

第8章 作业管理与用户接口

从用户使用和系统管理两个方面出发，讨论用户如何将各种作业任务提交到计算机系统中，再设法通过各级调度并进入处理器运行，然后将处理得到的运行结果进行输出。

主要内容

8.1 作业管理概述

8.2 批处理作业的管理

8.3 批处理作业调度算法

8.4 交互式作业的控制与管理

8.5 操作系统与用户的接口

8.6 Linux系统接口

8.7 本章小结

本章要点

- ◆ 了解作业的基本概念。
- ◆ 理解批处理作业和交互式作业的管理过程，以及操作系统提供给用户的各种接口。
- ◆ 掌握单道批处理作业和多道批处理作业的主要调度算法。

8.1 作业管理概述

◆ 作业

- ▣ 把用户在一次解题过程中要求计算机所做工作的集合称为一个作业。
- ▣ 作业运行步骤：编辑、编译、连接装配、运行。

◆ 作业步

- ▣ 把作业的每一个加工步骤称为一个“作业步”，具体如图。

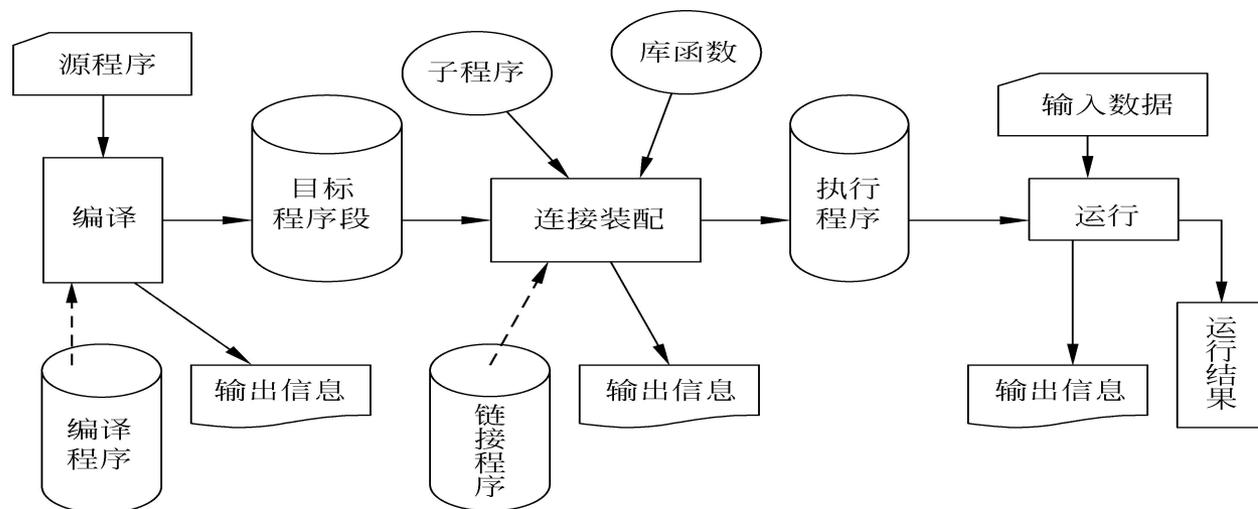


图8-1 作业控制过程

8.1 作业管理概述

◆ 作业控制方式

- 指用户根据操作系统提供的手段来说明作业加工步骤的方式。
- **批处理控制方式**
 - 用户使用“作业控制语言”对作业执行的控制意图编写成一份“作业控制说明书”，连同该作业的源程序和初始数据一同提交给计算机系统，操作系统将按照作业控制说明书来控制作业的执行。
 - 采用批处理控制方式的作业称“批处理作业”，又称“脱机作业”。
- **交互式控制方式**
 - 用户使用“操作控制命令”来表达对作业执行的控制意图。执行时，用户逐条输入命令，操作系统逐条把命令执行情况通知用户且让用户输入下一条命令，直至作业执行结束。
 - 相应作业称为“交互式作业”，又称“联机作业”或“终端作业”。

8.2 批处理作业的管理

◆ 作业管理功能

- **作业输入**：把作业装入辅存输入井中，并按照进入输入井的先后顺序形成后备作业队列的过程。
- **作业调度**：按某种调度策略选择后备作业队列中的若干作业装入主存运行的过程。
- **作业控制**：在操作系统控制下，用户如何组织他的作业并控制作业进入处理器运行的过程。

