Internet-Draft
Document: draft-boschi-export-perpktinfo-01.txt
Expires: April 2006

E. Boschi
Hitachi Europe
L. Mark
Fraunhofer FOKUS

October 2005

Use of IPFIX for Export of Per-Packet Information
draft-boschi-export-perpktinfo-01.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on April 27, 2006.

Copyright Notice

Use of IPFIX for Export of Per-Packet Information

Abstract

This document describes the usage of the IP Flow Information Export (IPFIX) protocol for the case of exporting and processing per-packet information.
The main idea is to separate the export of the information about packets and flows those packets belong to, using two different records. The association between the records is kept using unique Flow or Template Identifiers.

Table of Contents

1. Introduction

In the scope of passive QoS Measurements, there is often the need to exchange and export measurement data in a finer granularity then per flows. One typical application is passive One-Way-Delay measurement; this draft takes it as example when demonstrating the need for information export on a per-packet basis.

The IPFIX protocol however, has been designed to export flow records. A possible approach to export packet records using IPFIX could be exporting flow records containing information about single packets. This method has been proposed by the PSAMP working group in [PSAMP-PROTO]. Exporting flow related information per-packet introduces a high degree of redundancy. This draft shows how packet information and flow information can be efficiently exported and related using IPFIX.

2. Terminology

Collecting Process
: The collecting process receives records of flow or packet information. The data is stored for later processing (by the calculation process)

Exporting Process
: The exporting processes send flow and packet records to the collecting processes. The records are generated by the measurement process.

Filtering
: Filtering selects a subset of packets by applying deterministic functions on parts of the packet content like header fields or parts of the payload. A filtering process

needs to process the packet (look at packet header and/or payload) in order to make the selection decision.

Measurement Device
: A measurement device has access to at least one observation point. It is hosting at least one measurement process and one export process.

Metering/Measurement Process
: The measurement process generates records of packet and flow information. Packets passing the observation point are captured, time stamped, filtered and classified. The measurement process calculates the packet Ids.

Passive One-Way-Delay Measurement
: Abbreviated: POWD Measurement

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

3. General Problem Statement

In [IPFIX-PROTO] the IPFIX working group has defined a protocol to transport measurement data containing flow information.

The main purpose of the protocol is to exchange information about IP traffic flows. In this scope a flow is defined by a set of key attributes (source/destination address, source/destination port, Layer3 Protocol Type, TOS/DSCP byte, interface of the flow exporting network element). As such, a flow is a collection of packets that share a set of common attributes.

However, for a number of metrics there is a need to export per-packet data.

A single packet could be considered a special case of a flow and thus, per-packet information could be exported using flow records. Doing this though would have consequences on the efficiency of the exporting procedure, as it would mean additional overhead. Packets belonging to the same flow share common attributes, i.e. source address, destination address, etc. Exporting these attributes on a per-packet basis, each time with a different packet ID, would be redundant information.

There are cases however, where it is desirable to keep flow information along with the per-packet information, that is, when

analyzing packet characteristics while observing flows. This document proposes a solution that reduces the overhead caused by the flow properties while keeping a link to flow information.

The proposed method does not need any changes to the IPFIX protocol.

4. Export Per-Packet Information

Figure 1 depicts three packets belonging to flow A and one packet belonging to flow B, respectively. It shows export records containing packet information plus flow information (source and destination address). Undoubtedly, the flow information introduces a huge amount of redundancy, as it is repeated for every packet in every record. Minimizing the redundancy is a common problem in relational data base design

and we apply here similar solutions to those proposed in that area.

In Figure 2 we separate flow from packet information. In order to maintain the relation between Packet Properties and Flow Properties we introduce indices (idxA and idxB) for the Flow Properties that are unique for all Flow Property entries. The purpose of the indices is to serve as "primary key" that identifies rows of the Flow Properties. More details about these indices will be given in section 5. The rows are then referenced by the Packet Properties by using the appropriate value for the flow identifier. The linkage of one packet and flow B (srcB, dstB, idxB) is explicitly drawn.

