

《大数据技术原理与应用》

图书基本信息

书名：《大数据技术原理与应用》

13位ISBN编号：9787115392870

出版时间：2015-8-1

作者：林子雨

页数：250

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu111.com

《大数据技术原理与应用》

内容概要

大数据作为继云计算、物联网之后IT行业又一颠覆性的技术，备受关注。大数据处不在，包括金融、汽车、零售、餐饮、电信、能源、政务、医疗、体育、娱乐等在内的社会各行各业，都融入了大数据的印迹，大数据对人类的社会生产和生活必将产生重大而深远的影响。

大数据时代的到来，迫切需要高校及时建立大数据技术课程体系，为社会培养和输送一大批具备大数据专业素养的高级人才，满足社会对大数据人才日益旺盛的需求。本书定位为大数据技术入门教材，为读者搭建起通向“大数据知识空间”的桥梁和纽带。本书将系统梳理总结大数据相关技术，介绍大数据技术的基本原理和大数据主要应用，帮助读者形成对大数据知识体系及其应用领域的轮廓性认识，为读者在大数据领域“深耕细作”奠定基础、指明方向。在本书的基础上，感兴趣的读者可以通过其他诸如《Hadoop权威指南》等工具书，继续深入学习和实践大数据相关技术。

《大数据技术原理与应用》

作者简介

林子雨，博士，厦门大学计算机科学系老师，中国高校首个“数字教师”的提出者和建设者。在数据库、数据仓库、数据挖掘、大数据、云计算和物联网等领域有着十多年的知识积累，对各个领域知识都有比较深入的了解，有比较宽泛的视野，将相关大数据知识综合成一本适合本科和研究生教学的教材。

书籍目录

第一篇大数据基础

第1章大数据概述

1.1 大数据时代

1.1.1 第三次信息化浪潮

1.1.2 信息科技为大数据时代

提供技术支撑

1.1.3 数据产生方式的变革促成

大数据时代的来临

1.1.4 大数据的发展历程

1.2 大数据的概念

1.2.1 数据量大

1.2.2 数据类型繁多

1.2.3 处理速度快

1.2.4 价值密度低

1.3 大数据的影响

1.3.1 大数据对科学研究的影响

1.3.2 大数据对思维方式的影响

1.3.3 大数据对社会发展的影响

1.3.4 大数据对就业市场的影响

1.3.5 大数据对人才培养的影响

1.4 大数据的应用

1.5 大数据关键技术

1.6 大数据计算模式

1.6.1 批处理计算

1.6.2 流计算

1.6.3 图计算

1.6.4 查询分析计算

1.7 大数据产业

1.8 大数据与云计算、物联网

1.8.1 云计算

1.8.2 物联网

1.8.3 大数据与云计算、物联网的关系

1.9 本章小结

1.10 习题

第2章大数据处理架构Hadoop

2.1 概述

2.1.1 Hadoop简介

2.1.2 Hadoop的发展简史

2.1.3 Hadoop的特性

2.1.4 Hadoop的应用现状

2.2 Hadoop的项目结构

2.2.1 Common

2.2.2 Avro

2.2.3 HDFS

2.2.4 HBase

2.2.5 MapReduce

2.2.6 Zookeeper

- 2.2.7 Hive
- 2.2.8 Pig
- 2.2.9 Sqoop
- 2.2.10 Chukwa
- 2.3 Hadoop的安装与使用
 - 2.3.1 创建Hadoop用户
 - 2.3.2 Java的安装
 - 2.3.3 SSH登录权限设置
 - 2.3.4 安装单机Hadoop
 - 2.3.5 Hadoop伪分布式安装
- 2.4 本章小结
- 2.5 习题
- 第二篇大数据存储
- 第3章 Hadoop分布式文件系统
 - 3.1 分布式文件系统
 - 3.1.1 计算机集群结构
 - 3.1.2 分布式文件系统的结构
 - 3.1.3 分布式文件系统的设计需求
 - 3.2 HDFS简介
 - 3.3 HDFS的相关概念
 - 3.3.1 块
 - 3.3.2 名称节点和数据节点
 - 3.4 HDFS体系结构
 - 3.4.1 概述
 - 3.4.2 HDFS命名空间管理
 - 3.4.3 通信协议
 - 3.4.4 客户端
 - 3.4.5 HDFS体系结构的局限性
 - 3.5 HDFS的存储原理
 - 3.5.1 冗余数据的保存
 - 3.5.2 数据存取策略
 - 3.5.3 数据错误与恢复
 - 3.6 HDFS的数据读写过程
 - 3.6.1 读数据的过程
 - 3.6.2 写数据的过程
 - 3.7 HDFS编程实践
 - 3.7.1 HDFS常用命令
 - 3.7.2 HDFS的Web界面
 - 3.7.3 HDFS常用Java API及应用实例
 - 3.8 本章小结
 - 3.9 习题
- 第4章 分布式数据库HBase
 - 4.1 概述
 - 4.1.1 从BigTable说起
 - 4.1.2 HBase简介
 - 4.1.3 HBase与传统关系数据库的对比分析
 - 4.2 HBase访问接口
 - 4.3 HBase数据模型

