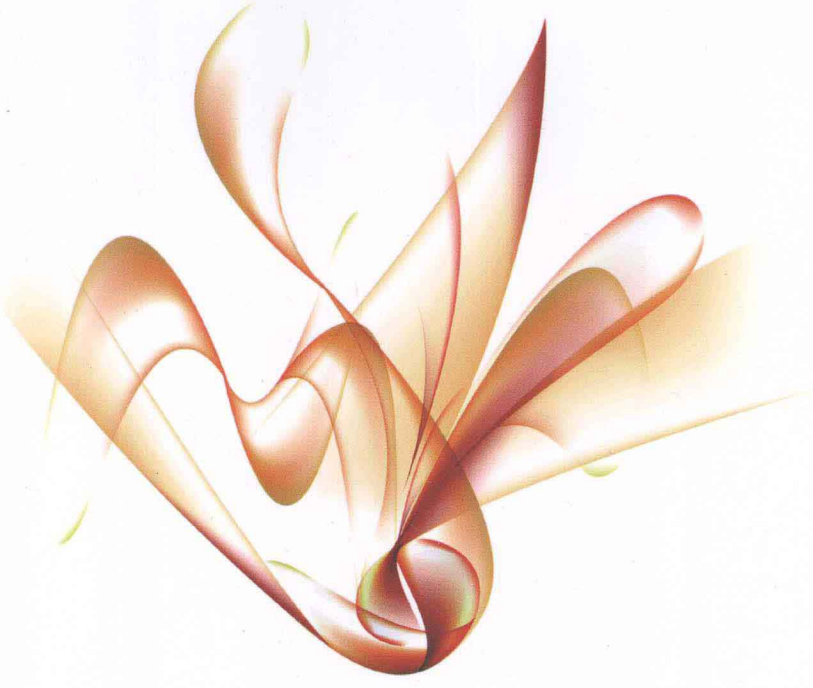
HZ BOOKS
华章科技

资深Hadoop技术专家撰写，从开发者角度系统深入地讲解了Hadoop分布式文件系统、Hadoop文件I/O、Hive、HBase、Mahout，以及MapReduce的工作原理、编程方法和高级应用

内容细致，包含大量基于实际生产环境的案例，实战性强

大数据技术丛书



Programming Hadoop

Hadoop应用开发技术详解

刘 刚◎著

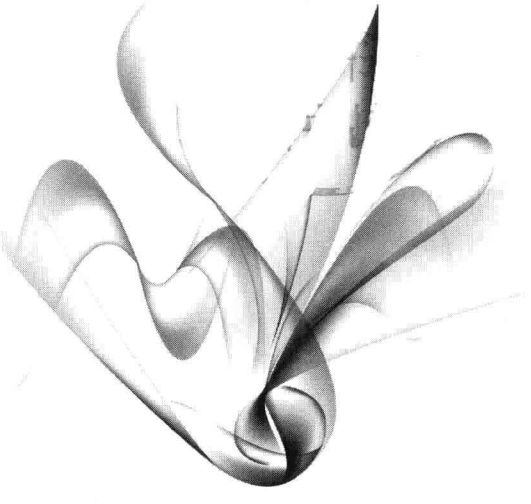
机械工业出版社
China Machine Press

技术丛书

Programming Hadoop

Hadoop应用开发技术详解

刘 刚◎著



机械工业出版社

图书在版编目（CIP）数据

Hadoop 应用开发技术详解 / 刘刚著. — 北京：机械工业出版社，2014.1
（大数据技术丛书）

ISBN 978-7-111-45244-7

I. H… II. 刘… III. 数据处理软件—程序设计 IV. TP274

中国版本图书馆 CIP 数据核字（2013）第 309446 号

本书由资深 Hadoop 技术专家撰写，系统、全面、深入地讲解了 Hadoop 开发者需要掌握的技术和知识，包括 HDFS 的原理和应用、Hadoop 文件 I/O 的原理和应用、MapReduce 的原理和高级应用、MapReduce 的编程方法和技巧，以及 Hive、HBase 和 Mahout 等技术和工具的使用。并且提供大量基于实际生产环境的案例，实战性非常强。

全书一共 12 章。第 1 ～ 2 章详细地介绍了 Hadoop 的生态系统、关键技术以及安装和配置；第 3 章是 MapReduce 的使用入门，让读者了解整个开发过程；第 4 ～ 5 章详细讲解了分布式文件系统 HDFS 和 Hadoop 的文件 I/O；第 6 章分析了 MapReduce 的工作原理；第 7 章讲解了如何利用 Eclipse 来编译 Hadoop 的源代码，以及如何对 Hadoop 应用进行测试和调试；第 8 ～ 9 章细致地讲解了 MapReduce 的开发方法和高级应用；第 10 ～ 12 章系统地讲解了 Hive、HBase 和 Mahout。

