新世纪高职高专实用规划教材

计算机系列

操作系统教程

（第2版）

韩　劼　编著

赠送电子课件

清华大学出版社

新世纪高职高专实用规划教材　计算机系列

操作系统教程

(第2版)

韩　劼　编　著

清华大学出版社
北　京

内 容 简 介

操作系统是计算机系统最重要的系统软件，操作系统的理论和常用微机操作系统的系统管理技术是高等职业技术教育计算机技术与应用专业学生必须掌握的一门重要的专业基础课。

本书主要内容包括：操作系统的整体概念；作业界面；进程管理；存储管理；设备管理；文件管理；网络操作系统的主要概念和Linux操作系统。

根据几年来高职高专课程教学的实践，作者对原有操作系统课程体系与讲授方法进行了多方面改进，形成了本书的特色。本书除适合作为高等职业技术教育计算机技术与应用专业学生的教材以外，还适合相关专业大学本科学生或参加自考、自学的读者使用。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。
版权所有，侵权必究。侵权举报电话：010-62782989　13701121933

图书在版编目(CIP)数据

操作系统教程/韩劼编著. —2版. —北京：清华大学出版社，2008.7
(新世纪高职高专实用规划教材　计算机系列)
ISBN 978-7-302-18033-3

Ⅰ. 操…　Ⅱ. 韩…　Ⅲ. 操作系统—高等学校：技术学校—教材　Ⅳ. TP316

中国版本图书馆CIP数据核字(2008)第096948号

责任编辑：朱　颖
封面设计：山鹰工作室
版式设计：北京东方人华科技有限公司
责任校对：李玉萍
责任印制：王秀菊
出版发行：清华大学出版社　　**地　　址**：北京清华大学学研大厦A座
http://www.tup.com.cn　　**邮　　编**：100084
社　总　机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质　量　反　馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 刷 者：北京鑫海金澳胶印有限公司
装 订 者：北京市密云县京文制本装订厂
经　　销：全国新华书店
开　　本：185×260　**印　张**：22.25　**字　数**：536千字
版　　次：2008年7月第2版　　**印　　次**：2008年7月第1次印刷
印　　数：1～4000
定　　价：32.00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。联系电话：(010)62770177转3103　　产品编号：030488－01

前　言

近几年，我国高等职业技术教育迅速发展，计算机技术与应用类专业更成为热门专业。众所周知，应用能力与实用技能的培养是高等职业技术教育的重点与特色，那么，像“操作系统”这样的理论性、概念性相当强的课程，在高职教育中是不是还要开设、应当怎样开设，确实是需要认真探讨的问题。我们认为，这需要从该门课程本身的性质、高职学生学习它的目的以及高职学生的知识基础等方面综合考虑。

操作系统是计算机系统中最重要的系统软件，是整个计算机系统的核心，也是用户使用计算机必然接触到的界面。从这个方面来看，学习计算机技术与应用就不能不学习操作系统。操作系统的理论和系统管理技术是高等职业技术教育计算机技术与应用专业学生必须掌握的、重要的专业基础知识。

高等职业技术教育计算机技术与应用专业学生学习操作系统的理论与概念，目的是什么呢？显然，不是为了研究或开发操作系统。经过多年教学实践，我们的体会是要通过学习，使学生具有：

- 对计算机进行系统管理的能力，能对系统性能做简单分析并做必要的调整、优化，而不仅仅是在某种操作系统环境中工作的一般应用型能力。
- 在操作系统平台上开发应用程序的能力，能在程序中调用操作系统提供的一些内核功能，如实现进程间的通信与同步等，而不仅仅是只会编写一些简单的、一般功能的程序。
- 在应用程序设计中借鉴操作系统中某些技术思想的能力，提高自己的编程水平。

但是，也必须看到，无论从学生知识基础、学习年限，还是从课程体系安排、与先修课程的衔接等问题来看，高职教育都与大学本科计算机专业有明显差别。所以，决不能把大学本科计算机专业的课程完全照搬过来，也不能仅仅在本科专业课程的基础上删去一些内容、减少一些课时。

所以，为高职高专学生讲解操作系统的知识时，既要有一定的理论内容，不能只是讲解常用计算机操作系统用户界面的使用方法；又要注意理论内容的深度、广度适当，不但要比大学本科专业课的内容浅显，而且要与计算机常用操作系统的实际相结合。注重于理解(而不是深入研究)这些计算机操作系统是怎样管理有关的系统资源，以及作为系统管理人员应当怎样运用相应的命令和工具软件去配置系统、了解系统工作状况、优化系统。具体来说，我们在编写这本教材时着重考虑了这样几个问题。

