PARALLEL DISTRIBUTED SEISMIC IMAGING ALGORITHMS ON PARAM 10000

S. PHADKE, R. RASTOGI, S. YERNENI and S. CHAKRABORTY

Centre for Development of Advanced Computing, Pune University Campus, Pune 411007, India

Summary

Obtaining high-resolution images of the underground geological structures using seismic reflection data in prestack or poststack domain is crucial for exploration of oil and gas deposits. In the last decade the development of parallel distributed computing platforms, related system software and programming environments have made it possible to use parallel codes for high resolution imaging. Centre for Development of Advanced Computing (C-DAC) located at Pune, developed the OpenFrame architecture for scalable parallel computing applications. Several seismic migration and modelling algorithms were developed and implemented for imaging purposes. In this presentation we shall discuss several wave equation based 2D and 3D seismic migration and modelling algorithms and their parallel implementation using MPI message passing environment. Large-scale problems can be solved by implementation of highly efficient and scalable codes. These codes can be easily ported across cluster of workstations.

Introduction

High-performance computers are now essentials tools in scientific and technological research and development. With their high-speed processing capability, large-scale storage capacity and efficient I/O, computers are now important tools for simulation experiments and for processing large volumes of data. Future developments in high-speed and large-scale supercomputers will play a significant role in the research and development of advanced technology for the 21st century. Parallel processing is the key technology to make large-scale processing capability possible.

Seismic imaging is a form of echo-reconstructive technique based on experiments, in which a certain earth volume is illuminated by an explosive or vibratory source, and the backscattered energy by the inhomogeneties of the medium is recorded on the surface in digital form. The inhomogeneties act as reflecting surfaces, which cause signal echoing; the echoes are then recorded at the surface and processed through a “computational lens” defined by a propagation model to yield an image of the inhomogeneties.

By far the widest commercial application of non-intrusive imaging and that, for which the algorithms are most sophisticated, is seismic exploration for oil and gas. The seismic experiments are generally based on wave propagation, where ray paths are strongly curved by variations of compressional and shear wave velocities with depth. The wavefield can be trapped and multiply reflected between lithological layers. By solving the one-way scalar wave equation or full scalar wave equation with recorded data as the initial condition, we solve for the complexities of the wave propagation, leading to an image of the reflecting surfaces (interfaces in the velocity field). This technique is known as seismic migration.

Forward modelling, where the synthetic data is generated for a given earth model, is a key step in the process of seismic inversion, where one tries to estimate the physical properties of the earth. 80 to 90 percent of the computer time in an inversion algorithm is spent on generating synthetic data. Parallel and efficient algorithms are therefore necessary for this purpose.

In this paper we shall describe several migration and modelling algorithms that are developed and parallelized for a distributed memory machine. Performance and efficiency is achieved by proper restructuring of the codes. All the imaging algorithms have been tested for both synthetic and real data sets.

Parallel Computing and Seismic Data Processing

There is a strong consensus amongst the computer professionals, that the greatest gains in price/performance can only be achieved through multiple processor parallel systems. Parallel computers are characterized by two or more processing elements and memory, tied together by some interconnection network. Abundance of relatively slow processors, working together to solve one problem, provides the necessary performance.

The trend in parallel computing is to move away from specialized traditional supercomputing platforms, such as Cray / SGI T3E, to cheaper and general purpose systems consisting of loosely coupled components built up from single or multiprocessor PCs or workstations. This approach has a number of advantages, including being able to build a platform for a given budget, which is suitable for a large class of applications and workloads.

The hardware technology and economic forces are right for an explosion of parallel processing into the market at all levels. Parallel processing, or concurrent computing as it is sometimes termed, is not conceptually new. The jobs that can be broken into multiple tasks that in turn be handed out to individual workers for simultaneous execution, are most suitable for parallel machines.

Recently, cluster of workstations or network of workstations has gained popularity as they provide a very cost-effective parallel-computing environment. Most of these clusters use Network File System (NFS) and MPI (Message Passing Interface) as message passing library. MPI calls allow us to communicate and synchronize between the processors. One limitation of NFS is that the I/O nodes are driven by standard UNIX read and write calls, which are blocking requests. This is not a problem for applications with small volume of I/O, but as the volume increases (as in 3D seismic acquisition), it is necessary to be able to overlap computations with the I/O to maintain efficient operation (Olfield et al., 1998, Poole, 1994). In the present study we have used both MPI and MPI I/O to improve the performance and efficiency of the codes (Bhardwaj et. al. 2000).

Conceptually, MPI consists of distributed support software that executes on participating UNIX / LINUX hosts on a network, allowing them to interconnect and cooperate in a parallel distributed computing environment. MPI offers an inexpensive platform for developing and running application. Heterogeneous machines can be used in a networked environment. The MPI model is a set of message passing routines, which allows data to be exchanged between tasks by sending and receiving messages.

