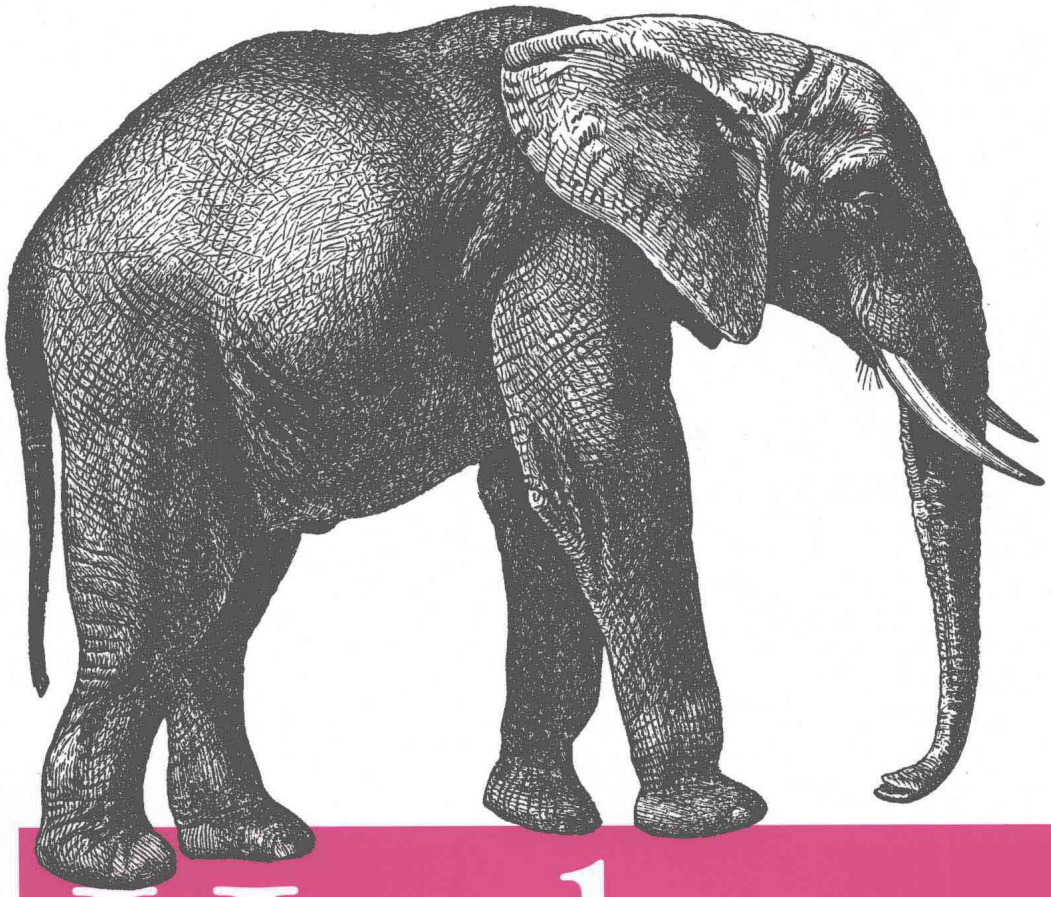
Hadoop权威指南（影印版）

第二版
修订版



# Hadoop

*The Definitive Guide*

O'REILLY® | YAHOO! PRESS
東南大學出版社

*Tom White* 著

（第2版）

# Hadoop权威指南（影印版）

# Hadoop: The Definitive Guide

*Tom White*

*foreword by Doug Cutting*

O'REILLY®

Beijing · Cambridge · Farnham · Köln · Sebastopol · Tokyo

O'Reilly Media, Inc.授权东南大学出版社出版

东南大学出版社

**图书在版编目（CIP）数据**

Hadoop权威指南：英文/（美）怀特（White, T.）著.—影印本.—南京：东南大学出版社，2011.5
书名原文：Hadoop: The Definitive Guide, 2E
ISBN 978-7-5641-2676-6

I. ①H… II. ①怀… III. ①数据处理－应用软件－指南－英文 IV. ①TP274-62

中国版本图书馆CIP数据核字（2011）第042756号

江苏省版权局著作权合同登记
图字：10-2010-438号

©2010 by O'Reilly Media, Inc.

