

UNIX系统

初级教程

刘彦明 编



第 010 字登健(夾)

刘彦明 编

西安电子科技大学出版社

1996

(陕)新登字 010 号

内 容 简 介

本书主要介绍 UNIX 系统的基础知识,包括三部分内容:操作系统的基础知识、UNIX 系统的基本使用方法(UNIX 系统入门知识)和 UNIX 系统最基本的编程方法。

本书以初学 UNIX 系统的读者为对象,深入浅出地介绍 UNIX 系统最基本的使用环境。为便于读者理解、掌握,书中安排了大量的应用实例。

本书既可作为科技工作者掌握 UNIX 系统的入门读物,也可作为高等学校高年级学生学习 UNIX 系统的参考书。

UNIX 系统初级教程

刘彦明 编

责任编辑 杨 兵 云立实

西安电子科技大学出版社出版发行

西安长青印刷厂印刷

新华书店经销

开本 787×1092 1/16 印张 23 2/16 字数 551 千字

1996 年 5 月第 1 版 1996 年 6 月第 1 次印刷 印数 1-6 000

ISBN7-5606-0442-0/TP·0191 定价:23.00 元

前 言

在计算机技术飞速发展的今天,软件的使用在计算机应用中变得越来越重要和受人关注。UNIX 系统作为一种操作系统软件,从一开始就以其高效的资源管理能力、简洁的用户界面、良好的安全性等特点受到广大计算机用户的欢迎,特别是 UNIX 系统的开放性,使其发展十分迅速。此外,UNIX 系统几乎可以安装到现有的所有硬件平台上,加之在 UNIX 系统上开发的应用软件非常丰富,因而 UNIX 系统的应用十分广泛。

由于 UNIX 业界的统一协调,使得 UNIX 系统在工作站以上档次的计算机系统中的地位十分稳定。

鉴于上述种种理由,并为进一步推广普及 UNIX 系统,作者编写了本书。本书以 UNIX System V Release 4、BSD 4. x 和 SunOS 4. 1. x 为背景介绍 UNIX 系统使用的基础知识,旨在把读者引入 UNIX 系统的世界。

UNIX 系统是一个较庞大的操作系统,要掌握它必须采用循序渐进的方法,因此本书将介绍 UNIX 系统中最简单、最常用的部分。全书内容分成两篇来介绍。

绪论,主要介绍操作系统的基础知识。

第一篇,UNIX 系统初级用户指南。主要介绍作为 UNIX 系统的初级用户如何进入/退出 UNIX 系统,并介绍一些 UNIX 系统最常用的实用程序(也称为命令),它包括 7 章:

第一章,UNIX 系统概述。主要介绍 UNIX 系统的组成。

第二章,UNIX 系统用户入门。主要介绍 UNIX 系统的用户基础知识和如何进入(注册)/退出(注销)UNIX 系统。

第三章,文件系统的使用。主要介绍 UNIX 系统中的文件系统及其使用。

第四章,UNIX 系统常用命令。主要介绍 UNIX 系统中最常用也是必须掌握的实用程序(命令)的使用方法。

第五章,vi 编辑工具。主要介绍在字符终端上,如何使用 vi 编辑器编辑文本文件(包括语言源程序)。

第六章,shell 编程。主要介绍如何使用 shell 管理文件、处理文件内容,以及将命令组合成 shell 程序。

第七章,UNIX 系统通信。主要介绍如何使用 UNIX 系统提供的命令完成 UNIX 系统间的数据通信。

第二篇,UNIX 系统初级程序员指南。主要介绍在 UNIX 系统环境下,如何用 C 语言编写程序,它包括 8 章:

第八章,UNIX 系统上的 C 编程环境。主要介绍如何使用 C 语言编译器 cc 来编译 C 语言源程序;如何使用命令行参数和如何取得 UNIX 系统的环境变量等内容。

