Mecánica Computacional Vol. XXI, pp. 2967–2976
S. R. Idelsohn, V. E. Sonzogni and A. Cardona (Eds.)
Santa Fe–Paraná, Argentina, October 2002

RESERVOIR SIMULATIONS IN CLUSTERS OF PCS

Adriano A. M. Soares[*], Ézio da R. Araújo[†]

[*] PRH 26 – ANP, Pós-Graduação em Engenharia Civil, CTG
Universidade Federal de Pernambuco
Cidade Universitária, Recife, Pernambuco, Brasil
e-mail: amucarbel11@hotmail.com

[†] Departamento de Engenharia Civil, CTG
Universidade Federal de Pernambuco
Cidade Universitária, Recife, Pernambuco, Brasil
e-mail: ezio@npd.ufpe.br

Key words: reservoir simulation, parallel computing, clusters of PCs, large-scale problems.

Abstract. *In recent years, clusters of PCs has becoming an economical and attractive alternative to the use of traditional high performance computers, since these computers can provide a price/performance ratio up to ten times fewer then current proprietary parallel computers, making possible to small and medium officers, centers and scientific institutions, looking up to this technology for the solution their problems. This work describes the application of clusters of PCs for reservoir simulations, showing some results obtained with a parallel simulator developed for distributed memory computers. In the simulator, a two-phase Black-Oil model was implemented, and the partial differential equations governing the problem are discretized using finite differences in a 3D block centered grid, using a Cartesian coordinate system. IMPES and fully implicit procedures can be used as numerical solution, and solvers from software PETSc, witch employ MPI for communication between processors, handles the resulting systems of nonlinear equations, by means of the Inexact-Newton method using several parallel Krylov subspaces methods and preconditioners. The performance of the simulator running in a cluster of PCs is analyzed for solution of some reservoir problems. Speedup and efficiency measurements show de applicability of these computers in the solution of large-scale problems in reservoir simulation and the parallel programming model used in the parallel simulator.*

1 INTRODUCTION

In recent years, reservoir simulation has become an important tool to model multiphase fluid flow in petroleum reservoirs, been essential to help the decision made about better production and management strategies to be taken before and during exploration of a petroleum field.

Many simulation problems use very refined grids to represent with detail the physical model heterogeneities and complexities resulting of high resolution geological models. In the same sense, the mathematical model that describe these problems results in the successive solution of highly non-linear equation systems with the order of tens million equations, exceeding the computational capacity of conventional sequential machines.

Traditionally, there are basically two manners to handle these large scale problems in reservoir simulation. The first one consists in using upscaling techniques, where coarser grids are used companied with quantify and manage of uncertainties risks. Another alternative consists in using high performance computing – vector and parallel computers – to handle more refined grids representing the high resolution models. These two trends are important in different stages of a reservoir simulation project, depending on the study objectives and the available data to describe the reservoir model.

Last years, the industry has shown an increasing need for reservoir simulations with more detailed and complex reservoir models, mainly due the recent advances in reservoir characterization methods. Ever more, with the advances in network and micro-processing technologies, clusters of PCs has become an efficient and attractive alternative to parallel computing, appearing as a good tool to officers and research centers around the world to the solution of large scale problems in several sciences and engineering areas, including reservoir simulation[1, 2].

In order to use high resolution models in reservoir simulations and considering the advent of PC clusters technologies for parallel processing, this work discusses the advantages that are companied with these technologies, showing the main difficulties involved in handling the solution of large scale problems and how the use of PC clusters has helped in overcoming these problems. Based on these ideas some details about the development of a parallel Black-Oil simulator[3] are shown, presenting performance results with the simulator running in a PC cluster for problems with up to half million equations.

2 SOLUTION OF LARGE SCALE PROBLEMS

The main difficulties involved in the solution of large scale problems in reservoir simulations, which normally can result in grids with millions gridblocks, are basically the following[4]:

a) More time consuming - longer simulations run;
b) More RAM memory and disk space requirements;
c) The need for expensive supercomputers and powerful workstations;

These factors have motivated most of petroleum engineers and scientists to apply alternative methods, such as upscaling techniques, to handle these large scale problems with

smaller computational resources. The question is that, in some situations, these methods present inadequate results, mainly when dealing with production rates in reservoirs with large number of wells.

