

计算机应用与教学系列

第二版

计算机系统结构

孙强南
孙昱东
编著

科学出版社



计 算 机 系 统 结 构

第二版

孙强南 孙昱东 编著

科 学 出 版 社

2 0 0 0

内 容 简 介

本书系统地介绍了计算机系统结构的基本概念及其组成原理。全书共七章,由局部到全局顺序介绍了存储系统,输入输出系统,指令系统的系统结构,单指令流单数据流的单处理机系统,以及单指令流多数据流和多指令流多数据流的多处理机系统,最后介绍了几种非传统计算机系统结构。

根据计算机系统结构技术的最新发展,本书对第一版进行了较多更新和补充,同时为适应科技人员和管理人员深入学习高性能微型计算机的需求,本书在把巨、大、中、小、微型计算机结合在一起的基础上,尽量以微型机为例,参照高等学校统编教材进行编写。

本书内容全面、重点突出、概念清晰、系统性强,既可以作为大学计算机系本科或自学高考教材,也可供从事计算机硬件工作及从事计算机应用和软件开发的科技人员参考。

图书在版编目(CIP)数据

计算机系统结构/孙强南,孙昱东编著. -2版. -北京:科学出版社,2000.

ISBN 7-03-008349-0

I. 计… II. ①孙…②孙… III. 电子计算机-系统结构
IV. TP303

中国版本图书馆 CIP 数据核字(2000)第 04257 号

科 学 出 版 社 出 版

北京东黄城根北街16号
邮政编码:100717

新 蕾 印 刷 厂 印 刷

新华书店北京发行所发行 各地新华书店经售

*

1992年10月第 一 版 开本:787×1092 1/16

2000年 7月第 二 版 印张:21 1/4

2000年 7月第七次印刷 字数:483 000

印数:18 739—22 238

定价:30.00 元

(如有印装质量问题,我社负责调换(环伟))

第二版前言

本书第一版脱稿于1990年,1992年10月第一次印刷,它是我从事计算机设计和教学工作三十多年的总结.出版后已为一些高等学校和成人教育选作教材,年年重印.

自第一版出版以来,计算机系统结构技术又有了很多新的发展.正如我在第一版前言中一开始所说的那样:“在集成电路技术的发展和计算机应用需要的推动下,计算机系统结构发展迅速,新的系统结构技术不断涌现,并迅速用于高性能的计算机产品”,这种迅速发展的势头有增无减.例如,在本书第一版中用作例子的80386处理器芯片,它的晶体管数是27.5万个,这在当时还算是比较新的.但在成书后不久出现的Pentium处理机,其芯片上的晶体管数即增加到310万个.目前,芯片上的晶体管数已突破千万大关.在此物质基础之上,微处理机中所采用的系统结构技术日新月异.

在处理机系统方面.在写第一版时,超级标量计算机刚刚出现,市场上还没有采用这项技术的微处理机产品,现在,超级标量已成为一切高性能微处理机都采用的常规技术.当时,对RISC机和CISC机还在进行孰优孰劣的争论,现在则事过境迁,两种技术互相渗透,当年的那种纯RISC机和纯CISC机都已成为明日黄花,RISC机的指令系统更加复杂,而CISC机也已发展成了CISC/RISC机.

在并行计算机系统方面.当时,并行处理机与巨型和大型计算机之争也不明朗.现在,即使是最为传统的IBM大型计算机也采用了并行处理机的技术,并已停止使用双极型电路,改用微处理器芯片和多处理机系统结构.20世纪90年代初以来,一些高性能的计算机系统和服务器的基本上都已采用了多处理机系统结构,很多多处理机技术都已进入商品化实用阶段.

在存储系统方面.过去仅在大、中型计算机中使用的高速缓冲存储器(cache),现在不仅在微处理机中普及,而且随着64位微处理机的出现,由微处理机率先对高速缓冲存储器技术作出不少新的发展,例如采用多级cache、虚地址cache等.虚拟存储器也随着虚地址的加长而发展了分层页表、反向页表等技术.

在其他方面,也有很多新的发展.