第九章,文件基本操作。主要介绍如何使用 UNIX 系统提供的文件操作系统调用,实施对文件的读写等基本操作。

第十章,文件操作的进一步讨论。主要介绍文件的保护和控制,以及目录文件的管理等内容。

第十一章,进程。主要介绍进程的创建和控制,包括 fork、system、exec、wait 等与进程

有关的系统调用。

第十二章，UNIX 系统的管道。主要介绍 UNIX 系统的管道通信机构，以及如何利用 UNIX 系统提供的系统调用实现管道上的编程。

第十三章，信号。主要介绍 UNIX 系统的信号机构，以及如何利用 UNIX 系统提供的有关系统调用实现其上的程序设计。

第十四章，进程通信。主要介绍如何使用 UNIX 系统提供的系统调用实现进程间的通信。进程通信包括三种通信机制：消息通信、信号量通信和共享内存段通信。

第十五章，UNIX 系统的源代码调试器——dbx。主要介绍如何使用 dbx 来调试程序，找到程序出现错误的位置。

在本书的最后给出了几个附录。

本书在编写上注重实用性，由浅入深，力求使读者在最短的时间内掌握 UNIX 系统最基本的知识。此外，为便于读者学习，书中给出了大量的应用实例。

本书在编写过程中得到了西安电子科技大学 ISN 国家重点实验室的大力支持，其中 ISN 国家重点实验室副主任张高樵副教授给予了热情帮助，在此表示衷心的感谢！

鉴于作者水平有限，难免有不妥之处，欢迎读者指正。

编者

1995 年 10 月

目 录

绪论	1
----------	---

第一篇 UNIX 系统初级用户指南

第一章 UNIX 系统概述	13
§ 1.1 UNIX 系统的硬件环境	13
§ 1.2 UNIX 系统的组成	14
§ 1.3 UNIX 系统核心的体系结构	15
§ 1.4 文件系统	16
§ 1.5 命令行句法与命令执行过程	17
§ 1.6 UNIX 系统提供的服务	18
§ 1.7 UNIX 系统的设计思想	18
§ 1.8 UNIX 系统的历史与发展	19
§ 1.9 UNIX System V Release 4 简介	19

第二章 UNIX 系统用户入门	22
§ 2.1 入门	22
§ 2.2 终端	22
§ 2.2.1 所需终端的设置	22
§ 2.2.2 键盘特性	23
§ 2.2.3 键入约定	23
§ 2.3 进入系统前的准备	26
§ 2.4 获得注册名	27
§ 2.5 与 UNIX 系统建立联系	27
§ 2.5.1 注册过程	27
§ 2.5.2 口令	27
§ 2.5.3 注册时的故障排除	29
§ 2.5.4 注销	30

第三章 文件系统的使用	31
§ 3.1 文件系统的构成	31
§ 3.2 用户在文件系统中的位置	31
§ 3.2.1 用户的起始目录	31
§ 3.2.2 用户的当前目录	32
§ 3.2.3 路径名	32
§ 3.3 目录组织	33
§ 3.3.1 建立目录	33
§ 3.3.2 列出目录内容	34
§ 3.3.3 改变当前工作目录	35
§ 3.3.4 删除目录	35
§ 3.4 文件操作的常用命令	35
§ 3.4.1 查找文件——find	35

§ 3.4.2 显示文件内容	37
§ 3.4.3 产生文件备份	42
§ 3.4.4 文件移动和改名	43
§ 3.4.5 删除文件	44
§ 3.4.6 计算文件行数、词数和 字符数	45
§ 3.4.7 文件保护	46
§ 3.5 文件操作的高级命令	49
§ 3.5.1 识别文件间的差别	50
§ 3.5.2 文件中模式搜索	50
§ 3.5.3 排序和合并文件	51

第四章 UNIX 系统常用命令	53
§ 4.1 口令的设置	53
§ 4.2 谁正在使用 UNIX 系统	54
§ 4.3 通信命令	54
§ 4.3.1 使用 write 发送报文	54
§ 4.3.2 使用 talk 与另一用户对话	56
§ 4.3.3 使用 msg 拒绝和允许 接收报文	59
§ 4.3.4 使用 mail 发送和接收 电子邮件	59
§ 4.4 打印文件	65
§ 4.4.1 发送一个打印请求	65
§ 4.4.2 取消打印请求	68
§ 4.5 查询打印服务状态	68
§ 4.6 状态信息查询	70
§ 4.6.1 使用 date 显示时间和日期	70
§ 4.6.2 报告磁盘利用情况	71
§ 4.6.3 进程及其状态	73
§ 4.6.4 终止进程执行	76
§ 4.7 在终端上显示正文或 变量的内容	77
§ 4.8 显示系统命令的使用手册	78
§ 4.9 文件传输	78
第五章 vi 编辑工具	86
§ 5.1 vi 的工作模式	86

§ 5.2 一个完整的编辑过程	87	§ 6.1.3 输入输出改向	102
§ 5.3 启动 vi	88	§ 6.1.4 执行、停止和重新启动进程	103
§ 5.4 基本命令模式	89	§ 6.2 shell 编程	105
§ 5.4.1 移动光标	89	§ 6.2.1 shell 程序	105
§ 5.4.2 搜索命令	91	§ 6.2.2 变量	106
§ 5.4.3 删除正文	92	§ 6.2.3 shell 编辑结构	109
§ 5.4.4 替代命令	93	§ 6.2.4 shell 程序排错	116
§ 5.4.5 修改正文	94	§ 6.3 shell 程序举例	117
§ 5.4.6 进入插入模式	95	第七章 UNIX 系统通信	125
§ 5.4.7 进入交互式命令模式	95	§ 7.1 发送文件	125
§ 5.4.8 其它命令	95	§ 7.1.1 uucp 命令	125
§ 5.5 交互式命令模式	96	§ 7.1.2 uuto 命令	128
§ 5.5.1 针对当前编辑文件的命令	96	§ 7.1.3 uustat 命令	129
§ 5.5.2 使用 shell 命令	96	§ 7.1.4 uupick 命令	129
§ 5.5.3 退出 vi	97	§ 7.2 联网	130
§ 5.5.4 其它命令	97	§ 7.2.1 联接远程终端——ct 命令	131
第六章 shell 编程	99	§ 7.2.2 呼叫另一 UNIX 系统 ——cu 命令	132
§ 6.1 shell 命令语言	99	§ 7.2.3 在远程系统上工作 ——uux 命令	134
§ 6.1.1 元字符	99		
§ 6.1.2 特殊字符	100		

