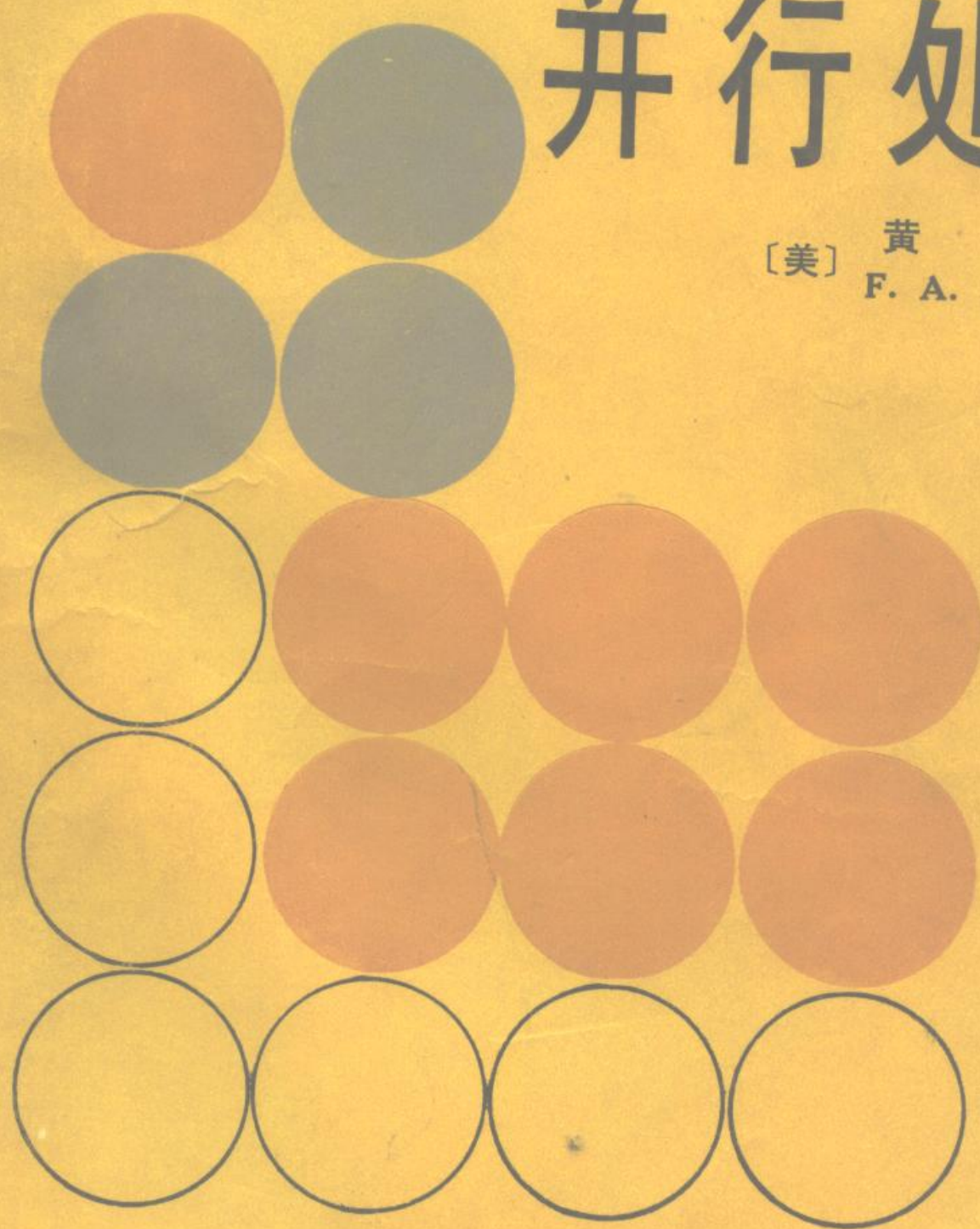


计算机结构与 并行处理

[美] 黄 铠 著
F. A. 布里格斯



科学出版社

计算机结构与并行处理

[美] 黄 铠 著
F. A. 布里格斯

金 兰 郑衍衡 王鼎兴 译
沈美明 郑伟民

金 兰 校

科 学 出 版 社

内 容 简 介

本书全面系统地介绍了当代最为流行的并行处理的概念及理论研究和工程实践成果。理论方面主要涉及高等计算机系统结构、并行计算理论和算法、并发程序设计语言、操作系统在多重处理下的资源调度、进程同步和系统控制、并行处理系统的结构原理、性能评估和改善等重要内容。实践方面主要介绍流水线向量计算机、阵列处理机、多处理机、数据流和 VLSI 计算机等先进计算机实例。本书内容丰富全面，既介绍了当今科学家及作者本人的重要研究成果，也引述了各大计算机公司和高等学校在超级计算机设计开发方面的大量实践经验。全书编排层次合理，叙述由浅入深。为了便于读者学习和研究，各章结尾还附有大量习题和文献注释。