机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码 100037）
责任编辑：白 宇
三河市杨庄长鸣印刷装订厂印刷
2014 年 1 月第 1 版第 1 次印刷
186mm × 240 mm · 26.25 印张
标准书号：ISBN 978-7-111-45244-7
定 价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991 88361066 投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259 读者信箱：hzjsj@hzbook.com

华章计算机
HZBOOKS | Computer Science and Technology

前　言

为什么要写这本书

三年前接触 Hadoop 的时候，Hadoop 在国内还没有普及，学习的资料也是少之又少，只能看英文文档一步一步操作，出现错误只能自己解决或者查看英文资料，有时候一个问题好几天才解决，学习起来非常的麻烦。

古时候，人们用牛来拉重物，当一头牛拉不动一根圆木的时候，他们不曾想过培育更强壮的牛。同样，我们也不需要尝试更大的计算机，而是应该开发更多的计算系统。简单说，数据量越来越大，我们的大型机负担不了，因此集群就产生了，而 Hadoop 就是为分布式计算和分布式存储而诞生的。最近两年，随着互联网的快速发展，公司的日志量成数量级增长，从而导致使用 Hadoop 的公司也越来越多，Hadoop 人才也就越来越供不应求，尤其是有丰富工作经验的人才。

在开源的 Hadoop 社区里，很多人问我怎么学习 Hadoop，学习 Hadoop 需要什么样的基础，如何才能成为一个 Hadoop 的技术大牛。为了帮助更多的人，我决定写本书，将我对 Hadoop 的理解和实战经验与大家分享。希望帮助读者在学习 Hadoop 的道路上少走弯路，避免一些不必要的时间付出。

读者对象

本书适合以下读者阅读。

- ❑ 大数据爱好者
- ❑ Hadoop 入门者
- ❑ 海量数据处理的从业人员
- ❑ 有 Java 基础，想学习 Hadoop 的开发者
- ❑ Hive、HBase 和 Mahout 的爱好者
- ❑ 具有 1 ～ 3 年使用 Hadoop 经验，想进一步提高能力的使用者
- ❑ 开设相关课程的大专院校

如何阅读本书

本书共 12 章。

第 1 章是 Hadoop 概述，互联网时代什么最重要——数据！大数据时代什么工具最流行——Hadoop！本章主要从大体上介绍 Hadoop，让读者对 Hadoop 有个大体的印象。

第 2 章主要是对 Hadoop 安装进行一步一步地详解，包括 Hadoop 的单机模式、伪分布式模式以及分布式模式，让读者了解 Hadoop 的安装过程。

第 3 章是 MapReduce 快速入门，通过一个实例介绍 MapReduce 开发的整个过程。包括环境的准备、Mapper 和 Reducer 代码的编写、打包、部署和运行的一整套流程。这套流程对以后 MapReduce 的开发奠定了基础。此外，介绍了 MapReduce 开发遇到的常见错误，以及解决方法。

第 4 章本章详细讲解了 Hadoop 分布式文件系统（HDFS），让读者对 HDFS 有一个全面的了解。HDFS 是 Hadoop 里面的一个核心组件，任何基于 Hadoop 应用的工具都要用到 HDFS，学习本章对 Hadoop 系统的维护有很大的帮助。

第 5 章本章主要讲 Hadoop I/O 底层的知识，让读者了解 Hadoop 底层的一些原理，比如数据在 Hadoop 上是怎么传输的。学完这章对 Hadoop 的输入和输出、序列化和反序列化、Writable 类型等有进一步的理解。

第 6 章主要介绍 MapReduce 的工作原理，讲解了 JobTracker、TaskTracker、任务的调度、任务工作的原理以及 MapReduce 的 Shuffle 和 Sort 机制等，此外还介绍第二代 MapReduce（YARN）。通过本章内容的学习，将会对日后 MapReduce 编程开发奠定基础。

第 7 章主要介绍三部分内容，Eclipse 的 MapReduce 插件安装、利用插件调试 MapReduce 代码，以及 MapReduce 单元测试。目的是提高读者编写 MapReduce 的调试、测试的效率。选择和使用一个好的工具是项目成功的一半。

第 8 章详细讲解了 MapReduce 编程开发用到的知识点，如 Combiner、Partitioner、DistributedCache 等。全面介绍了 Mapper 和 Reducer 两个抽象类，最后介绍了如何使用 Hadoop Streaming 接口开发 MapReduce 程序。

