

文章编号:1001-9081(2010)06-1687-03

## 混合并行技术在激光化学反应模拟中的应用

李鸿健<sup>1,2</sup>, 白明泽<sup>2</sup>, 唐红<sup>2</sup>, 孙世新<sup>1</sup>

(1. 电子科技大学 计算机科学与工程学院, 成都 610054; 2. 重庆邮电大学 计算化学研究所, 重庆 400065)

(richardlee@uestc.edu.cn)

**摘要:**为提高激光化学反应模拟效率,在半经典分子动力学模拟中引入混合并行技术和双层并行思想。基于MPI + OpenMP混合模型设计并实现激光化学反应双层并行模拟算法,上层基于MPI实现节点间的原子分解并行,下层基于OpenMP实现节点内的多线程矩阵并行乘法。在SMP集群中测试表明,模拟大分子体系激光化学反应并行效率可达60%以上。因此,应用混合并行技术可有效提高激光化学反应模拟效率。

**关键词:**半经典分子动力学;混合并行技术;SMP集群;双层并行

**中图分类号:** TP338 **文献标志码:** A

### Application of hybrid parallel algorithm for simulating photochemical reaction

LI Hong-jian<sup>1,2</sup>, BAI Ming-ze<sup>2</sup>, TANG Hong<sup>2</sup>, SUN Shi-xin<sup>1</sup>

(1. School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu Sichuan 610054, China;

2. Institute of Computational Chemistry, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** A high performance hybrid parallel algorithm for simulating photochemical reaction was developed by introducing the concept of two layers of parallel capacity. The hybrid parallel algorithm was designed based on MPI + OpenMP parallel model. MPI was used to achieve the atomic decomposition, while the OpenMP multithreading was used to achieve the matrix multiplication. The efficiency of the photochemical reaction simulating with large scale atoms can achieve 60% by being tested on SMP cluster. It was proved that this method is a feasible and efficient parallel algorithm for simulating the photochemical reactions.

**Key words:** semiclassical molecular dynamics; hybrid parallel technique; Symmetrical Multi-Processing (SMP) cluster; double-level parallelism

## 0 引言

激光化学反应模拟是指利用计算机模拟激光脉冲诱导的光化学反应过程,在原子水平上观察光化学反应,追踪分子的变化细节,为实验提供激光控制化学反应的直接信息<sup>[1]</sup>。激光化学反应模拟研究是目前研究的热门课题。目前应用最为广泛的激光化学反应模拟方法是半经典分子动力学模拟方法<sup>[2]</sup>,当原子和价电子数量增加时,半经典动力学模拟计算量巨大,因此,并行技术应用于半经典动力学模拟并提高模拟效率具有重要意义。

激光化学反应计算的主要部分是分子动力学模拟,目前对于分子动力学并行计算主要采用原子分解算法、空间分解算法和作用力分解算法3种算法<sup>[3]</sup>。原子分解算法是指为每一个处理器或节点分配一个原子子集,无需考虑原子所处的实际物理位置。如CHARMM和GROMOS等大型分子动力学软件就是采用原子分解的方法<sup>[4]</sup>。空间分解算法将模拟计算的单胞划分成多个子空间,每个子空间被分配给一个处理器,该处理器负责计算分配的子空间中原子的受力及位置变化<sup>[5-6]</sup>。作用力分解算法将计算原子受力的矩阵进行按块分割,从而相对减少了内存的消耗和相互间通信的处理器数目<sup>[7-8]</sup>。分子动力学模拟的各种算法均存在不同的缺陷:原子分解需要全局通信<sup>[4]</sup>,空间分解实现负载均衡较为困

难<sup>[5]</sup>,力分解算法无法处理大规模分子体系<sup>[3]</sup>。上述分子动力学模拟均采用单层并行,单层并行的并行效率在SMP集群中提高有限,结合半经典分子动力学模型和SMP集群的特征采用双层并行设计模拟算法可有效提高并行效率。

