Bull. Nov. Comp. Center, Comp. Science, 22 (2005), 1–11

Siberian SuperComputer Center (SSCC). Prospects for development*

A.S. Alekseev, B.M. Glinsky, S.P. Kotelevsky, N.V. Kuchin, V.E. Malyshkin, A.V. Selikhov

Abstract. The history of development of the SSCC computer resources and the organization of solving large-scale modeling problems are considered. The results of the project for the development of a multicomputer system SIBIR in the 80-s, the current SSCC structure, the main results of its operation in 2004, and prospects for its development are presented.

1. Introduction

This volume of the Bulletin of the Computing Center RAS (at present, the Institute of Computational Mathematics and Mathematical Geophysics — ICM&MG), Computer Science series, opens a series of the annual issues of the Siberian SuperComputer Center, where scientific results of the current year are presented. Conventional works of the Computing Center in parallel computing which were substantially retarded in the 90-s are by now restored and continued in the SSCC. The recent years problems were overcome, the first large-scale problems in numerical modeling were solved, exploitation of hardware and system software was organized, a team of qualified programmers and engineers for the hardware and the system software service was formed.

At the present time, computing technologies are being intensively developed, especially in science, where they were created in the mid-twenties century. These technologies are being continuously stimulated by rapidly growing needs in mathematical modeling in various fundamental and applied sciences.

2. The SSCC prehistory

Since the creation of the Siberian Branch of the USSR Academy of Sciences, Academicians M.A. Lavrentiev [1] and G.I. Marchuk set the highest priority to the development of hardware and numerical modeling methods

*Partially supported by the grants: Dutch-Russian Project "High Performance Simulation on the Grid", Contract NWO-RFBS 047.016.007; Dutch-Russian Project "Project Investigation of Plasma Induced Cluster Formation and Thin Film Deposition", Contract NWO-RFBS 047.016.018, RFBR 03-07-90302.

in various sciences. They always provided the Siberian Branch, including the Computing Center of the Novosibirsk Scientific Center (NSC) with the most advanced Russian computer facilities.

In 1975, the first network for the remote access from different research institutes to the main frames in the Computing Center was organized. Then, the first resource sharing network system consisting of three BESM-6 computers with a shared disc space was created. Telephone channels of the city telephone station and self-made modems constituted the communication net. In 1979, a project (that of the Resource Sharing Computing Center) for the creation of a corporate high-speed cable network to connect practically all the NSC institutes with three computer systems of the BESM-6 type and three ES-1060 computers of the Computing Center was developed. The implementation of this corporate network of the SB RAS with a trench cable laying between the NSC institute buildings played an important role with the advent of fiber-optic communication systems between computers and with the development of Internet with high-speed optical channels [2].

Another important project in computing started in 1987. Its purpose was to create a multiprocessor computer system called SIBIR on the basis of computers of the ES series: the main frame ES-1068 and eight attached processors ES-2706. The peak performance of the system reached 100 megaflops, which was fairly good for that time. All necessary system parallel software, including an extension of the operating system and the parallel programming system, was developed for this system. The system was oriented, first of all, to the processing of seismic data. About 30 systems were manufactured for the industry. In particular, the first parallel 3D PIC (Particle-In-Cell method) code was developed. When working on the SIBIR project, it was realized that a new generation of system parallel software should be developed. The parallel programming language and the system INYA (Fortran extension, 1989), specially developed for SIBIR, automatically provided the dynamic tunability of application programs to all accessible multicomputer resources.

The Computing Centers in Krasnoyarsk (KCC) and Irkutsk (ICC) were also actively involved in creation of the computer basis of the Siberian Branch of the USSR Academy of Sciences and in the development of the network technologies for the numerical modeling. These Computing Centers were the first Siberian institutes to work in the field of computing technologies [3].

