Form № 24

Multifunctional Information and Computing Complex LIT JINR

MICC

05-6-1118-2014/2019

Authors and main researchers:

LIT

Adam Gh., Adamov G., Aleksandrov E.I., Aleksandrov I.N., Angelov K.N., Astakhov N.S., Baginyan A.S., Balandin A.I., Balashov N.A., Baranov A.V., Belov S.D., Belyakov D.V., Bondyakov A.S., Butenko Yu.A., Chashchin S.V., Churin A.I., Dolbilov A.G., Gavrilov S.V., Gavrish A.P., Galaktionov V.V., Goloskokova T.M., Golunov A.O., Grafov E.A., Gromova N.I., Gushchin A.E., Kadochnikov I.S., Kamenskiy A.S., Kapitonov V.A., Kashunin I.A., Kondratiev A.O., Korenkov V.V., Korobova G.A., Kulpin E.Yu., Kutovskiy N.A., Lavrentiev A.A., Marchenko S.V., Matveev M.A., Mitsyn V.V., Mitsyn S.V., Nechaevskiy A.V., Oleinik D.A., Ososkov G.A., Pelevanyuk I.S., Petrosyan A.Sh., Plyashkevich M.S., Podgainy D.V., Popov L.A., Pryakhina D.I., Rozenberg Ya.I., Sapozhnikova T.F., Semenov R.N., Streltsova O.I., Strizh T.A., Trofimov V.V., Uzhinskiy A.V., Shishmakov M.L., Voytishin N.N., Vorontsov A.S, Zakomoldin A.Yu., Zhiltsov V.E., Zrelov P.V., Zuev M.I.

VBLHEP: Potrebenikov Yu.K., Shchinov B.G., Minaev Yu.I., Rogachevsky O.V., Shmatov S.V.

FLNP: Sukhomlinov G.A.

LRB: Chausov V.N.

FLNR: Sorokoumov V.V., Polyakov A.G.

DLNP: Ivanov Yu.P.

BLTP: Sazonov A.A.

UC: Semenyushkin I.N.

Participating countries and organizations:

Azerbaijan (IP ANAS),

Armenia (IIAP NAS RA, YSU),

Belarus (NC PHEP BSU, JIPNR-Sosny),
Bulgaria (INRNE BAS, SU),
CERN,
China (IHEP, Beijing),
Cuba (USI, Havana),
Czech Republic (IP ASCR),
Egypt (CU),
France (CPPM),
Georgia (GRENA, TSU, GTU),
Germany (DESY, GSI, KIT),
Italy (INFN, Bologna),
Kazakhstan (ENU, Astana),
Moldova (ASM, IMCS ASM, RENAM),
Mongolia (NUM),
Poland (CYFRONET),
Romania (IFIN-HH, IFA, INCDTIM),
Russia (RSCC, IITP RAS, ISP RAS, ITEP, MSU, MPEI, RCC MSU, SINP MSU, NRC KI, RIPN, INR RAS, PNPI, Dubna Univ., SEZ "Dubna", SCC "Dubna", UNN, INP CO RAS, PSI RAS, IHEP, IMPB RAS, SSAU, SPbSU, PSUE "Informatics and management", SPbSPU, SCC IPCP RAS, LITP RAS),
Slovakia (IEP SAS),
Taiwan (ASGC),
USA (BNL, UTA, Fermilab),
Ukraine (BITP NASU, NTUU KPI).

Collaborations:
WLCG, RDMS CMS, RDIG

PROJECT LEADER Korenkov V.V.

PROJECT DEPUTY LEADER Strizh T.A.

DATE OF SUBMISSION OF PROPOSAL OF PROJECT TO SOD ________

DATE OF THE LABORATORY STC April 8, 2019 DOCUMENT NUMBER № 80

STARTING DATE OF PROJECT January 1, 2017
(FOR EXTENSION OF PROJECT — DATE OF ITS FIRST APPROVAL)

Form № 25

Multifunctional Information and Computing Complex LIT JINR

MICC

05-6-1118-2014/2019

Korenkov Vladimir Vasilievich

APPROVED BY JINR DIRECTOR	SIGNATURE	DATE
ENDORSED BY		
JINR VICE-DIRECTOR	SIGNATURE	DATE
CHIEF SCIENTIFIC SECRETARY	SIGNATURE	DATE
CHIEF ENGINEER	SIGNATURE	DATE
HEAD OF SCIENCE ORGANIZATION DEPARTMENT	SIGNATURE	DATE
LABORATORY DIRECTOR	SIGNATURE	DATE
LABORATORY CHIEF ENGINEER	SIGNATURE	DATE
PROJECT LEADER	SIGNATURE	DATE
PROJECT DEPUTY LEADER	SIGNATURE	DATE
ENDORSED		
RESPECTIVE PAC	SIGNATURE	DATE

Content

Abstract

The implementation of the project of the JINR LIT Multifunctional Information and Computing Complex (MICC) in 2017-2019 laid foundation for its further development and evolution with regard to new requirements to the computing infrastructure for scientific research at JINR on the basis of modern information technologies in accordance with the Seven-Year Plan for the development of JINR for the 2017-2023 time period.

The rapid development of information technologies and new user requirements stimulate the development of all MICC components and platforms. Multi-functionality, high reliability and availability in a 24x7 mode, scalability and high performance, a reliable data storage system, information security and a customized software environment for different user groups are the main requirements, which the MICC should meet as a modern scientific computing complex.

To fulfill these requirements it is necessary to provide a high-speed telecommunication and network infrastructure as well as a reliable engineering infrastructure.

The MICC project prolongation is aimed at the modernization and development of major hardware-software components of the computing complex, the creation of a state-of-the-art software platform in order to develop methods and algorithms of machine/deep learning (ML/DL) for the solution of a wide range of tasks and applications.

The JINR computing infrastructure consists of numerous computing components and IT-technologies to solve JINR current tasks, from theoretical studies to experimental data processing, storage and analysis. They are the IT-ecosystem for the NICA project (BM@N, MPD, SPD), Tier1 of the CMS experiment at JINR, Tier2/CICC providing support to experiments at the LHC (ATLAS, ALICE, CMS), FAIR (CBM, PANDA) and other large-scale experiments as well as support to users of JINR Laboratories and the JINR Member States (MPD/NICA, BESIII, LRB, FLNR, DLNP, BLTP, FLNP, VBLHEP); the integrated cloud environment of the JINR Member States for support of JINR users and experiments (NICA, ALICE, BESIII, NOvA, Daya Bay, JUNO, etc.); the HybriLIT platform with the GOVORUN supercomputer as a major resource for high-performance hybrid computing.

Given the previous experience of the MICC component development, the financial support requested on the project implementation for 2020-2023 comprises 29867,5 kUSD. The financial support for an additional (above planned in the project) expansion of data storage systems for the neutrino program and the NICA project will be provided within the budget of specific experiments.

Introduction

The direction “Networking, Computing, Computational Physics” in the JINR activities represents one the defining vectors of the development of the entire Institute, which is determined by the growing significant dependence of scientific research on the state and development of information technologies. The development of scientific studies at JINR determine the requirements to computing infrastructures. The JINR LIT Multifunctional Information and Computing Complex (MICC) is the key element of this infrastructure and plays a decisive role in research, which requires modern computing power and data storage systems. The MICC is regarded as the JINR basic facility “Networks and ECM”, which is a combination of complexes, subsystems and other organizational units, which include data processing centers of the Tier1 and Tier2 levels; the cloud infrastructure, the GOVORUN supercomputer, data storage; the data transmission network; the MICC engineering infrastructure; the monitoring system.

A modern computing complex should meet the following requirements: multi-functionality, high performance, high reliability, fault-tolerance and availability, information security, scalability as well as customized software environment for different user groups. To fulfill these requirements in a 24x7 mode, the MICC needs a constant upgrade and expansion of the complex’s possibilities. The rapid development of information technologies does not allow one to fully define specific solutions that will determine the MICC development for the coming years; however, the trends for this upgrade are clear enough. The implementation of the MICC project in 2017-2019 laid foundation for the further development of the complex. The MICC project prolongation is related to the necessity to work in the following directions:

1. Development and improvement of the JINR telecommunication and network infrastructure.
2. Stage-by-stage modernization of the JINR MICC engineering infrastructure.
3. Modernization and development of the IT-infrastructure of the NICA project.
4. Extension of the performance and capacity of storage systems of the Tier1 data processing center for the CMS experiment.
5. Modernization and development of the resources being part of the Tier2/CICC integral component which provides support for the experiments using the grid environment and cooperating with physical groups in JINR as well as for non-grid JINR users and its Member States (MPD/NICA, BESIII, LRB, FLNR, DLNP, BLTP, FLNP, VBLHEP).
6. Extension of the cloud component in order to enlarge a range of services provided to users as well as to create an integrated cloud environment for experiments of JINR (NICA, ALICE, BESIII, NOvA, Daya Bay, JUNO, etc.) and its Member States using the containerization technology.
7. Enlargement of the HybriLIT heterogeneous platform with the GOVORUN supercomputer.
8. Significant extension of resources of the MICC components to meet requirements of neutrino experiments.
9. Development of a unified system for computing resource management aimed at big data processing.
10. Development of a unified data management system for all MICC components (JINR data lake).

