

“十二五”国家重点图书出版规划项目

# 大数据技术与应用

丛书策划

上海大数据产业技术创新战略联盟（上海产业技术研究院）

上海市数据科学重点实验室（复旦大学）

丛书主编

朱扬勇 吴俊伟

# Big Data

Technology and Application Series

童维勤 黄林鹏

主编

# 数据密集型 计算和模型



上海科学技术出版社



大数据技术与应用

# 数据密集型计算和模型

童维勤 黄林鹏  
主编

上海科学技术出版社

图书在版编目(CIP)数据

数据密集型计算和模型/童维勤,黄林鹏主编. —  
上海:上海科学技术出版社,2015.1(2015.1重印)  
(大数据技术与应用)  
ISBN 978-7-5478-2269-2  
I. ①数... II. ①童... ②黄... III. ①计算模  
型 IV. ①TP301.6  
中国版本图书馆CIP数据核字(2014)第133107号

数据密集型计算和模型

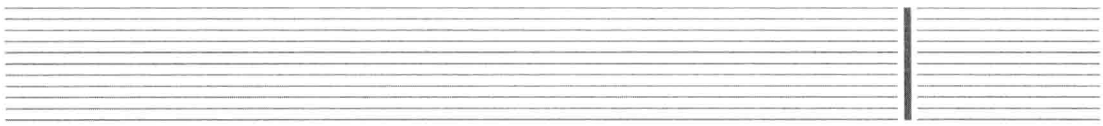
童维勤 黄林鹏 主编

上海世纪出版股份有限公司 出版  
上海科学技术出版社  
(上海钦州南路71号 邮政编码200235)

上海世纪出版股份有限公司发行中心发行  
200001 上海福建中路193号 www.ewen.co  
苏州望电印刷有限公司印刷  
开本 787×1092 1/16 印张 15  
字数:330千字  
2015年1月第1版 2015年1月第2次印刷  
ISBN 978-7-5478-2269-2/TP·26  
定价:58.00元

本书如有缺页、错装或损坏等严重质量问题,请向工厂联系调换

# 内容提要



本书涵盖了数据密集型计算的体系结构、计算模型和编程方法,内容系统全面,着重介绍原理和方法,并配以图片,便于读者理解。本书共分7章,第1章介绍数据密集型计算的概况,及其与高性能计算和云计算的异同、应用领域以及面临的挑战;第2章介绍海量数据时代的计算机组织体系和技术;第3章介绍内存计算组织体系和技术;第4章~第6章介绍几个常见的计算模型,如 MapReduce 模型、BSP 模型和 Dryad 模型等;第7章综合介绍一些专门领域的计算模型,如 All-Pairs 模型等。

本书主要读者对象是信息技术领域的从业者以及广大的计算机学科及其相关学科的师生。

大数据技术与应用

## 学术顾问



中国工程院院士 邬江兴

中国科学院院士 梅 宏

中国科学院院士 金 力

教授, 博士生导师 温孚江

教授, 博士生导师 王晓阳

教授, 博士生导师 管海兵

教授, 博士生导师 顾君忠

教授, 博士生导师 乐嘉锦

研究员 史一兵

大数据技术与应用  
**编撰委员会**



**主任**

朱扬勇 吴俊伟

**委员**

(以姓氏笔画为序)

于广军 朱扬勇 刘振宇 孙景乐 李光亚 李光耀 杨 丽  
杨佳泓 吴俊伟 何 承 张鹏翥 陈 云 武 星 黄林鹏  
童维勤 蔡立志

# 本书编委会



## 主 编

上海大学 童维勤

上海交通大学 黄林鹏

## 编 委

国防科学技术大学 谢伦国

上海大学 支小莉

上海科学院 吴俊伟

中国电信股份有限公司上海分公司 廖文昭

# 前 言

---

---

---

---

---

---

---

---

---

---

---

---

随着大数据时代的到来,适应数据密集型应用的计算模型与体系结构的研究以及平台应用技术的发展均呈现出快速增长态势。

大数据所特有的海量规模、急剧增长、移动变化和结构多样等特征,给计算机体系结构的计算、存储、网络等部件的设计提出了诸多新的挑战,众核架构、混合异构、分布存储、内存计算、软件定义基础设施等概念和技术层出不穷。蓬勃发展和不断演变的各种新技术给科学工作者和工程技术人员全面了解大数据时代的计算模型和体系结构带来了许多困难。

另外,针对不同的数据类型特点和数据处理流程,新的计算模型和新的计算平台不断涌现。目前,数据密集型计算模型繁多,那么,这些模型产生的动机是什么?它们之间又有着什么样的联系?它们分别适合于哪些类型数据的处理?从业者为了熟练掌握数据密集型计算技术需要进行大量的摸索和实践,但目前系统介绍这方面技术和知识的资料文献还比较欠缺。

