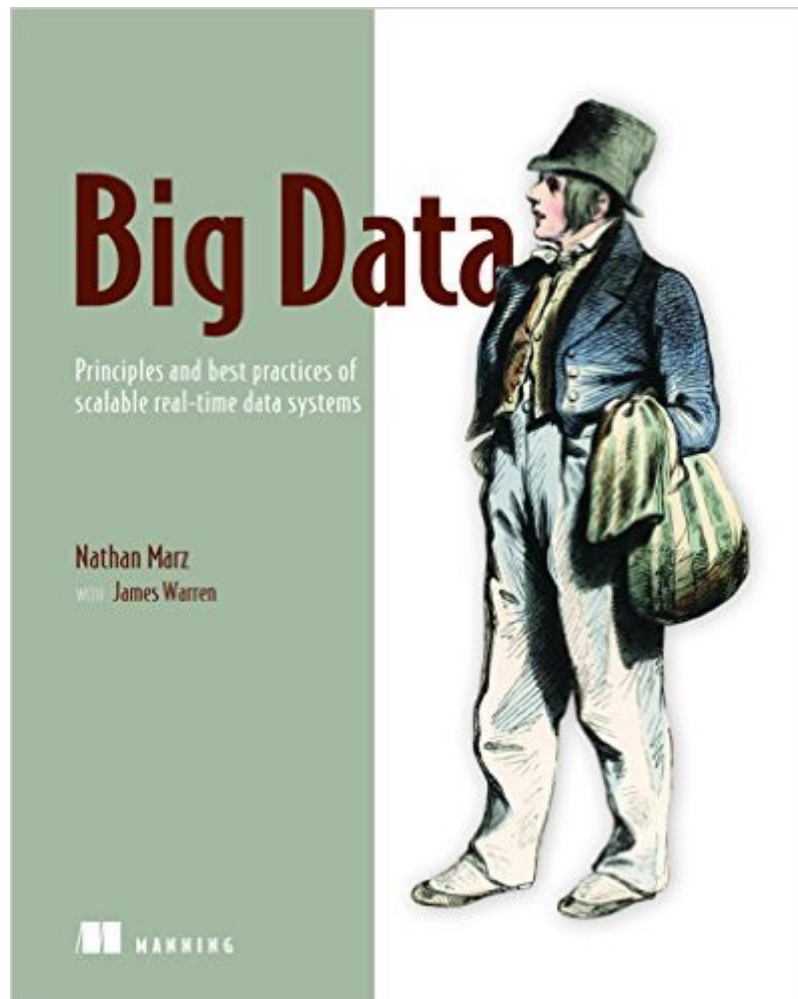


The book was found

Big Data: Principles And Best Practices Of Scalable Realtime Data Systems



Synopsis

SummaryBig Data teaches you to build big data systems using an architecture that takes advantage of clustered hardware along with new tools designed specifically to capture and analyze web-scale data. It describes a scalable, easy-to-understand approach to big data systems that can be built and run by a small team. Following a realistic example, this book guides readers through the theory of big data systems, how to implement them in practice, and how to deploy and operate them once they're built.Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.About the BookWeb-scale applications like social networks, real-time analytics, or e-commerce sites deal with a lot of data, whose volume and velocity exceed the limits of traditional database systems. These applications require architectures built around clusters of machines to store and process data of any size, or speed. Fortunately, scale and simplicity are not mutually exclusive.Big Data teaches you to build big data systems using an architecture designed specifically to capture and analyze web-scale data. This book presents the Lambda Architecture, a scalable, easy-to-understand approach that can be built and run by a small team. You'll explore the theory of big data systems and how to implement them in practice. In addition to discovering a general framework for processing big data, you'll learn specific technologies like Hadoop, Storm, and NoSQL databases.This book requires no previous exposure to large-scale data analysis or NoSQL tools. Familiarity with traditional databases is helpful.What's InsideIntroduction to big data systemsReal-time processing of web-scale dataTools like Hadoop, Cassandra, and StormExtensions to traditional database skillsAbout the AuthorsNathan Marz is the creator of Apache Storm and the originator of the Lambda Architecture for big data systems. James Warren is an analytics architect with a background in machine learning and scientific computing.Table of ContentsA new paradigm for Big DataPART 1 BATCH LAYERData model for Big DataData model for Big Data: IllustrationData storage on the batch layerData storage on the batch layer: IllustrationBatch layerBatch layer: IllustrationAn example batch layer: Architecture and algorithmsAn example batch layer: ImplementationPART 2 SERVING LAYERServing layerServing layer: IllustrationPART 3 SPEED LAYERRealtime viewsRealtime views: IllustrationQueuing and stream processingQueuing and stream processing: IllustrationMicro-batch stream processingMicro-batch stream processing: IllustrationLambda Architecture in depth

