КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени АЛЬ-ФАРАБИ
ИНСТИТУТ ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ
СИБИРСКОГО ОТДЕЛЕНИЯ РАН

ISSN 1560-7534
ISSN 1563-0285

СОВМЕСТНЫЙ ВЫПУСК

по материалам международной научной конференции
"Вычислительные и информационные технологии в науке, технике и образовании"
(CITech-2015)
(24-27 сентября 2015 года)

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

Том 20

ВЕСТНИК КАЗНУ им. АЛЬ-ФАРАБИ

Серия математика, механика и информатика № 3 (86)

ЧАСТЬ I

АЛМАТЫ – НОВОСИБИРСК, 2015

Вычислительные Технологии

2015

Том 20

Computational Technologies

2015
Vol 20

ВЕСТНИК КАЗНУ

2015

Серия математика, механика и информатика

№ 3 (86)

Научное издание

Вестник КазНУ

Серия математика, механика, информатика

№ 3(86) 2015

ИБ № 8514

Подписано 14.04.2015 г. Формат 60x84 1/8. Бумага офсетная.
Печать цифровая. Объем 30,5 п.л. Тираж 500 экз. Заказ № 2617.
Казахского национального университета им. аль-Фараби.
050040, г. Алматы, пр. аль-Фараби, 71, КазНУ.
Отпечатано в типографии издательского дома «Қазақ университеті».
Издательский дом «Қазақ университеті»

The BULLETIN of KAZNU

Mathematics, Mechanics and Informatics Issue

2015

№ 3 (86)

Table of Contents

Session I. Technological Process Automation and Control

Session II. Information Management, Processing and Security

A Scalable Parallel Algorithm and Software for 3D Seismic Simulation on Clusters with Intel Xeon Phi Coprocessors

Dmitry Karavaev, Boris Glinsky, and Valery Kovalevsky

Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
Prospect Akademika Lavrentieva 6, 630090 Novosibirsk, Russia
kda@opg.sscc.ru, gbm@opg.sscc.ru, kovalevsky@sscc.ru
www.sscc.ru

Abstract. In this paper, we present the results of the research in to the development of a scalable parallel algorithm for solving large problems of the forward modeling in geophysics. The problem to be solved is the system of equations of elastic theory representing the wave propagation in elastic 3D media. We have developed a scalable parallel algorythm and a program for the 3D seismic wave simulation on modern multi-core clusters with a hybrid architecture based on Intel Xeon Phi coprocessor. We present this parallel algorythm for solving the above-mentioned problem and the results of the parallel algorithm behavior on the Xeon Phi based cluster for different tests of the parallel program code. In addition, we compare implementation of the proposed parallel algorithm on different computing devices.

Keywords: parallel algorithm, seismic simulation, scalability, Xeon Phi, hybrid cluster.

Introduction

One of the methods for solving inverse geophysical problems is solving the forward problem for a various number of models, which are different in geometry structure and elastic parameters values [1]. Thus, carrying out the simulation, varying elastic parameters and establishing the correspondence with natural geophysical data, one can find a more appropriate geometrical structure and elastic parameters values of a geophysical object under investigation. In addition, one of the useful and well-known methods for solving a 3D seismic simulation problem is a difference method based on 3D grids [2]. The most useful difference methods can be a second or a fourth order of approximation [4,4,6] and have application in modeling elastic or viscoelastic media [7]. The more difficult are geophysical models the more difficult is to calculate them. Because of using the difference method one should handle with large 3D arrays and a great volume of data. Only a 3D grid model of 3D isotropic geophysical media is described with the three parameters: density and two velocities of elastic waves. In this paper, we deal only with isotropic 3D elastic media. The problem of 3D elastic wave propagation is presented in terms of velocity and stress. Therefore, we need to calculate nine 3D values that are 3D arrays. When using the 3D explicit difference scheme with the iterative technique one should use the values at two time steps of iteration. To carry out calculation for 3D models with a detailed representation is a difficult task because of dealing with large 3D arrays that are to be placed in the operation memory of a computer. In such a case, researchers use powerful multi-core calculation systems. Such systems can have different architectures. Modern cluster systems that are in the first places in TOP 500 rating have a hybrid architecture. This means that these systems have special computing devices that are presented by Nvidia GPU or Intel Xeon Phi cluster. Some examples of such clusters are NKS-30T+GPU cluster of the Siberian Supercomputer Center and MVS-10P cluster of the Joint Supercomputer Center. With the use of such computing systems, one can solve large 3D models in parallel. So each of the computing device solved its part of 3D data, all together covering the whole 3D area under study. Therefore, we need to develop a scalable

