INTERNATIONAL INTERGOVERNMENTAL ORGANIZATION
МЕЖДУНАРОДНАЯ МЕЖПРАВИТЕЛЬСТВЕННАЯ ОРГАНИЗАЦИЯ

JOINT INSTITUTE FOR NUCLEAR RESEARCH
ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

Исторические этапы, статус и перспективы развития компьютерной инфраструктуры ЛИТ ОИЯИ

Кореньков Владимир Васильевич

Научный руководитель Лаборатории информационных технологий имени М.Г. Мещерякова ОИЯИ

Нижний Новгород, 25-26 сентября 2023 г.

Эволюция ИТ

- менялись концепции
- круг и сложность решаемых задач
- возникал новый технологический набор
- углублялась специализация разработчиков
- сокращалось время ввода новых продуктов и сервисов

первые цифровые платформы поддержки, работающих в реальном времени

Первое поколение (1960-е годы) — мейнфреймы

Второе поколение (1970-е годы) — универсальные ЭВМ

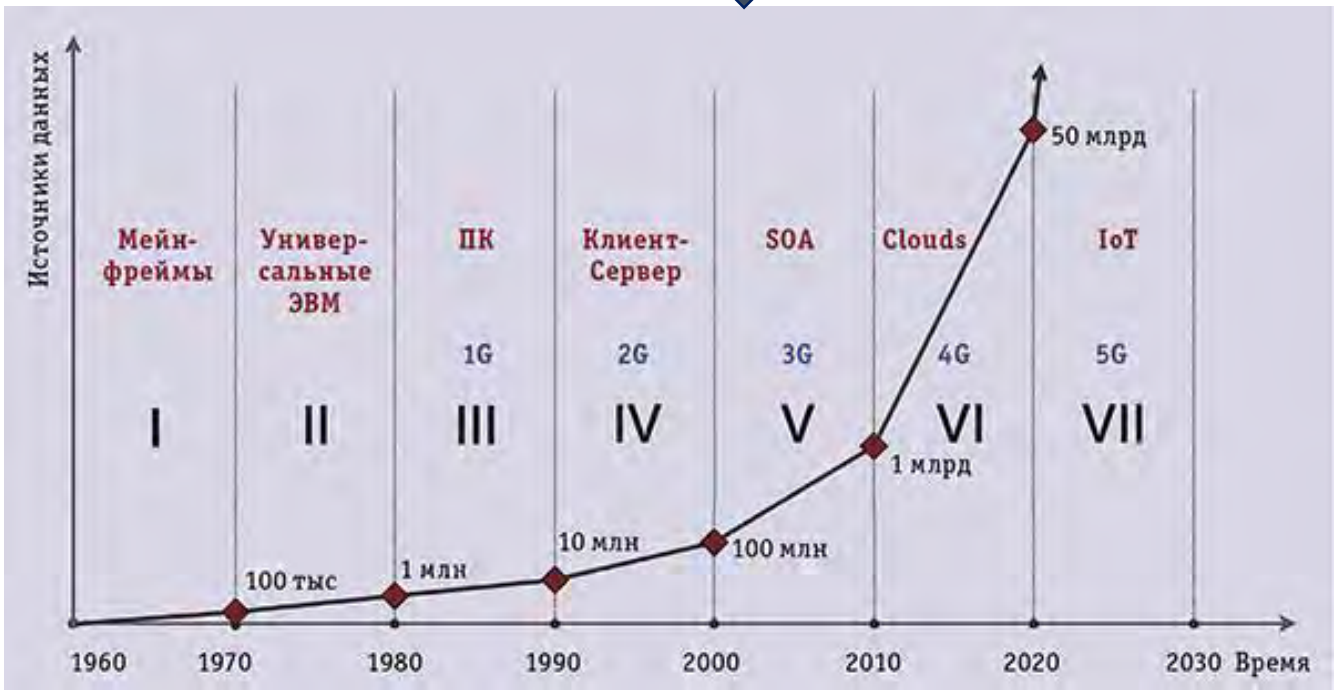
Третье поколение (1980-е годы) — персональные компьютеры

Четвертое поколение (1990-е годы) — клиент-сервер

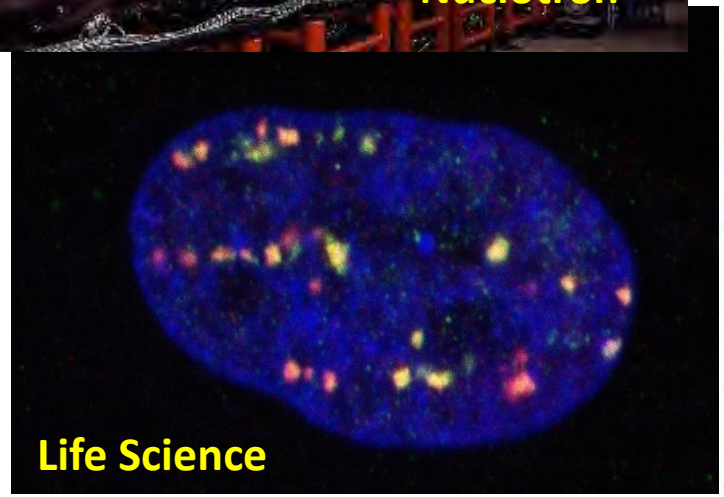
Пятое поколение (2000-е годы) — сервисная архитектура

Шестое поколение (2010-е годы) — облака

Седьмое поколение (2020-е годы) — IoT, искусственный интеллект, квантовые вычисления



Базовые установки ОИЯИ



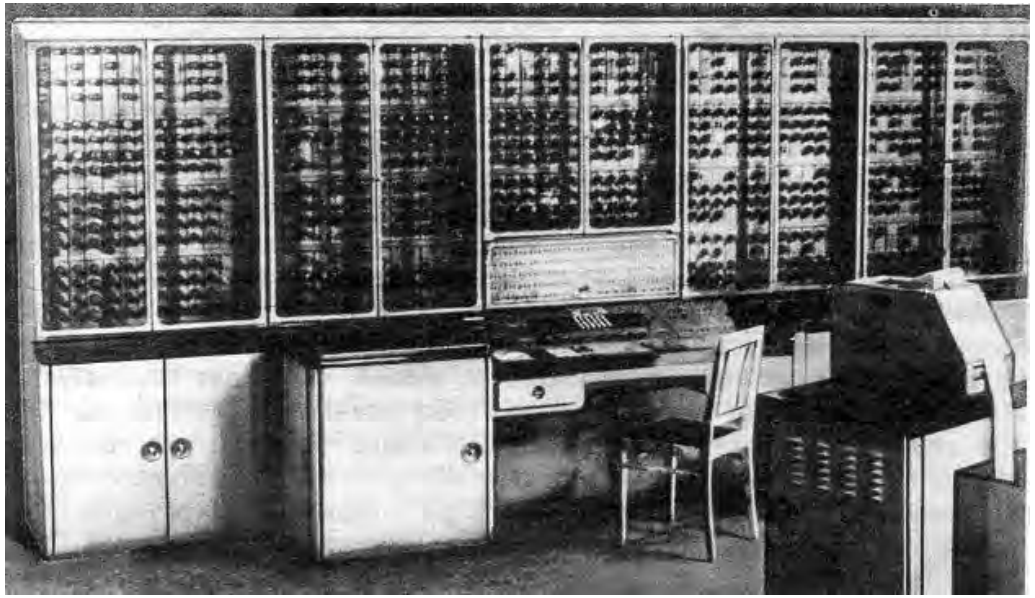
Meshcheryakov Laboratory of Information Technologies

M.G. Mesheryakov
(17.09.1910 - 24.05.1994)

N.N. Govorun
(18.03.1930 - 21.07.1989)

Meshcheryakov Laboratory of Information Technologies 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.

Первые шаги по созданию ЛВТА



1958-й г. – Первая ЭВМ в ОИЯИ "Урал-1" производительностью 100 операций/с и памятью на магнитном барабане

В 1962 г. был сделан первый реальный шаг на пути построения многомашинного комплекса, включающего уровень ЭВМ накопления и предварительной обработки данных с аппаратуры физического эксперимента

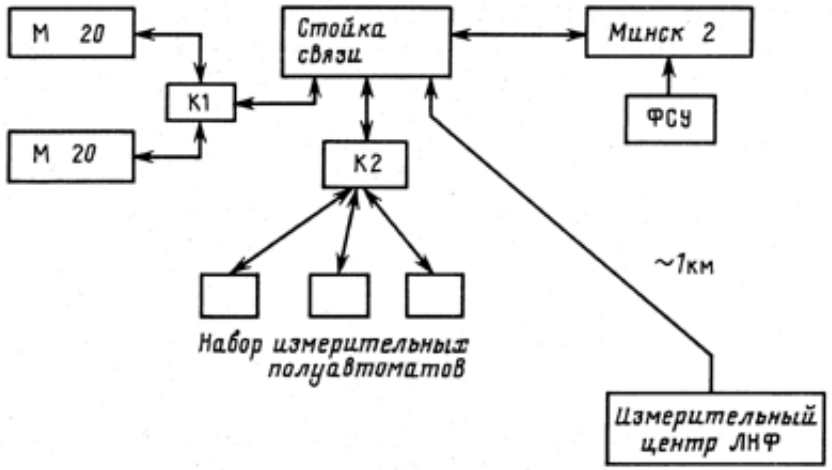


Рис. 1

Эпоха БЭСМ-6

1967-1968 гг. серийно выпускается самая быстрая в Европе отечественная ЭВМ БЭСМ-6. Под руководством Н.Н. Говоруна запущена в эксплуатацию в ОИЯИ БЭСМ-6. Создается мониторная система "Дубна" для БЭСМ-6, в которую включается язык ФОРТРАН и другие языки программирования. Следующий этап – создание операционной системы «Дубна»

Н.Н.Говорун знакомит гостей из ЦЕРН с БЭСМ-6

Центральный вычислительный комплекс (ЦВК ОИЯИ)

1968-1972 – в ОИЯИ создается многомашинный комплекс, который оснащается, наряду с отечественными вычислительными машинами, ЭВМ фирмы Control Data Corporation (CDC), используемыми в ЦЕРНе. Создана трехуровневая компьютерная среда (ЦВК ОИЯИ в ЛВТА, ИВК в Лабораториях ОИЯИ, малые ЭВМ на линии с физическими установками.

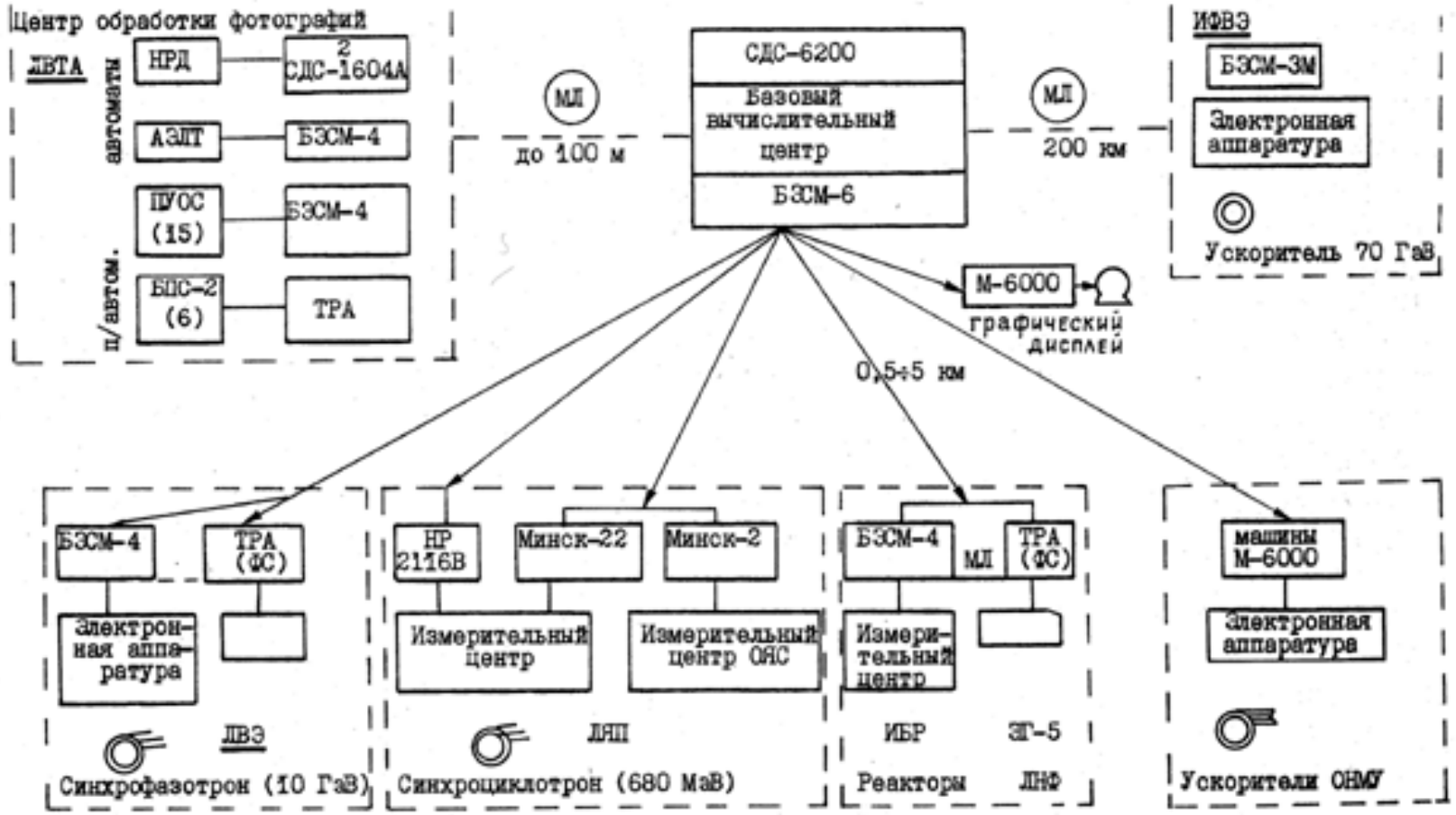
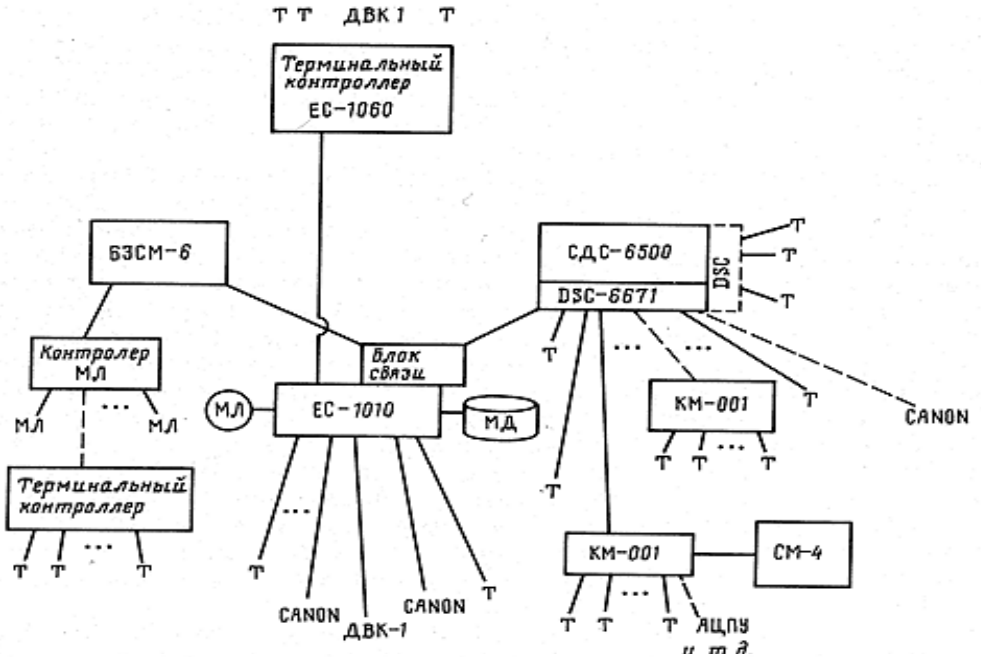
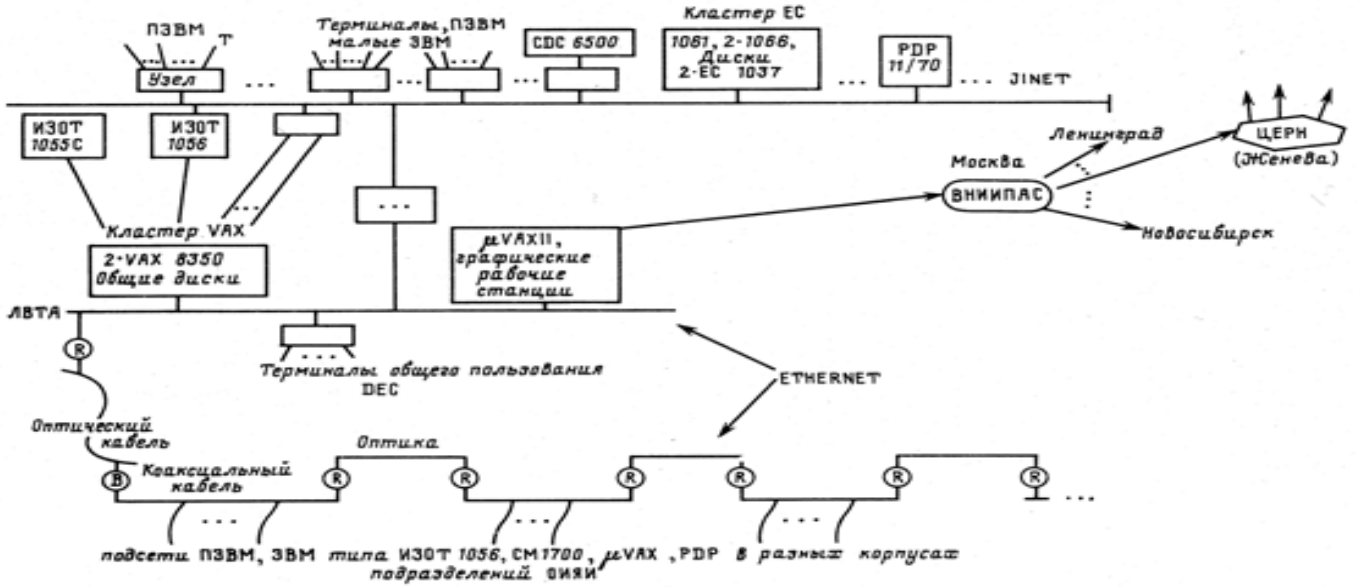


