

并行效率的初步研究*

喻 炜 颜 宏 金之雁

(国家气象中心,北京 100081)

提 要

用有限区域中尺度差分网格模式,采用分区域并行计算方法,在 TRANSPUTER 并行计算机上研究大规模并行计算中计算问题复杂度与并行计算效率的关系.结果表明,采用适当的并行算法,增加处理机个数和计算问题的复杂度可使并行加速比增大;并行效率则随着处理机个数的增加而降低,随着计算问题复杂度的增加而提高;大规模并行计算能明显地提高计算速度,很适合于解决大量计算的问题,在数值预报方面具有广阔的发展前景.

关键词:并行效率;计算速度;数值预报.

引 言

由于大气运动的复杂性,使得数值模拟的难度增加,计算量巨大.为了进一步提高数值预报的准确率,很好地模拟复杂的气候系统,必须首先解决计算机资源问题.

并行计算机的问世和发展为解决计算机资源的相对不足展示了新天地,并给气象模拟与预测技术的发展带来了一个极好的机遇.大气运动的物理特征对于并行化处理是极为有利的.大气的垂直结构、水平结构都有其相对独立性.但是,这些相对独立的可分性不是完全可分的.随着所研究的形势和天气尺度系统的时间尺度和空间尺度的改变,系统之间的耦合性(如中小尺度天气系统里的湍流和层流的共存和转化)变得十分重要^[1].而大气运动的耦合特征则是并行计算中信息传递(message passing)方法的物理根据之一.也就是说,大气科学问题的可分性和耦合性是大气科学并行计算的物理基础.

本文从中尺度模式物理过程复杂程度和模式预报区域分辨率高低两个方面来研究中尺度数值预报中计算量与并行效率的关系,并分析了并行算法对并行效率的影响,从而进一步研究大规模并行计算在中尺度数值天气预报中的应用及其潜力.

1 试验模式

本文的试验模式是金之雁等在颜宏等人在80年代研制的复杂地形条件下细网格嵌套模式(简称YH模式^[2])的基础上完成的,并在由TRANSPUTER处理机(CPU)构成的

并行计算机上运行.它是采用显示时间积分方案的有限区域中尺度差分网格并行模式(以下简称 YHP 模式).

该模式采用向 FORTRAN 语言程序(YH 模式)中植入消息通讯和数据并行功能的办法,在 EXPRESS 软件环境下采用分区域并行计算和 message passing 信息传递方法,使模式可以在并行机上运行.本文用 YHP 模式来研究大规模并行计算的并行效率及其在大气科学中的应用前景,为推出新的并行算法和优化并行模式作些基础性研究.

2 并行计算的算法分析

2.1 并行计算基本概念简述

为便于讨论,对文中涉及的一些有关并行计算基本概念作如下约定.

2.1.1 并行计算时间 把在给定预报时间和预报区域的前提下,模式在一个处理机上串行运算所需要的时间 A 定义为计算量(单位:s)(本文所作的试验均在单用户状态下完成的,因而计算时间用墙钟时间表示). A 由两部分组成:可并行处理部分的计算时间 A_p 和不能并行处理部分的计算时间 A_s .考虑理想情况下,各处理机的计算量近似平衡,这时各处理机的可并行部分、不可并行部分和系统开销可分别近似地看作是相同的(以下未作特别说明的,均满足该假设条件).模式预报所用机时 T 可表示为:

$$T = A_s + A_p/N + T_o \quad (1)$$

式中 T_o 为并行计算中单个处理机的系统开销, N 为处理机个数.

2.1.2 加速比 并行计算比传统的串行计算快,这种计算速度提高的幅度用加速比来定义.并行加速比(S_p)是模式在一个处理机上串行计算所用时间与模式并行计算所用时间之比.即

$$S_p = (A_s + A_p)/T \quad (2)$$

2.1.3 并行效率 并行效率定义为

$$P = S_p/N \quad (3)$$

由式(3)看到加速比越大并行效率越高,而且 $P \leq 1$.

在数值预报模式中 A_p 随积分时间急剧增加.而 A_s 部分随时间增加的量相对于 A_p 而言是很小的.在积分时间足够长的情况下,使 (A_s/A_p) 相对于 (T_o/A_p) 很小,可以忽略,这时并行效率可近似为

$$P = 1/[1 + T_o/(A_p/N)] \quad (4)$$

很明显,要提高并行效率的关键在于减少系统开销,增加可并行部分的计算时间.

