Internet Engineering Task Force (IETF)
Request for Comments: 6526
Category: Standards Track
ISSN: 2070-1721

B. Claise
P. Aitken
A. Johnson
Cisco Systems, Inc.
G. Muenz
TU Muenchen

March 2012

IP Flow Information Export (IPFIX) Per Stream Control Transmission Protocol (SCTP) Stream

Abstract

This document specifies an extension to the specifications in RFC 5101, IP Flow Information Export (IPFIX), when using the Partial Reliability extension of SCTP (PR-SCTP, Partial Reliability Stream Control Transmission Protocol).

When implemented at both the Exporting Process and Collecting Process, this method offers several advantages, such as the ability to calculate Data Record losses for PR-SCTP per Template, immediate export of Template Withdrawal Messages, immediate reuse of Template IDs within an SCTP stream, reduced likelihood of Data Record loss, and reduced demands on the Collecting Process. When implemented in only the Collecting Process or Exporting Process, then normal IPFIX behavior will be seen without all of the additional benefits.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6526.

Table of Contents

1. Introduction

The IPFIX protocol [RFC5101] has the goal of exporting Flow information. This protocol is designed to export information about IP traffic Flows and related measurement data, where a Flow is defined by a set of key attributes (e.g., source and destination IP address, source and destination port, etc.). However, thanks to its Template mechanism, the IPFIX protocol can export any type of

information, as long as the relevant Information Element is specified in the IPFIX information model [RFC5102], registered with IANA [IANA], or specified as an enterprise-specific Information Element.

The IPFIX protocol [RFC5101] specifies that traffic measurements for Flows are exported using a TLV (Type, Length, Value) format. The information is exported using a Template Record, which is sent once to export the {Type, Length} pairs that define the data format for the Information Elements in a Flow. The Data Records specify values for each Flow.

The IPFIX protocol [RFC5101] is flexible: It foresees the usage of multiple SCTP streams per association; it allows the transmission of Data Sets, Template Sets, and/or Options Template Sets on any SCTP stream; it offers full and partially reliable export of Data Sets; it specifies both ordered and out-of-order delivery of Data Sets. However, due to bandwidth restrictions and packet losses in the network as well as resource constraints on the Exporter and Collector (e.g., limited buffer sizes), it is not always possible to export all Data Sets in a reliable way.

This document specifies a method for exporting a Template Record and its associated Data Sets in a single SCTP stream, limiting each Template ID to a single SCTP stream if possible, and imposing in-order transmission.

This method offers several advantages over IPFIX export as specified in [RFC5101], such as the ability to calculate Data Record losses for PR-SCTP per Template, immediate export of Template Withdrawal Messages, immediate reuse of Template IDs within an SCTP stream, reduced likelihood of Data Record loss, and reduced demands on the Collecting Process.

1.1. Relationship with IPFIX and PSAMP

The specifications in this document apply to the IPFIX protocol specifications [RFC5101]. However, they only apply to the SCTP transport protocol [RFC4960] option of the IPFIX protocol specifications (see Section 10 of [RFC5101]), specifically if the Partial Reliability extension [RFC3758] is used. All specifications from [RFC5101] apply, unless specified otherwise in this document.

As the Packet Sampling (PSAMP) protocol specifications [RFC5476] are based on the IPFIX protocol specifications, the specifications in this document are also valid for the PSAMP protocol.

1.2. Applicability

The specifications contained in this document are applicable to cases where application requirements include knowing how many Data Records of a certain type (i.e., from a certain Template) were lost. A typical example is a router exporting billing records, where the Exporting Process cannot afford to export all the Flow Records reliably, due to limited resources to buffer a large number of Flow Records. Such a situation may occur if Data Sets are generated at a higher rate at the Exporter than can be transferred to the Collector because of bandwidth limitations in the network or slow reception at the Collector.

To be more precise, the specification applicability is the case where multiple Templates are simultaneously active within a single SCTP Transport Session and the calculation of the Data Record loss for a particular Template is required. Indeed, with the current IPFIX specifications [RFC5101], if an IPFIX Message is lost (UDP or SCTP partially reliable), it is not possible to determine to which Template(s) the lost Data Records belong.

Exporting Processes following the specifications in this document will interoperate with existing Collecting Processes that comply with [RFC5101]; no changes are required at the Collecting Process to receive data from an Exporting Process compliant with this method. However, Collecting Processes may implement additional support for per-stream export specified in this document in order to realize all the benefits of the approach specified herein. Since the specifications in this document mandate in-order transmission of (Options) Templates and associated Data Records, late arrival of (Options) Templates at the Collecting Process is avoided, which means that there are no Data Records that need to be dropped or buffered.

