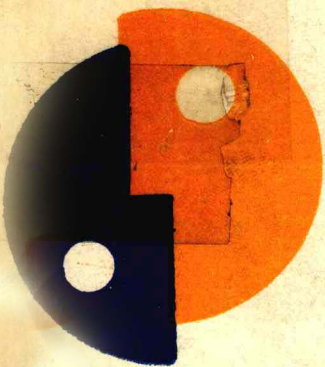




# 操作 系统 原理

华中理工大学出版社

庞丽萍 编



# 操作系统原理

庞丽萍 编

华中理工大学出版社

操作系统原理

庞丽萍 编

责任编辑 唐元瑜

华中理工大学出版社出版发行

《武昌喻家山》

新华书店湖北发行所经销

华中理工大学出版社沔阳印刷厂印刷

开本: 787×1092 1/16 印张: 17.75 字数: 400 000

1988年6月第1版 1990年5月第3次印刷

印数: 15 061—20 000

ISBN 7-5609-0196-4/TP·18

定价: 5.95元

## 内 容 简 介

本书阐述了操作系统的基本原理、主要功能及实现技术，给出了操作系统完整的逻辑结构和用户界面，并以 PC-DOS 为例介绍了操作系统的层次结构和系统调用方法。全书共分十章，主要内容包括：描述并发活动的基本概念——进程，进程控制及进程的交互作用；计算机系统各类资源的分配、调度策略及资源管理功能的具体实施方法。

本书可供高等院校计算机类本科、专科各专业和电视大学有关专业作为教材，亦可供从事计算机科学、工程和应用等方面工作的科技人员自学和参考。

## 序 言

操作系统是计算机系统中最重要系统软件之一。人们使用计算机必须和操作系统打交道，它往往给人一种神秘且威力无边的感觉。为了能得心应手地使用计算机和开发更高层的软件，必须了解和掌握操作系统的奥秘。本书旨在揭示操作系统的基本原理、主要功能、实现技术及逻辑结构，并以IBM-PC机上配置的PC-DOS为例剖析操作系统的结构和简介其系统调用方法。

若干年来我一直在华中理工大学计算机系担任“操作系统原理”这门课程的主讲工作。在教学实践中深感操作系统的内容庞杂、学生不易掌握其实质和形成有机的整体概念。为此，我在教材内容的选择和组织方面作了一些探索，对在本科生中使用多年的自编讲义反复进行修改，不断在教学实践中检验，从而形成了现在的《操作系统原理》一书。

为了揭示操作系统的奥秘并易被人们所理解，我编写本书的指导思想是，从计算机体系结构出发，以提高效率、实现并行处理为目标提出现代操作系统所必然具备的并发共享的特性，这一特性应始终贯穿在操作系统的各功能模块之中。即以实现并发为主导，以资源共享为手段把操作系统的内容组织成一个完整的逻辑体系。对计算机系统各类资源的管理是大量和复杂的，为了克服内容零散和庞杂的弊病，本书首先提出对资源共性的认识，把对资源的管理从资源数据描述、分配策略和管理方法高度上统一起来，形成对操作系统资源管理全局的清晰了解。在此基础上再依各类资源的个性分别进行描述，然后给出一个完整的操作系统逻辑结构，并强调操作系统和用户的接口，因为这一部分是挖掘其内部潜力、打开操作系统大门的钥匙。最后给出PC-DOS实例，简介其系统调用和开发CC-DOS的方法，使读者能从感性上认识一个具体实用的操作系统模型，初步了解操作系统的接口并知道如何灵活地运用它们。

随着信息革命的深入，从事计算机科学工作的人会越来越多，他们都需要学习操作系统，但不同层次的人所需掌握的深度是不同的。为此，我将本书的内容按两种层次编排，使之各成一个完整的体系。本书中除“\*”号之外的内容是作为计算机科学的电大、大专学生以及具有相同水平的科技工作者所必须掌握的。对于计算机科学的本科生则应掌握本书的全部内容。

