

IMPLEMENTATION OF THE CONDITION DATABASE FOR THE EXPERIMENTS OF THE NICA COMPLEX

**K.V. Gertsenberger^{1,a}, A.I. Chebotov¹, P.A. Klimai², I.N. Alexandrov¹,
E.I. Alexandrov¹, I.A. Filozova¹, A.A. Moshkin¹**

¹ *Joint Institute for Nuclear Research, 6 Joliot-Curie, Dubna, Moscow region, 141980, Russia*

² *Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow region, 141701, Russia*

E-mail: ^agertsen@jinr.ru

The processing and analysis of experimental and simulated data are an integral part of all modern high-energy physics experiments. These tasks are of particular importance in the experiments of the NICA project at the Joint Institute for Nuclear Research due to the high interaction rate and large particle multiplicity of ion collision events, therefore, the task of automating the considered processes for the NICA complex has particular relevance. The paper describes a new information system based on the Condition Database, as well as related services to automate storing, managing, searching and using various parameters and operation modes of the on-going and future NICA experiments for experimental and simulated data processing. The implemented system provides the necessary information for event processing and physics analysis tasks, and organises the transparent, unified access to and management of the required parameter data in scientific research. The scheme and purposes of the Condition Database, its attributes, key aspects of the development are described. In addition, a common deployment system of the developed database and related services for the NICA experiments is noted.

Keywords: condition database, NICA experiments, event data processing, information system, web services

Konstantin Gertsenberger, Alexander Chebotov, Peter Klimai, Igor Alexandrov,
Evgeny Alexandrov, Irina Filozova, Andrey Moshkin

Copyright © 2021 for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Introduction

According to the program of the Joint Institute for Nuclear Research (JINR, Dubna) for the study of heavy ion collisions, the construction of the NICA (Nuclotron-based Ion Collider fAcility) accelerator complex [1] will be completed in the coming years to collide particles in the range of atomic masses $A = 1-197$ at energies in the center-of-mass system up to 11 GeV for gold ions and up to 27 GeV for protons. There are two points of particle interactions in the storage rings of the NICA collider, where each interaction point will have a detector oriented to its own physics research program, namely, MultiPurpose Detector (MPD) and Spin Physics Detector (SPD). The MPD detector is optimized for a comprehensive study of the properties of hot and dense baryonic matter produced in heavy ion collisions and search for a critical point of phase transitions. Spin Physics Detector is aimed at studying the spin structure of nucleons in high-intensity polarized beams of light nuclei. Furthermore, one of the main elements of the first stage of the NICA project is an experiment on a fixed target, i.e. BM@N (Baryonic Matter at Nuclotron), which is carried out on particle beams extracted from the Nuclotron into the experimental hall, and whose technical runs have been conducted since 2015.

Modern large scientific research, such as the NICA experiments, is characterized by long duration, high labor intensity, complexity, and processing a large amount of data obtained during the experiment. In this regard, the task of automating the processes of collecting, storing, processing and analysing experimental and simulated data of the NICA complex acquires particular relevance. The automation of modern particle collision experiments is impossible without the use of specialized information and computing systems, which allow collecting, storing and processing event data, managing the experiment during runs, simultaneously servicing a large number of data requests and performing other actions being necessary to obtain high-quality physics results. Databases of various purposes and information services related to their use and support are an integral part of such software for online and offline data processing in experiments.

At present, large high-energy physics experiments are not employed without the creation and use of automated information systems, therefore, within the framework of the RFBR grant 18-02-40125, a complex of information systems for online and offline data processing, including a database of operation conditions and states of experimental systems (hereinafter the Condition Database), is developed for the NICA experiments. The next section presents general information on the architecture of the developed system and key aspects of its implementation.

2. Development of the Condition Database for the NICA experiments

Information systems built on top of modern databases and offering various user services for the transparent access to and management of stored information on the experiment are an important part of not only online, but also offline data flow tasks of experiments, including the processing and analysis of obtained physics data. One of such widely used software systems, the condition database, is aimed at storing, managing and using various parameters and operation modes of detectors and experimental subsystems for online and offline data processing, including algorithms for the reconstruction and physics analyses of particle collision events. The corresponding information system also solves problems of the convenient access to and control of the required parameters of the facility operation to employ them at all stages of experimental and simulated event processing.

The Condition Database for the NICA experiments (also called Unified Database) has been developed as a central storage of information on the operation conditions of various systems, as well as the parameters of the experiment being conducted or simulated. These data are necessary for the reconstruction and analysis of the recorded collision events, detector simulation, and are a key element at the stage of processing the obtained data. The implemented database provides the management of stored data and unified access to them for online and offline data processing systems, ensuring correct multi-user processing, actuality, integrity and consistency of the information. Furthermore, the

automatic regular backup of the stored data ensures recoverability in the case of software or hardware malfunctions.