1.3. Limitations

When multiple Templates are required, this method requires multiple SCTP streams in the association between the Exporting Process and Collecting Process, ideally one stream per Template. To properly handle the transmission of additional Templates during the Transport Session, additional SCTP streams are sometimes required. These SCTP streams can only be added within the existing SCTP association if the specifications in [RFC6525] are supported.

2. Terminology

IPFIX-specific terminology used in this document is defined in Section 2 of [RFC5101]. As in [RFC5101], these IPFIX-specific terms have the first letter of a word capitalized when used in this document.

Note that, in this document, "(Options) Template" is used to refer to Templates and Options Templates. Unless otherwise specified, "Template" alone refers to Templates exclusive of Options Templates.

Template Reuse Delay

The time the Exporting Process needs to wait after sending the last Data Set described by a given Template before sending a Template Withdrawal Message for the Template. A suitable default value is 5 seconds, as specified in [RFC5101].

2.1. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2.2. IPFIX Documents Overview

The IPFIX protocol [RFC5101] provides network administrators with access to Flow information.

The architecture for the export of measured Flow information out of an IPFIX Exporting Process to a Collecting Process is defined in the IPFIX architecture [RFC5470], per the requirements defined in [RFC3917].

The IPFIX architecture [RFC5470] specifies how IPFIX Data Records and Templates are carried via a congestion-aware transport protocol from IPFIX Exporting Processes to IPFIX Collecting Processes.

IPFIX has a formal description of IPFIX Information Elements, their names, their types, and additional semantic information, as specified in the IPFIX information model [RFC5102].

Finally, the IPFIX applicability statement [RFC5472] describes what types of applications can use the IPFIX protocol and how they can use the information provided. Furthermore, it shows how the IPFIX framework relates to other architectures and frameworks.

2.3. PSAMP Documents Overview

The document "A Framework for Packet Selection and Reporting" [RFC5474] describes the Packet Sampling (PSAMP) framework for network elements to select subsets of packets by statistical and other methods, and to export a stream of reports on the selected packets to a Collector.

The set of packet selection techniques (sampling, filtering, and hashing) supported by PSAMP are described in "Sampling and Filtering Techniques for IP Packet Selection" [RFC5475].

The PSAMP protocol [RFC5476] specifies the export of packet information from a PSAMP Exporting Process to a PSAMP Collecting Process. Like IPFIX, PSAMP has a formal description of its Information Elements, their names, their types, and additional semantic information. The PSAMP information model is defined in [RFC5477].

3. IPFIX Protocol Specifications: Limitations and Improvements

For three specific topics ("Data Record Loss Calculated Per Template", "Immediate Template Withdrawal and Reuse", and "Requirement for Data Set Buffering"), this section explains the limitations of the IPFIX protocol specifications on the one hand, and the advantages of the method specified in this document on the other.

3.1. Data Record Loss Calculated Per Template

3.1.1. IPFIX Protocol Specifications: Limitation

Section 6.3.2 of [RFC3917], "Requirements for IP Flow Information Export" discusses the data transfer reliability issues:

> Loss of flow records during the data transfer from the Exporting Process to the Collecting Process must be indicated at the Collecting Process.

However, in some cases, it may be important to know how many Data Records of a certain type were lost (e.g., in the case of billing), and IPFIX does not conventionally provide this information.

A Collecting Process can detect out-of-sequence, dropped, or duplicate IPFIX Messages by tracking the Sequence Number [RFC5101]. Note that the Sequence Number field in the IPFIX Message header increases with the number of IPFIX Data Records within the SCTP stream, so loss will be detected per stream.

The IPFIX protocol specifications [RFC5101] specify that Data Records defined by any Template may be sent on any SCTP stream. As such, if there is more than one Template defined within the whole SCTP association, then there is no way of knowing with which Template any lost Data Record is associated. This is true, no matter what convention the Exporting Process uses to send Data Records on different SCTP streams, as the protocol makes no guarantees.

Note that a workaround allowed by the IPFIX specifications in [RFC5101] is to use only one Template Record per SCTP Transport Session, at the cost of multiplying the number of SCTP Transport Sessions when multiple Template Records are required.

3.1.2. IPFIX Export Per SCTP Stream: Advantage

Using the specifications in this document, it is guaranteed that any lost Data Records will be associated only with the Templates that are defined on that SCTP stream. By defining only one Template per SCTP stream, it is ensured that any loss is associated with that single Template. So, by exporting each Template and its corresponding Data Records in a separate SCTP stream from other Templates and Data Records, the loss pertaining to each specific Template can be deduced from the Sequence Number field in the IPFIX Message headers.

