Large Research Infrastructure Project

Multifunctional Information and Computing Complex (MICC)

Leaders: **Korenkov V.V. ,**
Shmatov S.V.

Deputy Leaders: Dolbilov A.G.,
Podgainy D.V.,
Strizh T.A.

Information Technologies

Large research infrastructure project

Multifunctional Information and Computing Complex

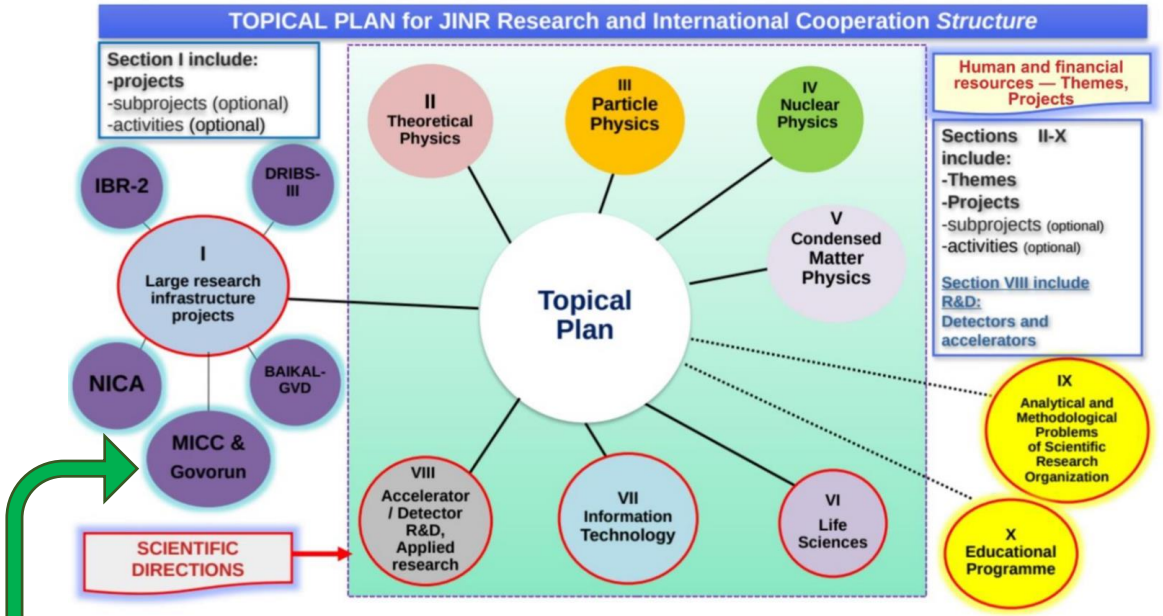
Leaders: Korenkov V.V. , Shmatov S.V.
Deputy Leaders: Dolbilov A.G., Podgainy D.V., Strizh T.A.

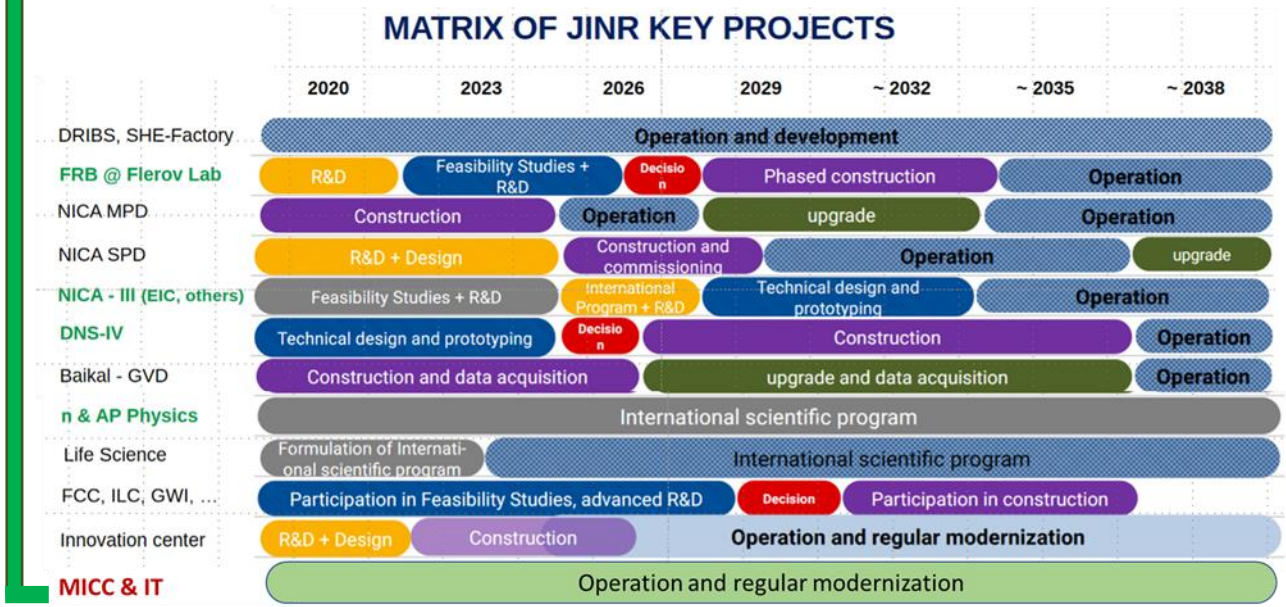
Activity
Multi-purpose hardware and software platform for Big Data analytics

Leader: Zrelov P.V .

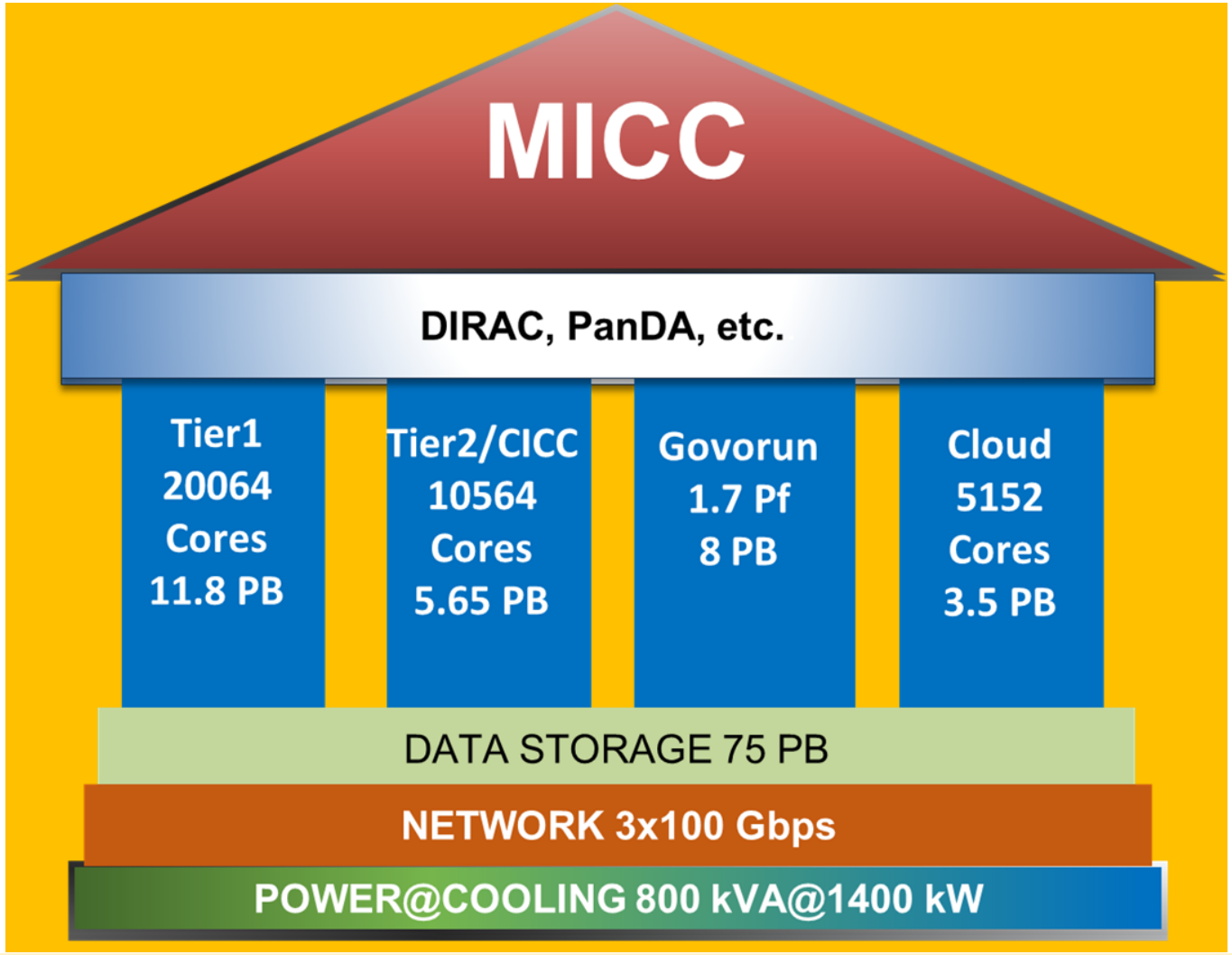
Activity
Digital ecosystem (Digital JINR)

Leaders: Korenkov V.V., Belov S.D.





Multifunctional Information and Computing Complex (MICC)



4 advanced software and hardware components

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

Distributed multi-layer data storage system

- Disks
- Robotized tape library

Network

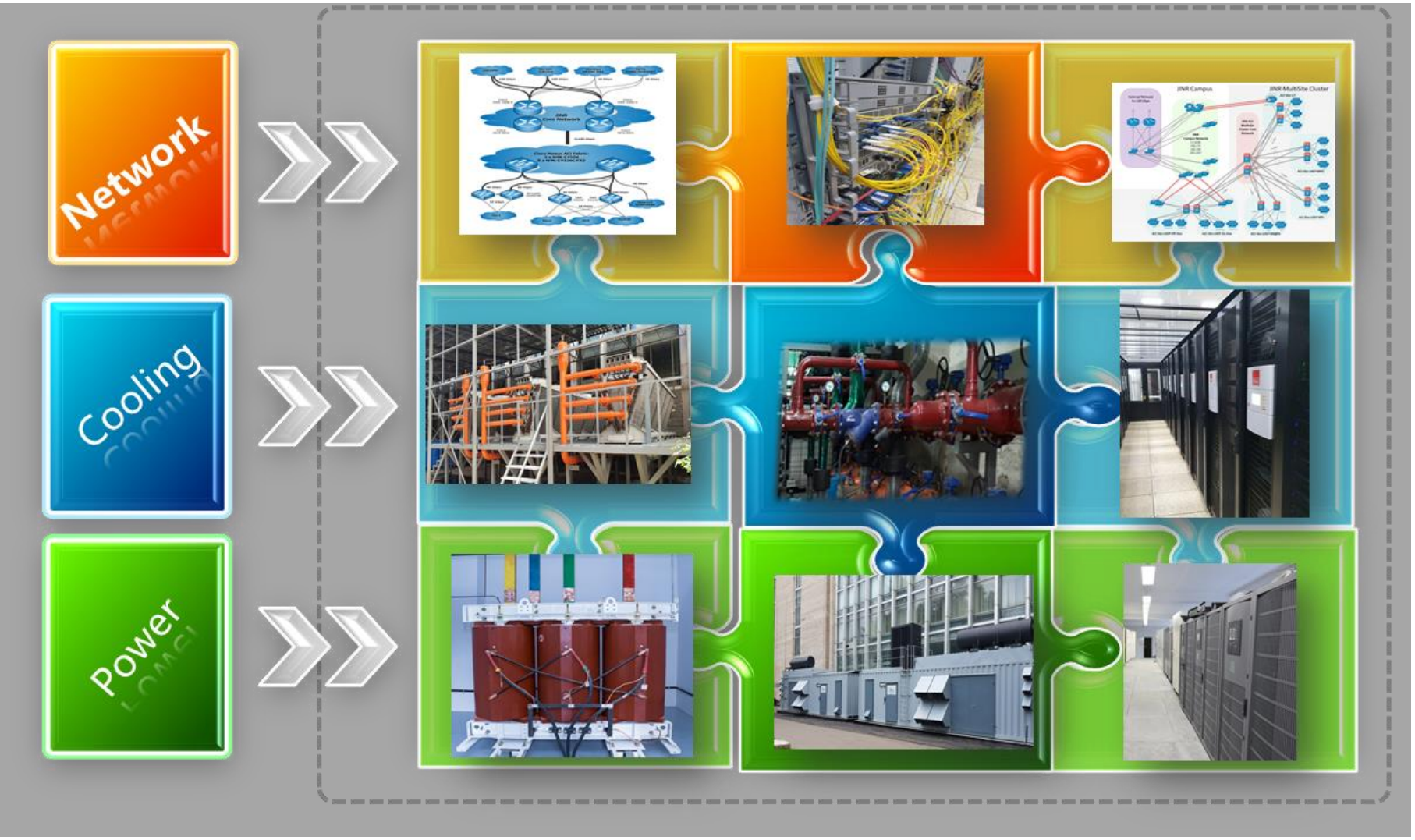
- Wide Area Network
- Local Area Network

