OPSAWG
Internet-Draft
Intended status: Standards Track
Expires: 3 April 2026

Y. Liu
T. Zhou
ZTE Corporation

30 September 2025

Export of Encapsulation Layer Information in IPFIX
draft-liu-opsawg-ipfix-muti-layer-00

Abstract

This document introduces new IPFIX IEs for encapsulation layer indication to help with the scenario when monitoring flow with multi-layer network encapsulations.

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 3 April 2026.

Copyright Notice

Table of Contents

1. Introduction

A packet may have multi-layer network encapsulations, and each layer may use the same or different network encapsulation headers. e.g, IP in IP encapsulation [RFC2003], which is an IP tunneling mechanism that encapsulates one IP packet in another IP packet, typical IP-in-IP scenario includes IPv4-in-IPv4, IPv6-in-IPv6, IPv4-in-IPv6 and IPv6-in-IPv4.

With the deployment of SRv6, the appearance of packets with IPv4-in-IPv6 or IPv6-in-IPv6 encapsulation becomes more and more common in the network. And there may be more than two network encapsulation layers in one packet as analyzed in section 3.1.

When monitoring a traffic flow with multiple encapsulations, e.g IP-in-IP, a typical use case is to answer the following questions:

* Which tunnel are the packets steered into (e.g, identified by the outmost IP header) ?

* What are the details of the inner packet (e.g, identified by the innermost IP header) ?

However, based on the existing IPFIX mechanisms, it is not easy to differentiate between IEs of different encapsulation layers.

This document aims to solve this problem by introducing new IPFIX IEs for encapsulation layer indication.

2. 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.

This document makes use of the terms defined in [RFC7011], [RFC8402] and [RFC8754].

The following terms are used as defined in [RFC7011]:

* IPFIX
* IPFIX Information Elements
* Metering Process
* Template
* Collector
* IPFIX Device

The following terms are used as defined in [RFC8402]:

* Segment Routing (SR)
* Segment List
* SRv6

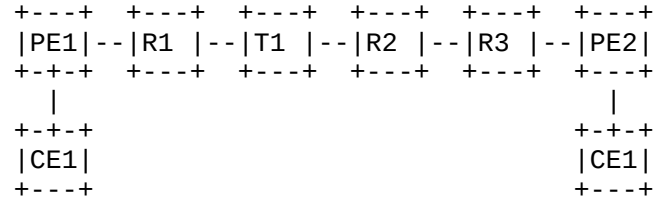
The following terms are used as defined in [RFC8754]:

* SRH

3. Use Cases and Requirements

3.1. Use Cases

There may be more than two network encapsulation layers in one packet. As shown in the figure below.



CE1 sends IPv6 packets to CE2.

Packet leaving CE1: IPv6SA=CE1,DA=CE2>

PE1 performs SRv6 function H.Encaps with SRv6 Policy, and the corresponding SID list is <SID-R1,SID-T1,SID-PE2>, wherein SID-T1 is a BSID initiated on node T1.

Packet leaving PE1: IPv6<SA=PE1,DA=SID-R1><SID-R1,SID-T1,SID-PE2>IPv6<SA=CE1,DA=CE2>

When the packet arrives at T1, based on BSID SID-T1 , T1 performs End.B6.Encaps [RFC8986] and encapsulate the third IPv6 header onto the packet with the corresponding SID list <SID-R2,SID-R3>.

Packet leaving T1:IPv6<SA=T1,DA=R2><SID-R2,SID-R3>IPv6<SA=PE1,DA=SID-PE2>lt;SID-R1,SID-T1,SID-PE2>IPv6<SA=CE1,DA=CE2>

So the packet observed at node R2 has three IPv6 headers in it.

Based on different goals of the network monitor, the data collection scenario may include one of the following:

* The network monitor wants to collect the information of the outmost SRH and the inner SRH of the same packet.

* The network monitor only wants to collect the information of the outmost IPv6 header.

* The network monitor only wants to collect the information of the innermost IPv6 header.

* The network monitor wants to collect the SA of the outmost IPv6 header (i.e, the starting point of the tunnel) and the DA of the innermost IPv6 header(i.e, the final destination of the packet)

3.2. Requirements

Based on the scenarios described in section 3.1, the information collection requirements in IPFIX for multi-layer encapsulated packets include:

* Req-a: Collecting the same fields from both the outer and inner packet headers at the same time.

* Req-b: Collecting only the fields from the inner packet header.

* Req-c: Collecting only the fields from the outer packet header

* Req-d: Collecting different fields from the outer and inner packet headers at the same time.

Req-a can be fulfilled leveraging the existing mechanism. As described in [RFC7011], if the same element appears multiple times in an IPFIX template, it should be processed in order. For example, when exporting the two source IP addresses of an IPv6-in-IPv6 packet, the first sourceIPv6Address Information Element occurrence should be the IPv6 address of the outer header, while the second occurrence should be the address of the inner header.

