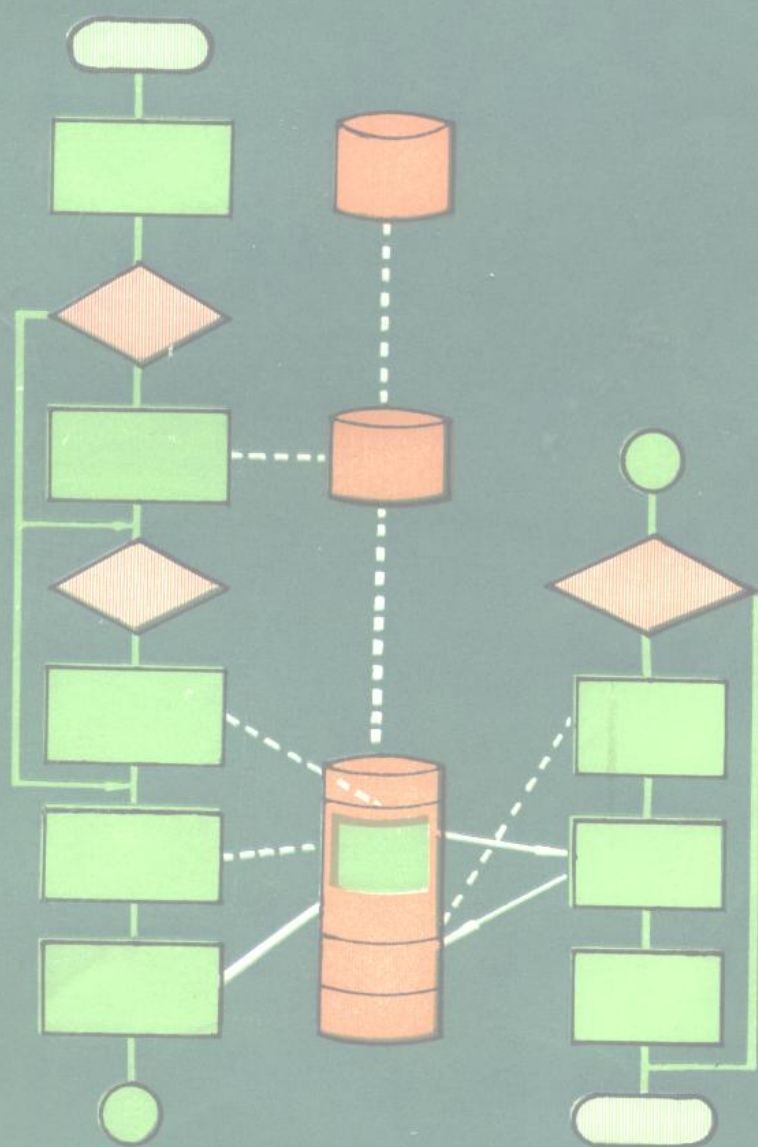


计算机文件的 结构和处理

赵立平 张治 编

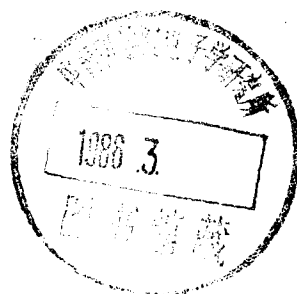


新 时 代 出 版 社

7318
774

计算机文件的结构和处理

赵立平
张 治 编



新时代出版社

8610194

内 容 简 介

本书首先介绍文件管理的基本概念和存储介质的特性,然后分述了顺序文件、分区文件、直接文件、索引顺序文件、虚拟存储文件的结构和处理方法,对常用的顺序文件和直接文件作了较详细的描述。

本书适合于大专院校计算机软件专业的师生、计算站软件人员和广大用户阅读参考。

计算机文件的结构和处理

赵立平 编
张 治

新时代出版社出版 新华书店北京发行所发行

国防工业出版社印刷厂印刷

787×1092毫米 16开本 11.5印张 264千字

1985年10月第1版 1985年10月北京第1次印刷

印数: 0001—8100册

统一书号: 15241·63 定价: 2.15元

前 言

在社会生活现代化的进程中，电子计算机已成为处理信息的有效工具。文件是计算机系统存储和加工数据的基本手段，本书旨在介绍计算机文件的结构和主要处理方式。

文件管理是计算机系统的主要组成部分之一，它统一管理存储在各种介质上的数据，使之能简便地存取，并保证安全可靠、共享和保密。用户应了解所使用文件的抽象组织形式（逻辑结构），软件人员则应了解各种文件在每类介质上的具体组织情况（物理结构）。文件管理实现这两种结构之间的映象和转换，并按需要对文件进行处理，如编制文件，检索、更新、追加或删除记录，以及再编制文件等。

近年来，我国陆续引进国外先进的大、中型计算机，它们已逐渐在国民经济的各行各业中发挥作用。为了进一步提高计算机的使用效率，软件人员和广大用户必须掌握计算机文件的结构和处理方法，以便得心应手地设计文件和调用文件管理提供的各种功能。本书主要以HITAC M-160 II计算机的VOS 2操作系统（与IBM/370兼容）为背景，着重从使用的角度介绍文件的逻辑结构，并适度地介绍一部分物理结构。全书由浅入深地引入各种概念，通过汇编语言、FORTRAN语言和COBOL语言的程序例题来说明对文件的各种处理方法。

