Mastering Apache Solr 7 X An Expert Guide To Advancing Optimizing And Scaling Your Enterprise Search

Learn everything you need to build highly scalable, robust web applications using Angular release 4 About This Book Apply best practices and design patterns to achieve higher scalability in your Angular applications Understand the latest features of Angular and create your own components Get acquainted with powerful, advanced techniques in Angular to build professional web applications Who This Book Is For This book is for JavaScript developers with some prior exposure to Angular, at least through basic examples. We assume that you've got working knowledge of HTML, CSS, and JavaScript. What You Will Learn Implement asynchronous programming using Angular Beautify your application with the UI components built to the material design specification Secure your web application from unauthorized users Create complex forms, taking full advantage of 2-way data binding Test your Angular applications using the Jasmine and Protractor frameworks for better efficiency Learn how to integrate Angular with Bootstrap to create compelling web applications Use Angular built-in classes to apply animation in your app In Detail Got some experience of Angular under your belt? Want to learn everything about using advanced features for developing websites? This book is everything you need for the deep understanding of Angular that will set you apart from the developer crowd. Angular has introduced a new way to build applications. Creating complex and rich web applications, with a lighter resource footprint, has never been easier or faster. Angular is now at release 4, with significant changes through previous versions. This book has been written and tested for Angular release 4. Angular is a mature technology, and you'll likely have applications built with earlier versions. This book starts by showing you best practices and approaches to migrating your existing Angular applications so that you can be immediately up-to-date. You will take an in-depth look at components and see how to control the user journey in your applications by implementing routing and navigation. You will learn how to work with asynchronous programming by using Observables. To easily build applications that look great, you will learn all about template syntax and how to beautify applications with Material Design. Mastering forms and data binding will further speed up your application development time. Learning about managing services and animations will help you to progressively enhance your applications. Next you'll use native directives to integrate Bootstrap with Angular. You will see the best ways to test your application with the leading options such as Jasmine and Protractor. At the end of the book, you'll learn how to apply design patterns in Angular, and see the benefits they will bring to your development. Style and approach This book provides comprehensive coverage of all aspects of development with Angular. You will learn about all the most powerful Angular concepts, with examples and best practices. This book is everything you need for the deep understanding of Angular that will set you apart from the developer crowd.

Topic: In the open source, full-text search community, a leader emerges – Apache Solr. Apache Solr enables you to index and access documents orders of magnitude faster than classical databases and thereby provides a first-class search experience to your end users. Brief Description: Mastering Apache Solr is a practical, hands-on guide containing crisp, relevant, systematically arranged, and progressive chapters. These chapters contain a wealth of information presented in a direct and easy-to-understand manner. This book covers key technical concepts, highlighting Solr's supremacy over classical databases in full-text search, which will help you accelerate your progress in the Solr world. Detailed Description: Mastering Apache Solr starts with an introduction to Apache Solr, its underlying technologies, the main differences between the classical database engines, and gradually moves to more advance topics like boosting performance. In this book, we will look under the hood of a large number of topics and discuss answers to pertinent questions like why denormalize data, how to import classical databases' data inside Apache Solr, how to serve Solr through five different web servers, how to optimize them to serve Solr even faster. An important and major topic covered in this book is Solr's querying mechanism, which will prove to be a strong ally in our journey through this book. We then look at boosting performance and deploying Solr using several servlet servers. Finally, we cover how to communicate with Solr using different programming languages, before deploying it in a cloud-based environment. Who this book is for: Mastering Apache Solr has been written for developers, programmers, and data specialists who want to take a leap towards the future of full-text storage and search and offer a world-class experience to their users. The reader is expected to have a working knowledge of traditional databases, Linux-based operating systems, and XML configuration files. Style and Approach: Mastering Apache Solr is written lucidly and has a dynamically simple approach. From the first page to the last, the book remains practical and focuses on the most important topics used in the world of Apache Solr without neglecting important theoretical fundamentals that help you build a strong foundation. Conclusion: Mastering Apache Solr will empower you to provide a world-class search experience to your end users through the discovery of the powerful mechanisms presented in this book.

Over 100 practical recipes to make Apache Solr faster, more reliable and return better results.

This book is part of Packt's Cookbook series; each chapter looks at a different aspect of working with Apache Solr. The recipes deal with common problems of working with Solr by using easy-to-understand, real-life examples. The book is not in any way a complete Apache Solr reference and you should see it as a helping hand when things get rough on your journey with Apache Solr. Developers who are working with Apache Solr and would like to know how to combat common problems will find this book of great use. Knowledge of Apache Lucene would be a bonus but is not required.

A step-by-step tutorial on implementing Liferay- based portals to learn performance best practices.The book is good for Liferay portal developers and architects who want to learn performance best practices for implementing Liferay- based solutions. It is assumed that you have a working knowledge of the Liferay portal.

JSON is becoming the backbone for meaningful data interchange over the internet. This format is now supported by an

entire ecosystem of standards, tools, and technologies for building truly elegant, useful, and efficient applications. With this hands-on guide, author and architect Tom Marrs shows you how to build enterprise-class applications and services by leveraging JSON tooling and message/document design. JSON at Work provides application architects and developers with guidelines, best practices, and use cases, along with lots of real-world examples and code samples. You'll start with a comprehensive JSON overview, explore the JSON ecosystem, and then dive into JSON's use in the enterprise. Get acquainted with JSON basics and learn how to model JSON data Learn how to use JSON with Node.js, Ruby on Rails, and Java Structure JSON documents with JSON Schema to design and test APIs Search the contents of JSON documents with JSON Search tools Convert JSON documents to other data formats with JSON Transform tools Compare JSON-based hypermedia formats, including HAL and jsonapi Leverage MongoDB to store and access JSON documents Use Apache Kafka to exchange JSON-based messages between services

This book is aimed at developers, designers, and architects who would like to build big data enterprise search solutions for their customers or organizations. No prior knowledge of Apache Hadoop and Apache Solr/Lucene technologies is required.

