

Science Innovation Review: Developers of Supercomputers in Russia

Andrey Molyakov

Institute of Information Technologies and Cybersecurity, Russian State University for the Humanities, Moscow, Russia

Email address:

andrei_molyakov@mail.ru

To cite this article:

Andrey Molyakov. Science Innovation Review: Developers of Supercomputers in Russia. *Journal of Electrical and Electronic Engineering*. Special Issue: *Science Innovation*. Vol. 7, No. 5, 2019, pp. 107-112. doi: 10.11648/j.jeeec.20190705.12

Received: August 18, 2019; **Accepted:** September 23, 2019; **Published:** October 14, 2019

Abstract: Author describes developers of supercomputers in Russia. The situation of a significant predominance of the field of science and education in the use of supercomputers in Russia was created artificially. Author believes that it will gradually improve in the natural process of state development. Due to the fact that the limit of miniaturization of the elemental base on the basis of silicon complementary metal-oxide-semiconductor (CMOS) technologies is approaching, i.e. the end of Moore's law is approaching (in the region of 2024, currently 22nm, the expected limit is 5 nm in 2020), in foreign countries, parallel to exaflops projects, work is underway on the promising element-design base of the Post-Moore's era. New application areas may require the creation of cluster supercomputers that differ from most modern ones oriented to high-performance computing, as well as supercomputers that differ from cluster ones, i.e. on a special element base and other special solutions, both in hardware and in software. This, in fact, is happening now in the United States and China. Author shows examples of the development in the world and in Russia of new applied areas for the use of supercomputer resources and these resources themselves. A complete solution to the protection problem is the development of supercomputers with hardware-based security levels for programs and data.

Keywords: Petaflops, Angara, Information Security, Supercomputer, Security Descriptor

1. Introduction

In foreign reviews on supercomputer topics, T-platforms and Russian companies (Russian supercomputers) are mentioned as the main developers. This is true, but in reality the largest suppliers of supercomputers according to official data are really T-platforms, but also IBM, Hewlett Packard [1]. Current open data can be taken from the site of the Top50 supercomputers (<http://top50.supercomputers.ru>). The vast majority of supercomputers (38 out of 50) are manufactured by foreign companies IBM and Hewlett Packard.

The highest result on the peak summer 2011 list was Lomonosov's supercomputer (Moscow State University, development of the T-platform), peak performance was 1.373 TFlops, and on Linpack test it was 674.11 TFlops.

The second result is the Hewlett Packard supercomputer (Interdepartmental Supercomputer Center RAS). Peak performance is 123.65 TFlops and Linpack test has 101.21 TFlops. The Russian supercomputer of the lowest performance on this list is in 48th place, its manufacturer is

IBM, the peak performance is 8.06 TFlops, and on the Linpack test 6.34 TFlops.

If we consider the total computing power attributable to different manufacturers, the situation looks somewhat better, since T-platforms carries out very large projects. However, the share of foreign firms is high, and Russian firms work using foreign technologies. Top50 is an official picture, but this open statistics did not include large supercomputers and medium-sized supercomputers of closed departments. By the way, behind these unknown supercomputers are the respective developers who have a rich history and strong scientific and technical potential. The successes of the young T-platform company are significant, largely determined by the brilliant management and support of government agencies, but we need to more critically evaluate its scientific and technical potential. Is he really great (the company is not even ten years old, there is no scientific school) to consider this company as a monopolist (what it claims to be) that alone can cope with the competition with foreign companies in the Russian market and determine the direction of work in

The Russian Supercomputer industry, the existence of which they are talking about, but even after analyzing the Top50, there are strong doubts about this. For these reasons, in order to objectively assess the situation in the field of supercomputer technologies in Russia and give recommendations for further development, it is necessary to consider not so public developers. There is such material, but in this work, the authors considered its inclusion unnecessary [2].

