

Software Solution for Direct Seismic Prospecting Task On Parallel Resources

A.Nedzved, A.Belotserkovsky, P.Lukashevich, A.Dosin

Abstract— Conceptual scheme of hybrid technology for solving the direct seismic task using parallel resources for normal incidence for plane-parallel layer of rocks is presented here. Proposed software architecture connects thin client part with HPC cluster on the basis of SKIF supercomputer. Resulting synthetic profile of seismograms allows to verify a proposed geological model visually in real time.

Keywords— seismic prospecting task, parallel resources, Python, MPI, cluster technology.

I. INTRODUCTION

One of important IT-task in geophysics is to create an intelligent computer environment for the formalization of professional knowledge in the field [1,2]. The developing of a conceptual model and tools helps to realize the process of preparing a geological model, gathering and analyzing data presented in collected materials. To implement such tasks, it is advisable to use computer interactive tools for modeling and design processes in the "human-computer" interaction by supercomputer facilities.

An improvement of known modeling methods is very crucial for the formalization of professional knowledge. Efficiency of technological processes and systems is supported by complexity and multivariate construction since organizing of computational processes is associated with the data processing when complex technological systems are designing. It requires interactive and graphical tools to organize the process. The choice of solutions is usually based on technological characteristics and parameters of the designed facilities [3-5].

It is known that such tasks are of high information complexity due to processed objects. It leads to the decomposition of the structures into component parts, with the separate processing of each part and their subsequent synthesis. High computational intensity and low level of formalization of initial data can be compensated with complexing of interactive tools with parallel computing.

II. DIRECT TASK OF SEISMIC EXPLORATION

The solution of direct task of seismic research on parallel resources for normal incidence for plane-parallel rocks layer is a good example for demonstration of application parallel resources of the cluster technology (such as SKIF supercomputer [7]). Direct task of seismic prospecting is the calculation of arrival times t and amplitudes A for a separate wave in a known seismogeological section. This task is solved when following data is known: density, depth of occurrence, size of certain geological objects (layers), speed of distribution of elastic waves and location, shape of a rock source [8,9]. A simple solution of direct dynamic seismic problems of inhomogeneous data is realized by solving of the following wave equation:

A. Nedzved, Belorussian State University, Minsk, Belarus, NedzvedA@tut.by.

A. Belotserkovsky, United Institute of Informatics Problems, Minsk, Belarus, Alex.belot@gmail.com.

P. Lucashevich, United Institute of Informatics Problems, Minsk, Belarus, pavel.lukashevich@newman.bas-net.by.

A. Dosin, United Institute of Informatics Problems, Minsk, Belarus.

$$\frac{1}{V^2} \cdot \frac{\partial A}{\partial t} = \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2},$$

where V – is velocity of wave, $A(t, x, y, z)$ – is magnitude of signal, propagation through a rock with coordinates (x, y, z) for different value of time t after initial stage.

The solution of the equation uses the boundary conditions. It can be implemented for simple rock models. In this way the kinematic problem is much simpler. It allows to determine the arrival time of a wave (direct, reflected, refracted, etc.) for a defined model. Such solution exists only for preset position of the source and the time of initiation of the elastic wave. Traditionally, the simplest result of solving of a direct task is obtaining by the hodograph equation. It is analytical expression for the function $t(x)$. As result the hodograph is a graph of time dependence of the arrival of a wave from the distance (x) . It is ranked from the source point to the point of reception.

Seismic research uses the method of reflected waves in a situation when the depth of rock layer is much bigger than the maximum distances between the source and the receiver in the case of normal incidence of waves. If the angles of incidence are not exceeding the critical value, the pulses of the exchange waves have the same shape as the pulses of the incident waves. Mixed and exchanged waves at normal incidence angle are not created. When decision is making, we generate a number of simplifying assumptions. However, the solution of the direct seismic survey task for normal incidence of waves in a thin-layered rock model improve studying of the wave characteristics of fields corresponding in real environments [9].

