IT solutions for JINR tasks at “Govorun” supercomputer

M.I. Zuev
on behalf of HybriLIT team

Meshcheryakov Laboratory of Information Technologies
Joint Institute for Nuclear Research

9th International Conference "Distributed Computing and Grid-technologies in Science and Education” (GRID’2021)

5 – 9 July 2021

Development of the heterogeneous HybriLIT platform

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

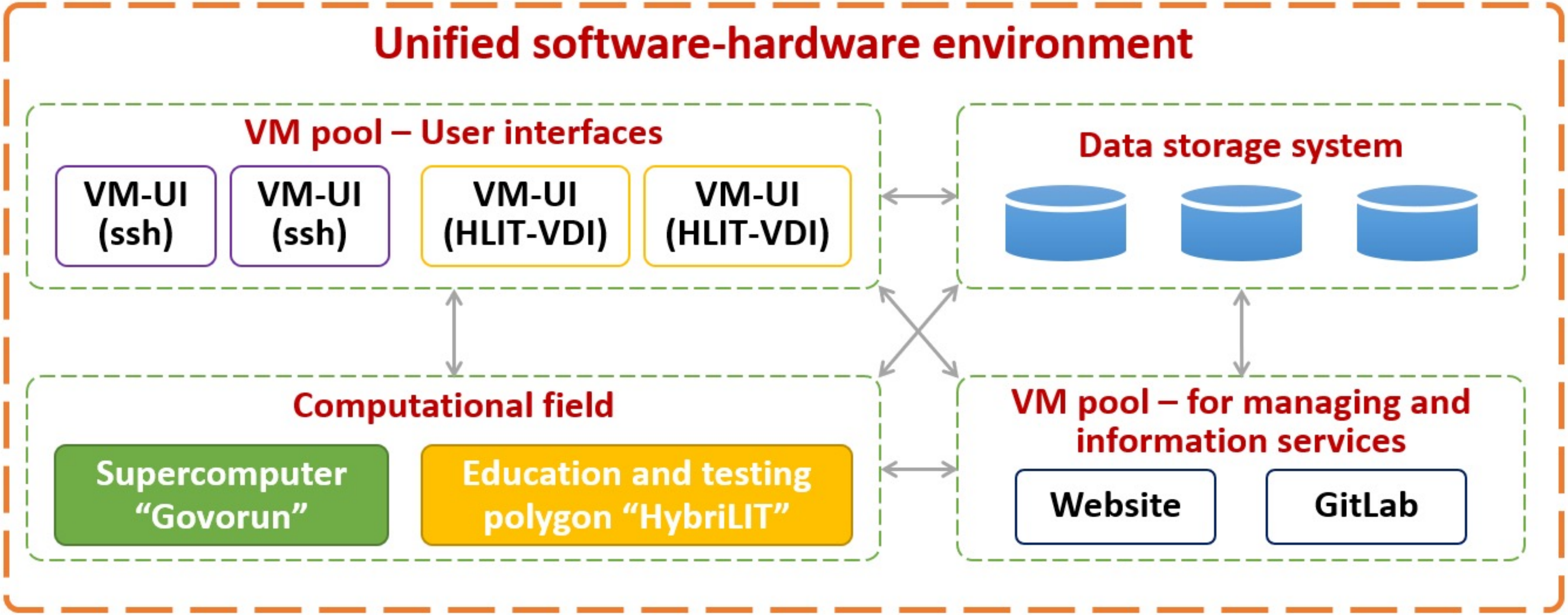
#18 в Top50

“Govorun” supercomputer
First stage 2018:
Full peak performance :
1 PFlops for single precision
500 TFlops for double precision

#10 в Top50

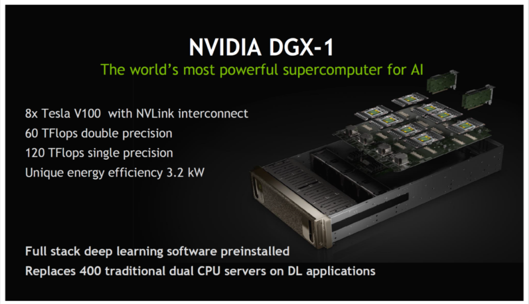
“Govorun” supercomputer
Wecond stage 2019:
Full peak performance :
1.7 PFlops for single precision
860 TFlops for double precision
288 ТБ ССХД with I/O speed >300 Gb/s

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.

“Govorun” supercomputer



The “Govorun” supercomputer is a hyper-converged software-defined system and occupied **26th and 31st** in the current edition of the **IO500** list (July 2021). For the high-speed data storage system, RSC Group has received the prestigious Russian DC Awards 2020 in “The Best IT Solution for Data Centers” nomination at the awards ceremony held on 10 December 2020 in Moscow.

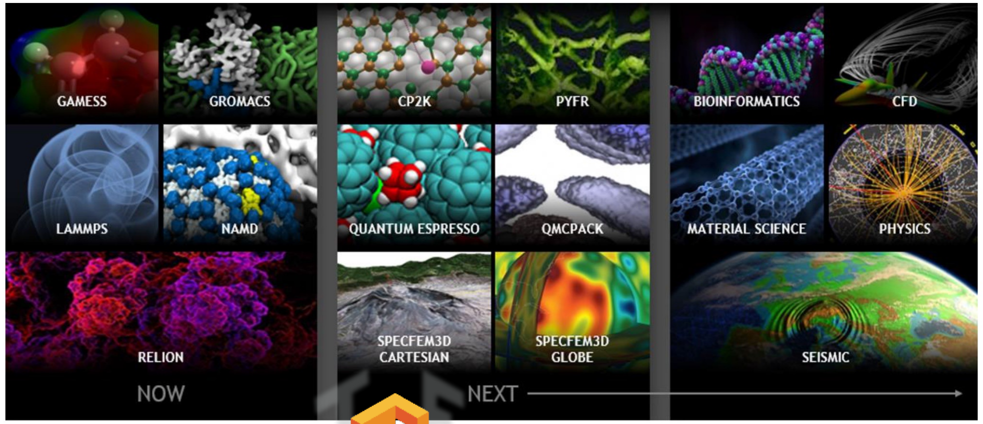
GPU-component
based on NVIDIA DGX-1 Volta.

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

Total peak performance:
1.7 PFLOPS SP
860 TFLOPS DP
300 Gb/s Data IO rate

The GPU-component of the "Govorun" supercomputer

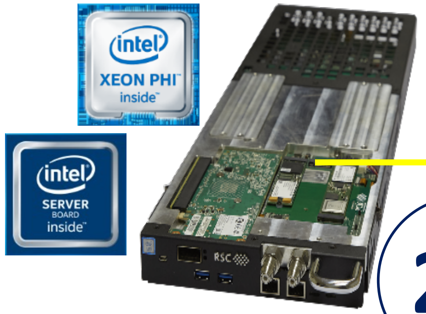
The GPU-component consists of **5 NVIDIA DGX-1 servers**. Each server has **8 GPU NVIDIA Tesla V100** based on the latest architecture NVIDIA Volta. Moreover, one server NVIDIA DGX-1 has **40960** cores CUDA, which are equivalent to 800 high-performance central processors. A whole number of novel technologies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



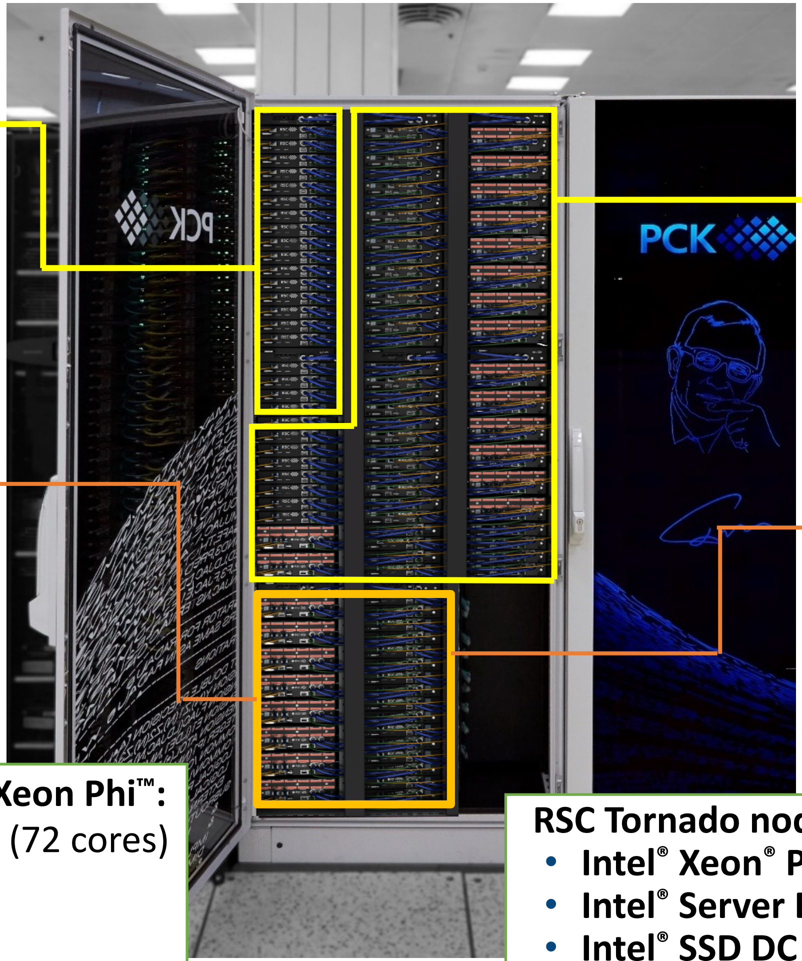
The GPU-component gives a users of the supercomputer a possibility to allow as massively parallel computation for general-purpose tasks using such technologies as CUDA and OpenCL, as well as use applications already adapted for this architecture. Also, GPU-component allow to use machine learning and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

TensorFlow PYTORCH theano APACHE mxnet The Microsoft Cognitive Toolkit Caffe

The CPU-component of the “Govorun” supercomputer



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



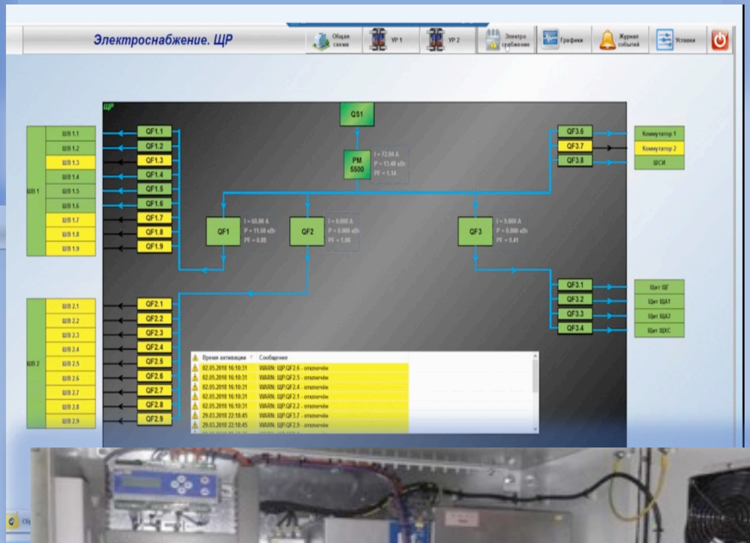
RSC Tornado nodes based on Intel® Xeon Phi™:

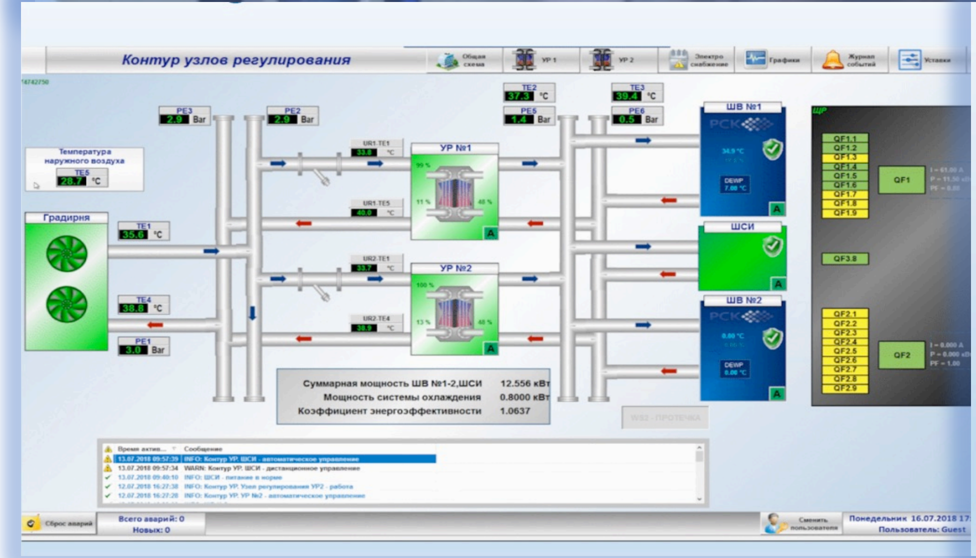
- **Intel® Xeon Phi™ 7290** processors (72 cores)
- **Intel® Server Board S7200AP**
- **Intel®** SSD DC S3520 (SATA, M.2)
- 96GB DDR4 2400 GHz RAM
- **Intel® Omni-Path 100 Gb/s adapter**

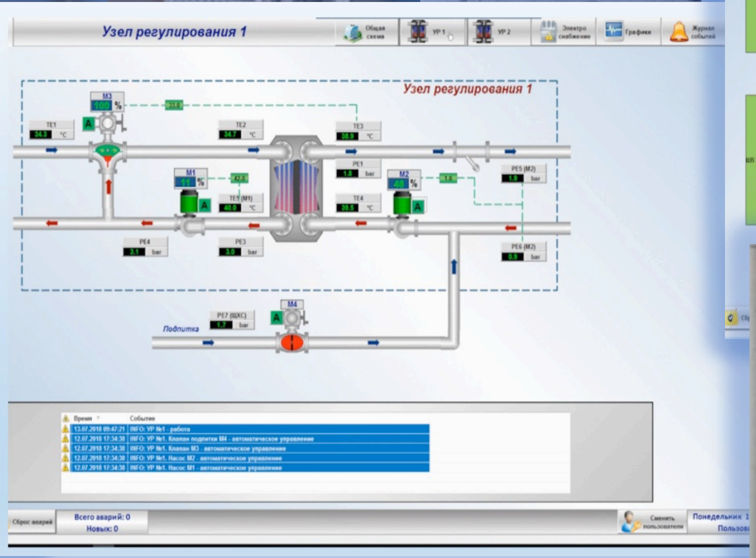
RSC Tornado nodes based on Intel® Xeon® Scalable gen 2:

- **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**

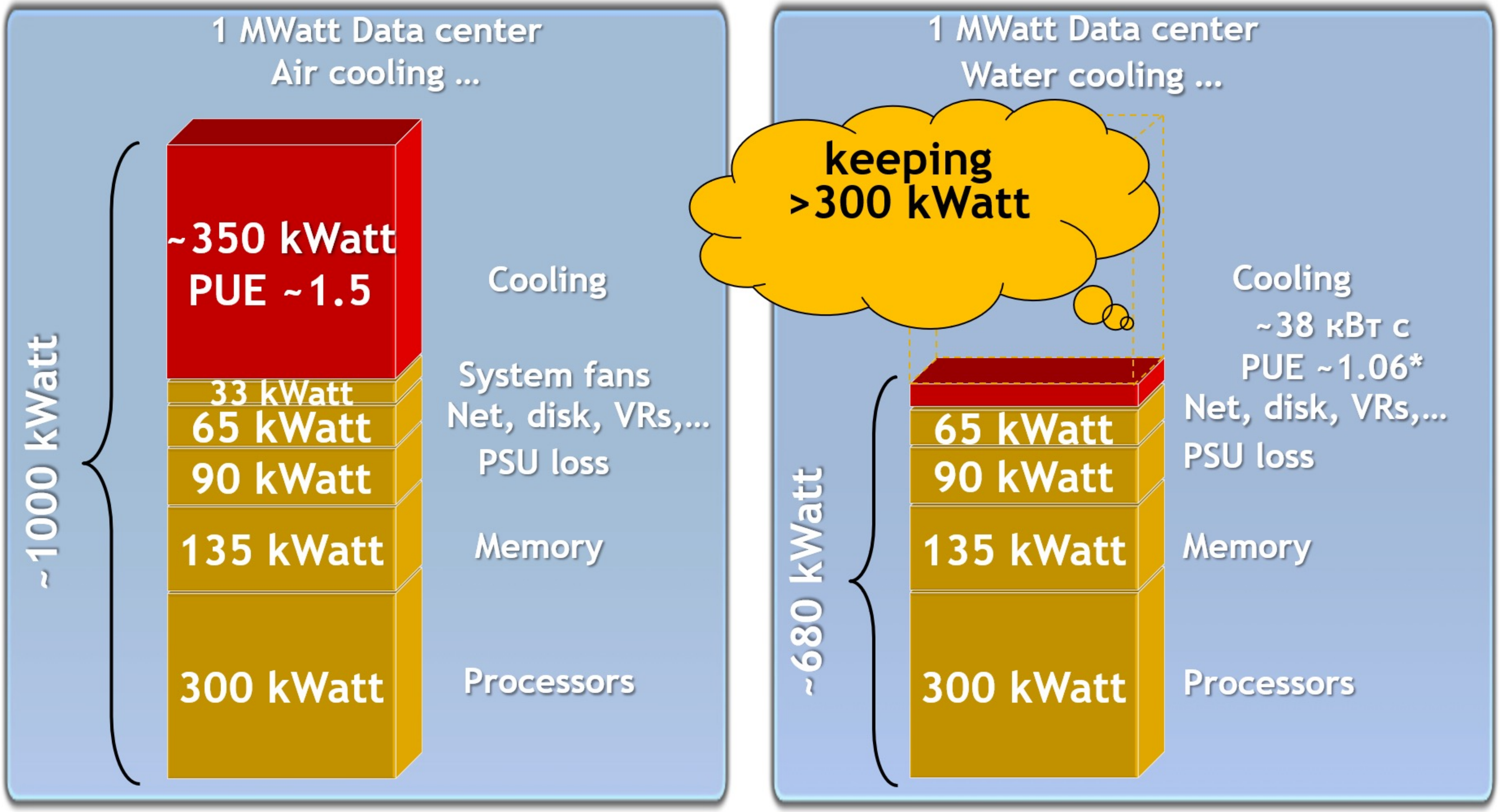
Engineering infrastructure







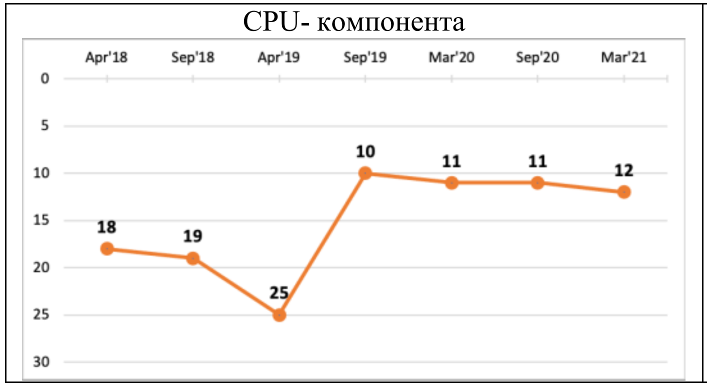
Cooling efficiency and cost

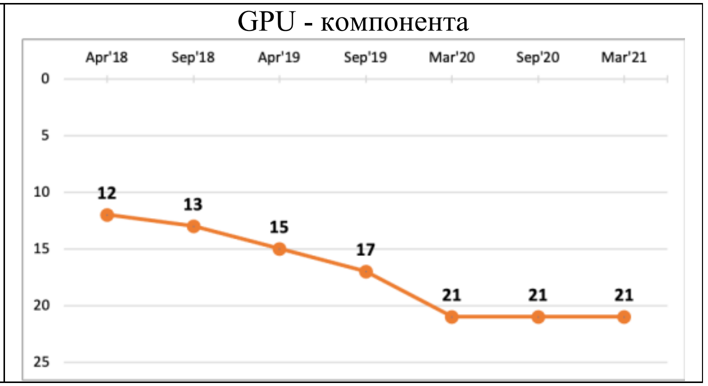


Current ratings of the “Govorun” supercomputer

№	Site	System	CPU-cores Number of acselerators	Rmax (Tflops/s)	Rpeak (Tflops/s)
1	SberCloud, Moscow	Christofari	3600 1200 NVIDIA Tesla V100	6669.0	8789.76
2	MSU, Moscow	Lomonosov 2	23424 1536 NVIDIA Tesla K40M 320 NVIDIA Tesla P100	2478.0	4946.79
3	Moscow	ФГБУ «ГВЦ Росгидромета»	35136	1200.35	1293.0
12	**JINR, Dubna**	**“Govorun” supercomputer CPU-component**	**5192 21 Intel Xeon Phi 7290**	**312.62**	**463.26**
21	**JINR, Dubna**	**“Govorun” supercomputer GPU-component**	**200 40 NVIDIA Tesla V100**	**175.13**	**319.0**

Rank	Institution	System	io500		
			Score	BW GiB/s	MD kIOP/s
1	Pengcheng Laboratory	Pengcheng Cloudbrain-II on Atlas 900	36,850.37	3,421.62	396,872.82
2	Intel	Endeavour	1,859.56	398.77	8,671.65
3	Intel	Wolf	1792.98	371.67	8649.57
26	**JINR**	**Govorun (DAOS)**	**132.06**	**20.19**	**863.69**
27	Joint Supercomputer Center of the Russian Academy of Sciences	MVS10POP2	125.50	45.31	347.61
31	**JINR**	**Govorun (Lustre)**	**90.87**	**35.61**	**231.88**
36	SPbPU	Polytechnic RSC Tornado	64.29	21.56	191.73





The “Govorun” supercomputer is ranked on the 26nd and 31st place in the current edition of the IO500 list (July 2021) and is the second in terms of the data processing rate among Russian supercomputers. The CPU and GPU components rank 12th and 22nd in the current TOP50 list, respectively

Hierarchical data processing and storage system for HPC

Problem: Divide several tens of PBs of data into several levels as efficiently as possible: Very hot, hot, warm, cold, very cold to support the NICA experiment and others with the required speed of data assimilation and processing.

