IPFIX/PSAMP: What Future Standards can Offer to Network Security

Tanja Zseby[1], Elisa Boschi[2], Thomas Hirsch[1], Lutz Mark[1]
[1]*Fraunhofer Fokus*
{zseby, ,hirsch, mark}@fokus.fraunhofer.de
[2]*Hitachi Europe*
elisa.boschi@hitachi-eu.com

Abstract

Network security often requires the surveillance of the actual traffic in the network. Methods like signature-based attack detection or the detection of traffic anomalies require input from network measurements. The IETF currently standardizes the IP Flow Information Export (IPFIX) protocol for exporting flow information from routers and probes. The packet sampling (PSAMP) group extends the information model of IPFIX with the ability to report per packet information including parts of the payload. With this IPFIX and PSAMP provide valuable tools for detecting anomalies and security incidents in IP networks. Whereas the basic IPFIX and PSAMP documents are currently finalized, new drafts emerge that provide recommendations and IPFIX extensions. This paper shows how IPFIX and PSAMP can be used to support network security. Furthermore it is shown which extensions are useful and can provide further features for network security.

1. IPFIX and PSAMP

IPFIX defines a format and a protocol for the export of flow information from routers or measurement probes [1]. IPIFX uses a push-based data export, from IPFIX exporters to IPFIX collectors, and can run over TCP, UDP and SCTP. Figure 1 shows the process of measurement and export of IPFIX and PSAMP data. Core functions are always part of the measurement process. Optional functions can be placed in the processing sequence for different operations like post processing or data selection.

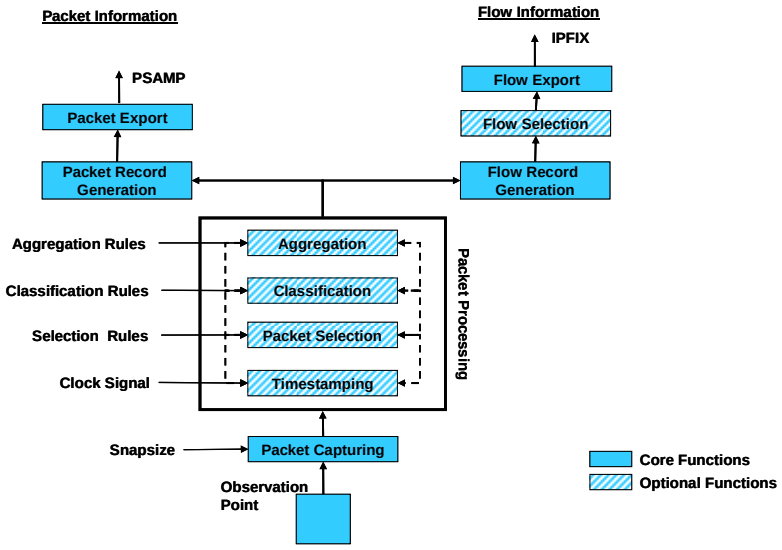

Figure 0: Measurement Model

[2] specifies observation point, flows, exporting and collecting process. The document also defines a metering process that consists of packet header capturing, timestamping, classifying, sampling and maintaining flow records. IPFIX Information Elements (IEs) for distinguishing flows and for reporting flow characteristics are defined in [3]. PSAMP extends the IPFIX information model by IEs for packet header and payload [4]. Furthermore it defines packet selection methods like filtering and sampling.

2. Metrics of Interest for Attack Detection

Network anomaly detection aims at discovering malicious behavior in a network by an analysis of the traffic profile. Using statistical means to detect unusual behavior patterns in the network or on the target machines, indicators for abuse or attacks are collected. While it can not be known in advance how precisely an attack will look like, experience has shown that certain information in packet headers provides better evidence than other. Also, some popular attack types can be easily detected by metering specific aspects of flow behavior.

Many detection algorithms base on very simple metrics. A successful detection of brute force attacks can be successfully done on data available from IPFIX. Early approaches like [5, 6] use simple flow definitions; and [7] successfully detects intrusions using only packet, ICMP packet and byte count using artificial immune systems. Most systems however face two difficulties: on the one hand, more advanced network intrusions are not easily detected by observing lower layers only. On the other, regular network events may produce a legitimate anomaly. The most common example is the so-called “flash crowds”, a sudden increases in traffic caused by a reference from another high-volume Internet service or news site. These events have to be distinguished from malicious attacks. Here, many algorithms rely on specialized metrics for popular types of attacks. The difference between TCP SYN and FIN packets for example is a clear indicator of ongoing SYN flooding attacks [8, 9].