8.2.1 批处理作业输入

- ◆ 每个批处理作业输入信息包括源程序、初始数据和作业控制说明书三部分，其中说明书用作业控制语言书写。
- ◆ **作业控制语言**
 - 由若干作业控制语句组成的集合，在脱机方式下，作业控制语言将作业的控制要求编写成作业控制说明书，提供以下功能：
 - ① 作业的提交。
 - ② 控制作业和作业步的执行。
 - ③ 各种软硬件资源的使用。

8.2.1 批处理作业输入

◆ 作业控制块(JCB)

- 批处理作业存在的标志，保存于辅存中，一般应包含：作业名、作业状态、作业类别、作业优先数、作业控制方式、资源需求量、进入系统时间、开始运行时间、运行时间、作业完成时间和所需主存地址及外设种类及台数等。系统输入程序、作业调度程序、作业控制程序和系统输出程序等需要访问JCB完成相应过程。

◆ 作业表

- 由所有作业的作业控制块构成，存放在辅存中，长度固定。

◆ 批处理作业的建立

- 作业控制块的建立
- 作业的输入

8.2.2 批处理作业调度

◆ 作业调度程序

- 完成作业调度功能的控制程序，通常要完成下述工作：
 - ① 按照某种调度算法从后备作业队列中选取作业。
 - ② 为被选中的作业分配必要的主存和外设资源。
 - ③ 为选中的作业开始运行做好一切准备工作。
 - ④ 在作业运行完成或由于某种原因需要撤离系统时，作业调度程序还要完成作业的善后处理工作。

◆ 作业状态

- 作业从提交到系统直到它完成后离开系统前的整个活动可划分为若干阶段，作业在每个阶段所处的状态称为作业的状态，共分**提交状态**、**后备状态**、**运行状态**、**完成状态**四种。

8.2.2 批处理作业调度

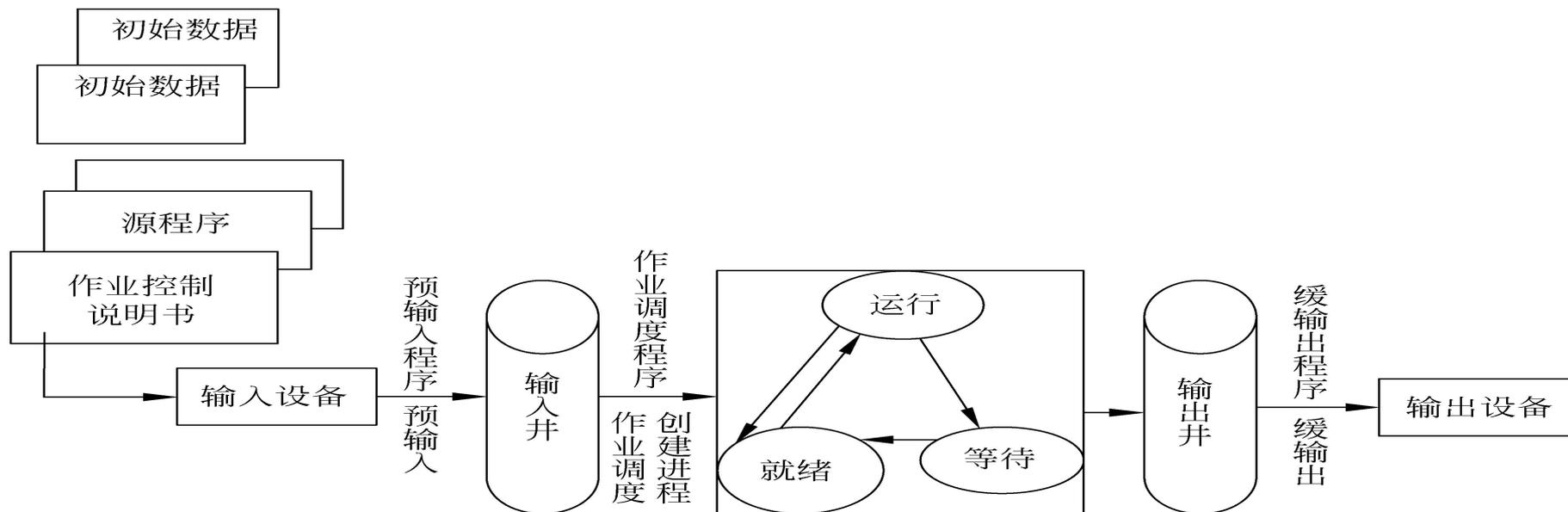


图8-2 作业状态转换图

8.2.2 批处理作业调度

◆ 作业调度的影响因素

- ① 公平性
- ② 均衡使用资源
- ③ 提高系统吞吐量
- ④ 平衡系统和用户需求

8.2.2 批处理作业调度

◆ 作业调度的主要性能指标

- ① **CPU利用率**: CPU的有效运行时间与总的运行时间之比
- ② **吞吐能力**: 吞吐能力是指单位时间内完成作业的数量。
- ③ **周转时间**: 从作业进入输入井开始, 到执行完成的时间间隔。
- ④ **平均周转时间**:

$$T = \frac{1}{n} \sum_{i=1}^{i=n} T_i$$

- ⑤ **带权平均周转时间**:

$$W = \frac{1}{n} \sum_{i=1}^{i=n} W_i$$

8.2.2 批处理作业调度

◆ 批处理作业调度阶段

- ① 高级调度(作业调度)
- ② 低级调度(进程调度)
- ③ 中级调度

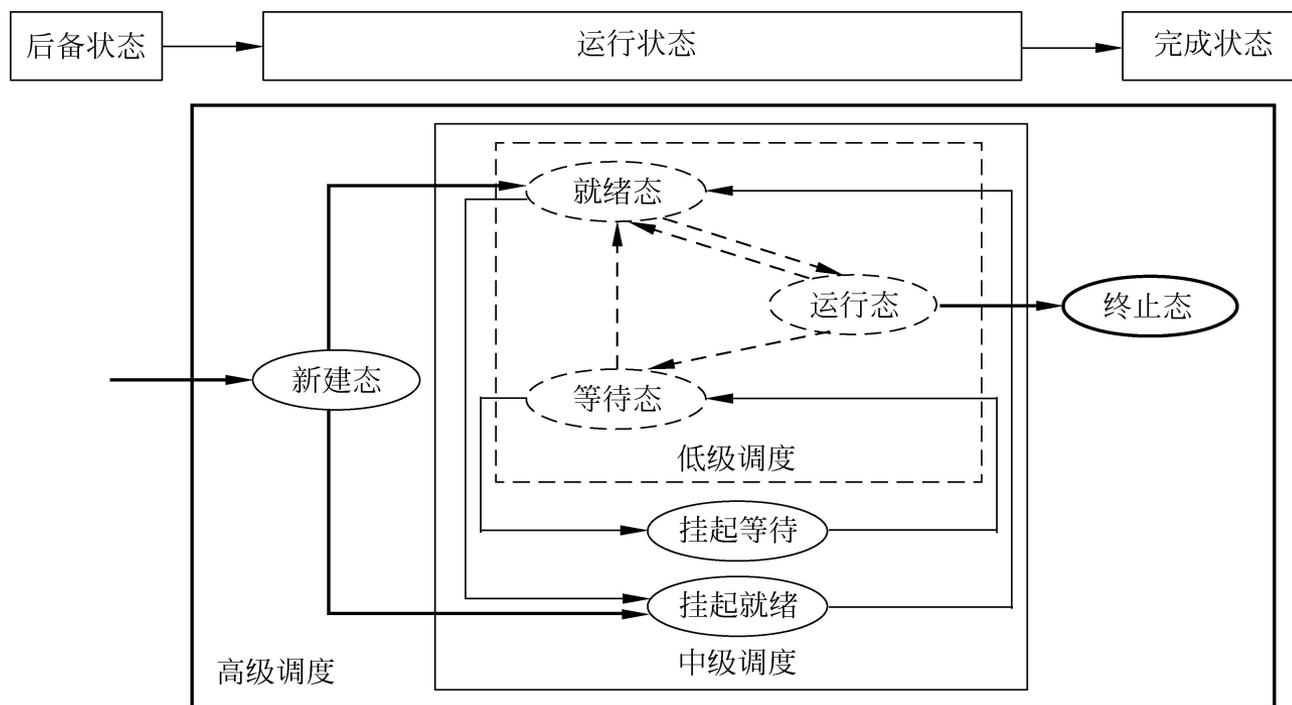


图8-3 三级调度层次关系

8.2.3 批处理作业控制

◆ 批处理作业的执行

- 当作业被选中转为运行态时，作业调度程序为其建立一个作业控制进程，由该进程具体控制作业运行。

◆ 批处理作业的终止与撤销

- 每当一个作业运行终止而被撤销后，系统又会再进行下一次作业调度，直至全部作业调度完毕。

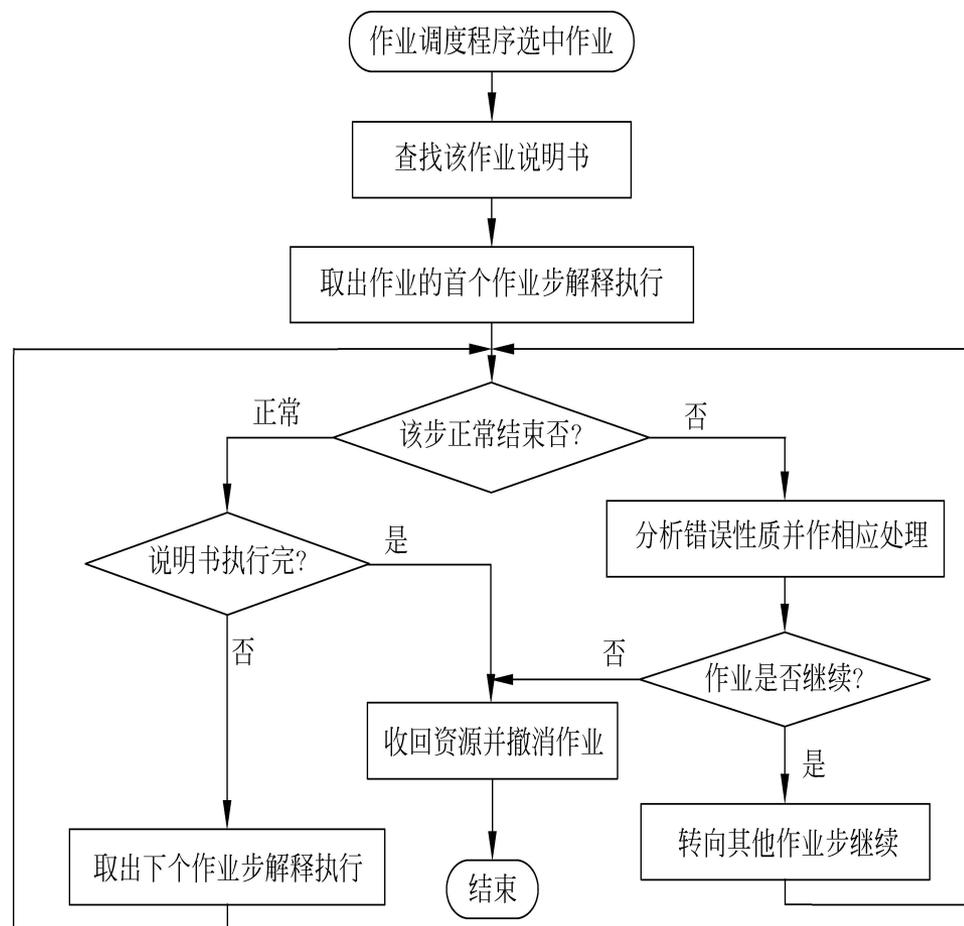


图8-4 作业的执行流程

8.3 批处理作业调度算法

- ◆ 作业调度算法是指依照某种原则或策略从后备作业队列中选取作业的方法。作业调度必须遵循一个基本原则，即系统现有的尚未分配的资源可以满足被选中作业的资源请求。