鉴于以上情况,促使我决定对本书作一次较大的修订,对内容进行更新和补充.由于第一版中所采取的那种由部件到系统的编写顺序经过实践证明容易为读者接受,适合于读者的学习理解过程,因而在第二版中因袭了第一版的体例,仍然由原次序排列的七章组成,这样做也便于已采用本书作为教材的教师们在教学中向修订本过渡.概括起来,第二版对各章的更新和补充主要有以下内容:

1. 在第一章绪论中增加了“微处理机的发展”和“计算机与通信技术结合引起的发展”这两节,藉以反映进入90年代后微处理机已成为计算机的重要部分,以及计算机与通信技术的紧密结合,使计算机系统发生重大变化的现实情况.

2. 长期以来,存储器件速度的提高落后于处理机的发展,存储器同处理器间的速度差距不断扩大,为此在存储器中发展了很多加速工作速度的措施以事弥补.为反映这一情况,在第二章存储系统中补充了“高速DRAM的进展”一节,并加强了高速缓冲存储器的内容,增加了分裂型高速缓冲存储器和二级高速缓冲存储器两小节.同时,为反映微处理

机向 64 位机发展的趋势,还增加了分层页表、反向页表、虚地址 cache 等小节。

3. 在第三章输入输出系统中充实了总线的内容,增加了系统总线的实例,还专门增加了介绍磁盘存储器系统结构的章节,内容包括磁盘高速缓冲存储器和磁盘阵列两部分。在现在的高性能计算机中这两种技术也是常用的。

4. 在第四章指令系统的系统结构中增加了“指令系统对多媒体应用的支持”一节,介绍已在微处理机中广泛采用的多媒体扩展指令集。

5. 对第五章处理机系统有较多的更新和补充,较多地充实了超级标量计算机的内容,补充了超级流水线计算机的实例,还增加了 CISC/RISC 系统结构的新内容。由于流水线技术在微处理机中的作用更加突出,在这一章中还充实了流水线中的相关问题和流水线中的精确断点这两部分的内容。

6. 第六章并行处理机系统变动比较大。在这一章中对 SIMD 和 MIMD 的内容都有了充实,引用了一些较新的商品化的实用计算机系统作为实例。此外,专门增加了有关分布共享存储器多处理机和工作站机群的章节和实例,前者以其良好的可缩放性已成为紧耦合多处理机中的一个重要机种,后者也已被认为是一种新的发展趋势。由于互联网络在多处理机中的地位已从一个普通部件上升为影响整个系统全局性能的一个重要方面,本章中对互联网络的内容也进行了充实,并增加了路由选择和流控制一节。

7. 第七章非传统计算机系统结构的内容没有变化,但删除了智能机(Lisp 机)的内容。

第二版对实例也作了更新和补充。在有的实例中,对一种类型的机器介绍了多个常见产品,这样做是为了向读者提供选择的余地,这些实例不一定全学,教学者可以根据教学条件,自学者可以根据个人需要选学其中一种。有的实例中同时包括了初期的机种和近期的机种,这是因为初期的机种所采用的技术比较朴素直观,便于初学者掌握,而近期的机种在功能上有诸多发展,使同样的技术演变得比较复杂,但它可以提供想多学一些内容或想知道当前机器中实际情况的读者们学习。

第二版中各章的学习重点基本未变,但略有调整。

第一章 绪论。未变。

第二章 存储系统。新增的分层页表、反向页表和虚地址 cache 可以作为选学的内容。

第三章 输入输出系统。调整为使读者建立输入输出系统发展过程中出现的几种输入输出方式的完整概念,以及总线和磁盘存储器的基本结构。重点学习输入输出通道和磁盘阵列。

第四章 指令系统的系统结构。应增加指令系统对多媒体应用的支持。

第五章 处理机系统。应改为流水线和超级标量计算机。本章中处理机的实例较多,叙述亦较详尽,可以根据具体情况选学,不必全学。

第六章 并行处理机系统。重点学习的紧耦合多处理机应是对称多处理机和分布共享存储器多处理机这两部分。本章内容较多,实例亦较多,可以根据具体情况选学。

第七章 非传统计算机系统结构。按情况选学。长期以来,人们想彻底突破传统的冯·诺依曼机模式而研制的计算机虽然还都未能达到产品化的地步,但从中产生的一些新概念,对计算机系统结构的发展有很多新的启示。如有时间学习这一章,对拓宽思路是

