Joint NCC & IIS Bull., Comp. Science, 4 (1996), 43–53

Parallel processing in Siberia: towards distributed parallel computing*

O.G. Monakhov

The history of the Siberian researches in parallel processing is surveyed and present project of the Siberian Network Supercomputing Center (SNSC) is considered.

1. Introduction

In the paper the history of the Siberian researches in parallel processing is surveyed and present project of the Siberian Network Supercomputing Center (SNSC) is considered. At present the main feature of the progress in the parallel computing in Siberia is the development of the concept of large block system design and multiple parallel systems is based on the approach in which not chips but large modules represented by special purpose processors including those with mass parallelism and even general purpose computers and parallel computer systems act as building blocks. The Siberian Network Supercomputing Center is based on distributed parallel computing model (similar Network Centric Computing), where the Network Supercomputing Center can be viewed as a single computing resource and parallel applications can be executed in an integrated fashion in a heterogeneous, multivendor environment.

Since beginning of the 60s, the Siberian Scientific Center, also known as "Novosibirsk Academgorodok", became an important center of parallel processing in Russia. The main concepts and projects in parallel processing developed in Academgorodok are presented below.

2. Homogeneous computer systems

In 1962 Edward Yevreinov suggested the concepts of a Homogeneous Universal High Performance Computer System (HUCS), Homogeneous Computing Structures and Homogeneous Computing Media [1], based on the model of a computer collective. The main principles of the model are opposite to model of single (von Neumann) computer: parallelism of operation, programmability of interconnection structure, modularity, homogeneity of elements and

*Supported by the Russian Foundation for Basic Research under Grant 96-01-01632.

interconnections. The model of computer collective has the following properties: universality, high performance, reconfigurability and programmability, reliability and fault-tolerance, low cost/performance ratio.

Homogeneous Computer System is a set of identical Elementary Machines (EM) connected by homogeneous reconfigurable network. EM consists of general purpose (universal) computer (UC) and commutator – system unit (SU) connected neighbouring EMs. The topology (interconnection structure) of EMs and number of EM in the system can be programmable changed.

Several classes of interconnection networks, operating algorithms and systems, parallel languages, parallel application algorithms based on large block parallelizing concept, types of exchange between EMs have been proposed for HUCSs. Several projects of homogeneous parallel computer systems (concentrated and distributed) have been developed in Novosibirsk in the 60s – 70s, namely, Minsk-222, Summa, Minimax, Astra [2].

3. Distributed computing in cellular structures

The Homogeneous Computing Media (HCM) [1] is a logical network consisting of identical and identically interconnected cells. The cell is a universal automata with a logical, memory and interconnection functions. The HCM has the following properties: homogeneity, universality, myopic interconnections and tuning of structure. The main idea of computing media is embedding of arbitrary automata into planar homogeneous cellular structure.

A specific approach to distributed cellular computations was proposed in Parallel Substitution Algorithms (PSA) [3]. It represent an abstract automata model providing for mapping of fine-grain parallel algorithms into cellular arrays. The following fundamental concepts form the background of the Parallel Substitution Algorithm.

1. Fine-grained parallelism. Each data item is introduced being attached to a point in the computation space represented as a countable discrete set of names. The computation is an iterative procedure over a data array in this space. At each step certain data subarrays are replaced by other ones, these actions being done all over the whole data array.
2. Decentralized control: no order of operation execution is defined in the model, each substitution is performed when and where the ready conditions coincide with a data pattern in the space.
3. Synchronous mode of execution: it all operations ready to be executed should be executed simultaneously with the first coming clock pulse.

4. Interpretability by automata nets. A set of substitutions representing the cellular computation admits the direct mapping onto a net of automata. This allows one to construct methods and tools for architectural design of hardware implementation of cellular algorithms.

The direction of the development of the PSA theory is stipulated by the objective of its creation: it to constitute the fundamentals for methods of synthesis of algorithm-oriented architectures of cellular processors.

