IPFIX Working Group
Internet-Draft
Intended status: Informational
Expires: February 1, 2010

A. Kobayashi, Ed.
NTT PF Lab.
B. Claise, Ed.
Cisco Systems, Inc.
July 31, 2009

IPFIX Mediation: Problem Statement
draft-ietf-ipfix-mediators-problem-statement-05

Status of this Memo

Abstract

Flow-based measurement is a popular method for various network monitoring usages. The sharing of flow-based information for monitoring applications having different requirements raises some open issues in terms of measurement system scalability, flow-based measurement flexibility, and export reliability that IPFIX Mediation may help resolve. This document describes the IPFIX Mediation applicability examples, along with some problems that network administrators have been facing.

Table of Contents

1. Introduction

One advantage of Flow-based measurement results from easily offering the traffic monitoring of a huge amount of traffic. While the usage is applied to any networks and to multiple measurement applications, network administrators need to optimize the resource of metering devices and of multiple measurement applications. IP traffic growth and a wide variety of measurement application make the optimization further difficult. To achieve system optimization, an intermediate device can generally be applied to the system platform.

The IPFIX requirements defined in [RFC3917] mention examples of intermediate devices, such as IPFIX Proxies or Concentrators, there are no documents defining a generalized concept for such intermediate devices. This document addresses that issue by defining IPFIX Mediation, a generalized intermediate device concept for IPFIX, and examining in detail the motivations behind its application.

This document is structured as follows: section 2 describes the terminology used in this document, section 3 gives an IPFIX/PSAMP document overview, section 4 introduces general problems related to flow-based measurement, section 5 describes some applicability examples where IPFIX Mediations would be beneficial, and, finally, section 6 describes some problems an IPFIX Mediation implementation might face.

2. Terminology and Definitions

The terms in this section are in line with those in the IPFIX Protocol specifications [RFC5101] and the PSAMP specification document [RFC5476]. The terms Observation Point, Observation Domain, Flow, Flow Key, Flow Record, Data Record, Exporting Process, Exporter, IPFIX Device, Collecting Process, Collector, IPFIX Message, Metering Process, Transport Session, Information Element, and Template Withdrawal Message, are defined in the IPFIX protocol specifications [RFC5101]. The terms Packet Report, Sampling, Filtering, PSAMP Device, and Configured Selection Fraction are defined in the PSAMP specification document [RFC5476].

Furthermore, new terminology to be used in the context of IPFIX Mediation is defined in this section. All the words in these terms are started with a capital letter in this document.

In this document, we use the generic term "Data Records" for IPFIX Flow Records, PSAMP Packet Reports, and Data Records defined by Options Templates, unless an explicit distinction is required.

Original Exporter

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

IPFIX Mediation

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

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 transformation 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 on to an Exporting Process, IPFIX File Writer, or another Intermediate Process, 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 in IPFIX Messages via an Exporting Process. In the common case, an IPFIX Mediator receives a record stream from a Collecting Process but could also receive a record stream from data sources not encoded using IPFIX, e.g., in the case of conversion from NetFlow V9 protocol [RFC3954] to IPFIX protocol.

Specific types of IPFIX Mediators are defined below.

IPFIX Proxy

An IPFIX Proxy is an IPFIX Mediator that converts a record stream for the purpose of protocol conversion.

IPFIX Concentrator

An IPFIX Concentrator is an IPFIX Mediator that receives a record stream from one or more Exporters and performs correlation, aggregation, and/or modification.

IPFIX Distributor

An IPFIX Distributor is an IPFIX Mediator that receives a record stream from one or more Exporters and exports each record to one or more Collectors, deciding which Collector(s) to export each record depending on the decision of an Intermediate Process.

IPFIX Masquerading Proxy

An IPFIX Masquerading Proxy is an IPFIX Mediator that receives a record stream from one or more Exporters to screen out parts of records according to configured policies, in order to protect the privacy of the network's end users or to retain sensitive data of the exporting organization.

3. IPFIX/PSAMP Documents Overview

3.1. IPFIX Documents Overview

The IPFIX protocol [RFC5101] 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 [RFC5470], per the requirements defined in [RFC3917]. The IPFIX protocol [RFC5101] specifies how IPFIX Data Records and Templates are carried via a number of transport protocols from IPFIX Exporting Processes to IPFIX Collecting Processes. IPFIX has a formal description of IPFIX Information Elements, their names, types, and additional semantic information, as specified in [RFC5102]. [I-D.ietf-ipfix-mib] specifies the IPFIX Management Information Base. Finally, [RFC5472] describes what types 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. The storage of IPFIX Messages in a file is specified in [I-D.ietf-ipfix-file].

