

МЕЖДУНАРОДНАЯ НАУЧНО-ТЕХНИЧЕСКАЯ ОРГАНИЗАЦИЯ
JOINT INSTITUTE FOR NUCLEAR RESEARCH
ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ



Методы и технологии обработки данных в гетерогенных вычислительных средах

Кореньков Владимир Васильевич

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

«Международная конференция «Математика в созвездии наук»

Филиал МГУ в Дубне, ОИЯИ, ЛИТ

2 апреля 2024 года

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

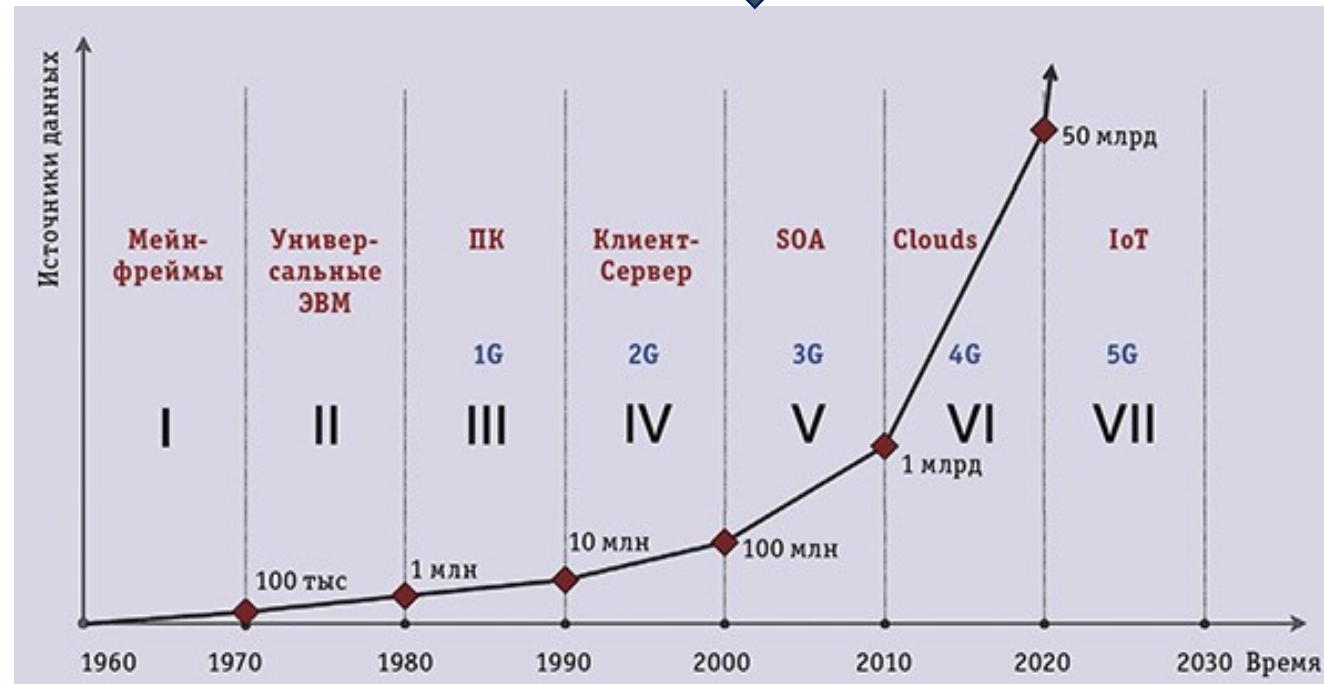
Третье поколение (1980-е годы) — персональные компьютеры

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

Пятое поколение (2000-е годы) — сервисная архитектура

Шестое поколение (2010-е годы) — облака

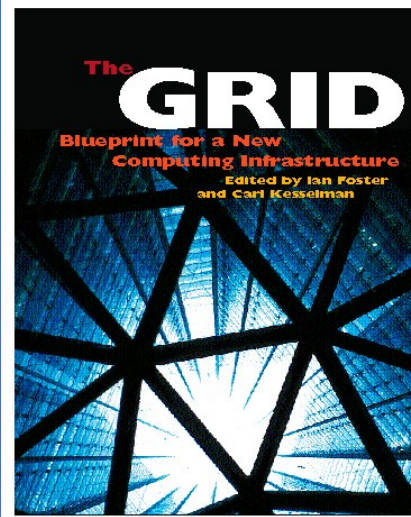
Седьмое поколение (2020-е годы) — IoT, искусственный интеллект, квантовые вычисления



Grids, clouds, fog, edge, supercomputers...

Grids

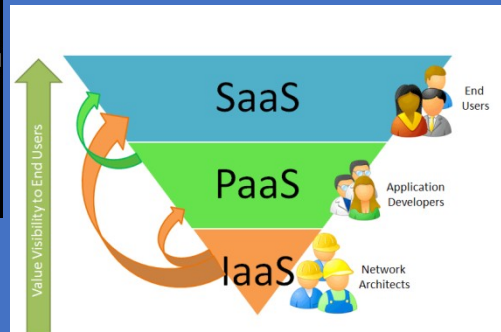
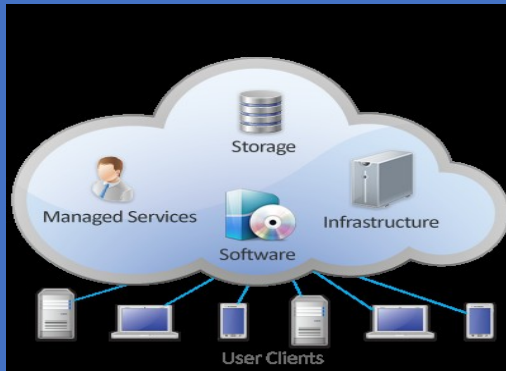
- Collaborative environment
- Distributed resources



Supercomputers



Clouds



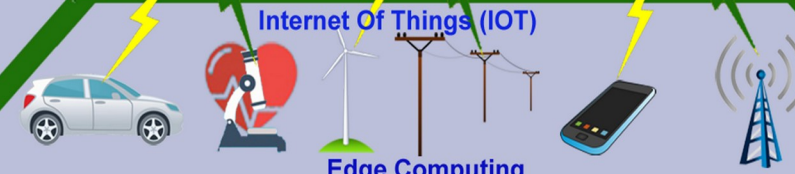
Cloud Node/Internet/Data center



Fog Nodes /Local Data storage



Edge Nodes /Devices /PLC/PAC



Edge Computing

Грид технологии – путь к успеху

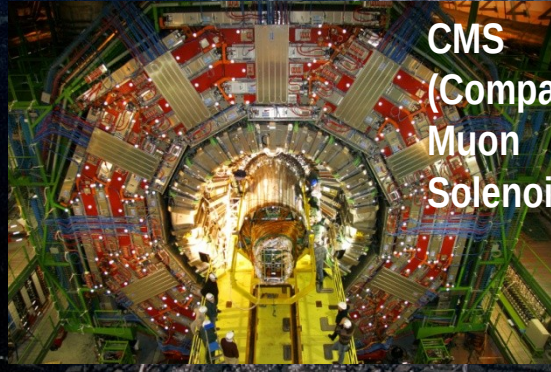


На торжестве по поводу получения Нобелевской премии за открытие бозона Хиггса директор ЦЕРНа Рольф Хойер прямо назвал **Грид-технологии** одним из **трех столпов успеха** (наряду с ускорителем LHC и физическими установками). **Без организации грид-инфраструктуры на LHC было бы невозможно обрабатывать и хранить колоссальный объем данных, поступающих с коллайдера, а значит, совершать научные открытия. Сегодня уже ни один крупный проект не осуществим без**



Large Hadron Collider

The Large Hadron Collider (**LHC**), one of the largest and truly global scientific projects ever, is the most exciting turning point in particle physics.



CMS
(Compact Muon Solenoid)



LHCb
(Large Hadron Collider beauty experiment)

Data flow to permanent storage: 4-6 GB/sec

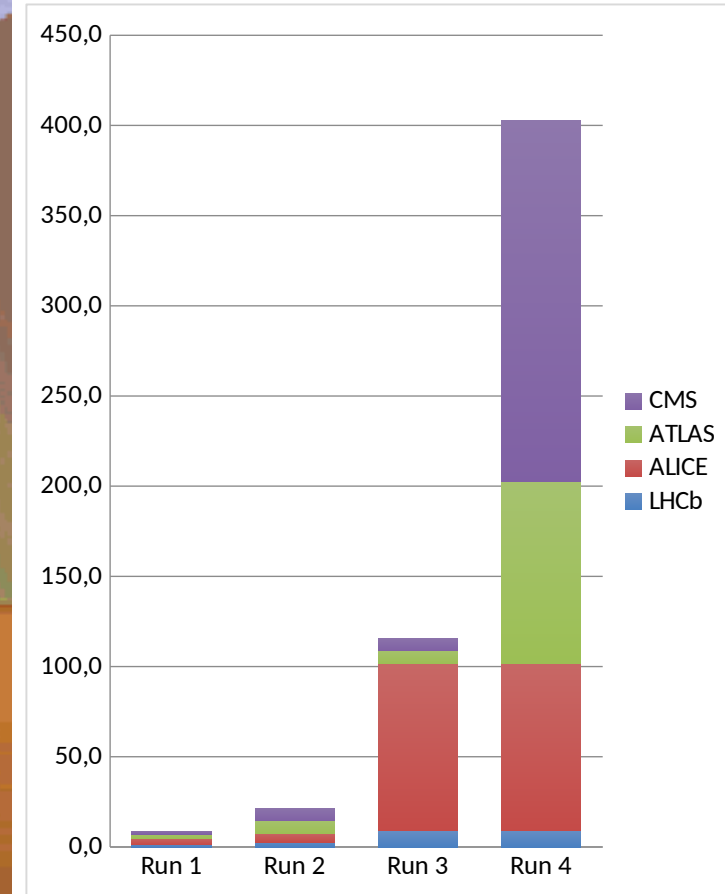
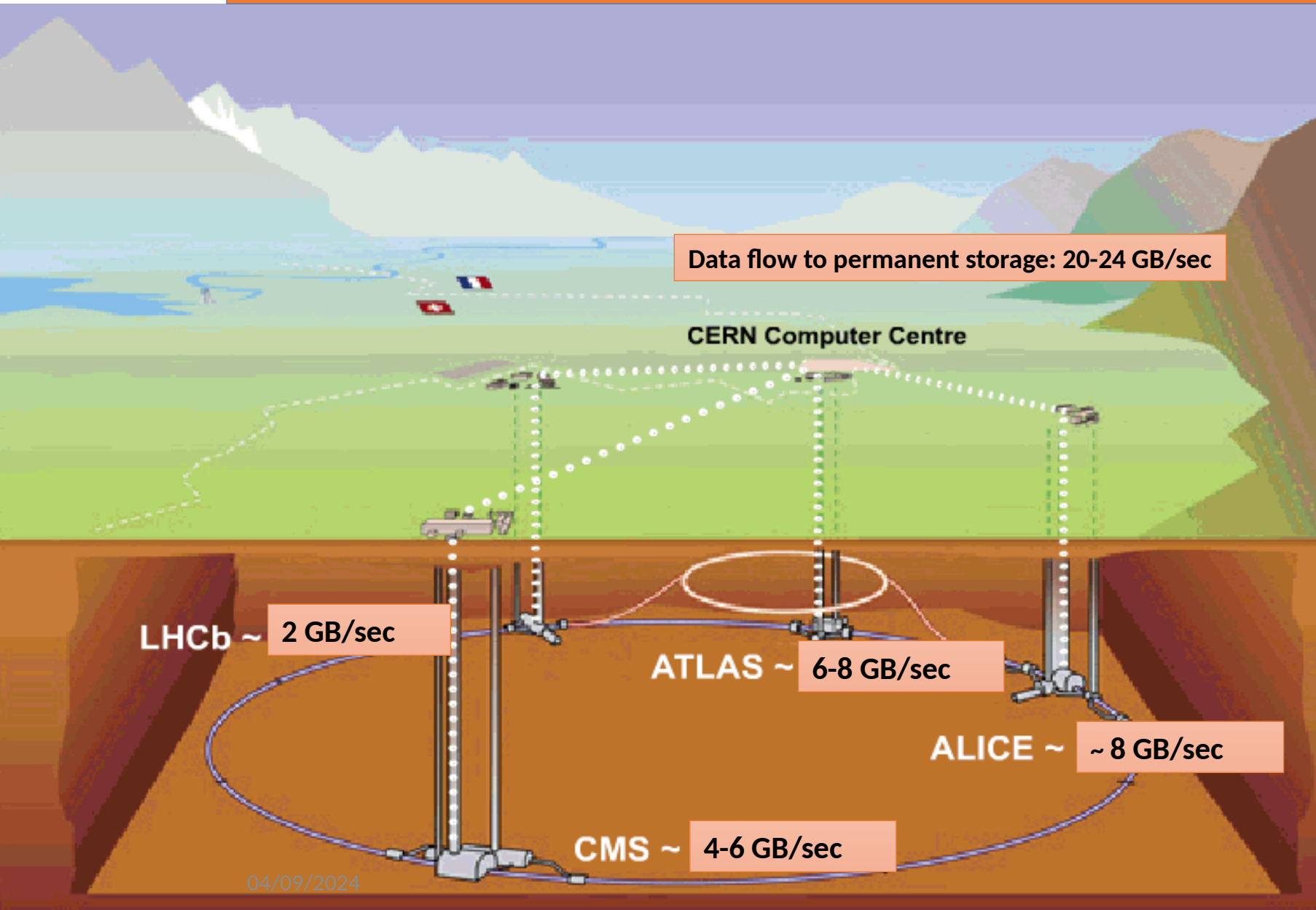


ALICE
(A Large Ion Collider Experiment)

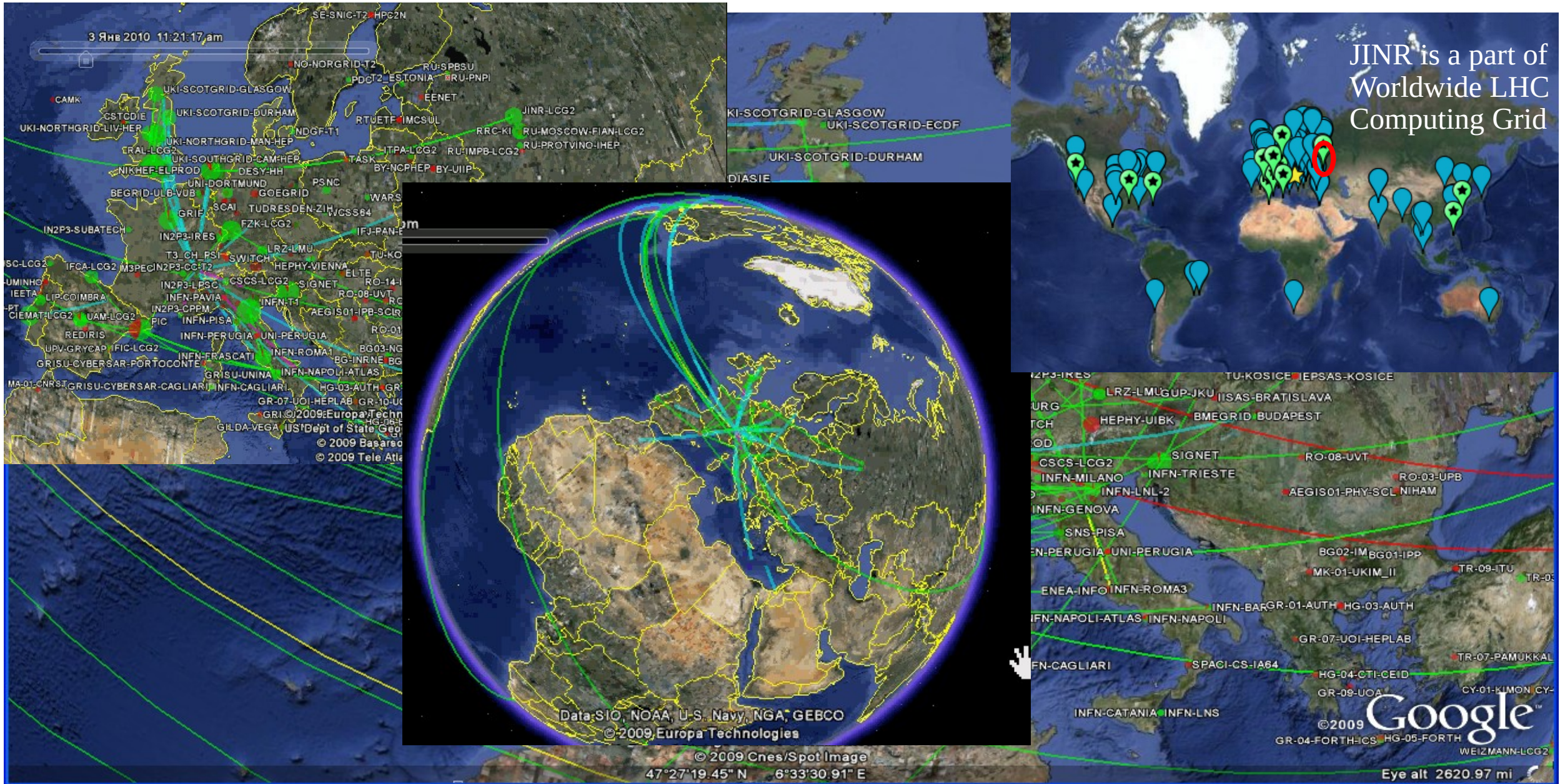


ATLAS
(A Toroidal LHC ApparatuS)

