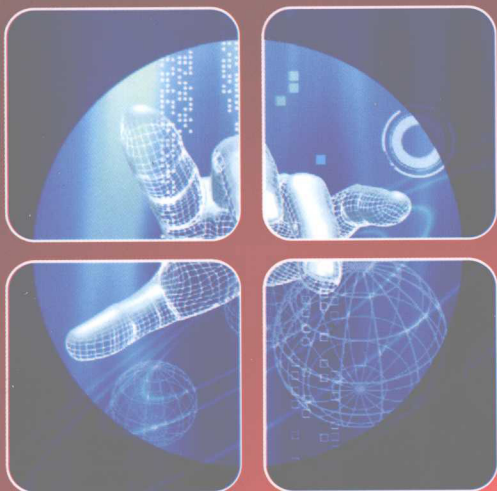


中国仿真科学与技术书系

“十一五”国家重点图书出版规划



SIMULATION SCIENCE

并行与分布仿真系统

Parallel and Distributed Simulation Systems

[美] Richard M. Fujimoto 著
李 革 刘宝宏 张耀程 译



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

“十一五”国家重点图书出版规划
中国仿真科学与技术书系

Parallel and Distributed Simulation Systems
并行与分布仿真系统

(美) Richard M. Fujimoto 著
李 革 刘宝宏 张耀程 译

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

本书为软件开发者提供了最前沿的技术,使用这些技术可以加速仿真软件在多处理器上的运行速度,处理广域网包括互联网上的数据分发。本书的重点是并行与分布式离散事件仿真技术,本书作者把该领域最近 20 年的研究成果进行了汇编和梳理,讨论了并行与分布式计算机在系统行为的建模与分析,以及分布式虚拟环境的建立方面的应用。

并行与分布仿真方面的其他著作专注于应用方面,而本书清晰地描述了如何实现并行与分布仿真技术。本书详细地解释了正确实现仿真所必需的同步算法,包括对时间弯曲算法和先进乐观技术的透彻讨论。最后,本书提供了非常丰富的参考文献、图例、表格以及当前的系统实例,如美国国防部的高层体系结构 HLA,HLA 已经成为美国国防项目的标准体系结构。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

Parallel and Distributed Simulation Systems

978-0-471-18383-9

Richard M. Fujimoto

Original English Edition Copyright © 2000 by John Wiley & Sons, Inc.

All rights reserved.

This translation published under license.

Authorized Translation of the Edition Published by John Wiley & Sons, Inc., New York, Chichester, Brisbane, Singapore and Toronto.

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

Copies of this book sold without a Wiley sticker on the back cover are unauthorized and illegal.

本书中文简体版专有出版权由 John Wiley & Sons, Inc. 公司授予电子工业出版社。未经许可,不得以任何手段和形式复制或抄袭本书内容。

本书封底贴有 John Wiley & Sons, Inc. 防伪标签,无标签者不得销售。

版权贸易合同登记号 图字: 01-2009-2869

图书在版编目(CIP)数据

并行与分布仿真系统 / (美) 藤本 (Fujimoto, R.M.) 著; 李革, 刘宝宏, 张耀程译. —北京: 电子工业出版社, 2010.3

(中国仿真科学与技术书系)

书名原文: Parallel and Distributed Simulation Systems

ISBN 978-7-121-10048-2

I. 并… II. ①藤…②李…③刘…④张… III. ①并行计算机—仿真系统②分布式计算机—仿真系统 IV. TP338.6 TP338.8

中国版本图书馆 CIP 数据核字 (2009) 第 226177 号

策划编辑: 徐 静

责任编辑: 康 霞

印 刷: 北京智力达印刷有限公司

装 订: 北京中新伟业印刷有限公司

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

开 本: 787×1 092 1/16 印张: 17 字数: 397 千字

印 次: 2010 年 3 月第 1 次印刷

印 数: 4 000 册 定价: 42.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

INTRODUCTION

The rapid expansion of the Internet and commodity parallel computers has made parallel and distributed simulation (PADS) a hot technology indeed. Applications abound not only in the analysis of complex systems such as transportation or the next-generation Internet, but also in computer-generated virtual worlds for military and professional training, interactive computer games, and the entertainment industry.

In this book, PADS expert Richard M. Fujimoto provides software developers with cutting-edge techniques for speeding up the execution of simulations across multiple processors and dealing with data distribution over wide area networks ,including the Internet. With an emphasis on parallel and distributed discrete event simulation technologies, Dr. Fujimoto compiles and consolidates research results in the field spanning the last twenty years, discussing the use of parallel and distributed computers in both the modeling and analysis of system behavior and the creation of distributed virtual environments.

