



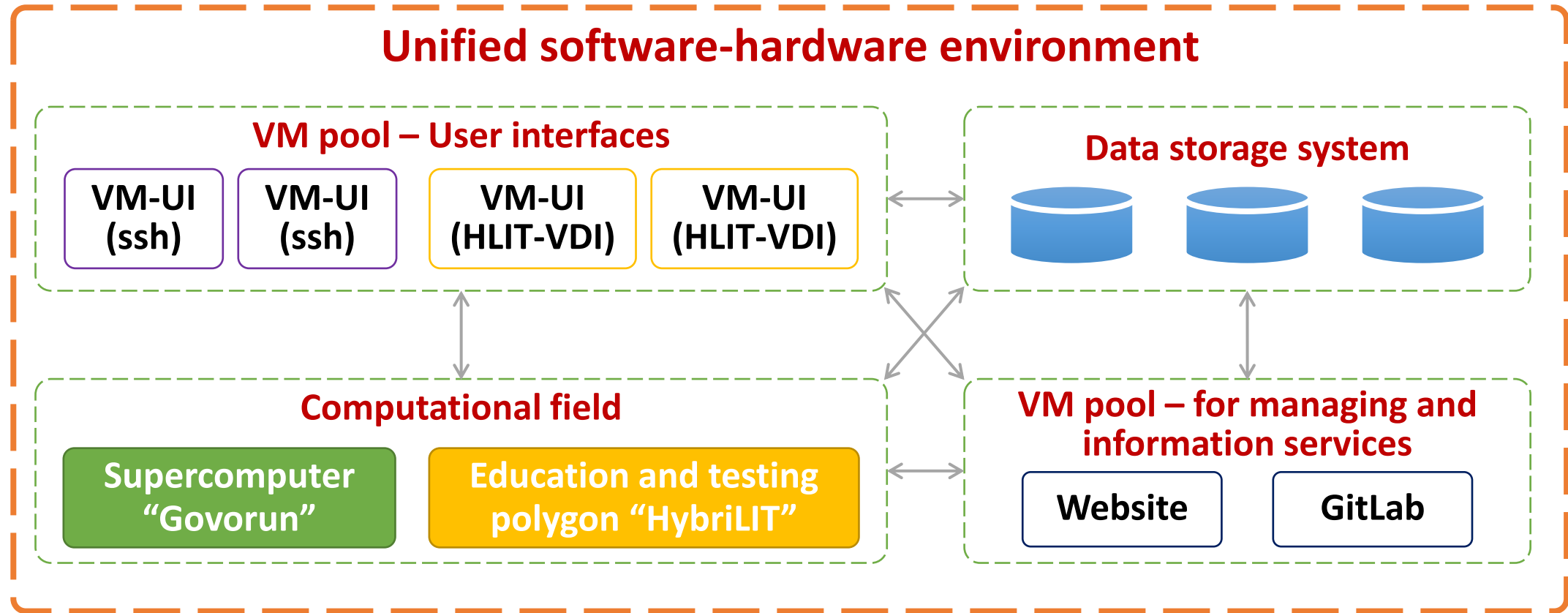
# Задачи, решаемые на суперкомпьютере «Говорун»

Дмитрий Подгайный

Лаборатория информационных технологий им. М.Г. Мещерякова

Осенняя школа по информационным технологиям ОИЯИ  
14 - 19 ноября 2022

# MICC component: HybriLIT platform

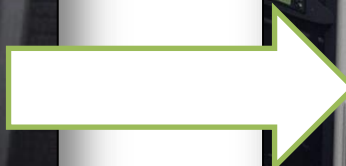


The **unified software and information environment** of the HybriLIT platform allows users to use the education and testing polygon is aimed at exploring the possibilities of novel computing architectures, IT-solutions, to develop and debug their applications, furthermore, carry out calculations on the supercomputer, which allows them to effectively use the supercomputer resources.

# Development of the heterogeneous HybriLIT platform

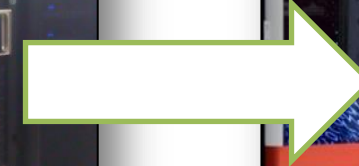


**Cluster HybriLIT 2014:**  
Full peak performance:  
**140 TFlops** for single precision;  
**50 TFlops** for double precision



**#18 в Top50**

**“Govorun” supercomputer**  
First stage **2018:**  
Full peak performance :  
**1 PFlops** for single precision  
**500 TFlops** for double precision  
**9th** in the current edition of the  
**IO500** list (July 2018)



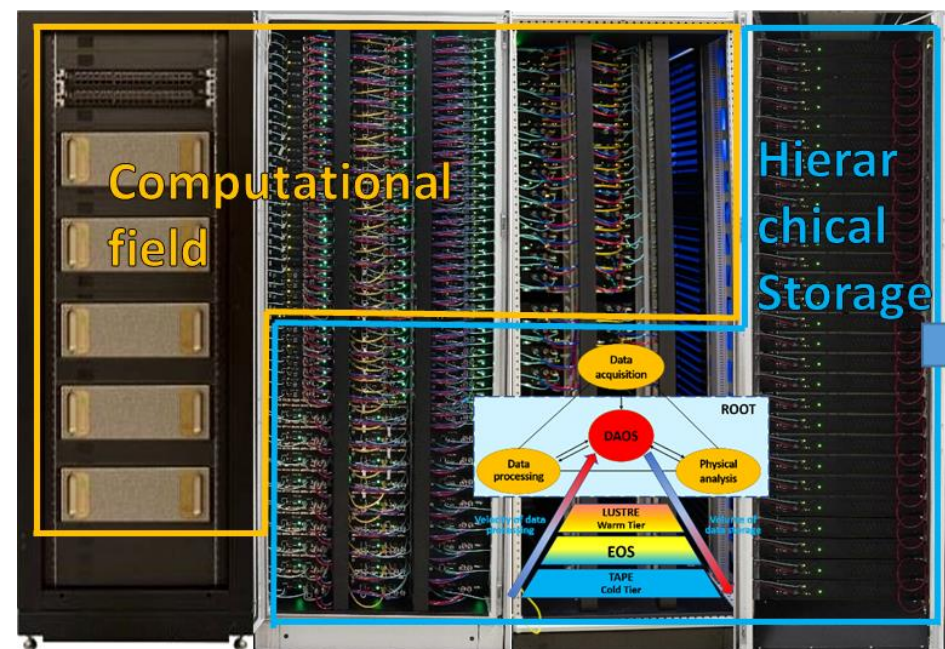
**#10 в Top50**

**“Govorun” supercomputer**  
Second stage **2019:**  
Full peak performance :  
**1.7 PFlops** for single precision  
**860 TFlops** for double precision  
**288 TB CCXD** with I/O speed **>300 Gb/s**  
**17th** in the current edition of the **IO500**  
list (July 2020)

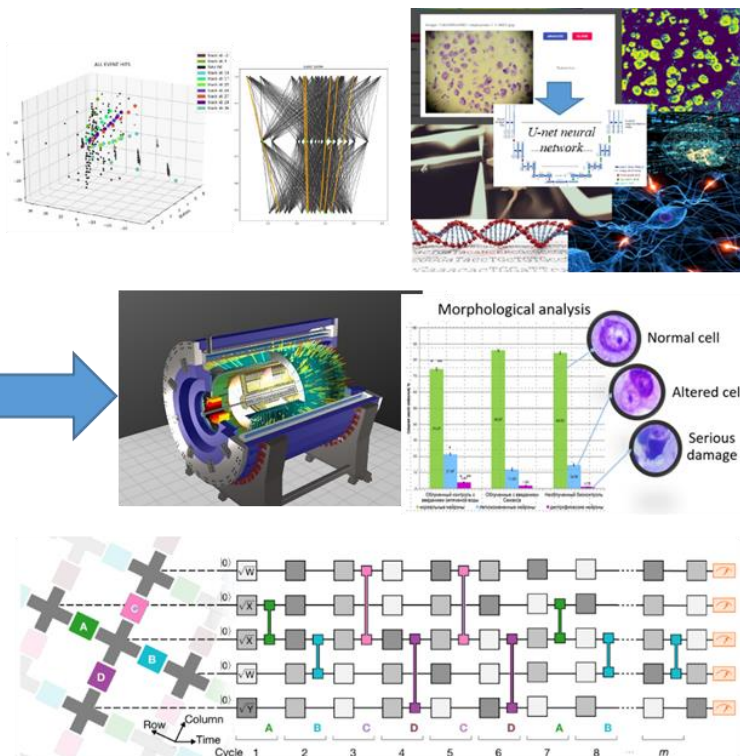
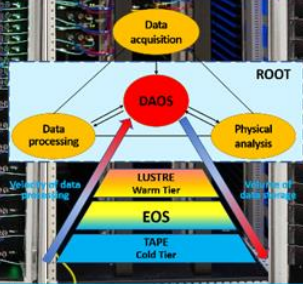


**Russian DC Awards 2020 in**  
**“The Best IT Solution for**  
**Data Centers”**

# Модернизация СК «Говорун» 2022 года



2022



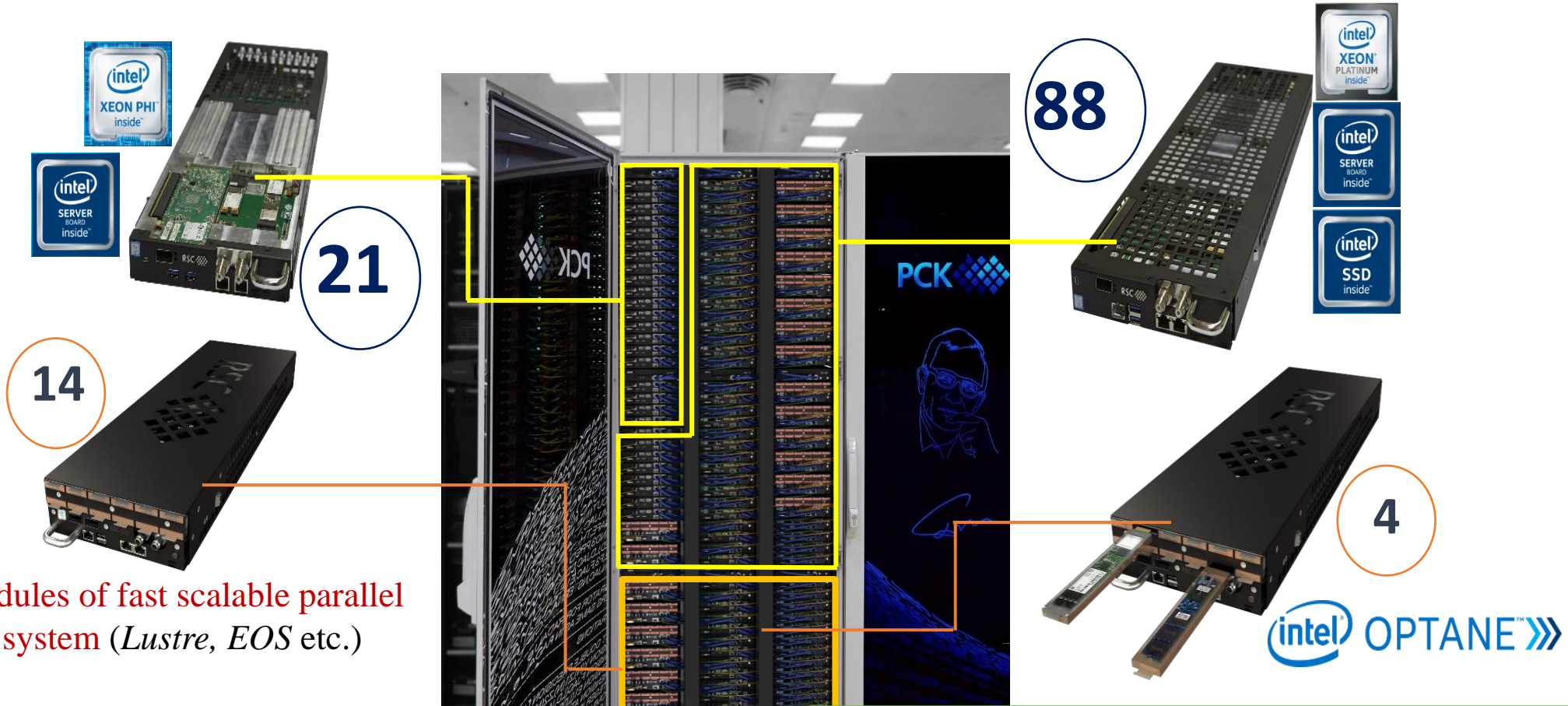
Непрерывный рост числа пользователей и расширение круга решаемых задач для обеспечения решения теоретических и экспериментальных задач ОИЯИ потребовал:

- Расширения вычислительных ресурсов СК «Говорун»;
- Внедрения новейших ИТ-решений;
- Создания среды для суперкомпьютерного моделирования и решения ресурсоемких задач с большими объемами вычислений и данных.

Расширение суперкомпьютера «Говорун» на 32 гиперконвергентных вычислительных узла и 8 узлами распределенной СХД позволит:

- повысить производительность на 239 ТФЛОПС (суммарно 1,1 ПФ);
- увеличить подсистему обработки и хранения данных DAOS до 1,6 Пбайт;
- расширить объем подсистемы хранения «теплых данных» на 8 Пбайт с поддержкой создания динамических систем хранения Lustre, DAOS, а также EOS, NFS.

# The CPU-component of the "Govorun" Supercomputer



Modules of fast scalable parallel file system (*Lustre, EOS* etc.)

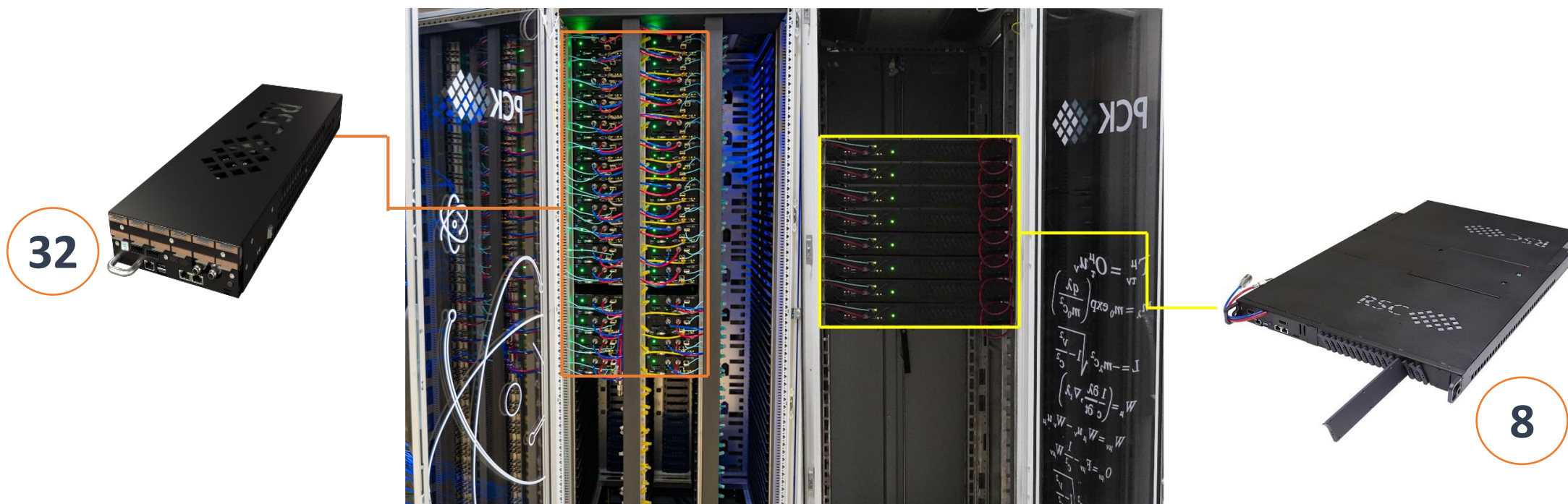
## RSC Tornado nodes based on Intel® Xeon Phi™:

- Intel® Xeon Phi™ 7290 processors (72 cores)
- Intel® Server Board S7200AP
- Intel® SSD DC S3520 (SATA, M.2)
- 96GB DDR4 2400 GHz RAM
- Intel® Omni-Path 100 Gb/s adapter

## RSC Tornado nodes based on Intel® Xeon® Scalable gen 2:

- Intel® Xeon® Platinum 8268 processors (24 cores)
- Intel® Server Board S2600BP
- Intel® SSD DC S4510 (SATA, M.2), 2x Intel® SSD DC P4511 (NVMe, M.2) 2TB
- 192GB DDR4 2933 GHz RAM
- Intel® Omni-Path 100 Gb/s adapter

# The CPU-component of the “Govorun” Supercomputer



## RSC Tornado nodes based on Intel® Xeon® Scalable gen 3:

- Intel® Xeon® Platinum 8368Q processors (38 cores)
- DDR-T PMEM 200 Series, 256 GB - 8 pc. total 2 048 GB
- Intel® SSD EDSFF E1.S 4TB – 4 pc. total 16 TB
- 256GB DDR4 2933 GHz RAM
- Intel® Omni-Path 200 Gb/s adapter

## RSC Tornado nodes based on Intel® Xeon® Scalable gen 2:

- Intel® Xeon® Gold 6248R processors (24 cores)
- Intel® Server Board S2600BP
- Intel® SSD NVMe EDSFF E1.L (RULER) 30,72 TB - 32 pc., total 983 TB
- 384 GB DDR4 2933 GHz RAM
- Intel® Omni-Path 200 Gb/s adapter

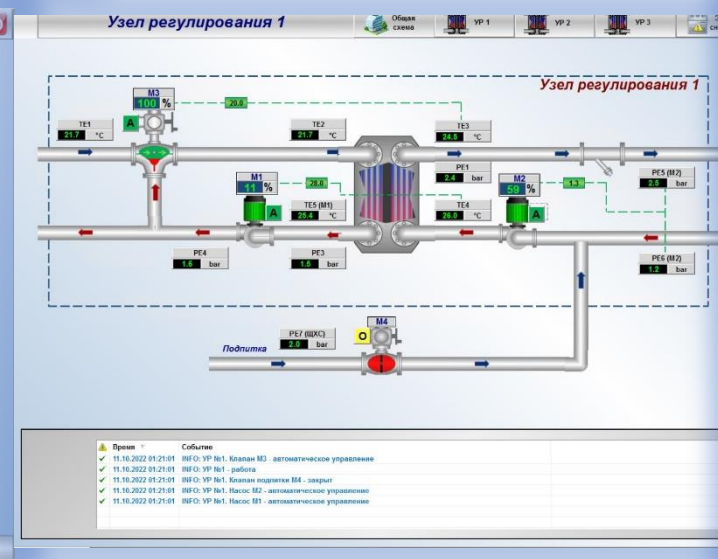
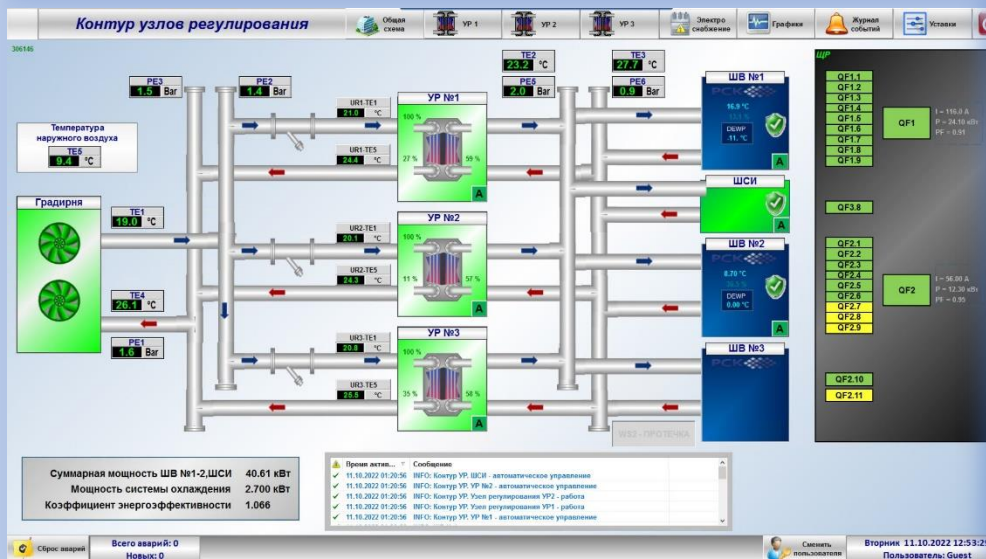
# Engineering infrastructure



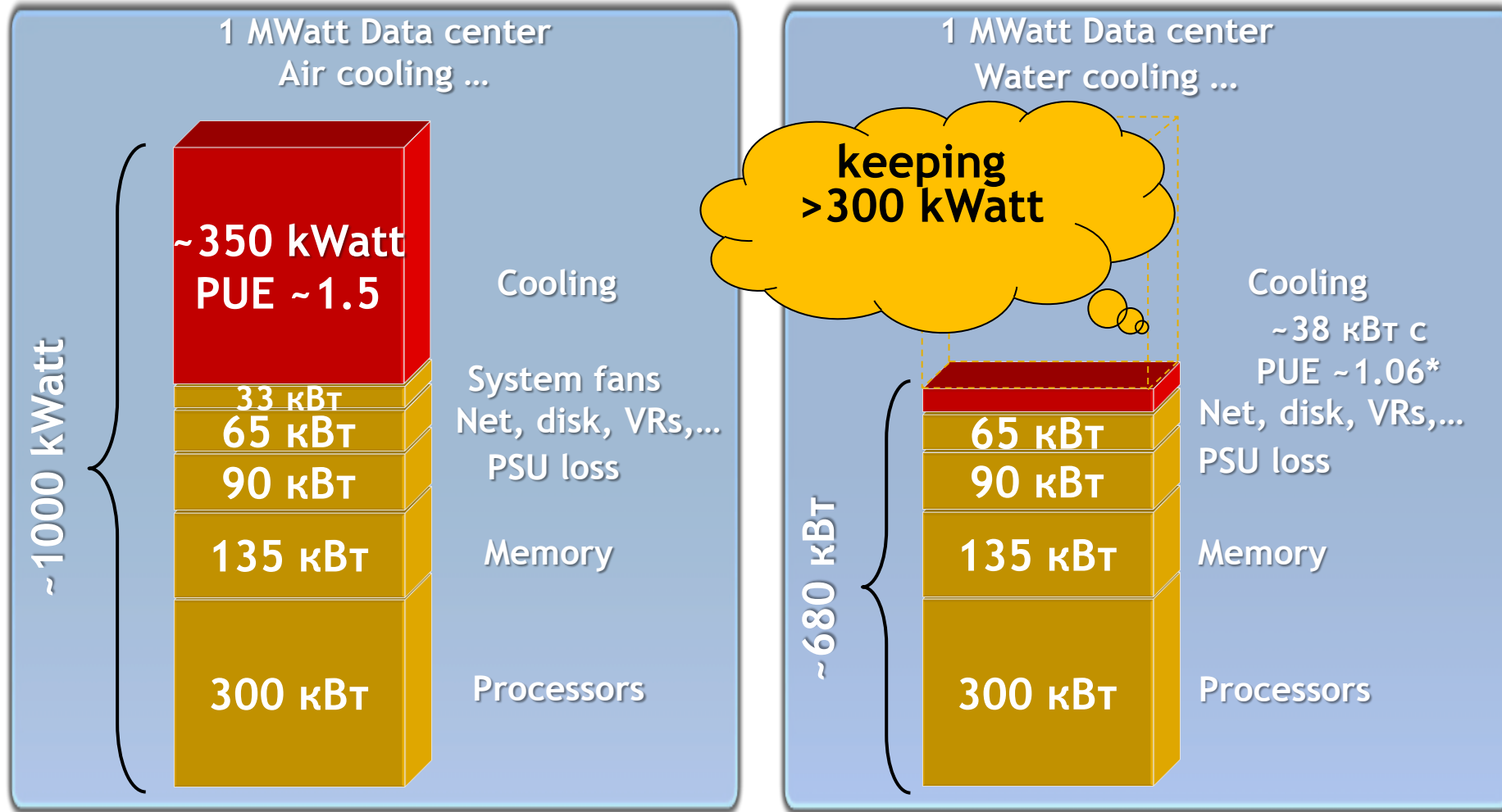
free cooling  
24x7x365



PUE ~ 1,06



# Cooling efficiency and cost





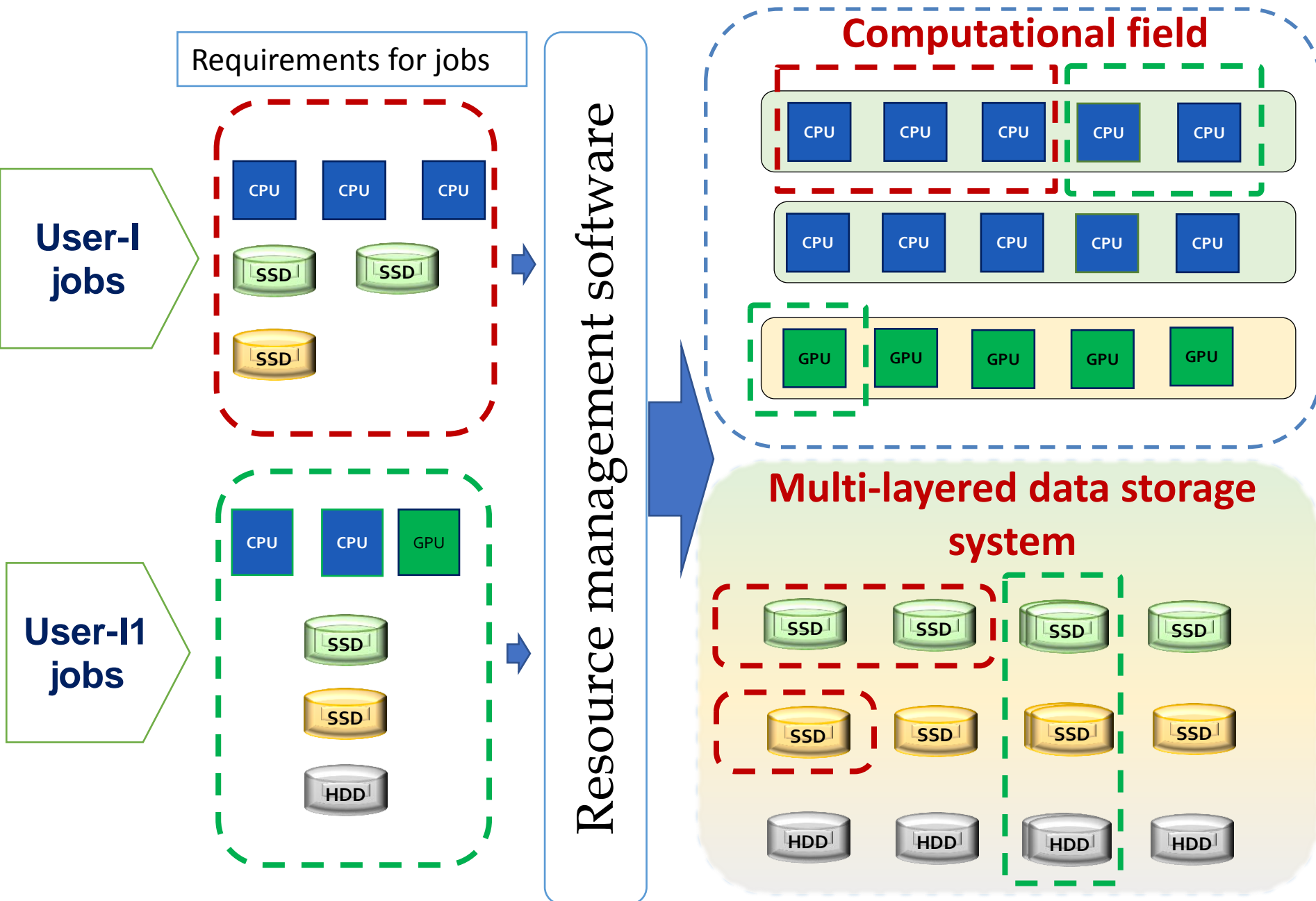
# The GPU-component of the “Govorun” Supercomputer



The **GPU-component** consists of **5 NVIDIA DGX-1 servers**. Each server has **8 GPU NVIDIA Tesla V100** based on the latest architecture NVIDIA Volta. Moreover, one server NVIDIA DGX-1 has **40960** cores CUDA, which are equivalent to 800 high-performance central processors. A whole number of novel technologies are used in DGX-1, including the NVLink 2.0 wire with the bandwidth up to 300 Gb/s.

The GPU-component gives a users of the supercomputer a possibility to allow as massively parallel computation for general-purpose tasks using such technologies as CUDA and OpenCL, as well as use applications already adapted for this architecture. Also, GPU-component allow to use machine learning and deep learning algorithms for solving applied problems by neural network approach.



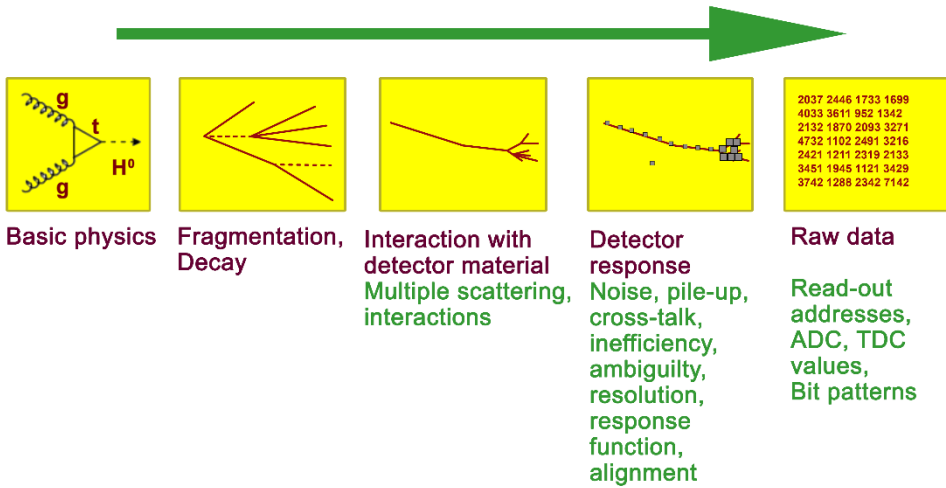


The "Govorun" supercomputer has unique properties for the flexibility of customizing the user's job.

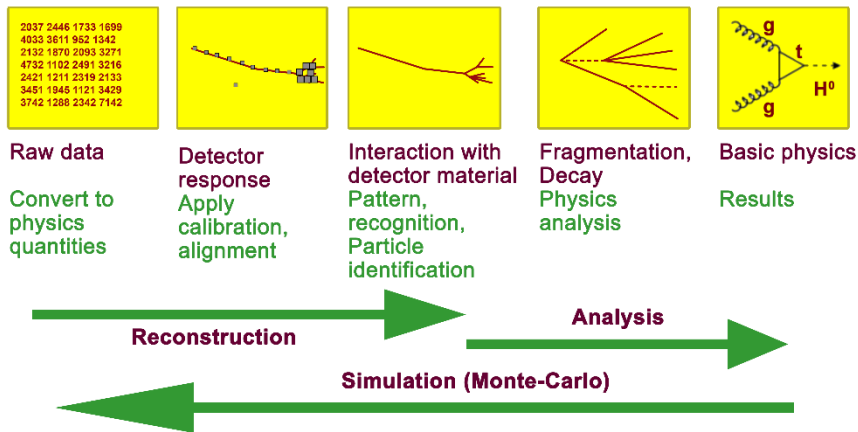
For his job the user can allocate the required number and type of computing nodes and the required volume and type of data storage systems.

