UDC 001.89:004.42:519.85

DEVELOPING THE OPEN SCIENCE INFRASTRUCTURE: A SUPERCOMPUTING PLATFORM FOR LARGE MATRIX COMPUTING

Sidko Alla
National University of "Kyiv-Mohyla Academy",
ORCID iD: 0000-0001-7035-9447,
E-mail: a.sidko@ukma.edu.ua

Malaschonok Gennadi
Doctor of Physical and Mathematical Sciences, Professor, National University of "Kyiv-Mohyla Academy",
ORCID iD: 0000-0002-9698-6374,
E-mail: malaschonok@ukma.edu.ua

Abstract. *It is proposed to use a new DAP supercomputer runtime for the creation of a component of the Open Science infrastructure in the form of access to a supercomputer with a special library of matrix algorithms. This will allow solving problems with large matrices in many application areas.*

Keywords: *Library of matrix algorithms, supercomputer runtime, DAP, Open Science infrastructure, large size matrices.*

Introduction

A researcher's access to high-performance computing resources is one of the necessary conditions for their successful scientific work. For example, packages similar to OpenFOAM can be used to successfully solve tasks in the field of aerodynamics and hydrodynamics, heat transfer, radiation, and combustion. Such a package should be installed on a computing cluster as a component of the Open Science infrastructure. Another example is the well-known Linpack package, which allows you to efficiently solve problems of linear algebra.

Materials and methods

We propose the creation of another component of the Open Science infrastructure in the form of access to a supercomputer that supports a special library for matrix algorithms. The library will work with a new runtime environment for tasks execution on a supercomputer with distributed memory. This environment is designed for recursive matrix algorithms for both dense and sparse matrices. It provides dynamic decentralized control of the computational process and significantly expands the size of matrix problems that can be successfully solved. This topic is relevant since the modern world relies on Big Data computations. In this environment, such matrix algorithms as matrix multiplication, finding the inverse and pseudo-inverse matrix, QR and LU decomposition of matrices, the Cholesky decomposition, finding the adjoint matrix and the kernel of a linear operator have already been implemented. The SVD algorithm is also being developed, which is actively used in recommendation systems.

Research results

The new mechanism for managing the computing process on the cluster is used. This mechanism can be applied to any recursive algorithms with both dense and sparse data. The new control scheme is called the DAP scheme [1]. Its main difference is that it sequentially deploys functions in depth, saving all states at any nesting level until all computations in the current computation subtree are complete. This mechanism allows any processor to switch tasks without waiting for the current task to finish. It can be used only for block-recursive algorithms. In such algorithms, independent separate subtasks are applied to blocks, so it is easy to organize decentralized control over the entire computational process.

Protection in case of node failure during the computational process is an important feature of the execution environment. The parent node that sent the data to the child node should receive the result. However, it may receive child failure messages. In this case, it should forward this task to another node. No additional changes should be made on other nodes. Only one subtree that matched this data will be lost and recomputed. The scheme was implemented using the Java programming language and OpenMPI [2].

Conclusion

Since Ukraine does not have its own high-performance supercomputers, international cooperation is needed to create such an Open Science infrastructure. It is necessary to create a joint team together with those organizations in the EU that have supercomputers of great power and have scientific teams working in areas close to ours. Our report aims to find partners in the EU who are interested in creating a joint team.

References

1. Malaschonok G, Sidko A. Distributed computing: DAP-technology for parallelizing recursive algorithms. Scientific notes of NaUKMA. Computer Science. 2018;1:25-32.
2. Malaschonok GI, Sidko AA. Supercomputer Environment for Recursive Matrix Algorithms. Program Comput Soft. 2022;48:90–101. DOI: https://doi.org/10.1134/S0361768822020086

РОЗРОБКА ІНФРАСТРУКТУРИ ВІДКРИТОЇ НАУКИ: СУПЕРКОМП'ЮТЕРНА ПЛАТФОРМА ДЛЯ ВЕЛИКИХ МАТРИЧНИХ ОБЧИСЛЕНЬ

Сідько Алла
асистент, Національний університет «Києво-Могилянська академія»,
ORCID iD: 0000-0001-7035-9447,
E-mail: a.sidko@ukma.edu.ua

Малашонок Геннадій
д-р.фіз.-мат.наук, проф., Національний університет «Києво-Могилянська академія»,
ORCID iD: 0000-0002-9698-6374,
E-mail: malaschonok@ukma.edu.ua

Анотація. *Для створення компонента інфраструктури Open Science у вигляді доступу до суперкомп'ютера зі спеціальною бібліотекою матричних алгоритмів пропонується використовувати нове середовище виконання суперкомп'ютера DAP. Це дозволить вирішувати проблеми з великими матрицями в багатьох областях застосування.*

Ключові слова: *Бібліотека матричних алгоритмів, середовище виконання суперкомп'ютера, DAP, інфраструктура Open Science, матриці великого розміру.*