

3.1 计算机系统的基本概念

冯·诺依曼型计算机系统由硬件系统和软件系统两大部分组成,如图 3.1 所示。

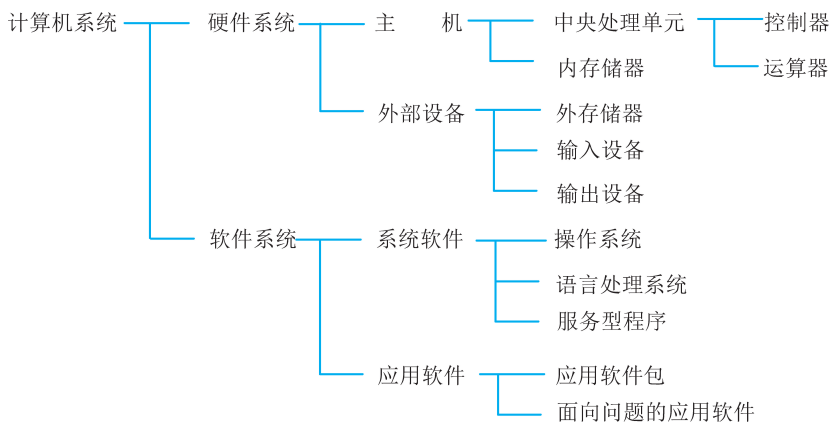


图 3.1 计算机系统的组成



3.1.1 计算机硬件系统

计算机硬件是指那些看得见、摸得着的部件,是构成计算机的物理装置。构成计算机的所有部件称为硬件(Hardware),这些硬件的整体结合称为硬件系统(Hardware System),是计算机系统的物理实现。如果按照层次抽象,计算机硬件系统由以下 4 个子系统组成。

1. 处理器子系统

处理器子系统是计算系统的核心部件,由运算器、控制器、寄存器和指令系统组成。处理器子系统的功能是实行指令控制、操作控制、时序控制和数据处理,它是实现冯·诺依曼“存储程序控制”的指挥系统,由指令系统发出控制命令,指挥各有关部件有条不紊地执行各项操作。

2. 存储器子系统

存储器子系统包括内存存储器和外存储器,是保存程序代码和数据的物理载体,也是计算机中存放程序和数据的各种存储设备、控制部件以及管理信息调度的设备和算法的总称。随着计算机应用领域越来越广泛,应用要求越来越高,数据自动存储能力成为现代计算机的重要指标。

3. 输入/输出子系统

输入/输出子系统用来实现输入设备、内存存储器、处理器、输出设备之间的相互连接和不同信息形式的转换,控制外部设备与内存存储器、外部设备与处理器之间进行数据交换。输入/输出子系统包括多种类型的输入设备、输出设备,以及连接这些设备与处理器、存储器进行信息交换的接口电路。

4. 总线子系统

总线子系统是多个功能部件之间进行数据传送的公共通路,是构成计算机系统的互联机构。总线子系统的主要特征是多个部件共享传输介质,一个部件发出的信号可以被连接到总线上的其他部件接收。采用总线传输方式使各子系统部件间实现地址、数据和控制信息的传递与交换,从而大大减少信息传送线的数量,增强系统的灵活性。

思考与探索

硬件系统是用正确的、低复杂度的芯片电路组合成高复杂度的芯片,逐渐组合,功能越来越强,这种层次化、构造化的思维是计算及自动化的基本思维之一。

3.1.2 计算机软件系统

1. 软件的定义

软件是为使计算机高效地工作所配置的各种程序及相关的文档资料的总称。其中:程序是经过组织的计算机指令序列,指令是组成计算机程序的基本单位;文档资料包括软件开发过程中的需求分析、方案设计、编程方法等的文档及使用说明书、用户手册、维护手册等。

1983年,IEEE对软件给出了一个较为新颖的定义:软件是计算机程序、方法、规范及其相应的文稿及在计算机上运行时必需的数据。这个定义在学术上有重要参考价值,它将程序与软件开发方法、程序设计规范及其相应的文档联系在一起,将程序与其在计算机上运行时必需的数据联系在一起。

2. 软件的分类

根据软件的功能作用,可将其分为系统软件、支撑软件和应用软件三类。

(1) 系统软件(System Software)。它是计算机厂家为实现对计算机系统的管理、调度、监视、服务、维护,以及扩充硬件功能等而提供给用户的软件。操作系统、翻译程序、服务程序都属于系统软件。系统软件的显著特点是与具体应用领域无关。

(2) 支撑软件(Supporting Software)。随着数据库应用系统开发和网络应用的不断拓展逐渐形成的软件。随着计算机技术的发展,软件的开发、维护与运行的代价在整个计算机系统中所占的比重越来越大,远远超过硬件系统。目前,常用的支撑软件有数据库管理系统、各类工具软件(如系统诊断、图像处理)、网络软件、软件开发环境、中间件(独立的系统软件或服务程序)等。

(3) 应用软件(Application Software)。它是相对于系统软件而言的,是用户针对各种具体应用问题而开发的一类专用程序或软件的总称。例如,计算机辅助设计(CAD)、计算机辅助测试(CAT)、计算机辅助制造(CAM)、计算机辅助教学(CAID)、专家系统(Expert System)、各类信息管理系统(Manage System)、科学计算(Scientific Computing)等,这些应

用软件在各有关领域大显神通,给传统产业注入了新的活力。

用户与计算机系统各层次之间的关系如图 3.2 所示。

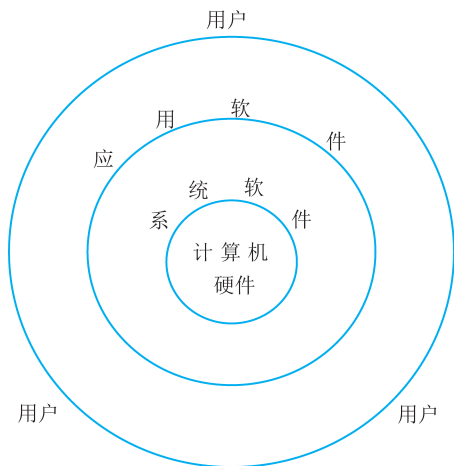


图 3.2 用户与计算机系统各层次之间的关系

思考与探索

为深刻理解和快速掌握软件的相关概念及其操作,系统软件的学习应从计算机硬件系统入手,应用软件的学习应从系统软件入手。软件的设计与开发是从特殊到一般的抽象和归纳思维,而软件的应用是从一般到特殊的具体化和演绎思维。

3. 软件系统中的交互方式

操作系统与大部分软件都提供程序式和交互式两种接口,本节主要介绍交互方式。

1) 操作系统中的交互方式

人类交互的最自然方式是通过语言、文字、图形、图像、声音和影像等表达自己的思想,因此,计算机软件也在不断地努力使其尽可能实现自然交互的方式。在操作系统中,桌面以及桌面上的各种形象化的图标的设计,都体现了软件操作界面的自然化模拟。

操作系统中,有一个专门处理交互方式的软件模块,称为操作系统的外壳(Shell),与之对应地,操作系统的核心功能部分称为内核(Kernel)。交互方式一般由软件的外壳软件来实现,外壳提供给用户的交互方式一般有两种:命令式和菜单式。

(1) 命令式交互方式。命令式交互方式的基本思想:人们通过简单的语言(命令)与计算机进行交互,请求计算机为我们解决各种问题。

Windows 操作系统中,单击图标和使用快捷键的这些方式可以看成命令式。

命令语言有肯定句和一般疑问句两种句型。基本格式一般包括动词、宾语和参数三部分。动词表示要做的具体任务,宾语表示任务对象。

例如,在 Windows 10 系统中,可以采用两种方式进入命令式交互方式:在“开始”菜单旁边的搜索框中输入 cmd 后按 Enter 键;或按下 Win+R 组合键,在弹出的对话框中输入 cmd,单击“确定”按钮。

