



新 编 计 算 机 类 本 科 规 划 教 材

计算机系统结构

蔡启先 编著



電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

新编计算机类本科规划教材

计算机系统结构

蔡启先 编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书主要讲述计算机系统结构的基本概念、系统构成技术和性能分析方法。全书共 9 章, 内容涵盖计算机系统结构的基本概念和基本分析原理, 包括 RISC、流水线技术、向量处理技术、存储系统、输入/输出系统; 并行技术和计算机系统结构的发展, 包括并行计算机(含互连网络)、多处理器系统等; 以及现代 PC 的系统结构和 DLX 虚拟处理器及其实验。每章之后均提供大量习题。本书努力反映现代计算机系统结构的最新理念和最新成果, 内容新颖丰富。其独特之处是结合现代 PC 系统结构进行分析和评测, 并提供基于 PC 平台开展的计算机系统结构实验。

本书可作为高等学校计算机类专业及电气信息类专业《计算机系统结构》课程的通用教材, 也可作为相关专业研究生、教师和科技人员的参考书。

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

计算机系统结构 / 蔡启先编著. —北京: 电子工业出版社, 2009.1

新编计算机类本科规划教材

ISBN 978-7-121-07724-1

I. 计… II. 蔡… III. 计算机系统结构—高等学校—教材 IV. TP303

中国版本图书馆 CIP 数据核字(2008)第 175179 号

策划编辑: 何 雄

责任编辑: 冯小贝

印 刷: 涿州市京南印刷厂

装 订: 涿州市桃园装订有限公司

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: 787×1092 1/16 印张: 20 字数: 512 千字

印 次: 2009 年 1 月第 1 次印刷

印 数: 4000 册 定价: 29.00 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。

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

服务热线: (010)88258888。

前 言

计算机系统结构历来为高校计算机科学与技术专业必修的主干专业基础课程。计算机的发展历史说明,计算机性能的不断提高必须依靠器件的变革和系统结构的改进。今天,在器件潜力几乎达到极限的情况下,计算机系统结构的改进尤为重要。该门课程主要反映现代计算机在系统结构上的新思想、新技术,如流水线处理、向量处理、并行处理、多处理器结构等,要求学生从分析和评测的角度把握计算机系统的设计。这对培养学生掌握和应用现代计算机系统来处理复杂计算问题具有重要意义,也为学生今后从事计算机系统软、硬件开发打下良好的基础。

国家教育部在《高等教育面向 21 世纪教学内容和课程体系改革计划》中明确指出,要“面向 21 世纪社会、经济、科技、文化的发展,改革我国高等教育中与其不相适应的教学内容和课程体系”,该计划把“基础课程、核心课程的教学内容体系及教材和教学手段、教学方法的创新”列为改革的主要内容。因此,在现代计算机技术日新月异、计算机应用迅猛扩展深入的今天,在高校教育逐步由精英教育向大众化教育转移的今天,在国家经济发展大量地、普遍地需要工程应用型计算机人才的今天,必须重新审视这门课程的现状和特点,按照新型人才的培养目标和定位,构建新的计算机系统结构课程体系,从而提高该课程的教学质量。

计算机系统结构课程具有下述特点:

1. 综合性强。计算机系统结构一般安排为计算机专业高年级课程,它需要用到几乎所有的计算机专业基础知识和相关的前继专业课程知识,主要需要计算机组成原理、汇编语言程序设计、高级语言程序设计、数据结构、操作系统、编译原理等课程的知识。教学中要求学生对各课程的知识融会贯通,教学的难度较大。

2. 理论性强。由于课程内容抽象繁杂、概念多,如果教学处理不好,那么容易让学生感到学习乏味,明显增加了教学难度。

3. 本课程教材内容多是针对大中型计算机系统描述的,常以一般学生接触不到的机型作为系统结构举例,而结合 PC 应用实际、结合现代 PC 系统结构技术的内容几乎没有,严重脱离学生的实际情况。

4. 缺乏实验环境,学生无法获得对计算机系统结构性能改进的直观认识。如果真正要求学生对其理论与技术有直观的接触,则需要深入到现代大中型计算机内部进行实践,这对我国绝大多数高校来说是不可能做到的。因此,长期以来,我国高校在开设这门重要课程时,仅仅停留在理论讲授上,相应的实践教学尚是空白,学生面对枯燥理论,学习兴趣大减,这对提高教学质量带来极大障碍。

