Abstract

This document specifies new IP Flow Information Export (IPFIX) Information Elements (IEs) to solve some issues with existing ipv6ExtensionHeaders and tcpOptions IPFIX IEs, especially the ability to export any observed IPv6 extension headers or TCP options.

Discussion Venues

This note is to be removed before publishing as an RFC.

Discussion of this document takes place on the Operations and Management Area Working Group Working Group mailing list (opsawg@ietf.org), which is archived at https://mailarchive.ietf.org/arch/browse/opsawg/.

Source for this draft and an issue tracker can be found at https://github.com/boucadair/ipfix-tcpoptions-and-v6eh.

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 20 April 2024.
1. Introduction

This document specifies new IP Flow Information Export (IPFIX) Information Elements (IEs) to solve a set of issues encountered with the specifications of ipv6ExtensionHeaders (to export IPv6 extension...
headers) and tcpOptions (to export TCP options) IEs. More details about these issues are provided in the following sub-sections.

1.1. Issues with ipv6ExtensionHeaders Information Element

The specification of ipv6ExtensionHeaders IPFIX IE does not:

* Cover the full extension headers range ([RFC8200]).

* Specify the procedure to follow when all bits are exhausted.

* Specify a means to export the order and the number of occurrences of a given extension header.

* Specify how to automatically update the IANA IPFIX registry ([IANA-IPFIX]) when a new value is assigned in [IANA-EH]. Only a frozen set of extension headers can be exported using the ipv6ExtensionHeaders IE.

* Specify whether the exported values match the full enclosed values or only up to a limit imposed by hardware or software (e.g., Section 1.1 of [RFC8883]).

* Specify how to report the length of IPv6 extension headers.

* Optimize the encoding.

* Explain the reasoning for reporting values which do not correspond to extension headers (e.g., "Unknown Layer 4 header" or "Payload compression header").

Section 3 addresses these issues.

1.2. Issues with tcpOptions Information Element

The specification of tcpOptions IPFIX IE does not:

* Describe how any observed TCP option in a Flow can be exported using IPFIX. Only TCP options having a kind <= 63 can be exported in a tcpOptions IPFIX IE.

* Support means to report the observed Experimental Identifiers (ExIDs) that are carried in shared TCP options (kind=253 or 254) ([RFC6994]).

* Optimize the encoding.

Section 4 addresses these issues.
2. Conventions and Definitions

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], when, and only when, they appear in all capitals, as shown here.

This document uses the IPFIX-specific terminology (Information Element, Template, Collector, Data Record, Flow, Flow Record, Exporting Process, Collecting Process, etc.) defined in Section 2 of [RFC7011]. As in [RFC7011], these IPFIX-specific terms have the first letter of a word capitalized.

Also, the document uses the terms defined in [RFC8200] and [RFC9293].

In addition, the document makes use of the following term:

**Extension header chain:** Refers to the chain of extension headers that are present in an IPv6 packet.

This term should not be confused with the IPv6 header chain, which includes the IPv6 header, zero or more IPv6 extension headers, and zero or a single Upper-Layer Header.

3. Information Elements for IPv6 Extension Headers

The definition of the ipv6ExtensionHeaders IE is updated in [I-D.ietf-opsawg-ipfix-fixes] to address some of the issues listed in Section 1.1. Because some of these limitations can't be addressed by simple updates to ipv6ExtensionHeaders, this section specifies a set of new IEs to address all the ipv6ExtensionHeaders IE limitations.

3.1. ipv6ExtensionHeadersFull Information Element

**Name:** ipv6ExtensionHeadersFull

**ElementID:** TBD1

**Description:** IPv6 extension headers observed in packets of this Flow. The information is encoded in a set of bit fields. For each IPv6 extension header, there is a bit in this set. The bit is set to 1 if any observed packet of this Flow contains the corresponding IPv6 extension header. Otherwise, if no observed
packet of this Flow contained the respective IPv6 extension header, the value of the corresponding bit is 0.

The IPv6 extension header associated with each bit is provided in [NEW_IPFIX_IPv6EH_SUBREGISTRY]. Bit 0 corresponds to the least-significant bit in the ipv6ExtensionHeadersFull IE while bit 254 corresponds to the most-significant bit of the IE. In doing so, few octets will be needed to encode common IPv6 extension headers when observed in a Flow.

The "No Next Header" (59) value is used if there is no upper-layer header in an IPv6 packet. Even if the value is not considered as an extension header as such, the corresponding bit is set in the ipv6ExtensionHeadersFull IE whenever that value is encountered in the Flow.

Several extension header chains may be observed in a Flow. These extension headers may be aggregated in one single ipv6ExtensionHeadersFull Information Element or be exported in separate ipv6ExtensionHeadersFull IEs, one for each extension header chain.

This Information Element **SHOULD NOT** be exported if ipv6ExtensionHeaderCount Information Element is also present.

