

# 记录可寻址随机访问文件的设计与实现

钟志新 宋君强

(国防科技大学电子计算机系 长沙 410073)

**摘要** 记录可寻址随机访问文件是一种新型 FORTRAN 文件, 它集中了顺序与直接文件两者的优点, 为用户随机访问文件中不等长记录提供了更为有效的手段。本文不仅介绍了其使用方法, 而且还讨论了实现中的有关问题。

**关键词** 记录, 文件, 随机访问

**分类号** TP311

就访问记录的方式而言, 标准 FORTRAN 语言只为用户提供了两种外部文件: 顺序文件与直接文件。前者的优点是文件中的记录可以不等长, 但欲访问第  $n$  个记录必须先访问第  $n-1$  个记录; 后者的优点是文件中的记录可被随机访问, 但记录的大小必须等长。不难发现, 如果用户求解问题时需要随机访问文件中不等长的记录, 那么上述两种文件均不能很好地满足用户要求。这种访问记录的方式实际上是要文件同时具备顺序与直接文件两者的优点, 这在标准 FORTRAN 语言中是没有的。为弥补此一不足, 我们提出了记录可寻址随机访问 (Record-Addressable, Random Access, 简称 RARA) 文件的概念。实践表明, RARA 文件能有效地解决用户随机访问文件中不等长记录的问题。在计算机数值天气预报处理等领域有着广泛的应用。

## 1 RARA 文件及其操作

从用户角度看, RARA 文件是一种新型 FORTRAN 外部文件。它的记录都是无格式的, 长度可以不等, 记录由标识符 (即记录号或记录名) 唯一标识并通过它来随机访问记录。在 RARA 文件上可以进行: 打开、关闭、读写、检测错、置异/同步 I/O 方式、等 I/O 完成等 8 种操作。

以什么形式给出这些操作呢? 一是在文本中引入‘方言’, 即新增或扩充 I/O 语句。这势必要修改文本、编译器和 I/O 库, 同时还会破坏用户所编程序的可移植性, 实不可取; 二是设计一组 FORTRAN 可调用专用子程序, 这只要扩充 I/O 库且用户易于接受, 不失为两全齐美之事。

RARA 文件的操作涉及如下二个新概念:

索引及其类型：索引是一张空间由用户提供、管理由 I/O 库负责的能反映 RARA 文件中记录有关信息的表格，由索引项组成，一个记录对应一个索引项，内容包括：记录标识、状态、大小以及它在文件中的位置等。当记录标识符采用记录号形式时，索引项占一个内存单元，此时索引是数字式的；当记录标识符采用记录名形式时，索引项占两个内存单元，此时索引是名字式的。数字/名字索引统称为索引类型。

重写控制：同直接文件一样，RARA 文件中的记录可以被重写。但直接文件的记录因受等长限制，重写时，新记录的长度不得超过规定的记录长度；RARA 文件的记录可以不等长，因而允许新记录的长度 (NRL) 大于老记录的长度 (ORL)，所以，重写时应指明控制方式，例如：

RRC1 方式：新记录总是写在当前文件的尾部；

RRC2 方式：若  $NRL \leq ORL$ ，则新记录写在老记录位置，否则按 RRC1 方式处理；

RRC3 方式：若  $NRL \leq ORL$ ，则按 RRC2 方式处理，否则报错。

下面给出的四个子程序（调用形式、功能、参数说明）能较全面地体现 RARA 文件的特色。

#### (1) 打开一个 RARA 文件

调用形式：CALL OPENYH(fn,index,il,it,iomode[,err])

功 能：为用户打开一个 RARA 文件，并将索引表置成有效状态

参数说明：fn 文件名或部件号

index 索引数组名，大小由 il 和 it 确定，须为盘块大小的倍数

il 索引实用长度，由项数和 it 确定

it 索引类型， $\phi$  表示数字或索引，1 表示名字式索引

iomode I/O 操作方式， $\phi$  表示同步，1 表示异步

err 出错标志，可选参数，当系统认为有错时，返回一错误码

#### (2) 写 RARA 文件的记录

调用形式：CALL WRITYH(fn,ub,m,rf,rrc[,err])

