



新世纪高职高专实用规划教材

• 计算机系列

操作系统教程

CAOZUO XITONG JIAOCHENG

韩 劼 编著



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

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重要的专业基础知识。

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。

版权所有，翻印必究。举报电话：010-62782989 13501256678 13801310933

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。

本书防伪标签采用特殊防伪技术，用户可通过下述方法识别真伪：在图案表面涂抹清水，图案消失，水干后图案复现；将表面膜揭下，放在白纸上用彩笔涂抹，图案在白纸上再现。

图书在版编目(CIP)数据

操作系统教程/韩劼编著. —北京：清华大学出版社，2005.7

(新世纪高职高专实用规划教材 计算机系列)

ISBN 7-302-11288-6

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(2005)第 070924 号

出版者：清华大学出版社 地 址：北京清华大学学研大厦
http://www.tup.com.cn 邮 编：100084
社总机：010-62770175 客户服务：010-62776969

组稿编辑：王景先

文稿编辑：闫光龙

封面设计：陈刘源

排版人员：王 婷

印刷者：北京市清华园胶印厂

装订者：三河市新茂装订有限公司

发行者：新华书店总店北京发行所

开 本：185×260 印张：19.5 字数：462 千字

版 次：2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

书 号：ISBN 7-302-11288-6/TP·7435

印 数：1~4000

定 价：25.00 元

《新世纪高职高专实用规划教材》序

编写目的

目前,随着教育的不断深入,高等职业教育发展迅速,进入到一个新的历史阶段。学校规模之大,数量之众,专业设置之广,办学条件之好和招生人数之多,都大大超过了历史上任何一个时期。然而,作为高职院校核心建设项目之一的教材建设,却远远滞后于高等职业教育发展的步伐,以至于许多高职院校的学生缺乏适用的教材,这势必影响高职院校的教育质量,也不利于高职教育的进一步发展。

目前,高职教材建设面临着新的契机和挑战:

(1) 高等职业教育发展迅猛,相应教材在编写、出版等环节需要在保证质量的前提下加快步伐,跟上节奏。

(2) 新型人才的需求,对教材提出了更高的要求,即教材要充分体现科学性、先进性和实用性。

(3) 高职高专教育自身的特点是强调学生的实践能力和动手能力,教材的取材和内容设置必须满足不断发展的教学需求,突出理论和实践的紧密结合。

有鉴于此,清华大学出版社在相关主管部门的大力支持下,组织部分高等职业技术学院的优秀教师以及相关行业的工程师,推出了一系列切合当前教育改革需要的高质量的面向就业的职业技术实用型教材。

系列教材

本系列教材主要涵盖以下领域:

- 计算机基础及其应用
- 计算机网络
- 计算机图形图像处理与多媒体
- 电子商务
- 计算机编程
- 电子电工
- 机械
- 数控技术及模具设计
- 土木建筑
- 经济与管理
- 金融与保险

另外,系列教材还包括大学英语、大学语文、高等数学、大学物理、大学生心理健康等基础教材。所有教材都有相关的配套用书,如实训教材、辅导教材、习题集等。

教材特点

为了完善高等职业技术教育的教材体系,全面提高学生的动手能力、实践能力和职业技术素质,特意聘请有实践经验的高级工程师参与系列教材的编写,采用了一线工程技术人员与在校教师联合编写的模式,使课堂教学与实际操作紧密结合。本系列丛书的特点如下:

(1) 打破以往教科书的编写套路,在兼顾基础知识的同时,强调实用性和可操作性。

(2) 突出概念和应用,相关课程配有上机指导及习题,帮助读者对所学内容进行总结和提高。

(3) 设计了“注意”、“提示”、“技巧”等带有醒目标记的特色段落,使读者更容易得到有益的提示与应用技巧。

(4) 增加了全新的、实用的内容和知识点,并采取由浅入深、循序渐进、层次清楚、步骤详尽的写作方式,突出实践技能和动手能力。

读者定位

本系列教材针对职业教育,主要面向高职高专院校,同时也适用于同等学历的职业教育和继续教育。本丛书以三年制高职为主,同时也适用于两年制高职。

本系列教材的编写和出版是高职教育办学体制和运作机制改革的产物,在后期的推广使用过程中将紧紧跟随职业技术教育发展的步伐,不断吸取新型办学模式、课程改革的思路和方法,为促进职业培训和继续教育的社会需求奉献我们的力量。