这时,输入命令就可以与计算机进行交互。例如,“dir c:/p”命令可以分页查看 C 盘中的内容,如图 3.3 所示;输入“msconfig”命令就能打开“系统配置”对话框等。

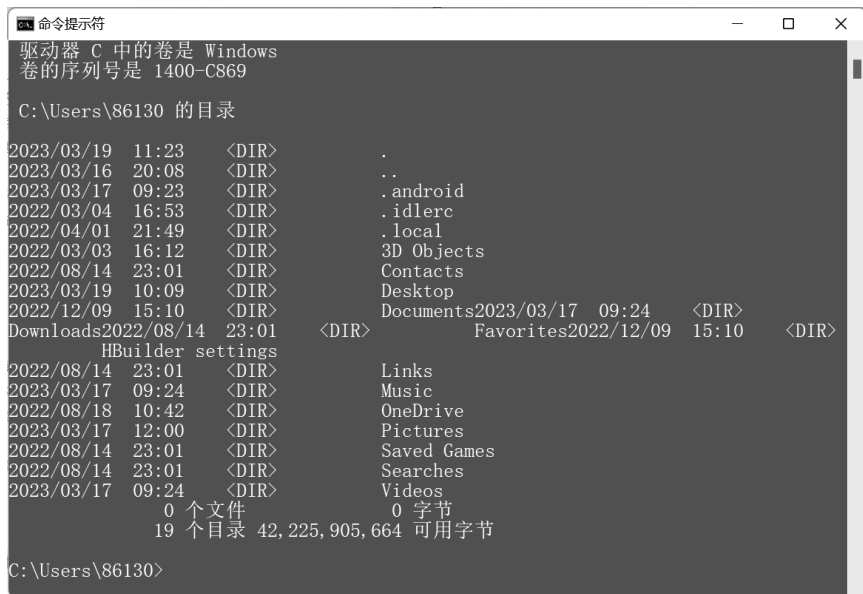


图 3.3 命令式交互方式

(2) 菜单式交互方式。菜单式与命令式本质上相同。菜单本身也是按树状结构思想分类组织的。在菜单式中,用户选择某个菜单项,就会调用与其对应的任务处理程序,这时菜单项(命令动词)使用对话框中的若干选项(各种命令参数),选择处理的具体对象(命令宾语),从而实现命令的解释和实现。

思考与探索

体验感受图形用户界面技术: 计算机应用之所以能够如此迅速地进入各行各业、千家万户,其中一个很重要的原因是 Windows 操作系统及其应用软件采用了图形化用户界面。图形化用户界面技术具有多窗口技术、菜单技术、联机帮助技术等特点。

以数据为中心和软件复用的思想: 交互式方式由以命令为主到以菜单为主的发展变迁,不仅反映了图形用户界面技术的优越性,而且反映了软件技术由功能型为主向数据型为主的转变。通过菜单将同样的功能运用到不同的数据集上,这种方式体现了以数据为中心和软件复用的思想。

2) 应用软件中的交互方式

应用软件运行在系统软件的基础之上,其启动和退出是通过系统软件的相关操作来完成的。启动是指将程序从计算机硬盘读到内存固定区域,并让其开始执行。关闭是指处于工作状态的软件停止运行,并正确地删除,释放所占内存空间。

应用软件的核心——应用程序,作为一种软件资源,一定存放在存储介质中,同一个应用软件可以运行多次,系统软件会把它们看作不同的多个任务来处理——多任务机制。一般地,常用的启动方式有以下几种。

(1) 基于查找的方式。基于查找的方式是指通过打开资源管理的树状目录结构,逐层地查找或通过搜索的方式找到应用程序,然后将其打开。

(2) 快捷方式。可以给任何文件、文件夹添加快捷方式。快捷方式是访问某个常用项

目的捷径。双击快捷方式图标可立刻运行这个应用程序、完成打开这个文档或文件夹的操作。例如,用户如果已经为打印机创建了快捷方式,那么以后要打印文件时,只需将该文件的图标拖到打印机图标上即可。

注意:快捷方式图标并不是对象本身,而是它的一个指针,此指针通过快捷方式文件(.lnk)与该对象联系。因此,对快捷方式的移动、复制、更名或删除只影响快捷方式文件,而不会改变原来的对象。

(3) 基于文件类型的方式。我们知道,系统软件中约定了一些专用文件的扩展名,表明了不同的文件类型。这时我们只要打开该类型的某一个个体文件即可启动其应用软件。比如,打开一个名为“我的大学规划.doc”文件即可启动 WPS 或 Word 应用软件。

3) 应用软件的操作模式

应用软件已逐渐趋于国际化软件的模式,通常它们有很多相似的操作模式,学会触类旁通,将大大提高学习和使用软件的效率。

(1) 菜单栏的设置模式。各种应用软件在不同的应用领域具有其显著的优势和特色,所以对于不同软件的使用学习,重点掌握其优势之处,再通过其与其他软件的共性学习,就会快速把握该软件的精髓,并且能够在适当的应用领域选择不同的应用软件来解决问题,从而进一步明白软件为什么要这样设计,为什么要提供这些功能,为今后设计软件奠定基础。

比如,菜单栏的设计,一般包含与资源管理有关的操作、编辑修改的操作、查看方面的操作、高级自定义设置方面的操作、自身软件的优势和特色、窗口布局有关的操作和联机帮助,一般命名为文件(File)、编辑(Edit)、查看或视图(View)、工具(Tool)、自身软件的优势和特色、窗口(Windows)、帮助(Help)。

(2) 快捷菜单——“右击无处不在”。一般地,在计算机屏幕的任何地方,使用鼠标右击,都会弹出一个快捷菜单——“右击无处不在”,该菜单包含右击对象在当前状态下的常用命令。快捷菜单具有针对性、实时性和快捷性,一般软件的常用功能均可以通过快捷菜单来完成。

(3) 快捷键和访问键。很多菜单项后面伴有带下画线的字母,表示该选项具有访问键,对于顶层菜单,按 Alt+访问键就可执行该项操作;对于子菜单,用户打开菜单后直接输入该字母即可执行。

有的很常用的菜单项后面跟着组合键,表示该选项具有快捷键,用户不必打开菜单,直接按下此快捷键,就可执行该项操作。比如,菜单项“复制(C) Ctrl +C”。

因为有些软件诞生于西方,所以这些快捷键和访问键往往使用该菜单项的英文单词本意的首字母,像复制(C)就是 Copy 的首字母、打开(O)就是 Open 的首字母等,使用单词本意来学习,会更好地触类旁通,实现知识的迁移。

(4) 文档格式设置策略。常用的文字处理软件和电子表格处理软件中文档的格式设置策略,体现和应用了正向(演绎)思维和反向(归纳)思维。

- 正向(演绎)思维:先指定整个文档的各种格式,然后再输入具体内容。
- 反向(归纳)思维:先输入具体内容,然后再设置整个文档的各种格式。这种方法较为普遍。

大家在实际应用中,可以根据情况自主选择,也可以将两者结合起来使用。

(5) 对象的嵌入与链接技术。对象的嵌入与链接又称为 OLE。嵌入和链接的主要区

别在于数据的存放位置以及将其插入目标文件后的更新方式的不同。

链接对象是指在修改源文件之后,链接对象的信息会随着更新。链接的数据只保存在源文件中,目标文件中只保存源文件的位置,并显示代表链接数据的标识。如果需要缩小文件大小,应使用链接对象。

