微型计算机系统

1.1 走进计算机

计算机是 20 世纪人类最伟大的发明之一,在从诞生起至今的半个多世纪中,它由最初的"计算"工具,迅速发展成为应用于各行各业的信息处理设备,成为人类工作和生活中不可缺少的助手。现在,几乎每个年轻人都会使用计算机。但每位使用者是不是都了解计算机呢?

我们平时说的计算机,确切地说是计算机系统。因为没有人会认为不带显示器和键盘、鼠标的机器能称为"计算机"。其实,就计算机系统而言,它可以说是一个"广义"的概念。因为现代的计算机系统,巨型机和微型机还是有比较多的差别,若融入网络技术和辅助的软件技术,如并行机、阵列机、机群系统、甚至今天比较"时髦"的云计算等,宏观上讲也都可以称为一个系统。

1.1.1 计算机系统构成

"计算机"是对一类系统的总称。它既可以指我们常见的个人计算机(Personal Computer)或称微型计算机(如图 1-1 所示),也可以是计算速度达每秒几亿亿次的超级计算机(如图 1-2 所示)^[1]。系统即包含可以看得见摸得着的硬件,也包含看得见却摸不着的各种软件。比如我们开机就可以看到的操作系统,或是象文字编辑、游戏等各类应用软件。也就是说,计算机系统不仅包含物理上能够看得见的硬件实体,还包含运行于实体之上的、可实现各种操作功能的软件。即:计算机系统是由硬件系统和软件系统两大部分组成的,其整体概念结构图如图 1-3 所示。由于我们更经常、更直接接触和使用的计算机大多数是微型计算机(Microcomputer),也称个人计算机(Personal Computer,PC)。所以,以下如无特殊说明,本书中所说的计算机均特指微型计算机,所介绍的计算机结构及基本结构和工作原理也均以微型计算机为蓝本。





亿亿次双精度浮点运算,2013年6月在德国莱比锡召开的国际超级计算机大会上,以其优异性能位居全球榜首。

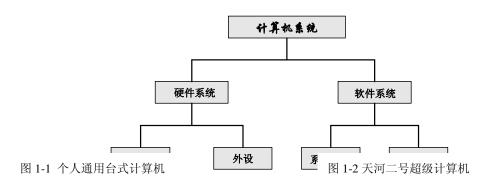


图 1-3 计算机系统概念结构

1. 硬件系统

微机硬件系统包括主机和能够与计算机进行信息交换的外部设备两部分。主机位于我们日常看到的主机箱内,主要包括微处理器(CPU)、内存储器、I/0接口、总线和电源等。其中,微处理器是整个系统的核心。能否与处理器进行直接信息交换是能够成为主机部件的重要标志。所谓"直接信息交换",就是不需通过任何中间环节(用专业术语说是接口),就能够实现从处理器接收数据或向处理器发送数据。典型的如内存,与处理器间的数据传输就是直接进行的。事实上,计算机正在运行的所有程序和数据,不论其曾经存放在哪里,在运行前都必须送入内存(也就是说,虽然你把文件都存放在硬盘上,但计算机操作这些文件时,它们是在内存中的,只是你不知道它们是怎样"自动"进入的而已),因为只有内存中的内容,处理器可以直接去"拿"(读取)、去"放"(写入),从而保证计算机工作的高速度。

今天,如果有人告诉你他买了一台计算机,你一定会清楚他不是只抱了一台主机箱回来, 而至少还包括显示器、键盘和鼠标。这些,都称为计算机的基本外部设备。

所谓外部设备,是指所有能够与计算机进行信息交换的设备(当然,这种信息交换需要通过接口)。它们既包括使用计算机所必须的基本外部设备(如上述的键盘、鼠标、显示器等),也包括其它各种能够连接到计算机、能够接受计算机所发送出的各种信息或向计算机发送信息的各类设备、控制仪器等。我们把用于向计算机输入信息的称为输入设备,如键盘、鼠标器、扫描仪等;接收计算机输出信息的设备则称为输出设备,如显示器、打印机、绘图仪等。当然,也有些设备既能接收计算机输出的信息,也能向系统输入信息,如数码摄像机、硬磁盘(考虑一下:硬盘为什么是外部设备呢?)等。亦即它们兼具了输入设备和输出设备的功能,具体担当何种角色,则视其在某个时刻传送数据的方向。

