



“十二五”普通高等教育本科国家级规划教材



普通高等教育“十一五”国家级规划教材

21世纪大学本科计算机专业系列教材



袁春风 主编

杨若瑜 王帅 唐杰 编著

计算机组成与系统结构 (第2版)

- 根据教育部“高等学校计算机科学与技术专业规范”组织编写
- 与美国 ACM 和 IEEE CS *Computing Curricula* 最新进展同步
- 教育部—微软精品课程教材
- 远程教育国家精品课程教材



清华大学出版社



“十二五”普通高等教育本科国家级规划教材



普通高等教育“十一五”国家级规划教材

21世纪大学本科计算机专业系列教材

计算机组成与系统结构 (第2版)

袁春风 主编

杨若瑜 王帅 唐杰 编著



清华大学出版社

北京

内 容 简 介

本书主要介绍计算机组成与系统结构涉及的相关概念、理论和技术内容,主要包括指令集体系结构、数据的表示和存储,以及实现指令集体系结构的计算机各部件的内部工作原理、组成结构及其相互连接关系。全书共9章:第1章对计算机系统及其性能评价进行概述性介绍,第2、3章主要介绍数据的机器级表示、运算以及运算部件的结构与设计,第4~6章介绍指令系统以及各种CPU设计技术,第7章主要介绍包含主存、cache和虚拟存储器在内的存储器分层体系结构,第8章介绍总线互连以及输入输出系统,第9章介绍并行处理计算系统的基本硬件结构和并行程序设计编程模型。

本书内容详尽,概念清楚,通俗易懂,实例丰富,并提供大量典型习题供读者练习。本书可以作为计算机专业本科或大专院校学生计算机组成原理与系统结构课程的教材,也可以作为有关专业研究生或计算机技术人员的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

计算机组成与系统结构/袁春风主编. —2版. —北京:清华大学出版社,2015

21世纪大学本科计算机专业系列教材

ISBN 978-7-302-40864-2

I. ①计… II. ①袁… III. ①计算机组成原理—高等学校—教材 ②计算机体系结构—高等学校—教材 IV. ①TP30

中国版本图书馆CIP数据核字(2015)第153330号

责任编辑:张瑞庆 战晓雷

封面设计:常雪影

责任校对:李建庄

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市市中晟雅豪印务有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:24.5

字 数:608千字

版 次:2010年4月第1版 2015年8月第2版

印 次:2015年8月第1次印刷

印 数:1~2000

定 价:44.50元

产品编号:065354-01

21 世纪大学本科计算机专业系列教材编委会

名誉主任：陈火旺

主任：李晓明

副主任：钱德沛 焦金生

委员：(按姓氏笔画排序)

马殿富 王志英 王晓东 宁 洪 刘 辰

孙茂松 李大友 李仲麟 吴朝晖 何炎祥

宋方敏 张大方 张长海 周兴社 侯文永

袁开榜 钱乐秋 黄国兴 蒋宗礼 曾 明

廖明宏 樊孝忠

秘 书：张瑞庆

计算机组成(computer organization)是指计算机主要功能部件的组成结构、逻辑设计及功能部件间的相互连接关系。计算机系统结构(computer architecture)的经典定义是指程序设计者(主要指低级语言程序员或编译程序设计者)所看到的计算机系统的属性,即计算机的功能特性和概念性结构,也称指令集体系结构(Instruction Set Architecture, ISA),包括数据类型及数据格式,指令格式,寻址方式和可访问空间大小,程序可访问的寄存器个数、位数和编号,控制寄存器的定义,I/O空间的编址方式,中断结构,机器工作状态的定义和切换,输入输出数据传送方式,存储保护方式等。

本书主要介绍计算机组成与系统结构涉及的相关内容。在计算机系统层次结构中,这些内容位于软件和硬件的结合处,不仅涉及计算机硬件设计和指令系统设计,还涉及操作系统、编译程序和程序设计等部分软件设计技术,是整个计算机系统中最核心的部分。