Reprint of the English Edition, jointly published by O'Reilly Media, Inc. and Southeast University Press, 2011. Authorized reprint of the original English edition, 2010 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由O'Reilly Media, Inc.出版2010。

英文影印版由东南大学出版社出版2011。此影印版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc.的许可。

版权所有，未得书面许可，本书的任何部分和全部不得以任何形式重制。

Hadoop权威指南（影印版）

出版发行：东南大学出版社
地　　址：南京四牌楼2号　　邮编：210096
出 版 人：江建中
网　　址：http://www.seupress.com
电子邮件：press@seupress.com
印　　刷：扬中市印刷有限公司
开　　本：787毫米×980毫米　16开本
印　　张：39.25印张
字　　数：972千字
版　　次：2011年5月第1版
印　　次：2011年5月第1次印刷
书　　号：ISBN 978-7-5641-2676-6
定　　价：98.00元（册）

本社图书若有印装质量问题，请直接与读者服务部联系。电话（传真）：025-83792328

# Foreword

Hadoop got its start in Nutch. A few of us were attempting to build an open source web search engine and having trouble managing computations running on even a handful of computers. Once Google published its GFS and MapReduce papers, the route became clear. They'd devised systems to solve precisely the problems we were having with Nutch. So we started, two of us, half-time, to try to re-create these systems as a part of Nutch.

We managed to get Nutch limping along on 20 machines, but it soon became clear that to handle the Web's massive scale, we'd need to run it on thousands of machines and, moreover, that the job was bigger than two half-time developers could handle.

Around that time, Yahoo! got interested, and quickly put together a team that I joined. We split off the distributed computing part of Nutch, naming it Hadoop. With the help of Yahoo!, Hadoop soon grew into a technology that could truly scale to the Web.

In 2006, Tom White started contributing to Hadoop. I already knew Tom through an excellent article he'd written about Nutch, so I knew he could present complex ideas in clear prose. I soon learned that he could also develop software that was as pleasant to read as his prose.

From the beginning, Tom's contributions to Hadoop showed his concern for users and for the project. Unlike most open source contributors, Tom is not primarily interested in tweaking the system to better meet his own needs, but rather in making it easier for anyone to use.

Initially, Tom specialized in making Hadoop run well on Amazon's EC2 and S3 services. Then he moved on to tackle a wide variety of problems, including improving the MapReduce APIs, enhancing the website, and devising an object serialization framework. In all cases, Tom presented his ideas precisely. In short order, Tom earned the role of Hadoop committer and soon thereafter became a member of the Hadoop Project Management Committee.

Tom is now a respected senior member of the Hadoop developer community. Though he's an expert in many technical corners of the project, his specialty is making Hadoop easier to use and understand.

Given this, I was very pleased when I learned that Tom intended to write a book about Hadoop. Who could be better qualified? Now you have the opportunity to learn about Hadoop from a master—not only of the technology, but also of common sense and plain talk.

—Doug Cutting
Shed in the Yard, California



试读结束，需要全本PDF请购买 www.ertongbook.com

# Preface

Martin Gardner, the mathematics and science writer, once said in an interview:

> Beyond calculus, I am lost. That was the secret of my column's success. It took me so long to understand what I was writing about that I knew how to write in a way most readers would understand.*

In many ways, this is how I feel about Hadoop. Its inner workings are complex, resting as they do on a mixture of distributed systems theory, practical engineering, and common sense. And to the uninitiated, Hadoop can appear alien.

