

# 第 1 章 操作系统引论

计算机系统由硬件和软件两部分组成，操作系统（Operation System, OS）是配置在计算机硬件上的第一层软件，是对硬件的第一次扩充。它在计算机系统中占据重要的地位，为其他所有软件奠定了重要基础。操作系统已成为所有的计算机必须配置的软件。

## 1.1 操作系统概述

### 1.1.1 计算机系统

计算机系统就是按人的要求接收和存储信息，自动进行数据的处理和计算，并输出结果信息的机器系统，它是一个相当复杂的系统，即使相当普及的个人计算机也是如此。一个计算机系统由硬件（子）系统和软件（子）系统组成。其中，硬件系统是借助电、磁、光和机械等原理构成的各种物理设备的有机结合，是系统赖以工作的实体；软件系统是各种程序和文件，用于指挥整个系统按照指定的要求工作的。

图 1-1 是一般计算机系统的层次结构。硬件是进行信息处理的实际物理装置。最外层是使用计算机的人，即用户。人与硬件系统的接口是软件系统，它大致分为系统软件、实用软件和应用软件 3 层。系统软件包括操作系统和各种语言处理程序等。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软件改造的计算机。这个层次结构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。

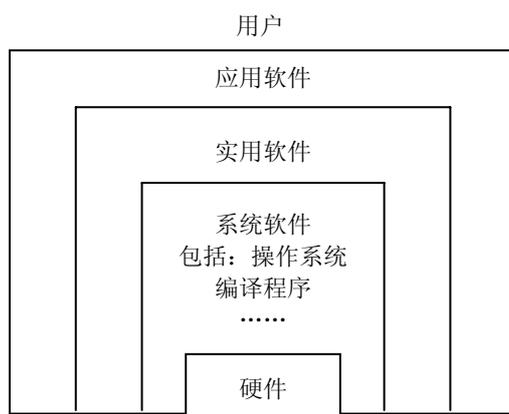


图 1-1 计算机系统的层次结构

### 1.1.1.1 计算机硬件

硬件是计算机系统中各种物理设备的总称,是计算机系统快速、可靠地自动完成工作任务的基础。其主要逻辑功能是完成信息变换、信息存储、信息传送和信息处理等。硬件系统主要由运算器、控制器、主存储器、输入设备/输出设备(I/O设备)、辅助存储器和总线等功能部件组成。

(1) 运算器。其主要功能是进行算术运算和逻辑运算。操作时,运算器从主存储器取得数据,执行指令指定的运算,将所得的运算结果留在运算器以备下次运算时使用,或存入主存储器。整个运算过程是在控制器的控制下自动进行的。

(2) 主存储器。主要功能是储存二进制信息。主存储器与运算器、控制器等快速部件直接交换信息。从主存储器应能快速读出信息,并送到其他功能部件中,或将其他功能部件处理过的信息快速写入主存储器。

(3) 控制器。主要功能是按照机器代码程序的要求,控制计算机各功能部件协调一致地动作,即从主存储器取出程序中的指令,对该指令进行分析和解释,然后向其他功能部件发出执行该指令所需要的各种时序控制信号。如此连续运行下去,直到程序执行完毕。控制器与运算器一起构成中央处理器(CPU),中央处理器与主存储器一起构成主机。

(4) 输入设备。主要功能是将用户信息(数据、程序等)变换为计算机能够识别和处理的二进制信息形式。信息载体通过相应的输入设备,将信息(如文字、图像、影像、语音等)变换为电信号被计算机接收、存储和处理。输入设备的种类很多,如键盘、鼠标、扫描仪等。

(5) 输出设备。主要功能正好与输入设备相反,是将计算机能够识别和处理的二进制信息变换为用户需要的信息形式。输出信息的形式有多种,如文字、图像、表格等。输出设备的形式也有多种,如打印机、绘图仪、显示器等。

(6) 辅助存储器。主要功能是存储主存储器难以容纳但又为程序所需的大量信息。它的特点是存储容量很大,存储成本很低,但存取速度较慢。它不能与中央处理器直接交换信息,而是必须借助主存储器。辅助存储器一般为磁带、磁盘、光盘等。

(7) 总线。是将数据从一个部件传送到另一个部件的一束连接线。在一台计算机系统中可能有多条传输速度和功能都不同的总线。总线包括总线自身和总线控制器,在一条总线上的某个时刻传输什么数据,以及把数据传送到哪里去,都是由总线控制器控制的。微型计算机系统普遍采用总线结构。CPU通过系统总线(含地址、数据和控制信息)与主存储器、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,IBMPC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位地址总线,IBMPC/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位地址总线。

(8) I/O通道。主要是控制I/O设备的工作过程。具体功能是:向I/O设备发出动作命令;控制数据的输入/输出;检查I/O设备的状态等。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,以不同的工作方式分别控制低速、中速和高速I/O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。目前大、中、小型计算机经常采用I/O通道,即这些计算机系统是以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。

当代计算机硬件正向微型化、智能化方向发展。硬件系统的组织结构也在不断地改进,如多机系统、分布式系统、计算机网络以及嵌入式系统等。

### 1.1.1.2 计算机软件

计算机软件是指程序和与程序相关的文档的集合，是计算机系统的重要组成部分。按功能划分，计算机软件可分为系统软件、实用软件和应用软件 3 种。

(1) 系统软件。指由计算机生产厂家提供、具有通用功能的那些软件，比如：操作系统、语言编译程序、文本编辑程序、调试程序、连接程序和系统维护程序等。系统软件的功能是为上层软件提供服务，并为管理员提供对系统进行日常维护的手段。例如：Microsoft 32 位 Windows 操作系统的基础是 32 位内核软件，由内核进行系统调度和内存管理。另外由于 32 位操作系统支持 32 位设备驱动器，因此操作系统和设备间的通信就更为迅速。32 位 Windows 操作系统的其他许多特性来源于它对 Win32 API 的支持。这个 API 集合只能由 32 位内核系统，如 Windows NT 和 Windows 95/98 方能完全实现。

(2) 实用软件。指支撑其他软件的编制和维护的软件。由一组实用程序组成，如数据库管理系统、各种接口软件和工具组。当然，操作系统和编译程序等系统软件也可以算做实用软件。随着计算机科学技术的发展，软件的编制和维护代价在整个计算机系统中所占的比重越来越大，远远超过硬件。因此，实用软件的研究具有重要的意义。

(3) 应用软件。指为了解决实际问题而研制的那些软件，它涉及计算机应用的各个领域，比如：各种管理软件、用于工程计算的软件包、辅助设计软件以及过程控制软件等。

### 1.1.2 操作系统的作用

在计算机系统的层次结构中，操作系统是计算机硬件上加载的第一层软件，是对计算机硬件功能的首次扩充，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。操作系统统一管理和支持各种软件的运行，其他软件必须建立在操作系统之上。因此，操作系统在整个计算机系统中具有极其重要的作用。操作系统的作用可以从以下两方面来说。一方面，操作系统是系统资源的管理者。系统资源包括硬件资源和软件资源，归纳起来可分为 4 类：处理器、存储器、I/O 设备以及信息（数据和程序）。相应地，操作系统的主要作用就是对这 4 类资源进行有效的管理。如：在多用户多任务系统中同时存在着多个程序，它们在运行过程中会使用系统中的各种资源。操作系统就是系统资源的管理者和仲裁者，它负责在各个程序间调度和分配资源，保证系统中各种资源得到有效利用。再如：操作系统要管理文件系统，如文件在磁盘上的位置，所占磁盘空间的大小等。其实这就是操作系统在管理信息，因为文件系统是由计算机中的程序和数据构成的。所以，操作系统是其他所有软件的管理者。为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。

在这里，操作系统管理的含义是多层次的，它对每一种资源的管理必须完成以下几项工作。

(1) 监视资源。该种资源共有多少，资源的状态如何，已经分配出去的资源有多少，都分配给了谁，可供分配的还有多少，使用资源的历史纪录等。

(2) 确定某种资源的分配策略。即决定将资源分配给谁，何时分配，分配多少，如何回收资源等。

(3) 分配资源。按已决定的分配策略分配资源，并进行相应的管理事务处理。

(4) 回收资源。在使用者放弃这种资源后，对资源进行处理。如果是可重复使用的资源，则回收、整理，进而进行再利用。

另一方面，操作系统是用户与计算机系统之间的接口。因为操作系统处于用户和计算机硬件系统之间，所以用户通过操作系统来使用计算机。从用户的角度看，计算机系统配置了操作系统后，由于操作系统隐蔽了硬件的复杂细节，用户会感到机器使用起来简单、方便、快捷且安全可靠了。或者说，正是操作系统为用户提供了一台功能经过扩展了的机器——“虚拟机”，使用户感觉使用机器更简单，更容易。一个好的操作系统一定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。应当注意，操作系统是个系统软件，因而这个接口是个软件接口。用户可以通过命令方式、图形方式和系统调用方式使用计算机。命令方式是指用户向作业发出命令以控制作业的运行。图形方式是指计算机系统通过简单的图标将系统的各项功能、各种程序和文件直观地表示出来，用户可通过鼠标、菜单和对话框来完成对应用程序和文件的操作。系统调用方式是为用户程序在执行中访问系统资源而设置的，是用户程序取得操作系统服务的惟一途径。

至此，可以把操作系统定义为：“操作系统是控制和管理计算机硬件和软件资源、合理地组织计算机工作流程以方便用户有效地使用计算机的程序集合。”

### 1.1.3 操作系统的功能

前面已经从计算机系统层次结构的观点，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、应用软件与用户的关系，说明了引入操作系统的主要目的是最大限度地发挥计算机系统资源的使用效率和方便用户的使用。为了实现上述目的，操作系统通常具备以下几个方面的重要功能。

#### 1.1.3.1 处理机管理

处理机管理的主要任务是对处理机进行分配，并对其进行有效的控制和管理。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。或者当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。在多道程序环境下，处理机的分配和运行是以进程为单位的，因而，对处理机的管理可归结为对进程的管理。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。对进程的管理主要包括：进程控制、进程同步、进程通信和进程调度等方面，在以后章节中会详细介绍。

#### 1.1.3.2 存储管理

存储器管理的主要任务是为多道程序的运行提供良好的环境，方便用户使用存储器，提高存储器的利用率，以及能从逻辑上扩充内存。内存（又称主存）是计算机系统中另一个宝贵的硬件资源。虽然内存芯片的集成度不断地提高、价格不断地下降，但因需求量大，所以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共享内存，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，都是存储管理的范畴。因此，存储管理应具有以下功能：内存分配、地址映射、内存保护和内存扩充。

#### 1.1.3.3 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，就能够方便地对这些设备进行操作。设备管理的另一任务就是利用中断技术，DMA 技术和通道技术，使外围设备尽可能地与 CPU 并行工作，以提高设备的使用效率并提高整个计算机系统的运行速度。为实现上述任务，设备管理应该具有缓冲管理、设备分配、设备驱动、设备独立性和虚拟设备等功能。

#### 1.1.3.4 文件管理

程序和数据以文件的形式存放在外存储器（如磁盘、光盘、磁带）中，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，而且各种系统软件甚至操作系统程序本身也是以文件的形式存储的。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。为此，在操作系统中必须配置文件管理机构，有效的组织、存储、保护文件，以使用户方便、安全地访问它们。总的来说，文件管理应具有文件存储空间管理、目录管理、文件的读写和存取控制等功能。

#### 1.1.3.5 作业管理

作业管理的任务是为用户提供一个环境良好的使用系统，使用户有效地组织自己的工作流，并使整个系统高效运行。所谓作业，就是用户在一次算题过程中，或一次事务处理中，要求计算机系统所做工作的集合。例如，用 C 语言编制一个算题程序，要经过下列步骤：把源程序读入系统，对源程序进行编译、链接，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。

操作系统的上述几种功能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。可以用一个基本层次式的操作系统模型来描述它们之间的关系，如图 1-2 所示。其中较高层的管理功能要依赖于较低层的管理功能。另外，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。



图 1-2 操作系统核心层次结构模型

#### 1.1.3.6 其他功能

随着社会的发展和技术的进步，对操作系统的要求也越来越高。除了上述 5 项管理功能

之外，操作系统还应该向用户提供其他一些必不可少的功能，如系统安全和网络通信等。

安全功能的主要任务是操作系统为用户提供一套信息保护的手段，保证系统中的信息不会被非法或未授权的用户窃取、修改或破坏。在当今社会，人们常常使用计算机保存、加工、处理和传送大量重要信息，操作系统的安全性则直接影响到这些信息的安全，因此安全性已经成为衡量操作系统性能的重要指标。以前，在早期的 DOS 操作系统中，几乎没有安全性措施。而现在的 Windows 操作系统中采取了很多安全性措施，如开机密码、屏保口令等。需要说明的是，由于不同计算机系统对安全性的要求不同，操作系统应该采取不同的安全措施。对安全性要求较高的系统，操作系统就要采取较多的安全措施，而对安全性要求较低的系统，操作系统可采取较少的安全措施。总之以够用为原则，因为采取的安全措施越多，操作系统的工作就会越复杂，操作系统的工作效率就会越低，进而导致整个计算机系统工作效率的降低。

网络通信功能主要指操作系统为用户提供一组支持网络通信协议的程序和接口，保证用户能快捷、安全地进行网络通信。目前处在 Internet 时代，任何计算机系统都离不开网络通信，因此操作系统必须提供网络通信支持。如，在 20 世纪 70 年代，在 UNIX 操作系统中第一次实现了早期的 TCP/IP 协议。Windows 等现代操作系统中不仅支持基于 TCP/IP 协议的通信，还支持 UDP 等多种通信方式。

从系统层次结构和资源管理两个角度考察操作系统后，就可以较完整地了解操作系统基本含义了。

## 1.2 操作系统发展历史

操作系统操作系统的形成迄今已有约 50 年的时间，在 20 世纪 50 年代中期出现了第一个简单的批处理操作系统 (Batch Processing System)，到 20 世纪 60 年代中期产生了多道批处理操作系统，不久又出现了基于多道的分时操作系统 (Time-Sharing System)。20 世纪 80 年代是计算机操作系统和网络操作系统形成和大发展的年代。

