MESHCHERYAKOV LABORATORY of INFORMATION TECHNOLOGIES

The activity of the Meshcheryakov Laboratory of Information Technologies (MLIT) in 2024 was focused on ensuring the reliable functioning and growth of the JINR network, information and computing infrastructure within the large research infrastructure project 06-6-1118-2014/2030 "Multifunctional Information and Computing Complex" (MICC), as well as on developing mathematical support and software for the research and production activities of the Institute and the JINR Member States (theme 06-6-1119-2014 "Methods, Algorithms and Software for Modeling Physical Systems, Mathematical Processing and Analysis of Experimental Data"). A distinctive feature of ongoing research directions is close cooperation with all the Laboratories of the Institute, institutes of the JINR Member States and other countries.

In 2024, the MLIT staff published over 200 scientific papers, four monographs and more than 100 articles within international collaborations, presented over 150 reports at international and Russian conferences.

MULTIFUNCTIONAL INFORMATION AND COMPUTING COMPLEX OF JINR

The development of scientific research at JINR defines the requirements for the computing infrastructure. The Multifunctional Information and Computing Complex represents a key element of this infrastructure and plays a decisive role in scientific research that entails advanced computing

2 April. Scientific section of the International Conference "Mathematics in the Constellation of Sciences"

power and data storage systems. The MICC is considered as a large research infrastructure project, being a set of hardware and software complexes, systems and subsystems, which embrace the Tier-1 and Tier-2 data processing grid centres, the cloud infrastructure, the HybriLIT platform and the Govorun supercomputer, the data storage, the data transfer system, the network infrastructure, the engineering infrastructure, and the monitoring system. To attain the main objectives of JINR flagship projects, it is needed to ensure the high performance, reliability and availability in 24 × 7 × 365 mode of all MICC components.

In 2024, work to modernize and enhance the performance of the hyperconverged Govorun supercomputer, distributed computing and data storage systems based on grid technologies and cloud computing was underway. The work was based on reliable engineering components and the state-of-the-art network infrastructure with a bandwidth of up to 4 × 100 Gb/s.

The Tier-1 grid site for the CMS experiment at the LHC continued to be a leader among seven similar sites worldwide. Tier-2/CICC provided data processing for the experiments at the LHC, NICA and other large-scale experiments, as well as support for users from the JINR Laboratories and Member States. The cloud environment of JINR and its Member States was mainly used for computing within the JINR neutrino programme.

JINR Network Infrastructure

The development of information technology and the MICC project is directly related to the further development of the JINR network infrastructure, without which the creation of distributed data processing and storage systems within the JINR research programme is inconceivable. The network infrastructure within the MICC project ensures external telecommunication channels, communication between MICC users through the JINR local area network, communication and data exchange through the MICC local area network. The network infrastructure is an intricate complex of multifunctional network equipment and specialized software, which is the foundation for the JINR information and computing infrastructure that has been created and is constantly developing. It consists of the following functional components: the external optical telecommunication data transmission channel; the backbone of the JINR local area network; the local area networks of the Institute's subdivisions; the MICC local area network.

In 2024, the functioning of the JINR telecommunication channels was ensured. First of all, this is the reliable operation of the Moscow backup channel with a bandwidth of 4 × 100 Gb/s. To operate the Tier-1 grid site, one must be a full member of the LHCOPN network to communicate with Tier-0 (CERN) and other Tier-1 sites. This connection is provided by the 100-Gb/s JINR–CERN direct channel and its 100-Gb/s backup channel passing through Moscow and Amsterdam. The JINR Tier-2 connectivity is ensured by the LHCONE external overlay network designed for Tier-2 grid sites. The National Research Computer Network of Russia (NIKS), created as a result of the integration of the federal university computer network RUNNet (Russian UNiversity Network) and the network of organizations of the Russian Academy of Sciences RASNet (Russian Academy of Science Network), provides communication with Russian scientific and educational organizations, as well as integration with individual National Research and Education Networks (NREN) and with the Internet.

The distribution of the incoming (exceeding 25 TB) and outgoing traffics by the subdivisions in 2024 is shown in Table 1.

Table 1

Subdivision	Incoming traffic, TB	Outgoing traffic, TB
MLIT	1240.0	227.1
HRC	502.25	90.35
VBLHEP	465.3	228.01
DLNP	292.28	130.73
FLNP	177.07	69.03
FLNR	176.82	36.16
Dubna State University	163.33	44.18
JINR Directorate	127.81	76.6
Public Access Servers	106.57	97.64
Remote Access Node	88.4	12.57
Medical Unit No. 9	63.81	11
UC	60.39	12.23
BLTP	38.66	24.67
SIMO	35.04	6.38
LRB	30.65	2.85
CPED	26.65	1.96

The overall incoming traffic of JINR, including the general-purpose servers, Tier-1, Tier-2, the Govorun supercomputer and cloud computing, amounted to 42.53 PB in 2024, while the overall outgoing traffic reached 20.62 PB. The traffic with the scientific and educational networks, accounting for 94.05% of the total, is overwhelming.

The local area network (LAN) is based on the JINR backbone network with a bandwidth of 2 × 100 Gb/s and the distributed multinode cluster network between the DLNP and VBLHEP sites (4 × 100 Gb/s) to ensure the reliable transfer of physical data received from the main nodes of the NICA Complex computing equipment for its processing and analysis on the MICC components.

Throughout the year, the e-mail service processed 4 million incoming messages and 850 thousand outgoing messages. The average message processing time was 3 s.

In 2024, over 1700 JINR user requests concerning the operation of the network and services were processed. More than 140 units of various equipment were configured/checked/installed. About 70 incidents of network security violations and about 130 cases of copyright infringement were identified and processed.

The JINR LAN comprises 13 496 network elements, 22 538 IP addresses in IPv4 format, 1422 IP addresses in IPv6 format, 5718 users (including 5586 JINR staff members), 4773 @jinr.ru email addresses, 1153 users of electronic libraries, 899 users of the remote access service and 147 users of the EDUROAM service. The registration procedures were supplemented with support for new user categories, namely, associated personnel and students.

MICC Engineering Infrastructure

In 2024, work on the replacement and enhancement of the MICC engineering infrastructure [1], designed to ensure the reliable, uninterrupted and fault-tolerant operation of the information and computing systems and the data storage resources, was in progress.

During the year, work to install new servers at the MICC was carried out: 12 × Asus, 2 × Huawei, 30 × SILA CP2-1627, 6 × SILA CP1-1626, and two network switches SILA CK3-630A-32Q. As part of the development of the disk storage of the MICC cloud component (for the JUNO experiment), 8 × ASUS RS720-E10-RS12 were installed.

In 2024, significant work to modernize the power supply system in the MICC modules was performed. The transition to uninterruptible three-phase power supply in the Tier-2 modules was carried out. Cabinets for clean power supply from Galaxy 7000 UPS were installed and put into operation. Due to the fact that standard equipment is not suitable for MICC tasks because of its size and power, special power distribution modules were elaborated, assembled, installed and connected. Special horizontal devices for power distribution in racks were also developed and assembled.

JINR Grid Environment (Tier-1 and Tier-2 Sites)

In 2024, the successful operation of the JINR grid sites continued. Both grid sites provided data processing and analysis within JINR's participation in the LHC projects at CERN, as well as tasks on modeling, processing and storing data from the BM@N, MPD and SPD experiments at the NICA Accelerator Complex.

