

计算机科学与技术系列教材

计算机系统结构

高辉 张玉萍 编著



全国优秀出版社
武汉大学出版社

计算机科学与技术系列教材

计算机系统结构

高辉 张玉萍 编著

武汉大学出版社

图书在版编目(CIP)数据

计算机系统结构/高辉,张玉萍编著. —武汉:武汉大学出版社,
2004.8

(计算机科学与技术系列教材)

ISBN 7-307-04246-0

I. 计… II. ①高… ②张… 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(2004)第 047914 号

责任编辑:黄金文 责任校对:程小宜 版式设计:支 笛

出版发行:武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件:wdp4@whu.edu.cn 网址:www.wdp.whu.edu.cn)

印刷:湖北省通山县印刷厂

开本:787×980 1/16 印张:27 字数:550 千字

版次:2004 年 8 月第 1 版 2004 年 8 月第 1 次印刷

ISBN 7-307-04246-0/TP·152 定价:39.00 元

版权所有,不得翻印;凡购我社的图书,如有缺页、倒页、脱页等质量问题,请与当地图书销售部门联系调换。

计算机科学与技术系列教材

编 委 会

- 顾 问：陈火旺（中国工程院院士，国防科技大学教授）
刘经南（中国工程院院士，武汉大学校长）
- 主 任：何炎祥（中国计算机学会常务理事，武汉大学教授）
- 副 主 任：张焕国（中国密码学会理事，武汉大学教授）
江建勤（武汉大学出版社社长，教授）
- 委 员：王振宇（中船七〇九研究所教授）
卢正鼎（华中科技大学教授）
肖德宝（华中师范大学教授）
熊前兴（武汉理工大学教授）
陈莘萌（武汉大学教授）
周洞汝（武汉大学教授）
贾小华（香港城市大学教授，“长江学者计划”特聘教授）
孟 波（武汉大学教授）
李元香（软件工程国家重点实验室教授）
胡瑞敏（武汉大学教授）
黄竞伟（武汉大学教授）
苏光奎（武汉大学教授）
毋国庆（武汉大学教授）
陈世鸿（国家多媒体软件工程技术研究中心教授）
郭学理（武汉大学教授）
吴产乐（国家多媒体软件工程技术研究中心教授）
曹加恒（武汉大学教授）
黄传河（武汉大学教授）
梁意文（武汉大学教授）
章登义（武汉大学教授）
彭智勇（软件工程国家重点实验室教授）
- 秘 书：黄金文



前言

本书作者已经多年且仍然正在从事计算机专业之“计算机系统结构”课程教学工作，这本教材是在多年教学经验的基础上，结合教学科研成果以及多种计算机课程参考书综合编著而成的。

当今的计算机系统是一个硬件和软件相结合的一个复杂系统。要全面地、正确地了解和掌握计算机系统，必须了解和掌握计算机软、硬件交界面及其上、下功能分配。本课程是在学习完《计算机组成原理》、《高级语言程序设计》、《数字逻辑》、《计算机操作系统》等专业基础课之后，再进一步学习计算机专业知识而设置的。为了使读者尽快掌握这门知识，本教材在编写上具有如下两个显著特点：一是深入浅出地介绍计算机系统的基本概念、基本原理、基本结构和基本分析方法。语言表达尽量通俗易懂，概念明确，关键之处以例题说明，达到提高学习效率的目的；二是结合当今流行的微型计算机体系结构的发展。以前只能在大、中型计算机上实现的关键技术，目前已经部分“下移”至微型计算机上。针对这一特点，教材中已在各章后部介绍了 Pentium 微处理器的体系结构的知识与特点，使读者能够有“理论联系实际”之新感受。

本教材共分为 7 章。第一章介绍计算机系统的基本概念。包括主计算机系统结构的定义，计算机系统的分类、设计方法、性能评价以及 Pentium 微处理器的体系结构。第二章介绍数据表示与指令系统。包括基本的数据表示和高级数据表示，指令的寻址方式，指令格式的优化设计和指令设计的两种风格，DLX 指令集结构。第三章介绍输入输出系统。包括输入输出原理，总线设计，中断系统，Pentium 系列微计算机的中断系统与 APIC 技术，I/O 通道与 I/O 处理机。第四章介绍存储系统。包括存储系统的基本原理，虚拟存储技术，Cache 存储器技术，Pentium 微处理器中虚拟存储技术和 Cache 存储器。第五章介绍流水和向量处理技术。包括标量流水技术，标量流水线的性能分析，标量流水中的障碍及控制，标量非线性流水线的调度，超标量与超流水技术，Pentium 微处理器中流水技术，指令的动态执行技术，向量流水处理机。第六章介绍并行处理技术。包括并行处理的基本概念，SIMD 计算机阵列基本结构，SIMD 计算机的互连网络，SIMD 计算机举例，多处理机系统。第八章介绍几种新型计算机系统。

