并行计算系列丛书

并行计算机体系结构

陈国良　吴俊敏　章　锋　章隆兵　编著

高等教育出版社
HIGHER EDUCATION PRESS

国内著名计算机专家，中国科学技术大学教授

陈国良　主编

并行计算系列丛书

□	并行计算——结构·算法·编程	陈国良　编著
□	并行算法的设计与分析（修订版）	陈国良　编著
■	并行计算机体系结构	陈国良　等编著
□	并行算法实践	陈国良　等编著

本套系列丛书中的《并行算法的设计与分析》、《并行计算机体系结构》和《并行算法实践》构成了并行计算三部曲，而《并行计算——结构·算法·编程》为三部曲之序曲。

- 《并行计算——结构·算法·编程》以并行计算为主题，主要讨论了并行计算的硬件平台（并行计算机）、并行计算的理论基础（并行算法）和并行计算的软件支撑（并行程序设计），强调融并行计算机结构、并行算法设计和并行编程为一体，书中内容具有相当的广度。
- 《并行算法的设计与分析》以并行计算模型为主线，系统深入地讨论了计算机科学中诸多常用的数值和非数值计算问题的并行算法设计和分析方法，同时也力图反映本学科的最新成就和发展趋势，书中内容具有相当的深度。
- 《并行计算机体系结构》以当代可扩放并行计算机系统结构为主题，着重讨论对称多处理机、大规模并行处理机、机群系统和分布共享存储多处理机系统的组成原理、结构特性、设计方法、性能分析以及相应系统实例，书中内容强调了软件与硬件相结合。
- 《并行算法实践》以基于消息传递的MPI并行编程语言为主，详细介绍了典型的数值并行算法和非数值并行算法的MPI具体编程实现过程，以及并行编程环境和开发工具的综合运用，书中内容体现了设计与实现相结合。

ISBN 7-04-011558-1

9 787040 115581 >

定价 35.00 元

并行计算系列丛书

并行计算机体系结构

陈国良　吴俊敏　章　锋　章隆兵 编著

高等教育出版社

图书在版编目（CIP）数据

并行计算机体系结构/陈国良等编著. —北京：高等教育出版社，2002. 9

ISBN 7-04-011558-1

Ⅰ. 并… Ⅱ. 陈… Ⅲ. 并行计算机-计算机体系结构 Ⅳ. TP338.6

中国版本图书馆 CIP 数据核字（2002）第 063438 号

并行计算机体系结构

陈国良 吴俊敏 章 锋 章隆兵 编著

出版发行	高等教育出版社	**购书热线**	010-64054588
社 址	北京市东城区沙滩后街 55 号	**免费咨询**	800-810-0598
邮政编码	100009	**网 址**	http://www.hep.edu.cn
传 真	010-64014048		http://www.hep.com.cn
经 销	新华书店北京发行所		
印 刷	人民教育出版社印刷厂		
开 本	787×1092 1/16	**版 次**	2002 年 9 月第 1 版
印 张	30.25	**印 次**	2002 年 9 月第 1 次印刷
字 数	570 000	**定 价**	35.00 元

本书如有缺页、倒页、脱页等 门联系调换。

0761024

作者介绍

陈国良，中国科学技术大学教授，博士生导师，1938 年 6 月生于安徽省颍上县，1961 年毕业于西安交通大学无线电系计算机专业。1981—1983 年在美国普度大学作访问学者，1984 年至今曾多次应邀赴东京大学、普度大学、澳大利亚国立大学、新南威尔士大学、昆士兰大学、格里福斯大学、堪萨斯城市大学、依阿华大学、威斯康星大学、Maharish 国际大学、香港理工大学、澳门大学、北京大学、国防科技大学等讲学交流。现任国家高性能计算中心（合肥）主任，国际高性能计算（亚洲）常务理事，全国高等教育电子、电工和信息类专业自考指导委员会副主任，中国计算机学会开放系统专业委员会副主任，中国数学会计算数学并行计算专业委员会委员，全国自然科学名词审定委员会委员。曾任国家教育部高等学校计算机科学与技术教学指导委员会副主任，安徽省高校计算机基础课程教学指导委员会副主任，中国计算机学会理事，安徽省计算机学会理事长，中国科学技术大学计算机系主任。