3.2. PSAMP Documents Overview

The framework for packet selection and reporting [RFC5474] enables network elements to select subsets of packets by statistical and other methods and to export a stream of reports on the selected packets to a Collector. The set of packet selection techniques (Sampling, Filtering, and Hashing) standardized by PSAMP are described in [RFC5475]. The PSAMP protocol [RFC5476] specifies the export of packet information from a PSAMP Exporting Process to a Collector. Like IPFIX, PSAMP has a formal description of its Information Elements, their names, types and additional semantic information. The PSAMP information model is defined in [RFC5477]. [I-D.ietf-psamp-mib] describes the PSAMP Management Information Base.

4. Problem Statement

Network administrators generally face the problems of measurement system scalability, flow-based measurement flexibility, and export reliability, even if some techniques, such as Sampling, Filtering, Data Records aggregation and export replication, have already been developed. The problems consist of optimizing the resources of the measurement system while fulfilling appropriate conditions: data accuracy, flow granularity, and export reliability. These conditions depend on two factors.

- measurement system capacity:
  This consists of the bandwidth of the management network, the storage capacity, and the performances of the collecting devices and exporting devices.

- application requirements:
  Different applications, such as traffic engineering, detecting traffic anomalies, and accounting, etc., impose different Flow Record granularities, and data accuracies.

The sustained growth of IP traffic has been overwhelming the measurement system capacities. Furthermore, a large variety of applications (e.g., QoS measurement, traffic engineering, security monitoring) and the deployment of measurement system in heterogeneous environments have been increasing the demand and complexity of IP traffic measurements.

4.1. Coping with IP Traffic Growth

Enterprise or service provider networks already have multiple 10 Gb/s links, their total traffic exceeding 100 Gb/s. In the near future, broadband users' traffic will increase by approximately 40% every year according to [TRAFGRW]. When operators monitor traffic of 500 Gb/s with a packet sampling rate of 1/1000, the amount of exported Flow Records from Exporters could exceed 50 kFlows/s. This value is beyond the ability of a single Collector.

To deal with this problem, current data reduction techniques (Sampling and Filtering in [RFC5475], and aggregation of measurement data) have been generally implemented on Exporters. Note that Sampling technique leads to potential loss of small Flows. With both Sampling and aggregation techniques, administrators might no longer be able to detect and investigate subtle traffic changes and anomalies as this requires detailed Flow information. With Filtering, only a subset of the Data Records are exported.

Considering the potential drawbacks of Sampling, Filtering, and Data

Records aggregation, there is a need for a large-scale collecting infrastructure that does not rely on data reduction techniques.

4.2. Coping with Multipurpose Traffic Measurement

Different monitoring applications impose different requirements on the monitoring infrastructure. Some of them require traffic monitoring at a Flow level while others need information about individual packets or just Flow aggregates.

To fulfill these divers requirements, an Exporter would need to perform various complex metering tasks in parallel, which is a problem due to limited resources. Hence, it can be advantageous to run the Exporter with a much simpler setup and to perform appropriate post-processing of the exported Data Records at a later stage.

4.3. Coping with Heterogeneous Environments

Network administrators use IPFIX Devices and PSAMP Devices from various vendors, various software versions, various device types (router, switch, or probe) in a single network domain. Even legacy flow export protocols are still deployed in current network. This heterogeneous environment leads to differences in Metering Process capabilities, Exporting Process capacity (export rate, cache memory, etc.), and data format. For example, probes and switches cannot retrieve some derived packet properties in [RFC5102] from a routing table.

To deal with this problem, the measurement system needs to mediate the differences. However, equipping all collecting devices with this absorption function is difficult.

4.4. Summary

In optimizing the resources of a measurement system, it is important to use traffic data reduction techniques as early as possible, e.g., at the Exporter. However, this implementation is made difficult by heterogeneous environment of exporting devices.

This implies that a new Mediation function is required in typical Exporter-Collector architectures. Based on some applicability examples, the next section shows the limitation of the typical Exporter-Collector architecture model and the IPFIX Mediation benefits.

5. Mediation Applicability Examples

5.1. Adjusting Flow Granularity

A set of common properties of simplest Flow type is a fixed 5-tuple of protocol, source and destination IP addresses, and source and destination port numbers. A shorter set of common properties, such as a triple, a double, or a single property, (for example network prefix, peering autonomous system number, or BGP Next-Hop fields), creates more aggregated Flow Records. This is especially useful for measuring traffic exchange in an entire network domain and for easily adjusting the performance of Exporters and Collectors.

Implementation analysis:

Implementations for this case depend on where Flow granularity is adjusted. More suitable implementations use configurable Metering Processes in Original Exporters. The cache in the Metering Process can specify its own set of common properties (Flow Keys) and extra fields. The Original Exporter thus creates directly aggregated Flow Records.