本书论述详尽、逻辑清晰、例证有力、插图丰富。除第九、十章外均配有习题，其中有的习题可作为讨论题（打“\*”号的习题电大及专科的学生可以不做）。第十章中的一些例子可上机验证。本书被中央电视大学选作教材，授课54学时，拟讲授除“\*”号之外的全部内容（若课堂讲授则需60学时）。作为高等院校本科生教材建议按70~80学时讲授。

我在历年教学和编写本书过程中，广泛参考了国内外操作系统方面的大量著作，并从中汲取了不少营养、得到了很多的启发和帮助。这些著作主要有：Kaisler, Stephen H. 的“The design of operating systems for small computer systems”；Peterson, James Lyle 的“Operating system concepts”；〔英〕利斯科著 徐良贤 尤晋元 王志良译的《操作系统原理》；孙钟秀等编著的《操作系统原理》；钱培德编的《CC-BIOS分析》等。

在本书的编写过程中,得到了华中理工大学计算机系刘键教授的指导;何双成老师认真审阅了部分章节并提出了宝贵的意见;本书责任编辑唐元瑜同志原是我的同行,他对本书的编辑和阅稿使我受益不浅。另外,华中理工大学计算机系和软件教研室的领导以及校教材科的同志一直关心这本书的写作和出版。在此,对所有为这本书的编写、审阅和出版给予支持和帮助的同志一并表示衷心的感谢。

计算机科学的发展是迅速的,它所涉及的内容在技术和概念上都在急剧地变化。《操作系统原理》这本书也不例外,它的内容也需不断更新。另外,本书中一定还存有少许谬误之处,敬请读者能将反馈信息告诉我,使我能汲取更多的营养,以便进一步提高这本书的质量。

作者

1987年9月

# 目 录

<b>第一章 绪论</b> .....	( 1 )
1.1 存贮程序式计算机 .....	( 1 )
1.1.1 存贮程序式计算机的结构和特点 .....	( 1 )
1.1.2 计算机系统结构与操作系统的关系 .....	( 2 )
1.2 操作系统的形成和发展 .....	( 3 )
1.2.1 早期计算机 .....	( 3 )
1.2.2 批处理 .....	( 4 )
1.2.3 多道程序设计与多道成批系统 .....	( 6 )
1.2.4 分时系统 .....	( 8 )
1.2.5 实时系统 .....	( 10 )
1.2.6 计算机网络 .....	( 11 )
1.2.7 分布式系统 .....	( 13 )
1.3 操作系统的基本概念 .....	( 14 )
1.3.1 操作系统的定义及其在计算机系统中的地位 .....	( 14 )
1.3.2 操作系统功能要点 .....	( 16 )
*1.3.3 定义程序运行环境 .....	( 16 )
*1.3.4 确定联系方式 .....	( 17 )
*1.3.5 建立操作界面 .....	( 17 )
1.3.6 管理系统资源 .....	( 18 )
1.4 操作系统的特性及其应解决的基本问题 .....	( 20 )
1.4.1 操作系统的特性 .....	( 20 )
1.4.2 操作系统应解决的基本问题 .....	( 20 )
习题一 .....	( 22 )
<b>第二章 用户界面</b> .....	( 23 )
2.1 用户工作环境 .....	( 23 )
2.1.1 用户环境 .....	( 23 )
2.1.2 运行一个用户程序的过程 .....	( 24 )
2.2 用户与操作系统的接口 .....	( 25 )
2.2.1 操作系统提供的两个接口 .....	( 25 )
2.2.2 作业控制语言和键盘命令 .....	( 26 )
2.3 系统服务请求 .....	( 27 )
2.3.1 系统调用命令 .....	( 28 )
2.3.2 系统调用的实现 .....	( 28 )
2.3.3 系统服务处理举例 .....	( 30 )
习题二 .....	( 33 )
<b>第三章 并发处理</b> .....	( 34 )
3.1 并发活动——进程的引入 .....	( 34 )

