GPU-accelerated encrypted network traffic inspection

Eva Papadogiannaki
epapado@ics.forth.gr
ICS-FORTH
Greece

Sotiris Ioannidis
sotiris@ics.forth.gr
ICS-FORTH
Greece

Statistics show that more than 60% of the Internet traffic is now encrypted, while this percentage is constantly increasing. The majority of communications are secured using common encryption protocols such as SSL/TLS and IPsec in order to ensure security and protect the privacy of Internet users. Traditionally, Internet traffic analysis and monitoring is based on techniques like deep packet inspection (DPI). The core functionality of such DPI implementations is based on pattern matching, that enables searching for specific strings or regular expressions inside the packet contents. Common applications of DPI include but are not limited to firewalls, intrusion detection and prevention systems, L7 filtering and packet forwarding. With the widespread adoption of network encryption though, DPI tools that rely on packet content are becoming less effective, demanding the development of more sophisticated techniques in order not to become obsolete. Traditional DPI implementations can only extract very coarse-grained information for the majority of encrypted traffic, even though its analysis is a core operation for many network systems. Apparently, network inspection systems need to be improved and adapted to current encryption trends.

An approach to inspect encrypted network traffic is the generation of signatures based on packet metadata, such as the packet timestamp, size and direction. These metadata can be usable even with encrypted traffic, since they can be easily extracted from packet headers. Recent related work has proven that revealing the traffic nature in encrypted communication channels is feasible. For instance, Conti et al. proposed a system to analyse encrypted network traffic to identify user actions on Android devices, such as email exchange, interactions over social network, etc [3]. Their framework leverages information that is available in TCP/IP packets, like IP addresses and ports, among with other features, like packet size, direction and timing. Using machine learning techniques, they conduct their experiments that show that the system can achieve accuracy and precision higher than 95% for a number of user actions. Papadogiannaki et al. proposed a pattern language to describe packet trains for fine-grained identification of application-level events in encrypted network traffic. They provided an efficient implementation of this language, namely OTTer, based on an extended version of the Aho-Corasick algorithm [4]. This approach is tested against real traffic and presents a minor CPU overhead when integrated with a proprietary DPI engine. Current solutions that focus on detecting malicious network traffic include Symantec's Encrypted Traffic Management (ETM) and Cisco's Encrypted Traffic Analytics (ETA) tools. ETM gains visibility into encrypted traffic to stop threats. Yet, this approach could violate user privacy since traffic is decrypted using SSL visibility appliances [2]. ETA uses a more sophisticated technique that combines many different features of traffic. Still, this solution remains proprietary [1].

In this work, we investigate the utilization of hardware accelerators, such as GPGPUs, for high performance metadata matching against network traffic. The benefits for such an implementation is the high processing throughput as well as the low cost of powerful commodity high-end GPUs (in contrast to expensive server setups) [5]. Since GPUs offer stream processing, real-time traffic inspection can be achieved. Fast metadata matching can enhance the implementation of numerous applications tailored for encrypted networks, such as traffic monitoring and intrusion detection. In addition, such system can be utilized by service providers for analytics extraction in order to ensure quality of service for their clients. In Figure 1 we present a high level overview of our engine. The signatures that are extracted through an analysis phase, are compiled into an Aho-Corasick automaton that enables simultaneous multi-pattern matching. The incoming network traffic is grouped into batches and then transferred to the device memory space. Our engine is able to report suspicious behaviour during the pattern matching period against incoming network traffic.

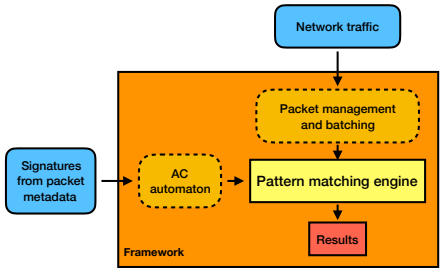

Figure 1: Architecture Overview

ACKNOWLEDGEMENTS

This work has received funding from the European Union's Horizon 2020 Research and Innovation programs under grant agreements No 780787 and No 830927.

REFERENCES

[1] [n. d.]. Cisco Encrypted Traffic Analytics. https://www.cisco.com/c/en/us/solutions/enterprise-networks/enterprise-network-security/eta.html. Accessed: 2019-04-11.
[2] [n. d.]. Symantec Encrypted Traffic Management. https://www.symantec.com/products/encrypted-traffic-management. Accessed: 2019-04-11.
[3] Mauro Conti, Luigi Vincenzo Mancini, Riccardo Spolaor, and Nino Vincenzo Verde. 2016. Analyzing android encrypted network traffic to identify user actions. *IEEE Transactions on Information Forensics and Security* 11, 1 (2016), 114–125.
[4] Eva Papadogiannaki, Constantinos Halevidis, Periklis Akritidis, and Lazaros Koromilas. 2018. OTTer: A Scalable High-Resolution Encrypted Traffic Identification Engine. In *International Symposium on Research in Attacks, Intrusions, and Defenses*. Springer, 315–334.
[5] Giorgos Vasiliadis, Lazaros Koromilas, Michalis Polychronakis, and Sotiris Ioannidis. 2014. {GASPP}: A GPU-Accelerated Stateful Packet Processing Framework. In *2014 {USENIX} Annual Technical Conference ({USENIX}{ATC} 14)*. 321–332.