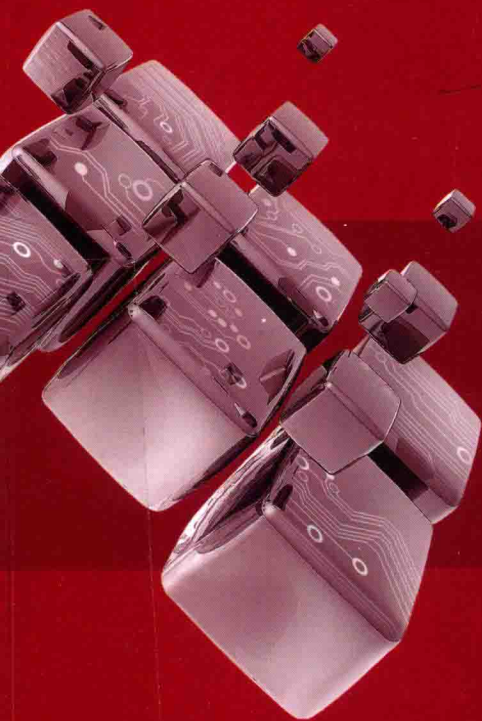


高等学校计算机规划教材

计算机组成与结构

◆ 桂盛霖 陈爱国 肖 堃 编著



中国工信出版集团



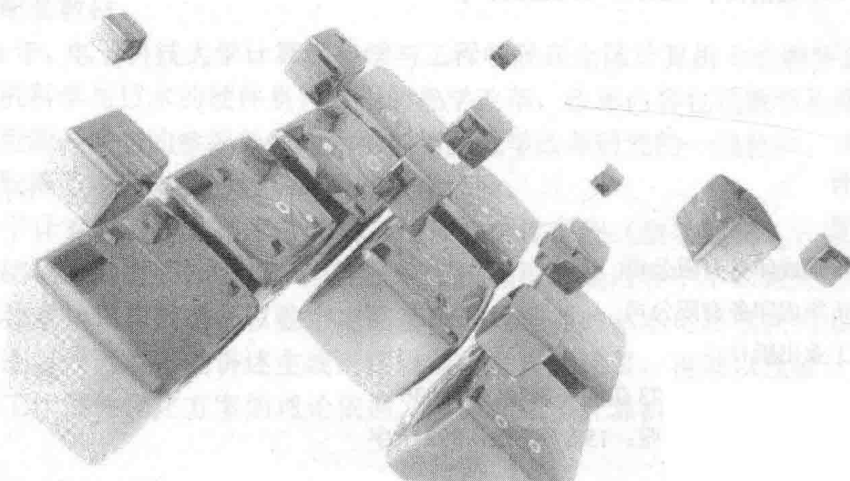
电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

高

十

计算机组成与结构

桂盛霖 陈爱国 肖堃 编著



電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

计算机组成与结构是计算机相关专业的重要专业核心课程。本书融合了计算机科学与技术专业的三门核心硬件课程（数字逻辑、计算机组成原理和计算机系统结构）的重要知识点，以数字电路基础、处理器的基本电路模块构成、指令的数据通路和控制逻辑作为本书的讲述主线，全面介绍计算机系统的基本原理、设计方法和实现。全书分为9章，内容包括：计算机系统概述，计算机的数值和编码，计算机芯片的数字电路基础，计算机芯片的基本电路组成，现代处理器基础，现代处理器的高级实现技术，存储系统，I/O系统，多核、多处理器与集群。

本书可作为高等院校计算机、信息安全、通信、电子、自动化、电气工程等专业计算机组成与结构课程的教材，也可供计算机系统的技术人员学习和参考。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

计算机组成与结构 / 桂盛霖等编著. —北京：电子工业出版社，2017.9

ISBN 978-7-121-32539-7

I. ①计… II. ①桂… III. ①计算机体系结构—高等学校—教材 IV. ①TP303

中国版本图书馆 CIP 数据核字（2017）第 203881 号

策划编辑：冉 哲

责任编辑：冉 哲

印 刷：三河市华成印务有限公司

装 订：三河市华成印务有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本：787×1 092 1/16 印张：15.5 字数：392 千字

