IPFIX Working Group   B. Claise
Internet-Draft   Cisco Systems, Inc.
Intended status: Standards Track   A. Kobayashi
Expires: July 8, 2013   NTT
                                                      B. Trammell
                                                       ETH Zurich
                                                  January 4, 2013

      Operation of the IP Flow Information Export (IPFIX) Protocol on IPFIX
                              Mediators
              draft-ietf-ipfix-mediation-protocol-03.txt

Abstract

   This document specifies the the operation of the IP Flow Information
   Export (IPFIX) protocol specific to IPFIX Mediators, including
   Template and Observation Point management, timing considerations, and
   other Mediator-specific concerns.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

   The IPFIX architectural components in [RFC5470] consist of IPFIX
   Devices and IPFIX Collectors communicating using the IPFIX protocol
   [I-D.ietf-ipfix-protocol-rfc5101bis], which specifies how to export
   IP 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
   [I-D.ietf-ipfix-information-model-rfc5102bis], registered with IANA,
   or specified as an enterprise-specific Information Element. The
   specifications in the IPFIX protocol
   [I-D.ietf-ipfix-protocol-rfc5101bis] have not been defined in the
   context of an IPFIX Mediator receiving, aggregating, correlating,
   anonymizing, etc... Flow Records from the one or multiple Exporters.
   Indeed, the IPFIX protocol must be adapted for Intermediate
   Processes, as defined in the IPFIX Mediation Reference Model as
   specified in Figure A of [RFC6183], which is based on the IPFIX
   Mediation Problem Statement [RFC5982].

   This document specifies the IP Flow Information Export (IPFIX)
   protocol in the context of the implementation and deployment of IPFIX
   Mediators. The use of the IPFIX protocol within a Mediator -- a
   device which contains both as a Collecting Process and an Exporting
   Process -- has an impact on the technical details of the usage of the
   protocol. An overview of the technical problem is covered in section
   6 of [RFC5982]: loss of original exporter information, loss of base
   time information, transport sessions management, loss of Options
   Template Information, Template Id management, considerations for
   network considerations for aggregation.

   The specifications in this document are based on the IPFIX protocol
   specifications [I-D.ietf-ipfix-protocol-rfc5101bis] but adapted
   according to the IPFIX Mediation Framework [RFC6183].

1.1. IPFIX Documents Overview

   The IPFIX Protocol [I-D.ietf-ipfix-protocol-rfc5101bis] provides
   network administrators with access to IP Flow information.

   The architecture for the export of measured IP Flow information out
   of an IPFIX Exporting Process to a Collecting Process is defined in
   the IPFIX Architecture [RFC5470], per the requirements defined in the
   IPFIX Requirement doc, [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
name, type and additional semantic information, as specified in the
IPFIX Information Model
[I-D.ietf-ipfix-information-model-rfc5102bis].

The IPFIX Applicability Statement [RFC5472] describes what type of
applications can use the IPFIX protocol and how they can use the
information provided. It furthermore shows how the IPFIX framework
relates to other architectures and frameworks.

"IPFIX Mediation: Problem Statement" [RFC5982], describing the IPFIX
Mediation applicability examples, along with some problems that
network administrators have been facing, is the basis for the "IPFIX
Mediation: Framework" [RFC6183]. This framework details the IPFIX
Mediation reference model and the components of an IPFIX Mediator.

1.2. IPFIX Mediator Documents Overview

The "IPFIX Mediation: Problem Statement" [RFC5982] provides an
overview of the applicability of Mediators, and defines requirements
for Mediators in general terms. This document is of use largely to
define the problems to be solved through the deployment of IPFIX
Mediators, and to provide scope to the role of Mediators within an
IPFIX collection infrastructure.

The "IPFIX Mediation: Framework" [RFC6183] provides more
architectural details of the arrangement of Intermediate Processes
within a Mediator.

The details of specific Intermediate Processes, when these have
additional export specifications (e.g., metadata about the
intermediate processing conveyed through IPFIX Options Templates),
are each treated in their own document (e.g., the "IP Flow
Anonymization Support" [RFC6235]). Documents specifying the
operations of specific Intermediate Processes cover the operation of
these Processes within the Mediator framework, and comply with the
specifications given in this document; they may additionally specify
the operation of the process independently, outside the context of a
Mediator, when this is appropriate. As of today, these documents
are:

1. "IP Flow Anonymization Support", [RFC6235], which describes
    Anonymization techniques for IP flow data and the export of
    Anonymized data using the IPFIX protocol.

2.   "Flow Selection Techniques" [I-D.ietf-ipfix-flow-selection-tech],
     which describes the process of selecting a subset of flows from
     all flows observed at an observation point, the flow selection
     motivations, and some specific flow selection techniques.

3.   "Exporting Aggregated Flow Data using IP Flow Information Export"
     [I-D.ietf-ipfix-a9n] which describes Aggregated Flow export
     within the framework of IPFIX Mediators and defines an
     interoperable, implementation-independent method for Aggregated
     Flow export.

This document specifies the IP Flow Information Export (IPFIX)
protocol specific to Mediation, i.e. the specifications that all
Intermediate Processes type must comply to. Some extra
specifications might be required per Intermediate Process type (In
which case, the Intermediate Process specific document would cover
those).

1.3. Relationship with IPFIX and PSAMP

The specification in this document applies to the IPFIX protocol
specifications [I-D.ietf-ipfix-protocol-rfc5101bis]. All
specifications from [I-D.ietf-ipfix-protocol-rfc5101bis] 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. Therefore, the
method specified by this document also applies to PSAMP.

2. Terminology

[EDITOR'S NOTE: change to only define terms in this section that are
actually used in the document.]

[EDITOR'S NOTE: Definition change proposal for the Intermediate
Process, Intermediate Conversion Process, Intermediate Selection
Process, Intermediate Anonymization Process, and IPFIX Mediator. See
http://www.ietf.org/mail-archive/web/ipfix/current/msg05969.html.
However, the definitions are copied over verbatim from RFC6183. Also
note that Intermediate Anonymization Process in this document is not
in line with the RFC6235.]

