Parallelization of Thermal Recovery Simulation based on PETSc

Lin Liu[1] Wei Xue[2] Jiwu Shu[2] Li Ma[1]
1 College of Computer and Information Technology, Liaoning Normal University.
Liaoning, China, 116029,
2 Department of Computer Science and Technology, Tsinghua University.
Beijing, China, 100084.

Abstract: Numerical reservoir simulation is one of the effective tools for complex fluid analysis in the oil industry. This paper analyzed the principle of thermal recovery simulation, and designed and implemented a parallel program of the thermal recovery simulation based on PESTc toolkit. The improved domain decomposition with load-balance control was presented to partition the computation into different processing units efficiently. Compared with the serial program, our implementation can gain 8.16 times speedup using 16 cores over a large-scale reservoir test case.

Keyword: parallel, numerical reservoir simulation, load-balance, PETSc

1 Introduction

Numerical reservoir simulation is one of the effective tools for complex fluid analysis in the oil industry. Currently, China's old oilfields have come into the high water-cut stage, requiring fine reservoir simulation to analyze the distribution of the remaining oils. Meanwhile China is rich in heavy oil resources, and is proved to have accumulated 13. 3×10^8t heavy oil resources at the end of 1994, accounting about 15% of the proven oil reserves of the nation [1]. The most effective method of heavy oil recovery is steam injection thermal recovery technology [2]. Therefore, the development of numerical simulation for the heavy oil reservoirs by thermal recovery is becoming very critical. In the real reservoir simulation, the number of simultaneous equations and grids is usually hundreds of thousands or even millions. The high computational complexity and long computational time (one reservoir simulation can be 1.15 days) hinder the efficiency of the oil analysis severely. So, the study on the parallelization of thermal recovery numerical simulation is extremely important.

At present, the research on parallelization of numerical reservoir simulation mainly focuses on the designing of the parallel algorithms and the parallel software. Jianwen Cao [6] [7] [8] has studied and compared the various solutions of parallel linear equations in numerical reservoir simulation. Jiwu Shu[2] improves the domain decomposition method to get load-balance task partitioning for a multiple layers reservoir simulation. Feng Pan[5] has achieved the parallel reservoir simulation for one million grids cases. Yuanle Ma[4] has implemented the parallel software which based on the distributed-memory system. However, because of the tectonic difficulty of boundary condition, the algorithm convergence is heavily affected by the boundary numerical characteristics, making it difficult to adapt to the needs of high-permeability reservoir simulation. Therefore, this paper uses the PETSc toolkit to design and implement the parallel software of thermal recovery simulation which is based on serial software. Testing shows that the parallel software can get correct results according to the serial simulator. The implementation can gain 8.16 times speedup using 16 cores over a large-scale reservoir test case.

978-0-7695-4190-7/10 $26.00 © 2010 IEEE
DOI 10.1109/ISPA.2010.18

2 Description of thermal recovery simulation

Thermal recovery model describes the flow of oil, gas and water phases under the action of viscosity, the gravity and capillary pressure, the impact of temperature on various physical properties of liquid and solid media, and the balance of each group in the vapor-liquid two-phase. At the same time, this model also considers the heat conduction in the reservoir, convection, and heat loss in the upper and lower cladding, simulate fracturing phenomenon that might arise when injection of steam and the compaction phenomenon which might arise in the exploitation. The model's basic equations are as follows :

1) mass conservation equation:

$$\sum_{j=1}^{N_p}\iiint_V \frac{\partial}{\partial T}(\varphi\rho_j S_j \chi_{ij})dv + \sum_{j=1}^{N_p}\oiint_S (\rho_j S_j \chi_{ij} \upsilon_j)ds$$

$$+q_i = 0, \quad i = 1,2,...,N_c$$

φ is the rock porosity, ρ_j is the density of j-phase meterial, S_j is the saturation of j-phase meterial, υ_j is the apparent velocity of j-phase meterial , χ_{ij} is the molar of component I in j-phase, N_p is the number of phases.

