

普通高等学校计算机专业特色教材

计算机组成与体系结构

裘雪红 李伯成 车向泉 刘凯



高等教育出版社
Higher Education Press

普通高等学校计算机专业特色教材

计算机组成与体系结构

裘雪红 李伯成 车向泉 刘 凯

高等教育出版社

内 容 提 要

本书主要讲授计算机的基本体系结构、组成原理和实现方法,涉及的内容从计算机最底层的 CPU 核心直到最上层的并行系统架构,从单机、多机到机群的各种计算机系统中采用的指令设计与流水线处理技术、存储体系与存储技术、输入输出系统与 I/O 技术、并行体系结构及互连技术等成熟技术与新技术,并结合新产品、新系统说明了各种技术的应用。

本书力求语言精练、深入浅出、通俗易懂、重点突出;在强调原理的同时,注重技术与实例的结合;在强调基础知识的同时,注重新技术的融入。

本书与教育部高等学校计算机专业教学指导分委员会制定的《计算机科学与技术专业规范》中建议的“计算机组织与体系结构”教学大纲和 ACM/IEEE-CS 课程指南吻合,涵盖全国硕士研究生入学考试计算机科学与技术学科联考大纲“计算机组成原理”课程内容,适用于高校计算机及相关专业“计算机组织与体系结构”和“计算机组成原理”课程的教学与自学,能够帮助学生建立完整的计算机组成与体系结构的基本概念和知识体系。

图书在版编目(CIP)数据

计算机组成与体系结构 / 裘雪红等. —北京: 高等教育出版社, 2009. 6

ISBN 978-7-04-026471-5

I. 计… II. 裘… III. 计算机体系结构-高等学校-教材
IV. TP303

中国版本图书馆 CIP 数据核字(2009)第 088163 号

策划编辑 倪文慧 责任编辑 郭福生 封面设计 张志奇 责任绘图 尹 莉
版式设计 张 岚 责任校对 金 辉 责任印制 韩 刚

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 北京中科印刷有限公司

开 本 787×1092 1/16
印 张 35.25
字 数 800 000

购书热线 010-58581118
免费咨询 400-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>
畅想教育 <http://www.widedu.com>

版 次 2009 年 7 月第 1 版
印 次 2009 年 7 月第 1 次印刷
定 价 38.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 26471-00

前 言

为了将“计算机组成原理”与“计算机系统结构”这两门紧密关联的课程有机地联系在一起,也为了避免两门独立课程中部分内容重叠造成课时浪费,已有一些院校的计算机专业选择将这两门课整合为一门课程,本书就是为这种课程提供的配套教材。另外,本书也完全适合作为“计算机组成原理”课程教材及研究生入学考试学习用书。

目前关于计算机组成与体系结构的中外教材有不少,但在教学一线的实践感受是,仍然需要一本内容能够比较完整地覆盖计算机组成与体系结构,原理性比较强又结合新技术,适合教学需要的教材,这也是促成我们编写本书的最直接原因。编写本书时,我们努力做到以普适的基本原理应对多变的计算机机型,并以实际的计算机产品或机器作为实例,使读者能够将原理与实践联系起来。

本书主要讲述计算机硬件的基本组成和典型的并行体系结构。第1章给出了对计算机的概述,可以帮助对计算机不太了解的读者先有一个感性的认识。第2章介绍计算机系统中的数据表示,包括数值、非数值表示及校验码,这是计算机设计、实践的基础,也是认识计算机信息的基础。第3章是运算方法与运算器,重点讲述定点数与浮点数的加、减、乘、除运算方法和具体实现,读者可以通过本章了解计算机是如何进行计算的以及如何设计算术逻辑部件(ALU)。第4章讲述存储系统,涉及存储系统的基本概念、内存与外存,读者可以从中了解到存储系统各层次的存储器工作原理及存储技术。第5章是指令系统,在一般性介绍指令系统的基础上,选择 Pentium 指令系统和 SPARC 指令系统作为 CISC 与 RISC 指令系统的实例,让读者对 CISC 与 RISC 有一个基本的认识。第6章描述的是 CPU,包括 CPU 结构、控制器设计以及 CPU 性能、新技术和实例,本章将从内到外较为全面地介绍 CPU。第7章详细介绍了流水线技术及指令级并行概念,包括浮点运算流水线、指令流水线、流水线性能量度、指令流水线的性能提高、提高指令级并行的技术、多发射处理器等内容,从中可以学习和掌握关于流水线较为完整的知识和技术。第8章是总线与输入输出系统,涉及总线基本知识、仲裁、实例和程序查询、中断、DMA、I/O 通道等输入输出技术,本章有助于读者掌握 I/O 设备接入计算机系统的连接技术。第9章介绍并行体系结构,重点讨论了计算机体系结构的并行性、多处理器与多计算机系统体系结构以及互连网络等并行计算机系统中普遍关注的问题,读者可以从本章了解当今高性能计算机的基本结构及相关技术。第10章讨论计算机系统设计问题,这是在国内同类教材中较少涉及、而笔者认为很有用的内容,特别是在强调突出实践教学环节的理念下,本章可以让读者了解利用硬件描述语言(HDL)进行 CPU、计算机设计的基本方法,很有参考价值。