陈国良教授长期从事计算机科学技术的研究与教学工作。主要研究领域为并行算法、并行计算机体系结构和智能计算等。先后承担 10 多项国家 863 计划、国家攀登计划、国家自然科学基金、国家 973 计划、教育部博士点基金等科研项目。取得了多项被国内外广泛引用、达国际先进水平的科研成果，发表论文 150 多篇，出版著作 8 部、译著 5 部，参与主编计算机类辞典、词汇 5 部，主审、主编计算机类各种教材 8 部。曾获国家科技进步二等奖、国家级教学成果二等奖、国家教育部科技进步一等奖、中国科学院科技进步二等奖和自然科学三等奖、全国优秀教材一等奖、全国学术著作优秀奖、安徽省科技进步二等奖、国家科委高技术研究与发展计划三等奖、国家教委科技进步三等奖共 16 项，并获 2001 年度“国家 863 计划 15 周年先进个人奖”。

陈国良教授在中国科学技术大学执教 30 年。长期以来，围绕着并行算法的教学与研究，逐渐形成了一套完整的“**算法理论—算法设计—算法实现—算法应用**”的并行算法学科体系，营造了我国并行算法类的教学基地。他先后指导培养研究生 100 多名，其中博士生 60 名，为我国培养了一批在国内外从事算法研究的高级人才。曾荣获安徽省教育系统劳动模范、安徽省优秀教师称号和 2001 年度宝钢教育基金优秀教师特等奖。

陈国良教授是我国非数值并行算法研究的学科带头人。他率先创建的我国第一个国家高性能计算中心是我国并行算法研究、环境科学与工程计算软件的重要基地，在学术界和教育界有一定的影响和地位。

序　言

高性能计算机是一个国家经济和科技实力的综合体现，也是促进经济、科技发展，社会进步和国防安全的重要工具，已成为世界各国竞相争夺的战略制高点。一些发达国家纷纷制定战略计划，提出很高目标，投入大量资金，加速研究开发步伐。多年来，随着大规模集成电路技术的不断进步，以及 CPU 为基础的高性能并行计算机得到了迅速的发展，其高端系统正向百万亿次、千万亿次迈进。我国近十年来，对高性能并行计算的研究开发也给予了很大重视，取得了长足进步和可贵经验，研制出了具有相当水平的并行机系统，但与发达国家相比，差距仍然甚大，在高性能并行计算的应用开发与相关的人才培养教育方面尤现不足。如何使高性能并行机系统深入充分地在国民经济、科研和社会应用的发展中发挥作用，实为当务之急，引起人们的普遍关心。

由中国科技大学陈国良教授主编的这套丛书，正适应了我国高性能并行计算研究、开发、应用、教育之需。本丛书由《并行算法的设计与分析》、《并行计算机体系结构》和《并行算法实践》三大部分组成，而以《并行计算——结构·算法·编程》为全丛书之提要。该丛书以并行计算为主题，对并行计算的硬件平台（当代主流并行计算机系统）、并行计算的理论基础（并行算法的设计与分析）和并行计算的软件支撑（并行程序设计）全面系统地展开了讨论，内容丰富，取材新近，具有相当的深度和广度，涵盖了并行计算机体系结构和并行算法的理论、设计和实践的各个方面，是国内外不多见的优秀著作。

陈国良教授是国家高性能计算中心（合肥）主任，长期从事并行算法和并行计算机体系结构的研究，本套丛书是作者几十年从事教学与科研工作的结晶，是目前国内该领域内容涵盖最为全面的系列著作。它的出版必将对进一步推动我国并行计算学科的发展与应用推广产生深远的影响。

张效祥

2002 年 8 月

前　言

写作背景　随着半导体工艺和通信网络技术的进步，多CPU的并行计算机系统和网络机群系统得到了飞速的发展，这为当今科学技术的发展提供了定量化和精确化的计算手段。但为了高效地使用并行计算机解决给定的问题，除了要学习求解问题的并行算法及其并行编程实现外，还要学习并行计算机体系结构的基本原理、组织方式、关键技术和设计方法等，因为任何并行算法均要通过并行程序最终运行在具体的并行计算机上。

