Mastering Apache Cassandra Second Edition

c5e34f1302021f4ed71c49f735dc0add

Mastering C++ Programming

Mastering vRealize Operations Manager

Learn Azure in a Month of Lunches, Second Edition

Mastering Julia

Imagine what you could do if scalability wasn't a problem. With this hands-on guide, you'll learn how the Cassandra database management system handles hundreds of terabytes of data while remaining highly available across multiple data centers. This expanded second edition—updated for Cassandra 3.0—provides the technical details and practical examples you need to put this database to work in a production environment. Authors Jeff Carpenter and Eben Hewitt demonstrate the advantages of Cassandra’s non-relational design, with special attention to data modeling. If you’re a developer, DBA, or application architect looking to solve a database scaling issue or future-proof your application, this guide helps you harness Cassandra’s speed and flexibility. Understand Cassandra’s distributed and decentralized structure. Use the Cassandra Query Language (CQL) and cqlsh—the CQL shell. Create a working data model and compare it with an equivalent relational model. Develop sample applications using client drivers for languages including Java, Python, and Node.js. Explore cluster topology and learn how nodes exchange data. Maintain a high level of performance in your cluster. Deploy Cassandra on site, in the Cloud, or with Docker. Integrate Cassandra with Spark, Hadoop, Elasticsearch, Solr, and Lucene.
Mastering Cloud Computing is designed for undergraduate students learning to develop cloud computing applications. Tomorrow’s applications won’t live on a single computer but will be deployed from and reside on a virtual server, accessible anywhere, any time. Tomorrow’s application developers need to understand the requirements of building apps for these virtual systems, including concurrent programming, high-performance computing, and data-intensive systems. The book introduces the principles of distributed and parallel computing underlying cloud architectures and specifically focuses on virtualization, thread programming, task programming, and map-reduce programming. There are examples demonstrating all of these and more, with exercises and labs throughout. Explains how to make design choices and tradeoffs to consider when building applications to run in a virtual cloud environment Real-world case studies include scientific, business, and energy-efficiency considerations

Mastering Node.js

This book takes you through tried and tested approaches to building distributed systems and implementing microservices architecture. It follows a single real-world project from start to finish, using Spring Boot, Spring Cloud, and a full suite of related tools and frameworks for development, security, testing, and deployment.

Mastering Apache Cassandra - Second Edition

Updated for the latest database management systems -- including MySQL 6.0, Oracle 11g, and Microsoft's SQL Server 2008 -- this introductory guide will get you up and running with SQL quickly. Whether you need to write database applications, perform administrative tasks, or generate reports, Learning SQL, Second Edition, will help you easily master all the SQL fundamentals. Each chapter presents a self-contained lesson on a key SQL concept or technique, with numerous illustrations and annotated examples. Exercises at the end of each chapter let you practice the skills you learn. With this book, you will: Move quickly through SQL basics and learn several advanced features Use SQL data statements to generate, manipulate, and retrieve data Create database objects, such as tables, indexes, and constraints, using SQL schema statements Learn how data sets interact with queries, and understand the importance of subqueries Convert and manipulate data with SQL’s built-in functions, and use conditional logic in data statements Knowledge of SQL is a must for interacting with data. With Learning SQL, you’ll quickly learn how to put the power and flexibility of this language to work.

Cassandra: The Definitive Guide

This book contains an extensive set of practical examples and an easy-to-follow approach to creating 3D objects. This book is great for anyone who already knows JavaScript and who wants to start creating 3D graphics that run in any browser. You don’t need to know anything about advanced math or WebGL; all that is needed is a general knowledge of JavaScript and HTML. The required materials and examples can be freely downloaded and all tools used in this book are open source.

