



网格技术在中尺度数值天气预报 MM5 模式中的应用

熊聪聪, 刘娜, 宋鹏

(天津科技大学计算机科学与信息工程学院, 天津 300222)

摘要: 以中尺度数值天气预报 MM5 在网格环境下运行为例, 分析了基于网格计算的应用系统在资源共享方面的应用. 介绍了中尺度数值预报模式的差分近似算法, 并以基于网格计算的中尺度数值天气预报 MM5 为例, 给出了在网格环境下的实验流程和结果. 实验表明, 使用 Globus 技术使得原本只能在集群或小型机上运行的科学计算程序能够在普通 PC 上运行起来, 并且随着加入网格机器的增多, 可以提高运行速度和吞吐量.

关键词: 中尺度数值天气预报; MM5; 网格计算

中图分类号: TP301.6 文献标识码: A 文章编号: 1672-6510 (2008) 01-0067-04

Application of Grids Technique to Mesoscale Numerical Weather Prediction MM5 Model

XIONG Cong-cong, LIU Na, SONG Peng

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

Abstract: As an example of the mesoscale numerical weather prediction MM5 in the grid environment operating, the application system based on grid computing in the resource sharing was analyzed. The approximate algorithm of mesoscale numerical weather prediction model was introduced. As an example of MM5 based on grid computing, the procedure and results in the grid environment were given. Experiment results show that Globus technology benefits the computer on ordinary computer, which originally only can be realized in clusters or minicomputer. And increasing of grid machines can enhance speed and throughput.

Keywords: mesoscale numerical weather prediction; MM5 model; grid computation

中尺度数值预报模式 MM5 是目前气象领域应用最广泛的中尺度预报模式之一^[1], 在我国已经建成的有限区域数值天气预报业务系统中, 北京气象局和天津气象局等均采用该模式作为业务模式. MM5 是基于天气运动变化的非线性变化偏微分方程组^[2], 处理大密度的数据、进行复杂的运算. 随着探测资料的不断丰富, 近年来社会发展对气象预报的要求也日益提高, 如 2008 年奥运会要求预报精度由 15 km 提高到 1 km, 有数据表明预报精度提高 1 倍, 计算量提高 16 倍^[3]. 提高预报精度的需求使得高性能计算环境开展区域数值天气预报的需求不断增长.

近年来正在兴起的网格计算技术在解决上述问题方面有着良好的应用前景^[4]. 本文研究了如何把传统的分布式科学计算移植到网格平台, 并以数值天气预报中尺度模型 MM5 为例, 详细阐述了系统的设计和实现, 并对系统性能作了评估.

1 中尺度数值预报模式的差分近似算法^[5]

MM5 中尺度数值模式控制方程组都是非线性的, 目前通常采用差分近似的方法, 把微分方程化为差分方程, 然后在网格上求数值解.

收稿日期: 2007-03-14; 修回日期: 2007-11-22

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: 熊聪聪 (1961—), 女, 四川泸州人, 教授.

现是使用 RSL 描述一个任务,即用户写一个标明资源(如计算机)、特定要求(如 CPU 数量、内存和执行时间等)和参数(如执行点、命令行参数、环境变量等)的 RSL 脚本. 基于在脚本中的信息,应用管理器调用由 Globus Toolkit 分布的联合分配库, DUROC 调度和启动由用户指定的多个计算机间的应用. DUROC 库使用 GRAM 的 API 和协议启动并管理一系列子计算. 对每一个子计算, DUROC 创建一个请求到远程的 GRAM 服务器,认证用户,执行本地授权并与本地调度机交互以初始化计算. DUROC 和相连的 MPICH-G2 库把大量的子计算绑定到单个 MPI 计算中.

(3) 如果连接成功, GRAM 将使用 GASS (Globus Access to Secondary Storage) 标明远端的节点可执行. 一旦应用启动, GASS 也用于把标准输出和错误流定向到用户终端并提供对文件的访问而不需要考虑其位置. 因此屏蔽了除性能之外的所有地理上的分布细节. 一旦应用启动后,应用管理器尽可能地在任意两个进程间选择最有效的计算方法.

