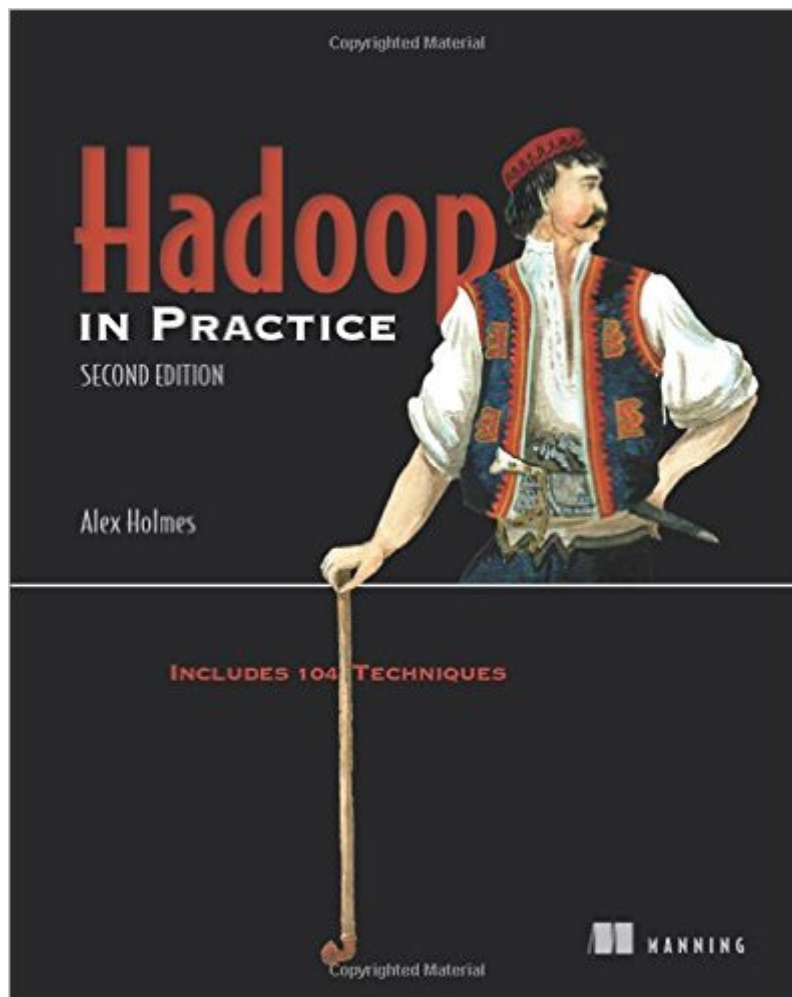


The book was found

# Hadoop In Practice



## Synopsis

SummaryHadoop in Practice, Second Edition provides over 100 tested, instantly useful techniques that will help you conquer big data, using Hadoop. This revised new edition covers changes and new features in the Hadoop core architecture, including MapReduce 2. Brand new chapters cover YARN and integrating Kafka, Impala, and Spark SQL with Hadoop. You'll also get new and updated techniques for Flume, Sqoop, and Mahout, all of which have seen major new versions recently. In short, this is the most practical, up-to-date coverage of Hadoop available anywhere. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Book It's always a good time to upgrade your Hadoop skills! Hadoop in Practice, Second Edition provides a collection of 104 tested, instantly useful techniques for analyzing real-time streams, moving data securely, machine learning, managing large-scale clusters, and taming big data using Hadoop. This completely revised edition covers changes and new features in Hadoop core, including MapReduce 2 and YARN. You'll pick up hands-on best practices for integrating Spark, Kafka, and Impala with Hadoop, and get new and updated techniques for the latest versions of Flume, Sqoop, and Mahout. In short, this is the most practical, up-to-date coverage of Hadoop available. Readers need to know a programming language like Java and have basic familiarity with Hadoop. What's Inside Thoroughly updated for Hadoop 2 How to write YARN applications Integrate real-time technologies like Storm, Impala, and Spark Predictive analytics using Mahout and RR Readers need to know a programming language like Java and have basic familiarity with Hadoop. About the Author Alex Holmes works on tough big-data problems. He is a software engineer, author, speaker, and blogger specializing in large-scale Hadoop projects. Table of Contents PART 1 BACKGROUND AND FUNDAMENTALS Hadoop in a heartbeat Introduction to YARN PART 2 DATA LOGISTICS Data serialization "working with text and beyond Organizing and optimizing data in HDFS Moving data into and out of Hadoop PART 3 BIG DATA PATTERNS Applying MapReduce patterns to big data Utilizing data structures and algorithms at scale Tuning, debugging, and testing PART 4 BEYOND MAPREDUCE SQL on Hadoop Writing a YARN application

## Book Information

Paperback: 512 pages

Publisher: Manning Publications; 2 edition (October 12, 2014)

Language: English

ISBN-10: 1617292222

ISBN-13: 978-1617292224

Product Dimensions: 1.2 x 7.5 x 9.2 inches

Shipping Weight: 1.9 pounds (View shipping rates and policies)