本书试图阐述大数据时代数据密集型计算技术对于数据分析和处理的核心支撑作用,并系统地介绍当前主流的适合数据密集型计算的体系结构和计算模型,以使相关领域的从业者能一书在手、了然于胸,亦使广大计算机以及相关专业的教师和学生能在短时间内通过阅读本书而对信息技术的这个前沿话题有个大致了解,这既是本书的目的,也是编者的心愿。

在计算机体系结构方面,本书主要介绍计算、存储、网络等功能部件,并从软件定义基础设施的角度进行统一阐述,此外,还讨论了抽象资源管理技术,并对当前较为流行的虚拟资源管理系统进行了横向对比。本书还将重点介绍目前备受关注的内存计算技术。近年来,材料技术的巨大进步使得过去停留在概念阶段的内存计算技术得到了前



所未有的发展和应用,本书将详细阐述内存计算技术给大数据处理带来的性能提高并讨论其对系统架构和系统软件设计的影响。

在计算模型方面,本书重点介绍了成熟的 MapReduce、BSP (Bulk Synchronous Parallel)、Dryad 等计算模型并对其适用性进行分析。MapReduce 模型适用于处理用键值对表示的密集型数据,本书阐述了 MapReduce 模型的基本原理及执行机制,并详细介绍了针对 MapReduce 模型的若干改进以及基于 MapReduce 模型的编程框架。BSP 模型是一种桥接模型,用来架起并行体系结构和并行编程语言之间的桥梁,本书介绍了 BSP 模型的产生背景、基本概念、基本原理和发展概况,分析了 BSP 模型的优缺点,并重点介绍了基于 BSP 模型用于大数据分析和处理的并行编程框架。针对 Dryad 架构,本书详细讨论分布式文件系统 Cosmos、执行引擎 Dryad 以及运行在 Dryad 之上的编程语言 DryadLINQ 和 SCOPE。另外,本书第 7 章还对一些用于专门领域的新兴并行计算模型进行了介绍。

本书第 1 章~第 3 章由童维勤、黄林鹏、谢伦国、廖文昭和吴俊伟组织编写,第 4 章~第 7 章由童维勤、黄林鹏和支小莉组织编写,全书由童维勤、黄林鹏组织统稿。上海大学计算机学院的博士研究生刘晓东、高强、申超,硕士研究生林凯、胡超、张利平、万浩然、鲁圣鹏做了大量的素材收集和整理工作,在此一并致谢。

本书虽数易其稿,几经增删,但由于编者水平有限,错误和不当之处在所难免,恳请广大读者朋友批评指正。

作者

# 目 录



<b>第1章 绪论</b>	<b>1</b>
• 1.1 数据密集型计算概念	2
• 1.2 大数据时代的数据密集型计算技术	4
• 1.3 数据密集型计算与高性能计算、云计算的关系	5
1.3.1 数据密集型计算与高性能计算	5
1.3.2 数据密集型计算与云计算	6
• 1.4 数据密集型计算的应用领域	7
1.4.1 社会管理领域	7
1.4.2 教育领域	9
1.4.3 医疗领域	10
1.4.4 交通领域	11
1.4.5 农业领域	12
1.4.6 金融领域	13
1.4.7 其他领域	14
• 1.5 大数据带来的挑战	15
1.5.1 大数据集成	16
1.5.2 大数据分析	16
1.5.3 计算机体系结构面临的挑战	17

1.5.4	编程模型面临的挑战	18
	参考文献	18
<b>第2章 大数据时代的计算机体系结构</b>		<b>21</b>
<hr/>		
• 2.1	计算部件	22
2.1.1	多核	22
2.1.2	众核	23
2.1.3	GPU+CPU 混合异构	25
2.1.4	集群	27
<hr/>		
• 2.2	存储部件	28
2.2.1	片上存储	29
2.2.2	本地存储	29
2.2.3	分布式存储	31
<hr/>		
• 2.3	网络部件	33
2.3.1	片上通信	34
2.3.2	数据中心网络	35
2.3.3	数据中心互联网络	39
<hr/>		
• 2.4	软件定义部件	42
2.4.1	软件定义计算	42
2.4.2	软件定义存储	46
2.4.3	软件定义网络	53
2.4.4	软件定义数据中心	56
<hr/>		
• 2.5	虚拟资源管理系统	60
2.5.1	典型虚拟资源管理系统	60
2.5.2	云计算一体机	70
	参考文献	71
<b>第3章 内存计算</b>		<b>73</b>
<hr/>		
• 3.1	内存计算的概念	74