- ◆ 若现有资源既能满足作业A的要求，也能满足作业B的要求时，却不能同时满足两个作业的要求，这时到底选择哪个直接取决于作业调度算法，目标不同，选择调度算法的侧重点也就不同。
- ◆ 作业调度算法根据操作系统的性质可区别为
 - 单道批处理作业调度算法
 - 多道批处理作业调度算法

8.3.1 单道批处理作业调度算法

- ◆ 先来先服务算法
- ◆ 短作业优先算法
- ◆ 响应比最高优先算法
- ◆ 优先数调度算法
- ◆ 分类调度算法

8.3.1 单道批处理作业调度算法

◆ 先来先服务算法

- FCFS（First Come First Server），每次调度都是按照作业进入“输入井”后备作业队列的先后次序来挑选作业，优先将最先进入队列的作业调入主存，分配所需资源，创建相应的进程，放入进程就绪队列准备运行。
- FCFS算法是一种非剥夺式调度算法，容易实现，体现了公平，但效率不高，只顾及到作业等候时间，而没考虑作业要求服务时间的长短。显然这不利于短作业而优待了长作业，或者说有利于CPU繁忙型作业而不利于I/O繁忙型作业。而大多数的作业是I/O繁忙的短作业，所以先来先服务算法作为主调度算法是不常用的。

