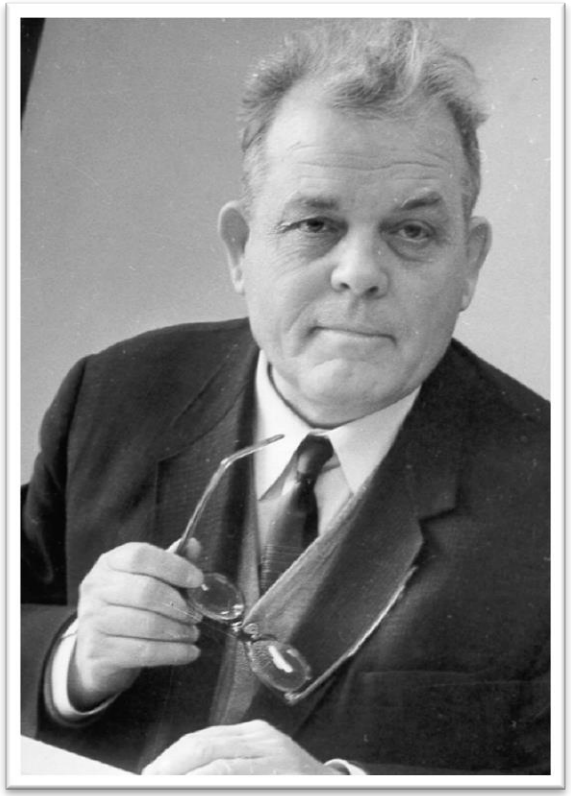
Meshcheryakov Laboratory of Information Technologies



History

The Laboratory of Computing Techniques and Automation of the Joint Institute for Nuclear Research in Dubna was founded in August 1966.

The main directions of the activities at the Laboratory are connected with the provision of networks, computer and information resources, as well as mathematical support of a wide range of research at JINR in high energy physics, nuclear physics, condensed matter physics, etc.

Computing is an integral part of theory, experiment, technology development

(17.09.1910 – 24.05.1994)

(18.03.1930 – 21.07.1989)

On 25 March 2021 the Committee of Plenipotentiary Representatives of the Governments of the JINR Member States **decided to name the Laboratory of Information Technologies after M. G. Meshcheryakov** for his outstanding contribution to the creation and development of the network infrastructure and the Information and Computing Complex of the Laboratory, the Institute, and the Member States.

MLIT today: Scientific IT-ecosystem

Staff: 297
Scientists: 100
Doctors of Science: 25
Candidates of Science: 62
Campus network 2x100 Gbps
Multisite network 4x100 Gbps
Telecommunication channel 3x100 Gbps
Grid Tier1 and Tier2 for global data processing
JINR Cloud computing
JINR Member States' Cloud environment
"Govorun" supercomputer

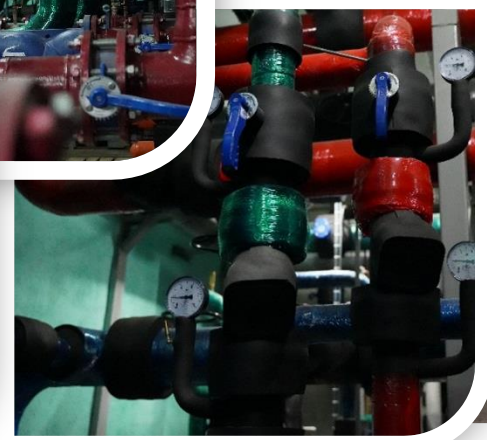
- The coordinated development of interconnected IT technologies and computational methods
- Providing the IT services necessary for the fulfillment of the JINR Topical Plan on Research and International Cooperation in an efficient manner.
- Building world-class competence in IT and computational physics.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

Engineering Infrastructure



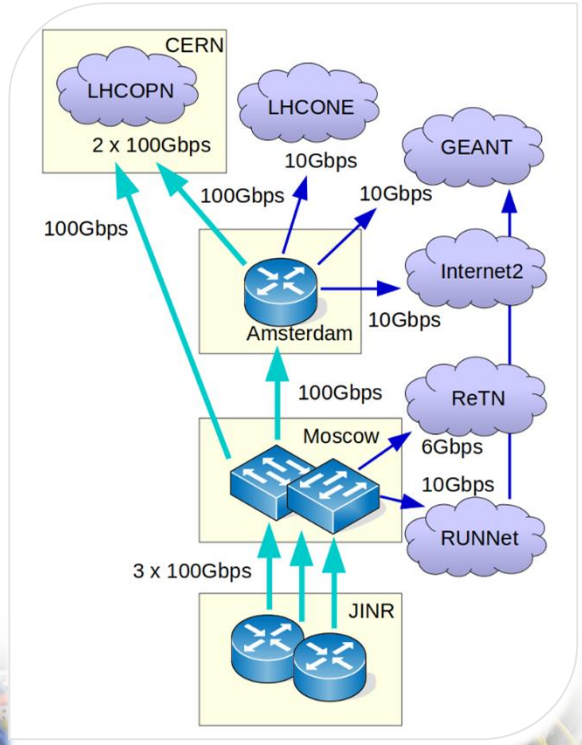




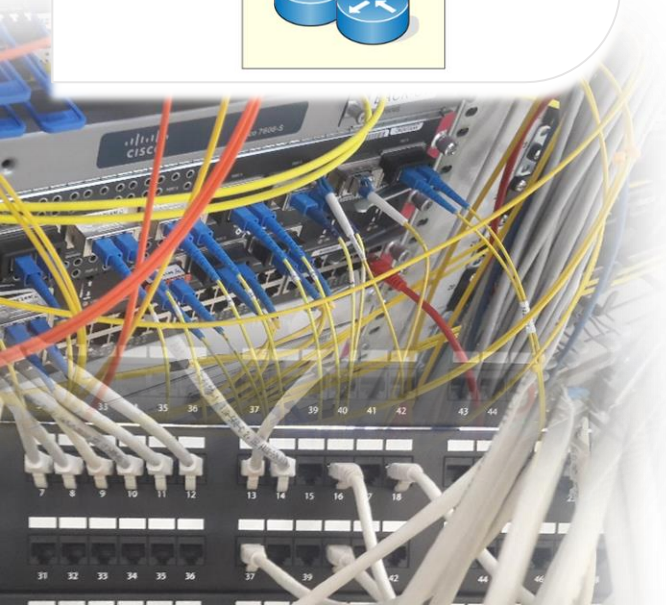


- ✓ Power supply expansion
- ✓ Cooling system for the MICC machine hall
- ✓ 100% “hot water” cooling system of the “Govorun” supercomputer
- ✓ Guaranteed power supply using diesel generators and uninterruptible power supplies

Networking



- JINR-Moscow 3x100 Gbit/s
- JINR-CERN - 100 Gbit/s and JINR-Amsterdam 100 Gbit/s for LHCOPN, LHCONE, GEANT networks
- Direct channels up to 100 Gbit/s for communication using RU-VRF technology with the collaboration of RUHEP research centers and with Runnet, ReTN networks
- The multi-site cluster network with a bandwidth 4x100 Gbit/s between VBLHEP and MLIT



The JINR LAN comprises:

8768 network elements

17602 IP-addresses

6377 users registered within the network

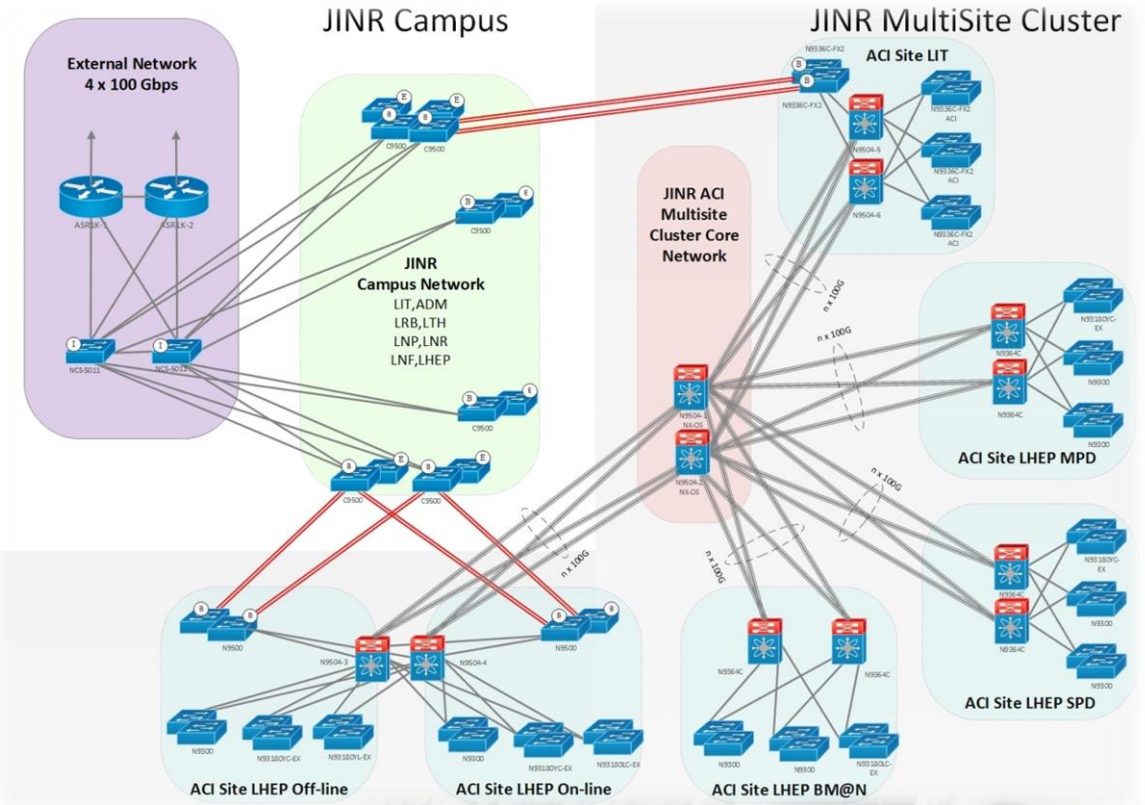
4203 *.jinr.ru service users

1419 digital library users

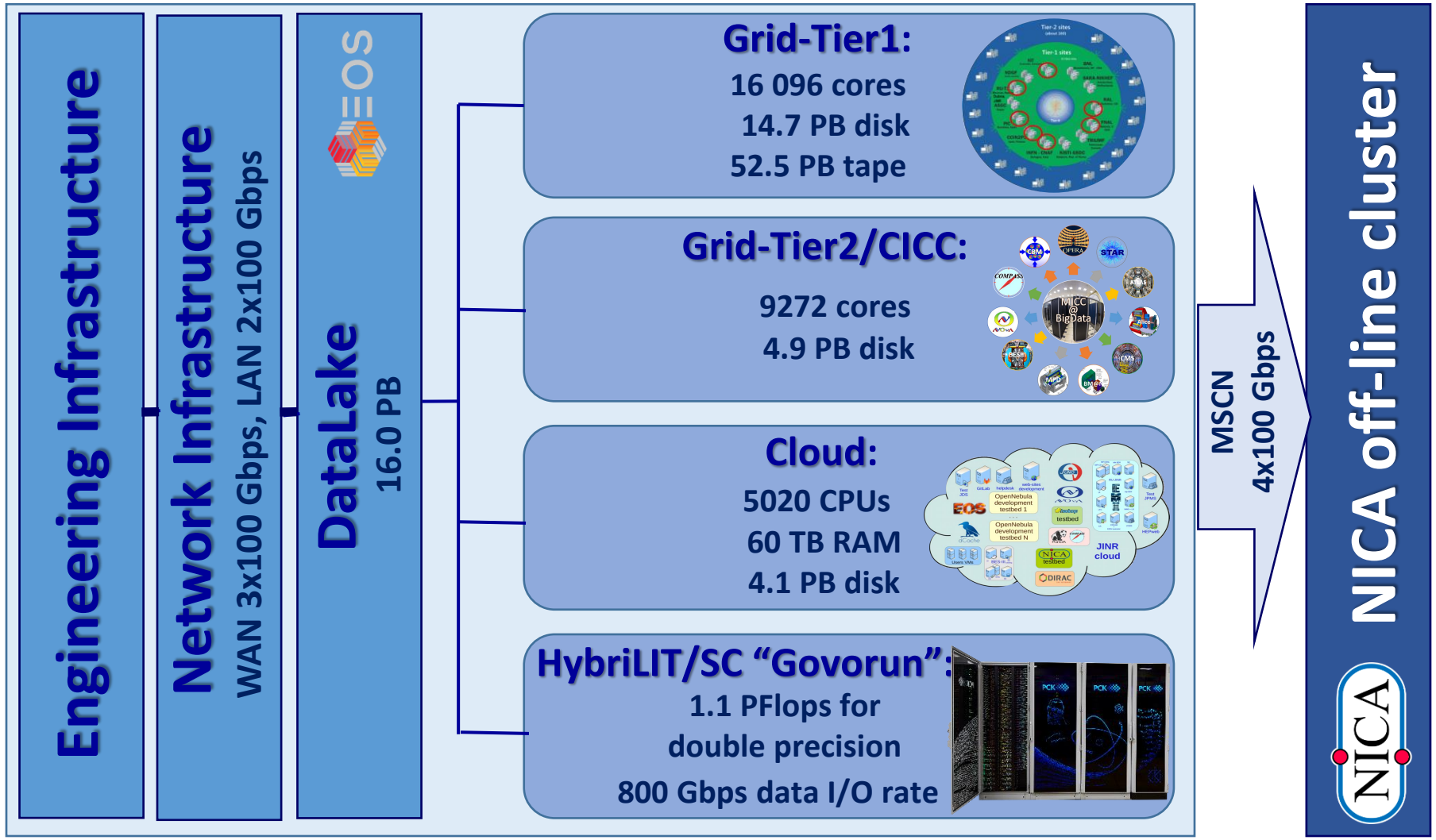
504 remote VPN and EDUROAM users

network traffic in 2021

- **33.23 PB - input**
- **35.86 PB - output**



Multifunctional Information and Computing Complex at JINR



The **MICC** meets the requirements for a modern highly performant scientific 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.