在编写本书的过程中,我们参考了较多的国内外同类教材及网上的相关资料,特别是参考文

献中提及的几本国外经典教材对本书的编写有着直接的影响,考虑到它们的权威性,本书的某些较前沿的内容取材于这些经典教材,在此,要特别感谢这些作者和他们的精彩著作。另外要感谢网上的参考文献资料、电子教案等的作者和网站(尽管有些作者的大名我们并不知晓),是他们给了我们许多灵感和素材。最后,还要感谢高等教育出版社为我们提供了出版本书的机会,感谢Sun公司通过我们的“计算机组织与体系结构”精品课程建设对本书编写提供的支持,感谢“西安电子科技大学教材建设基金资助项目”的资助。

本书的第1、3、4、10章由李伯成教授编写,第2章由车向泉副教授编写,第5章由刘凯副教授编写,第6~9章由裘雪红教授编写,全书由裘雪红教授统稿。由于实践条件及自身水平的限制,本书中有些新技术无法实际验证,所以一定存在对某些新技术理解的不透彻,书中难免有不妥之处,恳请专家、同行和读者给予批评指正,不胜感激。

作 者

E-mail: qiuxh0699@sina.com

2008年10月

目 录

第 1 章 绪论	1	3.1.3 除法运算	96
1.1 计算机的发展史	1	3.2 算术逻辑部件	104
1.1.1 发展经历	1	3.2.1 单元电路	104
1.1.2 摩尔定律	2	3.2.2 算术逻辑部件	106
1.2 计算机的基本组成	4	3.2.3 运算器的结构	108
1.2.1 硬件系统	4	3.3 浮点运算	109
1.2.2 软件系统	8	3.3.1 加减运算	109
1.3 计算机系统的组织结构	8	3.3.2 乘除运算	112
1.3.1 计算机系统的层次结构	9	3.3.3 浮点运算的实现	115
1.3.2 计算机系统结构、组成与实现	10	习题	116
1.4 计算机的分类及性能描述	11	第 4 章 存储系统	120
1.4.1 计算机的分类	11	4.1 存储系统概述	120
1.4.2 计算机的性能描述	14	4.1.1 存储系统的层次结构	120
习题	20	4.1.2 存储器分类	120
第 2 章 计算机系统中的数据表示	21	4.1.3 存储器的性能指标	121
2.1 数据表示	21	4.2 内存储器	123
2.2 数据编码	22	4.2.1 随机存储器	123
2.2.1 数值数据的编码	22	4.2.2 只读存储器	131
2.2.2 非数值数据的编码	36	4.2.3 动态存储器	133
2.3 数据的定点与浮点表示	44	4.2.4 主存储器校验	136
2.3.1 定点数表示	44	4.2.5 其他存储器	139
2.3.2 浮点数表示	46	4.3 高速缓冲存储器	143
2.4 检错与纠错码	58	4.3.1 工作原理	143
2.4.1 奇偶校验码	58	4.3.2 地址映射	144
2.4.2 循环冗余校验码	61	4.3.3 替换算法	149
2.4.3 汉明码	64	4.3.4 主存与 Cache 内容的一致性 问题	150
习题	70	4.3.5 Cache 性能分析	151
第 3 章 运算方法与运算器	75	4.3.6 Pentium 的 Cache	153
3.1 定点数运算	75	4.4 虚拟存储器	154
3.1.1 加减运算	75	4.4.1 虚拟存储器的概念	154
3.1.2 乘法运算	84	4.4.2 页式虚拟存储器	155