2.2 计算量与内部信息交换量相对大小对并行效率的影响

前人的研究表明大气运动具有很好的可分性^[3],因而可根据式(4)成立的条件来讨论并行计算效率和并行加速比.这时 A_p 的大小近似为模式的计算量,则 A_p/N 可近似地表示处理机的内点计算量.当 $T_o/(A_p/N) \ll 1$ 时,对式(4)右端作泰勒级数展开,取其前3项:

$$P = 1 - T_o/(A_p/N) + [T_o/(A_p/N)]^2 \quad (5)$$

令 $x = T_o/(A_p/N)$, $y = P$, 则: $y = 1 - x + x^2$. 对 x 求导:

$$y' = -1 + 2x \quad (x \ll 1) \quad (6)$$

因为 $x \ll 1$, 所以 $y' < 0$, y 随 x 的增加而减小, 如图 1a 虚线所示. 图 1a 表明分区域内点计算量相对于数据交换量越大, x 越小, 并行效率越高. 当 $x = 0$ 时, 相当于一个 CPU 串行计算, 这时并行效率最高, 为 1.

当 $(A_p/N)/T_o \ll 1$ 时, 式(4)可写成:

$$P = [(A_p/N)/T_o] \cdot 1/[1 + (A_p/N)/T_o] \quad (7)$$

令 $x = (A_p/N)/T_o$, $Y = P$. 对 $1/[1 + (A_p/N)/T_o]$ 作泰勒级数展开, 取其前 3 项, 则式(7)可写成:

$$y = x - x^2 + x^3 \quad (8)$$

式(8)对 x 求导: $y' = 3x^2 + (1 - 2x) \quad (x \ll 1)$

因为 $x \ll 1$, 所以 $y' > 0$, 即 y 随 x 的增加而增大(图 1a, 实线). 换言之, 数据交换量相对于处理机的内点计算量越大, x 越小, 并行计算效率越低. 当 $x = 0$ 时, 相当于死锁现象(即程序不再继续执行下去, 处于停滞状态, 系统开销为无穷大), 这时并行效率为 0.

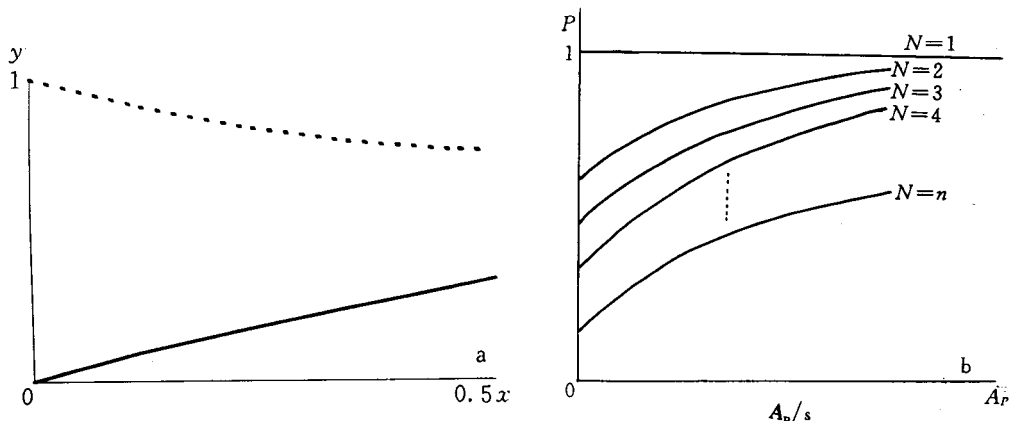


图 1 (a)子区域内点计算量(y)远远大于数据交换量时(虚线)和远远小于数据交换量时(实线)对并行效率的影响以及(b)并行效率(P)随计算量(A_p)大小的变化.

(横坐标 x : 数据交换量与子区域内点计算量的比值, 计算量单位: s , N : 处理机数目)

Fig. 1 (a) The relation of parallel efficiency (y) with computation intensity (x)

(dashed line: the computation for inner domain far more than the data exchange between subdomains; solid line: the computation for inner domain much less than the data exchange between subdomains)

(Where: x : ratio of inner-domain computation amount to data exchange amount)

(b) The relation of parallel efficiency (P) with computation amount (A_p)

(A_p : computation amount (unit: s) N : number of CPU)

由此可知, 在处理机总数 N 一定的情况下, 区域计算量远远大于处理机间的内部信息交换量时, 并行效率最高, 加速比最大. 对于计算量很小的简单问题, 或算法不好, 而使得 T_o/A_p 较大, 此时 N 的增加不仅可使并行效率降低, 而且有可能使并行加速比变小(甚至小于 1). 由此可看出, 并行计算很适合于解决巨量计算的问题.

