国家自然科学基金资助项目

陈国良 编著

并行算法

排序和选择

SIMD
MIMD
VLSI
SIMD
MIMD
VLSI

中国科学技术大学出版社

并 行 算 法

——排 序 和 选 择

陈国良 编著

国家自然科学基金资助项目

中国科学技术大学出版社
1990·合肥

并　行　算　法

——排序和选择

陈国良　编著

责任编辑：黄　德　　封面设计：王瑞荣

*

中国科学技术大学出版社出版

（安徽省合肥市金寨路96号，邮政编码230026）

中国科学院开封印刷厂印刷

安徽省新华书店发行

*

开本：787×1092/16　印张：16.25　字数：382千

1990年6月第1版　1990年6月第1次印刷

印数：1-4000册

ISBN 7-312-00135-1/TP·10　定价：7.00 元

内 容 简 介

本书较系统、深入和全面地介绍了各种专用和通用并行计算模型上的并行排序和选择算法。书中以最基本、最常用而又最具有代表性的典型算法为主，同时又兼顾到尽量能反映本学科的最新成就和发展趋势。取材内容具有一定的广度和深度，是目前国内第一本关于并行排序和选择算法方面的教材和专著。

本书共分十章。内容包括：并行算法基础，比较器网络、VLSI 专用结构上的排序和选择算法，固定连接和共享存贮的 SIMD 和 MIMD 机器上的排序和选择算法。

本书既可作为高等院校计算机科学、计算机工程、信息工程类的高年级学生和研究生的教材，也可作为从事计算机科学理论和算法研究的科技人员的参考书。

前　言

算法一词，对大部分读者并不陌生，学过算法的读者大都可以自然地接受并行算法这一术语．并行处理技术和并行计算机的出现和发展，为在并行计算机上求解问题(即并行计算)创造了条件．为达到高效的并行计算所发展的各种问题的精确求解方法就称之为并行算法．

并行算法和串行算法虽然有不少相似之处，但我们不能认为它只是串行算法的简单推广和扩充．并行算法的一大特点就是它和并行计算机的结构密切相关，以后我们将会看到，不同结构的并行计算机将会导致不同风格的并行算法．但是，我们不可能一一研究各种并行计算机上的并行算法，而只能从现有的并行计算机中抽象出几种典型的所谓并行计算模型，然后围绕着它展开讨论各种并行算法．

并行算法研究的范围很广，我们可以根据其基本运算特性将其分为：基于代数数值运算的一类数值计算的并行算法；基于比较关系运算的一类非数值计算的并行算法．矩阵运算和解三角形线性系就是典型的数值计算；而排序和选择则是典型的非数值计算．

排序和选择这一研究领域，在计算机学科中占有相当的地位．这不仅是由于它们有着广阔的应用前景，而且其本身也具有理论研究的意义．在计算机计算和处理数据的过程中，都会直接或间接地涉及到数据的排序和选择；在系统软件或应用软件中，也不可避免地会遇到排序和选择问题；在数据库、知识库管理系统中，排序和选择更是应用广泛；在高级计算机体系结构中，非数值计算所占的比重将越来越大，而排序和选择操作却占了非数值计算的相当分量．非旦如此，在已发现的各种著名的算法中，排序和选择本身展示出它们是值得深入研究和认真剖析的有趣的课题，其对问题的研究方法和解决思路、算法设计和分析的技巧，对设计其它问题的并行算法都颇值得鉴借．正是因为如此，在过去和现在，不少人都从排序和选择问题开始学习和研究算法．

由于排序和选择问题在计算机学科中占有如此突出地位，所以对它们的研究曾一度热闹非凡，而且研究的兴趣至今不衰，年年都涌现出数以百计的学术论文和研究报告．在这浩如烟海的文献中，摄取哪些材料作为教材的基本内容，最能反映这一学科的最新成就和发展趋势，乃是本书写作的宗旨和奋斗目标．作者通过对计算机专业高年级学生和研究生的教学实践，结合作者从事这方面研究工作的心得、体会，在阅读有关文献资料的基础上，按照自己的理解，对其进行了提炼、总结、归纳和整理，最后编著成书．

本书共分十章．前三章是并行算法基础，包括并行计算机和并行计算、并行算法的设计和递归方程的求解．有基础的读者可以浏览一下或跳过这些章节．第四章和第九章是讨论专用并行计算模型——比较器网络和 VLSI 专用结构上的排序和选择算法．第五、六、七章，集中讨论固定连接和共享存贮的 SIMD 及 MIMD 计算模型上的排序和选择算法．第八章是并行外排序，内容比较独立，可以选读．第十章讨论概率并行算法，虽然涉及到稍多的数学知识，但没有数学爱好的读者也能够理解性地阅读本章．最后附录 A 和附录 B 的内容仅供读者阅读时查阅．

本书的内容可作为计算机专业高年级学生和研究生一学期的教材使用。对于专门从事计算机科学理论和算法研究的科技工作者，他们可以选择各自所兴趣的章节深入地学习和研究.