However, Req-b,Req-b and Req-d can not be meet using standard method by far. When receiving a IPFIX message with a certain IE(e.g, sourceIPv6Address) from the IPFIX Device, the collector is not able to tell which layer this IE belongs to for traffic flow with multi-layer encapsulations.

4. IPFIX IEs for Encapsulation Layer Information

This section defines several Encapsulation Layer IEs for network encapsulation layer indication. When there is no need to differentiate between these IEs in this document, they will be collectively referred to as "encapsulation layer IE".

4.1. encapLayerTop

A new IE "encapLayerTop" is defined in this section to indicate which IEs in the IPFIX messages belongs to the outmost network encapsulation layer.

Name: encapLayerTop

ElementID: TBD1

Description: A 16-bit identifier indicating that the IEs follows

immediately after it till the next Encapsulation Layer IE belong to the outmost network encapsulation layer (e.g, from the outmost Ethernet header to the first IP header). If there's not any other Encapsulation Layer IE exists in the Template, it means that all the IEs following encapLayerTop belong to the outmost network encapsulation layer. This IE has a fixed value of 0xF.

Abstract Data Type: unsigned16

Data Type Semantics: identifier

Reference: This document.

4.2. encapLayer2

A new IE "encapLayer2" is defined in this section to indicate which IEs in the IPFIX messages belongs to the second network encapsulation layer.

Name: encapLayer2

ElementID: TBD2

Description: A 16-bit identifier indicating that the IEs follows immediately after it till the next Encapsulation Layer IE belong to the second network encapsulation layer. If there's not any other Encapsulation Layer IE exists in the Template, it means that all the IEs following encapLayer2 belong to the second network encapsulation layer. This IE has a fixed value of 0xF.

Abstract Data Type: unsigned16

Data Type Semantics: identifier

Reference: This document.

4.3. encapLayer3

A new IE "encapLayer3" is defined in this section to indicate which IEs in the IPFIX messages belongs to the third network encapsulation layer.

Name: encapLayer3

ElementID: TBD2

Description: A 16-bit identifier indicating that the IEs follows

immediately after it till the next Encapsulation Layer IE belong to the third network encapsulation layer. If there's not any other Encapsulation Layer IE exists in the Template, it means that all the IEs following encapLayer3 belong to the third network encapsulation layer. This IE has a fixed value of 0xF.

Abstract Data Type: unsigned16

Data Type Semantics: identifier

Reference: This document.

5. Operational Considerations

To generate Flow Records with IEs for encapsulation layer included, the metering process SHOULD recognize the encapsulation layer of the corresponding fields in the packet. This is mainly based on local implementation and the details are out of the scope of this document.

Each encapsulation layer IE SHALL NOT appear more than once more in a Template. If there's more than one encapsulation layer IE of the same type in the Template, the Collecting Process MUST ignore the Template and the Collecting Process SHOULD log the error.

As in [RFC5012], the Information Elements are derived from fields of packets or from packet treatment. For IEs that are not related with header fields, whether they are covered by scope of the encapsulation layer IE, they SHOULD be processed following the existing specifications.

For IEs of Header Fields that are not in the scope of encapsulation layer IE, e.g, there're IEs of Header Fields in the Template before the appearance of Encapsulation Layer IEs, they SHOULD be processed properly based on the default behavior of the Collector, how the Collector would process them is out of the scope of this document.

6. Security Considerations

There are no additional security considerations regarding allocation of these new IPFIX IEs compared to [RFC7012].

7. IANA Considerations

This document requests IANA to create new IEs under the "IPFIX Information Elements" registry [RFC7012] available at [IANA-IPFIX].

Element ID	Name	Reference
TBD1	encapLayerTop	This document
TBD2	encapLayer2	This document
TBD3	encapLayer3	This document

8. References

8.1. Normative References

[IANA-IPFIX] IANA, "IP Flow Information Export (IPFIX) Entities", <https://www.iana.org/assignments/ipfix>.

[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>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC7012] Claise, B., Ed. and B. Trammell, Ed., "Information Model for IP Flow Information Export (IPFIX)", RFC 7012, DOI 10.17487/RFC7012, September 2013, <https://www.rfc-editor.org/info/rfc7012>.

[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>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

8.2. Informative References

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, DOI 10.17487/RFC2003, October 1996, <https://www.rfc-editor.org/info/rfc2003>.

[RFC5012] Schulzrinne, H. and R. Marshall, Ed., "Requirements for Emergency Context Resolution with Internet Technologies", RFC 5012, DOI 10.17487/RFC5012, January 2008, <https://www.rfc-editor.org/info/rfc5012>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

Authors' Addresses

Yao Liu
ZTE Corporation
China
Email: liu.yao71@zte.com.cn

Taoran Zhou
ZTE Corporation
China
Email: zhou.taoran@zte.com.cn