Data Collection and Archiving at CERN



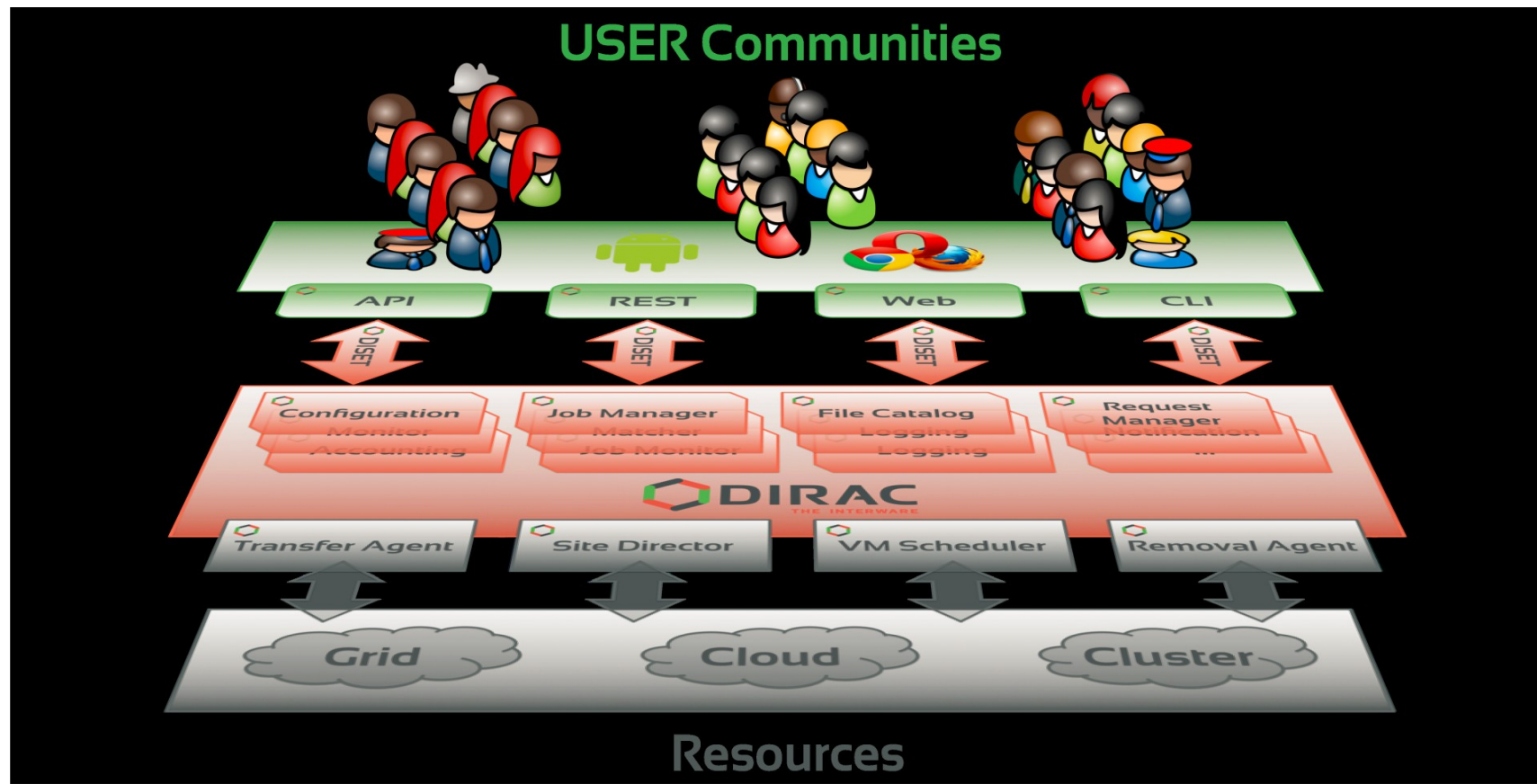
The Worldwide LHC Computing Grid (WLCG)



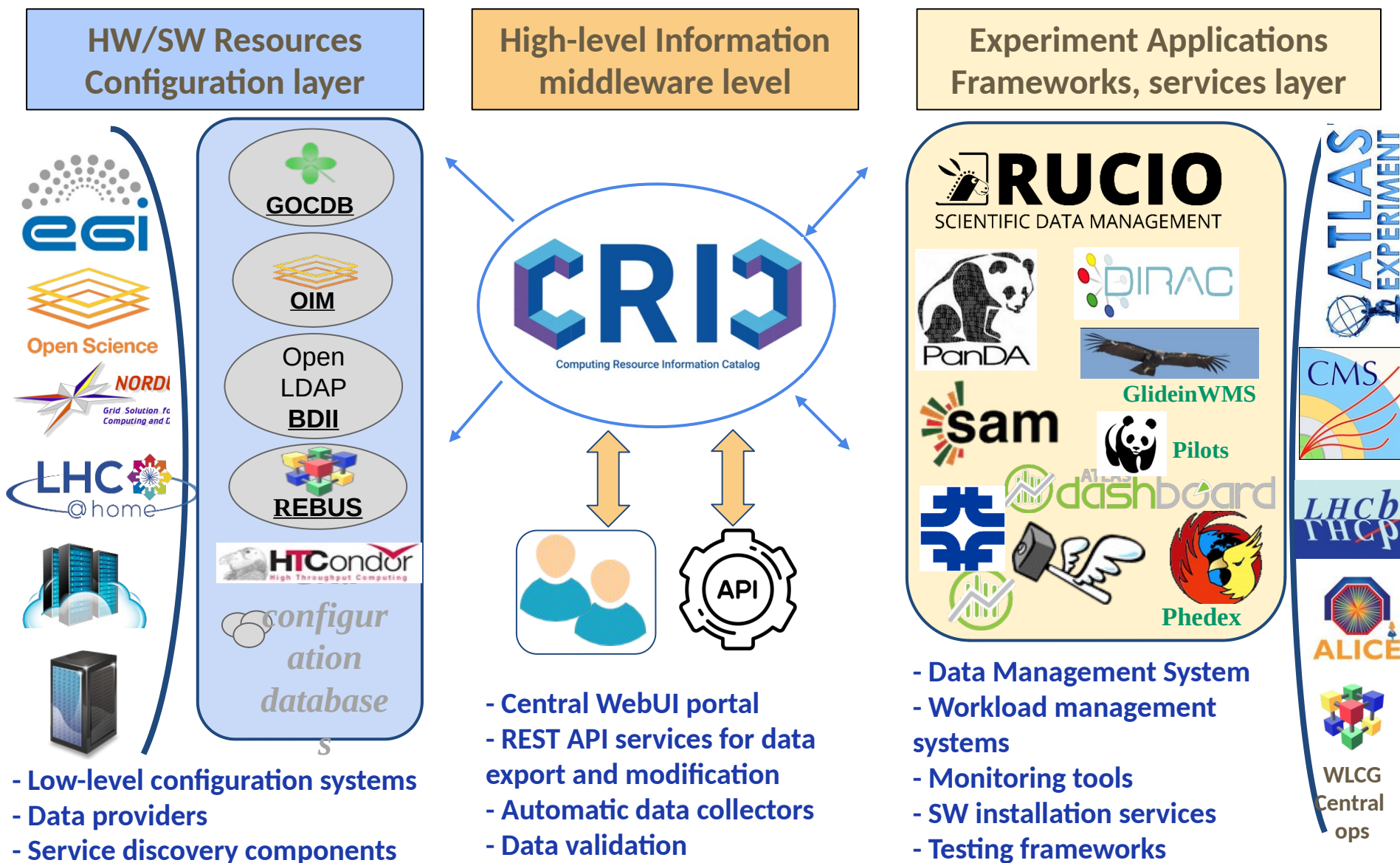
- Расширение компьютерных ресурсов за счет использования внешних невыделенных ресурсов (HLT, Clouds, HPC...)
- Изменения модели компьютеринга в каждом эксперименте, с целью оптимизации использования ресурсов
- Значительные усилия вкладываются в развитие программного обеспечения, чтобы улучшить общую производительность при использовании современных архитектур (многоядерность, GPU...)
- Оптимизации процессов обработки, количество хранящихся реплик данных и др.

Платформа DIRAC

- DIRAC has all the necessary components to build ad-hoc grid infrastructures **interconnecting** computing resources of different types, allowing **interoperability** and simplifying **interfaces**.
- This allows to speak about the DIRAC *interware*.



CRIC: a unified topology system for a large scale, heterogeneous and dynamic computing infrastructure



Distributed Storage and scientific data management

EOS Open Storage - Disk Storage at CERN

dCache - Distributed Storage for scientific data

CTA - the CERN Tape Archive

CEPH - Object & Application Storage

CERNBox - Sync & Share platform for collaboration

Filesystems - AFS, DFS, Samba, NFS3/4

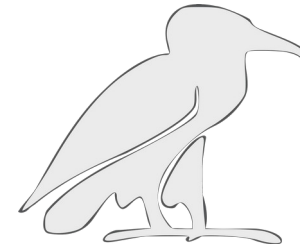
CernVM FS - CernVM File System

FTS & XRootd- File Transfer Service & File Access

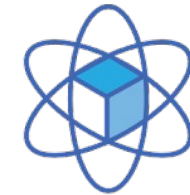
Rucio - Scientific Data Management

DAOS - high-performance storage system

Lustre - open-source, parallel file system



ceph



CERNBox

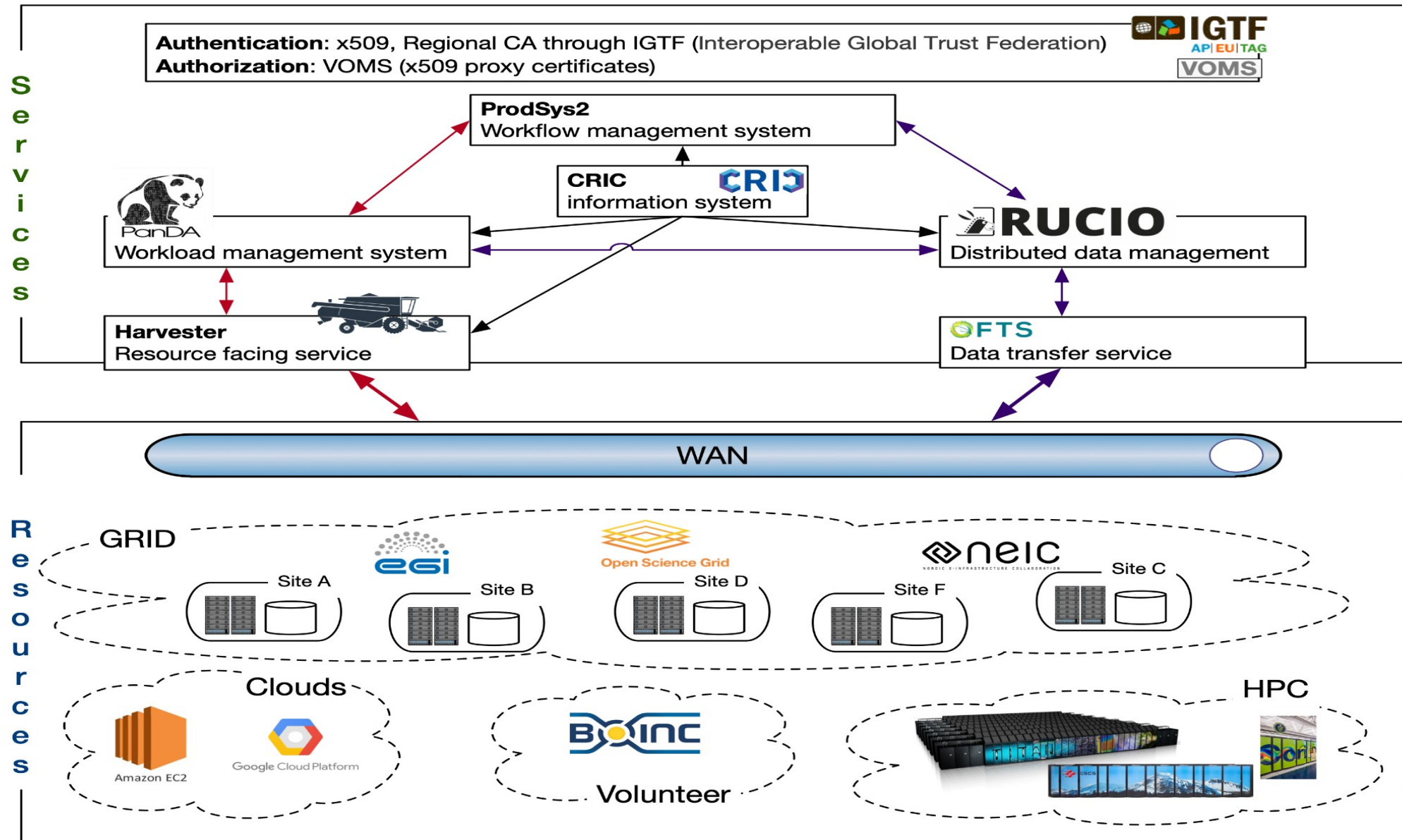


CERN
Tape Archive

DAOS

l·u·s·t·r·e®

ATLAS computing



The Worldwide LHC Computing Grid



WLCG: an International collaboration to distribute and analyse LHC data. Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists

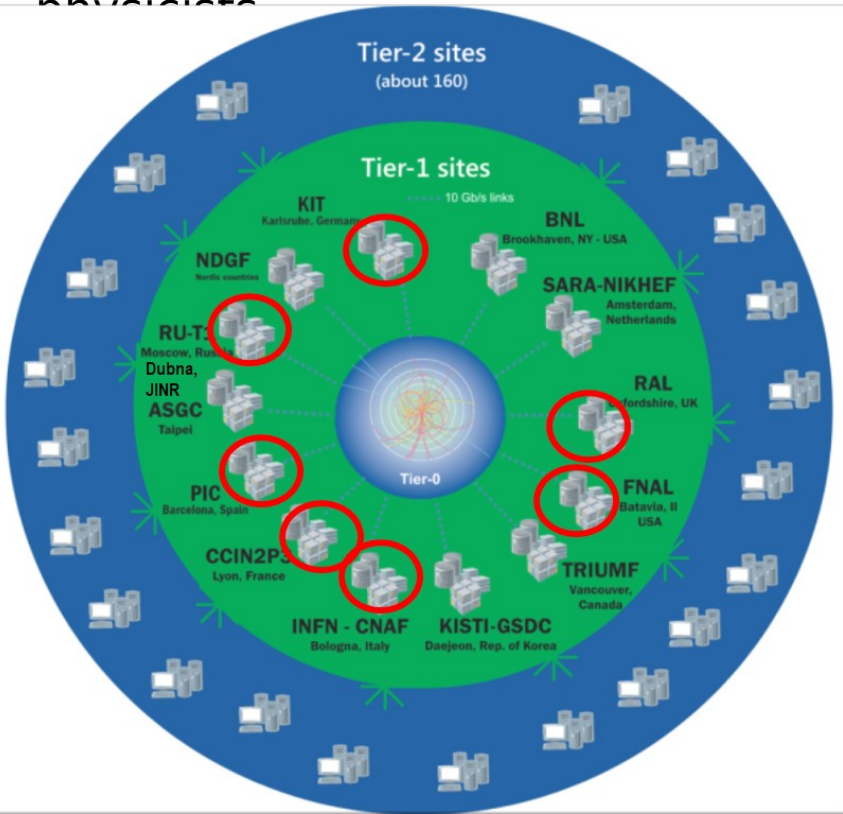
The mission of the WLCG project is to provide global computing resources to store, distribute and analyze the **~250-300 Petabytes** of data expected every year of operations from the Large Hadron Collider.

WLCG computing enabled physicists to announce the discovery of the Higgs Boson.

- 180 sites**
- 42 countries**
- > 12k physicists**
- ~1.6 M CPU cores**
- ~2 EB of storage (1 EB - CERN)**
- > 3 million jobs/day**
- 100-400 Gb/s links**



Worldwide LHC Computing Grid - 2023



Tier0 (CERN): data recording, reconstruction and

Tier1: permanent storage, re-processing, analysis

Tier2: Simulation, end-user analysis

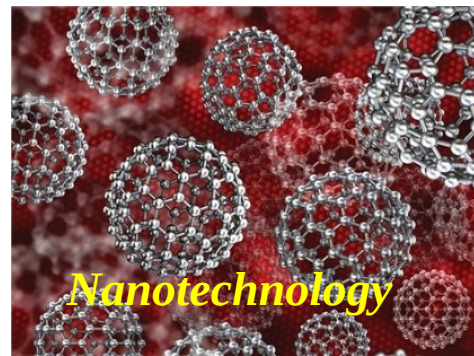
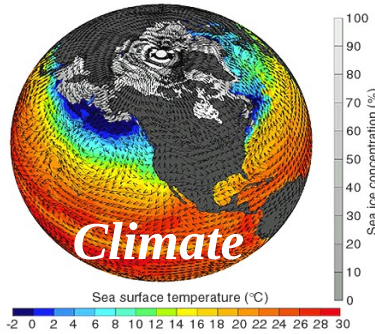
HPC+Big Data+Artificial intelligence



High Energy Physics



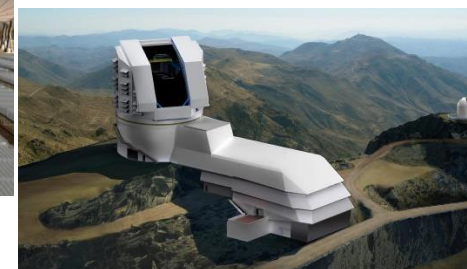
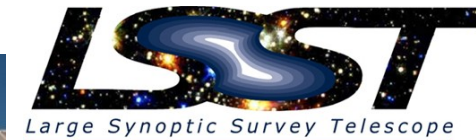
CERN Large Hadron Collider > 600 Pb/Year



Square Kilometer Array radio telescope
> 1 Eb/Year raw d (estimation)



International radiotelescope for the 21st century



Large Synoptic Survey Telescope > 10 Pb/Year (estimation)



... et cetera

International Large-scale projects



Russian research institutes and universities actively participate in international large-scale projects:

- LHC, CERN (experiments: ATLAS, ALICE, LHCb, CMS)
- XFEL, DESY (European free electron laser)
- ESRF, France (European synchrotron center)
- FAIR, GSI, Germany (CBM, PANDA experiments)
- ITER, France ...

International large-scale projects are being prepared in Russia:

- NICA, JINR, Dubna (proton and heavy ion collider)
- PIK, PNPI, Gatchina (high-flow reactor complex)
- SKIF, INP SB RAS Novosibirsk (Siberian ring photon source)
- Super S-Tau Fabric, Sarov (electron-positron collider)
- Нейтринная программа (Байкал, JUNO, NOVA, DUNE ...)
- синхротронно-нейтронная программа, науки о жизни

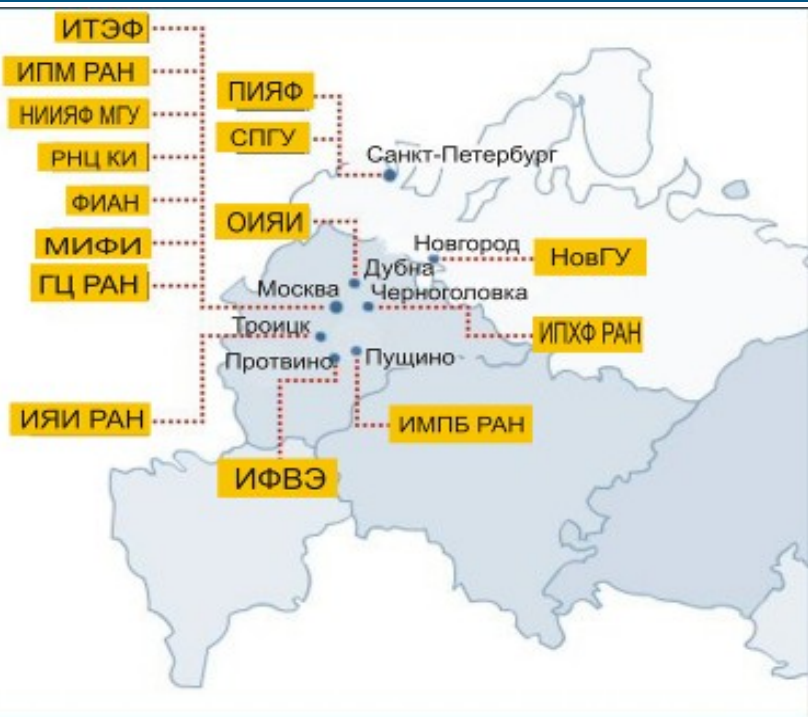


Joint Institute for Nuclear
Research
SCIENCE BRINGING NATIONS
TOGETHER



Институт ядерной физики
имени Г. И. Будкера СО РАН

From RDIG to RDIG-M



The Russian consortium RDIG (Russian Data Intensive GRID) was set up in September 2003 as a national federation in the EGEE project.