本教材第二、三章由张玉萍编写，其他各章由高辉编写，全书由高辉统稿。

由于作者水平有限，难免有错漏之处，敬请读者批评指正。



为了帮助广大读者更好地掌握这门知识，我们随后将推出与本教材配套的《计算机系统结构学习辅导及习题解答》。

在本教材的编写过程中，得到了武汉大学教务部和武汉大学出版社的大力帮助，在此表示衷心地感谢。

作 者

2004 年于武汉大学

目 录

前 言	1
第一章 计算机系统结构的设计基础	1
1. 1 计算机系统结构的基本概念	2
1. 1. 1 计算机系统的层次结构	2
1. 1. 2 计算机系统结构	3
1. 1. 3 计算机组成与实现	5
1. 1. 4 计算机系统结构的分类	6
1. 2 计算机系统设计技术	9
1. 2. 1 计算机系统设计原理	9
1. 2. 2 计算机系统设计的方法	11
1. 3 计算机系统的性能评价	13
1. 3. 1 CPU 性能	14
1. 3. 2 MIPS 和 MFLOPS	15
1. 3. 3 基准测试程序	16
1. 3. 4 性能评价结果的统计和比较	17
1. 3. 5 Intel 微处理器性能评价	19
1. 4 计算机系统结构的发展	24
1. 4. 1 计算机系统结构的演变	24
1. 4. 2 软件、应用和器件对系统结构发展的影响	26
习 题	34
第二章 数据表示与指令系统	38
2. 1 浮点数据表示和 IEEE754 标准	38
2. 1. 1 数据表示、数据类型、数据结构的关系	38
2. 1. 2 浮点数据表示	40
2. 1. 3 IEEE754 标准浮点数表示	47
2. 2 高级数据表示	49
2. 2. 1 自定义数据表示	50



2.2.2 向量数据表示	54
2.2.3 堆栈数据表示	56
2.3 寻址方式与指令格式的优化设计	58
2.3.1 寻址方式	58
2.3.2 程序定位技术	64
2.3.3 指令格式的优化与设计	66
2.4 指令系统设计的两种风格	72
2.4.1 指令系统的功能设计	72
2.4.2 复杂指令系统计算机 (CISC) 设计风格	74
2.4.3 精简指令系统计算机 (RISC) 设计风格	80
2.4.4 CISC 机和 RISC 机的比较	90
2.5 DLX 指令集结构	92
2.5.1 DLX 指令集结构	92
2.5.2 DLX 指令集结构效能分析	99
习题	100
第三章 输入输出系统	104
3.1 输入输出系统原理	104
3.1.1 输入输出系统的特点	105
3.1.2 输入输出系统的基本方式	107
3.2 总线设计	108
3.2.1 总线的类型	109
3.2.2 总线的控制方式	112
3.2.3 总线通信技术	116
3.2.4 总线设计	119
3.2.5 Pentium 微处理器的总线系统	121
3.2.6 Pentium 系列微计算机系统的输入输出总线 (USB 和 IEEE 1394)	123
3.3 中断系统	128
3.3.1 中断系统的分类与分级	128
3.3.2 中断系统软、硬件功能分配	132
3.3.3 中断响应与中断屏蔽	135
3.3.4 Pentium 系列微计算机的中断系统	138
3.3.5 APIC 技术简介	142
3.4 通道处理机	153
3.4.1 通道的功能	153
3.4.2 通道的工作原理	156