### 1.2.1 操作系统发展的主要动因

在短短的 50 年中，操作系统得到了飞速发展，其主要动因可归纳为以下几个方面。

#### 1.2.1.1 不断提高计算机各种资源利用率的需要

提高系统资源利用率是操作系统追求的基本目标之一。在没有配置操作系统的计算机系统中，由于人需要手工操作计算机，而人的操作速度与计算机中各种资源的运行速度相差很多，这使得诸如 CPU、I/O 设备等各种资源经常会处于空闲状态，而且当时系统资源特别昂贵，因此人们千方百计的提高系统各种资源的利用率。配置了操作系统后，能尽可能使 CPU 和 I/O 设备保持忙碌状态进而得到充分利用。另外，操作系统还可以采取适当的策略合理地组织计算机工作流程，从而可以进一步提高系统资源的利用率。

#### 1.2.1.2 方便用户的需要

方便用户是操作系统追求的另一个基本目标。未配置操作系统的计算机系统极难使用，因为用户必须用计算机能够识别的 0 和 1 代码编写程序，程序才能在计算机上运行。另外程序运行过程中需要用户输入的数据和运行结果也都是二进制形式。这些不仅使用户编程麻烦，而且非常容易出错。配置操作系统后，人们可以用类自然语言编写高级语言程序，而不用管

计算机是否能理解，编译程序能够把高级语言程序翻译成计算机能够识别的二进制代码，不仅使计算机变得易学易懂，而且维护程序的代价降低很多，极大地方便了用户。后来，人们又不断地在方便使用计算机方面做出努力，出现了使人与计算机能及时交互的分时操作系统，出现了使人能透明地使用整个计算机网络资源的网络操作系统和分布式操作系统。

#### 1.2.1.3 计算机硬件的快速发展

计算机硬件的不断更新换代，推动了操作系统不断发展。自 1946 年诞生第一台计算机至今的短短 50 多年中，计算机硬件技术得到了突飞猛进的发展。主要分为 4 个阶段：第 1 代是电子管计算机时代（1946 年~1957 年）。这一代计算机的运算速度约为几千至几万次/秒，体积大，成本高，可靠性低。第 2 代是晶体管时代（1958 年~1964 年）。这一代计算机的运算速度提高到几万至几十万次/秒；体积缩小，成本降低，可靠性提高。第 3 代是集成电路时代（1965 年~1970 年）。这一代计算机的运算速度提高到几十万至几百万次/秒；体积进一步缩小，成本进一步降低，可靠性进一步提高。第 4 代是大规模集成电路时代（1971 年~至今）。这一代计算机的运算速度提高到几千万至几百亿次/秒甚至更高，可靠性、体积、成本等方面都得到更进一步改善。这就使得计算机的性能不断提高，计算机的应用场合急剧扩大，甚至到了当今计算机无所不在，从而推动操作系统的性能也必须不断提高，功能也必须不断完善。如，当微机从 8 位发展到 16 位，进而又发展到 32 位时，微机操作系统也相应的由 8 位微机操作系统发展为 16 位微机操作系统，进而发展到 32 位微机操作系统。操作系统的各方面功能都有了显著提高，最明显的就是操作系统管理的内存空间从  $2^8$  字节扩大到  $2^{32}$  字节。

#### 1.2.1.4 计算机体系结构的不断发展

计算机体系结构的不断发展也是推动操作系统发展的动力。体系结构是指计算机硬件的系统结构，是机器语言、汇编语言程序设计者看到的计算机硬件的抽象。很明显，计算机体系结构不同，操作系统的管理功能就不同。例如，在总线体系机构中，操作系统要管理总线和总线控制器，保证总线能够正常传输数据。在阵列处理机结构中，操作系统要管理阵列中的多个处理机，使它们能在一定条件下并行工作。在多处理机系统中，就要有多处理机操作系统，控制和协调多个处理机的工作。

总之，不同的硬件结构、不同的应用环境、不同的用户需求，就应有不同类型的操作系统，以实现不同的追求目标。

### 1.2.2 手工操作阶段

在电子管计算机时代，数字计算机是由成千上万个电子管和许多开关组成的庞然大物。此时并没有操作系统，程序设计全部用机器语言，由手工控制作业的输入输出，通过控制台开关启动程序运行。到 20 世纪 50 年代，出现了穿孔卡片和纸带，程序员可以把机器语言程序穿在纸带或纸卡上而不再使用插板。

此时用户使用计算机的过程大致如下：先把程序纸带装上输入机，启动输入机把程序和数据送入计算机，然后通过控制台开关启动程序运行，计算完毕后，用户拿走打印结果，并卸下纸带。

手工操作阶段的特点如下。

- (1) 程序设计直接编制二进制目标程序。
- (2) 输入输出设备主要是纸带和卡片。

(3) 程序员自己上机操作, 程序的启动与结果处理都以手工方式进行。

(4) 单用户方式, 用户独占 CPU 和系统所有资源。

(5) CPU 和系统资源等待人工操作, 利用率很低。

可见, 手工操作方式严重影响计算机资源利用率, 即所谓的人机矛盾。而且随着计算机硬件和体系结构的发展, 人机矛盾日趋严重。

### 1.2.3 单道批处理系统

进入晶体管计算机时代后, 计算机硬件的性能和可靠性有较大提高, FORTRAN 等高级语言应运而生, 程序设计人员、编程人员、操作人员和维护人员间有了明确分工。这些进展都使得人们希望计算机能高效地做一些更复杂更有用的工作。而慢速的人工操作与快速的计算机处理能力之间越来越不协调, 例如, 有一道程序通过 3 分钟的安装等手工操作后, 在运算速度为 1 万次/秒的计算机上用 1 小时得到了结果。这时手工操作与程序运行时间之比为 1:20。把这道程序拿到此时速度为 60 万次/秒的机器上运行, 它只需花费 CPU 时间 1 分钟即可得到结果。如果在这种高速的机器上仍然坚持手工操作, 那么这时手工操作与程序运行时间之比为 3:1。这种比例是难以让人接受的。软件设计人员首先想到的是, 应该减少程序间转换的时间, 于是提出了“让计算机自动控制用户作业的运行, 废除上、下机手工交接”的要求。于是就出现了单道批处理系统监控程序, 它的功能是从磁盘上读入第一个作业的作业说明书, 按照它的规定控制该作业执行。这个作业运行结束后, 它又从磁盘上读入第二个作业的作业说明书, 继而执行之。这一过程一直进行到提交给系统的一批作业全部执行完毕时为止, 如图 1-3 所示。由于这种系统一次集中处理一批用户作业, 故被称为“批处理系统”, 该监控程序就是现今操作系统的雏形。

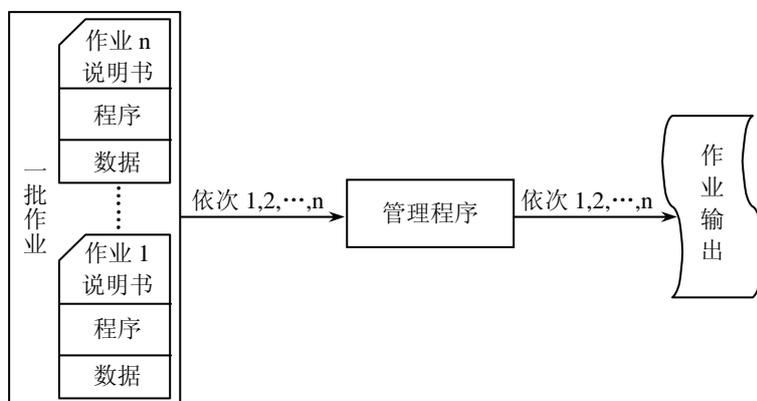


图 1-3 单道批处理系统示意图

在单道批处理阶段, 使用计算机的过程是: 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、调试各作业的运行并输出结果。正因为一批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。

单道批处理阶段的特点如上。

(1) 自动性。有专职操作员, 程序员不在现场, 作业(步)间不需要人的干预。

(2) 监控程序常驻内存。开机后第一个进入内存, 直到关机一直驻留在内存中。

(3) 单道性。一个作业调入内存后，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。

(4) 顺序性。由于是单道批处理，各道作业完成的顺序与它们进入主存的顺序一致。

(5) 监控程序只为一个计算机系统设计，通用性差。

此时，作业总是一个一个地顺序运行，当一个作业运行时它便独占了全机的所有资源。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。而且监控程序的通用性差。

#### 1.2.4 多道批处理系统

1964 年以后，计算机进入了集成电路和大规模集成电路时代。这时，硬件才有了长足的发展，中断和通道技术的出现，为输入/输出设备和 CPU 并行操作奠定了物质基础。另外，随着计算机应用的日益广泛，也要求进一步发展和扩大管理程序的功能，希望它能够最大限度地挖掘计算机系统本身的潜在能力。这时，人们开始把 CPU、存储器、外部设备以及各种软件都视为计算机系统的“资源”，提出不仅要合理地使用这些资源，而且要高效地使用这些资源。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，这样便相应地出了多道技术。多道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 I/O 设备高度并行工作，使 CPU 和 I/O 设备利用率显著提高，同时也提高了系统的吞吐量。系统吞吐量是指系统在单位时间完成的总工作量。可以用图 1-4 来说明。

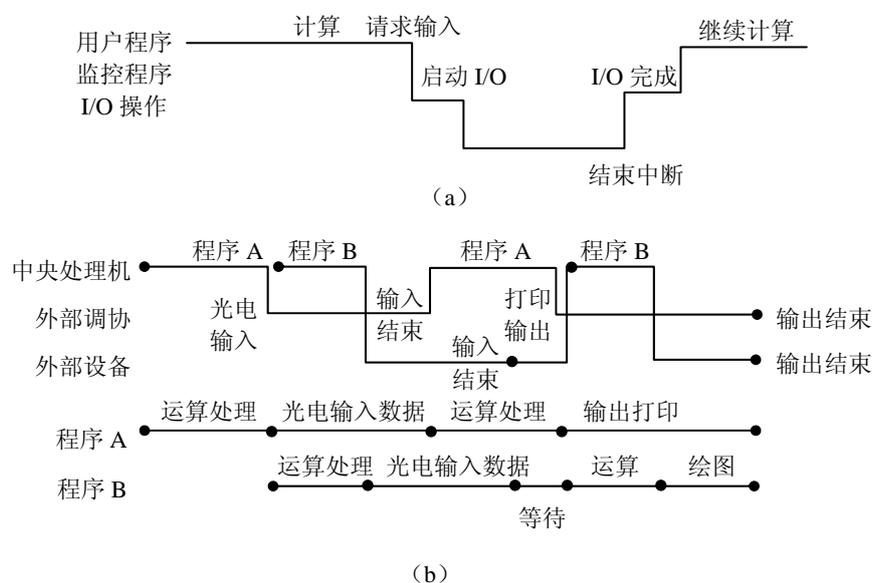


图 1-4 单道批处理和ado道批处理工作比较

(a) 单道程序工作示例；(b) 多道程序工作示例

从图 1-4 中可以看到，CPU 和 I/O 设备利用率都有了很大提高，CPU 和 I/O 设备能并行工

作，而且系统吞吐量也提高了。

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。将多道技术引入到批处理系统，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。可以说多道批处理操作系统是操作系统发展史上的革命性变革。

在多道批处理系统中工作流程大致如下：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，它们不断地轮流占用 CPU 来执行，并同时使用各自所需的外部设备，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，系统就使内存中另一作业占有 CPU，投入运行；若某个作业运行完毕，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，形成一个不断的作业流。这样的系统，系统资源利用率高，作业吞吐量大。

多道批处理操作系统有如下特点。

(1) 多路性。每次允许多个用户程序进入内存，它们交替地使用 CPU。某一段时间上，各道程序不同程度地向前推进。

(2) 共享性。整个系统资源被进入内存的多个程序共享使用。

(3) 无序性。多个作业完成的顺序与它们进入内存的顺序可能不一致。

(4) 封闭性。在一批作业处理过程中，用户不得干预系统的工作。即便是某个程序执行中出现一个很小的错误，也只能等到这一批作业全部处理完毕后，才能进行修改。

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。即使配置多道技术，此时操作系统的本质仍然是批处理。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。

### 1.2.5 分时操作系统

为了实现用户和计算机系统的及时交互、为了使多个用户共享主机、为了方便用户上机，将多道程序设计技术与分时技术结合在一起，就产生了分时操作系统。配有分时操作系统的计算机系统称为分时系统。如果说推动多道批处理系统形成和发展的动力是提高资源利用率和系统吞吐量，那么推动分时操作系统形成和发展的主要动力是方便用户的需要。

所谓分时系统，即一台计算机与多个终端设备连接，最简单的终端可以由一个显示器和一个键盘组成。每个用户通过终端向系统发出命令，请求系统为其完成某项工作。系统根据用户的请求完成指定的任务，并把执行结果返回。这样用户可以根据运行结果，再次通过终端向系统提出下一步请求。重复这种交互会话过程，直至每个用户实现自己的预定目标。计算机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。图 1-5 为分时系统工作过程的示意图。

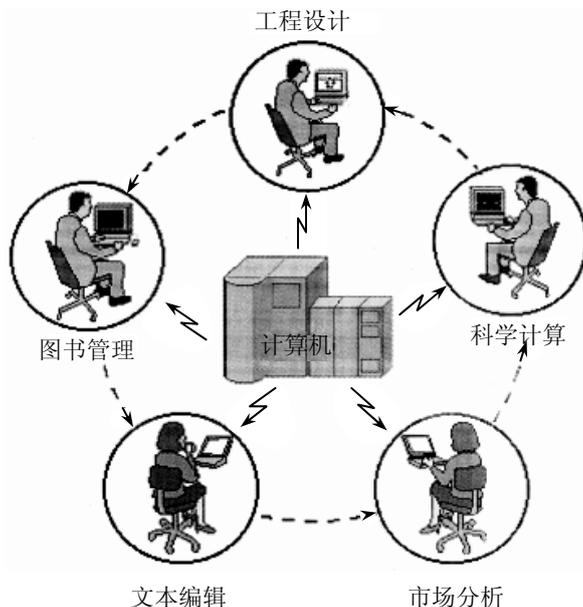


图 1-5 分时系统工作过程示意图

分时的概念并不陌生。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件、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。也就是说，分时系统之所以能在较短的时间内响应用户的请求，同时为多个终端用户提供服务，主要是因为分时操作系统中采用了“时间片轮转”的处理机调度策略。这种调度策略是把处理机时间划分成一个个很短的“时间片”，对提出请求的每个联机用户终端，系统轮流分配一个时间片给其使用。若在一个时间片内，用户所请求的工作未能全部做完，就会被暂时中断执行，等待下一轮循环再继续做，让出的 CPU 被分配给另一个终端使用。由于计算机的处理速度很快，只要时间片的间隔取得适当，那么用户就不会感觉到从一个时间片跨越到另一个时间片之间的“停顿”，就好像整个系统全由他“独占”使用似的。例如，若时间片为 100ms，系统中有 10 个用户终端分享 CPU，并假定忽略操作系统为实现用户终端之间的切换所需耗费的时间，那么每个用户平均响应时间（即从用完一个时间片到获得下一个时间片所需的时间间隔）为 1s。这 1s 的“停顿”，用户是完全感觉不出来的。