相对于主机,外部设备的主要特点就是不能与处理器直接进行数据输入和输出,数据的传输必须通过接口。如硬磁盘,虽然安装在主机箱内,但不属于主机系统,因为它与处理器的通信需要通过专用接口进行。而至于什么是接口,会在本书的1.1.2中给出简单的介绍。

有关计算机常用外设的基本工作原理,请参见本书附录 A。

2. 软件系统

硬件系统是计算机工作的物理基础,但要使其正常工作并完成各种任务,还必须要有相应的软件支撑。所谓软件,不仅仅是一般概念中的程序,而是程序、数据、以及相关文档的总称。这里,数据是程序处理的对象,文档是指与程序开发、维护和使用有关的各种图文资料。软件可以分为两大类:系统软件和应用软件。

系统软件是管理、监控和维护计算机软硬件资源的软件,由计算机设计者提供,包括操作系统和各种系统应用程序。操作系统(OS,Operating System)是配置在计算机硬件上的第一层软件,是其它软件运行的基础。其主要功能是管理计算机系统中的各种硬件和软件资源(如存储器管理、文件管理、进程管理、设备管理等),并为用户提供与计算机硬件系统之间的接口(如通过键盘发出命令控制作业运行等)。在计算机上运行的其它所有的系统软件(如编译程序、数据库管理系统、网络管理系统等)及各种应用程序,都要依赖于操作系统的支持。因此,操作系统是计算机中必须配置的软件,在计算机系统中占据着及其重要的位置。目前较为流行的操作系统有Windows系列、Unix、Linux、以及苹果计算机采用的Mac系列等。

系统应用程序运行于操作系统之上,是为应用程序的开发和运行提供支持的软件平台。 主要包括:

- 编译程序。用于将用各种计算机语言(汇编语言或各种高级语言)编写的程序翻译成计算机硬件能够直接识别的用二进制码表示的机器语言。由于计算机硬件是由各种逻辑器件构成,只能识别电脉冲信号,也就是"0"合"1"组成的二进制码,这种由二进制码组成的计算机语言称为机器语言,人类很难理解和记忆。目前广泛使用的计算机程序设计语言都是接近人类自然语言的高级语言,为了使计算机都够理解,必须要经过一个翻译的过程,而这类程序的功能就是实现这样的翻译。
- 计算机的监控管理程序(Monitor)、故障检测和诊断程序,以及调试程序(Debug)。 它们负责监控和管理计算机资源,并为应用程序提供必要的调试环境。
- 各类支撑软件,如数据库管理系统及各种工具软件等。

应用软件是应用程序员利用各种程序设计语言编写出的、面向各行各业实现不同功能的应用软件。如工程设计程序、数据处理程序、自动控制程序、企业管理程序等。目前,软件的设计还没有摆脱手工操作的模式,但随着软件技术的进步,应用软件也在逐渐地向标准化、模块化方向发展,目前已形成了部分用于解决某些典型问题的应用程序组合,称为软件包(Package)。

软件系统的核心是系统软件,而系统软件的核心则是操作系统。

计算机系统是硬件和软件的结果体,硬件和软件相辅相成,缺一不可。硬件是计算机工作的物质基础,而软件是计算机的灵魂。没有硬件,软件就失去了运行的基础和指挥对象;而没有软件,计算机就不能工作,其效能就不能充分发挥出来。

