## The 8th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID 2018)



Contribution ID: 199

Type: Plenary reports

## The ATLAS EventIndex and its evolution based on Apache Kudu storage

Tuesday, 11 September 2018 11:20 (30 minutes)

The ATLAS experiment produced hundreds of petabytes of data and expects to have one order of magnitude more in the future. This data are spread among hundreds of computing Grid sites around the world. The EventIndex catalogues the basic elements of these data - real and simulated events. It provides the means to select and access event data in the ATLAS distributed storage system, and provides support for completeness and consistency checks and data overlap studies. The EventIndex employs various data handling technologies like Hadoop and Oracle databases, and is integrated with other elements of the ATLAS distributed computing infrastructure, including systems for data, metadata, and production management (AMI, Rucio and PANDA). The project is in operation since the start of LHC Run 2 in 2015, and is in permanent development in order to fit the analysis and production demands and follow technology evolutions. The main data store in Hadoop, based on

MapFiles and HBase, can work for the rest of Run 2 but new solutions are explored for the future. Kudu offers an interesting environment, with a mixture of BigData and relational database features, which looked promising at the design level and is now used to build a prototype to measure the scaling capabilities as a function of data input rates, total data volumes and data query and retrieval rates. An extension of the EventIndex functionalities to support the concept of Virtual Datasets produced additional requirements that are tested on the same Kudu prototype, in order to estimate the system performance and response times for different internal data organisations. This talk reports on the current system performance and on the first measurements of the new prototype based on Kudu.

Primary author: Prof. BARBERIS, Dario (University and INFN Genova (Italy))Presenter: Prof. BARBERIS, Dario (University and INFN Genova (Italy))Session Classification: Plenary reports