不难看出，分时系统的基本特征如下。

(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的现象。

(3) 及时性。系统可对用户的输入及时作出响应，及时是指响应时间在人能接受时间范围内。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，并在可能的条件下尽量提高系统资源的利用率。目前广为流行的 UNIX 操作系统和 Windows 操作系统都是分时操作系统。

### 1.2.6 实时系统

计算机应用范围日益扩大，打破了只把计算机用于科学计算和数据处理等方面的格局。比如用计算机系统来控制飞机飞行、导弹发射以及冶炼轧钢等生产过程，用计算机系统实现飞机订票、银行业务等实时信息处理系统。而多道批处理系统和分时系统都不能满足它们的要求，因而出现了实时操作系统。

所谓“实时”，是指能够及时响应随机发生的外部事件、并对事件做出快速处理的一种能力。而“外部事件”，是指与计算机相连接的设备向计算机发出的各种服务请求。实时操作系统可以说成是能对来自外部的请求和信号在限定的时间范围内做出及时响应的那种操作系统。

实时操作系统主要应用于下面两类场合。

(1) 实时控制。把计算机用于生产过程的控制，形成以计算机系统为中心的控制系统。控制系统能随时采集数据，并对得到的数据进行分析 and 处理，进而控制相应的执行机构。如用计算机系统控制飞机飞行、导弹发射等，要求计算机能尽快处理测量系统测得的数据，及时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。再如，用计算机系统用于轧钢、石化、机械加工等工业生产过程时，也要求计算机能及时处理由各类传感器送来的数据，然后控制响应的部件或机构。

(2) 实时信息处理。在一台或多台主机和多个终端通过通信线路连接起来的网络中，主机接受远程终端发来的服务请求，对信息进行检索和处理，并及时给用户正确的回答。如把计算机用于预订飞机票、查询有关航班、航线、票价等事宜，或把计算机用于银行系统、情报检索系统等，这些都是典型的实时信息处理系统。

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、高安全性。它对系统资源利用率要求并不高，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如下不同。

(1) 通用性。无论批处理系统还是分时系统，都是属于通用操作系统，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) 及时性。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的，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的。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、时钟管理机制和快速的任务切换机制。

(3) 交互性。对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在的程序。当然实时系统的交互性比批处理系统强的多。

(4) 高可靠性。高可靠性对实时系统至关重要，因为一点点差错都可能造成不可估量的损失。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。

最后要说明一点，批处理系统、分时系统、实时系统是 3 种基本的操作系统类型。一个实际的操作系统可能兼有 3 者或其中两者的功能。如 VAX-11 系列机上配置的 VMS 操作系统，就兼有批处理、分时、实时系统的功能。

## 1.3 操作系统的分类

从不同的角度出发，可以对操作系统进行不同的分类。从历史发展的观点，操作系统可以分为传统操作系统和现代操作系统；从操作系统管理的硬件规模出发，可分为微机操作系统、中小型机操作系统和大型机操作系统；按照资源共享的级别，可分为单任务操作系统、多任务操作系统、单用户操作系统、多用户操作系统、单道操作系统、多道操作系统、多处理机操作系统和网络操作系统。按所允许的交互类型可分为批处理操作系统、分时操作系统和实时操作系统。本节介绍几种常用的操作系统类型。

### 1.3.1 多处理机操作系统

在 20 世纪 70 年代，出现了多处理机系统，它试图从计算机体系结构上改善计算机系统的性能。引入多处理机后，使计算机系统的性能有了较大幅度的提高，如，系统吞吐量成倍提高，因为多台处理机可以协调工作，完成更多的任务；系统可靠性也有很大提高，因为在多处理机系统中通常具有系统重构功能，当一台处理机发生故障时，多处理机操作系统能立即将该处理机上的任务迁移到其他处理机上运行，保证整个系统仍然正常运行。另外，在相同处理能力下，多处理机系统比多台单处理机系统构成的集合要节省资源。如 20 世纪 90 年代 IBM 公司开发的 IBM OS/390 大型机操作系统就是多处理机操作系统。

多处理机操作系统可以按下面两种模式来配置：

(1) 非对称多处理模式，即主从模式。把处理机分为主处理机和从处理机两类。主处理机只有一个，从处理机可有多个。主处理机上配置操作系统，管理整个系统的资源，并为从处理机分配任务。从处理机只能按照主处理机的管理去工作。这种操作系统比较容易管理，但主处理机构成系统的瓶颈。

(2) 对称多处理模式。所有处理机都相同，操作系统可根据需要运行于其中的任一处理机上。这种操作系统允许多个进程同时运行，系统资源利用率较高，但比较难控制。

### 1.3.2 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。在网络范围内，用于管理网络通信和共享资源，协调各计算机上任务的运行，并向用户提供统一的、有效方便的网络接口的程序集合，就称为网络操作系统。要注意的是，在网络中各独立计算机仍有自己的操作系统，由它管理着自身的资源。只有在它们要进行相互间的信息传递、要使用网络中的可共享资源时，才会涉及到网络操作系统。

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、网络管理和提供各种网络服务。

依据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百千米到几千千米，甚至上万千米，可以覆盖一个地区、一个国家，甚至跨洲。局域网的地理范围是几千米到几十千米，一个企业或一个大学内都可组建局域网。广域网和局域网是一个相对概念。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 3 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。

### 1.3.3 分布式操作系统

在以往的系统，其处理和控制在高度集中在一台主机上，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。所谓分布式系统，是指由多个分散的处理单元经网络的连接而形成的系统。在分布式处理系统中，系统的处理和控制在分散在系统中的各个处理单元上。系统中的所有任务可以动态地分配到各个处理单元中去。在分布式系统上配置的操作系统称为分布式操作系统。

分布式操作系统和网络操作系统有许多相似之处，但又各自有自己的特点，主要从以下几方面加以比较。

(1) 分布性。在分布式系统中只有一个分布式操作系统，它均匀地分布在系统的各个节点上，操作系统的处理和功能是分布的；而在网络系统中各个节点可以有不同的网络操作系统，控制各自的资源，而网络中共享资源的控制功能只集中在网络服务器中，控制是集中式的。

(2) 并行性。分布式系统中可以将一个或多个任务动态分配到不同的单元上；而在网络系统中每个用户的一个或多个任务通常都在本地处理，网络操作系统通常无任务分配能力。

(3) 透明性。分布式系统具有透明性，如用户要访问某文件，他只需要知道文件名，无需知道文件在哪个站点。对网络系统来说，用户要访问一个文件时要知道文件名及位置。

(4) 共享性。在分布式系统中，各站点的所有资源都可供全系统共享，在网络系统中一

般仅有服务器上的部分资源可供全网共享。

(5) 健壮性。分布式系统有健壮性，当某站点出现故障时，正在该站点处理的任务能被自动迁移到好的站点。而在网络系统中，服务器是系统的瓶颈，若服务器出现故障往往会导致全网瘫痪。

### 1.3.4 嵌入式操作系统

嵌入式操作系统是指嵌入在各种设备、装置或系统中，控制和协调系统各种部件操作，完成特定功能的系统软件。这些设备、装置或系统是更大的设备、装置或系统中的一部分，这个大设备、装置或系统可以不是计算机。嵌入式操作系统通常工作在反应快或对处理时间有较严格要求的环境中，如电话交换设备、移动通信设备、印刷机、机顶盒、零售设备、智能卡等。

VxWorks 是当今嵌入式操作系统的代表。VxWorks 支持各种工业标准，包括 POSIX、ANSI C 和 TCP/IP 网络协议。VxWorks 运行系统的核心是一个高效率的微内核，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。VxWorks 可运行在 Windows 9x、Windows NT、Sun Solaris、Sun OS、HP-UX 等多种系统上。在“极地登陆者”号、“深空二号”和火星气候轨道器等登陆火星探测器上，就采用了 VxWorks 操作系统。VxWorks 负责火星探测器全部飞行控制，包括飞行纠正、载体自旋和降落时的高度控制等，而且还负责数据收集和与地球的通信工作。VxWorks 目前在国内也占据嵌入式开发系统市场的主要份额。

### 1.3.5 通用操作系统

前面介绍的批处理系统、分时系统和实时系统是操作系统的 3 种基本类型。如果一个系统兼有批处理、分时处理和实时处理三者，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分时用户较少时，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，而需要及时响应的用户作业作为前台。前台和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。如 UNIX、Microsoft Windows NT、Windows 95/98 等，都是通用操作系统。

衡量一个通用操作系统性能时，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，但这也是衡量操作系统性能优劣的一个重要方面。

### 1.3.6 微机操作系统

随着大规模集成电路发展,个人计算机时代到来了,各种类型的个人计算机和软件层出不穷,最早出现的微机操作系统,是8位微机上的CP/M。后来出现16位微机操作系统——MS-DOS。当微机发展到32位时,又出现了32位微机操作系统,如Windows和UNIX等。

#### 1.3.6.1 CP/M

CP/M是单用户单任务8位微机操作系统,是Digital Research公司1973年推出的。CP/M配置在Intel 8080、8085、Z80为主的芯片上,可支持软盘系统。1979年又推出CP/M 2.2版,到了1981年,CP/M操作系统成为世界上流行最广的8位操作系统之一。CP/M操作系统有较好的层次结构。它的BIOS把操作系统的其他模块与硬件配置分隔开,所以它的可移植性好,具有较好的可适应性和易学易用性。

#### 1.3.6.2 MS-DOS 操作系统

MS-DOS属于单用户单任务16位微机操作系统。它是美国微软公司的产品,主设计人是Tim Paterson。在1981年,IBM公司与Microsoft公司签定协议,使用MS-DOS作为IBM个人计算机的操作系统。MS-DOS操作系统对CP/M进行了较大扩展,增加了许多内部命令和外部命令,而且MS-DOS有优良的文件系统。又因为MS-DOS可广泛地配置在IBM-PC及其兼容机上,因此非常畅销,实际上它成为16位微机单用户单任务操作系统的标准。但MS-DOS操作系统受到Intel x86体系结构的限制,并且缺乏以硬件为基础的存储保护机制。从1981的MS-DOS 1.0版到1998年在Windows 95/98之下的7.0版,MS-DOS历经了16个年头。迄今仍有MS-DOS爱好者继续开发各种DOS软件产品。

#### 1.3.6.3 Windows 操作系统

MS-DOS提供的是一种以字符为基础的用户界面,不了解硬件和操作系统的话,便难以称心如意地使用PC。人们企望能把PC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专用的Windows/386。随后,1990年5月推出Windows 3.0版,1992年3月推出了Windows 3.1版,1995年推出了Windows 95版,1998年推出了Windows 98版及Windows NT/2000版。引起计算机界的强烈反响,使得计算机(特别是PC)的操作方法和软件开发过程产生了革命性的变化。Windows在PC系列微型机操作系统中占主导地位。

Windows 9.x是一个图形窗口操作环境软件。它代替了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) 丰富多采的图形操作界面,使PC易于掌握和使用。
- (2) 支持多任务运行,各任务之间即易于转换,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。
- (4) 提供了程序管理、文件管理、打印管理、控制面板等方便的管理工具,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、画图软件、终端通信软件、多媒体、网络以及时钟、日历、计算器、卡片、便笺、记录器等一套桌面办公用具软件。

Windows 9.x 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, 为确保持续器的连续, 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 本身占据的内存也不少, 因此, 若想同时运行几个应用程序, 并且可对较大的文本文件进行操作, Windows 必须突破原内存只有 640KB 的限制, 而能够访问高档 PC 所提供的扩充内存和延伸内存。

Windows 9.x 具备的多任务能力, 基本上是协作式的, 它使用协作方式为 Windows 应用程序获得较高效率和较佳的多任务控制。Windows 具有动态数据交换(DDE)和动态链接库(DLL)功能。DDE 是 Windows 应用程序之间彼此通信的一种手段, 以服务器客户方式进行通信。DLL 是调用时加载到内存的例行程序的库文件, 它可作为应用程序的一部分来使用, 改写 DLL 能扩充应用程序的功能, 也能利用其他程序的 DLL。用 DDE 和 DLL, 可实现多个 Windows 应用程序的协作运行、联合动作。

#### 1.3.6.4 Windows NT/2000

在 1993 年, Windows NT 投放市场。Windows NT 完全从头编写并采用了与 UNIX 相类似的微内核类型结构, 借此 Windows NT 实现了多任务。另外 Windows NT 还利用了 Carnage Mellon 的 MACH 操作系统的一个流行概念——线程, 以支持对称多处理。

在微内核基础上, 又加入大量特征和服务, 其中包括集成安全子系统、抽象的虚拟化硬件接口、强大的多协议网络支持、容错性和集成化 GUI 管理工具等。当 Windows NT 最终发行时, Microsoft 拿出了两个不同版本: Windows NT 3.1 和 Windows NT Advanced Server 3.1, 它们在取得巨大成功的同时也暴露了许多不足, 其中包括与已有 Windows 程序兼容性差, 对硬件要求过高。后者使它在 Novell 的 NetWare 占优势的网络操作系统市场中未能有惊人之举。

