Report of the third workshop on the usage of NetFlow/IPFIX in network management

J Netw Syst Manage (2011) 19:529–535
DOI 10.1007/s10922-011-9204-2

REPORT

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

Idilio Drago · Ramin Sadre · Aiko Pras

Received: 21 July 2011 / Accepted: 27 July 2011 / Published online: 6 August 2011

Abstract The Network Management Research Group (NMRG) organized in 2010 the *Third Workshop on the Usage of NetFlow/IPFIX in Network Management*, as part of the 78th IETF Meeting in Maastricht. Yearly organized since 2007, the workshop is an opportunity for people from both academia and industry to discuss the latest developments of the protocol, possibilities for new applications, and practical experiences. This report summarizes the presentations and the main conclusions of the workshop.

Keywords NetFlow · IPFIX · NMRG

1 Introduction

The Network Management Research Group (NMRG) organized in 2010 a one-day workshop for discussing the use of NetFlow/IPFIX in Network Management. As in the two previous editions, already reported in this journal [3, 5], the goal was to create a forum where people from both industry and academia could discuss the development of the protocol, as well as emerging applications. The third edition was held together with the 78th IETF Meeting in Maastricht, The Netherlands. The major strength of this edition was the presence of researchers interested in both the NMRG meeting and the IPFIX Working Group of the IETF. The presence of people from both groups helped to deepen the discussions.

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

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

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

In total, 45 people attended the workshop, which was composed of 9 presentations grouped into four themes:

1. New methods for flow-based application recognition
2. Flow-based application monitoring
3. The use of NetFlow/IPFIX in network security
4. New tools for NetFlow/IPFIX research

In addition to those topics, Paul Aitken and Benoit Claise, both from Cisco, presented their thoughts about the current stage of IPFIX and the challenges in the future of the protocol. In the following sections, we summarize all presentations and discussions of the workshop.

2 NetFlow/IPFIX: Challenges and Future Directions

Paul Aitken and Benoit Claise (Cisco) started their presentation discussing current issues related to NetFlow/IPFIX adoption. As new applications for NetFlow/IPFIX emerge, as well as new functions are standardized, the complexity of the protocol increases. They listed three possible problems for the deployment of NetFlow/IPFIX:

1. The performance of NetFlow/IPFIX monitoring systems is a bottleneck. Several improvements in the performance of metering devices have been made as, for example, the use of hardware acceleration and package sampling. However, there will be new bottlenecks (e.g. in collectors) as the speed of networks increases.
2. The Options Template is overloaded. The Options Template has been used in different situations as, for example, for carrying statistics about the protocol. However, this overload of functions is increasing the complexity of collectors.
3. Flexible templates negatively impact the performance of the metering process. This impact can jeopardizes the primary function of the device (forwarding packages).

Regarding the future of NetFlow/IPFIX, Aitken and Claise listed some new flow-based applications that are already emerging both in IPFIX research and at the IETF:

1. Application layer visibility: This would allow different quality of service treatment per application.
2. Permanent cache: Some flows could reside permanently at the metering cache, and periodically be exported. From the functional point of view, this can be seen as a user-defined (pushed) MIB.
3. IPFIX as a scalable logging service: Given the performance limitations of syslog in high speed environments, IPFIX could be a replacement for syslog.
4. Mediation function inside metering equipment: This would provide data aggregation and reduction even before exporting flow data.

These new applications would require further development in the IETF Working Group to redefine the IPFIX applicability (RFC 5472) and to define the mediation

protocol. They expect that some of those functions will be supported by equipment in the market soon (e.g. by Cisco Flexible NetFlow).

3 Flow-Based Application Recognition

One of the new flow-based applications that Aitken and Claise cited was further discussed in two other presentations in the workshop. A first requirement for application visibility is the ability to recognize the traffic (flows) created by a protocol. Pavel Piskac (Masaryk University/INVEA-TECH) is studying the use of time characteristics in NetFlow data for improving protocol detection. Although the need for protocol detection is widely accepted and several methods for that have already been proposed, time characteristics are not commonly employed. The main reason for that is the lack of support for precise time exporting in metering equipment (e.g. with microsecond resolution).

Piskac implemented detection methods to classify flow data, and used the SSHv2 protocol as a case study. In general, he was not able to classify interactive SSH traffic, because interactive SSH sessions generate traces similar to HTTP and IMAP, for example. However, using only time characteristics, dictionary attacks against SSH servers could be identified. He expects to improve his method when more precise time characteristics become available.

Nikolay Melnikov (Jacobs University Bremen) presented his progresses on identifying users based on their network activity (represented by NetFlow records). For that purpose, he evaluated if it is possible to define and to identify signatures for some of the most popular applications. Some applications generate typical traffic, even when users are not interacting with them. For example, the latest version of Google Chrome contacts a Google server immediately after being opened in order to download an updated list of phishing and malware web sites. From that behavior, Google Chrome could be identified in network traces. Assuming that other applications also have a typical behavior, and that users in a network use a distinctive set of applications, the identification of users would be possible.