Seismic Data Processing occupies a significant role in the exploration of oil and natural gases. Over the last two decades the computational requirements of the SDP activities have grown up many folds due to the increase in the data volume as well as the development in the mathematical algorithms. Three dimensional data acquisition has become routine as it has become necessary to look at the minor details of the underground geology.

Wave equation based methods (Phadke et.al. 1998) are gaining more and more popularity in recent years as they provide finer detailed geological features than other conventional methods as well as they preserve amplitude information. Advanced techniques are distinguished primarily by their use of wave equation. The most common advanced techniques include seismic migration and forward modelling. Finite difference methods are most suitable for migration and modelling as they offer most direct solution to the problem in terms of the basic equation and initial and boundary conditions.

By nature most seismic problems carry an inherent parallelism in subdivision by source, receivers, frequency or wave number. Indeed the problem decomposition in several domains is possible. With the change in demand it has become very difficult for a processing facility build around a serial architecture machine to cope up with increase in data volume. The I/O problems are also better solved in parallel processing. The wave equation based methods are computationally more expensive but suitable for parallelization. The seismic processing industries all over the world have found parallel processing as the only solution to the challenges in probing the earth's interior for natural resources.

The digital data that needs to be processed before obtaining an interpretable image of the subsurface geological structures in enormous, amounting to 100s of GB (Giga Bytes) or a few TB (Tera Bytes) for 3D acquisition. All this numerical input will be passed perhaps 10 to 20 times through a major computer facility, and only after the complex numerical operations, the final processed sections are examined by geophysicists and geologists to formulate an initial or penultimate interpretation. Parallel processing is the only answer to cope with increase in data volume and changes in processing methodology. we are fortunate that Seismic Data Processing (SDP) is an ideal application for parallel architecture machines.

Migration Algorithms

The stacking of seismic data is a form of data compression, which improves signal-to-noise ratio and produces idealized seismic traces simulating a coincident source-receiver experiment. Migration of the resultant data set, called the zero-offset seismic section or the post-stack time section, is known as post-stack migration. Migration can also be carried out in the prestack domain and the results obtained are more accurate than that of poststack domain. However the computational requirements of prestack migration algorithms is orders of magnitude more than that of poststack migration algorithms. Processor speed, memory and I/O play a crucial role in the implementation of these algorithms.

Most of the migration methods comprise of two steps, extrapolation and imaging. In the extrapolation step the wavefield is downward continued using some form of the acoustic wave equation. At each depth the image is formed at t = 0. The extrapolation of the wavefield can be carried out in t-x-y, ω-x-y or ω-k_x-k_y domain. Here we shall describe the implementation of migration in ω-x-y and ω-k_x-k_y domains. Another technique, Reverse Time Migration (RTM), which makes use of the full wave equation is also developed and implemented on PARAM.

3D Depth Migration in ω-x-y domain

For 3D depth migration, the extrapolation equation in ω-x-y domain is a parabolic partial differential equation (Claerbout 1985) consisting of a diffraction term and a thin lens term. The thin lens term, which accounts for lateral velocity variations, is usually ignored in time migration. The diffraction term is numerically solved by the method of splitting, which is the basis for the onepass approach. A Crank-Nikolson finite difference scheme with absorbing boundary conditions on the sides of the model is used for the solution. The thin lens term is solved analytically. Imaging is the summation of all the frequencies at t=0 for each depth.

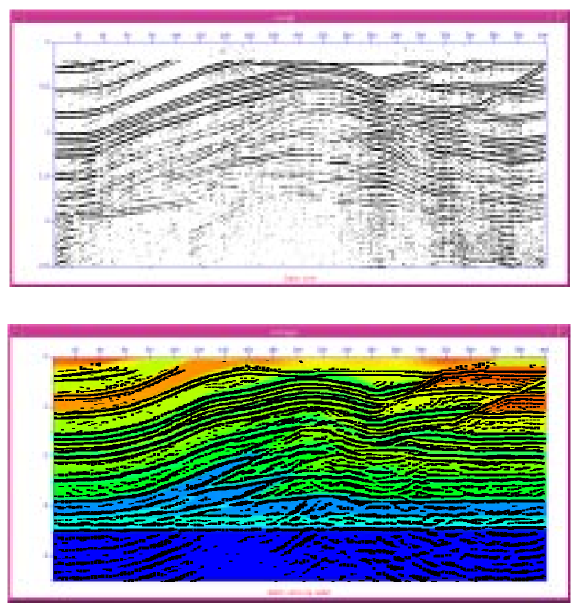

Figure 1: (a) Zero-offset section of a line from 3D volume of SEG/EAGE overthrust model. (b) 3D depth migrated section. The velocity model is also superimposed on the migrated section.