In the case where the Original Exporter contains a Metering Process that creates fixed tuple Flow Records (no ability to change the Flow Keys), or PSAMP Packet Reports, an IPFIX Concentrator can aggregate Data Records based on a new set of Flow Keys. Even in the case where the Original Exporter contains a Metering Process for which the Flow Keys can be configured, an IPFIX Concentrator can further aggregate the Flow Records.

5.2. Hierarchical Collecting Infrastructure

The increase of IPFIX Exporters, the increase of the traffic, and the variety of treatments expected to be performed over the Data Records is more and more difficult to handle within a single Collector. Hence to increase the collecting (e.g., the bandwidth capacity) and processing capacity, distributed Collectors must be deployed. As a possible approach, a hierarchical structure is useful for increasing the measurement systems capacity, both in export bandwidth capacity and in collecting capacity.

Implementation analysis:

To cope with the increase of IPFIX Exporters and traffic, one possible implementation uses IPFIX Concentrators to build a hierarchical collection system. To cope with the variety of treatments, one possible implementation uses IPFIX Distributors to build a distributed collection system. More specific cases are

described in section 5.9.

5.3. Correlation for Data Records

The correlation amongst Data Records or between Data Record and meta data provides new metrics or information, including the following.

- One-to-one correlation between Data Records
  - One way delay from the correlation of PSAMP Packet Reports from different Exporters along a specific path, packet inter-arrival time, etc.
  - Treatment from the correlation of Data Records with the common properties, observed at incoming/outgoing interfaces. Examples are the rate-limiting ratio, the compression ratio, the optimization ratio, etc.
- Correlation amongst Data Records

  Average/maximum/minimum values from correlating multiple Data Records. Examples are the average/maximum/minimum number of packets of the measured Flows, the average/maximum/minimum one way delay, the average/maximum/minimum number of lost packets, etc.
- Correlation between Data Record and other meta data

  Examples are some BGP attributes associated with Data Record by looking up the routing table.

Implementation analysis:

One possible implementation for this case uses an IPFIX Concentrator located between the Metering Processes and Exporting Processes on the Original Exporter, or alternatively a separate IPFIX Concentrator located between the Original Exporters and IPFIX Collectors.

5.4. Time Composition

Time composition is defined as the aggregation of consecutive Data Records with common properties. It leads to the same output as setting a longer active interval timer on Original Exporters with one advantage: the creation of new metrics such as average, maximum and minimum values from Flow Records with a shorter time interval enables administrators to keep track of changes that might have happened during the time interval.

Implementation analysis:

One possible implementation for this case uses an IPFIX Concentrator located between the Metering Processes and Exporting Processes on the Original Exporter, or alternatively a separate IPFIX Concentrator located between the Original Exporters and IPFIX Collectors.

5.5. Spatial Composition

Spatial composition is defined as the aggregation of Data Records in a set of Observation Points within an Observation Domain, across multiple Observation Domains from a single Exporter, or even across multiple Exporters. The spatial composition is divided into four types.

- Case 1: Spatial Composition within one Observation Domain

  For example, in the case where a link aggregation exists, Data Records metered at physical interfaces belonging to the same trunk can be merged.

- Case 2: Spatial Composition across Observation Domains, but within a single Exporter

  For example, in the case where a link aggregation exists, Data Records metered at physical interfaces belonging to a same trunk grouping beyond the line interface module can be merged.

- Case 3: Spatial Composition across Exporters

  Data Records metered within an administrative domain, such as the west area and east area of an ISP network, can be merged.

- Case 4: Spatial Composition across administrative domains

  Data Records metered across administrative domains, such as across different customer networks or different ISP networks, can be merged.

Implementation analysis:

One possible implementation for the cases 1 and 2 uses an IPFIX Concentrator located between the Metering Processes and Exporting Processes on the Original Exporter. A separate IPFIX Concentrator located between the Original Exporters and IPFIX Collector is a valid solution for the cases 1, 2, 3, and 4.

5.6. Data Record Anonymization

IPFIX exports across administrative domains can be used to measure traffic for wide-area traffic engineering or to analyze Internet traffic trends, as described in the spatial composition across administrative domains in the previous subsection.
In such a case, administrators need to adhere to privacy protection policies and prevent access to confidential traffic measurements by other people. Typically, anonymization techniques enables the provision of traffic data to other people without violating these policies.

Generally, anonymization modifies a data set to protect the identity of the people or entities described by the data set from being disclosed. It also attempts to preserve sets of network traffic properties useful for a given analysis while ensuring the data cannot be traced back to the specific networks, hosts, or users generating the traffic. For example, IP address anonymization is particularly important for avoiding the identification of the users, hosts, and routers. As another example, when ISP provides a traffic monitoring service to end customers by their own Exporters, even in case of exporting interface index fields, network administrators take care of anonymizing its fields to avoid disclosing the vulnerability.

