

第一章

操作系统引论

本章对操作系统作一般性的介绍。通过对批量处理系统、分时系统和实时系统的介绍,来了解计算机操作系统所要解决的主要问题以及解决这些问题的方法,并使读者对操作系统中的一些基本概念有一初步了解。本章还简要介绍了 CP/M 操作系统,因为 MS-DOS 是在 CP/M 操作系统基础上发展起来的,先初步认识 CP/M 对深入理解 MS-DOS 是有益的。最后介绍 MS-DOS 版本演变情况和微型机多任务操作系统的概况。

1.1 什么是操作系统

1.1.1 操作系统的地位

现代计算机系统是一个相当复杂的系统,即使相当普及的个人电脑也是如此。我们都知道,一个计算机系统由系统硬件和系统软件组成,系统硬件包括中央处理器(CPU)、内部存储器、输入/输出设备等,系统软件包括操作系统、各种语言处理程序、系统实用程序等。一台没有任何软件支持的计算机称之为裸机,用户直接使用裸机来编制和运行程序是相当困难的,几乎是不可能的。必须有计算机厂商提供的系统软件的支持,计算机系统才能为用户程序提供一个良好的编制与运行环境。因此,实际呈现在用户面前的计算机系统已是经过若干层次软件改造的计算机。

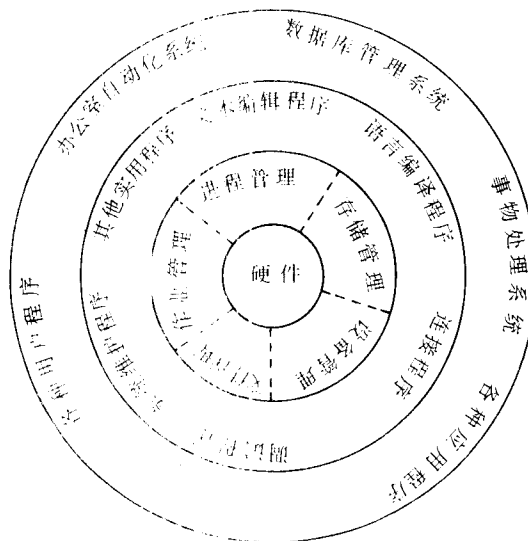


图 1.1 计算机系统层次结构

正如图 1.1 所示,可把整个计算机系统,按功能上划分为四个层次,即硬件、操作系统、系统实用软件和应用软件。这四个层次表现为一种单向服务关系,即外层可以使用内层提供的服务,反之则不行。

系统硬件是计算机系统的基础,操作系统及其他软件最终还是要以机器指令来访问和

控制各种硬件资源。硬件系统的组织结构也在不断地改进。图 1.2 给出的是目前大、中、小型计算机经常采用的一种硬件系统组织方式。这种结构以主存为中心,CPU 和 I/O 系统充分地并行工作,并且通过这个双端口的内存相互通信。I/O 通道实际上是一台专用的 I/O 处理机,它接收 CPU 的委托,独立地执行自己的通道程序,以字节多路通道、成组多路通道和选择通道分别控制低速、中速和高速外围设备的工作。I/O 通道使 CPU 摆脱了对各种 I/O 设备的繁杂控制,而且还可使各种外围设备之间并行工作。即这种大、中、小型计算机系统是以 CPU 和 I/O 系统之间以及各 I/O 设备之间尽可能地并行工作来组织硬件系统,以换取更高的性能/价格比的。

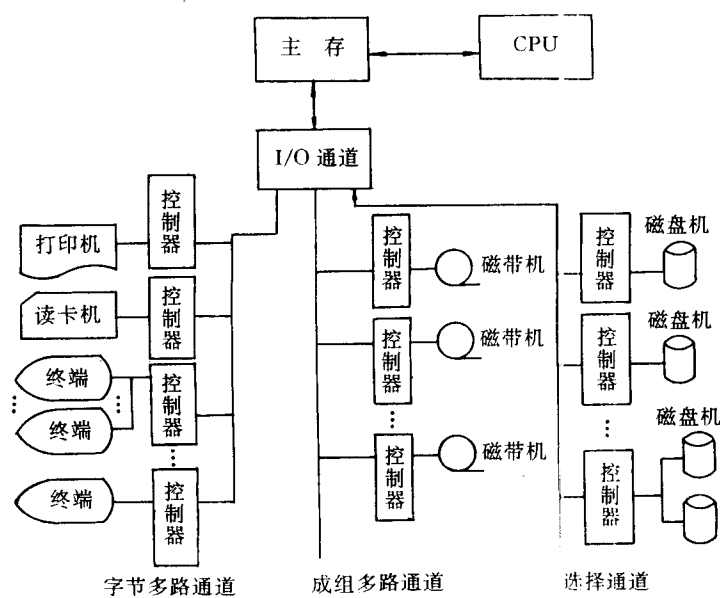


图 1.2 大、中、小型计算机系统硬件结构

微型计算机系统普遍采用单总线结构,以使扩充灵活、维护方便。CPU 通过系统总线(含地址、数据和控制信息)与存储器、I/O 接口相连,各种外围设备通过 I/O 接口挂接到系统总线上。例如,IBM PC 和 PC/XT 微型计算机使用的标准 PC 总线(62 引脚)支持 8 位数据传输和 20 位地址,IBM PC/AT 微型计算机使用的扩展 PC 总线(62+36 引脚)支持 16 位数据传输和 24 位地址。在 IBM PC 系列微机中,是将接口控制卡(适配器卡)插入机箱内的“I/O 扩展槽”与系统总线连接,I/O 扩展槽也称为 I/O 通道,但它实际上只是系统总线的延伸。这如图 1.3 所示。

在计算机系统的层次结构中,包围着系统硬件的一层就是操作系统(Operating System,简称为 OS)。它是最基本的系统软件,控制和管理着系统硬件(处理机、内存和外围设备),向上层的实用程序和用户应用程序提供一个屏蔽硬件工作细节的良好使用环境。即正是操作系统把一个裸机变成了一个可“操作”的、方便灵活的计算机系统。另一方面,因为计算机中的程序、数据大多以文件形式存放在外存储器中,而构成文件系统,接受操作系统的管理。所以,尽管操作系统处于系统软件的最底层,但却是其他所有软件的管理者。我们看

