



高等学校电子信息类专业规划教材

计算机操作系统

——基于UNIX系统V核心

王育勤 谢余强 汪 淼 编著



清华大学出版社
<http://www.tup.tsinghua.edu.cn>



北京交通大学出版社
<http://press.bjtu.edu.cn>



21 世纪高等学校电子信息类专业规划教材

计 算 机 操 作 系 统

——基于 UNIX 系统 V 核心

王育勤 谢余强 汪 淼 编著

清 华 大 学 出 版 社
北 京 交 通 大 学 出 版 社

• 北 京 •

内 容 简 介

本教材是在多年教学实践的基础上,吸收国内外新的理论和技术,依据操作系统教学大纲的要求进行编写的。

本教材重点讲述了操作系统的一般原理和实施技术。在讲授方法上,注意理论与实际的结合,特别是以当代最流行的 UNIX 操作系统为实例,介绍了操作系统中主要服务功能的实施技术和技巧;在内容安排上,由一般到具体,先介绍操作系统的概念和服务功能,然后以 UNIX 系统 V 为例讲述这些功能的具体实现算法,同时介绍了 UNIX 系统 V 中的流机制和 IPC 通信机制,最后以 SCO UNIX 操作系统为基础,讨论了 UNIX 系统的安装、应用及其网络与通信,以其加强实践性环节。

本书共分 12 章。第 1 章概述操作系统的发展历史、分类、功能、体系结构及 UNIX 系统的特点。第 2 章介绍进程的基本概念、有关进程的操作、进程通信及中断处理。第 3 章介绍处理机管理。第 4 章至第 6 章分别介绍存储管理、设备管理、文件系统的概念、功能及其主要实现技术。第 7 章介绍死锁的概念、预防与检测。第 8、9 章介绍了 STREAMS (流) 机制及流的基本操作和 UNIX 系统 V 的 IPC。第 10 章至第 12 章分别介绍了 UNIX 系统的安装、应用及其网络与通信。

本书可作为高等学校计算机专业、电子信息类专业本科生和非计算机专业研究生课程的教材,也可供有关专业人员参考。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

计算机操作系统:基于 UNIX 系统 V 核心 / 王育勤,谢余强,汪淼编著. —北京:清华大学出版社;北京交通大学出版社,2004.4

(21 世纪高等学校电子信息类专业规划教材)

ISBN 7-81082-107-5

I. 计… II. ①王… ②谢… ③汪… III. 操作系统-高等学校-教材 IV. TP 316

中国版本图书馆 CIP 数据核字(2004)第 022285 号

责任编辑:陈川

出版发行:清华大学出版社 邮编:100084 电话:010-62776969

北京交通大学出版社 邮编:100044 电话:010-51686045, 62237564

印刷者:北京鑫海金澳胶印有限公司

发行者:新华书店总店北京发行所

开 本:185×260 印张:19.75 字数:502 千字

版 次:2004 年 4 月第 1 版 2004 年 4 月第 1 次印刷

书 号:ISBN 7-81082-107-5/TP·111

印 数:1~5 000 册 定价:26.00 元

前 言

操作系统是计算机系统的基本组成部分，它在整个计算机系统软件中占据核心地位。对操作系统的概念、理论和方法的研究及对操作系统的使用、分析、开发和设计，历来是计算机领域中最主要的课题和任务之一。因此，操作系统也就成为计算机科学教育的基本课程之一。它涉及到对各种资源（包括硬件和软件资源）的有效管理，为用户及高层软件的运行提供良好的工作环境，起到承上启下、纵横贯通的作用。