1994 年秋, Windows NT 3.5 发行, 它作了大量重要的改动与增强, 如利用点到点协议(PPP)的多协议远程访问服务, 简化内存脚本, 扩展错误修复, 重新编写 TCP/IP 堆栈等。另外 Microsoft 还将它们更名为 Windows NT Workstation 和 Windows NT Server。同时, 两者的内部结构进一步优化, 这更有助于准确地定义各自充当的角色。

Windows NT 3.5 的发行标志着 Microsoft 正式进入服务器市场, 甚至有人开始预测 Windows NT Server 将向 Novell 支配的网络市场发起挑战。

Windows NT 4.0 标明 Windows NT 在 Microsoft 的努力下达到了外界对它要求的标准。Windows NT Server 4.0 不仅仅引入了 Windows 95 的用户界面, 还包括其他许多特征, 如 NetWork OLE、Internet Information Server(IIS)2.0、RAS 多联接和 RAS 自动拨号、点到点隧道协议(PPTP)、DNS 和 WINS 的全面集成、集成化多协议路由器、支持扩展驱动器、性能的改进等。

Windows 2000 是 Windows NT 的升级版本, 包括 Windows 2000 Professional 和 Windows 2000 Server 两个版本。其 Server 版本继承了 Windows NT 的所有服务器管理特性, 增加并改进了一些网络服务, 因而成为当前最流行的网络操作系统之一。

Windows NT/2000 是一种 32 位网络操作系统, 是面向分布式图形应用程序的完整的平台系统。Windows NT/2000 具有工作站和小型网络操作系统具有的所有功能。主要包括: 文件及

文件管理系统、具有优先级的多任务/多线程环境、支持对称的多机处理系统、拥有兼容于分布计算的环境。由于 Windows NT/2000 具有良好的用户界面，并且操作方便，所以世界各著名软硬件生产厂商都宣布支持 Windows NT/2000。

Windows NT/2000 的主要特点如下。

(1) 采用 32 位操作系统。Windows NT/2000 是用 32 位寻址方式访问对象，可使 NT 寻址到多达 4GB 的内存。另外，Windows NT/2000 引入一种新的文件系统——NTFS，使可访问的磁盘容量大大增加。

(2) 不再需要 DOS 操作系统。Windows NT/2000 最伟大的成就恐怕是对 DOS 的完全摒弃！虽然在 Windows NT/2000 操作系统中没有再保留 DOS 代码，每件事都是通过仿真标准 DOS 调用来完成的，但绝大多数 DOS 程序依旧能在 Windows NT/2000 上运行，而且比在标准的 DOS 下运行还要快。

(3) 客户/服务器操作系统。Windows NT/2000 操作系统是按照客户/服务器方式来设计的。整个操作系统分成许多自含系统单元，并相互传递定义良好的各种消息。

(4) 受保护内存模式。Windows NT/2000 引入的另外一个概念就是受保护内存模式。在 Windows NT/2000 的内存模式下所有进程都有它们各自的 32 位地址空间。4GB 的空间被分成两半，而应用程序能真正使用的是低端的 2GB 空间，高端的 2GB 空间则是用于系统其他部分的连接。利用这种方法，每个进程都会认为只有它在单独运行，从而不会使得某个进程有意或无意地去读或写不属于它本身的内存空间。这就带来两个积极效果，首先它避免了 Windows 3.x 中发生的 90% 的冲突；其次它为每个进程提供了安全保证。

(5) 抢先式多任务。Windows NT/2000 采用抢先式多任务。抢先式多任务最广为人知的是微内核设计。利用抢先式多任务方式，微内核始终保持了对系统的控制。它给应用程序分配时间片段使其运行，在指定时间结束之际，微内核抢先运行进程并将控制移交给下一进程。这种抢先式多任务方式避免了由于某个应用程序出错所造成的整个系统就会挂起。

(6) 可移植性。可移植性是 Windows NT/2000 的设计目标之一。Windows NT/2000 可移植性的关键部分是硬件抽象层 (HAL)，它在高层操作系统软件中隐藏了硬件的实际差别。Windows NT/2000 的可移植性使它不仅可在 Intel x86 微处理器上运行，也可在 RISC 芯片上运行，如 DEC Alpha AXP、MIPS R4400 和 Motorola Power PC。

(7) 可伸缩性。Windows NT/2000 的可伸缩性是指 NT 能在单独一个系统中充分利用多个处理器的工作的能力。NT 可伸缩性的关键是对称多处理 (SMP)，Windows NT/2000 Server 的 SMP 设计能让系统的处理器个数为 1~32 个不等。NT 能动态分配系统和应用程序线程，是它们在不同的微处理器上运行。

(8) 本地化。Microsoft 意识到了将产品投入国际市场的价值与重要性。Windows NT/2000 有下列国家的本地化版本——巴西、中国、丹麦、荷兰、芬兰、法国、德国、意大利、日本、韩国、挪威、葡萄牙、俄罗斯、西班牙和瑞典。在以上每个版本中，Microsoft 不仅努力保证某一特定语言中的通信，同时还采用标准习惯，使用正确的列表、日期、时间、数字和货币的格式输出。为支持本地化，Windows NT/2000 在存储和处理所有内部数据，显示符号时采用的不是 ASCII 而是使用 Unicode 标准。从而使得 Microsoft 在不同国家和地区迅速而轻松地实现 Windows NT/2000 的本地化，包括全面支持日本的片假名和中国的简、繁体汉字（注：Windows 95 也支持 Unicode。尽管 Windows 95 完全使用 Unicode 保存内部数据，但 Windows 95

依旧用到了大量非 Unicode 领域的替代品，以保证向后兼容)。

(9) 安全性。Windows NT/2000 符合美国国家安全机构 (NSA) 的 C2 级检测标准。强壮的安全模式适合于操作系统的各个级别，而不像在其他操作系统 (包括 Windows 95) 上提供的薄弱的安全性。在良好的安全模型下开发的 Windows NT/2000 保证了它能胜任大多数有安全性要求的场合。

(10) 容错性。Windows NT/2000 提供了多级系统容错能力，包括 NT 日志式的容错特征列表、可恢复文件系统 (NTFS)、磁盘镜像和磁盘奇偶检验条带化 (RAID1 和 RAID5)、磁盘扇区备用以及对不间断电源 (UPS) 的支持。使 Windows NT/2000 成为企业工作站和服务器产品。

最后说明的是，Windows NT/2000 既是一种操作系统，也是一种网络操作系统。在 Windows NT/2000 中同时集成操作系统和网络操作系统的强大功能。

#### 1.3.6.5 UNIX/XENIX 操作系统

UNIX 是一个多用户、多任务的分时操作系统。最早是由美国电话与电报公司 (AT&T) 贝尔实验室 (Bell Lab) 的 Ken Thompson 和 Dennis Ritchie 两人在 DEC 的 PDP-7 机上开始设计的。从 1962 年至今，它不断地发展、演变并被广泛地应用于小型机、超级小型机、大型机甚至超大型机。20 世纪 80 年代以来又凭借其性能的完善和可移植性，在微型机上也日益流行起来。UNIX 名扬计算机界，众多用户争先恐后地使用它。由于 UNIX 的巨大成功和它对计算机科学所作的贡献，两位主设计人曾获得国际计算机界的“诺贝尔奖”——ACM 的图灵奖。

UNIX 系统在不太长时间内取得很大成功的根本原因在于 UNIX 本身的性能和特点。正如图灵奖评选委员会对 UNIX 的评价中指出的那样：“UNIX 系统的成功在于它对一些关键思想所作的恰如其分的选择和精悍的实现。UNIX 系统关于程序设计的新的思想方法成了整整一代软件设计师的楷模。UNIX 的天才在于它为程序员提供了一种可以利用别人工作成果的机构”。

具体地说，UNIX 系统有如下特点。

(1) 短小精悍的内核与核外程序的有机结合。UNIX 系统在结构上分成两大层：内核和核外程序，即在 1.1 节所提到的操作系统层和系统实用程序层。内核包括进程管理、存储管理、设备管理和文件管理。UNIX 系统内核设计得非常精巧，合理的取舍使之提供了最基本的服务。核外程序充分利用内核的支持，向用户提供大量的服务，甚至终端命令解释程序也放在核外程序层。核外程序与普通的用户程序被一样看待，它们都作为文件被保存在文件系统当中。把常驻内存的内核与不必常驻内存的核外程序分开而又有机地结合，不仅使核心不庞大繁杂，便于使用和维护，也使 UNIX 用户能不断把一些优秀程序加到核外程序层中去，使 UNIX 系统便于扩充。

(2) 采用树形结构的文件系统。文件有普通文件、目录文件、特殊文件之分。一个文件系统保持有一个根目录，其下有若干文件和目录，每个目录下都可拥有若干文件和子目录，等等。这样的文件组织方式不仅便于文件进行分类和查找，而且容易实现文件的保护和保密。UNIX 系统还允许用户在自己的可装卸的文件存储设备上建立一个 (子) 文件系统，并把它连接到原有文件系统的某个末端节点上，从而成为一个子树。当用户不用它时，还可把此 (子) 文件系统卸下来。

(3) 把设备如同文件一样看待。系统中所配置的每一种设备，包括磁盘、磁带、终端、打印机、通信线路等等，UNIX 都有一个特殊文件与之一一对应。用户可使用普通的文件操作

手段。对设备进行 I/O 操作。例如, 用户可用文件复制命令, 把磁盘的某个文件内容复制到打印机这一特殊文件上, 从而由打印机输出这个文件的内容。特殊文件与普通文件在用户面前有相同的语法和语义, 使用相同的机制, 这即简化了系统设计又便于用户使用。

(4) UNIX 是一个真正的多用户、多任务操作系统。系统初启时, 引导程序把系统内核放入内存低地址。然后经过内部的初启程序为系统建立了进程 0 和进程 1。进程 0 是所有进程的祖先, 也是系统中惟一的核心态进程, 它负责把盘上准备运行的进程换入内存, 有时也把它称为交换进程。进程 1 负责为每个终端建立一个进程, 执行 Shell 解释程序。每个终端的 Shell 进程等待用户输入命令, 一旦有命令输入就要对其分析, 找到相应的命令执行文件, 并为之建立一个子进程来执行这个命令, 命令执行完, 相应的子进程即被撤销。用户还可以指定一个命令在后台运行, 同时在前台执行其他命令。

(5) UNIX 向用户提供了一个良好的界面。这包括两种界面: 一种是在用户在终端上通过使用命令和系统进行交互作用的界面, 称为用户界面; 另一种是面向用户程序的界面, 称为系统调用。

UNIX 系统的用户界面就是操作系统的外壳 Shell。Shell 既起着命令解释程序的作用, 同时又是一种程序设计语言, 具有许多高级语言所具备的复杂控制结构与变量运算功能。因此, 也可用来编写程序, 即所谓 Shell 编程。

所谓系统调用界面, 是指操作系统内核提供了一组诸如文件读写、I/O 设备操作、进程控制等功能子程序, 用户程序通过一些特殊的指令来调用这些子程序, 从而访问系统的各种硬件资源和软件资源并取得操作系统的服务。UNIX 不仅在汇编语言级, 而且也在 C 语言一级中提供了系统调用的手段, 给程序设计带来了很大方便。

(6) 良好的可移植性。与完全用汇编语言写成的 MS-DOS 不同, UNIX 系统的全部系统实用程序以及内核程序的 90% 都是用 C 语言编写的。由于 C 语言编译程序有良好的可移植性, 因此用 C 语言编制的 UNIX 系统也具有良好的可移植性。这不仅意味着 UNIX 系统易于从一种硬件系统移植到另一种硬件系统上, 而且在某一种硬件系统上开发的 UNIX 应用程序也易于移植到其他配置了 UNIX 的系统上去。这些正是 UNIX 系统得以普及和取得成功的重要原因之一。

UNIX 系统的各种版本比较多。从 1970 年至 1978 年, 不断改进而推出新的版本。从 1981 年 AT&T 发表 UNIX System III (S3) 开始, UNIX 不再采用版本 (Version) 号排列, 而改为按系统 (System) 号排列。最近的版本是 1989 年推出的 UNIX System v 的 4.0 版。另外一个系列是美国加利福尼亚大学伯克利分校开发的带有虚拟存储功能的 UNIX 系统, 它们是 1.0 BSD, 2.0 BSD, ……直到 1983 年的 4.2 BSD 等系统。

进入 20 世纪 80 年代以来, UNIX 开始进入微型计算机市场。1980 年 Microsoft 公司在 UNIX V7 基础上, 根据微型机的特点对 UNIX 进行了修改和扩充。1983 年 Microsoft 公司又根据 UNIX 系统 III 相应地改写了 XENIX, 发表了 XENIX 3.10。1984 年随着 IBM PC/AT 机的推出, Microsoft 公司又发表了 PC/AT XENIX 1.0。1986 年 Microsoft 公司又根据 UNIX 系统 V, 发表了 PC/AT XENIX V。另一家软件公司也发表了 SCO XENIX 5.2/2。1987 年, AT&T 公司和 Intel 公司联合推出 UNIX 系统 V/386 3.0 版。与此同时, Microsoft 公司也发表了 XENIX V/386。后来, AT&T 公司与 Microsoft 公司又联合推出 UNIX 系统 V/386 3.2 版。

XENIX 与 UNIX 虽有一些差别, 但核内差别大而核外差别小。从用户使用的角度看, Shell

命令解释程序、基本命令和主要实用程序的用法几乎完全一样。XENIX 只是在微型计算机上运行的 UNIX，两者本质上没有什么不同。

在提供多任务处理、调度、内存管理和磁盘信息管理等许多方面，微机多任务操作系统都类似于传统的多任务操作系统。但是，微机的应用环境不同于大、中、小型计算机的应用环境，因此，它们的追求目标也不同，传统的多任务操作系统追求的是系统资源利用率和作业的吞吐量，而微机多任务操作系统追求的是充分提高对用户的响应能力及使用的方便性。

## 1.4 Windows Server 2003 简介

自 2003 年起，Microsoft 将其服务器产品统一在 Windows Server System 品牌下，这是一个全面、集成而且具有互操作能力的服务器基础结构，提供了一个公用体系结构：

- (1) IT 基础设施。用于部署和操作。
- (2) 应用程序基础设施。用于设计和开发。
- (3) 信息工作者基础设施。用于通信和协同。

Windows Server System 的基础是 Microsoft Windows Server 2003 操作系统，它为以上各方面提供了核心基础设施和公用的服务。该服务器平台提供了：

