Contribution ID: **6**   Type: **Oral Presentation**

Russian National Data Lake Prototype

Thursday, 27 August 2020 11:50 (20 minutes)

The evolution of the computing facilities and the way storage will be organized and consolidated will play a key role in how this possible shortage of resources will be addressed by the LHC experiments. The need for an effective distributed data storage has been identified as fundamental from the beginning of LHC, and this topic has became particularly vital in the light of the preparation for the HL-LHC run. WLCG has started an R&D within DOMA project and in this contribution we will report the recent results related to the Russian federated data storage systems configuration and testing. We will describe different system configurations and various approaches to test data storage federation. We are considering EOS and dCache storage systems as a backbone software for data federation and xCache for data caching. We'll also report about synthetic tests and experiments specific tests developed by ATLAS and ALICE for federated storage prototype in Russia. Data Lake project has been launched in Russian Federation in 2019 to set up a National Data Lake prototype for HENP and to consolidate geographically distributed data storage systems connected by fast network with low latency, we will report the project objectives and status.

Primary author: Mr KIRYANOV, Andrey (NRC "Kurchatov Institute")

Presenter: Mr KIRYANOV, Andrey (NRC "Kurchatov Institute")

Session Classification: Data Management Session

Track Classification: Data Management & Big Data