The depth migration algorithm in ω-x-y domain is inherently parallel in terms of frequencies. The parabolic approximation of the wave equation in frequency-space domain has decomposed the wave field into monochromatic plane waves that are propagating downwards. Therefore, each frequency harmonic can be extrapolated in depth independently on each processor and there is no need of inter-task communication.

One can introduce parallel task allocation into each frequency harmonic component with the ultimate goal being to have as many processors as frequencies. At each depth step all frequency components after extrapolation are summed up (Imaging Condition) to give the migrated image. The summation is carried out by automatic merging using MPI_Reduce. MPI I/O is used for reading and writing input data, velocity data and output data.

We first tested the migration algorithm for the data set of SEG/EAGE (1997) Overthrust model. The original data had 101X25 CDP traces with inline spacing of 100m and crossline spacing of 100m. We interpolated this data volume to 401X97 CDP traces to make both inline and crossline spacing 25m for avoiding spatial aliasing. The input Fourier Transformed data size was of the order of 46MB. This data set was migrated with a depth step of 25m for 161 depth steps. Figure 1 shows the zero-offset section for one of the lines and the 3D migrated data for the same line. The velocity model is also superimposed on the migrated data to show the accuracy of the migration algorithm. Figure 2 illustrates the execution time as a function of number of processors. Since the problem size is small the speedup is not linear.

The second data set used for testing comprised of 950X665 CDPs. The inline spacing was 25m, the crossline spacing was 37.5m, and the depth step size was 12.5m. The data was migrated for 480 depth steps. Table 1 shows all the other parameters and the time required to migrate this data set with 64 processors. It is not possible to carry out a speedup analysis on this data volume, since there is not enough memory available on a smaller number of processors and the execution time required will also be very large.

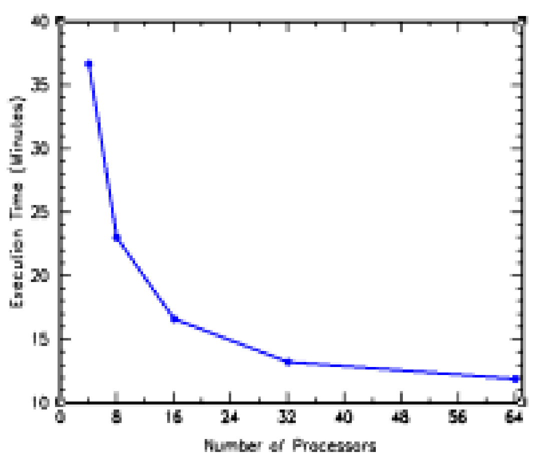

Figure 2: Number of processors versus execution time chart for SEG/EAGE Overthrust model.

Size of FFT data	1.3 GB
Size of Velocity model	1.2 GB
Frequency band	5 - 40 Hz
Number of Processors	64
Total Execution time with MPI-IO	7 hrs 44 mins

Table 1: Problem size for the second data set and the execution time on 64 processors.

3D Depth migration with PSPI Algorithm

The phase-shift migration method (Gazdag, 1978) downward continues the wavefield in wavenumber-frequency domain, under the horizontally layered velocity assumption. If the migration velocity has no horizontal variations, the phase-shift method extrapolates the wavefield exactly by rotating the phases of each Fourier component. In the presence of lateral velocity variations, the exact extrapolation equation is no longer valid. PSPI (Phase Shift Plus Interpolation) method circumvents the problem of lateral changes in migration velocity by downward extrapolating the wavefield with several reference velocities and then interpolating the wavefield for the correct velocity (Gazdag and Sguazzero, 1984).

The parallel implementation of the PSPI method is also straightforward. The method is inherently parallel in terms of frequency. Here also the data if first Fourier transformed and then different processors read and migrate their share of frequencies. At each depth step phase-shift are applied for the reference velocities and then wavefield is interpolated for the actual velocity. One of the processors, which act as the master, collects and images the data. The method was developed and implemented on PARAM 10000 and was tested by applying it to both synthetic and real data sets.

Reverse Time Migration (RTM)

Reverse time migration technique solves the full wave equation by extrapolation in time, allowing both the upgoing and downgoing wave to propagate. The full wave equation is solved using finite-differences and the wavefield recorded at the surface is used as boundary condition. McMechan (1983) has given the description of the method in detail and demonstrated its ability to image all dips with great accuracy. Time marching of the wavefield is similar to any modelling algorithm. The parallelization is carried out using domain decomposition scheme. A good description of wave propagation using finite differences is given in the next section on modelling algorithms.

