OPSAWG
Internet-Draft
Intended status: Standards Track
Expires: 23 November 2025

T. Graf
Swisscom
G. Fioccola
T. Zhou
Huawei
Y. Zhu
China Telecom
M. Cociglio
Telecom Italia
22 May 2025

IP Flow Information Export (IPFIX) Alternate-Marking Information Elements
draft-ietf-opsawg-ipfix-alt-mark-03

Abstract

This document specifies the IP Flow Information Export (IPFIX) Information Elements (IEs) to export Alternate Marking measurement data.

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 23 November 2025.

Copyright Notice

Table of Contents

1. Introduction

Alternate-Marking Method (AltMark) [RFC9341] [RFC9342] is a technique used to measure packet loss, delay, and jitter on in-flight packets.

[I-D.ietf-ippm-alt-mark-deployment] provides a framework for Alternate Marking deployments and includes considerations and guidance for application and methodology. The IP Flow Information Export (IPFIX) protocol [RFC7011] [RFC7012] is considered for data export in Section 6.1 of [I-D.ietf-ippm-alt-mark-deployment].

[RFC7012] defines the data types and management policy for the information model of the IPFIX protocol [RFC7011]. This document defines the new IPFIX Information Elements (IEs) for the Alternate Marking Method.

1.1. Requirements 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. AltMark IPFIX Information Elements

This section describes existing IEs [IANA-IPFIX] that are relevant for the Alternate Marking application and also introduces new IEs.

With AltMark [RFC9341] [RFC9342], each node needs to export the packet counters and timestamps at each period for the monitored flow, according to AltMark operation. To identify and export telemetry data for an AltMark monitored flow, it is needed a combination of already existing IEs and new IEs, which are introduced in this document. A flow can be identified using IEs such as source address, destination address, protocol and ports. But, according to [RFC9343] and any other AltMark protocol extensions, it is also needed to define new IEs (the flow identifier FlowMonID, the Period Number, the Loss flag and the Delay flag) to complete the AltMark information to be exported.

2.1. Flow Decomposition

Data decomposition can be achieved on an Alternate-Marking-aware node where IPFIX is exported or on the IPFIX data collection.

The ipPayloadPacketSection(IE314) Information Element (IE) carries a series of n octets from the IP payload, starting sectionOffset(IE409) octets into the IP payload.

When decomposed at the data collection, the packet header sections, as example the IPv6 options type header described in Section 3.1 of [RFC9343] or the Segment Routing header TLV as described in Section 3.1 of [I-D.fz-spring-srv6-alt-mark] containing the FlowMonID, Loss and Delay flag are being exposed as part of ipPayloadPacketSection(IE314), defined in Section 4.2 of [RFC7133].

The IPv4 payload is that part of the packet that follows the IPv4 header and any options. The IPv6 payload is the rest of the packet following the 40-octet IPv6 header. Note that any extension headers present are considered part of the payload. The sectionExportedOctets(IE410) expresses how much data was observed, while the remainder is padding.

2.2. Flow Aggregation

An Aggregated Flow is simply an IPFIX Flow generated from Original Flows by an Intermediate Aggregation Process.

When being decomposed on an Alternate-Marking-aware node, new IPFIX entities for FlowMonID (TBD1), LossFlag (TBD2) and DelayFlag (TBD3) are needed so that the data can now be aggregated according to Section 5 of [RFC7015].

According to Section 4 of [RFC7015] new Flow Keys may be derived from existing Flow Keys or "promoted" from specific non-key fields.

Therefore FlowMonID, LossFlag and DelayFlag are considered Flow Key fields.

2.3. Flow Correlation

The following IPFIX entities are of interest to describe the relationship to the forwarding topology and the control-plane.

* Hostname, ingressInterface(IE10) and egressInterface(IE14) describes on which node which logical ingress and egress interfaces have been used to forward the packet.

* Hostname and egressPhysicalInterface(IE253) describes on which node which physical egress interfaces have been used to forward the packet.

* Hostname and ipNextHopIPv4Address(IE15) or ipNextHopIPv6Address(IE62), describes the forwarding path to which next-hop IP address the packets are forwarded to.

* Hostname and mplsTopLabelIPv4Address(IE47) or srhActiveSegmentIPv6 (IE495) describes the forwarding path to which MPLS top label IPv4 address or SRv6 active segment the packets are forwarded to.

* BGP communities [RFC1997] are often used for setting a path priority or service selection. bgpDestinationExtendedCommunityList(IE488) or bgpDestinationCommunityList(IE485) or bgpDestinationLargeCommunityList(IE491) describes which group of prefixes have been used to forward the packet.

* Hostname, sourceIPv4Address(IE8) or sourceIPv6Address(IE27), sourceTransportPort(IE7), destinationIPv4Address(IE12) or destinationIPv6Address(IE28), destinationTransportPort(IE11), protocolIdentifier(IE4) describe the forwarding path on each node from each IPv4 or IPv6 source address to a specific application in the network.

