



普通高等教育“十一五”国家级规划教材

(修订版)

计算机组成与设计

JISUANJI ZUCHENG YU SHEJI

◆ 潘雪增 平玲娣 编著
王志英 主审



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

普通高等教育“十一五”国家级规划教材

计算机组成与设计

[修订版]

潘雪增 编著
平玲娣
王志英 主审

浙江大学出版社

图书在版编目 (CIP) 数据

计算机组成与设计/潘雪增,平玲娣编著. —杭州: 浙江大学出版社, 2003. 12 (2007 重印)
ISBN 978-7-308-03523-1

I. 计… II. ①潘…②平… III. ①计算机体系结构②电子计算机—设计 IV. TP30

中国版本图书馆 CIP 数据核字 (2003) 第 105870 号

内容简介

本书论述当代计算机的基本组成原理和设计 CPU 的方法, 内容包括组成计算机的基本逻辑部件与定时方法, 指令系统设计的基本原则, 与对现代编译器的支持接口, CPU 的构成和控制器的设计技术方法, 存储体系的构成原理和设计, 输入输出系统设计和磁盘阵列可靠性技术, 以及片上系统设计案例分析等。书中的举例紧密结合当今计算机先进技术, 取材先进、重点突出、叙述深入浅出。

本书可作为理工科大学学生学习《计算机组成与设计》课程, 或《计算机组成原理》课程的教科书, 也可提供从事计算机设计或计算机系统设计的工程技术人员参考。

责任编辑 杜希武

封面设计 刘依群

出版发行 浙江大学出版社

(杭州天目山路 148 号邮政编码 310028)

(E-mail: zupress@mail.hz.zj.cn)

(网址: <http://www.zjupress.com>)

排 版 杭州好友排版工作室

印 刷 浙江大学同力教育彩印有限公司

开 本 16

印 张 28

字 数 681 千字

版 印 次 2007 年 5 月第 2 版 2007 年 5 月第 3 次印刷

印 数 5001—8000

书 号 ISBN 978-7-308-03523-1

定 价 45.00 元

版权所有 翻印必究 印装差错 负责调换

前 言

本书是为计算机专业的学生，和从事计算机事业的工程技术设计人员编写的，本书的读者还包括有点汇编语言或数字逻辑设计经验而需要进一步学习掌握现代计算机的基本组成结构和设计实现的科技工作者，或者有志于我国计算机事业进步的自学者，因此，也适合非计算机专业的学生使用。本书着重基本原理讲述，强调示例启发，力图贯彻少而精。本书使读者通过大量先进的范例学习，掌握计算机系统成功设计的关键技术，包括决定计算机能力、性能的主要因素。

现代计算机技术要求每一位计算机专业人员，既要懂硬件又要懂软件，因为硬件与软件在多个层次上交叉融合，相互影响。并且提供了理解计算机系统的基础框架，不管读者的兴趣在于计算机科学技术或电子工程设计，计算机组成和设计的主旋律都是相同的。因此，本书特别强调在设计计算机硬件时与高层软件的接口，同时注重当代计算机原理的描述。传统上汇编语言单独讲述。本书则把汇编语言融入指令系统讲解，并且介绍指令系统时重点放在为设计 CPU 准备必要知识，着重分析指令系统及格式设计对 CPU 性能和成本的影响，即为设计服务，这有别于以前大多数《计算机组成原理》的课本。

全书共分七章，第一章是计算机概论，主要介绍计算机的五大功能构成：输入、输出、存储器、数据通路（运算器）和控制器。控制器和数据通路合在一起，有时称为处理器或 CPU。同时包括计算机的层次抽象结构、软件分类，大规模集成电路的制造技术及过程；计算机的技术年代划分，和计算机发展历史，体系结构演变等。

第二章为缺乏数字电路设计基础的读者而写，也为虽然学过数字逻辑设计而已生疏的读者复习而写，如果已掌握了这方面知识的读者可以跳过这一章，教学过程中，可以视实际情况而决定是否选择讲授。