嵌入对象是指即使更改了源文件,目标文件中的信息也不会发生变化。嵌入的对象是目标文件的一部分,而且嵌入之后,就不再和源文件发生联系。双击嵌入对象,将在源应用程序中打开该对象。

文档和文档间、应用程序和应用程序间通过 OLE 技术,自身的功能大大丰富和扩充了,而且这也是递归思想的体现。

思考与探索

各种应用软件在不同的应用领域具有其显著的优势和特色,注重捕捉软件及其使用过程中的经验规律和模式,掌握使用该软件的精髓,并在学习和使用软件的过程中注重总结归纳其共性,会大大提高学习、使用和设计软件的能力。比如,Shift 键配合鼠标往往实现多个连续对象的选择;Ctrl 键配合鼠标往往实现多个不连续对象的选择。又如,用鼠标从任意方向包围需要选择的对象,往往能够实现选择多个对象。大家可以在 Windows、WPS 等多个软件中体会。这种总结归纳的思维方式对于提升终身学习能力很有益处,这也是一种知识迁移的思维方式。

3.2 计算机操作系统

3.2.1 操作系统的基本概念与作用

1. 操作系统的定义

操作系统(Operating System, OS)是随着硬件和软件不断发展而逐渐形成的一套大型程序。从功能上,它是用户与计算机硬件之间的接口,是用户和其他软件与计算机硬件之间的桥梁;从作用上,它为用户操作和使用计算机提供了一个良好的操作与管理环境,使计算机的使用效率成倍地提高,并且为用户提供了方便的使用手段和令人满意的服务质量。因此,可将其定义为:操作系统是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,并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。

2. 操作系统的作用

现在呈现在用户面前的计算机,是经过若干抽象的计算系统,我们可把整个计算系统按功能划分为 4 个层次,即硬件系统、操作系统、支撑软件和应用软件。这 4 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。其中,操作系统密切地依赖于硬件系统,其功能作用主要体现在以下 3 个方面。

(1) 操作系统是用户与计算机硬件之间的接口:用户需要通过操作系统来使用计算机,操作系统使用户不需要过多了解硬件的情况下,方便地控制计算机中的资源,并且能够根据用户需求对硬件进行改造和扩充。因此,可把操作系统看作人-机交互(Human-

Computer Interaction)的接口或界面。

(2) 操作系统是计算机系统的资源管理者:基本任务是管理计算机系统软硬件资源,是软硬件之间的协调者。一方面,它控制和管理着系统硬件,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,把一个裸机变成可“操作”且方便灵活的计算机系统;另一方面,计算机中的程序和数据以文件形式存放在外存储器中,如何与内存储器进行数据交换,都由操作系统来实现。操作系统的管理包括处理器管理、存储管理、文件管理、设备管理,最大限度地提高资源的利用效率和实现资源共享,操作系统使计算机成为一台“海纳百川”的机器。

(3) 操作系统为用户提供了一个虚拟机(Virtual Machine):计算机硬件系统的功能是有限的,但通过操作系统及其相关软件,可以完成多种多样、复杂多变的任务,很多需要硬件实现的功能可以由软件实现,许多硬件实现不了的功能也由软件来实现。为了方便用户开发更为复杂的程序,操作系统提供了更容易理解的、任务相关的、控制硬件的命令,被称为应用程序接口(Application Program Interface, API),它将对硬件控制的具体细节封装起来,通过在计算机裸机上加上一层又一层的软件来组成一个完整的计算机系统,为用户提供了一台扩展基本功能、使用更为方便的机器,称为虚拟机(Virtual Machine),它是一个功能强大、安全可靠、效率极高的计算机系统。

3.2.2 操作系统用户接口

用户接口(User Interface)是为方便用户操作使用计算机而提供的人-机交互接口,分为3种接口方式:命令接口、程序接口和图形接口。

(1) 命令接口。为了便于用户直接或间接控制自己的程序,操作系统提供了命令接口,用户通过该接口向计算机发出命令,执行功能操作。

(2) 程序接口。它是为用户程序访问系统资源而设置的接口,为程序中使用操作系统提供的系统调用命令请求操作系统服务,也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,例如,DOS是以系统功能调用的方式提供程序接口,可以在编写汇编语言程序时直接调用。Windows操作系统是以API的方式提供程序接口,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。在可视化编程环境(VB、VC++、Delphi等)中,提供了大量的类库和各种控件,如微软基础类(Microsoft Foundation Classes, MFC),这些类库和控件构建在Win API函数之上,并提供了方便的调用方法,极大地简化了Windows应用程序的开发。

(3) 图形接口。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,但要求用户严格按照规定的格式输入命令。显然,这不便于操作使用。于是,图形用户接口(Graphical User Interface, GUI)应运而生。

GUI采用图形化的操作界面,用户利用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。在Windows操作系统中,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。此时用户不必像使用命令接口那样去记住命令名及格式,只要单击鼠标就能实现很多功能,从而使用户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。这也是Windows操作系统受到用户欢迎并得以迅速发展的原因。

3.2.3 常用的操作系统

常用的操作系统有以下几种。

(1) MS-DOS。MS-DOS 是 Microsoft 公司推出的配置在 16 位字长 PC 上命令行界面的单用户单任务的操作系统,它对硬件要求低,现已逐渐被 Windows 替代。

(2) Windows。Windows 是 Microsoft 公司推出的基于图形界面的单用户多任务的操作系统,是 20 世纪 90 年代以后使用率最高的一种操作系统。Windows 采用了图形化模式 GUI,比起从前的 DOS 需要键入指令使用的方式更为人性化。

(3) UNIX。UNIX 是一种运用较早、使用率较高的网络操作系统之一,是通用、交互式、多用户、多任务的操作系统,是在科学领域和高端工作站上应用最广泛的操作系统。它采用 C 语言编写,可移植性强,由于 UNIX 强大的功能和优良的性能,使之成为被业界公认的工业化标准的操作系统。

(4) Linux。Linux 是 20 世纪 90 年代由芬兰赫尔辛基大学的学生 Linus Torvalds 创建并由众多软件爱好者共同开发的操作系统。它是由 UNIX 衍生出来的,性能与 UNIX 接近,最大特点是它的所有源代码都开放,用户可以免费获取 Linux 及其生成工具的源代码,并可以进行修改,建立自己的 Linux 开发平台,开发 Linux 软件。Linux 与 UNIX 兼容,能支持多任务、多进程、多 CPU 和多种网络协议,是一个性能稳定的多用户网络操作系统。

(5) 移动设备的操作系统。智能手机、平板电脑和掌上电脑(Personal Digital Assistant, PDA),我们都称为移动设备。移动设备里的操作系统我们自然也称为“移动设备操作系统”,它在传统 PC 操作系统的基础上又加入了触摸屏、移动电话、蓝牙、Wi-Fi、GPS、近场通信等功能模块,以满足移动设备所特有的需求。

主流的移动设备的操作系统有 Android、苹果的 iOS 和塞班等。Android 是一种基于 Linux 的自由及开放源代码的操作系统,主要使用于移动设备,如智能手机和平板电脑。

