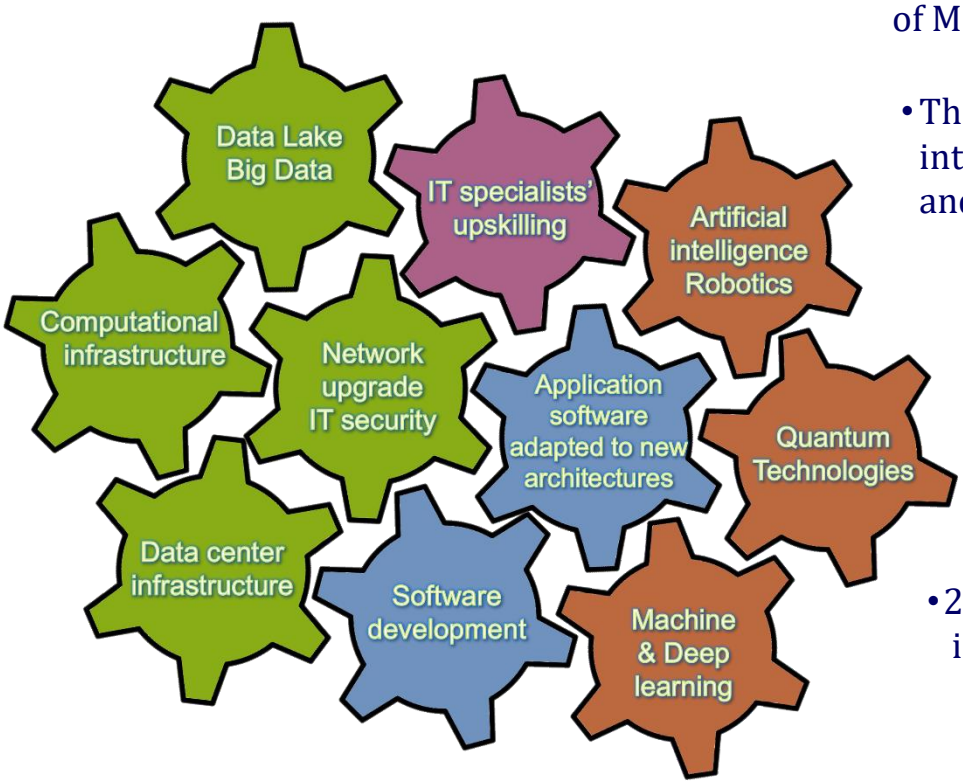
Meshcheryakov Laboratory of Information Technologies

Joint Institute for Nuclear Research

2024

Scientific IT Ecosystem

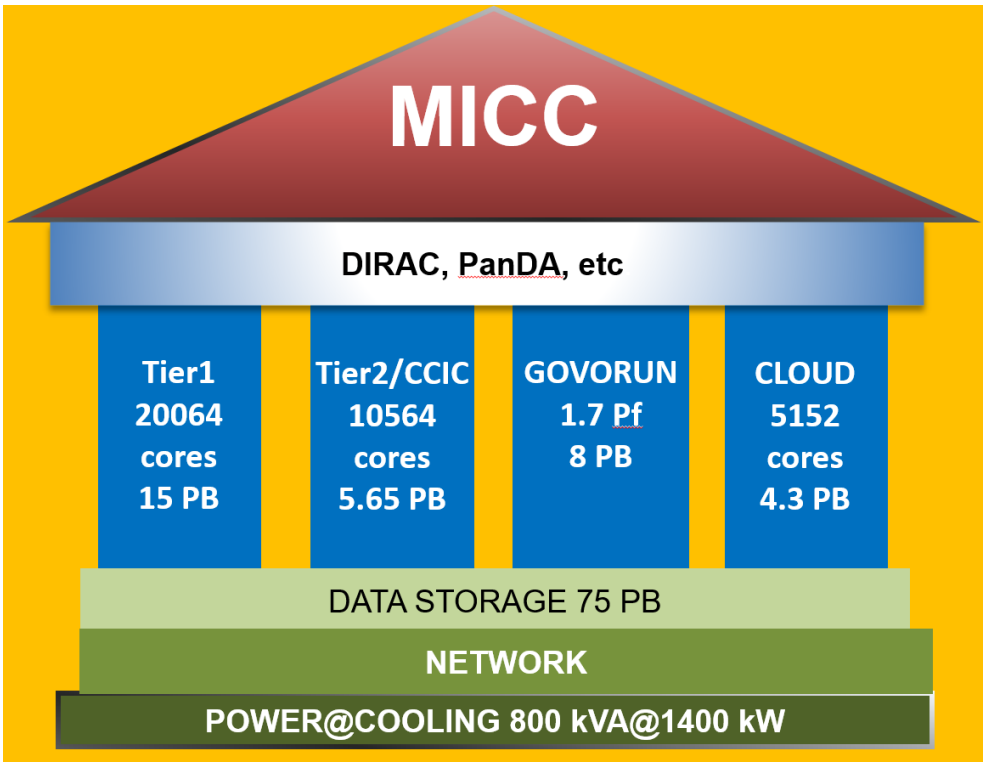
The Meshcheryakov Laboratory of Information Technologies (MLIT, former Laboratory of Computing Techniques and Automation) was founded in August 1966. The main directions of MLIT activities are connected with



- The coordinated development of interconnected IT technologies and computational methods
- Providing the IT services necessary for the fulfillment of the JINR Topical Plan on Research and International Cooperation in an efficient manner.
- 24x7 support of the computing infrastructure and services.

Multifunctional Information and Computing Complex

- multi-functionality,
- high performance,
- task-adapted data storage system,
- high reliability and availability,
- information security,
- scalability,
- customized software environment for different user groups,
- high-performance telecommunications and modern local network.

