Contribution ID: **265**   Type: **Sectional reports**

A distributed data warehouse system for astroparticle physics

Tuesday, 11 September 2018 15:45 (15 minutes)

A distributed data warehouse system is one of the actual issues in the field of astroparticle physics. Famous experiments, such as Tunka, Taiga, produce tens of terabytes of data measured by their instruments. It is critical to have a smart data warehouse system on-site to store the collected data for further distribution effectively. It is also vital to provide scientists with a handy and user-friendly interface to access the collected data with proper permissions not only on-site but also online. The latter case is handy when scientists need to combine data from different experiments for analysis. In this work, we describe an approach to implementing a distributed data warehouse system that allows scientists to acquire just the necessary data from different experiments via the Internet on demand. The implementation is based on the CERN CVMFS with additional components developed by us to search through the whole available data sets and deliver their subsets to users' computers.

Primary authors: Dr KRYUKOV, Alexander (SINP MSU); Mr NGUYEN, Minh Duc (Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University)

Presenter: Mr NGUYEN, Minh Duc (Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University)

Session Classification: 10. Databases, Distributed Storage systems, Datalakes

Track Classification: 10. Databases, Distributed Storage systems, Datalakes