

一个完整的计算机硬件系统包括 CPU、内存、硬盘、键盘、鼠标、显示器、打印机、网络接口等部件和设备。如果每位使用者或应用程序员都需要了解硬件系统的细节,将会使计算机的使用和程序编写变得很困难。操作系统的出现和功能不断完善,给人们使用计算机和编写应用程序带来了很大的方便。有了计算机硬件系统后,首先需要安装的计算机软件就是操作系统,操作系统是最靠近硬件的一种系统软件,是所有其他软件开发和运行的基础。在操作系统的支持下,开发了大量的使用简单、功能强大的软件系统,使一般用户能够方便地使用计算机进行网上购物、线上办公、网上聊天、收发电子邮件、参加视频会议等。

## 5.1 操作系统的形成与发展

在计算机的发展进程中,计算机的性能越来越高,使用越来越方便。虽然计算机性能提高的物质基础是计算机硬件技术的快速发展,但操作系统的发展保证了硬件性能的充分发挥。早期的计算机只有计算机专业人员才能使用,而现在的计算机已广泛普及,进入上亿家庭,操作系统的出现与不断发展起了非常重要的作用。包括操作系统在内的计算机软硬件技术的快速发展,使计算机的功能越来越强大,使用越来越方便。



微课视频

### 5.1.1 操作系统概念

#### 1. 计算机系统

计算机是一种能够按照程序对数据进行自动处理的电子设备。计算机系统的基本层次结构如图 5.1 所示。计算机系统由硬件和软件组成,软件又分为系统软件和应用软件。直

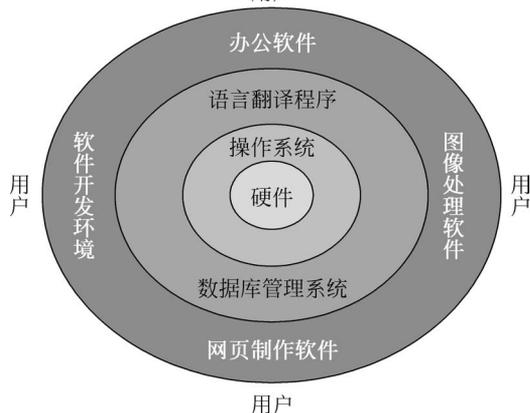


图 5.1 计算机系统的基本层次结构

接面向用户、能够解决实际问题的软件是应用软件,功能各异的应用软件帮助人们完成各种实际工作,在人们的日常工作与生活中发挥了十分重要的作用,是计算机应用的最终体现。系统软件为应用软件的开发与运行提供支持。在系统软件中,最重要的是操作系统,操作系统是其他系统软件和应用软件有效运行的基础。

## 2. 操作系统

操作系统是最靠近硬件的软件,有没有高性能是由计算机硬件决定的,能否把高性能发挥出来,操作系统起着决定性的作用。从微型计算机到超级计算机都必须在其硬件平台上加载相应的操作系统之后,才能构成一个可以协调运转的计算机系统。只有在操作系统的指挥控制下,各种计算机资源才能得到合理分配与高效使用;也只有有操作系统的支持下,其他系统软件和各种应用软件才能开发和运行。如果操作系统的功能不强,计算机硬件、其他系统软件和应用软件的功能很难充分体现。

操作系统(operating system, OS)可定义为:有效地组织和管理计算机系统中的硬件和软件资源,合理地组织计算机工作流程、控制程序的执行,提供多种服务功能及友好界面,方便用户使用计算机的系统软件。简单地说,操作系统就是管理计算机资源、控制程序执行、提供多种服务、方便用户使用。

操作系统有多种类型,不同类型的操作系统其侧重的目标有所不同,但共同的一般性目标主要有有效性、方便性、可扩充性、开放性、可靠性和可移植性等。对用户来说,其中的方便性和有效性是最主要的。对于超级计算机、大型计算机和服务器,由于价格昂贵,使用的人少且多为专业人员,比较强调有效性。微型计算机出现后,计算机使用者越来越多,而且多为非专业人员,方便性成为操作系统关注的重点。Windows 系列操作系统之所以广受欢迎,一个重要因素就是其学习和使用的方便性。

(1) 有效性。在未配置操作系统的计算机系统中,中央处理器等资源会经常处于空闲状态而得不到充分利用;存储器中存放的数据由于无序而浪费了存储空间。配置了高性能操作系统后,可使中央处理器等部件由于减少等待时间而得到更为有效的利用,使存储器中存放的数据有序而节省存储空间。此外,操作系统还可以通过合理地组织计算机的工作流程,进一步改善系统的资源利用率及提高系统的输入输出效率。

(2) 方便性。没有操作系统,就只能通过控制台输入控制命令。以这种方式操作使用计算机是让人非常头疼的一件事情,只适用于专业人员。有了操作系统,特别是有了像 Windows 这类功能强大、界面友好的操作系统,使使用计算机变得非常容易和方便,通过鼠标和键盘就能实现很多功能。

### 5.1.2 操作系统的形成

计算机硬件技术的发展目标是提高计算机的性能,主要指标就是运算速度。早期主要是提高单处理器的性能,20世纪70年代后出现了多核CPU和多处理器系统,通过多个处理器的合作来提高计算机的整体性能。操作系统的主要目标是提高处理器的利用率,使计算机的硬件特性能够充分发挥出来。1956年出现了第一个操作系统GM-NAA I/O,该系统是鲍勃·帕特里克(Bob Patrick)在美国通用汽车的系统监督(system monitor)程序的基础上,为美国通用汽车和北美航空公司在IBM 704计算机上设计的基本输入输出系统,可以成批地处理作业,在一个作业结束之后,它会自动执行新的作业。

## 1. 人工操作

1956年之前的计算机没有操作系统,专业人员用汇编语言或机器语言编写程序,程序员或操作员用手工方式直接控制和使用计算机硬件系统,用穿孔机将编写好的程序及相应的数据穿孔在纸带/卡片上,通过纸带/卡片机输入计算机。然后启动计算机执行程序,通过控制台上的开关、按钮和指示灯来操作和控制程序的执行,程序执行完并取走计算结果后,下一个用户才可以使用该计算机。

人工操作方式有两个主要缺点:

(1) 用户独占整个计算机。一台计算机的全部资源由一个用户独占使用。

(2) CPU 利用率偏低。当用户装卸纸带/卡片、纸带/卡片机运行、打印输出结果时,CPU 处于空闲状态。

相对于 CPU 的运行速度,手工操作和纸带/卡片机、打印机等外设的速度是很慢的,使得高速的 CPU 绝大部分时间在等待慢速的手工操作和外设运行,计算机资源得不到有效利用。要知道这一时期都是价格昂贵的大型计算机。

## 2. 批处理操作系统

批处理操作系统指操作员将用户提供的若干作业以“成批”的方式,同时交给计算机系统。作业(job)指用户在一次数据处理中要求计算机所做的全部工作的总和,由用户程序、数据和作业说明书组成。作业曾经是早期操作系统的一个重要概念,现代操作系统中已经很少使用这个概念。批处理操作系统分为单道批处理操作系统和多道批处理操作系统。

### 1) 单道批处理操作系统

操作员把接收到的一批用户作业存放在外存,由操作系统自动地一次调用一个作业进入内存运行。这种处理方法减少了人工操作的干预时间,提高了 CPU 的利用率。但是一个作业在运行时,若提出输入或输出请求,那么 CPU 就必须等待输入或输出的完成,这就意味着 CPU 仍可能长时间空闲。这就是早期的单道批处理操作系统的工作模式,虽然减少了等待人工操作的时间,但仍需要等待输入或输出操作时间。

单道批处理操作系统的代表是 FMS(FORTRAN monitor system,FORTRAN 监控系统)和 IBM 公司为 IBM 7094 计算机配置的 IBM SYS 操作系统。

### 2) 多道批处理操作系统

多道批处理操作系统改进了单道批处理操作系统的不足。从外存中把多个作业同时调入内存,当某个作业需要输入或输出时,CPU 为该作业启动相应的输入或输出操作后就转去执行内存中的下一个作业。这样,第二个作业的执行与第一个作业的输入或输出并行工作,从而进一步减少 CPU 的等待时间。但多道批处理操作系统仍有其不足,用户不能干预作业的执行,即使发现错误也不能及时改正,即没有人机交互。

用于 IBM 4300 系列大型计算机的 IBM DOS/VS 和 IBM DOS/VES 操作系统是多道批处理操作系统的代表。

把一批作业放入外存(当时主要用的是磁带)可以有脱机方式和联机方式。脱机方式是指事先将描述作业的纸带/卡片装入纸带/卡片输入机,在一台外围机(价格较低、功能较弱的计算机)的控制下把纸带/卡片上的作业输入到磁带上。当 CPU 需要这些作业时,再从磁带调入内存。类似地,当 CPU 需要输出时,可由 CPU 直接把数据从内存送到磁带上,然后在另一台外围机的控制下,将磁带上的结果送到相应的输出设备输出。由于作业的输入