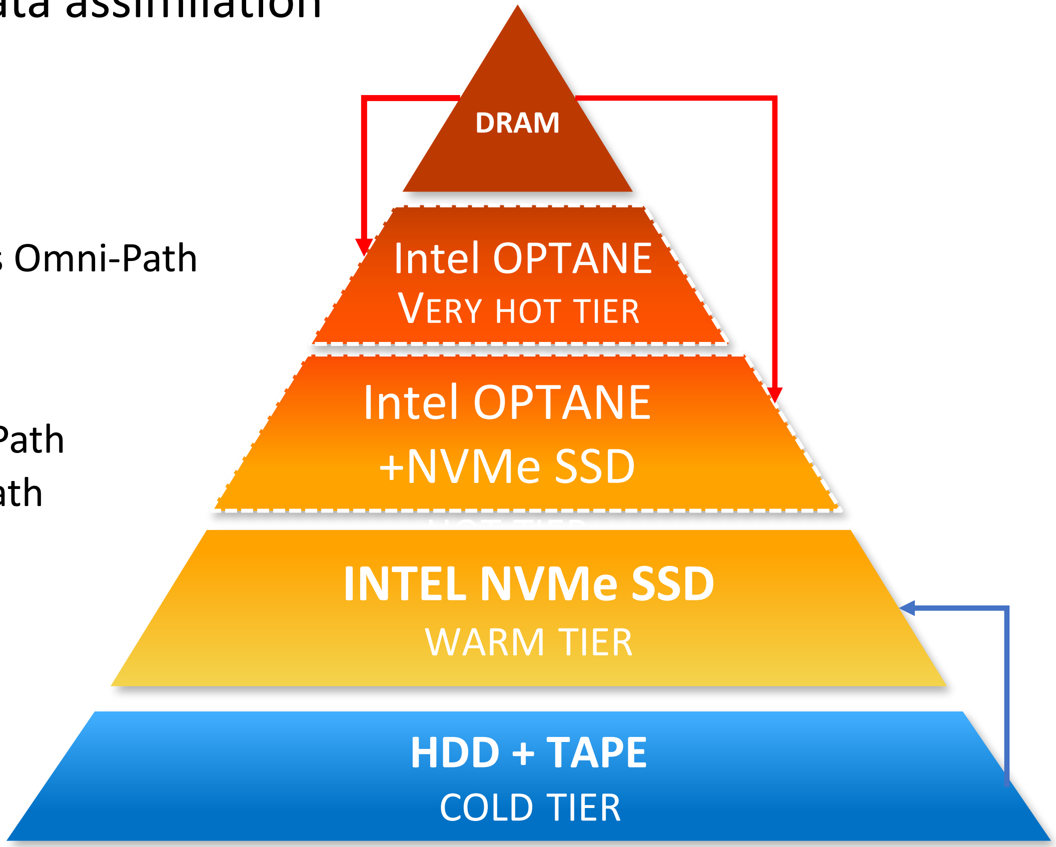
Result: Complete solution includes:

Very hot:
2xIntel Xeon 8268/12x**Intel Optane P4801X 375GB+IMDT**/100Gbs Omni-Path

Hot:
2xIntel Xeon 8268/12x**Optane P4801X 375GB**/100Gbs Omni-Path
2xIntel Xeon 8268/12x**Intel NVMe SSD P4511 2TB**/100Gbs Omni-Path
2xIntel Xeon 8268/2x**Intel NVMe SSD P4511 2TB**/100Gbs Omni-Path

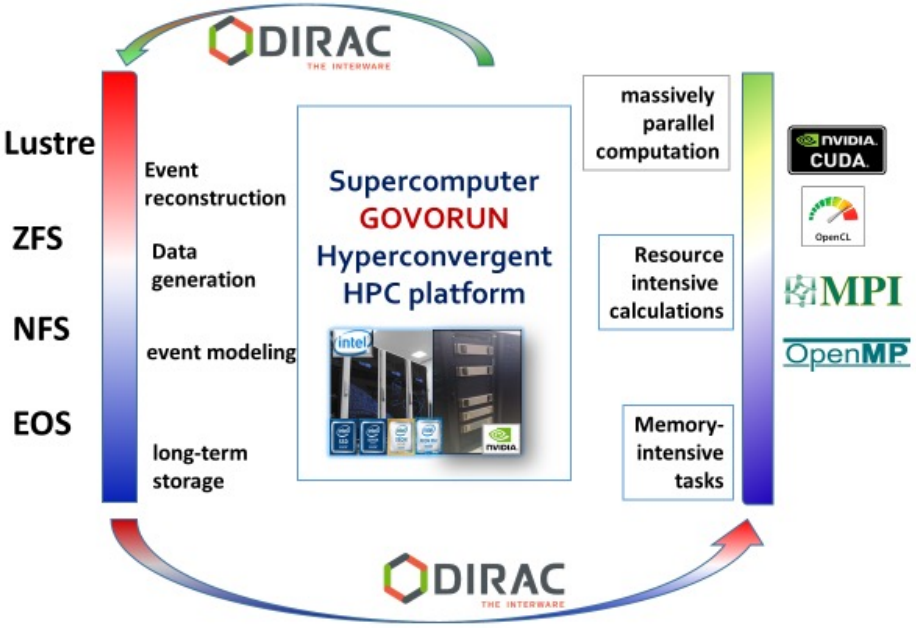
Warm:
2xIntel Xeon 8268/12x**Intel NVMe SSD Ruler**/100Gbs Omni-Path

Cold:
2xIntel Xeon 6xxx/24x14TB HDD/10Gbs Ethernet and Tape

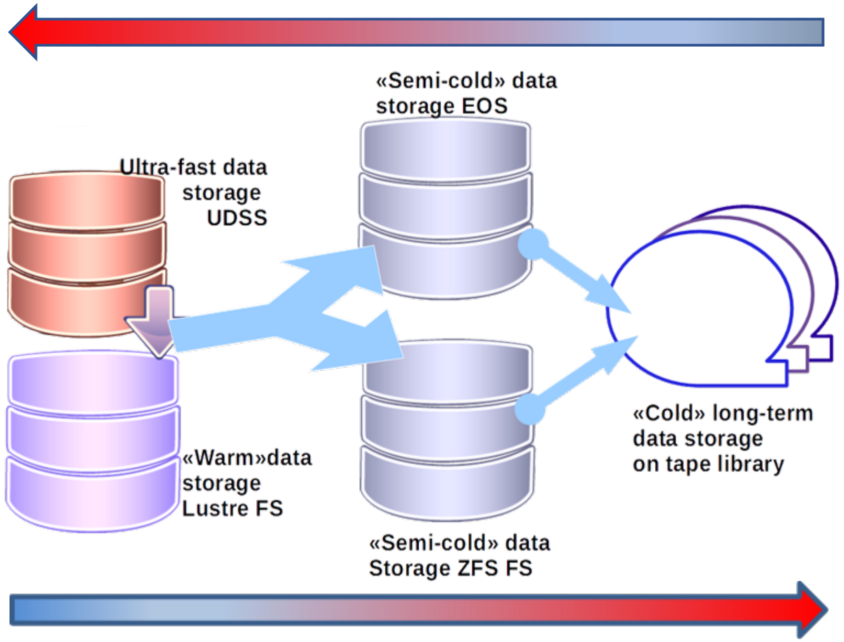


Current state: More than **300 Gb/s Data IO rate**

Technologies for storing, processing and analyzing experimental data in the NICA megasience project

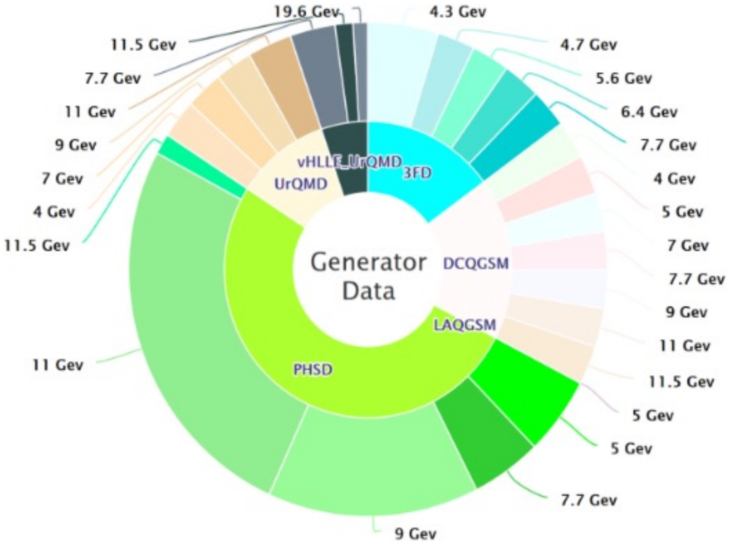


Velocity of data processing



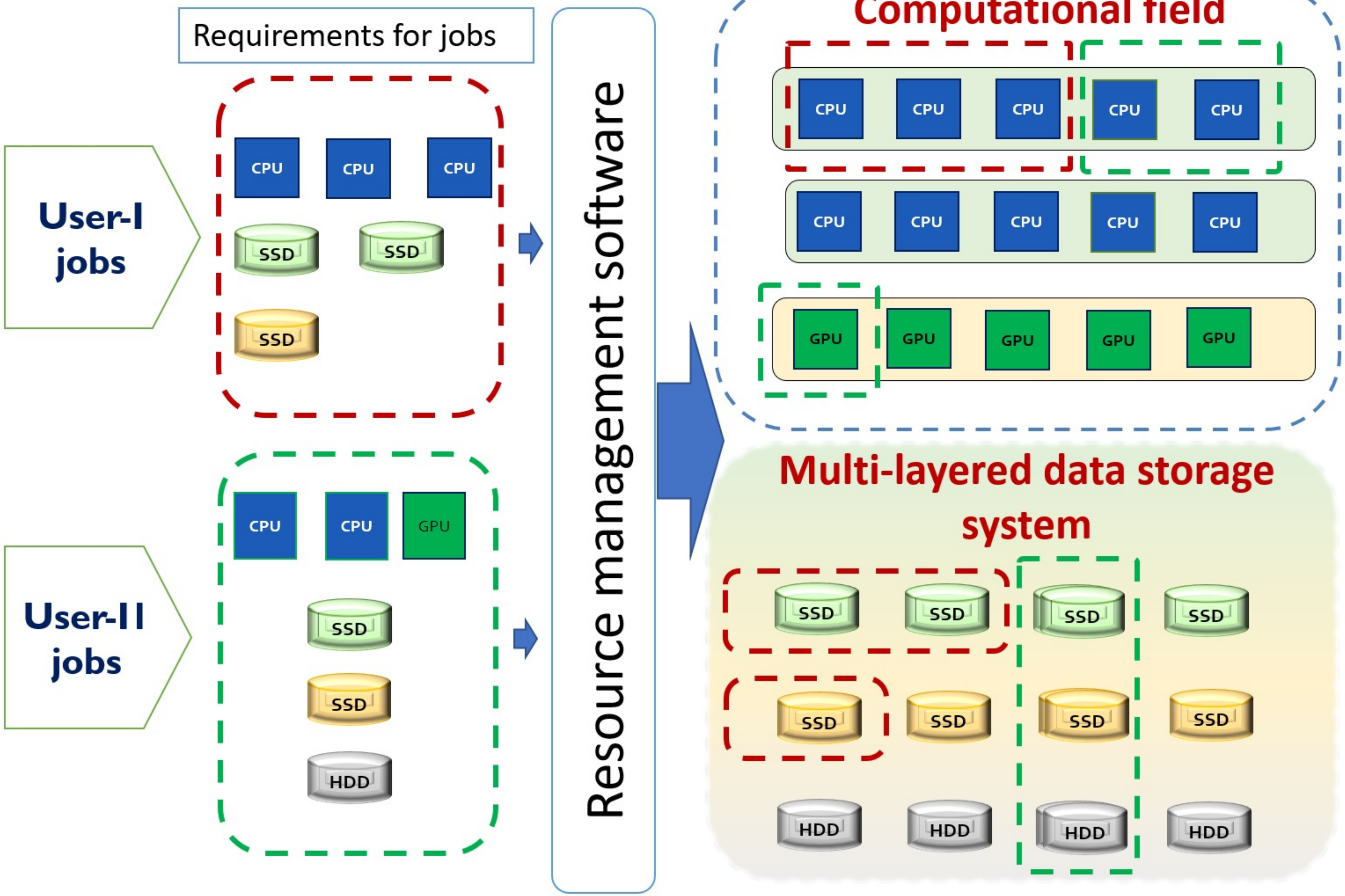
Volume of data storage

Events of the MPD experiment are simulated and reconstructed on ultrafast data storage system under the **FS Lustre** management with a subsequent transfer to semi-cold storages (**FS ZFS, EOS**) and to the tape library for long-term storage.