Engineering infrastructure

- Power
- Cooling

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



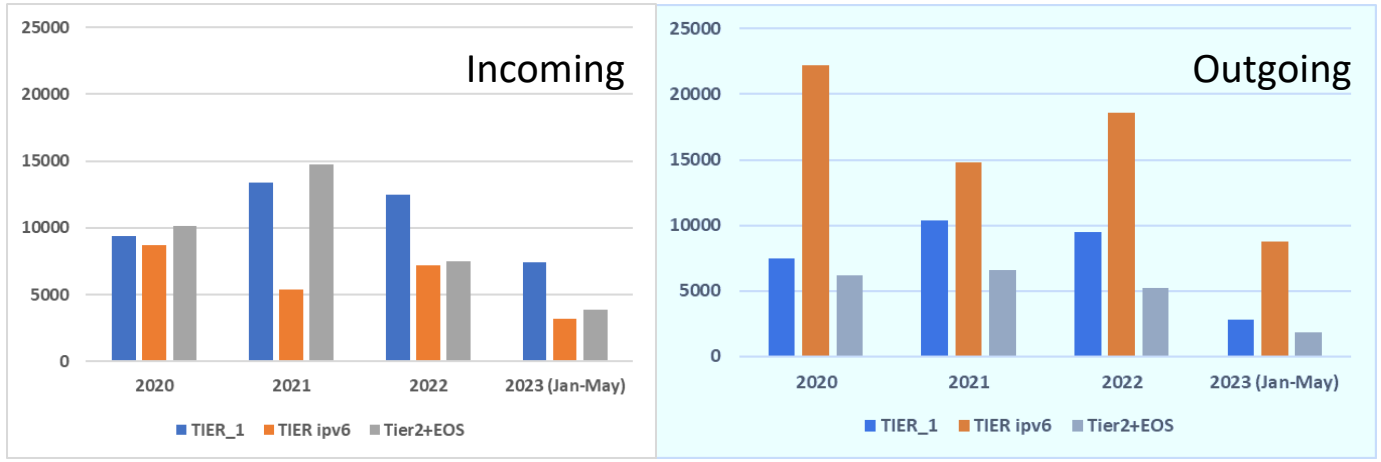
Wide Area Network 3x100 Gbps
Cluster Backbone 4x100 Gbps
Campus Backbone 2x100 Gbps

Dry chillers
In-Row systems
Total cooling 1400 kW

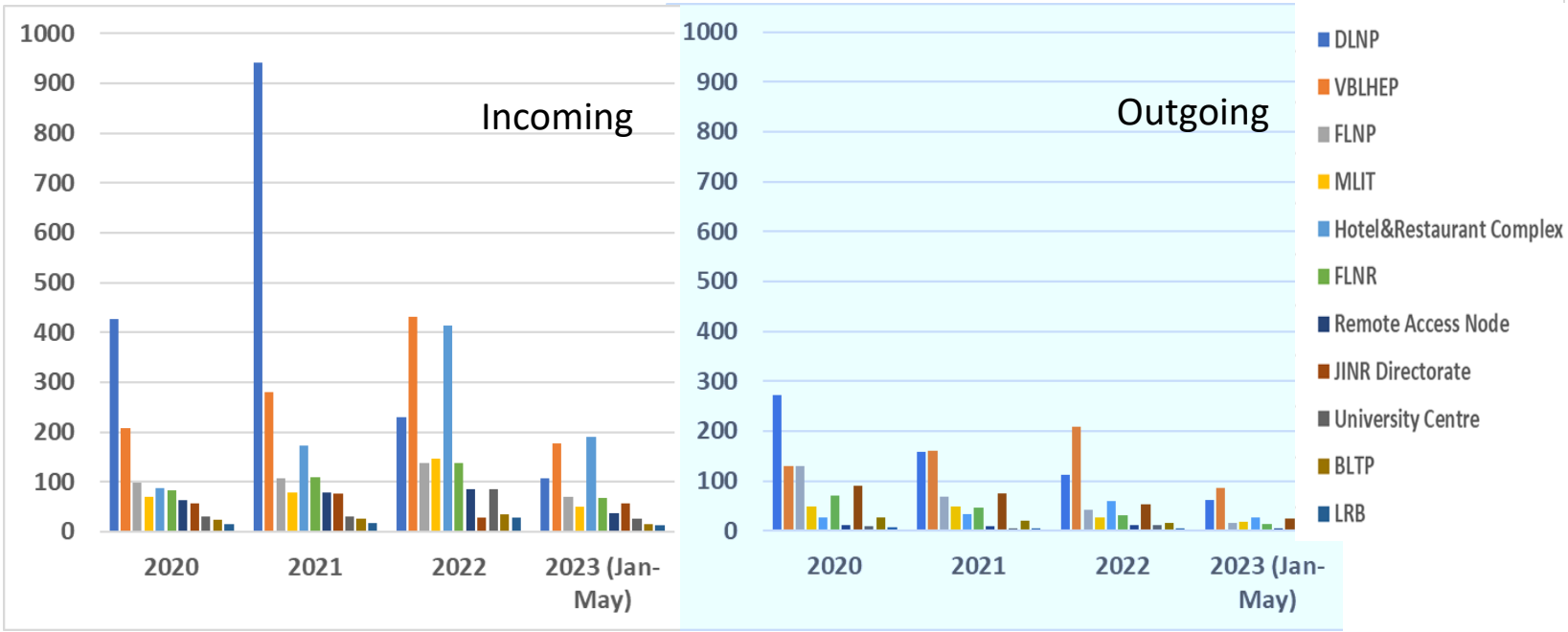
Uninterruptible power supplies (UPS) 8x300 kVA
Diesel-generator units (DGU) 2x1500 kVA
Transformers 2x2500 kVA

Networking @ Traffic

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

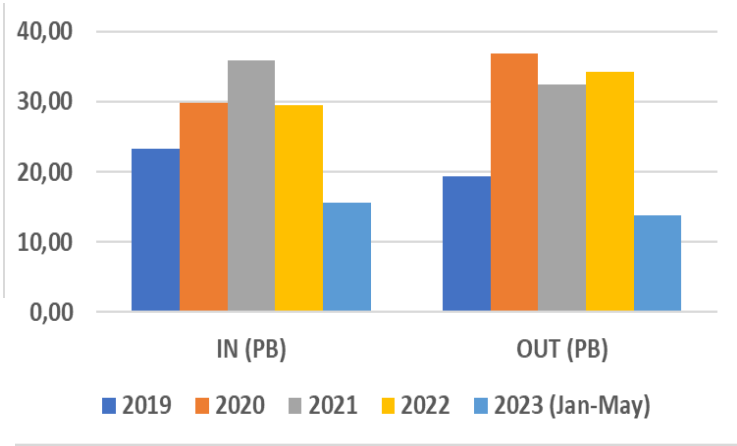


Distribution of the incoming and outgoing traffics by the JINR subdivisions in 2020-2023 (TB)



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

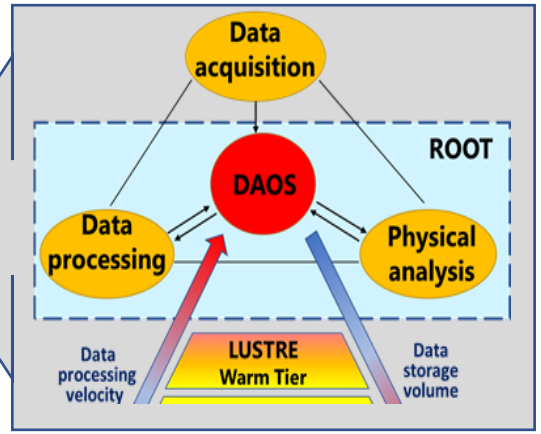
JINR Traffic in PB



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

Distributed Multi-layer Data Storage System

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



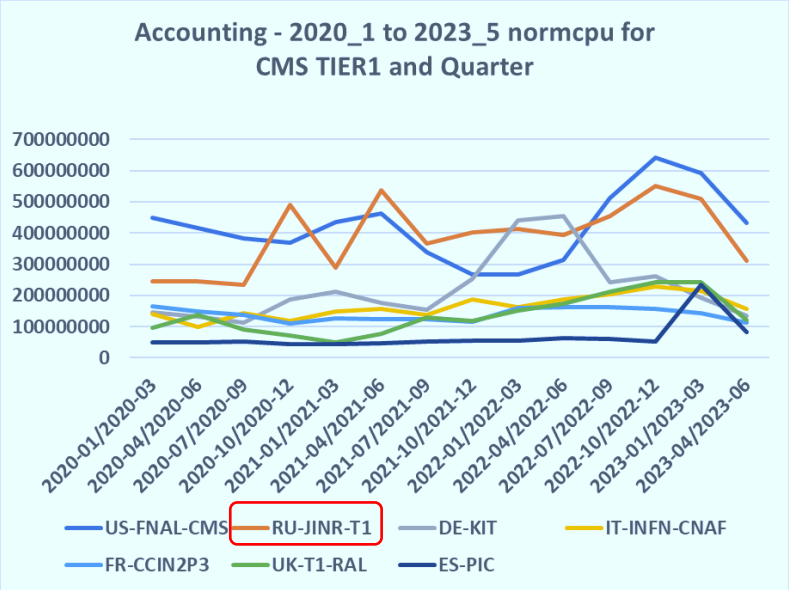
A special hierarchical data processing and storage system with a software-defined architecture was developed and implemented on the “Govorun” supercomputer.