全书共分七章，第一章介绍文件管理的基本功能，阐述与文件有关的基本术语和概念，从纵的方向简述文件的各种组织方式、存取法和处理方法；第二章描述常用的磁盘、磁带和软盘介质；第三章至第七章，从横的方向分述顺序文件、分区文件、直接文件、索引顺序文件和虚拟存储文件的结构特点、存取法和处理方法。

由于编者水平有限，缺点和错误之处难免，敬请读者批评指正。

目 录

第一章 文件管理概述	1
1.1 计算机系统与文件	1
1.1.1 计算机系统的构成	1
1.1.2 文件的概念及文件管理的作用	2
1.1.3 作业控制语言简述	3
1.2 基本概念和术语	6
1.2.1 记录	6
1.2.2 块与缓冲区	6
1.2.3 记录的形式	7
1.2.4 记录组合的分析	11
1.2.5 文件的介质	12
1.3 磁盘文件的空间分配	16
1.4 文件的组织与存取法	18
1.4.1 文件的组织	18
1.4.2 存取法的功能和分类	20
1.5 文件的定义方法	20
1.5.1 什么是数据控制块 (DCB)	21
1.5.2 DCB的信息来源	21
1.5.3 DCB的完成	22
1.5.4 出口例行程序	23
1.6 文件的目录管理	24
1.6.1 目录的功能和分类	24
1.6.2 目录的结构和文件名	25
1.6.3 目录的维护	27
1.7 文件的处理	30
1.7.1 打开文件	30
1.7.2 关闭文件	30
1.7.3 文件的使用处理	30
1.7.4 文件的保护	32
第二章 磁盘、磁带和软盘	33
2.1 磁盘	33
2.1.1 磁盘机的主要性能	33
2.1.2 磁盘组的构造和动作方式	34
2.1.3 磁道的形式	35
2.1.4 磁盘卷的处理	36
2.1.5 文件容量的计算	41
2.2 磁带	42
2.2.1 磁带机概况和磁带卷的形式	42
2.2.2 磁带卷的文件形式和标准标号	44
2.2.3 磁带的处理方式	47
2.3 软盘	47

2.3.1 软盘的构造	47
2.3.2 软盘内的文件形式	49
2.3.3 软盘的处理	50
第三章 顺序文件	51
3.1 顺序文件概述	51
3.2 顺序文件的存取法	52
3.2.1 排队顺序存取法QSAM	52
3.2.2 基本顺序存取法BSAM	59
3.3 顺序文件的处理方法	63
3.3.1 顺序文件的编制	63
3.3.2 顺序文件的检索	72
3.3.3 顺序文件的维护	74
第四章 分区文件	84
4.1 分区文件概述	84
4.2 分区文件的存取法	86
4.2.1 用顺序存取法处理一个成员	86
4.2.2 基本分区存取法BPAM	86
4.3 分区文件的处理方法	89
4.3.1 分区文件的编制	89
4.3.2 分区文件的维护	92
4.3.3 利用实用程序处理分区文件	94
第五章 直接文件	98
5.1 直接文件概述	98
5.1.1 直接文件的特点	98
5.1.2 直接文件的记录地址	98
5.2 直接文件记录的地址变换	99
5.2.1 地址变换的方法	99
5.2.2 处理冲突的方法	102
5.3 直接文件的存取法	104
5.3.1 用BSAM处理直接文件	104
5.3.2 基本直接存取法BDAM	106
5.4 直接文件的处理方法	110
5.4.1 直接文件的编制	110
5.4.2 直接文件的记录检索	117
5.4.3 直接文件的维护	117
第六章 索引顺序文件	126
6.1 索引顺序文件概述	126
6.2 索引顺序文件的结构	126
6.3 索引顺序文件的处理方式	128
6.4 排队索引顺序存取法QISAM	129
6.5 基本索引顺序存取法BISAM	132
第七章 虚拟存储文件	135
7.1 VSAM文件的管理和分配	135
7.1.1 VSAM目录	135
7.1.2 VSAM数据空间	135

7.2	VSAM文件的组织	137
7.2.1	构成VSAM文件的基本要素	137
7.2.2	键顺序文件KSDS	141
7.2.3	输入顺序文件ESDS	143
7.2.4	相对记录文件RRDS	143
7.3	VSAM文件的存取法	144
7.3.1	键顺序文件的存取法	145
7.3.2	输入顺序文件的存取法	148
7.3.3	相对记录文件的存取法	148
7.3.4	缓冲区的管理和数据传送方式	149
7.4	VSAM文件的保护	150
7.4.1	机密保护功能	150
7.4.2	文件的保存日期和删除后的保护功能	151
7.5	VSAM文件的处理	151
7.5.1	VSAM文件的编制	151
7.5.2	VSAM文件的维护	154
7.6	更替索引	156
7.6.1	更替索引的构造	157
7.6.2	更替索引的存取	158
7.6.3	更替索引的编制	160
附录 I	EBCDIC 代码表	161
附录 II	常用的外设名称和表示	162
附录 III	常用作业控制语句及其意义	163
附录 IV	VTOC一览表	169
附录 V	磁带标准标号的形式和内容	172
附录 VI	汇编指令说明	174
附录 VII	英汉用语对照表	178