A protocol between CERN, Russia and JINR on participation in the LCG project was signed in 2003. MoU on participation in the WLCG project was signed in 2007.
Consortium RDIG-M – Russian Data Intensive GRID for Megascience projects



Создание консорциума для IT-обеспечения исследовательской инфраструктуры класса «мегасайенс»

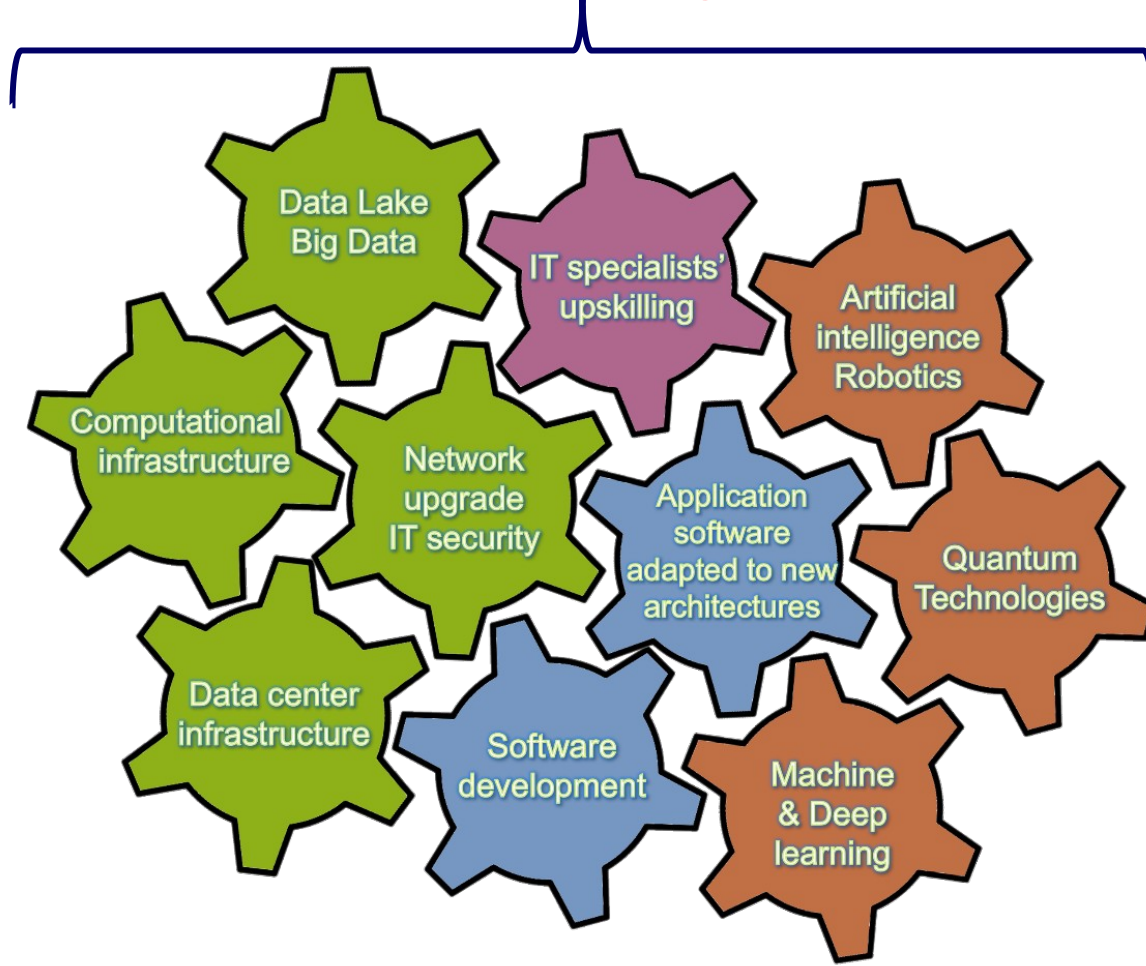


- Консорциум Российский ГРИД для интенсивных операций с данными (РДИГ) был создан в 2003 году для активного участия в распределенной обработке данных экспериментов на Большом адронной коллайдере LHC в рамках научной коллаборации LHC WLCG (Worldwide LHC Computing Grid). Созданная инфраструктура RDIG имеет огромное значение для эффективного участия ученых России в научной программе экспериментов на LHC.
- В России реализуется программа масштабных научных проектов, важнейшей частью которых является развитие распределенных гетерогенных компьютерных систем (включая системы с экстремальным параллелизмом) для обработки, хранения, анализа экспериментальных данных, разработка и внедрение эффективных методов, алгоритмов и программного обеспечения для моделирования физических систем, математической обработки и анализа экспериментальных данных, развитие методов машинного обучения, искусственного интеллекта, квантовых вычислений.
- Для решения этой масштабной задачи необходимо развивать распределенную компьютерную инфраструктуру, объединяющую ключевые научные и образовательные институты, участвующие в проектах мегасайенс - РДИГ-М. Созданный консорциум на базе ОИЯИ, НИЦ Курчатовский институт, ИСП РАН должен стать ядром для IT-обеспечения исследовательской инфраструктуры класса «мегасайенс».

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

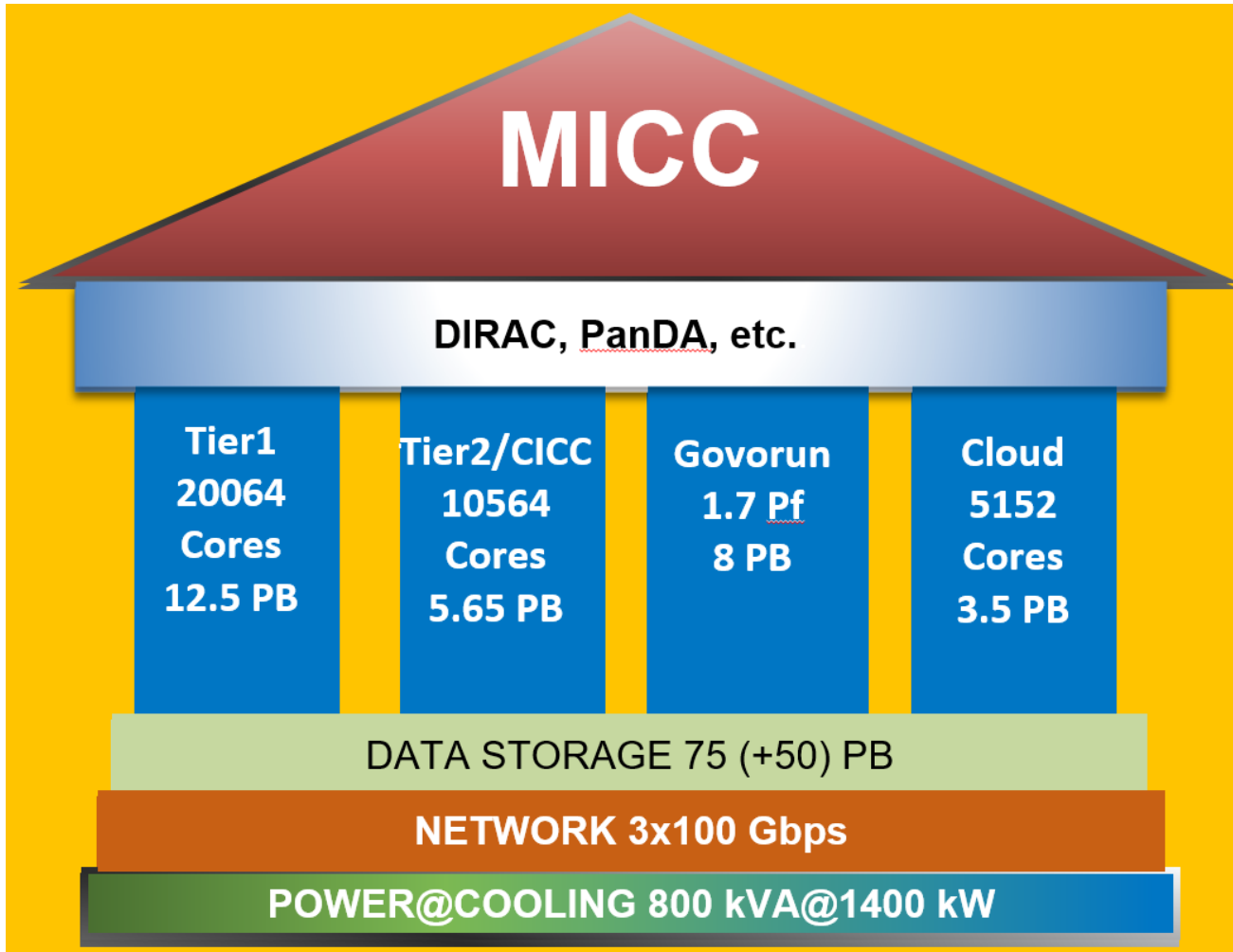
- Networking (**Tb/s** range),
- Computing infrastructure within the **Multifunctional Information & Computing Complex (MICC)** and
- “Govorun” Supercomputer,
- Data center infrastructure,
- Data Lake & long-term storage for all experiments.

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

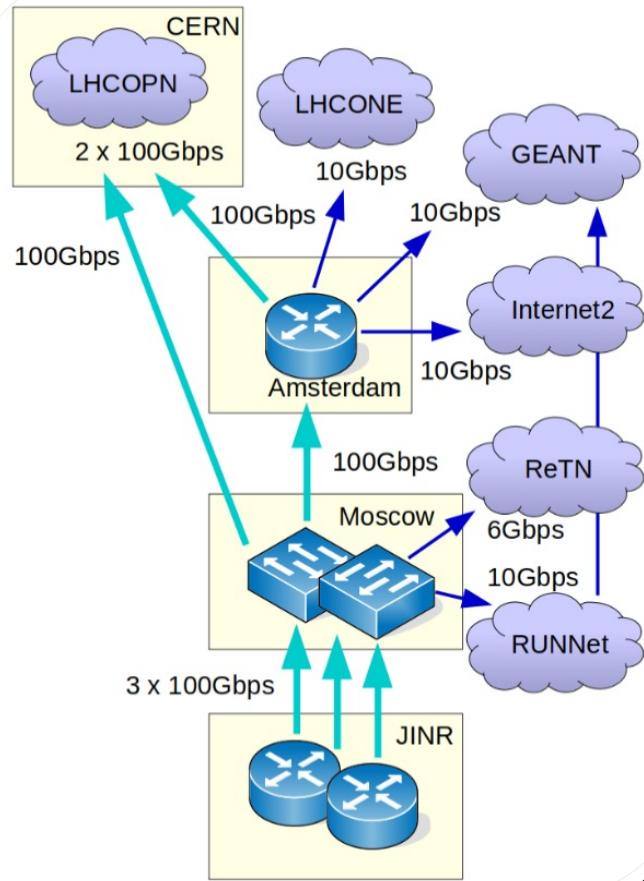
- Wide Area Network
- Local Area Network

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

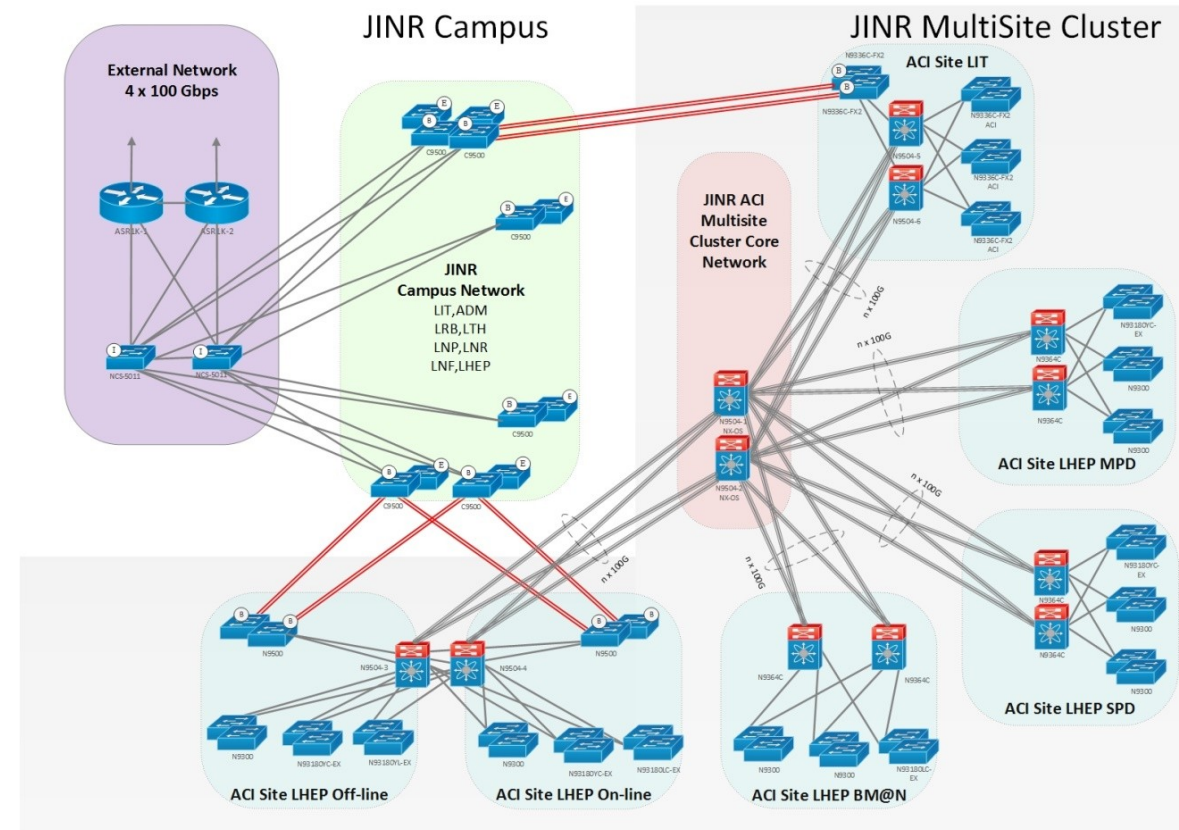
Networking



- 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 with NIKS networks
- The multi-site cluster network with a bandwidth **4x100 Gbit/s** between VBLHEP and MLIT



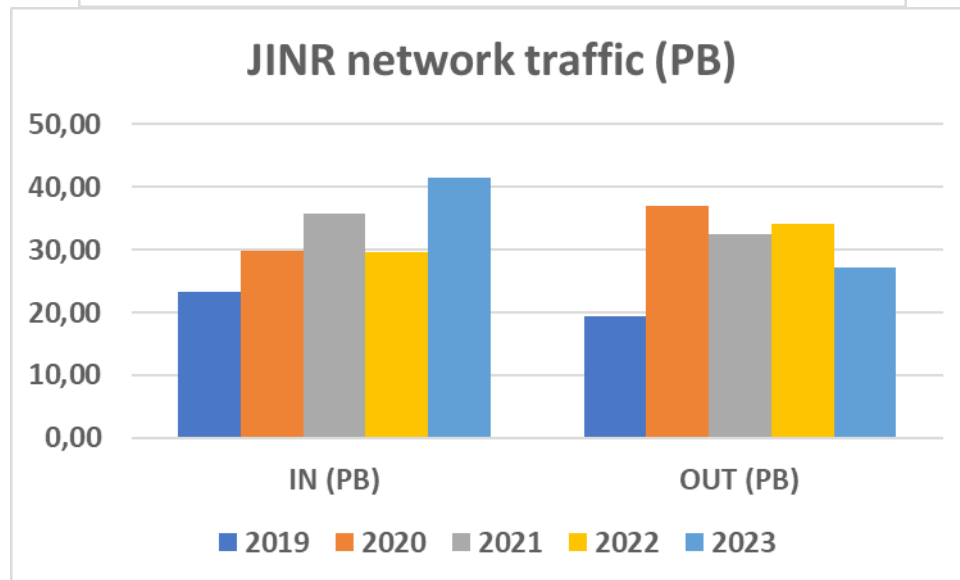
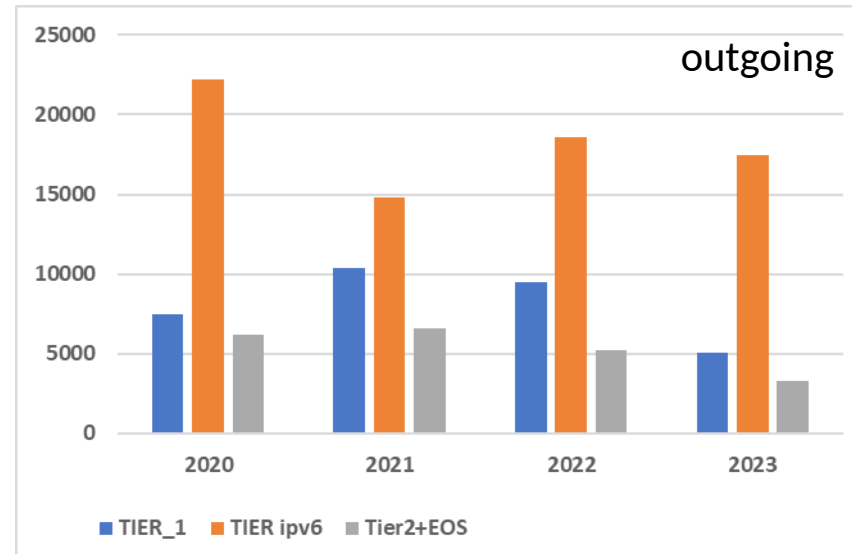
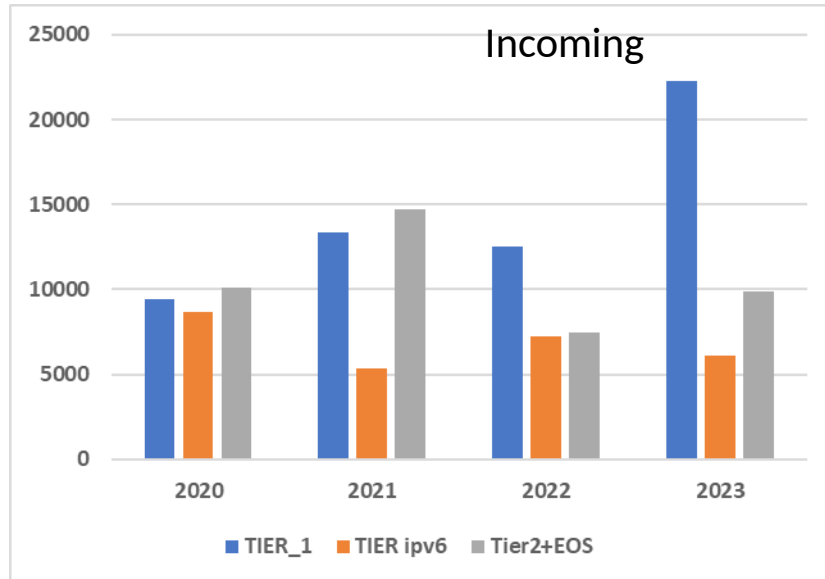
9327 network elements
18163 IP-addresses
6355 users
1464 E-library
911 remote VPN
121 VOIP
116 EDUROAM
4579 Email @jinr.ru



Networking @ Traffic



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

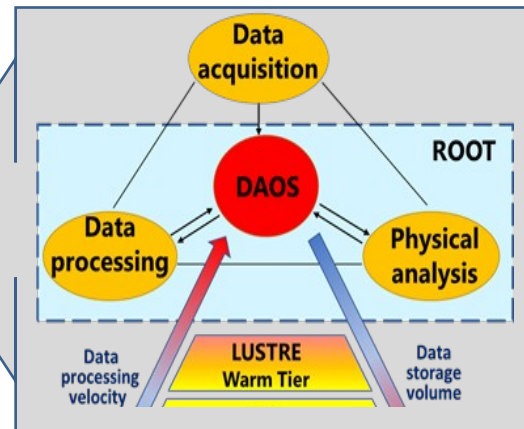


Общий входящий трафик ОИЯИ, включая сервера общего назначения, Tier1, Tier2, СК «Говорун» и облачные вычисления, составил в 2023 году 41,5 ПБ, общий исходящий – 27,5 ПБ.