State of research

The first years of the implementation of the MICC project, being the JINR basic facility, laid foundation for its further development and evolution with regard to new requirements to the computing infrastructure for scientific research underway at JINR. It is noteworthy that in 2018, together with BLTP, the project on the introduction of the GOVORUN supercomputer into the MICC supported by the JINR Directorate was implemented; its goal is to significantly speed up complex theoretical and experimental studies as well as to develop the NICA megaproject computing.

The implementation of this project on the most state-of-the-art computing architectures and IT-solutions provided users with the opportunity not only to effectively carry out parallel calculations, but also to develop methods and algorithms of machine learning and deep learning (ML/DL) for encompassing a wide range of tasks, including experimental data processing with an introduction of the neural network approach, and to elaborate applications with the artificial intelligence element.

Thus, the JINR computing infrastructure combines a wide spectrum of computing components and IT-technologies providing the opportunity to solve various scientific and engineering tasks facing the Institute, from theoretical studies to experimental data processing, storage and analysis.

It should be noted that the JINR research program for the next decade is aimed at conducting ambiguous and large-scale experiments on the Institute basic facilities and in frames of the worldwide cooperation. This program is connected with the implementation of the NICA megaproject, the construction of new experimental facilities, the JINR neutrino program, the modernization of the LHC experimental facilities (CMS, ATLAS, ALICE), the programs on condensed matter physics and nuclear physics. The implementation of the mentioned above projects requires adequate and commensurable investments in the systems providing the processing and storage of increasing data volumes. The experience of recent years shows that the progress in obtaining research results directly depends on the performance and efficiency of computing resources. In this regard, the further development and performance extension of the JINR MICC as well as the provision of novel IT-solutions to the Complex users and the increase in its operation efficiency are the uppermost tasks of the Laboratory of Information Technologies.

The following main directions can be defined in the implementation of this challenge.

The first one is related to the efficiency extension of the MICC resources being a combination of various computing complexes and infrastructure units, namely the LHC data processing centers of the Tier1 and Tier2 levels, the cloud infrastructure, the GOVORUN supercomputer, the big data processing cluster, the multi-component data storage, the data transmission network, the MICC engineering infrastructure and the monitoring system. The implementation of this direction is impossible without creating a unified resource management system designed to provide:

- a unified interface to computing resources and data storage systems (user and software);
- a unified user authorization system with support for user groups and virtual organizations, integrated with external authentication systems;
- the ability to manage quotas and priorities of the resource usage.

Another important direction in the MICC development plan is the modernization of data storage systems; it is connected with the significant increase of information volumes expected in 2020-2021, which are to be stored and processed. In this regard, it is necessary to meet the following requirements:

- to provide a sufficient resource for storage and fast access to the information during processing;

- to provide a constantly expanding resource for long-term data storage. The volume and speed of its enlargement should be balanced with the flow of information for long-term storage;
- to provide the ability to use a data management system that automates the processes of interaction with storage systems;
- to automate support for the storage system to optimize and minimize costs.

Similar tasks for the modernization of data storage systems lie before researchers' communities in the field of high energy physics (HEP) [1-3] due to the modernization of LHC at the High-Luminosity Large Hadron Collider (HL-LHC), which is expected to launch in the middle of 2026. Corresponding tasks are also relevant for other large-scale experiments such as DUNE in the field of neutrino physics and other sciences (e.g. the SKA project in astronomy) requiring the amount of resources comparable to the LHC, taking into account the successful experience of the WLCG project. To meet the given requirements an evolution of solutions is needed, where for almost two decades HEP, in frames of the Worldwide LHC Computing Grid (WLCG) project, has demonstrated unique opportunities of the distributed computing infrastructure for LHC experiments in managing hundreds of petabytes of data, scaling computing and data storage resources, ensuring transparent user access to the system.

At present, the work on elaborating the development strategy of computing in HEP is in progress. The evolution strategy relies on three issues, i.e. the use of the existing HEP computing infrastructure and its development as a unified computing system for HEP and other sciences; the development of facilities and services for the creation of the HEP data "lake"; the work in the field of creating unified software and software methods.

It is noteworthy that JINR, being its full participant, has taken an active part in the WLCG project since the first day of its operation, and therefore the development and modernization of the MICC computing infrastructure is in the trends of the development of this project.

Another direction of the MICC development is the creation of an effective, unified software and hardware environment for introducing the neural network approach, methods and algorithms of machine learning (ML) and deep learning (DL), big data analytics for solving a wide range of challenges facing the Institute. The active implementation of these approaches including experimental data processing in HEP [4], neutrino experiments [5], astronomy [6], etc. is defined by many factors. The development of computing architectures, especially while using DL methods for training convolutional neural networks, the development of libraries, in which various algorithms are implemented, and frameworks, which allow building different models of neural networks [4-5], can be referred to the main factors. To provide all the possibilities both for developing mathematical models and algorithms and carrying out resource-intensive calculations including graphics accelerators, which significantly reduce the calculation time, a two-component ecosystem for tasks of ML/DL and big data analysis is proposed to develop.

1. WLCG Strategy towards HL-LHC http://cds.cern.ch/record/2621698/files/LHCC-G-169.pdf?version=1
2. A Roadmap for HEP Software and Computing R&D for the 2020s https://arxiv.org/pdf/1712.06982.pdf
3. Evolution of Scientific Computing in the next decade: HEP and beyond http://wlcg-docs.web.cern.ch/wlcg-docs/technical_documents/HEP-Computing-Evolution.pdf
4. Dan Guest, Kyle Cranmer, Daniel Whiteson, Deep Learning and its Application to LHC Physics, 10.1146/annurev-nucl-101917-021019, arXiv:1806.11484 [hep-ex]; S. Farrell et al, "The HEP.TrkX Project: deep neural networks for HL-LHC online and offline tracking", EPJ Web of Conferences 150, 00003 (2017)
5. Alexander Radovic, Mike Williams, David Rousseau, Michael Kagan, Daniele Bonacorsi, Alexander Himmel, Adam Aurisano, Kazuhiro Terao & Taritree Wongjirad, Machine learning at the energy and intensity frontiers of particle physics, Nature, volume 560, pages41–48 (2018)
6. Rodrigo Carrasco-Davis, Guillermo Cabrera-Vives, Francisco Förster, Pablo A. Estévez, Pablo Huijse, Pavlos Protopapas, Ignacio Reyes, Jorge Martínez-Palomera, Cristóbal Donoso, Deep Learning for Image Sequence Classification of Astronomical Events, arXiv:1807.03869v3 [astro-ph.IM]

Description of research

The development of scientific studies at JINR defines the requirements to computing infrastructures. The JINR LIT Multifunctional Information and Computing Complex is the key element of this infrastructure and plays a defining role in research, which requires modern computing power and data storage systems. The MICC is the infrastructure project and aimed at providing users with state-of-the-art computing means to carry out research in a more effective way. The rapid development of information technologies that underlie the MICC, on the one hand, and new user requirements, on the other, are reasons for the continuous modernization of the existing MICC components and determine the necessity of its further development. As a modern highly performant scientific computing complex, the MICC should meet the following requirements:

- multi-functionality,
- high reliability and availability in a 24x7 mode,
- scalability,
- high performance,
- task adapted data storage system,
- information security,
- customized software environment for different user groups.

To fulfill these requirements it is necessary to provide high-speed telecommunications and a modern local network infrastructure as well as a reliable engineering infrastructure that ensures the guaranteed energy supply and air conditioning of the server equipment.

Network infrastructure

First years of the MICC project implementation laid foundation for the further development of the JINR network infrastructure, in particular, the projects related to increasing the bandwidth of the Moscow-JINR telecommunication channel to 100 Gb/s, installing and configuring the equipment of the Institute backbone network to 2 x 100 Gb/s and the distributed computing cluster network between JINR sites to 400 Gb/s were carried out.

The following main directions can be defined in the implementation of the development of the JINR network infrastructure:

1. Development of external telecommunication channels;
2. Development of the JINR local area network;
3. Development of the network infrastructure of the MICC and the NICA megaproject;
4. Development of the LIT laboratory network.

The solution of the first task related to scalability, reliability and fault-tolerance of data transmission means, requires the creation of at least two independent telecommunication channels from Dubna and the modern channel equipment (DWDM) capable of meeting strict requirements in the constantly growing needs of JINR in the capacity of external telecommunication channels.

The work on expanding the bandwidth of the external optical channel will consist of several successive stages:

- first stage: continuation of the modernization of the existing channel-forming telecommunication equipment and 100 Gbps transponders;
- second stage: scaling of the Moscow (MGTS-9) telecommunication segment on the MPLS (MultiProtocol Label Switching) and VRF (Virtual Routing Forwarding) technologies;
- third stage: optimization of the high-speed data transfer routes with participants of high-energy physics collaborations via the networks GEANT, LHCOPN, LHCONE, etc.

The second direction is associated with the development of the Institute inter-laboratory backup network and includes the development of the 100 Gbps segment within the laboratories. Internal laboratory networks should have the same bandwidth as the network channels leading to the laboratories, i.e. 100 Gbps. To carry out this task, it is necessary to replenish central switches with 100 Gbps interfaces. It is supposed to purchase the network inter-laboratory equipment at the expense of the laboratories.

In frames of the third direction, the following works are presupposed:

- development of the MICC local network, i.e. transfer to 100 Gbps interfaces;
- development of the communication component of network subsystems of the MICC and the NICA megaproject using the Multisite cluster network technologies at a speed of at least 400 Gbps;
- creation of the network infrastructure for receiving and transmitting data between the BM@N, MPD, SPD facilities and on/off-line clusters. The given direction presupposes laying new optical networks between DLNP and VBLHEP sites of JINR as well as ensuring maximum reliability and fault tolerance of this network.

