IPFIX Working Group
Internet-Draft
Expires: December 21, 2006

B. Trammell
CERT/NetSA
E. Boschi
Hitachi Europe

June 19, 2006

Bidirectional Flow Export using IPFIX
draft-trammell-ipfix-biflow-02.txt

Status of this Memo

Copyright Notice

Abstract

This document describes an efficient method for exporting bidirectional flow (Biflow) information using the IP Flow Information Export (IPFIX) protocol, representing each Biflow using a single Flow Record. It proposes two alternatives for information model extensions to support this method, for the consideration of the IPFIX Working Group.

Table of Contents

1. Introduction

Many flow analysis tasks benefit from association of the upstream and downstream flows of a bidirectional communication, e.g., separating answered and unanswered TCP requests, calculating round trip times, etc. Metering processes that are not part of an asymmetric routing infrastructure, especially those deployed within a single Observation Domain through which bidirectional traffic flows, are well positioned to observe bidirectional flows (Biflows). In such topologies, the total resource requirements for Biflow assembly are often lower if the Biflows are assembled at the Metering Process as opposed to the Collecting Process. IPFIX requires only information model extensions to be complete as a solution for exporting Biflow data.

To that end, we propose a Single Record Biflow export method in section 5 of this document. This method requires additional Information Elements to represent the reverse direction of each biflow; so Section 5 also presents two alternatives for policies that may be used to allocate these Information Elements. This method is motivated by an exploration of other possible methods of Biflow export; indeed, IPFIX may currently be used to export Biflow data without information model extensions at all, but the methods for doing so have important drawbacks. We describe these methods, their advantages, and their disadvantages in section 4.

2. Terminology

The terms in this section are in line with the Terminology section of the IPFIX Protocol [I-D.ietf-ipfix-protocol].

Flow: A Flow is a set of IP packets passing an Observation Point in the network during a certain time interval. All packets belonging to a particular Flow have a set of common properties. Each property is defined as the result of applying a function to the values of:

1. One or more packet header fields, transport header fields, or application header fields.

2. One or more characteristics of the packet itself.

3. One or more fields derived from packet treatment.

A packet is said to belong to a Flow if it completely satisfies all the defined properties of the Flow. This definition covers the range from a Flow containing all packets observed at a network interface to a Flow consisting of just a single packet between two

applications. It includes packets selected by a sampling mechanism.

Flow Key: Each of the fields which

1. Belong to the packet header (e.g. destination IP address)

2. Are a property of the packet itself (e.g. packet length)

3. Are derived from packet treatment (e.g. AS number)

and which are used to define a Flow are termed Flow Keys.

Directional Key Field: A Directional Key Field is a single field in a Flow Key as defined in the IPFIX Protocol [I-D.ietf-ipfix-protocol] that is specifically associated with a single endpoint of the flow. sourceIPv4Address and destinationTransportPort are example directional key fields.

Non-directional Key Field: A Non-directional Key Field is a single field within a Flow Key as defined in the IPFIX Protocol [I-D.ietf-ipfix-protocol] that is not specifically associated with either endpoint of the flow. protocolIdentifier is an example non-directional key field.

Uniflow (unidirectional flow): A Uniflow is a Flow, as above, restricted such that the Flow must be composed only of packets sent from a single endpoint to another single endpoint.

Biflow (bidirectional flow): A Biflow is a Flow composed of packets sent in both directions between two endpoints. A Biflow may also be defined as composed from two Uniflows such that:

1. each Non-directional Key Field of each Uniflow is identical to its counterpart in the other

2. each Directional Key Field of each Uniflow is identical to its reverse direction counterpart in the other

3. Biflow Semantics

As stated in the Terminology section above, a Biflow is simply a Flow representing packets flowing in both directions between two endpoints on a network. There are compelling reasons to treat Biflows as single entities within IPFIX. First, as most network communication is inherently bidirectional, a Biflow-based data model more accurately represents the behavior of the network, and enables easier

application of flow data to answering interesting questions about network behavior. Second, exporting Biflow data can result in improved export efficiency, by eliminating the duplication of Flow Key data in an IPFIX message stream.

Considering Biflows as single entities does introduce some additional semantic considerations within the IPFIX information model. When handling Uniflows, the semantics of "source" and "destination" Information Elements are clearly defined by the semantics of the underlying packet header data. When grouping Biflows into single IPFIX Data Records, the definitions of "source" and "destination" become less clear.