But it doesn't need to be like this. Stripped to its core, the tools that Hadoop provides for building distributed systems—for data storage, data analysis, and coordination—are simple. If there's a common theme, it is about raising the level of abstraction—to create building blocks for programmers who just happen to have lots of data to store, or lots of data to analyze, or lots of machines to coordinate, and who don't have the time, the skill, or the inclination to become distributed systems experts to build the infrastructure to handle it.

With such a simple and generally applicable feature set, it seemed obvious to me when I started using it that Hadoop deserved to be widely used. However, at the time (in early 2006), setting up, configuring, and writing programs to use Hadoop was an art. Things have certainly improved since then: there is more documentation, there are more examples, and there are thriving mailing lists to go to when you have questions. And yet the biggest hurdle for newcomers is understanding what this technology is capable of, where it excels, and how to use it. That is why I wrote this book.

The Apache Hadoop community has come a long way. Over the course of three years, the Hadoop project has blossomed and spun off half a dozen subprojects. In this time, the software has made great leaps in performance, reliability, scalability, and manageability. To gain even wider adoption, however, I believe we need to make Hadoop even easier to use. This will involve writing more tools; integrating with more systems; and

* "The science of fun," Alex Bellos, *The Guardian*, May 31, 2008, *http://www.guardian.co.uk/science/2008/may/31/maths.science*.

writing new, improved APIs. I'm looking forward to being a part of this, and I hope this book will encourage and enable others to do so, too.

## Administrative Notes

During discussion of a particular Java class in the text, I often omit its package name, to reduce clutter. If you need to know which package a class is in, you can easily look it up in Hadoop's Java API documentation for the relevant subproject, linked to from the Apache Hadoop home page at *http://hadoop.apache.org/*. Or if you're using an IDE, it can help using its auto-complete mechanism.

Similarly, although it deviates from usual style guidelines, program listings that import multiple classes from the same package may use the asterisk wildcard character to save space (for example: `import org.apache.hadoop.io.*`).

The sample programs in this book are available for download from the website that accompanies this book: *http://www.hadoopbook.com/*. You will also find instructions there for obtaining the datasets that are used in examples throughout the book, as well as further notes for running the programs in the book, and links to updates, additional resources, and my blog.

## What's in This Book?

The rest of this book is organized as follows. Chapter 1 emphasizes the need for Hadoop and sketches the history of the project. Chapter 2 provides an introduction to MapReduce. Chapter 3 looks at Hadoop filesystems, and in particular HDFS, in depth. Chapter 4 covers the fundamentals of I/O in Hadoop: data integrity, compression, serialization, and file-based data structures.

The next four chapters cover MapReduce in depth. Chapter 5 goes through the practical steps needed to develop a MapReduce application. Chapter 6 looks at how MapReduce is implemented in Hadoop, from the point of view of a user. Chapter 7 is about the MapReduce programming model, and the various data formats that MapReduce can work with. Chapter 8 is on advanced MapReduce topics, including sorting and joining data.

Chapters 9 and 10 are for Hadoop administrators, and describe how to set up and maintain a Hadoop cluster running HDFS and MapReduce.

Later chapters are dedicated to projects that build on Hadoop or are related to it. Chapters 11 and 12 present Pig and Hive, which are analytics platforms built on HDFS and MapReduce, whereas Chapters 13, 14, and 15 cover HBase, ZooKeeper, and Sqoop, respectively.

Finally, Chapter 16 is a collection of case studies contributed by members of the Apache Hadoop community.

## What's New in the Second Edition?

The second edition has two new chapters on Hive and Sqoop (Chapters 12 and 15), a new section covering Avro (in Chapter 4), an introduction to the new security features in Hadoop (in Chapter 9), and a new case study on analyzing massive network graphs using Hadoop (in Chapter 16).

This edition continues to describe the 0.20 release series of Apache Hadoop, since this was the latest stable release at the time of writing. New features from later releases are occasionally mentioned in the text, however, with reference to the version that they were introduced in.

## Conventions Used in This Book

The following typographical conventions are used in this book:

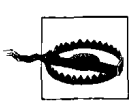
*Italic*
: Indicates new terms, URLs, email addresses, filenames, and file extensions.