- (1) 底层的安全模型、目录服务以及操作和管理服务，以支持 IT 基础设施。
- (2) 作为应用程序基础设施之基础的核心应用程序定义和编程模型。
- (3) 支持信息工作者基础设施的核心数据和协同服务。

图 1-6 显示了 Windows Server System 的功能范围。

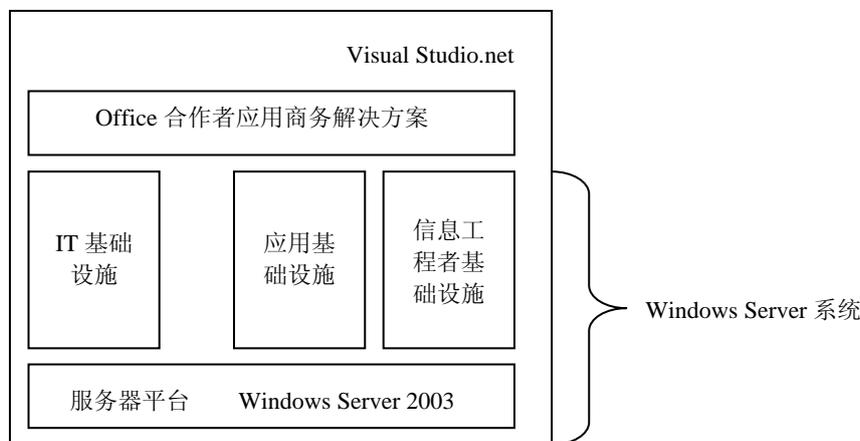


图 1-6 Windows Server 系统功能

Windows Server 2003 系列沿用了 Windows Server 2000 的先进技术，并且使服务器更易于部署、管理和使用。

### 1.4.1 服务器角色

Windows Server 2003 是一个多任务操作系统。它能够按照需要，以集中或分布的方式处

理各种服务器角色。其中的一些服务器角色包括：

- (1) 文件和打印服务器。
- (2) Web 服务器和 Web 应用程序服务器。
- (3) 邮件服务器。
- (4) 终端服务器。
- (5) 远程访问/虚拟专用网络 (VPN) 服务器。
- (6) 目录服务器、域名系统 (DNS)、动态主机配置协议 (DHCP) 服务器和 Windows 因特网命名服务 (WINS)。
- (7) 流媒体服务器。

#### 1.4.2 Windows Server 2003 的核心技术

Windows Server 2003 包含了基于 Windows Server 2000 构建的核心技术, 尽力实现具有可靠、可用、高效和联网功能的, 经济划算的优质服务器操作系统。

##### 1.4.2.1 可靠

Windows Server 2003 具有可靠性、可用性、可伸缩性和安全性, 这使其成为高度可靠的平台。

(1) 可靠性。Windows Server 2003 对系统可靠性的增强主要体现在如下方面: 把应用平台配备在可扩展操作系统的权利之上, 该应用平台包括了传统应用服务器的功能; 对涉及安全信息处理的体系结构进行一体化集成, 从而增强了事务信息的安全和访问控制。

(2) 可用性。Windows Server 2003 系列增强了集群支持, 从而提高了其可用性、可伸缩性和易管理性。Windows Server 2003 系列支持多达 8 个节点的服务器集群。如果集群中某个节点由于故障或者维护而不能使用, 另一节点会立即提供服务, 这一过程即为故障转移。Windows Server 2003 还支持网络负载平衡 (NLB), 它在集群中各个节点之间进行平衡传输通信。

(3) 可伸缩性。Windows Server 2003 系列通过由对称多处理技术 (SMP) 支持的向上扩展和由集群支持的向外扩展来提供可伸缩性。测试表明, 与 Windows Server 2000 相比, Windows Server 2003 在文件系统方面提供了更高的性能 (提高了 140%), 其他功能的性能也显著提高。Windows Server 2003 从单处理机解决方案一直扩展到 32 路系统, 它同时支持 32 位和 64 位处理机。

(4) 安全性。通过将内联网、外联网和因特网站点结合起来, 系统安全问题比以往任何时候都更为严峻。Windows Server 2003 在安全性方面提供了许多重要的新功能和改进, 包括:

- 公共语言运行库: 本软件引擎是 Windows Server 2003 的关键部分, 它提高了可靠性并有助于保证计算环境的安全。它降低了错误数量, 并减少了由常见的编程错误引起的安全漏洞。因此, 攻击者能够利用的弱点就更少了。公共语言运行库还验证应用程序是否可以无错误运行, 并检查适当的安全性权限以确保代码只执行适当的操作。
- Internet Information Services 6.0: 为了增强 Web 服务器的安全性, Internet Information Services (IIS) 6.0 在交付时的配置可获得最大安全性。IIS 6.0 为 Windows Server 2003 提供了最可靠、最高效、连接最通畅以及集成度最高的 Web 服务器解决方案, 该方案具有容错性、请求队列、应用程序状态监控、自动应用程序循环、高速缓存以及

其他更多功能。这些功能是 IIS 6.0 中许多新功能的一部分，它们使您得以在 Web 上安全地执行业务。

#### 1.4.2.2 高效

Windows Server 2003 在如下许多方面都具有可以提高工作效率的能力。

(1) 文件和打印服务。提供了智能的文件和打印服务，其性能和功能都得到提高。

(2) Active Directory。Active Directory 是 Windows Server 2003 系列的目录服务。它存储了有关网络上对象的信息，并且通过提供目录信息的逻辑分层组织，使管理员和用户易于找到该信息。

(3) 管理服务。通过自动化来减少日常维护是降低操作成本的关键。Windows Server 2003 新增了几套重要的自动管理工具来帮助实现自动部署。

(4) 存储服务。Windows Server 2003 在存储管理方面引入了新的增强功能，这使得管理及维护磁盘和卷、备份和恢复数据以及连接存储区域网络 (SAN) 更为简易和可靠。

(5) 终端服务。终端服务可以将基于 Windows 的应用程序或 Windows 桌面本身传送到几乎任何类型的计算设备上，包括那些不能运行 Windows 的设备。

#### 1.4.2.3 联网

Windows Server 2003 包含许多新功能和改进，以确保组织和用户保持连接状态。

(1) XML Web 服务。IIS 6.0 是 Windows Server 2003 系列的重要组件。

(2) 联网和通信。联网改进和新增功能扩展了网络结构的多功能性、可管理性和可靠性。

(3) Enterprise UDDI 服务。基于标准的解决方案使公司能够运行自己的内部 UDDI 服务，以供内联网和外联网使用。

(4) Windows 媒体服务。包括数字流媒体服务，还包括新版的 Windows 媒体播放器、Windows 媒体编辑器、音频/视频编码解码器以及 Windows 媒体软件开发工具包。

#### 1.4.2.4 XML Web 服务和 .NET

Microsoft .NET 已与 Windows Server 2003 系列紧密集成。分散、模块化的应用程序通过因特网互相连接并与其他大型应用程序相连接。.NET 提供了通过 XML Web 服务迅速可靠地构建、托管、部署和使用安全的联网解决方案的能力。这些 XML Web 服务提供了基于行业标准构建的可再次使用的组件，这些组件调用其他应用程序的功能，调用的方法独立于创建应用程序，操作系统、平台或设备用于访问它们的方法。利用 XML Web 服务开发人员可以在企业内部集成应用程序，并跨网络连接合作伙伴和客户。

### 1.4.3 产品系列比较

Windows Server 2003 系列能够满足从小公司到数据中心等不同规模组织的需要。

(1) Windows Server 2003 Datacenter 版用于实现最高级别的可伸缩性和可靠性。针对要求最高级别的可伸缩性、可用性和可靠性的企业而设计的 Windows Server 2003 Datacenter 版可以为数据库、企业资源规划软件、大容量实时事务处理以及服务器合并提供使命关键的解决方案。Datacenter 版可在最新硬件上使用，它同时有 32 位版本和 64 位版本，从而保证了灵活性和可伸缩性。它的优化功能使要求极为严格的应用程序和服务从中获益。它与 Windows Server 2003 企业版的主要区别是：支持更强大的多处理方式和更大的内存。

(2) Windows Server 2003 企业版用于最关键的服务器任务。针对大中型企业而设计的

Windows Server 2003 企业版是运行包括联网、消息传递、清单和顾客服务系统、数据库、电子商务 Web 站点以及文件和打印服务器等应用程序的服务器应该使用的操作系统。企业版可在最新硬件上使用，它同时有 32 位版本和 64 位版本，从而保证了灵活性和可伸缩性。各机构可从针对关系到业务的应用程序和服务而进行的优化结构中获益。

(3) Windows Server 2003 标准版用于部门任务和标准任务。Windows Server 2003 标准版是为小型企业单位和部门使用而设计的，它提供的功能包括智能文件和打印机共享、安全因特网连接、集中式的桌面应用程序部署以及连接职员，合作伙伴和顾客的 Web 解决方案等。Windows Server 2003 标准版提供了较高的可靠性、可伸缩性和安全性。

(4) Windows Server 2003 Web 版用于 Web 服务和托管服务。旨在生成并承载 Web 应用程序、Web 页和 XML Web 服务，Windows Server 2003 Web 版为因特网服务提供商 (ISP)、应用程序开发人员以及其他只想使用或部署特定 Web 功能的用户提供了一个单用途的解决方案。Windows Server 2003 Web 版利用了 Internet Information Services 6.0 (IIS 6.0)、Microsoft ASP.NET 以及 Microsoft .NET 框架的改进功能。

(5) Windows Small Business Server 2003 是一个用于小公司的服务器。

## 1.5 Windows Server 2003 模块结构

Windows Server 经过多年的发展已经从最初功能十分有限的 32 位操作系统演化成了现在的界面友好、管理方便、功能强大的操作系统家族，它之所以有这样大的发展，除了微软强大的商业推动力量之外，Windows Server 本身的技术因素也是重要的原因，尤其是它自身的体系结构所具有的可扩充性、可执行性、健壮性、兼容性和高效性，正是这些特性让 Windows Server 2003 不断得以进步。

### 1.5.1 Windows Server 2003 结构纵览

在讨论 Windows Server 2003 结构时，有两个词需要牢记：模块化和客户/服务器。

模块的意思是它内部分成许多小而独立的单元，每个单元的任务明确，定义清晰，模块化是计算机编程追求的目标之一，操作系统也不例外。模块化代码更易于维护，因为它意图清楚，并且整个代码段可被取代，而对要调用它的所有其他子程序不产生影响。

模块设计的概念与早期操作系统用到的分立设计方法相反。在分立设计中，操作系统运行在特许处理器模式下，代码块提供多个功能，但缺乏清晰分工，这样虽然可以取得更小、更紧凑的代码，但整个系统可移植性变差。

当人们听说 Windows Server 2003 是客户/服务器方式的操作系统时，他们通常认为这是指 Windows Server 2003 可用于客户/服务器数据库，或可用作网络系统。尽管 Windows Server 2003 是当作这种应用的明智选择，但从 Windows Server 2003 的结构来看，则并不确切，真正的含义是指 Windows Server 2003 内部的各部分是以客户/服务器方式来进行通信的。说的更为规范些，是指 Windows Server 2003 的模块成员是按照客户/服务器方式来设计开发的，如图 1-7 所示。当其中某一个小块代码有需求时，它是作为客户的；而能满足该需求的代码则是服务器。譬如要求在屏幕上画一幅图的用户程序就是个客户，它利用定义明确的消息向另一段代码发出请求（在本例中可能是指 Win32 子系统）要求画图，此时 Win32 子系统就是服务器——这

样构成了客户/服务器。

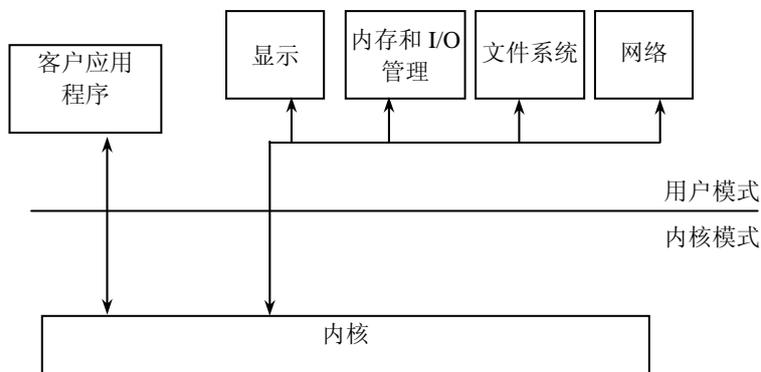


图 1-7 客户/服务器操作系统的设计方案

### 1.5.2 用户模式与内核模式的比较

Windows Server 2003 与当今先进的操作系统一样，采用将操作系统任务多级化的处理方案，这一点是和微处理器支持的实际模式相联系的。现在大多数微处理器支持的多重模式（有时称为环）程序就在其间运行。这些模式给运行在其中的程序不同级别的特许权访问硬件，并让其他一些程序在系统中运行。Windows Server 2003 采用特许模式和非特许模式，即通常所说的内核模式和用户模式。运行在内核模式下的程序可直接访问硬件和系统中的软件资源。在 Windows Server 2003 中只有操作系统的关键片段才允许在内核模式下运行，这样做是为了保证系统的安全性和可靠性。Windows Server 2003 的内核包括微内核、硬件抽象层及设备驱动程序，是 Windows Server 2003 中惟一能在微处理器特许内核模式下运行的部件。

所有不运行在内核模式下的程序都运行在用户模式下。大部分 Windows Server 2003 代码以用户模式运行，包括环境子系统（如 Win32 子系统和 POSIX 子系统）和所有用户应用程序。这些程序只能访问它们自己的 32 位地址和与系统其他部分的接口。这一切都是通过客户/服务器实现的，这将在以后介绍。

对于 Windows Server 2003，设计者尽可能地使操作系统运行在用户模式下，这样有助于确保系统可靠性和安全性。与此同时，当他们对基本部件修改时，也会变得简单。