The most basic method for classifying the two addresses in a Biflow is to define the source and destination addresses of the flow as the source and destination addresses of the packet initiating the flow, respectively. This can be roughly approximated by a Metering Process by simply assuming the first packet seen in a given Biflow is the packet initiating the flow. Some metering technologies may improve upon this method using some knowledge of the transport or application protocols (e.g., TCP flags, DNS question/answer counts) to better approximate the flow-initiating packet. These techniques are especially useful when assembling Biflows from lossy packet sources.

Other methods of assigning direction exist. One alternate way to classify Biflow addresses is by perimeter; in this method, the Metering Process discriminates between "inside" and "outside" a network of interest, and defines the source address as the address on one side of this perimeter (generally the "outside" address; defining source loosely as "attacker"). This approach is popular in security-focused flow collection tools.

In any case, the design is the same: one of the Uniflow halves is assumed to be in the "forward" direction, and one in the "reverse" direction; which is the "forward" half is selected based upon some characteristic of the connection itself. Note that as long as these directions are assigned consistently, and there exists sufficient information in the flow record for the Collecting Process to make its own determination as to the flow's direction, the Metering Process' assignment of flow direction is irrelevant. However, for the sake of simplicity and consistency, we recommend the flow initiator method of direction assignment.

Note that, by the definition of Observation Domain in section 2 of the IPFIX Protocol [I-D.ietf-ipfix-protocol], Biflows may be composed only of packets observed within the same Observation Domain. This implies that Metering Processes that build Biflows out of Uniflow halves must ensure that the two Uniflow halves were observed within

the same Observation Domain.

4. Existing Biflow Implementation Strategies

This section describes methods by which Biflow data may be presently exported using IPFIX, without any IPFIX information model extensions. We do not recommend these approaches, but present them in order to explore the advantages and drawbacks of these methods to provide a motivation for the proposed Single Record Biflow export method.

4.1. Two-Record Biflow Export using Record Adjacency

The simplest presently available way for an Exporting Process that uses a Biflow-based internal data model to implement flow export using IPFIX is simply to split the Biflow into two Uniflow records at export time. The two Uniflow sides of the Biflow can then be placed adjacent to each other in the IPFIX Message, with the initiating Uniflow (the first Uniflow seen by the Metering Process) appearing in the message first. This simple arrangement provides enough information for a Collecting Process which uses a Biflow-based internal data model to reassemble the Biflow without requiring any computationally-intensive flow matching. When using this method the order of the Uniflow records becomes crucial, and must be maintained by both the Exporting and Collecting Processes.

This method does have the benefit of extreme simplicity. However, it also has the disadvantage of extreme simplicity. It is not a protocol so much as an informal arrangement; Collecting Processes with Biflow-based internal data models cannot rely on the courtesy of the Exporting Process to arrange Biflow halves adjacently in the flow record stream and so must support computationally-intensive flow matching anyway. No explicit association is made between the two uniflow half records. It is also record space inefficient in that every key field in the first Uniflow, whether directional or non-directional, is duplicated in the second Uniflow record.

Additionally, because UDP and SCTP Unreliable transports may drop and/or reorder packets, if these transports are used and the Exporting Process is using record adjacency for biflow export, the Exporting Process is responsible for ensuring that the two Flow Records describing the two Uniflow halves are not divided by a message boundary.

While the Record Adjacency method is simple, and is presently available, its relative export size inefficiency and lack of any actual association between Uniflows suggest the need for a better Biflow export method.

4.2. Record Adjacency Example

Assuming that each Uniflow record is described by the following simple template:

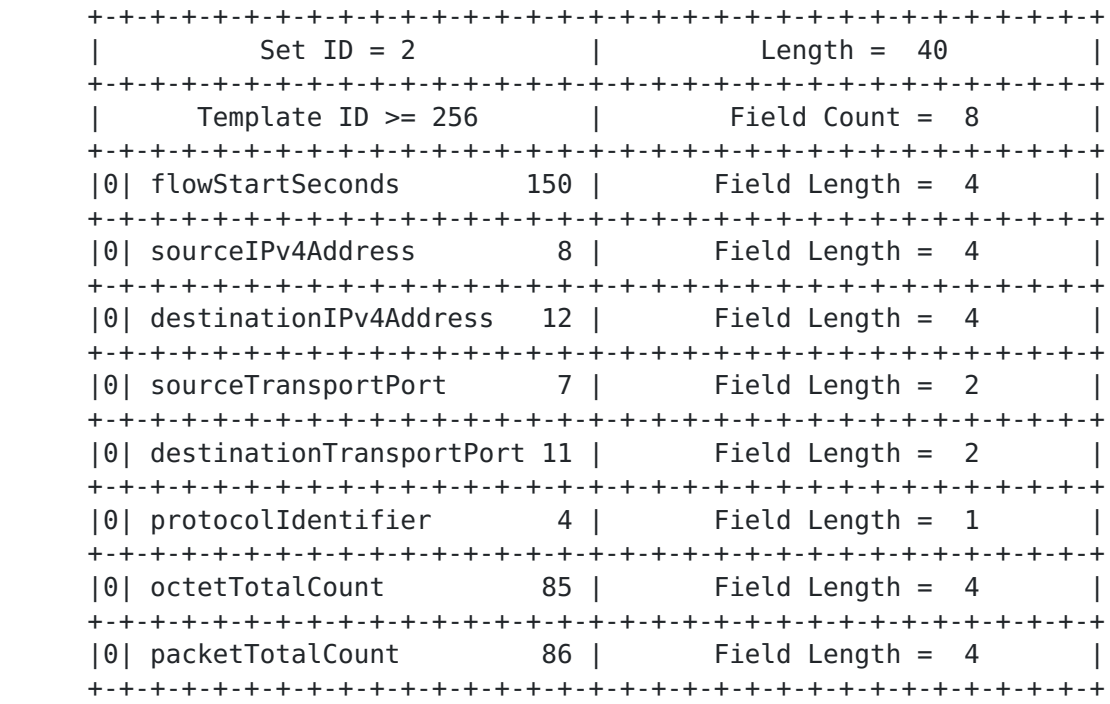

Figure 1: Record Adjacency Template Set Example

a two-record adjacent Biflow counting octets and packets in a typical HTTP transaction might look like the following:

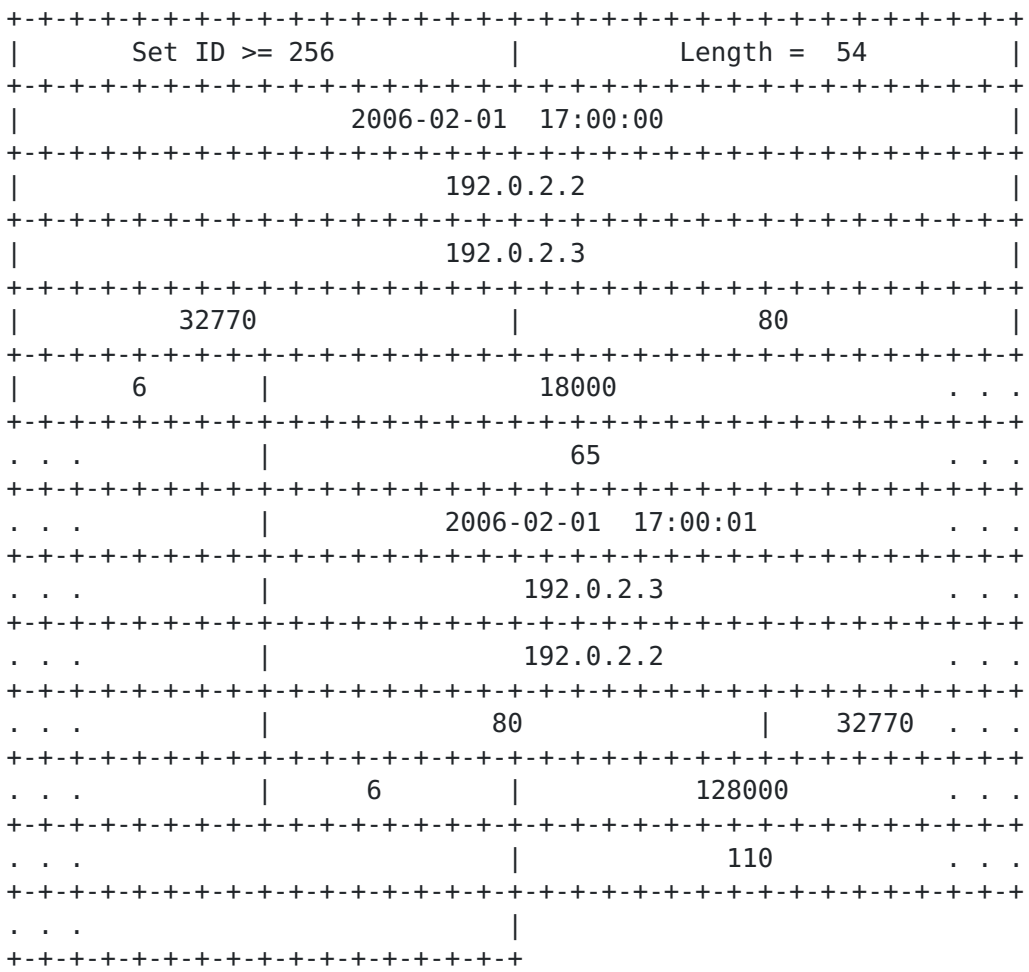

Figure 2: Record Adjacency Data Set Example

Notice that this trivial example duplicates 13 bytes of key information per Biflow.

4.3. Key-Value Separation using commonPropertiesId

The method described in Reducing Redundancy in IPFIX and PSAMP Reports [I-D.boschi-ipfix-reducing-redundancy] can be used to improve Biflow export bandwidth efficiency over the record adjacency method. This method separates the export of information common to a set of flow records from the export of the individual flow records, linking them with an index, the commonPropertiesID Information Element.

For Biflow export, the Flow Keys and any other fields common to the Biflow are exported within a data record defined by an Option Template, containing a commonPropertiesID Information Element as scope. This commonPropertiesID uniquely identifies that set of

properties and hence the Biflow itself. The individual Uniflow properties are then exported using one Flow Record per direction, each referencing the Biflow keys by the unique commonPopertiesID.

While this solution has the potential for significant bandwidth efficiency gains, and is widely applicable to a variety of bandwidth-reduction use cases, it is not yet an optimal method for Biflow export. First, the management of the commonPropertiesID for each biflow requires additional resources at both the Exporting Process and the Collecting Process. Second, instead of two records per Biflow as in the Record Adjacency method, the Reducing Redundancy method requires three. The set headers required to switch between common properties and specific properties templates also add slight bandwidth overhead.

5. Single Record Biflows

The most direct method for exporting Biflows using IPFIX is to use a single Flow Record to represent each Biflow. Each of these Flow Records will contain the Flow Key fields once, and both forward and reverse direction information elements for each non-key field. This proposal requires extending the IPFIX Information Model to provide for reverse value fields. This extension will cover most or all of the information model, creating a "reverse" Information Element counterpart to each presently defined "forward" Information Element, because any Information Element that may be a non-key field in a Biflow will require a counterpart.

The semantics of these single-record Biflows are outlined in section 3, above. Metering Process implementations using single-record biflow export SHOULD assign the forward and reverse direction such that the forward direction treats the flow initiator as source, to the best ability of the metering process to determine the flow initiator.

If a flow has no reverse direction -- that is, it is composed of a single Uniflow without another Uniflow in response -- it may only be represented as a single record Biflow if its only reverse value fields are counters. This is because the IPFIX Information Model makes no distinction between zeroes and null values. Exporting processes SHOULD switch to a template containing no reverse Information Elements when exporting flows without a reverse direction. Note that Flow Records containing no directional key fields (e.g., Flow Records representing aggregate octet counts by protocolIdentifier) cannot, by definition, have a reverse direction.

We have identified two possible methods for extending the information

model to include the required reverse Information Elements. The first and simpler method would be to add one new "reverse" Information Element to the information model for each Information Element subject to reversal in a single-record Biflow. The second and more convenient method would be to assign a special Private Enterprise Number (PEN) that creates a new reverse information element number space. Note that the choice between these methods impacts template representation of information elements only; the Data Records in which single-record Biflows are exported are identical with either assignment method. These methods are described in more detail below. The intent is to select one of these two methods; we present both here to promote discussion.

5.1. New Reverse Information Elements

As every Information Element in the information model that may appear as a non-key field in a Flow Record is subject to reversal, and the information model does not generally restrict Information Elements to key or non-key roles, single-record biflow export will require a great number of new reverse information elements. Only certain identifiers (flowId, templateId, and sourceId, from section 5.1 in the IPFIX Information Model [I-D.ietf-ipfix-info]), and metering and export process properties (section 5.2 in the IPFIX Information Model [I-D.ietf-ipfix-info]) are not subject to reversal.

