

基于深腾1800机群系统的分子动力学并行仿真研究

郭晓光*, 郭东明, 康仁科, 金洙吉

(大连理工大学 精密与特种加工教育部重点实验室, 辽宁 大连 116024)

摘要: 介绍了分子动力学并行仿真计算的软硬件环境, 分析了现有的几种并行算法, 确定采用区域分解法作为并行算法, 并在此基础上提出了基于区域二次划分的分子动力学并行仿真算法. 另外, 阐述了原子链、原子近邻表和原子亲属表的概念, 提出了基于永久序号的消息传递策略. 最后, 设计了分子动力学并行仿真程序, 并分别在1、2、3、4台结点机上进行了实验, 运行结果表明: 加速比随着结点数的增加而增加, 并行效率虽略有下降但都在87.5%以上, 并行效率并没有随着结点数的增加有明显的降低, 说明并行程序具有很好的扩展性.

关键词: 分子动力学仿真; 并行算法; 区域分解法

中图分类号: TG580.1; TP311.1 **文献标识码:** A

0 引言

分子动力学方法是在原子级、分子级模拟固态、液态及气态物质的主要计算方法, 它提供了一条从系统的微观细节探索实验过程宏观特性的捷径, 其具有可实验性、安全性, 可进行超前研究, 减少实验量以及节省大量人力物力资源等特点, 目前在物理、化学、生物、医药以及材料等多个研究领域得到了广泛应用, 并逐渐渗透到机械加工领域. 国内应用分子动力学方法研究机械加工机理的工作起步较晚, 只有清华大学、天津大学和哈尔滨工业大学等几所大学开展了分子动力学仿真的研究, 尚处于起步阶段^[1~3]. 国外方面, 美国、日本和澳大利亚等国的研究者对切削、压痕等过程进行了深入研究, 但在磨削方面研究得比较少^[4~6]. 总的来说, 分子动力学仿真在机械加工领域虽然已经取得了很大进展, 但对于单晶硅的超精密磨削过程机理研究还没有一个统一的完整认识, 尚需要进一步深入研究.

另外实际研究表明, 分子动力学仿真结果的可靠性主要取决于仿真系统所选取的计算分子间相互作用的势能函数和仿真系统的粒子数目, 因为在计算过程中, 粒子个数只有在尽可能多时, 才能比较真实地反映物质的宏观行为, 但是分子动

力学模拟属于计算机模拟中典型的计算瓶颈问题, 应用并行计算可将仿真规模扩大几十甚至上百倍, 因此分子动力学仿真计算是分子动力学仿真的必由之路. 国内外很多学者于20世纪80年代末90年代初开始致力于分子动力学的并行计算研究^[7~10], 但应用深腾1800针对单晶硅磨削过程进行分子动力学仿真未见报道.

本文设计以区域分解法为基础的基于区域二次划分的分子动力学仿真并行算法, 提出原子亲属表和基于永久序号的消息传递等策略和方法, 编制分子动力学并行仿真程序, 并在深腾1800机群系统上进行仿真计算, 对仿真结果进行可视化处理、加速比和并行效率的分析.

1 并行算法研究

1.1 并行计算环境

并行计算环境一般由硬件平台和软件平台两个部分组成, 其中硬件平台是指计算机的硬件系统, 而软件平台是指编程和运行的环境.

1.1.1 硬件环境 目前的硬件平台主要是并行机或机群系统, 如共享存储对称多处理机(SMP)、分布式共享存储多处理机(DSM)、大规模并行处理机(MPP)、并行向量多处理并行机

收稿日期: 2006-03-19; 修回日期: 2007-09-15.

基金项目: 国家自然科学基金资助项目(重大项目50390061); 国家杰出青年科学基金资助项目(50325518).

作者简介: 郭晓光*(1976-), 女, 博士生, 讲师.