2.3 大规模并行计算在实际应用中的优势

2.3.1 大规模并行计算在负载均衡的情况下, 并行效率又可表示为

$$P = A/(A_p \cdot N + T_o \cdot N + A_p) \quad (9)$$

令 $y = P$, $x = N$, $T_o = K \cdot T(x)$ (K 为比例系数, 仅随时间变化), 则式(9)可写成:

$$y = A/[A_s \cdot x + x \cdot T(x) + A_p] \quad (10)$$

式(10)对 x 求导:

$$y' = -A \cdot [A_s + T' \cdot x + T(x)]/[A_s \cdot x + x \cdot T(x) + A_p]^2 \quad (11)$$

在 $(0, \infty)$ 上, 如果 $T' < 0$, 则 $T(x)$ 随 x 的增加而减小, 即系统开销随处理机个数的增加而减少, 这显然不对, 因而有 $T' > 0$. 这时, $y' < 0$, 也就是说, y 随 x 的增加而减少. 这表明对于一定算法的大规模并行计算, 并行效率随处理机个数的增加而降低.

对于式(5), 假设 T_0 随 A_p 和 N 的变化很小, 这时可以得到 P 随 A_p 变化的一族曲线(图1b). 图1b表明在 $T_0/(A_p/N) \ll 1$ 的情况下, 并行效率随着处理机数目的增加而降低, 随着计算问题复杂度的增加而提高.

对于并行加速比, 又可表示为

$$S_p = A/(A_s + T_0 + A_p/N) \quad (12)$$

令 $y = S_p, x = N, T_0 = T(x)$, 则有

$$y = A/(A_s + T(x) + A_p/x) \quad (13)$$

式(13)对 x 求导:

$$y' = A \cdot (A_p/x^2 - T')/(A_s + T(x) + A_p/x)^2 \quad (14)$$

假定 $T_0(x) = C \cdot x$ (C 为比例系数, 与 x 无关), 则在 $(0, \infty)$ 上有

$$y' = A \cdot (A_p/x^2 - C)/(A_s + Cx + A_p/x)^2 \quad (15a)$$

式(15a)对 x 求导:

$$y'' = \frac{-A \cdot (2A_p \cdot A_s/x^3 + 6A_p \cdot C/x^2 - 2C^2)}{(A_s + C \cdot x + A_p/x)^3} \quad (15b)$$

当 $x \leq \sqrt{A_p/C}$ 时, $y'' < 0$, 即 y 在 $[0, \sqrt{A_p/C}]$ 上有极大值. 当 $x = \sqrt{A_p/C}$ 时, 取最大值.

当 $x < \sqrt{A_p/C}$ 时, $y' > 0$, 这时, y 随 x 的增加而增大, 即

$$N < \sqrt{A_p/C} \quad (16a)$$

式(16a)表示并行加速比随处理机个数的增加而增大.

当 $N = \sqrt{A_p/C}$ (16b)

成立时, 并行加速比最大. 由式(16b)可以看出, 在最佳并行状态下 (S_p 最大时), 计算量越大 (A_p 越大), 并行计算规模也越大. 换言之, 计算规模越大, 越能发挥大规模并行计算的威力.

2.3.2 考虑完善模式的物理过程所引起的计算量的差异在大规模并行计算中的变化特点, 则有:

$$T' = A'_s + A'_p/N + T'_0 \quad (17)$$

式中: $A'_s = A_s + \Delta A_s$ (18a)

$$A'_p = A_p + \Delta A_p \quad (18b)$$

$$T'_0 = T_0 + \Delta T_0 \quad (18c)$$

把式(18a)、(18b)、(18c)代入式(17)得:

$$T' = (A_s + A_p/N + T_0) + \Delta A_s + \Delta A_p/N + \Delta T_0 \quad (19)$$

式(19)表明在计算量急剧增加的情况下,应尽可能降低系统开销增加的幅度,即假定 $\Delta A_p/N \gg \Delta T$ 。(使处理机内点计算量的增加远远大于相应的系统开销的增加).考虑到 $\Delta A_p/N \gg \Delta A$,式(19)可以简化为