第 9 章介绍了 MapReduce 的一些高级应用、算法等知识点，算是对第 8 章学的知识进行应用，本章内容是 MapReduce 开发和 Hadoop 系统调优常用的知识点，希望读者能够掌握，为 MapReduce 的算法开发和调优打下基础。

第 10 章对 Hive 的知识点和使用做了介绍。Hive 是基于 Hadoop 的一个数据仓库，支持类似 SQL 语句来替代 MapReduce，这使不会写代码的数据分析师也能利用 Hadoop 平台对海量数据进行分析，如 DBA。Hive 工具现在已经非常成熟，使用的人也越来越多。

第 11 章介绍了 HBase 的安装和 Shell 操作、MapReduce 操作 HBase 等。HBase 是一个基于列存储的 NoSQL 数据库，不同于一般的 NoSQL 数据库，HBase 是一个适合于非结构化数据存储的数据库。

第 12 章介绍了 Mahout 算法框架的安装和使用。本章介绍的常用知识是入门的最佳实战。Mahout 是非常优秀的基于 MapReduce 的算法框架。Mahout 工具在商品推荐、电影推荐、广告推荐等领域应用非常广泛。

附录 A 为 Hive 内置操作符与函数。附录 B 为 HBase 默认配置解释。附录 C 为 Hadoop

三个配置文件的参数含义说明。

其中第8章和第9章侧重于Java开发MapReduce的用户，如果你是一名有经验的Hadoop用户，并且想学习MapReduce开发，可以直接阅读这两章的内容。但是如果你是一名初学者，请一定从第1章的基础理论知识开始学习。

勘误和支持

除封面署名外，参加本书编写工作的还有：向磊、艾男、曾泽、胡畔、郭洁、童小军、王虎、李晋博、林斌、闫越等。由于作者的水平有限，编写时间仓促，书中难免会出现一些错误或者不准确的地方，恳请读者批评指正。书中的全部源文件可以从华章网站[⊖]下载，我也会定期将相应的功能更新及时更正出来。可以将书中的错误和遇到的任何问题，发邮件到jaylg2010@gmail.com。如果你有更多的宝贵意见，也欢迎发送邮件，期待能够得到你们的真挚反馈。

致谢

首先要感谢Doug Cutting，他开创了一款影响整个互联网领域的开源平台——Hadoop。

感谢机械工业出版社华章公司的杨福川老师，在这一年多的时间始终支持我的写作，你的鼓励和帮助引导我顺利完成全部书稿。感谢编辑白宇对稿件的认真修改，你提出的一些写书注意事项，让我受益匪浅。