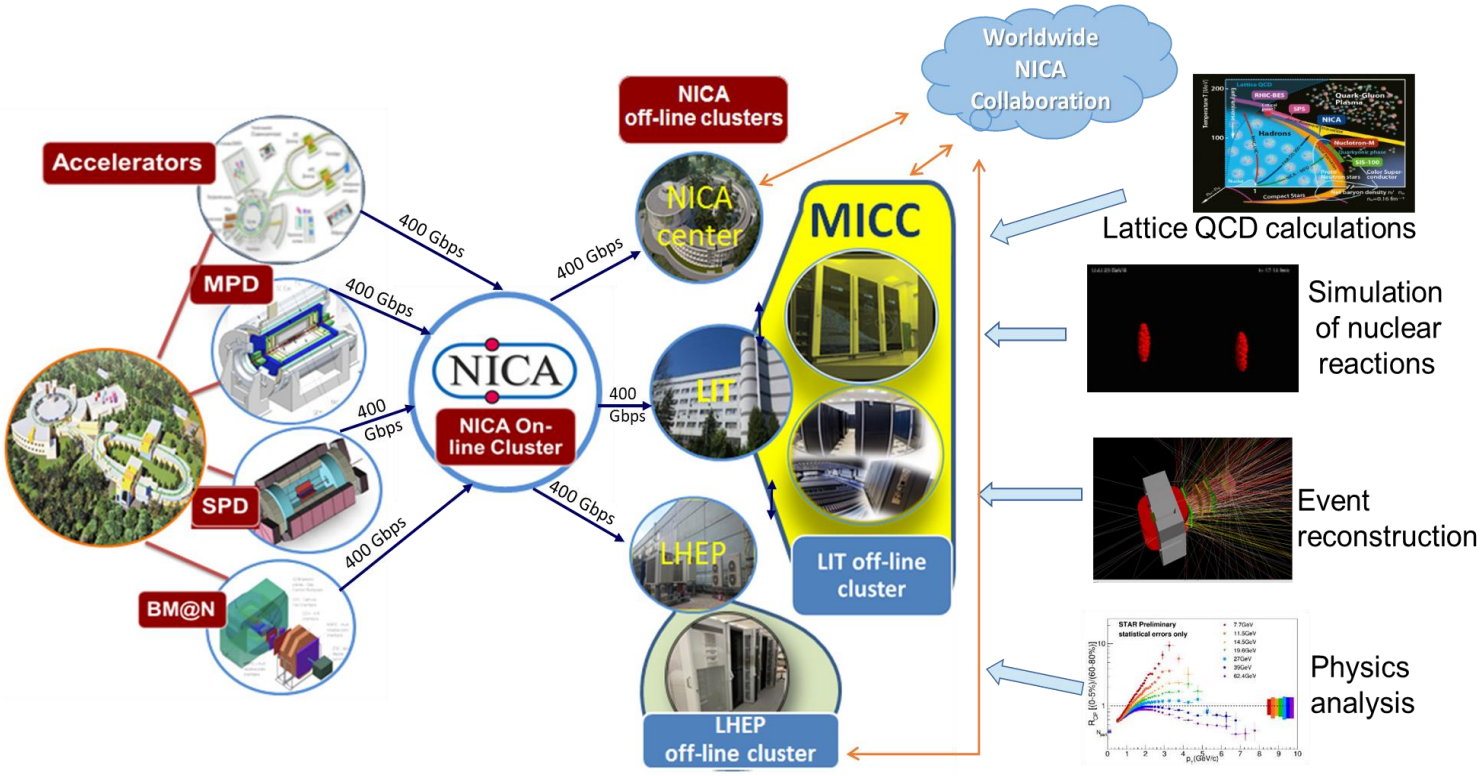


The IT infrastructure is one of JINR's basic facilities.

micc.jinr.ru

NICA Computing

An off-line computer complex for data modeling, processing, analysis and storage within the NICA project consists of territorially distributed on-line and off-line clusters connected by the high-speed computer network. The NICA computing and information off-line cluster in MLIT was organized on the basis of the MICC as a distributed scalable hybrid cluster.



Support for the JINR neutrino program

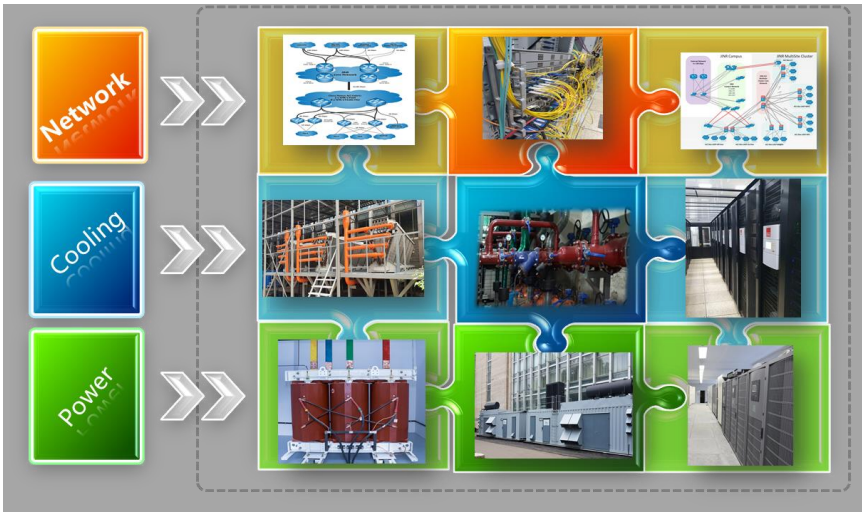
MLIT contribution: engineering infrastructure (electricity, UPS, cooling, network, racks, manpower)

DLNP contribution: computing and storage resources (CPUs/GPUs&disks)

Computational resources and support of users for the **JINR neutrino program** using the cloud infrastructure of MICC.

The **NOvA, Baikal-GVD and JUNO** experiments are the **major users** of the cloud infrastructure.

MICC Power @ Cooling @ Network

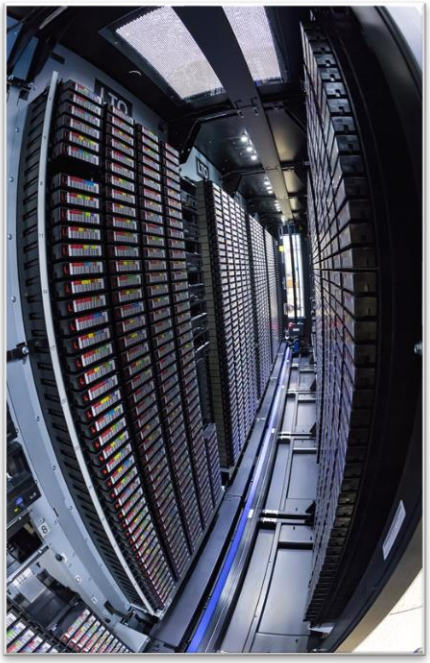


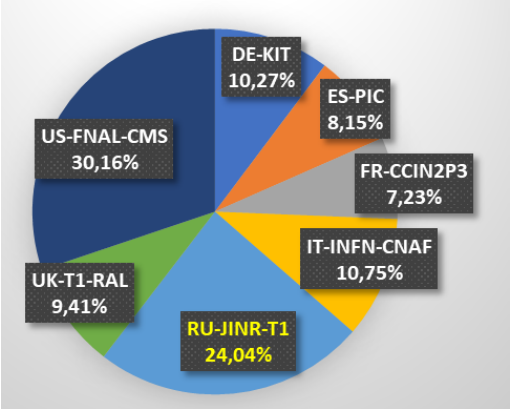
Wide Area Network 3x100 Gbps
Cluster Backbone 4x100 Gbps
Campus Backbone 2x100 Gbps

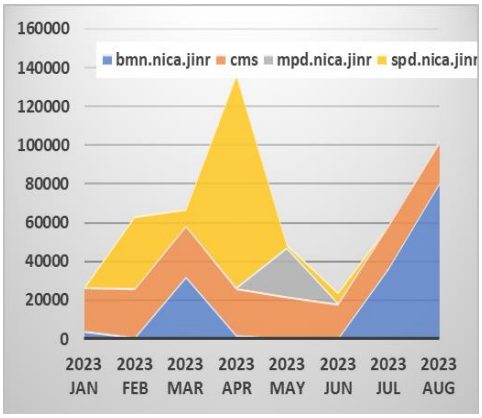
Dry chillers
In-Row systems
Total cooling 1400 kW

Uninterruptible power supplies (UPS) 8x300 kVA
Diesel-generator units (DGU) 2x1500 kVA
Transformers 2x2500 kVA

Grid Infrastructure

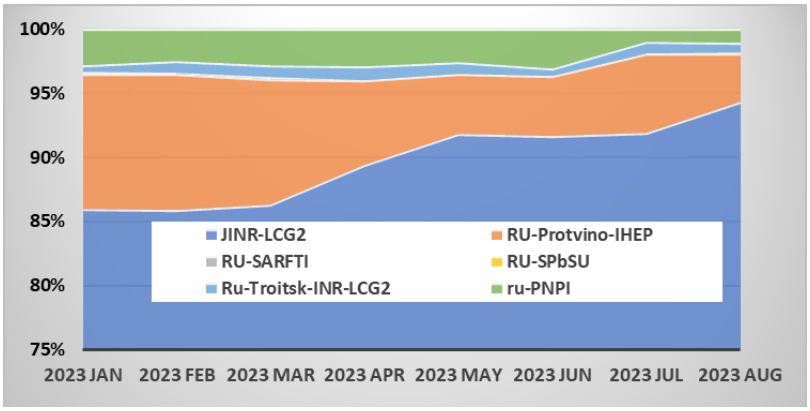






The CMS Tier1 centre at JINR has demonstrated stable work through the entire period since its launch into full operation in 2015 and is also used for the NICA experiments.
JINR Tier1 is regularly ranked on top among world Tier1 sites that process data from the CMS experiment at the LHC.

