

第 1 部分

基 础 篇

1.1 计算机概述

计算机是 20 世纪人类最伟大的科学技术发明之一,对人类的生产和社会生活产生了极大的影响。计算机是一种能够根据程序指令对复杂任务进行自动、高速、精确处理的电子设备。通常所说的计算机主要是指电子计算机,它在人们的日常生活中几乎无处不在。现代电子计算机虽然只经历了短短的几十年,但是它却彻底改变了人类的生产和生活方式。

1.1.1 计算机发展历史

1946 年 2 月,美国宾夕法尼亚大学研制成了大型电子数字积分计算机(Electronic Numerical Integrator and Computer, ENIAC),它最初专门用于火炮弹道计算,后经多次改进成为能进行各种科学计算的通用计算机,如天气预报、原子核能、风洞试验设计等。ENIAC 是个宽度约为 1m,长度为 30.5m,总重量达 30t 的庞然大物,如图 1-1 所示,它每秒可以执行 5000 次加法或 400 次乘法运算,是手工计算速度的 20 万倍,只需要 3s 就可以完成此前需要 200 人手工计算两个月的弹道计算。ENIAC 是计算机发展史上的一个里程碑,也是公认的世界上第一台现代电子计算机。

伴随电子技术的发展,计算机所采用的元件经历了从电子管到晶体管,从分离元件到集成电路,再到高集成度的微处理器阶段。每次物理元件的变革都是一次新的突破,促使计算机性能出现新的飞跃。概括地说,自 1946 年以来,根据所采用的电子元件不同,可以将电子计算机的发展划分为四代。

第一代——电子管计算机(1946—1959 年)。第一代计算机的逻辑器件采用电子管,如图 1-2 所示。主存储器有水银延迟线存储器、阴极射线示波管、静电存储器等类型,内存储器(简称内存)大小仅几千字节,外存储器(简称外存)使用磁带、磁鼓、纸带、卡片等。运算速度为每秒几千次至几万次。第一代计算机没有系统软件,使用机器语言和汇编语言编程。这个时

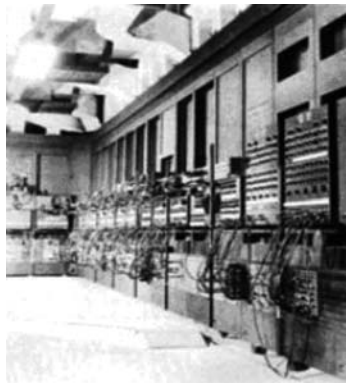


图 1-1 世界上第一台现代电子计算机 ENIAC

期的计算机主要用于科学计算,只被运用于少数尖端领域。第一代计算机的体积庞大、运算速度慢、存储容量小、可靠性低,但它们奠定了以后计算机技术发展的基础,对计算机的发展产生了深远的影响。

第二代——晶体管计算机(1959—1964 年)。第二代计算机的逻辑器件采用晶体管,如图 1-3 所示。主存储器均采用磁心存储器,内存容量扩大到几万字节,磁鼓和磁盘开始用作主要的辅助存储器,利用 I/O(输入输出)处理机进行输入输出处理。运算速度明显提高,每秒可以执行几万次到几十万次的加法运算。计算机中出现了操作系统,配置了子程序库和批处理管理程序,还出现了高级语言,如 Fortran、COBOL、ALGOL 等。计算机不仅继续大量用于科学计算,还被用于数据处理的工业过程控制。中小型计算机开始大量生产并逐渐被工商企业用于商务处理。与第一代计算机相比,第二代计算机的晶体管体积小、寿命长、发热小、功耗低、价格便宜,使得计算机电子线路的结构大有改观,存储容量大为增加,运算速度也得到大幅提高。



图 1-2 电子管



图 1-3 晶体管

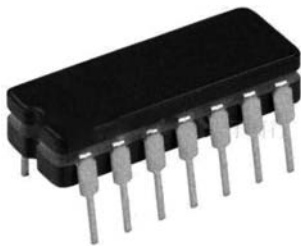


图 1-4 中小规模集成电路

第三代——中小规模集成电路计算机(1964—1970 年)。第三代计算机的逻辑器件采用中小规模集成电路,如图 1-4 所示。与晶体管电路相比,集成电路计算机的体积、重量、功耗都进一步减小,运算速度、逻辑运算功能和可靠性进一步提高。半导体存储器逐步取代了磁心存储器的主存储器地位,内存容量大幅度提高,磁盘成了不可缺少的辅助存储器,并且开始普遍采用虚拟存储技术。运算速度达到每秒几百万次。操作系统软件在规模和功能上发展很快,功能日趋成熟和完善;软件技术进一步提高,提出了结构化、模块化的程序设计思想,出现了结构化的程序设计语言 Pascal;软件开始形成产业,出现了大量面向用户的应用程序。第三代计算机的应用进入了更多的科学技术领域和工业生产领域。

第四代——大规模、超大规模集成电路计算机(1970 年至今)。20 世纪 70 年代以来,集成电路的集成度迅速从中、小规模发展到大规模、超大规模的水平,如图 1-5 所示。微处理器和微型计算机应运而生,各类计算机的性能迅速提高。金属氧化物半导体电路(Metal Oxide Silicon, MOS)的出现,使计算机的主存储器由半导体存储器完全替代了服役达 20 年之久的磁心存储器,主存储器的功能和可靠性进一步提高,存储容量向百兆字节、千兆字节发展;外

存储器除了软盘和硬盘外,还出现了光盘。运算速度向每秒十万亿次、每秒百万亿次及更高速度发展。这个时期,操作系统不断地完善,应用软件成为现代工业中的一个重要产业,计算机的发展进入网络时代。

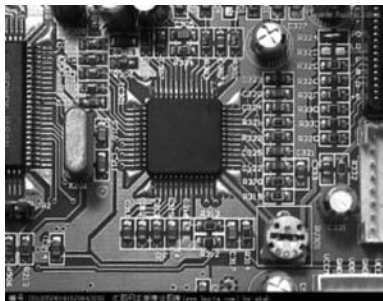


图 1-5 大规模及超大规模集成电路

计算机在提高性能、降低成本、普及和深化应用等方面的发展趋势仍在继续,节奏进一步加快,学术界和工业界不再沿用“第 X 代计算机”的说法。人们正在研究开发的计算机系统,主要着力于智能化,它以知识处理为核心,可以模拟或部分替代人的智能活动,具有自然的人机通信能力。当然,这是一个需要持续努力才能逐步实现的目标。

1.1.2 计算机分类

计算机及相关技术的迅速发展带动计算机的类型也不断分化,形成了各种不同种类的计算机。计算机按结构原理可分为模拟计算机、数字计算机和混合式计算机;按用途可分为专用计算机和通用计算机。较为普遍的一种分法是按照计算机的运算速度、字长、存储容量等综合性能指标,将计算机分为巨型机、大型机、小型机、微型机和嵌入式计算机等。

