第7卷 第4期 应用气象学报 Vol. 7, No. 4
1996年11月 QUARTERLY JOURNAL OF APPLIED METEOROLOGY November, 1996

中期预报模式T63L16在银河-2巨型机上的并行计算*

杨学胜 宋君强
(国家气象中心,北京 100081) (国防科技大学,长沙 410073)

皇甫雪官
(国家气象中心,北京 100081)

提 要

多任务中期预报模式T63L16[1]主要利用了银河-2(以下简称为YH-2)巨型计算机提供的多任务并行计算环境,在原欧洲中期天气预报中心(ECMWF)T106L19模式版本的基础上,针对YH-2计算机的特点,实现了多任务并行计算,且获得了较高的加速比.

关键词:中期预报模式;多任务并行计算;加速比.

前 言

限制数值预报模式制作中期数值业务预报的因素之一就是计算机执行模式程序所需的时间. 中期数值预报系统包括资料预处理,四维同化,初值化,预报模式,后处理,场库的生成,传真图的制作和分发,格点报的生成等,所有这些子系统都要在规定的时段内完成,以供全国各级气象台站使用. 为尽早生成数值预报产品,国外许多气象机构作了大量的工作,其中之一就是在计算机上实现预报模式的多任务并行计算[2].

1 中期数值预报谱模式T63L16

目前国家气象中心运行的中期数值预报模式是ECMWF 80年代末的业务预报模式,代表了当今中期预报模式的先进水平. 该模式含近14万行FORTRAN源程序,398个子程序模块.

1.1 单任务T63谱模式的计算组织

T63模式采用半隐式时间积分,时间步长为22.5 min. 模式每积分一步,对每个文件都要从头到尾进行二次扫描处理. 在第二次扫描(子程序N2SC2)计算中,把对称和反

* 85-906-03-04课题资助.
1994-12-07收到,1995-08-31收到修改稿.

对称的傅里叶系数(以下简称为傅氏系数)写到工作文件中去;在第一次扫描(子程序NNSC1)计算中,从文件中读出这些傅氏系数,另外还读出其它的高斯格点值,以便在傅氏空间和格点空间进行计算,再把新的格点值写到工作文件中去,并利用傅氏空间和格点空间的增量去改变谱系数值. 在调用资料扫描子程序的过程中还完成了模式动力学的谱计算.

1.2 单任务下T63的输入输出

如果把所有预报模式的中间计算结果都保留在计算机内存里,则需要计算机内存8000兆字节. 因此必须把中间计算值储存在工作文件中,并利用输入输出方案,进行内存与工作文件之间的数据交换. 目前在模式中采用了两类工作文件:一类文件存放与模式动力学部分相联系的变量的傅氏系数(对称与反对称值);另一类文件存放模式物理过程和时间积分用到的格点值. 由于一组对称和反对称的傅氏系数能产生两个纬圈格点上的值(一个纬圈在北半球,另一个纬圈在南半球,且此两纬圈相对于赤道是对称的),根据纬圈剖面存取的方式,这两类文件的数据结构分别由96个纬圈格点记录和48个傅氏系数记录组成. 图1给出了两次扫描结构与模式文件的对应关系.

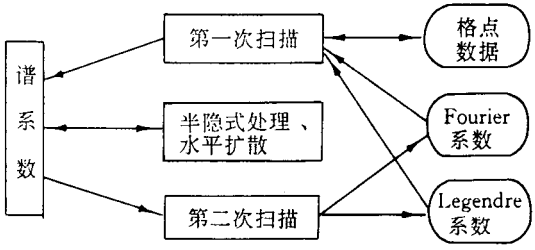

图1 模式的两次扫描结构

Fig. 1 The two scan structure of the model

为了便于重新启动计算及采用最快速磁盘输入输出方案,要求工作文件是有序文件,而不是随机或直接文件,每个工作文件都复制成两份. 数据从第一个文件读出,然后写到第二个工作文件中去;象这样全部扫描整个文件一次后,这两个文件的功能就相互交换,即输入文件变成输出文件,反之亦然.