8.3.1 单道批处理作业调度算法

- ◆ **【例】** 在一个单道系统中，有4个作业要执行，每个作业进入后备队列的时间、运行时间如下表所示，假设系统当前可满足任何一个作业的请求，忽略调度和I/O时间，计算作业的平均周转时间和带权平均周转时间。

作业名	进入时间	运行时间 (min)
JOB1	8:00	120
JOB2	8:50	50
JOB3	9:00	10
JOB4	9:50	20

8.3.1 单道批处理作业调度算法

FCFS调度算法运行结果

作业名	进入时间	运行时间 (min)	开始时间	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	120	8:00	10:00	120	1
JOB2	8:50	50	10:00	10:50	120	2.4
JOB3	9:00	10	10:50	11:00	120	12
JOB4	9:50	20	11:00	11:20	90	4.5
作业平均周转时间 $T=112.5\text{min}$ 作业平均带权周转时间 $W=4.975$					450	19.9

8.3.1 单道批处理作业调度算法

◆ 短作业优先算法

- 短作业优先（Shortest Job First, SJF），进行作业调度时以作业运行时间长短作为标准进行调度，从后备作业队列中选取运行时间最短的作业装入主存运行。

SJF调度算法运行结果

作业名	进入时间	运行时间 (min)	开始时间	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	120	8:00	10:00	120	1
JOB2	8:50	50	10:30	11:20	150	3
JOB3	9:00	10	10:00	10:10	70	7
JOB4	9:50	20	10:10	10:30	40	2
作业平均周转时间 $T=95\text{min}$ 作业平均带权周转时间 $W=3.25$					380	13

8.3.1 单道批处理作业调度算法

◆ 响应比最高者优先算法

- 响应比最高者优先（Highest Response Ratio First, HRRF），既考虑了作业等待作业，又考虑了作业的运行时间，调度时对输入井中的所有作业计算其响应比，选择响应比最高的作业优先装入主存运行。
- 响应比 = (作业等待时间 + 作业运行时间) / 作业运行时间

HRRF调度算法运行结果

作业名	进入时间	运行时间 (min)	开始时间	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	120	8:00	10:00	120	1
JOB2	8:50	50	10:10	11:00	130	2.6
JOB3	9:00	10	10:00	10:10	70	7
JOB4	9:50	20	11:00	11:20	90	4.5
作业平均周转时间T=102.5min 作业平均带权周转时间W=3.775					410	15.1

8.3.1 单道批处理作业调度算法

◆ 优先数调度算法

- 为每个作业确定一个优先数，根据优先数的不同让作业排成多个队列，调度时从后备队列中优先选取资源能满足且优先数最高的作业装入主存运行。
- 作业优先数的确定原则可参照下列两条：
 - ① 对于某些时间要求紧迫的作业赋予较高的优先数；
 - ② 为了充分利用系统资源，对于I/O量大的作业给予较高的优先数，对于CPU量大的作业给予较低的优先数。

◆ 分类调度算法

- 又称均衡调度算法，根据系统运行情况和作业对资源的需求先将作业进行分类，然后作业调度程序轮流从不同的作业类中去挑选作业，尽可能使得使用不同资源的作业同时执行，资源能得到均衡利用。

8.3.2 多道批处理作业调度算法

- ◆ 在单道批处理系统中，作业运行的时候，占用所有的计算机资源，系统的资源利用率不高。
- ◆ 现代操作系统都支持多道程序设计技术，可同时让多个满足装入要求的作业同时装入计算机系统，从而可以让这些作业分别使用处理器运行，或者访问外部设备完成数据的输入和输出，这样就可大大提高系统资源的利用率。

8.3.2 多道批处理作业调度算法

- ◆ **【综合实例1】** 若某系统采用可变分区方式管理主存中的用户空间，供用户使用的最大主存空间为**100K**，主存分配算法为最先适应分配法，系统配有**4台**磁带机，一批初始作业如下表所示。

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	磁带机需求量 (台)
JOB1	8:00	40	35	3
JOB2	8:10	30	70	1
JOB3	8:15	20	50	3
JOB4	8:35	10	25	2
JOB5	8:40	5	20	2

8.3.2 多道批处理作业调度算法

- ◆ 分析：由于主存采用可变分区分配方式，且允许移动作业汇聚空闲分区，则**只要主存空闲空间总量大于作业需要，主存就可满足作业对主存的需求**(必要时通过移动技术汇聚)；由于磁带机采用静态分配方式，则**只有系统剩余的磁带机数大于等于作业的申请数才能满足**。这两个条件是作业调度必须满足的前提条件。
- ◆ 另外，由于该系统采用多道程序设计技术，因此，每次作业调度时若有多个作业同时满足调度条件，则可同时选中多个作业装入主存，然后创建相应进程插入到就绪队列中。由于忽略了设备工作时间和系统进行调度所花的时间，则可认为一个作业一旦获得处理器后连续工作，可在给出的作业运行时间内执行完并让出处理器给下一个作业执行，这样作业调度可选择下一个批处理作业。

