Workgroup: MPLS WG
Internet-Draft:
draft-lm-mpls-sfc-path-verification-03
Updates: RFC8595 (if approved)
Published: 11 June 2022
Intended Status: Standards Track
Expires: 13 December 2022
Authors: Y. Liu G. Mirsky
ZTE Ericsson

MPLS-based Service Function Path(SFP) Consistency Verification

Abstract

This document describes extensions to MPLS LSP ping mechanisms to support verification between the control/management plane and the data plane state for SR-MPLS service programming and MPLS-based NSH SFC.

This document defines the signaling of the Generic Associated Channel (G-ACh) over a Service Function Path (SFP) with an MPLS forwarding plane using the basic unit defined in RFC 8595. The document updates RFC 8595 in respect to SFF's handling TTL expiration. The document also describes the processing of the G-ACh by the elements of the SFP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Service Function Chain (SFC) defined in [RFC7665] as an ordered set of service functions (SFs) to be applied to packets and/or frames, and/or flows selected as a result of classification.

SFC can be achieved through a variety of encapsulation methods, such as NSH [RFC8300], SR service programming [I-D.ietf-spring-sr-service-programming] and MPLS-based NSH SFC [RFC8595].

This document describes extensions to MPLS LSP ping [RFC8029] mechanisms to support verification between the control/management plane and the data plane state for both SR-MPLS service programming and MPLS-based NSH SFC.

An MPLS LSP ping is a component of the MPLS Operation, Administration, and Maintenance (OAM) toolset. OAM packets used to monitor a specific Service Function Path (SFP) can be transported over a Generic Associated Channel (G-ACh). This document defines the signaling of the G-ACh over an SFP with an MPLS forwarding plane using the basic unit defined in [RFC8595]. The document updates [RFC8595] in respect to SFF's handling TTL expiration. The document also describes the processing of the G-ACh by the elements of the SFP.

[Editor's note:] LSP ping for SR-SFC will be discussed in a separate draft in the future, leaving this document focusing on LSP ping for SFC-MPLS. As for LSP ping for SFC-MPLS, although GAL and G-Ach are used currently, the proposal will follow the architecture of MNA [I-D.andersson-mpls-mna-fwk] in the future version.

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].

2.2. Terminology and Acronyms

SFC: Service Function Chain

SFF: Service Function Forwarder

SF: Service Function

SFI: Instance of an SF

SFP: Service Function Path

RSP: Rendered Service Path

SFC-MPLS: SFC over an MPLS forwarding plane introduced in [RFC8595]

SR-SFC: SFC achieved by SR service programming [I-D.ietf-spring-sr-service-programming]

NSH-SR: SFC based on the integration of Network Service Header (NSH) and SR for SFC [I-D.ietf-spring-nsh-sr]

SPL: Special-Purpose Label

bSPL: Base SPL

eSPL: Extended SPL

GAL: Generic Associated Channel Label

ELI: Entropy Label Indicator

OAM: Operation, Administration, and Maintenance

G-ACh: Generic Associated Channel

GAL: Generic Associated Channel Label

3. MPLS-based SFP Consistency Verification

MPLS echo request and reply messages [RFC8029] can be extended to support the verification of the consistency of an MPLS-based Service Function Path (SFP).

SR-MPLS/MPLS can be used to realize an SFP. Two methods have been defined:

* [I-D.ietf-spring-sr-service-programming] describes how to achieve service function chaining in SR-enabled MPLS and IPv6 networks. In an SR-MPLS network, each SF is associated with an MPLS label. As a result, an SFP can be encoded as a stack of MPLS labels and pushed on top of the packet.

* [RFC8595] provides another method to realize SFC in an MPLS network by means of using a logical representation of the Network Service Header (NSH) in an MPLS label stack. This method, throughout this document, is referred to as SFC over an MPLS data plane (SFC-MPLS). When an MPLS label stack is used to carry a logical NSH, a basic unit of representation is used, which can be present one or more times in the label stack. This unit comprises two MPLS labels, one carries a label to provide a context within the SFC scope (the SFC Context Label), and the other carries a label to show which SF is to be enacted (the SF Label). SFC forwarding can be achieved by label swapping, label stacking, or the mix of both. When an SFF receives a packet containing an MPLS label stack, it examines the top basic unit of the MPLS label stack for SFC, {SPI, SI} or {context label, SFI index}, to determine where to send the packet next.