- ✓ 在内容取舍与体系安排上，总体仍然是首先介绍操作系统的整体概念(第 1 章)与作业界面(第 2 章)，然后分别对操作系统的几大部分深入讲述，即：进程管理(第 3、4 章)，存储管理(第 5 章)，设备管理(第 6 章)，文件管理(第 7 章)。但是在一些局部也做了调整，目的在于突出重点，便于学生尽早接触并领会操作系统的核心技术思想，同时又分散难点。最后还简单介绍了网络操作系统的主要概念(第 8 章)和 Linux 操作系统(第 9 章)。
- ✓ 在讲述形式上、叙述方法上力图把有关知识条理化、通俗化，避免出现那些对于

高职学生来说可能过于艰深的抽象论述。

- ✓ 操作系统不仅是理论性很强的课程，而且也是实践性很强的课程。考虑到目前社会上实际情况、高职院校的实验条件以及先修语言、程序设计等因素，我们选择 Windows 操作系统作为实例，结合理论讲解 Windows 系统的具体实现方法，同时举出若干采用系统 API 函数的 Visual Basic 程序示例。而没有采用一般操作系统课程中普遍使用的 UNIX 系统、C 语言程序设计的做法。对于具有相应条件的教学单位，我们在最后一章对 Linux 操作系统作了简单介绍，作为一个选修内容。
- ✓ 尽管是在第 3、4、5、6、7 各章对操作系统的几大管理分别讨论，但是注意了整体和局部功能的的关系，提倡把有关内容前后串联、对比、逐步系统化的思路，避免把知识割裂开来、变成一个个散块。

当然由于编写者水平有限，加之时间仓促，本书难免会有缺点和疏漏，还望读者批评指正，以使我们不断改进。

在本书编写过程中，天津大学边奠英教授、南开大学刘璟教授、天津工业大学李兰友教授、天津商学院高福成教授等给予了热情的指导与帮助，对此表示衷心的感谢。

编 者

2008 年 5 月

试读结束：需要全本请在线购买： www.ertongbook.com

目　录

第1章 引 论

教学提示：在引论中，首先要说明操作系统在计算机系统中的作用与地位，然后以操作系统发展的简要历史作为引导，学习操作系统的基本技术思想——多道程序设计思想，进而学习进程、资源、并发、共享等基本概念以及操作系统的功能。由于操作系统是最贴近硬件的系统软件，是计算机硬件的“第一次扩充”，所以掌握它所依赖的硬件环境十分重要。必须讲清中断机制的有关概念、中断的工作过程，以及它在计算机系统，特别是操作系统工作过程中的作用，讲清为什么处理器要有核心态和用户态这两种工作模式，然后在此基础上理解操作系统的结构和工作机制。

教学目标：掌握操作系统的概念、作用与分类、操作系统的资源管理观点、操作系统的功能与结构。理解有关进程、并发、资源共享等概念。掌握中断、处理器工作模式及相关概念，为以后各章的学习奠定理论基础。

1.1 操作系统的定义与作用

当今计算机系统的使用都离不开操作系统。可以说，每位计算机用户实际上都是通过某种操作系统去使用计算机的，都需要基本掌握在某种操作系统上的操作方法，由此可见操作系统的地位。那么，到底什么是操作系统呢？

一般来说，学习一个概念总是需要首先搞清楚它的定义。但是要想给出操作系统的准确定义是很困难的，许多关于操作系统的论著中有着不同的提法。一个操作系统包括哪些部分、不包括哪些部分，也没有统一的规定。现在操作系统已经成为一种软件产品、一种商品，对于其范围，各生产厂商也有不同的界定。可以说，要想讲清楚“操作系统是什么”，还不如讲清楚“操作系统是做什么的”、“操作系统有什么作用”来得更方便。

1. 操作系统在计算机系统中的地位

从计算机系统组成的角度看，一个完整的计算机系统是由硬件系统和软件系统两大部分组成的。其中软件系统又按其所面向的对象和着眼点的不同而分为两大类：面向计算机本身功能进行组织管理、维护，从而简化用户在各个环节上的工作的软件，称为系统软件；而面向用户、为用户解决各种具体实际问题的软件，称为应用软件。应用软件需要在系统软件创造的适当环境下运行。操作系统(Operating System)就是系统软件中的一种，而且可以说是系统软件的核心。除操作系统以外的其他系统软件可以称之为系统实用软件。