IPFIX-specific terms, such as Observation Domain, Flow, Flow Key,
Metering Process, Exporting Process, Exporter, IPFIX Device,
Collecting Process, Collector, Template, IPFIX Message, Message
Header, Template Record, Data Record, Options Template Record, Set,

Data Set, Information Element, and Transport Session, used in this
document are defined in [I-D.ietf-ipfix-protocol-rfc5101bis]. The
PSAMP-specific terms used in this document, such as Filtering and
Sampling, are defined in [RFC5476].

IPFIX Mediation terms related to aggregation, such as the Interval,
Aggregated Flow, and Aggregated Function are defined in
[I-D.ietf-ipfix-a9n].

The IPFIX Mediation-specific terminology used in this document is
defined in "IPFIX Mediation: Problem Statement" [RFC5982], and reused
in "IPFIX Mediation: Framework" [RFC6183]. However, since both of
those documents are an informational RFCs, the definitions have been
reproduced here along with additional definitions.

Similarly, since [RFC6235] is an experimental RFC, the Anonymization
Record, Anonymized Data Record, and Intermediate Anonymization
Process terms, specified in [RFC6235], are also reproduced here.

In this document, as in [I-D.ietf-ipfix-protocol-rfc5101bis],
[RFC5476], [I-D.ietf-ipfix-a9n], and [RFC6235], the first letter of
each IPFIX-specific and PSAMP-specific term is capitalized along with
the IPFIX Mediation-specific term defined here. In this document, we
call a stream of records carrying flow- or packet-based information a
"record stream". The records may be encoded as IPFIX Data Records of
any other format.

Transport Session Information: The Transport Session is specified
   in [I-D.ietf-ipfix-protocol-rfc5101bis]. In SCTP, the Transport
   Session Information is the SCTP association. In TCP and UDP, the
   Transport Session Information corresponds to a 5-tuple {Exporter
   IP address, Collector IP address, Exporter transport port,
   Collector transport port, transport protocol}.

Original Exporter: An Original Exporter is an IPFIX Device that
   hosts the Observation Points where the metered IP packets are
   observed.

Original Observation Point: An Observation Point of the Original
   Exporter. In the case of the Intermediate Aggregation Process on
   an IPFIX Mediator, the Original Observation Point can be composed
   of, but not limited to, a (set of) specific exporter(s), a (set
   of) specific interface(s) on an Exporter, a (set of) line card(s)
   on an Exporter, or any combinations of these.

IPFIX Mediation: IPFIX Mediation is the manipulation and conversion
   of a record stream for subsequent export using the IPFIX protocol.

Template Mapping: A mapping from Template Records and/or Options
   Template Records received by a Mediator to Template Records and/or
   Options Template Records sent by that IPFIX Mediator. Each entry
   in a Template Mapping is scoped by incoming or outgoing Transport
   Session and Observation Domain, as with Templates and Options
   Templates in the IPFIX Protocol.

Anonymization Record: A record that defines the properties of the
   anonymization applied to a single Information Element within a
   single Template or Options Template, as in [RFC6235].

Anonymized Data Record: A Data Record within a Data Set containing
   at least one Information Element with Anonymized values. The
   Information Element(s) within the Template or Options Template
   describing this Data Record SHOULD have a corresponding
   Anonymization Record, as in [RFC6235].

The following terms are used in this document to describe the
architectural entities used by IPFIX Mediation.

Intermediate Process: An Intermediate Process takes a record stream
   as its input from Collecting Processes, Metering Processes, IPFIX
   File Readers, other Intermediate Processes, or other record
   sources; performs some transformations on this stream, based upon
   the content of each record, states maintained across multiple
   records, or other data sources; and passes the transformed record
   stream as its output to Exporting Processes, IPFIX File Writers,
   or other Intermediate Processes, in order to perform IPFIX
   Mediation. Typically, an Intermediate Process is hosted by an
   IPFIX Mediator. Alternatively, an Intermediate Process may be
   hosted by an Original Exporter.

IPFIX Mediator: An IPFIX Mediator is an IPFIX Device that provides
   IPFIX Mediation by receiving a record stream from some data
   sources, hosting one or more Intermediate Processes to transform
   that stream, and exporting the transformed record stream into
   IPFIX Messages via an Exporting Process. In the common case, an
   IPFIX Mediator receives a record stream from a Collecting Process,
   but it could also receive a record stream from data sources not
   encoded using IPFIX, e.g., in the case of conversion from the
   NetFlow V9 protocol [RFC3954] to IPFIX protocol.

Specific Intermediate Processes are described below. However, this
is not an exhaustive list.

   Intermediate Conversion Process: An Intermediate Conversion Process
      is an Intermediate Process that transforms non-IPFIX into IPFIX,
      or manages the relation among Templates and states of incoming/
      outgoing Transport Sessions (or equivalent for non IPFIX
      protocols) in the case of transport protocol conversion (e.g.,
      from UDP to SCTP).

   Intermediate Aggregation Process: An Intermediate Aggregation
      Process is an Intermediate Process that aggregates records based
      upon a set of Flow Keys or functions applied to fields from the
      record (e.g., binning and subnet aggregation).

   Intermediate Correlation Process: An Intermediate Correlation
      Process is an Intermediate Process that adds information to
      records, noting correlations among them, or generates new records
      with correlated data from multiple records (e.g., the production
      of bidirectional flow records from unidirectional flow records).

   Intermediate Selection Process: An Intermediate Selection Process
      is an Intermediate Process that selects records from a sequence
      based upon criteria-evaluated record values and passes only those
      records that match the criteria (e.g., Filtering only records from
      a given network to a given Collector).

   Intermediate Anonymization Process: An Intermediate Anonymization
      Process is an Intermediate Process that transforms records in
      order to anonymize them, to protect the identity of the entities
      described by the records (e.g., by applying prefix-preserving
      pseudonymization of IP addresses).