2 YH-2并行计算环境

2.1 YH-2计算机并行环境

YH-2巨型计算机是紧耦合的MIMD多处理机系统,其主频为50兆赫,最多可装配4个中央处理机(CPU),内存为32兆字,字长64位,每个中央处理机有4K字的局部存储器,其输入输出系统由磁盘子系统支持,可连接32个磁盘.

YH-2提供了多种使用多CPU的技术,即作业级、子程序级和循环级的并行. 作业级的并行是指多个CPU同时执行不同的作业,由操作系统支持. 子程序级的并行是指多个CPU同时执行不同的子程序的调用,由宏任务库支持. 循环级的并行是多个CPU同时执行同一个循环的不同迭代,由微任务库支持. 宏任务和微任务库是目前YH-2提供的两种支持多任务并行计算的技术. 采用何种并行形式,应依赖于具体应用问题. 一般来说,只有在应用程序多任务并行计算比串行计算更快时,并行计算才具有实际意义. 多任务并行计算的加速比可表示为:

$$S_p = T_s/T_p$$

其中：T_s 为应用程序单 CPU 运行的时间，T_p 为并行计算运行的时间. 需要指出的是，在并行计算的情况下，多任务及操作系统组成的对并行计算的控制系统存在必要的额外开销. 因此，仅当可执行的 CPU 数目>1 时，S_p 的值才会超过 1.

通常多任务化一个应用程序最简单的方法是对 FORTRAN 源程序的 DO 循环多任务化，这种方法只需要分析该 DO 循环的数据相关性. 但这种方法得到的性能改进也仅限于该循环的范围. 多任务的最有效使用是在数学物理模型或程序的高层次上开发进行，这需要宏任务库的支持. 因此下面仅介绍 T63 多任务化时用到的 YH-2 宏任务库.

2.2 YH-2 宏任务库

应用宏任务库实现多任务化是指通过调用宏任务库的子程序而实现并行计算的技术. 宏任务库的子程序提供了将同一程序的多个子程序调度到多个 CPU 上同时运行的手段. 宏任务子程序可分为以下四类，它提供了并行实体(任务)的控制机制，同步机制和互斥机制.

(1) 任务类子程序：用于处理任务的创建和同步以及管理与任务有关的一些信息.

(2) 事件类子程序：主要用于实现任务间的同步.

(3) 锁类子程序：主要用于代码的临界区和共享数据的保护，锁是一个特殊整型变量，只有两种状态，'LOCKED'和'UNLOCKED'.

(4) 会合类子程序：用于支持一种常见的同步方式——会合，即几个并发任务，在会合点会合后，再恢复运行.

2.3 宏任务库的使用

使用宏任务库时，当任务的粒度小时，由于创建任务和任务之间同步的开销大，效果会很差. 当任务的大小不一时，由于工作量的不平衡，会使某些处理机忙闲不均，产生比预期更低的加速比. 因此，对应用问题的多任务化，有下列基本要求必须考虑：首先，该应用问题可以分成若干个任务，并且这些任务能平衡地分配到各 CPU 上；其次，任务要足够长，值得进行多任务并行计算.

3 T63L16 谱模式多任务并行计算的方案设计

下面以划分两个任务为例，介绍 T63L16 多任务化方案.

3.1 计算的组织

首先多任务并行计算的目标是子程序 NNSC1 中的格点计算部分. 由于有足够的计算机内存存放格点值，我们将一对南北半球纬圈分成两个任务进行并行计算，即一个任务处理一个纬圈. 其次考虑在傅氏空间进行并行计算. 在第一次扫描中，包括傅氏空间，格点空间，傅氏正逆变换和勒让德正变换. 对第一次扫描的多任务化的基本考虑是：对一对南北半球纬圈进行并行计算，但是在计算傅氏系数的对称和反对称部分时，南北半球对应的两个纬圈格点值对它们均有贡献，因此将第一次扫描分为两个计算子段. 第一个子段如图 2(图中方框内的英文字母表示 T63L16 模式的子程序)所示，包括对称与反对称系数的合并、计算纬向导数、傅氏逆变换、动力学计算、傅氏正变换及半隐式调整