于是计算机系统就形成了这样一种层次结构，从里向外(或说从最底层开始直到用户)分别是：硬件、操作系统、系统实用软件、应用软件。操作系统是最靠近硬件的一个层次，它控制和管理着在它内层的硬件系统，也控制和管理着在它外层的系统实用软件和用户应用软件，为其他软件提供了良好的开发与运行环境，并与各种系统实用软件协作，从而使各种应用软件得以开发和正常、高效率地运行。而从用户的角度讲，操作系统则是用户与计算机之间的接口。上述计算机系统各部分的作用和关系如图 1.1 所示，这些层次既

相互独立，又紧密相连、互相依赖，形成完整的计算机系统，完成各种信息处理任务。

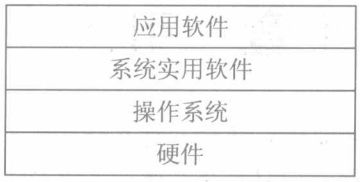


图 1.1 计算机系统的层次关系

2. 操作系统的作用

在这样一种层次结构中，操作系统的作用可以归纳为 3 点：提高效率、扩展功能、方便用户。

- 合理充分地控制和利用各种软、硬件资源，合理地组织系统的工作流程，提高系统资源的利用率，最大限度地满足用户的使用需要。
- 提供软件的开发与运行环境，使计算机系统的功能从最基本的硬件(所谓“裸机”)开始不断得到扩充。用户是无法使用没有任何软件的裸机的，而各种软件的运行离不开操作系统的支持，其他的系统程序及应用程序都是在操作系统提供的操作界面下，依赖操作系统提供的硬件服务和输入/输出控制，才得以建立或运行的。所以，操作系统是开发和运行其他软件的一个平台。在不同的操作系统上开发出来的软件，只有在该操作系统环境下才能正常运行。
- 提供了方便友好的用户操作界面，使用户可以很容易地使用和操控计算机。操作系统的一系列程序规定了计算机从启动到各种操作的过程和方式，用户只要掌握操作系统的工作过程及其提供的操作命令，就能够直接控制计算机完成各种复杂的信息处理任务。从用户的角度讲，操作系统就是用户与计算机之间的接口。

3. 操作系统的定义

现在回到本节开始提到的问题：“操作系统是什么”。通常根据操作系统的以上 3 点作用，给出操作系统“非形式化的定义”如下。

操作系统是一个大型系统软件，由大量的程序模块和数据结构集合而成，它全面地统一控制和管理着计算机系统的所有软、硬件资源，合理地组织计算机的工作流程，以便高效率地利用这些资源为用户服务，使用户有一个功能强大且可扩展的工作环境，向用户提供方便友好的操作界面。

提示： 操作系统的这个“定义”，并不必死记硬背，还是从根本上理解操作系统的地位与作用为好。而且，这种理解需要随着本课程学习的逐步展开而加深。

1.2 操作系统的形成与发展

为了进一步理解操作系统的作用，就需要了解它的形成与发展过程，因为操作系统并不是随着计算机的问世而同时出现的，它是计算机技术发展的产物。

1. 操作系统的形成

1)　手工操作阶段

20 世纪 50 年代末以前，人们是以手工操作的方式来使用计算机的，也就是说那时还没有“操作系统”这样的软件。用户自己启动输入/输出设备、输入程序和数据、启动程序运行，并在运行结束后取出结果。运行过程中如果有什么问题，也要由用户运用硬件的指示信号和操作按钮加以观察、判断和调整。于是，一个用户在完成一项工作(作业)的过程中，所有操作都是“联机”、“顺序”(串行)进行的，这台计算机的全部资源从始至终都由他独占(不论他是否始终在使用)，而且其中夹着许多“人工干预”。于是，快速的 CPU 和慢速的外设、人工操作的矛盾，特别是人机工作速度差异太大的矛盾，严重制约着整个作业速度。如果说，早期计算机的运算速度本来也不高，输入/输出设备的速度更低，手工操作的矛盾还不算突出。那么，随着计算机技术的迅速发展，计算机运算速度和输入/输出设备的工作速度都在迅速提高，人机速度矛盾越来越严重。人们急需提高计算机作业的速度，而“手工操作”却在严重地拖着后腿。所以，从 20 世纪 50 年代末开始，着手解决这个矛盾的任务被提到了日程上。