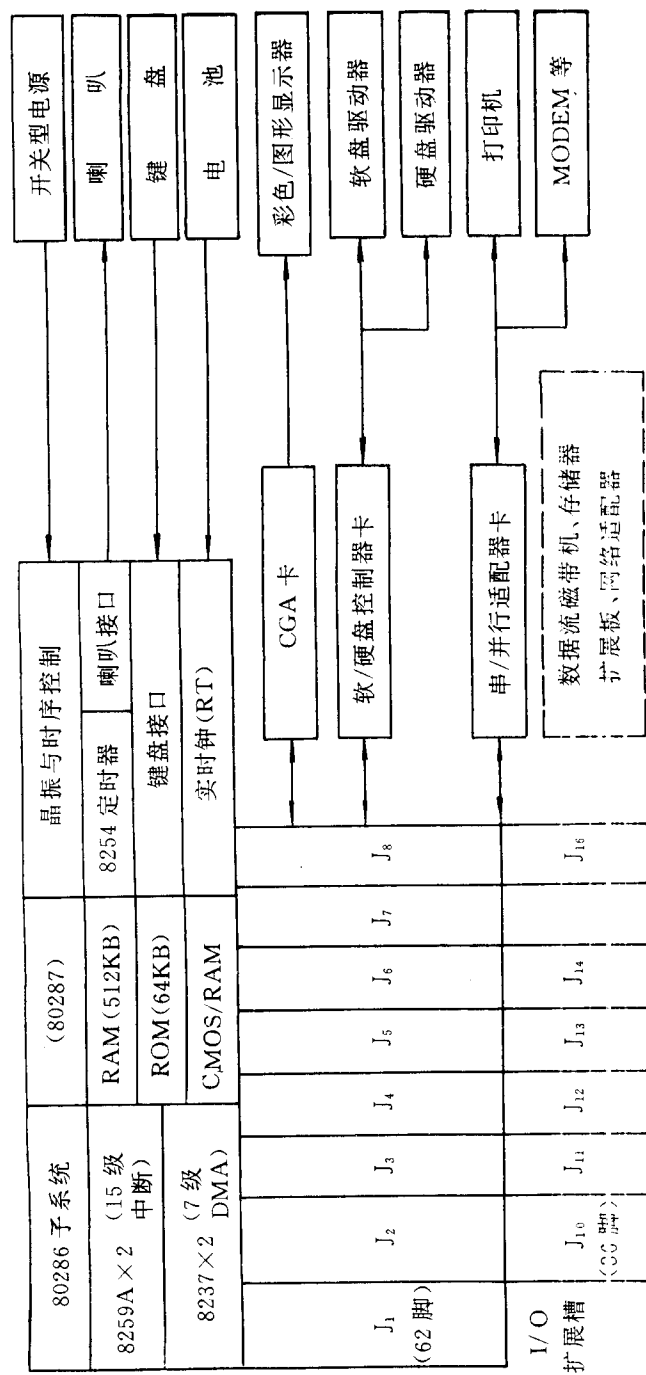


图 1.3 IBM PC/AT 硬件结构

到,操作系统层在计算机系统层次结构中是特殊的、极为重要的一层,它不仅接受硬件层提供的服务并向上层的系统实用软件层、应用软件层提供服务,而且还管理着全系统的硬件和软件资源。

系统实用软件层由一组系统实用程序(utility)组成,如语言编译程序(compiler)、文本编辑程序(text editor)、调试程序(debugging utility)、连接程序(linker)和系统维护程序(maintenance program)等。系统实用程序的功能是为应用软件以及用户加工自己的程序和数据提供服务,并为管理员对系统进行日常维护提供手段。系统实用软件和操作系统一起,应由计算机厂商提供。但使用时与操作系统不同,系统启动时操作系统即由外存储器调入内存并常驻内存,而系统实用程序则是需要时才由外存储器装入内存。也有一种观点,把这些系统实用程序也看作操作系统的一部分,本书不采用此观点。

应用软件层中的各类软件,如数据库管理系统、办公自动化系统、事务处理系统等等,可能是商售软件,也可能是用户自己编制的软件。

1.1.2 操作系统的功能

前面我们从计算机系统的层次结构的观点,讨论了操作系统这个最基本的系统软件与系统硬件以及上层的系统实用软件、应用软件的关系。现在再从资源管理的观点,进一步看看操作系统应具备哪些功能。

引入操作系统的目的有如下两方面:

(1) 操作系统要方便用户使用。一个好的操作系统应给用户使用计算机提供一个良好的界面,使用户不必了解硬件和系统软件的细节就可方便地使用计算机。这里的“用户”是一个广义的概念,不仅包括系统的一般用户、系统管理员,还应包括系统实用软件的设计者。

(2) 操作系统应最大限度地发挥计算机系统资源的使用效率。这里的系统资源既包括CPU、内存、外围设备等硬件资源,也包括程序和数据等软件资源(亦称信息资源)。应合理地组织计算机的工作流程,使这些硬、软件资源为多个用户所共享。

为了实现上述目标,操作系统通常具备以下五个方面的功能。

1. 进程管理

进程管理主要是对处理机进行管理。CPU是计算机系统中最宝贵的硬件资源。为了提高它的利用率,采用了多道程序技术。如果一个程序因等待某一条件而不能运行下去时,就把处理机占用权转交给另一个可运行的程序。或者出现了一个比当前运行的程序更重要的可运行的程序时,后者应能抢占CPU。为了描述多道程序的并发执行,就要引入进程的概念。通过进程管理协调多道程序之间的关系,以使CPU资源得到最充分的利用。

2. 存储管理

内存(又称主存)是另一宝贵的硬件资源。虽然RAM芯片的集成度不断地提高、价格不断地下降,但需求量大,内存整体的价格仍较昂贵,而且受CPU寻址能力的限制,内存的容量也有限。因此,当多个程序共享有限的内存资源时,如何为它们分配内存空间,使它们既彼此隔离、互不侵扰,又能保证在一定条件下调用,尤其是当内存不够用时,如何把当前未运行的程序及数据及时调出内存,要运行时再从外存调入内存等等,都是存储管理的任务。

3. 设备管理

设备管理是指计算机系统中除了CPU和内存以外的所有输入、输出设备的管理。除了

进行实际 I/O 操作的设备外,还包括诸如设备控制器、DMA 控制器、通道等支持设备。外围设备的种类繁多、功能差异很大。这样,设备管理的首要任务是为这些设备提供驱动程序或控制程序,以使用户不必详细了解设备及接口的技术细节,就可方便地对这些设备进行操作。另一任务就是利用中断技术,DMA 技术和通道技术,使外围设备尽可能与 CPU 并行工作,以提高设备的使用效率并提高整个系统的运行速度。

4. 文件管理

程序和数据是以文件形式存放在外存储器(如磁盘、磁鼓、磁带)的,需要时再把它们装入内存。文件包括的范围很广,例如用户作业、源程序、目标程序、初始数据、结果数据等等,而且各种系统软件甚至操作系统本身也是文件。因此,文件是计算机系统中除 CPU、内存、外围设备以外的另一类资源,即软件资源。有效地组织、存储、保护文件,以使用户方便、安全地访问它们,是操作系统文件管理的任务。

