06-6-1118-2014/2030

Multifunctional Information and Computing Complex (MICC)

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, Italy, Kazakhstan, Mexico, Moldova, Mongolia, Russia, Slovakia, South Africa, Taiwan, USA, 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:

1. 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.
2. 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:

	Name of the project	Project Leaders	Project code
1.	Multifunctional Information and Computing Complex (MICC)	V.V. Korenkov S.V. Shmatov *Deputies:* A.G. Dolbilov D.V. Podgainy T.A. Strizh	06-6-1118-1-2014/2030

Project:

	Name of the project	Project Leaders	Status
	Laboratory (Subdivision)	Responsible from laboratories	
1.	**Multifunctional Information and Computing Complex (MICC)**	**V.V. Korenkov** **S.V. Shmatov** *Deputies:* **A.G. Dolbilov** **D.V. Podgainy** **T.A. Strizh**	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, A.S. Bondyakov, Yu. A. Butenko, S.V. Chashchin, A.I. Churin, O.Yu. Derenovskaia, V.P. Dergunov, 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.A. Kokorev, A.O. Kondratiev, 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, I.I. Lensky, 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, I.K. Nekrasova, V.N. Nekrasov, A.V. Nechaevsky, 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, E. V. Solovieva, I.G. Sorokin, I.N. Stamat, V.P. Sheiko, D.A. Shpotya, B.B. Stepanov, A.M. Shvalev, M.L. Shishmakov, O.I. Streltsova, I.A. Sokolov, E.V. Toneeva, 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, Yu.I. Minaev, A.N. Moshkin, O.V. Rogachevsky, I.P. Slepov
BLTP	A.A. Sazonov
FLNP	G.A. Sukhomlinov
FLNR	A.S. Baginyan, A.G. Polyakov, V.V. Sorokoumov
DLNP	A.S. Zhemchugov, Yu.P. Ivanov, V.A. Kapitonov
LRB	V.N. Chausov
UC	I.N. Semenyushkin

Brief annotation 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 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, 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:
1. Modernization of the JINR MICC engineering infrastructure (reconstruction in accordance with modern requirements of the machine hall of the 4th floor of MLIT).
2. Modernization and development of the distributed computing platform for the NICA project with the involvement of the computing centres of the NICA collaboration.

3. 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.
4. 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.
5. Expansion of the HybriLIT heterogeneous platform, including the "Govorun" supercomputer, as a hyperconverged software-defined environment with a hierarchical data storage and processing system.
6. 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.
7. 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 of the project in the current year:

1. Maintenance of the stable, safe and integral functioning of the JINR information and telecommunication network (backbone network (2x100 Gbps); the transport network of the NICA megaproject (4x100 Gbps); the MLIT mesh network (100 Gbps); backbone external telecommunication channels (3x100 Gbps); the Wi-Fi network at the Institute's sites in 24x7x365 mode. Support of standard network services: email (SMTP, IMAP, POP3, WebMail), file sharing (ftp, scp, sftp, http, https), security (ssh, https, TACACS authentication, dns, SSO), user database support, IPDB network element database support, etc.
2. Maintenance and operation of the full-scale and optimal functioning of the guaranteed power supply (diesel generators, uninterruptible power supplies) and climate control systems (chillers, dry coolers, inter-row air conditioners, etc.) of the MICC computing infrastructure in 24x7x365 mode. Commissioning of a new fire safety system of the MICC infrastructure. Project elaboration and start of modernization of the server room in the hall of the 4th floor of the MLIT building.
3. Expansion of the performance and storage system of the MICC basic components, namely, the Tier1 centre up to 22,000 CPU cores and 14,500 TB, Tier2/CICC up to 11,000 CPU cores, the EOS system up to 27 PB. Enlargement of the total volume of the robotic tape storage up to 70 PB. Support and maintenance of user work with the EOS system. Support of the system of access to the home directories of JINR users (AFS). Development and support of the unified storage and access system for common software (CVMFS). Support of the software system for working with tape robots (CTA). Creation and update of a polygon for debugging and testing new software for the MICC uppermost components. Support and maintenance of the operation of WLCG virtual organizations, the NICA, COMPASS, NOvA, ILC and other experiments, local user groups on the MICC Tier1 and Tier2 resources.
4. Extension of the number of users and participants of the distributed information and computing environment (DICE) on the basis of the cloud resources of the JINR Member States' organizations. Development of tools for the functional monitoring of cloud resources connected to the DICE. Enlargement of the resources of the MICC cloud, including at the expense of the resources acquired by the Baikal-GVD, JUNO, NOvA/DUNE experiments, and their maintenance. Migration of the server operating systems (OS) of all components of the JINR cloud, as well as HTCondor and JupyterHub virtual clusters deployed in the cloud, to a new OS due to the end of the lifespan of CentOS Linux 7 in June 2024. Development and implementation of the JupyterHub cluster monitoring system. Implementation of the HTCondor cluster monitoring system based on the htcondor-exporter metrics collector of own design.
5. Transition to a new version of DIRAC: DIRAC 8. Development and implementation of a system for analysing the performance of resources integrated into the DIRAC-based distributed heterogeneous computing environment. Development of tools and approaches to data transfer monitoring.
6. Development of a distributed system for storing and processing hot data under the management of parallel low-latency file systems (Luster/BeeGFS), as well as the DAOS distributed data storage, and introduction of this system to the hierarchical data processing and storage structure of the "Govorun" supercomputer and the NICA offline cluster to model and reconstruct events for the NICA experiments.
7. Putting into trial operation the components of the prototype of a data processing system in the distributed computing environment for the SPD experiment (SPD offline computing). Trial operation of the data management system and testing of its interaction with the data processing management system. Development of specialized services typical for Tier0 centers.
8. Development and support of the current MICC monitoring and accounting system, inclusion of monitoring the parameters of new computing and engineering elements in the list of monitored services and hardware. Development of new scripts to automate data acquisition processes. Creation of a script of data transfer failure notification for the dCache file system. Within the creation of a control room for engineering systems (power supply and climate control systems), special information displays with schematic representations of these systems will be elaborated. Development of analytical systems capable of notifying of the most critical issues of the MICC in real time.

Activities:

	Name of the activity	Leaders	Implementation period
	Laboratory (Subdivision)	Responsible from laboratories	
1.	**JINR Digital EcoSystem**	**V.V. Korenkov** **S.D. Belov**	2024-2026
	MLIT	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, I.K. Nekrasova, M.M. Pashkova, L.V. Popkova, A.V. Prikhodko, T.F. Sapozhnikova, V.S. Semashko, S.V. Semashko, I.A. Sokolov, E.V. Sheiko, G.V. Shestakova, T.S. Syresina, D.Yu. Usov, P.V. Ustenko, T.N. Zaikina	
	VBLHEP	V.V. Morozov, I.V. Slepnev, A.V. Trubnikov	
	DSDD	A.V. Sheiko	

Brief annotation 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:

1. Creation of a hardware-software and methodological basis for the functioning of the Institute-wide digital platform.
2. Development and implementation of digital services for distributed access to resources (information, computing, administrative, organizational ones) in a unified environment.
3. Transition of the processes of getting permits, approvals and applications of different types into a digital form.
4. 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:

1. Creation and mutual integration of existing basic digital infrastructure services: authentication, management and control of roles and access rights, data exchange bus, notification system, automated data catalogue, distributed storage.
2. Commissioning of the ecosystem's user interface, including mechanisms and methods for integrating services into it, the organization of feedback from users on the basis of electronic requests, a system of notifications and alerts. Development of administrative mechanisms for maintaining the functioning of the DES, including the distribution of roles and responsibilities, as well as coordination chains for services and the DES as a whole. Creation of a set of specifications and programme templates for developers of digital services.
3. Putting into trial operation a prototype of a service of the institutional repository of JINR staff members' publications, which enables to create and update data on the profile of the author and structural subdivisions, to receive bibliographic metadata from external sources, to upload metadata to the repository in automatic mode with reference to authors' profiles.
4. Putting into trial operation a prototype of a service for storing scientific documentation, which provides the possibility of centralized storage and exchange of different types of documentation between users of the system. Transition to the service and integration of data from existing outdated disparate databases of scientific documentation.
5. Current maintenance and development of the "Dubna" EDMS, including the creation of a subsystem for the archival storage of documents, the elaboration of new electronic documents and reports and modification of existing ones in accordance with the Institute's orders and user requests, the expansion of the scope of the "Dubna" EDMS for getting permits, approvals and applications of different types.