第一章 文件管理概述

本章主要介绍文件管理的作用、作业控制语言的基本内容以及文件管理的基本功能。

1.1 计算机系统与文件

1.1.1 计算机系统的构成

计算机系统是硬件与软件的总称。硬件是指电子数字计算机的各部分设备，主要包括中央处理机、主存储器、通道和外部设备。中央处理机（简称CPU）通过翻译和执行指令，控制整个计算机的运行。其中备有十六个通用寄存器（简称寄存器），编号为0^{*}，1^{*}，…，15^{*}。主存储器（简称主存或内存）是计算机存储程序或数据的装置，它可以直接与中央处理机交换信息。八个二进制位构成一个字节，作为信息处理的最小单位。通常，可以用一个字节中的各二进制位不同的组合状态来表示各种字符（数字、字母或特殊符号），这些组合状态称为代码。计算机中常用的代码之一是EBCDIC码（见附录I）。一个字节可以存放两个十六进制数，两个字节组成半字，四个字节组成一个字。如图1.1所示，图中两个字的内容是字符'FILE16'。

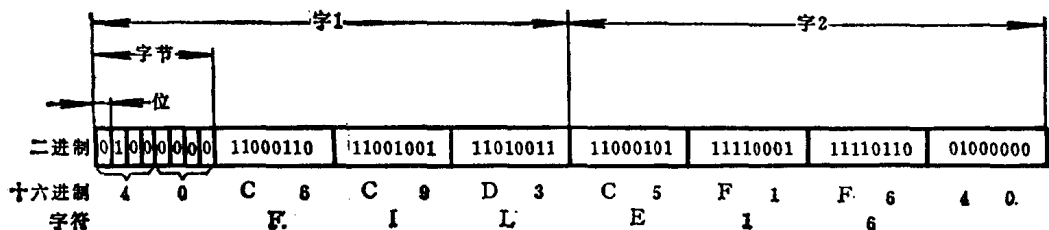


图1.1 位、字节、字和EBCDIC代码

外存储器是存放大量信息的存储介质，例如卡片、纸带、打印纸、磁盘和磁带等。它们的存储容量大，但存取速度比主存慢，用于存放永久性信息。外部设备是外存储器和输入输出设备（卡片输入机、纸带输入机、打印机、磁盘机和磁带机等）的总称，简称外设。常用外设的名称和图形表示见附录II。通道是主存与外设之间进行信息传输的控制装置。

计算机软件是各种程序的总称，其目的在于方便用户和充分发挥计算机的效能。操作系统是软件的主要组成部分，它包括控制程序和处理程序。控制程序负责监督、管理和协调作业的执行并及时处理特殊情况。在控制程序的监督和调度下，处理程序对数据进行具体的加工处理。控制程序按功能分为管理程序、作业管理、文件管理、恢复管理和系统生成等部分。

1.1.2 文件的概念及文件管理的作用

在用途、形式、结构和内容上彼此相似的数据的集合叫做文件。赋予文件的名字叫文件名。不管是原始数据、处理后得到的计算结果、还是一段程序，从计算机处理和外部介质存储的角度来看，都是按一定的代码规则变换后存储在某种介质上，成为计算机系统的处理对象，它们统称为数据或信息。通常把存储在磁盘上的一组数据称为磁盘文件，存储在磁带上的一组数据称为磁带文件，一组数据的穿孔卡片称为卡片文件。在 VOS 2 系统中，文件又称为数据集，文件名又称为数据集名 (DSN)。例如，图书馆索引卡文件取名为 INDEX 7，它由每本书的序号、索引号、书名、编著者和出版日期等数据组成，如图 1.2 所示。

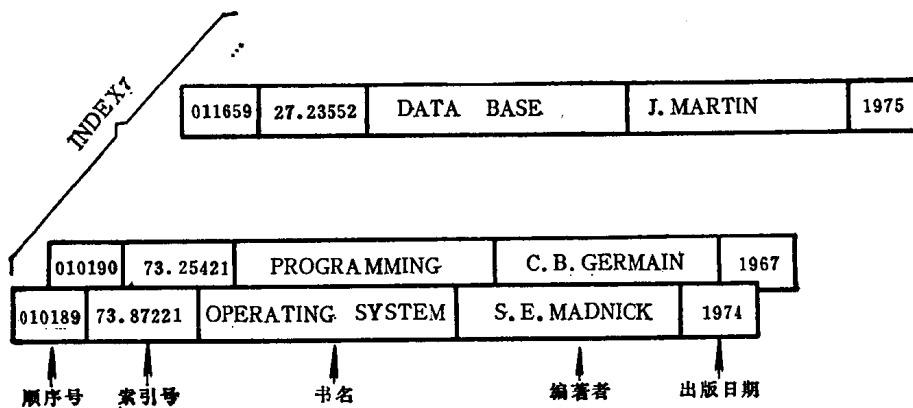


图1.2 图书索引卡文件

我们可以从两种角度来研究文件的结构。一种是文件的逻辑结构，即用户想象中的抽象文件，这是从处理数据的角度出发，为便于用户进行构思和设计而提供的层次清晰、使用方便的结构形式。另一种是文件的物理结构，主要研究文件在介质上的具体组织形式、存储位置和存取方式。文件管理的重要作用之一就是要在用户的逻辑文件和相应介质上的物理文件之间建立映象关系，实现两者之间的相互转换。

