

解非均匀区域油藏模拟问题的一种有效并行计算方法

刘青昆¹, 归丽忠¹, 范建平², 赵金熙¹, 张德富¹

(1. 南京大学计算机软件新技术国家重点实验室, 南京大学计算机科学与技术系, 江苏南京 210093;

2. 中石油勘探开发公司, 北京 100029)

摘 要: 运用区域分解方法求解油藏模拟问题, 结合灵活的区域划分和负载平衡策略, 可以提高解题速度。但一般区域分解方法在应用中存在算法收敛慢, 计算量增加, 虽有一定加速比, 但提高不大。特别是对非均匀区域的应用问题, 子区域收敛速度不同引起的负载不均衡进一步导致性能下降。本文基于分布式并行系统, 运用整体预处理改进区域分解方法, 降低了非均匀油藏模拟问题的计算复杂性, 提高了模拟计算的速度, 并根据全局和本地负载状态调整区域划分, 结合有效的负载平衡策略, 达到对一类黑油模型油藏模拟问题高效的并行求解。实验结果表明, 模拟速度有较大提高, 并取得较高加速比。

关键词: 区域分解; 并行计算; 油藏数值模拟; 负载平衡

中图分类号: TP301.6; TE119

文献标识码: A

文章编号: 0372-2112 (2004) 02-0342-04

Valid Parallel Algorithm for Nonuniform Domain Simulation of Reservoir

LIU Qing-kun¹, GUI Li-zhong¹, FAN Jian-ping², ZHAO Jin-xi¹, ZHANG De-fu¹

(1. State Key Laboratory for Novel Software Technology at Nanjing University, Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China; 2. CNPC, Beijing 100029, China)

Abstract: The domain decomposition method, when combined with flexible decomposition and load-balancing strategies, can improve the parallel performance of iterative computations and save the costs in large-scale numerical reservoir simulations. However, the ordinary domain decomposition method was found to have problems of increased computational load and slowed convergence, especially in working with non-uniform domain numerical simulations. Here, we report the results of an efficient parallel approach that we have recently designed and implemented. This approach used preconditioned domain decomposition method and load-balancing strategy based on the distributed parallel environment to reduce the computational complexity of non-uniform numerical reservoir simulations and enhance the computational speed. Our computational results showed that the simulation speed was improved largely with a high speedup.

Key words: numerical simulation of reservoir; parallel computing; domain decomposition; load-balancing

1 引言

作为石油行业计算机应用主要方面之一的油藏数值模拟, 随着计算机技术的发展, 网络速度的提高, 变得更加复杂。这类问题的计算规模大, 计算时间长, 如何提高模拟速度和精度, 一直为油藏数值模拟软件工作者所关注^[1,2]。区域分解算法是一种偏微分方程数值解技术, 它具有大问题化小问题的特点, 使原问题求解转化为各个子问题的并行求解, 这不仅加快解题速度, 还可以结合区域的灵活划分和负载分配策略更易达到负载的均衡。目前, 油藏数值模拟较多采用该方法进行并行计算^[1,3,4]。

近年来, 国内外已开发出不同粒度的并行油藏模拟器^[1,5]。如: Wheeler 等研制了基于局部网格加密油藏模拟器, Shiralkar 等研制的分布式内存油藏模拟器, 但在并行化过程中

存在程序结构、模型的并行线性求解方法优化等问题。在运用区域分解方法求解时, 由于子区域达到收敛条件的迭代次数变化很大, 往往出现等待个别处理机计算任务的完成, 这样导致处理机资源的浪费^[1,3,6]。如何提高并行求解速度, 克服负载不平衡带来的性能下降成为油藏模拟问题求解的重要研究方向^[7]。

本文基于分布式并行系统, 对一类非均匀区域黑油模型油藏模拟问题进行分区并行求解。文中采用预处理技术加快收敛速度和提高计算精度, 并结合有效的负载平衡策略, 根据全局和本地负载状态调整区域划分, 来达到并行计算的负载均衡。在整个区域模拟中计算速度有较大提高, 解非均匀区域油藏问题取得较佳的并行效率, 目前该方法已应用于实际模拟计算中。