1. 巨型机

巨型机是一种超大型电子计算机,具有很强的计算和处理数据的能力,其主要特点表现为高速度和大容量,配有多种外围设备及丰富的、高性能的软件系统,如图 1-6 所示。巨型计算机实际上是一个巨大的计算机系统,主要用来承担重大的科学研究、国防尖端技术和国民经济领域的大型计算课题及数据处理任务。例如,大范围天气预报、卫星照片整理、原子核物理探索、洲际导弹、宇宙飞船研究等。



图 1-6 巨型机

目前,我国最快的超级计算机——神威太湖之光,安装在无锡国家超级计算中心。线性系统软件包基准测试测得其运行速度达到每秒 93 千万亿次浮点运算。神威太湖之光拥有 10 649 600 个计算核心,包括 40 960 个节点,速度是“天河二号”的 2 倍,效率是其 3 倍。

2. 大型机

大型机也称为大型主机。大型机使用专用的处理器指令集、操作系统和应用软件。大型机最初是指装在非常大的带框铁盒子里的大型计算机系统,以用来与小一些的迷你机和微型机相区别,如图 1-7 所示。

大型机和巨型机的主要区别如下。

(1) 大型机使用专用的指令系统和操作系统;巨型机使用通用处理器及 UNIX 或类 UNIX 操作系统(如 Linux)。

(2) 大型机擅长非数值计算(数据处理);巨型机擅长数值计算(科学计算)。

(3) 大型机主要用于商业领域,如银行和电信;而巨型机用于尖端科学领域,特别是国防领域。

(4) 大型机大量使用冗余等技术确保其安全性及稳定性,所以内部结构通常有两套;而巨型机使用大量处理器,通常由多个机柜组成。

(5) 为了确保兼容性,大型机的部分技术较为保守。

3. 小型机

小型机是指采用精简指令集处理器,性能和价格介于 PC 服务器和大型机之间的一种高性能计算机,如图 1-8 所示。在中国,小型机习惯上是指 UNIX 服务器。



图 1-7 大型机



图 1-8 小型机

UNIX 服务器具有区别于 x86 服务器和大型机的特有体系结构,基本上各厂家的 UNIX 服务器均使用自家 UNIX 版本的操作系统和专属处理器。使用小型机的用户一般是看 UNIX 操作系统和专用服务器的安全性、可靠性、纵向扩展性以及高并发访问下的出色处理能力。

目前,生产 UNIX 服务器的厂商主要有 IBM、HP、浪潮、富士通和甲骨文(收购了 SUN 公司)。典型机器如 IBM 公司曾经生产的 RS/6000、HP 公司的 Superdome、浪潮公司的天梭 K1950 等。

4. 微型机

微型机又称微型计算机、微机、微电脑。微型机是由大规模集成电路组成的体积较小的电子计算机。它是以微处理器为基础,配以内存储器及输入输出接口电路和相应的辅助电路而构成的裸机。

微型机的特点是体积小、灵活性大、价格便宜、使用方便。自 1981 年美国 IBM 公司推出第一代微型机(IBM-PC)以来,微型机以其执行结果精确、处理速度快、性价比高、轻便小巧等

特点迅速进入社会各个领域,且技术不断更新,产品快速换代,从单纯的计算工具发展成为能够处理数字、符号、文字、语言、图形、图像、音频、视频等多种信息的强大多媒体工具。如今的微型机产品无论从运算速度、多媒体功能、软硬件支持,还是易用性等方面都比早期产品有了很大飞跃。

微型机主要包括台式机、电脑一体机、笔记本电脑、平板电脑和智能手机等。

1) 台式机

台式机也称桌面机,是一种主机、显示器等设备相对独立的计算机,相较于笔记本电脑,其体积较大,一般需要放置在电脑桌或者专门的工作台上,因此命名为台式机。台式机的性能相对来说比笔记本电脑强。

台式机一般具有如下特点。

(1) 散热性。台式机具有笔记本所无法比拟的优点,台式机的机箱因空间大、通风条件好等因素而一直被人们广泛使用。

(2) 扩展性。台式机的机箱方便用户进行硬件升级,如台式机机箱的硬盘驱动器插槽是4~5个,非常方便日后的硬件升级。

(3) 保护性。台式机全方位保护硬件不受灰尘的侵害。

(4) 明确性。台式机机箱的开关键、重启键、USB口、音频接口都在机箱前置面板上,使用方便。

2) 电脑一体机

电脑一体机是由一台显示器、一个键盘和一个鼠标组成的计算机,如图1-9所示。它的芯片、主板与显示器集成在一起,显示器就是一台电脑,因此只要将键盘和鼠标连接到显示器上,机器就能使用。随着无线技术的发展,电脑一体机的键盘、鼠标与显示器可实现无线连接,机器只有一根电源线。这就解决了一直为人诟病的台式机线缆多而杂的问题。有的电脑一体机还具有电视接收、AV功能。

3) 笔记本电脑

笔记本电脑也称手提电脑或膝上型电脑,是一种小型、可携带的个人计算机,通常重1~3kg。它和台式机架构类似,但是提供了更好的便携性,如图1-10所示。



图 1-9 电脑一体机



图 1-10 笔记本电脑

笔记本电脑大体上可以分为6类:商务型、时尚型、多媒体应用型、上网型、学习型、特殊用途型。商务型笔记本电脑一般移动性强、电池续航时间长、商务软件多;时尚型笔记本电脑外观时尚,主要针对时尚女性;多媒体应用型笔记本电脑则有较强的图形、图像处理能力和多媒体能力,尤其是播放能力,为享受型产品;上网型笔记本电脑即上网本,轻便、配置低,具备上网、收发邮件以及即时信息(Instant Messaging, IM)等功能,可以实现流畅播放流媒体和音

乐；学习型笔记本电脑机身设计为笔记本外形，全面整合学习机、电子词典、复读机、学生电脑等多种机器的功能；特殊用途型笔记本电脑服务于专业人士，可以在酷暑、严寒、低气压、战争等恶劣环境下使用，如奥运会前期在“华硕珠峰大本营 IT 服务区”使用的华硕笔记本电脑。

4) 平板电脑

平板电脑是一款无须翻盖、没有键盘、大小不等、形状各异但功能完整的计算机。其构成组件与笔记本电脑基本相同，但它是利用触控笔或数字笔在屏幕上书写，而不是使用键盘和鼠标输入，并且打破了笔记本电脑键盘与屏幕垂直的 L 型设计模式，如图 1-11 所示。它除了拥有笔记本电脑的所有功能外，还支持手写输入或语音输入，移动性和便携性更胜一筹。

5) 智能手机

智能手机是由掌上电脑演变而来的。最早的掌上电脑并不具备手机通话功能，但是随着用户对掌上电脑的个人信息处理方面功能的依赖，厂商将掌上电脑的系统移植到手机中，于是出现了智能手机。智能手机比传统手机具有更多的综合性处理能力。