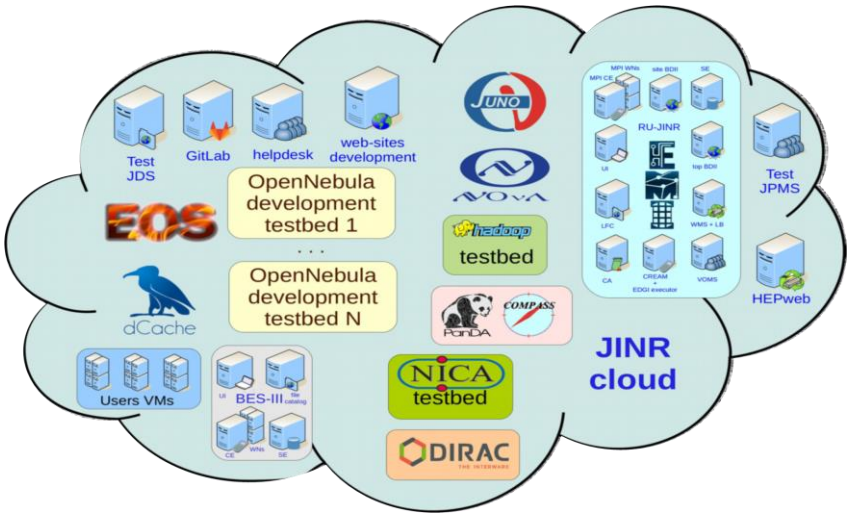
The Tier2 center is used to process data from the NICA, LHC, ILC, NOvA experiments and others, as well as by JINR local users.



The JINR Tier2 output is the highest (89.3%) in the Russian Data Intensive Grid (RDIG) Federation.

Cloud infrastructure

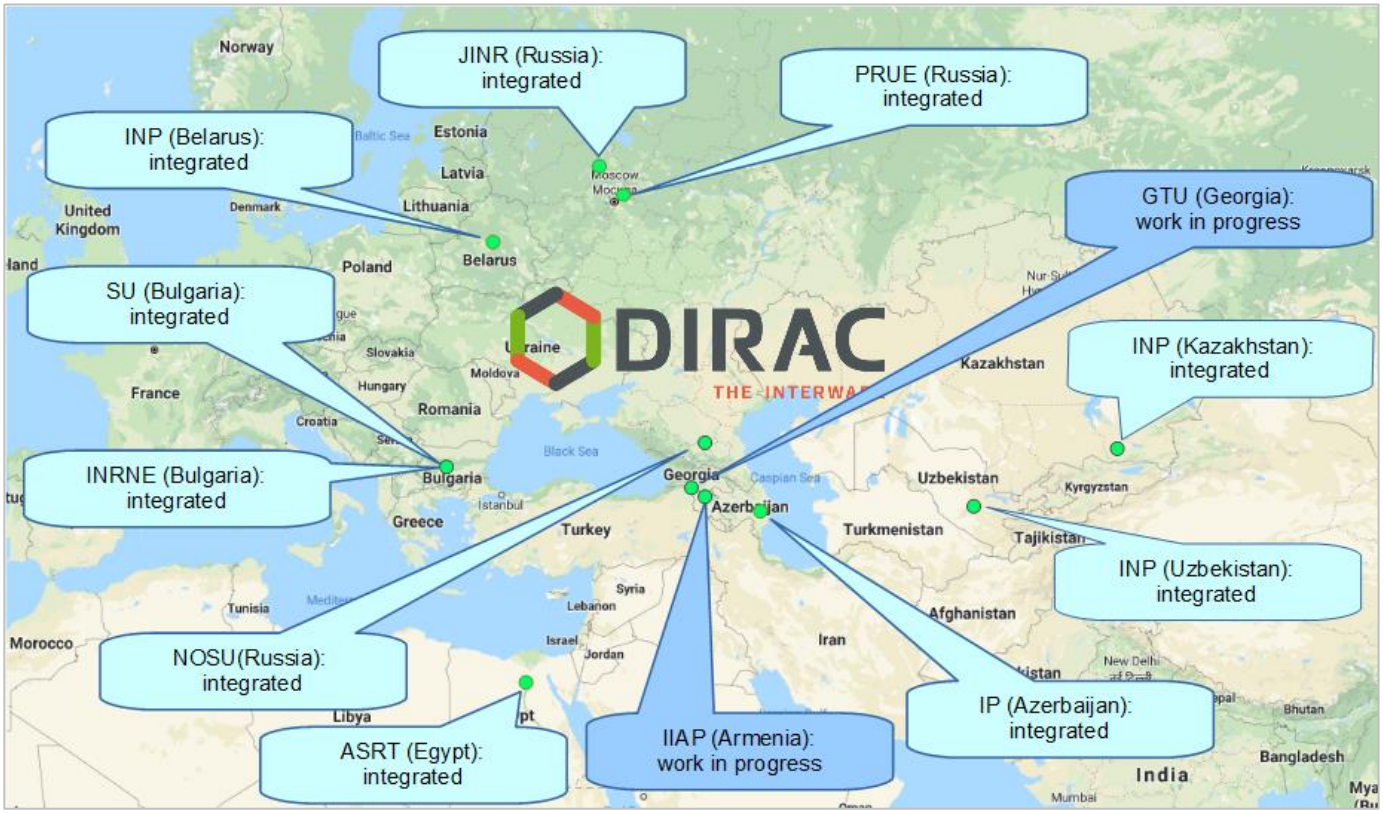
❖ Cloud Platform – OpenNebula
Virtualization – KVM
Storage (Local disks, Ceph)
Total Resources:
- 5,000 CPU cores
- 60 TB RAM
- 3.1 PB of raw ceph-based storage



❖ A universal computing resource that supports individual scientists and provides a variety of common services, ranging from simple web sites to complex multi-user computational system.

❖ Utilization:
- VMs for JINR users
- Computational resources for neutrino experiments
- Testbeds for research and development in IT
- COMPASS production system services
- Data management system of the UNECE ICP Vegetation
- Scientific and engineering computations
- Service for data visualization and execute nodes for it
- Gitlab and some others

❖ **DIRAC-based distributed information and computing environment integrated the JINR Member State organizations' clouds.**