版 次：2017 年 9 月第 1 版

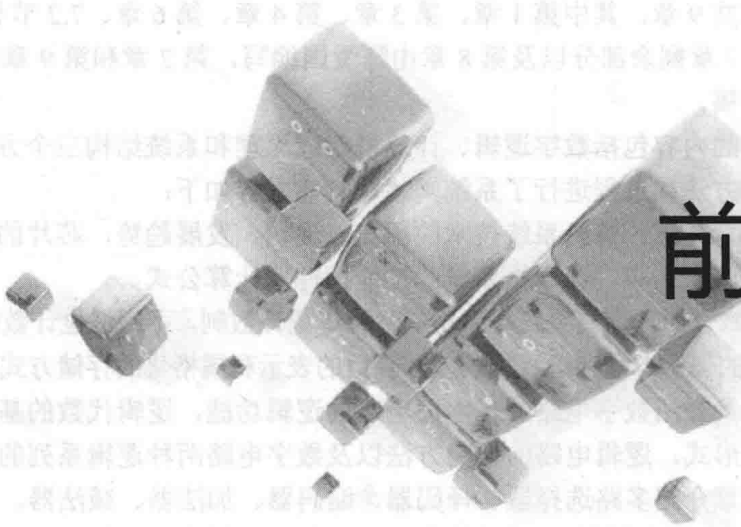
印 次：2017 年 9 月第 1 次印刷

定 价：35.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询方式：ran@phei.com.cn。



前言

电子科技大学于2016年启动了“互联网+”复合型精英人才培养计划，以“互联网+”复合培养专业为载体，旨在培养具有良好的人文精神和互联网思维、扎实的数学与自然科学知识、宽厚的移动互联网知识和信息深度分析与应用能力、扎实的“互联网+”核心知识与能力，以及良好的国际视野和创新能力，面向“互联网+”国家战略需求的复合型精英人才。本书的参编教师承担了该培养计划中平台核心课程“计算机组成与结构”的建设工作，因此本书也作为该课程的配套教材。

早在2010年，电子科技大学计算机科学与工程学院在全国计算机专业教学改革浪潮中也开始了对计算机科学与技术的硬件系列课程的教学改革，改革内容包括教学思路、内容和教学方法的全面更新。本书的参编教师均为承担相关教学改革研究的一线教师，具有多年的丰富教学经验，取得了较为丰硕的教学成果和教学经验。

本书融合了计算机科学与技术专业的三门核心硬件课程（数字逻辑、计算机组成原理和计算机系统结构）的重要知识点，摒弃了这三门课程传统内容中不太相关的琐碎知识点，引入了最新的技术知识和数据，以数字电路基础、处理器的基本电路模块构成、指令的数据通路和控制逻辑作为本书的讲述主线，逻辑清晰且自然连贯，再辅以性能计算公式贯穿全书，作为各工作部件优化方案的理论依据，全面介绍了计算机系统中的基本原理、设计方法和实现。

本书在编写过程中还注重对考研要求的相关考点进行尽量多的覆盖，并对处理器的相关章节从基本原理到设计方法再到代码实现的完整过程进行了系统性的详细讲述。在正文中还穿插了“快速练习”等提示，提示读者进行进一步的思考。部分章节还配有课后阅读材料，对正文中的相关内容进行补充或进一步的解释。

作者充分查阅和对比了目前国内外主流的组成原理类和系统结构类教材以及相关材料，进行了多方面的研讨，从而确定了内容的编排和编写的分工。除了教学工作外，作者还承担了繁重的科研工作，因此，书中难免出现疏漏和错误，恳请读者理解和海涵。

教材的内容

全书共 9 章, 其中第 1 章、第 3 章、第 4 章、第 6 章、7.2 节和 7.4 节由桂盛霖编写, 第 5 章、第 7 章剩余部分以及第 8 章由陈爱国编写, 第 2 章和第 9 章由肖堃编写。全书由桂盛霖负责统稿。

本书的内容包括数字逻辑、计算机组成原理和系统结构三个方面的内容, 围绕现代处理器的设计方法和过程进行了系统阐述, 具体内容如下:

第 1 章介绍计算机系统技术的历史、现状、发展趋势, 芯片的主要制造过程, 现代计算机系统的内部构成, 计算机系统的性能指标及计算公式。

第 2 章介绍计算机系统中常用的几种进位计数制, 不同进位计数制之间的转换计算方法, 二进制整数的表示方法和运算规则, 浮点数的表示和规格化的存储方式, 以及数据校验的原理。

第 3 章介绍数字电路的基本门电路的逻辑功能, 逻辑代数的基本公式和基本定理, 逻辑函数标准形式, 逻辑电路的化简方法以及数字电路两种逻辑系列的实现机制。

第 4 章介绍多路选择器、译码器、编码器、加法器、减法器、移位器等的功能、输入/输出接口、内部实现的电路图及代码, 以及 S_n-R_n 锁存器和触发器的原理与实现, 给出了时钟同步时序电路的功能分析方法和寄存器的设计与实现方法。

第 5 章介绍处理器中指令的格式、寻址方式、寄存器、地址空间分配, 对软件三个方面的重要支持, MIPS 风格的单周期模型机的设计和实现过程, 处理器的异常和中断概念及其处理机制。

第 6 章介绍流水线的基本概念, 单周期模型机的流水线扩展, 解决结构冒险、数据冒险和控制冒险的策略与实现, 以及两类更先进的指令级并行的流水线架构。

第 7 章介绍存储器的分类和存储系统的层次结构, 包括 Cache、主存、虚拟存储器、外存和 ROM 的工作机制与性能指标。

第 8 章介绍 I/O 设备的相关概念及属性指标, 总线的概念和分类, 总线仲裁的原理和方式, I/O 接口的功能、结构、编址和访问方式, 以及三类常见的 I/O 数据传送控制方式。

