Contribution ID: **310** Type: **Poster**

Enhanced Data Analytics capabilities in the ELK Stack - a review of the premium features and their benefit to a Scientific Compute Facility

Thursday 27 October 2022 16:10 (30 minutes)

In real-time computing facilities - system, network, and security monitoring are core components to run efficiently and effectively. As there are many diverse functions that can go awry, such as load, network, processes, and power issues, having a well-functioning monitoring system is imperative. In many facilities you will see the standard set of tools such as Ganglia, Grafana, Nagios, etc. While these are noteworthy, the diversity of tools used clearly points to an adequacy gap (none is self-sufficient) and furthermore, they lack in their alerting and anomaly detection capabilities beyond the binary events.

The ELK stack (Elasticsearch, Logstash, & Kibana) is the combination of three open-source projects to ingest, search, and visualize logs and data. The basic free license of ELK enables these features but overall is limited for use in a real-time facility. Instead, by leveraging the full capabilities of ELK, the gained features are significant. ELK offerings provide many enhancements from single sign-on and means to control Authorization for security, including alerting for unusual events, Machine Learning capabilities, and many other tools that are useful for advanced data analytics.

With the advanced set of Machine Learning techniques, the ELK toolbox adds features such as clustering, time series decomposition, and correlation analysis. For example, these Machine Learning techniques can be applied to alerts, providing you with the details of events for an unusual uptick in resource usage, if there is rare or high process activity, or unusual port activity. A standard monitoring tool would typically not have such capability.

In this report, will discuss the details and features of how a facility could benefit from the open source and premium versions of the ELK stack. We will provide procedures and details for configuring these tools, and how it benefits compute facility monitoring postures within a scientific based environment.

Experiment context, if any

References

Flexible visualization of a 3rd party Intrusion Prevention (Security) tool: A use case with the ELK stack - https://indico.cern.ch/event/855454/contributions/4604980/

Significance

Focus on the ELK offerings with use of their Machine Learning techniques to provide scientific compute facilities specialized alerting and monitoring.

Primary author: POAT, Michael

Co-author: Dr LAURET, Jerome (Brookhaven National Laboratory)

Presenter: POAT, Michael

Session Classification: Poster session with coffee break