2)　监控程序阶段

首先人们想到在作业过程中避免人工干预。这里所说的“作业”，是指用户要使用计算机完成的一个独立的、完整的计算任务或事务处理任务。为了解决人机之间的矛盾，编制了标准的输入/输出程序和对用户程序进行装配、对运行进行控制、对运行出错进行处理等的程序。这些程序逐渐发展壮大，形成了“程序库”和“监控程序”(Monitor)。用户可以事先在外围处理机上准备好若干个作业(称为“脱机方式”)，由监控程序把各作业的控制命令和作业本身(程序与数据)输入到磁带上，然后把磁带装到主机上，由主机逐个地调入作业，进行编译、装配、运行、输出。于是，在一个作业内的操作和在作业之间的转换都实现了自动化，这是一种成批处理的系统。其中，监控程序对成批作业进行管理，这就是操作系统最早的形式，为了区别于后来的批处理系统，可以称之为“早期的批处理系统”或“单道批处理系统”。但是随着人机速度矛盾的初步解决，计算机主机与外部设备工作速度的差异过大成为主要矛盾。而且，监控程序与各个用户程序轮流工作，缺少安全保护措施。

3)　执行系统阶段

监控程序并不能完全解决高速的 CPU 和低速的外设之间的矛盾。人们注意到，一个程序在运行过程中经常会停下来等待某些输入输出操作的完成。所以很自然地就会想到，如果能在内存中同时存放几个用户程序，当一个程序停下来等待外设传送数据时立即让另一个程序执行，就可以使主机 CPU 的利用率得以提高。特别是不同作业的程序往往有不同的特点，有的以输入/输出操作为主，有的以在内存中运算为主，如果它们搭配得当，就能使主机和外设都几乎不停地工作，这就是“多道程序”的主导思想，即在内存中同时存放若干个作业的程序，宏观上它们在同时工作着。当然，为了实现这一点，还需要有计算机硬件方面的条件。随着计算机技术的发展，特别是进入 20 世纪 60 年代以后，硬件方面出现了“通道”和“中断”等技术，实现了主机 CPU 和输入/输出设备的并行(parallel，意指肩并肩地一起前进)工作，使操作系统进入了“执行系统”(Executive)阶段，对操作系

试读结束：需要全本请在线购买：www.ertongbook.com

统的真正形成起了重要的推动作用。

2. 操作系统的成熟

在计算机内存中同时存在着多个互相独立的程序，它们在宏观上都在运行着，但是由于只有一个 CPU(在单处理机系统中)，所以在微观上它们还是在轮流使用 CPU 执行各自的指令序列，这称为多道程序的并发(concurrence，意指同时存在或发生)。并发的多道程序可以充分地利用主机和外设的并行操作能力，所以使系统资源利用率大大提高。

注意：在有关操作系统的著作中经常出现“并发”与“并行”这两个术语，严格说来它们的含义是有差别的。“并行”是指多个操作、处理在真正(每一时刻都)、同时地进行，在本书中主要用于描述计算机的不同部件、设备在同时工作；“并发”则是指在一段时间内多个事物同时存在、发生，在本书中主要用于描述多个程序同时投入运行。

在原先的批处理系统中采用多道程序设计技术就产生了“多道批处理系统”，随后又出现了“分时系统”、“实时系统”。20 世纪 60 年代，上述 3 种系统先后被研究出来，标志着“操作系统”的正式形成。操作系统迅速完善、普及，直至进入了“通用操作系统”的阶段。所谓通用操作系统是指同时具有多道批处理、分时、实时等两种或 3 种处理能力的系统，其典型代表有 IBM 的 OS/360 系统。随着操作系统功能越来越复杂、规模越来越大，一些共同性的、较深层次的问题逐渐被归纳出来，使操作系统的理论研究从 20 世纪 60 年代末开始广泛、深入地开展起来，其成果又反过来进一步促进了操作系统的进步和成熟，20 世纪 70 年代开发并迅速发展起来的 UNIX 就是一个优秀的、被广泛使用的操作系统。