Microservices with Spring Boot and Spring Cloud

Congratulations! You completed the MongoDB application within the given tight timeframe and there is a party to celebrate your application’s release into production. Although people are congratulating you at the celebration, you are feeling some uneasiness inside. To complete the project on time required making a lot of assumptions about the data, such as what terms meant and how calculations are derived. In addition, the poor documentation about the application will be of limited use to the support team, and not investigating all of the inherent rules in the data may eventually lead to poorly-performing structures in the not-so-distant future. Now, what if you had a time machine and could go back and read this book. You would learn that even NoSQL databases like MongoDB require some level of data modeling. Data modeling is the process of learning about the data, and regardless of technology, this process must be performed for a successful application. You would learn the value of conceptual, logical, and physical data modeling and how each stage increases our knowledge of the data and reduces assumptions and poor design decisions. Read this book to learn how to do data modeling for MongoDB applications, and accomplish these five objectives: Understand how data modeling contributes to the process of learning about the data, and is, therefore, a required technique, even when the resulting database is not relational. That is, NoSQL does not mean NoDataModeling! Know how NoSQL databases differ from traditional relational databases, and where MongoDB fits. Explore each MongoDB object and comprehend how each compares to their data modeling and traditional relational database counterparts, and learn the basics of adding, querying, updating, and deleting data in MongoDB. Practice a streamlined, template-driven approach to performing conceptual, logical, and physical data modeling. Recognize that data modeling does not always have to lead to traditional data models! Distinguish top-down from bottom-up development approaches and complete a top-down case study which ties all of the modeling techniques together. This book is written for anyone who is working with, or will be working with MongoDB, including business analysts, data modelers, database administrators, developers, project managers, and data scientists. There are three sections: In Section I, Getting Started, we will reveal the power of data modeling and the tight connections to data models that exist when designing any type of database (Chapter 1), compare NoSQL with traditional relational databases and where MongoDB fits (Chapter 2), explore each MongoDB object and comprehend how each compares to their data modeling and traditional relational database counterparts (Chapter 3), and explain the basics of adding, querying, updating, and deleting data in MongoDB (Chapter 4). In Section II, Levels of Granularity, we cover Conceptual Data Modeling (Chapter 5), Logical Data Modeling (Chapter 6), and Physical Data Modeling (Chapter 7).
Notice the “ing” at the end of each of these chapters. We focus on the process of building each of these models, which is where we gain essential business knowledge. In Section III, Case Study, we will explain both top down and bottom up development approaches and go through a top down case study where we start with business requirements and end with the MongoDB database. This case study will tie together all of the techniques in the previous seven chapters. Nike Senior Data Architect Ryan Smith wrote the foreword. Key points are included at the end of each chapter as a way to reinforce concepts. In addition, this book is loaded with hands-on exercises, along with their answers provided in Appendix A. Appendix B contains all of the book’s references and Appendix C contains a glossary of the terms used throughout the text.

Mastering MariaDB

This book is intended for intermediate users who want to learn how to administrate a MariaDB server or a set of servers. It is aimed at MariaDB users, and hence working knowledge of MariaDB is a prerequisite.

Getting Started with Hazelcast

If you are an administrator of a virtual environment and have used vROps before but want to gain a professional understanding by implementing complex tasks easily with it, then this book is for you.

Mastering Kubernetes

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You’ll explore the basic operations and common functions of Spark’s structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib. Spark’s scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark’s core APIs—through worked examples Dive into Spark’s low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming. Spark’s stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation.

Data Modeling for MongoDB

Every enterprise application creates data, whether it’s log messages, metrics, user activity, outgoing messages, or something else. And how to move all of this data becomes nearly as important as the data itself. If you’re an application architect, developer, or production engineer new to Apache Kafka, this practical guide shows you how to use this open source streaming platform to handle real-time data feeds. Engineers from Confluent and LinkedIn who are responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications with this platform. Through detailed examples, you’ll learn Kafka’s design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the controller, and the storage layer. Understand publish-subscribe messaging and how it fits in the big data ecosystem. Explore Kafka producers and consumers for writing and reading messages Understand Kafka patterns and use-case requirements to ensure reliable data delivery Get best practices for building data pipelines and applications with Kafka Manage Kafka in production, and learn to perform monitoring, tuning, and maintenance tasks Learn the most critical metrics among Kafka’s operational measurements Explore how Kafka’s stream delivery capabilities make it a perfect source for stream processing systems.

Mastering Akka

Summary MongoDB in Action, Second Edition is a completely revised and updated version. It introduces MongoDB 3.0 and the document-oriented database model. This perfectly paced book gives you both the big picture you’ll need as a developer and enough low-level detail to satisfy system engineers. Purchase of the print book includes a free eBook in PDF, Kindle, and ePUB formats from Manning Publications. About the Technology This document-oriented database was built for high availability, supports rich, dynamic schemas, and lets you easily distribute data across multiple servers. MongoDB 3.0 is flexible, scalable, and very fast, even with big data loads. About the Book MongoDB in Action, Second Edition is a completely revised and updated version. It introduces MongoDB 3.0 and the document-oriented database model. This perfectly paced book gives you both the big picture you’ll need as a developer and enough low-level detail to satisfy system engineers. Lots of examples will help you develop confidence in the crucial area of data modeling. You’ll also love the deep explanations of each feature, including replication, auto-sharding, and deployment. What’s Inside Indexes, queries, and standard DB operations Aggregation and text searching Map-reduce for custom aggregations and reporting Deploying for scale and high availability Updated for Mongo 3.0 About the Reader Written for developers. No previous MongoDB or NoSQL experience is assumed. About the Authors After working at MongoDB, Kyle Banker is now at a startup. Peter Bakkum is a developer with MongoDB expertise. Shaun Verch has worked on the core server team at MongoDB. A Genentech engineer, Doug Garrett is one of the winners of the MongoDB Innovation Award for Analytics.
software architect, Tim Hawkins has led search engineering at Yahoo Europe. Technical Contributor: Wouter Thielen. Technical Editor: Mihalis Tsoukalos. Table of Contents PART 1 GETTING STARTED A database for the modern web MongoDB through the JavaScript shell Writing programs using MongoDB PART 2 APPLICATION DEVELOPMENT IN MONGODB Document-oriented data Constructing queries Aggregation Updates, atomic operations, and deletes PART 3 MONGODB MASTERY Indexing and query optimization Text search WiredTiger and pluggable storage Replication Scaling your system with sharding Deployment and administration