本教材是在近几年教学实践的基础上，吸收国内外较新的理论和技术，依据操作系统教学大纲的要求编写的。

本教材重点讲述了操作系统的一般原理和实现技术。在理论上力求系统、完整，尽量体现当前先进研究成果；在讲授方法上，注意理论与实际的结合，特别是以当代最流行的 UNIX 操作系统为实例，介绍了操作系统中主要服务功能的技术实现和技巧；在内容安排上，注意由浅入深，由一般到具体，先介绍操作系统的概念和服务功能，然后讲述这些功能的实现算法，并在最后以 SCO UNIX 操作系统为基础，讨论了 UNIX 系统的安装、应用及其网络与通信，以其加强实践性环节。

本书共分 12 章。第 1 章概述操作系统的发展历史、分类、功能、体系结构及 UNIX 系统的特点。第 2 章介绍进程的基本概念、有关进程的操作、进程通信及中断处理。第 3 章介绍处理机管理。第 4 章至第 6 章分别介绍存储管理、设备管理、文件系统的概念、功能及其主要实现技术。第 7 章介绍死锁的概念、预防与检测。第 8、9 章介绍了 STREAMS（流）机制及流的基本操作和 UNIX 系统 V 的 IPC。第 10 章至第 12 章分别介绍了 UNIX 系统的安装、应用及其网络与通信。每章的开头，先交代所要讨论的问题、环境和意义，然后逐步展开论述。在讲授理论的基础上，辅以 UNIX 系统的实例，从而加深对概念的理解和形象化思维。第 1 章至第 7 章后面附有大量习题，这些有代表性的习题对学习巩固正文中的知识是有益的。本教材在介绍 UNIX 系统 V 各功能模块的实现方法时，重点、难点突出。结合以往教学实践体会，对其难以理解的部分给出了较为详细的说明，便于学生自学复习有关内容，并为学生在 UNIX 系统环境下，上机实习、应用开发提供了指南。

本教材适用于计算机专业、电子信息类专业本科生和非计算机专业研究生。操作系统课程在 50 至 80 授课学时，教师可根据实际情况作选学处理。

在本教材编写过程中，得到了解放军信息工程大学信息科技研究所所长沈建京教授的大力支持，在此表示衷心感谢！

由于编者水平有限，时间紧迫，书中一定有错误和不当之处，恳请广大读者给予指正。

编者

2004 年 4 月

目 录

第 1 章 操作系统概述.....	1
1.1 操作系统的发展过程.....	1
1.1.1 手工操作阶段.....	1
1.1.2 早期批处理阶段.....	2
1.1.3 执行系统.....	3
1.1.4 多道程序系统阶段.....	3
1.2 什么是操作系统.....	4
1.2.1 概念.....	4
1.2.2 设置 OS 的目的.....	5
1.2.3 操作系统的主要功能.....	5
1.2.4 操作系统的服务功能.....	7
1.3 操作系统的结构.....	8
1.3.1 外结构（环境）.....	8
1.3.2 内结构（体系结构）.....	9
1.4 操作系统的分类.....	9
1.4.1 多道批处理系统.....	10
1.4.2 分时系统.....	11
1.4.3 实时系统.....	11
1.4.4 网络操作系统.....	13
1.4.5 区别.....	13
1.5 操作系统的特征.....	13
1.5.1 并发.....	13
1.5.2 共享.....	14
1.5.3 长期信息存储.....	14
1.5.4 不确定性.....	14
1.6 操作系统的性能——即如何评价一个操作系统.....	14
1.6.1 效率.....	14
1.6.2 可靠性.....	15
1.6.3 可维护性.....	15
1.7 当前比较流行的几种微机操作系统.....	15
1.7.1 当前微机上的主流操作系统.....	15