3.1.1	程序的顺序执行 .....	( 34 )
3.1.2	程序的并发执行 .....	( 35 )
3.1.3	并发执行实例——誊抄 .....	( 37 )
3.1.4	与时间有关的错误 .....	( 40 )
3.1.5	并发程序的特点 .....	( 42 )
3.2	进程概念 .....	( 44 )
3.2.1	进程的定义 .....	( 44 )
3.2.2	进程的类型 .....	( 45 )
3.2.3	进程的状态 .....	( 46 )
3.2.4	进程的描述——进程控制块 .....	( 48 )
3.3	进程控制 .....	( 49 )
3.3.1	进程控制的概念 .....	( 49 )
3.3.2	进程创建 .....	( 50 )
*3.3.3	进程撤消 .....	( 51 )
3.3.4	进程挂起 .....	( 54 )
*3.3.5	进程唤醒 .....	( 54 )
*3.3.6	进程延迟 .....	( 55 )
3.4	进程的相互制约关系 .....	( 58 )
3.4.1	资源共享 .....	( 58 )
3.4.2	进程合作 .....	( 58 )
3.5	进程互斥 .....	( 59 )
3.5.1	互斥的概念 .....	( 59 )
3.5.2	锁和LOCK、UNLOCK操作 .....	( 61 )
3.5.3	用上锁原语和开锁原语实现进程互斥 .....	( 62 )
3.6	信号灯和P、V操作 .....	( 63 )
3.6.1	信号灯的概念 .....	( 63 )
3.6.2	P、V操作 .....	( 64 )
3.6.3	用信号灯实现进程互斥 .....	( 65 )
3.7	进程同步 .....	( 66 )
3.7.1	同步的概念 .....	( 66 )
3.7.2	同步的例子 .....	( 66 )
3.7.3	用信号灯实现进程同步 .....	( 66 )
3.7.4	生产者-消费者问题 .....	( 70 )
3.8	进程通信 .....	( 72 )
3.8.1	进程通信的概念 .....	( 72 )
3.8.2	消息地址匹配式通信 .....	( 72 )
*3.8.3	消息缓冲通信 .....	( 74 )
*3.8.4	信箱通信 .....	( 77 )
习题三	.....	( 80 )
<b>第四章</b>	<b>资源分配与调度 .....</b>	<b>( 83 )</b>
4.1	资源管理概述 .....	( 83 )
4.1.1	资源管理的目的和任务 .....	( 83 )
4.1.2	资源的几种分类方法 .....	( 84 )



4.1.3	资源管理的机构和策略	( 85 )
4.2	资源分配机构	( 86 )
4.2.1	资源描述器	( 86 )
4.2.2	资源信息块	( 86 )
4.3	系统资源描述	( 87 )
4.3.1	中央处理机	( 87 )
4.3.2	主存	( 88 )
4.3.3	外部设备	( 89 )
4.3.4	后备存储器	( 90 )
4.3.5	文件	( 90 )
4.4	资源分配策略	( 90 )
4.4.1	概述	( 90 )
4.4.2	先请求先服务	( 91 )
4.4.3	优先调度	( 92 )
*4.4.4	适应调度	( 93 )
*4.4.5	均衡调度	( 93 )
*4.4.6	针对设备特性的调度	( 93 )
4.5	死锁	( 94 )
4.5.1	死锁的概念	( 94 )
4.5.2	死锁的起因	( 96 )
4.5.3	解决死锁问题的策略	( 98 )
4.5.4	死锁的预防	( 101 )
*4.5.5	死锁的避免	( 101 )
*4.5.6	死锁的检测与恢复	( 103 )
	习题四	( 104 )
<b>第五章</b>	<b>处理机调度</b>	( 106 )
5.1	处理机的二级调度	( 106 )
5.2	作业调度	( 106 )
5.2.1	作业调度的功能	( 106 )
5.2.2	作业控制块JCB	( 107 )
5.2.3	调度性能的衡量	( 108 )
5.2.4	先来先服务调度算法	( 110 )
5.2.5	短作业优先调度算法	( 110 )
*5.2.6	响应比高者优先调度算法	( 111 )
*5.2.7	优先数调度算法	( 112 )
*5.2.8	均衡调度算法	( 112 )
5.3	进程调度	( 113 )
5.3.1	调度/分派结构	( 113 )
5.3.2	进程调度的功能	( 114 )
5.3.3	调度方式	( 114 )
5.3.4	调度用的进程状态变迁图	( 115 )
5.3.5	进程优先数调度算法	( 116 )
5.3.6	循环轮转调度	( 117 )

