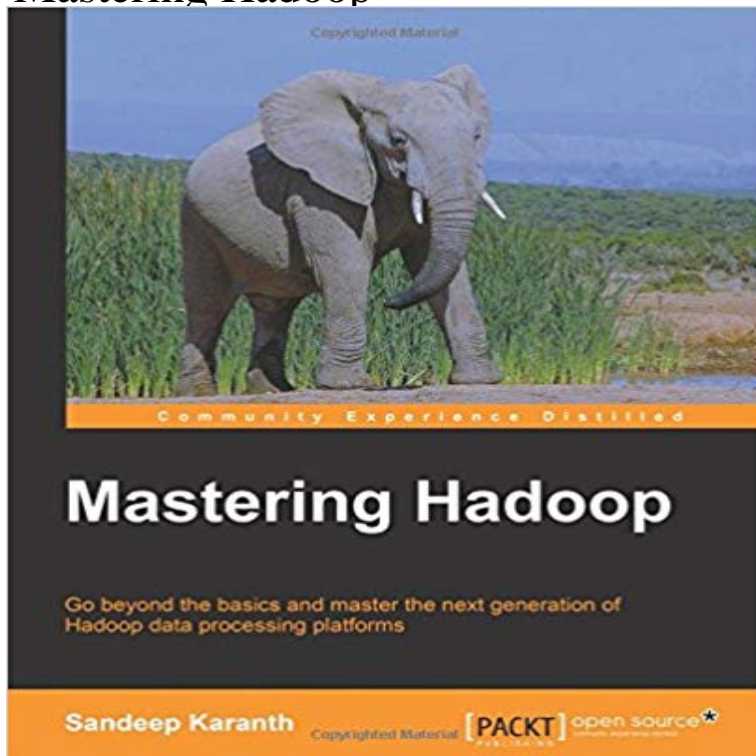


# Mastering Hadoop



Go beyond the basics and master the next generation of Hadoop data processing platforms

About This Book

Learn how to optimize Hadoop MapReduce, Pig and Hive

Dive into YARN and learn how it can integrate Storm with Hadoop

Understand how Hadoop can be deployed on the cloud and gain insights into analytics with Hadoop

Who This Book Is For

Do you want to broaden your Hadoop skill set and take your knowledge to the next level? Do you wish to enhance your knowledge of Hadoop to solve challenging data processing problems? Are your Hadoop jobs, Pig scripts, or Hive queries not working as fast as you intend? Are you looking to understand the benefits of upgrading Hadoop? If the answer is yes to any of these, this book is for you. It assumes novice-level familiarity with Hadoop.

What You Will Learn

- Understand the changes involved in the process in the move from Hadoop 1.0 to Hadoop 2.0
- Customize and optimize MapReduce jobs in Hadoop 2.0
- Explore Hadoop I/O and different data formats
- Dive into YARN and Storm and use YARN to integrate Storm with Hadoop
- Deploy Hadoop on Amazon Elastic MapReduce
- Discover HDFS replacements and learn about HDFS Federation
- Get to grips with Hadoop's main security aspects
- Utilize Mahout and RHadoop for Hadoop analytics

In Detail

Hadoop is synonymous with Big Data processing. Its simple programming model, code once and deploy at any scale paradigm, and an ever-growing ecosystem makes Hadoop an all-encompassing platform for programmers with different levels of expertise. This book explores the industry guidelines to optimize MapReduce jobs and higher-level abstractions such as Pig and Hive in Hadoop 2.0. Then, it dives deep into Hadoop 2.0 specific features such as YARN and HDFS Federation. This book is a step-by-step guide that focuses on advanced Hadoop concepts and aims to

take your Hadoop knowledge and skill set to the next level. The data processing flow dictates the order of the concepts in each chapter, and each chapter is illustrated with code fragments or schematic diagrams.

