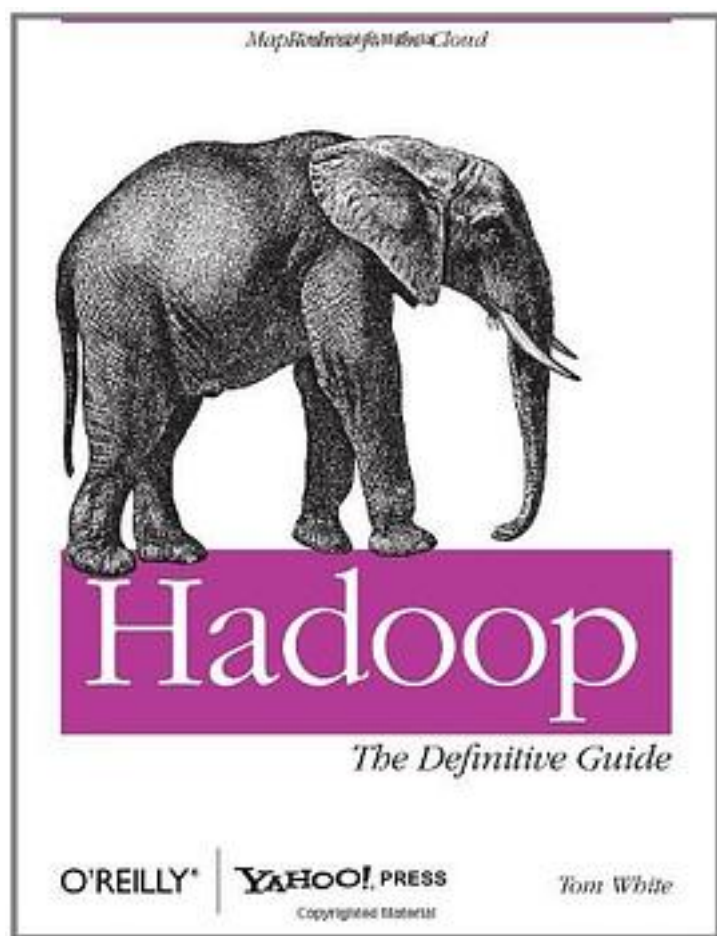


Hadoop



[Hadoop_下载链接1_](#)

著者:Tom White

出版者:O'Reilly Media

出版时间:2012-5-26

装帧:Paperback

isbn:9781449311520

Ready to unleash the power of your massive dataset? With the latest edition of this comprehensive resource, you'll learn how to use Apache Hadoop to build and maintain reliable, scalable, distributed systems. It's ideal for programmers looking to

analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. This third edition covers recent changes to Hadoop, including new material on the new MapReduce API, as well as version 2 of the MapReduce runtime (YARN) and its more flexible execution model. You'll also find illuminating case studies that demonstrate how Hadoop is used to solve specific problems. * Store large datasets with the Hadoop Distributed File System (HDFS), then run distributed computations with MapReduce * Use Hadoop's data and I/O building blocks for compression, data integrity, serialization (including Avro), and persistence * Discover common pitfalls and advanced features for writing real-world MapReduce programs * Design, build, and administer a dedicated Hadoop cluster, or run Hadoop in the cloud * Use Pig, a high-level query language for large-scale data processing * Analyze datasets with Hive, Hadoop's data warehousing system * Load data from relational databases into HDFS, using Sqoop * Take advantage of HBase, the database for structured and semi-structured data * Use ZooKeeper, the toolkit for building distributed systems

作者介绍:

目录:

[Hadoop 下载链接1](#)

标签

Hadoop

分布式

并行计算

数据挖掘

大数据

计算机

O'Reilly

编程

评论

很细致，跟着入门一遍。对这种快速演进的技术，还是老老实实看原版吧..中文版落后太多，等新翻译出来黄花菜都凉了

这书到后面已经神游了，没这环境先不玩

中英结合疗效好，不过还是建议中文为主，英文为辅。

第三版

读的是第4版，和第2版的中文版比起来，加了一些最新的内容。

看完了Hadoop部分，看了部分可选模块章节。真心写的挺仔细的。 17.3月 171216-17
看完了PART III Hadoop Operations

看了前两部分，讲的比较清楚

过了一遍，只知道个大概结构。细节还不是很懂

这段时间学完了大数据，看这本书复习一遍还是很棒的

可以当做概览

终于读完了。。。这本是本系列的第三版了，好评如潮，无需我再费口舌了。

主要看设计思想。设计思想。。设计思想。。。

告别小白

Hadoop权威指南英文版，非常给力

感觉还行，讲的比较细

The system of Big Data, all focus on the Scalability, Fault tolerance, Scheduler, Shuffle.