输出都是在外围机的控制下完成的,或者说它们是在脱离主机的情况下进行的,故称为脱机输入输出方式。与此对应,在主机的直接控制下进行输入输出的方式称为联机输入输出方式。

脱机输入输出方式的主要优点如下。

(1) 减少了主机 CPU 的空闲时间。装卸纸带/卡片、作业从低速的纸带/卡片机送到高速的磁带上(当时磁带的存取速度远高于纸带/卡片的存取速度),都是在脱机情况下进行的,不占用主机时间,从而有效地减少了主机 CPU 的空闲时间。

(2) 提高了输入输出速度。当主机在运行中需要输入或输出数据时,是直接通过高速的磁带机进行,节省了输入输出时间,进一步减少了主机 CPU 的空闲时间。

总之,脱机方式有效地提高了主机 CPU 的利用率。

### 3. 分时操作系统

所谓分时,就是把计算机的硬件资源(主要是 CPU)在时间上加以分割,形成一个个微小的时间段,每个微小时间段称为一个时间片,每个用户(程序)依次使用一个时间片,从而可以将 CPU 的工作时间轮流地提供给多个用户使用。一台计算机可以连接多个控制台或终端,从而使多个用户通过这些控制台或终端共享一台计算机。由于时间片非常小(一般为几十毫秒),用户感觉不到有其他用户在和自己共享计算机,就如同自己的专用计算机一样。同时每个用户可以通过控制台或终端控制自己程序的运行,具有人机交互功能,这就克服了批处理系统的不足。需要说明的是,根据计算机的性能高低,对连接终端的个数是有一定限制的,超出这个限制,用户会感觉计算机的速度很慢。

1961年,美国麻省理工学院开发出第一个分时操作系统 CTSS,运行在 IBM 709 和 IBM 7094 大型计算机上,更有名的分时操作系统是麻省理工学院、贝尔实验室和通用电气公司共同开发的 MULTICS。

### 4. 实时操作系统

实时系统是一种能在限定时间内对外部事件做出响应和处理的计算机系统,可以分为实时控制系统和实时信息系统。

实时控制系统是以计算机为中心的过程控制系统,如自动化生产线控制系统、无人侦察机控制系统等。实时控制系统属于硬实时任务,系统必须满足对限定时间的要求,否则会产生“失之毫厘,差之千里”的严重后果。实时信息系统通常是指实时信息处理系统,如航空订票系统、信息检索系统等。实时信息系统属于软实时任务,系统对限定时间的要求并不十分严格,偶尔的超时也不会产生严重后果。

能够满足实时系统要求的操作系统称为实时操作系统,实时系统的最重要特性就是及时(响应)性和可靠性。当然,一个计算机系统的实时性既依赖于实时操作系统,又和计算机的硬件性能密切相关。

现在的嵌入式操作系统都是实时操作系统,如 VxWorks、QNX、Windows CE 等。

### 5. 通用操作系统

同时具有分时、实时和批处理功能的操作系统称作通用操作系统。显然,通用操作系统规模更为庞大,结构更为复杂,功能也更为强大。开发通用操作系统的目的是为用户提供多模式的服务,同时进一步提高系统资源的利用效率。

在通用操作系统中,可能同时存在 3 类任务:实时任务、分时任务和批处理任务。这 3

类任务通常按照其紧迫程度加以分组:实时任务级别最高,分时任务次之,批处理任务级别最低。当有实时请求时,系统优先处理;当没有实时任务时,系统为分时用户服务;仅当既无实时任务又无分时任务时,系统才执行批处理任务。在智能手机上,手机操作系统为电话、微信语音和视频播放设置了从高到低的优先级,看视频时如果有微信语音进来,系统停止播放视频,优先处理微信语音,接听微信语音时,如果又有电话打进来,停止微信语音,优先处理电话。

UNIX 的早期版本就是当时通用操作系统的典型代表,现在常用的操作系统 Windows、UNIX 和 Linux 都属于通用操作系统。

### 5.1.3 操作系统的发展

操作系统的发展已有近 70 年的历史。经过 20 世纪 60~70 年代的大发展,到 80 年代已趋于成熟。但随着超大规模集成电路的发展和计算机体系结构的变化,操作系统也在不断发展和完善,先后出现了微机操作系统、多处理器操作系统、网络操作系统、分布式操作系统和嵌入式操作系统等。

#### 1. 微机操作系统

配置在微机上的操作系统称为微机操作系统。最早出现的微机操作系统是使用在 8 位微机上的 CP/M。一般把微机操作系统分为单用户单任务操作系统、单用户多任务操作系统和多用户多任务操作系统。

单用户单任务的含义是,只允许一个用户使用计算机,且只允许该用户运行一个程序。这是一种最简单的微机操作系统,主要配置在早期的 8 位微机和 16 位微机上。代表性的单用户单任务操作系统是 CP/M 和 MS-DOS。

单用户多任务的含义是,只允许一个用户使用计算机,但允许该用户提交多个程序并发执行,即可以同时完成多个任务,从而有效地改善系统的性能。目前,在微机上配置的操作系统大多数是单用户多任务操作系统,其中代表性的是 Windows 10/11,用户可以一边调试执行程序一边听音乐,还可以同时收发电子邮件。

多用户多任务的含义是,允许多个用户通过各自的终端,使用同一台主机,共享主机系统中的各类资源,而每个用户又可以提交几个程序,使它们并发执行,从而进一步提高资源利用率和增加系统吞吐量。本来多用户多任务操作系统一般是用于大型计算机的,随着微型计算机性能的不提高,在高档微机上也可以安装多用户多任务操作系统。其中,代表性的是微机版的 UNIX 和 Linux、Windows 10 等操作系统。把微机用作服务器连接多个客户机。

#### 2. 多处理器操作系统

从计算机的发展历史可以看出,提高计算机硬件性能的主要途径有两个:一是提高单个处理器的运算速度;二是改进计算机系统的体系结构。早期的计算机系统基本上都是单处理器系统,重点在于提高单个处理器的运算速度及相关器件的性能。20 世纪 70 年代出现了多核处理器和多处理器系统(multi-processor system, MPS),试图通过改进计算机体系结构来提高系统性能。近年来推出的超级计算机和大型计算机,大多采用多处理器结构,高档微机使用多核处理器。

根据多个处理器之间耦合的紧密程度,可把 MPS 分为两类:紧耦合 MPS 和松耦合



微课视频

MPS。紧耦合 MPS 是指多个处理器通过高速线路互连,共享内存、外存和外设,多处理器系统一般是指紧耦合 MPS;松耦合 MPS 是指每个处理器有各自的内存、外存和外设,实际上是构成了一台独立的计算机,多台计算机通过通信线路互连,松耦合 MPS 也可以称为多计算机系统或计算机网络。

为多处理器系统配置的操作系统称为多处理器操作系统,主要有两种模式:

(1) 非对称模式,又称为主从模式。在非对称模式中,把处理器分为主处理器和从处理器两类。主处理器只有一个,其上安装有操作系统,用于管理整个系统的资源,并负责为各从处理器分配任务及协调从处理器的运行。从处理器可有多个,它们执行预先规定的任务及由主处理器所分配的任务。主从式操作系统易于实现,但资源利用率比较低。

(2) 对称模式。在对称模式中,所有处理器的地位都是相同的。在每个处理器上运行一个相同的备份操作系统,用它来管理本地资源和控制进程的运行以及各处理器之间的通信。对称模式操作系统资源利用率和整体性能比较高,但实现起来比较复杂。

代表性的多处理器操作系统有 SUN 公司的 Solaris、AT&T 公司的 UNIX System V4.0 MP 版本、DG 公司的 DG/UX 等。

### 3. 网络操作系统

计算机网络可以定义为自主计算机的互连集合。自主计算机是指一台独立的计算机,互连是表示计算机之间能够实现相互通信和资源共享。计算机网络是在计算机技术和通信技术快速发展与相互结合的基础上发展起来的。

运行在计算机网络环境上的网络操作系统应具有如下 4 方面的功能。

(1) 网络通信。这是网络最基本的功能,其任务是在源主机和目标主机之间实现无差错的数据传输。为此,应有的主要功能包括建立和拆除通信链路、传输控制、差错控制、流量控制和路由选择等。

(2) 资源管理。对网络中可共享的软硬件资源实施有效的管理,协调和控制各用户对共享资源的使用,保证数据的安全性和一致性。常用的共享资源有硬盘、打印机、软件和数据文件等。

(3) 网络服务。这是在网络通信和资源管理的基础上,为了方便用户而直接向用户提供的多种有效服务,主要有电子邮件、文件传输、网络新闻、信息检索、即时通信和电子商务等服务。

(4) 网络管理。网络管理最基本的任务是安全管理,通过存取控制技术来确保存取数据的安全性,通过容错技术来保证系统出现故障时数据的安全性,通过反病毒技术、入侵检测技术和防火墙技术等来确保计算机系统免受非法攻击。此外,还应对网络性能进行监测,对使用情况进行统计分析,以便为网络性能优化和网络维护等提供必要的信息。

