

**В. А. РАЙХЛИН**

**СИСТЕМЫ  
ПАРАЛЛЕЛЬНОЙ  
ОБРАБОТКИ ДАННЫХ**



**Казань 2010**

УДК 681.3  
ББК 32  
Р 11

**Райхлин В.А.**

**Р 11 Системы параллельной обработки данных – Казань: Изд-во «Фэн» («Наука»), 2010. – 268 с.**

**ISBN 978-5-9690-0157-2**

*Анализируются – ретроспектива и действительность параллельных систем, различные способы их классификации и оценок производительности, предметные предпосылки параллелизма. Даётся знакомство – со спецификой ассоциативных вычислений и систем битовой обработки в целом, с архитектурой современных мультипроцессоров и мультикомпьютеров типов майнфрейм и кластер, с параллельными системами управления консервативными базами данных на кластерной платформе. Рассматриваются – организация коммутационных сетей и главной памяти параллельных систем, их скоростные характеристики, механизмы поддержки когерентности кэшей, специфика RAID-массивов внешней памяти. Излагаются – вопросы построения операционных сред и процессорных матриц на их основе, параллельных алгоритмов для таких матриц. Показываются – особенности графических и Cell-процессоров. Представляется – комплексный характер разработок специализированных параллельных систем.*

Для студентов вузов направления «Информатика и вычислительная техника» и их преподавателей, научных работников и аспирантов.

Табл. 26. Ил. 86. Библиогр.: 89 назв.

© Издательство «Фэн» («Наука»), 2010

© В.А. Райхлин, 2010

## ПРЕДИСЛОВИЕ

Наблюдаемый бум активности в области высокопроизводительных информационно-вычислительных систем вполне закономерен. Он обусловлен тем, что успехи информационных технологий хорошо отвечают сейчас потребностям человеческого сообщества в целом. Сочетания «высокопроизводительные вычисления» и «параллельные вычисления» синонимичны в единстве архитектуры, алгоритмов и программирования. В книге рассматриваются архитектурно-алгоритмические вопросы параллельных вычислений во взаимосвязи с насущными задачами современности. Проблематика параллельного программирования затрагивается в минимальной степени.

Кратко о содержании книги.

Раздел 1 поясняет различные трактовки понятия СуперЭВМ, ретроспективу и действительность параллельных систем. Дает начальные представления о способах классификации и оценок производительности параллельных систем, о некоторых предметных предпосылках параллелизма, связанных с цифровой обработкой сигналов, обработкой изображений и символьной информации.

Раздел 2 уделяет достаточное внимание ассоциативным вычислениям и связанным с ними так называемыми системами вертикальной обработки (СВО). То и другое можно смело отнести к «неувядающей» классике параллельных систем. По ряду специальных приложений СВО до сих пор – вне конкуренции. В другой части раздела на конкретных примерах обсуждаются особенности построения мультипроцессоров и мультикомпьютеров типов «мэн-фрейм» и «клuster», реализации параллельных систем управления консервативными базами данных на клустерной платформе. Пример параллельной СУБД наглядно иллюстрирует ограниченную масштабируемость клустерных систем.

Раздел 3 знакомит читателя с вопросами организации коммутационных сетей и главной памяти параллельных систем, с обеспечением когерентности по данным многопроцессорных систем и построением RAID-массивов – эффективным способом повышения надежности и производительности систем хранения информации.

*Раздел 4* посвящен микропрограммируемым матрицам и особенностям параллельных алгоритмов для таких матриц – во многом специализированных систем на основе операционных логико-запоминающих сред. Рассмотренные примеры конкретных разработок ориентированы на использование технологии ПЛИС. Они убедительно демонстрируют достоинства специализации и объективность относительной оценки производительности параллельных систем на множестве представительских программ.

*Раздел 5* показывает архитектурные особенности суперпроцессоров. Понятие суперпроцессора (процессора сверхвысокой производительности) было впервые «озвучено» в связи с выпуском компанией Motorola процессора PowerPC 7400, содержащего блок векторных вычислений AltiVec. Архитектура Cell – очередной скачок в развитии PowerPC. Говоря о графических процессорах GeForce и процессорах Cell BE как о революционных средствах параллельных высокопроизводительных вычислений, пришлось немалое внимание уделить их истории и основному «игровому» назначению.

