

---

# Access Free Pdf Edition Second Cassandra Apache Mastering

---

This is likewise one of the factors by obtaining the soft documents of this **Pdf Edition Second Cassandra Apache Mastering** by online. You might not require more era to spend to go to the books opening as without difficulty as search for them. In some cases, you likewise reach not discover the declaration Pdf Edition Second Cassandra Apache Mastering that you are looking for. It will totally squander the time.

However below, later you visit this web page, it will be fittingly unconditionally easy to get as well as download guide Pdf Edition Second Cassandra Apache Mastering

It will not understand many get older as we tell before. You can pull off it while conduct yourself something else at house and even in your workplace. thus easy! So, are you question? Just exercise just what we give below as capably as evaluation **Pdf Edition Second Cassandra Apache Mastering** what you afterward to read!

---

## KEY=SECOND - ROMAN LEXI

---

**Mastering Apache Cassandra - Second Edition** *Packt Publishing Ltd* **The book is aimed at intermediate developers with an understanding of core database concepts who want to become a master at implementing Cassandra for their application. Mastering Apache Cassandra 3.x An expert guide to improving database scalability and availability without compromising performance, 3rd Edition** *Packt Publishing Ltd* **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: The Definitive Guide Big Data Processing Made Simple "O'Reilly Media, Inc."**  
**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**  
**Mastering Apache Cassandra 3.x - Third Edition**  
**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.

*Learning Spark O'Reilly Media* 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

*Big Data SMACK A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka Apress* 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, MLib, 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 Learn *Azure in a Month of Lunches, Second Edition* *Manning Publications* *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 Publications. About the technology Microsoft Azure is vast and powerful, offering virtual servers, application templates, and prebuilt services for everything from data storage to AI. To navigate it all, you need a trustworthy guide. In this book, Microsoft engineer and Azure trainer Iain Foulds focuses on core skills for creating cloud-based applications. About the book *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. What's inside Understanding Azure beyond point-and-click Securing applications and data Automating your environment Azure services for machine learning, containers, and more About the reader This book is for readers who can write and deploy simple web or client/server applications. About the author

Iain Foulds is an engineer and senior content developer with Microsoft. **Table of Contents PART 1 - AZURE CORE SERVICES 1 Before you begin 2 Creating a virtual machine 3 Azure Web Apps 4 Introduction to Azure Storage 5 Azure Networking basics PART 2 - HIGH AVAILABILITY AND SCALE 6 Azure Resource Manager 7 High availability and redundancy 8 Load-balancing applications 9 Applications that scale 10 Global databases with Cosmos DB 11 Managing network traffic and routing 12 Monitoring and troubleshooting PART 3 - SECURE BY DEFAULT 13 Backup, recovery, and replication 14 Data encryption 15 Securing information with Azure Key Vault 16 Azure Security Center and updates PART 4 - THE COOL STUFF 17 Machine learning and artificial intelligence 18 Azure Automation 19 Azure containers 20 Azure and the Internet of Things 21 Serverless computing** *Mastering Node.js Packt Publishing Ltd* 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. **Kafka: The Definitive Guide Real-Time Data and Stream Processing at Scale "O'Reilly Media, Inc."** Learn how to take full advantage of Apache Kafka, the distributed, publish-subscribe queue for handling real-time data feeds. With this comprehensive book, you will understand how Kafka works and how it is designed. Authors Neha Narkhede, Gwen Shapira, and Todd Palino show you how to deploy production Kafka clusters; secure, tune, and monitor them; write rock-solid applications that use Kafka; and build scalable stream-processing applications. Learn how Kafka compares to other queues, and where it fits in the big data ecosystem. Dive into Kafka's internal design Pick up best practices for developing applications that use Kafka. Understand the best way to deploy Kafka in production monitoring, tuning, and maintenance tasks. Learn how to secure a Kafka cluster. *Learning Apache Cassandra Packt Publishing Ltd* If you're an application developer familiar with SQL databases such as MySQL or Postgres, and you want to explore distributed databases such as Cassandra, this is the perfect guide for you. Even if you've never worked with a distributed database before, Cassandra's intuitive programming interface coupled with the step-by-step examples in this book will have you building highly scalable persistence layers for your applications in no time. *Mastering Kubernetes Packt Publishing Ltd* Master the art of container management utilizing the power of Kubernetes. **About This Book** This practical guide demystifies Kubernetes and ensures that your clusters are always available, scalable, and up to date Discover new features such as autoscaling, rolling updates, resource quotas, and cluster size Master the skills of designing and deploying large clusters on various cloud platforms **Who This Book Is For** The book is for system administrators and developers who have intermediate level of knowledge with Kubernetes and are now waiting to master its advanced

