J Netw Syst Manage (2011) 19:298–304
DOI 10.1007/s10922-010-9163-z

REPORT

Report of the Second Workshop on the Usage of NetFlow/IPFIX in Network Management

Idilio Drago · Rafael R. R. Barbosa · Ramin Sadre · Aiko Pras · Jürgen Schönwälder

Published online: 25 March 2010

Abstract Following the success of the *First Workshop on the Usage of NetFlow/IPFIX* (Pras et al. in J Netw Syst Manag 17(4), 2009) in 2008, the European EMANICS Network of Excellence organized a second workshop in October 2009, held at Jacobs University Bremen. This report summarizes the workshop and presents its main conclusions.

Keywords NetFlow · IPFIX · EMANICS

1 Introduction

NetFlow is a protocol developed by Cisco Systems to monitor Internet traffic that flows through network elements [2]. A flow is a unidirectional stream of packets that pass through a network element and share a common set of attributes [3, 4]. In early versions of NetFlow, a flow was defined by a fixed set of seven fields: source and destination IP addresses, source and destination port numbers, protocol type, type of service and logical interface (ifIndex). Since NetFlow version 9, flow definitions

I. Drago (✉) · R. R. R. Barbosa · R. Sadre · A. Pras
University of Twente, Enschede, The Netherlands
e-mail: i.drago@utwente.nl

R. R. R. Barbosa
e-mail: r.r.r.barbosa@utwente.nl

R. Sadre
e-mail: r.sadre@utwente.nl

A. Pras
e-mail: a.pras@utwente.nl

J. Schönwälder
Jacobs University Bremen, Bremen, Germany
e-mail: j.schoenwaelder@jacobs-university.de

have become flexible and can be specified via templates. IPFIX (IP Flow Information Export) is an effort by the IETF (Internet Engineering Task Force) to create standard protocols to collect and export IP flows. IPFIX has been under discussion in the IETF since 2001, and in 2004 NetFlow version 9 was chosen to be the basis for the IPFIX specification [10]. Several improvements to NetFlow version 9 have already been added to IPFIX, including enterprise-defined fields [1] and bidirectional flow-exporting functions [16].

For the second consecutive year, the European EMANICS Network of Excellence organized an one-day workshop on the use of NetFlow/IPFIX for network management. The aim of this workshop series is to provide a forum in which researchers, operators and device manufacturers can discuss the latest development and exchange practical experiences in this area. As reported earlier [13] in this Journal, the first edition of the workshop discussed technologies to capture and analyze flow data, and the effects of sampling and aggregation techniques on the accuracy of results. This second edition was organized into three sessions and addressed the following questions:

- What is the current stage of IPFIX standardization and what are the challenges in adopting the protocol?
- What are potential new applications for NetFlow/IPFIX?
- How is NetFlow/IPFIX used in practice?

This report presents a summary of the various presentations at the workshop along with their main conclusions.

2 IPFIX Standardization and Open Issues

In the first session, the status of IPFIX standardization was discussed, together with some other problems that have impact on the practical implementation of the protocol. The session was opened with a presentation on the status of IPFIX in the IETF. As in the previous workshop, Benoit Claise (Cisco Systems) gave an overview about the history of IPFIX and the main differences between IPFIX and NetFlow Version 9. In addition, he compared IPFIX with PSAMP [5], showing that these protocols are complementary. He finished by summarizing the current work in the IETF:

- The *IPFIX File Format* specification, which defines a format for storing IPFIX data, has been completed.
- There are three network management related drafts under discussion: (1) *Definitions of Management Objects for IP Flow Information Export*, (2) *Definitions of Managed Objects for Package Sampling*, and (3) *Configuration Data Model for IPFIX and PSAMP*.
- There is still work in progress related to *IPFIX Structured Data*, *Mediation Function* and *IPFIX Export per SCTP Stream*.
- New items have been added to the charter: *Flow Anonymization*, *Flow Selection* and *IPFIX Benchmarking*.

More information about the current work in the IETF can be found in the IPFIX status pages [15].