*Приложение* имеет целью показать комплексный характер разработок специализированных параллельных систем на примере исследований по вопросам позитивизма маскирования бинарных объектов.

Основу книги составили материалы лекций, прочитанных автором студентам и сотрудникам Казанского государственного технического университета им. А.Н. Туполева (КАИ) в 1995-2010 г. В них отражены (с необходимыми ссылками) и результаты оригинальных исследований. Автор неизменно признателен своим соотечественникам по разработке и моделированию матричного процессора ассоциативного типа (ПМА), параллельной СУБД Clusterix, системы Security Map-Point Cluster за их творческий вклад в сделанное.

Несколько иная по компоновке версия первых трех разделов (без RAID-массивов и Clusterix, но с ПМА) уже публиковалась ранее (Райхлин В.А. Начала параллельных вычислений. – Казань: Изд-во КГТУ, 2008). Благодарю доктора М.В. Яковлевского (ИММ РАН) за ее внимательное прочтение. Его замечания по коммуникационной трудоемкости параллельных алгоритмов были учтены в процессе работы над этой книгой.

*Автор*

# СОДЕРЖАНИЕ

<b>I. НАЧАЛЬНЫЕ ПОНЯТИЯ И ПРЕДПОСЫЛКИ .....</b>	9
<b>1.1. Необходимость, ретроспектива и тенденции развития параллельных систем .....</b>	9
Адекватность параллельной обработки современным задачам ...	10
Некоторые задачи. Основа повышения производительности	
Историческая ретроспектива .....	12
Быстро действующие скалярные компьютеры. Конвейерные векторные ЭВМ. Алгоритмические матричные процессоры. Процессорные матрицы. Ортогональные и ассоциативные ЭВМ.	
Современные тенденции .....	19
Векторно-конвейерные суперкомпьютеры. SMP-системы. Numa-технологии. MPP-системы. Кластерные системы.	
<b>1.2. Классификация параллельных систем .....</b>	21
Первые классификации .....	22
Систематика М.Флина. Классификация Дж.Шора.	
Морфологическая классификация .....	25
Структурная нотация. Структурная систематика.	
<b>1.3. Показатели производительности .....</b>	31
Абстрактные оценки производительности .....	31
Производительность векторных операций. Производительность программ.	
Системная производительность .....	37
Предварительные замечания. Концепция времени маршрутизации. Относительная оценка системной производительности. Тестовые оценки производительности.	
<b>1.4. Предметные предпосылки параллелизма .....</b>	45
Цифровая обработка сигналов .....	46
Дискретное и быстрое преобразования Фурье. Алгоритм Кули-Тьюки.	
Примеры использования БПФ. Цифровая фильтрация и устройства ЦОС.	
Обработка изображений .....	54
Основные понятия. Улучшение изображений. Кодирование изображений.	
Обработка графической информации.	
Обработка символов .....	62
Обработка цепочек символов. Обработка естественных языков.	
<b>II. ПРЕДСТАВИТЕЛИ ПАРАЛЛЕЛЬНЫХ СИСТЕМ .....</b>	68
<b>2.1. Ассоциативные параллельные процессоры .....</b>	68