2. Methods and Trends of Developing Supercomputer Network Centers

Currently, the main users of supercomputers of the highest performance range continue to be those organizations that were the main consumers of high-performance equipment in Soviet times:

- a) Two nuclear centers (Sarov and Snezhinsk), as well as the parent enterprises of Rosatom State Corporation;
- b) Research physical and biological centers (Obninsk, Pushchino, etc.);
- c) Leading institutes of the Russian Academy of Sciences (IPM named after Keldysh of the Russian Academy of Sciences);
- d) The head organizations of the aerospace and shipbuilding complex (CIAM named after Baranov, Central Research Institute named after Krylov and others);
- e) Hydrometeorological center and related structure;
- f) Mining industry - the tasks of geological exploration, optimization of oil and gas production and transportation;
- g) Institutes and facilities of the Ministry of Defense of the Russian Federation (for example, facilities providing air defense and missile defense) and the Ministry of Emergencies;
- h) Special services.

In the past 10 years, leading educational institutions have become leaders in consumers of supercomputer technology.

First of all, it is Moscow State University named after Lomonosov, as well as such universities as: Ulyanovsk, Tomsk, Perm and etc. Groups have appeared in these organizations that perform settlements in the interests of Russian and foreign customers. At the same time, however, foreign application packages are actively used. Often the capacity of their own computing facilities for these groups is not enough, and they use the power of supercomputers from large Russian centers, as well as foreign supercomputers. This means that resources are already scarce. The university environment turned out to be the most active in the development of new technologies for high-performance computing and in conjunction with many other factors (the decline of industry, the great activity in financing universities, not industry and RAS). It was through universities that the modern distribution of computing power in Russia was obtained, which is very different from what is

available in the world [3].

This is a clear bias and when financing in the future, this should be taken into account, to understand the reasons. According to the authors, this is another consequence of the lack of funding for organizations that have traditionally been engaged in computer technology and have historically been associated with industrial and other fields. Young and active companies have mastered the new market of university applicants - in the end, they got such a picture. Such a detailed review of the situation in the field of supercomputer resources in Russia was necessary to develop a policy to protect supercomputers and supercomputer centers from cyber-attacks. It turns out that the main share of supercomputers is foreign-made, the rest are created using the foreign component base (elemental base and constructs, foreign system and embedded software), most of the capacities are at universities.

Thus, according to the state of resources and their exploitation, there is no acute problem of national cybersecurity support so far, this can concern only certain closed centers. However, as already noted, leading Russian research institutes and universities in Moscow and St. Petersburg have already been attacked by Chinese hackers with the aim of collecting information for now, which went unnoticed, as a random check showed [4].

2.1. Conception of Creating Supercomputers

Several groups of organizations that compete with each other can be distinguished, although they are tied to their traditional different application areas.

The most powerful petaflops supercomputer was recently developed at the Federal State Unitary Enterprise VNIIEF (Sarov). This organization is a real leader in the segment of supercomputers in the highest performance range. It is obvious that the Government has made a bet on this organization in the matter of creating supercomputers at the levels of 10, 100 and 1000 Pflops. This is an enterprise of the nuclear weapons complex and the nuclear industry, which explains everything. Rosatom State Corporation prepared the Exaflops Technologies Conception in May 2011, several organizations participated in its preparation (Department for the Development of the Scientific and Production Unit of the Nuclear Weapon Complex, Department of Development of the Ministry of Education and Science of Russia, Research Institute Kvant, IPM named after MV Keldysh RAS, IPS named after AK Ailamazyan RAS, NIISI RAS, NIIMM named after N. G. Chebotarev at KSU, Nizhny Novgorod State University named after Lobachevsky, Moscow State University named after Lomonosov, Moscow Aviation Institute, MSTU named after Bauman and VNIIEF).

T-platform is another group of developers, they recently signed a Memorandum of Intent on cooperation in the field of creating supercomputers of a new generation of exaflops level.

IPS RAS was the lead organization of several supercomputer programs of the Union State, and a notable modern development is the SKIF-AURORA supercomputer

(3rd place in the current Top50 list). IPS RAS also took upon itself a great organizational burden for the formation of the National Supercomputer Platform (jointly with Moscow State University). Research Institute Kvant is organization with a long history of several decades of joint works.

Recently, VT-Consulting, E-tronic, SPbSPU, Federal State Unitary Enterprise VO Vneshtekhnika, the Center for Engineering Development of the Physical Faculty of Moscow State University joined this group. There has always been a collaboration with IPS them. A. K. Ailamazyan of the Russian Academy of Sciences and NICEVT. Examples of large commercial companies are Craftway, SKIF, Open Technologies.

There are a number of small companies having, first of all, good relations with foreign suppliers of commercially available components and small-sized groups of engineers for adjusting equipment and software, these are Niagara, Krok.