features. You should also have basic networking knowledge. This advanced-level book provides a pathway to master Kubernetes. What You Will Learn Architect a robust Kubernetes cluster for long-time operation Discover the advantages of running Kubernetes on GCE, AWS, Azure, and bare metal See the identity model of Kubernetes and options for cluster federation Monitor and troubleshoot Kubernetes clusters and run a highly available Kubernetes Create and configure custom Kubernetes resources and use third-party resources in your automation workflows Discover the art of running complex stateful applications in your container environment Deliver applications as standard packages In Detail Kubernetes is an open source system to automate the deployment, scaling, and management of containerized applications. If you are running more than just a few containers or want automated management of your containers, you need Kubernetes. This book mainly focuses on the advanced management of Kubernetes clusters. It covers problems that arise when you start using container orchestration in production. We start by giving you an overview of the guiding principles in Kubernetes design and show you the best practises in the fields of security, high availability, and cluster federation. You will discover how to run complex stateful microservices on Kubernetes including advanced features as horizontal pod autoscaling, rolling updates, resource quotas, and persistent storage back ends. Using real-world use cases, we explain the options for network configuration and provides guidelines on how to set up, operate, and troubleshoot various Kubernetes networking plugins. Finally, we cover custom resource development and utilization in automation and maintenance workflows. By the end of this book, you'll know everything you need to know to go from intermediate to advanced level. Style and approach Delving into the design of the Kubernetes platform, the reader will be exposed to the advanced features and best practices of Kubernetes. This book will be an advanced level book which will provide a pathway to master Kubernetes Cassandra: The Definitive Guide Distributed Data at Web Scale "O'Reilly Media, Inc." 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 Professional NoSQL *John Wiley & Sons* 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 Datastore 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. Learning SQL Master SQL Fundamentals *O'Reilly Media* 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. Mastering ServiceNow *Packt Publishing Ltd* Unleash the full potential of ServiceNow from foundations  
to advanced functions, with this hands-on expert guide fully revised for the Helsinki version About This Book Give your  
ServiceNow Helsinki implementation a powerful kick-start by understanding the deep capabilities of the platform,