我们希望,通过本系列教材的编写和推广应用,不仅有利于提高职业技术教育的整体水平,而且有助于加快改进职业技术教育的办学模式、课程体系和教学培训方法,形成具有特色的职业技术教育的新体系。

教材编委会

新世纪高职高专实用规划教材

计算机系列编委会

顾问 吴文虎

主编 边莫英

副主编 刘 璟 李兰友

委员 (以姓氏笔画为序)

王景先 王温君 刘光然 许洪杰

曲建民 迟丽华 李 平 汪 洋

林章波 张 炜 张 玲 赵家俊

高福成 傅连仲 韩 劼 喻 梅

詹青龙 魏则燊

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操作系统。操作系统的理论和常用计算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、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。
- 操作系统虽然是理论性很强的课程,但同时也是实践性很强的课程。考虑到目前社会上的实际情况、高职院校的实验条件以及先修语言、程序设计等因素,我们选择了 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。
- 尽管在第 3、4、5、6、7 各章是对操作系统的几大管理的分别讨论,但是注意了整体和局部功能的关系,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

2005 年 5 月

目 录

第 1 章 引论	1	2.3 Windows 的用户界面.....	42
1.1 操作系统的定义与作用.....	1	2.3.1 操作命令接口.....	42
1.2 操作系统的形成与发展.....	3	2.3.2 编程接口.....	43
1.3 操作系统的基本概念.....	4	2.4 习题.....	45
1.3.1 多道程序设计的思想.....	5	第 3 章 进程管理	48
1.3.2 进程与资源.....	7	3.1 进程.....	48
1.3.3 操作系统依赖的硬件环境.....	9	3.1.1 进程的基本特征.....	48
1.3.4 当前操作系统的主要分类.....	13	3.1.2 进程状态及其转换.....	49
1.3.5 研究分析操作系统的 几种观点.....	16	3.1.3 进程的描述.....	51
1.3.6 操作系统的功能.....	18	3.1.4 进程控制.....	54
1.4 目前微机常用操作系统的特点.....	20	3.2 线程.....	60
1.4.1 DOS.....	20	3.2.1 线程的概念.....	60
1.4.2 Windows.....	21	3.2.2 线程的种类与实现.....	62
1.4.3 UNIX.....	23	3.2.3 Windows 系统中的进程 与线程.....	62
1.4.4 Linux.....	24	3.3 处理器调度.....	66
1.5 操作系统的组成与工作机制.....	26	3.3.1 处理器调度的 3 种类型.....	66
1.5.1 操作系统的组成结构.....	26	3.3.2 进程调度算法.....	68
1.5.2 操作系统的引导.....	28	3.3.3 进程调度的时机.....	71
1.5.3 操作系统的基本 工作机制.....	28	3.3.4 进程调度的操作内容.....	71
1.6 习题.....	29	3.3.5 Windows 系统的线程调度.....	72
第 2 章 作业管理与用户界面	32	3.4 习题.....	76
2.1 作业及其管理.....	32	第 4 章 进程通信与死锁	79
2.1.1 作业管理的一般概念.....	32	4.1 死锁.....	79
2.1.2 批量型作业的管理.....	34	4.1.1 死锁的基本概念.....	79
2.1.3 终端型作业的管理.....	38	4.1.2 死锁的预防.....	82
2.2 系统调用.....	39	4.1.3 死锁的避免.....	82
2.2.1 系统调用的一般概念.....	39	4.1.4 死锁的检测与解除.....	84
2.2.2 系统调用的执行过程与 使用方法.....	41	4.2 进程之间的同步与互斥.....	84
		4.2.1 进程之间的关系.....	84
		4.2.2 同步与互斥.....	85