There is another organization that, since 1999, by the end of the 2000s, has significantly restored its position as the main developer of computing systems of the Soviet era - this is NICEVT. The first two 32-processor clusters of this program at the end of 1999 and the beginning of 2000 were created at NICEVT. This program, in essence, spawned a T-platform company, providing it with proven cluster assembly technologies based on the SCI communications network.

Further, in the NICEVT program of the Union State program TRIADA was developed, which was carried out in the Union State in the mid 2000s, and the NICEVT was the lead contractor from Russia. At the same time, research work was started at NICEVT, which then transferred to the development project Angara, a supercomputer with promising architecture, which was the only response in Russia to the DARPA HPCS program. This topic was supported by the Federal Security Service (FSS) and industry, but in its scale it required not a departmental, but a federal level of implementation. Despite the positive decisions on this project, it was not supported, because there was strong opposition from other enterprise-developers of supercomputers on cluster technologies. This influenced a certain curtailment of work in 2009, which was reinforced by the resignation of the Director General (summer 2009), thanks to which this organization was restored in 10 previous years. Most of the active part of the team working on the project was lost. Nevertheless, a significant part of the staff was able to be transferred to Kvant. In the same year, NICEVT was transferred to the Vega (Radio Concern), which, apparently, influenced the change of subject and the weakening of work in the supercomputer direction.

The concept of creating exaflops systems was officially presented by a group of organizations working under the leadership of Rosatom State Corporation. It can be divided into two main areas:

K-N1 (E) is an evolutionary direction, it uses foreign commercially available element base, hybrid architectures, it is basic for Rosatom State Corporation, it is possible to develop its own communication network;

K-H2 (I) is an innovative direction, focused on the use of its own elemental base, which must be created, the proposal of the NIISI RAS.

The total cost of implementing the Concept (H1 + H2) is (without capital construction) 42,382 million rubles. In particular:

- a) The cost of creating exaflops computing technology is 19872 million rubles, of which 9000 million rubles are components and special equipment;
- b) The cost of creating a domestic electronic component base is 25510 million rubles, of which 24902.5 million rubles are components and special equipment.

2.2. Trends in the Development of Russian Supercomputing

The situation of a significant predominance of the field of science and education in the use of supercomputers in Russia was created, in our opinion, artificially. We believe that it will gradually improve in the natural process of state development. New application areas may require the creation of cluster supercomputers that differ from most modern ones oriented to high-performance computing, as well as supercomputers that differ from cluster ones, i.e. on a special element base and other special solutions, both in hardware and in software. This, in fact, is happening now in the United States and China. We give examples of the development in the world and in Russia of new applied areas for the use of supercomputer resources and these resources themselves.

Example 1. The development of the use of supercomputers in the financial sector. Judging by the Top50 list, the share of supercomputers in this area has grown from 18% to 30%. One of the main functions for this area is working with large databases, large network structures.

Example 2. Began increasing the use of supercomputers in government. This is an area of information and information management systems, it is again databases and network applications. This area of application is important not only for the civilian sector, but also for ensuring national security and defense capability (sixth generation wars - network-centric wars). In these applications, conventional cluster-oriented computing-only systems are unacceptable, because, according to experts, these information-type tasks are characterized by processing data volumes from dozens of petabytes to exabytes, intensive irregular work with graph-type data, and work with sets of input-output data streams. As a matter of fact, the new DARPA UHPC program (previously it was the DARPA HPCS program) was aimed at developing a new generation of supercomputers to solve these problems. In the USA and China, hybrid massive multi-thread supercomputers with globally addressable memory are being developed for this area (they were previously discussed in the sections on the USA and China). In Russia, the Angara project had such an orientation. In China, this class of supercomputers is now called HTC computers (High Throughput Computers), the most important area that they have identified along with HPC computers (High Performance Computers) and HRC (High-Reliability Computers) [5].

Example 3. Application areas related to molecular modeling, biotechnology, pharmaceuticals, living systems, healthcare and medicine began to develop. These applications require both a large amount of computation and the ability to work with large amounts of data.