智能手机同传统手机的外观和操作方式类似，但是传统手机使用的都是生产厂商自行开发的封闭式操作系统，所实现的功能非常有限，不具备智能手机的扩展性。智能手机可以像计算机那样随意安装和卸载应用软件，具有独立的操作系统、独立的运行空间，用户可以自行安装软件、游戏、导航等第三方服务商提供的程序，并可以通过移动通信网络来实现无线网络接入，如图 1-12 所示。



图 1-11 平板电脑



图 1-12 智能手机

智能手机作为一种新型的移动终端，也可以归入微型机一类，但是由于手机要求体积非常小，便于携带，因此它与普通计算机在硬件设计上有很大的不同。

5. 嵌入式计算机



图 1-13 嵌入式计算机

嵌入式技术是针对某个特定的应用，如针对网络、通信、音频、视频、工业控制等的“专用”计算机技术。嵌入式计算机一般由嵌入式微处理器、外围硬件设备、嵌入式操作系统以及用户的应用程序四部分组成，如图 1-13 所示。

嵌入式计算机在应用数量上远远超过了各种通用计算机，制造工业、过程控制、通信、仪器、仪表、汽车、船舶、航空、航天、军事装备、消费类产品等均是嵌入式计算机的应用领域。



1.2 微电子技术

1.2.1 集成电路

集成电路(Integrated Circuit, IC)又称微电路(Microcircuit)、微芯片(Microchip)、芯片(Chip),是把一定数量的常用电子元件,如电阻、电容、晶体管等,以及这些元件之间的连线,通过半导体工艺集成在一起的具有特定功能的电路。

单块芯片上所容纳的元件数目称为集成度。一般来说,集成度越高,性能越强大。集成电路按集成度高低的不同可分为如下6种。

- (1) 小规模集成电路(Small Scale Integrated circuits,SSI): 集成度 <100 。
- (2) 中规模集成电路(Medium Scale Integrated circuits,MSI): $100 < \text{集成度} < 1000$ 。
- (3) 大规模集成电路(Large Scale Integrated circuits,LSI): $1000 < \text{集成度} < 10 \text{ 万}$ 。
- (4) 超大规模集成电路(Very Large Scale Integrated circuits,VLSI): $10 \text{ 万} < \text{集成度} < 100 \text{ 万}$ 。
- (5) 特大规模集成电路(Ultra Large Scale Integrated circuits,ULSI): $100 \text{ 万} < \text{集成度} < 1 \text{ 亿}$ 。
- (6) 极大规模集成电路(Giga Scale Integration,GSI): 集成度 $>1 \text{ 亿}$ 。

不过需要注意的是,对于超大规模以上的集成电路,有时候人们不那么严格地区分超大、特大和极大规模的区别,而是笼统地称为超大规模集成电路。

第四代计算机中的主要部件几乎都和集成电路有关,如CPU、显卡、主板、内存、声卡、网卡、光驱等,并且最新技术把越来越多的元件集成到一块电路板上,使计算机拥有了更多功能,在此基础上产生了许多新型计算机,如掌上电脑、指纹识别计算机、声控计算机等。

集成电路在通信中的应用也非常广泛,如通信卫星、手机、雷达等,尤其是我国自主研发的北斗导航系统就是其中典型的例子。北斗导航系统是我国具有自主知识产权的卫星定位系统,与美国的GPS、俄罗斯的格洛纳斯、欧盟的伽利略系统并称为全球四大卫星导航系统。北斗导航系统的研究成功,打破了卫星定位导航应用市场由国外GPS垄断的局面。

除此之外,集成电路技术在日常生活中的其他各个领域都有广泛应用。如在汽车上,微控制器、功率半导体器件、电源管理器件、LED驱动器等汽车集成电路器件的应用使得汽车能够处于最佳工作状态;再如在热能动力工程领域中的应用,最简单的莫过于温控计;当然,火电厂中的信息管理系统也离不开集成电路技术。总之,集成电路技术的发展使人们的生活越来越美好,越来越便利。

1.2.2 摩尔定律

摩尔定律是由Intel(英特尔)创始人之一戈登·摩尔(Gordon Moore)提出来的。其内容为:当价格不变时,集成电路上可容纳的元件的数目每隔18~24个月便会增加一倍,性能也将提升一倍。这种指数级的增长,促使20世纪70年代的大型家庭计算机转化成20世纪80年代、90年代更先进的机器,然后又孕育出了高速度的互联网、智能手机,以及现在的车联网、智能冰箱和自动调温器等。

摩尔定律可以说是整个计算机行业最重要的定律,它其实是一个预言,这个看起来自然而然的进程,实际很大程度上也是人类有意控制的结果。芯片制造商有意按照摩尔定律预测的

轨迹发展,软件开发商的新软件产品也日益挑战现有设备的芯片处理能力,消费者需要更新配置更高的设备,设备制造商赶忙去生产可以满足处理要求的下一代芯片……

20 世纪 90 年代以来,半导体行业每两年就会发布一份行业研发规划蓝图,协调成百上千家芯片制造商、供应商跟着摩尔定律走。这份规划蓝图使整个计算机行业跟着摩尔定律按部就班地发展。

但是现在,这种发展轨迹可能要告一段落了。由于同样小的空间里集成越来越多的硅电路,产生的热量也越来越大,这种原本两年处理能力加倍的速度已经慢慢下滑。此外,还逐渐出现了更多、更大的问题。5nm 工艺是硅芯片的一个技术分水岭,因为电子的行为受限于量子的不确定性,晶体管将变得不可靠,并且芯片制造的核心设备光刻机需要从深紫外光(Deep Ultraviolet Light, DUV)技术向极紫外光(Extreme Ultraviolet Light, EUV)技术发展,目前全球仅有荷兰阿斯麦尔(ASML)公司掌握该项技术。EUV 光刻机是一种高尖端的机器,里面有 80 000 多个精密零部件,许多零件非常复杂,甚至有些关键部件在全球只有一两家公司能够制造。虽然中国有着强大的制造业,且模仿能力超强,但是国内配套零部件的生产在短期内还无法追上,技术经验和人才储备更是远远不足,这些成为我国芯片制造发展的阻碍。但是,中国对芯片半导体的自主研发突击已经开始,中国的技术实力比当初要强得多,全面突破只是时间问题。



1.3 计算机的组成与工作原理

1.3.1 冯·诺依曼体系结构

冯·诺依曼(1912—1957 年),布达佩斯大学数学博士,美籍匈牙利数学家,在 ENIAC 的研制中期,冯·诺依曼参与了原子弹的研制工作,他带着原子弹研制过程中遇到的大量计算问题加入计算机的研制工作中。

ENIAC 是世界上第一台现代电子计算机,但是 ENIAC 有两个致命的缺陷:一是采用十进制运算,逻辑元件多,结构复杂,可靠性低;二是没有内部存储器,操纵运算的指令分散存储在许多电路部件内,这些运算部件如同一副积木,解题时必须像搭积木一样用人工把大量运算部件搭配成各种解题的布局,每算一题都要搭配一次,非常麻烦且费时。

