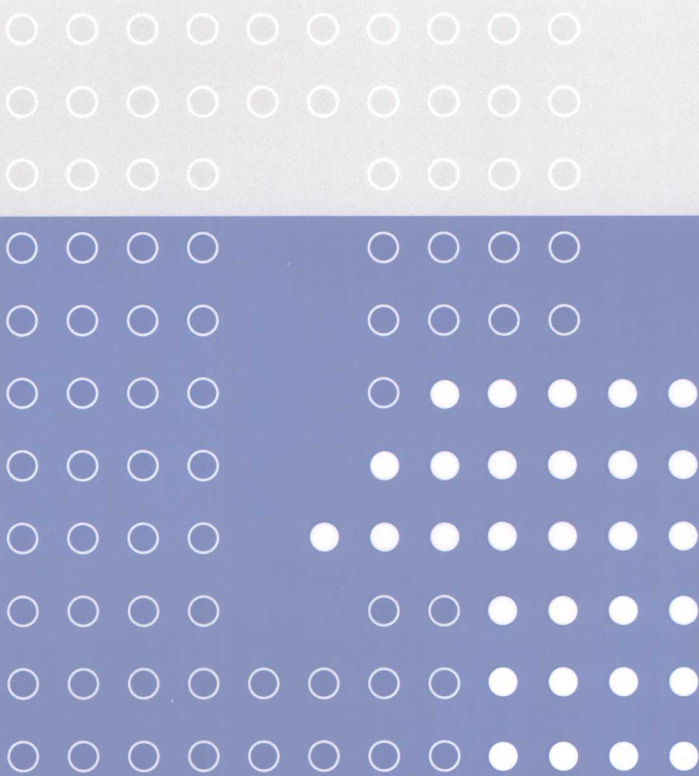




普通高等教育“十一五”国家级规划教材 计算机系列教材

计算机组织与体系结构

(第4版·立体化教材)



白中英 主编
戴志涛 杨春武 张天乐 于艳丽 编著
陈鸿安 主审



内附光盘



清华大学出版社



普通高等教育“十一五”国家级规划教材 计算机系列教材

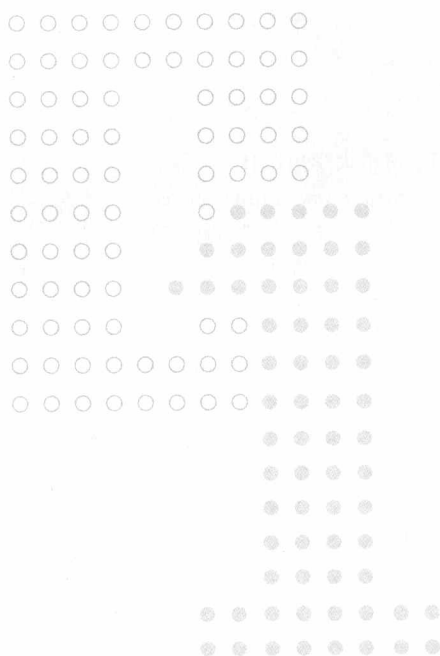
白中英 主编

戴志涛 杨春武 张天乐 于艳丽 编著

陈鸿安 主审

计算机组织与体系结构

(第4版·立体化教材)



清华大学出版社
北京

内 容 简 介

本书重点讲授计算机单处理机系统的组成和工作原理。在此基础上,扩展讲授并行计算机的体系结构。书中内容分为13章:(1)计算机系统概论;(2)运算方法和运算器;(3)内部存储器;(4)指令系统;(5)中央处理机;(6)总线系统;(7)外围设备;(8)输入/输出系统;(9)操作系统支持;(10)安腾高性能处理机体系结构;(11)并行体系结构;(12)教学实验设计;(13)课程综合设计。在附录A中介绍了配套光盘与教学设备。

本书是作者对“计算机组织与体系结构”课程体系、教学内容、教学方法、教学手段进行综合改革的具体成果。

本书特色:基础性、时代性、系统性、实践性、实用性、启发性融为一体,文字教材、多媒体CAI软件、教学课件、习题答案库、自测试题库、教学仪器、实验设计、课程设计综合配套,形成“理论、实验、设计”三个过程相统一的立体化教学体系。本书文字流畅、通俗易懂,有广泛的适应面,可作为高等学校计算机系的教材,也可作为成人自学考试和全国计算机等级考试NCRE(四级)用书。

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

计算机组织与体系结构 / 白中英主编;戴志涛等编著. —4版. —北京:清华大学出版社, 2008.8

(计算机系列教材)