parallel algorithm and a program code for using such a device in calculations. It is not only a programmable problem but also a researcherв Ђ™s problem. When developing a scalable parallel algorithm we should make special tests of a program code on one computing device to tune it for the 3D elastic wave simulation using difference method to watch the behavior of a program on different number of computing devices and models with different volume of data. There are many program codes realizing difference methods for the seismic wave propagation modeling on clusters with GPUs [8,11]. Using the Intel Xeon Phi coprocessors in simulation is a new and modern approach for parallel computation. Such systems can allow researchers using OpenMP parallel tools make developing programs for large-scale simulation easier. Our main purpose is to describe how the difference method for numerical simulation of seismic wave propagation is implemented on supercomputers with Intel Xeon Phi coprocessors. Section 2 gives a brief description of the problem statement and numerical method. In Section 3 we describe parallel implementation in detail. Section 4 presents computational experiments for different test of a parallel code. We discuss the implementation of the simulation code for the single coprocessor case to tune the script parameters. Second we discuss the multi-device case in order to study parallel algorithm behavior for large-scale problems. Section 5 concludes the main results.

Problem Statement

We solve the forward geophysical problem of the 3D elastic wave propagation and deal with isotropic and elastic material [9]. Before carrying out the calculation we have to construct the 3D grid model of a geophysical medium under study. Such an object is described by parameters of density, shear wave velocity and longitudinal wave velocity. Thus, we need three material parameters (LamBee coefficients and density) in a difference scheme. All the geometries of elastic media have plane free surfaces. Such a medium can have a difficult geometrical structure and different values of elastic parameters at each point of 3D grid. A problem of the 3D elastic wave propagation is described in terms of components of velocities of displacements $\mathbf{u} = (U, V, W)^T$ and components of a stress tensor $\sigma = (\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz})^T$. The problem is to be solved with appropriate initial conditions and boundary values. We apply a free-surface condition at the top boundary. We use the Cartesian coordinate system. To numerically solve the simulation problem, we use the difference method [4]. This method is based on using staggered grids. This means that different values are placed at different points of a grid cell. The difference scheme is of second order of approximation with respect to time and space. The government equations of difference scheme will be in the form of (1).

$$\rho\frac{\partial \mathbf{u}}{\partial t} = [A]\sigma + \mathbf{F}(t,x,y,z), \frac{\partial \sigma}{\partial t} = [B]\mathbf{u}; \tag{1}$$

$$A = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} & 0 \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} & 0 & \frac{\partial}{\partial z} \\ 0 & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix}, B = \begin{bmatrix} (\lambda+2\mu)\frac{\partial}{\partial x} & \lambda\frac{\partial}{\partial y} & \lambda\frac{\partial}{\partial z} \\ \lambda\frac{\partial}{\partial z} & (\lambda+2\mu)\frac{\partial}{\partial y} & \lambda\frac{\partial}{\partial z} \\ \lambda\frac{\partial}{\partial z} & \lambda\frac{\partial}{\partial y} & (\lambda+2\mu)\frac{\partial}{\partial z} \\ \mu\frac{\partial}{\partial y} & \mu\frac{\partial}{\partial x} & 0 \\ \mu\frac{\partial}{\partial z} & 0 & \mu\frac{\partial}{\partial x} \\ 0 & \mu\frac{\partial}{\partial z} & \mu\frac{\partial}{\partial y} \end{bmatrix}$$

Our modification for the elastic wave propagation simulation is that we use the calculated coefficients in the developed program. This means that coefficients from the problem statement differ from those we use in the difference scheme, including all the summations and multiplications, and are placed in special 3D arrays.