The creation of the SSCC was the next step in the development of a resource-sharing computer center. At the initial stage of creation of the SSCC (during the first four years), Dr. G.N. Erokhin played an active role. He organized a highly qualified team of engineers and system programmers. These specialists created the NSC Resource Sharing Center. In 2000, the Presidium of SB RAS has supported the purchase of the MVS-1000M cluster composed of 32 processors. By the end of 2003, the MVS-1000M computer has reached the planned configuration.

Formally, the SSCC was created in 2001 by the Order of the Presidium of SB RAS No. 100 of 06.06.2001 "On Creation of Resource Sharing Siberian Supercomputer Center of SB RAS" on the basis of ICM&MG SB RAS. The SSCC is a supernumerary structure combining the Department of Computer Systems and Networks of the ICM&MG and the supernumerary scientific-technical group called "Parallel" consisting of specialists from the ICM&MG and other institutions of SB RAS. The SSCC web-site is http://www2.sscc.ru. Academician A.S. Alekseev is its scientific director.

Since 2001, the SSCC has been in uninterrupted 24-hour year-round exploitation. More than 100 thousand hours of processor time have been used by now. 14 institutes of SB RAS used the SSCC services. At present, 22 organizations are the SSCC users. More than 20 large problems were solved to date, and more than 50 problems are under debugging. An important feature of this period was advancement of the parallel programming technology.

To solve large-scale problems successfully, the efforts of scientists and experienced engineers should be combined. The high-performance SSCC equipment is exploited by professionals who had worked on the BESM-6 system and the ES computers (Department of Computer Systems and Networks). This is, however, not enough for solving large and superlarge problems, because to solve such problems it is also necessary to solve complex scientific problems in the field of parallel computations.

Parallel programs are executed on a complicated parallel equipment, and they must possess some new dynamic properties that are difficult to implement. Therefore, in order to develop good-quality parallel programs, it is necessary to know the architectures of parallel computers, parallel programming, and system parallel software. Suffice it to say that the performance of an average statistically parallel program of numerical modeling written in the language C or Fortran can be increased by a factor of three and, practically always, more (by a factor of several tens) if the peculiarities of the multicomputer architecture are taken into account. Errors in the design of parallel algorithms have even a greater effect on the quality of programs.

3. The Supercomputer Software Department (SSD)

Scientific research in the field of parallel computations is concentrated in the SSD. The department continues the traditions and draws on the experience of those scientists who started in the 60-s the investigations of parallel computations in the Institute of Mathematics and Computing Center in Novosibirsk.

In the 60-s of the past century, investigations headed by Dr. E.V. Evreinov (the Institute of Mathematics) started the research aimed at a considerable increase in the computer performance by the parallel execution of a large number of operations were started. Thus, a move from individual computers to computer systems was begun [9].

An idea of a homogeneous universal computer system consisting of identical universal elementary computers connected with each other in some way, with a possibility of programmable change of the structure, both the communication system and the main parameters of elementary computers, was put forward.

However, at that time the technology of electronic components manufacturing could not support the construction of real high-performance parallel computer systems on this basis. This idea, however, served as a background for the development of modern parallel computer technologies, in particular, methods to solve large problems on multiprocessor supercomputers. These ideas constitute the so-called large-block parallelism. Theoretical and experimental investigations of the 60–80-s underline the modern achievements in the construction of parallel computer systems and creation of software for them. In the 60-s, for instance, experiments on creation of computer systems assembled of several computers were performed. Also, a software for the interaction between elementary computers of the system was developed. Languages for the description of algorithms to solve problems of the numerical modeling were also created. At present, such systems represent multicomputers of various types, and the next step of generalization of this idea is the GRID computers. In 1984, this scientific trend was brought from the Institute of Mathematics SB RAS to the Supercomputer Software Department of the Computing Center by a group of researches.

Other researchers of the Department are disciples of A.P. Ershov and V.E. Kotov. They continue investigation of the large-block computations and, first of all, of problems of parallel programming of multicomputers [10, 11]. These specialists determined the comprehensive nature of investigations necessary for the creation of modern parallel computing technologies. At present, there are various projects on the development of systems for debugging and monitoring of the execution of parallel programs, asynchronous programming, software for GRID computations, assembling technology of parallel programming, development of large-scale numerical models and recommendations for effective parallel programming of application problems, visualization of the results of modeling.