<b>Архитектурные основы</b>	68
Ассоциативная память. Команды АПП.	
<b>Алгоритмы для АПП</b>	71
Логические алгоритмы. Арифметические алгоритмы.	
<b>Ортогональная память</b>	76
<b>2.2. Процессорные матрицы</b>	77
<b>Общие замечания</b>	78
В порядке дискуссии. Альтернатива «СВО – СГО». Децентрализация управления	
<b>Системы вертикальной обработки</b>	83
Обобщенная структурная схема. Способ управления. Матрица процессорных элементов. Акселераторы. Внешние устройства.	
<b>2.3. Майн-фреймовые архитектуры</b>	90
<b>Примеры архитектур</b>	90
Идея Cray*. Система BBN Butterfly. Система Hewlett-Packard (HP) Superdome. Семейство Cray T3D/T3E.	
<b>2.4. Кластерные архитектуры</b>	95
<b>Особенности кластерных архитектур</b>	96
Достиоинства кластерных систем. Технологии соединения компьютеров в кластер. Beowulf-технологии. Задачи кластерных систем.	
<b>Прикладные вопросы</b>	103
Вопросы разработки параллельных программ.	
<b>Параллельные системы управления консервативными базами данных на Beowulf-кластерах</b>	108
Регулярный план обработки запросов. Принятая гипотеза. Варианты архитектуры.	
<b>III. ИНТЕРКОННЕКТ И ПАМЯТЬ</b>	114
<b>3.1. Сети коммутации</b>	114
<b>Элементы теории коммутационных сетей</b>	114
Некоторые перестановки. Коммутации в процессорных матрицах.	
<b>Сети коммутации в майнфреймах</b>	120
<b>Коммуникационная трудоемкость параллельных алгоритмов</b>	123
Некоторые понятия. Временные параметры сетевых передач. Способы коммуникации. Трудоемкость операций передачи данных. Случай кластерных систем.	
<b>3.2. Организация главной памяти</b>	127
<b>Случай ОКМД-систем</b>	128
Варианты организации. Исключение коллизий.	
<b>Память типа DDR– SDRAM</b>	130
<b>Поддержка когерентности по данным</b>	133

*Кэш-память. Понятие когерентности для данных. Случай общей памяти.  
Случай распределенной памяти. Алгоритм DASH. Масштабируемый когерентный интерфейс SCI.*

<b>3.3. RAID-массивы</b>	139
Понятие RAID	140
Уровни RAID	142
RAID 0. RAID 1 ( <i>Mirrored disk</i> ). RAID 2. RAID 3. RAID 4. RAID 5.	
Практическая реализация	148
<b>IV. МИКРОПРОГРАММИРУЕМЫЕ МАТРИЦЫ</b>	150
<b>4.1. Операционные среды</b>	150
Понятие операционной среды	151
<i>Определение. Связь с автоматами. Преобразование тактности. Случай бесконечной области. Связь с битовой обработкой.</i>	
Подход к синтезу	155
<i>Рассматриваемые виды обработки. Суть подхода.</i>	
Примеры	158
<i>Среда числового поиска. Поиск минимального покрытия.</i>	
Обсуждение	165
<i>Конструктивность рассмотренного подхода. Связь с ассоциативной обработкой. Перспективы применения.</i>	
<b>4.2. Матрицы распознавания</b>	169
Однотактное распознавание	169
Многотактное распознавание	173
<i>Параллельный алгоритм. Структура спецпроцессора. Сравнительные оценки производительности.</i>	
<b>4.3. Матричный процессор ассоциативного типа</b>	182
Архитектура	182
<i>Регистровая структура ПМА. Характеристика состава команд. Особенности обработки по полям.</i>	
Процессорный элемент	187
Функционирование и эффективность	189
<i>Особенности функционирования. Показатели эффективности.</i>	
<b>4.4. Параллельные алгоритмы для ПМА</b>	192
Сортировка	192
<i>Первый проход. Второй проход. Третий проход. Свойство кусочной линейности времени сортировки.</i>	
Распознавание двоичных образов	195
<i>Решаемая задача. Первый этап. Второй этап.</i>	
Алгоритм распознавания для последовательной ЭВМ	200
<i>Первый этап. Второй этап.</i>	

