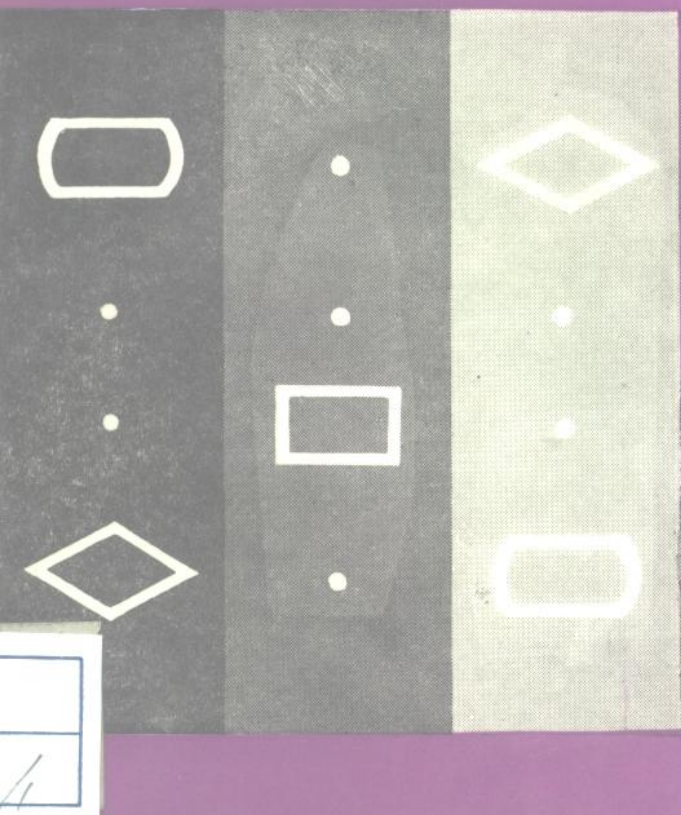


计算机基础自学丛书

操作系统

金千方 史美林 编著



机械工业出版社

计算机基础自学丛书

操作系统

金千方 史美林 编著

机械工业出版社

操作系统是现代计算机系统中必不可少的基本系统软件。本书介绍操作系统的基本概念和一般原理。全书分为九章，第一章概述操作系统的作用、功能和基本类型。第二章介绍与硬件的预备知识。第三章介绍操作系统的对外接口。第四章至第八章分别介绍进程管理、处理机管理、存贮管理、文件系统和设备管理。第九章介绍一个操作系统实例——UNIX 操作系统。

本书供从事计算机应用的科技人员阅读，也可作为计算机有关专业的教学参考书。

操 作 系 统

金千方 史美林 编著

责任编辑：邱锦来 版式设计：罗文莉
封面设计：方 芬 责任校对：熊天荣

机械工业出版社出版（北京阜成门外百万庄南街一号）

（北京市书刊出版业营业登记证出字第 117 号）

机械工业出版社印刷厂印刷

新华书店北京发行所发行·新华书店经售

开本 850×1168¹/₃₂·印张 8¹/₄·插页 1·字数 217 千字

1989年5月北京第一版·1989年5月北京第一次印刷

印数 0,001—3,070·定价：7.70 元

ISBN 7-111-00601-1/TP·42

出版说明

计算机的广泛应用极大地提高了人类认识和改造客观世界的的能力，特别是微电子及大规模集成电路技术的发展和微型计算机的广泛应用，为各国工业技术的改造提供了重要而有效的途径。一场以解放人类智力为标志的新技术革命开始了。其规模和深度都将是空前的，不仅涉及到科学、工业、经济、教育等各个领域，同时也将会引起社会结构和生活方式的改变。

预计 90 年代将出现以知识处理为基础的智能技术与智能系统，并得到广泛的应用。例如，智能机器人，智能信息处理系统，智能计算机辅助设计系统，智能办公室自动化系统等。它将开创新的计算机应用领域。传统的数据处理、计算机控制等的应用更加向前推进。

为了适应新技术革命的需要，帮助各个领域的科技人员进一步普及和提高计算机基础科学知识，使计算机得到广泛的更有效的应用。我们特为全国机电专业及计算机应用专业中专以上文化水平的读者编写了这一套丛书。对广大读者只要具有基本的数字电路及逻辑设计基础知识，均可自学这套丛书。“丛书”是着眼于计算机的应用，搜集归纳了国内外大量的资料，并在长期的教学和科研工作基础上进行编写的。丛书力求做到选材上的先进性、系统性和完整性。由浅入深、全面系统地阐述了计算机科学的基础知识，同时对计算机的主要应用领域的先进技术也作了必要的介绍。