常见的网络操作系统有 Windows Server、网络版的 UNIX 和 Linux 等。

### 4. 分布式操作系统

在分布式概念提出之前的计算机系统中,其处理和控制在功能上都高度地集中在一台主机上,所有的任务都由主机处理,这样的系统称为集中式处理系统。

在分布式处理系统中,系统的处理和控制在功能上分散在系统的各个处理单元上。系统中的所有任务也可动态地被分配到各个处理单元上去,使它们并行执行,实现分布处理。分布式处理系统最基本的特征是处理上的分布性,而分布处理的实质是资源、功能、任务和控制

都是分布的。所谓分布式处理系统(distributed processing system),指由多个分散的处理单元经互连网络的连接而形成的系统,简称分布式系统。其中,每个处理单元既具有高度的自治性,又与其他处理单元相互协同,能在系统范围内实现资源管理,动态地分配任务,并能并行地运行分布式程序。在分布式系统中,如果每个处理单元都是计算机,则可称为分布式计算机系统,它通常就是计算机网络。分布式系统一般是指其处理单元只是由处理器和局部存储器组成。

在分布式系统上配置的操作系统,称为分布式操作系统。

代表性的分布式操作系统有荷兰自由大学的 Amoeba、法国 INRIA 学会的 Chorus、国产 Laxcus 等。

### 5. 嵌入式操作系统

嵌入式操作系统指运行在嵌入式计算机中的操作系统,嵌入式计算机及其操作系统与相关设备有机地结合成一体。嵌入式计算机广泛应用于网络通信、国防安全、航空航天、智能家电、无人驾驶等多个领域,如火星探测器、扫地机器人、无人机等。与一般操作系统相比,嵌入式操作系统具有微型化、可定制、实时性好、可靠性高和易移植等特点。嵌入式操作系统也是实时操作系统。

常用的嵌入式操作系统有嵌入式 Linux、VxWorks 和 QNX 等。

## 5.2 操作系统的特征

虽然不同类型的操作系统各有其特点,但一般都具有并发性、共享性、虚拟性和异步性等共同的基本特征。

### 1. 并发性

并发(concurrency)指两个或多个事件在同一时间段内发生,而并行指两个或多个事件在同一时刻发生。并发和并行是有区别的,在多处理器系统中,可以有多个进程并行执行,一个处理器执行一个进程。在单处理器系统中,多个进程是不可能并行执行的,但可以并发执行,即多个进程在一段时间内同时运行,但在每一时刻,只能有一个进程在运行,多个并发的进程在交替地使用处理器运行,操作系统负责这些进程之间的执行切换。什么是进程呢?简单地说,进程指程序的一次执行过程。

并发性改进了在一段时间内一个进程对 CPU 的独占,可以让多个进程交替地使用 CPU,从而有效提高系统资源的利用率,提高系统的处理能力,但也使系统管理变得复杂,操作系统要具备高效控制和管理各种并发活动的的能力。

### 2. 共享性

共享(sharing)指系统中的资源可供多个并发执行的进程共同使用。共享可以提高系统资源的利用率,为每个进程分别提供其所需的所有资源是非常浪费的,也没这个必要。

并发性和共享性是操作系统的两个最基本的特征,它们互为存在条件。一方面,资源共享是以进程的并发执行为条件的,若系统不允许进程并发执行,也就不存在资源共享问题;另一方面,若操作系统不能对资源共享实施有效管理,则必将影响到进程正确地并发执行,甚至根本无法并发执行。



微课视频

### 3. 虚拟性

操作系统中的虚拟(virtual)指通过某种技术把一个物理实体变成若干逻辑上的对应物。物理实体是实际存在的,对应物是虚的,是用户感觉到的。例如,在分时系统中,虽然只有一个CPU,但每个终端用户都认为有一个CPU在专门为自己服务,即利用分时技术可以把物理上的一个CPU虚拟为逻辑上的多个CPU,逻辑上的CPU称为虚拟处理器。类似地,也可以把一台物理输入输出设备虚拟为多台逻辑上的输入输出设备(虚拟设备),把一条物理信道虚拟为多条逻辑信道(虚拟信道)。在操作系统中,虚拟主要是通过分时使用的方式实现的。

### 4. 异步性

在多道程序环境下,允许多个进程并发执行,但由于资源及控制方式等因素的限制,进程的执行并非一次性地连续执行完,通常是以“断断续续”的方式进行的。内存中的每个进程在何时执行、何时暂停、以怎样的速度向前推进、每个进程总共需要多长时间才能完成,都是不可预知的。先进入内存的进程不一定先完成,而后进入内存的进程也不一定后完成,即进程是以异步(asynchronism)方式运行的。操作系统要严格保证,只要运行环境相同,多次运行同一进程,都应获得完全相同的结果。

## 5.3 操作系统的功能

操作系统的作用是管理计算机资源、方便用户使用计算机。由于计算机的资源主要包括处理器(CPU)、存储器(内存)、输入输出设备、外存(以文件形式存储软件和数据)、网络设备等,因此操作系统的主要功能包括处理器管理功能、存储器管理功能、设备管理功能、文件管理功能和网络与通信管理功能。此外,为了方便用户通过操作系统使用计算机,还需向用户提供一个使用方便的用户接口。

### 5.3.1 处理器管理功能

处理器管理的主要任务是对中央处理器进行分配,对其运行进行有效的控制和管理,最大限度地提高处理器的利用率,减少其空闲时间。在多道程序环境下,处理器的分配和运行都是以进程为基本单位的,因而处理器管理也可称为进程管理。进程(process)指程序的一次执行过程。进程是操作系统中最基本、最重要的一个概念,是在多道程序系统出现后,为描述程序运行的动态特性而引入的概念。进程是一个动态概念,其静态实体需要一种数据结构来表示,描述程序运行中的状态。

在多道程序环境下,处理器在多个进程之间切换,一个进程的执行可能是断断续续的,经历多次的执行、等待交替才能完成整个进程的执行,进程由执行变为等待,再由等待变为执行,接着上次执行的断点继续执行,断点处的状态信息要保存,需要有合适的数据结构。在多道程序环境中,允许有多个进程并发执行,就可能出现一个程序被多次调用而参与到并发执行中,一个程序对应多个进程,需要保存每个进程的状态信息,也需要合适的数据结构。

处理器管理要保证处理器在多个进程间进行有效的切换,既保证各进程执行的正确,又保证处理器具有比较高的利用率。处理器管理主要包括进程控制、进程同步、进程通信和处理器调度4方面。



微课视频

### 1. 进程控制

在多道程序环境下,要使程序运行,必须先为它创建一个或几个进程,并为之分配必要的资源(如处理器、内存等)。当进程运行结束时,要立即撤销该进程,以便及时回收该进程所占用的各类资源。进程控制的主要任务是为程序创建进程,撤销已结束的进程,以及控制进程在运行过程中的状态转换。进程有3个状态:运行状态、就绪状态和等待状态。运行状态是进程占用处理器运行的状态;就绪状态是进程具备运行条件,只要分配给处理器就能运行的状态;等待状态是进程不具备运行条件,正在等待某个事件完成的状态。

### 2. 进程同步

一般来说,相互无关的多个进程是以异步方式运行的,并以人们不可预知的速度向前推进。有时多个进程也存在一定的制约关系,如多个进程共享同一独占型资源或多个进程协作完成同一项任务,为了保证这些相互有关的进程能够正确地执行,系统中必须设置进程同步机制。进程同步的主要任务是对存在制约关系的多个进程的运行进行协调,主要有同步和互斥两种协调方式。

(1) 进程同步方式。多个进程协作完成同一项任务,进程之间在执行次序上有制约关系,应有同步机制对这些进程的执行加以协调,保证按正确的先后次序进行。

(2) 进程互斥方式。多个进程在对独占型资源进行共享访问时,按照一定的策略逐次使用资源,如先来先服务、短者优先等策略。互斥也可以看作一种特殊的同步方式,是对进程访问资源顺序的一种协调。

实现进程互斥常用加锁机制,进程P1获得独占型资源R的使用权,就对资源R加锁,此时进程P2也想使用资源R,必须等待,直到P1使用完并释放R。这种加锁机制也有副作用,有时会导致死锁或饥饿现象,死锁(deadlock)指一组分别占有有一定资源的进程相互等待其他进程的资源而永远也得不到,导致各进程都无法执行。饥饿(starvation)指由于资源分配策略不当等原因,使某一进程永远也得不到所需资源而导致无法执行。所以,操作系统在引入加锁机制的同时,还要有相应的解决死锁或饥饿问题的有效方法,完全避免死锁或饥饿是很困难的,只能是尽可能减少死锁或饥饿的发生。