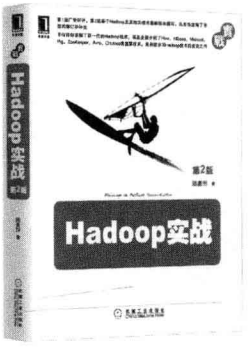
感谢胡畔老师对我的支持和鼓励。

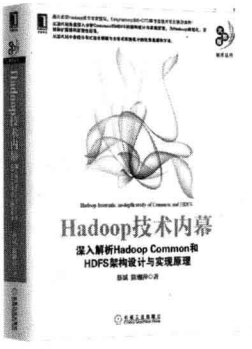
最后感谢我的爸爸、妈妈，感谢你们将我培养成人，并时时刻刻为我灌输着信心和力量！

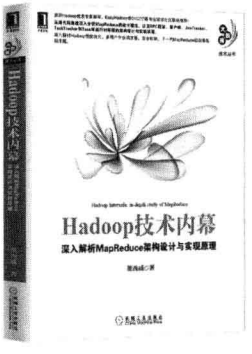
谨以此书献给我最亲爱的家人，以及众多热爱Hadoop的朋友们！

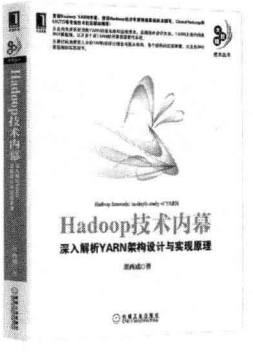
⊖ 参见华章网站www.hzbook.com。——编辑注

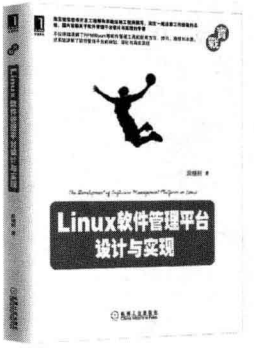
推荐阅读

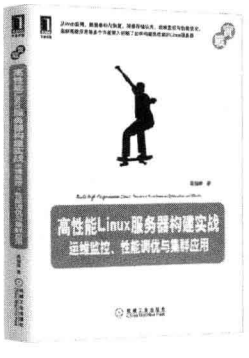

Hadoop实战
第2版
Hadoop实战

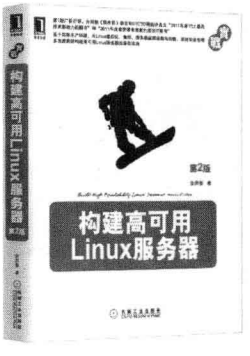

Hadoop技术内幕
深入解析Hadoop Common和HDFS架构设计与实现原理

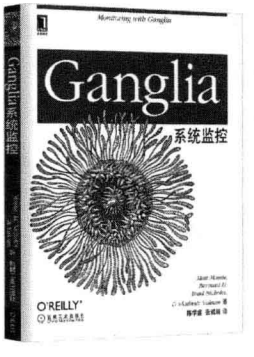

Hadoop技术内幕
深入解析MapReduce架构设计与实现原理


Hadoop技术内幕
深入解析YARN架构设计与实现原理

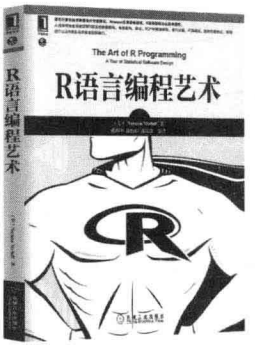

Linux软件管理平台
设计与实现


高性能Linux服务器构建实战
运维监控、性能调优与集群应用

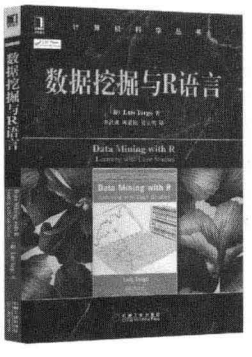

构建高可用Linux服务器
第2版
构建高可用
Linux服务器

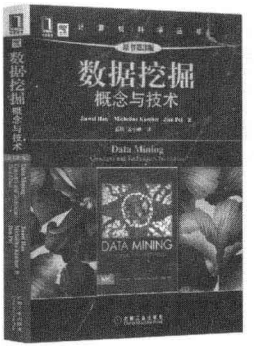

Ganglia
系统监控
O'REILLY

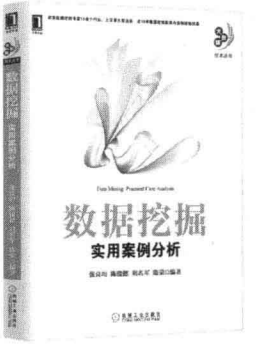
推荐阅读


The Art of R Programming
R语言编程艺术

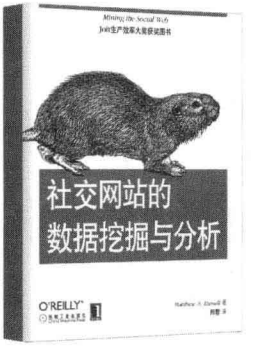

R语言经典实例
O'REILLY

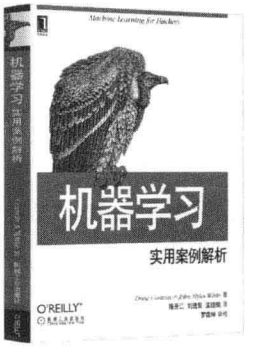

数据挖掘与R语言
Data Mining with R


数据挖掘
概念与技术
Data Mining
DATA MINING


数据挖掘
实用案例分析


数据挖掘与数据化运营实战
思路、方法、技巧与应用


社交网站的
数据挖掘与分析
O'REILLY


机器学习
实用案例解析
O'REILLY

目　录

第1章　Hadoop 概述

Hadoop 是一个开发和运行处理大规模数据的软件平台，是 Apache 的一个用 Java 语言实现开源软件框架，实现在大量计算机组成的集群中对海量数据进行分布式计算。

1.1　Hadoop 起源

Hadoop 框架中最核心设计就是：HDFS 和 MapReduce。HDFS 提供了海量数据的存储，MapReduce 提供了对数据的计算。

1.1.1　Google 与 Hadoop 模块

Google 的数据中心使用廉价的 Linux PC 机组成集群，在上面运行各种应用。即使是分布式开发的新手也可以迅速使用 Google 的基础设施。Hadoop 核心组件与 Google 对应的组件对应关系如表 1-1 所示。