第 9 章讨论并行硬件的基本分类, 常见的并行技术, 多处理器的互连方式, Cache 一致性问题, 以及多核微处理器和云平台的架构。

由于作者水平有限, 成稿时间较短, 书中难免有错误和不当之处, 恳请各位专家和广大读者批评指正, 我们不胜感激。如有问题请直接与作者邮件联系: shenglin_gui@uestc.edu.cn。

致谢

本书在编写过程中得到了电子科技大学教务处、“互联网+”专业和计算机科学与工程学院的相关领导和老师的大力支持和鼓励, 还有电子工业出版社对本教材出版工作的积极配合和辛勤工作, 在此一并表示诚挚的谢意。

此外还感谢实验室裴亚琳、方丹、刘一飞等研究生在教材配图和习题等方面的制作和整理。

作者
于电子科技大学

反侵权盗版声明

电子工业出版社依法对本作品享有专有出版权。任何未经权利人书面许可，复制、销售或通过信息网络传播本作品的行为，歪曲、篡改、剽窃本作品的行为，均违反《中华人民共和国著作权法》，其行为人应承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。

为了维护市场秩序，保护权利人的合法权益，我社将依法查处和打击侵权盗版的单位和个人。欢迎社会各界人士积极举报侵权盗版行为，本社将奖励举报有功人员，并保证举报人的信息不被泄露。

举报电话：(010) 88254396; (010) 88258888

传 真：(010) 88254397

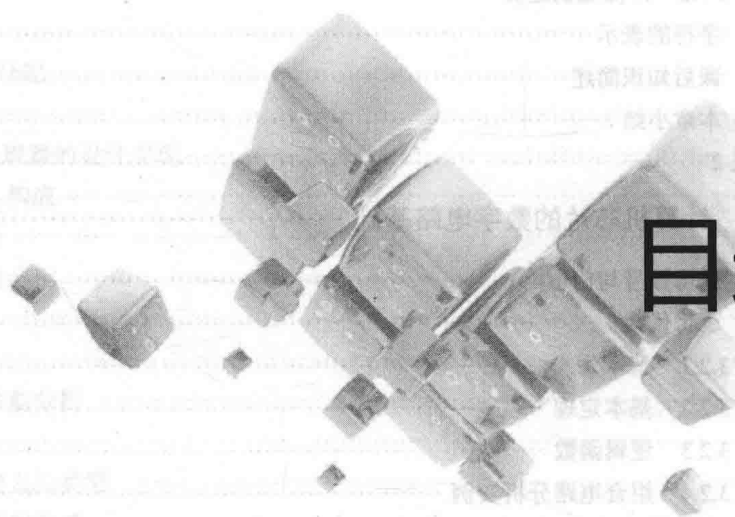
E-mail: dbqq@phei.com.cn

通信地址：北京市海淀区万寿路 173 信箱

电子工业出版社总编办公室

邮 编：100036

目录



第 1 章 计算机系统概述	1
1.1 引言	1
1.2 现代计算机	5
1.2.1 计算机的分类	5
1.2.2 计算机的组成	6
1.3 计算机的性能	8
1.3.1 什么是性能	9
1.3.2 性能的计算	9
1.3.3 性能的测量	12
1.3.4 性能的改进	13
1.4 课后知识简述	16
1.5 本章小结	17
习题 1	17
第 2 章 计算机的数值和编码	19
2.1 进位计数制	19
2.1.1 二进制和十六进制	19
2.1.2 不同进制间的转换	20
2.2 二进制数的表示和运算	22
2.2.1 二进制数的基本加/减法运算	22
2.2.2 二进制数的补码表示法	23
2.2.3 二进制数的加/减法运算	25
2.2.4 二进制数的乘/除法运算	27
2.2.5 二进制数的逻辑运算	28

2.3 浮点数的表示和运算	31
2.3.1 浮点数的表示	31
2.3.2 浮点数的运算	33
2.4 字符的表示	34
2.5 课后知识简述	36
2.6 本章小结	38
习题 2	38
第 3 章 计算机芯片的数字电路基础	39
3.1 逻辑信号与门电路	39
3.2 逻辑代数	42
3.2.1 基本公式	43
3.2.2 基本定理	44
3.2.3 逻辑函数	45
3.2.4 组合电路分析实例	50
3.3 逻辑系列	50
3.3.1 CMOS 逻辑	51
3.3.2 TTL 逻辑	61
3.4 门电路的代码实现	65
3.5 本章小结	65
习题 3	66
第 4 章 计算机芯片的基本电路组成	67
4.1 组合逻辑电路	67
4.1.1 多路选择器	68
4.1.2 译码器	70
4.1.3 编码器	71
4.1.4 加法器	73
4.1.5 减法器	77
4.1.6 移位器	77
4.1.7 其他组合逻辑电路	79
4.2 时序逻辑电路	80
4.2.1 锁存器和触发器	81
4.2.2 时钟同步时序电路分析	87
4.2.3 通用寄存器的设计和实现	89
4.3 本章小结	89
习题 4	90
第 5 章 现代处理器基础	91
5.1 指令集概述	91
5.1.1 机器指令	92