Abstract Data Type: unsigned

Data Type Semantics: flags

Additional Information: See the assigned bits to each IPv6 extension header type in [NEW_IPFIX_IPv6EH_SUBREGISTRY].

See [IANA-EH] for assigned extension header types.

See Section 4 of [RFC8200] for the general definition of IPv6 extension headers.

Reference: This Document

Note to the RFC Editor: Please replace [NEW_IPFIX_IPv6EH_SUBREGISTRY] with the link to the "ipv6ExtensionHeaders Bits" registry created by [I-D.ietf-opsawg-ipfix-fixes].

3.2. ipv6ExtensionHeaderCount Information Element

Name: ipv6ExtensionHeaderCount

ElementID: TBD2
Description:
As per Section 4.1 of [RFC8200], IPv6 nodes must accept and attempt to process extension headers in occurring any number of times in the same packet. This Information Element echoes the order of extension headers and number of consecutive occurrences of the same extension header type in a Flow.

If several extension header chains are observed in a Flow, each header chain MUST be exported in a separate ipv6ExtensionHeaderCount IE.

The same extension header type may appear several times in an ipv6ExtensionHeaderCount Information Element. For example, if an IPv6 packet of a Flow includes a Hop-by-Hop Options header, a Destination Options header, a Fragment header, and Destination Options header, the ipv6ExtensionHeaderCount Information Element will report two counts of the Destination Options header: the occurrences that are observed before the Fragment header and the occurrences right after the Fragment header.

Abstract Data Type: unsigned64

Data Type Semantics: identifier

Additional Information: See the assigned IPv6 extension header types in [IANA-EH].

See [RFC8200] for the general definition of IPv6 extension headers.

Reference: This-Document

3.3. ipv6ExtensionHeadersLimit Information Element

Name: ipv6ExtensionHeadersLimit

ElementID: TBD3

Description: When set to "true", this Information Element indicates that the exported extension headers information (e.g., ipv6ExtensionHeaderFull or ipv6ExtensionHeaderCount) does not
match the full enclosed extension headers, but only up to a limit that is typically set by hardware or software.

When set to "false", this Information Element indicates that the exported extension header information matches the full enclosed extension headers.

When this Information Element is absent, this is equivalent to returning an ipv6ExtensionHeadersLimit Information Element with a value set to "false".

Abstract Data Type: boolean

Data Type Semantics: default

Additional Information: See Section 4 of [RFC8200] for the general definition of IPv6 extension headers.

See [RFC8883] for an example of IPv6 packets processing due to limits on extension headers.

Reference: This-Document

3.4. ipv6ExtensionHeadersChainLength Information Element

Name: ipv6ExtensionHeadersChainLength

ElementID: TBD4

Description: In theory, there are no limits on the number of IPv6 extension headers that may be present in a packet other than the path MTU. However, it was regularly reported that IPv6 packets with extension headers are often dropped in the Internet.

As discussed in Section 1.2 of [RFC8883], some hardware devices implement a parsing buffer of a fixed size to process packets, including all the headers. When the aggregate length of headers of an IPv6 packet exceeds that size, the packet will be discarded or deferred to a slow path.

The ipv6ExtensionHeadersChainLength IE is used to report, in octets, the length of an extension header chain observed in a Flow. The length is the sum of the length of all extension headers of the chain. Exporting such information may help identifying root causes of performance degradation, including packet drops.

Abstract Data Type: unsigned

Data Type Semantics: identifier
4. Information Elements for TCP Options

The definition of the tcpOptions IE is updated in [I-D.ietf-opsawg-ipfix-fixes] to address some of the issues listed in Section 1.2. Because some of these limitations can't be addressed by simple updates to tcpOptions, this section specifies a set of new IEs to address all the tcpOptions IE limitations.

4.1. tcpOptionsFull Information Element

This section specifies a new IE to cover the full TCP options range.

Name: tcpOptionsFull

ElementID: TBD5

Description: TCP options in packets of this Flow. The information is encoded in a set of bit fields. For each TCP option, there is a bit in this set. The bit is set to 1 if any observed packet of this Flow contains the corresponding TCP option. Otherwise, if no observed packet of this Flow contained the respective TCP option, the value of the corresponding bit is 0.

Options are mapped to bits according to their option numbers. Option number X is mapped to bit position "254 - X". This approach allows an observer to export any observed TCP option even if it does support that option and without requiring updating a mapping table.

Abstract Data Type: unsigned

Data Type Semantics: flags

Additional Information: See the assigned TCP option kinds at [IANA-TCP].

See [RFC9293] for the general definition of TCP options.
4.2. tcpSharedOptionExID Information Element

Name: tcpSharedOptionExID

ElementID: TBD6

Description: Observed 2-byte (or 4-byte) Experiments IDs (ExIDs) in a shared TCP option (Kind=253 or 254) in a Flow. The information is encoded in a set of 16-bit (or 32-bit) fields. Each 16-bit (or 32-bit) field carries the observed 2-byte (or 4-byte) ExID in a shared option.