1945 年 6 月底,冯·诺依曼执笔写出了 EDVAC(Electronic Discrete Variable Automatic Computer,离散变量自动电子计算机)计划草案,提出了在计算机中采用二进制算法和设置内存存储器的理论,并明确规定了电子计算机必须由运算器、控制器、存储器、输入设备和输出设备五大部分组成基本结构形式。他认为,计算机采用二进制算法和内存存储器后,指令和数据便可以一起存放在存储器中,可以使计算机的结构大大简化,并且为实现运算控制自动化和提高运算速度提供了良好的条件。

EDVAC(图 1-14)于 1952 年建成,它的运算速度与 ENIAC 相似,而使用的电子管却只有 5900 多个,比 ENIAC 少得多。EDVAC 的诞生,使计算机技术出现了一个新的飞跃。EDVAC 是世界上第一台采用冯·诺依曼体系结构的通用计算机,它奠定了现代电子计算机的基本结构,标志着电子计算机时代的真正开始。

从第一台通用计算机诞生到今天已经过去将近 70 年,计算机的技术与性能也都发生了巨大的变化,但整个主流体系结构依然是冯·诺依曼体系结构。由于冯·诺依曼对现代计算机

技术的突出贡献,因此他被称为“计算机之父”。

冯·诺依曼的主要贡献是他提出了“存储程序控制”的工作原理。该原理的要点是:程序由二进制指令构成,所有指令都以操作码和地址码的形式存放在存储器中,以运算器和控制器为中心,顺序执行指令所规定的操作。冯·诺依曼设计思想可以简要概括为以下4点。

(1) 计算机应包括运算器、存储器、控制器3个核心部件,以及输入设备和输出设备。输入设备负责把人工编制的指令以及需要处理的数据输入存储器中;输出设备负责把存储器里的计算结果输出(显示)。

(2) 计算机的数制采用二进制。

(3) 程序的每条指令一般具有一个操作码和一个地址码。操作码表示运算性质,如加法或者除法;地址码指出操作数在存储器中的位置。

(4) 将编好的程序和原始数据送入存储器,然后启动计算机工作。计算机可以在不需要操作人员干预的情况下,自动逐条取出指令和执行指令,并最终完成整个任务。



图 1-14 第一台通用计算机 EDVAC

1.3.2 五大基本部件

冯·诺依曼机的硬件系统主要由五大基本部分组成:运算器、控制器、存储器、输入设备和输出设备。这五大部分通过系统总线完成指令所传达的操作,当计算机接受指令后,由控制器指挥,将数据从输入设备传送到存储器中存放,再由控制器将需要参加运算的数据传送到运算器中,由运算器进行处理,处理后的结果由输出设备输出。

下面简要介绍计算机的五大基本部分。

1. 运算器

运算器又称算术逻辑部件(Arithmetic and Logic Unit, ALU)。运算器的主要任务是执行各种算术运算和逻辑运算。算术运算是指各种数值运算,如加、减、乘、除等。逻辑运算是进行逻辑判断的非数值运算,如与、或、非、比较、移位等。计算机所完成的全部运算都是在运算器中进行的。根据指令规定的寻址方式,运算器从存储器或寄存器中取得操作数,进行计算后,送回指令所指定的寄存器中。运算器的核心部件是加法器和若干寄存器,加法器用于运算,寄存器用于存储参加运算的各种数据以及运算后的结果。

2. 控制器

控制器对输入的指令进行分析,并统一控制计算机的各个部分完成一定任务的部件。它一般由指令寄存器、状态寄存器、指令译码器、时序电路和控制电路组成。计算机的工作方式是执行程序,程序就是为完成某一任务所编制的特定的指令序列,各种指令按一定的时间关系有序安排,控制器产生各种最基本的不可再分的微操作命令信号,即微命令,以指挥整个计算机有条不紊地工作。当计算机执行程序时,控制器首先从程序计数器中取得指令的地址,并将下一条指令的地址存入程序计数器中,然后从存储器中取出指令,由指令译码器对指令进行译码后产生控制信号,用以驱动相应的硬件完成指令操作。简言之,控制器就是协调指挥计算机各部分工作的部件,它的基本任务就是根据指令的要求,综合有关逻辑条件与时间条件产生相

应的微命令。

运算器和控制器是计算机的核心部件,现代计算机通常把运算器、控制器和若干寄存器集中在一块芯片上,这块芯片称为中央处理器(Central Processing Unit,CPU)。微型计算机的CPU又称为微处理器。计算机以CPU为中心,输入设备和输出设备与存储器之间的数据传输和处理都通过CPU来控制执行。

3. 存储器

存储器由大量的记忆单元组成,记忆单元是一种具有两个稳定状态的物理器件,可用来表示二进制的0和1,这种物理器件一般由半导体器件或磁性材料等构成。存储器分为内存(简称内存或主存)、外存储器(简称外存或辅存)和缓冲存储器(简称缓存)。

内存一般由半导体存储器构成,通常装在主板上,主要用来存放计算机正在执行的或经常使用的程序和数据。CPU可以直接访问内存,执行程序时就是从内存中读取指令,并且在内存中存取数据的。内存的特点是存取速度快,但容量有限,大小受到地址总线位数的限制。

外存储器用来存放不经常使用的程序和数据,CPU不能直接访问它。外存储器属于计算机的外围设备,是为弥补内存容量不足而配置的。它的特点是容量大、成本低,但存取速度慢,通常使用DMA(Direct Memory Access)技术和IOP(I/O Processor)技术来实现内存和外存储器之间的数据直接传送。

缓冲存储器位于内存与CPU之间,其存取速度非常快,但存储容量更小,一般用来解决存取速度与存储容量之间的矛盾,以提高整个系统的运行速度。

在现代计算机中,存储系统是一个具有不同容量、不同访问速度的存储设备的层次结构。整个存储系统中包括CPU寄存器、缓冲存储器(内部Cache和外部Cache)、内存、辅助存储器(辅助存储器和大容量辅助存储器),如图1-15所示。在存储系统的层次结构中,层次越高,速度越快,但是价格越高;而层次越低,速度越慢,同时价格越低。这样就能做到在性能和价格之间的一个很好的平衡。

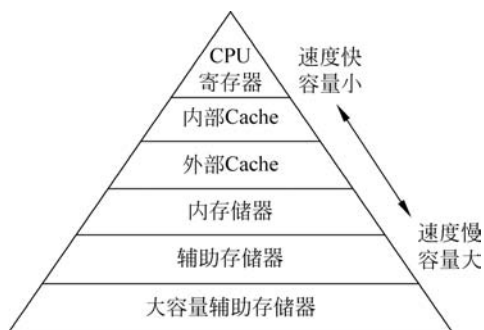


图 1-15 存储系统的层次结构

4. 输入设备

输入设备用来接受用户输入的原始数据和程序,并将它们变为计算机能识别的二进制信息存入内存中。常用的输入设备有键盘、鼠标、扫描仪、光笔等。

5. 输出设备

输出设备用于将内存中的由计算机处理的结果转变为人们能接受的形式输出。常用

的输出设备有显示器、打印机、绘图仪等。

1.3.3 计算机的工作原理

1. 计算机的基本工作原理

计算机的基本工作原理是存储程序和程序控制。程序与数据都存储在内存存储器中,CPU按照程序编排的顺序,一步一步地取出指令,自动完成指令规定的操作,这是计算机的基本工作原理,如图 1-16 所示。