8.3.2 多道批处理作业调度算法

◆ FCFS调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	磁带机需求量 (台)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	40	35	3	8:00	8:00	8:40	40	1
JOB2	8:10	30	70	1	8:40	8:40	9:10	60	2
JOB3	8:15	20	50	3	9:25	9:25	9:45	90	4.5
JOB4	8:35	10	25	2	8:40	9:10	9:20	45	4.5
JOB5	8:40	5	20	2	9:10	9:20	9:25	45	9
作业平均周转时间 $T=56\text{min}$ 作业平均带权周转时间 $W=4.2$								280	21

8.3.2 多道批处理作业调度算法

◆ SJF调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	磁带机需求量 (台)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	40	35	3	8:00	8:00	8:40	40	1
JOB2	8:10	30	70	1	8:45	8:55	9:25	75	2.5
JOB3	8:15	20	50	3	9:25	9:25	9:45	90	4.5
JOB4	8:35	10	25	2	8:40	8:45	8:55	20	2
JOB5	8:40	5	20	2	8:40	8:40	8:45	5	1
作业平均周转时间 $T=46\text{min}$ 作业平均带权周转时间 $W=2.2$								230	11

8.3.2 多道批处理作业调度算法

◆ HRRF调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	磁带机需求量 (台)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:00	40	35	3	8:00	8:00	8:40	40	1
JOB2	8:10	30	70	1	9:05	9:15	9:45	95	3.17
JOB3	8:15	20	50	3	8:40	8:40	9:00	45	2.25
JOB4	8:35	10	25	2	9:00	9:05	9:15	40	4
JOB5	8:40	5	20	2	9:00	9:00	9:05	25	5
作业平均周转时间 $T=49\text{min}$ 作业平均带权周转时间 $W=3.08$								245	15.42

8.3.2 多道批处理作业调度算法

- ◆ **【综合实例2】** 有一个多道程序设计系统，仍采用可变分区方式管理主存中的用户空间，但不允许移动已在主存中的任何作业。假设用户可使用的最大主存空间为100K，主存分配算法为**最先适应分配法**，作业序列如表所示。

作业名	进入时间	运行时间 (min)	主存需求量 (KB)
JOB1	8:06	42	55
JOB2	8:20	30	40
JOB3	8:30	24	35
JOB4	8:36	15	25
JOB5	8:42	12	20

8.3.2 多道批处理作业调度算法

◆ FCFS调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:06	42	55	8:06	8:06	8:48	42	1
JOB2	8:20	30	40	8:20	8:48	9:18	58	1.93
JOB3	8:30	24	35	8:48	9:18	9:42	72	3
JOB4	8:36	12	25	9:18	9:54	10:09	93	6.2
JOB5	8:42	15	20	8:48	9:42	9:54	72	6
作业平均周转时间 $T=67.4\text{min}$ 作业平均带权周转时间 $W=3.63$							337	18.13

8.3.2 多道批处理作业调度算法

◆ SJF调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:06	42	55	8:06	8:06	8:48	42	1
JOB2	8:20	30	40	8:20	8:48	9:18	58	1.93
JOB3	8:30	24	35	9:18	9:45	10:09	99	4.13
JOB4	8:36	12	25	8:48	9:30	9:45	69	4.6
JOB5	8:42	15	20	8:48	9:18	9:30	48	4
作业平均周转时间 $T=63.2\text{min}$ 作业平均带权周转时间 $W=3.13$							316	15.66

8.3.2 多道批处理作业调度算法

◆ HRRF调度算法运行结果

作业名	进入时间	运行时间 (min)	主存需求量 (KB)	作业调度	进程调度	完成时间	周转时间 (min)	带权周转时间
JOB1	8:06	42	55	8:06	8:06	8:48	42	1
JOB2	8:20	30	40	8:20	8:48	9:18	58	1.93
JOB3	8:30	24	35	9:18	9:45	10:09	99	4.13
JOB4	8:36	12	25	8:48	9:18	9:33	57	3.8
JOB5	8:42	15	20	8:48	9:33	9:45	63	5.25
作业平均周转时间 $T=63.8\text{min}$ 作业平均带权周转时间 $W=3.22$							319	16.11