本书既可作为高等院校计算机专业高年级学生和研究生学习“高等计算机系统结构”和“并行处理”等课程的教科书或参考书，也可供计算机专业科技人员学习参考。

Kai Hwang, Faye A. Briggs

COMPUTER ARCHITECTURE AND PARALLEL PROCESSING

McGraw-Hill Book Company, 1984

计算机结构与并行处理

[美] 黄 铠 著
F. A. 布里格斯

金 兰 郑衍衡 王鼎兴 译
沈美明 郑纬民

金 兰 校
责任编辑 黄岁新

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码: 100707

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1990 年 6 月 第 一 版 开本: 787×1092 1/16

1990 年 6 月 第 一 次 印 刷 印张: 43

印数: 平 1—1 140 插页: 插 2

精 1—550 字数: 994 000

ISBN 7-03-001538-X/TP·104 (平)

ISBN 7-03-001702-1/TP·127 (精)

定 价: 平 装 39.80 元

布 脊 精 装 42.00 元

译 者 的 话

《计算机结构与并行处理》中译本就要出版了。我怀着兴奋的心情记述一段友谊的经历,并代表译者写几句话。

我初次认识本书作者——美国黄铠教授和布里格斯教授——是在 1979 年春。特别是黄铠教授,我们一见面就在编著一本关于并行处理的教科书上有了共同语言。此后我们曾就编写大纲交换过多次意见。虽然我们的合作写书计划最终未成现实,但是黄铠教授的待人热忱、治学严谨、学术思想活跃、开拓和著述工作的精益求精,给我留下了深刻的印象。我想这就是他的这本著作受到国际科技界重视而成为一本畅销书的根本原因。

黄铠教授还为我们编写的一本并行处理教科书给予了有力的支持,并且委托我们与科学出版社洽商在国内独家翻译和出版他的名作,这同样体现了他和我们之间的特殊友谊。我 1984 至 1986 年执教于美国,黄铠教授曾经邀请我访问他所在的南加州大学,并让我在他家就榻和欢聚一周。我在感激之余,也因未能及早完成本书译稿而感到不安。后来,我邀集几位同仁加紧了本书的翻译工作,他们是上海工业大学的郑衍衡教授(第一、二、十章)、清华大学的王鼎兴教授(第五、六章)、沈美明副教授(第三、四章)和郑纬民副教授(第七、八、九章)。我回国后伏案数月完成了全部译稿的统编校订工作,又加译了黄铠教授寄来的最新补充材料,编成本书第十一章。

我们衷心希望本书能使读者满意,并在国内能成为一本受计算机科技工作者欢迎的书。如果能做到这一点,首先应归功于原著质量的高水平,也要感谢科学出版社的大力支持,而我们自己只不过尽了译者应负的一份责任,尤其希望读者给我们提出批评和意见。

金 兰

1988 年 10 月

中 文 版 序

拙著“Computer Architecture and Parallel Processing”是 1984 年由纽约市的 McGraw-Hill 图书公司以英文出版的。至 1987 年初,此书在欧美已经印刷了五次。现在由清华大学计算机科学与技术系的同仁,金兰教授、王鼎兴教授、沈美明副教授、郑纬民副教授以及上海工业大学的郑衍衡教授联合翻译为中文,以飨国内读者之需。我非常感激他们在百忙中抽空完成这件工作。

中文版中,我们增添了一些新的材料,这样国内的读者就能预先看到此书第二版的一些新材料。著者在此致谢科学出版社与清华大学对中文版问世所做的大量工作。

黄 钊

1987 年 3 月 24 日

于美国南加州大学

原 版 序

在许多科学和工程应用领域中,对高性能计算机的需求日益增长,这些领域包括结构分析、气象预报、石油勘探、核聚变能研究、医疗诊断、气动力学模拟、人工智能、专家系统、工业自动化、遥感、军事防御、遗传工程、社会经济等。没有超强功能的计算机,上述许多推动人类文明进步的复杂课题便不可能在合理的时间内取得进展。要获得高性能,不但要依赖于使用更快速可靠的器件,而且取决于在计算机结构和处理技术上取得的重大进展。本书主要论述高等计算机结构、并行计算理论、最佳资源分配、快速算法、高效程序设计语言以及为满足上述需求对高性能价格比计算机系统的应用要求。