针对这门课程的特点和原有的不足,我们从以下三方面对教材进行了重新规划与调整。

1. 修改教学要求,明确教学目的。我们根据全国大多数高校的定位,把对现代计算机系统结构的分析和评测作为构建的主要目标。很难想象,一个生活、工作在计算机广泛应用

的信息化时代的计算机高级技术人员，如果对计算机系统结构知之甚少，又怎么能够用好计算机呢？因此，在认识和了解一般计算机系统结构新技术的基础上，加强对计算机系统结构的学习是完全必要的。

2. 在总的教学目标的要求下，大刀阔斧地对教学内容进行重组，构建符合教学目标、既有理论又有实践的新的课程体系。我们积极开展本课程教学内容改革，注意吸收国内外同类教材的教学思路，结合我国当代社会发展对计算机人才知识结构的要求，在教学内容上进行压缩和扩充，精简了有关大中型计算机的过多介绍，增加和突出 PC 上常用的新的系统结构，突出基本知识，注意和前继课程的内容贯通，形成具有一定特色的教学内容和教学大纲。

理论上，既讲清一般原理，又紧密结合个人计算机系统结构。例如，结合 80x86 指令系统，阐述从 CISC 到 RISC 的发展；在介绍各项新技术时，尽量列举 PC 上应用相关技术的实例，并专辟一章介绍 PC 的系统结构，重点是当前主流 PC 的微结构及其发展。应该说，结合系统结构的原理，可以比较全面地介绍与学生密切相关的 PC 的系统结构。

如何在常规的条件下利用普通的 PC，让学生直接用到有关计算机系统结构的新思想和新技术，得到实践学习的机会，从而大大提高学生对现代计算机系统的认识，大大提高这门重要课程的教学效果，值得我们花大气力进行研究。为此，我们引进了 DLX 虚拟处理器实验。DLX 处理器是 1995 年美国斯坦福大学的 John L. Hennessy 和加利福尼亚大学伯克利分校的 David A. Patterson 在其 *Computer Architecture: A Quantitative Approach* 一书中首次提出的一个虚拟的 32 位处理器。该处理器不仅体现了当今多种机器（如 AMD 29K、DEC station 3100、HP 850、IBM 801、Intel i860、MIPS M/120A、MIPS M/1000、Motorola 88K、RISC I、SGI 4D/60、SPARCstation-1、Sun-4/110、Sun-4/260 等）系统结构的共同特点，还将体现未来一些机器的设计思想。特别是 DLX 虚拟处理器提供了一个基于 PC 的研究平台，使读者可以在 PC 上模拟新的处理器技术。

3. 我们在全面了解 DLX 虚拟处理器的结构和工作原理的基础上，探讨利用 DLX 虚拟处理器进行处理器指令系统的设计、流水线的设计与实现、并行处理的设计与实现等带有新一代处理器思想和技术的实践教学体系，总结出有典型教学意义的注重设计性和综合性训练的实验。例如，让学生通过实验，采用旁路技术消除数据相关，采用增加运算单元的方法消除控制相关，采用优化程序的方法来提高流水线的性能等，这是一项难度较大的、具有开创意义的教学研究工作。

全书共分 9 章，第 1 章“计算机系统结构的基础”概述了计算机系统结构的概念，分析了相关原理、评价标准及系统结构的发展，为后续章节打下基础。第 2 章“计算机指令系统”从数据表示讲到指令系统的设计和优化，为读者学习流水线技术做好准备。第 3 章“存储系统”叙述存储系统的结构和性能提高方法，并介绍了现代 RAID 和 SAN 新技术。第 4 章“输入/输出系统”简要介绍了现代输入/输出系统，包括总线、通道和外围处理机。第 5 章“流水线技术与向量处理技术”从流水线基础（包括 RISC 技术、先行控制技术等）开始，讲述流水线原理和性能分析，进而介绍向量处理机。第 6 章“并行处理机”强调并行处理概念，介绍了 SIMD 计算机、阵列计算机、相联计算机，以及现代超标量流水线、超级流水线和超长指令字处理机，并对实现并行计算所必不可少的互连网络原理也进行了描述。第 7 章“PC 的系统结构”如前所述，集中地介绍了 PC 的发展、系统结构和相关的主板芯片组知识。

第 8 章“多处理器系统”对多处理器系统的基本结构、多处理器线程级并行处理技术和并行程序设计等进行了介绍，并以介绍我国自主研发的曙光 5000 超级计算机作为结束。第 9 章“DLX 虚拟处理器及其实验”是本书的实践基础，它将帮助读者综合地了解和运用有关处理器指令系统的设计、流水线的设计与实现等方面的知识，有助于对本书前面章节所述内容的理解。各章之后都有大量的习题供读者练习。

