

广州区域数值预报模式并行化计算*

袁金南¹⁾ 王在志¹⁾ 薛纪善²⁾

¹⁾(中国气象局广州热带海洋气象研究所,广州 510080)

²⁾(中国气象科学研究院,北京 100081)

摘 要

采用消息传递方式(MPI)对最近发展的广州区域数值预报模式进行了并行化计算研究。根据模式的结构和计算过程特点,模式适合采用水平分区方案进行并行计算处理。在曙光3000并行计算机上分别采用一维和二维分区并行方案实现了模式的并行化计算,并对模式的并行效率、并行加速比和并行通讯时间百分比等做了测试。对测试结果的分析表明:采用8个CPU时,两种方案都能在1h内完成72h的预报,一维分区方案的并行效率则保持在90%左右,可以满足业务运行需要。当模式使用8个以上CPU时,通讯时间迅速增加并超过了计算时间的50%,模式并行效率明显下降。CPU相同时,模式一维分区并行方案比二维分区并行方案并行效率高且实现起来简单。

关键词:广州区域数值预报模式 曙光3000并行计算机 并行计算

引 言

随着计算机和计算方法的飞速发展,大部分学科通过采用定量化的方法变得更加精确,计算机成为科学研究等不可缺少的工具。与此同时人们对计算机的性能要求越来越高,而任何高性能计算和超级计算都离不开使用并行技术^[1]。

社会经济的发展对气象预报的要求越来越高,公众需要定点、定时、定量的气象服务,高分辨率数值预报模式是提供这种需求有效的方法之一。由于高分辨率数值预报模式计算量非常大,同时要求预报的时效性,对计算机性能要求非常高,在并行计算机上采用并行计算技术可以有效地解决数值预报模式计算量大、时效性强这个问题。

广州区域数值预报模式是广州热带海洋气象研究所自主开发的国内最早用于业务预报的有限区域原始方程模式^[2],也是中国台风专家组准许发布台风预报的数值模式之一。早期的模式分辨率较低,物理过程也不完善。为了满足业务需要,模式在空间分辨率和物理过程等方面进行了升级,升级后模式垂直方向分为10层,水平分辨率为111.1 km($\sim 1^\circ$),时间积分采用省时显式积分方案(EES)^[3],对热带气旋中心位置的预报有了明显的改进^[4]。为进一步提高模式的业务预报能力,特别是提高气旋路径及强度的预报,最近又提高了模式的空间分辨率,使模式垂直方向达到22层,水平分辨率达到55.6 km(\sim

* 本文得到“国家智能计算机研究开发中心和国家高性能计算基金”课题和广东省科技攻关项目“珠江三角洲热带气旋暴雨预警信号发布关键技术研究”资助。

2003-01-27收到,2004-03-17收到修改稿。

0.5°);同时模式预报区域也要扩大,以适应预报时间由原来的 48 h 延长到 72 h,减少模式边界的影响。这使得模式的计算量迅速增加,串行模式已经不能满足业务的时效性要求,因此模式必须采取并行化计算。计算机条件的改善也为并行化的实现提供了基础。本文针对广州区域数值预报模式的并行化计算进行了一些研究和探讨。

1 模式基本介绍

1.1 模式基本方程

模式采用 σ 坐标,水平方向采用 Mercator 投影,模式的水平动量方程为

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{\pi u}{m} \right) + m \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} u \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} u \right) \right] + \frac{\partial}{\partial \sigma} \left(\frac{\pi \sigma}{m} u \right) + \\ \pi \frac{\partial \phi'}{\partial x} - \sigma \frac{\partial \phi'}{\partial \sigma} \frac{\partial \pi}{\partial x} - f^* \frac{\pi v}{m} = \frac{\pi F_u}{m} \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{\pi v}{m} \right) + m \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} v \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} v \right) \right] + \frac{\partial}{\partial \sigma} \left(\frac{\pi \sigma}{m} v \right) + \\ \pi \frac{\partial \phi'}{\partial y} - \sigma \frac{\partial \phi'}{\partial \sigma} \frac{\partial \pi}{\partial y} + f^* \frac{\pi u}{m} = \frac{\pi F_v}{m} \end{aligned} \quad (2)$$