看完了第一篇关于hadoop,Hdfs,yarn的部分（我看的是第四版），算是对于hadoop有个初步认识吧。这本书写的很有逻辑，值得推荐。但是hadoop暂时还不是我的侧重点，所以先就此别过，以后有机会再回头细读吧。

Hadoop必读

入门最佳，此书最好结合实际操作去理解体会，map/reduce原理及应用部分应该是最权威的，重点看了，后面部分看其他书籍了，平常工作可以当作工具书，有空再多温习下最佳～

简单过了一遍

[Hadoop 下载链接1](#)

书评

买了第一版，时间太紧，没来得及看，后来出了个号称修订升级的第二版，毫不犹豫又买了，后来听说第二版比第一版翻译得好，心中窃喜，再后来看了第二版，我震惊了，我TM就是一傻子，放着好好的英文版不看，赶什么时髦买中文版呢。在这个神奇的国度，牛奶里放的是三聚氰胺，火腿...

其实也不算全部读完了，读它主要是为了技术选型，考虑升级持久层架构、提高系统可扩展性，仔细研读了前几章，对Hadoop、MapReduce、HDFS的模型、机制、使用场景有了一定了解。后面几章及其生态圈内的其他项目抱着了解的心态简单浏览了一下。整体感觉还行，至少从我看过的章节来...

中文版412页：
所以理论上，任何东西都可以表示成二进制形式，然后转化成为长整型的字符串或直接对数据结构进行序列化，来作为键值。原文460页： ..., so theoretically anything can serve as row key, from strings to binary representations of long or even serialized ...

-- china-pub 赠书活动 -- <http://www.douban.com/group/topic/20965935/>
一直比较忙，整本书还没读完，只是粗略翻了个大概，其中有两三章细读了一遍。先做个大体评价吧，有时间全部细读后再评论。
从书的内容上来讲，大致上与网上该书的内容介绍一致。简单点概括：这本书对...

参加豆瓣China-pub抽奖，比较幸运的得到这本Hadoop权威指南中文第二版，拿来与第一版相比，发现新加入了Hive和Sqoop章节，译文质量也提高了不少，并且保留了英文索引。
这本书对Hadoop的介绍还算全面，有实践冲动的朋友基本可以拿着书、配合Google百

度马上实现梦想。个人感觉 “...

看了几章中文版的，各种错误，太低级，实在是看不下去了。建议还是看原版吧。译者们的脸皮可真厚，英文译不明白也就罢了，中文都组织的不通顺，好意思吗！！什么叫“但是，.....，但是”啊，“但是体”啊。

很好的Hadoop教程，比Apache和Yahoo!网页版guide详细很多，很多想不明白的Hadoop实现细节都可以在这本书里找到。

详见：<http://www.cnblogs.com/aprilrain/archive/2013/03/07/2947664.html>

很多地方翻译的不行，需要对照英文看才能明白。。。不过对于快速学习，仍然是不错的选择。建议译者看看每部分内容的重要性，不重要的瞎翻翻就算了，重要的部分还是好好花点功夫，不要本末倒置了。比如第三章的数据流部分，这么经典的地方居然被翻译烂的一塌糊涂。不知道译者会...

书中没有透露太多实现架构方面的细节，更多的是从使用者的角度上介绍了Hadoop的各种知识，包括MapReduce, HDFS, Hive, Pig, HBase, ZooKeeper。几乎涉及了Hadoop的所有关于使用方面的知识，包括安装和使用。你甚至可以直接在自己的电脑上装上一个Hadoop，对着书中的例子实际演...

是我遇到过的翻译最烂的一本书，在译者的“妙语连珠”里折腾了半个钟头就再也没兴趣了。略举几例如下： P.6 任然 -> 仍然 P.21 输入键（为什么不像后面那样有个“的”？），输入的值，输出的键…… P. 27 “计数器” (Counter)，译文附原文；"Context Object"(上下文对象)，原...

专门登录来评论的，翻译也太烂了吧，真的真的建议强烈英语阅读能力好的人去读原版书，不要花冤枉钱在这上面，除了文字错误外，里边的图居然也有错，就比如260页的图最后两个年份应该是1901结果这里竟然是1900，我是真滴服了，一本神书被翻译成这样，作者得气死。zsbd zsbd zsbd...

你的履历添了一笔<hadoop权威指南>译者,但是你不配 这是我见过的最不用心的翻译,字里行间行文不通顺,请别勉强自己,map reduce shuffle机制都没翻译的好 虽然原作者写作功底也实在是一般 第 1 2 5 6 7 这几章 翻译的实在是太烂了 请不要呐Google翻译糊弄人阿 误人子弟 ...

首先,翻译太差,很多句子就是瞎翻,根本不通顺,很多时候你要停下来断句,慢慢去理解。然后,这本书是很多人去翻译的,很多人连代码都不懂,曾经一段代码看到我蒙圈,去看了一下源代码,好家伙,四行有五个错误。另外,从代码瞎缩进也可以看出这是群没写过代码的人翻的,而且...

Cobub Razor APP数据统计分析工具官网上有篇文章是讲Hadoop Yarn调度器的选择和使用的,我觉得写的挺好的,推荐<http://www.cobub.com/the-selection-and-use-of-hadoop-yarn-scheduler/>

[Hadoop_下载链接1](#)