高等计算机结构的中心是并行处理概念。现代并行计算机系统可以按特性分为三种结构类型: **流水线计算机、阵列处理机和多处理机系统**。要发展和应用这些计算机系统,需要对它们的基础硬件和软件结构有广阔的知识,并且了解并行算法和最佳机器资源分配之间的密切作用关系。本书为读者提供必要的知识去设计新的计算机系统;改进已有的计算机系统;发展快速算法;以及在求解大型计算问题时分配硬、软件资源。

本书的读者对象是计算机科学、电机工程、工业工程以及其它有关专业的大学四年级学生和研究生,这些读者需要为解决特定的应用问题而使用高性能大中型计算机、附属式向量处理机、科学超级计算机和多处理机。并行处理既可以在硬件/软件系统级起作用,也可以在算法和程序设计级起作用,它要求在计算机中并发执行多个程序。并行处理计算机提供一种高性能价格比的手段,它能够通过并发活动获得高系统性能。

计算机科学工作者、系统设计师、高级程序员、应用工程师、计算实验工作者以及计算机专业人员都能够在本书中找到他们各自工作中所需要的资料。并行处理计算机的许多研究人员、设计人员和用户所取得的成果均已收入本书。本书提供的资料是由两位作者在普度大学和莱斯大学讲授的计算机结构和高级计算等几门课程发展而来。

全书组织成十章*,第一章介绍并行处理的基本概念和计算机结构,并对计算机系统内的并行性和各种应用领域进行综述,以给读者提供预备知识。第二章叙述并行处理计算机需用的存储器层次结构和输入输出子系统。

第三和第四章专门讨论流水线超级计算机的设计原理和应用。将要研究的系统包括 Star-100, TI-ASC, Cray-1, Cyber-205, Fujitsu VP-200, CDC-NASF, 以及附属科学处理机 AP-120B (FPS-164), IBM 3838 和 Datawest MATP。以最佳向量化为目标对语言和编译程序的要求进行评估。此外,还叙述流水线向量处理方法和流水线计算机的性能评价。

第五和第六章叙述处理机阵列的互连结构。这种阵列处理机可以对多数据流进行单指令流处理。还为阵列处理机和关联处理机叙述几种并行算法。供实例研究的系统包括 Illiac-IV, Burroughs 科学处理机 BSP, STARAN, PEPE 以及大规模并行处理机 MPP。

* 中译本为十一章,请见第十一章脚注。——译者注

还为同步阵列处理机提出了改进性能的方法。

在第七、第八和第九章中,我们研究多处理机系统的硬件系统结构、操作系统控制、并行算法和性能评价。将要叙述三种探索研究性多处理机 C. mmp, S-1 和 Cm* 的设计经验.将要研究的商品化多处理机包括 IBM 370/168 MP, 3033 和 3081, Univac 1100/80 和 1100/90, Cray X-MP, Tandem/16 以及 Denelcor 的 HEP. 与设计紧密耦合的多处理机有关研究课题也将予以讨论。

第十章专门研究新的计算概念以及它们的实现方法. 介绍数据驱动计算原理, 概述函数式语言和已有的数据流计算机结构。最后, 我们将研究适合于用超大规模集成电路硬件实现的并行算法, 叙述 VLSI 结构对图象处理的应用。

为阅读本书所需的先修课程是一门关于计算机组织和程序设计的大学初级课程。由于大多数向量超级计算机都是 Fortran 机器, 所以我们在第三至第六章中使用 Fortran 及其扩展语言. 对于第七至第九章的多处理机, 我们使用并发 Pascal 作为说明性语言. 用星号*标注的各节是面向研究的题目. 在学习这些研究题目时, 希望读者具有某些离散数学和概率论的知识。在第一遍阅读时, 可以跳过这些困难的章节而不致失掉连续性。家庭作业题目主要提供读者深入思考和通过实际练习获得设计、应用和评价并行计算机的经验。

并行处理和计算机结构是两个宽阔开放的研究和发展领域。每一章的后面都附有文献注释, 以帮助有兴趣的读者为进一步研究找到补充的参考资料。如有任何错误或遗漏, 应由作者负全责。对某些计算机专家的创造性工作未列入本书, 作者向他们表示歉意. 计算机领域的发展如此迅速, 以至于没有一本书是可能概括每一项最新成就的。但是, 我们仍欢迎从读者那里获得意见和批评。请读者们将你们的意见直接寄给作者, 以便在将来重印或修订本书时加以改进。

