For the JINR Scientific Papers Competition for 2024

SUBMISSION

Submitted work:

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

Section: **Scientific methodical and technical works**

Team of co-authors:

1. Alexandrov E. I.
2. Alexandrov I. N.
3. Balashov N. A.
4. Gertsenberger K. V.
5. Moshkin A. A.
6. Pelevanyuk I. S.
7. Filozova I. A.
8. Chebotov A. I.
9. Shestakova G. V.
10. Klimai P. A.

The presented series of the works includes 25 publications:

1. K. Gertsenberger, I. Pelevanyuk, P. Klimai, and A. Chebotov, "*Computing Software Architecture for the BM@N Experiment*", Phys. Part. Nuclei **55**, 338–342 (2024)
2. K. Gertsenberger, I. Pelevanyuk, "*BM@N Run 8 Data Processing on a Distributed Infrastructure with DIRAC*", Phys. Part. Nucl. Lett. **21**. 778–781 (2024)
3. N. Balashov, "*JINR Container Distribution Service*", Phys. Part. Nuclei **55**, 482–484 (2024)
4. K. Gertsenberger, P. Klimai, O. Nemova, "*Development of Monitoring Service for BM@N Information Systems*", Phys. Part. Nucl. Lett. **21**, 793–796 (2024)
5. E. Alexandrov, I. Alexandrov, A. Chebotov, K. Gertsenberger, I. Filozova, D. Priakhina, G. Shestakova, and A. Yakovlev, "*Development of the Online Configuration System for the BM@N Experiment* ", Phys. Part. Nuclei **55**, 433–436 (2024)
6. E. Alexandrov, I. Alexandrov, A. Chebotov, A. Degtyarev, I. Filozova, K. Gertsenberger, P. Klimai, and A. Yakovlev, "*Implementation of the Event Metadata System for physics analysis in the NICA experiments*", Journal of Physics: Conference Series **2438**, 012046 (2023)
7. E. Alexandrov, I. Alexandrov, A. Chebotov, K. Gertsenberger, I. Filozova., D. Priakhina, and G. Shestakova, "*Configuration Information System for online processing and data monitoring in the NICA experiments*", Journal of Physics: Conference Series **2438**, 012019 (2023)

8. A. Chebotov, A. Degtyarev, K. Gertsenberger, and P. Klimai, “*REST API and Web Interface for the Event Metadata System of the BM@N Experiment*”, Phys. Part. Nucl. Lett. **20**, 1527–1530 (2023)
9. A. Chebotov, K. Gertsenberger, A. Moshkin, and I. Slepov, “*Common Deployment Complex for the Information Systems of the BM@N Experiment*”, Phys. Part. Nucl. Lett. **20**, 1269–1271 (2023)
10. K. Gertsenberger, P. Klimai, M. Zelenyi, “*Auxiliary Services for the Condition Database of the BM@N Experiment at NICA*”, Phys. Part. Nucl. Lett. **20**, 1217–1219 (2023)
11. E. Alexandrov, I. Alexandrov, A. Chebotov, K. Gertsenberger, I. Filozova, D. Priakhina, and G. Shestakova, “*Status of the Configuration Information System for the NICA experiments* “, Phys. Part. Nucl. Lett. **19**, 543–546 (2022)
12. A. Degtyarev, K. Gertsenberger, P. Klimai, “*Usage of Apache Cassandra for Prototyping the Event Metadata System of the NICA Experiments*”, Phys. Part. Nucl. Lett. **19**, 562–565 (2022)
13. A. Chebotov, K. Gertsenberger, P. Klimai, and A. Moshkin, “*Information System Based on the Condition Database for the NICA Experiments, User WEB Application, and Related Services*”, Phys. Part. Nucl. Lett. **19**, 558–561 (2022)
14. K. Gertsenberger, I. Alexandrov, I. Filozova, E. Alexandrov, A. Moshkin, A. Chebotov, M. Mineev, D. Pryahina, G. Shestakova, A. Yakovlev, A. Nozik, and P. Klimai, “*Development of Information Systems for Online and Offline Data Processing in the NICA Experiments*”, Phys. Part. Nuclei **52**, 801-807 (2021)
15. A. Chebotov, K. Gertsenberger, I. Slepov, and A. Moshkin, “*Electronic Logbook platform for NICA experiments*”, AIP Conf. Proc. **2377**, 040003 (2021)
16. E. Akishina, E. Alexandrov, I. Alexandrov, I. Filozova, K. Gertsenberger, and V. Ivanov, “*Development of a Geometry Database and Related Services for the NICA experiments*”, Phys. Part. Nuclei **52**, 842-846 (2021)
17. E. Alexandrov, I. Alexandrov, A. Degtyarev, K. Gertsenberger, I. Filozova, P. Klimai, A. Nozik, and A. Yakovlev, “*Design of the Event Metadata System for the Experiments at NICA*”, Phys. Part. Nucl. Lett. **18**, 603-616 (2021)
18. K. Gertsenberger, A. Chebotov, P. Klimai, I. Alexandrov, E. Alexandrov, I. Filozova, and A. Moshkin, “*Implementation of the Condition Database for the Experiments of the NICA Complex*”, CEUR Workshop Proceedings **3041**, 128–132 (2021)
19. E. Akishina, E. Alexandrov, I. Alexandrov, I. Filozova, K. Gertsenberger, V. Ivanov, D. Priakhina, and G. Shestakova, “*Development of the Geometry Database for the BM@N Experiment of the NICA Project*”, EPJ Web Conf. **226**, 03001 (2020)
20. K. Gertsenberger, A. Chebotov, I. Alexandrov, I. Filozova., and E. Alexandrov, “*Design of the Condition Database for online and offline data processing in experimental setups of the NICA complex*” (in russian), Izvestiya SFedU. Engineering Sciences **217**, no.7, 172-180 (2020)
21. A. Chebotov, K. Gertsenberger, “*Development of web-service for Unified Database of the BM@N experiment at NICA*”, AIP Conf. Proc. **2163**, 040002 (2019)
22. K. Gertsenberger, A. Moshkin, A. Chebotov, “*Development of the Electronic Logbook for the BM@N Experiment at NICA*”, CEUR Workshop Proceedings **2507**, 175–179 (2019)
23. E. Alexandrov, I. Alexandrov, K. Gertsenberger, M. Mineev, A. Moshkin, D. Pryahina, I. Filozova, A. Chebotov, G. Shestakova, and A. Yakovlev, “*Information Systems for Online and Offline Data Processing in Modern High-energy Physics Experiments*” (in

russian), International scientific journal «Modern Information Technologies and IT-Education» **15**, no.3, 654-671 (2019)

24. K. Gertsenberger, O. Rogachevsky, "*The Unified Database for BM@N experiment data handling*", EPJ Web Conf. **177**, 05001 (2018)
25. E. Akishina, E. Alexandrov, I. Alexandrov, I. Filozova, V. Friese, K. Gertsenberger, V. Ivanov, and O. Rogachevsky, "*Geometry Database for the CBM experiment and its first application to the experiments of the NICA project*", CEUR Workshop Proceedings **2267**, 504–508 (2018)

The work submitted for the competition is presented in the form of a series of the articles published from 2018 to 2024 in peer-reviewed scientific journals within the framework of Topic 1065: "NICA Complex: creation of a complex of accelerators, a collider and experimental facilities on colliding and extracted ion beams for studying dense baryonic matter, the spin structure of nucleons and light nuclei, and conducting applied and innovative works", 02-1-1065-2007/2026, BM@N project (02-1-1065-2-2012/2026).