### 1.5.3 Windows Server 2003 结构部件

为理解 Windows Server 2003 如何运行，以及为何如此运行，观察一下操作系统不同成分及其接口是十分重要的。如今已对 Windows Server 2003 有了一些了解，让我们在深入一层进行了解。图 1-8 显示了 Windows Server 2003 中各关键层及其逻辑联系。

Windows Server 2003 结构中的 4 个重要组成部分是：硬件抽象层（HAL）、微内核、Windows Server 2003 执行体、环境子系统。模块中的每一部分各司其职，Windows Server 2003 也就借此运行。

#### 1.5.3.1 硬件抽象层

硬件抽象层（HAL）是硬件与操作系统其余部分的一个软件接口，它通过动态链接库（DLL）实现，并负责屏蔽掉硬件特征对 Windows Server 2003 其余部分的影响，如中断控制

和 I/O 接口。这种抽象使 Windows Server 2003 更易于移植，因为此时操作系统其余部分不必再关心它所处的硬件平台是什么了。每个运行 Windows Server 2003 的硬件平台要有一特定的 HAL。这样设计的好处是，当 Windows Server 2003 移植到一种新型处理器结构上时，只需重新为该新型处理器编写 HAL，而其余部分不需进行重新编译。这使 Windows Server 2003 的移植极为容易。

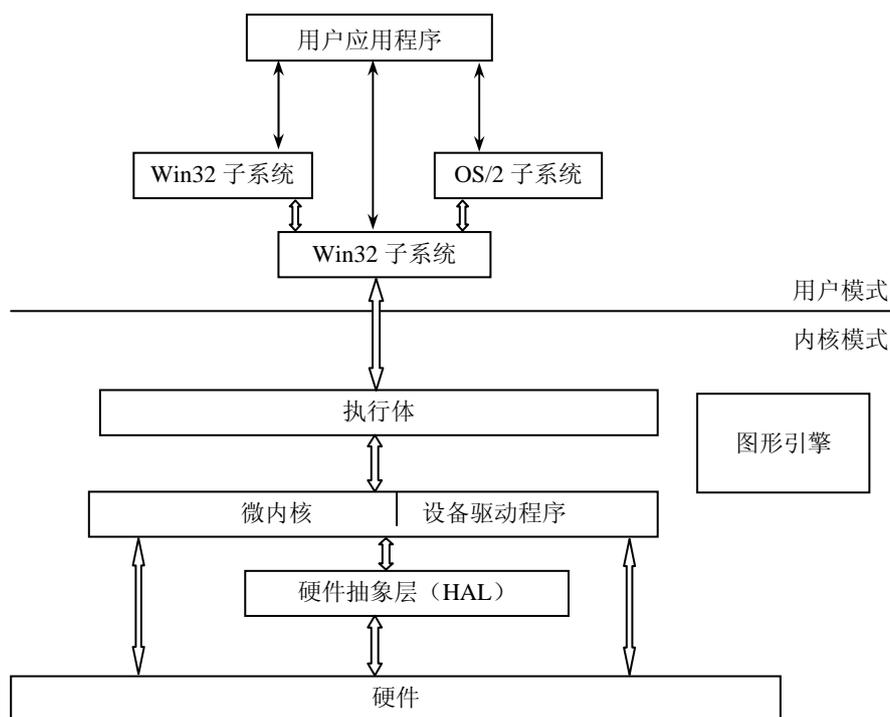


图 1-8 Windows Server 2003 中各关键层及其逻辑联系

HAL 还提供了对称多处理 (SMP) 接口。Windows Server 2003 为每种处理器结构 (Intel、MIPS、Power PC 和 Alpha) 都提供了两种 HAL，第 1 种 HAL 用于支持单个处理器；而第 2 种 HAL 用于支持多达 4 个处理器。硬件销售商还可提供其他的 HAL，实现对 Windows Server 2003 多达 32 个处理器的支持。

对计算机中每个实际的处理器，HAL 提供给微内核一个虚拟处理器，借此将处理器所有的特殊特征隐藏起来，从而使得各个处理器对操作系统来说并无区别。举例来说，如果有两套多处理器系统，其中之一是用 Intel Pentium 处理器；另一套是用 DEC Alpha AXP 处理器，虽然每套系统上的 HAL 可能不同，但 HAL 提供给微内核的虚拟处理器则是完全一样的。在 SMP 系统中，HAL 为每个物理处理器向微内核提供一个虚拟处理器，再由微内核将其等效为有 3 个处理器的 Intel Pentium 系统。

尽管 HAL 尽量减少硬件依赖性而提高 Windows Server 2003 的可移植性，事实上并不那么简单！Windows Server 2003 设计师们将其对硬件依赖性降至最低这一努力，在将 TVT 移植到新的平台所需的时间与精力已大为减少。（注：在 Windows Server 2003 最初开发阶段，所有初始代码是按 Intel 的 I860 RISC 芯片而写的。然而由于 Intel 对该芯片支持减小，及其开发过程

中遇到的设计问题而放弃了该芯片，而将开发重新转移到 MIPS 芯片。这也是 Windows Server 2003 操作系统良好移植性的有力证明)。

HAL 只能由 Server 2003 执行体直接访问，而用户程序永远无法直接进行调用。同时 HAL 也是 Windows Server 2003 系统中惟一被允许直接与硬件打交道的软件。这样做的好处是破坏性程序无法向硬件进行有意或无意的信息写入，从而避免了系统冲突；同时禁止直接从硬件读取信息，将有助于支持 Windows Server 2003 的安全模式。

尽管 Windows Server 2003 的目标是让所有的与硬件有关的调用都通过 HAL，但事实是少量的设备驱动程序和内核调用绕过了 HAL，而直接与硬件打交道。

HAL 模式最大的副作用是导致 Windows Server 2003 与原有 DOS 和 Windows 版本的程序不兼容，因为这些程序直接对硬件进行读写。但这种不兼容性只是 HAL 在为我们带来安全性和易移植性的同时所收取的微小代价罢了。

### 1.5.3.2 微内核

Windows Server 2003 微内核（简称内核）的权力无边。内核负责处理系统的所有操作，而且几乎是所有的系统功能都要经过内核。内核主要提供下列功能：线程安排和调度；陷阱处理和异常调度；中断处理和调度；多处理机同步；供执行体使用的基本内核对象（在某些情况下可以导出到用户态）。

Windows Server 2003 的内核始终运行在核心态，代码短小紧凑，可移植性也很好。一般来说，除了中断服务例程，正在运行的线程是不能抢先内核的。Windows Server 2003 的微内核的本质意思是执行某个功能时所必须的内核。（注：请勿混淆内核（或微内核）和内核模式。尽管它们是有联系的，但不是同一回事儿。内核是指各个独立的代码片段，这些代码片段组成操作系统的核心。内核模式是指处理器支持一种特许操作状态。在 Windows Server 2003 中，微内核运行在内核模式下）。

Windows Server 2003 中的微内核设计将许多操作分派给执行体，除了执行线程调度外，几乎将所有的策略制定都留给了执行体。在内核以外的系统组件，处理它们的资源分配、安全认证等都要执行体付出不可忽略的策略开销。为了降低这种开销，内核提供了一组严格定义的、可预测的、使操作系统得以工作的基础设施，这为执行体的高级组件提供了必需的低级功能接口。这一点充分体现了 Windows Server 2003 将策略与机制分离的设计思想。

内核通过一组被称作“内核对象”的简单对象帮助控制、处理并支持执行体对象的创建。大多数执行体级别的对象都封装了一个或多个内核对象。如，一个称作“控制对象”的内核对象集合为控制各种操作系统功能建立了语义。这个对象集合包括内核进程对象、异步过程调用对象、延迟过程调用对象和几个由 I/O 系统使用的对象（如中断对象）。再如，另一个称作“调度程序对象”的内核对象集合负责同步操作并影响线程调度。调度程序对象包括内核线程、互斥体、事件、内核事件对、信号量、定时器和可等待定时器。执行体使用内核函数创建内核对象的实例，使用它们构造更复杂的对象提供给用户态。

另外内核还处理处理器异常现象。当要求处理器进行某些不允许的操作，如向内存锁定区写入或被 0 除时，就会产生异常。Windows Server 2003 内核的最后一个用处是支持掉电恢复。如果 Windows Server 2003 系统装备了智能型不间断电源（UPS），则当电源掉电时，内核会作标记并按顺序切断系统，包括向 I/O 设备通知掉电消息和允许它们复位。（注：因为内核几乎囊括了 Windows Server 2003 系统上的所有操作，所以内核的关键部分是用汇编语言写的，

这保证了速度和效率。鉴于此，当 Windows Server 2003 移植到不同结构上时，内核优化就成了性能的关键因素。

在多处理器系统中，每个处理器上真正运行的是内核的备份，这些内核段能保持系统共享资源的相关性，而这些系统共享资源是所有进程上运行的线程。内核同时也处理物理设备（如 I/O 设备、处理器时钟或定时器等）产生的中断。正常情况下，当有系统中断产生时，内核会抢先运行处理该中断的线程。

### 1.5.3.3 Server 2003 执行体 (Server 2003 Executive)

若将 Windows Server 2003 内核比作总统，那么执行体就是他的执行部门 (Server 2003 执行体是执行部门的头，那么内核就是 Server 2003 执行体的头)。Server 2003 执行体负责处理与整个系统休戚相关的而内核由于太忙了而无暇直接指挥的重大任务。一个明确而又简练的定义是，Server 2003 执行体是所有其他程序在此系统上可以运行的操作系统基础。它包括的服务有：对象管理、虚拟内存管理、I/O 管理和处理器管理等。（记住：Windows Server 2003 内核实际上只是 Server 2003 执行体的一部分）。

Server 2003 执行体在内核模式下独享运行，并且当保护环境子系统需要服务时就调用它。用户应用程序并不直接调用 Server 2003 执行体片段，而是向诸如 Win32 和 POSIX 之类的子系统请求服务，再由这些环境子系统调用 Server 2003 执行体组件。Server 2003 执行体中有些功能并没有被现有的 API 开发利用，这是因为 Windows Server 2003 设计者们想在该操作系统中保留一些内容，为日后发展留下余地。除了内核本身以外，Server 2003 执行体还包括以下主要片段：对象管理器 (Object Manager)、进程管理器 (Process Manager)、虚拟内存管理器 (Virtual Memory Manager)、本地过程调用功能 (Local Procedure Call Facility)、安全参考监视器 (Security Reference Monitor)、I/O 管理器 (I/O Manager)。

### 1.5.3.4 保护下的环境子系统

在 Windows Server 2003 的设计目标中，有两个设计目标是个性化和兼容性的设计，这两点是通过保护下的环境子系统实现的。个性化是指 Windows Server 2003 提供多组应用程序接口 (API)，并且能有效地承担另一种操作系统的角色。Windows Server 2003 除了 Win32、Win16 和 DOS 个性外还包括了 POSIX 和 OS/2 个性。这是操作系统保持兼容的有效途径。Windows Server 2003 若仅能完全运行已有 DOS 和 Windows 软件的话，就不会有如此大的成功了。

在 Windows Server 2003 中有 3 种保护的环境子系统：Win32 子系统、POSIX 子系统、OS/2 子系统（注：尽管 Win16 和 DOS 也属于 Windows Server 2003 个性，但它们不属于保护下的环境子系统，因为实际上它们是 Win32 子系统的一部分）。保护下的子系统是用户级应用程序和 Server 2003 执行体的中间媒介。

记住我们所说的，Server 2003 执行体及其部件是在内核模式下运作的，而其他的任何程序基本上都是处于用户模式的，其中包括所有环境子系统，它们是完全工作在用户模式下的。当某应用程序要调用某环境子系统时，它是通过系统服务层传送到 Server 2003 执行体的。每个环境子系统保留了各自进程的磁道并与其他子系统分开独立工作，每个应用程序也只能在其设计时所处的相应子系统下运行。当 Windows Server 2003 中引导应用程序时，Windows Server 2003 察看文件映像头后，再决定该应用程序在哪种子系统中运行。

#### 1. Win32 子系统

Win32 是 Windows Server 2003 的基本子系统，该系统是以 Win32 系列的 API 为基础的，

它是在 Windows Server 2003 产品开发阶段所编写的，因此，这些 API 的许多内容是 Win16 中相应部分的直接扩展。（注：Win32 是 Win32 API 有关的调用中各种服务所用到的 API 及 Windows Server 2003 子系统的统称。在 Windows Server 2003 开始设计的一年半中，曾打算将 OS/2 Presentation Manager 作为其默认主子系统，然而随着 Windows 3.x 的成功，Microsoft 决定将 Windows 界面及其相关 API 作为基本个性）。

在先前讨论的客户/服务器模型中，Win32 子系统充当 Windows Server 2003 中其余环境子系统的服务器，而其余环境子系统担任客户角色，并将它们的 API 调用传送给适当的 Win32 API。Win32 API 由 Win32 子系统提供服务。

Win32 子系统负责所有的用户输入和输出，包括显示器、键盘和鼠标。当其他子系统（如 OS/2 或 POSIX）需要使用这些设备时，它们向 Win32 子系统发出请求，要求提供服务。最初设计 Win32 子系统时，Windows Server 2003 建设者们尽量让其和 Windows 3.x 尽可能类似地工作。结果，设计分成 5 大块：窗口管理器（称为 USER）、图形设备接口（GDI）、操作台、操作系统子程序和 Win32 图形设备驱动程序。

然而在 Windows Server 2003 中，对 Win32 的结构组织作了改动。图形设备接口（GDI）和窗口管理器（USER）是在 Win32 子系统内。然而，已经强调过，所有其他子系统是利用 Win32 API 处理用户输入和输出的，因此 Windows Server 2003 工作小组将 Win32 子系统中的 GDI 和 USER 这两部分移到了 Server 2003 执行体中。这样做减小了系统所有进程的开销，并充分发挥了使用这些代码的服务的功效。

这种转移对 Windows Server 2003 可靠性所产生的影响早就引起了人们的思考。许多人批评说这样做就会允许更多代码运行在内核模式下，而内核模式运行是可以访问系统资源的，从而使系统崩溃可能性加大。对原先模式而言，GDI 和 USER 都位于 Win32 子系统中，那么当它们出问题时，要用到 GDI 或 USER 服务的所有程序将不再能作响应，从而导致整个用户界面锁定。然而将它们移到 Server 2003 执行体后，内核能对它们进行监视。当它们不产生响应时，不再将系统锁定挂起，而是内核进行故障修复，并允许系统进行重引导。