3. Handling IPFIX Message Headers

   The format of the IPFIX Message Header as exported by an IPFIX
   Mediator is shown in Figure 1. Note that the format is compatible
   with the IPFIX Message Header defined in
   [I-D.ietf-ipfix-protocol-rfc5101bis], with some field definitions
   (for the example, the Export Time) updated in the context of the
   IPFIX Mediator.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Version           |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Export Time                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sequence Number                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Observation Domain ID                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
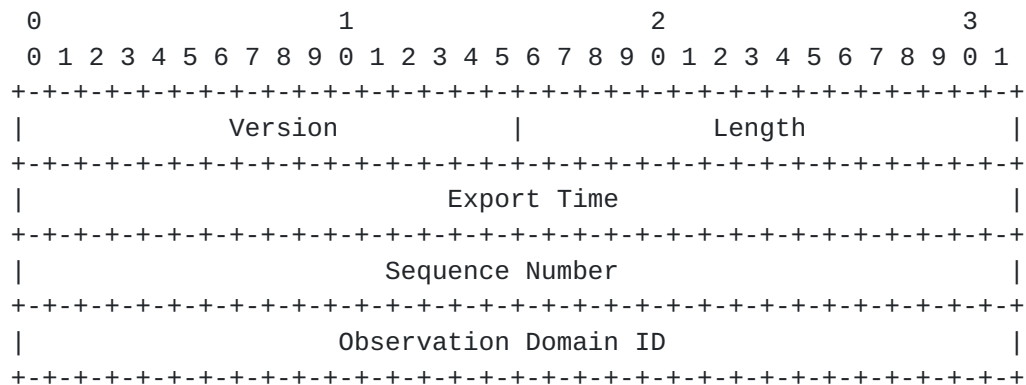
                 Figure 1: IP Message Header format

The header fields as exported by an IPFIX Mediator are describe
below.

Version: Version of Flow Record format exported in this message.
   The value of this field is 0x000a for the current version,
   incrementing by one the version used in the NetFlow services
   export version 9 [RFC3954].

Length: Total length of the IPFIX Message, measured in octets,
   including Message Header and Set(s).

Export Time: Time at which the IPFIX Message leaves the Mediator,
   expressed in seconds since the UNIX epoch of 1 January 1970 at
   00:00 UTC, as defined in [POSIX.1] encoded as an unsigned 32-bit
   integer. However, in the specific case of an IPFIX Mediator
   containing an Intermediate Conversion Process, the IPFIX Mediator
   MAY keep the export time received from the incoming Transport
   Session.

Sequence Number: Incremental sequence counter modulo 2^32 of all
   IPFIX Data Records sent in a the current stream from the current
   Observation Domain by the Exporting Process. Each SCTP Stream
   counts sequence numbers separately, while all messages in a TCP
   connection or UDP transport session are considered to be part of
   the same stream. This value SHOULD be used by the Collecting
   Process to identify whether any IPFIX Data Records have been
   missed. Template and Options Template Records do not increase the
   Sequence Number.

Observation Domain ID: A 32-bit identifier of the Observation
   Domain that is locally unique to the Exporting Process. The
   Exporting Process uses the Observation Domain ID to uniquely
   identify to the Collecting Process the Observation Domain that
   metered the Flows. It is RECOMMENDED that this identifier should

be unique per IPFIX Device. Collecting Processes SHOULD use the
Transport Session and the Observation Domain ID field to separate
different export streams originating from the same Exporting
Process. The Observation Domain ID SHOULD be 0 when no specific
Observation Domain ID is relevant for the entire IPFIX Message.
For example, when exporting the Exporting Process Statistics, or
in case of hierarchy of Collector when aggregated Data Records are
exported. See Section 4.1 for special considerations for
Observation Domain management while passing unmodified templates
through a Mediator, and Section 5 for guidelines for preservation
of original Observation Domain information at a Mediator.

4. Template Management

How a Mediator handles the Templates it receives from the Original
Exporter depends entirely on the nature of the Intermediate Process
running on that Mediator. For Mediators which pass substantially the
same Data Records from the Original Exporter downstream, (e.g., an
Intermediate Selection Process), the templates can be passed
unmodified as described in Section 4.1; this section describes a
Template Mapping required to make this work in the general case.
Mediators which export Data Records which are substantially changed
from the Data Records received from the Original Exporter follow the
guidelines in Section 4.1 instead.

Subsequent subsections deal with specific issues in Template
management that may occur at Mediators.

4.1. Passing Unmodified Templates through a Mediator

[EDITOR'S NOTE: the definition of template mappings seems really
implementation specific -- why not notionally just map IDs on each
socket to a base template? on the other hand, if we're providing a
real example, it should have concrete content in each field.
reformatting is held off until this issue is resolved.]

The first case is a situation where the IPFIX Mediator doesn't modify
the (Options) Template Record(s) content. A typical example is an
Intermediate Selection Process acting as distributor, which collects
Flow Records from one or more Exporters, and based on the Information
Elements content, redirects the Flow Records to the appropriate
Collector. This example is a typical case of a single network
operation center managing multiple universities: an unique IPFIX
Collector collects all Flow Records for the common infrastructure,
but might be re-exporting specific university Flow Records to the
responsible system administrator.

As specified in [I-D.ietf-ipfix-protocol-rfc5101bis], the Template
IDs are unique per Exporter, per Transport Session, and per
Observation Domain. As there is no guarantee that, for similar
Template Records, the Template IDs received on the incoming Transport
Session and exported to the outgoing Transport Session would be same,
the IPFIX Mediator MUST maintain a Template Mapping composed of
related received and exported (Options) Template Records:

o for each received (Options) Template Record: Template Record Flow
   Keys and non Flow Keys, Template ID, Observation Domain Id, and
   Transport Session Information

o for each exported (Options) Template Record: Template Record Flow
   Keys and non Flow Keys, Template ID, Collector, Observation Domain
   Id, and Transport Session Information

If an IPFIX Mediator receives an IPFIX Withdrawal Message for a
(Options) Template Record that is not used anymore in any other
Template Mappings, the IPFIX Mediator SHOULD export the appropriate
IPFIX Withdrawal Message(s) on the outgoing Transport Session, and
remove the corresponding entry in the Template Mapping.

If a (Options) Template Record is not used anymore in an outgoing
Transport Session, it MUST be withdrawn with an IPFIX Template
Withdrawal Message on that specific outgoing Transport Session, and
its entry MUST be removed from the Template Mapping.