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.

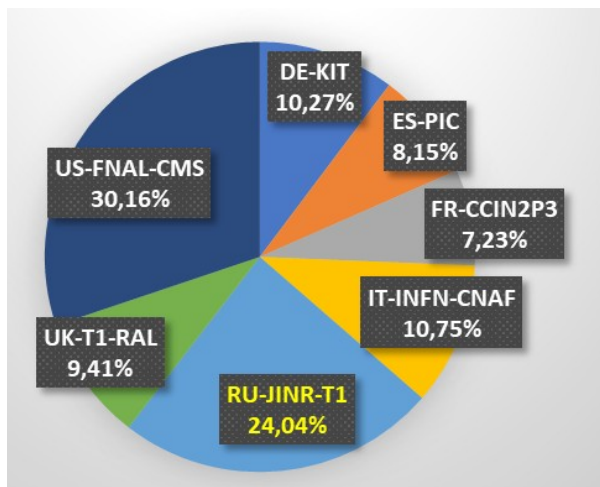
According to the speed of accessing data there are next layers:

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

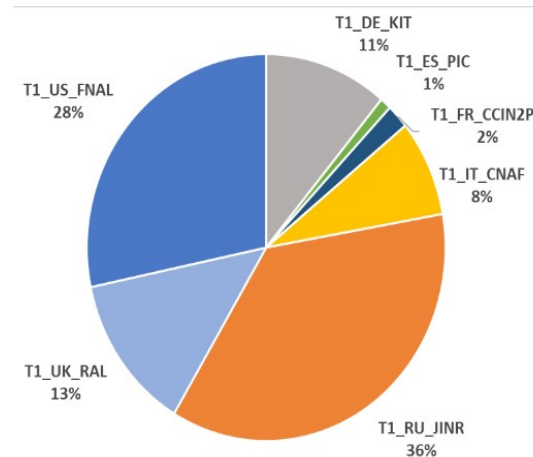
JINR Tier1 for CMS (LHC) and NICA



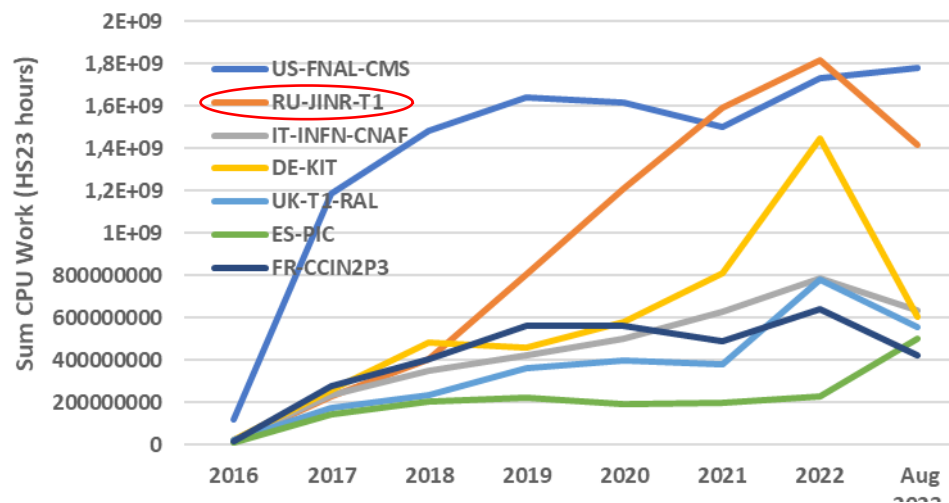
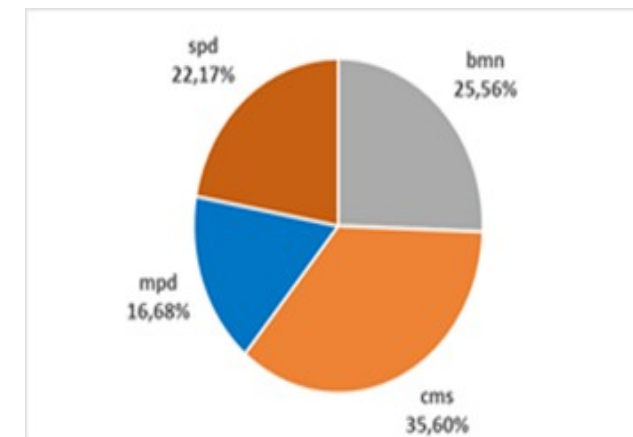
Вклад мировых Tier1 центров в обработку экспериментальных данных CMS за 2023 год:



Количество обработанных событий эксперимента CMS за 2023 год



Распределение по числу задач выполненных на Tier1 экспериментами CMS, BM@N, MPD и SPD в 2023 году



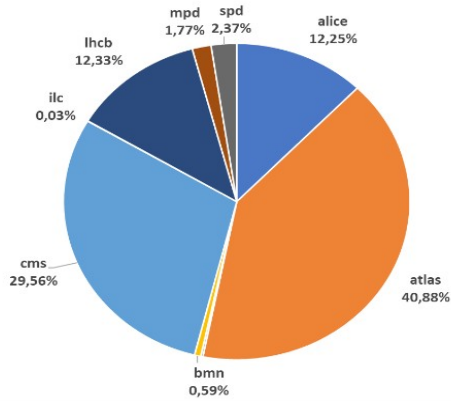
- 20064 cores
- 360 kHS06
- 12.5 PB disks
- 50.6 (+ 50) PB tapes
- 100% reliability and availability



Tier2 at JINR

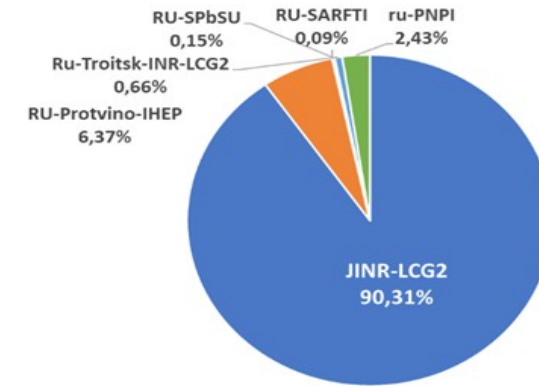


Использование Tier2 сайта ОИЯИ (JINR-LCG2) виртуальными организациями в рамках грид-проектов



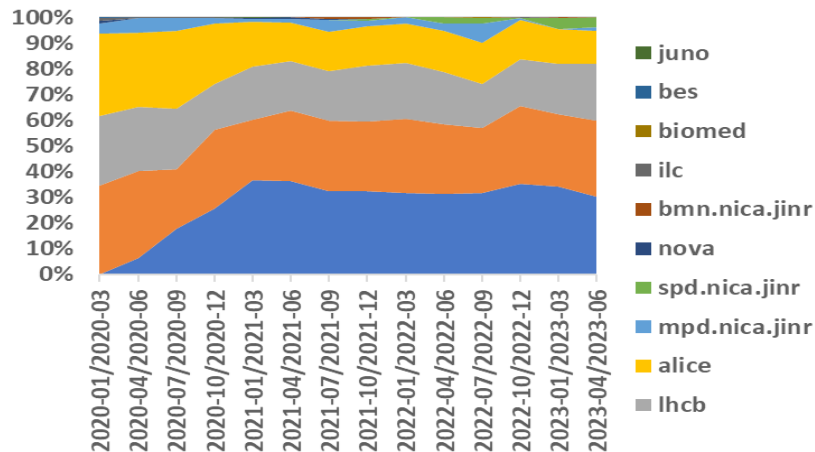
Tier2 at JINR provides computing power and data storage and access systems for the majority of JINR users and user groups, as well as for users of virtual organizations (VOs) of the grid environment (LHC, NICA, FAIR, etc.).

Распределение выполненных на грид-сайтах RDIG задач

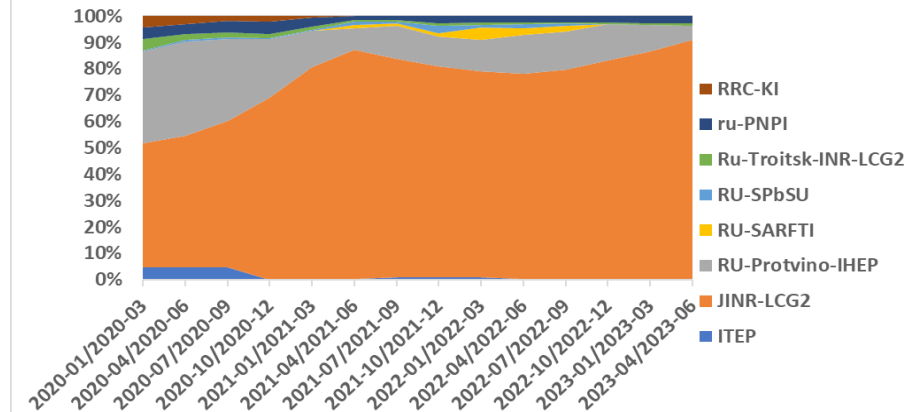


The JINR Tier2 output is the highest (90.31%) in the Russian Data Intensive Grid (RDIG) Federation.

Accounting - 2020_1 to 2023_5 normcpu on JINR Tier2 for VO and Quarter



Accounting - 2020_1 to 2023_5 normcpu for RDIG Tier2 and Quarter



Cloud Infrastructure



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



- Cloud Platform - OpenNebula
- Virtualization - KVM
- Storage (Local disks, Ceph)
- Total Resources
~ **5,152** CPU cores; 80 TB RAM;
3.5 PB of raw ceph-based storage

- Computational resources for neutrino experiments
- Testbeds for research and development in IT
- COMPASS production system services
- Data management system of the UNECE ICP Vegetation
- Scientific and engineering computing
- Service for data visualization
- VMs for JINR users

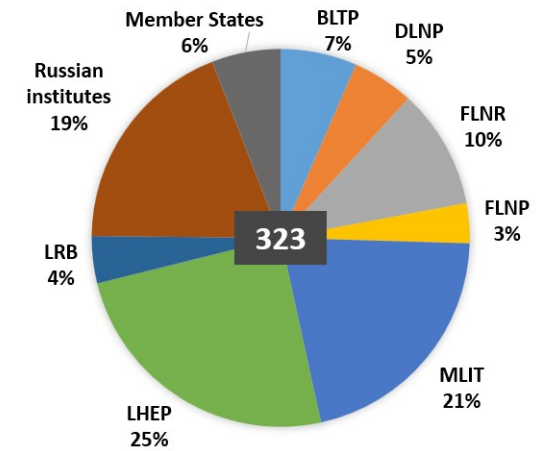
“Govorun” Supercomputer



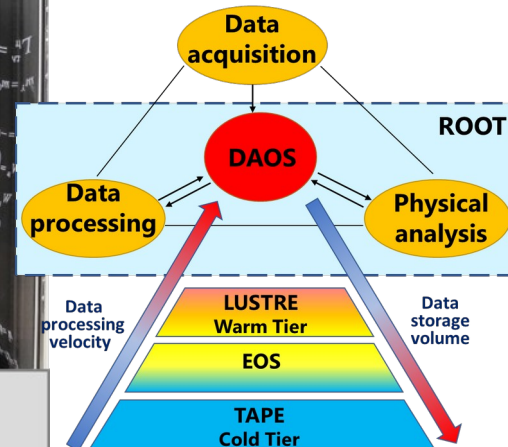
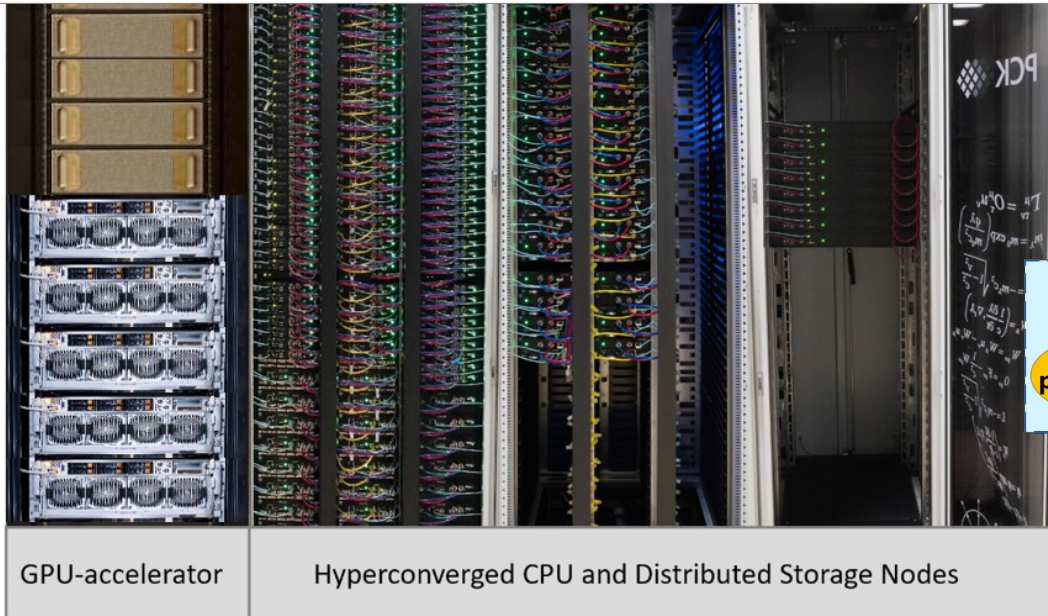
- Hyper-converged software-defined system
- Hierarchical data processing and storage system
- Scalable solution Storage-on-demand
- Total peak performance: 1.7 PFlops DP
- GPU component based on NVIDIA Tesla V100&A100
- CPU component based on RSC “Tornado” liquid cooling solutions
- The most energy-efficient center in Russia (PUE = 1.06)
- Storage performance >300 GB/s

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.