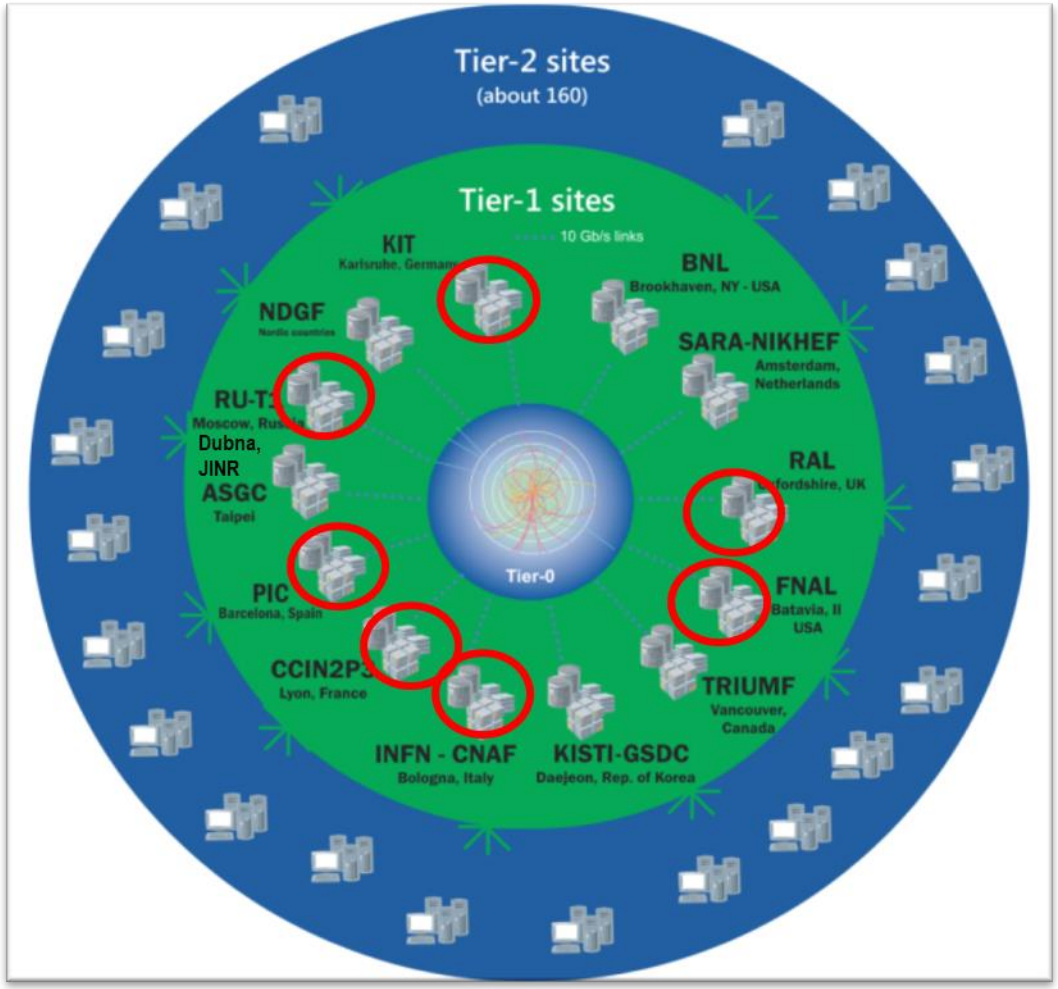
The Worldwide LHC Computing Grid

The mission of the **WLCG** is to provide global computing resources for the storage, distribution and analysis of the data generated by the LHC. **Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists**

Tier0 (CERN): data recording, reconstruction and distribution

Tier1: permanent storage, re-processing, analysis

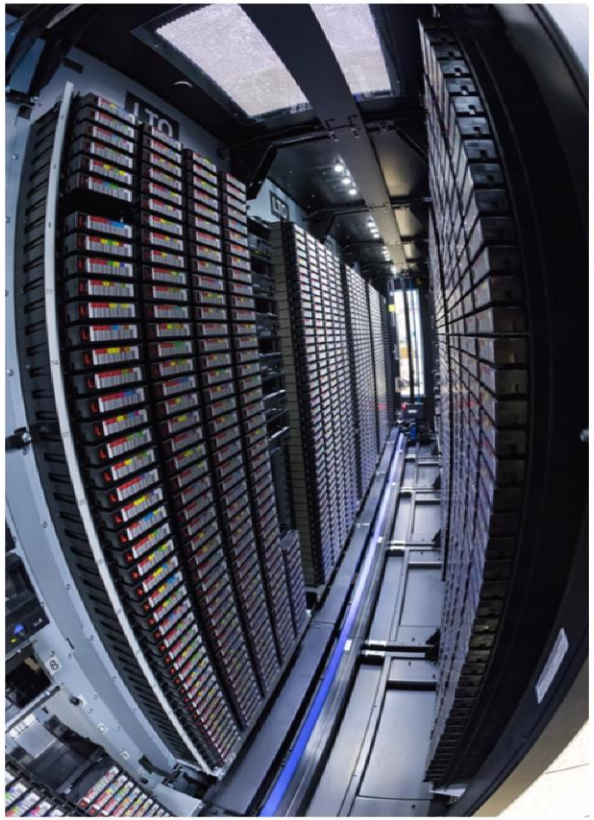
Tier2: Simulation, end-user analysis



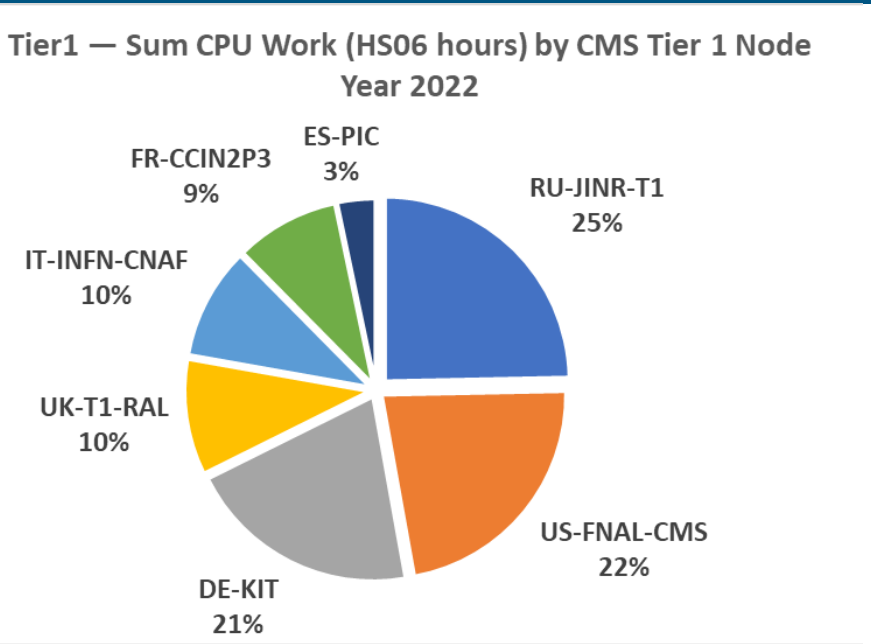
WLCG computing enabled physicists to announce the discovery of the Higgs Boson on 4 July 2012.

170 sites
42 countries
> 12k physicists
~1.4 M CPU cores
1.5 EB of storage
> 2 million jobs/day
100-250 Gb/s links

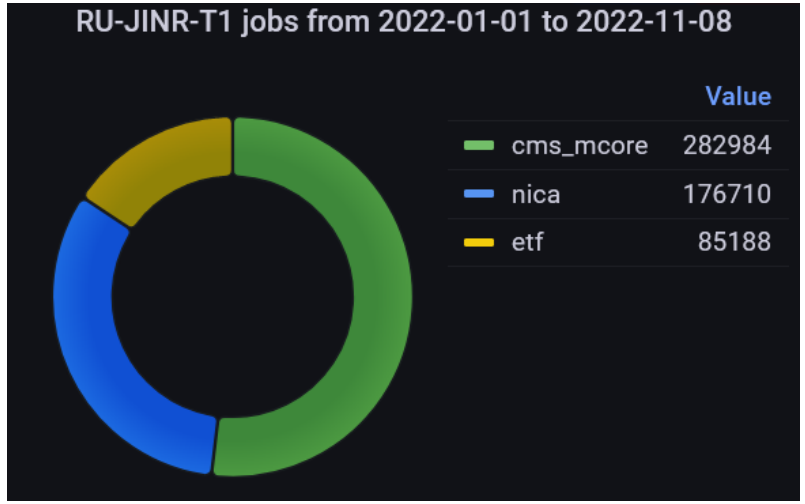
JINR Tier1 for CMS



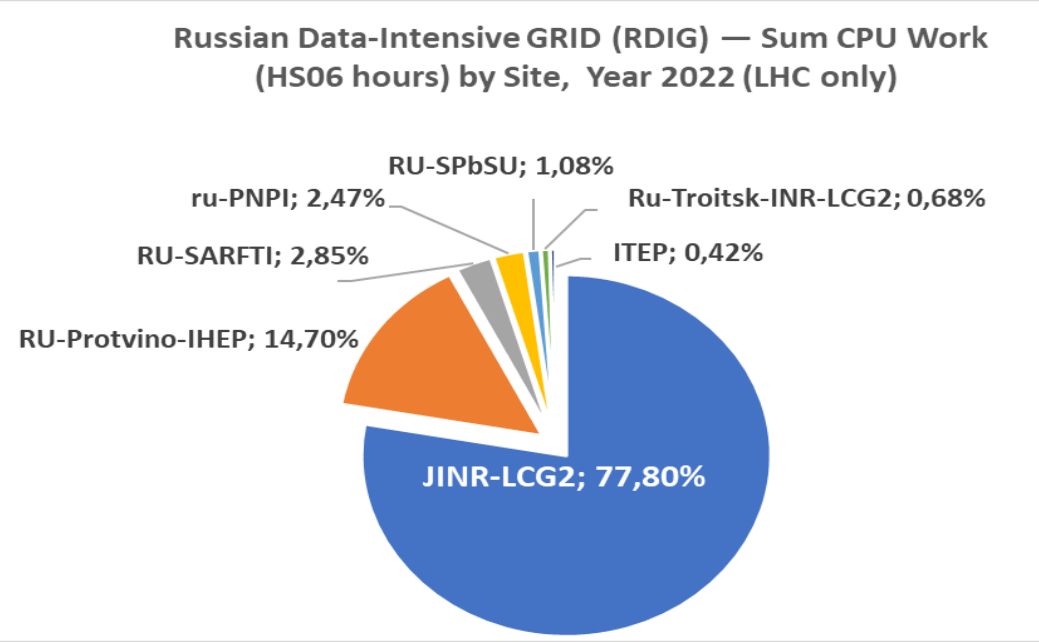
- 16096 cores
- 260 kHS06
- 14 PB disk
- 50.6 PB tapes
- 100% reliability and availability

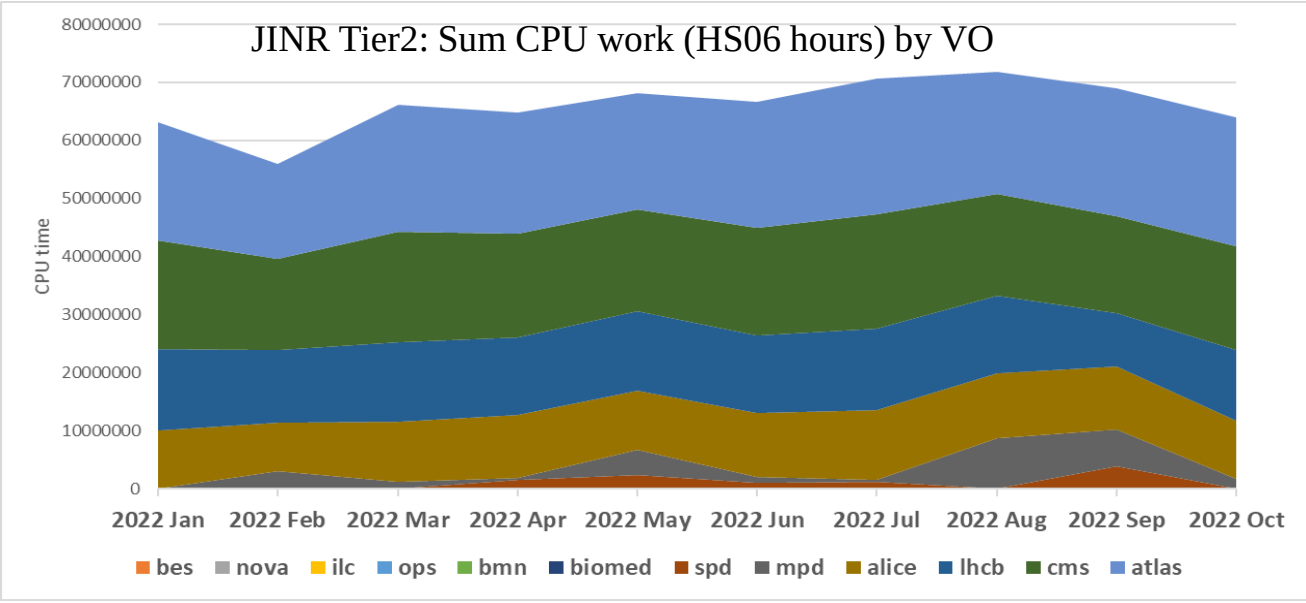


- The Tier1 center at JINR has demonstrated stable work through the entire period since its launch into full operation not only for CMS (LHC), but also for MPD (NICA).
- The Tier1 site for CMS is ranked first among world centers for CMS.
- 30% of all jobs executed at Tier1 JINR are NICA MPD jobs



Tier2 for Experiments and JINR Laboratories

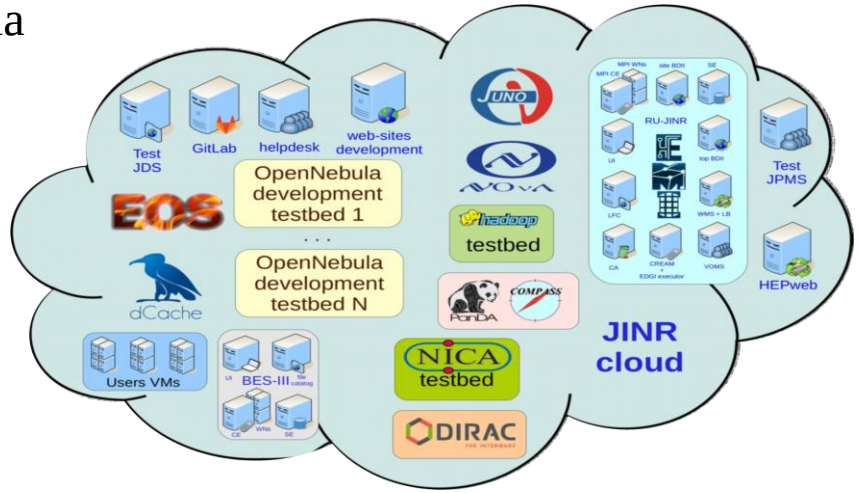




Tier2 for Alice, ATLAS, CMS, LHCb, BES, BIOMED, COMPASS, MPD, NOvA, STAR, ILC, etc. is recognized the best in the Russian Data Intensive Grid (RDIG) Federation.

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



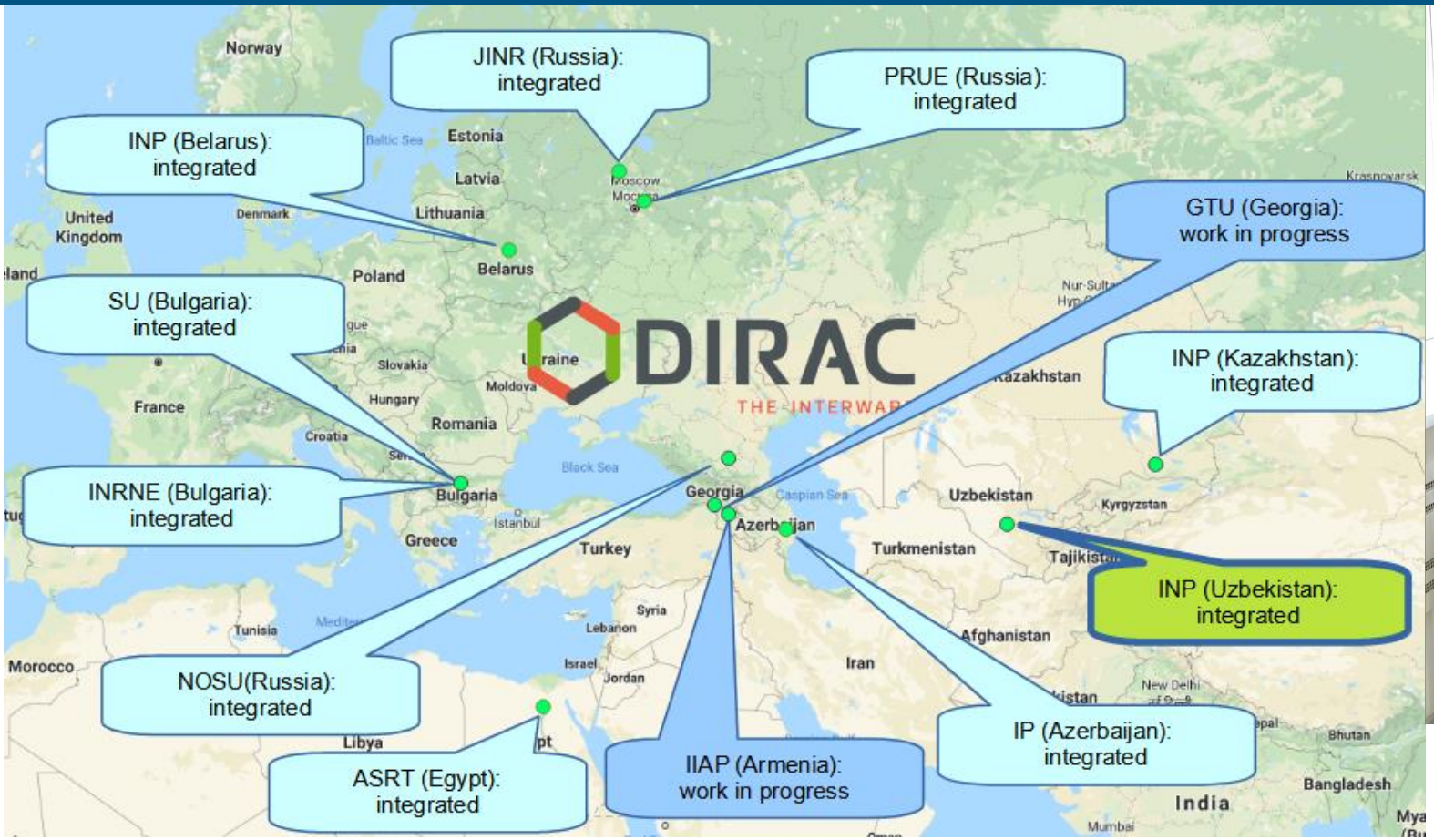
– VMs for JINR users,
– Computational resources for neutrino experiments,
– Testbeds for R&D in IT,
– COMPASS production system services,
– Data management system of the UNECE ICP Vegetation,
– Scientific and engineering computing,
– Service for data visualization based on Grafana, jupyterhub head and execute nodes for it,
– Gitlab and its runners, as well as some others.

