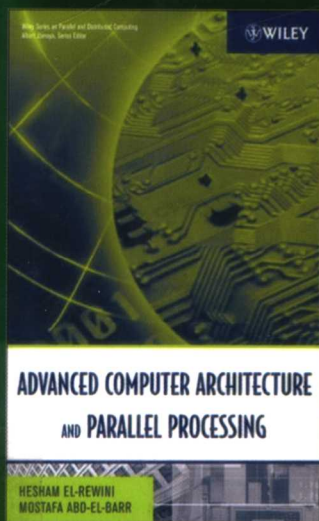


国外计算机科学教材系列



先进计算机体系结构 与并行处理

Advanced Computer Architecture and Parallel Processing



[美] Hesham El-Rewini 著
[科] Mostafa Abd-El-Barr

陆鑫达 林新华 翁楚良 译



电子工业出版社

Publishing House of Electronics Industry
<http://www.phei.com.cn>

国外计算机科学教材系列

先进计算机 体系结构与并行处理

Advanced Computer
Architecture and Parallel Processing

[美] Hesham El-Rewini 著
[科] Mostafa Abd-El-Barr
陆鑫达 林新华 翁楚良 译

Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书主要论述计算机体系结构和并行计算中的深层次问题,主要包括:先进计算机体系结构和并行处理导论、多处理器互连网络、多处理器体系结构的性能分析、共享存储器多处理器体系结构、消息传递多处理器体系结构、并行计算的抽象模型、网络计算、并行程序设计的环境和工具(包括并行虚拟机中的并行编程和消息传递接口),以及并行计算中的任务调度和分配问题。本书取材合理,侧重实例分析和解决实际问题的能力。

本书可作为计算机专业及电气工程专业大学高年级本科生和研究生的教材,对从业工程师和程序设计人员也是一本有用的参考书。

Hesham El-Rewini, Mostafa Abd-El-Barr: **Advanced Computer Architecture and Parallel Processing.**

ISBN 0-471-46740-5

Copyright © 2005, John Wiley & Sons, Inc.

All Rights Reserved. Authorized translation from the English language edition published by Wiley Publishing, Inc.

No part of this book may be reproduced in any form without the written permission of Wiley Publishing, Inc.

Simplified Chinese translation edition Copyright © 2005 by Publishing House of Electronics Industry.

本书中文简体字翻译版由 Wiley Publishing, Inc 授予电子工业出版社。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号 图字: 01-2004-5756

图书在版编目(CIP)数据

先进计算机体系结构与并行处理 / (美) 莱维尼 (Rewini, H. E.) 等著; 陆鑫达等译.

北京: 电子工业出版社, 2005.12

(国外计算机科学教材系列)

书名原文: Advanced Computer Architecture and Parallel Processing

ISBN 7-121-01988-4

I. 先... II. ①莱... ②陆... III. ①计算机体系结构-教材 ②并行处理-教材 IV. ①TP303 ②TP274

中国版本图书馆CIP数据核字(2005)第138813号

责任编辑: 贺瑞君

印 刷: 北京牛山世兴印刷厂

出版发行: 电子工业出版社

北京市海淀区万寿路173信箱 邮编: 100036

经 销: 各地新华书店

开 本: 787 × 1092 1/16 印张: 12.75 字数: 359千字

印 次: 2005年12月第1次印刷

定 价: 22.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换;若书店售缺,请与本社发行部联系。联系电话:(010)68279077。质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

出版说明

21世纪初的5至10年是我国国民经济和社会发展的关键时期,也是信息产业快速发展的关键时期。在我国加入WTO后的今天,培养一支适应国际化竞争的一流IT人才队伍是我国高等教育的重要任务之一。信息科学和技术方面人才的优劣与多寡,是我国面对国际竞争时成败的关键因素。

当前,正值我国高等教育特别是信息科学领域的教育调整、变革的重大时期,为使我国教育体制与国际化接轨,有条件的高等院校正在为某些信息学科和技术课程使用国外优秀教材和优秀原版教材,以使我国在计算机教学上尽快赶上国际先进水平。

电子工业出版社秉承多年来引进国外优秀图书的经验,翻译出版了“国外计算机科学教材系列”丛书,这套教材覆盖学科范围广、领域宽、层次多,既有本科专业课程教材,也有研究生课程教材,以适应不同院系、不同专业、不同层次的师生对教材的需求,广大师生可自由选择 and 自由组合使用。这些教材涉及的学科方向包括网络与通信、操作系统、计算机组织与结构、算法与数据结构、数据库与信息处理、编程语言、图形图像与多媒体、软件工程等。同时,我们也适当引进了一些优秀英文原版教材,本着翻译版本和英文原版并重的原则,对重点图书既提供英文原版又提供相应的翻译版本。

