Using the resources of the Supercomputer Center of Voronezh State University in learning processes and scientific researches

S.V. Borzunov, S.D. Kurgalin

Voronezh State University

The work describes the supercomputer resources of Voronezh State University and their use in training of IT-specialists and for scientific and technical calculations.

Keywords: supercomputer, parallel computing, software, learning and methodological support of the IT-specialists' training, training courses, scientific researches

Solution of topical scientific, technical and applied problems, including simulation of physical, chemical, biological and other complex processes and objects, forecasting in economy and financial sector, management of large industrial and transport complexes etc., necessarily require high-performance computer systems employing the parallel programming methods. Such problems cannot be solved on conventional sequential computers because of a too long program execution time and a limited size of the available memory. This is why the availability in a large university of a high-performance computing system – supercomputer – is undoubtedly the essential condition for organization of scientific researches of fundamental and applied nature, as well as for maintaining the learning process for high-quality training of the graduates strongly sought-for on the labor market.

In 2016, Voronezh State University (VSU) launched a new supercomputer with performance of 39 Tflops, according to LINPACK test – 28 Tflops. It consists of 10 nodes each having two 12-core processors (the total of 240 processor cores), each possessing 128 Gbytes of random-access memory and a 256-Gbyte SSD. 7 nodes contain 2 Intel Xeon Phi accelerators each, and the remaining 3 contain 2 Tesla accelerators each. All the nodes are combined into 40 Gbit/s InfiniBand high-speed network. Apart from this, it includes a master node that is also a file storage server. The latter has two 6-core processors, 64 Gbytes of RAM and a 48-Tbyte disk subsystem.

The computing equipment development in VSU associated with the use of the parallel computing and supercomputer technologies was initiated in 2002, when at the Digital Technologies Department of the Computer Science Faculty of VSU was installed the first in the region 20-processor high-performance cluster with the peak performance of 28 Gflops. At that time, the computer cluster in its parameters ranked among Top50 cluster installations in Russia and the CIS.

In the same year 2002, the Computer Science Faculty started training of specialists in parallel computing, which has been continuously performed since then.

Over the next 5 years, the cluster's performance had increased several times, and VSU Supercomputer Center was established. In 2010, on the basis of the achieved level of supercomputer developments, VSU was admitted to the Supercomputer Consortium of Universities of Russia and has been its permanent member since then.

Today the renovated Supercomputer Center (SCC) includes two computer clusters:

- the main – the new supercomputer;

- the second computer cluster used for training purposes.

Presently, the Center's software includes:

- set of compilers GCC and Intel, including C, C++ and Fortran compilers;

- program parallel start tools Intel®MPI;

- program for computing the structure and properties of molecular systems Gaussian 09;

- compiler Python 2.7 with packages NumPy, SciPy, Scikit-learn, Mathplotlib etc.

Software is also used: Abinit, ASE/GPAW, Quantum Espresso, ELK, Exciting, Siesta, Gromacs, and for quantum-chemical calculations – libraries of exchange-correlation functional libxc and libvdwxc, programs for analysis of charge distribution and electronic density Bader, Wannier90, which have support for Python. The share of open source software distributed under the GPL and its derivatives is about 90%.

The computing system runs under OS Linux CentOS 7, which allows installing a wide range of academic and application programs. The supercomputer resources are accessed remotely through the VSU network or the Internet.

When selecting the hardware and system software, account was taken of the system's purpose (scientific and technical developments and training process), medium-term material and technical up-to-dateness, serviceability, and other parameters. Since many computing systems from the top lines in the list of the world's most high-performance supercomputers Top500 belong to the hybrid type, i.e. combine both classic central processing units (CPU) and graphic subsystems (GPU), it was decided to provide the supercomputer with NVIDIA GPU accelerator based on Tesla architecture, as well as Intel Xeon Phi accelerators.

A special middleware is used for distribution of the resources of the cluster and supercomputer systems. On VSU supercomputer is installed one of the most popular representatives of such software – Slurm Workload Manager – the task scheduler for the clusters run under the Linux operating system, oriented to working with modern high-performance computer systems. The basic functions of Slurm are: organization of the user access to the cluster's computing resources in accordance with a certain resource allocation policy; execution of the parallel tasks on the system nodes dedicated for work; formation and support of the user task queue.

The supercomputer is actively applied in the educational process of the Computer Science Faculty for training bachelors and masters in such fields as «Information Systems and Technologies» and «Mathematics and Computer Sciences», as well as in the educational process of the Faculty of Physics and the Faculty of Applied Mathematics, Informatics and Mechanics.

On the bachelor program level of field «Mathematics and Computer Sciences», group course «Parallel programming» and optional course «Parallel computing technology» have presently been taught.