RTM has the same problems with the stability and numerical dispersion that finite-difference (FD) modelling has, and it is straightforward (but computationally expensive) to control these problems. We have implemented a central difference FD scheme for RTM on PARAM 1000 using domain decomposition. The application of the method to both the synthetic and real data sets will be shown during presentation.

Modelling Algorithms

A basic problem in theoretical seismology is to determine the wave response of a given earth model to the excitation of an impulsive source by solving the wave equation. In scalar approximation, the acoustic wave equation may be solved to evaluate the waveform but only compressional waves are considered. A more complete approach is to study the vector displacement field using the full elastic wave equation for modelling both, compressional waves and shear waves. However, important wave properties such as attenuation and dispersion require a more sophisticated set of equations. These properties will be incorporated in the future versions of codes.

2D Acoustic / Elastic Wave Modelling

The mathematical model for elastic wave propagation in 2D heterogeneous media consists of coupled second order partial differential equations governing motions in x- and z-directions

$$\rho \frac{\partial \dot{u}}{\partial t} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \quad (1)$$

$$\rho \frac{\partial \dot{w}}{\partial t} = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \quad (2)$$

and the stress-strain relations are given by

$$\sigma_{xx} = (\lambda + 2\mu)\frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial z} \quad (3)$$

$$\sigma_{xz} = \lambda\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) \quad (4)$$

$$\sigma_{zz} = \lambda \frac{\partial u}{\partial x} + (\lambda + 2\mu)\frac{\partial w}{\partial z} \quad (5)$$

Where u and w are horizontal and vertical displacements, $\dot{u}$ and $\dot{w}$ are the horizontal and vertical particle velocities, σ_{xx} , σ_{zz} and σ_{xz} are the stress components, λ and μ are the Lamé parameters and ρ is the density.

Instead of solving these second order coupled partial differential equations we formulate them as a first order hyperbolic system (Virieux 1986, Vafidis 1988, Dai et al. 1996):

$$\frac{\partial Q}{\partial t} = A \frac{\partial Q}{\partial x} + B \frac{\partial Q}{\partial z} \quad (6)$$

Where,

$$Q = \begin{bmatrix} \dot{u} \\ \dot{w} \\ \sigma_{xx} \\ \sigma_{zz} \\ \sigma_{xz} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 & \rho^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho^{-1} \\ \lambda + 2\mu & 0 & 0 & 0 & 0 \\ \lambda & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 \end{bmatrix}$$

and $$B = \begin{bmatrix} 0 & 0 & 0 & 0 & \rho^{-1} \\ 0 & 0 & 0 & \rho^{-1} & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & \lambda + 2\mu & 0 & 0 & 0 \\ \mu & 0 & 0 & 0 & 0 \end{bmatrix}$$

When we move from elastic to acoustic media, the value of μ becomes zero. By substituting $\mu = 0$ in the above equation we get a first order system of hyperbolic partial differential equations which governs the acoustic wave propagation.

$$Q = \begin{bmatrix} p \\ \dot{u} \\ \dot{w} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & K & 0 \\ \rho^{-1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 & K \\ 0 & 0 & 0 \\ \rho^{-1} & 0 & 0 \end{bmatrix} \quad (7)$$

Where p is the negative pressure wavefield, and $K=\lambda$ is the incompressibility.

For solving the first order hyperbolic system (6), we use the method of splitting in time (Vafidis 1988). An explicit finite difference method based on the MacCormack scheme is used for the numerical solution (Mitchell and Griffiths, 1981). This scheme is fourth order accurate in space and second order accurate in time. The model discretization is based upon regular grid. Sponge boundary conditions are used for attenuating the reflected energy from the left, right and bottom edges of the model (Sochaki et al. 1987). Free-surface boundary condition is used for top edge.

The parallel implementation of an algorithm involves the division of total workload into a number of smaller tasks, which can be assigned to different processors and executed concurrently. This allows us to solve a large problem more quickly. The most important part in parallelization is to map out a problem on a multiprocessor environment. The choice of an approach to the problem decomposition depends upon the computational scheme. Here we have implemented a domain decomposition scheme.

The idea of this scheme is simple. First, the problem domain is divided into a number of subdomains that are assigned to separate processors. The upper part of Figure 3 shows an example of the division of problem domain into nine subdomains. Depending upon the number of available processors and the problem, one can divide the problem domain into any number of subdomains. Since MacCormack scheme uses a nine-point difference star, the calculation of the wavefield at an advanced time level for any grid point, requires the knowledge of the wavefield at 9 grid points of the current time level. For grid points along the boundaries of the subdomain, the information about the neighbouring grid points comes from the adjacent subdomains. Therefore after each time step the subdomains have to exchange some wavefield data. Lower part of Figure 3 shows the required memory space for each 2D array of the subdomain and the communication between two adjacent subdomains. The data in the darker region is sent to the lighter region of the neighbouring subdomian using MPI message passing calls.