1. 本书的写作思路和内容组织

计算机组成与系统结构这两部分涉及的内容相互融合,密不可分。无论是国内还是国外,很多高校都逐渐把计算机组成原理和系统结构课程的内容有机结合起来;甚至国外一些经典教材还把与之密切相关的软件设计的内容也融合进来。这种方式可以加深读者对计算机软、硬件系统的整体化理解,并有效地增强对学生的计算机系统设计能力的培养。

本书在总结和借鉴国外著名高校使用的教材、教案、教学理念和教学方法的基础上,力图以“培养学生现代计算机系统设计能力”为目标,贯彻“从程序设计视角出发、强调软硬件关联与协同、以CPU设计为核心”的组织思路,试图改变国内同类教材通常的就硬件讲硬件、软硬件分离的传统内容组织方式,以系统化观点全面地介绍计算机组成和系统结构的相关知识和技术。

为了体现以上目标和思路,本书在以下几个方面进行了重点考虑和内容组织:

(1) 首先基于“高级语言程序→汇编语言程序→机器指令序列→控制信号”的路线展现程序从编程设计、转换翻译到最终在CPU上运行的整个过程;在此基础上,用计算机系统层次化的观点阐述计算机组成与系统结构课程在整个计算机系统的位置、内容和作用,从而为清晰了解本课程的内容和作用、为全面建立计算机软件和硬件系统的整体概念打下基础。

(2) 将指令执行过程和异常、中断、存储访问、I/O访问等重要概念和技术结合起来进行介绍,力求清晰地说明CPU执行指令过程中硬件与操作系统相互切换和协同工作的处理过程,使读者深刻理解软硬件系统之间的关系与协同工作过程。

(3) 在讲述与程序设计有密切关系的体系结构内容(如数据表示、信息存放、操作数寻址、过程调用、程序访问局部性等)时,试图通过对硬件设计与程序设计的关系的说明,使读者建立“从程序员视角理解计算机硬件系统设计,从硬件设计的视角理解程序设计与执行”的思想,力图在提高读者硬件设计能力的同时,也增强其进行高效的和系统化的程序设计的能力。

(4) 在国内“计算机组成原理”教材传统内容基础上增加指令流水线设计的详细内容,依照“最简单的 IAS 计算机 CPU→总线式 CPU→单周期 CPU→多周期 CPU→基本流水线 CPU→动态超标量超流水线 CPU”的次序,循序渐进地介绍 CPU 设计技术及其发展过程,以 MIPS 处理器和 Pentium 4 处理器为蓝本,力图使读者全面深入地掌握现代计算机的 CPU 设计技术。

(5) 结合指令流水线技术介绍基于流水线的编译优化技术,使读者对编译技术与指令流水线实现技术之间的密切关系有一定的认识和理解。

2. 各章节主要内容

本书共有 9 章,各章主要内容如下:

第 1 章(计算机系统概述)主要介绍冯·诺依曼结构的特点、计算机硬件的基本组成、计算机软件设计和执行过程、计算机系统层次结构以及系统性能评价方法。

第 2 章(数据的机器级表示)主要介绍无符号数和带符号整数的表示、IEEE 754 浮点标准、西文字符和汉字的编码表示、大端/小端存放顺序及对齐方式以及常用检/纠错码表示与使用方法等。

第 3 章(运算方法和运算部件)主要介绍各类定点数和浮点数的运算方法和相应的运算部件,以及核心运算部件 ALU 的功能和设计实现等。

第 4 章(指令系统)主要介绍高级语言与低级语言的关系、指令格式、操作数类型、寻址方式、操作类型、硬件对过程的支持、用户程序在虚存空间的配置和划分等技术。

第 5 章(中央处理器)主要介绍 CPU 的基本功能和内部结构、指令执行过程、数据通路的基本组成和定时、单周期和多周期数据通路、硬连线路和微程序控制器、异常和中断等概念和技术。