The BM@N experiment is the first, already ongoing experiment at the NICA accelerator-storage complex, designed to study the interaction of beams of relativistic heavy ions with energies up to 6 GeV per nucleon with fixed targets. The physics program of the experiment is aimed at studying dense baryonic matter produced as a result of such collisions, including study of the equation of state of matter at extreme densities, matter with strong isospin asymmetry, mixed phase at a first-order phase transition, the production of hyperons and hypernuclei in the given energy range, hadron femtoscopy, and event-by-event fluctuations.

Since 2015, seven technical runs had been conducted in which beams of deuterons, carbon, argon and krypton ions collided with different types of targets, and in the winter of 2022-2023 the first physics run of the experiment was successfully conducted, in which beams of xenon ions collided with a cesium-iodine target at the energies of 3.8 and 3 GeV per nucleon. Only in the last run, about 600 million events were collected, which occupied about 400 TB of raw data. Moreover, when the experiment reaches the design parameters, the volume of data obtained is expected to increase by an order of magnitude. Since all particle collision events of the experiment with such a large volume of data must be quickly processed and necessary physics analysis must be performed, then the realization of the BM@N physics program requires a comprehensive approach with a properly designed software architecture for distributed data processing on the computing platforms provided to the experiment. An important role in the architecture, in addition to the systems and services directly related to distributed processing, is played by information systems that ensure the collection, storage, organization of convenient access and management of information being necessary for processing and analyzing data of the experiment.

The results obtained within the framework of the presented cycle of the works:

1. Developed architecture of the software complex for distributed data processing at the BM@N experiment [1]. In order to determine which software systems had to be implemented to automate the execution of tasks of big data processing in the experiment, at the first stage a study of the experimental data processing flow of the facility was conducted (Fig. 1), and event data processing model was developed for BM@N. As a result of the analysis, necessary software systems and services were determined, including information systems required for data processing both online, that is during the experiment runs, for quality assurance of incoming data, and after the runs (offline) for data decoding, event reconstruction and physics analysis.

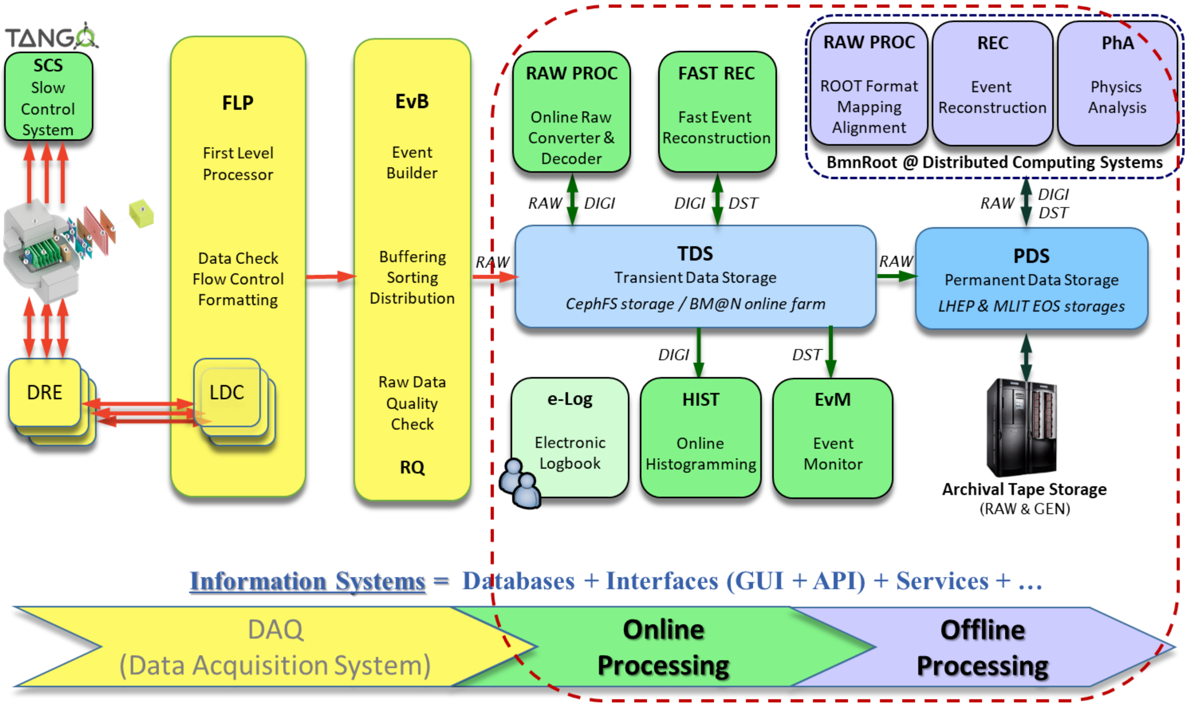

Fig. 1. Schematic diagram of BM@N data processing flow (the part that has been automated by the team of coauthors is highlighted by the red dotted rectangle)

As a result, to solve the issue of fast and high-quality processing of both experimental and simulated data of BM@N collision events with the purpose of timely obtaining scientific results, a comprehensive architecture of the software complex for distributed BM@N data processing has been developed (Fig. 2).

