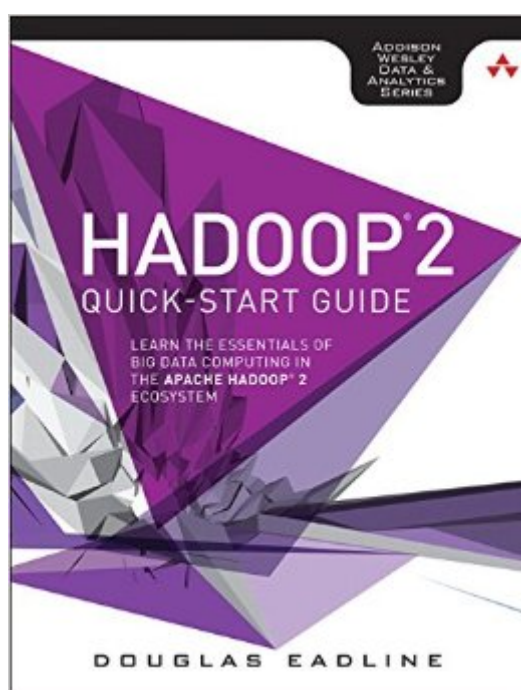


The book was found

Hadoop 2 Quick-Start Guide: Learn The Essentials Of Big Data Computing In The Apache Hadoop 2 Ecosystem (Addison-Wesley Data & Analytics)



Synopsis

Get Started Fast with Apache Hadoop® 2, YARN, and Today's Hadoop Ecosystem

With Hadoop 2.x and YARN, Hadoop moves beyond MapReduce to become practical for virtually any type of data processing. Hadoop 2.x and the Data Lake concept represent a radical shift away from conventional approaches to data usage and storage. Hadoop 2.x installations offer unmatched scalability and breakthrough extensibility that supports new and existing Big Data analytics processing methods and models.

Hadoop® 2 Quick-Start Guide is the first easy, accessible guide to Apache Hadoop 2.x, YARN, and the modern Hadoop ecosystem. Building on his unsurpassed experience teaching Hadoop and Big Data, author Douglas Eadline covers all the basics you need to know to install and use Hadoop 2 on personal computers or servers, and to navigate the powerful technologies that complement it.

Eadline concisely introduces and explains every key Hadoop 2 concept, tool, and service, illustrating each with a simple beginning-to-end example and identifying trustworthy, up-to-date resources for learning more.

This guide is ideal if you want to learn about Hadoop 2 without getting mired in technical details. Douglas Eadline will bring you up to speed quickly, whether you're a user, admin, devops specialist, programmer, architect, analyst, or data scientist.

Coverage Includes

- Understanding what Hadoop 2 and YARN do, and how they improve on Hadoop 1 with MapReduce
- Understanding Hadoop-based Data Lakes versus RDBMS Data Warehouses
- Installing Hadoop 2 and core services on Linux machines, virtualized sandboxes, or clusters
- Exploring the Hadoop Distributed File System (HDFS)
- Understanding the essentials of MapReduce and YARN application programming
- Simplifying programming and data movement with Apache Pig, Hive, Sqoop, Flume, Oozie, and HBase
- Observing application progress, controlling jobs, and managing workflows
- Managing Hadoop efficiently with Apache Ambari – including recipes for HDFS to NFSv3 gateway, HDFS snapshots, and YARN configuration
- Learning basic Hadoop 2 troubleshooting, and installing Apache Hue and Apache Spark

Book Information

Series: Addison-Wesley Data & Analytics

Paperback: 304 pages

Publisher: Addison-Wesley Professional; 1 edition (November 5, 2015)

Language: English

ISBN-10: 0134049942

ISBN-13: 978-0134049946

Product Dimensions: 6.9 x 0.8 x 9 inches

Shipping Weight: 12.6 ounces (View shipping rates and policies)

Average Customer Review: 4.8 out of 5 stars See all reviews (5 customer reviews)

Best Sellers Rank: #268,628 in Books (See Top 100 in Books) #15 in Books > Computers & Technology > Web Development & Design > Website Analytics #162 in Books > Computers & Technology > Databases & Big Data > Data Mining #289 in Books > Textbooks > Computer Science > Database Storage & Design

Customer Reviews

Having never installed or played around with a Hadoop environment myself, I was on the look out for an intro style book that would give me the basics and enough info to start me off. When browsing this one caught my eye as I didn't even realise there was a Hadoop 2 and the title was pretty much spot on for what I was looking for so decided to give it a shot. Overall, I enjoyed the book and it was spot on for what I was looking for. It's a traditional tutorial/walk through type of book on how to get a Hadoop cluster up and running and how to admin/interact with it, but it also covers enough theory that you don't need to have any prior experience with Hadoop to follow along. However, I would say that I think it's overpriced in the paper edition and retail price ebook so if you're interested in this book, try and read it on Safari or get a Kindle edition to make it affordable. Other than that definitely recommended. The book starts off with a really good overview of what Hadoop is, the MapReduce pattern and the changes in Hadoop 2. Good intro material. The next chapter is a more traditional walk through on how to install Hadoop uses both the Hortonworks distribution and the Apache sources. It also covers use of Ambari for a simple web based admin console for your cluster. Nothing too detailed is explained here as it's covered off later, but it's a straight forward walk through so is spot on for that. The third chapter gives a really good intro to how HDFS works, covering the nodes involved, their roles and the approach taken to replication and then some basic file system commands.

[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) 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) Big Data,

MapReduce, Hadoop, and Spark with Python: Master Big Data Analytics and Data Wrangling with MapReduce Fundamentals using Hadoop, Spark, and Python Data Just Right: Introduction to Large-Scale Data & Analytics (Addison-Wesley Data and Analytics) R for Everyone: Advanced Analytics and Graphics (Addison-Wesley Data & Analytics Series) R for Everyone: Advanced Analytics and Graphics (Addison-Wesley Data and Analytics) Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business Leveraging the Power of Data Analytics, Data Science, ... (Hacking Freedom and Data Driven Book 2) Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference (Addison-Wesley Data & Analytics) Big Data For Beginners: Understanding SMART Big Data, Data Mining & Data Analytics For improved Business Performance, Life Decisions & More! TCP/IP Illustrated, Vol. 1: The Protocols (Addison-Wesley Professional Computing Series) TCP/IP Illustrated, Volume 1: The Protocols (Addison-Wesley Professional Computing Series) TCP/IP Illustrated, Volume 1: The Protocols (2nd Edition) (Addison-Wesley Professional Computing Series) TCP/IP Illustrated, Vol. 2: The Implementation (Addison-Wesley Professional Computing Series) Advanced Programming in the UNIX Environment (Addison-Wesley Professional Computing Series) The Go Programming Language (Addison-Wesley Professional Computing Series) Advanced Programming in the UNIX(R) Environment (Addison-Wesley Professional Computing Series) Apache Cookbook: Solutions and Examples for Apache Administrators Ruby on Rails Tutorial: Learn Web Development with Rails (Addison-Wesley Professional Ruby Series)

[Dmca](#)