The development of the Condition Database has been carried out on the PostgreSQL relational database management system. In the designed architecture of the database [2], three component parts can be emphasized: storing information on the conducted experiment runs and corresponding experimental data files, on generated simulation files, and storing the parameters of arbitrary formats for the subsystems of the experiment. During the design, the structure of the database was simplified as much as possible and unified to have a general scheme that can be used in various physics experiments on particle collisions, in particular, the NICA experiments. The following groups of the parameters stored in the Condition Database can be highlighted: configuration data concerned with the operation modes of the experiment systems, for example, programmable parameters for electronics, calibration data describing the calibration and alignment of the subdetectors, parameter data presenting the states of the detector systems and algorithm data that specify the way algorithms process events, for instance, cuts for the selection of particle tracks.

An important property of the parameters stored in condition databases is their changing during runs of experiments [3]. A single value or a set of values for a parameter describes the state of the system for a limited period of time and is used only for analysing events received in this interval of operation. In this regard, the stored parameters are characterized by their validity interval. As a result, the validity of parameters is usually defined as a time period or a run interval, and the corresponding time value or run number is used as a key to retrieve the parameter. In the developed Condition Database, the validity of parameters is determined by the run interval, for which they are valid. Furthermore, parameter values are retained in the case of updating (called *historical preservation*), which allows one to repeat event data processing with outdated parameter values used in the past. In addition, a key feature of the database is that the developed architecture provides storing parameters of any predefined types in the form of a binary object serialized to one database field through a programming interface.

3. Application programming and user interfaces for the Condition Database

To use the parameters stored in the Condition Database for event data processing of the NICA experiments, two application programming interfaces (API) have been developed. The first one, REST API service, provides users and other systems with an HTTP format of requests to search, manage and use information stored in the database. The REST interface allows querying and responding to GET, PUT, POST and DELETE commands, which refers to the reading and searching, updating, creating and deleting of operations concerning the parameters required for experimental or simulated data processing. The implemented REST API service has been deployed on the NICA cluster and provides a special URL as a REST endpoint. The second API is an accustomed C++ interface for integrating the Condition Database with the software frameworks of the NICA experiments, which are all based on the CERN ROOT package [4], for using stored information in online and offline data processing. The C++ interface allows accessing and managing necessary condition data as attributes of the objects in the experiment software.

In addition, a user-friendly web service for the Condition Database has been developed to simplify viewing, searching and managing information on the experiments by collaboration members via a web browser. It provides the following features depending on the role of the authenticated user: convenient viewing, adding or changing information on the experiment in a tabular view, multi-column sorting and data filtering, easy searching for necessary parameters according to specified criteria, visualizing summary information on stored data and many other functions. Figure 1 shows the deployed web service of the Condition Database for the BM@N experiment. After authorization, users can add and modify records directly from their web browser. It is crucial that the implemented web service has a configurable view, so it is easily adjusted to use in any NICA experiment.

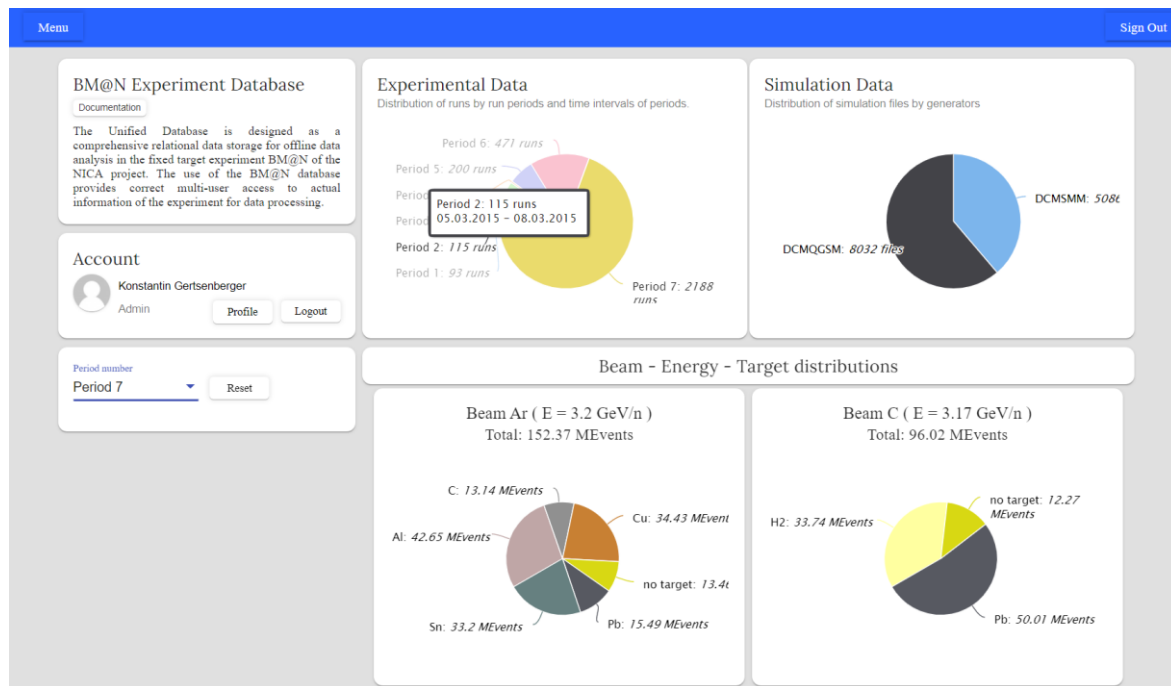


Figure 1. Home page of the web service for the Condition Database of the BM@N experiment