在开设计算机结构和并行处理两门课程序列时, 可使用本书作为教材。每一门课包含 45 讲, 每讲 50 分钟。我们建议在这两门课程序列的第一门课程中包含下列材料, 其余各节则留到第二门课程讲授。

第一门课程建议安排下列材料:

章	节 和 小 节
1	1.1, 1.2, 1.3, 1-3, 1.3.5, 1.4.1, 1.5
2	2.1.1, 2.1.3, 2.2.1-2, 2.4.1-2, 2.5.1
3	3.1.1-2, 3.2.1-2, 3.3.1-2, 3.4.1
4	4.1, 4.2.1, 4.3.1, 4.4.1-3
5	5.1, 5.2.1-2, 5.3.1, 5.4.1
6	6.1, 6.2.1-2, 6.3.1, 6.4.1
7	7.1, 7.2.1-2, 7.4.1, 7.5.1
8	8.1.1, 8.2.1, 8.3.1, 8.4.1
9	9.1, 9.2.1, 9.5.1-2, 9.6.1
10	10.1.1, 10.2.1, 10.3.1

第一门课程适合于大学四年级学生和一年级研究生。第二门课程则主要是为了研究生。第一门课程是第二门课程的先修课。如果使用本书只开设一门课, 则教师可从第二门课移动某些章节到第一门课, 以便更完全地覆盖某些根据教师和学生兴趣选择的题目。

这便需要在上表某些章节以及第二门课的选加章节之间进行权衡取舍。本书还有一本题解手册,将由 McGraw-Hill 出版公司只向教师提供。该手册包含全部习题的解答,外加一些适用于课程设计的设计作业。欢迎教师直接和作者或 McGraw-Hill 出版公司的代表联系,提出建议,或者交流使用本书作为必读教材或研究参考书的经验。

作者要向许多人表示感谢,是他们在专业上的鼓励与协助才使得本书漫长的编写和生产过程成为一次很愉快的努力。特别要感谢提出过宝贵的意见与建议的。Harold S. Stone, Jean Loup Baer, C. V. Ramamoorthy, Tse-yun Feng, John P. Hayes, King-Sun Fu, Clarence L. Coates, V. Carl Hamacher, Herschel H. Loomis, Jr., Bart Sinclair, J. Robert Jump, Edward S. Davidson, H. J. Siegel, Tom Mowbray, Wolfgang Händler, Kenichi Miura, Lionel M. Ni, Michel Dubois 和 Shun-Piao Su. 许多我们的学生帮助我们改进了书稿,特别要感谢 Chi-Yuan Chin, Zhiwei Xu 和 William Carlson 的帮助。对于打字、绘图和秘书工作上的协助,予以感谢 Andy Hughes, Wanda Booth, Linda Stovall, Vicki Johnson, Pat Loomis, Mickey Krebs, Sharon Katz, Nancy Lein 和 D. Ringe. 最后但不是最少地,我们赞赏 McGraw-Hill 图书公司的编辑和生产人员为本书的生产所作的出色工作。没有上面许多人不懈的努力,本书也许仍只处在它的准备阶段。

黄 铠

F. A. 布里格斯

目 录

译者的话

中文版序

原版序

第一章 并行处理引论	1
1.1 计算机系统的进展	1
✓ 1.1.1 计算机系统经历的各代	1
1.1.2 朝向并行处理的发展趋势	3
1.2 单处理机系统中的并行性	6
✓ 1.2.1 基本的单处理机系统结构	6
1.2.2 并行处理机构	7
1.2.3 子系统带宽的平衡	9
1.2.4 多道程序和分时	11
1.3 并行计算机结构	13
✓ 1.3.1 流水线计算机	13
✓ 1.3.2 阵列计算机	16
✓ 1.3.3 多处理机系统	17
1.3.4 并行计算机的性能	18
1.3.5 数据流和新概念	20
1.4 系统结构的分类方法	21
1.4.1 指令-数据流的多重性	21
1.4.2 串行和并行处理	24
1.4.3 并行性和流水线处理	26
1.5 并行处理的应用	28
1.5.1 预测模型的构造和模拟	29
1.5.2 工程设计和自动化	30
1.5.3 能源勘探	32
1.5.4 医学、军事和基础研究	33
1.6 参考文献注释与习题	34
第二章 存储器和输入输出子系统	37
2.1 层次式存储器结构	37
✓ 2.1.1 存储器的层次结构	37
2.1.2 存储器层次结构的优化	40
✓ 2.1.3 主存的寻址方式	41
2.2 虚拟存储器系统	43
✓ 2.2.1 虚拟存储器的概念	43
✓ 2.2.2 分页存储器系统	46
2.2.3 分段存储器系统	52