*5.3.7 多重时间片循环调度	(118)
习题五	(119)
<b>第六章 存贮管理</b>	(121)
6.1 主存共享特征——空间分片	(121)
6.2 主存管理的功能	(121)
6.2.1 主存映射	(122)
6.2.2 程序的逻辑组织	(124)
6.2.3 主存分配	(125)
6.2.4 提供虚存	(125)
6.2.5 存贮保护	(127)
6.3 分区存贮管理	(129)
6.3.1 概述	(129)
6.3.2 用基址寄存器实现动态地址映射	(131)
6.3.3 分区分配机构	(131)
6.3.4 分区的分配和回收	(132)
6.3.5 几种基本的放置策略	(137)
6.3.6 碎片问题及拼接技术	(140)
6.4 页式存贮管理	(141)
6.4.1 页式系统应解决的问题	(141)
6.4.2 页式地址变换	(142)
6.4.3 清调策略	(145)
6.4.4 淘汰策略	(148)
6.4.5 几种置换算法	(148)
*6.4.6 页式系统的存贮分配	(152)
6.4.7 放置策略	(154)
6.4.8 页式系统中的共享	(154)
*6.4.9 页式系统的保护措施	(155)
*6.4.10 工作集模型	(156)
6.5 段式系统	(153)
6.5.1 段式系统的特点	(158)
6.5.2 段式地址变换	(158)
6.5.3 扩充段表功能	(159)
习题六	(159)
<b>第七章 输入输出管理</b>	(162)
7.1 I/O管理概念	(162)
7.1.1 引言	(162)
7.1.2 设计目标	(163)
7.1.3 I/O管理功能	(164)
7.1.4 设备独立性	(164)
7.1.5 设备控制块	(167)
7.2 中断技术	(168)
7.2.1 中断概念	(168)

7.2.2	中断进入 .....	(169)
7.2.3	两种类型的中断机制 .....	(170)
7.2.4	中断处理 .....	(173)
7.3	缓冲技术 .....	(176)
7.3.1	缓冲概述 .....	(176)
7.3.2	双缓冲 .....	(176)
7.3.3	环形缓冲 .....	(177)
*7.3.4	缓冲池 .....	(178)
7.4	设备分配 .....	(180)
7.4.1	设备分配原则 .....	(180)
7.4.2	独享分配 .....	(182)
7.4.3	共享分配 .....	(182)
7.4.4	虚拟设备技术 .....	(182)
7.4.5	SPOOLING系统 .....	(184)
*7.4.6	慢速字符设备的分配 .....	(185)
7.5	输入输出控制 .....	(187)
7.5.1	输入输出控制概念 .....	(187)
7.5.2	输入输出过程 .....	(189)
7.5.3	设备处理进程 .....	(190)
7.5.4	用户进程调用外部设备的过程 .....	(193)
	习题七 .....	(193)

**第八章 文件系统 .....** (194)

8.1	文件系统的概念 .....	(194)
8.1.1	引言 .....	(194)
8.1.2	文件 .....	(195)
8.1.3	文件系统 .....	(197)
8.2	文件的逻辑组织与存取方法 .....	(198)
8.2.1	文件的组织 .....	(198)
8.2.2	文件的逻辑结构 .....	(199)
8.2.3	存取方法 .....	(200)
8.3	文件的物理结构 .....	(202)
8.3.1	连续文件 .....	(202)
8.3.2	串联文件 .....	(203)
8.3.3	文件映照 .....	(206)
8.3.4	随机文件 .....	(207)
8.3.5	文件物理结构比较 .....	(210)
8.4	存贮技术 .....	(210)
8.4.1	文件存贮器的主要参数 .....	(210)
*8.4.2	磁带存贮技术 .....	(212)
*8.4.3	磁鼓存贮技术 .....	(213)
8.4.4	磁盘存贮技术 .....	(214)
8.5	文件存贮空间的管理 .....	(215)
8.5.1	空白文件目录 .....	(215)