**Hands-On Deep Learning with Apache Spark**

Deep Learning is a subset of Machine Learning where data sets with several layers of complexity can be processed. This book teaches you the different techniques using which deep learning solutions can be implemented at scale, on Apache Spark. This will help you gain experience of implementing your deep learning models in many real-world use cases.

**Kafka Streams in Action**

Gain expertise in processing and storing data by using advanced techniques with Apache SparkAbout 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 functionalitiesWho This Book Is ForIf 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 approachThis 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.

**Big Data – BigData 2019**

Data is bigger, arrives faster, and comes in a variety of formats—and it all needs to be processed at scale for analytics or machine learning. But how can you process such varied workloads efficiently? Enter Apache Spark. Updated to include Spark 3.0, this second edition shows data engineers and data scientists why structure and unification in Spark matters. Specifically, this book explains how to perform simple and complex data analytics and employ machine learning algorithms. Through step-by-step walk-throughs, code snippets, and notebooks, you’ll be able to: Learn Python, SQL, Scala, or Java high-level Structured APIs Understand Spark operations and SQL Engine Inspect, tune, and debug Spark operations with Spark configurations and Spark UI Connect to data sources: JSON, Parquet, CSV, Avro, ORC, Hive, S3, or Kafka Perform analytics on batch and streaming data using Structured Streaming Build reliable data pipelines with open source Delta Lake and Spark Develop machine learning pipelines with MLlib and productionize models using MLflow

**Mastering Spark with R**

Write efficient GIS applications using PostGIS - from data creation to data consumption About This Book Learn how you can use PostGIS for spatial data analysis and manipulation Optimize your queries and build custom functionalities for your GIS application A comprehensive guide with hands-on examples to help you master PostGIS with ease Who This Book Is For If you are a GIS developer or analyst who wants to master PostGIS to build efficient, scalable GIS applications, this book is for you. If you want to conduct advanced analysis of spatial data, this book will also help you. The book assumes that you have a working installation of PostGIS in place, and have working experience with PostgreSQL. What You Will Learn Refresh your knowledge of the PostGIS concepts and spatial databases Solve spatial problems with the use of SQL in real-world scenarios Practical walkthroughs of application development examples using Postgis, GeoServer and OpenLayers. Extract, transform and load your spatial data Expose data directly or through web services. Consume your data in both desktop and web clients In Detail PostGIS is open source extension on PostgreSQL object-relational database system that allows GIS objects to be stored and allows querying for information and location services. The aim of this book is to help you master the functionalities offered by PostGIS - from data creation, analysis and output, to ETL and live edits. The book begins with an overview of the key concepts related to spatial database systems and how it applies to Spatial RMDs. You will learn to load different formats into your Postgres instance, investigate the spatial nature of your raster data, and finally export it using built-in functionalities or 3rd party tools for
Backup or representational purposes. Through the course of this book, you will be presented with many examples on how to interact with the database using JavaScript and Node.js. Sample web-based applications interacting with backend PostGIS will also be presented throughout the book, so you can get comfortable with the modern ways of consuming and modifying your spatial data.

**Mastering Cloud Computing**

This updated and expanded second edition of the Mastering Apache Cassandra · Second Edition provides a user-friendly introduction to the subject. Taking a clear structural framework, it guides the reader through the subject’s core elements. A flowing writing style combines with the use of illustrations and diagrams throughout the text to ensure the reader understands even the most complex of concepts. This succinct and enlightening overview is a required reading for all those interested in the subject. We hope you find this book useful in shaping your future career & Business.