Book Information

Paperback: 328 pages

Publisher: Manning Publications; 1 edition (May 10, 2015)

Language: English

ISBN-10: 1617290343

ISBN-13: 978-1617290343

Product Dimensions: 7.3 x 0.6 x 9.1 inches

Shipping Weight: 1.8 pounds (View shipping rates and policies)

Average Customer Review: 4.4 out of 5 stars [See all reviews](#) (32 customer reviews)

Best Sellers Rank: #32,670 in Books (See Top 100 in Books) #3 in [Books > Computers & Technology > Web Development & Design > User Generated Content](#) #9 in [Books > Computers & Technology > Business Technology > Management Information Systems](#) #19 in [Books > Computers & Technology > Databases & Big Data > Data Mining](#)

Customer Reviews

Here's my bottom line: Get this book, whether you are new to working with Big Data or now an old hand at dealing with Big Data's seemingly never-ending (and steadily expanding) complexities. You may not agree with all that the authors offer or contend in this well-written "theory" text. But Nathan Marz's Lambda Architecture is well worth serious consideration, especially if you are now trying to come up with more reliable and more efficient approaches to processing and mining Big Data. The writers' explanations of some of the power, problems, and possibilities of Big Data systems are among the clearest and best I have read. "More than 30,000 gigabytes of data are generated every second, and the rate of data creation is only accelerating," Marz and Warren point out. Thus, previous "solutions" for working with Big Data are now getting overwhelmed, not only by the sheer volume of information pouring in but by greater system complexities and failures of overworked hardware that now plague many outmoded systems. The authors have structured their book to show "how to approach building a solution to any Big Data problem. The principles you'll learn hold true regardless of the tooling in the current landscape, and you can use these principles to rigorously choose what tools are appropriate for your application." • In other words, they write, you will learn how to fish, not just how to use a particular fishing rod. • However, a particular Big Data architecture is featured, as well: Marz's Lambda Architecture. It is, the two authors explain, "an architecture that takes advantage of clustered hardware along with new tools designed specifically to capture and analyze web-scale data."

[Download to continue reading...](#)

Big Data: Principles and best practices of scalable realtime data systems
Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business,

Predictive Analysis, Big Data) Florida Real Estate Principles, Practices & Law (Florida Real Estate Principles, Practices and Law) Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business. Leveraging the Power of Data Analytics, Data ... (Hacking Freedom and Data Driven) (Volume 2) RealTime Physics Active Learning Laboratories, Module 1: Mechanics Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications) Best of the Best from Big Sky Cookbook: Selected Recipes from the Favorite Cookbooks of Montana and Wyoming (Best of the Best Cookbook Series) Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data From Big Data to Big Profits: Success with Data and Analytics Java for the Web with Servlets, JSP, and EJB: A Developer's Guide to J2EE Solutions: A Developer's Guide to Scalable Solutions Serverless Single Page Apps: Fast, Scalable, and Available Building Scalable Apps with Redis and Node.js FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics (Springer Tracts in Advanced Robotics) The Art of Scalability: Scalable Web Architecture, Processes, and Organizations for the Modern Enterprise Scalable Innovation: A Guide for Inventors, Entrepreneurs, and IP Professionals Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results CCNP Building Scalable Internetworks (BSCI 642-901) Lab Portfolio (Cisco Networking Academy) The Startup Checklist: 25 Steps to a Scalable, High-Growth Business Industrial Steam Systems: Fundamentals and Best Design Practices

[Dmca](#)