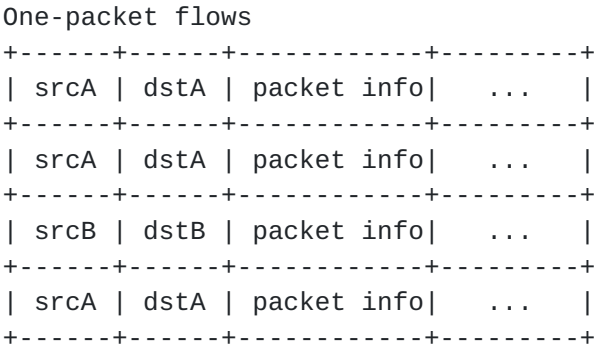

Figure 1: Flow and packet information represented in one-packet flows

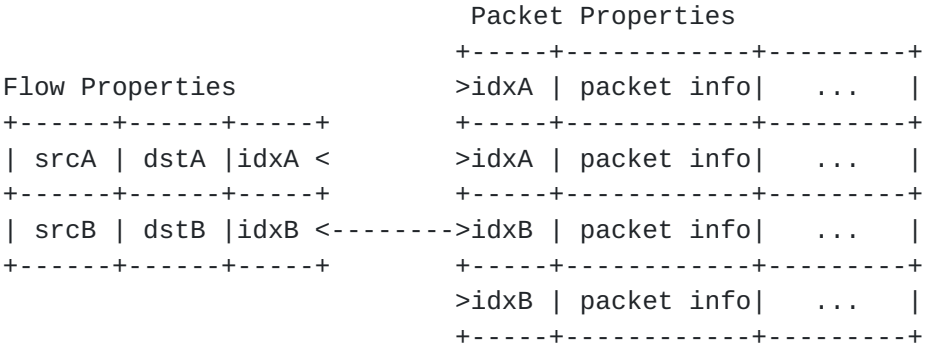

Figure 2: Flow information and packet information

The IPFIX protocol is template based like NetFlow version 9. For a complete description of features of IPFIX refer to [IPFIX-PROTO].
Templates define the structure of data to be exported, describing data fields together with their type and meaning. IPFIX specifies two types of records to export data: data

records and option data records. These records are defined via template records and option template records. To export per-packet-information we define two different templates: an option template for Flow Properties and a template for Packet Properties.

Figure 3 shows the relation between template and data sets for packet and flow properties. The Flow Properties option template defines the attributes for a flow; e.g. IP source and destination address and the flowID. The flowID is a unique identifier for a flow; this field allows packet records to reference flow attributes. Subsequent option data records of this template define the actual flows. The reference could be alternatively provided by the TemplateID, as explained in Section 5.

The format for the information related to single packets is defined in the Packet Properties template. This information is packet specific and normally not shared between many packets. Otherwise one would rather consider the information as flow related and therefore it needs not be exported in every record.