III. DATA PREPARATION FOR TRACE SIMULATION

The seismic geological model is constructed from seismic layers (seams). Its basic feature is the limitation in space by simple boundaries. It is interpreted as soles and roof of the layer. Such model differs from the enclosing layers with elastic properties. Physical properties of a homogeneous layer are constant. For gradient layers they have monotonic changings. In practice, homogeneous models of layers have usually priority in investigation. Thus, the simplest models of horizontal layers are mostly used. The sections of the sedimentary cover in many areas of seismic exploration are well approximated for them. For layered gradient rocks, practical applications are solved by models with gradients of vertical velocity that reflect the most general pattern of changes of elastic properties in sedimentary deposits [9].

Characteristics of the vertical coordinate of the layer determine the basic parts for the hodograph equation. These characteristics are determined from the lines of geological horizon. It limits the layer. Properties of lines are defined interactively by user at the “tablet” (main interactive windows for seismic data visualization) of the analysis of seismic profile, as shown in Fig. 1.

The line of horizon has following characteristics: name, velocity of wave propagation, density and absorption coefficient. The name of the horizon is an identification parameter, which is settled by the user when the model is prepared by researcher.

The density of the rock is also determined with additional geological prospecting data by the user. It includes characteristics that are obtained from analysis of a geological well. In this case the density is defined as the mass in a unit of volume for the solid phase of the rock. So, it affects to the velocity of wave propagation in the rock.