2) energy conservation equation:

$$\iiint_V \frac{\partial}{\partial T}\left[\varphi\sum_{j=1}^{N_p}\rho_j S_j u_j + (1-\varphi)\rho_{rock}C_p(T-T_i)\right]dv +$$

$$\oiint_S (q_h + q_c)ds + Q_c + Q_h = 0$$

u_j is the internal energy of j-phase meterial, ρ_{rock} is the rock density, C_p is the rock specific heat, T is the temperature, T_i is the temperature of component i, q_h is the flow rate of enthalpy, q_c is the flow rate of heat, Q_c is the heat source items, Q_h is the enthalpy source term. .

3) Mole fraction of constraint equations:

$$\sum_{i=1}^{N_c} \chi_{ij} = 1.0$$

4) Saturation constraint equation:

$$\sum_{i=1}^{N_p} S_j = 1.0$$

The basic process of research on the above mentioned issue with numerical simulation method includes two major steps: mathematical model discretization and non-linear equations solving. Take the serial software with steam injection oil recovery, which offers a variety of grid ordering methods and grid difference method, as an example. This software uses Newton method to solve non-linear equation and iterative method to solve linear equation after discretization. Steam injection oil recovery software is divided into two parts: the initialization module and the main module(as figure 1). Initialization module is used to read data. It reads data only once, thus is less time-consuming while complex for programming; Main module is the core of the whole software, including equations formation and solving. As showed in the analysis of the serial software, about 30% of the computation time is consumed by forming the coefficient matrix of linear equation, while 50%~60% by solving linear equation; Other computation time occupies less than 20%. Therefore, the parallelization of the thermal recovery simulation equation is essential.

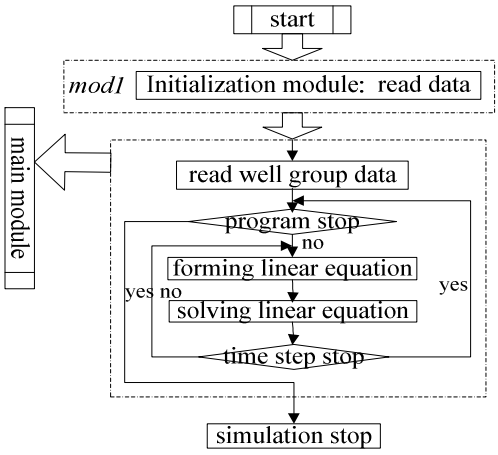

Figure 1 the serial software's flow chart of thermal recovery simulation

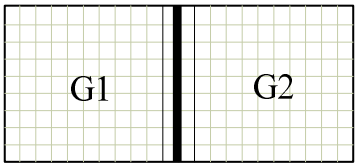

Figure 2 overlapping domain decomposition technology

Take Lotus 2 oilfield of Huanxiling as the example, the ratio of the time consumed in different part is about 2:3:5. The serial computation took nearly 27.5 hours under the main server environment(four-way eight-core SMP). Thus it can be seen that the main work of parallelization should be the solution of linear equation.

3 Parallelization of thermal recovery simulation

3. 1 Design of the parallel simulator

Existing parallel programming adopts parallel methods to solve linear equations in the numerical reservoir simulation[8] [9] [10]. However these methods are subjected to poor efficiency by wasting about 40%~50% computation on parts other than equation solving. Domain decomposition method provides a new solution for reservoir numerical simulation. It could divide the region into several sub-regions(Figure 2). Each processor is responsible for a smaller part of grid's solution, reducing computation load as well as improving efficiency. This paper uses overlapping domain decomposition method to divide the entire region, uses the iterative method which provides by PETSc toolkit to solve the linear equation, and uses the special technology to control the consistency of sub-regions, enhancing the parallel efficiency and ensuring the accuracy of numerical simulation. The parallel software's flow chart of thermal recovery process is as follows:

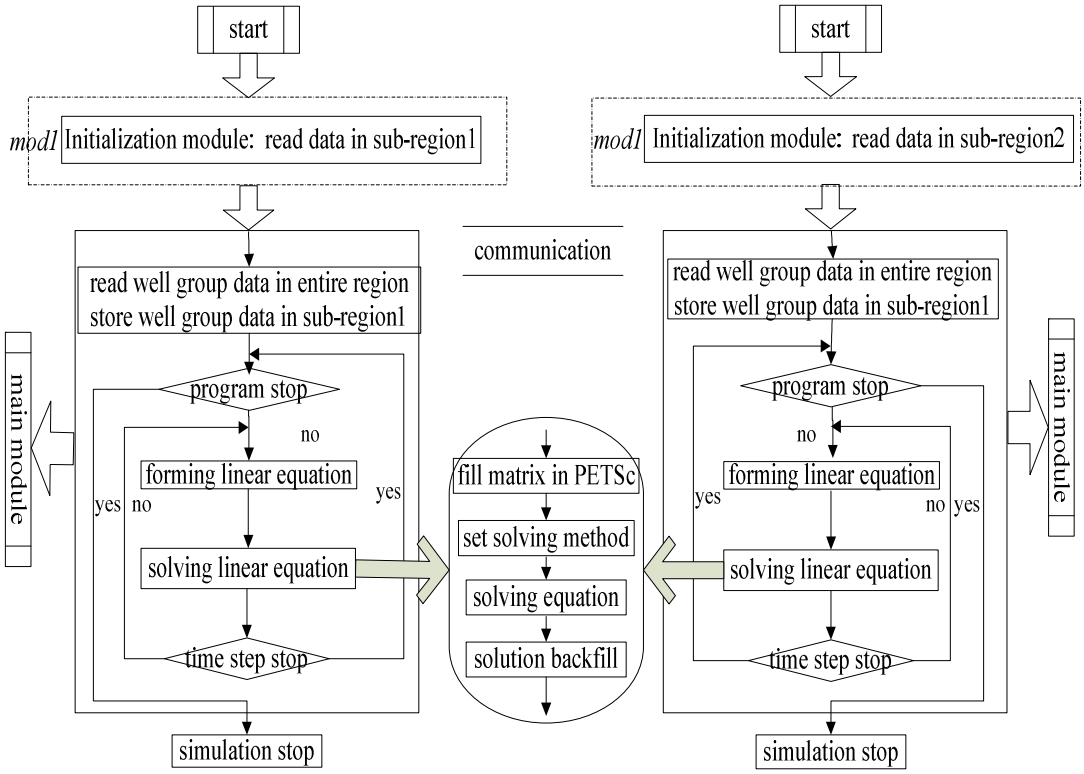

Figure 3 the parallel software's flow chart of thermal recovery simulation

3. 2 Domain decomposition technology

Using domain decomposition method could not ensure energy and mass conservation. Yuanle Ma analyzed the numerical reservoir simulation and the relationship between neighboring grids. Then overlapping domain decomposition method is used to realize the thermal recovery simulation parallelization program. Analysis shows that the overlapping domain decomposition provides reasonable boundary conditions for each sub-region. It could ensure the material and energy conservation.

It is well known that a job of parallel program is not finished until the computation on the slowest processor[3] is over. So how to ensure load-balance is important. In the realistic reservoir simulation, the entire region always contains a lot of death nodes（the nodes that do not require iterative calculation）. If "rolling" method is used, each processor will solve different scale of active grids, causing inefficient. Considering the actual reservoir's characteristics, we divide by the active grid's number of the entire region: First, determine the grid's number under average condition according to the active grids and the processors' number. Then count grid's number from previous divided column. If the gap between them is consistent with the scope allowed by processor, then this column is the dividing line or else continuance count. This method could ensure the load-balance and minimize the differences between each processor's computation time.

The serial software of thermal recovery simulation provides many kinds of grid's ordering: D4 ordering, nature ordering and etc. This paper uses nature ordering(as figure 4) due to domain decomposition in order to reduce the modification of the program and increase calculation efficiency. And it is related to the division order: if divided along the x-axis, then the grid's ordering is along y-z-x; if divided along the y-axis, then the grid's ordering is along x-z-y, facilitating data communication.