Other general capabilities of interest for attack detection with IPFIX are: Flow separation by transport (e.g., TCP, UDP) or application layer (e.g., HTTP, FTP) protocols [10] or the retrieval of information from higher-level protocol headers such as TCP/IP [11] or information from MIB-II [12]. Further approaches use specialized statistics for attack detection that model real user behaviour more closely; [2] for example defines a question as any number of consecutive packets going from the client to the server. The number of questions (and answers resp.) per second is used as parameter for configuring self-organized maps. Finally, samples of the full payload information [13] allow further insight into transactions.

3. Measurement Requirements and what IPFIX and PSAMP can Offer

The detection of traffic anomalies requires passive measurements of the traffic in the network. IPFIX and PSAMP can be implemented on routers or probes and provide a standardized method to export flow and packet information from different points in the network. A variety of metrics are of interest for anomaly detection (see section 2). Currently IPFIX defines IEs for all IPv4 header fields (except checksum), the main IPv6 header fields (addresses, next header, flow label, etc.), the main transport header fields (UDP, TCP ports, sequence numbers, ICMP types), and some sub IP header fields (MAC addresses, MPLS labels, etc.). For reporting of flow statistics it defines a variety of counters (e.g. bytes, packets, delta and total counters), timestamps (flow start, end, duration) and basic statistics (min/max pktlength, min/max TTL, TCP flags, options). PSAMP extends the information model by adding IEs for reporting the full header and payload information. A useful information element for attack detection would be a counter to report the number of packets with specific flags (e.g. SYN, FIN) in a flow (e.g. to a specific destination address). This is currently not provided by IPFIX; IPFIX only supports the IE tcpCcontrolBits, which is a bitfield with all TCP

flags, where bits are set if a particular flag was observed for the flow. Nevertheless, the information model can be easily extended to support this counter.

In order to detect unusual behavior at different granularities or timescales, traffic needs to be observed at different aggregation levels. IPFIX provides an extremely flexible flow definition; a flow is defined as a set of packets with common properties. Each property is defined as a result of applying a function to one or more packet header fields (e.g. destination IP address), to further packet properties (e.g. number of MPLS labels) or to values derived from packet treatment (e.g. output IF). IEs defined in [3] can be used as flow keys to distinguish flows.

The analysis of the connection status (e.g. for TCP connections) requires a mapping of both directions of a communication. IPFIX currently reports each direction of a flow separately. With some additional effort a mapping of both directions is possible without IPFIX extensions. A more efficient way that introduces IEs for forward and backward direction is discussed in [14].

Distributed metrics often outperform metrics collected at a single observation point (see e.g., [15], [16], [17]), therefore data from multiple observation points should be correlated. Using identical flow keys at the observation points provides a network-wide picture about the flow situation. Synchronized clocks at the involved observation points allow the calculation of time-related metrics like one-way delay. If packet data needs to be correlated packet arrival events need to be recognised at different observation points. This can be done based on the packet content (header and optionally payload). For this only fields should be used that are immutable during transport but highly variable between different packets. In case the packet content itself is not needed (e.g. when calculating delay etc.) a packet ID based on those fields can be generated and post processing will be based only on this ID (see [18], [19], [20]). This significantly reduces the traffic that needs to be reported.

Post-incident analysis (network forensics) requires the storage of past data. This is also useful for sharing information among providers and to provide training data with “normal” behavior. In [21] requirements for an IPFIX file format are discussed and existing solutions for storage of flow information are investigated. It is planned to propose an IPFIX file format based on this study.

The ability to calculate specific metrics (e.g. packet ratios, statistics, etc.) directly on the router is a desired feature, since even if it consumes processing power on routers, it increases the speed at which incidents can be detected. Furthermore the reporting of derived metrics requires fewer transport resources than the export of all raw data. A disadvantage is the inability to derive arbitrary other metrics. If one does not know what to look for one can apply different methods on captured raw data. This is not possible if only derived metrics are reported. The reporting of derived metrics can be realized by extending the information model with new IEs as described in [22]. IPFIX is a push-based protocol. Currently the sending of flow records is triggered by flow termination criteria (e.g. flow idle time, TCP FIN, etc.) or resource limitations (cache full). If attack detection metrics are calculated directly on the router thresholds on these metrics could be used to trigger flow export. This would allow to reduce flow export to only those cases were suspicious behavior was observed.