The infrastructure and services of the Tier-1 (JINR-T1) and Tier-2 (JINR-LCG2) sites ensure computing, data storage, grid support service, data transfer, distributed computing management systems and information services (monitoring, information sites).

One of the major functions of the Tier-1 site is to receive and responsibly store unique experimental data from the CMS experiment transferred from the Tier-0 site to CERN. In addition, the site provides consistent and continuous data processing and reprocessing using new software or new calibration constants, as well as access to various datasets to Tier-1 and Tier-2 sites involved in processing CMS experiment data, etc. For the NICA Complex, simulation tasks for the MPD and SPD experiments are performed on the grid sites. The BM@N experiment has the ability to conduct full experimental data processing using the JINR grid infrastructure.

To organize computing in the grid environment, the Advanced Resource Connector (ARC), middleware for grid computing, and the Slurm workload manager are employed.

Power distribution modules (PDMs) developed at MLIT and installed at the MICC

In 2024, continuous support for lower-level services, which are important for the reliable functioning of both grid sites and the entire MICC as a whole, was ensured. It is noteworthy that without the uninterrupted operation of these services, the operation of the MICC and the JINR IT infrastructure is impossible. These include: DNS and IPDB, services for registering and resolving addresses and names of network elements; a time synchronization service for all JINR machines; Kerberos and LDAP, services for registering IT infrastructure users and for authenticating and authorizing users (in particular, the SSO service); AFS, a distributed file system, a storage of user home directories; CVMFS and GIT, systems for the distributed access and organization of software versions for collaborations and user groups; EOS, a system for storing and accessing large data volumes; a number of services for the JINR grid and international collaborations.

In 2024, a large amount of work on the transition of the Tier-1 and Tier-2 sites to the AlmaLinux 9 operating system (OS) due to the end of the life cycle of the CentOS 7 OS was completed.

One of the uppermost elements of the JINR grid infrastructure, as well as the entire MICC, is the data storage system. The dCache and EOS systems are utilized as the main data storage systems for disk storage, and Enstore and CTA are used for tape storage. In 2024, a number of works related to the technical re-equipment, installation and adjustment of unique hardware and software for the JINR MICC, as well as to the operation and development of the dCache–Enstore data storage system, were completed. Significant work to modify Enstore was done. This comprises converting code from Python 2 to Python 3, organizing the parallel execution of requests for mounting/unmounting tapes in tape recorders, which considerably enhanced the time of these operations, and increasing the throughput of processing a large flow of requests for data transfer.

EOS is viewed as a common distributed storage system for all MICC users with a total capacity of 23.3 PB. The EOS system is used by 32 experiments/user groups to store their data. The distribution by the volume of data stored in the system as of December 2024 (exceeding 50 TB) is shown in Fig. 1.

The EOS–CTA storage system of the TS3500 robot was configured to store data from experiments not

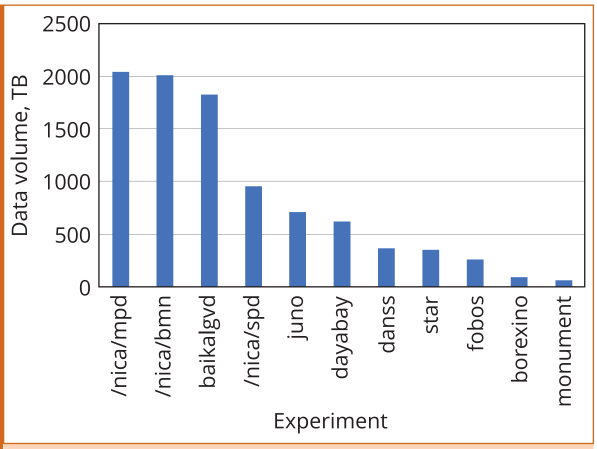

Fig. 1. EOS storage usage by experiments as of December 2024 (data volume exceeding 50 TB)

March. Scientific school for physics teachers. Introductory excursion to the Laboratory included in the WLCG (Worldwide LHC Computing Grid). The Baikal-GVD experiment was connected to the EOS–CTA system.

Currently, Tier-1 embraces 468 compute nodes (20 096 cores) with a performance of 32.4 kHS23. The launch of tasks for CMS experiment data processing is carried out by 16 nuclear pilots, and all Tier-1 computing resources are available to them. Data storage is provided by the 12.4-PB dCache system and the robotic tape storage with a capacity of 90 PB. The TS4500 robot runs the Enstore and dCache software. To work with tapes, a 2.65-PB disk array is used to cache data.

In 2024, the JINR Tier-1 site took first place in terms of the total normalized CPU time (HS23 hour) for data processed in the ranking of world Tier-1 sites that process data from the CMS experiment at the LHC (Table 2) according to the statistics available at https://accounting.egi.eu.

Table 2

Tier-1	Total CPU time (kHS23 h)	Total actual time (kHS23 h)
RU-JINR-T1	1 888 913.532	2 432 207.483
US-FNAL-CMS	1 665 321.019	3 979 538.749
UK-T1-RAL	1 132 299.978	1 341 419.038
DE-KIT	1 096 517.306	1 349 110.949
FR-CCIN2P3	807 778.212	859 384.319
ES-PIC	671 670.217	701 553.254
IT-INFN-CNAF	618 136.609	692 168.453

The Tier-2 site embraces 485 compute nodes (10 356 cores) with a total performance of 166.8 kHS23. Data storage is provided by the 5.62-PB dCache system. The TS3500 11-PB robotic tape library runs the CTA software and is used for the backup storage of data from local experiments.

The JINR Tier-2 output is the highest in the Russian grid segment (RDIG). In 2024, 2 892 483 tasks were completed on it, which amounted to 90.7% of the total CPU time of the WLCG Russian segment (Russian Data Intensive Grid, RDIG).

The software of the MICC grid sites and storage systems is regularly updated to the latest versions. These include: EOS, CVMFS, Rucio, ALICE VObox, XRootD, UMD, VOMS, the WLCG standard program stack, ARC-CE, the BDII top, the BDII site, OpenAFS, CentOS Scientific Linux, AlmaLinux, GCC: gcc (GCC), C++: g++ (GCC), GNU Fortran (GCC), dCache, Enstore, CTA.

For the simpler and more reliable updating of various MICC system software, test setups for most key components, including the Slurm batch processing system, dCache, EOS, CTA, were installed and configured. This allows for the necessary software updates to be carried out quickly and without significant problems in the main setups.

The key element in the organization of distributed heterogeneous computing infrastructures is connecting middleware (platform), which enables the joint operation of various information and computing systems. Today, the major type of such software at JINR is the DIRAC Interware, a tool for integrating heterogeneous computing and data storage resources into a single platform. Resource integration is based on the use of standard data access protocols (XRootD, GridFTP, etc.) and pilot tasks. Thanks to this, the user is provided with a single environment for launching tasks, managing data, building processes, and controlling their execution.

In 2024, DIRAC (Distributed Infrastructure with Remote Agent Control) continued to be used to solve the tasks of collaborations of all three experiments at the NICA Accelerator Complex. Together with DIRAC platform developers, a method for using the DIRAC pilot pre-installed in CVMFS was proposed and implemented. The implementation of this method entailed making changes to the source code of the DIRAC platform and creating a pilot configuration program at JINR using the version pre-installed in CVMFS. This significantly reduced the load on the local storages of the computing resources. A method for assessing the integrity of files under the management of the DIRAC platform was worked out. The method is based on adding to the files a field of metadata related to the integrity check status. The application of this approach made it possible to check all files of the BM@N and MPD experiments. The peak data transfer rate amounted to 12 GB/s. It is noteworthy that access to the tape robot based on CTA was integrated into the DIRAC infrastructure. This enabled to organize the backup of the data obtained during the 8th session of the BM@N experiment on tapes. In addition, the MPD experiment began backing up some of its most critical datasets.