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

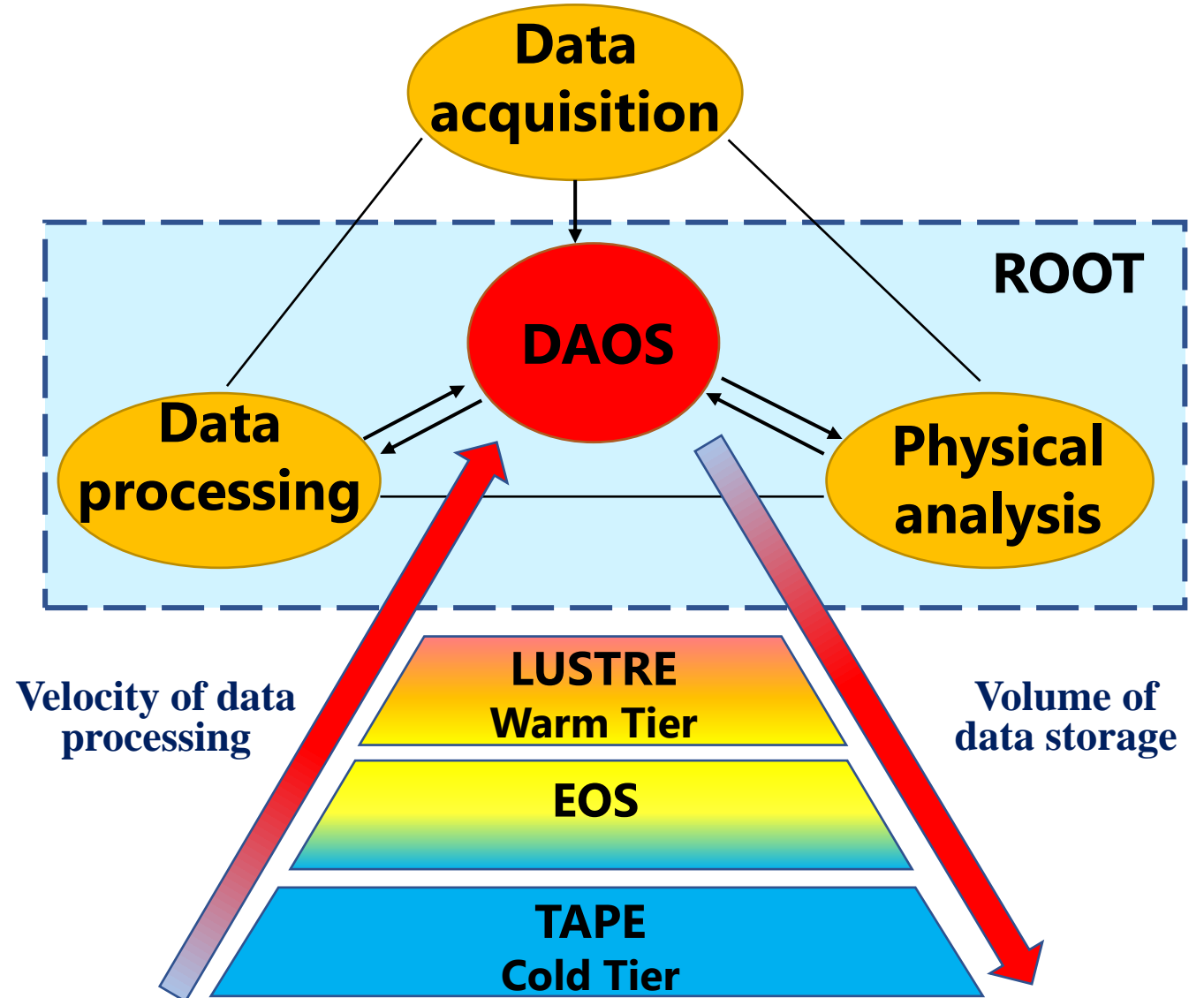
## From Physics to raw data



## From raw data to Physics



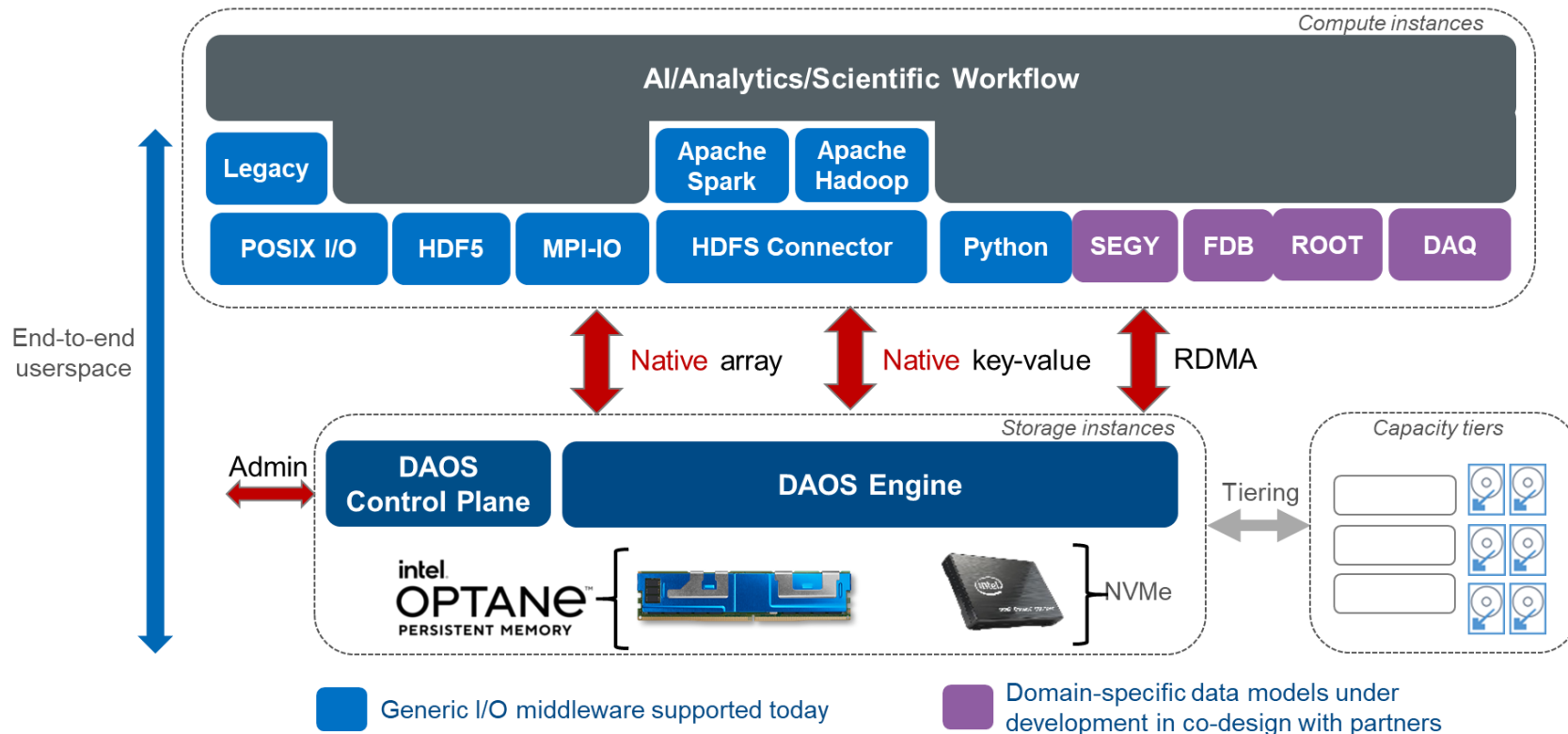
We need to go from raw data back to physics reconstruction + analysis of the event(s)



# DAOS: Promising technology for HPC, Big Data, AI



## DAOS (Distributed Asynchronous Object Storage) Software Ecosystem



- Complex approach to build a hierarchical storage system
- DAOS is significant part of data acquisition and processing
- Different types of containers are used for different data processing stages
- No need of POSIX file system for most data operations
- Great system performance even for a few DAOS clients
- RSC Storage on-Demand software offers unique flexibility, speed, and convenience for DAOS users

The DAOS polygon on the supercomputer "Govorun" take the **1<sup>st</sup>** place among Russian supercomputers in the current **IO500** list

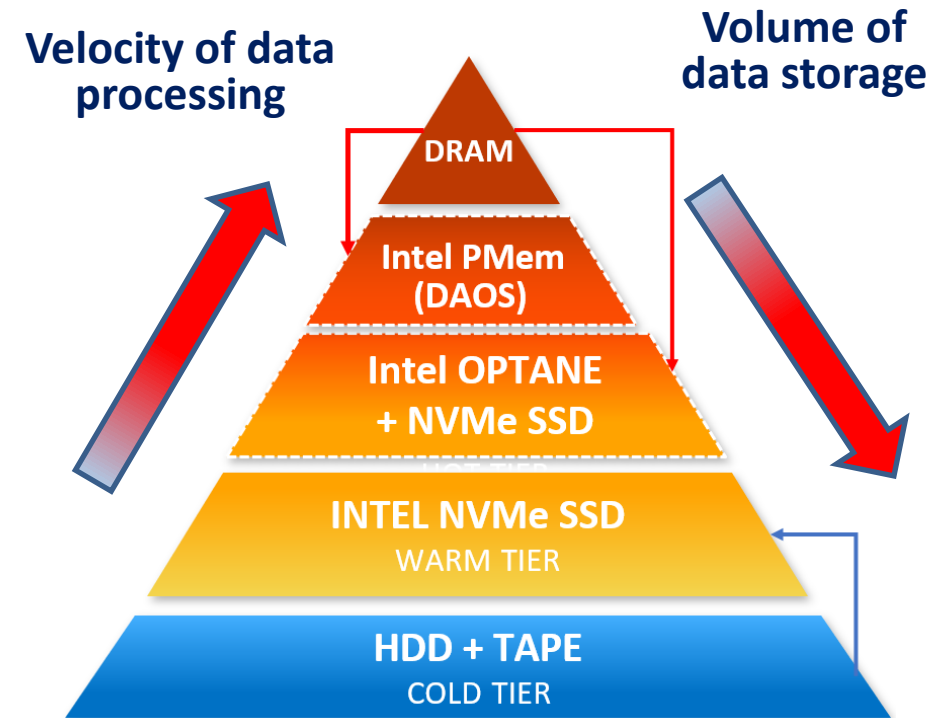
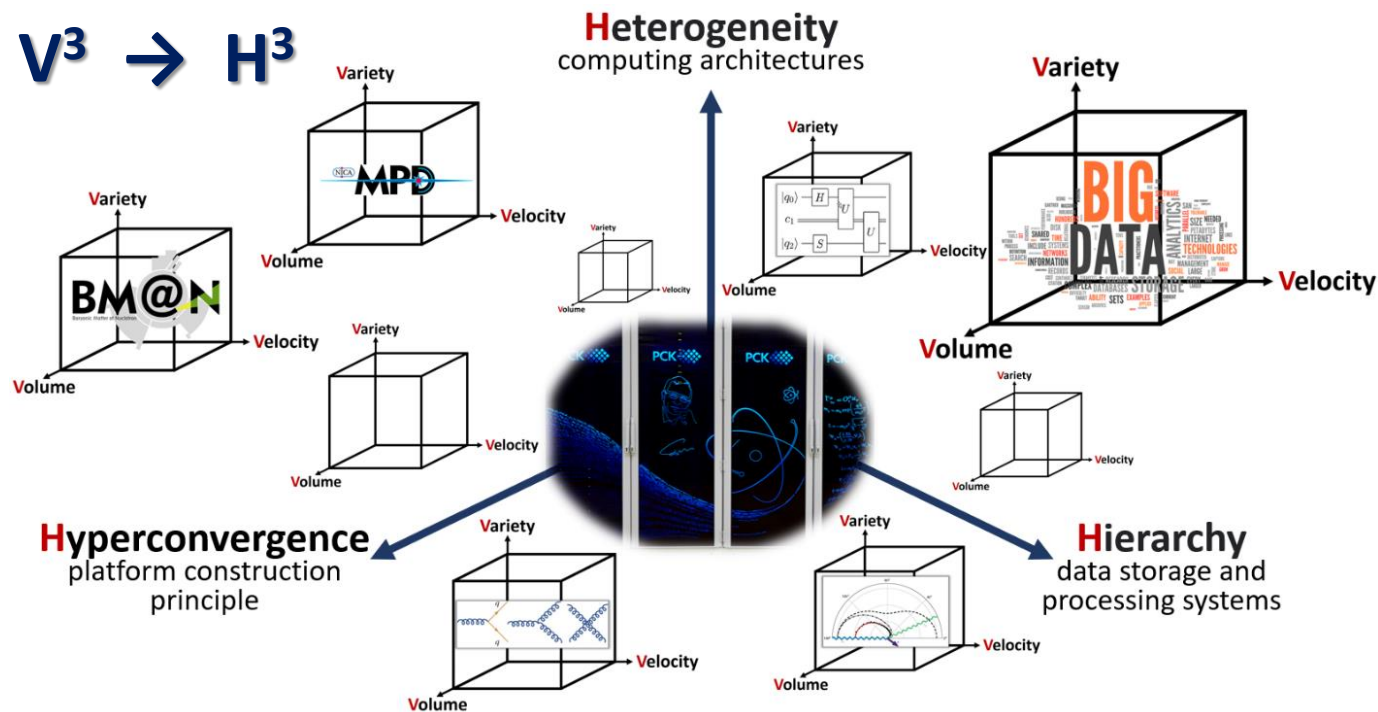
# Big Data on the "Govorun" Supercomputer for NICA megaproject



The DAOS polygon of the "Govorun" supercomputer takes the **1<sup>st</sup>** place among Russian supercomputers in terms of the data processing rate in the current **IO500 list**.