ISBN 978-7-302-18332-7

I. 计… II. ①白… ②戴… 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(2008)第151820号

责任编辑:马瑛珺 焦 虹

责任校对:梁 毅

责任印制:何 芊

出版发行:清华大学出版社

地 址:北京清华大学学研大厦A座

<http://www.tup.com.cn>

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

印 刷 者:清华大学印刷厂

装 订 者:三河市新茂装订有限公司

经 销:全国新华书店

开 本:185×260 印 张:26.75 彩 插:1 字 数:633千字

附光盘1张

版 次:2008年8月第4版

印 次:2008年8月第1次印刷

印 数:1~5000

定 价:43.00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。
联系电话:010-62770177 转 3103 产品编号:028790-01

主 任：周立柱

副 主 任：王志英 李晓明

编委委员：（按姓氏笔画为序）

汤志忠 孙吉贵 杨 波

岳丽华 钱德沛 谢长生

蒋宗礼 廖明宏 樊晓桢

责任编辑：马瑛珺

E D I T O R S

责任编委：王志英

“计算机组织与体系结构”是计算机科学与技术系的一门硬件核心专业基础课程。从课程地位来说,它在先导课和后续课之间起着承上启下的作用。

“计算机组织与体系结构”重点讲授单处理机系统的组织和工作原理,在此基础上扩充讲授并行计算机系统的体系结构。课程教学具有知识面广、内容多、难度大、更新快等特点。另一方面,体现课程特点的教材对于提高教学水平和培养人才的质量起着十分重要的作用。作者认为,一本好的“计算机组织与体系结构”教材主要应具备下述特点:

- (1) 基础性强,知识结构合理,为学生建立终生知识体系打下良好基础。
- (2) 系统性强,知识模块彼此交互,使学生能清晰地建立计算机整机概念。
- (3) 时代性强,及时反映前沿方向,以适应计算机科学技术快速发展的需要。
- (4) 实践性强,理论教学与实践教学结合,注重学生的智力开发和能力培养。
- (5) 应用性强,有较广的适应面,以适应学生在各类计算机上从事开发应用的需要。
- (6) 启发性强,结合计算机科学技术的重大进展,培养学生的创新思维和创新意识。

按 CCC2005 教学计划,本课程的先修课是“数字逻辑与数字系统”。

作者根据多年从事计算机硬件课程理论教学和实践教学经验,从传授知识和培养能力的目标出发,结合本课程教学的特点、难点和要点,在这次新版教材编写中又进行了课程体系、教学内容、教学方法和教学手段的改革,使文字教材、多媒体 CAI 软件、教学课件、习题库、试题库、教学仪器、实验设计、课程设计综合配套,力求形成“理论、实验、设计”三个过程相统一的立体化教学体系。

理论教学计划 64 学时。授课学时建议:第 1 章 3(学时,下同),第 2 章 9,第 3 章 9,第 4 章 4,第 5 章 12,第 6 章 4,第 7 章 4,第 8 章 6,第 9 章 4,第 10 章 4,第 11 章 5。实验教学 16 学时,与理论教学同步进行;另外,可在小学期集中安排 2 周时间的课程综合设计实践。

倪辉、覃健诚、张杰、靳秀国、杨秦、陈楠、冯一兵、祁之力、王军德、张振华、刘俊荣、李娇娇、宗华丽、王晓梅、胡文发、李贞、王坤山、肖炜、崔洪浚、白媛、郝静、吴璇、杨光辉、杨孟柯等参与了第 4 版文字教材、CAI 软件、教学课件、习题库、试题库、教学仪器、实验设计、课程设计等的编写和研制工作,限于版面,未能在封面上一一署名。

Age Group	Percentage of Respondents
18-29	~65%
30-49	~75%
50-69	~80%
70+	~85%

2008 年 8 月于北京

F O R W A R D

IV

《计算机组织与体系结构(第4版·立体化教材)》目录

第1章 计算机系统概论 /1

- 1.1 计算机的分类 /1
- 1.2 计算机的发展简史 /2
 - 1.2.1 计算机的五代变化 /2
 - 1.2.2 半导体存储器的发展 /3
 - 1.2.3 微处理器的发展 /3
 - 1.2.4 计算机的性能指标 /5
- 1.3 计算机的硬件 /5
 - 1.3.1 硬件组成要素 /5
 - 1.3.2 运算器 /7
 - 1.3.3 存储器 /7
 - 1.3.4 控制器 /8
 - 1.3.5 适配器与输入输出设备 /11
- 1.4 计算机的软件 /11
 - 1.4.1 软件的组成与分类 /11
 - 1.4.2 软件的发展演变 /12
- 1.5 计算机系统的层次结构 /13
 - 1.5.1 多级组成的计算机系统 /13
 - 1.5.2 软件与硬件的逻辑等价性 /14
 - 1.5.3 三个常用术语的物理概念 /15
 - 1.5.4 计算机体系结构的分类 /15