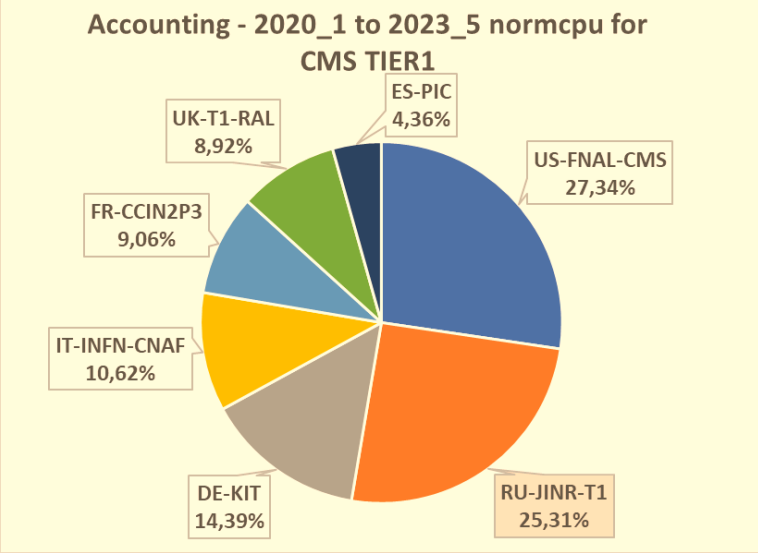
According to the speed of accessing data, there are the following layers:

- ✓ 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

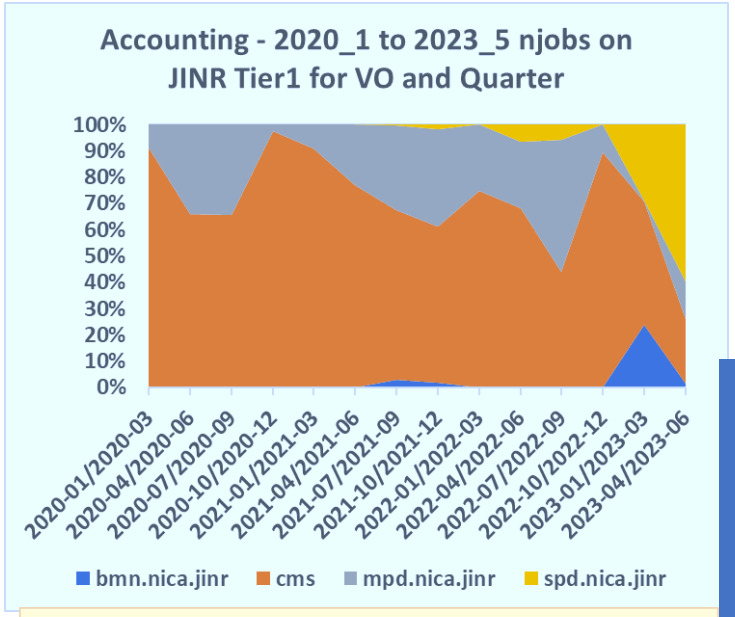


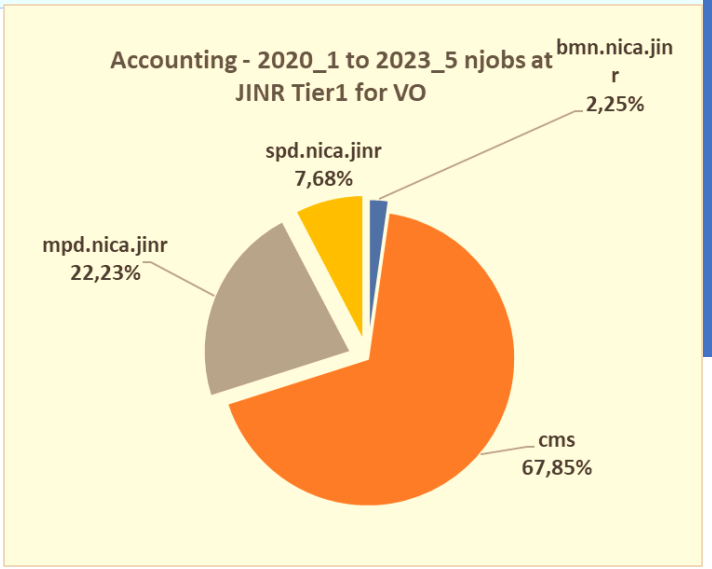


Since the beginning of 2015, a full-scale WLCG Tier1 site for the CMS experiment has been operating at MLIT JINR.

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.

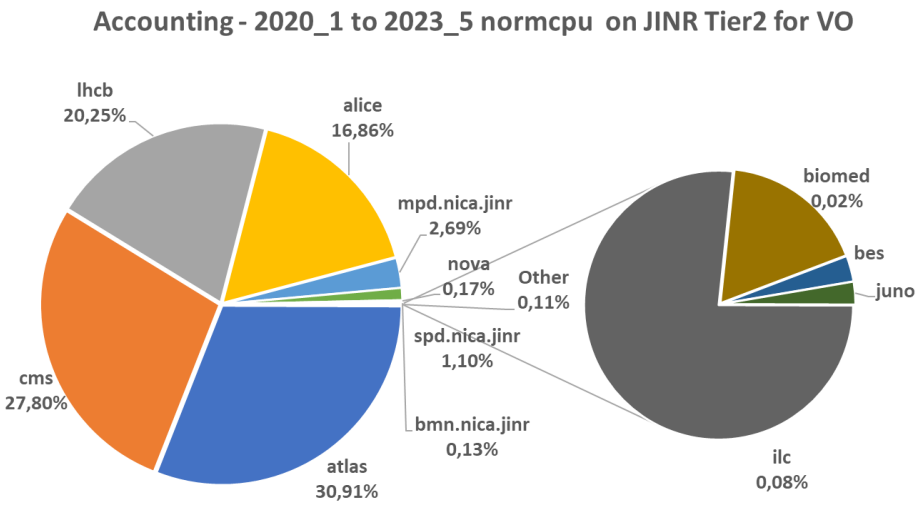
JINR Tier1 is regularly ranked on top among world Tier1 sites that process data from the CMS experiment at the LHC.



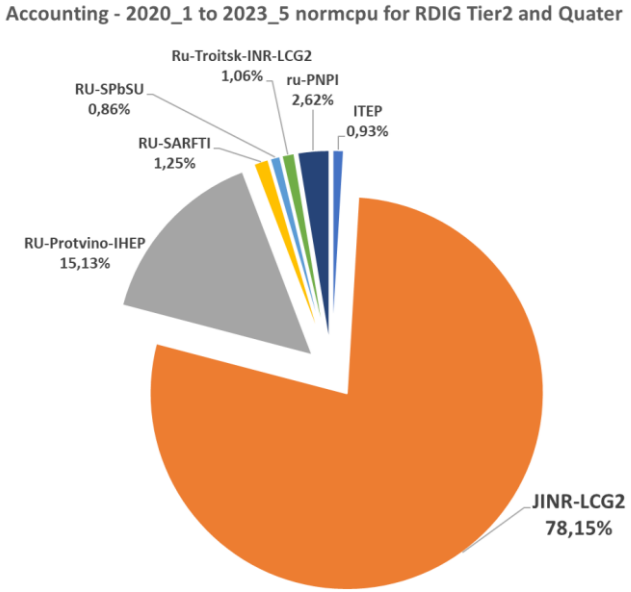


Since 2021, the JINR Tier1 center has demonstrated stable operation not only for CMS (LHC), but also for the NICA experiments.

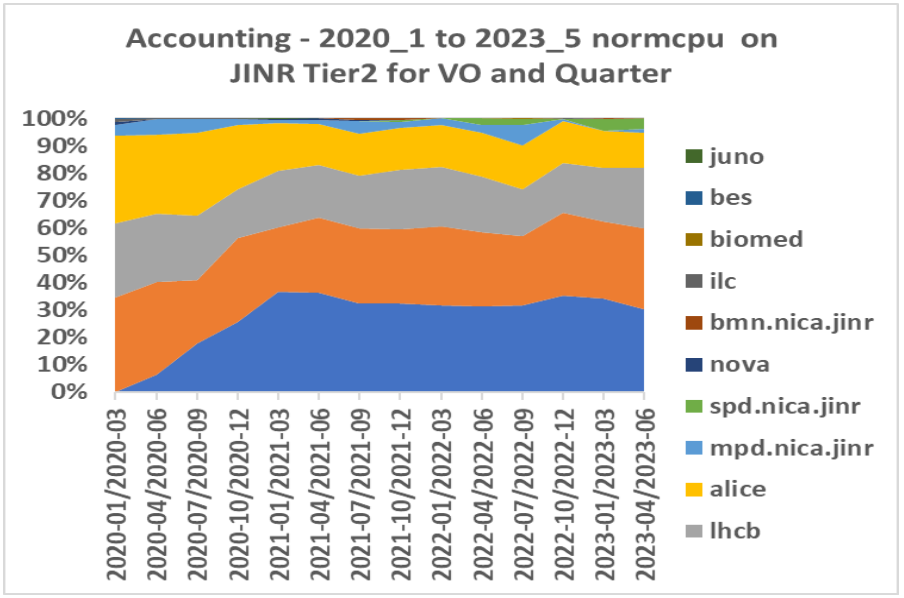
JINR Tier2 in WLCG & RDIG

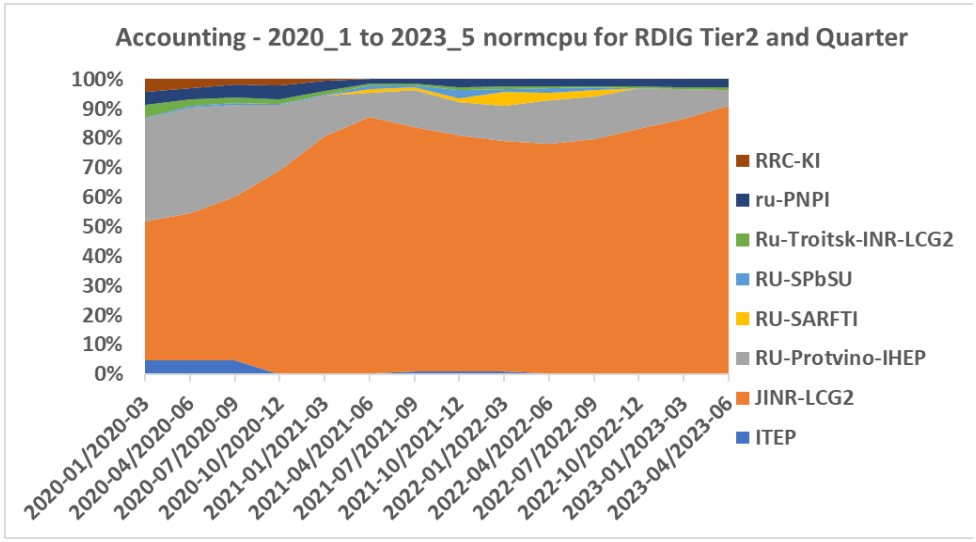


Tier2 at JINR provides computing power and data storage and access systems for the majority of JINR users and user groups, as well as for users of virtual organizations (VOs) of the grid environment (LHC, NICA, etc.).

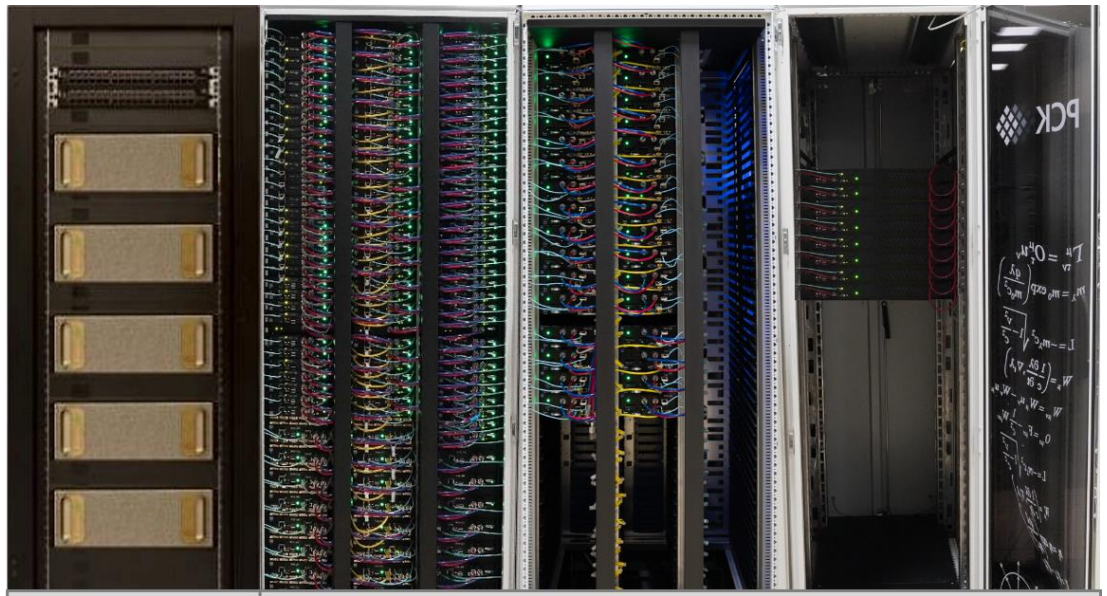


JINR Tier2 is the most productive in the Russian Data Intensive Grid (RDIG) Federation. More than 80% of the total CPU time in the RDIG is used for computing on our site.