**MongoDB in Action**

This volume constitutes the proceedings of the 8th International Congress on BIGDATA 2019, held as Part of SCF 2019 in San Diego, CA, USA in June 2019. The 9 full papers presented in this volume were carefully reviewed and selected from 14 submissions. They cover topics such as: Big Data Models and Algorithms; Big Data Architectures; Big Data Management; Big Data Protection, Integrity and Privacy; Security Applications of Big Data; Big Data Search and Mining; Big Data for Enterprise, Government and Society.

**Kafka: The Definitive Guide**

Every enterprise application creates data, including log messages, metrics, user activity, and outgoing messages. Learning how to move these items is almost as important as the data itself. If you’re an application architect, developer, or production engineer new to Apache Pulsar, this practical guide shows you how to use this open source event streaming platform to handle real-time data feeds. Jowanza Joseph, staff software engineer at Finicity, explains how to deploy production Pulsar clusters, write reliable event streaming applications, and build scalable real-time data pipelines with this platform. Through detailed examples, you'll learn Pulsar's design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the load manager, and the storage layer. This book helps you understand how event streaming fits in the big data ecosystem Explore Pulsar producers, consumers, and readers for writing and reading events Build scalable data pipelines by connecting Pulsar with external systems Simplify event-streaming application building with Pulsar Functions Manage Pulsar to perform monitoring, tuning, and maintenance tasks Use Pulsar's operational measurements to secure a production cluster Process event streams using Flink and query event streams using Presto

**Mastering JavaFX 8 Controls**

Take your C++ coding to the next level by leveraging the latest features and advanced techniques to building high performing, reliable applications. About This Book Get acquainted with the latest features in C++ 17 Take advantage of the myriad of features and possibilities that C++ offers to build real-world applications Write clear and expressive code in C++, and get insights into how to keep your code error-free Who This Book Is For This book is for experienced C++ developers. If you are a novice C++ developer, then it's highly recommended that you get a solid understanding of the C++ language before reading this book What You Will Learn Write modular C++ applications in terms of the existing and newly introduced features Identify code-smells, clean up, and re-factor legacy C++ applications Leverage the possibilities provided by Cucumber and Google Test/Mock to automate test cases Test frameworks with C++ Get acquainted with the new C++17 features Develop GUI applications in C++ Build portable cross-platform applications using standard C++ features In Detail C++ has come a long way and has now been adopted in several contexts. Its key strengths are its software infrastructure and resource-constrained applications. The C++ 17 release will change the way developers write code, and this book will help you master your developing skills with C++. With real-world, practical examples explaining each concept, the book will begin by introducing you to the latest features in C++ 17. It encourages clean code practices in C++ in general, and demonstrates the GUI app-development options in C++. You'll get tips on avoiding memory leaks using smart-pointers, Next, you’ll see how multi-threaded programming can help you achieve concurrency in your applications. Moving on, you'll get an in-depth understanding of the C++ Standard Template Library. We show you the concepts of implementing TDD and BDD in your C++ programs, and explore template-based generic programming, giving you the expertise to build powerful applications. Finally, we’ll round up with debugging techniques and best practices. By the end of the book, you'll have an in-depth understanding of the language and its various facets. Style and approach This straightforward guide will help you level up your skills in C++ programming, be it for enterprise software or for low-latency applications like games. Filled with real-world, practical examples, this book will take you gradually up the steep learning curve that is C++.

**Spark: The Definitive Guide**

A hands-on guide to leveraging NoSQL databases NoSQL databases are an efficient and powerful tool for storing and
manipulating vast quantities of data. Most NoSQL databases scale well as data grows. In addition, they are often malleable and flexible enough to accommodate semi-structured and sparse data sets. This comprehensive hands-on guide presents fundamental concepts and practical solutions for getting you ready to use NoSQL databases. Expert author Shashank Tiwari begins with a helpful introduction on the subject of NoSQL, explains its characteristics and typical uses, and looks at where it fits in the application stack. Unique insights help you choose which NoSQL solutions are best for solving your specific data storage needs. Professional NoSQL: Demystifies the concepts that relate to NoSQL databases, including column-family oriented stores, key/value databases, and document databases. Delves into installing and configuring a number of NoSQL products and the Hadoop family of products. Explains ways of storing, accessing, and querying data in NoSQL databases through examples that use MongoDB, HBase, Cassandra, Redis, CouchDB, Google App Engine Datatstore and more. Looks at architecture and internals. Provides guidelines for optimal usage, performance tuning, and scalable configurations. Presents a number of tools and utilities relating to NoSQL, distributed platforms, and scalable processing, including Hive, Pig, RRDtool, Nagios, and more.