Increasing the performance of the network segment of the MICC/Tier2 system will allow speeding up access to data stored in the databases of the LHC experiments.

The forth direction related to the modernization of the LIT local network presupposes the installation of a new network equipment in the 134 building with support of network technologies 100 Gbps Ethernet.

The planned increase in the bandwidth of the external telecommunication channels, the JINR backup and local network is shown in Table 1.

Table 1.

	2020	2021	2022	2023
External Channel	2 x 100 Gbps	4 x 100 Gbps	4 x 100 Gbps	6 x 100 Gbps
Backup Network	2 x 100 Gbps	2 x 100 Gbps	2 x 100 Gbps	2 x 100 Gbps
MultiSite Cluster	4 x 100 Gbps	4 x 100 Gbps	8 x 100 Gbps	8 x 100 Gbps

It is noteworthy that the increase in the bandwidth of external and internal telecommunication channels above the plan will be defined by the needs of network-connected facilities, including the needs of the NICA megaproject, and elaborated due to the financial support from the budgets of experiments, joint grants and other sources.

Network services

One more important direction is the development of network services supporting the network functioning, namely Network services (NOC Cluster) of e-mail, name management (DNS), data caching (Proxy), resource management (IPDB), authorization (Radius, Tacacs, Kerberos), monitoring (NMIS); Service of Single Sign-On (SSO); Information security system; System of testing of users for the knowledge of the operation rules in the JINR computing network, etc.

Network services are configured on two hardware and software server clusters, i.e. the e-mail cluster and the network cluster of virtual services. The e-mail cluster is built on 5 server nodes and functionally divided into two clusters, namely the disk cluster based on CEPH (Ceph is a free software object storage network providing both file and block access interfaces) and the virtual cluster based on proxmox. Proxmox Virtual Environment (Proxmox VE) is a virtualization system with an open source code based on Debian GNU/Linux. The reliability of the e-mail service will significantly increase with the implementation of the project of a high-speed backup storage system based on the disk fault-tolerant cluster (DRDB, CEPH) for organizing “cold” backups (VM backup images of the e-mail and virtual clusters, mailboxes of the mail.jinr.ru service with “deep storage”, backup images of the physical NOC servers, storage of statistical and journal data of long-term storage).

The virtual network cluster includes the following virtual services: name management (DNS), data caching (Proxy), resource management (IPDB), authorization (Radius, Tacacs, Kerberos), monitoring (NMIS); Service of Single Sign-On (SSO); Information security system; System of testing of users for the knowledge of the operation rules in the JINR computing network, etc.

The development of the project “My account” will be in progress; it will provide users with access to general Institute services and the ability to request computing and information resources required for their professional activity. Some resources can have certain policies for providing them to various categories of users. Upon users' request, this resource will be provided to them, or users will get information about the rules of its provision.

The implementation of these tasks requires the modernization and purchase of a new network equipment and software.

Engineering infrastructure – guaranteed power supply, climate control system

The MICC engineering infrastructure is designed to ensure the reliable, uninterrupted and fault-tolerant operation of the information and computing system and the network infrastructure. The use of the integrated approach to building the MICC engineering infrastructure allowed one to elaborate algorithms of the equipment operation and interaction of separate systems both in a normal operation mode and in emergencies, which ensured the uninterrupted performance regardless of external factors.

Attention should be paid to the fact that another integral feature of the engineering infrastructure is its scalability, the possibility of which was defined on the basis of the analysis of computing equipment growth prospects for 3-4 years.

The main task of the next stage of the modernization and development of the MICC engineering infrastructure relies on providing the uninterrupted power supply, air-conditioning and ventilation of all the components of the complex in accordance with the growth of computing powers. The modernization of the fire safety complex, the automated system of dispatching and management of the engineering infrastructure is also planned in 2020-2023.

Guaranteed power supply:

In 2017-2019 the modernization of the power supply system of the 134 LIT JINR building, together with the replacement of the transformers with a nominal value of 1000 kVA with new transformers of 2500kVA, was performed, and the transfer to the sharing use of the uninterruptible power supplies (UPS) and diesel-generator units (DGU) was carried out, which meets the strictest reliability standards imposed on the MICC functioning in a 24x7x365 mode by the 1st class. The guaranteed power supply system (GPSS) created during first years of the MICC project implementation ensures:

- guaranteed power supply of connected consumers;
- automatic start of the diesel-generator;
- automatic load switching from the major external power supply network to the diesel-generator and backward;
- sending an alarm to the dispatcher's post in case of an emergency with the DGU equipment.

The complex of the system consisting of UPS and DGU is a guaranteed power supply system, which ensures the complete energy independence of consumers from the external network.

The further modernization of GPSS should provide all MICC components with the high-quality power supply in conditions of the quality decrease of short-term loss of the main network via uninterruptible power supplies (UPS). At present, UPS operating on the principle of double (or on-line) conversion are widely used. The principle of the operation of UPS of this class relies on the following: the input alternating voltage is converted into direct-current voltage by the rectifier, and then back into alternating by the inverter. To increase the MICC GPSS fault tolerance a system of reservation (2N) is applied in the uninterrupted power supply system. In this regard, it is planned to expand the power of uninterruptible power supplies to 2.4 mVA based on eight APC Galaxy 7000, which will fully provide the MICC halls on the 2nd and 4th floors with electricity until 2023. APC Galaxy is a double conversion system that ensures the following: voltage and frequency stabilization, phase continuity of output voltage in all modes, absence of load impact on the main network, full filtering of pulses and noise of the main network, high information security, which provides the most reliable protection of the connected equipment from power network failures.

Together with this, the modernization of distribution panels for the guaranteed power supply and of load distribution sections for own needs of the building will be performed, as well as a power supply system for four air-cooling modules and cabinets for the network and computing equipment of the MICC complex will be created. The MICC GPSS development plan presupposes:

- Creation of a power supply system for four air-cooling modules and cabinets for the network and computing equipment of the MICC complex;
- Development and installation of the power supply system of the hall on the 4th floor;
- Elaboration of the project and connection of an additional source of power supply from the Dubna HPS (hydroelectric power station) to provide the MICC with the power supply of the first category;
- Completion of the reconstruction of the power supply and control of air handling units;
- Modernization of schemes and backbones for the power supply of the MICC monitoring and control rooms on a new element base.

Climate control system:

The MICC existing refrigeration equipment is a complex of interconnected installations of various air and liquid cooling schemes, with the help of which the corresponding temperature regime ensuring the MICC functioning in a 24x7x365 mode is created. Heat exchangers are evaporators and condensers as well as recuperators. They provide the implementation of important heat exchange processes involving a refrigerant and a coolant. Pumping stations are controlled automatically, which minimizes losses in case of emergencies in the operation of pumps, mainly circulation ones. It should be noted that pipeline communications and refrigeration units are quickly exposed to the corrosive process and need the restoration of separate damaged elements from time to time.

The development and modernization of the climate control system should rely on novel technological solutions applied in modern computing centers to create a required microclimate inside the building and should satisfy the needs of the development of all MICC components. At present, the MICC climate control system has the following components: free cooling of the equipment with cooled air of the machine hall; raised floor supply of cold air with a forced exhaust of hot air by ventilation panels; cooling of the cold corridor of the module by in-row conditioners; liquid cooling of the computing machine elements. According to the type of heat removal, the MICC climate control system refers to the mixed type of performance that combines systems with the evaporation of a refrigerant and systems with an intermediate coolant.

The further development of the climate control system is connected with the creation of a centralized cooling system with dry (for the winter period) and wet (for the summer period) cooling towers in the 134 LIT JINR building with a capacity of up to 1.8 mW and a possibility of extension up to 3 mW. The implementation of the MICC climate control system will ensure the connection of the refrigeration equipment of any type and principle of operation, including energy-efficient cooling systems with free cooling, which, with proper design and setup, allow increasing the energy efficiency of the computing complex and reducing costs of energy supply.

To develop and satisfy the growing needs of all MICC components, the creation of climate control systems with cold, warm and hot corridors in the hall of the 2nd floor, the installation of four new modules with in-row conditioners in the machine hall of the 2nd floor and the organization of a cooling system in the machine hall of the 4th floor are required. The installation of external cooling towers and the organization of a two-circuit cooling system are planned. In the external circuit connecting cooling towels and intermediate heat exchangers, the coolant is a water solution of ethylene glycol, and in the internal circuit, pure water circulates between heat exchangers and air conditioners (cabinet and/or in-row).The use of ethylene glycol in the external circuit is explained by a low temperature of freezing of this substance. The use of cooling towers will ensure a function of “uninterrupted cooling”, moreover, in case of a power outage the system is able to provide cooling of the computing complex until the DGU launch due to cold water stored in battery tanks. It is possible owing to the fact that to maintain the operation of the air-conditioning system it is sufficient to ensure the functioning of pumps, ventilators and conditioners from UPS. Heat removal for cooling the MICC high-loaded racks (more than 30 kW) presupposes the use of in-row conditioners.

Another technology, which will be used during the modernization of the MICC climate control system, is heat recuperation (a type of free cooling) playing a key role in the creation of an effective air cooling system for the machine hall. The efficiency of this solution is defined by the PUE coefficient, which is the ratio of the total amount of consumed electricity (servers + cooling system) to the amount of electricity consumed by servers. When introducing recuperation installations of a centralized and local type into the ventilation system, the expected PUE indicator will be 1,2.