Total number of users : 323



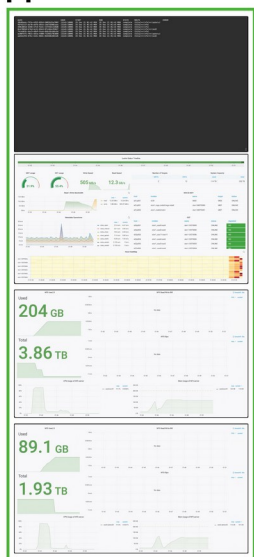
> 250 user papers (two in Nature Physics)

Unified Scalable Supercomputer Research Infrastructure

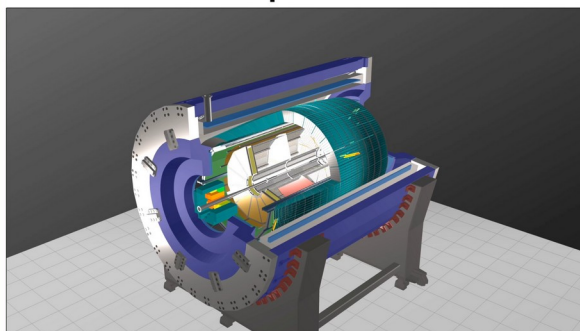


Based on the integration of the supercomputers of JINR, of the Interdepartmental Supercomputer Center of the Russian Academy of Sciences and of Peter the Great St. Petersburg Polytechnic University, a **unified scalable supercomputer research infrastructure** based on the National Research Computer Network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

ДАнные



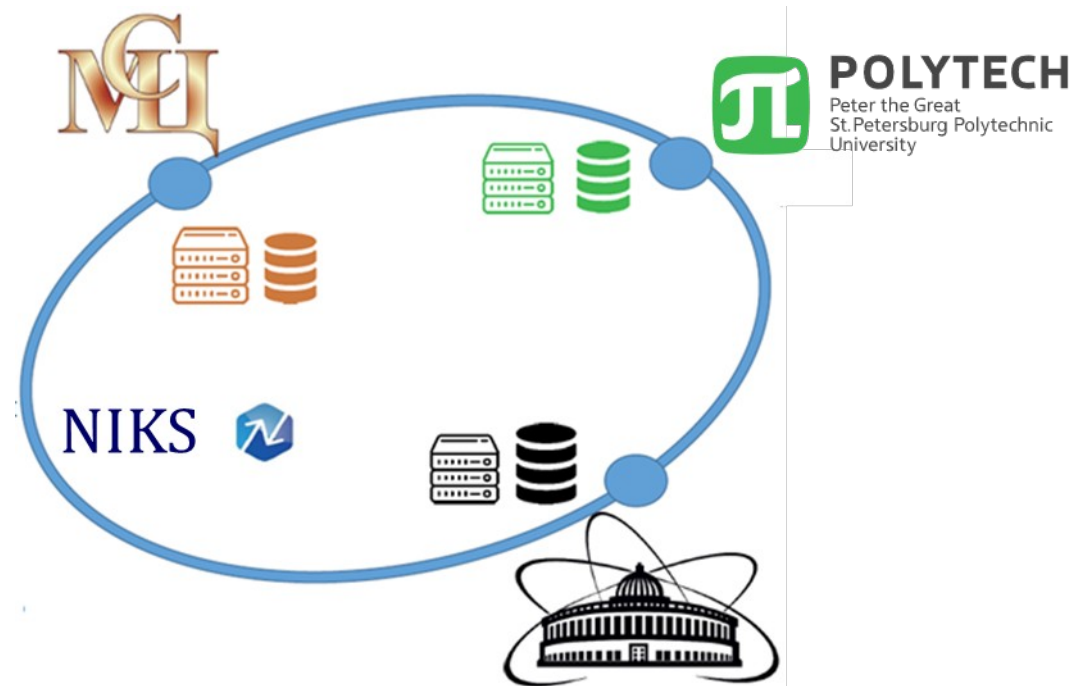
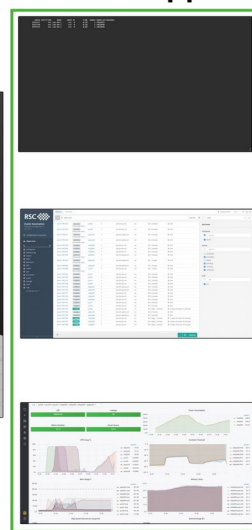
Центр управления виртуальным экспериментом Multi-Purpose Detector



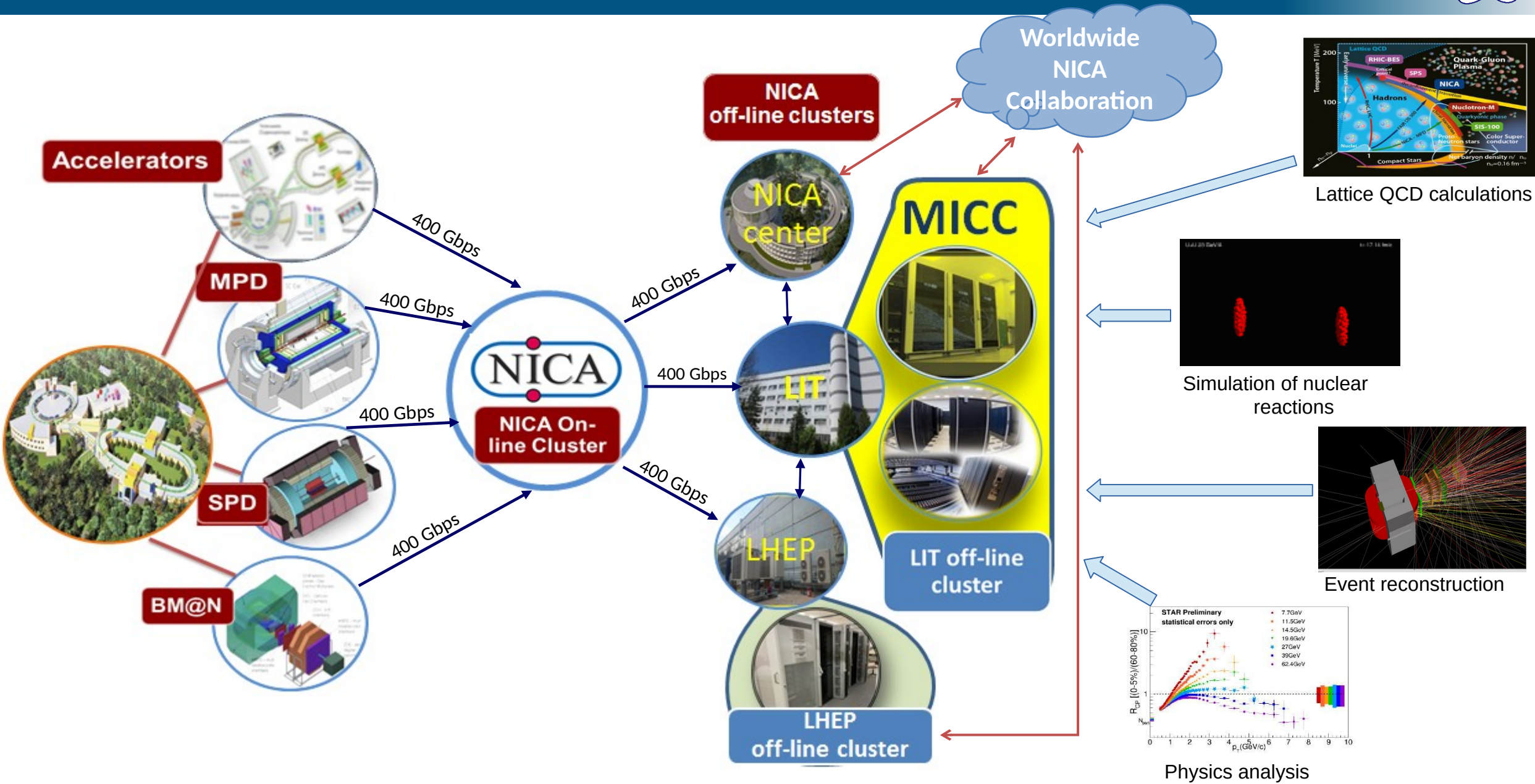
00:00:44:19



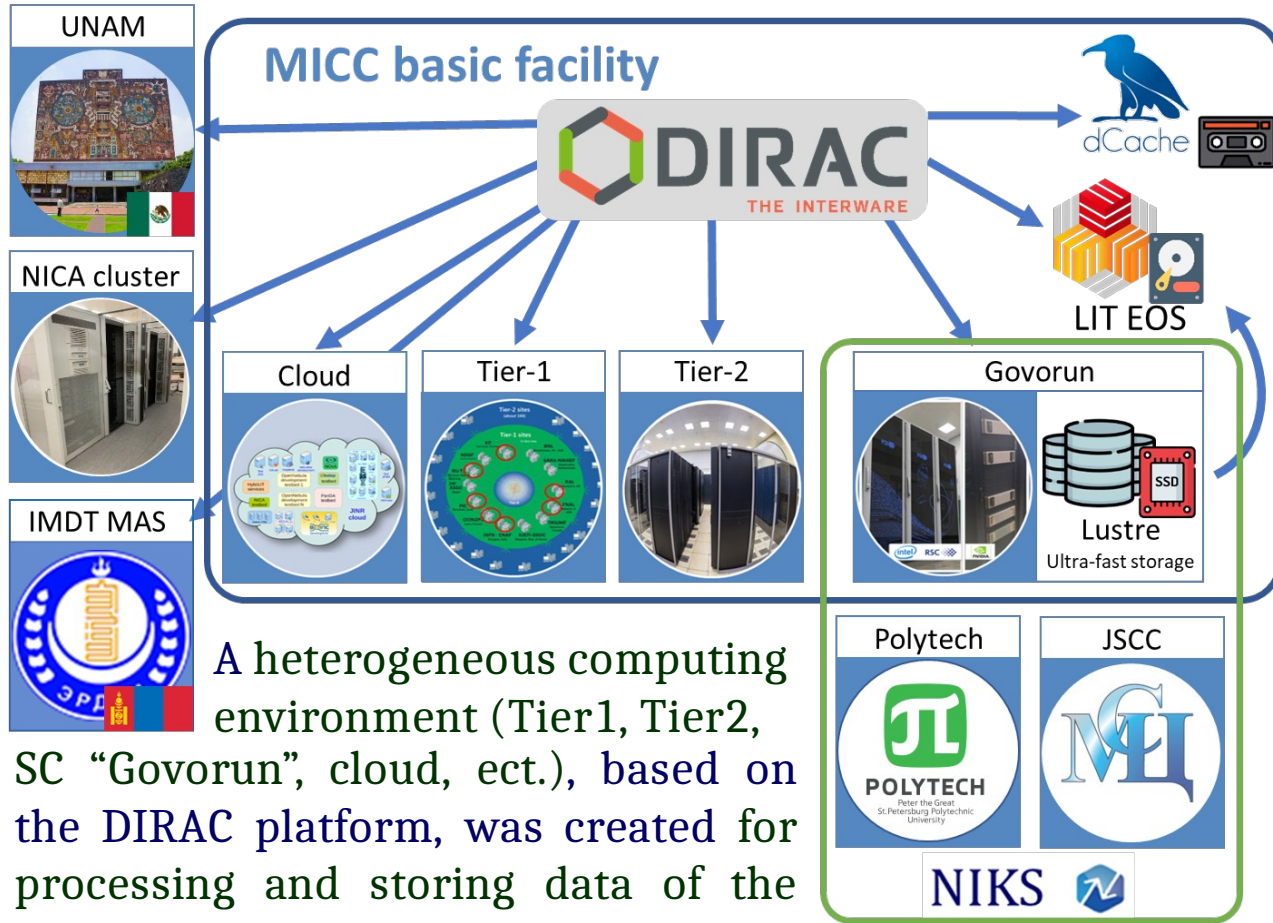
ЗАДАЧИ



NICA Computing Concept & Challenges



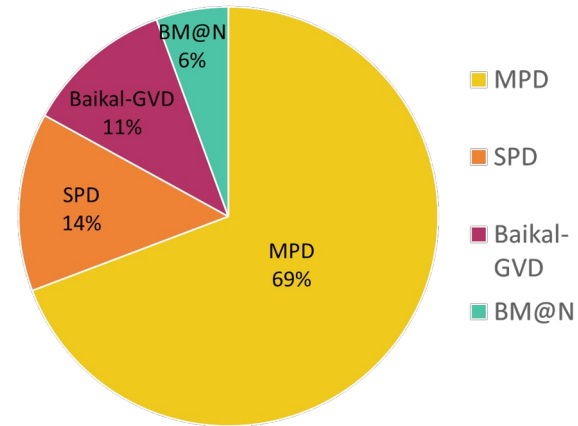
DIRAC-based distributed heterogeneous environment



A heterogeneous computing environment (Tier1, Tier2, SC "Govorun", cloud, ect.), based on the DIRAC platform, was created for processing and storing data of the experiments conducted at JINR.

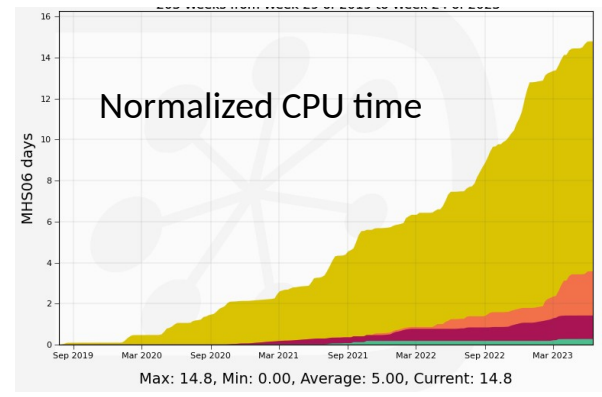
The distributed infrastructure is used by the MPD, Baikal-GVD, BM@N, SPD.

Use of DIRAC platform by experiments in 2019-2022

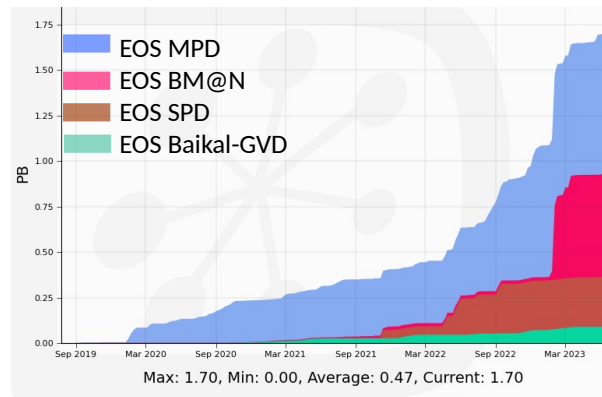


Total Number of executed jobs

The major user of the distributed platform is the MPD experiment



Data processed by experiments



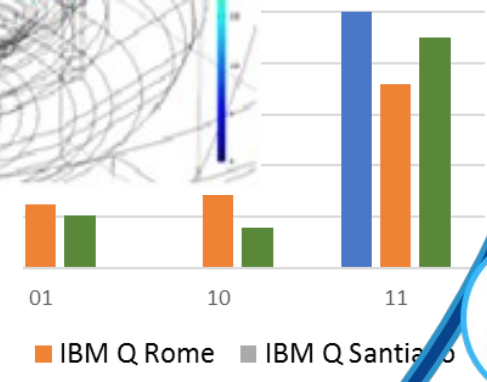
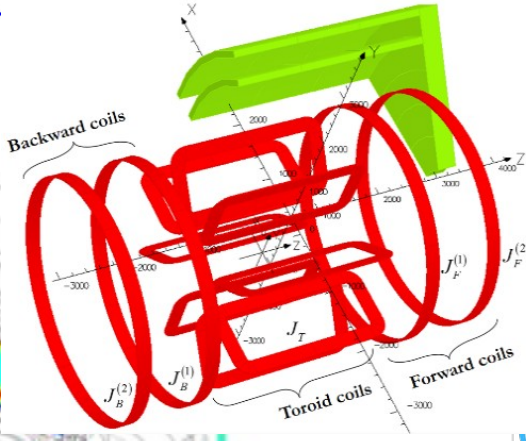
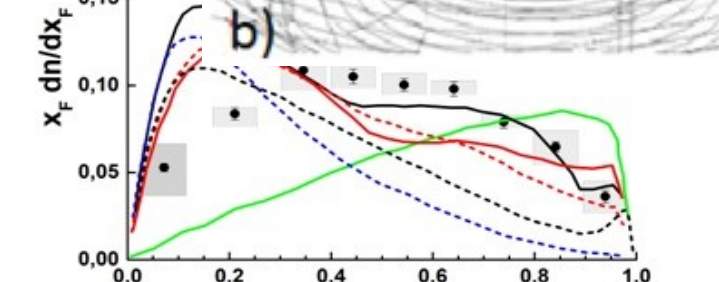
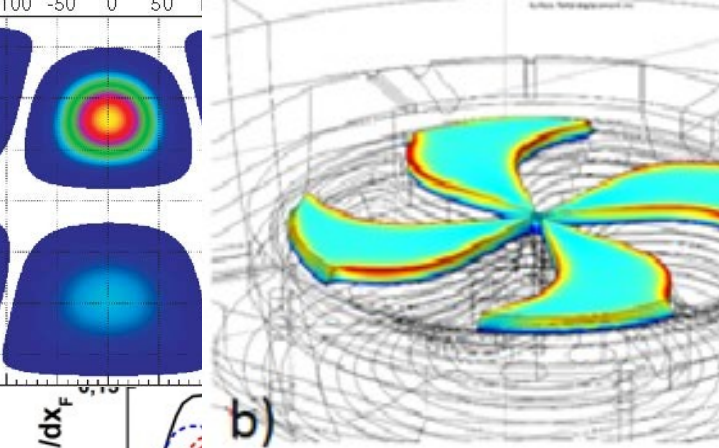
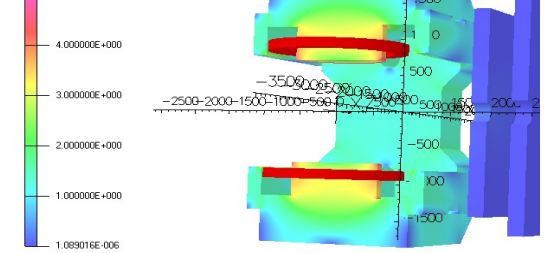
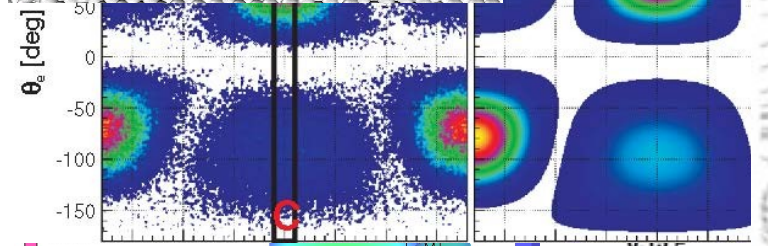
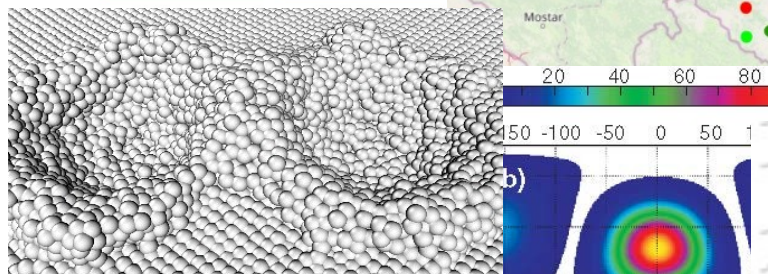
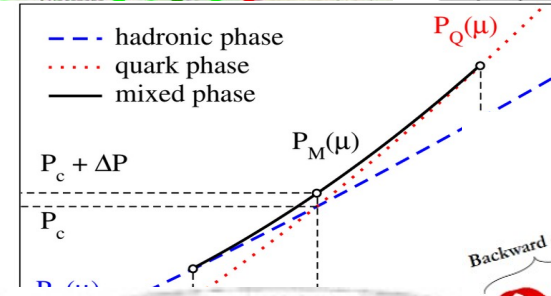
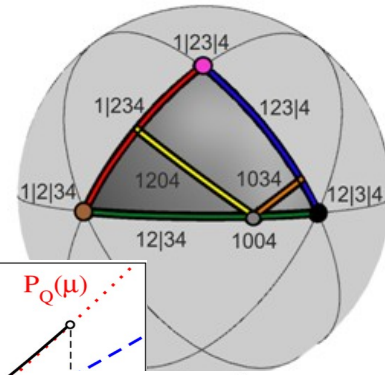
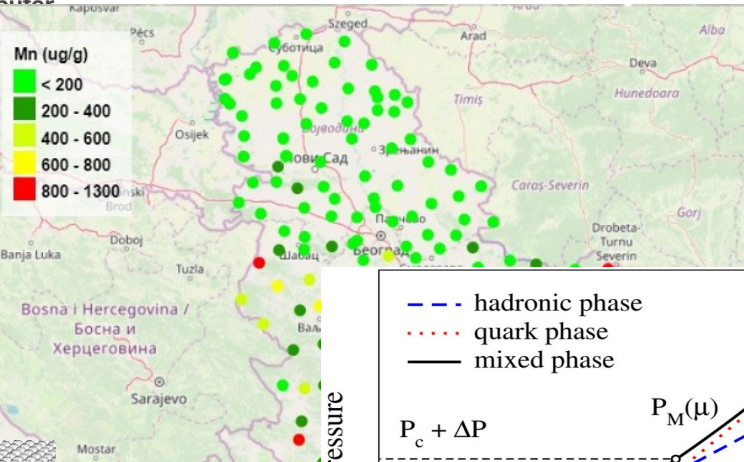
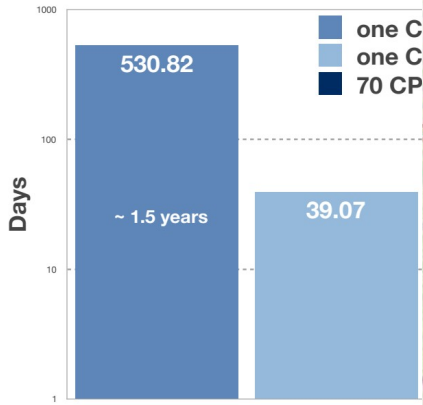
Summary statistics of using the DIRAC platform for MPD tasks in 2019-2022