The IPFIX Information Model has more than adequate number space for official information element expansion - 32768 IANA-managed information elements are available, of which less than 250 have been allocated or reserved. The addition of fewer than 250 new reverse elements would not place significant strain on available number space. However, the additional reverse information elements are not so much a discrete list of new Information Elements as a new dimension in the information model. This increases the effort required to manage the future extension of the IPFIX Information Model, adding a new task to this process: that of evaluating the reversibility of each new proposed Information Element and ensuring that every new Information Element that should have a reverse counterpart does. It also effectively reduces the available IANA-managed information element number space by half.

A complete list of reverse IEs required to implement this method will appear in a future revision of this draft if working group consensus moves toward this method.

5.2. Reverse Information Element Private Enterprise Number

Concerns have been raised in past deliberations of the IPFIX Working Group about adding information model dimensions; a real solution

would probably require protocol changes and hence is outside the scope of this draft. However, another more elegant short term solution may be possible by leveraging private enterprise information elements.

Instead of defining multiple new reverse information elements, it would also be possible to have IANA assign a single PEN to this draft, and to define that PEN to signify "IPFIX Reverse Information Element" (the Reverse PEN). This reverse PEN would serve as a "reverse direction flag" in the template; each Information Element number within this PEN space would be assigned to the reverse counterpart of the corresponding IANA-assigned public Information Element number. In other words, to generate a reverse information element in a template corresponding to a given forward information element, simply set the enterprise bit and define the Information Element within the Reverse PEN space, as in the figure below.

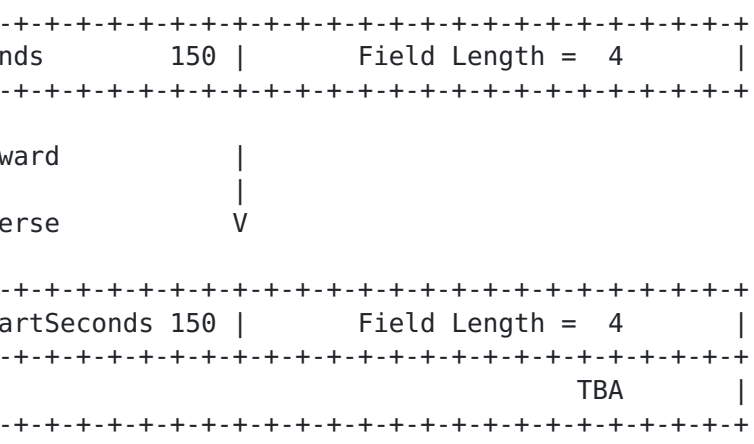

Figure 3: Example Mapping between Forward and Reverse IEs using Reverse PEN

This approach has the advantage of flexibility. It treats a new number space dimension explicitly as a dimension. New Information Elements can be added freely to the IANA-managed space without concern for whether a reverse element should also be added. Aside from the initial allocation of an enterprise number for this purpose, there is no additional maintenance overhead for supporting reverse information elements in the information model. The approach is also parallel with early proposals to add explicit information model dimensioning in a future revision of the IPFIX Protocol.

The primary drawback of this method is that it may slightly abuse the intent of the IANA Enterprise Number registry; this concern is detailed in the IANA Considerations section below.

5.3. Single Record Biflow Examples

The following template describes a simple Biflow record equivalent to the Record Adjacency example in section 4.2, using new information elements for the reverse-direction fields; these new information elements are denoted as TBA, as they have not been assigned by IANA.

