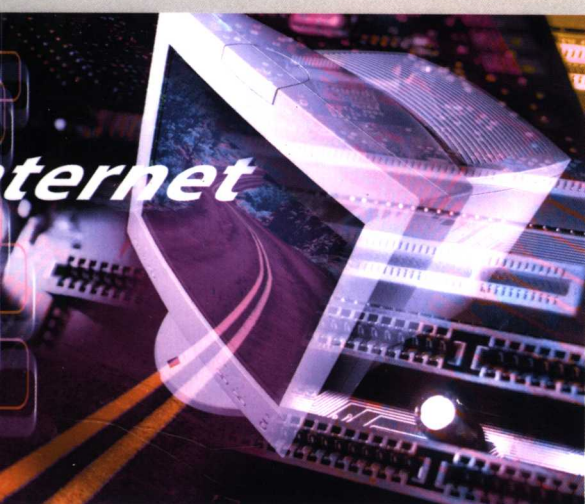


21世纪

高等院校计算机科学与技术规划教材



计算机系统结构

刘超 主编



中国水利水电出版社
www.waterpub.com.cn

21 世纪高等院校计算机科学与技术规划教材

计算机系统结构

刘 超 主编

中国水利水电出版社

内 容 提 要

本书以并行处理技术为主线,介绍计算机系统结构基本概念、基本理论和分析设计的技术方法及其应用实现的逻辑结构模型,讨论指令级高度并行的处理机、向量处理机、阵列处理机、多处理机系统等实现的技术基础、基本结构和性能分析。全书共8章,可分为三部分,第一部分介绍计算机系统结构的基本概念、基本理论和计算机系统结构发展的方向,第二部分介绍流水线技术、指令优化与调度技术、存储组织技术和互连网络技术,第三部分介绍指令级高度并行的处理机、阵列处理机、多处理机系统等。

本书结构新颖、内容实用、逻辑性强、重点突出、语言精炼,可作为高等院校计算机各专业及相关专业本科生的教材,同时也可作为相关方向的研究生或研究人员的参考书。建议本书教学用时为70~80学时。

本书配有免费电子教案,读者可以从中国水利水电出版社网站上下载,网址为:
<http://www.waterpub.com.cn/softdown/>。

图书在版编目(CIP)数据

计算机系统结构 / 刘超主编. —北京:中国水利水电出版社, 2005

(21世纪高等院校计算机科学与技术规划教材)

ISBN 7-5084-3224-X

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

中国版本图书馆CIP数据核字(2005)第098824号

书 名	计算机系统结构
主 编	刘超
出版 发行	中国水利水电出版社(北京市三里河路6号 100044) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 63202266 (总机)、68331835 (营销中心)、82562819 (万水)
经 售	全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	北京蓝空印刷厂
规 格	787mm×1092mm 16开本 17.75印张 429千字
版 次	2005年9月第1版 2005年9月第1次印刷
印 数	0001—5000册
定 价	25.00元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

前 言

“数字逻辑与数字系统”、“计算机组成原理”、“计算机系统结构”和“微机原理及其接口技术”是计算机科学与技术专业本科生硬件一条线的必修课程,其中“计算机系统结构”属于专业课。通过本课程的学习,学生应认识到计算机科学与技术的发展是计算机性能不断提高的过程,是计算机系统结构不断改善的过程,是并行处理技术不断进步的过程;应系统理解计算机系统结构的基本概念、组织结构、分析设计方法,以及目前计算机系统结构发展的新理论、新技术。以此培养学生的组织结构思维能力和概念理论创新能力。

“计算机系统结构”作为计算机科学与技术发展的先导性学科,其主要特点有以下几个方面。一是有许多内容既不完善,也不具体,教材及相应的参考书不多,教材间无论是思想内容,还是侧重点,差异都较大;二是内容本身是以概念、理论和方法为主,与其他硬件课程相比,理论性强、抽象度大;三是内容广泛,涉及面广,它不仅与大量软硬件课程有密切的关系,还与系统学、结构学等有密切的关系,但其关系主要不是知识结构间的关系,而是认识思维间的关系;四是计算机系统结构的发展迅速,应及时补充新的内容。

本书作者长期从事“计算机系统结构”等相关课程的教学和科研工作,在查阅和综合分析相关资料的基础上,编写了本书。本书在系统介绍计算机系统结构基本概念、基本性能和基本理论的基础上,突出并行处理技术是改善计算机系统结构的关键技术。从并行性实现的技术途径——时间重叠、资源重复和资源共享受出发,全面分析流水线技术、指令优化与调度技术、精简指令技术、向量处理技术、存储组织技术和互连网络技术等的概念、理论方法、实现的逻辑结构要求、提高计算机性能的效果及其技术应用的典型的计算机结构。阐述流水线处理机、向量处理机、阵列处理机、多处理机系统和机群系统等实现的技术基础、基本结构和性能分析。指出了目前计算机系统结构存在的问题和未来计算机系统结构发展的基本框架,对数据流计算机的系统结构作了较全面的介绍。