3. 操作系统的发展

随着计算机科学技术的飞速发展，个人微型计算机、多处理器计算机和计算机网络的出现，促使操作系统的理论和实践进一步发展。从 20 世纪 70 年代后期开始，出现了微机操作系统、网络操作系统、分布式操作系统等。

1.3 操作系统的基本概念

学习操作系统的形成与发展，当然不是单纯地为了学历史。从中能够看出，“多道程序设计”的思想可以说是操作系统理论的一个基本思想。领会了这一点，才能真正理解操作系统最重要的概念：“进程”、“资源”，才能真正理解操作系统最基本的特征：“并发”、“共享”和“随机”。

1.3.1 多道程序设计的思想

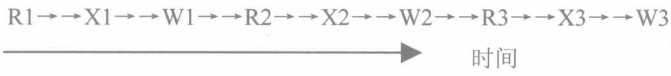
“多道程序设计”(multiprogramming)的思想是从原先的、现在被称为“单道”的程序顺序执行思想发展而来的。那么，它们各自有什么特点呢？

1. 程序的顺序执行

冯·诺依曼结构的计算机主要技术特点之一就是“程序的顺序存储与控制”。简单说来，程序就是按照一定顺序排列的指令集合。从微观上说，组成程序的一条条指令在存储器中按照一定顺序存放，程序运行时这些指令则被依次送入控制器执行，转换为一个个动作。一个具有独立功能的程序在独占了 CPU 后直到其运行完毕得到最终结果的过程就称为程序的顺序执行，早期的计算机都是这样工作的。假设一个程序按照

输入数据(R)→计算处理(X)→输出结果(W)

这样几大部分的次序编制和运行，如果有 1、2、3 三个程序要运行，只能依次进行，每个程序在执行时都独占 CPU 和其他所有系统资源，执行结束后再把全部资源交给下一个程序。

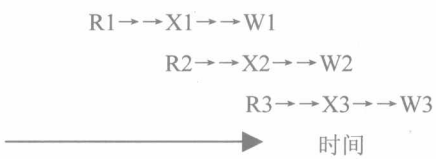


在一个处理机上多个程序的顺序执行具有如下的特点。

- 顺序性。计算机严格地按照程序规定的顺序去执行程序所指定的动作，只有前一个动作完成了，后一个动作才能开始。也就是说，程序本身(一条条指令)和执行程序的活动(一个个动作)严格地一一对应。
- 封闭性。计算机工作所涉及的“处理机”、“指令”和“环境”这几个要素都仅仅与当前运行的一个程序有关，处理机被这个程序独占，它执行的是这个程序的指令，环境(包括所使用的内存空间、各有关寄存器如指令计数器等的内容和状态)只由这个程序本身决定，不受任何外界影响。即使在两个动作之间由于人为干预而出现暂时停顿，但程序计数器内容不变，所以停顿后仍能继续按原顺序执行。无论怎样，一个程序执行的最终结果只取决于这个程序本身及其初始条件。
- 可再现性(结果确定性)。当重复运行一个程序时，只要原始数据(初始条件)不变，则运行的全部过程可以再现，也必定得到同样的最终结果。即使程序中有错误而造成计算结果有误，也能重现这个错误的发生情况。

2. 多道程序的并发执行

如前所述，为了提高处理机的工作效率，引入了多道程序的思想。仍以同时存在 3 个程序为例，不必等到第一个程序执行完，就可以开始第二个程序的运行。



这就叫做多道程序的并发执行，即通过计算机不同硬件部分并行工作，使几个程序运行时间有重叠，从宏观上看好像在同时工作。

可以举一个比较形象的例子来说明这种执行方式的效果。设有甲、乙两个程序，甲程序在执行时将要用到的资源与时间顺序为：CPU—5 秒，设备 A—10 秒，CPU—4 秒，设备 B—12 秒，CPU—7 秒；乙程序则顺序为：CPU—3 秒，设备 A—12 秒，CPU—6 秒，设备 B—5 秒，CPU—4 秒。如果按照顺序方式执行，甲程序的执行共需 38 秒，其中 CPU

工作时间16秒，利用率为42%；乙程序的执行共需30秒，其中CPU工作时间13秒，利用率为43%。无论哪个程序先执行，总共执行68秒，其中CPU工作时间29秒，利用率为42.6%。如果按照多道程序并发的方式，甲程序先开始执行，情况如图1.2所示(图中X为等待时间，注意这里没有考虑状态切换所消耗的时间，事实上一个程序申请I/O或得到I/O结束的通知都需要消耗CPU一段时间，从不占有CPU到占有CPU也需要消耗CPU一段时间，当然这些消耗的时间都非常短)。

