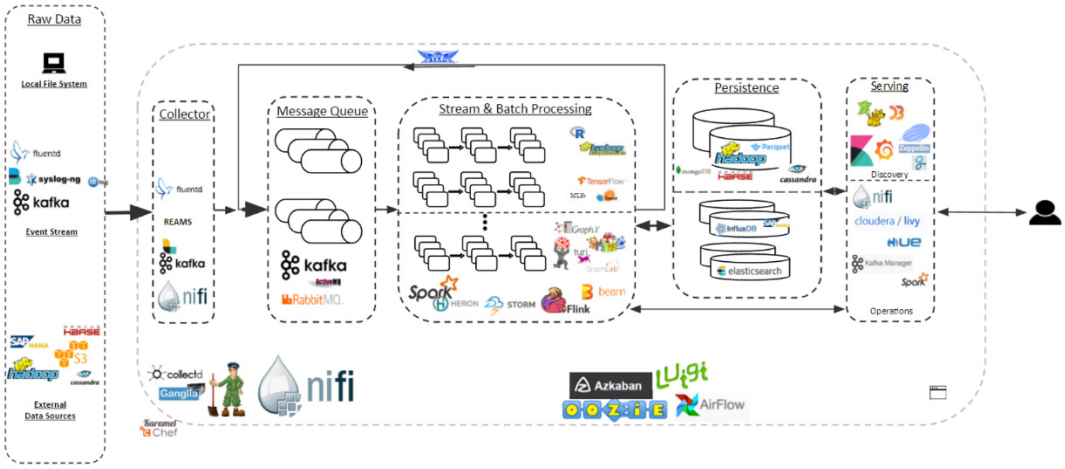
Prof. Dr. Christoph Meinel
Internet-Technologies & -Systems

Bachelor Project Winter Semester 2017/2018

A Big Security Data Analytical Framework

Background

Modern enterprises have invested more and more to collect, store and analyze data from their productive IT infrastructure. The goal is to identify valuable facts/information by mining in the huge amount of data they've collected. Performance, in terms of both accuracy and speed, is always a challenging task while the Big Data approaches are being applied in these practical use cases. Within this project, it is expected that the team can study the existing Big Data analytical approaches, including theories, technologies, and open source tools, etc., and explore the potentials to apply these approaches in the domain of security analytics. The team will work closely with the IT-Security Engineering team at HPI as well as the relevant project teams of both collaborative partners, i.e., SAP SE (Walldorf, Germany) and Shell International B.V. (The Hague, Netherlands).



Concrete Tasks

The suggested tasks of the project are described in (but not limited to) the following list:

1) Big Security Data Analytics: Back-End (Platform)
2) Big Security Data Analytics: Front-End (Serving)
3) Big Security Data Analytics: Arsenal
    a. HPI Threat Intelligence Platform
    b. Advanced Machine Learning, Data Mining, and Graph Analytics

Deliverables

- A running prototype
- An integrated technical report and several bachelor theses on individual research topics within the project

Requirements

- Java/Python
- Basic knowledge on IT security as well as data science

- Experiences with R, Apache Spark, Kafka, Hadoop, MapReduce, Jupyter/Zepplin, ELK Stack (Elasticsearch, Logstash, Kibana), SIEM and log management systems (e.g., Splunk, arcsight, …) (Preferred)

References

- Data Science Labs:
  - https://cloud.google.com/datalab/
  - https://azure.microsoft.com/en-us/solutions/big-data/
  - https://aws.amazon.com/big-data/
  - https://datascience.ibm.com/
- Databricks: https://databricks.com/
- Hops: http://www.hops.io/

Contacts

- Prof. Dr. Christoph Meinel, Dr. Feng Cheng (HPI)
  - Email: security-analytics@hpi.de
  - Room: H-1.13 / Tel: +331-5509-519
- Industrial Partners: SAP SE, Shell International B.V.