本书的主要特点是:①在章节和内容安排上思路新,将计算机系统结构的设计技术与典型结构尽量分开,突出前者来满足不同学时需要;②强调基本概念、基本理论和基本设计技术的系统性,与许多教材强调结构体系的多样性不同;③注重基本概念与基本理论的一般解释,而不是典型解释或狭意解释;④突出硬件结构性,对于软件实现的内容少讲或不讲;⑤在保持知识体系完整的前提下,尽量避免与“计算机组成原理”、“操作系统”等课程内容的重复。

本书结构新颖、内容实用、逻辑性强、重点突出、语言精炼,可作为高等院校计算机各专业及有关专业本科生的教材,同时也可作为相关方向的研究生或研究人员的参考书。建议本书教学用时为70~80学时。

本书由刘超主编,参与本书部分章节编写及大纲讨论的还有万承兴、衷尔英等。本书在编写与出版过程中,得到了江西师范大学薛锦云、甘登文、谢旭升等教授关心与帮助,在此表示衷心感谢。

由于作者水平有限,加之时间仓促,书中错误在所难免,敬请各位专家、读者批评指正。

编者

2005年7月

目 录

前言

第 1 章 计算机系统结构导论	1
1.1 计算机系统结构的基本概念	1
1.1.1 提高计算机系统性能的硬件因素	1
1.1.2 计算机系统中的有关术语	2
1.1.3 计算机系统结构与组成及其实现	3
1.1.4 计算机系统结构的特性	5
1.2 计算机系统结构的发展及其影响因素	10
1.2.1 计算模型及其驱动方式	10
1.2.2 冯·诺依曼型计算机系统结构及其发展	11
1.2.3 归约型计算机	13
1.2.4 智能型计算机	13
1.2.5 计算机系统结构发展的影响因素	14
1.2.6 计算机系统结构的生命周期	16
1.3 计算机系统结构中的并行性及其发展	16
1.3.1 并行性及其等级的划分	16
1.3.2 提高计算机系统并行性的技术途径	18
1.3.3 多机系统及其耦合度	18
1.3.4 计算机系统结构向并行处理系统的发展	21
1.3.5 计算机系统结构的分类	23
1.4 计算机系统结构的设计	27
1.4.1 计算机系统结构设计的主要方法	27
1.4.2 计算机系统结构设计中软硬件取舍的基本原则	28
1.4.3 软件可移植性及其实现的基本方法	28
1.4.4 并行处理及其实现的技术问题	30
1.4.5 并行计算机及其性能	33
1.5 计算机系统结构的定量分析	33
1.5.1 系统结构设计的定量原理	33
1.5.2 计算机系统结构的评价标准	36
1.5.3 程序执行时间的测定	38
1.6 数据流计算机	39
1.6.1 数据驱动原理	39

1.6.2	数据流计算机的指令结构及其执行过程.....	40
1.6.3	数据流计算机的结构模型.....	41
1.6.4	数据流计算机的优点与存在的问题.....	43
1.6.5	数据流计算机的发展趋势.....	45
习题一		46
第 2 章	数据表示、指令系统与 I/O 系统的优选技术	49
2.1	数据表示的选择	49
2.1.1	数据表示的基本概念	49
2.1.2	引入数据表示的基本原则.....	50
2.1.3	自定义数据表示	51
2.1.4	浮点数据表示	54
2.2	指令集结构格式的优化设计.....	56
2.2.1	指令集结构的分类	56
2.2.2	指令集结构格式的选择	58
2.2.3	操作码的优化表示	60
2.2.4	缩短地址码长度的方法	64
2.3	指令集结构的功能设计.....	65
2.3.1	指令集结构功能的发展方向.....	65
2.3.2	CISC 指令集结构的功能设计.....	65
2.3.3	RISC 计算机指令集结构的功能设计.....	66
2.3.4	精简指令系统计算机	67
2.4	编译技术与指令集结构的设计.....	70
2.4.1	现代编译器的组成结构	70
2.4.2	现代编译技术与指令集结构之间的影响.....	72
2.5	总线的设计	74
2.5.1	总线的类型	74
2.5.2	通用总线的控制方式	74
2.5.3	数据宽度与总线线数	76
2.6	输入输出系统的设计	77
2.6.1	输入输出系统的基本概念.....	77
2.6.2	输入输出设备与主机信息传送的控制方式.....	78
2.6.3	中断系统中的软硬件功能分配.....	79
习题二		80
第 3 章	流水线技术	83
3.1	流水线的基本概念	83
3.1.1	什么是流水线	83
3.1.2	流水线的表示方法	84