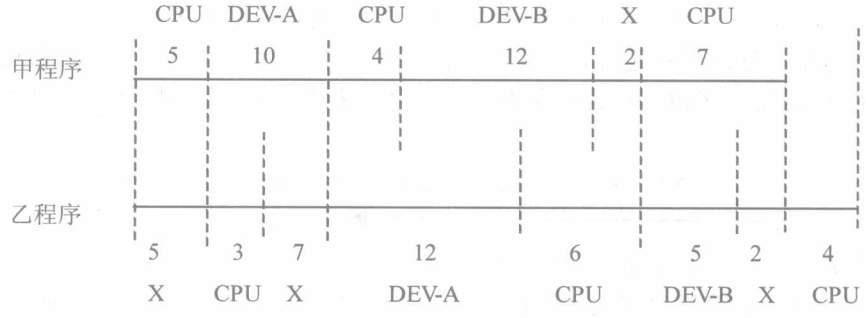


图1.2 按甲、乙顺序开始投入并发执行

总共执行44秒，其中CPU工作时间29秒，利用率为66%。如果是乙程序先开始执行，情况如图1.3所示。总共执行48秒，其中CPU工作时间29秒，利用率为60%。可见，在多道程序并发的方式下，总的执行时间减少，CPU利用率得以提高。

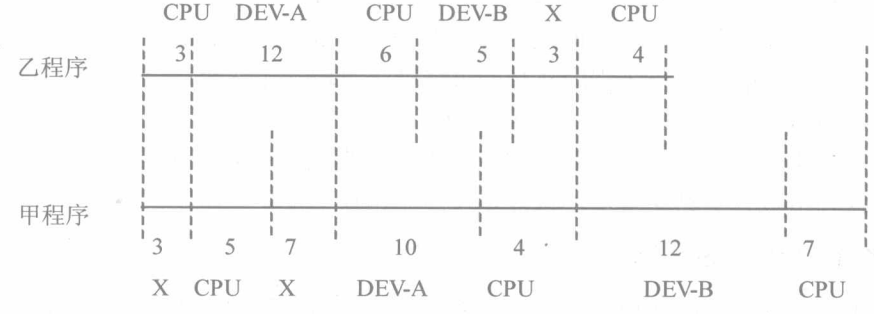


图1.3 按乙、甲顺序开始投入并发执行

这种多道程序的执行方式使得程序顺序执行方式下的封闭性和可再现性将不复存在，而是出现了新的特点。

- 并发性(concurrence)。计算机系统中同时有两个或两个以上程序存在，它们都已经开始运行而且都还没有结束运行。宏观上它们在同时运行，但微观上，同一个系统硬件(例如处理器)还是被几个程序轮流使用。只不过处理器在执行完某个程序的某一条指令后，并不一定就接着执行该程序的下一条指令，而是很可能转而去执行另一个程序的一条指令。如果使用系统的不同硬件，则几个程序的不同类型的动作可能同时工作，如前面所示意的W1、X2和R3。
- 共享性(sharing)，即资源共享。这时系统的各项资源都不再由一个程序独占，而

是由多个程序共同使用。不仅硬件资源(处理器、内存、外设等)被共享，软件资源(如编译、连接等实用软件等)也是如此。有共享就有竞争，例如处理器，即使在多处理器系统中，处理器的个数总是要少于多道程序的道数，于是各个程序就要争相使用。

- 随机性(randomness)。系统中运行的各程序的行为和状态是随机的，它们何时开始运行、何时做输入/输出、何时由 CPU 执行其指令等都无法在事先确切地知道，因为它们在“并发”，在竞争使用各种资源。系统中的硬件设备工作状态也有类似的情况。

注意： 切不可把“随机性”理解成随意性，以为在多道系统中程序执行的结果都是随心所欲的。事实上，操作系统的工作就在于控制、协调好各个程序的运行和资源的使用，为正常工作提供保证。

多道程序的并发执行是现代操作系统中最基本、最重要的技术。所以上述特点也就是现代操作系统的基本特征。并发性提高了系统资源的利用效率，但另一方面，资源的共享、竞争和状态的随机性也为多道程序设计带来了一系列必须加以解决的问题。为了深入理解这些问题，首先必须明了两个基本的概念：进程与资源。