3.2. Immediate Template Withdrawal and Reuse

3.2.1. IPFIX Protocol Specifications: Limitation

A Collecting Process must have received the Template Record associated with the Data Records to be able to decode the information in the Data Records. [RFC5101] specifies the following:

> The Exporting Process SHOULD transmit the Template Set and Options Template Set in advance of any Data Sets that use that (Options) Template ID, to help ensure that the Collecting Process has the Template Record before receiving the first Data Record.

The fact that the Collecting Process cannot decode the Data Records without the corresponding Template Record may result in Data Records being discarded by the Collecting Process, as specified in [RFC5101]:

> The Collecting Process normally receives Template Records from the Exporting Process before receiving Data Records. The Data Records are then decoded and stored by the Collector. If the Template Records have not been received at the time Data Records are received, the Collecting Process MAY store the Data Records for a short period of time and decode them after the Template Records are received.

3.2.2. IPFIX Export Per SCTP Stream: Advantages

By exporting each Template Record and the corresponding Data Records within a single SCTP stream and imposing in-order transmission, the Template Record will always arrive before the associated Data Records. Therefore, there is no risk that the Collecting Process discards Data Records while waiting for the Template Record to arrive.

Furthermore, when reusing a Template ID within an SCTP stream, the Template Withdrawal Message will be guaranteed to arrive before the new definition of the Template, and therefore the Template Record may be sent directly after the Template Withdrawal Message. In other words, the Template Reuse Delay restriction (5 seconds by default, as specified in [RFC5101]) does not need to be applied to Template ID reuse within the same SCTP stream.

Another advantage of the new specifications in this document is a reduced load on the Collecting Process. Indeed, the Collecting Process doesn't have to store the Data Records while waiting for the Template Record, as the transmission order is always guaranteed. This way, extra reliability of the Data Records is achieved without extra burden on the Collecting Process.

3.3. Requirement for Data Set Buffering

3.3.1. IPFIX Protocol Specifications: Limitation

The fact that the protocol specifications in [RFC5101] are flexible in terms of SCTP stream(s) on which the Template Set, Options Template Set, and corresponding Data Sets are exported implies that the (Options) Template Record might be exported on a different SCTP stream than the corresponding Data Records. This might cause Data Record loss in the Collecting Process, as ordered transmission across SCTP streams is not guaranteed.

For example, a Template Record may be blocked pending reliable transmission on one SCTP stream while the corresponding Data Records may be transmitted immediately in another SCTP stream. Also, due to different levels of SCTP stream congestion, it is possible that even if the Template Record and corresponding Data Records are sent reliably, Data Records sent on a different SCTP stream than the Template Record might still arrive before the Template Record.

3.3.2. IPFIX Export Per SCTP Stream: Advantages

By exporting each Template Record and all corresponding Data Records within a single SCTP stream, and imposing in-order transmission, the issue of ordered transmission across multiple SCTP streams is avoided.

By exporting all corresponding Data Records within the same ordered SCTP stream as the Template Record, each SCTP stream is independent and self-contained, and the interaction between SCTP streams is limited to that of the Options Template and associated Data Records sent in different streams. This has several advantageous consequences, including order preservation that does not result in the blocking of unrelated data, and load reduction on the Collecting Process (as the Template Records are guaranteed to be delivered before the associated Data Records, there is no need for the buffering of Data Sets that correspond with Templates that are missing).

4. Specifications

This section specifies Exporting Process and Collecting Process behavior different from that in [RFC5101] in order to realize the benefits of per-stream export. Note that Exporting Processes following these specifications will interoperate with [RFC5101]-compliant Collecting Processes, but that Collecting Processes will have to follow additional non-interoperable specifications to realize the full benefits of the technique. These new specifications, which add to those in [RFC5101], are described with the key words defined in [RFC2119].

4.1. New Information Element

dataRecordsReliability

Description:
The export reliability of Data Records, within this SCTP stream, for the element(s) in the Options Template scope. A typical example of an element for which the export reliability will be reported is the Template ID, as specified in the Data Records Reliability Options Template. A value of 'True' means that the Exporting Process MUST send any Data Records associated with the element(s) reliably within this SCTP stream. A value of 'False' means that the Exporting Process MAY send any Data Records associated with the element(s) unreliably within this SCTP stream.

```
Abstract Data Type: boolean
Data Type Semantics: identifier
ElementId: 276
Status: current
```

Per Section 6.1.5 of [RFC5101], the boolean data type is encoded as a single octet, with the value of 1 for True and the value of 2 for False.