About 50 million events were generated for the MPD experiment using the hierarchical structure of working with data. The unique composition of the “Govorun” supercomputer equipment, which includes a super-fast data access system and computing nodes with a large amount of RAM (3 TB per node), made it possible to process the same number of events on almost half the number of computing cores as on other available computing resources.

Orchestration and hyperconvergence on the "Govorun" supercomputer



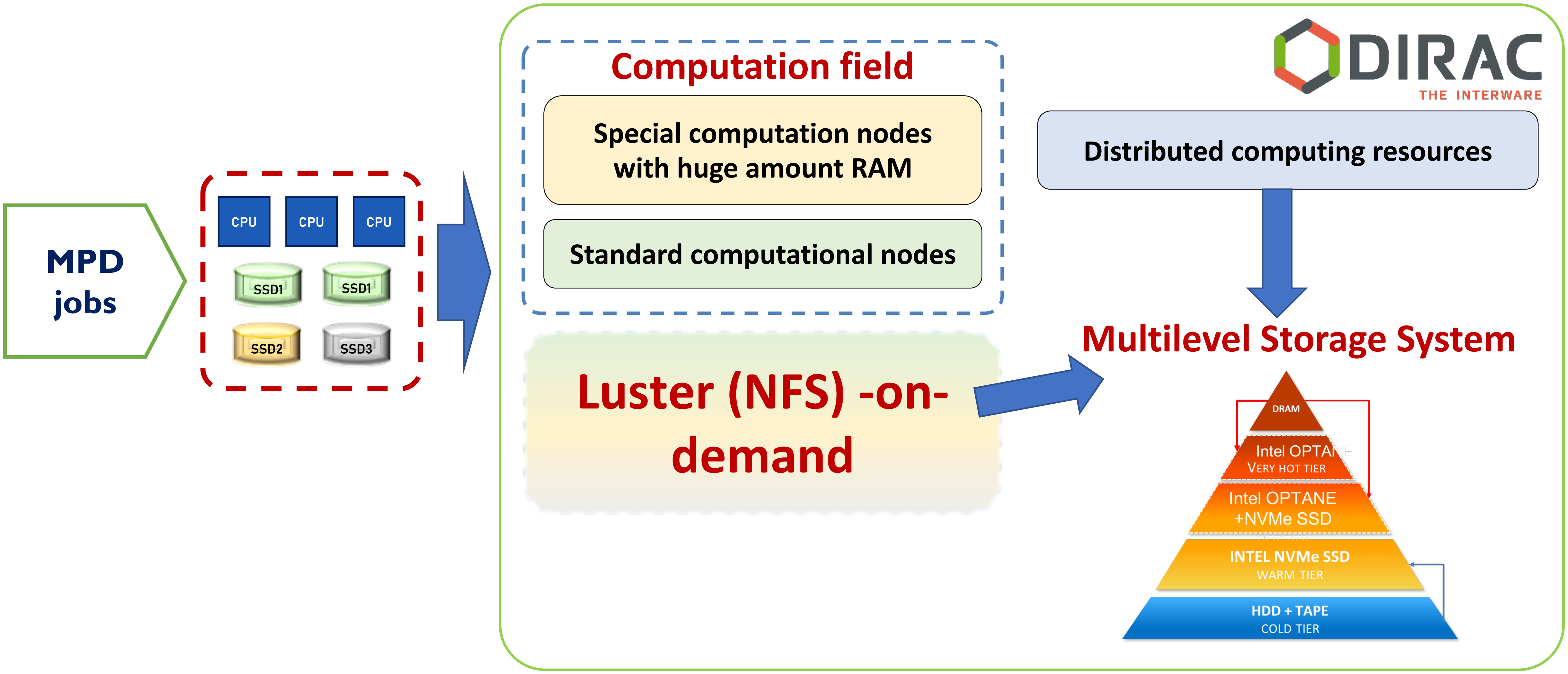
The "Govorun" supercomputer has unique properties for the flexibility of customizing the user's job.

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.

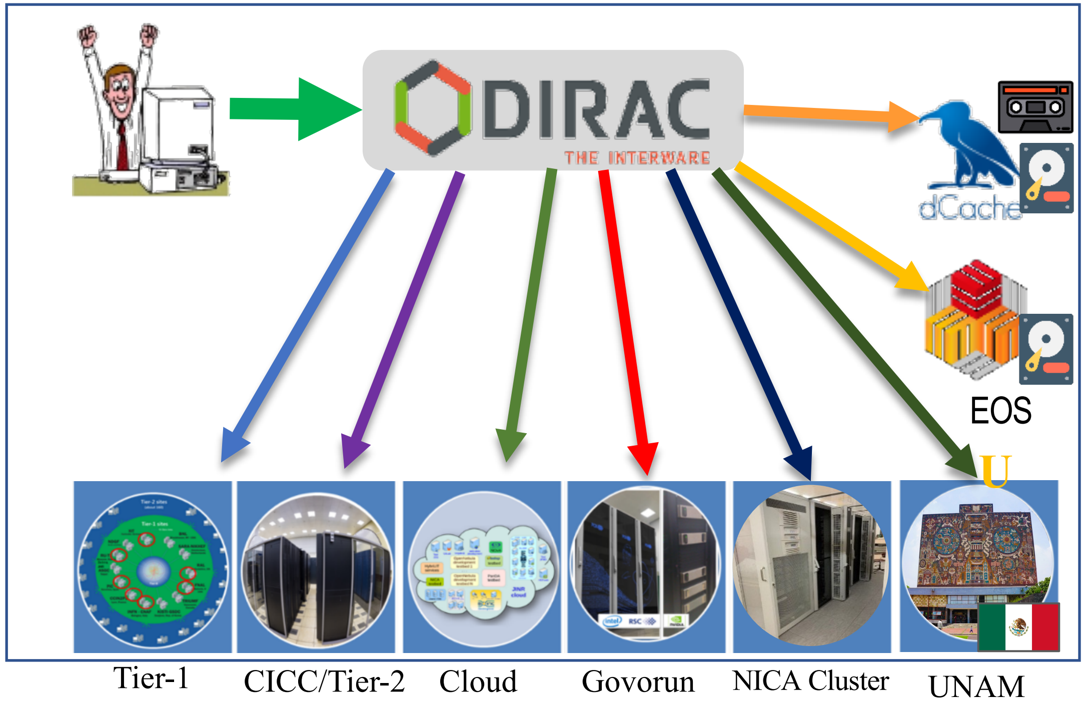
This property enables the effective solution of different tasks, which makes the "Govorun" supercomputer a unique tool for research underway at JINR.

Research results obtained using the SC "Govorun" resources in 2020 are presented in 65 publications, 6 of them in Q1, 7 in Q2.

System-on-demand for MPD

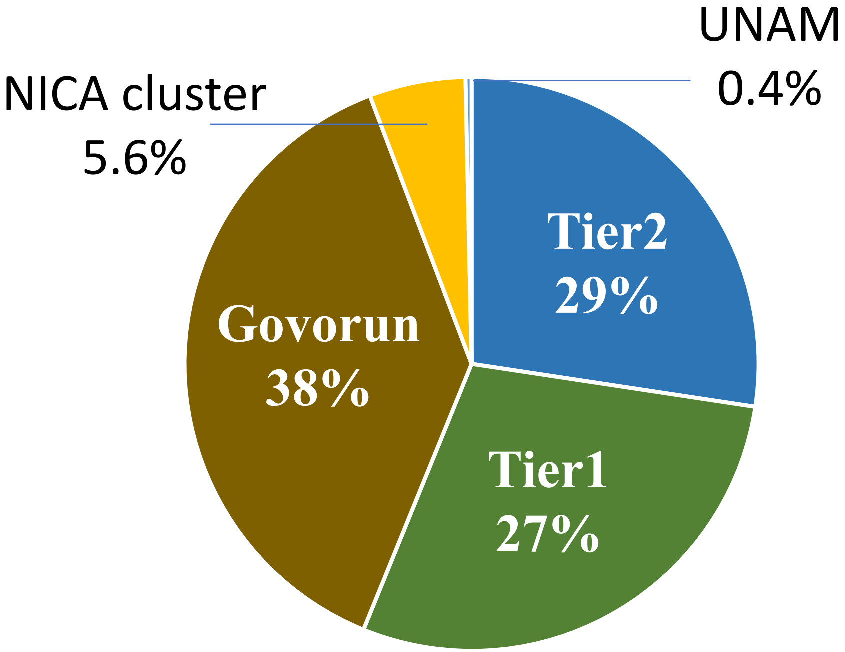


JINR Computing resources integration for the MPD experiment



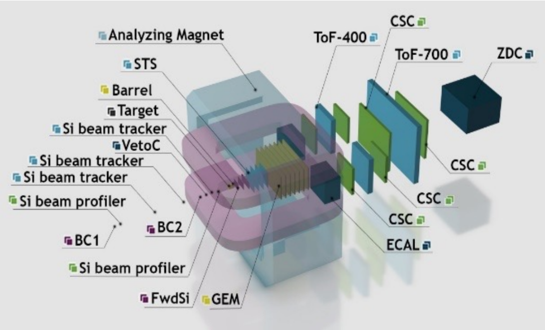
The computing resources of all JINR Multifunctional Information and Computing Complex, i.e. Tier1/Tier2, the “Govorun” supercomputer and storage resources, JINR and Member States cloud resources, NICA Cluster and the cluster of the National Autonomous University of Mexico (UNAM, within cooperation on the MPD project) were combined using the DIRAC Interware.

More than **500,000 MPD jobs** were performed on the Tier1/Tier2 components, the “Govorun” supercomputer, the NICA cluster (fully integrated in July) and the UNAM cluster using the DIRAC platform in the framework of Monte-Carlo data simulation for the MPD experiment.