1.3.2　进程与资源

1. 为什么需要引进“进程”这样一个新概念

正如上面所分析的，顺序执行的程序与并发执行的程序有着完全不同的特点。并发程序失去了顺序程序的封闭性和可再现性。在多道系统中，处理器和系统环境不是被一个程序独占，而是被几个程序交替地掌握，这就是“不封闭性”。各个程序开始运行的时间是随机的，每个程序开始运行后的前进速度也是不可预测的(因为这期间还有若干其他运行的程序“插”进来占有系统资源)，于是程序运行的结果也就有可能无法确保再现。比如有两个程序，甲程序运行过程中在一定时机要修改某个数据的值，而乙程序运行过程中则在一定时机要读取这个数据的值，当这两个程序与其他程序一起“并发”工作时，如果不采取特殊措施就不能保证甲乙两个程序的同步，对这个公共的数据谁先读谁先写就是不确定的，因而结果也就是不确定的。这种影响，当然是计算机用户所不能接受的。

所以，操作系统必须管理和控制并发程序的运行，消除上述影响。在此首先遇到的问题就是，“程序”这个在理解计算机工作原理时极为重要的概念，在多道系统中就显得不够清晰了。在过去程序顺序执行(相对于多道来说也可以称为单道)的情况下，只要把程序本身设计好，并掌握了把它投入运行的初始条件(包括原始数据)，就可以把它的整个运行过程及结果完全掌握了。而在多道系统中，只掌握程序本身和初始条件，还不能完全解释它在运行中所发生的各种情况。这是因为：

一方面，还有其他程序与这个程序“并发”工作，大家“共享”资源，互相竞争，于是每个程序运行的“前进”速度就不只取决于自己，还会受到其他程序运行情况的影响。

另一方面，如果这个程序和另一个程序在逻辑功能上还有联系的话，那么它们的运行情况更是会互相影响，就像前面所举的甲、乙两个程序例子那样。

程序是个静态的概念，而程序的运行具有动态的特征。当一个程序要运行时，操作系统要考虑的不仅仅是这个程序本身有什么功能，而是必须同时考虑：是哪个用户要执行这个程序，它的具体处理对象是一组什么数据，是否允许该用户执行这个程序，如果允许的话该用户的重要程度如何，可以分配给他多少系统资源，允许他运用这个程序做什么，不允许做什么，等等。显然，同一个程序，如果有3个用户分别执行它，操作系统应该是作为三件事来对待，而且是在3种工作环境下运行。这就是在操作系统理论中引进“进程”概念的必要性。

2. 进程的定义

关于进程(Process)的定义，目前的说法较多，通常可以从以下几个方面来理解。简单地说，一个进程，就是一个程序针对某个数据集合(处理对象)的一次执行过程。也就是说，进程既包括程序，还包括程序运行所使用的存储空间(程序代码、程序处理的数据、程序运行中临时所用的存储单元等都放在这个空间内)，还包括程序运行时的现场场景(各个寄存器的内容，堆栈的内容等)。进程是系统进行资源分配与调度的一个独立单位。每当要运行一个程序时，系统就要为之建立一个进程，指定一个进程名和一个进程号(标识码)，分配给它一定的资源，为它形成相应的工作环境。当程序运行结束时，进程也就被取消，分配给它的资源由系统收回。

可见，程序与进程既有联系又有区别。

- 程序是个静态的事物，是若干指令的有序集合；而进程是程序的一次运行活动，是一条条指令按顺序执行的过程，是个动态的概念。
- 程序存放在一定的存储介质上，一般(只要不删除)是长期存在的；而进程不但占用存储空间还要占用CPU，只有一个很短的存在时间(生命期)。
- 同一个程序可以被多个进程同时执行，而一个进程也可以包括执行多个程序。

有了进程这个概念，就容易理解操作系统的工作情况了。正是进程，而不是程序，才包括了程序运行时的全部有关信息，才是一个个可以并发工作、共享资源的独立的基本单位。操作系统面向的是进程，而不是程序。