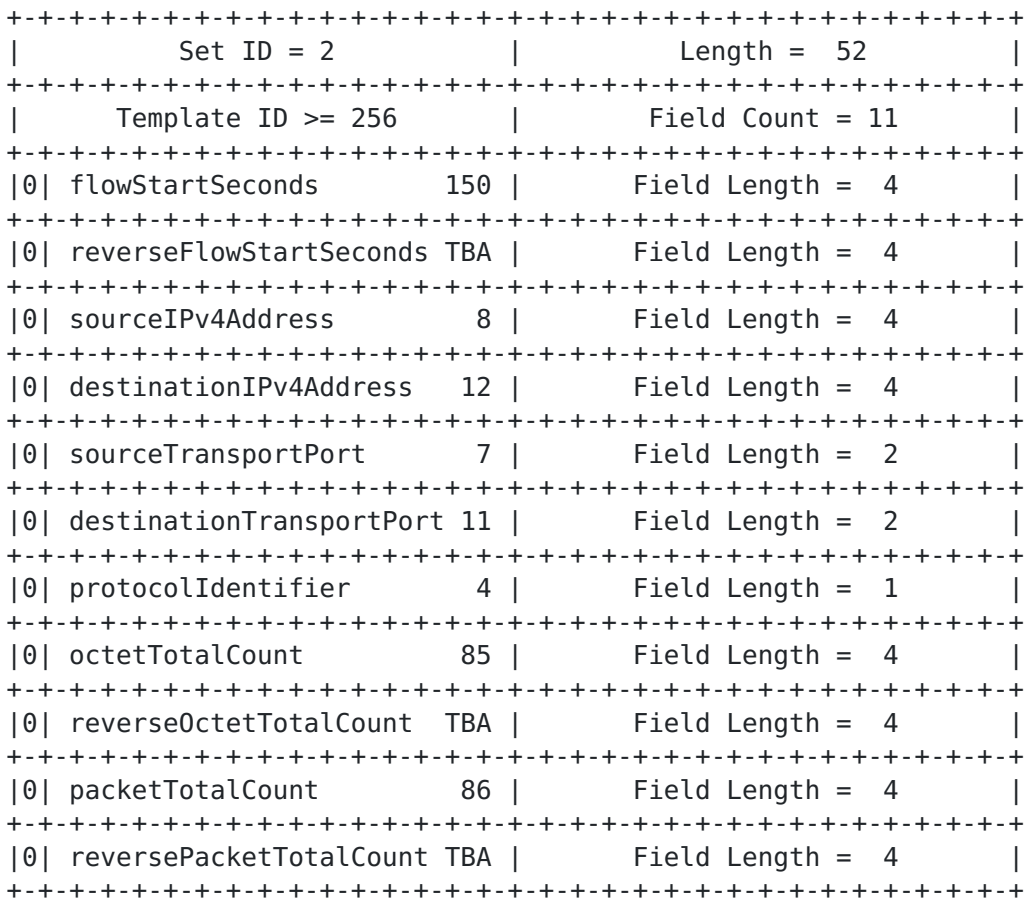

Figure 4: Single Record Biflow Template Set with New IEs

The following template describes the same data record as the previous one, but using the "IPFIX Reverse Information Element" PEN assigned for the purpose of differentiating forward from reverse information elements. This private enterprise number is denoted as TBA, as it has not yet been assigned by IANA.

