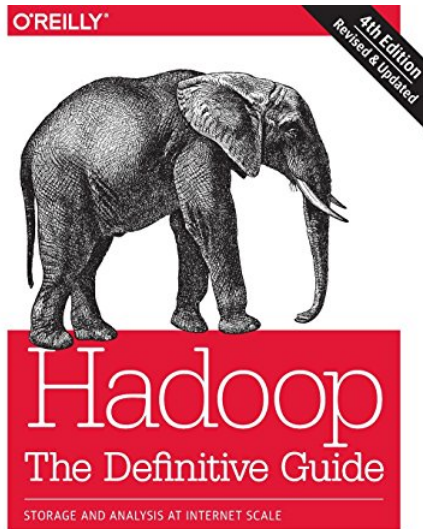


yU73g [Free PDF File] Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale



Tom White

CONTINUE ►

Get ready to unlock the power of your data. With the fourth edition of this comprehensive guide, you'll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. Using Hadoop 2 exclusively, author Tom White presents new chapters on YARN and several Hadoop-related projects such as Parquet, Flume, Crunch, and Spark. You'll learn about recent changes to Hadoop, and explore new case studies on Hadoop's role in healthcare systems and genomics data processing. Learn fundamental components such as MapReduce, HDFS, and YARN. Explore MapReduce in depth, including steps for developing applications with it. Set up and maintain a Hadoop cluster running HDFS and MapReduce on YARN. Learn two data formats: Avro for data serialization and Parquet for nested data. Use data ingestion tools such as Flume (for streaming data) and Sqoop (for bulk data transfer). Understand how high-level data processing tools like Pig, Hive, Crunch, and Spark work with Hadoop. Learn the HBase distributed database and the ZooKeeper distributed configuration service.

Get ready to unlock the power of your data. With the fourth edition of this comprehensive guide, you'll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. Using Hadoop 2 exclusively, author Tom White presents new chapters on YARN and several Hadoop-related projects such as Parquet, Flume, Crunch, and Spark. You'll learn about recent changes to Hadoop, and explore new case studies on Hadoop's role in healthcare systems and genomics data processing. Learn fundamental components such as MapReduce, HDFS, and YARN. Explore MapReduce in depth, including steps for developing applications with it. Set up and maintain a Hadoop cluster running HDFS and MapReduce on YARN. Learn two data formats: Avro for data serialization and Parquet for nested data. Use data ingestion tools such as Flume (for streaming data) and Sqoop (for bulk data transfer). Understand how high-level data processing tools like Pig, Hive, Crunch, and Spark work with Hadoop. Learn the HBase distributed database and the ZooKeeper distributed configuration service.

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale pdf free

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale epub download

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale online

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale epub download

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale epub vk

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale pdf download

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale read online

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale epub

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale vk

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale pdf

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale amazon

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale free download pdf

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale mobi

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale PDF - KINDLE - EPUB - MOBI

Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale download ebook PDF EPUB, book in english language
[download] Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale in format PDF
Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale download free of book in format