4.2. Template Management

To take advantage of per-stream export, Exporting Processes MUST follow the specification in this section in addition to Section 8, "Template Management", of [RFC5101].

As specified in [RFC5101], Template Sets and Options Template Sets MUST be sent reliably.

Any Data Sets associated with a Template Record MUST be sent on the same SCTP stream on which the Template Record was sent.

The Data Records Reliability Options Template is used to explicitly inform the Collecting Process which Templates will be used in each SCTP stream and whether each set of associated Data Records will be sent reliably or unreliably. After defining a Template ID and before sending any associated Data Records on an SCTP stream, the Exporting Process MUST notify the Collecting Process of its intention to send those Data Records reliably or unreliably within that SCTP stream. It does this by sending a Data Record defined by the Data Records Reliability Options Template for the Template associated with the Data Records to be sent. If it does not, then the Collecting Process MUST disable this extension for the SCTP association. The one exception to this rule is that the Data Records associated with the Data Records Reliability Options Template don’t require an explicit notification, as these MUST always be sent reliably.

The Data Records Reliability Options Template MUST contain the following Information Elements:

```
Scope:     Template ID
Non-scope: dataRecordsReliability
```

After sending a value of ’True’ for the dataRecordsReliability Element, the Exporting Process MUST send any Data Records associated with the currently defined Template ID reliably within this SCTP stream. After sending a value of ’False’ for the

dataRecordsReliability Element, the Exporting Process MAY send any Data Records associated with the Template ID unreliably within this SCTP stream.

If the Exporting Process wants to change the Data Records Reliability value (from reliable to unreliable, or vice versa) for Data Records on an SCTP stream, the Template MUST be withdrawn, and a new Template MUST be used.

The Data Records Reliability Options Template MAY contain other non-scope Information Elements associated with the (Options) Template.

When an Options Template (including the Data Records Reliability Options Template) and associated Data Records are sent in the same SCTP stream, the first associated Data Record can follow the Options Template immediately. When the Options Template and associated Data Records are sent in different SCTP streams, the Exporting Process SHOULD transmit the Options Template in advance of any Data Sets that use it, to help ensure that the Collector has received the Options Template Record before receiving the first associated Data Record.

It is RECOMMENDED that the Exporting Process only sends a single Template and corresponding Data Sets within a single SCTP stream in order to enable calculation of the potential Data Record loss for this Template. The Exporting Process MAY group related (Options) Templates and their associated Data Records within a single SCTP stream so that loss statistics are calculated for the group of Templates that are being sent unreliably within the SCTP stream. This is suitable in cases where there are only slight variations among the Templates in a group (e.g., the omission of unavailable fields for export efficiency) and may be necessary if the SCTP association does not support enough SCTP streams to export each Template in its own SCTP stream.

If an SCTP stream contains a mixture of Data Records defined by Template Records and by Options Template Records, the Data Records defined by the Options Template Records SHOULD be sent reliably so that the Collecting Process does not consider any loss to be associated with the Options Data Records.

4.3. SCTP

To take advantage of per-stream export, Exporting Processes MUST manage SCTP streams according to the specification in this section, in addition to Section 10.2.4.3, "Stream", of [RFC5101].

PR-SCTP [RFC3758] MUST be implemented by all compliant implementations.

All IPFIX Messages in an SCTP stream MUST be sent in order.

As specified in [RFC5101], depending on the requirements of the application, the Exporting Process may send Data Sets with full or partial reliability.

If the Exporting Process is required to export a new Template Record but there are no more free SCTP streams available, it SHOULD attempt to increase the number of outbound SCTP streams to which it is able to send, per [RFC6525]. Alternatively, the Exporting Process MAY add the Template Set and Data Records to an existing SCTP stream at the cost of diluting the granularity of any Data Record loss attribution. An alternative that may result in the loss of Flow Records (for example, due to lack of buffering on the Exporter) is to restart the SCTP association with an increased number of SCTP streams.

4.4. Template Withdrawal Message

To take advantage of per-stream export, Exporting Processes MUST send Template Withdrawal Messages according to the specification in this section, in addition to Section 8, "Template Management", of [RFC5101].

As specified in [RFC5101], Templates that are no longer in use SHOULD be deleted. Before reusing a Template ID, the Template MUST be deleted. In order to delete an allocated Template, the Template is withdrawn through the use of a Template Withdrawal Message.

The Template Withdrawal Message MUST be sent on the same SCTP stream as the associated Template Record.