2.2.4 段页式存储器	54
2.3 存储分配与管理	57
2.3.1 存储器策略的分类	57
2.3.2 最佳负荷控制	62
2.3.3 存储管理策略	65
2.4 高速缓冲存储器及其管理	70
✓2.4.1 高速缓冲存储器的特性	70
✓2.4.2 高速缓冲存储器组织	73
2.4.3 读取策略和主存更新策略	81
2.4.4 存储块替换策略	83
2.5 输入输出子系统	85
✓2.5.1 I/O 子系统的特性	86
2.5.2 中断机构和特殊硬件	88
2.5.3 I/O 处理和 I/O 通道	93
2.6 参考文献注释与习题	101
第三章 流水线和向量处理原理	105
3.1 流水线: 一种重叠的并行性	105
✓3.1.1 线性流水线原理	105
✓3.1.2 流水线处理机的分类	109
3.1.3 通用流水线和预约表	111
3.1.4 交叉存储器组织	113
3.2 指令和运算流水线	118
✓3.2.1 流水线指令部件的设计	119
✓3.2.2 运算流水线设计举例	120
3.2.3 多功能与阵列流水线	130
3.3 设计流水线处理机的原理	135
✓3.3.1 指令预取和转移处理	135
✓3.3.2 数据缓冲及总线结构	139
3.3.3 内部定向和寄存器标记	142
3.3.4 冒险的测试与分解	145
3.3.5 作业顺序与冲突防止	146
*3.3.6 动态流水线与可重构性	150
3.4 向量处理的要求	153
✓3.4.1 向量处理的特性	153
*3.4.2 多向量任务调度	158
3.4.3 流水线向量处理方法	163
3.5 参考文献注释与习题	166
第四章 流水线计算机及向量化方法	169
✓4.1 流水线计算机的领域	169
4.1.1 向量超级计算机	169
4.1.2 科学附属处理机	170
4.2 早期的向量处理机	172

✓ 4.2.1	Star-100 和 TI-ASC 的系统结构	172
4.2.2	流方式的向量处理	177
4.3	科学附属处理机	180
✓ 4.3.1	AP-120B 的系统结构	180
4.3.2	后端向量计算	184
4.3.3	FPS-164、IBM 3838 和 Data west 公司的 MATP	187
4.4	近期的向量处理机	191
✓ 4.4.1	Cray-1 的系统结构	192
✓ 4.4.2	流水线链接和向量循环	197
✓ 4.4.3	Cyber-205 的系统结构	203
4.4.4	Cyber-205 和 CDC-NASF 的向量处理	207
4.4.5	Fujitsu 的 VP-200 及其特性	211
4.5	向量化和优化方法	217
4.5.1	向量处理的并行语言	217
4.5.2	向量化编译程序的设计	220
4.5.3	向量功能的优化	222
4.5.4	流水线计算机的性能评价	227
4.6	参考文献注释与习题	232
第五章	阵列处理机的结构和算法	235
✓ 5.1	SIMD 阵列处理机	235
5.1.1	SIMD 计算机组织	235
5.1.2	屏蔽与数据传送寻径机制	237
5.1.3	PE 间的通信	240
5.2	SIMD 互连网络	241
✓ 5.2.1	静态和动态网络	241
✓ 5.2.2	网格连接的 Illiac 网络	244
5.2.3	立方体互连网络	246
5.2.4	循环移数网络与数据变换网络	248
5.2.5	混洗交换与 Omega 网络	253
5.3	阵列处理机的并行算法	255
✓ 5.3.1	SIMD 矩阵乘法	255
5.3.2	在阵列处理机上实现并行排序	259
5.3.3	SIMD 快速傅里叶变换	263
5.3.4	SIMD 处理的连接问题	267
5.4	关联阵列处理	268
✓ 5.4.1	关联存储器组织	268
5.4.2	关联处理机 (PEPE 和 STARAN)	272
* 5.4.3	关联查找算法	275
5.5	参考文献注释与习题	277
第六章	SIMD 计算机及其提高性能的方法	281
✓ 6.1	SIMD 计算机的类型	281
6.1.1	阵列与关联处理机	282
6.1.2	SIMD 计算机展望	283