**Heterogeneity**  
**Hierarchy**  
**Hyperconvergence** → provide → **Variety**  
**Velocity**  
**Volume**

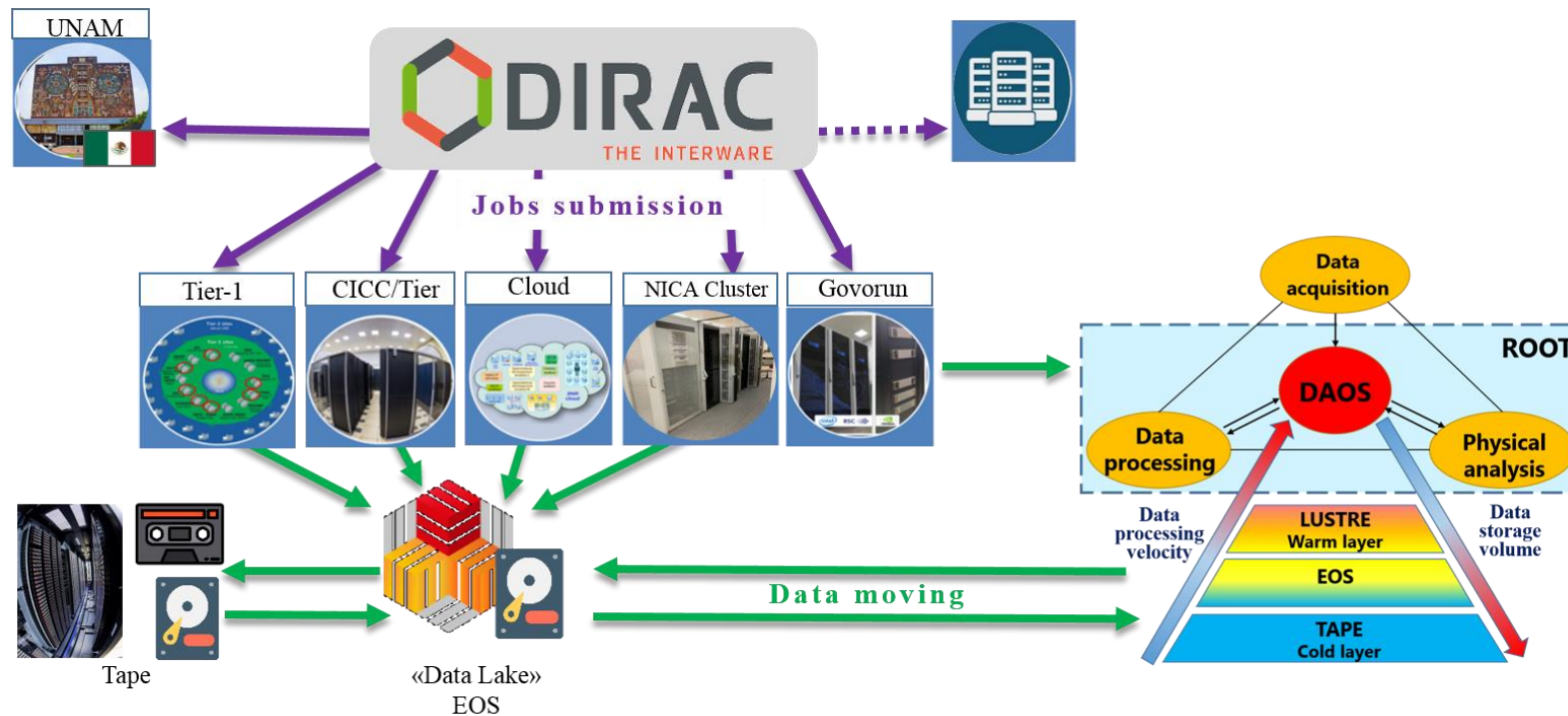
$V^3 \rightarrow H^3$



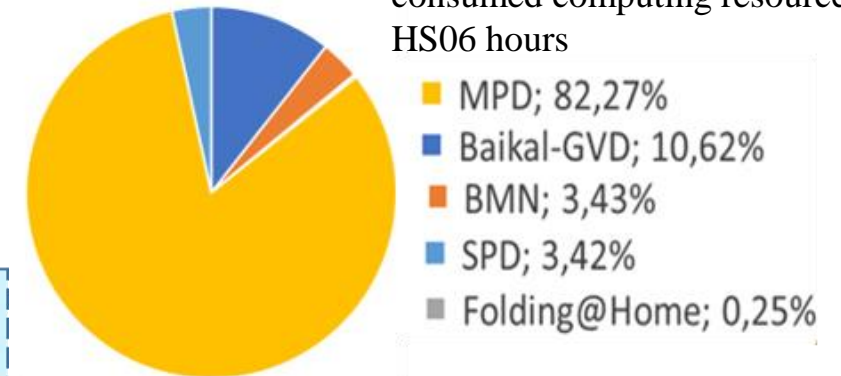
# Heterogeneous distributed computing environment



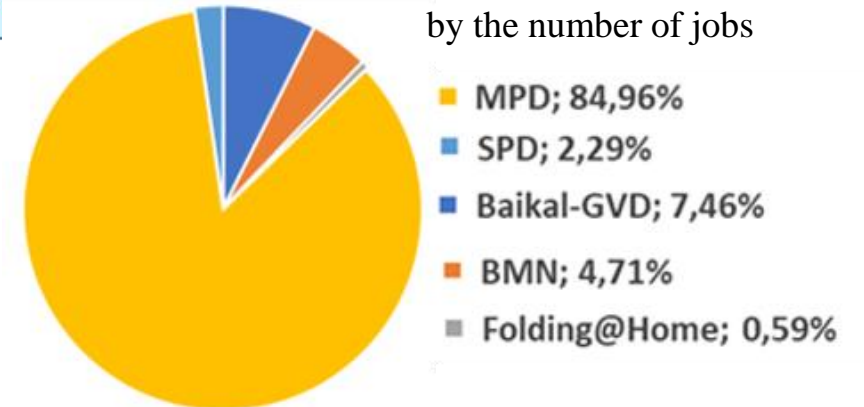
A heterogeneous computing environment, based on the DIRAC platform, was created. for processing and storing data of the experiments conducted at JINR. By the end of 2021, all the MICC components, the clouds of the JINR Member States, the NICA cluster, as well as the cluster of the National Autonomous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



Usage of resources integrated into DIRAC by main users:  
consumed computing resources in HS06 hours

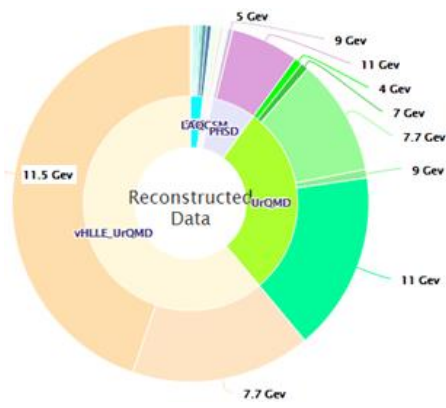
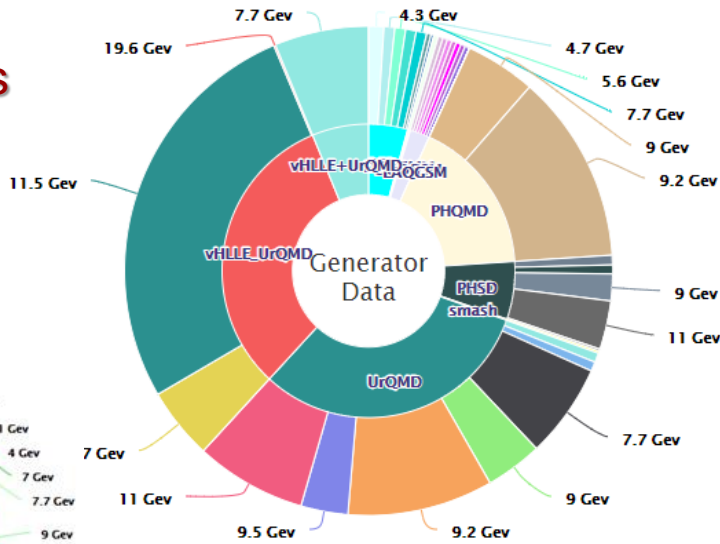


by the number of jobs



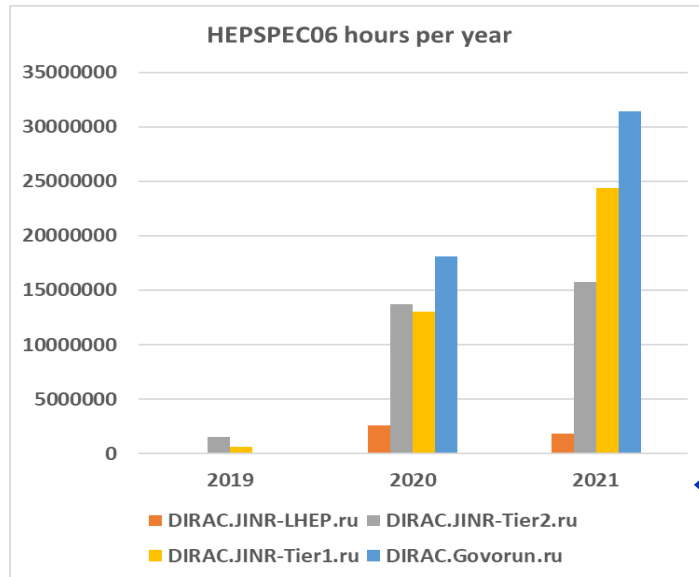
# Heterogeneous distributed computing environment for the MPD experiment

✓  $1200 * 10^6$  events were generated using *UrQMD*, *PHQMD*, *PHSD* and other models



✓  $392 * 10^6$  events were reconstructed

“Govorun” up to 40%

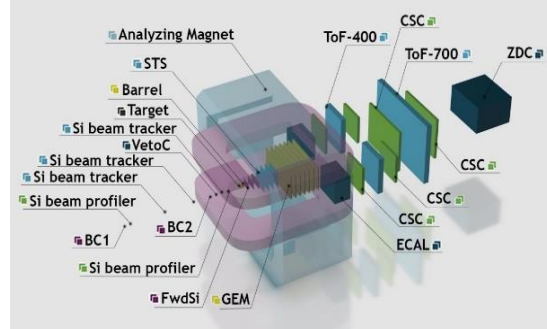


Available resources of the DIRAC platform for the MPD experiment:

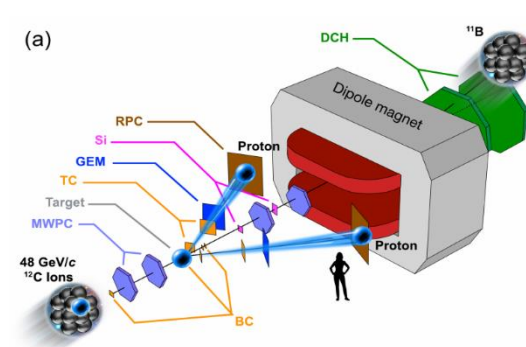
- “Govorun” supercomputer: up to 1,586 cores in the latest production
- Tier1: 920 cores
- Tier2: 1,000 cores
- Clouds (JINR and JINR Member States): 70 cores
- NICA offline cluster: 300 cores (limit for users)
- UNAM (Mexico University): 100 cores
- National Research Computer Network of Russia (NIKS, now resources from SPBTU and JSCC): 672 cores – New resource, added in 12.2021.

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

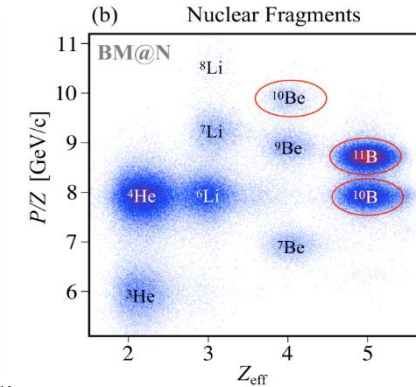
The histogram illustrates the accounting data from the centers. The metric shown is Sum CPU Work, grouped by center and year.



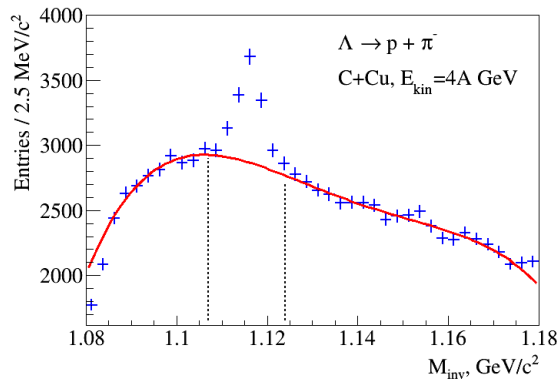
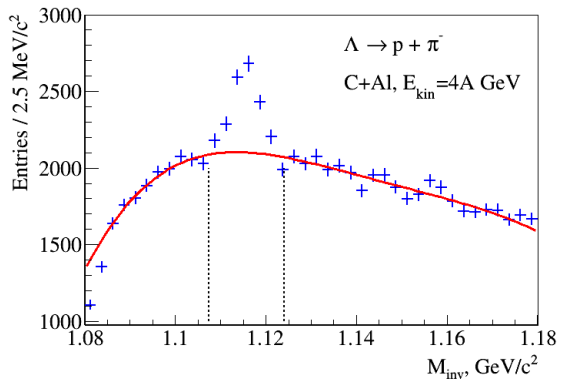
Full BM@N configuration for heavy ion studies in 2018.



(a) BM@N configuration for SRC studies.



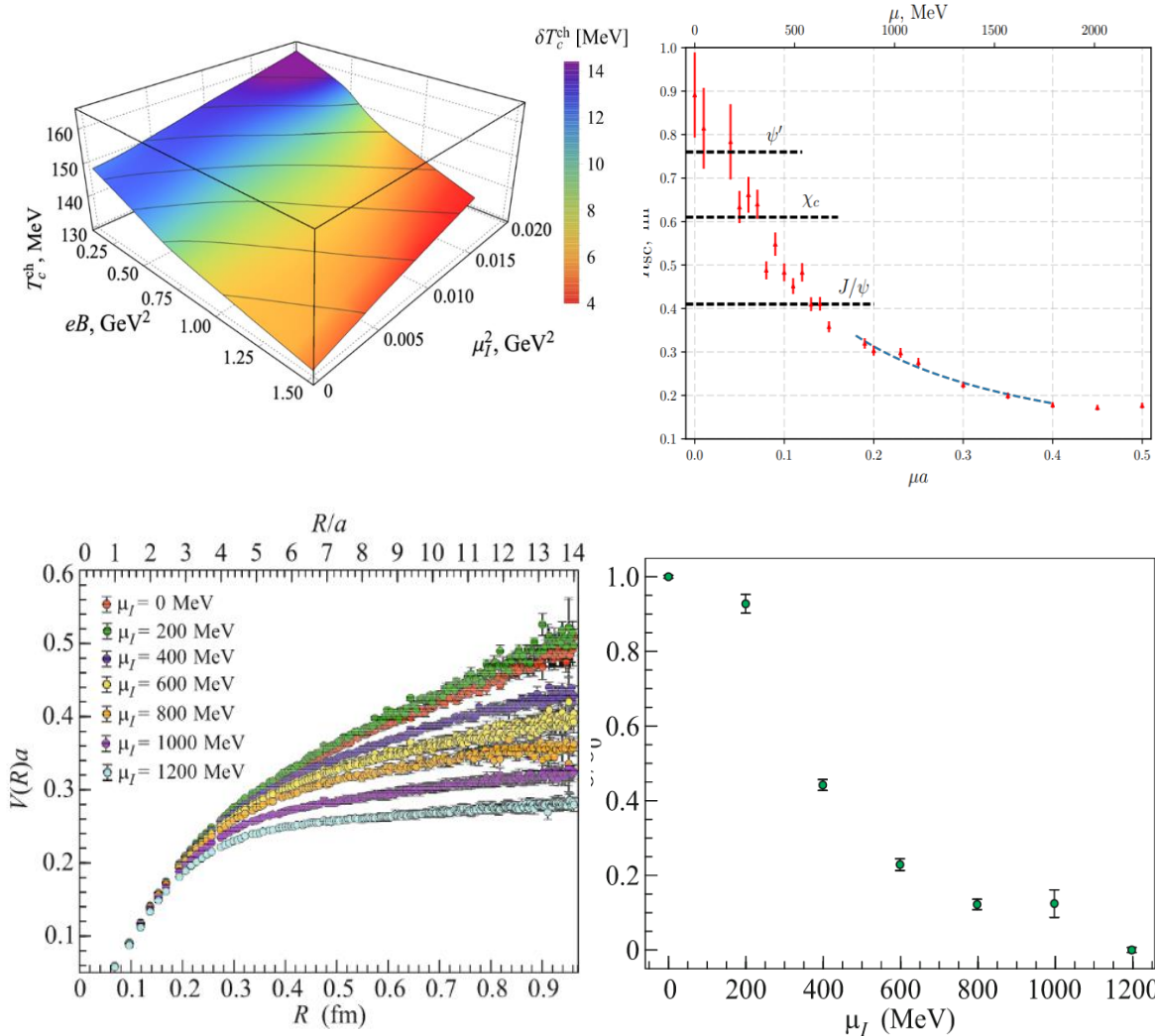
(b) isolation of nuclear fragments in the experiment under the SRC program.



Signals of  $\Lambda$ -hyperons in the spectra of invariant masses ( $p, \pi^-$ ) measured in C+Al and C+Cu interactions.