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

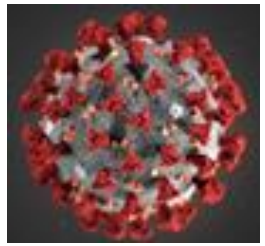
MLIT contribution:
engineering infrastructure (electricity, UPS, cooling, network, racks, manpower)

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

Clouds integration



The goal of the integration of the JINR member State organizations' clouds into the DIRAC-based distributed information and computing environment (DICE) is to join resources for solving common tasks, as well as to distribute a peak load across resources of partner organizations.



Idle resources of the DICE JINR were involved in research on the SARS-CoV-2 virus within the Folding@Home project, which uses distributed computing to perform computer modelling of protein molecule coagulation.

Development of the heterogeneous HybriLIT platform

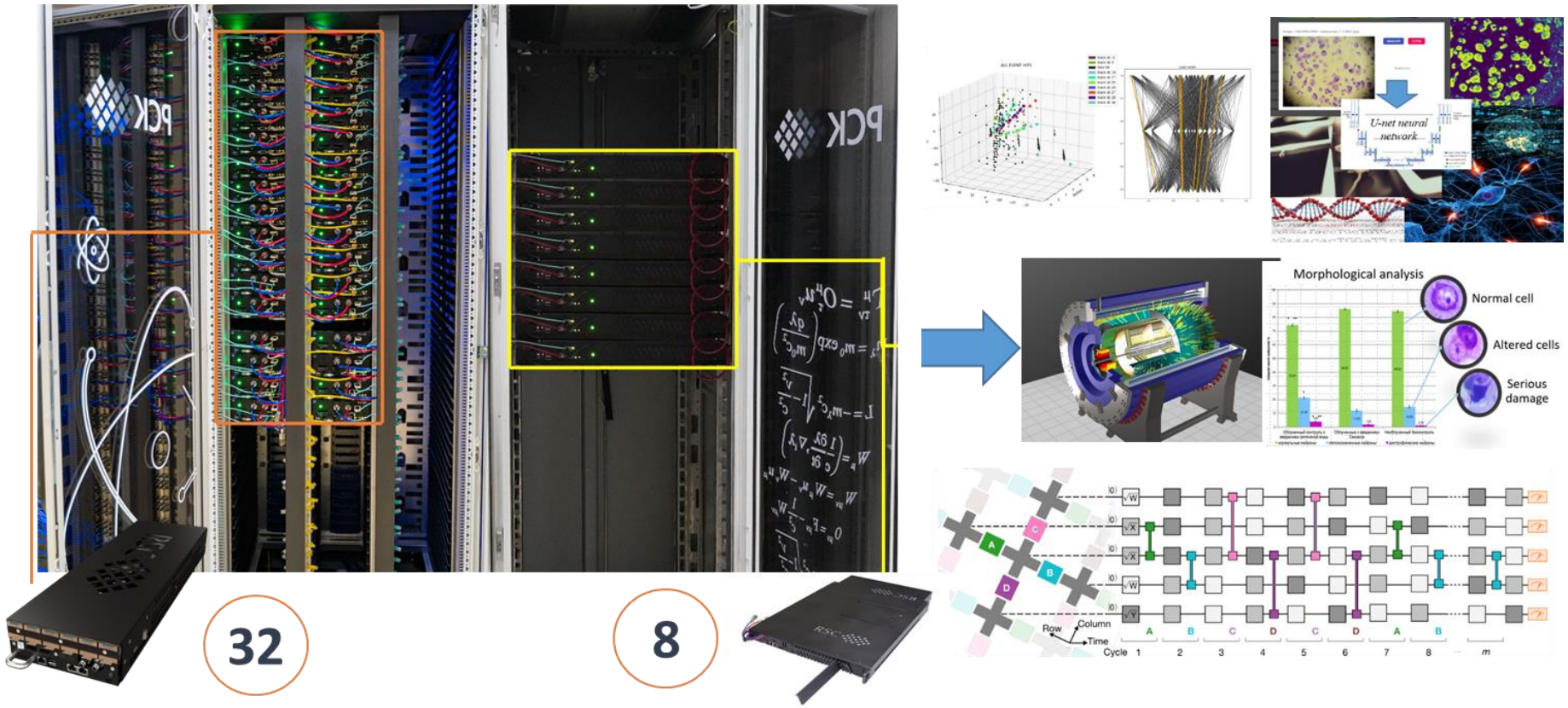
Cluster HybriLIT 2014:
Full peak performance:
50 TFlops for double precision

#18 в Top50
"Govorun" supercomputer
First stage 2018:
Full peak performance :
500 TFlops for double precision
9th in the current edition of the IO500 list (July 2018)

#10 в Top50
"Govorun" supercomputer
Second stage 2019:
Full peak performance :
860 TFlops for double precision
288 ТВ ССХД with I/O speed >300 Gb/s
17th in the current edition of the IO500 list (July 2020)

Russian DC Awards 2020 in "The Best IT Solution for Data Centers"

"Govorun" supercomputer modernization 2022



The expansion of the "Govorun" supercomputer by 32 hyperconverged compute nodes and 8 distributed storage nodes made it possible to:

- enhance its performance by 239 Tflops (Total peak performance: 1.1 PFlops DP);
- increase the DAOS data processing and storage subsystem to 1.6 PB;
- enlarge the volume of the "warm data" storage subsystem by 8 PB with support for the creation of dynamic storage systems such as Luster, DAOS, EOS, dCache, NFS.

The continuous increase in the expansion of the range of tasks to be solved in order to ensure the solution of the theoretical and experimental tasks of JINR has required

- The constant re-equipment of the "Govorun" SC with computing resources;
- Permanent implementation of the novel IT solutions;
- Creating an environment for supercomputer modeling and the solution of compute-intensive and data-intensive tasks.

“Govorun” Supercomputer Today

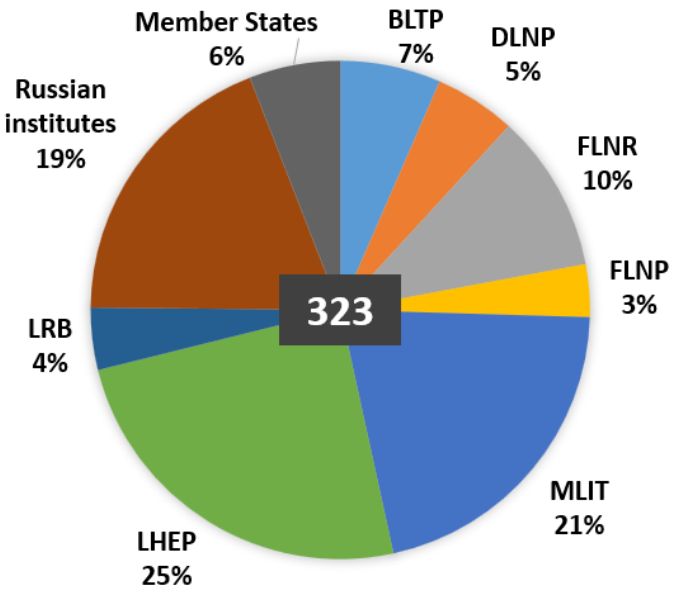
GPU-component
based on NVIDIA DGX-1 Volta

CPU-component
based on the newest Intel architectures:
Intel Xeon Phi gen.2
Intel Xeon gen.2 and gen.3

- Hyper-converged software-defined system
- Hierarchical data processing and storage system
- Scalable solution Storage-on-demand
- Total peak performance: 1.1 PFlops DP
- GPU component based on the NVIDIA
- CPU component based on RSC “Tornado” liquid cooling solutions
- The most energy-efficient center in Russia (PUE = 1,06)
- Storage performance >300 GB/s

“Govorun” supercomputer

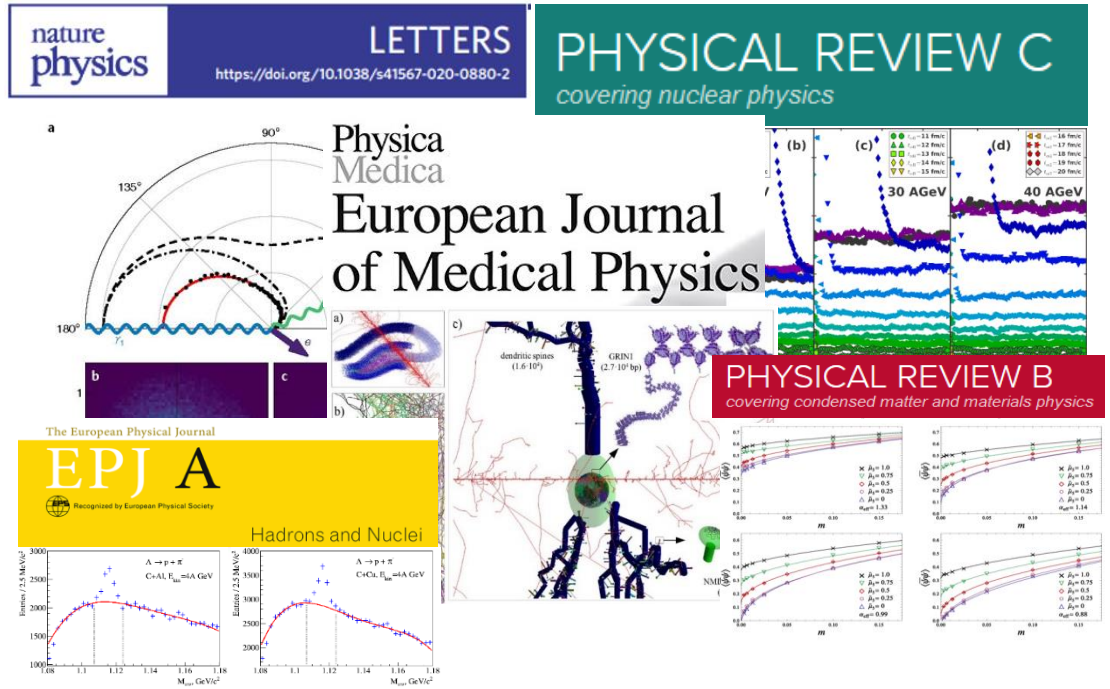
Total number of users : 323



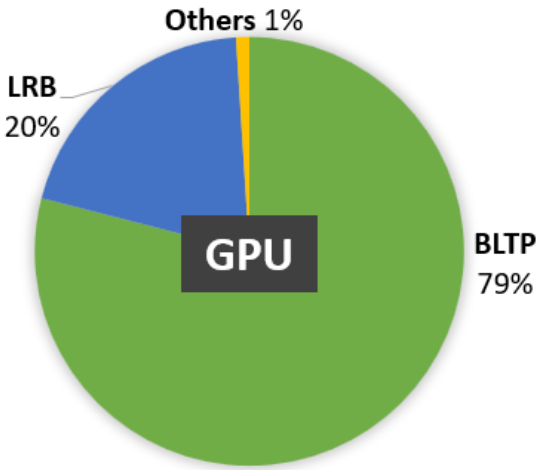
Key projects that use the resources of the “Govorun” supercomputer:

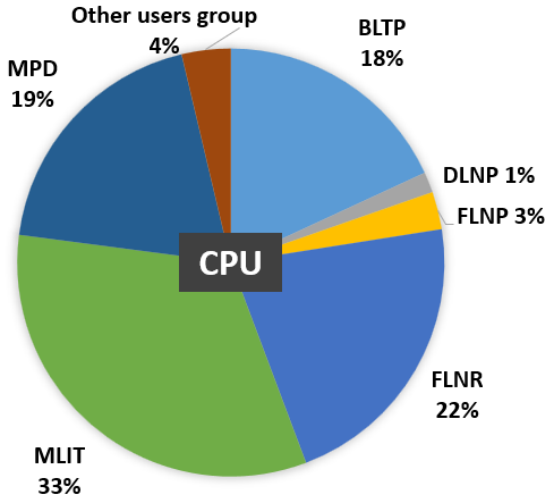
- NICA megaproject,
- calculations of lattice quantum chromodynamics,
- computations of the properties of atoms of superheavy elements,
- studies in the field of radiation biology,
- calculations of the radiation safety of JINR’s facilities.

The results obtained using the resources of the “Govorun” SC since its launch into operation are reflected in 204 user publications.



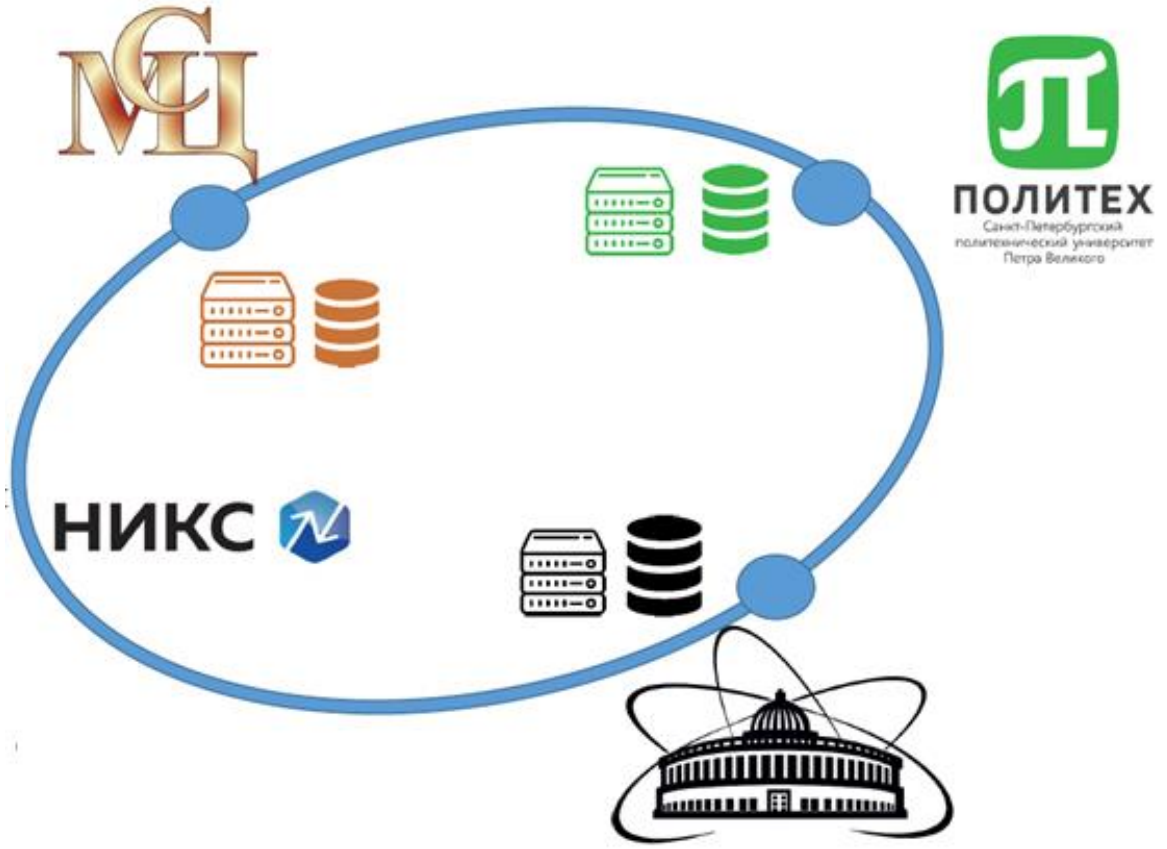
Distribution of the GPU and CPU resources by user group





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 scalable research infrastructure of a new level was created.
Such an infrastructure is in demand for the tasks of the NICA megascience project.