在图书选题上,我们大都选择国外著名出版公司出版的高校教材,如Pearson Education培生教育出版集团、麦格劳-希尔教育出版集团、麻省理工学院出版社、剑桥大学出版社等。撰写教材的许多作者都是蜚声世界的教授、学者,如道格拉斯·科默(Douglas E. Comer)、威廉·斯托林斯(William Stallings)、哈维·戴特尔(Harvey M. Deitel)、尤利斯·布莱克(Uyless Black)等。

为确保教材的选题质量和翻译质量,我们约请了清华大学、北京大学、北京航空航天大学、复旦大学、上海交通大学、南京大学、浙江大学、哈尔滨工业大学、华中科技大学、西安交通大学、国防科学技术大学、解放军理工大学等著名高校的教授和骨干教师参与了本系列教材的选题、翻译和审校工作。他们中既有讲授同类教材的骨干教师、博士,也有积累了几十年教学经验的老教授和博士生导师。

在该系列教材的选题、翻译和编辑加工过程中,为提高教材质量,我们做了大量细致的工作,包括对所选教材进行全面论证;选择编辑时力求达到专业对口;对排版、印制质量进行严格把关。对于英文教材中出现的错误,我们通过与作者联络和网上下载勘误表等方式,逐一进行了修订。

此外,我们还将与国外著名出版公司合作,提供一些教材的教学支持资料,希望能为授课老师提供帮助。今后,我们将继续加强与各高校教师的密切联系,为广大师生引进更多的国外优秀教材和参考书,为我国计算机科学教学体系与国际教学体系的接轨做出努力。

电子工业出版社

教材出版委员会

- | | | |
|----|-----|---|
| 主任 | 杨芙清 | 北京大学教授
中国科学院院士
北京大学信息与工程学部主任
北京大学软件工程研究所所长 |
| 委员 | 王 珊 | 中国人民大学信息学院院长、教授 |
| | 胡道元 | 清华大学计算机科学与技术系教授
国际信息处理联合会通信系统中国代表 |
| | 钟玉琢 | 清华大学计算机科学与技术系教授
中国计算机学会多媒体专业委员会主任 |
| | 谢希仁 | 中国人民解放军理工大学教授
全军网络技术研究中心主任、博士生导师 |
| | 尤晋元 | 上海交通大学计算机科学与工程系教授
上海分布计算技术中心主任 |
| | 施伯乐 | 上海国际数据库研究中心主任、复旦大学教授
中国计算机学会常务理事、上海市计算机学会理事长 |
| | 邹 鹏 | 国防科学技术大学计算机学院教授、博士生导师
教育部计算机基础课程教学指导委员会副主任委员 |
| | 张昆藏 | 青岛大学信息工程学院教授 |

译 者 序

本书是一本有关先进计算机体系结构与并行处理的教科书。在计算机体系结构的基础上,对并行计算中的先进计算机体系结构进行了更深层次的讨论。本书的主要内容包括:并行计算中的共享存储、消息传递体系结构及它们的互连网络,多处理器体系结构及其性能分析,并行计算的抽象模型,网络计算,并行程序设计环境和工具,以及多处理机系统中的任务调度和分配。内容涉及到当前并行计算中最关键的问题。本书的特点之一是从硬件和软件两方面来探讨有关的关键问题,并注重它们之间的交互影响,从而使读者能更好地掌握有关问题的真谛所在。本书的另一特点是注重实际举例和对先进而又成熟的技术进行介绍。本书的第三个特点是取材较为先进。

本书章节组合逻辑性强,能使读者较快地掌握有关的专业知识点。此外,本书的每一章都有小结以及适量的习题,有助于读者加深对各章节主要概念的理解和掌握。

本书适合作为计算机科学、计算机工程和电气工程本科高年级学生的计算机专业课教材,也可作为研究生的教材。对已就业的工程师和程序员而言,本书也是一本极有价值的参考书。本书的翻译工作由陆鑫达教授主持,并翻译了目录、前言、第1~4章以及索引。林新华助教翻译了第5~7章,翁楚良讲师翻译了第8~10章。译稿全文由陆鑫达教授进行了统一和校对。由于翻译时间较紧,译文中若有错误或不妥之处,敬请广大读者不吝指正。

最后,要感谢电子工业出版社的有关策划和编辑人员,正是他们的大力支持才使本书得以顺利出版。

前 言

单处理器超级计算机已经达到极大速度,并已将硬件工艺推向芯片制造的物理极限。由于物理和体系结构的约束,单处理器系统能够达到的最大计算能力将受到限制,因而这一发展趋势将很快走到尽头。在本书中,我们将研究利用多处理部件并行性的先进计算机体系结构。内部互连处理器的并行计算成为利用并行性主要形式的同时,计算机网络中的进展已经带来了一种连网自治计算机的新型并行性,与将所有部件放在一个盒子且紧耦合处理器与存储器,以获取并行性的方式不同,因特网通过在盒子外松耦合所有部件来获取并行性。为从一个计算机系统采掘最多的内部和外部并行性,设计者和软件开发者必须了解计算机系统中硬件和软件间的交互。这就是我们撰写本书的原因。我们要求读者能知晓多处理机系统的威力和局限性。我们的目的是要告诉读者先进的体系结构和并行性具有优越性,也带来了不小的挑战,具有双重性。本书的内容被组织成如下的十章。