"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

“New cores”/”old cores” performance increase by more than 1.5 times

\+8 distributed storage nodes

Lustre, EOS increase: +8 PB

DAOS increase: +1.6 PB

+0.4 PB for 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

“Govorun” SC total peak performance: 1.7 PFlops DP

Total capacity of Hierarchical Storage: 8.6 PB

Data IO rate: 300 Gb/s

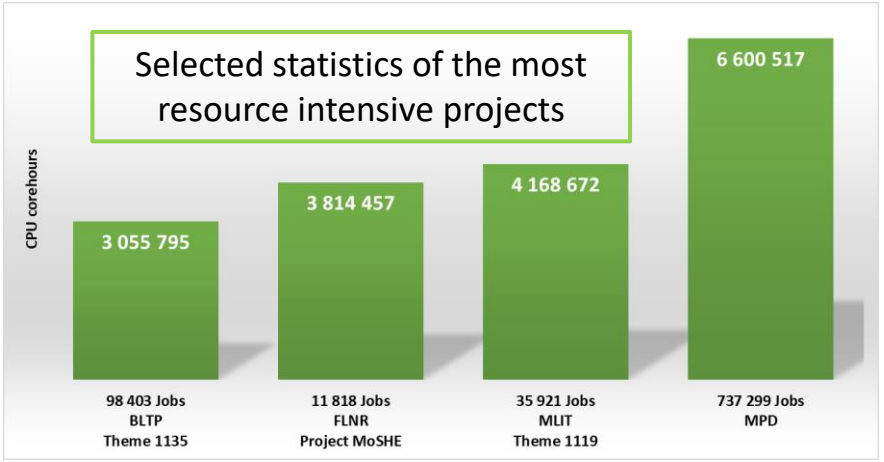
“Govorun” Supercomputer for JINR tasks in 2022

Projects that mostly intensively use the CPU resources of the “Govorun” supercomputer:

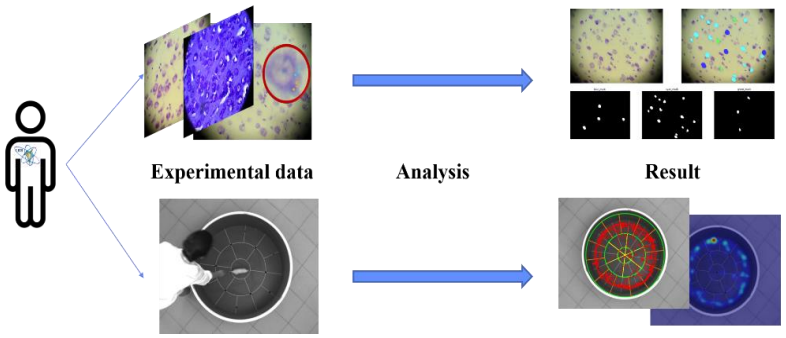
- NICA megaproject,
- simulation of complex physical systems,
- computations of the properties of atoms of superheavy elements,
- calculations of lattice quantum chromodynamics.

The GPU component is actively used for solving applied tasks by the neural network approach:

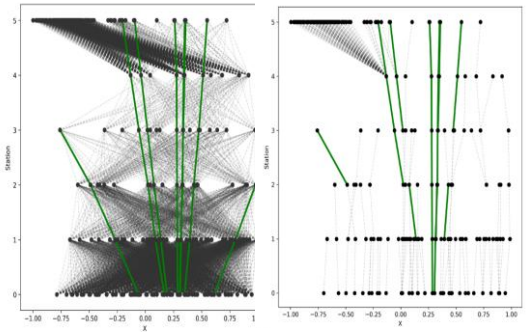
- processing of data from experiments at LRB,
- data processing and analysis at the NICA accelerator complex, etc.



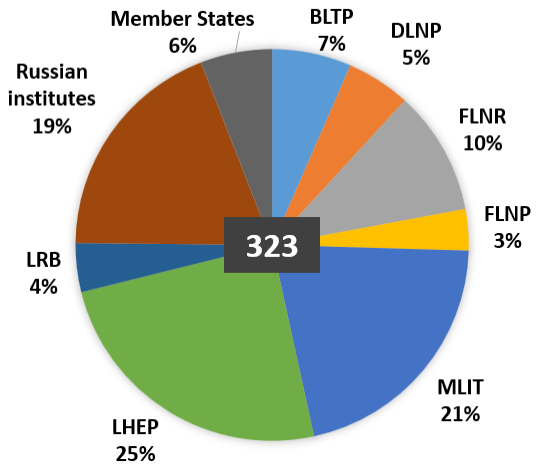
During 2022, **890 911** jobs were performed on the **CPU** component of the “Govorun” supercomputer, which corresponds to **18 543 076** core hours.



Information system for radiation biology tasks



Neural network for data analysis

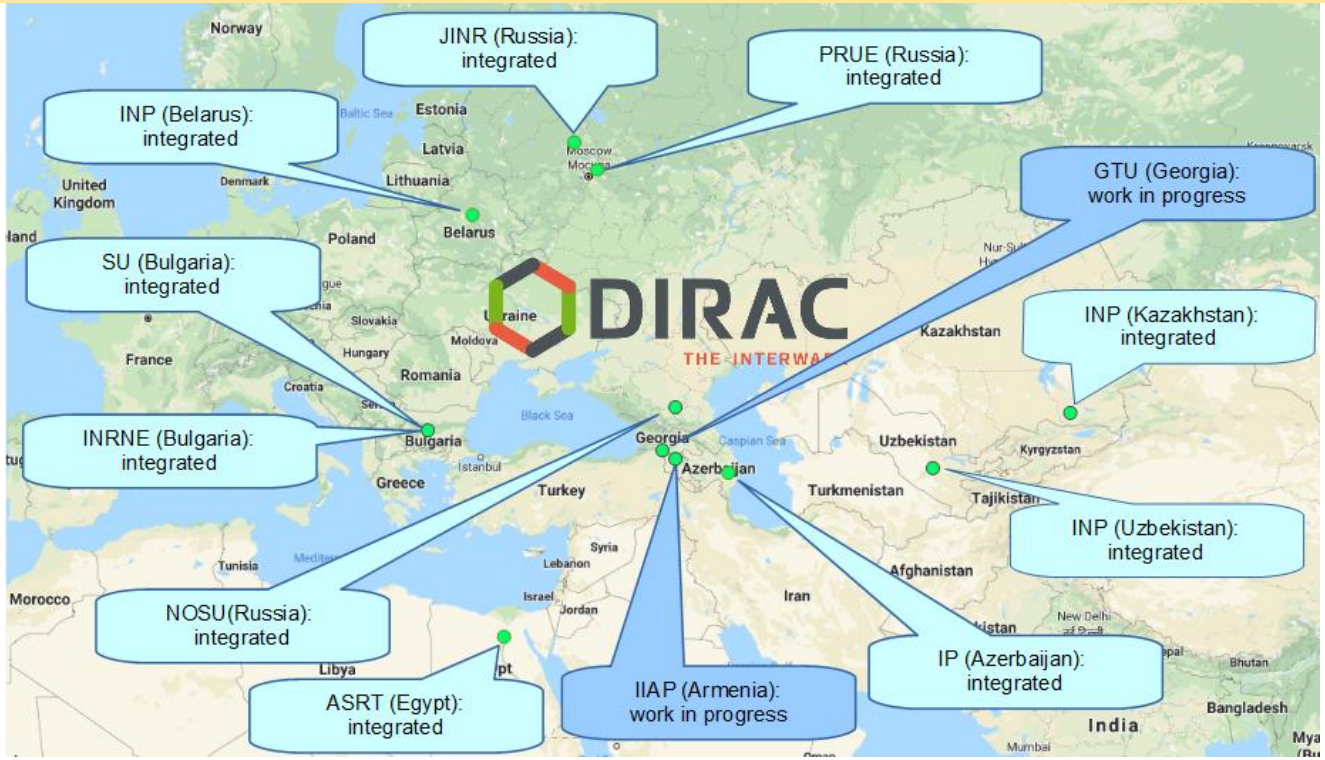


The resources of the “Govorun” spercomputer are used by scientific groups from all the Laboratories of the Institute within **25 themes** of the JINR Topical Plan.

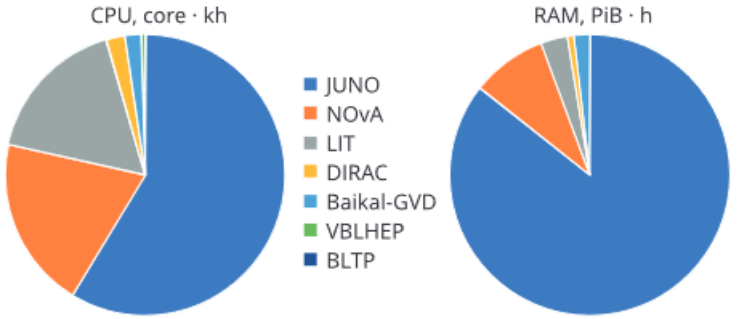
Cloud Infrastructure

- 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



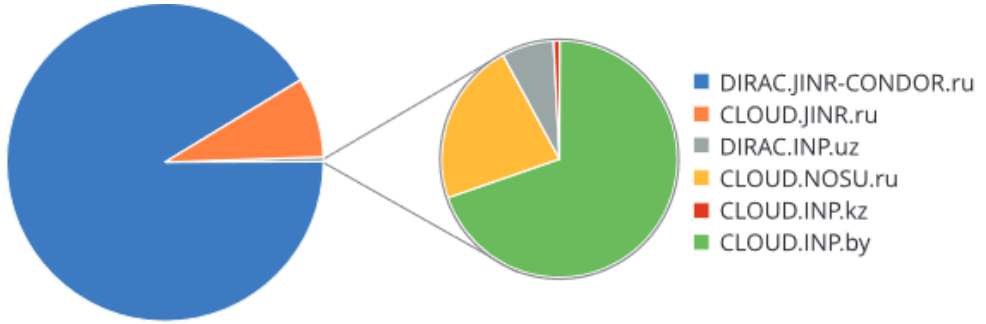
The Baikal-GVD, NOvA and JUNO experiments are the major users of the cloud infrastructure.



Usage 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



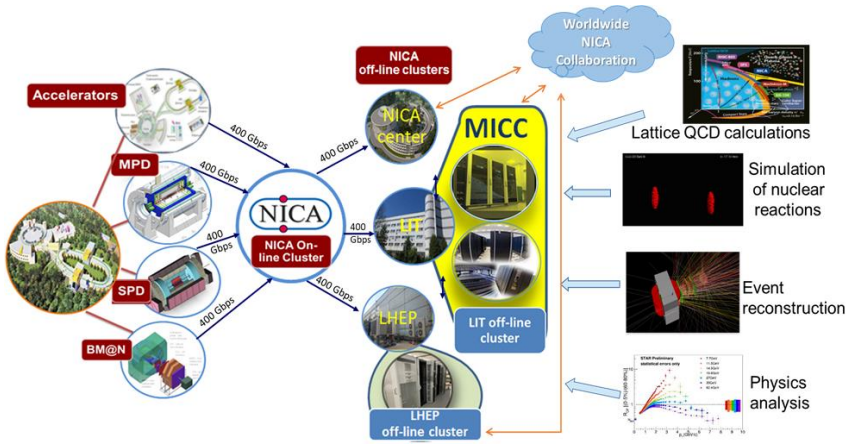
The main consumer of the JINR DICE resources in 2022 was the Baikal-GVD experiment (96%).