4. Computer system with programmable structure MICROS

The multimicrocomputer system MICROS is an experimental realization of the concept of computer system (CS) with programmable structure. In 80s MICROS was built from the Soviet family of personal computers "Electronica" in Novosibirsk Scientific Center. The system can be characterized as MIMD multicomputer with programmable interconnection network, distributed operating system, and modular design [4–6].

The MICROS system is based on the following architecture principles allowing high performance, reliability and survivability: parallelism (of control processes, access to data and their processing); decentralization (distribution of control facilities and data); programmability (reconfiguring) of a interconnection structure; modularity and constructive homogeneity; asynchronization of modules and processes interactions; fault-tolerance, static and dynamic system reconfiguration.

In the design of MICROS, main attention has been given to the development of the concept of virtual subsystems, programming structure tools and distributed operating system, to provide the possibility of system applications both for parallel computations and distributed processing in terms of the local network. The software environment with distributed jobs processing, information flow management without deadlocks, programming language Parallel Pascal has been developed for the system and some results of MICROS application in scientific computing and manufacture management systems are also presented [7, 8, 9].

5. High performance heterogeneous system "Siberia"

The architecture and software of high performance computer system (HCS) "Siberia" were designed under the direction of Proffesor N. Mirenkov at the Computing Center of Academy of Sciences in Novosibirsk [10]. It is intended

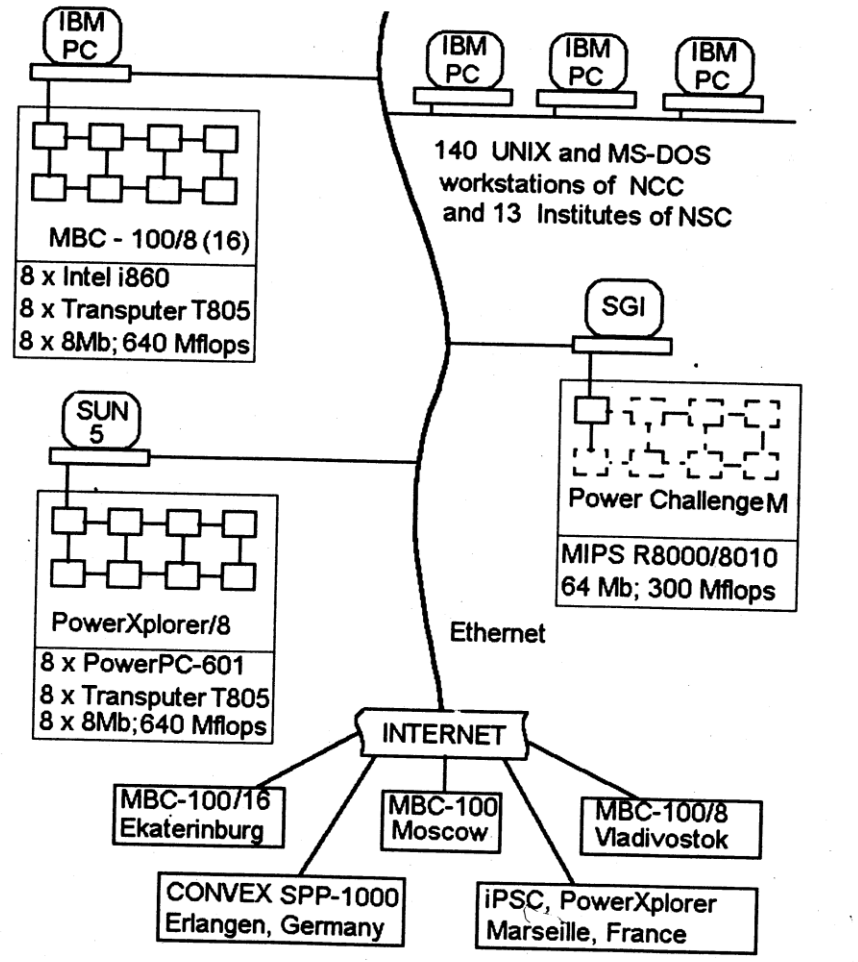
for the support of operation of the supercomputing center for global simulation and large scale computations. The system is based on the large block system design concept with various special purpose processors, multiport memory modules, small computers and even mainframes serving as building blocks. The HCS includes central multimachine kernel surrounded by a set of special purpose subsystems of vector pipeline, vector parallel and associative types. The HCS operation and application is supported by special parallel programming systems designed for representation of the computer intensive problems as communicating processes tuning to a number of accessible processors; by resource monitor for controlling heterogeneous modules with different interfaces and by interactive service system for improving programmer's environment [11].