第三章讲述机器语言——即计算机指令系统，它是计算机硬件与高层编译系统的接口，也是程序员所见到的计算机体系结构。着重介绍设计精简指令系统与格式的基本准则，及其与 CPU 硬件性能和成本的密切关系；当代计算机发明的存储程序概念原理，编译优化技术对计算机性能提高的重要影响；并结合指令功能设计，通过大量例子，讲解了符号汇编程序的设计方法。

第四章讲述计算机算术运算，及各类运算器的设计算法和实现技术。二进数数在计算机中的各种表示，包括表示为指令，有符号整数、无符号整数、浮点数等；机器数与算术概念的数的主要差别；用超大规模集成电路技术，如何设计快算乘法器和快算除法器的方法。

第五章讲述处理器的设计原理及实现技术，包括 CPU 的有限状态机与微程控制设计技术，具有挑战性的 CPU 中中断事件处理的设计方法，逻辑方程、真值表的表示，PLA 的结构化设计等。讲述用 HDL 语言(Verilog)设计单时钟 CPU，和多时钟 CPU 的设计案例。

第六章讲述多级存储体系结构, 包括单级与多级 cache 设计原理, 虚拟存储系统及设计技术, 重点是地址映射变换技术, 快查表设计技术, 现代多级 Cache 设计方法, 存储体系的框架技术。

第七章讲述输入输出系统, 重点讲述与处理器及外部设备的接口技术, 包括 I/O 系统的性能评价、I/O 设备的行为特征、接口的组成、总线协议与新的总线标准、I/O 设备数据传输的控制方式、大型服务器中磁盘可靠性设计技术、以及 I/O 系统的系统设计技术、设计 I/O 系统的定量分析方法。

本书另有中文课件, 有助于教师上课备课, 也便于学生课后复习, 更能帮助自我学习。

本书的第一、二、三、四章由潘雪增编写, 第五、六、七章由平玲娣编写。参加本书第二版部分整理编写的还有孙康、王继民和姜励。全书由潘雪增统编。

参加本书编写的都是浙江大学计算机学院的教师, 他们在不同程度上从事过多年到几十年有关计算机的教学、设计、研制工作。

感谢浙江大学出版社杜希武在本书出版编辑过程中所提的宝贵意见。

希望广大读者、同行们以及使用本教科书的师生们给我们提出宝贵的意见。

潘雪增

2006 年 8 月

目 录

第 1 章 计算机组成及其技术概论	1
1.1 引言.....	1
1.1.1 计算机应用分类及其特点	2
1.1.2 读者能从本书中学到什么	4
1.2 计算机语言与软件系统.....	5
1.3 计算机硬件系统.....	8
1.3.1 鼠标	8
1.3.2 图形显示器	9
1.3.3 主板及其所安装的硬件.....	10
1.3.4 外存储器.....	12
1.3.5 计算机网络.....	14
1.4 集成电路及其发展	15
1.5 计算机发展历史及发展阶段划分	17
1.5.1 第一台电子计算机.....	17
1.5.2 计算机的商业发展.....	19
1.5.3 嵌入式计算机的发展.....	19
1.5.4 计算机发展阶段及其年代划分.....	20
1.6 本章小结	21
1.7 关键术语中英文对照表	22
1.8 习题	23
第 2 章 基本逻辑电路.....	27
2.1 概述	27
2.2 门电路, 真值表, 以及逻辑等式	27
2.2.1 真值表.....	28
2.2.2 逻辑代数.....	28
2.2.3 门电路.....	29

2.3 组合逻辑	30
2.3.1 译码器	30
2.3.2 编码器	31
2.3.3 多路选择器	31
2.3.4 可编程逻辑阵列 PLA (programmable logic array) —两级与或逻辑实现	32
2.3.5 ROM	35
2.3.6 无关项	36
2.3.7 逻辑元件阵列	37
2.4 时钟	38
2.5 存储器与存储元器件	40
2.5.1 触发器和锁存器	40
2.5.2 寄存器文件	42
2.5.3 SRAM	44
2.5.4 DRAM	47
2.5.5 差错的检查和纠正	49
2.6 有限状态机	50
2.6.1 有限状态机	50
2.7 定时方法学	53
2.7.1 电平触发定时	55
2.7.2 异步输入和同步器	55
2.8 小结	57
2.9 关键术语中英文对照表	57
2.10 习题	58
 第3章 指令系统	 61
3.1 概述	61
3.2 计算机硬件的操作指令	61
3.3 计算机硬件的操作数	63
3.3.1 操作数在寄存器中	63
3.3.2 存储器操作数—操作数在存储器中	63
3.3.3 立即数操作数—指令格式中的常数	66
3.4 计算机指令格式	68
3.5 逻辑操作	73