“丛书”编写者是清华大学计算机科学与技术系长期在第一线从事教学、科研工作的副教授、高级工程师。丛书由清华大学周远清主编，叶乃萃为全套丛书的编写、定稿和出版作了大量的工作。

这套丛书共分九册。

IV

第一册计算机概论 阐述计算机的体系结构及各部分工作原理。

第二册高级语言程序设计 介绍 PASCAL 语言及 C 语言的基本结构以及语言的使用。

第三册 PDP-11 机器语言与汇编语言程序设计 介绍 PDP-11 机的基本组成及编程技术。

第四册数据结构 阐述非数值程序中有关结构以及排序和查找等问题。

第五册操作系统 阐述操作系统的有关基本概念以及管理系统资源，提供用户应用环境。

第六册汉字微型计算机与汉字识别 介绍汉字微型计算机系统的组成、输入编码方案、汉字处理和识别技术以及汉字微型计算机系统的应用。

第七册计算机控制系统及应用、主要分析计算机控制系统的原理、设计方法（连续、离散、状态空间），通道接口及应用实例。

第八册计算机辅助图形设计 介绍计算机辅助设计基础——二维三维图形设计的基本算法以及交互式图形软件的设计。

第九册办公室自动化基础 阐述数据库的基本概念、办公室自动化基础知识及技术。

在编写这套丛书时，我们既注意到它的系统性和完整性，同时又使各册之间有相对的独立性，以便于读者能够根据自己从事工作的领域以及自己的兴趣，阅读全套丛书或选读其中几册。

对于需要使用语言进行程序设计的初学者，可以首先阅读第一、二、三册。在有关领域中从事计算机应用的科技人员，可以分别选读第六、七、八、九册。对于需要系统深入了解和掌握计算机科学知识的读者，建议进一步阅读第四、五册。

书中所选实例均经过上机验证，有的还附有习题，希望读者能完成练习，并上机进行实践。

这套丛书既可作广大读者的普及读物和科技参考书，又可作

为大学教科书。我们希望，这套丛书能为我国计算机的普及与应用起到一定的作用。

由于编者水平有限，书中定有不当之处和错误的地方，欢迎读者批评指正。

清华大学计算机科学与技术系
计算机基础自学丛书编写组

前 言

操作系统是现代计算机系统中必不可少的基本系统软件。它控制和管理整个计算机系统中各种硬件、软件资源，并为用户使用计算机提供一个方便灵活、安全可靠的工作环境。操作系统原理不仅是计算机有关专业的必修课程也是一般从事计算机应用的人员所不可缺少的。

从 1978 年开始，编者在清华大学计算机科学与技术系讲授操作系统原理课程。融会近十年教学实践经验，并分析和研究了若干个操作系统，又考虑到本丛书的主要对象是非计算机专业但从事计算机应用的技术人员，此次成书时，在阐明计算机操作系统的基本概念、一般工作原理和实现方法的同时，力求深入浅出。在取材上，力求基本、实用，以利初学。在编排上，既注意系统性，又在有关章节适当补充介绍一些相关课程的基础知识，便于读者掌握和理解。

全书共分九章，第一章概括介绍什么是操作系统、操作系统的分类及功能。第二章从操作系统的角度介绍必备的硬件知识，主要是通道和中断的概念。第三章介绍操作系统的对外接口，即从使用的角度来看操作系统的外貌，说明用户如何通过操作系统来使用计算机。第四章阐述进程的基本概念、状态、控制及进程间通信。第五至八章介绍各种资源的管理，它们是存贮管理、处理机管理、文件系统和设备管理。第九章介绍一个操作系统实例——UNIX 操作系统，目的是帮助读者加深理解前面各章介绍的若干主要概念，并对操作系统能有一个完整的认识。