在章节编排上，本书注意中心知识的贯通。第 1 章是全书的概要，其中提出的计算机系统结构分析原理和评价原则贯穿全书。第 2 章~第 5 章主要围绕流水线技术展开，并逐步突出并行性结构思想，此后则以并行技术为主线叙述各章，并扩展到 PC 和现代集群系统。对与前继课程如计算机组成原理可能重叠的内容，注意从系统结构这一新的角度去阐述，而对学生已经熟悉的知识则给予提示或简要描述。本书的讲解深入浅出、突出重点，并精选了一些实例。

在撰写本书的过程中，参考了大量书籍和网络上的信息，本人对这些信息的作者致以诚挚的谢意。同时本书的编著出版得到了电子工业出版社、广西工学院的大力支持，谨致谢忱。

作者深感计算机技术的发展神速，虽然尽力反映最新的系统结构技术，但限于作者水平，确实很难全面反映其进展。遗漏之处和不当之处敬请读者批评指正。

蔡启先
2008 年 11 月

目 录

第 1 章 计算机系统结构的基础	(1)
1.1 计算机的发展及其分类	(1)
1.1.1 计算机的发展	(1)
1.1.2 计算机的分类	(2)
1.2 计算机系统结构的概念	(5)
1.2.1 计算机系统的层次结构	(5)
1.2.2 计算机系统结构的定义	(6)
1.2.3 计算机系统设计的主要方法	(9)
1.3 计算机系统的评价标准	(10)
1.3.1 计算机系统设计的定量分析原理	(10)
1.3.2 计算机系统的评价	(17)
1.4 器件、软件、应用对计算机系统结构的影响	(20)
1.4.1 计算机系统结构的演变	(20)
1.4.2 器件对计算机系统结构的影响	(22)
1.4.3 软件对系统结构的影响	(23)
1.4.4 应用对系统结构的影响	(25)
1.5 计算机系统结构的分类	(26)
1.5.1 弗林分类法	(26)
1.5.2 其他分类法	(27)
本章小结	(28)
习题 1	(29)
第 2 章 计算机指令系统	(32)
2.1 数据类型	(32)
2.1.1 数据类型	(32)
2.1.2 数据表示与数据结构	(32)
2.1.3 浮点数据表示	(33)
2.1.4 自定义数据表示	(38)
2.1.5 其他数据表示	(41)
2.1.6 引入数据表示的原则	(44)
2.2 寻址技术	(45)
2.2.1 编址方式	(45)
2.2.2 寻址方式	(48)
2.2.3 程序装入与定位方式	(49)

2.3	指令系统的设计	(51)
2.3.1	指令格式的优化设计	(51)
2.3.2	指令功能的设计	(56)
2.4	指令系统的改进	(57)
2.4.1	复杂指令系统 (CISC)	(57)
2.4.2	精简指令系统 (RISC)	(59)
2.4.3	指令系统的优化发展方向	(62)
	本章小结	(63)
	习题 2	(64)
第 3 章	存储系统	(66)
3.1	存储系统原理	(66)
3.1.1	存储系统的概念	(66)
3.1.2	存储器的层次结构	(68)
3.2	虚拟存储系统	(68)
3.2.1	虚拟存储器的地址映像和地址变换	(69)
3.2.2	页面替换算法	(74)
3.2.3	提高主存命中率的方法	(77)
3.3	高速缓冲存储器	(79)
3.3.1	Cache 存储器的基本工作原理	(79)
3.3.2	Cache 的一致性及性能分析	(85)
3.3.3	11 种先进的 Cache 性能优化方法	(87)
3.4	三级存储系统	(90)
3.5	并行存储器	(91)
3.5.1	存储器的频带平衡	(91)
3.5.2	并行存储器	(91)
3.6	RAID 系统	(96)
3.7	存储域网络	(99)
	本章小结	(100)
	习题 3	(102)
第 4 章	输入/输出系统	(106)
4.1	输入/输出系统概述	(106)
4.1.1	输入/输出系统的特点	(106)
4.1.2	基本的输入/输出方式	(107)
4.2	输入/输出总线	(108)
4.2.1	总线概述	(108)
4.2.2	输入/输出总线的设计	(110)
4.3	I/O 处理机	(113)
4.3.1	通道	(113)

