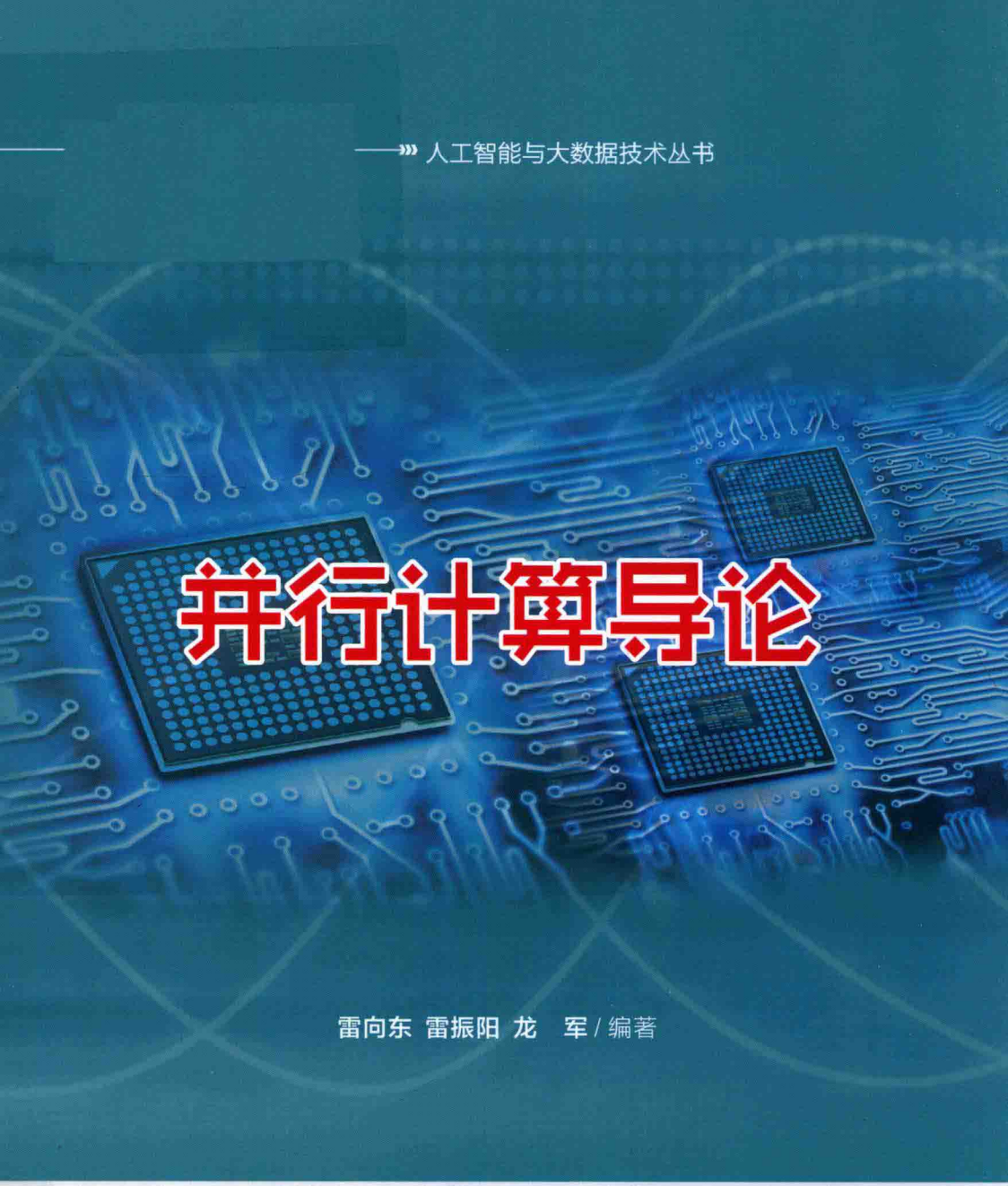
人工智能与大数据技术丛书



并行计算导论

雷向东 雷振阳 龙 军 / 编著

BINGXING JISUAN DAOLUN

中南大学出版社
www.csupress.com.cn

人工智能与大数据前沿技术丛书

并行计算导论

雷向东 雷振阳 龙 军 / 编著

BINGXING JISUAN DAOLUN

中南大學出版社
www.csupress.com.cn
·长沙·

图书在版编目（CIP）数据

并行计算导论／雷向东，雷振阳，龙军编著. --长沙：中南大学出版社，2018.11

ISBN 978-7-5487-3373-7

Ⅰ.①并… Ⅱ.①雷… ②雷… ③龙… Ⅲ.①并行算法 ②并行程序－程序设计 Ⅳ.①TP301.6 ②TP311.11

中国版本图书馆 CIP 数据核字(2018)第 200725 号

并行计算导论

雷向东　雷振阳　龙　军　编著

□责任编辑　韩　雪
□责任印制　易红卫
□出版发行　中南大学出版社
社址：长沙市麓山南路　　邮编：410083
发行科电话：0731-88876770　　传真：0731-88710482
□印　　装　长沙雅鑫印务有限公司