5.1.2 寄存器	92
5.1.3 寻址方式	94
5.2 MIPS 指令集结构	97
5.2.1 MIPS 指令格式	97
5.2.2 MIPS 地址空间分配	101
5.2.3 对软件的支持	101
5.3 MIPS 风格的单周期处理器的设计实现	104
5.3.1 数据通路的基本构成	105
5.3.2 数据通路的设计	110
5.3.3 指令的数据通路分析	116
5.3.4 控制部件设计	121
5.3.5 CPU 封装	123
5.3.6 单周期方式的性能分析	124
5.4 异常和中断设计	126
5.4.1 异常和中断的定义与类型	126
5.4.2 响应异常和中断的方式	127
5.4.3 异常和中断管理	129
5.4.4 带有异常和中断处理功能的单周期模型机扩展	131
5.5 课后知识	136
5.6 本章小结	137
习题 5	137
第 6 章 现代处理器的高级实现技术	139
6.1 流水线的基本概念	139
6.2 流水线模型机的基本扩展	141
6.2.1 基本的流水线模型机	141
6.2.2 流水线的分析	149
6.2.3 基本流水线的实现	153
6.3 数据冒险的解决策略	155
6.3.1 寄存器堆的写操作提前半个时钟周期	155
6.3.2 内部前推	156
6.3.3 lw 指令的数据冒险	160
6.4 控制冒险的解决策略	162
6.4.1 缩短分支的延迟	162
6.4.2 减少性能损失的其他方法	164
6.5 流水线中的异常和中断	165
6.6 指令集并行的高级实现技术	167
6.6.1 静态多发射处理器	167
6.6.2 动态多发射处理器	170
6.6.3 浮点数流水线的扩展	172
6.7 本章小结	173
习题 6	173

第7章 存储系统	175
7.1 引言	175
7.2 Cache	178
7.2.1 Cache 的块映射	179
7.2.2 Cache 的块查找	180
7.2.3 Cache 块的访问	181
7.2.4 Cache 的性能分析	183
7.2.5 Cache 的实现原理	184
7.3 主存储器	185
7.3.1 主存的结构	185
7.3.2 主存性能指标	187
7.4 虚拟存储器	189
7.4.1 分页机制	189
7.4.2 与 Cache 的关系	193
7.4.3 对进程的保护	195
7.5 外存储器	196
7.5.1 磁盘存储器	196
7.5.2 快闪式存储器	199
7.6 课后阅读材料	200
7.7 本章小结	201
习题7	201
第8章 I/O 系统	203
8.1 I/O 设备	203
8.1.1 概述	203
8.1.2 属性指标	204
8.2 I/O 互连与总线	206
8.2.1 I/O 互连方式	206
8.2.2 总线的分类	208
8.2.3 总线的仲裁	210
8.3 I/O 访问	211
8.3.1 I/O 接口功能	212
8.3.2 I/O 接口控制方式	213
8.4 本章小结	217
习题8	217
第9章 多核、多处理器与集群	219
9.1 概述	219
9.1.1 并行硬件的基本分类	219
9.1.2 常见的并行技术	221

9.1.3 多处理器网络拓扑	223
9.1.4 Cache 一致性问题	224
9.2 多核微处理器	225
9.2.1 多核架构	225
9.2.2 多核实例	226
9.3 云计算平台	230
9.3.1 云计算概念	230
9.3.2 云计算服务及部署类型	231
9.3.3 云计算的实现	232
9.4 本章小结	234
习题9	234
参考文献	236

第1章 计算机系统的组成

计算机是一种能够按照事先存储的程序自动地、连续地进行各种运算和逻辑操作的计算设备。有人将其定义为“能够按事先存储的程序（Program），自动地在高速电子器件上，对数据进行各种运算和逻辑操作，并能按照事先存储的程序，自动地控制整个系统的运行”。计算机系统的组成包括硬件和软件两大部分。硬件是指计算机系统中所有物理设备的总称，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。软件是指计算机系统中所有程序的总称，包括操作系统、应用程序、数据库系统等。计算机系统的组成是一个复杂的系统，涉及到多个学科的知识，如电子学、计算机科学、数学等。本章将介绍计算机系统的组成，包括硬件和软件两大部分。首先，我们将介绍计算机系统的组成框图，然后分别介绍硬件和软件的组成。最后，我们将介绍计算机系统的组成原理。

图 1.1 给出了计算机系统的组成框图。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。图 1.2 给出了计算机系统的组成原理。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。图 1.3 给出了计算机系统的组成原理。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。

图 1.4 给出了计算机系统的组成原理。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。

图 1.5 给出了计算机系统的组成原理。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。

图 1.6 给出了计算机系统的组成原理。图中可以看到，计算机系统由多个部分组成，包括中央处理器（CPU）、总线系统、内存系统、输入输出设备等。这些部分通过总线系统相互连接，形成一个完整的计算机系统。