4. Information System based on the Condition Database

Information on experiment runs, on obtained experimental and simulated files, as well as various conditions and parameters of the system operation, is recorded to the Condition Database both in an online and offline mode. Then the stored information is used in developed algorithms for processing event data of the NICA experiments, including simulation, raw data digitization, the reconstruction of the obtained particle collision events and their further physics analyses. Experimental and simulated data processing is performed on distributed computing systems of the NICA project via experiment frameworks, which access and use the required parameters through the implemented APIs.

The Condition Database is being constantly improved via the development of various additional services. To have only one account for all the information systems of the experiments, FreeIPA authentication has been added. Once a collaboration member has registered an account, the member can use the login to sign in to all the information systems, including the web service of the Condition Database. The FreeIPA system provides centralized authentication, authorization and account information by storing data about collaboration members, their groups and other objects being necessary to manage security aspects. Three groups for administrators (full access), editors (e.g., detector team) and readers (regular users) have been created in the FreeIPA system to restrict user access to the Condition Database.

It is important to monitor the state of the information systems in order to minimize the response time in the case of failures being especially critical during experiment runs. The Monitoring Service has been developed for the information system, it tracks server hosts and the Condition Database, stores their state parameters and visualizes them through the Grafana system on a special web site. In the case of any malfunctions, the Monitoring Service sends notifications to given emails of responsible people.

The Condition Database is in demand for all the NICA experiments, so the information system was made generic and highly customizable. A Configuration and Deployment System has been implemented for convenient installation for any NICA experiment. The main concept of the common deployment system was to develop a unified procedure for configuration and installation steps on deploying all the components of the Condition Database and related information services. It uses

Docker containers and shell scripts, which allows installing the database, web service, REST API and other components for the experiments of the NICA project, taking into account some specifics of the experiments. The result of the configurable deployment of the information system for the NICA experiments is the created database and services, including the web service deployed on the host, thus collaboration members can immediately start to work with the Condition Database.

5. Conclusion

The Condition Database implemented on PostgreSQL 12 is an important component of the complex of information systems for the NICA experiments. It provides a centralized storage of the necessary conditions and parameters and ensures unified access to them for online and offline experimental systems to process and analyse data from the facilities of the NICA complex, as well as allows collaboration members to manage up-to-date information on the experiments. The Condition Database has been deployed on the NICA cluster, and at present, it is actively used in distributed data processing of the BM@N experiment, the first experiment of the NICA project. The REST API service and ROOT C++ interface have been implemented to use information stored in the Condition Database for processing simulated and experimental data in other software systems of the experiments. The web service, single authorization and monitoring systems have been developed to simplify viewing and managing information on the experiments by collaboration members over the Web. The Common Deployment System and additional services have been implemented to install the Condition Database for all the experiments of the NICA project, taking into account some specifics of the experiments. The information system makes a significant contribution to solving the problem of automating the acquisition, storage, processing and analysis of obtained data, and is also a necessary element for obtaining timely high-quality physics results.

6. Acknowledgement

The reported work was funded by the Russian Foundation for Basic Research (RFBR) according to the research project No.18-02-40125.

References

- [1] NICA White paper. Searching for a QCD mixed phase at the Nuclotron-based ion collider facility. Available at: http://mpd.jinr.ru/wp-content/uploads/2016/04/WhitePaper_10.01.pdf. (accessed 10.09.2021)
- [2] Gertsenberger K., Chebotov A., Alexandrov I., Filozova I., Alexandrov E. Design of the Condition Database for online and offline data processing in experimental setups of the NICA complex // *Izvestiya SFedU. Engineering Sciences*. 2020. No. 7 (217). pp. 172-180.
- [3] Laycock P., Dykstra D., Formica A., Govi G., Pfeiffer A., Roe S. and Sipos R. A Conditions Data Management System for HEP Experiments // *Journal of Physics: Conference Series*. 2018. Vol. 1085. 032040.
- [4] R. Brun and F. Rademakers. ROOT – An Object Oriented Data Analysis Framework // *Proceedings AIHENP'96 Workshop. – Nucl. Inst. & Meth. in Phys. Res. A*. 1997. Vol. 389. pp. 81-86.