D.V. Belyakov[1], A.S. Vorontsov[1], E.A. Druzhinin[2], M.I. Zuev[1], V.V. Korenkov[1], Yu.M. Migal[2], A.A. Moshkin[3], D.V. Podgainy[1], T.A. Strizh[1], O.I. Streltsova[1]

[1] Meshcheryakov Laboratory of Information Technologies, JINR
[2] CJSC “RSC Technologies”
[3] Veksler and Baldin Laboratory of High Energy Physics, JINR

A series of works **“Hyperconverged “Govorun” supercomputer for the implementation of the JINR scientific program”** is being proposed for JINR’s competition in the scientific-methodological and scientific-technical research section.

The creation of the “Govorun” supercomputer at JINR is an essential technological achievement being of great importance for the implementation of the JINR scientific program and international cooperation.

The “Govorun” supercomputer was created in 2018 on top of the experience gained during the operation of the HybriLIT heterogeneous cluster, which is part of the JINR MLIT Multifunctional Information and Computing Complex [1]. HybriLIT has shown its relevance in solving tasks of QCD on lattices, radiation biology, applied research, etc. [2] The continuous growth in the number of users and the expansion of the range of tasks to be solved entailed not only a significant increase in the computing capabilities of the cluster, but the development and implementation of new technologies, which resulted in the creation of a new computing system, the “Govorun” supercomputer. The “Govorun” supercomputer was created as a high-performance, scalable liquid-cooled system with a hyperconverged and software-defined architecture. The current configuration of the “Govorun” supercomputer involves computing modules containing GPU and CPU components, as well as a hierarchical data processing and storage system [3]. The total peak performance of the “Govorun” supercomputer is 1.1 PFlops for double-precision calculations (2.2 PFlops for single-precision calculations) with a read/write speed of 300 GB/sec for the hierarchical data processing and storage system.

The technology of direct liquid cooling of CJSC “RSC Technologies”, which is the leading Russian developer and integrator of the “full cycle” of supercomputer solutions and has a number of its own innovative developments, was chosen for the CPU component of the supercomputer [4, 5]. Thanks to the introduction of these technologies the “Govorun” supercomputer managed to achieve a record density of placement of compute nodes per rack (153 nodes vs 25 nodes for air cooling), and the operation in the “hot water” cooling mode made it possible to use the year-round free cooling mode (24x7x365). In addition to high-energy efficiency, this approach enabled to significantly simplify the infrastructure of the supercomputer center, i.e., the cooling system of the “Govorun” supercomputer was created using only dry cooling towers that cool the liquid using ambient air. Due to liquid cooling, the average annual PUE indicator of the system, reflecting the level of energy efficiency, is less than 1.06. That is, less than 6% of the total electricity consumed is spent on cooling, which is an outstanding result for the HPC industry. The given system is the first system in the world with 100% liquid cooling; all components, namely, compute nodes, network switches and the data storage system, are cooled.

Another technology underlying the “Govorun” supercomputer is a hyperconverged approach to building a computing complex, which enables creating computing environments whose hardware and software configuration is optimized for specific user tasks, without changing the hardware of the compute nodes themselves. Hyperconvergence allows orchestrating computing resources and data storage elements, as well as creating computing systems on demand, taking into account the needs of user applications, with the help of the RSC BasIS software. The notion “orchestration” means the logical disintegration of a compute node into separate components, such as compute nodes, data storage elements (SSDs), with their subsequent integration into the configuration. Thus, computing elements (CPU cores and graphics accelerators) and data storage elements (SSDs) form independent sets of resources (pools). Due to orchestration, the user can allocate for his task the required number and type of compute nodes (including the required number of graphics accelerators), the required volume and type of data storage systems, as well as automatically configure the required software, including parallel file systems. After the task is completed, the compute nodes and storage elements are returned to their corresponding pools and are ready for the next use. This feature allows one to effectively solve user tasks of different types, to enhance the level of confidentiality of working with data and avoid system errors that occur when crossing the resources for different user tasks. The storage-on-demand system implemented on the hyperconverged nodes of the first modification under the management of the Luster file system allowed the “Govorun” supercomputer to take 9th place in the IO500 world rating (June 2018) for HPC storage systems.