第1章以导论方式综述计算机体系结构研究领域。我们首先讨论计算的进展,以及通过并行化而获得高性的能计算的演变。接下来我们介绍比较流行的计算机系统的弗林分类法,此外还对单指令多数据(SIMD)和多指令多数据(MIMD)系统进行介绍。本章还将叙述共享存储器和消息传递系统,以及它们的互连网络。

第2章通过对一些多处理器的系统构造进行叙述,来讨论用来互连多处理器的不同拓扑。介绍基于它们拓扑的互连网络的分类、研究动态和静态互连方式。总线、交叉开关和多级拓扑将作为动态互连进行介绍。而在静态互连方式方面将介绍三种主要机制,它们是超立方体拓扑、网格拓扑以及 k 元 n 立方体拓扑。本章还将介绍一些有关的性能指标,包括成本、时延、直径、结点度和对称性。

第3章介绍性能。当在事实上并行计算重新定义传统的性能衡量标准时,如每秒百万次指令(MIPS)和每秒百万次浮点运算(MFLOPS),我们应如何刻画一个计算机系统的性能?为此我们将讨论用加速作为新的衡量性能的方法。在这一章中,我们将讨论几种加速的版本,以及其他的性能衡量方法和基准测试程序。

第4章和第5章分别讨论共享存储器和消息传递系统。共享存储器系统的挑战在于因竞争和高速缓存一致性问题而导致的系统性能衰退。当连接处理器到全局存储器的互连网络成为瓶颈时,共享存储器的性能将成为问题。通常使用本地高速缓存来缓和这一瓶颈问题。但可扩展性仍是该系统的主要缺点。高速缓存的引入导致了高速缓存间的一致性问题以及存储器和高速缓存之间的一致性问题。在第4章中,我们将介绍几种高速缓存一致性协议,这些协议可被分类成监听协议或基于目录的协议。由于共享存储器系统难于扩展以支持具有大量处理器的系统,消息传递系统成为惟一能有效获取可扩展性的方法。在第5章中我们将讨论消息传递系统的体系结构和网络模型。我们将介绍路由和网络交换技术。第5章将以共享存储器和消息传递系统之间的对比结束。

第6章将涉及抽象模型、算法和复杂性分析。我们将讨论一个共享存储器模型(PRAM),该模型可用来研究并行算法并评估它们的复杂性。我们也将简述同步模型下的消息传递系统形

式化模型的基本成分。我们将设计和讨论上述两个模型中的算法复杂性分析。

第7章到第10章将讨论一些有关网络计算的问题,在网络计算中每个结点是一台独立的计算机,它们可以通过交换机、局域网或因特网连接起来。第7章将介绍网络计算的基本概念,包括客户/服务器模式、机群计算和网格计算。第8章阐述并行虚拟机(PVM)编程,展示如何在异构机的网络上编写程序。第9章涉及消息传递接口(MPI)标准,按此标准就可开发可移植的分布式并行程序。第10章论及向处理部件分配任务的问题。对调度问题将涉及若干种它的变体。我们将综述这些重要问题的解决方法。我们将论及程序和系统的模型、优化算法、启发式算法、调度及分配技术,以及同构和异构环境。

计算机工程、计算机科学以及电气工程的大学生将受益于本书。本书也可作为高级计算机体系结构和并行处理的研究生课程教材。各章节可灵活抽取搭配用做不同侧重点的专题课程教材。本书也可作为从业者(工程师、程序员和技术人员)的综合参考书。此外,本书的一部分也可用做从业者培训教材。不同的章节可用来满足不同的需求。例如,第1章到第5章和第7章、第8章可作为高等计算机体系结构的一学期课程,而并行处理的课程可选择第1章到第4章以及第6章、第9章、第10章。

本书业已经过两位作者的课堂教学实践。事实上本书是从SMU的CSE8380和CSE8383的教学笔记以及Saskatchewan大学CMPT740和KFUPM的COE520的教学笔记演化而来的。这些教学经验已融合进本书中。我们的学生纠正了不少错误,并改进了本书的组成,我们谨向这些班级的学生表示感谢。我们要感谢许多学生和同事,他们为本书的出版做出了贡献。Chuck Mann, Yehia Amer, Habib Ammari, Abdul Aziz, Clay Breshears, Jahanzeb Faizan, Michael A. Langston 和 A. Naseer 阅读了本书的草稿,他们均对原稿的改进做出了贡献。Ted Lewis 对某些章节的早期版本有很大贡献。我们要感谢 John Wiley 出版社安排的评审者的建议和校正。我们谨向丛书编辑 Albert Y. Zomaya 以及 John Wiley 出版社的 Val Moliere, Kirsten Rohstedt 和 Christine Punzo 表示衷心的感谢,正是他们的帮助才使本书得以出版。当然,错误和不一致的责任将由我们承担。

