one SFF (i.e., more than one SFC forwarding instance), then the SFF Label can be used to select the proper SFF, by having the receiving node advertise more than one SFF Label to its upstream sending nodes as appropriate.

The method used by the downstream receiving node to advertise SFF Labels to the upstream sending node is out of scope of this document.

That said, a number of methods are possible, such as via a protocol exchange, or via a controller that manages both the sender and the receiver using NETCONF/YANG, BGP, PCEP, etc. One such BGP-based method has already been defined, and is documented in [[I-D.ietf-bess-nsh-bgp-control-plane](#)]. This does not constrain the further definition of other such advertisement methods in the future.

While the SFF label will usually be at the bottom of the label stack, there may be cases where there are additional label stack entries beneath it. For example, when an Associated Channel Header (ACH) is carried that applies to the SFF, a Generic Associated Channel Label (GAL) [[RFC5586](#)] will be in the label stack below the SFF. Similarly, an Entropy Label Indicator/Entropy Label (ELI/EL) [[RFC6790](#)] may be carried below the SFF in the label stack. This is identical to the situation with VPN labels.

This document does not define a use for the Traffic Class (TC) field [[RFC5462](#)] (formerly known as the Experimental Use (EXP) bits [[RFC3032](#)]) in the SFF Label.

2.1. MPLS Label Stack Construction at the Sending Node

When one SFF wishes to send an SFC packet with a NSH to another SFF over an MPLS transport network, a label stack needs to be constructed by the MPLS node that contains the sending SFF in order to transport the packet to the destination MPLS node that contains the receiving SFF. The label stack is constructed as follows:

1. Push zero or more labels that are interpreted by the destination MPLS node on to the packet, such as the Generic Associated Channel [[RFC5586](#)] label (see [Section 4](#)). The TTL for these labels is set according to the relevant standards that define these labels.
2. Push the SFF Label to identify the desired SFF in the receiving MPLS node. The TTL for this MPLS label MUST be set to one to avoid mis-forwarding.
3. Push zero or more additional labels such that (a) the resulting label stack will cause the packet to be transported to the destination MPLS node, and (b) when the packet arrives at the destination node, either:
 - * the SFF Label will be at the top of the label stack (this is typically the case when penultimate hop popping is used at the penultimate node, or the source and destination nodes are direct neighbors), or

- * as a part of normal MPLS processing, the SFF Label becomes the top label in the stack before the packet is forwarded to another node and before the packet is dispatched to a higher layer.

The TTL for these labels is set by configuration, or set to the defaults for normal MPLS operation in the network.

2.2. SFF Label Processing at the Destination Node

The destination MPLS node performs a lookup on the SFF label to retrieve the next-hop context between the SFF and SF, e.g. to retrieve the destination MAC address in the case where native Ethernet encapsulation is used between SFF and SF. How the next-hop context is populated is out of the scope of this document.

The receiving SFF SHOULD check that the received SFF label has a TTL of 1 upon receipt. Any other values indicate a likely error condition and SHOULD result in discarding the packet.

The receiving MPLS node then pops the SFF Label (and any labels beneath it) so that the destination SFF receives the SFC packet with the NSH is at the top of the packet.

3. Equal Cost Multipath (ECMP) Considerations

As discussed in [[RFC4928](#)] and [[RFC7325](#)], there are ECMP considerations for payloads carried by MPLS.

Many existing routers use deep packet inspection to examine the payload of an MPLS packet, and if the first nibble of the payload is equal to 0x4 or 0x6, these routers (sometimes incorrectly, as discussed in [[RFC4928](#)]) assume that the payload is IPv4 or IPv6 respectively, and as a result, perform ECMP load balancing based on (presumed) information present in IP/TCP/UDP payload headers or in a combination of MPLS label stack and (presumed) IP/TCP/UDP payload headers in the packet.

For SFC, ECMP may or may not be desirable. To prevent ECMP when it is not desired, the NSH Base Header was carefully constructed so that the NSH could not look like IPv4 or IPv6 based on its first nibble. See [Section 2.2 of \[\[RFC8300\]\(#\)\]](#) for further details. Accordingly, the default behavior for MPLS-encapsulated SFC is to not use ECMP.