8.5.2	空白块链	(216)
8.5.3	位示图	(216)
8.5.4	分配策略	(217)
8.6	文件目录	(217)
8.6.1	文件目录及内容	(217)
8.6.2	一级文件目录	(218)
8.6.3	二级文件目录	(219)
8.6.4	多级文件目录	(220)
8.7	共享与安全	(221)
8.7.1	文件共享与安全性的关系	(221)
8.7.2	建立“当前目录”实现文件共享	(222)
8.7.3	采用“链接技术”实现文件共享	(222)
*8.7.4	基本文件目录和符号文件目录	(222)
8.7.5	存取权限的类型及其验证方法	(223)
8.8	文件的完整性	(227)
8.9	文件操作	(228)
	习题八	(229)
<b>第九章</b>	<b>操作系统的逻辑结构</b>	<b>(230)</b>
9.1	概述	(230)
9.2	基本硬件机构	(232)
9.3	系统核	(234)
9.4	核外各层	(236)
<b>第十章</b>	<b>PC-DOS简介</b>	<b>(238)</b>
10.1	PC-DOS的硬件环境	(238)
10.1.1	IBM-PC的微处理器	(238)
10.1.2	存储器组织	(238)
10.1.3	寄存器结构	(240)
10.1.4	中断和中断向量表	(241)
10.1.5	堆栈	(243)
10.1.6	输入/输出结构	(243)
10.2	PC-DOS的结构	(243)
10.2.1	PC-DOS的层次结构	(243)
10.2.2	基本输入/输出系统	(244)
10.2.3	引导程序	(245)
10.2.4	输入/输出接口模块	(245)
10.2.5	文件管理程序	(246)
10.2.6	命令处理程序	(246)
10.3	PC-DOS与用户程序的接口	(247)
10.3.1	PC-DOS的软中断	(247)
10.3.2	PC-DOS的功能调用过程	(248)
10.3.3	PC-DOS的功能调用分类	(249)
10.4	PC-DOS磁盘空间分配	(254)

10.4.1	存贮格式.....	(254)
10.4.2	PC-DOS的盘区分配.....	(254)
10.4.3	文件位置分配表——FAT表.....	(255)
10.4.4	磁盘空间分配策略.....	(258)
10.5	PC-DOS的文件目录.....	(258)
10.5.1	根目录的目录项内容.....	(258)
10.5.2	文件控制块.....	(260)
10.5.3	树形目录结构.....	(261)
10.5.4	文件的路径名和当前目录.....	(261)
10.6	IBM-PC机上开发的汉字操作系统——CC-DOS .....	(262)
10.6.1	CC-DOS系统概述 .....	(262)
10.6.2	CC-BIOS的形成 .....	(263)
10.6.3	CC-DOS的自举过程 .....	(264)
10.6.4	FILE1. EXE的功能 .....	(265)
10.6.5	CCCC. EXE的功能 .....	(266)
*10.6.6	CC-BIOS 的打印模块的形成.....	(267)
<b>附录</b>	.....	<b>(268)</b>

# 第一章 绪 论

## 1.1 存贮程序式计算机

### 1.1.1 存贮程序式计算机的结构和特点

人们在科学实验、生产斗争和社会实践中有大量问题需要求解，如科学计算、数据处理及各种管理问题等。要解决这些问题，首先需要分析所研究的对象，提出对问题的形式化定义和给出求解方法的形式描述。对问题的形式化定义叫做数学模型，而对问题求解方法的形式描述称为算法。其次是必须具备实现算法的工具或设施。我们将一个算法的实现叫做一次计算。显然，一个计算既与算法有关，也与实现该算法的工具有关。算法和实现算法的工具是密切联系在一起的，二者互相影响、互相促进。

人们在生产活动和商业交易中最早需要解决的问题是算术四则运算问题，开始是凭人们的大脑和手来进行计算，随后使用算盘，再后用计算器，这些计算工具可以进行加、减、乘、除运算。当人们要解决某一问题时，只有将问题的求解方法归结为四则运算问题后，才能使用算盘之类的工具进行计算。由此可见，算法和计算工具是相互影响的。因为算法是四则运算，所以计算工具必须具备加、减、乘、除功能。当遇到一个复杂的算法时，如求解一个微分方程，若计算工具仍然只能进行四则运算，则必须把微分方程的解法转化为数值解法。