Heterogeneous Infrastructure

In 2024, the software and information environment of the HybriLIT platform, which is a component of the JINR MICC for massively parallel and resource-intensive computing, was actively developing [2, 3]. The hierarchical data processing and storage system of the Govorun supercomputer was enhanced. It comprises two high-performance servers, each containing two Intel Xeon Platinum 8458P processors with 512 GB of RAM and 32 NVMe Ruler SSDs with a capacity of 30.72 TB, which enabled to expand the volume of the warm data storage subsystem by 2 PB. As a result, the total capacity of the hierarchical storage of the Govorun supercomputer reached 10.6 PB.

Work on developing the polygon with the geographically distributed parallel file system Lustre as a system for simultaneous data processing for computing clusters located at the MLIT and VBLHEP sites was in progress. The polygon embraces the physical nodes of the HybriLIT heterogeneous platform, the VBLHEP NCX cluster, and the Govorun supercomputer with a total disk space of 2.1 PB. Its diagram is demonstrated in Fig. 2.

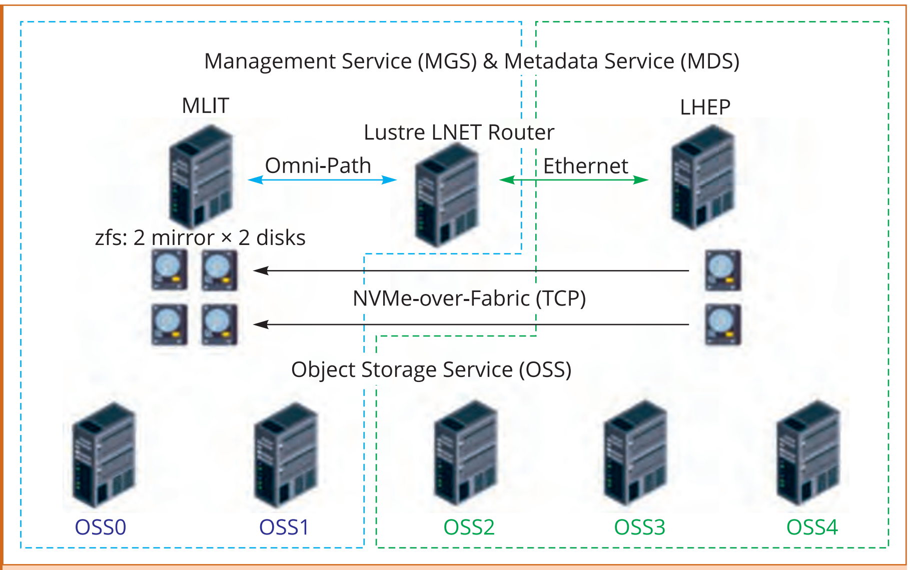

Fig. 2. Geographically distributed parallel file system Lustre located at the MLIT and VBLHEP sites

This prototype consists of two segments that utilize different types of network connections (Omni-Path, Ethernet) and are combined into the common Lustre network using the LNET Router service. The MGS/MDS management services were launched on two servers in High Availability mode. The MGS/MDS data storage section was assembled on a RAID array (mirror), containing a pair of local disks and a pair of network disks connected via the NVMe-over-Fabric protocol (TCP). The OSS data storage servers are distributed geographically, two of them are located in the network segment of the Govorun supercomputer (OSS0 and OSS1 in Fig. 2), the other three are connected to the NCX computing cluster. A number of load tests and a study of the performance of the created polygon, including the usage of the Interleaved Or Random (IOR) Benchmark tool together with MPI technology, were performed [4]. The obtained results showed an increase in the productivity of solving tasks by 30% using the created polygon compared to using local Lustre file systems located at MLIT and VBLHEP with subsequent data exchange between them.

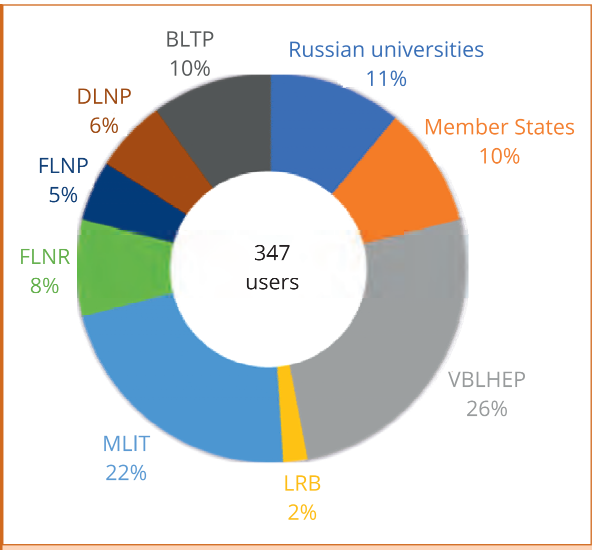

Fig. 3. Distribution of the Govorun supercomputer users by the Laboratories and the Member States

As part of the modernization of the ML/DL/HPC ecosystem for artificial intelligence methods and high-performance computing, an environment for the development of neural network algorithms was prepared to support the activity on the joint project of MLIT and LRB, BioHLIT. It was deployed on a server with eight NVidia A100 graphics accelerators and the AlmaLinux 9.4 operating system. The developed environment embraces JupyterHub 4.0.11 (for the joint development of programs in the Python programming language), the Mercury 2.4.3 framework (for publishing web applications on top of Jupyter Notebook and the Python programming language), the Writer AI 0.7.5 toolkit (for the rapid development of web applications in the Python programming language) and is available at https://mostlit.jinr.ru.

The total number of the Govorun supercomputer users is currently 347. The distribution of users by the Laboratories is shown in Fig. 3.

Throughout 2024, all groups of Govorun supercomputer users computed 6.2 million tasks on the CPU component, which corresponds to 32 million core hours. Figure 4 demonstrates the distribution among the most resource-intensive projects.

The GPU component of the HybriLIT platform is employed to solve high-performance tasks using CUDA technology, for example, for computations in lattice quantum chromodynamics, the development of quantum algorithms and computing with quantum simulators, the elaboration of neural network algorithms and data annotation, the training of neural network models and the deployment of GPU-oriented web services using a neural network ap-

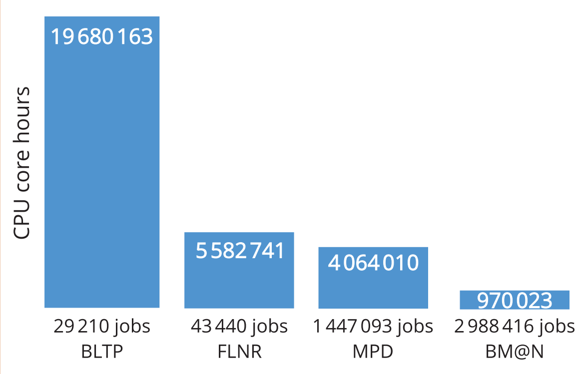

Fig. 4. Statistics on the usage of the Govorun supercomputer resources by the most resource-intensive projects