8.4 交互式作业的控制与管理

◆ 交互式作业的控制

◆ 交互式作业的控制方式采用交互式控制方式。对于交互式作业来说，用户通过输入操作控制命令直接控制作业的执行。

◆ 1.操作控制命令

▣ 命令的一般格式为：命令名 参数1， 参数2， ---， 参数n

◆ 2.菜单方式

◆ 3.视窗方式

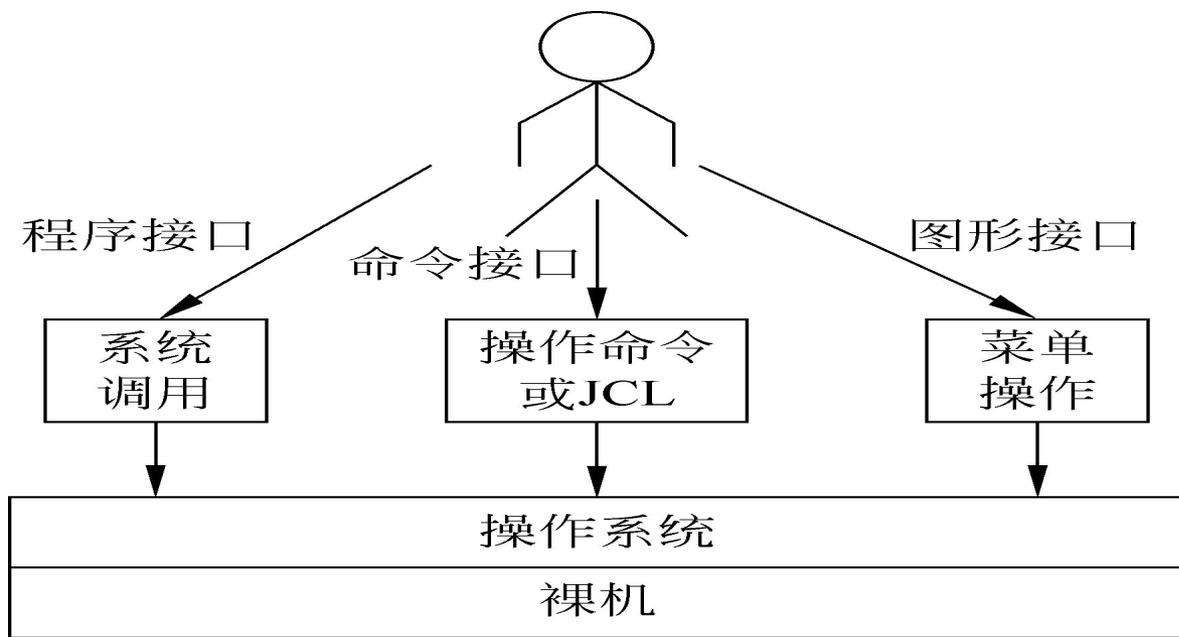
8.4 交互式作业的控制与管理

◆ 交互式作业的管理

- ◆ 一般地，控制终端作业的执行大致分成四个阶段，即：
 - ① 终端连接
 - ② 用户注册
 - ③ 作业控制
 - ④ 用户退出。

8.5 操作系统与用户的接口

- ◆ 操作系统是用户与计算机系统之间的接口，用户通过用户接口向操作系统提出服务请求，操作系统则通过用户接口将服务结果返回给用户。
- ◆ 现代操作系统向用户提供了三种典型接口。



8.5 操作系统与用户的接口

◆ 命令接口(作业级接口)

- 用户通过输入设备发出一系列操作控制命令告诉操作系统执行所需的功能。可将命令接口又分成联机 and 脱机命令接口。

- **1. 联机命令接口(也称交互式命令接口)**

- 由一组键盘操作命令及命令解释程序组成，用于完成联机作业的控制，可分成文件、目录、磁盘操作类和系统访问、权限管理等。
- 命令格式：Command arg1,arg2,...,argn,[option 1,...option k]