If an incoming or outgoing Transport Session is gracefully shutdown
or reset, the (Options) Template Records corresponding to that
Transport Session MUST be removed from the Template Mapping.

For example, Figure 2 displays an example of an Intermediate
Selection Process, re-distributing Data Records to Collectors on the
basis of customer networks, i.e. the Route Distinguisher (RD). In
this example, the Template Record received from the Exporter #1 is
reused towards Collector #1, Collector #2, and Collector #3.

```
                                 Tmpl.  .---------.
                                 ID 256 |         |
                                  .---->|Collector|<==>Customer
                                  |     |#1       |     A
                                  |     |         |
                               RD=100:1 '---------'
    .---------.Templ.  .---------.   |
    |         |Id      |         |----'      .---------.
    |         |258     |         | RD=100:2  |         |
    |IPFIX    |------->|IPFIX    |---------->|Collector|<==>Customer
    |Exporter |        |Mediator | Tmpl.     |#2       |     B
    |#1       |        |         | ID 257    |         |
    |         |        |         |----.      '---------'
    '---------'        '---------'    |
                                 RD=100:3
                              Tmpl. |      .---------.
                              ID    |      |         |
                              257   '----->|Collector|<==>Customer
                                           |#3       |     C
                                           |         |
                                           '---------'
```
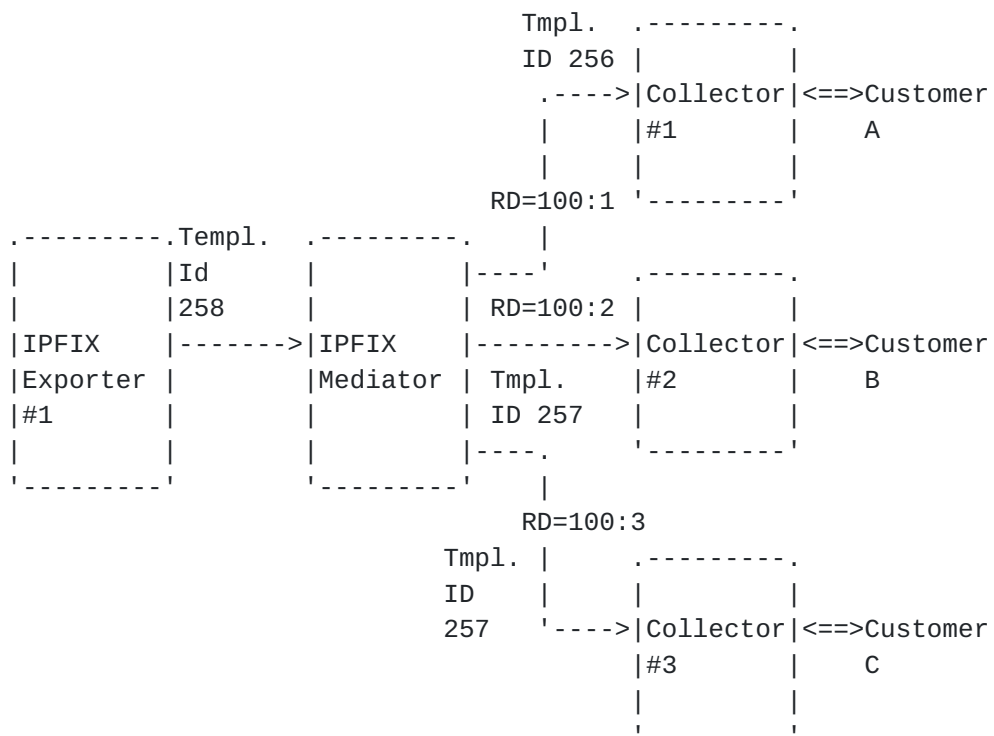
             Figure 2: Intermediate Selection Process example

      Figure 3 shows the Template Mapping for the system shown in Figure 2.

```
Template Entry A:
Incoming Transport Session Information (from Exporter#1):
  Source IP: <Exporter#1 export IP address>
  Destination IP: <IPFIX Mediator IP address>
  Protocol: SCTP
  Source Port: <source port>
  Destination Port: 4739 (IPFIX)
Observation Domain Id: <Observation Domain ID>
Template Id: 258
Flow Keys: <series of Flow Keys>
Non Flow Keys: <series of non Flow Keys>

Template Entry B:
Outgoing Transport Session Information (to Collector#1):
  Source IP: <IPFIX Mediator IP address>
  Destination IP: <IPFIX Collector#1 IP address>
  Protocol: SCTP
  Source Port: <source port>
  Destination Port: 4739 (IPFIX)
Observation Domain Id: <Observation Domain ID>
Template Id: 256
Flow Keys: <series of Flow Keys>
Non Flow Keys: <series of non Flow Keys>

Template Entry C:
Outgoing Transport Session Information (to Collector#2):
  Source IP: <IPFIX Mediator IP address>
  Destination IP: <IPFIX Collector#2 IP address>
  Protocol: SCTP
  Source Port: <source port>
  Destination Port: 4739 (IPFIX)
Observation Domain Id: <Observation Domain ID>
Template Id: 257
Flow Keys: <series of Flow Keys>
Non Flow Keys: <series of non Flow Keys>

Template Entry D:
Outgoing Transport Session Information (to Collector#3):
  Source IP: <IPFIX Mediator IP address>
  Destination IP: <IPFIX Collector#3 IP address>
  Protocol: SCTP
  Source Port: <source port>
  Destination Port: 4739 (IPFIX)
Observation Domain Id: <Observation Domain ID>
  Template Id: 257
Flow Keys: <series of Flow Keys>
Non Flow Keys: <series of non Flow Keys>
```

Figure 3: Template Mapping example: templates

The Template Mapping corresponding to figure B can be displayed as:

```
Template Entry A   <----> Template Entry B
Template Entry A   <----> Template Entry C
Template Entry A   <----> Template Entry D
```

                   Template Mapping example: mappings

Alternatively, the Template Mapping may be optimized as:

```
                   +--> Template Entry B
                   |
Template Entry A   <--+--> Template Entry C
                   |
                   +--> Template Entry D
```

                   Template Mapping example: mappings

Note that all examples use Transport Sessions based on the SCTP
protocol, as simplified use cases. However, the protocol would be
important in situations such as an Intermediate Conversion Process
doing transport protocol conversion.