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

Methods, Algorithms and Software



Govorun Supercomputer



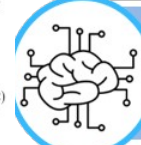
Numerical modeling of complex physical systems



Experimental data processing and analysis



Big Data



Machine and Deep learning



AI and robotics



Computer algebra



Quantum computing

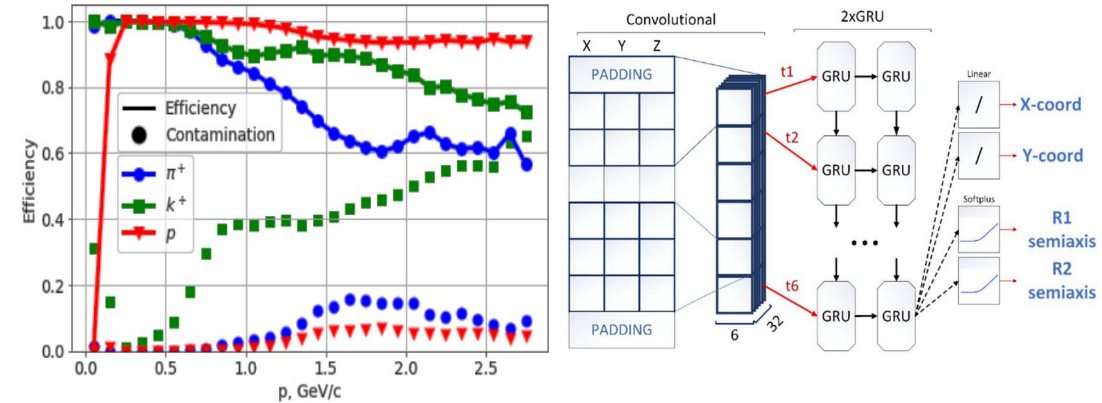
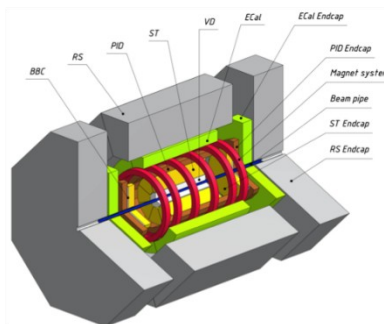
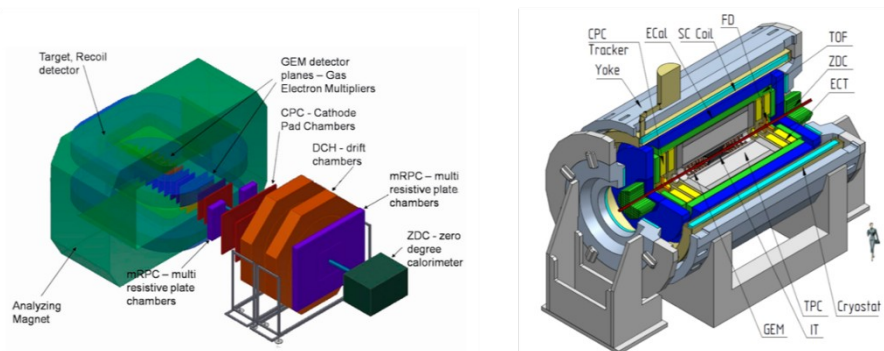
Implementation of ML/DL Methods in Data Processing and Analysis at the NICA Experiments: BM@N, MPD and SPD



Scientific and practical significance: **expanding the scope of machine learning methods**, in particular, in high energy physics; **software for experimental data processing and analysis** at the NICA accelerator complex; corresponding development of root-frameworks.

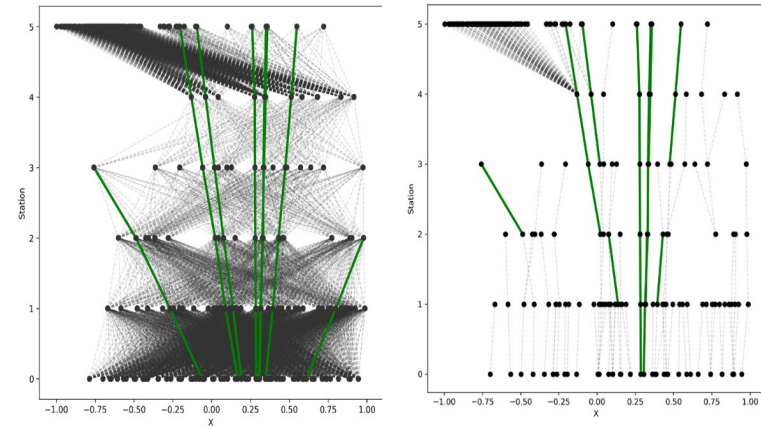
Possible areas for ML/DL application: hit finding, tracking, particle identification, decay reconstruction, global tracking.

The main ML/DL methods: Recurrent Neural Networks, Graph Neural Networks, Convolutional Neural Networks, Decision Trees, Gradient Boosting, etc.



Gradient-boosted decision trees for PID in MPD

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



Graph Neural Networks for Tracking

The participants are presented by members of all targeted international collaborations: BM@N, MPD, SPD.

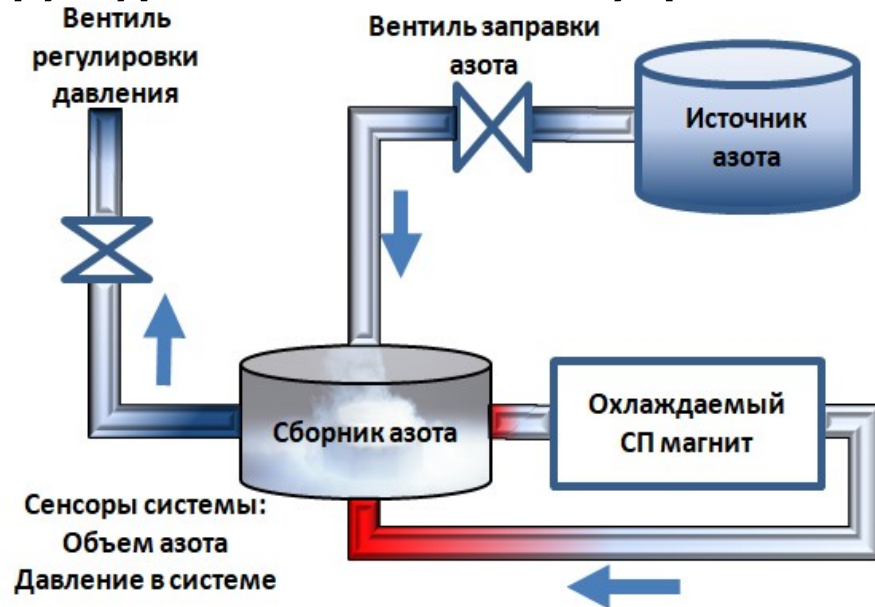
ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ДАВЛЕНИЕМ АЗОТА В КРИОГЕННОЙ УСТАНОВКЕ

Представлены результаты внедрения интеллектуальных систем управления на основе нечёткой логики, нейронных сетей, генетических и квантовых алгоритмов к задаче стабилизации давления азота в криогенной системе испытательного стенда фабрики магнитов ЛФВЭ ОИЯИ.

Апробированы все этапы технологии проектирования встраиваемых интеллектуальных систем управления на основе мягких и квантовых вычислений.

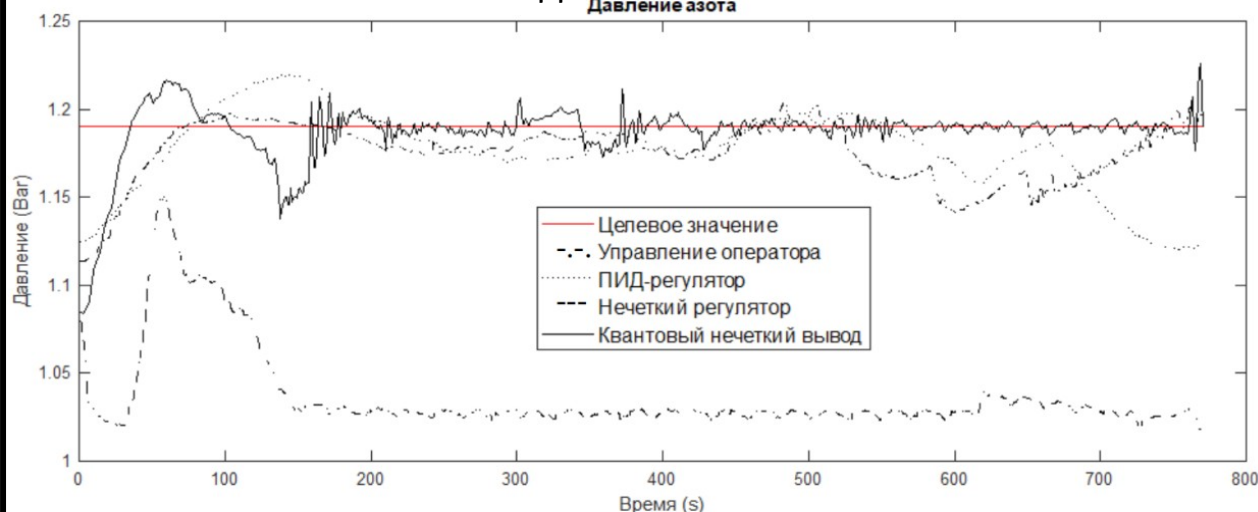
Не изменяя аппаратный уровень установки, повышена эффективность функционирования, уменьшен расход полезного ресурса.

Структурная схема объекта управления

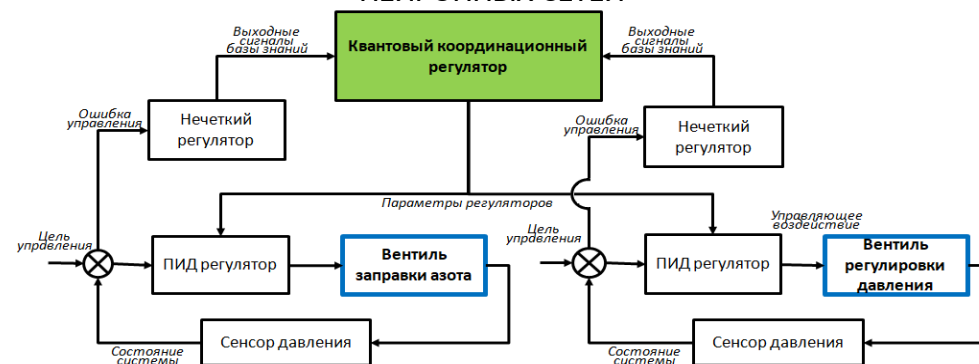


РОБАСТНОЕ УПРАВЛЕНИЕ ДАВЛЕНИЕМ ВО ВРЕМЯ ЗАПРАВКИ АЗОТА

Давление азота

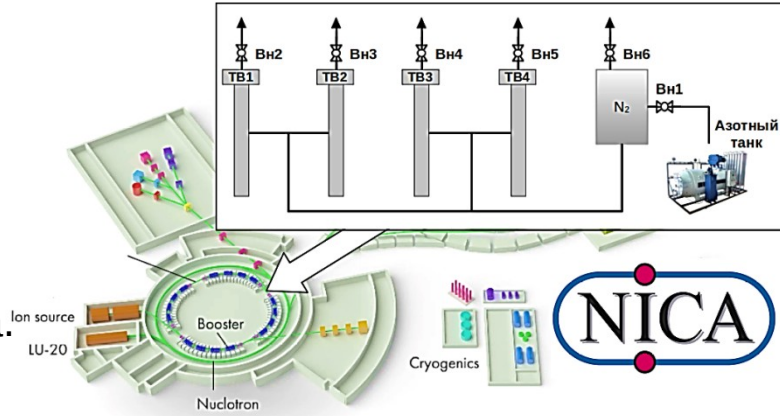


КВАНТОВЫЙ КООРДИНАЦИОННЫЙ РЕГУЛЯТОР НА ОСНОВЕ НЕСКОЛЬКИХ НЕЙРОННЫХ СЕТЕЙ



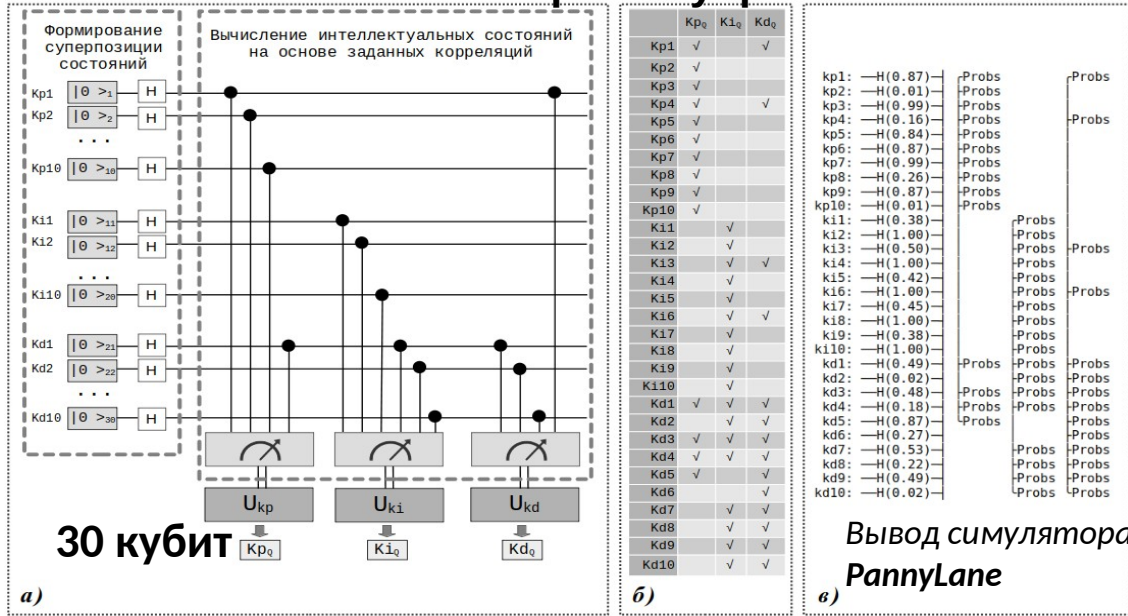
Моделирование квантового нечёткого вывода для скоординированного управления 5 вентилями

- Координационное управление пятью контроллерами;
- 3 выходных значения из каждого регулятора.

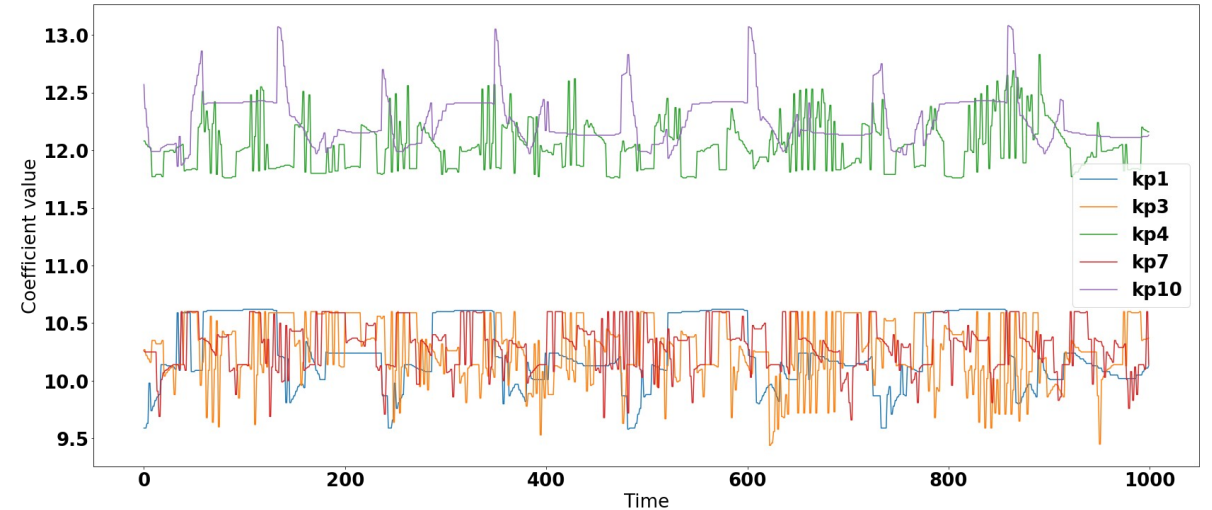


Моделирование работы квантового алгоритма проводилась с помощью симулятора PannyLane на СК "Говорун".