Let us consider the content of the course «Parallel programming». The first lecture offers an overview of the basic parallel programming means, classification of computing system architecture, and examples of modern supercomputers. Several following lessons are devoted to the basic parallel computation laws. Within the PRAM model are discussed the quantitative characteristics of efficiency of parallel algorithms. Here, the notion of «operation–operand» graph is introduced and Bernstein's conditions are formulated. The next lectures contain the description of widely used parallel programming technologies. The shared memory computing system programming environment OpenMP is discussed, and the computing model used when working with this environment. By the example of the OpenMP technology, implementations of the basic parallel algorithms are demonstrated, including the algorithm for array element summation, data sorting, basic matrix operations, and Monte Carlo method.

The concluding part of the course is devoted to the description of the distributed memory system programming technology, MPI. The essential functions of MPI are discussed, the operations of pairwise message exchange in blocking and non-blocking modes, and collective data passing operations. The C algorithmic language is used. As convenient teaching aids are used both well-known textbooks [1,2] and specially developed teaching aids [3-5].

The laboratory workshop on «Parallel programming» includes accomplishing by the students of 6-7 laboratory works within a semester. The workshop methodological materials were presented in 2017 at the International Conference-Competition «Innovative Information and Pedagogical Technologies in IT Education» (Moscow, MSU) and awarded a laureate diploma for work «Laboratory workshop on «Parallel programming».

On the master program level of field «Mathematics and computer sciences», the course «Parallel and GRID technologies» is taught. For the master's degree students, an optional course «Parallel computing on graphic processors» is provided. The curriculum of the bachelors in field «Information systems and Technologies» includes an elective discipline «Parallel data processing algorithms».

Let us enumerate the basic objectives of the courses associated with the supercomputer technologies:

- familiarizing the students with the basic trends in the development of high-performance computing systems;

- giving an overview of the parallel programming means;

- discussing the parallel programming methods using the message passing interface;

- studying the parallel program functioning models;

- obtaining practical skills in applying the technology of distributed computations and data processing.

Implementation of the said objectives helps establishing general professional and professional competences. For field «Mathematics and computer sciences», these competences are formed as follows [6]:

- ability to use the fundamental knowledge in the sphere of mathematical analysis, complex and functional analysis, algebra, analytical geometry, differential geometry and topology, differential equations, discrete mathematics and mathematical logic, probability theory, mathematical statistics and random processes, numerical methods and theoretical mechanics in the future professional activity;

- ability to accomplish standard professional activity tasks based on the informational and bibliographic culture using the information and communication technologies and taking into account the basic information security requirements ;

- ability to perform independent research and development work;

- ability to find, analyze, programmatically implement and practically use the mathematical algorithms, including those involving the modern computing systems;

- ability to use the mathematical and algorithmic simulation methods when solving theoretical and applied problems;

- ability to transfer the results of the accomplished physico-mathematical and applied researches in the form of specific recommendations expressed in terms of the domain under study.

The supercomputer technologies find application in training foreign students. In particular, in 2017, an Iraqi student defended the master's thesis in field «Mathematics and computer sciences» on the subject «Parallel implementation of Numerov's method with the help of the CUDA technology».

Owing to installing a new supercomputer in VSU Supercomputer Center in 2016, a problem was posed to develop new teaching aids taking into account the upgraded level of the information technologies being used and the necessity to ensure efficiency of training of modern specialists. The future IT specialists should have competences associated with application of supercomputer computations for accomplishing practical tasks. And having considered the above requirements, in the last two years (2016-2018) VSU created a kit of teaching aids on supercomputer technologies.

The teaching aid kit consists of three books [3-5]. It is noted for being oriented to the practical application of knowledge and formation of competences in solving real computing problems. The tutorials were aimed at the students of different training levels. The first one [3], which is the basis for the whole kit, is meant for the students quite trained in the sphere of information and communication technologies and mathematical methods: senior students, master's degree students, postgraduates, research fellows and specialists. The second one [4] is aimed at the equally skilled hearers, mostly students. The third one [5] is for beginners in the supercomputer technologies. All the tutorials include special best practices making the self-study possible. All the aids have appendices, detailed references, name and subject indexes. The tutorials also discuss some adjacent yet very important issues such as theoretical methods of algorithm analysis.

The basic tutorial [3] of the kit is somewhat different from other teaching aids on parallel programming and high-performance computing.

Firstly, the book is intended for a wide range of the readers' levels – from the juniors beginning to familiarize themselves with the subject to the master's degree students, postgraduates, research fellows

and specialists. Availability in the tutorial of both theoretical materials and computing problem solution examples discussed in detail, as well as the tasks for individual work, makes it useful for both teaching parallel programming courses in higher education institutions and using by the researchers and programmers for computations on parallel computer systems.