1.7.2 用户如何选用操作系统.....	18
1.8 UNIX 系统的特点和结构	19
1.8.1 外部原因	19
1.8.2 内部特点	19
1.8.3 结构	20
习题	22
第 2 章 进程管理	23
2.1 进程管理的概念	23
2.1.1 程序的顺序执行	23
2.1.2 程序的并发执行和资源共享	23
2.1.3 程序并发执行的特性	25
2.1.4 进程	27
2.1.5 用进程概念说明操作系统的并发性和不确定性	28
2.1.6 进程的状态及其变迁	29
2.1.7 进程的组成	31
2.1.8 UNIX 系统的进程映像	33
2.2 有关进程的操作	36
2.2.1 进程的创建	36
2.2.2 进程终止和父/子进程的同步	39
2.3 进程间的相互作用和通信	42
2.3.1 同步	42
2.3.2 互斥	43
2.3.3 进程的临界区和临界资源	44
2.3.4 实施临界区互斥的锁操作法	44
2.3.5 信号量与 P、V 操作	48
2.3.6 高级通讯机构	53
2.4 中断处理	57
2.4.1 中断及其一般处理过程	57
2.4.2 中断优先级和多重中断	60
2.4.3 中断屏蔽	60
2.4.4 中断在操作系统中的地位	60
2.4.5 UNIX 系统对中断和陷入的处理	61
习题	64
第 3 章 处理机管理	67
3.1 概述	67
3.1.1 CPU 调度的三级实现	67

3.1.2 进程的执行方式	70
3.1.3 CPU 调度的基本方式	70
3.2 作业概念	71
3.2.1 作业管理的概念	71
3.2.2 作业管理的功能	73
3.3 常用调度算法	77
3.3.1 先来先服务 FCFS	78
3.3.2 最短周期优先 SBF	78
3.3.3 优先级	80
3.3.4 轮转法	80
3.3.5 可变时间片轮转法	82
3.3.6 多队列轮转法和多级反馈队列法	82
3.4 UNIX 系统中的进程调度	84
习题	89
第 4 章 存储管理	91
4.1 引言	91
4.1.1 二级存储器及信息传送	91
4.1.2 存储器分配	91
4.1.3 存储管理的基本任务	92
4.1.4 存储空间的地址问题	93
4.1.5 用户程序的装入	94
4.1.6 存储管理的功能	96
4.1.7 内存的扩充技术	96
4.2 分区式管理技术	99
4.2.1 固定分区法	99
4.2.2 可变分区法	100
4.2.3 硬件支持	103
4.3 可重定位分区分配	103
4.4 多道程序对换技术	104
4.5 请求分页存储管理	105
4.5.1 分页管理	105
4.5.2 请求分页管理	109
4.6 段式存储管理	116
4.6.1 分段和分段的地址空间	117
4.6.2 分段管理的实现	118
4.6.3 分段共享	118

4.6.4 段的动态链接	119
4.7 段页式存储管理	122
4.7.1 基本思想	122
4.7.2 实现过程	122
4.8 UNIX 系统的存储管理	124
4.8.1 对换	124
4.8.2 请求分页	126
习题	131
第 5 章 设备管理	133
5.1 概述	133
5.1.1 设备分类	133
5.1.2 设备管理的目标 and 功能	133
5.1.3 通道技术	134
5.1.4 缓冲技术	137
5.2 设备分配技术与 Spooling 系统	138
5.2.1 设备分配技术	138
5.2.2 Spooling 系统	139
5.3 RK 磁盘设备	140
5.4 UNIX 系统的设备管理	141
5.4.1 UNIX 设备管理的特点	141
5.4.2 与设备驱动有关的接口	142
5.4.3 块设备管理中的缓冲技术	142
5.4.4 块设备的读写	146
5.4.5 字符设备管理	151
习题	153
第 6 章 文件系统	154
6.1 概述	154
6.1.1 文件及其分类	154
6.1.2 文件系统的功能	155
6.1.3 文件系统的用户界面	156
6.1.4 文件系统的层次结构	157
6.2 文件的组织和存取方法	158
6.2.1 文件的逻辑组织和物理组织	158
6.2.2 文件的存取方式	162
6.3 目录结构	162
6.3.1 一级目录结构	163

