



【应用技术】

中尺度大气模式 MM5 在传统并行和网格环境上的对比

范凤岐, 熊聪聪, 刘 娜

(天津科技大学计算机科学与信息工程学院, 天津 300222)

摘要: 为了将中尺度大气模式 MM5 应用到并行和网格计算领域, 对中尺度大气模式 MM5 的原理进行研究, 搭建了并行和网格环境, 给出了需要的组件和配置流程, 对网格和并行计算进行了比较. 结果表明, CPU 个数不再是提升 MM5 运行效率的瓶颈; 在网络环境较差的条件下, gridftp 比 ftp 更安全和稳定.

关键词: 网格; Globus Toolkit; MM5

中图分类号: TP393.09 **文献标识码:** A **文章编号:** 1672-6510 (2007) 02-0084-03

Comparison of MM5 Based on Parallel Model and Grid System

FAN Feng-qi, XIONG Cong-cong, LIU Na

(College of Computer Science and Information Engineering, Tianjin University of Science & Technology,
Tianjin 300222, China)

Abstract: In order to implement MM5 module in traditional parallel computing and grid system, MM5 was studied. Operational environment of MM5 and parallel computing were constructed. Essential condition and configuration procedure were given. Parallel computing was compared with MM5 by experiments. Results show that number of CPU is no longer the bottleneck of running MM5. Gridftp is more safe and stable than ftp under bad network computing environment.

Keywords: grid computing; Globus Toolkit; mesoscale model version 5

数值天气预报具有其他预报方法不可替代的地位和作用^[1]. MM5(Mesoscale Model Version 5)是目前气象领域应用最广泛的中尺度数值天气预报模式之一, 它是由美国宾州大学和国家大气研究中心研制, 该模式具备了描述较小空间尺度而变化强烈的天气系统能力^[2]. MM5 有两种运行模式^[1], 一种是串行模式, 另一种是并行模式. 串行模式, 主要包括 6 个模块, 其详细内容可以阅读参考文献[3]. MM5 一直在扩展其在诸如 IBM SP, Cray T3E, Fujitsu VPP, 由 PC 和 workstation 组成的 Beowulf 集群^[4]以及多处理器的分布式集群并行计算机上的模块^[5].

网格计算是一种比较经济的运算方法. 据调查^[6], 1997 年北京市政府投资 800 万元用于购置 SGI 分布存储并行处理机, 1999 年上海市气象局准备购置银河 III 并行计算机, 分别解决当地中尺度预报问题. 这样的预算对于大部分地区都是无法接受的, 网格是解决该问题的一种经济的解决方案.

既然 MM5 是一种应用广泛的计算模式, 而网格计算可以提供一种更加廉价的要求. 能否将 MM5 模式应用到网格领域是一个需要研究的课题. 本文将 MM5 分别在传统并行环境下和在网格环境下进行了实验, 探究了两者各自的优缺点, 希望对今后这方面的研究起到一定的借鉴作用.

1 网格计算技术

网格计算是利用网格把分散的资源组织成一个巨大的“虚拟超级计算机”, 以实现计算资源、存储资源、数据资源、信息资源、知识资源、专家资源的全面共享. 这样的“虚拟超级计算机”有两大优势: 超强的数据处理能力和充分利用网上闲置的处理能力. 在我国国内网格计算目前大多只停留在理论分析阶段, 能否将网格计算的种种优势应用于 MM5 是一个新的课题.

收稿日期: 2006-07-03; 修回日期: 2007-03-16

基金项目: 天津市高等学校科技发展基金资助项目 (20051507)

作者简介: 范凤岐 (1982—), 男, 天津人, 硕士研究生.

2 搭建并行环境与网格环境

2.1 MM5 并行模式的实验

2.1.1 需要的组件

通过试验,确定组件的信息:

节点的操作系统采用RedHat9和Solaris9, Fortran编译器使用Portland公司的PGI5.1.用于并行运算的MPI使用MPICH-1.2.7,绘图软件采用NCARG-4.4.0.

需要注意的是RedHat Advanced Server 4(AS4)不能够很好的执行MM5的运算,主要是因为PGI5.1版本对AS4的支持不好,无法完全安装,所以采用了Redhat 9的版本. PGI主要是用在MM5进行数值计算时使用的编译器.试验表明,其最新的6.0版本可以通过TERRAIN单元,但是在REGRID单元的REGRIDDER模块时无法继续运算,而其5.1版就可以顺利通过,因此使用了5.1版.其他组件都无特殊要求.