本书第一至六章由金千方编写，第七至九章由史美林编写。全书由王诚初审，最后由周远清总审，在此我们表示诚挚的感谢。

由于编者水平所限，书中难免会有不少错误和不妥之处，恳请广大读者予以批评指正。

目 录

第一章 引论	1
§ 1-1 什么是操作系统	1
§ 1-2 操作系统分类	2
§ 1-3 操作系统功能	8
第二章 预备知识	12
§ 2-1 引言	12
§ 2-2 通道机构	12
§ 2-3 中断与捕俘	14
第三章 操作系统对外接口	23
§ 3-1 引言	23
§ 3-2 作业组织和作业控制	24
§ 3-3 广义指令	34
第四章 进程管理	39
§ 4-1 进程的概念	39
§ 4-2 进程的描述	44
§ 4-3 进程的状态和转换	46
§ 4-4 进程通信	47
§ 4-5 进程控制	56
§ 4-6 死锁	57
第五章 处理机管理	66
§ 5-1 引言	66
§ 5-2 分级调度	67
§ 5-3 调度目标和调度准则	71
§ 5-4 调度算法	72
第六章 存贮管理	82
§ 6-1 引言	82

§ 6-2	单道环境下的存贮管理方法	87
§ 6-3	分区存贮管理	89
§ 6-4	内存扩充技术	101
§ 6-5	页式存贮管理	104
§ 6-6	段式存贮管理	118
§ 6-7	段页式存贮管理	126
§ 6-8	各种存贮管理方法的比较	129
第七章	文件系统	130
§ 7-1	概述	130
§ 7-2	文件的逻辑结构、物理结构和存取方法	133
§ 7-3	文件目录及其管理	139
§ 7-4	文件的共享和存取控制	146
§ 7-5	文件存贮空间的管理	149
§ 7-6	文件的使用	151
§ 7-7	文件系统层次模型	153
第八章	设备管理	156
§ 8-1	引言	156
§ 8-2	通道技术	159
§ 8-3	缓冲技术	162
§ 8-4	设备分配和管理	164
第九章	操作系统实例——UNIX 操作系统简介	171
§ 9-1	概述	171
§ 9-2	UNIX 操作系统的进程管理	178
§ 9-3	UNIX 操作系统的文件系统	196
§ 9-4	UNIX 操作系统的 I/O 系统和设备管理	213
§ 9-5	UNIX 系统的 Shell 命令语言	244
参考文献	256

第一章 引 论

§ 1-1 什么是操作系统

计算机发展到今天，从巨型计算机到个人计算机，几乎都配备有一种或多种操作系统。如果要用用户去使用一台没有操作系统的现代计算机，那是难以想象的。即使他只要求做一些最简单的计算，也是非常困难的。例如，他如何把自己的程序和数据输入到机器的内存贮器中，又如何把计算结果输出来呢？这在普通的计算机用户看来，是件再简单不过的事了。只需在终端键盘上把程序和数据敲进去，用打印机把结果打出来即可。其实，这里已经把输入输出工作交给了操作系统去完成。如果没有操作系统，就需要用户自己去编制输入输出处理程序。为此，用户必须了解输入输出设备的物理特性，以及使其工作的许多细节。这对一般用户来说，是相当陌生和厌烦的。

那么，什么是操作系统呢？简单地说，操作系统是负责控制和管理计算机系统中所有硬件和软件的一些程序模块。

通常我们所说的计算机系统包括：

系统硬件，如 CPU (Central Processing Unit)、存贮器、输入/输出设备等。

系统软件，如操作系统、各种语言处理程序、系统实用程序等。

一台没有任何软件支持的计算机称之为裸机(Bare Machine)而现在实际呈现在用户面前的计算机系统是经过若干层软件改造的计算机。操作系统和硬件与软件的关系可粗略地用图 1-1 来表示。

由图 1-1 可看出，操作系统是用户和计算机之间的接口（又称作界面），用户是通过操作系统来使用计算机的。

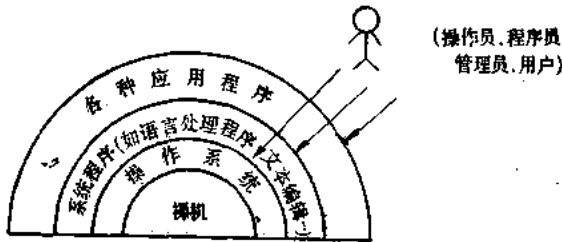


图1-1 操作系统与硬、软件的关系

引入操作系统的目的有两方面：

1. 从系统管理人员的观点来看，引入操作系统是为了合理地去组织计算机工作流程，使计算机系统硬件、软件资源能为多个用户所共享，最大限度地发挥计算机系统资源的效率。

2. 从用户的观点来看，引入操作系统是为了给用户使用计算机提供一个良好的界面，以使用户无需了解许多有关硬件和系统软件的细节，就能方便地、灵活地使用计算机。

这两个目标既有联系而又有矛盾。通常，要使用户用起来方便灵活，相对说来，计算机系统的资源利用率就会受到影响。操作系统的任务就是要在使用简便和代价低之间尽可能提供最佳平衡。