2 黑油模型方程

黑油模型在油藏数值模拟研究中得到广泛的研究和应用. 其压力计算模型如下:

$$\begin{cases} (B_o - R_{so}B_g) [\nabla T_a (\nabla P_o - \nabla r_o Z) - (q_w /_{sc})] \\ + (B_w - R_{sw}B_g) [\nabla T_w (\nabla P_o - \nabla r_w Z - \nabla P_{caw}) - (q_w /_{wsc})] \\ + B_g [\nabla T_g (\nabla P_o - \nabla r_g Z - \nabla P_{cgo}) + \nabla R_{so} T_o (\nabla P_o - \nabla r_o Z) \\ + \nabla R_{sw} T_w (\nabla P_o - \nabla r_w Z - \nabla P_{caw}) - (q_g /_{sc})] \\ = C_t \frac{\partial P_a}{\partial t} \quad () \\ P(x, y, z, 0) = P^{(0)} \\ L_1 P_o = f \quad (\partial) \end{cases} \quad (1)$$

$$\begin{aligned} \text{其中: } C_o &= -\frac{1}{B_o} \frac{\partial B_o}{\partial P_o} + \frac{B_g}{B_o} \frac{\partial R_{so}}{\partial P_o} \\ C_w &= -\frac{1}{B_w} \frac{\partial B_w}{\partial P_o} + \frac{B_g}{B_w} \frac{\partial R_{sw}}{\partial P_o} \\ C_g &= -\frac{1}{B_g} \frac{\partial B_g}{\partial P_o}, \quad C_r = -\frac{1}{B_g} \frac{\partial}{\partial P_o} \\ C_t &= C_r + C_{os} + C_{ws} + C_{gs} \end{aligned}$$

这里:下标 o, w, g 分别表示油、水、气相, B_l 表示 l 相的体积系数, R_{so} 为溶解气油比, l 表示 l 相的注入或采出的体积流量, r_w 为井底半径, Z 为油层深度, g 为重力加速度, P_o 为油层压力, P_{cgo}, P_{caw} 分别为油气毛管力和油水毛管力, ∇ 为油层孔隙度, l 为 l 相的密度, 下标 sc 表示在标准条件下, t 为时间, L_1 为边界条件算子, f 是关于时间的压力函数.

式(1)是非线性偏微分方程组,采用隐式压力显式饱和度方法(IMPES)求解,经过差分离散化以及线性化后,一般可化成大型非对称稀疏七对角线性方程组,表达式如下:

$$\begin{aligned} g_{ijk} P_{i,j,k-1} + b_{ijk} P_{i,j-1,k} + c_{ijk} P_{i,j,k} + d_{ijk} P_{i,j,k+1} \\ + e_{ijk} P_{i,j+1,k} + h_{ijk} P_{i,j,k+1} = f_{ijk} \end{aligned} \quad (2)$$

其中 $g_{ijk}, b_{ijk}, d_{ijk}, c_{ijk}, d_{ijk}, h_{ijk}$ 为七对角线性方程组的系数, f_{ijk} 为方程右端项,下标 i, j, k 表示网格所在位置. 设 I 为 X 方向的网格点数, J 为 Y 方向的网格点数, K 为 Z 方向的网格点数,对应黑油模型油藏问题总网格点数为 $n = I \times J \times K$, 式(2)可写成矩阵形式 $AP = f$.

该方程组具有大型稀疏、病态等特点.

3 并行方案及负载均衡策略

对于黑油模型油藏模拟问题求解采用 IMPES 方法,具有所占内存小,计算工作量少,方法简便等特点.从总体上说压力与饱和度求解是交替进行,每个阶段结束需要收集计算结果以供下阶段使用.求解黑油模型,计算时间主要花在解压力方程组上,其串行计算流程如图1所示.

运用区域分解方法对压力方程进行并行化,将大的区域分成若干小的区域,每个处理器同时并行处理一个或多个区域,各个子区域之间只在区域迭代完成时交换一次信息,通信量不大.我们分别采用子结构和交替 Schwarz 方法求解压力方程,发现子结构算法计算量大,交替 Schwarz 算法收敛较慢,这

些使得两种算法在并行时,虽有一定加速比,但提高不大.为此,对于前者利用子矩阵 A_{ii} 的不完全 LU 分解得到 $M_{ii}, i = 1, \dots, m$, 由 M_{ii} 代替容量矩阵 $S = A_{BB} - \sum_{i=1}^m A_{Bi} A_{ii}^{-1} A_{iB}$ 中的 A_{ii} , 由于 M_{ii}^{-1} 比 A_{ii}^{-1} 容易计算, $i = 1, \dots, m$, 所以算法的计算复杂性降低,而原直接方法则转化为迭代方法;对于后者预处理整个区域上的离散问题,预处理

后的系数矩阵的条件比原系数矩阵的要好,较大提高了交替 Schwarz 算法的收敛速度.在实际模型计算中,当收敛条件较宽时,整体预处理交替 Schwarz 算法比迭代子结构算法得到更高的并行加速比,反之,整体预处理交替 Schwarz 算法收敛,而迭代子结构算法不收敛.随着问题规模的增大,前者加速比增大,后者加速比下降.下面给出求解实际黑油模型的整体预处理交替 Schwarz 算法离散化形式.