2.1.2 配置流程

第一,安装PGI5.1.以root身份登陆,执行tar zxvf PGI5.tar.gz对安装文件解压,并运行./install,进入交互式安装,安装后须将PGI的安装目录加入环境变量;第二,下载配置好MM5各个模块^[2];第三,安装NCARG,这是一个MM5运算时需要用到的图形软件;之后是安装配置MPI,解压后运行./configure prefix=/usr/local/mpi with-common=shared rsh=ssh从而生成makefile文件,再依次运行make和make install命令进行安装.完成后要配置每一个节点的/usr/local/mpich/share目录下的machines.LINUX,加入参加每一个运算的节点名字.都配置好后进入到MM5模块,运行mpirun-np 再加节点个数和要执行的文件名,这样就可以进行运算了.

2.2 MM5 在网格上的实验

2.2.1 需要的组件

除了首先安装配置上面提到的各种组件之外,还要进行网格基础平台的选择.按照GT官方网站上的建议,实验中使用了Globus Toolkit 4.0.2(以下简称GT)版本.目前大多数网格项目都是采用基于Globus Toolkit所提供的协议及服务建设的.

2.2.2 系统拓扑图

网格由3台PC组成:两台机器的操作系统是Redhat 9,域名分别是grid3.grid.org和grid2.grid.org,都作为运算节点,一台机器的操作系统是Solaris 9,域名为grid4.grid.org,是一个主控节点.网络的拓扑结构如图1所示.

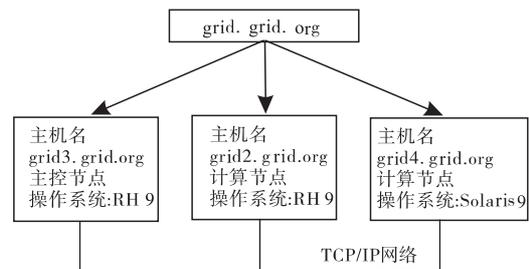


图1 系统拓扑结构

Fig.1 System topology structure

2.2.3 配置流程

第一,安装 j2sdk, ant, perl, postgres 等软件,它们是安装 GT 的前提.安装 GT,以 globus 用户登录,执行 tar zxvf gt-4.0.2.tar.gz,对安装文件解压,在 linux 下运行 ./configure-prefix=/usr/local/globus-4.0.2-with-iodbc=/usr/lib 来生成配置文件,需要注意的是,在 solaris 下需要在上面的语句后添加--with-flavor=gcc;之后再运行 make 命令生成安装文件,最后运行 make install 命令将安装文件安装到系统中,从而完成安装.

第二,安装 gridftp,分别通过两种方法进行文件的远程传输,最后进行比较.实验中分别对两者的传输时间,传输性能和安全性进行了测试.

第三,在网格上部署 MM5.首先要对用户的权限进行设置,使得该用户可以不用认证,就可以通过 SSH 的方式远程登录到集群上的每一个节点,其次编写 xml 格式的任务描述文件,用来描述提交任务的执行方法;最后调用 GT 的 GRAM 执行单元,完成在网格下部署 MM5 的过程.

3 实验结果与分析

3.1 单台PC与集群的对比

MM5 并行模式的使用同 MM5 串行运算相比速度得到了大幅度的提高.图 2 是 MM5 在单台 PC 和集群上运行时间的对比图.将 MM5 数值计算的程序分别

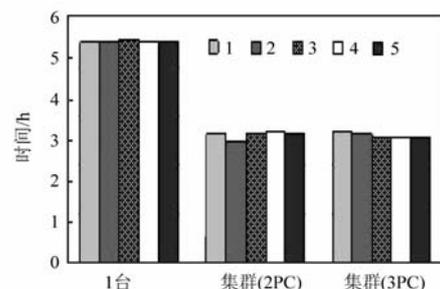


图2 MM5 在单台 PC 与集群上运行时间

Fig.2 Runtime of one computer and cluster

放在一台 PC,由两台 PC 组成的集群和由三台 PC 组成的集群上,为了减小误差影响,每一次测试都进行了 5 次. 实验结果说明单台 PC 的运算时间小于集群的运算时间,但是当集群个数由 2 变成 3 时,运算时间并没有大幅度减少.

3.2 gridftp和ftp的对比

对 GT 的 gridftp 和 ftp(采用二进制传输模式)进行了对比传输测试,传输三个大小不同的文件,测试其传输速率. 为了减小误差影响,每次实验作了 5 次. 结果如图 3 所示.