3.6 决策指令	77
3.7 计算机硬件对子程序的支持	83
3.8 计算机对字符的处理	90
3.9 其他类型的 MIPS 寻址方式	93
3.10 程序的运行过程	100
3.10.1 编译.....	101
3.10.2 汇编.....	101
3.10.3 链接目标代码.....	102
3.10.4 把程序装入内存中.....	104
3.10.5 动态链接库.....	105
3.11 编译器优化	107
3.11.1 高层次优化.....	108
3.11.2 局部优化和全局优化.....	108
3.11.3 理解程序性能.....	110
3.11.4 全局代码优化.....	110
3.11.5 优化技术总结.....	110
3.12 综合程序范例.....	111
3.12.1 子程序 swap.....	111
3.12.2 子程序 sort.....	113
3.13 数组和指针的比较	117
3.14 实例分析: IA32 指令集.....	120
3.14.1 IBM/Motorola PowerPC.....	120
3.14.2 Intel IA32.....	122
3.15 指令集发展历史	128
3.15.1 累加器体系结构.....	128
3.15.2 通用寄存器体系结构.....	129
3.15.3 紧凑编码和堆栈体系结构.....	130
3.15.4 高级语言计算机体系结构.....	131
3.15.5 精简指令集计算机体系结构.....	131
3.15.6 80x86 的历史简介.....	131
3.16 本章小结.....	132
3.17 关键术语中英文对照表.....	134
3.18 习题.....	134



第 4 章 计算机算术	143
4.1 概述	143
4.2 有符号和无符号数	143
4.3 加减法运算	149
4.4 构造算术逻辑单元	153
4.4.1 1 位 ALU	154
4.4.2 32 位 ALU	156
4.4.3 先行进位	162
4.5 乘法	166
4.5.1 乘法运算基本原理	166
4.5.2 有符号数乘法	172
4.5.3 快速乘法	176
4.5.4 MIPS 的乘法运算	177
4.6 除法	178
4.6.1 第一代除法算法和硬件实现	179
4.6.2 第二代除法算法和硬件实现	181
4.6.3 第三代除法算法和硬件实现	182
4.6.4 有符号数除法	183
4.6.5 快速除法	184
4.6.6 MIPS 除法指令	184
4.7 浮点数运算	187
4.7.1 浮点数的表示法	187
4.7.2 浮点数加法	191
4.7.3 浮点数乘法	194
4.7.4 MIPS 的浮点数指令	196
4.7.5 精确运算	201
4.7.6 浮点运算小结	202
4.8 本章小结	204
4.9 关键术语中英文对照表	207
4.10 习题	207
第 5 章 处理器：数据通路和控制器	213
5.1 引言	213

5.1.1 指令的执行过程	214
5.1.2 逻辑约定和时钟	216
5.1.3 MIPS 指令子集的实现	216
5.2 建立数据通路	217
5.2.1 所需要的主要功能部件	217
5.2.2 执行 R 类型指令所需的功能部件	218
5.2.3 存取指令所需的功能部件	219
5.2.4 条件转移指令所需的功能部件	221
5.2.5 一个复合功能的数据通路	222
5.3 一个简单的实施方案	224
5.3.1 ALU 控制单元设计	225
5.3.2 设计主控制器	228
5.3.3 指令执行过程分析——数据通路的操作	231
5.3.4 设计单时钟周期控制器	240
5.3.5 用门电路实现主控函数	243
5.3.6 为什么不使用单时钟周期实现	243
5.4 多时钟周期的实现	245
5.4.1 多时钟周期数据通路及控制信号	245
5.4.2 多时钟周期的指令执行过程	252
5.4.3 多时钟周期 CPU 控制器设计	256
5.5 微程序设计	270
5.5.1 微指令的格式定义	271
5.5.2 编写微程序	273
5.5.3 实现微程序	276
5.5.4 使用序列发生器实现后继状态函数	277
5.5.5 微程序的硬件实现	280
5.6 异常	282
5.6.1 异常的概念	282
5.6.2 异常的处理	282
5.6.3 控制器对异常的检测	285
5.7 本章小结	286
5.8 关键字	288
5.9 习题	288