The two most important issues in this implementation are (1) to balance workload (2) to minimize the communication time. In a homogeneous multiprocessor environment, as in our case, the load balancing is assured if all the subdomains are of the same size. Minimizing the perimeters of the subdomain boundaries minimizes communication.

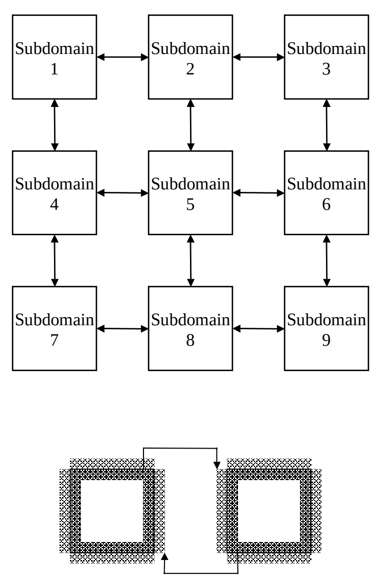

Figure 3: The upper picture shows the division of problem domain into a number of subdomains. The lower picture shows the communication between two adjacent tasks.

In the MPI implementation of the modelling codes there is a master task and there are a number of worker tasks. The main job of master task is to divide the model domain into subdomains and distribute them to worker tasks. The worker tasks perform time marching and communicate after each time step. As per the requirement of the user the snapshot and synthetic seismogram data are collected by the master and written out on the disk.

The wave propagation described by equation 6 is valid for both acoustic and elastic media. This is because when the Poisson's ratio becomes 0.5 the medium becomes acoustic (Phadke et. al. 2000). The upper part of Figure 4 shows the P-wave velocity model used for calculating the synthetic data in a marine environment. There is a water layer at the top. The water bottom is quite undulating. Poisson's ratio and density in other layers are 0.25 and 2.2gm/cc respectively. The snapshots of the wave propagation through this model are also shown in Figure 4. The synthetic seismogram data for this model are shown in Figure 5. A gain function is applied for display purposes. Since the free-surface boundary condition is used for the top edge, all kinds of multiples are also modeled. The example clearly demonstrates the capability of this approach for generating synthetic seismograms in realistic marine models. Another advantage of this approach is that acoustic and elastic wave propagation is modeled by the same code ($\mu = 0$ for acoustic wave propagation).

3D Acoustic wave Modelling

The acoustic wave equation in a 3D heterogeneous medium is given by

$$\frac{1}{K}\frac{\partial^2 p}{\partial t^2} = \frac{\partial}{\partial x}\left[\frac{1}{\rho}\frac{\partial p}{\partial x}\right] + \frac{\partial}{\partial y}\left[\frac{1}{\rho}\frac{\partial p}{\partial y}\right] + \frac{\partial}{\partial z}\left[\frac{1}{\rho}\frac{\partial p}{\partial z}\right] \quad (8)$$

Where, *p* is the negative pressure wavefield, , ρ is the density and *K* is the incompressibility.

We divide the 3D geological model into a grid of I X J X K points. In order to obtain finite difference approximation to equations (1), let us introduce a set of indices *i, j, k* and *n* such that

$$x = i\ \Delta x \quad i = 0,1,2,\ldots\ldots I, \qquad y = j\ \Delta y \quad j = 0,1,2,\ldots\ldots J$$

$$z = k\ \Delta z \quad k = 0,1,2,\ldots\ldots K, \qquad t = n\ \Delta t \quad n = 0,1,2,\ldots\ldots N$$

where Δx, Δy and Δz are the grid spacing and *I ,J* and *K* are the number of grid points in x- y- and z- directions respectively, Δt is the time step and *N* is the total number of time steps. Physical parameters, density ρ(i,j,k) and incompressibility *K*(i,j,k) are specified at each grid point.

Substituting central difference approximations of the derivatives in equation (1), an expression is obtained for calculating the wavefield $p^{n+1}_{i,j,k}$ from the knowledge of the wavefield at previous time levels i.e. $p^{n}_{i,j,k}$ and $p^{n-1}_{i,j,k}$ as

$$\begin{aligned} p^{n+1}_{i,j,k} = {} & 2p^{n}_{i,j,k} - p^{n-1}_{i,j,k} + A_{i,j,k}p^{n}_{i-1,j,k} + B_{i,j,k}p^{n}_{i+1,j,k} \\ & + D_{i,j,k}p^{n}_{i,j-1,k} + E_{i,j,k}p^{n}_{i,j+1,k} + F_{i,j,k}p^{n}_{i,j,k-1} \\ & + H_{i,j,k}p^{n}_{i,j,k+1} - (A_{i,j,k} + B_{i,j,k} + D_{i,j,k} \\ & + E_{i,j,k} + F_{i,j,k} + H_{i,j,k})p^{n}_{i,j,k} \end{aligned} \quad (9)$$

where *A, B, D, E, F,* and *G* are the functions of physical parameters *K* and ρ (Phadke et. al. 2000).