While other books on PADS concentrate on applications, *Parallel and Distributed Simulation Systems* clearly shows how to implement the technology. It explains in detail the synchronization algorithms needed to properly realize the simulations, including an in-depth discussion of time warp and advanced optimistic techniques. Finally, the book is richly supplemented with references, tables and illustrations, and examples of contemporary systems such as the Department of Defense's High Level Architecture (HLA), which has become the standard architecture for defense programs in the United States.

谨以此书献给

Jan, Emily 和 Alex

“中国仿真科学与技术书系” 编委会

主编： 黄柯棣

副主编： 庞国峰 李 革

编委会成员：（按拼音排序）

毕红哲	方胜良	郭齐胜	龚建华	胡晓峰	吕跃广
李 群	李世忠	王维平	王雪松	王中杰	卫军胡
肖田元	杨 峰	杨瑞平	杨西龙	朱一凡	

前 言

并行与分布仿真领域目前正处于令人激动的时候。经过大学和工业实验室多年的研究和开发，并行与分布仿真领域得到了蓬勃发展，目前已经在现实世界的许多系统中得到了应用。写本书的目的是对离散事件仿真程序在由通过网络互联的多个处理器组成的计算平台上执行所涉及的技术问题进行深入论述。这种平台可以从位于房间内或单个机柜内紧耦合的多处理器计算机系统，到地理上分布的个人计算机或分布于全球的专用模拟器（例如视频游戏系统）。该技术能够用来加快大型仿真的执行速度，例如下一代互联网的仿真，或用来建立用于训练或娱乐的分布式综合环境。

作者写本书的目的是将散落于无数期刊和会议论文中关于并行与分布仿真系统的基本原理汇集于一册，适用的读者包括参与分布式仿真系统研究或开发的管理者和从业者。本书可以作为高年级本科生或研究生计算机科学课程的教科书，尽管本书的重点是并行与分布计算问题，其他学科也可能对本书感兴趣（例如工业工程或运筹学）。如果您具有离散事件仿真、并行或分布式计算的预备知识，将有助于您对本书的学习，不过这并不重要，因为本书将对这些知识进行简单介绍。

内容

本书分为三部分，第一部分对该领域进行了介绍。第1章描述了并行与分布仿真技术的典型应用，通过历史回顾描绘了开发和改进该技术的团体，复习了并行与分布仿真系统的背景信息。第2章复习了离散事件仿真的基本原理，为本书后续章节的学习提供了公共的基础和术语。

第二部分主要涉及仿真程序的并行与分布执行，主要用于分析领域的应用，例如大型、复杂系统的设计，目标是采用多处理器加快执行速度。这4章的大部分内容与同步算法有关，同步算法用于保证仿真程序并行执行产生的结果与串行执行相同，但是更快（希望如此）！处理同步问题的两个主要方法被称为保守同步和乐观同步，第3章涉及前者，第4章和第5章涉及后者。第6章涉及处理并行执行的一种完全不同的方法，称为时间并行执行，时间并行执行仅仅适用于某类仿真问题，但是一旦得到应用，能够使性能大幅提高。

第三部分涉及分布式虚拟环境（Distributed Virtual Environments, DVEs），强调的是实时仿真，即建立人可以嵌入其中的虚拟环境，例如用于训练或娱乐的虚拟环境。第7章对该领域进行了介绍，主要集中介绍国防团队的两个成就，即分布式交互仿真（Distributed Interactive Simulation, DIS）和高层体系结构（High Level Architecture, HLA），这项技术大部分在国防领域开发出来并得到应用。第8章和第9章涉及分布式虚拟环境的两个具体问题。第8章讨论将数据有效分发到所有分布式虚拟环境参与者的问题，本章前半部分介绍了为分布式虚拟

环境底层通信提供支持的计算机网络，后半部分涉及有效利用网络基础设施的技术，特别是在具有许多交互组件的大规模仿真中。最后，第 9 章再次涉及分布式虚拟环境中的时间同步问题，以及保证所有参加仿真的不同计算机具有正确同步时钟的问题。

第一部分将为本书的余下部分打基础，因此必须首先阅读，第二部分和第三部分可以按照任意顺序阅读。在为乔治亚理工学院为期 10 周的并行与分布仿真课程授课时，我采用了本书作为教科书，当大学一个学期改为 15 周时我仍计划使用本书。作为选择，本书可以作为离散事件系统仿真课程的一部分。当使用本书时，教师为了更加简化地处理主题素材，可以跳过第 5 章、第 6 章以及第 8 章的前半部分。