对上述四种资源的管理,其彼此之间并非是完全独立的,它们之间存在着相互依赖的关系。对每一种资源的管理,操作系统都要做到:记录资源的使用情况;以某种策略分配资源;回收资源。操作系统常借助于一些表、队列等数据结构来实施管理功能。

除了上述四项功能之外,操作系统还应该向用户直接提供使用它自己的手段,这就是操作系统的作业管理功能。

5. 作业管理

所谓作业,就是用户在一次算题过程中,或一次事物处理中,要求计算机系统所做工作的集合。例如,我们用 FORTRAN 语言编制了一个算题程序,那么要经过下列步骤:把源程序读入系统,对源程序进行编译,运行目标程序,打印输出结果。上述每一步骤称为作业步,它们的总和完成了一个作业。用户应该如何向系统提交作业,操作系统应该如何组织和调度它们的运行来提高整个系统的运行效率?这就是作业管理的任务。

以系统层次结构和资源管理两个角度考察操作系统后,我们就可以完整地回答什么是操作系统了:

操作系统是直接控制和管理计算机硬件、软件资源的最基本的系统软件,用以方便用户充分、有效地利用这些资源并增强整个计算机的处理能力。

1.2 操作系统分类

不同的硬件结构,尤其是不同的应用环境,应有不同类型的操作系统,以实现不同的追求目标。通常,将操作系统分成如下三类:

- 批处理系统(Batch Processing System)
- 分时系统(Time-sharing System)
- 实时系统(Real-time System)

1.2.1 批处理系统

批处理系统的设计目标是提高系统资源的使用效率和作业吞吐量(单位时间里处理作业的个数)。

这类操作系统要求用户使用系统提供的作业控制语言,来描述自己对作业运行的控制意图,并将这些控制信息连同自己的程序和数据一起作为一个作业提交给操作员。操作员启动有关程序将一批作业输入到计算机外存,由操作系统去控制、调度各作业的运行并输出结果。由于作业进入计算机系统后,用户不再对作业的运行进行人工干预,从而提高了系统的运行效率。

早期的批处理系统是单道的,一个作业调入内存后要直到正常运行结束或因某种原因运行不下去而中途退出后,系统才去调入下一个作业。作业总是一个接着一个地顺序运行,当一个作业运行时它便独占了全机所有资源。显然,尽管这种单道批处理系统减少了作业交接时的机时浪费,但整个系统资源利用率和运行效率还是很低的。其主要原因是处理机和输入/输出设备的串行工作方式,当作业进行输入/输出操作时处理机在空闲着,等待输入/输出操作的完成,这就浪费了大量的处理机时间。

为提高处理机资源的利用率,应尽可能地使处理机与输入/输出设备并行工作,这样便相应地出现了多道技术。多道(multiprogramming)是指在内存中同时存放若干道用户作业,允许这些作业在系统中交替地运行。如果这些作业搭配得合理,就能使处理机与输入/输出设备高度并行工作,当然,这还需要通道和中断机构等硬件的支持。多道技术看似简单,实践起来并不容易,如何使内存中的多道作业有条不紊地运行,如何为它们分配资源,这促进了操作系统的重大进步。对此,第三章将作进一步阐述。

将多道技术引入批处理系统,就成为一个多道批处理系统。现代计算机上的批处理系统,差不多都是多道批处理系统。其工作流程大致如下:各用户使用操作系统提供的作业控制语言,描述作业运行时的控制意图以及对资源的需求,然后将程序、数据以及这些信息一并交给操作员;操作员可随时把一批批作业交给系统。在外存中存放大量后备作业,系统根据一定的调度原则从后备作业中选择搭配合理的若干作业调入内存。搭配合理主要是指作业的选择既能充分利用系统各类资源又能满足不同用户的响应时间要求。在内存中的作业按多道方式组织它们的运行。例如,某个作业因等待输入输出操作的完成而暂时不能继续运行下去时,系统就使另一作业投入运行。这样,内存中的多个作业交替运行,某个作业运行完毕,系统输出它的运行结果并回收分配给它的资源,系统再从外存中调入另一个作业。作业不断地流入系统,经处理后又退出系统,形成一个源源不断的作业流。这样的系统,系统资源利用率高,作业吞吐量。

批处理系统的主要缺点是用户使用不太方便。用户把作业提交给系统后,无法对作业运行中可能出现的意外情况进行干预,于是可能是程序中一个很小的错误就导致了它无法继续运行。这种情况特别不利于程序的调试。

1.2.2 分时系统

一台计算机连接多个终端,用户通过各自的终端把作业送入计算机,计算机又通过终端向各用户报告其作业的运行情况。这种计算机能分时轮流地为各终端用户服务并能及时地对用户服务请求予以响应,这就构成了分时系统。

分时系统的主要优点是它和多个终端用户的交互会话工作方式,方便了用户使用计算机并可大大加快程序的调试过程。因此,分时系统设计的主要目标是对用户请求的及时响应,并在可能条件下尽量提高系统资源的利用率。

分时概念并不陌生。例如,硬件设计技术中的 I/O 通道与 CPU 分时使用内存;又如,多道程序技术中的多道程序分时共享硬、软件资源。分时系统中的分时则不然,它是把 CPU 的运行时间划分成一个个长短相等(或基本相等)的时间片,并把这些时间片依次轮流地分配给各终端用户程序。系统中的终端数目是有限的,只要时间片的长短确定得适当,那么一个终端用户程序从放弃 CPU 到下次再得到 CPU,只需经过不太长的而且可确定的一段时间(例如 2~3s)。

多道批处理系统中,用户无法干预程序的运行,其中一个原因就是用户不知道系统何时把他的程序投入运行。在分时系统中,用户程序从微观上看是“走走停停”,但从宏观上看却总是在“走”,用户的任何干预总能在 2~3s 时间内得到及时响应。用户的感受好像是一台 CPU 速度变慢的计算机在专门为他服务。

不难看出,分时系统的基本特征是:

(1) 同时性。若干个用户同时使用一台计算机。微观上看是各用户轮流使用计算机;宏观上看是各用户在并行工作。

(2) 独立性。用户之间可以相互独立操作,互不干涉;系统保证各用户程序运行的完整性,不会发生相互混淆或破坏现象。

(3) 及时性。系统可对用户的输入及时作出响应。分时系统性能的主要指标之一是响应时间,是指从终端发出命令到系统予以应答所需的时间。

(4) 交互性。用户可根据系统对请求的响应结果,进一步向系统提出新的请求。这种能使用户与系统进行人一机对话的工作方式,明显地有别于批处理系统,因而分时系统又被称为交互式系统。