The Template Withdrawal Message MUST be sent reliably, using SCTP-ordered delivery per [RFC5101]. As all IPFIX Messages are sent in order within an SCTP stream (per the specifications in this document), the IPFIX Message containing the Template Withdrawal Message will not arrive at the Collecting Process before any associated and previously sent Data Record. As a consequence, no Data Records will be lost due to delayed arrival at the Collecting Process.

The Template ID from a withdrawn Template MAY be reused on the same SCTP stream immediately after the Template Withdrawal Message is sent. This case is equivalent to the use of a Template Reuse Delay value of 0.

After reusing the Template ID, the Exporting Process MUST send a Data Record associated with the Data Records Reliability Options Template to specify the reliability level of the Data Records associated with the new Template.

If the Template ID is to be reused on a different SCTP stream, the new Template Record MUST NOT be sent before the Template Reuse Delay interval.

A Template Withdrawal Message to withdraw all Templates for the Observation Domain ID specified in the IPFIX Message header MUST NOT be used.

Multiple Template IDs MAY be withdrawn with a single Template Withdrawal Message under the condition that all the Template IDs in the Template Withdrawal Message are used on the same SCTP stream as the Template Withdrawal Message.

4.5. The Collecting Process's Side

Collecting Processes must operate in a fashion slightly contrary to [RFC5101] in order to realize the full benefits of per-stream export. However, the specification in this section contains a mechanism that allows per-stream-capable Collecting Processes to selectively enable per-stream export, in order to ensure interoperability of per-stream-capable Collecting Processes with Exporting Processes that do not implement per-stream export.

4.5.1. SCTP

As specified in [RFC5101], the Collecting Process SHOULD listen for a new association request from the Exporting Process. The Exporting Process will request a number of SCTP streams to use for export.

A Collecting Process SHOULD support the procedure for the addition of an SCTP stream specified in [RFC6525].

4.5.2. Enabling the Per-SCTP-Stream Extension

In IPFIX, there is no explicit notification of the Exporting Process's capabilities. There is also no return channel for the Collecting Process to communicate its capabilities.

When the Exporting Process is sending according to the per-SCTP-stream extension, the first Data Record received by the Collecting Process will be associated with the Data Records Reliability Options

Template. In this case, the Collecting Process enables the extension for this Transport Session. Otherwise, the Collecting Process MUST NOT enable the extension for this Transport Session.

The Collecting Process MUST accept other non-scope Information Elements in the Data Records Reliability Options Template.

4.5.3. Disabling the Per-SCTP-Stream Extension

Nothing prevents an implementation that does not meet the specification of the per-SCTP-stream extension from sending a Template that looks like a dataRecordsReliability Options Template. Therefore, a Collecting Process MUST detect if the Exporting Process fails to meet the specification fully. If any of the conditions below is met, the Exporting Process does not properly use the per-SCTP-stream extension, and the Collecting Process MUST log an error message and disable this extension for the SCTP association.

1. A Data Record is received before the appropriate Data Record associated with the Data Records Reliability Options Template has been received on the same SCTP stream (see Section 4.2). Note: Data Records associated with the Data Records Reliability Options Template are an exception to this rule.

2. A Data Record associated with a Data Records Reliability Options Template is received on an SCTP stream for a (non-Options) Template that was defined on a different SCTP stream.

3. A second Data Record associated with the Data Records Reliability Options Template is received for the same (Options) Template.

4. A Data Record or a Template Withdrawal Message is associated with a Template that was defined on a different SCTP stream.

5. Loss of Data Records is detected within a stream where a Data Record associated with the Data Records Reliability Options Template indicating unreliable transmission for any Template has not been received.

6. A message is received with the SCTP U(nordered) flag set to 1 (i.e., the message was sent unordered), even if it is processed in order.

4.5.4. Calculating Data Record Loss Per Template

As specified in [RFC5101], the IPFIX protocol has a Sequence Number field in the IPFIX Message header that increases with the number of IPFIX Data Records in the IPFIX Message. A Collecting Process may detect out-of-sequence, dropped, or duplicate IPFIX Messages by tracking the Sequence Number.

When one or more sequential IPFIX Messages are considered lost, the number of lost Data Records is equal to the Sequence Number of the first IPFIX Message Header following the lost packets (S2) minus the Sequence Number of the first lost IPFIX Message (S1). The Sequence Number of the first lost IPFIX Message can be calculated as the Sequence Number of the last IPFIX Message before the sequence of lost IPFIX Messages (S0) plus the number of Data Records in that IPFIX Message (N0).

```
S1 = S0 + N0
loss = (S2 - S1) (mod(2^32))
     = (S2 - (S0 + N0)) (mod(2^32))
```

Note that modulo 2^32 arithmetic is required, since the Sequence Number may wrap within the series of lost IPFIX Messages. If less than 2^32 Data Records are lost in a sequence (which can be assumed in practice), the above equation returns the exact number of lost Data Records.

