06-6-1118-2014/2030

MICC
Multifunctional Information and Computing Complex

Leaders: V.V. Korenkov
S.V. Shmatov

Deputies: A.G. Dolbilov
D.V. Podgainy
T.A. Strizh

Participating countries and international organizations:
Armenia, Azerbaijan, Belarus, Bulgaria, CERN, China, Egypt, France, Georgia, Kazakhstan, Mexico, Mongolia, Russia, South Africa, Uzbekistan.

The problem under study and the main purpose of the research:
The main objective of the MICC is to meet the needs of the JINR scientific community to the maximum extent possible in order to solve urgent tasks, from theoretical research and experimental data processing, storage and analysis to the solution of applied tasks in the field of life sciences. The tasks of the NICA project, the neutrino programme, the tasks of processing data from the experiments at the LHC and other large-scale experiments, as well as support for users of the JINR Laboratories and its Member States will be the priorities.

The project presupposes the inclusion of two activities, which, like the project, are aimed at meeting the requirements of a large number of research and administrative personnel:
– development of the digital platform "JINR Digital EcoSystem", which integrates existing and future services to support scientific, administrative and social activities, as well as to maintain the engineering and IT infrastructures of the Institute, which in turn will provide reliable and secure access to different types of data and enable a comprehensive analysis of information using modern technologies of Big Data and artificial intelligence;

– creation of a multi-purpose hardware and software platform for Big Data analytics based on hybrid hardware accelerators; machine learning algorithms; tools for analytics, reports and visualization; support of user interfaces and tasks.

Project of the infrastructure:

Name of the project	Project leaders	Project code
Laboratory Responsible from laboratories		**Status**
1. MICC Multifunctional Information and Computing Complex	**V.V. Korenkov** **S.V. Shmatov** *Deputies:* **A.G. Dolbilov** **D.V. Podgainy** **T.A. Strizh**	06-6-1118-1-2014/2030 Realization

MLIT K.N. Angelov, A.I. Anikina, O.A. Antonova, A.I. Balandin, N.A. Balashov, A.V. Baranov, D.V. Belyakov, T.Zh. Bezhanyan, S.V. Chashchin, A.I. Churin, O.Yu. Derenovskaia, V.P. Dergunov, A.T. Dzakhoev, A.V. Evlanov, V.Ya. Fariseev, M.Yu. Fetisov, S.V. Gavrilov, A.P. Gavrish, T.M. Goloskokova, A.O. Golunov, L.I. Gorodnicheva, E.A. Grafov, E.N. Grafova, N.I. Gromova, A.E. Gushchin, A.V. Ilyina, N.N. Karpenko, I.I. Kalagin, A.S. Kamensky, I.A. Kashunin, M.Kh. Kirakosyan, A.K. Kiryanov, A.A. Kokorev, G.A. Korobova, S.A. Kretova, N.A. Kutovsky, I.V. Kudasova, O.N. Kudryashova, E.Yu. Kulpin, A.E. Klochiev, A.V. Komkov, V.I. Kulakov, A.A. Lavrentiev, A.M. Levitin, Yu.M. Legashchev, M.A. Lyubimova, M.A. Maksimov, V.N. Markov, S.V. Marchenko, M. A. Matveev, A.N. Makhalkin, Ye. Mazhitova, A.A. Medyantsev, V.V. Mitsyn, N.N. Mishchenko, A.N. Mityukhin, A.N. Moibenko, I.K. Nekrasova, V.N. Nekrasov, D.A. Oleinik, V.V. Ovechkin, S.S. Parzhitsky, I.S. Pelevanyuk, D.I. Pryakhina, A.Sh. Petrosyan, D.S. Polezhaev, L.A. Popov, T.V. Rozhkova, Ya.I. Rozenberg, D.V. Rogozin, R.N. Semenov, A.S. Smolnikova, I.A. Sokolov, E.V. Solovieva, I.G. Sorokin, I.N. Stamat, V.P. Sheiko, D.A. Shpotya, B.B. Stepanov, A.M. Shvalev, O.I. Streltsova, Sh.G. Torosyan, V.V. Trofimov, N.V. Trubchaninov, E.O. Tsamtsurov, V.Yu. Usachev, S.I. Vedrov, A.S. Vorontsov, N.N. Voytishin, A.Yu. Zakomoldin, S.E. Zhabkova, M.I. Zuev