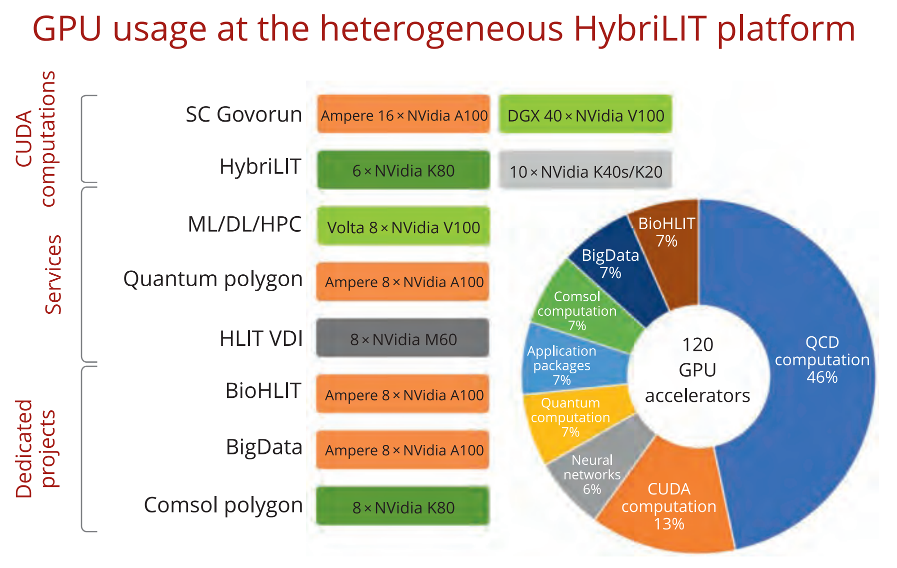

Fig. 5. Structure of employing the HybriLIT platform resources for tasks solved on the platform

proach [3, 5]. The structure of utilizing the graphics accelerators of the HybriLIT platform is displayed in Fig. 5.

In 2024, HybriLIT platform users published 54 papers, including nine in Q1 journals, five in Q2 journals, and eight in Q3 journals.

Cloud Infrastructure

In 2024, the operating systems on cloud infrastructure servers were updated from CentOS 7.9.2009 to AlmaLinux 9.5, and the cloud software was updated from 5.12.0.4 Community Edition (CE) to the stable version OpenNebula 6.10.0 CE. The OS was updated on some cloud virtual machines (VMs) of users and services.

Cloud storages were cleared of obsolete data and virtual machine images, and the transfer of VMs from the decommissioned subnet to a private routed subnet allocated by the network service for the cloud infrastructure was completed.

Within the enlargement of the range of cloud services, a new service for website traffic analysis on top of the Matomo platform, webanalytics.jinr.ru, was deployed. The service was created as a replacement for Google Analytics, which was previously used on the JINR website. It ensures the confidential storage of all collected data in the Institute's local storages and does not require the creation of third-party accounts. The new service can be used to collect statistics on visits to any JINR website. For example, the JINR GitLab service was already connected to it. In addition, it was integrated with the JINR single sign-on system, which provides centralized access to the system web interface.

In 2024, 12 125 668 CPU hours were consumed on the resources of the Neutrino Computing Platform (NCP), which is a segment of the JINR cloud infrastructure. Figure 6 illustrates information on the major consumers.

In reference to the distributed information and computing environment (DICE) based on the re-

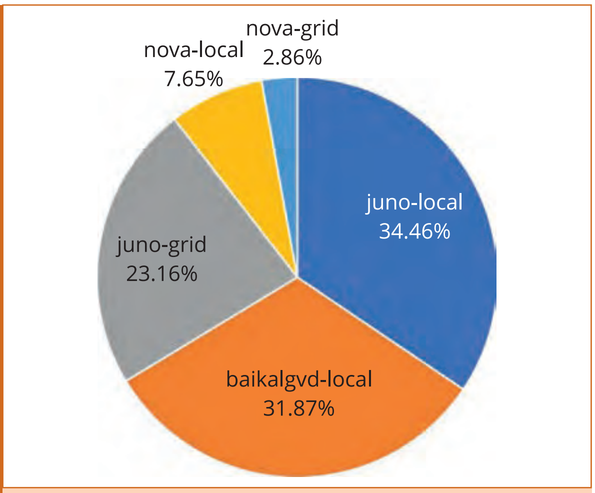

Fig. 6. Number of CPU hours consumed in 2024 on the NCP resources: juno-local, baikalgvd-local, nova-local — by JINR user tasks launched within the JUNO, Baikal-GVD, NOvA experiments by means of the local batch system (i.e., outside the grid); juno-grid, nova-grid — by tasks launched within the grid infrastructure

sources of the JINR Member States' organizations, the main work was related to maintaining the operability and availability of these resources. In addition, MLIT provided technical support to the resources of the Institute of Nuclear Physics (Kazakhstan), Khetagurov North Ossetian State University, and the Institute of Nuclear Physics (Uzbekistan).

Monitoring System

As part of the transition to the AlmaLinux 9 operating system, data acquisition scripts for the monitoring system were rewritten in Python 3. To implement the monitoring of the power supply system, specialized software was developed for collecting data, namely, input voltage, current, consumed power for the IR-23 controller of room 110, as well as for collecting data from the main switchboard of room 110a; an information panel visualizing the main parameters of the power supply systems was created for MICC operators. A similar approach was implemented to monitor the operation of the wet cooling tower (Fig. 7).

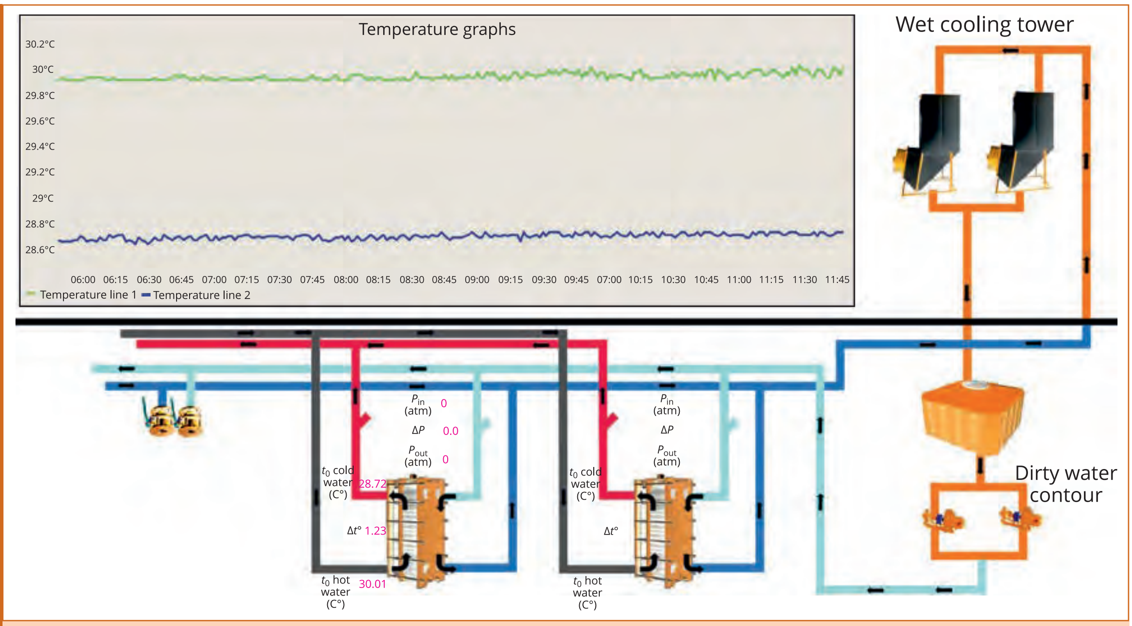

Fig. 7. Information screen of the wet cooling tower monitoring system of the MICC

MULTIPURPOSE HARDWARE AND SOFTWARE PLATFORM FOR BIG DATA ANALYTICS

In 2024, work to create a computing component of the Big Data intelligent analysis platform was carried out on the basis of a research infrastructure that contains computing resources on top of hardware accelerators (GPUs) of several models (NVidia A100, V100, T4). Within the created environment, computing polygons were deployed on Apache Spark and Dask using the RAPIDS software ecosystem; their operability for practical computing was confirmed. In particular, a study on the performance and scalability of the RDataFrame declarative analysis tool of the ROOT package in the distributed Apache Spark environment was conducted.