(6) 实时嵌入式操作系统。单片机的操作系统软件叫嵌入式系统。嵌入式系统是用于控制、监视或者辅助操作机器和设备的装置。

嵌入式系统几乎都是实时操作系统。实时系统比普通操作系统的响应速度更快。系统收到信息后,没有丝毫延迟,马上就能做出反应,因此才称为“实时”。实时系统被广泛用于对时间精度要求非常苛刻的领域,例如,工业控制系统、数字机床、电网设备监测、交通管理中的全球定位系统(Global Positioning System, GPS)、科学实验的精准控制、医疗图像系统、飞机和导弹中的导航系统、商业自动化设备(如自动售货机、收银机等)、家用电器设备(如微波炉、洗衣机、电视机、空调等)、通信设备(如手机、网络设备等)等,都需要用到实时嵌入式系统。

(7) 分布式操作系统。随着网络技术的出现与发展,大量联网的计算机可以通过网络通信,相互协调一致,共同组成一个大的运算系统——分布式计算系统,系统中的每台计算机都有独立的运算能力,各个计算机内运行的分布式程序之间相互传递信息,彼此协调,共同完成特定的运算任务,称为分布式操作系统。分布式系统具有可靠性高和扩展性好的优点。系统中任何一台(或多台)机器发生故障,都不会影响到整个系统的正常运转。同时,整个系统的结构是可以动态变化的,也就是说,随时可以有新的计算机加入系统中来,也随时可以有机器从系统中被移除。而且,系统中的计算机可以是多种多样的,网络连接形式也可

以是多种多样的。

3.2.4 操作系统的管理功能

操作系统是一个庞大的管理控制程序,一个操作系统通常包括进程管理、中断处理、内存管理、文件系统、设备驱动、网络协议、系统安全和输入/输出等功能模块。

1. 进程管理

计算机程序通常有两种存在形式:一种是人(通常是程序员)能够读懂的“源程序”形式。源程序经过某种处理(行话叫“编译”)就得到了程序的另一种形式,也就是我们常说的“可执行程序”,或者叫“应用软件”。源程序是给人看的:程序员阅读、学习、修改源程序,然后可以生成新的更好的可执行程序。可执行程序通常是人看不懂的,但计算机能读懂它,并按照它里面的指令做事情,以完成一个运算任务。

那么什么是进程呢?一个运行着的程序,我们就把它称为“进程”。具体点说,程序是保存在硬盘上的源代码和可执行文件,当我们要运行它的时候,如当你要运行浏览器程序的时候,会在浏览器图标上双击,这个浏览器程序的可执行文件就被操作系统加载到了内存中,一个浏览器进程就此诞生了。之后,CPU 会逐行逐句地读取其中的指令,这也就是所谓的“运行”程序了。直到你上网累了,关闭了浏览器窗口,这个进程也就终止了。但浏览器程序(源代码和可执行文件)还原封不动地保存在硬盘上。

一个运转着的计算机系统就像一个小社会,每个进程都是这个社会中活生生的人,而操作系统就像是政府,它负责维持社会秩序,并为每一个进程提供服务。进程管理就是操作系统的重要工作之一,包括为进程分配运行所需的资源,帮助进程实现彼此间的信息交换,确保一个进程的资源不会被其他进程侵犯,确保运行中的进程之间不会发生冲突。

进程的产生和终止、进程的调度、死锁的预防和处理……这些都是操作系统对于进程的管理工作。

2. 中断处理

操作系统的工作会经常被打断,而且被打断得非常频繁。任何一个进程如果要请求操作系统帮它做点什么,如读写磁盘上的文件,都要去“敲操作系统的门”,也就是要操作系统中断手头的工作,来为它服务。针对不同的服务请求,操作系统会调用不同的“中断处理程序”来处理。操作系统里有数以百计的“中断处理程序”来处理各种各样的服务请求。

操作系统中断处理分为硬件中断和软件中断两类。

(1) 硬件中断。硬件中断就是“硬件来敲门,请求服务”,是外围硬件设备(如键盘、鼠标、磁盘控制器等)发给 CPU 的一个电信号。在键盘上每按下一个键,都会触发一个硬件中断,于是 CPU 就要立即来处理,把我们敲的字母显示到屏幕上。

(2) 软件中断。软件就是一系列指令。所谓“软件运行”,就是 CPU 逐行地读取并执行这些指令。在一个软件程序中,通常有很多指令都会请求操作系统提供某种服务。由这些程序指令触发的中断就叫“软件中断”。比如说,一个进程要产生子进程、要读写磁盘上的文件、要建立或删除文件等,这些任务都要在操作系统的帮助下才能完成。

另外,系统运行过程中出现的硬件和软件故障也会向操作系统发出中断信号,以便这些

意外情况能及时得到处理。

3. 内存管理

在计算机里有很多可以存放程序和数据的地方,按从里向外的层次依次有:寄存器、缓存区、内存、硬盘、光盘、U 盘等。

一个程序如果要运行起来,必须先把它加载到内存中。为什么呢?因为寄存器和缓存区太小,通常放不下一个程序。而硬盘又太慢,如果让 CPU 直接从硬盘里读指令的话,速度将是内存里读指令的速度的百万分之一。因此,内存这个速度较快,而且又能容纳下不少东西的地方,就成了我们加载程序的唯一选择。

内存管理是操作系统的重要工作。操作系统是计算机内硬件资源的管理者,而内存就是最为抢手的计算机硬件资源之一。大大小小的程序如果要运行,必须由操作系统给它们分配一定的内存空间。内存空间的分配是否合理直接关系到计算机的运行速度和效率。

(1) 操作系统内存管理的主要任务如下:

- 随时知道内存中的哪些地方被分配出去了,还有哪些空间可用;
- 给将要运行的程序分配空间;
- 如果有程序结束了,就把它占用的空间收回,以便分配给新的进程;
- 保护一个进程的空间不会被其他进程非法闯入;
- 为相关进程提供内存空间共享的服务。

(2) 虚拟内存。实际内存的使用情况只有操作系统才需要知道,用户进程看到的内存并不是真正的物理内存,而是一个“虚拟大内存”,大到系统能支持的上限。对于 32 位系统来说,这个上限是 2^{32} B,也就是 4GB。而 64 位操作系统的寻址能力就是 2^{64} B,也就是 17 179 869 184GB,当然这只是理论值,实际中不可能用到这么大的内存,目前 64 位 Windows 系统最大只支持 128GB。

即使实际可用的物理内存小于用户进程所需空间,进程也可以运行,因为用户进程并不需要 100% 被加载到内存中。实际上,一个程序经常包含一些极少被用到的功能模块。比如,用于出错处理的功能模块,如果程序不出错的话,这部分功能模块就没必要加载到内存中。一旦需要加载程序剩余的部分,而找不到可用空间的话,操作系统可以“拆东墙补西墙”,把暂时不运行的进程挪出内存,以腾出空间加载正要运行的程序。

上述内存管理方式采用的就是“虚拟内存”技术。它将用户进程和物理内存隔离开来,给用户进程一个“虚拟内存”的概念,是内存管理的一大飞跃。虚拟内存不仅提高了系统的安全性,还可以让更多的进程同时运行,使内存的使用效率大大提高。同时,程序员在编程时,不必考虑物理内存有多大,这也降低了编程的复杂度。

4. 设备驱动

操作系统和硬件设备打交道依靠的就是设备驱动程序。操作系统内部有很多设备驱动程序。例如,上网要用网卡;听音乐要用声卡等。键盘、鼠标、硬盘……所有这些硬件设备都必须有相应的驱动程序才能正常工作。

现代通用操作系统,如 Windows、Linux 等都会提供一个 I/O 模型,允许设备厂商按照此模型编写设备驱动程序,并加载到操作系统中。目前的 Windows 和 Linux 操作系统都支持即插即用,即插即用是一种使用户可以快速简易安装某硬件设备而无须安装设备驱动程