作者在过去几年曾陆续撰写了几本有关并行算法及并行计算方面的书籍，它们都是针对如何学习设计和分析并行算法的。而并行算法的主要特点是与具体的并行计算机体系结构有关的，所以，要想设计一个高效的并行算法必须对具体的并行机体系结构特点有充分的了解，尽管在那些书中照例都讲了一些并行机体系结构的内容，但限于篇幅，它们都是从并行算法设计的角度以高度抽象的方式来介绍一些最基本的和结论性的知识，显然，为了深入研究并行算法，它们是远远不够的。因此，出于为并行算法的研究提供雄厚的硬件基础的动机，促使作者编写了此书。尽管这是写作此书的初衷，但此书的写作仍严格遵循着其学科自身的完备性、系统性和内容的相对独立性，所以它仍是一本完整的并行计算机体系结构的教材。

章节内容　本书主要围绕着当代可扩放的并行计算机体系结构，从硬件和软件的角度，对对称多处理机系统、大规模并行处理系统、机群系统和分布共享存储系统等的组成原理、结构特性、关键技术、性能分析、设计方法及相应的系统实例等进行了讨论。书中取材先进、内容简练、体系完整，基本上涵盖了并行计算机体系结构的主要研究内容和主要研究方面。

全书内容可分为三个单元，共八章：第一单元为并行计算机体系结构的基础部分，包括绪论（第一章）、性能评测（第二章）和互连网络（第三章）；第二单元为当代主流并行计算机系统，包括对称多处理机系统（第四章）、大规模并行处理机系统（第五章）和机群系统（第六章）；第三单元为并行计算机体系结构的较深入的内容，包括分布共享存储系统（第七章）和并行机中的通信与延迟问题（第八章）。

使用方法　并行计算机体系结构是并行算法和并行程序设计两门课程的硬件基础，是计算机科学与技术一级学科的硕士研究生的学位选修课，是计算机系统结构专业的硕士研究生学位必修课。学生应在学习过计算机体系结构、操作系统和编译原理等之后学习本课程。作为必须讲授和最低60学时的教学要求，建议各章节讲授的学时分配为：第一章讲8学时，第二章讲6学时，第三章讲4学时，第四章讲10学时，第五章讲4学时，第六章讲6学时，第七章讲10学时，第八章讲10学时。书中带*号的部分是建议阅读的，它们或是预备性的知识（希望不熟悉此内容的读者课前预习），或是深入研究性内容（鼓励面向研究的

读者深入阅读)。每章之后均附有适量的习题;同时开列了本章正文中所引用的主要参考文献。全书最后还提供了专业术语中英对照及索引,以方便读者查阅。

相关书目 作者深知,一部著作应该内容深广和学科先进。而作为教材,要在充分考虑适应国情和便于教学使用以及发扬国人著书谨严、简练之特色的同时,更应该广泛吸取当今国内外相关教材中的先进精彩部分,以丰满自身的内容。所以作者在撰写此书时,及时地参阅了下列的著作:《**Parallel Computer Architecture: A Hardware/Software Approach**》(David E.Culler, Jaswinder Pal Singh and Anoop Gupta, Morgan Kaufmann Publishers, 1998);《**Scalable Parallel Computing: Technology, Architecture, Programming**》(Kai Hwang and Zhiwei Xu, WCB McGraw-Hill, 1998;中译本:《**可扩展并行计算:技术、结构与编程**》,黄铠、徐志伟著,陆鑫达等译,机械工业出版社,2000.5);《**共享存储系统结构**》(胡伟武,高等教育出版社,2001.7)和《**高等计算机体系结构——并行性,可扩展性,可编程性**》(Kai Hwang 著,王鼎兴、郑纬民、沈美明、温冬婵译,清华大学出版社和广西科学技术出版社,1995.8)等。如果读者在阅读本教材时,也能配合阅读它们,将是非常有益处的。