4.3.2 外围处理机	(115)
本章小结	(117)
习题 4	(117)
第 5 章 流水线技术与向量处理技术	(119)
5.1 流水线基础	(119)
5.1.1 指令的重叠执行	(119)
5.1.2 RISC 对流水线技术的支持	(121)
5.1.3 先行控制技术	(125)
5.2 流水线技术	(135)
5.2.1 流水线工作原理	(135)
5.2.2 流水线的分类	(137)
5.2.3 流水线的主要性能及其分析	(140)
5.3 流水线的相关性分析及处理	(151)
5.3.1 局部相关及处理	(151)
5.3.2 全局相关及处理	(153)
5.4 向量的流水线处理与向量流水线处理机	(156)
5.4.1 向量处理方式	(156)
5.4.2 向量流水线处理机	(158)
5.4.3 提高向量流水线处理机性能的常用技术	(160)
本章小结	(162)
习题 5	(164)
第 6 章 并行处理机	(168)
6.1 阵列处理机	(168)
6.1.1 SIMD 计算机的基本概念和模型	(168)
6.1.2 阵列处理机的特点	(169)
6.2 阵列机中并行存储器的无冲突访问	(170)
6.3 相联处理机	(172)
6.3.1 相联处理机和相联存储器	(172)
6.3.2 相联存储器的基本组成和工作原理	(173)
6.3.3 相联检索算法	(174)
6.4 脉动阵列机	(176)
6.4.1 脉动阵列机的结构原理	(176)
6.4.2 通用脉动阵列机的发展	(178)
6.5 互连网络	(178)
6.5.1 静态互连网络	(178)
6.5.2 动态互连网络	(183)
6.5.3 互连网络的通信问题	(187)
6.6 超标量流水线和超级流水线	(190)

6.6.1 超标量处理机	(190)
6.6.2 超流水线处理机	(192)
6.6.3 超标量超流水线处理机	(193)
6.7 超长指令字处理机	(195)
本章小结	(196)
习题 6	(198)
第 7 章 PC 的系统结构	(199)
7.1 PC 的发展回顾	(199)
7.1.1 PC 发展的简单回顾	(199)
7.1.2 影响 PC 处理器发展的两个重大事件	(204)
7.2 Intel Core 微架构	(205)
7.3 AMD 64 处理器架构	(209)
7.4 主板芯片组	(211)
7.4.1 主板芯片组的概念	(211)
7.4.2 前端总线	(213)
7.5 未来 PC 处理器的发展方向	(214)
本章小结	(215)
习题 7	(215)
第 8 章 多处理器系统	(216)
8.1 指令级并行性的限制和超线程技术	(216)
8.1.1 指令级并行性的限制	(216)
8.1.2 超线程技术	(217)
8.1.3 限制单处理器发展的其他主要因素	(218)
8.2 多处理器构成的系统结构	(219)
8.2.1 多核处理器结构	(219)
8.2.2 多处理器结构	(221)
8.2.3 多处理器并行处理定量分析和所遇到的问题	(225)
8.3 多处理器的 Cache 一致性	(227)
8.3.1 对称式共享存储器系统和多处理器的 Cache 一致性	(227)
8.3.2 分布式共享存储器系统和基于目录的 Cache 一致性	(231)
8.4 多处理器系统的同步机制	(234)
8.4.1 基本硬件原语	(234)
8.4.2 同步机制的实现	(236)
8.4.3 多线程同步机制带来的问题	(239)
8.5 集群系统	(240)
8.5.1 集群系统及其特点	(240)
8.5.2 集群系统的关键技术	(243)
8.6 多处理器系统的程序并行性	(246)

8.6.1	程序的并行性挖掘	(246)
8.6.2	支持并行程序的软件工具	(251)
8.7	曙光 5000 超级计算机	(254)
本章小结		(255)
习题 8		(257)
第 9 章	DLX 虚拟处理器及其实验	(258)
9.1	WinDLX 虚拟处理器概述	(258)
9.2	WinDLX 虚拟处理器的指令系统	(259)
9.2.1	DLX 指令集结构	(259)
9.2.2	DLX 指令集	(262)
9.3	WinDLX 汇编	(265)
9.3.1	数据定义及存储器分配	(265)
9.3.2	陷入操作	(267)
9.4	WinDLX 的流水线结构	(269)
9.4.1	WinDLX 的基本结构	(270)
9.4.2	引入流水线操作的 DLX 处理器	(272)
9.4.3	DLX 基本流水线面临的问题分析	(275)
9.5	WinDLX 虚拟处理器上机操作	(287)
9.5.1	WinDLX 的启动和配置	(287)
9.5.2	装载测试程序	(288)
9.5.3	模拟运行程序	(291)
实验 1	熟悉 WinDLX 的使用	(298)
实验 2	DLX 流水线结构相关	(300)
实验 3	DLX 流水线数据相关	(301)
实验 4	DLX 流水线指令调度	(303)
参考文献		(308)