- The analysis of experimental data acquired during the Nuclotron runs in 2016-2018 was performed. Special attention is paid to the study of interactions between beams of carbon and argon ions with fixed targets of different types. The reconstruction of particle tracks was carried out using the method of “cellular automata”.
- The modeling of the work of the experiment using generators of physical models, such as DCM-QGSM and URQMD, and the embedding procedure were performed.
- The staff of the BM@N collaboration from Russia, the USA, Israel, Germany, France and JINR, working on the program for the study of short-range correlations (SRC) of nucleons in nuclei, developed and applied a new experimental method for investigating the internal structure of the atomic nucleus in carbon-hydrogen interactions. [A publication based on the results of the SRC program of the BM@N experiment was sent to the scientific journal Nature.](#)
- The polarization of  $\Lambda$ -hyperons was studied using the model data of the DCM-QGSM generator of the BM@N experiment.





The resources of the “Govorun” supercomputer were used to study the properties of quantum chromodynamics (QCD) and Dirac semimetals in a tight-binding mode under extreme external conditions using lattice modeling. The given study entails the inversion of large matrices, which is performed on video cards (GPU), as well as massive parallel CPU calculations, to implement the quantum Monte-Carlo method:

- The influence of the magnetic field on the confinement/deconfinement transition and the chiral transition at finite temperature and zero baryon density were investigated using the numerical modeling of lattice QCD with a physical quark mass.
- Quantum chromodynamics with non-zero isospin density taking into account dynamical u- d-, s- quarks in the Kogut-Susskind formulation was studied.
- The potential of the interaction between a static quark-antiquark pair in dense two-color QCD was investigated, and the confinement/deconfinement phenomenon was studied.
- The effect of the non-zero chiral chemical potential on dynamical chiral symmetry breaking for Dirac semimetals was studied.
- The influence of the external magnetic field on the electromagnetic conductivity of quark-gluon plasma was investigated.

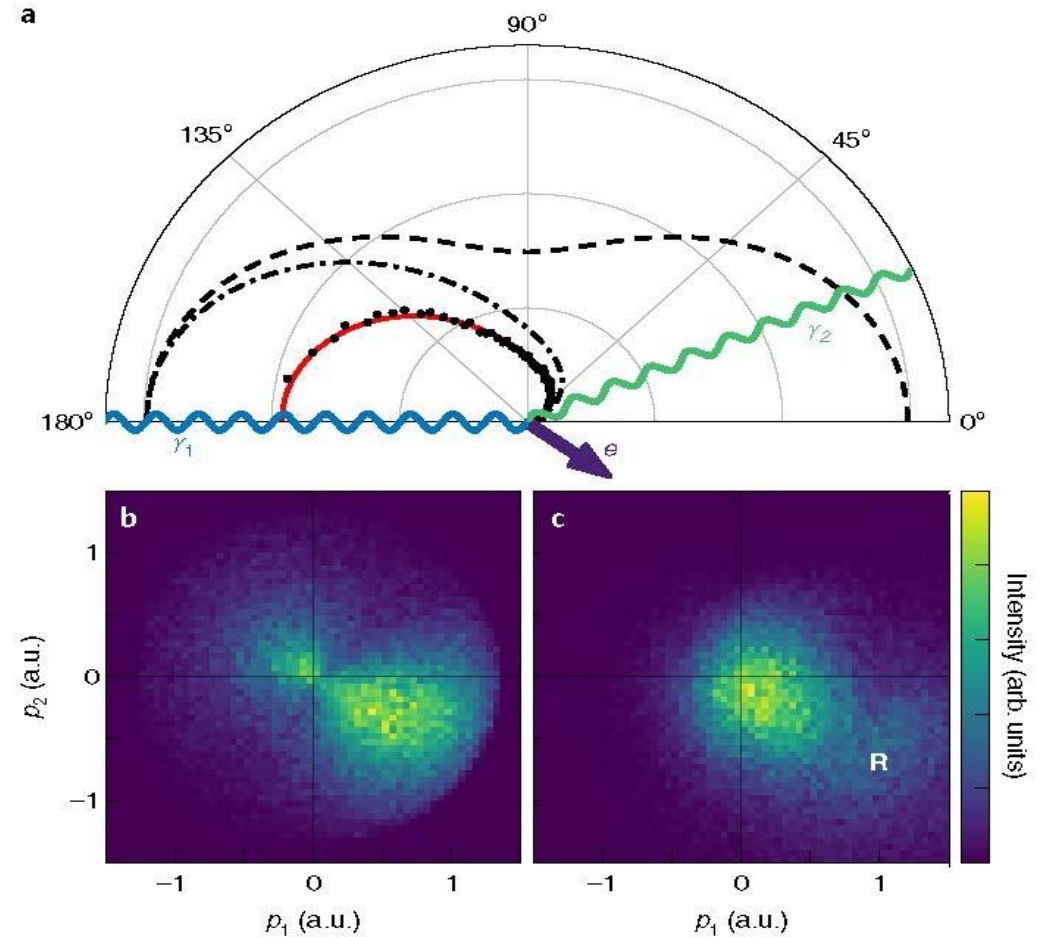
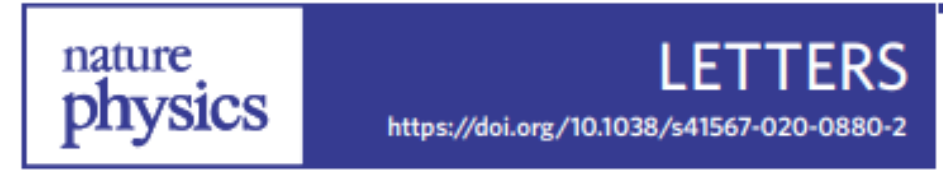
### The results are published in the articles:

1. V. V. Braguta, M. N. Chernodub, A. Yu. Kotov, A. V. Molochkov, and A. A. Nikolaev, Phys. Rev. D 100 (2019), 114503, DOI: 10.1103/PhysRevD.100.114503, arXiv:1909.09547
2. V.V. Braguta , A.Yu. Kotov, A.A. Nikolaev, JETP Lett. 110 (2019) no.1, 1-4, DOI: 10.1134/S0021364019130083 (JETP Letters, 110 (2019) no.1, 3-6)
3. N. Astrakhantsev, V. Bornyakov, V. Braguta, E.M. Ilgenfritz, A.Y. Kotov, A. Nikolaev, A. Rothkopf, PoS Confinement2018 (2019), 154, DOI: 10.22323/1.336.0154
4. V. V. Braguta, M. I. Katsnelson, A. Yu. Kotov, and A. M. Trunin, Phys.Rev. B100 (2019), 085117, DOI: 10.1103/PhysRevB.100.085117 , e-Print: arXiv:1904.07003
5. N. Yu. Astrakhantsev, V. G. Bornyakov, V. V. Braguta, E.-M. Ilgenfritz, A. Yu. Kotov, A. A. Nikolaev, A. Rothkopf, JHEP 1905 (2019) 171, DOI: 10.1007/JHEP05(2019)171,e-Print: arXiv:1808.06466
6. <https://arxiv.org/abs/1902.09325>
7. <http://arxiv.org/abs/1910.08516>

# Kinematically complete experimental study of Compton scattering at helium atoms near the threshold

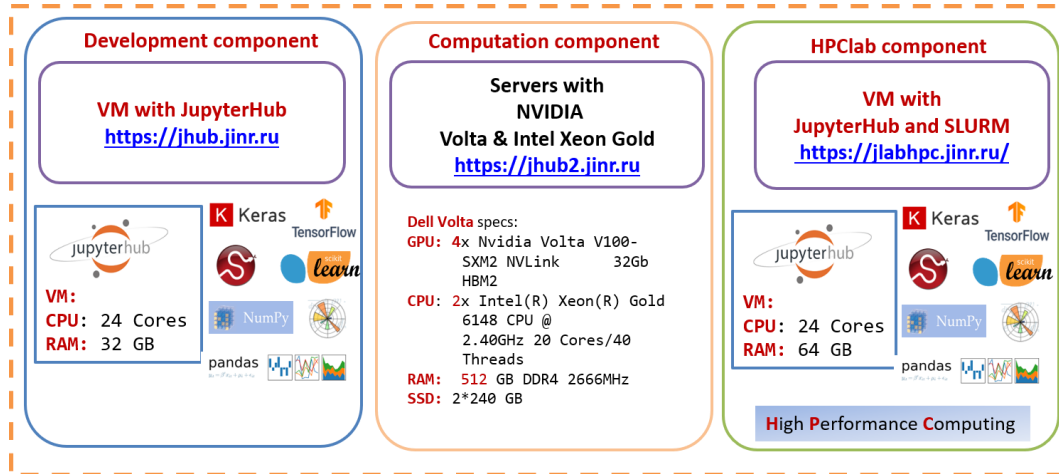


On 13 April, a scientific paper by an international scientific group was published in **Nature Physics**. A LIT staff member *O. Chuluunbaatar* and a BLTP employee *Yu. V. Popov* were members of the group in the frames of JINR international cooperation. The group conducted a kinematically complete experimental measurement of characteristics of Compton scattering at free atoms using the highly efficient method of COLD Target Recoil Ion Momentum Spectroscopy (COLTRIMS). The group also provided a **relevant theoretical description** of it which was carried out at the supercomputer “Govorun”.



Scheme of ionization by Compton scattering at  $h\nu=2.1$  keV

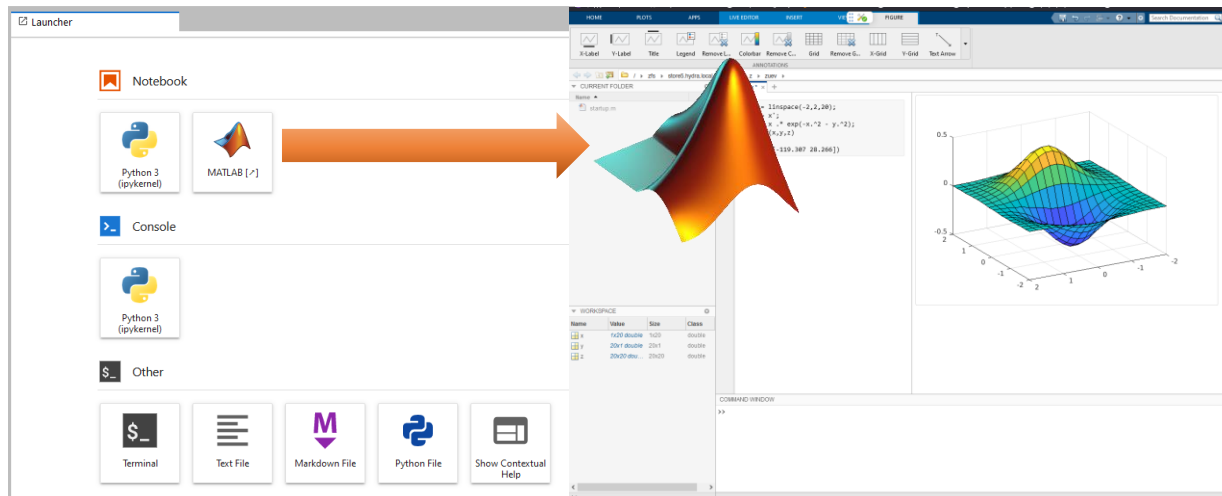
# ML/DL/HPC Ecosystem of the HybriLIT Heterogeneous Platform: New Opportunities for Applied Research



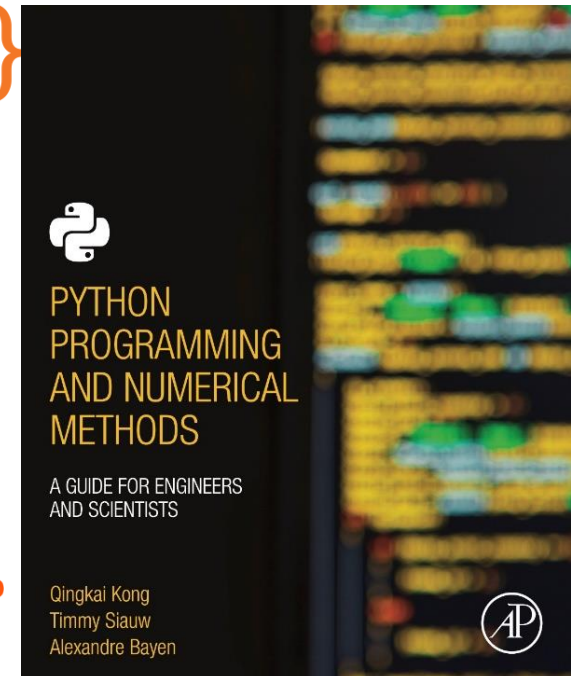
The ML/DL/HPC ecosystem is now actively used for machine and deep learning tasks. At the same time, the accumulated tools and libraries can be more widely used for scientific research, including:

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

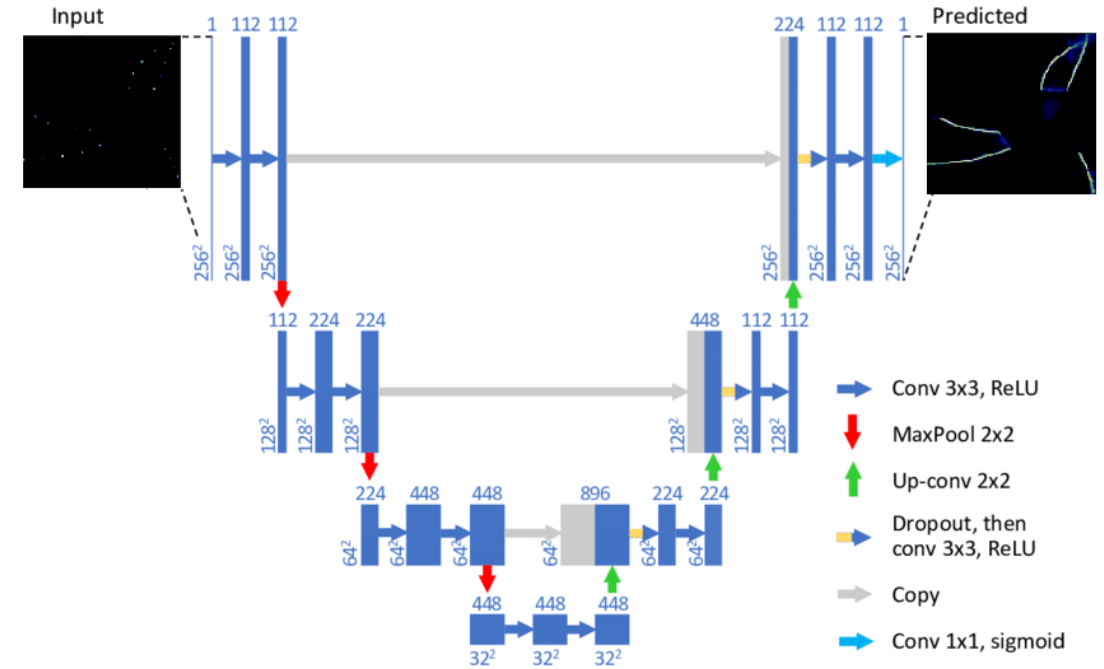
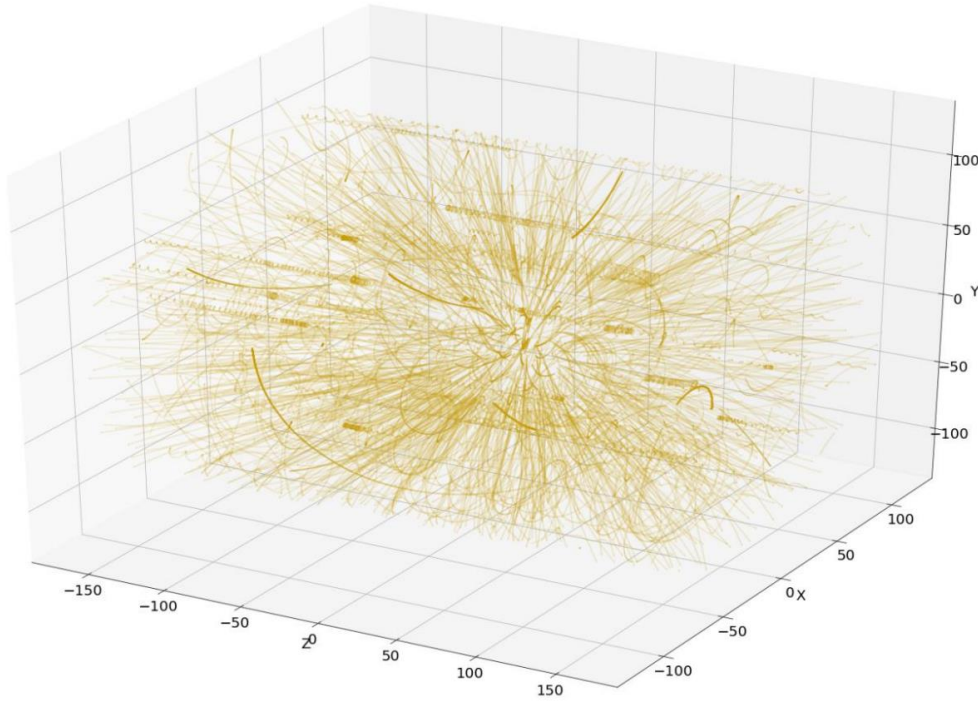
In 2022, on the ML/DL/HPC ecosystem, it became possible to run the MATLAB code in Jupyter Notebook, which allows one to effectively perform applied and scientific computations.