第 6 章(指令流水线)主要介绍指令流水线的基本原理、流水段寄存器的概念、流水线数据通路的设计、流水线的控制信号、结构冒险及其处理、数据冒险及其处理、转发技术、控制冒险及其处理、分支预测原理、超标量和动态流水线的概念和技术。

第 7 章(存储器分层体系结构)主要介绍存储器分层结构的概念、半导体存储器的组织、多模块存储器、cache 的基本原理、cache 和主存间的映射关系及替换算法、虚拟存储器的基本概念、页表结构、缺页异常、TLB 的实现技术。

第 8 章(互连及输入输出组织)主要介绍常用输入输出外设和磁盘存储器的工作原理、现代计算机内部互连结构、I/O 接口的结构、I/O 端口编址方式、程序查询 I/O 方式、中断 I/O 方式、DMA 方式以及内核 I/O 软件概述等。

第 9 章(并行处理系统)主要介绍 UMA(对称)多处理器、CC-NUMA 多处理器、片级多处理器和多线程技术、集群多计算机、网格多计算机、向量处理机和 SIMD 技术,还对并行处理编程模型进行了概述。

3. 第 2 版修订内容

近几年来,计算机组成与系统结构领域发生了一些变革,一些技术有了新的发展。为了

反映技术的进步,拓宽本领域知识的覆盖面,并更加合理地构建知识框架,本书第2版在第1版的基础上进行了若干修订。

首先,对第1版的章节顺序进行了一些调整。因为数据的运算、指令系统、中央处理器以及指令流水线的关系比较密切,所以将第1版中原来的第4章存储器层次结构的内容调到了指令流水线的后面。

此外,在内容上对一些知识点进行了增、删、改,主要包括以下几个方面。

(1) 程序的开发和执行过程。第1章中增加了对hello程序执行过程的描述,并举例对程序和指令的关系进行了说明,同时还增加了一个冯·诺依曼结构模型机,利用该模型机对指令的执行过程进行了详细讲解。

(2) 运算方法和运算部件。对第3章进行了较大的调整。为了突出ALU中算术运算所生成的标志信息的作用,增加了对带标志加法器的描述;简化了对单级先行进位加法器、两级先行进位加法器、原码加减法、移码加减法、原码两位乘法、补码两位乘法、原码除法、十进制加减运算等的描述。

(3) 标志信息的生成与使用。第4章中增强了对标志信息处理问题的说明,增加了IA-32指令系统中的条件转移指令功能描述,结合第3章中ALU生成的标志信息,比较透彻地说明了标志信息的生成与使用问题。

(4) 指令系统举例。第4章中保留了目前最流行的PC处理器指令系统IA-32,增加了嵌入式处理器指令系统ARM,并将Intel处理器架构中的MMX、SSE等SIMD指令集移到了新增加的第9章中,作为并行处理技术的一部分来介绍。

(5) DRAM芯片技术。第7章中增加了对DDR2和DDR3 SDRAM芯片的描述。

(6) 存储器芯片的扩展及其与CPU的连接。对第7章中的这部分内容进行了大幅度修订,删去了陈旧的内容和举例,增加了目前系统中实际使用情况的描述。

(7) cache结构举例。第7章中增加了新的多核处理器芯片Core i7中cache结构的举例。

(8) 系统总线。系统内部的互连技术发展非常快,很多传统的概念已经过时,为此,对第1版中第8章(系统总线)的内容进行了大幅度修改,并将其合并到后面的输入输出组织一章中。总线裁决、总线定时、PCI总线标准和SCSI接口等内容被取消或弱化。增加了有关PCI-Express、QPI等互连技术的介绍。

(9) 外部设备。第8章中删除或弱化了CRT显示器、磁带存储器和光盘存储器等内容。

(10) I/O接口。第8章中删除有关I/O接口的分类、并行接口和串行接口、I/O接口的举例的内容。

(11) I/O控制方式。第8章中删除了关于通道方式和I/O处理器方式的内容。增加了对I/O子系统的介绍,简要描述了从高级语言层到I/O设备层之间的一个转换过程。