第 1 章

计算机系统概述

1.1 引言

计算机是一种能按照事先编制好的程序自动化地进行算术和逻辑操作的计算设备。自人类第一台数字可编程计算机 Colossus 于 20 世纪 40 年代发明开始, 计算机的计算性能和物理形态都发生了天翻地覆的变化。图 1-1 展示了由工程师 Tommy Flower 于 1943 年在英国建造的世界上第一台可编程电子计算机 Colossus, 它用于破译第二次世界大战中德军的密码。与 Colossus 齐名的同时期的数字计算机还有美国的 ENIAC (Electronic Numerical Integrator And Computer)。由于 Colossus 的严格保密, 因此许多不知道 Colossus 的教科书误将 ENIAC 认作世界第一台电子计算机。这一时期的计算机使用了上万个真空管, MTBF (Mean Time Before Failure, 平均故障间隔时间) 非常短, 发生两次故障之间的时间间隔一般为数分钟, 可靠性非常低, 而且功耗巨大。

图 1-2 展示了 1981 年第一台 IBM 个人计算机, 这个时期的计算机开始采用超大规模集成电路 (Very Large Scale Integration, VLSI), 将整个 CPU 集成到一块硅芯片上, 这样具有更高的可靠性和更低的功耗, 运算速度也更快, 价格更低。图 1-3 展示了苹果公司的智能手表产品 Apple Watch。现代的计算机系统已经变得更小更移动化, 功耗更低, 智能化程度更高。通过这三个不同时期的计算机产品对比, 可以直观地看出计算机制造技术的进步趋势: 体积越来越小, 功耗越来越小, 价格越来越低, 计算速度越来越快, 智能化程度越来越高。近几十年推动计算机制造技术的发展因素是多方面的, 但总体符合由英特尔创始人之一 Gordon Earle Moore 提出来的摩尔定律: 集成电路上可容纳的晶体管数量, 约每隔 18 个月便会增加一倍。

下面总结了计算机技术发展的几个主要阶段。

1. 非集成电路年代

1883 年, 爱迪生发现了热的灯丝发射电荷的现象, 并被称之为“爱迪生效应”。1904 年, 英国伦敦大学教授弗莱明 (John Ambrose Fleming) 基于爱迪生效应, 研制出检测电波

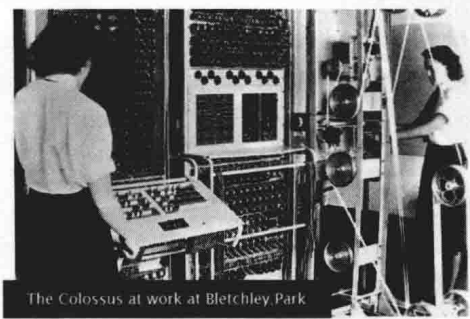


图 1-1 在 Bletchley Park 运行的 Colossus



图 1-2 第一台 IBM 个人计算机



图 1-3 Apple Watch

用的第一只真空二极管，如图 1-4 所示，从而宣告人类第一只电子二极管的诞生。1906 年，美国发明家德福雷斯特（De Forest Lee）在对弗莱明（John Ambrose Fleming）的二极管进行实验时发现，若在阳极 A 和阴极 K 的中间加上栅状的电极 G，在 G 上加负的偏压就可以使阳极电流为零。如果改变栅极电压，就可以使阳极电流发生相应的变化。根据这个实验结果，德福雷斯特在二极管的两个电极之间增加了一个形状像栅栏的电极，从而制造出真空三极管，可用作电流开关和信号放大器。如前所说，ENIAC 使用了 18000 个真空电子管，重达 30 吨，占地超过 1000 平方英尺（1 英尺=0.3048 米），用于计算弹道方程以计算火炮射程表。

由于真空电子管在体积、功耗、速度等方面的约束，人们开始探索使用半导体材料制作和真空管功效相当的晶体管。1947 年，贝尔实验室的三位物理学家肖克利（William Shockley）、巴丁（John Bardeen）、布拉坦（Walter Brattain）制成了世界上第一个固体放大器晶体三极管，如图 1-5 所示。晶体管的发现在电子学发展史上具有划时代的意义，此后，晶体管的体积不断变小，与之配套的电阻、电容、线圈、继电器、开关等元件，也沿着小型化的道路被压缩成微型电子元器件。