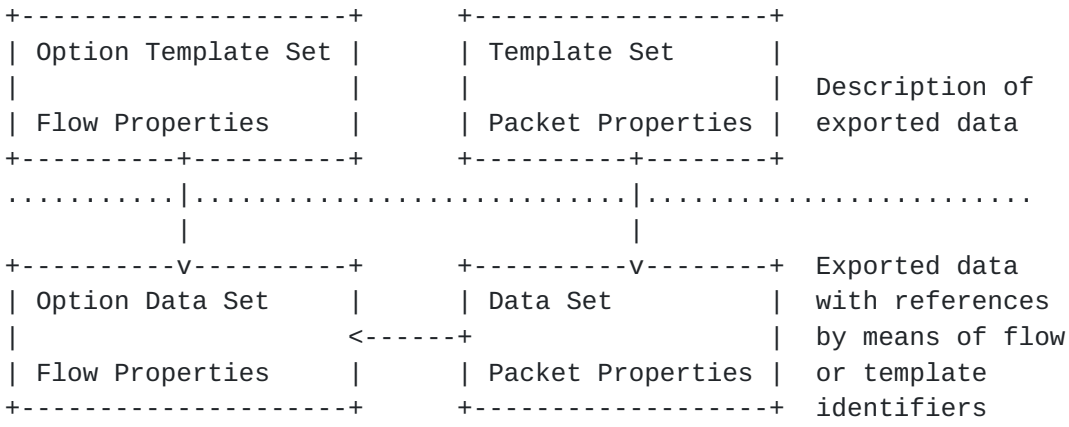

Figure 3: Template FlowSet and Data FlowSet dependencies

The Flow Properties option data records SHOULD be sent prior to the corresponding Packet Properties data records.

5. Using Scopes

Flow Properties are sent via IPFIX option records. IPFIX option records contain one or more scope fields. The Flow Properties record can contain the FlowID and/or the TemplateID as scope fields. There are three options:

1) Use FlowID as scope
   The flow properties are valid for all data records containing that flowID. This solution limits the number of different flows that can be exported at the same tame in the same observation domain to 2**32 (using 32 bits flowIDs)
2) Use FlowID and TemplateID as scope
   The flow properties are valid only for data records referring to the template specified by the TemplateIDand containing that flowID. This allows the export up to 2**32 flows per template. The solution is to be chosen when the number of flows to be exported is expected to be very high (and beyond the limit posed by solution 1)
3) Use TemplateID as scope
   The flow properties are valid for all data records of the specified template. In this case flowIDs are not needed but the exporting process requires a templateID per flow. In the

general case this solution is not scalable but can be suitable for certain applications (e.g. flow aggregation).

6. FlowID Management

The management of FlowIDs is very similar to the management of TemplateIDs described in [IPFIX-PROTO]. The Exporting Process assigns and maintains the FlowIDs for the exporter's Observation Domains. Like templateIDs, a FlowID MUST be unique per Observation Domain (source identifier in the IPFIX header). Different Observation Domains from the same exporter may use the same FlowID value to refer to different flows.
There are no constraints regarding the order of the Flow ID allocation. When limiting the scope to special templates, the flowIDs have to be unique per Observation Domain and template.

Using 32 bit flow IDs allows the export of 2**32 active flows in parallel. FlowIDs have a certain lifetime inside which they cannot be reused. After that time a FlowID can be assigned to another flow. FlowID whose lifetime has expired from longer SHOULD be preferred. The lifetime MUST be configurable.

The collecting process associates a lifetime with each flowID. The lifetime MUST be configurable. The mapping of data records to flow properties uses the most recent flow definition of the specified FlowID. If there is no flow definition of that FlowID or the lifetime of the flow definition has been expired, no mapping is possible. In this case the collecting process SHOULD log an error.

When IPFIX uses an unreliable transport protocol to export the option data records containing the flow properties and the flowIDs these records MUST be re-sent at regular intervals, whose frequency MUST be configurable.

When using a connection oriented transport protocol the flow properties have to be re-sent after a connection re-establishment in prior to the corresponding Packet Properties data records.

7. Example of Per-Packet Information Export

To demonstrate how to use IPFIX efficiently to export per-packet information, this section proposes how to use the IPFIX protocol for exporting flow information and per-packet information (in this case related to a long-lived flow) for OWD computation.

In order to acquire a One-Way path delay information, two measurement points with synchronized clocks must exist, one at each end of the path under examination. Both measurement points will capture packets, assign them timestamps and generate an identifier for a packet passing that point. Each measurement point will export its measurement data to a collecting process where the data are correlated based on the packet identifiers and timestamps and then the delay is calculated as a difference of two timestamps of a packet pair.