4.2. Creating New Templates at a Mediator

The second case is a situation where the IPFIX Mediator generates new
(Options) Template Records as a result of the Intermediate Process.

In this situation, the IPFIX Mediator doesn't need to maintain a
Template Mapping, as it generates its own series of (Options)
Template Records. However, the following special case might still
require a Template Mapping, i.e. a situation where the IPFIX
Mediator, typically containing an Intermediate Conversion Process,
Intermediate Aggregation Process [I-D.ietf-ipfix-a9n], or
Intermediate Anonymization Process in case of black-marker
Anonymization [RFC6235], generates new (Options) Template Records
based on what it receives from the Exporter(s), and based on the
Intermediate Process function. In such a case, it's important to
keep the correlation between the received (Options) Template Records
and exported Derived (Options) Template Records in the Template
Mapping. These template mappings would be kept as in Section 4.1,
except that the export template would not be identical to the
collection template.

4.3. Information Element Ordering within Templates

   [EDITOR'S NOTE: address the following: What Paul Aikten would like to
   see in section 3.5 (See
   http://www.ietf.org/mail-archive/web/ipfix/current/msg05969.html):
   What about IE ordering? May an exporter re-order received fields?
   eg, two devices sending the same information, though with the fields
   in a different order. Or the mediator is extracting the same
   information from two sources. That seems to be a valid scenario. eg,
   this reduces the number of templates received at the collector.]

4.4. Handling Unknown Information Elements

   [EDITOR'S NOTE: also from Paul Aitken: What should a mediator do with
   a field which it doesn't know/understand? Inevitably, exporters will
   be updated without mediators keeping in step. It's also very likely
   that mediators will see Enterprise-specific IEs. May a mediator re-
   export unknown IEs unchanged, or should it drop them? Presumably a
   mediator may report received Enterprise-specific IEs even from
   multiple different Enterprises. What if an unknown field depends on
   the field ordering? eg, it's a bitfield like flowKeyIndicator. Re-
   ordering, adding or removing fields breaks the meaning of this field,
   so it can't be passed on. It can only be used if the received fields
   are reported unchanged.]

5. Preserving Original Observation Point Information

   [EDITOR'S NOTE: Decide whether we want to address export of
   observation point information without 6313. Review this section to
   make sure it adequately explains how original Observation Point
   information can get so complicated.]

   Depending on the use case, the Collector in an Exporter - Mediator -
   Collector structure may need to receive information about the
   Original Observation Point(s), otherwise it may wrongly conclude that
   the IPFIX Device exporting the Flow Records, i.e. the IPFIX Mediator,
   directly observed the packets that generated the Flow Records. Two
   new Information Elements are introduced in the subsections below to
   address this use case: originalExporterIPv4Address and
   originalExporterIPv6Address. Practically, the Original Exporters
   will not exporting these Information Elements. Therefore, the
   Intermediate Process SHOULD report the Original Observation Point(s)
   to the best of its knowledge. Note that the Configuration Data Model
   for IPFIX and PSAMP [RFC6728] may help.

   In the IPFIX Mediator, the Observation Point(s) may be represented
   by:

o A single Original Exporter (represented by the
   originalExporterIPv4Address or originalExporterIPv6Address
   Information Elements)

o A list of Original Exporters (represented by the
   originalExporterIPv4Address or originalExporterIPv6Address
   Information Elements).

o Any combination or list of Information Elements representing
   Observation Points. For example:

   * A list of Original Exporter interface(s) (represented by the
      originalExporterIPv4Address or originalExporterIPv6Address, the
      ingressInterface and/or egressInterface Information Elements,
      respectively)

   * A list of Original Exporter line card (represented by the
      originalExporterIPv4Address or originalExporterIPv6Address, the
      lineCardId Information Elements, respectively)

Some Information Elements characterizing the Observation Point may be
added. For example, the flowDirection Information Element specifies
the direction of the observation, and, as such, characterizes the
Observation Point.

Any combination of the above representations is possible. For
example, in case of an Intermediate Aggregation Process, an Original
Observation Point could be composed of:

exporterIPv4Address 192.0.2.1
exporterIPv4Address 192.0.2.2,
  interface ethernet 0, direction ingress
  interface ethernet 1, direction ingress
  interface serial 1, direction egress
  interface serial 2, direction egress
exporterIPv4Address 192.0.2.3,
  lineCardId 1, direction ingress

        Figure 4: Complex Observation Point Definition Example

If the Original Observation Point is composed of a list, then the
IPFIX Structured Data [RFC6313] MUST be used to export it from the
IPFIX Mediator.

The most generic way to export the Original Observation Point is to
use a subTemplateMultiList, with the semantic "exactlyOneOf". Taking
the previous example, the following encoding can be used:

```
   Template Record 257: exporterIPv4Address
   Template Record 258: exporterIPv4Address,
                        basicList of ingressInterface, flowDirection
   Template Record 259: exporterIPv4Address, lineCardId, flowDirection
```

     Figure 5: Complex Observation Point Definition Example: Templates

   The Original Observation Point is modeled with the Data Records
   corresponding to either Template Record 1, Template Record 2, or
   Template Record 3 but not more than one of these ("exactlyOneOf"
   semantic). This implies that the Flow was observed at exactly one of
   the Observation Points reported.

   When an IPFIX Mediator receives Flow Records containing the Original
   Observation Point Information Element, i.e.
   originalExporterIPv6Address or originalExporterIPv4Address, the IPFIX
   Mediator SHOULD NOT modify its value(s) when composing new Flow
   Records in the general case. Known exceptions include anonymization
   per [RFC6235] section 7.2.4 and an Intermediate Correlation Process
   rewriting addresses across NAT. In other words, the Original
   Observation Point should not be replaced with the IPFIX Mediator
   Observation Point. The daisy chain of (Exporter, Observation Point)
   representing the path the Flow Records took from the Exporter to the
   top Collector in the Exporter - Mediator(s) - Collector structure
   model is out of the scope of this specification.