4.4.3 段式虚拟存储器	157	6.2 硬布线控制器设计	233
4.4.4 段页式虚拟存储器	158	6.3 微程序控制器设计	240
4.4.5 有关虚拟存储器的几点说明	159	6.3.1 微程序控制原理	240
4.4.6 Pentium 虚拟存储器	160	6.3.2 微指令设计	243
4.5 外存储器	161	6.3.3 微程序设计	253
4.5.1 磁表面存储器	162	6.3.4 微程序控制器设计	257
4.5.2 光存储器	172	6.4 微程序控制器与硬布线 控制器的比较	259
习题	176	6.5 CPU性能的测量与提高	259
第5章 指令系统	182	6.5.1 计算机系统性能测量	259
5.1 指令格式	182	6.5.2 CPU性能测量	260
5.1.1 机器指令特征	182	6.5.3 提高CPU速度的策略	268
5.1.2 指令格式及设计	183	6.6 CPU中的新技术	269
5.1.3 指令长度设计	186	6.6.1 多核技术	269
5.2 数据存储与寻址方式	187	6.6.2 多线程技术	270
5.2.1 数据存储方式	187	6.6.3 多核+多线程技术	273
5.2.2 寻址方式	189	6.7 典型的CPU	274
5.3 指令类型	197	6.7.1 Intel的CPU	275
5.3.1 数据传送类	197	6.7.2 SUN的CPU	276
5.3.2 算术运算类	197	6.7.3 MIPS的CPU	280
5.3.3 逻辑运算类	197	习题	281
5.3.4 数据转换类	199	第7章 流水线技术与指令级并行	287
5.3.5 输入输出类	199	7.1 流水线处理	287
5.3.6 系统控制类	199	7.1.1 流水线的一般结构	287
5.3.7 控制传递类	199	7.1.2 流水线类型	288
5.4 典型处理器的指令系统	201	7.2 浮点运算流水线	292
5.4.1 Pentium指令格式及核心指令	202	7.2.1 浮点加减法器流水线	292
5.4.2 SPARC指令格式及核心指令	209	7.2.2 浮点乘除法器流水线	295
5.5 CISC与RISC	215	7.3 指令流水线	298
5.5.1 CISC与RISC	216	7.3.1 基本的指令流水线	298
5.5.2 RISC的寄存器窗口技术	218	7.3.2 指令流水线策略	301
习题	220	7.4 流水线性能度量	305
第6章 中央处理器	222	7.4.1 时-空图	305
6.1 CPU结构和微操作	222	7.4.2 吞吐率	306
6.1.1 CPU的功能与结构	222	7.4.3 加速比	308
6.1.2 指令周期	224	7.4.4 效率	309
6.1.3 微操作	225	7.4.5 吞吐率、加速比和效率的关系	310
6.1.4 控制器的组成	232		

7.4.6 流水线性能分析	310	9.2 计算机体系结构的分类	398
7.5 指令流水线的性能提高	313	9.3 阵列处理机和向量处理机	400
7.5.1 流水线的基本性能问题	314	9.3.1 阵列处理机	400
7.5.2 结构相关	315	9.3.2 向量处理机	402
7.5.3 数据相关	315	9.4 互连网络	407
7.5.4 控制相关	317	9.4.1 基本概念	407
7.6 指令级并行概念	322	9.4.2 网络拓扑结构	408
7.6.1 指令流水线的限制	322	9.4.3 常用的互连模式和互连网络	424
7.6.2 突破限制的途径	323	9.5 多处理器系统	430
7.7 提高指令级并行的技术	324	9.5.1 UMA 对称多处理器系统	430
7.7.1 程序的相关	324	9.5.2 NUMA 对称多处理器	439
7.7.2 指令调度	325	9.5.3 Sun T1 多处理器	442
7.7.3 乱序执行和寄存器重命名	326	9.6 多计算机系统	442
7.7.4 推测执行	331	9.6.1 多计算机的概念	443
7.8 多发射处理器	332	9.6.2 MPP	447
7.8.1 超标量处理器	334	9.6.3 机群	451
7.8.2 超长指令字处理器	336	9.7 网格	457
7.8.3 多发射处理器的限制	339	9.7.1 网格概述	457
7.9 指令级并行的限制	339	9.7.2 Globus 体系结构	460
习题	340	9.7.3 网格计算用户案例	461
第 8 章 总线与输入输出系统	344	9.7.4 网格的难题	463
8.1 总线与输入输出系统概述	344	9.8 并行处理面临的挑战	464
8.2 总线	346	习题	465
8.2.1 总线类型与结构	346	第 10 章 计算机系统设计	470
8.2.2 总线的信息传输方式	352	10.1 概述	470
8.2.3 总线仲裁	357	10.1.1 数字系统设计的发展	470
8.2.4 典型的总线	362	10.1.2 IP 核	471
8.3 输入输出接口	366	10.1.3 数字系统的设计方法	472
8.4 输入输出技术	369	10.2 设计语言与工具	473
8.4.1 程序查询方式	369	10.2.1 EDA 工具软件分类	473
8.4.2 中断方式	370	10.2.2 硬件描述语言 HDL	474
8.4.3 直接存储器存取方式	381	10.2.3 SystemC	484
8.4.4 I/O 通道方式	386	10.2.4 Quartus II	486
8.4.5 操作系统的支持	391	10.3 基于 VHDL 的 CPU 设计	487
习题	393	10.3.1 单元电路的设计	487
第 9 章 并行体系结构	395	10.3.2 CPU 设计概要	500
9.1 计算机体系结构的并行性	395	10.4 SoC 设计	523