第1章 计算机系统结构的基础

什么是计算机系统结构？为什么要学习计算机系统结构？很多读者打开本书后都有这样的疑问。本章将从计算机系统的多级层次结构入手，提出计算机系统结构的概念，并说明系统结构、计算机组成和计算机实现的含义与它们之间的关系。在此基础上，重点介绍计算机系统结构的分析原理和评价标准，并讨论器件、软件、应用等因素对系统结构发展的影响；最后介绍计算机系统结构的分类。在本章结尾的小结中，将得到现代计算机系统性能的提高主要依赖系统结构的改进这一结论。

1.1 计算机的发展及其分类

1.1.1 计算机的发展

自 1946 年第一台电子计算机诞生以来，计算机技术在 60 余年的时间里得到了突飞猛进的发展。今天，价值不到 500 美元的个人计算机就比 1985 年花 100 多万美元购买的大型计算机具有更高的性能、更大的存储器和磁盘空间。计算机的高速发展不仅在于器件技术的进步，更离不开计算机系统结构的创新。

在大规模集成电路出现之前，器件的发展和对冯·诺依曼系统结构的改进使计算机性能大约以每年 25% 的速度提高。20 世纪 70 年代，VLSI（超大规模集成电路）技术和微处理器的出现加速了计算机的发展，其性能约以每年 35% 的速度提高，并逐步转向以微处理器为基础来组成计算机系统。在软件上，人们已极少使用汇编语言对计算机进行编程，从而降低了目标代码兼容性的要求。通用、跨平台操作系统如 UNIX 和 Linux 的出现与发展，使得新的计算机系统结构的成本和风险进一步降低。这促使了 20 世纪 80 年代 RISC（精简指令系统计算机）的出现和迅速发展，并使指令级并行（从采用流水线到多指令发射）和 Cache（高速缓存）技术的使用得以推广。Intel 及时跟上了发展的脚步，在其处理器内部将 x86（或 IA32）指令翻译成类 RISC 指令，并大量采用多种 RISC 新技术。而红极一时的 CISC（复杂指令系统计算机）VAX 系统结构则惨遭淘汰，被 RISC 系统结构所取代。

1986 年之后的 16 年，计算机系统性能的提高更以每年 50% 的速度持续增长。进入 21 世纪以来，基于微处理器的计算机占据了统治地位，工作站和个人计算机（PC）成为计算机产品的主流。基于微处理器的服务器取代了基于逻辑电路或门阵列的小型机，由多个微处理器构成的多处理器系统已经取代了大型机，甚至以复杂计算为主的高端的超级计算机也由多个高性能微处理器构成。

例如，2007 年名列世界超级计算机第一的 IBM 蓝色基因/P，其中共有 73 728 颗 PowerPC 450 处理器（主频 850 MHz，单芯 4 核）、共计 294 912 个计算核心，根据 Linpack 基准测试，其最高运算速度达到了 1 petaFLOPS（1 petaFLOPS = 1000 TFLOPS；FLOPS 即 Floating-point Operation Per Second，浮点操作每秒），即每秒 1000 万亿次！其后起之秀即 2008 年上半年名

列第一的 IBM Roadrunner (走鹃) 计算机系统 (部署于美国能源部的 Los Alamos 国家实验室), 就由 IBM PowerXCell 3.2 GHz 和 AMD Opteron DC 1.8 GHz 两种处理器组成, 拥有 122 400 个计算核心, 其最高性能为 1.026 petaFLOPS, 峰值性能为 1.375 78 petaFLOPS。

然而, 自 2002 年开始, 处理器性能的增长降到了每年约 20% 的水平。其原因大致有 3 个:

- 器件集成度的提高带来器件内部热耗的迅速增大。
- 半导体器件集成度提高的余地不多。
- 系统结构上没有多少可以有效发掘的指令级并行, 很难设法降低存储器延迟。

2004 年, Intel 取消了其高性能单一处理器的研究计划, 转而加入 IBM 和 Sun 的阵营, 宣称将通过发展多核处理器来进一步提高计算机性能。这标志着一个历史性的转折, 处理器的性能不再只依赖于指令级并行, 而应更关注线程级并行和数据级并行。一般来说, 指令级并行属于隐式并行, 它主要利用编译器和硬件技术, 可以不需要程序员的干预。而线程级并行和数据级并行是显式并行, 要求程序员编写并行代码来改善系统性能。

1.1.2 计算机的分类

可以从多种角度对计算机进行分类, 例如, 根据处理的对象, 可以划分为数字计算机、模拟计算机及数字/模拟混合计算机。由于模拟计算机几乎绝迹, 这种分法已毫无意义; 也可按照用途将计算机分为通用型计算机和专用型计算机两种, 但是这种分法粒度过粗, 难以概括计算机的基本特征。