文件管理又称为数据管理，它是操作系统的重要组成部分。它有效而方便地管理外部介质上的系统信息和用户数据（文件），自动分配文件区域，提供不依赖于外设具体特性的各种存取文件的功能。文件管理给用户带来的好处是：

(1) 使用方便 用户给出文件名，使用有关输入输出宏指令即可存取各种介质上的数据，文件管理自动处理存取细节。

(2) 安全可靠 文件管理设立了多种控制表，对各种文件的区域分配和使用状态进行严格的控制和管理，使得诸文件的区域不受侵犯，相互之间不受干扰，保证文件在外部介质上安全可靠。

(3) 共享与保密 文件管理设置了一些保密措施，使得未经文件所有者明确授权的用户不能存取文件，被授权的用户则可在规定的权限内简便地使用其文件。此外，当多个用户需要同时访问某一文件时，使文件具有共享的特性。

1.1.3 作业控制语言简述

1. 作业和作业步的概念

作业是用户准备好的、向计算机系统提交处理工作的单位，赋给作业的名称称为作业名。一般情况下，除了通过存储在外部介质上的文件交换信息外，各作业之间是相互独立的。一个作业往往可以分为若干个步骤来完成，每一步骤称为一个作业步，赋给作业步的名称称为作业步名。一个作业步就是通常所说的一个处理程序，因此作业也可以看作是连续处理的、一系列程序的集合。一个作业中的各个作业步可以相互关联，也可以各自独立，操作系统将按照输入的先后顺序来处理各作业步。例如一个用 FORTRAN 语言编写的源程序，可以作为一个作业提交给计算机系统，通常它含有编译、连接编辑和执行三个作业步，前一个作业步的输出结果是后一个作业步的输入数据，只有完成前一作业步，并确认完全正确后，进行下一个作业步才有意义。一般来说，用汇编语言或其它高级程序设计语言编写的源程序，也都具有这三个作业步。

2. 作业控制语句

在多道运行的计算机系统中，用户不能直接控制作业的运行过程，必须将有关作业的各种控制信息（如程序名，必要的外设类型、数量以及有关文件等）通知计算机系统。用来记述作业控制信息的语言叫作业控制语言（JCL）。作业控制语言由一系列的作业控制语句组成。用户提交作业时，把作业控制语句和程序、数据一起交给计算机系统。常用的作业控制语句有作业（JOB）语句、执行（EXEC）语句、数据定义（DD）语句、分段语句和空语句等。它们都是以卡片形式记述的，当一张卡片不够时，可使用续卡，续卡的头两个字符必须都是'/'。下面简述常用作业控制语句的格式和意义，各种语句的细则请参阅附录Ⅱ。

（1）作业语句（JOB语句）

格式：

//作业名	JOB	操作数
-------	-----	-----

作业语句表示一个作业的开始，用来指定作业的名称和作业的属性。作业的属性通过操作数给出，包括作业级、作业优先数、系统信息输出级、系统信息输出量、作业的执行时间、程序员名以及帐号等。作业语句必须放在一个作业的最前面，它是一个作业不可缺少的控制语句。

（2）执行语句（EXEC语句）

格式：

//作业步名	EXEC	操作数
--------	------	-----

执行语句表示作业步的开始，用它来指定作业步名和作业步的属性。作业步的属性通过操作数给出，包括在该作业步中执行程序的名称或编目过程名[●]、作业步的执行条件、帐号和执行时间等。

● 把一组作业控制语句登记在系统文件（编目过程库）中，称为编目过程，赋予编目过程的名字叫编目过程名。编目过程中的一个作业步叫过程步，赋予它的名字叫过程步名。

(3) 数据定义语句 (DD语句)

格式:

//DD名	DD	操作数
-------	----	-----

数据定义语句对作业或作业步中使用的文件进行定义。它用数据定义语句名 (DD名) 来识别, 通过操作数给出文件名、文件所在外部设备的种类和数量、外部设备的卷序列号、文件空间分配方式和分配量、文件的属性 (组织形式、记录形式与记录长度等)、文件标号类型、文件的状态及作业或作业步结束后的处理等。一个数据定义语句对应于一个文件。

(4) 分段语句

格式:

/*

表示从卡片输入机输入数据的结束。

(5) 空语句

格式:

//

表示作业的结束。

一个作业的范围是从作业语句开始到空语句为止或者到下一个作业的作业语句之前。一个作业步的范围是从执行语句开始到下一个作业步的执行语句之前, 或者到空语句或分段语句为止, 见图1.3。作业步内的程序结束或中断时, 为了表示结束或中断的原因及状态, 系统把某个十六进制数放在15*寄存器[●]中, 这个数称为作业或作业步的返回