### 3. 进程通信

在多道程序环境下,可由系统建立多个进程,这些进程相互合作去完成一共同任务,在这些相互合作的进程之间,往往需要交换信息。例如,有3个相互合作的进程,它们分别是输入进程、计算进程和打印进程。输入进程负责输入数据并传送给计算进程,计算进程对输入数据进行计算,并把计算结果传送给打印进程,由打印进程把结果打印出来。进程通信的任务是实现相互合作进程之间的信息交换,如输入进程把输入的数据传送给计算进程,计算进程把计算结果传送给打印进程。

当相互合作的进程处于同一计算机系统时,通常采用直接通信方式。由源进程利用发送命令的方式直接将消息挂到目标进程的消息队列上,以后由目标进程利用接收命令从其消息队列中取出消息。

当相互合作的进程处于不同的计算机系统中时,常采用间接通信方式,由源进程利用发送命令将消息送入一个存放消息的中间实体中,之后由目标进程利用接收命令从中间实体中取走消息。该中间实体通常称为邮箱,相应的通信系统称为电子邮件系统。

#### 4. 处理器调度

处理器调度的主要任务是为并发执行的多个进程分配处理器,分为 3 级:高级调度、中级调度和低级调度。

高级调度也称为作业调度。作业调度的基本任务是从存放在外存中的后备作业队列中,按照一定的算法选择若干作业调入内存准备执行。

中级调度也称为交换调度。根据进程的当前状态决定外存和内存的进程交换,当内存容量不足时,把暂不执行的进程从内存调至外存(这种作用的外存称为虚拟内存),将需要执行的进程调入内存。这种方式可以提高内存资源的利用率。

低级调度也称为进程调度。进程调度的任务则是从进程的就绪队列中,按照一定的算法选出一个进程,把处理器分配给它,并为它设置运行环境,使该进程进入运行状态。

作业调度是多道批处理系统的重要功能,现代操作系统中不再有作业调度,只有交换调度和进程调度。

进程作为资源分配和并发调度的基本单位,提高了 CPU 等系统资源的利用率。但由于系统的地址空间和资源有限,限制了系统中所允许的并发进程的个数。同时,并发进程之间的执行切换也消耗了比较多的处理器时间。为进一步提高并发程度和减少进程切换的时间消耗,20 世纪 80 年代中期提出了线程的概念,线程(thread)指进程内部一个可独立执行的实体。在一个进程中可以创建多个线程,实现多个线程的并发执行,即一个进程的多个部分可以并发执行,进一步提高了 CPU 的利用率,减少了进程切换次数及时间消耗。

### 5.3.2 存储器管理功能

存储器管理的主要任务是管理内存资源,为并发进程的执行提供内存空间;提高内存空间的利用率,并能从逻辑上扩充内存空间以适应大进程和更多进程并发执行的需要。存储器管理应具有内存分配、内存保护、地址映射和内存扩充等功能。

#### 1. 内存分配

内存分配的主要任务是为需要执行的进程分配适当的内存空间,及时回收执行完的进程所释放的内存空间;尽量减少不可用的内存空间,提高内存空间的利用率。

#### 2. 内存保护

内存保护的主要任务是确保并发执行的每个进程都在自己的内存空间中执行,互不干扰,防止一个进程访问其他进程的内存空间进而影响那个进程的正常执行。特别是不允许用户进程访问操作系统所占用的内存区域,否则可能造成整个系统的瘫痪。

为了确保每个进程只在自己的内存区内执行,操作系统要有内存保护机制。一种比较简单的方法是使用两个寄存器,分别存放一个进程占用内存空间的上界和下界。对进程中每条指令所访问的内存地址进行检查,如果超出了所分配的内存空间范围,便停止该进程的执行。

#### 3. 地址映射

高级语言源程序经编译和连接后形成可装入内存的机器语言程序,这时程序中第一条指令的地址是从 0 开始的,后面的指令依次编址,由这些指令在程序中的地址所构成的地址范围称为地址空间,其中的地址称为逻辑地址或相对地址。由内存中的若干存储单元所构成的地址范围称为内存空间,其中的地址称为物理地址。

在进程并发执行的环境下,地址空间中的逻辑地址和内存空间中的物理地址是不一致的,地址映射功能是将地址空间中的逻辑地址转换为内存空间中的物理地址,即实现逻辑地址到物理地址的变换,这样才能保证每个进程都分配到合适的内存空间并正确执行。

#### 4. 内存扩充

虽然随着计算机硬件技术的快速发展,计算机的内存容量也相应地有了大幅度提高,但计算机要完成任务的规模也在逐渐增大,内存容量满足不了用户(解决大规模问题)需要的可能性一直存在。内存扩充的任务是借助于虚拟存储技术(把一部分外存虚拟成内存使用)从逻辑上扩充内存容量,而不是真正增加物理内存的容量。这种虚拟存储技术在不增加硬件成本的前提下,扩充了逻辑内存,能够执行更大的进程或使更多的进程能并发执行,提高了系统性能。

这实际上是改变了初始的进程执行模式:要执行的进程及相应的数据需要全部调入内存才能执行。现在的模式是,先调入部分指令和数据就能启动进程执行,在执行过程中,根据需要逐步把后续指令和数据调入内存,同时把暂时不需要的已经执行过的指令和数据调至特定的外存区域。这样就能在不增加物理内存容量的前提下,执行更大的进程或使更多的进程并发执行。与内存配合的特定外存区域称为逻辑内存或虚拟内存。

### 5.3.3 设备管理功能

为方便使用计算机,计算机要配备键盘、鼠标、显示器、打印机等输入输出设备。设备管理的主要任务是响应用户提出的输入输出请求,为其分配相应的输入输出设备;提高 CPU 和输入输出设备的使用效率,提高输入输出速度;方便用户使用输入输出设备。为有效完成上述任务,设备管理应具有缓冲区管理、设备分配、设备驱动调度、设备独立性和虚拟设备等功能。

#### 1. 缓冲区管理

缓冲区管理的基本任务是管理好各种类型的缓冲区。缓冲区指内存中的一块特定存储区域或设备本身自有的存储空间,用以缓和 CPU 和输入输出设备速度不匹配的矛盾,目的是提高 CPU 和输入输出设备的利用率。例如,需要打印输出时,可以把打印内容放入缓冲区,供打印机取出打印,此时 CPU 可以继续执行其他任务,避免了高速的 CPU 等待低速的打印机打印,实际上是 CPU 在和打印机并行工作。

#### 2. 设备分配

设备分配的基本任务是根据用户的输入输出请求,为之分配相应的输入输出设备。

为了实现设备的有效分配,系统中应设置设备控制表等数据结构,记录设备的标识符、类型、地址和状态等信息,用以表示该设备的唯一标识、是否空闲等,作为设备分配的依据。设备使用完后,系统要及时回收以便其他用户使用。

#### 3. 设备驱动调度

设备驱动调度的基本任务是把用户提交的输入输出请求转换为实际的输入输出操作,完成用户的输入输出请求。由 CPU 向设备控制器发出输入输出指令,启动输入输出设备完成指定的输入输出操作,并能接收由设备控制器发来的中断请求,给予及时的响应和相应的处理。设备驱动调度通过设备驱动程序来完成,设备驱动程序与硬件密切相关,其中部分代码可能需要用汇编语言编写。

#### 4. 设备独立性

设备独立性指应用程序独立于具体的物理设备,与实际使用的物理设备无关。设备独立性不仅能提高用户程序的适应性,使程序不局限于某个具体的物理设备,而且易于实现输入输出的重定向,易于应对输入输出设备故障。

#### 5. 虚拟设备

虚拟设备指通过某种方法(如分时方法)把一台独占型物理设备改造成能供多个用户共享使用的逻辑设备,这种逻辑设备称为虚拟设备。虚拟设备技术能够有效提高设备的利用率,使每个共享使用设备的用户都感觉自己在独自使用该设备。

### 5.3.4 文件管理功能

要执行一个程序,需要将这个程序送入内存;要编辑修改一个数据文件(如一个 Word 文档),需要把这个文件送入内存。暂时不需要执行的程序或不用的文件要存放在硬盘等外存上,以备需要时直接调入内存。文件(file)是存放在计算机外存上的相关数据的集合。操作系统要具备文件管理功能,对存放在外存上的大量文件(程序也是一种文件)进行有效的管理,以方便用户操作使用这些文件,并保证文件内容的安全。在规定文件命名规则、按层次组织文件的基础上,文件管理应具有文件存储空间管理、目录管理、文件的读写管理以及文件的安全保护等功能。

#### 1. 文件命名

同一外存上可能有很多文件,为了便于对文件的识别和管理,要给每个文件规定一个唯一的文件名。例如,班主任老师把所带班级每个学生的基本情况和联系方式输入计算机中,保存在自己的 U 盘上,起一个名字:“学生基本情况与联系方式”,这样就生成了一个学生联系方式文件,需要时就可以在计算机上打开这个文件,查看某个学生的基本情况或联系方式。

严格说来,一个规范的文件名包括主文件名和扩展名两部分,格式如下:

<主文件名>[.扩展名]