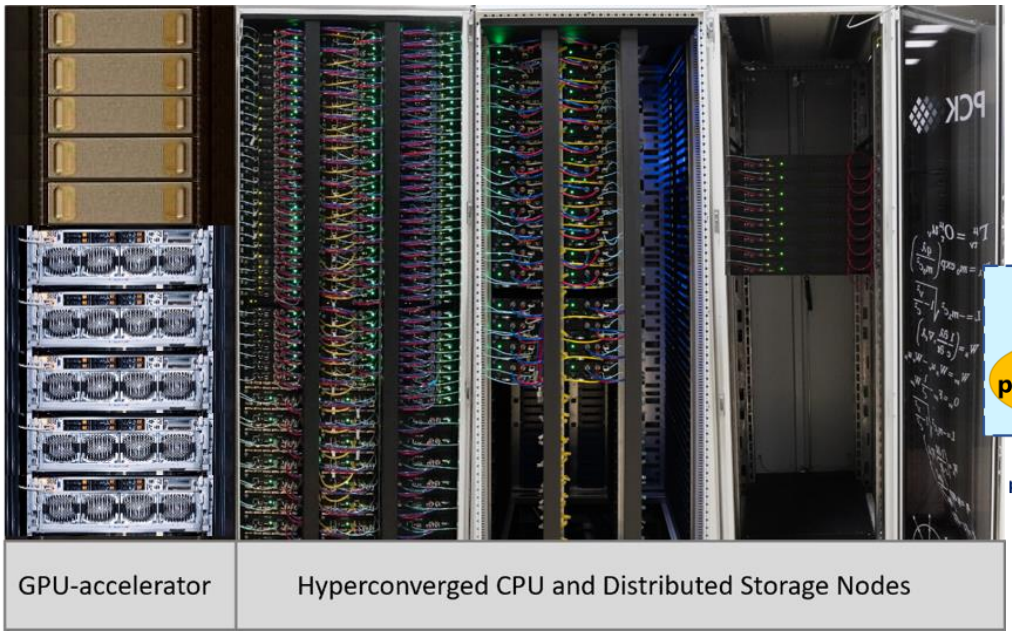


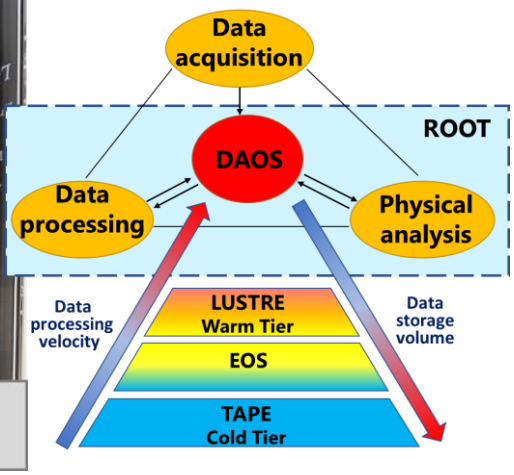
HybriLIT platform and "Govorun" supercomputer

The HybriLIT platform consists of two elements: the education and testing polygon and the "Govorun" supercomputer, combined by a unified software and information environment. The supercomputer was named after N.N. Govorun.

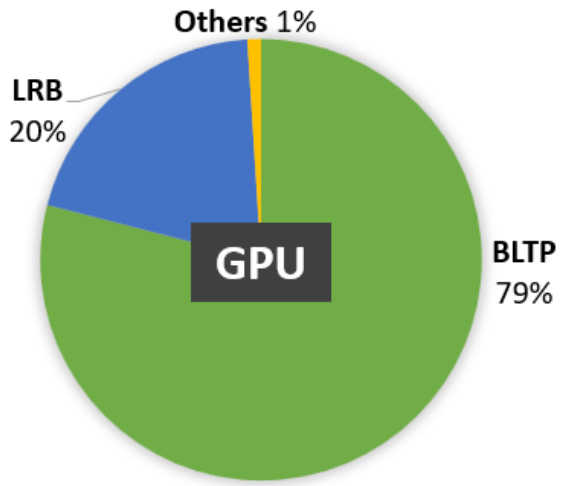
- Hyper-converged software-defined system
- Hierarchical data processing and storage system
- Scalable solution Storage-on-demand
- Total peak performance: 1.7 PFlops DP
- GPU component based on the NVIDIA Tesla V100&A100
- CPU component based on RSC "Tornado" liquid cooling solutions
- The most energy-efficient center in Russia (PUE = 1,06)
- Storage performance >300 GB/s

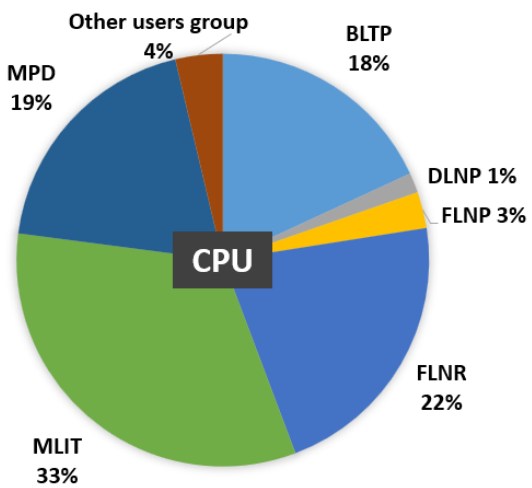
hlit.jinr.ru





The resources of the "Govorun" supercomputer are used by scientific groups from all the Laboratories of the Institute within 25 themes of the JINR Topical Plan for solving a wide range of tasks in the field of theoretical physics, as well as for the physical modeling and experimental data processing.

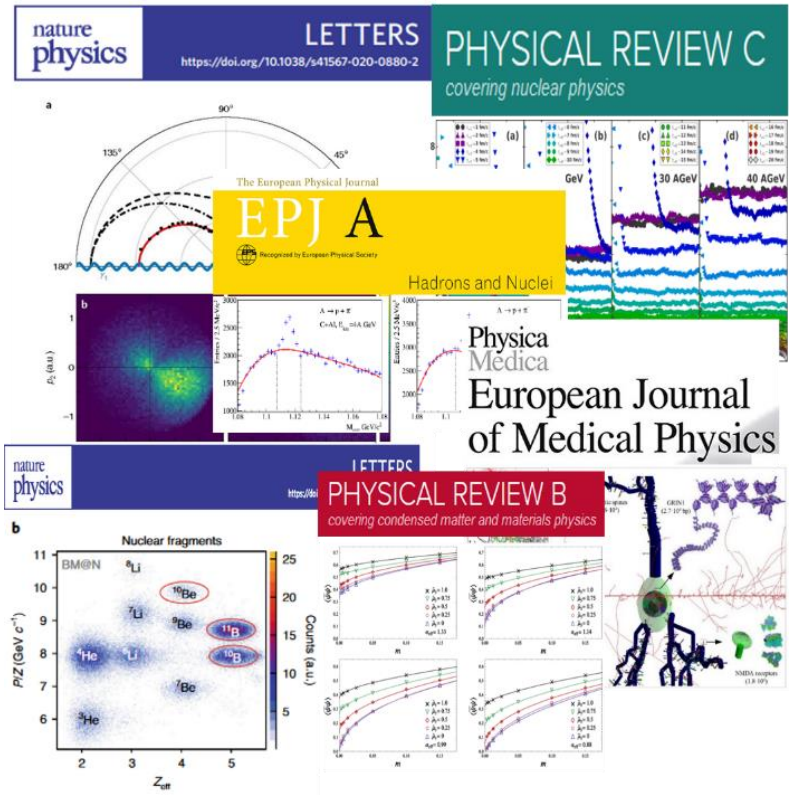




Key projects that use the resources of the "Govorun" supercomputer:

- NICA megaproject,
- calculations of lattice quantum chromodynamics,
- computations of the properties of atoms of superheavy elements,
- studies in the field of radiation biology,
- calculations of the radiation safety of JINR's facilities.