4.2.3 生产者与消费者问题.....	87	5.5 Windows 系统的存储管理.....	128
4.2.4 进程互斥的一种实现方法.....	88	5.5.1 地址空间.....	128
4.3 进程间的低级通信—— 信号量及其操作.....	89	5.5.2 页表与地址转换.....	131
4.3.1 信号量与 P、V 操作.....	89	5.5.3 调页.....	132
4.3.2 运用信号量实现同步 与互斥.....	90	5.5.4 内存页帧的状态与队列.....	133
4.3.3 经典的进程同步 问题示例.....	93	5.6 段式与段页式存储管理.....	135
4.3.4 Windows 系统中的 同步与互斥.....	95	5.6.1 段式存储管理.....	135
4.4 进程间的高级通信.....	101	5.6.2 段页式存储管理.....	137
4.4.1 消息缓冲通信.....	101	5.7 习题.....	138
4.4.2 信箱通信.....	103	第 6 章 设备管理	142
4.4.3 管道.....	104	6.1 设备管理的基本概念.....	142
4.5 习题.....	105	6.1.1 设备的分类.....	142
第 5 章 存储管理	108	6.1.2 设备管理的任务.....	143
5.1 存储管理的基本概念.....	108	6.2 设备管理的基本概念.....	144
5.1.1 存储系统的层次组织.....	108	6.2.1 设备的连接与控制.....	144
5.1.2 程序及其运行与存储器 地址的关系.....	110	6.2.2 设备的控制方式.....	145
5.1.3 存储管理的基本任务.....	112	6.2.3 缓冲区的管理.....	149
5.2 分区存储管理.....	113	6.3 I/O 软件原理.....	150
5.2.1 分区存储管理技术.....	113	6.3.1 设备处理程序.....	150
5.2.2 固定分区.....	114	6.3.2 物理设备与逻辑设备.....	153
5.2.3 可变分区.....	114	6.3.3 I/O 进程的工作过程.....	154
5.2.4 覆盖技术.....	116	6.4 设备的分配与回收.....	154
5.3 简单页式存储管理.....	116	6.4.1 设备分配的一般问题.....	155
5.3.1 页面.....	116	6.4.2 虚拟设备与 SPOOLING 技术.....	155
5.3.2 页表与地址映射.....	117	6.4.3 磁盘调度问题.....	156
5.3.3 快表与关联寄存器.....	119	6.5 习题.....	160
5.3.4 交换技术.....	119	第 7 章 文件管理	163
5.4 请求页式虚拟存储管理.....	120	7.1 文件系统的基本概念.....	163
5.4.1 虚拟存储的一般概念.....	120	7.1.1 文件与文件系统.....	163
5.4.2 请求页式存储管理技术.....	121	7.1.2 外存设备的存储特点.....	164
5.4.3 调页与页面淘汰.....	122	7.1.3 文件的逻辑结构.....	165
5.4.4 页面置换算法.....	124	7.1.4 文件的物理结构.....	167
5.4.5 局部性原理与抖动现象.....	126	7.1.5 文件的存取方式.....	169
		7.2 文件系统的实现.....	170
		7.2.1 文件目录.....	170
		7.2.2 盘图文件.....	173

7.2.3 Windows 的 FAT 文件系统	173	8.3.4 传输层通信接口 Winsock	220
7.2.4 Windows 的 NTFS 文件系统	182	8.4 习题.....	227
7.3 文件共享与安全.....	185	第 9 章 Linux 操作系统简介	229
7.3.1 文件的共享	186	9.1 Linux 概述	229
7.3.2 文件的保密	187	9.1.1 系统构成与引导.....	229
7.3.3 文件的保护	187	9.1.2 基本应用知识.....	237
7.3.4 NTFS 系统的安全性 措施.....	188	9.1.3 文件与目录.....	242
7.4 文件操作的实现过程.....	189	9.1.4 创建与维护用户帐户 和组帐户	246
7.4.1 文件系统的功能模块.....	189	9.1.5 系统调用.....	249
7.4.2 文件操作的基本内容 与过程	189	9.2 Linux 的进程管理	251
7.4.3 Windows 文件系统的 层次结构	190	9.2.1 有关进程的一些概念	251
7.5 习题	191	9.2.2 进程控制.....	257
第 8 章 网络操作系统简介	194	9.2.3 进程调度.....	263
8.1 计算机网络与网络操作系统.....	194	9.2.4 进程通信.....	265
8.1.1 计算机网络	194	9.3 Linux 的存储管理	270
8.1.2 网络体系结构与协议.....	195	9.3.1 虚空间与实空间的映射	270
8.1.3 网络操作系统.....	198	9.3.2 物理空间页帧的管理	272
8.2 局域网网络操作系统.....	199	9.3.3 交换.....	273
8.2.1 局域网的工作模式.....	199	9.4 Linux 的设备与文件管理	275
8.2.2 局域网网络操作系统 的组成	200	9.4.1 设备管理.....	275
8.2.3 主从网中的服务器 操作系统	201	9.4.2 磁盘空间管理	277
8.2.4 Windows 服务器 操作系统	202	9.4.3 EXT2 文件系统	279
8.3 操作系统对互联网的支持.....	207	9.4.4 文件系统是可装卸的	284
8.3.1 网络互联技术的特点.....	207	9.4.5 Linux 通过 VFS 能支持 多种文件系统.....	287
8.3.2 TCP/IP 协议	208	9.4.6 文件的共享与保护	288
8.3.3 Windows 对互联网 的支持	212	9.5 习题.....	290
		参考答案	293
		参考文献	298