第 6 章 存储器层次结构	293
6.1 引言	293
6.2 Cache 存储器的基本原理	295
6.2.1 直接映射	296
6.2.2 Cache 的访问	297
6.2.3 空间局部性原理的利用——多字块 cache	301
6.2.4 处理 Cache 失配	303
6.2.5 写一致性处理策略	304
6.2.6 Cache 系统实例分析	305
6.2.7 支持 cache 的存储器系统的设计方案	306
6.3 评测与改进 Cache 性能	309
6.3.1 Cache 对处理器性能影响的定量分析	309
6.3.2 Cache/主存映射方式	311
6.3.3 定位 Cache 中的块	315
6.3.4 Cache 替换策略	316
6.3.5 多级 Cache	316
6.4 虚拟存储器	319
6.4.1 地址变换	322
6.4.2 缺页中断	324
6.4.3 写策略	326
6.4.4 快查表	326
6.4.5 实例分析: Intrinsity FastMATH TLB	328
6.4.6 虚拟存储器、TLB、Cache 之间的联系	329
6.4.7 存储保护	332
6.4.8 处理缺页中断和 TLB 失配	333
6.4.9 小结	338
6.5 存储层次的通用结构	338
6.5.1 问题 1: 如何存放块?	339
6.5.2 问题 2: 如何寻找块?	340
6.5.3 问题 3: 发生 Cache 失配时选择哪个块被替换?	341
6.5.4 问题 4: 如何处理写操作?	342
6.5.5 Cs 模型	343
6.6 Pentium P4 与 AMD Opteron 存储体系实例分析	345

6.6.1 P4 和 Opteron 的存储体系	345
6.6.2 减少失配损失的技巧	346
6.7 本章小结	347
6.8 习题	350
第 7 章 处理器和外部设备接口	356
7.1 引言	356
7.2 I/O 设备系统	358
7.2.1 鼠标	358
7.2.2 磁盘存储器和可靠性	359
7.3 网络	367
7.3.1 局域网	367
7.3.2 长距离网络	367
7.3.3 无线局域网	369
7.4 总线	370
7.4.1 总线分类	372
7.4.2 同步总线与异步总线	374
7.4.3 增加总线带宽	376
7.4.4 总线标准	377
7.4.5 奔腾 4 总线与网络案例分析	381
7.5 I/O 设备与系统的接口	382
7.5.1 向 I/O 设备发送命令与接口的主要功能	383
7.5.2 I/O 设备与处理器数据传输控制方式	384
7.6 程序中断输入输出方式	393
7.6.1 中断的作用、产生和响应	393
7.6.2 中断处理	395
7.7 DMA 输入输出方式	398
7.7.1 DMA 三种工作方式	399
7.7.2 DMA 控制器组成	399
7.7.3 DMA 的数据传送过程	400
7.8 I/O 性能测量	401
7.8.1 巨型计算机 I/O 基准测试程序	402
7.8.2 事务处理基准测试程序	402



计算机组成与设计

7.8.3 文件系统 I/O 基准测试程序	402
7.8.4 I/O 性能与处理器性能比较	403
7.9 设计 I/O 系统	403
7.10 一个典型的 I/O 系统.....	405
7.11 本章小结	407
7.12 习题	408
 部分参考答案	 413
 附录	 427

第一章 计算机组成及其技术概论

1.1 引言

计算机的飞速发展所引起的信息革命被人们称作第三次社会大变革，其影响远远超过了以前的农业革命和工业革命。它对科学产生了巨大的影响，现在出现了一种新型的科学研究，就是计算科学家同理论科学家、实验科学家一起合作，共同开发天文学、生物学、化学、物理等学科的新的前沿领域。