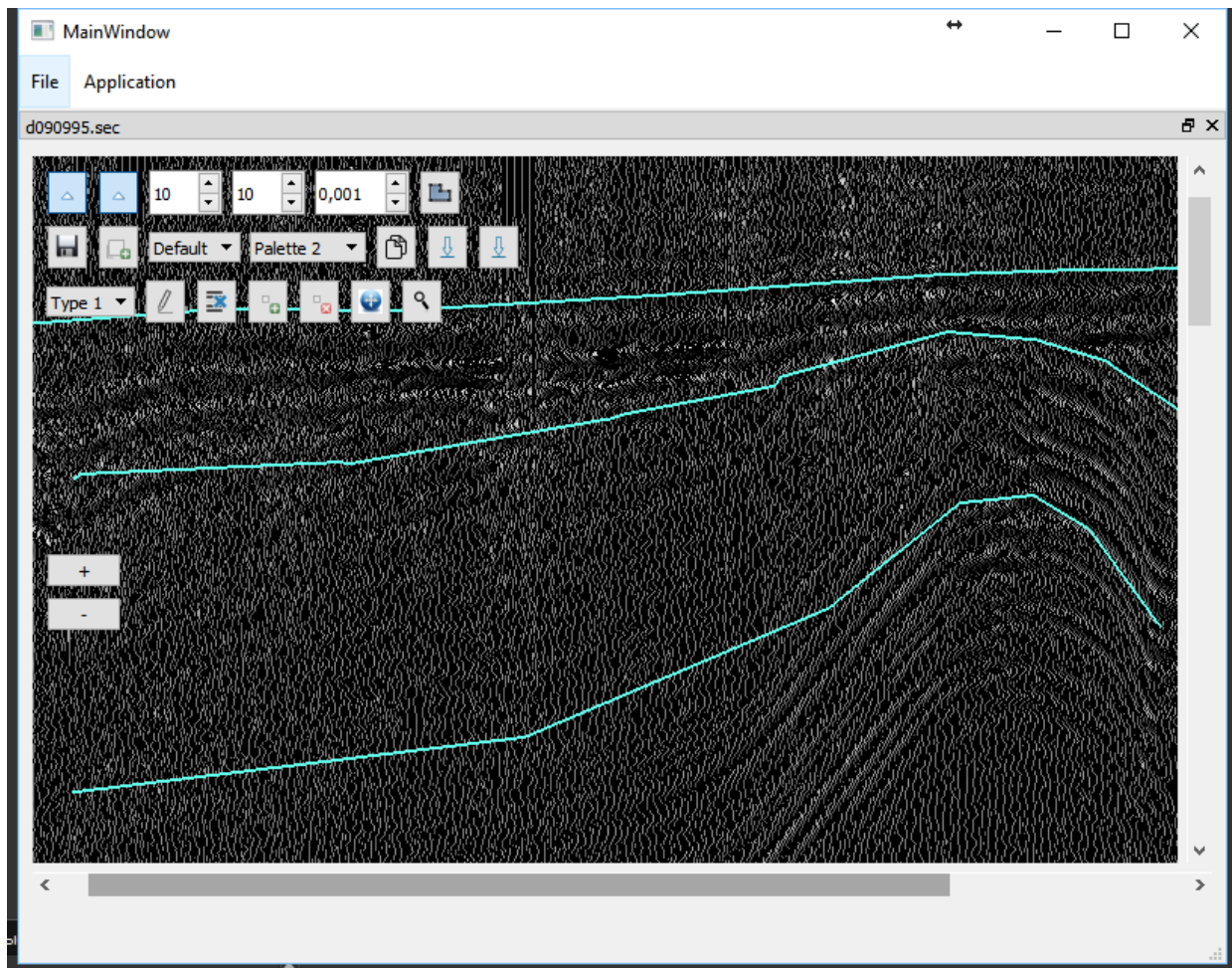


Fig. 1 Screenshot of the “tablet” with seismic profile where lines of limitation of layer is prepared

The time of the wave propagation in the rock is determined from point of its reflection in boundaries of the layer to the point of a recording sensor. In the time section, the reflecting horizon for bottom layer of the rock has a tracing time. Then, the depth from the horizontal boundary is calculated as the multiplication of half the tracing time to the wave speed, that is defined as average value of velocity of tracing from the boundary. The tracing time characterizes the depth of occurrence of the rock.

The absorption of seismic waves depends on the lithological composition and physical state of the rock. The absorption coefficient characterizes one of the physical properties of the rock and can be used for geological interpretation of seismic observations. In the algorithm for normal incidence for parallel plane layers, this coefficient is not used.

The next set of raw data is used to implement the solution of the direct seismic task at normal wave incidence. These are characteristics of impulse that are used to construct of syntetic trace. As a rule, it must correspond to the signal, which is used in geological research of the seismogram registration process. These impulses have different localization in the frequency-time domain. Therefore, the impulse is preliminarily chosen for calculation. It should look like as a research signal.