1989 年 11 月, 美国电气与电子工程师协会 (IEEE) 的科学巨型机委员会提出把计算机划分为巨型机、小巨型机、大型机、小型机、工作站和个人计算机 6 类。这种分类方法兼顾了计算机的性能、应用和发展趋势, 在当时是非常合理的, 也为世人普遍采用。直到今天, 在大众媒体上, 在多种流行的计算机基础教材中, 包括最新出版的有关计算机基础的书籍、高校教材等仍然普遍采用了这种分类方法或与之相近的方法。

进入 21 世纪以来, 随着计算机技术的发展, 特别是计算机网络、多媒体技术的蓬勃发展及计算机系统结构的极大改进, 本书认为, 19 年前提出的这种分类方法已不能正确反映当前计算机在性能、应用和发展趋势等方面的现状, 原因如下所述。

① 所谓巨型、大、中、小、微型计算机之间的界限逐渐模糊。例如, 现代价格低廉的个人计算机已具有过去大、中型甚至巨型计算机的性能。有人曾根据 SPECint 基准程序测试数据, 做过这样的性能比较: 以 1978 年的小型机 VAX-11/780 为基准 (性能取 1), 则 1984 年的 VAX-11/785 为 1.5, 1990 年的 IBM RS6000/540 为 24, 而现在的处理器 Intel Pentium III 1.0 GHz 达到了 1779 (2000 年), 2002 年的 Intel Pentium 4 (3.0 GHz) 更达到了 4195。

② 现代计算机的硬件处理核心基本相同。无论是高性能的超级计算机还是 PC, 甚至包括排名在世界计算机前 10 强的超级计算机, 它们的处理核心无一例外, 几乎都由相似的处理核心 (过去常称为微处理器) 组成; 只不过在 PC 上多采用低端处理器而已。例如, 2006 年曾名列世界第 3 的 Gray “红色风暴” (Red Storm), 它包含了 26 544 颗 AMD Opteron 2.4 GHz 双核处理器 (部署于美国 Sandia 国家实验室), 浮点运算速度为每秒 101.4 万亿次。2006 年排名第 7 的 Tera-10-NovaScale 5160, 采用了 9968 颗 Intel 安腾 2 (Itanium 2) 1.6 GHz 处理器。

事实上,从 20 世纪 80 年代开始,计算机的设计者逐步转向以微处理器为基础来组成计算机系统。其原因,一是器件集成度的提高,有可能使微处理器包容更复杂的结构。二是许多在巨型机和大、中、小型计算机上用到的技术也下移到低一级或微型计算机上,如 RISC 技术、流水线技术、超标量技术、超流水线技术、超标量超流水线技术、超长指令字技术、Cache 技术乃至线程级并行技术、多核(多处理器)技术等,已普遍用于现代处理器上,使得处理器的性能快速提高。这样,拥有单处理器的价格低廉的 PC 也具有大、中型甚至巨型计算机的性能。三是多处理器系统中同类型处理器应用的批量优势和多线程并行优势又大大降低了机器成本并提升了机器性能。因此,进入 21 世纪以来,基于微处理器的计算机占据了统治地位,工作站和 PC 成为计算机产品的主流。以处理器为核心的服务器取代了基于逻辑电路或门阵列的中、小型机,由多个微处理器构成的多处理器系统已经取代了大型机和巨型机。

③ 出现于 20 世纪 90 年代的计算机网络和多媒体技术,扩大和深化了计算机的应用范围。对于通用计算机而言,无论何种机型,都存在实现网络高带宽和高速处理多媒体信息的要求,应用促使了处理器的融合。同时,由一个规模较大的主机带多个用户终端的分时式多用户系统,要么很快由松散的、相互可频繁进行消息传递的网络计算机结点所替代,要么由基于网络的分布式计算机系统甚至集群来构成紧耦合系统。这种现象削弱并消除了大、中、小型机的功能划分,规模较大的、拥有更多处理器的计算机,更多地成为承担网络中心任务的服务器角色。

也许有人说,巨型、大、中、小、微型计算机之间的划分是会随着计算机的发展而产生动态变化的,但这种分法并没有真正体现出计算机的本质特征。例如,今天的计算机几乎都基于网络,在传统的分类上看不出它们在网络应用上的区别。又如,在规模上,现代多处理器系统在构成处理器阵列上具有一定的灵活性,可依据计算的规模来组织参与计算的处理器数目,它究竟属于何种类型的计算机就很难下定论了。