模式热力学方程为

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{\pi T'}{m} \right) + m \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} T' \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} T' \right) \right] + \frac{\partial}{\partial \sigma} \left(\frac{\pi \sigma}{m} T' \right) + \\ \left(- \frac{C_0^2 \pi}{R p} - \frac{\pi \kappa T'}{p} \right) \left(\frac{\pi \sigma}{m} + \frac{\sigma \pi}{m} \right) = \frac{\pi}{m} F_{T'} + \frac{\pi Q}{m} \end{aligned} \quad (3)$$

模式连续方程为

$$\frac{\partial \pi}{\partial t} + \int_0^1 m^2 \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} \right) \right] d\sigma = 0 \quad (4)$$

模式状态方程为

$$\sigma \frac{\partial \phi'}{\partial \sigma} = - \frac{R \pi \sigma}{p} T' \quad (5)$$

其中: $\pi \sigma = - \int_0^\sigma m^2 \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} \right) \right] d\sigma + \int_0^1 m^2 \left[\frac{\partial}{\partial x} \left(\frac{\pi u}{m} \right) + \frac{\partial}{\partial y} \left(\frac{\pi v}{m} \right) \right] d\sigma$

1.2 时间积分方案

模式起步采用 Euler 后差积分,即

$$f^{*n+1} = f^n + \Delta t \cdot \left(\frac{\partial f}{\partial t} \right)^n \quad (6)$$

$$f^{n+1} = f^n + \Delta t \cdot \left(\frac{\partial f}{\partial t} \right)^{*n+1} \quad (7)$$

这里 f^{*n+1} 是采用向前差分求得 $(n+1)$ 时刻的第一近似值, f^{n+1} 是采用向后差分求得 $(n+1)$ 时刻的值, $\left(\frac{\partial f}{\partial t} \right)^{*n+1}$ 是指由第一近似值 f^{*n+1} 所计算的倾向值。

模式大部分时间积分采用省时显式积分方案(EES)。设定两个时间步长 Δt_a 和 Δt_b , $\Delta t_a = m \cdot \Delta t_b$, m 取 4 或 6,时间积分按 Δt_b 进行,但所有的非线性项在 $2\Delta t_a$ 时段内保持

不变。以 τ 表示短步长的时间变量, 向量 $x = (u, v, t, \pi, A)$ 表示预报量值, 则有

$$\frac{\partial x}{\partial \tau} = N(x^n) + L(x) \quad (8)$$

这里 N 表示非线性项, 而 L 表示线性项。长时步与短时步预报结果的联结关系是

$$x^n(0) = x^{n-1}, x^{n+1} = x^n(2m) \quad (9)$$

这里 $x^n(m)$ 表示短时步第 m 个时间层上的值。

短时步积分过程中, 速度和高度、温度定义在不同的时间层上, 并采用“时间跳点”方案。这时有

$$\begin{aligned} T'^n(1) &= T'^n(0) + \left[\frac{\delta T'^n(0)}{\delta \tau} \right] \cdot \Delta t_b \\ u'^n(2) &= u'^n(0) + \left[\frac{\delta u'^n(1)}{\delta \tau} \right] \cdot 2\Delta t_b \\ T'^n(3) &= T'^n(1) + \left[\frac{\delta T'^n(2)}{\delta \tau} \right] \cdot 2\Delta t_b \\ u'^n(4) &= u'^n(2) + \left[\frac{\delta u'^n(3)}{\delta \tau} \right] \cdot 2\Delta t_b \\ T'^n(2m) &= T'^n(2m-1) + \left[\frac{\delta T'^n(2m-1)}{\delta \tau} \right] \cdot \Delta t_b \\ u'^n(2m) &= u'^n(2m-2) + \left[\frac{\delta u'^n(2m-1)}{\delta \tau} \right] \cdot 2\Delta t_b \end{aligned} \quad (10)$$