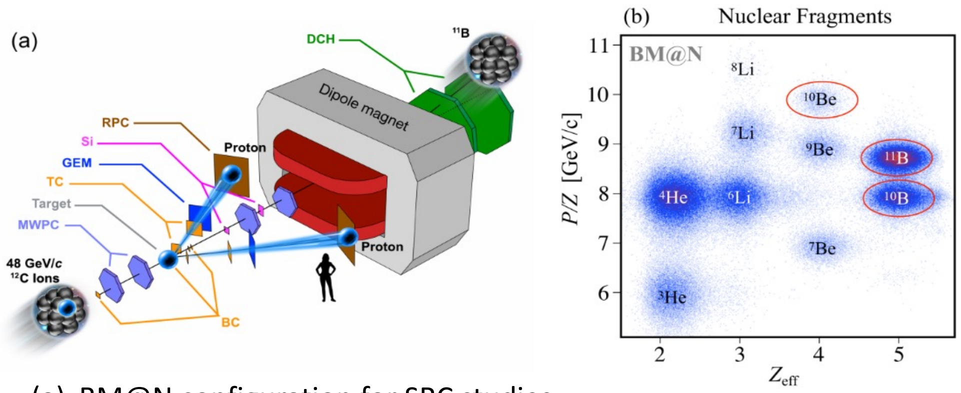


Distribution of simulation jobs by the computing resources via DIRAC

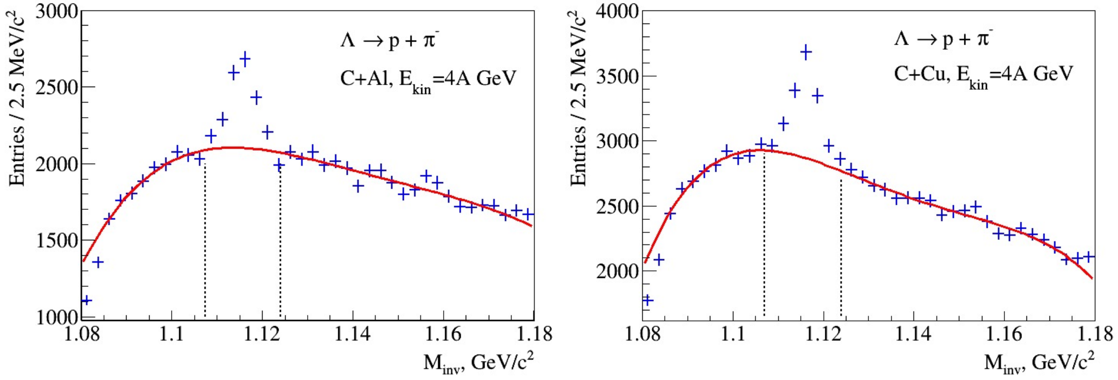
Computing for the NICA megaproject
“Govorun” supercomputer for BM@N tasks



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.

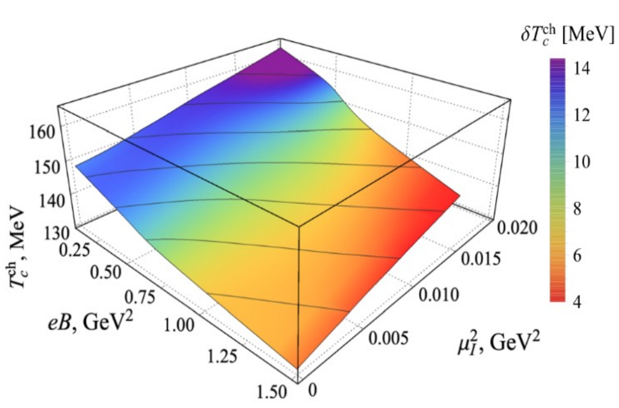


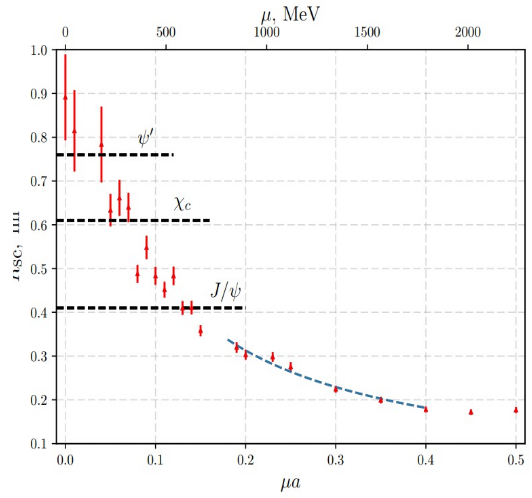
Signals of Λ-hyperons in the spectra of invariant masses (p,π-) measured in C+Al and C+Cu interactions.

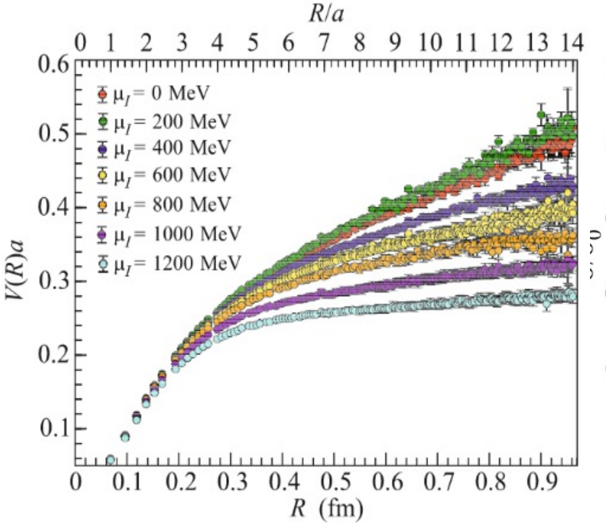
BM@N Collaboration. Production of Λ hyperons in 4 and 4.5 AGeV 2 carbon-nucleus interactions at the Nuclotron // The European Physical Journal A (awaiting publication)

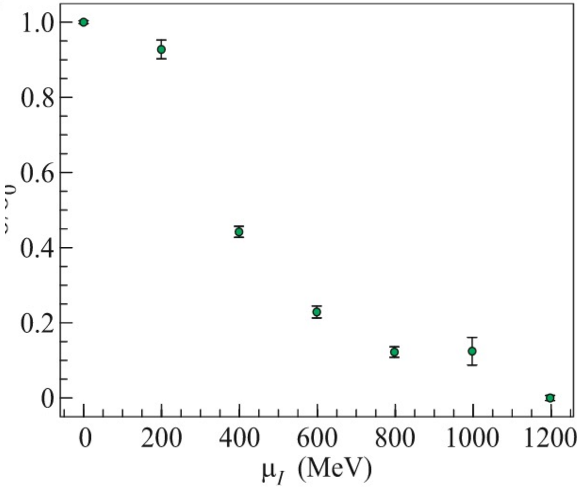
- The analysis of experimental data acquired during the Nuclotron runs in 2016-2018 was performed. Special attention is paid to the study of interactions between beams of carbon and argon ions with fixed targets of different types. The reconstruction of particle tracks was carried out using the method 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 Λ-hyperons was studied using the model data of the DCM-QGSM generator of the BM@N experiment.

“Govorun” supercomputer for QCD tasks









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

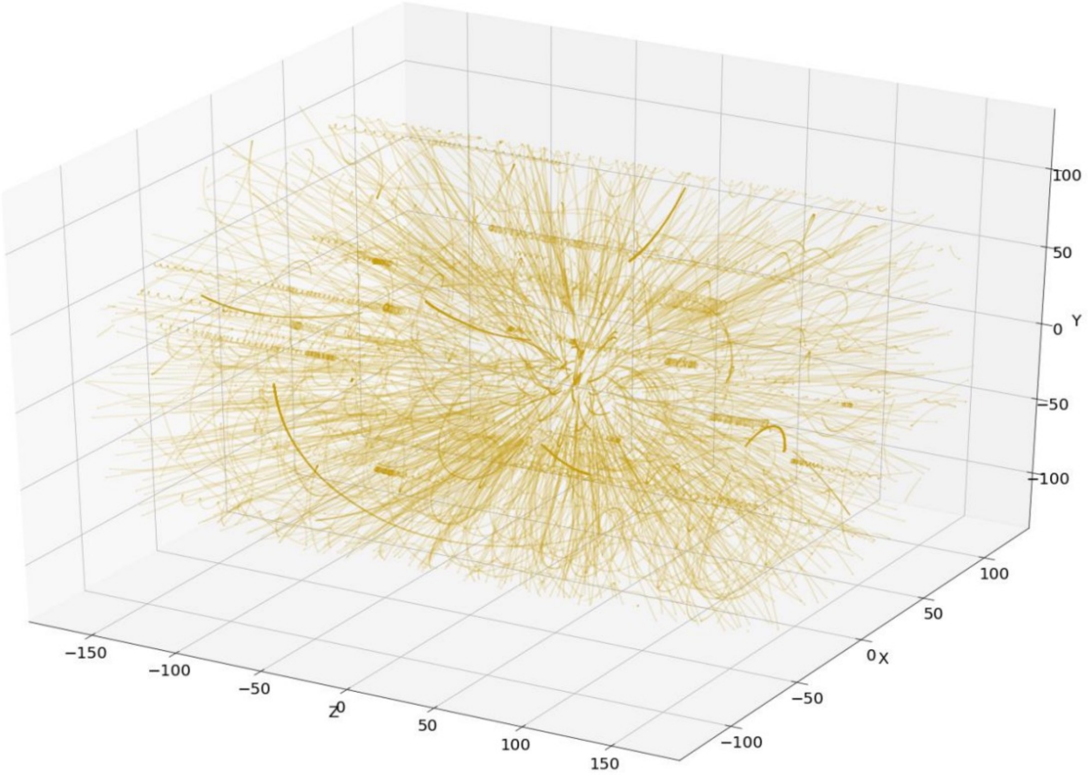
5. N. Yu. Astrakhantsev, V. G. Bornyakov, V. V. Braguta, E.-M. Ilgenfritz, A. Yu. Kotov, A. A. Nikolaev, A. Rothkopf, JHEP 1905 (2019) 171, DOI: 10.1007/JHEP05(2019)171,e-Print: arXiv:1808.06466