Carsten Schmoll (Fraunhofer FOKUS) proposed a solution for making transmission of IPFIX data more secure. Network flow data must be treated as confidential, since they contain information that can, for example, be misused during attacks. His solution addresses two major threats:

- Anonymity disclosure: NetFlow/IPFIX records contain information about active flows, addresses of involved nodes, and traffic patterns in the network. Such information can be used by attackers to identify users' behavior and reveal details about the network infrastructure, easing attacks against other network elements.
- Attacks against the measurement system: applications that depend on network flow data can be affected if the measurement structure is damaged. For example, unprotected collectors are vulnerable to *flooding attacks*, which can disrupt accounting systems.

Schmoll proposed to encrypt exported IPFIX data and to decrypt them only when strictly necessary. His solution uses a different encryption key for each collector device, allowing exporting devices to decide which collectors will decrypt which portion of the data. All communication for key exchange is protected by standard TLS (Transport Layer Security), and all standard security measures–such as protection by firewalls and access control policies–should also be in place. However, a comprehensive evaluation of the effectiveness of his approach is still to be performed.

Cristian Morariu's presentation targeted the bottlenecks in handling NetFlow/IPFIX data. Since NetFlow/IPFIX meters are often used in high-speed network environments, the infrastructure to transport and to process those data must be designed to support heavy workloads. Bottlenecks can occur if, for example, NetFlow/IPFIX data arrive at a collector device at rates higher than the writing speed of the storage hardware, if the bandwidth available in the network is not sufficient, or if the time required to process a NetFlow/IPFIX record is longer than the inter-arrival time of such records. These bottlenecks are normally addressed at the metering point, by sampling packets or flows before exporting any data. However, some applications require highly accurate measurements, and sampling approaches may have a negative impact on that.

Morariu proposed a new architecture, suitable for situations in which sampling is not acceptable. His solution, based on the Kademlia distributed hash table [11], aims to increase the number of flows that can be processed, by distributing the workload across several network nodes. Furthermore, his solution is more robust, since peer-to-peer networks provide redundancy and avoid single points of failure. Although a prototypical implementation already exists, further analysis is needed to ensure its feasibility.

3 New Applications for IPFIX

In the second session, new applications for NetFlow/IPFIX were discussed. The session was opened by Nikolay Melnikov (Jacobs University Bremen), who is researching methods for the identification of users, on the basis of network flow analysis. Each user has his or her own individual browsing style that leads to specific patterns of flow data. Melnikov's goal is to identify users automatically, comparing unknown NetFlow/IPFIX records to known examples of users' flow signatures. Even though this research is still in its early stages, first results are promising: using cross-correlation as proximity measure, the duration distribution of flows from four volunteer users could be differentiated. The next steps of his research will include the creation of more features from original data and the selection of the most discriminative features. However, all those tasks depend on the availability of data from new volunteers.

Tim Kleefass (SWITCH/University of Stuttgart) presented a new application for NetFlow/IPFIX data that can help network operators to identify remote connectivity problems. Network operators need to know about problems before their customers notice them, even when those problems are not directly related to their services or products. As an example, all traffic going to YouTube was lost for several minutes in 2008, due to a wrong configuration created by a telecommunication company in Pakistan. Users all around the world were affected, but without knowing exactly who caused the problem. Kleefass proposed to compare the number of flows passing edge routers within a network to and from a remote location. Using those numbers, he updates a connectivity matrix every 5 min. In the event of remote connectivity problems, those numbers will get unbalanced, indicating an abnormal situation. Some applications that have unbalanced traffic in normal situations, such as Skype, must be filtered out to avoid false alarms. Using logs from the Swiss Research and Educational Network (SWITCHlan), Kleefass showed that this approach successfully detects outages at remote Internet locations.