应该说,将计算机分为巨型机、小巨型机、大型机、小型机、工作站和个人计算机的分法在 20 世纪还是基本正确的。而进入 21 世纪后,基于前述的理由,应该修正现代计算机的分类方式。现在,一般将现代计算机分为 4 种:桌面计算机、服务器、超级计算机和嵌入式计算机。表 1.1 列出了这 4 种计算机的主要特征。

表 1.1 4 种主流计算机的主要特征

	桌面计算机	服 务 器	超级计算机	嵌入式计算机
系统价格	500~5000 美元	5000~5 000 000 美元	1 千万~1 亿美元	10~100 000 美元 (含高端网络路由器)
微处理器价格	50~500 美元	200~10 000 美元	200~10 000 美元	0.01~100 美元
系统设计的关键问题	性价比、图形功能	吞吐量、可用性、可 量测性	吞吐量、浮点计算、 可靠性	价格、能耗、专用性能

1. 桌面计算机

桌面计算机(desktop computer)崛起于 20 世纪 80 年代,分为个人计算机(Personal Computer, PC)和工作站(workstation)两种。PC 就是人们常说的微型计算机,通常是单机系统,主要面向个人应用。目前,较为普遍的是 IBM 或 X86 PC 系列机及其兼容机(多

使用 Intel 和 AMD 的处理器芯片); 以及使用 IBM、Apple 和 Motorola 联合研制的 PowerPC 芯片的机器, 如苹果公司的 Macintosh; 还有 DEC 公司推出的使用自己的 Alpha 芯片的机器。工作站是外围设备配套较完整的单机系统或多机系统, 机型与个人计算机相同, 采用了较为高端的处理器芯片, 主要面向单位团体应用。

计算机网络的出现, 导致了桌面计算机取代了传统的分时系统。这类计算机市场在目前是最大的, 市场需求趋势是提高性价比, 其中多媒体和网络应用性能是用户最为关注的。顺便提一下, 微型计算机是过时的提法, 实际上正确的提法应该是个人计算机或 PC。另外, 这里所说的“桌面”也并非特指“摆在桌面上”的含义, 虽然这一类计算机大部分是摆在桌面上应用的。

2. 服务器

计算机网络的发展, 特别是 WWW (万维网) 的普及, 服务器 (server) 在提供大量数据的集中存储、处理和管理方面的作用日趋明显。服务器基本上取代了传统的大、中型机而成为高可靠性、高数据安全性和对企业进行大规模信息处理的中枢。服务器的主要设计目标就是高效的吞吐量, 即以每分钟处理的事务数或每秒提供的页面数来衡量。服务器的一个关键特性是可靠性, 很难想象一个管理重要网站的服务器系统出现故障会带来怎样的灾难性的损失。另一个是可扩展性, 服务需求或功能需求的增长, 要求服务器在存储容量、输入/输出 (I/O) 带宽、处理能力等方面应能有效升级。

对服务器分类目前尚无明确的提法, 以拥有处理器多少来划分规模是一种方式, 例如小型的服务器可能只有一个高性能处理器, 中型服务器具有数十个处理器, 大型服务器拥有多达数百个处理器。但处理器价格的降低、多核处理器的出现, 使得越来越多的服务器采用规模更大的多处理器结构, 因而人们对这种以处理器数目作为划分标准的时间有效性提出了质疑。

3. 超级计算机

超级计算机 (supercomputer) 拥有最强的并行计算能力, 主要用于科学计算, 在气象、军事、能源、航天、探矿等领域承担大规模、高速度的计算任务。在结构上, 虽然超级计算机和服务器都可能是多处理器系统, 二者并无实质区别; 但是现代超级计算机较多采用集群系统, 更注重浮点运算的性能, 可将其看成是一种专注于科学计算的高性能服务器, 而且价格非常昂贵。

4. 嵌入式计算机

嵌入式计算机 (embedded computer), 即嵌入式系统 (embedded system), 是一种以应用为中心、以微处理器为基础且软、硬件可裁剪的精简计算机系统。它是计算机市场中增长最快的领域, 也是种类繁多、形态多种多样的计算机系统。小到日常生活随处可见的智能电器、PDA (手持数据设备), 大到某些昂贵的工业和医用控制装置, 都有它们的身影。

嵌入式系统的核心部件是嵌入式处理器, 分成 4 类: 嵌入式微控制器 (MicroController Unit, MCU; 俗称单片机), 嵌入式微处理器 (MicroProcessor Unit, MPU), 嵌入式数字信号处理器 (Digital Signal Processor, DSP), 以及嵌入式片上系统 (System on Chip, SoC)。

嵌入式系统对系统的功能、成本、可靠性、实时性、体积、功耗等有严格要求，系统结构和组成按需而定，例如根据用途尽量简化操作系统、最小化存储器、最小化接口和最小化功耗，从而以较低的成本来满足应用的要求。