6.3.2	二级目录结构	163
6.3.3	多级目录结构	164
6.4	文件存储空间的管理	166
6.4.1	记住空间分配现状的数据结构	167
6.4.2	存储空间分配程序	169
6.5	文件保护	169
6.5.1	文件系统的完整性	170
6.5.2	文件的共享与保护保密	170
6.6	对文件的主要操作	171
6.6.1	创建文件	172
6.6.2	文件的连接与解除连接	172
6.6.3	文件的打开和关闭	173
6.6.4	文件的读写	173
6.7	文件系统的执行过程	174
6.8	UNIX 文件系统的内部实现	176
6.8.1	i 节点	176
6.8.2	活动 i 节点的分配与释放	177
6.8.3	用户打开文件表和系统打开文件表	178
6.8.4	目录项和检索目录文件	181
6.8.5	文件的索引结构	183
6.8.6	文件卷和卷专用块	184
6.8.7	空闲 i 节点的管理	185
6.8.8	空闲存储块的管理	187
6.8.9	子文件系统装卸和装配块表	188
6.8.10	各主要数据结构之间的联系	190
6.8.11	管道文件 (pipe)	191
6.9	系统调用的实施举例	195
	习题	198
第 7 章	死锁	199
7.1	死锁的基本概念	199
7.1.1	什么是死锁	199
7.1.2	死锁的表示	200
7.1.3	死锁判定法则	201
7.2	死锁的预防	203
7.3	死锁的避免	204
7.3.1	资源分配状态 RAS	204

7.3.2	系统安全状态	205
7.3.3	死锁避免算法	206
7.3.4	对单体资源类的简化算法	207
7.4	死锁的检测和清除	208
7.4.1	死锁的检测	208
7.4.2	死锁的解除	210
	习题	211
第 8 章	STREAMS (流) 机制	212
8.1	STREAMS 概述	212
8.1.1	STREAMS 机制的产生	212
8.1.2	STREAMS 的特征和结构	212
8.1.3	STREAMS 的应用	214
8.2	流基本操作	214
8.2.1	流的建立和关闭	215
8.2.2	流的读/写	216
8.3	流的高级操作	217
8.3.1	流组操作	218
8.3.2	多路流操作	219
8.3.3	消息处理	221
第 9 章	UNIX 系统 V 的 IPC	223
9.1	概述	223
9.2	消息	224
9.2.1	系统调用 msgget	224
9.2.2	系统调用 msgsnd	225
9.2.3	系统调用 msgrcv	227
9.2.4	系统调用 msgctl	229
9.3	共享存储区	230
9.3.1	系统调用 shmget	230
9.3.2	系统调用 shmat	232
9.3.3	系统调用 shmctl	234
9.4	信号量	235
9.4.1	系统调用 semget	235
9.4.2	系统调用 semop	238
9.4.3	系统调用 semctl	242
9.5	总的评价	243
第 10 章	UNIX 操作系统的基本应用	244

10.1	UNIX 操作系统的安装、启动和关闭	244
10.1.1	硬件配置	244
10.1.2	软件考虑	245
10.1.3	安装过程	246
10.1.4	增加用户账号	247
10.1.5	启动和关闭系统	248
10.2	UNIX 操作系统的基本操作命令	250
10.2.1	文件和目录的命名规则	250
10.2.2	UNIX 的输入和输出命令	250
10.2.3	目录管理	251
10.2.4	文件管理	253
10.2.5	改变文件及目录的属主和存取权	259
10.2.6	屏幕编辑程序 Vi	261
10.2.7	特殊命令	264
10.3	UNIX 操作系统和 DOS 操作系统的兼容操作	265
10.3.1	DOS 和 UNIX 共存系统的转换	265
10.3.2	DOS 格式的操作命令	265
10.3.3	用 DOS 文件系统	267
第 11 章	UNIX 操作系统的系统管理和实用工具程序	269
11.1	UNIX 操作系统的系统管理	269
11.1.1	系统管理软件 sysadmsh	269
11.1.2	数据转储和恢复	270
11.1.3	文件系统管理	272
11.1.4	设备管理	274
11.1.5	进程控制	277
11.2	UNIX 操作系统的实用工具程序	279
11.2.1	用户之间相互通讯	279
11.2.2	文本处理工具	281
第 12 章	UNIX 操作系统的网络与通讯	284
12.1	本地网络	284
12.1.1	建立一个网络	284
12.1.2	启动网络	285
12.1.3	测试 Micnet 网络	285
12.1.4	停止网络	285
12.1.5	网络命令使用	285
12.2	远程网络	286