Traditional areas of computing applications have become complex, multiphysical. It also requires the ability to work with huge amounts of data, and in different modes of spatio-temporal localization. Before specialized supercomputers appeared for solving such problems, solutions began to appear using a commercially available element base, i.e. on the line of cluster-type supercomputers. This can be seen

from the results of the analysis of supercomputers used in the USA, it began to penetrate into domestic practice. For example, Table 1 shows the specification of the data center of the Telecommunication Center of St. Petersburg State University. This center was created by HP specialists in 2010. It can be seen that it consists of cluster segments composed of nodes of various types. In the light of the considered problem of ensuring cybersecurity, which is strongly associated with the ability to process large data arrays, we draw attention to the presence of a segment with *heavy* 8-socket nodes.

Table 1. Specification of the Computing Center of St. Petersburg State University."

Computer complex SPbSU
Software:
System virtualization support: VMware 5;
Windows, Linux, special software on the research areas;
Compilers: Intel, Lahey, NAG, PGI, GCC.
Libraries: IMKL, MPI, Platform MPI.
Cluster of virtual machines:
60 hosts, 120 CPU, 720 cores, 5.76TB RAM
BL460-G6/BL460c-G7, 8.61/7.40 TFlops
Multiprocessing systems with common memory: 3 hosts, 24 CPU, 192 cores, 3TB RAM (2TB/512GB), 1.7 TFlops
Hybrid clusters: 24 hosts, 48 CPU/112 GPU, 288/1536 cores, 2.3TB RAM, 59.6/17.66 TFlops, Compute cluster:
48 hosts, 96 processors, 384 cores, 768GB RAM, 7,68TB HDD, 20Gbps IB, 3.07 TFlops
Compute cluster:
48 hosts, 96 processors, 384 cores, 768GB RAM, 7,68TB HDD, 20Gbps IB, 3.07 TFlops
Storage system:
4 × HP X9300, 2 × MDS6000 HP 4800 62
HDD 480TB, HDD 63TB
Backup system:
HP D2D4106i Backup System
Common network (LAN):
QDR Infiniband/Gigabit Ethernet/Fast Ethernet

The world practice is such that the networks of supercomputers should not be just a common field of supercomputer resources to which users are connected, they should be infrastructures with the administration and the program of events. Examples of such foreign networks are the network of US military supercomputer centers funded by the HPCMP program, the network of university centers under the XSEDE program. In Russia, the creation of this type of network was planned in the mid-2000s when developing the concept of the TRIADA Union State program, but the Russian Ministry of Finance did not miss this part. They returned to this in a slightly modified form in the SKIF-

GRID program. Now, as far as the authors know, they plan to implement this in the new ORBISS program. The role of ensuring cybersecurity in such networks will increase, especially since, as already mentioned, attacks on several domestic supercomputers have already been recorded, while others became known through other channels.

3. Result

A general comparison of SC implementation results is presented in Table 2.

Table 2. Matching Productivity Growth Russian and foreign supercomputers.

Countries	Years of creation				
	2011	2012	2013	2014-2015	2016-2020
USA	10	20	35	100	1000
Japan	8	-	-	100	
China	5-10	20	35	100	1000
France	3	7	11	30	
Russia					
RAS ⁽³⁾	0,14	0,3	8	20	
ROSATOM ⁽³⁾	1	3	5	10 (2014)	100 (2017) 1000 (2020)
NIISI RAS	-	-	0.5 ⁽³⁾	10 ⁽¹⁾ (2015)	100 ⁽²⁾ (2017) 500 (2020)
T-Platforms	0.5	1.7	50	200 (2015-2016)	1000

- 1) It is planned to use its own microprocessor with 16 superscalar cores and 128 streaming (clock frequency 1 GHz, peak performance 1 TFlops), the main module - 6-processor, with 256 GB of memory, will also have its own communication network.
- 2) It is believed that a domestic supercomputer optimized for a fixed class of tasks with a peak performance of 100 Pflops with an efficiency of 70-80% will be able to solve problems of the exaflops class, that is, tasks available to a non-specialized supercomputer with a peak performance of 1 Exaflops.
- 3) A commercially available element base is used, they are guided by hybrid architectures - superscalar architectures with coprocessors - accelerators.