`Constant width`
: Used for program listings, as well as within paragraphs to refer to program elements such as variable or function names, databases, data types, environment variables, statements, and keywords.

**`Constant width bold`**
: Shows commands or other text that should be typed literally by the user.

*`Constant width italic`*
: Shows text that should be replaced with user-supplied values or by values determined by context.

This icon signifies a tip, suggestion, or general note.



This icon indicates a warning or caution.

## Using Code Examples

This book is here to help you get your job done. In general, you may use the code in this book in your programs and documentation. You do not need to contact us for permission unless you're reproducing a significant portion of the code. For example, writing a program that uses several chunks of code from this book does not require permission. Selling or distributing a CD-ROM of examples from O'Reilly books does

require permission. Answering a question by citing this book and quoting example code does not require permission. Incorporating a significant amount of example code from this book into your product's documentation does require permission.

We appreciate, but do not require, attribution. An attribution usually includes the title, author, publisher, and ISBN. For example: "*Hadoop: The Definitive Guide*, Second Edition, by Tom White. Copyright 2011 Tom White, 978-1-449-38973-4."

If you feel your use of code examples falls outside fair use or the permission given above, feel free to contact us at *permissions@oreilly.com*.

## Safari® Books Online

Safari Books Online is an on-demand digital library that lets you easily search over 7,500 technology and creative reference books and videos to find the answers you need quickly.

With a subscription, you can read any page and watch any video from our library online. Read books on your cell phone and mobile devices. Access new titles before they are available for print, and get exclusive access to manuscripts in development and post feedback for the authors. Copy and paste code samples, organize your favorites, download chapters, bookmark key sections, create notes, print out pages, and benefit from tons of other time-saving features.

O'Reilly Media has uploaded this book to the Safari Books Online service. To have full digital access to this book and others on similar topics from O'Reilly and other publishers, sign up for free at *http://my.safaribooksonline.com*.

## How to Contact Us

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
800-998-9938 (in the United States or Canada)
707-829-0515 (international or local)
707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at:

*http://oreilly.com/catalog/0636920010388/*

The author also has a site for this book at:

*http://www.hadoopbook.com/*

To comment or ask technical questions about this book, send email to:

*bookquestions@oreilly.com*

For more information about our books, conferences, Resource Centers, and the O'Reilly Network, see our website at:

*http://www.oreilly.com*

## Acknowledgments

I have relied on many people, both directly and indirectly, in writing this book. I would like to thank the Hadoop community, from whom I have learned, and continue to learn, a great deal.

In particular, I would like to thank Michael Stack and Jonathan Gray for writing the chapter on HBase. Also thanks go to Adrian Woodhead, Marc de Palol, Joydeep Sen Sarma, Ashish Thusoo, Andrzej Białecki, Stu Hood, Chris K. Wensel, and Owen O'Malley for contributing case studies for Chapter 16.

I would like to thank the following reviewers who contributed many helpful suggestions and improvements to my drafts: Raghu Angadi, Matt Biddulph, Christophe Bisciglia, Ryan Cox, Devaraj Das, Alex Dorman, Chris Douglas, Alan Gates, Lars George, Patrick Hunt, Aaron Kimball, Peter Krey, Hairong Kuang, Simon Maxen, Olga Natkovich, Benjamin Reed, Konstantin Shvachko, Allen Wittenauer, Matei Zaharia, and Philip Zeyliger. Ajay Anand kept the review process flowing smoothly. Philip ("flip") Kromer kindly helped me with the NCDC weather dataset featured in the examples in this book. Special thanks to Owen O'Malley and Arun C. Murthy for explaining the intricacies of the MapReduce shuffle to me. Any errors that remain are, of course, to be laid at my door.