Thanks to its hyperconvergence, the “Govorun” supercomputer has a flexible architecture that enables to create software-defined HPC subsystems, which qualitatively distinguishes it from other supercomputers having, as a rule, a “rigid” architecture and designed to effectively solve highly specialized classes of tasks [6].

The experience of operating the first modification of the “Govorun” supercomputer revealed the need not only to enlarge computing resources, which was defined by its demand for solving JINR tasks and the growth in the number of users, but also the need to create tools for working with Big Data, primarily for the NICA megaproject [7, 8]. In this regard, a hierarchical hyperconverged 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, the system is divided into layers, namely, very hot data, the most demanded data, to which it is currently required to provide the fastest access, hot data and warm data. Each layer of the developed data storage system can be used both independently and as part of data processing workflows. At the moment, as a layer of very hot data, the latest DAOS (Distributed Asynchronous Object Storage) technology for Big Data processing, which has shown its promise for deep learning tasks and for the operation of quantum simulators to emulate a larger number of qubits, is being implemented on the “Govorun” supercomputer. For the high-speed data processing and storage system, the “Govorun” supercomputer received the prestigious Russian DC Awards 2020 in “the Best IT Solution for Data Centers” nomination.

The tasks of mass generation and data reconstruction within the NICA MPD experiment actively use the hierarchical data processing and storage system of the “Govorun” supercomputer [9]. At the same time, at different stages of event reconstruction and simulation workflows, there is a need for different access rates to the data; for example, for long-term storage tasks, access speed is not an important factor, however, for reconstruction tasks, speed plays a relevant role. In addition, for a number of MPD tasks, there was a need for a large amount of RAM, which resulted in the introduction of hyperconverged nodes with a large amount of memory in the supercomputer architecture. Thus, methodologically, to ensure all workflows for the tasks of the NICA megaproject, a system that combines both computing architectures of different types and the developed hierarchical data processing and storage system was created on the “Govorun” supercomputer. The computing resources and the hierarchical data processing and storage system of the “Govorun” supercomputer were integrated into a DIRAC-based distributed heterogeneous environment that includes the resources of JINR and its Member States [10]. The experience of using different computing resources of JINR and other MPD collaboration institutes has shown that at present, the use of the “Govorun” supercomputer resources is the most efficient [11].

The implementation of the above technologies on the "Govorun" supercomputer made it possible to perform a number of complex resource-intensive calculations in the field of lattice quantum chromodynamics to study the properties of hadronic matter at high energy density and baryon charge and in the presence of supramaximal electromagnetic fields, to qualitatively increase the efficiency of modeling the dynamics of collisions of relativistic heavy ions, to speed up the process of event generation and reconstruction for conducting experiments within the NICA megaproject implementation, to carry out computations of the radiation safety of JINR experimental facilities and enhance the efficiency of solving applied tasks [12]. The technologies implemented on the "Govorun" supercomputer enabled the development of IT solutions, such as the ML/DL/HPC ecosystem, which provide opportunities not only for solving tasks in the field of machine and deep learning, but also for the convenient organization of calculations and analysis of results. Examples of such solutions are the developed information and computing system for a joint project with BLTP to study theoretical models of Josephson junctions [12] and the information system for a joint project with LRB for processing, analyzing and visualizing data of radiobiological studies [13].