Within the research infrastructure, software environments and shells for working with open large language models (LLMs) were configured on virtual servers with GPUs. It is planned to employ LLMs for analyzing and abstracting scientific articles, extracting key ideas, forming thematic digests and bulletins, as well as developing recommendation systems, API interfaces and creating various AI assistants of the DES (navigator for JINR publications and scientific results; assistant for regulatory and financial documents, contracts; formulation of procurement requests; support for DES users, etc.).

JINR DIGITAL ECOSYSTEM

In 2024, within the development of the JINR Digital EcoSystem, several institute-wide digital services were put into operation: a document development management service, a calendar for collaborative work, etc. A corporate data bus prototype to integrate digital services and manage information flows was created. Some of the existing digital services were actively developed, and they were integrated with each other and with the basic DES services. The services put into operation are available via the DES interface https://digital.jinr.ru.

The transition to deeply redesigned PIN-2 was completed, it implements new functionality, additional information about the employee (for example, identifiers in citation systems), and data from the previous PIN version was transferred. For user authentication when accessing from foreign networks, the two-factor authorization functionality implemented in the JINR single sign-on system (JINR SSO) was connected.

The JINR Publications Repository was put into pilot operation [6]. The automatic collection of information on publications is performed from three main sources, namely, ELibrary, INSPIRE HEP, and Scopus. Publications are linked to the profiles of users, i.e., JINR staff members. Information on them is available both via the repository web interface and in PIN-2. Currently, the repository contains data on 1890 publications for 2024 and on 9074 publications since 2020.

To organize joint work on documents, the SciDocsCloud software platform, designed to replace the outdated DocDB system, was elaborated.

The functionality for reserving halls and rooms for meetings, conferences and other events was configured in the Indico conference management system.

Work on the current development and support of the following services was in progress: the Dubna EDMS, ISSC, DES Shell, Document Database, Advance Reports, the EDMS “JINR Staff at CERN”, ADB2, NICA EVM, Data Exchange Gateway. A hacker attack detection system was developed and implemented in test operation for such services as the Dubna EDMS, PIN, PIN-2, Advance Reports, CERN DB, ISSC, DES Shell.

27–28 May. The Workshop on Mathematical Problems of Quantum Information Technologies (MPQIT-2024)

One of the main activities of MLIT is to provide mathematical, algorithmic and software support for experimental and theoretical research underway at JINR. In 2024, within Theme 1119, a number of works and investigations aimed at the elaboration and enhancement of mathematical methods and software for modeling physical processes and experimental facilities, processing and analyzing data from experiments in elementary particle physics, nuclear physics, neutrino physics, radiobiology, etc. were performed. Highly needed and appreciated contributions were brought to the solution of specific tasks within the BM@N, MPD and SPD projects at the NICA Complex, the CMS and ATLAS projects at the LHC, and the projects of the JINR neutrino programme (Baikal-GVD and JUNO).

A summary of selected results is presented below.

Development and Enhancement of Physical Event Reconstruction Algorithms and Particle Identification Methods

Based on the analysis of data from the BM@N Run 8 (2022–2023) session, software for a detailed description of the geometry of coordinate detectors was developed. Algorithms for modeling realistic responses of track detectors and reconstructing spatial coordinates on microstrip planes were enhanced [7].

To eliminate false measurements in the reconstruction of charged particle tracks in the SPD experiment, a transformer-based neural network architecture was proposed (Fig. 8). An effective use of an attention layer for simulated data was demonstrated with the help of a voxelization procedure [8]. To disentangle overlapping events in one time slice, an approach based on a deep Siamese neural network [9] with a loss function in the form of triplets was developed.

Based on the gradient boosting decision tree (GBDT) algorithm, a method for identifying charged particles produced by the interaction of bismuth nuclei at $\sqrt{s_{NN}}$ = 9.2 GeV was elaborated in the MPD experiment [10]. The use of the XGBoost library showed an advantage over the traditional n-sigma approach, especially for high-momentum particles. The method was adopted by the MPD Collaboration for integration into MPDRoot to solve particle identification tasks (Fig. 9).

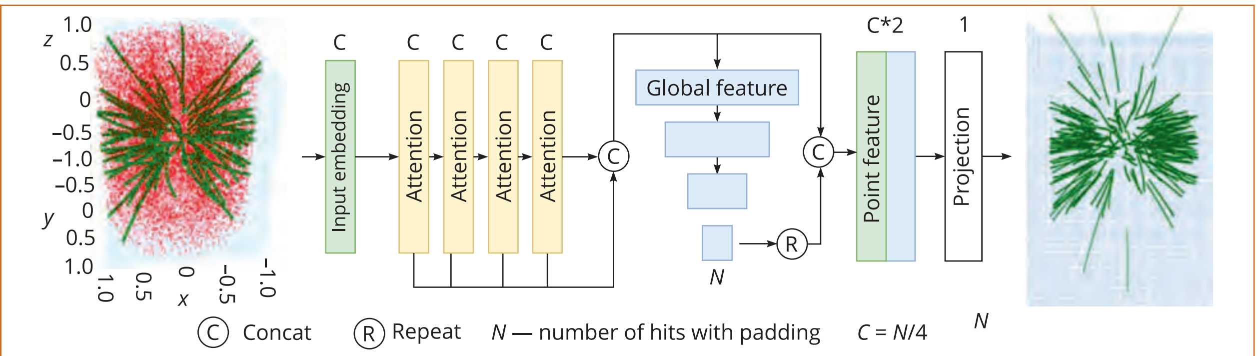

Fig. 8. Transformer-based neural network architecture used for particle track reconstruction