对某项具体任务,通常可以既用硬件完成,也能通过软件完成。从理论上讲,任何软件算法都能用硬件实现,反之亦然,这就是软件与硬件的逻辑等价性。设计计算机系统或是在现有的计算机系统上增加功能时,具体采用硬件还是软件实现,取决于价格、速度、可靠性等因素。早期的计算机受技术合成本的限制,硬件都相对简单。如今,随着超大规模集成电路技术的发展,以前由软件实现的功能现在更多地直接用硬件实现,为的是提高系统的运行速度和效率。另外,在软件和硬件之间还出现了所谓的固件(firmware),它们在形式上类似

硬件,但从功能上又像软件,可以编程和修改,这种趋势称为软件的硬化和固化。

1.1.2 主机与主机板

1. 主机

PC 机的硬件系统除了各种类型的外部设备之外,最主要的就是主机系统了,它是组成 微型计算机的主体。主机系统的主要部件包括中央处理单元(Central Processing Unit, CPU)、内存储器、总线和输入输出接口。其基本组成如图 1-4 所示。

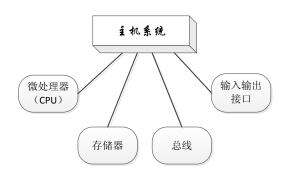


图 1-4 主机系统

1) CPU

中央处理单元(Central Processing Unit, CPU)在微机中也称为微处理器(Microprocessor), 是微型计算机的核心芯片,也是整个系统的运算和指挥控制中心。它最基本的功能就是按照 程序一条一条地执行指令。每一条指令都规定了一个基本的操作,并且与相应的物理电路对 应,可以说这就是计算机硬件和软件之间最基本的连接,也是计算机软、硬件协同的基础。 有关什么是指令、以及 CPU 的内部结构,将在本书的 3.3 节中做具体的介绍。

CPU 是人类借助超大规模集成电路技术生产的最复杂、最精细的产品之一。CPU 的种类很多,除了微型计算机中的微处理器之外,各种网络服务器、巨型机等中的高性能处理器,则是计算能力更强的 CPU。另外,安装在各种现代化仪器设备、通信设备内的处理器,则称为"嵌入式 CPU",目前几乎所有的高档电器内部都装备了一片或几片这种"嵌入式 CPU"。

无论哪一种 CPU, 其内部总体上都主要包括三大部分,即:运算器、控制逻辑单元(或称控制器)和内部寄存器组(如图 1-5 所示),各部分通过 CPU 内部总线连接在一起,现代微处理器中还包含有高速缓冲存储器(Cache)。

运算器的主要部件是算术逻辑单元 ALU (Arithmetic Logical Unit),它是运算器的主体。 ALU 的主要功能就是在控制信号的作用下可完成加、减、乘、除等算术运算、移位操作及各种逻辑运算(扩大一点讲就是执行指令),现代新型 CPU 的运算器还可完成各种浮点运算。运算产生的中间结果可以存放在 CPU 的内部寄存器中。

内部寄存器组是 CPU 内部的若干个用于暂时存放数据的存储单元。包括多个通用寄存器和若干专用寄存器。寄存器的功能按其字面意思可理解为是用来暂时存放数据的部件(当然,这里的"数据"

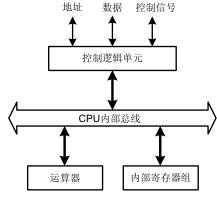


图 1-5 CPU 基本结构

是广义的,并非仅仅是数值。),"暂存数据"这一任务主要由通用寄存器完成。设置它们的目的是在执行程序过程中,对某些需要重复使用的操作数据或中间运算结果,可将它们暂时存放在寄存器中,以避免对存储器的频繁访问,从而缩短执行时间。专用寄存器的作用则是固定的。

为了便于本书后续内容的学习,有一个专用寄存器需要在这里先做初步的介绍,这就是程序计数器(Program Counter,PC)。程序计数器用于指示下一条要取指令的地址。不论是程序员编写完成的程序,还是我们日常编写的各种文档,通常都是存放在外存储器(如硬盘)中,它们在被CPU 执行之前,都需要先进入到内存^[1]。而CPU 在执行的时候,为什么就能够按照我们要求的顺序去执行程序、处理文档呢?靠的就是程序计数器 PC 的控制。所以,它是非常重要的一个内部寄存器^[2]。