第二篇 UNIX 系统初级程序员指南

第八章 UNIX 系统上的 C 编程环境	137	第九章 文件基本操作	161
§ 8.1 编译和运行 C 程序	137	§ 9.1 文件基本操作简介	161
§ 8.2 C 语言编译器——cc	138	§ 9.2 文件的创建和删除	162
§ 8.2.1 cc 简介	138	§ 9.2.1 creat 系统调用	162
§ 8.2.2 cc 编译器的选择项	138	§ 9.2.2 unlink 系统调用	164
§ 8.2.3 举例	139	§ 9.3 文件的打开和关闭	165
§ 8.3 程序环境	140	§ 9.3.1 open 系统调用	165
§ 8.3.1 命令行参数	140	§ 9.3.2 close 系统调用	168
§ 8.3.2 存取环境变量	141	§ 9.3.3 应用实例	168
§ 8.4 C 语言库	143	§ 9.4 文件的读/写操作	170
§ 8.4.1 头文件 stdio.h	143	§ 9.4.1 read 系统调用	170
§ 8.4.2 头文件 math.h	145	§ 9.4.2 write 系统调用	170
§ 8.4.3 头文件 ctype.h	148	§ 9.4.3 应用实例	171
§ 8.4.4 头文件 fcntl.h	148	§ 9.5 文件的随机存取	173
§ 8.4.5 头文件 memory.h	152	§ 9.5.1 lseek 系统调用	173
§ 8.4.6 头文件 malloc.h	152	§ 9.5.2 lseek 应用实例	173
§ 8.4.7 头文件 time.h	153	§ 9.6 文件基本操作应用举例	176
§ 8.4.8 头文件 stdlib.h	154	§ 9.7 标准 I/O 库	181
§ 8.4.9 头文件 string.h	154	§ 9.7.1 流和 FILE 结构	182
§ 8.5 C 语言程序示例	156	§ 9.7.2 标准 I/O 库例行程序的使用	183

§ 9.8 实例程序	191	第十三章 信号	268
第十章 文件操作的进一步讨论	194	§ 13.1 信号及其类型	268
§ 10.1 文件系统结构	194	§ 13.2 信号的处理	270
§ 10.1.1 磁盘文件系统结构	194	§ 13.3 kill、alarm 和 pause 系统调用	275
§ 10.1.2 文件系统的数据结构	196	§ 13.3.1 kill 系统调用	275
§ 10.2 文件的保护和控制	201	§ 13.3.2 系统调用 alarm	277
§ 10.2.1 文件保护	201	§ 13.3.3 系统调用 pause	279
§ 10.2.2 文件控制	205	§ 13.3.4 例行程序 setjmp 和 longjmp	281
§ 10.3 目录文件管理	211	§ 13.4 应用举例	282
§ 10.3.1 目录的创建和删除	211	第十四章 进程通信	285
§ 10.3.2 目录的改变和链接	214	§ 14.1 进程通信基础知识	285
§ 10.3.3 目录的读取	216	§ 14.2 消息通信	286
§ 10.4 文件信息查询	218	§ 14.3 信号量	298
§ 10.4.1 文件状态信息的获取	219	§ 14.4 共享内存段	314
§ 10.4.2 stat、fstat 应用实例	220	§ 14.5 综合实例	327
§ 10.5 综合应用实例	221	第十五章 UNIX 的源代码调试器	
§ 10.5.1 文件服务实用工具设计	221	——dbx	336
§ 10.5.2 UNIX 中误删文件的恢复	228	§ 15.1 dbx 概述	336
第十一章 进程	232	§ 15.1.1 dbx 的使用格式	336
§ 11.1 进程的基本概念	232	§ 15.1.2 dbx 中的命令	337
§ 11.1.1 进程图像的组成	232	§ 15.1.3 dbx 的调试方式	337
§ 11.1.2 进程状态及其转换	233	§ 15.2 dbx 使用前的准备	338
§ 11.1.3 进程控制	234	§ 15.3 启动 dbx	338
§ 11.2 system 系统调用	234	§ 15.4 显示源文件(源程序)	340
§ 11.3 进程创建	235	§ 15.5 显示活动或历史过程	340
§ 11.4 低级进程运行	236	§ 15.6 命名和显示数据	341
§ 11.4.1 exec 系统调用	236	§ 15.7 设置断点	342
§ 11.4.2 exec 应用实例	238	§ 15.8 运行和跟踪程序	343
§ 11.5 进程控制	242	§ 15.9 存取源程序和改变	
§ 11.6 应用实例	243	路径(当前目录)	344
第十二章 UNIX 系统的管道	254	§ 15.10 其它命令	345
§ 12.1 pipe 管道通信设备	255	§ 15.11 dbx 使用举例	346
§ 12.1.1 pipe 系统调用	255	附录 A UNIX 系统的主要命令	351
§ 12.2 pipe 应用实例	2256	附录 B UNIX 系统的程序设计环境	356
§ 12.3 FIFO 管道通信机制	263	附录 C 软盘的使用	358
§ 12.3.1 FIFO 及其应用	263	附录 D UNIX 系统的引导和初始化	360
§ 12.3.2 FIFO 程序设计	264	参考文献	362
§ 12.3.3 FIFO 应用实例	265		