5.1. originalExporterIPv4Address Information Element

   Description: The IPv4 address used by the Exporting Process on an
      Original Exporter, as seen by the Collecting Process on an IPFIX
      Mediator. Used to provide information about the Original
      Observation Points to a downstream Collector.

   Data Type: ipv4Address

   ElementId: TBD1

5.2. originalExporterIPv6Address Information Element

   Description: The IPv6 address used by the Exporting Process on an
      Original Exporter, as seen by the Collecting Process on an IPFIX
      Mediator. Used to provide information about the Original
      Observation Points to a downstream Collector.

   Data Type: ipv6Address

ElementId: TBD2

6. Managing Observation Domain IDs

In any case, the Observation Domain ID of any IPFIX Message
containing Flow Records relevant to no particular Observation Domain,
or to multiple Observation Domains, MUST have an Observation Domain
ID of 0, as in Section 3 above, and section 3.1 of
[I-D.ietf-ipfix-protocol-rfc5101bis].

IPFIX Mediators that do not change (Options) Template Records MUST
maintain a Template Mapping, as detailed in Section 4.1, to ensure
that the combination of Observation Domain IDs and Template IDs do
not collide on export.

For IPFIX Mediators that export New (Options) Template Records, as in
Section 4.2, there are two options for Observation Domain ID
management. The first and simplest of these is to completely
decouple exported Observation Domain IDs from received Observation
Domain IDs; the IPFIX Mediator, in this case, comprises its own set
of Observation Domain(s) independent of the Observation Domain(s) of
the Original Exporters.

The second option is to provide or maintain a Template Mapping for
received (Options) Template Records and exported inferred (Options)
Template Records, along with the appropriate Observation Domain IDs
per Transport Session, which ensures that the combination of
Observation Domain IDs and Template IDs do not collide on export.

In some cases where the IPFIX Message Header can't contain a
consistent Observation Domain for the entire IPFIX Message, but the
Flow Records exported from the IPFIX Mediator should anyway contain
the Observation Domain of the Original Exporter, the (Options)
Template Record must contain the originalObservationDomainId
Information Element. When an IPFIX Mediator receives Flow Records
containing the originalObservationDomainId Information Element, the
IPFIX Mediator MUST NOT modify its value(s) when composing new Flow
Records with the originalObservationDomainId Information Element.

6.1. originalObservationDomainId Information Element

Description: The Observation Domain ID reported by the Exporting
   Process on an Original Exporter, as seen by the Collecting Process
   on an IPFIX Mediator. Used to provide information about the
   Original Observation Domain to a downstream Collector.

   Data Type: unsigned32

   Data Type Semantics: identifier

   ElementId: TBD3

7. **Timing Considerations**

   The IPFIX Message Header "Export Time" field is the time in seconds
   since 0000 UTC Jan 1, 1970, at which the IPFIX Message leaves the
   IPFIX Mediator. However, in the specific case of an IPFIX Mediator
   containing an Intermediate Conversion Process, the IPFIX Mediator MAY
   keep the export time received from the incoming Transport Session.

   It is RECOMMENDED that Mediators handle time using absolute
   timestamps (e.g. flowStartSeconds, flowStartMilliseconds,
   flowStartNanoseconds), which are specified relative to the UNIX epoch
   (00:00 UTC 1 Jan 1970), where possible, rather than relative
   timestamps (e.g. flowStartSysUpTime, flowStartDeltaMicroseconds),
   which are specified relative to protocol structures such as system
   initialization or message export time.

   The latter are difficult to manage for two reasons. First, they
   require constant translation, as the system initialization time of an
   intermediate system and the export time of an intermediate message
   will change across mediation operations. Further, relative
   timestamps introduce range problems. For example, when using the
   flowStartDeltaMicroseconds and flowEndDeltaMicroseconds Information
   Elements [iana-ipfix-assignments], the Data Record must be exported
   within a maximum of 71 minutes after its creation. Otherwise, the
   32-bit counter would not be sufficient to contain the flow start time
   offset. Those time constraints might be incompatible with some of
   the application requirements of some Intermediate Processes.

   Intermediate Processes MUST NOT assume that received records appear
   in flowStartTime, flowEndTime, or observationTime order. An
   Intermediate Process processing timing information (e.g., an
   Intermediate Aggregation Process) MAY ignore records that are
   significantly out of order, in order to meet application-specific
   state and latency requirements, but SHOULD report that records were
   dropped.

   When an Intermediate Process aggregates information from different
   Flow Records, the timestamps on exported records SHOULD be the
   minimum of the start times and the maximum of the end times in the
   general case. However, if the Flow Records do not overlap, i.e. if
   there is a time gap between the times in the Flow Records, then the

report may be inaccurate. The IPFIX Mediator is only reporting what
it knows, on the basis of the information made available to it - and
there may not have been any data to observe during the gap. Then
again, if there is an overlap in timestamps, there's the potential of
double-accounting: different Observation Points may have observed the
same traffic simultaneously. Therefore, as there is not a single
rule that fits all different situations, a complete specification of
the precise rules of applying Flow Record timestamps at IPFIX
Mediators is out of the scope of this document.

Note that [I-D.ietf-ipfix-a9n] provides additional specifications for
handling of timestamps at an Intermediate Aggregation Process.

8. Transport Considerations

SCTP [RFC4960] using the PR-SCTP extension specified in [RFC3758]
MUST be implemented by all compliant IPFIX Mediator implementations.
TCP [RFC0793] MAY also be implemented by IPFIX Mediator compliant
implementations. UDP [RFC0768] MAY also be implemented by compliant
IPFIX Mediator implementations. Transport-specific considerations
for IPFIX Exporters as specified in sections 8.3, 8.4, 9.1, 9.2, and
10 of [I-D.ietf-ipfix-protocol-rfc5101bis] apply to IPFIX Mediators
as well.

SCTP SHOULD be used in deployments where IPFIX Mediators and
Collectors are communicating over links that are susceptible to
congestion. SCTP is capable of providing any required degree of
reliability. TCP MAY be used in deployments where IPFIX Mediators
and Collectors communicate over links that are susceptible to
congestion, but SCTP is preferred due to its ability to limit back
pressure on Exporters and its message versus stream orientation. UDP
MAY be used, although it is not a congestion-aware protocol.
However, in this case, the IPFIX traffic between IPFIX Mediator and
Collector MUST run in an environment where IPFIX traffic has been
provisioned for, or is contained through some other means.