3.4.3 通道的类型	158
3.4.4 通道流量分析	161
3.5 外围处理机	164
3.5.1 外围处理机的功能	165
3.5.2 外围处理机的特点	166
3.5.3 外围处理机的分类	168
习 题	169
第四章 存储系统	173
4.1 存储系统的原理	173
4.1.1 存储系统的意义	173
4.1.2 存储系统的性能指标	174
4.1.3 “Cache-主存”和“主存-辅存”层次	176
4.1.4 主存频宽的平衡与提高	178
4.2 虚拟存储器	185
4.2.1 虚拟存储器的管理方式	185
4.2.2 页式虚拟存储器的构成	192
4.2.3 加快页式虚拟存储器地址变换的方法	201
4.2.4 提高主存命中率的方法	206
4.2.5 虚拟存储器的保护技术	209
4.2.6 Pentium 微处理器的虚拟存储器	210
4.3 高速缓冲存储器 (Cache)	216
4.3.1 Cache 工作原理	217
4.3.2 地址映像与地址变换	218
4.3.3 Cache 替换算法及其实现	222
4.3.4 Cache 一致性与写策略	227
4.3.5 Cache 性能分析	232
4.3.6 Pentium PC 的 Cache	249
习 题	254
第五章 流水技术与向量处理	260
5.1 标量流水工作原理	260
5.1.1 指令的重叠解释方式	260
5.1.2 先行控制技术	263
5.1.3 标量流水工作原理	264
5.1.4 标量流水线的分类	266



5.1.5 标量流水线性能分析	269
5.2 标量流水中的障碍及控制	274
5.2.1 局部性相关及处理	275
5.2.2 全局性相关及处理	278
5.2.3 流水线的中断及处理	282
5.3 流水线的调度技术	283
5.3.1 非线性流水线的静态调度技术	283
5.3.2 流水线的动态调度技术	285
5.4 先进的流水技术	290
5.4.1 超标量流水线技术	290
5.4.2 超流水线技术	296
5.4.3 超标量超流水线技术	298
5.4.4 超长指令字 (VLIW) 技术	299
5.5 Pentium 微处理器中的流水技术	301
5.5.1 Pentium 微处理器的超标量流水线	301
5.5.2 Pentium 微处理器 U, V 流水线指令配对	305
5.5.3 Pentium 微处理器中的 BTB	307
5.5.4 Pentium II / III 微处理器中动态执行技术	309
5.6 向量流水技术	314
5.6.1 向量流水的基本概念	315
5.6.2 CRAY-1 型向量流水处理机	317
5.6.3 增强向量处理性能的方法	319
习 题	324
 第六章 并行处理技术	 330
6.1 并行处理技术的基本概念	330
6.2 SIMD 并行计算机 (阵列处理机)	331
6.2.1 阵列机的基本结构	331
6.2.2 阵列机的主要特点	333
6.2.3 典型 SIMD 计算机举例	334
6.3 SIMD 并行计算机算法	340
6.3.1 矩阵加	341
6.3.2 矩阵乘	342
6.3.3 累加和	344
6.4 SIMD 计算机的互连网络	346
6.4.1 互连网络的设计目标	346



6.4.2 互连函数	347
6.4.3 互连网络的分类和结构参数	351
6.4.4 静态互连网络	354
6.4.5 动态互连网络	358
6.5 多处理机	371
6.5.1 多处理机的特点	371
6.5.2 多处理机的分类	372
6.5.3 多处理机间的互连方式	375
6.5.4 多处理机系统中并行性开发	380
6.5.5 多处理机操作系统	388
6.5.6 多处理机的调度策略	390
习 题	393
第七章 新型计算机结构	397
7.1 脉动阵列计算机	397
7.2 数据流计算机	398
7.2.1 数据流计算机的基本工作原理	399
7.2.2 数据流程图和数据流语言	399
7.2.3 数据流计算机的基本结构	403
7.2.4 数据流计算机存在的主要问题	407
7.3 归约机	407
7.3.1 函数式程序设计语言	407
7.3.2 归约机的结构特点	408
7.3.3 面向函数式语言的归约机	408
7.4 人工智能计算机	409
7.4.1 人工智能计算特征	410
7.4.2 AI 计算机的分类和设计方法	411
7.4.3 PROLOG 推理机	412
7.4.4 AI 计算机的研究进展	412
7.4.5 RWC 研究计划	413
习 题	415
主要参考文献	417



第一章 计算机系统结构的设计基础

世界上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。由于 ENIAC 输入和更换程序特别繁杂, 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。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 3 000 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。

50 多年来, 计算机系统性能得到了大幅度地提高, 价格却大幅度地下降。计算机的发展已经历了四次更新换代, 现在正处于第五代。依据半导体技术发展水平, 这五代划分应该是:

第一代: 1945 ~ 1954 年, 电子管和继电器;

第二代: 1955 ~ 1964 年, 晶体管和磁芯存储器;

第三代: 1965 ~ 1974 年, 中、小规模集成电路 (MSI-SSI);

第四代: 1975 ~ 1990 年, LSI/VLSI 和半导体存储器;