Квантовая схема алгоритма управления



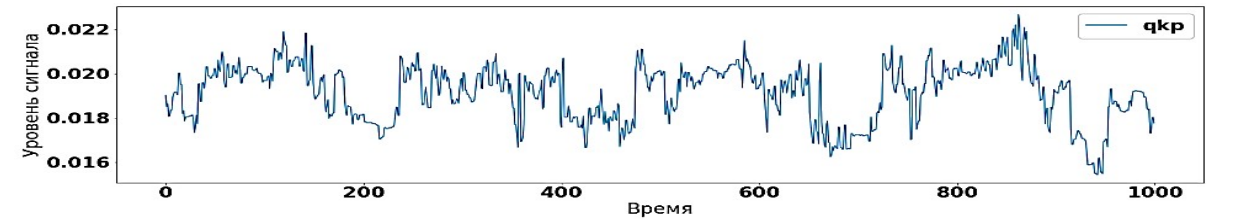
Changing the coefficient value over time



Изменение значения пропорционального коэффициента нечётких регуляторов



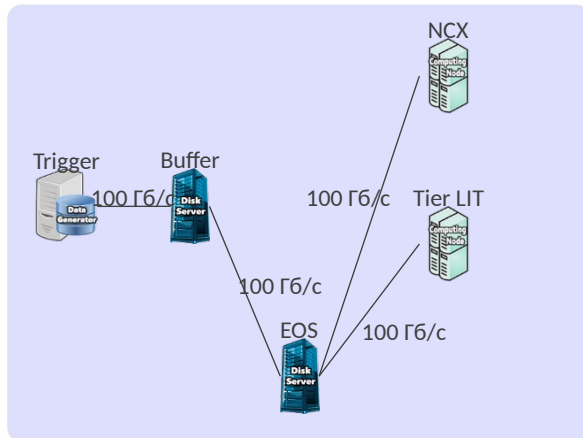
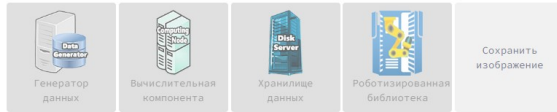
Получение результирующего сигнала



Программный комплекс для создания цифровых двойников распределенных центров сбора, хранения и обработки данных

Цифровой двойник РЦОД

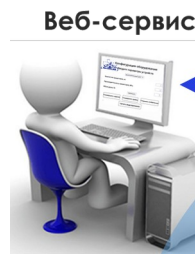
Построение инфраструктуры центра сбора, хранения и обработки данных



Примеры применения для экспериментов комплекса NICA

- Построение цифрового двойника вычислительной инфраструктуры эксперимента VM@N.
- Построение цифрового двойника вычислительной системы онлайн-фильтра данных эксперимента SPD.

- Редактировать инфраструктуру
- Настроить каналы связи
- Настроить потоки данных
- Настроить потоки задач
- Создать цифровой двойник



Особенности

- ✓ Универсальность.
- ✓ Учитываются важные функциональные параметры распределенных центров:
 - характеристики оборудования;
 - характеристики потоков данных и задач;
 - вероятности сбоев, отказов и изменений в производительности оборудования и других процессов, происходящих в системе.

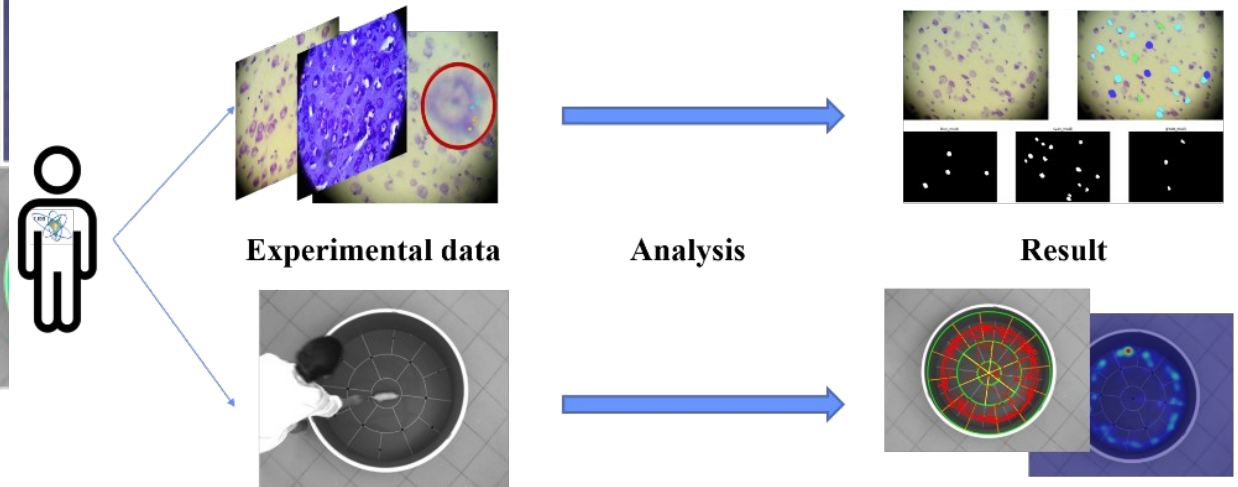
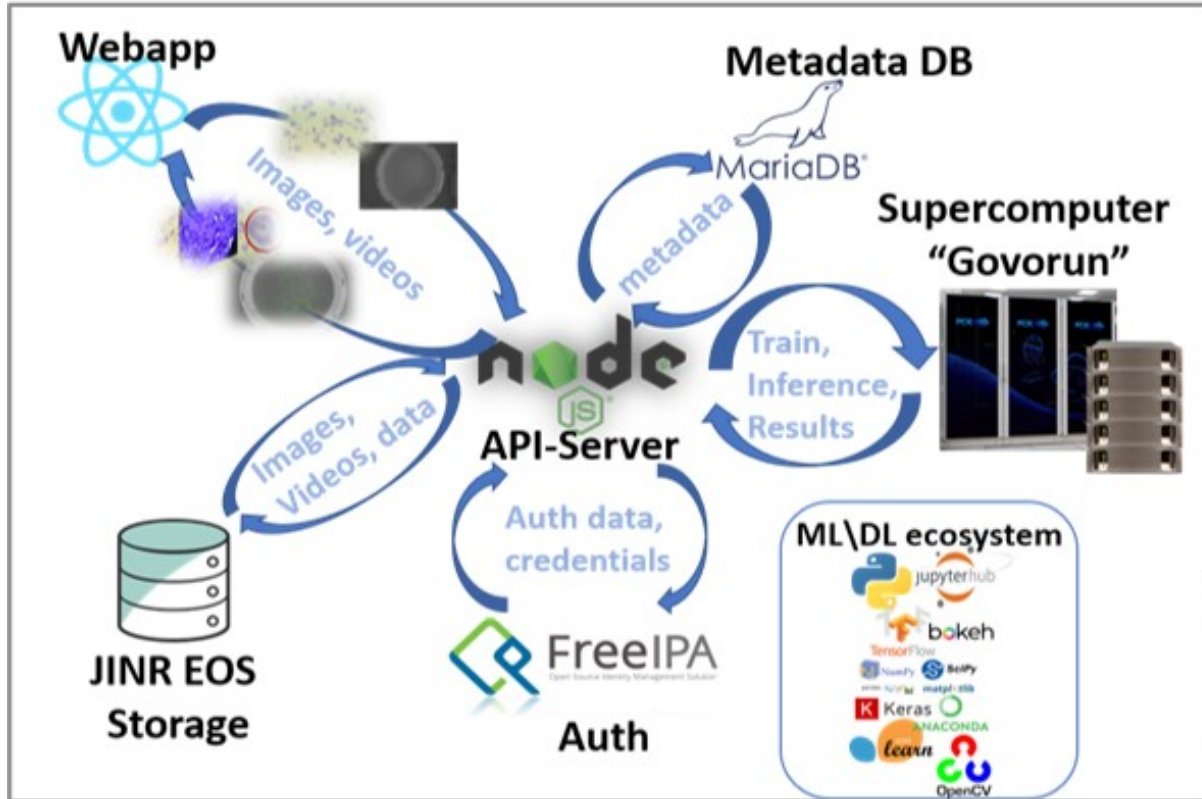
Результаты работы отличаются от результатов работы существующего распределенного центра не более, чем на 20%.

Information System for Radiation Biology Tasks



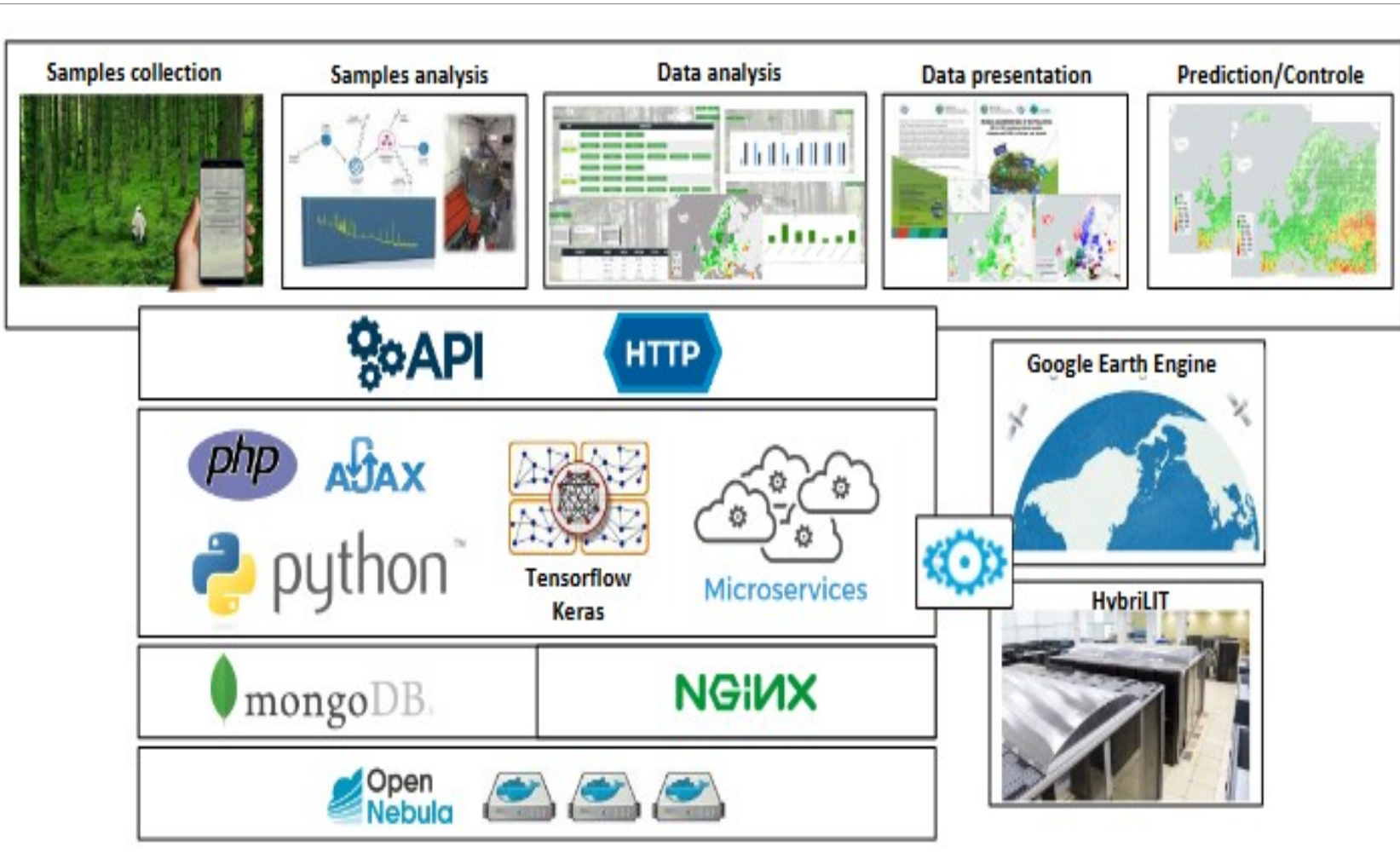
The joint project of MLIT and LRB is focused on creating an Information System (IS) as a set of IT solutions.

The information system allows one to store, quickly access and process data from experiments at LRB 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



Conceptual scheme of the service

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 technologies of machine and deep learning is in progress. On the MLIT cloud platform, the Data Management System (DMS) of the UNECE ICP Vegetation was created to provide its participants with a modern unified system of collecting, analyzing and processing biological monitoring data.

Artificial Intelligent in Agriculture



Доктор для растений в вашем телефоне

Диагностика по фото
Сфотографируйте или загрузите с устройства

Выберите наиболее подходящий вариант

Лечение
Получите советы и рекомендации от лучших агрономов

ТОЛСТЯНКА, СУККУЛЕНТ, ЛАВАНДА, МАНГО, ПЕТУНИЯ, КАКТУС, КЛУДНИКА, АГЛАОНЕМА, ФИКУС, ПОМИДОРЫ, АГЛАОНЕМА, БАРБАРИС, СПАТИФИЛЛУМ, ОГУРЦЫ, ПАПРОТНИК, КУКУРУЗА, МОРКОВЬ, ХРИЗАНТЕМА, СМОРОДИНА, ВИНОГРАД, РОЗМАРИН, ТУЯ, ГАМЕДОРЕЯ, ОРХИДЕЯ, ЛОБЛОН, РОЗМАРИН, МОНСТЕРА, ГЕРНИКА, АНТУРИУМ, ЗАМКОКУЛЬКАС, ДЕКАБРИСТ, КАЛЕУС, СВЕКЛА, БЕГОНИЯ, АЛМОКОУЛЬКАС, СТОРОБИТУМ, РОЗЫ, НИОНЫ, ТЮЛЬПАН, БАЗИЛИС, АСТРА, МАРАНТА, ДРОЗДИ, ПШЕНИЦА, АВОКАДО, КАПУСТА, ПЕРЦЕЦ, ГЕРАНЬ, КАРТОФЕЛЬ, ГЕОРГИНЫ, МАНИОКА, САНСЕВИЕРИЯ, ДИФФЕНБАХИЯ, ГОРОХ, ПЕТРУШКА, АЛОЭ, ТРАДЕСКАНЦИЯ, ВИШНЯ, БАЛЬЗАМ, ДАМБУК, ЕЖЕВИКА, ГИДИСКУС, УРАЦЕНА, КЛОНОК, САЛАТ, ФИЯЛКА

```
{
  "error": 0,
  "general_predictions": [
    { },
    { },
    { }
  ],
  "custom_predictions": [
    { },
    {
      "type": "rose",
      "prediction": [
        {
          "name": "Мозаика",
          "sample": "http://userbb.ru/botainika/diseases/rmv1.jpg",
          "description": "<ор>Это наиболее распространенное заболевание вирусного типа. Узнать его очень просто - листья покрываются хлоротичными пятнами и узорами, которые могут перейти в прожилковой хлороз. Листья деформируются (узколистность, курчавость, искривленность) и постепенно опадают. Рост побегов замедляется, и они не вызревают. Побеги, которые больны, вырезают. При дальнейшем распространении болезни растение уничтожают"
        }
      ]
    }
  ]
}
```

MLIT scientists have developed an online platform pdd.jinr.ru and mobile application (DoctorP) to detect diseases of indoor and agricultural plants. The neural network architecture of the platform can detect 68 classes of various diseases and pests with an accuracy of more than 95%. Experts have collected more than 4,000 images in the database and received more than 70,000 requests from users. Everyone can use the platform interface, starting from agricultural holdings and ending with novice gardeners.

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

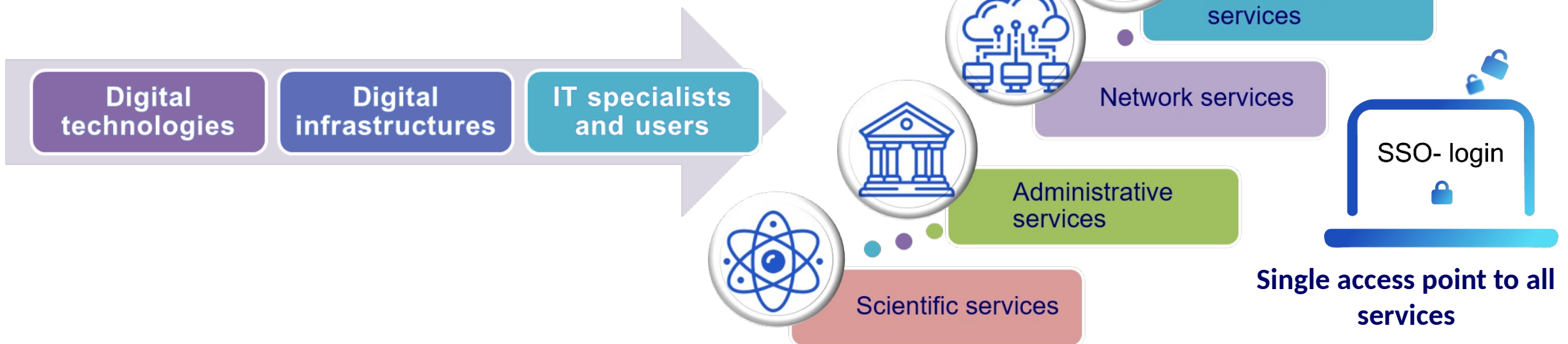
to enable

a comprehensive analysis of information

using

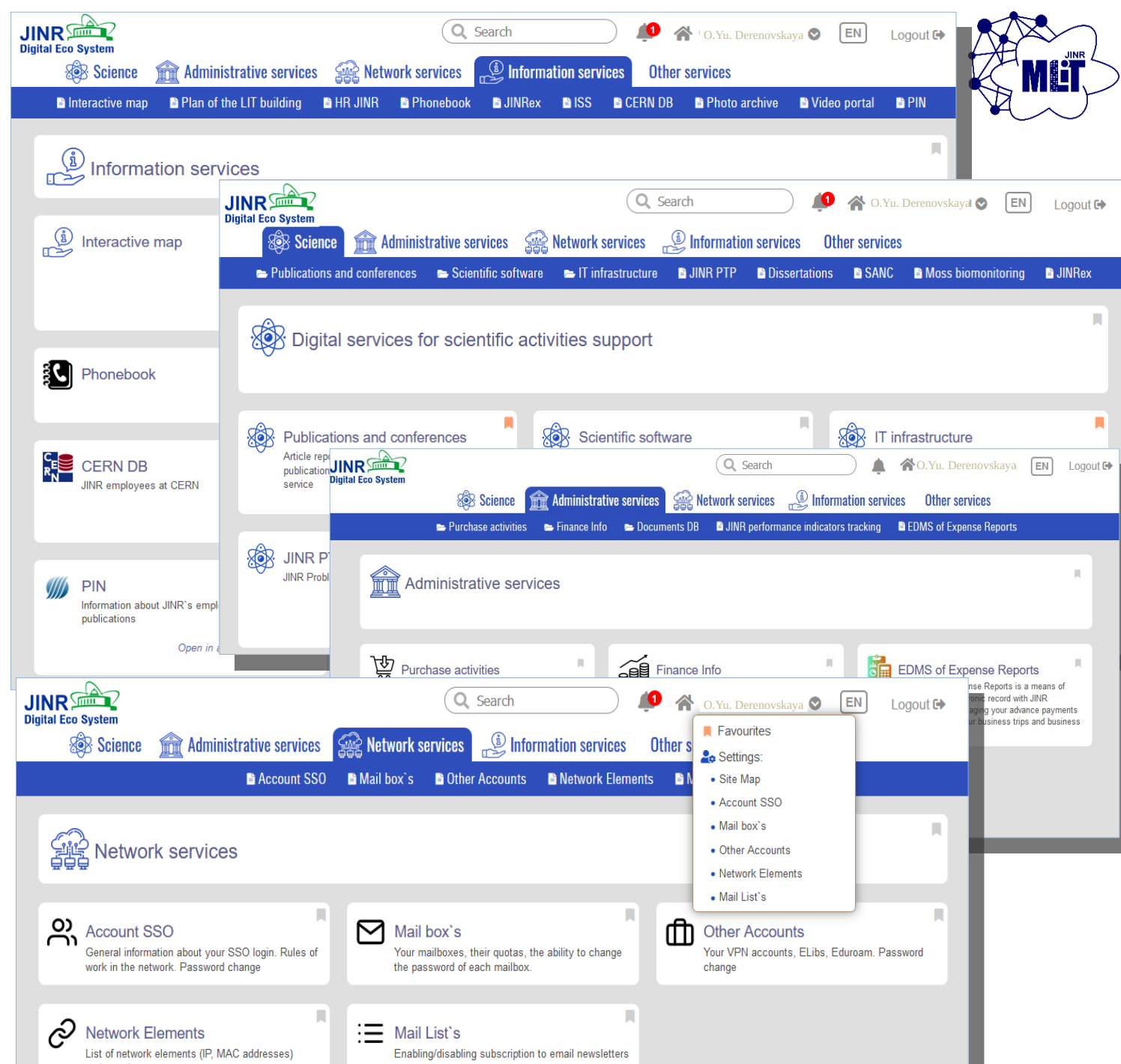
modern Big Data technologies and artificial intelligence.