一个文件可以有扩展名,也可以没有扩展名(用方括号表示),但必须要有主文件名(用尖括号表示),如果有扩展名,要用点(.)与主文件名分开。

主文件名代表文件的特点,由用户根据文件内容命名,主文件名最好能反映代表的文件内容,做到见名知义,便于对文件的查找和管理。特别是管理的文件很多时尤显重要,如我们撰写的实验报告,主文件名分别命名为操作系统实验报告、数据结构实验报告和 Python 语言实验报告等,以后找起来就非常方便,看到名字就知道内容了。如果分别命名为 T11、T12 和 T13 等,找起来就比较麻烦,需要逐一打开这些文件查看内容,才能找到需要的那个文件。

扩展名代表文件属于哪一类,一般使用计算机系统已经规定好的一些名字,在使用一些软件系统建立文件时使用系统默认的扩展名即可,用户不必自己命名扩展名。表 5.1 列举了一些常用的文件扩展名。

表 5.1 常用的文件扩展名

| 扩展名      | 文件类型                        |
|----------|-----------------------------|
| doc/docx | Microsoft Word 文档文件         |
| xls/xlsx | Microsoft Excel 电子表格文件      |
| ppt/pptx | Microsoft PowerPoint 演示文稿文件 |
| pdf      | Adobe 文档文件                  |
| c        | C 语言程序源文件                   |
| cpp      | C++ 程序源文件                   |
| py       | Python 程序源文件                |
| exe      | 可执行程序文件                     |
| bmp      | 位图格式图片文件                    |
| jpg      | JPEG 格式图片文件                 |
| gif      | GIF 格式图片文件                  |
| wav      | Microsoft Windows 声音文件      |
| avi      | Microsoft Windows 视频文件      |
| mpg      | MPEG 格式视频文件                 |
| zip      | ZIP 格式压缩文件                  |
| rar      | RAR 格式压缩文件                  |

对于需要用户自己命名的主文件名部分,原来的 DOS 平台限制比较多,现在的 Windows 平台限制就少多了,汉字、英文字母和除 /、\、:、\*、?、“”、<、>、| 之外的其他符号都可以使用。

当文件比较多时,需要建立文件夹(子目录),把不同性质的文件分门别类地存放在不同的文件夹中,在加上文件命名时遵循见名知义的原则,会大大提高文件管理的效率,提高工作效率。需要注意的是,在同一个文件夹中,不允许有文件名完全相同的文件(主文件名和扩展文件名都相同),否则新文件的建立会覆盖旧文件,导致旧文件内容的丢失。

## 2. 按层次组织文件

外存的容量一般是比较大的,可以存放成千上万个文件,这么多的文件如果没有一个好的组织结构,会导致文件管理效率低下,如在上万个 Word 文档中找出“2023 年工作计划”文件并不是一件很容易的事情。如果记得住文件名,还可以用文件搜索的方式;如果连文件名都没有记住,就是一件很困难的事情了。

按层次组织文件,会大大提高文件管理效率,特别是文件查找效率。

**【例 5.1】** 张教授既承担教学工作,又承担科研工作。教学工作包括本科生教学工作和研究生教学工作,还要指导本科生和研究生提交的论文。科研工作包括科研项目和学术论文,论文有待发表论文和已发表论文。几年下来,光是 Word 文档就会积累下成百上千个,采用层次结构,会帮助张教授有效管理这些文件,提高工作效率。

在 Windows 系统中,可以通过逐层建立文件夹,并把不同文件放入不同文件夹的方式来实现文件的层次化管理,张教授可以建立的层次文件夹如图 5.2 所示。

Windows 等图形用户界面操作系统都包含有文件管理器之类的实用工具,文件管理器能够帮助用户容易地在各文件夹之间进行文件的移动、复制、重命名和删除等操作。

## 3. 文件存储空间管理

建立一个新的文件时,系统要为其分配相应的存储空间;删除一个文件时,系统要及时

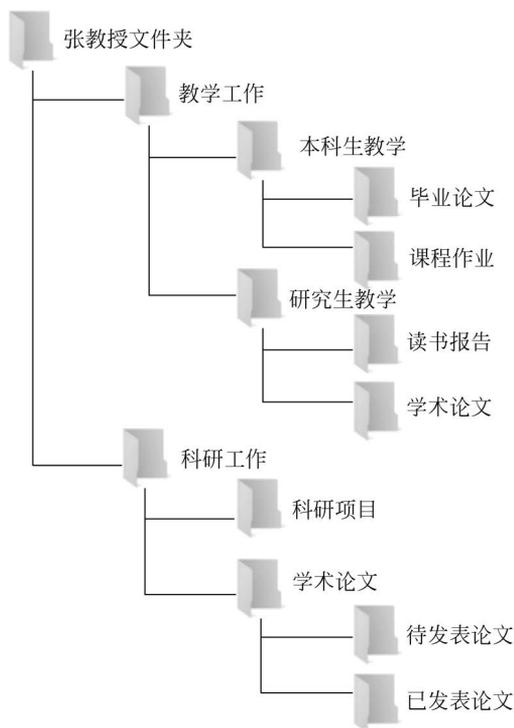


图 5.2 文件的层次结构

收回其所占用的空间。为了实现文件存储空间的管理,系统应设置相应的数据结构,用于记录存储空间的使用情况,作为新建文件分配存储空间的依据。为了提高存储空间的利用率和空间分配效率,对存储空间的分配通常是采用非连续分配方式,并以块为基本分配单位,块的大小通常为 512B~4KB,甚至更大。一个文件的内容可能存放在多段物理存储区域中,系统要有一种良好的机制把它们从逻辑上连接起来。

#### 4. 目录管理

一个文件夹对应外存中的一块存储区域,文件夹中的文件内容就存放在这块区域中,这块存储区域分成两个部分:目录区和数据区。目录区用于存放文件的目录项,每个文件有一个目录项,包含文件名、文件属性、文件大小、建立或修改日期、文件在外存上的开始位置等信息。数据区用于存放文件的实际内容。目录管理的主要任务是为每个文件建立目录项,并对由目录项组成的目录区进行管理,能有效提高文件操作效率。例如,只检索目录区就能知道某个特定的文件是否存在。删除一个文件只在该文件的目录项上做一个标记,告诉文件管理器该文件所占用的存储空间(包括目录区和内容区)可以被其他文件使用了。如果文件刚删除,还没有写入新的文件或还没有占用删除文件所使用的区域,这个被删除的文件是可以恢复的,用有关工具软件把删除标记改过来即可。

#### 5. 文件的读写管理

文件的读写管理就是根据用户的请求,从文件中读出数据或将数据写入文件。在进行文件读(写)时,首先根据用户给出的文件名,去查看文件目录区,找到该文件在外存中的开始存放位置;然后,对文件进行相应的读(写)操作。文件读写也称为文件存取。

## 6. 文件的安全保护

为了防止文件内容被非法读取和篡改,保证文件的安全,文件系统需要提供有效的安全保护机制。一般采取多级安全控制措施:一是系统级控制,没有合法账号和密码的用户不能进入计算机系统,自然也就无法访问系统中的文件;二是用户级控制,对有合法账号和密码的用户分配适当的文件存取权限,使其只能访问有访问权限的文件;三是文件级控制,通过设置文件属性(如只读)、密码保护、文件加密等措施来进一步限制用户对文件的存取。

文件管理功能由操作系统中的文件系统提供。Windows 文件系统主要有文件分配表(file allocation table, FAT)和新技术文件系统(new technology file system, NTFS)两种格式。在 Windows 9x 中,FAT16 支持的硬盘分区最大为 2GB;在 Windows 2000、Windows XP 中,FAT32 支持的硬盘分区最大为 32GB,但单个文件大小不能超过 4GB。NTFS 兼顾了磁盘空间的使用和访问效率,单个文件大小可以超过 4GB,硬盘分区可达到 2TB,在文件和文件夹权限设置、文件加密、设置磁盘配额和文件压缩等方面具有更好的性能。为解决 FAT32 不支持 4GB 以上文件的限制,引入扩展 FAT 文件系统(extended FAT, exFAT),exFAT 只适用于闪存等移动存储设备。

### 5.3.5 网络与通信管理功能

随着计算机网络的快速发展与普及,操作系统要具备网络与通信管理功能,以保证网络功能的正常、高效实现,主要包括资源管理、通信管理和网络管理等。

资源管理要保证网络资源的共享,管理用户对资源的访问,保证信息资源的安全性和完整性。通信管理就是通过通信软件,按照通信协议的规定,完成网络上计算机之间的信息传递。网络管理就是保证网络的安全、高效运行,并对出现的网络故障有合适的应对技术,包括故障管理、安全管理、性能管理、日志管理和配置管理等。

### 5.3.6 用户接口

为了方便用户使用操作系统,操作系统应向用户提供一个友好的接口。该接口通常是以命令或系统调用的形式供用户使用,前者提供给用户在直接操作时使用,后者则提供给用户在编程时使用。在 Windows 等操作系统中,又向用户提供了图形用户接口。

