

T42 准业务预报系统中的并行计算

皇甫雪官 张建春
(国家气象局国家气象中心)

提 要

本文描述了国家气象中心 T42 同化预报准业务系统中模式计算和后处理在 M360-AP 的双 CPU 计算机上实现作业级的多任务并行计算。计算结果表明:采用并行计算技术使日常的 5 天数天气业务作业至少提前 60 分钟完成,其加速因子为 1.23。

一、引 言

利用数值模式制作业务天气预报时必须考虑的重要因素之一是预报的及时性。即使计算机执行模式程序及其后处理程序所需要的时间尽量地少,使之在限定的时间里完成预报作业,这样业务预报才具有实用意义。近年来,国外气象预报业务机构,如欧洲中期天气预报中心,英国气象局等均装备了每秒亿次浮点计算的多 CPU 超级向量巨型计算机,对预报作业采取了各种级别的并行计算技术,以减少作业计算的挂钟时间。国家气象中心在 1985 年 5 月从日本富士通购进一台单 CPU 处理机的 M360-R 计算机。在 1989 年此机扩体升级为 M360-AP(Attached Processor)双 CPU 处理机。本文即描述了国家气象中心 T42 同化预报准业务系统中模式计算和后处理在 M360-AP 的双 CPU 计算机上实现作业级上的多任务并行计算。

二、预报模式和后处理在作业级上的多任务并行计算

并行计算是指在同一时间里处理两个或两个以上软件整体的能力。按文献[1],有如下软件并行处理级别:

- 第一级:作业级(JOB)并行。相互依赖的多个作业并行,每个作业占有一个处理机;
- 第二级:同一作业中多个相关的作业步并行,即 JOB STEPS 并行;
- 第三级:子程序级并行(SUBROUTINE);
- 第四级:循环体或程序块并行(DO LOOPS, BLOCK);

第五级;语句(STATEMENT)并行。

上述多级别的并行计算需配有多处理机支持并行计算的软件,如 CRAY-XMP 系列巨型计算机的操作系统提供了多任务并行计算程序库(Multi-Tasking Library,简称 MTL)。该 MTL 中的子程序有分裂任务,可起始并行计算(TSKSTART)、协调各处理机计算进程并使之等待同步(TSKWAIT)、并对某些计算区域加以保护(LOCKON,LOCKOFF)等功能。

为了使 FORTRAN 语言能直接调用,仿照 MTL 的功能,文献[2]提供了几个子程序(如 ¥START, ¥WAIT),以达到预报模式及其结果后处理在计算机 M360-AP 上进行作业级的并行计算。工作原理如下:

1. ¥START 激发后处理作业

批处理作业的运行需提交作业控制语句(JCL),以通知计算机操作系统,再由系统内部读程序(INTRDR)来接收 JCL。因此,在运行预报模式的主作业程序中,用预报时效定时调用子程序 ¥START,可把后处理作业的 JCL 当作数据流输入到内部读程序,启动后处理作业在另一个 CPU 处理机上作为多任务并行运行。

子程序 ¥START(A, N)的功能是将主作业程序中定义的次作业 JCL 过程输入到内部读程序 INTRDR 中去,启动次作业运行。 A 和 N 是子程序 ¥START 的形参,按它们的取值,可分两种情况。

第一种情况: $N = 0, A$ 为任意数组。

在主作业的 JCL 之后,加上如下两条作业控制语句:

//GO. FT91F001 DD DSN=A. B. C(D),DISP=SHR (1)

//GO. FT90F001 DD SYSOUT=(,INTRDR) (2)

设备号 90 和 91 是任意定义的,磁盘内数据集名称采用了复合形式。语句(1)定义了预先作好的次作业 JCL,存放在数据集 $A. B. C$ 的某个成员 D 中。语句(2)是调用系统的内部读程序 INTRDR,以接收提交的次作业 JCL。因此,当主作业的程序中调用 ¥START 时,就可启动次作业运行。

第二种情况: $0 < N \leq n, N$ 为次作业 JCL 的语句条数, n 为存放次作业 JCL 的字符型数组 A 的长度,即 CHARACTER * 80 $A(n)$

