

Cassandra The Definitive Guide

This is likewise one of the factors by obtaining the soft documents of this **Cassandra The Definitive Guide** by online. You might not require more become old to spend to go to the books commencement as well as search for them. In some cases, you likewise pull off not discover the message Cassandra The Definitive Guide that you are looking for. It will unquestionably squander the time.

However below, in imitation of you visit this web page, it will be in view of that entirely easy to get as competently as download guide Cassandra The Definitive Guide

It will not acknowledge many period as we notify before. You can complete it even though play a role something else at home and even in your workplace. as a result easy! So, are you question? Just exercise just what we offer under as without difficulty as review **Cassandra The Definitive Guide** what you afterward to read!

Semantic Software Design Eben Hewitt 2019-09-25 With this practical book, architects, CTOs, and CIOs will learn a set of patterns for the practice of architecture, including analysis, documentation, and communication. Author Eben Hewitt shows you how to create holistic and thoughtful technology plans, communicate them clearly, lead people toward the vision, and become a great architect or Chief Architect. This book covers each key aspect of architecture comprehensively, including how to incorporate business architecture, information architecture, data architecture, application (software) architecture together to have the best chance for the system's success. Get a practical set of proven architecture practices focused on shipping great products using architecture Learn how architecture works effectively with development teams, management, and product management teams through the value chain Find updated special coverage on machine learning architecture Get usable templates to start incorporating into your teams immediately Incorporate

business architecture, information architecture, data architecture, and application (software) architecture together **Cassandra: The Definitive Guide** Eben Hewitt 2010-11-12 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

[Cassandra: The Definitive Guide, \(Revised\) Third Edition](#) Jeff Carpenter 2022-01-24 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 revised third edition--updated for Cassandra 4.0 and new developments in the Cassandra ecosystem, including deployments in Kubernetes with K8ssandra--provides technical details and practical examples to help you put this database to work in a production environment. Authors Jeff Carpenter and Eben Hewitt demonstrate the advantages of Cassandra's nonrelational design, with special attention to data modeling. Developers, DBAs, and application architects looking to solve a database scaling issue or future-proof an application will learn how to 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 Design and develop applications using client drivers Explore cluster topology and learn how nodes exchange data Maintain a high level of performance in your cluster Deploy Cassandra onsite, in the cloud, or with Docker and Kubernetes Integrate Cassandra with Spark, Kafka, Elasticsearch, Solr, and Lucene

Programming Hive Edward Capriolo 2012-09-26 Describes the features and functions of Apache Hive, the data infrastructure for Hadoop.

CouchDB: The Definitive Guide J. Chris Anderson 2010-01-19 Three of CouchDB's creators show you how to use this document-oriented database as a standalone application framework or with high-volume, distributed applications. With its simple model for storing, processing, and accessing data, CouchDB is ideal for web applications that handle huge amounts of loosely structured data. That alone would stretch the limits of a relational database, yet CouchDB offers an open source solution that's reliable, scales easily, and responds quickly. CouchDB works with self-contained data that has loose or ad-hoc connections. It's a model that fits many real-world items, such as contacts, invoices, and receipts, but you'll discover that this database can easily handle data of any kind. With this book, you'll learn how to work with CouchDB through its RESTful web interface, and become familiar with key features such as simple document CRUD (create, read, update, delete), advanced MapReduce, deployment tuning, and more. Understand the basics of document-oriented storage and manipulation Interact with CouchDB entirely through HTTP using its RESTful interface Model data as self-contained JSON documents Handle evolving data schemas naturally Query and aggregate data in CouchDB using MapReduce views Replicate data between nodes Tune CouchDB for increased performance and reliability

The JHipster Mini-Book Matt Raible 2016-11-25 The things you need to do to set up a new software project can be daunting. First, you have to select the back-end framework to create your API, choose your database, set up security, and choose your build tool. Then you have to choose the tools to create your front end: select a UI framework, configure a build tool, set up Sass processing, configure your browser to auto-refresh when you make changes, and configure the client and server so they work in unison. If you're building a new application using Spring Boot and Angular, you can save days by using JHipster. JHipster generates a complete and modern web app, unifying: - A high-

performance and robust Java stack on the server side with Spring Boot - A sleek, modern, mobile-first front-end with Angular and Bootstrap - A robust microservice architecture with the JHipster Registry, Netflix OSS, the ELK stack, and Docker - A powerful workflow to build your application with Yeoman, Webpack, and Maven/Gradle