(12) 并行处理系统。增加了一章专门介绍各类并行处理系统的硬件结构以及涉及的硬件互连、软件通信及并行编程等技术问题。

4. 关于本书使用的一些建议

本书可作为“计算机组成原理”课程的教材,也可作为“计算机组成原理实验”课程和“计算机系统结构”课程的教学参考书。特别是对于不专门开设“计算机系统结构”课程的院校,使用本书作为“计算机组成与系统结构”课程的教材是比较合适的。

对于本书的使用,具体建议如下:

(1) 课堂教学应以主干内容为主,力求完整地给出知识框架体系,并着重讲清楚相关概念之间的联系。

(2) 标注为*的内容是可以跳过而不影响阅读连贯性的部分,主要有以下三类:简单易懂的基础性内容、具体实现方面的细节内容和在技术层面上更加深入的内容。这些内容对深入理解课程的整体核心内容是非常有帮助的。因此,在课时允许的情况下,可以选择其中的一部分进行课堂讲解;在课时不允许的情况下,也尽量安排学生进行课后阅读。

(3) 书中每个重要的知识点和概念后面都有一些例子,可选择部分重要的、难懂的例子在课堂上讲解,而大部分可留给学生自学。

(4) 习题中列出的概念术语基本涵盖了相应章节的主要概念,可以让学生对照检查是否全部清楚其含义;习题中列出的简答题是相应章节重要的基本问题,可以通过对照检查以判断学生对相应章节内容的掌握程度;对于综合运用题,如果与程序设计相关,则可用编程方式来求解或验证,这样做,对学生深刻理解课程内容有帮助。

(5) 本书在CPU设计方面给出了比较具体的实现方案,相关内容可以作为基于FPGA和硬件描述语言进行CPU设计实验的参考资料。

(6) 书后给出了部分国际一流大学相关课程网站网址,可以到这些网站找到课堂讲义、习题参考答案以及更多的相关教辅资料。

5. 作者及致谢

参与本书编写工作的有袁春风、杨若瑜、王帅、唐杰、杨杰和樊士庆。全书由袁春风担任主编,负责全书内容的组织、统稿和编审。

在本书的编写过程中,得到了张福炎教授的悉心指导;黄宜华教授从书稿的篇章结构到内容各方面都提出了许多宝贵的意见,进行了修改,并对全书内容进行了全面细致的审核和校对;书中有关CPU设计的最初图稿和内容组织思路由陈贵海教授提供;此外,武港山教授、杨若瑜、俞建新、吴海军、张泽生、蔡晓燕等老师也对本书提出了许多宝贵的意见;杨晓亮、肖韬、翁基伟、刘长辉、宗恒、莫志刚、叶俊杰等研究生对相关章节的内容和习题分别进行了校对和试做,并提出了许多宝贵的意见和修改建议。在此对以上各位老师和研究生一并表示衷心的感谢。

本书是作者在南京大学从事“计算机组成与系统结构”课程教学20多年来所积累的讲稿内容的基础上编写而成的,感谢各位同仁和各届学生对讲稿内容所提出的宝贵的反馈和改进意见,使得本教材的内容得以不断地改进和完善。

6. 结束语

本书广泛参考了国内外相关的经典教材和教案,在内容上力求做到取材先进并反映技术发展现状;在内容的组织和描述上力求概念准确、语言通俗易懂、实例深入浅出,并尽量利用图示和实例来解释和说明问题。但是,由于计算机组成与系统结构相关的基础理论和技术在不断发展,新的思想、概念、技术和方法不断涌现,加之作者水平有限,在编写中难免存在不当或遗漏之处,恳请广大读者对本书的不足之处给予指正,以便在后续的版本中予以改进。