Note that using an unsigned32 type for the loss would automatically take care of the mod(2^32) operation.

As this Sequence Number is incremented per SCTP stream, the loss of Data Records sent in that SCTP stream can be calculated in the case of partially reliable export. This loss can be attributed to the Data Records sent for the (Options) Template(s) whose records are being sent unreliably within that SCTP stream.

5. Resource Impact

Although adding the new SCTP streams requires a message exchange, it is more lightweight to set up additional SCTP streams than to set up a new SCTP association, since the only overhead of adding SCTP stream(s) to an existing SCTP association is the addition of 16-24 more bytes (allocated in the SCTP association, a single time), whereas setting up a new SCTP association requires more overhead.

In terms of throughput impact, the fact that these specifications discourage multiplexing Templates and Data Records of different Template IDs may lead to a slightly larger IPFIX Message overhead.

If the Data Record rate is low for a specific Template (and hence a specific SCTP stream), the Exporting Process might not be able to fill the IPFIX Messages with Data Records associated with other Templates. In such a situation, there is a potential overhead due to additional IPFIX Message headers and SCTP chunk headers.

Finally, with respect to the processing overhead on the Exporter, a lot of state information must be stored when a large number of SCTP streams are used within an SCTP association. However, no comparison of the performance impact of multiple streams within an SCTP association versus opening the same number of independent SCTP associations is available.

6. Examples

Figure 1 shows an example where SCTP stream 10 carries a Template Record with Template ID 257 transmitted with full reliability (FR), together with associated Data Records transmitted with partial reliability (PR). The Data Records Reliability Options Template with Template ID 256 is transmitted with full reliability. Its corresponding Data Set contains one Data Record.

Record 1:

- Scope: Template ID = 257
- Non-scope: dataRecordsReliability = False

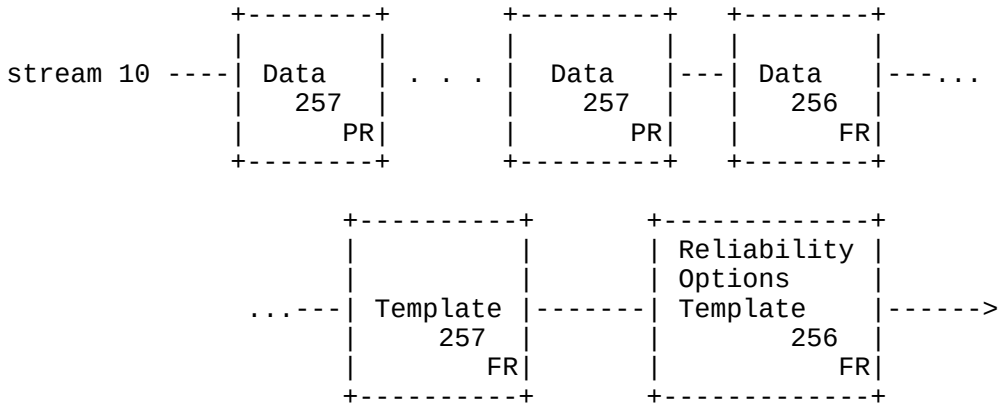


Figure 1

Note that Template 257 will always be processed before the Data Records by the Collecting Process, because all IPFIX Messages are sent in order within an SCTP stream. Therefore, the job of the Collecting Process is simplified. Furthermore, the Data Record loss for Template 257 can easily be calculated by the Collecting Process.

If an Options Template is necessary to understand the content of a Data Record (i.e., the scope in the Options Template Record is an Information Element contained in the Data Record or associated with the Data Record), the Options Template Record should be sent in the same SCTP stream, as displayed in Figure 2.