Implementation analysis:

One possible implementation for this case uses an anonymization function at the Original Exporter. However, this increases the load on the Original Exporter. A more flexible implementation uses a separate IPFIX Masquerading Proxy between the Original Exporter and Collector.

5.7. Data Retention

Data retention refers to the storage of traffic data by service providers and commercial organizations. Legislative regulations often require that network operators retain both IP traffic data and call detail records, in wired and wireless networks, generated by end users while using a service provider's services. The traffic data is required for the purpose of the investigation, detection and prosecution of serious crime, if necessary. Data retention services examples are the following:

- Fixed telephony (includes fixed voice calls, voicemail, and conference and data calls)

- Mobile telephony (includes mobile voice calls, voicemail, conference and data calls, SMS, and MMS)

- Internet telephony (includes every multimedia session associated with IP multimedia services)
- Internet e-mail
- Internet access

Data retention for Internet access services in particular requires a measurement system with reliable export and huge storage as the data must be available for a long period of time, typically at least six months.

Implementation analysis:

Regarding export reliability requirement, the most suitable implementation uses the SCTP transport protocol between the Original Exporter and Collector. If an unreliable transport protocol such as UDP is used, a legacy exporting device exports Data Records to a nearby IPFIX Proxy through UDP, and then an IPFIX Proxy could reliably export them to the IPFIX Collector through SCTP. If an unreliable transport protocol such as UDP is used and if there is no IPFIX Proxy, the legacy exporting device should duplicate the exports to several Collectors to lower the probability of loosing Flow Records. However, it might result in network congestion, unless dedicated export links are used.

Regarding huge storage requirement, one possible implementation adopts a distributed measurement system to increase the storage capacity, by locating Collectors closer to the Exporters. In such a case, those Collectors would become IPFIX Mediators, re-exporting Data Records on demand to a centralized application.

5.8. IPFIX Export from a Branch Office

Generally, in large enterprise networks, Data Records from branch offices are gathered in a central office. However, in the long distance branch office case, the bandwidth for transport IPFIX is limited. Therefore, even if multiple Data Record types should be of interest to the Collector (e.g., IPFIX Flow Records in both directions, IPFIX Flow Records before and after WAN optimization techniques, performance metrics associated with the IPFIX Flow Records exported on regular interval, etc.), the export bandwidth limitation is an important factor to pay attention to.

Implementation analysis:

One possible implementation for this case uses an IPFIX Concentrator located in a branch office. The IPFIX Concentrator

would aggregate and correlate Data Records to cope with the export bandwidth limitation.

5.9. Distributing Data Records

Recently, several networks have shifted towards integrated networks, such as the pure IP and MPLS networks, which includes IPv4, IPv6, and VPN traffic. Data Record types (IPv4, IPv6, MPLS, and VPN) need to be analyzed separately and from different perspectives for different organizations. A single Collector handling all Data Record types might become a bottleneck in the collecting infrastructure. Data Records distributed based on their respective types can be exported to the appropriate Collector, resulting in the load distribution amongst multiple Collectors.

Implementation analysis:

One possible implementation for this case uses the replications of the IPFIX Message in an Original Exporter for multiple IPFIX Collectors. Each Collector then extracts the Data Record required by its own applications. However, the replication increases the load of the Exporting Process and the waste of the bandwidth between the Exporter and Collector.

A more sophisticated implementation uses an IPFIX Distributor located between the Metering Processes and Exporting Processes in an Original Exporter. The IPFIX Distributor determines to which Collector a Data Record is exported depending on certain field values. If a Original Exporter does not have IPFIX Distributor capability, it exports Data Records to a nearby separate IPFIX Distributor, and then the IPFIX Distributor could distribute them to the appropriate IPFIX Collectors.

For example, in the case of distributing a specific customer's Data Records, an IPFIX Distributor needs to identify the customer networks. The Route Distinguisher (RD), ingress interface, peering AS number, or BGP Next-Hop, or simply the network prefix may be evaluated to distinguish different customer networks. In the following figure, the IPFIX Distributor reroutes Data Records on the basis of the RD value. This system enables each customer's traffic to be inspected independently.

