ABSTRACT
for the dissertation of Nurislam Kassymbek
submitted for the degree of Doctor of Philosophy (PhD) in the educational program 8D06103 – Computer Science
on the topic: "Development of effective parallel algorithms in the form of active knowledge to solve the problem of nonequilibrium filtering (compositional model) and parallel program execution algorithms"

Relevance. In the modern world, high-performance computing is rapidly developing, which allows researchers to solve various new complex scientific and applied problems related to the modeling of different processes, big data processing, and, in general, tasks that require a large amount of computation. One such task is solving systems of differential equations using numerical methods. When solving systems of differential equations numerically, the required computational resources, including computation time, increase as the desired accuracy of the solution increases. Computing on supercomputers allows for significantly accelerating the solution of many problems.

In the field of solving numerical problems, there is an issue of developing efficient parallel algorithms that run on a supercomputer. Efficiency refers to reducing the execution time and memory usage of parallel programs, dynamically distributing the load among computing nodes, and so on. Achieving good efficiency in parallel programs can present certain challenges for the parallel program developer. In this area, the programmer must take into account the features of the chosen numerical method, input data, computing system, load distribution among computing nodes; implement not only computation but also communication between nodes. The complexity of writing efficient programs is caused by the heterogeneity of computing systems (presence of GPUs and other accelerators).

Programs for high-performance computing systems can be implemented manually or using ready-made solutions. One such ready-made solution is the LuNA system, which addresses the issues of developing efficient parallel programs.

LuNA is a fragmented programming system that enables the automation of the process of parallel execution of various programs on a supercomputer. The process of achieving the aforementioned properties is automated within this system. The quality of automatically generated programs may be inferior to manually implemented programs by certain criteria, but automation reduces the effort required for development, debugging, and maintenance of the code.

The goal of the dissertation is the development of a fragmented algorithm for solving the problem of modeling the flow of a multiphase multicomponent fluid in a form that can be integrated into the LuNA active knowledge base.

Research objectives:

1) Numerical solution of the problem of modeling multicomponent multiphase flow using Newton and GMRES methods in oil production problems;
2) Development of a fragmented algorithm for solving the oil production problem;
3) Implementation of the fragmented algorithm using MPI;

4) Implementation of the fragmented algorithm in the LuNA language;

Improvement of the developed algorithms and presentation of them in a form that can be integrated into the LuNA system.

The object of the research is geophysical processes in an oil reservoir, high-performance computing, and the active knowledge base.

The subject of the research is parallel and fragmented algorithms for solving the problem of nonequilibrium filtration.

Scientific novelty of the obtained results

- For the first time, a fragmented algorithm for the numerical solution of the problem of modeling the flow of a multiphase multicomponent fluid based on the Newton method has been developed;

- A library subroutine has been developed — an algorithm with the described properties for reuse in solving a class of problems involving systems of equations with nonsymmetric and non-diagonally dominant matrices, which allows for dynamic control enabling dynamic resource distribution and load balancing among computing nodes.

Provisions submitted for defense:

- Developed and implemented large-block fragmented algorithms for solving the problem of modeling the flow of a multiphase multicomponent fluid based on the Newton method;

- Library subroutines ensuring the implementation of dynamic properties (resource and computation management) for solving the filtration problem.

Theoretical and practical significance of the research

The theoretical significance of the work lies in the study of methods for developing parallel/distributed algorithms for solving filtration problems with linearization and subsequent solution of linear equations.

The practical significance of the work lies in the applicability of the developed parallel/distributed algorithms for solving many oil production problems, as well as problems in sparse matrix algebra. The developed fragmented algorithm is implemented in the LuNA environment, which allows it to be included in the active knowledge base, ensuring automatic and portable execution on various computing architectures. This significantly reduces the labor required for developing parallel programs and enables the algorithm to be reused multiple times in problems of a similar structure.

The reliability of the scientific results is ensured by the use of proven methods (parallel programming, fragmented programming, load balancing), the consistency of computation results from various parallel algorithms on heterogeneous computing resources, the internal consistency of the proposed algorithms, and the agreement of the obtained results with those of other authors.

The topic of the dissertation is aligned with the plans of the following research programs:

- "Development of an active knowledge management system for automating the construction of high-performance parallel programs for unstructured data processing and numerical modeling in filtration problems" — state registration No.

0118RK00741, conducted at the Research Institute of Mathematics and Mechanics under grant funding, planned for 2018–2020.

- "High-performance computer modeling of multiphase fluid flow in a porous medium under uncertainty" — state registration No. 0121RK00431, conducted at the National Engineering Academy of the Republic of Kazakhstan under grant funding, planned for 2021–2023.

Structure and Content of the Dissertation

The introduction substantiates the relevance of the dissertation topic, formulates the objectives and goals of the research, scientific novelty, research methods, practical significance of the work, and presents a brief overview of the dissertation chapters.