Narrative Coaching David B. Drake 2017-12-05 REAL CHANGE IN REAL TIME--THE ULTIMATE GUIDE TO WORKING WITH PEOPLE'S STORIES IN COACHING This is a rare book; it is grounded in both a deep academic rigor and a deep personal understanding of how people change. It is a treasure chest of information and insights based in over twenty years of experience. It will enable you to get to the crux of people's issues in less time and help them make significant shifts in the moment. This book is an indispensable resource for anyone who works with people's stories and wants to develop themselves so they have more impact. The tools and models are presented in simple and clear language. However, there is a depth here that offers a limitless guide for your learning. Narrative Coaching is timely because it works at the level of identities, addresses the collective narratives that shape our stories, and expands the roles and modalities we can use to bring about transformational change with individuals and teams. What is new in this edition: It goes deeper into attachment theory and applied mindfulness It offers design thinking as a framework for adult development It shows how change is a naturally human and integrative process It offers more examples and cases, e.g., how to coach without goals This book will both challenge you and inspire you to think in new ways about what is possible in your life and in your practice.

The Fortune-Teller's Bible Jane Struthers 2007 A practical handbook designed to teach readers how to develop their divination skills looks at a wide variety of fortune-telling devices and techniques, including tarot, the I Ching, and scrying, and features skill-building exercises and instruction to teach

newcomers how to use both Western and Chinese astrology, numerology, palmistry, runes, and more to forecast the future. Original.

Cassandra High Availability Robbie Strickland 2014-12-29 If you are a developer or DevOps engineer who understands the basics of Cassandra and are ready to take your knowledge to the next level, then this book is for you. An understanding of the essentials of Cassandra is needed.

Data Pipelines with Apache Airflow Julian de Ruyter 2021-04-05 "An Airflow bible. Useful for all kinds of users, from novice to expert." - Rambabu Posa, Sai Aashika Consultancy Data Pipelines with Apache Airflow teaches you how to build and maintain effective data pipelines. A successful pipeline moves data efficiently, minimizing pauses and blockages between tasks, keeping every process along the way operational. Apache Airflow provides a single customizable environment for building and managing data pipelines, eliminating the need for a hodgepodge collection of tools, snowflake code, and homegrown processes. Using real-world scenarios and examples, Data Pipelines with Apache Airflow teaches you how to simplify and automate data pipelines, reduce operational overhead, and smoothly integrate all the technologies in your stack. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Data pipelines manage the flow of data from initial collection through consolidation, cleaning, analysis, visualization, and more. Apache Airflow provides a single platform you can use to design, implement, monitor, and maintain your pipelines. Its easy-to-use UI, plug-and-play options, and flexible Python scripting make Airflow perfect for any data management task. About the book Data Pipelines with Apache Airflow teaches you how to build and maintain effective data pipelines. You'll explore the most common usage patterns, including aggregating multiple data sources, connecting to and from data lakes, and cloud deployment. Part

reference and part tutorial, this practical guide covers every aspect of the directed acyclic graphs (DAGs) that power Airflow, and how to customize them for your pipeline's needs. What's inside Build, test, and deploy Airflow pipelines as DAGs Automate moving and transforming data Analyze historical datasets using backfilling Develop custom components Set up Airflow in production environments About the reader For DevOps, data engineers, machine learning engineers, and sysadmins with intermediate Python skills. About the author Bas Harenslak and Julian de Ruiter are data engineers with extensive experience using Airflow to develop pipelines for major companies. Bas is also an Airflow committer. Table of Contents PART 1 - GETTING STARTED 1 Meet Apache Airflow 2 Anatomy of an Airflow DAG 3 Scheduling in Airflow 4 Templating tasks using the Airflow context 5 Defining dependencies between tasks PART 2 - BEYOND THE BASICS 6 Triggering workflows 7 Communicating with external systems 8 Building custom components 9 Testing 10 Running tasks in containers PART 3 - AIRFLOW IN PRACTICE 11 Best practices 12 Operating Airflow in production 13 Securing Airflow 14 Project: Finding the fastest way to get around NYC PART 4 - IN THE CLOUDS 15 Airflow in the clouds 16 Airflow on AWS 17 Airflow on Azure 18 Airflow in GCP