Due to the expected enlargement of computing powers of the GOVORUN supercomputer, the expansion of the hot water cooling system to 600 kW is planned.

The planned extension of power consumption and cooling supply is shown in Table 2.

Table 2.

	2020	**2021**	**2022**	**2023**
Power consumption	500 kVA	600 kVA	800 kVA	1000 kVA
Cooling system	300 kW	500 kW	650 kW	800 kW

The funding of the development and modernization of the engineering infrastructure will be elaborated due to the financial support from the budget of the Laboratory and other sources.

Off-line NICA cluster within MICC

Work on the construction of a unique NICA accelerator complex is in progress; it requires new approaches to the implementation of a distributed infrastructure for experimental data processing, analysis and storage.

The computing for the NICA megaproject should provide data acquisition from detectors and data transmission for processing and analysis. To perform the given tasks, computing has certain requirements including requirements to the network infrastructure, computing architectures, storage systems as well as appropriate software of the system and for data processing and analysis. Developed computing models should take into account the trends in the development of network solutions, computing architectures and IT-solutions, which allow combining supercomputer (heterogeneous), grid and cloud technologies and creating distributed, software-configured HPC platforms on its basis. The use of such solutions for data processing and analysis requires the creation of software environments, which provide a necessary code abstraction enabling to implement the required functionality for a wide range of computing tools.

One of the uppermost components of the NICA computing complex is the computer network combining clusters and servers inside clusters into a unified computing infrastructure of the complex.

The distributed information and computing cluster of the NICA complex in its basic configuration should provide the processing and storage of up to 10 petabytes of data per year. The complex consists of territorially distributed on-line and three off-line clusters (Off-line cluster 216, Off-line cluster «NICA center», LIT Off-line cluster). Thus, as all large-scale centers of experimental data processing and storage, the computing complex is territorially distributed and combines all the components located in VBLHEP and DLNP sites of JINR in a unified high-speed local multi-site computer network with 4 x 100 Gb/s.

The network equipment of the central telecommunication node, i.e. the core of the switching and routing system of on-line and off-line clusters, is located in two JINR sites and implemented on four multifunctional switches Cisco Nexus 9504 with the full-mesh topology for maximum reliability and performance.

The main task of the LIT off-line cluster is the creation of a two-level (disk-tape) storage system for NICA experiments, as after the first stage of these experiments, significant storage volumes will be required (from 2,5 PB to 70 PB per year).

One of the key components of the computer unit of the project “NICA Complex” is a cluster file system. At present, several cluster file systems are used, namely GPFS, Lustre, dCache, Ceph, EOS, GlusterFS, etc. The most appropriate one for on-line and off-line clusters is the EOS file system. EOS is a distributed, parallel, linearly scalable file system with the opportunity to protect it from failures.

Special attention is paid to novel perspective directions in creating distributed data storages (Data Lake), integrating Big data and supercomputer technologies, methods of machine learning.

The implementation of various computing models of the NICA megaproject requires confirmation of the model’s performance, i.e. meeting the requirements for temporal characteristics of data acquisition from detectors with their subsequent transfer to processing, analysis and storage, as well as the requirements for the effectiveness of modeling and processing of events in the experiment. For these purposes, it is necessary to carry out tests in a real software and hardware environment, which should contain all required components. The GOVORUN supercomputer, commissioned in LIT, can become such an environment; it contains the latest computing resources and data storage resources including the ultrafast data storage system, which provides a high speed of data acquisition up to hundreds of gigabytes per second with the possibility of a linear extension of performance and capacity of the system up to 1000 times.

Table 3 illustrates the data on the main subsystems of the NICA complex (event size, transmission rate and total data volume per day) used to calculate the major characteristics of the computing unit of the NICA complex.

Table 3.

Subsystems of the NICA complex	Data acquisition rate (GB/s)	Event generation rate (kHz)	Size of the event (MB)	Total size of events per second (GB/s)	Average data transfer rate (Gb/s)	Data volume per 24 hours (TB/24 hours)
Accelerators						
2019 - 2020	0,5				0,1	4
>2020	1,5				0,3	10

BM@N						
2019-2020		30	0,5	15	20	100
>2020		50	0,7	35	100	300
MPD						
2021-2022		0,1	1	0,1	10	200
>2022		6	2	12	100	600
SPD						
>2023		50	0.5	25	100	1000

The network infrastructure of computer clusters of the NICA complex refers to a basic configuration and is elaborated due to the financial support from the budgets of the NICA complex experiments and other sources.

Information and computing environment for neutrino experiments with the JINR participation

The needs of modern neutrino experiments in data storage volumes and computing power required to obtain relevant scientific results have increased significantly. To effectively use resources of the information and computing environment of neutrino experiments, in which JINR employees participate, a decision on the creation of a unified neutrino information and computing platform (environment) based on the MICC resources was made by the directorate of LIT and DLNP.

At present and over the next several years, the following neutrino scientific experiments with the JINR participation require storage resources and computing power: Baikal-GVD, JUNO, NOvA.

Baikal-GVD

The Baikal-GVD experiment uses the JINR cloud infrastructure resources as computing ones and the cloud storage resources as data storage ones.

The local group of the experiment plans to increase the amount of cloud resources by 300 CPU cores, 3000 GB of RAM and 200 TB of disk space per year.

The total expected amount of resources of the experiment by year is presented in Table 4.

Table 4.

	2020	2021	2022	2023
Total number of cores	684	984	1284	1584
Total RAM volume, TB	6.75	9.75	12.75	15.75
Disk space, TB	400	600	800	1000

JUNO

In frames of the JUNO experiment, after the launch of the installation, primary data in the amount of 2 PB/year are expected. The experiment is scheduled for 20 years. Accordingly, the total amount of raw data at the end of the JUNO operation is estimated at 40 PB. Raw data will be permanently stored on tape drives.

In frames of the JUNO collaboration, it was agreed that raw data would be stored at least in two data centers, namely one in China (IHEP) and one in Europe. It is planned to use the LHCONE infrastructure for transmission.

It is expected to store one full copy of raw data in the JINR MICC.

Apart from raw data stored on tape drives, it is necessary to provide disk storage used in data processing. The estimated volumes by data type are presented in Table 5.

Table 5.

	Required volume, TB		
Data type	Permanent part	For every year	Total (for 20 years)
Reconstruction from raw data		10	200
Calibration data		1	20
MC simulation		150	3000
Data analysis	100		100
Total	100	161	3500

As seen from Table 5, the main part of the space is required for storing simulation data (MC). The given data will be updated twice a year, and their volume will be proportional to the volume of acquired data.

The other data types require a much smaller volume. The additional amount of space, required for storing secondary data in addition to raw data, is estimated at 3.5 PB at the end of the experiment.

The needs of the entire JUNO experiment in computing servers are estimated at 12000 CPU cores, which will be used for uninterrupted raw data processing and modeling as well as for analysis.

The JINR preliminary contribution to the end of the experiment is estimated at 4000 CPU cores provided to the JUNO collaboration.

Furthermore, the possibility of using graphics coprocessors (GPU) for experimental data processing is actively studied. In particular, GPUs can be used to reconstruct the position and energy of the event. GPUs will be used by the GNA software developed in the Laboratory of Nuclear Problems for the needs of neutrino experiments.

The use of GPU for the given tasks is under consideration, that is why at the time of writing the document it is difficult to estimate the number of GPU cores required for the analysis of JUNO data.

The total expected amount of the experiment resources by year is shown in Table 6.

Table 6.

	2020	**2021**	**2022**	**2023**
Total number of cores	2096	3096	4096	5096
Total RAM volume, TB	32.5	48.5	64.5	80.5
Disk space, PB	1.13	1.63	2.13	2.63
Capacity of tape storage, PB	0	4	6	8
Disk space for tape storage cache, PB	0	0.4	0.6	0.8

NOvA

For several years the NOvA experiment has successfully used the JINR cloud infrastructure resources as computing ones and the cloud storage resources and the dCache storage as data storage resources.

The local group of the experiment plans to increase the amount of cloud resources by 120 CPU cores, 640 GB of RAM and 100 TB of disk space per year.

The total expected amount of resources of the experiment by year is presented in Table 7.

Table 7.

	2020	2021	2022	2023
Total number of cores	660	780	900	1030
Total RAM volume, TB	3.14	3.78	4.42	5.06
Disk space, TB	388	488	588	688

All hardware resources for neutrino experiments are planned to be purchased due to the budgets of the experiments. The infrastructure support (racks, power supply and cooling) is provided by the Laboratory of Information Technologies.