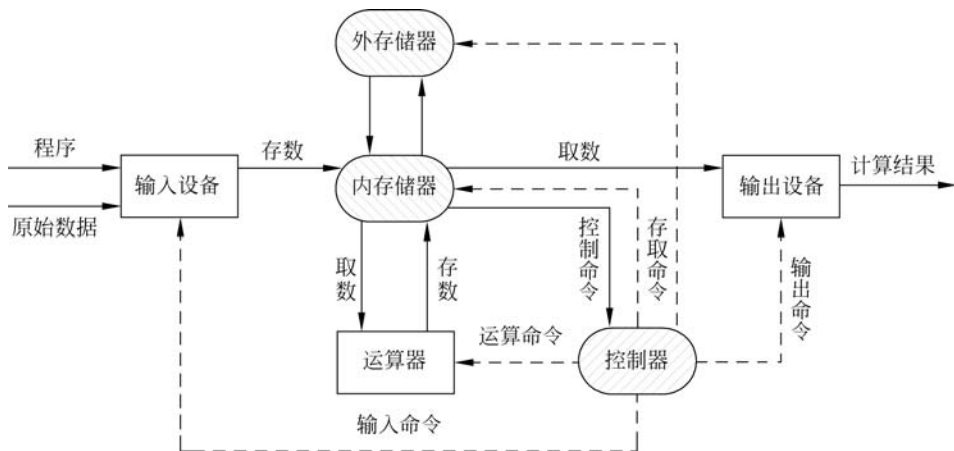


图 1-16 计算机的基本工作原理

具体描述如下。

- (1) 将程序和原始数据通过输入设备送入存储器。
- (2) 启动运行后,计算机从存储器中取出程序指令送到控制器去识别,分析该指令要做什么事情。
- (3) 控制器根据指令的含义发出相应的命令(如加法、减法),将存储单元中存放的操作数据取出送往运算器进行运算,再把运算结果送回存储器指定的单元。
- (4) 当运算任务完成后,就可以根据指令将结果通过输出设备输出。

因此,计算机的工作过程实际上就是快速执行指令的过程。指令执行时,必须先装入计算机内存存储器,CPU 负责从内存存储器中逐条取出指令,并对指令分析译码,判断该条指令要完成的操作,向各部件发出完成操作的控制信号,从而完成一条指令的执行。总之,计算机的基本工作过程就是不断地重复取指令、分析指令及取数、执行指令等过程,如此周而复始,直到遇到停机指令或外来事件的干预为止。

在计算机执行指令过程中有两种信息在流动:数据流和控制流。数据流包括原始数据、中间结果、结果数据和源程序等,这些信息从存储器读入运算器进行运算,所得的计算结果再存入存储器或传送到输出设备。控制流是由控制器对指令进行分析、解释后向各部件发出的控制命令,指挥各部件协调地工作。

2. 指令及指令系统

计算机工作的过程就是执行程序的过程。为了解决某一问题,程序设计人员将一条条指令进行有序的排列,只要在计算机上执行这一指令序列,便可完成预定的任务。因此,程序是

一系列有序指令的集合,计算机执行程序就是执行这一系列的有序指令。

指令是一种能被计算机识别并执行的二进制代码,它规定了计算机能完成的某一种操作。一条指令通常由操作码和操作数两部分组成。

(1) 操作码:指明该指令要完成的操作类型或性质,如加、减、取数或输出数据等。

(2) 操作数:指明操作对象的内容或所在的单元地址,大多数情况下操作数是地址码。

通常一台计算机有许多条作用不同的指令,所有指令的集合称为该计算机的指令系统。一般来说,无论是哪一种类型的计算机,都具有表 1-1 所示的指令。

表 1-1 常用指令

指 令	说 明
数据传送型指令	实现主存和寄存器之间,或寄存器和寄存器之间的数据传送
数据处理型指令	主要用于定点数或浮点数的算术运算和逻辑运算
程序控制型指令	主要用于控制程序的流向
输入输出型指令	用于主机与外围设备之间交换信息
其他指令	除以上各类指令外,较少被用到的一些指令包括字符串操作指令、堆栈指令、停机指令等

不同种类的计算机,其指令系统的指令数目与格式也不相同。CPU 的指令系统反映了计算机对数据处理的能力。由于每种 CPU 都有自己独特的指令系统,因此在某类计算机上可以执行的机器语言程序难以在其他不同类型的计算机上使用,这是由于不同类型的 CPU 采用的指令相互不兼容。

通常,同一 CPU 生产厂家在开发新的 CPU 产品时,既会设计增加一些高效的新指令,又同时“向下兼容”,使新的处理器可以正确执行老处理器中的所有指令。“向下兼容”的开发方式使用户在升级计算机硬件时不必担心原有的软件会被作废,但这也使得采用“向下兼容”方式开发的 CPU 指令系统越来越庞大和越来越复杂。

根据指令系统设计架构的不同,产生了复杂指令系统计算机(Complex Instruction Set Computing,CISC)和精简指令系统计算机(Reduced Instruction Set Computer,RISC)。

3. 指令的执行过程

按照存储程序的原理,计算机在执行程序时必须先将要执行的相关程序和数据放入内存存储器中,在执行程序时 CPU 根据当前程序指令寄存器的内容取出指令并执行指令,然后再取出下一条指令并执行,如此循环下去,直到程序结束指令时才停止执行。整个工作过程就是不断地取指令和执行指令的过程,最后将计算的结果放入指令指定的存储器地址中。指令执行过程中所涉及的部件主要有程序计数器、指令寄存器、指令译码器、通用寄存器和运算器等,如图 1-17 所示。

一条指令的执行过程按时间顺序可分为以下 4 个步骤。

(1) 取指令。当某个程序开始执行时,控制器根据程序计数器中的内容,向内存存储器的相应存储单元发出读请求,内存存储器将相应存储单元的指令读取后,通过总线送到指令寄存器中。

(2) 分析指令及取操作数。取出指令后,机器立即进入分析指令及取数阶段,指令译码器可识别和区分不同的指令类型及各种获取操作数的方法。指令译码器根据指令的内容分析出对应的操作类型,并产生相应的控制电信号。如果当前指令中的操作数需要从通用寄存器或