序或重新配置系统的标准。即插即用需要硬件和软件两方面支持,因此主要是看计算机配件是否支持即插即用。如具备即插即用功能,安装硬件就更为简易。

3.2.5 操作系统中的计算思维

1. 树状目录结构与资源管理

在 Windows 中常常利用“资源管理器”和“此电脑”进行信息资源管理,在查看和显示信息时,我们用到了树状目录结构。利用树状目录结构来进行资源管理是计算机中一个重要思想,一般地,涉及资源管理的操作都会使用树状结构。

这种树状结构的设计思想在我们日常生活中也常常用来进行信息的分类组织,比如,我们表达一个单位机构的层次结构等。另外,树状目录结构中,随着当前盘和当前文件夹的转换,也就是当前视点的不断转换,我们是不是感受到了一种层次化结构性的跳跃性思维?这种思维方式是计算机学科一个重要的思维特征。比如,网络域名管理、面向对象的分析与设计方法中的类及其继承体系的应用、Java 中的包管理及引用、程序三大结构的理解、网络规划与设计等。

2. 信息共享机制

Windows 内部出色地提供了各种机制,使应用程序能够迅速而方便地共享数据和信息,这些机制包括剪贴板、对象链接与嵌入、动态数据交换等。

剪贴板是由操作系统维护的一块内存区域,是在 Windows 程序和文件之间传递信息的临时存储区,可以存储正文、图像、声音等多种多样的信息,可以实现不同应用程序间的信息交换。

对象链接与嵌入技术(Object Linking and Embedding, OLE)可以用来创建复合文档,复合文档可以包含来源于不同源应用程序的文件,有着不同类型的数据,即可以把文字、声音、图像、表格、视频、应用程序等组合在一起。OLE 是一种面向对象的技术,对象被赋予了智能属性,即参与链接和嵌入的对象本身带有计算机指令,所以导致它具有明显的缺点:缓慢而且庞大。

动态数据交换(Dynamic Data Exchange, DDE)是一种允许数据在程序间被共享或者通信的技术,可以用来协调操作系统的应用程序之间的数据交换及命令调用。

3. 回收站——恢复机制

“回收站”是计算机硬盘中的一个名为“Recycled”的文件夹,用于存放被删除的文件、文件夹和快捷方式等对象,处于被回收状态。回收站中的对象仍然占用磁盘空间,但是可以恢复,给用户提供一种“后悔药”。这是一种通过纠错方式,在最坏情况下进行预防、保护和恢复的思维,是一种常用的工程思维。在软件设计时应应对可能发生的种种意外故障采取措施。软件是很脆弱的,很可能因为一个微小的错误而引发严重的事故,所以必须加强防范。

思考与探索

在电子类和机械类等工程中,哪些设计属于“回收站”式的工程思维下的产物呢?

3.2.6 我国操作系统的发展

我国在 20 世纪末期就开始研究国产的操作系统,过去 20 年,诞生超过 20 多个不同的

版本,红旗 Red Flag、深度 Deepin、中标麒麟 Neokylin、银河麒麟 Kylin、中兴新支点等都被人所熟知,基本上都是基于 Linux 内核进行开发的。随着技术的成熟,国产操作系统已经从“可用”阶段向“好用”阶段发展,同时也可以支持绝大部分的常用办公软件。基本可以满足企业日常办公、电子政务、智慧城市、国防军工、教育、能源交通、党政机关、医疗、电信、教育、金融、生产作业系统以及安全可靠等多个领域应用需求。

在 2019 年 8 月 9 日,华为正式发布鸿蒙操作系统。2020 年 9 月 10 日,华为鸿蒙系统升级至华为鸿蒙系统 2.0 版本,华为鸿蒙系统是一款全新的面向全场景的分布式操作系统,创造一个超级虚拟终端互联的世界,将人、设备、场景有机地联系在一起,将消费者在全场景生活中接触的多种智能终端实现极速发现、极速连接、硬件互助、资源共享,用最合适的设备提供最佳的场景体验。

华为鸿蒙系统能够把手机的内核级安全能力扩展到其他终端,进而提升全场景设备的安全性。通过设备能力互助,共同抵御攻击,保障智能家居网络安全;通过定义数据和设备的安全级别,对数据和设备都进行了分类分级保护,确保数据流通安全可靠。

思考与探索

递归思想:操作系统作为最基本的系统软件,是计算机硬件系统和软件系统(包括系统软件和应用软件)的组织者和管理者,因此,操作系统作为软件也要受到其自身的管理和控制。这是不是和“老和尚给小和尚讲故事”很相似?“从前有座山,山里有个庙,庙里有个老和尚,给小和尚讲故事。故事讲的是:从前有座山,山里有个庙,庙里有个老和尚,给小和尚讲故事……”,这就是著名的递归思想,即整体由局部组成,整体又可以作为局部。在日常生活中,递归一词较常用于描述以自相似方法重复事物的过程。例如,当两面镜子相互之间近似平行时,镜中嵌套的图像是以无限递归的形式出现的。还有哪些计算机应用体现或使用了递归思想呢?

3.3 文件管理系统

为了管理外存储器上的文件,操作系统中设置了管理文件的功能模块——文件管理系统,简称为文件系统。它负责管理文件,并为用户提供对文件进行存取、共享及保护的手段,这不仅方便了用户,保证了文件的安全性,还可以有效地提高系统资源利用率,实现对磁盘文件“长治久安”。

3.3.1 文件与文件系统

操作系统对文件管理的实质是对存放文件的磁盘存储空间管理。具体说,就是当建立一个新的文件时,文件系统要为其分配相应的存储空间;删除一个文件时,要及时收回其所占用的空间。

1. 文件

逻辑上具有完整意义的信息集合称为文件。计算机中,所有的程序和数据都是以文件的形式进行存放和管理的。文件通常由若干记录组成,而记录是一些相关数据项的集合,其数据项是数据组织中可以命名的最小逻辑单位。例如,一位职工的信息记录由姓名、性别、

籍贯、出生日期、学历、职称等数据项组成,而一个单位职工信息记录的集合就组成了一个文件。

在计算机系统中,一个完整的程序或一组完整的数据记录都是一个文件,它是操作系统实行信息管理的基本单位。文件中记录的符号元素不仅可以是文字、公式、表格、插图信息,还可以是视频信息和音频信息等,而承载这些信息的媒介是外存储器(各类磁盘),并将磁盘中的文件称为磁盘文件。

2. 文件系统

文件系统是操作系统中与文件管理有关的软件和数据相结合,操作系统中负责存取和管理信息的模块,是对文件的存储空间进行组织和分配,负责文件的存储,并对存入文件进行保护和检索的系统。它用统一的方式管理用户和系统信息的存储、检索、更新、共享和保护,并为用户提供一套高效的文件使用方法,是对文件组织和管理的层次抽象。

文件系统由三部分组成:与文件管理有关的软件、被管理的文件、实施管理所需的数据结构。

(1) 逻辑文件系统层。处理文件及记录的相关操作。例如,允许用户利用符号文件名访问文件及文件中的记录,实现对文件及记录的保护,实现目录操作等。

(2) 文件组织模块层。又称为基本 I/O 管理程序层,完成与磁盘 I/O 有关的工作,包括选择文件所在设备,进行文件逻辑块号到物理块号的转换,对文件空闲存储空间管理,指定 I/O 缓冲区。

(3) 基本文件系统层。又称为物理 I/O 层,负责处理内存和外存之间的数据块交换,只关心数据块在辅助存储设备和在主存缓冲区中的位置,不需了解所传送数据块的内容或文件结构。