上面所谈的计算是一种手工计算方式，而算盘或计算器是手工计算的一种工具。在这种计算方式中，人们按照预先确定的一种计算方案，先输入原始数据，然后按操作步骤做第一步计算，记下中间结果，再做第二步计算，直到算出最终结果，最后把结果记录在纸上。这里，一切都是依靠人的操作，即无论是输入原始数据，执行运算操作，还是中间结果的存贮和最终结果的抄录都是依靠人的操作，所以这一计算过程是手工操作过程。

著名数学家 Von Neumann 总结了手工操作的规律以及前人研究计算机的经验教训后，提出了“存贮程序式计算机”方案，从而使计算初步实现了自动化。要使计算机能够自动地计算，必须使机器可以“看到”计算方案即计算机程序，能够“理解”程序语言的含义并顺序执行指定的操作，可以及时取得初始数据和中间数据，能够自动地输出结果。于是，机器必须：有一个存贮器，用来存贮程序和数据；有一个运算器，用以执行指定的操作；要有一个控制部件以便实现自动操作。此外，还要有输入输出部件，以便输入原始数据和输出计算结果。它们就构成了“存贮程序式计算机”或称“Von Neumann计算机”。

综上所述，存贮程序式计算机由以下五类部件组成：控制器，运算器，存贮器，输入装置，输出装置，如图1.1所示。时至今日，人们基本上还是依照这一结构来构造计算机。

人们通常把控制器和计算器做在一起，称为中央处理机或中央处理部件(CPU)。CPU是计算机的“大脑”，能控制指挥各个部件的工作。它是一种能够解释指令、执行指令并控制操作顺序的硬设备。在CPU中，控制器负责从主存贮器提取指令、并分析其类型，运算器则完成了为实现该指令所需进行的操作。CPU还包含一个小的快速存贮器，用来存贮一些暂时的结果和其它控制信息。这个存贮器由若干个寄存器组成，每一个都具有某种功能，其中很重要的一个寄存器是程序计数器(PC)，它指示下一条应该执行的指令。

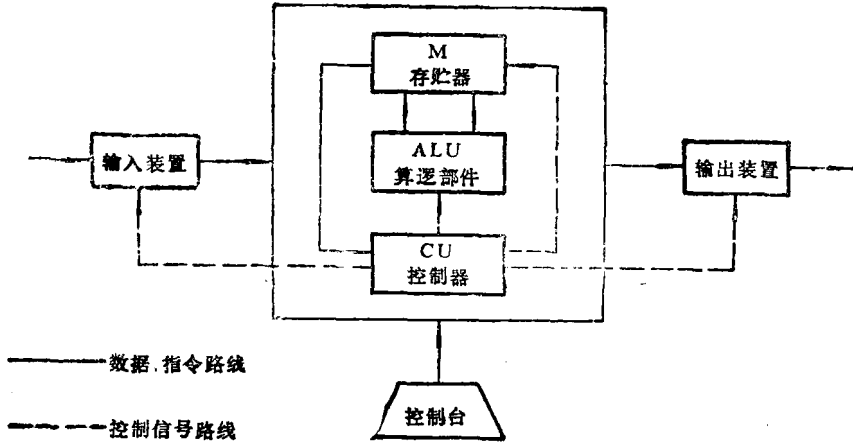


图1.1 存储程序式计算机

存储器是计算机存储程序和数据部件。如果没有一个使中央处理机能直接读、写信息的存储器，那就不存在我们所熟悉的可存储程序的数字计算机了。早期的主存储器是由磁芯做成的，价格比较昂贵。因此，大部分计算机还有一个存取速度较慢、价格较便宜、容量大得多的辅助存储器，用于保存大量的数据信息。

输入输出设备则是完成信息传输任务的。当某一个问题需要计算机处理时，必须给定程序和初始数据，它们是通过输入设备进入计算机的。当得出解答后，计算机必须把计算结果通知用户，这是通过输出设备实现的。而控制台则是操作员用来实施控制和发布命令的设备。

