Luca Magnoni
CERN IT-SDC/MI

PhD Project Proposal:

“ Scalable data store and analytics platform for monitoring WLCG, a distributed data-intensive scientific infrastructure “

Introduction and context

The Worldwide LHC Computing Grid (**WLCG**) provides the computing resources to store, distribute and analyse the ~25 Petabytes of data annually generated by the Large Hadron Collider (LHC), serving a community of more than 8,000 physicists distributed among 150 computing centres around the world.

Every day, thousands of files are transferred and hundreds of thousands of processing jobs are executed on the different computing sites to perform scientific analysis on LHC data. Monitoring such a distributed, geographically-sparse and data-intensive infrastructure is a core functionality and one of the greatest challenges to provide a reliable scientific platform. The development and the operation of the WLCG monitoring infrastructure is responsibility of the Support for Distributed Computing (**SDC**) group, at the CERN IT department.

The current monitoring system has proven to be a solid and reliable solution to support WLCG functions and operations during LHC data-taking years. A variety of data coming from different services and experiment-specific frameworks is gathered, processed and archived and a generic web-based dashboard provides a uniform and customizable monitoring interface for scientists and sites.

Problem and objectives

In the next future, the WLCG monitoring infrastructure have to cope with an extension of the **volume** (e.g. higher LHC luminosity) and the **variety** (e.g. new data-transfer protocols and new resource-types, as cloud-computing) of the monitoring data. Nevertheless, traditional architecture where relational database systems are used to store, to process and to serve monitoring events has clear limitations. Scalability is difficult to achieve, technique like sharding push **complexity at application level**, leading to higher maintenance and operational costs and to human-faults. Moreover, effective monitoring requires **low-latency read access** to real-time data, while built-in database procedures impose constraints on the format, granularity and timeliness of monitoring information.

In recent years, the challenge of handling big volume of data has been taken by many companies, particularly on the internet domain, leading to a full paradigm shift on data archiving, processing and visualization. A stack of new technologies appeared, each one targeting specific aspects on big-scale distributed data-processing. However, all these technologies, such as batch computation system and non-structured databases, can handle very large data-volumes with little cost but with **serious trade-offs**. For example, Hadoop can run large-scale batch computations on very large data volume, but with high latency. Another example, non-relational database systems as Cassandra can offer scalability but only supporting a very limited data model with no or relaxed consistency.

The student will join the WLCG monitoring team which is working on the research, the design and the development of the new data store and analytics platform for the evolution of the WLCG monitoring, able to cope with the scalability, flexibility and fault-tolerance requirements foreseen in the long-term WLCG scenario. The task requires a sound knowledge of distributed systems theory and concepts together with a deep understanding of the new technology stack for large-scale data analysis. The project will be done in collaboration with the Agile Monitoring initiative of the CERN IT department. The goal is to architect the new platform in a tool-chain approach leveraging on the most appropriate technologies and computing techniques for the WLCG use case.

The project can be decomposed in three main objectives and areas of work. The first objective is the **batch layer,** to store a constantly growing dataset providing the ability to compute arbitrary functions on it. The second objective is the **serving layer**, to store the batch-processed views, using indexing techniques to make them efficiently queryable. The third objective is the **real-time processing layer** able to perform analytics on fresh data with incremental algorithms to compensate for batch-processing latency. Moreover, the student will investigate how to use the real-time analytics layer as input for **active-reaction,** adopting classical pattern matching approach to promptly detect errors and failures on the stream of monitoring events.

The work will start with the analysis of the state of the art technology and of the WLCG requirements. Then, the work will concentrate on the design of the architecture and the data model, followed by a development and prototyping phase validated on historical WLCG monitoring datasets. The final result of the project should be a production-quality data store and analytics platform ready to be used by the evolution of the WLCG monitoring infrastructure for the coming years of LHC data-taking.