软件

感兴趣的读者可能希望对本书讨论的一些算法进行试验。尽管本书不包括软件，但还是可以获取到的，特别是第 5 章讨论的 GTW (Georgia Tech Time Warp) 软件和 HLA-RTI 子集的一个实现都可以免费应用于教育和研究中。软件信息见 <http://www.cc.gatech.edu/computing/pads>，若要获取这些软件包的复制，可以通过电子邮件 fujimoto@cc.gatech.edu 与我联系。

感谢

显然，如果没有无数学术界和工业界人士做出的许多技术贡献，本书无法写成。我努力在补充阅读材料和参考文献中尽量多地列出这些贡献者，遗憾的是该领域已经发展到了不可能列出全部贡献者的程度。

我感谢对本书直接做出贡献的许多人，特别是我在乔治亚理工学院讲授并行与分布仿真研究生课程时学生对早期草稿提出的许多修改建议，也感谢 Glenn Oberhauser 和 Katherine Morse 明确、详细的意见。感谢资助我开展并行与分布仿真研究的投资机构，部分研究成果已经包括在本书中，这些机构包括弹道导弹防空组织 (BMDO)、国防先期研究计划局 (DARPA)、国防建模与仿真办公室 (DMSO)、国家科学基金会、SAIC 公司、Mitre 公司、Bellcore 公司、陆军研究办公室、海军研究办公室和战略导弹防空司令部。

最后，我还要感谢我的家人，由于这个写书计划，我无数个夜晚和周末不能与她们相伴，尽管如此家人还是给予我无尽的支持和理解。如果没有她们的爱和奉献，本书不会完成。

目 录

第一部分 总 论

第 1 章 背景与应用	3
1.1 为什么要用分布/并行仿真	5
1.2 分析仿真与虚拟环境	6
1.3 历史回顾	8
1.3.1 高性能计算团体	8
1.3.2 国防团体	9
1.3.3 交互式游戏和互联网团体	9
1.4 应用	10
1.4.1 军事应用	11
1.4.2 娱乐	11
1.4.3 社会交往和商业合作	12
1.4.4 教育和训练	12
1.4.5 远程通信网络	12
1.4.6 数字逻辑电路和计算机系统	13
1.4.7 交通	14
1.5 基础技术	14
1.6 硬件平台	14
1.6.1 并行计算机与分布式计算机	14
1.6.2 共享存储多处理器	16
1.6.3 分布式存储多计算机	17
1.6.4 单指令流、多数据流计算机	18
1.6.5 分布式计算机	18
1.7 总结	19
1.8 补充阅读材料	19
第 2 章 离散事件仿真基本原理	23
2.1 时间	25
2.2 实时、比例实时与尽可能快执行	26
2.3 状态改变和时间流机制	27
2.3.1 时间步进执行	27
2.3.2 事件驱动执行	29
2.4 离散事件仿真程序	29
2.5 一个应用例子	31

2.6	仿真的启动与终止	34
2.7	并行/分布式仿真例子	34
2.8	世界观和面向对象仿真	35
2.8.1	仿真进程	36
2.8.2	基于对象仿真和面向对象仿真	38
2.8.3	查询事件和推拉处理	39
2.8.4	事件取消	40
2.9	开发并发执行的其他途径	40
2.10	补充阅读材料	41

第二部分 并行与分布式离散事件仿真

第 3 章	保守同步算法	45
3.1	同步问题	47
3.2	使用空消息避免死锁	49
3.3	前瞻量和仿真模型	52
3.4	死锁的检测和恢复	53
3.4.1	死锁检测	53
3.4.2	死锁恢复	55
3.5	同步执行	57
3.5.1	集中式栅障	58
3.5.2	树栅障	58
3.5.3	蝴蝶栅障	59
3.5.4	暂态消息	60
3.5.5	一个简单的同步协议	63
3.5.6	逻辑进程间的距离	64
3.6	受限滞后	67
3.7	条件信息与无条件信息	69
3.8	动态进程及互连	69
3.9	可重复性和并发事件	71
3.9.1	使用时戳的隐藏域对并发事件排序	71
3.9.2	优先数	72
3.9.3	接收方指定排序	72
3.10	保守机制的性能	72
3.11	保守机制的总结和评论	77
3.12	补充阅读材料	78