This book is a step-by-step tutorial that will enable you to leverage the flexible search functionality of Apache Solr together with the Big Data power of Apache Hadoop.Scaling Big Data with Hadoop and Solr provides guidance to developers who wish to build high-speed enterprise search platforms using Hadoop and Solr. This book is primarily aimed at Java programmers who wish to extend the Hadoop platform to make it run as an enterprise search without any prior knowledge of Apache Hadoop and Solr.

Start using Kubernetes in complex big data and enterprise applications, including Docker containers. Starting with installing Kubernetes on a single node, the book introduces Kubernetes with a simple Hello example and discusses using environment variables in Kubernetes. Next, Kubernetes Microservices with Docker discusses using Kubernetes with all major groups of technologies such as relational databases, NoSQL databases, and in the Apache Hadoop ecosystem. The book concludes with using multi container pods and installing Kubernetes on a multi node cluster. /div "a concise but clear introduction to containers, Docker and Kubernetes, using simple real-world examples to pass on the core concepts, via repetition, and is a very useful enabler." 10/10 Dave Hay MBCS CITP: review for BCS, The Chartered Institute for IT (http://www.bcs.org/content/conWebDoc/58512) What You Will Learn Install Kubernetes on a single node Set environment variables Create multi-container pods using Docker Use volumes Use Kubernetes with the Apache Hadoop ecosystem, NoSQL databases, and RDBMSs Install Kubernetes on a multi-node cluster Who This Book Is For Application developers including Apache Hadoop developers, database developers and NoSQL developers.

Spring 5.0 brings major advancements in the rich APIs provided by the Spring framework and thus creates a need for developers to master its tools and techniques to achieve high-performing applications. This book will help you improve the speed of your code and optimize the performance of your apps.

Unleash the data processing and analytics capability of Apache Spark with the language of choice: Java About This Book Perform big data processing with Spark—without having to learn Scala! Use the Spark Java API to implement efficient enterprise-grade applications for data processing and analytics Go beyond mainstream data processing by adding querying capability, Machine Learning, and graph processing using Spark Who This Book Is For If you are a Java developer interested in learning to use the popular Apache Spark framework, this book is the resource you need to get started. Apache Spark developers who are looking to build enterprise-grade applications in Java will also find this book very useful. What You Will Learn Process data using different file formats such as XML, JSON, CSV, and plain and delimited text, using the Spark core Library. Perform analytics on data from various data sources such as Kafka, and Flume using Spark Streaming Library Learn SQL schema creation and the analysis of structured data using various SQL functions including Windowing functions in the Spark SQL Library Explore Spark Mlib APIs while implementing Machine Learning techniques to solve real-world problems Get to know Spark GraphX so you understand various graph-based analytics that can be performed with Spark In Detail Apache Spark is the buzzword in the big data industry right now, especially with the increasing need for real-time streaming and data processing. While Spark is built on Scala, the Spark Java API exposes all the Spark features available in the Scala version for Java developers. This book will show you how you can implement various functionalities of the Apache Spark framework in Java, without stepping out of your comfort zone. The book starts with an introduction to the Apache Spark 2.x ecosystem, followed by explaining how to install and configure Spark, and refreshes the Java concepts that will be useful to you when consuming Apache Spark's APIs. You will explore RDD and its associated common Action and Transformation Java APIs, set up a production-like clustered environment, and work with Spark SQL. Moving on, you will perform near-real-time processing with Spark streaming, Machine Learning analytics with Spark MLlib, and graph processing with GraphX, all using various Java packages. By the end of the book, you will have a solid foundation in implementing components in the Spark framework in Java to build fast, real-time applications. Style and approach This practical guide teaches readers the fundamentals of the Apache Spark framework and how to implement components using the Java language. It is a unique blend of theory and practical examples, and is written in a way that will gradually build your knowledge of Apache Spark.

How prepared are you to build fast and efficient web applications? This eloquent book provides what every web developer should know about the network, from fundamental limitations that affect performance to major innovations for building even more powerful browser applications—including HTTP 2.0 and XHR improvements, Server-Sent Events (SSE), WebSocket, and WebRTC. Author Ilya Grigorik, a web performance engineer at Google, demonstrates performance optimization best practices for TCP, UDP, and TLS protocols, and explains unique wireless and mobile network optimization requirements. You'll then dive into performance characteristics of technologies such as HTTP 2.0, client-side network scripting with XHR, real-time streaming with SSE and WebSocket, and P2P communication with

WebRTC. Deliver superlative TCP, UDP, and TLS performance Speed up network performance over 3G/4G mobile networks Develop fast and energy-efficient mobile applications Address bottlenecks in HTTP 1.x and other browser protocols Plan for and deliver the best HTTP 2.0 performance Enable efficient real-time streaming in the browser Create efficient peer-to-peer videoconferencing and low-latency applications with real-time WebRTC transports