Development of the NICA Information and Computer Complex

The Seven-Year Plan provides for the creation of a long-term data storage center on the MICC resources at MLIT (Tier0). The process of modeling, processing and analyzing experimental data obtained from the BM@N, MPD and SPD detectors will be implemented in a distributed computing environment based on the MICC and the computing centers of VBLHEP and collaboration member countries.

The information and computer unit of the NICA complex embraces:

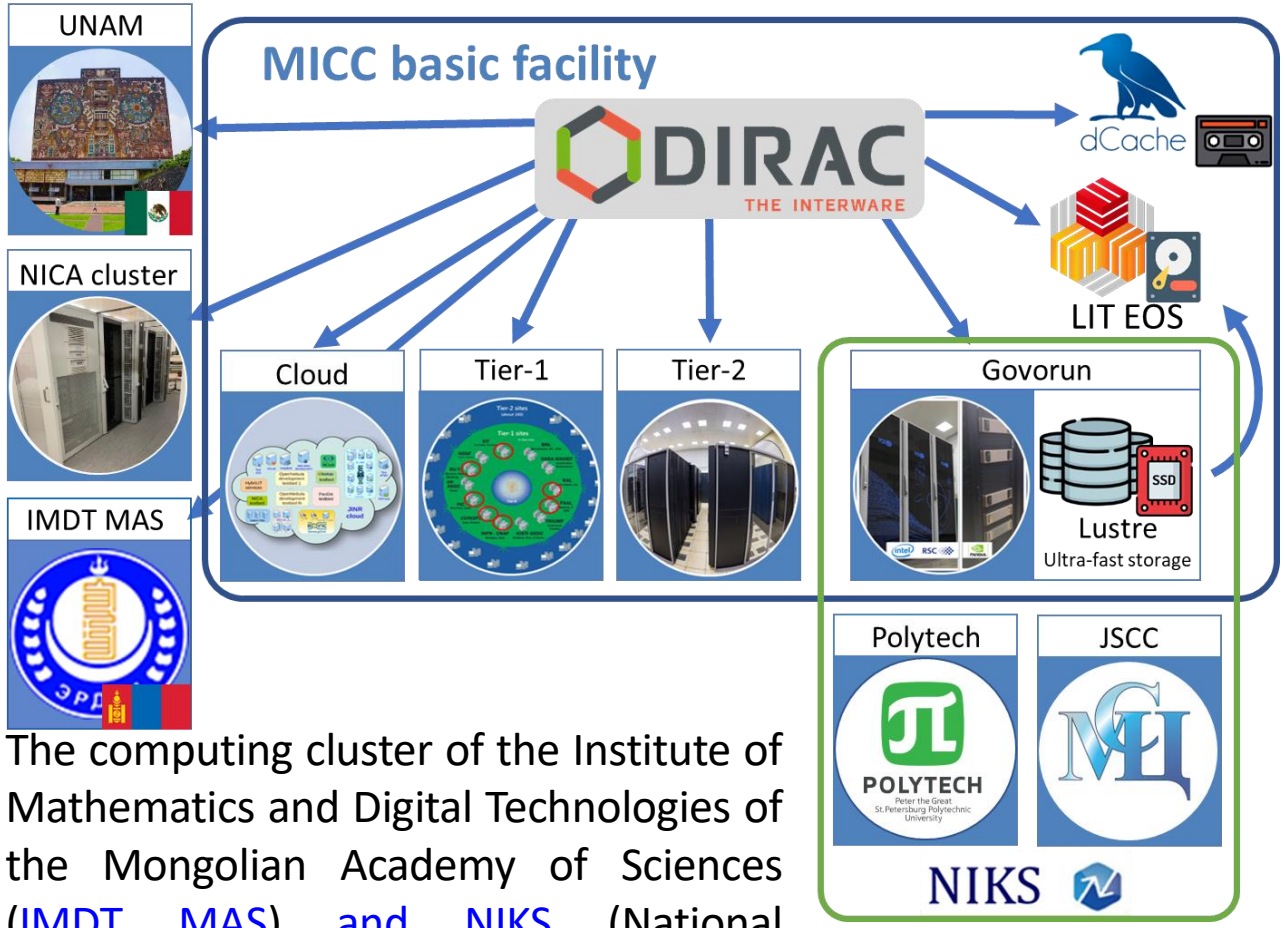
1. **online NICA cluster**,
2. **offline NICA cluster at VBLHEP**,
3. **all MICC components** (Tier0, Tier1, Tier2, "Govorun" supercomputer, cloud computing),
4. multi-layer **data storage system**,
5. **distributed computing network**.



NICA Tier 0,1,2	2024	2025	2026	2027	2028	2029	2030
CPU (PFlops)	2.2	2.6	8.6	8.6	15.6	15.6	15.6
DISK (PB)	17	24	47	75	96	119	142
TAPE (PB)	45	88	170	226	352	444	536
NETWORK (Gbps)	400	400	800	800	800	1000	1000

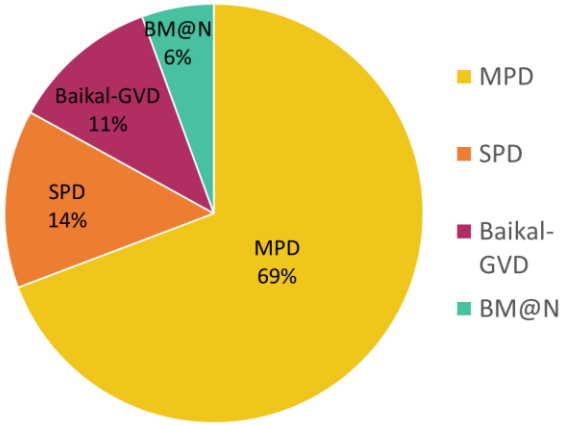
It should be underlined that the resources given in the table can be approximately satisfied by 20-25% of the budget allocated for the MICC.

DIRAC-based distributed heterogeneous environment

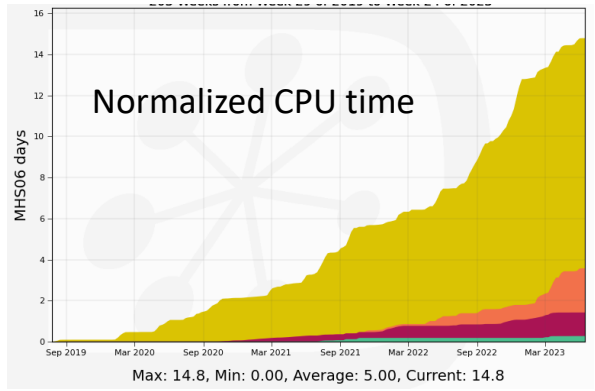


The computing cluster of the Institute of Mathematics and Digital Technologies of the Mongolian Academy of Sciences (IMDT MAS) and NIKS (National Research Computer Network, Russia's largest research and education network) were integrated into the heterogeneous distributed environment based on the DIRAC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

Usage of the DIRAC platform by experiments in 2019-2022

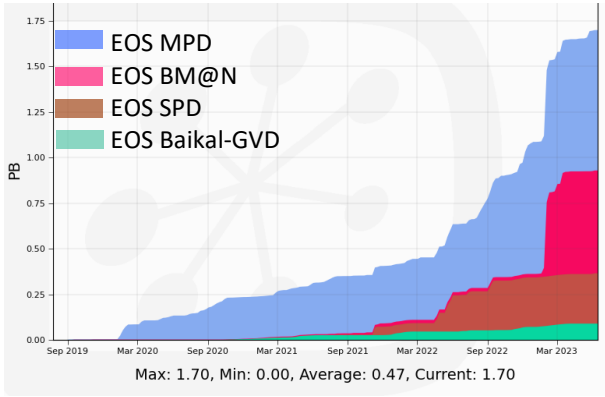


Total number of executed jobs



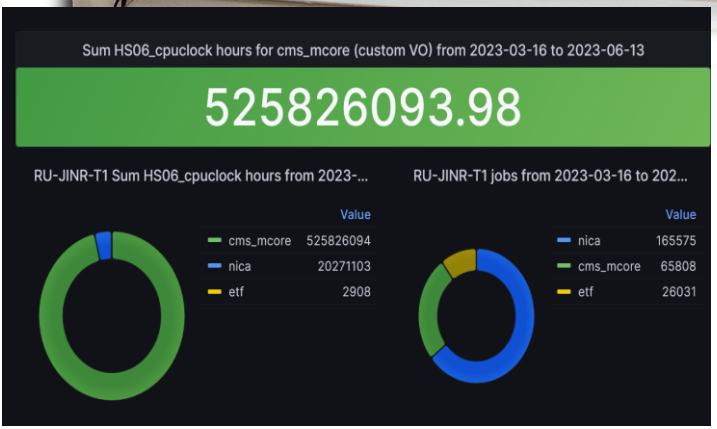
The major user of the distributed platform is the MPD experiment.

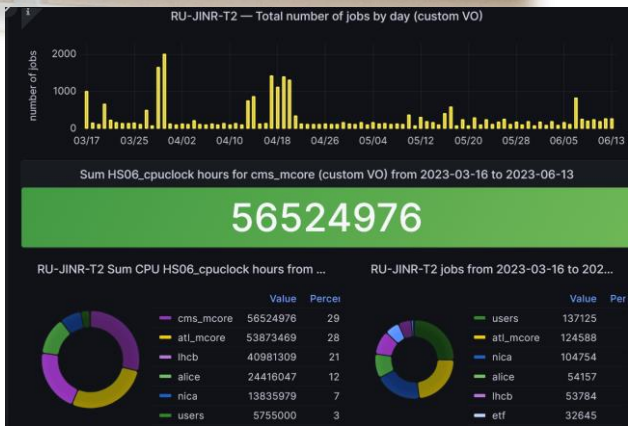
Data processed by experiments



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

MICC Monitoring @Accounting





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

The successful functioning of the computing complex is ensured by the system that monitors all MICC components. We must

- ➢ 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);
- ➢ organize the monitoring of the cooling system (cooling towers, pumps, hot and cold water circuits, heat exchangers, chillers);
- ➢ create an engineering infrastructure control center (special information panels for visualizing all statuses of the MICC engineering infrastructure in a single access point);
- ➢ account each user job on each MICC component.

It is required 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.

Estimation of the Resources of the MICC Components

	2024	2025	2026	2027	2028	2029	2030
HybriLIT heterogeneous platform. "Govorun" supercomputer.							
Total number of CPU cores	11000	11000	11000	14000	14000	14000	17000
Total number of GPU accelerators	40	64	64	64	64	88	88
Total volume of the hierarchical data processing and storage system, PB	8	8	14	14	20	20	20
Tier1 grid site							
Tier1 performance HEPS06	350000	400000	500000	550000	650000	750000	850000
Total number of CPU cores	22000	23000	30000	32000	38000	45000	50000
Total data storage capacity, TB	14500	16000	18000	20000	22000	23000	25000
Tier2 grid site							
Tier2 performance HEPS06l	187000	204000	221000	238000	306000	408000	510000
Total number of CPU cores	11000	12000	13000	14000	18000	24000	30000
Data storage system							
Total volume of the Data Lake on EOS, PB	27	35	38	53	58	71	83
Total robotic tape storage capacity, PB	70	90	130	130	170	170	190
Cloud computing							
Total number of CPU cores	2072	3072	4072	5072	6072	7072	8072
SSD-based ceph storage capacity, TB	868	968	1068	1168	1268	1368	1468

Prices for equipment in 2022-2023 are taken into account.