最后,也是最重要的,要向我们的妻子和孩子们表示感谢,是他们容忍我们长时间地专注于本书的写作。Hesham 还感谢 Ted Lewis 和 Bruce Shriver 多年来的友谊、关爱和指导。

HESHAM EL-REWINI
MOSTAFA ABD-EL BARR
2004年5月

目 录

第 1 章 先进计算机体系结构与并行处理导论	1
1.1 计算的四个年代	1
1.2 计算机体系结构的弗林分类方法	3
1.3 SIMD 体系结构	4
1.4 MIMD 体系结构	5
1.5 互连网络	8
1.6 小结	11
习题	11
参考文献和推荐读物	13
第 2 章 多处理器互连网络	15
2.1 互连网络的分类	15
2.2 基于总线的动态互连网络	16
2.3 基于交换的互连网络	19
2.4 静态互连网络	26
2.5 分析和性能指标	31
2.6 小结	34
习题	34
参考文献和推荐读物	35
第 3 章 多处理器体系结构的性能分析	37
3.1 计算模型	37
3.2 关于并行体系结构的争论	39
3.3 互连网络的性能问题	41
3.4 并行体系结构的可扩展性	45
3.5 基准测试程序性能	48
3.6 小结	52
习题	52
参考文献和推荐读物	53
第 4 章 共享存储器体系结构	56
4.1 共享存储器系统的分类	56
4.2 基于总线的对称多处理机系统	58
4.3 基本的高速缓存一致性方法	59
4.4 监听协议	60

4.5 基于目录的协议	65
4.6 共享存储器的编程	69
4.7 小结	71
习题	72
参考文献和推荐读物	73
第 5 章 消息传递体系结构	74
5.1 消息传递简介	74
5.2 消息传递网络中的路由	75
5.3 消息传递中的交换机制	78
5.4 消息传递编程模型	81
5.5 支持消息传递的处理器	83
5.6 消息传递体系结构实例	84
5.7 消息传递体系结构与共享存储器体系结构	87
5.8 小结	88
习题	88
参考文献和推荐读物	89
参考网站	90
第 6 章 抽象模型	91
6.1 PRAM 模型及其变体	91
6.2 在 EREW PRAM 上模拟多个访问	92
6.3 并行算法的分析	93
6.4 计算和与所有和	95
6.5 矩阵乘法	98
6.6 排序	100
6.7 消息传递模型	101
6.8 领导者选举问题	105
6.9 同步环中的领导者选举	106
6.10 小结	111
习题	111
参考文献和推荐读物	111
第 7 章 网络计算	113
7.1 计算机网络基础	113
7.2 客户/服务器系统	115
7.3 机群	119
7.4 互连网络	122
7.5 机群实例	126
7.6 网格计算	127
7.7 小结	128
习题	128

参考文献和推荐读物	129
参考网站	129
第 8 章 并行虚拟机的并行编程	130
8.1 PVM 环境和应用结构	130
8.2 创建任务	133
8.3 任务组	135
8.4 任务间的通信	136
8.5 任务同步	141
8.6 归约操作	142
8.7 工作分配	143
8.8 小结	145
习题	145
参考文献和推荐读物	146
参考网站	146
第 9 章 消息传递接口	147
9.1 通信	147
9.2 虚拟拓扑	150
9.3 任务通信	152
9.4 同步	155
9.5 集合操作	157
9.6 任务创建	161
9.7 单边通信	163
9.8 小结	165
习题	165
参考文献和推荐读物	167
第 10 章 调度与任务分配	168
10.1 调度问题	168
10.2 未考虑通信的 DAG 调度	170
10.3 通信模型	173
10.4 考虑通信的 DAG 调度	174
10.5 调度问题的 NP 完全性	177
10.6 启发式算法	178
10.7 任务分配	183
10.8 异构环境下的调度	187
10.9 小结	187
习题	188
参考文献	188

第 1 章 先进计算机体系结构与并行处理导论

计算机体系结构学者总是不断努力去增加他们所设计的计算机体系结构的性能。高性能可以来自快速高密度电路、封装技术和并行性。单处理器超级计算机已经达到前所未有的速度,并一直在将硬件工艺推向芯片制造的物理极限。本书中我们将研究利用多处理部件所提供并行性的先进计算机体系结构。