这里还需要说明的是，在本书中，我们把操作系统定义为那些负责控制和管理系统中所有硬件和软件的程序模块。这些模块有时也称作监督程序、管理程序等等，但是我们常常也会看到有人把一些系统实用程序，如连接装配程序、编辑程序、辅助调试程序、库程序以及各种语言处理程序等等也作为操作系统的成分。这时，我们可以把后面这些软件看成是操作系统的子系统，它们同样也由操作系统来控制和管理。并且它们的运行环境和普通用户程序是一样的。

§ 1-2 操作系统分类

早期的计算机没有操作系统。用户直接用机器语言编制程序，并在上机时独占全部资源。用户既是程序员，又是操作员，

上机完全是手工操作。随着计算机硬件、软件的发展，操作系统逐渐形成和完善，产生了各种各样的操作系统。按照对作业处理的不同方式，可以将操作系统划分成三种基本类型：

一、批处理系统 (Batch Processing System)

在一般计算中心的计算机上常常配置批处理类型的操作系统，这类系统是以提高系统资源的效率和作业吞吐量（单位时间里处理作业的个数）为其设计目标。

所谓批处理是指各用户用操作系统提供的作业控制语言来描述自己对作业运行的控制意图，并将这些控制信息连同自己的程序和数据一起提交给操作员。操作员启动有关程序，将一批作业输入到计算机外存（如磁带、磁鼓、磁盘）上，然后由操作系统去控制，让各作业运行并输出运行结果。由于各用户作业进入计算机系统后，不再对它进行人工干预，而是由操作系统按照用户提供的有关作业控制信息来对作业的运行进行控制和管理，从而提高了系统的效率。

（一）早期批处理系统

早期的批处理系统是单道的，即一个作业被调入内存运行直到运行完毕，或是因某种原因执行不下去而不得不终止退出时，操作系统才去调下一个作业来运行。也就是说，一个作业在运行时，它独占全机所有资源。如果前一个作业未退出系统，下一个作业就不能开始运行。对于这种单道批处理系统，虽然一批作业能够自动地一个接一个地顺序去运行，节省了作业交接时人工操作的时间，解决了人工操作和机器运行速度不匹配这一矛盾，但整个系统的效率还是很低。主要原因在于处理机和输入/输出设备串行工作，当作业在进行输入/输出操作时，处理机空闲，在等待输入/输出操作完成，因此浪费了大量的处理机时间。特别是随着处理机速度大幅度提高，处理机和输入/输出设备间的速度不匹配的矛盾也就更为突出。为解决这一矛盾，出现了多道程序设计技术。

（二）多道批处理系统

所谓“多道”是指在内存中同时保持若干道作业，并使这些作业交替地执行。例如，当其中某一道程序在等待外设传输时，就让另一道程序运行，以尽可能减少主机空闲等待时间。如果这些作业搭配得合理，就能保证主机和外设高度并行工作。当然，这还需要有硬件的支持。这种硬件就是 60 年代初期出现的通道和中断机构。我们将在下一章进一步说明。

“多道”和“批处理”相结合，就构成了多道批处理系统。它兼有“多道”和“批处理”两者的特征。其工作流程大致如下：各用户将自己的作业及有关作业的资源要求和运行时的控制意图等信息一并交给系统操作员；操作员随时把一批批作业提交给操作系统。这样，在系统的外存贮器中，可堆积大量的等待运行的后备作业。操作系统根据一定的调度策略，从后备作业中选出若干个作业并调入内存，按多道组织它们运行。这些作业共享系统资源。当某个作业因某种原因（如等待输入输出完成）暂时工作不下去时，系统就转去执行另一个作业，即让这些作业交替执行，直至某个作业工作完成。这时，系统输出该作业的运行结果，并回收该作业所占用的全部资源。于是操作系统又去调另一个作业投入运行，使得作业不断地流入系统，经处理后又退出系统，形成一个源源不断的作业流。