3.1.3	流水线的分类	86
3.1.4	流水线的特点	89
3.2	流水线实现的基本结构	90
3.2.1	重叠执行方式的基本结构	90
3.2.2	先行控制方式的基本结构	91
3.2.3	运算操作流水线和宏流水线的基本结构	93
3.3	线性流水线的性能分析	93
3.3.1	吞吐率 (TP, Through put rate)	93
3.3.2	加速比 (S, Speedup ratio)	96
3.3.3	效率 (E, Efficiency)	97
3.3.4	吞吐率、加速比和效率之间的关系	98
3.3.5	流水线最佳段数的选择	98
3.4	流水线的相关及其处理	101
3.4.1	流水线中相关的基本概念	101
3.4.2	数据相关的形成	102
3.4.3	数据相关的处理	103
3.4.4	控制相关的形成	104
3.4.5	条件转移的处理	104
3.4.6	中断转移的处理	108
3.5	非线性流水线的调度技术	108
3.5.1	问题的提出	108
3.5.2	非线性流水线调度的相关术语及其计算方法	110
3.5.3	最小启动循环调度策略的求解及其实现	112
3.5.4	非线性流水线的优化调度方法	112
	习题三	114
第 4 章	存储组织技术	118
4.1	存储系统的层次结构	118
4.1.1	存储系统的基本概念	118
4.1.2	存储系统的层次结构	119
4.1.3	存储系统的基本结构与典型结构	120
4.1.4	存储系统的性能指标	121
4.2	并行存储器与相联存储器	123
4.2.1	什么是并行存储器	123
4.2.2	单体多字并行存储器	123
4.2.3	交叉访问存储器	124
4.2.4	无访问冲突存储器	126
4.2.5	相联存储器	128

4.3	高速缓冲存储器 (Cache)	129
4.3.1	Cache 存储器的工作原理	129
4.3.2	几种地址映像和地址变换的方法	130
4.3.3	Cache 系统的性能指标	136
4.3.4	影响命中率的基本因素	137
4.3.5	Cache 的预取算法	138
4.3.6	Cache 的一致性问题的	138
4.4	虚拟存储器	140
4.4.1	虚拟存储器的工作原理	140
4.4.2	虚拟存储器的管理方式	140
4.4.3	加快地址变换的方法	145
4.5	存储系统的替换算法及其实现	147
4.5.1	问题的提出	147
4.5.2	替换算法及其实现	148
4.5.3	堆栈型替换算法	149
4.6	存储系统的组织与虚拟地址 Cache	153
4.6.1	三级存储系统的组织方式	153
4.6.2	虚拟地址 Cache	155
	习题四	156
第 5 章	互连网络技术	160
5.1	互连网络的基本概念	160
5.1.1	互连网络的功能和特征	160
5.1.2	互连网络的描述工具	161
5.1.3	常用的基本互连函数	161
5.1.4	互连网络结构特性和传输性能参数	165
5.1.5	互连网络的分类	166
5.2	静态互连网络	167
5.2.1	静态互连网络的概念	167
5.2.2	静态互连网络的种类与结构	167
5.2.3	静态互连网络特性	168
5.3	动态互连网络	169
5.3.1	动态互连网络的概念	169
5.3.2	动态互连网络的互连形式	169
5.3.3	几种常用的多级互连网络	173
5.4	互连网络的消息传递机制	181
5.4.1	路由选择的基本概念	181
5.4.2	消息传递的格式与方式	182

5.4.3	死锁和虚拟通道	184
5.4.4	流量控制策略	186
5.4.5	互连网络的通信模式	187
习题五		187
第 6 章	指令级高度并行处理机	190
6.1	指令级高度并行处理机的基本概念	190
6.1.1	标量指令与标量处理机	190
6.1.2	标量指令级并行的实现方法及其衡量指标	190
6.1.3	指令调度及其方法	192
6.1.4	动态调度技术	192
6.1.5	向量处理的基本概念	195
6.2	超标量和超流水线处理机	196
6.2.1	超标量处理机 (Superscalar Processor)	196
6.2.2	超流水线处理机 (Superpipelining Processor)	199
6.2.3	超标量超流水线处理机	202
6.2.4	四种不同类型处理机的性能比较	203
6.3	超长指令字处理机	204
6.3.1	超长指令字处理机 (VLIW: Very Long Instruction Word) 的时空图	204
6.3.2	超长指令字处理机工作的基本原理	204
6.3.3	VLIW 方式的特征	206
6.4	向量处理机	207
6.4.1	向量指令	207
6.4.2	向量处理机的结构	209
6.4.3	向量协处理器	212
6.4.4	向量处理机的性能	213
习题六		213
第 7 章	阵列处理机	215
7.1	阵列处理机的基本概念	215
7.1.1	阵列处理机的操作模型	215
7.1.2	阵列处理机的特点	215
7.1.3	阵列处理机的基本结构	216
7.2	阵列处理机的并行算法	218
7.2.1	阵列处理机的系统结构与并行算法的关系	218
7.2.2	阵列处理机的算法	219
7.3	几种典型的阵列处理机	222
7.3.1	Illiac IV 阵列处理机	222
7.3.2	MasPar MP-1 阵列处理机	224