作者

2015年6月于南京

目 录

CONTENTS

第 1 章 计算机系统概述	1
1.1 计算机的发展历程	1
*1.1.1 通用电子计算机的诞生	1
*1.1.2 第一代计算机	2
*1.1.3 第二代计算机	2
*1.1.4 第三代计算机	3
*1.1.5 第四代计算机	3
1.2 计算机系统的基本组成	4
1.2.1 计算机硬件	4
1.2.2 计算机软件	8
1.3 计算机系统的层次结构	8
1.3.1 计算机系统抽象层的转换	9
1.3.2 计算机系统的不同用户	10
1.4 程序开发与执行过程	12
1.4.1 从源程序到可执行程序	12
1.4.2 可执行文件的启动和执行	14
1.4.3 程序与指令的关系	15
1.4.4 指令的执行过程	16
1.5 计算机系统性能评价	17
1.5.1 计算机性能的定义	18
1.5.2 计算机性能测试	18
1.5.3 用指令执行速度进行性能评估	20
1.5.4 用基准程序进行性能评估	21
1.6 本章小结	22
习题	23
第 2 章 数据的机器级表示	25
2.1 数制和编码	25

*2.1.1	信息的二进制编码	25
*2.1.2	进位记数制	27
2.1.3	定点与浮点表示	31
2.1.4	定点数的编码表示	31
2.2	整数的表示	36
2.2.1	无符号整数的表示	36
2.2.2	带符号整数的表示	36
2.2.3	C语言中的整数类型	36
2.3	实数的表示	38
2.3.1	浮点数的表示格式	38
2.3.2	浮点数的规格化	39
2.3.3	IEEE 754 浮点数标准	39
2.3.4	C语言中的浮点数类型	44
2.4	十进制数的表示	45
*2.4.1	用 ASCII 码字符表示	45
*2.4.2	用 BCD 码表示	46
2.5	非数值数据的编码表示	46
2.5.1	逻辑值	46
*2.5.2	西文字符	47
*2.5.3	汉字字符	48
2.6	数据的宽度和存储	50
2.6.1	数据的宽度和单位	50
2.6.2	数据的存储和排列顺序	51
2.7	数据校验码	54
2.7.1	奇偶校验码	55
2.7.2	海明校验码	56
2.7.3	循环冗余校验码	60
2.8	本章小结	62
	习题	62
第3章	运算方法和运算部件	66
3.1	高级语言和机器指令中的运算	66
*3.1.1	C程序中涉及的运算	66
*3.1.2	MIPS指令中涉及的运算	69
3.2	基本运算部件	70
3.2.1	串行进位加法器	71
3.2.2	并行进位加法器	72
3.2.3	带标志加法器	73
3.2.4	算术逻辑部件	74

3.3	定点数运算	75
3.3.1	补码加减运算	75
*3.3.2	原码加减运算	77
*3.3.3	移码加减运算	77
3.3.4	原码乘法运算	78
3.3.5	补码乘法运算	81
*3.3.6	快速乘法器	83
3.3.7	原码除法运算	85
*3.3.8	补码除法运算	90
3.4	浮点数运算	93
3.4.1	浮点数加减运算	93
*3.4.2	浮点数乘除运算	97
3.5	运算部件的组成	100
*3.5.1	定点运算部件	100
*3.5.2	浮点运算部件	102
3.6	本章小结	103
	习题	104
第4章	指令系统	107
4.1	指令格式设计	107
4.1.1	指令地址码的个数	107
4.1.2	指令格式设计原则	108
4.2	指令系统设计	108
4.2.1	基本设计问题	109
4.2.2	操作数类型	110
4.2.3	寻址方式	110
4.2.4	操作类型	112
4.2.5	操作码编码	114
4.2.6	标志信息的生成与使用	116
4.2.7	指令系统设计风格	118
4.3	指令系统实例	120
*4.3.1	IA-32 指令系统	120
*4.3.2	ARM 指令系统	123
4.4	程序的机器级表示	125
4.4.1	MIPS 汇编语言和机器语言	125
*4.4.2	选择结构的机器代码表示	129
*4.4.3	循环结构的机器代码表示	130
*4.4.4	过程调用的机器代码表示	131
4.5	本章小结	138