其中

$$\begin{aligned} \frac{\delta T'^n(2k)}{\delta \tau} &= \left| \frac{m}{\pi} \left(\frac{\partial}{\partial t} \left(\frac{\pi T}{m} \right) \right) - \frac{T}{m} \frac{\partial \pi}{\partial t} \right|^n + \left(\frac{C_0^2}{R\sigma} + \frac{\kappa T_e^n}{\sigma} \right) \omega'(2k) \\ \frac{\delta u'^n(2k-1)}{\delta \tau} &= \left| \frac{m}{\pi} \left(\frac{\partial}{\partial t} \left(\frac{\pi u}{m} \right) \right) - \frac{u}{m} \frac{\partial \pi}{\partial t} \right|^n - m \frac{\partial \phi^n(2k-1)}{\partial x} \end{aligned} \quad (11)$$

预报时 $\left| \frac{m}{\pi} \left(\frac{\partial}{\partial t} \left(\frac{\pi u}{m} \right) \right) - \frac{u}{m} \frac{\partial \pi}{\partial t} \right|^n$ 和 $\left| \frac{m}{\pi} \left(\frac{\partial}{\partial t} \left(\frac{\pi T}{m} \right) \right) - \frac{T}{m} \frac{\partial \pi}{\partial t} \right|^n$ 项在整个长时步 $2\Delta t_a$ 内保持不变。

1.3 模式垂直网格和水平网格

模式采用 σ 坐标, 如图 1 所示, 实线代表整层, 虚线代表半层。模式水平方向采用半跳点的 Arakawa B 网格, 如图 2 所示, 图中 V 表示矢量格点如 u, v ; F 表示标量格点如 ϕ, T', π, σ 。

1.4 模式侧边界

模式侧边界处理采用开放边界方法和边界松弛方法。

开放边界方法, 即在边界上假设

$$\frac{\partial f}{\partial t} = -C_x \frac{\partial f}{\partial x} - C_y \frac{\partial f}{\partial y} \quad (12)$$

其中 C_x, C_y 根据内点的变化倾向来估计。

边界松弛方法即在模式边界采用如下处理

$$\left(\frac{\partial f}{\partial t} \right)_p = (1 - \alpha) \left(\frac{\partial f}{\partial t} \right)_c + \alpha \left(\frac{\partial f}{\partial t} \right)_f \quad (13)$$

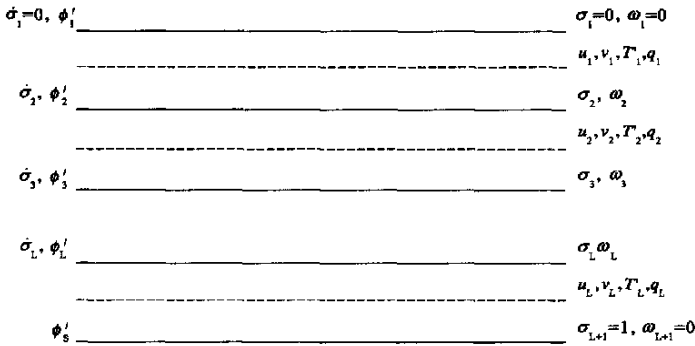


图 1 模式的垂直分层示意图

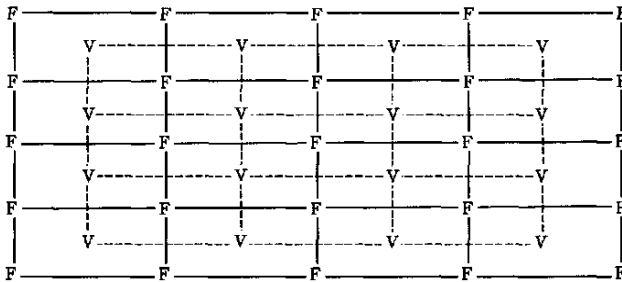


图 2 模式变量水平网格示意图

这里 $(\frac{\partial f}{\partial t})_p$ 是变量 f 的最终预报倾向值, $(\frac{\partial f}{\partial t})_c$ 表示更大尺度模式提供的预报倾向值, $(\frac{\partial f}{\partial t})_f$ 表示模式预报的倾向值, α 是边界平滑权重系数。