Cassandra ji shu shou ce □□□ 2017

Database Internals Alex Petrov 2019-09-13 When it comes to choosing, using, and maintaining a database, understanding its internals is essential. But with so many distributed databases and tools available today, it's often difficult to understand what each one offers and how they differ. With this practical guide, Alex Petrov guides developers through the concepts behind modern database and storage engine internals. Throughout the book, you'll explore relevant material gleaned from numerous books, papers, blog posts, and the source code of several open source databases. These resources are listed at the end of parts one and two. You'll discover that the most significant distinctions among

many modern databases reside in subsystems that determine how storage is organized and how data is distributed. This book examines: Storage engines: Explore storage classification and taxonomy, and dive into B-Tree-based and immutable Log Structured storage engines, with differences and use-cases for each Storage building blocks: Learn how database files are organized to build efficient storage, using auxiliary data structures such as Page Cache, Buffer Pool and Write-Ahead Log Distributed systems: Learn step-by-step how nodes and processes connect and build complex communication patterns Database clusters: Which consistency models are commonly used by modern databases and how distributed storage systems achieve consistency

Ant: The Definitive Guide Steve Holzner 2005-04-13 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's been writing about Java topics since the language first appeared; his books have sold more than 1.5 million copies worldwide.

Routledge Handbook of Applied Sport Psychology Stephanie J. Hanrahan 2010-10-18 Now available in paperback, the Routledge Handbook of Applied Sport Psychology is a definitive guide to the theory and practice of applied sport psychology. It goes further than any other book in surveying the full variety of issues that practising sport psychologists will confront in their working lives. It introduces the most important tools and skills that psychologists will need to be truly helpful to their clients, and it also adopts a holistic definition of the role of the sport psychologist, explaining how effective counseling, assessment, and therapeutic models can add important extra dimensions to professional practice. The book is divided into seven thematic sections, addressing: counseling; assessment; theoretical and therapeutic models; issues for the individual athlete, from injury and overtraining to depression; issues for teams, from conflict resolution to travel; working with special populations; mental skills, such as imagery, goal setting, and concentration. Moving beyond the traditional tracks of clinical psychology and performance enhancement, the authors in this book argue convincingly that psychologists would benefit from attempting to understand athletes' social and familial contexts, their health, happiness, and interpersonal dynamics in the broadest sense, if they are to serve their clients' best interests. With contributions from many of the world's leading sport psychologists, and with clear descriptions of best practice in each chapter, the Routledge Handbook of Applied Sport Psychology is essential reading for all serious students and practitioners of sport psychology,

counseling, applied sport science, health psychology, and related fields.

Cassandra Jeff Carpenter 2016-07-08 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 **Trino: The Definitive Guide** Matt Fuller 2021-04-14 Perform fast interactive analytics against different data sources using the Trino high-performance distributed SQL query engine. With this practical guide, you'll learn how to conduct analytics on data where it lives, whether it's Hive, Cassandra, a relational database, or a proprietary data store. Analysts, software engineers, and production engineers will learn how to manage, use, and even develop with Trino. Initially developed by Facebook, open source Trino is now used by Netflix, Airbnb, LinkedIn, Twitter, Uber, and many other companies. Matt Fuller,

Manfred Moser, and Martin Traverso show you how a single Trino query can combine data from multiple sources to allow for analytics across your entire organization. Get started: Explore Trino's use cases and learn about tools that will help you connect to Trino and query data Go deeper: Learn Trino's internal workings, including how to connect to and query data sources with support for SQL statements, operators, functions, and more Put Trino in production: Secure Trino, monitor workloads, tune queries, and connect more applications; learn how other organizations apply Trino

Spark: The Definitive Guide Bill Chambers 2018-02-08 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

The Tai Chi Bible D J DOCHERTY Ltd 2017-05-25 The complete illustrated guide to Tai Chi for beginners to advanced level.

Cassandra: The Definitive Guide Jeff Carpenter 2016-06-29 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

Cassandra Jeff Carpenter 2022 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 revised third edition--updated for Cassandra 4.0 and new developments in the Cassandra ecosystem, including deployments in Kubernetes with K8ssandra--provides technical details and practical examples to help you put this database to work in a production environment. Authors Jeff Carpenter and Eben Hewitt demonstrate the advantages of Cassandra's nonrelational design, with special attention to data modeling. Developers, DBAs, and application architects looking to

solve a database scaling issue or future-proof an application will learn how to 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 Design and develop applications using client drivers Explore cluster topology and learn how nodes exchange data Maintain a high level of performance in your cluster Deploy Cassandra onsite, in the cloud, or with Docker and Kubernetes Integrate Cassandra with Spark, Kafka, Elasticsearch, Solr, and Lucene.