In 2002 and 2003, the spring and the autumn lecture courses on parallel algorithms and programs for the SSCC users were organized. The second revised version of the MPI manual was issued. The 7th International Conference on Parallel Computing Technologies was held in Nizhni Novgorod.

Highly qualified and experienced researchers of the Department provide not only the scientific component of the SSCC activities, but, also, advanced training in parallel computing. Although specialists in parallel programming are in permanent demand on the labor market and they often leave the institute, new staff members-graduates of the two chairs based on the department, namely, the NTSU Chair of Parallel Computing Technologies and the NSU Chair of Parallel Computing, come to the department.

4. The main SSCC objectives

- To provide institutes of SB RAS and Novosibirsk State University (NSU) with high-performance computing technologies, technical facilities, and high-quality services for numerical modeling, fundamental and applied research in mechanics, physics, chemistry, geology, biology, and other disciplines;
- To organize training in methods of parallel computations on supercomputers for SB RAS specialists, students and graduate students of the NSU and Novosibirsk State Technical University (NSTU);
- To provide networking with other supercomputer centers of SB RAS, Moscow, other Russian cities, and other countries, the joint development of technologies of distributed computations.

5. The SSCC computational resources and facilities

1. The multicomputer MVS-1000/50H manufactured in 2001 by the state research institute “Kwant” (Russia) with a peak performance of 50 Gflops. The system includes:

- 16 computational modules (2 × DEC Alpha 21264 / 823 MHz / 4 Mb SLC, 2 Gb RAM, 9 Gb HDD SCSI);
- host computer (DEC Alpha 21264 / 633 MHz / 4 Mb SLC, 2 Gb RAM, 36 Gb HDD SCSI-2);
- 2 commutators: Intel EtherExpress E510T (24 ports FastEthernet and 1 port GigabitEthernet) and Myricom M2M-SW16 (16 ports with a speed of 1.2 Gb/s);
- an operating system LINUX RedHat 7.2, a parallel programming library MPICH-1.2.5, a parallel debugger TotalView, compilers C, C++, Fortran (Compaq), a shared access subsystem of the Keldysh Research Institute of Applied Mathematics of RAS.

2. The computer system of two servers, RM600-E30, manufactured by Siemens (Germany) in 1999, with a total peak performance exceeding 5 Gflops. The SMP-servers of the cc-NUMA architecture with a shared

RAM field (a maximum capacity: 8 Gb and 4 Gb) based on the processors R10000 / 250MHz / 4Mb SLC manufactured by the MIPS (the maximum number of processors: 8 and 5) are connected via adapters SCI-2 (1 Gb/s). The operating system Reliant UNIX 5.44, compilers C, C++, Fortran 77, EPC-Fortran, MPI libraries, and the system ADAPTOR of version 7.0, as well as DQS 3.2.7.

3. The graphic server Power Challenge M manufactured by Silicon Graphics Inc. (USA) in 1996 (R7000 / 75 MHz / 2 Mb SLC manufactured by MIPS, RAM 4 Gb, HDD 16 Gb). The operating system IRIX, the database management system Oracle 7.1, the geoinformation systems (GIS) Arcinfo 7.0 and Grass, and the system of scientific visualization IRIS Explorer 2.0.

4. The file server manufactured in 2002 (Athlon MP / 1 GHz, 2 Gb RAM, a controller RAID-5 / IDE, HDD 7×120 Gb = 840 Gb). The operating system Windows 2000.

5. The WWW-server manufactured in 2001 ($2 \times$ Intel PentiumIII / 1000MHz, 2Gb RAM, HDD 36 Gb / SCSI). The operating system Windows 2000.

