

# 并行多核体系结构基础



[并行多核体系结构基础\\_下载链接1](#)

著者:[美]汤孟岩 (Yan Solihin)

出版者:机械工业出版社

出版时间:2018-10-25

装帧:平装

isbn:9787111610410

虽然多核现在是主流架构，但很少有教科书涵盖并行多核体系结构。本书填补了这一空

白，为研究生或高级本科体系结构课程提供了所有材料，重点是多核处理器的体系结构。这本书也适合作为从事多核编程或多核芯片设计的专业人员的参考书。

## 本书特点

从不同的视角向读者介绍共享存储并程序是如何编写的。

帮助读者理解共享存储多核和多处理器所需要的软件基础及硬件支持。

讨论了存储层次、设计共享存储并行多处理器时的基本问题、高速缓存一致性、存储一致性、同步、互连网络，并向读者展示不同概念是如何交互和适配的。

探索图形处理单元（GPU）系统中经常使用的单指令流多线程（SIMT）编程模型。

## 作者简介:

[美]汤孟岩 (Yan Solihin) 著: 汤孟岩 (Yan Solihin)

北卡罗来纳州立大学电子与计算机工程系教授，长期从事计算机体系结构方向的研究工作。研究兴趣包括计算机体系结构、计算机系统建模方法和图像处理，在计算机体系结构和性能建模领域发表过大量高水平论文，相关研究受到美国国家自然科学基金、Intel、IBM、Samsung、Tekelec、Sun Microsystems和HP的资助。他于2017年被选为IEEE会士，并入选了高性能计算机体系结构国际会议（HPCA）名人堂（2015年）。此外，他还长期从事计算机体系结构的教学工作，具有丰富的教学经验。创立和领导了针对性能、可靠性和安全的体系结构研究小组，并且开源了大量针对多核体系结构性能建模和性能优化的软件工具。