**Apache Mahout - Mastering Hadoop** YARN can bring in other computing paradigms to Hadoop. In Hadoop 2.X, MapReduce, Pig, and Hive are all Application Master libraries and their **The First 3 Steps to Mastering Hadoop, Big Data and NoSQL - NAYA** Apache Mahout is a scalable machine learning library. It is an open source library under the Apache Software Foundation. It supports algorithms for **Summary - Mastering Hadoop** Go beyond the basics and master the next generation of Hadoop data processing platforms. : **Buy Mastering Hadoop Book Online at Low Prices in** RHadoop Summary for Microsoft Windows Deploying Hadoop on Microsoft Windows Prerequisites Building Hadoop Configuring Hadoop Deploying **Data serialization in Hadoop - Mastering Hadoop** Chapter 4. Advanced HiveSQL is a popular data-processing language that has been around for four decades. There are scores of people who are already **Advanced Hive - Mastering Hadoop** Big data processing involves data representation either in storage or in transit over the network. Compact representation, fast transformations, **Hadoop on the cloud - Mastering Hadoop** Based on the situation, it might be a necessity to replace HDFS with a filesystem of your choice. Hadoop provides out-of-the-box support for a few **Buy Mastering Hadoop Book Online at Low Prices in India** Like Hadoop, YARN has a script that provides commands to manage YARN. The commands are of the following two kinds: User commands: These are **Implementing a filesystem in Hadoop - Mastering Hadoop** Interactive mode: In this mode, a grunt shell is provided to the user. Users can type in Pig commands in an interactive session with Pig and the Hadoop cluster. **Developing YARN applications - Mastering Hadoop** Though we see data in a structured form, the raw form of data is a sequence or stream of bits. This raw form of data is the one that travels over the network and is **Mastering Hadoop: Sandeep Karanth: 9781783983643: Amazon** Mastering Big Data with real-world Hadoop use cases. This e-book contains 10 real world use cases with complete source code and explanation. **Hadoops - source url** Hadoop is synonymous with Big Data processing. Its simple programming model, code once and deploy at any scale paradigm, and an ever-growing **Mastering-Hadoop-Example-Code/3643OS\_12\_ at master** Like Hadoop, YARN has a script that provides commands to manage YARN. The commands are of the following two kinds: User commands: These are Go beyond the basics and master the next generation of Hadoop data processing platforms In Detail Hadoop is synonymous with Big Data processing. Its simple **Real-world Hadoop Use Cases E-Book AcadGild Blog** Buy Mastering Hadoop on ? FREE SHIPPING on qualified orders. **Summary - Mastering Hadoop** HDFS is a great filesystem for MapReduce workloads. But its sequential access pattern and non-compliance with POSIX interfaces make it tedious to work with **Summary - Mastering Hadoop** Buy the Paperback Book Mastering Hadoop by Sandeep Karanth at , Canadas largest

bookstore. + Get Free Shipping on Computers **Mastering Hadoop - Google Books Result** I came across a book Mastering Hadoop published by Packt and authored by Sandeep Karath. Here is my detail review about the book- **Mastering Hadoop PACKT Books - Packt Publishing** - Buy Mastering Hadoop book online at best prices in India on Amazon.in. Read Mastering Hadoop book reviews & author details and more at **Mastering Hadoop - Packt Publishing** Authentication in Hadoop can either be simple or in Kerberos. Hadoop also allows you to have your custom authentication scheme. In this section, we will **Different modes of execution - Mastering Hadoop** In this chapter, you saw optimizations at different stages of the Hadoop MapReduce pipeline. With the join example, we saw a few other advanced features **Authentication in Hadoop - Mastering Hadoop** All major cloud service providers have Hadoop as a PaaS offering. Amazon with Elastic MapReduce, Microsoft with their HDInsight offering, and Google with **Summary - Mastering Hadoop** Security becomes a primary feature in a multitenant and distributed environment. Networks and resource sharing can potentially lead to information leaks via **Mastering Hadoop by Sandeep Karanth OverDrive: eBooks** - Buy Mastering Hadoop book online at best prices in india on Amazon.in. Read Mastering Hadoop book reviews & author details and more at **Mastering Hadoop** Mastering Hadoop Example Code. Contribute to Mastering-Hadoop-Example-Code development by creating an account on GitHub. **Mi Gems: Mastering Hadoop: Book Review** Mastering Hadoop. Go beyond the basics and master the next generation of Hadoop data processing platforms. Sandeep Karanth. BIRMINGHAM - MUMBAI. **Mastering Hadoop - OReilly Media** Go beyond the basics and master the next generation of Hadoop data processing platforms About This Book Learn how to optimize Hadoop MapReduce, Pig **YARN commands - Mastering Hadoop** Do you want to broaden your Hadoop skill set and take your knowledge to the next level? Do you wish to enhance your knowledge of Hadoop to solve