最早的分时系统是美国麻省理工学院在 1963 年研制的 CTSS(The Compatible Time-Sharing System),用于 IBM 7904 机。他们以 CTSS 为开发环境,于 1969 年又推出第二代分时系统 MULTICS(Multiplexed Information and Computing Service)。IBM 公司于 1968 年推出的 TSS/360 也是一个成功的分时系统。目前广为流行的 UNIX 系统也是一个分时操作系统,但它是每隔一秒就产生一次时钟中断来重新计算所有进程的优先数,按动态优先数来分配 CPU,而不是采用固定时间片轮转法。

1.2.3 实时系统

实时系统的应用环境是需要对外部事件及时响应并处理的那些场合。实时系统可分成两类:

(1) 实时控制系统。把计算机用于飞机飞行、导弹发射等的自动控制时,要求计算机能尽快处理测量系统测得的数据,及时地对飞机或导弹进行控制,或将有关信息通过显示终端提供给决策人员。把计算机用于轧钢、石化、机加工等工业生产过程控制时,也要求计算机能及时处理由各类传感器送来的数据,然后控制相应的执行机构。

(2) 实时信息处理系统。把计算机用于预订飞机票、查询有关航班、航线、票价等事宜时,或把计算机用于银行系统、情报检索系统时,都要求计算机能对终端设备发来的服务请求及时予以正确的回答。

实时系统设计的目标是实时响应及处理的能力和高可靠性。对系统资源利用率,实时系统要求不高,甚至为保证高可靠性而在硬件上采用冗余措施。

实时系统与批处理系统和分时系统有何不同呢?

(1) 无论批处理系统还是分时系统,都是属于处理用户作业的通用系统,系统本身只起管理、调度、服务的作用。而许多实时系统则是为某种应用而专门设计的专用系统,系统本身就包含有实时控制过程或实时信息处理的专用程序。有些实时系统也有一定范围内的通用性,但应具有较强的系统生成能力,以便根据需要来剪裁或贴补系统功能,与实际应用环境相适应。

(2) 分时系统也要求及时响应,但那是以人所能接受的等待时间来定,一般为 2~3s,稍长、稍短一些都不会带来“灾难性”后果。而实时系统则不然,及时响应要求是以具体的控制过程或信息处理过程所能接受的延迟时间来定。对于某些实时控制过程,及时响应要求的时间可达毫秒甚至微秒数量级。实时系统,特别是实时控制系统,应具备强有力的中断机制、实时钟管理机制和快速的任务切换机制。

(3) 在一些实时信息处理系统,如民航订票系统中也有多个终端用户的问题。但与分时系统中的多终端用户相比,其与系统的交互作用要受到限制。实时系统仅允许终端操作员访问有限的专用程序,而不能编写程序或修改已存程序。

(4) 高可靠性对实时系统至关重要。实时系统常采用双工体制,即两台完全相同的计算机,一台作为值班机一台作为后备机,两机并行运行,任一时刻都有着完全相同的 CPU 现场,一旦值班机发生故障,后备机立即切入。操作系统软件也要特别可靠,并应具备某种防护机构来保证任务过载时系统仍能正常运行。与批处理系统、分时系统相比,实时系统的资源利用率一般较低。

1.2.4 通用操作系统

上面介绍的批处理系统、分时系统和实时系统是操作系统的三种基本类型。如果一个系统兼有批量处理、分时处理和实时处理三者,或其中两者的功能,那就形成了通用操作系统。

例如,把批处理与分时处理相结合,系统有分时用户时,系统及时对他们的请求作出响应,而当系统暂时没有分时用户或用户较少时,就可利用一些空闲的时间片去处理无需及时响应的批作业。类似地,也可把批处理与实时处理相结合,有实时请求时进行实时处理,没有实时请求时运行批作业。这都可提高系统资源的利用率。

在通用操作系统中,往往把批处理作业作为后台(background),而需要及时响应的用户作业作为前台(foreground)。前、后台作业的区别在于:只有前台作业不需要使用处理机时,后台作业才能得到处理机的控制权;一旦前台作业可以开始工作时,后台作业要立即让出处理机。

衡量一个操作系统性能时,常采用如下一些指标:

- (1) 系统的 RAS。RAS 是可靠性、可维护性和可用性三者的总称;
- (2) 系统的吞吐率;
- (3) 系统的响应时间;
- (4) 系统资源利用率;
- (5) 可移植性,即把一个操作系统从一种硬件环境移植到另一种硬件环境所需要的工作量。

至于操作系统是否方便用户使用,有无良好的用户界面,是无法定量评判的,但这却是

操作系统性能优劣的一个重要方面。

1.2.5 网络操作系统

将地理位置不同、具有独立功能的多个计算机系统,通过通信设备和通信线路连接起来,在功能完善的网络系统软件(网络协议、信息交换方式、控制程序和网络操作系统)的支持下,以实现更加广泛的硬件资源、软件资源的共享,这就是计算机网络。

单机操作系统是封闭的,而网络操作系统恰恰相反,它是开放系统。因为只有“开放”,才适应网络中多用户之间的交往和全网资源的共享。一个计算机系统入网后,不但大大扩大了本机用户的可用资源范围,同时也使该机的用户范围从本机用户扩大到全网用户。这就要求网络环境下的操作系统,既要为本机用户提供有效地使用网络资源的手段,又要为网络用户使用本机资源提供服务。因此,网络操作系统除了应具有单机操作系统的功能之外,还应有网络管理模块,其主要功能是支持网络通信和提供各种网络服务。

依网络作用的地理范围不同,可将计算机网络分成两大类:广域网和局域网。广域网的地理范围从几百公里到几千公里,甚至上万公里,可以覆盖一个地区、一个国家,甚至跨洲,其通信装置和线路往往由电信部门提供。局域网的地理范围是几公里到几十公里,一个企业或一个大学内都可组建局域网。

局域网的系统软件,通常由网络协议及协议软件、网络通信软件和网络操作系统三部分组成。局域网的网络操作系统属于网络层次结构的高层协议,其主要功能是实现系统共享资源的管理。

就目前大多数微机局域网而言,其网络操作系统是建筑在单机操作系统之上的,是网络用户程序与单机操作系统之间的软件接口。局域网络操作系统并不直接对网上各台计算机上的共享资源进行具体操作和维护,而是依靠单机操作系统对其所属资源进行管理,局域网络操作系统只负责协调用户进程与单机操作系统的交互作用,以及多个网络用户争用共享资源时的分配和管理。本书第十三章将介绍几种典型的局域网络操作系统。

