
Apache Kafka

When somebody should go to the books stores, search commencement by shop, shelf by shelf, it is truly problematic. This is why we offer the book compilations in this website. It will unquestionably ease you to look guide **Apache Kafka** as you such as.

By searching the title, publisher, or authors of guide you essentially want, you can discover them rapidly. In the house, workplace, or perhaps in your method can be every best place within net connections. If you intend to download and install the Apache Kafka, it is unconditionally easy then, back currently we extend the colleague to purchase and make bargains to download and install Apache Kafka hence simple!

Apache Kafka

2022-10-03

MCDOWELL KAYDEN

Kafka Streams - Real-time Stream Processing Springer Nature

Gain deep insight into real-time analytics, including the features of these systems and the problems they solve. With this practical book, data engineers at organizations that use event-processing systems such as Kafka, Google Pub/Sub, and AWS Kinesis will learn how to analyze data streams in real time. The faster you derive insights, the quicker you can spot changes in your business and act accordingly. Author Mark Needham from StarTree provides an overview of the real-time analytics space and an understanding

of what goes into building real-time applications. The book's second part offers a series of hands-on tutorials that show you how to combine multiple software products to build real-time analytics applications for an imaginary pizza delivery service. You will: Learn common architectures for real-time analytics Discover how event processing differs from real-time analytics Ingest event data from Apache Kafka into Apache Pinot Combine event streams with OLTP data using Debezium and Kafka Streams Write real-time queries against event data stored in Apache Pinot Build a real-time dashboard and order tracking app Learn how Uber, Stripe, and Just Eat use real-time analytics

Apache Kafka Cybellium Ltd

Process large volumes of data in real-time while building high performance and robust data stream processing pipeline using the latest Apache Kafka 2.0 Key Features Solve practical large data and processing challenges with Kafka Tackle data processing challenges like late events, windowing, and watermarking Understand real-time streaming applications processing using Schema registry, Kafka connect, Kafka streams, and KSQL Book Description Apache Kafka is a great open source platform for handling your real-time data pipeline to ensure high-speed filtering and pattern matching on the fly. In this book, you will learn how to use Apache Kafka for

efficient processing of distributed applications and will get familiar with solving everyday problems in fast data and processing pipelines. This book focuses on programming rather than the configuration management of Kafka clusters or DevOps. It starts off with the installation and setting up the development environment, before quickly moving on to performing fundamental messaging operations such as validation and enrichment. Here you will learn about message composition with pure Kafka API and Kafka Streams. You will look into the transformation of messages in different formats, such as text, binary, XML, JSON, and AVRO. Next, you will learn how to expose the schemas contained in Kafka with the Schema Registry. You will then learn how to work with all relevant connectors with Kafka Connect. While working with Kafka Streams, you will perform various interesting operations on streams, such as windowing, joins, and aggregations. Finally, through KSQL, you will learn how to retrieve, insert, modify, and delete data streams, and how to manipulate watermarks and windows. What you will learn

with Kafka
Add information to existing data flows
Generate new information through message composition
Perform data validation and versioning with the Schema Registry
How to perform message Serialization and Deserialization
How to perform message Serialization and Deserialization
Process data streams with Kafka Streams
Understand the duality between tables and streams with KSQL
Who this book is for
This book is for developers who want to quickly master the practical concepts behind Apache Kafka. The audience need not have come across Apache Kafka previously; however, a familiarity of Java or any JVM language will be helpful in understanding the code in this book.

Learning Apache OpenWhisk jideon francisco marques
Gain deep insight into real-time analytics, including the features of these systems and the problems they solve. With this practical book, data engineers at organizations that use event-processing systems such as Kafka, Google Pub/Sub, and AWS Kinesis will learn how to analyze data streams in real time. The faster you derive insights, the quicker you can spot

changes in your business and act accordingly. Author Mark Needham from StarTree provides an overview of the real-time analytics space and an understanding of what goes into building real-time applications. The book's second part offers a series of hands-on tutorials that show you how to combine multiple software products to build real-time analytics applications for an imaginary pizza delivery service. You will:
Learn common architectures for real-time analytics
Discover how event processing differs from real-time analytics
Ingest event data from Apache Kafka into Apache Pinot
Combine event streams with OLTP data using Debezium and Kafka Streams
Write real-time queries against event data stored in Apache Pinot
Build a real-time dashboard and order tracking app
Learn how Uber, Stripe, and Just Eat use real-time analytics
[Building Data Streaming Applications with Apache Kafka](#) Springer
Summary
Kafka Streams in Action teaches you everything you need to know to implement stream processing on data flowing into your Kafka platform, allowing you to focus on getting more from your

data without sacrificing time or effort. Foreword by Neha Narkhede, Cocreator of Apache Kafka Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Not all stream-based applications require a dedicated processing cluster. The lightweight Kafka Streams library provides exactly the power and simplicity you need for message handling in microservices and real-time event processing. With the Kafka Streams API, you filter and transform data streams with just Kafka and your application. About the Book Kafka Streams in Action teaches you to implement stream processing within the Kafka platform. In this easy-to-follow book, you'll explore real-world examples to collect, transform, and aggregate data, work with multiple processors, and handle real-time events. You'll even dive into streaming SQL with KSQL! Practical to the very end, it finishes with testing and operational aspects, such as monitoring and debugging. What's inside Using the KStreams API Filtering, transforming, and splitting data Working with the Processor API Integrating with external systems About the Reader