绪 论

计算机系统由硬件和软件两大部分组成，通常把未配置软件的计算机称为裸机。计算机软件分为系统软件和应用软件两类。系统软件是指操作系统以及与计算机语言相关的编辑、编译系统，应用软件是指由用户自己开发的应用程序，它们之间的关系如图 0-1 所示。

从图 0-1 可以看出，操作系统是裸机上的第一层软件，它是对计算机硬件系统功能的首次扩充。操作系统在计算机系统中占有特殊和重要的地位，所有其它软件，如编辑、编译等系统软件以及大量的应用软件都是建立在操作系统之上，并得到它的支持和取得它的服务而实现的。

从用户的角度来看，当计算机配置了操作系统之后，用户不再直接使用计算机的硬件资源，而是通过操作系统提供的命令和服务去操纵计算机，因此操作系统被看作是用户和计算机之间的接口。

事实上，操作系统已成为现代计算机系统中必不可少的最重要的组成部分。可以这样说，没有操作系统的计算机是一堆“垃圾”。

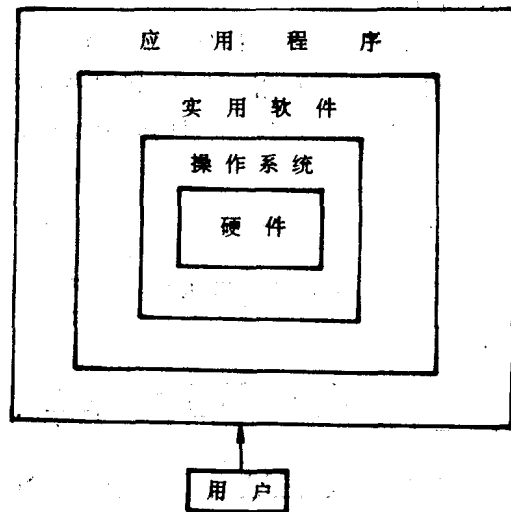


图 0-1 计算机系统组成

一、什么是操作系统

在计算机上配置操作系统的目的有三个：

- ① 提供一个计算机用户与计算机硬件之间的接口，使计算机系统更易于使用。
- ② 有效地管理和控制计算机系统内的各种软件和硬件资源，使之得到更有效的利用。
- ③ 合理地组织计算机系统的工作流程，以改善系统的性能。

据此，我们可以给操作系统下一定义：操作系统是指管理和控制计算机资源（包括硬件和软件资源），合理组织计算机系统的工作流程，方便用户使用的计算机软件集合。

二、如何评价操作系统

衡量操作系统的标准有以下几点：

1. 方便性

方便性是指用户使用计算机系统的方便程度。用户使用计算机的目的是要让计算机为

用户服务, 因此操作系统必须提供方便的用户使用接口。操作系统为用户提供了下列服务:

- ① 建立程序。通过编辑、编译系统建立可执行程序。
- ② 执行程序。操作系统提供了在系统提示符下执行可执行程序的服务。
- ③ 设备访问。操作系统支持用户访问(使用)输入输出设备(包括显示器、键盘、磁盘、打印机等)。
- ④ 系统访问。用户可以使用操作系统提供的软件资源(如系统调用和库例程等)。
- ⑤ 系统检测。操作系统可以检测到软件和硬件错误, 并给出检测结果。
- ⑥ 系统记录。收集资源的使用情况, 监视计算机系统的性能。

2. 有效性

操作系统作为计算机硬件资源和软件资源的管理者, 必须对它们进行有效的控制和管理, 使之得到有效的利用。计算机硬件资源包括内存、CPU 和外设。图 0-2 说明了操作系统作为用户和计算机之间接口的分层结构, 图 0-3 说明了操作系统作为资源管理者的作用。