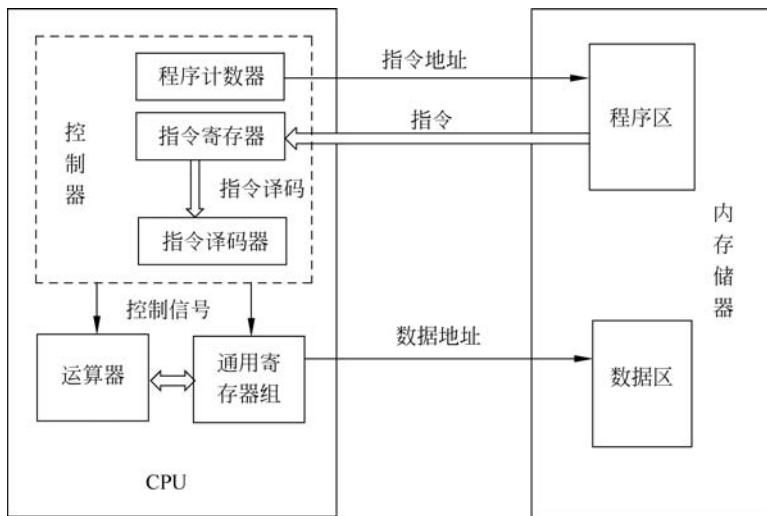


图 1-17 与指令执行有关的 CPU 部件

内存存储器获取,则控制器将先向相关部件发送读数据的请求,取到操作数后,再向相关部件发送完成指令操作相关的控制电信号。由于各种指令功能不同,寻址方式也不同,因此分析指令及取数阶段的操作是不同的,甚至会有很大的区别。

(3) 指令执行。由控制器发出完成该操作所需要的一系列控制信息,相关部件根据控制信号,完成当前指令所要求的操作。

(4) 写回数据及转下条指令。当前指令操作完成后,可能会有运算结果。控制器根据指令中操作结果的存放位置(通用寄存器或内存存储器),向相关部件发送“写数据”的请求,写回结果数据。一条指令执行完毕后,程序计数器加 1 或将转移地址码送入程序计数器,然后回到步骤(1),开始执行下一条指令。



1.4 PC 的组件

人们经常使用的台式机,简单地从外观上看,其硬件包括两部分:主机系统和外围设备(简称外设)。主机是指安装在 PC 机箱内部的一个整体,包括主板、硬盘、光驱、电源和风扇等。主板上安装了 CPU、内存、总线和 I/O 控制器等。

1.4.1 主板

主板(Motherboard 或 Mainboard)又称主机板、系统板、逻辑板、母板或底板等,是构成复杂电子系统(如电子计算机)的中心或者主电路板。

1. 主板概述

主板安装在机箱内,是微型机的最基本也是最重要的部件之一。主板的性能影响着整个微型机系统的性能,在整个微型机系统中扮演着举足轻重的角色。可以说,主板的类型和档次决定整个微型机系统的类型和档次。

主板一般为矩形电路板,能提供一系列接合点,供处理器、显卡、声卡、硬盘、存储器、外部设备等部件连接,如图 1-18 所示。

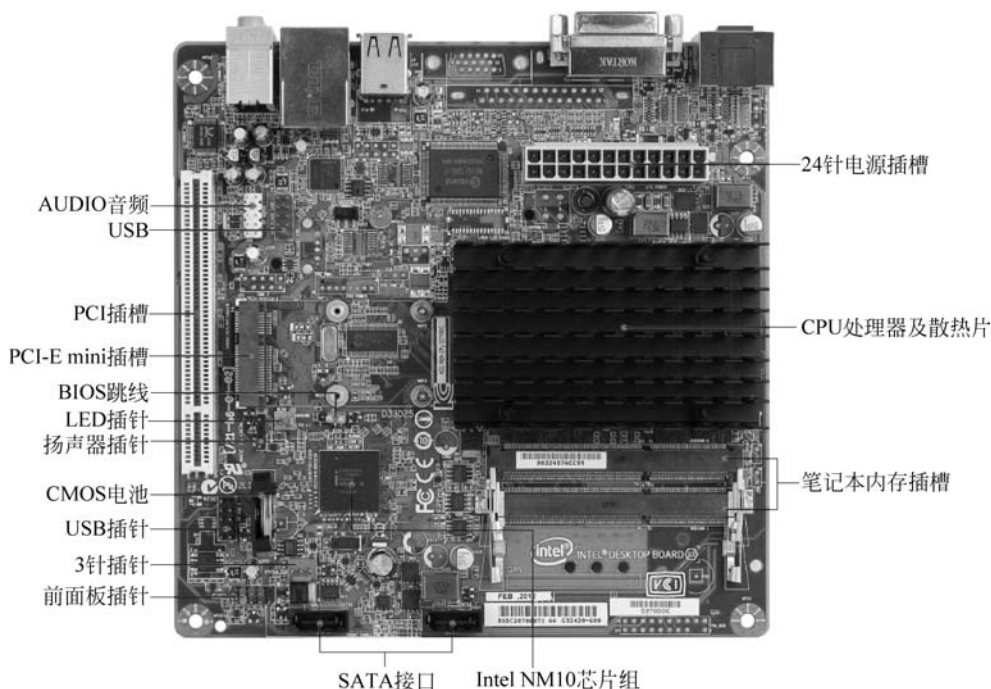


图 1-18 主板

主板采用开放式结构,一般提供 6~15 个扩展插槽,供 PC 外围设备的控制卡(适配器)插接。通过更换这些插卡,可以对微型机的相应子系统进行局部升级,使厂家和用户在配置机型方面有更大的灵活性。

2. 主板的重要芯片

主板功能的实现,很大程度上依赖于主板上各类芯片的作用。面对主板上密密麻麻的芯片时,大家经常会感到一阵阵的疑惑。这些芯片都是用来干什么的?彼此之间有什么区别?

1) 芯片组

芯片组(Chipset)是主板的核心组成部分,几乎决定了这块主板的功能,进而影响整个计算机系统性能的发挥。芯片组性能的优劣,决定了主板性能的好坏与级别的高低。芯片组通常由北桥和南桥组成,也有些以单片设计,增强其性能。

北桥芯片又称为主桥(Host Bridge),在计算机中起着主导作用。一般来说,芯片组的名称是以北桥芯片的名称来命名的。北桥芯片负责与 CPU 的联系并控制内存储器、PCI-E 数据在北桥内部传输,提供对 CPU 的类型和主频、系统的前端总线频率、内存储器的类型和最大容量、AGP/PCI-E 插槽、ECC 纠错等支持,整合型芯片组的北桥芯片还集成了显示核心。

北桥芯片是主板上离 CPU 最近的芯片,这主要是考虑到北桥芯片与处理器之间的通信最密切,为了提高通信性能而缩短传输距离。北桥芯片的数据处理量非常大,发热量也越来越大,因此北桥芯片都覆盖着散热片,有些主板的北桥芯片还会配合风扇进行散热。

南桥芯片负责 I/O 总线之间的通信,如 PCI 总线、USB、LAN、ATA 总线、SATA、音频控制器、键盘控制器、实时时钟控制器、高级电源管理等,这些技术一般相对来说比较稳定,所以不同芯片组中可能南桥芯片是一样的,不同的只是北桥芯片。