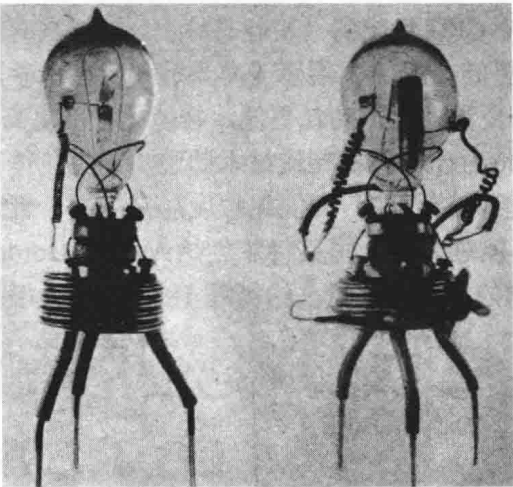


图 1-4 1904 年出现的第一只实用型的真空二极管

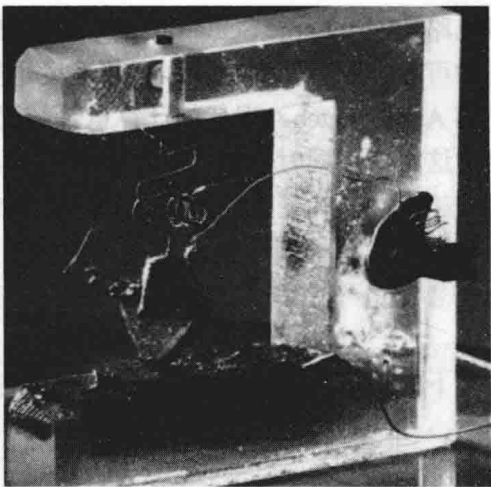


图 1-5 1947 年第一只晶体管

2. 集成电路年代

1958年,美国德州仪器公司的青年工程师基尔比(Jack Kilby),大胆提出了用一块半导体硅晶片制作一个完整功能电路的新方案。他在研制微型组件的晶体管中频放大器时,用一块硅晶制成了包括电阻、电容在内的分立元件实验电路。到1958年底,已经解决了半导体阻容元件和电路制作中的许多具体工艺问题,确定了集成电路的标准封装尺寸,为大规模工业化生产做好了各项准备。1959年,美国仙童公司的诺伊斯(Robert Noyce)研究出一种二氧化硅的扩散技术和PN结的隔离技术,从而完成了集成电路制作的全部工艺。紧接着,光刻技术和其他技术也相继发明,使得人们可以把晶体管和其他功能的电子元件压缩到一小块半导体硅晶片上。

1964年4月7日,IBM公司研制出世界上第一个采用集成电路的通用计算机IBM 360系统,随后计算机技术的发展进入了集成电路时代。之后随着大直径硅单晶材料性能的提高以及离子束和新隔离技术的应用,特别是光刻工艺精度的不断提高,使晶片上的电子元件的几何尺寸越来越小,人们开始把一个线路系统或一台电子设备所包含的所有晶体管和其他电子元件都制作在一块晶片上,从而大大缩小了体积并提高了可靠性。

1971年,Intel发布第一款微处理器4004,如图1-6所示,时钟频率108kHz,集成了2250个晶体管,制造工艺^①10 μm ;1978年,Intel微处理器8086发布,如图1-7所示,时钟频率10MHz,集成了2.9万个晶体管,制造工艺3 μm ;1985年,Intel微处理器80386发布,如图1-8所示,时钟频率33MHz,集成了27.5万个晶体管,制造工艺1.5 μm ;2008年,Intel微处理器Core i7(4核)发布,如图1-9所示,时钟频率2.5GHz,制造工艺14~45nm。

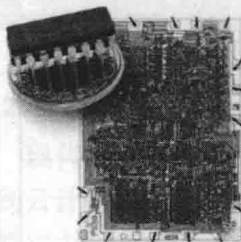


图 1-6 Intel 4004

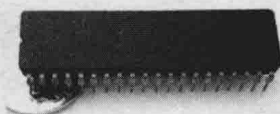


图 1-7 Intel 8086

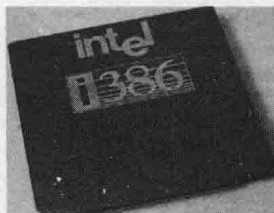


图 1-8 Intel 80386

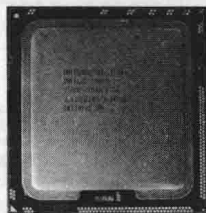


图 1-9 Intel i7

图1-10展示了从1960年到2010年间单个芯片上晶体管数变化的趋势,与摩尔定律保持一致。推动计算机运算能力发展的因素除了制造工艺外,还有计算机体系结构的创新。从20世纪80年代中期到21世纪初出现的指令级并行技术、浮点计算架构等新的先进处理器架构技术,进一步推动处理器的性能以每年52%的速度增长。2003年之后,由于功耗、指令级并行程度和存储器延迟等因素限制,单核处理器的性能增速放缓,约每年增长22%,如图1-11所示。从2006年开始,单片微处理器芯片中开始加入多个处理器,从追求更快的主频和更短的程序响应时间转变为追求更大的任务吞吐率的多核处理器。表1-1给出了近两年发布的三款微处理器的时钟频率、核数、工艺、功耗等参数。