Melnikov evaluated 13 applications, grouped into 4 categories: web browsers, instant messaging clients, mail clients and media players. Whenever possible, Linux and Windows versions were considered. For each application, he defined a flow-based signature using a special query language called Flowy [4]. In his first experimental evaluation, he was able to identify 10 voluntary users using 5 applications, with only 2 application mismatches. Melnikov is planning to continue his research creating signatures for more applications, doing experiments in uncontrolled environments (real traffic), and extending his analysis to also consider the date/time in which each application is used.

4 Flow-Based Application Monitoring

A consequence of IPFIX application visibility is that flow-based application monitoring becomes possible. Shingo Kashima (NTT) demonstrated his solution for

IPTV monitoring. Multicast is finally in use in some provider networks, especially because of IPTV. Current monitoring tools are not appropriate for large-scale multicast monitoring, either because they are not focused on real traffic (as in the case of multicast ping and traceroute), or because they are not scalable (as in the case of full traffic capture). According to Kashima, a solution for IPTV monitoring has to fulfill the following 4 requirements: (1) It must be able to detect service deterioration (like loss and delay variation); (2) it must identify users and channels affected by a failure; (3) it must identify the location of a failure; and (4) it must be ready for immediate deployment. Kashima used IPFIX/PSAMP for IPTV traffic monitoring. His solution (called QCast) uses Property Match Filtering and Systematic Time-Based Sampling, both from PSAMP, to collect the input packets, and IPFIX enterprise-specific information elements to carry data from RTP headers. The traffic received by IPFIX collectors is processed and alerts are triggered to operators. Kashima demonstrated his software during his presentation. The tool, however, is not available for public download.

Jochen Kögel (University of Stuttgart) is researching methods to extract network characteristics (like one-way delay, round trip time and packet loss) from the correlation among flows representing the same communication, but collected in different metering points. He identified several accuracy problems in the NetFlow exporters. Kögel collected data for his experiments in a real network connecting two locations through five routers—three of them under control of his institution. Kögel noted problems both in the precision of flow data and in the consistency of data collected in different metering points. Loss and duplication of information (NetFlow records), clock inaccuracy and imprecision in the packet/byte counters were the main issues. For example, in one of his experiments, he showed that the counters of different routers were precisely recording the number of packets, but not the number of bytes. Some deeper analyzes revealed that one router was rounding the byte counter.

From those experiments, he concluded that each router behaves in a peculiar manner, and that corrections are needed for better results. When routers with more accurate measurement capabilities are available, the characteristics of the network can be estimated. In general, a profile describing the artificial effects introduced by the exporter is necessary. Although such a profile could be provided by the manufacturer, a calibration step would be more realistic. The inclusion of profile information into IPFIX records was suggested as a measure to correct bias on the monitoring system.

5 NetFlow/IPFIX and Network Security

Two presentations covered the use of NetFlow/IPFIX in network security. Anna Sperotto (University of Twente) opened this session presenting her experience in creating a labeled data set for flow-based intrusion detection. Although several flow-based intrusion detection approaches have already been proposed in the literature, most of them were validated using proprietary traces, which makes it difficult to compare different approaches. Sperotto created the first public data set of NetFlow

records with ground truth, from traces collected at a honeypot at the University of Twente. The log files of the server hosting the honeypot were used to trigger alerts, which were later correlated to flow traces. Besides showing details of this data set, Sperotto also described the steps for creating it, which involved several non-automatic activities to label the flows. The data set has been anonymized using Nfdump, and is publicly available at [2].

Pavel Celeda (Masaryk University) is researching methods for malware detection based on the analysis of the network behavior. According to Celeda, as traffic acquisition and storage becomes easier, security analysis becomes more important and challenging. Neflow-based approaches for malware detection are promising because they are scalable and suitable for early detection of malwares and vulnerabilities (zero day exploits). The *Chuck Norris Botnet*, discovered at Masaryk University on 2 December 2009, was used as an example. The *Chuck Norris Botnet* is a Linux malware programmed to propagate itself in ADSL modems and routers with weak passwords. The infection occurs through brute force password discovery (dictionary attack) on the TELNET server of ADSL equipment. As soon as the attacker discovers the password, a malware is downloaded from a command host. The Botnet was used for DDoS attacks and DNS spoofing, aiming the infection of hosts using the infected ADLS devices. The Botnet was estimated to have infected around 33,000 distinct hosts from October 2009 to February 2010—most of them in South America. The activity of this Botnet was interrupted on February 2010.

From this experience with the *Chuck Norris Botnet*, the researchers of Masaryk University developed a plug-in for Nfsen able to detect botnets with similar behavior. Although the plug-in in its current version does not allow the detection a variety of botnets, a new version is already being prepared, making it as generic as possible.