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

Summary Taming Text, winner of the 2013 Jolt Awards for Productivity, is a hands-on, example-driven guide to working with unstructured text in the context of real-world applications. This book explores how to automatically organize text using approaches such as full-text search, proper name recognition, clustering, tagging, information extraction, and summarization. The book guides you through examples illustrating each of these topics, as well as the foundations upon which they are built. About this Book There is so much text in our lives, we are practically drowningin it. Fortunately, there are innovative tools and techniquesfor managing unstructured information that can throw thesmart developer a much-needed lifeline. You'll find them in thisbook. Taming Text is a practical, example-driven guide to working withtext in real applications. This book introduces you to useful techniques like full-text search, proper name recognition,clustering, tagging, information extraction, and summarization.You'll explore real use cases as you systematically absorb thefoundations upon which they are built.Written in a clear and concise style, this book avoids jargon, explainingthe subject in terms you can understand without a backgroundin statistics or natural language processing. Examples arein Java, but the concepts can be applied in any language. Written for Java developers, the book requires no prior knowledge of GWT. 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. Winner of 2013 Jolt Awards: The Best Books—one of five notable books every serious programmer should read. What's Inside When to use text-taming techniques Important open-source libraries like Solr and Mahout How to build text-processing applications About the Authors Grant Ingersoll is an engineer, speaker, and trainer, a Lucenecommitter, and a cofounder of the Mahout machine-learning project. Thomas Morton is the primary developer of OpenNLP and Maximum Entropy. Drew Farris is a technology consultant, software developer, and contributor to Mahout,Lucene, and Solr. "Takes the mystery out of verycomplex processes."—From the Foreword by Liz Liddy, Dean, iSchool, Syracuse University Table of Contents Getting started taming text Foundations of taming text Searching Fuzzy string matching Identifying people, places, and things Clustering text Classification, categorization, and tagging Building an example question answering system Untamed text: exploring the next frontier

Your one stop solution to create highly scalable enterprise grade Java applications with WildFly. About This Book Master Java EE development with the latest WildFly 10 application server. Integrate with JSF and JMS and use efficient load balancing techniques to create real-time apps Integrate your backend JavaScript code seamlessly into Java applications Who This Book Is For If you are a Java developer with at least basic knowledge of Java EE, then this book is for you. No previous knowledge of WildFly is required. What You Will Learn Configure the development environment along with native and cloud installation of WildFly Write a DB schema and the relative entities and how to use the relationships between the entities Analyze with examples all the java annotations to manage the EJB and the configuration to get better performances Write different REST services through the EJB Implement Web sockets 1.0 and know why and when use the web sockets Work with Active MQ and write JMS clients to manage the authentication and authorization in the clients Configure the mail server through the wildfly console Learn how and when to use a new feature JAX-RS 2.0, which is the asynchronous call through REST Use the new JSF features of Wildfly 10 such as Mojarra 2.2, JSF 2.2, Richfaces 4.5 In Detail Packed with rich assets and APIs, Wildfly 10 allows you to create state-of-the-art Java applications. This book will help you take your understanding of Java EE to the next level by creating distributed Java applications using Wildfly. The book begins by showing how to get started with a native installation of WildFly and it ends with a cloud installation. After setting up the development environment, you will implement and work with different WildFly features, such as implementing JavaServer Pages. You will also learn how you can use clustering so that your apps can handle a high volume of data traffic. You will also work with enterprise JavaBeans, solve issues related to failover, and implement Java Message Service integration. Moving ahead, you will be working with Java Naming and Directory Interface, Java Transaction API, and use ActiveMQ for message relay and message querying. This book will also show you how you can use your existing backend JavaScript code in your application. By the end of the book, you'll have gained the knowledge to implement the latest Wildfly features in your Java applications. Style and approach Each part of this book shows you how to use different features of WildFly 10 to create enterprise grade Java applications as easily as possible.

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

Learn about the fastest-growing open source project in the world, and find out how it revolutionizes big data analytics About This Book Exclusive guide that covers how to get up and running with fast data processing using Apache Spark Explore and exploit various possibilities with Apache Spark using real-world use cases in this book Want to perform efficient data processing at real time? This book will be your one-stop solution. Who This Book Is For This guide appeals to big data engineers, analysts, architects, software engineers, even technical managers who need to perform efficient data processing on Hadoop at real time. Basic familiarity with Java or Scala will be helpful. The assumption is that readers will be from a mixed background, but would be typically people with background in engineering/data science with no prior Spark experience and want to understand how Spark can help them on their analytics journey. What You Will Learn Get an overview of big data analytics and its importance for organizations and data professionals Delve into Spark to see how it is different from existing processing platforms Understand the intricacies of various file formats, and how to process them with Apache Spark. Realize how to deploy Spark with YARN, MESOS or a Stand-alone cluster manager. Learn the concepts of Spark SQL, SchemaRDD, Caching and working with Hive and Parquet file formats Understand the architecture of Spark MLLib while discussing some of the off-the-shelf algorithms that come with Spark. Introduce yourself to the deployment and usage of SparkR. Walk through the importance of Graph computation and the graph processing systems available in the market Check the real world example of Spark by building a recommendation engine with Spark using ALS. Use a Telco data set, to predict customer churn using Random Forests. In Detail Spark juggernaut keeps on rolling and getting more and more momentum each day. Spark provides key capabilities in the form of Spark SQL, Spark Streaming, Spark ML and Graph X all accessible via Java, Scala, Python and R. Deploying the key capabilities is crucial whether it is on a Standalone framework or as a part of existing Hadoop installation and configuring with Yarn and Mesos. The next part of the journey after installation is using key components, APIs, Clustering, machine learning APIs, data pipelines, parallel programming. It is important to understand why each framework component is key, how widely it is being used, its stability and pertinent use cases. Once we understand the individual components, we will take a couple of real life advanced analytics examples such as 'Building a Recommendation system', 'Predicting customer churn' and so on. The objective of these real life examples is to give the reader confidence of using Spark for real-world problems. Style and approach With the help of practical examples and real-world use cases, this guide will take you from scratch to building efficient data applications using Apache Spark. You will learn all about this excellent data processing engine in a step-by-step manner, taking one aspect of it at a time. This highly practical guide will include how to work with data pipelines, dataframes, clustering, SparkSQL, parallel programming, and such insightful topics with the help of real-world use cases.