- 4.3.1 数据模型概述
- 4.3.2 数据模型的相关概念
- 4.3.3 数据坐标
- 4.3.4 概念视图
- 4.3.5 物理视图
- 4.3.6 面向列的存储
- 4.4 HBase的实现原理
 - 4.4.1 HBase的功能组件
 - 4.4.2 表和Region
 - 4.4.3 Region的定位
- 4.5 HBase运行机制
 - 4.5.1 HBase系统架构
 - 4.5.2 Region服务器的工作原理
 - 4.5.3 Store的工作原理
 - 4.5.4 HLog的工作原理
- 4.6 HBase编程实践
 - 4.6.1 HBase常用的Shell命令
 - 4.6.2 HBase常用的JavaAPI及应用实例
- 4.7 本章小结
- 4.8 习题
- 第5章 NoSQL数据库
 - 5.1 NoSQL简介
 - 5.2 NoSQL兴起的原因
 - 5.2.1 关系数据库无法满足Web2.0的需求
 - 5.2.2 关系数据库的关键特性在Web2.0时代成为"鸡肋"
 - 5.3 NoSQL与关系数据库的比较
 - 5.4 NoSQL的四大类型
 - 5.4.1 键值数据库
 - 5.4.2 列族数据库
 - 5.4.3 文档数据库
 - 5.4.4 图形数据库
 - 5.5 NoSQL的三大基石
 - 5.5.1 CAP
 - 5.5.2 BASE
 - 5.5.3 最终一致性
 - 5.6 从NoSQL到NewSQL数据库
 - 5.7 本章小结
 - 5.8 习题
- 第6章 云数据库
 - 6.1 云数据库概述
 - 6.1.1 云计算是云数据库兴起的基础
 - 6.1.2 云数据库的概念
 - 6.1.3 云数据库的特性
 - 6.1.4 云数据库是个性化数据存储需求的理想选择

- 6.1.5 云数据库与其他数据库的关系
 - 6.2 云数据库产品
 - 6.2.1 云数据库厂商概述
 - 6.2.2 Amazon的云数据库产品
 - 6.2.3 Google的云数据库产品
 - 6.2.4 微软的云数据库产品
 - 6.2.5 其他云数据库产品
 - 6.3 云数据库系统架构
 - 6.3.1 UMP系统概述
 - 6.3.2 UMP系统架构
 - 6.3.3 UMP系统功能
 - 6.4 云数据库实践
 - 6.4.1 阿里云RDS简介
 - 6.4.2 RDS中的概念
 - 6.4.3 购买和使用RDS数据库
 - 6.4.4 将本地数据库迁移到云端RDS数据库
 - 6.5 本章小结
 - 6.6 习题
- 第三篇大数据处理与分析
- 第7章 MapReduce
- 7.1 概述
 - 7.1.1 分布式并行编程
 - 7.1.2 MapReduce模型简介
 - 7.1.3 Map和Reduce函数
 - 7.2 MapReduce的工作流程
 - 7.2.1 工作流程概述
 - 7.2.2 MapReduce的各个执行阶段
 - 7.2.3 Shuffle过程详解
 - 7.3 实例分析：WordCount
 - 7.3.1 WordCount的程序任务
 - 7.3.2 WordCount的设计思路
 - 7.3.3 MapReduce的具体执行过程
 - 7.3.4 一个WordCount执行过程的实例
 - 7.4 MapReduce的具体应用
 - 7.4.1 MapReduce在关系代数运算中的应用
 - 7.4.2 分组与聚合运算
 - 7.4.3 矩阵-向量乘法
 - 7.4.4 矩阵乘法
 - 7.5 MapReduce编程实践
 - 7.5.1 任务要求
 - 7.5.2 编写Map处理逻辑
 - 7.5.3 编写Reduce处理逻辑
 - 7.5.4 编写main方法
 - 7.5.5 编译打包代码以及运行程序
 - 7.6 本章小结
 - 7.7 习题
- 第8章 流计算

- 8.1 流计算概述
 - 8.1.1 静态数据和流数据
 - 8.1.2 批量计算和实时计算
 - 8.1.3 流计算的概念
 - 8.1.4 流计算与Hadoop
 - 8.1.5 流计算框架
- 8.2 流计算的处理流程
 - 8.2.1 概述
 - 8.2.2 数据实时采集
 - 8.2.3 数据实时计算
 - 8.2.4 实时查询服务
- 8.3 流计算的应用
 - 8.3.1 应用场景1：实时分析
 - 8.3.2 应用场景2：实时交通
- 8.4 开源流计算框架Storm
 - 8.4.1 Storm简介

《大数据技术原理与应用》

精彩短评

- 1、非常不错的一本大数据入门书，林子雨老师治学还是非常严谨的，配套的网站和ppt也有很多非常有用的资料，很赞！
- 2、大数据101
- 3、big data era1

《大数据技术原理与应用》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu111.com