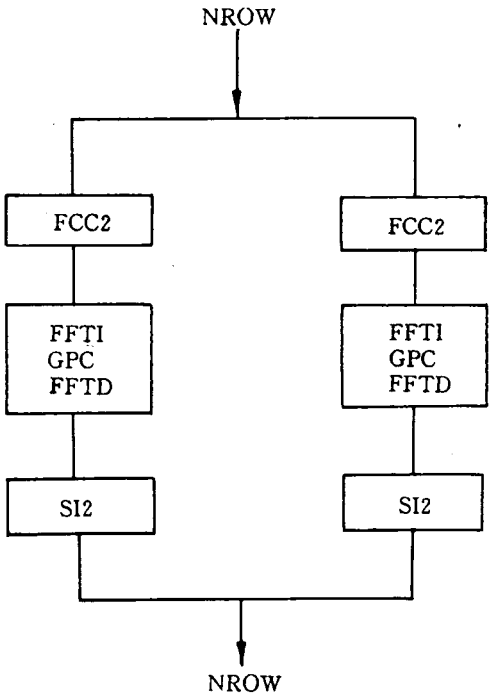

图 2 第一次扫描的第一子段多任务(FCC2：Fourier 空间计算，FFTI：FFT 逆变换，GPC：格点计算，FFTD：FFT 正变换，SI2：Fourier 空间计算)

Fig. 2 The first section of scan 1 (FCC2: computations in Fourier space; FFTI, GPC, FFTD: Inverse FFT computations, grid point computations and direct FFT computations, respectively; FFT —— Fast Fourier Transforms; SI2: computations in Fourier space)

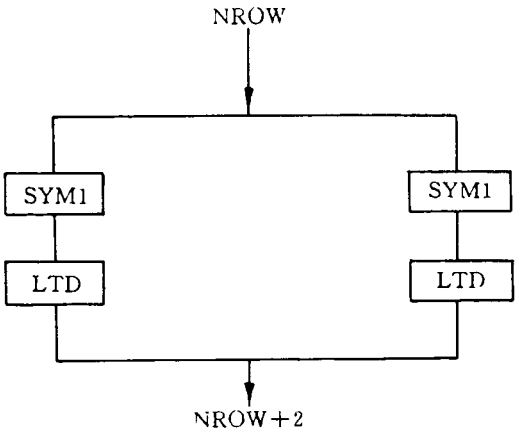

图 3 第一次扫描的第二子段 (SYM1：Fourier 空间计算，LTD：Legendre 正变换)

Fig. 3 The second section of scan 1 (SYM1: computations in Fourier space; LTD: Direct Legendre transforms)

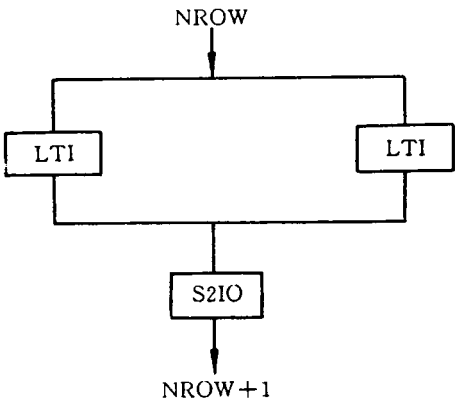

图 4 第二次扫描 (LTI：Legendre 逆变换，S2IO：写 Fourier 系数)

Fig. 4 The structure of scan 2 (LTI: Inverse Legendre transforms; S2IO: write-up of Fourier coefficients)

的第一和第二部分. 第二个子段的计算如图 3 所示，包括将傅氏系数分为对称和反对称部分，勒让德正变换.

第二次扫描的多任务划分比较简单，如图 4 所示，两个处理机同时处理两个纬圈的傅氏系数的对称和反对称部分，然后将傅氏系数写到工作文件中去.

3.2 多任务下的输入输出

T63 支持下列两种输入输出方式：

(1) 无固态存储器　为了对两个纬圈格点值记录进行并行处理，需要在同一时间内读入或写出两个纬圈的格点值记录. 为此我们建立了具有存放两个纬圈格点记录的两倍缓冲区空间的工作文件，以满足并行要求.