6.2 ILLIAC-IV 和 BSP 系统	284
✓6.2.1 Illiac-IV 系统结构	285
✓6.2.2 Illiac-IV 的应用	288
6.2.3 BSP 的系统结构	292
6.2.4 质数存储系统	295
6.2.5 BSP Fortran 向量化程序	297
6.3 大规模并行处理机	301
✓6.3.1 MPP 系统结构	301
6.3.2 处理阵列、存储器和控制	304
6.3.3 MPP 的图象处理	307
6.4 提高性能的方法	309
✓6.4.1 并行存储器的分配	310
6.4.2 阵列处理语言	314
*6.4.3 阵列处理机的性能分析	317
*6.4.4 多 SIMD 计算机组织	320
6.5 参考文献注释与习题	324
第七章 多处理机的结构和程序设计	328
7.1 功能结构	328
7.1.1 松散耦合的多处理机	328
7.1.2 紧密耦合的多处理机	335
7.1.3 多重处理对处理机特性的要求	342
7.2 互连网络	344
✓7.2.1 时分总线或公共总线	344
✓7.2.2 交叉开关和多端口存储器	349
7.2.3 多处理机的多级网络	353
*7.2.4 互连网络的性能	360
7.3 并行存储器组织	364
7.3.1 交叉访问的存储器构形	365
*7.3.2 存储器组织的性能权衡	369
7.3.3 多高速缓冲存储器的问题及其解决方法	372
7.4 多处理机的操作系统	379
✓7.4.1 多处理机操作系统的分类	379
7.4.2 多处理机的软件要求	381
7.4.3 操作系统的要求	383
7.5 多重处理并行性的开发	384
✓7.5.1 用以开发并行性的语言特征	385
7.5.2 程序并行性的检测	391
*7.5.3 程序和算法的重构	394
7.6 参考文献注释与习题	398
第八章 多重处理的控制和算法	404
8.1 进程间通信机构	404
✓8.1.1 进程同步机构	404
8.1.2 利用信号灯实现同步	410

8.1.3 条件临界区和管理	415
8.2 系统死锁和保护	419
✓ 8.2.1 系统死锁问题	419
8.2.2 死锁的预防和避免	421
8.2.3 死锁检测和恢复	423
8.2.4 保护方法	423
8.3 多处理机的调度策略	428
✓ 8.3.1 多处理机管理的范围	428
8.3.2 确定的调度模型	432
* 8.3.3 随机调度模型	439
8.4 多处理机的并行算法	444
✓ 8.4.1 并行算法的分类	445
8.4.2 同步并行算法	446
8.4.3 异步并行算法	452
* 8.4.4 并行算法的性能	456
8.5 参考文献注释与习题	463
第九章 多处理机系统的例子	468
9.1 多处理机系统领域	468
9.1.1 供研究的多处理机系统	468
9.1.2 商品化的多处理机	468
9.2 C.mmp 多处理机系统	469
9.2.1 C.mmp 系统的系统结构	469
9.2.2 Hydra 操作系统	473
* 9.2.3 C.mmp 的性能	476
9.3 S-1 多处理机系统	479
9.3.1 S-1 系统结构	479
9.3.2 多重处理的单处理机	481
9.3.3 S-1 软件的开发	486
9.4 HEP 多处理机系统	486
9.4.1 HEP 系统的系统结构	487
9.4.2 进程执行模块	489
9.4.3 HEP 的并行处理	494
9.5 大型多处理机系统	497
9.5.1 IBM 370/168MP, 3033 和 3081	497
9.5.2 IBM 多处理机的操作系统	503
9.5.3 Univac 1100/80 和 Univac 1100/90 系列	504
9.5.4 Tandem Nonstop 系统	512
9.6 Cray X-MP 和 Cray-2	518
9.6.1 Cray X-MP 系统结构	518
9.6.2 Cray X-MP 的多重任务	520
* 9.6.3 Cray X-MP 的性能	524
9.7 参考文献注释与习题	528