控制逻辑单元主要用于控制和协调整个 CPU 的工作,是整个 CPU 的指挥控制中心。我们日常的上下班是以时间为基准的,CPU 的工作基准是时钟信号。这是一组周期恒定的脉冲信号。不同的时刻 CPU 做不同的工作,它们在时间上有着严格的关系,这就是时序。时序信号由控制器产生,控制 CPU 的各个部件按照一定的时间关系有条不紊地完成指令要求的操作。

各种数据在 CPU 中的传送要借助于 CPU 的内部通道,这个通道称为总线(Bus)。

从上世纪七十年代第一片处理器诞生至今,CPU 的性能不断提高。近年来,人们又开发了一种具有新型体系结构的 CPU——多核 CPU。所谓多核,指的是在一个芯片上集成多个物理的 CPU 运算内核(即运算器),这些运算内核可以并行、协同地工作(这就相当于将一件复杂的工作分给多个人同时工作一样),从而使计算机的处理能力大大增强。多核 CPU 和以前的单核 CPU 在外观上并没有太大的区别,但其内部结构已是大不相同了。图 1-6 是 Intel Pentium 4 单核处理器 Intel Core i7 四核处理器。





图 1-6 微处理器外观

2) 存储器

存储器(Memory)是计算机中用于存储各种信息的部件,总体上可以分为内存和外存两种类型。对存储器的操作有两种,即"读"和"写"。"读"表示从存储器中输出数据,也称为读取;"写"表示向存储器输入数据,也称为写入。对存储器的读写,可以按字节、字或块。

内存储器由半导体材料制成(所以也称半导体存储器),属于主机部分,用于存放计算机当前运行的程序(包括确保计算机运行所必须的程序)及运算的数据。我们常说的内存条

^[1] 有关编写的程序如何进入到内存以及为什么要进入内存才能被求理部分的描述。



^[2] 要真正理解程序计数器 PC 的作用,需要深入理解微型计算机的工作尽理。作为订具机基础取材,4节 仅在第3章对微型机的基本工作原理做一点初步的介绍,若希望深入理解,需要进一步学习其它书籍。

(如图 1-7 所示)就是主要的内存储器,称为主存。

相对于外存储器,内存的主要特点有:

- 可以与 CPU 直接进行信息交换 (所以,它属于主机部分);
- 存取速度快、容量小、单位字节容量价格较高;

图 1-7 内存条

● 需要后备电源,当断电时,其上存放的信息将丢失。

如同一栋大楼是由若干个房间组成一样,内存由若干个单元组成(如图 1-8 所示)。大楼中的每个房间都有门牌号码,且每个号码在楼内都是惟一的,目的是便与寻找;同样,内存中的每个单元也有"门牌号码",称为地址码,每个单元的地址在内存中也是惟一的。由于计算机只能识别二进制,所以内存中的地址码都是用二进制表示。地址码的长度依内存单元的个数(称为容量)而定。比如:4个单元的内存的地址码只需要2位二进制就可以表示(明白为什么吗?),而4G个单元的内存中,每个单元的地址则需要32位二进制码来表示。

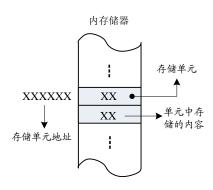


图 1-8 内存结构示意图

在微机系统中,内存的每个单元都存放 8 位二进制码,即 1B 数据。内存的容量就是指它具有的单元数。如常说的 2GB 内存,意思就是该内存有 2G(2³⁰)个单元,每单元中有 1 字节数据。