第 4 章 时间弯曲算法	81
4.1 预备知识	83
4.2 本地控制机制	84
4.2.1 回退状态变量	85
4.2.2 取消消息发送	86
4.2.3 零前瞻量、并发事件及可重复性	91
4.3 全局控制机制	92
4.3.1 化石回收	93
4.3.2 错误处理	94
4.4 计算全局虚拟时间	95
4.4.1 暂态消息问题	95
4.4.2 同步报告问题	97
4.4.3 Samadi 的 GVT 算法	98
4.4.4 Mattern 的 GVT 算法	98
4.5 其他机制	102
4.5.1 动态内存分配	102
4.5.2 周期性状态保存	104
4.5.3 指定保存内容	105
4.5.4 事件取消	107
4.5.5 懒惰撤销	108
4.5.6 懒惰再评估	110
4.6 调度逻辑进程	111
4.7 总结	112
4.8 补充阅读材料	112
第 5 章 先进乐观技术	115
5.1 时间弯曲算法中的内存使用	117
5.1.1 预备内容: 状态向量和消息发送时戳	118
5.1.2 内存管理机制和消息返回	119
5.1.3 存储最优化	120
5.1.4 回撤协议	122
5.1.5 人工回退协议	123
5.1.6 回剪协议	124
5.1.7 基于内存的流控制协议	125
5.1.8 性能和内存需求之间的权衡	126
5.2 时间弯曲算法中的性能风险	127
5.2.1 追逐错误计算	127
5.2.2 回退反射	129

5.3 其他乐观同步算法	130
5.3.1 移动时间窗	131
5.3.2 基于前瞻量的阻塞机制	131
5.3.3 本地回退算法	132
5.3.4 周期时间桶算法	132
5.3.5 “狼来了”算法	133
5.3.6 概率回退算法	134
5.3.7 时空仿真	134
5.3.8 小结	135
5.4 综合: GTW 仿真环境	135
5.4.1 编程接口	135
5.4.2 I/O 及动态内存分配操作	136
5.4.3 GTW 的数据结构	137
5.4.4 直接撤销	139
5.4.5 事件处理循环	139
5.4.6 缓冲区管理	140
5.4.7 流控制	141
5.4.8 GVT 计算和化石回收	141
5.4.9 增量式状态保存	142
5.4.10 本地消息发送	142
5.4.11 消息复制	142
5.4.12 事件批处理	142
5.4.13 性能评测	143
5.5 小结	144
5.6 比较乐观与保守同步协议	144
5.7 补充阅读材料	146
第 6 章 时间并行仿真	149
6.1 使用修正计算方法进行时间并行缓存仿真	152
6.2 使用再生点进行 ATM 多工器仿真	155
6.3 使用并行前缀的队列仿真	159
6.4 小结	161
6.5 补充阅读材料	161

第三部分 分布式虚拟环境

第 7 章 分布式虚拟环境介绍	165
7.1 目标	167

7.2	分布式虚拟环境与并行离散事件仿真系统对照	167
7.3	服务器结构与无服务器结构	168
7.4	分布式交互仿真	170
7.4.1	DIS 设计原则	170
7.4.2	DIS PDUs	171
7.4.3	时间约束	172
7.5	航位推测法	173
7.5.1	DR 模型	175
7.5.2	时间补偿	175
7.5.3	平滑	176
7.6	高层体系结构	177
7.6.1	历史回顾	177
7.6.2	HLA 述评	178
7.6.3	HLA 规则	179
7.6.4	对象模型和对象模型模板	180
7.6.5	接口规范	183
7.6.6	典型的联邦执行	184
7.7	总结	185
7.8	补充阅读材料	185
第 8 章	联网技术和数据分发	187
8.1	消息传输服务	189
8.1.1	可靠传输	189
8.1.2	消息排序	189
8.1.3	面向连接的与无连接的通信	190
8.1.4	单点传输与组通信	190
8.1.5	举例: DIS 和 NPSNet	190
8.2	联网要求	191
8.3	联网技术	192
8.3.1	局域网技术	192
8.3.2	广域网技术	195
8.3.3	服务的质量	196
8.4	通信协议	197
8.4.1	OSI 协议栈	197
8.4.2	ATM 协议栈	199
8.4.3	网际互联与互联网协议	201
8.5	组通信	203
8.5.1	组和组通信原语	203

8.5.2 传输机制	204
8.6 数据分发	204
8.6.1 数据分发系统的接口	204
8.6.2 例子: HLA 中的数据分发	206
8.6.3 实现问题	209
8.6.4 动态组管理	211
8.6.5 广域观察者和快速移动的实体	212
8.7 小结	212
8.8 补充阅读材料	213
第 9 章 时间管理和事件排序	215
9.1 问题描述	217
9.2 消息排序服务	218
9.2.1 因果排序	218
9.2.2 因果全排序	221
9.2.3 Delta 因果关系	222
9.2.4 时戳排序	223
9.3 同步墙上时钟	224
9.3.1 时间和时钟源	224
9.3.2 时钟同步算法	225
9.3.3 纠正时钟同步错误	227
9.3.4 网络时间协议	227
9.4 小结	228
9.5 补充阅读材料	229
缩略语	231
参考文献	235