In MPLS Label Switched Paths (LSPs), MPLS LSP ping [RFC8029] is used to check the correctness of the data plane functioning and to verify the data plane against the control plane.

The proposed extension of MPLS LSP ping allows verification of the correlation between the control/management (if data model-based central controller used) plane and the data plane state in SR-MPLS/MPLS-based SFC.

As for NSH-SR, OAM defined for NSH in [draft-ietf-sfc-multi-layer-oam] can be re-used and it is out of the scope of this document.

4. LSP Ping in SFC-MPLS

In SFC-MPLS, SFFs are responsible for MPLS echo request processing. there're two reasons:

* In SFC-MPLS, the packet forwarding decision is made by SFFs based on the basic unit. SFs are not aware of the FEC of the basic unit.

* Generally, except for the designed specific functions, the packet processing functions supported by SFs are limited. SFs may not support control and/or management protocols operated over the G-ACh defined in [RFC5586], e.g., MPLS OAM protocols like LSP ping. Such packets may be mishandled.

To support that processing, the basic unit can use the mechanism described in Section 4.1.

4.1. Special-purpose Label in SFC-MPLS Environment

When an SFC-MPLS is used, an SFF needs to identify an OAM packet with the SFP scope. To achieve that, this specification first defines the use of a base special-purpose label (bSPL) [RFC3032] or an extended special-purpose label (eSPL) [RFC7274] (referred to in this document as SPL Unit) with the basic unit defined in [RFC8595]. And based on that, the use of Generic Associated Channel Label (GAL) [RFC5586] with the basic unit in the SFC-MPLS environment.

Special-purpose label (SPL), whether bSPL or eSPL, has special significance in the data and control planes. An ability to use an SPL in the basic unit allows for a closer functional match between the NSH-based SFC and SFC-MPLS. For example, Entropy Label Indicator (ELI) [RFC6790] with the basic unit can be used as the Flow ID TLV [I-D.ietf-sfc-nsh-tlv] to allow an SFF to balance SFC flows among SFs of the same type. An SPL MAY be used with the basic unit in SFC-MPLS, as displayed in Figure 1. Note that an SPL unit MAY be present in one or more basic units when MPLS label stacking is used to carry the SFC information.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           SFC Context Label           | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            SPL Unit                           |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           SF Label                    | TC  |S|      TTL      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Special-purpose Label Unit with the Basic Unit of MPLS Label Stack for SFC

4.1.1. G-ACh over SFC-MPLS

SFC-MPLS environment could include instances of an SF (SFI) or SFC proxies that cannot properly process control and/or management protocol messages that are exchanged between nodes over the G-ACh associated with the particular SFP. To support OAM over G-ACh, it is beneficial to avoid handing over a test packet to the SFI or SFC proxy. Hence, this specification defines that if the Generic Associated Channel Label (GAL) immediately follows the SFC Context label [RFC8595], then the packet is recognized as an SFP OAM packet.

Below are the processing rules of an SFP OAM packet by an SFF:

* An SFF MUST NOT pass the packet to a local SFI or SFC proxy.

* The SFF MUST decrement SF Label entry's TTL value. If the resulting value equals zero, the SFF MUST pass the SFP OAM packet to the control plane for processing. An implementation that supports this specification MUST provide control to limit the rate of SFP OAM packets passed to the control plane for processing.

* If the TTL value is not zero, the SFP OAM packet is processed as defined in Section 6, Section 7, and Section 8 [RFC8595], according to the type of MPLS forwarding used in the SFP.

4.2. SFC Basic Unit FEC Sub-TLV

Unlike standard MPLS forwarding, based on a single label, packet forwarding defined in [RFC8595] is based on the basic unit of MPLS label stack for SFC(SFC Context Label+SF Label). A new SFC Basic Unit FEC sub-TLV with Type value (TBA1) is defined in this document. The SFC Basic Unit FEC sub-TLV MAY be used to carry the corresponding FEC of the basic unit.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Route Distinguisher (RD)                     |
|                      (8 octets)                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         SF Type               |       Reserved                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: SFC Basic Unit sub-TLV

The format of the basic unit sub-TLV is shown in Figure 2 and includes the following fields:

*Route Distinguisher (RD): 8 octets field in SFIR Route Type specific NLRI [I-D.ietf-bess-nsh-bgp-control-plane].

