Contribution ID: **38** Type: **Oral Presentation**

Building a large scale Intrusion Detection System using Big Data technologies

Thursday, 22 March 2018 14:00 (20 minutes)

Computer security threats have always been a major concern and continue to increase in frequency and complexity. The nature and techniques of the attacks evolve rapidly over time, making the detection of attacks more difficult, therefore the means and tools used to deal with them need to evolve at the same pace if not faster.

In this paper a system for intrusion detection (IDS) both at the network (NIDS) and host (HIDS) level is presented. The system is currently processing in real time approximately half a TB of data per day, with the final goal of coping with 2.5 TB. In order to accomplish this goal firstly an infrastructure to collect data from sources such as system logs, web server logs and the network based Intrusion Detection System logs has been developed making use of technologies such as Apache Flume and Apache Kafka. Once the data is collected it needs to be processed in search of malicious activity: the data is consumed by Apache Spark jobs which compare in real time this data with known signatures of malicious activities. These are known as IoC or Indicator of Compromise, they are published by many security experts and centralized in a local MISP (Malware Information Sharing Platform) instance.

Nonetheless, detecting an intrusion is not enough. There is a need to understand what happened and why. In order to gain knowledge on the context of the detected intrusion the data is also enriched in real time when it is passing through the pipeline. For example DNS resolution and IP geolocation are applied to the it. Therefore, a system generic enough to process any kind of data in JSON format is enriching the data in order to get full context of what is happening and finally looking for indicators of compromise to detect possible intrusions, making use of the latest technologies in the Big Data ecosystem.

Primary author: Mr PANERO, Pablo (CERN)

Co-authors: SCHUSZTER, Cristian (CERN); Mr VALSAN, Liviu (CERN); WARTEL, Romain (CERN); Mr BRILLAULT, Vincent (CERN/EGI)

Presenter: Mr VALSAN, Liviu (CERN)

Session Classification: Networking, Security, Infrastructure & Operation Session

Track Classification: Networking, Security, Infrastructure & Operations