> 250 user papers (two in Nature Physics)



ML/DL/HPC ecosystem

The ML/DL/HPC ecosystem is used for machine and deep learning tasks. At the same time, the accumulated tools and libraries can be more widely used for scientific research, including:

- numerical computations;
- parallel computing on CPUs and GPUs;
- results visualization;
- accompanying numerical results with the necessary formulas and explanations.

jupyter

Keras

scikit learn

NumPy

pandas

matplotlib

HPClab component — High Performance Computing

VM with JupyterHub and SLURM
[**https://jlabhpc.jinr.ru**]

- ❑ Intel Xeon Gold 6126 (24 Cores @ 2.6 GHz)
- ❑ 32 GB RAM

Educational component

JupyterLab Server [**https://studhub.jinr.ru**]

- ❑ 2x Intel Xeon Gold 6152 (22 Cores @ 2.1 GHz)
- ❑ 512 GB RAM

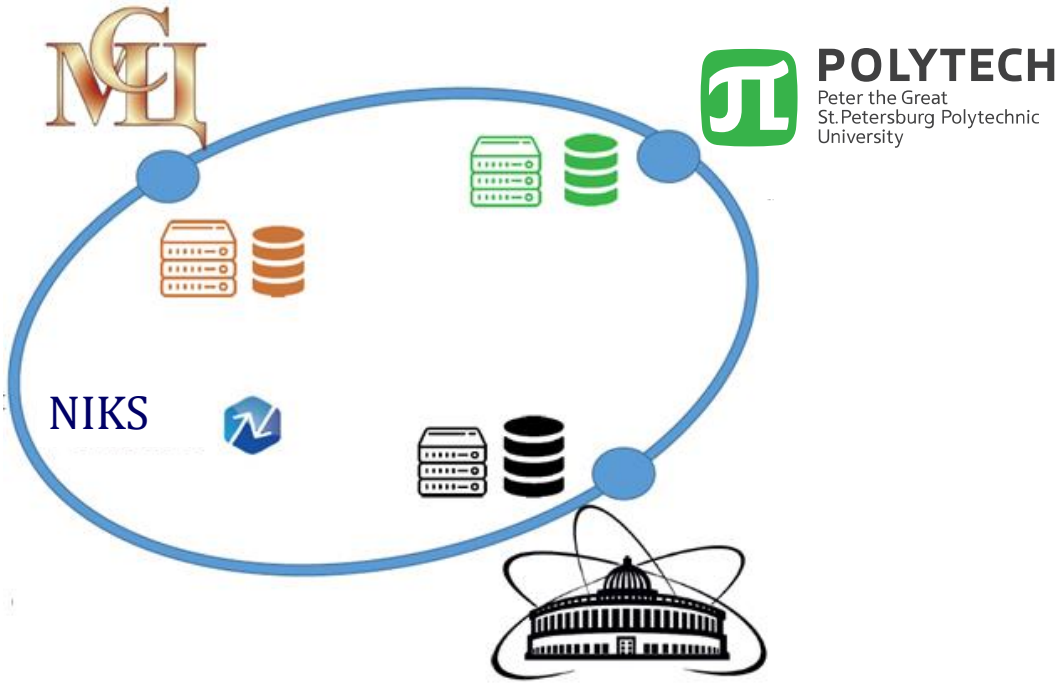
Computation component

Server with NVIDIA Volta
[**https://jhub2.jinr.ru**]

- ❑ 2x Intel Xeon Gold 6148 (20 Cores @ 2.4 GHz)
- ❑ 4x NVIDIA Tesla V100 SXM2 32 GB HBM2
- ❑ 512 GB RAM

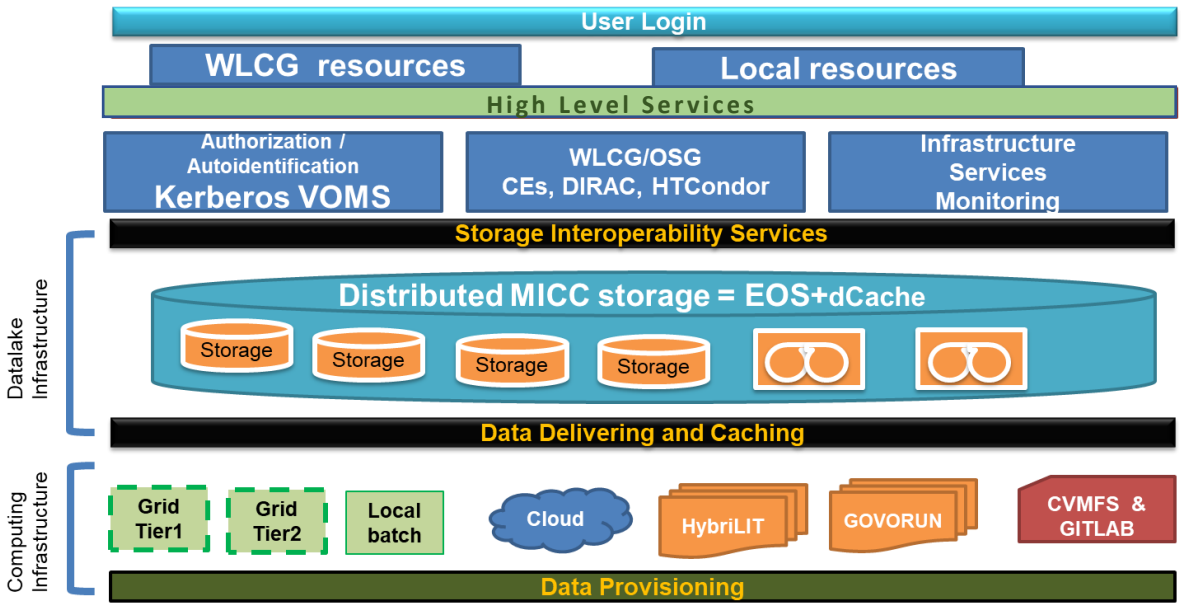
Unified scalable supercomputer research infrastructure

Based on the integration of the supercomputers of JINR, of the Interdepartmental Supercomputer Center of the Russian Academy of Sciences and of Peter the Great St. Petersburg Polytechnic University, a unified scalable supercomputer research infrastructure based on the National Research Computer Network of Russia (NIKS) was created. Such an infrastructure is in demand for the tasks of the NICA megaproject.