2 并行算法设计

模式的时间积分均为显式积分,即变量后一时步的求解只利用到前面时步的已知值,因此模式采用并行化计算从理论上可行^[5]。模式侧边界处理也采用的是显式方案,同样可以采用并行化计算,侧边界处理采用先进行数据传递后进行计算的方法。

模式在垂直各层上计算与水平方向上的时间积分不相关,因此模式可以在水平方向上采用分区并行计算,而在垂直方向上不适合进行分区并行。模式中最高偏导数为二阶偏导,采用格点差分求解时,某一内部格点上值的求解一般只用到周围 8 个格点上的值^[6],在分区时区域内部的分区边界至少要外扩一圈,模式中的标量和矢量采用两套不同的网格。图 3(a)和 3(b)分别为模式一维 Y 方向分区和二维 X-Y 方向分区并行方案。

如图 3(a)和 3(b)所示,双箭头表示每次积分时各分区边界要进行数据的相互交换情况。模式一维 Y 方向分区方案在并行方法上相对比较简单,数据交换时仅交换边界上的一行数据。这里一维并行设计采用自动定义各分区长度的方法,并行计算时可以任选一

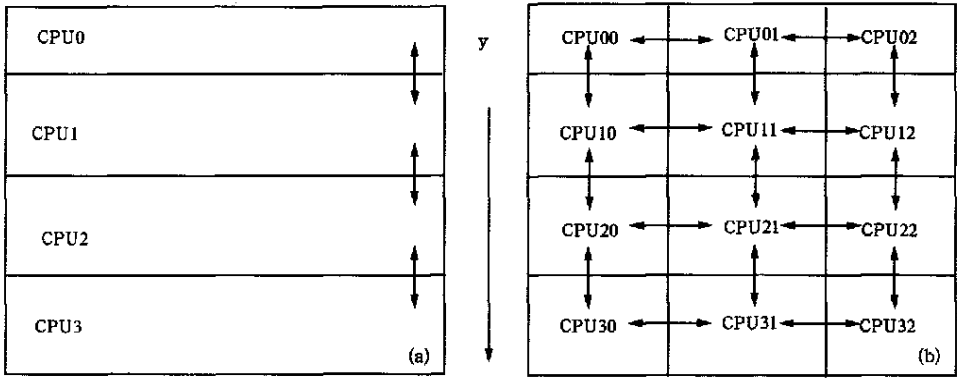


图3 模式一维 Y 方向分区并行方案(a)和二维 X-Y 方向分区并行方案(b)

个到多个 CPU 运行,但要求每个分区在 X 和 Y 方向均不少于 5 个格点。一般情况下模式在水平方向上的格点数是固定的,采用一维分区并行计算时各分区在 Y 方向上的格点长度不一定相同,有的分区可能会多一行格点。

二维 X-Y 方向分区方案跟一维 Y 方向分区方案基本相同,只是它在 X 和 Y 方向同时进行分区。采用二维分区并行计算时各分区在 X 和 Y 方向上的格点长度不一定相同,有的分区可能多一行或多一列格点。采用二维分区方案在数据交换时比一维分区方案复杂很多,其中有的分区要进行四个面的边界数据交换,因此模式实现二维并行比实现一维并行要复杂。另外,本模式采用二维分区并行方案时,通讯时间占总运算时间的百分比可能会随着分区数增加而明显增加。

模式采用一维 Y 方向分区和二维 X-Y 方向分区并行计算时模式积分和计算过程中没有采用隐式和迭代计算,模式并行计算只进行了分区处理,因此模式并行计算和串行计算的结果应该完全相同。这里把模式并行计算和串行计算的所有输出预报结果采用二进制文件进行比较,结果显示模式并行和串行计算输出的二进制文件完全相同,这表明模式并行和串行计算结果完全相同。这里模式并行采用 MPI(Message Passing Interface) 并行计算环境。