If ECMP is desired when SFC is used with an MPLS transport network, there are two possible options, Entropy [[RFC6790](#)] and Flow-Aware Transport [[RFC6391](#)] labels. A recommendation between these options, and their proper placement in the label stack, is for future study.

4. Operations, Administration, and Maintenance (OAM) Considerations

OAM at the SFC Layer is handled by SFC-defined mechanisms [[RFC8300](#)]. However, OAM may be required at the MPLS transport layer. If so, then standard MPLS-layer OAM mechanisms may be used at the transport label layer (the labels above the SFF label).

5. IANA Considerations

This document does not request any actions from IANA.

Editorial note to RFC Editor: This section may be removed at your discretion.

6. Security Considerations

This document describes a method for transporting SFC packets using the NSH over an MPLS transport network. It follows well-established MPLS procedures in widespread operational use and does not define any new protocol elements or allocate any new code points, and is no more or less secure than carrying any other protocol over MPLS. To the MPLS network, the NSH and its contents is simply an opaque payload.

In addition, the security considerations in [[I-D.ietf-mpls-sfc](#)] also apply to this document.

7. Acknowledgements

The authors would like to thank Jim Guichard, Eric Rosen, Med Boucadair, Sasha Vainshtein, Jeff Tantsura, Anoop Ghanwani, John Drake, Loa Andersson, Carlos Pignataro, Christian Hopps, and Benjamin Kaduk for their reviews and comments.

8. References

8.1. Normative References

[I-D.ietf-mpls-sfc]

Farrel, A., Bryant, S., and J. Drake, "An MPLS-Based Forwarding Plane for Service Function Chaining", [draft-ietf-mpls-sfc-07](#) (work in progress), March 2019.

[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>>.

- [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>>.
- [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>>.
- [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>>.
- [RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", [RFC 7665](#), DOI 10.17487/RFC7665, October 2015, <<https://www.rfc-editor.org/info/rfc7665>>.
- [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>>.
- [RFC8300] Quinn, P., Ed., Elzur, U., Ed., and C. Pignataro, Ed., "Network Service Header (NSH)", [RFC 8300](#), DOI 10.17487/RFC8300, January 2018, <<https://www.rfc-editor.org/info/rfc8300>>.

[8.2.](#) Informative References

- [I-D.ietf-bess-nsh-bgp-control-plane] Farrel, A., Drake, J., Rosen, E., Uttaro, J., and L. Jalil, "BGP Control Plane for NSH SFC", [draft-ietf-bess-nsh-bgp-control-plane-09](#) (work in progress), March 2019.
- [RFC4928] Swallow, G., Bryant, S., and L. Andersson, "Avoiding Equal Cost Multipath Treatment in MPLS Networks", [BCP 128](#), [RFC 4928](#), DOI 10.17487/RFC4928, June 2007, <<https://www.rfc-editor.org/info/rfc4928>>.
- [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>>.

- [RFC6391] Bryant, S., Ed., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network", [RFC 6391](#), DOI 10.17487/RFC6391, November 2011, <<https://www.rfc-editor.org/info/rfc6391>>.
- [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", [RFC 6790](#), DOI 10.17487/RFC6790, November 2012, <<https://www.rfc-editor.org/info/rfc6790>>.
- [RFC7325] Villamizar, C., Ed., Kompella, K., Amante, S., Malis, A., and C. Pignataro, "MPLS Forwarding Compliance and Performance Requirements", [RFC 7325](#), DOI 10.17487/RFC7325, August 2014, <<https://www.rfc-editor.org/info/rfc7325>>.

Authors' Addresses

Andrew G. Malis
Huawei Technologies

Email: agmalis@gmail.com

Stewart Bryant
Huawei Technologies

Email: stewart.bryant@gmail.com

Joel M. Halpern
Ericsson

Email: joel.halpern@ericsson.com

Wim Henderickx
Nokia

Email: wim.henderickx@nokia.com