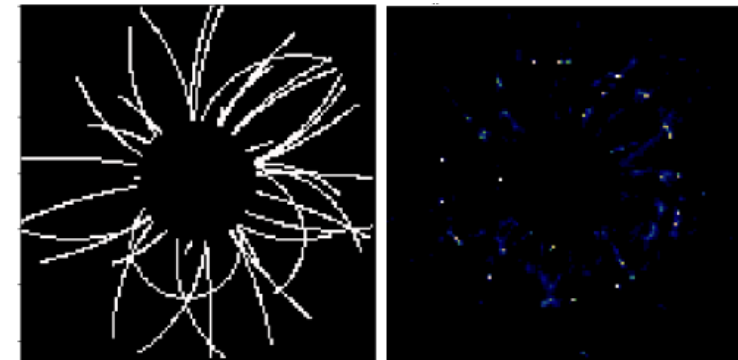
## Python Numerical Methods



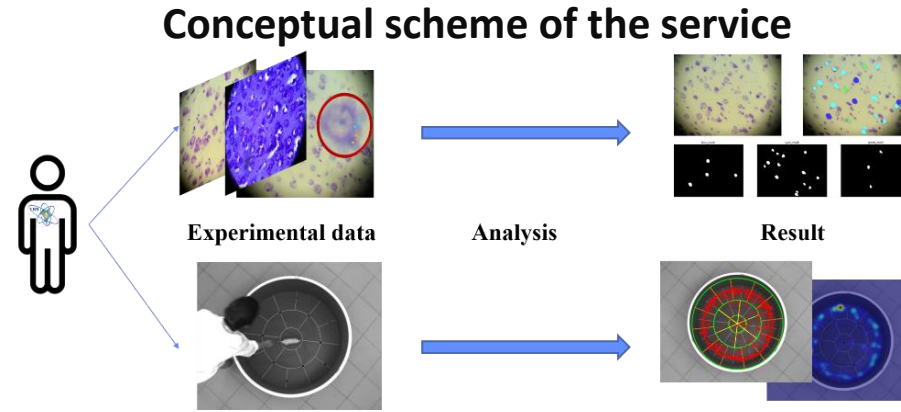
A large number of tracks in events requires the development of approaches that have constant computational complexity regardless of the number of tracks in an event. The use of deep neural network architectures allows developing tracking one-pass algorithms that work in just single step.



Model experiments show that neural network models are capable of both interpolating tracks and creating an internal model to represent the results in the phase space of the track parameters.

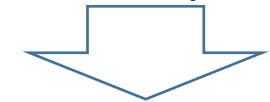


The information system allows one to store, quickly access and process data using a stack of neural network and classical algorithms of computer vision, providing a wide range of possibilities for automating routine tasks. It gives an increase in productivity, quality and speed of obtaining results.



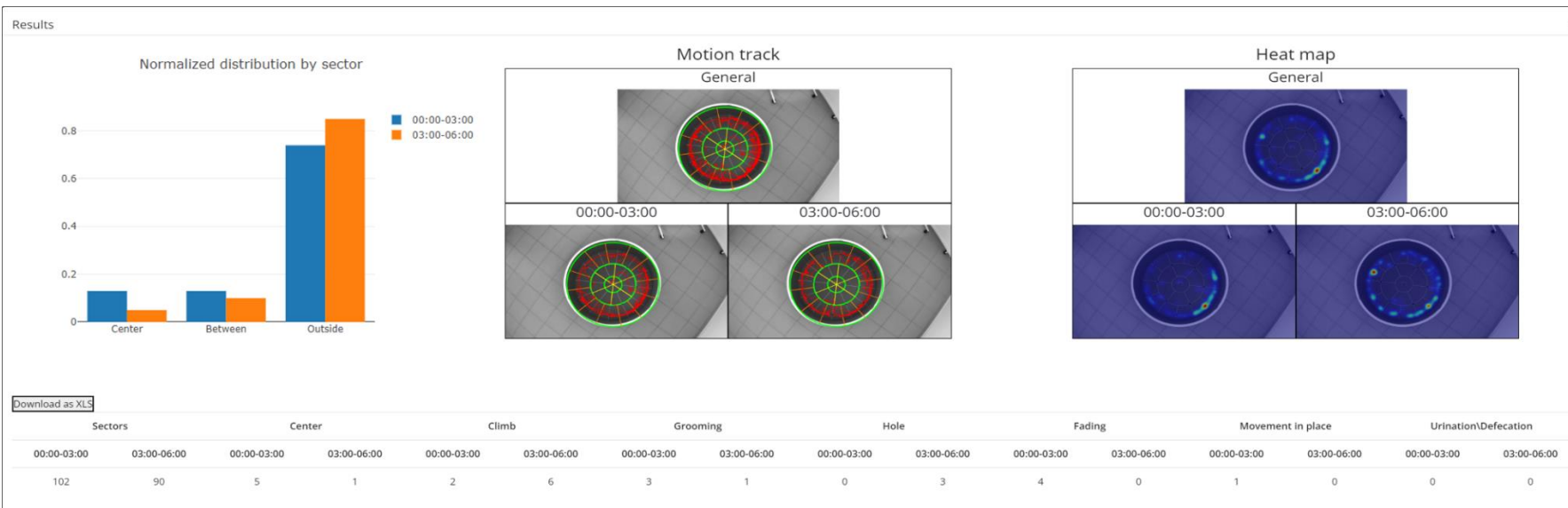
## Developed algorithms:

- algorithms for the automated marking of the field of experimental setups,
- algorithms for tracking the animal's position in experimental setups of different types,
- algorithms for evaluating the animal's behavioral patterns.



The obtained information is stored in different forms:

- visualized track of the animal's movement,
- video file with tracking the animal's position,
- heat map by sectors,
- file that stores all the information for subsequent statistical analysis.



# Study the dynamics of magnetization in a Phi-0 Josephson Junction



The dynamics of the magnetic moment  $M$  of the system under consideration is described by the Landau-Lifshitz-Gilbert equation:

$$\frac{dm_x}{dt} = -\frac{1}{1 + M^2 \alpha^2} \{m_y H_z - m_z H_y + \alpha [m_x (M, H) - H_x]\},$$

$$\frac{dm_y}{dt} = -\frac{1}{1 + M^2 \alpha^2} \{m_z H_x - m_x H_z + \alpha [m_y (M, H) - H_y]\}$$

$$\frac{dm_z}{dt} = -\frac{1}{1 + M^2 \alpha^2} \{m_x H_y - m_y H_x + \alpha [m_z (M, H) - H_z]\},$$

$M = [m_x, m_y, m_z]$  are the magnetic moment components; the effective field components  $H = [H_x, H_y, H_z]$  depend on the Josephson phase difference  $\phi$  and are defined as follows:

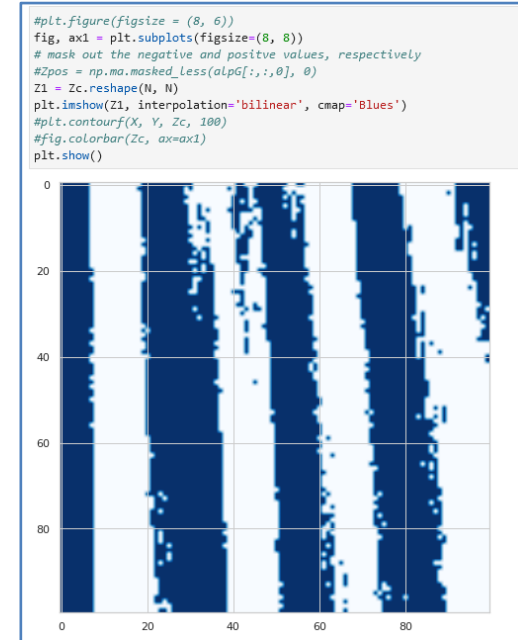
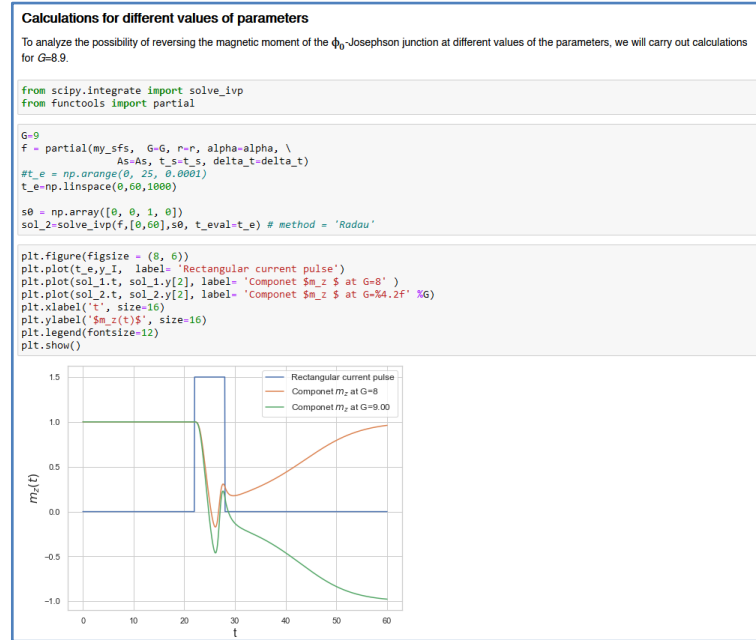
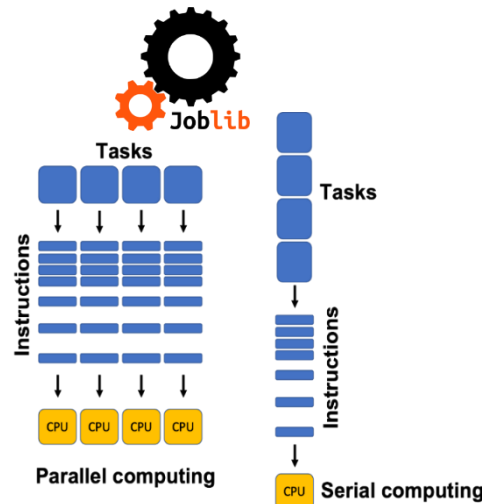
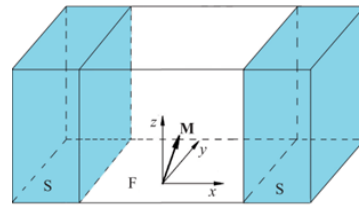
$$H_x(t) = 0,$$

$$H_y = Gr \sin(\phi(t) - tm_y(t)),$$

$$H_z(t) = m_z(t).$$

The equation for the Josephson phase difference  $\phi(t)$  is determined from the equation for the electric current  $I$  flowing through the Josephson junction, measured in units of the critical current  $I_c$ :

$$\frac{d\phi}{dt} = -\frac{1}{w} \left( \sin(\phi - rm_y) + r \frac{dm_y}{dt} \right) + \frac{1}{w} I,$$



**Define a function called by each process**

```

from joblib import Parallel, delayed
import numpy as np

def funk_parallel(k):
    i=k//N
    j=k//N
    m_z_sol=0
    G=G0*delta_G*1
    alpha=alpha0*delta_alpha*j
    f = partial(my_sfs, G=G, r=r, alpha=alpha, \
                As=As, t_s=t_s, delta_t=delta_t)
    t_e = np.linspace(0, 60, 1000)
    s0 = np.array([0, 0, 1, 0])
    sol = solve_ivp(f, [0, 60], s0, t_eval=t_e) # method = 'Radau'
    if sol.y[2][999] < 0:
        m_z_sol = -1
    # alpGxy[i+j*N,2] += -1
    return m_z_sol
    
```

**Serial mode calculation**

```

t0 = time.time()
rez = Parallel(n_jobs=1)\
      (delayed(funk_parallel)(k) for k in range(N*N))
t1 = time.time()
print(f'Execution time (t1 - t0) s')
    
```

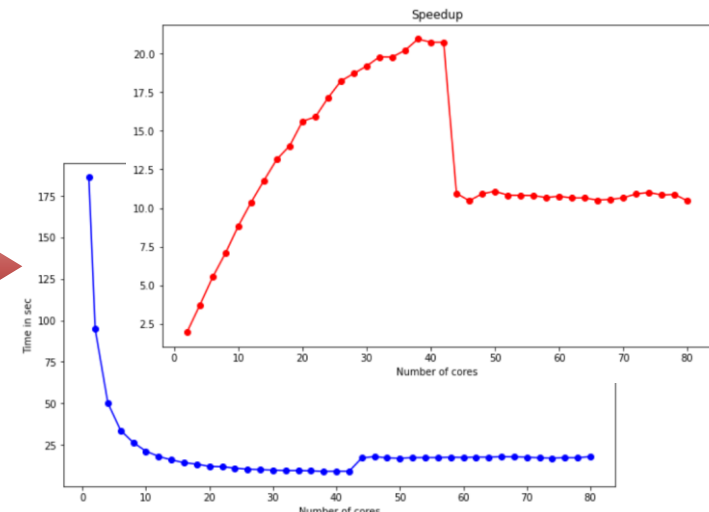
Execution time 159.9254457950592 s

**Computing in Parallel Mode**

```

t0 = time.time()
rez = Parallel(n_jobs=6)\
      (delayed(funk_parallel)(k) for k in range(N*N))
t1 = time.time()
print(f'Execution time (t1 - t0) s')
    
```