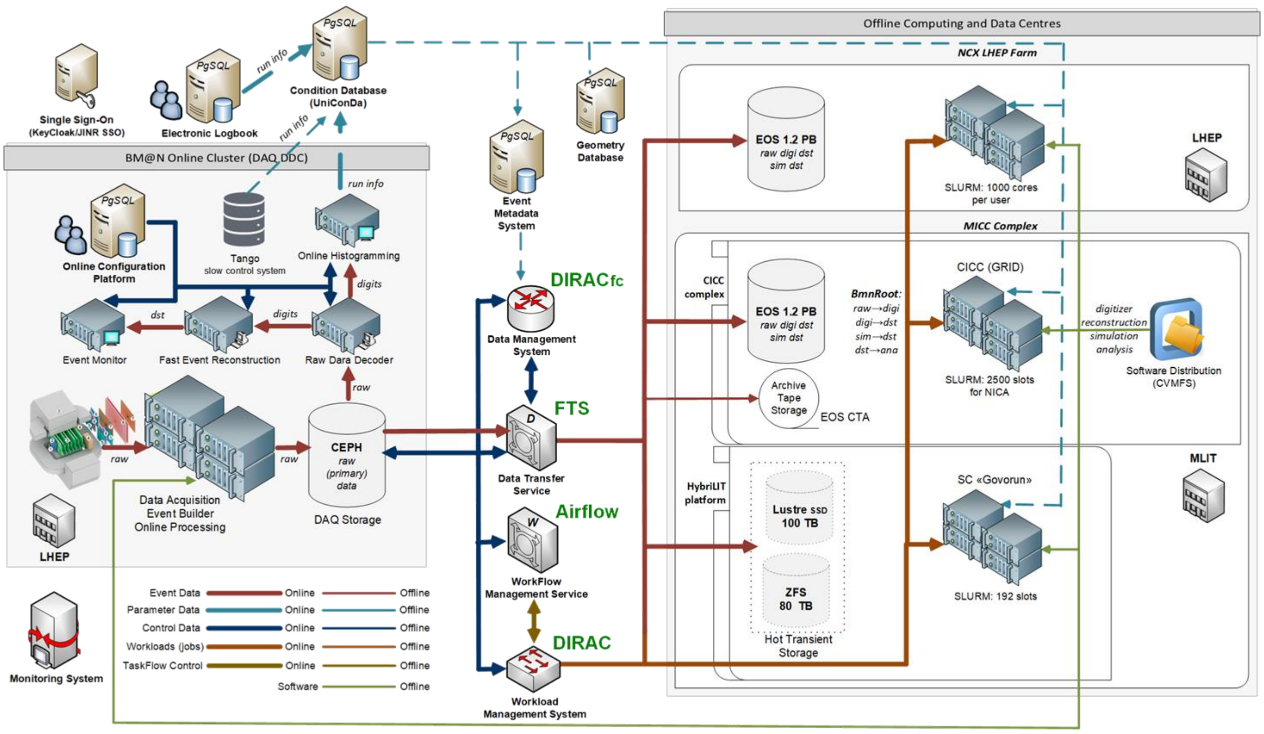

Fig. 2. Developed architecture of the software complex for distributed data processing and storage at the BM@N experiment

The peculiarity of the hardware computing platforms provided for the experiment, namely: the NICA cluster at LHEP, the Central Information and Computing Complex (CICC) of LIT, the HybriLIT platform with the Govorun supercomputer and BM@N online cluster (DAQ Data Center, DDC), is the separation of their computing resources and data storages. To solve the issue, the presented architecture includes a set of software systems for combining all the distributed computing and storage resources into a single computing system with a single storage space, ensuring management of big data processing flows on all the available resources, thereby reducing the time required to obtain physics results. The complex contains the Workload Management System using the DIRAC [A] platform, Data Management System implemented through the DIRAC File Catalog, Data Transfer Service and Workflow Management Service built on the Apache Airflow [B] solution.

An important point for data processing both during and after the experiment runs is the need to use a large number of various parameters and information on the experiment systems, therefore the developed architecture also contains a complex of implemented information systems and services that ensure high-quality management and provision of the information required for processing to BM@N software systems for its use at all stages, including raw data decoding, event reconstruction and physics analysis, as well as detector simulation. The Complex [14, 23] includes the Electronic Logbook System, Configuration Online Platform, Geometry Information System, an information system based on the Condition Database, and the Event Metadata System. The developed information systems are built on modern databases and provide user services for transparent access and management of stored data and information on the experiment, allowing for simultaneous servicing of a large number of requests from software systems and users, and also provide automatic regular data backup in case of software or hardware failures. In addition, the software systems shown in the figure, such as the software distribution system, unified authentication and authorization service, as well as monitoring service improve the efficiency and reliability of the developed software architecture.

2. The complex of the software systems for automating the execution of distributed data processing tasks in the BM@N experiment.

2.1. The central Workload Management System using DIRAC Interware for distributed BM@N data processing [1, 2]. For distributed processing of experimental and simulated data of the BM@N experiment, the DIRAC Interware platform has been integrated and is successfully used as a central workload manager. It provides necessary components for building a data processing infrastructure interconnecting computing resources of different types. The DIRAC system is deployed on the JINR computing infrastructure and combines the resources of the NICA cluster, the Tier1 and Tier2 centers of the CICC, the Govorun supercomputer, and the BM@N online cluster, which is successfully used for offline data processing between experiment runs. In addition to computing resources, the BM@N experiment employs storage systems integrated into DIRAC: its storage element on the EOS file system and tape storage with the CTA (CERN Tape Archive) solution. Moreover, when the collaboration gains cloud resources or computing resources of the external collaborators, DIRAC will make it easy to add them to the single processing system. The central Workload Management System provides graphical Web and command line interfaces used to run and manage thousands of tasks for processing data of the experiment on all the provided computing resources.

As a result, the DIRAC workload manager is actively used by the collaboration for mass processing of experimental and simulated data of the BM@N experiment, allowing, for instance, both to decode all the raw data of the last run for 35 hours and to reconstruct the particle collision events with a beam energy of 3.8 GeV per nucleon for 7 days. From the moment of receiving BM@N data of the last run, more than 200 years of processor time, calculated per one computing core, have been used for tasks related to processing the data (Fig. 3). The system also demonstrates a high data transfer rate over the network using the XRootD protocol, which reaches up to 2 GB/s, and copes with a peak load of almost 8 GB/s.

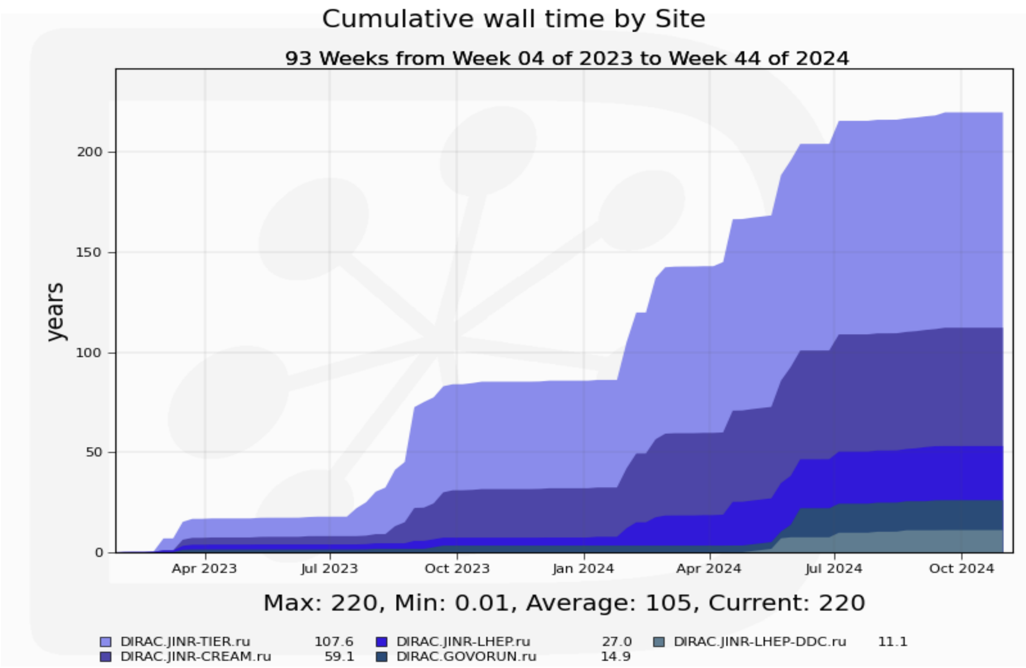

Fig. 3. Cumulative data processing time of the last run on the CICC Tier1 and Tier2 centers (DIRAC.JINR-TIER.ru and DIRAC.JINR-CREAM.ru), NICA cluster (DIRAC.JINR-LHEP.ru), Govorun supercomputer (DIRAC.GOVORUN.ru), and online cluster (DIRAC.JINR-LHEP-DDC.ru)

Since February 2023, 13 large campaigns have been conducted to process and generate data for the BM@N collaboration, and more than 20 test runs have been performed on small data samples. Monitoring and analytics systems have been tested, which allow comparing data processing campaigns, monitoring the correctness of the software operation, and analyzing the efficiency of resource use.