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

Learn the art of efficient web scraping and crawling with Python About This Book Extract data from any source to perform real time analytics. Full of techniques and examples to help you crawl websites and extract data within hours. A hands-on guide to web scraping and crawling with real-life problems and solutions Who This Book Is For If you are a software developer, data scientist, NLP or machine-learning enthusiast or just need to migrate your company's wiki from a legacy platform, then this book is for you. It is perfect for someone , who needs instant access to large amounts of semi-structured data effortlessly. What You Will Learn Understand HTML pages and write XPath to extract the data you need Write Scrapy spiders with simple Python and do web crawls Push your data into any database, search engine or analytics system Configure your spider to download files, images and use proxies Create efficient pipelines that shape data in precisely the form you want Use Twisted Asynchronous API to process hundreds of items concurrently Make your crawler super-fast by learning how to tune Scrapy's performance Perform large scale distributed crawls with scrapyd and scrapinghub In Detail This book covers the long awaited Scrapy v 1.0 that empowers you to extract useful data from virtually any source with very little effort. It starts off by explaining the fundamentals of Scrapy framework, followed by a thorough description of how to extract data from any source, clean it up, shape it as per your requirement using Python and 3rd party APIs. Next you will be familiarised with the process of storing the scrapped data in databases as well as search engines and performing real time analytics on them with Spark Streaming. By the end of this book, you will perfect the art of scarping data for your applications with ease Style and approach It is a hands on guide, with first few chapters written as a tutorial, aiming to motivate you and get you started quickly. As the book progresses, more advanced features are explained with real world examples that can be reffered while developing your own web applications.

Summary A developer-focused guide to writing applications using Spring Boot. You'll learn how to bypass the tedious configuration steps so that you can concentrate on your application's behavior. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology The Spring Framework simplifies enterprise Java development, but it does require lots of tedious configuration work. Spring Boot radically streamlines spinning up a Spring application. You get automatic configuration and a model with established conventions for build-time and runtime dependencies. You also get a handy command-line interface you can use to write

scripts in Groovy. Developers who use Spring Boot often say that they can't imagine going back to hand configuring their applications. About the Book Spring Boot in Action is a developer-focused guide to writing applications using Spring Boot. In it, you'll learn how to bypass configuration steps so you can focus on your application's behavior. Spring expert Craig Walls uses interesting and practical examples to teach you both how to use the default settings effectively and how to override and customize Spring Boot for your unique environment. Along the way, you'll pick up insights from Craig's years of Spring development experience. What's Inside Develop Spring apps more efficiently Minimal to no configuration Runtime metrics with the Actuator Covers Spring Boot 1.3 About the Reader Written for readers familiar with the Spring Framework. About the Author Craig Walls is a software developer, author of the popular book Spring in Action, Fourth Edition, and a frequent speaker at conferences. Table of Contents Bootstarting Spring Developing your first Spring Boot application Customizing configuration Testing with Spring Boot Getting Groovy with the Spring Boot CLI Applying Grails in Spring Boot Taking a peek inside with the Actuator Deploying Spring Boot applications APPENDIXES Spring Boot developer tools Spring Boot starters Configuration properties Spring Boot dependencies

Accelerate your enterprise search engine and bring relevancy in your search analytics Key Features A practical guide in building expertise with Indexing, Faceting, Clustering and Pagination Master the management and administration of Enterprise Search Applications and services seamlessly Handle multiple data inputs such as JSON, xml, pdf, doc, xls,ppt, csv and much more. Book Description Apache Solr is the only standalone enterprise search server with a REST-like application interface. providing highly scalable, distributed search and index replication for many of the world's largest internet sites. To begin with, you would be introduced to how you perform full text search, multiple filter search, perform dynamic clustering and so on helping you to brush up the basics of Apache Solr. You will also explore the new features and advanced options released in Apache Solr 7.x which will get you numerous performance aspects and making data investigation simpler, easier and powerful. You will learn to build complex queries, extensive filters and how are they compiled in your system to bring relevance in your search tools. You will learn to carry out Solr scoring, elements affecting the document score and how you can optimize or tune the score for the application at hand. You will learn to extract features of documents, writing complex queries in re-ranking the documents. You will also learn advanced options helping you to know what content is indexed and how the extracted content is indexed. Throughout the book, you would go through complex problems with solutions along with varied approaches to tackle your business needs. By the end of this book, you will gain advanced proficiency to build out-of-box smart search solutions for your enterprise demands. What you will learn Design schema using schema API to access data in the database Advance querying and fine-tuning techniques for better performance Get to grips with indexing using Client API Set up a fault tolerant and highly available server with newer distributed capabilities, SolrCloud Explore Apache Tika to upload data with Solr Cell Understand different data operations that can be done while indexing Master advanced querying through Velocity Search UI, faceting and Query Re-ranking, pagination and spatial search Learn to use JavaScript, Python, SolrJ and Ruby for interacting with Solr Who this book is for The book would rightly appeal to developers, software engineers, data engineers and database architects who are building or seeking to build enterprise-wide effective search engines for business intelligence. Prior experience of Apache Solr or Java programming is must to take the best of this book.