However, lately, with advances observed in mathematical algorithms and computing technologies[5], these difficulties have been bypassed making the solution of these problems more accessible to the industry and scientific community. The following sections discuss each one of these questions, emphasizing the use of clusters of PCs.

2.1 Time Processing

Normally, increasing the problem size in reservoir simulations leads to higher time processing, and this behavior can be directly observed when using sequential machines to perform the simulations, as shown in Figure 1. Since computational time is an important and decisive factor for commercial purposes to the industry, increasing this time can result in serious problems.

Using parallel computers, one can add more processors as the problem size increases, distributing the computational work among processors and reducing the time processing spent in the simulations. But, depending of some factors such as hardware, software and parallel efficiency of mathematical algorithms, the behavior of the variation in computational time can be nonscalable, as shown in Figure 1, resulting in increasing time processing as either the problem size and the number of processors increase, which is not desirable.

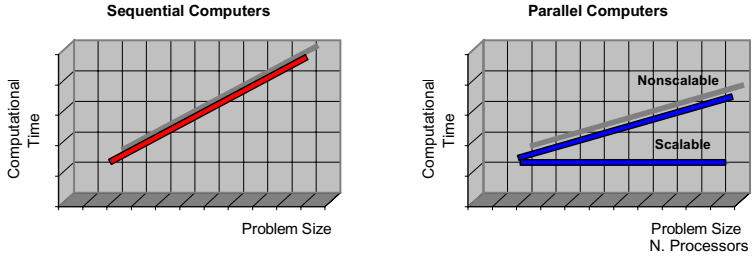

Figure 1: Behavior of computational time in sequential and parallel machines.

So, using parallel computers with efficient and scalable mathematical algorithms, one can eliminate the difficulty related to time consuming when handling high resolution problems.

2.2 Expensive Supercomputers

Another problem that arises when handling large scale problems is the need to use high performance computers, such as traditional proprietary supercomputers and powerful workstations, which normally are very expensive and present difficult maintenance, being accessible only to minority big officers and research centers.

Cluster of PCs are a low cost and efficient alternative to parallel computing, presenting

cost/performance ratios up to ten times smaller than traditional supercomputers. Figure 2 shows the difference in cost presented by these machines for clusters with up to 64 nodes, considering a cost of $650,00 p/ node.

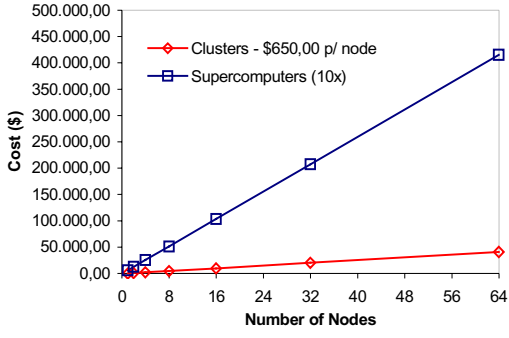

Figure 2: Comparisons between the cost of PC clusters and supercomputers.

Therefore, instead of traditional supercomputers, clusters of PCs are a more accessible alternative to the solution of large scale problems either to medium and small officers and research institutions.

2.3 Memory Resources

In order to increase the amount of data to describe a petroleum reservoir model, one need more computational memory to storage these data. This is an important factor that has limited the use of high resolution models in reservoir simulations.

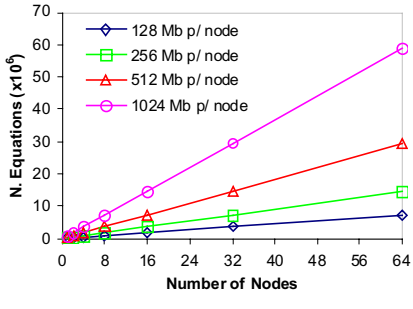

Figure 3: Order of the problem size of reservoir simulation problems that can be solved in clusters of PCs with up to 64 nodes, for different RAM memory configurations.

Clusters of PCs have solved the problem of memory requirements in a simpler way than traditional parallel computers, not only due to the easy in adding more RAM memory and disk space in a single PC, but also considering that adding more nodes in the solution, either the number of processors and the resources memory available increase.

To show the capacity of a PC cluster with up to 64 nodes and considering several RAM memory configurations, some projections based in memory resources necessary to some simulations with a Black-Oil simulator were done, as Figure 3 shows. So, using clusters with 64 nodes one can solve problems with up to 60 million equations, which are very big problems, greatly attending the industry need to perform reservoir simulations with high resolution models.