VBLHEP	K.V. Gertsenberger, A.O. Golunov, Yu.I. Minaev, A.N. Moshkin, O.V. Rogachevsky, I.V. Slepnev, I.P. Slepov
BLTP	A.A. Sazonov
FLNP	G.A. Sukhomlinov
FLNR	A.S. Baginyan, A.G. Polyakov, V.V. Sorokoumov
DLNP	Yu.P. Ivanov, V.A. Kapitonov, A.S. Zhemchugov
LRB	V.N. Chausov
UC	I.N. Semenyushkin
Assosiated personnel MICC	A.V. Anisenkov

Abstract and scientific rationale:
To attain the major goals of JINR's flagship projects, it will be required to process a huge amount of experimental data. According to a very rough estimate, these are tens of thousands of processor cores and hundreds of petabytes of experimental data. The experiments of the NICA project and the JINR neutrino programme (Baikal-GVD, JUNO, etc.) entail Tier0, Tier1 and Tier2 grid infrastructures. To achieve these goals, it is essential to develop distributed multi-layer heterogeneous computing environments, including on top of the resources of the participants of other projects and collaborations.

The concept of the development of information technology, scientific computing and Data Science in the JINR Seven-Year Plan provides for the creation of a scientific IT infrastructure that combines a multitude of various technological solutions, trends and methods. The IT infrastructure implies the coordinated development of interconnected IT technologies and computational methods aimed at maximizing the number of JINR strategic tasks to be solved that require intensive data computing. The large research infrastructure project "Multifunctional Information and Computing Complex" holds a special place in this concept.

The MICC LRIP main objective for 2024-2030 is to perform a set of actions aimed at the modernization and development of the major hardware and software components of the computing complex, the creation of a state-of-the-art software platform enabling the solution of a wide range of research and applied tasks in accordance with the JINR Seven-Year Plan. The rapid development of information technology and new user requirements stimulate the development of all MICC components and platforms. The MICC computing infrastructure encompasses four advanced software and hardware components, namely, the Tier1 and Tier2 grid sites, the hyperconverged "Govorun" supercomputer, the cloud infrastructure and the distributed multi-layer data storage system. This set of components ensures the uniqueness of the MICC on the world landscape and allows the scientific community of JINR and its Member States to use all progressive computing technologies within one computing complex that provides multifunctionality, scalability, high performance, reliability and availability in 24x7x365 mode with the multi-layer data storage system for different user groups.

Within the MICC LRIP, it is provided to support the operation of all MICC hardware and software components, i.e., the Tier1 and Tier2 grid sites, the cloud infrastructure, the hyperconverged "Govorun" supercomputer, the multi-layer data storage system, the network infrastructure, the power supply and climate control systems, as well as to modernize/reconstruct the above components in accordance with new trends in the development of IT technologies and user requirements. In addition, it is required to ensure high-speed telecommunications, a modern local area network infrastructure and a reliable engineering infrastructure that provides guaranteed power supply and air conditioning for the server equipment.

Expected results upon completion of the project:
Modernization of the JINR MICC engineering infrastructure (reconstruction in accordance with modern requirements of the machine hall of the 4^{th} floor of MLIT).

Modernization and development of the distributed computing platform for the NICA project with the involvement of the computing centres of the NICA collaboration.

Creation of a Tier0 grid cluster for the experiments of the NICA megaproject to store experimental and simulated data.

Expansion of the performance and storage capacity of the Tier1 and Tier2 grid clusters as data centres for the experiments of the NICA megaproject, the JINR neutrino programme and the experiments at the LHC.

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.

Expansion of the HybriLIT heterogeneous platform, including the "Govorun" supercomputer, as a hyperconverged software-defined environment with a hierarchical data storage and processing system.

Design and elaboration of a distributed software-defined high-performance computing platform that combines supercomputer (heterogeneous), grid and cloud technologies for the effective use of novel computing architectures.

Development of a computer infrastructure protection system based on fundamentally new paradigms, including quantum cryptography, neurocognitive principles of data organization and data object interaction, global integration of information systems, universal access to applications, new Internet protocols, virtualization, social networks, mobile device data and geolocation.

Expected results for the project in the current year:
Operation and provision of the reliable, safe and integral functioning of backbone external telecommunication channels (3x100 Gbps), the backbone network (2x100 Gbps), the transport network of the NICA megaproject (4x100 Gbps), the MLIT mesh network (100 Gbps), and the Wi-Fi network at the Institute's sites in 24x7x365 mode. Provision of standard network services (email, file sharing, security), maintenance, support and development of user databases, network elements, etc. Introduction of dual authorization and certification authorities in the JINR network. Elaboration of alternative routes for external telecommunication channels with Chinese centers and other organizations participating in the JUNO and NICA projects.