(PVP) 和机群系统(COW)。一般并行机价格比较昂贵,扩展性差,因此应用受到极大限制。

目前机群系统应用非常广泛,其具有良好的扩展性,计算性能可相当甚至超过现有的超级计算机(如 MPP 机),并具有非常高的性能价格比,是并行计算环境今后的发展趋势。

基于上述原因并根据现有的硬件设备,本文决定采用大连理工大学工业装备结构分析国家重点实验室新购进的联想深腾 1800 机群系统。其采用主流的 Cluster 结构作为其体系结构,结点机采用基于 Intel 架构的商用服务器,通过高速通信网络实现结点间的互联,对外提供单一系统映像,能够很好地支持大规模的科学工程计算,运算速度可达每秒数万亿次。

1.1.2 软件环境 对于软件平台而言,根据其编程机制,主要分为 3 类^[11]:消息传递机制、共享存储机制和数据并行机制。消息传递机制以 MPI 为代表,PVM 是消息传递模式的一个变种。共享存储机制以 OpenMP 为代表,主要是利用添加并行化指令到顺序程序中,由编译器完成自动并行化,但其成本高、扩展性差。数据并行机制比较少见,但以其独特的处理方式受到特定用户群的喜欢。共享存储机制和数据并行机制基于细粒度并行,仅被 SMP、DSM 和 MPP 并行机所支持,可移植性不如消息传递机制,但它们支持数据的共享存储,所以并行编程的难度较小,一般情况下,当处理机个数较多时,其并行性能明显不如消息传递机制。而消息传递机制基于大粒度的进程级并行,具有很好的可移植性和可扩展性,几乎被当前流行的各类并行机所支持,因此近年来基于消息传递的并行处理方式越来越受到重视。

文中使用的联想深腾 1800 机群系统,提供了标准的并行计算环境 PVM 和 MPI,由于分子动力学仿真并行计算是进程级大粒度的并行计算,本文中采用了 MPI 的编程环境,算法实现的编程语言为 C 语言。

1.2 并行算法分析

分子间相互作用力的计算占用了分子动力学仿真的大部分时间,因此,在分子动力学并行算法程序的设计中,主要解决的问题就是如何减少分子间相互作用力的计算时间。根据计算载荷分配的不同,分子动力学并行算法主要分为 3 种:原子分解法(atom decomposition)、原子间作用力分解法(force decomposition) 和区域分解法(domain

decomposition)^[12~14]。

对比这 3 种算法,原子分解并行算法适用于现有的串程序的并行化,并且易于实现负载均衡问题,但这种算法的缺点是需要全局通信,导致通信成本高,因此仅适合规模比较小或长距离作用力问题,不易扩展。作用力分解算法是原子分解法的一种改进,具有原子分解法的优点,且通信成本较低,节省了大量内存空间,因此它比原子分解法计算规模要大,但仍不适合计算规模大以及处理机台数较多的并行机。与原子分解法和作用力分解法相比较,区域分解法只与近邻处理机之间进行通信处理,通信时间耗费少,内存占用少,可以明显增加分子动力学模拟的原子数量,并且具有很好的扩展性,适合计算规模大和处理机台数多的情况,故最终确定区域分解法作为本文分子动力学并行仿真设计的算法基础。

1.3 空间区域的划分

区域分解法是本文分子动力学并行仿真的基本算法,其空间区域的划分必须要根据仿真模型的具体情况和并行计算的需要来进行。在本文的并行程序设计过程中,将整个仿真模型中第一次划分的区域称为计算区(每个计算区对应一个计算进程),将每个计算区内第二次划分的区域称为计算单元。

基于以上因素本文提出了空间区域(计算区和计算单元)的二次划分策略,如图 1 所示,即首先对整个仿真模型进行一维的计算区的划分,然后采用立方网格区域法将各个计算区均分为各计算单元。具体来说,对整个分子动力学仿真模型进行区域划分时(设有 P 个计算结点),首先在 x 方向平均划分出 P 个大小基本相同的计算区,每个计算区对应一个计算结点,整个模型中的原子根据其坐标被分配于各个计算区中,此次为第一次区域划分。根据同样原则,在每一个计算区中再进行第二次划分,从而划分出计算单元。每一

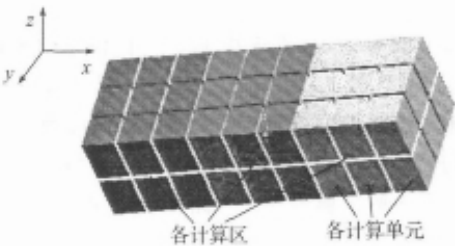


图 1 区域划分示意图

Fig. 1 Schematic diagram of domain partitioning

个计算单元均为一个长方体,其最小边长应大于或等于截断半径,所以每个计算区中计算单元的数量取决于计算区本身尺度的大小. 通过以上划分可以很方便地以一个计算单元为基本单位进行消息的传递,有效地减少了通信开销.