计算机革命仍在继续。计算机的成本不断在降低，而它的普及程度也不断在提高。过去很多在经济上不可行的应用，现在也逐一实现了。比如下面的应用，在不久前曾经属于“计算机科学幻想”，如今已是很常见的了：

1. 自动出纳机 (automatic teller machines, ATM): 这是一种置于银行墙上，用来分发和收集现金的计算机。这在 20 世纪 50 年代是无法想象的，因为那时候最便宜的计算机也要值 50 万美元以上，并且有一辆汽车那么大。

2. 汽车计算机 (computers in automobiles): 由计算机对汽车进行控制，这在过去也是无法理解的。直到 20 世纪 80 年代早期，微处理器的性能取得了突破性的进展，价格也大幅度下降，它才成为可能。计算机减少了汽车的污染，通过引擎控制提高了燃料利用率，通过防止危险刹车增加了安全系数，并能在撞车时迅速弹出气袋来保护乘车人员。

3. 膝上型计算机 (laptop computers): 过去的人做梦也想不到，计算机会发展到可以让我们随身携带的膝上型计算机，使我们几乎可以在任何地方使用计算机。

4. 人类基因工程 (human genome project): 现在绘制人类 DNA 序列所需计算机设备的成本高达几百万美元，但往前推十年二十年，成本可能是现在的 10 倍甚至 100 倍，那时是没人会考虑这个项目的。

5. 万维网 (World Wide Web, WWW): 现在网络正在改变我们的生活。利用网络，我们可以实现资源共享，分发新闻、发送鲜花、在线购物、遨游世界、聊天交友等等。

计算机技术的发展已经影响到社会的各个方面。硬件的发展允许程序员编出很多精彩的实用软件，也使得计算机更加普及。如今，人们对计算机应用又提出了新的科学幻想，如电子图书馆、无货币社会、自动信息高速公路，以及计算机的完全普及（真正意义上的无所不在，到处都有计算机，人们不再需要携带计算机）。

让用户尽快得到结果，是软件开发成功的一个关键因素。在 20 世纪 60 和 70 年代，计算机性能的一个主要制约因素是其主存储器的大小。因而那时的程序员常常遵循一个简单的信条：在最小的主存储空间内使程序快速执行。但近年来，计算机设计和存储技术的发



展已使主存的容量得到很大提高, 存储容量的限制大大降低。程序员应当懂得现在的存储器模型和以前的简单模型的区别, 以及存储器的层次性和处理器的并行性。要编出有竞争力的各类软件, 程序员应该增强对计算机组织结构的理解和认识。

1.1.1 计算机应用分类及其特点

尽管从智能家电到手机乃至大型超级计算机都涉及计算机硬件技术, 但这些不同的应用具有不同的设计需求, 使用硬件的方式也不同。计算机应用可归纳为三种不同类型。

桌面计算机 桌面机是为个人设计的计算机, 通常带有图形显示系统, 键盘和鼠标。桌面机可能是最著名的计算系统, 它的特点是个人计算机, 本书的大部分读者可能都使用过桌面计算机。桌面机强调为低端的个人用户提供高性能计算服务, 通常运行第三方软件, 也称作拆封授权软件 (shrink-wrap software)。桌面计算是计算机最大的市场之一, 许多计算技术的进步都源于桌面计算发展, 该领域只有大约 30 年的发展历史。

服务器 服务器是用于运行面向多用户的大型程序的计算机, 这些用户通常会通过网络并发访问服务器。服务器是过去的大型机、小型机、巨型计算机的现代化称谓。服务器用于承担大型的工作负载, 这类工作可以是单一的复杂应用, 如科学和工程计算应用, 也可以是大量小规模任务, 如一个大型的万维网服务器要应付成千上万的访问需求。这些应用通常基于来自另一方的软件 (如数据库或仿真系统), 但这些软件通常需经修改或者自定义以具有某些特定功能。服务器和桌面计算机技术相同, 但在计算能力和输入/输出容量方面进行了扩展。相对于其他计算机, 服务器更强调可靠性, 因为服务器承担工作往往复杂而重要, 工作负载更大, 它的故障比起单用户的桌面计算机会引起更大的损失。