**Big Data SMACK**

“Eric and Russell were early adopters of Cassandra at SimpleReach. In Practical Cassandra, you benefit from their experience in the trenches administering Cassandra, developing against it, and building one of the first CQL drivers. If you are deploying Cassandra soon, or you inherited a Cassandra cluster to tend, spend some time with the deployment, performance tuning, and maintenance chapters... If you are new to Cassandra, I highly recommend the chapters on data modeling and CQL.” –From the Foreword by Jonathon Ellis, Apache Cassandra Chair

**Build and Deploy Massively Scalable, Super-fast Data Management Applications with Apache Cassandra**

Practical Cassandra is the first hands-on developer’s guide to building Cassandra systems and applications that deliver breakthrough speed, scalability, reliability, and performance. Fully up to date, it reflects the latest versions of Cassandra—including Cassandra Query Language (CQL), which dramatically lowers the learning curve for Cassandra developers. Pioneering Cassandra developers and Datastax MVPs Russell Bradberry and Eric Lubow walk you through every step of building a real production application that can store enormous amounts of structured, semi-structured, and unstructured data. Drawing on their exceptional expertise, Bradberry and Lubow share practical insights into issues ranging from querying to deployment, management, maintenance, monitoring, and troubleshooting. The authors cover key issues, from architecture to migration, and guide you through crucial decisions about configuration and data modeling. They provide tested sample code, detailed explanations of how Cassandra works “under the covers,” and new case studies from three cutting-edge users: Ooyala, Hailo, and eBay. Coverage includes Understanding Cassandra’s approach, architecture, key concepts, and primary use cases- and why it’s so blazingly fast Getting Cassandra up and running on single nodes and large clusters Applying the new design patterns, philosophies, and features that make Cassandra such a powerful data store Leveraging CQL to simplify your transition from SQL-based RDBMSes Deploying and provisioning through the cloud or on bare-metal hardware Choosing the right configuration options for each type of workload Tweaking Cassandra to get maximum performance from your hardware, OS, and JVM

**Cassandra: The Definitive Guide**

By introducing in-memory persistent storage, Apache Spark eliminates the need to store intermediate data in filesystems, thereby increasing processing speed by up to 100 times. This book will focus on how to analyze large and complex sets of data. Starting with installing and configuring Apache Spark with various cluster managers, you will cover setting up development environments. You will then cover various recipes to perform interactive queries using Spark SQL and real-time streaming with various sources such as Twitter Stream and Apache Kafka. You will then focus on machine learning, including supervised learning, unsupervised learning, and recommendation engine algorithms. After mastering graph processing using GraphX, you will cover various recipes for cluster optimization and troubleshooting.

**Practical Cassandra**

Follow this handbook to build, configure, tune, and secure Apache Cassandra databases. Start with the installation of Cassandra and move on to the creation of a single instance, and then a cluster of Cassandra databases. Cassandra is increasingly a key player in many big data environments, and this book shows you how to use Cassandra with Apache Spark, a popular big data processing framework. Also covered are day-to-day topics of importance such as the backup and recovery of Cassandra databases, using the right compression and compaction strategies, and loading and unloading data. Expert Apache Cassandra Administration provides numerous step-by-step examples starting with the basics of a Cassandra database, and going all the way through backup and recovery, performance optimization, and monitoring and securing the data. The book serves as an authoritative and comprehensive guide to the building and management of simple to complex Cassandra databases. The book: Takes you through building a Cassandra database from installation of the software and creation of a single database, through to complex clusters and data centers Provides numerous examples of actual commands in a real-life Cassandra environment that show how to confidently configure, manage, troubleshoot, and tune Cassandra databases Shows how to use the Cassandra configuration properties to build a highly stable, available, and secure Cassandra database that always operates at peak efficiency What You'll Learn Install the Cassandra software and create your first database Understand the Cassandra data model, and the internal architecture of a Cassandra database Create your own Cassandra cluster, step-by-step Run a Cassandra cluster on Docker Work with Apache Spark by connecting to a Cassandra database Deploy Cassandra clusters in your data center, or on Amazon EC2
instances Back up and restore mission-critical Cassandra databases Monitor, troubleshoot, and tune production Cassandra databases, and cut your spending on resources such as memory, servers, and storage Who This Book Is For Database administrators, developers, and architects who are looking for an authoritative and comprehensive single volume for all their Cassandra administration needs. Also for administrators who are tasked with setting up and maintaining highly reliable and high-performance Cassandra databases. An excellent choice for big data administrators, database administrators, architects, and developers who use Cassandra as their key data store, to support high volume online transactions, or as a decentralized, elastic data store.