10.4.1	概述.....	523	10.5.1	问题的由来.....	545
10.4.2	SoC 片内总线	528	10.5.2	Tensilica Xtensa 可配置处理器	546
10.4.3	IP 核设计与复用	536	习题	553	
10.4.4	SoC 设计举例	542	参考文献	554	
10.5	基于可配置处理器的 SoC 设计 ...	545			

第 1 章 绪 论

本章主要描述计算机的构成及性能,使读者在开始了解全书内容之前,首先对计算机有一个概念上的认识。

1.1 计算机的发展史

电子计算机的发明,无疑是人类科学技术发展史上最伟大的发明之一。计算机的出现深刻地影响着人类的精神文明和物质文明的发展。

在 20 世纪 70 年代前,电子计算机包括电子模拟计算机和电子数字计算机。前者是由模拟电子器件(如模拟加法器、减法器、乘法器、除法等)构成,它早已完成了其历史使命,再无人提及。因此,本书所描述的就是电子数字计算机,简称计算机。电子数字计算机是指能对离散逻辑符号表示的数据或信息进行自动处理的电子装置。

►► 1.1.1 发展经历

1. 第 0 代:机械计算器

早在二千多年前,我们的祖先就发明了算盘,这是最简单实用的机械计算器。在漫长的历史中,算盘传遍了全世界,为人类社会做出了巨大贡献。

近代,人们发明了计算尺及手摇机械计算器。在电子计算器(机)普及之后,它们也逐步退出了历史舞台。

2. 第 1 代:电子管计算机

第 1 代计算机所在的年代为 1946—1954 年。这一代计算机由电子管、电磁继电器等器件构成。软件主要使用机器语言编程。典型机器有 ENIAC、IBM 701 等。

3. 第 2 代:晶体管计算机

第 2 代计算机所在的年代为 1955—1964 年。这一代计算机由晶体管、磁芯存储器等构成。在软件方面有了监控程序对计算机进行管理,并且开始使用高级语言。这个时期的计算机有很多种,如 IBM 7030、Univac LARC 等。

4. 第 3 代:集成电路计算机

这一代计算机所在的年代为 1965—1974 年。这一代计算机由小规模及中规模集成电路芯

片、多层印制电路板及磁芯存储器等构成。在软件方面,高级语言迅速发展并出现了分时操作系统。在这个时期,计算机的应用领域不断扩展,开始向国民经济各部门及军事领域渗透。典型机器如 IBM 360、370,DEC PDP-8 等。

5. 第4代:超大规模集成电路计算机

1975—1990年出现的是第4代计算机,这一代计算机由大规模、超大规模集成电路构成,其主存也是由超大规模集成电路构成的半导体存储器来实现。这一代计算机在结构上有了很大的发展,在性能上有了很大的提高。

同时,在这一时期由于微细加工技术的发展、超净环境的实现、超纯材料的研制成功,超大规模集成技术得到了迅速发展。于是,出现了依赖于这种技术的微型计算机、单片微型计算机等。

在硬件发展的同时,这一代计算机的软件飞速发展。出现了许多著名的操作系统,如 DOS、Windows、UNIX 等。

这一时期出现了一些典型的计算机,如 IBM 3090、VAX 9000 等。而这一时期应用最多、最广的还是个人微型计算机,诸如 IBM PC 兼容机、苹果 Macintosh 机等。

6. 第5代:高性能智能计算机

从1991年开始,进入了计算机发展的第5代,这一代计算机采用超大规模、超高速集成电路构成。在结构上,计算机的构成已从单处理器向多处理器发展。即使构成微型机也采用多核处理器,目前常见的是双核处理器和4核处理器。用这样的多核处理器构成计算机可获得很高的性能。此前,Intel公司已可以做出一块芯片,内含80个核心的多核处理器,用这样的一块80核处理器芯片构成的计算机,其运算速度已超过每秒10 000亿次。

可以想象,若用几百、几千甚至上万块双核(或多核)处理器芯片构成一台计算机,如机群系统,那么,这种计算机系统的性能将是非常优异的。例如,目前用这种思路做出来的计算机系统,其运算速度可达到每秒1 600万亿次。