十分有益的。

这次的修订工作是由孙强南执笔完成的。修订版中更新和增补的新图达 117 幅之多，这些新图是由孙晟东工程师协助用计算机精心制作的。

在本书的修订过程中，继续得到前辈和朋友们的鼓励和帮助，谨向他们表示深切的谢意。

本书第一版盛销八年，已得到了广大读者的认可。这次以崭新面貌再度奉献给读者，如能再次有所贡献，则荣幸之至。

孙强南

2000 年 1 月 15 日于北京

第一版前言

近年来,在集成电路技术的发展和计算机应用需要的推动下,计算机系统结构发展迅速,新的系统结构技术不断涌现,并迅速用于高性能的计算机产品.同时,随着超大规模集成电路技术的长足进步,一度曾为巨型机、大型机所特有的计算机系统结构技术也正在迅速下移到小型计算机和微型计算机中,成为计算机专业人员必须掌握的基本知识.本书就是在这种新的系统结构技术不断涌现并迅速下移的形势下,为适应各方面的需要而编写的.它参照了统编教材的教学大纲,内容更加充实和新颖.因而,本书不仅适合于作大学计算机系本科生的教材,也可以作为有关专业研究生的参考书.对于从事计算机应用系统开发和软件开发的科技人员,对于想对高档微型计算机了解得较为深入的其他专业的科技人员,它也是一本适用的参考书.

本书在编写中力求反映出计算机系统结构的最新发展.书中增加了简化指令系统计算机、超级标量机、超长指令字计算机、脉动阵列结构等内容,并加强了多处理机部分,以便读者能掌握一些最新计算机产品中的新技术.最后还介绍了一些系统结构的新发展,如数据流机、人工智能机、归约机等,以扩大读者的视野和思路.

本书在编写中还力求做到理论与实际相结合,注意加强基础,注重联系实际,照顾到深入学习高档微型计算机的人员的需要,并把巨、大、中、小、微型计算机的系统结构有机地结合在一起.凡在小型计算机、微型计算机中已采用的技术,尽量以小型机或微型机为例叙述,并尽量以我国应用较多的主流机种为例.

计算机系统结构的内容不仅涉及硬件,而且涉及到很多软件技术.在系统结构研究的问题中,很多是关于硬软件间的关系和硬软件间功能分配的问题,因而它不仅是计算机总体设计人员和硬件设计人员所必须掌握的,也应是高层次的应用系统开发人员和系统软件开发人员所必须掌握的专业基础知识.因此在考虑读者对象方面,本书照顾到了计算机应用和软件专业的需要,把计算机系统结构作为一门连接软件和硬件的课程来对待.

在一些新型的计算机系统中,编译程序同系统结构的关系已变得十分密切.计算机硬件功能的发挥在很大程度上取决于优化的编译程序.因此,在本书中增加了一些在硬软件结合点上的问题,如指令调度、软件流水线等.

本书力求做到内容精练、重点突出、概念清晰.重点问题力求讲透,一般问题也要使读者能建立完整概念.全书可在50~70学时内讲授完毕.

本书的章节布局采取“由里到外”,即由部件到系统的编写顺序.先讲两个相对独立的分系统(存储系统和输入输出系统),然后再讲对计算机系统起全局作用的指令系统的系统结构和处理机系统.对处理机系统,先讲SISD(单指令流单数据流)单处理机系统,再讲SIMD(单指令流多数据流)和MIMD(多指令流多数据流)多处理机系统.最后介绍系统结构的新发展方向.这种由局部到全局的讲授方法容易为读者接受,并使读者在学习中能逐步增加对计算机系统的概括能力和想象能力.

根据这一原则,全书共分七章,其基本内容和学习重点如下:

第一章 绪论.使读者了解计算机系统结构的含义、发展历史、促进发展的因素.重点学习计算机系统结构的发展和了解改革冯·诺依曼计算机的必要性.

第二章 存储系统. 使读者建立存储系统的层次结构、存储管理和存储保护的完整概念. 重点学习虚拟存储器和高速缓冲存储器.

第三章 输入输出系统. 使读者建立输入输出系统发展过程中出现的几种输入输出方式的完整概念. 重点学习输入输出通道.

第四章 指令系统的系统结构 (ISA). 使读者对计算机的指令系统和中断系统的一般情况建立全面的概念. 重点学习指令系统对高级语言和操作系统的支持, 以及中断系统的作用.