Re-configuration of measurement processes is useful to zoom in or out based on the actual situation. Since the IPFIX group wanted first to concentrate on the protocol, the configuration of IPFIX functions was out of scope. Now that the IPFIX protocol is finished, several proposals for IPFIX configuration emerged. A first draft for an IPIFX MIB was described in [23]. An XML data model for configuration of IPFIX processes was proposed in [24]. Furthermore the Next Steps in Signaling (NSIS) group proposed a draft for path-coupled dynamic configuration of metering entities [25]. This framework can be used to configure parameters for IPFIX processes. A further desired feature is cost efficiency. Resources can be reduced by using filtering or sampling

techniques as described in [26]. [27] and [28] describe methods for aggregation and sharing of flow key information among data records.

An interoperation of measurement functions with AAA functions provides further features for network security [22]. AAA Functions may be able to map the traffic to specific users (e.g. by using the src address) and can stop network access for suspicious systems or users. Furthermore AAA provides secure channels to neighbor AAA servers and can inform neighbors about incidents or suspicious observations. Although most providers are still reluctant to information sharing, the ability to share information with neighbor domains is a useful feature. IPFIX provides the means to do that: TCP or SCTP can be used as transport protocol to ensure congestion-awareness and IPsec and TLS can be used as described in [1] to provide security features.

Table 1 summarizes the measurement requirements and shows how IPFIX, PSAMP and/or IPFIX extensions support specific features.

Measurement Requirement	**IPFIX support**	**PSAMP support**	**IPFIX extensions**
Network-wide passive measurements	Passive flow measurements integrated in routers	Packet capturing integrated in routers	-
Different aggregation levels	Flexible flow definition	Packet selection methods	[27], [28]
Variety of metrics	IEs for flow statistics, extensible info model	IEs for packet capturing, extensible info model	New IEs can be easily added
Analysis of connections	TCP flags bitmap	Header and payload information	[14]
Correlation from multiple observation points	Header fields for packet ID generation	Header and payload info for packet ID generation	[22]
Storage of past data	-	-	[21]
Export of derived metrics	-	-	[22], further planned
(Re-)configurability	-	Configuration of packet selection methods	[23], [24], [25]
Cost efficiency	Aggregation, packet selection	Packet selection methods	[27], [28]
Link to AAA functions	-	-	[22]
Inter-domain data exchange	Standard format, congestion-aware (TCP, SCTP), secure (IPsec, TLS)	Standard format, congestion-aware(TCP, SCTP), secure (IPsec, TLS)	[22]

Table 1: IPFIX and PSAMP Support for Anomaly Detection

5. Conclusions

IPFIX and PSAMP provide standardized measurement methods to support network security applications like attack and anomaly detection. A variety of relevant metrics can be derived from IPFIX and PSAMP data. Useful IPIFX extensions for correlation, aggregation and storage of IPIFX data have been proposed already within the IETF. Approaches for IPFIX configuration are underway.

Fraunhofer FOKUS has developed an open source IPIFX implementation. Besides the standard IPFIX IEs it supports proprietary IEs for reporting QoS metrics (loss, delay, jitter), TCP flag counters and packet IDs. The FOKUS IPFIX implementation is available at [29].

References

[1] B. Claise (Editor), "IPFIX Protocol Specification", Internet Draft, work in progress, June 2006

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

[3] J. Quittek, S. Bryant, J. Meyer, "Information Model for IP Flow Information Export", Internet Draft, work in progress, July 2006

[4] T. Dietz, F. Dressler, G. Carle, B. Claise, "Information Model for Packet Sampling Exports", Internet Draft, work in progress, June 2006

[5] Information-theoretic measures for anomaly detection, Wenke Lee; Dong Xiang, IEEE Symposium on Security and Privacy, S&P 200. 14-16, May 2001, Pages:130 - 143

[6] P. Barford, J. Kline, D. Plonka, A. Ron, "A signal analysis of network traffic anomalies", Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement, 2002 Marseille, France. Pages: 71 - 8

[7] M. Ostaszewski, F. Seredynski, and P. Bouvry. "A nonself space approach to network anomaly detection". In 20th International Parallel and Distributed Processing Symposium (IPDPS), NIDISC, Rhodes, Greece, April 2006.