第5代计算机不仅在速度等性能上不断提高,而且希望计算机更加人性化,具备能听、会看、会说等技能。

第5代计算机的发展必定对软件提出更高的要求,因此也必然会促使包括操作系统、应用软件等在内的各种软件的快速发展。

►► 1.1.2 摩尔定律

1. 摩尔定律的由来

1965年4月,《电子学》杂志刊登了戈登·摩尔(Gordon Moore)撰写的一篇文章。戈登·摩尔当时是飞兆半导体公司研发部门的主管。文章中讲述了他将50个晶体管集成在一块芯片中;并且预言,到1975年,就可能将6.5万只这样的元件密植在一块芯片上,制成高度复杂的集成电路。

当时,集成电路问世才6年,摩尔的预测听起来不可思议。但那篇文章的核心思想——即预

测集成电路芯片内可集成的晶体管数量差不多每年可增加一倍,在后来的技术发展过程中证明是正确的。现在人们根据几十年走过的技术历程将“摩尔定律”描述为:集成电路芯片的集成度每 18 个月翻一番。经过了四十多年,到今天摩尔定律依然有效,而且许多人确信该定律在未来很多年仍将成立。

摩尔的预言不仅对他本人,而且对整个社会而言都是意义深远的。后来摩尔与他人共同成立了 Intel 公司,并通过他所开创的技术创造了巨额财富。

摩尔定律并不是一个物理定律(定律是放之四海而皆准),而是一种预言。但它鞭策着工业界不断改进,并努力去实现它。从基本上讲,摩尔定律是一种自我激励的机制,它让人们无法抗拒,并努力追赶。从人们认识摩尔定律开始,无论是 Intel 公司、Motorola 公司还是其他的半导体器件公司,无一不是在不断地努力去实现摩尔定律,促使各厂家不断地推出集成度更高的产品。在 20 世纪 90 年代中期,Intel 公司利用 350 nm 技术制造出集成了 120 万只晶体管的 80486,但很快线宽就实现了 250 nm、180 nm、130 nm、90 nm、65 nm。今天,已经可以用 45 nm 的生产线制造处理器、DRAM 等器件。今天的处理器的集成度已超过 10 亿。根据摩尔定律,芯片的集成度还会迅速提高。有人曾经说过,集成度提高 100 倍,则相对价格可以降至 1/100,性能可以提高 100 倍,可靠性也可以提高 100 倍。当然,也许不一定是 100 倍,但是,随着集成度的提高,性能及可靠性都会大大提高、价格会大大降低将是毋庸置疑的。正是摩尔定律使得计算机日新月异地发展。

归根结底,历时四十多年的实践证明摩尔定律有利于工业的发展及人类的需求。直至今日,半导体工业还是按照 DRAM 每 18 个月、微处理器每 24 个月集成度翻倍的规律发展着。

如果按照旧有方式制造电路,即将晶体管、电阻器和电容器安装在电路板上构成电子设备,那么个人计算机、移动电话、计算机辅助设计软件等都不可能问世。

2. 摩尔定律的未来

2004 年以后,半导体工业无论从技术上还是成本上的挑战都越来越激烈,各种关键问题综合在一起,而且几乎要同时得到解决。如硅片尺寸的继续扩大,新材料、新工艺和新电路结构的采用,都使得工业制造难度显著提高。同时为了实现以上的要求,固定资产的投入必然加大,这就导致电路成本和市场风险急剧增加。

晶体管是处理数字信号的微电子开关。其中栅极用来打开或闭合晶体管,而栅极介质是用来将栅极从电流通道隔离出来的绝缘体底层。过去数十年来,芯片厂商一直采用二氧化硅作为栅极介质,采用多晶硅作为栅极。但是,使用现有的材料,晶体管的尺寸缩小几乎已达到极限,如 Intel 公司在其 65 nm 制程技术中,已经成功地 will 将二氧化硅栅极的厚度缩小至 1.2 nm(相当于 5 个原子的厚度),但是栅极的漏电流也随之迅速增加,这就导致功耗和发热等问题日益严重,让芯片厂商头痛不已。

在最近的处理器中,随着刻蚀线宽从 65 nm 过渡到 45 nm,晶体管的尺寸愈来愈小,其漏电及功耗问题越来越严重,性能和功耗之间的平衡也越来越困难。这是过去近十年里摩尔定律遇到的最大的技术难题,这也引起了人们的疑虑,芯片的集成度还能继续增加吗?摩尔定律是否已走

到了尽头?