7.3.3	CM-2 阵列处理机	227
7.4	大规模并行处理机	230
7.4.1	大规模并行处理机的基本概念	230
7.4.2	大规模并行处理机系统软件设计的新概念	230
习题七		231
第 8 章	多处理机系统	233
8.1	多处理机系统的基本概念	233
8.1.1	多处理机系统及其分类	233
8.1.2	一般性多处理机系统的模型	233
8.1.3	多处理机系统的基本结构	235
8.1.4	多处理机系统的特点	236
8.1.5	多处理机系统的操作系统	237
8.2	多处理机的 Cache 一致性	238
8.2.1	多处理机 Cache 不一致性的由来	238
8.2.2	解决多处理机 Cache 不一致的方法	239
8.2.3	总线监听协议法	240
8.2.4	基于目录协议法	242
8.3	多处理机并行性的实现技术	245
8.3.1	并行算法	246
8.3.2	程序结构中的数据相关及其处理	247
8.3.3	并行程序设计语言	249
8.3.4	并行优化编译器	250
8.4	多处理机系统性能	252
8.4.1	性能指标及其计算模型	252
8.4.2	任务粒度对性能的影响	255
8.5	机群系统	259
8.5.1	机群系统的概念	259
8.5.2	机群系统的关键技术	259
8.5.3	机群系统的通信技术	261
8.5.4	并行程序设计语言	262
8.5.5	并行程序调试技术	262
8.5.6	负载均衡技术	263
8.6	典型的多处理机系统	264
8.6.1	CM-5 系统	264
8.6.2	SGI Origin 2000 系列服务器	267
习题八		270
参考文献		272

第 1 章 计算机系统结构导论

本章介绍计算机系统结构的基本概念、相关术语、主要特性及冯·诺依曼型计算机的发展,讨论影响计算机系统结构发展的主要因素和计算机系统并行性发展的技术途径与分类,分析计算机系统结构设计的技术方法与定量分析,阐述新型的计算机系统结构。

1.1 计算机系统结构的基本概念

1.1.1 提高计算机系统性能的硬件因素

1.1.1.1 器件的换代是计算机系统换代的突出标志

自第一台电子计算机问世以来,以器件的发展为标志,计算机经历了电子管、晶体管、小规模集成电路和大规模超大规模集成电路四代。由于器件的迅速发展,使得计算机在体积、重量、速度、可靠性和稳定性等性能方面有了极大的改善,而且价格不断降低。器件技术的突飞猛进,为计算机的发展提供了必不可少的物质基础,它是推动计算机发展的基本因素。

1.1.1.2 计算机系统结构的不断改进是现代计算机系统发展的关键因素

从计算机系统结构来看,计算机系统已经历了五次更新换代。第一代计算机(1945-1954)将电子管和继电器存储器用绝缘导线互连一起。由单个 CPU 构成,CPU 用程序计数器和累加器顺序完成定点运算。采用机器语言或汇编语言,用 CPU 程序控制 I/O。第二代计算机(1955-1964)采用分立式晶体管和铁氧体磁芯存储器,用印刷电路将它们互连起来。采用了变址寄存器、浮点运算、多路存储器和 I/O 处理机。采用有编译程序的高级语言、子程序库和批处理监控程序。第三代计算机(1965-1974)采用小规模或中规模集成电路和多层印刷电路,微程序控制在这一代开始普及。采用了流水线、高速缓存和先行处理技术,采用多道程序设计和分时操作系统。具有代表性的系统有 IBM360/370 系列和 CDC6600/7600 系列等。第四代计算机(1974-1991)采用大规模或超大规模集成电路和半导体存储器。出现了用共享存储器、分布存储器或向量硬件构成的不同结构的并行计算机,开发了用于并行处理的多处理机操作系统、专用语言和编译器,以及用于并行处理或分布处理的软件工具和环境。代表性的系统有 VAX9000、Cray X-MP 和 IBM3090VF 等。第五代计算机(1991-现在)采用 VLSI 工艺,具有更完善的高密度、高速度的处理机和存储器芯片。它最重要的特点是进行大规模并行处理,采用可扩展和容许时延的系统结构。具有代表性的系统有 Fujitsu 的 VPP500、Cray Research 的 MPP、Thinking Machines 公司的 CM-5、Intel 超级计算机系统 Paragon、SGI 的 Origin2000 和 Sun 公司的 10000 服务器等。

由此可以看到,计算机系统结构是计算机换代的又一个标志。研究如何最合理地组织器件,如何最大限度地发挥器件的作用,如何构成综合性能最佳的系统,是提高计算机性能的重要途径。恩斯洛(P.H.Enslow)曾经比较了 1965 年至 1975 年间,器件的更新使器件延迟时间降低至原来的十分之一,但计算机指令执行时间却降低至原来的百分之一。由此可见,在这 10 年

中, 计算机性能提高的幅度比器件性能提高的幅度大得多。

1.1.2 计算机系统中的有关术语

1.1.2.1 什么是计算机系统 (computer system)

系统一词本身来自希腊文, 是由部分组成整体的意思。从技术角度看, 系统 (system) 是指为完成特定任务而由相关部件或要素组成的有机整体。而人们经常使用的计算机系统这个术语, 不难理解又容易混淆, 它的含义一般有 3 种指向。

一是指狭义的计算机系统。认为计算机系统是由运算器、控制器、存储器、输入设备和输出设备五个部件组成。其中运算器和控制器合称为中央处理器 (CPU), 存储器又分为内存 (又称主存 memory) 和外存 (又称辅存 storage) 两种, 所以这种说法又可简化为 I/O—CPU—M/S 模式。