Equation (9) is programmed to calculate the wave propagation in heterogeneous media. This approximation is second order accurate in both space and time. Grid dispersion is minimized by keeping the grid spacing smaller than one tenth of the shortest wavelength. The finite difference approximation (2) is stable if

$$\Delta t \le \frac{\min(\Delta x, \Delta y, \Delta z)}{\sqrt{2}\, V_{max}} \quad (10)$$

where $(V = K/\rho)$ and V_{max} is the maximum wave velocity in the medium.

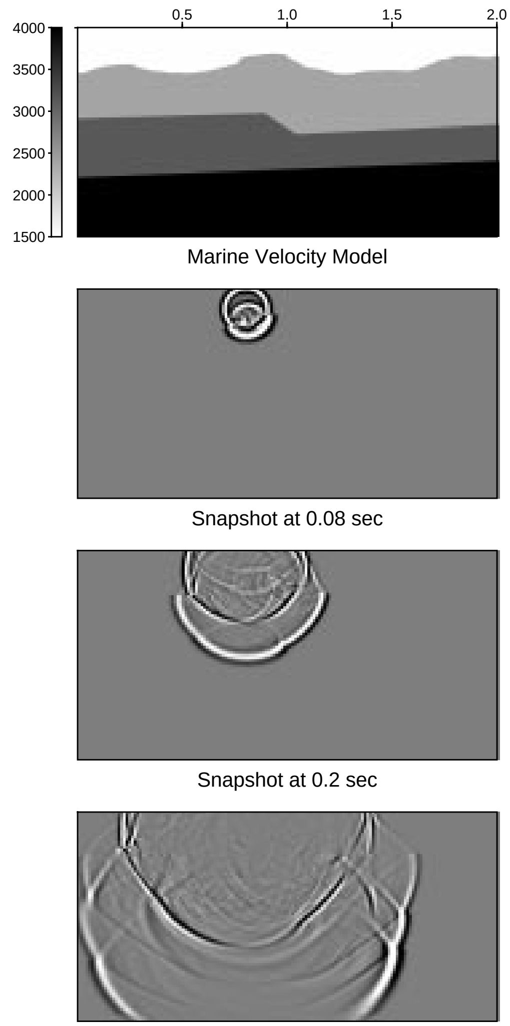

Figure 4: Snapshots of the wave propagation through the marine velocity model. Free surface boundary condition is applied on the top edge of the model and absorbing boundary conditions are applied to left, right and bottom edges of the model.

Since a digital computer has finite memory capabilities, we have to restrict the model size to a fixed number of grid points. This introduces artificial boundaries at the edges of the model. In reality the earth is infinite and therefore all the energy impinging on these boundaries must be absorbed. For the finite difference scheme presented here a sponge boundary condition as described by Sochacki et.al. 1987, is used for attenuating the energy impinging on the left, right, bottom, front and back edges of the model. To implement sponge boundary condition extra grid points are added to gradually attenuate the energy. The free-surface condition is applied to the top boundary.

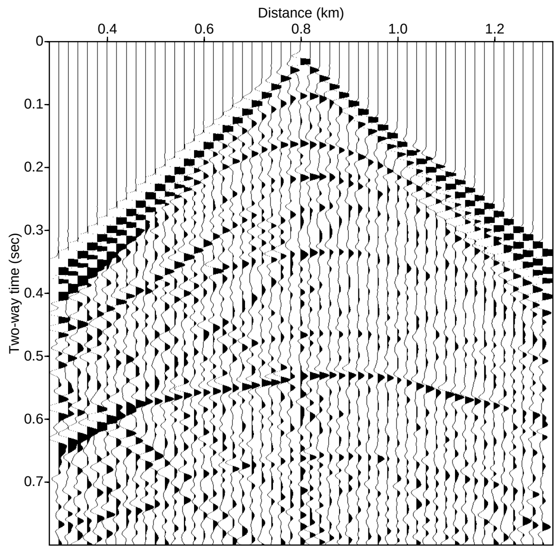

Figure 5: Synthetic seismogram for the marine model. A uniform gain function is applied for plotting purposes.