本书在撰写中，曾直接或间接地引用了许多专家、学者的文献；在本书内容的取舍和章节的安排等方面，唐策善副教授提出了很多宝贵意见；书稿付梓前，承蒙王鼎兴教授审阅，作者谨此一并表示衷心感谢.

顾乃杰同志帮助我整理了第三章、第十章和附录B的素材；沈鸿、熊焰、姚新、张永民、梁维发、唐锡南、张士恒同志，在资料的收集方面作了大量工作. 他们在听我的讲授中又提出过可贵的建议，从而丰富和充实了本书的内容. 对于他们的辛勤劳动和良好的愿望，作者尤为感谢.

当国内外系统、深入阐述这方面内容的教材或专著尚为鲜见的时候，作者能向国内广大读者奉献此书而感到高兴，但也深深感到，限于作者水平，书中定有不妥和错误之处，恳请读者批评指正.

陈国良

于中国科学技术大学

1989年4月15日

目　　录

第一章　并行计算机与并行计算

本章主要讨论并行算法的基础：包括并行计算机与并行计算；并行计算模型；并行计算的若干理论问题以及并行算法的一般概念.

1.1　并行计算机及其分类

1.1.1　并行计算机的发展[1,2]

计算机发展的趋势是越来越先进，从数据和信息处理到知识处理，最终到智能高处理，每前进一步，均要求增强计算机的处理能力．计算机发展的历史表明，为了达到处性能，除了必须提高元器件的速度外，系统结构也必须不断改进，特别是当元器件的速度达到极限时，后者将变成问题的焦点．计算机系统结构的改进，主要围绕着增加同一时间间隔内的操作数量，即所谓**并行处理** (Parallel Processing) **技术**；而为并行处理所设计的计算机统称之为**并行计算机** (Parallel Computer)；在并行计算机上求解问题的过程称之为**并行计算** (Parallel Computing)；凡适合于在各种并行计算机上求解问题和处理数据的算法就称之为**并行算法** (Parallel Algorithm).

并行计算是一比较年轻的领域：因为著名的 IlliacⅣ (阵列处理机，也叫阵列处理器，下同) 1975 年才开始运行；第一台 Cray-1(流水线向量处理机) 1976 年才交付使用；低成本的多处理机(也称多处理器，下同)1984年以后也才开始流行.

并行计算机的发展主要是某些大型应用领域的要求：例如气象预报，空气动力学，人工智能，卫星图象处理，核反应堆以及军事应用等．为了达到上述高性能的要求，除了提高线路速度外，主要是改进计算机的系统结构，例如：

①**引入 I/O 通道**：将费时的 I/O 操作交给 I/O 处理器(即通道)去作，从而解脱了 CPU 的负担；

②**交叉存贮**：将存贮体分成多个模块，以达到并行存取和减少冲突之目的；

③**高速缓存**：减少处理器和主存中的数据交换时间，平滑其间的数据流动速率；

④**指令先行**：一次取出好几条下一次待执行的指令，致使取指令可以和指令译码同样快；

⑤**多功能单元**：在中央处理器中设置多个功能单元 (加法、乘法器等) 可以提高单一程序的吞吐量，缩短周转时间；

⑥**指令重叠和流水线技术**：在同一时间允许多条指令在不同的阶段按流水线方式执行不同的操作；

⑦**向量处理技术**：一条向量指令可同时处理向量的 n 个分量，它们可以同时在各个处理器中进行运算，也可以连续地送入流水线中进行重叠处理；

⑧**多道程序设计**：同时把若干个作业放在内存中，允许在同一时间有多道程序处于运行状态；

⑨**分时系统**：允许多个终端用户同时交互地使用一台计算机；

⑩**数据驱动**：与常规的冯·诺依曼计算机不同，它不是采用指令驱动操作，而是采用操作数(据)驱动操作，这样如果有多个操作数准备就绪时，就可以彼此并行地执行而不受指令顺序执行的限制，从而有可能充分开拓并行度.