Learn by doing with an extended, comprehensive example, creating a feature-rich, secure and automated application from the ground up Interact with your whole organization by integrating with REST web services and build a custom Service Portal interface Who This Book Is For This book is aimed at advanced ServiceNow System Administrators and developers who would like to gain greater control of ServiceNow and its architecture. The book expects you to be new to ServiceNow, but have a good grounding in internet and computing technologies, like HTML, JSON, REST and database systems. Readers should be especially familiar with JavaScript, and be keen to extend and alter the platform. With this book, they will be able to develop a new application for their company. What You Will Learn Build custom scoped applications that access the full ServiceNow API Build a modern, responsive self-service interface with Service Portal Design feature-rich, responsive, automated workflow systems Design powerful data-driven applications Control information flow and apply business logic with Business Rules Write efficient and effective client-side JavaScript Learn how to authenticate and secure Web Services Integrate and exchange data with people and systems Create and secure your systems with proper access control In Detail ServiceNow is a SaaS application that provides workflow form-based applications. It is an ideal platform for creating enterprise-level applications giving requesters and fulfillers improved visibility and access to a process. ServiceNow-based applications often replace email by providing a better way to get work done. The book steps through the main aspects of the ServiceNow platform, from the ground up. It starts by exploring the core architecture of ServiceNow, including building the right data structure. To add business logic and control data, and interactivity to user interaction, you will be shown how to code on both server and the client. You will then learn more about the power of tasks, events and notifications. The book will then focus on using web services and other mechanisms to integrate ServiceNow with other systems. Furthermore, you will learn how to secure applications and data, and understand how ServiceNow performs logging and error reporting. You will then be shown how to package your applications and changes, so they can be installed elsewhere and ways to maintain them easily. If you wish to create an alternative simple interface, then explore ways to make ServiceNow beautiful using Service Portal. By the end of the book, you will know the fundamentals of the ServiceNow platform, helping you be a better ServiceNow System Administrator or developer. Style and approach Explore how to implement business logic and automated workflows and write effective code by flexible choices for client-side scripting Mastering Cloud Computing Foundations and Applications Programming *Newnes* 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 Introducing MLOps *O'Reilly Media* More than half of the analytics and machine learning (ML) models created by organizations today never make it into production. Some of the challenges and barriers to operationalization are technical, but others are organizational. Either way, the bottom line is that models not in production can't provide business impact. This book introduces the key concepts of MLOps to help data scientists and application engineers not only operationalize ML models to drive real business change but also maintain and improve those models over time. Through lessons based on numerous MLOps applications around the world, nine experts in machine learning provide insights into the five steps of the model life cycle--Build, Preproduction, Deployment, Monitoring, and Governance--uncovering how robust MLOps processes can be infused throughout. This book helps you: Fulfill data science value by reducing friction throughout ML pipelines and workflows Refine ML models through retraining, periodic tuning, and complete remodeling to ensure long-term accuracy Design the MLOps life cycle to minimize organizational risks with models that are unbiased, fair, and explainable Operationalize ML models for pipeline deployment and for external business systems that are more complex and less standardized Programming Hive "*O'Reilly Media, Inc.*" Describes the features and functions of Apache Hive, the data infrastructure for Hadoop. Practical Data Science A Guide to Building the Technology Stack for Turning Data Lakes into Business Assets *Apress* Learn how to build a data science technology stack and perform good data science with repeatable methods. You will learn how to turn data lakes into business assets. The data science technology stack demonstrated in Practical Data Science is built from components in general use in the industry. Data scientist Andreas Vermeulen demonstrates in detail how to build and provision a technology stack to yield repeatable results. He shows you how to apply practical methods to extract actionable business knowledge from data lakes consisting of data from a polyglot of data types and dimensions. What You'll Learn Become fluent in the essential concepts and terminology of data science and data engineering Build and use a technology stack that meets industry criteria Master the methods for retrieving actionable business knowledge Coordinate the handling of polyglot data types in a data lake for repeatable results Who This Book Is For Data scientists and data engineers who are required to convert data from a data lake into actionable knowledge for their business, and students who aspire to be data scientists and data engineers Ant: The Definitive Guide

**Complete Build Management for Java "O'Reilly Media, Inc."** Soon after its launch, Ant succeeded in taking the Java world by storm, becoming the most widely used tool for building applications in Java environments. Like most popular technologies, Ant quickly went through a series of early revision cycles. With each new version, more functionality was added, and more complexity was introduced. Ant evolved from a simple-to-learn build tool into a full-fledged testing and deployment environment. **Ant: The Definitive Guide** has been reworked, revised and expanded upon to reflect this evolution. It documents the new ways that Ant is being applied, as well as the array of optional tasks that Ant supports. In fact, this new second edition covers everything about this extraordinary build management tool from downloading and installing, to using Ant to test code. Here are just a few of the features you'll find detailed in this comprehensive, must-have guide: Developing conditional builds, and handling error conditions Automatically retrieving source code from version control systems Using Ant with XML files Using Ant with JavaServer Pages to build Web applications Using Ant with Enterprise JavaBeans to build enterprise applications Far exceeding its predecessor in terms of information and detail, **Ant: The Definitive Guide, 2nd Edition** is a must-have for Java developers unfamiliar with the latest advancements in Ant technology. With this book at your side, you'll soon be up to speed on the premiere tool for cross-platform development. **Author Steve Holzner** is an award-winning author who has been writing about Java topics since the language first appeared; his books have sold more than 1.5 million copies worldwide. **Mastering Microservices with Java 9** Build domain-driven microservice-based applications with Spring, Spring Cloud, and Angular *Packt Publishing Ltd* Master the art of implementing scalable microservices in your production environment with ease **About This Book** Use domain-driven design to build microservices Use Spring Cloud to use Service Discovery and Registration Use Kafka, Avro and Spring Streams for implementing event based microservices **Who This Book Is For** This book is for Java developers who are familiar with the microservices architecture and now wants to take a deeper dive into effectively implementing microservices at an enterprise level. A reasonable knowledge level and understanding of core microservice elements and applications is expected. **What You Will Learn** Use domain-driven design to design and implement microservices Secure microservices using Spring Security Learn to develop REST service development Deploy and test microservices Troubleshoot and debug the issues faced during development Learning best practices and common principals about microservices **In Detail** Microservices are the next big thing in designing scalable, easy-to-maintain applications. It not only makes app development easier, but also offers great flexibility to utilize various resources optimally. If you want to build an enterprise-ready implementation of the microservices architecture, then this is the book for you! Starting off by understanding the core concepts and framework, you will then focus on the high-level design of large software projects. You will gradually move on to