6. The disc array MA8000 manufactured by Compaq Computer Corporation (USA) with a total capacity of 648 Gb (a maximum capacity: 3 Tb), a RAID-controller, and a commutator FC on 8 optical ports (2 Gb/s).

7. The graphic work station Blade2000 manufactured by the Sun Microsystem (USA) in 2003 ($2 \times$ UltraSPARC III / 1.2 GHz / 8 MB SLC, RAM 8 Gb (maximum), HDD $2 \times$ 73GB / 10000 RPM / SCSI-2, a graphic accelerator SUN XVR1000 / 320 Mb, LAN 1000Base-SX). Two monitors: LCD 24″ (1920x1200) and CRT 21″. The operating system Solaris 9, the package SUN Star Office (a text editor, an editor of vector images, and presentation facilities), a package Forte Compiler (C, C++ and Fortran), the libraries Open GL and QT, and the development environment SUN One Studio (Java).

8. The commutator equipment consists of the following commutators manufactured by Cisco Systems (USA): three commutators Catalyst 2950 (24 ports 100Base-T) in 2002 and Catalyst 3550-12T (10 ports 1000Base-T and 2 ports 1000Base-LX) in 2003.

9. The UPS includes the modules PW-9170 manufactured by Power Wave Inc. (USA) in 2002 (18kVA), 4 modules Smart-UPS 1500 manufactured by APC (USA) in 2001, and 2 modules Masterguard manufactured by Siemens (Germany) in 1999 (3kVA and 6kVA).

10. The conditioning system includes: 2 air-conditioners FTY600GV1B manufactured by DAIKIN (Japan) in 2000 (6 kW) and an air-conditioner OSA-51C-H manufactured by "Tecnair" (Italy) in 2003 (15 kW).

In accordance with the SSCC regulations, organizations and institutes of SB RAS, as well as those from other cities, may use SSCC servers free of charge. Access to the servers is provided via the SSCC corporate network or through the Internet. The users must fill in the registration form available on the www-server at http://www2.sscc.ru/Information/Information_ru.htm. Then, the login and the password are provided by the server administrator. The users obtain the results of computations through the SSCC or the Internet network.

Research problems of 14 SB RAS institutes solved by the SSCC in 2003 included some large ones: modeling of a 3D high-speed flow past aerospace vehicles, evolution of a protoplanetary dust cloud, evolution of the Earth's crust, seismic wave propagation, the particles transfer in physics, oil filtration in elastic-porous media, etc.

6. Development of computer resources

An analysis of large-scale problems of the numerical modeling performed in SB RAS and prediction of the demand for computer resources reveals the following regularity: the number of processors and the RAM volume should be increased every two years by a factor of 2–3. For instance, one needs no less than 100 processors for the problem of modeling of prebiotic investigations with allowance for the chemical reactions between gases*.

Integrating the recommendations and experience of the developers of the big size problems and the administration, the Inter-institute Council on Supercomputations and the SSCC management have proposed the following program of development of the computer resources:

Next year, a 128-processor multicomputer DEC Alpha 21264 / 633 MHz / 4 Mb SLC (64 doubled computational modules with an operative memory of 2 Gb and a disk memory of 9 Gb) and the commutators FastEthernet and Myrinet with the operating system LINUX RedHat 7.2 will be installed;

A 32-processor multiprocessor computer system based on Itanium-II microprocessors (64 doubled nodes with an operative memory of 4 Gb) and the commutators GigabitEthernet and Myrinet with the operating system LINUX will be installed.

To test on real problems the computer performance of the new servers with free software and commercial products, the SSCC has established contacts with firms-producers. For instance, the servers *hp Integrity rx5670* with four processors Intel Itanium II, 1.3 GHz, 3 Mb cache, and 4 Gb of operative memory (can be extended to 96 Gb) and 2-processor servers *Celestica A2210* with microprocessors AMD Opteron 2 GHz, 1 Mb cache, and 4 Gb of operative memory were taken for approbation.