6. Siberian Network Supercomputing Center

The aim of the project is combination of the efforts of Siberian Mathematical Schools and distributed data bases with modern technologies for distributed parallel data processing. The base of that is creation in Novosibirsk Scientific Center of the Siberian Network Supercomputing Center (SNSC) which oriented to both wide scale computation experiments in multidiscipline researches and mutual interaction with distributed data base and knowledge base, which are used in Internet. The Siberian Network Supercomputing Center is based on distributed parallel computing model (similar Network Centric Computing), where the Center can be viewed as a single computing resource and parallel applications can be executed in an integrated fashion in a heterogeneous, multivendor environment.

The SNSC is built in accordance with concept of large block system design in which large modules represented by special purpose processors including those with mass parallelism and even general purpose computers and parallel computer systems act as building blocks. The Center consists of the following three subsystems: supercomputer ChallengeM (SGI) and parallel multiprocessor systems Power Explorer (Parsytec) and MBC-100 which the Computing Center has now. The latter two systems are Homogeneous Computer Systems based on PowerPC and Intel860 processors with MIMD architecture. All three main subsystems interconnected by LAN and Internet and form the most powerful Center in Siberia. Now the SNSC is in fact common used Center for 13 scientific institutions of Novosibirsk Scientific Center. The Center works with Geo Information System (GIS) (ArcInfo 7.01, Ermapper 2.0, Erdass), various data base system (Informix, Oracle) and another software.

The Center is connected via Internet to other Supercomputing Centers in Russia: in Moscow, Ekaterinburg and Vladivostok. Also, the Center is



connected via Internet to Convex SSP-1000 at the University of Erlangen-Nurnberg (Germany) and to Supercomputing Center in Marselle, France.

The aim of the project is creation of Supercomputing Center, which consists of high performance computing systems, its software, and mathematical and information achievements that can permit to solve new scientific and industrial problems in heterogeneous parallel distributed environment. The result of the project will be the possibilities to solve the problems of analysis and predictions of various regions dynamic of development (especially of Siberia).

The Center will be used by all institutions of Novosibirsk Scientific Center that permit to combine the efforts in the using of geophysics and ecological

information resources, to create new methods and algorithms for parallel processing and develop extra ones [12, 13]. Additional achievements will be integration of specialists of Siberian Region in modern information environment that will be associate to both economy and ecology situation in the region.

7. Software development for parallel distributed systems

The software for parallel distributed systems is developed by researchers of the Laboratory of Parallel Algorithms and Structures under the direction of author in the following areas:

(1) development of the parallel distributed algorithms for mapping of the parallel programs into parallel computer systems (CS);

(2) development new methods and software tools for programming (configuring) of the system structure;

(3) research and development concepts, methods and algorithms for distributed resource management in parallel CS;

(4) development of new classes of macrostructure (interconnection network) for the parallel computer systems;

(5) implementation of the proposed concepts, methods, algorithms and tools for design and creating of the parallel application algorithms.

7.1. Parallel program mapping on parallel architectures

A problem of mapping of a graph of algorithm into parallel computer system is considered [12, 14–16]. Three parallel distributed algorithms for optimal or suboptimal solution of the mapping problem are proposed, the objective functions for mapping and necessary conditions of optimal mapping are developed. The first parallel mapping algorithm is developed for process allocation onto an arbitrary processor network topology through local search [14]. The second algorithm is a parallel recursive algorithm for mapping into multiprocessor systems with pyramidal structure [15]. The third mapping algorithm is based on the self-organizing neural networks [16]. The experimental research of the algorithms was conducted on multitransputer systems, MBC-100, pyramidal multiprocessor system MEMSY and CONVEX SPP-1000 (Germany).