Von Neumann 计算机是人类历史上第一次实现自动计算的计算机，可以真正称得上是一架自动机。该机是人类历史上第一次出现的作为人脑延伸的智能工具，它的影响是十分深远的。它具有逻辑判断能力和自动连续运算能力。它的计算模型是顺序过程计算模型，其主要特点是：集中顺序过程控制，即控制部件根据程序对整个计算机的活动实行集中过程控制，并根据程序规定的顺序依次执行每一个操作。计算是过程性的，故这种计算机是模拟人们的手工计算的产物。即首先取原始数据，执行一个操作，将中间结果保存起来，再取一个数，和中间结果一起又执行一个操作，如此计算下去。在遇到有多个可能同时执行的分支时，也是先执行完一个分支，然后再执行第二个分支，直到计算完毕。由于 Von Neumann 计算机的计算模型是顺序过程计算模型，所以它具备的根本特点是集中顺序过程控制。

### 1.1.2 计算机系统结构与操作系统的关系

计算机系统的硬件基础是 Von Neumann 机的五大部件，而构成计算机系统的另一个重要的系统软件是操作系统，它负责管理计算机系统的硬、软资源和整个计算机的工作流程。顺序过程计算模型决定着 Von Neumann 型计算机的根本特点——集中顺序过程控制，操作系统既是计算机系统中的一个重要组成部分，当然也不能不反映这一特点。最早产生的单用户操作系统正是如此。它只允许一个用户使用计算机，该用户独占计算机系统的各种资源，整个系统为他的程序运行提供服务。在这里，除了CPU和外部设备有可能提供并行操作外，其余的都是顺序操作，这种单用户操作系统简单明了，容易实现。但在这样的系统中，昂贵的计算机硬部件并没有得到充分利用，计算机的性能特别是资源利用率大大低于可能达到的

程度。

为了提高资源利用率，人们作了极大的努力把单处理机系统改造成逻辑上的多处理机系统，而且使之能进行并行处理。让多个用户共用一个计算机系统，这就必须解决多个用户的算题任务共享计算机系统资源的问题，也需要解决系统如何控制多个算题任务的共同执行。为此，出现了一系列新的软件技术，如多道程序设计技术、分时技术，以及解决资源分配和调度、进程及进程间的交互作用等问题的技术。这些技术已经载入了操作系统发展的光荣史册，并被人们誉为60年代至70年代计算机科学的奇迹。在CPU和存贮器都十分昂贵的情况下，这些技术的应用取得了可观的经济效益。由于计算机系统的计算模型是顺序计算模型，其特点是集中顺序过程控制，因而导致操作系统必然具备并发、共享的特征。而这种特征正蕴藏着一对矛盾，即硬件结构是顺序过程计算模型，而操作系统却是并行处理（计算）模型，这二者是十分不协调的。这种尖锐的矛盾，使操作系统变得非常复杂、不易理解，成为一个庞然大物，且其效果并不一定很理想。

随着计算机技术的迅猛发展和计算机应用的日益广泛，操作系统出现了多种形式，有批量操作系统、分时操作系统、实时操作系统、单用户磁盘操作系统、计算机网络和正处在研制阶段的分布式操作系统等。其中，用得最广泛的一种是配置在微型机上的磁盘操作系统。这种操作系统的文件系统功能很强，并且提供各种方便用户使用计算机的手段，人们用起来得心应手，很受欢迎。如果某系统想共享其它系统的硬件或软件资源，也可考虑联网使用，这就是现在发展极快的计算机网络。

另一方面，人们也正在研究与并行计算模型一致的计算机系统结构，并使得具有并行处理能力的操作系统具有更强的生命力。在人们研究的多种并行处理结构中，有多指令流单数据流的流水线机，有单指令流多数据流的阵列机，还有多指令流多数据流的多处理机系统。现在有两种结构特别值得注意，一是数据流结构，就是按照数据相关关系用数据来驱动操作的执行，它能获得最大的并行性；另一种是分布式系统，它实行分布式控制并进行并行处理。正在酝酿的第五代计算机，后面这两种结构将是它的主要结构形式。