2.4 Some Advantages of Using PC Clusters

Summarizing, some advantages of using PC clusters are listed bellow:

- Smaller cost/performance ratios compared to traditional supercomputers;
- Technology either accessible to medium and small officers and research centers;
- Free distribution software (ex: Linux), which reduce the costs;
- Hardware with a relative easy maintenance;

In fact, the advantages of using clusters of PCs with off-the-shelf components are very clear, being a technology accessible for most of scientific computing community.

3 PARALLEL BLACK-OIL SIMULATOR

Based on the need to perform simulations using high resolution models and considering the advantages and accessibility of clusters of PCs, a parallel reservoir simulator for distributed memory computers was developed. Some characteristics of the simulator are described in the following sections[3].

3.1 Mathematical Formulation

A two-phase Black-Oil model was implemented, which governing equations are obtained combining mass conservation law, Darcy equation and equations of state[6,7]. Considering an immiscible two-phase oil(o)/water(w) system, the following non-linear system of partial differential equations arises:

$$\begin{aligned} \nabla \cdot \left(\mathbf{K} \frac{k_{ro}}{\mu_o B_o} (\nabla p_o - g\rho_o \nabla Z) \right) &= \frac{\partial}{\partial t} \left(\frac{\phi S_o}{B_o} \right) - Q_{osc} \\ \nabla \cdot \left(\mathbf{K} \frac{k_{rw}}{\mu_w B_w} (\nabla p_w - g\rho_w \nabla Z) \right) &= \frac{\partial}{\partial t} \left(\frac{\phi S_w}{B_w} \right) - Q_{wsc} \end{aligned} \tag{1}$$

where $\mathbf{K}$ is the absolute permeability tensor; $\phi(p)$ is the rock porosity; $k_r(S_w)$, $\mu_l(p_l)$, $B_l(p_l)$ and $\rho_l(p_l)$ are the relative permeability, viscosity, formation volume factor and density of phase $l=o,w$, respectively; p_l and S_l are the pressure and saturation of phase l; Z is the depth, positive

downward; g is the gravity acceleration; Q_{lsc} is the volumetric well rate at standard conditions of phase l.

To complete the mathematical formulation of the problem some constitutive equations are necessary, which are the equation of saturation restraint,

$$S_o + S_w = 1 \tag{2}$$

and the capillary pressure relationship,

$$P_{cow}(S_w) = p_o - p_w . \tag{3}$$

Combining these equations results in a system with two equations and two primary unknowns, oil pressure, p_o, and water saturations, S_w.

Initial conditions are considered computing the primary variables at the initial simulation time. The boundary conditions are handled using no-flow conditions at the external boundaries of the reservoir, and specifying well rates or bottom hole pressure in the production or injection wells.

3.2 Numerical Formulation

The numerical discretization of the system defined by equations (1) to (3) is handled by the finite difference method, using a regular block-centered tridimensional grid, in the Cartesian coordinate system. Recall that finite difference method is the *in facto* standard numerical method in petroleum industry, but the finite element method could be used as well.

The resulting systems of highly non-linear equations are linearyzed using *fully implicit* or *IMPES* formulations - both implemented in the simulator, resulting in the solution of successive systems of linear equations that can be solved by means of several direct or iterative solution methods[8].

3.3 Parallelization of the Simulator

In order to the efficient application of the simulator in distributed memory parallel computers, such as clusters of PCs, the simulator code was developed using specific parallel programming techniques.

The *domain decomposition method* was used in the parallelization of the computational code. The main idea of the method consists in partition and distribute the data related to the grid between processors in order to each processor performs operations only over the data stored in local memory.

In some instants, to perform the computations each processor will need copies of values from neighborhoods processors. During the simulation, these copies, called ghost values, need to be updated, requiring communication between processors. Minimization of the time spent in these communications is very important to the parallel efficiency of the simulator.

Figure 4 and 5 show the partition and distribution of data between processors for some finite difference problems, using 5-point stencil for 2D grids and 7-point stencil for 3D grids, for different decomposition schemes.

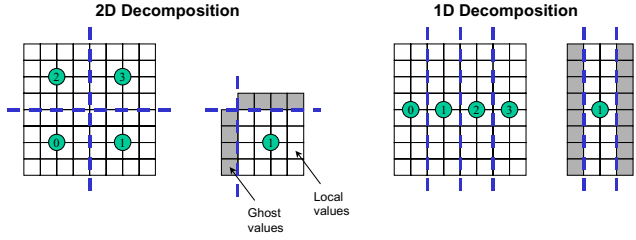

Figure 4: Decomposition schemes in 2D grids partitioned between 4 processors, showing local and ghost values of the processor 1.