Creation and commissioning of a server hall on the 4th floor of the MLIT building.

Ensuring the uninterrupted operation of the MICC engineering complex in 24x7x365 mode, including guaranteed power supply systems (diesel generators, uninterruptible power supplies), climate control systems (chillers, dry coolers, inter-row air conditioners, etc.), and fire safety systems for the machine hall on the 2nd floor of the MLIT building. Comprehensive maintenance and optimization of all engineering equipment to ensure the ultimate efficiency of the computing infrastructure. Modernization of Tier2 Module 1 (installation of inter-row air conditioners and additional telecommunication cabinets) within the development of the MICC engineering infrastructure. Modernization of the power supply system of Tier1 Module 0. Organization of new Module 3 based on existing telecommunication cabinets.

Expansion of the performance and storage system of the MICC basic components, namely, the Tier1 center up to 25,000 CPU cores and 25,000 TB, Tier2/CICC up to 13,000 CPU cores, the EOS-based data lake up to 60 PB, including dedicated EOS systems for MPD and SPD. Maintenance of the unified storage and access system for common software (CVMFS). Commissioning of additional drives for the connection of CTA and Enstore with the TS4500 tape library. Support and maintenance of the operation of WLCG virtual organizations, the NICA experiments, etc., local user groups on the MICC CICC/Tier1 and Tier2 resources. Elaboration of a regional center for the JUNO experiment on top of the MICC resources.

Development of the prototype of a fully functional Tier0/Tier1 center for the experiments at the NICA accelerator complex.

Enhancement of the distributed information and computing environment (DICE) on the basis of the cloud resources of the JINR Member States' organizations and the neutrino platform at the expense of resources acquired by the Baikal-GVD, JUNO, and NOvA experiments.

Modernization of performance metrics visualization tools for the cloud infrastructure and the failure notification system. Implementation of Ceph Dashboard, a new web interface for the centralized management of Ceph-based cloud storage.

Implementation of automation tools for launching tasks in a distributed heterogeneous environment created on top of the DIRAC platform. Integration of the FTS3 system for managing data transfer within the existing distributed environment. Conducting BM@N mass data production sessions, technical support for launching MPD experiment tasks.

Enlargement of the CPU component of the "Govorun" supercomputer to meet current user needs, including within event modeling for the experiments at the NICA complex. Automation of the processes of OS network loading and cluster software configuration using XCAT, Puppet, etc. software packages on the compute nodes of the "Govorun" supercomputer.

Implementation of a system for collecting, visualizing and analyzing statistics on the utilization of computing resources and software by users of the HybriLIT platform and the "Govorun" supercomputer.

Testing and implementation of parallel and distributed data storage and processing systems, such as Apache Ignite, Deepseek 3FS (Fire-Flyer File System), etc., to enhance the efficiency of working with model and experimental data on the HybriLIT platform and the "Govorun" supercomputer.

Enhancement of the ML/DL/HPC ecosystem, including the quantum computing polygon, and implementation of work on integrating the computing resources of the “Govorun” supercomputer on top of the DASK software product to solve massively parallel tasks related to machine and deep learning algorithms.

Maintenance of the data storage and processing system for the SPD experiment using the MICC resources (cloud infrastructure for hosting middleware services, CICC computing infrastructure for performing tasks). Support for carrying out mass modeling within the preparation stage for SPD experiment data taking, processing, and analysis.

Development of a hardware and software platform for monitoring the MICC engineering and computing infrastructure. Enhancement of the LITmon monitoring system: elaboration of a monitoring system for the server infrastructure to store SPD experiment data; modernization of monitoring the engineering infrastructure of the MICC cooling system; modernization of the serial console monitoring system; creation of a platform for the automatic deployment of the monitoring system.