3 试验结果及并行效率分析

这里测试的并行模式格点数为:181 × 109 × 22,水平范围为:75.0° ~ 165.0°E 和 0.0° ~ 47.4°N,水平分辨率为 55.6 km(~ 0.5°)。模式积分步长为 120 s,积分时间为 72 h。这里所有测试采用同一个例。

试验采用广州曙光 3000 并行计算机,该并行计算机拥有 4 个节点,每个节点有 4 个 CPU 共享内存,每个 CPU 为 1.5 Gflops 2.0 G 内存,节点间通讯网为 1000 M/s。模式一维 Y 方向分区和二维 X-Y 方向分区的并行测试结果如表 1 所示。

这里,加速比 s_p 的定义为

$$s_p = \frac{T_l}{T_p} \tag{14}$$

表 1 广州曙光 3000 并行机上模式一维 Y 方向分区和二维 X-Y 方向分区的并行测试结果

CPU 个数	72 h 预报墙钟时间(s)	加速比	并行效率(%)	通讯时间百分比
1 × 1	22089 .01	1 .0000	100 .000	0 .0000
1 × 2	11308 .40	1 .9533	97 .665	9 .8296
1 × 4	5559 .17	3 .9734	99 .336	18 .8648
1 × 8	3083 .52	7 .1636	89 .545	44 .2653
1 × 12	2730 .09	8 .0909	67 .424	57 .7702
1 × 16	2573 .26	8 .5841	53 .650	68 .8048
2 × 2	6041 .40	3 .6563	91 .407	23 .2745
2 × 4	3562 .55	6 .2003	77 .504	52 .3181
3 × 4	3657 .85	6 .0388	50 .323	69 .0932
4 × 4	4011 .54	5 .5064	34 .415	78 .5606

注:CPU 个数 = X × Y 方向分区数,当某方向分区数为 1 时表示该方向没有进行分区,其中 CPU 个数 = 1 时为串行计算。

其中, T_1 是单机上求解一个问题所需要的时间; T_p 是在具有 p 个(与单机相同)处理机的并行系统上,并行求解同一问题所需要的时间。并行效率 E_p 的定义为

$$E_p = \frac{S_p}{p} \times 100\% \quad (15)$$

其中, s_p 为 p 个处理机的加速比; p 为处理机的个数。

通讯时间百分比的定义为

$$\text{通讯时间百分比} = \frac{\text{通讯时间}}{\text{总运算时间}} \times 100\% \quad (16)$$

这里,总运算时间由通讯时间和实际计算时间两部分组成,其中通讯时间通过并行计算中的 MPF-WTIME() 计时函数累计模式中数据通讯所用的时间得到。

表 1 给出了广州曙光 3000 并行计算机上广州区域数值预报模式两种并行化方案的测试结果。可以看出,模式并行化以后采用多个 CPU 并行计算比模式串行计算的计算时间明显缩短,模式在使用相同个数的 CPU 时,一维分区并行方案比二维分区并行方案的并行效率要高。另外,模式采用一维分区并行方案时使用 2~8 个 CPU 的并行效率比较高,而使用 12~16 个 CPU 时由于通讯时间百分比的增加,模式并行效率明显降低,其主要是由于使用 12~16 个 CPU 时通讯时间百分比超过了 50%,即通讯时间比实际计算时间要长,模式并行运算的时间主要花在了数据通讯上面。这里值得注意的是模式采用二维分区并行方案时,使用 12~16 个 CPU 时比使用 8 个 CPU 时的运算时间要长,而且使用 CPU 的个数越多运算时间越长,出现这种现象的直接原因是由于 CPU 个数的增加,通讯时间明显增加,导致了总运算时间的增加,这可能跟模式的二维分区并行技术方案等有关。从并行测试结果可以看到,模式并行时不一定是使用 CPU 个数越多越好,只有在通讯时间百分比比较小的情况下,模式的并行效果才比较好。