Parallel Implementation

We consider hybrid parallel implementation, using both CPUs and Intel Xeon Phi coprocessors for computation. We have developed a scalable parallel algorithm based on the difference method with 3D grids and the program code for a cluster with a hybrid architecture and Intel Xeon Phi coprocessors. The developed parallel scheme includes use of technologies for parallel computing such as Message Passing Interface (MPI) and a software for Intel Xeon Phi coprocessor programming. We use MPI and OpenMP, respectively, for parallel computations. Our parallel realization has a data distributed character. We divide a large 3D model into smaller 3D subdomains, Fig. 1. Each of them is calculated independently and in the parallel manner. For the computations, we use the multi-core computing system with Intel Xeon Phi coprocessor. Several CPUs and several Xeon Phi coprocessors (devices) are placed at the computing nodes of such a system. Each Xeon Phi coprocessor can be treated as SMP (Symmetric Multiprocessing) machine. Such a device has 8GB DDR5 memory, 60 computing cores based on x86 architecture, 4HW threads/core, IP addressable, have Linux OS. We can employ up to 240 parallel threads with such a device. The cluster consists of 207 computing nodes with 2 Xeon E5-2650 processors and 2 Intel Xeon Phi 7110X coprocessors, www.jscc.ru. For running the program code on it, one should recompile a program code. We take the 3D model data on the CPUs and initialize necessary 3D arrays on the computing devices. After that, we copy the model data into the computing devices. Then we can carrying out computations using a parallel algorithm. All the calculations for 3D subdomains are conducted only on devices. The CPU is used only for device manipulation and for making exchanges between data placed in the devices. The Xeon Phi coprocessors are used in the offload mode. Thus, direct communication between coprocessors is not available. The data must be sent from coprocessor to the host CPU in order that the data be exchanged with the other coprocessors. In our parallel realization, to make the next time step we should make data exchange among neighbor devices that can be either at one computing node or at different computing nodes of the cluster. We use non-blocking MPI Send/Receive procedures that take place for the data exchange among the computing devices, Fig. 1. Since we use the 3D domain decomposition, we first compute for the points on the sides of subdomains. Second, we start the data copying from Xeon PhiЂ™s to CPU cores and run exchange procedures with use of special designed buffers and MPI functions. Then we do the computation for the remaining internal grid points of 3D subdomains and the communication procedures simultaneously. After that, we verify whether all the exchanges have been done and make a data copy from buffers placed at CPUs into buffers at Xeon Phi coprocessors and then into 3D arrays. After Then we proceed to the next time step. Therefore, we overlap the communication and computation by using the non-blocking MPI functions for data transfer. We do the data transfer between the CPUs at nodes concurrent with the computations at Xeon Phi coprocessors.

All the program code was written using C++ language.

Studying the Work of Parallel Algorithm on Xeon Phi Cluster

In this section, we present the results of the parallel algorithm behavior for a hybrid cluster architecture with Intel Xeon Phi coprocessor. We have carried out experiments on one computing device to choose appropriate options for large 3D models. We made a comparison of programs running on different computing devices for calculations that is using only CPUs or only Xeon Phi coprocessors. All the program codes were developed by the authors with the use of the proposed parallel algorithm and the designed parallel scheme. In addition, we present the results for different tests for the parallel program code. The first one is a scalability test that reveals

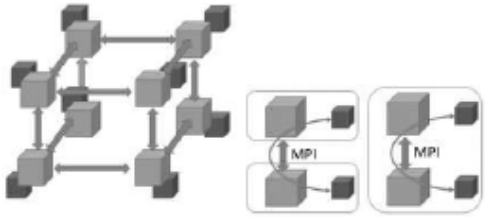

Fig. 1. A parallel computing scheme.

that the calculation time should not vary strongly if we do calculations taking into consideration the fact that the number of points in a 3D grid model grows proportional to the number of devices. This means that each computing device will do calculations for the same number of points. Another test is speed-up one. In this case, the number of points in a 3D model is fixed and we show the program behavior running on a different number of computing devices. All the results were carried out using the NKS-30T+GPU cluster of the Siberian Supercomputer Center (SSCC SB RAS), www2.sscc.ru, and the MVS-10P cluster of the Joint Supercomputer Center of RAS, www.jscc.ru. On one cluster node and on one device, we have carried out experiments with different options of affinity and a different number of threads per core. The 3D model under study has parameters 308x308x308 and 11 iterations. The affinity option has been taken in two versions: B«not declaredB» or B«balancedB». The results of such a research is presented at Fig. 2. The most appropriate is the B«balancedB» option and using 60 cores with a 3 threads per core.

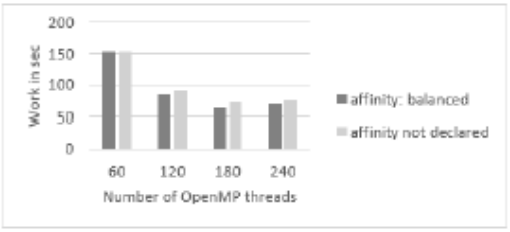

Fig. 2. A test with affinity option and the number of parallel threads on one device.

The performance of developed multi-coprocessor code is shown in Fig. 3 and Fig. 4. The results of scalability tests presented at figure show the well-done program behavior. When we scale a 3D model as great as 2-fold along each spatial coordinate and the number of devices as great as 8-fold, the program shows a good behavior. In these tests, we use a 3D subdomain size with 308x308x308 grid points and 11 iterations for one device. Maximum eight subdomains were used. In this case we have used 2 x 2 x 2 grid of computing devices. From Fig. 3 we can see that the program was effectively parallelized and the ratio of CPU/Xeon Phi is about x5.7.