^① 制造工艺是指晶体管的尺寸,即晶体管栅极的长度。

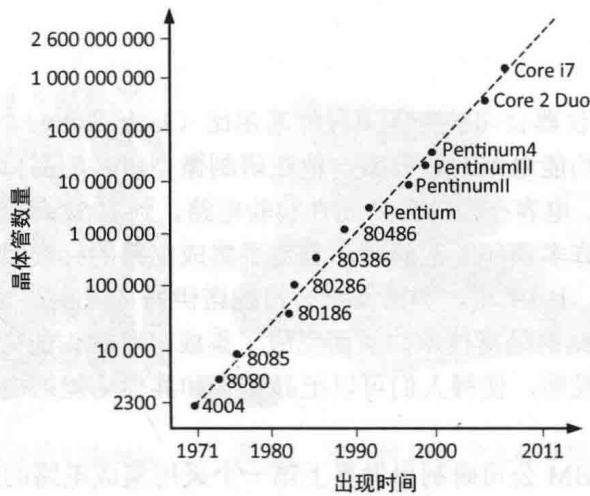


图 1-10 单个芯片上的晶体管数变化趋势

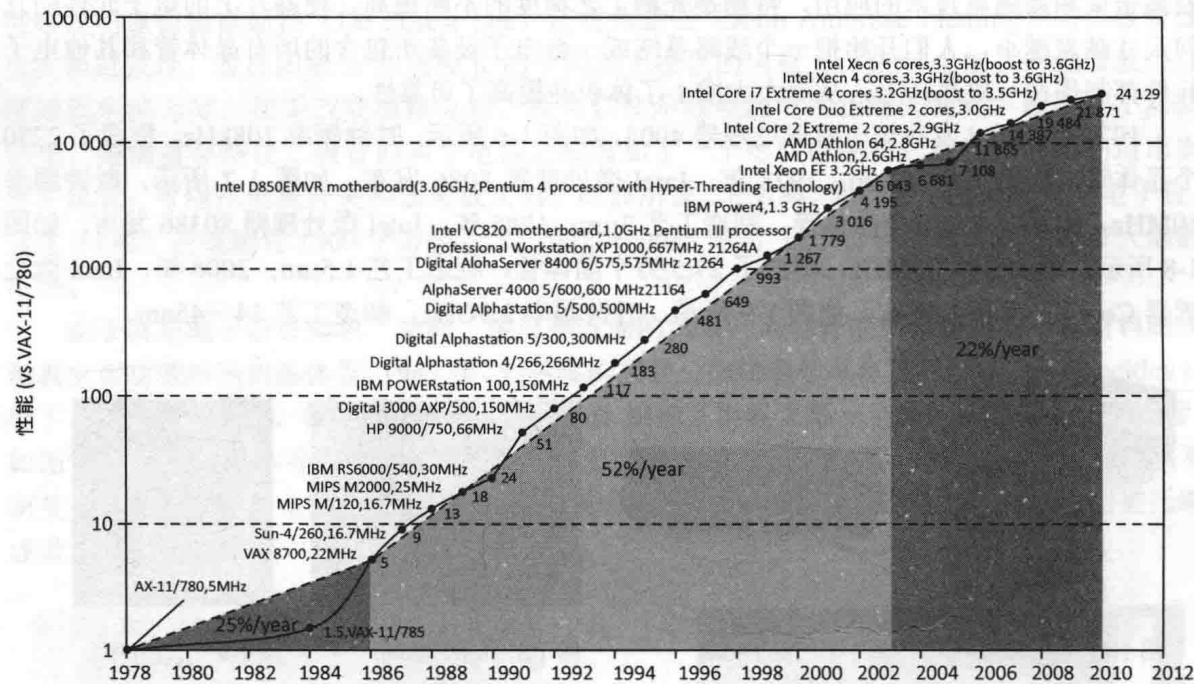


图 1-11 处理器性能增长的过程

表 1-1 三款微处理器的参数

产 品	Exynos 7420	Intel i3-4350	Intel i7-4790
发布时间	2015 年	2014 年	2014 年
时钟频率	2.1GHz	3.6GHz	3.6GHz
处理器位数	64 位	64 位	64 位
制造工艺	14nm	22nm	22nm
核数	8	2	4
功耗	小于 10W	89W	127W

表 1-1 中提到了功耗,目前在微处理器制造过程中占统治地位的集成电路技术是 CMOS (互补性金属氧化半导体),其主要功耗来源是动态功耗,即晶体管在开关过程中产生的功耗。影响微处理器动态功耗的因素有:负载电容、工作电压和晶体管的开关频率。

$$\text{功耗} = \text{负载电容} \times \text{工作电压}^2 \times \text{开关频率}$$

其中,负载电容是指芯片上导线和晶体管的电容,它是连接到输出上的晶体管数量和制造工艺的函数;开关频率是时钟频率的函数。从微处理器的动态功耗公式可以看出,尽管微处理器的时钟频率不断增长,但由于每次工艺换代时,电压可以降低约 15%,因此微处理器功耗的增长幅度较为平滑。另外,从表 1-1 中还可以看出,采用 ARM 架构的三星 Exynos 7420 处理器的功耗远低于两款同期的 Intel 处理器,这也是 ARM 架构处理器成为移动市场上主流处理器的一个重要原因。