目前在市场上销售的计算机，大部分仍然采用 Von Neumann 式计算机的结构，在可见的将来也仍然是如此。因此，我们必须学好当前计算机系统上配置的操作系统，另外也要关心计算机系统结构发展的新趋势。从计算机体系结构的角度出发去分析操作系统，就比较容易理解操作系统的功能和特点。通过这样的分析，我们不但可以学到对当前有用的知识，而且可以鉴别，哪些是合理的，哪些是将来仍然有用的、哪些是需要改造的。只有深刻地了解过去和现在，才能更好地迎接未来。下面我们首先简单地回顾一下操作系统的形成和发展过程。

## 1.2 操作系统的形成和发展

### 1.2.1 早期计算机

操作系统在现代计算机中起着相当重要的作用。它是由于客观的需要而产生，随着计算机技术的发展和计算机应用的日益广泛而逐渐发展和完善的。它的功能由弱到强，在计算机系统中的地位也不断提高，以至成为系统的核心。我们研究操作系统的形成和发展是用一种历史的观点去分析操作系统，以便从中体会到操作系统产生的必然性和促使它发展的根本原因。



第一台电子计算机于1946年问世，此后，计算机在其运算速度、存储容量、外部设备的功能和种类等方面都有了惊人的发展和提高。人们通常按照元件工艺的演变把计算机的发展过程分为四个阶段：

1946年~1958年：第一代(电子管时代)；

1958年~1964年：第二代(晶体管时代)；

1964年~1974年：第三代(集成电路时代)；

70年代中期至今：第四代(大规模集成电路时代)。现在正向着巨型、微型、网络化、智能化几个方向发展。

在第一代计算机时期，构成计算机的主要元件是电子管，其运算速度较慢(只有几千次/秒)。早期计算机由主机(运控部件、内存)、输入设备(如纸带输入机、卡片阅读机)、输出设备(如打印机)和控制台组成。

人们利用这样的计算机解题只能采用手工操作方式。在手工操作情况下，用户一个挨一个地轮流使用计算机。每个用户的使用过程大致如下：先把程序纸带(或卡片)装上输入机，然后经手工操作把程序和数据输入计算机，接着通过控制台开关启动程序运行。计算完毕，用户拿走打印结果，并卸下纸带(或卡片)。在这个过程中需要人工装纸带、人工控制程序运行、人工卸纸带，这些都是手工操作，即所谓“人工干预”。这种由一道程序独占机器的情况，在计算机速度较慢的时候是允许的，因为此时计算所需的时间相对而言较长，手工操作时间所占比例还不会很大。

当计算机进入第二代——晶体管时代后，计算机的速度、容量、外设的品种和数量等方面和第一代计算机相比都有了很大的进展，比如，计算机的速度从每秒几千次、几万次发展到每秒几十万次、上百万次。由于计算机速度有几十倍、上百倍的提高，故使得手工操作的慢速度和计算机运算的高速度之间形成了一对矛盾，即所谓人-机矛盾。表 1.1 说明了人-机矛盾的严重性。

表1.1 操作时间与运行时间的关系

机器速度	作业在机器上计算所需时间	人工操作时间	操作时间与机器有效运行时间之比
1万次/秒	1小时	3分钟	1:20
60万次/秒	1分钟	3分钟	3:1

说明：作业是为完成用户的解题任务，计算机所需进行的各项工作。

随着计算机速度的提高，人-机矛盾已到了不可容忍的地步。为了解决这一矛盾，只有设法去掉人工干预，实现作业的自动过渡，这样就出现了成批处理。

### 1.2.2 批处理

在计算机发展的早期阶段，系统是让用户独占使用的，即在其使用期间，用户可以建立、运行他的作业，并最后作结尾处理。由于当时软件处于初期阶段，用于管理的软件还没有产生，因此所有的运行管理和具体操作都由用户自己承担。

引入批量监督程序是为了实现作业建立和作业过渡的自动化。监督程序是一个常驻内存的小的核心代码。每一种语言翻译程序(汇编语言或某种高级语言的编译程序)，或实用程序(如连接程序)都作为监督程序的子例程。