2.2. The DIRAC File Catalog for organizing a single storage space in the BM@N experiment [1, 2]. The issue of using separated data storages in the experiment led to the need for a specialized data management system, also called a single file catalog system, which maps logical filenames to physical ones on the data storages, allowing users to work with the separated storages as with a single file catalog. To solve the problem, a solution based on the built-in file catalog of the DIRAC platform, DIRAC File Catalog has been successfully implemented for the BM@N experiment. The integrated file catalog provides wide functionality, including data replication and use of file metadata, i.e. attributes containing summary information on a file.

The deployed BM@N file catalog provides a single file namespace used for distributed data processing and currently contains a list of reconstructed files of the physics run with a large set of relevant metadata, such as a run number, acquisition time, beam and target particle types, beam energy, magnetic field value, a number of events in the file, and others. The developed application programming interface (Fig. 4) in the form of a service on the REST [F] architecture provides the ability to search using the specified metadata and obtain a list of only those files that are necessary for a particular physics analysis of the reconstructed data.

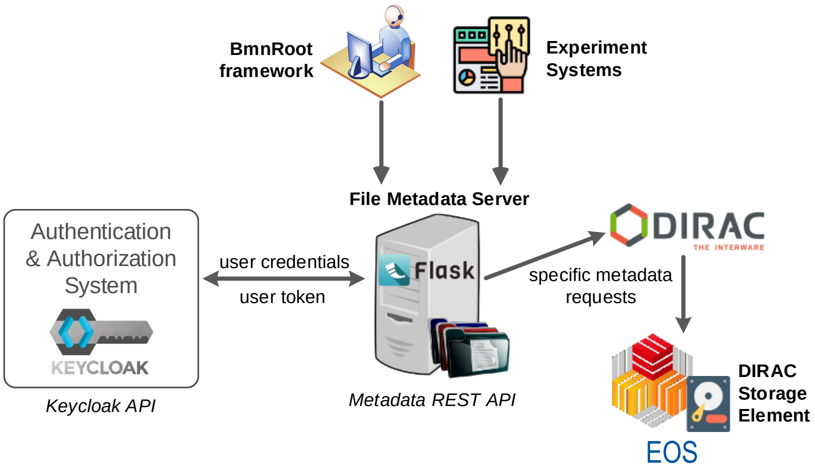

Fig. 4. Using metadata REST API to process requests to the DIRAC File Catalog

2.3. Workflow Management Service based on Apache Airflow for the BM@N experiment [1]. To automate the execution of recurring sets of tasks for processing big data of the BM@N experiment, the Workflow Management Service of the BM@N experiment has been implemented based on the Apache Airflow solution, which provides a platform for creating, scheduling and monitoring the execution of a predefined set of workflows. Such solutions are also called orchestration systems, because they manage the execution of recurring sets of tasks using various software systems of the experiment, combining them together into a single execution flow. In accordance with a described set of tasks with dependencies between them in the form of a Directed Acyclic Graph, the manager forms chains of workflows required for the data processing, starts, manages and monitors their execution (Fig. 5).

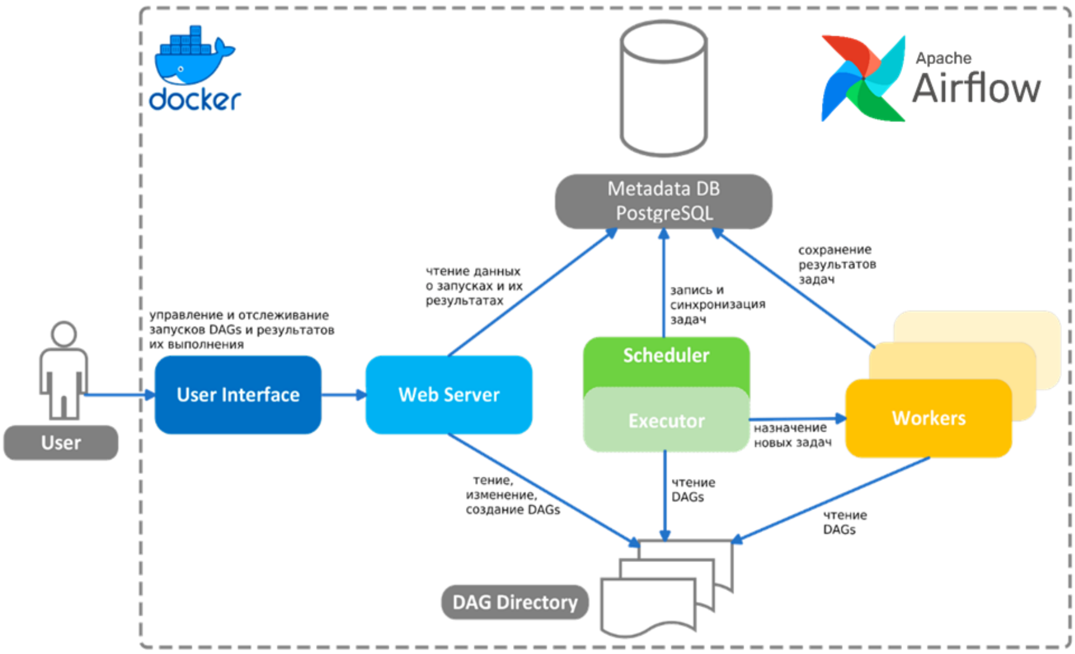

Fig. 5. Interaction of the Workflow Management Service components

The deployed solution was already used during the last run to automatically transfer incoming experimental data from the online DDC farm to the main offline storages at JINR LHEP and LIT with integrity checking of the transferred files. It is planned to expand the use of the orchestration system to automatically run mass processing of incoming data via the DIRAC platform, as well as automated archiving of raw experimental data to the reliable BM@N tape storage at LIT.

2.4. Software build, testing and distribution system for the BM@N experiment [1, 3]. To ensure the availability of the BM@N software on all the computing platforms, two software storages have been organized: a central software repository based on the CernVM-FS network file system [C], and a container registry. Containerization of the runtime environment is used for the BM@N experiment to simplify the process of preparing applications for operation in a heterogeneous software environment on the working nodes of the computing clusters and personal computers. In addition, isolation of the runtime environment from the system software of the computing nodes reduces administrative overhead for the computing infrastructure, since the support of the containerization system is only required.

The software build, testing and distribution system has been implemented using GitLab CI/CD [D] tools: a new pipeline has been developed, containing tasks that automatically perform these operations in automatic mode whenever changes are made to the main software framework of the BM@N experiment – BmnRoot. Since all the tasks are very resource-intensive, the JINR cloud is employed as a computing resource to perform them. The pipeline also builds and publishes BmnRoot container images in Docker format [E], which can be used for BM@N data processing on the computing clusters (converted to Apptainer format) and on personal computers. Both container images and software are distributed through the central software repository hosted in CernVM-FS, which is used for distributed data processing by all the computing platforms.

3. Developed complex of the information systems that ensure the collection, storage and provision of information required for BM@N data processing.

