

PC 集群的建立与 MPI 并行环境的实现及其应用

王莹 屈一新

(北京化工大学化学工程学院, 北京 100029)

摘要: 通过网络集线器, 采用以太网形式, 建立了由多台微机(PC) 组建的 PC 集群。采用 TURBO-LINUX 操作系统, 使 MPI 并行环境在该 PC 机群上得以实现。利用了该系统实现了分子动力学程序的并行计算。

关键词: 并行计算; PC 集群; MPI 并行环境; 分子动力学

中图分类号: TP 316.8

1 PC 集群的建立

计算机集群系统, 也称为计算机机群系统^[1], 简称集群, 是互相连接的多个独立计算机的集合。这些计算机可以是单机或多处理器系统(PC 或工作站), 每个结点都有自己的存储器, I/O 设备和操作系统。

集群的性能取决于下面几个因素: 节点本身; 节点之间互联设备; 底层通信软件(包括互联设备的驱动); 全局资源管理系统以及并行计算环境等等。

1.1 PC 集群节点的组成

经过努力, 成功地自行设计并组建了四台 PC 机组成的 PC 集群, 本集群节点主要配置如下:

CPU 主节点 P 400 MHz, 从节点 P 400 MHz;

内存 主节点 64 M, 从节点 64 M;

网卡 主节点 Intel (R) PRO/100 + PSI Adapter, 从节点 Intel(R) PRO/100 + PSI Adapter;

硬盘 主节点 4 G, 从节点 4 G;

显示卡 主节点 S3 Trio64V + PCI (765), 从节点 S3 Trio64V2-DX/ GX (775/785)。

1.2 底层通信设备与全局资源管理系统

本研究中底层通信设备采用 Super stack speed hub 500, 底层通信软件采用以太网(Ethernet)。

本集群采用 TURBO-LINUX 作为操作系统, 它实际是一种能在 PC 机上运行的 UNIX 的变体^[2],

是一个以 Intel 系列 CPU 为平台的、高性能的 32 位多任务、多用户的完全免费的 UNIX 兼容系统。

2 MPI 的实现

国际上大部分并行工作是在可移植网络并行环境 PVM 和可移植消息传递界面 MPI 并行环境下进行的。MPI 发展较快, 本研究采用了可移植消息传递界面 MPI 并行环境。

2.1 MPI 在 LINUX 系统上的实现

虽然免费下载的 MPICH 称其可在 LINUX 上实现, 但在具体实施中需对 LINUX 的一些配置文件进行修改。

2.1.1 建立各处理机之间的信任关系 在 MPI 程序运行之前, 必须通过以下步骤在各处理机之间建立信任关系。

第一步建立用户帐号, 即系统管理员以 root 登录, 建立用户帐号。

第二步是修改密码文件 passwd。当系统利用隐蔽口令系统时, 本字段被设置为“*”。由于“*”并不是密码字段的有效字符, 所以“*”阻止注册访问帐号。这个字段(在建立口令之后)永远不能进行手工编辑。笔者用文本编辑器进入 passwd 文件删除该字段。

第三步修改密码文件 shadow。一般系统都利用隐蔽口令系统, 这时系统中存在一个含有密码的单独文件, 名叫 shadow, 被存放在 /etc/ 子目录中。由于保证在各个处理机之间可以执行 rlogin 命令, 以及无身份验证的 rsh 命令是网络并行计算平台 MPI 可以运行的必要前提, 因此需要把要运行 MPI 程序的用户设置为无口令的用户, 即以 root 的身份把用户的口令删除。修改 shadow 与修改 passwd 的

收稿日期: 2001-03-13

基金项目: 国家自然科学基金资助项目(29806001)

第一作者: 女, 1976 年生, 硕士生

方法相类似。

第四步修改网间守护程序 `inetd.conf` 文件。如果上述删除操作已完成,但还无法执行 `rlogin` 命令或无须口令验证的 `rsh` 命令等 `r` 命令,则问题有可能出现在网间守护程序(Internet Deamon)上。因此,需要修改 `inetd.conf` 文件增加 `rlogin` 和 `rsh` 服务。