为了顺利实现处理器制程由现在的 65 nm 向 45 nm 转变, Intel 公司采用了新的基于铪的高-K 绝缘体材料, 用来取代二氧化硅和多晶硅两种化合物, 从而可以大大减少电流的泄漏。Intel 公司认为, 新材料和新技术的使用可以提升处理器芯片的性能, 并能达到降低功耗、减小噪声和降低成本的要求。由于 45 nm 晶体管远小于上一代晶体管, 因此, 晶体管开关所需电量也大为减少, 使晶体管功耗大约降低了 30%, 栅漏电流减少至原来的十几分之一。随着晶体管新材料的改进, 既可以通过大幅提升 CPU 主频来实现更高性能, 也可以通过降低主频来提升能源效率。

同时, 微细加工技术的线宽仍在不断减小, 目前生产线上所用的是 45 nm 线宽, 而且 30 nm 线宽也已开始应用。据报道, 当前实验室所能做到的最小线宽为 9 nm。

最新发展的纳米技术使得制造性能最佳的计算机成为可能, 而这已经突破了硅技术所能达到的极限。这项新技术通过以碳为成分的纳米管来制造元件, 而纳米管的直径只有 10 个原子那么大, 是现在的基于硅成分的晶体管体积的 1/500。

该技术的出现向制造分子级电子设备的目标迈进了一大步。研究显示以碳为主要成分的纳米管在性能上不会逊色于硅晶体管, 还因为它们的体积要小得多, 所以有很大希望成为将来纳米电子技术的基础。纳米技术的前景非常广阔, 这样的技术会使摩尔定律得以延续。

可以看到, 新技术、新材料的出现, 使芯片的集成度进一步提高成为可能。尽管集成度不可能无限地增加下去, 但预计今后 10 年、20 年甚至 30 年, 芯片的制造依旧会继续遵循摩尔定律发展。摩尔定律将会继续激励人们向更高的高度攀登。

1.2 计算机的基本组成

▶▶ 1.2.1 硬件系统

早期的冯·诺依曼计算机由两大部分——硬件系统和软件系统组成。硬件系统是指计算机中那些看得见、摸得着的物理实体。

(1) 硬件组成

早期的计算机硬件结构如图 1.1 所示。图 1.1 所示的计算机结构最早是在 1946 年由冯·诺依曼提出的。同时, 在此硬件结构的基础上, 冯·诺依曼还提出了计算机是依据存储程序、程序控制的方式工作。这就是冯·诺依曼的计算机设计思想。

当时的计算机由运算器、控制器、存储器、输入设备和输出设备 5 大部分构成。运算器用来实现算术运算和逻辑运算; 控制器根据指令的功能产生相应的控制信号, 控制其他部分的工作以便实现指令的功能; 存储器用来存放数据和程序; 输入设备可将外部的信息输入到计算机中; 输出设备可将结果显示或记录下来。

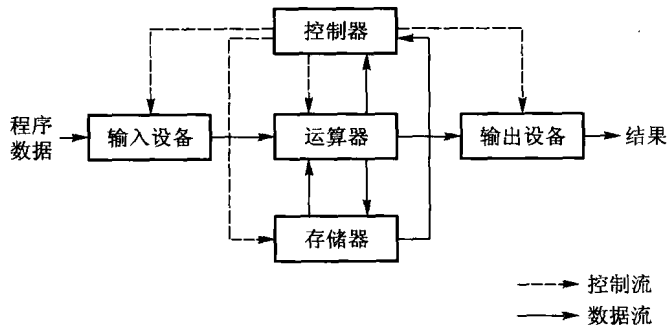


图 1.1 早期计算机(硬件)的组成

(2) 冯·诺依曼计算机的特点

冯·诺依曼计算机工作的基本思想就是将计算机要处理的问题用指令编成程序,并将程序存放在存储器中;而后,在控制器的控制下,将指令从存储器中逐条取出来执行;通过执行程序最终解决计算机所要处理的问题。尽管经历了几十年的发展,又出现了新的设计思想,但冯·诺依曼的这种思路直到今天仍然还在广泛地应用。

在冯·诺依曼计算机的工作过程中,总是一条指令接一条指令地执行,执行指令会产生控制流,在控制流的驱动下完成指令的功能。在此过程中,数据(流)则是被动地调用。

冯·诺依曼计算机的另一特点就是进入计算机的指令、数据及其他信息均是用二进制编码来表示的。使用二进制编码不仅电路简单、使用方便而且抗干扰性强。因此,二进制编码一直沿用至今。

(3) PC 的结构

计算机的发展已走过了 60 多年,尤其是在最近 30 年,其发展更是日新月异。本书的后面将会介绍一系列的计算机结构形式,从中可以看到,在过去的年代里计算机从最简单的冯·诺依曼计算机到现在都有哪些发展与改变,这些发展和改变是由什么原因引起的,是如何实现这些发展和改变的,实现的依据又是怎样的。这些问题正是本书要呈现给读者的内容。