注意： “进程”是操作系统中最核心的概念，在学习操作系统的过程中，应当尽量早地理解它。可以说，本书以下所有内容几乎都与这个概念有关。

3. 资源及其特点

操作系统全面地管理计算机硬件、软件资源，把它们分配给各个并发的进程。其中硬件资源主要有处理器、内存储器、外部设备(包括外存)。软件资源主要是存放在外存储器(对于微机来讲主要是磁盘)上的信息，即程序文件和数据文件。它们是计算机系统中全部活动的物质基础和环境。

对于各种资源，从不同的角度上看，就会分别具有不同的特点，可分为不同的种类。

从“是否能被几个作业同时使用”的角度看，分为“共享资源”(例如内存空间就可以划分成若干部分，被几个作业分别、同时使用)和“独享资源”(例如处理器，每一时刻只能由一个作业独占，多个作业只能轮流使用，在宏观上“共享”)。

从“当被一个作业使用时是否还能被其他作业抢夺去使用”的角度看，分为“可抢占

的资源”(例如处理器、内存空间)和“不可抢占的资源”(例如打印机、磁带机，只有等当前作业完成了输出、读写操作后才能释放出来)。

从“是否能被反复使用”的角度看，分为“永久资源”(例如各种硬件资源、外存上的数据文件和可执行程序等)和“临时资源”(例如作业之间的通信内容，使用后即被撤销，不能再次使用了)。

如前所述，操作系统的基本特征之一就是“共享” 资源。资源总是有限的，而并发的进程要竞争使用它们，这就是操作系统要解决的基本矛盾。

1.3.3　操作系统依赖的硬件环境

操作系统是最靠近硬件层的软件，其工作是直接依赖硬件环境的。那么，从操作系统的角度来看，有哪些硬件问题是特别需要关注的呢？

1. 硬件概述

计算机主机的硬件主要包括中央处理器(CPU)、主存储器、外部设备控制器，它们之间由总线连接，如图 1.4 所示。

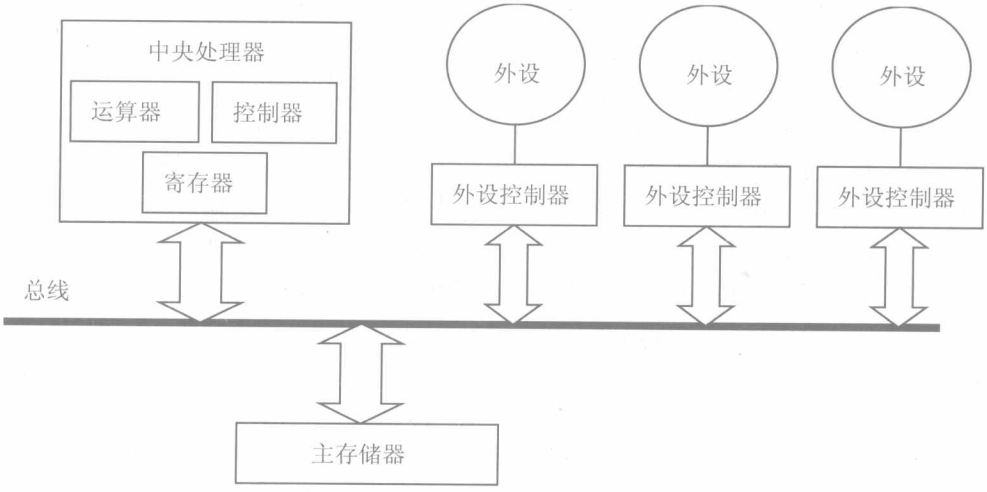


图 1.4　计算机硬件组成

数字式电子计算机是自动化的信息处理机。在计算机内，计算对象和计算步骤都是以二进制形式出现的，被计算、存储和传送的都是二进制数码。计算对象——“数据”就是以二进制数码形式存入计算机的各种字符、数值、图形图像、声音……。计算步骤就是程序，即指令序列。指令是人们控制计算机进行各种操作、运算的二进制代码，一条指令对应着一个基本操作，程序则对应着一系列有序的操作，也就是完成一个任务的步骤。

1)　中央处理器(CPU)

CPU 是执行指令的核心部件，它由控制器、运算器和若干寄存器组成。控制器根据指令的内容产生指挥运算器和其他部件协调工作的控制信号。运算器完成二进制数码的各种运算。CPU 的基本功能是从内存储器中依次取出组成某程序的一条条指令，解释该指令并执行该指令，整个计算机的工作过程可以看成是 CPU 在重复一个个“取指、译码、

试读结束：需要全本请在线购买：　www.ertongbook.com