Learning SQL

This book is an easy-to-follow, hands-on introduction that guides you through this innovative new technology. It covers everything from data grids to the simple-to-use distributed data storage collections. Queuing and topic messaging capabilities, as well as locking and transaction support to guard against concurrency race-conditions, are some of the topics that we will cover. We will then move on to distributed task execution, in-place data manipulations and big data analytical processing using MapReduce. At the end of all this, you will be armed with everything you need to bring amazing power and data scalability to your applications, as well as making them truly global and ready for a worldwide audience.

Mastering Data Analysis with R

Master the art of creating scalable, concurrent, and reactive applications using Akka About This Book This book will help you cure anemic models with domain-driven design We cover major Akka programming concepts such as concurrency, scalability, and reactivity You will learn concepts like Event Sourcing and CQRS via Akka Persistence, Akka Streams, Akka Http as well as Akka Clustering Who This Book Is For If you want to use the Lightbend platform to create highly performant reactive applications, then this book is for you. If you are a Scala developer looking for techniques to use all features of the new Akka release and want to incorporate these solutions in your current or new projects, then this book is for you. Expert Java developers who want to build scalable, concurrent, and reactive application will find this book helpful. What You Will Learn Use Akka actors to enable parallel execution Build out domain-driven design based components like entities and aggregates Respond to command requests on that aggregate root that affect the internal state Leverage Akka Persistence, protobuf and Cassandra to save the persistent state of your entities Build out complex processing graphs with the Graph Builder DSL Understand the dynamic push/pull nature of backpressure handling within Akka Streams Route HTTP requests to an actor and return a response Deploy actor instances across a set of nodes via ConductR for high availability In Detail For a programmer, writing multi-threaded applications is critical as it is important to break large tasks into smaller ones and run them simultaneously. Akka is a distributed computing toolkit that uses the abstraction of the Actor model, enabling developers to build correct, concurrent, and distributed applications using Java and Scala with ease. The book begins with a quick introduction that simplifies concurrent programming with actors. We then proceed to master all aspects of domain-driven design. We’ll teach you how to scale out with Akka Remoting/Clustering. Finally, we introduce ConductR as a means to deploy to and manage microservices across a cluster. Style and approach This comprehensive, fast-paced guide is packed with several real-world use cases that will help you understand concepts, issues, and resolutions while using Akka to create highly performant, scalable, and concurrency-proof reactive applications.

Mastering PostGIS

Build, manage, and configure high-performing, reliable NoSQL database for your applications with Cassandra Key Features Write programs more efficiently using Cassandra's features with the help of examples Configure Cassandra and fine-tune its parameters depending on your needs Integrate Cassandra database with Apache Spark and build strong data analytics pipeline Book Description With ever-increasing rates of data creation, the demand for storing data fast and reliably becomes a need. Apache Cassandra is the perfect choice for building fault-tolerant and scalable databases. Mastering Apache Cassandra 3.x teaches you how to build and architect your clusters, configure and work with your nodes, and program in a high-throughput environment, helping you understand the power of Cassandra as per the new features. Once you've covered a brief recap of the basics, you'll move on to deploying and monitoring a production setup and optimizing and integrating it with other software. You'll work with the advanced features of CQL and the new storage engine in order to understand how they function on the server-side. You'll explore the integration and interaction of Cassandra components, followed by discovering features such as token allocation algorithm, CQL3, vnodes, lightweight transactions, and data modelling in detail. Last but not least you will get to grips with Apache Spark. By the end of this book, you'll be able to analyse big data, and build and manage high-performance databases for your application. What you will learn Write programs more efficiently using Cassandra's features more efficiently Exploit the given infrastructure, improve performance, and tweak the Java Virtual Machine (JVM) Use CQL3 in your application in order to simplify working with Cassandra Configure Cassandra and fine-tune its parameters depending on your needs Set up a cluster and learn how to scale it Monitor a Cassandra cluster in different ways Use Apache Spark and other big data processing tools Who this book is for Mastering Apache Cassandra 3.x is for you if you are a big data administrator, database administrator, architect, or developer who wants to build a high-performing, scalable, and fault-tolerant database. Prior knowledge of core concepts of databases is required.

