Title of the series of works

"Development of the software complex for the implementation of a unified architecture for distributed data processing and storage at the BM@N/NICA experiment"

Abstract

One of the key elements of the first stage of the JINR flagship project on creation of the NICA accelerator-storage complex is an already ongoing experiment on a fixed target, BM@N (Baryonic Matter at Nuclotron). Since 2015, seven technical runs of the experiment had been conducted, and from December 2022 to February 2023 the first physics run was successfully performed, in which about 600 million events of xenon ion beam collided with a cesium-iodine target were collected, occupying about 400 TB of disk space. Moreover, when the experiment reaches the design parameters, the volume of data obtained is expected to increase by an order of magnitude. To process and perform necessary physics analyses of such a large volume of data, a comprehensive approach with a properly designed architecture of software complex is required for distributed data processing on the computing platforms provided to the experiment.

The presented series of the articles describes a complex of the developed software systems for organizing distributed data processing at the BM@N experiment and for unified storing and management of information required for the processing at all stages, including raw data decoding, reconstruction, physics analysis, as well as detector simulation. The implemented architecture includes both software systems (such as a central workload management system, single file catalogue, workflow management service) that solve the issue of combining separated resources of the experiment into a single processing and storage system to automate the execution of BM@N task flows, and original information systems (the Electronic Logbook, Online Configuration Platform, Geometry and Condition Databases, the Event Metadata System) that ensure collection, storage and organization of access to information required for processing and analysing the obtained data throughout the entire life cycle of scientific research at the BM@N experiment. In addition, a set of auxiliary services, such as a software distribution system, single authentication and authorization service, and monitoring service, improves the efficiency and reliability of the developed architecture. The implemented complex of the software systems for distributed data processing and storage at the BM@N experiment is successfully used in solving issues of processing, storing and analysing the collected data, and it is also an essential element for high-quality, timely obtaining of physics results in terms of working with big data flows.