In the second order central difference scheme implemented here, one can observe that the calculation of the wavefield at a grid point at an advanced time level involves the knowledge of the wavefield at five grid points of the current time level and one grid point of the previous level. Therefore, it is a seven point differencing star. Therefore, if we use a domain decomposition scheme for solving this problem only first order neighbors will be involved in communication for central difference scheme.

The parallel implementation of the algorithm is based on domain decomposition. Domain decomposition involves assigning subdomains of the computational domain to different processors and solving the equations for each subdomain concurrently. The problem domain is a cuboid as shown in the figure 6,

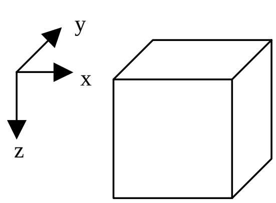

Figure 6: Problem domain (grid point size: I X J X K).

This domain can be partitioned in three ways viz., stripe, hybrid stripe and checkerboard (Phadke et. al., 2000). The checkerboard partitioning involves the least communications and therefore is the most efficient.

Again a message passing paradigm, MPI, is used for implementation of the 3D modelling algorithm. The present implementation is analogous to a Master-Worker system, where master works as the manager and assigns tasks to his workers. The main job of master is to provide the required data to all the workers and distribute workload properly, so that the idle time of the workers is minimized. Also, at the end, the master collects the completed work from all the workers, compiles it and writes it on the disk in a proper manner.

Finite-difference computation of the snapshots can help in our understanding of wave propagation in the medium. We have used a constant velocity model as a numerical example for generating snapshots of 3D acoustic wave propagation. Source is placed at the center of the cubic model. For simplicity sake there is no density variation within the model. However, the algorithm can handle density variations. The source wavelet used for calculation of snapshots is the second derivative of a Gaussian function with a dominant frequency of 30Hz. Figure 7 shows the snapshots of the 3D acoustic wave propagation through the constant velocity model.

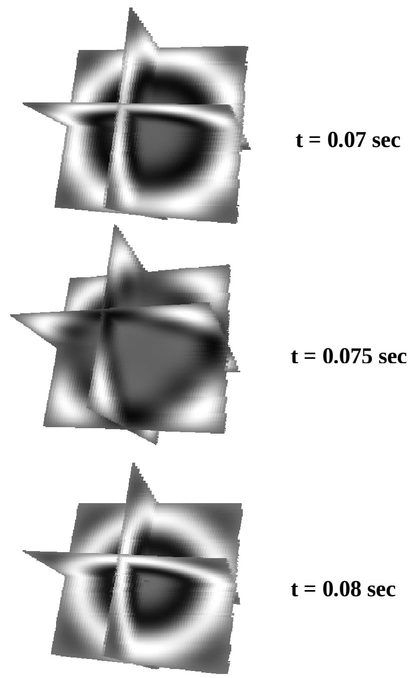

Figure 7: Snapshots of the 3D acoustic wave propagation through the constant velocity model.

We performed the benchmark tests of the parallel algorithm for problem size of 400 X 400 X 400 and a smaller problem size of 200 X 200 X 200. The grid spacing in all three directions was 2m. A time step of 0.0001sec was used and the wave propagation was carried out for 0.1sec. Since the model size 400 X 400 X 400 is too large to fit into the processor memory of PARAM 10000, the test was performed using minimum of 8 processors for the bigger model.

We have used three types of partitioning for the domain decomposition and have experimented with all the three types. For implementation point of view all three types of partitioning play an important role on the basis of memory access pattern. Theoretically, checkerboard partitioning has the best memory access pattern as the partitioned data can reside in the first level of the cache available. In the case of stripe and hybrid stripe partitioning, the access of data from memory may require swapping between first and second levels of cache, which is an expensive operation. Hybrid stripe partitioning has better access patter as compared to stripe partitioning. Bar charts for execution time verses number of processors for 3D acoustic wave modeling, shown in Figure 8a for two different problem sizes, support this statement.

A speedup analysis for the two model sizes (Figure 8b) shows a sub-linear speedup as we increase the number of processors. For a fixed model size the compute to communication ratio decreases with the increase in the number of processors. Therefore if we increase the size of the problem, better speedup can be achieved for large number of processes.

Conclusions

In this paper we have presented several migration and modelling algorithms for seismic imaging on a parallel distributed computer. PSPI algorithm and *ω-x-y* algorithm are both parallelized in frequency domain. RTM algorithm is parallelized by domain decomposition. Highly efficient and scalable codes were developed for these algorithms and implemented on PARAM 1000. The algorithms were tested for both synthetic and real data sets. Modelling algorithms for wave propagation in heterogeneous media were developed and parallelized using a domain decomposition scheme. Efficient codes for both acoustic and elastic wave propagation were developed. These codes form an integral part of the seismic inversion algorithms for estimating the physical properties of the subsurface.