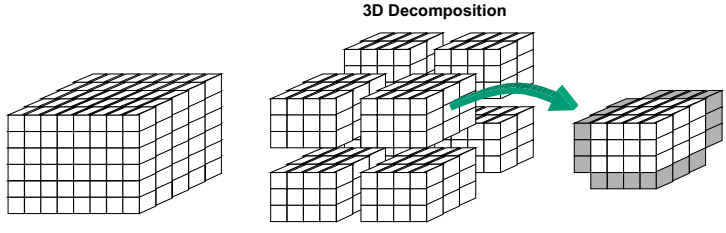

Figure 5: 3D grid partitioned between 8 processors using a tridimensional decomposition scheme.

In the simulator, all communication is handled by MPI – *Message Passing Interface*[9], developed to be efficient and portable among several parallel computers architectures.

Aiming at application of PC clusters, and in order to reduce the time spent in developing new basic numerical subroutines, subroutines and solvers from software PETSc - *Portable, Extensible Toolkit for Scientific Computation*[10] - were used in the simulator, since this software provides high level interfaces to parallelization of numerical codes, and presents several linear, nonlinear and time-dependent solvers, employing different parallel iterative Krylov subspace methods and preconditioners to the solution of the linear equation systems.

4 NUMERICAL APPLICATIONS

Some performance results with the parallel simulator running in a PC cluster with 4 cpus were obtained. The cluster has 4 homogeneous nodes, each one with a single Pentium II - 350 MHz, 128 Mb of RAM memory, a 3Com FastEtherlink XL-PCI 10/100BASE-TX network card, connected to an Intel Express 140T Standalone hub working at 100 Mbps.

The problem analyzed is the water injection in a ¼ -5-spot reservoir, where measurements of time processing were done considering four timesteps in the simulation.

Grids with different number of gridblocks were considered, increasing the size of the problem as the number of processors in the cluster increases, maintaining nearly constant the

number of equations per node, and completely occupying the RAM memory of the nodes in the simulations. The results were obtained using the fully implicit and IMPES formulations, maintaining the gridblocks dimensions constants in all simulations, with 100x100x20 ft.

In order to compare the parallel efficiency of the PETSc solvers, the following iterative methods were used: *Biconjugate Gradient* (bicg), *Biconjugate Gradient Stabilized* (bcgs) and *Generalized Minimum Residual* with restart at each 10 iterations (gmres(10)). In all simulations the Block-Jacobi preconditioner was used, considering one block per node with ILU(0) per block.

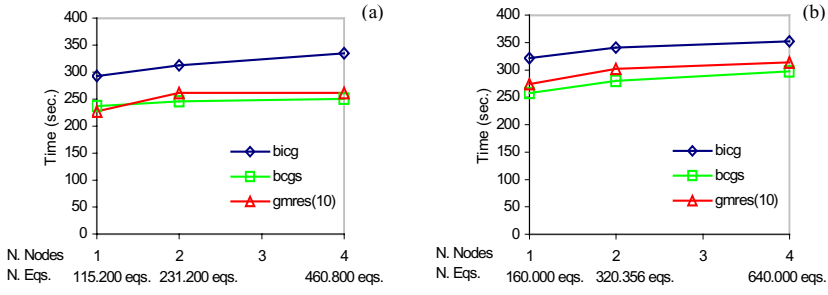

Figure 6: Time processing obtained for simulations with different grids, using (a)fully implicit and (b)IMPES formulations, maintaining nearly constant the number of equations per node (after Soares and Araújo, 2002)[3].

		N. Nodes (N. Equations)						
		1 (115.200)	2 (231.200)			4 (460.800)		
	Method	***T***	***T***	***S***	***E***	***T***	***S***	***E***
Fully Implicit	bicg	293	313	1,87	0,94	335	3,50	0,87
	bcgs	237	246	1,93	0,96	251	3,78	0,94
	gmres(10)	227	252	1,80	0,90	262	3,47	0,87

T = time processing (sec.); *S* = Speedup; *E* = Efficiency.

Table 1: Performance results obtained with fully implicit formulation.