3.1. The Electronic Logbook System [15, 22]. For the interpretation of data collected during experiment runs, not only the experimental data acquired from the detectors are important, but also logbook entries recorded by shift operators, which describe the operation modes and states of the various subsystems, detectors, and the obtained events. To solve the task, the Electronic Logbook System has been developed and is successfully used. It provides shift operators with a web interface (*bmn-elog.jinr.ru*) for recording information on the current parameters and operation modes of the subsystems, the current events, occurred problems and taken actions during experiment runs. It is also designed for convenient viewing, changing and searching for the required information in the logbook by the collaboration members. The Electronic Logbook ensures automatic regular backup of the logbook data and provides the following functions depending on the role of the authenticated user: viewing, adding or changing records, attaching text and graphic files, sorting and filtering the data, convenient search by parameters, as well as an important function, being especially critical during runs, – subscribing to notifications for certain types of events.

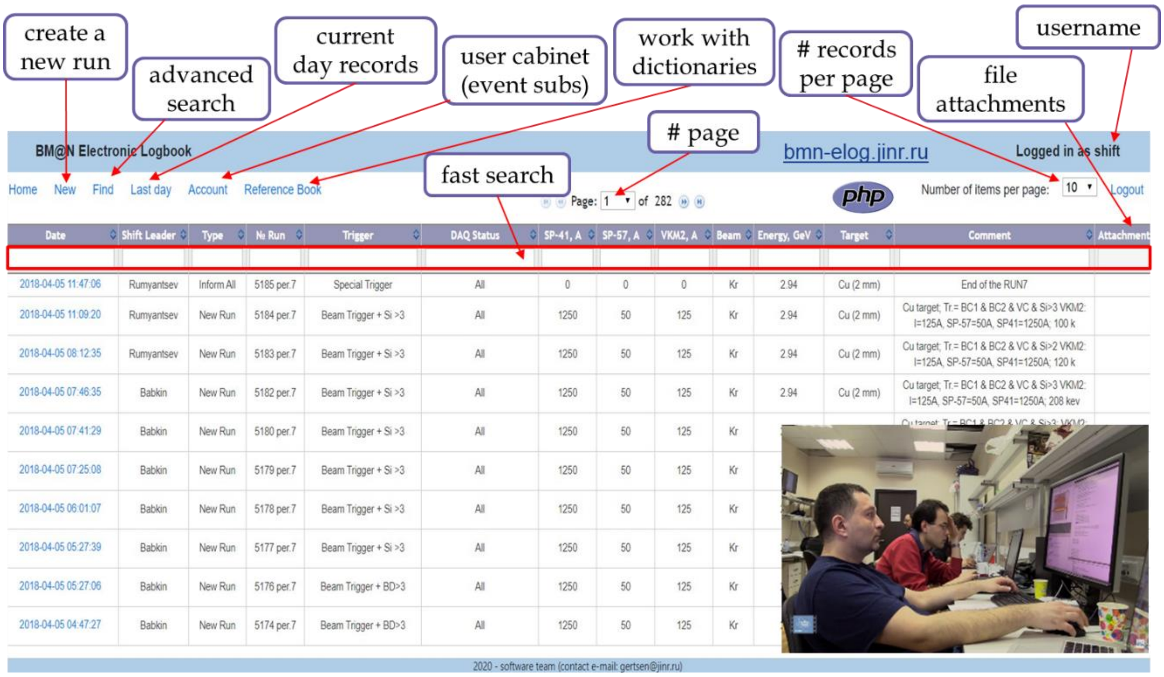

Fig. 6. Web interface of the Electronic Logbook for the BM@N experiment with the marked functionality (in the photo – BM@N shift operators)

The implemented Electronic Logbook System includes a database for storing the experiment logbook and working with it, and for subsequent use of the logbook data by other software systems, including algorithms for decoding, processing and analyzing particle collision events, an application programming interface, as well as a set of auxiliary services, have been developed. For instance, one of the services made it possible to automatically generate daily statistics on the collected particle collision events of different types in the last run of the BM@N experiment.

3.2. The Configuration Online Platform of the BM@N experiment [5, 7, 11]. Another problem that required the implementation of a new software system is the configuration and management of a required set of online processing tasks during the experiment runs. To check the quality of incoming experimental data, the software tasks for raw data decoding, online histogramming, fast event reconstruction, and event monitoring must be continuously executed online and exchange data with each other. In general, it is possible to manually control one online process, but if there is the full set of the necessary processes that must work simultaneously on distributed resources and dependent on each other, then it is impossible to effectively control them manually. To automate the online processing of BM@N data, the Configuration Online Platform (Fig. 7) has been implemented to store and provide data on the configuration of BM@N hardware systems and software tasks to be executed online.

The developed Configuration Database stores both a set of necessary configuration parameters and a sequence of tasks that must be run and executed during an experiment session. The operator describes the required online tasks in the Configuration Designer, the system loads the sequence into the Configuration Database, and the Central Manager reads the generated topology, starts all the necessary processes with required parameters on specified distributed nodes, manages the processes, and automatically restarts them in case of failure. The DDS (Dynamic Deployment System) [G] was selected to run online processing tasks on distributed resources and manage their execution, as well as to ensure their interaction with each other. The Web interface (*bmn-*

online.jinr.ru) of the platform, in addition to conveniently defining and managing the process topology, includes the monitor of active tasks, providing information on their states and output.