本章小结 /16

习题 /16

第2章 运算方法和运算器 /18

- 2.1 数据与文字的表示方法 /18
 - 2.1.1 数据格式 /18
 - 2.1.2 数的机器码表示 /21
 - 2.1.3 字符与字符串的表示方法 /25
 - 2.1.4 汉字的表示方法 /27
 - 2.1.5 校验码 /28
- 2.2 定点加法、减法运算 /29
 - 2.2.1 补码加法 /29
 - 2.2.2 补码减法 /30
 - 2.2.3 溢出概念与检测方法 /31

2.2.4 基本的二进制加法/减法器 /32

2.3 定点乘法运算 /34

- 2.3.1 原码并行乘法 /34
- * 2.3.2 直接补码并行乘法 /39

2.4 定点除法运算 /41

- 2.4.1 原码除法算法原理 /41
- 2.4.2 并行除法器 /43

2.5 定点运算器的组成 /45

- 2.5.1 逻辑运算 /45
- 2.5.2 多功能算术/逻辑运算单元 /47

2.5.3 内部总线 /51

2.5.4 定点运算器的基本结构 /52

2.6 浮点运算方法和浮点运算器 /53

- 2.6.1 浮点加法、减法运算 /53
- 2.6.2 浮点乘法、除法运算 /56
- 2.6.3 浮点运算流水线 /58
- 2.6.4 浮点运算器实例 /61

本章小结 /63

习题 /63

第3章 内部存储器 /66

3.1 存储器概述 /66

- 3.1.1 存储器的分类 /66
- 3.1.2 存储器的分级 /67
- 3.1.3 主存储器的技术指标 /68

3.2 SRAM 存储器 /68

- 3.2.1 基本的静态存储元阵列 /68
- 3.2.2 基本的SRAM逻辑结构 /69
- 3.2.3 读/写周期波形图 /70

3.3 DRAM 存储器 /71

- 3.3.1 DRAM存储元的记忆原理 /71
- 3.3.2 DRAM芯片的逻辑结构 /72
- 3.3.3 读/写周期、刷新周期 /73
- * 3.3.4 存储器容量的扩充 /74

3.3.5	高级的 DRAM 结构	/76
3.3.6	DRAM 主存读/写的正确性校验	/80
3.4	只读存储器和闪速存储器	/81
3.4.1	只读存储器 ROM	/81
3.4.2	FLASH 存储器	/84
3.5	并行存储器	/87
3.5.1	双端口存储器	/87
3.5.2	多模块交叉存储器	/90
3.6	cache 存储器	/93
3.6.1	cache 基本原理	/93
3.6.2	主存与 cache 的地址映射	/95
3.6.3	替换策略	/99
3.6.4	cache 的写操作策略	/100
3.6.5	Pentium 4 的 cache 组织	/100

本章小结 /101

习题 /102

第4章 指令系统 /104

4.1	指令系统的发展与性能要求	/104
4.1.1	指令系统的发展	/104
4.1.2	对指令系统性能的要求	/105
4.1.3	低级语言与硬件结构的关系	/105
4.2	指令格式	/106
4.2.1	操作码	/107
4.2.2	地址码	/107
4.2.3	指令字长度	/108
4.2.4	指令助记符	/109
4.2.5	指令格式举例	/109
4.3	操作数类型	/112
4.3.1	一般的数据类型	/112
4.3.2	Pentium 数据类型	/112
4.3.3	Power PC 数据类型	/113
4.4	指令和数据的寻址方式	/113
4.4.1	指令的寻址方式	/113

4.4.2	操作数基本寻址方式	/114
4.4.3	寻址方式举例	/118
4.5	典型指令	/122
4.5.1	指令的分类	/122
4.5.2	基本指令系统的操作	/123
4.5.3	精简指令系统	/125

本章小结 /126

习题 /127

第5章 中央处理机 /129

5.1	CPU 的功能和组成	/129
5.1.1	CPU 的功能	/129
5.1.2	CPU 的基本组成	/129
5.1.3	CPU 中的主要寄存器	/131
5.1.4	操作控制器与时序产生器	/132
5.2	指令周期	/132
5.2.1	指令周期的基本概念	/132
5.2.2	MOV 指令的指令周期	/134
5.2.3	LAD 指令的指令周期	/136
5.2.4	ADD 指令的指令周期	/137
5.2.5	STO 指令的指令周期	/138
5.2.6	JMP 指令的指令周期	/140
5.2.7	用方框图语言表示指令周期	/141
5.3	时序产生器和控制方式	/143
5.3.1	时序信号的作用和体制	/143
5.3.2	时序信号产生器	/144
5.3.3	控制方式	/146
5.4	微程序控制器	/147
5.4.1	微程序控制原理	/147
5.4.2	微程序设计技术	/153
5.5	硬连线控制器	/157
5.6	传统 CPU	/159
5.6.1	Intel 8088 CPU	/159
5.6.2	IBM 370 系列 CPU	/161
5.7	流水 CPU	/162