对内存的读/写操作通常按"字"进行,不同的系统"字"的长度不同。目前的微型机 多为 64 位机,其在一个周期中能够对内存读出或写入 8B 数据。

外存储器包括联机外存和脱机外存两种,用于存放 CPU 不直接执行、但可以长期保留的数据。脱机外存有光驱、磁带、移动存储器等,由复合材料(如光盘)、磁性材料或半导体材料(如优盘)构成。它们可以脱离计算机而存在,所以理论上可以存放无限多的数据。联机外存就是我们常说的硬磁盘。



图 1-9 硬磁盘及其及内部结构

硬盘是微机中主要且必备的存储部件,由多片磁性材料制造的盘片叠加在一起构成(如图 1-9 所示)。每个盘片有两个记录面(每个记录面对应一个磁头,所以也用磁头数表示记录面数),每个记录面上是一系列称为磁道的同心圆(多个记录面上的同心圆叠放在一起就构成柱面),每个磁道又被划分为若干个扇区(Sector)。硬磁盘就是按记录面、磁道和扇区来对数据进行组织的。对硬磁盘进行读/写操作的基本单位是扇区,每个扇区的容量为 512 字

磁头数×柱面数×扇区数×512(B)

(1, 1)

【例 1-1】设已知磁头数为 16, 柱面数为 4096, 扇区数为 64, 求该硬盘的容量。 由式 1.1 可得:

该硬盘的容量=16×4096×64×512=2147483648B=2GB

外存储器的主要作用是保存各种希望由计算机保存和处理的信息。相对于内存,外存具有存储容量大(目前微机硬磁盘的常规配置为 1TB,而内存通常为 4GB 或 8GB)、速度慢、单位字节容量价格低、不能与处理器直接进行信息交换等特点。外存储器虽然也安装在主机箱中,但属于外部设备的范畴。理由是外存与处理器之间的信息交换需要通过输入/输出接口。目前微型机中最常用的硬盘接口标准是 SATA(Serial Advanced Technology Attachment),它们定义了外存储器(如硬盘、光盘等)与主机的物理接口。

3)输入输出接口

上文中已提到,凡与处理器之间的信息交换需要通过输入/输出接口的设备都称为外部设备(比如硬磁盘)或输入输出设备。我们很容易想到,不能连接外部设备的计算机是没有意义的(你能够想象没有显示器、鼠标和键盘的计算机吗?)。

所有能够与计算机通信的设备都可以称为输入输出设备,它们虽然千差万别,种类繁多,结构、原理各异,但都有一个共同的特点,就是:要想接收来自 CPU 的信息或将信息送入 CPU 去处理,必须通过一个中间环节,就是输入/输出接口(Input/Output Interface,简称 I/0 接口),也称 I/0 适配器(I/0 Adapter)。

I/0接口是将外设连接到系统总线上的一组逻辑电路的总称,也称为外设接口。其在系统中的作用如图1-10所示。在一个实际的计算机控制系统中,CPU与外部设备之间常需要进行频繁的信息交换,包括数据的输入输出、外部设备状态信息的读取及控制命令的传送等,这些都是通过接口来实现的。由I/0接口在系统中的位置,使得接口电路应解决如下问题,也是接口应具有的功能:

- ① CPU与外设的速度匹配。CPU与外设之间的工作时序和速度差异很大,要使两者之间能够正确进行数据传送,需要接口做"适配"。接口电路应具有信息缓冲能力,不仅应缓存CPU送给外设的信息,也要缓存外设送给CPU的信息,以实现CPU与外设之间信息交换的同步。
- ② 信息的输入输出。通过I/0接口,CPU可以从外部设备输入各种信息,也可将处理结果输出到外设。同时,为保证数据传输的正确性,需要有一定的监测、管理、驱动等能力。
- ③ 信息的转换。外部设备种类繁多,其信号类型、电平形式等与CPU都可能存在差异。 I/0接口应具有信息格式变换、电平转换、码制转换、传送管理以及联络控制等功能。
- ④ 总线隔离。为防止干扰,I/0接口还应具备一定的信号隔离作用,使各种干扰信号 不影响CPU的工作。



