Referee's report
on the large research infrastructure project
"Multifunctional Information and Computing Complex"

The considered large research infrastructure project "Multifunctional Information and Computing Complex (MICC) of JINR" is an integral part of the Seven-Year Plan for the development of JINR for 2024-2030 within the "Networking, Computing, Computational Physics" direction. The main objective of this direction is to ensure the continuous development of the information and computing infrastructure, which will allow JINR to stay at the forefront of scientific research in different fields that the Institute is conducting and will conduct in the coming years. The MICC is one of JINR's basic facilities. Classifying this Project as a large research infrastructure project is justified and timely, given the decisive importance of computing for obtaining physics results within the experiments of the NICA megaproject and the JINR neutrino program, as well as other theoretical and experimental studies, according to the Seven-Year Plan for the development of JINR.

The JINR MICC was created at the Meshcheryakov Laboratory of Information Technologies and successfully developed from year to year in accordance with the rapid development of information technologies, computing equipment and computing technologies, satisfying user needs. Thus, the participation of JINR scientists in the experiments at the Large Hadron Collider (LHC) at CERN entailed the creation of computing clusters at JINR on the basis of grid technologies, which are integrated into a distributed computing environment for processing and storing hundreds of petabytes of data obtained from the LHC experimental facilities. In this regard, the Tier1 center for the CMS experiment, which is one of the best in the world, and the Tier2 center, which is the best in Russia and serves all experiments at the LHC and other studies with JINR scientists' participation using grid technologies, should be particularly noted. At present, both of these centers serves the experiments at the NICA complex. At the next stage in the MICC development, a cloud infrastructure was created, it broadened the range of services provided to users by the MLIT computing center and ensured flexible configuration, a guaranteed level of service, reliable data storage and accessibility from an arbitrary computer with a network connection.

The growing interest of MICC users in parallel computing and the emergence of different computing architectures, such as multi-core processors, coprocessors and graphics processing units, resulted in the creation of the HybriLIT heterogeneous cluster for parallel computing, being part of the HybriLIT heterogeneous platform, the major component of which is the "Govorun" supercomputer.

In parallel, within the MICC, to ensure mass data storage tasks, a multi-layer data storage system from the operational layer to the long-term layer, implemented on robotic tape libraries, was developed.

In general, the Project's leaders and team coped with all tasks at the previous stages of the MICC development. This is evidenced by the creation of Tier1, the Member States' cloud infrastructure and the "Govorun" supercomputer, aimed not only at massively parallel computing, but also at tasks with intensive work with large data volumes. It is also noteworthy that the "Govorun" supercomputer is built on liquid cooling, which ensures its high-energy efficiency, and the hierarchical data processing and storage system, which was created on it following the initiative of the MLIT team, made it possible to effectively solve the tasks of the MPD experiment at NICA at the simulation stage and process data from the BM@N experiment.

To date, the JINR MICC is a unique computing complex that integrates grid, cloud and HPC technologies, as well as data storage systems. Thus, a technological platform for scientific research conducted by the JINR scientific community, combining engineering and computing structures built on different technological solutions, concepts and methods, was created at MLIT.

The present Project is aimed at ensuring the further development of the JINR network and computing infrastructures to carry out scientific studies of the Institute and its Member

States on the basis of advanced information technologies. The investment nature of the MICC project, which requires the regular necessary modernization of systems, retrofitting or major replacement of equipment, should be underlined. The main goal of the Project is to provide multifunctionality, high performance, extended data storage, high reliability and availability, information security, scalability, a state-of-the-art software environment for different user groups, high-speed telecommunications and a wide range of services in 24x7x365 mode to conduct scientific research within the JINR Topical Plan.

To meet these requirements, it is needed to solve the following key tasks from year to year:

1. Development and enhancement of the JINR telecommunication and network infrastructure.
2. Modernization of the JINR MICC engineering infrastructure, taking into account the requirements for power supply and climate control systems of new server equipment.
3. Modernization and development of the offline distributed computing platform for the NICA project, involving NICA collaboration computing centers.
4. Creation of a Tier0 grid cluster for the experiments of the NICA megaproject to process and store experimental and modeled data. Expansion of the performance and capacity of the storage systems of the Tier1 and integrated Tier2/CICC grid clusters as data centers for the experiments of the NICA megaproject, the JINR neutrino program and the LHC experiments.
5. Enlargement of the JINR cloud infrastructure to broaden the range of services provided to users on the basis of containerization technologies. Automation of the deployment of cloud technologies in the JINR Member States' organizations.
6. Enhancement of the HybriLIT heterogeneous platform, including the "Govorun" supercomputer, as a hyperconverged software-defined environment with a hierarchical data storage and processing system.

Undoubtedly, the strength of the Project is the purposeful integration of all computing center support systems from engineering and network to server solutions. Such an approach enables to ensure the unified monitoring of all infrastructures and the required technical work on the non-stop modernization and enlargement of the equipment.

A promising direction for further development is the creation of a hardware and software complex to support the activity of the geographically distributed research environment that provides MICC services for the remote use of its infrastructure in areas of cooperation with Russian and foreign research groups.

In addition, it is noteworthy that to ensure the fulfillment of the above tasks, the Project provides for the solution of organizational and technological issues and the conduct of scientific research as part of the activity on the creation of a multi-purpose hardware and software platform for Big Data analytics based on hybrid hardware accelerators (GPU, FPGA, quantum systems). One of the goals of creating this platform is to enhance the efficiency of the MICC operation on the basis of the methodology of project management in the field of information technology and software.

A certainly promising direction of the Project is the further expansion of the use of the DIRAC and PanDA middleware to organize distributed computing and provide computing for megascience projects and not only. As the experience of computing for the LHC at CERN has shown, successful experimental data processing has been ensured and is currently ensured thanks to computing resources in a distributed environment. These systems enable the integration of computing resources, including networks, clouds, supercomputers, storage resources. It is the heterogeneity of the combined resources that is attractive for collaborations, since there are no requirements to purchase the same type of equipment.

I would like to point out that the team of authors of the Project is well integrated into the international community. The introduction of grid technologies for Tier1, cloud solutions for integrating the Member States' infrastructures and a solution in the field of unifying three supercomputers in Russia are especially impressive. The choice of technologies and middleware is made for the major components and is consistent with the software development process.

To my mind, the Project's team consists of a group of computer professionals and computer scientists who achieved outstanding success in the past, such as the best grid infrastructure in the Member States, the cloud infrastructure of JINR and its Member States, as well as the hyperconverged computing environment, i.e., the "Govorun" supercomputer.

The team has detected a number of technical risks that the Project may face and should have strategies in place to mitigate and eliminate them.

It should be mentioned that the funding suggested for the Project meets the basic tasks of the Project; however, attention should be paid to the fair indication of the authors that the budgetary funding of the Project provided by the Seven-Year Plan cannot fully satisfy the requirements of JINR's flagship projects, in particular, Tier0 for NICA and computing for the neutrino program (Baikal-GVD, JUNO projects), in terms of computing power and storage systems. Therefore, special attention should be paid to the collective work on the issue of the joint financing of computing with the authors of the above projects and programs.

In general, I believe that the proposed Project for the development of the JINR Multifunctional Information and Computing Complex deserves full approval and continuous financial support from the JINR Directorate, collaborations and other JINR projects that need advanced computing.

06.04.2023

A.V.Bogdanov

Professor, D.Sc.,
Professor of Department of Fundamental
Informatics and Distributed Systems
Saint Petersburg State University