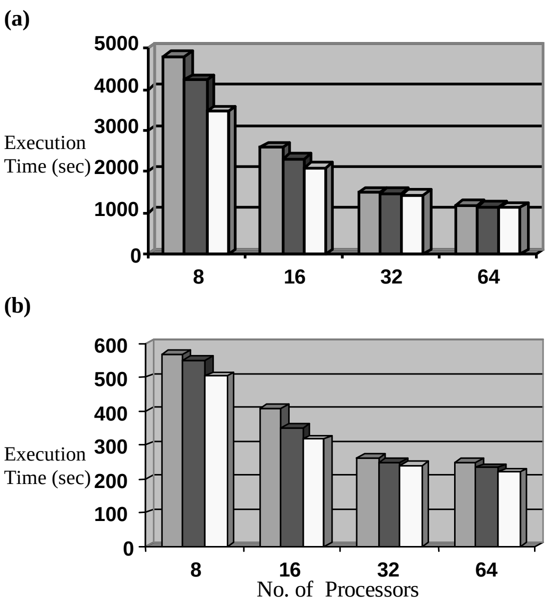

Figure 8: Comparison of execution time for Stripe, Hybrid-Stripe and Checkerboard partitioning for 3-D acoustic wave modeling for two model sizes viz., (a) 400 X 400 X 400, (b) 200 X 200 X 200.

Acknowledgements

Authors wish to thank the Executive Director, C-DAC for providing computational facility on PARAM 10000 and permission to publish this work. We are also thankful to Department of Science and Technology, Government of India, for funding a part of this study under DCS (Deep Continental Studies) program. Discussions with the scientists of GEOPIC, ONGC, Dehradun, were also helpful in improving the quality of the codes.

References

Bhardwaj, D., S. Yerneni, and S. Phadke, 2000, Efficient parallel I/O for seismic imaging in a distributed computing environment, In proceedings of 3rd Conference and Exposition on Petroleum Geophysics (SPG'2000), P. 105-108.

Bhardwaj, D., S. Phadke and Sudhakar Yerneni, 2000, On improving performance of migration algorithms using MPI and MPI-IO, Expanded Abstracts, Society of Exploration Geophysicists.

Claerbout, J. F., 1985, Imaging the Earth's interior, Blackwell Scientific Publications

Dai, N., Vafidis, A., and Kanasewich, E. R., 1996, Seismic migration and absorbing boundaries with a one way wave system for heterogeneous media, Geophysical Prospecting, 44, 719-739.

Gazdag, J, 1978, Wave equation migration with the phase-shift method, Geophysics, 43, P. 1342-1351.

Gazdag, J, and Sguazzero, P, 1984, Migration of seismic data by phase shift plus interpolation, Geophysics, 49, P. 124-131.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values, Geophysical Prospecting, 31, P. 413-420.

Mitchell, A. R. and Griffiths, D. F., 1981, The finite difference method in partial differential equations: John Wiley & Sons Inc.

Oldfield, R. A., D. E. Womble and C. C. Ober, Efficient Parallel I/O in Seismic Imaging, The Int. J. of High Performance Computing Applications, Vol. 12, No. 3, Fall 1998, pp 333-344.

Phadke, S., Bhardwaj D. & Yerneni, S., 1998, Wave equation based migration and modelling algorithms on parallel computers, Proc. of SPG (Society of Petroleum Geophysicists) second conference, pp. 55 - 59.

Phadke, S., D. Bhardwaj and S. Yerneni, 2000, 3D seismic modeling in a message passing environment, In proceedings of 3rd Conference and Exposition on Petroleum Geophysics (SPG'2000), P. 168-172.

Phadke, S., and D. Bharadwaj and Sudhakar Yerneni, 2000, Marine synthetic seismograms using elastic wave equation, Expanded Abstracts, Society of Exploration Geophysicists.

Poole, J., Preliminary survey of I/O intensive applications, Technical Report CCSF-38, Scalable I/O initiative, Caltech Concurrent supercomputing facilities, California Institute of Technology, Pasadena, 1995.

SEG/EAGE 3-d modeling series No. 1, 1997, 3-D salt and overthrust models, SEG publications.

Sochacki, J., Kubichek, R., George, J., Fletcher, W. R. and Smithson, S., 1987, Absorbing boundary conditions and surface waves: Geophysics, 52, 60-71.

Vafidis, A., 1988, Supercomputer finite difference methods for seismic wave propagation: Ph.D. Thesis, University of Alberta, Edmonton, Canada.

Virieux, J.,1986, P-SV wave propagation in heterogeneous media: velocity stress finite difference method: Geophysics, 51, 889-90.