Spark Cookbook
Learn how to integrate full-stack open source big data architecture and 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 data sets 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. This 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, MLlib, Streaming, GraphX) The container: Mesos, Docker The view: Akka The storage: Cassandra The message broker: Kafka What You Will Learn: Make big data architecture without using complex Greek letter architectures Build a cheap but effective cluster infrastructure Make queries, reports, and graphs that business demands Manage and exploit unstructured and No-SQL data sources Use tools to monitor the performance of your architecture Integrate all technologies and decide which ones replace and which ones reinforce Who This Book Is For: Developers, data architects, and data scientists looking to integrate the most successful big data open stack architecture and to choose the correct technology in every layer

Stream Processing with Apache Flink

Go beyond simply learning Kubernetes fundamentals and its deployment, and explore more advanced concepts, including serverless computing and service meshes with the latest updates Key Features Master Kubernetes architecture and design to build and deploy secure distributed applications Learn advanced concepts like autoscaling, cluster federation, serverless computing, and service mesh integration for observability Explore Kubernetes 1.18 features and its rich ecosystem of tools like Kubect, Knative, and Helm Book Description The third edition of Mastering Kubernetes is updated with the latest tools and code enabling you to learn Kubernetes 1.18’s latest features. This book primarily concentrates on diving deeply into complex concepts and Kubernetes best practices to help you master the skills of designing and deploying large clusters on various cloud platforms. The book trains you to run complex stateful microservices on Kubernetes including advanced features such as horizontal pod autoscaling, rolling updates, resource quotas, and persistent storage backend. With the two new chapters, you will gain expertise in serverless computing and utilizing service meshes. As you proceed through the chapters, you will explore different options for network configuration and learn to set up, operate, and troubleshoot Kubernetes networking plugins through real-world use cases. Furthermore, you will understand the mechanisms of custom resource development and its utilization in automation and maintenance workflows. By the end of this Kubernetes book, you will graduate from an intermediate to advanced Kubernetes professional. What will you learn Master the fundamentals of Kubernetes architecture and design Build and run stateful applications and complex microservices on Kubernetes Use tools like Kubect, secrets, and Helm to manage resources and storage Master Kubernetes Networking with load balancing options like Ingress Achieve high-availability Kubernetes clusters Improve Kubernetes observability with tools like Prometheus, Grafana, and Jaeger Extend Kubernetes working with Kubernetes API, plugins, and webhooks Who this book is for If you are a system administrator or a cloud developer with working knowledge of Kubernetes and are keen to master its advanced features, along with learning everything from building microservices to utilizing service meshes, Mastering Kubernetes is for you. Basic familiarity with networking concepts will be helpful.

Mastering Apache Pulsar

What could you do with data if scalability wasn’t a problem? With this hands-on guide, you’ll learn how Apache Cassandra handles hundreds of terabytes of data while remaining highly available across multiple data centers -- capabilities that have attracted Facebook, Twitter, and other data-intensive companies. Cassandra: The Definitive Guide provides the technical details and practical examples you need to assess this database management system and put it to work in a production environment. Author Eben Hewitt demonstrates the advantages of Cassandra's nonrelational design, and pays special attention to data modeling. If you're a developer, DBA, application architect, or manager looking to solve a database scaling issue or future-proof your application, this guide shows you how to harness Cassandra's speed and flexibility. Understand the tenets of Cassandra's column-oriented structure Learn how to write, update, and read Cassandra data Discover how to add or remove nodes from the cluster as your application requires Examine a working application that translates from a relational model to Cassandra's data model Use examples for writing clients in Java, Python, and C# Use the JMX interface to monitor a cluster's usage, memory patterns, and more Tune memory settings, data storage, and caching for better performance

Mastering Apache Cassandra 3.x - Third Edition

Learn Azure in a Month of Lunches, Second Edition, is a tutorial on writing, deploying, and running applications in Azure. In it, you’ll work through 21 short lessons that give you real-world experience. Each lesson includes a hands-on lab so you can try out and lock in your new skills. Summary You can be incredibly productive with Azure without mastering every feature, function, and service. Learn Azure in a Month of Lunches, Second Edition gets you up and running quickly, teaching you the most important concepts and tasks in 21 practical bite-sized lessons. As you explore the examples, exercises, and labs, you'll pick up valuable skills immediately and take your first steps to Azure mastery! This fully revised new edition covers core changes to the Azure UI, new Azure features, Azure containers, and the upgraded Azure Kubernetes Service. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning
Julia is a well-constructed programming language with fast execution speed, eliminating the classic problem of performing analysis in one language and translating it for performance into a second. This book will help you develop and enhance your programming skills in Julia to solve real-world automation challenges. This book starts off with a refresher on installing and running Julia on different platforms. Next, you will compare the different ways of working with Julia and explore Julia's key features in-depth by looking at design and build. You will see how data works using simple statistics and analytics, and discover Julia's speed, its real strength, which makes it particularly useful in highly intensive real-time data processing, and creating custom Spark extensions.
computing tasks and analyze how Julia can cooperate with external processes in order to enhance graphics and data visualization. Finally, you will look into meta-programming and learn how it adds great power to the language and establish networking and distributed computing with Julia.

