High performance encrypted network traffic inspection using hardware accelerators

Eva Papadogiannaki
FORTH-ICS

Giorgos Vasiliadis
FORTH-ICS

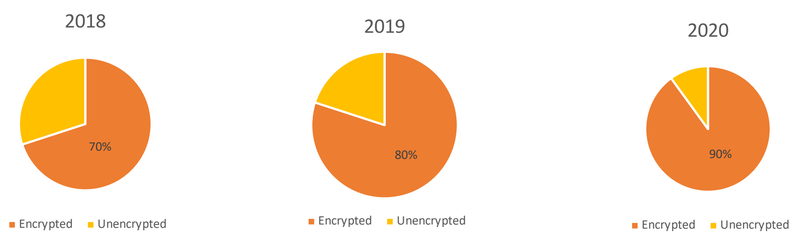
Sotiris Ioannidis
FORTH-ICS

Motivation

Background

Internet traffic analysis is commonly based on techniques like *deep packet inspection (DPI)*. The core of traditional DPI implementations is based on *pattern matching*, that enables searching for specific *strings* or *regular expressions* inside the packet content. With the widespread adoption of network encryption, DPI tools that rely on plain text pattern matching become less effective, demanding the development of more sophisticated techniques. Traditional DPI implementations can only extract very coarse grained information for the majority of such traffic.

By 2019, 80% of all network traffic will be encrypted.



Are we prepared for handling fully encrypted network traffic?

State of the art

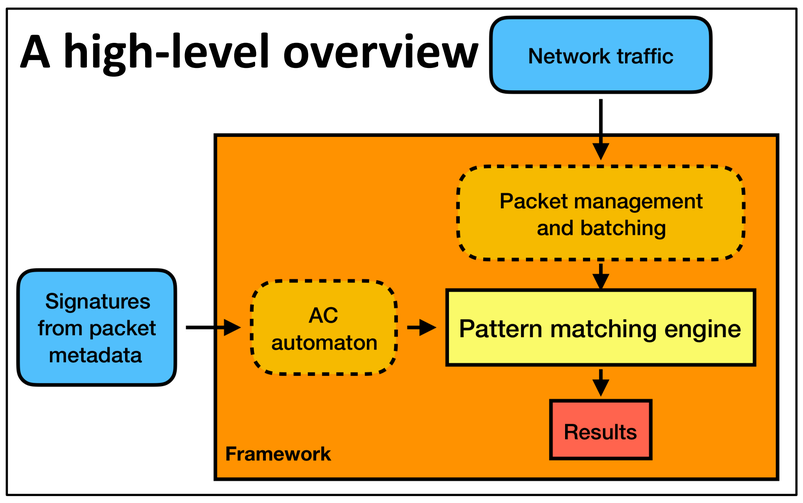
- Traffic decryption and inspection
  - BlindBox [1], Symantec's ETM [2]
  - **Could cause privacy violations / Expensive processing**
- Traffic classification using ML techniques that examine accuracy focusing on packet metadata (e.g. packet sizes, timestamps, direction).
  - Conti et al. [3]
  - **No real implementation**
- Real implementation of encrypted network inspection systems using packet metadata.
  - OTTer [4], Cisco's ETA [5]
  - **Proprietary source code**

Example applications based on DPI	Application domain
Firewall, Antivirus	Network security
Intrusion detection and prevention system	Network security
Network usage characterization	Network analytics
Traffic monitoring and classification	Network analytics
Network optimization	Network performance

Our solution

A framework that offers the functionality of a pattern matching engine tailored for packet metadata, such as packet sizes. This framework can be used to build different network related applications, such as firewalls, L7 filtering, or an Intrusion Detection System.
The signatures are processed into an automaton. The generation of the automaton is based on the Aho-Corasick algorithm that offers simultaneous multi-pattern matching. The implementation of the automaton is DFA-based and refactored to search for integers.