计算机从出现到现在，可以按照器件工艺、系统结构、处理模式和计算机语言，将计算机分成若干个代(Generations)．绝大部分作者均把1980年之前的计算机划归为第一代到第四代的计算机：第一代的计算机由电子管构成，程序设计语言采用机器语言；第二代的计算机由晶体管构成，开始使用汇编语言和高级语言 (Fortran, Algol, Cobol 等)；第三代计算机广泛采用中、小规模集成电路，使用多道程序设计技术；第四代计算机使用 VLSI 作为逻辑和存贮单元，采用了分时操作系统和虚拟存贮技术．目前正在兴起的第五代计算机，对其尚无明确的定义和界线.

1.1.2 并行计算机简介

因为并行计算机是并行算法的基础，所以我们首先简单介绍一些现存的、典型的并行计算机．对于它们的讨论，不求全面，而是从算法的角度，略去一些技术细节，抽象地讨论一下有关的并行计算机的类型、结构特点和运算方式.

1. 阵列处理机 (Array Processor)

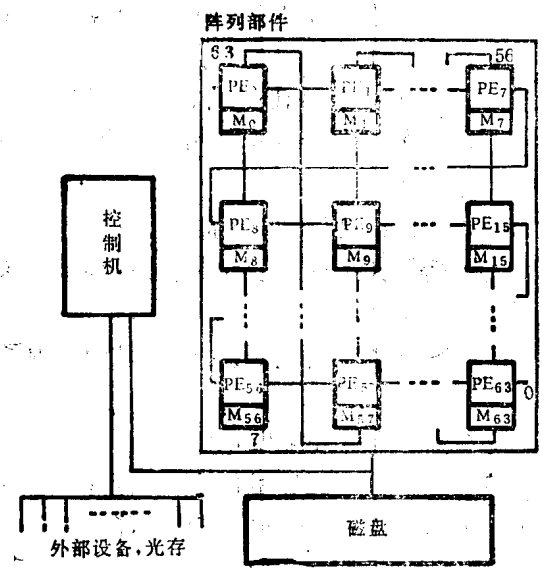

图 1.1 Illiac Ⅳ机器框图

有名的并行处理机 Illiac Ⅳ 就是阵列结构的．如图1.1所示，64台PE (Processing Element)各带有自己的私有存贮器M，它们排成 8×8 的阵列(因而得名为阵列处理机)．每个 PE 均可和其上、下、左、右四个近邻相连.所有的 PE 在同一控制器控制之下，按同一指令的要求，对同一数组中的不同数据同步地进行操作，从而达到操作级并行．这样的机器对有限差分、矩阵运算和快速富氏变换 (FFT) 等计算具有很高的效率，在诸如图象信息处理一类的专门应用领域中也卓有成效[3]．今后我们所讨论的一大类有关SIMD机器上的同步并行算法就是以此种机器为基础的.

注意，**关联处理机** (Associative Processor) 也是一种特殊的阵列机，它按存贮内容寻址，对于信息搜索之类的处理尤为合适.

2. 流水线处理机 (Pipeline Processor)

将生产流水线装配技术应用于计算机结构中，把计算机的运算部件或控制部件等装配成一些有序的子部件，利用功能部件分离与时间重叠的办法，使每个被操作的对象处在整个操作流程的不同功能部件中，且保持在不同的完成阶段，从而达到操作级的并行，这种结构的计算机称为流水线处理机，亦叫做向量机．流水线思想首先应用于指令

的操作上；继之在功能划分的基础上，以流水线方式组成高速中央处理器；进而出现了在一台机器的中央处理器中设置多条专用流水线，让它们并行工作或协同完成一些复合操作．这种系统对向量加工甚为有效，是目前解决大型数值计算问题的巨型机的主要型式．

Star-100是典型的流水线向量处理机[4]．如图1.2所示，它是由中央处理器，主存贮器，外围机及盘，站和外部设备等组成．

3．多处理机 (Multiprocessor) 和多计算机 (Multicomputer)

多处理机(有时讨论算法时也称为多处理器)和多计算机是由一些可编程的且均可各自执行自己程序的多台处理器组成．其中，多处理机以各处理器共享公共存贮器为特征；而多计算机以各处理器经通信链路传递信息为特征．它们与阵列机的根本区别在于：阵列机中的每台处理器只能执行中央处理器的指令；而前者的每台处理器仅接收中央处理器分给它的任务，它执行自己的指令，所以可达到指令、任务级并行．