setting up the development environment and configuring it before implementing continuous integration to deploy your microservice architecture. Using Spring security, you will secure microservices and test them effectively using REST Java clients and other tools like RxJava 2.0. We'll show you the best patterns, practices and common principals of microservice design and you'll learn to troubleshoot and debug the issues faced during development. We'll show you how to design and implement reactive microservices. Finally, we'll show you how to migrate a monolithic application to microservices based application. By the end of the book, you will know how to build smaller, lighter, and faster services that can be implemented easily in a production environment. Style and approach This book starts from the basics, including environment setup and provides easy-to-follow steps to implement the sample project using microservices.

**Active Directory For Dummies** *John Wiley & Sons* Whether you're new to Active Directory (AD) or a savvy system administrator looking to brush up on your skills, **Active Directory for Dummies, 2nd Edition** will steer you in the right direction! Since its original release, Microsoft's implementation of the lightweight directory access protocol (LDAP) for the Windows Server line of networking software has become one of the most popular directory service products in the world. If you are involved with the design and support of Microsoft directory services and/or solutions, you really need this book! You'll understand the basics of AD and utilize its structures to simplify your life and secure your digital environment. You'll discover how to exert fine-grained control over groups, assets, security, permissions, and policies on a Windows network and efficiently configure, manage, and update the network. You'll find new and updated material on security improvements, significant user interface changes, and updates to the AD scripting engine, password policies, accidental object deletion protection, and more. You will learn how to: Navigate the functions and structures of AD Understand business and technical requirements and determine goals Become familiar with physical components like site links, network services, and site topology Manage and monitor new features, AD replication, and schema management Maintain AD databases Avoid common AD mistakes that can undermine network security Complete with lists of the ten most important points about AD, and ten cool Web resources, and ten troubleshooting tips, **Active Directory For Dummies, 2nd Edition** is your one-stop guide to setting up, working with, and making the most of Active Directory. Note: CD-ROM/DVD and other supplementary materials are not included as part of eBook file.

**Hadoop For Dummies** *John Wiley & Sons* Let **Hadoop For Dummies** help harness the power of your data and rein in the information overload Big data has become big business, and companies and organizations of all sizes are struggling to find ways to retrieve valuable information from their massive data sets with becoming overwhelmed. Enter Hadoop and this easy-to-understand **For Dummies** guide. **Hadoop For Dummies** helps readers understand the value of big data, make a business case for using Hadoop, navigate the Hadoop ecosystem, and build and manage