The templates that would be needed for exporting measurement data of this kind are illustrated in the figures below.
Figure 4 shows the option template containing the information concerning flows using the FlowID as scope.

In the Flow Properties template we export the following Information Elements:

- the source IPv4 Address, sourceIPv4Address [IPFIX-INFO], with a type of 8 and a length of 4 octets

- the destination IPv4 Address, destinationIPv4Address [IPFIX-INFO], with a type of 8 and a length of 4 octets

- the Class of Service field, ClassOfServiceIPv4 [IPFIX-INFO], with a type of 5 and a length of 1 octet

- source and destination ports, transportSourcePort and transportDestinationPort [IPFIX-INFO]with a type of 7 and 11 respectively, and a length of 2 octets each

The flow identifier, which is represented by the FlowId Information Element [IPFIX-INFO], is used as the Scope Field.

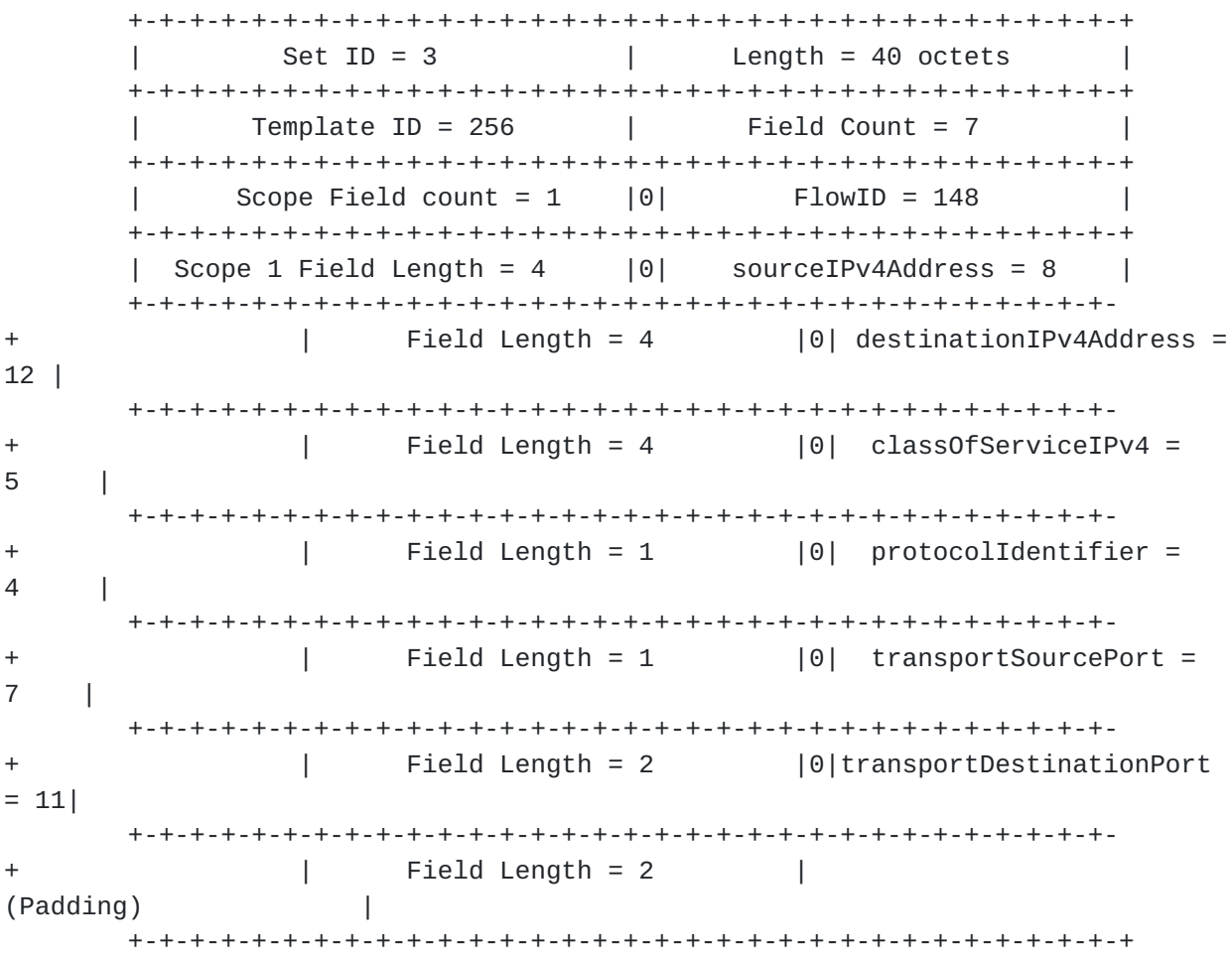

Figure 4: Example Flow Properties Template

For passive One-Way-Delay measurement, the Packet Properties template consists of at least Timestamp and Packet ID. Additionally, this template contains a flow identifier field. In packet records, the value of this field will contain one of the unique indices of the flow records exported before.