第十章 数据流计算机和 VLSI 计算	532
10.1 数据驱动计算及其语言	532
✓10.1.1 控制流和数据流计算机的比较	532
10.1.2 数据流图和语言	537
10.1.3 优点及存在的问题	541
10.2 数据流计算机的结构	543
✓10.2.1 静态数据流计算机	543
10.2.2 动态数据流计算机	547
10.2.3 其它数据流设计方案	553
10.3 VLSI 计算结构	556
✓10.3.1 搏动式阵列结构	557
*10.3.2 把算法映射到 VLSI 阵列	560
10.3.3 可重构的处理器阵列	564
10.4 VLSI 矩阵运算处理器	570
10.4.1 VLSI 运算模块	570
10.4.2 分块的矩阵算法	571
10.4.3 矩阵运算流水线	576
*10.4.4 实时图象处理	580
10.5 参考文献注释与习题	583
第十一章 高级并行处理及超级计算机系统结构	587
11.1 绪论	587
11.1.1 现代超级计算机的演变	587
11.1.2 性能度量和基准测试	589
11.1.3 高级系统结构选择	590
11.2 并行语言与编译技术	591
11.2.1 并发程序设计语言	596
11.2.2 向量化和迁移技术	598
11.2.3 多重处理与多任务	600
11.2.4 智能编译程序与伪指令	603
11.3 新系统结构与并发控制	606
11.3.1 处理机调度与活动控制	607
11.3.2 处理机间通信方法	608
11.3.3 存储器层次和访问方法	610
11.3.4 I/O 多重处理与前端机	614
11.3.5 进程迁移与负载平衡	615
11.4 程序设计和应用软件	618
11.4.1 对用户友好的程序设计环境	618
11.4.2 并发程序设计技术	619
11.4.3 算法向并行系统结构的映射	620
11.4.4 数学软件与应用程序包	624
11.5 超级计算机的未来	626
11.5.1 光计算机和神经元计算机	627

11.5.2 发展并行计算科学	629
11.6 结论	629
11.7 参考文献	630
参考文献	640
索引	659

第一章 并行处理引论

本章介绍在高性能计算机上并行处理的基本概念。我们将回顾系统结构的进展、探讨在现代计算机系统中并发活动的各种形式,并且估计并行处理计算机的一些尖端应用。并行计算机结构按特性可分为**流水线计算机**、**阵列处理机**和**多处理机系统**。本章还要介绍一些新的计算概念,包括数据流和 VLSI 方法。本章介绍的内容是全书涉及的各个领域的一个概况,目的是为详细研究以后各章中的并行计算理论、机器结构、系统控制、快速算法以及编程要求打下基础。

1.1 计算机系统的进展

在过去的 40 年间,计算机工业经历了四代的发展。在物理上的标志是组成部件的迅速变化:从继电器和真空管(40—50 年代)到分立的二极管和三极管(50—60 年代),再发展到小规模和中规模集成(SSI/MSI)电路(60—70 年代),最后到大规模和超大规模集成(LSI/VLSI)电路(70 年代以后)。器件速度和可靠性的增加,以及硬件价格和物理尺寸的减小大大增强了计算机的性能。然而,较好的器件不是产生高性能的唯一因素。自从有了冯·诺依曼的存储程序概念以来,就一直认为计算机不仅仅是一个硬件组织的问题。一个现代的计算机系统实际上是由处理器、存储器、功能部件、互连网络、编译程序、操作系统、外围设备、通信通道以及数据库等这些内容组合而成的。

要设计一个功能强、性能价格比高的计算机系统以及编制有效的程序来解决一个计算问题,人们必须了解基本的硬件和软件系统的结构,以及由面向用户的程序设计语言在这台机器上所实现的计算算法。这些学科构成了**计算机结构学**的技术范畴。计算机系统结构实际上是把硬件、软件、算法和语言合为一体来完成大型计算的一个系统概念。一个好的计算机设计师应精通所有这些学科。在过去的 40 年中,计算机性能的重大改善大多数是由于集成电路和系统结构的革命性进展。在这一节中,我们回顾计算机系统所经历的年代,并说明高性能计算机发展的普遍趋势。

1.1.1 计算机系统经历的各代

把计算机系统划分成几代是由于器件工艺技术、系统结构、处理模式以及使用的语言决定的。我们认为每一代大约经历 10 年的时间。如图 1.1 所示,相邻代之间可能有若干年的重叠。把时间划得较长的目的是为了包含在世界各地的计算机的研制和使用这两个方面。当前,我们正处于第四代,而第五代还没有实现。

1. 第一代(1938—1953) 1938 年第一台电子模拟计算机和 1946 年第一台电子数字计算机 ENIAC (Electronic Numerical Integrator and Computer) 标志着第一代计算机的开始。在 40 年代,用作开关器件的是电磁继电器;50 年代是真空管。这些器件由绝