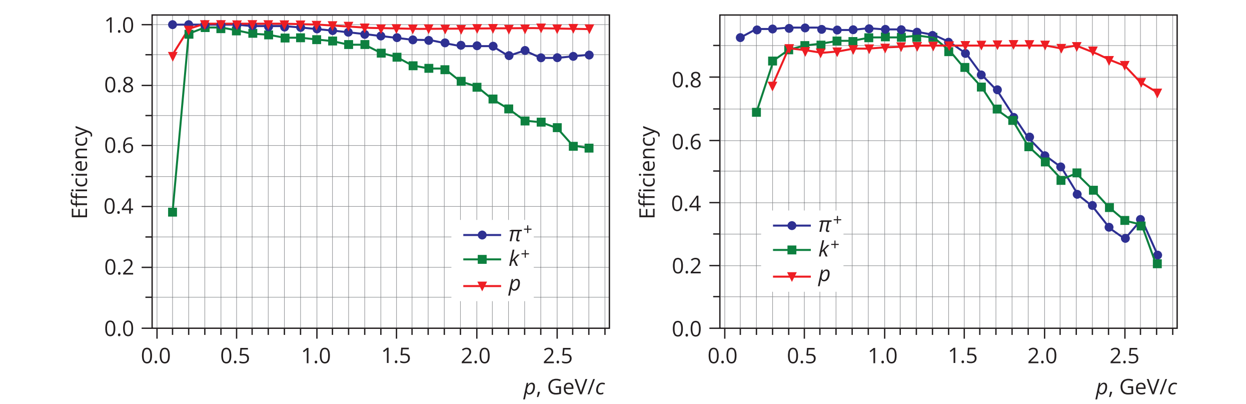

Fig. 9. Effectiveness of charged particle identification in the MPD experiment. Using the gradient boosting method (XGBoost) (left), using the n-sigma method (right)

Development and Enhancement of Methods for Physical Process Modeling and Data Analysis

In 2024, the development of interaction generators for modeling nuclear collision processes at NICA and LHC energies was underway. Analyzing NA61/SHINE data [11] on the production of $\pi^{\pm}$, $K^{\pm}$, protons and antiprotons in ^{40}Ar + ^{45}Sc interactions at P_{lab} = 13–150A GeV/c, it was found that the Geant4 FTF model significantly underestimated the meson yields. A new algorithm for taking into account diffraction dissociation in FTF was proposed, which enabled to increase meson multiplicity and successfully describe the data. The proposed approach gives decent results when applied to BM@N data on the yields of π^{+} mesons in argon–nucleus interactions at an energy of 3.2 GeV.

Using Monte Carlo data, the predictions of two dark matter models, namely, two-doublet Higgs expansion with an additional scalar and axial Higgs singlet, were analyzed. A full simulation of the response of CMS (Geant4) detectors at $\sqrt{s}$ = 13.6 TeV was performed for events with a Z boson and a pair of b quarks associated with a large fraction of the lost transverse energy [12]. The main background processes and optimal kinematic constraints for the maximum signal-to-background ratio were defined.

Within the development of OLVE-HERO experiment data analysis systems, a Monte Carlo assessment of the annual statistics of Earth orbit data for protons (1–45 GeV) was carried out using a spherical detector with a radius of 1.25 m [13]. It was discovered that the addition of 1–5% boron to the scintillator did not affect the registration of energy losses from protons (1, 10, 100, 1000 GeV). For a 5% additive, the energy resolution was estimated in the range of 1–100 TeV.

Within the interlaboratory working group (a joint project of MLIT, LRB, VBLHEP, FLNP, and DRS), studies to assess the radiation situation in the temporary control room of the NICA Accelerator Complex were conducted, namely, a detailed 3D model of the accelerator complex was built in the Geant4 format, hardware and software were configured to perform computing on the Govorun supercomputer, and the estimates of the effective dose rate and neutron spectra were obtained for the model source [14]. To ensure high-energy neutron dosimetry, two methods for neutron spectrum reconstruction based on readings from the Bonner multisphere spectrometer were developed: i) based on spectrum expansion in Legendre polynomials and the numerical solution of the Fredholm integral equation of the first kind using Tikhonov regularization, taking into account the detector "weight"; ii) based on the random forest machine learning algorithm.

Development of Data Processing Systems, Creation and Enhancement of Information and Computing Systems to Support JINR Research Projects

A prototype of a distributed data processing and analysis system integrating the resources of JINR and PNPI NRC "Kurchatov Institute" was deployed for the SPD experiment. In 2024, Monte Carlo simulation tasks were processed on the prepared platform: 200 million events with a total volume of 100 TB were generated [15]. Within the creation of a middleware complex for the real-time event selection system (SPD Online Filter), most of the functional requirements for the software complex were implemented, and work to formalize and implement nonfunctional requirements was performed. Together with colleagues from DLNP, a testbed for developing and debugging the components of the SPD facility data acquisition system was created and put into operation (Fig. 10). The testbed provides an opportunity to develop and conduct long-term tests of hardware

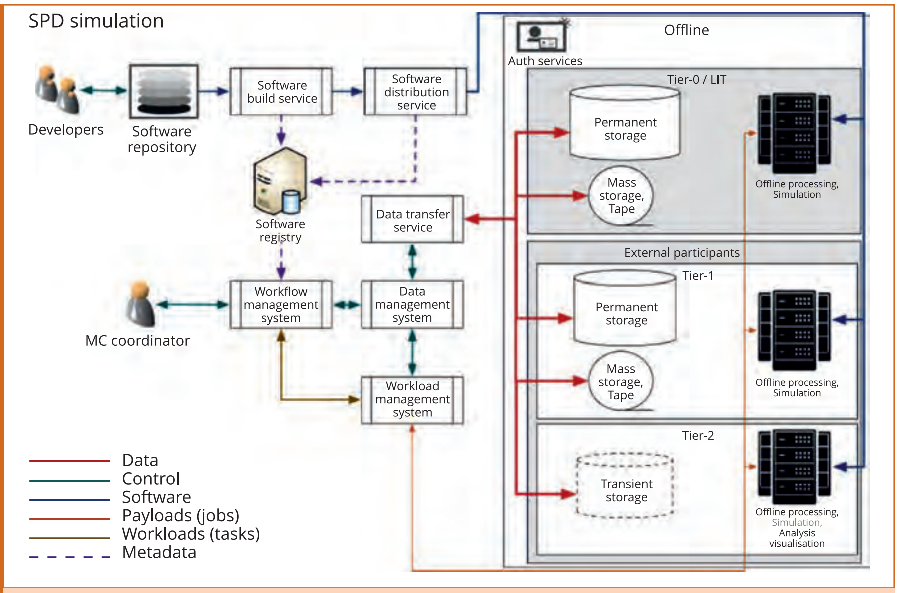

Fig. 10. Scheme of data processing and analysis within the SPD experiment

and software systems and comprises a "cold zone" for placing hardware.

Within the participation of MLIT specialists in the ATLAS experiment at CERN, the resource manager, a component of the TDAQ core, was enhanced. All components of the ATLAS Event Picking service [16] were modernized, and a service user account was created. The experience and knowledge gained in developing databases (DBs) for the ATLAS experiment were applied to the experiments at the NICA Accelerator Complex. Work to create, modernize and support information systems and databases for the BM@N experiment was carried out [17], namely, the configuration (put into production operation) and geometric (ready for test operation) information systems were improved. Within the MPD experiment, work on creating and implementing the Conditional DB and other DBs of the experiment was underway.

Development of Computational Physics Methods

Three monographs were published on the basis of the results of research in applied mathematics and computational physics. The monographs [18] and [19] are devoted to methods and programs for solving systems of equations of various types using the finite difference method, the high-order finite element method (FEM), and iterative schemes on top of the continuous analogue of Newton's method. In [20], the basic concepts, general techniques, and methods for working with matrices, including the calculation of determinants of various orders, are considered.

As part of the cooperation programme with the University of Cape Town (RSA), a study of the spherically symmetric time-periodic standing waves of the φ^4 model in a ball of finite radius was carried out. They are considered as an approximation of weakly radiating oscillons. For the 3D case, coexisting types of standing waves were found, and the dependence of the energy and stability intervals of the waves on the radius and frequency was investigated. Calculations were performed on the HybriLIT platform and the Govorun supercomputer. The frequency range where the variational method provides an accurate description of the (1 + 1)-dimensional oscillon was defined [21].