南桥芯片一般位于主板上离 CPU 插槽较远的下方、PCI 插槽的附近,这种布局是考虑到

它所连接的 I/O 总线较多,离处理器远一点有利于布线。相对于北桥芯片来说,其数据处理量并不算大,所以南桥芯片一般都没有覆盖散热片。南桥芯片不与处理器直接相连,而是通过一定的方式与北桥芯片相连,如图 1-19 所示。

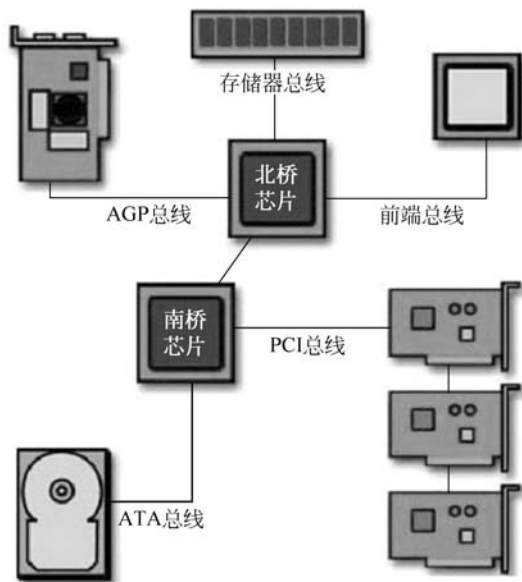


图 1-19 芯片组连接示意图

值得注意的是,近年来,随着处理器(CPU)的集成度越来越高,北桥芯片的大部分功能都已经整合进了处理器内,北桥芯片基本取消,剩余的功能都由南桥芯片提供。这样一来,主板上只剩一个南桥芯片,但是习惯上仍然叫作芯片组。

2) BIOS 芯片

BIOS(Basic Input Output System,基本输入输出系统)芯片是主板上的一块长方形或正方形芯片,一般是一块 32 针的双列直插式集成电路,上面印有 BIOS 字样。既然 BIOS 称为系统,那它就不只是一个简单的软件或一个硬件设备,而是软硬件结合在一起,把一组重要程序固化在主板上的一块 ROM 芯片中,人们把这种硬件化的软件称为固件。

早期 BIOS 使用的 ROM 都是在工厂里用特殊的方法把内容烧录进去的,用户只能读取而不能修改其中的内容。从奔腾机时代开始,主板一般都使用 Flash ROM 作为 BIOS 的存储芯片,能通过特定的写入程序实现 BIOS 的升级。BIOS 中主要包括 4 种程序。

(1) 加电自检程序。计算机接通电源后,系统将有一个对内部各个设备进行检查的过程,这是由一个称为 POST(Power On Self Test)的程序来完成的。完整的 POST 自检包括 CPU、640KB 基本内存、1MB 以上的扩展内存、ROM、主板、CMOS 存储器、串并口、显卡、软硬盘子系统及键盘测试。POST 自检中若发现问题,系统将给出提示信息或蜂鸣警告。

(2) 系统启动自检程序。当系统完成 POST 自检后,ROM BIOS 就按照系统 CMOS 设置中保存的启动顺序搜索软硬盘驱动器及 CD-ROM、U 盘、网络服务器等有效的启动驱动器,读入操作系统引导记录,然后将系统控制权交给引导记录,并由引导记录来完成系统的顺序启动。

(3) CMOS 设置程序。CMOS 设置程序只在开机时才可以进行设置。一般在计算机启动时按 F2 键或者 Delete 键进入设置,一些特殊机型按 F1 键、Esc 键、F12 键等进行设置。

CMOS 设置程序主要对计算机的基本输入输出系统进行管理和设置,使系统运行在最好状态下。使用 CMOS 设置程序还可以排除系统故障或者诊断系统问题。

(4) 主要 I/O 设备的驱动程序和中断服务程序。操作系统对软硬盘、光驱与键盘、显示器等外围设备的管理是建立在 BIOS 的基础之上的。基本输入输出的程序决定了主板是否支持某种 I/O 设备,如果 BIOS 中不包含某种 I/O 设备的驱动程序,则系统不支持此 I/O 设备。BIOS 中断服务程序是计算机系统软硬件之间的一个可编程接口,用于程序软件功能与计算机硬件实现的衔接。程序员可以通过对 INT 5、INT 13 等中断的访问直接调用 BIOS 中断服务程序。

3) CMOS 芯片

CMOS(Complementary Metal Oxide Semiconductor,互补金属氧化物半导体)是主板上的一块可读写的 RAM 芯片,主要用来存放 BIOS 中的设置信息以及系统时间和日期。如果 CMOS 数据损坏,则计算机将无法正常工作。为了确保 CMOS 数据不被损坏,主板厂商都在主板上设置了开关跳线,一般默认为关闭。当要对 CMOS 数据进行更新时,可将它设置为可改写。为了使计算机不丢失 CMOS 和系统时钟信息,在 CMOS 芯片的附近有一个电池给它持续供电。

3. 总线和 I/O 接口

1) 总线

总线(Bus)是计算机各种功能部件之间传送信息的公共通信干线,它是由导线组成的传输线束。按照计算机传输的信息种类,计算机的总线可以划分为数据总线、地址总线和控制总线,分别用来传输数据信号、地址信号和控制信号。

总线是一种内部结构,是 CPU、内存、输入输出设备传递信息的公用通道,主机的各个部件通过总线相连接,外围设备通过相应的接口电路再与总线连接,从而形成计算机的硬件系统,如图 1-20 所示。微型机是以总线结构来连接各个功能部件的。

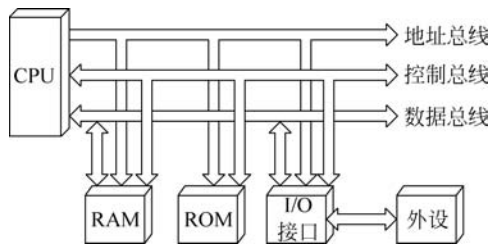


图 1-20 微型机总线结构

总线的主要技术指标有 3 个:总线位宽、总线工作频率和总线带宽。

(1) 总线位宽。总线位宽是指总线能够传送的二进制数据的位数。例如,32 位总线、64 位总线等。总线的位宽越宽,数据传输速率越大,总线带宽越宽。

(2) 总线工作频率。总线的工作频率以 MHz 为单位,工作频率越高,总线工作速度越快,总线带宽越宽。

(3) 总线带宽。总线带宽是指单位时间内总线上传送的数据量,反映了总线数据传送速率。总线带宽、总线位宽和总线工作频率之间的关系为

$$\text{总线带宽} = \text{总线工作频率} \times \text{总线位宽} \times \text{传输次数} / 8$$

其中,传送次数是指每个时钟周期的数据传输次数,一般为 1。

为了提高计算机的可拓展性,以及部件和设备的通用性,除了片内总线外,各个部件或设备都采用标准化的形式连接到总线上,并按照标准化的方式实现总线上的信息传输。而总线的这些标准化的连接形式及操作方式,统称为总线标准。常用的总线标准有 PCI 总线和 PCI-E 总线。