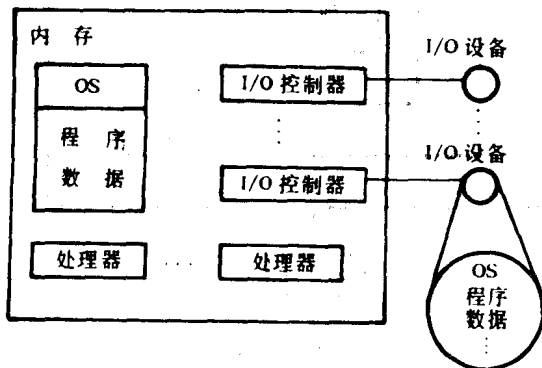


图 0-2 计算机系统的分层结构

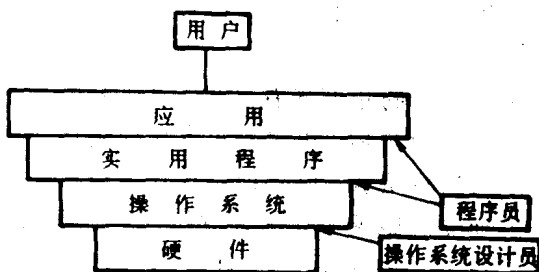


图 0-3 操作系统作为资源管理者

3. 易扩充性

由于操作系统是根据当时计算机硬件的水平来设计的, 因此当硬件发展时, 操作系统应能适应新的环境并作相应的改进。为此, 设计的操作系统必须具有易扩充性, 不改变操作系统的整体结构, 只对相应于硬件的部分加以改进, 即可适应新的硬件环境。其次, 随着时间的推移, 人们需要操作系统提供新的服务, 这时作为好的操作系统应能较容易地增加这些新的服务。第三, 由于操作系统设计时可能有缺陷, 因此操作系统应便于维护。

三、操作系统的特征

现代操作系统具有以下基本特征:

1. 并发性

并行性和并发性是既相似又有区别的两个概念。并行性是指两个或多个事件在同一时刻发生; 并发性是指两个或多个事件在同一时间间隔内发生。并发性是指宏观上在一段时间内有多个事件发生, 能并发执行的程序称为并发程序, 相应的计算机系统称为并发系统。

2. 共享性

共享是指系统中的软硬件资源不再为某个程序所独占,而是供多个程序共同使用。

并发和共享是操作系统的两个最基本的特征,而这两者之间又是互为存在的条件。一方面,资源共享是以程序的并发执行为条件,若计算机系统不允许程序并发执行,自然就不存在资源共享这一特征了;另一方面,若计算机系统不能对资源共享实施有效的管理,也必然影响程序的并发执行,甚至根本无法并发执行。

3. 虚拟性

在操作系统中,所谓虚拟是指把一个物理上的实体变为若干个逻辑上的对应物。前者是实际存在的,后者是虚的,只是使用者的一种感觉。例如,在多道程序系统(在计算机系统中同时可以有多个程序并发执行)中,虽然只有一个 CPU,某一时刻只能执行一个程序,但是通过分时使用 CPU,则可在一段时间间隔内有多个程序都处在其开始点和结束点之间。这样,宏观上这台处理机就能同时运行多个程序,给用户的感觉是每个程序都有一个 CPU 为其服务。

4. 不确定性

在操作系统中,不确定性有两种含义:

① 程序执行结果是不确定的。即对同一程序,使用相同的输入,在相同的运行环境下运行,却可能得到完全不同的结果,亦称程序是不可再现的。

② 多个程序在并发执行时是以异步方式进行的。换言之,每个程序何时执行,以及完成各个程序所需要的时间是不确定的,因此也是不可预知的。

上述第一种不确定性是绝对不允许的,因而也是操作系统必须解决的主要问题,第二种不确定性却是允许的。

四、操作系统的功能

在多道程序环境下,用户提交给系统的作业所需资源的总和远多于系统所拥有的资源,系统必然会因不能同时满足所有作业的要求而无法将它们都开动起来。于是,为了使自己能投入运行,这些作业之间必将争夺资源。为了使多道程序能有条不紊地运行,应使操作系统具有以下五个方面的功能。

1. 存储器管理

存储器管理的主要任务是:

- ① 为多道程序的运行提供良好的运行环境。
- ② 方便用户使用存储器。
- ③ 提高存储器的利用率。
- ④ 为尽量多的用户提供足够大的存储空间。

为实现上述任务,存储器管理应具有以下四种功能:

(1) 内存分配:多道程序能并发执行的首要条件,是各道程序都有自己的内存空间,因此为每道程序分配内存是存储管理的最基本功能。

(2) 内存保护:为保证每道程序都能在自己的内存空间运行而互不干扰,要求每道程序执行时能随时检查对内存的访问是否合法。必须防止因一道程序的错误而扰乱了其它程序,尤其应防止用户程序侵犯操作系统的内存区。

(3) 地址映射：由目标程序所确定的地址范围称为该程序的地址空间。地址空间中的地址是逻辑地址，内存空间是内存物理地址的集合。在多道程序的系统中，操作系统必须提供可以把程序地址空间中的逻辑地址转换为内存空间对应的物理地址的功能，我们把这种地址变换过程称为地址映射。显然，地址映射功能可使用户不必过问物理存储空间的分配细节，从而为用户编程提供了方便。

(4) 内存扩充：由于物理内存的大小可能限制了大型作业或多个作业的并发执行，为了满足用户的要求并改善系统性能，必须对内存加以扩充。但是，我们无法真正地增加内存空间，而是借助于虚拟存储技术获得这样的效果，从而使系统能运行内存要求量远比物理内存大得多的作业，或让更多的作业并发执行。

2. 处理机管理

处理机管理的主要任务是对处理机分配和运行实施有效的管理。在多道程序环境下，处理机的分配和运行都是以进程为单位进行的。因此，对处理机的管理可归结为对进程（程序的一次执行过程）的管理。进程管理应实现下述主要功能：