二是指本义的计算机系统。认为计算机系统是由硬件 (hardware) 和软件 (software) 两部分组成。由于技术飞速进步, 软件和硬件在逻辑功能上是等效的, 某些操作既可以用软件实现, 也可以用硬件实现。软件和硬件之间没有固定不变的界面, 主要受实际的需要和性能价格比所支配。具有相同功能计算机系统的软件与硬件之间的组合分配可能相差很大。这就是当今软硬结合的现实, 这种说法可以认为是 S—H 模式。

三是指广义的计算机系统。认为计算机系统是由人员 (people)、数据 (data)、设备 (equipment)、程序 (program) 和规程 (procedure) 五部分组成, 只有把它们有机地结合在一起才能完成各种任务, 这种说法又称为广义模式。

1.1.2.2 广义的计算机语言

计算机系统的根本功能在于对信息进行处理, 而一个信息的处理过程是指令序列 (程序) 的执行过程, 指令序列可用控制流程的概念来描述。指令序列中的指令仍然是一个信息的处理过程, 即任何一个信息的处理过程都是对某一层面而言的。例如, 软件用户信息的处理过程与软件开发者信息的处理过程是不同层面的。因此, 把描述控制流程的、有一定规则的字符集合称为计算机语言。而控制流程的实现有 3 种方法:

- (1) 全硬件的方法, 即用硬件组合逻辑线路实现。
- (2) 硬件和软件相结合的方法, 即部分流程由微程序实现, 而另一部分由硬件逻辑实现。
- (3) 全软件的方法, 即按流程算法编制程序实现。

可见, 计算机语言并不专属软件范畴, 它可以分属于计算机系统的各个层面, 而且有不同的作用, 这也就是计算机语言的广义性。

1.1.2.3 虚拟计算机

计算机语言描述的信息处理过程都必须编译或解释成具体的操作控制信号才能实现流程控制, 所以编译或解释程序是计算机系统不可分割的一部分。从这一角度来讲, 计算机只对某一层次的观察者存在, 通过该层次的广义语言来体现它的功能, 并为广义语言提供翻译手段, 作用于信息处理和控制对象上, 并从对象上获得必要的状态信息, 这就是虚拟计算机。简言之, 由软件实现的机器即为虚拟计算机。显然, 它是基于广义的计算机语言的观点而形成的一个概念, 其组成如图 1-1 所示。

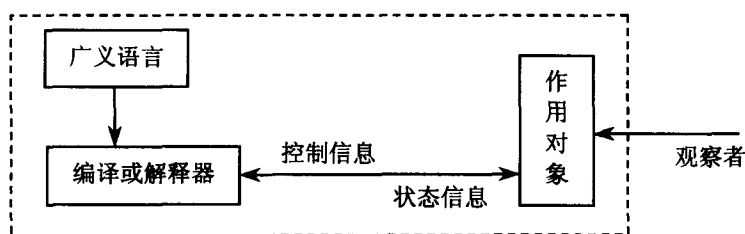


图 1-1 虚拟计算机的组成

1.1.2.4 透明性 (Transparency)

从某一层次的观察者来看，只是通过该层次的语言来了解和使用机器，不必关心其内层的那些机器是如何工作和如何实现各功能的，这一概念称为透明性 (Transparency)。即所谓透明性是指在计算机技术中，一种本来存在的事物或特性，但从某种角度看似乎不存在。它是基于虚拟计算机的观点而形成的一个概念。

1.1.3 计算机系统结构与组成及其实现

1.1.3.1 计算机系统结构

1. 什么是计算机系统结构

“计算机系统结构”这个词来源于英文 Computer architecture，也可翻译为“计算机体系结构”。Architecture 这个词原来用于建筑领域，其意义是“建筑学”、“建筑物的设计或式样”，是指一个系统的外貌。20 世纪 60 年代这个名词被引入计算机领域，20 世纪 70 年代开始被广泛采用，研究内容不但涉及计算机硬件，也涉及计算机软件，已成为一门学科。但是，由于计算机软硬件界面在动态地变化，至今有各种各样的理解，很难有一个通用的定义。

计算机系统结构这个词是 G.M.Amdahl 等人在 1964 年提出的。他们提出计算机系统结构是程序员看到的计算机的属性，即程序员为编写出能在机器上正确运行的程序所必须了解的机器概念性结构和功能特性。然而从计算机系统的层次结构来看，不同级程序员所看到的计算机属性显然不同，实际上，Amdahl 等人提出的程序员是指机器语言程序员和编译程序设计者，计算机属性是硬件系统的概念结构及其功能特性，是计算机系统的外特性。因此完整的计算机系统结构的定义是：机器语言程序员所必须了解的机器概念性结构和功能特性。

2. 计算机系统结构研究的问题及其内容

计算机系统结构作为一门学科，主要是研究软件与硬件功能分配和对软件与硬件界面的确定，即哪些功能由软件完成，哪些功能由硬件完成，以及如何最佳最合理地实现分配给硬件的功能。当然，在层次结构的各个级上都有它的系统结构，而且低层级的概念性结构和功能特性对高层级语言程序员来说是透明的。具体说来，计算机系统结构所包括的内容一般有：