*SF Type: 2 octets. It is defined in [I-D.ietf-bess-nsh-bgp-control-plane] and indicates the type of SF, such as DPI, firewall, etc.

Note: [I-D.ietf-bess-nsh-bgp-control-plane] covers the BGP control plane of MPLS-SFC as well.

A node that receives an LSP ping with the Target FEC Stack TLV and the SFC Basic Unit FEC Sub-TLV included will check if it is its Route Distinguisher and whether it advertised that Service Function Type. If the validation is not passed, the SFF will generate an MPLS echo reply with an error code as defined in [RFC8029].

4.3. SFC Basic Unit Nil FEC Sub-TLV

[RFC8029] is based on the premise that one label corresponds to one FEC sub-TLV. For example, in [RFC8029] section 4.4 step 4, before the FEC validation process of an intermediate node first the node should determine FEC-stack-depth from the Downstream Detailed Mapping TLV, and then if the number of FECs in the FEC stack is greater than or equal to FEC-stack-depth, FEC validation is triggered.

In SFC-MPLS OAM, since one basic unit is related to only one FEC sub-TLV, there may be situations that the label stack in Downstream Detailed Mapping TLV contains two labels, but there is only one FEC in the FEC stack.

The SFC Basic Unit Nil Sub-TLV(TBA2) is introduced in this document to ensure that the proper validation can still be performed.

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |           SFC Context Label           |         MBZ           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |           SF Label                    |         MBZ           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: SFC Basic Unit Nil sub-TLV

SFC Context Label and SF Label are the actual label values inserted in the label stack; the MBZ fields MUST be zero when sent and ignored on receipt.

The SFC Basic Unit Nil sub-TLV, when present, MUST be immediately followed by an SFC Basic Unit sub-TLV. During FEC validation, an SFF should skip the SFC Basic Unit Nil sub-TLV and use the following SFC Basic Unit sub-TLV to validate the FEC of the basic unit.

4.4. Theory of Operation

The MPLS echo request is sent with a label stack corresponding to the SFP being tested. To trace SFC-MPLS, the Generic Associated Channel Label (GAL), which immediately follows the SFC Context label is also included.

If FEC validation is required, the SFC Basic Unit sub-TLV SHOULD be carried in the FEC stack of the request packet, and the SFC Basic Unit Nil sub-TLV MAY also be carried. A Downstream Detailed Mapping TLV MAY be included in the MPLS echo request of the SFP.

Sending an SFC echo request to the control plane is triggered by one of the following packet processing exceptions: IP TTL expiration, MPLS TTL expiration, or the receiver is SFP's egress SFF.

As described in Section 4.1.1, the packet with GAL is recognized by the SFF as an SFP OAM packet. The SFF then decrements the SF Label entry's TTL value. If the resulting value equals zero, the SFF passes the SFP OAM packet to the control plane for processing. The system that supports this specification then generates a reply message.

In "traceroute" mode the TTL of the SF Label is set successively to 1, 2, and so on. After all SFFs on the SFP send back MPLS echo reply, the sender collects information about all traversed SFFs and SFs on the rendered service path (RSP).

But the TTL processing in SR-MPLS is defined in Section 6 of [RFC8595], as follows:

> If an SFF decrements the TTL to zero, it MUST NOT send the packet and MUST discard the packet.

and it excludes TTL expiration as the exception mechanism. As a result, tracing a path of an SFC-MPLS-based service chain is problematic.

To support the tracing of an SFC, it must be changed to allow punting an OAM packet to the control plane though under throttling control.

Hence, this document updates Section 6 of [RFC8595] to state that:

> If an SFF decrements the TTL to zero, an OAM packet MAY be sent to the control plane given it does not exceed the configured rate intended to protect the system from the possible denial-of-service attack.

5. LSP Ping in SR-SFC

In SR service programming, the packet forwarding decision is made based on every single SID/label. The SR proxy SHOULD process the OAM packet for the SF when the SF is not capable of doing so.

If only the SFP connectivity check is required, the current LSP Ping for SR-MPLS [RFC8287] is sufficient.

If operators want to check more information about the SFP(service segment related SF type, SR proxy type, etc.), new FEC sub-TLVs for the service segment should be defined.