(1) 进程控制：当用户作业要运行时，应为之建立一个或多个进程，并为它分配除处理机之外的所有资源，将它放入进程就绪队列。当进程运行完成时，应立即撤销该进程，以便及时释放其所占有的资源。进程控制的基本功能就是创建和撤销进程，以及控制进程的状态转换。

(2) 进程同步：所谓进程同步，是指系统对并发执行的进程进行协调。最基本的进程同步方式，是使诸进程以互斥的方式访问临界资源。此外，对于彼此合作去完成共同任务的进程，应由系统对它们的运行速度加以协调。

(3) 进程通信：对于互相合作的进程，在它们运行时相互之间往往要交换一定的信息，我们把这种进程间进行信息交换称为进程间的通信。

(4) 进程调度：当一个正在执行的进程已经完成，或因某件事而无法继续执行时，系统应进行进程调度，重新分配处理机。进程调度是指按一定算法，从进程就绪队列中选出一进程，把处理机分配给它，为该进程设置运行现场，并使之投入运行。

3. 设备管理

设备管理的主要任务是：

- ① 为用户程序分配 I/O 设备。
- ② 完成用户请求的 I/O 操作。
- ③ 提高 CPU 和 I/O 设备的利用率。
- ④ 改善人机界面。

为实现上述任务，设备管理程序应具有下述功能：

(1) 缓冲管理：几乎所有的外围设备与处理机交换信息时，都要利用缓冲来缓和 CPU 和 I/O 设备间速度不匹配的矛盾，提高 CPU 与设备、设备与设备间操作的并行程度，从而提高 CPU 和 I/O 设备的利用率。因此，在系统中设置了许多类型的缓冲，系统必须对它们进行有效的管理。

(2) 设备分配：系统根据用户请求的设备类型和所采用的分配算法对设备进行分配，并将未获得设备（进程所要求的设备）的进程放入相应设备的等待队列。

(3) 设备处理：启动指定的 I/O 设备，完成用户规定的 I/O 操作，并对由设备发来的中

断请求进行及时响应,根据中断类型进行相应的处理。

(4) 虚拟设备功能:通常,把一次仅允许一个进程所使用的设备称为独占设备。系统可通过某种技术使该设备成为能被多个进程共享的设备,以提高设备的利用率和加速程序的执行过程,使每个用户都感觉到自己在独占该设备。如前所述,这种只为用户感觉到而实际上并不存在的设备,称为逻辑设备或虚拟设备。

4. 文件管理

在现代计算机系统中,总是把大量信息以文件形式存放在外存中,以供所有的或指定的用户使用。系统也允许用户把处理结果保存在系统中,供以后使用。这样,操作系统中又必须配置文件系统,作为信息管理机构。

文件系统应具有下述功能:

(1) 文件存储空间的管理:通常,系统中的所有文件都存放在外存中,供多个用户使用并共享。若由用户自己对文件的存储进行管理,不仅极其困难,而且也是非常低效的,这就要求由文件系统对文件存储空间进行统一管理,包括对文件存储空间的分配和回收。

(2) 目录管理:为使用户能方便地找到其所需文件,通常由系统为每个文件建立一个文件目录项,包含文件名、文件属性、文件所在的物理位置及有关其它信息。若干个目录项又可以构成一个目录文件。对于大型文件系统,由于目录项极多,文件系统应妥善地对它们加以组织,以加速文件的查找,并获得令人满意的文件检索速度。

(3) 文件读、写管理:系统在对文件进行读操作时,根据用户给定的目的地址、传送字节数,把文件信息从外存读入缓冲区后,再复制到指定的用户区或系统区。文件的写操作是根据用户指定的源地址、传送字节数,把信息从指定的用户区写到磁盘上。

(4) 文件保护:为防止文件被偷窃或破坏,文件系统必须提供文件保护功能以实现下述目标:①防止未经核准的用户存取文件;②防止冒名顶替存取文件;③防止核准用户以不正确的方式使用文件。在一个文件系统中,可采取多种文件保护措施来达到上述目标,如用口令、存取权限以及对文件进行加密等。

(5) 向用户提供接口:为了方便用户,操作系统应向用户提供一个使用方便的接口,用户通过该接口取得文件系统的服务,如建立、打开、关闭、撤销(删除)文件,以及文件的读写操作等系统调用,供编程时使用。

5. 作业管理

作业管理的主要任务是根据系统条件和用户需要,对作业的运行进行合理的组织及相应的控制。

作业管理应具备两方面的功能:

(1) 作业调度:作业调度是指根据系统的能力和当前作业的运行情况,按一定策略从后备作业队列中选出一批作业,为它们分配所需的 I/O 设备和存储空间,将它们调入内存并为之建立相应的进程,使其成为具有获得处理机资格的候选进程。

(2) 作业控制:作业控制是指从作业进入系统开始,直到运行完成的整个过程中,用户可以通过某种形式向系统发出各种命令,对自己的作业进行控制和管理。在实际系统中,作业有两种类型:批量型作业和终端型作业。由于这两种作业有不同的输入和运行方式,相应地就形成了两种不同的作业控制方式,分别称之为脱机控制方式和联机控制方式。