The simplest way to choose an impulse is a determination of the spectrum amplitude, which most completely coincides to the spectrum amplitude of the autocorrelation function of a real seismic trace for a zero-phase realization.

In our software, the following types of impulses are implemented for using as the initial seismic signal (impulse):

- Riker's momentum;

- Puzyrev's pulse;
- Sinc in Window;
- Decaying sinusoid;
- Damped cosine;
- Wavelets:
 - a) DOG;
 - b) WAVE;
 - c) MHAT ("Mexican hat");
 - d) Mople real (the real part of the complex Moplet).

In addition, the impulse can be simulated by calculation program. In this case, it is determined by a numeric vector or one-dimensional matrix. Basic impulses used in the system are shown at Fig. 2.

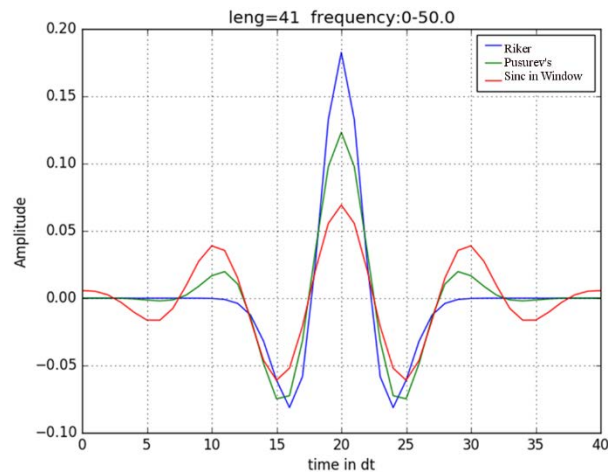


Fig. 2 Basic impulses for trace simulation

The next set of data determines the way of data processing for model trace simulation. It consists of a sequence of sets. These sets are describing the intersection points of the seismic trace and the lines of the horizons. Such set has digital data of characteristics and name of layer:

- the final processing time;
- step of discretization;
- impulse characteristics that determine its reading from a file or calculation set;
- wave characteristics for single and multiple reflection;
- the polarity of the mapping and the construction of the wave (right or left).

