ElasticSniff

Degree programme: BSc in Computer Science | Specialisation: Digital Business Systems
Thesis advisor: Dr. Kenneth Ritley

This thesis describes the results of a product analysis, technical design and a working proof-of-concept involving the Elasticsearch Cloud product to carry out an agentless assessment of the IT services running in a data-center.

Context and goals

The project idea comes from Swisscom IT Services Finance Custom Solutions Ltd, for which a solution design and a proof-of-concept including the creation of a working prototype was agreed. The main goal of the thesis is to support the technical sales department in identifying and winning potential new customers. Many companies are in the process of migrating their on-prem datacenter environments to the cloud, but they do not know which services and applications they have in place. The idea is to offer the customer a low cost, cloud-based "least invasive" assessment to support building a business case for moving to the cloud.
There are existing products on the market to carry out such datacenter assessments, but they all have the same drawbacks: they are difficult to use, they require manual installation in the datacenter environment, and they have very high costs. It must be possible to carry out this assessment as autonomously and cost-effectively as possible, as the plan is to do this as a free pre-sales offer. Therefore, a good approach to achieve all these goals is to develop an agentless and cloud-based solution. During the pre-analysis of the task, i.e. before starting the thesis, the version of Elasticsearch hosted in the cloud was evaluated and found to be the most promising.

Contents of the Thesis

The Thesis consists of 5 different phases. The first is the project initialization with the organization, planning and requirements analysis. After that, an analysis of the technical possibilities of Elastic Cloud is performed. Thereby the different tools are looked at and described. The next step is to see which data sources can be used for the analysis. Since the solution must be agentless, existing data or simple data to be collected must be used. The 4th step of the thesis contains the technical conception for the prototype. The findings of the previous phases are taken into account in order to develop a solution that is as complete as possible. The last phase contains the implementation of the prototype on the basis of the technical conception.

Conclusion

Elasticsearch provides all the tools for effective analysis of the collected data. With Elasticsearch, the entire life cycle of the data can be covered. From collection, transport and processing to analysis and visualization, Elasticsearch has solutions on offer. So the challenge is more with the data itself. The right data source must be evaluated in order to obtain the required information. In the thesis, DNS log data was evaluated as the most promising, since it is easy to collect and is not critical for security or privacy. This can give the customer a first overview of the existing services and their usage. In later steps, this solution can be extended with other data, whereby correlations can be used to generate even more information in the analysis.

Severin Thalmann