**Hadoop applications and clusters. Explains the origins of Hadoop, its economic benefits, and its functionality and practical applications Helps you find your way around the Hadoop ecosystem, program MapReduce, utilize design patterns, and get your Hadoop cluster up and running quickly and easily Details how to use Hadoop applications for data mining, web analytics and personalization, large-scale text processing, data science, and problem-solving Shows you how to improve the value of your Hadoop cluster, maximize your investment in Hadoop, and avoid common pitfalls when building your Hadoop cluster From programmers challenged with building and maintaining affordable, scalable data systems to administrators who must deal with huge volumes of information effectively and efficiently, this how-to has something to help you with Hadoop. Mastering Julia *Packt Publishing Ltd* 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 computing tasks and observe 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. Modern Java in Action Lambdas, streams, functional and reactive programming *Simon and Schuster* Summary Manning's bestselling Java 8 book has been revised for Java 9! In Modern Java in Action, you'll build on your existing Java language skills with the newest features and techniques. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Modern applications take advantage of innovative designs, including microservices, reactive architectures, and streaming data. Modern Java features like lambdas, streams, and the long-awaited Java Module System make implementing these designs significantly easier. It's time to upgrade your skills and meet these challenges head on! About the Book Modern Java in Action connects new features of the Java language with their practical applications. Using crystal-clear examples and careful attention to detail, this book respects your time. It will help you expand your existing knowledge of core Java as you master modern additions like the Streams API and the Java Module System, explore new approaches to concurrency, and learn how functional concepts can help you write code that's easier to read and maintain. What's inside Thoroughly revised edition of Manning's bestselling Java 8 in Action New features in Java 8, Java 9, and beyond Streaming data and reactive programming The Java Module System About the Reader Written for developers familiar**

with core Java features. About the Author Raoul-Gabriel Urma is CEO of Cambridge Spark. Mario Fusco is a senior software engineer at Red Hat. Alan Mycroft is a University of Cambridge computer science professor; he cofounded the Raspberry Pi Foundation. Table of Contents PART 1 - FUNDAMENTALS Java 8, 9, 10, and 11: what's happening? Passing code with behavior parameterization Lambda expressions PART 2 - FUNCTIONAL-STYLE DATA PROCESSING WITH STREAMS Introducing streams Working with streams Collecting data with streams Parallel data processing and performance PART 3 - EFFECTIVE PROGRAMMING WITH STREAMS AND LAMBDA Collection API enhancements Refactoring, testing, and debugging Domain-specific languages using lambdas PART 4 - EVERYDAY JAVA Using Optional as a better alternative to null New Date and Time API Default methods The Java Module System PART 5 - ENHANCED JAVA CONCURRENCY Concepts behind CompletableFuture and reactive programming CompletableFuture: composable asynchronous programming Reactive programming PART 6 - FUNCTIONAL PROGRAMMING AND FUTURE JAVA EVOLUTION Thinking functionally Functional programming techniques Blending OOP and FP: Comparing Java and Scala Conclusions and where next for Java Mastering PostGIS *Packt Publishing Ltd* 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 RDBMS. 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.

**Style and approach** This book is a comprehensive guide covering all the concepts you need to master PostGIS. Packed with hands-on examples, tips and tricks, even the most advanced concepts are explained in a very easy-to-follow manner. Every chapter in the book does not only focus on how each task is performed, but also why.

**Getting Started with Grails** *Lulu.com* Grails is an open-source, rapid web application development framework that provides a super-productive full-stack programming model based on the Groovy scripting language and built on top of Spring, Hibernate, and other standard Java frameworks. Ruby on Rails pioneered the innovative coupling of a powerful programming language and an opinionated framework that favors sensible defaults over complex configuration, but many organizations aren't yet ready to stray from the safety of Java or forgo their current Java investments. Grails makes it possible to achieve equivalent productivity in a Java-centric environment. Over the course of this book, the reader will explore the various aspects of Grails and also experience Grails by building a Grails app.

**Mongodb in Action** *Manning Publications* Application developers love MongoDB, a document-oriented NoSQL database, for its speed, flexibility, scalability, and ease of use. MongoDB is well-suited as a back-end for modern web applications. Its schema-free design encourages rapid application development, and built-in replication and auto-sharding architecture allow for massive parallel distribution. Production deployments at SourceForge, Foursquare, and Shutterfly demonstrate daily that MongoDB is up to real-world challenges. **MongoDB in Action, Second Edition** is a comprehensive guide to MongoDB version 2.6. It begins with a general overview of current database systems, explaining what makes MongoDB unique and describing its ideal use cases. Then, a series of tutorials lead into detailed examples for leveraging MongoDB in e-commerce, social networking, and other common applications. A reference section on schema design patterns helps ease the transition from the relational data model of SQL to MongoDB's document-based data model. Purchase of the print book comes with an offer of a free PDF, ePub, and Kindle eBook from Manning. Also available is all code from the book.