The experience of exploiting the SSCC computational resources to solve large-scale problems shows that the disc space volume for the online use should be increased. As mentioned above, the SSCC has 144 (16×9) Gb in

*V.N. Snytnikov (Institute of Catalysis SB RAS), V.A. Vshivkov, V.E. Malyshkin (ICM&MG SB RAS) the form of local disc memory nodes and 36 Gb at the host node of the MVS-1000M system, 117 Gb of disc memory of the RM-600 system, disc array of 840 Gb of the file server, and 648 Gb of the disc array MA8000, which makes 1785 Gb (∼1.7 Tb). However, the really accessible disc space volume for solving large-scale problems occupies a much smaller volume and is limited both by peculiarities of the communications between different parts of the total disc space and by the order of its allocation to the SSCC users. To further increase the effectiveness of the SSCC operation, it will be necessary to review both factors.

The main line of development of the SSCC disc memory system is the creation of a high-performance file server that controls the available disc space and allows a further increase of its volume. This file server must be functionally specialized and provide the maximum access rate at average effective loading typical of most computational problems. High performance of the file server is needed to maintain effective operation of the shared file system combining the entire accessible file space [4, 5, 6].

One of the most important current problems is the creation, within the SSCC, of a high-performance computer system on the basis of different computer clusters [8], in particular, clusters based on the microprocessors DEC Alpha 21264 and Itanium-II. For this, it will be necessary to solve the following problems:

- The choice of a system architecture (communication equipment, an operating system, a subsystem to manage the system resources, compilers, and libraries);
- Creation of a library of parallel application programs for heterogeneous clusters and purchase application packages with allowance for the user demands (AMBER, Gaussian 03, GROMACS, etc.);
- Organization and realization of large-scale numerical experiments on different clusters with the use of the Grid technology.

The experience of exploiting of high-performance computer systems shows that there always exist problems for which the power of one computer system is insufficient. The need for a more effective use of different available computer systems calls for a combination of computational and informational capabilities of the SSCC, the NSC, and of other cities (in particular, the Yugorsk Research Institute of Informational Technologies in Khanty–Mansiisk). Now, the main problem that does not make possible a wide use of Grid-technologies is a high complexity and limited functions of free software in this field. For a more effective use of various computer resources it is proposed: 1) to develop a concept and a prototype of the user interface for access to combined computer resources (a preferable basis for this is the web-interface); 2) to create a system of distributed dispatching of resources based on specialized servers for effective distribution of various types of problems among the computer resources, control over execution of programs, and interact with the users via this interface; 3) to develop a concept and create a prototype of an intercluster communication service for data transfer between clusters within the same program [6, 8].

7. Some trends of development of high-speed computations

The common experience gained by the Computing Centers in Novosibirsk, Krasnoyarsk, and Irkutsk has confirmed some general trends in the development of mathematical modeling in various disciplines. One of such trends, which is probably the main one, was clearly formulated by Ken Wilson, a well-known American physicist: "Scientific computations have become the third component of the scientific methods that is as important as the experiment and theory" (Computer Physics Communications // Elsevier Science. — 2001.— Vol. 1. — P. 1–6).

As early as in the 50-s, many mathematicians, and among them Academicians M.A. Lavrentyev, S.L. Sobolev, S.A. Khristianovich, and G.I. Marchuk, the founders of Siberian Branch, became aware of these trends resulting from the development of applied mathematics. Therefore, in the "pre-perestroika" period, numerical modeling, computational mathematics, and computer centers of the Siberian Branch of the USSR Academy of Sciences were given a high priority, were in high demand, and were developed rapidly.

Another important trend is realizing by the scientific community the fact that the efficiency of research work greatly depends on the development of quantitative methods [7]. Hence, there are two forms of self-organization of informal scientific teams:

- Formation of numerous multidisciplinary teams that include, in addition to experts in specific fields (that is, researchers such as physicists, chemists, and biologists), applied mathematicians and programmers for the quantitative formulation of problems (this is a tradition of SB RAS, which has been stimulated in recent years by the support of integration projects by the Presidium of SB RAS);
- Formation of online informational-computational interrelations of the interactive type between the theory, experiment, and scientific calculations (This trend corresponds to the following aphorism by R. Descartes: "If not everything is clear — let us compute").