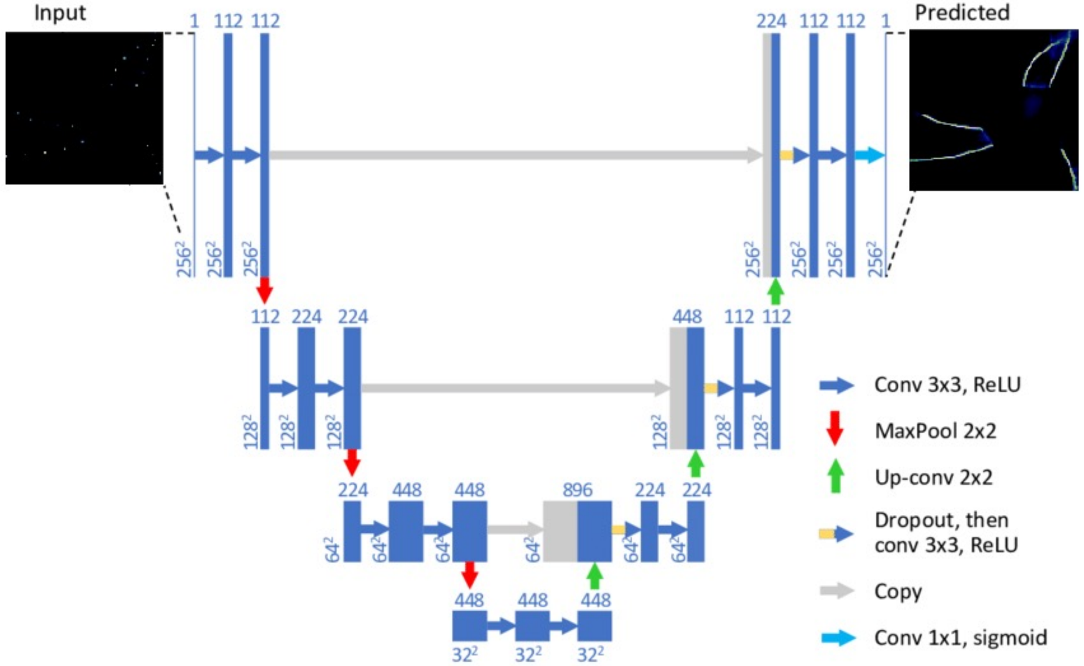
6. https://arxiv.org/abs/1902.09325

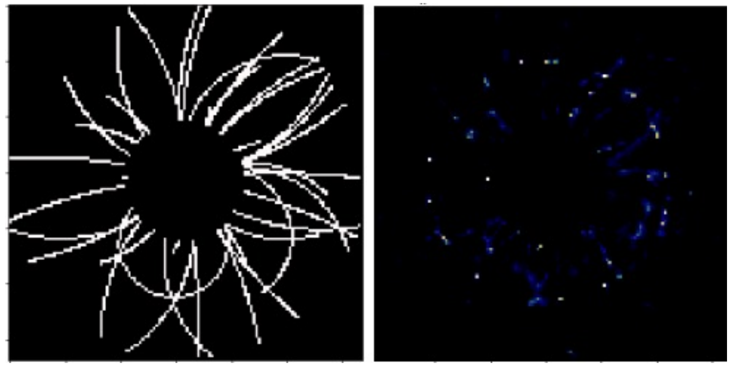
7. http://arxiv.org/abs/1910.08516

Computing for the NICA megaproject
Machine learning for MPD tracking tasks

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.

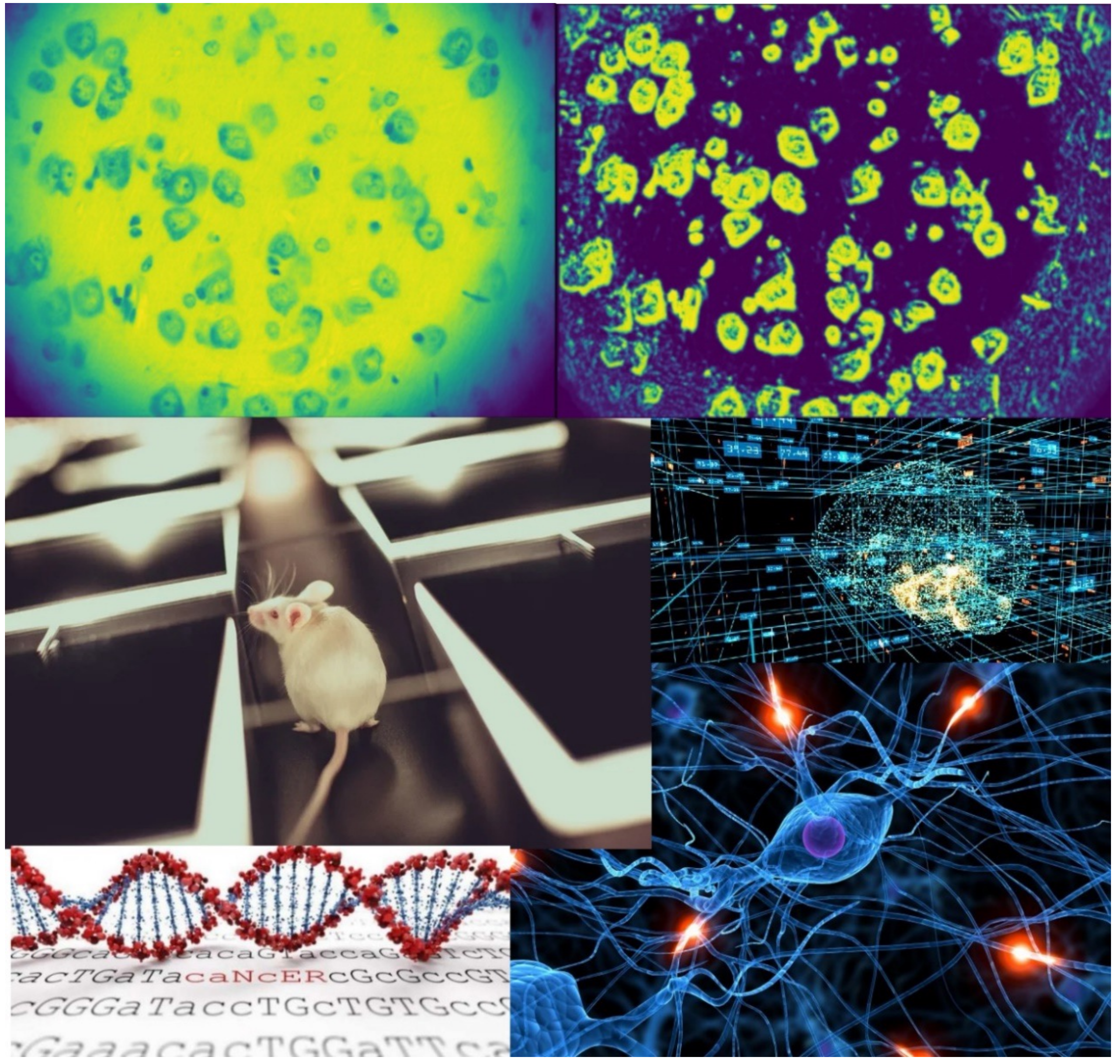






Model experiments show that neural network models are capable of both interpolating tracks and creating an internal model to represent the results in the phase space of the track parameters.

Information System for the tasks of radiation biology



The joint project of MLIT and LRB is focused on creating an Information System (IS) for analyzing behavioral and pathomorphological changes in the central the nervous system when studying the effects of ionizing radiation and other factors on biological objects.

IS is based on:

- computer vision algorithms on the basis of machine and deep learning technologies (ML/DL);
- modern IT solutions for data storage, processing and visualization.

IS will allow one to simplify and accelerate:

- accelerate the processing of experimental data through automation of morphological classification of neural cells;
- data analysis techniques using the latest neural network algorithms based on ML/DL;
- work with experimental data for different research groups;
- systematize experimental data and develop effective methods for preventing and countering the negative effects of ionizing radiation.

The studies are carried out using all the capabilities of the HybriLIT platform

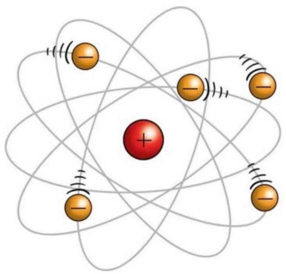
Quantum computing: step by step

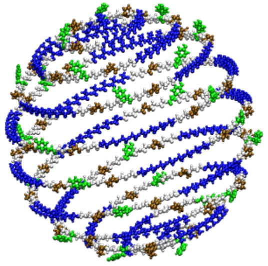
Creation of a physical quantum computer

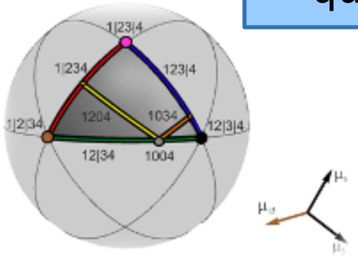
Creation of a simulator of a quantum computer

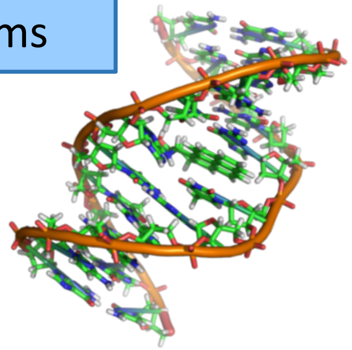
Development of applied quantum algorithms

Simulation of real quantum systems

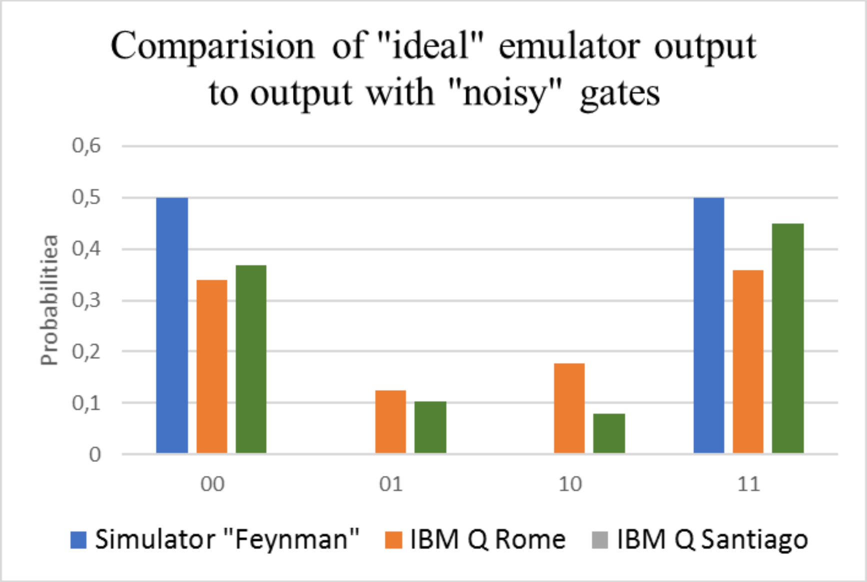






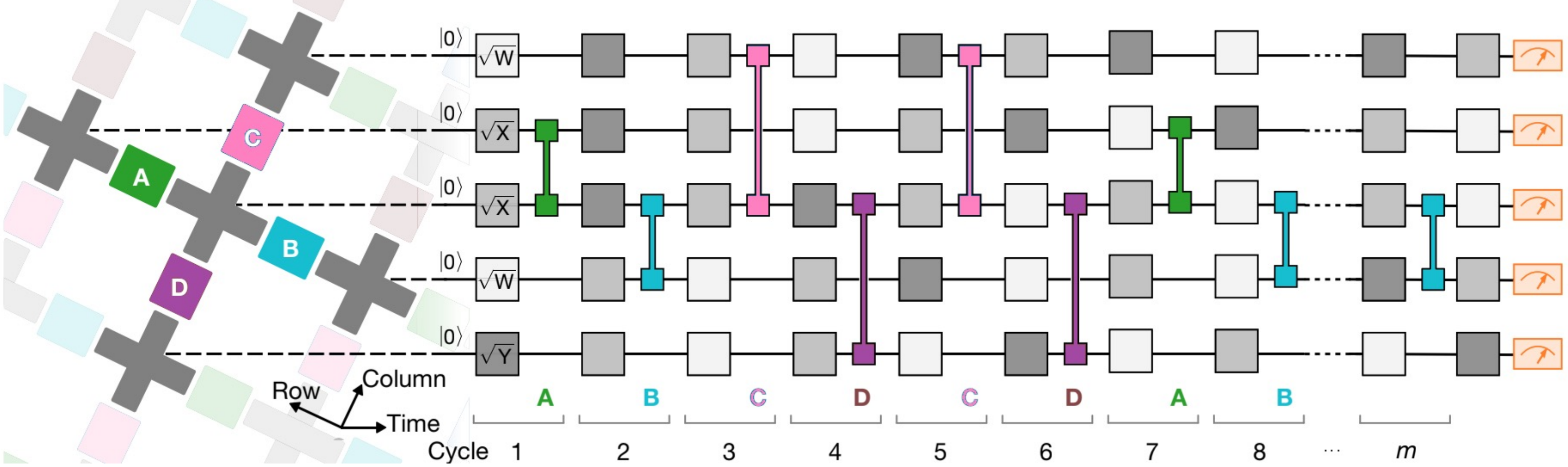


In the framework of the project “Superheavy nuclei and atoms: the limits of the masses of nuclei and the boundaries of D.V. Mendeleev’s Periodic Table” (grant of the Ministry of Education and Science of the Russian Federation No. 075-10-2020-117), an algorithm for quantum teleportation of the Bell states on different IBM five-qubit processors was implemented. To reduce arising errors, several modifications of the original teleportation protocol were proposed.