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应用

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。

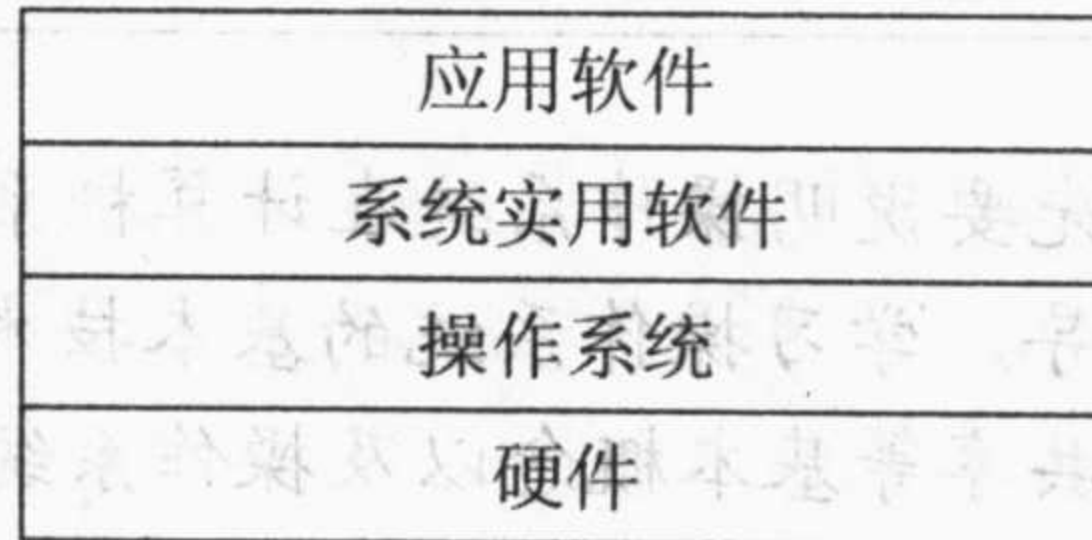


图 1.1 计算机系统的层次关系

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硬、软件资源，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(即所谓“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硬、软件资源，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

二十世纪五十年代末以前,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密。人们急需提高计算机作业的速度,而“手工操作”在严重地拖着后腿。所以,从五十年代末开始,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。这些程序逐渐丰富发展,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，特别是进入二十世纪六十年代以后，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”。二十世纪六十年代，上述 3 种系统先后被研究出来，标志着“操作系统”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六十年代末开始广泛、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，七十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。从二十世纪七十年代后期开始，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

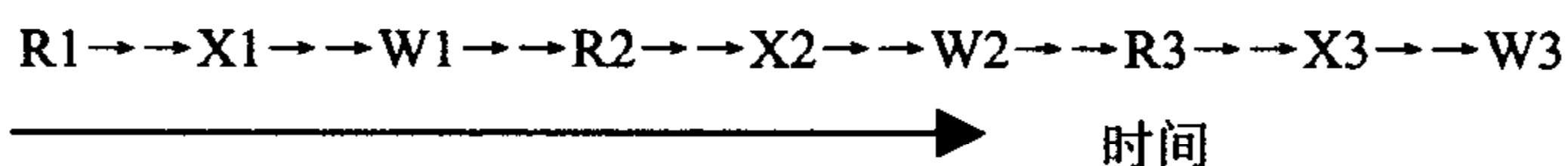
“多道程序设计(multiprogramming)”的思想是从原先的、现在被称为“单道”的程序顺序执行思想发展而来的。那么，它们各自有什么特点呢？