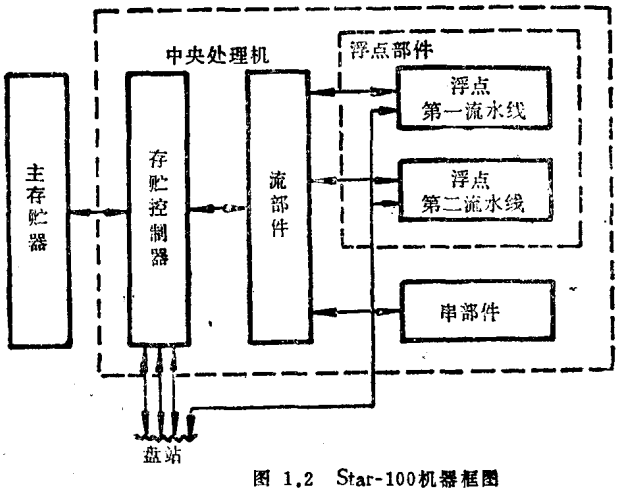

图 1.2 Star-100机器框图

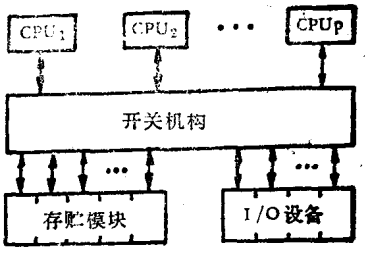

图 1.3 紧耦合的多处理机系统

如图1.3所示，在多处理机系统中，如果所有的处理器都是通过一中央开关机构(如公共总线、交叉开关、包开关等)去访问全局共享存贮器，则这种形式的多处理机称为**紧耦合多处理机** (Tightly Coupled Multiprocessor) 系统．Carnegie-Mellon 大学的 $C.mmp$ 就是一种有名的紧耦合多处理机系统[5]．和紧耦合多处理机系统不同，如果每个处理器都带各自的局存，将各个局存组合起来形成一个共享的地址空间，则这样的多处理机称为**松散耦合多处理机** (Loosely Coupled Multiprocessor) 系统．因为松散耦合的多处理机没有集中的开关机构，所以可以连接大量的处理器．Carnegie-Mellon 大学的 $C.m^*$ 就是一种有名的松散耦合的多处理机系统[6]．如图1.4所示，$C.m^*$ 采用了机群结构．目前已研制成的系统共包含 50 台 DEC LSI-11 微计算机，用三级总线(LSI-11 总线，Map 总线和群间双总线)连接起来，$K.map$ 负责把各个计算机模块所带有的容量为 28K 字的局存组织在一起，形成统一的具有 28 位地址的虚拟存贮空间．

今后，我们所讨论的一类有关 MIMD 机器上的异步并行算法和分布式算法就是以它们为基础的．

正在发展中的非冯·诺依曼计算机，例如**数据流计算机** (Data Flow Machine)，

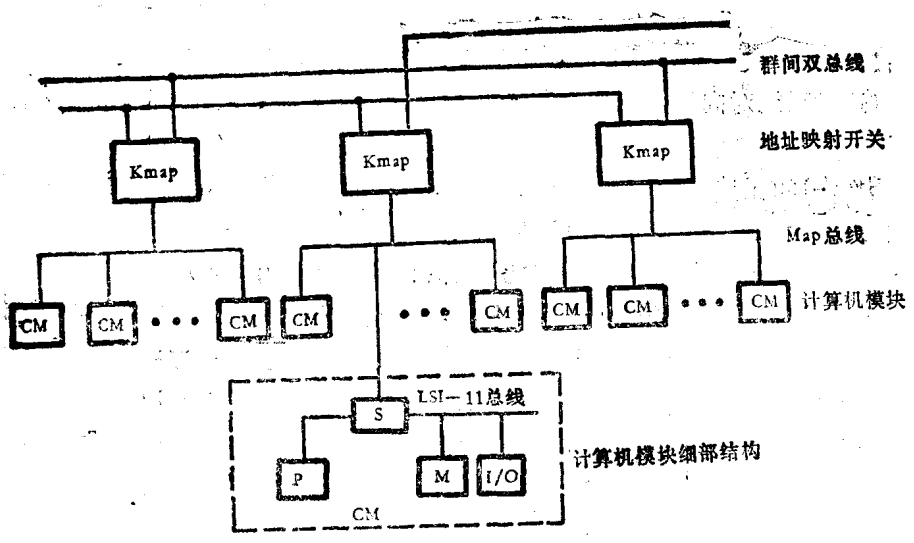

图 1.4 C_m^*多处理机结构框图

归约计算机 (Reduction Machine)，**推理计算机** (Inference Machine) 等，虽然都是新一代的并行计算机，但是由于在这些机器上并行算法的研究尚不成熟，所以在此不再讨论它们.