If it is possible to read a numerical series from the impulse section for its assignment, an impulse is formed from it. Otherwise, it is calculated on the basis of the characteristics of the seismogram. If the variable is equal to "All", all possible processing options are evaluated.

The last set of source data is the set of values of seismogram. This set includes general characteristics of traces and consists of the following data:

- The name of the seismic profile;
- The name of the picket or trace;
- The initial time for the trace;
- The final time for the trace;
- Step of discretization of the trace..
- Profile and picket names are string values.
- Other variables are characterized by numerical values. They are specified in the specific header before a set of discrete values, which correspond to the values of the trace.

In the software the special function is realized. It provides analysis of these data and constructs specific structure that allows us to divide data for parallel processing. Also it is shows in specific window with structures (Fig. 3).

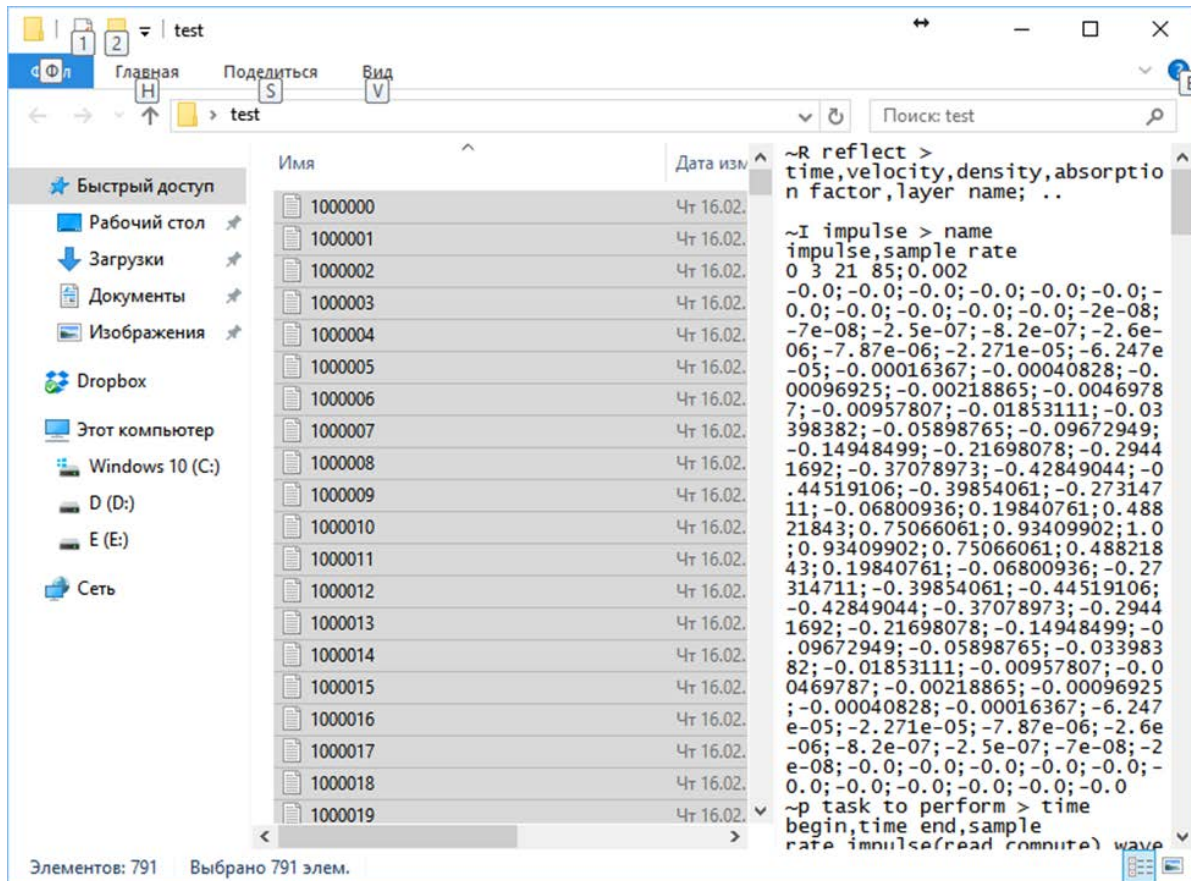


Fig. 3 Special window with data structure for parallel processing

This information describes only geological content that is used for construction of new profile of synthetic seismograms. It uses a special representation of the information, which is broken down into separate traces. As a result, the synthesis can be performed for each individual trace. Then the result can be combined in the profile of synthetic seismograms.

IV. PARALLEL SOLUTION FOR THE SEISMIC SURVEY TASK

Running a software for seismic prospecting on server will bring a delay of graphics activation due to a transfer rate of graphic content for seismogram. Using a network with common topology and average characteristics the visualization has meet a speed equal to 3fps. On the other hand, there are long delays in simulating of seismic traces, if the software is started from client side. Thus, the calculations of the traces simulation and interactive processing at graphics of seismograms profile representation can be executed on server and on client parts separately.

The basic problem of such realization is separation of information flows. The solution of the seismic task on parallel resources is executed at the computing architecture of SKIF family (cluster technology). It uses an operating system that based on the CentOS linux kernel. In such system the access to the cluster is organized via SSH protocol. To organize such access, the software used libssh 0.3 library.

The libssh library is designed to integrate supporting of SSH protocols into C source code. The library allows to run programs on a remote server, copy files through a secure connection. The library supports both standard scp and sftp transfer mechanism. It is implemented without external libraries, such as libcrypto (from OpenSSL) or libcrypt. Therefore, such developing of software is easier.

Libssh implements server and client functions, supports SSH1 and SSH2 protocols, AES-128, AES-192, AES-256, Blowfish, 3DES encryption modes in CBC, the ability to work with RSA and DSS public key servers. It supports data compression, co-working with ssh-agent and the organizing several sessions within a single connection.

The QSshSocket class is also implemented to interact with the software “tablets”. It is designed to create an asynchronous software interface for data exchange (socket) using the SSH and SCP protocols. The class is specially implemented to use it with the Qt library, which is cross-platform and corresponds to logics of ”tablets”. The class effectively uses the functions of the libSSH library.

The QSshSocket class is a child of the standard QThread class from the QT library. In addition to the standard features, it includes an additional set of constants and functions.

These functions allow the client and SKIF clusters [10] to interact through data flow (Fig. 4).