□开　　本　710×1000　1/16　□印张 18.75　□字数 377 千字
□版　　次　2018 年 11 月第 1 版　□2018 年 11 月第 1 次印刷
□书　　号　ISBN 978-7-5487-3373-7
□定　　价　48.00 元

前 言

并行处理是实现高性能、高可用计算机系统的主要途径。本书以并行计算为主题，主要介绍当代并行计算机系统及其结构模型、并行算法设计与并行程序的设计原理与方法，强调融并行机结构、并行算法和并行编程为一体，力图反映本学科的最新成就和发展趋势。通过本课程的学习，使学生从系统结构、算法、程序设计三个方面，初步了解并行处理的基本概念和涉及的各种学术和技术问题。通过本书的学习，使学生能把握并行处理技术的最新成就和发展趋势，掌握并行计算技术和方法。

全书内容被组织成如下 11 章。第 1 章并行计算机体系结构，介绍了高性能并行计算机的概念和系统组成，叙述了计算机系统的弗林分类法，详细讲解了 SIMD 和 MIMD 系统。第 2 章多处理器互连网络，介绍了静态互连网络和动态互连网络，详细讲解了并行计算机访存模型。第 3 章多处理器体系结构性能分析，介绍了并行计算性能评测，详细讲解了 Amdahl 定律、Gustafson 定律和 Sun - Ni 定律，并对并行系统的可扩展性和基准测试程序进行了讨论。第 4 章共享存储器系统，介绍了共享存储器系统高速缓存一致性，详细讲解了基于监听总线协议和基于目录的协议和共享存储器的编程模型。第 5 章消息传递系统，介绍了消息传递系统中路由，详细讲解了消息传递中的交换机制和消息传递系统编程模型。第 6 章多核构架，介绍了多核技术、多核中的并行性，以及多核处理器关键技术。第 7 章并行计算模型，介绍了 PRAM 模型、APRAM 模型、BSP 模型和 logP 模型等。第 8 章并行算法，包括数组求和、排序、矩阵运算、线性方程组求解。第 9 章并行程序设计方法，包括并行程序设计基础、共享存储系统并行编程、分布式存储系统并行编程和并行程序设计环境与工具。第 10 章 MPI 并行程序设计，详细讲解了 MPI 并行程序设计思想、方法和技术。第 11 章 OpenMP 并行程序设计，详细讲解了 OpenMP 并行程序设计思想、方法和技术。

本书涵盖了并行计算技术各个方面，包括并行计算机系统及其结构、并行计

算模型以及并行算法的设计方法和并行程序设计。每一章都附有大量的习题，根据教学进度和学时，合理选择书上习题，以达到进一步加深理解课堂讲授的内容的目的。

本书可选作高等院校《并行计算》课程的教材或参考书，主要读者为计算机科学与技术、数据科学与大数据技术、软件工程及相关学科的高年级本科生、研究生。由于本书具有内容充实、语言浅显、结构清晰、实例丰富等特点，所以本书的适用面非常广泛，如电子、自动控制等专业的高校学生，计算机应用开发人员，深入学习计算机应用技术的普通读者等。

由于作者水平有限，书中难免有不妥之处，恳请读者给予指正和提出修改意见。

编　者

2018 年 8 月

目 录

第 1 章　并行计算机体系结构

在快速解决计算量大、数据密集型问题时，人们越来越认识到并行处理是唯一节省成本的方法。价格低廉的并行计算机(如商用桌面多处理机和工作站机群)的出现，使得并行处理的适用范围越来越广。现在人们已经为可移植的并行程序设计制定了专门的软件标准，为并行软件的大幅度发展打好了基础。

事务处理、信息检索、数据挖掘和分析以及多媒体服务等数据密集型应用已经为当代的并行平台提供了新的挑战。计算生物学和纳米技术等新兴的领域对并行计算的算法和系统开发提供了前瞻性的启示，而体系结构、编程模型和应用中出现的变化，对如何使用户以网格服务形式得到并行平台也提供了一些启发。