5.7.1 流水 CPU 的结构 /162	第 7 章 外围设备 /207
5.7.2 流水线中的主要问题 /164	7.1 外围设备概述 /207
5.7.3 奔腾 CPU /166	7.1.1 外围设备的一般功能 /207
5.8 RISC CPU /171	7.1.2 外围设备的分类 /208
5.8.1 RISC CPU 的特点 /171	7.2 磁盘存储设备 /209
5.8.2 RISC CPU 实例 /172	7.2.1 磁记录原理 /209
5.9 多媒体 CPU /176	7.2.2 磁盘的组成和分类 /211
5.9.1 多媒体技术的主要问题 /176	7.2.3 磁盘驱动器和控制器 /212
5.9.2 MMX 技术 /177	7.2.4 磁盘上信息的分布 /214
本章小结 /179	7.2.5 磁盘存储器的技术指标 /215
习题 /180	7.3 磁盘存储设备的技术发展 /216
第 6 章 总线系统 /183	7.3.1 磁盘 cache /216
6.1 总线的概念和结构形态 /183	7.3.2 磁盘阵列 RAID /217
6.1.1 总线的基本概念 /183	7.3.3 可移动存储设备 /218
6.1.2 总线的连接方式 /184	* 7.4 磁带存储设备 /219
6.1.3 总线的内部结构 /186	7.5 光盘和磁光盘存储设备 /220
6.2 总线接口 /187	7.5.1 光盘存储设备 /220
6.2.1 信息传送方式 /187	7.5.2 磁光盘存储设备 /222
6.2.2 总线接口的基本概念 /189	* 7.6 显示设备 /223
6.3 总线的仲裁 /190	7.6.1 显示设备的分类与有关概念 /223
6.3.1 集中式仲裁 /191	7.6.2 字符/图形显示器 /224
6.3.2 分布式仲裁 /192	7.6.3 图像显示设备 /226
6.4 总线的定时和数据传送模式 /193	7.6.4 VESA 显示标准 /227
6.4.1 总线的定时 /193	* 7.7 输入设备和打印设备 /230
6.4.2 总线数据传送模式 /195	7.7.1 输入设备 /230
6.5 HOST 总线和 PCI 总线 /196	7.7.2 打印设备 /231
6.5.1 多总线结构 /196	本章小结 /232
6.5.2 PCI 总线信号 /197	习题 /233
6.5.3 总线周期类型 /198	第 8 章 输入/输出系统 /235
6.5.4 总线周期操作 /200	8.1 外围设备的速度分级与信息交换方式 /235
6.5.5 总线仲裁 /201	8.1.1 外围设备的速度分级 /235
6.6 InfiniBand 标准 /201	8.1.2 信息交换方式 /236
6.6.1 InfiniBand 的体系结构 /201	8.2 程序查询方式 /238
6.6.2 InfiniBand 的通信协议 /203	8.3 程序中断方式 /241
本章小结 /204	8.3.1 中断的基本概念 /241
习题 /205	

8.3.2	程序中中断方式的基本 I/O 接口	存储器	/285
	/243	9.4.4	虚存的替换算法 /288
8.3.3	单级中断和多级中断	9.5	存储保护 /289
8.3.4	中断控制器	9.5.1	存储区域保护 /289
8.3.5	Pentium 中断机制	9.5.2	访问方式保护 /291
8.4	DMA 方式	9.6	奔腾系列机的虚存组织 /291
8.4.1	DMA 的基本概念	9.6.1	存储器模型 /292
8.4.2	DMA 传送方式	9.6.2	虚地址模式 /292
8.4.3	基本的 DMA 控制器	9.6.3	分页模式下的地址转换 /293
8.4.4	选择型和多路型 DMA 控制器	本章小结	/294
	/257	习题	/295
8.5	通道方式	第 10 章	安腾高性能处理机体系结构 /297
8.5.1	通道的功能	10.1	高性能处理机体系结构的演变 /297
8.5.2	通道的类型	10.1.1	IA 体系结构的历史演变 /297
8.5.3	通道结构的发展	10.1.2	英特尔 64 位处理机的两种体系结构 /298
8.6	通用 I/O 标准接口	10.2	安腾体系结构的基本设计思想 /299
8.6.1	并行 I/O 标准接口 SCSI	10.3	安腾指令系统结构 /302
8.6.2	串行 I/O 标准接口 IEEE 1394	10.3.1	执行单元与指令类型 /302
	/264	10.3.2	安腾寄存器结构 /303
本章小结	/266	10.3.3	安腾指令格式 /304
习题	/267	10.3.4	安腾汇编语言格式 /305
第 9 章	操作系统支持 /270	10.4	指令级并行机制 /306
9.1	操作系统概述 /270	10.4.1	推断执行技术 /307
9.1.1	操作系统的概念	10.4.2	推测技术 /309
9.1.2	操作系统的功能	10.5	双核安腾处理机的组成 /312
9.1.3	操作系统的特性及其需要解决的问题	10.5.1	双核安腾处理机的基本特性 /312
9.1.4	操作系统的硬件环境	10.5.2	双核安腾处理机的组织结构 /314
9.2	调度 /276	本章小结	/317
9.2.1	进程	习题	/318
9.2.2	调度的层次	第 11 章	并行体系结构 /320
9.2.3	处理机调度的实现	11.1	体系结构中的并行性 /320
9.3	存储管理 /280	11.1.1	并行性的概念 /320
9.3.1	分区式存储管理	11.1.2	提高并行性的技术途径 /321
9.3.2	交换技术和分布技术	11.1.3	单处理机系统中并行性的发展
9.4	虚拟存储器 /281		
9.4.1	虚拟存储器的基本概念		
9.4.2	页式虚拟存储器		
9.4.3	段式虚拟存储器和段页式虚拟存		

