Contribution ID: **22** Type: **Presentation short (15 min)**

A high performance and real-time prototype for Anomaly Detection in a Datacenter

Wednesday, 20 October 2021 11:40 (25 minutes)

The amount of companies and DataCenters that handle large volumes of data using Computer Information Systems are constantly growing. Besides, the computing infrastructures become more complex integrating different types of architecture (HPC, Cloud Computing, GPU, low-latency networking, etc). In addition, hundreds or thousands of users sharing the same resources make the systems more vulnerable. Whenever a security issue arise due to an internal or external access, the consequences can be very serious for both suppliers and the customers sharing resources. In the worst case, this may cause the partial or total loss of data. Furthermore, the cybercriminals use new techniques that require complex security systems with many services involved, which become more difficult for the system administrators to manage. The use of monitoring systems and its continuous update are intended for controlling the network traffic at real time as well as avoiding the complex and unknown threats.
Intrusion Detection Systems (IDS) play an important role in a Datacenter with the purpose of generating alerts at the moment when a malicious event arises. By this way, the system administrator can immediately react to know what services are vulnerable. However, the system generates alerts based on known rules and it is not capable of detecting and classifying the anomaly traffic causing the administrator unaware of the attack. In recent research, the IDS are complemented with other tools and techniques of Machine Learning to allow the analysis and post-processing of those events. Supported by high-distributed and high-performance computing, the integration of those tools makes the processing of the data in real time as well as the detection of anomaly events as quickly as possible.
This presentation will introduce the most used IDS as well as a description of our current deployed architecture based on Suricata. Apart from that, a high-distributed and high-performance computing prototype architecture will be presented, where different tools of fault-tolerant message systems (like kafka) and processing systems like Spark Streaming, Hadoop and Data Lakehouses are integrated to collect the data from the IDS and transform it into a more efficient type of data automatically. This architecture will allow the post-processing of the traffic packets as well as the anomaly events using ML techniques to make a fast detection and correct classification of the anomaly events by the IDS.

Speaker bio: https://www.linkedin.com/in/aida-palacio-hoz-633bb45a/?originalSubdomain=es

Most suitable track

Innovating services together

By submitting my abstract, I agree that my personal data is being stored in accordance to conference Privacy Policy

Primary author: Ms PALACIO HOZ, Aida (IFCA)

Co-author: Dr LÓPEZ GARCÍA, Álvaro

Presenter: Ms PALACIO HOZ, Aida (IFCA)

Session Classification: Envisioning the Future - Presentations