从总体上说,其离散模型可化为系数矩阵为三对角块的线性方程组.形式如下:

$$Ap = \begin{bmatrix} A_1 & B_1 & & & \\ B_1^T & A_2 & \ddots & & 0 \\ & \ddots & \ddots & \ddots & \\ & 0 & \ddots & \ddots & B_{m-1} \\ & & & B_{m-1}^T & A_m \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \quad (3)$$

式中 A_1, A_2, \dots, A_m 分别由子区域 $1, 2, \dots, m$ 形成的矩阵, B_1, B_2, \dots, B_{m-1} 分别由子区域 $1, 2, \dots, m$ 的相互重叠区域形成的矩阵.选择 MLU 预处理矩阵 M , 改善 A 的条件数,提高外循环收敛速度.由此可给出整体预处理交替 Schwarz 算法计算步骤如下:

步骤1 计算 A 预处理阵 M .

步骤2 选择初始近似 P_{12}^0 ;

步骤3 并行计算子区域上的边值问题

(1) 解 1 上的边值问题对应的离散问题

$$(M^{-1}A)_1 P_1^1 = (M^{-1}f)_1 - (M^{-1}A)_{12} P_{12}^{0-1} \quad (4)$$

(2) 解 2 上的边值问题对应的离散问题

$$(M^{-1}A)_2 P_2^1 = (M^{-1}f)_2 - (M^{-1}A)_{21} P_{12}^0 \quad (5)$$

由于整体预处理增加了一次内迭代的计算量,但减少了外迭代的次数,若离散算子 A 条件较坏,块 Gauss-Seidel 迭代收敛缓慢,由于算法对 A 预处理,收敛速度提高,这时其优越性很明显.

运用区域分解方法求解,任务的完成是等待最慢结点机上的消息,解题时间取决于最后完成的子任务.在非均匀区域

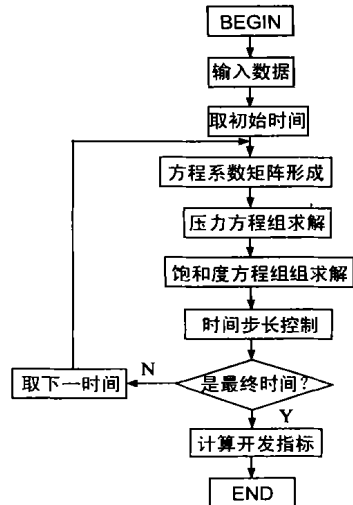


图1 黑油模型压力串行计算流程

油藏数值模拟中,子区域达到收敛条件的迭代次数变化很大,往往出现等待个别处理机计算任务的完成,这样导致处理机资源的极大浪费,为了很好地解决这个问题,根据整个区域的全局状态和本地处理机的局部状态信息动态地改变区域的划分,这里假设划分的区域数不大于处理机数。

设处理机数为 n , 将处理机 M_i 的一次迭代完成时间 T_i 计为它的任务量 ld_i , 采用负载均衡的分区并行算法可表述为:

- (1) 将区域均分为子区域 $1, \dots, n$, 并对应地分配到处理机 M_1, \dots, M_n 上. 令区域负载初始集为: $Set_s = \{1, 2, \dots, n\}$;
- (2) 应用区域分解算法在处理机上并行执行各子区域, 直至达到各子区域的收敛条件; 分别统计其迭代完成任务量 ld_i ;
- (3) 求解子区域间的边界节点或公共子区域;
- (4) 如果整个区域达到其收敛条件, 则终止; 否则, 转(5);
- (5) 从 Set_s 中选择满足 $\tilde{ld}_i = \max\{\tilde{ld}_m, m \in Set_s\}$ 的处理机 M_i , 及处理机 M_l, M_l 满足 $\tilde{ld}_l = \min\{\tilde{ld}_m, m \in Set_s\}$. 如果 $\tilde{ld}_i - \tilde{ld}_l$ (δ 为处理机间容许的负载差), 转(2); 否则, 转(6);
- (6) 在 M_i 的相邻处理机集中选择 M_j , 满足 $\tilde{ld}_j = \min\{\tilde{ld}_m, m \in R_i\}$ (R_i 的相邻区域集记为 R_i), 如果 $\tilde{ld}_i - \tilde{ld}_j > \delta$, 需在域 i 与其邻域 j 之间重新划分区域, 使得处理机 M_i 的部分任务迁移到 M_j 上, 迁移任务量约为: $ld_i - \tilde{ld}_k$; 否则, 转(7);
- (7) $Set_s = Set_s - \{i\}$, 如果 $Set_s = \emptyset$, 转(5); 否则, 重置 $Set_s = \{1, 2, \dots, n\}$, 转(2)。