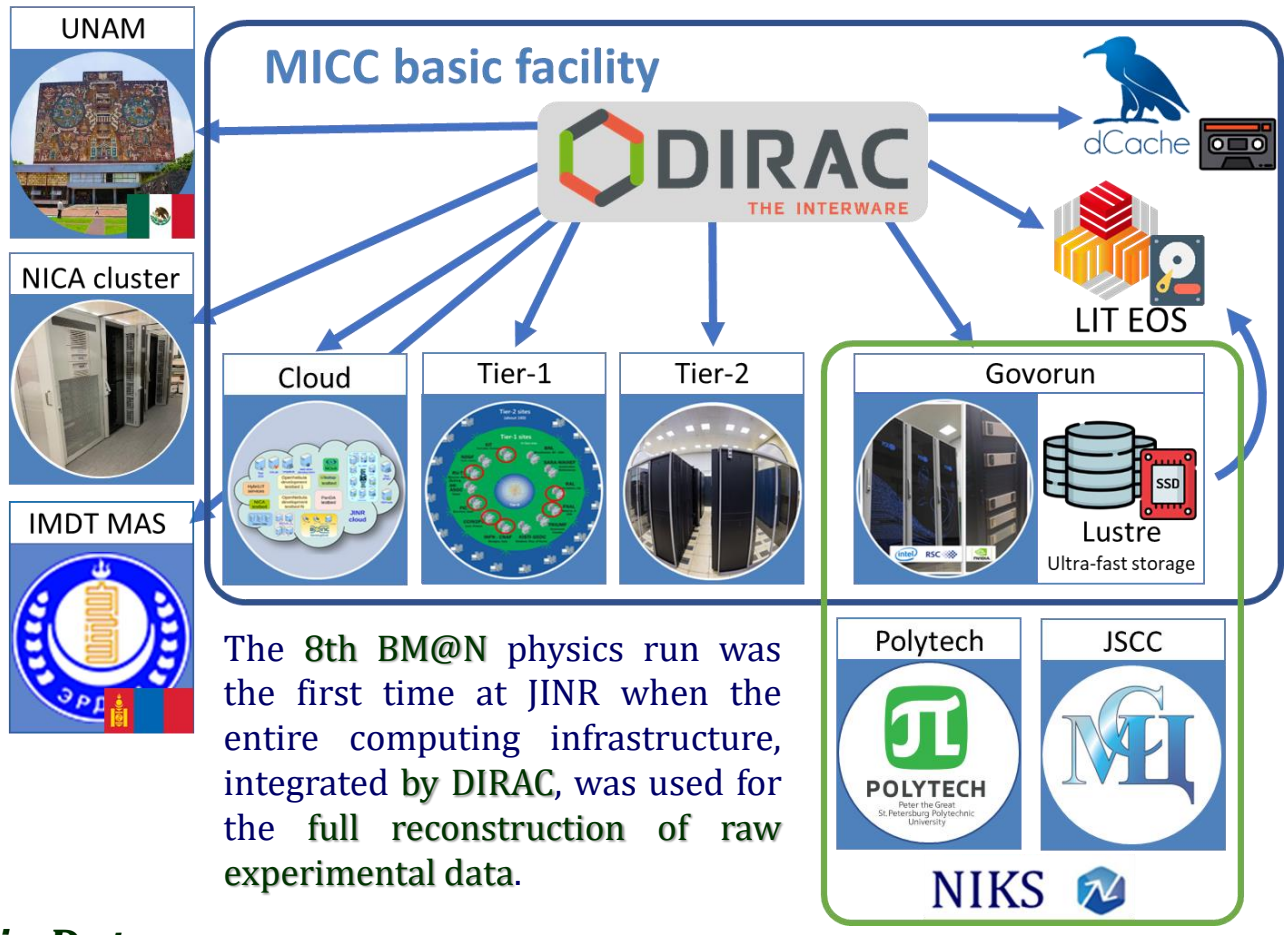
Data Lake

- ✓ The JINR data lake was built as a distributed EOS storage system.
- ✓ EOS is used for storing and accessing big arrays of information. It can be applied for collective data simulation, storage of raw data gathered from experimental setups, data processing and analysis.
- ✓ There is currently 23.9 PB of disk space available for EOS.
- ✓ Baikal-GVD, DANSS, FOBOS, JUNO, BM@N, MPD, SPD, PANDA are its major users.



Integration of heterogeneous computing resources

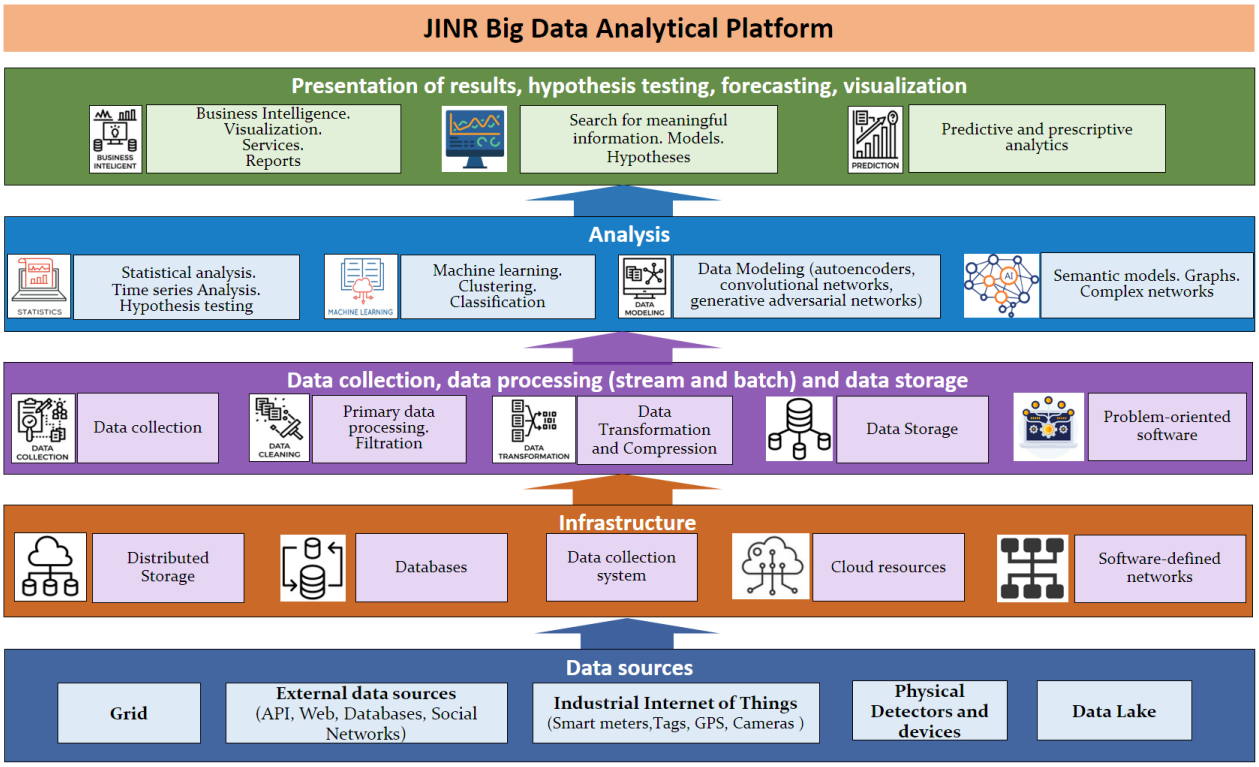
A heterogeneous computing environment (Tier1, Tier2, SC "Govorun", cloud, ect.), based on the DIRAC platform, was created for processing and storing data of the experiments conducted at JINR. The distributed infrastructure is used by the MPD, Baikal-GVD, BM@N, SPD.



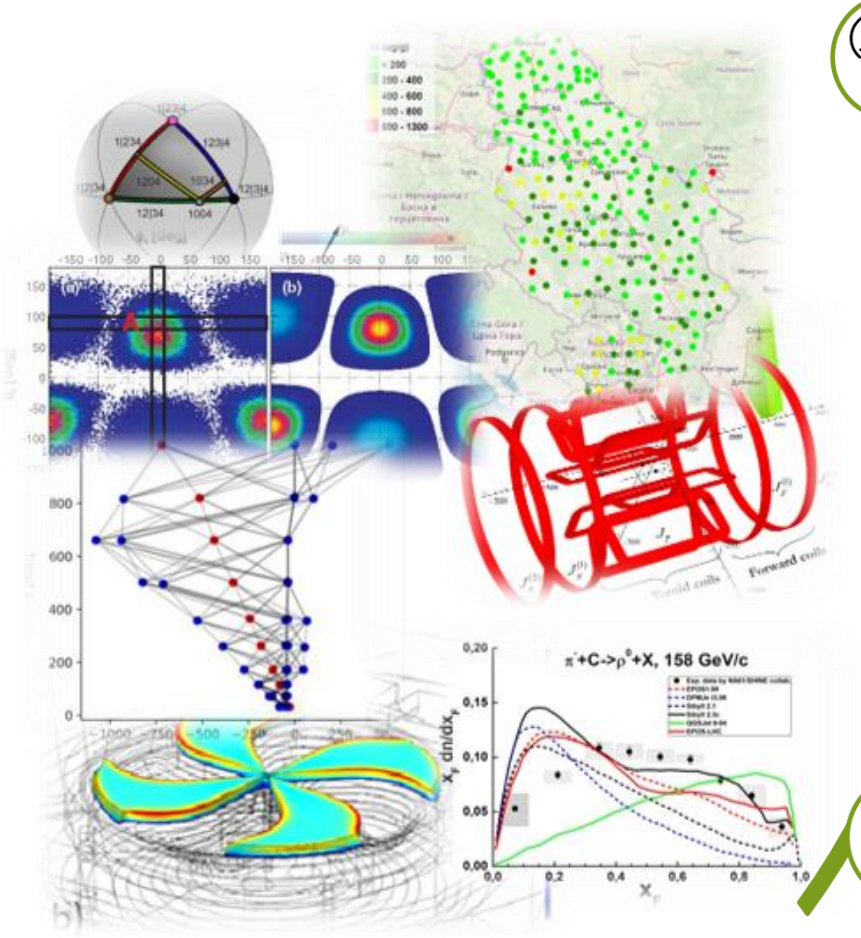
The 8th BM@N physics run was the first time at JINR when the entire computing infrastructure, integrated by DIRAC, was used for the full reconstruction of raw experimental data.

