Chair of Network Architectures and Services
School of Computation, Information, and Technology
Technical University of Munich

Thesis B.Sc. Thesis M.Sc. IDP

TLS Certificate Usage Evaluation

Motivation

We at the chairs research group GINO (Global Internet Observatory) [1] operate a weekly Internet-wide TLS scan [4]. The full IPv4 address space and more than 200M domain name-IP address pairs are probed each time. This provides us with a unique insight into the usage and movement of x.509 certificates. Therefore, we want to create a data structure which allows to check the presence of certificates inside our data archive with a simple and fast query. Moreover, we want to evaluate the presence and movement of certificates across network entities.
x.509 certificates are the main trust provider for the TLS ecosystem. Due to their cryptographic properties certificate are unique and in this thesis we want to use this fact. Our goal is to better understand the current deployment and use the uniqueness property of certificates to learn more about deployed systems.

Your Task

- Familiarize yourself with data structures to efficiently store big data (Apache Parquet [2], Apache ORC [3], . . .)
- Understand the requirements and their accompanying problems
- Develop a pipeline which setup and fulfills the following goals
  - Determine which data needs to be extracted and how it can be efficiently stored
  - Select a data storage format which fulfills our requirements
  - Convert existing archive data into the new format in a limited given time frame
  - Write a comprehensive and accurate documentation.

Requirements

- In-depth knowledge on the Unix command line and how to process data using built-in tools
- Programming knowledge in Python is highly useful
- Familiarity with GIYF-Based work approaches
- Good code quality in order for others to continue and understand your work
- Optionally: Experience with big data storage formats is useful

Bibliography

[1] https://net.in.tum.de/projects/gino/
[2] https://parquet.apache.org/
[3] https://orc.apache.org/
[4] Amann, Johanna, et al. ”Mission Accomplished? HTTPS Security after DigiNotar.” Proceedings of the 2017 Internet Measurement Conference.

Contact

Patrick Sattler sattler@net.in.tum.de
Johannes Zirngibl zirngibl@net.in.tum.de