不同的服务器成本和性能差别很大。低端的服务器也许就相当于一台增强的桌面计算机, 价格低廉。这种低端服务器可以用于文件存储, 小型商业应用或者简单的 Web 服务。高端服务器如巨型计算机, 可能会具有上千个处理器, GB 级到 TB 级的内存, TB 到 PB 级的外存, 造价昂贵。巨型计算机用于高端的科学和工程计算, 如天气预报, 石油勘探, 蛋白质结构确定, 以及其他大规模计算问题。尽管这类巨型计算机计算能力超强, 但它们在服务器和整个计算机市场中所占比例甚小。

嵌入式计算机 嵌入式计算机是指嵌入到其他设备内部并且执行预定功能或者程序的计算机。嵌入式计算机是拥有数目最多, 应用范围最广, 性能差异最大的计算机。在日常生活中, 洗衣机, 汽车, 手机, 个人数字助理, 游戏机, 数字电视, 工业中的控制系统, 各类仪器, 通信系统, 医疗领域的各类检测仪器, 治疗设备等等都存在着嵌入式计算机的身影。嵌入式系统通常将要执行的程序和硬件集成到一起, 以单一的系统出现, 这使得尽管嵌入式处理器大量存在, 但用户往往看不到他们在使用计算机。

嵌入式应用往往功能需求单一, 但对成本和功耗要求严格。例如一个手机处理器, 它的性能只要能够满足通信应用的需求就可以了, 但是作为移动设备和消费产品, 功耗和成本的要求却非常重要。尽管嵌入式处理器要求低成本, 但并非代表它们就是低性能, 嵌入式处理器要求在控制成本前提下, 尽量获得高性能, 一个要求就是嵌入式处理要有一定容错能力, 因为嵌入式处理的故障带来的后果有可能是灾难性的 (如汽车, 飞机甚至核反应堆的控制系统)。在面向消费的嵌入式产品中, 如数字家电, 嵌入式处理器的可靠性主要通

过简化设计来实现——强调高效稳定地完成某项单一功能。在大型嵌入式系统中，通常使用服务器中常见的冗余备份技术实现可靠性的保证。本书主要面对通用计算机系统，但是很多概念可以直接或者稍加修改即可适用于嵌入式计算机。所以，在某些地方，我们也会触及嵌入式计算机的话题。

图 1.1 显示了几年内嵌入式计算机数量的增长率（大约每年 40%），大大高于桌面计算机和服务器（每年 9%）。这里的嵌入式处理器主要包括手机，机顶盒，游戏机，数字电视等各类消费产品，不包括使用 8 位到 16 位处理器的低端嵌入式控制设备。