Note that, in case of Link Aggregation Group (LAG) interface, the ingressInterface IE and egressInterface IE can be used to refer the logical LAG port, while ingressPhysicalInterface IE and egressPhysicalInterface IE can be used to indicate the physical interfaces which are members of the LAG port.

2.4. Flow Measurements

To calculate loss, the packet count can be based upon octetDeltaCount(IE1) or packetDeltaCount(IE2).

While, to calculate delay, either flowStartSeconds(IE150), flowStartMilliseconds(IE152), flowStartMicroseconds(IE154) or flowStartNanoseconds(IE156), can be used depending on timestamp granularity requirements. It is also possible to use flowEndSeconds(IE151), flowEndMilliseconds(IE153), flowEndMicroseconds(IE155) or flowEndNanoseconds(IE157).

It is also defined the PeriodNumber (TBD4), which is needed for Alternate-Marking measurement correlation as per [I-D.ietf-ippm-alt-mark-deployment].

3. Performance Measurement Considerations

[I-D.ietf-ippm-alt-mark-deployment] describes how to manage and deploy the AltMark method and IPFIX can be used to export the telemetry data to the Network Management System (NMS).

An Alternate-Marking Domain consists of marking nodes, unmarking nodes, and transit nodes. These nodes are all IPFIX observation points, while the IPFIX observation domain is the AltMark measurement domain.

As specified in the previous sections, the Flow Keys used to define a flow are, at minimum, FlowMonID (TBD1), LossFlag (TBD2) and DelayFlag (TBD3). The traditional ’5-tuple’ Flow Key of source and destination IP address, source and destination transport port, and transport protocol can also be combined with the FlowMonID, LossFlag and DelayFlag.

The AltMark Flows are defined separately for loss measurements and delay measurements. In particular, the flow for loss measurements can be aggregated using the 5-tuple, the FlowMonID and the LossFlag; while the flow for delay measurements can be aggregated using the 5-tuple, the FlowMonID and the DelayFlag.

The Flow Record, which is observed at each Observation Point, contains the characteristic properties of the Flow together with the the measured properties (i.e. packet count and timestamps). Note that, as specified in [I-D.ietf-ippm-alt-mark-deployment], the PeriodNumber (TBD4) is calculated on each Observation Point as the modulo of the local time and the interval of the marking time period. The PeriodNumber is associated to each Flow Record.

The Flow Records are different for loss measurements and delay measurements. The flow record for loss measurements can be composed by the 5-tuple, the FlowMonID, the LossFlag, the PeriodNumber and the packetDeltaCount(IE2); while the flow record for delay measurements can be composed by the 5-tuple, the FlowMonID, the DelayFlag, the PeriodNumber and the flowStartMicroseconds(IE154), flowEndMicroseconds(IE155).

The IPFIX Metering Process parameters, like the IPFIX Template Record, that generate the Flow Records must be reported to provide the complete measurement context.

The AltMark marking, transit and unmarking nodes can be Exporters and export the data record. The periodicity or export policy is configurable and it must be in line with the AltMark period, as specified in [I-D.ietf-ippm-alt-mark-deployment].

The NMS acts as Collector. The loss and delay metrics are computed on the NMS by comparing the packet counts and timestamps at each AltMark period, according to the AltMark technique [RFC9341] [RFC9342].

4. IANA Considerations

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

The allocation policy of these new Information Elements is Expert Review (Section 4.5 of [RFC8126]).

The code points are defined in Table 1.

Element ID	Name	Reference
TBD1	FlowMonID	[RFC-to-be], RFC9341, RFC9342, RFC9343
TBD2	LossFlag	[RFC-to-be], RFC9341, RFC9342, RFC9343
TBD3	DelayFlag	[RFC-to-be], RFC9341, RFC9342, RFC9343
TBD4	PeriodNumber	[RFC-to-be], [I-D.ietf-ippm-alt-mark-deployment]

Table 1: "IPFIX Alternate-Marking" Registry

4.1. FlowMonID

Name: FlowMonID

Element ID: TBD1

Description: The Flow Monitoring Identification (FlowMonID) is described in [RFC9343] and identifies the monitored flows with AltMark. It is 20-bit unsigned integer encoded in the 20 least significant bits of the 32 bits, while the other bits are set to 0. It MUST be set pseudo-randomly by the source node or by a centralized controller. It is to be noted that a new element has been defined since the flowid (IE148) can be used for other purposes and simultaneously with FlowMonID.

Abstract Data Type: unsigned32

Data Type Semantics: identifier

4.2. Loss flag

Name: LossFlag

Element ID: TBD2

Description: Loss flag (L flag) for Packet Loss Measurement as described in [RFC9343].