第五代: 1990 年至今, ULSI/GSI (Giga-Scale Integration) 巨大规模集成电路。

计算机系统更新换代的标志主要是两个方面: 一是由于器件不断发展, 使计算机系统的工作速度、功能、集成度和可靠性等指标不断提高和价格不断降低而造成的; 二是得益于计算机系统结构的改进。有人统计在 1965 年到 1975 年期间, 计算机系统性能提高近 100 倍, 其中由于器件性能提高使其性能增加 10 倍, 而另外的 10 倍, 主要归功于系统结构改进。在 50 多年的发展进程中, 器件在技术上的改进是比较稳定的, 而系统结构的改进则有较大起伏。特别是近 20 年来, 由于计算机系统的设计对集成电路技术的依赖性大为增加, 从而使得在这一期间内, 各类不同的计算机系统的性能增长率有了差异。

巨型机的性能增长得益于器件技术和系统结构两方面的改进; 大型机的性能增长则主要靠器件工艺上的改进, 因为系统结构方面的改进没有新的突破; 小型机的发展, 一方面是由于计算机实现方法有了较大的改进, 另一方面是因为采用了许多大型机中行之有效的先进技术。然而这三类计算机, 在 1970 年到 1990 年期间, 每年计算机的性能平均增长率均只在 18% 左右。与之形成明显对照的是微型计算机的性能



增长非常快,每年的平均增长率约为 35%,这是因为微型计算机能从集成电路技术的进展中得到最为直接的好处。自 20 世纪 80 年代起,微处理机技术实际上已成为新系统结构和老系统结构更新时所选用的主要技术。

自 1985 年开始,一种具有新颖设计风格的系统结构,即 RISC 技术的系统结构,为计算机工业界所青睐。它将集成电路技术进展、编译技术改进和新的系统结构设计思想三者有机地结合起来,从而使得这种风格设计的计算机系统的性能以每年增长一倍的高速率加以改进。应该指出的是,这种改进的基础是通过以往计算机如何被使用的模拟实验数据进行定量分析后获得的。有的学者将这种设计风格称为定量分析的计算机系统结构设计风格,显然这比传统的定性设计风格要精确得多,开发 RISC 技术的两位先驱者,美国加州大学伯克莱分校的 D. Patterson 教授和斯坦福大学的 J. Hennessy 教授是这种定量分析设计方法的主要倡导者。

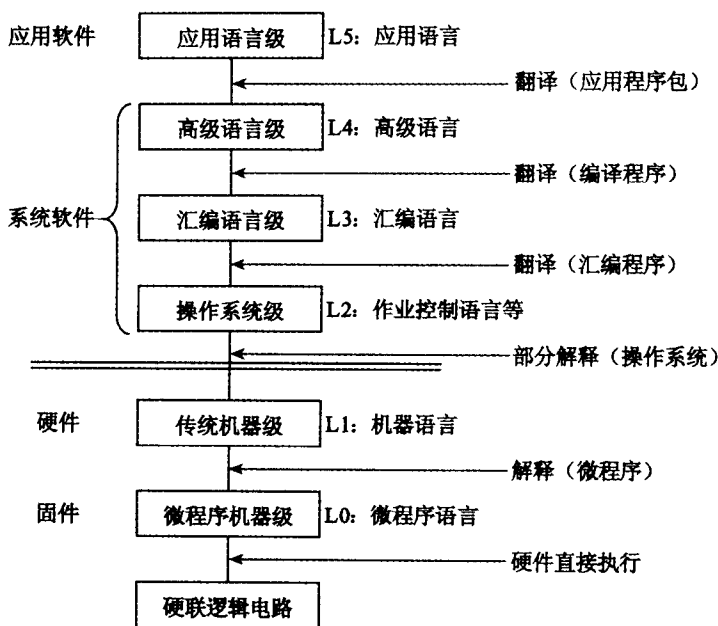
1.1 计算机系统结构的基本概念

1.1.1 计算机系统的层次结构

现代通用计算机系统是由硬件和软件组成的一个复杂系统,按其功能可划分为多级层次结构,如图 1.1 所示。层次结构由上往下依次为应用语言机器级、高级语言机器级、汇编语言机器级、操作系统机器级、传统机器级和微程序机器级。对于一个具体的计算机系统,层次的多少会有所不同。