The Reflexology Bible Louise Keet 2009 Find out how to give the healing touch! Reflexology is a safe and gentle therapy that combats stress, boosts the immune system, and stimulates our natural healing processes. From zone maps of the hands and feet to techniques for working the reflex points to illustrated step-by-step sequences, this guide provides all the information necessary for treating the whole body. A detailed directory provides instructions on relieving common ailments, and there are specialized treatments for children, couples, and the elderly.

Jenkins: The Definitive Guide John Ferguson Smart 2011-07-12 Streamline software development with Jenkins, the popular Java-based open source tool that has revolutionized the way teams think about Continuous Integration (CI). This complete guide shows you how to automate your build, integration, release, and deployment processes with Jenkins—and demonstrates how CI can save you time, money, and many headaches. Ideal for developers, software architects, and project managers, Jenkins: The Definitive Guide is both a CI tutorial and a comprehensive Jenkins reference. Through its wealth of best practices and real-world tips, you'll discover how easy it is to set up a CI service with Jenkins. Learn how to install, configure, and secure your Jenkins server Organize and monitor general-purpose build jobs Integrate automated tests to verify builds, and set up code quality

reporting Establish effective team notification strategies and techniques Configure build pipelines, parameterized jobs, matrix builds, and other advanced jobs Manage a farm of Jenkins servers to run distributed builds Implement automated deployment and continuous delivery

Cassandra Jeff Carpenter. Eben Hewitt 2016

Cassandra Admin DBA Guide Prasad Bagewadi 2019-08-15 What value this book provides? This book absolutely provides tremendous value in terms its usefulness. This book takes away the pain associated with learning and mastering Cassandra. All complexity has been digested by the author and simplified for the reader with very useful and practical information that can be absorbed just by glancing through the pages. Years of author's experience and practical knowledge in Cassandra has been gifted to the reader in this book with great diligence and generosity. If you are planning to undergo expensive Cassandra training, think again, spending few hours with this book will change your mind, this book has been written with great care to reduce the learning curve. The aim of this book is multi fold, just to mention a few: Provide confidence to the reader in Cassandra concepts and architecture Provide a flexible, practical framework and context customizable for various situations Provide practical guidance to manage Cassandra platforms of various hues, sizes, shapes Provide real world examples to reduce guess work Provide executable query statements and command line statements at every step Provide practical outcomes to help the reader to gain instance understanding of what to expect Provide insights into making Cassandra environment robust and scalable Provide tricks and tips to implement and maintain seamlessly Provide security and vulnerability mitigation tips and steps Provide best practices to follow for optimal Cassandra use There is no doubt - this book makes the reader very productive Cassandra professional in very short span of time. This book essentially bridges the training gap as the industry is moving fast to take full

advantage of what Cassandra can offer to fulfill emerging business needs. This book can be very helpful to Data administrators, Developers, Data modelers/Application Architects, Project Managers and Data Custodians. This book has range of topics, many are listed here: Cassandra concepts and architecture Cassandra Installation and Configuration Cassandra system architecture depicting gossip protocol, replication, consistency, tombstones, hinted handoff, compactions, repairs, memtables, commit log, read and write path functions Cassandra oriented data modelling Cassandra QL (CQL) tutorial Handling of Primary and Partition keys in Cassandra covering No joins, Static columns and TTL aspects Configuring authentication, authorization to access Cassandra in addition, steps to set up node-node and client-node SSL Configuring nodes addition, removal, decommission in single token and vnode setup modes in Cassandra Instructions to add new data center and delink the existing data center from a multi-dc cluster arrangement Cassandra backup and recovery functions with real examples of restoring tables after truncation events Cassandra utilities such as sstabledump, sstablemetadata, sstablesplit, cqlsh and cassandra-stress Troubleshooting methods such as Node down, Read latency and Recover truncated table Upgrading Cassandra to higher versions Additional Cassandra architecture II methods such as Read and Write path, Compactions and Repairs *The Book of Esports* William Collis 2020-08-04 The definitive guide to the modern world of competitive gaming and the official history of Esports™. Almost overnight, esports—or competitive video games—have exploded into the largest entertainment and sporting phenomenon in human history. The Book of Esports answers: What exactly are esports, and how did they become so popular so quickly? Why did blockbuster video games like League of Legends, Fortnite and Starcraft succeed? Where exactly is all this video gaming headed? What do gamers and college students need to know to position themselves for success in the industry?