**Mastering Apache Spark**

Design and Deploy High-Performance JavaFX Controls Deliver state-of-the-art applications with visually stunning UIs. Mastering JavaFX 8 Controls provides clear instructions, detailed examples, and ready-to-use code samples. Find out how to work with the latest JavaFX APIs, configure UI components, automatically generate FXML, build cutting-edge controls, and effectively apply CSS styling. Troubleshooting, tuning, and deployment are also covered in this Oracle Press guide. Understand the properties of JavaFX 8 controls and APIs Create custom controls, transformations, and layouts Work from JavaFX Scene Graph and Scene Builder Visualize data models using advanced control types Use ListView, TableView, and TreeView APIs Handle audio and video content using JavaFX media classes Maximize separation between UI and application logic using FXML Style controls and applications using CSS Extend functionality of Swing and SWT applications with JavaFX Code examples in the book are available for download.

**Mastering Apache Cassandra**

Gain sharp insights into your data and solve real-world data science problems with R—from data munging to modeling and visualization About This Book Handle your data with precision and care for optimal business intelligence Restructure and transform your data to inform decision-making Packed with practical advice and tips to help you get to grips with data mining Who This Book Is For If you are a data scientist or R developer who wants to explore and optimize your use of R's advanced features and tools, this is the book for you. A basic knowledge of R is required, along with an understanding of database logic. What You Will Learn Connect to and load data from R's range of powerful databases Successfully fetch and parse structured and unstructured data Transform and restructure your data with efficient R packages Define and build complex statistical models with glm Develop and train machine learning algorithms Visualize social networks and graph data Deploy supervised and unsupervised classification algorithms Discover how to visualize spatial data with R In Detail R is an essential language for sharp and successful data analysis. Its numerous features and ease of use make it a powerful way of mining, managing, and interpreting large sets of data. In a world where understanding big data has become key, mastering R will be able to deal with your data effectively and efficiently. This book will give you the guidance you need to build and develop your knowledge and expertise. Bridging the gap between theory and practice, this book will help you to understand and use data for a competitive advantage. Beginning with taking you through essential data mining and management tasks such as munging, fetching, cleaning, and restructuring, the book then explores different model designs and the core components of effective analysis. You will then discover how to optimize your use of machine learning algorithms for classification and recommendation systems beside the traditional and more recent statistical methods. Style and approach Covering the essential tasks and skills within data science, Mastering Data Analysis provides you with solutions to the challenges of data science. Each section gives you a theoretical overview before demonstrating how to put the theory to work with real-world use cases and hands-on examples.

**Expert Apache Cassandra Administration**

Summary Kafka Streams in Action teaches you everything you need to know to implement stream processing on data flowing into your Kafka platform, allowing you to focus on getting more from your data without sacrificing time or effort. Foreword by Neha Narkhede, Co-creator of Apache Kafka Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Not all stream-based applications require a dedicated processing cluster. The lightweight Kafka Streams library provides exactly the power and simplicity you need for message handling in microservices and real-time event processing. With the Kafka Streams API, you filter and transform data streams with just Kafka and your application. About the Book Kafka Streams in Action teaches you to implement stream processing within the Kafka platform. In this easy-to-follow book, you'll explore real-world examples to collect, transform, and aggregate data, work with multiple processors, and handle real-time events. You'll even dive into streaming SQL with KSQL! Practical to the very end, it finishes with testing and operational aspects, such as monitoring and debugging. What's inside Using the KStreams API Filtering, transforming, and splitting data Working with the Processor API Integrating with external systems About the Reader Assumes some experience with distributed systems. No knowledge of Kafka or streaming applications required. About the Author Bill Bejeck is a Kafka Streams contributor and Confluent engineer with over 15 years of software development experience. Table of Contents PART 1 - GETTING STARTED WITH KAFKA STREAMS Welcome to Kafka Streams Kafka quicklyPART 2 - KAFKA STREAMS DEVELOPMENT Developing Kafka Streams Streams and state The KTable API The Processor API PART 3 - ADMINISTERING KAFKA STREAMS Monitoring and performance Testing a Kafka Streams application PART 4 - ADVANCED CONCEPTS WITH KAFKA STREAMS Advanced applications with Kafka Streams APPENDIXES Appendix A - Additional configuration information Appendix B - Exactly once semantics

Copyright code : c5e34f1302021f4ed71c49f735d0add