(4) DUROC 和 GRAM 也彼此交互以监视和管理应用的执行. 每一个 GRAM 服务器监视它的子计算的生命期,当它经历挂起、运行和终止时, GRAM 把每一状态的转折都返回给 DUROC. DUROC 控制整个计算过程. 用户也可以在任何时间发出中止计算的请求,在 DUROC 和 GRAM 服务器

上,通过 GRAM 进程控制信息进行通信,中止所有进程.

(5) 在进程开始运行后,用管理器和 RSL 脚本指明的信息,创建基于底层网络拓扑的多层进程簇.

(6) 由于 VO 中的资源可以发出动态变化,网络也可出现拥塞或者中断连接,因此, NWS 负责搜集网络的负载情况,把信息通知 MDS. 同时, MDS 也定时查询 VO 中的资源使用情况并发送给应用管理器.

(7) 应用管理器在获得计算结果后,发送给 Web server 并存储在缓冲区中,考虑到用户的动态性,可使用推(push)或拉(pull)的方式方便用户取出数据.

2.3 算法模拟

把原本的数值天气预报计算平台移植到网格平台后,科学计算程序本身并没有变化,但是由于程序调用时增加了 XML parse 等过程,在一定程度上会影响效率,为了测试网格平台在调用时的延时,进行了两轮测试,第一轮主要测试服务创建时间,每次调用时创建一个服务实例并且销毁它. 输出结果第二轮测试随客户端的增多,每秒服务的次数相应增加.

2.4 数值实验结果

通过采用 SGE 管理集群,对数值天气预报 MM5 模式进行测试,分别记录了参与计算的多台计算机的运行时间,结果如表 1 所示.

表 1 多台计算机参与天气预报计算时间记录

Tab. 1 The records of multi computers participating in the weather forecast's calculation time

s

机器台数	实验序号										平均值
	1	2	3	4	5	6	7	8	9	10	
1	19 440	19 620	19 560	19 500	19 680	19 440	19 620	19 500	19 560	19 380	19 530
2	11 460	11 580	11 700	11 580	11 640	11 520	11 640	11 760	11 520	11 580	11 598
4	9 360	9 367	9 343	9 354	9 348	9 347	9 345	9 344	9 345	9 354	9 350.7

通过使用 LINPACK 中的 HPL 工具包对系统的并行性能进行测试. LINPACK 是一个用 Fortran 语言编写的线性代数软件包,主要用于求解线性方程和线性最小平方问题,测试结果如表 2 所示. 表中的时间为系统完成运算所需的时间.

表 2 系统并行性能测试

Tab. 2 The testing of system parallelism

系统平台	时间/s	浮点运算速度/次·s ⁻¹
天气预报网格—CPU 个数 1	9.79	1.790 × 10 ²
天气预报网格—CPU 个数 2	0.58	3.026 × 10 ³
天气预报网格—CPU 个数 4	0.09	1.999 × 10 ⁴
气象台并行机—CPU 个数 4	0.05	3.504 × 10 ⁴
南开之星—CPU 个数 4	0.03	5.841 × 10 ⁴

2.5 结果分析及性能评价

从表 2 中可看到,随着加入运算的 CPU 个数的增多,系统的性能得到极大提升. 并且通过对测试结果的分析发现,造成并行效率较低的原因及其解决方法有:

(1) 由于采用的是一个固定规模的计算量很小的算例,根据 Amdahl 定律,固定负载加速比为

$$S(p) = \frac{T(1)}{T(p)} = \frac{\omega}{\alpha\omega + (1-\alpha)(\omega/p)} = \frac{p}{1 + (p-1)\alpha} \xrightarrow{p \rightarrow \infty} \frac{1}{\alpha}$$

式中: ω 为一个问题的固定负载; $\alpha\omega$ 表示不可以并

行的串行工作量; $(1-\alpha)\omega$ 为并行工作量; p 为节点机个数, 不考虑其他所有开销. 该式说明加速比既与处理机数目 p 有关, 也与串行工作量所占比例 α 有关, 随着顺序瓶颈 α 和平均开销的增加而急剧下降. 该问题不能单纯用增加处理器数目来解决.