In January 2022, the first joint experiment on the use of the unified supercomputer infrastructure for the tasks of the NICA megascience project was successfully completed:

- ✓ 3,000 Monte-Carlo data generation and event reconstruction jobs were launched for the MPD experiment
- ✓ about 3 million events were generated and reconstructed
- ✓ the obtained data were transferred to Dubna for further processing and physics analysis.

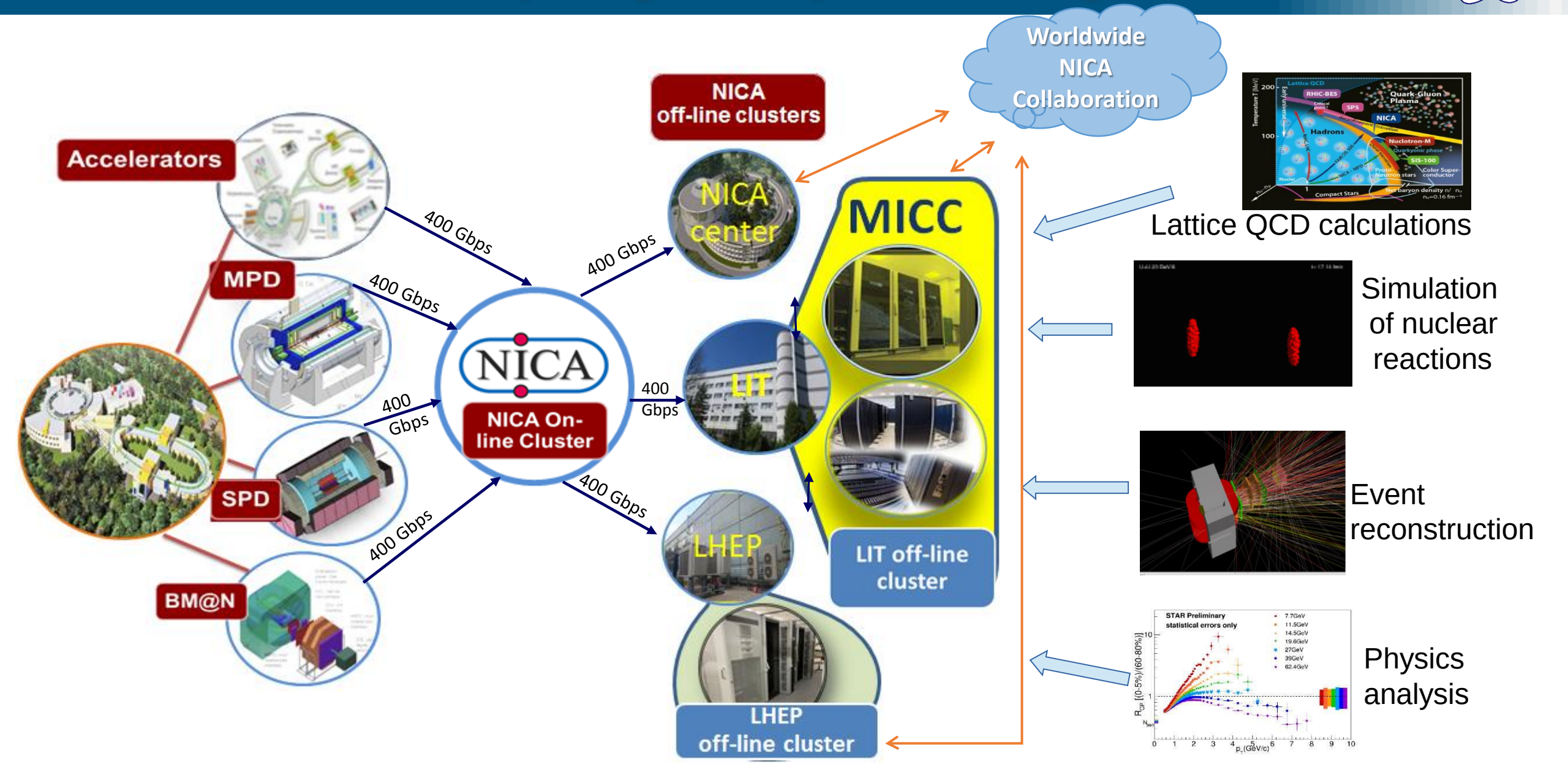
MICC Monitoring

For a robust performance of the complex it is necessary to monitor the state of all nodes and services - from the supply system to the robotized tape library.

- Global real time 24x7 survey of the state of the whole computing complex
- In case of emergency, alerts are sent to users via e-mail, SMS, etc.
- ~ 850 elements are under observation

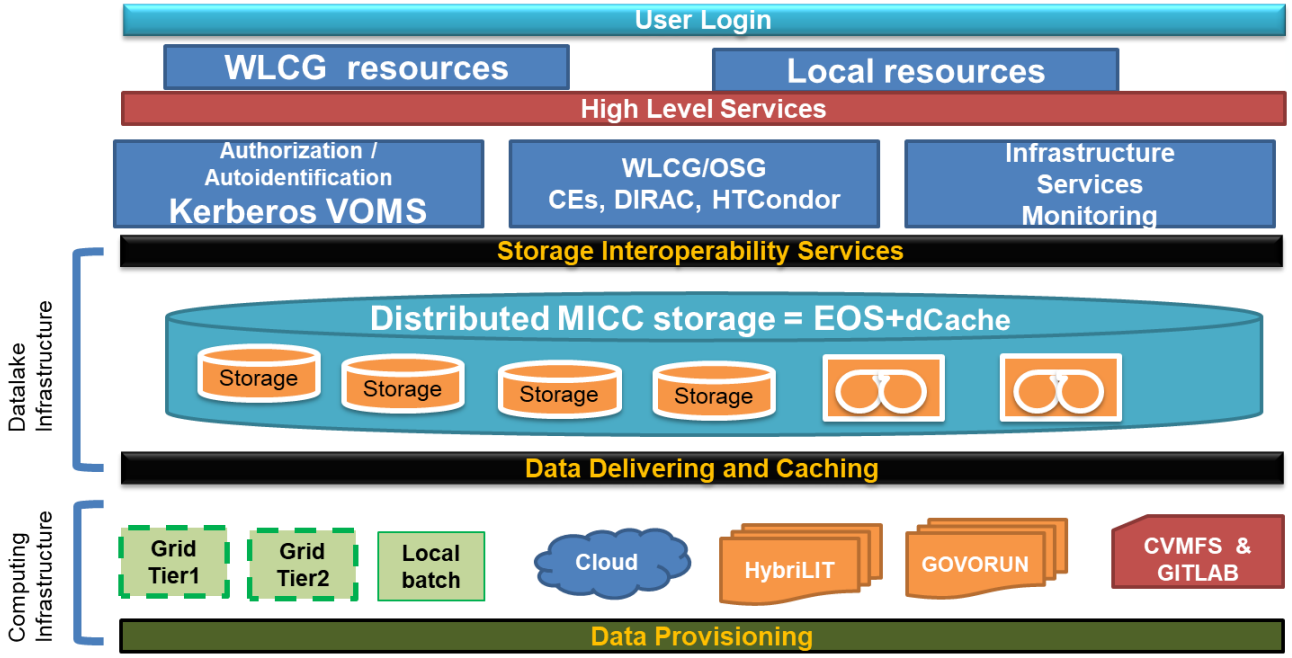

MICC Operational Center

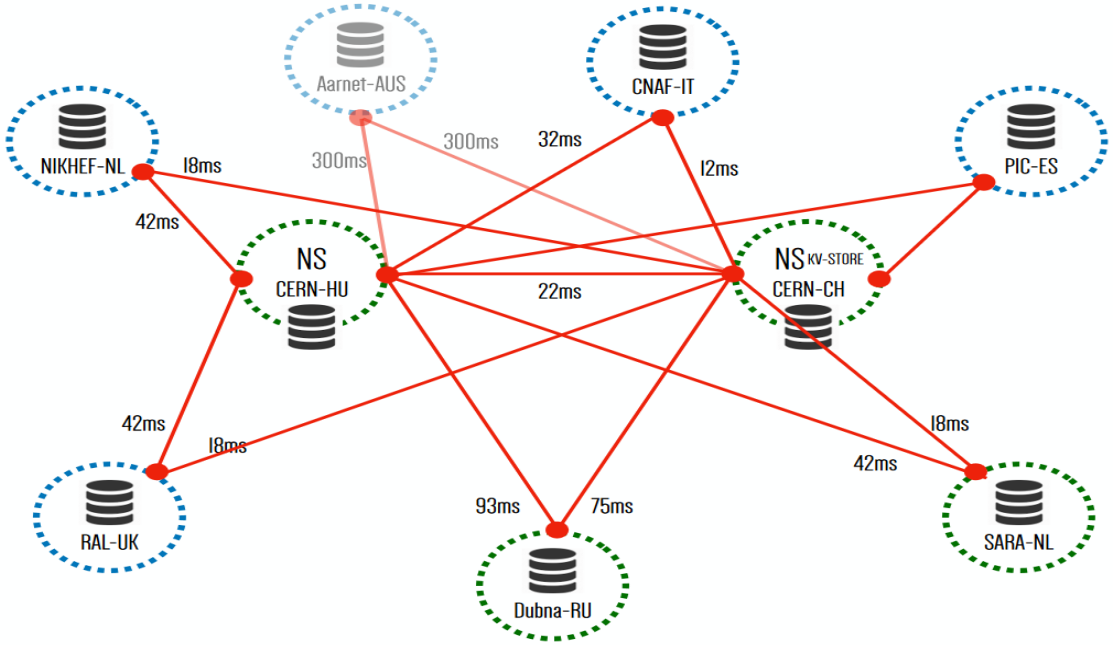
NICA Computing Concept & Challenges



JINR in DataLakes

- The JINR data lake prototype was built as a distributed EOS storage system.
- EOS was successfully integrated into the MICC structure.
- 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.

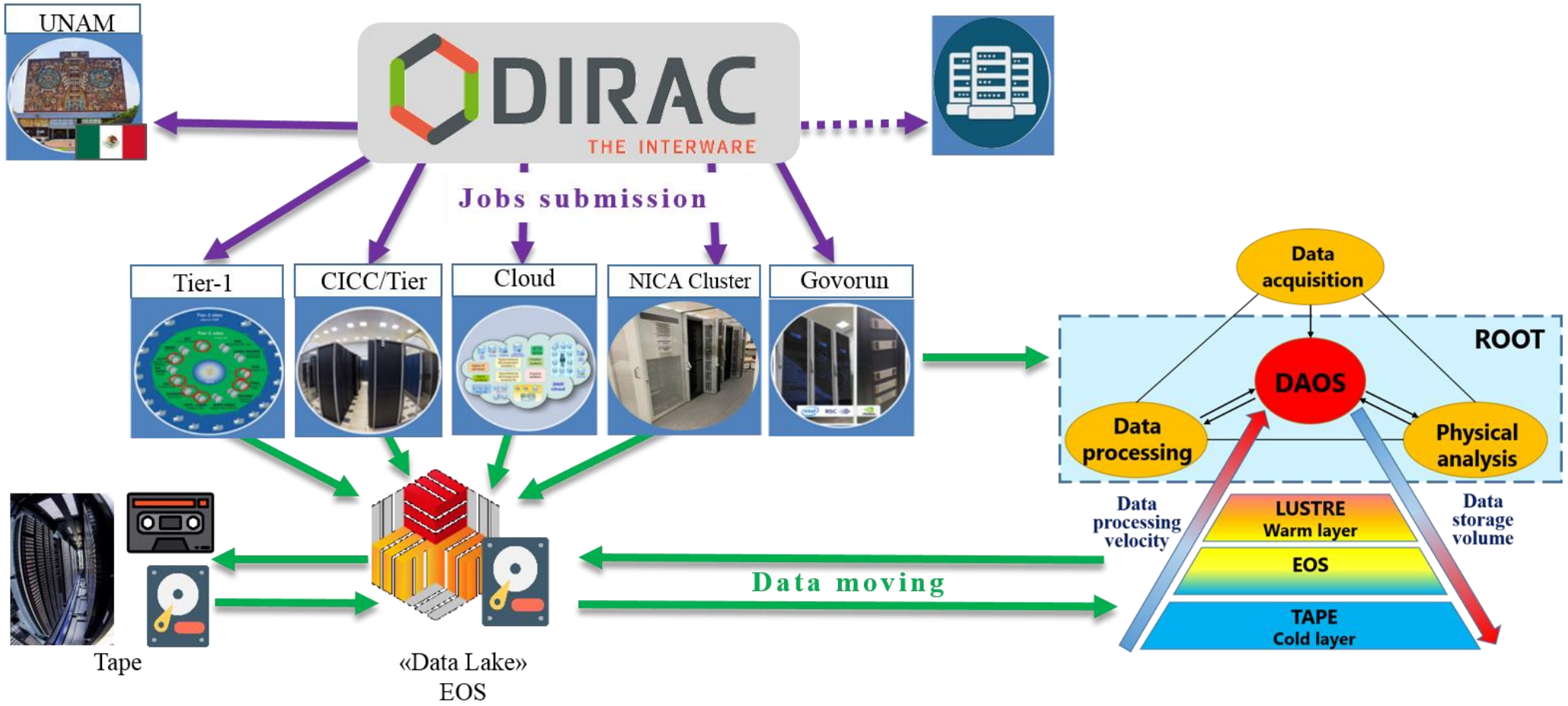




- There is currently 17 PB of disk space available for EOS.
- Baikal-GVD, DANSS, FOBOS, JUNO, BM@N, MPD, SPD, PANDA are its major users.
- EOS is visible as a local file system on the MICC working nodes and allows authorized users (by the kerberos5 protocol) to read out and record data.

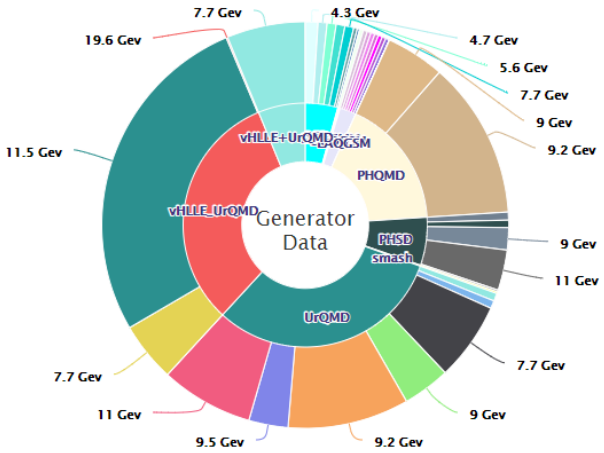
Heterogeneous distributed computing environment

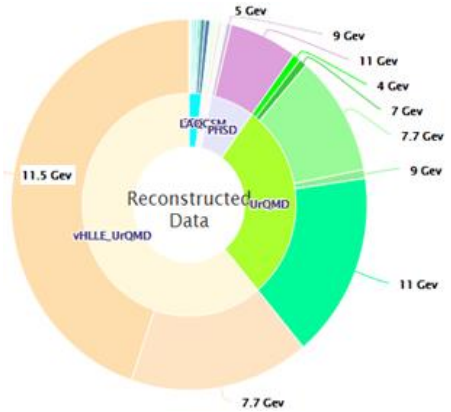
A heterogeneous computing environment, based on the DIRAC platform, was created for processing and storing data of the experiments conducted at JINR. By the end of 2021, Tier1, Tier2, the "Govorun" supercomputer, the clouds of the JINR Member States, the NICA cluster, as well as the cluster of the National Autonomous University of Mexico (NAUM, within the cooperation on the MPD project), were integrated into DIRAC. For the time being, the distributed infrastructure is used by the following experiments: MPD, Baikal-GVD, BM@N, SPD.