Development of the MICC grid component Tier1 of the CMS experiment and the MICC integral component Tier2/CICC

The grid center resources of the JINR MICC are part of the global grid infrastructure WLCG (Worldwide LHC Computing Grid), developed for the LHC experiments. JINR LIT actively participates in the WLCG global project, which aims at providing distributed computing resources for annual data processing, storage and analysis. About 30 PB of raw data is acquired by the LHC experiments, the volume of processed data is significantly higher. The work on the use of the grid infrastructure in frames on the WLCG project is carried out in cooperation with the collaborations such as CMS, ATLAS, ALICE and major international centers, which operate as Tier1 centers of the CMS experiment (CH-CERN, DE-KIT, ES-PIC, FR-CCIN2P3, IT-INFN-CNAF, US-FNAL-CMS) and as Tier2 grid centers located in more than 170 computing centers of 42 countries worldwide. Since the beginning of 2015, a full-scale WLCG Tier1 site for the CMS experiment at the LHC has been operating in JINR LIT. It has consistently occupying the 2nd place in the number of processed events and demonstrating almost 100% availability and reliability. The importance of increasing computational performance and data storage volume of the center is determined by the research program of the CMS experiment, in which JINR physicists take an active part in frames of the RDMS CMS collaboration. In terms of hardware, a linear increase in the characteristics of Tier1, Tier2/CICC is planned in accordance with the figures in the LIT Seven-Year Plan.

It is planned to launch Run3 at the LHC by the summer of 2021. The CMS and ATLAS plans do not require a cardinal increase in computing and storage resources at this stage. ALICE plans to increase its resource needs for Run3 almost twice. The fulfillment of the seven-year plan figures will ensure the required level of resources for all LHC collaborations at Tier1 and Tier2 in JINR.

In the coming years, it is necessary to increase the information storage capacity on the EOS system, which was put into operation at the beginning of 2019. This data storage and access system should become the major system for all MICC components, and later for all JINR computing resources. There are prospects of using EOS for WLCG collaborations.

The development of the data storage system on robot tape libraries is required. In addition to this installation for the CMS experiment at Tier1, it is required to create an installation for the NICA experiments in 2020-2021 as well as other WLCG collaborations and user groups. The update of the existing tape library IBM TS3500 to the TS4500 with an increase in its capacity up to 30-40 PB is needed.

Table 8 illustrates the planned amount of resources by year.

Table 8.

	2020	2021	2022	2023
Tier1 CPU, kHS06	200	240	300	350
Tier1 disk, PB	8.8	10.88	13.1	16.1
Tier1 MSS, PB	25	30	35	42
Tier2 CPU, kHS06	110	130	150	170
Tier2 disk, PB	4.3	4.6	4.7	5
EOS ALICE, PB	1.2	1.4	1.8	2
EOS MICC, PB	10	30	50	60
MSS/MICC, PB	10	20	30	40

Software of Tier1, Tier2/CICC requires a cardinal update in 2019 and further permanent support and upgrade. It concerns both the WLCG middleware at Tier1 and Tier2 and the CICC software.

It is necessary to install new batch processing systems and a job scheduler at the Tier1 and Tier2/CICC computing farms. The installed Torque and Maui systems are already outdated and unable to service the increasing number of computing slots/cores as well as the growing number of users and WLCG virtual organizations. Free software will be installed, as commercial software is extremely expensive. There are two perspective systems to install, i.e. HTCondor and Slurm. Both systems will be tested on our DevLab resources.

A unified system of access to software CVMFS (CernVM File System) should be further developed. CVMFS is regarded as EOS, i.e. a unified universal system of access to different software installed on one server and available on all interactive and computing machines of the MICC and JINR. For a long time CVMFS has been successfully used for these purposes by numerous WLCG virtual organizations. CVMFS will replace specialized software repositories in AFS in the nearest future.

JINR cloud

As the resource base of the JINR cloud develops, its computing resources and storage disk space are planned to be provided to individual JINR users as well as to various scientific projects, in which JINR has participation. A significant increase in the demand for cloud resources is expected from neutrino experiments (NOvA, JUNO, Daya Bay, Baikal-GVD, etc.). The transition of users of the mentioned above experiments to the use of centralized cloud resources will eliminate the need for local scientific groups to independently support computing resources and data storages, and they will be able to fully concentrate on their research.

By the end of 2020 in the JINR cloud, as cloud resources purchased on the LIT budget, it is planned to have about 2000 cores, 9.7 TB of RAM, 1.2 PB of disk space in the high-speed cloud storage based on Ceph with triple data replication.

Table 9 shows the approximate amount of resources that are planned to be purchased on the budget of the Laboratory of Information Technologies by the end of each year.

Table 9.

	2020	2021	2022	2023
Total number of cores	2000	2400	2800	3200
Total RAM volume, TB	10	12	14	16
Disk space capacity, PB	1.5	2	2.5	3

Extension of the HybriLIT heterogeneous platform including the GOVORUN supercomputer

The development of the HybriLIT heterogeneous platform will be carried out both in terms of expanding computing resources of the GOVORUN supercomputer and in terms of developing the services deployed on the platform.

The first direction is related to the increasing demands of JINR users, particularly in the field of theoretical studies within quantum chromodynamics on lattices, and the satisfaction of user requests from the JINR Member States as well as the growing involvement of the supercomputer into the implementation of the leading JINR projects, namely the NICA megaproject and the neutrino program. Expanding the computing power of the supercomputer will be performed both by increasing the number of nodes of the CPU-component implemented on the basis of liquid cooling and by increasing the number of graphics accelerators of calculations from Nvidia, taking into account the development of computing architectures.

The second direction is associated with the development of software and information services providing to users of the heterogeneous platform. In particular, it is planned to develop an ecosystem for tasks of ML/DL and big data analysis, which is primarily aimed at the development of algorithms based on recurrent and convolutional neural networks with deep training for solving problems of fast recognition of multiple tracks in particle physics experiments, including the NICA megaproject and neutrino experiments. Furthermore, it is planned to develop the HLIT-VDI service designed for calculations in frames of the packages of applied programs with a developed graphics interface such as Maple, Mathematica, Matlab, COMSOL Multiphysics, FLUKA, etc., which are actively used by the Institute Laboratories both for theoretical studies and engineering and applied tasks. To improve work with users and more effectively use the HybriLIT platform resources, a personal account will be introduced into the software and information environment; it will enable the user to create requests on solving different types of problems appearing during the work, e.g. installing the necessary software, troubleshooting technical issues, providing information on the system operation, the statistics on the use of the platform resources for all components, including the visualization of the workload state of computing nodes, as well as the current status of available resources in queues, and monitoring of the execution of own tasks.

In 2019 it is planned to replenish the GOVORUN supercomputer with “warm” storage based on the SSD "Ruler" under the management of the Lustre file system with a total volume of 384 TB with the transfer of users’ home directories and projects’ data (MPD and BM@N) to it, which, together with the fast DSS, will allow implementing an effective hierarchy of data processing.

The increase of the GOVORUN supercomputer performance, as well as the engineering supply including the liquid cooling system of the CPU-component, the uninterrupted power supply and the development of the platform network infrastructure will be carried out in accordance with the Seven-Year Plan for the development of JINR in the direction “Information Technologies”. The total increase in the performance of the CPU- and GPU-components will amount to 90 Tflops for double-precision operations per year. The planned increase in the performance of the GOVORUN supercomputer by components is presented in Table 10.

Table 10.

HybriLIT heterogeneous platform. GOVORUN supercomputer.	2020	2021	2022	2023
Performance of the CPU-component (Tflops, for double-precision operations)	260	320	380	450
Performance of the GPU-component (Tflops, for double-precision operations)	330	360	390	420
Total performance of the supercomputer (Tflops, for double-precision operations)	590	680	770	870

It is noteworthy that the enlargement of the supercomputer performance above the plan will be defined by user needs, including the needs of the NICA megaproject and the neutrino program, and elaborated due to the financial support from the budgets of experiments, joint grants and other sources.

Data storage system (JINR Data Lake)

The uppermost direction in the MICC development plan is the modernization of data storage systems. In 2020-2023 a significant increase in the amount of information is expected, which is needed to be stored and processed and the volumes of which were not entirely clear when forming the Seven-Year Plan for the development of JINR for 2017-2023. At the same time, it is very difficult to evaluate the requirements to storage systems due to the evolution of data and processing models. However, the following requirements should be met:

- to provide a sufficient resource for storage and fast access to the information during processing;

- to provide a constantly expanding resource for long-term data storage. The volume and speed of its enlargement should be balanced with the flow of information for long-term storage;

- to provide the ability to use a data management system that automates the processes of interaction with storage systems;

- to automate support for the storage system to optimize and minimize costs.

In the coming years, it is necessary to increase the information storage capacity on the EOS system, which was put into operation at the beginning of 2019. This data storage and access system should become the major system for all MICC components, and later for all JINR computing resources. There are prospects of using EOS for WLCG collaborations. The system has already been used to store data from BM@N. Storing uppermost data or data distributed territorially requires the creation of replicas, which reduces the available space multiple to the number of replicas. A part of the EOS disk resources can be purchased with funds from the experiments and JINR Laboratories; such servers can be installed not only at LIT, but also at the other JINR subdivisions, and combined into a system with unified management.

The forecast of volume changes for storages of WLCG, NICA and other virtual organizations is based on the organizational differences of the experiments. The pledged resource mechanism for WLCG allows one to define (set) its volume growth and temporal dynamics. For CMS Tier1 an increase in the data flow is expected starting from 2020. The CMS requests for data storage volumes will be limited to site offers. At present, CERN evaluates an increase in the data flow of 2021 to 2018 by a factor of 1.5. Analyzing the dynamics of the increase in 2015-2018 shows the progression that CERN describes as adiabat. Based on the fact that 10 PB for tapes and 10 PB for disks were claimed for 2019, by 2023 the figures should be 20-30 PB for tapes and 15-20 PB for disks. The expansion of storage in Tier2 from 2 to 4-5 PB by 2023 is expected. The enlargement of the dCache storage capacity in Tier1 and Tier2 is almost consistent with the LIT Seven-Year Plan. One of the options for developing Tier1 and Tier2 storages is to consolidate and import local installations of WLCG sites into a “Data Lake”; such projects are already under development and testing in WLCG. It does not mean that a constant expansion of the resources of certain WLCG sites is not needed, but it may smooth an unsteady increase in resources of various organizations. We support ALICE at the Tier2 level on a separate EOS installation. It is determined by the collaboration rules and cannot be changed. According to the experiment plans, starting from 2021 the intensity of the flow will increase by two times. We need to expand the storage capacity for ALICE in 1.5-2 times.