1.3 微型机操作系统

随着大规模集成电路(LSI)和超大规模集成电路(VLSI)技术的发展,从1971年的第一个微处理器问世到现在,微处理器已历经了:4位微处理器(如Intel 4040)—8位微处理器(如Intel 8080/8085)—16位微处理器(如Intel 8086/8088)—32位微处理器(如Intel 80386)整整四代。以微处理器为CPU的微型计算机广泛应用在各个领域中。高档微型机在功能上丝毫不比传统的小型机差,甚至有所超过。另一方面,小型机也在微型化。如今,高档微型机与小型机之间的界限正在消失。

衡量微型机操作系统性能时,常看它是支持单用户还是支持多用户;是支持单任务还是支持多任务。所谓多任务,是指在一台计算机上能同时运行多个应用程序的能力。

下面简要介绍几个著名的微型机操作系统。

1.3.1 CP/M 操作系统

CP/M(Control Program for Microcomputers)是为 8 位微型机设计的一个操作系统,由美国数字研究公司(Digital Research Company)于 1976 年推出,设计人是 Gary Kildall。CP/M 是一个单用户单任务操作系统,适合于以 Intel 8080 及 Z-80 为 CPU 的微型计算机,例如 TRS-80 Mode I 微型机。1981 年又推出了 CP/M-86,它是一个 16 位微型机的单用户单任务操作系统,可用于 IBM PC 微型机。CP/M 和 CP/M-86 为用户提供了较强的使用功能和灵活方便的使用环境,从而获得了广泛应用,成为当时世界上一个较标准的磁盘操作系统。

CP/M 系列还有: 并发 CP/M-86,这是在 CP/M-86 基础上设计的,可同时运行几个应用程序的新版本; MP/M,是与 CP/M-86 兼容的多用户多任务操作系统,它提供了实时和分时处理能力; CP/NET,是计算机网络操作系统,主机由 MP/M 管理,从机与主机的通信由 CP/M 来完成。

1. CP/M 的结构

CP/M 由三个主要的子系统组成(见图 1.4):

- 控制台命令处理程序 CCP(Console Command Processor)
- 基本输入/输出系统 BIOS(Basic Input/Output System)
- 基本磁盘操作系统 BDOS(Basic Disk Operating system)

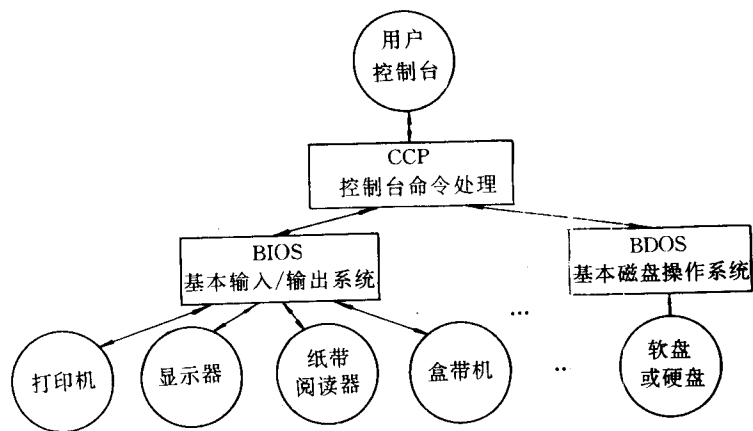


图 1.4 CP/M 的结构

CCP 负责解释用户输入的命令并负责应答。它调用 BIOS 和 BDOS 来完成实际的输入/输出和文件处理。CCP 向用户提供了一个友好的界面,使用户无需了解 CP/M 和计算机系统的内部操作过程,就可方便地使用计算机。

BIOS 包含几个设备驱动程序,它们负责向设备发送数据,由设备接收数据,并接收有关 I/O 操作成功或失败的状态信息。

BDOS 包含几个管理磁盘的实用例行程序。它使得磁盘文件管理对用户是透明的,这是

07/04/19

一切操作系统成功的关键,不能也不应当由用户对磁盘直接进行物理操作。BDOS 还管理磁盘空间的分配与释放。

2. CP/M 的内存分配

CP/M 是一个实存系统。内存的一定区域留给操作系统,大量的区域可被用户程序占用。

图 1.5 给出一个典型的 CP/M 系统的内存分配情况。第一个 256 字节的空间(第 0 页)保留给各种系统参数。操作系统驻留在内存的高端。其余部分称为暂驻程序区 TPA(Transient Program Area),可用来存放用户程序。在 S-100 总线的计算机,像北极星和 Cromemco 系统等中,TPA 在第 0 页之后。在其他系统中,如 Radio Shack 的 TRS-80,低端存储区有 ROM 程序,则 TPA 在 ROM 位置之后。

CP/M-86 的内存分配有所变动。最低端的 1KB 是系统参数区,它包括中断向量、系统缓冲区以及系统与 TPA 的公用区;系统参数区之上是 CCP 和 BDOS,再往上是 BIOS,这三部分构成了系统常驻部分;再往上才是 TPA 区。

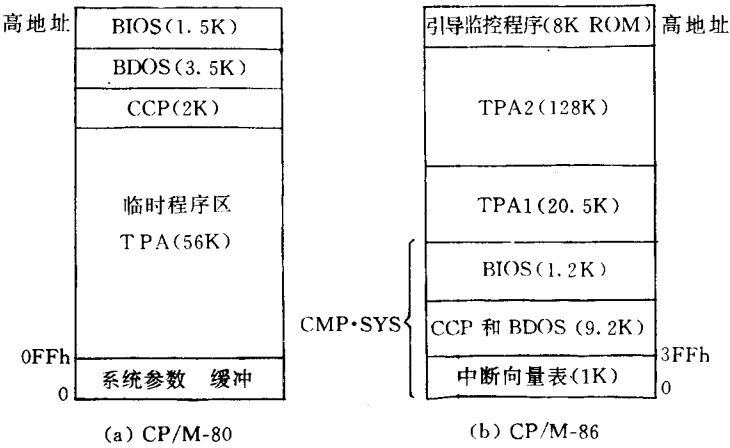


图 1.5 CP/M 内存分配

3. 逻辑设备对物理设备的映射

CP/M 把设备看成是逻辑设备或物理设备。物理设备是能完成实际输入/输出的实设备。逻辑设备是操作系统提供的一种概念,实际上是不存在的,但它可使用户界面变得更友好。用户使用逻辑设备来完成输入/输出。逻辑设备对物理设备的映射,可使用户的输入/输出请求由操作系统来满足。