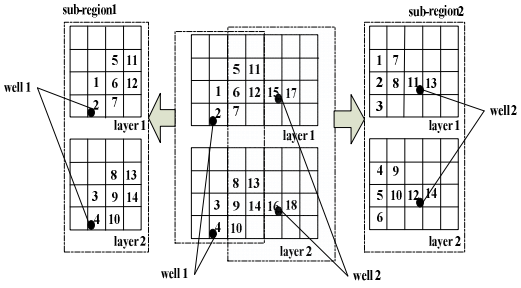

Figure 4 grid's ordering

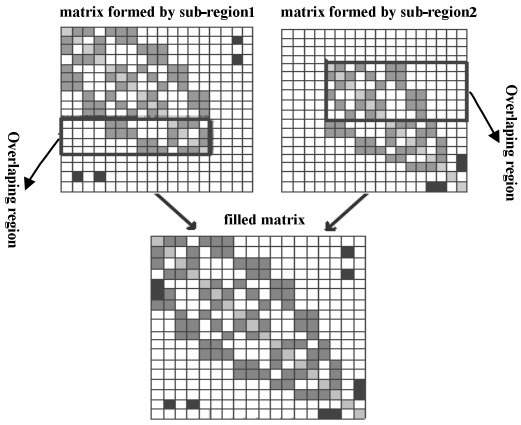

Figure 5 matrix filling in PETSc toolkit

3. 3 Forming linear equation

At present, the popular parallelization method in the numerical reservoir simulation is using parallel algorithm to solve linear equations. However it is difficult to program using this method. PETSc toolkit has made it possible to quickly realize parallel method to solve linear equation. Using PETSc toolkit to solve the linear equation, first is to form the equation, and distributed to each processor. In the realistic reservoir simulation, the dimension of solving linear equation is usually hundreds of thousands or even millions while the coefficient matrix is a highly sparse matrix. In order to reduce the storage space, we use the block compressing sparse matrix(BAIJ) which provided by PETSc toolkit. The non-zero block number of BAIJ is

needed to set before using. We can set it according to the grid's ordering and the difference method which determine the form of equation.

After being formed, the equation is needed to be filled. Using overlapping domain decomposition method, each processor is responsible for a sub-region's solution, forming part of the linear equation and distributing storage on each processor. Overlapping area is used to provide reasonable boundary condition in the formation of linear equation. During the equation being formed, the calculation of sub-regions only needs to correctly fill matrix and right-hand side corresponding to this region's grid; however in the end of calculation, the overlapping region needs communication. It is needed to generate the linear equation, rather than part of the equation. Therefore, the overlapping region is needed to remove when fill the matrix in PETSc. As shown in figure 5, the filled matrix is the combination of two sub-regions by removing the overlapping region.

If the equation is formed, then the pre-condition and Krylov method are used to solve the equation. PESTc provides a large number of pre-conditions and Krylov methods. Different combinations are suitable for solving different issues In this paper compared the various solutions provided by PETSc in order to find the best one for thermal recovery simulation. After tests, the combination of ASM and BCGS has the fastest speed.

3. 4 Communication optimization

Data communication is one of the key factors affecting parallel program's efficiency. Because this paper uses overlapping domain decomposition method to divide the entire region, it is need to ensure the consistency of each sub-region. The data communication needed by thermal recovery simulation is as followed: equation's communication, time step communication, conservation communication and data update step communication. And the equation's communication contains message exchange in the aggregation and the communication with the overlapping region. Each group of data communication contains many parameters. In order to improve performance and reduce communications, related data is grouped into one package.