A fast paced guide that will help you learn about Apache Hadoop 3 and its ecosystem Key Features Set up, configure and get started with Hadoop to get useful insights from large data sets Work with the different components of Hadoop such as MapReduce, HDFS and YARN Learn about the new features introduced in Hadoop 3 Book Description Apache Hadoop is a widely used distributed data platform. It enables large datasets to be efficiently processed instead of using one large computer to store and process the data. This book will get you started with the Hadoop ecosystem, and introduce you to the main technical topics, including MapReduce, YARN, and HDFS. The book begins with an overview of big data and Apache Hadoop. Then, you will set up a pseudo Hadoop development environment and a multi-node enterprise Hadoop cluster. You will see how the parallel programming paradigm, such as MapReduce, can solve many complex data processing problems. The book also covers the important aspects of the big data software development lifecycle, including quality assurance and control, performance, administration, and monitoring. You will then learn about the Hadoop ecosystem, and tools such as Kafka, Sqoop, Flume, Pig, Hive, and HBase. Finally, you will look at advanced topics, including real time streaming using Apache Storm, and data analytics using Apache Spark. By the end of the book, you will be well versed with different configurations of the Hadoop 3 cluster. What you will learn Store and analyze data at scale using HDFS, MapReduce and YARN Install and configure Hadoop 3 in different modes Use Yarn effectively to run different applications on Hadoop based platform Understand and monitor how Hadoop cluster is managed Consume streaming data using Storm, and then analyze it using Spark Explore Apache Hadoop ecosystem components, such as Flume, Sqoop, HBase, Hive, and Kafka Who this book is for Aspiring Big Data professionals who want to learn the essentials of Hadoop 3 will find this book to be useful. Existing Hadoop users who want to get up to speed with the new features introduced in Hadoop 3 will also benefit from this book. Having knowledge of Java programming will be an added advantage.

IBM® WatsonTM Content Analytics (Content Analytics) Version 3.0 (formerly known as IBM Content Analytics with Enterprise Search (ICAwES)) helps you to unlock the value of unstructured content to gain new actionable business insight and provides the enterprise search capability all in one product. Content Analytics comes with a set of tools and a robust user interface to empower you to better identify new revenue opportunities, improve customer satisfaction, detect problems early, and improve products, services, and offerings. To help you gain the most benefits from your unstructured content, this IBM Redbooks® publication provides in-depth information about the features and capabilities of Content Analytics, how the content analytics works, and how to perform effective and efficient content analytics on your content to discover actionable business insights. This book covers key concepts in content analytics, such as facets, frequency, deviation, correlation, trend, and sentimental analysis. It describes the content analytics miner, and guides you on performing content analytics using views, dictionary lookup, and customization. The book also covers using IBM Content Analytics Studio for domain-specific content analytics, integrating with IBM Content Classification to get categories and new metadata, and interfacing with IBM Cognos® Business Intelligence (BI) to add values in BI reporting and analysis, and customizing the content analytics miner with APIs. In addition, the book describes how to use the enterprise search capability for the discovery and retrieval of documents using various query and visual navigation techniques, and customization of crawling, parsing, indexing, and runtime search to improve search results. The target audience of this book is decision makers, business users, and IT architects and specialists who want to understand and analyze their enterprise content to improve and enhance their business operations. It is also intended as a technical how-to guide for use with the online IBM Knowledge Center for configuring and performing content analytics and enterprise search with Content Analytics.

Develop cloud native applications with microservices using Spring Boot, Spring Cloud, and Spring Cloud Data Flow About This Book Explore the new features and components in Spring Evolve towards micro services and cloud native applications Gain powerful insights into advanced concepts of Spring and Spring Boot to develop applications more effectively Understand the basics of Kotlin and use it to develop a quick service with Spring Boot Who This Book Is For This book is for an experienced Java developer who knows the basics of Spring, and wants to learn how to use Spring Boot to build applications and deploy them to the cloud. What You Will Learn Explore the new features in Spring Framework 5.0 Build microservices with Spring Boot Get to know the advanced features of Spring Boot in order to effectively develop and monitor applications Use Spring Cloud to deploy and manage applications on the Cloud Understand Spring Data and Spring Cloud Data Flow Understand the basics of reactive programming Get to know the best practices when developing applications with the Spring Framework Create a new project using Kotlin and implement a couple of basic services with unit and integration testing In Detail Spring 5.0 is

due to arrive with a myriad of new and exciting features that will change the way we've used the framework so far. This book will show you this evolution—from solving the problems of testable applications to building distributed applications on the cloud. The book begins with an insight into the new features in Spring 5.0 and shows you how to build an application using Spring MVC. You will realize how application architectures have evolved from monoliths to those built around microservices. You will then get a thorough understanding of how to build and extend microservices using Spring Boot. You will also understand how to build and deploy Cloud-Native microservices with Spring Cloud. The advanced features of Spring Boot will be illustrated through powerful examples. We will be introduced to a JVM language that's quickly gaining popularity - Kotlin. Also, we will discuss how to set up a Kotlin project in Eclipse. By the end of the book, you will be equipped with the knowledge and best practices required to develop microservices with the Spring Framework. Style and approach This book follows an end-to-end tutorial approach with lots of examples and sample applications, covering the major building blocks of the Spring framework.