Рис. 2

Локальная сеть ОИЯИ на базе JINET и ETHERNET





Этапы развития внешних коммуникаций ОИЯИ

1991 – 64 Кбит/с спутниковый канал связи с узлом сети HEPNET в Италии

1994 – 64 Кбит/с спутниковый канал связи с узлом сети DFN в Германии

1995 – 128 Кбит/с наземный канал связи с узлом INTERNET в Москве

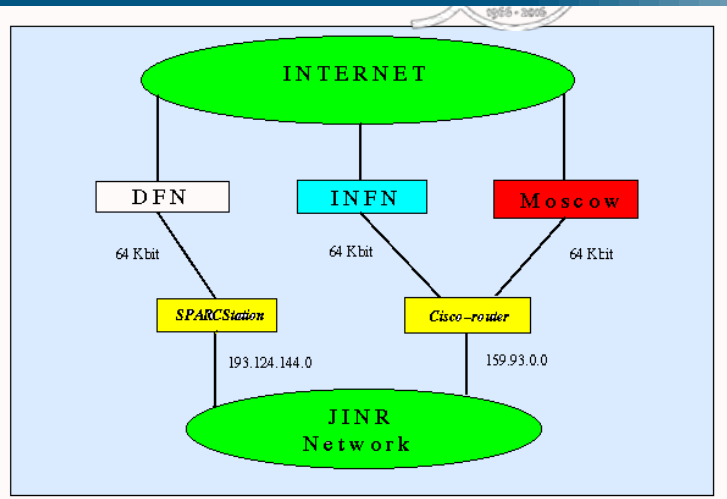
1997 – 2 Мбит/с оптический канал связи ОИЯИ-ЦКС «Дубна»-Шаболовка-М9. Узел сети RBNet в Дубне

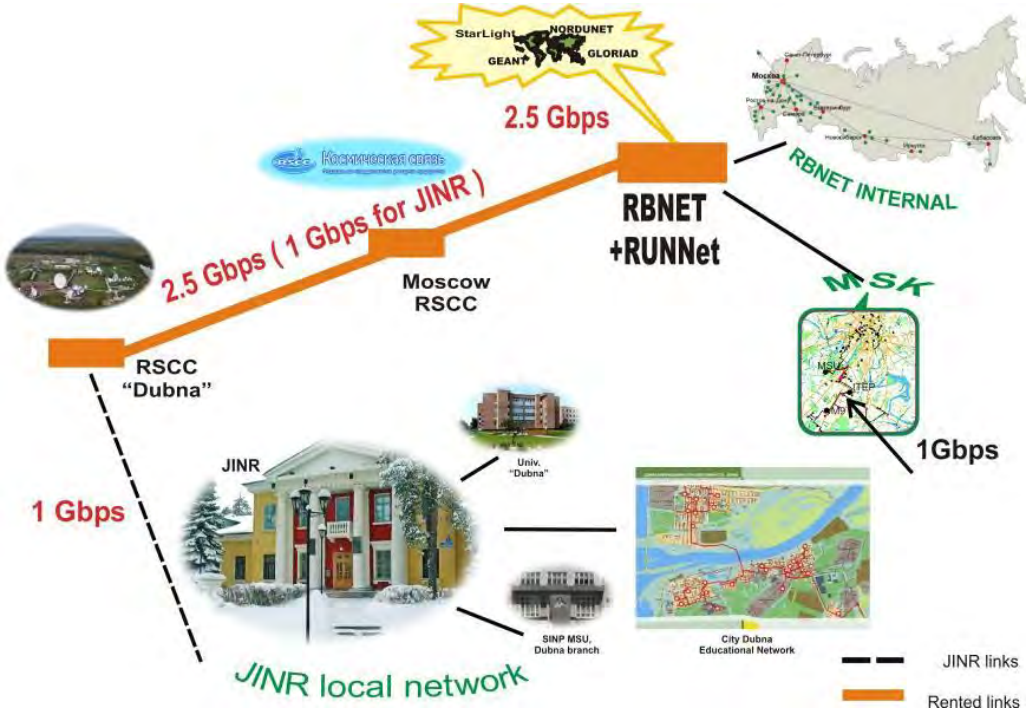
2001 – реализация проекта ATM канала связи Дубна-Москва емкостью 622 Мбит/с (155 Мбит/с для ОИЯИ)

2002 – спутниковый канал ОИЯИ-пансионат «Дубна» в Алуште (первый канал в Крыму)

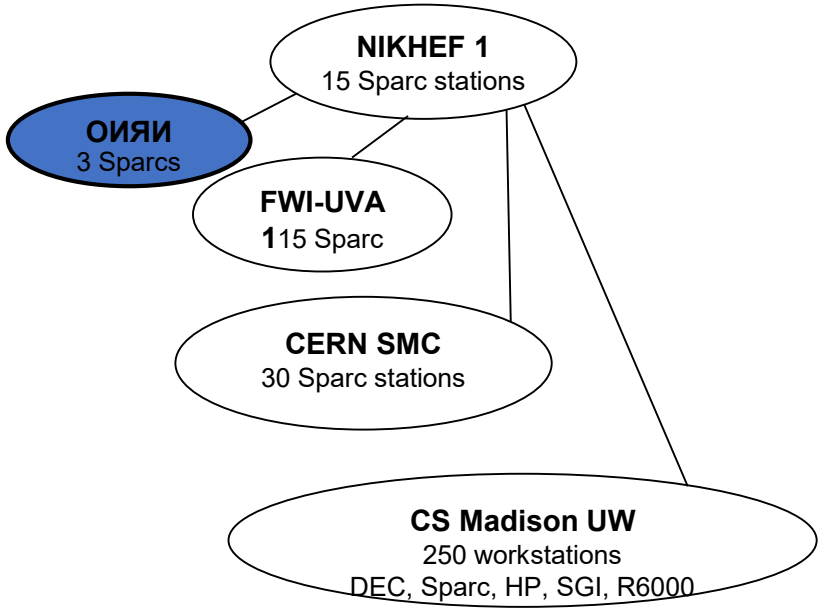
2005 – реализация проекта канала связи Дубна-Москва на основе технологии SDH емкостью 2.48 Гбит/с (1 Гбит/с для ОИЯИ)

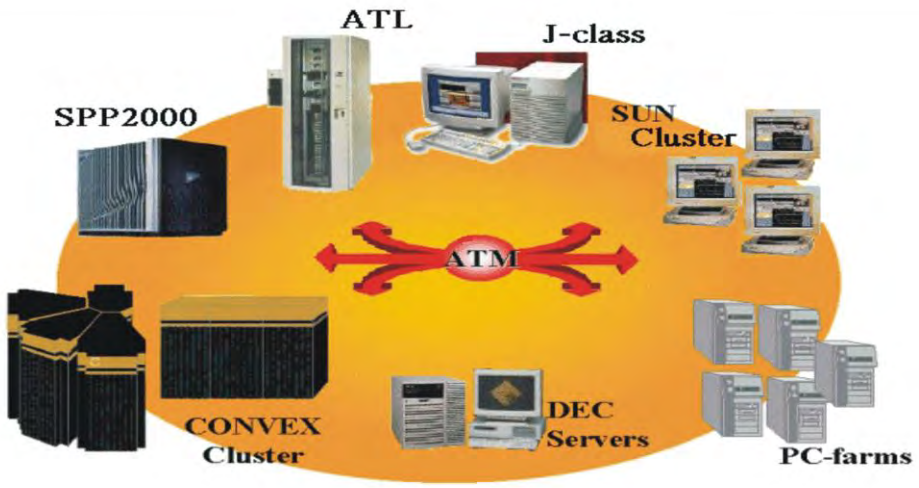
2009– реализация проекта канала связи Москва-Дубна на базе технологии DWDM (20 Гбит/с)

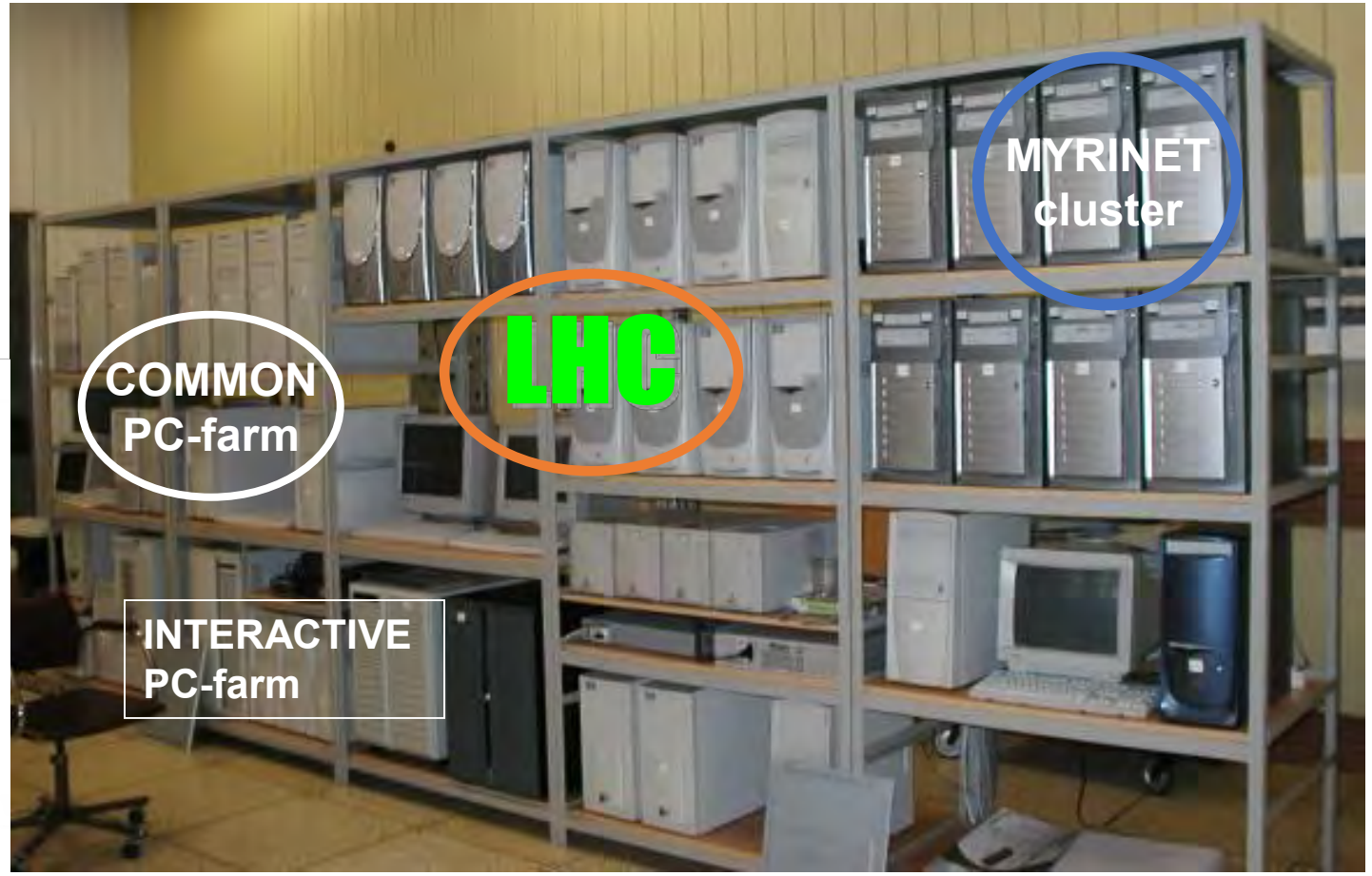




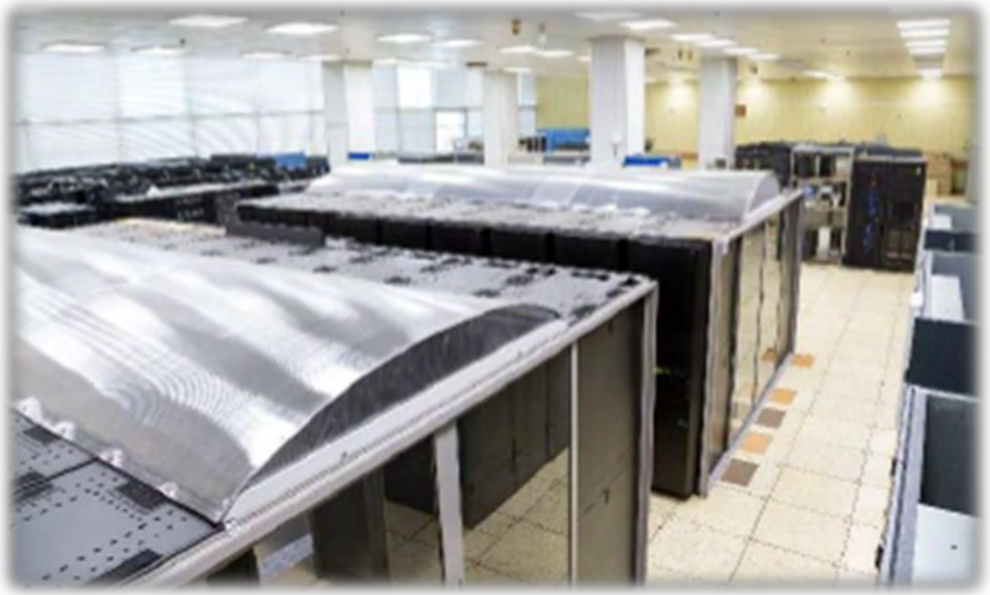
Развитие компьютерной инфраструктуры ОИЯИ с 1994-2004 год







MLIT today



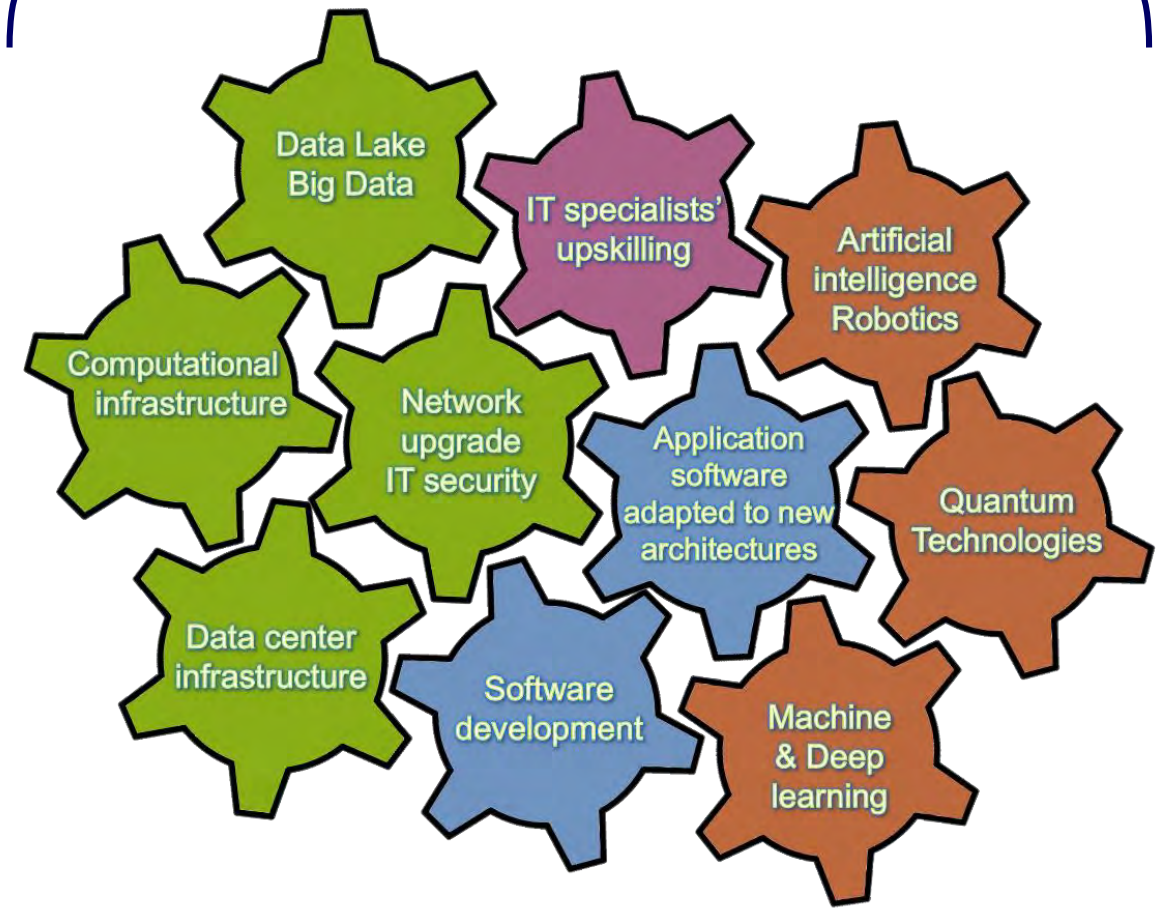
Staff: 325
Scientists: 100
Doctors of Science: 24
Candidates of Science: 61
Campus network 2x100 Gbps
Multisite network 4x100 Gbps
Telecommunication channel 3x100 Gbps
Grid Tier1 and Tier2 for global data processing
JINR Cloud computing
JINR Member States' Cloud environment
"Govorun" supercomputer

MLIT Fundamentals:

* **Provide** IT services necessary for the fulfillment of the JINR Topical Plan on Research and International Cooperation
* **Building** world-class competence in IT and computational physics
* **24/7** support of computing infrastructure and services such availability is called nonstop service.