修改 `inetd.conf` 文件的执行步骤如下(在终端中以 `root` 身份执行):

```
# cd /etc
# mv inetd.conf inetd.conf.org
# cat inetd.conf.org | sed s/ # shell/ shell/ >
inetd.conf
# killall - HUP inetd
# cd /etc
# mv inetd.conf inetd.conf.org
# cat inetd.conf.org | sed s/ # login/ login/ >
inetd.conf
```

```
# killall - HUP inetd
```

第五步定义可信任主机与可信任用户,可信任主机又称为“等价主机”^[3]。系统假设对可信任主机具有访问权的用户对本地主机也应具有等价的访问权,在两个主机上具有相同名字的用户帐号应由同一个用户拥有。因此这种验证系统需要定义可信任主机和可信任用户的数据库,它们为

`/etc/host.equiv` 它为整个系统定义可信任主机与可信任用户。

`.rhosts` 它为单个帐号定义可信任主机与可信任用户。

第六步修改 MPI 程序的机器名列表。用文本编译器修改 `...../mpich/util/machines/machines.LINUX` 文件,该文件中要列出所有的处理机的名称。这个文件的格式为

```
hostname2 hostname3 ..... hostname1
```

其中 `hostname1` 为主节点名。

经过这些步骤之后,就可以对 MPI 的并行程序进行编译和运行了。

2.2 MPI 并程序设计流程图

图 1 给出了 MPI 并程序设计流程图。其中,在所有 MPI 函数中, `MPI_Init` 必须被首先调用,而且只能调用一次,其作用是初始化进程,使之进入 MPI 系统,并可进一步调用其它 MPI 函数。同时,进程结束前,必须调用函数 `MPI_Finalize` 来通知

MPI 系统。

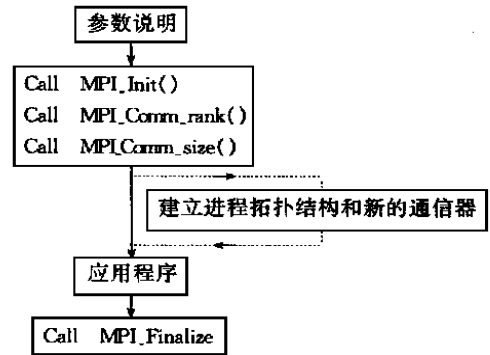


图 1 MPI 并程序设计流程图

Fig. 1 Design diagram of MPI parallel program

3 微机电网络并行计算实例

3.1 分子动力学程序的并行计算

计算机分子模拟^[4]即人们所说的计算机实验,是现代统计力学和分子热力学中一个重要的研究方法。它可分为 Monte Carlo (MC) 分子模拟及分子动力学(Molecular Dynamics, MD)模拟。前者主要用于通过模拟来得到体系的平衡性质,如密度、压力、体积、热力学能、自由能及反映体系内部结构的径向分布函数等;而后者能获得平衡性质的同时还可得到体系的传递性质(或运输性质),如动量、表面自由能、粘度及传递性质等。

MD 方法的特点确定了并行化计算比 MC 方法更为容易。MD 是通过解牛顿运动方程组求得各粒子在不同时刻下的位置、速度及其方向等微观信息,应用时间平均的方法得到宏观性质,因其计算的对象为粒子的作用力及位置,而这些计算是相互独立的。故可将粒子平均分配到并行计算机不同的节点上运算而同时完成体系中所有的粒子的计算。计算方法的并行度较好,并行效率较高。图 2 所示为 MD 模拟并行计算方法的一般性框图。

3.2 并行计算数据及分析

为了验证并行编程的可靠性,本文计算了对比密度($\rho^* = 0.9$)不同对比温度下体系的热力学能值和压力值,其结果如图 3 和图 4 所示。从图 3,4 可以看到本文的计算值与文献值^[5]吻合良好,表明模拟程序是正确可靠的。