9. Collecting Process Considerations

Any Collecting Process compliant with
[I-D.ietf-ipfix-protocol-rfc5101bis] can receive IPFIX Messages from
an IPFIX Mediator. If the IPFIX Mediator uses IPFIX Structured Data
[RFC6313] to export Original Exporter Information as in Section 5,
the Collecting Process MUST support [RFC6313].

10. Specific Reporting Requirements

IPFIX provides Options Templates for the reporting on the reliability
of processes within the IPFIX Architecture. As each Mediator
includes at least one IPFIX Exporting Process, they SHOULD use the
Exporting Process Reliability Statistics Options Template, as
specified in [I-D.ietf-ipfix-protocol-rfc5101bis].

Analogous to the Metering Process Reliability Statistics Options
Template, also specified in [I-D.ietf-ipfix-protocol-rfc5101bis],
Mediators SHOULD implement the Intermediate Process Reliability
Statistics Options Template, specified in the subsection below.

The Flow Keys Options Template, as specified in
[I-D.ietf-ipfix-protocol-rfc5101bis], may require special handling at
an IPFIX Mediator as described below.

In addition, each Intermediate Process may have its own specific
reporting requirements (e.g. Anonymization Records as in [RFC6235],
or the Aggregation Counter Distribution Options Template as in
[I-D.ietf-ipfix-a9n]); these SHOULD be implemented as necessary as
described in the specification for each Intermediate Process.

10.1. Intermediate Process Reliability Statistics Template

The Intermediate Process Statistics Options Template specifies the
structure of a Data Record for reporting Intermediate Process
statistics. It SHOULD contain the following Information Elements;
the intermediateProcessId Information Element is defined in
Section 10.3, and the ignoredRecordTotalCount Information Element is
defined in Section 10.4:

```
+-------------------------+-----------------------------------------+
| IE                      | Description                             |
+-------------------------+-----------------------------------------+
| observationDomainId     | An identifier of the Observation Domain |
| [scope]                 | (of messages exported by this           |
|                         | Mediator), locally unique to the        |
|                         | Intermediate Process, to which this     |
|                         | statistics record applies.              |
| intermediateProcessId   | An identifier for the Intermediate      |
| [scope]                 | Process to which this statistics record |
|                         | applies.                                |
| ignoredRecordTotalCount | The total number of Data Records        |
|                         | received but not processed by the       |
|                         | Intermediate Process.                   |
```

```
      | time first record      | The timestamp of the first record that  |
      | ignored                | was ignored by the Intermediate         |
      |                        | Process.  For Data Records containing    |
      |                        | timestamp ranges, this SHOULD be taken   |
      |                        | from the start timestamp of the range;   |
      |                        | for data records containing no timing    |
      |                        | information, this SHOULD be taken from    |
      |                        | the Export Time in the message header     |
      |                        | of the containing IPFIX Message.  For     |
      |                        | this timestamp, any of the following      |
      |                        | timestamp can be used:                    |
      |                        | observationTimeSeconds,                   |
      |                        | observationTimeMilliseconds,              |
      |                        | observationTimeMicroseconds, or           |
      |                        | observationTimeNanoseconds.               |
      | time last record       | The timestamp of the last record that    |
      | ignored                | was ignored by the Intermediate           |
      |                        | Process.  For Data Records containing     |
      |                        | timestamp ranges, this SHOULD be taken    |
      |                        | from the end timestamp of the range;      |
      |                        | for data records containing no timing     |
      |                        | information, this SHOULD be taken from    |
      |                        | the Export Time in the message header     |
      |                        | of the containing IPFIX Message.  For     |
      |                        | this timestamp, any of the following      |
      |                        | timestamp can be used:                    |
      |                        | observationTimeSeconds,                   |
      |                        | observationTimeMilliseconds,              |
      |                        | observationTimeMicroseconds, or           |
      |                        | observationTimeNanoseconds.               |
      +------------------------+------------------------------------------+
```

10.2. Flow Key Options Template

   The Flow Keys Option Template specifies the structure of a Data
   Record for reporting the Flow Keys of reported Flows. A Flow Keys
   Data Record extends a particular Template Record that is referenced
   by its templateId identifier. The Template Record is extended by
   specifying which of the Information Elements contained in the
   corresponding Data Records describe Flow properties that serve as
   Flow Keys of the reported Flow. This Options Template is defined in
   section 4.4 of [I-D.ietf-ipfix-protocol-rfc5101bis], and SHOULD be
   used by Mediators for export as defined there.

   When an Intermediate Process exports Data Records containing
   different Flow Keys from those received from the Original Exporter,
   and the Original Exporter sent a Flow Keys Options record to the
   Mediator, the Mediator MUST export a Flow Keys Options record

defining the the new set of Flow Keys.

10.3. intermediateProcessId Information Element

Description: An identifier of an Intermediate Process that is
   unique per IPFIX Device. Typically, this Information Element is
   used for limiting the scope of other Information Elements. Note
   that process identifiers may be assigned dynamically; ie., and
   Intermediate Process may be re-started with a different ID.

Data Type: unsigned32

Data Type Semantics: identifier

ElementId: TBD4

10.4. ignoredRecordTotalCount Information Element

Description: The total number of received Data Records that the
   Intermediate Process did not process since the (re-)initialization
   of the Intermediate Process; includes only Data Records not
   examined or otherwise handled by the Intermediate Process due to
   resource constraints, not Data Records which were examined or
   otherwise handled by the Intermediate Process but which merely do
   not contribute to any exported Data Record due to the operations
   performed by the Intermediate Process.