1. 程序的顺序执行

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。

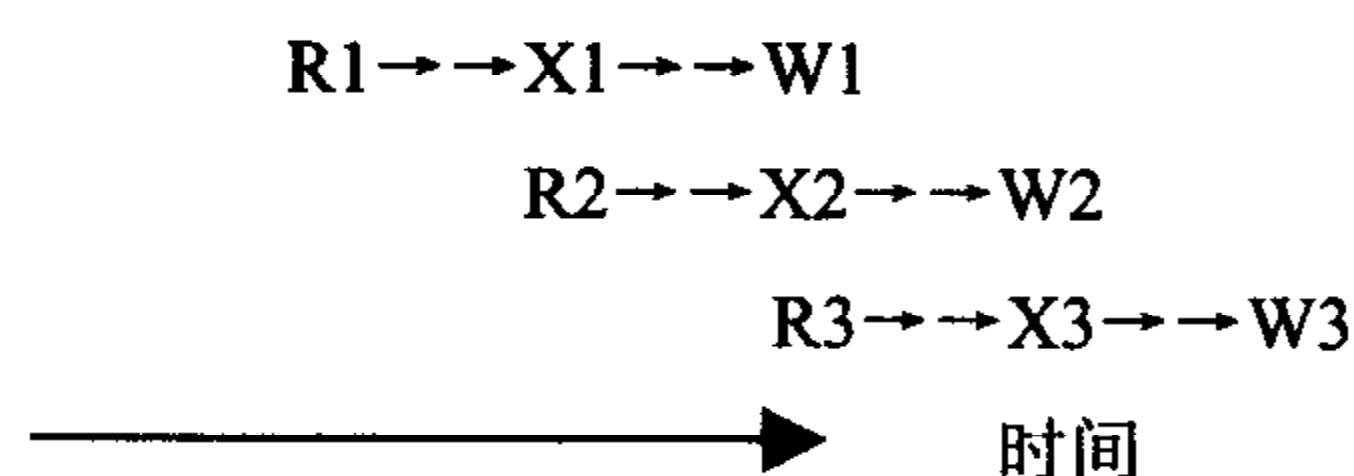


在一个处理机上多个程序的顺序执行具有如下的特点：

- 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。计算机工作所涉及到的“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。