(1) PCI 总线。

PCI(Peripheral Component Interconnect)总线是一种同步的独立于处理器的 32 位或 64 位局部总线,是一种局部并行总线标准。

PCI 总线可以在主板上和其他系统总线(如 ISA、EISA 或 MCA)相连接,系统中的高速设备挂接在 PCI 总线上,而低速设备仍然通过 ISA、EISA 等这些低速 I/O 总线支持。

从 1992 年创立规范到如今,PCI 总线已成为计算机的一种标准总线,广泛用于高档微型机、工作站,以及便携式微型机,主要用于连接显示卡、网卡、声卡。

(2) PCI-E 总线。

PCI-E(PCI-Express)是最新的总线和接口标准,它原来的名称是 Intel 提出的 3GIO,意思是第三代 I/O 接口标准。2002 年正式命名为 PCI-Express。它采用了目前业内流行的点对点串行连接,比起 PCI 以及更早期的计算机总线的共享并行架构,每个设备都有自己的专用连接,不需要向整个总线请求带宽,而且可以把数据传输速率提高到一个很高的频率,达到 PCI 不能提供的高带宽。

根据总线位宽的不同,PCI-Express 规格允许实现 X1、X2、X4、X8、X12、X16 和 X32 通道规格,有非常强的伸缩性,可以满足不同系统设备对数据传输带宽不同的需求。从形式上来看,PCI-Express X1 和 PCI-Express X16 已成为 PCI-Express 的主流规格,芯片组厂商在南桥芯片中添加了对 PCI-Express X1 的支持,在北桥芯片中添加对 PCI-Express X16 的支持。除去提供极高数据传输带宽之外,PCI-Express 因为采用串行数据包方式传递数据,所以 PCI-Express 接口每个引脚可以获得比传统 I/O 标准更多的带宽,这样就可以降低 PCI-Express 设备的生产成本和体积。另外,PCI-Express 也支持高阶电源管理、热插拔、数据同步传输,为优先传输数据进行带宽优化。

在兼容性方面,PCI-Express 在软件层面上兼容 PCI 技术和设备,也就是说驱动程序、操作系统无须推倒重来。PCI-Express 可以为带宽渴求型应用分配相应的带宽,大幅提高 CPU 和图形处理器(Graphics Processing Unit,GPU)之间的带宽。

2) I/O 接口

I/O(Input/Output Port)即输入输出接口。每个设备都会有一个专用的 I/O 地址来处理自己的输入输出信息。由于计算机的外围设备种类繁多,几乎都采用了机电传动设备,CPU 在与 I/O 设备进行数据交换时存在很多不匹配的问题,因此 CPU 与外围设备之间的数据交换必须通过接口来完成。I/O 接口的功能实现 CPU 通过系统总线把 I/O 电路和外围设备联系在一起。

I/O 接口是一个电子电路(以 IC 芯片或接口板的形式出现),其内由若干专用寄存器和相应的控制逻辑电路构成。它是 CPU 和 I/O 设备之间交换信息的媒介和桥梁。CPU 与外围设备、存储器的连接和数据交换都需要通过接口设备来实现,通常前者称为 I/O 接口,后者称为存储器接口。

计算机系统有很多不同种类的 I/O 设备,其相应的接口电路也各不相同,因此 I/O 接口也很多,下面对一些目前比较常见的接口做具体说明。

(1) SATA。SATA(Serial ATA, 串行 ATA)的主要功能是用作主板和大量存储设备(如硬盘及光盘驱动器)之间的数据传输。这是一种完全不同于传统 ATA(也就是并行 ATA)的新型硬盘接口类型,因采用串行方式传输数据而得名。SATA 总线使用嵌入式时钟信号,具备了更强的纠错能力,与以往相比其最大的区别在于能对传输指令(不仅是数据)进行检查,如果发现错误会自动矫正,这在很大程度上提高了数据传输的可靠性。串行接口还具有结构简单、支持热插拔的优点。图 1-21 为 SATA 接口。

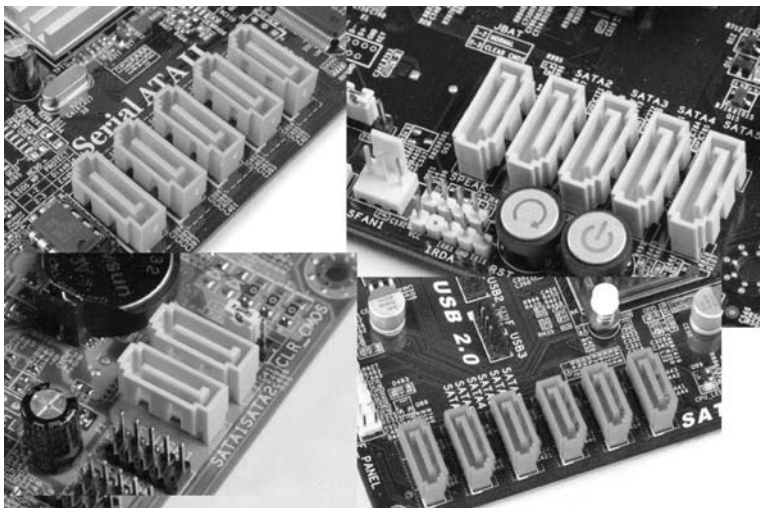


图 1-21 SATA 接口

(2) USB。USB(Universal Serial Bus, 通用串行总线)是由 Intel、Compaq、Digital、IBM、Microsoft、NEC、Northern Telecom 七家世界著名的计算机和通信公司共同推出的一种新型接口标准。它基于通用连接技术,实现外围设备的简单快速连接,达到方便用户、降低成本、扩展 PC 连接外围设备范围的目的。它可以为外围设备提供电源,而不像普通的使用串口和并口的设备那样需要单独的供电系统。另外,快速是 USB 技术的突出特点之一,而且 USB 还能支持多媒体。图 1-22 为 USB 接口。



图 1-22 USB 接口

目前,USB 接口的主流版本是 USB 3.1,传输速度为 10Gb/s,有三段式电压 5V/12V/20V,最大供电为 100W,而且新型 Type C 插型不再分正反面。USB 设备主要具有以下优点。

① 可以热插拔。用户在使用外接设备时,不需要关机再开机等动作,而是在计算机工作时,直接将 USB 插上使用。

② 携带方便。USB 设备大多以小、轻、薄见长,对用户来说,随身携带大量数据时,使用 USB 设备很方便。例如,USB 硬盘就是首选。

③ 标准统一。过去大家常见的设备是 IDE 接口的硬盘、串口的鼠标键盘、并口的打印机和扫描仪,有了 USB 接口之后,这些外围设备都可以用同样的标准与 PC 连接,这就有了 USB 硬盘、USB 鼠标、USB 打印机等。