

# Federated data storage system prototype for LHC experiments and data intensive science

*Thursday, 13 October 2016 16:30 (15 minutes)*

Rapid increase of data volume from the experiments running at the Large Hadron Collider (LHC) prompted national physics groups to evaluate new data handling and processing solutions. Russian grid sites and universities' clusters scattered over a large area aim at the task of uniting their resources for future productive work, at the same time giving an opportunity to support large physics collaborations.

In our project we address the fundamental problem of designing a computing architecture to integrate distributed storage resources for LHC experiments and other data-intensive science applications and to provide access to data from heterogeneous computing facilities. Studies include development and implementation of federated data storage prototype for Worldwide LHC Computing Grid (WLCG) centers of different levels and University clusters within one National Cloud. The prototype is based on computing resources located in Moscow, Dubna, St.-Petersburg, Gatchina and Geneva. This project intends to implement a federated distributed storage for all kind of operations such as read/write/transfer and access via WAN from Grid centers, university clusters, supercomputers, academic and commercial clouds. The efficiency and performance of the system are demonstrated using synthetic and experiment-specific tests including real data processing and analysis workflows from ATLAS and ALICE experiments, as well as compute-intensive bioinformatics applications (PALEOMIX) running on supercomputer.

We present topology and architecture of the designed system, report performance and statistics for different access patterns and show how federated data storage can be used efficiently by physicists and biologists. We also describe how sharing data on a widely distributed storage system can lead to a new computing model and reformations of computing style, for instance how bioinformatics program running on supercomputer can read/write data from federated storage.

## Primary Keyword (Mandatory)

Distributed data handling

## Secondary Keyword (Optional)

Storage systems

## Tertiary Keyword (Optional)

**Primary author:** KLIMENTOV, Alexei (Brookhaven National Laboratory (US))

**Co-authors:** KIRIANOV, Andrey (B.P. Konstantinov Petersburg Nuclear Physics Institute - PNPI ()); ZAROCHENTSEV, Andrey (St. Petersburg State University (RU)); KRASNOPEVTSEV, Dimitrii (National Research Nuclear University MEPhI (RU)); RYABINKIN, Eygene (National Research Centre Kurchatov Institute (RU))

**Presenter:** KIRIANOV, Andrey (B.P. Konstantinov Petersburg Nuclear Physics Institute - PNPI ( ))

**Session Classification:** Posters B / Break

**Track Classification:** Track 4: Data Handling