Jochen Kögel (Universtität Stuttgart) closed this session, explaining how NetFlow/IPFIX data can be used to extract performance metrics in enterprise networks. As described in RFC3917 [14] and RFC5472 [17], quality of service monitoring is one of the target applications of IPFIX, but a comprehensive research into what can be extracted from IPFIX data is still lacking. Kögel presented methods to extract RTT (Round Trip Time) and one-way delay figures from flow data. The RTT can be extracted from a single router if all traffic (in both directions) is routed through it. In that case, the RTT will be the difference between the starting times of a pair of related flows (request/response flows). Delay estimation, on the other hand, requires information from several routers on the path between the end nodes. The delay can be calculated based on the starting time of the same flow, but collected at different points (assuming that all devices have synchronized clocks). Experiments showed good results when comparing his calculations with active measurements.

4 NetFlow/IPFIX in Practice

The remaining presentations focused on NetFlow/IPFIX usage in practice, and in particular on (1) real-world experiences of a NetFlow tool-vendor, (2) a visualization tool for NetFlow data, and (3) an analysis of the bandwidth behavior of network flows.

Andreas Bourges (IsarNet) reported some experiences of using NetFlow from a tool-vendor's perspective. In the first part of his presentation, the IsarFlow monitoring tool [9] was described. Its capabilities include bandwidth monitoring, application discovery, and anomaly detection. Emphasis was placed on the approach implemented to deal with the large quantity of collected data. The second part of his presentation focused on how IsarNet's customers are using NetFlow. In their experience, customers typically use the tool to find out:

- Which protocols are causing the main load on a given link?
- Which IP addresses are causing the main load on a given link?
- Which TCP flags and port numbers are used between end nodes?
- What settings are used by a given protocol?

Most of IsarNet's customers are still using NetFlow version 5. However, the demand for version 9 is increasing. Bourges argued that Flexible NetFlow has several promising features, but version 5 is enough to meet the requirements of most customers. Moreover, the deployment of Flexible NetFlow faces a dilemma: tool vendors require that customers' hardware/software infrastructure supports this new version before they start providing products for it. However, customers are not willing to upgrade their infrastructure if there is no application that supports Flexible NetFlow.

Rick Hofstede (University of Twente) proposed a novel way to visualize flow data, based on geographic information. He implemented a plug-in for NfSen [12] that interfaces with IP2Location [8] and the Google Maps API [7]. The result is a web application that displays geographic information about the network traffic. His tool provides zoom levels that allow discrimination between various aspects of network information, while retaining an intuitive interface.

Ramin Sadre (University of Twente) presented his analysis of bandwidth behavior of large flows. This study was motivated by the interest of network managers in such flows. For example, if packet switching/optical networks are available, large flows can be moved from the IP to the optical level, in an attempt to increase network efficiency. The objective of his research is to go beyond the aggregated information provided by NetFlow records, such as start/end time and transmitted bytes, and to analyze how flows behave during their lifetime. Two research questions were presented:

- Do large flows, in general, have a constant throughput?
- What do we know about the (overall) throughput of a flow, after observing it for some time (e.g. 5 min)?

The study was based on data collected at the University of Twente during 2007. At this stage of his research, only those flows with more than 100 MB were considered. Sadre concluded that most flows in his data set have a constant throughput, but there are some large deviations. Furthermore, more precise throughput estimation is achieved as the observing time of a flow gets longer.

5 Conclusions

Like the first workshop, the *2nd EMANICS Workshop on NetFlow/IPFIX Usage* was a success. All presentations generated highly interactive discussions, resulting in valuable feedback for researchers. More information about the second workshop, including slides and contact information of all presenters, can be found on the EMANICS website [6]. A third workshop will be organized in 2010, probably in conjunction with the *78th IETF Meeting*, which will take place on July 25–30, in Maastricht, The Netherlands. More information about the next workshop can be obtained from the authors of this report.

Acknowledgments The 2nd Workshop on NetFlow/IPFIX Usage was partly supported by the IST Network of Excellence EMANICS, funded by the European Union under contract number FP6-2004-IST-026854-NoE. This report was partly supported by the Dutch Ministry of Economic Affairs via its agency SenterNovem, in the context of the IOP GenCom project Service Optimization and Quality (SeQual).