The forecast of volume changes for non-LHC experiments is a more sophisticated task. By 2020 all dCache Tier2/CICC will be replaced with EOS/MICC. The data of WLCG virtual organizations, local users and user groups will be transferred to EOS/MICC. It will reduce costs and efforts to support large data storages at LIT and JINR. EOS/MICC will be available throughout JINR and WLCG.

Due to the significant increase in needs of NICA experiments and the extended use of EOS/MICC, it is necessary to plan a substantial growth in the storage capacity. It is reasonable to plan a significant and permanent increase in the EOS/MICC volumes. It should be noted that this project was not included in the LIT Seven-Year Plan.

One should point out that EOS/MICC will store data not only of LHC collaborations, but also of other ones such as JUNO, BES, etc.

For long-term storage nothing but tapes is foreseen. At the beginning of the planning period, the cartridge capacity is equal to 12-20 TB for LTO8 – TS1160. By the end of the period, it will amount to 59-60. It means that the capacity of the existing library TS3500 can be upgraded to 50-60 PB, and the capacity of the planned library for 2020 will reach up to 120-150 PB.

For disk drives, by the beginning of the period, we have 16-20 TB per disk, while by the end 40 is expected, i.e. a 2-2.5 times increase in density. Although the price of the fast memory SSD/NAND drops fast, it is not so attractive for long-term storage of information with online access (HDD), however, it can be used for some applications.

For the 2020-2022 time period the volume of information for NICA experiments and the neutrino program is not defined.

On this basis, it is necessary to focus on creating a long-term storage infrastructure, i.e. the installation of new storages and the modernization of the existing ones. The expansion of the purchase of the required number of cartridges should be provided by the experiments (except for the LHC).

When planning the enlargement, one should keep in mind the following: disk memory is extended at the rate of the expansion of processors, while tape memory is enlarged in accordance with the growth of the data flow.

The approximate figures by the end of the planning period (2023) are presented in Table 11.

Table 11.

Name	Volume of disk pools, PB	To what extent the expansion will be, PB	Volume of tapes, PB	To what extent the expansion will be, PB
T1 Buffer	1	2-3	8	40
T1 DISK	8.3	15-20	-	

T2	2	7	-	
T2/CICC	0.14	Transfer in EOS/MICC	-	
EOS/MICC	4	60		
NICA tapes	0		0	30
Cloud	1.0	2.5	-	-
HybriLIT	0.05	0.5	-	-

Apart from expanding the storage system resources (Table 9), it is planned:

- to constantly develop software during the planning period;
- to constantly support dCache and transfer to new versions of Enstore;
- to fully authorize access to storages in frames of JINR including the NICA storage at VBLHEP;
- to introduce EOS/MICC into WLCG as an opportunistic data storage.

MICC unified resource management system

At present, the MICC consists of various computing and data storage resources. A natural heterogeneity of the given systems complicates the process of their effective operation. The concept of operation in this context includes the use of resources by users, support, monitoring and resource management by LIT. The lack of a unified management system significantly affects the assessment of resource efficiency.

The main objectives of a unified resource management system are:

- to provide the ability to process large amounts of data;
- to provide the possibility to organize massive computing tasks;
- to optimize the efficiency of the use of computing and storage resources;
- to effectively monitor the resource loading;
- to consolidate the accounting for the use of resources;
- to provide a unified interface of access to resources.

A resource management system should meet the following requirements:

- a unified interface to computing resources (user and software);
- a unified interface to data storage systems (user and software);
- a unified user authorization system with support for user groups and virtual organizations, integrated with external authentication systems;
- the ability to manage quotas and priorities of the resource usage.

A unified resource management system is multicomponent. Some of the components are autonomous systems, which can be used separately, the others are auxiliary. The major components of the system are:

- a management system of load on computing resources (tasks and jobs);
- a data management system;
- an information system;
- an authentication and authorization system.

The auxiliary components are:

- monitoring;
- an accounting system;
- a management service for configurations of working nodes and servers;
- a service of deployment of applied software.

The creation of a unified resource management system of the JINR MICC does not presuppose the development of all components “from scratch”; it implies the integration and, if necessary, the refinement of existing systems.

The JINR LIT staff took part and are experienced in the development and integration of complex software products to implement such a system.

Unified data flow management system

Dozens of physical experiments are conducted at JINR, each of which creates and supports its own computing infrastructure or uses the existing IT-infrastructure of JINR and its Member States. LIT provides experiments with a wide range of IT-services for computing: computing powers, services of storage, authentication and authorization, etc.; it plans to ensure a unified user and software interface to computing and storage resources, but it does not offer ready-made solutions for the organization of the processing model. Despite the fact that each experiment is unique, the processes of data processing for different experiments have similar stages. At present, it is widespread when each of the new experiments elaborates its own processing management system or adapts one of the systems originally developed for another experiment.

The problem of organizing calculations faces both the experiments, which operate small amounts of data, and the projects working with Big Data. For the former, the development of such a system is a real challenge, for which the experiments simply lack resources. For the latter, the elaboration of an effective data flow management system is a key to success of the experiment itself (Rolf Hoyer’s report at the Higgs workshop).

At present, there is not a single implementation of a workflow management system, which could unconditionally be called universal. Considering the number of experiments underway at JINR, the development of the given system, either as a universal product or an expandable platform, is an extremely promising direction. A unified data processing management system will simplify the process of launching the processing of data from new experiments, optimize the use of existing computing resources and ensure a cardinally new level of service. For LIT such a system will allow increasing the efficiency of the use of computing resources due to a better forecast of data flows.

The system should provide:

- the ability to ensure an isolated interface for each experiment;
- the possibility to form a chain of data processing in accordance with the experiment computing model;
- processing on available computing resources.

The system consists of:

- an interface for the formation of jobs, chains and groups of jobs;
- a database for storing job definitions;
- a job generation system based on certain parameters, the preparation of input data through the data management system of the MICC unified resource management system, the selection of suitable computing resources, sending jobs to the load management system of the MICC unified resource management system and ensuring control over their implementation.

Monitoring systems of MICC LIT JINR

The MICC further development is determined by the development of the level of services it provides. To maintain the level of quality of the provided services, an organization of functional and service monitoring is required. The present project encompasses the further development of the monitoring system and its expansion to an information and analytics system of monitoring, the diagnosis of system components and the determination of emergencies. In case of malfunctions, the elaborated system should retain all the information necessary to identify an emergency.

The tasks of service monitoring include:

1. Checking the functioning of services with the help of test jobs.
2. Monitoring the performance of services.
3. Sending notifications in case of emergencies.
4. Visualizing the state of services.

The implementation plan consists of the following steps:

1. Collecting the requirements for monitoring different services.
2. Designing a monitoring system.
3. Developing a system.
4. Adding services and commissioning.

In 2020, the monitoring of parameters of the JINR LIT external engineering infrastructure, including the following components: diesel generators, cooling towers, external elements of the cooling system, will be added to the system. The work on the expansion of the monitoring system by incorporating new nodes of the computing infrastructure of Tier1, Tier2, CICC, the GOVORUN supercomputer, will be in progress.

The work on the development of additional tools for analysis and visualization based on Graphana is planned for 2021. The work on the introduction of new computing resources and elements of the MICC storage system into the monitoring system will be in progress.

The creation of an algorithm to ensure the fault tolerance of the main node of the monitoring system is scheduled for 2022.

It is planned to develop new technological approaches for monitoring, analysis and optimization of distributed computer systems of data processing for large-scale scientific experiments based on the Big Data technology. The creation of the JINR MICC intelligent monitoring in frames of the complex operation in the WLCG distributed environment (in particular, the Tier2 and CMS Tier1 infrastructures) is planned as well; it allows detecting and predicting problems in the operation of basic services, equipment and engineering systems.

Information security

The tendency of threats to information security revealed the need for creating a new class of information security systems. In this regard, it is necessary to elaborate solutions allowing one to form and ensure information security policies for the work of IT-administrators, audits and users of JINR corporate information systems. The given solutions should help to follow IT-security requirements protecting private personal data.

The tasks of the security system are to monitor the actions of users and system administrators and evaluate the threads to their actions.

Currently, the JINR information security system is built on the Intrusion Detection Systems (IDS) technology. Due to the growing number of threads to information security, the system needs to be upgraded. To increase the efficiency of the system it is planned:

- to introduce a system service for the analysis of DPI data packages using a specialized gateway;
- to adapt the 6-tier structure of information security control at JINR diversified to:
  1) available public network,
  2) networks of computing complexes such as Tier-1, Tier-2, MICC, NICA,
  3) JINR Management network,
  4) access control network,
  5) networks of research physical installations,
  6) particularly important networks (IBR, NICA, etc.).