- (1) 数据表示，即硬件能直接识别和处理的数据类型和格式等。
- (2) 寻址方式，包括最小寻址单位，寻址方式的种类、表示和地址计算等。
- (3) 寄存器组织，包括操作数寄存器、变址寄存器、控制寄存器及某些专用寄存器的定义、数量和使用约定。
- (4) 指令系统，包括机器指令的操作类型和格式、指令间的排序方式和控制执行机构等。
- (5) 存储系统，包括最小编址单位、编址方式、存储容量、最大可编址空间等。

(6) 中断机构, 包括中断的类型、中断分级、中断处理程序的功能和入口地址等。

(7) 机器工作状态(如管态和目态等)的定义和切换。

(8) 机器级的 I/O 结构, 包括 I/O 系统的联结方式、设备的访问方式、处理机/存储器与 I/O 设备间数据传送的方式和格式、传送的数据量、I/O 操作的结束与出错标志等。

(9) 信息保护, 包括信息保护方式和硬件对信息保护的支持等。

这些都是程序员为了使其编写的程序能在机器上正确运行所需要了解和遵循的计算机属性。当然不包括基本的数据流和控制流的逻辑设计和物理实现等。

另外, Amdahl 等人对计算机系统结构定义的主要内容是指令系统及其执行模型, 认为只要指令系统兼容就能保证程序正确运行。但由于程序的执行要依赖于程序库、操作系统和其他在 Amdahl 等人的系统结构定义中没有涉及的因素, 这要求操作系统接口等其他层次的标准化。同时, 由于 VLSI 的迅速发展及其成本急剧下降, 有些系列机推出有新指令的机器, 例如 24 位地址的 IBM360 和 370 系统发展为 31 位地址的 370xA 系统, 16 位地址的 PDP-11 发展为 32 位地址的 VAX 系列。因此, 计算机系统结构定义的内容不是固定不变的。

1.1.3.2 计算机组成

1. 什么是计算机组成

在确定计算机系统结构以及分配给硬件子系统的功能及其概念结构之后, 则需要计算机组成来研究各组成部分的内部构造和相互联系, 以实现机器指令级的各种功能和特性。这种相互联系包括各功能部件的配置、相互连接和相互作用, 各功能部件的性能参数间的相互匹配, 因而相应地就有许多计算机组织方法。例如, 为了使存储器的容量大、速度快, 人们研究出层次存储系统和虚拟存储技术; 在层次存储系统中, 又有高速缓存、多模块交错工作、多寄存器组和堆栈等技术。为了使输入输出设备与处理机间的信息流量达到平衡, 人们研究出通道和外围处理机等方式。为了提高处理机速度, 人们研究出先行控制、流水线和多执行部件等方式。在各功能部件的内部结构研究方面, 产生了许多组合逻辑和时序逻辑的高效设计方法和结构, 例如在运算器方面, 出现了多种自动调度算法和结构等。

由此可见, 计算机组成是计算机系统结构的逻辑实现, 包括计算机内部的数据流和控制流的组成以及逻辑设计等。

2. 计算机逻辑设计的目的及其内容

计算机逻辑设计是按所希望达到的性能价格比, 最佳最合理地把各种设备和部件组成计算机, 以实现所确定的计算机系统结构。一般计算机组成设计包括数据通路宽度的确定、各种操作对功能部件的共享程度的确定、专用功能部件的确定、功能部件的并行性确定、缓冲器和排队的确 定、控制机构的设计及可靠性技术的确定等。对传统机器程序员来说, 计算机组成的设计内容一般是透明的。

1.1.3.3 计算机实现

计算机实现是指计算机组成的物理实现。它包括处理机和主存等部件的物理结构, 器件的集成度和速度, 信号传输, 器件、模块、插件和底板的划分与连接, 专用器件的设计, 电源、冷却和装配等技术以及有关的制造技术和工艺等。

1.1.3.4 计算机系统结构、组成与实现间的关系

1. 计算机系统结构、计算机组成和实现是 3 个不同的概念

系统结构是计算机系统软件与硬件的界面; 计算机组成是计算机系统结构逻辑实现; 计算

机实现是计算机组成的物理实现, 它们各自的目的、任务和包括的内容不同。

2. 计算机系统结构、计算机组成和实现相互影响

系统结构相同的计算机, 可以因速度不同而采用不同的组成。例如, 指令间既可以顺序执行, 也可以重叠执行来提高速度。同样, 同一种计算机组成可有多种不同的实现。例如, 主存器件可以用双极型的, 也可以用 MOS 型的。反之, 计算机组成也会影响系统结构。例如, 通过改变控制存储器中的微程序, 可以改变系统的机器指令和结构。计算机实现的技术方法不同, 计算机组成也会有所改变。例如, 机器字长为 32 位, 采用 16 位或 32 位的存储器, 组成逻辑就不同。

3. 计算机系统结构、计算机组成和实现关系密切

一是计算机系统结构、组成和实现包含的具体内容是随不同机器而变化的。有些计算机系统是作为系统结构的内容, 其他计算机系统可能是作为组成和实现的内容。开始是作为组成和实现提出来的设计思想, 到后来就可能被引入系统结构中。例如, 高速缓冲存储器一般是作为组成提出来的, 其中存放的信息全部由硬件自动管理, 对程序员来说是透明的。然而, 有的机器为了提高其使用效率, 设置了高速缓冲存储器的管理指令, 使程序员能参与高速缓冲存储器的管理。这样, 高速缓冲存储器又成为系统结构的一部分, 对程序员来说是不透明的。

