

An overview of technologies behind the Russian Scientific Data Lake prototype

Friday, 25 March 2022 11:30 (30 minutes)

A substantial data volume growth is expected with the start of the HL-LHC era. Even taking into account the hardware evolution it will require substantial changes to the ways data is managed and processed. The WLCG DOMA project was established to address the relevant research, and along with the national Data Lake R&Ds it studied the possible technology solutions for the organization of intelligent distributed federated storage. In this talk we will present one such R&D: the Russian Scientific Data Lake prototype, which is ongoing for the last three years, with the focus on software and technologies used for its deployment and testing.

Primary authors: ALEKSEEV, Alexandr; KIRYANOV, Andrey (NRC "Kurchatov Institute"); KLIMENTOV, Alexei; ZAROCHEVSEV, Andrey; KORCHUGANOVA, Tatiana

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

Session Classification: Data Management & Big Data

Track Classification: Track 6: Data Management & Big Data