第五章 处理机系统. 使读者对各种主要的单处理机机种的系统结构有全面了解, 建立起完整的概念. 重点学习流水线和 RISC 机.

第六章 并行处理机系统. 使读者建立起 SIMD 和 MIMD 计算机的一般概念. 重点学习紧耦合多处理机和作联机事务处理用的容错计算机.

第七章 非传统计算机系统结构. 介绍非冯·诺依曼型的数据流、归约机, 面向科学计算的脉动阵列结构, 以及人工智能计算机.

本书格局和内容由孙强南规划. 第一、二、三、四章由孙强南编写, 第七章由孙昱东编写, 第五、六章由两人合写, 最后由孙强南加工润色.

本书的编写得到了计算机世界报社、中国长城计算机集团公司和华北计算技术研究所的鼎力相助. 在写作中得到中国科学院学部委员、原电子工业部科技委副主任罗沛霖教授的勉励和指教, 罗老还亲自命笔为本书作序. 清华大学房家国教授仔细审阅了全稿, 并提出宝贵意见. 我们还得到许多同事和朋友们的帮助. 在这里谨向他们致以深深的谢意.

我投身于我国计算机事业不觉已进入第 35 个年头, 愿将我长期以来从事计算机设计和教学的一得之见奉献给读者, 如能对我国的计算机事业有所裨益, 则将是我本人莫大的幸事.

孙强南

1992 年 5 月 15 日于北京

目 录

第一章 绪论	1
1.1 计算机系统结构的含义	1
1.2 促进计算机系统结构发展的因素	2
1.2.1 冯·诺依曼计算机	2
1.2.2 改革冯·诺依曼机的必要性	3
1.3 计算机系统结构的发展	5
1.3.1 计算机系统的发展	5
1.3.2 微处理机的发展	7
1.3.3 计算机系统结构的发展	8
1.3.4 计算机与通信技术结合引起的发展	10
1.4 硬件与软件的功能分配	11
第二章 存储系统	14
2.1 存储系统的基本概念	14
2.1.1 存储器在计算机中的地位	14
2.1.2 存储系统的层次结构	14
2.1.3 存储器访问的局部性	16
2.1.4 层次结构中的命中率	17
2.1.5 哈佛系统结构	18
2.2 提高主存储器速度	19
2.2.1 高速 DRAM 的进展	19
2.2.2 并行主存系统	21
2.3 高速缓冲存储器	23
2.3.1 高速缓冲存储器的结构	23
2.3.2 地址映像	25
2.3.3 替换算法	29
2.3.4 写入策略	32
2.3.5 分裂型高速缓冲存储器	33
2.3.6 二级高速缓冲存储器	33
2.3.7 实例:Pentium Pro 的高速缓冲存储器	35
2.4 存储管理	40
2.4.1 主存储器分区管理的基本概念	40
2.4.2 页式管理	42
2.4.3 段式管理	44
2.4.4 段页式管理	45

2.5	虚拟存储器	45
2.5.1	按需调页	45
2.5.2	虚拟存储器的实现	47
2.5.3	实例:IBM 370 系统的虚拟存储器	49
2.5.4	实例:Pentium Pro 的虚拟存储器	52
2.5.5	缺页及其处理	58
2.5.6	虚拟存储器工作过程	59
2.5.7	分层页表	61
2.5.8	反向页表	64
2.6	高速缓冲存储器和虚拟存储器协同工作	65
2.6.1	cache-主存-辅存层次协同工作	65
2.6.2	实地址 cache 和虚地址 cache	67
2.7	存储保护	70
2.7.1	存储保护的必要性	70
2.7.2	工作状态的保护	70
2.7.3	存储区域的保护	70
2.7.4	访问方式的保护	73
2.8	相联存储器	73
2.8.1	信息访问方式	73
2.8.2	相联存储器	74
第三章	输入输出系统	78
3.1	输入输出系统基本概念	78
3.2	程序控制输入输出	80
3.3	中断驱动输入输出	80
3.4	直接存储器访问	81
3.4.1	基本原理	81
3.4.2	直接存储器访问的输入输出工作过程	83
3.4.3	直接存储器访问部件在计算机系统布局	84
3.4.4	DMA 实例	84
3.5	输入输出通道	87
3.5.1	基本原理	87
3.5.2	分类	91
3.5.3	工作过程举例	92
3.6	总线系统	94
3.6.1	基本概念	94
3.6.2	总线的定时	95
3.6.3	总线的判优	98
3.6.4	系统总线实例:PCI 总线	102
3.7	磁盘存储器	106

