

# 《大数据：技术与应用实践指南（第2版）》

## 图书基本信息

书名：《大数据：技术与应用实践指南（第2版）》

13位ISBN编号：9787121281503

出版时间：2016-3

作者：赵刚

页数：316

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.tushu000.com](http://www.tushu000.com)

# 《大数据：技术与应用实践指南（第2版）》

## 内容概要

大数据是互联网、移动应用、社交网络和物联网等技术发展的必然趋势，大数据应用成为当前最为热门的信息技术应用领域。《大数据：技术与应用实践指南（第2版）》由浅入深，首先概述性地分析了大数据的发展背景、基本概念，从业务的角度分析了大数据应用的主要业务价值和业务需求，在此基础上介绍大数据的技术架构和关键技术，结合应用实践，详细阐述了传统信息系统与大数据平台的整合策略，大数据应用实践的流程和方法，并介绍了主要的大数据应用产品和解决方案。最后，对大数据面临的挑战和未来的趋势进行了展望。

《大数据：技术与应用实践指南（第2版）》既具有技术深度，又具有很强的可操作性，提供了一个系统性、架构性的大数据应用实践指南，概要性地指导大数据应用实践，推动大数据技术在各个行业的广泛应用。

## 书籍目录

### 第1章 大数据的概念和发展背景

- 1
  - 1.1 大数据的发展背景
  - 1
  - 1.2 大数据的概念和特征
  - 4
  - 1.2.1 大数据的概念
  - 4
  - 1.2.2 大数据的特征
  - 4
  - 1.3 大数据的产生
  - 5
  - 1.3.1 数据产生由企业外部向企业内部扩展
  - 5
  - 1.3.2 数据产生从Web 1.0向Web 2.0、从互联网向移动互联网扩展
  - 6
  - 1.3.3 数据产生从计算机/互联网（IT）向物联网（IOT）扩展
  - 7
  - 1.4 数据的量级
  - 7
  - 1.4.1 数据大小的量级
  - 7
  - 1.4.2 大数据的量级
  - 8
  - 1.5 大量不同的数据类型
  - 8
  - 1.5.1 按照数据结构分类
  - 9
  - 1.5.2 按照产生主体分类
  - 12
  - 1.5.3 按照数据作用方式分类
  - 13
  - 1.6 大数据的速度
  - 14
  - 1.7 大数据的应用价值
  - 14
  - 1.8 大数据的挑战
  - 15
  - 1.8.1 业务视角不同带来的挑战
  - 15
  - 1.8.2 技术架构不同带来的挑战
  - 15
  - 1.8.3 管理策略不同带来的挑战
  - 16
- ### 第2章 大数据应用的业务需求
- 17
  - 2.1 大数据应用的业务流程

17	
2.1.1 产生数据	18
2.1.2 聚集数据	18
2.1.3 分析数据	19
2.1.4 利用数据	19
2.2 大数据应用的业务价值	19
2.2.1 发现大数据的潜在价值	20
2.2.2 发现动态行为数据的价值	20
2.2.3 实现大数据整合创新的价值	20
2.3 各行业大数据应用的个性需求	21
2.3.1 互联网与电子商务行业	21
2.3.2 零售业	26
2.3.3 金融业	28
2.3.4 政府	31
2.3.5 医疗业	34
2.3.6 能源业	35
2.3.7 制造业	37
2.3.8 电信运营	38
2.3.9 交通业	40
2.4 企业级大数据应用的共性需求	42
2.4.1 客户分析	42
2.4.2 绩效分析	46
2.4.3 欺诈和风险评估	47
2.5 以银行客户分析为例，分析一个大数据的应用场景	48
第3章 大数据应用的总体架构和关键技术	51

3.1 总体架构	51
3.1.1 业务目标	51
3.1.2 架构设计原则	52
3.1.3 总体架构参考模型	55
3.1.4 总体架构的特点	58
3.2 大数据存储和处理技术	59
3.2.1 Hadoop：分布式存储和计算平台	59
3.2.2 HDFS：分布式文件系统	65
3.2.3 MapReduce：分布式计算框架	72
3.2.4 NoSQL：分布式数据库	98
3.2.5 MPP：大规模并行处理系统	113
3.2.6 Spark：轻量级的分布式内存计算系统	117
3.2.7 S4和Storm：流计算框架	126
3.2.8 大数据存储和处理技术的比较分析	132
3.3 大数据查询和分析技术	133
3.3.1 Hive：基本的Hadoop查询和分析	134
3.3.2 Hive 2.0：Hive的优化和升级	144
3.3.3 实时互动的SQL：Impala和drill	147
3.3.4 基于PostgreSQL的SQL on Hadoop	153
3.4 大数据高级分析和可视化技术	154
3.4.1 传统数据仓库与联机分析处理技术	154
3.4.2 大数据对传统分析的挑战	157
3.4.3 大数据挖掘与高级分析	157
3.4.4 大数据挖掘与高级分析库	162
3.4.5 非结构化复杂数据分析	