习题	139
第5章 中央处理器	142
5.1 CPU概述	142
5.1.1 指令执行过程	142
5.1.2 CPU的基本功能和基本组成	143
5.1.3 数据通路的基本结构	144
5.2 单周期处理器设计	152
5.2.1 指令功能的描述	153
5.2.2 数据通路的设计	154
5.2.3 控制逻辑单元的设计	162
5.2.4 时钟周期的确定	168
5.3 多周期处理器设计	169
*5.3.1 信号竞争问题	169
*5.3.2 指令执行状态分析	170
*5.3.3 硬连线控制器设计	172
5.4 微程序控制器设计	175
5.4.1 微程序控制器的结构	175
*5.4.2 微命令编码	176
*5.4.3 微指令地址的确定	181
5.5 异常和中断处理	184
5.5.1 基本概念	184
5.5.2 异常处理过程	186
*5.5.3 带异常处理的处理器设计	187
5.6 本章小结	189
习题	190
第6章 指令流水线	194
6.1 流水线概述	194
6.1.1 流水线的执行效率	194
6.1.2 适合流水线的指令集特征	195
6.2 流水线处理器的实现	196
6.2.1 每条指令的流水段分析	196
6.2.2 流水线数据通路的设计	197
6.2.3 流水线控制器的设计	202
6.3 流水线冒险及其处理	203
6.3.1 结构冒险	203
6.3.2 数据冒险	205
6.3.3 控制冒险	211

6.4	高级流水线技术	219
*6.4.1	静态多发射处理器	220
*6.4.2	动态多发射处理器	223
*6.4.3	Pentium 4 处理器的流水线结构	227
6.5	本章小结	230
习题	232
第 7 章	存储器分层体系结构	235
7.1	存储器概述	235
7.1.1	存储器的分类	235
7.1.2	主存储器的组成和基本操作	236
7.1.3	存储器的主要性能指标	237
7.1.4	存储器的层次化结构	238
7.2	半导体随机存取存储器	239
7.2.1	基本存储元件	239
7.2.2	SRAM 芯片和 DRAM 芯片	240
7.2.3	SDRAM 芯片技术	244
7.3	存储器芯片的扩展及其与 CPU 的连接	245
7.3.1	内存条和内存条插槽	245
7.3.2	存储器芯片的扩展	247
7.4	半导体只读存储器和 Flash 存储器	248
7.4.1	半导体只读存储器	248
7.4.2	半导体 Flash 存储器	248
7.5	并行存储器结构技术	250
7.5.1	双口存储器	250
7.5.2	多模块存储器	250
7.6	高速缓冲存储器	252
7.6.1	程序访问的局部性	253
7.6.2	cache 的基本工作原理	255
7.6.3	cache 行和主存块之间的映射方式	257
7.6.4	cache 中主存块的替换算法	264
7.6.5	cache 的一致性问题的	267
*7.6.6	cache 性能评估	268
*7.6.7	影响 cache 性能的因素	269
*7.6.8	cache 结构举例	272
7.7	虚拟存储器	274
*7.7.1	进程与进程的上下文切换	274
*7.7.2	存储器管理	275
7.7.3	虚拟地址空间	277