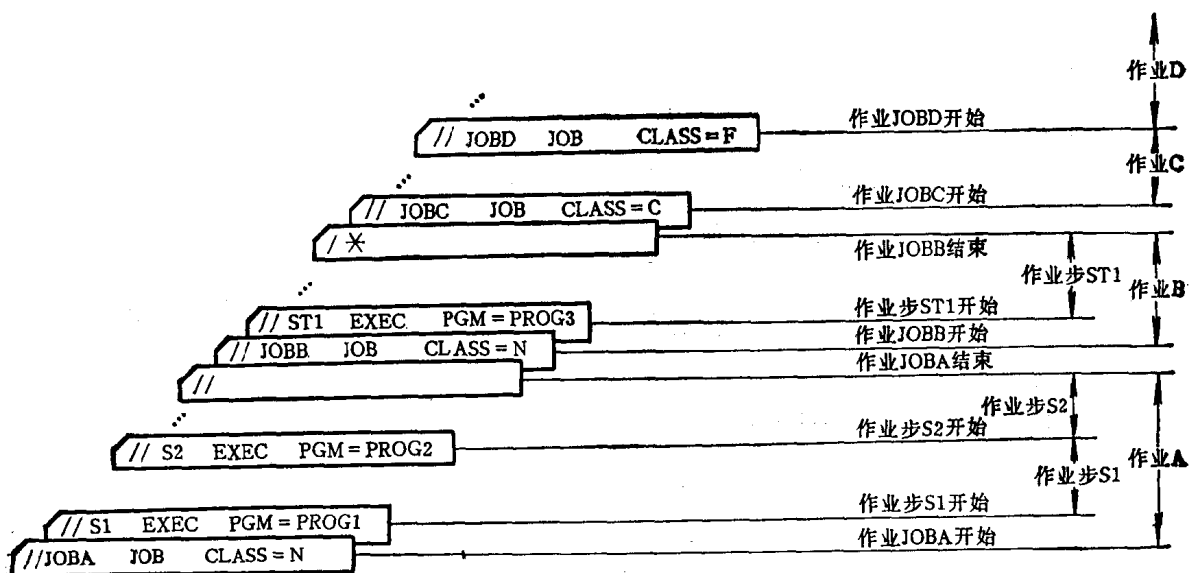


图1.3 作业和作业步的范围

- 在16个通用寄存器中, 0*、1*、13*、14*和15*寄存器通常由系统使用。13*寄存器中存放寄存器保护区地址, 14*寄存器中存放程序连接时的返回地址, 15*寄存器中存放程序的入口点和返回码。在文件管理中, 0*和1*寄存器由存取法的宏指令使用。除13*寄存器外, 用户程序可以临时使用0*、1*、14*和15*寄存器。其余寄存器可任意使用。

码。一般情况下，程序正常结束，返回码置为零。在作业控制语句中，用户可以指定如何根据返回码来控制以后作业步的执行。

下面以例说明作业控制语句的写法和安排。

例1.1

```
//SAMPLE JOB 【SX01, WANG, CLASS = B,
// PRTY = 10, MSGCLASS = A, MSGLEVEL = ( 1, 1 ),
// TIME = 5 .....①
//JOB LIB DD DSN = PGMLIB, UNIT = DISK,
// VOL = SER = D00007, DISP = (SHR, KEEP, KEEP) .....②
//STEP 1 EXEC PGM = TEXT01 .....③
//DA DD DSN = DAFILE, UNIT = DISK,
// VOL = SER = WDISK 1, DISP = (NEW, KEEP, DELETE),
// DCB = (RECFM = FB, LRECL = 80, BLKSIZE = 800),
// SPACE = (TRK, (10, 1)) .....④
//TAPE DD DSN = MTFILE, UNIT = MT,
// VOL = SER = WTAPE 1, DISP = OLD, LABEL = ( 1, SL) .....⑤
// .....⑥
```

相应序号的语句含义如下：

① 作业语句，作业名为SAMPLE，帐号为SX01，程序员名为WANG，作业级（CLASS）为B，作业优先数（PRTY）为10，系统信息输出级（MSGCLASS）为A，系统信息输出量（MSGLEVEL）为（1，1），即全部输出系统信息，作业最长执行时间（TIME）为5分钟。

② 该DD语句定义的是存放执行程序TEXT01（见③）的程序库，这种DD语句的名字约定为JOB LIB[●]，文件名（DSN）为PGMLIB，外部设备（UNIT）是磁盘（DISK），磁盘卷序列号（VOL = SER）为D00007，文件状态和处理（DISP）是：该文件已存在于磁盘中，并为多个用户共享（SHR），作业步正常或异常结束时，都保留在介质上（KEEP）。

③ 执行语句，作业步名为STEP1，执行程序名（PGM）是TEXT01。

④ 用名字为DA的DD语句定义该作业步内要编制的磁盘文件（UNIT = DISK），文件名为DAFILE，磁盘卷序列号为WDISK1，文件最初所占空间（SPACE）为10个磁道（TRK），当空间不足时，每次增加一个磁道。文件记录形式（RECFM）为固定长组合形式（FB），记录长（LRECL）为80字节，块长（BLKSIZE）为800字节。该文件的状态和处理是：它是本作业步内新编制的文件（NEW），作业步正常结束时，保留在介质上（KEEP），异常结束时，从介质上删除（DELETE）。

⑤ 用名字为TAPE的DD语句定义该作业步内要使用的磁带文件（UNIT = MT），文件名为MTFILE，所在磁带卷序列号为WTAPE1，在该作业步之前，文件已存在于磁带中（OLD）。磁带具有标准标号（SL），该文件是磁带内顺序排列的第一个文件。

⑥ 表示作业结束。

● 这种情况下，该文件（PGMLIB）为作业库，存放该作业内各作业步的执行程序，在整个作业内有效。若将该DD语句放在EXEC语句后面，DD名改为STEPLIB，则该文件称为作业步库，存放该作业步内的执行程序，仅在该作业步中有效。