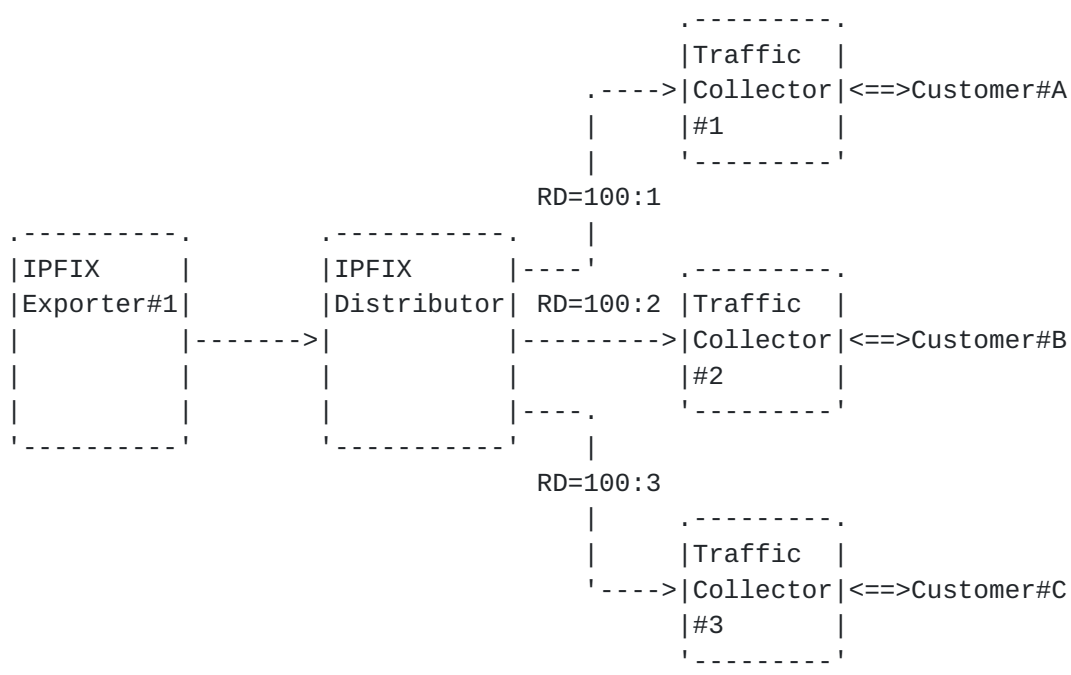

Figure A: Distributing Data Records to Collectors using IPFIX Distributor

5.10. Flow-based Sampling and Selection

Generally, the distribution of the number of packets per Flow seems to be heavy-tailed. Most types of Flow Records are likely to be small Flows consisting of a small number of packets. The measurement system is overwhelmed with a huge amount of these small Flows. If statistics information of small Flows is exported as merged data by applying a policy or threshold, the load on the Exporter is reduced. Furthermore, if the flow distribution is known, exporting only a subset of the Data Records might be sufficient.

Implementation analysis:

One possible implementation for this case uses an IPFIX Concentrator located between the Metering Processes and Exporting Processes on the Original Exporter, or alternatively a separate IPFIX Concentrator located between the Original Exporters and IPFIX Collectors. A set of IPFIX Mediation functions, such as filtering, selecting and aggregation is used in the IPFIX Concentrator.

5.11. Interoperability between Legacy Protocols and IPFIX

During the migration process from a legacy protocol such as NetFlow [RFC3954] to IPFIX, both NetFlow exporting devices and IPFIX Exporters are likely to coexist in the same network. Operators need to continue measuring the traffic data from legacy exporting devices, even after introducing IPFIX Collectors.

Implementation analysis:

One possible implementation for this case uses an IPFIX Proxy that converts a legacy protocol to IPFIX.

6. IPFIX Mediators Implementation Specific Problems

6.1. Loss of Original Exporter Information

Both the Exporter IP address indicated by the source IP address of the IPFIX Transport Session and the Observation Domain ID included in the IPFIX Message header are likely to be lost during IPFIX Mediation. In some cases, a IPFIX Masquerading Proxy might drop the information deliberately. In general, however, the Collector must recognize the origin of the measurement information, such as the IP address of the Original Exporter, the Observation Domain ID, or even the Observation Point ID. Note that, if an IPFIX Mediator can not communicate the Original Exporter IP address, then the IPFIX Collector will wrongly deduce that the IP address of the IPFIX Mediator is that of the Original Exporter.

In the following figure, a Collector can identify two IP addresses: 10.1.1.3 (IPFIX Mediator) and 10.1.1.2 (Exporter#2), respectively. The Collector, however, needs to somehow recognize both Exporter#1 and Exporter#2, which are the Original Exporters. The IPFIX Mediator must be able to notify the Collector about the IP address of the Original Exporter.