3.7.1	提高磁盘存储器速度的重要性	106
3.7.2	提高磁盘驱动器速度	106
3.7.3	磁盘高速缓冲存储器	107
3.7.4	磁盘阵列	108
第四章	指令系统的系统结构	117
4.1	引言	117
4.2	指令格式和寻址方式	118
4.2.1	指令格式和编码	118
4.2.2	指令格式举例	119
4.2.3	寻址方式	123
4.3	数据表示	126
4.3.1	浮点数表示	126
4.3.2	带标志符的数据表示	128
4.3.3	数据表示举例	131
4.3.4	字节顺序	132
4.4	指令系统的设计	133
4.4.1	指令系统设计过程	133
4.4.2	对指令系统的要求	133
4.5	指令系统对高级语言的支持	134
4.5.1	冯·诺依曼语义差距	134
4.5.2	加强对高级语言支持能力的途径	135
4.6	指令系统对操作系统的支持	137
4.7	指令系统对多媒体应用的支持	140
4.7.1	媒体处理机产生的背景	140
4.7.2	多媒体扩展指令集	140
4.7.3	多媒体扩展指令集实例:MMX	141
4.8	中断系统	143
4.8.1	中断的分类	144
4.8.2	中断响应的优先级	145
4.8.3	中断的屏蔽和中断处理次序	145
4.8.4	中断处理过程和硬软件功能分担	147
4.8.5	中断的断点	150
4.9	状态控制	150
第五章	处理机系统	153
5.1	计算机系统的分类	153
5.2	流水方式	154
5.2.1	指令流水线	154
5.2.2	运算流水线	157
5.2.3	运算流水线分类	159

5.2.4	流水线中的相关问题	160
5.2.5	流水线中的循环处理	168
5.2.6	流水线中的精确断点	171
5.2.7	先行控制处理机	173
5.3	多功能部件处理机	176
5.4	向量处理机	178
5.4.1	向量的流水处理	178
5.4.2	向量处理机的组成	179
5.4.3	提高运算速度的技术	181
5.4.4	向量化编译程序	185
5.5	附加式数组处理机	185
5.6	简化指令系统计算机	188
5.6.1	产生背景	188
5.6.2	主要特点	190
5.6.3	寄存器窗口	192
5.6.4	RISC 实例:R2000/R3000	194
5.6.5	RISC 实例:PowerPC601	197
5.7	指令级并行	201
5.7.1	指令级并行性	201
5.7.2	一个周期流出多条指令的处理机	203
5.8	超级流水线计算机	203
5.8.1	概述	203
5.8.2	超级流水线计算机实例:R4000	205
5.9	超级标量计算机	208
5.9.1	概述	208
5.9.2	超级标量计算机中的基本技术	209
5.9.3	超级标量计算机实例:PowerPC 620	214
5.9.4	超级标量计算机实例:R10000	218
5.9.5	超级标量计算机实例:Alpha 21264	221
5.9.6	CISC/RISC 系统结构及实例:Pentium Pro	224
5.10	超长指令字计算机	228
5.10.1	概述	228
5.10.2	超长指令字计算机实例:Cydra 5	229
5.11	软件流水线	232
第六章	并行处理机系统	238
6.1	引言	238
6.1.1	高度并行的计算机系统	238
6.1.2	多处理机系统的优点	239
6.2	SIMD 并行处理机	241