## 目录: 译者序

前言

缩写词表

第1章 多核体系结构概述 1

1.1 多核体系结构的由来 2

1.2 并行计算机概述 9

1.2.1 并行计算机的Flynn分类法 12

1.2.2 MIMD并行计算机分类 13

1.3 未来的多核体系结构 14

1.4 习题 18

课堂习题 18

课后习题 18

第2章 并行编程概述 20

2.1 并行程序性能的限制因素 20

2.2 并行编程模型 23

2.2.1 共享存储与消息传递模型的对比 25

2.2.2 一个简单的例子 26

2.2.3 其他编程模型 29

2.3 习题 37

课后习题 37

第3章 共享存储并行编程 39

3.1 并行编程的步骤 39

3.2 依赖分析 40

3.2.1 循环级依赖分析 42

3.2.2 迭代空间遍历图和循环传递依赖图 42

3.3 识别循环结构中的并行任务 45

|                                |     |
|--------------------------------|-----|
| 3.3.1 循环迭代间的并行和DOALL并行         | 45  |
| 3.3.2 DOACROSS: 循环迭代间的同步并行     | 46  |
| 3.3.3 循环中语句间的并行                | 48  |
| 3.3.4 DOPIPE: 循环中语句间的流水线并行     | 50  |
| 3.4 识别其他层面的并行                  | 51  |
| 3.5 通过算法知识识别并行                 | 53  |
| 3.6 确定变量的范围                    | 55  |
| 3.6.1 私有化                      | 56  |
| 3.6.2 归约变量和操作                  | 57  |
| 3.6.3 准则                       | 58  |
| 3.7 同步                         | 59  |
| 3.8 任务到线程的映射                   | 60  |
| 3.9 线程到处理器的映射                  | 64  |
| 3.10 OpenMP概述                  | 67  |
| 3.11 习题                        | 72  |
| 课堂习题                           | 72  |
| 课后习题                           | 77  |
| 第4章 针对链式数据结构的并行编程              | 79  |
| 4.1 LDS并行化所面临的挑战               | 79  |
| 4.2 LDS并行化技术                   | 80  |
| 4.2.1 计算并行化与遍历                 | 80  |
| 4.2.2 针对数据结构的操作并行化             | 82  |
| 4.3 针对链表的并行化技术                 | 89  |
| 4.3.1 读操作之间的并行                 | 89  |
| 4.3.2 LDS遍历中的并行                | 91  |
| 4.3.3 细粒度锁方法                   | 94  |
| 4.4 事务内存                       | 98  |
| 4.5 习题                         | 99  |
| 课堂习题                           | 99  |
| 课后习题                           | 101 |
| 第5章 存储层次结构概述                   | 103 |
| 5.1 存储层次的意义                    | 103 |
| 5.2 高速缓存体系结构基础                 | 104 |
| 5.2.1 数据放置策略                   | 105 |
| 5.2.2 数据替换策略                   | 109 |
| 5.2.3 数据写策略                    | 111 |
| 5.2.4 多级高速缓存中的包含策略             | 113 |
| 5.2.5 统一/分立/Banked高速缓存和高速缓存流水线 | 117 |
| 5.2.6 高速缓存寻址和旁路转换缓冲            | 119 |
| 5.2.7 非阻塞式高速缓存                 | 121 |
| 5.3 高速缓存性能                     | 122 |
| 5.3.1 高速缓存缺失的幂次定律              | 124 |
| 5.3.2 栈距离特性                    | 125 |
| 5.3.3 高速缓存性能指标                 | 126 |
| 5.4 预取                         | 127 |
| 5.4.1 步长预取和顺序预取                | 128 |
| 5.4.2 多处理器系统中的预取               | 130 |
| 5.5 多核体系结构中的高速缓存设计             | 130 |
| 5.6 高速缓存的物理组成                  | 131 |
| 5.6.1 集中式高速缓存                  | 131 |
| 5.6.2 分布式高速缓存                  | 133 |
| 5.6.3 混合式高速缓存                  | 133 |
| 5.7 高速缓存的逻辑组成                  | 135 |
| 5.7.1 散列函数                     | 139 |

|                                              |     |
|----------------------------------------------|-----|
| 5.7.2 改善共享高速缓存的距离局部性                         | 140 |
| 5.7.3 私有高速缓存结构中的容量共享                         | 141 |
| 5.8 案例分析                                     | 143 |
| 5.8.1 IBM Power7的存储层次                        | 143 |
| 5.8.2 AMD Shanghai和Intel Barcelona处理器存储层次的比较 | 146 |
| 5.9 习题                                       | 148 |
| 课堂习题                                         | 148 |
| 课后习题                                         | 150 |
| 第6章 共享存储多处理器简介                               | 152 |
| 6.1 缓存一致性问题                                  | 153 |
| 6.2 存储一致性问题                                  | 155 |
| 6.3 同步问题                                     | 156 |
| 6.4 习题                                       | 160 |
| 课堂习题                                         | 160 |
| 课后习题                                         | 161 |
| 第7章 缓存一致性基础                                  | 163 |
| 7.1 概述                                       | 164 |
| 7.2 基于总线的多处理器缓存一致性问题                         | 168 |
| 7.2.1 “写直达”缓存的一致性协议                          | 168 |
| 7.2.2 “写回”缓存的MSI协议                           | 170 |
| 7.2.3 “写回”缓存的MESI协议                          | 175 |
| 7.2.4 “写回”缓存的MOESI协议                         | 178 |
| 7.2.5 “写回”缓存基于更新的协议                          | 183 |
| 7.3 缓存设计对缓存一致性性能的影响                          | 186 |
| 7.4 性能及其他实际问题                                | 187 |
| 7.4.1 预取和一致性缺失                               | 187 |
| 7.4.2 多级缓存                                   | 187 |
| 7.4.3 侦听过滤                                   | 189 |
| 7.5 点对点互连网络上的广播式协议                           | 189 |
| 7.6 习题                                       | 203 |
| 课堂习题                                         | 203 |
| 课后习题                                         | 207 |
| 第8章 对同步的硬件支持                                 | 209 |
| 8.1 锁的实现                                     | 209 |
| 8.1.1 对锁实现性能的评估                              | 209 |
| 8.1.2 对原子指令的需求                               | 210 |
| 8.1.3 TS锁                                    | 212 |
| 8.1.4 TTSL                                   | 214 |
| 8.1.5 LL/SC锁                                 | 215 |
| 8.1.6 Ticket锁                                | 218 |
| 8.1.7 ABQL                                   | 219 |
| 8.1.8 各种锁实现的量化比较                             | 221 |
| 8.2 栅障的实现                                    | 222 |
| 8.2.1 翻转感应集中式栅障                              | 223 |
| 8.2.2 组合树栅障                                  | 225 |
| 8.2.3 硬件栅障实现                                 | 225 |
| 8.3 事务内存                                     | 227 |
| 8.4 练习                                       | 232 |
| 课堂习题                                         | 232 |
| 课后习题                                         | 236 |
| 第9章 存储一致性模型                                  | 238 |
| 9.1 程序员的直觉                                   | 238 |
| 9.2 保证顺序一致性的体系结构机制                           | 242 |
| 9.2.1 在基于总线的多处理器中基本的SC实现                     | 242 |