		N. Nodes. / N. Equations						
		1 (160.000)	2 (320.356)			4 (640.000)		
	Method	***T***	***T***	***S***	***E***	***T***	***S***	***E***
IMPES	bicg	322	341	1.89	0.94	353	3.65	0.91
	bcgs	258	280	1.84	0.92	298	3.46	0.87
	gmres(10)	274	302	1.81	0.91	314	3.49	0.87

Table 2: Performance results obtained with IMPES formulation.

Graphs in Figure 6 show the behavior of the variation in time processing obtained in the simulations. The methods analyzed presented nearly the same computational time, where the

bcgs presented the best performance with the fully implicit formulation, and the bicg presented the best performance with IMPES formulation, showing efficiencies of 94% and 91%, respectively.

Table 1 and 2 present additional results of time processing, speedup and efficiency for the two formulations, showing the good scalability of the parallel simulator for all methods analyzed, where efficiencies about 87% to 94% were obtained.

Based on these results, it's important to emphasize that using parallel computers with scalable and efficient mathematical algorithms, the problem size can be considerably increased just with the cost of adding more processors, maintaining the computational time spent in the simulations nearly constant and making possible to solve problems with more detailed and refined grids in place of smaller and coarser grids.

5 CONCLUSIONS

The need to use detailed models to increase the comprehension about the physical phenomena associated with the fluid flow in petroleum reservoirs, has motivated the industry to employ new technologies to reduce the cost and risks in the production process in petroleum fields.

This work showed how recent computing technologies have made the use of clusters of PCs accessible and attractive to the scientific community perform reservoir simulations with high resolution models, allowing the solution of large scale problems with the order of tens million equations.

The ideas and concepts discussed showed that the difficulties associated with the solution of these large scale problems can be bypassed using recent technologies and techniques, demystifying the use of high performance computing.

The parallel Black-Oil simulator developed presented good scalability running in a cluster of PCs, showing that the parallel programming model employed can be used as a reference framework for future developments and research works.

This work has been motivated by the design of a PC cluster with 128 cpus, being constructed in the Center of Technology and Geosciences, in Federal University of Pernambuco. The cluster has 64 dual nodes, with 65.0 GB of total RAM memory and 2.5 TB of total disk memory, using a 100 Mbps FastEthernet network for communication. This cluster will be used in the development of a parallel reservoir simulator framework for the efficient solution of large scale problems, with the order of millions gridblocks.

6 REFERENCES

[1] Wang, P., Balay, S., Sepehrnoori, K., Wheeler, J., Abate, J., Smith, B., Pope, G. A., *A Fully Implicit Parallel EOS Compositional Simulator for Large Scale Reservoir Simulation*, SPE 51885, Symposium on Reservoir Simulation, Houston-TX, (1999).

[2] Shiralkar, G. S., Stefhenson, R. E., Joubert, W., Lubeck, O., Waanders, B. van B., *Falcon: A Production Quality Distributed Memory Reservoir Simulator*, SPE 37975,

Symposium on Reservoir Simulation, Dallas-TX, (1997).
[3] Soares, A. A. M., Araújo, E. R., *Simulação de Reservatórios em Arquiteturas Paralelas com Memória Distribuída*, Revista Técnica de Energia, Petróleo e Gás; Nº 1, Ano 1, (2002);
[4] Gorell, S., Bassett, R., *Trends in Reservoir Simulation: Big Models, Scalable Models? Will you Please Make up Your Mind?*, SPE 71596, 2001 SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, (2001).
[5] Rivière, B., Jenkins, L., *In Pursuit of Better Models and Simulations, Oil Industry Looks to the Math Sciences*, SIAM News, Vol. 35, 1, (2002).
[6] Ertekin, T., Abou-Kassem, J. H., King, G. R., *Basic Applied Reservoir Simulation*, SPE Textbook Series, Vol. 7, (1991).
[7] Mattax, C. C., Kyte, R. L., *Reservoir Simulation*, SPE Monograph Series, Richardson-TX, (1990).
[8] Saad, Y., *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, (1996).
[9] Gropp, W., Lusk, E., Skjellum, A., *Using MPI: A Portable Parallel Programming with the Message Passing Interface*, MIT Press, 2nd Edition, (1999).
[10] Balay, S., Gropp, W. D., McInnes, L. C., Smith, B. F., PETSc Web page; http://www.mcs.anl.gov/petsc.