6.2.1	概述	241
6.2.2	SIMD 实例:Illiac IV	241
6.2.3	SIMD 实例:CM-2	247
6.3	MIMD 多处理机	250
6.3.1	概述	250
6.3.2	松耦合多处理机系统(多计算机系统)	250
6.3.3	紧耦合多处理机系统(多处理机系统)	252
6.3.4	UMA 和 NUMA 多处理机	254
6.3.5	对称多处理机	254
6.3.6	同构型和异构型多处理机	255
6.4	互连网络	256
6.4.1	互连网络的特征	256
6.4.2	静态拓扑结构互连网络	256
6.4.3	动态拓扑结构互连网络	259
6.4.4	静态互连网络的路由选择和流控制	266
6.5	cache 一致性	269
6.5.1	概述	269
6.5.2	以硬件为基础的方法	270
6.5.3	以软件为基础的方法	273
6.5.4	监视 cache 协议法实例:MESI	274
6.5.5	CC-NUMA 机	277
6.6	大中型计算机的多处理机系统	279
6.6.1	IBM370 系列多处理机系统	279
6.6.2	IBM S/390 G5 多处理机系统	282
6.7	多处理机系统实例	283
6.7.1	Cray T3D 系统	285
6.7.2	Intel Paragon XP/S 系统	287
6.8	容错计算机	288
6.8.1	基本概念	288
6.8.2	松耦合多处理机构成的容错计算机	289
6.8.3	紧耦合多处理机构成的容错计算机	291
6.8.4	容错计算机实例:Tandem 的 NonStop 系统	292
6.8.5	容错计算机实例:Stratus 系统	295
6.9	机群	297
6.9.1	基本概念	297
6.9.2	机群的分类	299
6.9.3	机群的系统结构	300
6.9.4	机群实例:IBM SP2	300
第七章	非传统计算机系统结构	304

7.1	引言	304
7.2	脉动阵列结构	304
7.2.1	基本原理	304
7.2.2	面向特定算法的脉动阵列结构	306
7.2.3	通用的脉动阵列结构	310
7.3	数据流计算机	311
7.3.1	基本原理	311
7.3.2	数据流图与数据流程序	312
7.3.3	数据流语言	315
7.3.4	数据流计算机的分类	317
7.4	归约机	321
7.4.1	函数式语言	322
7.4.2	归约模型	324
7.4.3	归约机	324

第一章 绪 论

1.1 计算机系统结构的含义

“计算机系统结构”这个名词来源于英文 computer architecture,我国也有译成“计算机体系结构”或者“计算机架构”的. architecture 这个字原来用于建筑领域,其意义是“建筑学”、“建筑术”,或者是“建筑物的设计或式样”. 20 世纪 60 年代这个名词被引入计算机领域. 近 40 年来,“计算机系统结构”一词已经得到普遍应用,它研究的内容不但涉及计算机硬件,也涉及计算机软件,已成为连接硬件和软件的一门学科. 但对“计算机系统结构”一词的含义仍是说法颇多^[1~12],并无统一定义.

早在 1964 年, Amdahl 等人在 IBM 期刊上著文介绍 IBM360 系统时就开始采用“系统结构”一词^[13],他们还为它下了个定义,认为计算机系统结构是“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”. 当时的这个定义所包含的内容比较窄,仅仅把系统结构作为从使用计算机机器语言或汇编语言的程序设计人员的角度所看到的计算机系统的属性. 这些属性只包括指令系统和实现指令系统的硬件,像通用寄存器、数据表示、程序状态字、存储器的组成和寻址技术、中断、I/O 机制等等,这些内容对于使用高级语言的人来说是看不到的,是“透明的”.

随着计算机技术的发展,出现了一系列对计算机系统有着重要影响的硬件技术,例如高速缓冲存储器(cache)技术,在大部分情况下,它们是软件设计人员所看不到的,但研究计算机系统结构的人却不能不研究这些硬件技术. 同时,现代计算机中硬件与软件的结合越来越紧密,例如超长指令字(VLIW)计算机和超级标量计算机的并行操作功能的发挥同优化编译程序的效能密切相关. 因而,研究系统结构也就越来越需要把硬件同软件结合起来研究,也就是说系统结构已经向硬件和软件两个方向扩大了它的研究范围. 现在的计算机系统结构的含义,已经比 Amdahl 等人的定义宽得多了.

计算机系统结构除了包含前述的从软件开发人员(指使用机器语言或汇编语言的软件设计人员)角度看到的系统属性外,还要包含那些从计算机设计人员角度看到的系统属性. 计算机设计人员所看到的属性是把软件开发人员所看到的属性加以具体实现的属性,是计算机中的各个部件及它们之间的相互连接关系,也就是部件的逻辑设计和物理实现. 这些属性都是软件开发人员所看不到的. 为了区别从两个不同角度看到的系统属性,人们把从计算机设计人员角度看到的计算机系统属性叫做“物理实现的系统结构(implementation architecture)”,也有称为“微系统结构”的,而把从软件开发人员角度看到的计算机系统属性叫做“系统的系统结构(system architecture)”,或称之为“宏系统结构”. 有的著作中又把前者叫做“计算机组成”,与之对应地把后者称作“计算机系统结构”.