/321	第12章 教学实验设计 /357
11.1.4 多处理机系统中并行性的发展 /322	12.1 教学实验仪器与测试工具 /357
11.1.5 并行处理机的体系结构类型 /323	12.1.1 TEC-8 计算机组成与体系结构实验系统 /357
11.2 超长指令字处理机 /324	12.1.2 逻辑测试笔 /358
11.2.1 VLIW 处理机的特点 /324	12.2 TEC-8 实验系统结构和操作 /359
11.2.2 VLIW 处理机的结构模型 /325	12.2.1 模型计算机时序信号 /359
11.2.3 典型处理机结构 /326	12.2.2 模型计算机组成 /359
11.3 多线程与超线程处理机	12.2.3 模型计算机指令系统 /362
11.3.1 指令级并行与线程级并行 /328	12.2.4 开关、按钮、指示灯 /363
11.3.2 同时多线程结构 /330	12.2.5 E ² PROM 中微代码的修改 /365
11.3.3 超线程处理机结构 /331	12.3 运算器组成实验 /371
11.4 向量处理机 /332	12.4 双端口存储器实验 /376
11.4.1 向量处理的基本概念 /332	12.5 数据通路实验 /380
11.4.2 向量处理机的结构 /335	12.6 微程序控制器实验 /385
11.4.3 并行向量处理机 /338	12.7 CPU 组成与机器指令执行实验 /391
11.5 对称多处理机 /340	12.8 中断原理实验 /395
11.5.1 多处理机系统的分类 /340	第13章 课程综合设计 /399
11.5.2 SMP 的基本概念 /341	13.1 模型机硬连线控制器设计 /399
11.5.3 SMP 的结构和实例 /342	13.2 模型机流水微程序控制器设计 /404
11.5.4 多处理机操作系统 /344	13.3 模型机流水硬连线控制器设计 /407
11.5.5 多处理机的 cache 一致性 /345	13.4 含有阵列乘法器的 ALU 设计 /409
11.6 机群系统 /347	附录 A 配套光盘与教学设备 /412
11.6.1 机群系统的定义和特点 /347	参考文献 /413
11.6.2 机群系统的体系结构 /348	
11.6.3 IBM SP2 系统 /350	
11.6.4 超级刀片系统 /351	
本章小结 /354	
习题 /355	

第 1 章 计算机系统概论

计算机系统不同于一般的电子设备,它是一个由硬件、软件组成的复杂的自动化设备。本章先说明计算机的分类,然后采用自上而下的方法,简要地介绍硬件、软件的概念和组成,目的在于使读者先有一个粗略的总体概念,以便于展开后续各章内容。

1.1 计算机的分类

计算机从总体上来说分为两大类。一类是模拟计算机。“模拟”就是相似的意思。例如计算尺是用长度来标示数值;时钟是用指针在表盘上转动来表示时间;电表是用角度来反映电量大小,这些都是模拟计算装置。模拟计算机的特点是数值由连续量来表示,运算过程也是连续的。

另一类是数字计算机,它是在算盘的基础上发展起来的,是用数字来表示数量的大小。数字计算机的主要特点是按位运算,并且不连续地跳动计算。表 1.1 列出了数字计算机与模拟计算机的主要区别。

表 1.1 数字计算机与模拟计算机的主要区别

比较内容	数字计算机	模拟计算机
数据表示方式	数字 0 和 1	电压
计算方式	数字计数	电压组合和测量值
控制方式	程序控制	盘上连线
精度	高	低
数据存储量	大	小
逻辑判断能力	强	无

模拟计算机由于精度和解题能力都有限,所以应用范围较小。数字计算机则与模拟计算机不同,它是以近似于人类的“思维过程”来进行工作的,所以有人把它叫做电脑。它的发明和发展是 20 世纪人类最伟大的科学技术成就之一,也是现代科学技术发展水平的主要标志。习惯上所称的计算机,一般是指现在广泛应用的数字计算机。