本文采用MPI + OpenMP的混合并行技术,设计并实现基于半经典分子动力学模型的激光化学反应并行算法。该算法充分利用了两种并行技术的优势,将任务划分为分布和共享双层并行,在SMP集群实现层次并行。实验测试表明,在计算节点内采用OpenMP并行具有显著优势,基于MPI + OpenMP混合并行计算可有效提高激光化学反应模拟效率。

## 1 激光化学反应模拟计算

### 1.1 计算模型

采用半经典动力学模型模拟激光化学反应,在该模型中,电子运动过程由量子力学描述,核运动过程由经典力学描述。半经典动力学模型需要进行复杂的计算,包括核动力学计算、电子运动计算、激光脉冲与电子耦合计算。原子核运动的力 $F$ 由广义的Ehrenfest方程来描述,而Ehrenfest方程中涉及到求取哈密顿矩阵 $H$ 、重叠矩阵 $S$ 和离子-离子排斥势 $U$ 相对于每一个原子位置的梯度。电子运动由含时薛定谔方程描述,该方程可以通过基于演化时间算符的西算法迭代方案<sup>[9]</sup>求解。哈密顿矩阵元 $H$ 是原子核位置向量的函数,而驱动原子核运动

收稿日期:2009-12-25;修回日期:2010-03-15。 基金项目:国家自然科学基金资助项目(20773168)。

**作者简介:**李鸿健(1981-),男,重庆人,讲师,博士研究生,主要研究方向:并行算法设计;白明泽(1982-),男,重庆人,讲师,博士研究生,主要研究方向:分子动力学并行算法;唐红(1957-),女,江西南昌人,教授,博士生导师,主要研究方向:网络计算;孙世新(1940-),男,湖北孝感人,教授,博士生导师,主要研究方向:并行与分布式计算。

的力由分子的电子状态确定,因此电子运动和原子核运动相互耦合。哈密顿矩阵元  $H$  的积分应用密度泛函约束束缚(Density Functional based Tight Bonding, DFTB)近似数值求解获得<sup>[10]</sup>。

基于上述模型,计算从输入分子构型和激光参数开始,到输出本次反应的产物为止<sup>[11]</sup>。计算过程分为3个步骤:激光辐射前分子构型平衡的计算、激光与分子相互作用的计算和激光源关闭后分子反应的计算。基于本模型的激光化学模拟计算是一个时间步循环,程序的主体是半经典分子动力学计算部分。

图1显示了C60在激光作用下发生裂解的化学反应模拟。激光化学反应模拟的主要计算部分为:计算核运动的力  $F$ 、哈密顿矩阵元  $H$  和求解哈密顿矩阵元  $H$  的特征值<sup>[12]</sup>,如图2所示。在上述计算中,核运动力  $F$  的计算占单步循环计算时间的90%以上<sup>[6]</sup>。因此并行设计并优化核运动力  $F$  的计算可有效提高激光化学反应模拟效率。

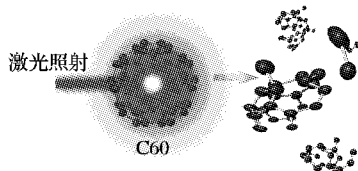


图1 C60 激光化学反应模拟

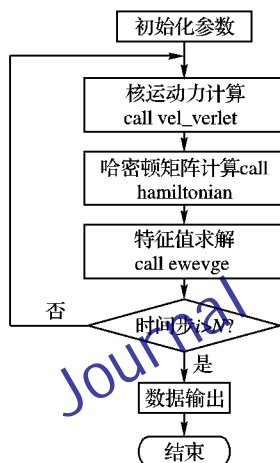


图2 激光化学反应模拟单步计算

## 1.2 核运动力 $F$ 的计算

核动力  $F$  的计算是对  $m$  个原子进行  $N \times N$  的矩阵计算,在三维立体空间中,外层循环次数为  $3m$ ,  $N$  表示分子系统中所有原子外层电子数总和。例如,模拟1000个C原子激光化学反应中,在单步计算中,核动力  $F$  计算需要进行3000次  $4000 \times 4000$  的矩阵计算,而在激光化学反应模拟计算中,每飞秒通常需要计算20步,完整进行一次激光化学反应模拟需要计算上千飞秒,因而计算十分费时。