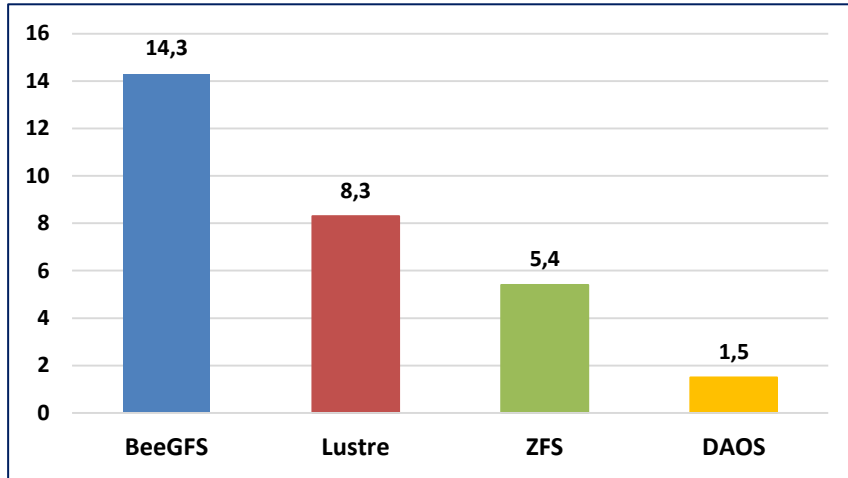
Execution time 34.51503801345825 s



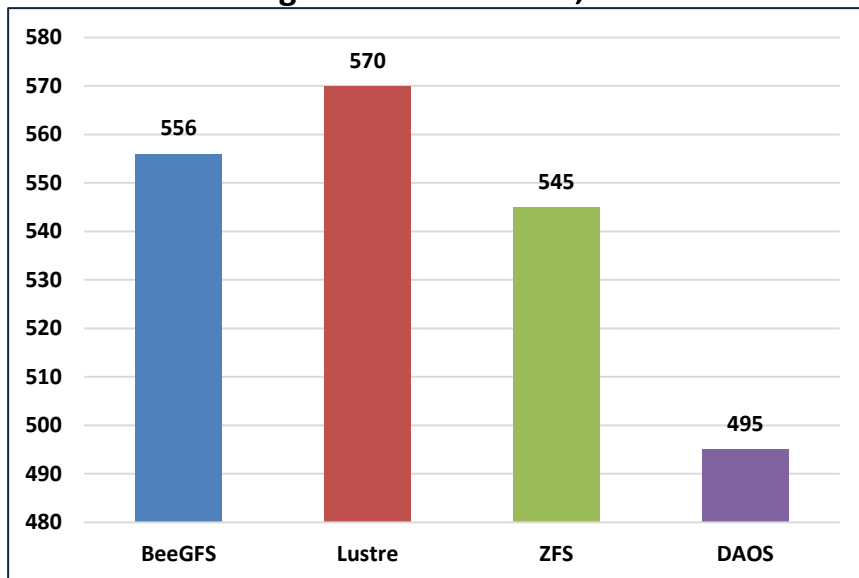
# DAOS Testing. Results



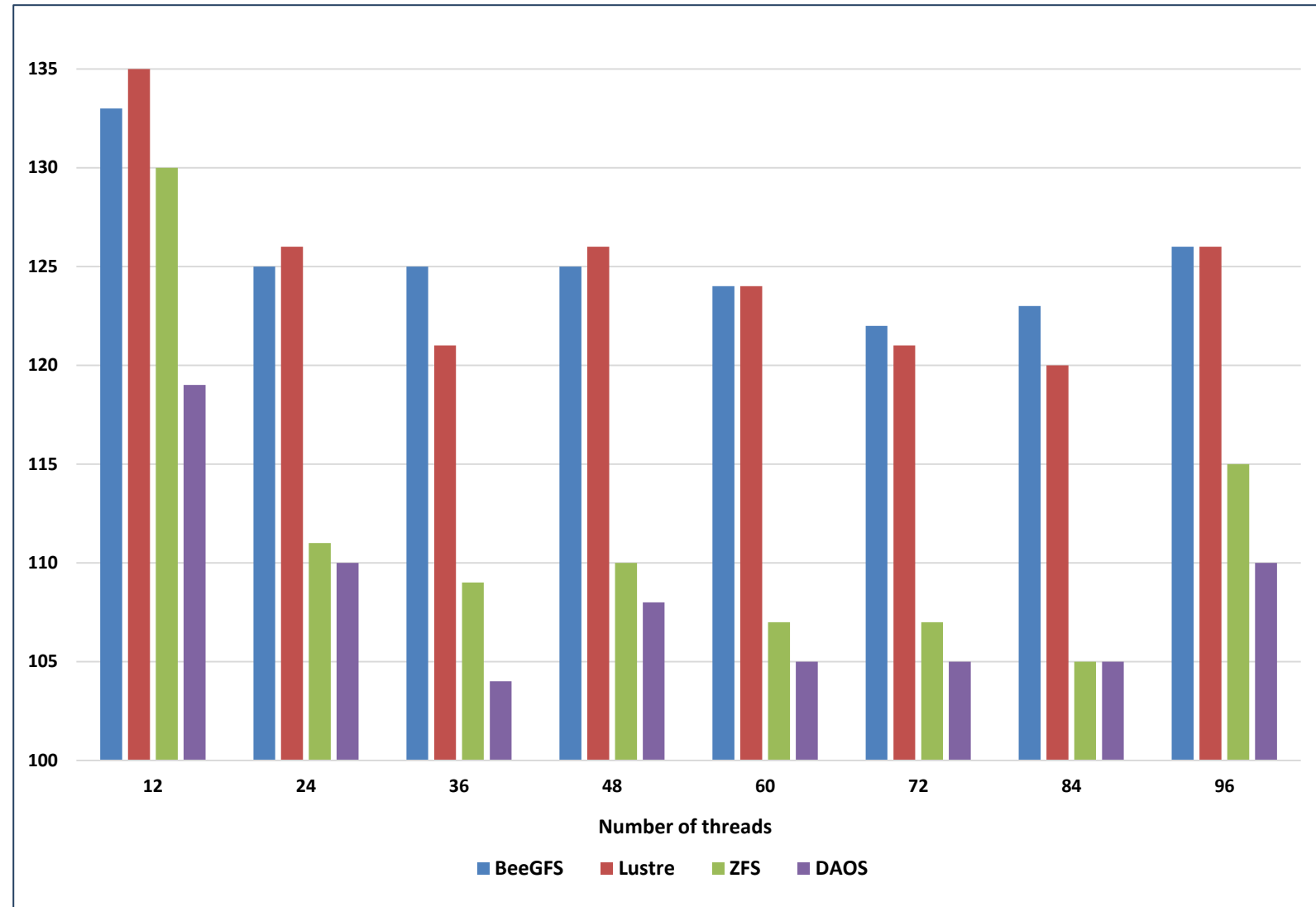
### Time to load data into the model, sec



### Average training time of one epoch in single-threaded mode, sec



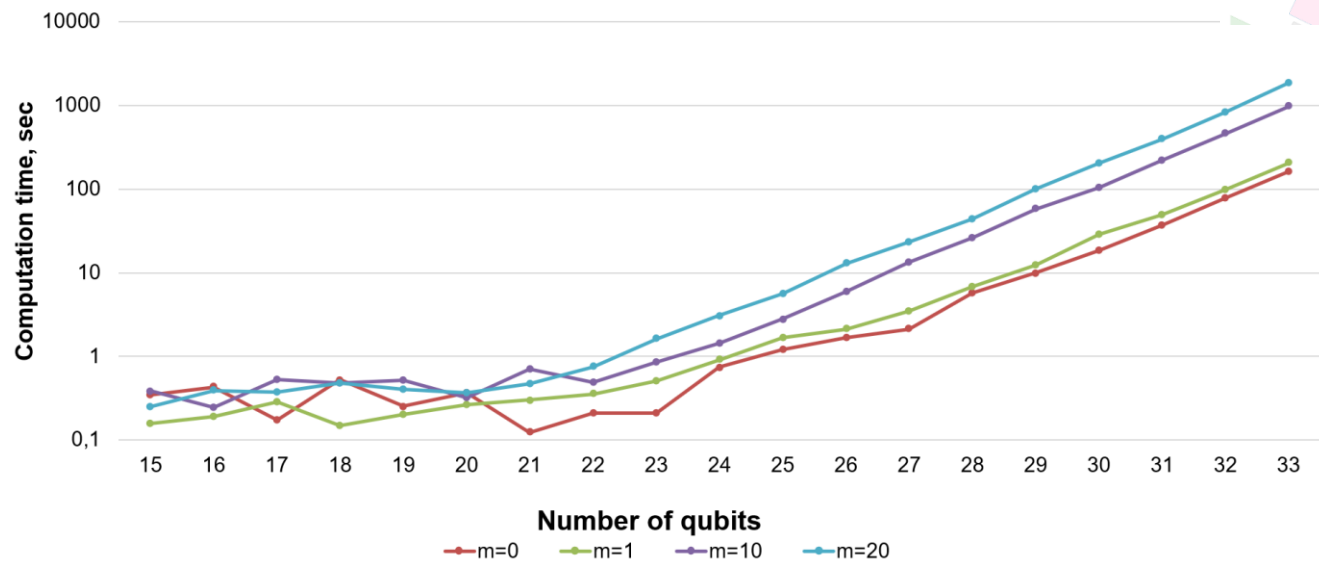
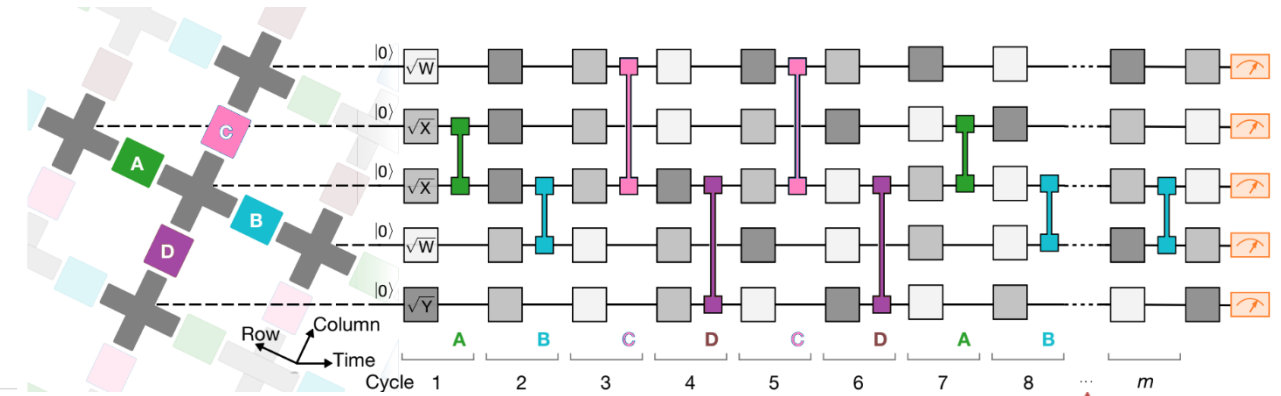
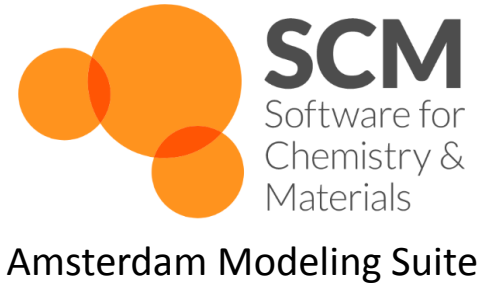
### Average training time of one training epoch, sec



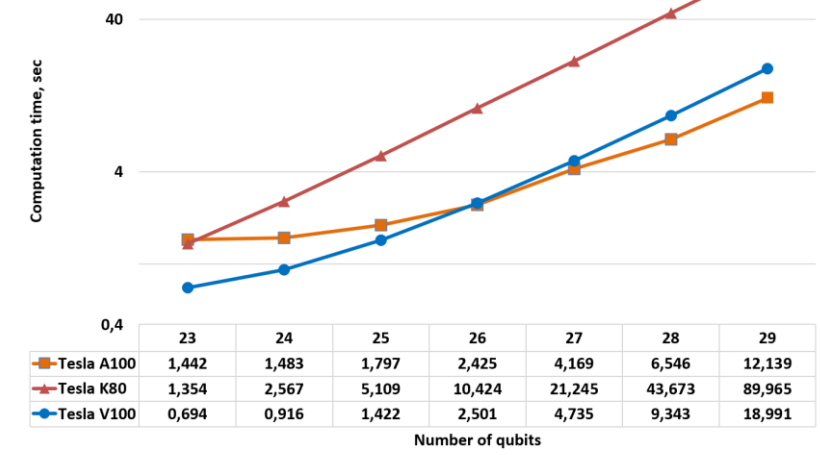
# Quantum polygon



Quantum Exact Simulation Toolkit  
<https://quest.qtechtheory.org/>



Time dependency on the number of qubits for CPU calculations with OpenMP technology

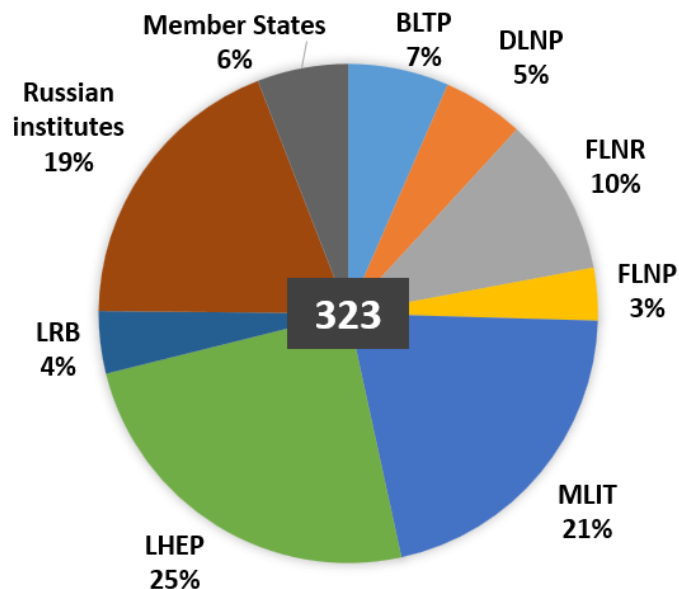


Time dependency calculations on the number of qubits for various GPUs

Для задач, решаемых в рамках Проекта была создана вычислительная система по требованию, содержащую 288 физических ядер (576 логических ядер) и файловое хранилище емкостью 7 ТБ под управлением файловой системы NFS. На этой системе проводились интенсивные расчеты с использованием ПО AMS, DIRAC, QuEST и др. для расчетов электронных свойств сверхтяжелых элементов. За время выполнения проекта было решено 4200 задач на которые было затрачено 740 000 ядро-часов



# “Govorun” supercomputer

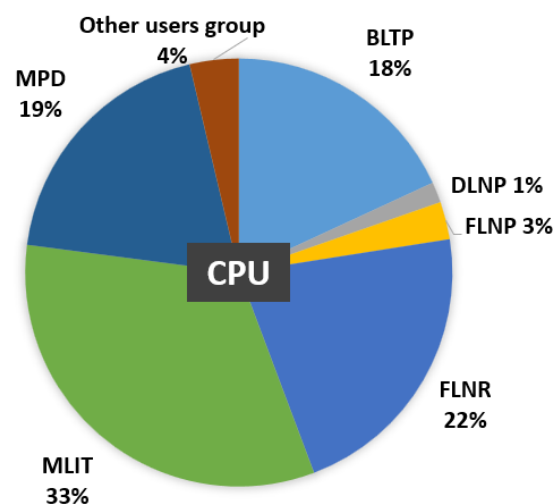
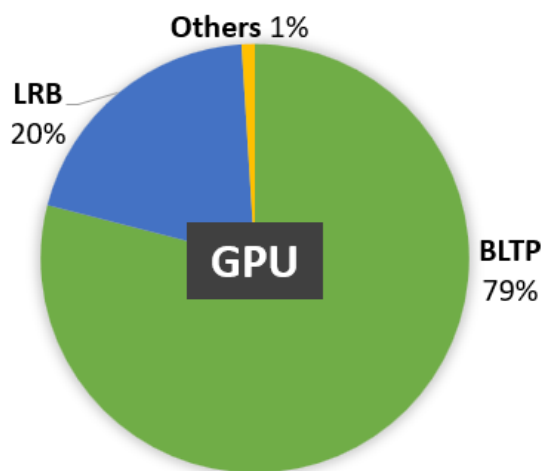


## Key projects that use the resources of the “Govorun” supercomputer:

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

Total number of users of the "Govorun" supercomputer: **323**

## Distribution of the GPU and CPU resources by user group



During February-August 2022, **555079** jobs were performed on the **CPU** component of the “Govorun” supercomputer, which corresponds to **~8M** core hours, and **455** jobs were carried out on its **GPU** component, which corresponds to **32890** GPU hours.