Molecular dynamics methods were employed to simulate the interaction of beta-amyloid peptide 25–35 with DMPC phospholipid membranes in the presence of a large number of K^+ and Cl^- ions in the solution. The results provide information on possible processes occurring during membrane destruction in the presence of peptides [22].

To describe the collective quadrupole model of the atomic nucleus and compute the rotational-vibrational spectra and probabilities of quadrupole transitions, the GCMFEM program to solve boundary value problems for a system of two-dimensional elliptic differential equations with mixed derivatives was elaborated [23]. For the isotopes ^{190}Os and ^{154}Gd, the results are consistent with experimental data.

Within the microscopic model of the optical potential, an analysis of proton–nucleus scattering data at energies of 200–1000 MeV was carried out. The estimates of the influence of the nuclear medium on the scattering amplitude depending on the energy of the incident proton and the atomic mass of the target nucleus were made [24].

A method for the refined calculation of the spread of energy losses in crystal lattices under irradiation with heavy ions with relativistic energies was proposed, it has an advantage in accuracy compared to existing approaches in the case of irradiation with highly charged ions [25].

The simulation of domains of reversal in the SFS model of the φ_0 junction with weak dissipation depending on the parameters that regulate the action of the external current was conducted. The parallel implementation enabled to reduce the computation time by up to 30 times [26].

A toolkit for modeling the physical properties of a superconducting quantum interference device (SQUID) with two Josephson junctions was developed [27]. Algorithms for calculating the SQUID current-voltage characteristics under the influence of an external magnetic field were implemented using the Numba library.

A scheme for the parallel numerical integration of the three-dimensional nonstationary Schrödinger equation in the representation of a discrete variable, providing a significant reduction in the computation time for such problems with a large number of grid nodes, was proposed and implemented [28].

A combined FEM-based procedure [29] was proposed, it allows reducing computational costs without loss in accuracy when solving 3D magnetostatics problems with complex geometry using a magnetic vector potential.

The FITTER_WEB web application [30], developed and deployed on the cloud infrastructure, was adapted to investigate the structure of phospholipid vesicular systems of various types using small-angle neutron and X-ray scattering data on the basis of the separated form factor model.

A method for extrapolating perturbative expansions in powers of asymptotically small parameters to the region of finite or infinite values of variables was proposed [31]. The extrapolation is performed using self-similar factor approximants. In a number of cases, the method enables to accurately reconstruct the desired functions from their weak-coupling asymptotic expansions.

A quantum-chemical cluster method to quantitatively substantiate effective spin models for various crystal structures of magnetic transition metal oxides was developed [32]. The method is applicable to wide families of new magnetic materials with complex chemical compositions.

Using the example of searching for the ground state in the Ising model with an external magnetic field, the quantum approximation optimization algo-

Yerevan, 21–25 October. 11th International Conference “Mathematical Modeling and Computational Physics” (MMCP 2024) dedicated to the 80th anniversary of the birth of Academician A. Sissakian.
Photo: *https://indico.jinr.ru/event/4467/page/2192-photos*

rithm was tested on the quantum polygon deployed on the HybriLIT platform. Comparative computations were performed on a CPU and a GPU using the CUDA and cuStateVec packages, they confirmed the efficiency of the cuStateVec package [5, 33].

Research in the field of using neural network models to classify images in conditions of a small training sample continued [34]. Software and hardware solutions to organize automated control and accounting in greenhouse complexes are developed. Methods and tools for organizing mobile object tracking complexes are worked out.

A web service based on machine learning (https://mostlit.jinr.ru) was developed, it allows for an automatic analysis of DNA double-strand breaks (DSBs) in mammalian and human cell nuclei relying on the counting of radiation-induced foci (RIF), which represent repair protein clusters at the DSB site. By processing a series of fluorescent images of cell nuclei, the service provides analytical information on key RIF parameters, such as the average number of RIF per cell nucleus, RIF area, and intensity.

EDUCATIONAL ACTIVITY

Within the educational activity, the JINR Spring and Autumn Schools of Information Technologies (IT Schools) were held in 2024. Ninety students from different Russian universities participated in the IT Schools. Since 2022, 18 people have become JINR employees according to the results of a series of IT Schools. More than 40 theses of postgraduates, Master’s and Bachelor’s students were prepared and defended in 2024 under the supervision and scientific advice of the MLIT staff. Over 20 non-graduate students work on JINR projects and come for practice. In 2024, the education and testing polygon of the HybriLIT platform was actively used for conducting semester-long training courses and within schools and workshops. Semester-long training courses on the IT disciplines “Architecture and Technologies of High-Performance Systems”, “Parallel Distributed Computing”, “Languages and Technologies of Data Analysis”, “High-Performance Computing Technologies”, “Mathematical Computing Software”, held at Dubna State University and Tver State University, were attended by 457 students. In addition, four Bachelor’s and two Master’s theses were prepared on the basis of the HybriLIT platform. The educational Master’s programme “Data Processing Methods and Technologies in Heterogeneous Computing Environments” was developed and licensed, the curriculum and work programmes of disciplines in direction 01.04.02 “Applied Mathematics and Computer Science” for the MSU Branch in Dubna were elaborated. To develop cooperation, seminars were held for students of such educational institutions as Lomonosov Moscow State University, Plekhanov Russian University of Economics (PRUE), and the PRUE Economic Lyceum.


15–16 April. JINR Spring School of Information Technologies

7–11 October. JINR Autumn School of Information Technologies

REFERENCES

1. *Vorontsov A. S., Evlanov A. V., Dolbilov A. G., Gavrish A. P.* JINR Multifunctional Information and Computing Complex. Engineering Infrastructure // Proc. of the All-Russian Conf. with Intern. Participation ITTMM 2024. 2024. P. 281–287 (in Russian).
2. *Anikina A., Belyakov D., Bezhanyan T. et al.* Capabilities of the Software and Information Environment of the HybriLIT Heterogeneous Computing Platform for JINR Tasks // Distributed Computer and Communication Networks. 2024. P. 244–249; https://2024.dccn.ru/papers/2628.
3. *Lyubimova M. A., Anikina A. I., Belyakov D. V. et al.* Software and Information Environment of the HybriLIT Heterogeneous Computing Platform // Proc. of the All-Russian Conf. with Intern. Participation ITTMM 2024. 2024. P. 315–321 (in Russian).
4. *Kokorev A. A., Belyakov D. V., Podgainy D. V., Moshkin A. A., Pelevanyuk I. S.* Distributed Parallel File System Lustre for Processing and Analyzing Data from High Energy Physics Experiments // Intern. Sci. Conf. "Russian Supercomputing Days 2024", Moscow, Russia, 23–24 Sept. 2024; https://russianscdays.org/files/2024/pdf/HPCCenters/3_AKokorev.pdf (in Russian).
5. *Belyakov D. V., Bogolubskaya A. A., Zuev M. I., Palii Yu. G., Podgainy D. V., Streltsova O. I., Yanovich D. A.* Polygon for Quantum Computing on the HybriLIT Heterogeneous Platform // Proc. of the All-Russian Conf. with Intern. Participation ITTMM 2024. 2024. P. 303–309 (in Russian).
6. *Filozova I., Shestakova G., Kondratyev A., Bondyakov A., Zaikina T., Nekrasova I.* DSpace Software Plat-