数字计算机进一步又可分为专用计算机和通用计算机。专用和通用是根据计算机的效率、速度、价格、运行的经济性和适应性来划分的。专用机是最有效、最经济和最快速的计算机,但是它的适应性最差。通用计算机适应性很强,但是牺牲了效率、速度和经济性。

通用计算机可分超级计算机、大型机、服务器、工作站、微型机和单片机六类,它们的区别在于体积、简易性、功率损耗、性能指标、数据存储容量、指令系统规模和机器价格,如图 1.1 所示。一般来说,超级计算机主要用于科学计算,其运算速度在每秒万亿次以上,数据存储容量很大,结构复杂,价格昂贵。而单片计算机是只用一片集成电路做成的计算机,体积小,结构简单,性能指标较低,价格便宜。介于超级计算机和单片机之间的是大型机、服务器、工作站和微型机,它们的结构规模和性能指标依次递减。但是随着超大规模集成电路的迅速发展,微型机、工作站等彼此之间的概念也在发生变化,因为今天的工作站可能就是明天的微型机,而今天的微型机可能就是明天的单片机。专用计算机是针对某一任务设计的计算机,一般来说,其结构要比通用机简单。目前,已经有多种型号的通用单片机及嵌入式单片机用于测试或控制。

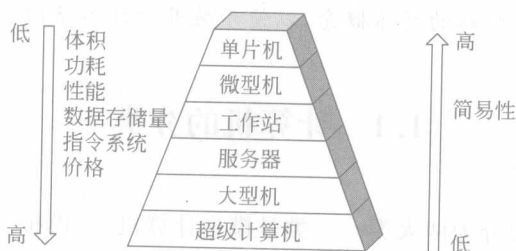


图 1.1 单片机、微型机、工作站、服务器、大型机、超级计算机之间的区别

1.2 计算机的发展简史

1.2.1 计算机的五代变化

世界上第一台电子数字计算机是 1946 年在美国宾夕法尼亚大学制成的。这台机器用了 18000 多个电子管,占用长度超过 30m 的房间,重量达 30t,而运算速度只有 5000 次每秒。从今天的眼光来看,这台计算机耗费既大又不完善,但却是科学史上一次划时代的创新,它奠定了电子计算机的基础。自从这台计算机问世 60 多年来,从使用的器件角度来说,计算机的发展大致经历了五代的变化。

第一代电子管计算机,1946—1957 年。计算机的运算速度为几千次至几万次每秒,体积庞大,成本很高,可靠性较低。在此期间,形成了计算机的基本体系,确定了程序设计的基本方法,数据处理机开始得到应用。

第二代晶体管计算机,1958—1964 年。计算机的运算速度提高到几万次至十几万次每秒,可靠性提高,体积缩小,成本降低。在此期间,工业控制机开始得到应用。

第三代中小规模集成电路计算机,1965—1971 年。计算机的可靠性进一步提高,体积进一步缩小,成本进一步下降,运算速度提高到几十万次至几百万次每秒。在此期间形成了机种多样化,生产系列化,使用系统化的特点,小型计算机开始出现。

第四代大规模和超大规模集成电路计算机,1972—1990 年。其可靠性更进一步提

高,体积更进一步缩小,成本更进一步降低,速度提高到 1000 万次至 1 亿次每秒。在此期间,由几片大规模集成电路组成的**微型计算机**开始出现。

第五代为 1991 年开始的巨大规模集成电路计算机,运算速度提高到 10 亿次每秒。由一片巨大规模集成电路实现的**单片计算机**开始出现。

总之,从 1946 年计算机诞生以来,大约每隔五年运算速度提高 10 倍,可靠性提高 10 倍,成本降低 10 倍,体积缩小 10 倍。而 20 世纪 70 年代以来,计算机的生产数量每年以 25% 的速度递增。

值得一提的是,计算机从第三代起,与集成电路技术的发展密切相关。由于 LSI 的采用,使得一块集成电路芯片上可以放置 1000 个元件;VLSI 使得每个芯片可放置 1 万个元件;现在的 ULSI 芯片更是超过了 100 万个元件。1965 年摩尔观察到芯片上的晶体管数量每年翻一番,1970 年这种态势减慢成每 18 个月翻一番,这就是人们所称的**摩尔定律**。直到目前,这个增长速率仍在持续下去。

1.2.2 半导体存储器的发展

20 世纪 50~60 年代,所有计算机存储器都是由微小的铁磁体环(磁芯)做成,每个磁芯直径约 1mm。这些小磁芯处在计算机内用三条细导线穿过的网格板上。每个磁芯的一种磁化方向代表一个 1,另一个磁化方向则代表一个 0。磁芯存储器速度相当快,读存储器中的一位只需 $1\mu\text{s}$ 。但是磁芯存储器价格昂贵,体积大,而且读出是破坏性的,因此必须有读出后立即重写数据的电路。不仅如此,其工艺也比较复杂,甚至需手工制作。