Uncover the power of MySQL 8 for Big Data About This Book Combine the powers of MySQL and Hadoop to build a solid Big Data solution for your organization Integrate MySQL with different NoSQL APIs and Big Data tools such as Apache Sqoop A comprehensive guide with practical examples on building a high performance Big Data pipeline with MySQL Who This Book Is For This book is intended for MySQL database administrators and Big Data professionals looking to integrate MySQL 8 and Hadoop to implement a high performance Big Data solution. Some previous experience with MySQL will be helpful, although the book will highlight the newer features introduced in MySQL 8. What You Will Learn Explore the features of MySQL 8 and how they can be leveraged to handle Big Data Unlock the new features of MySQL 8 for managing structured and unstructured Big Data Integrate MySQL 8 and Hadoop for efficient data processing Perform aggregation using MySQL 8 for optimum data utilization Explore different kinds of join and union in MySQL 8 to process Big Data efficiently Accelerate Big Data processing with Memcached Integrate MySQL with the NoSQL API Implement replication to build highly available solutions for Big Data In Detail With organizations handling large amounts of data on a regular basis, MySQL has become a popular solution to handle this structured Big Data. In this book, you will see how DBAs can use MySQL 8 to handle billions of records, and load and retrieve data with performance comparable or superior to commercial DB solutions with higher costs. Many organizations today depend on MySQL for their websites and a Big Data solution for their data archiving, storage, and analysis needs. However, integrating them can be challenging. This book will show you how to implement a successful Big Data strategy with Apache Hadoop and MySQL 8. It will cover real-time use case scenario to explain integration and achieve Big Data solutions using technologies such as Apache Hadoop, Apache Sqoop, and MySQL Applier. Also, the book includes case studies on Apache Sqoop and real-time event processing. By the end of this book, you will know how to efficiently use MySQL 8 to manage data for your Big Data applications. Style and approach Step by Step guide filled with real-world practical examples.

ElasticSearch is an open source search server built on Apache Lucene. It was built to provide a scalable search solution with built-in support for near real-time search and multi-tenancy.Jumping into the world of ElasticSearch by setting up your own custom cluster, this book will show you how to create a fast, scalable, and flexible search solution. By learning the ins-and-outs of data indexing and analysis, "ElasticSearch Server" will start you on your journey to mastering the powerful capabilities of ElasticSearch. With practical chapters covering how to search data, extend your search, and go deep into cluster administration and search analysis, this book is perfect for those new and experienced with search servers.In "ElasticSearch Server" you will learn how to revolutionize your website or application with faster, more accurate, and flexible search functionality. Starting with chapters on setting up your own ElasticSearch cluster and searching and extending your search parameters you will quickly be able to create a fast, scalable, and completely custom search solution.Building on your knowledge further you will learn about ElasticSearch's query API and become confident using powerful filtering and faceting capabilities. You will develop practical knowledge on how to make use of ElasticSearch's near real-time capabilities and support for multi-tenancy.Your journey then concludes with chapters that help you monitor and tune your ElasticSearch cluster as well as advanced topics such as shard allocation, gateway configuration, and the discovery module.

The organization pursuing digital transformation must embrace new ways to use and deploy integration technologies, so they can move quickly in a manner appropriate to the goals of multicloud, decentralization, and microservices. The integration layer must transform to allow organizations to move boldly in building new customer experiences, rather than forcing models for architecture and development that pull away from maximizing the organization's productivity. Many organizations have started embracing agile application techniques, such as microservice architecture, and are now seeing the benefits of that shift. This approach complements and accelerates an enterprise's API strategy. Businesses should also seek to use this approach to modernize their existing integration and messaging infrastructure to achieve more effective ways to manage and operate their integration services in their private or public cloud. This IBM® Redbooks® publication explores the merits of what we refer to as agile integration; a container-based, decentralized, and microservice-aligned approach for integration solutions that meets the demands of agility, scalability, and resilience required by digital transformation. It also discusses how the IBM Cloud Pak for Integration marks a significant leap forward in integration technology by embracing both a cloud-native approach and container technology to achieve the goals of agile integration. The target audiences for this book are cloud integration architects, IT specialists, and application developers.

Get the most out of Elasticsearch 7's new features to build, deploy, and manage efficient applications Key Features Discover the new features introduced in Elasticsearch 7 Explore techniques for distributed search, indexing, and clustering Gain hands-on knowledge of implementing Elasticsearch for your enterprise Book Description Elasticsearch is one of the most popular tools for distributed search and analytics. This Elasticsearch book highlights the latest features of Elasticsearch 7 and helps you understand how you can use them to build your own search applications with ease. Starting with an introduction to the Elastic Stack, this book will help you quickly get up to speed with using Elasticsearch.

You'll learn how to install, configure, manage, secure, and deploy Elasticsearch clusters, as well as how to use your deployment to develop powerful search and analytics solutions. As you progress, you'll also understand how to troubleshoot any issues that you may encounter along the way. Finally, the book will help you explore the inner workings of Elasticsearch and gain insights into queries, analyzers, mappings, and aggregations as you learn to work with search results. By the end of this book, you'll have a basic understanding of how to build and deploy effective search and analytics solutions using Elasticsearch. What you will learn Install Elasticsearch and use it to safely store data and retrieve it when needed Work with a variety of analyzers and filters Discover techniques to improve search results in Elasticsearch Understand how to perform metric and bucket aggregations Implement best practices for moving clusters and applications to production Explore various techniques to secure your Elasticsearch clusters Who this book is for This book is for software developers, engineers, data architects, system administrators, and anyone who wants to get up and running with Elasticsearch 7. No prior experience with Elasticsearch is required.