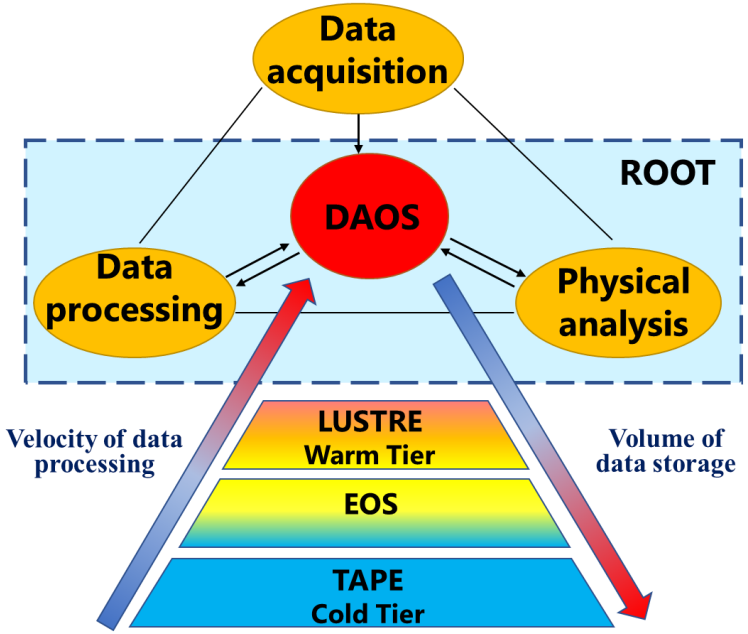
Heterogeneous distributed computing environment for the MPD experiment

✓ More than 1.1 billion events were generated using *UrQMD*, *LAQGSM*, *PHSD* and other models





✓ 270 millions events were reconstructed



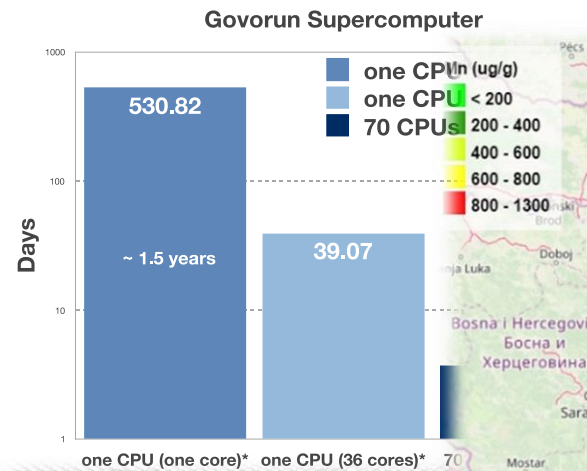
Available resources of the heterogeneous computing environment, based on the DIRAC platform, for the MPD experiment (about 5,000 cores):

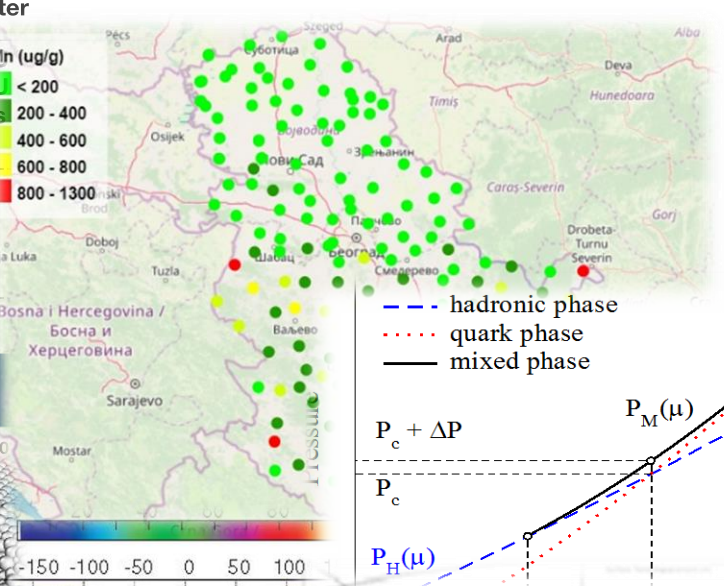
- "Govorun" SC: up to 1,586 cores in the latest production
- Tier1: 920 cores
- Tier2: 1,000 cores
- Clouds (JINR and JINR Member States): 70 cores
- NICA offline cluster: 250 cores (limit for users)
- UNAM (Mexico University): 100 cores
- National Research Computer Network of Russia (NIKS, now resources from SPBTU and JSCC): 672 cores – New resource, added in 12.2021.

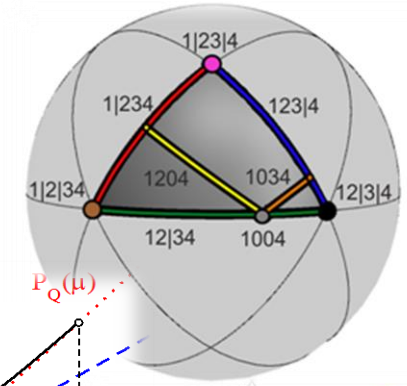
Mass production storages integrated into the Dirac File Catalog are 1.5 PB in size.

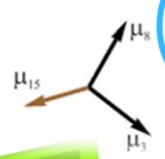
To work with Big Data, including for the NICA megaproject, a hierarchical data processing and storage system with a software-defined architecture was developed and implemented on the "Govorun" supercomputer.

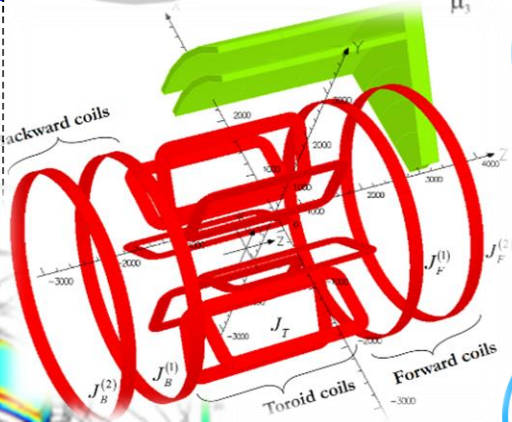
Methods, Algorithms and Software

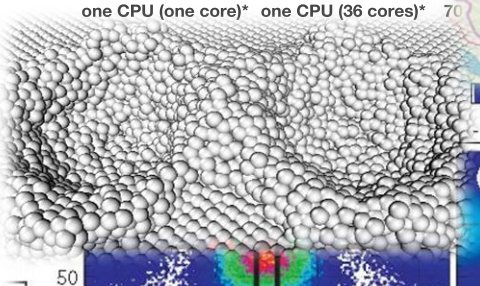


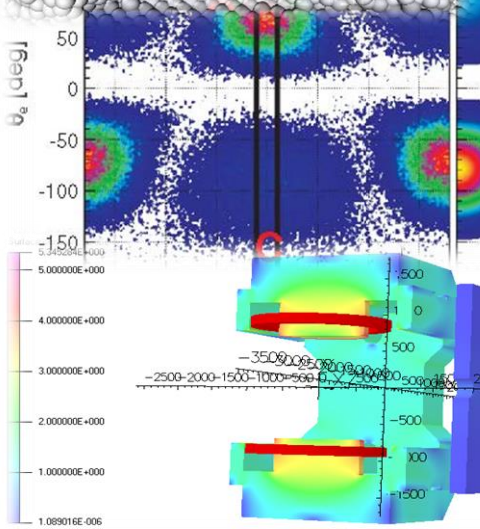


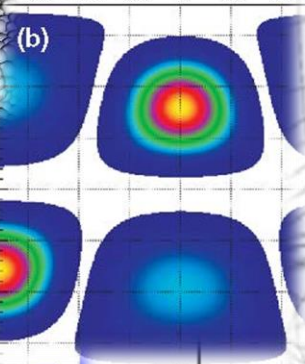


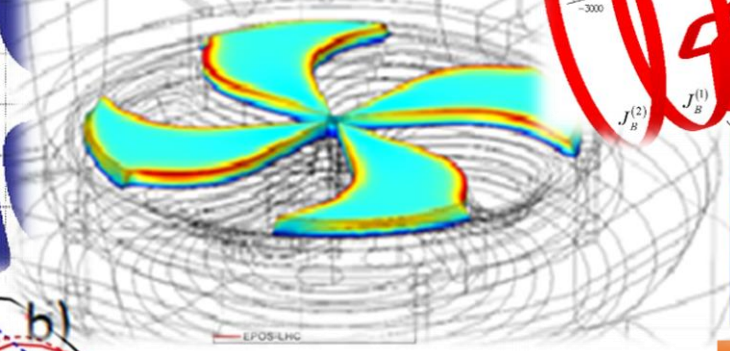


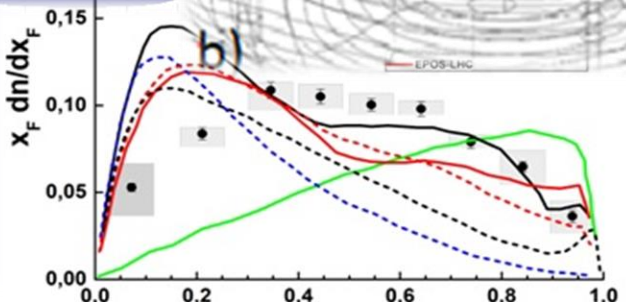


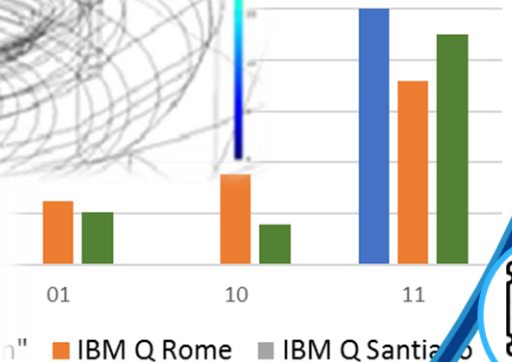








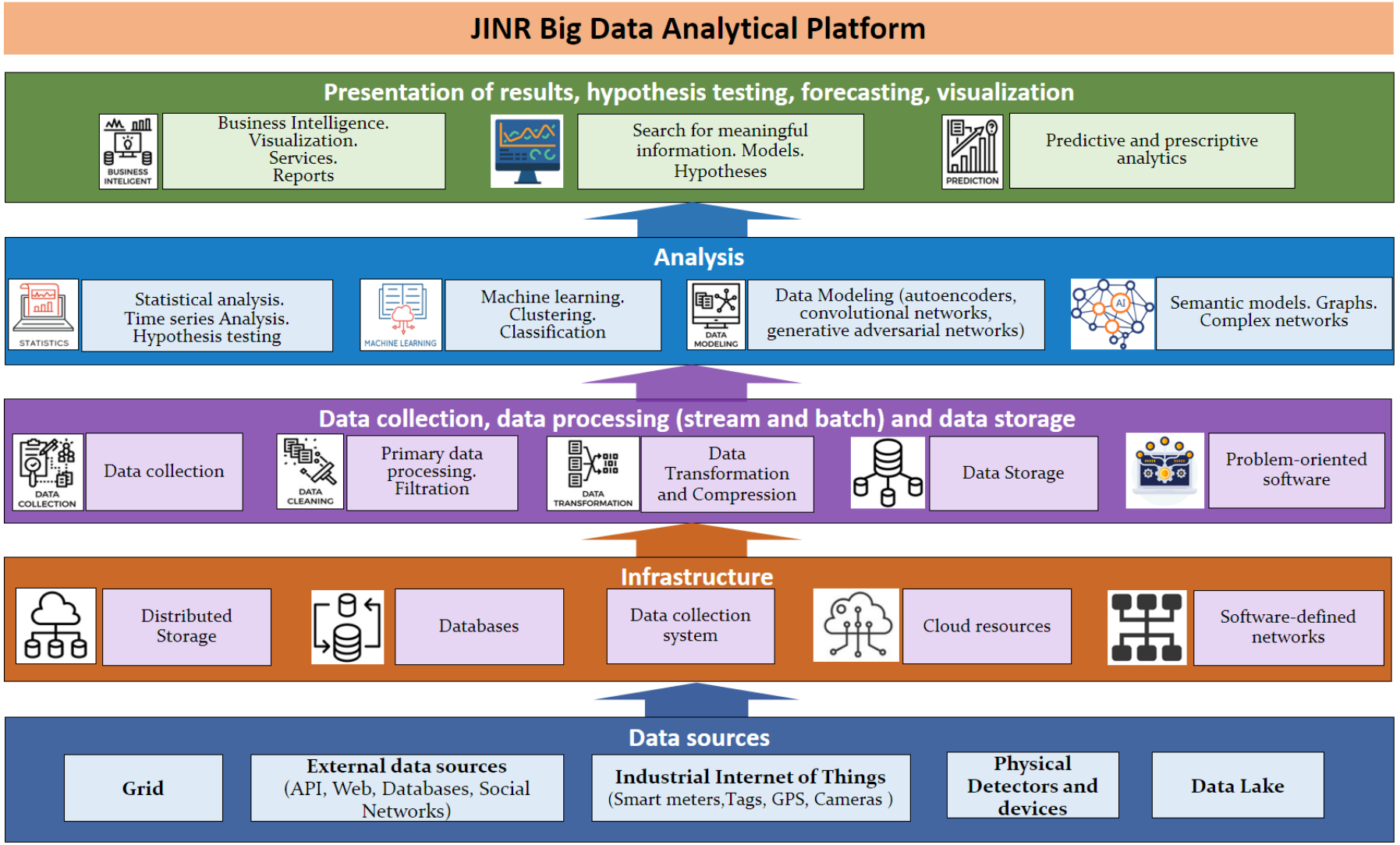




- Numerical modeling of complex physical systems
- Experimental data processing and analysis
- Big Data
- Machine and Deep learning
- AI and robotics
- Computer algebra
- Quantum computing

JINR Big Data Analytical Platform

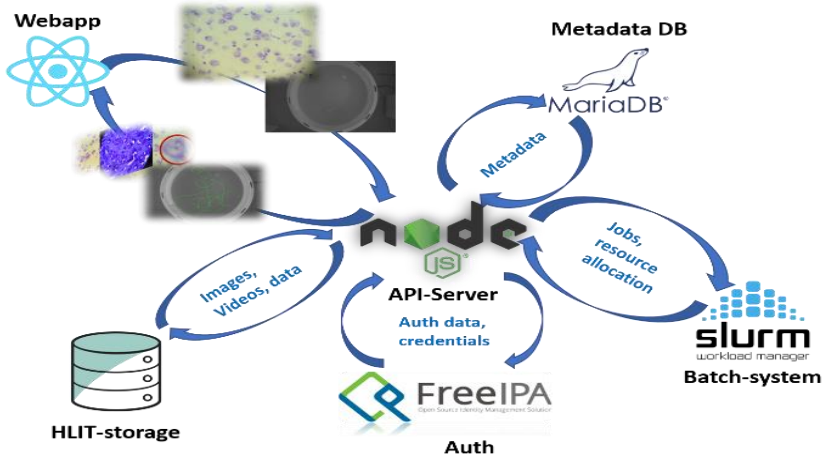
- Bringing best of Big Data approaches to JINR practices
- Providing the Big Data infrastructure for users



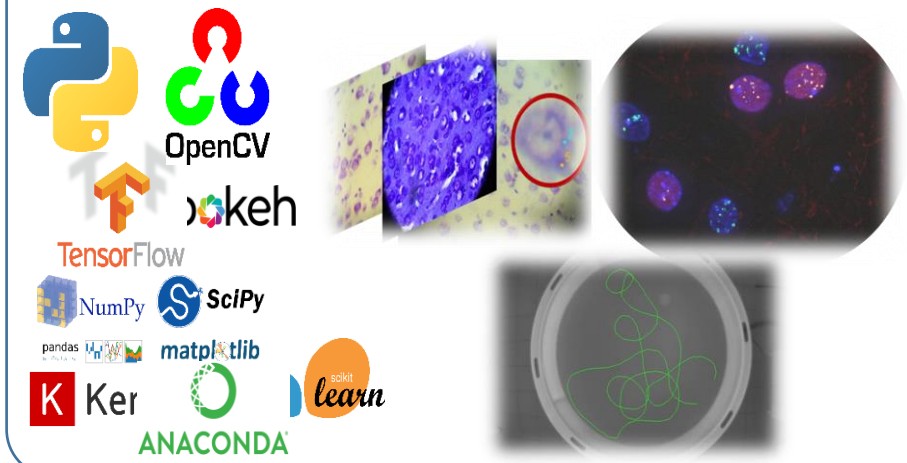
Information System for the tasks of Radiation Biology
https://bio.jinr.ru

The joint project of LIT and LRB is focused on creating an Information System (IS). The IS allows one to store, quickly access and process data using a stack of neural network and classical algorithms of computer vision, providing a wide range of possibilities for automating routine tasks. It gives an increase in productivity, quality and speed of obtaining results.