References

1. Boschi, E., Trammell, B., Mark, L., Zseby, T.: Exporting type information for IP flow information export (IPFIX) information elements (2009). http://www.ietf.org/rfc/rfc5610.txt
2. Cisco Systems: NetFlow services solution guide (2007). http://www.cisco.com/en/US/docs/ios/solutions_docs/netflow/nfwhite.html
3. Claise, B.: Cisco systems NetFlow services export version 9 (2004). http://www.ietf.org/rfc/rfc3954.txt
4. Claise, B.: Specification of the IP flow information export (IPFIX) protocol for the exchange of IP traffic flow information (2008). http://www.ietf.org/rfc/rfc5101.txt
5. Claise, B., Quittek, J., Johnson, A.: Packet sampling (PSAMP) protocol specifications (2009). http://www.ietf.org/rfc/rfc5476.txt
6. EMANICS: Emanics Network of Excellence (2010). http://www.emanics.org/
7. Google: Google Maps API (2010). http://www.code.google.com/apis/maps/
8. IP2Location.com: IP2Location Geolocation Service (2010). http://www.ip2location.com/
9. IsarNet Software Solutions GmbH: IsarFlow (2010). http://www.isarflow.com/
10. Leinen, S.: Evaluation of candidate protocols for IP flow information export (IPFIX) (2004). http://www.ietf.org/rfc/rfc3955.txt
11. Maymounkov, P., Mazières, D.: Kademlia: A peer-to-peer information system based on the XOR metric. In: Proceedings of the 1st International Workshop on Peer-to-peer Systems (IPTPS'02). Springer, Berlin (2002)
12. NfSen: Netflow Sensor (2010). http://www.nfsen.sourceforge.net/
13. Pras, A., Sadre, R., Sperotto, A., Fioreze, T., Hausheer, D., Schönwälder, J.: Using NetFlow/IPFIX for network management. J. Netw. Syst. Manag. **17**(4) (2009)
14. Quittek, J., Zseby, T., Claise, B., Zander, S.: Requirements for IP flow information export (IPFIX) (2004). http://www.ietf.org/rfc/rfc3917.txt
15. The Internet Engineering Task Force (IETF): IPFIX Status Pages (2010). http://www.tools.ietf.org/wg/ipfix/
16. Trammell, B., Boschi, E.: Bidirectional flow export using IP flow information export (IPFIX) (2008). http://www.ietf.org/rfc/rfc5103.txt
17. Zseby, T., Boschi, E., Brownlee, N., Claise, B.: IP flow information export (IPFIX) Applicability (2009). http://www.ietf.org/rfc/rfc5472.txt

Author Biographies

Idilio Drago is a Ph.D student at the DACS Group of the University of Twente, the Netherlands. He received his M.Sc in Computer Science from the Federal University of Espírito Santo, Brazil in 2007. He is currently researching the use of NetFlow for service performance monitoring.

Rafael R. R. Barbosa is a Ph.D student at the DACS Group of the University of Twente. In 2009 he received his M.Sc in Telematics from the same university. He is currently investigating how to perform intrusion detection based on the observation of network-wide behaviors. His topics of interest include intrusion detection and SCADA networks.

Ramin Sadre is a postdoctoral researcher at the DACS Group of the University of Twente. He is a Work Package leader within the European EMANICS Network of Excellence. He was technical program co-chair of AIMS 2009 and co-chair of two workshops on NetFlow/IPFIX usage in network management. His research interests include performance evaluation, traffic modelling, and intrusion detection.

Aiko Pras is Associate Professor at the University of Twente, where he is leading the DACS Group. He is chairing the IFIP Working Group 6.6 on "Management of Networks and Distributed Systems" and is Research Leader in the European EMANICS Network of Excellence. He is steering committee member of several conferences, including IM/NOMS and manweek, and series/associate editor of ComMag and IJNM.

Jürgen Schönwälder is Associate Professor at Jacobs University Bremen, Germany. He received his doctoral degree in 1996 from the Technical University Braunschweig. He is an active member of the Internet Engineering Task Force and chair of the Network Management Research Group of the Internet Research Task Force. He is a member of the editorial board of the JNSM and the IEEE TNSM.