The format of the new Service Segment Sub-TLV(TBA3) is shown in Figure 4.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Service Type             | Traffic Type |Func  Identifier|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Service Segment Sub-TLV

The Service Type and Traffic Type are taken from the Service Chaining (SC) TLV defined [I-D.ietf-idr-bgp-ls-sr-service-segments].

Func(Function) Identifier: 1 octets. Function Identifier, as described in [I-D.ietf-idr-bgp-ls-sr-service-segments], identifies the function of this SID, such as Static Proxy, Dynamic Proxy, Shared Memory Proxy, Masquerading Proxy, SR(-MPLS) Aware Service etc.

There's no definition for Function Identifier field of SR-MPLS-SFC in [I-D.ietf-idr-bgp-ls-sr-service-segments] yet. If the control plane defines the Function Identifier field in the future, this draft shall be consistent with its definition.

6. IANA Considerations

This document requests assigning three new sub-TLVs from the "sub-TLVs for TLV Types 1, 16, and 21" sub-registry of the "Multi-Protocol Label Switching(MPLS) Label Switched Paths (LSPs) Ping Parameters" registry according to Table 1.

Value	Description	Reference
TBA1	SFC Basic Unit	This document
TBA2	SFC Basic Unit Nil	This document
TBA3	Service Segment	This document

Table 1: Sub-TLV Values

7. Security Considerations

This specification defines the processing of an SFP OAM packet. Such packets could be used as an attack vector. A system that supports this specification MUST provide control to limit the rate of SFP OAM packets sent to the control plane for processing.

This document defines additional MPLS LSP Ping sub-TLVs and follows the mechanisms defined in [RFC8029]. All the security considerations defined in [RFC8029] will be applicable for this document.

8. References

8.1. Normative References

[I-D.ietf-bess-nsh-bgp-control-plane] Farrel, A., Drake, J., Rosen, E., Uttaro, J., and L. Jalil, "BGP Control Plane for the Network Service Header in Service Function Chaining", Work in Progress, Internet-Draft, draft-ietf-bess-nsh-bgp-control-plane-18, 21 August 2020, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-nsh-bgp-control-plane-18>.

[I-D.ietf-idr-bgp-ls-sr-service-segments] Dawra, G., Filsfils, C., Talaulikar, K., Clad, F., Bernier, D., Uttaro, J., Decraene, B., Elmalky, H., Xu, X., Guichard, J., and C. Li, "BGP-LS Advertisement of Segment Routing Service Segments", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-ls-sr-service-segments-01, 24 April 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ls-sr-service-segments-01>.

[I-D.ietf-spring-nsh-sr] Guichard, J. N. and J. Tantsura, "Integration of Network Service Header (NSH) and Segment Routing for Service Function Chaining (SFC)", Work in

Progress, Internet-Draft, draft-ietf-spring-nsh-sr-11, 20 April 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-nsh-sr-11>.

[I-D.ietf-spring-sr-service-programming]
Clad, F., Xu, X., Filsfils, C., Bernier, D., Li, C., Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and S. Salsano, "Service Programming with Segment Routing", Work in Progress, Internet-Draft, draft-ietf-spring-sr-service-programming-06, 9 June 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-service-programming-06>.

[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>.

[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>.

[RFC7274] Kompella, K., Andersson, L., and A. Farrel, "Allocating and Retiring Special-Purpose MPLS Labels", RFC 7274, DOI 10.17487/RFC7274, June 2014, <https://www.rfc-editor.org/info/rfc7274>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[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>.

[RFC8287] Kumar, N., Ed., Pignataro, C., Ed., Swallow, G., Akiya, N., Kini, S., and M. Chen, "Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes", RFC 8287, DOI 10.17487/RFC8287, December 2017, <https://www.rfc-editor.org/info/rfc8287>.

[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>.

[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>.

8.2. Informative References

[I-D.andersson-mpls-mna-fwk] Andersson, L., Bryant, S., Bocci, M., and T. Li, "MPLS Network Actions Framework", Work in Progress, Internet-Draft, draft-andersson-mpls-mna-fwk-03, 1 June 2022, <https://datatracker.ietf.org/doc/html/draft-andersson-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>.

[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>.

[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>.

Authors' Addresses

Yao Liu
ZTE
Nanjing
China

Email: liu.yao71@zte.com.cn

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com