“系统的系统结构”同“物理实现的系统结构”所关心的问题是不同的,前者关心的是怎样合理地进行硬软件功能的分配,为软件设计人员提供更适用的机器;后者关心的是怎样合理地实现分配给硬件的功能和指标,提高性能价格比. 例如,对于乘法,系统的系统结构考虑的是

“是否设乘法指令”，而物理实现的系统结构考虑的则是“是设专门的乘法器还是通过反复使用加法器来实现乘法”。这两者的差别可以 IBM 系列机为例来说明(图 1.1)^[14]。

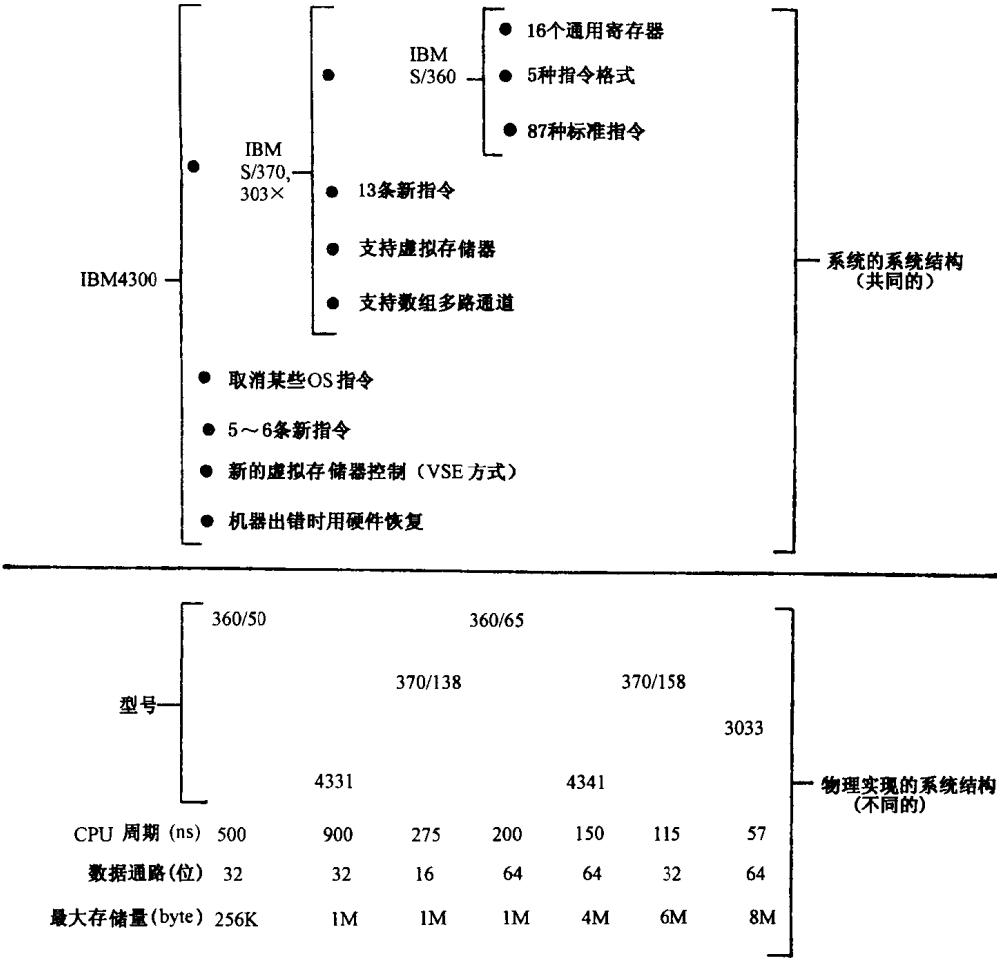


图 1.1 系统的系统结构同物理实现的系统结构的差别

从系统的系统结构来看，图 1.1 中所示的 IBM360,370 或 4300 系列都是相同的，但从物理实现的系统结构来看，各种机器都不同. 在同一个计算机系列中，各个机种由于采用的技术和工艺不同，性能可以相差 100 倍以上。