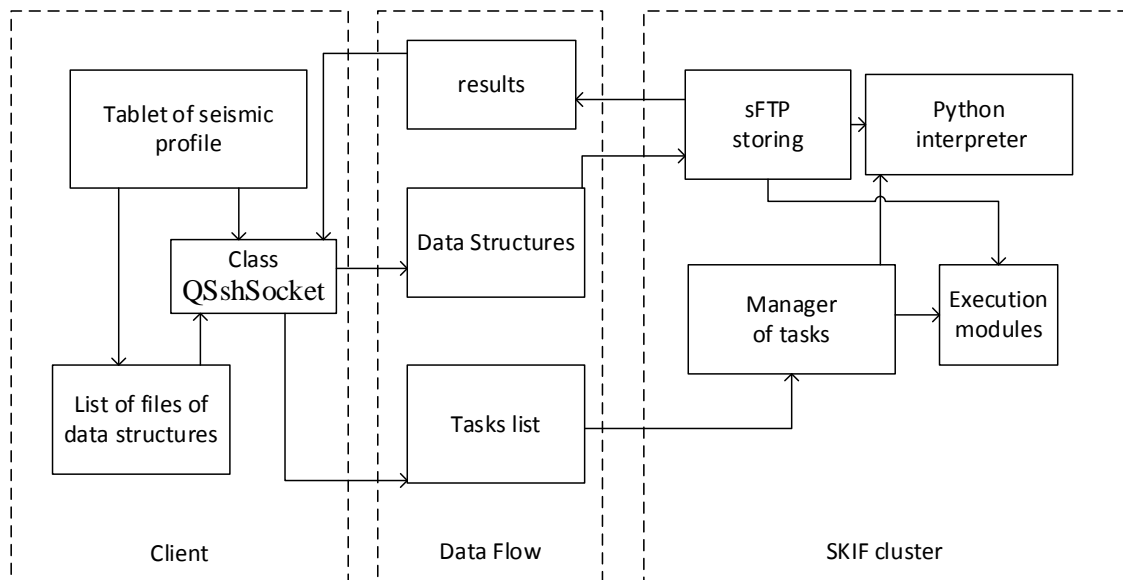


Fig. 4 Scheme of implementation of software for geological interpreting by simulation of direct seismic task

According to the scheme we divided the graphic and computational part. It allowed us to accelerate the interactive work with a graphical representation of the seismic profile.

A process management was realized by the Portable Batch System (PBS). The PSB package is designed to manage computers and system resources. It allows to increase an effectiveness of managing of calculation resources by refusing equipment downtime. PSB generates sets of task for execution, it stores, starts and controls of execution every task and sends information about implementation of calculation processes to user.

The clusters SKIF use free version of PBS – Torque, which is based on famous system OpenPBS. It used Maui package as free manager of tasks. PSB’s directives are transferring to computational system by separated command QSUB or through execution script. The software uses following features:

1. Starting of execution software from command prompt with argument, for example:

```
qsub -l walltime=00:20:00,nodes=3:mpp /opt/scali/bin/mpirun -machinefile nodes.sca \
$PBS_O_WORKDIR/python theogram.py 3806.txt.
```

This sample launches a python script to calculate a synthetic seismogram for separate picket.

2. Creating script of PBS for starting package of software and commands have parameters of tasks. They can be set by a string: “qsub <script_name>”, where <script_name> is name of script. For execution that string is generated by client part of software. This script is created by client part of software (“tablet” of seismogram profile). Commands from script are queued up as jobs in "dteam" ("batch" on the cluster).

In the simplest case, when a uniprocessor program is started on a Linux cluster, a script file is created with a set of commands as follows:

```
### PBS script to run one CPU task ###
#PBS -l walltime=00:10:00
#PBS -l nodes=1
#PBS -N python theogram.py
hostname
cd ~/test_pbs/
date >> ~/test_pbs/log
hostname >> ~/test_pbs/log-
./python theogram.py 3806-0001.txt
./python theogram.py 3806-0002.txt
./python theogram.py 3806-0003.txt
.....
.....
.....
./python theogram.py 3806-0748.txt.
```

This script initiates an environment of parallel processing by creating tasks' sequences to calculate a set of synthetic seismograms. A few variants of calculation are to be started by the next command:

```
qsub -l walltime=00:20:00,nodes=3:mpp /opt/scali/bin/mpirun -machinefile nodes.sca \
$PBS_O_WORKDIR/python mergetheogram.py 3806-0001.txt 3806-0748.txt.
```

As a result, a new profile of synthetic seismogram is created. A seismogram for every picket is calculated as a separate process at separate node. The same sequence of command is used to calculate each seismogram with corresponding initial data for each picket.

Seismograms from pickets are processed on the server and the software generates a folder with solution which is necessary to verify. It allows to sort resulting data by picket number. Calculated data sequence is merging to general array of seismogram and representing as new synthetic profile. Several synthetic profiles for different phases or impulses can be built. These profiles are sent to client part of software by SSH protocol and shown in separate graphics windows (Fig.5).