1.1　并行计算

并行机上所作的计算，称为并行计算，又称高性能计算或超级计算。并行计算的目的是提高计算速度及通过扩大问题求解规模来解决大型而复杂的计算问题。并行计算可分为时间上的并行和空间上的并行。时间上的并行就是指流水线技术，而空间上的并行则是指用多个处理器并发的执行计算，即通过网络将两个以上的处理机连接起来，达到同时计算同一个任务的不同部分，或者解决单个处理机无法解决的大型问题的目的。并行计算主要用于科学与工程计算、气象预报、油藏模拟、核武器数值模拟、航天器设计、基因测序等。

并行计算机就是由多个处理单元组成的计算机系统，这些处理单元相互通信和协作，能快速高效地求解大型的复杂问题。

网络计算是在工作站机群 COW 环境下进行的计算。主要特点是其结合了客户机/服务器结构的健壮性，Internet 面向全球的数据访问方式的通用性和分布式对象的灵活性，提供了统一的跨平台开发环境，基于开放的和事实上的标准，把应用和数据的复杂性从桌面转移到智能化的网络和基于网络的服务器，给用户提供了对应用和信息的通用、快速的访问方式。

分布式计算是一门计算机科学，它研究如何把一个需要非常巨大的计算能力才能解决的问题分成许多小的部分，然后把这些部分分配给许多计算机进行处理，最后把这些计算结果综合起来得到最终的结果。

集群计算是使用多个计算机(如典型的个人计算机或 UNIX 工作站)、多个存储设备进行冗余互联，来组成一个对用户来说单一的高具有可用性的系统。

网络计算与分布式计算和集群计算都是属于计算密集型、数据密集型和网络密集型应用。

1.2 计算机的四个发展阶段

大多数计算机科学工作者认同计算机的发展存在四个不同的计算模式或时期，即批处理时期、分时时期、台式机时期和网络时期。

1. 批处理时期

在批处理系统中加载在计算机上的一个系统软件控制下，计算机能够自动地、成批地处理一个或多个用户的作业(这作业包括程序、数据和命令)。直到 1965 年，IBM System/360 主机一直占据着公司计算机中心的市场批处理机器。它是极具代表性的批处理机器，该机器有一个操作系统，并配置多种编程语言。

2. 分时时期

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，当时，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。硬件技术的发展导致了小型计算机的来临。由于 CPU 速度不断提高和采用分时技术，一台计算机可同时连接多个用户终端，而每个用户可在自己的终端上联机使用计算机。处理机的运行时间分成很短的时间片，按时间片轮流把处理机分配给各联机作业使用。由于计算机运行速度很快，每位用户会觉得自己独占了一台计算机。而每位用户可以通过自己的终端向系统发出各种操作控制命令，在充分的人机交互情况下，完成作业的运行。在 20 世纪 70 年代出现了超级计算机，Cray 公司在 1976 年推出具有最好性价比的超级计算机 Cray - 1。

3. 台式机时期

个人计算机(PC)是由 Altair、Processor Technology、North Star、Tandy、Commodore、Apple 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。个人计算机的普及改变了计算的面貌。约在 1990 年，功能强大的个人计算机和工作站局域网(LAN)开始替代大型计算机和小型计算机。不久单个台式机被广域网(WAN)连接成更大的计算联合体。

4. 网络时期

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。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—90 年代面世了许多具有多处理器的商用并行计算机。它们可以分为两大类：①共享存储器系统；②分布式存储器系统。

计算机发展的一个明显走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，它们将能提供可依赖的、一致的以及廉价的对高端计算设施的访问。

1.3　高性能计算机的发展

从 20 世纪 70 年代产生第一代高性能计算机开始，经过几十年的发展，高性能计算机经历了向量机、MPP、集群等几个发展阶段。1974 年，CDC 公司推出了 CDC STAR－100，首先使用向量处理器。1982 年，Cray 公司生产的 Cray X－MP 是世界上第一部并行向量计算机。在 20 世纪 70 年代和 80 年代，并行向量计算处理充分利用了流水线和多功能部件，极大地提高了计算机运算速度。