Strategy for Information Technology and Scientific Computing at JINR

Scientific IT ecosystem:



Coordinated development of interconnected IT technologies and computational methods

It will be a **steady implementation/upgrades** of

- Networking (**Tb/s** range),
- Computing infrastructure within the Multifunctional Information & Computing Complex (MICC) and
- "Govorun" Supercomputer,
- Data center infrastructure,
- Data Lake & long-term storage

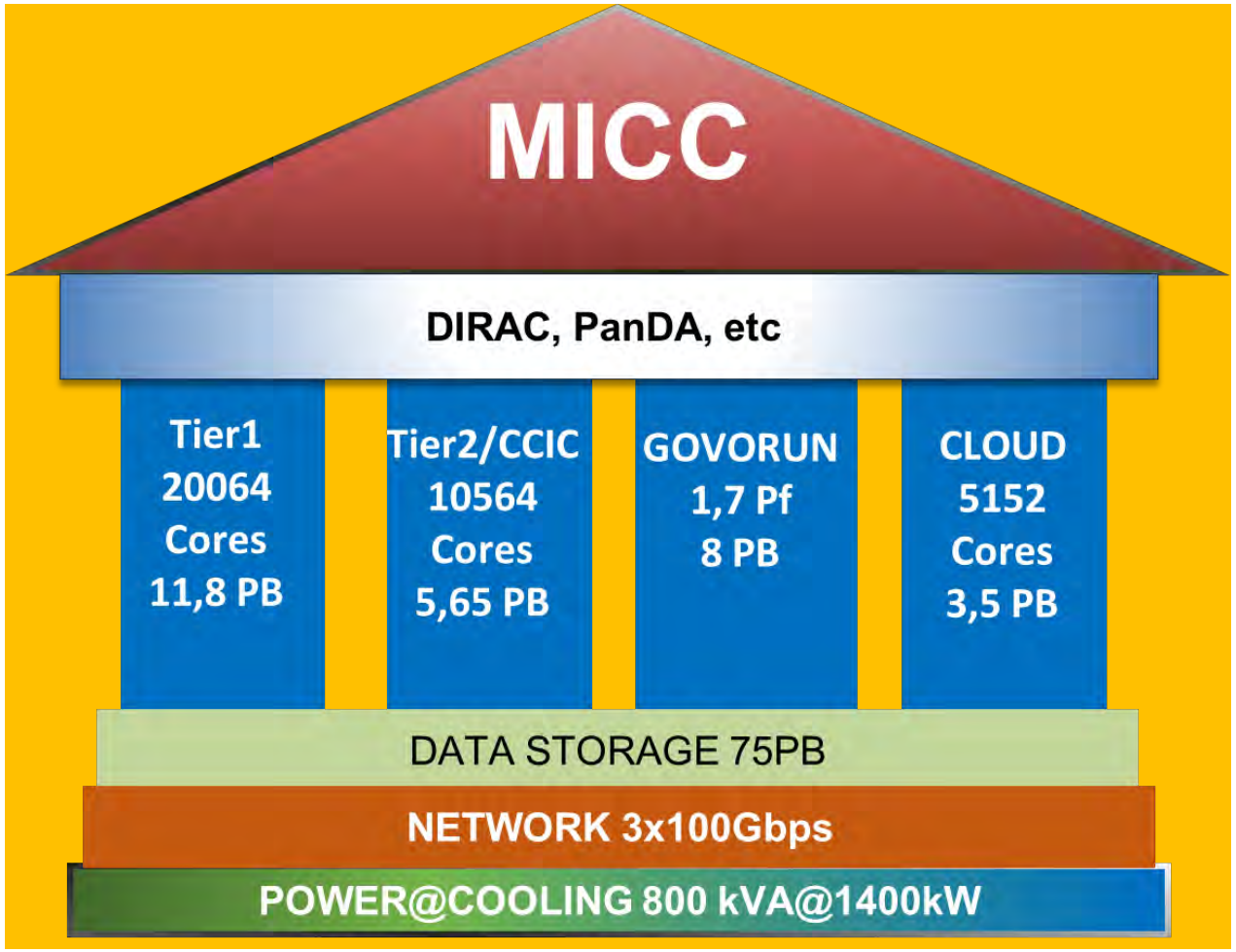
for all experiments.

The **development of new data processing and analysis algorithms** based on

- ML/DL,
- Artificial intelligence,
- Big Data
- Quantum technologies.

A variety of means will be used for IT specialists' upskilling.

Multifunctional Information and Computing Complex (MICC)



4 advanced software and hardware components

- Tier1 grid site
- Tier2 grid site
- hyperconverged “Govorun” supercomputer
- cloud infrastructure

Distributed multi-layer data storage system

- Disks
- Robotized tape library

Engineering infrastructure

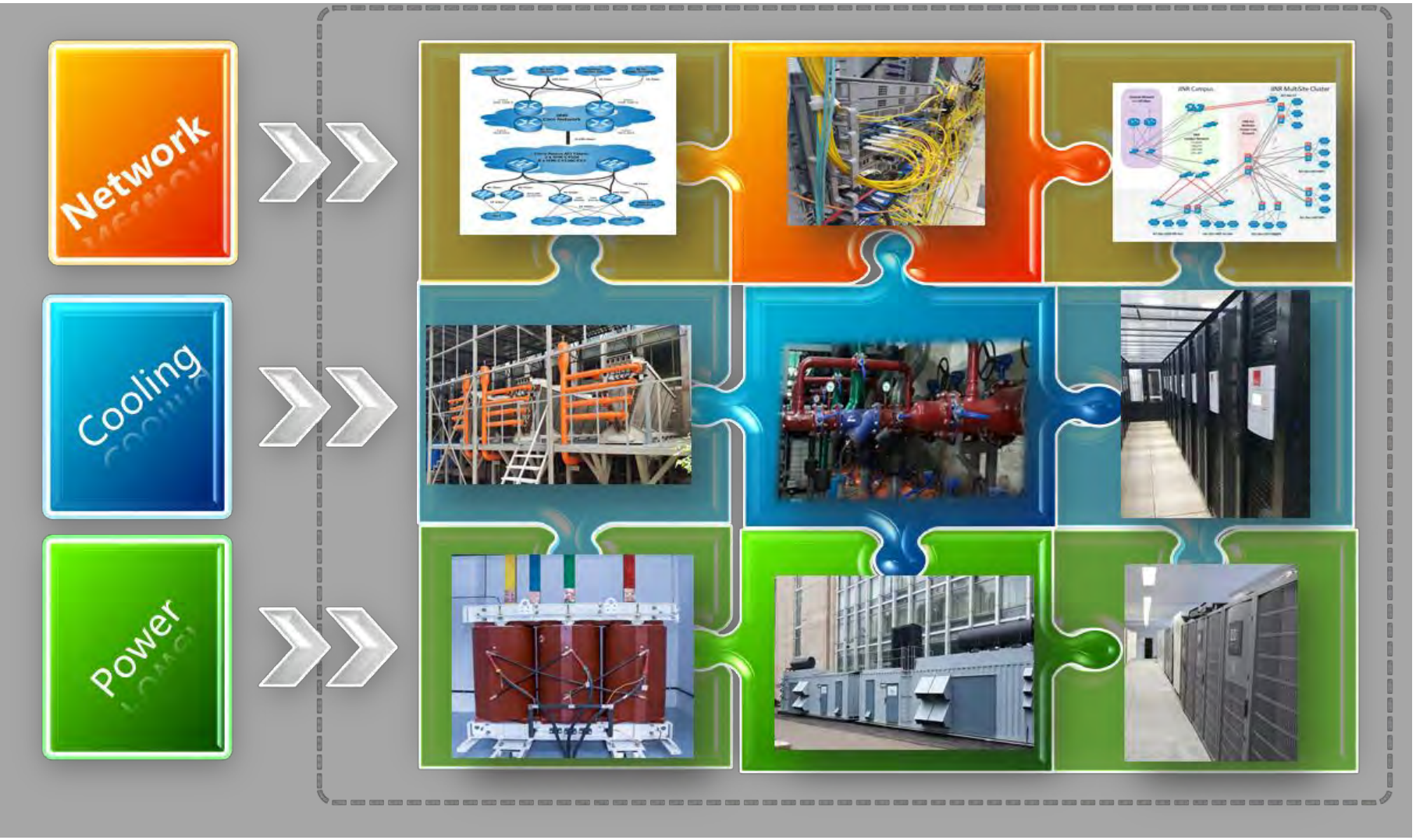
- Power
- Cooling

Network

- Wide Area Networkr
- Local Area Network

The main objective of the project is to ensure multifunctionality, scalability, high performance, reliability and availability in 24x7x365 mode for different user groups that carry out scientific studies within the JINR Topical Plan

MICC Power @ Cooling @ Network

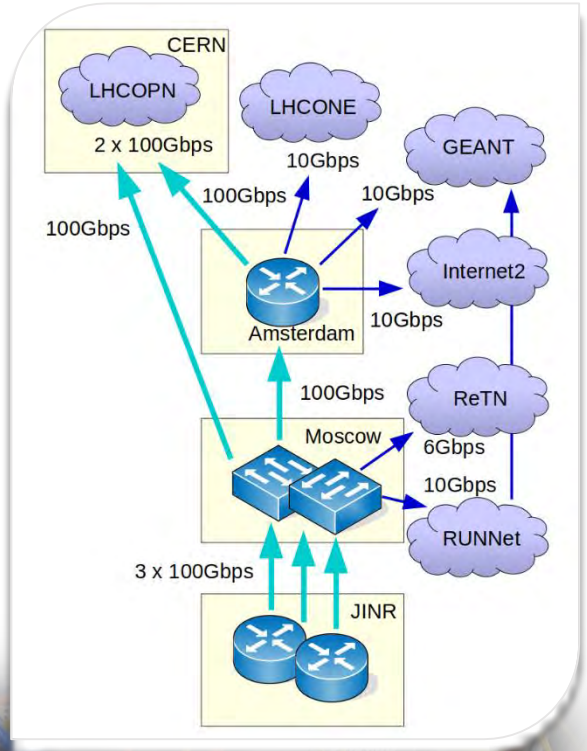


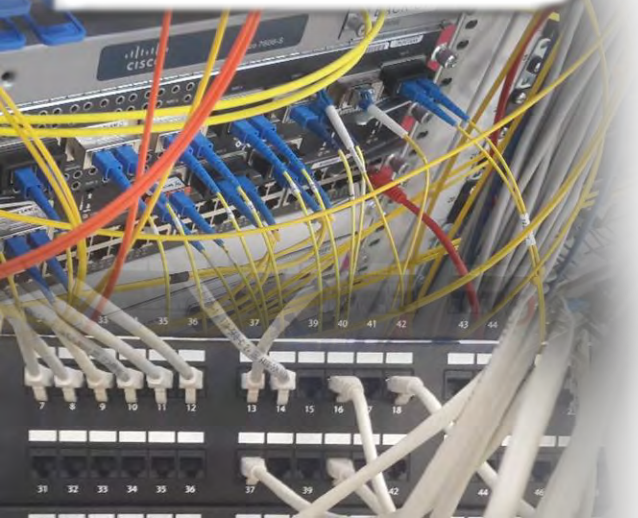
Wide Area Network 3x100 Gbps
Cluster Backbone 4x100 Gbps
Campus Backbone 2x100 Gbps

Dry chillers
InRow systems
Total cooling 1400 kW

Uninterruptible power supplies
8 x 300 kVA
Diesel-generator units (DGU)
2x1500 kVA
Transformers2x2500 kVA

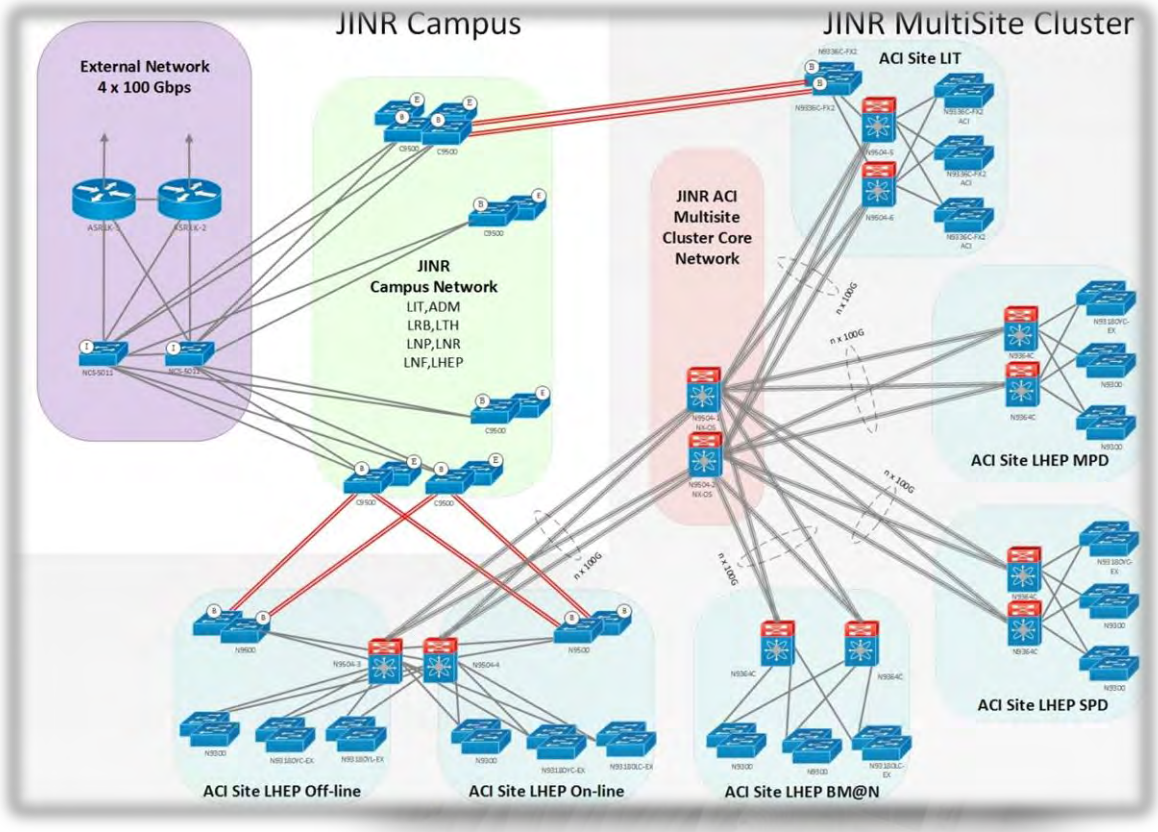
Networking





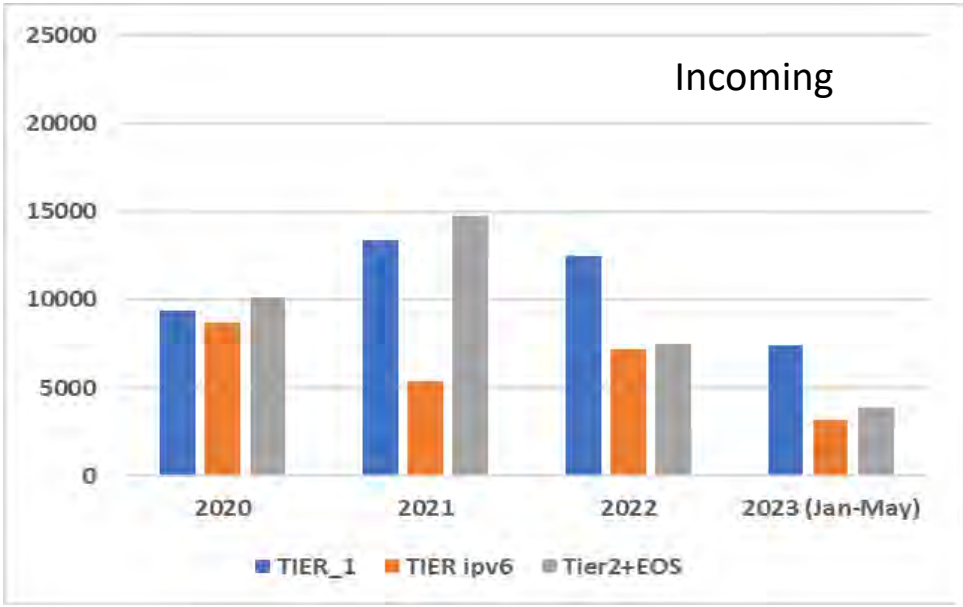
- JINR-Moscow 3x100 Gbit/s
- JINR-CERN - 100 Gbit/s and JINR-Amsterdam 100 Gbit/s for LHCOPN, LHCONE, GEANT networks
- Direct channels up to 100 Gbit/s for communication using RU-VRF technology with the collaboration of RUHEP research centers and with Runnet, ReTN networks
- The multi-site cluster network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

