

IP Flow Information Export
Internet-Draft
Intended status: Standards Track
Expires: January 28, 2021

C. Munukutla
S. Vaid
Juniper Networks, Inc.
A. Mahale
D. Patel
Google, Inc.
July 27, 2020

**IP Flow Information Export (IPFIX) Information Elements Extension for
Forwarding Exceptions
draft-mvmd-opsawg-ipfix-fwd-exceptions-00**

Abstract

This draft proposes couple of new Forwarding exceptions related Information Elements (IEs) and Templates for the IP Flow Information Export (IPFIX) protocol. These new Information Elements and Exception Template can be used to export information about any forwarding errors in a network. This essential information is adequate to correlate packet drops to any control plane entity and map it to an impacted service. Once exceptions are correlated to a particular entity, an action can be assigned to mitigate such problems essentially enabling self-driving networks.

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 January 28, 2021.

Copyright Notice

Copyright (c) 2020 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](#)
- [1.2.](#) Requirements Language [3](#)
- [2.](#) Scope [3](#)
- [3.](#) Information Elements [4](#)
- [4.](#) New Information Elements [6](#)
- [4.1.](#) Proposed New Information Elements [6](#)
- [4.2.](#) Definition of Exceptions [6](#)
- [5.](#) Exception Templates [7](#)
- [5.1.](#) IPFIX Exception Template 1 for Forwarding Exceptions . . [7](#)
- [5.2.](#) IPFIX Exception Template 2 for Forwarding Exceptions . . [8](#)
- [6.](#) IANA Considerations [9](#)
- [6.1.](#) Information Elements [9](#)
- [6.2.](#) Forwarding Exception Codes [10](#)
- [7.](#) Security Considerations [10](#)
- [8.](#) Contributors [10](#)
- [9.](#) References [11](#)
- [9.1.](#) Normative References [11](#)
- [9.2.](#) Informative References [11](#)
- Authors' Addresses [12](#)

1. Introduction

All networks are susceptible to traffic drops due to a number of factors. Traffic drops can go unnoticed unless they are service impacting. In a multi-layered network architecture, it is tedious manual work to localize and root cause traffic blackholing issues. Transient drops are even harder to detect. Existing methodologies that rely on periodically monitoring interfaces on several hosts in a network does not guarantee timely detection, and are not scalable for large networks.

In order to eliminate this tedious monitoring work-flow, objective is to simplify localization and build correlation of dropped packets to particular entity. The network entity shall identify the dropped packets by monitoring dropped counters or doing a deep packet

inspection of the packet discarded by the forwarding ASIC. The implementation of the method used to detect the drop is outside the scope of this document. Dropped packets will be sampled in the forwarding-path and sent to a host or software queue along with type of exception, in/out interface information and other relevant meta data. This will be a push model where the node encountering the error will emit the information about dropped packets and associated meta-data. Techniques for IP Packet Selection [[RFC5475](#)] describes Sampling and Filtering techniques for IP packet selection either using Systematic Sampling or Random Sampling.

The IPFIX Protocol Specification [[RFC7011](#)] defines a generic exchange mechanism for collecting flow information. It supports source-triggered export of information via the push model approach. The IPFIX Information Model [[IANA-IPFIX](#)] defines a list of standard Information Elements (IEs) which can be carried by the IPFIX protocol.

This document focuses on telemetry information for dropped packet exceptions, and proposes an extension to IPFIX message format for collecting sampled exceptions. Some of the IPFIX Information Elements (IEs) already exist, some will be defined along with corresponding formats. It is also possible to achieve sampling of the dropped packets by using sampling methods like SFLOW but details of other sampling methods are outside the scope of this document.

[1.1.](#) Terminology

IPFIX-specific terminology (e.g. Information Element, Template, Template Record, Options Template Record, Template Set, Collector, Exporter, Data Record) used in this document is defined in [Section 2 of \[RFC7011\]](#). As in [[RFC7011](#)] these IPFIX-specific terms have the first letter of a word capitalized. This document also makes use of the same terminology and definitions as [Section 2 of \[RFC5470\]](#).

[1.2.](#) 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.](#) Scope

This document specifies the information model used for reporting packet-based forwarding exceptions. [[RFC7011](#)] provides guidance on the choices of the transport protocols used for IPFIX and their

effects. Encoded IPFIX exception packets need to be reliably transported to the collector. The choice of the actual transport protocol is beyond the scope of this document.

This document assumes that all devices reporting exceptions will use existing IPFIX framework/module to send encoded packets to the collector. This would mean that the network device will specify the template that it is going to use for each of the events. The templates can be of varying length, and there could be multiple templates that a network device could use to encode the exceptions.

