


# 计算机系统结构

## ——奔腾PC (第二版)

杨厚俊 张公敬 张昆藏 编著

 科学出版社  
[www.sciencep.com](http://www.sciencep.com)

# 计算机系统结构——奔腾 PC

(第二版)

杨厚俊 张公敬 张昆藏 编著

科学出版社

北京

## 内 容 简 介

本书以奔腾 PC 为典型实例,介绍当代计算机系统结构的基本概念、原理及实现技术。本书是 1999 年版本的修订版,书中较详细地分析了以 Pentium 4 处理器为核心的当代奔腾 PC 系统结构,包括芯片组、主板、硬盘机、光盘、显示卡以及如 1394 和 USB 等新型接口总线,并专门介绍了 Intel Itanium (安腾) 64 位处理器和超线程技术等。本书不仅全面分析了以奔腾 PC 为典型实例的单处理机的体系结构,也较详细地介绍了如 CC-NUMA、MPP 和机群等并行计算机的体系结构。并且每章之后均附有习题。对广大 PC 用户,本书是一本理论与实用结合紧密的读物。

本书可作为高等院校计算机相关专业的计算机系统结构课程的教材,也可供从事个人计算机/工作站开发与应用的广大技术人员参考。

### 图书在版编目 (CIP) 数据

计算机系统结构: 奔腾 PC/杨厚俊等编著. —2 版. 北京: 科学出版社, 2004

ISBN 7-03-013214-9

I. 计… II. 杨… III. 计算机体系结构 IV. TP303

中国版本图书馆 CIP 数据核字 (2004) 第 025056 号

策划编辑: 赵卫江/责任编辑: 徐 忱  
责任印制: 吕春珉/封面设计: 王 浩

科学出版社 出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

\*

1999年3月第一版 开本: 787×1092 1/16

2004年5月第二版 印张: 19 1/2

2004年5月第五次印刷 字数: 448 000

印数: 7 501—12 500

定价: 26.00 元

(如有印装质量问题, 我社负责调换〈路通〉)

## 第二版前言

本书第一版自 1999 年出版以来，年年加印，受到读者的欢迎。为了反映计算机技术的发展并满足广大读者的要求，我们对本书的内容进行了修订，将原版中介绍到 Pentium II/III 处理器为止的内容推进到对当代主流处理器 Pentium 4 的介绍；每章还附加习题，以满足教学要求。

我们对第二版的内容版做了如下六个方面的重大改动：

(1) 加入了对 Pentium 4 处理器 (Willamette 版和 Northwood 版) 的介绍，既分析了它们的 Net Burst 体系结构也分析了它们的 20 段的超流水线。

(2) 主存储器的内容加入了对 RDRAM 和 DDR 新型内存芯片及模块的介绍；主板芯片组的内容加入了对 Intel 850 和 845 芯片组的介绍；对显卡及 AGP 接口总线的内容进行了改写，从而较完整地介绍了以 Pentium 4 为核心的当代奔腾 PC 的系统结构。

(3) 新增一章对 64 位处理器的介绍。较详细地剖析了 Intel Itanium (安腾) 处理器的 IA-64 体系结构和 EPIC (显式并行指令计算) 技术。同时介绍了 AMD 的 x86-64 体系结构和新兴的 Intel 超线程技术以及后 Pentium 4 的处理器。

(4) 考虑到计算机组成原理等课程已介绍了浮点数表示和浮点运算器，故删除了原书第 5 章的内容，也删除了原书 7.5 节未来总线 Futurebus<sup>+</sup> 的内容，因为它目前以及未来四五年内将不会出现在奔腾 PC 中，但第二版仍保留了对 Pentium 处理器 (P5) 的完整介绍，一是为便于教学；二是为理解 Pentium 4 奠定基础。

(5) 为符合计算机系统结构课程教学大纲的要求，我们将原书第 10 章并行计算机系统结构扩展成第二版的第 9、10、11 章。用三章的篇幅可更详细地分析并行计算机体系结构，而且，单机系统与并行机系统的篇幅之比 (7:3)，也符合当前对计算机系统结构知识的要求。

(6) 为每章编写了习题。

第二版全书共 11 章，作为教材仍可在 60~70 学时内讲授完毕。除 9.4 节和 10.3 节指定作为阅读材料不讲授之外，教师还可指定一些章节内容作为阅读或讨论课材料。视前期课程 (计算机组成原理、接口技术等) 的进展情况和学生状况，50 学时的讲授时间也是可以做到的。

第二版的材料搜集到 2002 年底，改写工作到 2003 年底完成。这样，2003 年出现的新的技术动向和技术标准仍未在第二版中反映出来，这是难免的遗憾。盼望广大读者继续关爱和支持本书，并将批评、指正和建议反馈给科学出版社。我们将继续本着对读者负责的宗旨，尽心尽力改进本书。