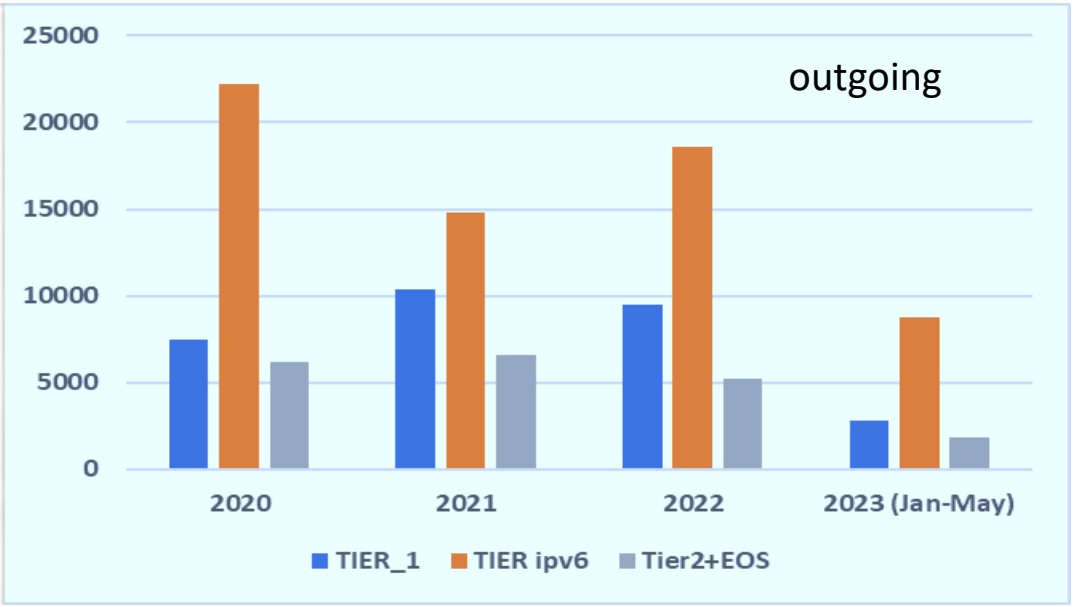
Users - 6353
Network elements - 9327
IP addresses - 18163
Remote access - 911
E-library- 1464
VOIP - 121
EDUROAM - 116
Email @jinr.ru - 4579



Networking @ Traffic

Distribution of the incoming and outgoing traffics by the JINR MICC in 2020-2023 (TB)

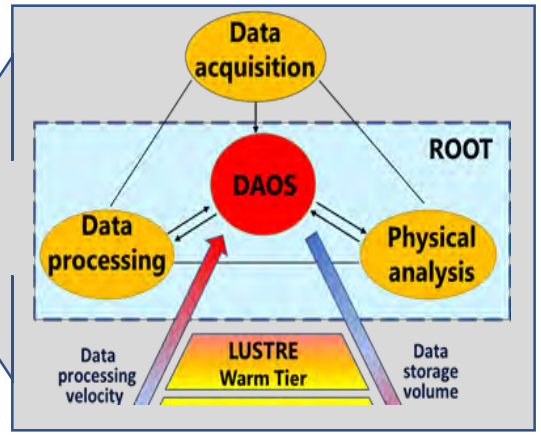




JINR Traffic in PB

Distributed Multilayered Data Storage System

- Limited data and short-term storage – to store OS itself, temporary user files
- AFS distributed global system – to store user home directories and software
- dCache is traditional for MICC grid sites – to large amounts of data (mainly LHC experiments) for middle-term period
- EOS is extended to all MICC resources – to store large amounts of data for middle-term period. At present, EOS is used for storage by BM@N, MPD, SPD, BaikalGVD, etc.
- Tape robotic systems – to store large amounts of data for long-term period. At present for CMS. BM@N, MPD, SPD, JUNO – in progress.



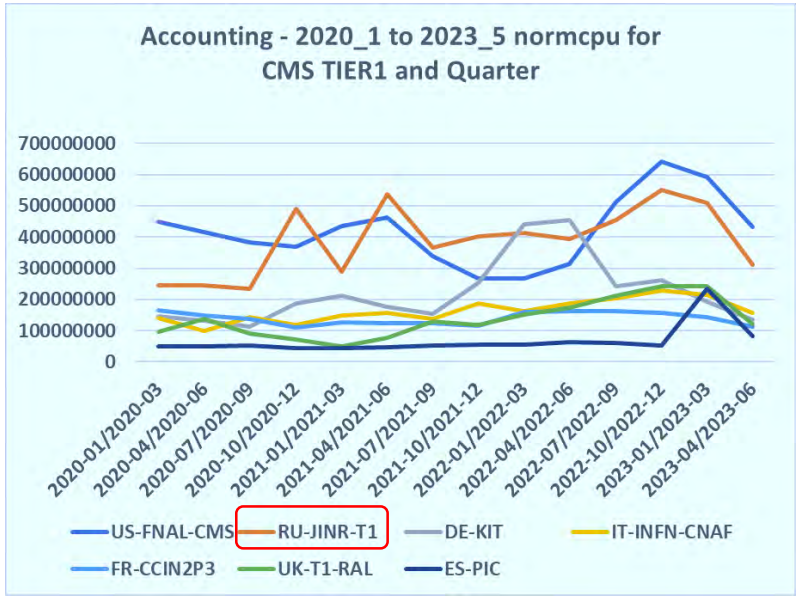
Special hierarchical data processing and storage system with a software-defined architecture was developed and implemented on the “Govorun” supercomputer.

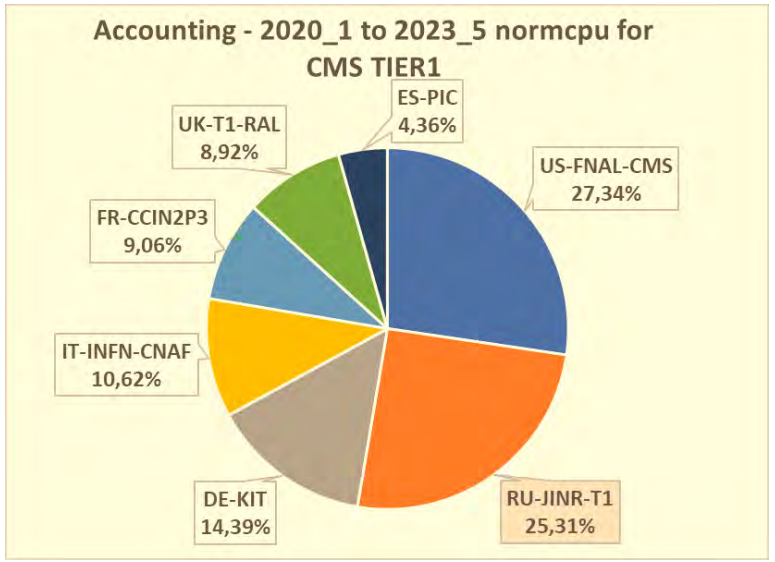
According to the speed of accessing data there are next layers:

- ✓ very hot data (DAOS (Distributed Asynchronous Object Storage)) ,
- ✓ the most demanded data (fastest access),
- ✓ hot data
- ✓ warm data (LUSTRE).

JINR Tier1 for CMS (LHC) and NICA

2020-2023

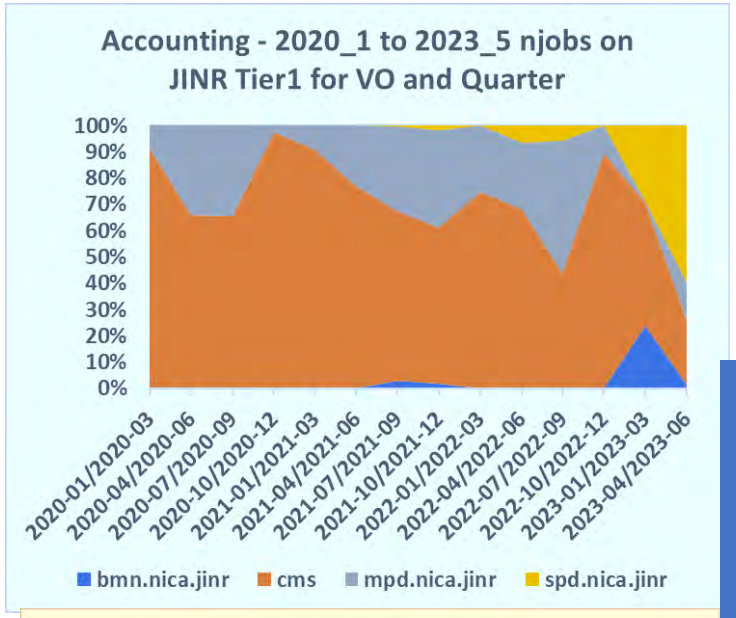


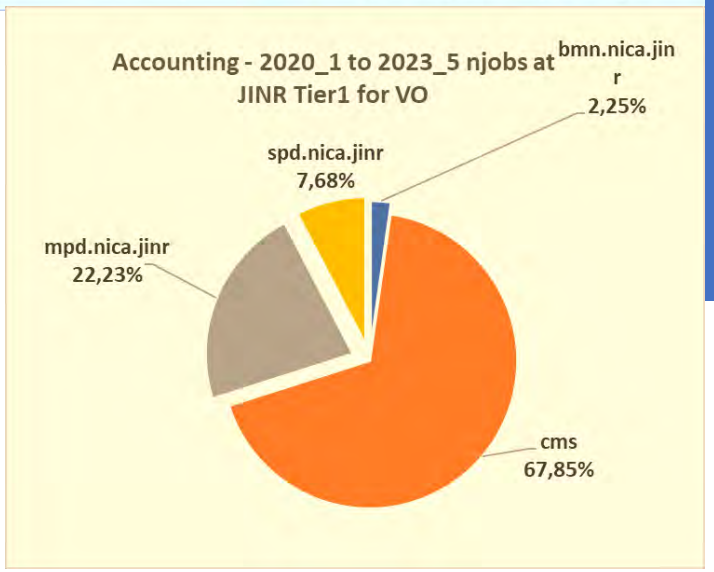


Since the beginning of 2015, a full-scale WLCG Tier1 site for the CMS experiment has been operating at MLIT JINR.

The importance of developing, modernizing and expanding the computing performance and data storage systems of this center is dictated by the research program of the CMS experiment, in which JINR physicists take an active part within the RDMS CMS collaboration.

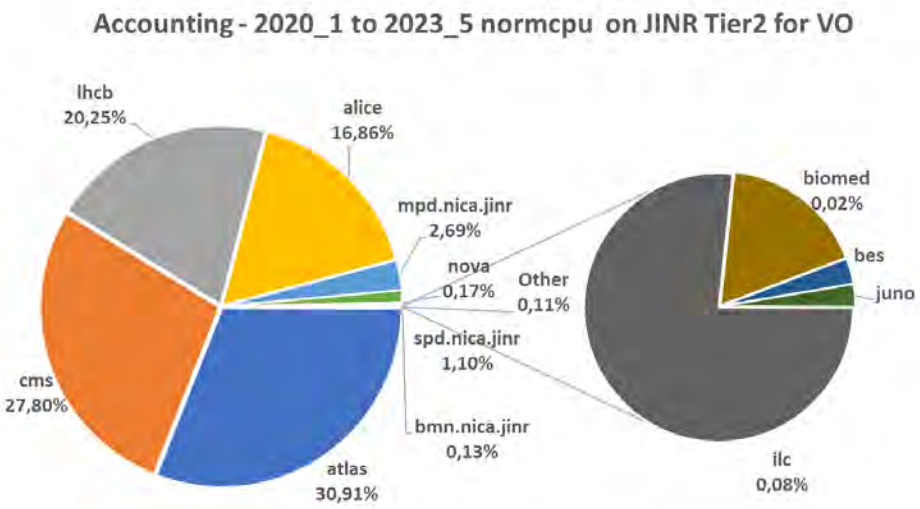
The JINR Tier1 is regularly ranked on top among world Tier1 sites that process data from the CMS experiment at the LHC.

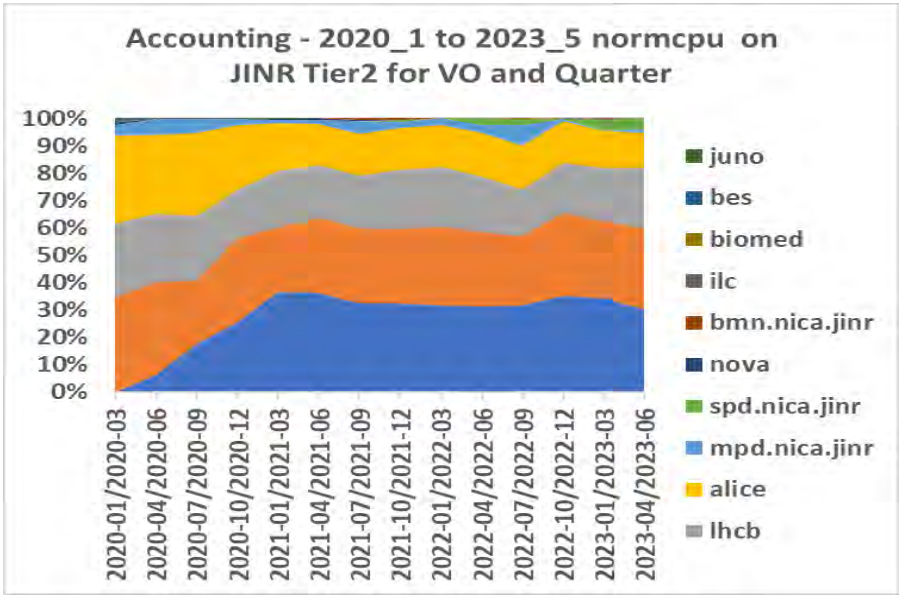




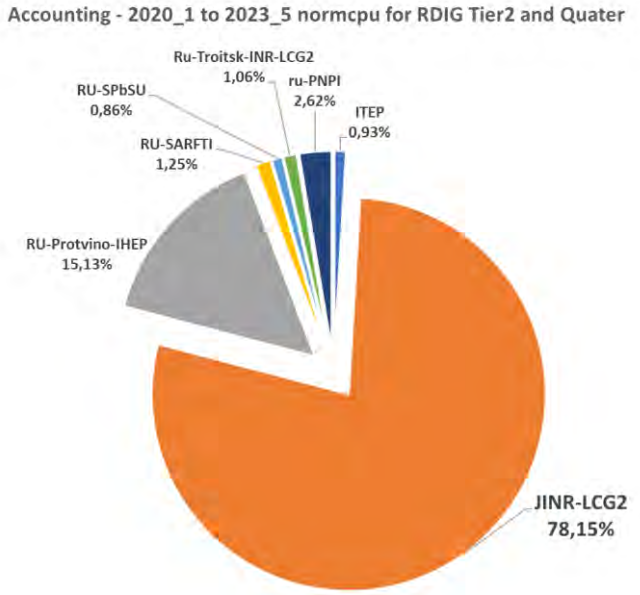
Since 2021 the JINR Tier1 center has demonstrated stable work not only for CMS (LHC), but also for NICA experiments.