五、操作系统的分类

操作系统的分类方法很多,我们根据操作系统所配置计算机拥有的处理机个数,把操作系统分成单处理机操作系统和多处理机操作系统。

1. 单处理机操作系统

单处理机操作系统是指这种操作系统只能运行在单个处理机的计算机系统上。单处理机操作系统又分为单用户操作系统和多用户操作系统。

(1) 单用户操作系统:单用户操作系统是指只能为单个用户使用的操作系统,这种操作系统一般运行在微机上,因此也称为微机操作系统。典型的单用户操作系统有以下几种:

① CP/M,它是由 Digital Research 公司推出的 8 位微机操作系统,一般运行在苹果机上。

② MS-DOS,它是由 Microsoft 公司推出的微机操作系统,一般运行在 PC 机上。

③ OS/2,它是由 IBM 公司推出的,运行在 16 位以上的微机上,是一种单用户、多任务的微机操作系统。一些专家指出,OS/2 将成为下一代主机操作系统之一。

(2) 多用户操作系统。多用户操作系统是指能为多个用户同时使用计算机系统提供支持的操作系统。一般来说,多用户操作系统也具有多任务的特性。典型的多用户操作系统有:

① 多用户微机操作系统,包括 MP/M、Xenix、Uniflex、Linux 等。

② UNIX 系统。该系统既有微机版本,又有小型机版本,也可以运行在大型机和巨型机上,是当今世界最流行的多用户操作系统。

③ Windows 95。最新推出的 Windows 95 是基于微机的多用户多任务操作系统,具有较好的图形用户界面和连网能力。

2. 多处理机操作系统

提高计算机性能和可靠性的重要途径,是改善计算机的体系结构。在 1975 年前后,终于打破了以单处理机体系结构为主的局面,形成了由多台处理机通过互连网络连接在一起的计算机系统。近年来所推出的大、中、小型机都在采用多处理机体系结构。

根据多个处理机之间耦合的紧密程度,可把多处理机结构分成两大类:

(1) 紧耦合多处理机结构:这种结构的多处理机通常采用高速总线或高速交叉开关进行处理机之间的互连,并共享主存储器,由所配置的操作系统统一控制所有的处理机及系统硬件。图 0-4 示出了利用高速交叉开关连接成的紧耦合多处理机系统。

(2) 松耦合多处理机结构:在松耦合系统中,多个处理机之间通常通过通道或通信线路实现互连,每个处理机有自己的存储器和操作系统,能独立地工作,必要时通过彼此间的通信来交换信息及协调它们之间的关系。图 0-5 示出了通过公用总线互连的松耦合多处理机系统。

一般来说,这两种体系结构的计算机系统所配置的操作系统分别是分布式操作系统和计算机网络操作系统。

3. 网络操作系统

(1) 计算机网络的类型:计算机网络是指通过数据通信系统,把地理上分散的计算机

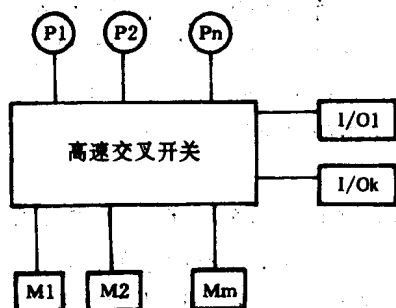


图 0-4 紧耦合多处理机系统

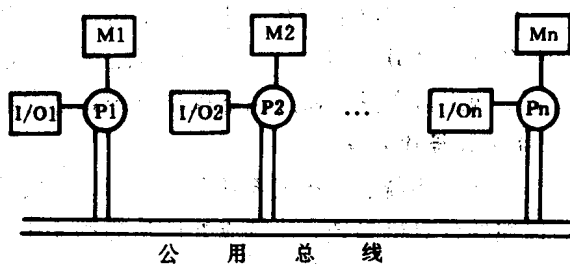


图 0-5 松耦合多处理机系统

群和终端设备连接起来,以达到数据通信和资源共享目的的一种计算机系统。可见,计算机网络是高度发展的计算机技术和通信技术相互结合的产物。按照网络所覆盖的地理范围和互连计算机之间的距离的不同,可把计算机网络分成广域网和局域网两种:

① 广域网(WAN)。广域网也称运程网,它所覆盖的地理范围常为几公里到几千公里,所连接的计算机大多是大、中、小型计算机。

② 局域网(LAN)。互连计算机之间的距离常为几十米至几公里,所连接的计算机主要是微机,也有少量小型机或中型机。

(2) 计算机网络的基本特征:

① 自治性。每台计算机有自己的内存、I/O 设备和操作系统,因而具有较强的自治能力,能独立承担分配给它的任务。

② 分布性。分布性包括三方面的含义:一是地理分布,即系统中的所有计算机在地理位置上是分散的;二是功能分布,即各个计算机可独立承担任务;三是任务分布,即可把一个大任务划分成若干个小任务,分别由不同的计算机去执行。