编者  
2003 年 11 月

## 前 言

自第一台电子计算机问世,半个多世纪已经过去了。半导体集成电路的迅速发展和计算机的应用广泛深入到社会生活各个领域,推动着计算机工业高速发展。

这种高速发展反映在计算机系统结构学科领域中,有两点变化是非常引人注目的。一是七八十年代大中型计算机所采用的主要系统结构技术,如虚拟存储、高速缓存、流水线、超标量和超流水、向量运算、I/O 通道等均已纳入个人计算机/工作站中。而且,一些巨型机已是微处理器阵列构成的高度并行机器。因此,进入 90 年代以来,传统的巨、大、中、小、微的计算机分类,只有性能高低、功能多少、规模大小之分,而技术的界线正在消失。二是个人计算机/工作站正在逐渐占据计算机工业的主导地位。从 1992 年起,个人计算机/工作站已占有计算机市场一半以上的份额。主要为个人计算机/工作站提供芯片和操作系统的 Intel 公司和 Microsoft 公司,已分别成为世界上最大的计算机芯片厂商和世界上最大的软件厂商。旺盛的市场需求、雄厚的资本和技术优势,已使个人计算机/工作站率先实现了多媒体等先进技术,今后也会带动计算机系统结构的全面革新与变化。

面对这种新局面,广大的个人计算机/工作站用户需要深入理解他们赖以编程的平台系统结构;广大的高校师生也需要有既能结合个人计算机/工作站实际又能相对完整地、系统地介绍计算机系统结构的新颖教材。本书的编写与出版,就是满足这些需要的一种尝试和努力。

全书共 9 章,作为教材可在 60~70 学时内讲授完毕。下面对各章的基本内容和重点予以导读:

第 1 章:绪论。使读者了解计算机系统结构的任务以及与计算机组成、实现的联系和区别。以 Wintel(Windows + Intel)平台介绍个人计算机软、硬件基本组成,以 RISC 技术特征介绍个人计算机/工作站的系统结构发展趋势。

第 2 章:cache 与主存储器。在建立存储层次结构概念后,重点介绍 cache(高速缓存)的工作原理和实现技术,并介绍了主存的交错组织和先进的 DRAM 芯片。本章亦给出 Pentium PC 的 cache 与主存结构。

第 3 章:虚拟存储器。结合操作系统知识介绍了虚拟存储器的工作原理及实现技术。重点在于 Pentium CPU 的实模式、保护模式和 V86 模式的三种工作模式介绍。第 2,3 章为一个学习单元。

第 4 章:指令流水线。在建立流水线模型之后,重点介绍基本指令流水线和超标量指令流水线,其中的转移预测技术、数据相关性问题及处理策略以及指令发射与完成策略等都极为重要。本章亦给出 Pentium CPU 的 U, V 流水线结构。

第 5 章:浮点运算和运算流水线。本章介绍了 IEEE 754 浮点标准和 Pentium CPU 的浮点运算器。本章介绍的运算流水线与存储流水线、指令流水线等一起,将使流水线知识完整化。本章的非线性流水线调度一节内容,已超出教学大纲范围,标以 \* 号作为选读内容。

第6章: Pentium CPU。及时对前面各章有关 Pentium CPU 的知识予以总结, 介绍它的指令系统、寄存器组织和中断机制, 并给出 Pentium 处理器的结构框图。第4, 5, 6章为一个学习单元。

第7章: 主板总线系统。介绍了 Pentium PC 的总线层次结构之后, 分别介绍了当代总线 PCI、遗留总线 ISA 和未来总线 Futurebus<sup>+</sup>, 三类总线标准及其定时协议和仲裁方式。其中, 应以 PCI 总线为学习重点。

第8章: 外围设备及 I/O 总线。根据对计算机系统结构有重要影响的原则, 本章有选择地介绍了硬、软磁盘、CD-ROM、图形显示器等外围设备及其适配器。其中, 阐述了 AGP(加速图形端口)这一当前热门技术话题。然后介绍了先进的 I/O 总线: 并行的 SCSI 和串行的 P1394。第7, 8章为一个学习单元。

第9章: 新一代 Pentium PC。前面各章有关 Pentium CPU 的介绍是基于普通 Pentium(P54C), 本章介绍了 Pentium MMX(P55C), Pentium Pro, Pentium II, 以及它们采用的多媒体扩展技术和动态执行技术, 最后给出新型的 Pentium II 的 PC 系统结构框图。同第1章一样, 本章也是一个单独的学习单元。

第10章: 并行计算机系统结构。首先以导论一节介绍了并行计算机系统分类、系统互连机构, 以及消息传递机制等基本知识。随后的两节分别以 CC-NUMA 为代表介绍多处理机, 以群集系统为代表介绍多计算机, 并将重点放在可编程性和可扩展性都比较好的, 当前受到普遍重视的 CC-NUMA 多处理机上。本章也是一个单独的学习单元。

作为一名高校教师, 笔者曾在教材改革上做过一些尝试和努力。其中值得提及的是, 1993年编写、1994年由清华大学出版社出版发行的《操作系统原理 DOS篇》一书。该书站在操作系统理论的高度上, 详细分析 MS-DOS 内部结构并实事求是地论述它的特色与不足。结果该书获得很大成功, 4年来发行量达9万册。