图 1-10 I/O 接口在系统中的作用示意图

2. 主机板

主机板也称主板(mainboard)或系统板(System board),是微型机的物理构成,在整个微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着整个微机系统的类型和档次,主板的性能影响着整个微机系统的性能。

主机板在结构上主要有 AT 主板、ATX 主板、NLX 主板和 BTX 主板等类型。它们之间的区别主要在于各部件在主板上的位置排列、电源的接口外形、控制方式、及尺寸等。不论哪种结构,均采用开放式结构。可以通过更换安装在扩展插槽上的外围设备控制卡(适配器),实现对微机相应子系统的局部升级。图 1-11 为一个实际的 ATX 主板的布局结构及外型图。

主板位于主机箱内,上面安装了组成计算机的主要电路系统,主要包括芯片、扩展槽和 对外接口三种类型的部件

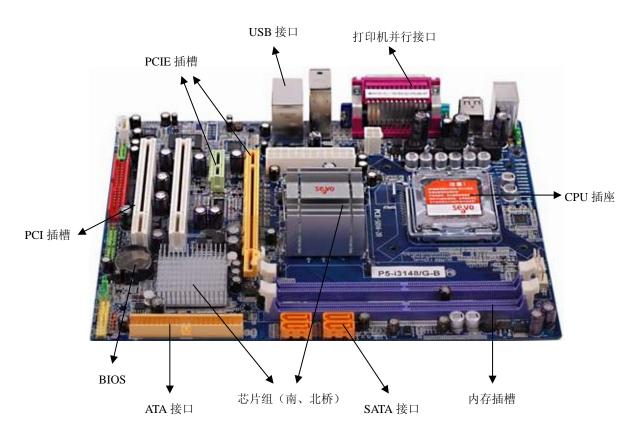


图 1-11 主机板

1) 芯片部分

这部分除微处理器(CPU)外,主要有控制芯片组和 BIOS。 芯片组是主板上一组超大规模集成电路芯片的总称,是主板的关键部件,用于控制和协 调计算机系统各部件的运行,它在很大程度上决定了主板的功能和性能。可以说,系统的芯片组一旦确定,整个系统的定型和选件变化范围也就随之确定。

典型的芯片组由北桥芯片和南桥芯片两部分(2 片芯片)组成,故也称南北桥芯片。图 1-11 中 CPU 插槽旁边被散热片盖住的就是北桥芯片。北桥芯片是芯片组的核心,主要负责处理 CPU、内存、显卡三者间"交通",由于发热量较大,故需加装散热片散热。南桥芯片主要负责硬盘等存储设备和 PCI 之间的数据流通。

需要说明的是,现在一些高端主板上已将南北桥芯片封装到一起,使"芯片组"在形式上只有一个芯片,提高了芯片组的功能。

BIOS 也称系统 BIOS 是一块方块状的存储器芯片,里面存有与该主板搭配的基本输入输出系统程序。能够让主板识别各种硬件,还可以设置引导系统的设备,调整 CPU 外频等。BIOS 芯片是可读写的只读存储器(EPROM 或 E²PROM)。机器关机后,其上存储的信息不会丢失。在需要更新 BIOS 版本时,还可方便地写入。当然,不利的一面便是会让主板遭受病毒的袭击。

系统 BIOS 程序主要包含以下几项功能:

- 上电自检(POST, Power-On Self Test)。在微机加电后, CPU 从地址为 0xFFFFFF0H 处读取和执行指令, 进入加电自检程序, 测试整个微机系统是否工作正常。
- 初始化。包括可编程接口芯片的初始化;设置中断向量表(一个专门用于存放中断程序入口地址的内存区域);设置 BIOS 中包含的中断服务程序的中断向量(即将这些中断程序入口地址放入中断向量表中);通过 BIOS 中的自举程序将操作系统中的初始引导程序装入内存,从而启动操作系统。
- 系统设置(Setup)。装入或更新 CMOS RAM 保存的信息。在系统加电后尚未进入操作系统时,按键(或其它热键)可进入 Setup 程序,修改各种配置参数或选择默认参数。