6 NetFlow/IPFIX Tools

In the last section, three tools that can help in NetFlow/IPFIX research were presented. Brian Trammell (ETH Zurich) presented his *ripfix*: an environment for fast prototyping and debugging of IPFIX. The tool is optimized for development, and therefore ideal for situations where running code is needed to speed up the solution of issues and the development of new ideas (like in the IETF IPFIX Working Group). The tool, developed in Ruby, supports most of what is specified in RFC 5101, RFC 5103 and RFC 5655, besides some other functions specified on draft standards. In the current version, TCP is used as transport, but SCTP support will be included in the future. Several other functionalities specified in current IETF drafts, like anonymization of IPFIX records and structured data, are planned to be included in the tool soon. The tool is open source and can be downloaded from [8].

Stefan Burschka (Swisscom/CSIT) presented the Tranalyzer and Traviz tools. Tranalyzer is an open source tool designed to help with network analysis. Traviz is a graphical interface for Tranalyzer. Besides extracting basic NetFlow information from raw network data, Tranalyzer provides several methods for

both flow and packet analysis, like TCP state machine reconstruction, malicious packet detection, and advanced statistics extraction. Burschka presented some scenarios where Tranalyzer could be employed as, for example, the detection of routing anomalies, and the detection of applications by automatic reverse engineering their state machines. Tranalyzer also has simple interfaces with several other tools, like Matlab, GnuPlot and Excel. Tranalyzer is available for downloading at [6].

Lucjan Janowski (AGH Poland) closed this session presenting his trace2netFlow tool. Although IPFIX has been in development for several years, there are few tools for converting package traces into IPFIX records. The trace2netFlow aims to fill part of this gap, providing easy and flexible exporting capabilities. Janowski presented in details how the tool can be used to convert binary network data into IPFIX records, with configurable output fields and sampling techniques. The result data set can be directly exported into MATLAB, for example. The Trace2netFlow is available for downloading at [1].

7 Summary

The Workshop on the Usage of NetFlow/IPFIX in Network Management has been considered a valuable place to discuss new ideas related to NetFlow/IPFIX. The third edition was remarkable for the interaction between people from the NMRG and the IETF IPFIX Working Group. The minutes and all presentations are available at the website of the 78th IETF [7]. A next edition of the Workshop is planned for the second half of 2011.

Acknowledgments This report was partly supported by the Dutch Ministry of Economic Affairs, Agriculture and Innovation via its agency Agentschap NL, in the context of the IOP GenCom project Service Optimization and Quality (SeQual).

References

1. Adamski, K., Rusek, K., Korczynski, M., Janowski, L.: Traces2Flow—a simple way to generate a IPFIX data from packet trace (2010). http://www.tracesplay.sourceforge.net/traces2flow/
2. Barbosa, R.R.R., Sadre, R., Pras, A., van de Meent, R.: Simpleweb/University of Twente Traffic Traces Data Repository. Technical Report TR-CTIT-10-19, Centre for Telematics and Information Technology University of Twente, Enschede (2010). http://www.traces.simpleweb.org/
3. Drago I., Barbosa R.R.R., Sadre R., Pras A., Schönwälder, J.: Report of the second workshop on the usage of NetFlow/IPFIX in network management. J. Netw. Syst. Manag. **19**, 298–304 (2011)
4. Marinov, V., Schönwälder, J.: Design of a stream-based ip flow record query language. In: Bartolini, C., Gaspary L. (eds.) Integrated Management of Systems, Services, Processes and People in IT, Lecture Notes in Computer Science, vol. 5841, pp. 15–28. Springer, Berlin (2009)
5. Pras A., Sadre R., Sperotto A., Fioreze T., Hausheer D., Schönwälder J.: Using NetFlow/IPFIX for network management. J. Netw. Syst. Manag. **17**, 482–487 (2009)

6. Rühl, T., Bühlmann, F., Burschka, S.: Tranalyzer—flow based traffic analyzer (2010). http://www.tranalyzer.sourceforge.net/
7. The Internet Engineering Task Force: IETF 78—Maastricht, Netherlands (2010). http://www.ietf.org/meeting/78/
8. Trammell, B.: ripfix, IPFIX for Ruby (2010). http://www.ripfix.rubyforge.org/

Author Biographies

Idilio Drago is a PhD student at the Design and Analysis of Communication Systems (DACS) Group of the University of Twente, the Netherlands. He received his MSc 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.

Ramin Sadre is a postdoctoral researcher at the DACS Group of the University of Twente. He is currently active in the European Integrated Project UniverSelf on the self-management of communication networks. Before that, he was 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 the design of network intrusion detection systems, statistical traffic modeling, and the performance evaluation of communication systems.

Aiko Pras is Associate Professor in the Departments of Electrical Engineering and Computer Science at the University of Twente, the Netherlands where he is leading the Design and Analysis of Communication Systems Group. He received a PhD degree for his thesis titled "Network Management Architectures". His research interests include network management technologies, network monitoring, measurements and security. He is chairing the IFIP Working Group 6.6 on "Management of Networks and Distributed Systems", and was Research Leader in the European Network of Excellence on "Management of the Internet and Complex Services" (EMANICS). He is a steering committee member of several conferences, including IM/NOMS and CNSM, and series/associate editor of ComMag and IJNM.