The implementation details of the collector application are beyond the scope of this document.

3. Information Elements

The Exception template could contain a subset of the IEs shown in Table 1, depending upon the exception reported.

Whenever packet drop happens inside forwarding plane, following information is key to understanding the issue: reason for packet drop, flow which encountered the drop (packet content), additional meta-data e.g. flow direction (ingress/egress), nexthop index, input interface, output interface, etc. on which this packet was flowing.

The following table includes all the existing IEs that a device reporting IPFIX Exceptions using Exception Template would typically need. The formats of IEs and IPFIX IDs are listed in the table below.

Field Name	Size (bits)	IANA IPFIX ID	Description
flowDirection	8	61	The direction of the Flow observed at Observation point.
ingressInterface	32	10	Index of IP interface where packets of this flow are being received.
egressInterface	32	14	Index of IP interface where packets of this flow are being sent.
dataLinkFrameSize	16	312	Specified length of data link frame.
dataLinkFrameSection	65535	315	Carries n octets from data link frame of selected frame.
commonPropertiesID	64	137	Identifier of a set of common properties that is unique per observation domain.

Table 1: Forwarding Exception Information Elements

Information Elements

4. New Information Elements

4.1. Proposed New Information Elements

The proposed new IEs that a device reporting Exceptions using Exception template would need are listed in Table 2 below.

Field Name	Abstract Data Type	Description
forwardingExceptionCode	unsigned32	Unique code for every exception
forwardingNextHopID	unsigned64	Forwarding NH - index associated with packet that encountered this exception

Table 2: New Information Elements

New Information Elements

The Information Elements defined in Figure 1 are proposed to be incorporated into the IANA IPFIX Information Elements registry [[IANA-IPFIX](#)]

4.2. Definition of Exceptions

Every network will encounter issues like packet loss, from time to time. Some of the causes for such a loss of traffic or a block in transmission of data packets include overloaded system conditions, misconfiguration, profiles and policies that restrict the bandwidth or priority of traffic, network outages, or disruption with physical cable faults. Packet loss could also happen because of incorrect stitching of the forwarding path or a mismatch between control plane and data plane state. Exception code entails the reason/error code due to which this packet has been dropped.

forwardingExceptionCode will be defined in "IPFIX Information Elements" registry. This list can be expanded in the future as necessary. The data record will have corresponding exception code value to indicate forwarding error that caused the traffic drop.

An implementation may choose to encode device internal exception codes as forwardingExceptionCode. In such scenarios, Enterprise Bit

MUST be set to 1 and corresponding Enterprise Number MUST be present as described in [[RFC7011](#)]

A list of commonly used forwarding Exception codes will be identified and listed as part of Table 3 below.

Forwarding Exception Code	Reason
1	FIREWALL_DISCARD
2	TTL_EXPIRY
3	DISCARD_ROUTE
4	BAD_IPV4_CHECKSUM
5	REJECT_ROUTE
6	BAD_IPV4_HEADER (Version incorrect or IHL < 5)
7	BAD_IPV6_HEADER (Version incorrect)
8	BAD_IPV4_HEADER_LENGTH (V4 frame is too short)
9	BAD_IPV6_HEADER_LENGTH
10	BAD_IPV6_OPTIONS_PACKET(too many option headers)
..	..

Table 3: Exception Codes

Reachability to any given destination inside the router is defined using a next-hop which is typically represented in the forwarding path as an index. The nexthop index uniquely identifies the egress path a packet would take to reach the destination. This could include information about the outgoing interface, forwarding features configured for the packet path etc.

An implementation may choose to report linecard and forwarding ASIC information on which an exception occurs, but mechanism to export these fields is out of the scope of this document.

5. Exception Templates

This section presents a list of templates for reporting exceptions using newly proposed IEs in addition to few existing IEs.

5.1. IPFIX Exception Template 1 for Forwarding Exceptions

Exception Template defined in Figure 1 may be used to export forwarding Exceptions.


```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|      Set ID = 2      |      Length = N octets      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|      Template ID = 256      |      Field Count = N      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0| forwardingExceptionCode      |      Field Length = 4      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0| forwardingNextHopId      |      Field Length = 8      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|      flowDirection      |      Field Length = 1      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|      ingressInterface      |      Field Length = 4      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|      egressInterface      |      Field Length = 4      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|      dataLinkFrameSize      |      Field Length = 2      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|      dataLinkFrameSection      |      Field Length = 65535      |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Padding (opt)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

IPFIX Exception Template for Forwarding Exceptions

5.2. IPFIX Exception Template 2 for Forwarding Exceptions

Alternatively, Exception Template defined in Figure 2 may be used. This includes Information Element 137 to represent following fields: forwardingNextHopId, ingressInterface, underlyingIngressInterface and egressInterface.