缘导线互相连接。硬件部件很贵，从而迫使 CPU 的结构必须**按位串行**：在定点的基础上按位进行运算，如在行波进位加法中那样，只用一个全加器和一个一位的进位标志。

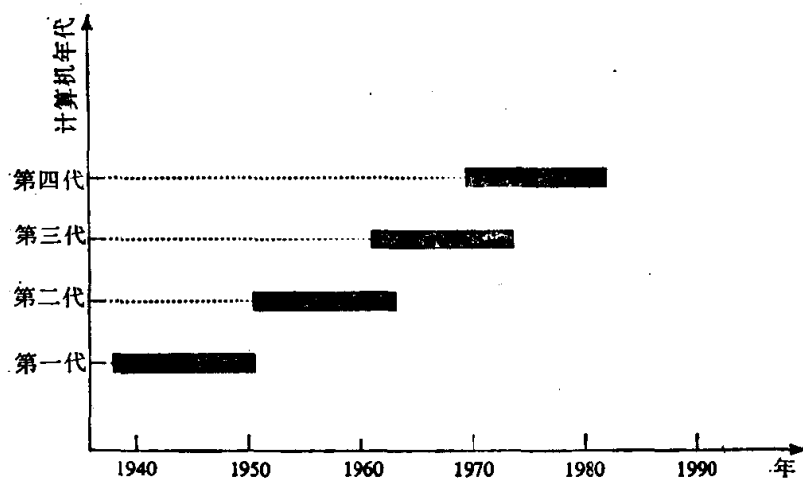


图 1.1 计算机系统的进展

在早期计算机中，只使用二进制码的机器语言。在 1950 年，诞生了第一台存储程序的计算机 EDVAC (Electronic Discrete Variable Automatic Computer)，这标志着使用系统软件的开始，从而把用户从低级程序设计的负担中解脱了出来。然而，不难想像在早期的

计算机中，硬件价格占着主导地位，而软件语言的性能都是相当原始的。到了 1952 年，IBM 推出了它的 701 电子计算机，这个系统使用的是威廉管 (Williams' tube) 存储器、磁鼓和磁带。

2. 第二代 (1952—1963) 1948 年发明了晶体管。第一台晶体管数字计算机 TRADIC (TRAnsistorized Digital Computer) 是在 1954 年由贝尔实验室制成的。组成元件是分立晶体管和二极管：在 TRADIC 中使用了 800 个晶体管。同时，出现了印制电路。这时，又研制出了电流重合磁芯存储器，并相继在许多机器中得到了使用。在 1956 年出现 Fortran (Formula Translation) 和 1960 年出现 Algol (Algorithmic Language) 这些高级语言之前，一直使用的是汇编语言。

1959 年，Sperry Rand 公司制造了 Larc 系统，IBM 公司也开始了它的 Stretch 计划。这是最初两台致力于改进系统结构的计算机。Larc 有一个独立的 I/O 处理器，它和一个或两个中央处理器并行运行。Stretch 的特点是指令先行和错误校正，这将在 1.2 节中讨论。第一台 IBM 科学计算用的晶体管计算机 IBM 1620，是在 1960 年诞生的。Cobol (Common Business Oriented Language) 出现于 1959 年。可互换磁盘组的引入是在 1963 年。批处理方式已经盛行，这种方式是顺序地执行用户程序，每次一个程序，直到全部执行完。

3. 第三代 (1962—1975) 这一代的标志是，基本组成部件由小规模集成 (SSI) 电路和中规模集成 (MSI) 电路构成。多层印制电路也已使用。虽然在 CDC-6600 和其它机器中还用磁芯存储器。但是，到 1968 年，许多快速计算机，如 CDC-7600，已经开始用固态存储器代替磁芯。在这期间，有了功能更强的编译程序，这大大改进了高级语言。

多道程序技术的发展使多个程序段可以同时和 I/O 操作交叉重叠运行。在 70 年代初研制出的高性能计算机有 IBM 360/91, Illiac IV, TI-ASC, Cyber-175, STAR-100 和 C.mmp，另外还有一些向量处理机。分时操作系统在 60 年代后期就开始进入实用阶段，还研制出了采用层次结构存储系统的虚拟存储器。

4. 第四代 (1972—至今) 这一代计算机在逻辑和存储器部分注意采用大规模集成 (LSI) 电路。这时出现了高密度组装。高级语言向既能处理标量，又能处理向量的方向