并行处理器是由借助某种互连网络而连接起来的多个处理部件,以及使这些部件能在一起工作的软件所组成的计算机系统。对这种计算机系统的分类主要取决于以下两个因素:处理部件本身,以及将它们结合在一起的互连网络。这些处理部件可使用共享存储器或消息传递方法进行通信和交互。共享存储器系统中的互连网络可被分为基于总线的和基于交换机的。在消息传递系统中,互连网络则被分为静态的和动态的。静态的连接有一个固定的拓扑,在程序运行期间它不会发生改变;动态的连接则在程序执行过程中才创建连接。

使用多处理器的主要出发点是只需简单地连接多个处理器就可创建功能强大的计算机。人们期待多处理机系统可以达到比最快的单处理器系统更快的速度。此外,可以期待由许多单处理器组成的多处理机系统比建造一台高性能单处理器系统更具性价比(cost-effective)。多处理机系统的另一个优点是具有容错能力。若一台处理器失效,其余的处理器仍能继续提供性能降级的服务。

1.1 计算的四个年代

大多数计算机科学工作者认同存在四个不同的计算模式或时期,即批处理、分时、台式机和网络。表 1.1 是 Lawrence Tesler 所提出的表的修改版本。该表中列出了自 20 世纪 60 年代开始与每个计算年代相关的不同计算模式的主要特征。

表 1.1 计算的四个年代

特征	批处理	分时	台式机	网络
年代	20 世纪 60 年代	20 世纪 70 年代	20 世纪 80 年代	20 世纪 90 年代
场所	计算机房	终端室	桌面	移动
用户	专家	专业人员	个人	小组
数据	字母,数字	文本,数字	字型,图	多媒体
对象	计算	访问	显示	通信
接口	穿孔卡片	键盘和 CRT	观看和点击	询问和告知
操作	处理	编辑	布局	协调
网络连接	无	外设电缆	局域网	因特网
拥有者	公司计算中心	分区的 IS 商店	部门终端用户	几乎所有人

1.1.1 批处理年代

直到 1965 年,IBM System/360 主机一直占据着公司计算机中心的市场。这是极具代表性的

批处理机器,带有穿孔卡阅读机、磁带以及磁盘驱动器,所有的连接均在一个计算机房内。这种以单个主机构成的大型集中式计算机作为一种标准的计算形式存在了几十年。IBM System/360有一个操作系统、多种编程语言及10 MB磁盘存储器。System/360的金属机柜和运行系统的人员充塞了整个机房。它的晶体管电路相当快。机器使用者最多可预订32位字的1 MB磁心存储器容量。该机器大到足以在存储器中同时支持许多程序,但中央处理器不得不在运行时从一个程序切换到另一个程序。

1.1.2 分时年代

批处理年代的大型机在20世纪60年代后期具有不可动摇的地位,在那时半导体技术的进展已使得固态存储器和集成电路变得可行。硬件技术的这些进展导致了小型计算机时代的来临。它们的小体积、快速和相对不很贵的价格,使它们能在公司的部门级单位大有用武之地。但是,要将它们交付给终端用户,则仍显得过于昂贵和难于使用。由DEC、Prime和Data General公司制造的小型计算机导致了新一类计算的出现:分时系统。在20世纪70年代很显然存在两种类型的商务及商业计算:(1)集中式数据处理主机,(2)分时小型计算机。与小型机同时期出现的是超级计算机。由Control Data公司推出的CDC 6600是第一台此类超级计算机,而Cray研究公司则在1976年推出了具有最好性价比的超级计算机Cray-1。

1.1.3 台式机年代

个人计算机(PC)是由Altair、Processor Technology、North Star、Tandy、Commodore、Apple以及许多其他公司于1977年推出的,PC机在许多部门增强了终端用户的生产率。由Compaq、Apple、IBM、Dell以及许多其他公司生产的个人计算机立刻普及开来,并改变了计算的面貌。

约在1990年,功能强大的个人计算机和 workstation 局域网(LAN)开始替代大型机和小型计算机。曾经具有最强计算能力的大型机,现在可用台式机的形式以其十分之一的价格获得相同的性能。然而这些单个的台式机不久就被广域网(WAN)连接成更大的计算联合体。

1.1.4 网络年代

第四个年代——网络计算年代——由于网络技术的飞速发展而非常活跃。在20世纪90年代的大部分时间内,网络技术的发展速度超过了处理器技术的发展速度。这就说明了表1.1中所列出的网络模式崛起的原因。网络的蓬勃发展使以处理器为中心的观点向以网络为中心的观点倾斜。

20世纪80年代和20世纪90年代面世了许多具有多处理器的商用并行计算机。它们一般可分为两大类:(1)共享存储器系统,(2)分布式存储器系统。在单机中所包含的处理器数从在共享存储器计算机中的几个增长到在大规模并行系统中的几百到几千个。这一时代中的并行机例子包括Sequent Symmetry、Intel iPSC、nCUBE、Intel Paragon、Thinking Machines (CM-2, CM-5)、MsPar (MP)、Fujitsu (VPP500)以及其他机器。