1.2. 基本概念和术语

本节从逻辑结构和物理结构两个角度, 介绍文件的构成、记录的各种形式和介质的存储特性。

1.2.1 记录

记录是由数据组成的, 它是用户对文件进行处理的基本单位, 又是文件逻辑结构的基本成分。在本书中记录系指逻辑记录。

在数据处理中, 表示一件事或一个人某些属性的数据就可构成一个记录。例如在一个学生成绩统计的文件中, 每个学生的学习成绩可表示为一个记录, 它由学号、姓名、各科成绩、总成绩、平均成绩、名次、备注等数据组成, 如图 1.4 所示。图中, 命名为学号、姓名、语文成绩等数据, 称为记录的项目 (简称项), 它由若干个二进制位或字节组成。各记录的学号是互不相同的, 它可用来识别记录, 我们称它为该记录的关键字或键。

学号	姓 名	各 科 成 绩							总成绩	平均成绩	名 次	备 注
		语文	数学	物理	化学	英语	政治	生物				
47	LI MING	90	95	89	86	92	95	91	638	91.1	7	GOOD

图1.4 学生成绩记录

1.2.2 块与缓冲区

外部介质上的文件与主存进行数据交换的基本单位, 称为块或物理记录。块可由一个或多个记录组成, 前者称为非组合记录 (或非成块记录), 后者称为组合记录 (或成块记录)。只有磁性介质上才允许有组合记录形式。

为了解决计算机中高速主存与慢速外部设备之间交换数据时速度不匹配的矛盾, 在主存中开辟输入输出缓冲区, 它是主存与外部介质上的块进行数据传输的暂存区域, 在主存与外设之间起接口的作用。缓冲区的长度与块的长度相对应。

这里, 我们以文件管理设立的缓冲区为例, 说明数据的传输过程 (图1.5)。数据输出 (从主存的角度看) 时, 用户使用输出宏指令, 将一个个记录送入缓冲区, 当缓冲区填满时, 文件管理自动地启动外部设备, 将缓冲区中组合起来的各记录传送到外部介质

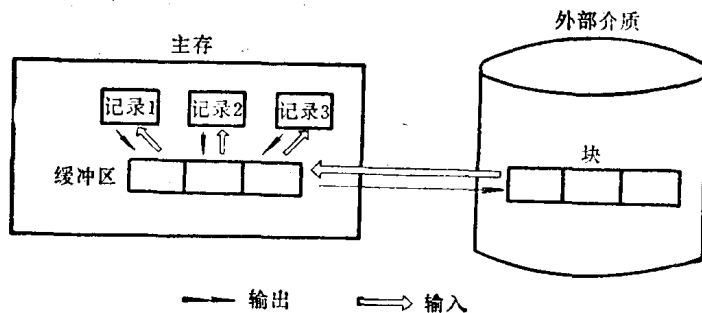


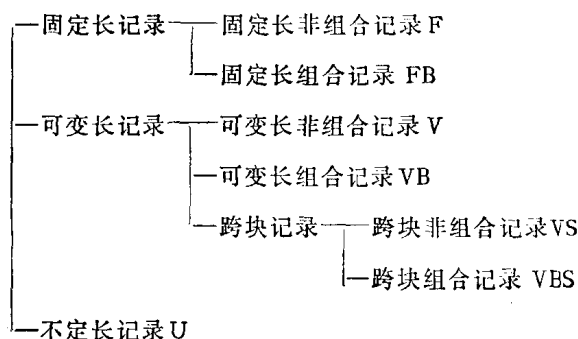
图1.5 数据传输过程

上作为一个块存储起来。数据输入（也是从主存的角度看）时，用户发出输入宏指令，文件管理自动地启动外部设备，将外部介质上的一块读取到缓冲区中，用户从缓冲区取出各记录，进行加工处理，当缓冲区内的诸记录被取用完后，文件管理又可自动地读取新的一块。

在实际使用时，CPU与通道传输是并行工作的，而CPU的运算速度比通道传输速度快得多，为了减少CPU对通道的等待时间，通常设立两个以上的缓冲区（称为缓冲池），以进一步提高整个计算机系统的效率。

1.2.3 记录的形式

根据记录的长度变化与否、记录组合成块的形式以及存储介质的不同，记录可分为下列七种形式：



根据各个操作系统内文件管理的功能，用户可使用其中全部或部分记录形式。

1. 固定长记录

所谓固定长记录，就是文件中全部记录的长度都相等。该记录形式又有下列两种：

(1) 固定长非组合记录 F 一个块只由一个记录构成，记录长与块长一致，各记录的长度都相等（图1.6）。

(2) 固定长组合记录FB 一个块是由两个以上的记录组合起来构成的，每个块中，各记录的长度都相等，记录的数目也都一致，因此块的长度也都相等（图1.7）。

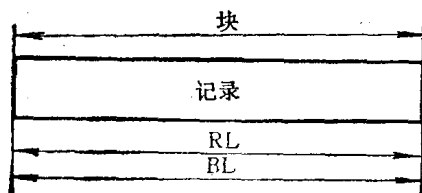


图1.6 固定长非组合记录形式

RL—记录长；BL—块长。

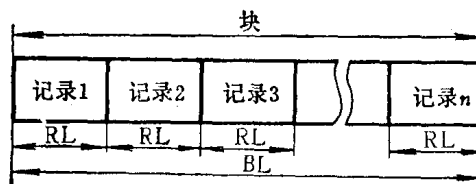


图1.7 固定长组合记录形式

2. 可变长记录

在某些数据处理中，由于各记录需用的项目个数或长度不尽相同，各记录的长度也就不相等。如果仍采用固定长记录形式（例如取最大记录长作为统一的记录长），就会造成介质的浪费。这时，可以把每个记录实际占用长度的信息附加在记录前部，把每个