CP/M 支持四类逻辑设备,命名为: CON:, RDR:, PUN: 和 LST: (冒号也是名字的一部分)。

(1) CON: —— 是用户与 CP/M 通信的低速设备。与它有关的是三个逻辑设备驱动程序: CONIN 每次从控制台输入设备输入一个字符; CONOUT 每次向控制台输出设备输出一个字符; CONST 检查控制台输入设备状态,确定字符是否就绪。

(2) RDR: —— 是用于由大容量存储设备输入的阅读设备,它的逻辑设备驱动程序是 READER。

(3) PUN:——是用于向大容量存储设备输出的穿孔设备,它的逻辑设备驱动程序是PUNCH。

(4) LST:——是列表设备,通常是输出到打印机,但某种情况下也可输出到非磁盘的大容量存储设备,它的逻辑设备驱动程序是LIST。

逻辑设备使用户编程方便,由操作系统指挥物理设备最终完成用户所要求的各种 I/O 操作。CP/M 的设备管理提供了设备独立性功能。如果使用 CP/M 的 IOBYTE 功能,一个逻辑设备每次可映射到 4 个物理设备中的某一个。

4. CP/M 的文件系统

CP/M 的文件处理由 CCP 和 BDOS 来管理; BIOS 工作在更下一层。CP/M 把文件的用户逻辑观点转换成系统的物理观点,具体的作法取决于所用的存储介质。CP/M 可支持 8 英寸软盘和 5 英寸软盘,可以是单密度的也可是双密度的,可以是硬分段的也可是软分段的;后来还支持温彻斯特(Winchester)硬盘。

• 物理磁盘结构

我们以 CP/M 1.4 版所支持的,具有 IBM 3740 格式的单面单密、软分段的 8 英寸软盘为例。此类磁盘有 77 个磁道,编号 0—76,每磁道 26 个扇区,每扇区有 128B,总容量为 256256B。

磁盘空间分配如图 1.6 所示。磁道 0 和 1 存放 CP/M 操作系统,包括一个小的系统引导装入程序。磁道 2 的 16 个扇区存放磁盘上所有文件的目录。磁道 2 的 10 个扇区和剩余的全部磁道(3—76)存放用户文件。

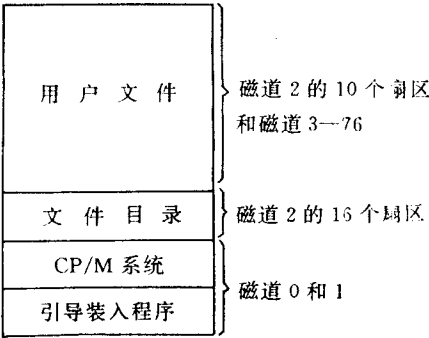


图 1.6 CP/M 磁盘空间分配

CP/M 为用户文件分配磁盘空间,以簇(cluster)为单位。1 簇是(逻辑)扇区号连续的 8 个扇区(1KB)。BDOS 为每个磁盘驱动器维持着一个 243 位的分配位示映象表(allocation bit map),如图 1.7 所示。表中的每一位代表 1 簇,若此位为 0 则未分配,若为 1 则表示此簇已分配,243 位代表 243 簇。映象表的位号(即簇号)用两位十六进制数表示,由 00h 到 F2h。图 1.7 给出的第一个可分配簇的簇号是 75h。

• 文件组织

CP/M 的文件逻辑组织是一个有记录结构的顺序文件,每个记录是 128B 长(相应于一个扇区的大小)。文件最大长度不能超过 256KB。CP/M 将文件分成若干个 16KB 长的单位(Unit),最后一个单位可能不是 16KB。CP/M 还为每一单位设置了一个文件控制块(File Control Block——FCB)。于是可多达 16 个 FCB 连接到一起,来描述 1 个 256KB 长的文件的盘簇分配情况和记录读写情况。

第一个数字	第二个数字															
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0													

图 1.7 CP/M 磁盘的分配位示映象表

每个 FCB 是 33 字节长,其结构如表 1.1 所示。它们存放在磁盘的目录区;打开文件时,该文件的各 FCB 装入内存的 TPA 区;关闭文件时又写回目录区。

表 1.1 CP/M 的 FCB 格式

字节位置	域名	意 义
0	ET	入口类型(Entry Type),CP/M 1.4 不使用
1—8	FN	文件名,8 个字符
9—11	FD	文件扩展名,3 个字符,表达文件的类型
12	EX	同一文件的目录项编号
13—14		未使用
15	RC	记录数,文件当前目录项(即本 FCB)中的记录数
16—31	DM	磁盘分配表,16 个字节长,每一字节表示一个簇号
32	NR	下一个要读写的记录号

1.3.2 MS-DOS 操作系统

自 1981 年 10 月 MS-DOS 版本 1.0 诞生,到 1991 年 7 月版本 5.0 问世,MS-DOS 已经

十易其版,表现出极活跃的生命力。当今世界上,MS-DOS 已拥有 6 千万用户,普及之广远远超过其他操作系统。

MS-DOS 是美国微软公司(Microsoft)的产品,主设计人是 Tim Paterson。IBM 公司与 Microsoft 公司签定协议,使用 MS-DOS 作为 IBM 个人计算机(Personal Computer)的操作系统,并更名为 PC-DOS。于是,除系统文件名称有些不同外,PC-DOS 与 MS-DOS 没有什么不同,二者的版本号也是对应的(有些版本的次版本号稍有差异)。进入 90 年代,两公司在发展策略上出现分歧,只有 MS-DOS 有 5.0 版。最近 Microsoft 公司又推出 6.0、6.2 版,并雄心勃勃地宣称,只要用户需要今后还有更新、更高功能的版本面世。特别值得一提的是,1994 年推出中文版的 MS-DOS 6.2,更具有十分深远的意义。

MS-DOS 是 IBM PC 系列微型机及其各种兼容机的主流操作系统。IBM 公司为其最初的 PC 机曾选择了三种操作系统,除 DOS 外,另两个操作系统是 CP/M-86 和 Softech 公司的 P 系统。竞争的结果,DOS 无可争议地显示了巨大成功。在 DOS 的发展过程中,又有一些兼容 DOS 的操作系统与之竞争。如 IGC 公司的 VM/386, Digital Research 公司的 DR DOS,其中 DR DOS 6.0 已流入我国成为 MS-DOS 5.0 的主要竞争者。

早期 PC 机,CPU 是 8088,主频是 4.77MHz,内存是 256KB,软盘容量是 180KB,显示器是 MDA。如今 PC/AT 及其各种 386,486 兼容机,CPU 已发展到 80386,80486,主频已到 33MHz 甚至更高;内存至少已是 1MB,一般已是 2MB 到 4MB,80486 微机内存最大可达 2GB;I/O 总线已达 32 位;软盘容量 360KB,1.2MB,1.44MB,直到 2.88MB;PC/XT 机上的 10MB 硬盘早已被淘汰,如今至少都是 40MB 硬盘了;显示卡也由 CGA,EGA,VGA 发展到今天普受欢迎的 TVGA。