在多道批处理系统中，由于系统资源为多个作业所共享，再加之系统的批处理工作方式是作业之间自动过渡，并在运行过程中用户不能干预自己的作业，从而大大提高了系统资源的利用率和作业吞吐量。现代计算机上的批处理系统，绝大多数都是多道批处理系统。例如：M-150F 机上的 OSIV/F2 操作系统，就是多道批处理类型的，最多可同时容纳八道作业。当然多道程序设计技术并不只限于用在批处理系统中。下面介绍的分时系统、实时系统，亦可采用多道程序设计技术。

附带要说明的是在上述的多道系统中，通常认为其计算机系统只配置一台 CPU，此时我们所说的多个作业在系统中同时运行是指从宏观上来看它们在“同时运行”，实际上是在交

替执行。有一种称之为多重处理的系统，它配置多台 CPU，因而该系统能够真正的同时执行多个程序。不要把一般多道(Multiprogramming)系统和多重处理(Multiprocessing)系统相混淆。一般说来，要想有效地使用多重处理系统，就必须采用多道程序设计技术，而多道程序设计则不一定要求有多重处理系统的支持。多重处理系统的硬件结构可以是多种多样的。它可以是由几个独立的计算机系统通过通信链路相互连接起来，亦可以是多个处理机共享一个公用内存。多重处理系统比起单处理机系统来说，虽然增加了硬件设施，但却换来了提高系统吞吐量、可靠性、计算能力及并行能力这些好处。

二、分时系统 (Time-sharing System)

在一个计算机系统上，借助于通信线路接入多个终端，各个用户通过自己的终端将作业送入计算机，计算机又通过各终端向用户报告其作业运行情况。计算机分时地轮流为各个终端用户服务，这就构成了一个分时系统。

分时系统的主要特点是它和用户交互会话的工作方式。分时系统中的每个用户，可以随时通过自己的终端向计算机提出各种服务请求，系统必须在比较短的时间内（通常是在两三秒以内）对此作出响应。在用户看来，这台计算机就好像仅供他一人使用一样。这是因为对分时系统来说，它虽然接入了多个终端，同时要为这些终端用户服务。但由于终端用户是手工操作，用键盘输入自己的请求命令，这相对于处理机的内部处理的速度来说实在是太慢了。因此，分时系统将处理机的时间划分成很短的时间片，轮流分配给各个终端用户，以便对终端的服务请求作出及时的响应。每个用户每次只是暂时的使用一下处理机，但很快又能轮上他，所以他们都觉得好像是自己独占了计算机。

分时系统的这种交互会话的工作方式给用户带来许多好处。首先，它使用方便，特别是对远程终端的用户，不必将自己的作业交付给机房，在自己的终端上就可以上机了。其次，用户可以在程序动态运行的情况下，对他的程序加以控制。这就有可能大大

加快程序调试的过程。他可以随时去检查和修改自己程序中的错误，而不像在批处理系统中那样，投入一个作业常常要在几小时甚至几天以后才能拿到结果。所以分时系统常常是被用户用来进行程序开发的工作。分时系统还有一个优点是为用户之间进行合作提供了方便的条件，他们可以通过文件系统去交换信息而进行合作。

最早的分时系统CTSS(Compatible Time-Sharing System)是在1963年由MIT研制成功的。当前十分流行的UNIX操作系统也是一个分时系统。

三、实时系统

随着计算机应用的发展，出现了实时系统。它主要用于实时控制和实时信息处理两个方面。实时控制是指计算机通过特定的外围设备，如传感器等，与被控制对象发生联系。这些外围设备将采集到的物理量经过模数转换送到计算机，计算机对其加工处理后，再经数模转换变成控制信号，通过执行机构去直接控制有关的对象。例如，对发电厂锅炉设备内各种参数的控制，可由各种量测装置测出锅炉内蒸气的压力、温度、负荷大小等，经模数转换后送入计算机，由计算机进行实时计算后确定应送进炉膛的燃料和空气量，再经数模转换后通过有关的执行机构去控制阀门的开启。再如用于军事指挥的防空警戒系统，也是一个实时系统。实时系统广泛用于生产过程控制、军事指挥、医疗等各个领域。