1970 年,仙童半导体公司生产出了第一个较大容量半导体存储器。一个相当于单个磁芯大小的芯片,包含了 256 位的存储器。这种芯片是非破坏性的,而且读写速度比磁芯快得多,读出一位只要 70ns,但是其价格比磁芯要贵。

1974 年,每位半导体存储器的价格已低于磁芯。这以后,存储器的价格持续快速下跌,但存储密度却不断增加,从而使得新的机器更小、更快,存储容量更大,价格更便宜。存储器技术的发展与处理器技术的发展一起,在不到 10 年间改变了计算机的生命力。虽然庞大昂贵的计算机仍然存在,但计算机已经走向了个人电脑时代。

从 1970 年起,半导体存储器经历了 11 代:单个芯片 1KB、4KB、16KB、64KB、256KB、1MB、4MB、16MB、64MB、256MB 和现在的 1GB。其中 $1\text{K}=2^{10}$, $1\text{M}=2^{20}$, $1\text{G}=2^{30}$ 。每一代比前一代的存储密度提高了 4 倍,而价格和存取时间都在下降。

1.2.3 微处理器的发展

同存储器芯片一样,处理器芯片的单元密度也在不断增加。随着时间的推移,每块芯片上的单元个数越来越多,因此构建一个计算机处理器所需的芯片越来越少。表 1.2 列出了 Intel 公司微处理器的演化。

1971 年 Intel 公司开发出 Intel 4004。这是第一个将 CPU 的所有元件都放入同一块芯片内的产品,于是,微处理器诞生了。

表 1.2 Intel 公司微处理器的演化

(a) 20 世纪 70 年代的处理器					
型 号	4004	8008	8080	8086	8088
发布时间	1971 年	1972 年	1974 年	1978 年	1979 年
时钟速率	108kHz	108kHz	2MHz	5MHz,8MHz,10MHz	5MHz,8MHz
总线宽度	4 位	8 位	8 位	16 位	8 位
晶体管数	2300	3500	6000	29 000	29 000
特征尺寸(μm)	10		6	3	3
可寻址存储器	640B	16KB	64KB	1MB	1MB
虚拟存储器	—	—	—	—	—
(b) 20 世纪 80 年代的处理器					
型 号	80286	386TM DX	386TM SX	486TM DX	
发布时间	1982 年	1985 年	1988 年	1989 年	
时钟速率	6~12.5MHz	16~33MHz	16~33MHz	25~50MHz	
总线宽度	16 位	32 位	16 位	32 位	
晶体管数	134 000	275 000	275 000	1 200 000	
特征尺寸(μm)	1.5	1	1	0.8~1	
可寻址存储器	16MB	4GB	16MB	4GB	
虚拟存储器	1GB	64TB	64TB	64TB	
(c) 20 世纪 90 年代的处理器					
型 号	486TM SX	Pentium	Pentium Pro	Pentium II	
发布时间	1991 年	1993 年	1995 年	1997 年	
时钟速率	16~33MHz	60~166MHz	150~220MHz	200~300MHz	
总线宽度	32 位	32 位	64 位	64 位	
晶体管数	1.185 百万	3.1 百万	5.5 百万	7.5 百万	
特征尺寸(μm)	1	0.8	0.6	0.35	
可寻址存储器	4MB	4GB	64GB	64GB	
虚拟存储器	64TB	64TB	64TB	64TB	
(d) 最近的处理器					
型 号	Pentium III	Pentium 4	Itanium	Itanium 2	
发布时间	1999 年	2000 年	2001 年	2002 年	
时钟速率	450~600MHz	1.3~1.8GHz	733~800MHz	0.9~1GHz	
总线宽度	64 位	64 位	64 位	64 位	
晶体管数	9.6 百万	42 百万	25 百万	220 百万	
特征尺寸(μm)	0.25	0.18	0.18	0.18	
可寻址存储器	64GB	64GB	64GB	64GB	
虚拟存储器	64TB	64TB	64TB	64TB	

4004 能完成两个 4 位数相加,通过重复相加能完成乘法。按今天的标准,4004 虽然过于简单,但它却是微处理器的功能不断发展的奠基者。

微处理器演变中的另一个主要进步是 1972 年出现的 Intel 8008。这是第一个 8 位微处理器,它比 4004 复杂一倍。

1974 年出现了 Intel 8080。这是第一个通用微处理器,而 4004 和 8008 是为特殊用途而设计的。8080 是为通用微机而设计的中央处理器。它与 8008 一样,都是 8 位微处理器,但 8080 更快,有更丰富的指令集和更强的寻址能力。