For the second edition, I owe a debt of gratitude for the detailed review and feedback from Jeff Bean, Doug Cutting, Glynn Durham, Alan Gates, Jeff Hammerbacher, Alex Kozlov, Ken Krugler, Jimmy Lin, Todd Lipcon, Sarah Sproehnle, Vinithra Varadharajan, and Ian Wrigley, as well as all the readers who submitted errata for the first edition. I would also like to thank Aaron Kimball for contributing the chapter on Sqoop, and Philip ("flip") Kromer for the case study on graph processing.

I am particularly grateful to Doug Cutting for his encouragement, support, and friendship, and for contributing the foreword.

Thanks also go to the many others with whom I have had conversations or email discussions over the course of writing the book.

Halfway through writing this book, I joined Cloudera, and I want to thank my colleagues for being incredibly supportive in allowing me the time to write, and to get it finished promptly.

I am grateful to my editor, Mike Loukides, and his colleagues at O'Reilly for their help in the preparation of this book. Mike has been there throughout to answer my questions, to read my first drafts, and to keep me on schedule.

Finally, the writing of this book has been a great deal of work, and I couldn't have done it without the constant support of my family. My wife, Eliane, not only kept the home going, but also stepped in to help review, edit, and chase case studies. My daughters, Emilia and Lottie, have been very understanding, and I'm looking forward to spending lots more time with all of them.



试读结束，需要全本PDF请购买 www.ertongbook.com

# Table of Contents

**Foreword** ..... xv

**Preface** ..... xvii

**1. Meet Hadoop** ..... 1
Data! 1
Data Storage and Analysis 3
Comparison with Other Systems 4
RDBMS 4
Grid Computing 6
Volunteer Computing 8
A Brief History of Hadoop 9
Apache Hadoop and the Hadoop Ecosystem 12

**2. MapReduce** ..... 15
A Weather Dataset 15
Data Format 15
Analyzing the Data with Unix Tools 17
Analyzing the Data with Hadoop 18
Map and Reduce 18
Java MapReduce 20
Scaling Out 27
Data Flow 28
Combiner Functions 30
Running a Distributed MapReduce Job 33
Hadoop Streaming 33
Ruby 33
Python 36
Hadoop Pipes 37
Compiling and Running 38

**3. The Hadoop Distributed Filesystem ..... 41**
The Design of HDFS 41
HDFS Concepts 43
Blocks 43
Namenodes and Datanodes 44
The Command-Line Interface 45
Basic Filesystem Operations 46
Hadoop Filesystems 47
Interfaces 49
The Java Interface 51
Reading Data from a Hadoop URL 51
Reading Data Using the FileSystem API 52
Writing Data 55
Directories 57
Querying the Filesystem 57
Deleting Data 62
Data Flow 62
Anatomy of a File Read 62
Anatomy of a File Write 65
Coherency Model 68
Parallel Copying with distcp 70
Keeping an HDFS Cluster Balanced 71
Hadoop Archives 71
Using Hadoop Archives 72
Limitations 73

**4. Hadoop I/O ..... 75**
Data Integrity 75
Data Integrity in HDFS 75
LocalFileSystem 76
ChecksumFileSystem 77
Compression 77
Codecs 78
Compression and Input Splits 83
Using Compression in MapReduce 84
Serialization 86
The Writable Interface 87
Writable Classes 89
Implementing a Custom Writable 96
Serialization Frameworks 101
Avro 103
File-Based Data Structures 116
SequenceFile 116

MapFile 123

**5. Developing a MapReduce Application ..... 129**
The Configuration API 130
Combining Resources 131
Variable Expansion 132
Configuring the Development Environment 132
Managing Configuration 132
GenericOptionsParser, Tool, and ToolRunner 135
Writing a Unit Test 138
Mapper 138
Reducer 140
Running Locally on Test Data 141
Running a Job in a Local Job Runner 141
Testing the Driver 145
Running on a Cluster 146
Packaging 146
Launching a Job 146
The MapReduce Web UI 148
Retrieving the Results 151
Debugging a Job 153
Using a Remote Debugger 158
Tuning a Job 160
Profiling Tasks 160
MapReduce Workflows 163
Decomposing a Problem into MapReduce Jobs 163
Running Dependent Jobs 165