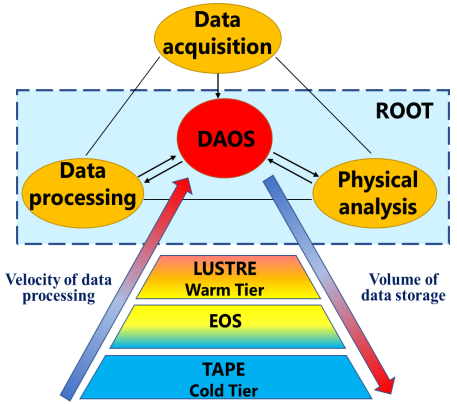
Web services for morphofunctional changes in the CNS and DNA damage analysis



Machine and deep learning methods, computer vision and neural network approaches



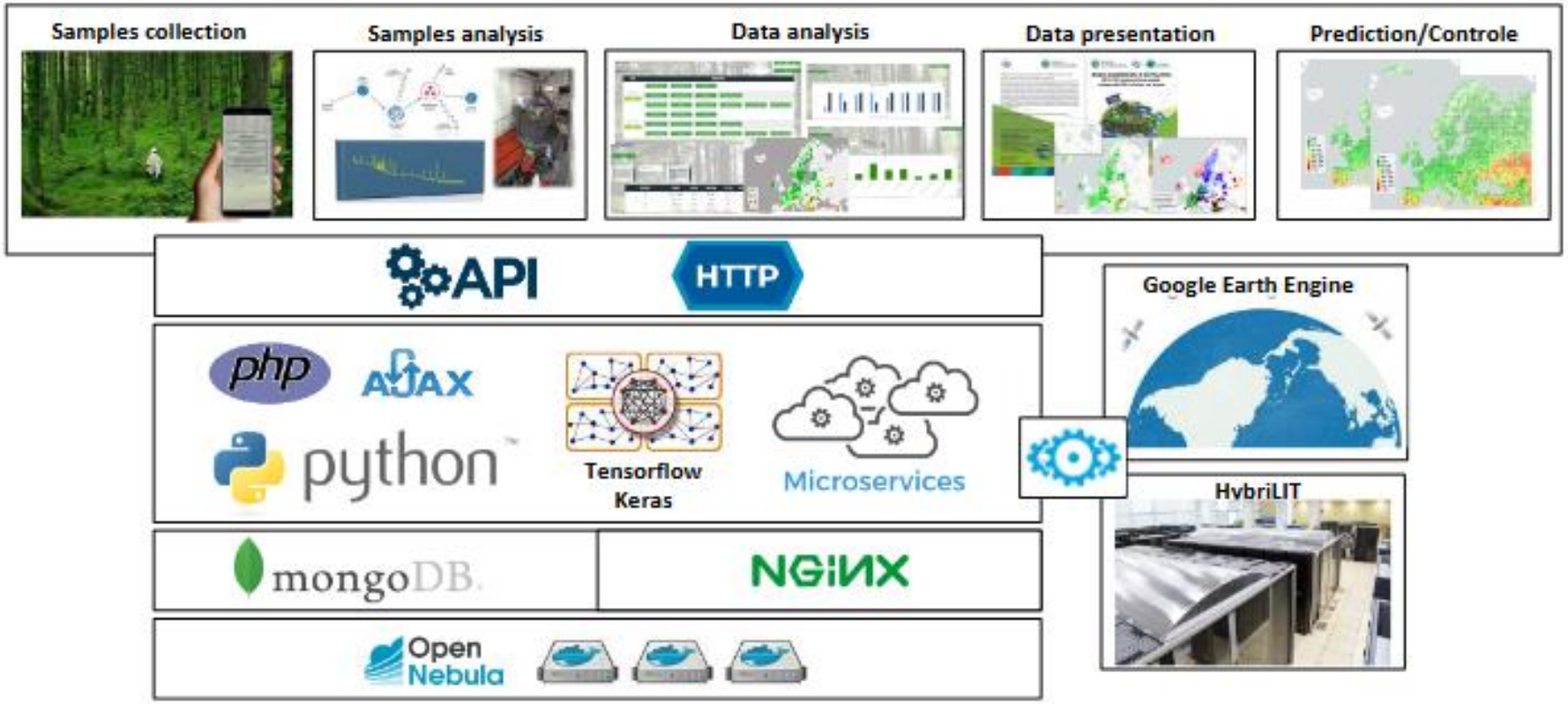
Computing resources and hierarchical storage system based on the “Govorun” SC



Intelligent Environmental Monitoring Platform

Within the framework of cooperation between MLIT and FLNP, the work on the prediction of air pollution by heavy metals using biomonitoring data, satellite imagery and different IT technologies is in progress.

On the MLIT cloud platform, a Data Management System (DMS) of the UNECE ICP Vegetation was created. DMS is intended to provide its participants with a modern unified system of collecting, analyzing and processing biological monitoring data**.** A combination of satellite data, biomonitoring measurements, and different machine and deep learning technologies was used to predict potentially toxic elements.

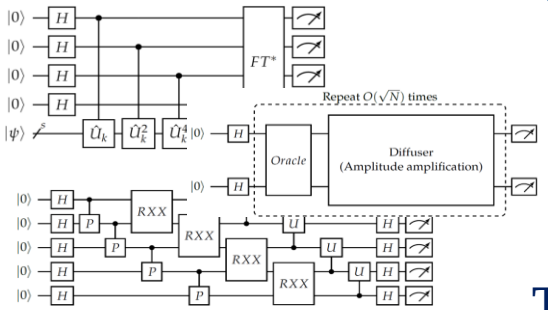


A. Uzhinskiy, M. Aničić Urošević, M. Frontasyeva. Prediction of air pollution by potentially toxic elements over urban area by combining satellite imagery, Moss Biomonitoring Data and Machine Learning. Ciencia e Tecnica Vitivinicola Journal

Quantum computing and quantum algorithms

Objective: development of quantum algorithms (QAs) to calculate complex atomic and molecular systems, taking into account the limiting capabilities of available computing resources.

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 exact solutions to urgent problems of quantum chemistry and study the chemical properties of heavy elements

Current result

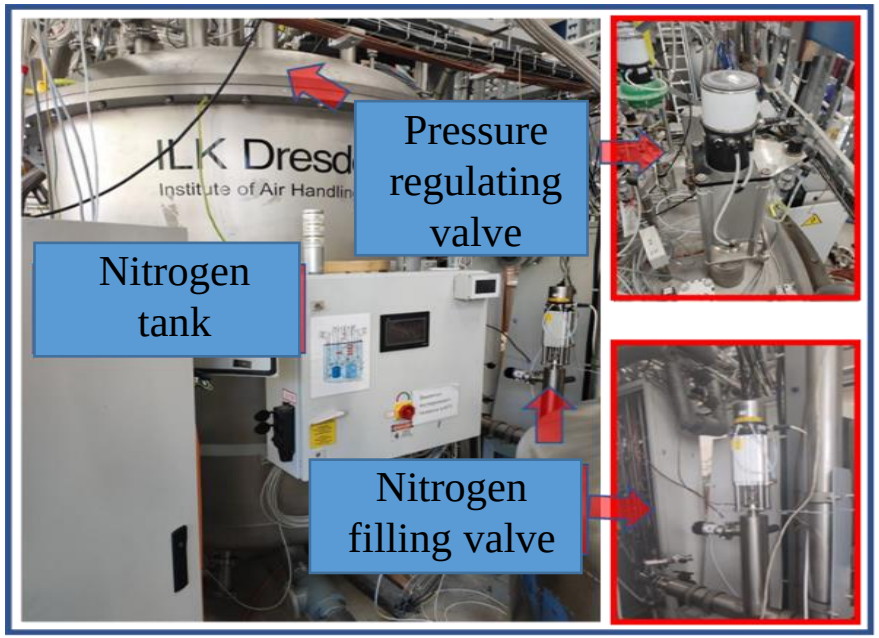
The limiting computing capacities of the "Govorun" supercomputer are revealed on the example of simulating quantum algorithms (quantum Fourier transform, quantum phase estimation, Grover's algorithm, test synthetic algorithm) using a different class of quantum circuits for the following simulators: QuEST, Qiskit, CuQuantum.

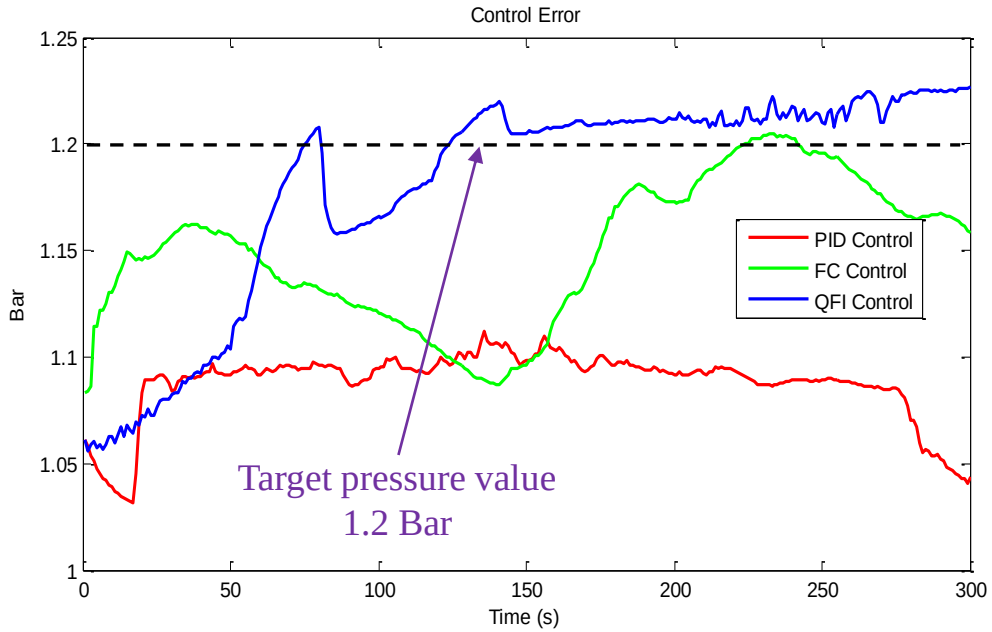
CPU	GPU
• 38 qubits	• 34 qubits

According to modern concepts, from 30 to 50 qubits are sufficient for the exact solution of most practically significant problems of quantum chemistry

Quantum intelligent control

Tests of an **intelligent automatic control system for the nitrogen collector** of the satellite helium refrigerator #1 at the site of the cryogenic testing of superconducting magnets at VBLHEP **on the basis of quantum algorithms** (QFI) are successfully completed.





Control of the process of reaching a predetermined pressure level in cooling mode

- The quantum controller (blue curve) is almost 5 times faster in reaching the target value than the closest controller on soft computing (green curve), while the PID-controller (red curve) does not reach the target value.
- The quantum controller demonstrates low overshoot and accuracy in achieving the control goal compared to other types of controllers.
- Automatic control based on the quantum controller reduces nitrogen consumption by 53%.

In the future, the system will be put into operation, and its regular operation will start.

JINR Digital EcoSystem

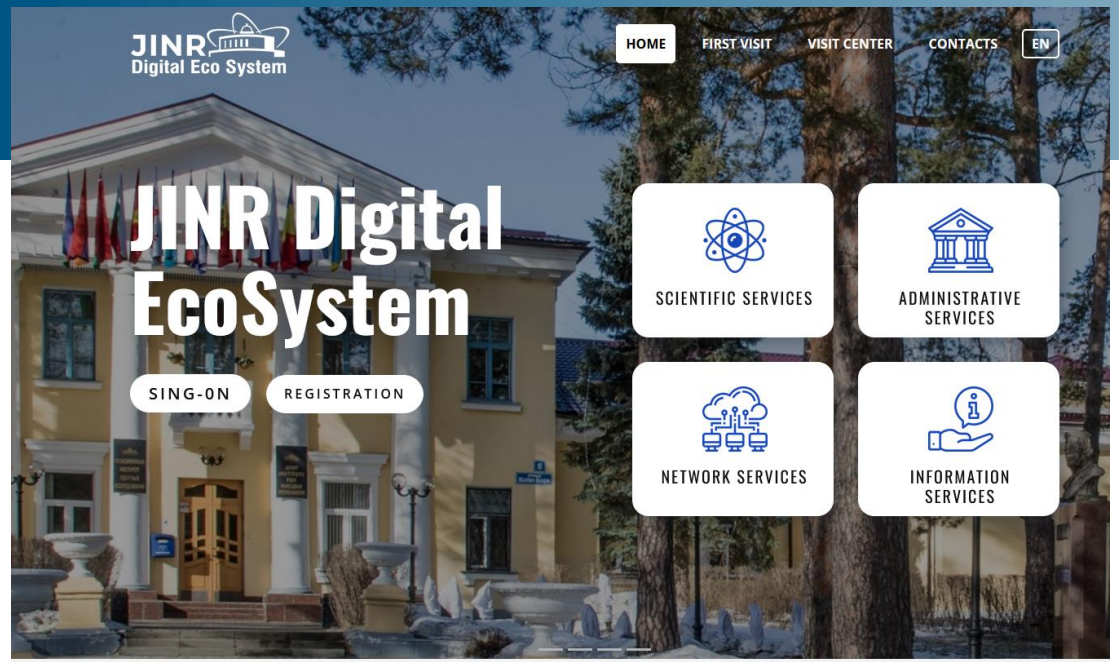
The digital platform “JINR Digital EcoSystem ” integrates existing and future services