1.1.5 当前发展趋势

计算发展的一个明显趋向是昂贵和专用的并行机将被更经济有效的工作站机群所替代。一个机群是用某种互连网络连接众多独立计算机的集合。此外,因特网的普及将使网络计算以

及近期的网格计算变得更加引人入胜。网格是地域分布式计算平台,它们将能提供可信赖的、一致的以及廉价的对高端计算设施的访问。

1.2 计算机体系结构的弗林分类方法

最流行的计算机体系结构分类方法是由弗林在 1976 年定义的。弗林的分类方法基于信息流的概念。处理器中存在两种类型的信息流:指令和数据。指令流被定义为由处理部件所完成的指令序列。数据流被定义为在存储器和处理部件间的数据通信。按照弗林的分类,指令流或数据流可以是单个的也可以是多个的。由此,计算机体系结构可分成为如下四种不同的类型:

- 单指令单数据流(SISD)。
- 单指令多数据流(SIMD)。
- 多指令单数据流(MISD)。
- 多指令多数据流(MIMD)。

传统的单处理器冯·诺依曼计算机被归为 SISD 系统。并行计算机可以归为 SIMD 或 MIMD 系统。当并行机中只有一个控制部件且所有处理器以同步方式执行相同指令时,就被归类为 SIMD。在 MIMD 机器中,每个处理器有自己的控制部件且能在不同的数据上执行不同的指令。在 MISD 类型中,相同的数据流流过执行不同指令的一个线性处理器阵列。实际中没有可行的 MISD 机;但是某些作者认为流水机(以及脉动阵列计算机)可以作为 MISD 的例子。图 1.1、图 1.2 和图 1.3 各自画出了 SISD、SIMD 和 MIMD 的结构图。

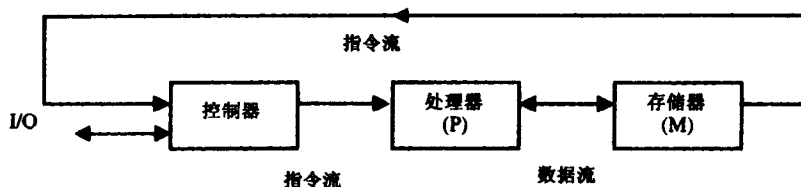


图 1.1 SISD 体系结构

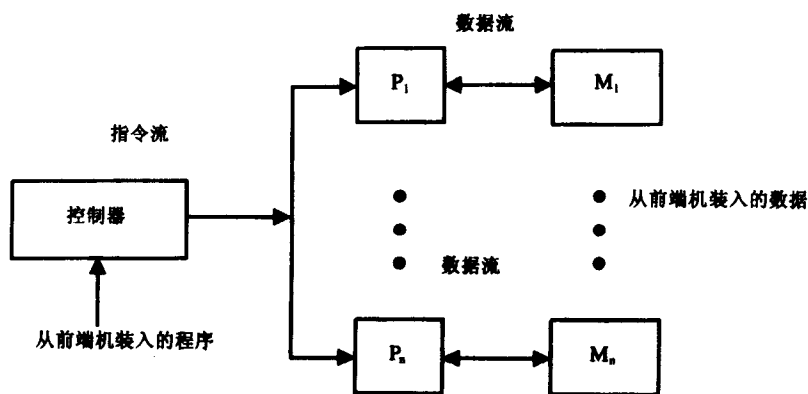


图 1.2 SIMD 体系结构

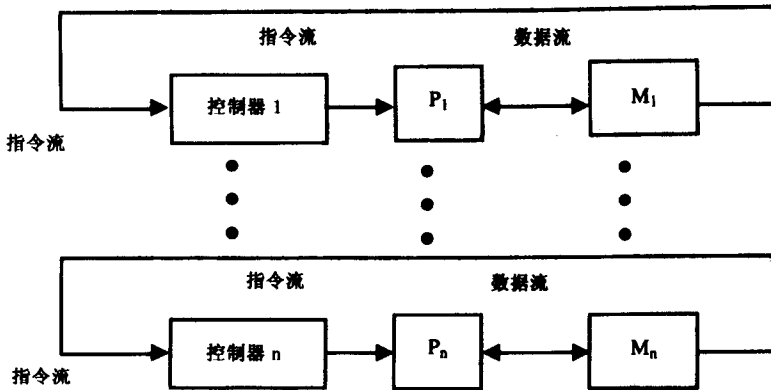


图 1.3 MIMD 体系结构

弗林分类法的一个扩展是由 D. J. Kuck 于 1978 年提出的。在 Kuck 的分类中,指令流进一步扩展成单(标量和数组)流和多(标量和数组)流,并将数据流称为执行流,而且也扩展成包括单(标量和数组)流和多(标量和数组)流。这些流的组合形成了总共 16 类的体系结构。