In Russia, the Concept of exaflops technologies has just been formulated (May 2011), large-scale targeted financing has not yet been opened, but financing of some small research projects has already begun. The number of publications on this topic in Russia is small. Judging by open data, the gap between Russia and the United States in the class of tasks with good spatial-temporal localization of memory accesses is about 10 times (Linpack test, Top500 rating), and in the class of tasks with poor spatial-temporal localization, i.e. intensive irregular work with memory or Data intensive tasks (DIS - tasks) – at least 100 times (BFS test, Graph500 rating, Random Access test, HPC Challenge rating). DIS-tasks are typical for the field of robotics and artificial intelligence, information tasks of collecting and processing information, analysis and management decisions typical of situational centers, an increasing number of critical technical tasks are also acquiring features of DIS-tasks.

Since tasks from the military field and ensuring national security are the most pronounced DIS-tasks, projects to create specialized exaflops military supercomputers that have massive multi-thread projects were launched in the USA, China and Japan, are based on development of globally addressable memory (tens of petabytes). They should be put on combat duty no later than 2018-2019. These supercomputers should be at least 100 times superior to conventional universal supercomputers, comparable in size, cost and energy consumption. The authors are not aware of the development of this type of supercomputer in Russia; before, such a supercomputer was developed within the framework of the Angara project, but it did not receive funding at the federal level and in 2010 was actually phased out.

Due to the fact that the limit of miniaturization of the elemental base on the basis of silicon CMOS technologies is approaching, i.e. the end of Moore's law is approaching (in the region of 2024, currently 22nm, the expected limit is 5nm in 2020), in foreign countries, parallel to exaflops projects, work is underway on the promising element-design base of the Post-Moore's era.

The implementation of exaflops projects and projects on the promising new element design base, for example, in the USA, is aimed at obtaining the following results:

1. In the civilian sector, an evolutionary exaflops supercomputer is expected to appear in 2018-2020 (for a narrow range of tasks, low energy efficiency), and an innovative supercomputer is expected after 2022 (for a wide range of tasks, energy efficient).

2. Specialized military supercomputers of exaflops performance level will be put on combat duty no later than 2018-2019. The using of a promising new element base and new principles for building supercomputers allows us to plan (discussed in the Japanese-American expert environment) the creation of specialized military supercomputers of the zettaflops level around 2020, and the yottaflops level - in 2024, moreover, the consumption of the latter is expected to reach 15MW.

4. Conclusion

The main threats to Russian supercomputers, according to the authors, come from the used imported element base, imported server boards, as well as software, especially virtualization and cloud computing:

- a) As noted, it is known about the work on embedding bookmarks in the GPU in China and Taiwan (TSMC), as well as the United States. The presence of such tabs on behavioral information about the operation of the GPU is difficult to detect because of its huge performance. When programming such bookmarks, they began to use the means of converting symbolic information that developed in the 80s. In particular, there is information on the study of the Russian language Refal and the American language Ambit in the framework of the Stork project in this regard in China.
- b) It is known about bookmarks already found at the BIOS and BMC level in imported server boards, in particular, of Chinese manufacture. Virtualization and cloud computing tools allow to legally control the computing process and legalize some performance loss, which, in fact, can occur due to the operation of malicious programs.
- c) The most terrible of these threats is the use of imported element base in supercomputers, in particular, GPU accelerators, and in the future - MIC.

Neither isolation from the outside world in the bunker nor isolation from global networks will save the supercomputers from the mentioned internal threats, since the danger lurks within themselves. As a partial solution, at least a multilevel protection of supercomputers is required, including multilevel control of their functioning. Analysis of the huge amounts of information coming from such a control system requires the use of special hardware and software technologies, including even specialized supercomputers. The technologies of mobile agents are extremely interesting here [6].

For such tasks, special massive multit-thread supercomputers are developed in which the well-known problem of large delays in executing memory accesses (the

“memory wall problem”) is solved by putting the work of applications and equipment in such a mode that high memory bandwidth can be used, while creating The “illusion” of performing calls with small delays. In the USA and China, such supercomputers and a special element base for them are being developed. Such developments are not conducted in Russia, but we hope that they will be restored, taking into account the new modern requirements. There are also more affordable ways of creating cluster supercomputers close to such special supercomputers on a commercially available element base and, possibly, using fragments on the FPGA [7-11].