$$T' = (A_s + A_p/N + T_o) + \Delta A_p/N \quad (20)$$

由式(20)可以看出,两种方案引起的计算量的差异,在并行计算中只相当于原来的 $(1/N)$. 引进更符合实际的物理过程,对于串行运算来说,会引起计算量大幅度的增加,影响预报时效.而对于大规模并行计算,在 N 很大的情况下,这种计算量幅度的增加相对而言是很小的.因而采用大规模并行计算,在一定程度上,可以尽可能地完善模式的物理过程,不必担心由此引起的计算量的增加以及它对预报时效的影响.

3 预报区域等面积划分并行计算试验方案

根据试验模式所提供的多种参数化方案设计了两组试验方案:①物理过程复杂程度对并行效率的影响(表 1);②预报模式水平分辨率的提高对并行效率的影响(表 2). 模式预报时间为 12h,时间步长为 4min.

表 1 预报区域等面积划分并行计算试验方案 I

Table 1 Parallel computation test scheme I with area-even-divided subdomain

试验编号	物理过程	处理机	模式复杂程度	试验编号	物理过程	处理机	模式复杂程度
GJ011	干绝热,简单 PBL	1	较小	FJ031	非绝热,简单 PBL	1	较湿绝热大
GJ012	干绝热,简单 PBL	2		FJ032	非绝热,简单 PBL	2	
GJ013	干绝热,简单 PBL	3		FJ033	非绝热,简单 PBL	3	
GJ014	干绝热,简单 PBL	4		FJ034	非绝热,简单 PBL	4	
GJ015	干绝热,简单 PBL	5		FJ035	非绝热,简单 PBL	5	
SJ021	湿绝热,简单 PBL	1	较干绝热大	FH041	非绝热,复杂 PBL	1	较非绝热 简单 PBL 大
SJ022	湿绝热,简单 PBL	2		FH042	非绝热,复杂 PBL	2	
SJ023	湿绝热,简单 PBL	3		FH043	非绝热,复杂 PBL	3	
SJ024	湿绝热,简单 PBL	4		FH044	非绝热,复杂 PBL	4	
SJ025	湿绝热,简单 PBL	5		FH045	非绝热,复杂 PBL	5	

注:试验方案的分辨率均为 $33 \times 17 \times 6$;1 个处理机为串行,2 个以上为并行.

表 2 预报区域等面积划分并行计算试验方案 II

Table 2 Parallel computation test scheme II

试验	分辨率	处理机	计算时间(s)	加速比
SJ101	$21 \times 21 \times 6$	1	1441	
SJ102	$21 \times 21 \times 6$	3	626	2.30
SJ201	$27 \times 27 \times 6$	1	2424	
SJ202	$27 \times 27 \times 6$	3	980	2.47
SJ301	$33 \times 33 \times 6$	1	3743	
SJ302	$33 \times 33 \times 6$	3	1432	2.60

注:模式物理过程均为湿绝热,1 个处理机为串行,2 个为并行.

3.1 物理过程复杂度对并行计算的影响

3.1.1 计算时间比较 图 2 给出了不同物理过程并行计算的时间.干绝热模式是描述大气运动状态的最简单的中尺度模式.它不考虑水汽和非绝热加热的影响,物理过程简单,计算速度较快.复杂 PBL 的计算量最大,并行计算的时间也相应地较多.从计算速度来看,在计算量相同的情况下,随着处理机个数的增加计算时间是减少的,5 个处理机并行