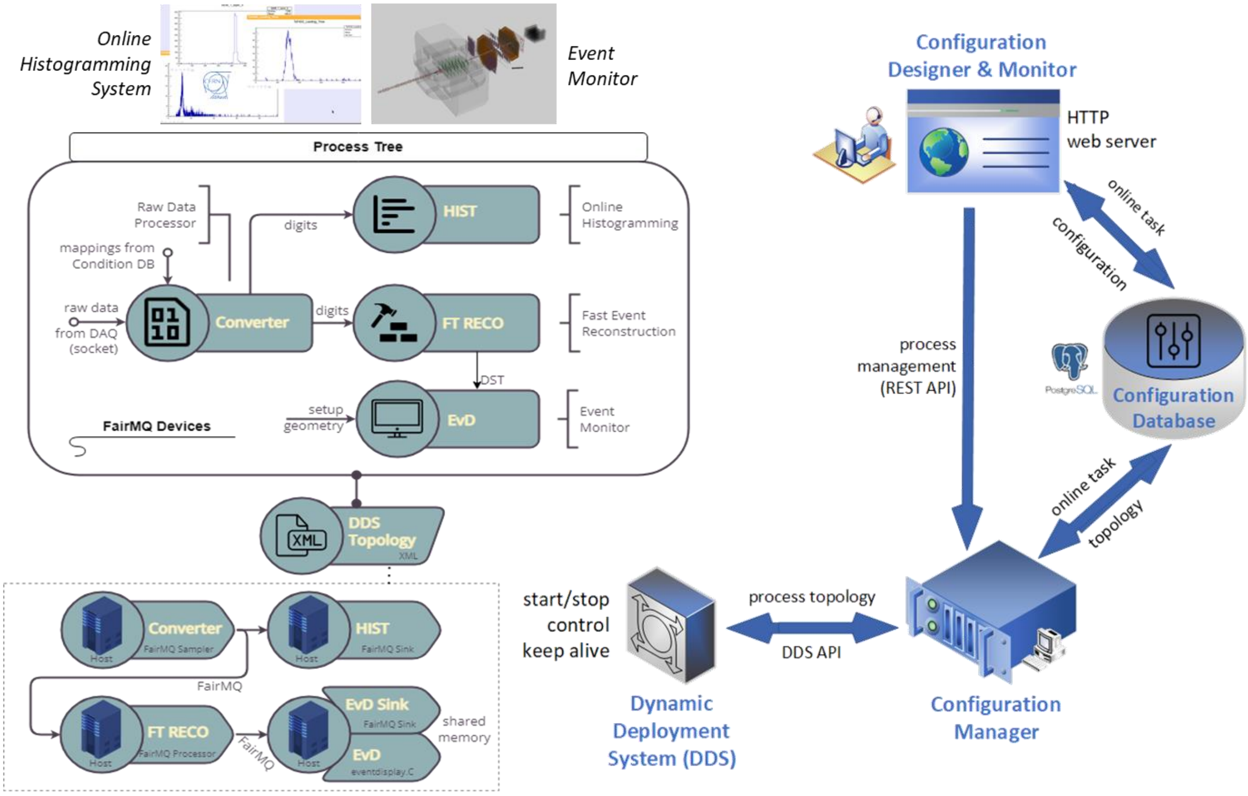

Fig. 7. BM@N data quality assurance using the Configuration Online Platform

3.3. The Condition Database of the BM@N experiment [13, 18, 20, 21, 24]. Another important problem in solving the tasks of processing the obtained data is the need to use a large number of various parameters of the experimental subsystems, therefore, within the framework of the architecture, an information system based on the parametric database, which is also called the Condition Database, has been implemented. The system has been developed and is actively used, providing collaboration members with unified access, search and management of parametric information on the experiment, necessary for processing experimental and simulated data, and ensuring relevance, consistency and automatic backup of the data. An important property of the stored parametric data is that the parameters are characterized by a time interval of their validity, so they are used to process data collected only during the validity period of the corresponding parameter. Furthermore, the developed architecture provides storing parameters of arbitrary structure.

The BM@N Condition Database includes the following main parts: storing information on simulation files obtained by event generators; on conducted runs of the experiment (run metadata) and the corresponding data files, as well as the main part – storing values of required parameters (configuration, calibration, algorithmic and others) of an arbitrary format, which are necessary for further data processing. For convenient viewing of the parametric data by collaboration members through a browser, a special web interface (*bmn-uniconda.jinr.ru*) has been developed, both displaying summary statistics on the stored data (Fig. 8) and providing a convenient tabular form for viewing, managing and searching for parametric information on the experiment. But the main implemented interfaces of the Condition Database are application programming interfaces (REST

API and C++), which are used, for example, to obtain parameters in BmnRoot necessary for simulation, raw data decoding, event reconstructions and physics analyses.

In addition to the development of comprehensive software systems within the architecture, a large set of various auxiliary services has been also implemented in the experiment [10]. As an example, a service (File Inspector) for checking the availability and integrity of experimental files and files with simulated events described in the Condition Database has been developed in Python. This is due to the large amount of data and storage time, as well as the need to quickly respond to data loss or damage. The developed service regularly checks files with BM@N events on distributed clusters: it verifies that the files are present, readable and have not been damaged. All information on these checks is displayed on the centralized website of the File Inspector.

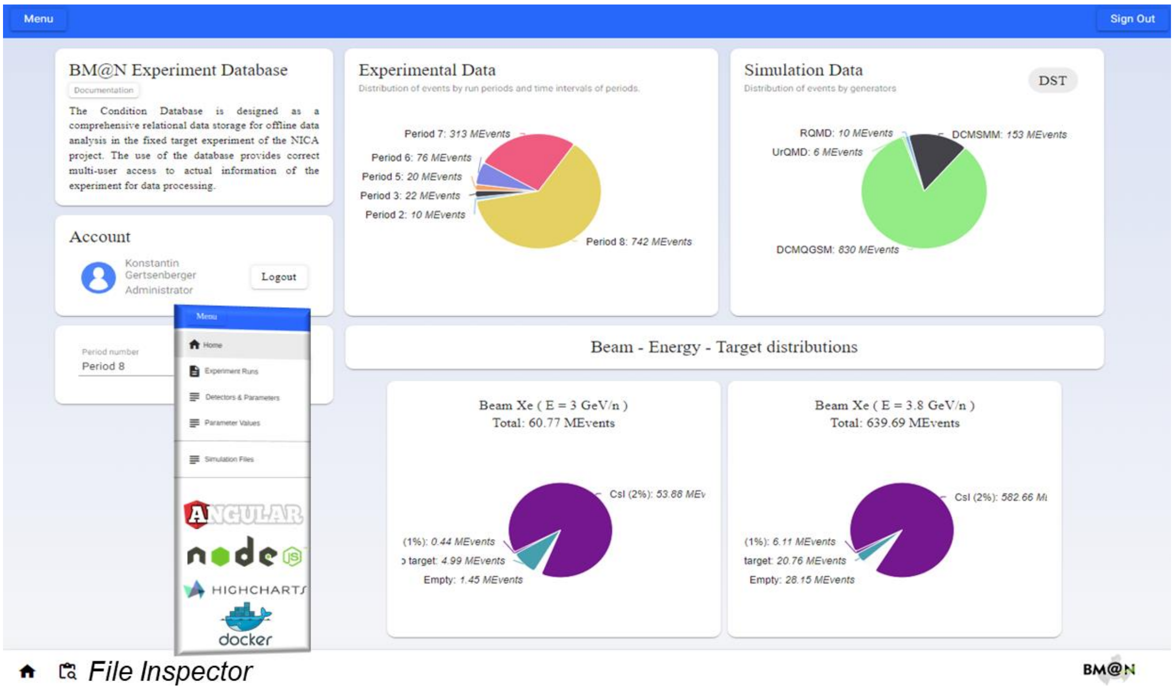

Fig. 8. Web interface of the BM@N Condition Database

3.4. Geometry Information System of the BM@N experiment [16, 19, 25]. The specified versions of the geometric components of the BM@N setup, which define the implementation of geometry objects in the software of the experiment, are of great importance for processing the data of particle collision events. Therefore, another implemented software solution is the information system for working with detector geometries of the setup, based on the developed Geometry Database. The Geometry System has been designed to store and manage information on the geometric models of the detectors, which are then used to process and analyze simulated and experimental data of the experiment. For each geometry module, an identifier, version, transformation matrix and a link to the parent module are stored. The versions of the full setup geometry are defined as a combination of the constituent geometry modules of the detectors, descriptions of the magnetic field, used materials and media, which are stored in the database. The developed Geometry Information System provides a centralized repository of the geometry information and a set of convenient tools for managing both different versions of individual geometry modules and assemblies of versions of the full setup.

The Geometry Information System is built using "client-server" architecture (Fig. 9), where the server part is represented by the central storage of geometries on the PostgreSQL DBMS, providing all the functionality, and the client part is implemented to work with local user replicas (on the SQLite DBMS) of the central database. The main interfaces for working with the Geometry System are implemented: an application programming interface to select and load setup geometries and its components into the software framework for their using in the algorithms of simulation, reconstruction and physics analysis of the obtained data, as well as a graphical interface in the form of the Web service (*bmn-geometry.jinr.ru*) for access to all functions of the system depending on the user category.