《操作系统原理 DOS篇》一书的成功, 说明广大读者迫切需要将计算机科学理论与当代实际应用紧密结合的新书。需求就是动力, 在科学出版社的热情支持下, 笔者萌生了将近年来在计算机系统结构教学与科研方面的成果予以总结提高, 编写本书的愿望。经过一年多的努力, 《计算机系统结构——奔腾 PC》一书与读者见面了。然而本书也有不足之处, 例如, Intel 430, 440 系列核心芯片组在 Pentium PC 系统结构中起着至关重要的作用, 但是由于厂家的技术封锁, 我们查询不到它的内部组成技术资料; 又如, 本打算将 Intel P7 处理器与其他公司的 64 位处理器一起作为第十章内容, 以期深入介绍 RISC 技术并将有关知识完整化, 但 P7 迟至今日仍未推出; 再如, 作为教材, 各章应附有习题, 但一年的熬灯夜战精力实所不支, 也只得暂且作罢。尽管如此, 笔者本着对读者负责的宗旨, 尽心尽力精心编写, 力图使本书具有将相对完整并系统的计算机系统结构理论与 Pentium PC 的实际应用相结合的特色。但愿广大读者能从中获得启迪和帮助。

本书在编写过程中, 西蒙-舒斯特图书公司北京代表处首席代表姜峰先生、北京科海培训中心夏非彼女士都给予了大力支持和帮助, 在此, 一并致以衷心的感谢。

限于作者的经验与水平, 书中错误和不当之处敬请读者批评指正。

张昆藏  
1999年1月

# 目 录

<b>第 1 章 绪论</b> .....	1
1.1 计算机系统结构基本概念 .....	1
1.1.1 计算机系统结构、组成与实现 .....	1
1.1.2 计算机系统结构的发展 .....	2
1.1.3 计算机系统结构的分类 .....	4
1.2 个人计算机和工作站 .....	7
1.2.1 何谓个人计算机/工作站 .....	7
1.2.2 Intel x86 处理器系列 .....	9
1.2.3 PC 总线技术 .....	12
1.3 并行计算机 .....	15
1.3.1 并行计算机系统分类 .....	16
1.3.2 并行性、可扩展性和可编程性 .....	21
习题 1 .....	23
<b>第 2 章 cache 与主存储器</b> .....	24
2.1 cache 工作原理 .....	24
2.2 cache 组织结构 .....	26
2.2.1 映射方式 .....	26
2.2.2 替换策略 .....	31
2.2.3 Look-through 与 Look-aside 结构 .....	33
2.3 cache 一致性与写策略 .....	34
2.3.1 cache 写策略 .....	35
2.3.2 MESI 协议 .....	36
2.4 Pentium 处理器的 cache 技术 .....	38
2.4.1 Pentium L1 cache 界面 .....	39
2.4.2 Pentium L1 cache MESI 协议 .....	41
2.4.3 Pentium L1、L2 cache 的关系 .....	44
2.5 主存储器的组织 .....	46
2.5.1 交错存储器组织 .....	46
2.5.2 猝发存取的 DRAM 芯片 .....	49
2.5.3 先进的 DRAM 芯片及模块 .....	51
习题 2 .....	54
<b>第 3 章 虚拟存储器</b> .....	56
3.1 虚拟存储技术导论 .....	56
3.1.1 请求页式管理 .....	56

3.1.2	虚拟存储器模型	58
3.1.3	段页式管理	59
3.2	地址转换机制	59
3.2.1	直接映像	59
3.2.2	倒置映像	61
3.3	页面替换策略	63
3.4	Pentium 的虚拟存储器支持	65
3.4.1	Pentium 的工作模式	65
3.4.2	保护模式的分段地址转换	67
3.4.3	保护模式的分页地址转换	69
3.4.4	特权级保护	71
3.5	Pentium 寄存器组织	73
3.5.1	基本结构寄存器组	74
3.5.2	系统级寄存器组	76
3.5.3	x87 型浮点部件寄存器组	78
	习题 3	80
<b>第 4 章</b>	<b>处理器执行结构</b>	<b>81</b>
4.1	CISC 与 RISC 结构	81
4.1.1	RISC 的提出	81
4.1.2	RISC 的特征	83
4.1.3	RISC 与 CISC 的竞争	86
4.2	指令流水线结构	87
4.2.1	流水线导论	87
4.2.2	指令流水线机制	90
4.2.3	转移处理技术	92
4.2.4	数据冒险及其处理技术	96
4.3	超标量指令流水线结构	100
4.3.1	超标量流水线与超标流水线	101
4.3.2	超标量流水线的发射策略	102
4.3.3	实例: 88110 处理器的超标量流水线	106
4.4	Pentium 的超标量结构	109
4.4.1	Pentium 处理器核心结构	109
4.4.2	U、V 指令流水线	114
4.4.3	转移目标缓冲器 BTB	117
	习题 4	119
<b>第 5 章</b>	<b>P6 和 Pentium 4 处理器结构和性能</b>	<b>120</b>
5.1	Pentium Pro、Pentium II/III 处理器结构和性能	120
5.1.1	由 Pentium Pro 到 Pentium II/III	120
5.1.2	iCOMP Index 3.0 性能指标	125