[8] R.B.Blazek, H.Kim, B.Rozovskii, A.Tartakovsky, "A novel approach to detection of DoS attacks via adaptive sequential and batch-sequential change-point methods", Proceedings of the 2001 IEEE Workshop on Information Assurance and Security

[9] H Wang, D Zhang, K G Shin, "Detecing SYN flooding attacks", in Proc IEEE INFOCOM, 2002

[10] Novikov, D.; Yampolskiy, R.V.; Reznik, L, ".Anomaly Detection Based Intrusion Detection", ITNG 2006. April 2006 Page(s):420 – 425

[11] Hixon, R.; Gruenbacher, M., "Evaluation of the Fisher discriminant and chi-square distance metric in network intrusion detection", Region 5 Conference: Annual Technical and Leadership Workshop, 2 April 2004 Pages:119-124

[12] Jun Li; Manikopoulos, C., "Early statistical anomaly intrusion detection of DOS attacks using MIB traffic parameters", Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society 18-20 June 2003 Page(s):53 - 59

[13] K. Wang and S. J. Stolf, "Anomalous Payload-Based Network Intrusion Detection", Recent Advances in Intrusion Detection: 7th International Symposium, RAID 2004, Sophia Antipolis, France, September 15 - 17, 2004.

[14] B. Trammell, E. Boschi, "Bidirectional Flow Export using IPFIX", Internet-Draft, work in progress, June 2006

[15] J. Yin, G. Zhang, Y. Chen, X. Fan, "Multi-events analysis for anomaly intrusion detection", International Conference on Machine Learning and Cybernetics, 2004

[16] Y. Zhang, Z. Xiong, X. Wang, "Distributed Intrusion Detection Based on Clustering", International Conference on Machine Learning and Cybernetics, 2005

[17] Y. Wang; H. Yang; X. Wang; R. Zhang; "Distributed intrusion detection system based on data fusion method", Fifth World Congress on Intelligent Control and Automation, WCICA 2004.

[18] I. D. Graham, S. F. Donnelly, S. Martin, J. Martens, J. G. Cleary, "Nonintrusive and Accurate Measurement of Unidirectional Delay and Delay Variation on the Internet", INET'98, Geneva, Switzerland, 21-24 July, 1998

[19] N. Duffield, M. Grossglauser, "Trajectory Sampling for Direct Traffic Observation", Proceedings of ACM SIGCOMM 2000, Stockholm, Sweden, August 28 - September 1, 2000

[20] T. Zseby, S. Zander, G. Carle. "Evaluation of Building Blocks for Passive One-way-delay Measurements". Passive and Active Measurement Workshop (PAM), Amsterdam, The Netherlands, April 23-24, 2001.

[21] B. Trammell, E. Boschi, L. Mark, T. Zseby, "An IPFIX-Based File Format", Internet-Draft, work in progress, June 2006

[22] T. Zseby, E. Boschi, N. Brownlee, B. Claise, "IPFIX Applicability", Internet-Draft, work in progress, June 2006

[23] T. Dietz, A. Kobayashi, B. Claise, "Definitions of Managed Objects for IP Flow Information Export", Internet Draft, work in progress, June 19, 2006

[24] G. Muenz, "IPFIX Configuration Data Model", Internet Draft, work in progress, June 2006

[25] A. Fessi, G. Carle, F. Dressler, J. Quittek, C. Kappler, H. Tschofenig, "NSLP for Metering Configuration Signaling", Internet Draft, work in progress, June 26, 2006.

[26] T. Zseby, M. Molina, N. Duffield, S. Niccolini, F. Raspall, "Sampling and Filtering Techniques for IP Packet Selection" Internet Draft, work in progress, July 2005

[27] F. Dressler, C. Sommer, G. Muenz, "IPFIX Aggregation", Internet Draft, work in progress, June 2006

[28] E. Boschi, L. Mark, B. Claise, "Reducing redundancy in IPFIX and PSAMP reports", Internet Draft, work in progress, June 2006

[29] FOKUS IPFIX Implementation, *http://ants.fokus.fraunhofer.de/ipfix/*