Thermal simulation software needs to read well group data at the beginning of each period, in order to judge the status of wells. Because each sub-region is responsible for different wells and each well's production dynamics are different, some parameters are different for different wells, resulting in parallel simulation error. If communication to ensure the consistency of the parallel program is used, it will lead to the parameters needing communication are different due to different oil. According to the characteristics of oil, point-to-point communication may be needed, while will increase the communication overhead and the complexity of the parallel program. Therefore, our program modified the procedure of well group data reading: each processor need to read all wells, but only to store the wells which locate in this sub-region. After this refinement, the differences of the parameters which caused by the difference of oils in the parallel program is avoided.

PETSc toolkit's storage strategies of parallel matrix and parallel vector use continuous distributed storage, which means the continuous neighboring matrix line and vector elements are stored in the same processor. According to the above principles, in order to improve the efficiency of parallel program and reduce the data exchange of each processor in the aggregation, we should store the calculating matrix and vector

data as much as possible which formed by local processor.

4 Testing results

Table 1 test environment

Four-way Eight-core SMP	
Cpu	Intel(R) Xeon(R) 7560
Memory Size	64GB
Operating System	RHEL5. 4
Intel® Package	3. 2. 2. 006
PETSc Version	petsc-3. 0. 0-p5
Solving Method	ASM+BCGS

Real reservoir test cases:

Experience case a2.dat: the grid's number is 96 * 120 * 10=115200, well's number is 64, the well's production perform dynamics no difference and the distance of each well is equal;

Lotus 2 oilfield of Huanxiling: the grid's number is 155 * 62 * 6=57660, well's number is 290, the well's production perform dynamics difference, and the distance of each well is not equal. there are a number of major faults, reservoir geological conditions vary widely.

Table 2 combined solution testing

Krylov	pre-condition	
	asm	bjacobi
fgmres	0. 8072	0. 4758
lgmres	0. 8992	0. 7251
gmres	0. 8627	0. 7180
tcqmr	1. 3278	1. 6414
bcgs	0. 9120	0. 9791
bcgsl	0. 9537	0. 9989
ibcgs	0. 9714	1. 0563
tfqmr	0. 8961	0. 9945
bicg	1. 2798	1. 1395
cgs	0. 9263	0. 9803

Table 3 choosing the best solution

pc	krylov	execution time of test case a2.dat
asm	fgmres	279. 4434
bjacobi	fgmres	321. 8175
asm	bcgs	271. 9680
bjacobi	bcgs	273. 6149

Table 2 is the once equation solving test of the parallel solution in PETSc toolkit. In order to find the best solution, we choose the fast Krylov method: fgmres. Because of lgmres and gmres are extances of fgmres, so we choose another Krylov method: bcgs. Then test these four combinations on experience case a2.dat with eight processors. The result shows that the combination of asm and bcgs is the best for solving the linear equation in thermal recovery simulation. We tested our parallel program using ASM+BCGS solving method on the same machine. The results are as follows.