1.2 计算机系统结构的概念

1.2.1 计算机系统的层次结构

计算机系统由硬件（hardware）和软件（software）组成。人们认识计算机，往往考虑它能起什么作用，用户该怎样使用它才能发挥其功能。事实上，从人们使用机器的角度，可以将系统看成是按功能划分的从第 0 级至第 6 级等 7 层机器级组成的层次结构，如图 1.1 所示。其中，第 0 级和第 1 级为“实际机器”，第 2 级至第 6 级为“虚拟机器”。



图 1.1 计算机系统的层次结构

对“虚拟机器”用户来说，他是通过该层次的语言来了解和使用计算机的。例如，高级语言程序员面向的就是高级语言计算机，只要熟悉和遵循高级语言的使用规定，所编写的程序总是能在这台机器上运行并得到结果，而用不着考虑这台机器是如何通过下一层的汇编语言或机器语言来实现程序的有关功能，更不用了解底层的硬件是如何工作的。在他看来，“机器”就是能存储和执行相应语言程序的算法和数据结构的集合体。

“虚拟机器”的特征是由软件实现机器功能，体现了由机器语言、汇编语言、高级语言到应用程序语言的层次性。一般来说，上层语言语句或命令的功能，是由下层语言的翻译或解释来实现的。翻译（translation）是先用转换程序将高一级的机器上的程序整体变换为低一级机器级上等效的程序，然后再在低一级机器级实现的技术；解释（interpretation）则是在低一级机器级上用它的一串语句或指令来仿真高级机器级上的一条语句或指令的功能，是通过在高一级的机器级语言程序中的每条语句或指令进行逐条解释来实现的技术。越往上层，越接近最终用户而远离实际机器。

“实际机器”的使用者，其主要任务是设计和维护机器。第 1 级机器由微程序（固件）实现。所谓固件（firmware）是一种具有软件功能的硬件，例如将软件固化在 ROM 上。一个微程序往往对应一条机器指令，它根据该指令操作所需要的控制时序，配备一套微指令，编写出微程序，控制信息在寄存器、运算器等之间的传送。第 0 级机器是具体的硬联逻辑电路，实现各个微指令的功能或微操作信号的要求。这些硬件，往往是大规模集成电路及其连接通路。

这里有两点要指出，一是上述层次的划分及其层次间的相互关系并非绝对的，它们之间可能有交叉。例如，高级语言程序经编译后可能跳过汇编语言程序直接以机器语言程序实现；操作系统程序有可能使用高级语言（如 C 语言）而不是汇编语言编写；有些机器指令可以直接使用组合逻辑电路实现，不一定非要由微程序解释实现。

二是软件和硬件在逻辑上是等效的。原理上，软件的功能可用硬件或固件完成，硬件的功能也可用软件模拟完成，只是性能、价格、实现的难易程度不同。例如，一条逻辑运算指令，可直接用组合电路实现；反过来，一次复杂的逻辑运算，即使没有具体对应的逻辑电路，仍可用一条或若干条机器指令来完成。又如，早期的 8 位计算机很多没有乘除法指令，实现乘除法必须用加减移位指令编程来实现。以后的计算机有了乘除法器硬件，就可以设置乘除法指令，直接执行乘除法运算。今天，很多复杂的计算机硬件电路，都可以采用软件方法编程来实现硬件或固件形式，如基于 FPGA（Field Programmable Gate Array）的 SOPC（System On Programmable Chip）技术。也就是说，计算机系统发展到今天，软、硬件的界限变得越来越模糊。具有相同功能的计算机系统，很难说哪些功能只能用软件实现，哪些功能只能用硬件实现。在满足应用的前提下，其软、硬件功能分配的比例，应从系统的应用、效率、速度、成本、资源状况等多个方面综合考虑，从而对软、硬件取舍进行综合平衡。

1.2.2 计算机系统结构的定义

1. 计算机系统结构的含义

计算机系统结构（computer architecture）一词最早由 Amdahl 等人在 1964 年提出。定义为由计算机程序设计者所看到的一个计算机系统的属性，即概念性结构和功能特性，这实际上是计算机系统的外特性。显然，不同机器级的使用者看到的计算机属性是不相同的。在计算机技术中，客观存在的事物或属性，从某个角度来看，它好像不存在，称之为透明性。例如，高级语言程序员所看到的计算机属性主要是高级语言系统，以及操作系统、数据库管理系统、网络软件等用户界面。那些汇编语言程序员所看到的计算机属性，例如通用寄存器、