• 3.2	内存计算的硬件结构	75
3.2.1	用于内存计算的专用内存	76
3.2.2	用于内存计算的计算机体系结构	79
• 3.3	内存计算的系统软件	88
3.3.1	内存文件系统	88
3.3.2	任务及数据的调度	98
• 3.4	内存数据库	102
3.4.1	内存数据库的结构	102
3.4.2	内存数据库的关键技术	103
3.4.3	主流内存数据库及其优缺点	107
	参考文献	111
<b>第4章</b>	<b>MapReduce 模型</b>	<b>115</b>
• 4.1	MapReduce 模型简介	116
4.1.1	MapReduce 模型概念及原理	116
4.1.2	MapReduce 模型工作机制	117
4.1.3	MapReduce 模型优缺点	124
• 4.2	基于 MapReduce 模型的实现	124
4.2.1	Hadoop	125
4.2.2	Phoenix	129
4.2.3	其他实现	131
• 4.3	MapReduce 模型的改进	133
4.3.1	Spark	133
4.3.2	Data Freeway 和 Puma	135
4.3.3	Storm	137
4.3.4	Nephele/PACTs	139
4.3.5	其他改进工作	140
	参考文献	143

<b>第5章 BSP 模型</b>	<b>147</b>
• 5.1 BSP 模型简介	148
5.1.1 BSP 模型概念	148
5.1.2 BSP 模型原理	148
5.1.3 BSP 模型优缺点	149
• 5.2 BSP 模型发展概况	149
5.2.1 BSP 模型初级阶段	149
5.2.2 多核 BSP	152
5.2.3 BSP 模型在云平台上的应用	152
5.2.4 BSP 模型在大数据时代的应用	153
• 5.3 基于 BSP 模型的编程框架	154
5.3.1 Pregel	154
5.3.2 HAMA	163
5.3.3 GPS	167
5.3.4 Giraph	172
参考文献	174
<b>第6章 Dryad 模型</b>	<b>177</b>
• 6.1 Dryad 简介	178
6.1.1 Dryad 系统概述	178
6.1.2 Dryad 图描述	179
6.1.3 Dryad 执行机制	181
• 6.2 SCOPE 脚本语言	182
• 6.3 DryadLINQ	184
• 6.4 Cosmos	186
6.4.1 Cosmos 存储系统	187
6.4.2 Cosmos 执行环境	187
• 6.5 MapReduce 与 Dryad 的比较	188
参考文献	188

---

<b>第7章 其他计算模型</b>	191
• 7.1 All - Pairs	192
7.1.1 All - Pairs 简介	192
7.1.2 All - Pairs 的应用	193
7.1.3 All - Pairs 面临的挑战	193
7.1.4 All - Pairs 的具体实现	194
• 7.2 DOT	196
7.2.1 DOT 简介	196
7.2.2 DOT 模型	197
7.2.3 DOT 展望	200
• 7.3 Pig Latin	200
7.3.1 Pig Latin 简介	201
7.3.2 Pig Latin 的功能及原理	201
7.3.3 Pig 数据模型	203
7.3.4 Pig 实现	205
• 7.4 GraphLab	206
7.4.1 GraphLab 出现背景	207
7.4.2 GraphLab 特性	207
7.4.3 GraphLab 框架	208
• 7.5 工作流	209
7.5.1 工作流简介	209
7.5.2 CloudWF 的系统设计	211
7.5.3 工作流的特性	214
参考文献	214
<b>附录 英文缩略语</b>	217

---

# 第1章

## 绪论

人类社会的快速发展使得反映人类生活和生产活动的的数据急剧增加,同时数据类型也日渐丰富。如何对海量的、非结构化的和快速增长的数据进行分析和处理成为一个亟待解决的问题。数据密集型计算作为大数据处理的一种核心支撑技术得到了广泛的应用,它基于并行计算模型和高性能计算机系统对海量数据进行分析处理,以满足数据应用的需求。本章将从数据密集型计算的概念开始,对大数据时代的数据密集型计算技术、数据密集型计算与高性能计算和云计算的关系、数据密集型计算的应用领域,以及大数据带来的挑战等内容进行介绍。

## 1.1 数据密集型计算概念

并行应用可以划分为计算密集型和数据密集型两类。计算密集型应用是指那些需要大量计算资源和计算的应用,这类应用往往处理的数据量较小,但需要大量的计算,它们把大部分执行时间花费在计算上而不是数据的输入/输出上。在对计算密集型应用进行并行处理时,首先需要对应用的算法并行化,之后整个处理过程将被分解成多个独立的任务,这些任务将在计算平台上并行执行,以达到比串行处理更高的性能。在计算密集型应用并行化运行过程中,多个计算作业同时进行,每个计算作业解决问题的一小部分,这通常被称作函数并行化或控制并行化<sup>[1]</sup>。