- 9.2.2 改善SC性能的技术 244
- 9.3 松弛的一致性模型 246
  - 9.3.1 安全网 246
  - 9.3.2 处理器一致性 246
  - 9.3.3 弱序 248
  - 9.3.4 释放一致性 250
  - 9.3.5 惰性释放一致性 253
- 9.4 不同存储一致性模型中的同步 254
- 9.5 习题 256
  - 课堂习题 256
  - 课后习题 260
- 第10章 高级缓存一致性设计 262
  - 10.1 目录式一致性协议 262
  - 10.2 目录式一致性协议概览 262
  - 10.3 目录式缓存一致性协议基础 268
  - 10.4 实现正确性和性能 272
    - 10.4.1 由目录状态不同步引起的竞争处理 272
    - 10.4.2 由对请求非实时处理引起的竞争处理 274
    - 10.4.3 写传播和事务串行化 280
    - 10.4.4 同步支持 281
    - 10.4.5 存储一致性模型 282
  - 10.5 当前设计问题 283
    - 10.5.1 处理不精确的目录信息 283
    - 10.5.2 一致性粒度 286
    - 10.5.3 系统划分 288
    - 10.5.4 加速线程迁移 289
  - 10.6 习题 291
    - 课堂习题 291
    - 课后习题 293
- 第11章 互连网络体系结构 295
  - 11.1 链路、信道和延迟 296
  - 11.2 网络拓扑 298
  - 11.3 路由策略和算法 302
  - 11.4 路由器体系结构 312
  - 11.5 案例研究：Alpha 21364网络体系结构 315
  - 11.6 多核设计的问题 317
  - 11.7 习题 319
    - 课堂习题 319
    - 课后习题 321
- 第12章 SIMT体系结构 323
  - 12.1 SIMT编程模型 323
  - 12.2 将SIMT工作负载映射到SIMT核上 325
  - 12.3 SIMT核体系结构 326
    - 12.3.1 标量ISA 326
    - 12.3.2 SIMD化/向量化：Warp构造 326
    - 12.3.3 细粒度多线程（Warp级并行） 328
    - 12.3.4 微体系结构 328
    - 12.3.5 流水线执行 329
    - 12.3.6 控制流处理 330
    - 12.3.7 内存系统 331
  - 12.4 习题 334
    - 课堂习题 334
    - 课后习题 335
- 第13章 专家访谈 337

参考文献 356  
索引 361  
· · · · · (收起)

[并行多核体系结构基础\\_下载链接1\\_](#)

## 标签

并行

多核

体系结构

计算机

并行计算

计算机科学

多线程

Concurrency

## 评论

更偏软件，我读过的对于cache consistency讲的最好的一本.

-----  
[并行多核体系结构基础\\_下载链接1\\_](#)

-----  
[并行多核体系结构基础\\_下载链接1](#)