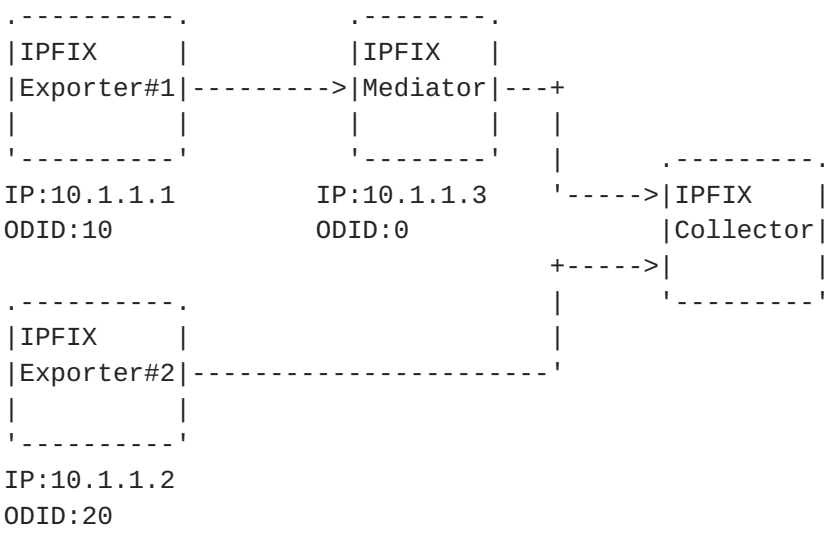

Figure B: Loss of Original Exporter Information.

6.2. Loss of Base Time Information

The Export Time field included in the IPFIX Message header represents a reference timestamp for Data Records. Some IPFIX Information Elements, described in [RFC5102], carry delta timestamps that indicate the time difference from the value of the Export Time field. If the Data Records include any delta time fields and the IPFIX

Mediator overwrites the Export Time field when sending IPFIX Messages, the delta time fields become meaningless and, because Collectors cannot recognize this situation, wrong time values are propagated.

6.3. Transport Sessions Management

Maintaining relationships between the incoming Transport Sessions and the outgoing ones depends on the Mediator's implementation. If an IPFIX Mediator relays multiple incoming Transport Sessions to a single outgoing Transport Session, and if the IPFIX Mediators shuts down its outgoing Transport Session, Data Records of the incoming Transport Sessions would not be relayed any more. In the case of resetting an incoming session, the behavior of the IPFIX Mediator needs to be specified.

6.4. Loss of Options Template Information

In some cases, depending on the implementation of the IPFIX Mediators, the information reported in the Data Records defined by Options Templates could also be lost. If, for example, the Sampling rate is not communicated from the Mediator to the Collector, the Collector would miscalculate the traffic volume. This might lead to crucial problems. Even if an IPFIX Mediator was to simply relay received Data Records defined by Options Templates, the values of its scope fields could become meaningless in the content of a different Transport Sessions. The minimal information to be communicated by an IPFIX Mediator must be specified.

6.5. Template ID Management

The Template ID is unique on the basis of the Transport Session and Observation Domain ID. If an IPFIX Mediation is not able to manage the relations amongst the Template IDs and the incoming Transport Session information, and if the Template ID is used in the Options Template scope, IPFIX Mediators would, for example, relay wrong values in the scope field and in the Template Withdrawal Message. The Collector would thus not be able to interpret the Template ID in the Template Withdrawal Message and in the Options Template scope. As a consequence, there is a risk that the Collector would then shut down the IPFIX Transport Session.

For example, an IPFIX Distributor must maintain the state of the incoming Transport Sessions in order to manage the Template ID on its outgoing Transport Session correctly. Even if the Exporter Transport Session re-initializes, the IPFIX Distributor must manage the association of Template IDs in specific Transport Session. In the following figure, the IPFIX Distributor exports three Templates (256,

257, and 258), received respectively from Exporter#3, Exporter#2, and Exporter#1. If the Exporter#1 re-initializes, and the Template ID value 258 is now replaced with 256, the IPFIX Distributor must correctly manage the new mapping of (incoming Transport Session, Template ID) and (outgoing Transport Session, Template ID) without shutting down its outgoing Transport Session.