Gerdt V.P., Kotkova E.A. // Communications in Computer and Information Science, vol 1337, pp 129-143, 2021

Quantum Algorithm Scheme

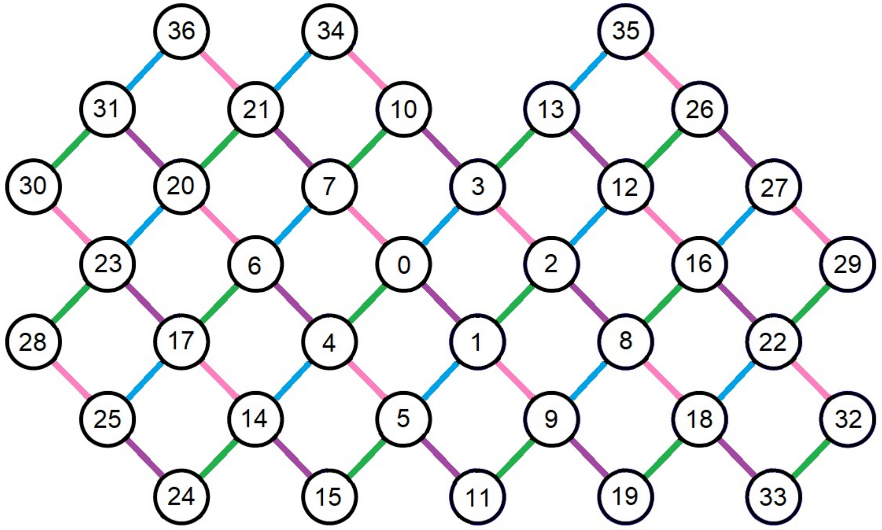


Two-qubit gate :

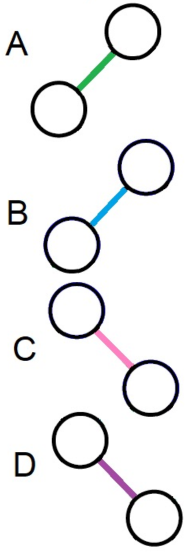
$$fSim\left(\frac{\pi}{2},\frac{\pi}{6}\right)=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & -i & 0\\ 0 & -i & 0 & 0\\ 0 & 0 & 0 & e^{-i\,\pi/6}\end{bmatrix}$$

Quantum algorithm for creating a randomized quantum scheme:

1. m steps of the algorithm consisting of two parts :
   1. Applying one-qubit gates to all qubits, which are randomly selected from the set $\{\sqrt{X},\sqrt{Y},\sqrt{W}\}$, where $W=(X+Y)/\sqrt{2}$.
   2. Depending on the step number, applying two-qubit gates according to the pattern ABCDCDAB: at the first step, gates are applied between qubits with numbers corresponding to the pattern A, on the second, pattern B, etc.
2. Repeat of step 1.1.
3. Measurement of all qubits.



Паттерны:



Computing in a quantum simulator QuEST

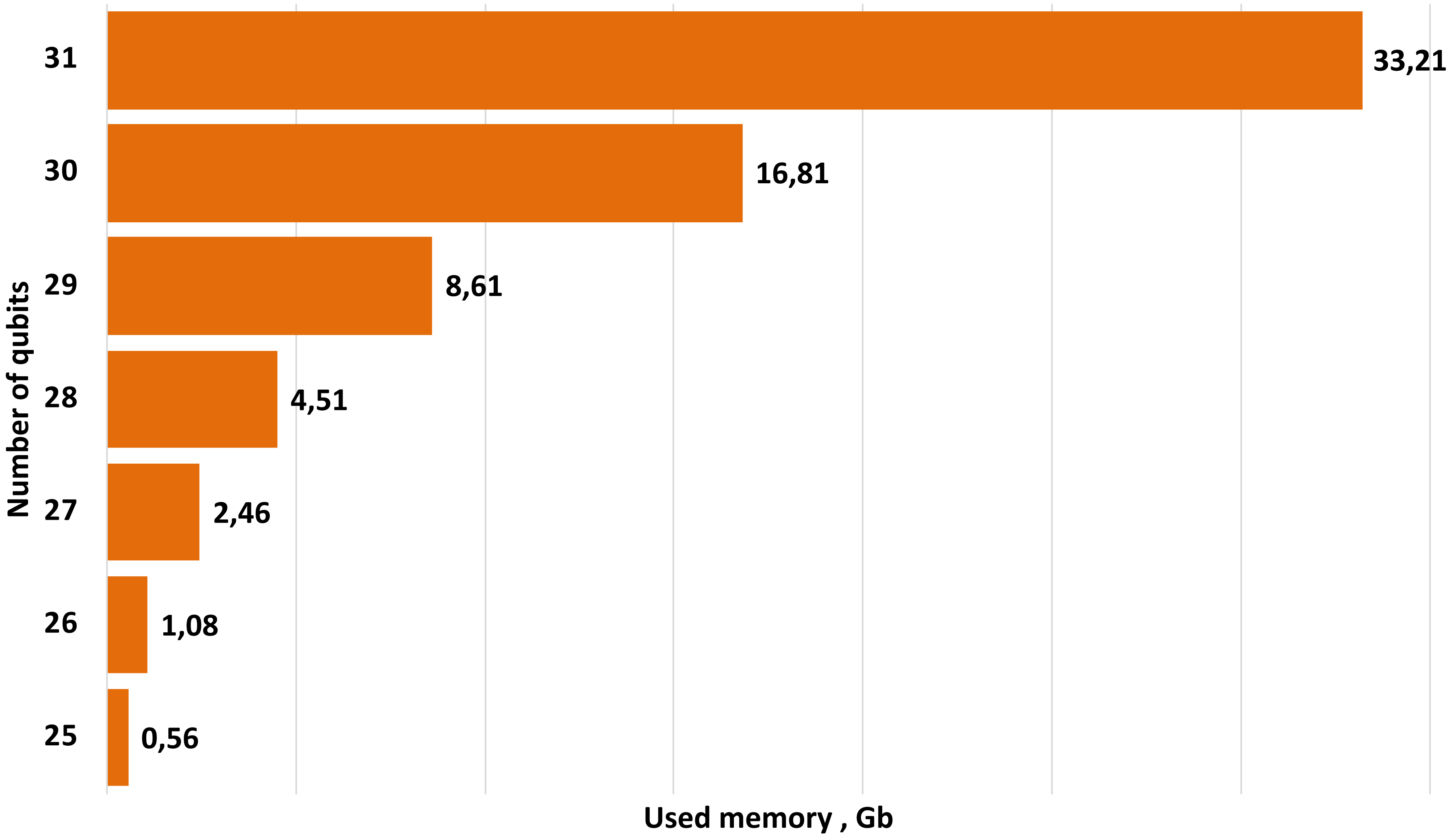
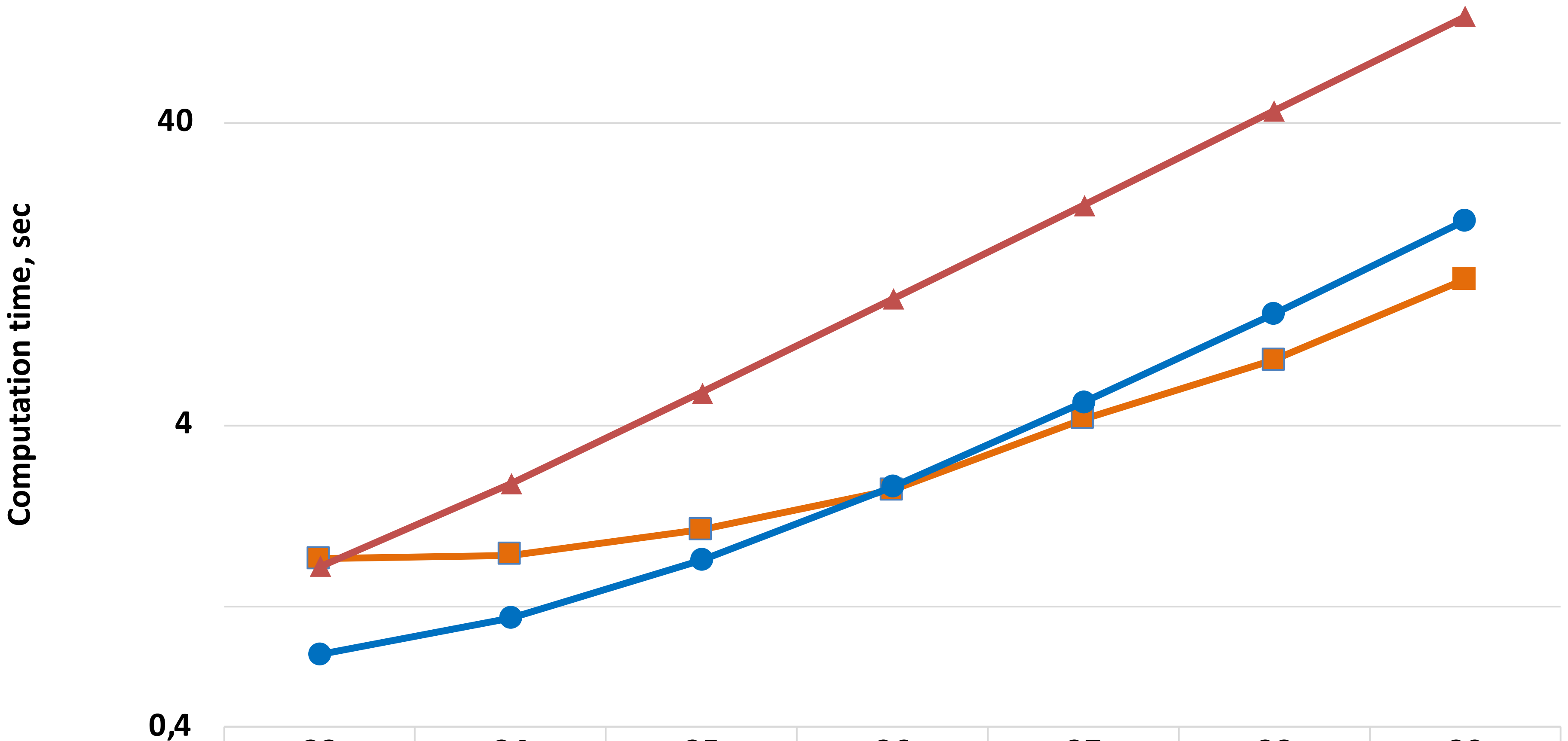