Step by step guide to monitor, manage, and secure your database engine Key Features Your companion to master all the administration-related tasks in MySQL 8 Ensure high performance and high availability of your MySQL solution using effective replication and backup techniques A comprehensive guide to performing query optimization, security and a whole host of other administrative tasks in MySQL 8 Book Description MySQL is one of the most popular and widely used relational databases in the world today. The recently released version 8.0 brings along some major advancements in the way your MySQL solution can be administered. This handbook will be your companion to understand the newly introduced features in MySQL and how you can leverage them to design a high-performance MySQL solution for your organization. This book starts with a brief introduction to the newly introduced features in MySQL 8, followed by quickly jumping onto the crucial administration topics that you will find useful in your day to day work. Topics such as migrating to MySQL 8, MySQL benchmarking, achieving high performance by implementing the indexing techniques, and optimizing your queries are covered in this book. You will also learn how to perform replication, scale your MySQL solution and implement effective security techniques. A special section on the common and not so common troubleshooting techniques for effective MySQL administration is also covered in this book. By the end of this highly practical book, you will have all the knowledge you need to tackle any problem you might encounter while administering your MySQL solution. What you will learn Understanding different MySQL 8 data types based on type of contents and storage requirements Best practices for optimal use of features in MySQL 8 Explore globalization configuration and caching techniques to improve performance Create custom storage engine as per system requirements Learn various ways of index implementation for flash memory storages Configure and implement replication along with approaches to use replication as solution Understand how to make your MySQL 8 solution highly available Troubleshoot common issues and identify error codes while using MySQL 8 Who this book is for This book is intended for MySQL administrators who are looking for a handy guide covering all the MySQL administration-related tasks. If you are a DBA looking to get started with MySQL administration, this book will also help you. Knowledge of the basic database concepts is required to get started with this book.

Master text-taming techniques and build effective text-processing applications with R About This Book Develop all the relevant skills for building text-mining apps with R with this easy-to-follow guide Gain in-depth understanding of the text mining process with lucid implementation in the R language Example-rich guide that lets you gain high-quality information from text data Who This Book Is For If you are an R programmer, analyst, or data scientist who wants to gain experience in performing text data mining and analytics with R, then this book is for you. Exposure to working with statistical methods and language processing would be helpful. What You Will Learn Get acquainted with some of the highly efficient R packages such as OpenNLP and RWeka to perform various steps in the text mining process Access and manipulate data from different sources such as JSON and HTTP Process text using regular expressions Get to know the different approaches of tagging texts, such as POS tagging, to get started with text analysis Explore different dimensionality reduction techniques, such as Principal Component Analysis (PCA), and understand its implementation in R Discover the underlying themes or topics that are present in an unstructured collection of documents, using common topic models such as Latent Dirichlet Allocation (LDA) Build a baseline sentence completing application Perform entity extraction and named entity recognition using R In Detail Text Mining (or text data mining or text analytics) is the process of extracting useful and high-quality information from text by devising patterns and trends. R provides an extensive ecosystem to mine text through its many frameworks and packages. Starting with basic information about the statistics concepts used in text mining, this book will teach you how to access, cleanse, and process text using the R language and will equip you with the tools and the associated knowledge about different tagging, chunking, and entailment approaches and their usage in natural language processing. Moving on, this book will teach you different dimensionality reduction techniques and their implementation in R. Next, we will cover pattern recognition in text data utilizing classification mechanisms, perform entity recognition, and develop an ontology learning framework. By the end of the book, you will develop a practical application from the concepts learned, and will understand how text mining can be leveraged to analyze the massively available data on social media. Style and approach This book takes a hands-on, example-driven approach to the text mining process with lucid implementation in R.

This book is for intermediate Solr Developers who are willing to learn and implement Pro-level practices, techniques, and solutions. This edition will specifically appeal to developers who wish to quickly get to grips with the changes and new features of Apache Solr 5.

This book will provide clear guidance on how to work through the most valuable design patterns effectively in Angular. You will explore some of the best ways to work with Angular to meet the performance required in the web development world. You will also learn the best practices to improve your productivity and the code base of your application.

This book is for Elasticsearch users who want to extend their knowledge and develop new skills. Prior knowledge of the Query DSL and data indexing is expected.

Highly illustrated and practical, Microhydro is the first complete book on the topic in many years. Covering both AC and DC systems, it first introduces the important principles on which microhydro is based, including the advantages and disadvantages of using small amounts of water to generate power. Along with a glossary of microhydro terms, further reading and resources - including websites and commerical suppliers - Microhydro includes all the information a homeowner needs to start generating clean, off-grid, and independent power.

Summary Elasticsearch in Action teaches you how to build scalable search applications using Elasticsearch. You'll ramp up fast, with an informative overview and an engaging introductory example. Within the first few chapters, you'll pick up the core concepts you need to implement basic searches and efficient indexing. With the fundamentals well in hand, you'll go on to gain an organized view of how to optimize your design. Perfect for developers and administrators building and managing search-oriented applications. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Modern search seems like magic—you type a few words and the search engine appears to know what you want. With the Elasticsearch real-time search and analytics engine, you can give your users this magical experience without having to do complex low-level programming or understand advanced data science algorithms. You just install it, tweak it, and get on with your work. About the Book Elasticsearch in Action teaches you how to write applications that deliver professional quality search. As you read, you'll learn to add basic search features to any application, enhance search results with predictive analysis and relevancy ranking, and use saved data from prior searches to give users a custom experience. This practical book focuses on Elasticsearch's REST API via HTTP. Code snippets are written mostly in bash using cURL, so they're easily translatable to other languages. What's Inside What is a great search application? Building scalable search solutions Using Elasticsearch with any language Configuration and tuning About the Reader For developers and administrators building and managing search-oriented applications. About the Authors Radu Gheorghe is a search consultant and software engineer. Matthew Lee Hinman develops highly available, cloud-based systems. Roy Russo is a specialist in predictive analytics. Table of Contents PART 1 CORE ELASTICSEARCH FUNCTIONALITY Introducing Elasticsearch Diving into the functionality Indexing, updating, and deleting data Searching your data Analyzing your data Searching with relevancy Exploring your data with aggregations Relations among documents PART 2 ADVANCED ELASTICSEARCH FUNCTIONALITY Scaling out Improving performance Administering your cluster