随着硬件的飞速发展,MS-DOS 也从简易的单用户单任务操作系统,发展成支持高级文件操作、局域网的完善的单用户单任务操作系统,进而向多任务方向发展,如今的 5.0 版已是支持任务切换、具备彩色图形用户界面的高级操作系统。

MS-DOS 巨大成功的原因,首先在于它的发展策略的正确。它总是不断地推出新版本,增添新功能以支持更新的硬件并满足用户的新需求;同时新版本兼容老版本的功能,决不抛弃老用户。DOS 各类版本的演变情况如表 1.2 所示。

由此发展简表可以看出,除 3.10 版是为了网络需要而推出外,其它版本主要是为了磁盘的升级而推出。这也正是 DOS 名为磁盘操作系统(Disk Operating System)的原因。DOS 1.x~DOS 3.x 均属于单用户、单任务操作系统,DOS 4.x 和 DOS 5.0 具有多任务功能。由于推出 DOS 4.0 期间,Microsoft 正投入大量人力、物力开发 OS/2 操作系统,只好由一个全新的编程小组去开发 DOS 4.0。因而失去了工作的继承性,尽管 DOS 4.0 版提供了几项新技术,但与以前的 DOS 版本的兼容性不好,匆忙推出的 DOS 4.01 版更是隐含有许多错误。现在,国内普遍使用的是 DOS 3.30 版或 DOS 3.31 版和推出不久的、深受用户青睐的 DOS 5.0 版。

MS-DOS 取得巨大成功的另一个原因,在于它的最初设计思想及其追求目标的正确和恰当,那就是为用户的上机操作和应用软件开发提供良好的外部环境。首先是用户可以非常方便地使用几十个 DOS 命令,或以命令行方式直接键入或在 DOS 4.0 及以上版本下以 DOS Shell 菜单驱动,都可完成上机所需的一切操作。其次在于用户可用汇编语言或 C 语言来调用 DOS 支持的十多个中断功能和成百个系统功能。用户通过这些 DOS 提供的服务功

能所开发的应用程序具有代码清晰、简洁和实用性强等优点。于是,人们偏爱 DOS,觉得它对用户特别友好。而众多的用户使用 DOS 开发的各类应用程序,其数量之多、内容之丰富本身就构成了一笔巨大的财富。可以说,任何一个想战胜 DOS 的竞争者都不能无视这笔财富,它不仅提供更优异的功能和更友好的用户界面,还一定要保持与 DOS 的兼容,绝不能使拥有这笔巨大财富的广大用户变得一贫如洗。

表 1.2 DOS 发展简表

版本号	推出时间	主要性能
DOS 1.0	1981. 10.	以单面软盘为基础的 PC 机的第一个操作系统
DOS 1.1	1982. 10.	支持 5.25 英寸双面软盘
DOS 2.0	1983. 3.	支持 10MB 固定盘、响应 PC/XT 的推出,采用树状文件结构
DOS 2.10	1984. 3.	支持对错误精确定位和 PCjr 的半高软盘驱动器,改进了国际支持
DOS 3.00	1984. 8.	支持以 80286 为 CPU 的 PC/AT 机,支持 1.2MB 软盘
DOS 3.10	1984. 11.	具有支持网络的功能,首先被用于 IBM PC Network 局域网(PC 宽带网)
DOS 3.20	1986. 1.	支持 3.5 英寸的 720KB 软盘
DOS 3.30	1987. 4.	支持 3.5 英寸的 1.44MB 软盘,硬盘 DOS 分区可达 32MB,使用磁盘高速缓存
DOS 3.31 (COMPAQ)	1988.	支持大于 32MB 的硬盘分区
DOS 4.0/4.01	1988. 8.	支持 2GB 硬盘分区,支持 EMS 4.0 扩充内存,有 DOS Shell
DOS 5.0	1991. 7.	支持 3.5 英寸的 2.88MB 软盘,支持扩充内存和延伸内存,完善的 DOS Shell,全屏幕编辑器和 QBasic
DOS 6.0	1993.	更为完善的硬盘管理和内存管理,提供了更丰富的外部命令,如反病毒程序 Anti-Virus,内存优化程序 MemMaker 等。
DOS 6.2		

但 DOS 还存在着很大的局限性。尽管 DOS 4.0,5.0 已具备一些多任务处理能力,但能力有限。所以说,迄今为止 MS-DOS 仍是基于单用户单任务的操作系统。在内存管理上采用的是静态分配,DOS 内核的不可重入性,I/O 控制和修改中断向量缺乏自我保护等方面都存在缺陷。最主要的局限在于,286,386,486 各类高档微机在 DOS 下其 CPU 工作于实模式,未能充分发挥 80286,80386,80486 CPU 的保护模式的优异性能。换句话说,DOS 仍是基于 8086/8088 的操作系统。因此,作为 386,486 高档微机的操作系统,DOS 面临着单用户、多任务操作系统(如 OS/2)和多用户、多任务操作系统(如 UNIX)的巨大挑战。

鉴于本书将要用大量篇幅详细介绍 DOS,这里只先简要介绍它的发展概况和它的成功与局限。

1.3.3 微型机多任务操作系统

在一台计算机上同时运行多个应用程序的能力,十年前还只是大、中、小型计算机系统才具备的。而今,随着 386,486 等高档微机的涌现,具备多任务能力的微机操作系统受到广大用户的欢迎。图形用户界面(GUI)、多任务能力是当今微机操作系统的发展趋势。

目前,已推出十几种微型机多任务操作系统。按系统结构它们可分成内在式(built-in)和贴附式(bolt-on)两类。内在式是指操作系统开始设计就把多任务功能构筑其中,如 OS/2 和 UNIX 操作系统。贴附式是指把多任务功能附加到一个单任务操作系统中而形成,如 Microsoft 的 Windows 就是在 DOS 基础上加入多任务功能而构成的多任务操作系统。按任务调度方式它们可分成先决式(preemptive)和协作式(cooperative)两类。先决式是指由系统内部时钟来决定 CPU 由一个任务转移去执行另一个任务的时刻,转移时刻是系统以某种分配策略预先确定的,应用程序无权干预。于是,任务之间切换频繁,切入切出需要保护和恢复的信息量大。OS/2 和 UNIX 都是先决式多任务操作系统的例子。所谓协作式是指,几个应用程序联合动作,通过某种通信方式来分享 CPU。于是,任务切换可选择在安全、方便的时刻进行,切换时需保存和恢复的信息量少;但它常要依赖于任务主动让出 CPU,遇到一个“贪婪”的程序不肯让出的话,则多任务运行将会失败。Windows 和 Novell 公司的 Netware 网络操作系统是以协作式来工作的。