由于加速比不高的原因在于没有足够的工作量提供给多个处理机, 若消除固定规模的限制, 增加并行工作负载, 降低串行部分所占比例, 即可达到较高加速比. 假设将工作负载扩大为

$$\omega' = \alpha\omega + (1-\alpha)p\omega$$

则相应的顺序瓶颈为

$$\alpha' = \frac{\alpha}{\alpha + (1-\alpha)p}$$

加速比为

$$S'(p) = \frac{T(l)}{T(p)} = \frac{\alpha(1-\alpha)p}{\alpha + \frac{1-\alpha}{p}} = \alpha + (1-\alpha)p$$

这就是 Gustafson 定律——当机器规模增大时, 使计算规模也扩大, 可获得加速比的改善.

(2) 气象预报中的计算流体动力学问题一般用有限差分法和有限元法进行求解, 把整个问题域化成网格, 将网格分解成子域, 并把这些子域分配到相应的节点, 各节点机对分配到本节点上的一个或多个子域进行有关计算, 相邻子域间边界上的值必须从相邻节点中取得该边界点在另一边的函数值. 因此, 在每步计算之前, 相邻的处理器间必须相互交换消息, 即通信.

由分析, 加速比为

$$S(p) = \frac{pT_{\text{cak}}}{T_{\text{cak}} + T_{\text{comm}}} = \frac{p}{1 + f_c}$$

并行效率为

$$E(p) = \frac{1}{1 + T_{\text{comm}}/T_{\text{cak}}} = \frac{1}{1 + f_c}$$

式中, $f_c = T_{\text{comm}}/T_{\text{cak}}$ 为通信计算比, f_c 越小就可得到越高的加速比和效率, 随着节点数的增加, 分配到每个节点上的子域越小, 即计算量越小, 相邻节点间的通信量也减少, 而通信时间基本不变 (因为短消息的发送时间主要取决于以太网的通信启动时间), 从而 f_c 变大, 造成并行效率的低下.

在不改变算法和区域划分的情况下, 该问题主要通过提高通信速度的方法来解决. 气象预报的计算特性决定了节点机间需进行频繁的通信, 但是以太网的

通信速度大大低于共享内存系统中的访存速度. 以太网有相当大的通信启动时间, 且随着节点数的增加, 资源竞争冲突增大, 导致通信速度很低. 采用其他的高速通信网络能进一步改善并行效率.

3 结 语

把原本在集群上运行的数值天气预报中尺度模型 MM5 移植到网格平台. 按照 GT 规范中规定的方式, 把 MM5 要计算的工作以.xml 的形式提交到网格上, 并用 Web Service 的方式发布, 这样网格中的节点就可以调用这个服务, 从而也就实现了分布式的调用.

使用 Globus 等技术, 提高了系统的灵活性和安全性, 增强了系统的可用性, 使得原本只能在集群或小型机上运行的科学计算程序能够在普通 PC 上运行起来, 并且随着加入网格机器的增多, 可以提高运行速度和吞吐量. 在增加网格服务后虽然肯定会对科学计算的性能有影响, 但是服务的创建时间属于毫秒级, 相对于整个科学计算过程几乎可以忽略. 系统在大负载下 (大于 100 并发用户) 对服务响应比较及时, 可以用于生产环境.

由于网格技术作为新兴技术, 还存在许多技术上的难点. 在本实验中, 随着加入运算的 CPU 个数的提升, 系统的性能得到极大提升. 但是, 随着参与计算的 CPU 数量的增加, 性能不会成线性的增长, 并行效率不高. 如何提高并行效率, 将成为今后研究的重点.

参 考 文 献:

- [1] Grell G A, Dudhia J, Stauffer D R. A description of the fifth-generation penn state/NCAR mesoscale model (MM5) [R]. National Center for Atmospheric Research, Tech Rep: NCAR/TN-398+STR, 1994.
- [2] 郑良杰. 中尺度天气系统诊断分析和数值模拟 [M]. 北京: 气象出版社, 1982.
- [3] 吕绍华, 龚 斌. 基于 UNICORE 数值天气预报系统的设计与实现 [D]. 济南: 山东大学, 2006.
- [4] 都志辉, 陈 渝, 刘 鹏. 网格计算 [M]. 北京: 清华大学出版社, 2002.
- [5] 解以杨, 庞 璐, 孟冬梅. 数值天气预报与并行计算 [J]. 天津师范大学学报, 2001, 21 (4): 67—68.
- [6] 翁 鸣, 梁俊斌, 苏德富. 基于 OGSA 的网格虚拟组织建立 [J]. 计算机工程与科学, 2006, 28 (2): 140—142.