5.2	实现动态执行的 Pentium II / III 核心结构	127
5.2.1	动态执行技术概述	127
5.2.2	Pentium II / III 的双独立总线结构	129
5.2.3	Pentium II / III 超标量流水线及其核心结构	130
5.3	单指令流多数据流技术	133
5.3.1	MMX 技术	133
5.3.2	MMX 应用举例	138
5.3.3	SSE 技术	140
5.4	Pentium 4 处理器结构和性能	143
5.4.1	Pentium 4 处理器性能	144
5.4.2	Pentium 4 超标量流水线及其核心结构	145
	习题 5	149
<b>第 6 章</b>	<b>PC 主板总线结构</b>	<b>150</b>
6.1	核心芯片组及主板结构	150
6.1.1	430 芯片组和早期 Pentium PC 的主板结构	150
6.1.2	440 芯片组和 Pentium II / III PC 的主板结构	152
6.1.3	850、845 芯片组和 Pentium 4 PC 的主板结构	156
6.2	处理器总线和内存总线	159
6.2.1	处理器总线	159
6.2.2	内存总线	164
6.3	PCI 总线	165
6.3.1	PCI 总线结构	165
6.3.2	PCI 总线周期操作	166
6.3.3	PCI 总线仲裁	169
6.4	遗留总线 ISA	170
6.4.1	ISA 总线结构	170
6.4.2	Super I/O 芯片和 LPC 总线	174
	习题 6	175
<b>第 7 章</b>	<b>外围设备及 I/O 总线</b>	<b>176</b>
7.1	外存储器	176
7.1.1	硬磁盘机和软磁盘驱动器	177
7.1.2	IDE/ATA 接口标准	180
7.1.3	CD-ROM 驱动器	185
7.2	显示适配器	189
7.2.1	显示标准	190
7.2.2	显示适配器结构	192
7.2.3	显示适配器带宽需求与 AGP	194
7.3	并行 I/O 总线 SCSI	195
7.3.1	SCSI 标准	196

7.3.2	SCSI 总线	198
7.4	串行 I/O 总线 1394 和 USB	202
7.4.1	PC 传统的串、并口	202
7.4.2	IEEE 1394 标准	203
7.4.3	通用串行总线 USB	206
习题 7		208
<b>第 8 章</b>	<b>新一代 64 位处理器</b>	<b>209</b>
8.1	IA-64 体系结构	209
8.1.1	IA-64 基本概念	209
8.1.2	IA-64 通常组织	211
8.1.3	IA-64 指令格式	212
8.2	EPIC 技术	214
8.2.1	断定式执行	214
8.2.2	推测装入	216
8.2.3	高级装入	217
8.3	Itanium 处理器	218
8.3.1	Itanium 处理器结构	218
8.3.2	Itanium 的硬、软件协合	220
8.4	x86-64 体系结构及其他	221
8.4.1	AMD 的 x86-64 体系结构	221
8.4.2	超线程技术	223
习题 8		225
<b>第 9 章</b>	<b>并行计算机的通信结构</b>	<b>226</b>
9.1	ICN 的组成和类型	226
9.1.1	ICN 的基本部件	226
9.1.2	静态互连网络	228
9.1.3	动态互连网络	230
9.2	ICN 的消息传送	232
9.2.1	消息格式和寻径方式	233
9.2.2	寻径算法	234
9.2.3	流量控制	237
9.3	ICN 的通信时延问题	240
9.3.1	ICN 性能指标	241
9.3.2	时延的减少和隐藏技术	243
*9.4	并行编程模型及其通信抽象	244
9.4.1	并行编程模型	244
9.4.2	通信抽象	246
习题 9		248

<b>第 10 章 共享存储器的多处理机</b> .....	250
10.1 UMA 和 NC-NUMA 多处理机 .....	250
10.1.1 Sun Enterprise 10000 多处理机 .....	251
10.1.2 $C_m$ 多处理机 .....	252
10.2 维护 cache 一致性的目录协议 .....	253
10.2.1 目录协议的特点 .....	254
10.2.2 Stanford Dash 多处理机 .....	255
10.2.3 目录协议的类型 .....	259
* 10.3 存储器一致性模型 .....	261
10.3.1 顺序一致性 .....	261
10.3.2 PRAM 一致性和处理器一致性 .....	264
10.3.3 弱一致性和释放一致性 .....	265
10.3.4 小结 .....	267
10.4 CC-NUMA 多处理机 .....	268
10.4.1 SGI Origin 服务器 .....	268
10.4.2 Origin 服务器的目录协议 .....	271
10.4.3 Sequent NUMA-Q 多处理机 .....	274
10.4.4 NUMA-Q 的目录协议 .....	278
10.4.5 COMA 模型的多处理机 .....	281
习题 10 .....	283
<b>第 11 章 消息传递的多计算机</b> .....	285
11.1 大规模并行处理机 MPP .....	285
11.1.1 IBM SP-2 系统 .....	286
11.1.2 Cray T3E 系统 .....	289
11.1.3 Intel/Sandia Option Red 系统 .....	291
11.2 机群系统 .....	292
11.2.1 机群系统的目标、分类和配置 .....	293
11.2.2 服务器机群 .....	295
11.2.3 COW 的通信网络 .....	296
习题 11 .....	299
<b>主要参考文献</b> .....	300