**JavaScript and jQuery Interactive Front-End Web Development** *John Wiley & Sons* A visual and accessible guide to JavaScript and jQuery in a built-to-last hardcover edition In JavaScript & jQuery renowned author Jon Duckett discards the traditional programming book template and approaches writing code in a more relevant, less intimidating way. Full-color and packed with instructional graphics and photos, his books have gained a loyal following by illustrating programming in a way both instructive for newcomers and invaluable for seasoned coders. By discussing JavaScript and jQuery in a single text, Duckett ensures you will quickly be writing your own working scripts. Yet the book doesn't assume you have experience in either JavaScript or jQuery. By making use of popular jQuery plugins, Duckett illustrates techniques that would require pages upon pages of detailed explanation if you were being taught to create

the plugins yourself. This durable and attractive hardcover edition is a book you will have open on your desk as a reference for years to come. A timeless and lasting version of a classic "Duckett" book in a hardbound, dust-jacketed edition Completely accessible to those who feel intimidated by the subject matter and relevant to all front-end designers and developers who need to understand JavaScript Discusses applying responsive design techniques, leveraging APIs to save you time and work in coding, and identifying when to use CSS transforms over pure JavaScript Beautiful 4-color illustrations and examples make this a terrific academic resource for those of all experience levels JavaScript & jQuery clearly explains the jargon of programming, addressing the vocabulary without making it a prerequisite for readers. Duckett speaks directly to readers, making this an invaluable resource. This book is also available as a set, Web Design with HTML, CSS, JavaScript and jQuery Set 978119038634 along with HTML and CSS Design and Build Websites 9781118871645. Mastering Apache Storm *Packt Publishing Ltd* Master the intricacies of Apache Storm and develop real-time stream processing applications with ease About This Book Exploit the various real-time processing functionalities offered by Apache Storm such as parallelism, data partitioning, and more Integrate Storm with other Big Data technologies like Hadoop, HBase, and Apache Kafka An easy-to-understand guide to effortlessly create distributed applications with Storm Who This Book Is For If you are a Java developer who wants to enter into the world of real-time stream processing applications using Apache Storm, then this book is for you. No previous experience in Storm is required as this book starts from the basics. After finishing this book, you will be able to develop not-so-complex Storm applications. What You Will Learn Understand the core concepts of Apache Storm and real-time processing Follow the steps to deploy multiple nodes of Storm Cluster Create Trident topologies to support various message-processing semantics Make your cluster sharing effective using Storm scheduling Integrate Apache Storm with other Big Data technologies such as Hadoop, HBase, Kafka, and more Monitor the health of your Storm cluster In Detail Apache Storm is a real-time Big Data processing framework that processes large amounts of data reliably, guaranteeing that every message will be processed. Storm allows you to scale your data as it grows, making it an excellent platform to solve your big data problems. This extensive guide will help you understand right from the basics to the advanced topics of Storm. The book begins with a detailed introduction to real-time processing and where Storm fits in to solve these problems. You'll get an understanding of deploying Storm on clusters by writing a basic Storm Hello World example. Next we'll introduce you to Trident and you'll get a clear understanding of how you can develop and deploy a trident topology. We cover topics such as monitoring, Storm Parallelism, scheduler and log processing, in a very easy to understand manner. You will also learn how to integrate Storm with other well-known Big Data technologies such as HBase, Redis, Kafka, and Hadoop to realize the full potential of Storm. With real-world

examples and clear explanations, this book will ensure you will have a thorough mastery of Apache Storm. You will be able to use this knowledge to develop efficient, distributed real-time applications to cater to your business needs.

**Style and approach** This easy-to-follow guide is full of examples and real-world applications to help you get an in-depth understanding of Apache Storm. This book covers the basics thoroughly and also delves into the intermediate and slightly advanced concepts of application development with Apache Storm.