How do you create a billion-dollar esports business? What strategic choices drive success in the modern gaming industry? Can video games really get your kid into college? (All expenses paid, of course...) Whether you are a lifelong gamer, a curious Fortnite parent, or a businessperson seeking to understand the marketing opportunities of this multibillion-dollar phenomenon, *The Book of Esports* charts the rise of this exciting new industry, for the first time ever crafting a comprehensive overview of esports and its implications for human competition—and even the future of humanity itself. Gaming luminary and Harvard MBA William Collis has painstakingly translated esports' mysteries into a detailed and accessible testament for today. Featuring select interviews from the biggest names in the industry, *The Book of Esports* weaves tales of trust, betrayal, and superhuman reflexes into predictive frameworks, explaining exactly why our industry looks the way it does, and how all this growth—and more—is inevitable as the divide between man and machine blurs into oblivion.

Kafka: The Definitive Guide Neha Narkhede 2017-08-31 Every enterprise application creates data, whether it's log messages, metrics, user activity, outgoing messages, or something else. And how to move all of this data becomes nearly as important as the data itself. If you're an application architect, developer, or production engineer new to Apache Kafka, this practical guide shows you how to use this open source streaming platform to handle real-time data feeds. Engineers from Confluent and LinkedIn who are responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications with this platform. Through detailed examples, you'll learn Kafka's design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the controller, and the storage layer. Understand publish-subscribe messaging and how it fits in the big data ecosystem. Explore

Kafka producers and consumers for writing and reading messages Understand Kafka patterns and use-case requirements to ensure reliable data delivery Get best practices for building data pipelines and applications with Kafka Manage Kafka in production, and learn to perform monitoring, tuning, and maintenance tasks Learn the most critical metrics among Kafka's operational measurements Explore how Kafka's stream delivery capabilities make it a perfect source for stream processing systems

Google BigQuery: The Definitive Guide Valliappa Lakshmanan 2019-10-23 Work with petabyte-scale datasets while building a collaborative, agile workplace in the process. This practical book is the canonical reference to Google BigQuery, the query engine that lets you conduct interactive analysis of large datasets. BigQuery enables enterprises to efficiently store, query, ingest, and learn from their data in a convenient framework. With this book, you'll examine how to analyze data at scale to derive insights from large datasets efficiently. Valliappa Lakshmanan, tech lead for Google Cloud Platform, and Jordan Tigani, engineering director for the BigQuery team, provide best practices for modern data warehousing within an autoscaled, serverless public cloud. Whether you want to explore parts of BigQuery you're not familiar with or prefer to focus on specific tasks, this reference is indispensable.

Programming Pig Alan Gates 2011-10-06 This guide is an ideal learning tool and reference for Apache Pig, the programming language that helps programmers describe and run large data projects on Hadoop. With Pig, they can analyze data without having to create a full-fledged application--making it easy for them to experiment with new data sets.

Cassandra: The Definitive Guide Jeff Carpenter 2016-06-29 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 **Cassandra: the Definitive Guide, (Revised) Third Edition** Jeff Carpenter 2022-02 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 revised third edition--updated for Cassandra 4.0 and new developments in the Cassandra ecosystem, including deployments in Kubernetes with K8ssandra--provides technical details and practical examples to help you put this database to work in a production environment. Authors Jeff Carpenter and Eben Hewitt demonstrate the advantages of Cassandra's nonrelational design, with special attention to data modeling. Developers, DBAs, and application architects looking to solve a database scaling issue or future-proof an application will learn how to 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
Design and develop applications using client drivers
Explore cluster topology and learn how nodes exchange data
Maintain a high level of performance in your cluster
Deploy Cassandra onsite, in the cloud, or with Docker and Kubernetes
Integrate Cassandra with Spark, Kafka, Elasticsearch, Solr, and Lucene