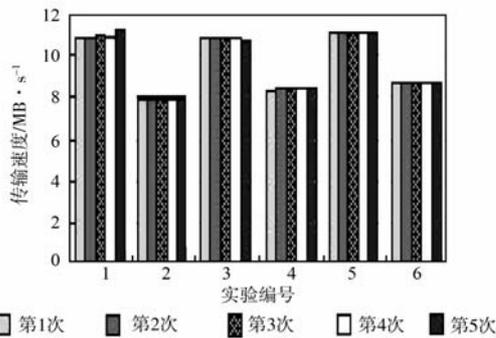


图 3 Ftp 与 gridftp 传输对比图
Fig.3 Transmission contrast between ftp and gridftp

横轴上的 1、2 分别表示文件大小为 182MB 时 ftp 和 gridftp 的传输速率;3、4 分别表示文件大小为 2.9GB 时 ftp 和 gridftp 的传输速率; 5、6 分别表示文件大小为 7.1GB 时 ftp 和 gridftp 的传输速率.

实验表明:(1)在传输速率上,gridftp 比传统的 ftp 要慢,这是因为 gridftp 是建立在 ftp 基础之上,另外它提供的可靠传输本身也消耗了一部分时间;(2)传输稳定性上,在中小文件上 gridftp 要比 ftp 更稳定.在多次传输同一文件的过程里,gridftp 的方差除了在传输 7.1GB 大小的文件时比 ftp 大外,其他情况下其方差比 ftp 都小,说明其抗干扰性要比 ftp 更好;(3)普通的 ftp 认证过程比较简单,而 gridftp 支持基于 X.509 的认证方式,更加安全.

3.3 并行和网格对比

对 MM5 在并行环境下和网格环境下针对 1998 年 07 月 31 日的天气进行预测所花费的时间进行了对比,结果如图 4 所示,为了减小偶然误差,实验进行了 5 次.

实验结果表明:在运行时间上,网格环境下并不比并行环境下慢,这是因为在运算量比较大的情况下,网格环境下的安全机制所消耗的资源可以忽略不计;在稳定性上,从图 4 中可以看到,网格运算每次花费的时间比较平均,而并行更容易受外界影响,其运

算时间的方差较网格运算更大;配置的过程上,网格运算的配置较并行计算更为繁琐,除了完成 MPI 本身的配置之外,还需要配置安全认证,用户认证,还要书写提交文件.

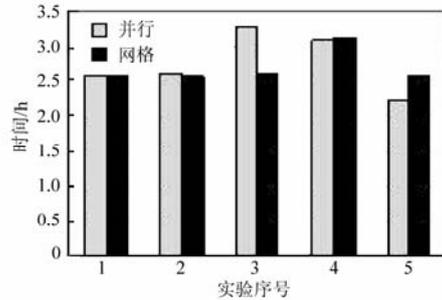


图 4 并行和网格运行时间对比图
Fig.4 Runtime contrast between parallel and grid

4 结 语

MM5 在实验环境中 PC 个数由 1 变 2 时,系统的性能有了较大的提高,但是由 2 变成 3 时,系统的效率没有很大的提高,这说明提升 MM5 运行效率的瓶颈已经不再是 CPU 个数,应该去查找其他影响运算速度的因素.

gridftp 与 ftp 的对比实验说明 gridftp 适用在网络环境较差的条件下使用,比 ftp 更安全、稳定. 网格适用于运算量较大,稳定性较高,并对于安全性有要求的应用.

参 考 文 献:

- [1] 朱小谦,张卫民,宋君强. 中尺度数值天气预报模式 MM5 分布式并行计算[J]. 国防科技大学学报, 2003, 25(2): 56—67.
- [2] Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research. MM5 modeling system overview[EB/OL]. (2003-04-02)[2006-06-21].<http://www.mmm.ucar.edu/mm5/overview.html>.
- [3] 戴进,余兴,岳治国. MM5 中尺度模式及其微物理过程[J]. 陕西气象, 2004(1): 11—14.
- [4] 许霞义,何旭丽,许红梅. 基于 Linux 的 PC 集群在地震资料处理中的应用与实现[J]. 计算机应用研究, 2004, 21(3): 154—204.
- [5] Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research. The distributed memory (DM) parallel option [EB/OL].(2006-06-02)[2006-06-21]. <http://www.mmm.ucar.edu/mm5/mpp.html>.
- [6] 任建军,张卫民. 网格门户下基于 GSI 的 MyProxy Login 实现[J]. 计算机工程与应用, 2004, 40(13): 36—39.