各机器级的实现主要靠翻译或解释,或者是这两者的结合。翻译(Translation)是先用转换程序将上一级机器级上的程序整个地变换成下一级机器级上可运行的等效程序,然后再在下一级机器级上去实现的技术。解释(Interpretation)则是在下一级机器级上用它的一串语句或指令来仿真上一级机器级上的一条语句或指令的功能,通过对上一级机器语言程序中的每条语句或指令逐条解释来实现的技术。

应用语言虚拟机器级是为了满足信息管理、人工智能、图像处理、辅助设计等专门的应用来设计的。使用面向某种应用环境的应用语言(L5)编写的程序一般是经应用程序包翻译成高级语言(L4)程序后,再逐级向下实现的。高级语言机器级上的程序可以先用编译程序整个地翻译成汇编语言(L3)程序或机器语言(L1)程序,再逐级或越级向下实现,也可以用汇编语言(L3)程序、机器语言(L1)程序,甚至微指令语言(L0)程序解释实现。对汇编语言(L3)源程序则先用汇编程序整个将其变换成等效的二进制机器语言(L1)目标程序,再在传统机器级上实现。操作系统程序虽然已发展成用高级语言(如面向编写操作系统软件的 C 语言)编写,但最终还是要用机器语言程序或微指令程序来解释的。它提供了传统机器级所没有,但为汇编语言和高级语言使用和实现所用的基本操作、命令和数据结构,例如,文件管理、存储管理、进程管理、多道程序共行、多重处理、作业控制等所用到的操作命令、语句和数据结构



等。因此,操作系统机器级放在传统机器级和汇编语言机器级之间是适宜的。传统机器级采用组合逻辑电路控制,其指令可直接用硬件来实现,也可以采用微程序控制,用微指令(L0)程序来解释实现。微指令直接控制硬件电路的动作。

就目前的状况来看,第0级用硬件实现,第1级用微程序(固件)实现,第2级到第5级大多用软件实现。我们称以软件为主实现的机器为虚拟机器,以区别于由硬件或固件实现的实际机器。虚拟机器不一定全都由软件实现,有些操作可以用固件或硬件实现。如操作系统中的某些命令可由比它低两级的微程序解释,或全部用硬件来实现。高级语言机器则直接用微程序解释或用硬件实现,没有编译软件。

计算机系统结构作为一门学科,主要研究软件、硬件功能分配和对软件、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。采用何种实现方式,要从整个计算机系统的效率、速度、价格、资源状况等方面考虑,对软件、硬件、固件取舍进行综合平衡。软件和硬件在逻辑功能上是等效的。原理上,软件实现的功能完全可以用硬件或固件完成,硬件实现的功能也可以由软件的模拟来完成,只是其性能、价格、实现的难易程度等有所不同。具有相同功能的计算机系统,其软、硬件功能分配比例可以在很宽的范围内变化。

1.1.2 计算机系统结构

“计算机系统结构”这个名词来源于英文 Computer Architecture, 也有译成“计算



机体系结构”的。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”,它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,“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,已成为一门学科。但对“计算机系统结构”一词的含义仍有多种说法,并无统一的定义。

计算机系统结构这个词是 Amdahl 等人在 1964 年提出的。他们把系统结构定义为由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。按照计算机层次结构,不同程序设计者所看到的计算机有不同的属性。例如,对于使用 FORTRAN 高级语言程序员来讲,一台 IBM3090 大型机、一台 VAX11/780 小型机或一台 PC 微型机,看起来都是一样的,因为在这三台计算机上运行他所编制的程序,所得到的结果是一样的。但对于使用汇编语言程序的程序员来讲,由于这三台机器的汇编语言指令完全不一样,他所面对的计算机的属性就会不一样。另外,即使对同一台机器来讲,处在不同级别的程序员,例如应用程序员、高级语言程序员、系统程序员和汇编程序员,他们所看到的计算机外特性也是完全不一样的。那么通常所讲的计算机系统结构的外特性应是处在哪一级的程序员所看到的外特性呢?比较一致的看法是机器语言程序员或编译程序编写者所看到的外特性,这种外特性是指由他们所看到的计算机基本属性,即计算机的概念性结构和功能特性,这是机器语言程序员或编译程序编写者为使其所编写、设计或生成的程序能在机器上正确运行所必须遵循的。由机器语言程序员或编译程序编写者所看到的计算机的基本属性是指传统机器级的系统结构,在传统机器级之上的功能被视为属于软件功能,而在其之下的则属于硬件和固件功能。因此,计算机系统的概念性结构和功能属性实际上是计算机系统中软、硬件之间的界面。