**6. How MapReduce Works ..... 167**
Anatomy of a MapReduce Job Run 167
Job Submission 167
Job Initialization 169
Task Assignment 169
Task Execution 170
Progress and Status Updates 170
Job Completion 172
Failures 173
Task Failure 173
Tasktracker Failure 175
Jobtracker Failure 175
Job Scheduling 175
The Fair Scheduler 176
The Capacity Scheduler 177

Shuffle and Sort 177
The Map Side 177
The Reduce Side 179
Configuration Tuning 180
Task Execution 183
Speculative Execution 183
Task JVM Reuse 184
Skipping Bad Records 185
The Task Execution Environment 186

**7. MapReduce Types and Formats ..... 189**
MapReduce Types 189
The Default MapReduce Job 191
Input Formats 198
Input Splits and Records 198
Text Input 209
Binary Input 213
Multiple Inputs 214
Database Input (and Output) 215
Output Formats 215
Text Output 216
Binary Output 216
Multiple Outputs 217
Lazy Output 224
Database Output 224

**8. MapReduce Features ..... 225**
Counters 225
Built-in Counters 225
User-Defined Java Counters 227
User-Defined Streaming Counters 232
Sorting 232
Preparation 232
Partial Sort 233
Total Sort 237
Secondary Sort 241
Joins 247
Map-Side Joins 247
Reduce-Side Joins 249
Side Data Distribution 252
Using the Job Configuration 252
Distributed Cache 253
MapReduce Library Classes 257

**9. Setting Up a Hadoop Cluster ..... 259**
Cluster Specification 259
Network Topology 261
Cluster Setup and Installation 263
Installing Java 264
Creating a Hadoop User 264
Installing Hadoop 264
Testing the Installation 265
SSH Configuration 265
Hadoop Configuration 266
Configuration Management 267
Environment Settings 269
Important Hadoop Daemon Properties 273
Hadoop Daemon Addresses and Ports 278
Other Hadoop Properties 279
User Account Creation 280
Security 281
Kerberos and Hadoop 282
Delegation Tokens 284
Other Security Enhancements 285
Benchmarking a Hadoop Cluster 286
Hadoop Benchmarks 287
User Jobs 289
Hadoop in the Cloud 289
Hadoop on Amazon EC2 290

**10. Administering Hadoop ..... 293**
HDFS 293
Persistent Data Structures 293
Safe Mode 298
Audit Logging 300
Tools 300
Monitoring 305
Logging 305
Metrics 306
Java Management Extensions 309
Maintenance 312
Routine Administration Procedures 312
Commissioning and Decommissioning Nodes 313
Upgrades 316

**11. Pig ..... 321**
Installing and Running Pig 322

Execution Types 322
Running Pig Programs 324
Grunt 324
Pig Latin Editors 325
An Example 325
Generating Examples 327
Comparison with Databases 328
Pig Latin 330
Structure 330
Statements 331
Expressions 335
Types 336
Schemas 338
Functions 342
User-Defined Functions 343
A Filter UDF 343
An Eval UDF 347
A Load UDF 348
Data Processing Operators 351
Loading and Storing Data 351
Filtering Data 352
Grouping and Joining Data 354
Sorting Data 359
Combining and Splitting Data 360
Pig in Practice 361
Parallelism 361
Parameter Substitution 362

**12. Hive ..... 365**
Installing Hive 366
The Hive Shell 367
An Example 368
Running Hive 369
Configuring Hive 369
Hive Services 371
The Metastore 373
Comparison with Traditional Databases 375
Schema on Read Versus Schema on Write 376
Updates, Transactions, and Indexes 376
HiveQL 377
Data Types 378
Operators and Functions 380
Tables 381



试读结束，需要全本PDF请购买 www.ertongbook.com