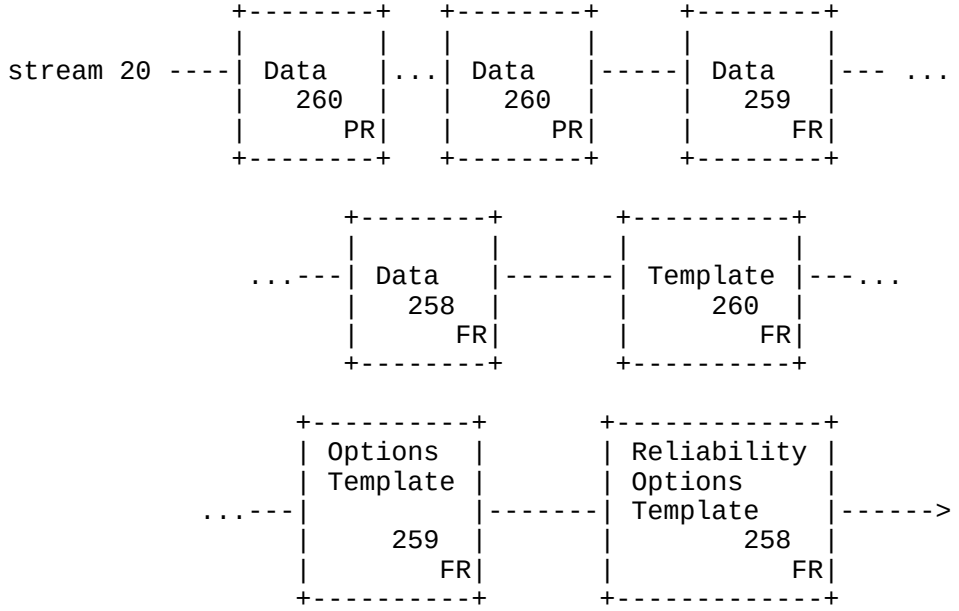


Figure 2

Figure 2 shows an example where SCTP stream 20 carries the following:

- a Data Records Reliability Options Template with Template ID 258, transmitted with full reliability.

- an Options Template Record with Template ID 259, transmitted with full reliability. This Options Template Record contains additional information related to the subsequent Data Records based on Template ID 260. Typical examples are the Common Properties information [RFC5473] or the Selector Report Interpretation [RFC5476].

- a Template Record with Template ID 260, transmitted with full reliability.

- a Data Set specified by the Reliability Options Template with Template ID 258, transmitted with full reliability.

  The Data Set contains three Data Records:

  Record 1:
  - o Scope: Template ID = 258
  - o Non-scope: dataRecordsReliability = True

  Record 2:
  - o Scope: Template ID = 259
  - o Non-scope: dataRecordsReliability = True

  Record 3:
  - o Scope: Template ID = 260
  - o Non-scope: dataRecordsReliability = False

  These Data Records inform the Collecting Process that the Data Records for Template IDs 258 and 259 are sent reliably, while the Data Records for Template ID 260 are not. Note that the first Data Record associated with the Data Record Reliability Options Template (Template ID 258) is not required and can be omitted.

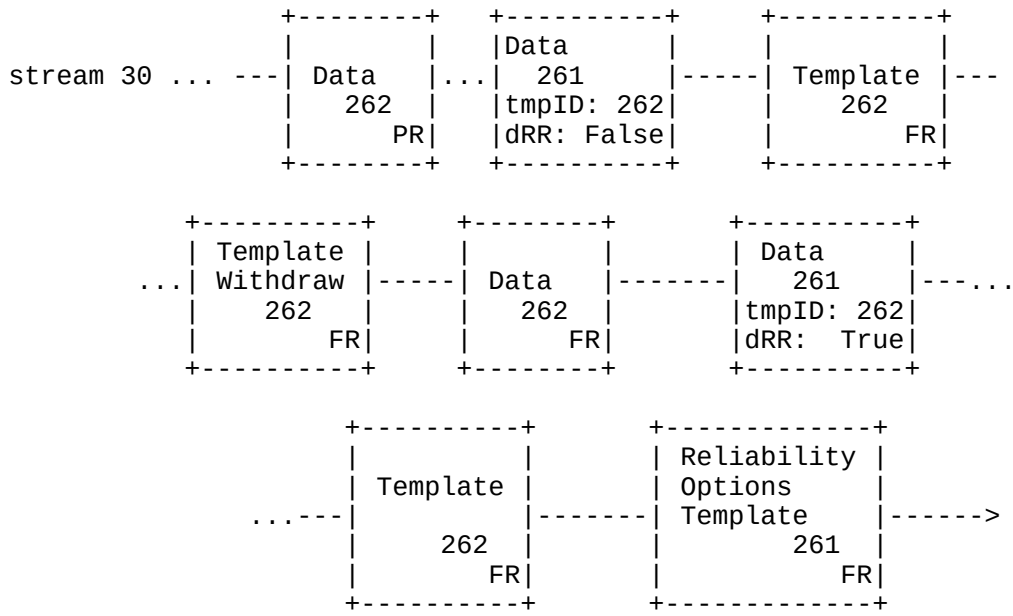
- a Data Record specified by Template ID 259, transmitted with full reliability.

- a Data Record specified by Template ID 260, transmitted with partial reliability.