表 1-1　Google 与 Hadoop 对应模块

Google	功能描述	对应 Hadoop 模块
GFS	分布式文件系统（Google File System），隐藏下层负载均衡、冗余复制等细节，对上层程序提供一个统一的文件系统 API 接口。Google 根据自己的需求对它进行了特别优化，包括超大文件的访问、读操作比例远超过写操作、PC 机极易发生故障造成节点失效等。GFS 把文件分成 64MB 的块，分布在集群的机器上，使用 Linux 的文件系统存放，同时每块文件至少有 3 份以上的冗余。中心是一个 Master 节点，根据文件索引找寻文件块	HDFS
MapReduce	Google 发现大多数分布式运算可以抽象为 MapReduce 操作。Map 是把输入 Input 分解成中间的 Key-Value 对，Reduce 把 Key-Value 合成最终输出 Output。这两个函数由程序员提供给系统，下层设施把 Map 和 Reduce 操作分布在集群上运行，并把结果存储在 GFS 上	MapReduce
BigTable	一个大型的分布式数据库，这个数据库不是关系式的数据库。像它的名字一样，就是一个巨大的表格，用来存储结构化的数据	HBase

1.1.2　为什么会有 Hadoop

随着互联网快速的发展，产生的日志数量级的增加，大量的日志给公司带来了很大的挑战。如日志存储问题、海量日志分析的效率问题、成本问题等。

下面我们来分析个问题。

一般网站把用户的访问行为记录以 Apache 日志的形式记录下来，这些日志中包含了用

户访问网站的所有信息，下面列举关键的信息，关键字段如下。

客服端 IP	用户标示	访问时间	访问 URL	关联的 URL	状态	流量	代理
client_ip	user_ip	access_time	url	refrence	status	page_size	agent

因为需要统一对数据进行离线计算，所以常常把它们全部移到同一个地方，如 Oracle 数据库等，每天产生的日志量大概计算一下，如下所示。

- ❑ 网站请求数：1000 万条 / 天
- ❑ 每天日志大小：450 字节 / 行 × 1000 万条 = 4.2 GB
- ❑ 日志存储周期：2 年

一天产生 4.5 GB 日志，2 年需要 4.2 GB × 2 × 365 = 3.0 TB。

怎么来解决 3.0 TB 的数据备份和容错的问题？解决方案如下

1）为了方便系统命令查看日志，不压缩，总共需要 3.0 TB 空间，刚好有一些 2U 的服务器，每台 1 TB 的磁盘空间。

2）为了避免系统盘坏掉影响服务器使用，对系统盘做 Raid1。

3）为了避免其他存放数据的盘坏掉导致数据无法恢复，对剩下的盘做 Raid5。

所有的数据都汇聚到这几台 LogBackUp 服务器上了。有了 LogBackUp 服务器，离线统计就可以全部在这些服务器上进行了。在这套架构上，用 wc、grep、sort、uniq、awk、sed 等系统命令，完成了很多的统计需求，比如统计访问频率较高的 client_ip，某个新上线的页面的 reference 主要是哪些网站。

当公司业务迅猛发展，网站流量爆发增长，日志量也就成指数级增长，日志对于一个互联网公司来说是非常重要的，互联网公司可以通过分析日志来获取用户的行为，如推荐系统、淘宝的用户买卖行为分析等。假如现在的日志量计算如下所示。

- ❑ 日志总行数：10 亿 / 天
- ❑ 每天日志大小：450 字节 / 行 × 10 亿 = 420 GB
- ❑ 日志种类：5 种

1 天产生 420 GB 日志，2 年需要 420 GB × 2 × 365 = 300 TB，这么大的数据量怎么来存储和分析？

1.1.3 Hadoop 版本介绍

最新发布的 Apache Hadoop 版本如图 1-1 所示。一些 Hadoop 入门者会问：这个版本的功能有哪些？基于哪个版本？后续的版本是什么？要解释这一点，我们应该从 Apache 项目发布的一些基本知识开始分析。一般来说，Apache 项目的新功能在主干（trunk）代码上开发。有时候，很大的特性也会有自己的开发分支（branch），它们期望后续会并入 trunk。新功能通常是在 trunk 发布之前就有的，一般质量或稳定性没有太大保证。候选的分支会定期从主干分支上分离出来发布。一旦一个候选分支发布，它通常停止获得新的功能。如果有 bug 修复，经过投票后，会针对这个特定的分支再发布一个新版本。社区的任何成员可以创建一个版本分支，并可随意命名。