#### 1. 命令接口

为了便于用户直接或间接地控制自己的程序,操作系统向用户提供了命令接口。用户可通过该接口向计算机发出命令以实现相应的功能。该接口又可进一步分为联机用户接口和脱机用户接口。

(1) 联机用户接口。由一组键盘操作命令及对应的命令解释程序所组成。当用户在终端(键盘)或控制台上输入一条命令后,系统便立即转入命令解释程序,对该命令进行解释并执行该命令。在完成指定功能后,控制权又返回到终端或控制台上,等待用户输入下一条命令。DOS 操作系统提供的就是联机用户接口。

(2) 脱机用户接口。该接口是为批处理作业的用户提供的,也称为批处理用户接口。它由一组作业控制语言组成。批处理作业的用户不能直接与自己的作业交互作用,只能委托系统代替用户对作业进行控制和干预。早期使用的批处理操作系统提供脱机用户接口。

## 2. 程序接口

程序接口是为用户程序访问系统资源而设置的,是用户程序取得操作系统服务的唯一途径。现在的操作系统都提供程序接口,如 DOS 操作系统是以系统功能调用的方式提供程序接口的,为用户提供的常用子程序有 80 多个,可以在编写汇编语言程序时直接调用。Windows 操作系统是以应用程序编程接口(application programming interface, API)的方式提供程序接口的,WIN API 提供了大量的具有各种功能的函数,直接调用这些函数就能编写出各种界面友好、功能强大的应用程序。例如,用 Python 语言编写程序时,通过第三方库 pywin32 可以调用 WIN API 函数,实现相关功能。

## 3. 图形用户接口

虽然用户可以通过联机用户接口来取得操作系统的服务,并控制自己的应用程序运行,但要求用户能熟记各种命令的名字和格式,并严格按照规定的格式输入命令,这既不方便又费时间。于是,图形用户接口应运而生。

图形用户接口(graphical user interface, GUI)采用了图形化的操作界面,用非常容易识别的各种图标将系统的各项功能、各种应用程序和文件直观、逼真地表示出来。可通过鼠标、菜单和对话框来完成对各种应用程序和文件的操作。此时用户已完全不必像使用命令接口那样去记住命令名及格式,使用鼠标就能实现很多功能。用户被从烦琐且单调的操作中解放出来,能够为更多的非专业人员使用。Windows、macOS 等操作系统因提供方便用户使用的图形用户接口而得到广泛应用。

# 5.4 操作系统的启动过程

目前的 Windows 操作系统启动过程主要有两种模式:一种是基于基本输入输出系统(basic input output system, BIOS)的传统启动模式;另一种是基于统一可扩展固件接口(unified extensible firmware interface, UEFI)的新型启动模式。提出统一可扩展固件接口的主要目的是提供一组在操作系统启动之前在所有平台上一致的、正确的启动服务,被看作 BIOS 的代替者。从 Windows 8 开始支持 UEFI 启动模式。UEFI 启动模式具有更高的开机速度和运行性能以及更好的兼容性和可扩展性,操作配置也更为简单方便。

下面以 Windows 10 为例,分别说明 BIOS 模式和 UEFI 模式的计算机启动过程。

### 5.4.1 BIOS 模式启动过程

计算机接通电源后,可在屏幕显示“按 DEL 键进入 SETUP 界面”时按下 Del 键(有的是 F2 键)进行操作系统启动设置。通过改变存储在互补金属氧化物半导体(complementary metal oxide semiconductor, CMOS)中的系统启动参数设置,可更改系统的日期和时间、CPU 和内存频率、风扇转速、启动顺序等内容。

计算机的启动首先是从执行 BIOS 中自检程序开始的(BIOS 存储在 ROM 中,ROM 中主要存储的是自检程序和引导程序),这时可以看到计算机面板上的指示灯依次闪烁。自检顺序为:先进行 CPU、内存等关键部件的诊断测试,接着识别并检查显卡、硬盘等外部设备。自检通过后,BIOS 引导程序开始执行,寻找磁盘上的主引导记录(master boot record, MBR)。MBR 位于硬盘的 0 柱面、0 磁头、1 扇区(称为主引导扇区),大小与普通扇区一样

占用 512 字节, MBR 中存储着计算机的磁盘分区表与主磁盘引导程序。其主要功能是当磁盘启动后, 将系统控制权转交给硬盘分区表中的某个操作系统。然后 MBR 开始读取扇区内的硬盘分区表(disk partition table, DPT), 并找到活动分区(Windows 10 往往在安装系统时会建立一个 100MB 的活动分区)中的分区引导记录(partition boot record, PBR), 然后把控制权交给 PBR。PBR 搜索活动分区中的启动管理器 BOOTMGR, 并将控制权交给 BOOTMGR。

BOOTMGR 寻找活动分区中 boot 文件夹下的 BCD 文件(启动配置参数), 如当主分区位于 C 盘时, BCD 文件默认位于 C:\boot 路径下。找到 BCD 后, BOOTMGR 首先从 BCD 中读取 BOOTMGR 菜单的语言版本信息, 然后再调用相应语言的启动菜单, 之后在显示器上显示多操作系统选择画面(如果存在多个操作系统且系统设置的等待时间不为 0, 那么屏幕上就会显示多个操作系统的选择界面)。如果只有 Windows 10 系统, 那么将直接进入 Windows 10 系统, 不显示选择界面。接着, BOOTMGR 就会读取 Windows 10 系统所在盘中的 Windows\System32\winload.exe 文件, 并且将控制权交给 winload.exe。

winload.exe 加载 Windows 10 内核、硬件、服务等, 然后加载个人设置、桌面等信息, 从而完成整个 Windows 10 系统的启动。

#### 5.4.2 UEFI 模式启动过程

在传统 BIOS 模式下启动系统时需要进行硬件自检、寻找主引导分区、读取硬盘分区表、寻找活动分区、寻找活动分区中的分区引导记录, 再启动 BOOTMGR、执行 BCD 的过程。

UEFI 模式相对于 BIOS 启动的一大特点在于其不必进行硬件自检, 直接查找 EFI 系统分区(EFI system partition, ESP)。ESP 中主要存储 UEFI 模式下系统的启动引导程序。BIOS 启动过程中分区表存于 MBR 扇区之中, 在 UEFI 模式下, 出于兼容 BIOS 模式的考虑, 通常让硬盘从第 2 个扇区开始存放 GUID 分区表(GUID partition table, GPT), 第一个扇区仍存储 MBR。BIOS 模式下启动过程中需要查找主分区、寻找活动分区的分区引导记录, 在 UEFI 模式下, 可直接在 ESP 分区中找到后续启动过程所需要的文件, 首先找到并执行 ESP 分区下的 EFI/Microsoft/Boot/bootmgfw.efi。bootmgfw.efi 会找到同一相对路径下的 BCD 文件, 与 BIOS 模式中相似, 从 BCD 文件中读取语言版本信息, 然后调用相应语言的启动菜单, 之后在显示器上显示多操作系统选择画面, 待选择系统进入后读取 Windows\System32 文件夹下的 winload.efi, 加载内核启动系统。而上述过程中的 .efi 文件中调用的即是可扩展固件接口, 而上面提到的 .efi 文件即为可以直接在 UEFI 下运行的程序。



微课视频

### 5.5 操作系统实例

最初的计算机没有操作系统, 人们通过各种按钮和开关来直接控制计算机运行, 自第一个操作系统出现到现在, 经过近 70 年的发展, 推出了众多的操作系统, 为使用各种计算机提供了非常大的方便。下面对几个著名的操作系统进行简要介绍。

### 5.5.1 CP/M 操作系统

最早的操作系统是出现在 1956 年的 GM-NAA I/O。微型计算机的第一个操作系统则是诞生于 1974 年的控制程序/监控程序(control program/monitor,CP/M)。

CP/M 是加里·基尔达尔(Gary Kildall,1942—1994)领导的 Digital Research 公司为 8 位微型计算机开发的操作系统,它能够进行文件管理,具有磁盘驱动功能,可以控制磁盘的读写、显示器的显示以及打印机的打印输出,它是当时操作系统的标准。CP/M 曾经有多个版本,如运行在 Intel 8080 CPU 上的 CP/M-80、运行在 8088/8086 CPU 上的 CP/M-86、运行在 Motorola 68000 CPU 上的 CP/M-68K 等。

### 5.5.2 DOS 操作系统

1981 年,IBM 公司首次推出了 IBM-PC,该机上安装了微软公司开发的 MS-DOS 操作系统。该操作系统在 CP/M 的基础上进行了较大的扩充,增加了许多内部和外部命令,使该操作系统具有较强的功能及性能优良的文件系统。又因为它是配置在 IBM-PC 上的,随着该机型及其兼容机的畅销,MS-DOS 操作系统也就成了事实上的 16 位微机单用户单任务操作系统的标准。

微软-磁盘操作系统(Microsoft-disk operating system,MS-DOS)最早的版本是 1981 年 8 月推出的 1.0 版,一直发展到 1995 年的 7.0 版。在 1990 年微软推出 Windows 3.0 之前,DOS 一直占据微机操作系统的霸主地位,在和 Windows 抗争了几年之后,从 1995 年推出的 Windows 95 开始,DOS 逐步退出了操作系统市场。