**60 Recipes for Apache CloudStack Using the CloudStack Ecosystem** "O'Reilly Media, Inc." Planning to deploy and maintain a public, private, or hybrid cloud service? This cookbook's handy how-to recipes help you quickly learn and install Apache CloudStack, along with several API clients, API wrappers, data architectures, and configuration management technologies that work as part of CloudStack's ecosystem. You'll learn how to use Vagrant, Ansible, Chef, Fluentd, Libcloud, and several other open source tools that let you build and operate CloudStack better and faster. If you're an experienced programmer, system administrator, or DevOps practitioner familiar with bash, Git, package management, and some Python, you're ready to go. Learn basic CloudStack installation from source, including features such as DevCloud, the CloudStack sandbox Get a step-by-step guide for installing CloudStack from packages on Ubuntu 14.04 using KVM Write your own applications on top of the CloudStack API, using CloudMonkey, Libcloud, jclouds, and CloStack Expose different APIs on CloudStack with the EC2Stack, Boto, and Eutester API wrappers Deploy applications easily, using Puppet, Salt, Ansible, Chef, and Vagrant Dive into cloud monitoring and storage with RiakCS, Fluentd, and Apache Whirr Deep Learning A Practitioner's Approach "O'Reilly Media, Inc." How can machine learning--especially deep neural networks--make a real difference in your organization? This hands-on guide not only provides practical information, but helps you get started building efficient deep learning networks. The authors provide the fundamentals of deep learning--tuning, parallelization, vectorization, and building pipelines--that are valid for any library before introducing the open source Deeplearning4j (DL4J) library for developing production-class workflows. Through real-world examples, you'll learn methods and strategies for training deep network architectures and running deep learning workflows on Spark and Hadoop with DL4J. Mastering Spark with R The Complete Guide to Large-Scale Analysis and Modeling "O'Reilly Media, Inc." If you're like most R users, you have deep knowledge and love for statistics. But as your organization continues to collect huge amounts of data, adding tools such as Apache Spark makes a lot of sense. With this practical book, data scientists and professionals working with large-scale data applications will learn how to use Spark from R to tackle big data and big compute problems. Authors Javier Luraschi, Kevin Kuo, and Edgar Ruiz show you how to use R with Spark to solve different data analysis problems. This book covers relevant data science topics, cluster computing, and issues that should interest even the most advanced users. Analyze, explore, transform, and visualize data in Apache Spark with R

Create statistical models to extract information and predict outcomes; automate the process in production-ready workflows Perform analysis and modeling across many machines using distributed computing techniques Use large-scale data from multiple sources and different formats with ease from within Spark Learn about alternative modeling frameworks for graph processing, geospatial analysis, and genomics at scale Dive into advanced topics including custom transformations, real-time data processing, and creating custom Spark extensions Mastering Akka *Packt Publishing Ltd* 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 you 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. Enterprise Java Microservices *Simon and Schuster* Summary Enterprise Java Microservices is an example-rich tutorial that shows how to design and manage large-scale Java applications as a collection of microservices. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Large

applications are easier to develop and maintain when you build them from small, simple components. Java developers now enjoy a wide range of tools that support microservices application development, including right-sized app servers, open source frameworks, and well-defined patterns. Best of all, you can build microservices applications using your existing Java skills. About the Book Enterprise Java Microservices teaches you to design and build JVM-based microservices applications. You'll start by learning how microservices designs compare to traditional Java EE applications. Always practical, author Ken Finnigan introduces big-picture concepts along with the tools and techniques you'll need to implement them. You'll discover ecosystem components like Netflix Hystrix for fault tolerance and master the Just enough Application Server (JeAS) approach. To ensure smooth operations, you'll also examine monitoring, security, testing, and deploying to the cloud. What's inside The microservices mental model Cloud-native development Strategies for fault tolerance and monitoring Securing your finished applications About the Reader This book is for Java developers familiar with Java EE. About the Author Ken Finnigan leads the Thorntail project at Red Hat, which seeks to make developing microservices for the cloud with Java and Java EE as easy as possible. Table of Contents PART 1 MICROSERVICES BASICS Enterprise Java microservices Developing a simple RESTful microservice Just enough Application Server for microservices Microservices testing Cloud native development PART 2 - IMPLEMENTING ENTERPRISE JAVA MICROSERVICES Consuming microservices Discovering microservices for consumption Strategies for fault tolerance and monitoring Securing a microservice Architecting a microservice hybrid Data streaming with Apache Kafka Microservice Architecture Aligning Principles, Practices, and Culture "O'Reilly Media, Inc." Have you heard about the tremendous success Amazon and Netflix have had by switching to a microservice architecture? Are you wondering how this can benefit your company? Or are you skeptical about how it might work? If you've answered yes to any of these questions, this practical book will benefit you. You'll learn how to take advantage of the microservice architectural style for building systems, and learn from the experiences of others to adopt and execute this approach most successfully. SQL Server Integration Services Design Patterns Apress SQL Server Integration Services Design Patterns is newly-revised for SQL Server 2014, and is a book of recipes for SQL Server Integration Services (SSIS). Design patterns in the book help to solve common problems encountered when developing data integration solutions. The patterns and solution examples in the book increase your efficiency as an SSIS developer, because you do not have to design and code from scratch with each new problem you face. The book's team of expert authors take you through numerous design patterns that you'll soon be using every day, providing the thought process and technical details needed to support their solutions. SQL Server Integration Services Design Patterns goes beyond the surface of the immediate problems to be solved, delving into why particular problems should