诚恳致谢 本书撰写时,曾直接或间接地引用了许多专家、学者的文献,不少材料也得益于上述所列的几本著作,作者向他们深表谢意;但也有很多作者的优秀论文未能被引用,作者深表歉意。书稿在付梓前承蒙王鼎兴先生进行了审阅,提出了很多中肯的修改意见,作者尤为感谢。

中国科学技术大学的历届学生们在听取本课程讲授中,曾提出过很多可贵意见,不断充实和完善了书稿的内容。特别是单久龙、何家华、陈勇、陈志辉、张青山、李辰等同学完成了本书的计算机绘图和计算机编辑工作,对于他们辛勤的劳动,作者一并表示感谢。

感谢中国科学技术大学教务处、计算机系、国家高性能计算中心(合肥)为本教材的写作所提供的支持和良好的工作条件。

陈国良教授拟定了全书章节内容,成稿后经过他反复修改而定稿。其中第一、二章由陈国良教授执笔,第三、四、六章由章隆兵博士完成初稿,第五、七、八章由章锋博士完成初稿,洪锦伟博士整理了全书的图稿,吴俊敏老师修订了第四、七、八章的内容和编制了索引。

本书的内容曾在中国科学技术大学计算机系讲授过多次,而定稿前的β版曾公布在中国科学技术大学国家高性能计算中心网站上由吴俊敏老师试用了一学期,并广泛地征求了意见。尽管这样,由于作者们学识有限,写作时间仓促,书中错误和片面之处在所难免,恳请读者不吝批评指正。

作者
中国科学技术大学
计算机科学技术系
国家高性能计算中心(合肥)
2002 年 6 月

内 容 提 要

本书以当代可扩放的并行计算机系统结构为主题，从硬件和软件的角度，着重讨论了对称多处理机系统、大规模并行处理机系统、机群系统和分布共享存储系统的组成原理、结构特性、关键技术、性能分析、设计方法及相应的系统实例等。

全书共八章，可分为三个单元：第一单元为并行计算机体系结构的基础部分，包括绪论（第一章）、性能评测（第二章）和互连网络（第三章）；第二单元为当代主流并行计算机系统，包括对称多处理机系统（第四章）、大规模并行处理机系统（第五章）和机群系统（第六章）；第三单元是并行计算机体系结构的较深入的内容，包括分布共享存储系统（第七章）和并行机中的通信与延迟问题（第八章）。

全书取材先进，内容精炼，体系完整，力图反映本学科的最新成就和发展趋势，可作为高等学校计算机及相关专业的本科高年级学生和研究生的教学用书；也可供从事计算机体系结构研究的科技人员阅读参考。

目　录

第一章 绪 论

本章首先简单介绍一下什么是并行计算机，为什么需要并行计算机以及如何学习并行计算机，然后详细讨论：并行计算机的发展背景，包括应用需求、技术进展和结构趋势；典型并行计算机系统介绍，包括阵列机、向量处理机、多处理机、多计算机和共享分布存储的多处理机；当代并行计算机体系结构，包括并行计算机的结构模型、并行计算机的访存模型和并行计算机存储层次结构及其一致性问题；并行计算机的应用基础，包括并行计算模型、并行程序设计模型以及并行计算机的同步和通信；我国曙光系列并行计算机的简介，包括曙光1号、曙光－1000和曙光－2000等。最后，小结简要讨论当今并行机体系结构研究的几个主要问题及其若干新技术。

本章的内容大体上概述了并行计算机体系结构的基本内容和研究的主要方面（其中互连网络和性能评测另辟两章单独讨论），对于那些不是从事并行计算机研究的读者，通读本章也是有益的。本章主要内容参考了文献[12]。

