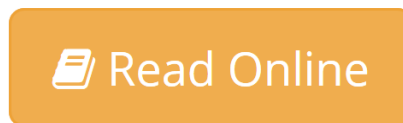


# Hadoop 2.x Administration Cookbook

By Gurmukh Singh



## Hadoop 2.x Administration Cookbook By Gurmukh Singh

### Key Features

- Become an expert Hadoop administrator and perform tasks to optimize your Hadoop Cluster
- Import and export data into Hive and use Oozie to manage workflow.
- Practical recipes will help you plan and secure your Hadoop cluster, and make it highly available

### Book Description

Hadoop enables the distributed storage and processing of large datasets across clusters of computers. Learning how to administer Hadoop is crucial to exploit its unique features. With this book, you will be able to overcome common problems encountered in Hadoop administration.

The book begins with laying the foundation by showing you the steps needed to set up a Hadoop cluster and its various nodes. You will get a better understanding of how to maintain Hadoop cluster, especially on the HDFS layer and using YARN and MapReduce. Further on, you will explore durability and high availability of a Hadoop cluster.

You'll get a better understanding of the schedulers in Hadoop and how to configure and use them for your tasks. You will also get hands-on experience with the backup and recovery options and the performance tuning aspects of Hadoop. Finally, you will get a better understanding of troubleshooting, diagnostics, and best practices in Hadoop administration.

By the end of this book, you will have a proper understanding of working with Hadoop clusters and will also be able to secure, encrypt it, and configure auditing for your Hadoop clusters.

### What you will learn

- Set up the Hadoop architecture to run a Hadoop cluster smoothly
- Maintain a Hadoop cluster on HDFS, YARN, and MapReduce
- Understand high availability with Zookeeper and Journal Node
- Configure Flume for data ingestion and Oozie to run various workflows
- Tune the Hadoop cluster for optimal performance

- Schedule jobs on a Hadoop cluster using the Fair and Capacity scheduler
- Secure your cluster and troubleshoot it for various common pain points

### About the Author

**Gurmukh Singh** is a seasoned technology professional with 14+ years of industry experience in infrastructure design, distributed systems, performance optimization, and networks. He has worked in big data domain for the last 5 years and provides consultancy and training on various technologies.

He has worked with companies such as HP, JP Morgan, and Yahoo.

He has authored Monitoring Hadoop by Packt Publishing

### Table of Contents

1. Hadoop Architecture and Deployment
2. Maintain Hadoop Cluster - HDFS
3. Maintain Hadoop Cluster -YARN and MapReduce
4. High Availability
5. Schedulers
6. Backup and Recovery
7. Data Ingestion and Workflow
8. Performance Tuning
9. Hbase and RDBMS
10. Cluster Planning
11. Troubleshooting, Diagnostics and Best practises
12. Security

 [Download Hadoop 2.x Administration Cookbook ...pdf](#)

 [Read Online Hadoop 2.x Administration Cookbook ...pdf](#)

# Hadoop 2.x Administration Cookbook

*By Gurmukh Singh*

**Hadoop 2.x Administration Cookbook** By Gurmukh Singh

## Key Features

- Become an expert Hadoop administrator and perform tasks to optimize your Hadoop Cluster
- Import and export data into Hive and use Oozie to manage workflow.
- Practical recipes will help you plan and secure your Hadoop cluster, and make it highly available

## Book Description

Hadoop enables the distributed storage and processing of large datasets across clusters of computers. Learning how to administer Hadoop is crucial to exploit its unique features. With this book, you will be able to overcome common problems encountered in Hadoop administration.

The book begins with laying the foundation by showing you the steps needed to set up a Hadoop cluster and its various nodes. You will get a better understanding of how to maintain Hadoop cluster, especially on the HDFS layer and using YARN and MapReduce. Further on, you will explore durability and high availability of a Hadoop cluster.

You'll get a better understanding of the schedulers in Hadoop and how to configure and use them for your tasks. You will also get hands-on experience with the backup and recovery options and the performance tuning aspects of Hadoop. Finally, you will get a better understanding of troubleshooting, diagnostics, and best practices in Hadoop administration.

By the end of this book, you will have a proper understanding of working with Hadoop clusters and will also be able to secure, encrypt it, and configure auditing for your Hadoop clusters.