(4) 设备驱动程序层。又称为基本 I/O 控制层,主要由磁盘驱动程序组成,负责启动设备 I/O 操作及对设备发来的中断信号进行处理。

3.3.2 文件的命名

一般地,文件名反映文件的内容和类型信息。给文件起名时,应尽可能“见名知义”,这样有助于记忆和查找。

(1) 文件名格式:主名.扩展名,主名表达文件内容,扩展名表达文件类型。

(2) 约定一些专用文件的扩展名,表明了不同的文件类型。常见的有:.exe(可执行文件)、.com(系统命令文件)、.sys(系统直接调用文件)、.bat(批处理文件)、.obj(目标程序文件)、.bak(备份文件)、.tmp(临时文件)、.txt(文本文件)、.doc(Word 文档)、.xls(Excel 工作簿文件)、.ppt(PowerPoint 演示文稿)等。在 Windows 操作系统中还给不同类型的文件赋以形象的图标。

(3) 一些常用的设备也作为文件处理。常见的设备文件名有:CON(键盘/屏幕)、PRN 或 LPT1(第一并行打印机)、LPT2(第二并行打印机)、AUX 或 COM1(第一串行口)、COM2(第二串行口)、NUL(虚拟外部设备)。用户在给文件命名时不能使用系统保留的这些设备名。

(4) 查找和显示时可以使用通配符“*”和“?”,“*”代表任意多个字符(包括 0 个);“?”代表任意一个字符。例如,“file*”可以代表“file123”“file1”“file2”“file.doc”;“file?”可以代

表“file1”“file2”; A *.doc 可以代表主文件名以 A 开头、扩展名为 DOC 的所有文件,像“ASTB.doc”“aBX.doc”“ADEF.doc”等。

1. Windows 文件的命名规定

(1) 文件名中可以是数字、大小写字母、汉字和多个其他的 ASCII 字符,最多可以有 255 个字符(包括空格),忽略文件名开头和结尾的空格。

(2) 不能有以下字符出现: \ / : * ? " < > |。

(3) 文件名中可以分别使用英文字母大写和小写,不会将它们转换成同一种字母,但认为大写和小写字母具有同样的意义。例如,MYFILE 和 myfile 认为是同一个文件名。

(4) 可以使用多个分隔符的名字。例如,“myfile.examples.2010”和“学习计划.2010.xls”等。

2. MS-DOS 文件的命名规定

MS-DOS 中文件的命名除符合文件的一般规定外,还有以下一些规定。

(1) 主文件名最多只允许 8 个字符,扩展名最多只允许 3 个字符,称为“8.3”型文件名。

(2) 这些字符可以是:大小写英文字母、0~9 个数字、汉字及一些特殊符号[如 \$、#、&、@、<、>、~、|、^、(、)、-、{、}等]。

3.3.3 文件目录与目录结构

磁盘文件管理采用了一种“逐级分化”的计算思维,即建立便于文件存放和查找的文件目录及其目录结构。

1. 文件目录

文件是信息的集合,文件与信息的关系形如书本的“节”与节中的符号元素,“节名”便是文件名,节中的文字、公式、表格、插图等信息便是文件的内容。书本是承载符号元素的载体。

为了便于文件存储和查找,可在磁盘中建立多个区域,然后在各区域中建立多个“文件夹”,而且允许在各文件夹中建立下一级文件夹,把不同文件分别存放到各相关文件夹中,该文件夹被称为文件目录,该目录名就是文件夹名。在每一个文件夹中,可以存放多个不同名字的文件;而在不同的文件夹中,可以使用(命名)相同名字的文件。

2. 目录结构

DOS 和 Windows 采用分层式的文件系统结构,即树状目录结构(Tree Directory Structure)。目录和文件的隶属关系像一棵倒置的树,树枝称为子目录,树的末梢称为文件,这里以 C 盘为例讲述文件系统的层次结构,如图 3.4 所示。

这种文件系统的树状结构有以下几个特点。

(1) 每一个磁盘只有唯一的一个根结点,称根文件,用“\”符号表示,如“C:\”。

(2) 根结点向外可以有若干个子结点,称为文件夹(Folder)。每个子结点都可以作为父结点,再向下分出若干个子结点,即文件夹中可以有若干个文件和子文件夹。

(3) 在同一个文件夹中,不允许有相同的文件名和子文件夹名。

在文件系统和层次结构中,一个文件的位置需要由 3 个因素来确定:文件存放的磁盘、存放的各级文件夹和文件名。在文件层次结构中,一个文件的完整定位为:

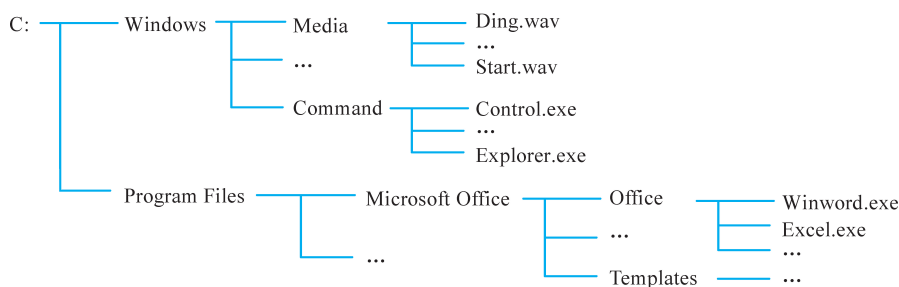


图 3.4 树状目录结构

[盘符][路径]主文件名[.扩展名]

说明：

(1) []表示可以省略。

(2) 盘符。用磁盘名加上一个“:”，如“C:”等。

(3) 路径。树状结构中，文件夹呈层次关系。当对某个文件或某一文件夹进行操作时，必须指出该文件或文件夹的存取路径。从某一级文件夹出发(可以是根文件夹，也可以是子文件夹)，去定位另一个文件夹或文件夹中的一个文件时，中间可能需要经过若干层次的文件夹才能到达，所经过的这些文件夹序列就称为路径。各文件夹后面要加一个“\”符号。

例如，在图 3.4 中，从根文件夹到“Control.exe”文件的路径表示为：“C:\Windows\Command\Control.exe”。

(4) 当前文件夹。引入多级文件夹后，对任何一个操作都需要知道当前系统所在的“位置”，也就是说要明确当前的操作是从哪一个文件夹出发的。把执行某一操作时系统所在的那个文件夹称为当前文件夹。

(5) 绝对路径。绝对路径是指从根文件夹出发表示的路径名。这种表示方法与当前文件夹无关。例如，在图 3.4 中，\Windows\Media\Ding.wav 表示从根文件夹出发定位文件 Ding.wav。其中：第一个字符“\”表示根文件夹，中间的“\”表示子文件夹之间或子文件夹与文件之间的分隔符。

(6) 相对路径。相对路径是指从当前文件夹开始表示的路径，这种表示方法与当前文件夹密切相关。

以图 3.4 为例，设当前目录为 Media，那么 Ding.wav 表示定位当前目录下的文件 Ding.wav；..\Command\Control.exe 表示先返回到当前目录的父目录(Windows 文件夹)，再向下定位文件 Control.exe。其中，符号“..”代表当前文件夹的父文件夹。操作系统在建立子文件夹的时候，自动生成两个文件夹，一个是“.”，代表当前文件夹；另一个是“..”，代表当前文件夹的上一级文件夹。