JINR Tier2 in WLCG & RDIG

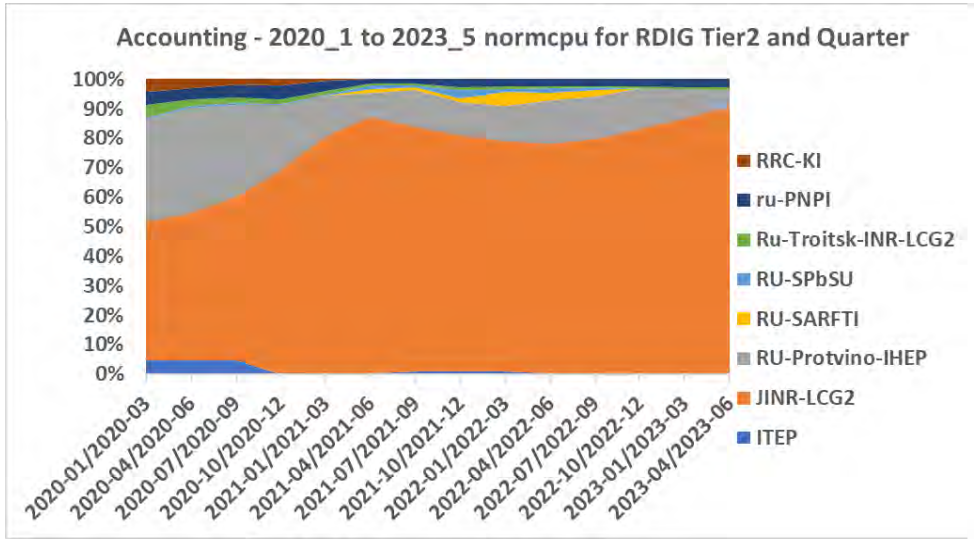




Tier2 at JINR provides computing power and data storage and access systems for the majority of JINR users and user groups, as well as for users of virtual organizations (VOs) of the grid environment (LHC, NICA etc.).



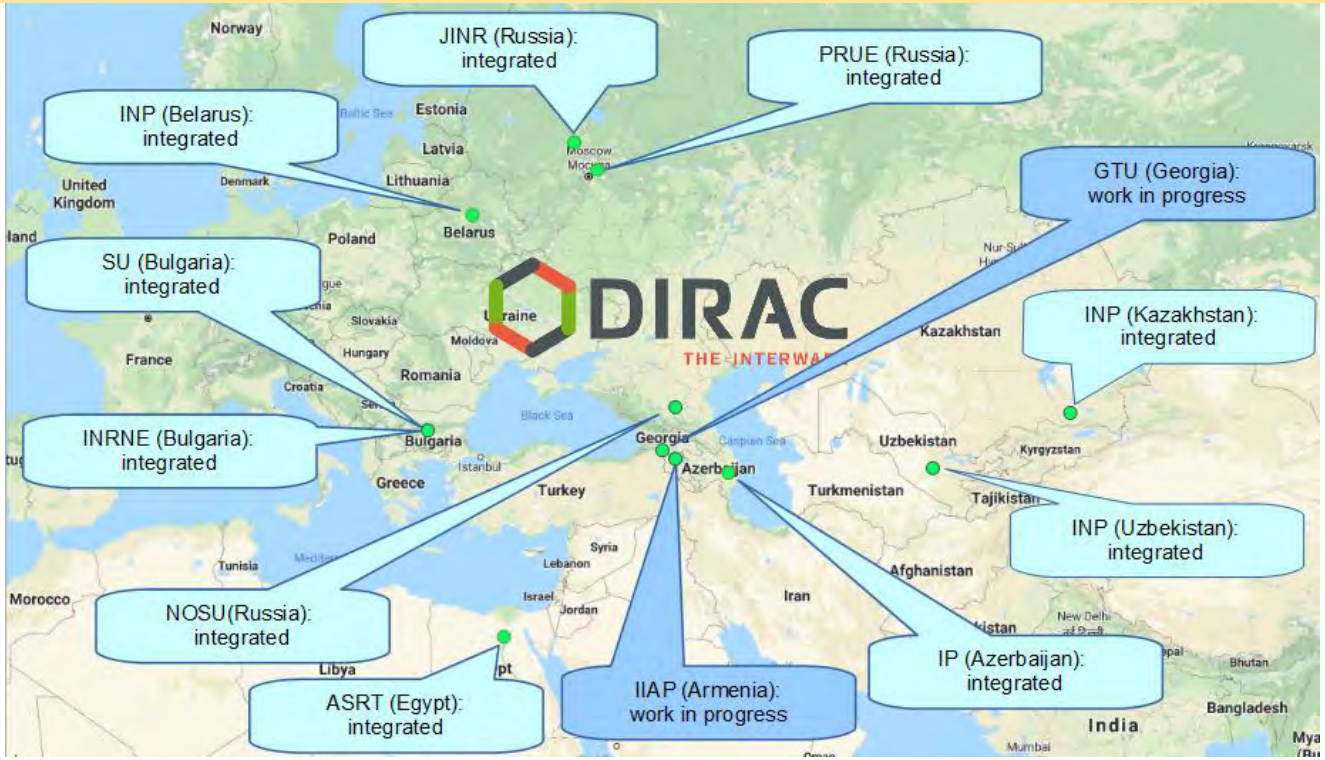
JINR Tier2 is the most productive in the Russian Data Intensive Grid (RDIG) Federation. More than 80% of the total CPU time in the RDIG is used for computing on our site.



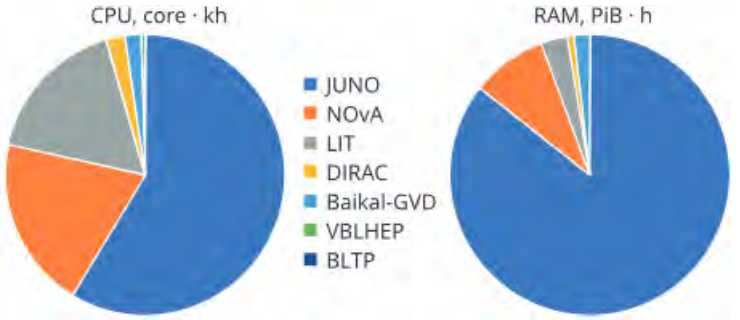
Cloud Infrastructure

- Computational resources for neutrino experiments:
- VMs for JINR users
- Testbeds for research and development in IT
- COMPASS production system services
- Data management system of the UNECE ICP Vegetation
- Service for data visualization, Gitlab and some others

DIRAC-based distributed information and computing environment (DICE) that integrates the JINR Member State organizations' clouds



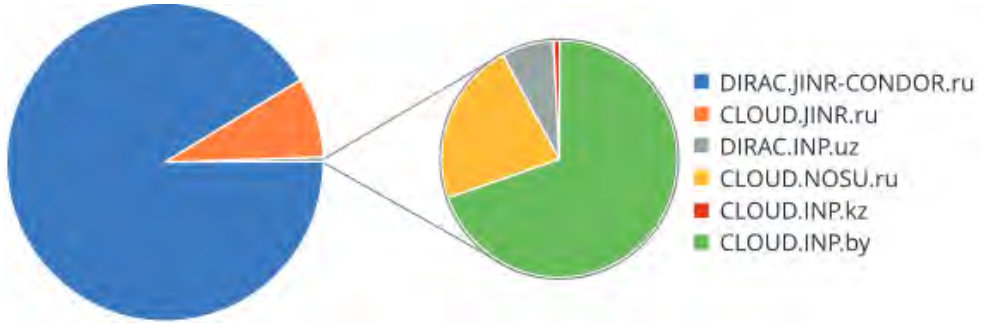
The Baikal-GVD, NOvA and JUNO experiments are the major users of the cloud infrastructure.



Use of cloud computing by experiments and JINR subdivisions in 2022

Most of the jobs in the JINR DICE in 2022 were performed on the neutrino computing platform (DIRAC.JINR-CONDOR.ru).

Distribution of the number of jobs completed in the JINR DICE by participants



The main consumer of the JINR DICE resources in 2022 was the Baikal-GVD experiment (96%).

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

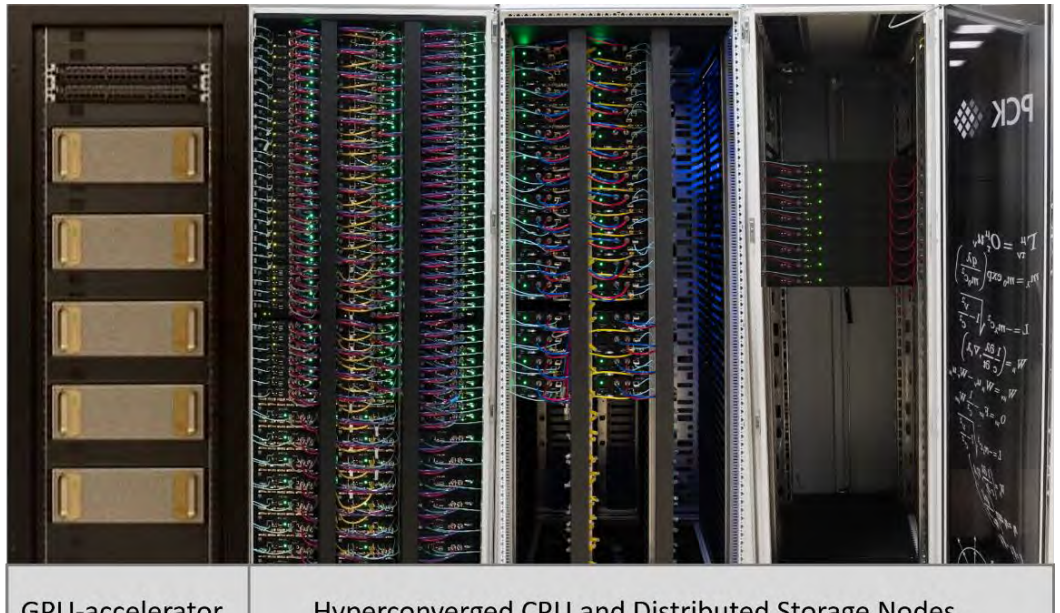
- Hyper-converged software-defined system
- Hierarchical data processing and storage system
- Scalable solution Storage-on-demand
- Total peak performance: 1.1 PFlops DP
- GPU component based on NVIDIA
- CPU component based on liquid cooling solutions
- The most energy-efficient center (PUE = 1.06)
- Storage performance >300 GB/s

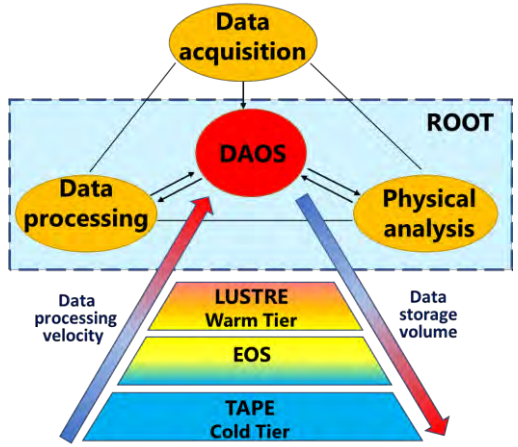
The expansion of the “Govorun” supercomputer by 32 hyperconverged compute nodes and 8 distributed storage nodes made it possible to:

- 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.

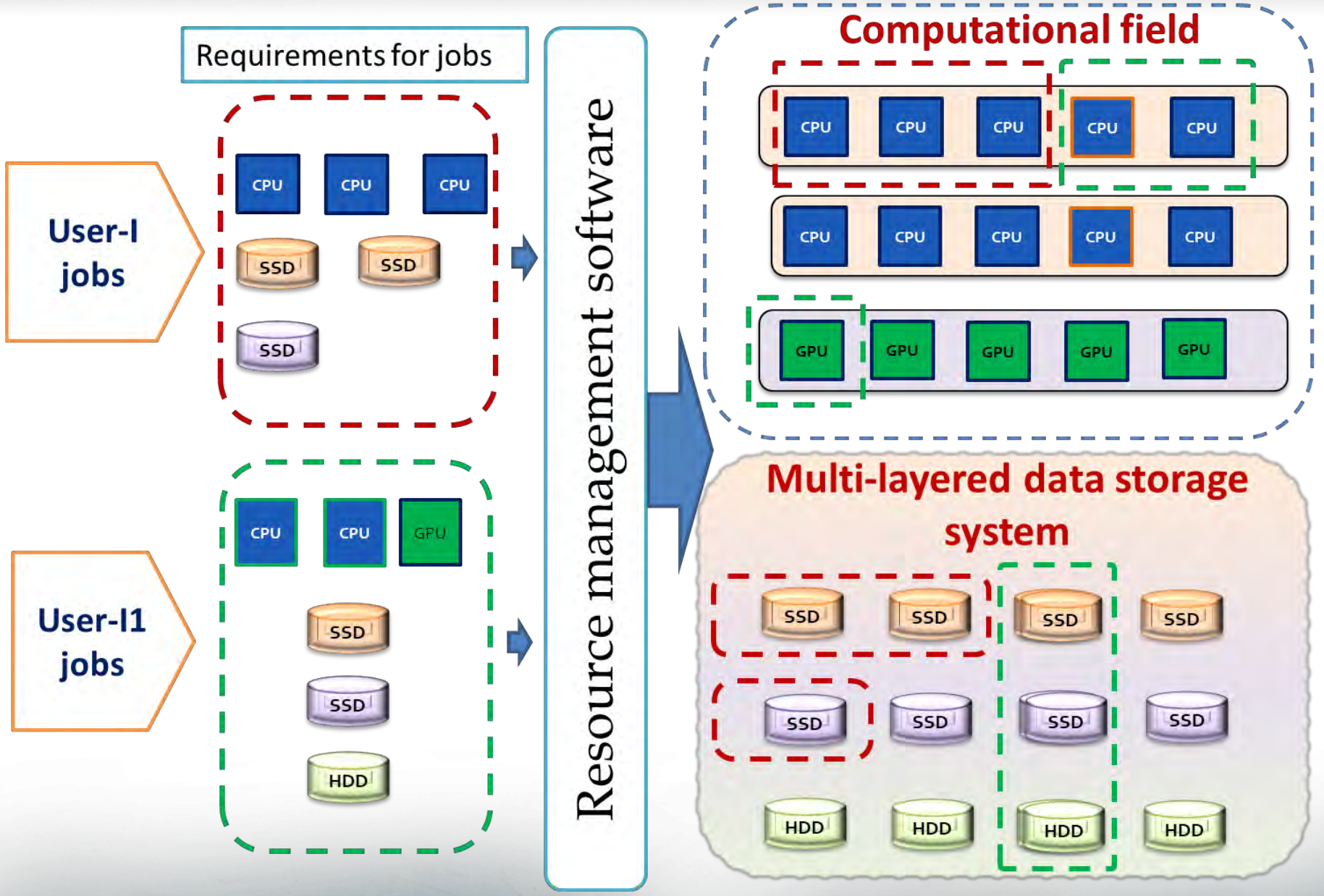
Key projects that use the resources of the SC “Govorun”:

- 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.





Orchestration and hyperconvergence on the SC “Govorun”



The SC “Govorun” has unique properties for the flexibility of customizing the user’s job.

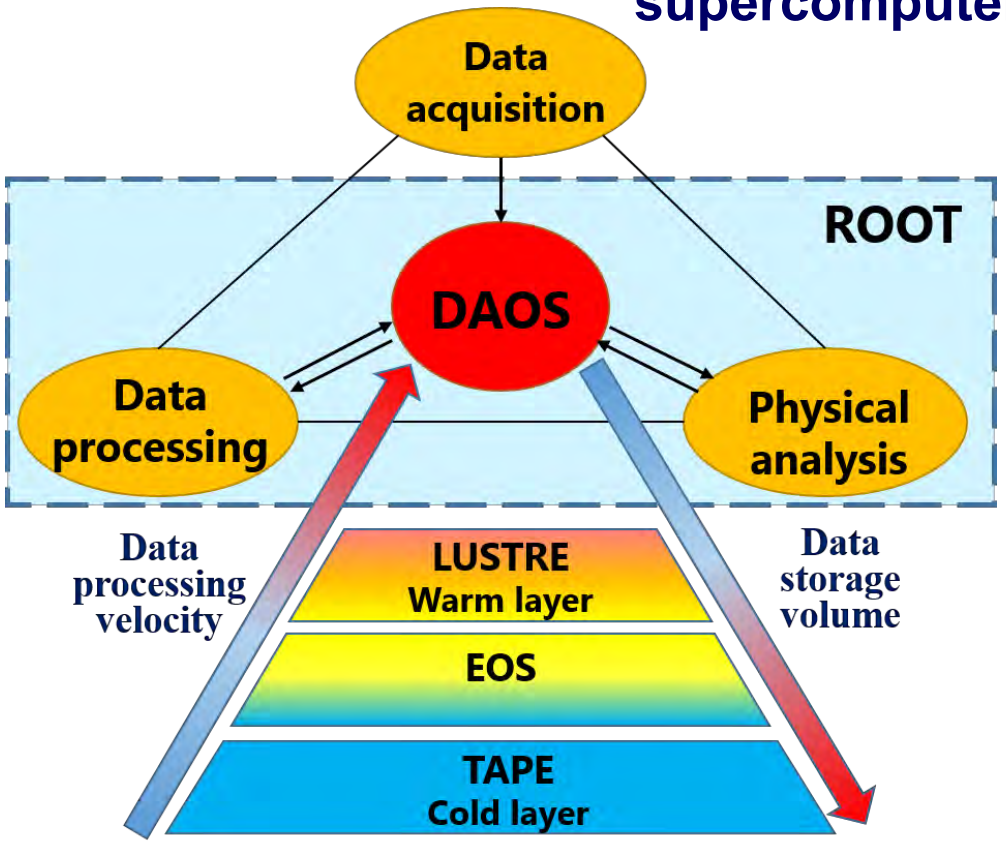
For his job the user can allocate the required number and type of computing nodes and the required volume and type of data storage systems.