# 第 1 章

---

## 绪 论

鉴于本书是以奔腾 PC 为主体介绍当代计算机系统结构的原理与技术，故本章先介绍计算机系统结构的基本概念，然后介绍 PC 的进展，尤其是 Intel x86 处理器的进展，最后简要介绍并行计算机的基本概念。

### 1.1 计算机系统结构基本概念

自第一台电子计算机问世以来，半个世纪已经过去了。随着科学技术的进步，计算机技术应用领域的扩大与深入，尤其是 VLSI 技术的新进展、算法与软件研究新成果的实现，使计算机系统性能得到大幅度提高，计算机系统结构也发生了重大变革。那么，什么是计算机系统结构？当今它又朝着什么方向发展？本节对此予以简要介绍。

#### 1.1.1 计算机系统结构、组成与实现

计算机系统结构 (computer architecture)，也称为计算机体系结构。1964 年，由 Amdahl 等人首先提出，并为计算机系统结构下了定义，即“程序员所看到的系统的一些属性：概念性的结构和功能上的表现，这些属性既不同于数据流和控制的组织，也不同于逻辑设计和物理实现”。按照制造 IBM 360 系统时的计算机技术水平，他们所指的程序员是使用机器语言或汇编语言的程序设计人员，而不是高级语言程序员。因此，所看到的计算机属性是硬件子系统的概念结构及功能特性，包括指令系统和实现指令系统的硬件，如寄存器定义和组织、存储器的组成和寻址方式、数据类型及表示、机器工作的状态及切换、中断以及输入输出机制等。

计算机组成 (computer organization)，也常称为计算机组织。在计算机系统结构确定了分配给硬件子系统的功能及其概念之后，计算机组成的任务是研究硬件子系统各部分的内部结构和相互联系，以实现机器指令级的各种功能和特性。它包括：数据通道宽度的确定，各种功能部件的相互连接及性能参数的匹配，功能部件的并行性确定，控制机构的设计，缓冲器和排队的使用，可靠性技术的采用等。

计算机实现 (computer implementation)，指的是计算机组成的物理实现。它包括处理机、主存等部件的物理结构，器件的集成度、速度和信号，器件、模块、插件、底板

的划分与连接,专用器件的设计,电源、冷却、装配等技术。

总之,按照上述的划分,计算机系统结构、计算机组成和计算机实现是三个不同的概念。计算机系统结构是指令系统及其执行模型;计算机组成是计算机系统结构的逻辑实现;计算机实现是计算机组成的物理实现。它们各自包含不同的内容和采用不同的技术,但又有紧密的关系。

有两点需要注意:

一是计算机系统结构、组成和实现之间的界限变得越来越模糊了,尤其是严格区分计算机系统结构和组成已不太可能,也没有太大的实际意义。随着 VLSI 技术的进步,新器件的不断涌现,当今计算机系统结构的设计所面临的问题与 Amdahl 所处的时期大不相同,就是与 30 年前也大不相同。例如,30 年前系统配置几十至几百 KB 的内存就很不错了,某些指令系统的设计中甚至有对存储器操作数直接进行加减的指令,不惜牺牲执行速度来珍惜宝贵、有限的内存资源。现在,存储器芯片的集成度大幅度提高而价格急剧下降,内存容量已不是计算机系统结构设计的主要问题了;如何组织存储器以提高存取速度,如何保证 CPU-内存之间的通道不致成为系统性能的瓶颈,是当代计算机系统结构设计必须考虑的问题。现在,一般已将功能模块设计移入计算机系统结构的考察范畴之内。

二是我们介绍了计算机系统结构、组成和实现三者之间的关系,但不要认为计算机系统结构设计就是硬件设计,两者不能混淆。操作系统、编译程序以及高级语言的发展都对计算机系统结构的设计有重要影响。计算机系统结构设计是在功能这一层次上考虑问题,当然要涉及到硬件,但它不是只包括硬件设计。例如,存储器管理功能可以由硬件和软件共同实现,它们之间的分工取决于当前硬件和软件的可用性、性能和价格。在 VLSI 发展的初期,存储器管理功能一般由软件实现;现在,存储器控制芯片已能实现存储器管理算法并自动维护存储器与高速缓存的一致性。因此,计算机系统结构设计的一个主要任务是研究软件、硬件功能分配和对软件、硬件界面的确定。