The results of the speed-up tests presented in the Fig. 4 show the program behavior with 308x308x308 grid points and 11 iterations for all devices. Figure 4 reveals that the ration of

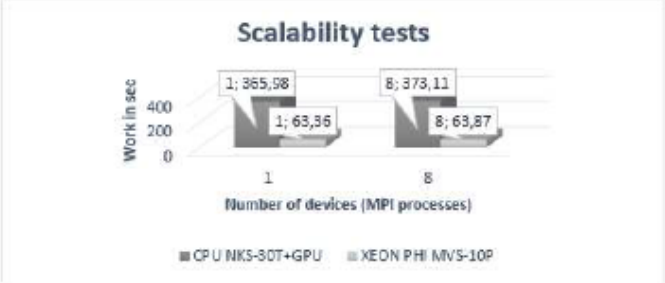

Fig. 3. A scalability test.

CPU/Xeon Phi is about x5.7 on one device and x3,6 other eight devices. The ratio of 1 to 8 devices for Xeon Phi is about x7.7 on eight devices.

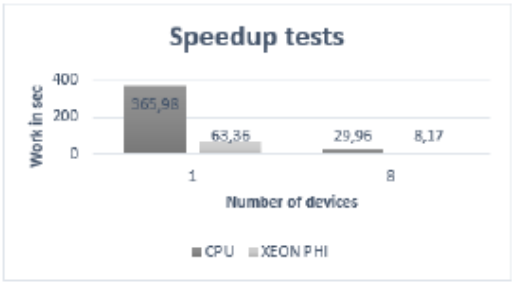

Fig. 4. A speedup test.

Based on the above-mentioned results, we conclude that carrying out computations on Intel Xeon Phi coprocessors for the large-scale seismic field simulation is a promising approach.

Conclusion

We presented the results of the research into developing a scalable parallel algorithm and program software. We proposed a new software for simulation of the elastic wave propagation in 3D isotropic elastic medium using hybrid supercomputers with Intel Xenon Phi coprocessors. We described the parallel implementation of the difference method based on 3D domain decomposition and using computing devices in offload mode. In the figures presented, the efficiency of a using such computing device for similar difference methods is shown. We have carried out computing experiments and investigated the behavior of the program on one Xeon Phi coprocessor to tune the script parameters to running the program for a greater number of computing devices placed at cluster nodes. It is shown that the 3D difference method with staggered grids can be well parallelized with Intel MIC architecture. We can use the discussed computing devices to simulate big size models. The results of the research done are important and can be of practical use in the field of developing scalable parallel algorithms for exzaflops

supercomputers [3] of the future and modeling its behavior on a greater number of computing cores in simulation systems.

Acknowledgments. This work was supported in part by the RFBR grants No. 14-07-00832, 14-05-00867, 15-07-06821, 15-31-20150 and MES RK 1760/GF4 and Intel Corporation.

References

1. Glinsky B.M., Karavaev D.A., Kovalevsky V.V., Martynov V.N: Numerical modeling and experimental research into B«Karabetova MountainB» mud volcano using vibroseismic methods. Vichislitelnie metody I Programmirovanie. Vol. 11, pp. 95–104 (in Russian)(2010)
2. R. W. Graves: Simulating seismic wave propagation in 3D elastic media using staggered grid finite differences. Bull. Seism. soc. Am., vol. 86, pp. 1091-1106 (1996)
3. A.R. Levander: Fourth-order finite-difference P-SV seismograms. Geophysics, vol. 53, issue 11, pp. 1425-1436 (1988)
4. Virieux J.: P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method . Geophysics, Volume 51,Number 4, pp. 889–901 (1986)
5. Moczo, P., Kristek, J. and Halada, L.: 3D fourth-order staggered-grid finite difference schemes: stability and grid dispersion. Bull. Seism. Soc. Am., Vol. 90, No. 3, pp. 587–603 (2000)
6. J.O. A, Robertsson, J.O. Blanch, and W.W. Symes: Viscoelastic finite-difference modeling. Geophysics, vol. 59, issue 9, pp. 1444-1456 (1994)
7. Dimitri Komatitsch: Fluid-solid coupling on a cluster of GPU graphics cards for seismic wave propagation. Comptes Rendus MГ©canique, Volume 339, Issues 2-3, pp. 125–135 (2011)
8. Dimitri Komatitsch / Dimitri Komatitsch, Gordon Erlebacher, Dominik GГ¶ddeke, David MichГ©a /: High-order finite-element seismic wave propagation modeling with MPI on a large GPU cluster. Journal of Computational Physics, Volume 229, Issue 20, pp. 7692–7714 (2010)
9. Landau L.D., Lifshitz E.M.: Theory of Elasticity. Third Edition, (1986)
10. Chernykh I., Glinskiy B., Kulikov I., Marchenko M., Rodionov A., Podkorytov D., Karavaev D.: Using Simulation System AGNES for Modeling Execution of Parallel Algorithms on Supercomputers. Computers, Automatic Control, Signal Processing and Systems Science. The 2014 Int. Conf. on Applied Mathematics and Computational Methods in Engineering, pp. 66–70 (2014)