大约在同时,16 位微机被开发出来。但是直到 20 世纪 70 年代末才出现强大的通用 16 位微处理器,8086 便是其中之一。这一发展趋势中的另一阶段是 1981 年,贝尔实验室和 HP 公司开发出了 32 位单片微处理器。Intel 于 1985 年推出了 32 位微处理器 Intel 80386。

1.2.4 计算机的性能指标

吞吐量: 表征一台计算机在某一时间间隔内能够处理的信息量。

响应时间: 表征从输入有效到系统产生响应之间的时间度量,用时间单位来度量。

利用率: 在给定的时间间隔内,系统被实际使用的时间所占的比率,用百分比表示。

处理机字长: 指处理机运算器中一次能够完成二进制数运算的位数,如 32 位、64 位。

总线宽度: 一般指 CPU 中运算器与存储器之间进行互连的内部总线二进制位数。

存储器容量: 存储器中所有存储单元的总数目,通常用 KB、MB、GB、TB 来表示。

存储器带宽: 单位时间内从存储器读出的二进制数信息量,一般用 B/s(字节/秒)表示。

主频/时钟周期: CPU 的工作节拍受主时钟控制,主时钟不断产生固定频率的时钟,主时钟的频率(f)叫 **CPU 的主频**。度量单位是 MHz(兆赫兹)、GHz(吉赫兹)。

主频的倒数称为 **CPU 时钟周期(T)**,即 $T=1/f$,度量单位是 μs (微秒)、 ns (纳秒)。

CPU 执行时间: 表示 CPU 执行一段程序所占用的 CPU 时间,可用下式计算:

$$\text{CPU 执行时间} = \text{CPU 时钟周期数} \times \text{CPU 时钟周期长}$$

CPI: 表示每条指令周期数,即执行一条指令所需的平均时钟周期数。用下式计算:

$$\text{CPI} = \text{执行某段程序所需的 CPU 时钟周期数} \div \text{该程序包含的指令条数}$$

MIPS: 每秒百万条指令数,即单位时间内执行的指令数。用下式计算:

$$\text{MIPS} = \text{指令条数} \div (\text{程序执行时间} \times 10^6)$$

MFLOPS: 每秒百万次浮点操作次数,用来衡量机器浮点操作的性能。用下式计算:

$$\text{MFLOPS} = \text{程序中的浮点操作次数} \div (\text{程序执行时间} \times 10^6)$$

1.3 计算机的硬件

1.3.1 硬件组成要素

要了解数字计算机的主要组成和工作原理,可从打算盘说起。假设给一个算盘、一张带有横格的纸和一支笔,要求我们计算 $y=ax+b-c$ 这样一个题目。为了和下面讲到的内容

进行比较,不妨按以下方法把使用算盘进行解题的步骤事先用笔详细地记录在带横格的纸上。

首先,将横格纸编上序号,每一行占一个序号,如1,2,3,⋯,n,如表1.3所示。其次,把计算式中给定的4个数 a, b, c 和 x 分别写到横格纸的第9,10,11,12行上,每一行只写一个数。接着详细列出给定题目的解题步骤,而解题步骤也需要记在横格纸上,每一步也只写一行。第一步写到横格纸的第一行,第二步写到第二行……依次类推。

根据表1.3所列的解题步骤,从第1行开始,一步一步进行计算,最后可得出所要求的结果。

表 1.3 解题步骤和数据记录在横格纸上

行 数	解题步骤和数据	说 明
1	取数 (9)→算盘	(9)表示第9行的数 a ,下同
2	乘法 (12)→算盘	完成 $a \cdot x$,结果在算盘上
3	加法 (10)→算盘	完成 $ax+b$,结果在算盘上
4	减法 (11)→算盘	完成 $y=ax+b-c$,结果在算盘上
5	存数 $y \rightarrow 13$	算盘上的 y 值记到第13行
6	输出	把算盘上的 y 值写出给人看
7	停止	运算完毕,暂停
8		
9	a	数据
10	b	数据
11	c	数据
12	x	数据
13	y	数据

在完成 $y=ax+b-c$ 的计算过程中,我们用到了什么东西呢?

首先,用到了带横格且编有序号的纸,我们把原始的数据以及解题步骤记录在纸上,即纸“存储”了算题的原始信息。其次,用到了算盘,它用来对数据进行加、减、乘、除等算术运算。再次,用到了笔,利用笔把原始数据和解题步骤记录到纸上,还可把计算结果写出来告诉别人。最后,用到了我们人本身,这主要是人的脑和手。在人的控制下,按照解题步骤一步一步进行操作,直到完成全部运算。

数字计算机进行解题的过程完全和我们人用算盘解题的情况相似,也必须有运算工具,解题步骤和原始数据的输入与存储,运算结果的输出以及整个计算过程的调度控制。和打算盘不同的是,以上这些部分都是由电子线路和其他设备自动进行的。在计算机里,相当于算盘功能的部件,称之为运算器;相当于纸那样具有“记忆”功能的部件,称之为存