1.3 SIMD 体系结构

并行计算的 SIMD 模型由两部分组成:一个具有常见的冯·诺依曼风格的前端计算机和一个处理器阵列,如图 1.4 所示。处理器阵列是一组相同的同步处理单元,它们能够在不同的数据上同时完成相同的操作。阵列中的每个处理器有一个小容量的局部存储器,分散的数据驻留在其上,它们将被并行处理。处理器阵列连接到前端机的存储器总线,这样前端机就能随机地访问处理器阵列中每个处理器的局部存储器,就好像这些局部存储器是它的另一个存储器。因此前端机能发出特定命令以使部分存储器同时操作或使数据在存储器中移动。可以用传统的顺序编程语言来开发程序,并在前端机上执行。前端机通常按串行方式执行应用程序,但前端机可向处理器阵列发出命令让它并行执行 SIMD 操作。这种串行和数据并行编程之间的类似性是数据并行性的优点之一。独立同步采用处理器间的锁步(lock-step)同步来实现,处理器要么什么都不做,要么同时做相同的操作。在 SIMD 体系结构中,借助在大数据集上同时进行操作来开发并行性。这一模式在求解需要大规模更新许多数据的问题时最为有用。在许多规则的数值计算中,这种模式特别有效。

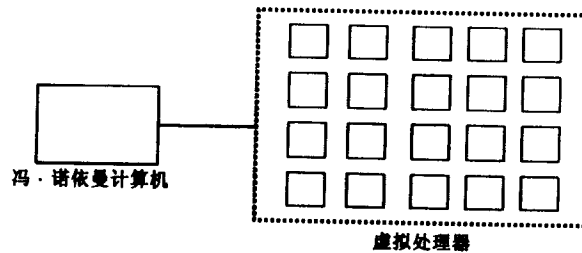


图 1.4 SIMD 体系结构模型

SIMD 机器中使用了两种主要的配置结构(参见图 1.5)。在第一种配置结构中,每个处理器都有自己的局部存储器。处理器可以通过互连网络相互进行通信。若互连网络在一对给定的

处理器间未提供直接连接,则这对处理器可通过一个中间处理器进行数据交换。ILLIAC IV 就使用这种互连方案。ILLIAC IV 中的互连网络允许每一个处理器在一个 8×8 的矩阵中直接与 4 个相邻的处理器通信,也就是说第 i 个处理器能直接与第 $(i-1)$ 个、第 $(i+1)$ 个、第 $(i-8)$ 个和第 $(i+8)$ 个处理器进行通信。在第二种 SIMD 配置结构中,处理器和存储器模块间的通信是通过互连网络进行的。两个处理器可通过中间的存储器模块进行相互通信,也可通过中间的处理器进行。BSP(Burroughs 公司的科学处理机)使用 SIMD 的第二种配置结构。

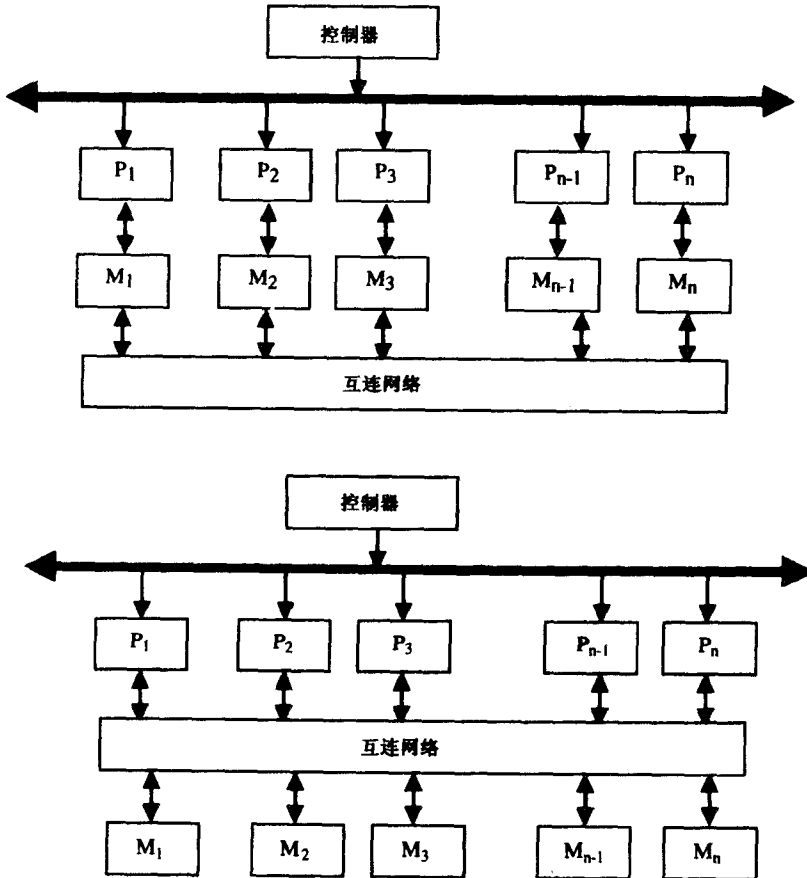


图 1.5 两种 SIMD 方案

1.4 MIMD 体系结构

多指令多数据流(MIMD)并行体系结构是由多处理器和多存储器模块借助某种互连网络连接在一起而构成的。它们被分为两大类:共享存储器或消息传递。图 1.6 说明了这两类的通用体系结构。在共享存储器系统中,处理器通过它们的中央共享存储器交换信息;而在消息传递系统中,信息的交换是通过它们的互连网络完成的。