- **2. 脱机命令接口(也称批处理命令接口)**

- 利用JCL中命令编写成作业控制说明书来完成脱机作业的自动控制执行，用户无法在计算机上进行干预，只能等待作业执行结束。

8.5 操作系统与用户的接口

◆ 程序接口

- 由若干系统调用命令组成。用户在运行程序中，利用系统调用在程序这一级提出资源申请和功能服务，从而获得操作系统的底层服务，使用或访问系统的各种软硬件资源。

□ 1. 系统调用

- 是操作系统为用户程序提供的一种服务，由若干不同功能的子程序组成，用户程序在执行中可以调用这些子程序，简称“系统调用”。“系统调用”是在核心态下执行的程序。

□ 2. 系统调用分类

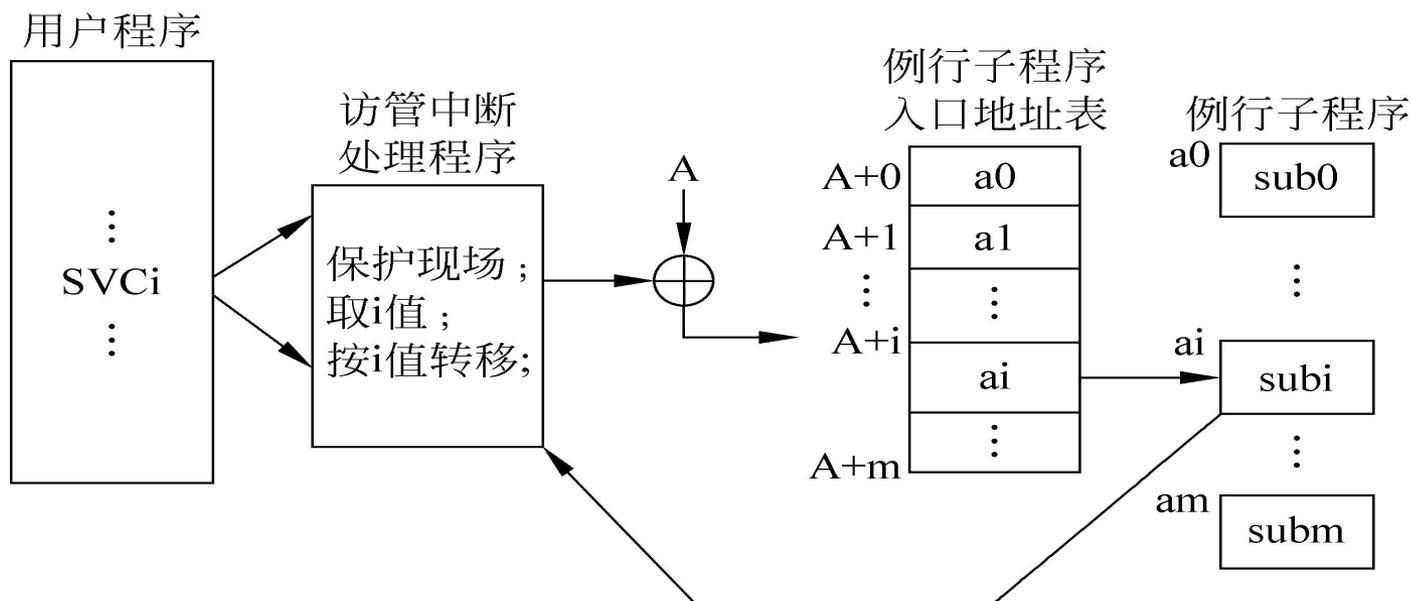
- 1)文件管理类、2)进程控制类、3)存储管理类、4)设备管理类、5)程序控制类、6)信息维护类等。

8.5 操作系统与用户的接口

◆ 程序接口

□ 3. 系统调用执行过程

- 设置系统调用参数
- 系统调用命令的一般性处理
- 系统调用具体过程



8.5 操作系统与用户的接口

◆ 图形接口

- 现代操作系统提供的接口在形式上和功能上发生了很大的变化，已经不再仅局限于常用的命令接口和程序接口了，能提供更友好的、更直观的、易懂的、图文并茂的、且具有一定智能的运行环境。
- 1.菜单驱动方式
- 2.图形驱动方式

8.7 本章小结

- ◆ 批处理作业的管理主要包括作业输入、作业调度和作业控制过程。
- ◆ 用户通过预输入过程输入并保存批处理作业到输入井的后备作业队列中；整个作业调度阶段需要经历三个阶段，即高级调度、低级调度和中级调度；通过详细对比批处理作业的**FCFS**、**SJF**、**HRRF**等调度算法实例，可以发现SJF调度算法综合性能最佳。
- ◆ 交互式作业的管理过程以交互式方式控制终端作业的执行，历经终端的连接、用户注册、作业控制和用户退出四个阶段。
- ◆ 为了使用户能够快速有效地访问计算机系统资源，操作系统向用户提供了三种典型接口：作业级接口、程序级接口和图形接口。