其中, $\tilde{ld}_i, \tilde{ld}_k$ 分别表示与处理机 M_i 相邻的处理机任务分布均值和 k 时刻所有处理机任务分布均值。

上述过程是根据上一时间阶段的计算历史, 将本时间阶段中子区域划分进行调整, 以期达到计算负载的平衡, 若负载差值小于经验值, 则只需传送与该层有关的初值, 可以利用上次区域划分信息, 减少通信量。在油藏数值模拟中, 这种负载均衡策略很有效, 各处理机的任务较先前的均衡, 由负载均衡引入的计算量不大。由于油区的物理性质一般是稳定的, 应用负载均衡分区方法后, 油区划分基本上逐渐趋于稳定, 而整个过程中增加的仅是油区边界的移动而引起的数据通信开销、计算时间的统计和均值计算, 子区域间仅交换重叠值, 通信量较少。采用分区负载均衡方法主要优点是在每时间阶段压力计算过程中没有处理机间通信, 只在整体预处理矩阵形成之后增加子区域划分、分配与收集计算结果并根据需要再分配的少量时间。对于每个子区域问题形成的稀疏非对称线性方程组的求解, 可以采用改进的预处理 GMRES(m) 算法^[8]。

4 测试结果

我们使用 10Mbps 以太网连接 SGI 工作站和 SUN 工作站 (主频 150M, 内存 512M) 构成一个含共享内存多处理机的分布式并行系统, 通过 PVM 连接成一个虚拟的网络计算环境。这里定义并行求解压力的加速比 S_p 为在单处理机上的求解时间与在分布式并行系统上的求解时间之比。

设 N_x, N_y, N_z 分别表示油层 X, Y, Z 方向上的网格划分。

模型 1 实际油藏第一层静态参数分布为地质模型, 动

态数据取自油藏实际数据, N_x, N_y, N_z 分别为 70、21、5, 节点规模为 7350。

模型 2 实际油藏模型, 油藏动静态数据完全取自现场实际。模型纵向分为 5 层, N_x, N_y, N_z 分别为 170、50、5, 节点规模为 42500。

上述两个模型形成的非对称线性方程组系数矩阵阶数分别是 7350 和 42500。因为 Z 方向的网格数最少, 形成的子区域并行度大, 所需增加的存储量也少, 所以, 当应用区域分解算法并行计算时, 模型沿平行于 Z 方向剖分区域, 这时整个区域收敛判别条件为: $p^{i+1} - p^i \leq 10^{-3}$, p^{i+1}, p^i 分别为第 $i+1$ 迭代步和第 i 迭代步的压力向量。各子区域采用改进的 MLU 预处理 GMRES(m) 求解, 取 $m = 10$, 子区域收敛判别条件为: $p_m - p_l \leq 10^{-3}$, p_m, p_l 分别为区域的每一迭代步中 GMRES(m) 算法的第 m 迭代步和第 l 迭代步的压力向量。测试环境为: 环境 1, 串行执行, 环境 2, 2CPU 共享并行系统, 环境 3, 4CPU 分布式并行系统。采用文中给出的并行计算方法和负载均衡策略求解压力方程, 此外为了在实际计算中进一步提高整体运行效率采取数据压缩存储与传输以及整体代码优化。两种模型在上述几种测试环境下实例运算结果如下:

表 1 解模型 1 时间对比 (计算 30 年生产历史, 下同)

时间/s	环境 1	环境 2	环境 3
压力求解时间/s	75.43	46.7	25.3
总运行时间/s	235.64	173.5	165.7

表 2 解模型 2 时间对比

时间/s	环境 1	环境 2	环境 3
压力求解时间/s	1600.83	874.7	507.8
总运行时间/s	4100.23	3010.48	2907.1

表 3 压力方程组求解加速比

	环境 2	环境 3
模型 1	1.62	2.98
模型 2	1.83	3.15

对黑油模型 2, 基于环境 2 解压力方程组, 并行解压力方程组的加速比可达 1.83 (表 3), 并行效率达 91.5%, 尤其是与原串行软件解压力方程组速度相比减少了近一半时间, 大大缩短了解压力方程组时间, 提高了油藏数值模拟计算速度和能力。