在大多数情况下更希望是这样的结果，因为当基本服务出问题而无法工作时，允许 Windows Server 2003 内核重引导比完全无法使用以及系统锁定要好。两种模型的孰是孰非目前仍在争论中。但这种新模式究竟是否可靠，最终还是要靠时间的检验和对 Windows Server 2003 的测试而定。

## 2. MS-DOS 和 Win16

Windows Server 2003 成功的关键之一是大部分 Windows 3.x 和 DOS 应用程序依旧能在其上运行。在 Windows Server 2003 的最初开发阶段，设计小组和 Microsoft 管理机构对于这些程序是否应能在 Windows Server 2003 上运行看法曾有些分歧。Microsoft 管理机构承认，如果 Windows Server 2003 不向后兼容，用户若想对他们已有软件作升级则投资过大。结果只会导致 Windows Server 2003 价格惊人，因此决定支持 16 位的 Windows 程序和 DOS 应用程序。支持这些个性的决定很快就被 Windows Server 2003 设计轻松采纳。其中一些目标为：

- (1) 允许 DOS 程序不作改动就可直接运行。
- (2) 允许绝大部份 16 位的 Windows 应用程序不用改动就可直接运行。
- (3) 使系统和其他 32 位应用程序不受 16 位及 DOS 程序的干扰。
- (4) 允许在 RISC 平台上运行 16 位 Windows 应用程序和 DOS 程序。

(5) 提供 32 位和 16 位 Windows 程序的数据共享机制。

许多人认为 Windows 3.x 是操作系统,但从技术眼光来看,它不是真正的操作系统,它只是 DOS 下的一个用户界面罢了, DOS 才是真正的操作系统。

因此保证兼容性的第一步是创建一个 DOS 环境。Windows Server 2003 中的 DOS 环境称为虚拟 DOS (Virtual DOS Machine, VDM), 或称为 Server 2003 VDM。VDM 是全 32 位用户模式应用程序,通常向 Win32 子系统请求服务,偶尔也直接向 Windows Server 2003 系统服务层申请。VDM 是建立在 DOS 5.0 基础上的兼容,因此也就只能提供到 DOS 5.0 为止的兼容。

Windows Server 2003 允许运行任意的 DOS 应用程序,想运行多少就运行多少,每个应用程序在其各自的 VDM 中运行。因为 VDM 只是 Windows Server 2003 上的一般进程而已,所以只要还有其他进程同时存在,系统就是抢先式多任务的 DOS 程序。

VDM 另一显著特点是它提供给用户的空闲“常规”内存超过 620KB,令人惊奇的是,与此同时它还向 DOS 应用程序提供完全的鼠标、网络和 CD-ROM 的支持。与等同的 DOS 系统相比,加载了相同的这些程序后的 VDM 提供了更多的自由内存。

像 Windows 3.x 依赖于 DOS 提供的服务一样,Windows Server 2003 上的 Win16 子系统同时依赖于 Windows Server 2003 VDM。VDM 中的 16 位 Windows 仿真器称为 WOW,它是 Windows on Win32 的缩写。因为它存在于 VDM 内部,它向 VDM 请求服务时绝大多数采用的是和标准 Windows 3.x 向 DOS 请求服务一样的方法,VDM 允许这些调用的大部分直接转换成对 Win32 子系统的调用。

用一个 16 位 Windows 程序产生一个 Win16 API 调用时,WOW 子系统利用称为转换的进程将该调用转换成等效的 Win32 API 调用,然后,再将其传送给 Win32 子系统。同理,Win32 调用后需返回给 Win16 应用程序的数据也必须经过该转换处理。

当将 16 位数据格式向 32 位数据格式转换时,必须要有一系列的标准。从 16 位向 32 位转换十分简便,只须在高位加上 16 个 0 即可,然而从 32 位向 16 位转换可不能简单将高端 16 位截掉,这样会引起数据丢失。转换进程实际上是在 Win32 子系统内进行的。

因为 Windows 3.x 环境使用共享内存模式,所以许多程序编写时假定系统提供共享地址空间,有些甚至必须要求有共享地址空间。为保留对这种应用程序的兼容,所有 16 位 Windows 应用程序在单个 VDM 内运行。WOW 子系统不是多线程的,同时 16 位 Windows 应用程序相互之间是排斥性多任务方式,就跟真正的 Windows 系统一样。(注:每次进行一个 16 位 Windows 应用程序时就会产生一新线程,因此从结果来看 WOW 是多线程的。然而微内核对这些线程的调度方式和对系统上其他线程的调度方式是不一样的)。

通常 Windows Server 2003 内核根据优先级调度线程。当一正在运行的线程被抢占时,内核根据优先级将控制移交给下一线程。内核对 WOW 线程处理方式则不同,当它抢占当前运行的 WOW 线程后,仅按常规将控制移交给另一线程,而并不会安排执行其他 WOW 线程,除非当前 WOW 线程放弃控制权。

因此尽管 WOW 实际上是多线程的,但它又不能利用这带来的优势。这并不是设计的过错,而是为了保证现存 16 位软件兼容性的谨慎做法,以避免抢先式多任务与 16 位应用程序冲突导致的系统瘫痪。

还有,因为所有 WOW 中的应用程序在单个共享内存空间中运行,所以如果某个 16 位 Windows 应用程序出错了,它会导致所有在该 WOW 子系统上运行的 16 位 Windows 应用程序

出错失效。但是它无论如何也不会影响到系统或其他任何正在运行的 32 位应用程序。(注：记住区别——所有 DOS 程序在各自独立分开的 VDM 上运行，是抢先式多任务；而另一方面，16 位 Windows 应用程序是在单个 VDM 上运行，相互之间不仅存在合作式多任务冲突；同时也和系统其余部分存在抢先式多任务冲突)。因为 DOS 和 Windows 3.x 的应用程序严重依赖 Intel 汇编语言，所以在 Windows Server 2003 支持的 RISC 系统中能够毫无改动地运行这些程序是要有些技巧的。VDM 将所有调用分解成指令执行单元 (Instruction Execution Unit) (IEU)。对 Intel 结构，这些调用可以在处理器上直接执行，而在 RISC 系统中，这些调用要经 Intel 仿真程序转换。Intel 仿真程序由 Insignia Solutions 有限公司编写，他们同时也编写了 Macintosh 的 SoftWindows。

在 Windows Server 2003 以前的版本中，仿真代码是基于 286 级处理器的。当 16 位 Windows 应用程序要求运行在 386 增强模式下时，在 RISC 处理器的 WOW 仿真方式下运行时会出现问题。例如 Microsoft Office 之类的程序要求在 386 增强方式下运行，它就无法运行在 RISC 平台上。Windows Server 2003 的一个重要改进是仿真程序已经升级，与 486 结构系列完全兼容。

### 3. POSIX

Microsoft 公司在开发 Windows Server 2003 时对各种开放系统标准都十分关注，他们意识到支持开放系统是进军先进操作系统这一新市场并取得成功的途径。Windows Server 2003 经常说起其所支持的标准是 POSIX 兼容。POSIX (可移植操作系统界面) 是由 IEEE 开发的用于支持 UNIX 平台应用程序的可移植性，然而现今 POSIX 已经集成到许多非 UNIX 系统中去了。POSIX 兼容性分为许多级，从 POSIX.0 到 POSIX.12 不等。这些等级包括一系列的提议，并不是所有建议都已经作为标准。

Windows Server 2000 中的 POSIX 子系统是与 POSIX.1 兼容的。POSIX.1 仅支持最基本的由 Windows Server 提供的服务。当一个 POSIX 应用程序在 Windows 2000 上运行时，载入该 POSIX 子系统并将 C 语言 API 调用转化成 Win32 API 调用 (这些 C 语言 API 是实现 POSIX.1 支持所必须的)，转化生成的 Win32 API，再在 Win32 子系统中实现该服务。POSIX.1 要求的一些特性如下：

- (1) 区分大小写的文件名。这是由 NTFS 文件系统支持实现的。
- (2) 多文件名支持 (硬链接)，同样由 NTFS 支持。

(3) 文件系统横向检验，这由用户 Bypass Tra-Rerse Checking (忽略横向检验) 来控制。为实现某个特定用户的 POSIX 兼容，必须将该用户的此权限取消。

因为 POSIX.1 只定义了一组有限的服务，所以单独的 POSIX 子系统并不是一个完整的编程环境。并且因为在 Windows 2000 上应用程序不能在子系统之间混合调用，所以 POSIX 应用程序被严格限制在 POSIX.1 中定义的一组服务。这种限制意味着在 Windows 2000 上可执行的 POSIX 不能创建线程或窗口，也不能使用远程过程调用 (RPC) 或套接字，然而可以在 Win32 应用程序中做所有这些事情。

这一情况在新一代的操作系统中有所改变。对于 Windows Server 2003 中为什么要引入 POSIX 子系统，不同的人有不同的看法，有些人认为这样做对终端用户来说提高了应用程序的使用率，有些人认为这符合开放系统标准；也有人认为这是为了验证 Windows Server 2003 和 UNIX 平台的互操作性。尽管 POSIX 子系统实现得不错，但有人还是认为将 POSIX 子系统加入 Windows Server 2003 中的全部理由是为了符合政府购买标准。对那些对此不熟悉的人来说，会觉得政府机关要求某一特殊操作系统必须符合多开放系统标准是很奇怪的。例证之一是政府信息处理标准 (FIPS)。

只有加入了 POSIX 子系统后, Windows Server 2003 才可在该巨大市场内出售; 若不支持这一标准它就会被排斥在外。因此这只是 Windows Server 2003 运行市场渗透的策略, 用以避免人们用贸易标准将 Windows Server 2003 排除在外。Windows Server 2003 是一个优秀的操作系统, 轻易地包容 POSIX 子系统只是 Windows Server 2003 功能和灵活性的一个体现。尽管如此, 不要受到欺骗而去大力使用该 POSIX 子系统。Windows Server 2003 实际上并不包含 POSIX 子系统。为了提供对 POSIX 标准的支持, 微软公司提供了一个独立于操作系统的软件即 Windows Services for UNIX (SfU), 也称为 Interix, 它是原 POSIX 子系统的超集。

在 Interix 的计划中, POSIX/UNIX 子系统得到了相当大的改善, 支持了主流的 UNIX 标准和 UNIX 应用 (例如 X 以及标准 UNIX Shell)。它几乎成为和 Win32 子系统平行的部分。

#### 4. OS/2

Windows Server 2003 曾一度就被认为是 OS/2 的下一代产品。那样的话, OS/2 Presentation Manager 界面将是 Windows Server 2003 的主界面, 所有标准 OS/2 程序将能在其上运行, 其中包括基于字符的和基于 GDI 的程序。然而当决定以 Windows 界面作为 Windows Server 2003 主界面并将 Windows Server 2003 作为 Windows 平台的进一步发展以后, 强调支持 OS/2 的观点逐渐消失了。最终结果是 OS/2 子系统可以运行标准 OS/2 1.x 版本的字符模式应用程序或视频 I/O 应用程序, 不能运行 OS/2 2.x 的与图形有关的应用程序。(注: OS/2 子系统只能工作在 Intel 芯片基础上, 无法在 RISC 平台上运行)。

OS/2 设计成为受保护环境子系统, 与 POSIX 十分相仿。它将 OS/2 API 调用转换成 Win32 API 调用, 后者再由 Win32 提供服务。OS/2 子系统和每个 OS/2 应用程序各自在其 32 位受保护内存空间运行, 彼此间是抢先式多任务方式, 与系统上其他正在运行的程序之间也是抢先式多任务方式关系。除了 OS/2 的核心集外, OS/2 子系统还包括许多 LAN Manager API, 其中有 Net-BIOS、邮件槽和命名管道。从这方面来看它是和 POSIX 不一样的, 后者没有网络上的 API 支持。

## 1.6 Server 2003 执行体中的重要组件

### 1.6.1 对象管理器

Server 2003 执行体的对象管理器用于所有系统要用到的对象的产生、修改和删除, 这些对象组成了 Server 2003 执行体。对象是抽象的数据类型, 代表操作系统资源。Server 2003 执行体同时还向操作系统其余部分提供对象状态信息。

对象可以是具体的 (如设备端口); 也可以比较抽象 (如线程)。当对象创建时会被赋予一个名称, 其他程序通过该名称就能访问该对象了。当另外一个进程想访问该对象时, 它会向对象管理器申请对象句柄, 对象句柄提供一指针, 用于实际对象的定位, 同时它还提供对象进行访问和访问控制信息。访问控制信息由 Windows Server 2003 安全子系统提供。

对象管理器还通过维护不同对象类型的引用来确保单个对象没有占用过多的资源 (一般指系统内存)。另外, 对象管理器还负责清除没有所有者的单独的对象, 这就是众所周知的垃圾回收技术。Windows 3.x 中缺少类似的程序是导致麻烦的重要原因。在 Windows 3.x 中, 若某个程序崩溃了, 或不能正确处理系统资源时, 程序所用到的系统资源就会无法正确地返回到系统池中, 从而产生缺少该系统资源的出错消息。实际上这是所谓的内存泄漏。

### 1.6.2 进程管理器

进程管理器负责所有进程和线程的创建、删除和修改，同时还向系统其余部分提供进程和线程状态信息。（注：在进程的定义中，包括一虚拟地址空间，一个或多个线程，一段可执行程序代码和一系列系统资源；一个线程是属于单个进程的一个可执行对象，它包括一个程序计数器，它指向本进程可执行代码段的当前位置，两个堆栈和一系列寄存器值）。

和其他所有 Server 2003 执行体成员一样，进程管理器在整个操作系统中扮演一个重要的角色。当启动一个应用程序时，相应地就要创建一个进程，这需调用进程管理器；又因为每个进程至少需要包含一个线程，因此进程管理器被再次调用产生一线程。

虽然进程管理器管理线程，但它本身没有权力决定何时创建进程和线程，也无权对它们进行调度，这由微内核本身决定。

### 1.6.3 虚拟内存管理器

虚拟内存管理器用于系统虚拟内存池的管理。虚拟内存是用磁盘资源模拟实际的系统内存，也即将内存中那些不使用的页移出到磁盘中，而当需要使用这些页时，再将它们恢复到内存中。