一、OS/2 操作系统

OS/2 是 Microsoft 和 IBM 公司根据 1985 年 8 月签署的联合开发协定,为 IBM 公司的 PS/2(Personal System/2)个人计算机系统配置的单用户多任务操作系统。它不仅可运行于 PS/2 的 50,60,80 机型上,也可运行在 PC/AT 机及其各种 286,386 兼容机上。

OS/2 的早期版本是 1.0 版,这是最基本的版本,它支持多任务和 16MB 的寻址能力,仍使用命令行键入方式,没有图形和窗口功能。在 1.0 版基础上增加了表示管理程序(Presentation Manager),并有了图形、窗口功能,而成为 1.1 版。以后,IBM 公司推出的有 IBM OS/2 1.2 版(标准版),Microsoft 公司推出的有 MS OS/2 1.21 版和 MS OS/2 2.0 版,新的版本充分利用了 Intel 80386 微处理器的性能,已是一个 32 位的版本,支持多个“虚拟 DOS 机”。

虚拟存储是 OS/2 的主要特征之一。它打破了 640KB 内存工作区的限制,应用程序可以访问 16MB 的实地址空间和 1GB 的虚拟地址空间。它利用段的换入换出在 RAM 和硬盘之间移动代码和数据。采用动态链接技术并力求代码部分公用等,使段的交换减少,而使 RAM 和硬盘交换系统的开销减少。

OS/2 使用完全先决的多任务方式,比协作方式能更好地隔离应用程序,也便于对存储器和 CPU 等资源进行全面控制。OS/2 为每个任务分配的优先级有三个级别,每个任务都有一个 CPU 时间片。若遇到在指定时间内不及时予以响应就会失败的程序,如网络通信程序,可获得“时间临界”(time critical)优先级。大多数的应用程序具有“一般”优先级。而那些平时被搁置在一边,只有系统中再没有可运行的其他任务时才被投入运行的应用程序,具有“空闲”优先级。OS/2 使用管道、队列、共享内存区、信号量等多种形式来实现进程间的通信。OS/2 的多任务能力是完善的和先进的。

OS/2 具有清晰的用户界面,它向用户提供的应用程序接口(API)是一组功能很强的调用命令。一般情况下,用户必须通过 API 使用系统资源。只有当用户申请并获得 I/O 特权后,方可直接控制 I/O 设备。OS/2 本身没有 ROM BIOS 系统。

OS/2 的窗口功能比较齐全,允许用户建立和删除窗口,允许用户定义窗口的位置、大小,以及观察多个作业的运行情况。还可在窗口或作业之间传送数据。OS/2 的窗口表示管

理程序(WPM)是以 Microsoft Windows 为基础的窗口控制系统。新版本的 OS/2 和 Windows 都做了相应的改进,两者的窗口、菜单及对话框的工作方式都几乎相同。

在 OS/2 支持下,可充分发挥全点可寻址设备的能力,绘制各种高质量的图形,乃至进行图象处理。OS/2 不用屏幕设备驱动程序,而用一组三个动态链接的程序包,向屏幕、鼠标器和键盘提供高性能的、与设备无关的接口。表示管理程序的图形程序设计接口(GPI)部分,提供了 300 多种不同性能的基本程序,以让用户方便地使用绘图功能。由于 OS/2 具有强大的绘图支持能力,又被称为面向图形处理的操作系统。

OS/2 支持 DOS 兼容环境。大多数 DOS 应用程序可运行在 OS/2 的 DOS 方式下,但时间性强的应用程序在前台运行有困难,后台运行会超时。一个双重引导程序为运行 DOS 应用程序提供了另一途径。安装此程序后,每当系统引导时,可由用户决定运行 DOS 还是运行 OS/2,但从 DOS 再转入 OS/2 时,必须重新引导系统。已投放市场的支持 32 位 80386 的 MS OS/2 2.0 版,提供的 DOS 兼容环境仍无显著的改进。

二、Windows 操作系统

MS-DOS 提供的是一种以字符为基础的用户接口,不了解硬件和操作系统的话,便难以称心如意地使用 PC 机。人们企望能把 PC 机变成一个更直观、易学、好用的工具。Microsoft 公司响应千百万 MS-DOS 用户的愿望,提供了一种图形用户界面(GUI)方式的新型操作系统,也就是 Windows。Windows 的早期版本有 2.0 版、2.1 版和 2.11 版,并在 1987 年推出了 386 PC 机专用的 Windows/386。但是,直到 1990 年 5 月推出 Windows 3.0,才引起计算机界的强烈反响,使得计算机(特别是 PC 机)的操作方法和软件开发过程产生了革命性的变化。1992 年 3 月推出了 3.1 版的 Windows,更完善了其性能,更确立了 Windows 在 PC 系列微型机操作系统中的主导地位。

Windows 3.x 是一个图形窗口操作环境软件。它替代了 DOS 环境下的命令行操作方式,而代之以“对话”、“图标”、“菜单”等图形画面和符号的全新操作方式。它有如下特点:

- (1) 丰富多彩的图形操作界面,使 PC 机易于掌握和使用。
- (2) 支持多任务运行,各任务之间既易于转换,又可方便地交换信息。
- (3) 突破了 DOS 内存 640KB 的限制,提供了实模式、标准模式、386 增强模式等操作模式,提供了虚拟存储管理能力,内存最大可达 16MB。
- (4) 提供了程序管理器、文件管理器、打印管理器、控制面板等功能强大、操作方便的管理工具,可完成任务、文件、设备等的并行管理工作。
- (5) 提供了多种方便实用的工具软件,包括:字处理程序 Write,画图软件 Paintbrush,终端通信软件 Terminal 以及时钟、日历、计算器、卡片、便笺、记录器等一套桌面办公用具软件。

Windows 最重要的改进之一是它使用的存储管理方式。与 MS-DOS 不同,Windows 采取动态内存管理方式。为了共存多个应用程序,它对应用程序以段为单位进行管理。在内存中保留程序运行所必要的段,废弃不使用的段。于是,为确保存储器的连续,必须经常进行段的再配置。另外,重复启动同一应用程序时,采取只保留个别数据段,而共用初始加载的代码段,被废弃的代码段在以后需要时,再重新从外存加载。显而易见,没有足够的内存空间就要频繁地进行段的废弃和重新装入,程序运行速度就会大大降低,再加上 Windows 本身占据