## 2 混合并行算法

### 2.1 混合并行模型

混合并行模型具有双层结构<sup>[13]</sup>:上层的MPI表示节点间的并行;下层的OpenMP表示节点内的多线程并行。MPI适用于粗粒度并行<sup>[14]</sup>,每个进程只能读写本地内存中的数据,进程间通过消息传递进行通信,实现对远程数据的访问。而对于MPI进程内的循环使用OpenMP共享结构并行,该并行方式也称为OpenMP细粒度并行<sup>[15-16]</sup>。

基于两个节点的集群的混合并行模型如图3所示:在每个

节点上运行一个MPI进程,MPI进程初始化后,每个节点上的MPI进程独立进行局部计算,节点间可以进行通信;在MPI进程内的主要计算部分采用OpenMP多线程并行计算;在OpenMP多线程计算完成后,MPI进程进行局部计算、通信或同步。

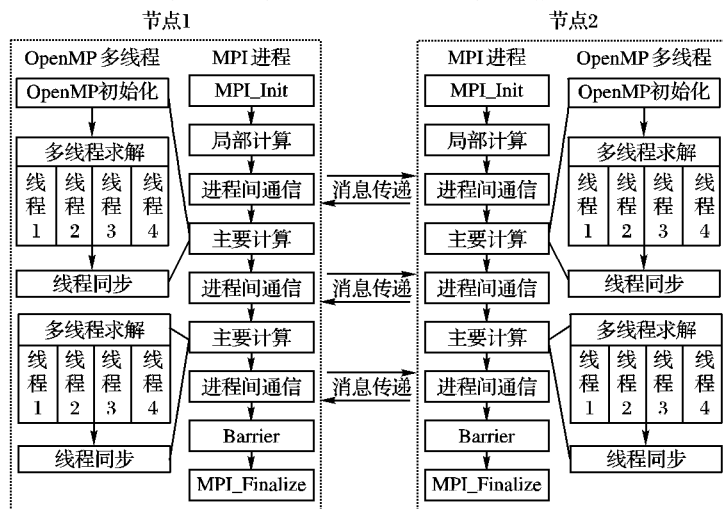


图3 混合并行模型

### 2.2 混合并行算法设计

激光化学反应并行模拟采用原子分解算法实现粗粒度并行。在单步循环中,不同原子核运动受力  $F$  的计算没有先后顺序,可采用原子分解算法进行并行处理。经典分子动力学并行模拟主要采用原子分解和空间分解实现并行,但对于激光化学反应模拟,模拟过程中原子区域分布不均匀,空间分解并不适合。另外,同经典分子动力学模拟相比,原子数量规模较小,单个原子的涉及的矩阵乘法计算,其计算量较大。因此原子分解并行算法能较好地实现各节点负载均衡。

激光化学反应并行模拟在节点内采用共享内存并行模式实现各原子涉及的矩阵乘法的并行计算,即细粒度并行,基于OpenMP并行接口,采用行列分解算法实现矩阵相乘。

图4显示了激光化学反应混合并行算法。首先进行原子分解。该方法实现简单,能够解决负载均衡问题,且不会由于通信开销较大而影响整体并行计算效率。再利用OpenMP实现节点内的多线程计算,将计算矩阵  $A$  按行进行划分给节点内  $k$  个处理器,使用  $k$  个线程计算矩阵乘法。