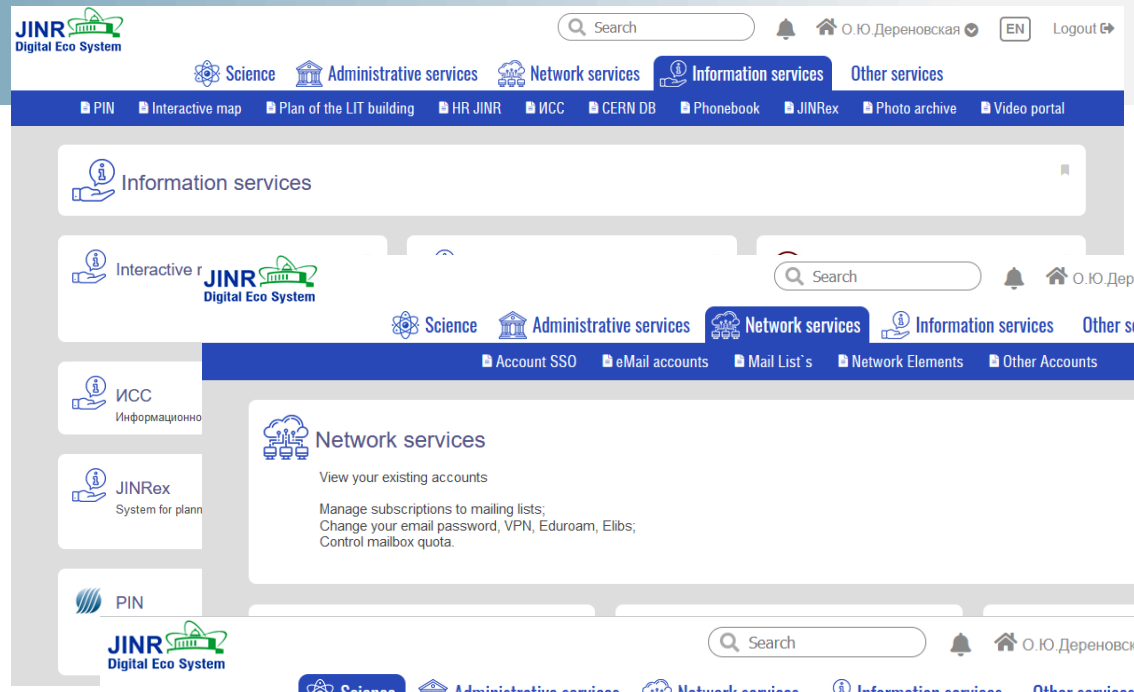
to support
scientific,
administrative and social activities,
maintenance of the engineering and IT infrastructures

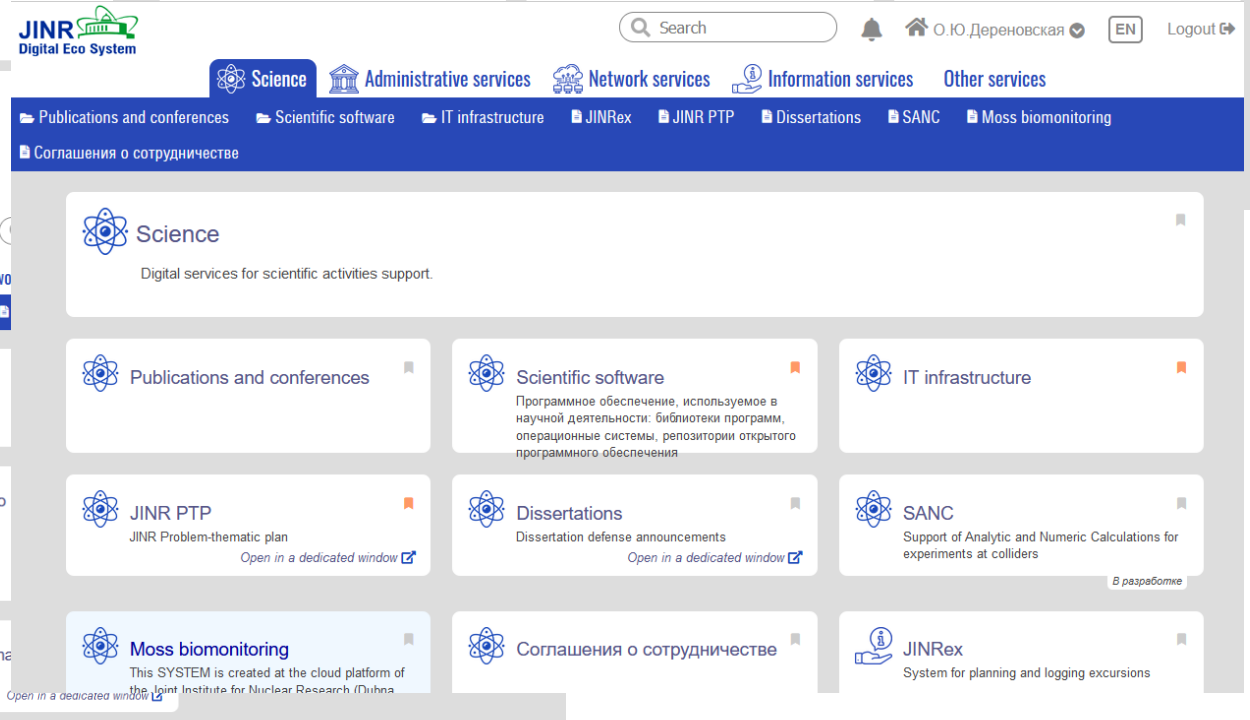
to provide
reliable and secure access to various types of data

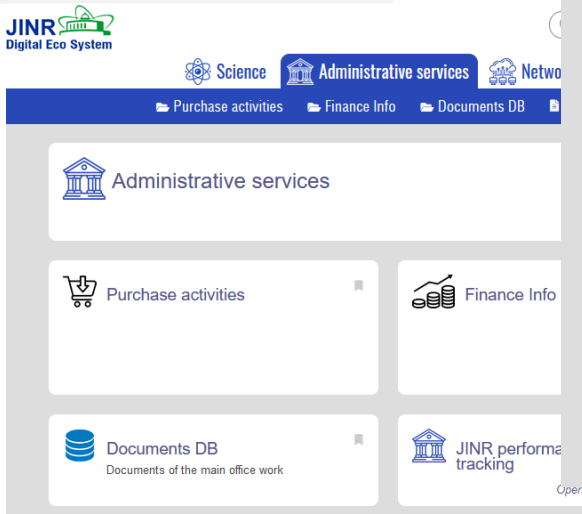
to enable
a comprehensive analysis of information

using
modern Big Data technologies and artificial intelligence.