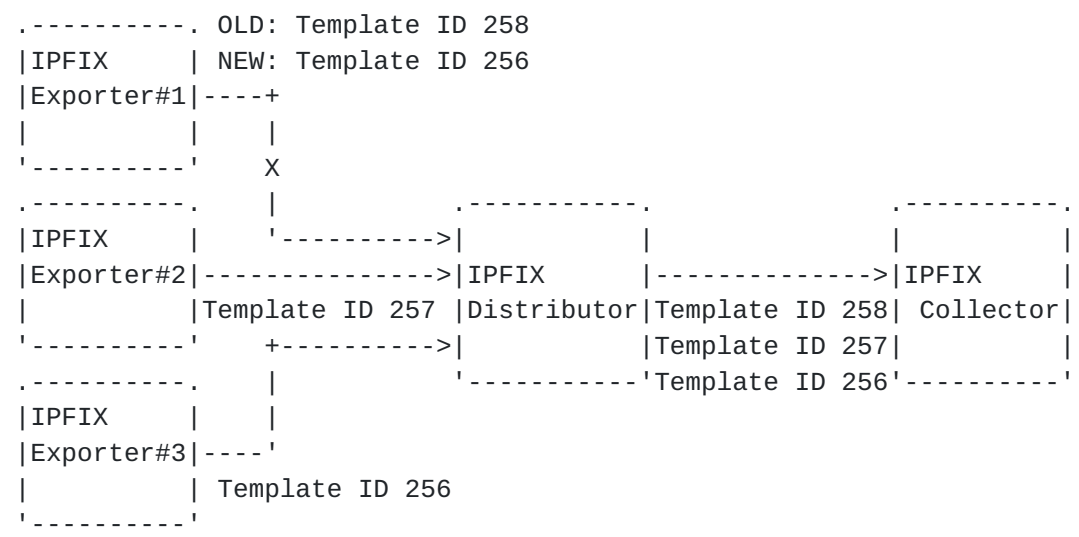

Figure C: Relaying from Multiple Transport Sessions to Single Transport Session.

6.6. Consideration for Network Topology

While IPFIX Mediation can be applied anywhere, caution should be taken as how to aggregate the counters, as there is a potential risk of double-counting. For example, if three Exporters export PSAMP Packet Reports related to the same Flow, the one-way delay can be calculated, while summing up the number of packets and bytes does not make sense. Alternatively, if three Exporters export Flow Records entering an administrative domain, then the sum of the packets and bytes is a valid operation. Therefore, the possible function to be applied to Flow Records must take into consideration the measurement topology. The information such as the network topology, or at least the Observation Point and measurement direction, is required for IPFIX Mediation.

6.7. Exporting the Function Item

In some case, the IPFIX Collector needs to recognize which specific function(s) the IPFIX Mediation has executed on the Data Records. The IPFIX Collector cannot distinguish between time composition, spatial composition, and Flow Key aggregation, if the IPFIX Mediator

does not export the applied function. Some parameters related to the function also would need to be exported. For example, in case of time composition, the active time of original Flow Records is required to interpret the minimum/maximum counter correctly. In case of spatial composition, spatial area information on which Data Records is aggregated is required.

6.8. Consideration for Aggregation

Whether the aggregation is based on time or spatial composition, caution should be taken on how to aggregate non-key fields in IPFIX Mediation. The IPFIX information model [RFC5102] specifies that the value of non-key fields, which are derived from fields of packets or from packet treatment and for which the value may change from packet to packet within a single Flow, is determined by the first packet observed for the corresponding Flow, unless the description of the Information Element explicitly specifies a different semantics.

However, this simple rule might not be appropriate when aggregating Flow Records which have different values in a non-key field. For example, if two Flows with identical Flow Key values are measured at different Observation Points, they may contain identical packets observed at different locations in the network and at different points in time. On their way from the first to the second Observation Point, some of the packet fields, such as the DSCP, may have changed. Hence, if the Information Element ipDiffServCodePoint is included as a non-key field, it can be useful to include the DSCP value observed at either the first or the second Observation Point in the resulting Flow Record, depending on the application.

Other potential solutions include: removing the Information Element ipDiffServCodePoint from the Data Record when re-exporting the aggregate Flow Record, changing the Information Element ipDiffServCodePoint from a non key-field to a Flow Key when re-exporting the aggregated Flow Record, or assigning a non valid value for the Information Element to express to the Collector that this Information Element is meaningless.

Furthermore, rules must be specify on how to aggregate the new Configured Selection Fraction an IPFIX Mediator should report when aggregating IPFIX Flow Records with different sampling rates. Finally, special care must be taken when aggregating Flow Records resulting from different Sampling techniques such as Systematic Count-Based Sampling and Random n-out-of-N Sampling for example.

7. Summary and Conclusion

This document described the problems that network administrators have been facing, the applicability of IPFIX Mediation to these problems, and the problems related to the implementation of IPFIX Mediators. To assist the operations of the Exporters and Collectors, there are various IPFIX Mediations from which the administrators may select. Examples of the applicability of IPFIX Mediation are as follows.

- Regarding large-scale measurement system, IPFIX Concentrators or IPFIX Distributors help to achieve traffic analysis with high data accuracy and fine flow granularity even as IP traffic grows. As IPFIX Mediation capabilities, Flow sampling, aggregation, and composition are effective.

- Regarding data retention, IPFIX Mediators enhance the export reliability, and the storage of the measurement system.

- Regarding the distribution of Data Records, IPFIX Distributors help to achieve multipurpose traffic analysis for different organizations, or help to achieve respective traffic analysis based on Data Record types(IPv4, IPv6, MPLS, and VPN).

- Regarding the IPFIX export across domains, IPFIX Masquerading Proxies help administrators to anonymize or filter Data Records, preventing privacy violations.