在计算机技术中,一种本来是存在的事物或属性,但从某种角度看似乎不存在,称为透明性现象。通常,在一个计算机系统中,低层机器级的概念性结构和功能特性,对高级语言程序员来说是透明的。由此看出,在层次结构的各个级上都有它的系统结构。

就目前的通用机来说,计算机系统结构的属性应包括以下几个方面:

- 硬件能直接识别和处理的数据类型和格式等的数据表示;
- 最小可寻址单位、寻址种类、地址计算等的寻址方式;
- 通用/专用寄存器的设置、数量、字长、使用约定等的寄存器组织;
- 二进制或汇编级指令的操作类型、格式、排序方式、控制机构等的指令系统;
- 内存的最小编址单位、编址方式、容量、最大可编址空间等的存储系统组织;
- 中断的分类与分级、中断处理程序功能及入口地址等的中断机构;
- 系统机器级的管态和用户态的定义和切换;
- 输出设备的连接、使用方式、流量、操作结束、出错指示等的机器级 I/O 结构;
- 各部分的信息保护方式和保护机构等。



1.1.3 计算机组成与实现

计算机组成(Computer Organization)指的是计算机系统结构的逻辑实现,也常称为计算机组织。它包括机器级内的数据流和控制流的组成以及逻辑设计等。它着眼于机器级内各事件的排序方式与控制机构、各部件的功能及各部件间的联系。计算机组成设计要解决的问题是在所希望达到的性能和价格下,怎样最佳、最合理地把各种设备和部件组织成计算机,以实现所确定的系统结构。计算机组成设计主要是围绕提高速度,着重从提高操作的并行度、重叠度以及分散功能和设置专用功能部件来进行的。

计算机组成设计要确定的方面一般应包括:

- 数据通路宽度(在数据总线上一次并行传送的信息位数多少)。
- 专用部件的设置(设置哪些专用部件,如乘法专用部件、浮点运算部件、字符处理部件、地址运算部件等,每种专用部件设置的数量,这些都与机器所需达到的速度、专用部件的使用频度高低及允许的价格等有关)。
- 各种操作对部件的共享程度(共享程度高,即使操作在逻辑上不相关也只能分时使用,限制了速度,但价格便宜,可以设置多个部件降低共享程度,提高操作并行度来提高速度,但价格也将提高)。
- 功能部件的并行度(功能部件的控制和处理方式是采用顺序串行,还是采用重叠、流水或分布处理)。
- 控制机构的组成方式(事件、操作的排序机构是采用硬联控制还是用微程序控制,是采用单机处理还是用多机处理或功能分布处理)。
- 缓冲和排队技术(在不同部件之间怎样设置及设置多大容量的缓冲器来弥补它们的速度差异;是采用随机方式,还是先进先出、先进后出、优先级或循环方式来安排等待处理事件的先后顺序)。
- 预估、预判技术(为优化性能和优化处理,采用什么原则来预测未来的行为)。
- 可靠性技术(采用什么样的冗余技术和容错技术来提高可靠性)。

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。

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,但又有紧密的联系。在学习和理解时有两点需要注意:

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 10 年前也大不相同。例如,10 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考察的范畴之内。

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这些存储器管理算法并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。

1.1.4 计算机系统结构的分类

研究计算机系统的分类方法有助于认识计算机系统结构的组成和特点,理解计算机系统的工作原理和性能。常用的计算机系统结构分类方法有三种:

1. Flynn 分类法

1966 年 M. J. Flynn 提出了按照指令流和数据流的多倍性概念进行分类的方法,其定义如下:

指令流(instruction stream)——机器执行的指令序列。

数据流(data stream)——由指令流调用的数据序列,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(瓶颈)上同时处于同一执行阶段的可并行执行的指令或数据的最大可能个数。

按照指令流和数据流的不同组织方式,把计算机系统的结构分为以下四类:

- ① 单指令流单数据流 SISD(Single Instruction stream Single Data stream);
- ② 单指令流多数据流 SIMD(Single Instruction stream Multiple Data stream);
- ③ 多指令流单数据流 MISD(Multiple Instruction stream Single Data stream);
- ④ 多指令流多数据流 MIMD(Multiple Instruction stream Multiple Data stream)。

对应于这四类计算机的基本结构框图,如图 1.2 所示。SISD 是传统的顺序处理计算机。SIMD 以阵列处理机或并行处理机为代表。MISD 在实际上代表何种计算机,也存在着不同的看法,有的文献把流水线结构机器看做是 MISD 结构,多处理机