The parallel mapping algorithms are implemented for the parallelizing of scientific software, distributed dynamic load balancing and the decentralized scheduling in massively parallel systems.

7.2. Virtual subsystems – tool for programming of parallel system structure

It is well-known in the field of parallel computing that any parallel algorithm has communication structure. Imposing certain restrictions on the graph structure often leads to more efficient algorithms. The important classes of communicational structure for many parallel algorithms are tree, line, ring, two-dimensional grid. This approach generates the problem of placement or embedding of given subsystem (with given communicational structure and given number of nodes) on the original interconnection network of CS.

A conception of virtual subsystems as a tool for programming (configuring) of parallel system structure is proposed [16, 17]. The parallel distributed generating algorithms of virtual subsystems and organization of communications (addressing and routing procedures) are developed.

The aspects of existence, possibility and conditions of efficient placement and embedding of given subsystem (such as a binary tree, line, ring, square grid) on the regular two-dimensional square mesh (with and without failures of some components) are studied.

Centralized and decentralized (distributed parallel) algorithms are developed for addressing and routing in subsystems, for placement and embedding such subsystems as spanning tree, line, ring, square and skew grid and subgraphs of the parallel system structure.

The virtual subsystem tools were realized in the parallel multicomputer system with programmable structure MICROS.

The conception of virtual subsystems allows to adjust system topology adequate to the structure of the parallel task being solved, to make user program independent of system topology and to organize efficient communications into subsystems.

7.3. Organization of distributed resource management in parallel systems

A hierarchical organization of distributed resource management in parallel CS with MIMD architecture is proposed. The organization is based on the hierarchical nested order of system resources. The parallel access protocols to resources for the asynchronous processes are developed. The protocols provide the polynomial access time to the system resources and prevent the deadlocks between the simultaneously executed processes [18].

A parallel method for adaptive decentralized allocation of the flows of independent tasks in computer systems is proposed. The objective functions for load balancing and for average task execution time are developed. Conditions for the optimal allocation plan and a convergence of the parallel distributed algorithm are proved [19, 20].

A parallel distributed method for solving nonlinear integer optimization problem of resource allocation in multicomputer systems is proposed. The objective functions and conditions for the optimal resource are developed. The proposed parallel method is applied for load balancing and for dynamic allocation of independent tasks in a parallel system [21, 22].

7.4. Synthesis of the optimal interconnection networks for the parallel computer systems

Three classes of regular graphs which present the structure of interconnection networks of CS are proposed and studied. These classes have different value of symmetry, but each of these classes can be characterized by the given number of nodes, given degree of nodes and two classes of graphs can be described in the reduced parametric style [23, 24].

The aspects of existence, connection and isomorphism of such graphs are studied.

Algorithms for synthesis of optimal networks in the proposed classes are developed. The optimal graphs minimize the delay in transit of message transmission and maximize structural reliability. Catalogs of optimal graphs are obtained for the practical use in designing the structure of parallel computers and multitransputer networks.

7.5. Parallel applied algorithms

The proposed concepts, methods, algorithms and tools was implemented for design and creating of the parallel applied algorithms for multitransputer systems, multiprocessor systems CONVEX SPP-1000 and MEMSY (Germany), parallel systems with i860 (MBC-100), Power Explorer (Parsytec) and MICROS in the following fields: geophysics, mathematical modeling, linear algebra, mathematical physics, optimization, image processing, financial operations, synergetic computing and etc. [25–28].

The parallel algorithms for image processing for the multicomputer system with MIMD architecture is investigated. The implementation of the algorithms have been made on the parallel system with programmable structure MICROS. The data structures used to represent the image and the parallel operations on images is developed. The execution times and speed-up of the algorithms are presented [29].