二是计算机系统结构、组成和实现密切相关, 难以明确区分。有人把计算机组成和计算机实现统称为计算机实现, 即包括计算机系统的逻辑实现和物理实现。计算机系统结构设计的任务是进行软硬件功能分配, 确定传统机器级的软硬件界面。但作为“计算机系统结构”这门学科来讲, 实际上包括了系统结构和组成两个方面的内容。因此, 计算机系统结构是研究软硬件的功能分配以及如何最佳最合理地实现分配给硬件的功能。

随着新器件的出现, 当今计算机设计者面临的问题与以前大不相同, 所以我们应当把计算机系统结构定义得更宽一些, 除了 Amdahl 等人定义的内容外, 还应包括功能模块的设计等。

1.1.4 计算机系统结构的特性

从不同角度看来, 计算机系统结构具有许多特性。但在功能上它就有以下几方面的特性, 其中层次性是其最重要的特性。

1.1.4.1 计算机系统结构的层次性

计算机系统是硬件和软件的统一体, 就直接反映了结构上的层次性。如果依据虚拟计算机的观点来看, 不同层次的观察者看到的计算机系统无论功能还是结构都不相同, 也就有如图 1-2 所示的计算机系统的层次结构。

从图 1-2 中可以看出, M0 级为硬联逻辑, 是用来实现微指令操作的控制时序的, 只需要很少的逻辑线路, 是所有逻辑线路的内核。M1 级为微程序控制, 是对机器指令进行译码的, 对应一个微指令序列, 给出微操作信号。M0 级和 M1 级实现了中央处理机功能。

M2 级为机器语言机器, 是以指令系统面向用户, 机器语言程序员编写的程序由 M1 级的微程序进行解释。M3 级为操作系统机器, 是为提高计算机系统功能和资源利用率而为用户提供的—个操作环境。该环境下有两类指令, 一类是机器语言指令, 仍由微程序解释; 另一类是操作系统指令, 由操作系统进行解释。