Functional capabilities of the ungraded security system should include the following components:

- centralized access control in real time;

— account management (including on target devices);
— providing a single entry point in the system Single Sign-On (SSO);
— activity tracking and record of sessions (such as RDP, SSH, Telnet, VNC, etc.);
— generating statistics and reports on user actions.

Conclusion

The above-planned work on the modernization and further expansion of the MICC resources is defined by the rapid development of information technologies and new requirements of the experiments conducted at JINR and with the JINR participation. Multi-functionality, high reliability and availability in a 24x7 mode, scalability and high performance, a reliable data storage system, information security and a customized software environment for different user groups are the main requirements, which the MICC should meet as a modern scientific computing complex.

It is noteworthy that the part of work, mainly on experiments at the LHC, is carried out in frames of the WLCG project and related to the operation of JINR computing resources and storage systems included in the grid environment and providing processing, storage and analysis of data from the LHC experiments. A full-scale WLCG Tier1 site for the CMS experiment at the LHC plays a special role in this infrastructure. The importance of developing, upgrading and expanding the computing performance and data storage systems of the center is dictated by the research program of the CMS experiment, in which JINR physicists take an active part in frames of the RDMS CMS collaboration. The fact of the creation and support of the work of the JINR Tier1 site demonstrates a high qualification level of the JINR LIT staff ensuring the functioning of this MICC component. There are only seven similar CMS centers in the world, and the JINR site is regularly ranked second in the number of processed events showing almost 100% level of availability and reliability.

The work on the creation of an integrated cloud environment with the JINR Member States was carried out by the project authors at its first stage. The GOVORUN supercomputer commissioned in March 2018 is successfully functioning for users.

The work performed at the first stage of the MICC project is presented in the project report, to which a list of publications for 2017-2019 is attached.

In the framework of the project, it is planned to defend at least four theses in 2019-2021. The work results were reported at the international conferences such as NEC 2017, CHEP 2018, GRID 2018, and at numerous international workshops.

Estimation of human resources

The executing team of the project consists of 62 employees of high qualification, which is proved by the uninterrupted functioning of the entire complex as well as by the effective introduction of a new equipment in a short period of time using the example of the GOVORUN supercomputer.

The diagram of Fig.1 shows the age distribution of the LIT staff involved in the project as principal performers.

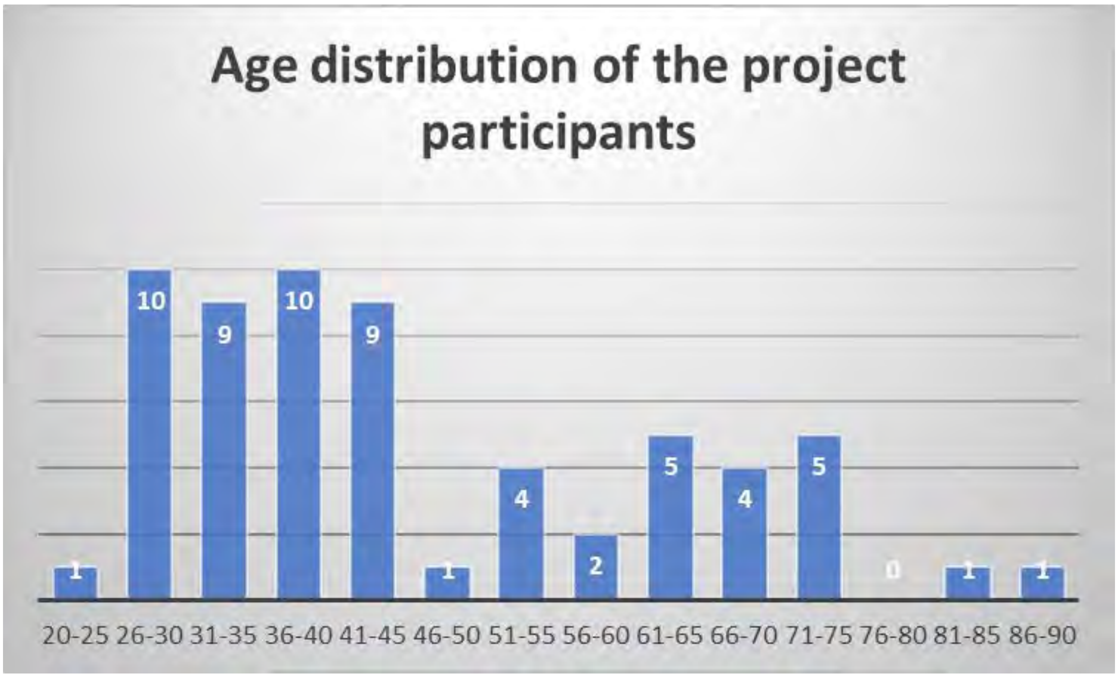

Fig. 1. Age distribution of the project participants

The distribution of the project performers by categories is shown in Fig.2.

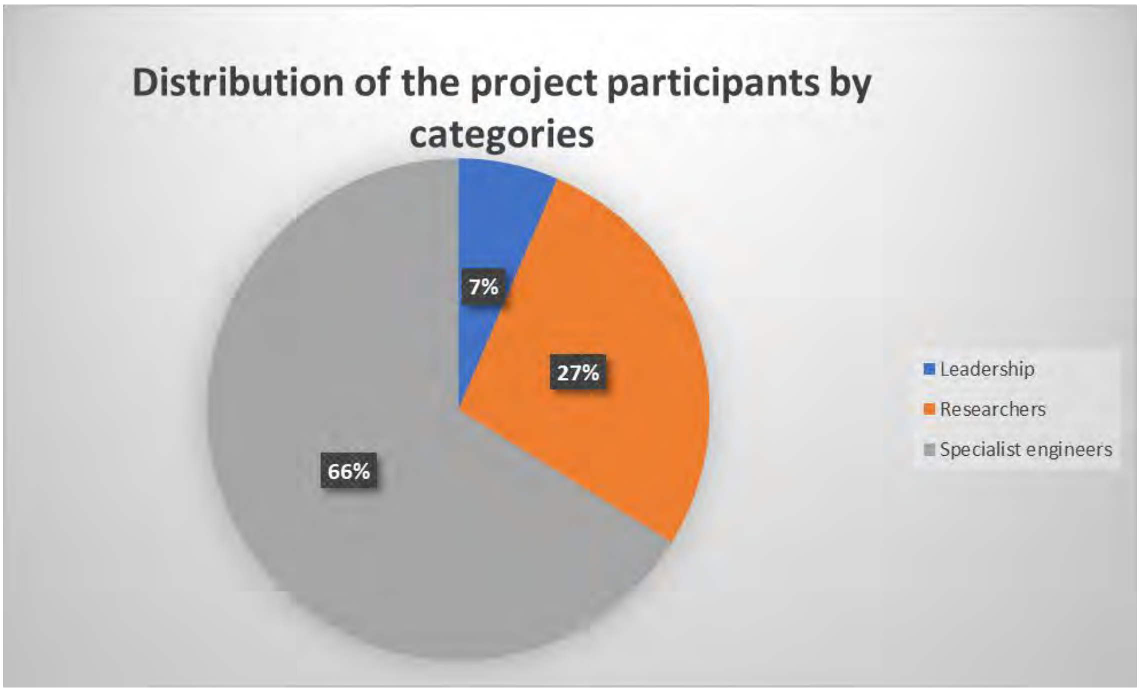

Fig.2. Distribution of the project participants by categories

The diagram of Fig.3 illustrates the distribution by the share of time that each participant will give to the work under the Project in relation to its full employment (FTE – Full Time Equivalent).