Another area of research, in which the resources of the "Govorun" supercomputer are involved, is the creation of a unified scalable research supercomputer infrastructure based on the National Research Computer Network of Russia (NIKS). At present, in addition to the "Govorun" supercomputer, this infrastructure combines the supercomputers of the Interdepartmental Supercomputer Center of the Russian Academy of Sciences and Peter the Great St. Petersburg Polytechnic University. The created infrastructure allows the participants to enlarge their local computing power, to provide access to the means for storing and processing large data volumes, to distributed data storages (data hubs), as well as to utilize each other's capacities in the case of peak loads. Such an infrastructure is in demand primarily for the tasks of the NICA megascience project. In 2022, the first joint experiment on the use of the unified supercomputer infrastructure for the tasks of the NICA megascience project was successfully completed. In total, 3,000 Monte Carlo data generation and event reconstruction tasks were launched for the MPD experiment. As a result, about 3 million events were generated and reconstructed. The obtained data were transferred to JINR for further processing and physics analysis.

The results obtained using the resources of the "Govorun" supercomputer from the moment it was put into operation in July 2018 to September 2022 are reflected in 204 user publications, two of them in the Nature Physics journal.

Thus, the operation experience of the "Govorun" supercomputer has shown the relevance and effectiveness of using both novel hyperconverged computing architectures and the hierarchical data processing and storage system being part of it. At present, the resources of the "Govorun" spercomputer are used by scientific groups from all the Laboratories of the Institute within 25 themes of the JINR Topical Plan. The number of users of the "Govorun" supercomputer is 323 people, of which 262 are JINR staff members, and 61 are from the Member States. Access to the supercomputer resources is provided only to those users who are directly involved in the implementation of the JINR Topical Plan.

The uppermost results are as follows:

1. A hyperconverged architecture for compute nodes of a supercomputer has been created and implemented for the first time worldwide. The hyperconvergence of compute nodes allows orchestrating computing resources and data storage elements, as well as creating computing systems on demand, using the RSC BasIS software. In addition to increasing the efficiency of solving user tasks of different types, this feature allows one to enhance the level of confidentiality of working with data and avoid system errors that occur when crossing the resources for different user tasks.

2. A hierarchical data processing and storage system, which is a single centrally controlled system divided into several layers of data storage, namely, very hot data, hot data and warm data, has been developed and implemented. The use of such a solution made it possible to formulate and implement the concept of working with Big Data on the "Govorun" supercomputer as the implementation of mapping the main characteristics of Big Data, V^3 (*Volume*, large amounts of data for processing and storage; *Velocity*, the need for high-speed data processing, *Variety*, data of different types) on the hardware and software characteristics of the supercomputer, H^3 (*Heterogeneity*, a set of calculators of different type; *Hierarchy*, multilayered organization of data access; *Hyperconvegence*, flexible organization of data storage systems). The implementation of the hierarchical data processing and storage system enables a significant increase in the efficiency of working with large data arrays, including for the NICA project.
3. The flexible architecture of the "Govorun" supercomputer allows one not only to carry out calculations, but also to use the supercomputer as a research polygon for developing software-hardware and IT solutions for tasks underway at JINR. This feature made it possible to deploy polygons for quantum computing and LRB experimental data processing, to integrate the resources of the "Govorun" supercomputer into a single heterogeneous environment based on the DIRAC platform for the NICA project and utilize its resources to implement the program of runs of data mass modeling within the MPD experiment. It is noteworthy that some tasks for modeling MPD experiment data can only be performed on the resources of the "Govorun" supercomputer.

List of publications:

[1] A. Baginyan, A. Balandin, N. Balashov, A. Dolbilov, A. Gavrish, A. Golunov, N. Gromova, I. Kashunin, V. Korenkov, N. Kutovskiy, V. Mitsyn, I. Pelevanyuk, D. Podgainy, O. Streltsova, T. Strizh, V. Trofimov, A. Vorontsov, N. Voytishin and M. Zuev: "Current Status of the MICC: an Overview" // CEUR Workshop proceedings, 2021, Vol. 3041, pp. 1-8.

[2] Gh. Adam, M. Bashashin, D. Belyakov, M. Kirakosyan, M. Matveev, D. Podgainy, T. Sapozhnikova, O. Streltsova, Sh. Torosyan, M. Vala, L. Valova, A. Vorontsov, T. Zaikina, E. Zemlyanaya and M. Zuev: "IT-ecosystem of the HybriLIT heterogeneous platform for high-performance computing and training of IT-specialists" // CEUR Workshop proceedings, 2018, Vol. 2267, pp. 638-644.