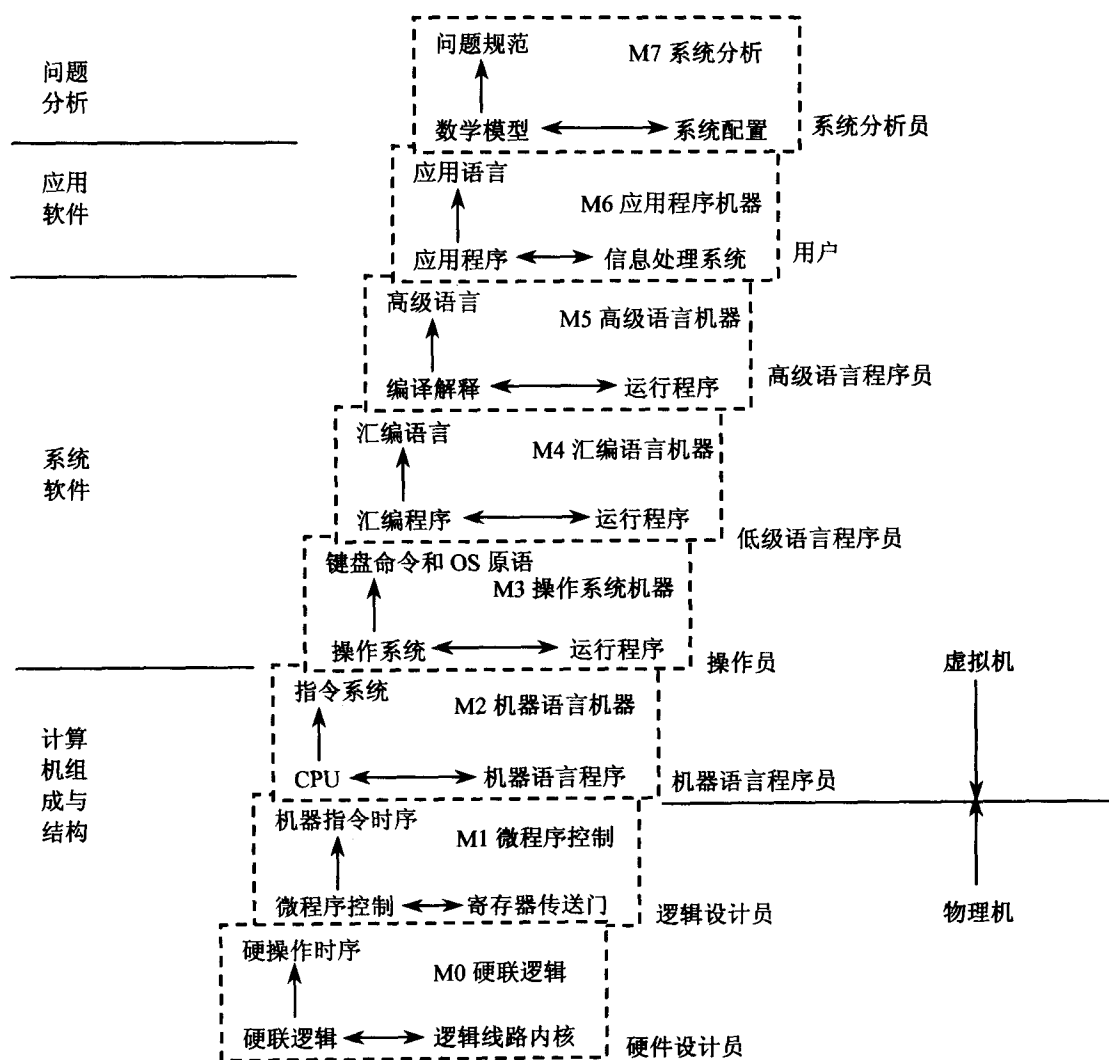


图 1-2 计算机系统的层次结构图

M4 级为汇编语言机器，是以汇编语言面向用户，汇编语言程序员编写的程序通过汇编程序翻译成 M3 级或 M2 级的语言，然后由相应层次的机器进行解释。M5 级为高级语言机器，是以高级语言面向用户，高级语言程序员编写的程序通过编译程序翻译成 M4 或 M3 级的语言，然后由相应层次的机器进行解释。M4 级和 M5 级合称为语言类计算机，供程序员使用，基本上脱离物理计算机。

M6 级为应用程序机器，是以应用语言面向用户，是用来供非计算机专业人员使用计算机的，用户只要简单操作就能进入 M6 级的信息处理系统。M7 级为应用系统分析与设计，建立数学模型和算法，确定系统的配置。M6 级和 M7 级完全脱离了物理计算机。

另外，从计算机系统结构领域来看，M0~M2 级是计算机组成与结构的范围，M3~M5 级是系统软件的范围，M6 与 M7 级是应用软件的范围。当然层次间存在某些交叉，例如，M2 级涉及汇编语言程序设计，M3 级处于硬件向软件过渡，此外，在特殊的计算机系统中，有些

级别可能不存在。

把计算机系统结构按功能划分为多个层次,有利于正确理解计算机系统的工作,明确软件、硬件和固件在计算机系统中的地位 and 作用,有利于理解各种语言的实现及其实质,有利于探索新的虚拟机实现方法,设计新的计算机系统。

1.1.4.2 计算机系统结构的等级性

1. 计算机系统结构等级的划分及其发展

通常把计算机系统结构分为巨型、大型、中型、小型和微型等若干等级。但是,随着技术进步,各等级的计算机性能指标不断提高,以致 30 年前的一台大型机的性能甚至比不上当今一台微型计算机的性能。因此,如果按性能指标来划分计算机等级,那么,一台计算机的等级将随时间而下移。显然,现在把 30 年前生产的大型机称为微型机是不合适的。各级机器的性能、价格随时间变化的趋势大致可以用图 1-3 示意说明,其中虚线称为等性能线。由图可见,各级机器所具备的性能随时间是在动态地下移的,但价格却在相当长一段时间内基本不变,因此,有人就主张用价格来划分巨型机到微型机的不同等级。

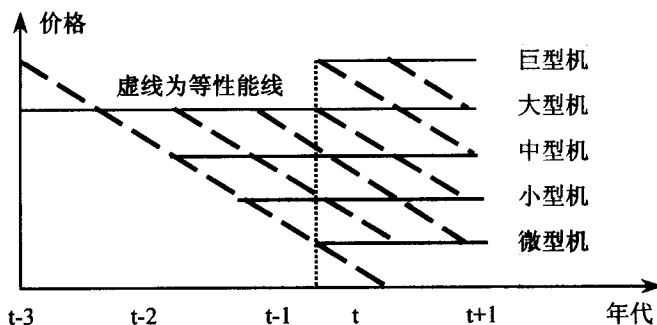


图 1-3 计算机系统结构等级与性能和价格的关系

各级计算机的性能随时间下移,实质上就是在低档(级)机上引用甚至照搬高档(级)机的系统结构和组成。这种低档机承袭高档机系统结构的情况正符合小型机和微型机的设计原则,即充分发挥器件技术的进步,以尽可能低的价格实现高档机已有的那些结构和组成,而不花费很大力量去研究和采用新的系统结构和组成技术。这有利于计算机工业的发展,有利于普及应用计算机。目前,系统结构和组成下移的速度越来越快,例如,超高速缓冲存储器和虚拟存储器从大型机下移到小型机所花时间不到 6 年,巨型阵列机问世不到 7 年小型机上就有了可扩充的高速阵列处理部件。

2. 计算机系统结构的等级设计思想

按等级来设计计算机系统结构应遵循以下三个基本思想。一是在本等级范围内以合理的价格获得尽可能好的性能,逐渐向高档机发展,称为最佳性能价格比设计。二是只求保持一定的使用性能而争取最低价格,称为最低价格设计,其结果往往是低档向下分化出新的计算机等级。三是以获取最高性能为主要目标而不惜增加价格,称为最高性能设计,其结果是产生当时最高等级的计算机。

第一类设计主要针对大中型计算机用户的需要,设计生产出性能价格比更好的中型计算机和超小型计算机。第二类设计以普及应用计算机为目标,设计生产数量众多的微型和小型计算