用 Lennard-Jones 模型进行计算,实验结果如下。

规模 500 个分子点。单台处理机 运算时间为 18 604.59 s;两台处理机 运行时间为 10 891.66 s。

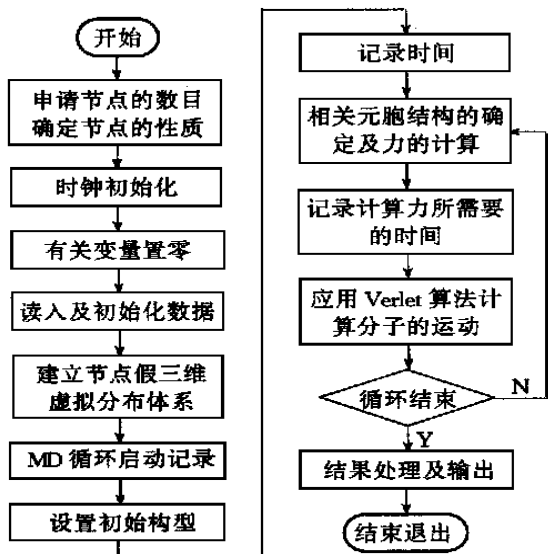


图 2 MD 模拟并行计算方法

Fig. 2 Diagram of MD simulation parallel computing

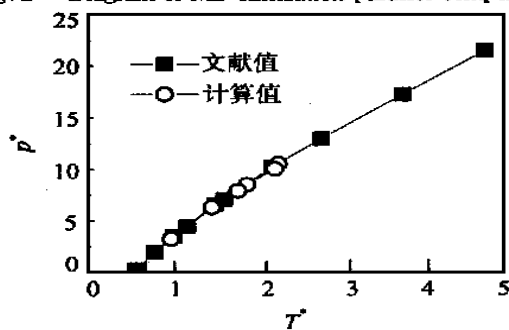


图 3 并行 MD 模拟压力值与文献值比较

Fig. 3 Comparison parallel MD simulated pressure values with literature values

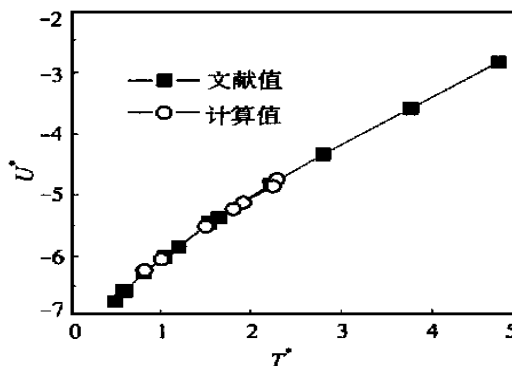


图 4 并行 MD 模拟热力学能值与文献值比较

Fig. 4 Relationship between reduced thermodynamic energy and reduced temperature

加速比为 1.708。并行效率为 0.854；四台处理机运行时间为 7 086.192 s。加速比为 2.625。并行效率为 0.656 3。

参 考 文 献

- [1] 钟 玲,孙强南. 机群:并行计算机的新类型. 北京:电子科技导报,1999
- [2] 杨文志. Linux 最佳实用手册. 北京:人民邮电出版社,1999
- [3] Hunt C. TCP/IP 网络管理. 翟炯,石祥生,石秋云译. 北京:电子工业出版社,1999
- [4] Frenkel D, Smit B. Understanding molecular simulation. Amsterdam:Academic Press,1996
- [5] Surridge M, Tildosley D L, Kong Y C, et al. A parallel dynamics simulation code for dialkyl cationics surfactants. Parallel Computing,1997,22:1053~1071

Cluster of PC and its development environment LINUX OS and MPI and a application

WANG Ying QU Yi-xin

(College of Chemical Engineering, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: The cluster of PC and its development environment LINUX OS and its platform MPI were introduced and how to install MPI on LINUX OS was the main feature in this paper. A program was given as examples of the parallel computation application.

Key words: parallel computation; cluster of PC; MPI parallel computation development environment; molecular dynamics