V. MERGING SYNTHETIC SEISMOGRAM OF PICKETS TO PROFILE

The United Institute of Informatics Problems of the National Academy of Science has organized the “State supercomputer multi-access center” (SSMAC), which includes the computational capabilities of cluster supercomputer family. The high speed approach to the resources connected to telecommunication network of the National academy of Sciences of Belarus BAS-NET from scientific networks that is connected to BAS-NET with fiber-optic. The network interaction of remote users with SSMAC is organized by protocol SSL.

We have tested our software by three cases of equipment:

- Desktop workstation;
- SKIF cluster by Xming that is an X Server running on Microsoft Windows;
- Hybrid scheme where interactive graphics is realized on PC-client and calculation on SKIF supercomputer.

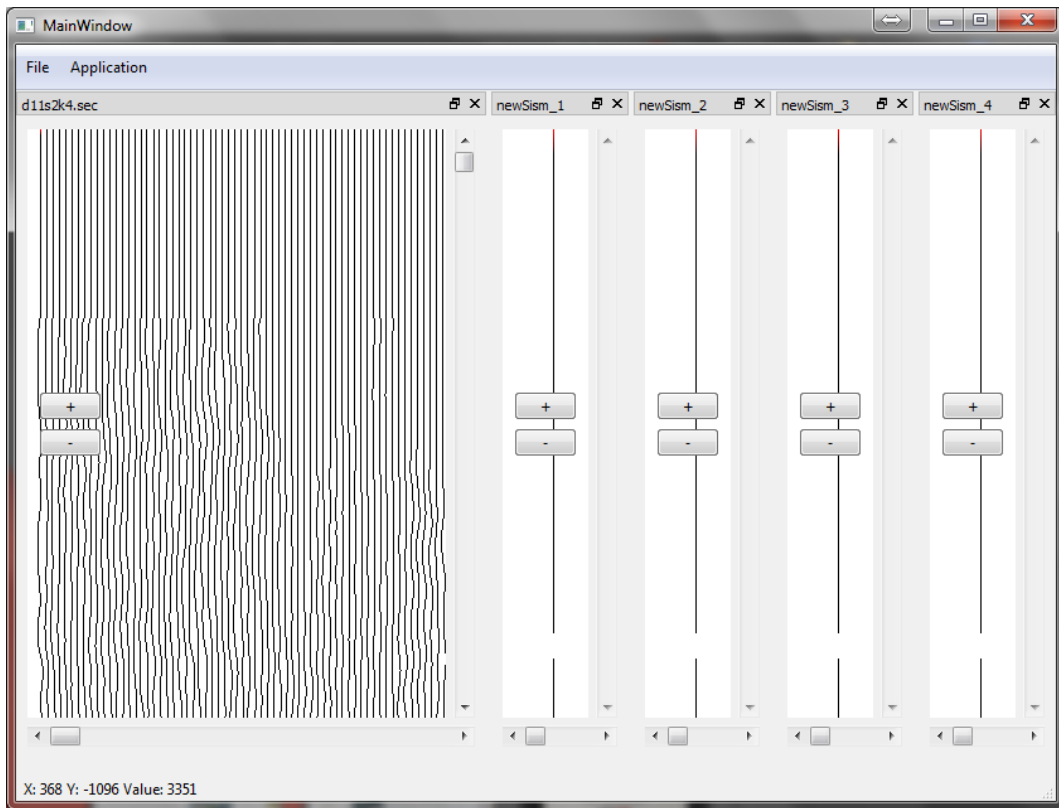


Fig. 5 The set of window with real and synthetic seismogram after seismic simulation.

The test was executed by specific sample. The source data for testing include real profile of seismogram, nine lines of geological model and sets of basic impulses. The profile of seismogram includes traces for 1240 pickets with 3000 discrete values. Geological model has about nine points of rocks layer for every picket. Sets of impulse include Riker's momentum. We use simple algorithm to solve a task for normal incidence for plane-parallel rocks layer. It is based on signal deconvolution.