163
3.4.6 实时预测分析
170
3.4.7 开源可视化工具：R语言
177
3.4.8 可视化技术
185
3.5 以银行客户分析为例的大数据应用体系架构
194
第4章 大数据与企业级应用的整合策略
196
4.1 大数据传输、接入、整合和流程管理平台
197
4.1.1 数据传输
197
4.1.2 数据接入
203
4.1.3 数据整合
207
4.1.4 流程管理
208
4.2 大数据与存储架构的整合
212
4.2.1 传统存储架构比较
212
4.2.2 大数据平台的存储架构的选择
214
4.2.3 集群存储的发展
214
4.2.4 基于HDFS的集群存储
216
4.2.5 固态硬盘（SSD）对内存计算的支持
218
4.2.6 软件定义存储（SDS）
218
4.2.7 超融合架构（HCI）
220
4.3 大数据与网络架构的发展
220
4.3.1 统一的以太网结构
222
4.3.2 软件定义网络（SDN）
223
4.3.3 网络功能虚拟化（NFV）
226
4.4 大数据与虚拟化技术的整合
228
4.5 大数据与Docker技术
230

4.5.1 Docker概述	230
4.5.2 Docker原理与总体架构	231
4.5.3 Docker与应用程序开发与管理	237
4.6 大数据与云计算	240
4.7 大数据安全	242
4.8 以银行客户分析为例，分析一个大数据的平台整合	244
第5章 大数据应用的实践方法与案例	246
5.1 实践方法论	246
5.1.1 业务需求定义	247
5.1.2 数据应用现状分析与标杆比较	248
5.1.3 大数据应用架构规划和设计	249
5.1.4 大数据技术切入与实施	250
5.1.5 大数据试用和评估	251
5.1.6 大数据应用推广	252
5.2 技术应用案例	252
5.2.1 Amazon和Google	252
5.2.2 Yahoo	255
5.2.3 Amazon	257
5.2.4 Facebook	259
5.2.5 Twitter	263
5.2.6 淘宝网	264
5.3 以银行客户分析为例的实施案例分析	266
5.3.1 银行基于大数据的客户分析的业务需求	266
5.3.2 银行基于大数据的客户分析的现状与标杆比较	267
5.3.3 银行基于大数据的客户分析的应用架构规划与设计	

269	
5.3.4	银行基于大数据的数据分析的实施、试点和推广
269	
第6章	大数据应用的主流解决方案
270	
6.1	产业链
270	
6.1.1	国际大数据产业生态
270	
6.1.2	国内大数据产业生态
273	
6.2	主流厂商解决方案
274	
6.2.1	Cloudera
275	
6.2.2	Hortonworks
276	
6.2.3	MapR
277	
6.2.4	IBM
278	
6.2.5	Oracle
280	
6.2.6	EMC
281	
6.2.7	Intel
282	
6.2.8	SAP
283	
6.2.9	Teradata
285	
第7章	大数据应用的未来挑战和趋势
286	
7.1	隐私保护
286	
7.1.1	法律保护
287	
7.1.2	技术保护
289	
7.1.3	理念革新
290	
7.2	技术标准
291	
7.2.1	ISO大数据标准化进展
291	
7.2.2	大数据基准和基准测试
293	
7.2.3	大数据处理分析标准套件
296	

## 7.3 大数据治理

296

### 7.3.1 数据治理框架

297

### 7.3.2 数据质量管理

298

### 7.3.3 大数据的组织、角色和责任

299

## 7.4 适应商业社会的未来趋势

300

### 7.4.1 从产品推销向数据营销的转变

300

### 7.4.2 从流程驱动到分析驱动的转变

300

### 7.4.3 从私有资源到公共服务的转变

301

## 精彩短评

1、这也能出书？

## 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.tushu000.com](http://www.tushu000.com)