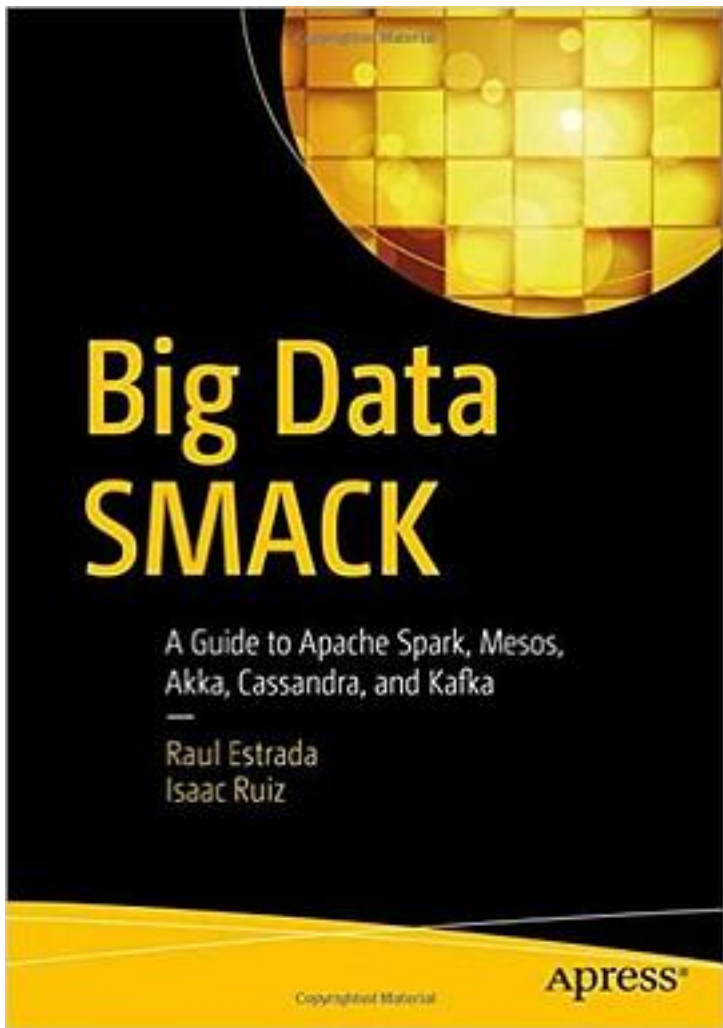


Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka



[Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka_下载链接1](#)

著者:Raul Estrada

出版者:Apress

出版时间:2016-10-31

装帧:Paperback

isbn:9781484221747

This book is about how to integrate full-stack open source big data architecture and how to choose the correct technology—Scala/Spark, Mesos, Akka, Cassandra, and Kafka—in every layer. Big data architecture is becoming a requirement for many different enterprises. So far, however, the focus has largely been on collecting, aggregating, and crunching large datasets in a timely manner. In many cases now, organizations need more than one paradigm to perform efficient analyses.

Big Data SMACK explains each of the full-stack technologies and, more importantly, how to best integrate them. It provides detailed coverage of the practical benefits of these technologies and incorporates real-world examples in every situation. The book focuses on the problems and scenarios solved by the architecture, as well as the solutions provided by every technology. It covers the six main concepts of big data architecture and how integrate, replace, and reinforce every layer:

The language: Scala
The engine: Spark (SQL, MLib, Streaming, GraphX)
The container: Mesos, Docker
The view: Akka
The storage: Cassandra
The message broker: Kafka

What you'll learn

How to make big data architecture without using complex Greek letter architectures.
How to build a cheap but effective cluster infrastructure.
How to make queries, reports, and graphs that business demands.
How to manage and exploit unstructured and No-SQL data sources.
How use tools to monitor the performance of your architecture.
How to integrate all technologies and decide which replace and which reinforce.

Who This Book Is For

This book is for developers, data architects, and data scientists looking for how to integrate the most successful big data open stack architecture and how to choose the correct technology in every layer.

作者介绍:

From the Back Cover

Integrate full-stack open-source fast data pipeline architecture and choose the correct technology—Spark, Mesos, Akka, Cassandra, and Kafka (SMACK)—in every layer. Fast data is becoming a requirement for many enterprises. So far, however, the focus has largely been on collecting, aggregating, and crunching large data sets in a timely manner. In many cases organizations need more than one paradigm to perform efficient analyses. Big Data SMACK explains each technology and, more importantly, how to integrate them. It provides detailed coverage of the practical benefits of these technologies and incorporates real-world examples. The book focuses on the problems and scenarios solved by the architecture, as well as the solutions provided by each technology. This book covers the five main concepts of data pipeline architecture and how to integrate, replace, and reinforce every layer:
The engine: Apache Spark
The container: Apache Mesos
The model: Akka
The storage: Apache Cassandra
The broker: Apache Kafka

Read more

About the Author

Raúl Estrada is the co-founder of Treu Technologies, an enterprise for Social Data Marketing and BigData research. Estrada is an Enterprise Architect with more than 15 years of experience in cluster management and Enterprise Software. Prior to founding Treu Technologies, Estrada worked as an Enterprise Architect in Application Servers & evangelist for Oracle Inc. Estrada loves functional languages like Elixir and Scala, and also has a Master degree on Computer Science. Isaac Ruiz is a Java programmer since 2001, and a consultant and architect since 2003. Ruiz had participated in projects of different areas and varied scopes (education, communications, retail, and others). Ruiz specializes in systems integration and has participated in projects mainly related to the financial sector. Ruiz is a supporter of free software. Ruiz like to experiment with new technologies (frameworks, languages, methods).

Read more

目录:

[Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka 下载链接1](#)

标签

数据科学

英文版

编程

DataScience

评论

主要是对Spark Mesos Akka Cassandra Kafka做一个简短的简介，以及这些组件如何组合使用。可以对大数据的架构有个简单的认识。

如果想了解些入门概念可以看看，否则严重不推荐，纯粹就是挂了smack的噱头写的一份上手指南。

[Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka_下载链接1](#)

书评

This book is about how to integrate full-stack open source big data architecture and how to choose the correct technology—Scala/Spark, Mesos, Akka, Cassandra, and Kafka—in every layer. Big data architecture is becoming a requirement for many different ent...

[Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka_下载链接1](#)