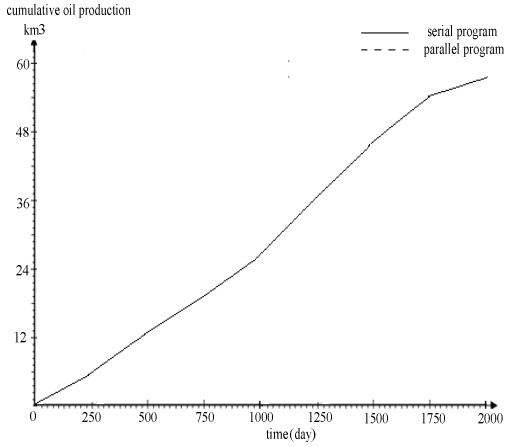

Figure 6 the forth well's history matching in experience case a2.dat

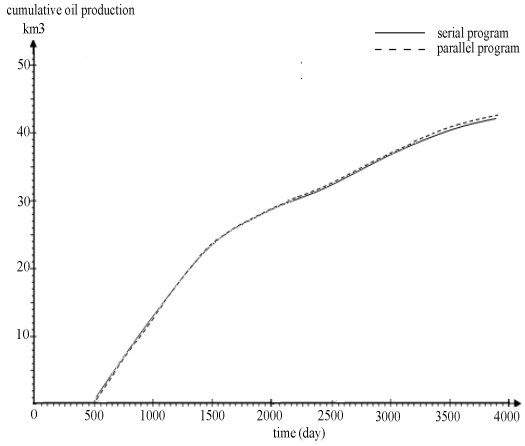

Figure 7 the forth well's history matching in Lotus 2 oilfield of Huanxiling

Figure 6 is the forth well's history matching in experience case a2.dat, Figure 7 is the forth well's history matching in Lotus 2 oilfield of Huanxiling. From the figure we can see that the parallel simulation and the serial simulation is almost no difference, the results are satisfactory and fitting the standard deviation.

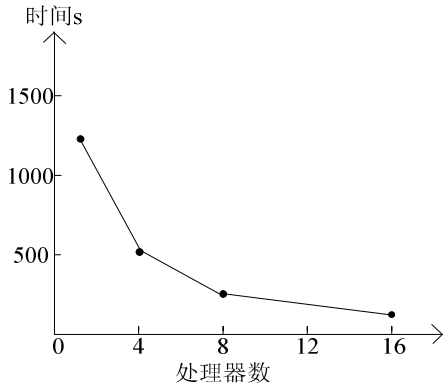

Figure 8 execution time of experience case a2.dat

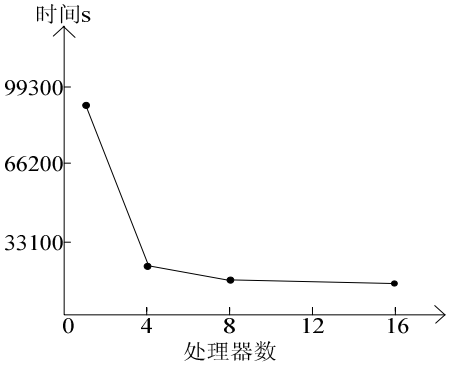

Figure 9 execution time of Lotus 2 oilfield of Huanxiling

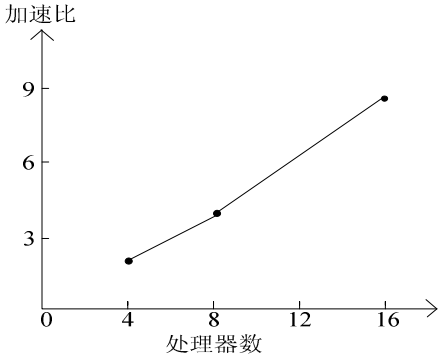

Figure 10 the speedup with experience case a2.dat

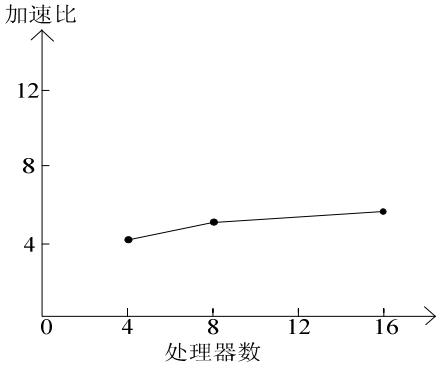

Figure 11 the speedup with Lotus 2 oilfield of Huanxiling

As shown in figure 8, the execution time of experience case a2.dat has a downward trend with the increase of the processor's number. At sixteen processors, the execution time is the least. As shown in figure 9, Lotus 2 oilfield of Huanxiling has the best effect with four processors, but when the number of processors increases, the time has no obvious reduction. As shown in figure 10 and figure 11, in the same grid density, with the increase in the number of parallel processors, the speedup has an upward trend, and the calculation with sixteen processors in experience case a2.dat can be accelerated by 8.16 times, which shows that in a certain range, with the increase of the hardware, the software can play a greater role. This could indicate the effective of parallel computing for thermal recovery heavy oil reservoirs.