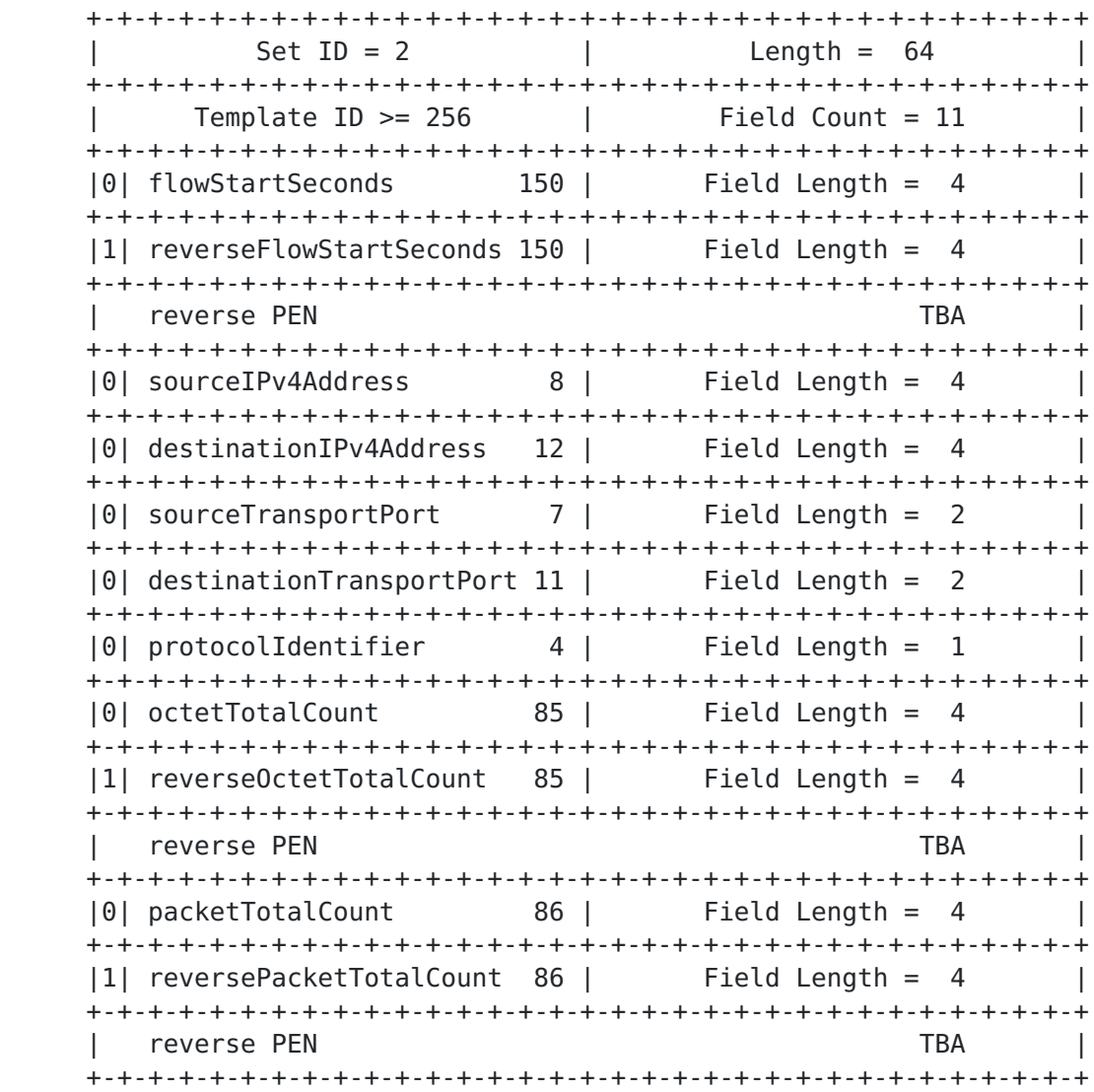

Figure 5: Single Record Biflow Template Set with Reverse PEN

Whether reverse information elements are assigned directly or implicitly by private enterprise number, both templates above describe the example single record Biflow below, which represents the same typical HTTP transaction as in example 4.2.

