

# DEVELOPMENT METHODOLOGY FOR ORGANIZING A UNIFIED ARCHITECTURE OF THE SOFTWARE COMPLEX FOR DISTRIBUTED DATA PROCESSING AND STORAGE IN THE BM@N EXPERIMENT

*K. V. Gertsenberger*<sup>1,\*</sup>

<sup>1</sup> Joint Institute for Nuclear Research, Dubna, Russia

From December 2022 to February 2023, the first physics run of the fixed target experiment, BM@N, of the NICA Accelerator Complex was successfully conducted, in which about 600 million events of xenon ion collisions with a cesium–iodine target were collected, occupying about 400 TB of disk space. To perform multiple processing and necessary physics analyses of such a large volume of data, a comprehensive architecture of software solutions for distributed data processing and storage is required. We present a development methodology and architecture of the software complex for organizing distributed processing of BM@N data and unified storage and management of information required for processing at all stages. The architecture implemented in accordance with the methodology includes both software systems that solve the problem of combining all the distributed resources of the experiment into a single computing complex with a single storage space to provide the automation of task processing flows, and original information systems that provide collection, storage and access to information being necessary for processing and analyses of the obtained data. In addition, a set of auxiliary solutions, such as a software distribution system, single authentication and authorization system, and monitoring system, improves the efficiency and reliability of the developed architecture.

С декабря 2022 по февраль 2023 г. был успешно проведен первый физический сеанс эксперимента на фиксированной мишени BM@N ускорительного комплекса NICA, в котором было набрано порядка 600 млн событий столкновения ионов ксенона с мишенью цезий–йод, занявших около 400 ТБ дискового пространства. Для многократной обработки и проведения необходимого физического анализа такого большого объема данных требуется комплексная архитектура программных решений для распределенной обработки и хранения данных. Представлены методика разработки и архитектура комплекса программных систем для организации распределенной обработки данных эксперимента и унифицированного хранения и управления информацией, требующейся для проведения обработки на всех этапах. Реализованная в соответствии с методикой архитектура включает в себя как программные системы, решающие задачу объединения распределенных ресурсов эксперимента в единую систему обработки с единым пространством

---

\* E-mail: gertsen@jinr.ru

хранения для автоматизации выполнения потока задач, так и оригинальные информационные системы, обеспечивающие сбор, хранение и организацию доступа к информации, необходимой для обработки и анализа полученных данных. Кроме того, набор вспомогательных программных решений, таких как система дистрибуции программного обеспечения, система единой аутентификации и авторизации и система мониторинга программных систем, повышает эффективность и надежность разработанной архитектуры.

PACS: 07.05.—t; 07.05.Bx; 07.05.Kf