Figure 5 displays the template with the packet properties. In this example we export the following Information Elements:

- FlowID [IPFIX-INFO] with a length of 4 octets

- packetTimestamp, packetID, and packetLength. Since packetID, packetLength and flowID are not (yet) IETF-defined information elements, we export them as enterprise-specific IEs. The three IEs have respectively a type of 220, 221, and 222 and a length of 8, 4, an 4 octets.

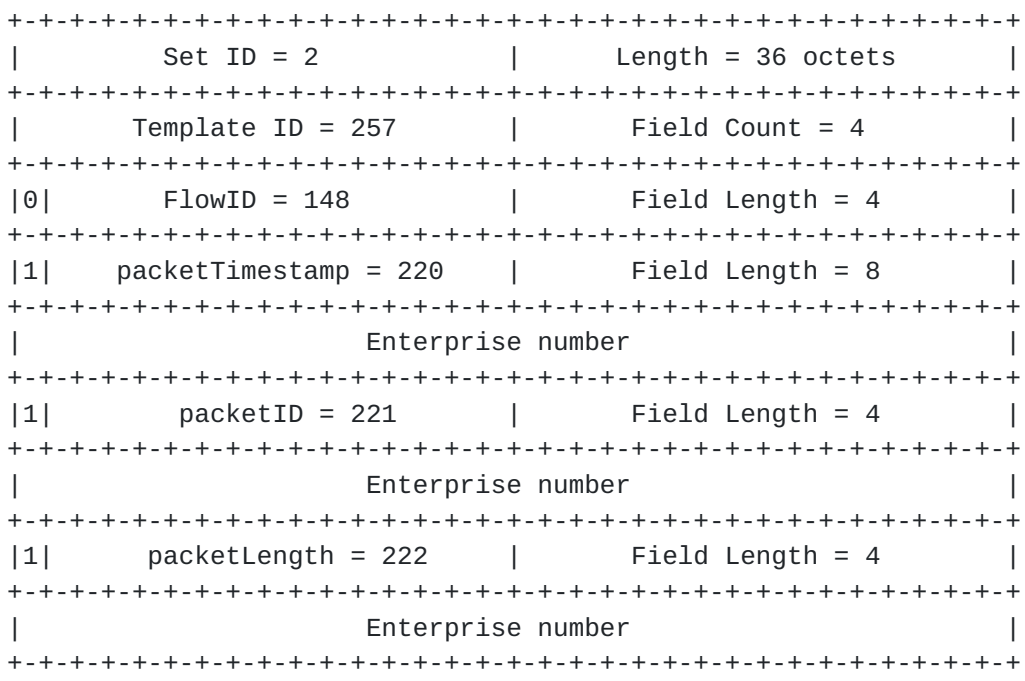

Figure 5: Example Packet Properties Template

The delay is derived by a calculation step: at the collection point packet records of two measurement points are gathered and correlated by means of the packet ID. The resulting delay data records are exported in a similar manner as the packet data have been. Especially, the linkage between delay data and flow information is represented with the discussed flow identifier fields. The OWD properties contain the Packet Pair ID (which is the packet ID matching that of the two contributing packet records), a timestamp (which is the timestamp of the packet passing the reference monitor point) in order to reconstruct a time series, the calculated delay value, and finally a flow identifier.