Average Customer Review: 4.6 out of 5 stars Â Â See all reviews Â (7 customer reviews)

Best Sellers Rank: #629,846 in Books (See Top 100 in Books) #97 in Â Books > Computers & Technology > Programming > Languages & Tools > XML #132 in Â Books > Computers & Technology > Programming > Software Design, Testing & Engineering > Tools #227 in Â Books > Textbooks > Computer Science > Object-Oriented Software Design

## Customer Reviews

Let me start by saying, if you want to maximize the benefit from reading "Hadoop in Practice," you better have a solid conceptual understanding of HDFS and MapReduce. The book does start with a primer of Hadoop, but it quickly delves into a deep detailed explanation of YARN and how YARN has helped evolve the Hadoop ecosystem. Chapter 2 in particular explains how the Hadoop architecture, including MapReduce, has been overhauled to use YARN applications. "Hadoop in Practice" is filled with excellent real-world examples that show Hadoop in motion. Those real-world examples are broken down in a consumable format that includes 85 different use cases with accompanying solutions. Having an intermediate-level knowledge of Java will go a long way as you traverse the examples and source code included with the book. My personal favorite chapter, Chapter 7, shows users how to leverage data structures and algorithms at scale. Whether it is describing and detailing the shortest-path algorithm or friends-of-friends algorithm, "Hadoop in Practice" simply does not skimp on providing legitimate use cases along with pseudo code before breaking down each implementation in MapReduce complete with Java code. Overall, "Hadoop in Practice" is structured well and every chapter begins with a high-level overview and ends with a thorough summary. Most chapters include detailed diagrams that I found useful when trying to understand some of the more complex issues in the book - and trust me, there are many.

The Hadoop world has undergone some big changes lately, and this hefty, updated edition offers some excellent coverage of a lot of what's new. If you currently work with Hadoop and MapReduce or are planning to take them up soon, give serious consideration to adding this well-written book to your technical library. A key feature of the book is its 104 techniques. These show how to accomplish practical and important tasks when working with Hadoop, MapReduce and their growing arrays of software friends. Of course, one book cannot possibly

cover everything you need to know about Hadoop, MapReduce, Parquet, Kafka, Camus, YARN and other technologies. And the book's software examples assume that you have had some experience with Java, XML and JSON. However, Hadoop in Practice, Second Edition gives a very good and reasonably deep overview of the Hadoop world, spanning such major topic categories as background and fundamentals, data logistics, Big Data patterns, and moving beyond MapReduce. I don't "love" Hadoop and MapReduce; my experiences with them have been a struggle. So I can't give this book five stars on the review scale. But I am happy to give it four solid stars, and I definitely recommend it to others.

this new second edition is what I was looking for - approximately 100 use case scenarios, including good diagrams and commented code to completely show the problem and solution. This book assumes the reader has some familiarity with Java, but is approachable for most professional coders. There's also 'enough computer science' in this book to arm you with the theory of what's happening during execution. if the table of contents lists stuff you want to learn about, this book is a very good choice.

The first book about Hadoop to make any sense ! I am new to Hadoop and went through a bunch of resources before finding this. The author just intuitively knows how to make an abstract concept easy to understand

[Download to continue reading...](#)

Hadoop in Practice Big Data Analytics with R and Hadoop Data Analytics with Hadoop: An Introduction for Data Scientists Agile Data Science: Building Data Analytics Applications with Hadoop Hadoop Application Architectures Pro Apache Hadoop MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems Practical Hive: A Guide to Hadoop's Data Warehouse System Starting a Medical Practice: The Physician's Handbook for Successful Practice Start-Up (Practice Success! Series) Ready, Set, Practice: Elements of Landscape Architecture Professional Practice Practice Makes Perfect Algebra (Practice Makes Perfect (McGraw-Hill)) MCAT Chemistry and Physics: Strategy and Practice: Timed Practice for the Revised MCAT MCAT Psychology and Sociology: Strategy and Practice (MCAT Strategy and Practice) CPC Practice Exam 2016: Includes 150 practice questions, answers with full rationale, exam study guide and the official proctor-to-examinee instructions McGraw-Hill Education 5 TEAS Practice Tests, 2nd Edition (Mcgraw Hill's 5 Teas Practice Tests) Adult CCRN Strategies, Practice, and Review with 2 Practice Tests (Kaplan Test Prep) Common Core Practice - 8th Grade English

Language Arts: Workbooks to Prepare for the PARCC or Smarter Balanced Test: CCSS Aligned (CCSS Standards Practice) (Volume 11) SSAT & ISEE 2017 Strategies, Practice & Review with 6 Practice Tests: For Private and Independent School Admissions (Kaplan Test Prep) NCLEX-RN 2016 Strategies, Practice and Review with Practice Test (Kaplan Test Prep) ACT 2016-2017 Strategies, Practice, and Review with 6 Practice Tests: Online + Book (Kaplan Test Prep)

[Dmca](#)