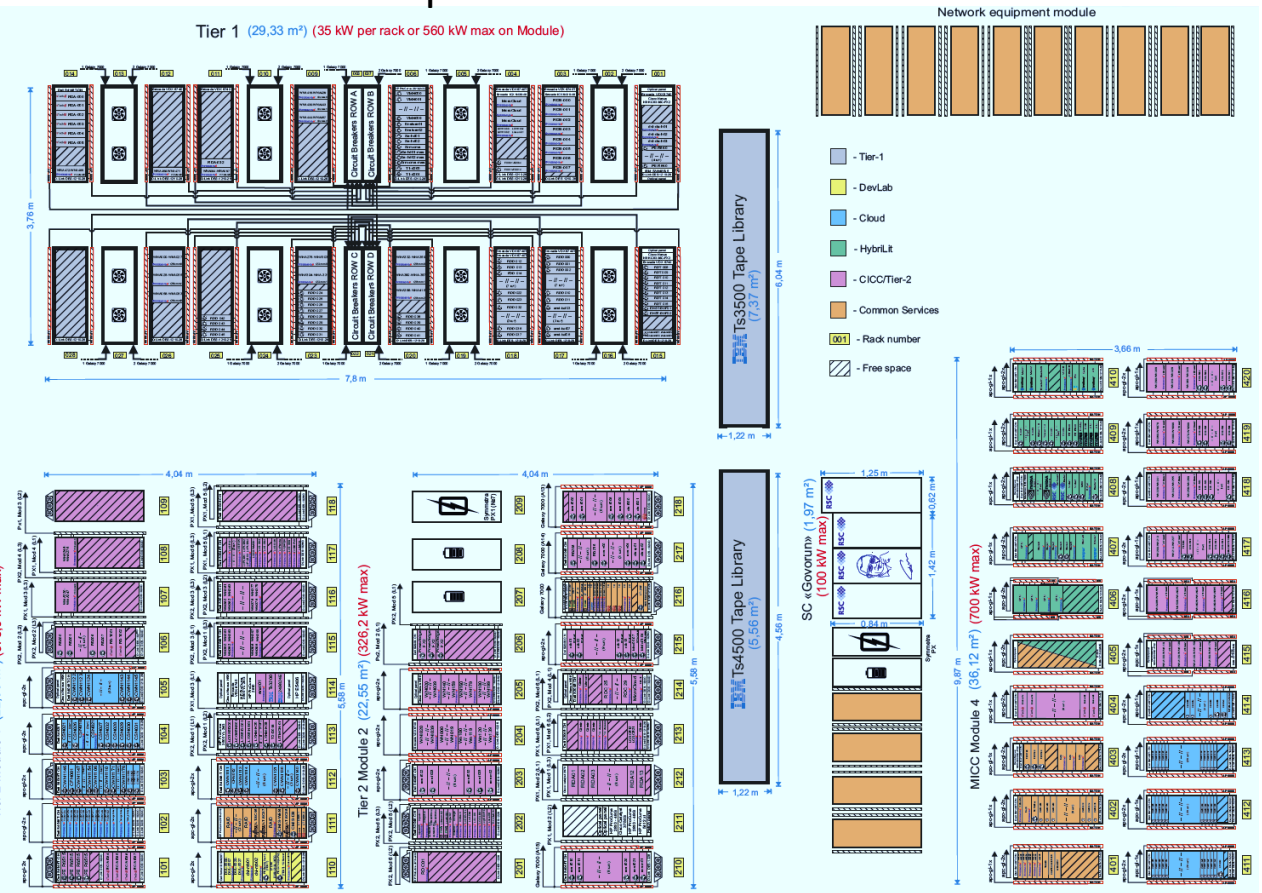
MICC Resources Development

- ✓ Increase in the computing resources of Tier1 up to **50 000 cores**
- ✓ Increase in computing resources of Tier2 up to **30 000 cores**
- ✓ Expansion of the storage system of Tier1 on disks up to **25 PB**
- ✓ Expansion of the MICC storage system on EOS up to **83 PB**
- ✓ Increase in Cloud total resourses up to **8720** cores, SSD-based ceph storage on **1.4 PB**
- ✓ Increase in computing resources of “Govorun” up to **17 000 CPU** cores, number of GPU up to **88,** volume of the hierarchical data processing and storage system, up to **20 PB**

MICC Server Halls

Present (1000 kW)

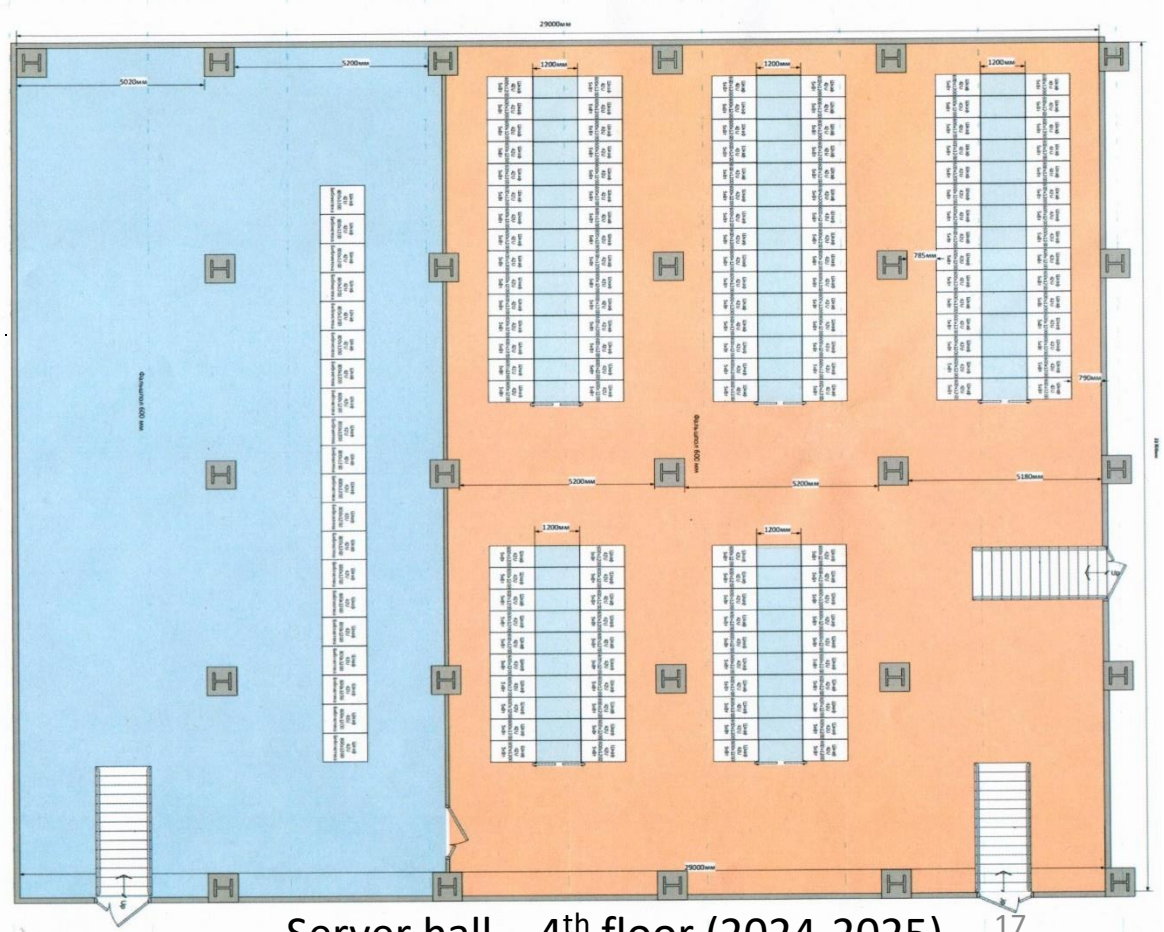
69 racks for servers
4 racks for the “Govorun” SC
10 racks for network equipment
4 racks for administrative services
2 robotic tape libraries



Server hall – 2nd floor (2023)

Planning for the future – new server hall for the MICC (600 kW)

containment area for robotic tape libraries
130 racks for servers

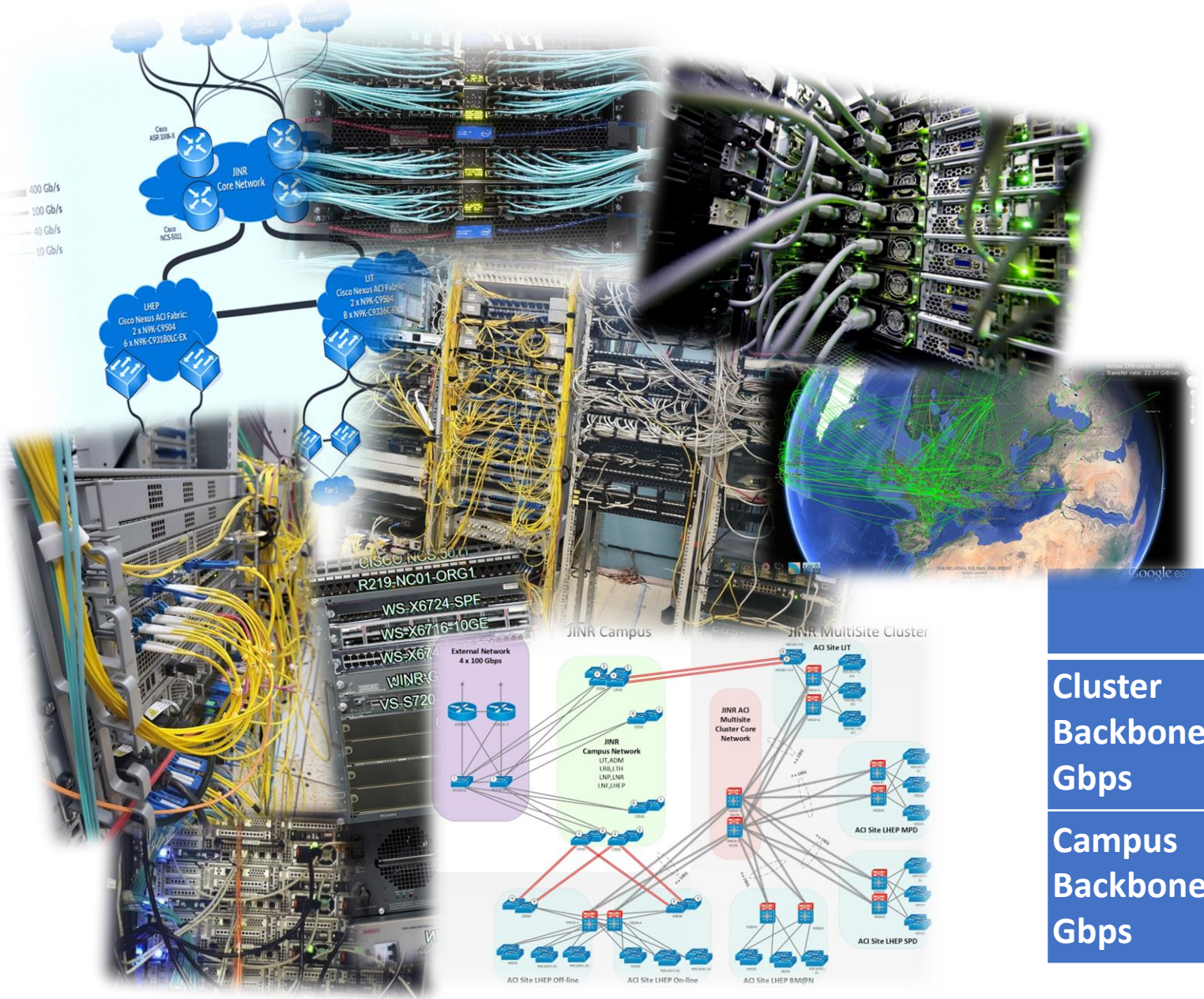


Server hall – 4th floor (2024-2025)

Power & Cooling

	2024	2025	2026	2027	2028	2029	2030
Power consumption, kVA	800	1000	1200	1400	1600	1800	2000
Cooling, kW	1400	1700	2000	2300	2600	2800	3000

Networking



Main goals:

- **support for state-of-the-art networking technologies**
- **software-defined networks (SDN)**
- **content delivery networks (CDN)**
- **named data networks (NDN)**
- **technologies for building distributed data centers - Data Center Interconnect (DCI).**

	2024	2025	2026	2027	2028	2029	2030
Cluster Backbone, Gbps	400	400	400	800	800	800	800
Campus Backbone, Gbps	200	200	200	200	400	400	400