The problems of computation design by statistical simulation of trajectories of the solution of systems of stochastic differential equations on transputer net are considered. The efficiency of the parallel programs for various numerical methods and calculated functionals of the solutions of stochastic differential equations is discussed. Results of numerical experiments are demonstrated [30].

The results of investigations of geophysics problems for seismic data parallel processing by the Common Depth Point (CDP) method and the Wave analogy of the Common Depth Point (WCDP) method on multitransputer networks and multiprocessor pyramidal architecture MEMSY System are presented [13, 31].

8. Conclusion

The short summaries of main concepts and projects of Siberian researchers in the field of the distributed parallel processing are overviewed. Some concepts are developed recently not only in Sibirea, but in other countries of the world.

References

[1] E.V. Yevreinov, Yu.G. Kosarev, *High-Performance Homogeneous Universal Computer Systems*, Nauka, Novosibirsk, 1966 (in Russian).

[2] N.N. Mirenkov, *Parallel Programming for Multimodule Computer Systems*, Moscow, Radio i svyas, 1989 (in Russian).

[3] S.M. Achasova, O.L. Bandman, V.P. Markova, S.V. Piskunov, *Parallel Substitiution Algorithm: Theory and Design*, World Sci. Publ., Singapore, 1994.

[4] O.G. Monakhov, *Multimicrocomputer system with programmable interconnection structure*, Proc. Internat. Conf. CONPAR-90, Vol. special tech. contrib., Zurich, 1990.

[5] V.V. Korneev, *Computer System Architecture with Programmable Structure*, Nauka, Novosibirsk, 1985 (in Russian).

[6] O.G. Monakhov, E.A. Monakhova, *Virtual subsystems – tool for programming of parallel system structure*, Proc. Internat. Conf. "Parallel Computing Technologies", Obninsk, 1993, 525–527.

[7] O.G. Monakhov, V.V. Korneev, M.S. Tarkov, *Operating system kernel of computer system with programmable structure*, Odnorodnye vychislitelnye sistemy, Computer systems, **90**, Novosibirsk, 1981, 22–42 (in Russian).

[8] E.A. Monakhova, V.V. Korneev, O.G. Monakhov, *Organization of distributed job management in MICROS computer system*, Upravlyayushchie sistemy i mashiny, No. 5, 1987, 31–33 (in Russian).

[9] O.G. Monakhov, E.A. Monakhova, *Performance evaluation of the information flow management algorithm in computer systems and networks*, Electronnoe modelirovanie, No. 6, 1989, 91–93 (in Russian).

[10] N.N. Mirenkov, *The Siberian approach for open-system high-performance computing architecture*, Computing and Control Engineering Journal, **3**, No. 3, 1992, 137–142.

[11] V.A. Anisimov, V.E. Malyshkin, *Assemble parallel programming system INYA*, Proc. Internat. Conf. "Parallel Computing Technologies", World Sci. Publ., Singapore, 1993, 316–324.

[12] O.G. Monakhov, *Parallel mapping of parallel program graphs into parallel computers*, Proc. Internat. Conf. "Parallel Computing 91", Elsevier Science Publishers, Amsterdam, 1992, 413–418.

[13] V. Elinov, O. Monakhov, I. Ushatinsky, *Seismic data processing: parallel and consrevativ computations on multitransputer system*, Proceedings "Parallel Scientific Computing, First International Workshop, PARA'94, Lyngby, Denmark, June 1994", Lecture Notes in Computer Science, **879** 190–193.

[14] O.G. Monakhov, *Parallel algorithm for mapping of program graphs into parallel computers*, Proc. Internat. Conf. "Parallel Computing Technologies", World Scientific, Singapore, 1991, 471–476.

[15] O. Monakhov, *Parallel algorithm for mapping parallel programs into pyramidal multiprocessor*, Proceedings "Applied Parallel Computing, International Workshop, PARA'95, Lyngby, Denmark, August 1995", Lecture Notes in Computer Science, **1041**, Springer–Verlag, 1995, 436–442.

