OPSAWG
Internet-Draft
Intended status: Informational
Expires: 4 September 2025

J. Evans
O. Pylypenko
K. Cheaito
Amazon

3 March 2025

Information Element for Flow Discard Classification
draft-evans-opsawg-ipfix-discard-class-ie-00

Abstract

This document defines a new IPFIX Information Element for classifying flow-level discards which aligns with the information model defined in [I-D.ietf-opsawg-discardmodel]. The flowDiscardClass Information Element provides consistent classification of packet discards across IPFIX implementations, enabling correlation between device and interface-level statistics and impacted flows.

Table of Contents

1. Introduction

For network operators, understanding both where and why packet loss occurs within a network is essential for effective operation. While certain types of packet loss, such as policy-based discards, are intentional and part of normal network operation, unintended packet loss can impact customer services. To automate network operations, operators must be able to detect customer-impacting packet loss, determine its root cause, and apply appropriate mitigation actions.

[I-D.ietf-opsawg-discardmodel] addresses this need by defining an information model that provides precise classification of packet loss, enabling accurate automated mitigation. While its YANG data model implementation provides device and interface-level statistics, effective automated triage often requires understanding which specific flows are impacted. For example, when mitigating congestion, operators may need to identify and trace the sources of elephant flows. This requires the ability to correlate device and interface-level discard classes with the specific flows being dropped.

Currently, [RFC7270] defines the forwardingStatus Information Element for reporting packet forwarding outcomes in IPFIX, including various reasons for packet drops. The defined drop reason codes lack the granularity and clarity needed for automated root cause analysis and impact mitigation, however. For instance, the "For us" reason code provides insufficient context to determine appropriate mitigation actions.

This document addresses these limitations by introducing a new Information Element, flowDiscardClass, to provide a consistent classification scheme for packet discards across IPFIX implementations. This new element aligns with the classification scheme defined in [I-D.ietf-opsawg-discardmodel] and enables:

1. Precise detection of unintended packet loss through clear distinction between intended and unintended discards

2. Accurate root cause analysis through detailed classification of discard reasons

3. Automated selection of mitigation actions based on discard type, rate, and duration

4. Consistent reporting across vendor implementations in both YANG and IPFIX data models

By providing this mapping between YANG and IPFIX implementations, this document enables operators to correlate device-level statistics with flow-level impacts, facilitating more effective automated network operations.

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.

A packet discard accounts for any instance where a packet is dropped by a device, regardless of whether the discard was intentional or unintentional.

Intended discards are packets dropped due to deliberate network policies or configurations designed to enforce security or quality of service. For example, packets dropped because they match an Access Control List (ACL) denying certain traffic types.

Unintended discards are packets that were dropped, which the network operator otherwise intended to deliver, i.e. which indicates an error state. There are many possible reasons for unintended packet loss, including: erroring links may corrupt packets in transit; incorrect routing tables may result in packets being dropped because they do not match a valid route; configuration errors may result in a valid packet incorrectly matching an ACL and being dropped.

3. Information Element

This Information Element has been specified in accordance with the guidelines in [RFC7013].

3.1. Design Rationale

The mapping between [I-D.ietf-opsawg-discardmodel] leaf nodes and IPFIX flowDiscardClass Information Element follows these principles to maintain consistency with the YANG model while leveraging existing IPFIX capabilities and minimise duplication of information:

1. The flowDiscardClass Information Element is specifically for reporting flow-level discard reasons, and therefore only represents the flow subtree from [I-D.ietf-opsawg-discardmodel]. The component is implicitly ’flow’ and the type is implicitly ’discards’, while other components (such as interface, device, and control-plane) are out of scope for this Information Element.

2. Leaf nodes that represent specific discard reasons are assigned unique sequential values to enable precise classification of drops.

3. While some information is also available through other IPFIX Information Elements, the flowDiscardClass maintains structural elements from the information model (such as layer) where needed to preserve the hierarchical classification.

4. Leaf nodes that can be represented by existing IPFIX Information Elements are not assigned reason codes to avoid redundancy. Specifically:

   a. Direction (ingress/egress) is handled by the flowDirection Information Element (IE 61)

   b. IP version is handled by the ipVersion Information Element (IE 60)

c. Unicast versus multicast classification is handled by examining the source and destination addresses (sourceIPv4Address (IE 8), destinationIPv4Address (IE 12), sourceIPv6Address (IE 27), destinationIPv6Address (IE 28))

d. QoS class information is handled by the ipDiffServCodePoint Information Element (IE 195)

5. Parent nodes in the YANG tree are assigned reason codes to enable both coarse and fine-grained reporting. For example:

a. errors/ (0) represents all error discards

b. errors/l3/rx/ (9) represents all L3 receive error discards

c. errors/l3/rx/checksum-error (10) represents specific L3 checksum error discards

While this draft takes the approach of leveraging existing IPFIX Information Elements where possible to avoid redundancy, an alternative approach would be to implement all leaves under the flow/discards branch from [I-D.ietf-opsawg-discardmodel] as distinct flowDiscardClass values. This would result in additional values for direction (ingress/egress), address family (IPv4/IPv6), cast type (unicast/multicast), and QoS class. This approach would provide a more complete mapping with the YANG model without dependencies, however, it would duplicate information already available through existing Information Elements.