Practical Cassandra Russell Bradberry 2013-12-17 "Eric and Russell were early adopters of Cassandra at SimpleReach. In Practical Cassandra, you benefit from their experience in the trenches administering Cassandra, developing against it, and building one of the first CQL drivers. If you are deploying Cassandra soon, or you inherited a Cassandra cluster to tend, spend some time with the deployment, performance tuning, and maintenance chapters... If you are new to Cassandra, I highly recommend the chapters on data modeling and CQL." -From the Foreword by Jonathon Ellis, Apache Cassandra Chair
Build and Deploy Massively Scalable, Super-fast Data Management Applications with Apache Cassandra
Practical Cassandra is the first hands-on developer's guide to building Cassandra systems and applications that deliver breakthrough speed, scalability, reliability, and performance. Fully up to date, it reflects the latest versions of Cassandra—including Cassandra Query Language (CQL), which dramatically lowers the learning curve for Cassandra developers. Pioneering Cassandra developers and Datastax MVPs Russell Bradberry and Eric Lubow walk you through every step of building a real production application that can store enormous amounts of structured, semi-structured, and unstructured data. Drawing on their exceptional expertise, Bradberry and Lubow share practical insights into issues ranging from querying to deployment, management, maintenance, monitoring, and troubleshooting. The authors cover key issues, from architecture to migration, and guide you through crucial

decisions about configuration and data modeling. They provide tested sample code, detailed explanations of how Cassandra works "under the covers," and new case studies from three cutting-edge users: Ooyala, Hailo, and eBay. Coverage includes Understanding Cassandra's approach, architecture, key concepts, and primary use cases— and why it's so blazingly fast
Getting Cassandra up and running on single nodes and large clusters
Applying the new design patterns, philosophies, and features that make Cassandra such a powerful data store
Leveraging CQL to simplify your transition from SQL-based RDBMSes
Deploying and provisioning through the cloud or on bare-metal hardware
Choosing the right configuration options for each type of workload
Tweaking Cassandra to get maximum performance from your hardware, OS, and JVM
Mastering Cassandra's essential tools for maintenance and monitoring
Efficiently solving the most common problems with Cassandra deployment, operation, and application development

Learning Spark Jules S. Damji 2020-07-16 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

The Definitive Guide to MongoDB Eelco Plugge 2015-12-11
The Definitive Guide to MongoDB, Third Edition, is updated for MongoDB 3 and includes all of the latest MongoDB features, including the aggregation framework introduced in version 2.2 and hashed indexes in version 2.4. The Third Edition also now includes Python. MongoDB is the most popular of the "Big Data" NoSQL database technologies, and it's still growing. David Hows from 10gen, along with experienced MongoDB authors Peter Membrey and Eelco Plugge, provide their expertise and experience in teaching you everything you need to know to become a MongoDB pro.

Technology Strategy Patterns Eben Hewitt 2018-10-15
Technologists who want their ideas heard, understood, and funded are often told to speak the language of business—without really knowing what that is. This book's toolkit provides architects, product managers, technology managers, and executives with a shared language—in the form of repeatable, practical patterns and templates—to produce great technology strategies. Author Eben Hewitt developed 39 patterns over the course of a decade in his work as CTO, CIO, and chief architect for several global tech companies. With these proven tools, you can define, create, elaborate, refine, and communicate your architecture goals, plans, and approach in a way that executives can readily understand, approve, and execute. This book covers:
Architecture and strategy: Adopt a strategic architectural mindset to make a meaningful material impact
Creating your strategy: Define the components of your technology strategy using proven patterns
Communicating the strategy: Convey your technology strategy in a compelling way to a variety of audiences
Bringing it all together: Employ patterns individually or in clusters for specific problems; use the complete framework for a comprehensive strategy

Cassandra Kieran L. Lee 2015-08-12 Thought-provoking and

accessible in approach, this updated and expanded second edition of the *Cassandra: The Definitive Guide* provides a user-friendly introduction to the subject, Taking a clear structural framework, it guides the reader through the subject's core elements. A flowing writing style combines with the use of illustrations and diagrams throughout the text to ensure the reader understands even the most complex of concepts. This succinct and enlightening overview is a required reading for advanced graduate-level students. We hope you find this book useful in shaping your future career. Feel free to send us your enquiries related to our publications to info@risepress.pw Rise Press
Cassandra: the Definitive Guide Jeff Carpenter 2020-07-31
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 third edition--updated for Cassandra 4.0--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 nonrelational 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
Weight Loss Burton Goldberg 2000 Explains how to use enzymes, aromatherapy, diet, nutrition, Ayurvedic medicine, glandular extracts, exercise, and supplements to help lose weight permanently.

Cassandra: the Definitive Guide Louie Bryant 2015-08-19
Cassandra High Performance Cookbook Edward Capriolo

2011-07-15 Over 150 recipes to design and optimize large scale
Apache Cassandra deployments.
CASSANDRA EBEN HEWITT. JEFF CARPENTER 2016