If the Collecting Process observes some Data Record loss using the Sequence Number, the loss can only stem from the Data Records associated with Template ID 260, as these are the only Data Records not exported reliably. Therefore, the calculation of loss per Template ID 260 is possible.

Note that Options Templates 258, 259, and 260 will always arrive before their associated Data Records, respectively, because all IPFIX Messages must be sent in order within an SCTP stream.

Figure 3 shows an example where SCTP stream 30 carries a Template Record with Template ID 262 transmitted with full reliability, an associated Data Record transmitted with full reliability, and a Template Withdrawal Message, followed by a redefinition of Template ID 262, and finally the Data Record associated with the new Template transmitted with partial reliability. The Template Withdrawal Message and the new definition of Template ID 262 are sent immediately, without waiting for the Template Reuse Delay interval.



dRR: Data Records Reliability

Figure 3

The second Data Record associated with the Data Records Reliability Options Template shows that the Data Records associated with the newly specified Template ID 262 will be sent unreliably.

7. IANA Considerations

According to the process defined in [RFC5102], IANA has allocated the dataRecordsReliability Information Element (defined in Section 4.1) in the "IPFIX Information Elements" registry [IANA].

8. Security Considerations

The same security considerations as for the IPFIX protocol [RFC5101] apply.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3758] Stewart, R., Ramalho, M., Xie, Q., Tuexen, M., and P. Conrad, "Stream Control Transmission Protocol (SCTP) Partial Reliability Extension", RFC 3758, May 2004.

[RFC4960] Stewart, R., Ed., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC5101] Claise, B., Ed., "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", RFC 5101, January 2008.

[RFC5102] Quittek, J., Bryant, S., Claise, B., Aitken, P., and J. Meyer, "Information Model for IP Flow Information Export", RFC 5102, January 2008.

[RFC5475] Zseby, T., Molina, M., Duffield, N., Niccolini, S., and F. Raspall, "Sampling and Filtering Techniques for IP Packet Selection", RFC 5475, March 2009.

[RFC6525] Stewart, R., Tuexen, M., and P. Lei, "Stream Control Transmission Protocol (SCTP) Stream Reconfiguration", RFC 6525, February 2012.

9.2. Informative References

[IANA] IPFIX Information Elements Registry, <http://www.iana.org/assignments/ipfix>.

[RFC3917] Quittek, J., Zseby, T., Claise, B., and S. Zander, "Requirements for IP Flow Information Export (IPFIX)", RFC 3917, October 2004.

[RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", RFC 5470, March 2009.

[RFC5472] Zseby, T., Boschi, E., Brownlee, N., and B. Claise, "IP Flow Information Export (IPFIX) Applicability", RFC 5472, March 2009.

[RFC5473] Boschi, E., Mark, L., and B. Claise, "Reducing Redundancy in IP Flow Information Export (IPFIX) and Packet Sampling (PSAMP) Reports", RFC 5473, March 2009.

[RFC5474] Duffield, N., Ed., Chiou, D., Claise, B., Greenberg, A., Grossglauser, M., and J. Rexford, "A Framework for Packet Selection and Reporting", RFC 5474, March 2009.

[RFC5476] Claise, B., Ed., Johnson, A., and J. Quittek, "Packet Sampling (PSAMP) Protocol Specifications", RFC 5476, March 2009.

[RFC5477] Dietz, T., Claise, B., Aitken, P., Dressler, F., and G. Carle, "Information Model for Packet Sampling Exports", RFC 5477, March 2009.

10. Acknowledgments

The authors would like to thank Brian Trammell for his expert feedback and continuous effort to improve the specifications; Elisa Boschi for her thorough reading; Randall Stewart, Peter Lei, and Michael Tuexen for their SCTP-related feedback and expertise; and Tobias Limmer.

Authors' Addresses

Benoit Claise
Cisco Systems, Inc.
De Kleetlaan 6a b1
Diegem 1813
Belgium

Phone: +32 2 704 5622
EMail: bclaise@cisco.com

Paul Aitken
Cisco Systems, Inc.
96 Commercial Quay
Commercial Street
Edinburgh, EH6 6LX, United Kingdom

Phone: +44 131 561 3616
EMail: paitken@cisco.com

Andrew Johnson
Cisco Systems, Inc.
96 Commercial Quay
Commercial Street
Edinburgh, EH6 6LX, United Kingdom

Phone: +44 131 561 3641
EMail: andrjohn@cisco.com

Gerhard Muenz
Technische Universitaet Muenchen
Department of Informatics - I8
Boltzmannstr. 3
Garching D-85748
DE

EMail: muenz@net.in.tum.de
URI: http://www.net.in.tum.de/~muenz