另外,应当特别提及的就是个人计算机——PC。从 1981 年 PC 诞生以来,由于其规模小、结构简单,人们称之为微型计算机。在 20 多年的时间里,PC 一代接一代地发展着。现在,PC 已遍布全世界。尽管今天的 PC 功能已十分强大,但人们仍然称之为微型计算机。目前读者所使用的、所看到的绝大多数都是 PC。

早期的 PC 结构非常简单,可以简化为如图 1.2 所示的框图。该框图也是那个年代微型计算机的基本结构形式。

由图 1.2 可以看到,中央处理器(CPU)是将图 1.1 中的运算器(算术逻辑部件,ALU)和控制器(还有一些寄存器)集成在一块集成电路芯片中。在执行指令的过程中,它可以产生地址信号、数据信号和一系列控制信号,从而形成系统总线信号。图 1.2 中的存储器也由大规模集成的半导体存储器芯片构成。外设包括输入设备、输出设备及外部存储器,均经过接口与系统总线相

连接。因此,图 1.2 与图 1.1 已有很大的不同,它们是不同的时代的计算机。同时,处理器厂家将图 1.2 中用虚线框起的部分集成在一块芯片中,这就是单片微型计算机。

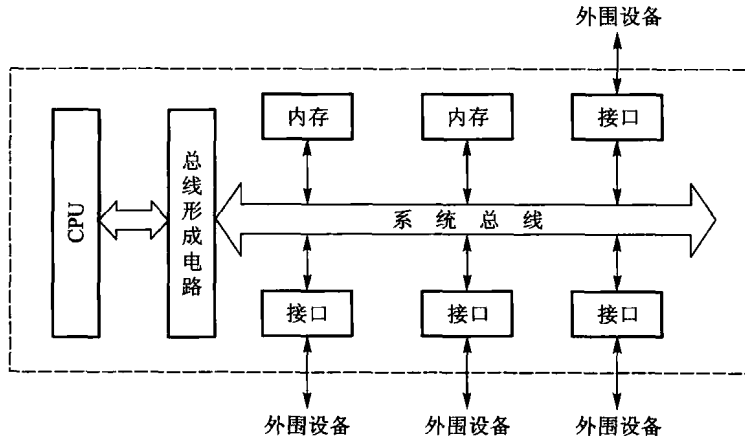


图 1.2 微型计算机结构框图

在摩尔定律的激励下,半导体集成电路,包括 CPU、半导体存储器、接口芯片以及外设所用到的芯片的集成度均以每 18~24 个月提高 1 倍的速度飞速发展。尤其值得说明的是 CPU,它也被简称为处理器。各处理器厂家不遗余力地、一代接一代、一种型号接一种型号地被开发研制出来。在这里仅将 Intel 公司的 80x86 系列处理器的发展历程列于表 1.1 中。由于本书篇幅的限制,表中并未列出各处理器的集成度及性能指标。

表 1.1 Intel 微处理器发展历程

年 份	产 品	年 份	产 品
1971 年	4004 微处理器	1999 年	Celeron 处理器
1972 年	8008 微处理器	1999 年	Pentium III 处理器
1974 年	8080 微处理器	1999 年	Pentium III Xeon 处理器
1978 年	8086、8088 微处理器	2000 年	Pentium 4 处理器
1982 年	80286 微处理器	2001 年	Xeon 处理器
1985 年	80386 微处理器	2001 年	Itanium 处理器
1989 年	80486 DX 微处理器	2002 年	Itanium 2 处理器
1993 年	Pentium 处理器	2003 年	Pentium M 处理器
1995 年	Pentium Pro 处理器	2005 年	Pentium D 处理器
1997 年	Pentium II 处理器	2006 年	Core 2 Duo 处理器
1998 年	Pentium Pro Xeon 处理器		

在处理器的发展过程中,过去用于大型计算机的许多技术,在摩尔定律的支持下不断地运用到处理器的发展过程中。例如,超标量、超流水、多级 Cache、虚拟技术、RISC、多机系统概念下的多核处理器等技术均已应用在近几年的处理器中。这些技术正是本书中所要讨论的重要内容。

随着处理器性能的不断提高,从 20 世纪 90 年代开始,PC 的结构发生了很大的变化。基本上是以芯片组为基础构成主板,而后插上相关的部件构成 PC。主板芯片组有 1~3 块芯片,经常采用两块芯片,分别称为北桥芯片和南桥芯片。随着处理器的不断更新换代,芯片组也在不断地更新换代。一般是每开发出一种新的处理器,就会开发出与之相对应的芯片组。目前常见的 PC 主板结构框图如图 1.3 所示。