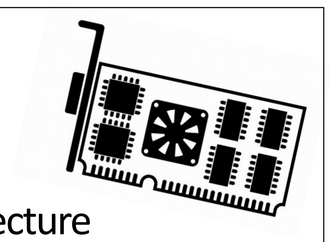
A high-level overview



Why Aho-Corasick?

- Preprocesses patterns to build a state machine
- Simultaneous multipattern searching
- Processes the input in a single pass
- Pattern searching can be transferred to GPU to achieve high throughput and real time results

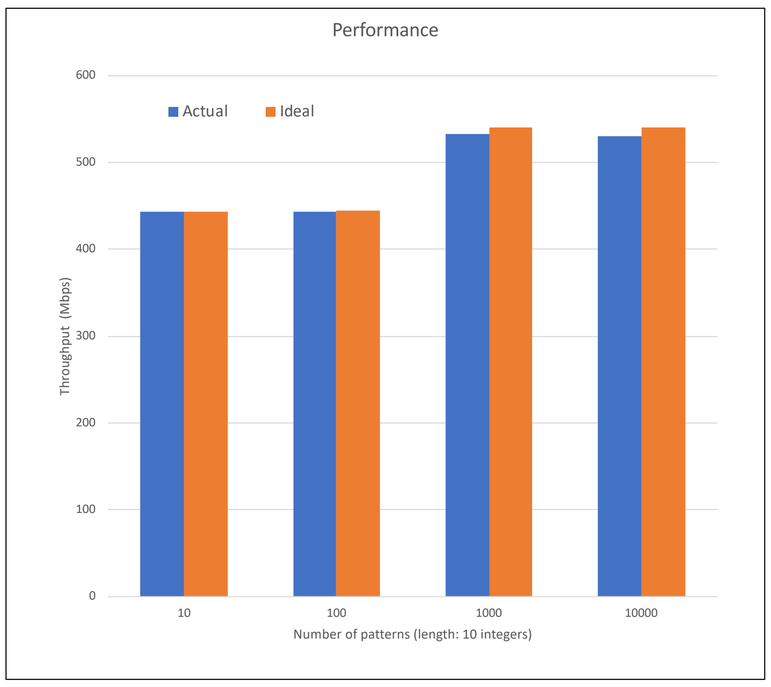
Use of GPUs



- Thousands of cores
- Highly parallel hardware architecture
- Low power consumption
- Up to 11GB GDDR6 memory
- High memory bandwidth (up to 616GB/s)
- Inexpensive commodity hardware

Packet processing and parallelization

- Assign one network traffic flow to one thread
- Packet batching to hide expensive transfers through PCIe
- Flow hashing to avoid packet reordering
- Filter out packet retransmissions
- Avoid costly packet copies and context switches with the netmap [6] module that operates in the user space (single buffer for efficient data sharing between the NIC & the GPU)



References

[1] Sherry, Justine, et al. "Blindbox: Deep packet inspection over encrypted traffic." ACM SIGCOMM Computer Communication Review 45.4 (2015): 213-226.
[2] Symantec's Encrypted Traffic Management (ETM). https://www.symantec.com/products/encrypted-traffic-management.
[3] Conti, Mauro, et al. "Analyzing android encrypted network traffic to identify user actions." IEEE Transactions on Information Forensics and Security 11.1 (2016): 114-125.
[4] Papadogiannaki, Eva, et al. "OTTer: A Scalable High-Resolution Encrypted Traffic Identification Engine." International Symposium on Research in Attacks, Intrusions, and Defenses. Springer, Cham, 2018.
[5] Cisco's Encrypted Traffic Analytics (ETA). https://www.cisco.com/c/en/us/solutions/enterprise-networks/enterprise-network-security/eta.html.
[6] Rizzo, Luigi. "Netmap: a novel framework for fast packet I/O." 21st USENIX Security Symposium (USENIX Security 12). 2012.

FORTH Foundation for Research & Technology - Hellas

This work has been partially supported by the **European Commission** under the **Horizon 2020** Program through the **I-BiDaaS** project under Grant Agreement **780787**.

I-BiDaaS