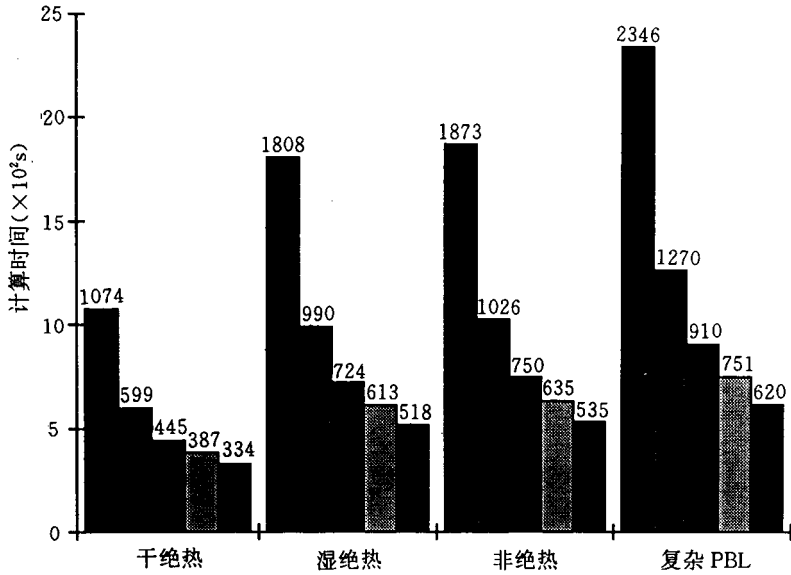


图 2 不同物理过程并行计算的时间(单位:s)

Fig. 2 The total computation time with different physical parameterization schemes (unit:s)

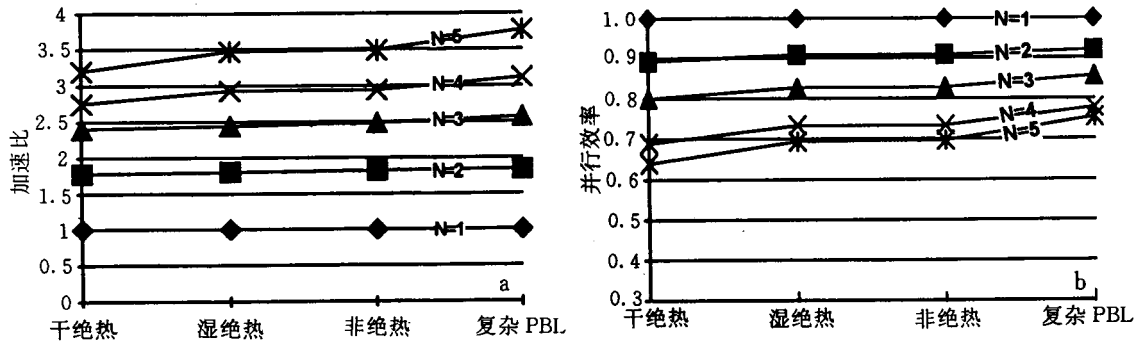


图 3 不同物理过程不同数量的处理机并行计算的(a)加速比(b)并行效率

Fig. 3 The speedup (a), and parallel efficiency (b) for different physical parameterization schemes (N, number of CPU)

计算的时间最短. 这表明, 处理机的个数越多, 并行计算的速度越快. 比较干绝热模式与复杂 PBL, 二者的计算量相差 1272s; 在采用并行计算后, 二者计算时间上的差距明显减少了, 5 个处理机并行计算时, 复杂 PBL 的计算时间仅比干绝热模式多 286s. 由此可见, 引进较为复杂的物理过程, 对于串行运算来说, 会引起计算时间大幅度的增加, 影响预报时效. 对于大规模并行计算, 随着处理机个数的增加, 这种计算时间增加的幅度相对而言是减小的, 这与式(20)的结论相一致. 因而采用大规模并行计算对于改进模式的物理过程, 提高计算速度是很有利的.

3.1.2 并行加速比 图 3a 给出了不同物理过程的并行加速比. 在物理过程相同的情况下, 并行加速比随处理机个数的增加而增大, 5 个处理机并行计算的加速比最大. 试验

GJ015 为 3.2, SJ025 为 3.47, FJ035 为 3.49, FH045 为 3.77. 结合式(16a)可以看到, 大规模并行计算能明显地提高计算速度.

在处理机个数相同的情况下, 随着计算量的增加, 并行加速比增大. 干绝热模式的加速比最小, 复杂 PBL 最大. 2 个处理机并行计算时, 干绝热模式的加速比约为 1.78, 复杂 PBL 的加速比约为 1.84, 增加了 0.06; 5 个处理机并行计算时, 干绝热模式的加速比约为 3.2, 复杂 PBL 的加速比约为 3.77, 增加了 0.57. 这与式(16b)的结果相同, 表明模式物理过程越复杂, 越能体现出大规模并行计算的优势.