在主作业程序中,把次作业的 JCL 采用赋值语句或 DATA 语句等形式定义数组 A ,然后调用子程序 ¥START(A, N),即可启动次作业运行。但次作业的 JCL 需按一般方法制作;而在主作业的 JCL 中,只需要加上语句(2)即可。在第二种情况下,启动多个次作业运行,只需定义多个字符型数组 A 。

若次作业的程序是已经过编译连结的二进制代码目标模块,则可省去次作业程序的编译连结时间。

2. ¥WAIT 可控制并行计算的同步

主作业和次作业对同一任务或不同任务的计算所耗费的计算机时间往往是不等的。为了控制它的在某一步计算处同步,一般设置一个无条件等待过程,这就是子程序 ¥WAIT(M) 即有这个功能。当一个作业接收不到对方发来的同步信息时,为使执行程序

中断一段时间,只需在该计算处加上

```
CALL ¥WAIT (M)
```

即可。其中 M 为等待时间,单位为 0.01 秒。

当引用 ¥START 和 ¥WAIT 时,主作业的作业控制语句,必须和存放这些子程序的库程序连结起来,即需加上如下三条语句:

```
//LKED.LIB DD DSN=USER.LINKLIB,DISP=SHR
//LKED.SYSIN DD *
INCLUDE USERLIB(¥START,¥WAIT)
```

3. 数据文件的处理

由于在作业级上进行并行计算,次作业的激发是当作批处理安排的,因此要求主次作业的数据文件保持独立,以免发生资源竞争,使计算结果不可预测。在 T42 同化预报准业务系统中,设计模式计算及其结果后处理在作业级上并行计算时,保持了它们的数据文件是独立的。在不影响模式计算条件下,仅制作一个小的数据文件,存放一个预报时效的模式输出结果,用于后处理作业的数据输入文件。当一次后处理作业完成后,该文件被下一个时效的模式预报结果所覆盖,继而进行后处理。

三、并行计算结果

目前国家气象中心的 T42 同化预报准业务系统是在 M360—AP 计算机上每天作二次全球资料同化和一次半球五天预报。同化和预报中的预报模式是 T42L9 谱模式。该模式在垂直方向上采用 SIGMA 地形坐标,等分为 9 层,并用能量守恒差分法在垂直方向上离散化;在水平方向上,变量用球谐函数展开,最大的三角形波截断波数为 42;时间积分为半隐式方法,步长取 30 分钟;物理过程中考虑了大地形作用、大尺度凝结降水、干湿对流调整、水平和垂直扩散、辐射以及地表计算。详细描述可参考文献[3]和文献[4]。后处理的内容包括:把预报模式的定时输出结果经过垂直和水平插值,得到各标准等压面上经纬度格距为 2.5 度的全球或北半球值,继而送入场库存档;再从场库中调出这些插值结果,在激光绘画机上输出 100 张天气预报图。

按照第二节的叙述,为使预报模式和后处理在作业级上并行计算,预报模式主作业的作业控制语句可写成下列形式:

```
//CS402HF JOB CLASS=0,MSGCLASS=J,REGION=7500K
//FCST EXEC FORT7CLG,TIME=300
// 提交模式程序的 JCL 语句
//LKED.USERLIB DD DSN=USER.LINKLIB,DISP=SHR
//LKED.SYSIN DD *
INCLUDE USERLIB(¥START,¥WAIT)
// 定义模式计算所需的工作文件和数据文件的 JCL 语句
//GO.FT91F001 DD DSN=CS402.NWP.FORT(POSTJ),DISP=SHR
```

```
//GO. FT90F001 DD SYSOUT=(,INTRDR)
//GO. FT92F001 DD DSN=NWP. PPP. DATA,DISP=SHR
//
```

相应地,在预报模式的定时输出预报结果的子程序 WSHIST 中加入调用启动次作业和等待子程序的语句:

```
SUBROUTINE WSHIST
:
CALL ¥WAIT(M)
WRITE(92) BUF
REWIRD 92
CALL ¥START(A,0)
:
RETURN
END
```