共享存储器系统通常通过为所有处理器所共享的一个全局存储器来完成处理器间的协调。典型的服务器系统中通信是通过总线和高速缓存控制器进行的。总线/高速缓存体系结构缓和了对昂贵的多端口存储器和接口电路的需求,以及当开发应用软件时采用消息传递模式的需

求。因为对共享存储器的访问是平衡的,故这类系统也称为 SMP(对称多处理器)系统。每个处理器读/写存储器的机会相等,有相同的访问速度。SMP 商用机的例子有 Sequent Computer 公司的 Balance 和 Symmetry, Sun Microsystem 公司的多处理器服务器,以及 Silicon Graphics 公司的多处理器服务器。

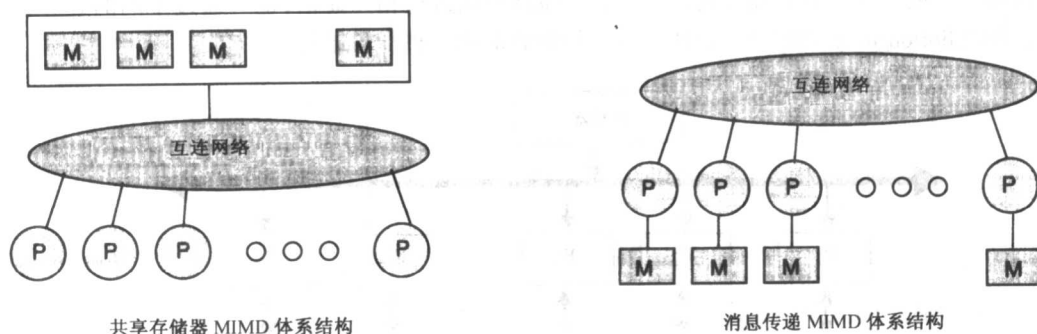


图 1.6 共享存储器与消息传递体系结构

消息传递系统(也称为分布式存储器)通常将局部存储器和处理器组合在互连网络的每个结点中。在这种系统中没有全局的存储器,所以必须借助消息传递将数据从一个局部存储器移动到另一个。通常用一对发送/接收命令来完成数据移动,但必须由程序员编写在应用程序中。为此,程序员必须学习消息传递模式,包括数据的复制以及处理一致性的事宜。消息传递体系结构商用机的例子在 1990 年时有 nCUBE、iPSC/2 以及各种基于 Transputer 的系统。这些系统最终都被因特网连接的系统替代了,其中的处理器/存储器结点可以是因特网服务器也可以是个人台式机客户机。

显然,分布式存储器是惟一能有效增加一个并行和分布式系统所管理处理器数目的方法。若可扩展性的继续发展导致越来越大的系统(由处理器的数目加以衡量),则系统不得不使用分布式存储器技术。这两股势力造成了以下矛盾:在共享存储器模型中编程较为容易,而在消息传递模型中所设计的系统提供了可扩展性。分布式共享存储器(DSM, Distributed-Shared Memory)体系结构开始出现在 SGI Origin2000 和其他的系统中。在这类系统中,存储器在物理上是分布的。例如,硬件体系结构追随消息传递学派的设计,而编程模型则追随共享存储器学派的思想。实际上,是软件在一定程度上封装了硬件。对程序员来讲,体系结构表现为一个共享存储器机器,但在软件下面存在一个消息传递体系结构。因此,DSM 机是利用了前面两种设计学派长处的混合系统。

1.4.1 共享存储器组成

在共享存储器模型中,处理器间的通信是通过读、写所有处理器平等访问的一个共享存储器中的单元来完成的。每个处理器可以有寄存器、缓冲器、高速缓存和局部存储器组作为附加的存储器资源。在设计共享存储器系统时必须考虑一些基本问题,其中包括访问控制、同步、保护和安全。访问控制决定了哪个进程可以访问哪些资源。访问控制模型将根据访问控制表的内容对每一个由处理器向存储器发出的访问请求加以必要的检查。访问控制表中含有决定每一次访问企图合法性的标志。若存在访问资源的企图,则直到所希望的访问完成之前,将阻塞所有不被允许访问的企图,以及不合法的进程。来自共享进程的请求在执行过程中可以改变