CONTENTS

Preface

PART I INTRODUCTION

1 Background and Applications	3
1.1 Why Parallel/Distributed Simulation?	5
1.2 Analytic Simulations versus Virtual Environments	6
1.3 Historical Perspective	8
1.3.1 High-Performance Computing Community	8
1.3.2 Defense Community	9
1.3.3 Interactive Gaming and Internet Communities	9
1.4 Applications	10
1.4.1 Military Applications	11
1.4.2 Entertainment	11
1.4.3 Social Interactions and Business Collaborations	12
1.4.4 Education and Training	12
1.4.5 Telecommunication Networks	12
1.4.6 Digital Logic Circuits and Computer Systems	13
1.4.7 Transportation	14
1.5 Underlying Technologies	14
1.6 Hardware Platforms	14
1.6.1 Parallel versus Distributed Computers	14
1.6.2 Shared-Memory Multiprocessors	16
1.6.3 Distributed-Memory Multicomputers	17
1.6.4 SIMD Machines	18
1.6.5 Distributed Computers	18
1.7 Summary	19
1.8 Additional Readings	19
2 Discrete Event Simulation Fundamentals	23
2.1 Time	25
2.2 Real-Time, Scaled Real-Time, and As-Fast-As-Possible Execution	26

2.3	State Changes and Time Flow Mechanisms	27
2.3.1	Time-Stepped Execution	27
2.3.2	Event-Driven Execution	29
2.4	Discrete-Event Simulation Programs	29
2.5	An Example Application	31
2.6	Starting and Stopping the Simulation	34
2.7	Parallel/Distributed Simulation Example	34
2.8	World Views and Object-Oriented Simulation	35
2.8.1	Simulation Processes	36
2.8.2	Object-Based and Object-Oriented Simulations	38
2.8.3	Query Events and Push versus Pull Processing	39
2.8.4	Event Retraction	40
2.9	Other Approaches to Exploiting Concurrent Execution	40
2.10	Additional Readings	41

PART II PARALLEL AND DISTRIBUTED DISCRETE-EVENT SIMULATION

3	Conservative Synchronization Algorithms	45
3.1	Synchronization Problem	47
3.2	Deadlock Avoidance Using Null Messages	49
3.3	Lookahead and the Simulation Model	52
3.4	Deadlock Detection and Recovery	53
3.4.1	Deadlock Detection	53
3.4.2	Deadlock Recovery	55
3.5	Synchronous Execution	57
3.5.1	Centralized Barriers	58
3.5.2	Tree Barrier	58
3.5.3	Butterfly Barrier	59
3.5.4	Transient Messages	60
3.5.5	A Simple Synchronous Protocol	63
3.5.6	Distance between Logical Processes	64
3.6	Bounded Lag	67
3.7	Conditional versus Unconditional Information	69
3.8	Dynamic Processes and Interconnections	69

3.9	Repeatability and Simultaneous Events	71
3.9.1	Using Hidden Time Stamp Fields to Order Simultaneous Events	71
3.9.2	Priority Numbers	72
3.9.3	Receiver-Specified Ordering	72
3.10	Performance of Conservative Mechanisms	72
3.11	Summary and Critique of Conservative Mechanisms	77
3.12	Additional Readings	78
4	Time Warp	81
4.1	Preliminaries	83
4.2	Local Control Mechanism	84
4.2.1	Rolling Back State Variables	85
4.2.2	Unsending Messages	86
4.2.3	Zero Lookahead, Simultaneous Events, and Repeatability	91
4.3	Global Control Mechanism	92
4.3.1	Fossil Collection	93
4.3.2	Error Handling	94
4.4	Computing Global Virtual Time	95
4.4.1	Transient Message Problem	95
4.4.2	Simultaneous Reporting Problem	97
4.4.3	Samadi's GVT Algorithm	98
4.4.4	Mattern's GVT Algorithm	98
4.5	Other Mechanisms	102
4.5.1	Dynamic Memory Allocation	102
4.5.2	Infrequent State Saving	104
4.5.3	Specifying What to Checkpoint	105
4.5.4	Event Retraction	107
4.5.5	Lazy Cancellation	108
4.5.6	Lazy Re-evaluation	110
4.6	Scheduling Logical Processes	111
4.7	Summary	112
4.8	Additional Readings	112
5	Advanced Optimistic Techniques	115
5.1	Memory Utilization in Time Warp	117