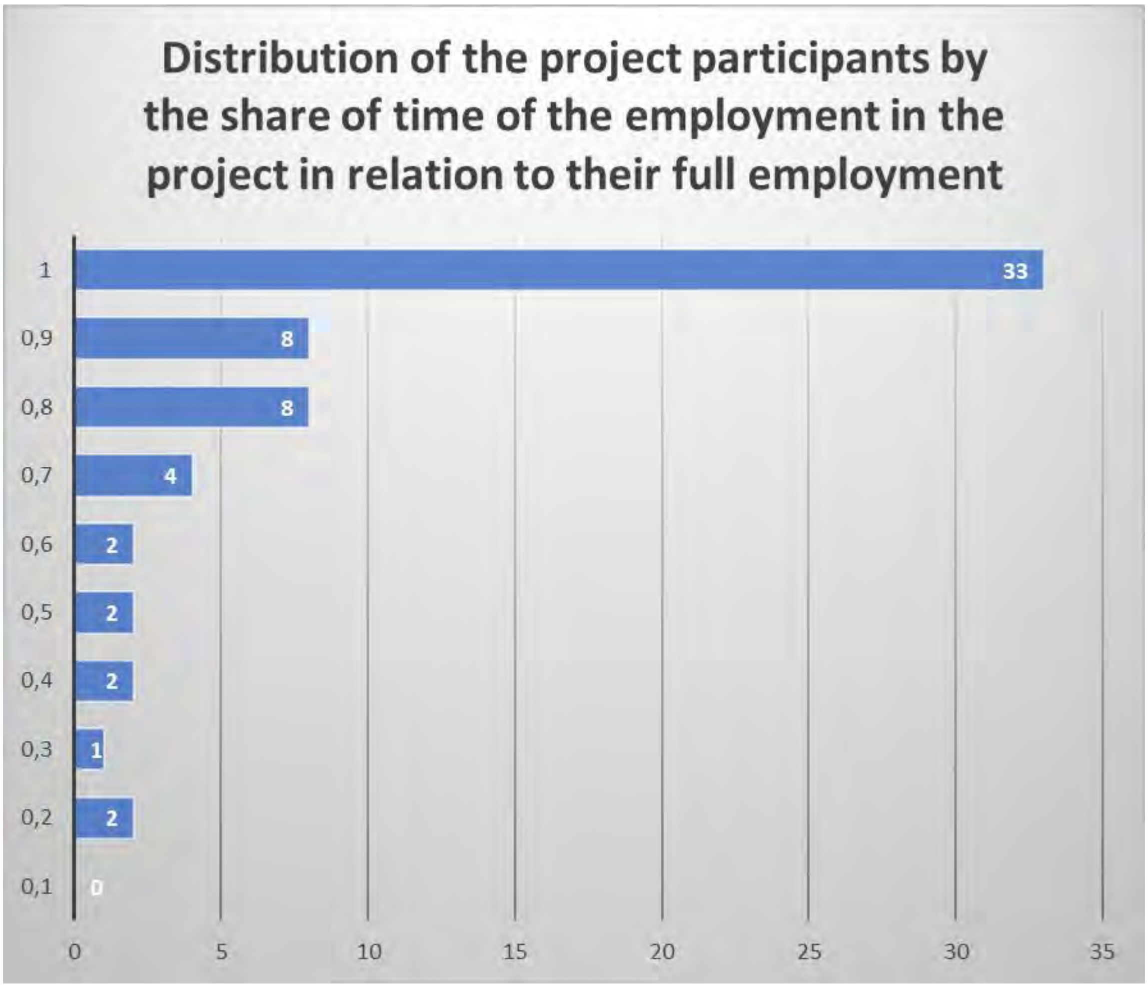

Fig.3. Distribution of the project participants by the share of time of the employment in the project in relation to their full employment

There are 4 doctors of science and 12 candidates of physical and mathematical sciences among the project participants. 4 defenses of theses of young employees are planned for the coming year. The average age of the project participants is 46; the age of engineers and programmers is 40.

Estimation of budgets

13700,3 kUSD from the JINR budget were spent on the MICC project in 2017-2018, and the budget for 2019 is planned in the amount of 8705,5 kUSD. The main funds were spent on the implementation of tasks and the expansion of the MICC computing powers planned by the project.

The increase in expenses in 2017 is related to the creation of the COVORUN supercomputer, which was commissioned in 2018 and developed as a joint project of the Bogoliubov Laboratory of Theoretical Physics and the Laboratory of Information Technologies, supported by the JINR Directorate. The project is aimed at a significant speed-up of complex theoretical and experimental studies in the field of nuclear physics and condensed matter physics underway at JINR, including for the NICA complex. The supercomputer CPU-component was created on the basis of the HPC-specialized engineering infrastructure on the contact liquid cooling technology implemented by the Russian company “RSC Technologies”. The company is the leader in the Russian market in the field of HPC solutions with liquid cooling, which are based on a number of its own unique developments that allow creating supercompact and energy-efficient HPC systems with high computing density. The supercomputer GPU-component was implemented on the basis of computing servers of the latest generation with graphics accelerators NVIDIA Volta. The total performance of the GOVORUN supercomputer at the time of its creation was 1 Pflops for single-precision calculations. The expenses on its creation comprised 2480 kUSD.

In 2018 additional funds from the grant of the JINR Directorate, in the amount of 650,1 kUSD, were allocated on the expansion of computing resources of one of the MICC components, i.e. the Tier1 center for the CMS experiment. At the expense of the JINR budget, in 2018 the equipment on the modernization and development of the JINR network infrastructure for a total value of 1285 kUSD was purchased. Due to these funds, the bandwidth of external telecommunication channels and the JINR backup local computing network was increased. An additional optical telecommunication channel MGTS-9 (Moscow) 2x100 Gbps was created, and the equipment for the transfer of the entire JINR backbone network to 100 Gbps was purchased.

In 2019 additional budget resources in the amount of 2500 kUSD presuppose the replenishment of the GOVORUN supercomputer with a ultrafast data storage system and the expansion of computing powers. It is planned to replenish the supercomputer with a full-size data storage system with a total capacity of 288 TB and expand the Skylake computing component on 20 computing nodes. Thus, the GOVORUN supercomputer performance will be increased by 200 Tflops for single-precision calculations. The purchase of a tape robot to create long-term storage for the BM@N experiment of the NICA project and in frames of the CMS experiment at the LHC, consisting of a tape robot with a capacity of 10PB and an intermediate disk array, in the amount of 1600 kUSD is planned from the budget for 2019. The extended budget for 2019 also presupposes the additional financing of works on the engineering infrastructure in the amount of 1046 kUSD.

The planned expenses on the MICC project are presented in the table, which shows the maximum figures on the financing of various project subsections. However, the adjustment of funding between subsections is allowed during the project implementation.

MICC project subsections	Description of works	2020	2021	2022	2023
		kUSD			
Engineering infrastructure – guaranteed power supply, climate control system	Stage-by-stage modernization of the climate control system and uninterruptible power supplies (UPS)	1000	1000	1000	1000

	Modernization of the 2^{nd} and 4^{th} floors	200	300	100	100
Network infrastructure	Equipment for the central core of the network infrastructure	150	150	150	175
	Modernization of the equipment of external telecommunication channels	100	100	100	100
	Local network infrastructure	400	400	400	550
Development of the MICC grid component Tier1 of the CMS experiment and the MICC integral component Tier2/CICC and computing resources of the NICA Off-line cluster in frames of the MICC	Computing resources	537	830,3	1 000	1 200
Data storage system (JINR data lake)	Disk servers	1400	2100	2200	2300
	Robot tape storage	1600	650	650	650
Extension of the HybriLIT heterogeneous platform including the GOVORUN supercomputer	Heterogeneous platform	470	470	470	470
JINR cloud	Cloud infrastructure	500	500	600	600
	Consumables, equipment and specialized licensed software	200	300	300	300
	Replacement of critical and obsolete equipment	350	350	350	400

Form № 26

Schedule proposal and resources required for the implementation of the Project Multifunctional Information and Computing Complex LIT JINR

(Project title)

Expenditures, resources, financing sources			Costs (k$) Resource requirements	Proposals of the Laboratory on the distribution of finances and resources			
				2020	2021	2022	2023
Expenditures		Main units of equipment, work towards its upgrade, adjustment etc.	28582,3	6747,0	6990,3	7160,0	7685,0
		Materials	640,0	160,0	160,0	160,0	160,0
		Computer communication	110,0	20,0	30,0	30,0	30,0
		Travel expenses	535,2	133,8	133,8	133,8	133,8
Required resources	Standard hour	Resources of – Laboratory design bureau; – JINR Experimental Workshop; – Laboratory experimental facilities division; – accelerator; – computer. Operating costs					
Financing sources	Budgetary resources	Budget expenditures including foreign-currency resources	29867,5	7060,8	7314,1	7483,8	8008,8
	External resources	Contributions by collaborators. Grants. Contributions by sponsors. Contracts. Other financial resources, etc.					

PROJECT LEADER

Form № 29

Estimated expenditures for the Project <u>Multifunctional Information and Computing Complex LIT JINR</u>
(Project title)

NN	**Expenditure items**	**Full cost**	**2020**	**2021**	**2022**	**2023**
	Direct expenses for the Project					
1.	Accelerator, reactor	h				
2.	Computers	h				
3.	Computer communication	k$	20,0	30,0	30,0	30,0
4.	Design bureau	standard hour				
5.	Materials	k$	160,0	160,0	160,0	160,0
6.	Equipment	k$	6747,0	6990,3	7160,0	7685,0
7.	Travel expenses, including: a) non-rouble zone countries b) rouble zone countries c) protocol-based	k$	133,8 98,3 10,0 25,5	133,8 98,3 10,0 25,5	133,8 98,3 10,0 25,5	133,8 98,3 10,0 25,5
	Total direct expenses:	29867,5	7060,8	7314,1	7483,8	8008,8

PROJECT LEADER

LABORATORY DIRECTOR

LABORATORY CHIEF ENGINEER-ECONOMIST

Concise SWOT[1] analysis

Strengths

The performers cope with the implementation of current tasks, the modernization and upgrade of computing components and data storage systems.

Multiyear successful experience in the WLCG project for processing of data from the LHC experiments.

The MICC components operate at the level of the best international standards in a 24x7 mode.

Modern hyper-convergent supercomputer built on liquid cooling and modern processors.

The network infrastructure is updated.

Established mechanism for monitoring the functioning of all MICC components.

Cooperation with MICC users.

Weaknesses

Weak control over users' actions.

Absence of a centralized user support system.

Procurement system.

Low rate of work on the modernization of the elements of the engineering and network infrastructure.

Unpredictable prices on the equipment from leading manufacturers of computing and related equipment in the region.

Opportunities

Understanding the need for investment and support of the developed IT-infrastructure by the JINR Directorate.

Resources of students from Dubna University and other institutes as a potential source of personnel for servicing the MICC components.

Threats

Rapid pace of moral obsolescence of the computer and network equipment.

Virus and hacker attacks outside and inside due to users' carelessness.

Depreciation and moral obsolescence of the engineering equipment, the modernization of which is delayed due to the excessive bureaucracy of the decision-making procedure.

[1]SWOT – Strengths, Weaknesses, Opportunities, Threats