③ 互连性。把地理上分布的资源(包括硬件资源和软件资源),在物理上和逻辑上连接在一起,并通过通信子网相互作用。

④ 统一性。在统一的网络操作系统的控制下,使分散的计算机系统能相互协调,实现网络通信和资源共享,完成各自的或共同的任务,并向网络用户提供统一的接口。

(3) 网络操作系统的功能和定义:网络操作系统(NOS)是网络范围的操作系统,其主要任务是用统一的方法管理整个网络中共享资源的使用和任务的处理,为此它应具有以下四个基本功能:

① 网络通信。网络通信的任务是实现计算机之间的无差错数据传输。其主要功能是为通信双方建立和拆除通信链路;对传输过程中的数据差错进行检查和校正,并使收发速率匹配。

② 资源管理。采用统一的、有效的方法协调诸用户对共享资源的使用,使用户使用远地资源也像使用本地资源一样。对资源的具体管理和控制,仍由主机操作系统实现。

③ 提供多种网络服务。向网络用户提供的主要服务有:电子邮件服务;文件传送、存取和管理服务;作业传送和操纵服务等。

④ 提供网络接口。向网络用户提供统一、经济地使用网络共享资源和取得网络服务的网络接口,使用户无须了解很多共享资源的属性和有关网络的知识。

综上所述,我们可以把网络操作系统定义为:网络操作系统是建立在主机操作系统基础之上,用于管理网络通信和共享资源,协调各主机上任务的运行,并向用户提供统一而有效的网络接口的软件集合。可见,网络操作系统是用户(或用户程序)和主机操作系统之间的接口。网络用户只有通过网络操作系统,方能取得网络所提供的各种服务。

4. 分布式操作系统

(1) 分布式系统:分布式系统是 70~80 年代发展起来的一种多处理机(或多计算机)系统。以往的计算机系统,计算或处理功能都集中在一台主机上,所有的任务也都由主机处理,所以称之为集中式处理系统。而在分布式处理系统中,计算和处理功能分散在构成分布式系统的各个处理单元上,相应地可把一个大任务划分成可以并行执行的多个子任务,并动态地把这些任务分配到各处理单元上去,由它们单独执行。可见,分布式处理系统最基本的特征是处理上的分布,而处理上的分布是功能和任务分布。分布式处理系统又可称为分布式系统。

由此可知,分布式系统是指由多个处理单元构成的系统,其中每个处理单元都包含有处理机和局部存储器,它们能独立承担系统分配给它们的任务。各个处理单元通过互连网络连接在一起,在统一的分布式操作系统控制和管理下,实现各处理单元间的通信、资源共享、动态分配任务和对任务进行处理。

在分布式系统中,如果处理单元是计算机,则该系统可称为分布式计算机系统,计算机网络是一种最常见的分布式计算机系统。如果处理单元是处理机和局部存储器,则只能称作分布式(处理)系统,当其为紧耦合结构时,则不是计算机网络。

(2) 分布式系统的特征:

① 分布性。分布式系统具有功能分布和任务分布的特征,对于松耦合系统还可能呈现地理分布特征。分布式系统与计算机网络相比较,更着重于任务的分布性,因而它具有更完善的任务分配功能。

② 自治性。不论分布式系统中的处理单元采用何种形式,它至少应包括处理机和局部存储器,具有独立执行任务的能力,因而每个处理单元都具有一定的自治性,但通常不具有各自独立的操作系统。

③ 模块性。分布式系统的结构已趋向于模块化,即系统由若干个结构完全相同、具有相同功能的处理单元构成,因此任务可以分配到其中的任一个处理单元上执行。模块性结构简化了系统的制作,提高了系统的性能和可靠性。

④ 并行性。由于有多个处理单元,因此可把一个作业的诸任务分配到多个处理单元上进行并行处理。在分布式系统中,还可按流水线方式使一道程序的各个程序段并行执行。

(3) 分布式操作系统的功能:

① 资源管理。能对并行工作的大量处理单元、存储器和 I/O 设备等资源进行有效的动态管理。由于系统的模块性,不仅简化了对系统资源的管理,而且提高了资源利用率。

② 任务分配。在分布式系统中,任务的分配不是以一个任务为单位,而是以一组能并行执行的任务集为单位,同时将它们分配到多个处理单元上,使之能并行执行。

③ 分布式进程同步和通信。系统以有效的方式控制诸进程的并行执行,采取分布式同步方式来保证不同处理机上的进程严格同步,实现它们之间的通信,以达到高度并行执行的目的。

④ 管理程序浮动。系统中的处理单元无主从之分，而是都具有执行管理程序的能力。因此，可指定一个或几个处理单元担任主处理机来执行管理程序，还可将管理程序由一台处理机浮动到另一台处理机上。

由上所述可知，所谓分布式操作系统是指能直接对系统中的各类资源进行动态分配和管理，有效地协调和控制诸任务的并行执行，允许系统中的处理单元无主从之分，并向用户提供统一而有效的使用接口的软件集合。