

Implementing an IETF IPFIX Meter

F. Raspall, S. Tartarelli, M. Molina, J. Quittek
NEC Europe Ltd, Network Laboratories

Extended Abstract

Flow measurements in the Internet are becoming increasingly important for a number of applications, including traffic engineering, network planning, accounting and attack detection. Monitoring of flows typically involves two separate entities: a meter and a collector. The former may be a sniffing probe or dedicated HW/SW in the routing equipment, while the collector is the recipient of the measurement data, transforming and/or storing it before it is processed by an application. Currently, a variety of IP flow export systems are used in the practice, making it difficult to achieve an acceptable level of interoperability among tools and hindering the development of generalised flow analysis tools. The IETF IPFIX Working Group has been established with the goal of standardizing the export of IP flow measurement results from the meter to the collector [1].

In this paper, we describe a prototype implementation of a traffic measurement tool, that measures flows at different granularities and that reports on the volume statistics (bytes/packets) of each observed flow in an IPFIX compliant manner. The tool consists of a meter and a collector, both implemented on a standard PC platform under the Linux operating system. This platform offers several advantages, including the possibility of acquiring fast enough hardware for a prototype implementation at a reasonable cost and the fact that the Linux operating system is particularly suitable for testing new protocols, as it is the case with the newly proposed and still under standardisation IPFIX protocol. The meter is implemented as a probe co-located to network routers' interfaces and the software is written in user-space and uses the BSD socket interface for external communications.

The meter's main building blocks are the packet filtering module, the flow cache and the IPFIX exporter [2][3].

The packet filter is the module that, for every observed packet, decides whether it needs to be accounted. This decision depends on a list of rules that can be remotely and dynamically added or removed. If a packet is found to match a rule, an indication to the flow cache is sent so that the proper flow record is updated.

The flow cache is the local data structure that holds the information about the observed traffic. The flow cache requires a very careful design for different reasons. Potentially it can be accessed several times upon arrival of every packet and it is also accessed whenever a reporting task is triggered. The other main concern related to the flow cache is its size. Keeping track of a large number of flow records can be either very expensive (in case of SRAM) or very slow (if using DRAM). In the prototype described in this paper we implemented a mechanism that, if enabled, allows maximising the amount of accounted traffic, in case the total number of flows observed exceeds the actual flow cache size. The idea is to try to select those flows that mostly contribute to the total throughput on a certain link. This solution derives from the observation that in the Internet the flow size distribution is "heavy tailed", with few "elephant" flows accounting for a good portion of the overall traffic, and a large number of "mice" flows with negligible contribution.

The last component is the IPFIX exporter, which reads the flow cache, builds the export packet and delivers it to the collector. To ensure the flexibility in the flow definition required by the IPFIX protocol, the format of the records relies on the use of templates, which indicate the features of the observed traffic (e.g. values of header fields) selected to define a flow and the reported statistics.

The system has been designed to work on top of either TCP or UDP. The former is the preferable choice in case the congestion awareness and reliability requirements are to be met.

References

- [1] IETF IP Flow Information eXport working group:
<http://www.ietf.org/html.charters/ipfix-charter.html>.
- [2] J. Quittek, T. Zseby, B. Claise, S. Zander:
Requirements for IP Flow Information Export,
Internet draft <http://www.ietf.org/internet-drafts/draft-ietf-ipfix-reqs-08.txt>, work in progress,
January 2003.
- [3] B. Claise: *Cisco Systems NetFlow Services Export Version 9*, Internet draft <http://www.ietf.org/internet-drafts/draft-bclaise-netflow-9-00.txt>, work in progress, June 2002.