6. The following features will be implemented in the geoinformation system to support JINR's technological services: accounting of engineering networks of different types with all the necessary attributes, land plots, real estate objects, landscaping and the infrastructure, linking of electronic documents (schemes, photos, etc.) to objects and information about ongoing works (reconstruction, repairs, etc.) with reference to time. Implementation of the possibility of editing the geometry of objects and their attributes, creating new objects. Creation of a role model for differentiating access to information on objects.

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

MLIT — S.D. Belov, I.A. Filozova, Yu.E. Gavrilenko, A.V. Ilyina, I.A. Kashunin, M.A. Matveev, I.S. Pelevanyuk, R.N. Semenov, T.M. Solovieva, E.V. Sheiko, V.A. Tarabrin, T.N. Zaikina, D.P. Zrelova

Brief annotation 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:

1. Creation of a universal core of a Big Data mining platform.
2. Development and implementation of a number of standard software solutions for different classes of tasks within the platform.
3. Elaboration and development of analytics tools for the JINR Digital EcoSystem.
4. Development of methods and creation of complex solutions for analysing the security of data and computer systems.
5. Development of artificial intelligence methods within the analytical platform and creation of a software environment for work with technical and scientific information.
6. 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:

1. Creation of a prototype of an infrastructure and a software-analytical platform for Big Data.
2. Methodology for the analysis of streaming data with a high arrival rate.
3. Elaboration of intelligent data marts based on the Big Data approach.

Collaboration

Country or International Organization	City	Institute or laboratory
Armenia	Yerevan	IIAP NAS RA
Azerbaijan	Baku	ADA
		IP ANAS
Belarus	Minsk	INP BSU
		JIPNR-Sosny NASB
		UIIP NASB
Bulgaria	Sofia	INRNE BAS
		SU
CERN	Geneva	CERN
China	Beijing	IHEP CAS
Egypt	Cairo	ASRT
	Giza	CU
France	Marseille	CPPM
Georgia	Tbilisi	GRENA
		GTU
		TSU
Italy	Bologna	INFN

Kazakhstan	Almaty	INP
	Astana	BA INP
Mexico	Mexico City	UNAM
Moldova	Chisinau	IMCS
		MSU
		RENAM
Mongolia	Ulaanbaatar	IMDT MAS
Russia	Chernogolovka	SCC IPCP RAS
	Dubna	Dubna State Univ.
		SCC “Dubna”
		SEZ “Dubna”
	Gatchina	NRC KI PNPI
	Moscow	FRC IM RAS
		IITP RAS
		ISP RAS
		ITEP
		JSCC RAS
		KIAM RAS
		MPEI
		MSK-IX
		MSU
		NRC KI
		NRU HSE
		PRUE
		RCC MSU
		RSCC
		SINP MSU
	Moscow, Troitsk	INR RAS
	Novosibirsk	BINP SB RAS
		ICMMG SB RAS
		SKIF
	Pereslavl-Zalesskiy	PSI RAS
	Protvino	IHEP
	Puschino	IMPB RAS
	Samara	SU
	Saint Petersburg	FIP
		ITMO Univ.
		SPbSPU
		SPbSU
	Vladikavkaz	NOSU
	Vladivostok	IACP FEB RAS
Slovakia	Kosice	IEP SAS
South Africa	Cape Town	UCT
Taiwan	Taipei	ASGCCA
USA	Arlington, TX	UTA
	Batavia, IL	Fermilab
	Upton, NY	BNL
Uzbekistan	Tashkent	AS RUz
		INP AS RUz