3. 文件管理

建立文件目录结构，为文件的管理和使用提供了极大方便。为了便于用户操作、管理和维护，在文件管理中采用了如下管理措施。

(1) 文件分类。为了便于管理和控制，可将文件分为多种类型。如果按用途划分，可分为系统文件、库文件和用户文件；如果按数据形式划分，可分为源文件、目标文件和可执行

文件。

(2) 文件命名。为了便于操作使用,每个文件必须有唯一的标记,称为文件名。文件名由主文件名和扩展名两部分组成:主文件名可以由多个连续的字符组成,类似一个人的姓名;扩展名是根据需要加上的,一般用来标识文件的性质(如系统文件、库文件、可执行文件等,类似于标明某人的性别)。扩展名跟在主文件名之后,以“.”开头,后跟1~3个连续的字符。

(3) 文件属性。操作系统为不同类型文件定义了独特性质,使每个文件能选取不同的属性:隐含(Hidden)、系统(System)、只读(Read only)、档案(Archive)。它们构成了属性组“HSRA”。

(4) 文件操作。为了便于实现文件管理,操作系统提供了许多文件操作功能,如创建文件、复制文件、修改文件、删除文件等。

3.4 处理器系统

现代计算机是一个高度自动化的计算系统,有人把计算机称为“电脑”,是因为它具有类似于人类大脑、按照人的意愿,指挥和控制计算机执行各种操作的功能,这个指挥和控制机构就是由运算器、控制器、寄存器、指令系统等构成的处理器系统,它是计算机硬件系统的核心。

3.4.1 处理器的结构组成

随着超大规模集成电路的高速发展,使得电子器件的体积越来越小,为了使主体器件具有一致性能,则将运算器和控制器集成在一块芯片上,称为中央处理器(Central Processor)或中央处理单元(Central Processing Unit,CPU),在微机中称为微处理器(Microprocessor)。

1. CPU 的基本组成

CPU 是计算机中的核心部件,用来实现运算和控制,并由运算器、控制器、控制线路等组成。

(1) 运算器(Arithmetical Unit)。运算器是用来完成算术运算和逻辑运算的部件,其功能:快速地对数据进行加、减、乘、除(包括变更数据符号)等算术运算;“与”“或”“非”等逻辑运算,逻辑左移、逻辑右移、算术左移、算术右移等移位操作;及时存放算术运算和逻辑运算过程中的中间结果(由通用寄存器组实现),实现挑选参与运算的操作数、选中执行的运算功能,并且把运算结果送到存储器中。

运算器由多功能算术逻辑运算部件(Arithmetic Logical Unit, ALU)、通用寄存器组(包括累加寄存器、数据缓冲寄存器、状态寄存器)及其控制线路组成。寄存器组用来存放运算器的工作信息和运算中间结果,减少访问存储器的次数,以提高运算器的速度。整个运算过程是在控制器的统一指挥下,对取自 RAM 的数据按照程序的编排进行算术或逻辑运算,然后将运算结果送到 RAM。

(2) 控制器(Control Unit)。控制器是计算机系统发布操作命令的部件,其功能是根据指令提供的信息,实现对系统各部件(不仅包括 CPU)的操作和控制。例如,计算机程序和原始数据的输入、CPU 内部的信息处理、处理结果的输出、外部设备与主机之间的信息交换

等,都是在控制器的控制下实现的。

2. CPU 中的主要部件

CPU 是计算机硬件系统的指挥中心,有人将它形容为人脑的神经中枢。CPU 的指挥控制功能由指令控制、操作控制、时间控制、数据加工等部件实现。

(1) 数据缓冲寄存器(Data Register,DR)。存放从 RAM 中取出的一条指令或数据。

(2) 指令寄存器(Instruction Register,IR)。存放从 RAM 中取出的将要执行的一条指令。

(3) 指令译码器(Instruction Decoder,ID)。执行 IR 中的指令,必须对指令的操作码进行检测,以便识别所要求的操作。

(4) 地址寄存器(Address Register,AR)。用来存放下一条将要执行的指令的地址码。

(5) 累加寄存器(Accumulator,AC)。简称位累加器,当 ALU 执行算术或逻辑运算时为 ALU 提供工作区。

(6) 状态寄存器(Flag Register,FR)。用来存放算术和逻辑运行或测试的结果建立的条件码内容,如运算进位标志、运算结果溢出标志、运算结果为零标志、运算结果为负标志等。

(7) 微操作控制单元和时序部件。根据指令操作码和时序信号产生微操作控制信号,对各种操作实施时间上的控制。

3.4.2 CPU 的主要性能指标

(1) 字与字长。前面讲过,计算机内部作为一个整体参与运算、处理和传送的一串二进制数,称为一个字。在计算机中,许多数据是以字为单位进行处理的,是数据处理的基本单位。字长越长,运算能力就越强,计算精度就越高。

(2) 主频。CPU 有主频、倍频、外频三个重要参数,它们的关系:主频=外频 $\times$ 倍频,主频是 CPU 内部的工作频率,即 CPU 的时钟频率(CPU Clock Speed)。外频是系统总线的工作频率,倍频是它们相差的倍数。CPU 的运行速度通常用主频表示,以 Hz 作为计量单位。主频越高,CPU 的运算速度越快。

(3) 时钟频率。即 CPU 的外部时钟频率(外频),它由计算机主板提供,直接影响 CPU 与内存之间的数据交换速度。

(4) 地址总线宽度。地址总线宽度决定了 CPU 可以访问的物理地址空间,即 CPU 能够使用多大容量的内存。假设 CPU 有 n 条地址线,则其可以访问的物理地址为 2^n 个。

(5) 数据总线宽度。数据总线宽度决定了整个系统的数据流量的大小,数据总线宽度决定了 CPU 与二级高速缓存、内存以及输入/输出设备之间一次数据传输的信息量。

3.4.3 计算机的基本工作原理

计算机的基本工作原理包括存储程序和程序控制。计算机工作时先要把程序和所需数据送入计算机内存,然后存储起来,这就是“存储程序”的原理。运行时,计算机根据事先存储的程序指令,在程序的控制下由控制器周而复始地取出指令、分析指令、执行指令,直至完成全部操作,这就是“程序控制”的原理,计算机基本工作原理如图 3.5 所示。

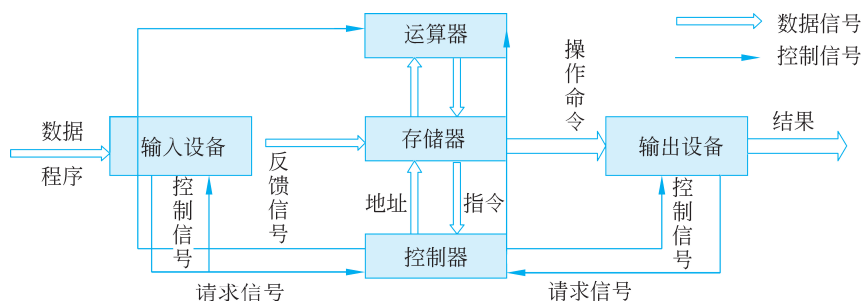


图 3.5 计算机的工作原理

1. 指令和指令系统

指令是指示计算机执行某种操作的命令,它由一串二进制数码组成。一条指令通常由两个部分组成:操作码+地址码。

(1) 操作码。操作码规定计算机完成什么样的操作,如算术运算、逻辑运算或输出数据等操作。

(2) 地址码。地址码是指明操作对象的内容或所在的存储单元地址,即指明操作对象是谁等信息。

一台计算机所能识别和执行的全部指令的集合称为这台计算机的指令系统。

指令按其完成的操作类型可分为数据传送指令(主机 $\longleftrightarrow$ 内存)、数据处理指令(算术和逻辑运算)、程序控制指令(顺序和跳转)、输入/输出指令(主机 $\longleftrightarrow$ I/O设备)和其他指令。