1.4 基于区域二次划分的分子动力学并程序序设计

在空间区域划分的并行算法基础上,本文提出了基于空间区域二次划分的分子动力学仿真并行算法,在算法中加入了原子链、原子近邻表、原子亲属表和基于永久序号的消息传递等策略和方法.

所谓原子链在程序中表现为存储原子信息的数组,本文的并行程序中为每一个计算单元都建立了原子链以储存该计算单元内的原子信息,同样,对于每个计算区也有一个原子链与之相对应以方便对整个计算区的访问.

原子近邻表即是对目标原子可能产生作用力的所有原子的集合. 在近邻表中通过截断半径(r_c)来判断有哪些原子对目标原子产生作用力进而计算作用力. 在分子动力学模拟时,为减少分子间作用力的计算时间,假定一个球形的截断半径 r_c . 当两个分子的间距 $r > r_c$ 时,则不考虑原子间的作用力,即两分子间的势能为零,从而避免对模拟系统中的所有分子进行逐个计算,一般取 $r_c = 2.5\delta$, δ 为粒子的直径,而且实际研究表明这个假设是合理的. 建立原子近邻表可以大大减少作用力计算过程中访问原子(即检验其是否对目标原子产生作用力)的数目,缩短计算时间,同时也减少了内存的使用量.

原子亲属表即原子作用力表,它是根据原子近邻表概念再细分出来的,是通过比较原子近邻表中原子与目标原子的距离和 r_c 的大小,从原子近邻表中挑选出与目标原子可以产生作用力的原子的集合. 对比原子近邻表,将原子作用力表称为原子亲属表. 建立原子亲属表,首先访问整个近邻表中的原子,计算表中原子与目标原子的距离 r_{nei} ,然后比较 r_{nei} 和 r_c 的大小,将 r_{nei} 小于 r_c 的原子加入到原子亲属表中,如图 2 所示,2 到 7 号原子与目标原子 1 的距离小于 r_c ,所以将它们加入到原子 1 的亲属表中. 因此原子作用力表是在原子近邻表的基础上建立起来的.

分子动力学并程序在通信的策略上,采取基于永久序号的方法. 永久序号的含义即是,从

原子位置的初始化开始,用从 1 到所有原子的总数 N 作为序号赋予每一个原子,每个原子被赋予序号之后,一直到整个计算结束其序号是恒定不变的,从而使得消息传递的操作变得简单,提高效率降低出错率. 在把计算区划分成若干个计算单元后,因为原子序号和其坐标、速度信息一一对应,就不用再特别对每个计算单元设置单独的原子链存储坐标和速度信息,这样大大节省了内存.

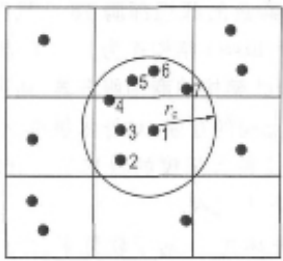


图 2 原子亲属表示意图

Fig. 2 Schematic diagram of atomic relative list

本文主要应用 PCAM 设计方法学,采用 MPI 的消息传递编程机制设计出基于空间区域二次划分的分子动力学仿真并程序,其逻辑流程如图 3 所示. 程序采用主从设计模式,即主进程的主要任务是位置和速度的初始化、计算区的划分,以及计算过程中和最终结果的收集和总结处理等;从进程则完成并行计算(包括力的计算、位置和速度的更新等主要部分),但主进程和从进程的运行节点并无主从区别. 各进程(节点)间通过消息传递来共享数据.

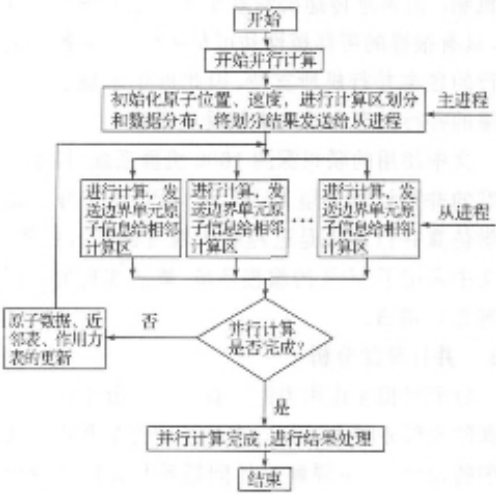


图 3 分子动力学仿真并程序序流程图

Fig. 3 Flow chart of the parallel program of the MD simulation

本文中所应用的单晶硅磨削过程三维分子动力学模型、势函数以及单晶硅温度转换模型等见文献[15].