Big Data

- Bringing best of Big Data approaches to JINR practices
- Providing the Big Data infrastructure for users

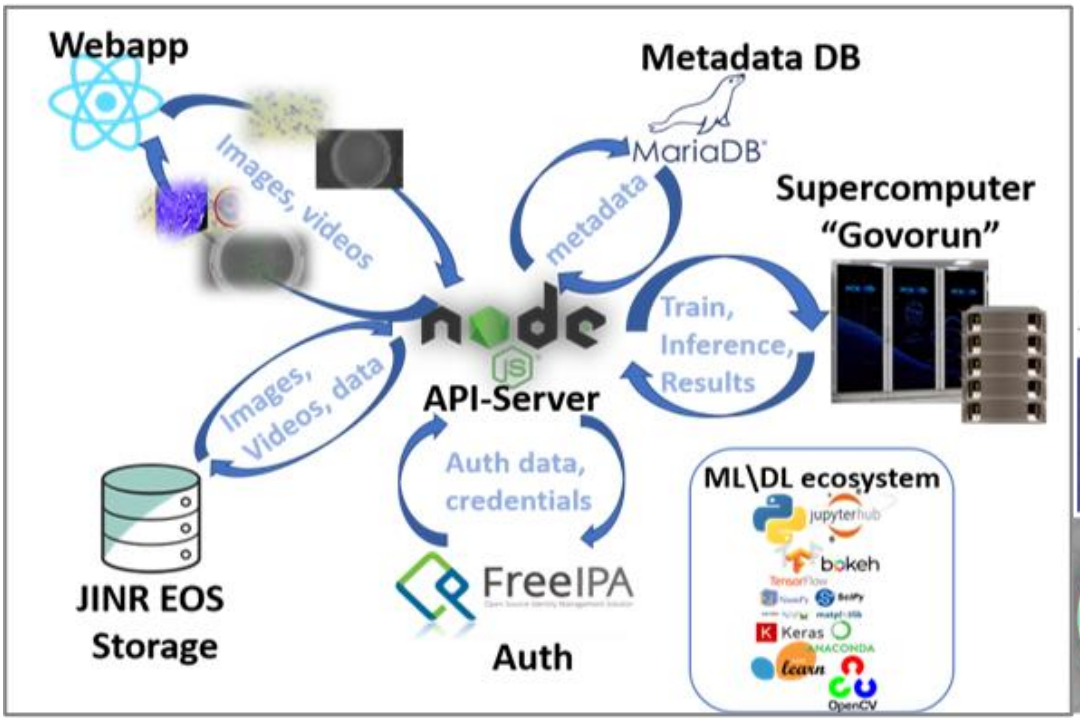


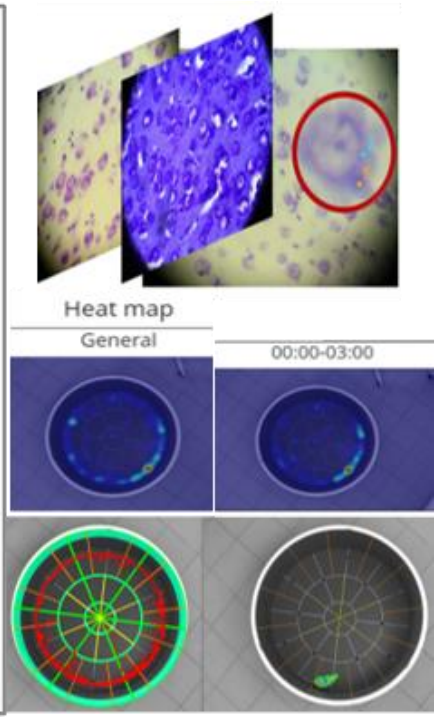
Methods, Algorithms and Software



- Numerical modeling of complex physical systems
- Experimental data processing and analysis
- Big Data
- Machine and Deep learning
- AI and robotics
- Computer algebra
- Quantum computing

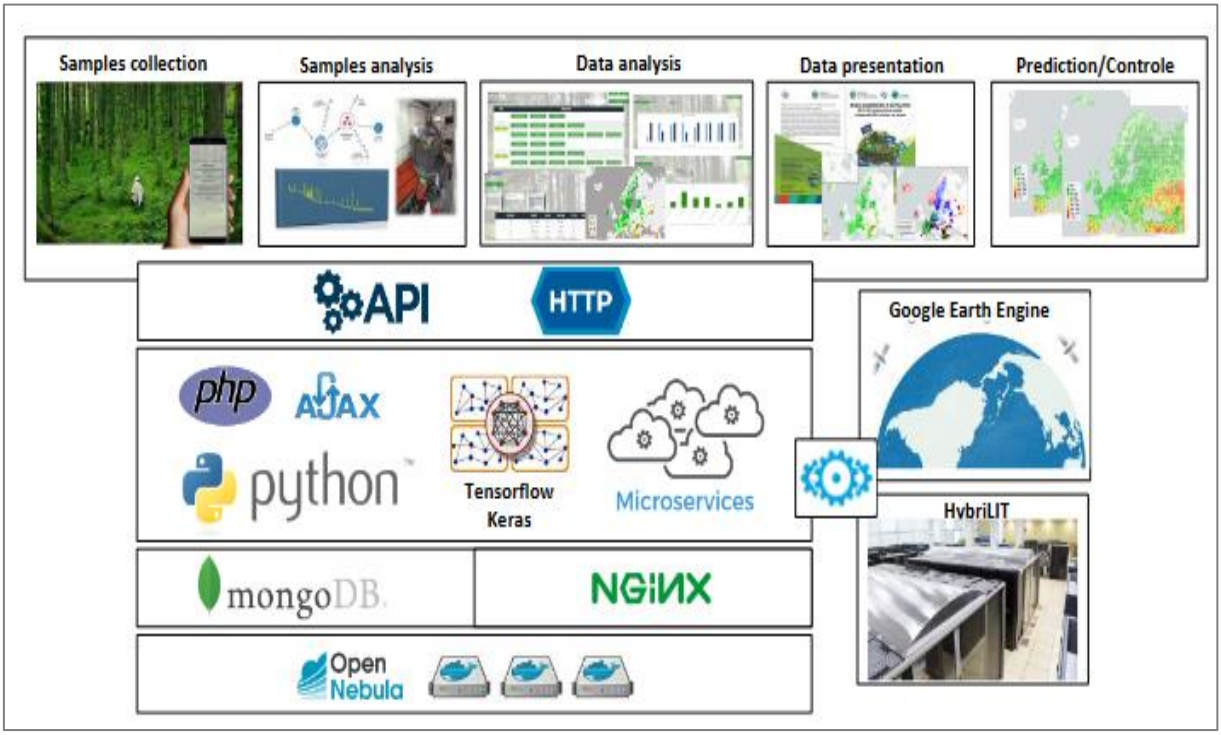
The joint project of MLIT and Laboratory of Radiation Biology is focused on creating an Information System (IS) as a set of IT solutions providing the storage, analysis and visualization of data from experiments at LRB. The IS is based on a stack of neural network and classical algorithms of computer vision.