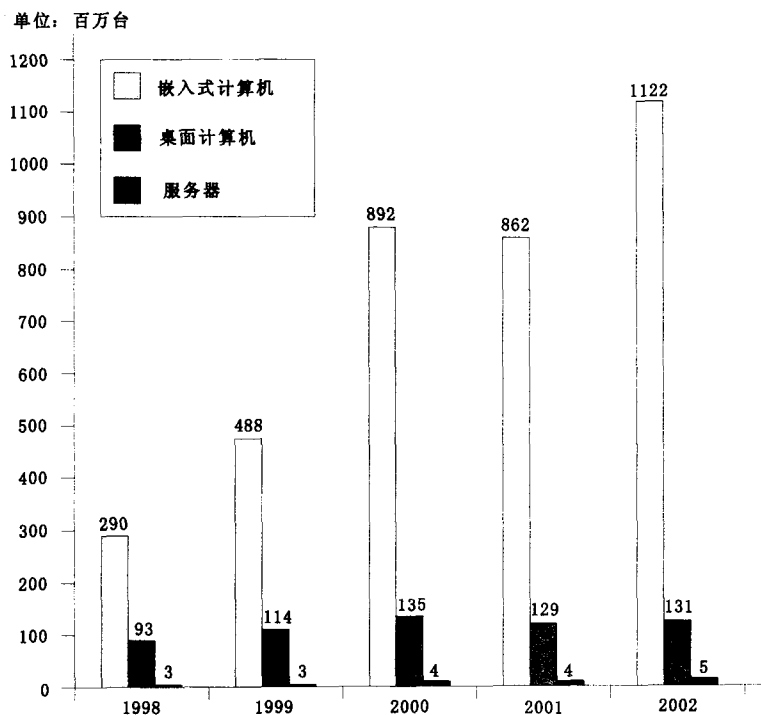


图 1.1 1998 年至 2002 年各类处理器销售数量

图 1.2 显示了几种主要架构在三类计算机（嵌入式，桌面，服务器）中的市场占有率。这里只包含了 32 位和 64 位的产品，其中 32 位处理器占了所有架构中的绝大多数。

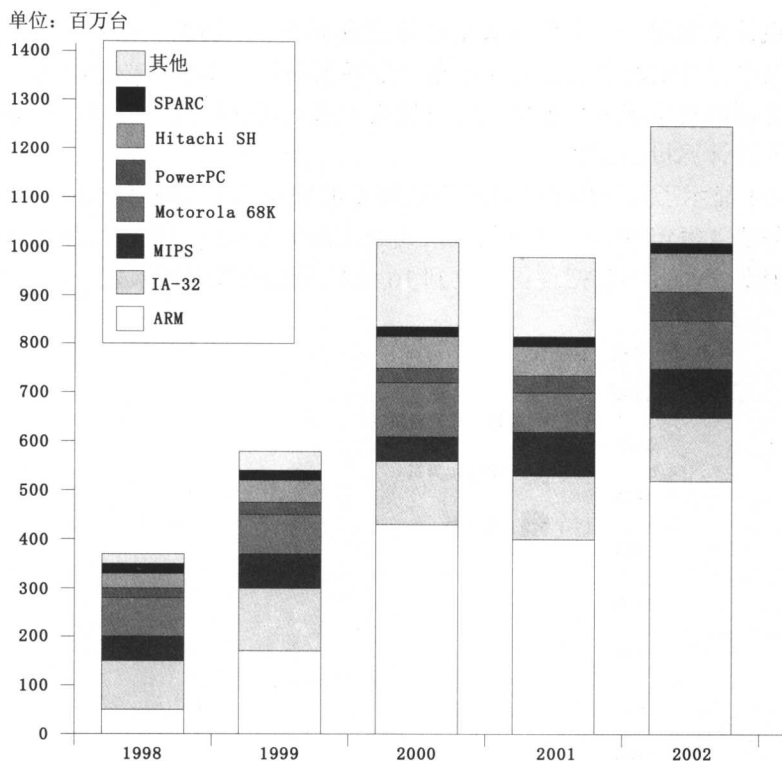


图 1.2 1998 年至 2002 年各类指令架构处理器销售数量

1.1.2 读者能从本书中学到什么

成功的程序员总要时刻考虑他们程序的性能, 因为让用户快速地得到结果对于软件的成功至关重要。在 1960 到 1970 年代, 限制计算机性能的瓶颈是内存容量。因此程序员总是遵循一个简单信条: 使用最小的内存空间取得最快的程序速度。在最近十年里, 计算机设计和内存技术进步大大降低了内存容量限制在嵌入式计算机系统之外其他计算机系统的影响。

对性能优化有兴趣的程序员现在要抛弃 1960 年代的内存模型, 转而理解新的模型: 内存的层次化结构和处理器的并行化。想要设计出高性能编译器, 操作系统, 数据库或者其他应用系统的程序员必须要增加他们的计算机组成知识。

本书的内容介绍了计算机内部的组织结构, 揭示了程序在机器中运行的机制和硬件的工作原理。通过阅读本书, 读者可以获得如下知识:

- 由高级语言如 C 或 Java 编写的程序如何被翻译成硬件能够接受的语言, 并且硬件如何执行这些程序。了解这些内容可以使读者认识到软件结构和硬件是如何共同对程序性能产生影响的。

- 了解什么是软件和硬件之间的接口, 软件的指令是如何指导硬件工作的。这对读者理解如何写出各类软件至关重要。

- 什么因素决定了程序性能, 怎样提高程序性能。这主要靠程序设计, 编译器性能