This property enables the effective solution of different tasks, which makes the SC “Govorun” a unique tool for research underway at JINR.

DAOS polygon of the “Govorun” supercomputer for data processing and storage in experimental high-energy physics

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. The fastest layer of the hierarchical system is based on the latest DAOS (Distributed Asynchronous Object Storage) technology.

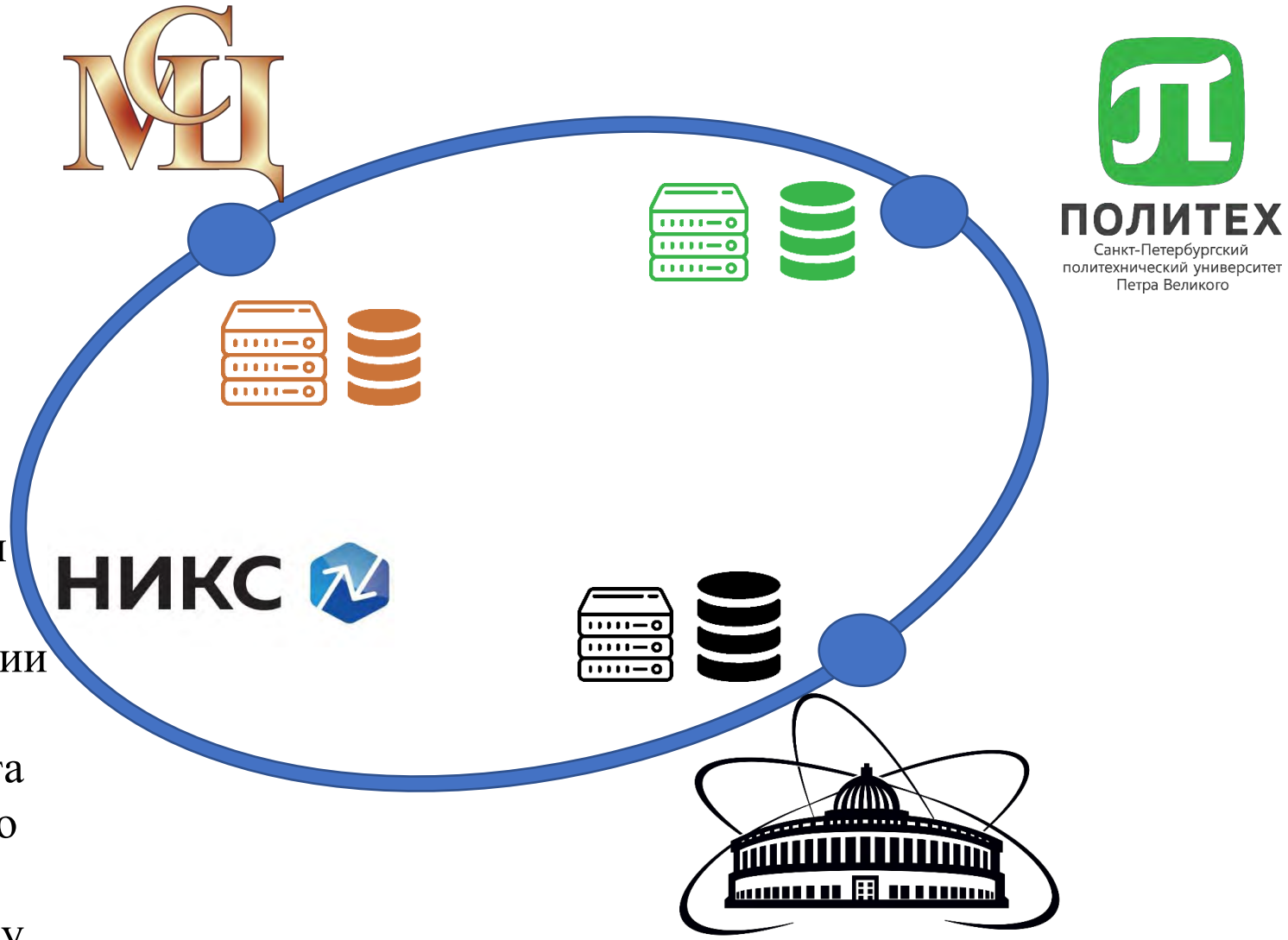
The DAOS polygon of the “Govorun” supercomputer takes the 1st place among Russian supercomputers in the current IO500 list.



The use of DAOS in high-energy physics enables to:

- Store and read multidimensional data structures of TB scale in a single address space
- Create a multi-user presentation layer for analyzing physics results
- Reduce hot storage costs in hundreds of times compared to using DDR (Double Data Rate) memory
- Significantly reduce the use of the GRID infrastructure (computing/storage/network) at the stage of physics analysis
- Easily integrate with other hot/warm storages

Объединенная географически распределенная суперкомпьютерная инфраструктура



В январе 2022 года успешно завершен первый совместный эксперимент по использованию объединенной суперкомпьютерной инфраструктуры для задач проекта NICA.

Всего было запущено 3000 задач генерации данных методом Монте-Карло и реконструкции событий для эксперимента MPD. Сгенерировано и реконструировано порядка 3 миллионов событий.

Полученные данные перемещены в Дубну для дальнейшей обработки и физического анализа

"Govorun" supercomputer modernization in 2022 - 2023

GPU-accelerator | Hyperconverged CPU and Distributed Storage Nodes

\+ 40 NVIDIA A100 GPU accelerators

Performance: + 600 Tflops DP

+32 hyperconverged compute nodes

+2 432 new computational cores

Performance: +239 Tflops DP

Performance “new cores”/”old cores” increase more than 1,5 times

+8 distributed storage nodes

Lustre, EOS increase: +8 PB

DAOS increase: +1.6 PB

+0,4 PB for the MPD mass production storages integrated into the DIRAC File Catalog

+1 PB for the MPD EOS storage

\+

5 servers with 8 NVidia A100 GPUs in each

\+

Computation field: +32 hyperconverged compute nodes

\+

Hierarchical Storage: +8 distributed storage nodes

SC “Govorun” total peak performance: 1.7 PFlops DP

Total capacity of Hierarchical Storage: 8.6 PB

Data IO rate: 300 Gb/s

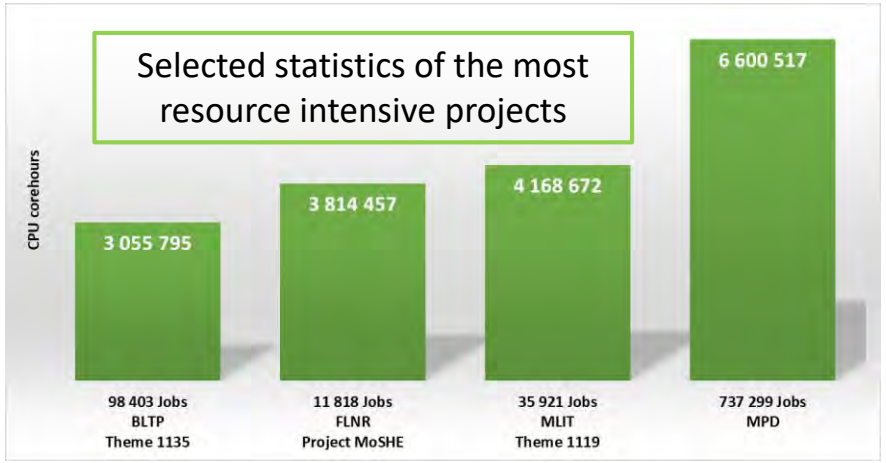
Using of the “Govorun” Supercomputer for JINR task in 2022

The projects that mostly intensive use the CPU resources of the “Govorun” supercomputer:

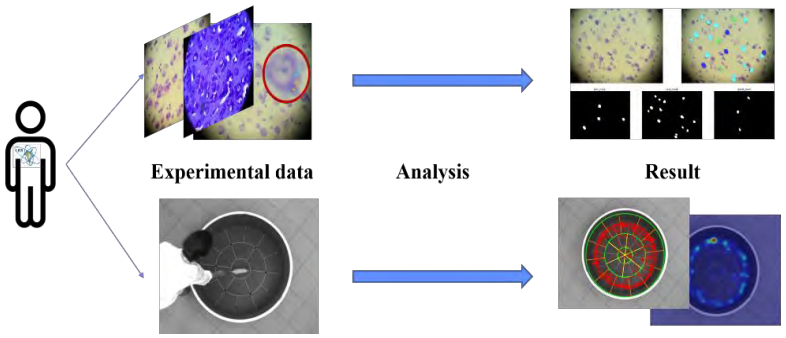
- NICA megaproject,
- simulation of complex physical systems,
- computations of the properties of atoms of superheavy elements,
- calculations of lattice quantum chromodynamics.

The GPU-component is activle used for solving applied problems by neural network approach:

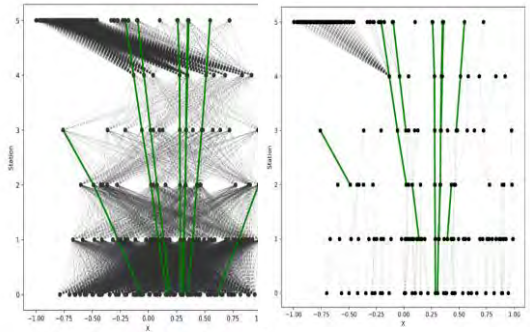
- process data from experiments at LRB,
- data processing and analysis at the NICA accelerator complex and ect.



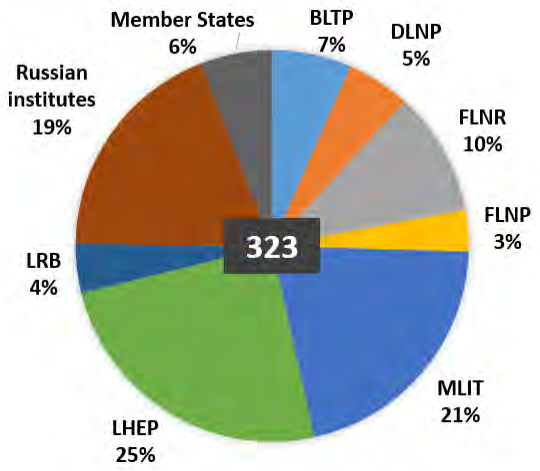
During 2022, **890 911** jobs were performed on the **CPU** component of the “Govorun” supercomputer, which corresponds to **18 543 076** core hours.



Information System for Radiation Biology Tasks

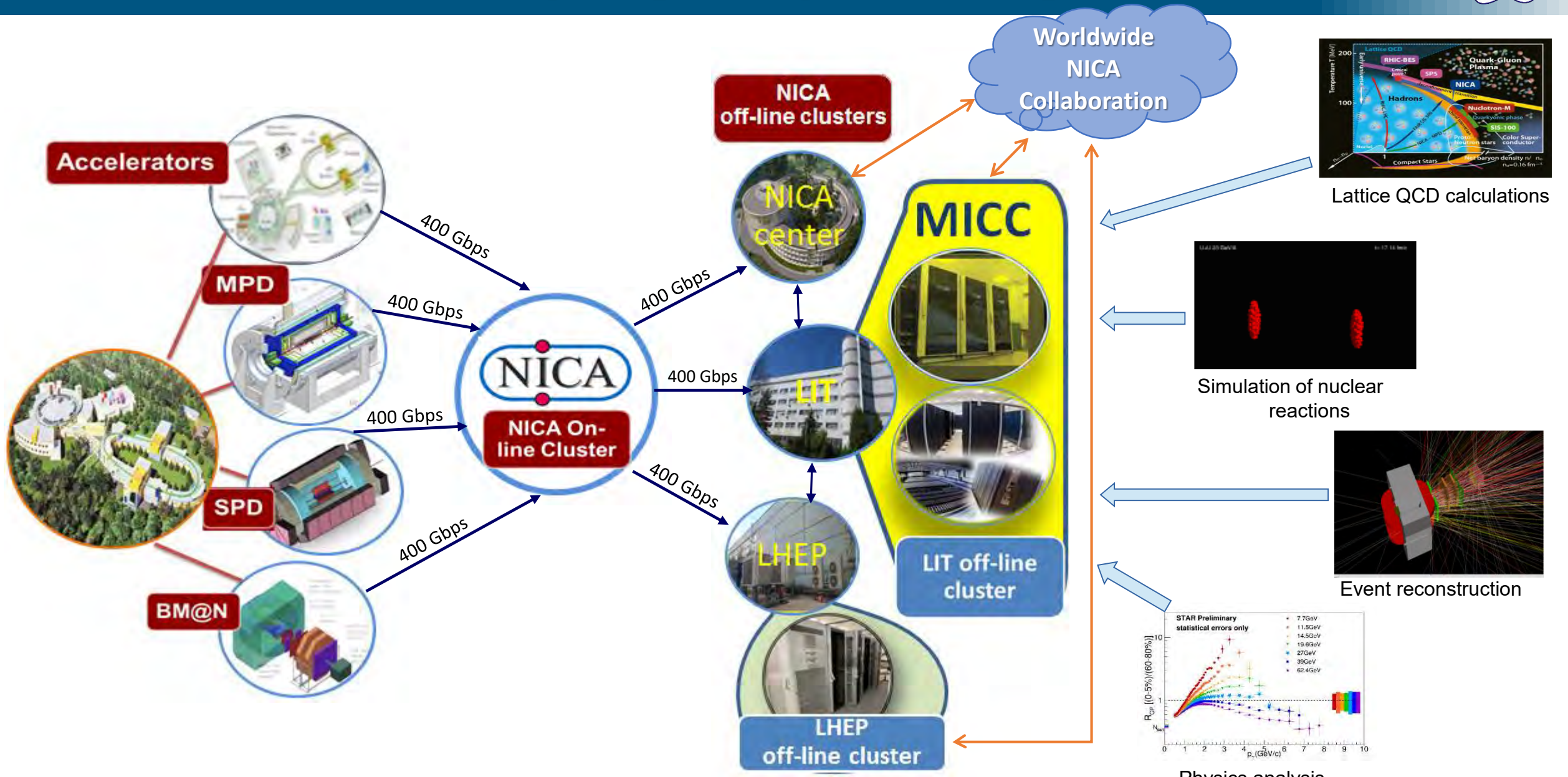


Neural network for data analysis

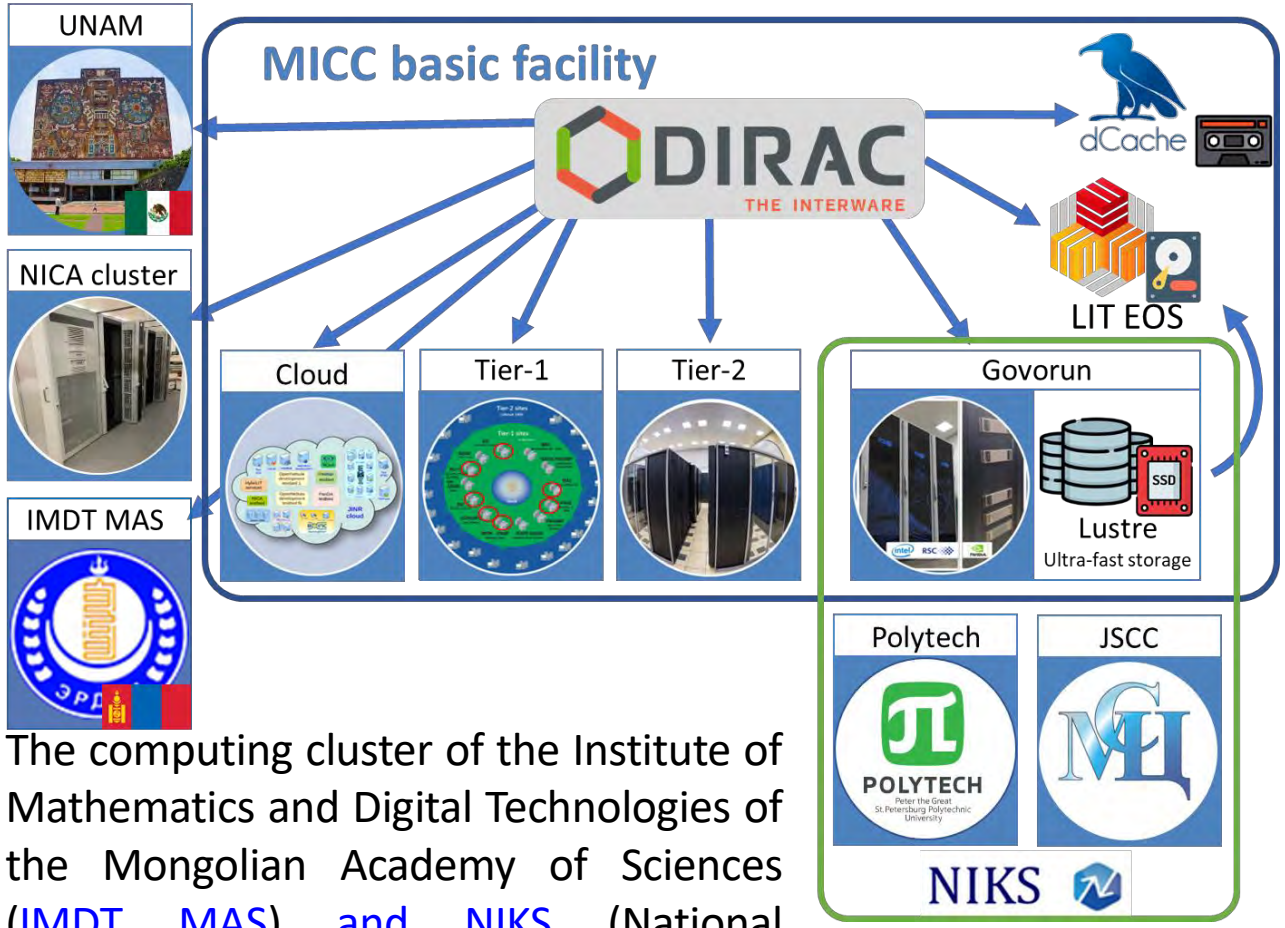


The resources of the “Govorun” spercomputer are used by scientific groups from all the Laboratories of the Institute within **25 themes** of the JINR Topical Plan.