2 数据处理和结果分析

2.1 数据可视化

并行数据的可视化即是通过绘图方式,将整个加工的状况和过程直观地呈现出来,目前能够可视化的软件有 Matlab 和 OpenGL. 对于三维的分子动力学仿真,Matlab 的三维模型显示比较模糊,而 OpenGL 则能够很逼真地展现出整个加工模型原子的三维图像,所以本文采用 C++ 和 OpenGL 编写三维图形程序对仿真结果进行可视化处理. 通过对结果的可视化,可以直观感性地看出原子的走向趋势、切屑的形成过程及形状等.

下边以一组图片为例,来观察磨削过程中原子瞬间位置的变化及模型整体的形貌变化. 模拟过程中认为金刚石磨粒是刚性的,即不变形、不磨损. 图 4 是同一时刻不同方向的瞬间原子位置图,从两张图中可以直观地看出磨粒磨削过程中切屑走向趋势. 切屑不仅在磨粒运行方向前面发生堆积而且在磨粒的两侧还有侧积,图像还显示出磨粒经过后已加工表面会有微弹.

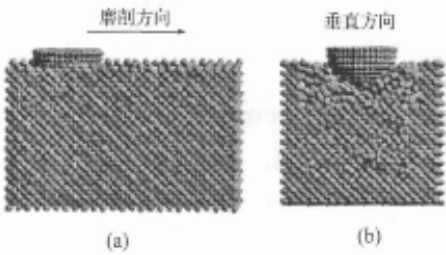


图 4 单晶硅并行仿真磨削过程中瞬间原子位置图

Fig. 4 Instantaneous distribution of monocrystal silicon atoms by the parallel simulation of the grinding process

2.2 加速比和并行效率分析

并行计算程序实现后,其中一个非常令人关心的问题就是程序的并行效率,即加速比和并行效率,故本文对并程序的并行效率做了初步的计算和分析. 加速比(S)是衡量一个并行系统和一台单机系统相对性能的标准,文中给出了使用多台结点机后,在速度上所获得的增加. 线性加速是并行算法加速比的“理想状态”,所谓线性加

速是指用 P 个处理器并行计算加速比值达到 P ,但是现实中是不可能达到这种状态的. 并行效率(E)一般可定义为实际加速比和理想加速比的比值,表示处理器用于计算的时间比例.

加速比和并行效率计算公式如下:

$$S = \frac{t_s}{t_p} \tag{1}$$

$$E = \frac{S}{P} \times 100\% = \frac{t_s}{t_p \times P} \tag{2}$$

式中: t_s 、 t_p 分别表示单机和并行机的执行时间.

表 1 是在 1、2、3、4 台结点机上运行相同规模的仿真结果,其中 a 为单晶硅的晶格常数,为 0.357 nm. 从表中可看出,加速比随着结点数的增加而增加,并行效率有微小下降但都在 87.5% 以上. 仿真规模扩大后,计算时间增加得不多,说明并程序的性能很好,并行效率并没有因为结点机的增加而出现明显降低的现象,故此并程序具有较好的扩展性.

表 1 分子动力学并行仿真结果

Tab. 1 The results of parallel MD simulation

结点数	工件尺寸	t/s	S	$E/\%$
1	$21a \times 9a \times 6a$	1.89×10^3		
2	$21a \times 9a \times 6a$	1.056×10^3	1.79	89.5
3	$21a \times 9a \times 6a$	0.71×10^3	2.66	88.7
4	$21a \times 9a \times 6a$	0.54×10^3	3.50	87.5
4	$26a \times 20a \times 10a$	1.59×10^3		

3 结 语

本文提出的基于区域二次划分的并行算法,进一步优化了区域分解法,并且兼顾了近邻列表方法和消息传递的优化设计,提高了作用力计算和消息传递的效率. 通过建立原子亲属表,降低了作用力计算时访问磨粒原子给编程带来的难度,同时解决了计算原子间 Tersoff 势时须考虑第三原子影响的要求,极大地简化了程序结构.