块实际占用长度的信息附加在块前部, 构成可变长记录。它具有下面两种形式:

(1) 可变长非组合记录V 一个块只由一个记录构成, 各块中记录的长度可以不相等, 因而块的长度也不尽相等 (图1.8)。

从图中可以看出, 每个记录的前面有四个字节存放记录实际占用长度 (简称记录长) 的信息, 这四个字节称为记录描述字 RDW。

记录长包括记录本身 (即组成记录的数据部分) 的长度和 RDW 所占的四个字节。每个块的前

部有四个字节存放块实际占用长度 (简称块长) 的信息, 这四个字节称为块描述字 BDW。块长也必须把 BDW 所占的四个字节计算在内。记录描述字和块描述字的详细内容如下:

RDW

字节 0 ~ 1 = $RL_{(16)}$ ● 记录长度的十六进制表示 ●

字节 2 ~ 3 = 0000₍₁₆₎

BDW

字节 0 ~ 1 = $BL_{(16)}$ 块长的十六进制表示

字节 2 ~ 3 = 0000₍₁₆₎

例如, 如果记录的数据部分长度为80字节, 那么, 记录描述字的第 0 ~ 1 字节和块描述字的第 0 ~ 1 字节将分别为 $0054_{(16)} = 84$ 和 $0058_{(16)} = 88$ 。

(2) 可变长组合记录VB 一个块是由两个以上的可变长记录组合起来构成的, 各块中记录的个数可以不同, 块的长度也不尽相等 (图 1.9)。

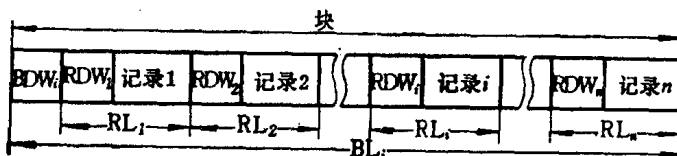


图1.9 可变长组合记录形式

这种记录形式与可变长非组合记录形式相似, 每个记录的前面有 4 个字节的记录描述字 RDW_i , 前两个字节是记录 i (包括 RDW_i) 长度 RL_i 的十六进制表示。每个块的前部是 4 个字节的块描述字 BDW_i , 前两个字节是块 i (包括 BDW_i) 长度 BL_i 的十六进制表示。

例如, 某个块由三个可变长记录组成, 三个记录数据部分的长度分别为80字节、100字节和120字节。那么三个记录的长度分别是 $RL_1 = 84$ 字节、 $RL_2 = 104$ 字节和 $RL_3 = 124$ 字节。这些信息以十六进制的形式存放在记录描述字的第 0 ~ 1 字节中, 分别为 $0054_{(16)}$ 、 $0068_{(16)}$ 和 $007C_{(16)}$ 。整个块的长度为 $BL = 316$ 字节, 这个信息也以十六进制的形式存放在块描述字的第 0 ~ 1 字节中, 即 $013C_{(16)}$ 。

● 本书中, 十六进制数在数的右下角用 (16) 表示, 而常用的十进制数采用一般的写法, 省写右下角 (10)。
● 单位均为字节。

3. 跨块记录

在某些情况下, 记录的长度比块长还要大, 记录被分成若干部分存放在相邻的几个块中, 这种跨越几个块存放的记录形式, 叫做跨块记录。我们把存放于块中某记录的那部分叫段。块长和段长都是可变的。跨块记录又分为两种形式:

(1) 跨块非组合记录 VS 一个记录跨越几个相邻的块, 一个块只是一个记录的一部分, 也就是说, 一个块是由一个段组成的。

图1.10表示记录跨3个块的情况。在每段的前面有4字节的段描述字SDW, 其中存放段的长度SL (包含段描述字本身4字节) 和该段在记录中的位置 (开始、中间或最后) 信息。在段描述字的前面有4字节的块描述字BDW, 其中存放该块长度的信息。

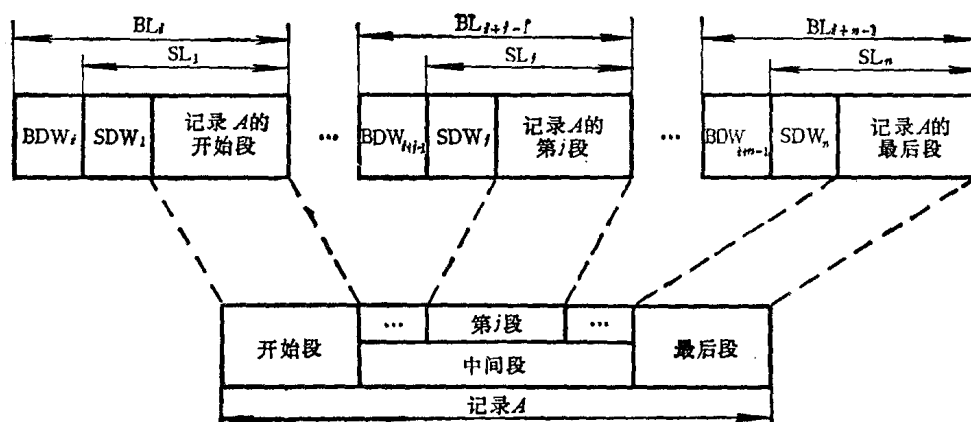


图1.10 跨块非组合记录形式
SDW_j—段j的描述字, SL_j—段j的长度。

块描述字BDW_i和段描述字SDW_j的详细内容如下:

BDW_i

字节 0~1 = BL_{i(16)}

字节 2~3 = 0000₍₁₆₎

SDW_j

字节 0~1 = SL_{j(16)}

字节 2~3 = 段控制码

$$\text{段控制码} = \begin{cases} 0000_{(16)}: & \text{全部记录存放在该块中} \\ 0100_{(16)}: & \text{记录的开始段} \\ 0200_{(16)}: & \text{记录的最后段} \\ 0300_{(16)}: & \text{记录的中间段} \end{cases}$$

例如, 若某记录的数据部分长度为1000字节, 块的最大长度为400字节, 该记录存放在三个相邻的块中, 如图1.11所示。

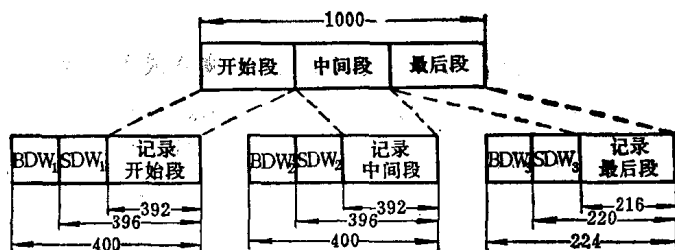


图1.11 跨块非组合记录形式实例

$BDW_1 = BDW_2 = 01900000_{(16)}$; $BDW_3 = 00E00000_{(16)}$; $SDW_1 = 018C0100_{(16)}$;
 $SDW_2 = 018C0300_{(16)}$; $SDW_3 = 00DC0200_{(16)}$ 。

(2) 跨块组合记录VBS 一个记录可以跨越几个相邻的块存放, 而一个块又可以包含一个记录的一个段或不同记录的多个段。图1.12表示这种记录的结构形式。

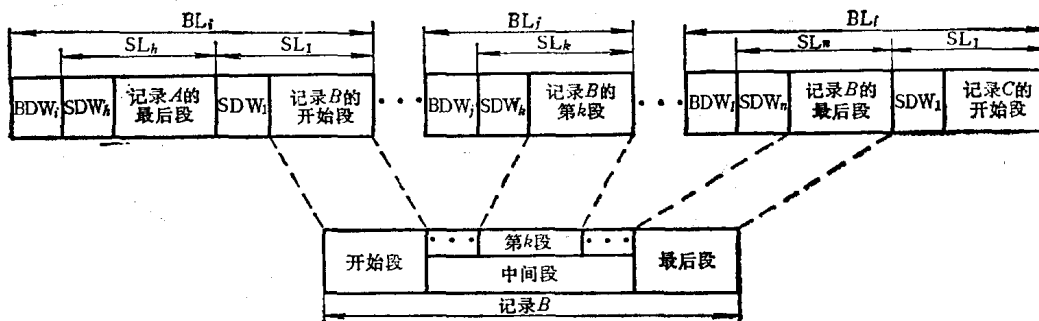


图1.12 跨块组合记录形式

例如, 三个记录的数据部分长度分别为256字节、850字节和132字节, 最大块长为400字节, 它们的存放形式如图1.13所示。

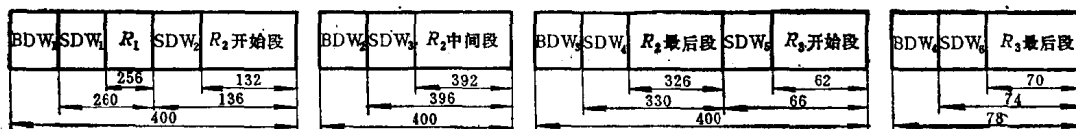


图1.13 跨块组合记录形式实例

$BDW_1 = BDW_2 = BDW_3 = 01900000_{(16)}$; $BDW_4 = 004E0000_{(16)}$; $SDW_1 = 01040000_{(16)}$;
 $SDW_2 = 00880100_{(16)}$; $SDW_3 = 018C0300_{(16)}$; $SDW_4 = 014A0200_{(16)}$; SDW_5
 $= 00420100_{(16)}$; $SDW_6 = 004A0200_{(16)}$ 。

4. 不定长记录U (图1.14)

这种记录形式和上述的几种形式不同, 它有以下特点:

- (1) 一个记录作为一个块进行处理;
- (2) 各个记录的长度可以不相等;
- (3) 在各个记录或块中, 没有表示块长或记录长的描述字。

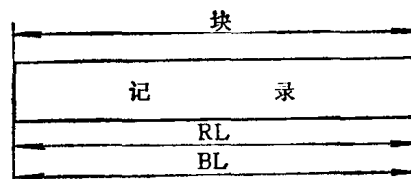


图1.14 不定长记录形式