[16] O. Monakhov, O. Chunikhin, *Parallel mapping of program graphs into parallel computers by self-organization algorithm*, Proceedings "Applied Parallel Computing in Industrial Problems and Optimization,, International Workshop, PARA'96, Lyngby, Denmark, August 1996", Lecture Notes in Computer Science, Springer–Verlag, 1996 (to appear).

[17] O.G. Monakhov, E.A. Monakhova, *Organization of intercomputer interaction in the computer system MICROS with programmable structure*, Raspredelennaya obrabotka informatsii, Computer systems, **105**, Novosibirsk, 1984, 85–104 (in Russian).

[18] O.G. Monakhov, *Organization of distributed resource management in computer system with programmable structure*, Programmirovanie, No. 2, 1984, 75–83 (in Russian).

[19] O.G. Monakhov, *Optimization of distributed resource management in computer system with programmable structure*, Programmirovanie, No. 5, 1985, 94–101 (in Russian).

[20] O.G. Monakhov, *Distributed dynamic resource management in computer systems with programmable structure*, Automatika i vychislitelnya tekhnika, No. 3, 1984, 9–17 (in Russian).

[21] O.G. Monakhov, *Analysis of parallel dynamic algorithms for decentralized tasks allocation in computer systems*, Raspredelennaya obrabotka informatsii, Computer systems, **105**, Novosibirsk, 1984, 45–89 (in Russian).

[22] O.G. Monakhov, *About the method of distributed task allocation in computer systems with programmable structure*, Automatika i vychislitelnya tekhnika, No. 1, 1983, 7–13 (in Russian).

[23] O.G. Monakhov, *Parametric description of graphs of intercomputer connections of homogeneous computer systems*, Voprosy teorii i postroenia vychislitelnykh sistem, Computer systems, **80**, Novosibirsk, 1979, 3–17 (in Russian).

[24] O.G. Monakhov, V.V. Korneev, *Graphs of interconnection networks of homogeneous computer systems*, Zadachi i teoriya Comp'yuternykh sistem, Computer systems, **73**, Novosibirsk, 1978, 93–106 (in Russian).

[25] V.D. Korneev, O.G. Monakhov, *Investigation and realization of operations with sets on multitransputer systems*, Proc. Internat. Conf. "Advanced Mathematics, Computations and Applications", Novosibirsk, 1995, p. 183.

[26] O.G. Monakhov, *Distributed system for data acquisition and processing in array "Extended Air Shower (EAS) - 1000"*, Parallel Computing Technologies, Novosibirsk, 1992, 91–121.

[27] S. Artemiev, S. Gusev, O. Monakhov, *Use of parallel computing for account of options by a Monte-Carlo method*, Proceedings "Applied Parallel Computing in Industrial Problems and Optimization, International Workshop, PARA'96, Lyngby, Denmark, August 1996", Lecture Notes in Computer Science, Springer–Verlag, 1996, (to appear).

[28] O.G. Monakhov, *Synergetic computing on multitransputer system*, Proc. All-Union Conf. "Transputer systems and its applications", Moscow, 1992.

[29] O.G. Monakhov, E.A. Monakhova, *Investigation of image processing algorithms on parallel computer system*, Autometria, No. 5, 1993, 89–95.

[30] S. Artemiev, S. Gusev, O. Monakhov, *Numerical solution of stochastic differential equations on transputer network*, Proceedings "Applied Parallel Computing, International Workshop, PARA'95, Lyngby, Denmark, August 1995", Lecture Notes in Computer Science, **1041**, Springer–Verlag, 1995, 33–37.

[31] A. Kremlev, O. Monakhov, T. Thiel, *Parallel seismic data processing method for MEMSY multiprocessor system*, Proc. Internat. Conf. "Parallel Computing Technologies", St.-Petersburg, Lecture Notes in Computer Science, **964**, Springer–Verlag, 1995, 434–438.