数据密集型应用是指那些需要处理大量数据和大量输入/输出的应用,这类应用要处理的数据量较大,它们把大部分执行时间用于数据的处理和输入/输出上。在对数据密集型应用进行并行处理时,首先需要把大量数据分割成多个小数据块,每块数据用统一的并行算法在计算平台上执行,最后每个数据块得出的执行结果将被合成为一个完整的输出结果<sup>[2]</sup>。因此,数据密集型应用中最重要的问题包括:怎样管理和处理以指数速度增长的数据、怎样缩短对大量数据分析的周期以支持实时应用,以及如何开发支持大量数据检索、处理的新算法<sup>[3]</sup>。

这两类应用可以用图 1-1 表示<sup>[4]</sup>。其中,右下象限表示计算密集型问题,大型的计算密集型应用目前主要使用高性能计算机来求解。而数据密集型计算的研究则主要针对图中上半部分两个象限的问题展开。与传统的高性能计算相比,数据密集型计算不仅要具备存储大规模数据集和高速 I/O 传输的能力,还要处理复杂的数据计算、数据分析和数据可视化问题。

数据密集型计算(Data Intensive Computing, DIC)目前尚未统一定义。在《数据密集



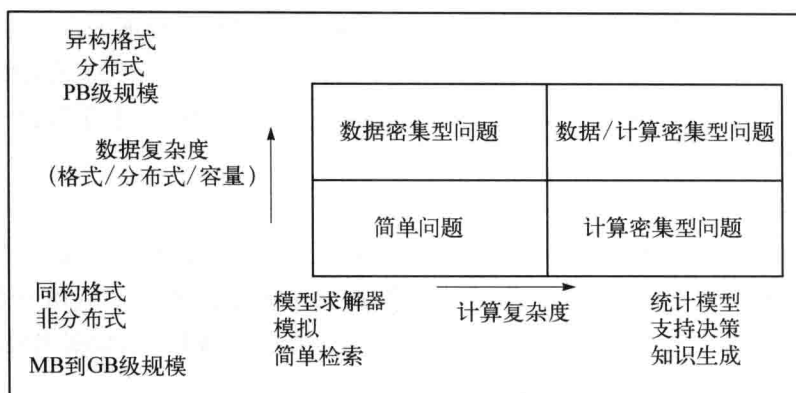


图 1-1 数据/计算密集型计算

型计算手册》这本书中,数据密集型计算是指能推动前沿技术发展的对海量和高速变化的数据的获取、管理、分析和理解<sup>[5]</sup>。维基百科(Wikipedia)将数据密集型计算定义为采用数据并行方法实现数量达 TB(太字节,1 TB=10<sup>12</sup> BYTE)或 PB(拍字节,1 PB=10<sup>15</sup> BYTE)级数据并行计算应用。本书如下定义数据密集型计算:

数据密集型计算是用可扩展的并行方式来解决大量数据的即时处理,以及支持这种并行处理的有关软硬件技术,包括算法、计算模型、软件平台、硬件体系结构和系统等。数据密集型计算是大数据的核心支撑技术。

数据密集型计算的特征体现在复杂的数据与计算的多样性、复杂的编程模型、非本地计算等方面,主要包括以下四点<sup>[6]</sup>:

### 1) 数据密集型计算处理的数据是大规模、快速变化、分布式和异构的

数据密集型应用的数据可达到 PB 级,所以传统的数据存储、索引技术已经不再适用。即使是执行非常简单的查询操作,在数据密集型计算下也会变得十分复杂,例如以 80 MB/s 的速度从 1 GB(吉字节,1GB=10<sup>9</sup> BYTE)数据中顺序扫描需要 12.5 s。如果数据量变成 1 PB,那么就需要 145 d,这对于任何一个计算任务来说都无法接受。这些数据呈现出地理上分布的、数据模型异构的、表达方式多样的等特征,这些都给访问、整合造成了诸多不便。数据的源头各不相同导致数据格式的多样性和异构性,也导致除了传统的结构化数据存在以外,半结构化和非结构化数据也大量涌现。另外,这些数据的快速变化对数据实时处理提出了要求,传统的静态数据库管理技术很难满足需求。

### 2) 计算具有多种含义

“计算”包括从数据管理到分析、理解的全过程<sup>[7]</sup>。一些科学研究领域,例如地球科学、天文学,会使用非常复杂的计算模型。在这些计算过程中,问题的局部优化计算和数据管理的计算复杂度对计算模型提出了新的挑战。

数据密集型计算中的数据处理操作除了包括搜索、查询和其他传统数据处理操作以外,还包括数据分析理解等智能处理操作。数据的分析理解过程不只表示简单的数据分析