但由于时钟周期已接近物理极限，进一步提高并行向量计算机的速度非常困难。在这样的背景下，一个全新的概念被提出来了，那就是大规模并行处理（massively parallel processing，MPP）。1992 年，Intel 公司推出 Paragon 超级计算机，它成为历史上第一台突破万亿次浮点计算屏障的超级计算机。紧接着，IBM 公司的 SP2、日立公司的 SR2201 和 SGI 公司的 Origin2000 超级计算机都先后出现，超级计算机也开始走上了真正的商用化道路。MPP 逐渐成为高性能计算机的主流。

20 世纪 90 代中期，随着局域网技术的快速发展，局域网在带宽和延迟上与传统高性能计算机所采用的专有网络的差距也日渐缩小，集群（cluster）系统出现。集群系统是使用高速通信网络将多台 PC 机、工作站或 SMP 连接在一起，构成一个统一的整体系统。与 SMP 和 MPP 相比，集群具有更高的可扩展性、可用性和易维护性，而且价格低、性价比高。但是，在最高端并行计算机中大多数还是采用 MPP 架构。

下面列出一些具有代表性的超级并行计算机：

- 1997，Intel 公司，ASCI Option Red，1Tflops。
- 1998，SGI 公司，Option Blue Mountain，3Tflops。
- 2001，IBM 公司，Option White，7Tflops。
- 2002，NEC 公司，Earth Simulator，35Tflops。
- 2005，IBM 公司，eServer Blue Gene Solution，280Tflops。
- 2002，中国联想，1Tflops。

- 2013，中国天河二号，54.9Pflops。
- 2016，神威·太湖之光，93Pflops。

国际 TOP500 组织 2017 年 11 月 13 日公布新一期全球超级计算机 500 强榜单，中国超算“神威·太湖之光”和“天河二号”连续第四次分列冠亚军。中国“神威·太湖之光”和“天河二号”的浮点运算速度分别为每秒 9.3 亿亿次和每秒 3.39 亿亿次。“神威·太湖之光”的浮点运算速度是“天河二号”的近 3 倍，更重要的是其核心部件全部为国产，凸显了中国在超算领域的自主研发能力。

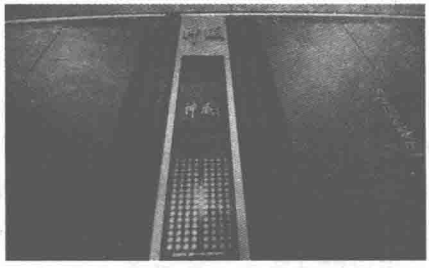


图 1.1 “神威·太湖之光”超级计算机

图 1.2 天河二号超级计算机

1.4 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(M. J. Flynn)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(single instruction stream single data stream，SISD)、单指令多数据流(single instruction stream multiple data stream，SIMD)、多指令单数据流(multiple instruction stream single data stream，MISD)、多指令多数据流(multiple instruction stream multiple data stream，MIMD)。

传统的单处理器冯·诺依曼被归为 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 机器中，相同的数据流流过执行不同指令的一个线性处理器阵列。现实中没有可靠的 MISD 机。

并行计算就是在并行计算或分布式计算机等高性能计算系统上所做的超级计

算。计算极大地增强了人们从事科学研究的能力，大大地加速了把科技转化为生产力的过程，深刻地改变着人类认识世界和改造世界的方法和途径。计算科学的理论和方法，作为新的研究手段和新的设计与创造技术的理论基础，正推动着当代科学与技术向纵深发展。并行计算的系统结构分两大类：单指令多数据流 SIMD 和多指令多数据流 MIMD；其中多指令多数据流 MIMD 包括：并行向量处理机 PVP、对称多处理机 SMP、大规模并行处理机 MPP、工作站机群 COW 和分布共享存储多处理机 DSM。

1.5　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。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，但前端相可向处理器阵列发出命令让它并行 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，以在巨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 1.3 所示为 SIMD 结构图。

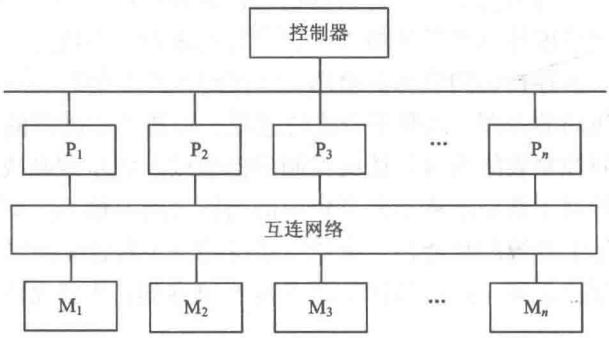


图 1.3　SIMD 结构