- Regarding interoperability, IPFIX Proxies provide interoperability between legacy protocols and IPFIX, even during the migration period to IPFIX.

As a result, the IPFIX Mediation benefits become apparent. However, there are still some open issues with the use of IPFIX Mediators.

- Both Observation Point and IPFIX Message header information, such as the Exporter IP address, Observation Domain ID, and Export Time field, might be lost. This data should therefore be communicated between the Original Exporter and Collector via the IPFIX Mediator.

- IPFIX Mediators are required to manage Transport Sessions, Template IDs, and Observation Domain IDs. Otherwise, anomalous IPFIX Messages could be created.

- Data Records defined by Options Templates, such as those reporting the Sampling rate and Sampling algorithm used, might be lost during IPFIX Mediation. If a Collector is not informed of current Sampling rates, traffic information might become worthless.

These problems stem from the fact that no standards regarding IPFIX Mediation have been set. In particular, the minimum set of information that should be communicated between Original Exporters and Collectors, the management between different IPFIX Transport Sessions, and the internal components of IPFIX Mediators should be standardized.

8. Security Considerations

A flow-based measurement system must prevent potential security threats: the disclosure of confidential traffic data, injection of incorrect data, and unauthorized access to traffic data. These security threats of the IPFIX protocol are covered by the security considerations section in [RFC5101] and are still valid for IPFIX Mediators.

And a measurement system must also prevent the following security threats related to IPFIX Mediation:

- Attacks against IPFIX Mediator

  IPFIX Mediators can be considered as a prime target for attacks, as an alternative to IPFIX Exporters and Collectors. IPFIX Proxies or Masquerading Proxies need to prevent unauthorized access or denial-of-service (DoS) attacks from untrusted public networks.

- Man-in-the-middle attack by untrusted IPFIX Mediator

  The Exporter-Mediator-Collector structure model would increase the risk of the man-in-the-middle attack.

- Configuration on IPFIX Mediation

  In the case of IPFIX Distributors and IPFIX Masquerading Proxies, an accidental misconfiguration and unauthorized access to configuration data could lead to the crucial problem of disclosure of confidential traffic data.

9. IANA Considerations

This document has no actions for IANA.

10. Acknowledgements

We would like to thank the following persons: Gerhard Muenz for the thorough detail review and significant contribution regarding the improvement of whole sections; Keisuke Ishibashi for contribution during the initial phases of the document; Brian Trammel for contribution regarding the improvement of terminologies section; Nevil Brownlee, Juergen Schoenwaelder, Motonori Shindo for the technical reviews and feedback.

11. References

11.1. Normative References

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

[RFC5476] Claise, B., "Packet Sampling (PSAMP) Protocol Specifications", March 2009.

11.2. Informative References

[I-D.ietf-ipfix-file] Trammell, B., Boschi, E., Mark, L., Zseby, T., and A. Wagner, "Specification of the IPFIX File Format", draft-ietf-ipfix-file-04 (work in progress) , July 2009.

[I-D.ietf-ipfix-mib] Dietz, T., Claise, B., and A. Kobayashi, "Definitions of Managed Objects for IP Flow Information Export", draft-ietf-ipfix-mib-06 (work in progress) , March 2009.

[I-D.ietf-psamp-mib] Dietz, T. and B. Claise, "Definitions of Managed Objects for Packet Sampling", draft-ietf-psamp-mib-06 (work in progress) , June 2006.

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

[RFC3954] Claise, B., "Cisco Systems NetFlow Services Export Version 9", October 2004.

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

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

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

[RFC5474] Duffield, N., "A Framework for Packet Selection and Reporting", March 2009.

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

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

[TRAFGRW] Cho, K., Fukuda, K., Esaki, H., and A. Kato, "The Impact and Implications of the Growth in Residential User-to-User Traffic", SIGCOMM2006, pp. 207-218, Pisa, Italy, September 2006. .

Authors' Addresses

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

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

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

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

Phone: +81-422-59-3978
Email: nishida.haruhiko@lab.ntt.co.jp

Christoph Sommer
University of Erlangen-Nuremberg
Department of Computer Science 7
Martensstr. 3
Erlangen 91058
Germany

Phone: +49 9131 85-27993
Email: christoph.sommer@informatik.uni-erlangen.de
URI: http://www7.informatik.uni-erlangen.de/~sommer/

Falko Dressler
University of Erlangen-Nuremberg
Department of Computer Science 7
Martensstr. 3
Erlangen 91058
Germany

Phone: +49 9131 85-27914
Email: dressler@informatik.uni-erlangen.de
URI: http://www7.informatik.uni-erlangen.de/~dressler/

Stephan Emile
France Telecom
2 avenue Pierre Marzin
Lannion, F-22307

Fax: +33 2 96 05 18 52
Email: emile.stephan@orange-ftgroup.com