<b>Сравнительные оценки быстродействия</b>	202
<b>V. СУПЕРПРОЦЕССОРЫ</b>	205
<b>5.1. Графические процессоры</b>	205
<b>Начальные сведения по графическим процессорам</b>	205
Понятие GPU. Графический конвейер. Вехи истории. Подход GPGPU.	
<b>Классика графических процессоров</b>	210
Принципы работы классического графического процессора. Недостатки классического графического конвейера. Понятие Microsoft DirectX.	
<b>Архитектура современных графических процессоров</b>	217
Кратко о возможностях GeForce 8800 и nForce 680i SLI. Унифицированный процессор. Архитектура ядра GPU NVIDIA GeForce 8800. Технологии, поддерживаемые процессором GeForce 8800. Семейство процессоров GeForce 8800.	
<b>5.2. Cell-процессоры</b>	226
<b>Общие сведения по процессору Cell</b>	227
Вехи истории. Современная версия Cell. Высокопроизводительные системы на Cell. Программная поддержка Cell. Модели программирования для Cell.	
<b>Cell Broadband Engine Architecture, CBEA</b>	233
Power Processor Element, PPE. Sinergistic Processor Element, SPE. Element Interconnect Bus, EIB. Memory Interface Controller, MIC. Bus Interface Controller, BIC. Расширение базовой Power Architecture. Memory Flow Control, MFC.	
<b>Кластер RoadRunner</b>	240
Минимальная ячейка RoadRunner. Средний уровень RoadRunner. Верхний уровень RoadRunner.	
<b>ПРИЛОЖЕНИЕ</b>	243
<b>Позитивизм двумерно-ассоциативного маскирования</b>	243
<b>Предпосылки защиты</b>	244
Функции маскирования. Суть рассматриваемого подхода. Замечания по стойкости защиты. «Пустые» объекты. Формулировка задач защиты. Сравнение с другими методами.	
<b>Базовый алгоритм маскирования</b>	253
Формулировка алгоритма. Свойства алгоритма.	
<b>Достижимая стойкость защиты</b>	257
Оценка уровня стойкости. Оценка числа ключей.	
<b>Вопросы построения систем управления защищенными картографическими базами данных кластерного типа</b>	259
Формирование ЗКБД. Процедуры управления ЗКБД. Проект системы Security Map-Point Cluster	
<b>ЛИТЕРАТУРА</b>	263

# I. НАЧАЛЬНЫЕ ПОНЯТИЯ И ПРЕДПОСЫЛКИ

---

С понятием высокопроизводительной вычислительной системы связываются синонимы: мэйнфрейм; многопроцессорный сервер; суперЭВМ, или супер-компьютер. По одному из давних определений (Оксфордский толковый словарь по вычислительной технике, 1986) суперЭВМ (мэйнфрейм) – это «очень мощная ЭВМ производительностью выше 10 MFLOPS» (миллионов операций с плавающей точкой в секунду). В начале 90-х годов прошлого века «планка производительности» суперЭВМ была поднята до 300 MFLOPS, после 1995 г. – до 5 GFLOPS (что отвечает «планке» для современного персонального компьютера), в настоящее время – до единиц TFLOPS.

СуперЭВМ – национальное достояние, приравниваемое стратегическому оружию. Их разработка – один из приоритетов государственной политики в экономически развитых странах. По определению Госдепартамента США (конец прошлого века) компьютеры производительностью выше 10 GTOPS (TOPS – теоретических операций в секунду) считаются суперкомпьютерами и подлежат запрету на поставки в другие страны.

В этом разделе дается краткий очерк развития параллельных систем обработки информации. Рассматриваются различные подходы к их классификации и оценке производительности. Дается начальное знакомство с рядом насущных задач современности, решение которых требует использования высокопроизводительных средств.

## 1.1. НЕОБХОДИМОСТЬ, РЕТРОСПЕКТИВА И ТЕНДЕНЦИИ РАЗВИТИЯ ПАРАЛЛЕЛЬНЫХ СИСТЕМ

Помимо высокой производительности, характерными признаками суперкомпьютеров являются:

- самый высокий (для своего времени) технологический уровень;
- специфичные архитектурные решения;
- высокая цена – обычно более 1–2 млн \$ (до 50 млн \$). Но имеются и т.н. мини суперЭВМ с улучшенными показателями цена/производительность стоимостью от нескольких сотен тысяч \$;

Научно-образовательное издание

РАЙХЛИН ВАДИМ АБРАМОВИЧ

СИСТЕМЫ ПАРАЛЛЕЛЬНОЙ  
ОБРАБОТКИ ДАННЫХ

ISBN 978-5-9690-0157-2



9 785969 001572 >

Работа печатается в авторской редакции

---

Подписано в печать 14 .06.10.

Формат 60 x 84 1/16. Бумага офсетная.  
Печ.л. 16,75. Усл.печ.л. 15,8. Усл.кр.-отт. 15,8. Уч.-изд.л. 16.  
Тираж 100.

---

Издательство «Фен» Академии наук Республики Татарстан  
420044, г.Казань, ул.Волгоградская,49

Отпечатано в множительном центре  
Института истории АН РТ