Activity: Digital Ecosystem (Digital JINR)

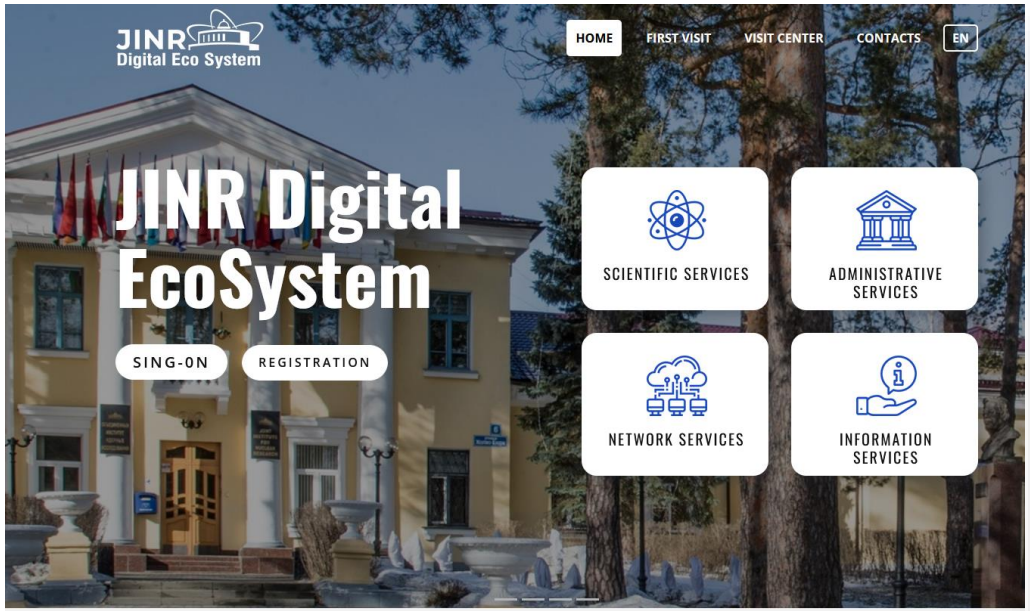
The digital platform "JINR Digital EcoSystem" integrates existing and future services

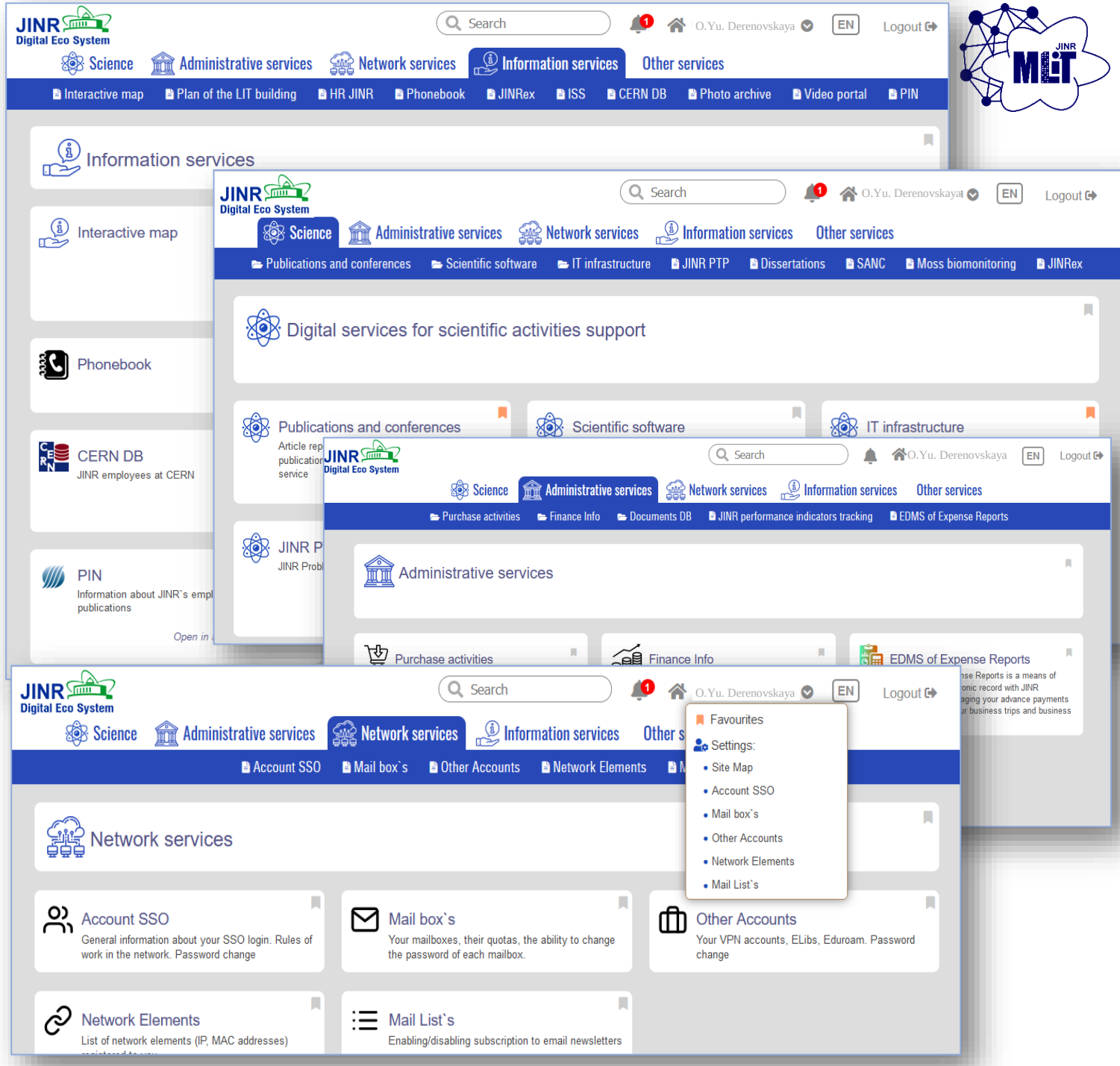
to support
scientific,
administrative and social activities,
maintenance of the engineering and IT infrastructures

to provide
reliable and secure access to various types of data

to enable
a comprehensive analysis of information

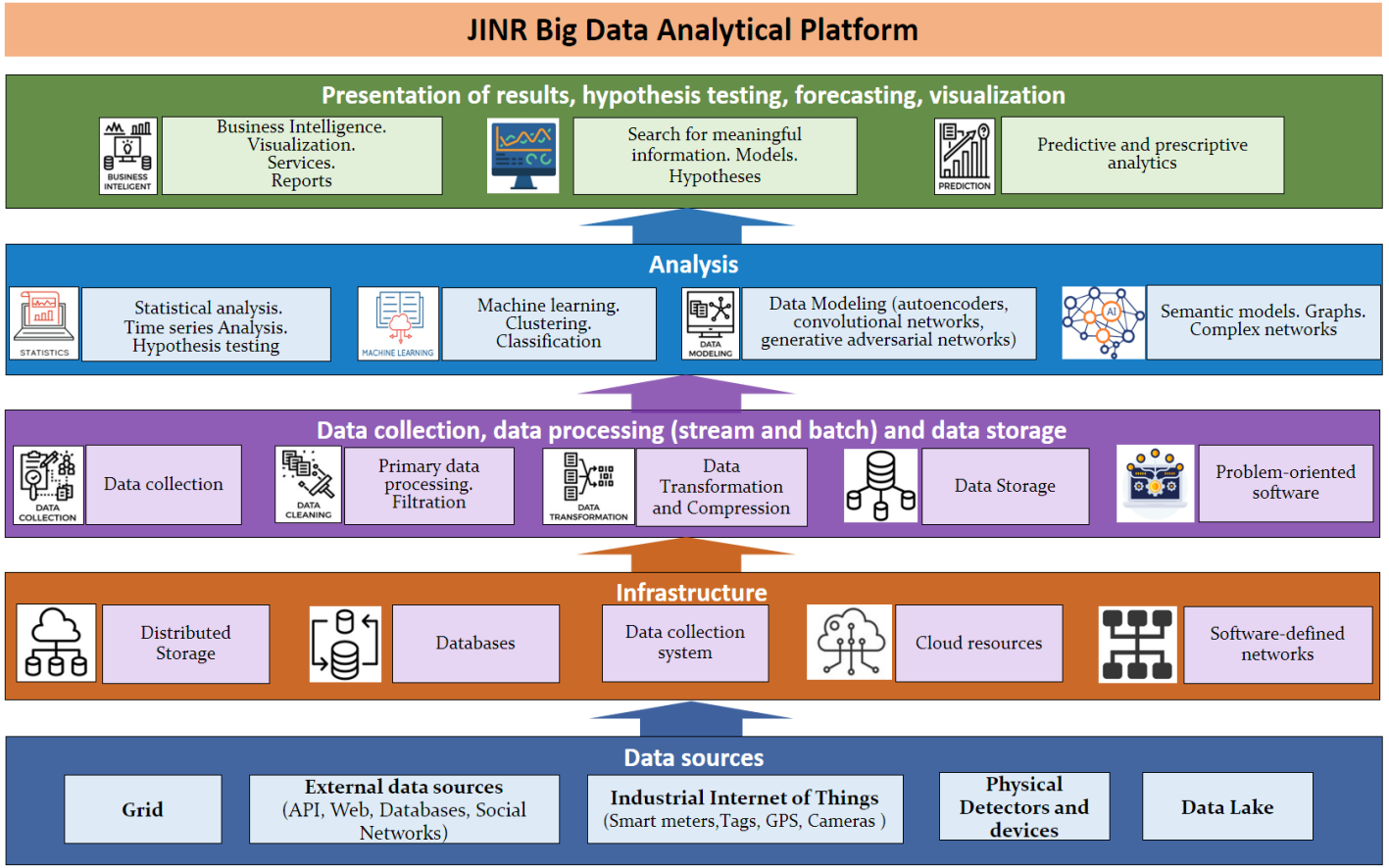
using
modern Big Data technologies and artificial intelligence.





- ✓ 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

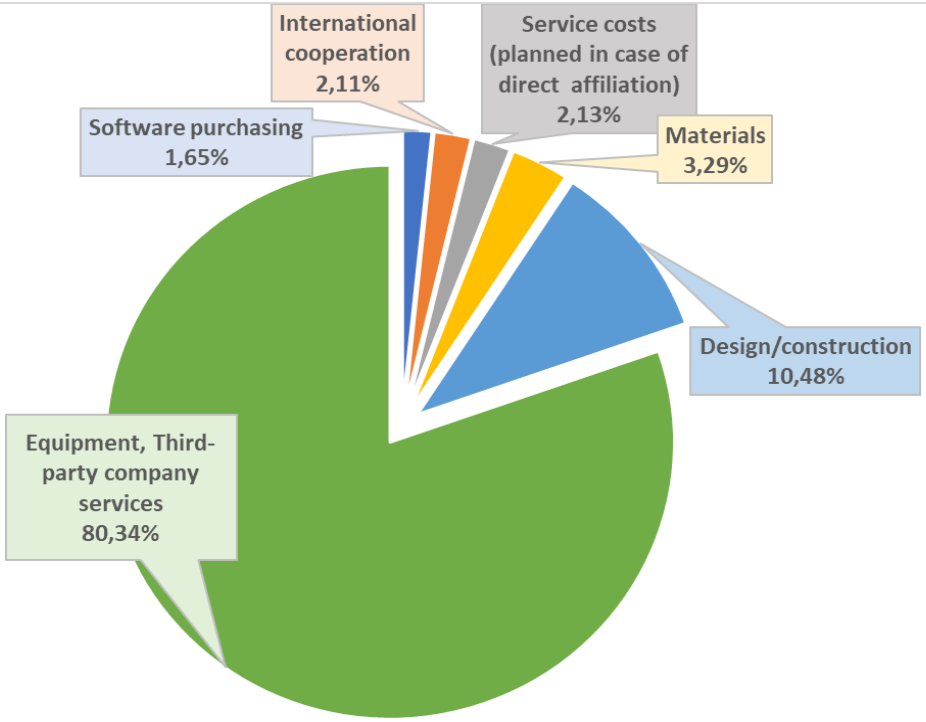
Activity: Multi-purpose Hardware and Software Platform for Big Data Analytics



Goal: 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.