Further analysis shows that the numerical calculation in the thermal recovery simulation, the main computational work is to solve linear equation. Parallel program is to transfer a high-dimensional issue into multiple low-dimensional issues to solve on different processors, which is the key to achieve high speedup. So in theory, the more sub-regions we divide, the more efficient it is. Such as shown in figure 10, the parallel efficiency significantly improved with the number of processors increases. The efficiency with the sixteen processors relative to eight processors of

Lotus 2 oilfield of Huanxiling has no obvious increase(figure 11). This is due to excessively small grain size under high parallelism and the switch of the well. However, this case is rare and also unforeseeable, thus may not affect the result of parallel software of thermal recovery simulation too much. It is needed for further research.

5 Conclusions

In this paper, we analyzed the principle of thermal recovery simulation in depth, and then designed and implemented a parallel software of the thermal recovery simulation based on PESTc toolkit. The improved domain decomposition with load-balance control was presented to partition the computation into different processing units efficiently. which leads to the efficient numerical simulation for large-scale reservior. With the experience case a2.dat, our parallel software achieved 8.16 times faster than the serial thermal recovery simulatior on a computer with 16 cores. At the same time, our parallel software can keep high accuracy according to the serial software. We believe that our work will benefit the analysis process of the thermal recovery technologies of reservoir development.

References

[1] Chuan-sheng Sun, Shun-long Peng, The Heavy Oil Reservoirs by Thermal Recovery in the District 9 of Karamay. Petroleum Industry Press. Beijng, China. 1998. 1

[2] Ji-wu Shu, Jin-xi Zhao, Wei-si Zhou, De-fu Zhang, Jian–fei Zhang, Lei Zhang, "Zonal Parallel Algorithm of Load-balance on Nonuniform Domain Simulation of Reservoir partitioning", Journal of Software. Vol. 10, No. 10, Oct. 1999, pp. 187-192

[3] Jian-wen Cao, Feng Pan , Ji-feng Yao , Jia-xu Sun, Guo-zhong Zhao, "The Implementation Of A Parallel Software Of Petroleum Reservoir Simulation And Its Application On Native High Performance Computers". Bijing: Journal Of Computer Research And Development, Beijing, China. Vol. 8, No. 39, 2002, pp. 971-980

[4] Yuan-le Ma, Yue-qiang Liu, Ping Lv. "Reservoir Thermal Recovery Simulation Software Parallelization". Journal of Tsinghua University. 2002. Vol. 42, No. 12, pp. 1612-1615

[5] FengPan JianwenCao, "Parallel Reservoir Integrated Simulation Platform for One Million Grid Blocks Cases". Journal of Communication and Compute, USA. Vol 2, 2005, pp. 29-42

[6] JianwenCao, jiachang sun, "An Efficient and Effective Nonlinear Solver in a Parallel Software for Large Scale Petroleum Reservoir Simulation". International Journal of Numerical Analysis and Modeling, Canada. Vol 2, 2005, pp. 15-27.

[7] JianwenCao, Choi-HongLai, "Numerical Experiments of Some Krylov Subspace Methods for Black 0il Model". An Intemational Journal of Computers and Mathematics with Applications, Elsevier. Vol. 44, 2002, pp. 125-141

[8] Jiachang Ssun, Jianwen Cao, "Numerical Petroleum Reservoir Simulation and Parallel Preconditioning Algorithms Research". Science in China, Beijing, China. Vol 47, 2004, pp. 32-40

[9] Qing-kun Liu, Li-zhong Gtui, Jian-ping Fan, Jin-xi Zhao, De-fu Zhang, "Valid Parallel Algorithm for Nonuniform Domain Simulation of Reservoir". Chinese Journal of Electronics, Beijing, China. Vol. 32, No.2, 2004, pp.342-345

[10] (USA) Turgay Ertekin. Jamal H, etc, Applied Reservoir Simulation. Technology Petroleum Industry Press, Beijing, China. 2004, pp. 1-736

[11] (Brazil)Eduardo Abreu, (USA)Frederico Furtado, (Brazil) Felipe Pereira, "Numerical Simulation of Three-Phase Flow in Heterogen-eous Porous Media". Springer Verlag Berlin Heidelberg. 2007, pp. 504-517