Mastering Apache Spark

Mike Frampton

Mastering Apache Spark

Mike Frampton

Mastering Apache Spark Mike Frampton

Gain expertise in processing and storing data by using advanced techniques with Apache Spark

About This Book

- Explore the integration of Apache Spark with third party applications such as H20, Databricks and Titan
- Evaluate how Cassandra and Hbase can be used for storage
- An advanced guide with a combination of instructions and practical examples to extend the most up-to date Spark functionalities

Who This Book Is For

If you are a developer with some experience with Spark and want to strengthen your knowledge of how to get around in the world of Spark, then this book is ideal for you. Basic knowledge of Linux, Hadoop and Spark is assumed. Reasonable knowledge of Scala is expected.

What You Will Learn

- Extend the tools available for processing and storage
- Examine clustering and classification using MLlib
- Discover Spark stream processing via Flume, HDFS
- Create a schema in Spark SQL, and learn how a Spark schema can be populated with data
- Study Spark based graph processing using Spark GraphX
- Combine Spark with H20 and deep learning and learn why it is useful
- Evaluate how graph storage works with Apache Spark, Titan, HBase and Cassandra
- Use Apache Spark in the cloud with Databricks and AWS

In Detail

Apache Spark is an in-memory cluster based parallel processing system that provides a wide range of functionality like graph processing, machine learning, stream processing and SQL. It operates at unprecedented speeds, is easy to use and offers a rich set of data transformations.

This book aims to take your limited knowledge of Spark to the next level by teaching you how to expand Spark functionality. The book commences with an overview of the Spark eco-system. You will learn how to use MLlib to create a fully working neural net for handwriting recognition. You will then discover how stream processing can be tuned for optimal performance and to ensure parallel processing. The book extends to show how to incorporate H20 for machine learning, Titan for graph based storage, Databricks for cloud-based Spark. Intermediate Scala based code examples are provided for Apache Spark module processing in a CentOS Linux and Databricks cloud environment.

Style and approach

This book is an extensive guide to Apache Spark modules and tools and shows how Spark's functionality can be extended for real-time processing and storage with worked examples.

Download and Read Free Online Mastering Apache Spark Mike Frampton

---

From reader reviews:

Shirley Glover:

Have you spare time for any day? What do you do when you have a lot more or little spare time? That's why, you can choose the suitable activity regarding spend your time. Any person spent their own spare time to take a walk, shopping, or went to the particular Mall. How about open as well as read a book allowed Mastering Apache Spark? Maybe it is to be best activity for you. You realize beside you can spend your time together with your favorite's book, you can cleverer than before. Do you agree with it has the opinion or you have some other opinion?

Aimee Nguyen:

The book Mastering Apache Spark can give more knowledge and also the precise product information about everything you want. Why then must we leave the great thing like a book Mastering Apache Spark? A number of you have a different opinion about publication. But one aim which book can give many data for us. It is absolutely proper. Right now, try to closer along with your book. Knowledge or data that you take for that, you could give for each other; you can share all of these. Book Mastering Apache Spark has simple shape however you know: it has great and big function for you. You can appearance the enormous world by open up and read a reserve. So it is very wonderful.

Joan Munoz:

Hey guys, do you really wants to finds a new book to learn? May be the book with the concept Mastering Apache Spark suitable to you? The book was written by well-known writer in this era. The actual book untitled Mastering Apache Sparkis the main of several books that will everyone read now. That book was inspired many men and women in the world. When you read this guide you will enter the new age that you ever know prior to. The author explained their concept in the simple way, consequently all of people can easily to understand the core of this e-book. This book will give you a lot of information about this world now. To help you to see the represented of the world in this particular book.

Jennifer David:

Do you really one of the book lovers? If so, do you ever feeling doubt when you are in the book store? Make an effort to pick one book that you find out the inside because don't judge book by its protect may doesn't work here is difficult job because you are scared that the inside maybe not because fantastic as in the outside appear likes. Maybe you answer could be Mastering Apache Spark why because the great cover that make you consider about the content will not disappoint you. The inside or content is actually fantastic as the outside as well as cover. Your reading 6th sense will directly show you to pick up this book.

Download and Read Online Mastering Apache Spark Mike Frampton #GIOX0N1BZV9

Read Mastering Apache Spark by Mike Frampton for online ebook

Mastering Apache Spark by Mike Frampton 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 Mastering Apache Spark by Mike Frampton books to read online.

Online Mastering Apache Spark by Mike Frampton ebook PDF download

Mastering Apache Spark by Mike Frampton Doc

Mastering Apache Spark by Mike Frampton Mobipocket

Mastering Apache Spark by Mike Frampton EPub