功 能：将用户区 ub 开始的 m 个字数据写入文件 fn 的 rf 记录中，重写控制按 rrc 处理

参数说明：ub 用户区始地

m 欲写数据的字数

rf 记录标识符

rrc 重写控制， $\phi$  表示 RRC1 方式，-1 表示 RRC2 方式，1 表示 RRC3 方式

#### (3) 读 RARA 文件的记录

调用形式：CALL READYH(fn,ub,m,rf[,err])

功 能：将文件 fn 中 rf 记录里的 m 个字数据读到用户区 ub 开始的一片单元中

参数说明：(同上)

#### (4) 关闭一个 RARA 文件

调用形式：CALL CLOSYH(fn[,err])

功能：将对应 fn 的索引从用户区写入 fn 的尾部，整理输出统计信息，然后按 FORTRAN 的 CLOSE 语句 (statu='keep') 关闭 fn

参数说明：(同上)

下面用 RARA 文件及有关操作来解决这样一个问题：将 K 个大小不超过 RS 的记录从终端文件中读进存于系统，以后若有老记录要更新，只须给出修改个数和新的记录即可，其中 K、RS 和盘块大小 BS 是已知的，记录的首字是其标识符。程序与注释如下：

第 4 句为定义索引表空间；第 6 句为打开 RARA 文件 F；第 7 句为读欲处理的记录个数，第一次作业时，M=K，以后的作业 M=修改个数；第 10 句为取记录标识符；第 11 句为写记录到文件 F 中，第一次作业完成存入系统，以后的作业完成新老记录的更新。

```
1 PROGRAM INI&UPD
2 PARAMETER(BS=...,K=...,RS=...)
3 PARAMETER(IS=(INT(2 * K/BS)+1) * BS)
4 INTEGER INDEX(IS)
5 INTEGER UB(RS)
6 CALL OPENYH('F',INDEX,2 * K,1,phi)
7 READ(*,'(1X,...)')M
8 DO I=1,M
9 READ(*,'(1X,...)')UB
10 RF=UB(1)
11 CALL WRITYH('F',UB,RS,RF,-1)
12 ENDDO
13 CALL CLOSYH('F')
14 END 'INI & UPD'
```

可见，编程简单高效，更新用户透明。同样的问题，若用顺序或直接文件来解决，不仅编程更为复杂，而且效率必然降低。

## 2 RARA 文件的实现

从系统的角度看，RARA 文件是一种没有 I/O 缓冲区的随机盘文件。它的创建、与主存用户区之间的直接数据传送，都是在 I/O 库与 OS 的协同控制下完成的。图 1 给出了实现 RARA 文件操作的逻辑关系。

其中，FPT 用于控制读写操作，表信息主要有：文件名、位置域（用户区首尾地址、记录在盘文件中的逻辑块号）、统计域（当前文件中记录的最大最小值、读写访问的次数）、索引控制字（由 OPENYH 所给参数形成）等；FST 用于反映文件状态，一个文件对应一个表项，包括一个相对 I/O 库而言的可知子项和一个相对用户程序而言的存在子项，每个子项均由文件名和 FPT 地址组成。