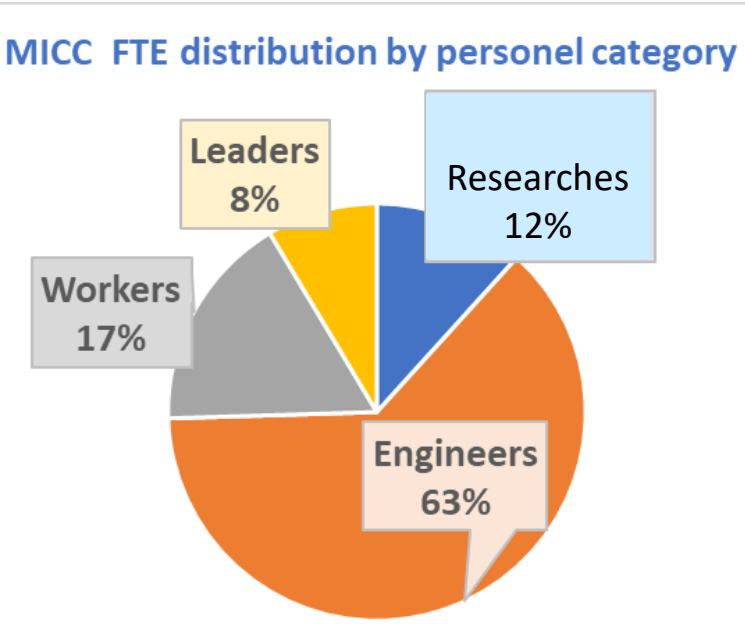
Total Estimated Cost of the MICC (including the operation cost)



No.	Items of expenditure	Cost	Expenditure per year (thousands of US dollars)						
			2024	2025	2026	2027	2028	2029	2030
1.	International cooperation	1120.0	160.0	160.0	160.0	160.0	160.0	160.0	160.0
2.	Materials	1750.0	250.0	250.0	250.0	250.0	250.0	250.0	250.0
3.	Equipment, Third-party company services	42683.0	5607.0	5675.0	5908.5	5575.0	6547.5	6185.0	7185.0
4.	Commissioning	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5.	R&D contracts with other research organizations	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6.	Software purchasing	875.0	110.0	115.0	120.0	125.0	130.0	135.0	140.0
7.	Design/construction	5568.1	646.6	760.1	658.6	1169.9	379.8	929.1	1024.0
8.	Service costs (planned in case of direct affiliation)	1132.1	150.0	153.7	157.6	161.5	165.6	169.7	174.0
TOTAL:		53128.2	6923.6	7113.8	7254.7	7441.4	7632.9	7828.8	8933.0

It should be pointed out that the increase in the number of computing cores, the number of GPU accelerators and the enlargement of the volume of the hierarchical data processing and storage system above the plan will be defined by the needs of users, including the needs of the NICA megaproject, and carried out by attracting financing from the budgets of the experiments, joint grants and other sources.

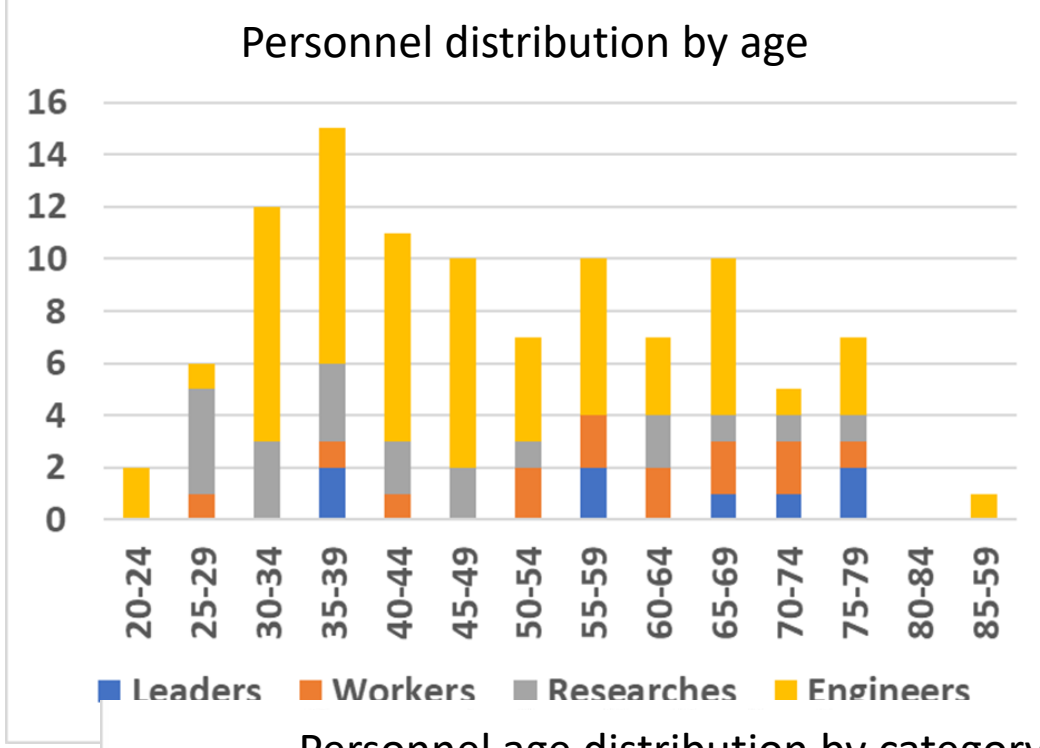
Personnel

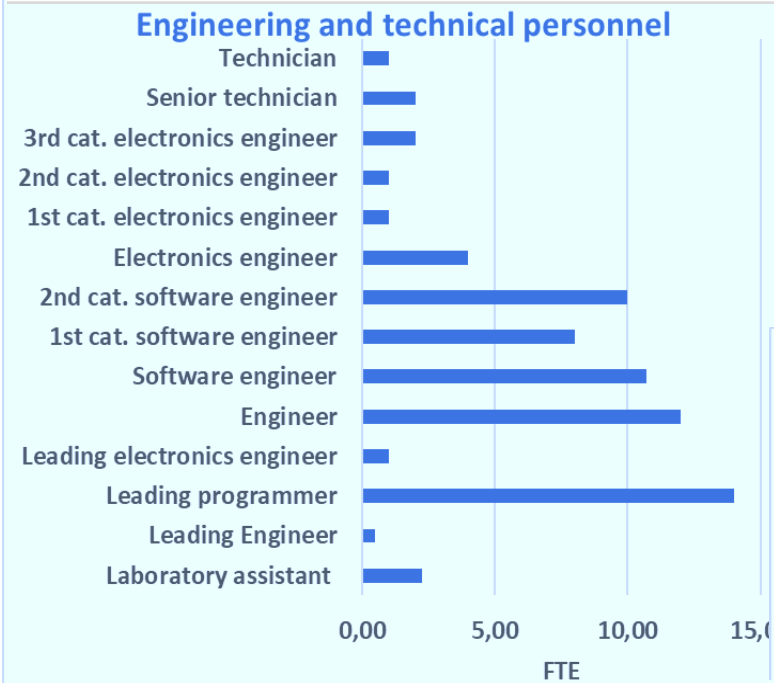


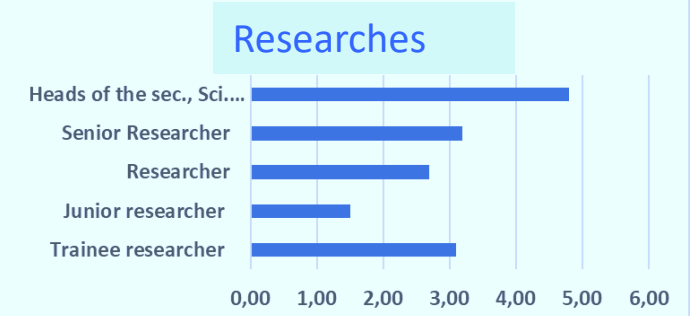
Manpower needs in the first year of the LRIP MICC implementation, including the activities

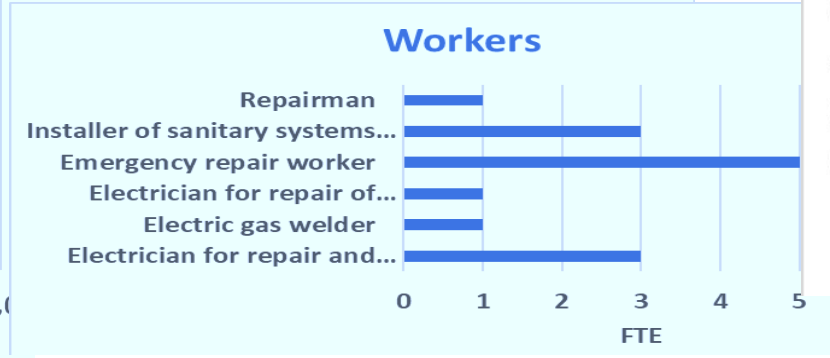
109.25 FTE

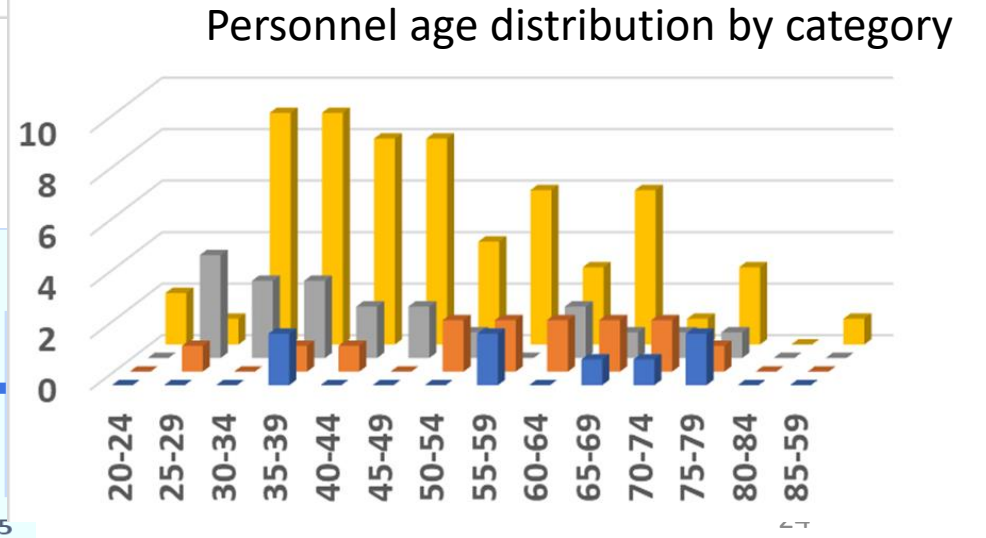
Collaborations with all JINR Laboratories, 62 institutes and universities from 19 countries











Risks

- ✓ Unpredictability of availability and prices of advanced equipment from leading manufacturers of computing architectures, low-latency network equipment and high-performance data storage elements.
- ✓ Rapid obsolescence of the computer and network equipment.
- ✓ Virus and hacker attacks from outside and inside due to user carelessness.
- ✓ Depreciation and moral obsolescence of the engineering equipment, the modernization of which is delayed due to the excessive bureaucratization of the decision-making procedure.

We ask the PAC to give a recommendation to transform the

theme 05-6-1118-2014/2023 "Information and Computing Infrastructure of JINR"

into a

large research infrastructure project

"Multifunctional Information and Computing Complex (MICC)"

for the period up to the end of 2030.

Thank you for your attention!