[3] D.V. Podgainy, D.V. Belaykov, A.V. Nechaevsky, O.I. Streltsova, A.V. Vorontsov and M.I. Zuev: "IT Solutions for JINR Tasks on the "GOVORUN" Supercomputer" // CEUR Workshop proceedings, 2021, Vol. 3041, pp. 612-618.

[4] E.A. Druzhinin, A.B. Shmelev, A.A. Moskovsky, V.V. Mironov, A. Semin, "Server Level Liquid Cooling: Do Higher System Temperatures Improve Energy Efficiency?" // Supercomputing frontiers and innovations, 2016, Vol. 3, № 1, pp. 67-73, DOI: 10.14529/jsfi160104

[5] E. Druzhinin, A. Shmelev, A. Moskovsky, Yu. Migal, V. Mironov, A. Semin, "High temperature coolant demonstrated for a computational cluster" // Proc. of 2016 International Conference on High Performance Computing & Simulation (HPCS), DOI: 10.1109/HPCSim.2016.7568418

[6] D. Belyakov, A. Nechaevskiy, I. Pelevanuk, D. Podgainy, A. Stadnik, O. Streltsova, A. Vorontsov, M. Zuev: "Govorun" Supercomputer for JINR Tasks" // CEUR Workshop proceedings, 2020, Vol. 2772, pp. 1-12.

[7] V. Korenkov, A. Dolbilov, V. Mitsyn, I. Kashunin, N. Kutovskiy, D. Podgainy, O. Streltsova, T. Strizh, V. Trofimov, and P. Zrelov: "The JINR distributed computing environment" // EPJ Web of Conferences, 2019, Vol. 214, p. 03009, DOI: https://doi.org/10.1051/epjconf/201921403009

[8] V.V. Korenkov: “Trends and prospects for the development of distributed computing and Big Data analytics to support megascience projects” // Nuclear Physics, 2020, Vol. 83, no. 6, pp. 534-538. (in Russian)
[9] D.V. Belyakov, AG. Dolbilov, A.A. Moshkin, I.S. Pelevanyuk, D.V. Podgainy, O.V. Rogachevsky, O.I. Streltsova and M.I. Zuev: “Using the “Govorun” Supercomputer for the NICA Megaproject” // CEUR Workshop proceedings, 2018, Vol. 2507, pp. 316-320.
[10] N. Kutovskiy, V. Mitsyn, A. Moshkin, I. Pelevanyuk, D. Podgayny, O. Rogachevsky, B. Shchinov, V. Trofimov and A. Tsaregorodtsev: “Integration of Distributed Heterogeneous Computing Resources for the MPD Experiment with DIRAC Interware”// Physics of Particles and Nuclei, 2021, Vol. 52 (4), pp. 835-841, DOI:10.1134/S1063779621040419
[11] A.A. Moshkin, I.S. Pelevanyuk, D.V. Podgainy, O.V. Rogachevsky, O.I. Streltsova and M.I. Zuev: “Approaches, services, and monitoring in a distributed heterogeneous computing environment for the MPD experiment” // Russian Supercomputing Days: Proceedings of the International Conference, 2021, pp. 4-11. DOI: https://doi.org/10.29003/m2454.RussianSCDays2021.
[12] Yu.A. Butenko, M.I. Zuev, M. Cosic, A.V. Nechaevsky, D.V. Podgainy, I.R. Rakhmonov, A.V. Stadnik, O.I. Streltsova. ML/DL/HPC ecosystem of the HybriLIT platform (MLIT JINR): new opportunities for applied research // (in press), 2022. (in Russian)
[13] I.A. Kolesnikova, AV. Nechaevskiy, D.V. Podgainy, A.V. Stadnik, A.I. Streltsov and O.I. Streltsova: “Information System for Radiobiological Studies” // CEUR Workshop proceedings, 2020, Vol. 2743, pp. 1-6.