Diagram of dependency
used GPU memory
on the number of qubits

Quantum simulator **QuEST**: https://github.com/QuEST-Kit/QuEST

Computing in a quantum simulator QuEST

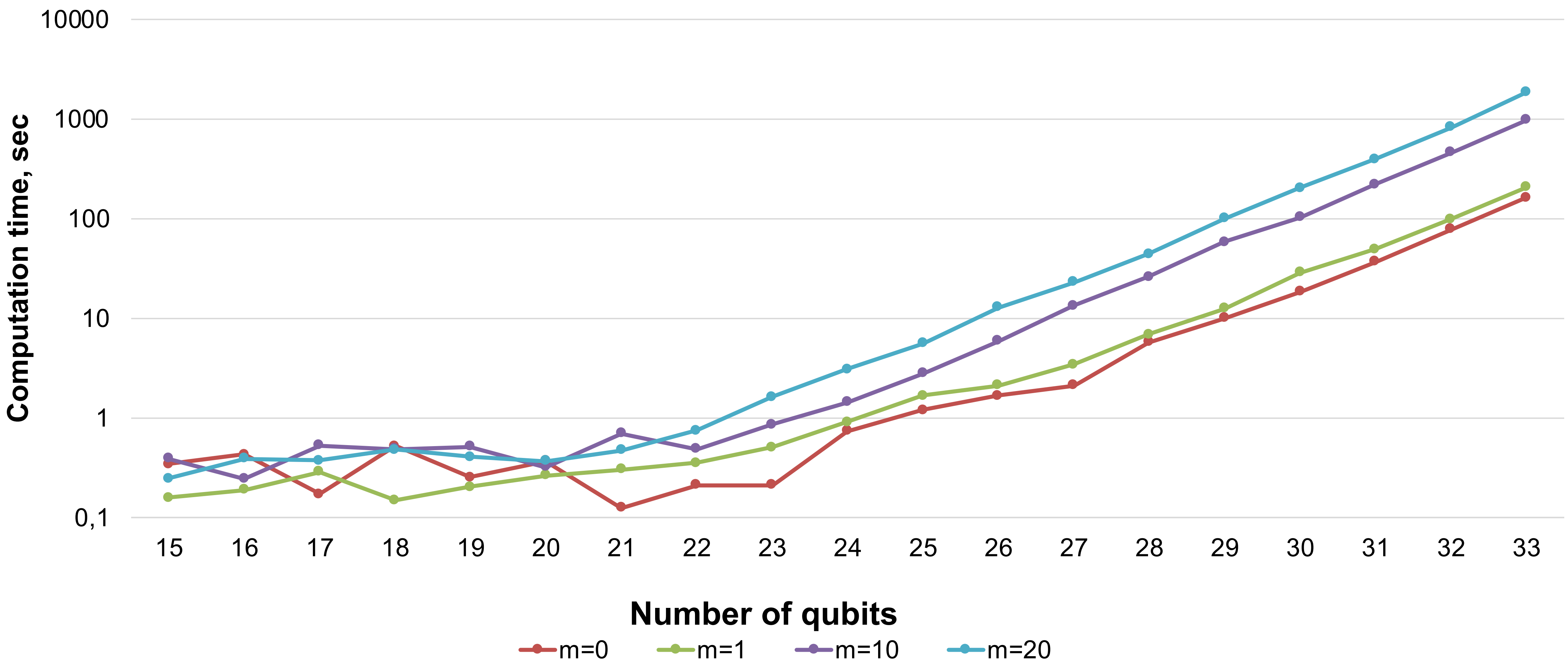


	23	24	25	26	27	28	29
Tesla A100	1,442	1,483	1,797	2,425	4,169	6,546	12,139
Tesla K80	1,354	2,567	5,109	10,424	21,245	43,673	89,965
Tesla V100	0,694	0,916	1,422	2,501	4,735	9,343	18,991

Number of qubits

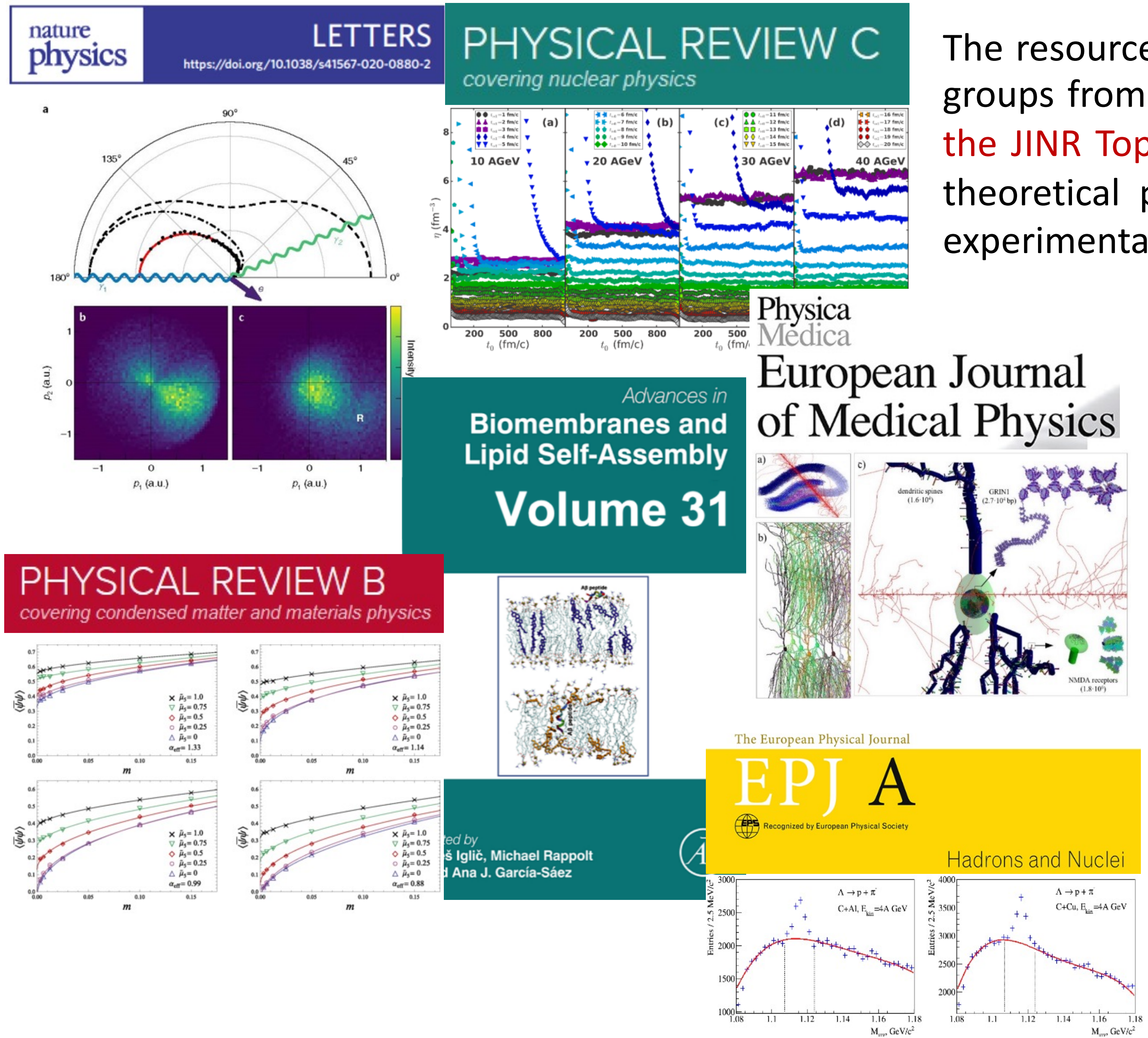
Time dependency calculations on the number of qubits for various GPUs

Computing in a quantum simulator QuEST



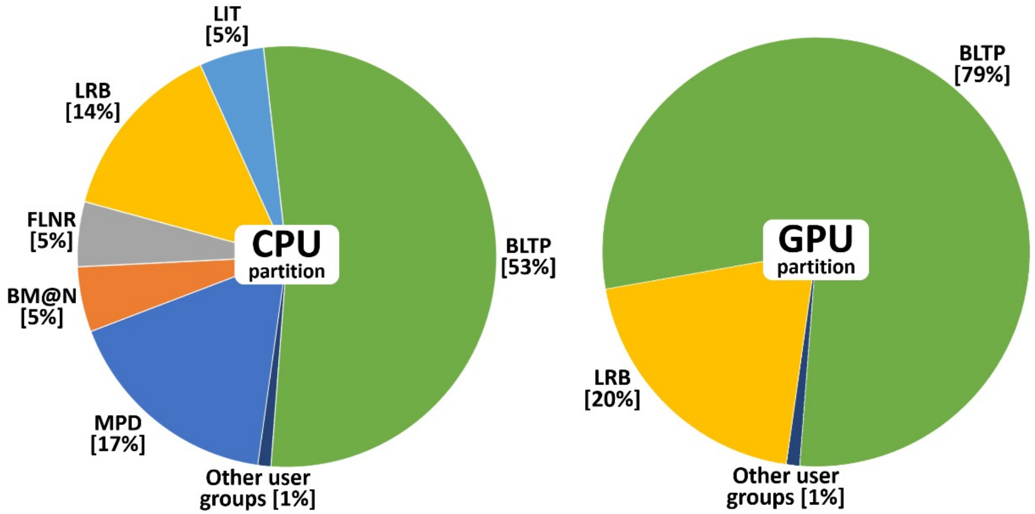
Time dependency on the number of qubits for CPU calculations with OpenMP technology

“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 in 2020 are presented in 65 publications, 6 of them in Q1, 7 in Q2.



Distribution of resources by user groups

Ecosystem for ML/DL tasks and HPC

Development component

VM with JupyterHub
https://jhub.jinr.ru

jupyterhub

VM:
CPU: 24 Cores
RAM: 32 GB

Keras
TensorFlow
scikit learn
NumPy
pandas

Computation component

Servers with NVIDIA Volta & Intel Xeon Gold
https://jhub2.jinr.ru

Dell Volta specs:
GPU: 4x Nvidia Volta V100-SXM2 NVLink 32Gb HBM2
CPU: 2x Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz 20 Cores/40 Threads
RAM: 512 GB DDR4 2666MHz
SSD: 2*240 GB

HPClab component

VM with JupyterHub and SLURM
https://jlabhpc.jinr.ru/

jupyterhub

VM:
CPU: 24 Cores
RAM: 64 GB

Keras
TensorFlow
scikit learn
NumPy
pandas

High **P**erformance **C**omputing

Training videos: **http://hlit.jinr.ru/training-video/**

Thank you for your attention

HybriLIT heterogeneous platform at MLIT JINR:

http://hlit.jinr.ru