总之,计算机系统结构完成各功能模块设计并把这些功能模块互连成一个完整的计算机系统,这些功能模块可以是硬件、软件或者两者的混合。它向编译程序、操作系统设计人员提供硬件子系统的概念结构及其功能特性。计算机系统结构的属性和特征与操作系统的属性和特征,一起构成计算机软件运行的平台。

### 1.1.2 计算机系统结构的发展

世界第一台通用电子数字计算机 ENIAC (Electronic Numerical Integrator And Computer), 于 1946 年在美国宾夕法尼亚大学建成。ENIAC 输入和更换程序特别繁冗。对此, ENIAC 课题组的顾问、著名数学家冯·诺依曼提出将程序的指令与指令所操作的数据一起存于存储器的概念。这个著名的存储式程序 (stored-program) 概念, 成为计算机工作的基本机理。这一概念也被图灵大约同时期提出。后来, 冯·诺依曼在普林斯顿主持设计实现这一思想的 IAS 计算机, 但迟至 1952 年才予完成。世界上第一台存储式程序计算机是 1949 年在英国剑桥大学建成的 EDSAC (Electronic Delay Storage Automatic Calculator) 计算机, 它使用 3000 只电子管, 每秒钟能完成 700 次加法运算。1953 年 IBM 公司制造出第一台电子存储式程序的商用计算机。

## 1. 冯·诺依曼机器

50多年来,计算机系统结构已取得重大进展,但大多数机器的计算机系统结构仍没能摆脱冯·诺依曼机器结构的范畴。冯·诺依曼型机器由运算器、控制器、存储器和输入/输出设备组成,如图1.1所示。它的计算机系统结构的基本特点可归纳为以下几点。

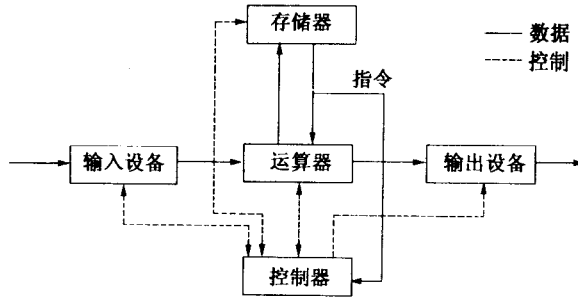


图 1.1 冯·诺依曼机器的结构

第一,采用存储程序方式,程序的指令和数据一起存放到存储器。存储器由线性编址的单元组成,每个单元的位数是固定的。

第二,指令由操作码和地址码组成,操作码指定操作类型,地址码指明操作数地址。操作数的数据类型由操作码确定,操作数本身不具有数据类型标志(如是定点数,还是浮点数等)。

第三,存储器中的指令和数据,从它们本身是区别不了的,它们都是以二进制编码表示的。只不过,指令的地址应由指令计数器给出。换句话说,以指令计数器的值为地址,由存储器读出的内容将被机器看成是指令,被送往控制器去解释和执行。

第四,每执行完一条指令,指令计数器一般自动加“1”,以指示下一顺序指令的地址。虽然执行顺序可以根据运算结果来改变,但解题算法依然是、也只能是顺序型的。

第五,机器以运算器为中心,采用二进制运算,输入输出设备与存储器之间的数据传送都要途经运算器。控制器实施对各部件的集中控制。

总之,冯·诺依曼机器实现了存储式程序这一基本原理;但指令的执行次序受指令计数器(也称为程序计数器)的控制,是一种控制驱动方式的机器。尤其是它的串行执行方式,使得解题算法和编程模型只能是顺序型的;它的更深远影响是使目前绝大多数程序设计语言、编译器、操作系统都是面向顺序式编程模型。

## 2. 计算机系统结构的发展

50年来,电子计算机已经历了四次更新换代,现在正处于第五代。各代的划分主要依据半导体技术水平,并有明显的硬件和软件技术为标志,如表1.1所示。

当前,计算机系统结构正在沿着两个方向发展:

一是改变冯·诺依曼机器的串行执行模式。如在处理器中加入更多的译码部件、更多的执行部件和更多的寄存器而实现的超标量处理器,以做到同时执行多条指令。又如,以多个处理器共享集中式存储器或分布式存储器的多处理机系统,以及大规模并行处理机(MPP)系统,它们在系统中注入强大的并行处理能力并提出了并行编程模型。