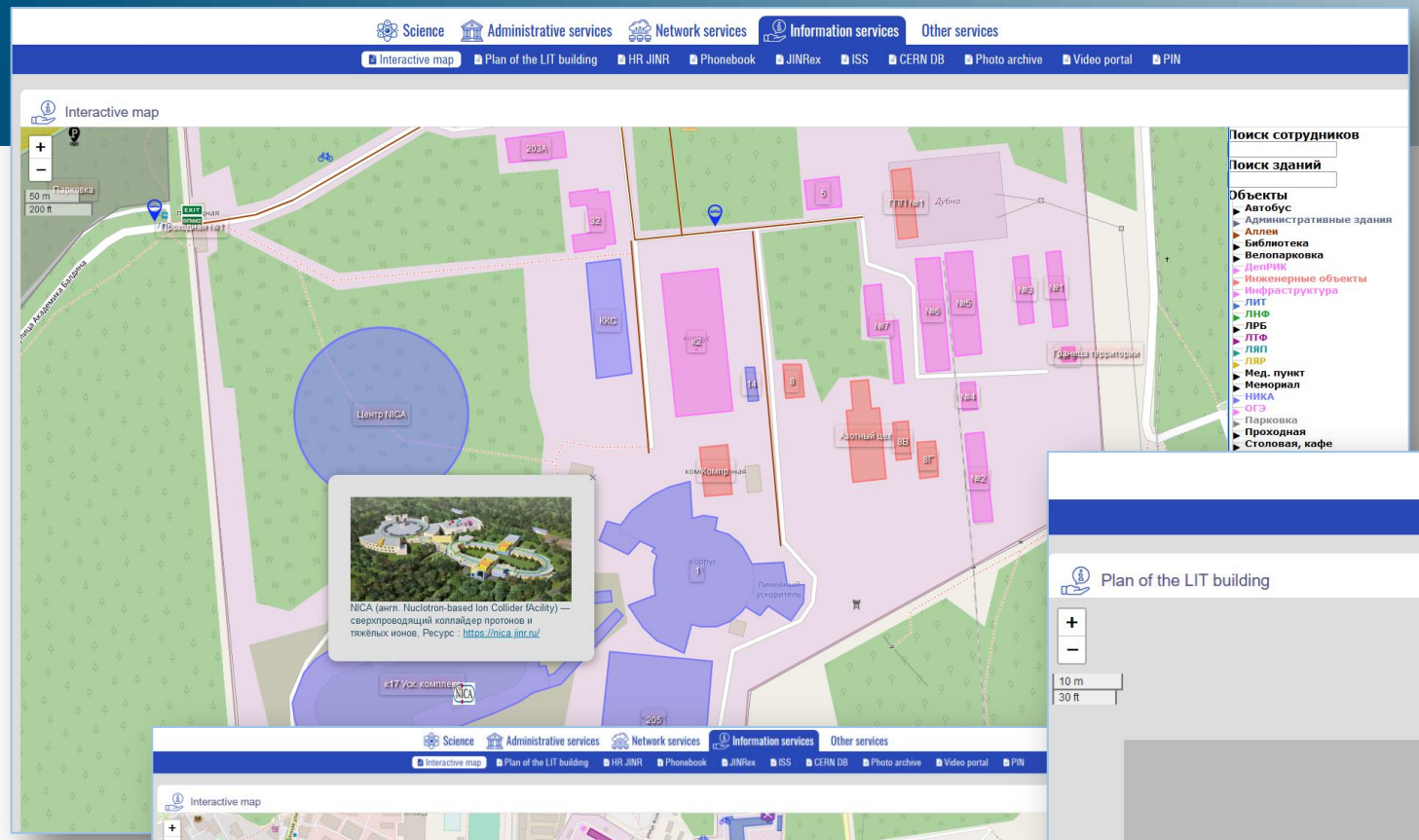


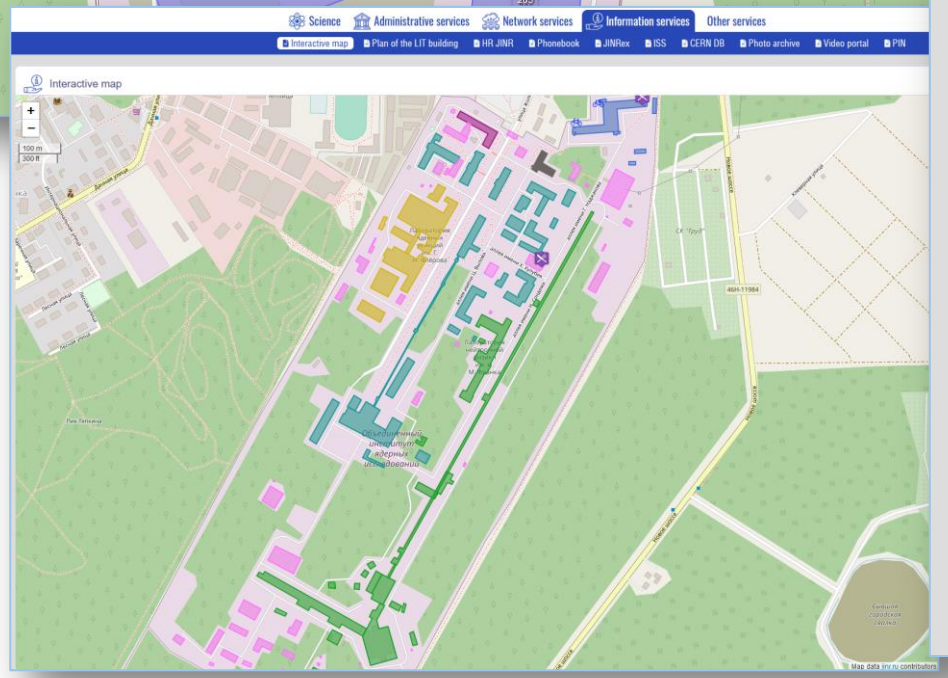


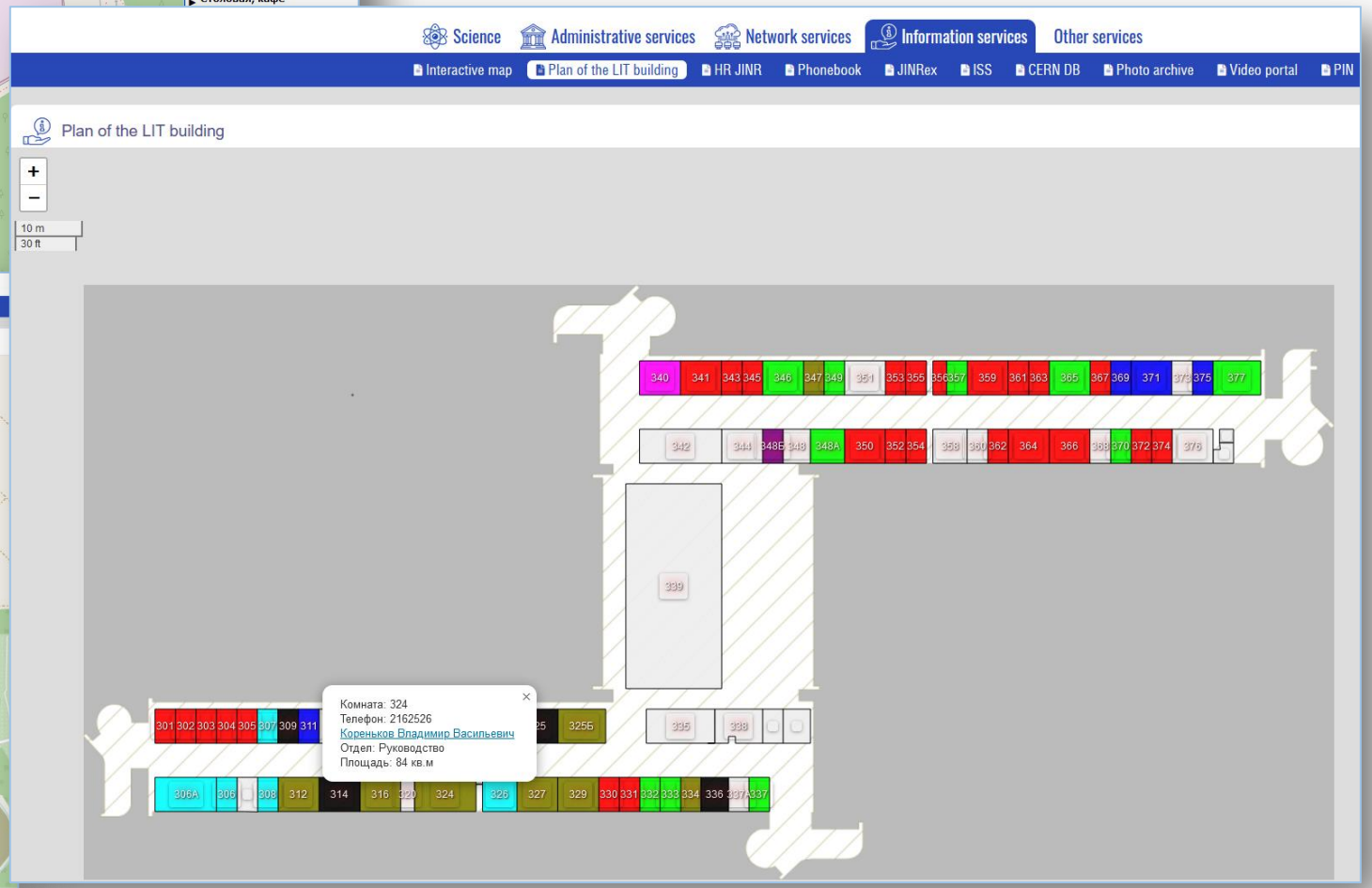




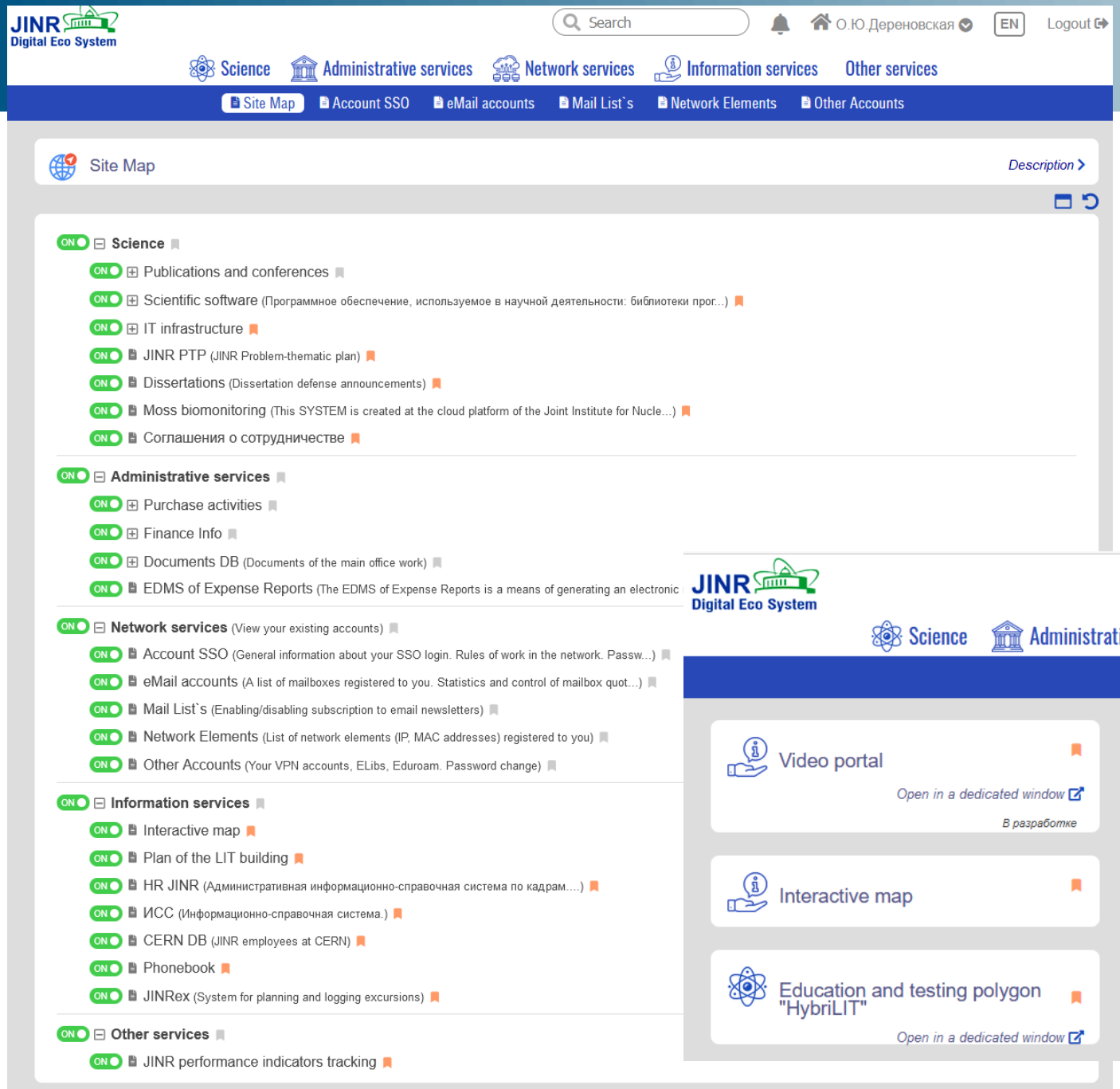
Easy access, convenient navigation and search for information on a large-scale network of a wide variety of JINR services



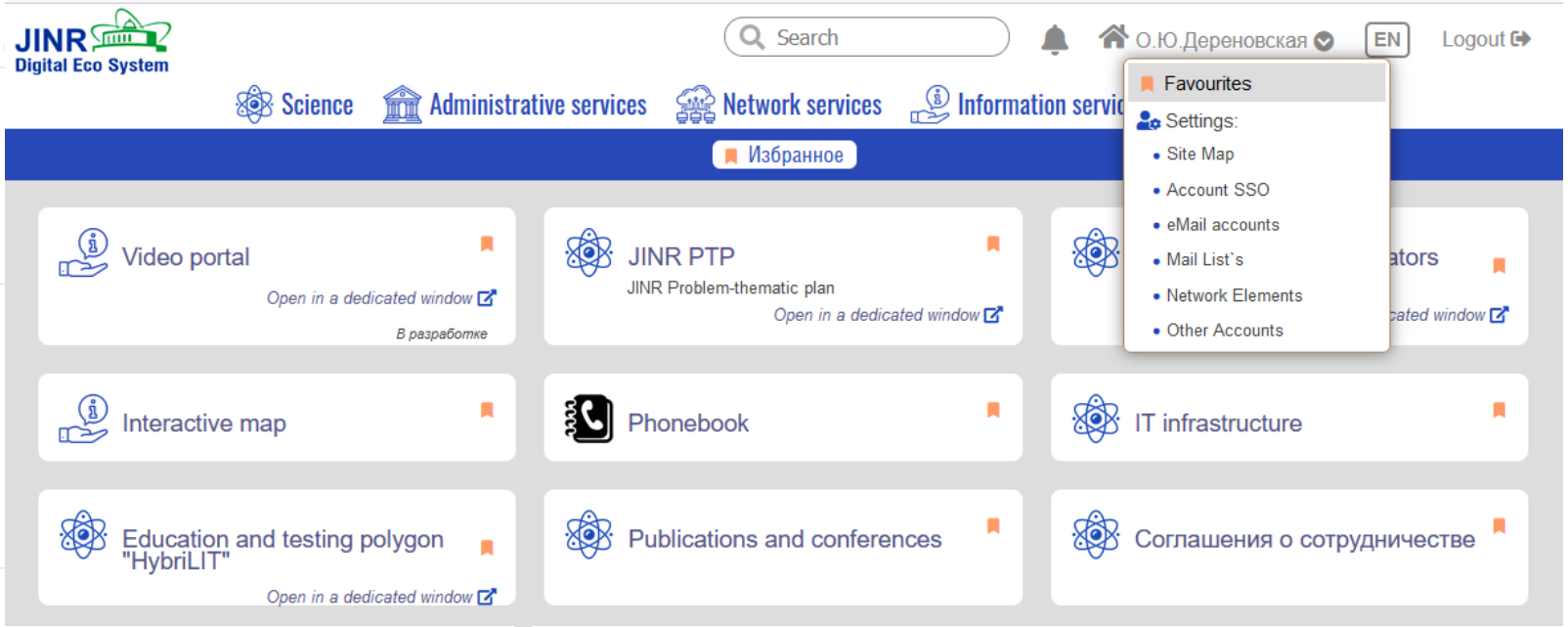




- ✓ Quick and easy search for information, both by services and by employees and buildings on an interactive JINR map



- ✓ Personal account of a JINR employee
- ✓ Notifications in your account
- ✓ Responsive interface, customizable by the user
- ✓ Part of the resources is available for unregistered users, for example, a phonebook, information on dissertations, scientific software, JINR map



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 Data Analysis Framework

MATLAB

GEANT4

Maple

Frameworks and tools for ML/DL tasks

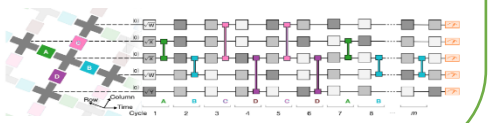
jupyterhub

TensorFlow

NumPy

scikit learn

Quantum algorithms, quantum programming and quantum control



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

- Distributed computing systems
- Computing for MegaScience Projects
- Distributed computing applications
- Data Management, Organisation and Access
- HPC
- Virtualization
- Big data Analytics and Machine learning
- Research infrastructure

NEC'2019

The International Symposium Nuclear Electronics and Computing

- Detector & Nuclear Electronics
- Triggering, Data Acquisition, Control Systems
- Distributed Computing, GRID and Cloud Computing
- Machine Learning Algorithms and Big Data Analytics new!
- Research Data Infrastructures
- Computations with Hybrid Systems (CPU, GPU, coprocessors)
- Computing for Large Scale Facilities (LHC, FAIR, NICA, SKA, PIC, XFEL, ELI, etc.)
- Innovative IT Education

MMCP

MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS

- ❑methods, software and program packages for data processing and analysis;
- ❑mathematical methods and tools for modeling complex physical and technical systems, computational biochemistry and bioinformatics;
- ❑methods of computer algebra, quantum computing and quantum information processing;
- ❑ machine learning and big data analytics;
- ❑ algorithms for parallel and hybrid calculations.

MLIT Schools

MMCP2019

MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS 2019

Stará Lesná, High Tatra Mountains, Slovakia
July 1 –5, 2019

in numbers

Total: 110 registrants

By counties: 15 states, including

- ARMENIA
- BELARUS
- CERN
- CZECH REPUBLIC
- EGYPT
- FINLAND
- GERMANY
- MOLDOVA
- ROMANIA
- SCOTLAND
- RUSSIA
- SLOVAKIA
- SWITZERLAND
- USA

By talks: 94, including

- Plenary: 19
- Sectional: 75

MMCP2019

MATHEMATICAL MODELING AND COMPUTATIONAL PHYSICS 2019

Stará Lesná, High Tatra Mountains, Slovakia
July 1 –5, 2019

Machine Learning, Parallel and Hybrid Computations & Big Data Analytics

Total: 25 registrants
From SLOVAKIA, RUSSIA

DATA SCIENCE
IT SCHOOL

Летняя компьютерная школа «Аналитика Больших данных Дубна-2019» Дубна, 6-13 июля 2019 г.

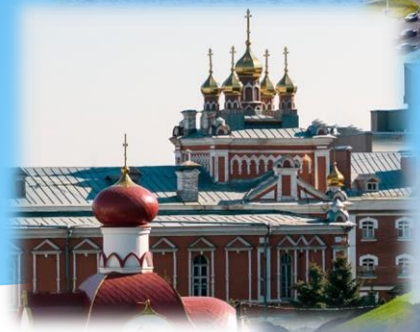
Регионы: 20 ВУЗов





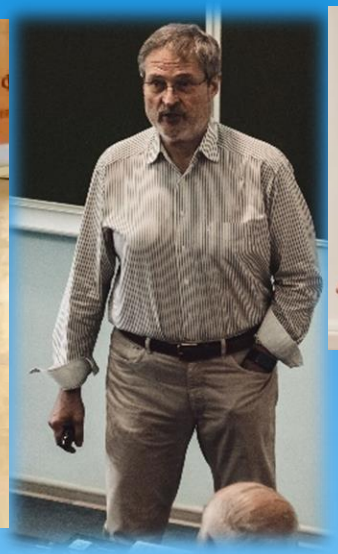







- БелГУ
- МАИ
- МГТУ
- МИЭМ НИУ ВШЭ
- НИУ МИЭТ
- НИЯУ МИФИ
- РУДН
- РЭУ
- ПГУТИ
- ПетрГУ
- РосНОУ
- СГАУ
- СКГМИ
- СКФУ
- СОГУ
- СПбГУ
- ТвГУ
- ТГУ
- Университет «Дубна»
- Федеральная политехническая школа Лозанны

Лекторы







18 лекций:

3 академика РАН

3 чл.-корр. РАН

14 докторов наук

6 кандидатов

Награждение победителей

Total: 193 registrants ***from 13 states,***

30 *institutions, including JINR, CERN*

SCHOOL NEC'2019

32 students from 10 universites

- Dubna State University
- Lomonosov Moscow State University
- National Research Nuclear University MEPhI
- North Ossetian State University
- Petrozavodsk State University
- Plekhanov Russian University of Economics
- RUDN University
- St Petersburg University
- Tomsk State University
- Tver State University

SCHOOL NEC'2019

Awarding "Best Students Reports"

2d Place Diploma

Volosnikov Vladislav	SPbGU
Rudenko Mikhail	Dubna

3d Place Diploma

Rogozhina Elizaveta	Dubna
Shaikhislamov Denis	MSU

Advanced Development

Solodilova Kseniia	Dubna
Ilina Anna	Dubna

Innovative Project

Shakhov Iurii	PRUE
Yurchenko Mikhail	TSU

Creative Presentation

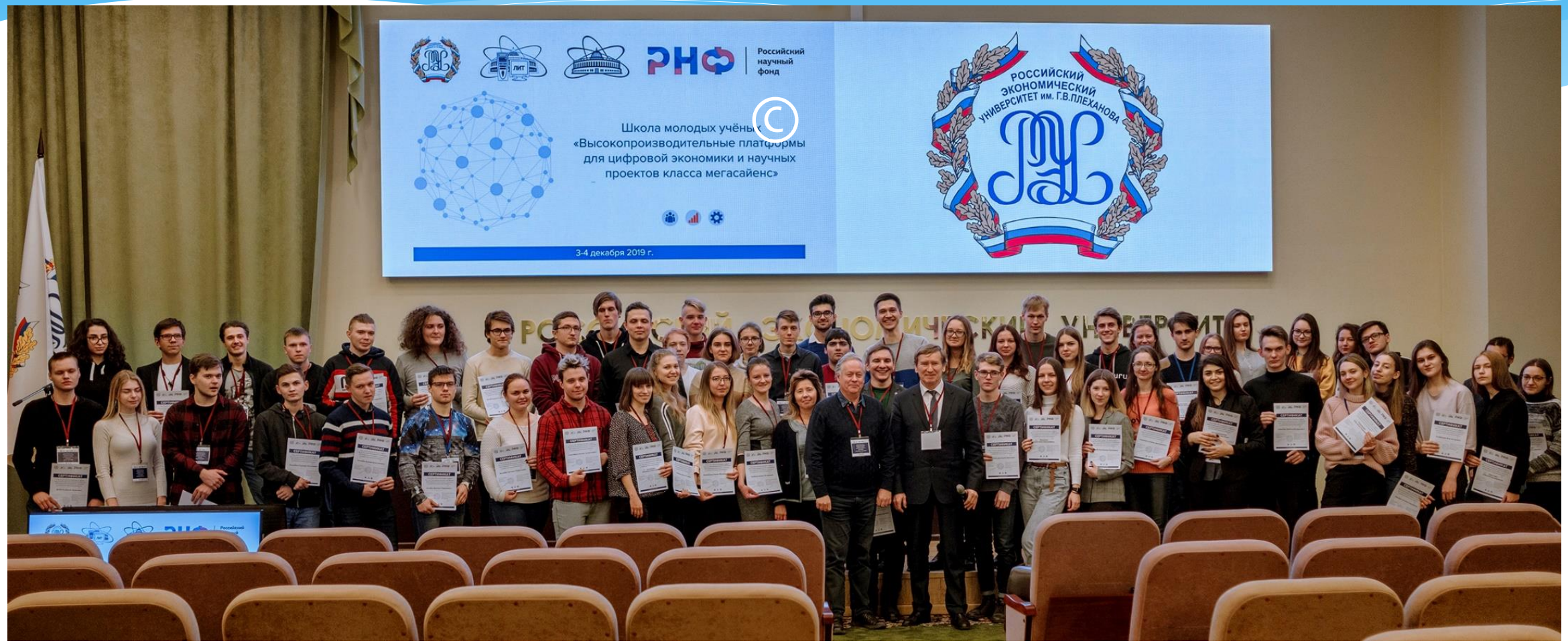
Antonov Evgenii	MEPhI
Shek Elena	PRUE

Promising Research

Safikanov Denis	MEPhI
Fomina Iuliia	MEPhI

Школа молодых ученых

Высокопроизводительные платформы для цифровой экономики и научных проектов класса мегасайенс»



3-4 декабря 2019, РЭУ им. Г.В. Плеханова, Москва, Россия

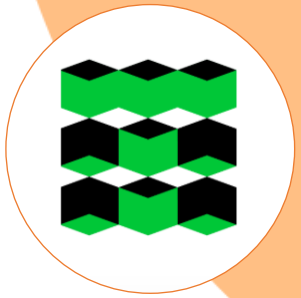
4-я Международная летняя школа молодых ученых «Современные информационные технологии для научных и прикладных задач»

29 июня — 1 июля 2022 г. Северо-Осетинский государственный университет, Владикавказ

JINR School of Information Technology 2022

60 students from 13 universities



- Dubna State University
- Far Eastern Federal University
- National Research Nuclear University MEPhI
- North Ossetian State University after K.L. Khetagurov
- Plekhanov Russian University of Economics
- St. Petersburg University
- The Bauman Moscow State Technical University
- The National University of Science and Technology (MISIS)
- The Peoples' Friendship University of Russia
- Tomsk Polytechnic University
- Tula State University
- Tver State University
- Vitus Bering Kamchatka State University