Contribution ID: **68**   Type: **Short Talk**

Prototype of the Russian Scientific Data Lake

Wednesday 19 May 2021 10:50 (13 minutes)

The High Luminosity phase of the LHC, which aims for a ten-fold increase in the luminosity of proton-proton collisions is expected to start operation in eight years. An unprecedented scientific data volume at the multi-exabyte scale will be delivered to particle physics experiments at CERN. This amount of data has to be stored and the corresponding technology must ensure fast and reliable data delivery for processing by the scientific community allover the world. The present LHC computing model will not be able to provide the required infrastructure growth even taking into account the expected hard-ware evolution. To address this challenge the Data Lake R&D project has been launched by the DOMA community in the fall of 2019. State-of-the-art data handling technologies are under active development, and their current status for the Russian Scientific Data Lake prototype is presented here.

Primary authors: Mr KIRIANOV, Andrey (NRC Kurchatov Institute PNPI (RU)); ZAROCHENTSEV, Andrey (St Petersburg State University (RU)); KLIMENTOV, Alexei (Brookhaven National Laboratory (US)); ALEKSEEV, Aleksandr (Universidad Andres Bello (CL)); MITSYN, Valeri (Joint Institute for Nuclear Research (RU)); Mr OLEYNIK, Danila (Joint Institute for Nuclear Research (RU)); ESPINAL, Xavier (CERN); JEZEQUEL, Stephane (LAPP-Annecy CNRS/USMB (FR)); KORCHUGANOVA, Tatiana (Universidad Andres Bello (CL)); SMIRNOV, Alexander (Plekhanov Russian University of Economics, Moscow, 117997, Russia); SMIRNOV, Sergei (National Research Nuclear University MEPhI, Moscow, 115409, Russia)

Presenter: Mr KIRIANOV, Andrey (NRC Kurchatov Institute PNPI (RU))

Session Classification: Storage

Track Classification: Distributed Computing, Data Management and Facilities