Activities of the infrastructure:

Name of the activity	Leaders	Implementation period
Laboratory Responsible from laboratories		**Status**
1. The digital ecosystem (Digital JINR)	**V.V. Korenkov** **S.D. Belov**	2024-2026 Realization

MLIT A.A. Artamonov, N.A. Balashov, N.E. Belyakova, O.V. Belyakova, A.S. Bondyakov, N.A. Davydova, I.A. Filozova, L.A. Kalmykova, E.N. Kapitonova, A.O. Kondratiev, E.S. Kuznetsova, E.K. Kuzmina, S.V. Kunyaev, L.D. Kuchugurnaya, D.V. Neapolitanskiy, I.K. Nekrasova, M.M. Pashkova, L.V. Popkova, Ya.I. Popova, A.V. Prikhodko, T.F. Sapozhnikova, V.S. Semashko, S.V. Semashko, E.V. Sheiko, I.A. Sokolov, G.V. Shestakova, L.A. Shiryaeva, T.S. Syresina, Yu.V. Trofimov, D.Yu. Usov, T.N. Zaikina

VBLHEP V.V. Morozov, I.V. Slepnev, A.V. Trubnikov

DLNP R.S. Kovalenko

DSDD A.V. Sheiko, M.P. Vasiliev

Abstract and scientific rationale:
The activity is related to the creation of an Institute-wide digital platform "JINR Digital EcoSystem". The main objective is the organization of a digital space with a single access and data exchange between electronic systems, as well as the transition of actions that previously required a personal or written request to a digital form. The platform is designed to ensure the integration of existing and future services to support scientific, administrative and social activities, as well as to maintain the engineering and IT infrastructures of the Institute.

Within the activity, two main directions of work are planned: the creation of the basic infrastructure of the digital platform (including the software-hardware and methodological support of its functioning) and different digital services. In addition to service support, digital services for scientific collaborations, whose activity is related to JINR’s basic facilities, will be developed and maintained for use by the Institute’s staff members.

Expected results upon completion of the activity:
Creation of a hardware-software and methodological basis for the functioning of the Institute-wide digital platform.

Development and implementation of digital services for distributed access to resources (information, computing, administrative, organizational ones) in a unified environment.

Transition of the processes of getting permits, approvals and applications of different types into a digital form.

Creation of a catalogue and a distributed storage of data related to the scientific and technical aspects of the Institute’s activity, as well as of tools for their analysis, presentation and the construction of predictive models.

Expected results in the activity in the current year:
Use of the unified environment for storing and managing data of basic and applied DES services, integration with the Big Data infrastructure to analyze the specified data.

Development of the institutional repository of publications of JINR staff members: transition to a new version of the repository software, connection of additional sources of information on articles, development of algorithms to link staff member profiles with articles, refinement of user interfaces, uploading articles of JINR staff members published before 2021 to the system.

Commissioning of a digital service for the centralized remote testing of JINR staff members' knowledge on radiation safety.

Implementation of electronic user registration on CICC resources.

Organization of an electronic system for accounting work and receiving requests from staff members for MLIT technical services.

Implementation of additional capabilities in the geoinformation system to support the activities of JINR technology services and departments upon their request. Integration of the geoinformation system with other DES services.

Ongoing support and development of the Dubna EDMS. Provision of procurement data for the document management system created by the Development of Digital Services Department.

Implementation of uploading electronic originals and copies of documents in the service of expense reports on business trips, optimization of request review processes.

Transition of the functionality of the baza.jinr.ru document database to the Dubna EDMS.

Development of digital collaboration services (scientific documentation database, project management, secure file sharing and collaboration service for documents of various types, etc.).

Development of the user support system and the electronic request system for various DES services, preparation of knowledge bases and documentation for users.

2. The multi-purpose hardware and software platform for Big Data analytics — **P.V. Zrelov** — 2024-2026

Realization

MLIT	A.A. Artamonov, D.A. Baranov, S.D. Belov, I.A. Filozova, Yu.E. Gavrilenko, O.V. Ivantsova, A.V. Ilyina, M.S. Katulin, I.A. Kashunin, E.A. Kuznetsov, M.A. Matveev, D.V. Neapolitanskiy, I.S. Pelevanyuk, N.V. Ryabov, R.N. Semenov, T.M. Solovieva, E.V. Sheiko, Yu.B. Starchenko, V.A. Tarabrin, T.N. Zaikina, D.P. Zrelova, A.V. Yakovlev
Assosiated personnel	A.V. Bogdanov, A.B. Degtyarev, Zu.U. Kiyamov, V.V. Korkhov, V.V. Mareev, N.L. Shchegoleva

Abstract and scientific rationale:
The activity provides for the creation of a multi-purpose hardware and software platform for Big Data analytics, which implements a full cycle of continuous processing, from data acquisition to the visualization of processing and analysis results, forecasts, recommendations and instructions, within the JINR MICC. One of the tasks planned to be solved using the platform is the elaboration of an analytical system for managing the MICC resources and data flows to enhance the efficiency of using computing and storage resources and optimize experimental data processing, as well as the development of the intelligent monitoring of distributed computing systems and data centres. Another essential task is the creation and development of analytics tools for the services of the JINR Digital EcoSystem.