NICA Computing Concept & Challenges

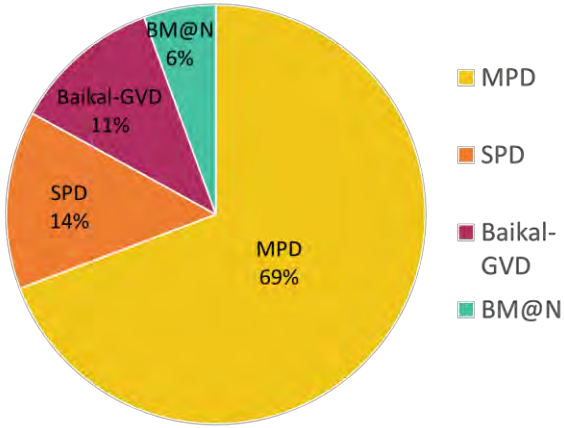


DIRAC-based distributed heterogeneous environment

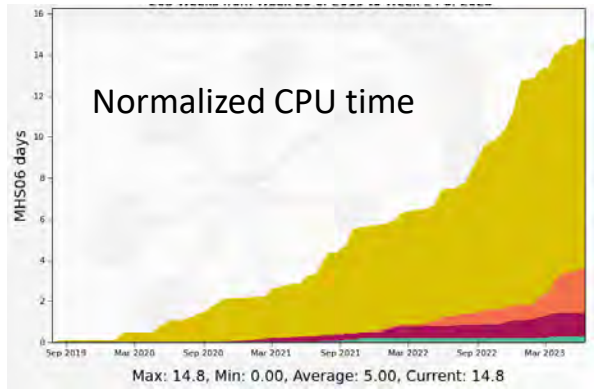


The computing cluster of the Institute of Mathematics and Digital Technologies of the Mongolian Academy of Sciences (IMDT MAS) and NIKS (National Research Computer Network, the Russia's largest research and education network) were integrated into the heterogeneous distributed environment based on the DIRAC platform.

Use of DIRAC platform by experiments in 2019-2022

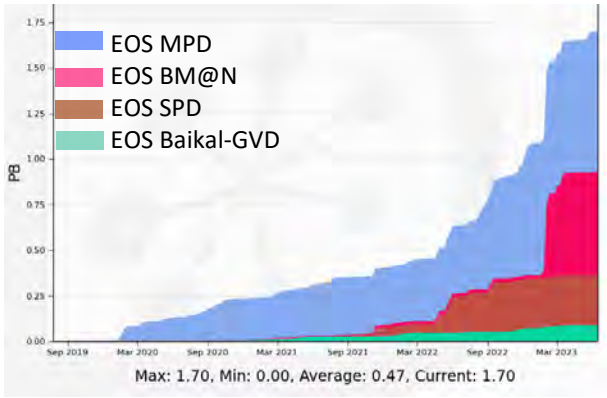


Total Number of executed jobs



The major user of the distributed platform is the MPD experiment

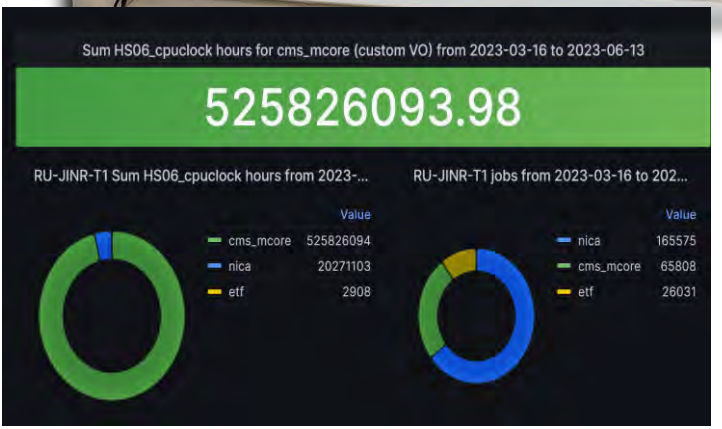
Data processed by experiments



Summary statistics of using the DIRAC platform for MPD tasks in 2019-2022

31 data generation campaigns	1.283 B events generated	440M events reconstructed	1.5M jobs completed	1557 years total computation time	1.3 PB MPD data produced

MICC Monitoring @Accounting



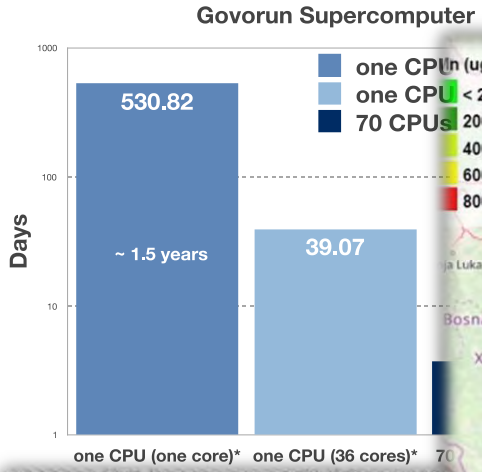
- ❖ 3 monitoring servers
- ❖ About 1800 nodes
- ❖ About 16000 service checks

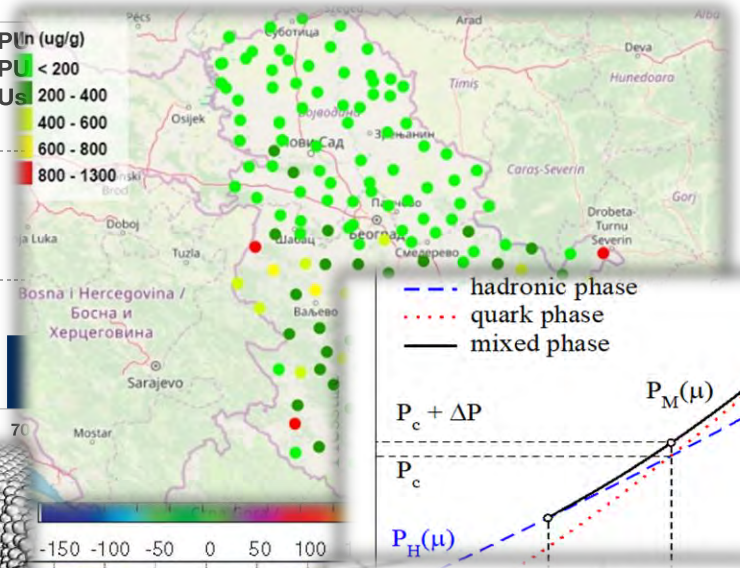
The successful functioning of the computing complex is ensured by the monitoring system of all MICC components/ We must

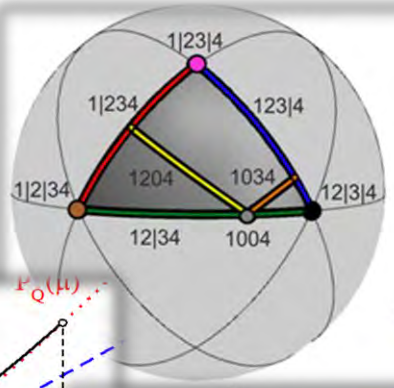
- ➢ to expand the monitoring system by integrating local monitoring systems for power supply systems into it (diesel generators, power distribution units, transformers and uninterruptible power supplies);
- ➢ to organize the monitoring of the cooling system (cooling towers, pumps, hot and cold water circuits, heat exchangers, chillers);
- ➢ to create an engineering infrastructure control center (special information panels for visualizing all statuses of the MICC engineering infrastructure in a single access point),
- ➢ to account every user job at every MICC component? account

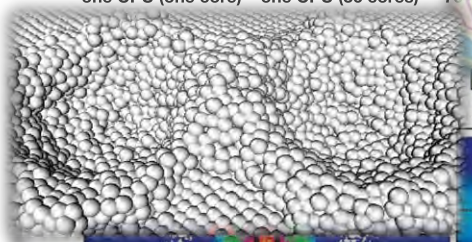
We must to develop intelligent systems that will enable to detect anomalies in time series on the basis of training samples, which will result in the need to create a special analytical system within the monitoring system to automate the process.

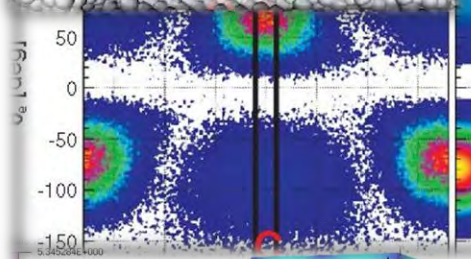
Methods, Algorithms and Software

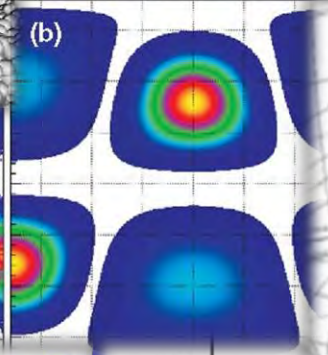


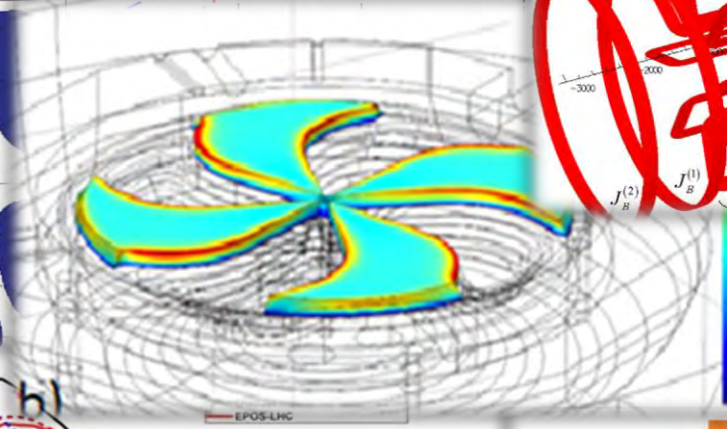


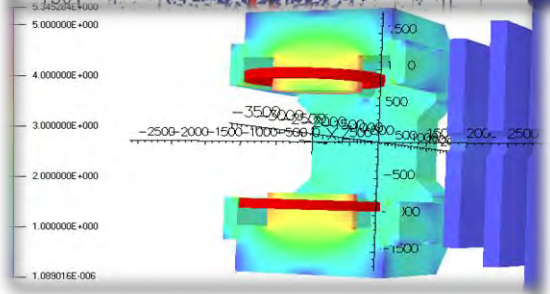


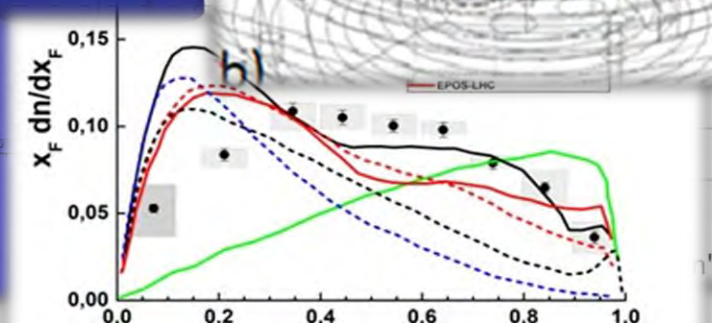


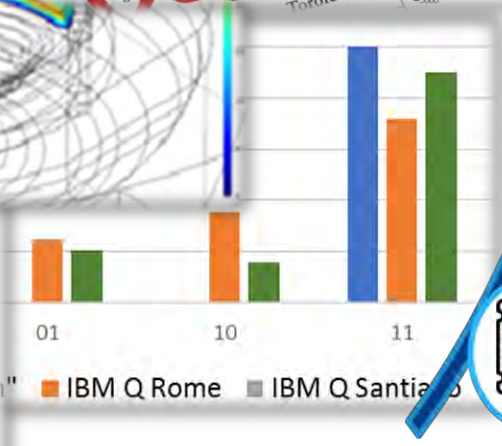












- Numerical modeling of complex physical systems
- Experimental data processing and analysis
- Big Data
- Machine and Deep learning
- AI and robotics
- Computer algebra
- Quantum computing

Activity: Digital ecosystem (Digital JINR)

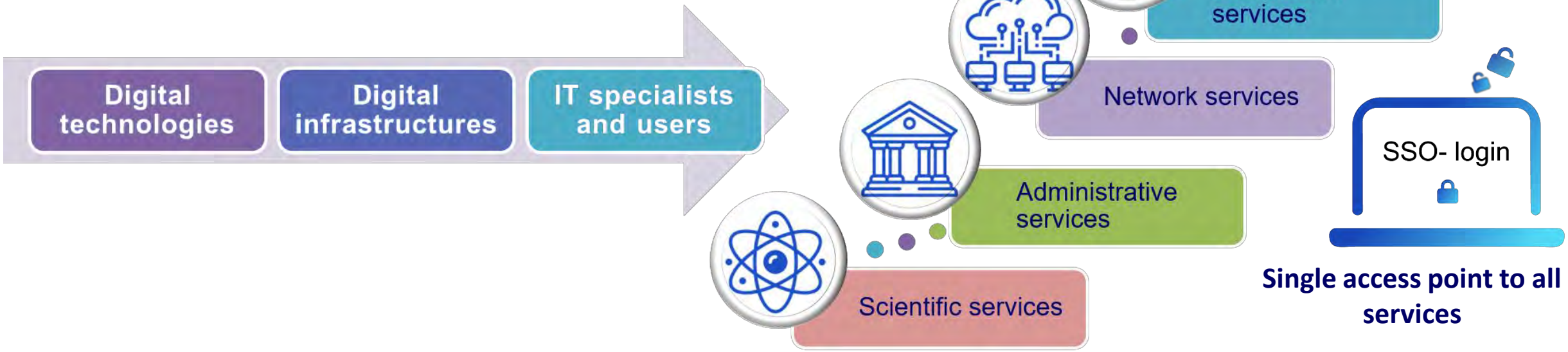
The digital platform "JINR Digital EcoSystem " integrates existing and future services