本书将系统性地为读者详细介绍计算机主要硬件模块的工作机制。在学习完本书之后,你将明白以下一些问题:

- 影响处理器及计算机系统性能的因素有哪些?如何进一步提升处理器及计算机系统的性能?如何量化评估改进的效果?
- 哪些操作系统的功能必须要有硬件的支持才能实现?软件如何与硬件进行协同工作?
- 处理器的语言是什么?C 或 Java 等高级程序语言是如何在处理器上执行的?为何理解处理器的工作机制对编写高级程序语言有帮助?
- 现代处理器有哪些主要工作部件?这些工作部件的工作原理和设计架构是什么样的?处理器的每个模块是如何实现并结合在一起工作的?

1.2 现代计算机

1.2.1 计算机的分类

现代计算机的应用领域已经进入到人们工作和生活的方方面面,从智能手机到个人电脑,再到云计算中心,这些不同的应用领域对计算机系统有不同的设计需求。现代计算机按应用类别大致可分为以下三类。

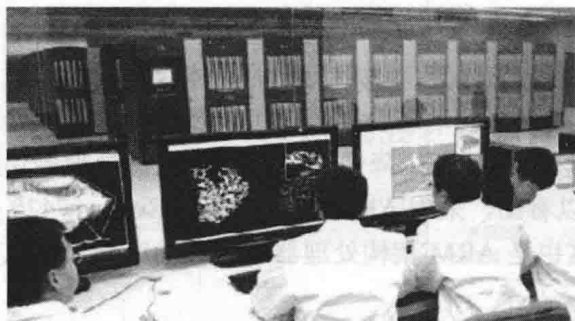
1. 桌面计算机

最典型的代表就是桌面计算机和笔记本电脑。这类计算机偏重对单用户提供良好的性能支持,满足用户工作需求或生活娱乐需求,价格相对低廉,一般需要使用桌面操作系统 Windows 或 Linux。最低端的上网本价格甚至在人民币千元以下。

2. 服务器

服务器涵盖大型机、小型机、超级计算机等,通常需要借助网络访问。服务器可支持大量用户同时访问,可执行单个定制化应用,也可执行大量的简单作业,如 Web 服务器。服务器的价格一般比桌面计算机贵,体积也比桌面计算机大,能提供更强的计算能力和更多的计算资源。服务器的设计通常要求更高的可靠性。高端的服务器又可分为两种类型:一种称为超级计算机,一般由成百上千个处理器组成,内存为 TeraByte 级,外存为 PetaByte 级,价格

从几百万到数亿美元不等,主要用于高端科学或工程计算,如人类基因序列分析、宇宙大爆炸过程模拟等;另一种由成千上万台服务器的集群构成,如 Google 等互联网公司所使用的数据中心。服务器集群的形态如图 1-12 所示。



(a) 我国天河二号超级计算机



(b) Google 数据中心

图 1-12 服务器集群的形态

3. 嵌入式计算机

嵌入式计算机的设计目标是在资源受限的条件下运行一个或一组应用程序,并且通常和硬件集成在一起交付给用户。嵌入式计算机的分布范围最为广泛,并且形态各异,从飞机和汽车的各类控制器到人们日常生活中的各类消费电子都属于嵌入式计算机范围。由于各类嵌入式计算机的应用需求存在显著差异,因此各类嵌入式计算机的设计目标存在差异:

- **移动终端。**人们对移动设备的需求量已经远远超过了对桌面计算机的需求。仅在 2016 年的第 2 季度,中国智能手机市场出货量超过 1 亿台,排名前三的厂商依次为华为、OPPO 和 Vivo。移动终端的敏感因素包括多个方面,如用户体验、生态环境支持、价格、便捷性、质量等。例如,诺基亚 Symbian 系统的衰败就是因为上述多方面原因综合起来的结果。
- **工业控制系统。**各种智能测量仪表、数控装置、现场总线仪表及控制系统、工业机器人、机电一体化设备中均包括了各种类型的嵌入式计算机。特别是飞机中的飞行控制等安全关键系统对实时性和可靠性要求更高,对于系统运行中出现的任何一个小问题,若处理不当,都可能引发严重的安全事故。

当然计算机还有其他的分类方法,例如按照计算机处理的数据形态不同可以分为数字计算机、模拟计算机和数模混合计算机;按计算机的使用范围不同可以分为专用计算机和通用计算机。本书的介绍重点将放在通用计算机上,但大多数概念和技术可直接或稍微修改后用于嵌入式计算机等其他类别的计算机。

1.2.2 计算机的组成

众所周知,计算机系统的组成分为两大部分:软件和硬件,如图 1-13 所示。与硬件紧密协作的软件是驱动程序,用于直接控制硬件的工作过程;除少数对价格特别敏感的单一用途的计算设备外,大多数通用计算机都会使用操作系统来管理计算机的各类资源,以增加应用