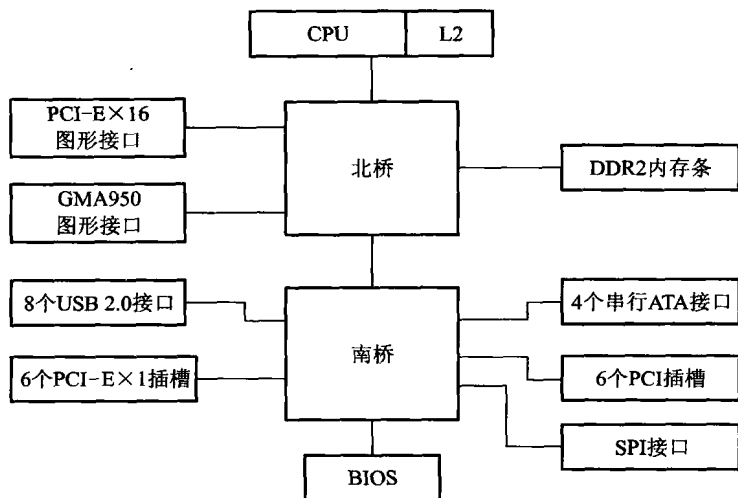


图 1.3 PC 主板结构框图

在图 1.3 中,北桥芯片是主板芯片组中最重要的组成部分,也称为主桥。北桥芯片主要决定主板的规格、对硬件的支持以及系统的性能,它连接着 CPU、内存、AGP 总线。主板支持哪种处理器,支持何种显卡,支持什么频率的内存条,都是北桥芯片决定的。北桥芯片往往有较高的工作频率,所以发热量较大,故需要为北桥加一个散热器。

南桥芯片主要决定主板的接口,该芯片所提供的各种接口(如串行接口、USB 接口)、PCI 总线、串行 ATA 接口(接硬盘、光驱)、PCI-E 总线(接声卡、RAID 卡、网卡等),都由南桥芯片控制。

南北桥间随时进行数据传递,需要一条通道,这条通道就是南北桥总线。南北桥总线越宽,数据传输越便捷。

目前的北桥都支持双核甚至 4 核等性能较高的处理器。

随着技术的发展和摩尔定律的继续生效,PC 这种双桥结构可望在不久的将来消失。到那时,处理器、北桥、南桥以至于内存条会集成在一块集成芯片中,构成所谓的片上系统(System on Chip, SoC),也有人称之为系统芯片。届时,PC 可能只需要一块片上系统,在此芯片上接上外设就可以构成。为此,在本书的最后一章将向读者介绍如何设计处理器及片上系统。

►► 1.2.2 软件系统

对计算机而言,只有上面提到的硬件系统,计算机是不能工作的,必须配上软件计算机才能工作。

计算机软件通常是指计算机所配置的各类程序和文件,由于它们是存放在内存或外存中的二进制编码信息,不能直接触摸而且修改相对比较容易,故称之为软件。在计算机系统中,各种软件相互配合,很好地支持计算机有条不紊地工作,这一系列软件就构成了计算机的软件系统。软件系统一般包括两大部分:系统软件和应用软件。

1. 系统软件

系统软件是一系列保障计算机良好运行的程序的集合。它们的功能是对系统的各种资源(硬件和软件)进行管理和调度,使计算机有条不紊地工作,为用户提供有效服务,充分发挥其效能。系统软件包括如下几类。

(1) 操作系统

操作系统是最重要的系统软件,它是管理计算机硬、软件资源,控制程序运行,改善人机交互并为应用软件提供支持的一种软件。通常,操作系统包括5大功能:处理器管理、存储管理、文件管理、设备管理及作业管理。

(2) 语言处理程序

每一台计算机都会配置多种语言以便于用户使用,包括各种高级语言及汇编语言等。当用户使用某种语言编写程序后,在该语言编译程序的支持下,可将用户的源程序转换为目标程序。

(3) 通用程序

在操作系统下提供给用户使用的一些通用程序。最常见的如在 Windows 下的“记事本”、“画图”等。

(4) 各种服务支持软件

这是一些帮助用户使用和维护计算机的软件。例如,各种调试程序、诊断程序、提示警告程序等。

2. 应用软件

应用软件是指用户在各自的应用中,为解决自己的有关问题而编写的程序。这是一类直接以用户的需求为目标的程序。由于用户的多样性(各行各业、各种部门)和用户需求多样性,使得这类软件也具有多样性。例如,用于科学计算、信息管理、过程控制、武器装备等方面的应用软件。

1.3 计算机系统的组织结构

前面内容可以看到,计算机是由硬件、软件系统构成的复杂的电子系统。不同的人,从不同的角度,出于不同的目的去理解和描述的计算机也是不一样的。本节将说明在不同的层次上对