2) 扩展槽

安装在扩展槽上的部件属于"可插拔"部件。所谓"可插拔"是指这类部件可以用"插"来安装,用"拔"来反安装。主板上的扩展槽包括内存插槽和总线接口插槽两大类。

内存插槽一般位于 CPU 插座下方,用于安装内存储器(也就是内存条)。通过在内存插槽上插入不同的内存条,就可方便地构成所需容量的内存储器。主板上内存插槽的数量和类型对系统主存的扩展能力及扩展方式有一定影响。

<u>总线接口插槽</u>是插接各种扩展接口卡的地方,是 CPU 通过系统总线与外部设备联系的通道,主要有 PCI 插槽、AGP 插槽或 PCI Express (PCIE) 插槽。

所谓**总线**,可以简单地说是计算机中传输信息的通道。微机中的总线按照其层次结构,可以分为 CPU 总线(或称前端总线)、系统总线和外设总线。

前端总线一般是指从 CPU 引脚上引出的连接线,用来实现 CPU 与主存储器、CPU 与 I/0 接口芯片、CPU 与控制芯片组等芯片之间的信息传输,也用于系统中多个 CPU 之间的连接。前端总线是生产厂家针对其具体的处理器设计的,与具体的处理器有直接的关系,没有统一的标准。

系统总线也称为 I/0 通道总线,是主机系统与外围设备之间的通信通道。在主板上,系统总线表现为与总线接口插槽(也称 I/0 插槽)的引线连接的一组逻辑电路和导线。I/0 插槽上可插入各种扩展板卡,它们作为各种外部设备的适配器与外设相连。为使各种接口卡能

够在各种系统中实现"即插即用",系统总线的设计要求与具体的CPU型号无关,而有自己统一的标准,各种外设适配卡可以按照这些标准进行设计。目前常见的总线标准有PCI总线、PCI-E总线等。

PCI 插槽多为乳白色,是主板的必备插槽,可以插入声卡、股票接受卡、网卡、多功能卡等设备。AGP 插槽的颜色多为深棕色,位于北桥芯片和 PCI 插槽之间,用于插入 AGP 显卡,有 $1\times$ 、 $2\times$ 、 $4\times$ 和 $8\times^{[1]}$ 之分。在 PCI Express 出现之前,AGP 显卡是主流显卡,其传输速度最高可达到 2133MB/s(AGP8×)。随着 3D 性能要求的不断提高,AGP 总线的传输速度已越来越不能满足视频数据处理的要求。在目前的主流主板上,显卡接口多转向 PCI Exprss。PCI Exprss 插槽有 $1\times$ 、 $2\times$ 、 $4\times$ 、 $8\times$ 和 $16\times$ 之分。

外设总线是指计算机主机与外部设备接口的总线,实际上是一种外设接口标准。目前在 微机系统中最常用的外设接口标准就是 USB(Universal Serial Bus,通用串行总线),可以用来连接多种外部设备。

图 1-12 所示为现代微型机中的总线结构示意图。

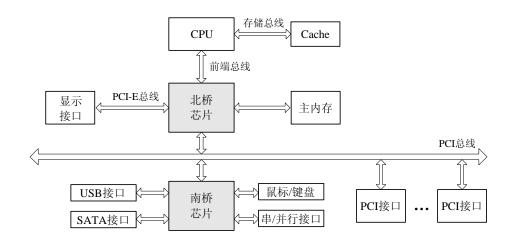


图 1-12 现代微机中的多总线结构

除上述这些主要部件外,主板上还有用于连接硬盘、光驱等的电缆插座、键盘/鼠标接口以及许多不可缺少的逻辑部件和跳线开关等。所有这些部件的密切联系、相互沟通,实现了整个微型机中各部件间的数据交流。

_

^[1] nX 表示 n 倍速,即对原来的时钟脉冲进行技术处理后,使时钟频率变成 n 倍频。