8. IPFIX for per-packet information export and PSAMP

In [PSAMP-PROTO] the PSAMP working group proposes to use IPFIX to export packet information from a PSAMP Exporting Process to a PSAMP Collecting Process. Even though no new version of the draft has been produced so far the solution seems to be accepted from the group.
While IPFIX is well suited for the purpose due to the good match between the IPFIX and PSAMP architectures and to the fact that IPFIX satisfies PSAMP requirements, the described approach has a high degree of redundancy. It proposes to treat packets as flows and export per-packet information using flow records.
We propose to use the solution described in this draft to efficiently export PSAMP packet information.

9. Export and evaluation considerations

The main advantage of this proposed method to export per-packet information is the reduced amount of measurement data that has to be transferred from the exporter to the collector. In addition there is less storage capacity needed at the collector side.

On the other hand there is some extra processing power needed on the exporter side to manage flow information and to assign packets to flows. The collector has to process records of two templates instead of just one but has to read and write less data. Additional effort is needed when post processing the measurement data, because now the correlation of flow and packet information is needed.

In the above example (see Figure 5) using IPFIX to export the measurement data for each received packet 30 bytes have to be transferred (sourceAddressV4=4, destinationAddressV4=4,

classOfServiceV4=1, protocolIdentifier=1, transportSourcePort=2, transportDestionationPort=2, packetTimestamp=8, packetID=4, packetLength=4). Disregarding the IPFIX protocol overhead a flow of 1000 packets produces 28000 bytes of measurement data. Using the proposed optimization each packet produces an export of only 20 bytes (packetTimestamp=8, packetID=4, packetLength=4, flowID=4). The export of the flow information produces 16 bytes (sourceAddressV4=4, destinationAddressV4=4, classOfServiceV4=1, protocolIdentifier=1, transportSourcePort=2, transportDestionationPort=2, flowID =4). For a flow of 1000 packet this sums up to 16016 bytes. This is a decrease of more than 40 percent.

10. IANA Consideration

This document does not imply any IANA action.

11. Security Considerations

For the proposed use of the IPFIX protocol for export of per-packet information the security considerations as for the IPFIX protocol apply.

12. References

12.1 Normative References

[IPFIX-PROTO] Benoit Claise et Al.: IPFIX Protocol Specification, IETF draft work in progress <draft-ietf-ipfix-protocol-19.txt>, September 2005

[IPFIX-INFO] J. Quittek, S.Bryant, B.Claise, J. Meyer: Information Model for IP Flow Information Export Internet-draft work in progress <draft-ietf-ipfix-info-11.txt>, September 2005

[PSAMP-PROTO] Benoit Claise: PSAMP Protocol Specification, Internet Draft <draft-ietf-psamp-protocol-01.txt>, February 2004

13. Author's Addresses

Elisa Boschi
Hitachi Europe SAS
Immeuble Le Theleme
1503 Route des Dolines
06560 Valbonne, France
Phone: +33 4 89874180
Email: elisa.boschi@hitachi-eu.com

Lutz Mark
Fraunhofer Institute for Open Communication Systems
Kaiserin-Augusta-Allee 31
10589 Berlin
Germany
Phone: +49-30-34 63 7306
Fax: +49-30-34 53 8306
Email: mark@fokus.fraunhofer.de

14. Intellectual Property Statement

license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be

required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

15. Copyright Statement

16. Disclaimer