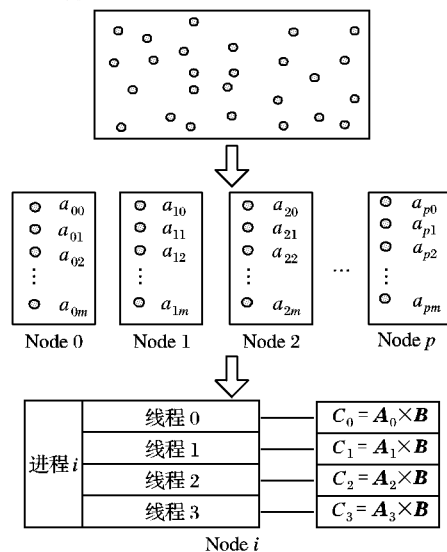


图4 激光化学反应混合并行算法

### 2.3 算法分析

混合并行算法主要特点是:1) 混合并行解决了单一消息传递模式不易实现负载均衡的问题,提高了并行程序的性能

和可扩展性;2)混合并行模式有效解决了细粒度并行带来的通信密集的问题,通过两级并行,有效提高并行效率;3)上述混合并行算法充分考虑了激光化学反应模拟的特点,即单步循环内各原子核运动受力 $F$ 的计算没有先后顺序,原子和涉及计算的矩阵可进行分割,结合两种并行计算模式的优势进行算法设计。

### 3 测试及分析

#### 3.1 测试环境

IBM BladeCenter 刀片服务器是 IBM 高性能运算领域的主力产品,广泛应用于科学计算、高端商用计算,以及高可用集群领域。测试系统配置一个 IBM BladeCenter H 刀片中心,14 个刀片服务器作为计算节点,每个节点均支持 SMP,配置 2.0 GHz 4 核处理器和 4 GB 内存。

#### 3.2 测试及分析

基于上节所述系统,首先测试单节点内消息传递模式与共享内存模式的并行效率,选取 200 至 500 个 C 原子组成的碳纳米管分子进行激光化学反应模拟。测试结果如表 1 所示, $N$  表示原子数量,在单个计算节点内,采用 MPI 并行并不能大幅减少运行时间,甚至当参与计算的 CPU 数量增加至 4 时,运行时间反而比单 CPU 运行时间更长。这是由于当采用 MPI 实现细粒度并行,每个进程占用一个 CPU,进程之间要求在每个时间步进行全局通信,当 CPU 数量增加时,通信开销增加较大,其并行效率迅速下降。OpenMP 并行模式适用于细粒度并行,当线程数量增加时,模拟时间迅速下降,有效提高了运行效率。因此,在节点内采用共享内存模式并行可有效提高单节点内运行效率。

表 1 OpenMP 与 MPI 并行模式下激光化学反应模拟时间/s

CPU 个数	$N=200$		$N=300$		$N=400$		$N=500$	
	OpenMP	MPI	OpenMP	MPI	OpenMP	MPI	OpenMP	MPI
1	396.6	396.6	2384.3	2384.3	4121.3	4121.3	6569.8	6569.8
2	272.0	291.7	2187.7	2294.6	4041.8	4023.1	6454.6	5909.1
3	212.2	318.6	1509.1	2094.8	2818.3	3576.3	4599.3	5592.8
4	187.8	812.3	1164.2	2426.5	1925.2	4167.0	3337.1	6557.9

利用混合并行进行大分子的激光诱导光化学反应进行模拟测试。加速比计算公式为  $S_N = T(1)/T(N)$ ,当模拟特定的分子体系时, $T(1)$  表示一个计算节点的运行时间, $T(N)$  表示  $N$  个计算节点参与的并行时间。如图 5 所示,随着计算节点增加,加速比均出现不同程度的增长,对于原子数量越多的体系,其通信时间在总运行时间中所占比例相对较少,其加速比上升速度更快。当原子数量等于 200 时,随着节点数量的增加,其加速比增长缓慢,计算效率下降。因此,在激光化学反应并行计算中,针对包含不同规模原子的分子体系,采用适当数量的节点进行计算可提高并行效率。在该测试中,对于  $N=200,300,400,500$  的分子体系,分别使用 2,4,6,8 个计算节点,均可获得 60% 以上的并行效率,因此该混合并行算法在保证并行效率前提下具有可扩展性。结果表明,应用混合并行技术可有效提高激光化学反应模拟效率。

### 4 结语

采用 MPI + OpenMP 的混合并行技术,设计并实现基于半经典分子动力学模型的激光化学反应并行算法,该算法充分利用了两种并行技术的优势,将任务划分为分布和共享双层并行,实验结果表明,应用混合并行技术的激光化学反应模拟

