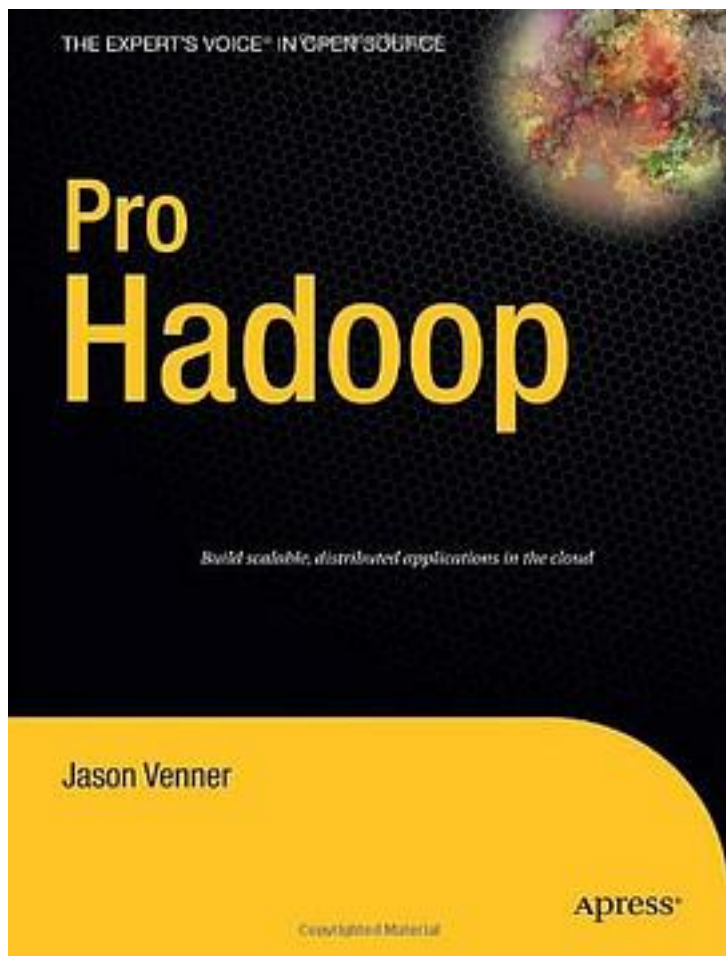# Pro Hadoop

[Pro Hadoop_下载链接1_](#)

著者:Jason Venner

出版者:Apress

出版时间:2009-6-21

装帧:Paperback

isbn:9781430219422

You've heard the hype about Hadoop: it runs petabyte--scale data mining tasks insanely fast, it runs gigantic tasks on clouds for absurdly cheap, it's been heavily committed to by tech giants like IBM, Yahoo!, and the Apache Project, and it's

completely open source (thus free). But what exactly is it, and more importantly, how do you even get a Hadoop cluster up and running? From Apress, the name you've come to trust for hands--on technical knowledge, Pro Hadoop brings you up to speed on Hadoop. You learn the ins and outs of MapReduce; how to structure a cluster, design, and implement the Hadoop file system; and how to build your first cloud--computing tasks using Hadoop. Learn how to let Hadoop take care of distributing and parallelizing your software--you just focus on the code, Hadoop takes care of the rest. Best of all, you'll learn from a tech professional who's been in the Hadoop scene since day one. Written from the perspective of a principal engineer with down--in--the--trenches knowledge of what to do wrong with Hadoop, you learn how to avoid the common, expensive first errors that everyone makes with creating their own Hadoop system or inheriting someone else's. Skip the novice stage and the expensive, hard--to--fix mistakes...go straight to seasoned pro on the hottest cloud--computing framework with Pro Hadoop. Your productivity will blow your managers away. What you'll learn * Set up a stand--alone Hadoop cluster the smart way, laid out simply and step by step so you can get up and running quickly to build your next data center, collaborative, data--intensive Internet services application, Software as a Service (SaaS), and more. * Optimize your Hadoop production tasks like an experienced pro. * Work with time--proven, bulletproof standard patterns that have been tested and debugged in high--volume production. * Understand just enough theoretical knowledge to know why something works in Hadoop, without getting bogged down in abstruse walls of theory. * Get detailed explanations of not only how to do something with Hadoop, but also why, from a front--line coder with years in the Hadoop game. * Turn someone else's expensive cluster--wide "wrong" into an orderly, productive "right" with professional--level debugging and testing. Who this book is for IT professionals interested in investigating Hadoop and implementing it in their organizations, and existing Hadoop users who want to deepen their professional toolkits. Table of Contents * Getting Started with Hadoop Core * The Basics of a MapReduce Job * The Basics of Multimachine Clusters * HDFS Details for Multimachine Clusters * MapReduce Details for Multimachine Clusters * Tuning Your MapReduce Jobs * Unit Testing and Debugging * Advanced and Alternate MapReduce Techniques * Solving Problems with Hadoop * Projects Based On Hadoop and Future Directions

作者介绍:

目录:

[Pro Hadoop_下载链接1_](#)

标签

hadoop

MapReduce

分布式

编程

云计算

Hadoop

架构

programming

评论

mapreduce only

不如另外一本好

很注重细节，哈哈。

也还不错吧

与另外一本"Hadoop: The Definite Guide"各有侧重，很不错的补充。

说实话，一般……APress的不要报太大希望。

Pro不是professional或者是promote，而是program。从开发者角度，详细说明hadoop的各种参数，竟然还贴出代码！感觉收获不如definitive hadoop多

Pro Hadoop_下载链接1_

书评

Pro Hadoop_下载链接1_

Pro不是professional或者是promote，而是program。从开发者角度，详细说明hadoop的各种参数，竟然还贴出代码！感觉收获不如definitive hadoop多

Pro Hadoop_下载链接1_