VMM 是 Windows Server 2003 的一个主要部分，它给每个进程分配 32 位地址空间，而不管系统的实际内存是多大。每个进程分配 4GB 虚拟内存空间，其中高端的 2GB 保留给系统使用，低端的 2GB 则为进程使用。进程对内存寻址时，会假定其间只有它本身在使用。VMM 负责将进程的内存地址转换成实际系统内存地址。如果进程的内存地址指向的内存段位于磁盘页面，VMM 将从磁盘检索该页。

### 1.6.4 本地过程调用功能

本地过程调用功能是一个用于高速信息传输的进程间通信机构，在 Win32 API 下它是不可用的，它是对 Windows Server 2003 组件有效的内部机制。如远程过程调用使用 LPC 在同一个系统中的进程之间通信。典型地，LPC 是 Windows Server 2003 的客户/服务器设计中所必备的功能，它是运行在本地 Windows Server 2003 系统上所有客户与服务器进程的接口。

LPC 结构和远程过程调用（RPC）极为相似，只是前者只支持本地机上客户/服务器进程之间的通信。更确切地说，LPC 是这样的一种机制，它使不同进程的两个线程能相互交换信息。

请记住：Win32 子系统是用户模式应用程序，且运行在它自己的内存空间里。当某个程序想和 Win32 子系统通信请求服务时，它从适当的 DLL 文件中调用一固定子程序，这个固定子程序再利用 LPC 功能向 Win32 子系统进程传递请求，请求经其处理并执行要求的动作后，再通过 LPC 返回有关请求完成的消息。

### 1.6.5 安全参考监视器

安全参考监视器是 Windows Server 2003 所有安全的基础，同时它还负责强化本地计算机的所有安全措施。SPM 是通过注册进程及本地安全权限，运行其子程序共同来实现的。当用户注册时，Windows Server 2003 会对其身份进行确认，注册进程子系统会给用户申请一安全访问令牌（Security Access Token, SAT）。SAT 也包含用户特许权及组成员关系列表，是用户在注册过程中的钥匙。无论何时用户想干点什么，SAT 经递交审核后就可决定该用户是否能执行该动作。

这也是安全参考监视器和对象管理器紧密工作之处。每当某用户想访问某对象时，对象管理器创建一句柄处理对象访问，同时调用安全参考监视器以决定该句柄可授予的访问级别。安全参考监视器要用到用户访问令牌中有关信息，将这些信息与对象中的访问控制列表比较，判断该用户是否能被授予该对象的访问级别。通过这种方法，安全参考监视器控制了 Windows Server 2003 中所有对象访问的安全性。

### 1.6.6 I/O 管理器

I/O 管理器 (I/O Manager) 负责处理和协调系统所有的输入与输出。它对设备驱动程序、可安装文件系统、网络重定向和系统超高速缓存进行监督。

I/O 管理器 (I/O Manager) 不仅负责各种设备间的对话，还要使它们能和平相处。它取消了传统的独立设计 I/O 设备驱动程序的做法，而采用支持混合及匹配设备的分层设计方法，而支持混合及匹配设备又是必要的。

## 1.7 本章小结

本章从较高的层次概述了什么是操作系统，操作系统的发展、功能、作用、分类及 Windows Server 2003 的简介。在章中可以看到 Windows Server 2003 代表了下一代的操作系统，它的 32 位设计使它也成为下一代的网络操作系统。通过本章的介绍还可对 Windows Server 2003 的历史有了一定了解，从而了解到 Windows Server 2003 设计师们是如何尽心尽力地将其设计成真正的企业解决之道。

Windows Server 2003 是一模块化的、精心设计的操作系统。它的灵活结构能满足今天大多数的要求；它除了实现了最初的设计目标外，也为日后发展留下了余地。本章粗浅介绍了 Windows Server 2003 结构的各个组成部分，并讨论了 Windows Server 2003 结构所需的一些概念，包括硬件抽象层 (HAL)、Server 2003 内核、Server 2003 执行体和受保护环境子系统等。Windows Server 2003 也是因其结构而达到优秀性能及高可靠性，这又使 Windows Server 2003 成为服务器和网络操作系统的首选。

## 习 题

1. 计算机系统包括哪些部分？
2. 什么是计算机的操作系统？
3. 简述操作系统 5 大管理的主要功能。
4. 批处理操作系统、分时操作系统和实时操作系统各有什么特点？
5. 何为多任务？微机多任务操作系统的设计目标是什么？
6. 目前操作系统主要有哪几类？各类的代表是什么？
7. DOS, Windows, UNIX 三种操作系统中谁具备虚拟存储能力？谁具备窗口功能？谁支持多用户？
8. UNIX 的特色有哪几个方面？
9. Windows Server 2003 操作系统的特点是什么？
10. Windows Server 2003 操作系统的 4 个重要组成部分是什么？

## 实验1 Windows Server 2003 操作系统的使用

### 一、实验目的

1. 练习 Windows Server 2003 操作系统的常用操作。
2. 练习文件夹、文件的属性设置和共享设置。

### 二、实验准备知识

Windows Server 2003 是 Windows Server 的最新版本。它和 Windows Server 2000 有很多相像的地方。下面简要说明 Windows Server 2003 基本操作。

(1) 查找界面。依次选择“开始”|“搜索”|“文件或文件夹”命令之后，打开“搜索结果”对话框，在左侧的窗格内填写、设置搜索选项，右侧的窗格显示搜索的结果。

(2) 菜单内容。在桌面上打开“开始”菜单，它的内容同 Windows NT 4.0 相比已经发生了很大的变化。“管理工具”出现在“程序”菜单的最前列，在默认的情况下，“管理工具”是由“认证管理器”、“计算机控制台”、“事件查看器”、“索引服务器控制台”、“可移动存储控制台”等 5 部分内容组成的，并且绝大部分系统管理工具都被集中在“计算机控制台”内，可以看到该控制台所具有的强大功能，本地计算机或域内所有计算机的资源都在它的管理范围之内。

在 Windows Server 2003 桌面上选择“开始”菜单，并逐级打开其中的选项，当选中最低一层的菜单命令时，菜单命令将高亮显示，将鼠标停留片刻，将出现一条提示信息，内容为该菜单命令的功能说明。例如，选择“开始”|“程序”|“管理工具”|“计算机管理控制台”命令时，提示信息将告诉用户使用计算机管理控制台来管理本地或者远程计算机。

依次选择“开始”|“程序”|“附件”|“系统工具”命令之后，将打开如图 1-9 所示的菜单选项。在此仅有“备份”、“磁盘清理”、“磁盘碎片整理程序”、“任务计划”与“字符映射表”、“系统信息”等 6 项内容，而其他的“磁盘空间管理”、“维护向导”、“压缩代理”、“计划任务”等都被分配到“计算机管理控制台”菜单内。



图 1-9 “系统工具”菜单

(3) 控制面板。“控制面板”是 Windows Server 2003 的功能控制中心，选择“开始”|“设置”|“控制面板”命令之后，将打开如图 1-10 所示的“控制面板”窗口。在地址栏内选择其他选项，可以导航到其他的资源。在控制面板中，新添加、改进的工具具有管理工具、硬件向导、网络和拨号连接、扫描仪和照相机、SQL 服务器客户端配置等。



图 1-10 “控制面板”窗口

仔细观察可以发现，用于配置网络属性的“网络”图标并没有出现在控制面板内，这是因为桌面上的“网络邻居”完全可以胜任网络的配置、名称的标识等操作。在 Windows Server 2003 的控制面板左侧有两个站点连接：“Windows Update”和“技术支持”，它们分别与 <http://windows.update.microsoft.com/default.com> 和 <http://support.microsoft.com/support> 两个主页保持链接，通过这两个页面可获得 Windows Server 2003 的最新版本和技术支持。

(4) 连接网络。在桌面上双击“网络邻居”图标之后，即可浏览和使用网上的计算机资源，而无需通过驱动器映射来连接某个网络。要共享某个文件夹或打印机时，可右击文件夹或打印机，执行“共享”命令，配置共享属性。通过共享口令、用户列表的设置，可将被访问用户进行有效的控制。需要配置网络时，可双击控制面板的“网络和拨号连接”图标，弹出“网络和拨号连接”窗口。在默认的情况下，该窗口包括“新建连接”和“本地连接”两个选项，前者用于创建新的拨号连接，后者代表本机所在局域网的配置情况。

### 三、实验步骤

按照下面步骤熟悉 Windows Server 2003 的常用操作。

资源管理器。Windows Server 2003 的资源管理器在标准工具栏上增加了 3 个按钮，即“搜索”、“文件夹”和“历史纪录”。这些按钮可以帮助用户更方便快捷地找到所要的文件，“历史纪录”栏显示出用户以前浏览过的文件或 web 网页，排列方式也有多种方案，可以按时间的先后、站点的名称，排序也可以按访问次数的多少排序。

资源管理器的“搜索”功能很强，可以根据用户的需要进行 3 种类型的搜索：搜索文件或文件夹、搜索计算机、搜索 Internet。

## 实验 2 Windows Server 2003 操作系统安装

利用光驱直接从 Windows Server 2003 光盘启动安装，这是在没有操作系统或者操作系统损坏的情况下常用的安装方式。当然，如果计算机里有一个完好的操作系统也可以使用这样的安装方式。下面是这种安装方式的详细过程。

## 一、从光盘引导

首先将计算机的启动顺序更改为从 CD-ROM 启动,然后将 Windows Server 2003 的安装光盘放入光驱,计算机会引导 Windows Server 2003 光盘上的启动文件并直接进入 Windows Setup 安装界面。

**注意:** 如果计算机的硬盘里没有任何操作系统,马上就会进入 Windows Server 2003 的安装界面。如果计算机的硬盘里有其他操作系统,则 Windows Server 2003 光盘引导程序会等待 5s,在这 5s 内只要按任意键也会进入安装界面,否则将进入硬盘中已有的操作系统。

根据系统提示,按下 Enter 键继续 Windows Server 2003 的安装。接下来,系统会提示关于 Windows 的授权协议,它包括许可证的授予、产品的维护及更新、知识产权等。

## 二、创建磁盘分区

这一步是选择将 Windows Server 2003 操作系统安装在哪个磁盘分区上,如果磁盘上没有任何分区,则必须新建一个。安装中的选择如下。

(1) ENTER=安装。用箭头键选择要安装 Windows Server 2003 的已经建立好的分区,然后按下 Enter 键。

(2) C=创建磁盘分区。用箭头键选择未划分的空间,然后按下 C 键,将创建新的磁盘分区。

(3) D=删除磁盘分区。用箭头键选择所要删除的现存分区,然后按下 D 键。

(4) F3=退出。退出安装程序。

在安装过程中所做出的选择,取决于硬盘的状态:

(1)如果硬盘未分区,则必须为 Windows Server 2003 创建分区并为设定适当的空间大小。

(2)如果硬盘已有分区,并有足够大的未分区空间,可以考虑为 Windows Server 2003 在未分区的空间上创建新的分区。

(3)如果硬盘中已经分区的空间足够大,那么可以利用已经分区的空间来安装 Windows Server 2003。

(4)如果硬盘中分区的空间都不能满足 Windows Server 2003 的需要,可以通过删除更多的分区来创建新的来分区磁盘空间。然后用它来安装 Windows Server 2003。

由于安装的是 Windows Server 2003 的企业版,所以安装系统的分区大小不应该小于 1.5GB,如果硬盘已经有了一个分区,则可以通过将它删除来创建更多的未分区的磁盘空间,然后用它来创建 Windows Server 2003 的分区。

**提示:** 也可以利用安装程序创建其他分区,但向大家推荐只创建要安装 Windows Server 2003 的分区,并适当地设置大小。其余的分区最好是在安装完 Windows Server 2003 之后,用 Windows Server 2003 自带的图形界面的磁盘管理工具来划分。

## 三、格式化磁盘分区并选择文件系统

选择新建立的磁盘分区,然后按下 Enter 键。由于此系统刚刚创建还没有格式化,这时系统会提示用那种文件系统格式化此分区。一共有 4 个选择,其中有两个是快速格式化。快速格式化相对常规模格式化的速度要快一些,尤其是分区空间特别大的时候更加明显。对于新硬盘或者怀疑有问题的硬盘建议使用常规模格式化,因为常规模格式化虽然速度慢,但会检测硬盘

是否有坏道，还会将坏道标志排除。

用箭头键选择“用 NTFS 文件系统格式化磁盘分区 (快)”，按下 Enter 键会出现磁盘格式化的界面。

**提示：**如果是在真正的服务器上，为了提高文件的安全性，建议使用 NTFS 文件系统，如果是学习或者做实验，建议使用 FAT32 文件系统。

#### 四、复制文件并重新启动

格式化分区结束以后，安装程序将文件复制到 Windows 的安装文件夹，然后出现红色等待 1.5s 的倒计时状态条，如果不想等待，只要按任意键就可以立即重新启动计算机。

这时候系统已经完成了收集硬盘信息的过程，接下来将进入 Windows Server 2003 的图形界面安装阶段。

#### 五、进入图形界面安装

计算机重新启动以后马上会进入图形安装界面，此时，系统提示安装程序已经运行到哪个步骤，还需要多少时间才能完成安装，并通过屏幕展示 Windows Server 2003 在各个方面的优点。

**注意：**如果计算机在重新启动以后，出现蓝屏状态，一般情况下是硬件导致的，解决方法是更换硬件，比如内存、硬盘。

**提示：**如果是中文版可以使用默认值，直接单击“下一步”按钮。如果是英文版则要选择语言，否则将来如果遇到中文汉字还需要重新配置才能支持。

安装程序总共完成 4 项任务：

- (1) 安装“开始”菜单。
- (2) 注册组件。
- (3) 保存设置。
- (4) 删除任何用过的临时文件。

当这 4 项任务完成以后，整个安装程序结束，整个安装过程因计算机的配置不同大概需要 40~60min 的时间。

重新启动后，按 Ctrl-Alt-Del 组合键，用户名是 Administrator，输入安装时设置的管理员密码，然后按回车键，就可以登录到 Windows Server 2003，然后就可以使用了。