```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|          Set ID = 2          |          Length = N octets          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|          Template ID = 256          |          Field Count = N          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0| forwardingExceptionCode          |          Field Length = 4          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          flowDirection          |          Field Length = 1          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          commonPropertiesId1          |          Field Length = 8          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          commonPropertiesId2          |          Field Length = 8          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          commonPropertiesId3          |          Field Length = 8          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          commonPropertiesId4          |          Field Length = 8          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          dataLinkFrameSize          |          Field Length = 2          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|0|          dataLinkFrameSection          |          Field Length = 65535          |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Padding (opt)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

IPFIX Exception Template 2 for Forwarding Exceptions

6. IANA Considerations

6.1. Information Elements

IANA manages the IPFIX Information Elements registry at [IANA-IPFIX]. This document introduces two new IPFIX Information Elements.

Name: forwardingExceptionCode
ElementID: TBD
Description: Exception code is an identifier uniquely describing cause of irregularity or traffic drop on a device.
Abstract Data Type: unsigned32
Data Type Semantics: identifier

Name: forwardingNextHopId
ElementID: TBD
Description: NextHop ID is a unique identifier for a NextHop on a device.
Abstract Data Type: unsigned64
Data Type Semantics: identifier

6.2. Forwarding Exception Codes

This document requests addition of a new registry for Forwarding Exception Codes.

Forwarding Exception Code	Reason
1	FIREWALL_DISCARD
2	TTL_EXPIRY
3	DISCARD_ROUTE
4	BAD_IPV4_CHECKSUM
5	REJECT_ROUTE
6	BAD_IPV4_HEADER (Version incorrect or IHL < 5)
7	BAD_IPV6_HEADER (Version incorrect)
8	BAD_IPV4_HEADER_LENGTH (V4 frame is too short)
9	BAD_IPV6_HEADER_LENGTH
10	BAD_IPV6_OPTIONS_PACKET(too many option headers)
..	..

Table 3: Exception Codes

All assignments in this registry are to be performed via Expert Review.

7. Security Considerations

Security Considerations listed in detail for IPFIX in [RFC7011] apply to this document as well. As described in [RFC7011], the IPFIX messages exchanged between network device and collector MUST be protected to provide confidentiality, integrity, and authenticity. Without those characteristics, the messages are subject to various kinds of attacks. These attacks are described in great detail in [RFC7011].

8. Contributors

Manikandan Musuvathi Poornachary
Juniper Networks, Inc.
Electra Exora Business Park~Marathahalli-Sarjapur Outer Ring Road,
Bangalore, KA - 560103
India
Email: mpoornachary@juniper.net

Vishnu Pavan Beeram
Juniper Networks, Inc.
1133 Innovation Way
Sunnyvale, CA 94089
USA
Email: vbeeram@juniper.net

Raveendra Torvi
Juniper Networks, Inc.
10 Technology Park Dr
Westford, MA 01886
USA
Email: rtorvi@juniper.net

9. References

9.1. 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>>.
- [RFC5475] Zseby, T., Molina, M., Duffield, N., Niccolini, S., and F. Raspall, "Sampling and Filtering Techniques for IP Packet Selection", [RFC 5475](#), DOI 10.17487/RFC5475, March 2009, <<https://www.rfc-editor.org/info/rfc5475>>.
- [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>>.
- [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>>.

9.2. Informative References

- [IANA-IPFIX] IANA, "IP Flow Information Export (IPFIX) Entities", <<https://www.iana.org/assignments/ipfix>>.
- [RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", [RFC 5470](#), DOI 10.17487/RFC5470, March 2009, <<https://www.rfc-editor.org/info/rfc5470>>.

Authors' Addresses

Venkata Naga Chaitanya Munukutla
Juniper Networks, Inc.
10 Technology Park Dr
Westford, MA 01886
USA

Email: vmunuku@juniper.net

Shivam Vaid
Juniper Networks, Inc.
Electra, Exora Business Park- Marathahalli-Sarjapur Outer Ring Road
Bangalore, Karnataka 560103
India

Email: shivamv@juniper.net

Aditya Mahale
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
USA

Email: amahale@google.com

Devang Patel
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
USA

Email: pateldevang@google.com