早期的 DOS 是不支持汉字处理的,为了能在微型计算机上处理汉字,1983 年我国电子工业部第六研究所推出了基于 MS-DOS 的汉字磁盘操作系统 CC-DOS,以后又推出了若干版本。

### 5.5.3 Windows 操作系统

微软公司从 1983 年开始研发 Windows 操作系统,当时的目的是在 DOS 的基础上增加一个多任务的图形用户界面。1985 年和 1987 年分别推出了 Windows 1.0 和 Windows 2.0,但并没有得到用户的广泛认可,Windows 的流行是从 3.0 版开始的。

1990 年由微软公司推出的 Windows 3.0,以其易学易用、友好的图形用户界面,并能支持多任务和虚拟内存的优点,得以很快地流行开来,开始逐步占领微型计算机操作系统市场。Windows 95 在 1995 年 8 月正式发布,这是第一个不要求使用者先安装 MS-DOS 的 Windows 版本。从此 Windows 9x 便取代 Windows 3. x 以及 MS-DOS 操作系统,成为个人计算机的主流操作系统。

Windows 家族的另一个重要分支是 Windows NT,它是一种面向高端微型计算机的操作系统,与支持个人应用的 Windows 9x 有根本的区别,采用客户机/服务器与层次式结合的模式,支持多进程并发,有较强的内置网络功能和较高的系统安全性,主要运行在小型计算机和服务器上。

Windows 2000 是在 Windows NT 5.0 的基础上修改和扩充而成的,分为 Windows 2000 Professional 和 Windows 2000 Sever 两种版本,前者是面向普通用户的,后者则是面

向网络服务器的,能够充分发挥 32 位微型计算机的硬件性能,使其在处理速度、存储能力、多任务和网络计算支持等方面具有小型计算机的性能。

2001 年 3 月,微软公司正式宣布把个人用版本 Windows 98、Windows ME 和商用版本 Windows 2000 合二为一,推出新的版本 Windows XP(eXPerience)。2003 年 3 月推出的 Windows Server 2003 是广泛应用于服务器的操作系统。

之后陆续推出了 Windows Vista、Windows 7、Windows 8、Windows 10、Windows 11 和 Windows Server 2008、Windows Server 2012、Windows Server 2016、Windows Server 2018、Windows Server 2019、Windows Server 2022 等版本。

自 DOS 退出操作系统市场后,Windows 成为人们使用最多的微机操作系统。根据 Statcounter 2021 年 7 月的统计,全球桌面操作系统领域中,Windows 各版本的市场占有率合计为 74.79%,位居第二名的是苹果公司的 macOS,其市占率为 16.16%。

#### 5.5.4 macOS 操作系统

macOS 是一种运行于苹果 Macintosh 系列计算机上的操作系统,它也是首个在商用领域成功运用图形用户界面的操作系统。2014 年发布的版本是 OS X 10.10 Yosemite 版,采用了扁平化的设计风格,增加了大量的新特性,针对跨设备跨平台无缝切换的需求,提供了全新的 Handoff(可实现在不同设备中持续同一工作)、iCloud Drive(云盘)等功能,大大提高了在 mac、iOS 多设备之间使用的连续性。2023 年 6 月 6 日,苹果公司在 2023 苹果全球开发者大会上推出 macOS 13.5 版本(macOS Sonoma),该版本新增功能:支持桌面端的小部件,允许用户自定义壁纸作为屏保;为 mac 设备引入了 Game Mode,该模式适用于所有 mac 平台上的游戏;进一步改进了视频会议效果,可以在会议中使用增强现实(AR)特效。

#### 5.5.5 UNIX 操作系统

UNIX 操作系统是一种典型的多用户多任务型操作系统,是一个能在微型计算机、工作站、大型计算机等各种机型上使用的操作系统。

UNIX 操作系统起源于美国电报电话公司(AT&T)贝尔实验室在 1969 年开发的一种分时操作系统,最早的工作集中在文件管理和进程控制上。1970 年将该系统移植到了小型计算机 PDP-11 上,吸收了分时操作系统 MULTICS 的技术精华,定名为 UNIX。1971 年 11 月 3 日,UNIX 第 1 版(UNIX-V1)正式诞生。1973 年 C 语言出现后,用 C 语言改写的 UNIX 具有非常好的可读性和可移植性,为其推广普及奠定了基础。20 世纪 70 年代中后期,UNIX 源代码的免费获取引起了大学和公司的兴趣,更多人的参与为 UNIX 的改进、完善和普及起了重要作用,最著名的是加州大学伯克利分校的 BSD 版本。从 1977 年开始,各公司陆续推出了多种 UNIX 的商业化版本,如 SUN 公司的 SUN-OS 和 Solaris、微软公司的 XENIX、DEC 公司的 ULTRIX、IBM 公司的 AIX、HP 公司的 HP/UX 以及 AT&T 公司的 UNIX System III、UNIX System V、UNIX SVR 4.0 和 UNIX SVR 4.2 等。众多 UNIX 版本的出现,促进了 UNIX 的快速发展和应用普及,但也出现互不兼容的问题,针对此问题制定了一些 UNIX 开发标准,促进了 UNIX 的标准化。

进入 20 世纪 90 年代后,由于多处理器系统和计算机网络技术的发展,UNIX 也在适应这一发展趋势,UNIX 开始支持多处理器系统和计算机网络,配置了图形用户界面,安全性

也得到进一步加强。

### 5.5.6 Linux 操作系统

Linux 是芬兰赫尔辛基大学的一个大学生莱纳斯·托瓦兹(Linus Torvalds)在 1991 年编写的一个操作系统内核,现在托瓦兹已成为芬兰著名的计算机科学家。托瓦兹在学习操作系统课程时自己编写了一个操作系统原型(这就是最早的 Linux),并把这个原型系统放在互联网上,允许自由下载,许多人对这个系统进行了改进、扩充和完善,他们上传的代码和评论对 Linux 的发展做出了重要贡献。于是,Linux 从最初的一个人的作品变成了在互联网上由无数志同道合的程序员们共同参与的一场软件开发活动。Linux 遵从国际上相关组织制定的 UNIX 标准 POSIX。它的结构、功能以及界面都与经典的 UNIX 并无两样。然而 Linux 的源代码完全是独立编写的,与 UNIX 源代码无任何关联。Linux 继承了 UNIX 的全部优点,而且还增加了一条其他操作系统不曾具备的优点,即 Linux 源代码全部开放,并能在网上自由下载。

现在,Linux 操作系统是一种得到广泛应用的多用户多任务操作系统,许多计算机公司如 IBM、Intel、Oracle、SUN 等都大力支持 Linux,各种常用软件纷纷移植到 Linux 平台上。Linux 和 Windows、UNIX 一起成为操作系统市场的主流产品。

### 5.5.7 VxWorks 操作系统

VxWorks 是嵌入式操作系统的优秀代表,是美国 Wind River 公司的产品。VxWorks 支持各种工业标准,包括 POSIX、ANSI-C 和 TCP/IP。VxWorks 的核心是一个高效率的微内核,支持各种实时功能,包括快速多任务处理、中断支持、抢占式和轮转式调度。微内核设计减轻了系统负载,并可快速响应外部事件。

2011 年 11 月 26 日发射并于 2012 年 8 月 6 日着陆(历经 8 个半月,在太空飞行 5.69 亿千米)的“好奇号”(Curiosity)火星探测器上使用的就是 VxWorks 操作系统。VxWorks 负责火星探测器的全部飞行控制,包括飞行纠正、载体自旋和降落时的高度控制等,还负责数据收集和与地球上控制中心的通信工作。

VxWorks 广泛应用于网络通信、医疗设备、消费电子产品、交通运输、工业控制、航空航天和多媒体设备等领域。

## 5.6 小 结

计算机之所以能够进入千家万户,逐渐成为人们学习、工作和娱乐的基本工具,主要得益于三方面:一是微型计算机的出现和快速发展,使计算机的成本快速下降;二是操作系统功能的不断完善和强大,使计算机的操作使用越来越简单方便;三是互联网的快速发展和网络服务的不断丰富,网络应用给人们带来了很大的方便。

作为最重要的系统软件,操作系统对于充分利用计算机资源、保证程序的高效正确执行、方便用户使用具有重要作用。操作系统在发展过程中,追求的一个主要目标是不断提高 CPU 等设备的利用率,以此来提升计算机的整体性能。操作系统的特征包括并发性、共享性、虚拟性和异步性。操作系统具体包括处理器管理、存储器管理、文件管理、设备管理和网

络管理等功能,并提供良好的用户接口。目前常用的通用操作系统有 Windows、macOS、UNIX、Linux 和 VxWorks 等。

## 拓展阅读: 比尔·盖茨与微软公司



1955年10月28日,比尔·盖茨(Bill Gates)出生于美国西北部华盛顿州的西雅图,自小酷爱数学和计算机。保罗·艾伦(Paul Alan,1953—)是他最好的中学校友,两人经常在学校的一台 PDP-8 小型计算机上玩三连棋的游戏。