实时系统用于实时信息处理方面的有飞机订票系统、情报检索系统、银行系统等等。在这些系统中，用户也是通过终端提出服务请求，系统完成服务后通过终端回答用户。从形式上看，用户使用它也是采用一种人-机会话工作方式，但它不同于一般分时系统。实时信息处理系统对用户提供的服务只限于该系统所管理的实时信息范围，如飞机订票系统，用户只能提出某月某日的某次航班有否空位，票价多少等等这一类问题。而分时系统提供给用户的是用于一般目的的计算机系统，用户主要是用来进行程序开发或是运行专门设计的应用程序。

实时系统的主要特点是提供即时响应。系统必须保证对这些实时信息的分析和处理的速度比这些信息进入系统的速度要快，而且系统本身安全可靠。因为象生产过程的实时控制系统、导弹发射系统等信息处理的延误或丢失往往会带来不堪设想的后果。为此，常常需要采取相应的硬件（如双机）和软件措施，以提高系统的可靠性。与批处理系统、分时系统相比，实时系统的资源利用率一般较低。

四、通用操作系统

上面所介绍的批处理系统、分时系统和实时系统是操作系统的三种基本类型。在此基础上又发展了以多种类型操作为特征的通用操作系统。它可以同时兼有批处理、分时、实时处理和多重处理的功能，或其中两种以上的功能。例如，将实时处理和批处理相结合构成实时批处理系统。在这样的系统中，它首先保证及时响应外部发来的实时信号，并在严格的时间范围内完成对实时设备的控制处理。但一般说来，实时信号并不是时时刻刻都有的。因此，系统除了管理实时控制任务外，还可能有空余的处理机时间。系统可利用这个时间来处理不需要及时响应的批量作业。通常把系统中保证优先处理的实时任务称为前台作业，而把插空去处理的批量作业称为后台作业。例如：国产 DJS-140 机上的 RDOS 操作系统，是一种前/后台两道实时磁盘操作系统。前台用于实时任务处理，后台可处理批量作业。

从 60 年代中期，国外开始研制一些大型的通用操作系统。这些系统功能齐全，可适应各种应用范围和操作方式变化多端的环境。当然，系统本身也就很庞大。研制它们付出了巨大的代价。而且由于系统过于复杂和庞大，在解决其可靠性、可维护性和可理解性等方面都遇到很大困难。相比之下，UNIX 系统却是一个例外。它首先建立一个精悍的核心，这就是 UNIX 操作系统。它只有一千多行汇编代码和一万多行 C 语言代码。尽管它规模很小，但其功能却足以与许多大型的操作系统相媲美。因此，它能支持该系统核心以外的庞大的软件系统。从而使 UNIX 系统成为

目前国际上十分流行的一个系统（详见第九章操作系统实例）。

五、网络操作系统

随着计算机技术和通信技术的发展和密切结合，产生了计算机网络。通俗地讲，计算机网就是把分布在不同地理位置的计算机系统，通过通信设备和通信线路连接起来，以实现更加广泛的硬件资源、软件资源的共享。根据网络作用的地理范围不同，可以将网络分成两大类：一类是远程计算机网，其地理范围从几百公里到几千公里，甚至可达上万公里；另一类是局部计算机网，其地理范围通常是几公里，例如一所大学、一个工厂或一个企业内的计算机可连成局部计算机网。

一个计算机系统入网后，不但大大地扩大了本机用户可用资源的范围，而且也使它自己的用户范围从本机用户扩大到网际用户。因此，它就要求网络环境中的操作系统既要为本机用户提供简便有效地使用网络资源的手段，而且又要为网络用户提供使用本机资源的服务。所以在网络环境下的操作系统，除了具有普通操作系统的功能外，还应增加网络管理模块，其主要功能是支持终端与计算机、计算机与计算机和计算机与网络之间的通信，以及提供各种网络服务。

§ 1-3 操作系统功能

如前所述，操作系统的职能是控制和管理计算机系统中所有的硬、软资源，并且为用户使用计算机提供一个良好的工作环境，计算机系统的主要硬件资源有处理机、存贮器和输入/输出设备。下面我们从资源管理的观点分五个方面来说明操作系统的基本功能。

一、处理机管理

在单道情况下，处理机为一个作业所独占，对处理机的管理十分简单。但在多道环境中，要组织多个作业同时运行，因此要由处理机管理模块来解决处理机的分配策略、分配实施以及资源的回收等问题。正是由于操作系统对处理机的管理策略不同，其