程序是由指令组成的有序集合。对一个计算机系统进行总体设计时,设计师必须根据要完成的总体功能设计一个指令系统。指令系统中包含许多指令。为了区别这些指令,每条指令用唯一的代码来表示其操作性质,这就是指令操作码。操作数表示指令所需要的数值或数值在内存中所存放的单元地址。

2. 计算机的工作过程

计算机的工作过程是计算机依次执行程序指令的过程。一条指令执行完毕后,控制器再取下一条指令执行,如此下去,直到程序执行完毕。计算机完成一条指令操作分为取指令、分析指令和执行指令三个阶段。

(1) 取指令。控制器根据程序计数器的内容(存放指令的内存单元地址)从内存中取出指令送到指令寄存器,同时修改程序计数器的值,使其指向下一条要执行的指令。

(2) 分析指令。对指令寄存器中的指令进行分析和译码。

(3) 执行指令。根据分析和译码实现本指令的操作功能。

思考与探索

计算机或计算系统可以被认为是由基本动作以及基本动作的各种组合所构成的。对这些基本动作的控制就是指令。指令的各种组合和数据组成了程序。指令和程序的思维是一种重要的计算思维。

3.5 存储器系统

现代计算机是以存储器为中心的,所有数据和程序都存放在存储器中,所以人们总是希望存储器的存储容量越大越好,存储速度越快越好,存储时间越长越好,存储器件的价格越低越好。然而,没有哪种存储器能够同时满足存储容量大、存储速度快、存储时间长、存储器件价格低的要求。为了适应系统需要,故采用不同性能的存储器件——内存储器和外存储器。为了满足应用需要,故将不同性能的存储器进行优化组合,构成一个存储系统,并由操作系统实行高效管理。

3.5.1 存储单位

存储单位用来表示存储容量的大小。计算机中所有的数据信息都是以二进制数的形式进行存储的,所以存储单位是指数据存放时占用的二进制位数,常用的存储单位有位、字节和字。

(1) 位(bit,b)。计算机中存储数据的最小单位,用来存放一位二进制数(0 或 1)。

一个二进制位只能表示 $2(2^1)$ 种状态,若要表示更多的信息,就得组合多个二进制位。

(2) 字节(Byte,B)。计算机中的一个存储单元(Memory Cell),ASCII 中的英文字母、阿拉伯数字、特殊符号和专用符号大约有 128~256 个,刚好可以用 8 个二进制位(1 字节)表示。

计算机中表示存储容量时通常用 KB、MB、GB、TB、PB、EB 等计量单位,换算关系如下:

$$1\text{KB}=1024\text{B}=2^{10}\text{B} \quad 1\text{MB}=1024\text{KB}=2^{20}\text{B} \quad 1\text{GB}=1024\text{MB}=2^{30}\text{B}$$

$$1\text{TB}=1024\text{GB}=2^{40}\text{B} \quad 1\text{PB}=1024\text{TB}=2^{50}\text{B} \quad 1\text{EB}=1024\text{PB}=2^{60}\text{B}$$

(3) 字(Word,W)。计算机在存储、传送或操作时,作为一个数据单位的一组二进制位称为一个计算机字,简称为“字”,每个字所包含的位数称为字长。一个字由若干字节组成,而字节是计算机进行数据处理和数据存储的基本单位,所以“字长”通常是“字节”的整数倍。

3.5.2 内存储器

内存储器(Internal Memory)是计算机中最主要的部件之一,用来存储计算机运行期间所需要的大量程序和数据。内存储器是直接和 CPU 相连并协同工作的存储器,包括只读存储器和随机存储器。随机存储器与 CPU 是计算机中最宝贵的硬件资源,是决定计算机性能的重要因素。内存储器主要由随机存储器(Random Access Memory, RAM)、只读存储器(Read Only Memory, ROM)和高速缓冲存储器(Cache)组成。

1. 随机存储器(RAM)

RAM 的作用是临时存放正在运行用户程序和数据及临时(从磁盘)调用的系统程序。其特点是 RAM 中的数据可以随机读出或者写入。关机或者停电时,其中的数据会丢失。

RAM 又可分为静态存储器(Static RAM, SRAM)和动态存储器(Dynamic RAM, DRAM)。

SRAM 的特点是工作速度快,只要电源不撤除,写入 SRAM 的信息就不会消失,不需要刷新电路,同时读取时不破坏原来存放的信息。信息一经写入可多次读出,但 SRAM 的

集成度较低功耗较大。SRAM 一般用来作为计算机中的高速缓冲存储器(Cache)。

DRAM 的优点是集成度较高,功耗也较低,其缺点是保存在 DRAM 中的信息,随着电容器的漏电会逐渐消失,一般信息保存时间为 2ms 左右,为了保存 DRAM 中的信息。必须每隔 1~2ms 对其刷新一次。因此采用 DRAM 的计算机必须配置动态刷新电路,以防信息丢失,DRAM 一般用作计算机中的主存储器,人们平常所说的内存就是 DRAM。

2. 只读存储器(ROM)

ROM 的作用是存放一些需要长期保留的程序和数据,如系统程序、控制时存放的控制程序等。其特点是只能读,一般不能改写,能长期保留其上的数据,即使断电也不会破坏。一般在系统主板上装有 ROM-BIOS,它是固化在 ROM 芯片中的系统引导程序,完成系统加电自检、引导和设置输入输出接口的任务。

3. 高速缓冲存储器(Cache)

CPU 在执行程序时,总是按指令地址或操作数地址访问主存的。理论上,主存的读、写速度与 CPU 的工作速度不仅越快越好,两者的速度还应该一致。然而,现代计算系统对 CPU 的性能要求越来越高,对主存容量要求越来越大。如果用同 CPU 一样性能的材料制作主存储器,则会使计算机的成本大幅度提高。因此,一般通用计算机的主存速度较 CPU 速度低,CPU 的时钟频率远远超过了主存的响应速度,故使得 CPU 的执行速度受主存的限制而不能充分发挥自身的快速作用,从而降低了计算机整体的运行速度。

为了协调主存与 CPU 速度上的差异,目前解决这个问题最有效办法是采用 Cache 技术。Cache 是一种在主存与 CPU 间起缓冲作用的存储器,所以称为高速缓冲存储器。Cache 是在 CPU 与主存两者之间增加一级在速度上与 CPU 相等,在功能上与主存相同的高速缓冲存储器,以实现在两个不同工作速度的部件之间,在交换信息的过程中起缓冲(协调)作用。

Cache 的容量一般为 512KB 左右,并嵌入 CPU。计算机开始运行后,将当前正要执行的一部分程序批量地从主存复制到 Cache,CPU 读取指令时,先到 Cache 中查找。若在 Cache 中找到,则直接从 Cache 中读取(称为命中),否则从主存中读取。

Cache 既提高了系统性能,又保持了造价低廉。现在一般微型机中都含有内部 Cache,否则其速度难以真正实现。

4. 内存的性能指标

(1) 存储容量。通常以 RAM 的存储容量来表示微型计算机的内存容量。常用单位有 KB、MB、GB 等。

(2) 存取周期。内存的存取周期是指存储器进行两次连续、独立的操作(存数的写操作和取数的读操作)之间所需要的最短时间,以 ns(纳秒)为单位,该值越小速度越快。常见的有 7ns、10ns、60ns 等。存储器的存取周期是衡量主存储器工作速度的重要指标。

(3) 功耗。它能反映存储器耗电量的大小,也反映了发热程度。功耗小,对存储器的工作稳定有利。

3.5.3 外存储器

外存储器(External Storage)是相对内存储器或主存储器而命名的,所以又称为辅助存