The book was found
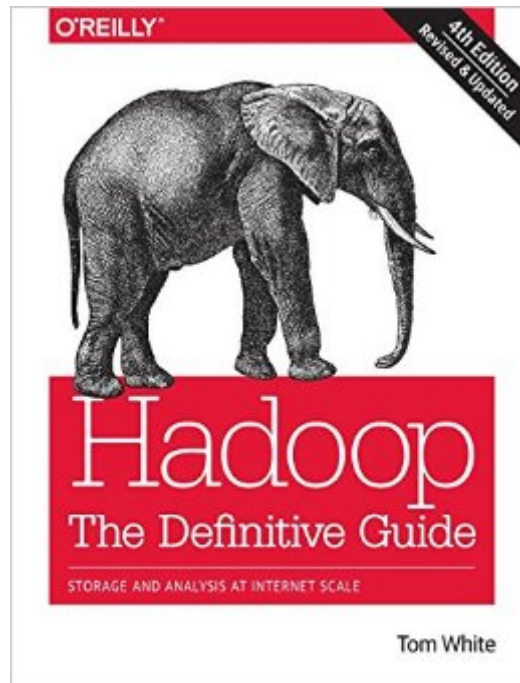
Hadoop: The Definitive Guide

Synopsis

Get ready to unlock the power of your data. With the fourth edition of this comprehensive guide, youâ ™ll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters.Using Hadoop 2 exclusively, author Tom White presents new chapters on YARN and several Hadoop-related projects such as Parquet, Flume, Crunch, and Spark. Youâ ™ll learn about recent changes to Hadoop, and explore new case studies on Hadoopâ ™s role in healthcare systems and genomics data processing.Learn fundamental components such as MapReduce, HDFS, and YARNExplore MapReduce in depth, including steps for developing applications with itSet up and maintain a Hadoop cluster running HDFS and MapReduce on YARNLearn two data formats: Avro for data serialization and Parquet for nested dataUse data ingestion tools such as Flume (for streaming data) and Sqoop (for bulk data transfer)Understand how high-level data processing tools like Pig, Hive, Crunch, and Spark work with HadoopLearn the HBase distributed database and the ZooKeeper distributed configuration service

Book Information

Paperback: 756 pages

Publisher: O'Reilly Media; 4 edition (April 11, 2015)

Language: English

ISBN-10: 1491901632

ISBN-13: 978-1491901632

Product Dimensions: 6.9 x 1.6 x 9.2 inches

Shipping Weight: 2.4 pounds (View shipping rates and policies)

Average Customer Review: 4.6 out of 5 starsÂ Â See all reviewsÂ (39 customer reviews)

Best Sellers Rank: #16,831 in Books (See Top 100 in Books)   #1 inÂ Books > Computers & Technology > Programming > Parallel Programming   #11 inÂ Books > Computers & Technology > Programming > Languages & Tools > Java   #11 inÂ Books > Textbooks > Computer Science > Database Storage & Design

Customer Reviews

I have around 14 years of java experience and this was my first book ever on Hadoop. So far I have been reading from internet. I agree that at times it's hard to understand things in one shot but once you re read them, they get clear. The best thing about this book is that it covers everything in new

API. I believe Tom White worked really hard to make the overall picture easy enough to understand.I would suggest one thing while reading this book, please have the computer and some IDE open while reading it otherwise you will get lost in theory (which is great nonetheless) and when you actually try to code later, things will go off the top of your head.Another thing is that, there are various things which you'll understand only after reaching at 4th or 5th chapter. If you are learning Hadoop (from scratch) and you have read its first few chapters only, you may find it difficult but once you reach at 4th or 5th chapters then I'm sure you will find this book amazing.

Nice update for an already great book. As other reviews have already mentioned, if you're new to Hadoop, you'll probably need to read a few chapters in order to really get a good picture of what Hadoop is, and how it works. This is not necessarily a problem with the book, but just the nature of the beast when learning as complex as Hadoop. If you're new to Hadoop, take the time to work through the first few chapters, and maybe expect to come back and re-read parts later. If you're an experienced engineer who is looking to get up-to-speed with what has changed in recent versions, you'll definitely find what you're looking for here.I appreciate that this book covers high-level concepts as well as dives deep into the technical details that you will need to know for the design, implementation and day-to-day running of Hadoop and its various associated technologies.

This huge guide has everything you need to know about Hadoop. I like that it has little asides that explain some concepts in more detail, as well as detailed examples. I've always had good experience with O'Reilly books. The thing I appreciate the most is that there are instructions on how to actually install Hadoop on your machine in the appendix.

Good overview of the Hadoop ecosystem. Of course, if you need to dive deeper, you will have to pick the component of interest for you and get additional material. This updated version covers the latest and greatest tools in the ecosystem, so, at least for me it was a good investment.

This book is so good that it forced it me to write my first book review on :). Everything is so clearly described and thoroughly explained both for Hadoop and related Apache projects. Read (and re-read) at least the whole section I thoroughly, it is very important for the understanding of the rest of the book.

Great book to clear Hadoop Developer certification

I've been a Java developer for about 20 years... And in code development for 25+ years. Don't get me wrong with the 3 stars, this is a good general book on the subject. I was considering giving it more stars... just be warned that the book expects you to already have a solid understanding on certain topics. Hadoop seems to be very parameter driven, and this book drives thru those parameters. For me, it was hard to take in all the information without multiple re-reads.

I'm a Hadoop newbey, purchase this book to get some inside view.I did like the book, since I had an objective of checking Hadoop and Spark, did jump some chapters, but consider o good book.

Download to continue reading...

Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem (Addison-Wesley Data & Analytics) Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem (Addison-Wesley Data & Analytics Series) Big Data, MapReduce, Hadoop, and Spark with Python: Master Big Data Analytics and Data Wrangling with MapReduce Fundamentals using Hadoop, Spark, and Python Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics Series) Hadoop: The Definitive Guide HTML & XHTML: The Definitive Guide: The Definitive Guide (Definitive Guides) Hadoop Operations 802.11 Wireless Networks: The Definitive Guide: The Definitive Guide Oracle SQL*Plus: The Definitive Guide (Definitive Guides) The Definitive Guide to GCC (Definitive Guides (Paperback)) Vertical Gardening: The Definitive Guide To Vertical Gardening For Beginners. (The Definitive Gardening Guides) The Definitive Guide to MongoDB: A complete guide to dealing with Big Data using MongoDB Raspberry Pi: 101 Beginners Guide: The Definitive Step by Step guide for what you need to know to get started (Raspberry Pi, Raspberry, Single Board Computers, ... Pi Programming, Raspberry Pi Projects) Raspberry Pi 2: 101 Beginners Guide: The Definitive Step by Step guide for what you need to know to get started Raspberry Pi 2: 101 Beginners Guide: The Definitive Step by Step guide for what you need to know to get started (Raspberry Pi 2, Raspberry, Single Board ... Pi Programming, Raspberry Pi Projects) The Champagne Guide 2016-2017: The Definitive Guide to Champagne Foraging: Your Beginner's Guide and Definitive Handbook to Foraging Wild Edible Plants and Herbs(foraging guide, foraging for beginners, foraging for survival, foraging tips, foraging wilderness) Woodworking: 101 Beginners Guide (The Definitive guide for what need to know to start your projects today) (Woodworking, Woodworking tools, Staining, Varnishing, Lacquering)

Alternative Medicine Definitive Guide to Headaches (Alternative Medicine Definative Guide)