(2) 有固态存储器　输入输出方案可改为一个缓冲区，并且是同步的和随机交换的操作方式.

上面的多任务化方案可直接推广到有任意偶数个任务的情况，对于第一次扫描，每一对任务处理南北一对纬圈傅氏系数的对称和反对称部分，可并行执行. 但在计算第一次扫描的第二子段时，有可能同时有几个任务同时去修改同一谱系数，为此当任务数>2 时，我们对谱系数加锁保护，以保证计算的正确性.

4 T63L16 在 YH-2 上的并行计算结果

在独占环境下，在 YH-2 单 CPU 系统，双 CPU 系统和 4 个 CPU 系统上，我们对 T63L16 的运行速度和正确性作了全面测试，速度测试结果如表 1、表 2 所示(表中时间为平均一天预报的开销). 从中可以看出，中期预报模式 T63 在 YH-2 上的并行计算效率比较高，双 CPU 墙上时间的加速比达到 1.67，4 个 CPU 墙上时间的加速比为 2.47. 这说明并行计算的设计方案是成功的.

表 1　T63L16 在 YH-2 上运行的速度测试（单位：h）

Table 1　The CPU and elapsed time of T63L16 running on YH-2 computer (unit: h)

单 CPU		双 CPU		4CPU	
CPU 时间	墙上时间	CPU 时间	墙上时间	CPU 时间	墙上时间
0.5044	0.5100	0.5096	0.3057	0.5286	0.2066

正确性的测试包括两个方面，一是 T63L16 在 YH-2 上运行的正确性，另一方面是 T63L16 的预报能力. 前一个方面的测试包括 T63L16 在 YH-2 上单 CPU，双 CPU，4 个 CPU 运行结果的相互比较，从结果来看完全一致. 表 3 给出了国家气象中心中期数值预报模式 T42L9，T63L16 的统计检验结果，可以看出，在预报能力方面，T63L16 系统的第六天预报其距平相关系数为 0.611，比 T42 系统延长了近一天.

表 2　T63L16 在 YH-2 上运行的各种加速比

Table 2　The speed-up factor of T63L16 on YH-2

墙上时间		CPU 时间	
单机/双 CPU	单机/4 CPU	单机/双 CPU	单机/4 CPU
1.67	2.47	0.99	0.95

表3 T63与T42 1993年11月中期预报检验结果

Table 3 The verification results of T63L16 in November 1993

	距平相关系数						
	(天)1	2	3	4	5	6	7
T63	0.953	0.910	0.856	0.786	0.714	0.661	0.540
T42	0.949	0.886	0.805	0.713	0.611	/	/

5 结 论

综上所述，T63L16中期预报模式在YH-2机上的并行计算方案的设计是成功的，7天预报可在1.5 h内完成，YH-2的4个CPU系统可以满足国家气象中心中期数值预报的业务要求，其预报能力比T42有明显改进.

参考文献

1 杨学胜，乌元康，等. 资料同化和中期数值预报. 北京：气象出版社，1991. 54～313.

2 Gibson J K and Dent D W. A memory manager for single and multi-tasking applications. ECMWF Tech. Mem., 1985, No. 104.

THE MULTI-TASKING CONCURRENT COMPUTING OF MEDIUM-RANGE NUMERICAL WEATHER PREDICTION MODEL T63L16 ON GALAXY-2 SUPERCOMPUTER

Yang Xuesheng Song Junqiang* Huangfu Xueguan

(*National Meteorological Center, Beijing* 100081)

*(*The University of Science and Technology for Defence, Changsha* 410073)

Abstract

The multi-tasking medium-range numerical weather prediction model T63L16 introduced from ECMWF's operational model was developed successfully on the Chinese Galaxy-2 supercomputer which provided a multi-tasking concurrent computer environment, and it was performed efficiently over 2 or 4 processors.

Key Words: Medium-range numerical weather prediction model; Multi-tasking concurrent computional mode; Speed-up factor.