form for Digital Repository of JINR Publications // Phys. Part. Nucl. Lett. 2024. V. 21, No. 4. P. 797–799.
7. *Baranov D.* Detailed Simulation of the Response of Inner Tracker Detectors for the First Physics Run in the BM@N Experiment // Phys. Part. Nucl. 2024. V. 55, No. 4. P. 1055–1060.
8. *Nikolskaya A., Goncharov P., Ososkov G., Rusov D., Starikov D.* Point Cloud Transformer for Elementary Particle Signals Segmentation // Ibid. No. 3. P. 458–460.
9. *Borisov M., Goncharov P., Ososkov G., Rusov D.* Unraveling Time-Slices of Events in the SPD Experiment // Ibid. No. 3. P. 453–455.
10. *Papoyan V., Aparin A., Ayriyan A., Grigorian H., Korobitsin A.* Gradient Boosted Decision Tree for Particle Identification Problem at MPD // Phys. Part. Nucl. Lett. (submitted).
11. *Galoyan A., Uzhinsky V.* Tuning the Geant4 FTF Model Using Experimental Data of the NA61/SHINE Collaboration // Phys. Part. Nucl. 2024. V. 55, No. 4. P. 962–967.
12. *Hayrapetyan A. A., Savina M. V., Tumasyan A. R., Shmatov S. V.* Search for Dark Matter Produced in Association with Standard Model Higgs Boson in *pp* Collisions at 13 TeV in the CMS (LHC) Experiment // Ibid. No. 1. P. 132–136.
13. *Karatash Kh., Satyshev I., Sholtan E.* Monte Carlo Simulation of the OLVE-HERO Orbital Experiment // Phys. Part. Nucl. 2025. V. 56, No. 2. P. 159–166 (in Russian).
14. *Chizhov K., Beskrovnaya L., Chizhov A.* Neutron Spectra Unfolding from Bonner Spectrometer Readings by the Regularization Method Using the Legendre Polynomials // Phys. Part. Nucl. 2024. V. 55, No. 3. P. 532–534.
15. *Abazov V. et al. (SPD Collab.).* Technical Design Report of the Spin Physics Detector at NICA // Natural Sci. Rev. 2024. No. 1. 1; arXiv:2404.08317.
16. *Gallas E. J., Alexandrov E., Alexandrov I. et al.* Deployment and Operation of the ATLAS EventIndex for LHC Run 3 // Eur. Phys. J. Web Conf. 2024. V. 295. 01018.
17. *Alexandrov E., Alexandrov I., Chebotov A. et al.* Development of the Online Data Processing System for the BM@N Experiment at NICA // Phys. Part. Nucl. Lett. V. 21, No. 4. P. 789–792.
18. *Vandandoo U., Zhanlav T., Chuluunbaatar O., Gusev A., Vinitsky S., Chuluunbaatar G.* High-Order Finite Difference and Finite-Element Methods for Solving Some Partial Differential Equations. Switzerland: Springer Nature, 2024. 114 p.
19. *Zhanlav T., Chuluunbaatar O.* New Developments of Newton-Type Iterations for Solving Nonlinear Problems. Switzerland: Springer Nature, 2024. 281 p.
20. *Kalinovskaya L. V., Kalinovsky Yu. L.* Matrices and Determinants. M.: Infra-Engineering, 2024. 164 p. (in Russian).
21. *Alexeeva N. V., Barashenkov I. V., Dika A., De Sousa R.* The Energy-Frequency Diagram of the (1 + 1)-Dimensional Φ^4 Oscillon // J. High Energy Phys. 2024. V. 10. 136.
22. *Kurakin S., Badreeva D., Dushanov E. et al.* Arrangement of Lipid Vesicles and Bicelle-Like Structures Formed in the Presence of Aβ (25–35) Peptide // Biochimica et Biophysica Acta (BBA) — Biomembranes. 2024. V. 1866, Iss. 1. 184237.
23. *Batgerel B., Vinitsky S. I., Chuluunbaatar O. et al.* Schemes of Finite Element Method for Solving Multidimensional Boundary Value Problems // J. Math. Sci. 2024. V. 279. P. 738–755.
24. *Abdul-Magead I. A. M., Lukyanov V. K., Zemlyanaya E. V., Lukyanov K. V.* Analysis of the Proton Amplitude of Scattering on Bounded Nuclear Nucleons Based on Proton–Nucleus Scattering Data // Int. J. Mod. Phys. E. 2024. 2441012; https://doi.org/10.1142/S021830132441012X.
25. *Kats P. B., Kudravets A. V., Rimashevskaya A. S., Voskresenskaya O. O.* Comparative Study of Some Rigorous and Approximate Methods for Calculating the Energy Loss Straggling // Rad. Phys. Chem. 2024. V. 222. 111860.
26. *Bashashin M., Zemlyanaya E., Rahmonov I.* Simulation of the Magnetization Reversal Effect Depending on the Current Pulse Duration within the ϕ_0 Josephson Junction Model Using MPI and OpenMP Parallel Computing Techniques // Phys. Part. Nucl. 2024. V. 55, No. 3. P. 498–501.
27. *Rahmonov I. R., Rahmonova A. R., Streltsova O. I., Zuev M. I.* Python Implementation of Algorithms and Tools for Modeling the Dynamics of a Superconducting Quantum Interferometer with Two Josephson Junctions (DC SQUID). http://studhub.jinr.ru:8080/jjbook/DC-SQUID.html (in Russian).
28. *Ayriyan A. S., Buša J., Jr., Melezhik V. S.* Parallelization of a Computational Scheme Based on Two-Dimensional DVR for Integrating Time-Dependent Three-Dimensional Schrödinger Equation // Phys. Part. Nucl. Lett. 2025. V. 22, No. 3. P. 617–621; doi: 10.1134/S1547477125700244.
29. *Chervyakov A. M.* On Finite-Element Modeling of Large-Scale Magnetization Problems with Combined Magnetic Vector and Scalar Potentials // Phys. Part. Nucl. Lett. V. 21, No. 5. P. 1074–1083.
30. *Soloviev A. G., Solovjeva T. M., Lukyanov K. V., Zemlyanaya E. V.* Web Interface Based on the JINR Cloud Infrastructure for Fitting Experimental Small-Angle Scattering Data Using ROOT Package Tools // Modern Information Technologies and IT-Education. 2024. V. 20, No. 3 (in Russian).
31. *Yukalov V. I., Yukalova E. P.* Strong-Coupling Limits Induced by Weak-Coupling Expansions // Ann. Phys. (N. Y.). 2024. V. 467. 169716.
32. *Siurakshina L., Yushankhai V.* Anisotropic Spin Models for Iridium Oxides: Justification in the Cluster Quantum-Chemical Approach // Bull. Russ. Acad. Sci.: Phys. (in press) (in Russian).
33. *Palii Yu. G., Bogolubskaya A. A., Yanovich D. A.* Modeling of the Operation of the Algorithm QAOA with the Software Library Cirq. JINR Preprint P11-2024-57. Dubna, 2024 (in Russian).
34. *Ososkov G. A., Uzhinskiy A. V., Nechaevskiy A. V.* Application of Deep Learning Methods to Solve Various Tasks in Agriculture // Collection “Dubna State University. 30 Years in Science”. Dubna, 2024. P. 289–302 (in Russian).