be solved in certain ways. You'll learn more about SSIS as a result, and you'll learn by practical example. Where appropriate, the book provides examples of alternative patterns and discusses when and where they should be used. Highlights of the book include sections on ETL Instrumentation, SSIS Frameworks, Business Intelligence Markup Language, and Dependency Services. Takes you through solutions to common data integration challenges Provides examples involving Business Intelligence Markup Language Teaches SSIS using practical examples Solr in Action *Simon and Schuster* Summary Solr in Action is a comprehensive guide to implementing scalable search using Apache Solr. This clearly written book walks you through well-documented examples ranging from basic keyword searching to scaling a system for billions of documents and queries. It will give you a deep understanding of how to implement core Solr capabilities. About the Book Whether you're handling big (or small) data, managing documents, or building a website, it is important to be able to quickly search through your content and discover meaning in it. Apache Solr is your tool: a ready-to-deploy, Lucene-based, open source, full-text search engine. Solr can scale across many servers to enable real-time queries and data analytics across billions of documents. Solr in Action teaches you to implement scalable search using Apache Solr. This easy-to-read guide balances conceptual discussions with practical examples to show you how to implement all of Solr's core capabilities. You'll master topics like text analysis, faceted search, hit highlighting, result grouping, query suggestions, multilingual search, advanced geospatial and data operations, and relevancy tuning. This book assumes basic knowledge of Java and standard database technology. No prior knowledge of Solr or Lucene is required. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. What's Inside How to scale Solr for big data Rich real-world examples Solr as a NoSQL data store Advanced multilingual, data, and relevancy tricks Coverage of versions through Solr 4.7 About the Authors Trey Grainger is a director of engineering at CareerBuilder. Timothy Potter is a senior member of the engineering team at LucidWorks. The authors work on the scalability and reliability of Solr, as well as on recommendation engine and big data analytics technologies. Table of Contents PART 1 MEET SOLR Introduction to Solr Getting to know Solr Key Solr concepts Configuring Solr Indexing Text analysis PART 2 CORE SOLR CAPABILITIES Performing queries and handling results Faceted search Hit highlighting Query suggestions Result grouping/field collapsing Taking Solr to production PART 3 TAKING SOLR TO THE NEXT LEVEL SolrCloud Multilingual search Complex query operations Mastering relevancy High Performance Python Practical Performant Programming for Humans *O'Reilly Media* Your Python code may run correctly, but you need it to run faster. Updated for Python 3, this expanded edition shows you how to locate performance bottlenecks and significantly speed up your code in high-data-volume programs. By exploring the fundamental theory behind design choices, High Performance Python helps you gain a deeper understanding of

Python's implementation. How do you take advantage of multicore architectures or clusters? Or build a system that scales up and down without losing reliability? Experienced Python programmers will learn concrete solutions to many issues, along with war stories from companies that use high-performance Python for social media analytics, productionized machine learning, and more. Get a better grasp of NumPy, Cython, and profilers Learn how Python abstracts the underlying computer architecture Use profiling to find bottlenecks in CPU time and memory usage Write efficient programs by choosing appropriate data structures Speed up matrix and vector computations Use tools to compile Python down to machine code Manage multiple I/O and computational operations concurrently Convert multiprocessing code to run on local or remote clusters Deploy code faster using tools like Docker Mastering Apache Spark 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 H2O, 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 H2O 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 H2O 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.