The value should be encoded in fewer octets as per the guidelines in Section 6.2 of [RFC7011].

Abstract Data Type: octetArray

Data Type Semantics: identifier

Additional Information: See assigned ExIDs at [IANA-TCP-EXIDs].

See [RFC9293] for the general definition of TCP options.

See [RFC6994] for the shared use of experimental TCP Options.

Reference: This-Document

5. Operational Considerations

5.1. IPv6 Extension Headers

The value of ipv6ExtensionHeadersFull and ipv6ExtensionHeaderCount IEs should be encoded in fewer octets as per the guidelines in Section 6.2 of [RFC7011].

If an implementation determines that it includes an extension header that it does no support, then the exact observed code of that extension header will be echoed in the ipv6ExtensionHeaderCount IE (Section 3.2). How an implementation disambiguates between unknown upper-layer protocols vs. extension headers is not IPFIX-specific. Readers may refer, for example, to Section 2.2 of [RFC8883] for a behavior of an intermediate nodes that encounters an unknown Next Header type. It is out of the scope of this document to discuss those considerations.

The ipv6ExtensionHeadersLimit IE (Section 3.3) may or may not be present when the ipv6ExtensionHeadersChainLength IE (Section 3.4) is also present as these IEs are targeting distinct properties of extension headers handling.
5.2. TCP Options

The value of tcpOptionsFull IE should be encoded in fewer octets as per the guidelines in Section 6.2 of [RFC7011].

If a TCP Flow contains packets with a mix of 2-byte and 4-byte Experiment IDs, two different Template Records may be used: one containing the tcpSharedOptionExID IE with a length of 2 bytes and another similar one containing the tcpSharedOptionExID IE with a length of 4 bytes.

6. Examples

6.1. IPv6 Extension Headers

This section provides few examples to illustrate the use of some IEs defined in the document.

**Figure 1** provides an example of reported values in an ipv6ExtensionHeadersFull IE for an IPv6 Flow in which only the IPv6 Destination Options header is observed. One octet is sufficient to report these observed options. Concretely, the ipv6ExtensionHeadersFull IE will be set to 1.

**Figure 1: A First Example of Extension Headers**

**Figure 2** provides another example of reported values in an ipv6ExtensionHeadersFull IE for an IPv6 Flow in which the IPv6 Hop-by-Hop Options, Routing, and Destination Options headers are observed. One octet is sufficient to report these observed options. Concretely, the ipv6ExtensionHeadersFull IE will be set to 19.

**Figure 2: A Second Example of Extension Headers**
6.2. TCP Options

Given TCP kind allocation practices and the option mapping defined in Section 4.1, fewer octets are likely to be used for Flows with common TCP options.

Figure 3 shows an example of reported values in a tcpOptionsFull IE for a TCP Flow in which End of Option List, Maximum Segment Size, and Window Scale options are observed. One octet is sufficient to report these observed options. Concretely, the tcpOptionsFull IE will be set to 15.

Figure 3: First Example of TCP Options

7. Security Considerations

IPFIX security considerations are discussed in Section 11 of [RFC7011]. This document does not add new security considerations for exporting IEs other than those already discussed in Section 8 of [RFC7012].

8. IANA Considerations

This document requests IANA to add the following new IPFIX IEs to the IANA IPFIX registry [IANA-IPFIX]:

<table>
<thead>
<tr>
<th>Value</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD1</td>
<td>ipv6ExtensionHeadersFull</td>
<td>Section 3.1 of This-Document</td>
</tr>
<tr>
<td>TBD2</td>
<td>ipv6ExtensionHeaderCount</td>
<td>Section 3.2 of This-Document</td>
</tr>
<tr>
<td>TBD3</td>
<td>ipv6ExtensionHeaderLimit</td>
<td>Section 3.3 of This-Document</td>
</tr>
<tr>
<td>TBD4</td>
<td>ipv6ExtensionHeadersChainLength</td>
<td>Section 3.4 of This-Document</td>
</tr>
<tr>
<td>TBD5</td>
<td>tcpOptionsFull</td>
<td>Section 4.1 of This-Document</td>
</tr>
<tr>
<td>TBD6</td>
<td>tcpSharedOptionExID</td>
<td>Section 4.2 of This-Document</td>
</tr>
</tbody>
</table>

Table 1: New IPFIX Information Elements

9. References

9.1. Normative References

[IANA-EH]


9.2. Informative References


Acknowledgments

Thanks to Paul Aitken and Eric Vyncke for the review and comments.

Authors' Addresses

Mohamed Boucadair
Orange

Email: mohamed.boucadair@orange.com

Benoit Claise
Huawei

Email: benoit.claise@huawei.com