JINR
Digital Eco System





- ✓ 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





Science Administrative services Network services Information services Other services

Interactive map Plan of the LIT building HR JINR Phonebook JINRex ISS CERN DB Photo archive Video portal PIN

Interactive map

Поиск сотрудников
Поиск зданий

Объекты

- Автобус
- Административные здания
- Аллея
- Библиотека
- Велопарковка
- ДепРНК
- Инженерные объекты
- Инфраструктура
- ЛИТ
- ЛНФ
- ЛРФ
- ЛТО
- ЛЯП
- Мед. пункт
- Мемориал
- НИКА
- ОГЭ
- Парковка
- Проходная
- Столовая, кафе

Центр NICA

Линейный ускоритель

ИТ/Ск. комплекс

50 m
200 ft

50 m
300 ft

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

Science Administrative services Network services Information services Other services

Interactive map Plan of the LIT building HR JINR Phonebook JINRex ISS CERN DB Photo archive Video portal PIN

Interactive map

100 m
300 ft

Science Administrative services Network services Information services Other services

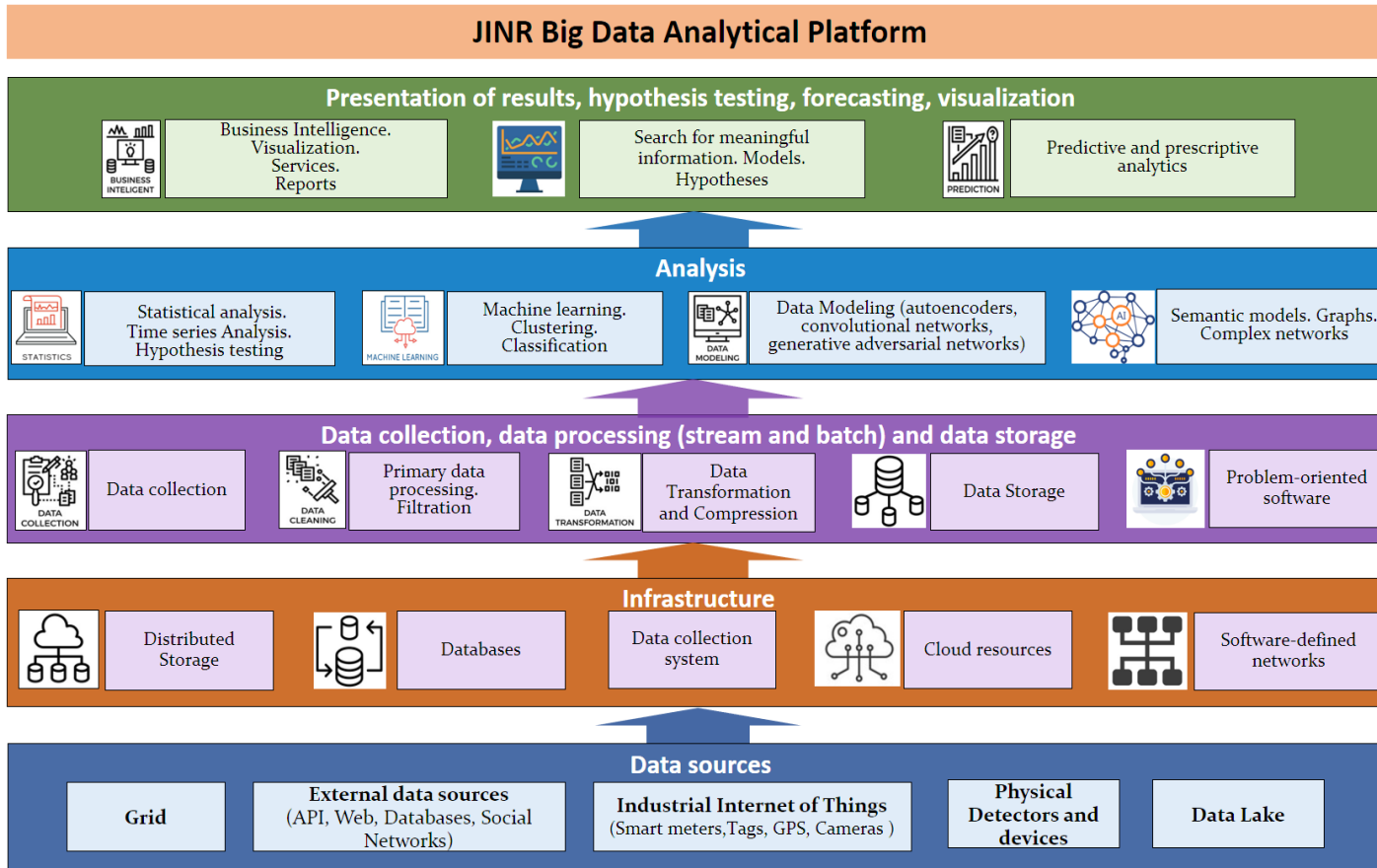
Interactive map Plan of the LIT building HR JINR Phonebook JINRex ISS CERN DB Photo archive Video portal PIN

Plan of the LIT building

10 m
30 ft

Комната: 324
Телефон: 2162526
Кореньков Владимир Васильевич
Отдел: Руководство
Площадь: 84 кв. м

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



Tools for debugging and profiling parallel applications



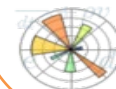
Work with applied software packages

COMSOL MULTIPHYSICS

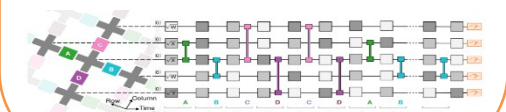
Wolfram Mathematica



Frameworks and tools for ML/DL tasks



Quantum algorithms, quantum programming and quantum control





International Conference “Distributed Computing and Grid Technologies in Science and Education”



- Distributed computing systems
- Computing for MegaScience Projects
- Distributed computing & HPC applications
- Distributed Storage Systems
- HPC
- Cloud technologies
- Big data Analytics and Machine learning
- Quantum informatics and computing

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.

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

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



10

GRID2023

3-7 July 2023

10th International Conference
“Distributed Computing and Grid Technologies in
Science and Education”



Conference Topics:

1. Distributed Computing Systems
2. HPC
3. Distributed Computing and HPC Application
4. Cloud Technologies
5. Computing for MegaScience Projects
6. Quantum Informatics and Computing
7. Big Data, M/D Learning, Artificial Intelligence
8. Student session

Workshop “Computing for radiobiology and medicine”

Workshop “Modern approaches to the modeling of research reactors, creation of the “digital twins” of complex systems”

Round table “RDIG-M - Russian distributed infrastructure for large-scale scientific projects in Russia”

Round table on IT technologies in education

More than **275** participants

In person - 216

Remotely - 60

30 Plenary reports

135 Sessional reports

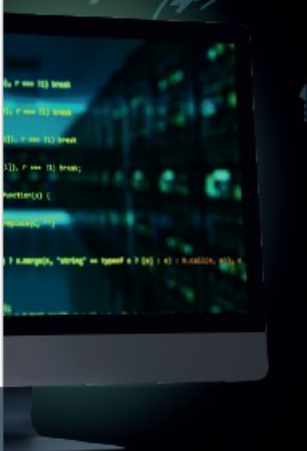
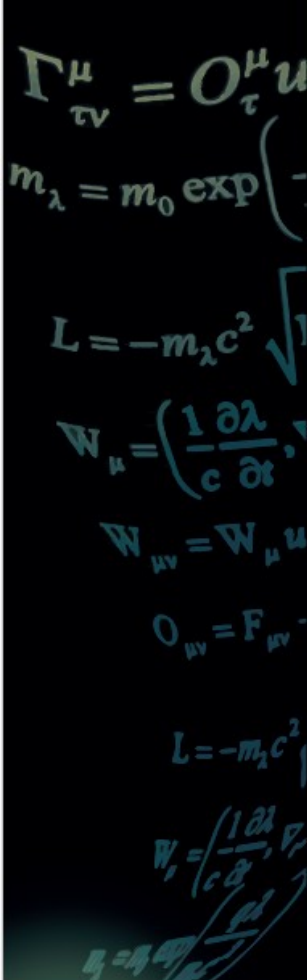
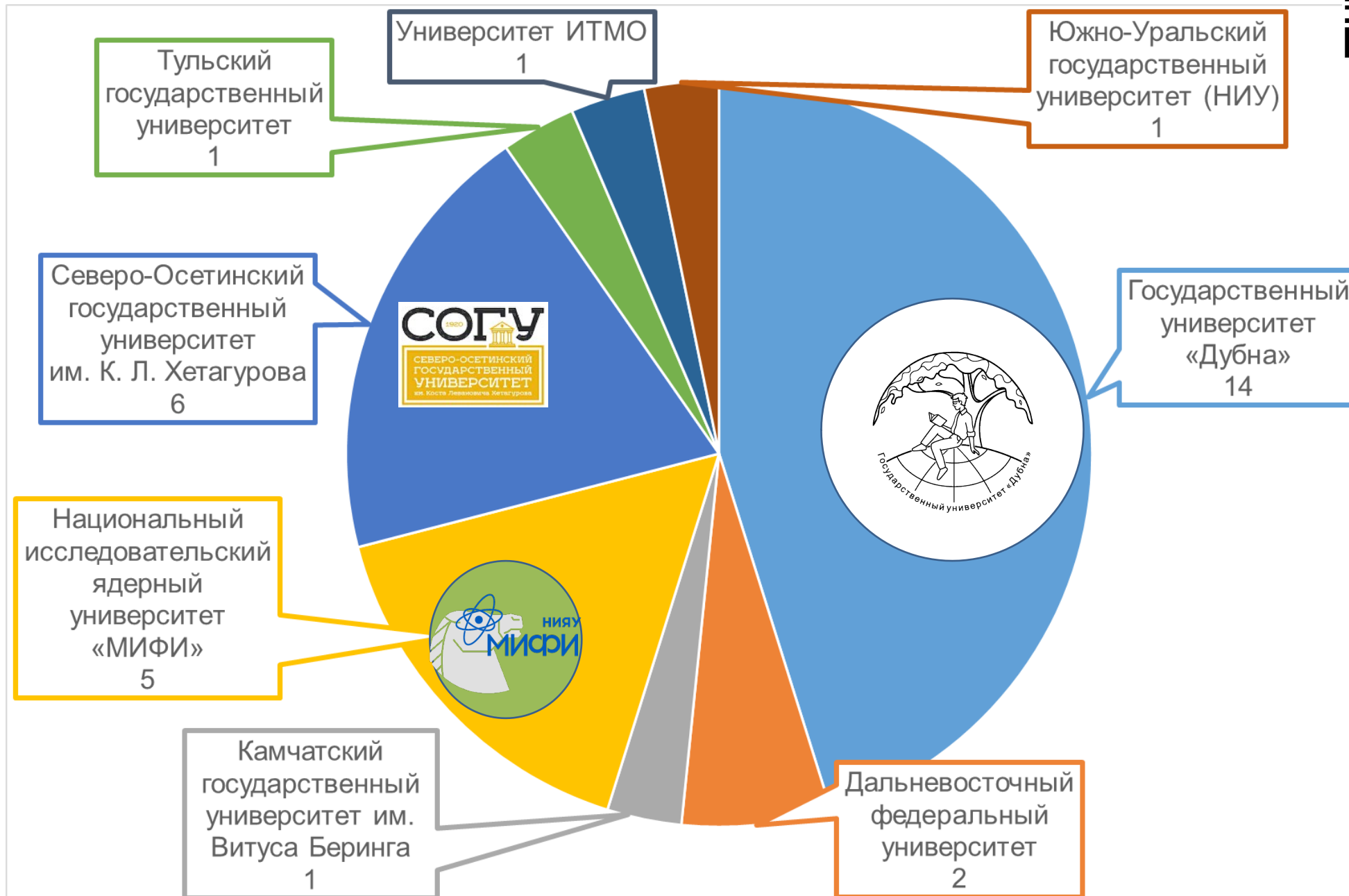
17 Countries: Azerbaijan, Armenia, Belarus, Bulgaria, the Czech Republic, Egypt, Germany, Georgia, Iran, Kazakhstan, Mexico, Moldova, Mongolia, Serbia, CERN and Uzbekistan. **Russia** was represented by participants **from 41 universities and research centers.**



JINR School of Information Technology 2023

50 students from 11 Russian universities







BMK MGU



Общие отцы-основатели ВМК МГУ и ЛВТА ОИЯИ (ныне ЛИТ им. М.Г. Мещерякова) внесли огромный вклад в развитие программного обеспечения для БЭСМ-6, численных методов, вычислительной физики и компьютеринга. И сегодня продолжает развиваться сотрудничество между МГУ и ЛИТ в области:

- математических методов и математического моделирования,
- грид-технологий в России,
- аналитики Больших данных,
- суперкомпьютерных технологий и методов параллельных вычислений,
- организации и проведения конференций и школ.



Академик Самарский Александр Андреевич



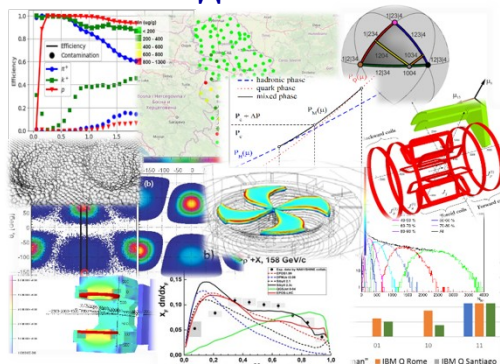
Член-корр. Говорун Николай Николаевич

Академик Тихонов Андрей Николаевич

Потребность в подготовке высококвалифицированных кадров в области математического моделирования и обработки данных проектов класса мегасайнс с применением методов аналитики Больших данных и искусственного интеллекта привела к идее создания на базе филиала МГУ в Дубне направления подготовки «Прикладная математика и информатика»

Магистерской программы

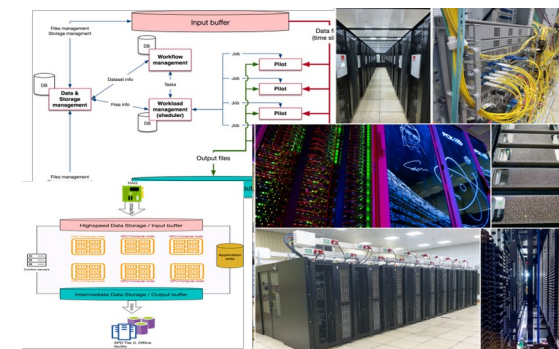
«Методы и технологии обработки данных в гетерогенных вычислительных средах»



Математическое моделирование, численные методы и комплексы программ



Глубокое машинное обучение и аналитика больших данных



Компьютинг (программные средства и модели) для проектов класса мегасайнс

Создана рабочая группа по подготовке магистерской программы, планируется создание координационного комитета совместно с ВМК МГУ

Магистерская программа филиала МГУ в Дубне «Методы и технологии обработки данных в гетерогенных вычислительных средах»

Дисциплины базовой части


1. Иностранный язык
2. Современная философия и методология науки
3. История и методология прикладной математики и информатики
4. Дифференциальные уравнения в прикладных задачах
5. Методы моделирования физических процессов
6. Вероятностные модели
7. Аналитика больших данных и искусственный интеллект
8. Машинное обучение и мягкие вычисления
9. Методы и технологии параллельных вычислений для повышения производительности компьютерного моделирования

Дисциплины вариативной части

1. Моделирование прохождения элементарных частиц через вещество
2. Языки и технологии анализа данных
3. Статистические методы обработки данных
4. Структуры и алгоритмы обработки данных
5. Модели компьютерного моделирования для обработки и анализа данных для крупных научных проектов
6. Распределенные вычисления и облачные технологии
7. Платформы, технологии и приложения больших данных
8. Введение в информационную безопасность
9. Квантовые вычисления
10. Научно-исследовательский семинар

Дисциплины по выбору

1. Информационные технологии в мегасайенс экспериментах (на английском языке)
2. Суперкомпьютерные системы и приложения (на английском языке)
3. Биоинформатика (на английском языке)
4. Методы и технологии машинного обучения в прикладных задачах.
5. Технологии анализа данных с применением открытых библиотек на языке программирования Python.
6. Web-технологии для решения научных и прикладных задач.
7. Платформы и инструменты для web-разработки.
8. Математические модели в радиационной биологии и ядерной медицине
9. Моделирование процессов физики высоких энергий.
10. Организация хранения и управления данными в крупных научных проектах.



Поздравляем с Юбилеем!
Московский университет - 270 лет
Садовничий Виктор Антонович - 85 лет