In recent years, large problems of the following types were formulated: the evolution of the protoplanetary substance and planets, the evolution of the Earth; and the evolution of life (in the quantitative statement); bioinformatics, models of climate and ecology, etc. In connection with these problems, which require the integration of knowledge from various fields, the process of forming of specific "societies for solving large-scale problems" is intensified. These inter-institute teams create models, numerical methods, parallel programs, technologies for computational experiments, their analysis, and theoretical generalization.

The demands for computational resources and software determined by these teams, must be satisfied in the first place. In this case, Centers of Supercomputing are a background for the growth of inter-institute research teams.

Now, on the basis of our experience and estimates of well-known international expert groups, we can say, that for solving problems in the fields of nuclear physics and nuclear weapons, molecular nanotechnologies, molecular chemistry, aerospace vehicles design, climate and environment modeling, prediction of storms, tsunami, earthquakes, etc., the following parameters of the computer system should be provided:

- Performance: 10^{15} floating point instructions per second (Pflops),
- Operative memory: thousands of gigabytes.

Systems with such parameters can provide functioning of realistic modern models for the above-mentioned problems within a reasonable time. The increase in the performance of the Top 500 computers shows that the PetaFlops performance could be attained in 2007-2008 and, most likely, in Japan or the USA.

References

[1] Alekseev A.S. M.A. Lavrentiev and informational-computational technologies // Nauka v Sibiri. — 2000. — No. 44/45 (In Russian).

[2] Alekseev A.S., Elepov B.S., Kotov V.E., Metlyayev Yu.V. On a program of works on the creation of a network of informational-computational systems (centers) in the Siberian Branch of USSR Acad. Sci. — Novosibirsk, 1987. — (Preprint / Computing Center SB RAS; 743) (In Russian).

[3] Alekseev A.S., Kuznetsov E.P., Metlyayev Yu.V. et al. Academic regional network of Siberia. — Novosibirsk, 1983. — (Preprint / Computing Center SB RAS; 467) (In Russian).

[4] Foster I., Kesselman C., Globus: a metacomputing infrastructure toolkit // Int. J. Supercomputer Applications. — 1997. — Vol. 11, No. 2. — P. 115–128.

[5] Foster I., Kesselman C., Nick J., Tuecke S. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration // Open Grid Service Infrastructure WG. Global Grid Forum, June 22, 2002.

[6] Selikhov A., Germain C. CMDE: a channel memory based dynamic environment for fault-tolerant message passing based on MPICH-V architecture // Proc. 7th Int. Conf. PaCT-2003. — Springer, 2003. — P. 528–537. — (LNCS; 2763).

[7] Alekseev A.S. On mathematics and informatics in the modern world // Science at the Frontier of Centuries, View from Siberia. — Novosibirsk: Publ. House of SB RAS, 2000. — P. 5–13 (In Russian).

[8] Fougere D., Malyshkin N.V., Malyhskin V.E., Roux B. The NumGRID metacomputing system // NCC Bulletin. Series Comp. Science. — Novosibirsk: NCC Publisher, 2004. — Iss. 21. — P. 57-69.

[9] Evreinov E.V., Kosarev Yu.G. Homogeneous High-Performance Universal Computer Systems. — Novosibirsk: Nauka, 1966 (In Russian).

[10] Kotov V.E., Narinyani A.S. Asynchronous computational processes on memory // Kibernetika. — 1966. — No. 3 (In Russian).

[11] Kotov V.E., Narinyani A.S. On transformation of sequential programs into asynchronous parallel programs // Proc. IFIP Congress 68, Edinburgh, 1968. — Amsterdam, 1969.