1.1 引 言

学习并行计算机，首先要知道什么是并行计算机，为什么需要并行计算机以及如何学习并行计算机。本节就从这些简单内容讲起。

1.1.1 什么是并行计算机

1. 并行计算机定义

简单地讲，并行计算机就是由多个处理单元(以下也称为处理器，或简称为CPU)组成的计算机系统，这些处理单元相互通信和协作，能快速、高效地求解大型复杂问题。

2. 并行机所涉及的问题

上述简单的定义中隐含着很多问题：例如处理单元有多少，这就涉及到系统是小规模的(十个或几十个)、中规模的(上百个)和大规模的(成千上万个)的问题；处理单元的功能有多强，这就涉及到系统的组织策略是"**蚁军法**"(Army of Ants)或"**象群法**"(Herd of Elephants)的问题；处理单元之间怎样连接，这就涉及到系统是按照什么样的拓扑结构彼此互连起来的问题；处理单元的数据是如何传输的，这就涉及到通信是按照共享变量方式还是消息传递方式的问题。至于各处理单元彼此相互协作共同求解大型复杂问题，则涉及到的问题更多，例如如何保证多处理单元操作的顺序性，这就涉及到同步互斥问题；如何确保共享数据的完整性问题，这就涉及到不同存储层次中的数据的一致性问题。此外，还有求解具体问题的并行程序的编写、调试、运行和性能分析等方面的问题。可见，并行计算机的定义虽很简单，但其内涵却相当丰富，具体实现亦相当复杂，而高效使用它也并非易事，所以必须认真仔细学习研究它。

3. 并行机的由来

并行计算机是相对串行计算机而言的，所谓**串行计算机**就是只有单个处理单元顺序执行计算程序的计算机，所以也称为**顺序计算机**。顺序计算机最早是从位串行操作到字并行操作、从定点运算到浮点运算改进过来的；然后它按照图1.1所示的过程逐步演变出各种并行计算机系统：从顺序**标量处理**(Scalar Processing)计算机开始，首先用**先行**(Look-Ahead)技术预取指令，达到重叠操作，实现功能并行；支持功能并行，可使用多功能部件和流水线两种方法；而流水线技术对处理

向量数据元素的重复相同的操作表现出强大的威力，从而产生了**向量流水线**(Vector-Pipelining)计算机(包括存储器到存储器和寄存器到寄存器两种结构)；不同于时间上并行的流水线计算机，另一分支的并行机是空间上并行的 SIMD(**单指令流多数据流**)并行机，它用同一控制器同步地控制所有处理器阵列，执行相同操作，来开发空间上的并行性；如果用不同的控制器异步地控制相应的处理单元，执行各自的操作，就派生出另一类非常主要的 MIMD(**多指令流多数据流**)并行机；其中，如果各处理单元通过公用存储器中的共享变量实现相互通信，就称为**多处理机**(Multiprocessors)；如果处理单元之间使用消息传递的方式来实现相互通信，就称为**多计算机**(Multicomputers)，它也是当今最流行的并行计算机，也是本书讨论的重点。