1.1.3 并行计算机的分类

1. Flynn 分类法[7]

分类是为了讨论问题的需要和方便，但任何分类法都不是尽善尽美的. 1966年美国M.J.Flynn，按照指令流和数据流将计算机系统分为四类：

①单指令流单数据流：SISD(Single Instruction stream Single Data stream)计算机. 这就是传统的顺序处理机.

②单指令流多数据流：SIMD(Single Instruction stream Multiple Data stream)计算机. 上节所述的阵列处理机，流水线处理机，关联处理机等均属此类计算机.

③多指令流单数据流：MISD(Multiple Instruction stream Single Data Stream)计算机. 此类计算机，实际上以何种机器为代表尚存疑议.

④多指令流多数据流：MIMD(Multiple Instruction stream Multiple Data stream) 计算机. 上节所述的多处理机和多计算机均属此类计算机. 这类计算机虽是并行处理机的理想结构，但在算法、程序、结构、操作系统等方面还有许多复杂的问题待解决.

2. Händler 分类法[8]

1977年，Händler 根据计算机系统中流水线和并行度出现的级别，将一台计算机表示为三对整数. 令 PCU 代表处理器控制单元，它相应于一处理器或 CPU；ALU 代表算术逻辑单元，它相应于功能单元或处理单元；BLC 代表位一级电路. 于是，按照 Händler 的方法，一台计算机可表示如下：

$$T(C) = \langle K \times K', D \times D', W \times W' \rangle \tag{1.1}$$

其中，K = PCU 的数目；

K' = 能够流水线执行的 PCU 的数目；

D = 每个 PCU 所控制的 ALU 的数目；

D' = 能够流水线执行的 ALU 的数目；

W = ALU 或处理单元 PE 中的位数；

W' = 在所有 ALU 或单一 PE 中流水线段数.

上式中，如果任一对整数的第二个元素值为 1，则就略去它. 在单一计算机系统中"×"号也用以连接不同种类的处理器描述.

对于 CDC6600 而言，它有一个单一的 CPU；ALU 有 10 个功能单元，它们都可以流水线执行；CDC6600 的字长 60 位；最多有 10 个外围 I/O 处理器，它们可以与 CPU 彼此并行工作；每个 I/O 处理器有一个单一的字长为 12 位的 ALU. 这样，

$$T(\text{CDC6600}) = T(\text{中央处理器}) \times T(\text{I/O处理器})$$
$$= \langle 1,1\times 10,60\rangle \times \langle 10,1,12\rangle$$

从研究并行算法的角度，大多数作者都乐于使用 Flynn 分类法.

1.2 处理器的互连方式[9]

在并行计算机中，如何将各类处理器或处理器与存贮器互连起来是个关键问题. 除了常见的总线连接和交叉开关连接外，本节将引入几种重要的互连方式，以便为下节并行计算模型直接引用.

1.2.1 一维线性连接

一维线性阵列 (Linear Array)，简记之为 LA，其连接方式或许是并行处理机中处理器之间一种最简单、最基本的互连方式. 其中每个处理器只与其左、右近邻相连(首、尾处理器例外)，所以也叫做二近邻连接. 这种连接方式是 Systolic 结构的最基本形式.

约定：处理器的数目 $n=2^m$，令其地址的二进制表示式为：$P=p_{m-1}p_{m-2}\cdots p_0$ (下同)，则线性阵列的连接函数 LC 可定义如下：

$$\left.\begin{aligned} &LC_{-1}(P)=P-1,\quad 1\leqslant P\leqslant n-1\\ &LC_{+1}(P)=P+1,\quad 0\leqslant P\leqslant n-2\end{aligned}\right\} \tag{1.2}$$

一维线性连接的变体使用方式是循环移位连接.

1.2.2 网孔结构

在**网孔连接**(Mesh-Connected)中，每个节点都排在 q-维网格中的格点上. 当 $q=2$ 时，就是熟知的四近邻连接，简记之为 MC^2，连接函数 MC^2 可定义如下：

$$\left.\begin{aligned} &MC^2_{-1}(P)=P-1\\ &MC^2_{+1}(P)=P+1\\ &MC^2_{-\sqrt{n}}(P)=P-\sqrt{n}\\ &MC^2_{+\sqrt{n}}(P)=P+\sqrt{n}\end{aligned}\right\} \bmod n,\quad 0\leqslant P\leqslant n-1 \tag{1.3}$$

上述连接函数，也可用置换的轮换表示法表示之：