12.2.1	UUCP 简介	286
12.2.2	用调制解调器连接远程 UUCP 系统	289
12.2.3	配置 UUCP	289
12.2.4	使用 UUCP	290
12.3	标准网络	293
12.3.1	显示网络中各计算机状态	293
12.3.2	显示网络中的用户	293
12.3.3	通过网络进行通信	295
12.3.4	通过网络在另一台主机上注册	296
12.3.5	在另一台主机上执行命令	296
12.3.6	在不同的计算机之间传送文件	297
12.3.7	使用 ftp 命令传送文件	298
	参考文献	301

第 1 章 操作系统概述

1.1 操作系统的发展过程

操作系统这门学科的产生和出现与其他任何新观点、新概念一样，并不是突然产生的，也不是刚有计算机就有了操作系统这门学科，它也有一个发展演变的过程。为了加深对这门课的了解，下面先来回顾一下操作系统的各个发展阶段，或者说操作系统这门学科是怎么产生的。

1.1.1 手工操作阶段

早期计算机是十分庞杂的由控制台“指挥”的机器，它使用的是一种很初级的人机交互方式。即先由人工在输入设备（纸带机、卡片机等）上把程序装入内存，然后启动执行它。通过控制台上的显示灯来监视程序的执行情况（如有错，则报错灯亮），并且直接由控制台对程序进行一些调试。其相互之间的控制关系如图 1-1 所示。

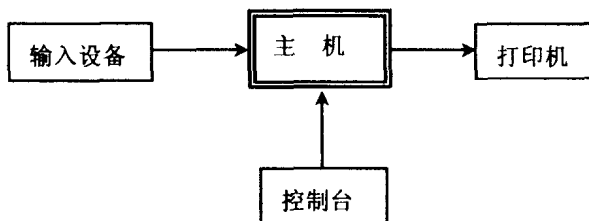


图 1-1 第一代计算机的控制关系

主机包括中央处理器（CPU）和主存储器两部分。这种使用方法最明显的特点是：①资源独占（即计算机的全部硬件资源都由一个程序独自占用）。②串行工作（人的操作与计算机的运行以及计算机各个部件之间都是按时间先后顺序工作的）。③人工干预（计算机是在人的直接联机干预下进行工作的）。

以上特点，决定它有非常严重的两个缺点：①资源浪费；②使用不便。这是为什么呢？因为资源独占和串行工作。很显然，一个程序在某一时刻不可能使用计算机的全部资源，它必须在程序全部装入内存之后，才能在主机上运行。所以当输入设备工作时，主机和输出设备都是空闲的，反之亦然。另外若是人工干预多的话，由于人工干预的操作速度远远低于主机的运算速度，而且在人工干预的操作过程中，主机必须是停止运算状态。这不仅大大降低了计算机的使用效率，也使用户感到使用不便。随着计算机运算速度大幅度提高，各种部件和设备日益增多，以及计算机应用范围的普及，上述缺点就越来越突出。

【例 1.1】 设有一程序在两个具有不同运算速度的甲、乙计算机上的运行情况如下。

	运算速度	运行时间	人工操作时间	比例
计算机甲:	1 000 OPS	1 h	5 min	12 : 1
计算机乙:	60 000 OPS	1 h	5 min	1 : 5