广州区域数值预报模式采用一维分区并行化方案相对比较简单,目前该并行方案已经在数值预报业务中得到应用。一般模式并行效率跟模式本身的特点和模式的并行技术方案等有关,因此如何进一步减少模式通讯时间百分比提高模式并行效率以及改进模式二维分区并行方案将有待于进一步探讨。

4 结 论

(1) 广州区域数值预报模式使用的是显式积分和计算,并且模式在水平方向上的计算和垂直方向上的计算相互独立,因此模式适合采用水平分区并行计算方法。模式并行化以后采用多个 CPU 并行计算比模式串行计算的计算时间明显缩短,模式采用水平方向一维分区并行方案相对比较简单,而采用二维分区并行方案比较复杂。

(2) 模式采用水平方向分区并行计算时在模式积分和计算过程中没有使用水平方向上的隐式或迭代求解,所以模式并行计算结果和串行计算结果完全相同。模式采用一维分区方案比二维分区方案并行效率要高,模式两种并行方案在使用 8 个以上 CPU 时由于通讯百分比明显增加,并行效率显著下降。一般模式并行效率跟模式本身的特点和模式的并行技术方案等有关,在模式并行时要提高模式并行效率,就要尽量减少通讯时间百分比。模式采用一维分区并行方案在使用 8 个 CPU 时并行效率保持在 90% 左右,可以满足业务需要。

致谢: 特别感谢国家智能计算机研究开发中心的石锦彩副研究员、乔香珍研究员等在技术上给予支持和帮助。

参 考 文 献

- 1 陈国良. 并行计算——结构·算法·编程. 北京: 高等教育出版社, 1999.
- 2 Xue Jishan, Wang Kangling, Wang Zhiming, et al. Test of a tropical limited area numerical prediction model including effect of real topography. *Advances in Atmospheric Sciences*, 1988, 5: 1 ~ 23.
- 3 薛纪善, 林俊杰, 王康玲. 高分辨率有限区域数值预报模式的省时显式积分. *大气科学*, 1995, 19(2): 165 ~ 175.
- 4 薛纪善, 王康玲, 何安国. 热带有限区域数值预报模式的升级及在台风预报中的应用. *热带气象学报*, 1996, 12(2): 97 ~ 104.
- 5 矫梅燕, 李国杰, 林洪. 原始方程模式多处理器的并行计算方案设计及其试验. *大气科学*, 1995, 19(5): 623 ~ 630.
- 6 张宝琳, 袁国兴, 刘兴平, 等. 偏微分方程并行有限差分方法. 北京: 科学出版社, 1994.

THE PARALLEL COMPUTATION OF GUANGZHOU AREA NUMERICAL PREDICTION MODEL

Yuan Jinnan¹⁾ Wang Zaizhi¹⁾ Xue Jishan²⁾

¹⁾ (*Guangzhou Institute of Tropical and Marine Meteorology, CMA, Guangzhou 510080*)

²⁾ (*Chinese Academy of Meteorological Sciences, Beijing 100081*)

Abstract

The parallel computation of the recently developed Guangzhou area numerical prediction model was researched by using Message Passing Interface (MPI) method. According to the characters of model structure and model computing process, the model is suited to parallel computation by using horizontal section method. The parallel computation of the model has been realized by using one dimensional and two dimensional section method, respectively, and parallel efficiency, parallel acceleration rate and percent of parallel communication time of the model have been tested on DAWN-3000 parallel computer. Results show that two parallel methods can finish 72-hour forecast in one hour by using 8 CPUs, the parallel efficiency of one dimensional section method keeps about 90%, which can meet the operational need. When more than 8 CPUs are used in the model, the communication time of the model increases rapidly and beyond 50% of the computation time, the parallel efficiency of the model decreases dramatically. One dimensional section parallel method of the model is simpler and more efficient by using the same CPU than two dimensional section parallel method of the model.

Key words: Guangzhou area numerical prediction model DAWN-3000 parallel computer
Parallel computation