另外,本文还设计了基于永久序号的原子信息调用策略,因为每一个原子的序号在整个程序运行过程中保持不变而且是惟一的,所以任何原子之间信息的调用都不会相互干扰. 结果表明,基于永久序号的原子信息调用极大地简化了消息传递、原子迁移和近邻列表等的设计,提高了并行计算效率,并且降低了编程难度.

仿真结果可视化处理及在结点机上并行计算表明,随着结点数的增加,加速比也随之增加,并

且在4台结点机上运算时获得了3.5的加速比和87.5%的并行效率,证明并行程序具有良好的可扩展性。

参考文献:

- [1] 温诗铸. 纳米摩擦学进展[M]. 北京:清华大学出版社,1996:69-80
- [2] 林 滨,韩雪松,于思远. 纳米磨削过程中分子动力学计算机仿真实验[J]. 天津大学学报,2000,33(5): 652-656
- [3] 唐玉兰,梁迎春,霍德鸿,等. 单晶铜纳米切削过程的研究[J]. 纳米技术与精密工程,2004,2(2): 132-135
- [4] KOMANDURI R, CHANDRASEKARAN N. MD simulation of indentation and scratching of single crystal aluminum [J]. *Wear*, 2000,240: 113-143
- [5] SHIMADA S, TANAKA H, IKAWA N. Atomistic mechanism of surface generation in micromachining of monocrystalline silicon [C] // *Proceeding of the 1st International Euspen Conference*. Bremen: Shaker Verlag,1999: 230-233
- [6] CHEONG W C D, ZHANG L C. Molecular dynamics simulation of phase transformations in silicon monocrystals due to nano-indentation [J]. *Nanotechnology*, 2000(11): 173-180
- [7] 庄昌文,林晓东,刘心松. 实现并行计算机的机群系统[J]. 计算机应用,1998,18(6):11-13
- [8] 曹小林,莫则尧,张景琳,等. 基于“块-单元”数据结构的分子动力学并行计算[J]. 计算物理,2004,21(5):377-385
- [9] 董科军,刘让苏,郑采星,等. 液态金属原子大系统凝固过程模拟的并行算法研究[J]. 自然科学进展,2001,2:187-193
- [10] 吴江涛,刘志刚,董小社,等. RS/6000机群系统中分子动力学并行算法的研究[J]. 工程热物理学报,2002,23(4):418-420
- [11] 陈国良. 并行计算结构[M] // 算法编程. 北京:高等教育出版社,1999
- [12] SHU Ji-wu, WANG Bing, CHEN M S. Optimization techniques for parallel force-decomposition algorithm in molecular dynamic simulations [J]. *Comp Phys Commun*, 2003,152(2):121-130
- [13] MURTY R, OKUNBOR D. Efficient parallel algorithms for molecular dynamics simulations [J]. *Parallel Comput*, 1999, 25:217-230
- [14] YAO Zhen-hua, WANG Jian-sheng, LIU Gui-rong. Improved neighbor list algorithm in molecular simulations using cell decomposition and data sorting method [J]. *Comp Phys Commun*, 2004,161(1-2):27-35
- [15] GUO Xiao-guang, GUO Dong-ming, KANG Ren-ke, *et al*. Theoretical and experimental analysis on super precision grinding of monocrystal silicon [J]. *Front Mech Eng China*, 2007, 2(2): 137-143

Study of MD parallel simulation based on Shenteng 1800 COW

GUO Xiao-guang*, GUO Dong-ming, KANG Ren-ke, JIN Zhu-ji

(Key Lab. for Precis. and Non-tradit. Mach. Technol. of Minist. of Edu.,
Dalian Univ. of Technol., Dalian 116024, China)

Abstract: The hardware and software environments for parallel algorithm of molecular dynamics (MD) simulation were discussed in detail firstly. Then some popular parallel algorithms were analyzed and the domain decomposition was selected for the parallel algorithm of MD simulation. Based on the domain decomposition, a twice domain partition algorithm was proposed for MD parallel simulation. Moreover, an atom chain, an atom neighbor list and an atom relative list were introduced, and an information visiting strategy based on the proposed permanent sequential number for each atom was designed. Lastly, MD parallel program was made and executed on 1,2,3,4 node computers respectively. The experimental results show that when the number of process increases, the accelerating ratio increases and the efficiency decreases a little but is still more than 87.5%. It means that the efficiency doesn't decrease obviously with the increase of the node number and the parallel program has good expansibility.

Key words: molecular dynamics simulation; parallel algorithm; domain decomposition