《深入理解大数据：大数据处理与编程实践》

图书基本信息

书名：《深入理解大数据：大数据处理与编程实践》

13位ISBN编号：9787111473256

出版时间：2014-8

作者：主编：黄宜华,副主编：苗凯翔

页数：520

《深入理解大数据：大数据处理与编程实践》

内容概要

【内容简介】
本书在总结多年来MapReduce并行处理技术课程教学经验和成果的基础上，与业界著名企业Intel公司的大数据技术和产品开发团队和资深工程师联合，以学术界的教学成果与业界高水平系统研发经验完美结合，在理论联系实际的基础上，在基础理论原理、实际算法设计方法以及业界深度技术三个层面上，精心组织材料编写而成。
全书的主要内容包括：
■ 大数据处理技术与Hadoop MapReduce简介
■ Hadoop系统的安装和操作管理
■ 大数据分布式文件系统HDFS
■ Hadoop MapReduce并行编程模型、框架与编程接口
■ 分布式数据库HBase
■ 分布式数据仓库Hive
■ Intel Hadoop系统优化与功能增强
■ MapReduce基础算法程序设计
■ MapReduce高级程序设计技术
■ MapReduce机器学习与数据挖掘基础算法
■ 大数据处理算法与应用编程案例
本书中算法设计章节的程序源码可在南京大学PASA大数据实验室（PASA：Parallel Algorithms，Systems，and Applications）网站上下载：
http://pasa-bigdata.nju.edu.cn/links.html
Intel Hadoop系统免费试用版下载地址：
http://www.intel.cn/idh
本书反馈意见发送邮箱：
feedback_bigdata@163.com。
【编辑推荐】
学术界与业界完美结合的结晶，从原理剖析到系统化算法设计与编程实践
多年来系统性教学实践和成果总结，一系列业界产品增强功能深度技术剖析
一系列大赛获奖算法、优秀课程设计以及来自科研课题及业界应用的实战案例
【媒体推荐】
从计算技术的角度看，大数据处理是一种涉及到几乎所有计算机技术层面的综合性计算技术，涉及到计算机软硬件技术的方方面面。大数据研究和应用已成为产业升级与新产业崛起的重要推动力量。
作为国内第一本经过多年课堂教学实践总结而成的大数据并行处理和编程技术书籍，本书全面地介绍了大数据处理相关的基本概念和原理，着重讲述了Hadoop MapReduce大数据处理系统的组成结构、工作原理和编程模型，分析了基于MapReduce的各种大数据并行处理算法和程序设计的思想方法。适合高等院校作为MapReduce大数据并行处理技术课程的教材，同时也很适合作为大数据处理应用开发和编程专业技术人员的参考手册。
我很高兴地看到，该书已纳入了教育部计算机类专业教学指导委员会制定的计算机类专业系统能力培养计划。大数据处理是一门综合性、最能体现计算机系统能力培养的课程。把大数据处理纳入计算机类专业系统能力培养课程体系中第三层次的核心课程，作为一门起到一定“收官”作用的综合性课程，这是在计算机系统能力培养方面的一个很好的尝试。
——中国工程院院士、中国计算机学会大数据专家委员会主任　李国杰

作为国内最早从事大数据技术研究和教学的团队之一，南京大学黄宜华教授和他的大数据实验室同仁们在大数据技术领域已经进行了多年系统深入的研究工作，取得了卓有成效的研究成果。英特尔作为一家全球领先的计算技术公司，长期以来始终以计算技术的创新为己任。在大数据处理技术方面，我们也竭尽全力发挥出我们在软硬件平台的组合优势引领大数据技术的全面发展和推广。
这本《深入理解大数据》的力作正是我们双方在大数据领域共同努力的结晶，是以学术界和业界完美

结合的方式，在融合了学术界系统化的研究教学工作和业界深度的系统和应用研发工作基础上，成功打造出的一本大数据技术佳作。相信这是一本适合软件技术人员和 IT 行业管理人员理解和掌握大数据技术的不可多得的技术书籍，也是一本适合于在校大学生和研究生学习和掌握大数据处理和编程技术的好教材。

—— 英特尔亚太研发有限公司总经理　何京翔

《深入理解大数据：大数据处理与编程实践》

作者简介

黄宜华博士，南京大学计算机科学与技术系教授、PASA大数据实验室学术带头人。中国计算机学会大数据专家委员会委员、副秘书长，江苏省计算机学会大数据专家委员会主任。于1983、1986和1997年获得南京大学计算机专业学士、硕士和博士学位。主要研究方向为大数据并行处理、云计算以及Web信息挖掘等，发表学术研究论文60多篇。2010年在Google公司资助下在本校创建并开设了“ MapReduce大数据并行处理技术 ”课程，成为全国最早开设该课程的院校之一。因在该课程教学和人才培养方面的出色成绩获得2012年Google奖教金。目前正在开展系统化的大数据并行处理技术研究工作，主持国家和省部级科研项目以及与美国Intel公司等业界的合作研究项目多项。
苗凯翔 (Kai X. Miao) 博士，英特尔中国大数据首席技术官，中国计算机学会大数据专家委员会委员。曾担任英特尔中国区系统集成部总监、信息技术研究部门亚洲地区总监、英特尔北美地区解决方案首席架构师。于2009荣获英特尔公司首席工程师职称。在加入英特尔以前，曾在美国Rutgers与DeVry大学任教。获得北方交通大学（北京）通信学士学位、美国辛辛那提大学电机工程硕士和博士学位。发表期刊和会议研究论文多篇，并拥有21项美国专利，在各种会议上发表过上百次主题演讲，曾参与IETF、ITU 和 MIT CFP等工业标准的制定，并于2006 年担任IEEE通信杂志的联合编辑。

《深入理解大数据：大数据处理与编程实践》

书籍目录

《深入理解大数据：大数据处理与编程实践》

精彩短评

1、勉强算是读完了吧，以后也懒得碰了。书很烂，要么就是介绍安装配置，要么就是贴代码。并没有深入介绍Hadoop的系统设计（介绍的也都是官方文档上的内容），不利于读者理解。另外，Hadoop一代已经过时了啊，Intel的Hadoop发行版已经放弃维护了吧？
2、推荐研究生看看

《深入理解大数据：大数据处理与编程实践》

版权说明