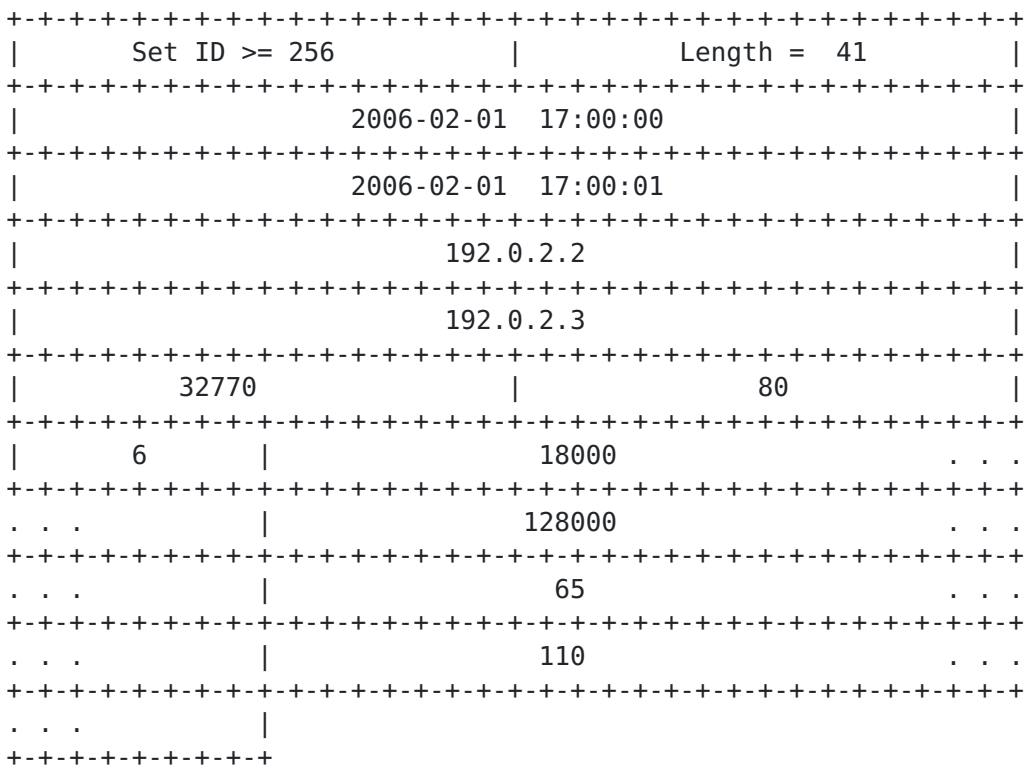

Figure 6: Single Record Biflow Data Set

6. IANA Considerations

As specified in the IPFIX Information Model [I-D.ietf-ipfix-info], IANA will create a new registry for IPFIX Information Element Numbers. New assignments for IPFIX Information Elements will be administered by IANA, on a First Come First Served basis, subject to Expert Review as per RFC 2434 [RFC2434], i.e. review by one of a group of expert designated by an IETF Operations and Management Area Director. The group of experts must double check the Information Element definitions against Information Elements already defined for completeness, accuracy, redundancy, and conformance to the naming conventions in the IPFIX Information Model [I-D.ietf-ipfix-info]. Those experts will initially be drawn from the Working Group Chairs and document editors of the IPFIX and PSAMP Working Groups, as noted in the IPFIX Information Model [I-D.ietf-ipfix-info].

This document proposes a set of new IPFIX Information Elements that extend those already defined in the information model; depending on the method selected to add these new Information Elements, either each Information Element or a single Reverse PEN must be assigned by IANA. Identifiers that have not yet been assigned by IANA are

denoted "TBA" (To Be Assigned) in this document.

If the consensus of the Working Group is to assign separate numbers to each reverse-direction Information Element as in Section 5.1, each of these reverse information elements will need to be assigned from the IANA IPFIX Information Element Number registry.

If the consensus of the Working Group is to assign a single Reverse PEN for reverse-direction Information Elements, this PEN will need to be assigned from the IANA Enterprise Number registry. The authors are in contact with IANA on this issue, and plan to have a PEN assigned to this draft, with the authors themselves as point of contact. A more definitive statement on the status of this Reverse PEN will appear in the IANA Considerations section of the next revision of this draft.

7. Security Considerations

The same security considerations as for the IPFIX Protocol [I-D.ietf-ipfix-protocol] apply.

8. Open Issues

We must select one policy for allocating reverse information elements. This single selection will appear in the next revision of this document [bht, eb].

9. Acknowledgements

We would like to thank Lutz Mark, Juergen Quittek, Andrew Johnson, Paul Aitken, Benoit Claise, and Carsten Schmoll for their contributions and comments. Special thanks to Michelle Cotton for her assistance in navigating the IANA process for enterprise number assignment.

10. References

10.1. Normative References

[I-D.ietf-ipfix-protocol]
Claise, B., "IPFIX Protocol Specification", draft-ietf-ipfix-protocol-21 (work in progress), April 2006.

[I-D.ietf-ipfix-info]
Quittek, J., "Information Model for IP Flow Information Export", draft-ietf-ipfix-info-11 (work in progress), September 2005.

[I-D.boschi-ipfix-reducing-redundancy]
Boschi, E. and L. Mark, "Reducing redundancy in IPFIX and PSAMP reports", draft-boschi-ipfix-reducing-redundancy-01 (work in progress), March 2006.

10.2. Informative References

[I-D.ietf-ipfix-reqs]
Quittek, J., "Requirements for IP Flow Information Export", draft-ietf-ipfix-reqs-16 (work in progress), June 2004.

[I-D.ietf-ipfix-as]
Zseby, T., "IPFIX Applicability", draft-ietf-ipfix-as-08 (work in progress), June 2006.

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

Authors' Addresses

Brian H. Trammell
CERT Network Situational Awareness
Software Engineering Institute
4500 Fifth Avenue
Pittsburgh, PA 15213
US

Phone: +1 412 268 9748
Email: bht@cert.org

Elisa Boschi
Hitachi Europe SAS
Immueble Le Theleme
1503 Route les Dolines
Valbonne 06560
France

Phone: +33 4 89874180
Email: elisa.boschi@hitachi-eu.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.