## What you will learn

- Set up the Hadoop architecture to run a Hadoop cluster smoothly
- Maintain a Hadoop cluster on HDFS, YARN, and MapReduce
- Understand high availability with Zookeeper and Journal Node
- Configure Flume for data ingestion and Oozie to run various workflows
- Tune the Hadoop cluster for optimal performance
- Schedule jobs on a Hadoop cluster using the Fair and Capacity scheduler
- Secure your cluster and troubleshoot it for various common pain points

## About the Author

**Gurmukh Singh** is a seasoned technology professional with 14+ years of industry experience in infrastructure design, distributed systems, performance optimization, and networks. He has worked in big data domain for the last 5 years and provides consultancy and training on various technologies.

He has worked with companies such as HP, JP Morgan, and Yahoo.

He has authored Monitoring Hadoop by Packt Publishing

### Table of Contents

1. Hadoop Architecture and Deployment
2. Maintain Hadoop Cluster - HDFS
3. Maintain Hadoop Cluster -YARN and MapReduce
4. High Availability
5. Schedulers
6. Backup and Recovery
7. Data Ingestion and Workflow
8. Performance Tuning
9. Hbase and RDBMS
10. Cluster Planning
11. Troubleshooting, Diagnostics and Best practises
12. Security

### Hadoop 2.x Administration Cookbook By Gurmukh Singh Bibliography

- Rank: #2292760 in eBooks
- Published on: 2017-05-26
- Released on: 2017-05-26
- Format: Kindle eBook

 [Download Hadoop 2.x Administration Cookbook ...pdf](#)

 [Read Online Hadoop 2.x Administration Cookbook ...pdf](#)

## Download and Read Free Online Hadoop 2.x Administration Cookbook By Gurmukh Singh

---

### Editorial Review

About the Author

#### Gurmukh Singh

Gurmukh Singh is a seasoned technology professional with 14+ years of industry experience in infrastructure design, distributed systems, performance optimization, and networks. He has worked in big data domain for the last 5 years and provides consultancy and training on various technologies. He has worked with companies such as HP, JP Morgan, and Yahoo. He has authored Monitoring Hadoop by Packt Publishing

### Users Review

From reader reviews:

#### William Herold:

The book Hadoop 2.x Administration Cookbook make one feel enjoy for your spare time. You need to use to make your capable more increase. Book can to get your best friend when you getting stress or having big problem with the subject. If you can make studying a book Hadoop 2.x Administration Cookbook to become your habit, you can get more advantages, like add your current capable, increase your knowledge about some or all subjects. You could know everything if you like available and read a e-book Hadoop 2.x Administration Cookbook. Kinds of book are a lot of. It means that, science book or encyclopedia or other individuals. So , how do you think about this book?

#### Kristine Toomey:

In this 21st millennium, people become competitive in most way. By being competitive currently, people have do something to make these survives, being in the middle of often the crowded place and notice by means of surrounding. One thing that oftentimes many people have underestimated the item for a while is reading. Yep, by reading a book your ability to survive improve then having chance to stand up than other is high. In your case who want to start reading any book, we give you this specific Hadoop 2.x Administration Cookbook book as beginner and daily reading publication. Why, because this book is usually more than just a book.

#### David Reed:

Hey guys, do you would like to finds a new book to study? May be the book with the name Hadoop 2.x Administration Cookbook suitable to you? Typically the book was written by renowned writer in this era. The actual book untitled Hadoop 2.x Administration Cookbookis the main one of several books that everyone read now. That book was inspired a number of people in the world. When you read this publication you will enter the new age that you ever know ahead of. The author explained their concept in the simple way, therefore all of people can easily to be aware of the core of this publication. This book will give you a great deal of information about this world now. So that you can see the represented of the world on this

book.

**Willie Briggs:**

A lot of book has printed but it is unique. You can get it by internet on social media. You can choose the most beneficial book for you, science, comedian, novel, or whatever through searching from it. It is named of book Hadoop 2.x Administration Cookbook. You can add your knowledge by it. Without leaving behind the printed book, it can add your knowledge and make you happier to read. It is most crucial that, you must aware about publication. It can bring you from one location to other place.

**Download and Read Online Hadoop 2.x Administration Cookbook  
By Gurmukh Singh #EUJIFKY6174**

## **Read Hadoop 2.x Administration Cookbook By Gurmukh Singh for online ebook**

Hadoop 2.x Administration Cookbook By Gurmukh Singh Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Hadoop 2.x Administration Cookbook By Gurmukh Singh books to read online.

### **Online Hadoop 2.x Administration Cookbook By Gurmukh Singh ebook PDF download**

**Hadoop 2.x Administration Cookbook By Gurmukh Singh Doc**

**Hadoop 2.x Administration Cookbook By Gurmukh Singh Mobipocket**

**Hadoop 2.x Administration Cookbook By Gurmukh Singh EPub**