在计算机甲上运行这个程序时,机时的浪费还可容忍,在计算机乙上运行这个程序时,机时的浪费就不可容忍了。

为此就迫使计算机的工程研究人员花费精力尽量克服这两个缺点。他们首先想到能否缩短建立作业和人工操作的时间,因而提出了从一个作业到下一个作业的自动转换方式,从而出现了早期批处理方式。

1.1.2 早期批处理阶段

早期批处理分为联机批处理和脱机批处理两种类型,一般将完成作业间自动转换工作的程序称为监督程序。

1. 早期联机批处理

在这种方式下,操作员(或用户)只需在输入设备上装入作业信息(或程序+数据+作业说明书),过一段时间后在打印机上取执行结果。其他操作都是由机器自动进行。如机器自动输入,编译和执行程序。当一个作业完成之后,由机器(监督程序)自动调入该批的第二个作业进行操作。因每次交给系统的作业是成批的,故称为批处理(就像流水线生产一样,进去一批作业,一个接一个完成后,又顺序地将结果打印出来)。

这种方式,比早期的手工操作要先进些,因为有了监督程序,能实现一个作业到另一个作业的自动转换,从而缩短建立作业和人工操作时间。但这种方式还是存在着很严重的缺点:即将程序调入内存→计算结果→打印输出,都是由中央处理器(CPU)直接控制完成的。我们知道,CPU的速度较之输入输出(I/O)设备要快得多,那么由CPU去直接控制I/O设备,势必出现许多空等时间,就相当于将CPU的速度降之为I/O设备的速度。这种慢速外设与高速主机间串行工作的矛盾随着计算机速度的提高越来越突出。为了克服这一缺点,人们在批处理中引入了脱机技术(因为这一缺点从表面上看是由联机造成的,自然人们就想到了脱机),从而形成了早期脱机批处理。

2. 早期脱机批处理

这种方式所用的方法,就是在主机之外另设一台小型机,称之为卫星机。由卫星机与外部设备打交道,而使主机腾出较多的时间来完成一些快速的业务。如图 1-2 所示。

工作过程:读卡机上的作业通过卫星机逐个地传送到输入磁带上,而主机只负责把作业从磁带上调入内存并运行它。作业完成后,主机把计算结果和其他信息记录在输出带上,由卫星机负责把输出带上的信息读出来,交由打印机打印。这样,就使主机与慢速设备间由串行工作变成了并行工作,而且主机是与较快速的磁带机进行信息交换,从而提高了主机效率。由于卫星机仅处理简单的输入输出工作,因而只需采用较小型的机器即可。

虽然批处理系统有很明显的优点(减少人工干预,提高主机效率),但它又带来一些新的问题,如保护问题。因在这种系统中,用户程序与监督程序等一些程序之间可以相互调

用，无主次之分。这就无法防止用户程序可能破坏监督程序和系统处理程序，所以就必须要想办法保护监督程序和系统处理程序，这就过渡到了后来的执行系统。

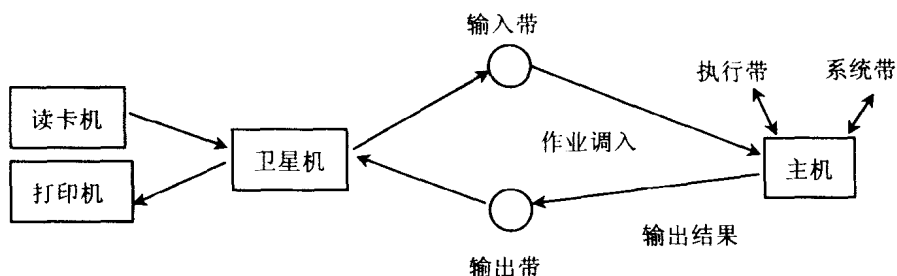


图 1-2 早期脱机批处理系统模型

1.1.3 执行系统

执行系统是伴随着通道和中断技术的出现而出现的，有了通道和中断之后，就迫使人们不得不编写中断处理程序和 I/O 控制程序来处理中断和控制通道工作。因这些程序对其他很多程序都起控制和指挥作用，需常驻内存。一般将具有这种执行程序（控制程序）的系统称为执行系统。