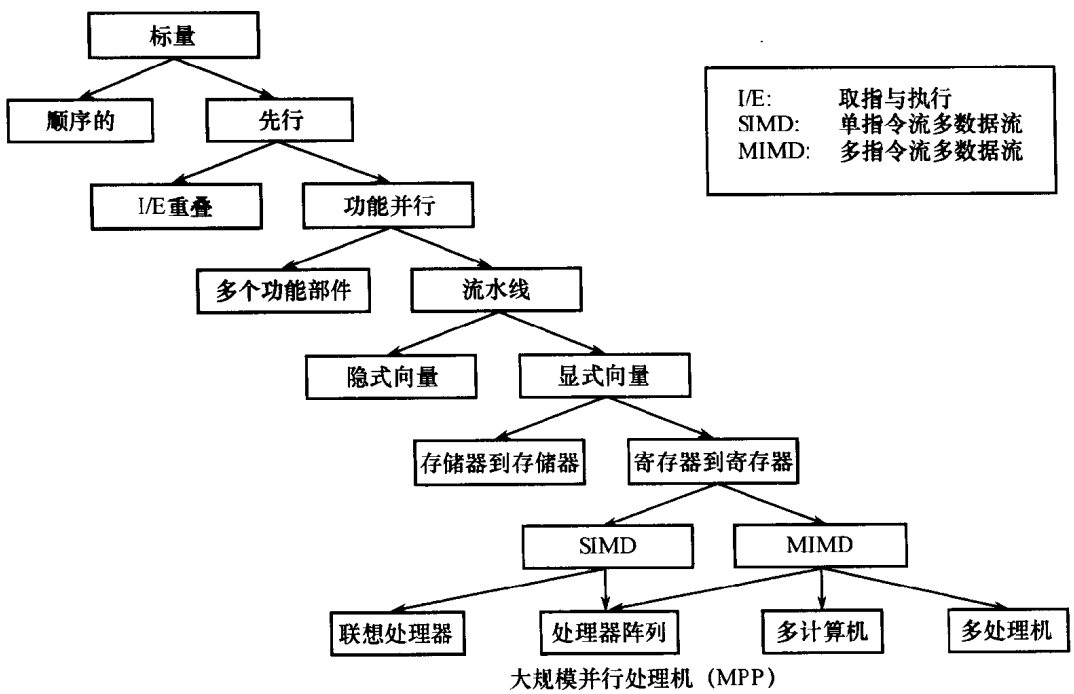

图 1.1 从标量处理计算机到向量流水线计算机和并行计算机的演变

4. Flynn 分类法

1966 年 Flynn[1]按照指令流和数据流的多倍性概念将计算机系统结构进行了分类。其中，指令流系指机器所执行的指令序列，数据流系指指令流所调用的数据序列，而多倍性系指机器的瓶颈部件上所可能并行执行的最大指令或数据的个数。根据指令流和数据流的不同组合，计算机系统可分为：如图 1.2(a)所示的**单指令**

流单数据流 SISD(Single Instruction Stream Single Data Stream),如图 1.2(b)所示的**单指令流多数据流 SIMD**(Single Instruction Stream Multiple Data Stream),如图 1.2(c)所示的**多指令流单数据流 MISD**(Multiple Instruction Stream Single Data Stream)和如图 1.2(d)所示的**多指令流多数据流 MIMD**(Multiple Instruction Stream Multiple Data Stream)。其中,**SISD** 就是传统的**单处理机**(又叫**串行机**或**顺序机**),**MISD** 是一种不太实际的计算机,但也有的学者把超标量机和**脉动阵列**(Systolic Array)机归属于此类,而 SIMD 和 MIMD 就是本书重点讨论的并行计算机。

5. 当代并行机系统

自 20 世纪 70 年代初到现在,并行计算机的发展已有 20 多年的历史。在此期间,出现了各种不同类型的并行机,包括历史上曾经风行一时的**并行向量处理机 PVP**(Parallel Vector Processor)和 SIMD 计算机,但它们现在均已衰落了下来,而 MIMD 类型的并行机却占了主导地位。当代的主流并行机是**可扩放并行计算机**(Scalable-Parallel Computer),包括:共享存储的**对称多处理机 SMP**(Symmetric Multiprocessor),分布存储的**大规模并行处理机 MPP**(Massively Parallel Processor),**分布式共享存储 DSM**(Distributed Shared Memory)多处理机和**工作站机群 COW**(Cluster of Workstations)以及刚刚兴起的跨地域性的、用高速网络将异构性计算节点连接起来满足用户分布式计算要求的所谓**网格计算环境 GCE**(Grid Computational Environment)。本书将重点讨论前四种当代可扩放的主流并行计算机。

6. 高性能计算机

在结束本小节之前,顺便讲一下并行计算机与高性能计算机的关系。其实,高性能计算机,并无明确严格的定义。因为性能可定义为求解问题所花费的时间的倒数,即求解问题的速度,所以按此意义,只要那些速度非常快的计算机都可认为是**高性能计算机**。当然,能高速求解问题的计算机,可以包括:**大型计算机**(Mainframe),如早期的 IBM370 系列;**超级计算机**(Supercomputer),如 Cray-1 向量计算机以及各种并行计算机。因为为了达到高性能,仅靠改进电路工艺,提高单机器件速度是有限的,所以使用并行计算机的方法则更为普遍和有效,于是并行计算机也就渐渐地变成了高性能计算机的同义词了,这种说法虽不严格,但已被普遍认可。

1.1.2 为什么需要并行计算机

为什么需要并行计算机?其最朴素的道理就是串行计算机满足不了求解问题的要求了,这些要求或是计算时间上的要求,或是计算精度上的要求,也或许是快