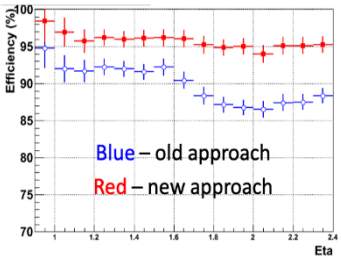
Intelligent Environmental Monitoring Platform

Within the framework of cooperation between MLIT and Frank Laboratory of Neutron Physics, the work on the prediction of the air pollution by heavy metals using biomonitoring data, satellite imagery and different technologies of machine and deep learning is in progress. On the MLIT cloud platform, the Data Management System of the UNECE ICP Vegetation was created to provide its participants with a modern unified system of collecting, analyzing and processing biological monitoring data.

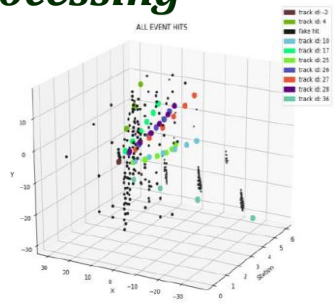


Methods and Software for Experimental Data Processing and Analysis

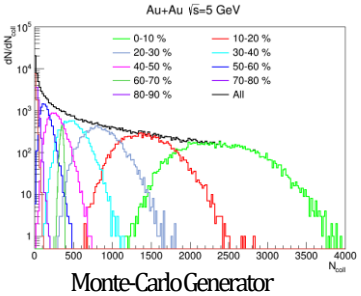
- ✓ Physical processes modeling
  - event simulation
  - GEANT-simulation of experimental setup
- ✓ Event reconstruction and data analysis
  - particle trajectory reconstruction
  - particle identification
  - physical processes reconstruction
  - data analysis
- ✓ Applied software and Data Bases
  - DBs for experimental services
  - experimental software frameworks
  - data model and data processing model
  - event visualization and monitoring



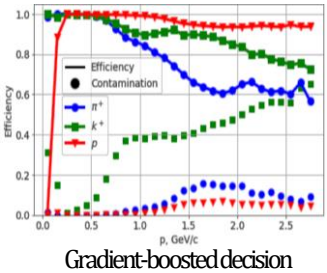
Effective algorithm for TeV muons reconstruction in CMS



Deep neural networks for solving tracking problems in BM@N, BESIII, SPD



Monte-Carlo Generator DCM-QGSM-SMM for NICA



Gradient-boosted decision trees for PID in MPD

End-to-end quantum intelligent computing

Development of an intelligent automatic control system for the control of the elements of the physical facility of the NICA complex using the methods of end-to-end quantum software engineering (together with VBLHEP).

Intelligent computing toolkit:
- Soft computing
- Quantum computing
- Quantum soft computing

Quantum software engineering:
- Quantum programming
- Quantum algorithms
- Quantum software simulators

Quantum deep machine learning:
- Quantum neural networks
- Quantum genetic algorithms
- Quantum approximated optimization algorithms
- Quantum information learning models

End-to-End Quantum Software Engineering

Quantum intelligent control:
- Quantum control algorithms
- Simulators of quantum control algorithms
- Quantum intelligent controllers and systems

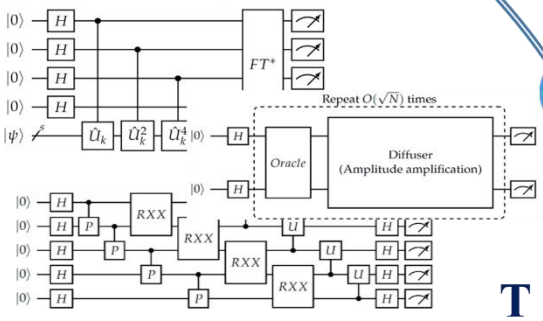
Intelligent cognitive robotics:
- Quantum control robotic algorithms
- Simulators of robotics models
- Software and hardware implementation of intelligent cognitive robotics

Self-organized intelligent control system design:
- Software development of self-organized knowledge bases
- Implementation of quantum self-organized controllers at JINR

Quantum computing and quantum algorithms

Software quantum simulators for computing on computers of a classical architecture using CPUs and GPUs is of particular interest for solving a number of problems in condensed matter, high-energy physics, quantum chemistry, AI and others.

Quantum algorithms



Quantum simulators

Qiskit
QuEST
qsim
CUQUANTUM

SC "Govorun"

TASKS

- Form a list of QAs required to solve tasks within the studied physical models
- Select the type of quantum simulator to simulate a classical architecture on computers
- Define resources for the selected quantum-limiting capabilities of available computing simulators (number of qubits and computation time)
- Search for solutions to urgent problems of quantum chemistry and study the chemical properties of heavy elements

Walt Platform For Web Application Development

The Web Application Lego Toolkit (WALT) is a template-oriented platform designed for the development of web applications of various degrees of complexity. Web applications developed using WALT are characterized by high performance and humble server resource requirements.

JINR corporate web applications developed using WALT	
ADB2	JINR's management accounting
PIN	Staff information
EDMS "Documents DB"	Electronic storage for administrative activity documents
NICA EVM	Project structure, workplans, expenses, Costbook, reports, etc.
EDMS "Dubna"	JINR electronic document management system
HR JINR	Staff administrative information
MAP JINR	Basic map of JINR's sites
Gateway	Universal gateway for data exchange between various systems
ISSC	Scientific attestation support system
CERN DB	JINR's staff at CERN: trips, accommodation, reports, etc.
EDMS "Advance reports"	Data preparation for business trip accounting reports
Checkpoint lists	Lists for access to JINR's sites in a limited access mode
DES	JINR Digital EcoSystem that integrates existing and future services

JINR Digital Eco System (DES)

JINR Digital Eco System

integrates existing and future services

to support
scientific,
administrative, and
social activities,
maintenance of the engineering
and IT infrastructures

to provide
reliable and secure access to
various types of data

to enable
a comprehensive analysis
of information

using
modern Big Data
technologies and
artificial
intelligence.

eco.jinr.ru

Development of the system for training and retraining IT specialists

Training courses, master classes and lectures

MLIT staff and leading scientists from JINR and its Member States

Leading manufacturers of modern computing architectures and software

Parallel programming technologies

OpenMP, MPI, OpenCL, NVIDIA CUDA

Tools for debugging and profiling parallel applications

Intel Cluster Studio

Work with applied software packages

COMSOL MULTIPHYSICS, Wolfram Mathematica, ROOT, MATLAB, GEANT4, Maple

Frameworks and tools for ML/DL tasks

jupyterhub, TensorFlow, NumPy, learn

Quantum algorithms, quantum programming and quantum control

it SCHOOL JINR

Involvement of young specialists in solving tasks that face JINR using state-of-the-art information technologies

International Conference "Distributed Computing and Grid Technologies in Science and Education"

- Distributed Computing Systems
- Computing for MegaScience Projects
- Data Management, Organisation and Access
- HPC
- Quantum Information Processing
- Big Data Analytics and Machine Learning
- Research Infrastructure

MMCP MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS

Methods, software and program packages for

- data processing and analysis
- modeling complex physical systems, computational bioinformatics
- computer algebra
- quantum computing
- machine learning and Big Data
- parallel and hybrid calculations

International Symposium on Nuclear Electronics and Computing

- Detector & Nuclear Electronics
- Triggering, Data Acquisition
- Computing for Large-Scale Facilities
- Distributed Storage Systems, Datalakes
- Distributed Computing
- GRID and Cloud Computing
- Machine Learning Algorithms
- Big Data Analytics
- Innovative IT Education

lit@jinr.ru

http://lit.jinr.ru

Russia, 141980
Dubna, Moscow Region,
Joliot-Curie 6, MLIT JINR