执行程序与监督程序的显著区别是：执行程序对其他程序拥有指挥控制权，它和用户程序之间的关系不是平等的调用关系，而是控制、被控制的关系，其他程序是在其指挥和控制下工作的。这样，系统就可以对不合法的要求进行检查，因而提高了系统的安全性。另一方面，由于执行系统发挥了通道和主机的并发性，亦提高了系统的效率（通道和中断在《计算机原理》课上已接触过，且在以后各章中还要介绍，故在此不介绍）。

不管是批处理系统还是执行系统，不管如何想办法以提高主机的效率，但由于是单道作业运行，所以主机时间的浪费仍然很严重地存在。治根要治本，要想提高主机的利用率，就必须在“单道”上下功夫，这就促进了系统的发展阶段进入到多道程序系统阶段。

1.1.4 多道程序系统阶段

所谓多道程序系统：就是能够控制多道程序并行的系统。多道程序设计的基本思想是在内存里同时存放若干道程序，它们可以并行地运行，也可以交替地运行。多道程序设计的特点是：多道程序并行。从宏观上看，多道程序都处在运行状态，它们之间是并行的；但从微观上看，每道程序又是交替地在 CPU 上运行，它们分时占有 CPU。

【例 1.2】当 CPU 对第一道程序进行处理后，需要输出时，CPU 在处理完它的 I/O 请求后就转去执行第二道程序的处理工作，这就使第一道程序的 I/O 操作与第二道程序的处理工作并行，第二道程序需要输出时，CPU 处理完它的 I/O 请求后又转向第三道程序，使第三道程序的处理工作与第一、二道程序的 I/O 操作并行……

在多道程序系统中，并发程序要共享系统内资源，使系统管理变得很复杂，从而对软硬件管理都提出了许多新课题，也促进了它们的进一步发展，如要解决系统保护、存储分

配和简单的动态地址翻译等问题。那么如何有条不紊地来管理计算机系统就形成一门新的学科脱颖而出，这就是我们正在学习的操作系统。下面解释什么叫操作系统。

1.2 什么是操作系统

1.2.1 概念

操作系统是由英文“Operating System”翻译而来的，它不是讲解操作方法的，望文生义可能会引起一些误解，平时我们常常将操作系统简称为“OS”。

一个计算机系统是由硬件和软件两大部分组成。硬件通常指诸如 CPU、存储器、外设等这样一类用以完成计算机功能的各种部件。硬件部分是计算机系统必须具备的部分，它是计算机赖以工作的基本部件，不存在无硬件部分的计算机。通常将只有硬件的计算机称为“裸机”。用户直接使用裸机是非常困难而且很不方便的。有人说：“没有软件的计算机是一堆废金属”。这话固然不很正确，但也反映了一些事实，即裸机没有什么适用范围。必须给裸机穿上“衣服”——即编制不同的软件，才能让它接待别人——即以更好的姿态面向用户。

计算机软件指为计算机编制的程序，加上执行程序时所需要的数据以及说明使用该程序的文档资料。因为程序是软件的核心部分，所以人们往往在介绍软件时只讲程序，即程序和软件是同意语。

计算机软件包括应用软件和系统软件两大部分。所谓应用软件是指针对某些特定应用领域所配置的软件。这些软件的应用范围往往要受到特定应用领域的限制（如计算机辅助设计软件 CAD，企业管理软件等）。而系统软件则不然，它是指计算机系统所必须配备的软件，通常是在各种应用领域都可通用的软件（如编译程序、连接程序、操作系统等），而操作系统又是这些软件中最基本的部分。