Secondly, the material is presented at a rather strict mathematical level; all statements (lemmas, theorems) are provided with proofs, or, in rare instances, with references to the specialized literature on the topic being discussed. Availability of the training material covering construction and analysis of parallel algorithms distinguished this book from many other tutorials on high-performance computing in Russian language.

Thirdly, the provided examples and problems pay much attention to ensuring the required computation accuracy. In all cases, the computation error bounds required from the computer program are strictly defined. The texts of the programs representing the solutions of both standard and more complex problems are usually given in full and unabridged form, which is especially important for first acquaintance with the book's material. Availability of the reference information allows the reader to address the specialized reference literature less often.

As is known, there exist two main approaches to programming of modern computing systems: multithreading-based parallelizing in the shared memory systems and application of the message passing technologies in the distributed systems. Both these approaches are discussed in detail in the tutorial. The thoroughly selected and consistent materials describe various sides of high-performance computations.

All parallel programs used as examples are provided with speedup curves for program code computing part depending on the number of threads in the shared memory system. The speedup values were calculated on the cluster of VSU Supercomputer Center. Visualization of the dependence of speedup on such a parameter as the number of threads allows the students to get deep understanding of the parallel algorithm operating principle.

The appendices accomplish the methodological task of forming in the students of skills required for working with the command shell responsible for communication between the user and the supercomputer system. It is only necessary to know the basic constructs of the C language, which realizes the algorithms and describes the problem solutions presented in the book. This language, on the one hand, is widely known, and on the other hand is one of the main languages for supercomputer computations.

The footnotes provide brief information about the scientists as they are first mentioned in the tutorial.

The tutorial discusses and gives examples of solutions of not only training problems, but also one rather complex real quantum-mechanical problem. This is due to the necessity to teach the students to overcome difficulties arising when passing from regular to parallel programming, and ensuring portability of program suites.

The methods of teaching supercomputer technologies and the educational material had been created for more than 15 years, starting from 2002, when the first parallel computer cluster was launched in VSU [7]. A scientific educational complex was created based on this cluster [8], and a system for protection of remote access to the cluster resources was developed [9]. Since then, training of students and upgrading in the sphere of supercomputer technologies have continuously been performed at the Computer Science Faculty of VSU [10]. Simultaneously, the works have been carried out towards mastering GRID technologies, training students in this field, and including the parallel cluster into the GRID network. A suite of parallel programs for simulation of physical, chemical and other processes has been created and applied [7,10]. Methods for using the cluster in distance training have been developed [11].

Since VSU is a member of the Supercomputer Consortium of Universities of Russia, the experience of other universities has been used when creating the teaching aid kit.

The complex has been successfully tested not only on students but also on the hearers of the advanced education and professional training system [10-12]. Following the orders of the Ministry of Education and Science, for more than 10 years VSU has been providing advanced training for the higher education institution teachers, research fellows and vocational college teachers of priority specialties. In the process of preparing the tutorials, feedbacks had been collected from the students and hearers of the advanced education system as to how the material should be arranged in methodologi-

cally correct manner, how much time should be devoted to certain topics, how the interactive means should be used, and what the ratio of class hours to individual work should be. All this was eventually reflected in the formation of the tutorials' materials. These tutorials are intended exactly for the students and learners of various training levels in the sphere of information and communication technologies, both specialists and non-specialists, as well as the hearers of different ages. They can be used for different training methods, both universal and special, for example person-centered ones, reflecting the peculiarities of the learner groups' compositions, being in a lecture hall or a computer class, with various relation between training with a teacher and individual work. They ensure more efficient adoption of supercomputer technologies, which has been proved by the actual educational process. The tutorials use «from-simple-to-complex» approach, this is why it can also be applied to the continuing education system and when realizing the concept of «Lifelong Learning», passing from a lower training level to a higher one.

The use of the resources of VSU SCC for scientific computations includes the solution of a range of important problems in physics and chemistry. Among such researches we can single out the exploring of atomic processes induced by the intense laser radiation fields; the research of the optical radiation high harmonics generation induced by the atom ionization by fast laser pulses; the calculation of the static electric dipole polarization and calculation of the molecule deformation by laser radiation; the simulation of hydrogen adsorption on metal surface; the investigation of the dynamic characteristics of macromolecules of linear and stellar structure in a melt by molecular dynamic method; the supercomputer simulation of the spatial and electronic structure of nanosystems; simulation in microphotonics (photonic band structure, electromagnetic state density in photonic-crystal systems); the simulation of vibrational modes and elastic characteristics in nanomechanics; the quantum-chemical simulation of ion transport in ion-exchange membranes; the simulation of artificial neural-like structures for the artificial intelligence systems; the implementation of the methods for simulation of formation of carbon nanostructures in plasma. The deformation of triatomic molecules by laser radiation is calculated, its influence on their tunneling ionization by this radiation is determined, and the polarizability of deformed molecules is calculated by multiconfiguration methods [13]. The vibrational parameters of the molecules were obtained by multiconfiguration methods in the absence of experimental values. The geometric characteristics and vibrational frequencies of molecules in a strong laser field are studied [14].