这就叫作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)。

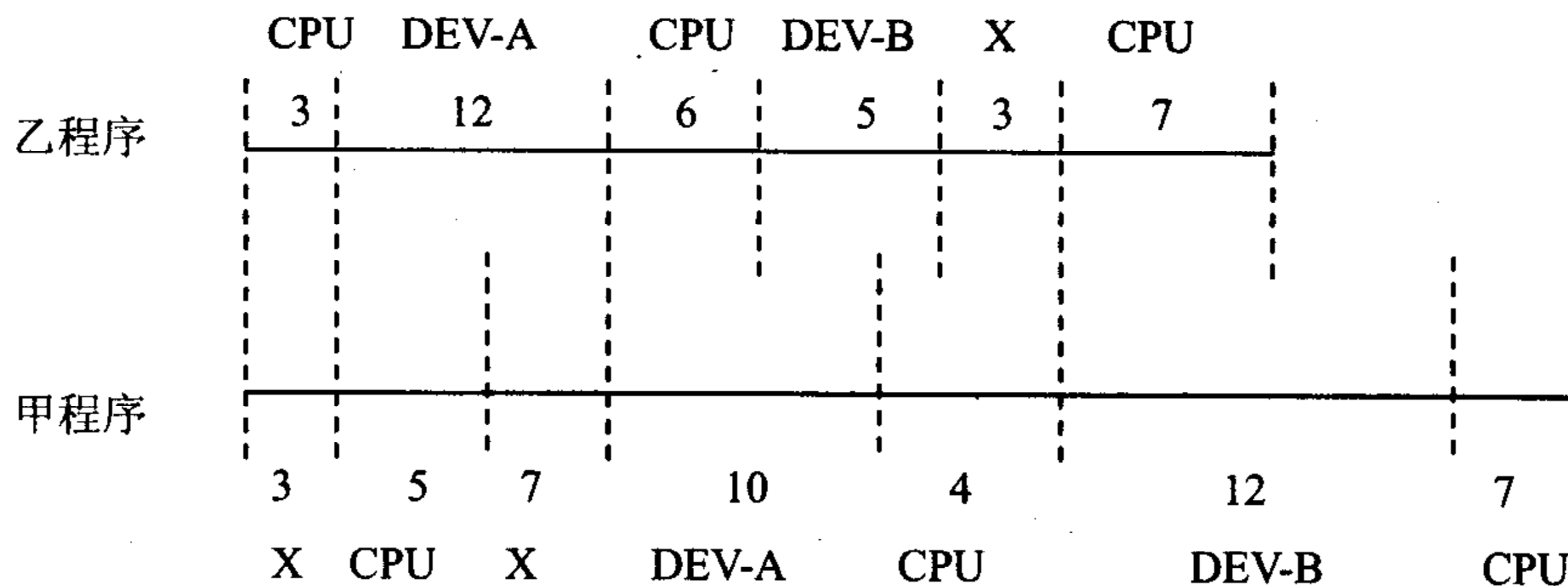


图 1.2 按甲、乙顺序开始投入并发执行

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。

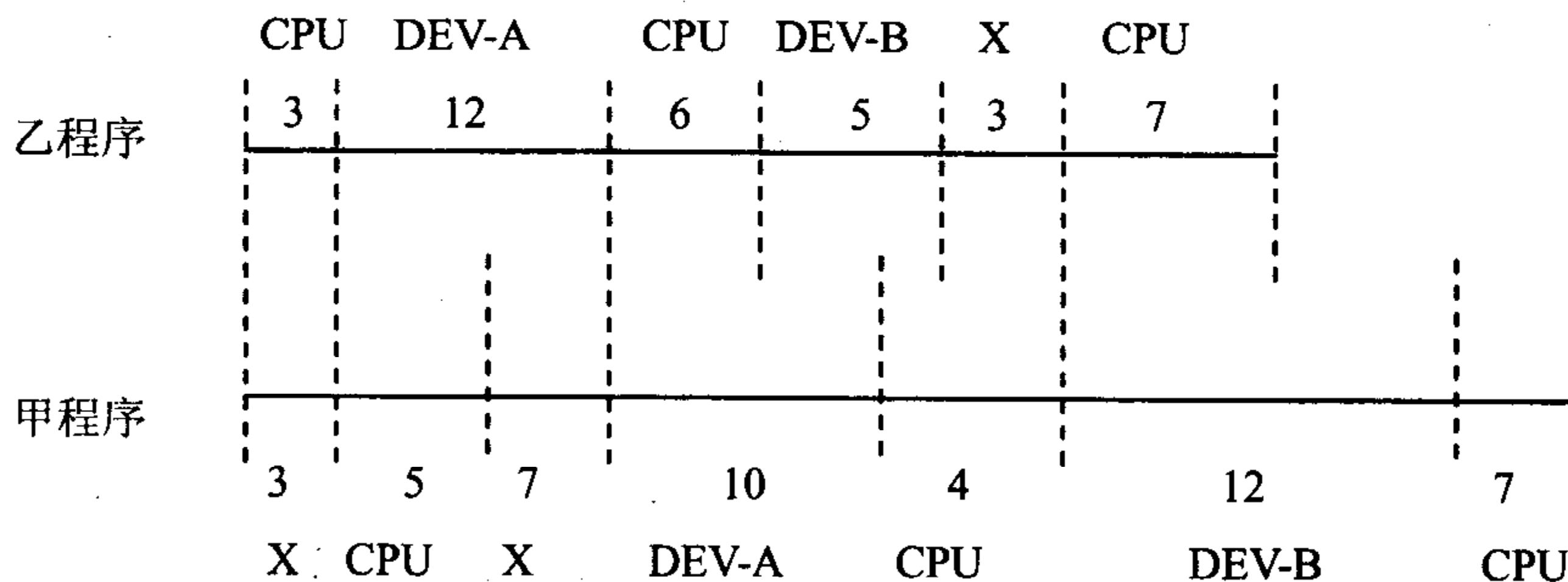


图 1.3 按乙、甲顺序开始投入并发执行

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(concurrency)。计算机系统中同时有两个或两个以上程序存在，它们都已