一种计算机系列，从开始设计最初的机种到最后全部被淘汰，往往要持续 10 年以上，在此期间系列机的系统的系统结构是不能任意改变的，这样才能保持其兼容性，也就是要使同一个系列中的所有机种间能够做到“向上兼容”(为某档机种编制的软件应能不加修改地运行于比它高的档次的机种上)和“向后兼容”(为某个时期投入市场的机种编制的软件应能不加修改地运行于在它之后投入市场的机种上)。

1.2 促进计算机系统结构发展的因素

1.2.1 冯·诺依曼计算机

1945 年冯·诺依曼等人发明了 IAS 计算机^[4]，提出了存储程序计算机方案，其后开

发的计算机基本上都采用这种方式,称为冯·诺依曼计算机或冯·诺依曼机.冯·诺依曼机由运算器、控制器、存储器和 I/O 设备四部分组成,如图 1.2 所示.

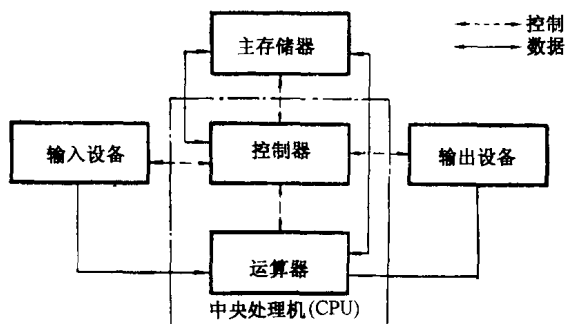


图 1.2 冯·诺依曼机结构

一般认为,冯·诺依曼机的基本特点如下^[10]:

- (1) 采用存储程序方式,程序和数据放在同一个存储器中,两者没有区别,指令同数据一样可以送到运算器进行运算,即由指令组成的程序是可以修改的.
- (2) 存储器是按地址访问的线性编址的一维结构,每个单元的位数是固定的.
- (3) 指令由操作码和地址码组成.操作码指明本指令的操作类型,地址码指明操作数的地址.操作数本身无数据类型的标志,它的数据类型由操作码确定.
- (4) 通过执行指令直接发出控制信号控制计算机的操作.指令在存储器中按其执行顺序存放,由指令计数器指明要执行的指令所在的单元地址.指令计数器只有一个,一般按顺序递增,但执行顺序可按运算结果或当时的外界条件而改变.
- (5) 机器以运算器为中心,I/O 设备与存储器间的数据传送都经过运算器.
- (6) 数据以二进制表示.

1.2.2 改革冯·诺依曼机的必要性

冯·诺依曼机是在 1945 年的技术水平和当时的环境下提出的,专为解决弹道等科学计算用.当时硬件价格昂贵,可靠性低.50 多年来,随着器件、软件技术的发展,以及非数值处理等新应用领域的开拓,对冯·诺依曼机作了很多改革,使计算机系统结构有了很多新发展.促使系统结构发展的主要因素如下^[9]:

(1) 要求增大处理能力.计算机应用领域的扩大对提高处理能力,提出了越来越高的要求.美国科学技术政策办公室的物理、数学和工程科学委员会在 1991 年提出了一个题为“最重要的挑战:高性能计算和通信(HPCC)”^[20]的研究报告.他们选出了一些需要有高性能计算机处理的科学和工程问题,称之为“最重要的挑战”(Grand Challenges),认为这些都是科学和工程方面的基本问题,这些问题的解决对于经济、政治和科学都会有广泛的影响.图 1.3 归纳了这项研究的结果.图中标出了每个阶段的一些有代表性的科学和工程计算问题.在右上角方框中列出的就是所谓“最重要的挑战”.图中还表示了每个阶段解决各类问题时需要计算机达到的运算速度和存储器容量的指标.可以看出,20 世纪 90 年代主要的挑战是要建立具有 10 吉字存储容量和每秒能进行 1 万亿次运算的计算机.

在提高处理能力的同时,对降低价格也提出了迫切的要求.从 20 世纪 50 年代到 80