1972年的一个夏天,他们从一本《电子学》杂志上得知 Intel 公司推出一种叫 8008 的微处理器芯片。两人不久就使用该芯片组装出一台机器,可以分析城市内交通监视器上的信息。1973年比尔·盖茨考入哈佛大学,保罗·艾伦则在波士顿一家叫“甜井”的计算机公司找到一份编写程序的工作,两人经常在一起探讨计算机的事情。1974年春天,当《电子学》杂志宣布 Intel 公司推出比 8008 芯片更快的 8080 芯片时,比尔和保罗预见到类似 PDP-8 的小型计算机的末日快到了。他们看到了新芯片背后适应性强、成本低的个人计算机的发展前景。

1975年1月的《大众电子学》杂志封面上 Altair 8080 微型计算机的图片深深地吸引了保罗·艾伦和比尔·盖茨。这台世界上最早的微型计算机,标志着计算机新时代的开端,这是一台基于 8080 微处理器的微型计算机。还在哈佛上学的盖茨看到了商机,他要给 Altair 开发 BASIC 语言,盖茨和艾伦在哈佛大学计算机中心奋战了 8 周,为 8080 配上 BASIC 语言,此前从未有人为微型计算机编过 BASIC 程序,艾伦亲赴 Altair 8080 的生产厂商 MITS 进行演示。这年春天,艾伦进入 MITS,担任软件部主管。学完大学二年级课程,盖茨也进入 MITS 工作。

微软(Microsoft)公司诞生于 1975 年,但当时微软公司与 MITS 公司之间的关系十分模糊,可以说微软公司“寄生”于 MITS 之上。1975 年 7 月下旬,他们与 MITS 签署了协议,为期 10 年,允许 MITS 公司在全世界范围内使用和转让 BASIC 及源代码。根据协议,盖茨他们最多可获利 18 万美元。借助 Altair 的风行,BASIC 语言也推广开来,同时,微软公司又赢得了另外两个大客户。盖茨和艾伦开始将更多的精力放在自己的公司上。正是 MITS,确定了盖茨和艾伦作为程序员的地位,跻身这个新兴行业。借助于 MITS 公司,积累了微软公司发展的第一批资金,同时他们目睹并参与了 MITS 公司从设计到生产、从宣传到销售服务的全过程,培养了市场意识。

艾伦离开 MITS 公司后不久的 1977 年元旦,盖茨退学了。

1980 年,IBM 公司准备进军 PC 市场,由 IBM 公司研制硬件系统,由微软公司开发一套方便用户使用 PC 的操作系统。1981 年 6 月,MS-DOS 的开发工作基本完成,8 月,IBM PC 问世,这台个人计算机主频是 4.77MHz,CPU 是 Intel 公司的 8088 芯片,内存为 64KB,操作系统就是微软的 MS-DOS。DOS 是磁盘操作系统(disk operation system)的简称,在 1981 年到 1995 年间占据 PC 操作系统的统治地位,版本从 1. x 发展到 7. x。

1985 年 6 月,微软公司和 IBM 公司达成协议,联合开发 OS/2 操作系统。根据协议,IBM 公司在自己的计算机上可免费安装,而允许微软公司向其他计算机厂商收取 OS/2 的

使用费。当时 IBM 公司在 PC 市场拥有绝对优势,兼容机份额极低,之后兼容机市场却逐步扩大,到 1989 年兼容机占据了市场 80% 的份额。微软公司在操作系统的许可费上,短短几年就赢利 20 亿美元。

相对于以前的操作系统,DOS 取得了很大的成功,但在使用过程中也逐渐暴露出其功能比较弱、安全性低、使用不方便的缺点,作为单用户单任务型操作系统,几乎没有安全性措施,使用者需要记忆大量的英文单词式的命令。微软公司从 1981 年就开始开发后来称之为 Windows 的操作系统。希望它能够成为基于 Intel x86 微处理芯片计算机上的标准图形用户接口操作系统。在 1985 年和 1987 年分别推出 Windows 1.0 版和 Windows 2.0 版。但是,由于当时硬件水平和 DOS 操作系统的风行,这两个版本并没有得到用户的广泛认可。此后,微软公司对 Windows 的内存管理、图形界面做了重大改进,使图形界面更加美观并支持虚拟内存。1990 年 5 月推出的 Windows 3.0 开始得到人们的认可。

一年之后推出的 Windows 3.1 对 Windows 3.0 作了一些改进,引入一种可缩放的 TrueType 字体技术,改进了系统的性能;还引入了一种新设计的文件管理程序,改进了系统的可靠性。Windows 3.0 和 Windows 3.1 都必须运行于 MS-DOS 操作系统之上。

几年的应用实践使用户逐渐熟悉和青睐于 Windows,可以与 DOS 分离了,1995 年微软公司推出新一代操作系统 Windows 95,它可以独立运行而无须 DOS 支持。Windows 95 是操作系统发展史上一个非常重要的版本,它对 Windows 3.1 版作了许多重大改进,包括更加优秀的、面向对象的图形用户界面,单击就能完成大部分操作,极大地方便了用户的学习和使用;全 32 位的高性能的抢先式多任务和多线程;内置的对 Internet 的支持;更加高级的多媒体支持,可以直接写屏并能很好地支持游戏;即插即用,简化用户配置硬件操作,并避免了硬件上的冲突;32 位线性寻址的内存管理和良好的向下兼容性等。

目前,微软公司的主要产品如下。

操作系统 Windows 系列: Windows 7、Windows 8、Windows 10、Windows 11、Windows Server 2016、Windows Server 2019、Windows Server 2022。

数据库管理系统 MS SQL Server: 一种可用于网络环境的大型数据库管理系统,新版本还具备一定的数据仓库和数据挖掘功能。

办公软件 Office 系列: 文字处理软件(Word)、电子表格软件(Excel)、桌面数据库(Access)、幻灯片制作软件(PowerPoint)、个人邮件管理软件(Outlook)、网页制作软件(FrontPage)等。

网页浏览器 Internet Explorer (IE)、Edge: IE 是 Windows 95 开始微软各版本 Windows 的默认浏览器,从 Windows 10 开始,自带浏览器改为 Edge。

Microsoft Azure 微软云计算: 为全球客户提供多种计算、数据服务、应用服务及网络服务,帮助个人开发者、初创公司、企业机构快速开发、部署、管理应用程序。

媒体播放器 Windows Media Player: 用于播放音频和视频。

开发工具包 Visual Studio: 包括 Visual Basic、Visual C++、Visual C# 等,目前已发布用于 .NET 环境的编程工具 Visual Studio .NET。

在线服务 MSN(Microsoft Network): 主要用于即时通信(网上聊天等)。

近几年,微软公司在向人工智能领域拓展,在认知服务、自然语言处理、语音识别、聊天机器人、智能化办公软件都有平台或产品推出。正在把 ChatGPT 技术逐步融入 Windows

11(包括 Edge)、Office 365 等软件中,使用户有更好的 AI 体验。2023 年 2 月,微软推出由 ChatGPT 支持的最新版本人工智能搜索引擎 Bing(必应)和 Edge 浏览器。

微软 2023 财年的营业收入为 2119.15 亿美元,与 2022 财年相比增长 7%,2023 财年的运营利润为 885.23 亿美元,与 2022 财年相比增长 6%。

## 习 题 5

### 一、填空题

1. 在通用操作系统中,可能同时存在 3 类任务: \_\_\_\_\_ 任务、\_\_\_\_\_ 任务和 \_\_\_\_\_ 任务, \_\_\_\_\_ 任务的优先级最高, \_\_\_\_\_ 任务的优先级最低。
2. 多处理器操作系统,主要有两种模式: \_\_\_\_\_ 模式和 \_\_\_\_\_ 模式。
3. 网络操作系统应具有 4 方面的功能: \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
4. 操作系统的特征包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
5. 处理器管理功能包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
6. 存储器管理功能包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
7. 设备管理功能包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
8. 文件管理功能包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
9. 一个规范的文件名包括两部分: \_\_\_\_\_ 和 \_\_\_\_\_。
10. 写出 5 种常用文件的扩展名: \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
11. 操作系统的接口包括 \_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
12. Windows 10 有两种启动模式: \_\_\_\_\_ 模式和 \_\_\_\_\_ 模式。

### 二、名词解释

操作系统、单用户单任务操作系统、单用户多任务操作系统、多用户多任务操作系统、GUI、API、虚拟内存、虚拟设备、进程、线程。

### 三、简答题

1. 简述操作系统的特征。
2. 简述操作系统的主要功能。
3. 对比说明几种目前常用的操作系统。

## 思 考 题 5

1. 对比说明网络操作系统、多处理器操作系统和分布式操作系统的概念。
2. 在操作系统的发展历程中,主要是在解决什么问题? 是如何解决的?