The first chapter of the dissertation discusses numerical methods for solving the problem of modeling multicomponent multiphase flow, parallel architectures, standards of parallel programming, and libraries of parallel subroutines. Based on the research results, numerical methods were selected to solve the problem under consideration, and a fragmented technology was chosen for the development of a parallel algorithm, which facilitates the implementation of algorithms and provides automatic balancing of computational processes on the computing machine.

The second chapter describes the mathematical model of multicomponent multiphase fluid flow in a porous medium. It considers the numerical solution and presents the difference scheme for the one-dimensional case, and describes the algorithm for solving linear equations with large sparse matrices.

The third chapter describes the development of the fragmented algorithm and its implementation in the LuNA system. A description of the fragmented algorithm for solving the chosen problem is provided. Ways to reduce the execution time of the fragmented program are shown.

The fourth chapter presents the testing results in terms of the efficiency of the implemented parallel algorithms. A comparison is made between the developed algorithms for solving the oil production problem based on the Newton method.

The conclusion summarizes the main theoretical and practical results obtained in the dissertation, highlights possible directions for further research, and presents the main results of this dissertation work.

Approval of the Work.

The results of the work were discussed at scientific seminars of the Department of Computer Science and the Faculty of Information Technologies of Al-Farabi Kazakh National University, as well as presented at the following conferences:

- IV International Scientific and Practical Conference "Informatics and Applied Mathematics" (Almaty, 2019);
- 17th European Conference on the Mathematics of Oil Recovery (ECMOR XVII), (online, 2020);
- Traditional International April Mathematical Conference in honor of the Day of Science Workers of the Republic of Kazakhstan, dedicated to the 75th

anniversary of Academician of the NAS RK Tynysbek Sharipovich Kalmenov (Almaty, 2021);

- The 18th European Conference on the Mathematics of Oil Recovery (ECMOR 2022), (Гаага, Нидерланды, 2022).

Publications. Eight printed works have been published on the topic of the dissertation, including 2 in publications recommended by the Committee for Quality Assurance in the Sphere of Science and Higher Education of the Ministry of Science and Higher Education of the Republic of Kazakhstan, 2 in peer-reviewed journals included in the international citation databases "Web of Science" and/or "SCOPUS", and 4 in proceedings of international conferences:

- Mukhambetzhanov, S. T., Lebedev D.V., Kassymbek N.M., Imankulov T.S., Matkerim B., Akhmed-Zaki D.Zh. (2020). GMRES based numerical simulation and parallel implementation of multicomponent multiphase flow in porous media. Cogent Engineering, 7(1), 1785189. https://doi.org/10.1080/23311916.2020.1785189
- Imankulov T.S., Kassymbek N.M., Lebedev D.V., Matkerim B., Daribayev B.S. Numerical Simulation of Multiphase Multicomponent Flow in Porous Media: Efficiency Analysis of Newton-Based Method. Fluids 2021, 6, 355. https://doi.org/10.3390/fluids6100355
- Kassymbek N.M., Akhmed-Zaki D.Zh., Imankulov T.S., Lebedev D.V., Kenzhebek E.G. High-performance and Intelligent Computational Models for Oil Production Problems // Bulletin of the National Engineering Academy of the Republic of Kazakhstan, 2020, 8 p.
- Kassymbek N.M., Lebedev D.V. MODELING THE MOVEMENT OF TWO-PHASE FLUID USING THE NEWTON METHOD // Bulletin of Abai KazNPU, Series "Physical and Mathematical Sciences", No. 4(76), 2021, pp. 38–44
- Kassymbek N.M., Matkerim B., Lebedev D.V., Imankulov T.S., Akhmed-Zaki D.Zh. Numerical Modeling and Parallel Implementation of Multicomponent Multiphase Flow in Porous Media Based on the GMRES Method. PROCEEDINGS of the IV International Scientific and Practical Conference "Informatics and Applied Mathematics", September 25–29, 2019, Almaty, Kazakhstan. – pp. 253–260.
- N. Kassymbek, B. Matkerim, D. Lebedev, T. Imankulov and D. Akhmed-Zaki, GMRES Based Numerical Simulation of Multicomponent Multiphase Flow in Porous Media on LuNA Fragmented Programming System, Conference Proceedings, ECMOR XVII, Sep 2020, Volume 2020, p. 1-10.
- Akhmed-Zaki D.Zh., Lebedev D.V., Qasymbek N.M. High-Performance Modeling of Multicomponent Multiphase Fluid Flow in Porous Media / Traditional International April Mathematical Conference in honor of the Day of Science Workers of the Republic of Kazakhstan, dedicated to the 75th anniversary of Academician of the NAS RK Tynysbek Sharipovich Kalmenov, International, KAZAKHSTAN, 05.04–08.04.2021
- Kassymbek, N., Lebedev, D., & Akhmed-Zaki, D. (2022). Optimization of a GMRES-Based Parallel Algorithm for Numerical Simulation of

Multicomponent Multiphase Flow in Porous Media. In ECMOR 2022. ECMOR 2022. European Association of Geoscientists & Engineers. https://doi.org/10.3997/2214-4609.202244055