表 1.1 计算机的发展

	技术和系统结构	软件和应用
第一代 (1945~1954)	电子管和继电器。单 CPU, 以程序计数器(PC)和累加器顺序完成定点运算	机器语言或汇编语言, 单用户, 用 CPU 程序控制 I/O
第二代 (1955~1964)	晶体管和磁芯存储器。用印刷电路互连。变址寄存器, 浮点运算; 多路存储器, I/O 处理机	有编译程序支持的高级语言, 子程序库, 批处理监控程序
第三代 (1965~1974)	中小规模集成电路 (MSI-SSI)。多层印刷电路。微程序设计, 流水线, 高速缓存, 先行处理机	多道程序设计, 分时操作系统, 多用户应用
第四代 (1975~1990)	大规模集成电路 (LSI/VLSI)。半导体存储器。多处理机, 多计算机, 向量超级计算机	用于并行处理的多处理机操作系统、专用语言和编译器; 并行处理或分布计算的软件工具和环境
第五代 (1991~今)	超大规模集成电路 (ULSI/GSI)。高密度高速处理器和存储器芯片, 可扩展系统结构, 因特网	大规模并行处理, Java 语言, 分布式操作系统, 万维网

尤其是在个人计算机/工作站迅速发展和计算机网络技术成熟和普及的情况下, 近年来成为研究热点的机群系统, 不仅具有良好的并行性, 而且具有良好的可扩展性和可编程序性, 以及良好的市场前景。

二是改变冯·诺依曼机器的控制驱动方式。自 20 世纪 70 年代以来, 提出了数据驱动、需求驱动和模式匹配驱动三种新型驱动方式。数据流计算机是一种数据驱动式系统结构的计算机, 只要指令所需的操作数已经齐备, 就可立即启动执行, 一条指令的运算结果又流向下一条指令, 作为下一条指令的操作数来驱动该条指令的启动执行。程序中各条指令的执行顺序仅由指令间的数据依赖关系决定。需求驱动方式是一个操作仅在要用到其输出结果时才开始启动, 如果此操作的操作数未到齐, 则它去启动能得到各输入数的操作, 需求链一直延伸下去, 直至遇到常数或外部已输入的数据为止, 然后再反方向地去执行运算。归约机就属于需求驱动系统结构的、使用函数式程序设计语言的计算机。在模式匹配驱动方式中, 计算的进行是由谓词模式匹配来驱动的, 而谓词是代表客体之间关系的一种字符串模式, 主要用来求解非数值的符号演算。面向智能的计算机, 如 Lisp 机、Prolog 机、神经网络等, 属于这种模式匹配驱动式系统结构的计算机。

近 30 年来, 前一种发展方向, 即控制驱动方式下并行处理系统结构的计算机, 已取得重大进展。不论是硬件技术还是软件技术都已相当成熟, 并有不少产品走向市场, 它们代表当前计算机系统结构发展的主流。而后一种发展方向, 除数据流计算机已有一些成型机之外, 大多还属于探索、研究阶段, 还有大量的工作待完成。

### 1.1.3 计算机系统结构的分类

曾普遍将计算机系统分为巨、大、中、小、微五类, 这是按规模、性能、速度以至价格的一种大致划分。

大、中型计算机, 国外常用 mainframe 一词表示, 它们大都是通用机, 在计算机工业中占很重要地位, 很多新的系统结构技术都是首先在大中型计算机上被采用。著名的产品, 早期有 IBM 360 系列、370 系列, 后来有 IBM 4300 系列等。

巨型计算机是为气象预报、国防工程、核物理研究等专门设计的具有极高运算速度和很大容量的计算机。其中, 著名的有 Cray-1 计算机, 它的向量运算速度达 8000 万

次/秒,并兼顾了一般的标量运算。1983年研制成功的CrayX-MP机向量运算速度达4亿次/秒。近年来,以微处理器为阵列结构的巨型机(常称为小巨型机)得到了发展,例如古德伊尔公司为美国宇航局研制的巨型计算机系统MPP,由16384个微处理器组成 $128 \times 128$ 方阵。我国自行研制的银河-I、银河-II型机也属于巨型机范畴。

因小型机规模小、结构简单,设计试制周期短,便于及时采用先进技术和工艺,硬、软件成本低等优点,曾在20世纪70年代被普遍采用。著名的有DEC公司的PDP-11系列和VAX-11系列。近年来曾在我国高校、银行、企事业单位中广泛流行的、用于Client/Server计算的AS/400亦属于小型机。

微型机的出现与发展,引发了世界范围的计算机大普及浪潮。1971年以Intel 4004的4位微处理器组成的MCS-4是世界第一台微型机。30多年来,微型机获得惊人的飞跃式发展,从4位、8位、16位到现在的32位机,目前正在向64位计算机发展。32位微型机已采用过去大中型计算机中所采用的技术,故现在的微型机性能已达到20世纪70年代大中型计算机的水平。

可见,这种按巨、大、中、小、微五类来划分计算机系统只能对同时期的计算机大致分类,规模、价格、尤其是性能和速度的指标随时间变化而变化。而且,这种分类法也不能反映计算机的系统结构特征。

1972年,Michael J. Flynn提出按指令流和数据流的多倍性对计算机系统结构进行分类。由于当前的计算机系统结构的主流发展方向是控制驱动方式下的并行处理,故这一分类法获得普遍赞同。Flynn分类法提出如下定义:

指令流(instruction stream):机器执行的指令序列。

数据流(data stream):指令流使用的数据序列,包括输入数据和中间结果。

多倍性(multiplicity):在系统最受限制的部件上,同时处于同一执行阶段的指令或数据的最大数目。

按指令流和数据流的多倍性,Flynn将计算机系统结构分成如下四类:

- 单指令流单数据流(SISD);
- 单指令流多数据流(SIMD);
- 多指令流单数据流(MISD);
- 多指令流多数据流(MIMD)。

图1.2分别表示了它们的基本结构(不包括I/O设备)。下面再对每类予以简要解释。

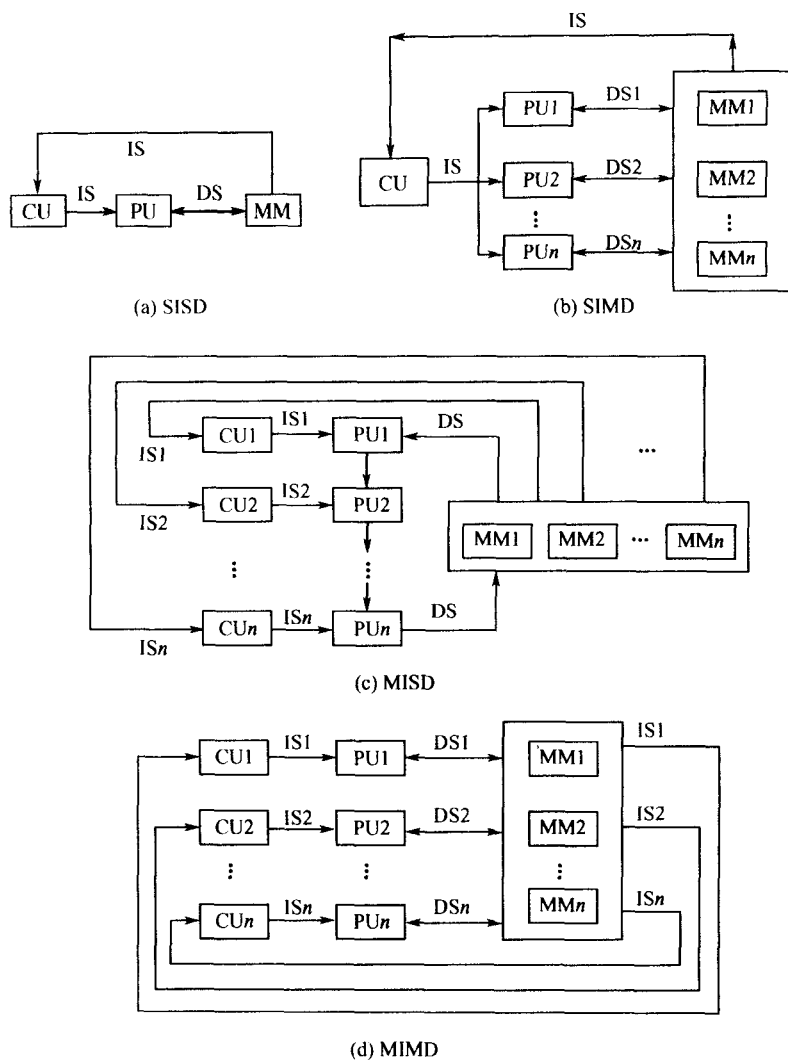
### 1. SISD系统结构

这种结构代表了传统的冯·诺依曼机器,即大多数的单机(单处理器)系统。处理器串行执行指令;或者处理器内采用指令流水线,以时间重叠技术实现了一定程度的指令并行执行;或者处理器是超标量处理器,内有几条指令流水线实现了更大程度上的指令并行执行。但它们都是以单一的指令流由存储器取指令,以单一的数据流由存储器取操作数和将结果写回存储器。

### 2. SIMD系统结构

这种系统结构有单一的控制部件(CU),但有多个处理部件(PU)。以CU由存储器取单一的指令流,一条指令同时作用到各个PU,指挥各PU对来自不同数据流的数据组进行操作。这种系统结构的典型代表是阵列处理机,有的学者认为应将向量处理机





CU= 控制部件; PU= 处理部件; MM= 存储器模块; IS= 指令流; DS= 数据流

图 1.2 Flynn 分类法各类机器结构

划入此类,但也有人持否定态度。值得一提的是,当前的处理器(如 Pentium II / III 和 Pentium 4)多具有多媒体指令功能,这类指令能对打包数据中的多个数据元素同时进行操作。这是一种 SIMD 的变异,也常简称为 SIMD 类的指令。

### 3. MISD 系统结构

这种系统结构中,有几个处理部件 PU,各配有相应的控制部件 CU。各个 PU 接收不同的指令,对来自同一数据流及其派生数据流(例如中间结果)进行操作。这已被证明是不可能的或不实际的,至少是目前没有这类机器。

### 4. MIMD 系统结构

这种系统结构中,也有几个 PU 并各配有相应的 CU,但各 PU 接收不同的指令对