其中 BUF 是存放预报结果的实型数组。当主作业运行到子程序 WSHIST 中的调用 ¥START 语句时,就在通道号 91 上输入数据流 POSTJ,并被系统内部读程序 INTRDR 接收,启动后处理作业在另一个 CPU 处理机上并行计算。POSTJ 就是后处理作业的 JCL 过程,它用一般方法编制而成,这里不一一列出其具体内容。为了建立起后处理的输入数据和预报模式定时计算结果输出二者之间的联系,在预报模式的主作业中最后一条 JCL 语句为:

```
//GO. FT92F001 DD DSN=NWP. PPP. DATA,DISP=SHR
```

它必须在 POSTJ 中出现。

日常的 T42 同化预报系统实时业务在计算机 M360-AP 上计算时,只允许业务作业运行,且其所定义的作业优先等级全为最高的特优等级 O。这样模式启动的后处理作业在另一个 CPU 处理机上运行,于是达到了它们在作业级上的并行计算。按业务规定,模式预报结果定时分八次输出,它们的预报时效依次为 0、12、24、36、48、72、96、120 小时;则后处理次作业相应地被启动八次。因模式进行五天半球天气预报的计算需要 CPU 时间约为 260 分钟,则每作 12 小时预报约用 26 分 CPU 时间,24 小时预报约用 52 分 CPU 时间。每次后处理的出图量不等(6~18),按其最多出图量计算,一次后处理作业的完成需 CPU 时间最多约为 12 分,从后处理作业开始到结束的经历时间最多约为 19 分钟,它小于作 12 小时预报计算的 CPU 时间。因此,后处理作业完全重迭在模式计算的过程中,子程序 WSHIST 中的调用 ¥WAIT(M) 语句不起作用,可以省去。

本文选了 1990 年 6 月 29 日 12 时(世界时)的例子,对并行计算的效率进行了检验。并行计算方案为目前正在使用的 T42 准业务预报系统中的业务方案,它完成模式五天半球预报及其后处理计算所需 CPU 时间为 268 分 39 秒。在同一资料及计算机条件下,采用的串行计算方案是八次后处理放在模式五天半球预报计算后一并进行。串行计算所需的 CPU 时间为 330 分 46 秒。若多任务并行计算的增益用下面的加速因子(Speed-up)来表示:

$$S = T_s/T_p$$

其中 T_s 为单任务串行计算所需要的 CPU 时间, T_p 为多任务并行计算所需要的 CPU 时间。那么, T42 准业务预报系统中模式计算和后处理在 M360-AP 双 CPU 计算机上并行计算的加速因子为 1.23。另外, 将此个例的并行计算与串行计算所需 CPU 时间相比较, 并行计算设计方案至少能使日常的五天半球数值业务预报作业提前 60 分钟。

参 考 文 献

- [1] Gibson, J. K. . A Production Multi-tasking NWP Model, Tech. Memorandum, No. 91, ECMWF, Aug. 1984.
- [2] 武家麟, 程序效率新探, 专题技术报告选编(内部材料), 1988 年。
- [3] 皇甫雪官, 中期数值预报绝热谱模式的数值试验, 《七·五》国家科技攻关课题《中期数值天气预报》成果汇编(一), 1990 年。
- [4] 乌元康等, 中期数值预报试验系统中物理过程参数化的引进和开发, 《七·五》国家科技攻关课题《中期数值天气预报》成果汇编(二), 1990 年。

PARALLEL COMPUTING IN THE T42 QUASI-OPERATIONAL SYSTEM

Huangfu Xueguan Zhang Jianchun

(National Meteorological Center, SMA)

Abstract

This paper describes that the model and postprocessing calculations are executed parallelly at JOB level on the computer M360-AP with two additional processors in the T42 assimilation forecasting quasi-operational system in the National Meteorological Center (Beijing). The results show that the multitasking parallel computing technique makes 5-day NWP completed at least 60 minutes ahead of time and its speed-up factor is equal to 1.23.