3.2. flowDiscardClass Definition

Name: flowDiscardClass

Description: Classifies the reason a packet was discarded in a flow, using the hierarchical classification scheme defined in [I-D.ietf-opsawg-discardmodel].

Abstract Data Type: unsigned8

Data Type Semantics: identifier

References: [I-D.ietf-opsawg-discardmodel]

ElementId: TBD

Status: current

3.3. flowDiscardClass Values

Table 1 defines the values for the flowDiscardClass Information Element mapped from the corresponding [I-D.ietf-opsawg-discardmodel] Discard Class:

Discard Class	flowDiscardClass Value
errors/	0
errors/internal/	1
errors/internal/parity-error	2
errors/l2/rx/	3
errors/l2/rx/crc-error	4
errors/l2/rx/invalid-mac	5
errors/l2/rx/invalid-vlan	6
errors/l2/rx/invalid-frame	7
errors/l2/tx	8
errors/l3/rx/	9
errors/l3/rx/checksum-error	10
errors/l3/rx/mtu-exceeded	11
errors/l3/rx/invalid-packet	12
errors/l3/ttl-expired	13
errors/l3/no-route	14
errors/l3/invalid-sid	15
errors/l3/invalid-label	16
errors/l3/tx	17
policy/	18
policy/l2/acl	19

policy/l3/acl	20
policy/l3/policer	21
policy/l3/null-route	22
policy/l3/rpf	23
policy/l3/ddos	24
no-buffer/class	25

Table 1: Flow discard classification values and corresponding discard classes

3.4. Usage with Existing Information Elements

When reporting flow-level discard statistics, the flowDiscardClass Information Element SHOULD be used in conjunction with the following existing Information Elements as defined in [RFC7012]:

YANG Path	IPFIX Information Element
flow/direction	flowDirection (IE 61)
.../address-family	ipVersion (IE 60)
.../unicast	sourceIPv4Address (IE 8), destinationIPv4Address (IE 12), sourceIPv6Address (IE 27), destinationIPv6Address (IE 28)
.../multicast	sourceIPv4Address (IE 8), destinationIPv4Address (IE 12), sourceIPv6Address (IE 27), destinationIPv6Address (IE 28)
.../qos/class	ipDiffServCodePoint (IE 195)

Table 2: Mapping between YANG model paths and IPFIX fields

4. Security Considerations

This document defines a new Information Element for flow-level discard classification to align with the information model defined in [I-D.ietf-opsawg-discardmodel]. As such, there are no security issues related to this document, which are additional to those discussed in [RFC7011], [RFC7012].

5. IANA Considerations

This document requests IANA to register the flowDiscardClass Information Element in the IANA IPFIX Information Elements registry.

6. References

6.1. 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/rfc/rfc2119>.

[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/rfc/rfc8174>.

6.2. Informative References

[I-D.ietf-opsawg-discardmodel] Evans, J., Pylypenko, O., Haas, J., Kadosh, A., and M. Boucadair, "An Information Model for Packet Discard Reporting", Work in Progress, Internet-Draft, draft-ietf-opsawg-discardmodel-04, 21 October 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-discardmodel-04>.

[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/rfc/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/rfc/rfc7012>.

[RFC7013] Trammell, B. and B. Claise, "Guidelines for Authors and Reviewers of IP Flow Information Export (IPFIX) Information Elements", BCP 184, RFC 7013, DOI 10.17487/RFC7013, September 2013, <https://www.rfc-editor.org/rfc/rfc7013>.

[RFC7270] Yourtchenko, A., Aitken, P., and B. Claise, "Cisco-Specific Information Elements Reused in IP Flow Information Export (IPFIX)", RFC 7270, DOI 10.17487/RFC7270, June 2014, <https://www.rfc-editor.org/rfc/rfc7270>.

Authors' Addresses

John Evans
Amazon
1 Principal Place, Worship Street
London
EC2A 2FA
United Kingdom
Email: jevanamz@amazon.co.uk

Oleksandr Pylypenko
Amazon
410 Terry Ave N
Seattle, WA 98109
United States of America
Email: opyl@amazon.com

Karim Cheaito
Amazon
410 Terry Ave N
Seattle, WA 98109
United States of America
Email: kcheaito@amazon.com