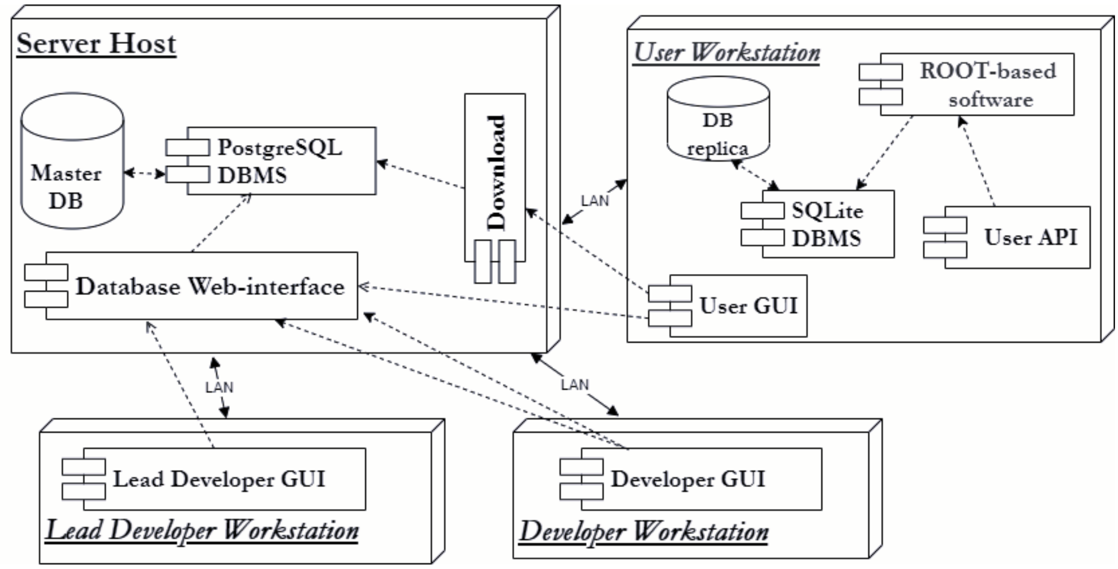

Fig. 9. Architecture of the BM@N Geometry Database

3.5. The Event Metadata System (Event Catalog) of the experiment [6, 8, 12, 17]. Another urgent issue solved in the BM@N experiment is the search and selection of only those particle collision events that are required to perform a particular physics analysis, for which purpose, the Event Metadata System has been implemented that contains summary parameters of the events, for example, a number of reconstructed tracks, which are called metadata and required for the selection. The system is based on a developed database, called the Event Catalog, for storing event metadata, versions of programs used in processing, location of the events at distributed storages, and allows users to quickly search using the parameters and receive a set of only those events that are required for a specific physics analysis. The corresponding events are identified by a unique link, which is a combination of a pointer to a data file in a distributed storage and an event number in the file.

The architecture of the Event Metadata System is quite comprehensive (Fig. 10) and includes, in particular, the developed user web interface (*bmn-event.jinr.ru*) for viewing and searching the event metadata stored in the catalog by collaboration members, as well as for requesting and receiving only necessary events, that satisfy the specified parameters, for their physics analysis. To improve the efficiency of the event selection process, information on experiment runs stored in the Condition Database is used, if possible, for preliminary selection of events before searching in the Event Catalog. An application programming interface has been implemented in the form of a REST API service, which is used to record information on new events received in the experiment and to execute requests from users and other software systems for obtaining necessary events using various search criteria. The implemented system is configurable to support an arbitrary set of event metadata for an experiment, and provides access control based on the roles of collaboration members and monitoring the operation of the system components.

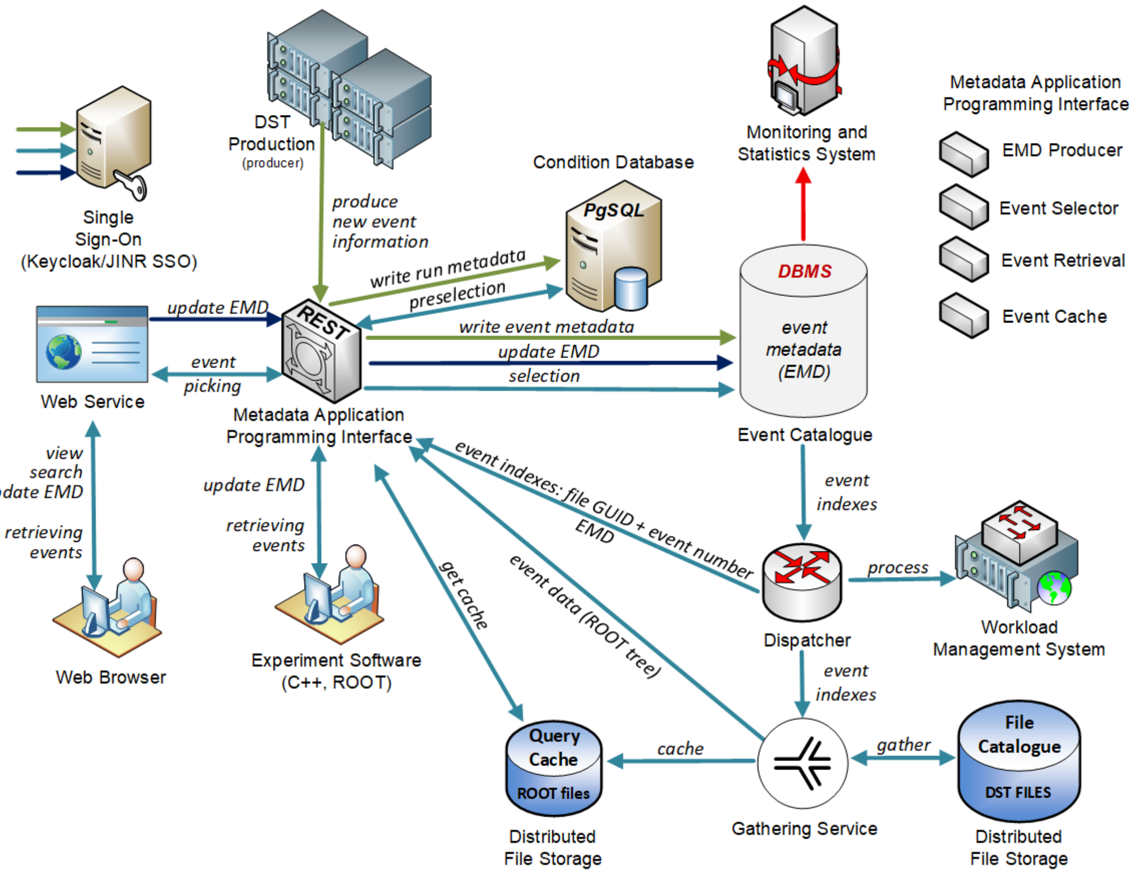

Fig. 10. Architecture of the BM@N Event Metadata System

3.6. Common deployment service for the developed information systems [9]. The information systems presented above may be in demand in other particle collision experiments, so they were made configurable for use in experiments with similar requirements. Moreover, a common configuration and deployment service has been implemented, providing for customized deployment of the systems. The developed information systems are configured using external parameters to set up both simple elements, including the logo and contact details, and more complex ones, such as fields in the electronic logbook or composition of the event metadata of the Event Catalog. The main concept of the created deployment system is to employ a configuration file with customizable parameters and one installation script that deploys using a multi-container architecture on several servers and configures all the components of the information system in accordance with the specifics of the experiment. The result of the configurable deployment is the created database, interfaces and services deployed in Docker containers on specified machines, as well as automated regular updates from the repository and backup of the stored data, after which collaborators can start working with the out-of-box system.