操作系统（OS）是管理计算机系统资源（硬件和软件）的系统软件，它为用户使用计算机提供方便、有效和安全可靠的工作环境。

对操作系统至今尚未有严格定义，但我们觉得这样一个解释，既说明了操作系统是什么，也说明了操作系统是干什么的，且说明了它的功能。但它并不十分全面，下面我们再对它作几点说明：

（1）操作系统是软件而不是硬件（但实际上它是一个软、硬结合的有机体）。操作系统这一软件的重要任务之一是管理计算机本身的机器硬件。因此，在操作系统运行和实现其功能的过程中，需要硬件强有力的支持，而且操作系统的一部分功能就是由硬件直接完成的（如中断系统中，有一部分功能就是由中断机构直接完成的）。从这个意义上讲，操作系统又不完全是软件，它是一个软、硬件结合的有机体——在软、硬件的相互配合下，共同完成操作系统所应完成的任务。因为由硬件直接完成的功能只占很少一部分，故一般还是说操作系统是软件。

(2) 操作系统是系统软件而不是应用软件,但又与其他系统软件不同。众所周知一个完善的计算机系统,通常都配有众多的系统软件,如编辑程序、编译程序等,所有这些程序,虽然与操作系统一样都属于系统软件,但它们都受 OS 的管理和控制,并得到 OS 的支持和服务。OS 可以说是这些系统软件的领导者(控制者)。

1.2.2 设置 OS 的目的

具有一定规模的现代计算机系统一般都配备有一个或几个 OS,而且 OS 的性能在很大程度上决定了计算机系统工作的优劣。那么在计算机系统中,设置 OS 的目的是什么呢?主要有两个目的:

1. 方便用户(即为用户创造良好的工作环境)

因为用户直接使用裸机非常困难且不方便。设想一下:如果所用的机器上没有装入 OS,那么将无法使用命令,不能利用应用程序,连设备、内存等都需要自己去亲自管理,这显然给用户带来一些非常繁琐的程序设计工作,这对多数用户往往是无法胜任的,用户必须掌握非常全面的计算机知识和具备很强的编程能力才行。但若在裸机的基础上设置了 OS,用户就可以用相当简便的方式,在 OS 的帮助下进行 I/O 操作,摆脱了非常繁琐的程序设计工作,所以从用户角度看,OS 是用户和裸机之间的一个界面。用户通过这一界面能方便地使用本来很难使用的计算机,也就是说 OS 向用户提供了一个方便而且强有力的使用环境。

2. 充分发挥计算机中各种资源的效率

这是从另一个观点(即资源管理观点)来看待 OS。因为 OS 是管理计算机系统中各种资源的软件,如果把一个计算机系统比作一个“国家”,则 OS 可以说是这个国家的“政府机构”。因现代计算机系统通常都是多道程序系统,所以 OS 就必须在多道程序之间合理地分配和回收各种资源,使资源得到合理有效的使用,使程序得以有条不紊地运行。

1.2.3 操作系统的主要功能

1. 操作系统的功能要点

操作系统既要管理资源,又要为用户服务,所以,系统资源管理和提供用户界面是操作系统的功能要点。在资源管理中,操作系统的任务是使各种系统资源(硬件和软件资源)得到充分的、合理的使用,解决用户作业因争夺资源而产生的矛盾。于是,对资源的监视、分配和回收就是每一个操作系统资源管理程序的设计目标。

1) 监视资源

操作系统作为用户作业的宏观调控者,它必须时刻保持系统资源分配的全局信息,了解系统资源的总数,已分配和未分配的资源情况,资源的增减和变动情况,每类资源所具有的特点和适应性。这些资源信息是通过操作系统中各类数据结构和表格记录下来,并且在系统运行过程中不断更新。

2) 分配资源

操作系统必须对来自用户和应用程序的资源使用请求作出快速的响应,适当地处理这