可获得有效提高并行效率。本文是在 SMP 集群上进行的一次 MPI + OpenMP 混合并行的应用,对于提高激光化学反应并行模拟效率的研究工作具有现实的参考意义。

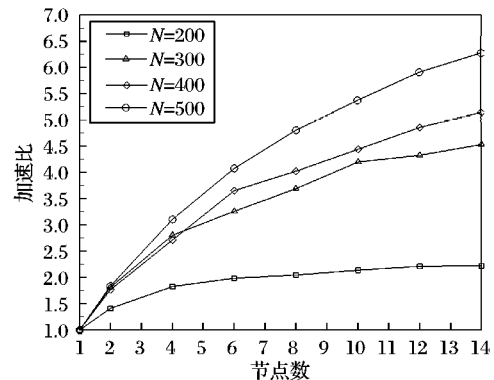


图 5 激光诱导光化学反应并行模拟加速比

#### 参考文献:

- [1] ZEWAIL A Z. Femtochemistry: Ultrafast dynamics of the chemical bond [M]. Singapore: World Scientific, 1998: 35–38.
- [2] WORTH G A, ROBB M A. Applying direct molecular dynamics to non-adiabatic systems [J]. Advances in Chemical Physics, 2002, 124(1): 355–432.
- [3] 徐伟,李玉忱,王丽. 分子动力学并行算法研究[J]. 计算机工程与应用, 2002, 38(14): 218–221.
- [4] 刘正华,杨决宽,陈云飞. 分子动力学并行算法的优化与应用[J]. 计算机应用研究, 2008, 25(3): 718–720.
- [5] BOWERS K J, EDMOND C, XU H F, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters [C]// Proceedings of the 2006 International Conference on Parallel Processing. New York: ACM, 2006: 84–96.
- [6] THOMPSON P, PLIMPTON S J, MATTSON W, et al. General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions [J]. Journal of Chemical Physics, 2009, 131(1): 154107.
- [7] SUMANTH J V, DAVID R S, HONG J. Adaptive load balancing for long-range md simulations in a distributed environment [C]// Proceedings of the 2006 International Conference on Parallel Processing. New York: ACM, 2006: 135–146.
- [8] SUMANTH J V, DAVID R S, HONG J. Adaptive load-balancing for force-decomposition based 3-body molecular dynamics simulations in a heterogeneous distributed environment with variable number of processors [C]// Proceedings of the 2007 International Conference on Parallel Processing. Washington, DC: IEEE Computer Society, 2007: 552–565.
- [9] RABITZ H, DE V R, MOTZKUS M, et al. Whither the future of controlling quantum phenomena [J]. Science, 2000, 288(5467): 824–828.
- [10] TULLY J. Mixed quantum-classical dynamics [J]. Farad Discuss, 1998, 110(1): 407–419.
- [11] 李鸿健,豆育升,唐红,等. 激光诱导化学反应的并行计算机模拟[J]. 中国激光, 2009, 36(2): 356–361.
- [12] LI H J, TANG H, DOU Y S, et al. A parallel algorithm for simulating photochemical reaction [C]// 2008 IFIP International Conference on Network and Parallel Computing. Washington, DC: IEEE Computer Society, 2008: 258–262.
- [13] 赵永华,迟学斌,程强. SMP 集群系统上矩阵特征问题并行求解器的有效算法[J]. 计算机研究与发展, 2007, 44(2): 334–340.
- [14] 宋伟,宋玉. 基于 SMP 集群系统的并行编程模式研究与分析[J]. 计算机技术与发展, 2007, 17(2): 164–171.
- [15] 王惠春,朱定局,曹学年,等. 基于 SMP 集群的混合并行编程模型研究[J]. 计算机工程, 2009, 35(3): 271–273.
- [16] 刘瑜,梁正,杨梓强. 混合并行技术在 FDTD 计算中的应用研究[J]. 电子科技大学学报, 2009, 38(2): 222–226.