4. Developed auxiliary services improving the efficiency and reliability of the implemented architecture.

4.1. Single authentication and authorization system [1]. The developed architecture of the software complex for organizing distributed data processing at the BM@N experiment also includes a set of important auxiliary systems of the experiment, one of which is a centralized authentication and authorization system for administration and delimitation of rights of BM@N participants in the software systems and collaboration services. The developed system is based on the modern Keycloak solution, which allows the participants to have one account for all the software systems of the experiment, as well as to use all the systems without re-entering account data. It provides centralized authentication and authorization and stores information on user accounts and groups that delimit the access of the collaboration members and determine available actions. As a result, requests for data of the BM@N information systems go through the developed interfaces interacting with the solution to authenticate the user and check his role and permissions for the requested operation.

4.2. Monitoring service for the software systems of the experiment [1, 4]. In order to minimize the response time in case of hardware or software failures of the implemented software systems of the experiment, it is important to continuously monitor their states, which is especially critical during BM@N runs. To solve the task, the monitoring service has been implemented, designed to track the existing BM@N software infrastructure, including the active servers, databases and web interfaces of the systems, to store their state parameters in the time-series database and visualize them using the Grafana [H] system on the JINR central web service (Fig. 11). In case of failures in operation or exceeding critical values of the tracking parameters, the monitoring service sends corresponding notifications to responsible persons, who need to take recovery actions, using specified email addresses and Telegram channel.

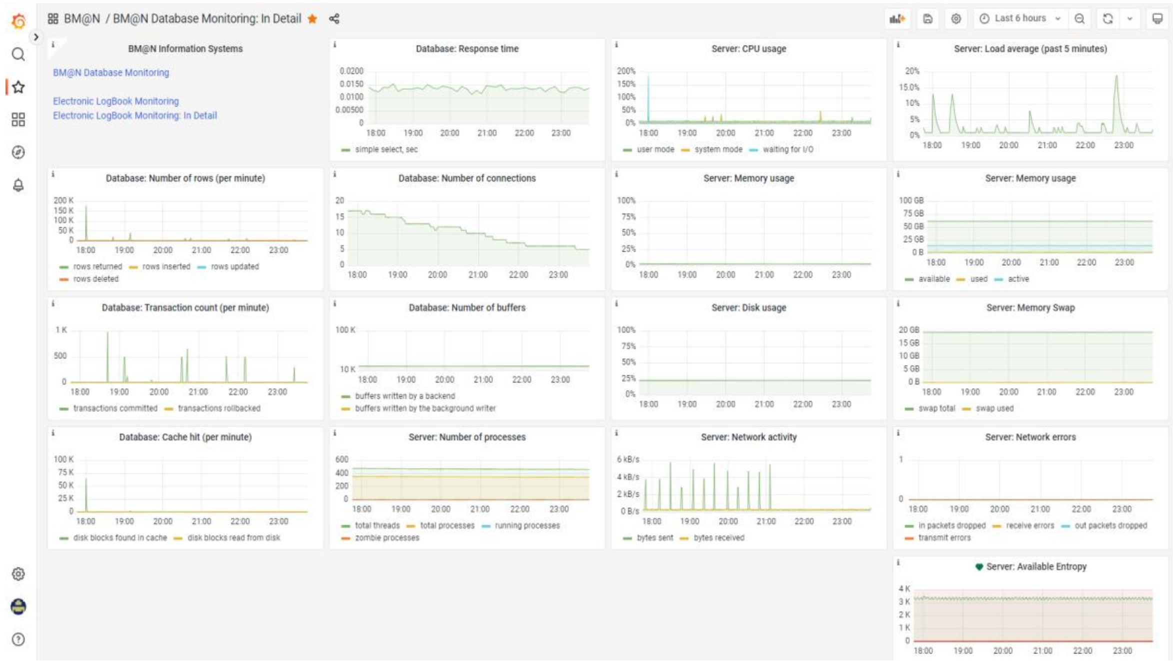

Fig. 11. Web interface of the software monitoring service for the BM@N experiment

Thus, within the framework of the work submitted for the competition, the team of coauthors has designed and implemented the architecture that includes a complex of modern software systems for organizing distributed processing of BM@N experiment data on the provided computing platforms, as well as for storage, management and unified access to information required for event data processing at all stages, including raw data decoding, event reconstruction and physics analysis, as well as detector simulation.

The implemented architecture includes both software systems, such as the Workload Manager, Single File Catalog and 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 processing task flows, and original information systems, such as the Electronic Logbook, Configuration Online Platform, Geometry Information System, Condition Database and Event Metadata System, that provide collection, storage, management, and organization of access to information necessary for processing and analyzing the obtained data throughout the life cycle of scientific research at the BM@N experiment. In addition, a set of auxiliary services, such as the software distribution system, single authentication and authorization service, and monitoring service, improves the efficiency and reliability of the developed software architecture.

The implemented software complex has been deployed on the existing BM@N infrastructure, successfully used for distributed processing, storage and physics analysis of collected experimental (as well as simulated) data, and it is also an essential element for high-quality data management and timely obtaining of physics results in terms of working with big data flows of the BM@N experiment.

The obtained results have been presented at the collaboration meetings of the BM@N experiment, the RDIG-M consortium meeting, and in 25 reports at scientific conferences.

References

[A] F. Stagni, A. Tsaregorodtsev, A. Sailer and C. Haen, *"The DIRAC interware: current, upcoming and planned capabilities and technologies",* EPJ Web Conf. **245**, 03035 (2020)

[B] B. Harenslak and J. Rutger de Ruiter, Data Pipelines with Apache Airflow, 480 p. (2021)

[C] J. Blomer, P. Buncic, R. Meusel, G. Ganis, I. Sfiligoi and D. Thain, "*The evolution of global scale filesystems for scientific software distribution*", CiSE **17**, no.6, 61-71 (2015)

[D] Ya. Jani, "*Implementing continuous integration and continuous deployment (ci/cd) in modern software development*", IJSR **12**, no.6, 2984-2987 (2023)

[E] D. Moreau, K. Wiebels and C. Boettiger, "*Containers for computational reproducibility*" Nat Rev Methods Primers **3**, no.1, 50 (2023)

[F] C. Pautasso, E. Wilde and R. Alarcon, REST: Advanced Research Topics and Practical Applications, 222 p. (2014)

[G] A. Lebedev, A. Manafov, "*DDS: The Dynamic Deployment System*", EPJ Web Conf. **214**, 01011 (2019)

[H] E. Salituro, Learn Grafana 7.0: A beginner's guide to getting well versed in analytics, interactive dashboards, and monitoring, 410 p. (2020)

Chairman of the VBLHEP Scientific and Technical Council E. A. Strokovsky

Scientific Secretary of the LHEP Scientific and Technical Council S. P. Merts