A comprehensive guide to mastering the most advanced Hadoop 3 concepts Key Features Get to grips with the newly introduced features and capabilities of Hadoop 3 Crunch and process data using MapReduce, YARN, and a host of tools within the Hadoop ecosystem Sharpen your Hadoop skills with real-world case studies and code Book Description Apache Hadoop is one of the most popular big data solutions for distributed storage and for processing large chunks of data. With Hadoop 3, Apache promises to provide a high-performance, more fault-tolerant, and highly efficient big data processing platform, with a focus on improved scalability and increased efficiency. With this guide, you'll understand advanced concepts of the Hadoop ecosystem tool. You'll learn how Hadoop works internally, study advanced concepts of different ecosystem tools, discover solutions to real-world use cases, and understand how to secure your cluster. It will then walk you through HDFS, YARN, MapReduce, and Hadoop 3 concepts. You'll be able to address common challenges like using Kafka efficiently, designing low latency, reliable message delivery Kafka systems, and handling high data volumes. As you advance, you'll discover how to address major challenges when building an enterprise-grade messaging system, and how to use different stream processing systems along with Kafka to fulfil your enterprise goals. By the end of this book, you'll have a complete understanding of how components in the Hadoop ecosystem are effectively integrated to implement a fast and reliable data pipeline, and you'll be equipped to tackle a range of real-world problems in data pipelines. What you will learn Gain an in-depth understanding of distributed computing using Hadoop 3 Develop enterprise-grade applications using Apache Spark, Flink, and more Build scalable and high-performance Hadoop data pipelines with security, monitoring, and data governance Explore batch data processing patterns and how to model data in Hadoop Master best practices for enterprises using, or planning to use, Hadoop 3 as a data platform Understand security aspects of Hadoop, including authorization and authentication Who this book is for If you want to become a big data professional by mastering the advanced concepts of Hadoop, this book is for you. You'll also find this book useful if you're a Hadoop professional looking to strengthen your knowledge of the Hadoop ecosystem. Fundamental knowledge of the Java programming language and basics of Hadoop is necessary to get started with this book.

The Earth-Sheltered Solar Greenhouse Book is the first to promote the benefits of both passive solar energy and earth sheltering in greenhouse design. This combination results in greenhouses which need no additional heating. The captured sun's energy and that which is stored in the earth is enough for successful year round harvest. It takes you step by step through the construction of an inexpensive greenhouse which may be built with either newly purchased or salvaged building materials for pennies on the dollar. It explains the author's unique Post/Shoring/Polyethylene construction methods and design techniques. The Earth-Sheltered Solar Greenhouse Book has 230 pages with nearly 200 illustrations, photos, diagrams, lists, charts and drawings. It contains all the information you need to free you from the pesticide, herbicide, fungicide, waxed, and E-coli laden, genetically modified and irradiated supermarket produce.

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.

Create clean code with Dependency Injection principles Key Features Use DI to make your code loosely coupled to manage and test your applications easily on Spring 5 and Google Guice Learn the best practices and methodologies to implement DI Write more maintainable Java code by decoupling your objects from their implementations Book Description Dependency Injection (DI) is a design pattern that allows us to remove the hard-coded dependencies and make our application loosely coupled, extendable, and maintainable. We can implement DI to move the dependency resolution from compile-time to runtime. This book will be your one stop guide to write loosely coupled code using the latest features of Java 9 with frameworks such as Spring 5 and Google Guice. We begin by explaining what DI is and teaching you about IoC containers. Then you'll learn about object compositions and their role in DI. You'll find out how to build a modular application and learn how to use DI to focus your efforts on the business logic unique to your application and let the framework handle the infrastructure work to put it all together. Moving on, you'll gain knowledge of Java 9's new features and modular framework and how DI works in Java 9. Next, we'll explore Spring and Guice,

the popular frameworks for DI. You'll see how to define injection keys and configure them at the framework-specific level. After that, you'll find out about the different types of scopes available in both popular frameworks. You'll see how to manage dependency of cross-cutting concerns while writing applications through aspect-oriented programming. Towards the end, you'll learn to integrate any third-party library in your DI-enabled application and explore common pitfalls and recommendations to build a solid application with the help of best practices, patterns, and anti-patterns in DI. What you will learn Understand the benefits of DI and fo from a tightly coupled design to a cleaner design organized around dependencies See Java 9's new features and modular framework Set up Guice and Spring in an application so that it can be used for DI Write integration tests for DI applications Use scopes to handle complex application scenarios Integrate any third-party library in your DI-enabled application Implement Aspect-Oriented Programming to handle common cross-cutting concerns such as logging, authentication, and transactions Understand IoC patterns and anti-patterns in DI Who this book is for This book is for Java developers who would like to implement DI in their application. Prior knowledge of the Spring and Guice frameworks and Java programming is assumed.