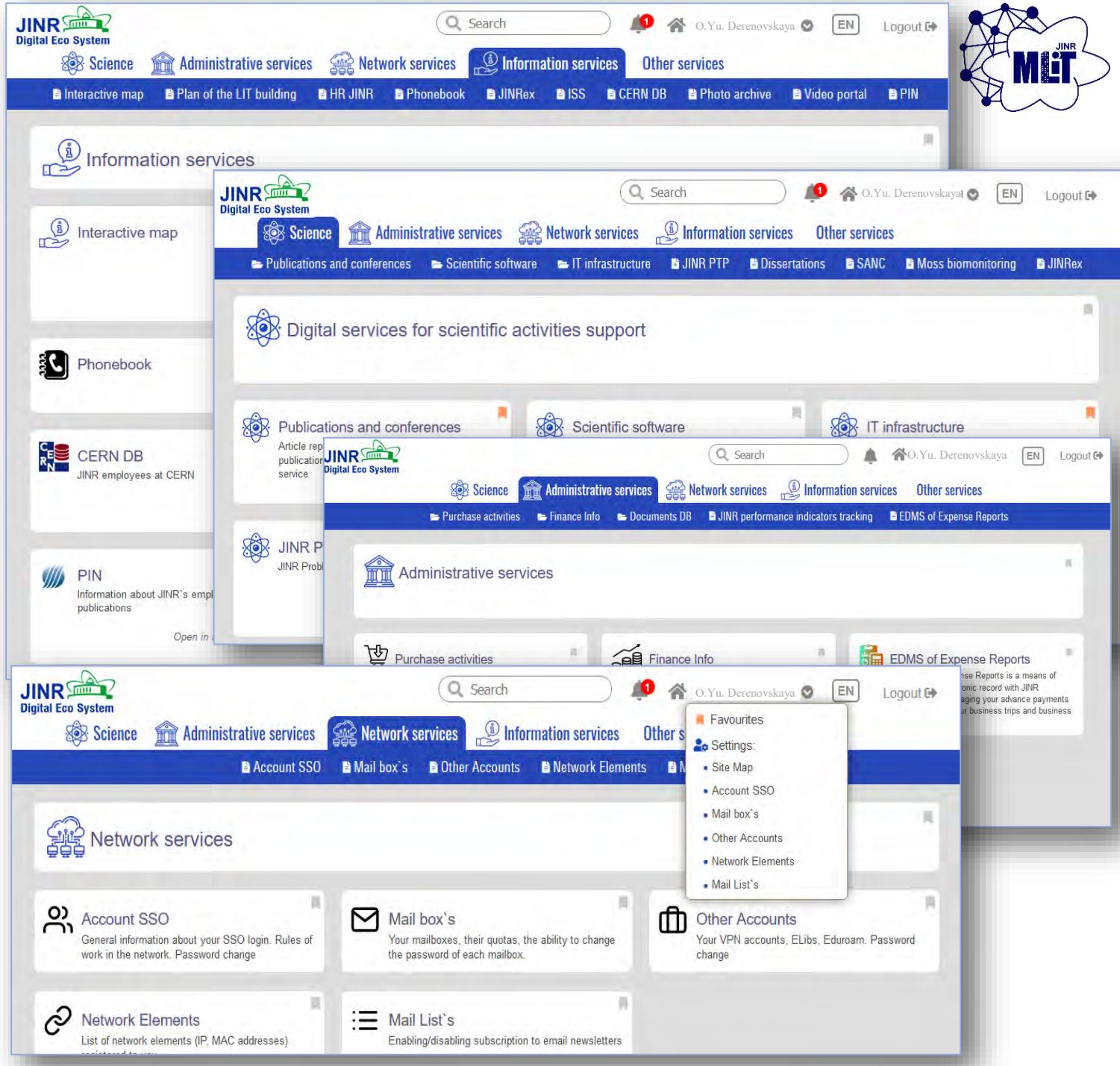
to support
scientific,
administrative and social activities,
maintenance of the engineering and IT infrastructures

to provide
reliable and secure access to various types of data

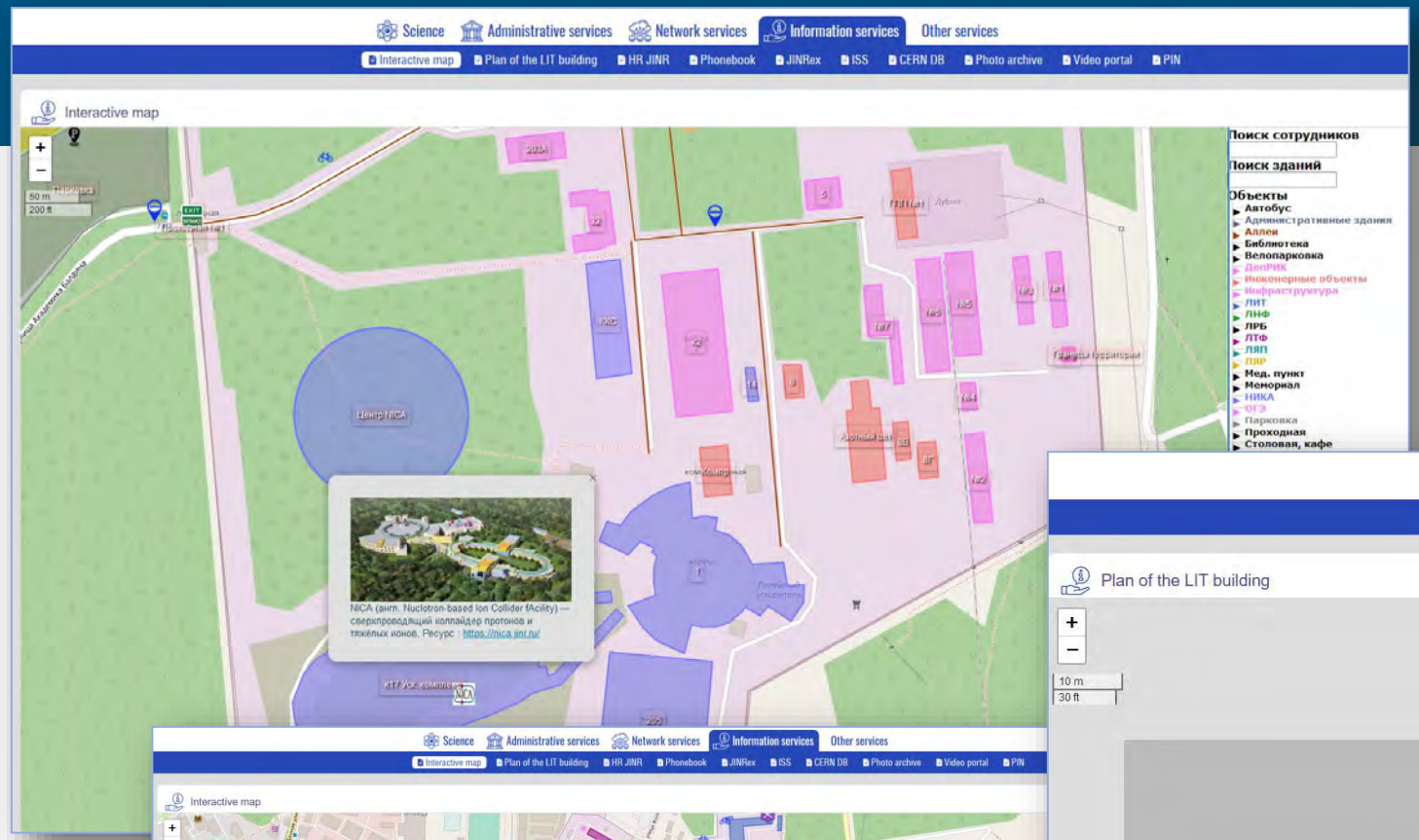
to enable
a comprehensive analysis of information

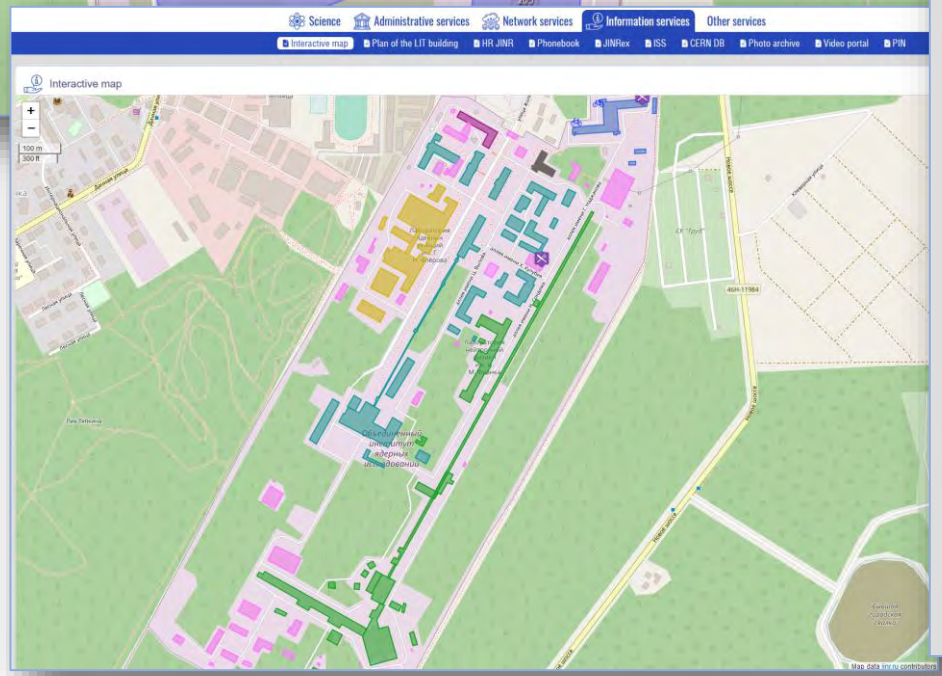
using
modern Big Data technologies and artificial intelligence.

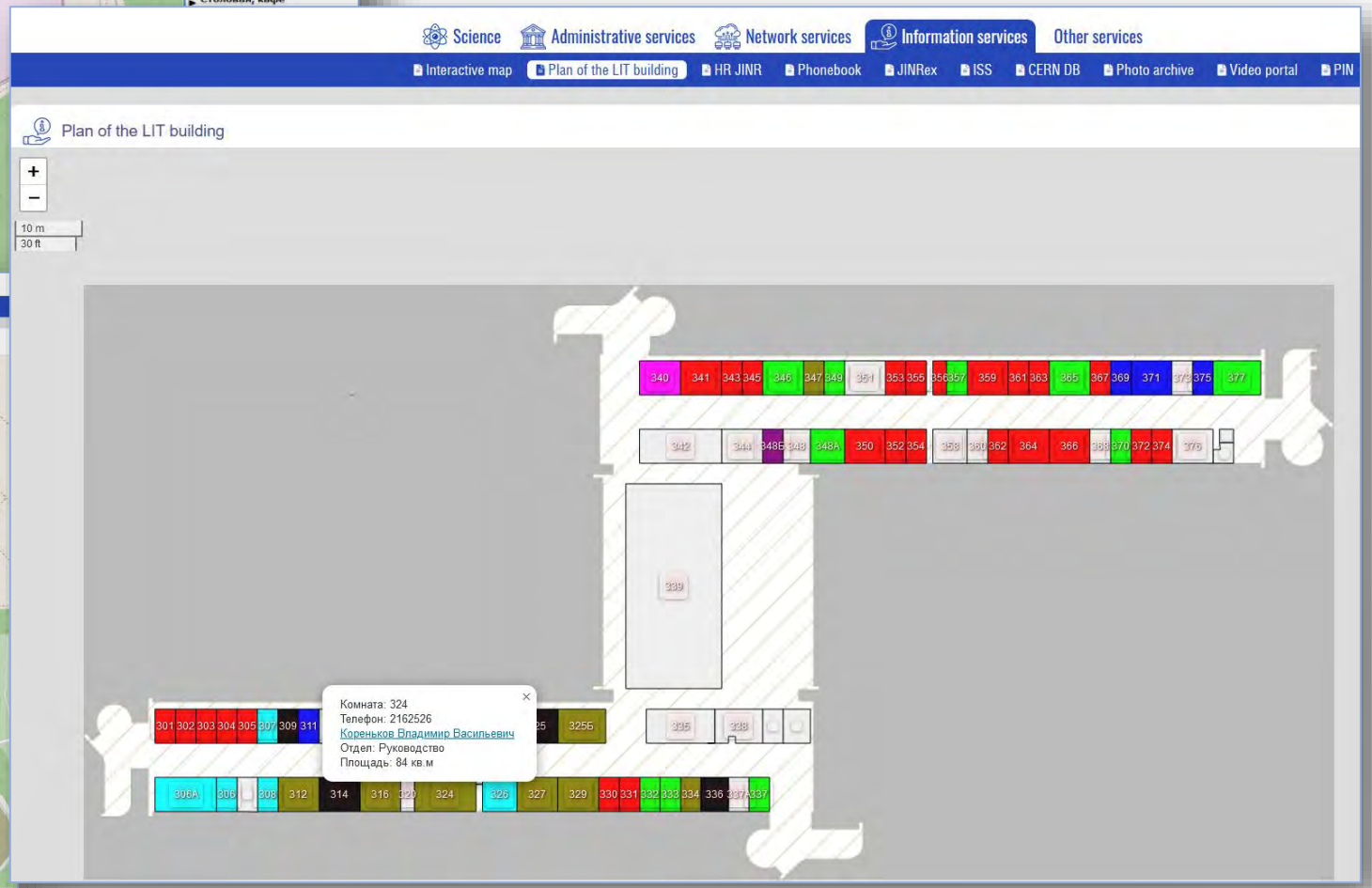




- ✓ Personal account of a JINR employee
- ✓ Notifications in a personal account
- ✓ Responsive interface, customizable by the user
- ✓ Easy access, convenient navigation and search for information on a large-scale network of a wide variety of JINR services

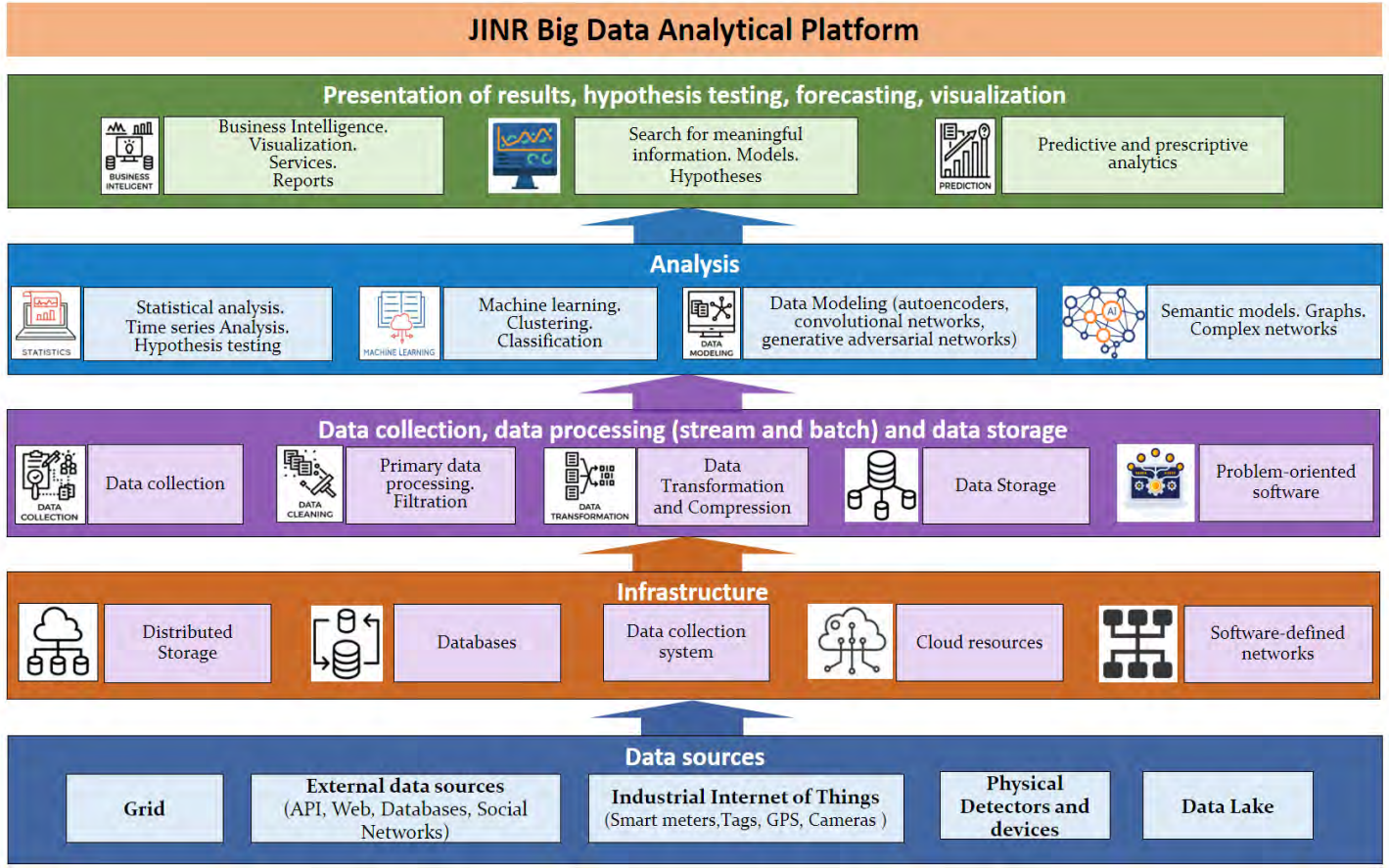






- ✓ Quick and easy search for information, both by services and by employees and buildings on an interactive JINR map

Activity: Multi-purpose Hardware and Software Platform for Big Data Analytics



Goal: the creation of a multi-purpose hardware and software platform for Big Data analytics based on hybrid hardware accelerators (GPU, FPGA, quantum systems); machine learning algorithms; tools for analytics, reports and visualization; support of user interfaces and tasks.

One of the tasks that is planned to be solved on the platform is the development of a unified analytical system for managing the MICC resources and data flows to enhance the efficiency of using computing and storage resources and simplify data processing within new experiments.

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

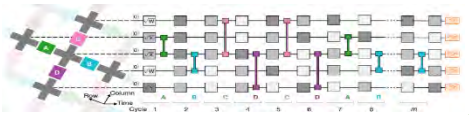
jupyterhub

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

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

Thank you for your attention

http://lit.jinr.ru