The average load of the CPU and GPU components amounted to 96.2% and 91.4% respectively.

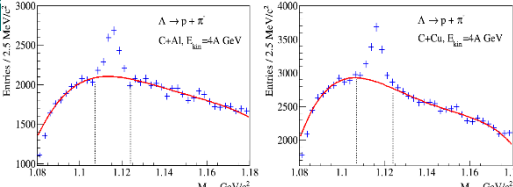
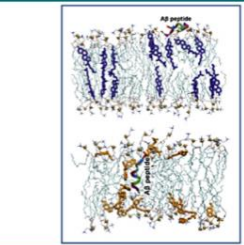
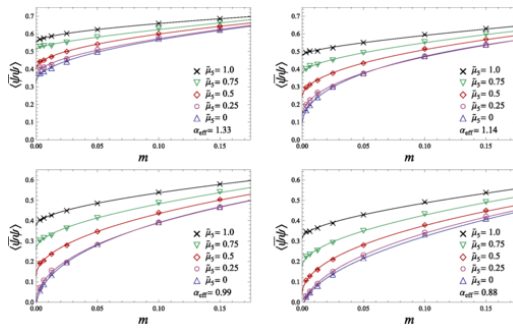
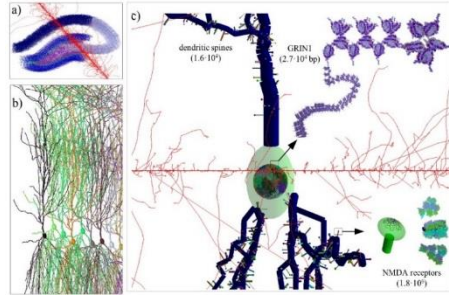
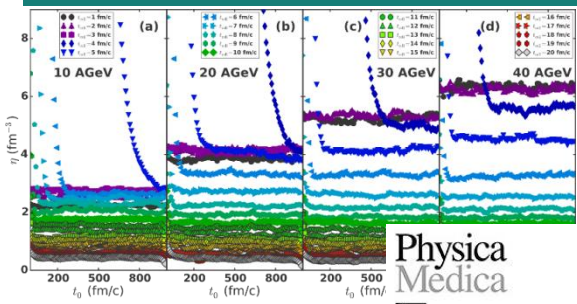
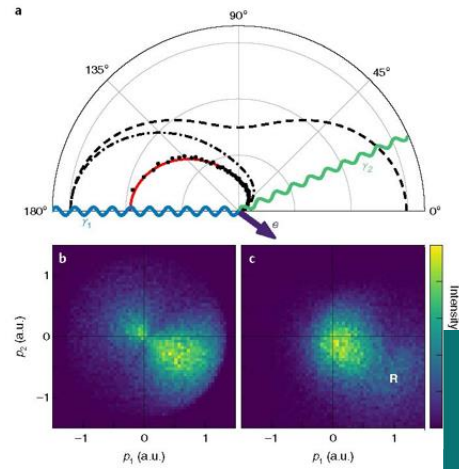
# “Govorun” Supercomputer for JINR Tasks

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

Research results obtained using the supercomputer resources are presented in **204** publications.

Using the results obtained at the Govorun SC, 2 publications were prepared in Nature Physics:

- M. Kircher ... , **O. Chuluunbaatar** et al. Kinematically complete experimental study of Compton scattering at helium atoms near the threshold. Vol. 16. № 4. Pp. 756-760
- BM@N Collaboration. Unperturbed inverse kinematics nucleon knockout measurements with a 48 GeV/c carbon beam. Vol. 17. Pp. 693-699



# SC Govorun included into unified supercomputer network of Russia



On 24 September, an agreement was signed in St. Petersburg on uniting three supercomputers, including the object of the scientific infrastructure of the JINR Member States – the “Govorun” supercomputer – into a single network. Its aim is to develop the National Research Computer Network of Russia.

Deputy Prime Minister of the Russian Federation Dmitry Chernyshenko



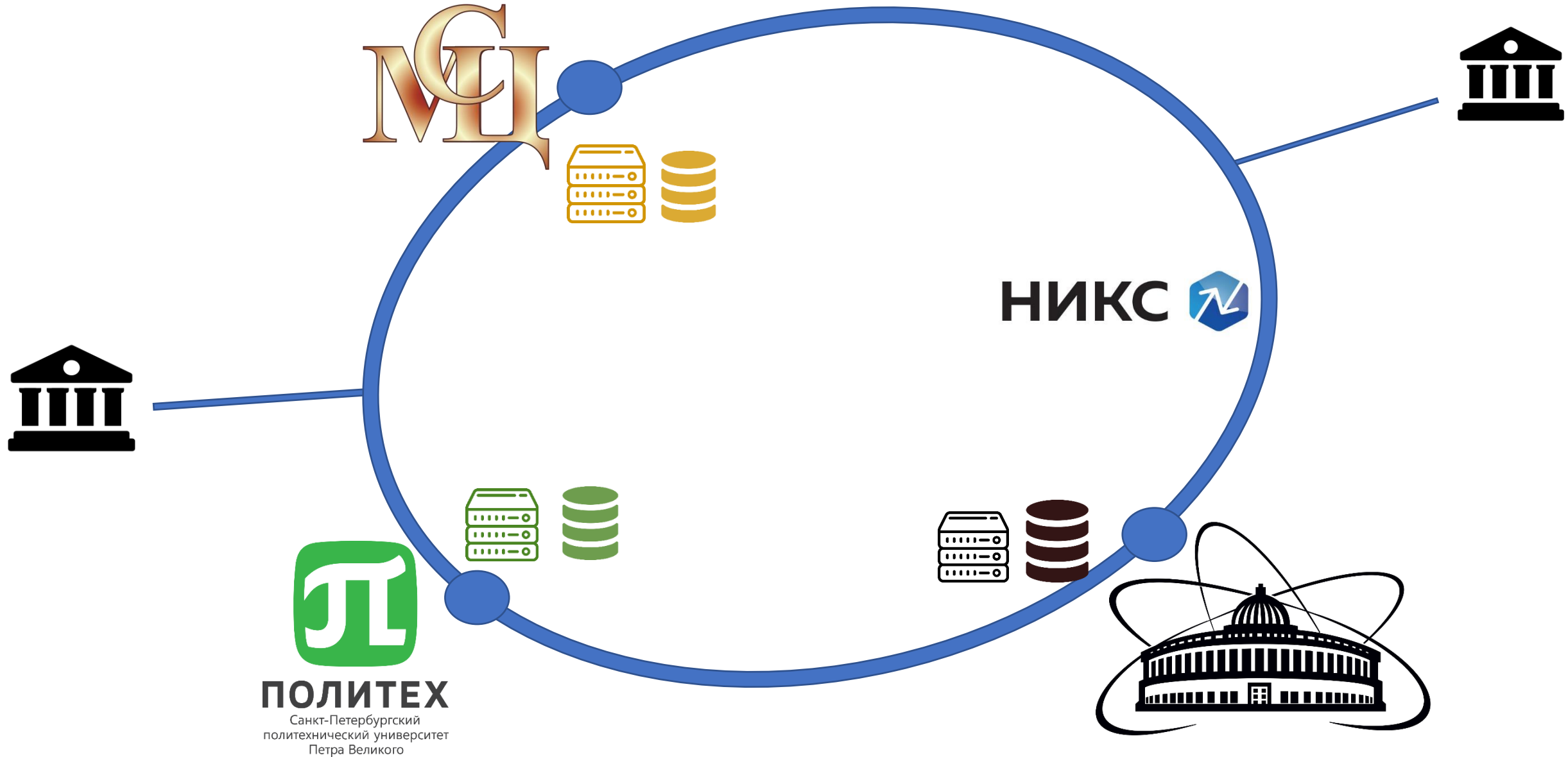
Director of the Meshcheryakov Laboratory of Information Technologies JINR Vladimir Korenkov

Rector of Peter the Great St. Petersburg Polytechnic University Andrey Rudskoi

Director of the Joint Supercomputer Centre of the Russian Academy of Sciences Boris Shabanov

There is a unified scientific and educational space of information technologies being formed in Russia. Scientific world-level centres, scientific-educational and engineering centres gain an opportunity for distributed work with big data at megascience scientific facilities in supercomputer centres. Researchers and developers will be provided with global access to services of machine learning, big data analysis, supercomputer resources.





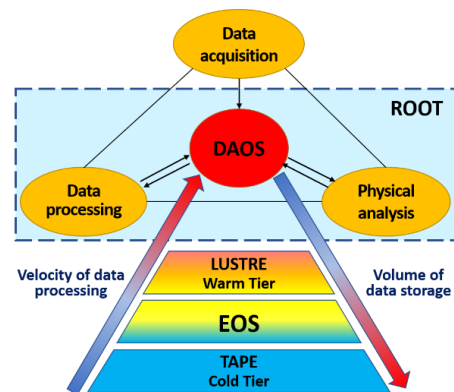
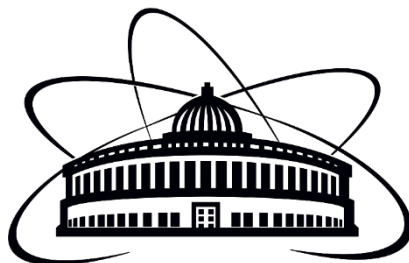
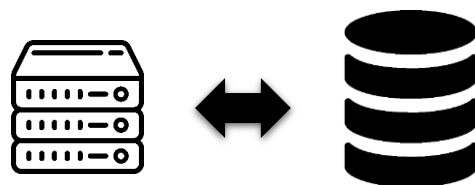
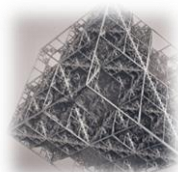
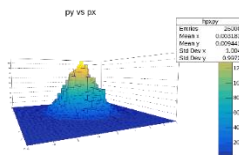
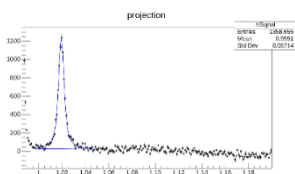
На основе интеграции суперкомпьютеров ОИЯИ, Межведомственного суперкомпьютерного центра РАН и Санкт-Петербургского политехнического университета Петра Великого создана масштабируемая исследовательская инфраструктура нового уровня. Такая инфраструктура востребована для задач меганауки NICA.



## Реконструкция траекторий, Физический анализ



**ПОЛИТЕХ**  
Санкт-Петербургский  
политехнический университет  
Петра Великого



Основными направлениями развития создаваемой инфраструктуры являются:

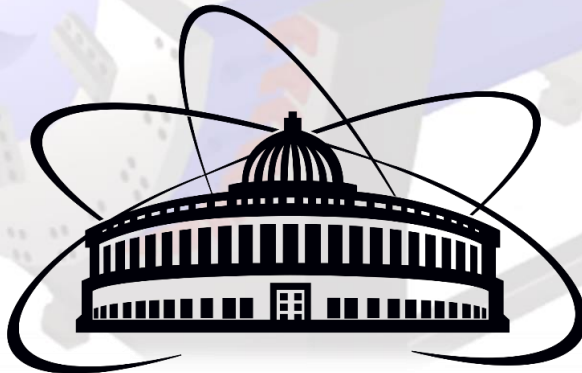
- объединение суперкомпьютерных ресурсов в интегрированную территориально распределённую сеть;
- создание экосистемы профессионального сообщества пользователей суперкомпьютерными ресурсами;
- совместное развитие средств хранения и обработки больших объемов данных;
- создание облачных цифровых сервисов для доступа к суперкомпьютерным ресурсам;
- создание сервисов машинного обучения и аналитики больших данных, распределенных витрин данных для пользователей научных и образовательных организаций.



Слияние физических данных



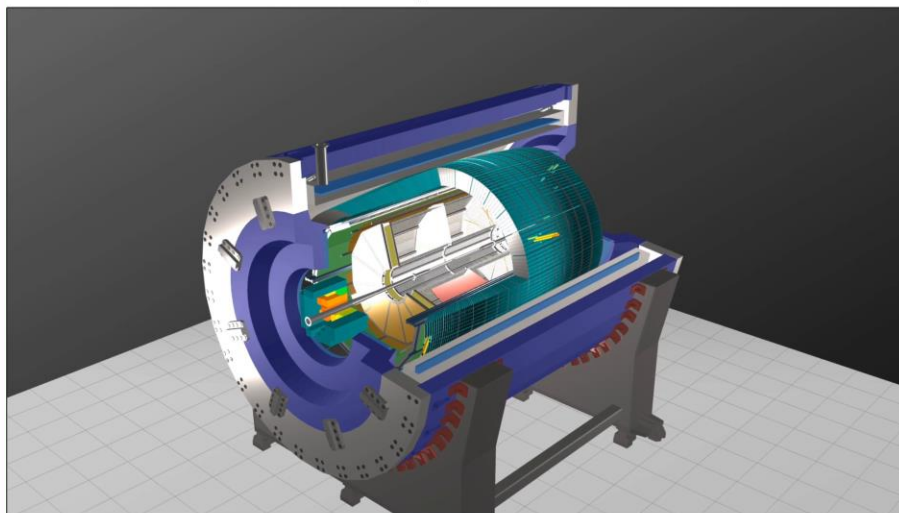
**ПОЛИТЕХ**  
Санкт-Петербургский  
политехнический университет  
Петра Великого



## ДАННЫЕ



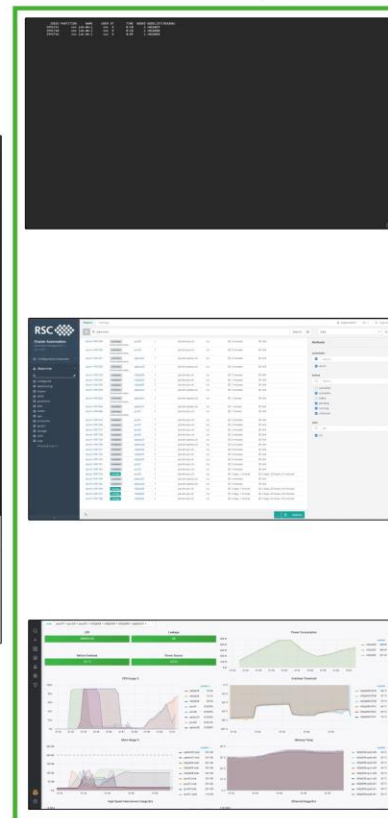
## Центр управления виртуальным экспериментом Multi-Purpose Detector



00:00:44:19



## ЗАДАЧИ



В январе 2022 года успешно завершен первый совместный эксперимент по использованию единой суперкомпьютерной инфраструктуры для задач мегасайенс-проекта NICA:

- ✓ для эксперимента MPD было запущено **3000** задач генерации данных;
- ✓ было сгенерировано порядка **3 миллиона событий**;
- ✓ полученные данные были переданы на СК «Говорун» для дальнейшей обработки и физического анализа.

1. Впервые в мире создана и внедрена гиперконвергентная архитектура для вычислительных узлов суперкомпьютера. Гиперконвергентность вычислительных узлов позволяет «оркестрировать» вычислительными ресурсами и элементами хранения данных и создавать, используя ПО РСК БазИС, вычислительные системы по требованию задач пользователя.
2. Разработана и внедрена иерархическая система обработки и хранения данных, представляющая собой единую централизованно управляемую систему и имеющая несколько уровней хранения данных - очень горячие данные, горячие данные и теплые данные. Внедрение иерархической системы обработки и хранения данных позволяет существенно повысить эффективность работы с большими массивами данных, в том числе для проекта NICA.
3. Гибкая архитектура СК «Говорун» позволяет не только проводить расчеты, но и использовать суперкомпьютер как научно-исследовательский полигон для выработки программно-аппаратных и ИТ-решений для задач, решаемых в ОИЯИ.





***Thank you for your attention***

**HYBRILIT HETEROGENEOUS PLATFORM at MLIT JINR:**

<http://hlit.jinr.ru>