The use of the resources of VSU Supercomputer Center has allowed accumulating considerable experience in managing high-performance systems, solving resource-intensive problems of natural sciences, and organizing the specialists' training in the sphere of information technologies at high scientific and methodological levels.

References

1. Voevodin V.V., Voevodin Vl.V. Parallel programming. Saint Petersburg: BHV-Petersburg, 2002. 608 p.
2. Gergel V.P. High-performance computations for multicore multiprocessor systems. Moscow: MSU Publishing, 2010. 544 p. (Series: Supercomputer education).
3. Borzunov S.V., Kurgalin S.D., Flegel A.V. Practical parallel programming. Saint Petersburg: BHV, 2017. – 236 p. – (Series: University textbooks).
4. Borzunov S.V., Kurgalin S.D. Parallel programming: problems and solutions. Voronezh: VSU Publishing, 2018. 113 p.
5. Borzunov S.V., Kurgalin S.D., Kutsov M.V. Practical parallel programming. Voronezh: VSU Publishing, 2016. 79 p.
6. Order of the Ministry of Education and Science of Russia of 07.08.2014 N.949 "On approval of the Federal state educational standard of higher education in the field of training 02.03.01 Mathematics and Computer Science (Bachelor's level)". URL: http://fgosvo.ru/uploadfiles/fgosvob/020301_mat_i_kom_nauki.pdf (access date: 27.05.2018).
7. Kurgalin S.D., Aleksandrov V.S., Pobezhimov S.V. High-performance computing cluster of Voronezh State University // High-performance parallel computations on cluster systems: materials of the 2nd International Research and Practice Seminar. (Nizhny Novgorod, 2002). P. 174-176.

8. Zapriagaev S.A., Kurgalin S.D. Establishing a regional scientific educational complex on the basis of a high-performance computer cluster // All-Russia Forum "Educational Environment – 2005". Educational Environment today and tomorrow: materials of the II All-Russia Scientific and Practice Conference. (Moscow, September 28 – October 1, 2005). Moscow, 2005. P. 309-310.

9. Kurgalin S.D., Kurakov A.N., Raskhozhev V.V. System for protection of remote access to the resources of the high-performance computer cluster // Telematika'2006: Proceedings of the XIII All-Russia Scientific and Methodological Conference. (Saint Petersburg, June 5-8, 2006). Saint Petersburg, 2006. V. 1. P. 207-208.

10. Kurgalin S.D. Technologies of high-performance parallel computations in scientific researches and educational process of a higher education institution // Izvestiya OrelSTU. Series «Information systems and technologies». 2004. N 2(3). P. 115-119.

11. Kurgalin S.D., Krivtsov A.N., Parshina E.N. Computer technologies of parallel calculations in scientific researches and distance learning system // Internet-Education-Science (IES – 2004): Proc. of the Fourth Intern. and Methodic Conf. (Azerbaijan, Ukraine, Bulgaria, 28 Sept.-16 Oct. 2004). Vinnytsia, Ukraine, 2004. V. 2. P. 433-436.

12. Kurgalin S. D. Advanced training of personnel at the Voronezh University // Higher education in Russia. 2008. N. 1. P. 122-125.

13. Kornev A.S., Chernov V.E., Zon B.A. Laser-induced deformation of triatomic molecules: Influence on tunnel ionization // Phys. Rev. A, 2017. Vol. 96. P. 053408.

14. Kornev A.S., Suvorov K.I., Chernov V.E., Zon B.A. Change of normal coordinates and geometrical parameters of a molecule in a laser field // Optics and Spectroscopy, 2018. Vol. 124. No. 1. P. 24-33.

JSCC RAS (Joint Supercomputer Center of the Russian Academy of Sciences) is non-profit research institute of the Russian Academy of Sciences. It was founded in 1996 by joint decision of Russian Academy of Sciences (RAS), Russian Federation Ministry of Science and Technologies, Russian Federation Ministry of Education and Russian Foundation for Basic Research (RFBR) as a state scientific institution. The center provides researchers from scientific institutions of the Russian Academy of Sciences (RAS) and federal scientific centers with modern computational, telecommunication and informational r...