Data Type: unsigned64

Data Type Semantics: totalCounter

ElementId: TBD5

11. Configuration Management

In general, using Mediators to combine information from multiple
Original Exporters requires a consistent configuration of the
Metering Processes behind these Original Exporters. The details of
this consistency are specific to each Intermediate Process.
Consistency of configuration should be verified out of band, with the
MIB modules ([RFC6615] and [RFC6727]) or with the Configuration Data
Model for IPFIX and PSAMP [RFC6728]

12. Security Considerations

As they act as both IPFIX Collecting Processes and Exporting

Processes, the Security Considerations for IPFIX Protocol
[I-D.ietf-ipfix-protocol-rfc5101bis] also apply to Mediators. The
Security Considerations for IPFIX Files [RFC5655] also apply to IPFIX
Mediators that write IPFIX Files or use them for internal storage.
However, there are a few specific considerations that IPFIX Mediator
implementations must also take into account.

By design, IPFIX Mediators are "men-in-the-middle": they intercede in
the communication between an Original Exporter (or another upstream
Mediator) and a downstream Collecting Process. This has two
important implications for the level of confidentiality provided
across an IPFIX Mediator, and the ability to protect data integrity
and Original Exporter authenticity across a Mediator. These are
addressed in more detail in the Security Considerations for Mediators
in [RFC6183].

Note that, while Mediators can use the exporterCertificate and
collectorCertificate Information Elements defined in [RFC5655] as
described in section 9.3 of [RFC6183] to export information about
X.509 identities in upstream TLS-protected Transport Sessions, this
mechanism cannot be used to provide true end-to-end assertions about
a chain of IPFIX Mediators: any Mediator in the chain can simply
falsify the information about upstream Transport Sessions In
situations where information about the chain of mediation is
important, it must be determined out of band.

13. IANA Considerations

This document specifies n new IPFIX Information Elements,
originalExporterIPv4Address in Section 5.1,
originalExporterIPv6Address in Section 5.2, and
originalObservationDomainId in Section 6.1, to be added to the IPFIX
Information Element registry [iana-ipfix-assignments]. [IANA NOTE:
please add the three Information Elements as specified in the
references subsections, and change TBD1, TBD2, and TBD3 in this
document to reflect the assigned identifiers.]

14. Acknowledgments

We would like to thank the IPFIX contributors, specifically Paul
Aitken for his thorough review and Rahul Patel for his feedback and
comments. This work is materially supported by the European Union
Seventh Framework Programme under grant agreement 257315 (DEMONS).

15. References

15.1. Normative References

   [I-D.ietf-ipfix-protocol-rfc5101bis]
              Claise, B. and B. Trammell, "Specification of the IP Flow
              Information eXport (IPFIX) Protocol for the Exchange of
              Flow Information", draft-ietf-ipfix-protocol-rfc5101bis-04
              (work in progress), December 2012.

   [I-D.ietf-ipfix-information-model-rfc5102bis]
              Claise, B. and B. Trammell, "Information Model for IP Flow
              Information eXport (IPFIX)",
              draft-ietf-ipfix-information-model-rfc5102bis-08 (work in
              progress), December 2012.

   [RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768,
              August 1980.

   [RFC0793] Postel, J., "Transmission Control Protocol", STD 7,
              RFC 793, September 1981.

   [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., "Stream Control Transmission Protocol",
              RFC 4960, September 2007.

   [RFC5655] Trammell, B., Boschi, E., Mark, L., Zseby, T., and A.
              Wagner, "Specification of the IP Flow Information Export
              (IPFIX) File Format", RFC 5655, October 2009.

   [RFC6313] Claise, B., Dhandapani, G., Aitken, P., and S. Yates,
              "Export of Structured Data in IP Flow Information Export
              (IPFIX)", RFC 6313, July 2011.

   [RFC6615] Dietz, T., Kobayashi, A., Claise, B., and G. Muenz,
              "Definitions of Managed Objects for IP Flow Information
              Export", RFC 6615, June 2012.

   [I-D.ietf-ipfix-flow-selection-tech]
              D'Antonio, S., Zseby, T., Henke, C., and L. Peluso, "Flow
              Selection Techniques",
              draft-ietf-ipfix-flow-selection-tech-12 (work in
              progress), September 2012.

   [I-D.ietf-ipfix-a9n]
              Trammell, B., Wagner, A., and B. Claise, "Flow Aggregation
              for the IP Flow Information Export (IPFIX) Protocol",
              draft-ietf-ipfix-a9n-08 (work in progress), November 2012.

   [RFC6727] Dietz, T., Claise, B., and J. Quittek, "Definitions of
              Managed Objects for Packet Sampling", RFC 6727,
              October 2012.

   [RFC6728] Muenz, G., Claise, B., and P. Aitken, "Configuration Data
              Model for the IP Flow Information Export (IPFIX) and
              Packet Sampling (PSAMP) Protocols", RFC 6728,
              October 2012.

15.2. Informative References

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

   [RFC3954] Claise, B., "Cisco Systems NetFlow Services Export Version
              9", RFC 3954, 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.

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

   [RFC5982] Kobayashi, A. and B. Claise, "IP Flow Information Export
              (IPFIX) Mediation: Problem Statement", RFC 5982,
              August 2010.

   [RFC6183] Kobayashi, A., Claise, B., Muenz, G., and K. Ishibashi,
              "IP Flow Information Export (IPFIX) Mediation: Framework",
              RFC 6183, April 2011.

   [RFC6235] Boschi, E. and B. Trammell, "IP Flow Anonymization
              Support", RFC 6235, May 2011.

   [iana-ipfix-assignments]
              Internet Assigned Numbers Authority, "IP Flow Information
              Export Information Elements

            (http://www.iana.org/assignments/ipfix/ipfix.xml)".

   [POSIX.1]   IEEE, "IEEE 1003.1-2008 - IEEE Standard for Information
               Technology - Portable Operating System Interface".

Authors' Addresses

   Benoit Claise
   Cisco Systems, Inc.
   De Kleetlaan 6a b1
   1831 Diegem
   Belgium

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

   Atsushi Kobayashi
   NTT Information Sharing Platform Laboratories
   3-9-11 Midori-cho
   Musashino-shi, Tokyo 180-8585
   Japan

   Phone: +81 422 59 3978
   Email: akoba@nttv6.net

   Brian Trammell
   Swiss Federal Institute of Technology Zurich
   Gloriastrasse 35
   8092 Zurich
   Switzerland

   Phone: +41 44 632 70 13
   Email: trammell@tik.ee.ethz.ch