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:

Helen Sullivan:

The e-book untitled Mastering Apache Spark is the book that recommended to you to read. You can see the quality of the reserve content that will be shown to you. The language that article author use to explained their ideas are easily to understand. The copy writer was did a lot of exploration when write the book, hence the information that they share for you is absolutely accurate. You also might get the e-book of Mastering Apache Spark from the publisher to make you a lot more enjoy free time.

Lloyd Schuler:

Reading can called brain hangout, why? Because when you find yourself reading a book specially book entitled Mastering Apache Spark your mind will drift away trough every dimension, wandering in most aspect that maybe unfamiliar for but surely will end up your mind friends. Imaging just about every word written in a reserve then become one type conclusion and explanation this maybe you never get before. The Mastering Apache Spark giving you one more experience more than blown away your head but also giving you useful details for your better life in this particular era. So now let us teach you the relaxing pattern this is your body and mind is going to be pleased when you are finished looking at it, like winning a sport. Do you want to try this extraordinary spending spare time activity?

Robert Murphy:

Do you really one of the book lovers? If yes, do you ever feeling doubt when you are in the book store? Try to pick one book that you find out the inside because don't evaluate book by its handle may doesn't work the following is difficult job because you are frightened that the inside maybe not seeing that fantastic as in the outside look likes. Maybe you answer might be Mastering Apache Spark why because the wonderful cover that make you consider regarding the content will not disappoint an individual. The inside or content is actually fantastic as the outside as well as cover. Your reading sixth sense will directly assist you to pick up this book.

Herbert Oakley:

In this era which is the greater person or who has ability in doing something more are more special than other. Do you want to become certainly one of it? It is just simple solution to have that. What you need to do is just spending your time almost no but quite enough to get a look at some books. One of many books in the top listing in your reading list will be Mastering Apache Spark. This book that is certainly qualified as The Hungry Inclines can get you closer in growing to be precious person. By looking right up and review this guide you can get many advantages.

Download and Read Online Mastering Apache Spark Mike Frampton #5ZRBT4GP62V

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