Such approaches are currently being applied in the United States, due to economic considerations - custom-made massive multi-thread supercomputers are still expensive and are intended only for strategic applications in intelligence and military affairs. A complete solution to the protection problem is the development of supercomputers with hardware-based security levels for programs and data. Currently, four levels of protection are implemented (user, OS executor, OS kernel, bootstrap program). For example, this is done in massive multi-thread supercomputers Cray XMT (USA) and CT-2 (China), designed to work in systems for analyzing large volumes of data and for active work with data from the network (applications of intelligence and military centers, business, science and sociology). Already considered options for 8-level protection [12-15].

References

- [1] James Bamford. The NSA Is Building the Country’s Biggest Spy Center (Watch What You Say). <http://www.infiltratednation.com/2012/03/nsa-is-building-country-s-biggest-spy.html>.
- [2] O’Melia Sean, Elbirt Adam J., Enhancing the performance of symmetric-key cryptography via instruction set extensions, IEEE Transactions on Very Large Scale Integration Systems, Vol. 18, No. 11, pp. 1505–1518, 2010.
- [3] Jinjiang Yang, Peng Cao, Jun Yang. Reconfigurable architecture with high area efficiency for block cipher algorithms. Journal of Southeast University, Sept 2016, Vol. 46, No 5, pp. 939-944.
- [4] Shahbazi Karim, Eshghi Mohammad, Mirzaee Reza Faghih. Design and Implementation of ASIP-based cryptography processor for AES, IDEA, and MD5. Engineering Science and Technology, an International Journal, 20, 2017, 1308-1317.
- [5] Baoning Zhan, Wei Ge, Zhen Wang. A distributed cross-domain register file for reconfigurable cryptographic processor. Journal of Southeast University, Sept 2017, Vol. 33, No 3, pp 260-265.
- [6] Balfour J. et al. An Energy-Efficient Processor Architecture for Embedded Systems. IEEE Computer Architecture Letters, Vol. 7, No 1, January-June 2008, pp. 29-32.
- [7] Nowatzki Tony, Gangadhar Vinay, Sankaralingam Karthikeyan, Wright Greg. Pushing the Limits of Accelerator Efficiency While Retaining Programmability, IEEE High performance computer architecture conference, 2016, 13 pp.
- [8] Molyakov, A. S. New Multilevel Architecture of Secured Supercomputers/A. S. Molyakov//Current Trends in Computer Sciences & Applications 1 (3) – 2019. – PP. 57-59. – ISSN: 2643-6744 – <https://lupinepublishers.com/computer-science-journal/special-issue/CTCSA.MS.ID.000112.pdf>. – DOI: 10.32474/CTCSA.2019.01.000112.
- [9] Molyakov, A. S. Technological Methods Analysis in the Field of Exaflops Supercomputers Development Approaching/A. S. Molyakov, L. K. Eisymont//Global Journal of Computer Science and Technology: Information & Technology. – 2017. – № 1 (17). – PP. 37-44.
- [10] Molyakov, A. S. A Prototype Computer with Non-von Neumann Architecture Based on Strategic Domestic J7 Microprocessor/A. S. Molyakov//Automatic Control and Computer Sciences. – 2016. – № 50 (8). – PP. 682-686.
- [11] Molyakov, A. S. Token Scanning as a New Scientific Approach in the Creation of Protected Systems: A New Generation OS MICROTEK/A. S. Molyakov//Automatic Control and Computer Sciences. – 2016. – № 50 (8). – PP. 687-692.
- [12] Molyakov, A. S. Model of hidden IT security threats in the cloud computing environment/A. S. Molyakov, V. S. Zaborovsky, A. A. Lukashin//Automatic Control and Computer Sciences. – 2015. – № 49 (8). – PP. 741-744.
- [13] Herr A. Y. et al. An 8-bit carry look-ahead adder with 150ps latency and sub-microwatt power dissipation at 10GHz. arXiv: 1212. 2994v1 [quant-ph] Dec 2012. 6 pp.
- [14] Herr A. Rapid Single Flux Quantum Logic. Northrop Grumman. March 2012. 23 slides.
- [15] Frank K. Gurkaynak, Kris Gaj, Beat Muheim, Ekawat Homsirikamol, Christoph Keller, Marcin Rogawski, Hubert Kaeslin, Jens-Peter Kaps. Lessons Learned from Designing a 65nm ASIC for Third Round SHA-3 Candidates. ETH Zurich - George Mason University, 22-23 March 2012, 65 slides.