Abstract Data Type: boolean

Data Type Semantics: flags

4.3. Delay flag

Name: DelayFlag

Element ID: TBD3

Description: Delay flag (D flag) for Single Packet Delay Measurement as described in [RFC9343].

Abstract Data Type: boolean

Data Type Semantics: flags

4.4. Period Number

Name: PeriodNumber

Element ID: TBD4

Description: The Period Number (PN), described in [I-D.ietf-ippm-alt-mark-deployment], is used to help to determine the packet counts related to the same block of markers, or the timestamps related to the same marked packet. The PN is associated with each packet count and timestamp reported.

Abstract Data Type: unsigned64

Data Type Semantics: identifier

5. Security Considerations

Alternate Marking [RFC9341] and Multipoint Alternate Marking [RFC9342] analyze different security concerns and related solutions. These aspects are valid and applicable also to this document. In particular the fundamental security requirement is that Alternate Marking MUST only be applied in a specific limited domain, as also mentioned in [RFC8799].

There are no additional security considerations regarding allocation of these new IPFIX IEs compared to [RFC7012]. The IEs described in this document export AltMark telemetry data. Applications and operators using the IEs described in this document must evaluate the sensitivity of this information in their implementation context and apply the storage guidance in Section 11.8 of [RFC7011] as appropriate.

6. Acknowledgements

The authors of this document would like to thank Greg Mirsky, Alex Huang Feng and Mohamed Boucadair for their comments and reviews.

7. Contributors

Fabio Bulgarella
Telecom Italia
Email: fabio.bulgarella@guest.telecomitalia.it

Massimo Nilo
Telecom Italia
Email: massimo.nilo@telecomitalia.it

Fabrizio Milan
Telecom Italia
Email: fabrizio.milan@telecomitalia.it

8. References

8.1. Normative References

[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>.

[RFC7015] Trammell, B., Wagner, A., and B. Claise, "Flow Aggregation for the IP Flow Information Export (IPFIX) Protocol", RFC 7015, DOI 10.17487/RFC7015, September 2013, <https://www.rfc-editor.org/info/rfc7015>.

[RFC7133] Kashima, S., Kobayashi, A., Ed., and P. Aitken, "Information Elements for Data Link Layer Traffic Measurement", RFC 7133, DOI 10.17487/RFC7133, May 2014, <https://www.rfc-editor.org/info/rfc7133>.

[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>.

[RFC9341] Fioccola, G., Ed., Cociglio, M., Mirsky, G., Mizrahi, T., and T. Zhou, "Alternate-Marking Method", RFC 9341, DOI 10.17487/RFC9341, December 2022, <https://www.rfc-editor.org/info/rfc9341>.

[RFC9342] Fioccola, G., Ed., Cociglio, M., Sapio, A., Sisto, R., and T. Zhou, "Clustered Alternate-Marking Method", RFC 9342, DOI 10.17487/RFC9342, December 2022, <https://www.rfc-editor.org/info/rfc9342>.

[RFC9343] Fioccola, G., Zhou, T., Cociglio, M., Qin, F., and R. Pang, "IPv6 Application of the Alternate-Marking Method", RFC 9343, DOI 10.17487/RFC9343, December 2022, <https://www.rfc-editor.org/info/rfc9343>.

8.2. Informative References

[I-D.fz-spring-srv6-alt-mark] Fioccola, G., Zhou, T., Cociglio, M., Mishra, G. S., wang, X., and G. Zhang, "Application of the Alternate Marking Method to the Segment Routing Header", Work in Progress, Internet-Draft, draft-fz-spring-srv6-alt-mark-11, 16 April 2025, <https://datatracker.ietf.org/doc/html/draft-fz-spring-srv6-alt-mark-11>.

[I-D.ietf-ippm-alt-mark-deployment] Fioccola, G., Zhu, K., Graf, T., Nilo, M., and L. Zhang, "Alternate Marking Deployment Framework", Work in Progress, Internet-Draft, draft-ietf-ippm-alt-mark-deployment-03, 27 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-ippm-alt-mark-deployment-03>.

[IANA-IPFIX] "IANA, IPFIX", <https://www.iana.org/assignments/ipfix/ipfix.xhtml>.

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <https://www.rfc-editor.org/info/rfc1997>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

Authors' Addresses

Thomas Graf
Swisscom
Binzring 17
CH-8045 Zurich
Switzerland
Email: thomas.graf@swisscom.com

Giuseppe Fioccola
Huawei
Palazzo Verrocchio, Centro Direzionale Milano 2
20054 Segrate (Milan)
Italy
Email: giuseppe.fioccola@huawei.com

Tianran Zhou
Huawei
156 Beiqing Rd.
Beijing
100095
China
Email: zhoutianran@huawei.com

Yongqing Zhu
China Telecom
China
Email: zhuyq8@chinatelecom.cn

Mauro Cociglio
Telecom Italia
Italy
Email: mauro.cociglio@outlook.com