从表 1 和表 2 中可以看出, 子区域内解压力方程组采用改进预处理 GMRES(m) 算法特别当问题规模增大时, 算法的优势更加明显。另外初始值以及问题的性态对改进预处理 GMRES(m) 算法影响并不明显。改进的预处理 GMRES(m) 算法收敛判别相差很大, 因为在区域分解算法一次区域迭代步中, 子区域采用改进的预处理 GMRES(m) 算法, 解得在 m 维超平面 $x_0 + \text{span}\{p_1, p_2, \dots, p_m\}$ 上的极小点, $p_m - p_l$ 与 m 的值有关, 而且改进的预处理 GMRES(m) 算法的每步迭代的下降幅度大得多。所以在解压力方程组时, 并行软件选取

GMRES(m) 算法收敛判别一般是原串行软件收敛判别的 $m \times 10^1 \sim m \times 10^2$ 倍。

在区域分解算法并行计算的每一迭代步中, GMRES(m) 算法参数 m 选取一般不益太大, 且随着迭代进行, m 可以减小, 这样能够减少不必要的迭代次数和计算量, 这是由于黑油模型数值模拟问题每一时间步跨度不大, 一般情况下每一计算时间步的初始值接近近似结果。此外, 一般情况下, 随着问题规模的增大, 最优的 m (记: M_{opt}) 也随之增大。若 $m \gg M_{opt}$, 可能会出现许多不必要的迭代, 既延长了运行时间, 又占用了较多存储空间; 若 $m \ll M_{opt}$, 程序中的控制迭代将频频发生作用, 强行终止程序的执行, 也降低求解精度。而我们取 GMRES(m) 算法参数 m 的初始值为 $m_0 \in [10, 20]$, 且随着迭代进行, m 减小, 一般第 i 步取 $m = m_0 \times 2^{-i}$, 当 m 小于 5 时, 令 $m = 5$ 。实验证实: 使用上述 m 的选取策略时, 可以较大地缩短运行时间。

整体预处理增加一次内迭代计算量, 但减少了外迭代次数; 同时, 由于预处理后的系数矩阵条件比原系数矩阵的更好, 较大提高了交替 Schwarz 算法的收敛速度, 这时其优越性很明显。在实际模型计算中, 当收敛条件较宽时, 整体预处理交替 Schwarz 算法得到更高的并行加速比。

感谢清华大学计算机系舒继武博士在研究工作中给予的帮助。

参考文献:

- [1] Wheeler M F, et al. A parallel multiblock/ multidomain approach for reservoir simulation[A]. SPE 51884, the 1999 SPE reservoir simulation symposium[C]. Houston, Texas, 1999. 51 - 61.
- [2] 蒋其垠. 21 世纪初石油信息技术应用展望[J]. 石油化工动态, 2000, 8(2): 12 - 16.
- [3] Lbragimov A L, et al. Domain decomposition methods and application to modeling of the performance of horizontal and slanting wells in anisotropy reservoirs[A]. SPE 51906, the 1999 SPE reservoir simulation symposium[C]. Houston, Texas, 1999. 269 - 278.
- [4] 莫则尧, 等. 优化和并行一个油藏数值模拟软件中的解法器[J]. 石油学报, 2000, 21(2): 56 - 61.
- [5] Shiralkar G S, et al. Falcon: A production quality distributed memory reservoir simulator[J]. SPE Reservoir Evaluation and Engineering, 1998, 10: 400 - 405.
- [6] Franklin M A, et al. General matrix iterative model for dynamic load balancing[J]. Parallel Computing, 1996, 22(7): 969 - 989.
- [7] 舒继武, 等. 非均匀区域油藏模拟负载均衡的分区并行算法[J]. 软件学报, 1999, 2: 187 - 192.
- [8] Morgan R B. A restarted GMRES method augmented with eigenvectors[J]. SIAM J Matrix Anal. 1995, 16: 1154 - 1171.

作者简介:



刘青昆 男, 1971 年生于青海, 1992 年在北京大学计算机系软件专业获学士学位, 1998 年在南京大学计算机系软件专业获硕士学位, 现为南京大学计算机系博士研究生, 主要研究兴趣为并行处理与分布式计算。

张德富 男, 1937 年生于江苏, 1962 年毕业于南京大学数学系计算技术专业, 现为南京大学计算机系教授、博士生导师, 主要从事计算机软件、并行处理与分布式计算等方面的研究。