We estimate two parameters: frame rate and speed of calculation. The frame rate is defined as count of frames per second (fps). This parameter allows to estimate properties of “tablet” of seismic profile.

The speed of calculation is defined through a time of script execution to solve a task for normal incidence for plane-parallel rocks layer.

The result of testing is represented in Table 1.

TABLE I
Units for Magnetic Properties

Type of platform	Execution marks		
	FPS	Time of estimation (s)	Common mark
Desktop	30	220	0.13
SKIF cluster	5	35	0,14

Type of platform	Execution marks		
	<i>FPS</i>	<i>Time of estimation (s)</i>	<i>Common mark</i>
Hybrid scheme	30	61	0,41

We define common mark of effectivity of software execution as relation FPS to time execution. Test show that hybrid technology is very effective for comfort of geological research of seismic data.

VI. CONCLUSION

In this paper, the hybrid technology is described for solving of direct seismic task at parallel resources for normal incidence for plane-parallel layer of rocks. The proposed software environment connects a “tablet” of seismogram profile for interactive modeling with calculations at server. It uses the SSH protocol for communication issue. The computational tasks are solved using the Python interpreter, which is located on the cluster of the SKIF supercomputer. The solution of the seismic problem is divided into tasks for each trace or picket for parallel execution at nodes of the supercomputer. For each task, a file and a command line are generated, which are transmitted by the SSH protocol to the cluster for execution. For finishing, the general solution of the task merged all result. As a result, a new synthetic profile of seismograms is created for visual verification of the geological mod

ACKNOWLEDGMENT

The research is supported by Joint Russian-Belarusian Scientific and Technical Programme “Research and Development of High-Performance Information and Computing Technologies for Increasing and Effective Usage of Resource-Intensive Hydrocarbon Resources of Union State (SKIF-NEDRA)”.

REFERENCES

- [1] Ponsa X., Masó J. A comprehensive open package format for preservation and distribution of geospatial data and metadata // *Computers & Geosciences*, vol. 97, 2016, pp. 89–97.
- [2] Vichalai C., Huy Giao P., Puttiwongrak A. Integrated analysis of high resolution seismic and geotechnical data using opensource software for an offshore drilling rig site // *Symp. on Imaging and Interpretation*, 2009, pp. 5-9.
- [3] Sato H, Fehler MC, Maeda T. *Seismic wave propagation and scattering in the heterogeneous earth*, 2nd edn. Berlin, Germany: Springer. 2012, P. 184
- [4] Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. // *Geophysics*, vol. 74, pp. 2009, pp.1-26.
- [5] Graves, R. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. // *Bulletin of the Seismological Society of America*, vol. 86, 1996, pp.1091-1106.
- [6] Ablameyko S., Abramov S., Anishchanka V., Paramonov N., Tchij O. 2008. SKIF Supercomputers and Their Applications // *Problems of Cybernetics and Informatics: proceedings of the Second International Conference*. September 10 -12, 2008, Baku, Azerbaijan. – Baku, Azerbaijan National Academy of Sciences, 2008, pp.3-7.
- [7] Avseth, P, T Mukerji and G Mavko. *Quantitative seismic interpretation*. // Cambridge University Press, Cambridge, 2005, p. 183
- [8] Louie J.N.; Pullammanappallil S. K. Advanced seismic imaging for geothermal development // *New Zealand Geothermal Workshop 2013 Proceedings*. 2013. pp.1-7.
- [9] Laine J.M. Seismic interpretation and stratigraphy using opensource software // *Proc.of Geological Society of America Abstracts with Programs: Vancouver Convention Centre-West*, vol. 46, no. 6, 2014, pp. 356.
- [10] Nedzved A., Belotserkovsky A., P. Lukashovich Building the regional system for seismic data processing // *Proc.on Information and Digital Technologies 2015 International Conference*. Zilina, 2015, pp. 261-265