成功地打开一个 RARA 文件，将在系统内形成如下一条关系链：

TFCT 表项→FST 可知子项→FST 存在子项→FPT→IT

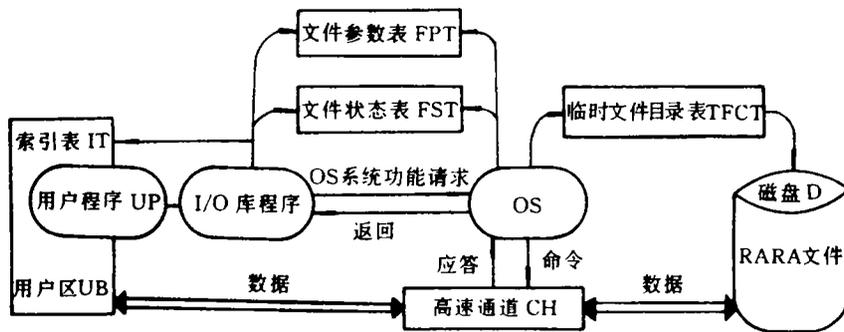


图 1 RARA 文件操作逻辑关系

当用户为打开一个 RARA 新文件而调用子程序 OPENYH 时，OPENYH 就用参数 FN 查 FST 以判明当前文件的状态。因为是新文件，FST 中无对应表项，所以，OPENYH 首先为该文件申请 FPT 表空间，接着以 FN 和 FPT 地址为参数向 OS 发创建打开该文件的请求，结果是 TFCT 中有了与 FN 对应的表项，FST 中有了与 FN 对应的可知子项；然后，OPENYH 根据所给的参数形成索引控制字并填入 FPT 中，将索引数组初始化（清  $\phi$ ），在 FST 中填入与 FN 对应的存在子项，以表明索引已有意义，可以对 FN 进行合理的读写等操作了。

读写操作要根据索引项的内容来确定记录在文件中的位置；索引项在写操作过程中形成；写操作受重写控制的约束；尽管用户以随机方式写记录，但它在盘文件中的存放次序主要还是顺序的。这些都是 RARA 文件与标准 FORTRAN 文件在处理上的主要不同之处。图 2 给出了读写子程序的主要处理流程。

### 3 三种文件的比较

作为一种新型 FORTRAN 外部文件，它以及其上的操作对用户提出了一些合理的要求，使用时应加以注意：

1) 它们不能同标准 FORTRAN 文件及其 I/O 语句交叉使用；2) 读写时所给的记录标识符必须与打开该文件时指明的索引类型一致；3) RARA 文件没有 I/O 缓冲区。它与用户区之间的直接数据传送物理上是以盘块大小为基本单位进行的。从而用户定义的索引数组大小、读操作时 UB 开的大小至少应为满足实用长度的盘块大小的整数倍，否则将破坏下邻内存单元的信息。

为从整体上加深 RARA 文件概念的印象，我们可以对顺序/直接/RARA 三种文件的有关属性作一个对比：

表 1

文件类型	记录标识	记录大小限制	记录格式	访问方式	记录在文件中的序放次序
顺 序	无	无	F/UF	顺序的	与写次序一致
直 接	记录号	等长	F/UF	随机的	与记录号一致
RARA	记录号/名	无	UF	随机的	一般与写次序一致
				重写控制	但 RRC2 重写可例外

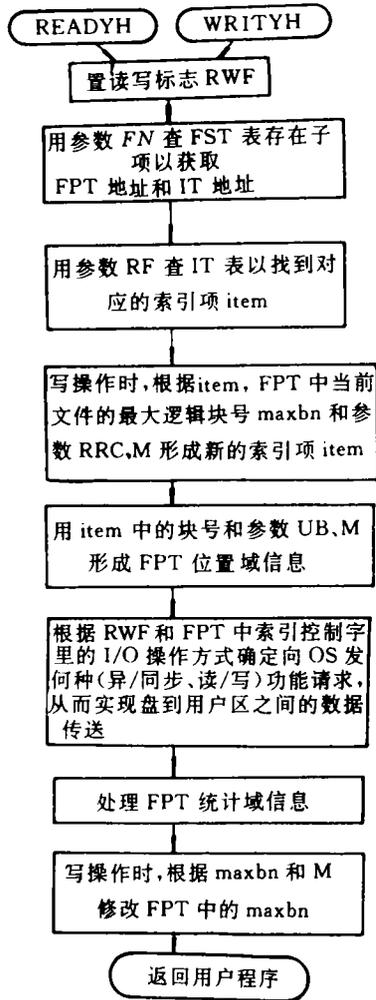


图 2 读写处理流程

# **The Design and Implementation of Record-Addressable, Random Access File**

Zhong Zhixin Song Junqiang

(Department of Computer Science, NUDT, Changsha 410073)

## **Abstract**

The Record-addressable random access file (RARA) is a new-type file in FORTRAN. It combines the advantages of the sequential and direct files, and provides a more efficient means for the users in their record unequal in length in the random access file. In this paper, we not only introduce the way to use RARA file, but also discuss the problems in implementation.

**Key words** record, file, random access