7.7.4 虚拟存储器的实现	279
*7.7.5 存储保护	286
7.8 本章小结	287
习题	288
第8章 互连及输入输出组织	294
8.1 外部设备的分类与特点	294
8.1.1 外设的分类	294
8.1.2 外设的特点	295
8.2 输入设备和输出设备	295
*8.2.1 键盘	295
*8.2.2 鼠标器	297
*8.2.3 打印机	297
*8.2.4 显示器	299
8.3 外部存储设备	301
8.3.1 磁表面存储原理	301
8.3.2 硬磁盘存储器	303
8.3.3 冗余磁盘阵列	307
8.3.4 U盘和固态硬盘	309
8.4 外设与CPU、主存的互连	310
8.4.1 总线概述	310
8.4.2 基于总线的互连结构	312
8.5 I/O接口	317
8.5.1 I/O接口的功能	317
8.5.2 I/O接口的通用结构	317
8.5.3 I/O端口及其编址	319
8.6 I/O数据传送控制方式	320
8.6.1 程序直接控制I/O方式	320
8.6.2 程序中断I/O方式	323
8.6.3 DMA方式	333
8.7 I/O子系统概述	336
8.8 内核空间I/O软件	339
8.8.1 与设备无关的I/O软件	339
8.8.2 设备驱动程序	341
8.8.3 中断服务程序	342
8.9 本章小结	344
习题	345
*第9章 并行处理系统	349
9.1 并行处理系统概述	349

9.1.1	并行处理的主要技术问题	349
9.1.2	并行处理系统的分类	351
9.2	多处理器系统	353
9.2.1	UMA 多处理器结构	354
9.2.2	NUMA 多处理器结构	356
9.2.3	多处理器系统中的互连网络	358
9.2.4	片级多处理器和多线程技术	359
9.3	多计算机系统	360
9.3.1	集群多计算机系统	360
9.3.2	网格多计算机系统	364
9.4	向量处理机和 SIMD 技术	364
9.4.1	向量处理机	364
9.4.2	Intel 架构中的 SIMD 技术	366
9.4.3	GPU 架构	368
9.5	并行处理编程模式简介	371
9.6	本章小结	374
	参考文献	376

第 1 章

计算机系统概述

本章主要介绍计算机的发展历程、计算机系统的基本组成和基本功能、计算机系统层次化结构以及计算机系统的性能评价。

1.1 计算机的发展历程

* 1.1.1 通用电子计算机的诞生

世界上第一台通用电子计算机是 1946 年在美国诞生的 ENIAC,其设计师是美国宾夕法尼亚大学的莫齐利(Mauchly)和他的学生艾克特(Eckert)。莫齐利 1932 年获得著名的霍普金斯大学物理学博士学位并留校任教,1941 年转入宾夕法尼亚大学,他常常为物理学研究中屡屡出现的大量枯燥、烦琐的数学计算而头痛,渴望电子计算机帮忙。一天,他偶然发现爱荷华州立大学的阿塔纳索夫教授正在试制电子计算机,莫齐利深感鼓舞,立即启程拜访。阿塔纳索夫教授热情地接待了这位志同道合的不速之客,毫无保留地介绍了研制情况,并无私地把有关电子计算机设计的珍贵笔记本郑重地交给了莫齐利。莫齐利认真研究了阿塔纳索夫的方案,凭着特有的聪明才智,加上雄厚的数学和物理基础以及电子学方面的丰富实践经验,于 1942 年写出了一份题为《高速电子管装置的使用》的报告。该报告很快引起了一个年轻人——23 岁的研究生艾克特——的兴趣,于是,师生密切协作,开始了电子计算机的研制。当时正值第二次世界大战期间,军方急需一种高速电子装置来解决弹道的复杂计算问题,莫齐利与艾克特的方案在 1943 年得到了军方的支持。在冯·诺依曼等人的帮助下,他们经过两年多的努力,终于研制成了第一台电子计算机。1946 年 2 月,美国陆军军械部与摩尔学院共同举行新闻发布会,宣布了第一台电子计算机 ENIAC (Electronic Numerical Integrator and Computer,电子数字积分机和计算机)研制成功的消息。

ENIAC 能进行每秒 5000 次加法运算、每秒 400 次乘法运算以及平方和立方、 \sin 和 \cos 函数数值运算。当时主要用它来进行弹道参数计算,60 秒钟射程的弹道计算时间由原来的 20 分钟一下子缩短到仅需 30 秒。ENIAC 是个庞然大物,耗资 40 多万美元,使用了 18 000 个真空管,重 30 吨,占地面积 170m^2 ,耗电 150 千瓦,第一次开机时甚至使整个费城地区的照明都闪烁变暗。该机正式运行到 1955 年 10 月 2 日,这十年间共运行了 80 223 个小时。