Expected results upon completion of the activity:
Creation of a universal core of a Big Data mining platform.

Development and implementation of a number of standard software solutions for different classes of tasks within the platform.

Elaboration and development of analytics tools for the JINR Digital EcoSystem.

Development of methods and creation of complex solutions for analysing the security of data and computer systems.

Development of artificial intelligence methods within the analytical platform and creation of a software environment for work with technical and scientific information.

Elaboration of common solutions based on Big Data analytics for expert and recommendation systems, including for the optimization of the processes of functioning of the MICC components.

Expected results of the activity in the current year:
Deployment, configuration and service of access to autonomous specialized generative AI models and tools for interacting with them to solve the tasks of developing information systems and analyzing data.

Development of a prototype of a system for collecting, analyzing and storing traffic data from the JINR local area network with implementation in the form of a Big Data processing pipeline to enhance the network's security.

Elaboration of a methodology for the intelligent processing of scientific and technical information (scientific publications, patents, program and database registration materials, texts of projects, educational programs, job offers, CVs, etc.). Examples of implementation on the Institute's topics.

Development of a prototype of a software data platform on top of a composition of Big Data analytics tools, namely, Apache Spark, Dask, and the ROOT package, using hardware computing accelerators (graphics processors).

Analytical tools, hardware and software infrastructure and methods for integrating and analyzing data from Digital EcoSystem services.

Elaboration of quantum-like analogues of classical data analysis methods, in particular, machine learning methods, including the application of quantum convolutional and quantum generative adversarial neural networks.

Analysis of the possibility of using Industrial Internet of Things (IIoT) technologies for the JINR MICC engineering infrastructure: a) analysis of requirements for the IIoT system (parameters of the environment, types of equipment, data exchange protocols, and security requirements), b) development of a prototype of the system architecture, including the logical and physical structure of interaction of diverse components.

Study of the possibility of creating multi-level decentralized storage and its optimization:
– comprehensive analysis of the existing architecture of storage and processing systems within the MICC infrastructure and determination of optimal strategies for the system's development:

– development of analytical methods and software tools for modeling the behavior of a distributed system based on the MICC infrastructure.

Collaboration 1118

Country or International Organization	City, region	Institute or laboratory	
Armenia	Yerevan, ER	IIAP NAS RA	
		RAU	
Azerbaijan	Baku, BA	ADA	
		IP	
Belarus	Minsk, MI	INP BSU	
		IP NASB	
		JIPNR-Sosny	
		UIIP NASB	
Bulgaria	Sofia	INRNE BAS	
		SU	
CERN	Geneva, CH	CERN	
China	Beijing, BJ	IHEP CAS	
Egypt	Cairo, C	ASRT	
France	Marseille, PAC	CPPM	
Georgia	Tbilisi, TB	GRENA	
		GTU	
		UG	
Kazakhstan	Almaty, ALA	INP	
	Astana, AST	AB INP	
		AITU	
Mexico	Mexico City, CDMX	UNAM	

Country or International Organization	City, region	Institute or laboratory
Mongolia	Ulaanbaatar	IMDT MAS
Russia	Chelyabinsk, CHE	SUSU
	Chernogolovka, MOS	SCC IPCP RAS
	Dubna, MOS	SCC "Dubna"
		SEZ "Dubna"
		Uni Dubna
	Gatchina, LEN	NRC KI PNPI
	Moscow, MOW	BMSTU
		FRC CSC RAS
		HSE
		IITP RAS
		ISP RAS
		KIAM RAS
		MEPhI
		MISIS
		MPEI
		MSU
		NRC KI
		PFUR
		PRUE
		RCC MSU
		RSCC
		SINP MSU
	Novosibirsk, NVS	BINP SB RAS
		ICMMG SB RAS
		SKIF
	Protvino, MOS	IHEP
	Pushchino, MOS	IMPB RAS
	Saint Petersburg, SPE	FIP
		ITMO
		SPbSPU
		SPbSU
	Samara, SAM	SNRU
	Taganrog, ROS	SRI MCS SFU
	Troitsk, MOW	INR RAS
	Vladikavkaz, SE	NOSU
	Vladivostok, PRI	IACP FEB RAS
	Voronezh, VOR	VSU
South Africa	Cape Town, WC	UCT
Uzbekistan	Tashkent, TK	AS RUz
		INP AS RUz