3.1.3 并行效率 图 3b 给出了不同物理过程的并行效率, 与图 1b 中 P 随 A_p 的变化曲线一致. 在物理过程相同的情况下, 并行效率随处理机个数的增加而降低, 5 个处理机并行计算时的并行效率最低, 这与式(11)的分析一致. 在处理机个数相同的情况下, 随着计算量的增加, 并行效率提高. 干绝热模式的并行效率最低, 复杂 PBL 最高. 2 个处理机并行计算时, 干绝热模式的并行效率约为 0.89, 复杂 PBL 的并行效率约为 0.92, 提高了 0.03; 5 个处理机并行计算时, 干绝热模式的并行效率约为 0.64, 复杂 PBL 的并行效率约为 0.75, 提高了 0.11. 这说明模式物理过程越复杂, 并行计算的效果越好.

综上所述, 在并行计算中, 同一物理过程, 随着处理机个数的增加, 模式计算速度提高, 并行加速比增大, 并行效率降低; 不同物理过程, 计算量较大的并行加速比较大, 并行效率也较高.

3.2 模式水平分辨率提高对并行计算的影响

预报区域分辨率的提高会引起计算量的急剧增加, 主要表现在预报区域的格点数目有可能成倍增加; 或格距减小. 因此, 为了满足计算稳定性条件, 必须相应地缩短时间步长.

通过 3 种水平分辨率试验计算时间的比较(表 2)可以看出: 预报区域的格点数增加引起的计算量的差异是很明显的. 采用并行计算后, 计算速度明显提高了. 也就是说, 预报区域格点数增多引起的计算时间延长, 在并行计算中这种差距表现得较小. 对格点数为 21×21 的预报区域用 1 个处理机串行计算的时间与格点数为 33×33 的预报区域用 3 个处理机并行计算的时间大致相同. 可见, 并行计算对于数值预报的发展具有很大的潜力.

比较 3 个处理机并行计算时它们的加速比(表 2)可以看到, 随着预报区域计算量的增加, 加速比增大. 并行效率也随着预报区域计算量的增加而提高.

4 结 论

通过以上的分析和研究, 对于显示时间积分的差分网格模式得到以下结论:

(1) 并行计算能够明显地提高计算速度. 在满足 $N \leq \sqrt{A_p/C}$ 时, 处理机个数越多, 加速比越大, 计算速度越快. 一般而言, 随着处理机个数的增加, 系统开销增大, 并行效率降低.

(2) 模式物理过程越完善, 水平分辨率越高, 计算量越大. 在处理机个数一定的情况下, 随着计算量的增加, 并行加速比增大, 并行效率提高. 计算量越大, 越能体现出大规模并行计算的优势.

(3)在理想情况下,采用并行计算后,由于提高模式的水平分辨率、改进其物理过程而引起的计算时间延长只相当于计算量增加的数值的 $1/N$. 从预报时效的角度来考虑,大规模并行计算对于完善模式的物理过程、提高水平分辨率是很有利的.

(4)要充分发挥大规模并行计算的优势,要求分区域的内点计算量远远大于系统开销,因而大规模并行计算更适合于解决巨量计算的问题.

参 考 文 献

- 1 王宗皓. 大气科学问题并行计算方法. 数值天气预报新技术讲义. 1993, 83~93.
- 2 颜 宏. 复杂地形条件下嵌套细网格模式的设计. (一)数值模式的基本原理. 高原气象, 1987, 6(7): 1~62.
- 3 Simon F B Tett. Implimentation of Atmospheric Models on Large Multi-Processor Surfaces. Topics in Atmospheric and Oceanic Sciences. 1990, 32~46.

THE PRELIMINARY STUDY ON THE EFFICIENCY OF PARALLEL MODEL COMPUTATION

Yu Wei¹⁾ Yan Hong²⁾ Jin Zhiyan¹⁾

¹⁾(National Meteorological Center, Beijing 100081)

²⁾(China Meteorological Administration, Beijing 100081)

Abstract

The suitability and relationship between complexity and efficiency of multiple parallel computation have been intensively studied by using the parallelized version of the limited-area fine-mesh grid-point model.

The preliminary study indicates that, with suitable algorithm, the speedup can be improved by increasing the number of CPU (the number of transputers) in general. And, the efficiency of parallel may decrease with the CPUs, and increase with higher complexity. The results suggest that massively or highly parallel computer systems are most privileged for solving computationally intensive problem. In the light of its scalability the system which has the unlimited potential to provide much faster and more powerful computer resources required for further development of NWP and numerical modeling.

Key words: Parallel efficiency; Computational speed; Numerical weather prediction.