自从第一台通用电子计算机 ENIAC 诞生后,人类社会进入了一个崭新的电子计算和信息化时代。计算机硬件早期的发展受电子开关器件的影响极大,为此,传统上人们以元器

件的更新作为计算机技术进步和划代的主要标志。

* 1.1.2 第一代计算机

第一代计算机(20世纪40年代中到20世纪50年代末)为电子管计算机,其逻辑元件采用电子管,存储器件为声延迟线或磁鼓,典型逻辑结构为定点运算。计算机“软件”一词尚未出现,编制程序所用工具为低级语言。电子管计算机体积大,速度慢(每秒千次或万次),存储器容量小。

1944年,冯·诺依曼在参加原子弹的研制工作中遇到了极为困难的计算问题。1944年夏的一天,他巧遇美国弹道实验室的军方负责人戈尔斯坦,他正参与ENIAC的研制工作,因此,冯·诺依曼被戈尔斯坦介绍加入ENIAC研制组。在设计和研制ENIAC过程中,研制小组意识到ENIAC还存在很多问题,例如,没有存储器,也没有采用二进制。1945年,在共同讨论的基础上,冯·诺依曼以《关于EDVAC的报告草案》为题,起草了长达101页的总结报告,发表了全新的“存储程序(stored-program)通用电子计算机方案”,宣告了现代计算机结构思想的诞生。

“存储程序”方式的基本思想是:必须将事先编好的程序和原始数据送入主存后才能执行程序,一旦程序被启动执行,计算机能在不需操作人员干预下自动完成逐条取出指令并执行的任务。

冯·诺依曼及其同事在普林斯顿高级研究院(Institute for Advance Study, IAS)于1946年开始设计“存储程序”计算机,该机被称为IAS计算机,它是后来通用计算机的原型。

IAS计算机基本结构如图1.1所示,包含5个部件:算术逻辑单元(即运算器)、程序控制器、主存储器、输入设备和输出设备。整个计算机以主存储器为中心,程序和数据首先输入到运算器,然后保存到主存。程序运行时,从主存调出指令,在程序控制器和运算器中执行,执行的中间结果送主存保存,最终结果送输出设备。

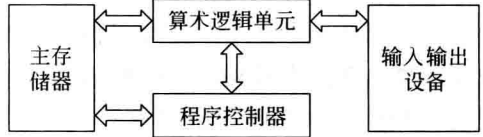


图 1.1 IAS 计算机基本结构

英国剑桥大学的 M. V. Wilkes 在 EDVAC 方案启发下,于 1949 年制造成功的 EDSAC 成为世界上第一台“存储程序”式的现代计算机,而 IAS 计算机直到 1951 年才告完成。此外,尚有 1951 年的 UMVAC-1 和 1956 年的 IBM 704 等也都属于第一代计算机。

* 1.1.3 第二代计算机

第二代计算机(20世纪50年代中后期到20世纪60年代中)为晶体管计算机。1947年,美国贝尔实验室的三位科学家 William Shockley、John Bardeen、Walter Brattain 发明了晶体管,为计算机的发展提供了新的技术基础。该实验室于 1954 年研制了晶体管计算机 TRADIC,而麻省理工学院于 1957 年完成的 TX-2 对晶体管计算机的发展起了重要作用。IBM 公司于 1955 年宣布的全晶体管计算机 7070 和 7090 开启了第二代计算机蓬勃发展的新时期,特别是 1959 年 IBM 公司推出的商用机 IBM 1401,更以其小巧价廉和面向数据处理的特性而获得广大用户的欢迎,从而促进了计算机工业的迅速发展。

这一代计算机除了逻辑元件采用晶体管以外,其内存采用磁芯存储器,外存采用磁鼓与