Assumes some experience with distributed systems. No knowledge of Kafka or streaming applications required. About the Author Bill Bejeck is a Kafka Streams contributor and Confluent engineer with over 15 years of software development experience. Table of Contents PART 1 - GETTING STARTED WITH KAFKA STREAMS Welcome to Kafka Streams Kafka quicklyPART 2 - KAFKA STREAMS DEVELOPMENT Developing Kafka Streams Streams and state The KTable API The Processor APIPART 3 - ADMINISTERING KAFKA STREAMS Monitoring and performance Testing a Kafka Streams applicationPART 4 - ADVANCED CONCEPTS WITH KAFKA STREAMS Advanced applications with Kafka StreamsAPPENDIXES Appendix A - Additional configuration information Appendix B - Exactly once semantics **Apache Kafka** Network Automation Nerds, LLC.

How this book is organized: A roadmap I organized this book into three sections with 13 chapters. Part 1 introduces IaC and how you, as an individual, write it. • Chapter 1 defines IaC and its benefits and principles. The chapter explains that the

book has examples in Python, run by HashiCorp Terraform, and deployed to Google Cloud Platform (GCP). I also discuss the tools and use cases you'll encounter in your IaC journey. • Chapter 2 dives into the principle of immutability and how you can migrate existing infrastructure resources to IaC. It also covers the practices of writing clean IaC. • Chapter 3 offers a few patterns for dividing and grouping infrastructure resources into modules. Each pattern includes an example and a list of use cases. • Chapter 4 covers how to manage dependencies among infrastructure resources and modules and decouple them with dependency injection and some common patterns. Part 2 describes how to write and collaborate on IaC as a team. • Chapter 5 organizes the practices and considerations for expressing IaC in different repository structures and sharing it across your team. • Chapter 6 provides an infrastructure testing strategy. It describes each type of test and how to write them for IaC. • Chapter 7 applies continuous delivery to IaC. It covers a high-level view of branching models and how your team can use them to change infrastructure. •

Chapter 8 provides techniques to build secure and compliant IaC, including testing and tagging. Part 3 covers how to manage IaC across your company. • Chapter 9 applies immutability to infrastructure changes, including an example for blue-green deployments. • Chapter 10 refactors a large body of IaC to improve its maintainability and mitigate the blast radius of failed changes to one codebase. • Chapter 11 describes reverting IaC and rolling forward changes to the system. • Chapter 12 addresses the use of IaC to manage cloud computing costs. It includes an example for cost estimation of IaC. • Chapter 13 completes the book with practices to manage and update IaC tools. You will find that many concepts build on each other throughout the book, and it may help to read the chapters in order if you have not previously practiced IaC. Otherwise, you can choose the sections that best apply to the challenges you face in your IaC practice.

I Heart Logs "O'Reilly Media, Inc."

This book, the third one of three volumes, focuses on data and the actions around data, like storage and processing. The

angle shifts over the volumes from a business-driven approach in “Disruption and DNA” to a strong technical focus in “Data Storage, Processing and Analysis”, leaving “Digitalization and Machine Learning Applications” with the business and technical aspects in-between. In the last volume of the series, “Data Storage, Processing and Analysis”, the shifts in the way we deal with data are addressed.

Apache Kafka Quick Start Guide Packt Publishing Ltd

Build a platform using Apache Kafka, Spark, and Storm to generate real-time data insights and view them through Dashboards. **KEY FEATURES** ● Extensive practical demonstration of Apache Kafka concepts, including producer and consumer examples. ● Includes graphical examples and explanations of implementing Kafka Producer and Kafka Consumer commands and methods. ● Covers integration and implementation of Spark-Kafka and Kafka-Storm architectures. **DESCRIPTION** Real-Time Streaming with Apache Kafka, Spark, and Storm is a book that provides an overview of the real-time streaming concepts and architectures of Apache Kafka, Storm, and

Spark. The readers will learn how to build systems that can process data streams in real time using these technologies. They will be able to process a large amount of real-time data and perform analytics or generate insights as a result of this. The architecture of Kafka and its various components are described in detail. A Kafka Cluster installation and configuration will be demonstrated. The Kafka publisher-subscriber system will be implemented in the Eclipse IDE using the Command Line and Java. The book discusses the architecture of Apache Storm, the concepts of Spout and Bolt, as well as their applications in a Transaction Alert System. It also describes Spark's core concepts, applications, and the use of Spark to implement a microservice. To learn about the process of integrating Kafka and Storm, two approaches to Spark and Kafka integration will be discussed. This book will assist a software engineer to transition to a Big Data engineer and Big Data architect by providing knowledge of big data processing and the architectures of Kafka, Storm, and Spark Streaming. **WHAT YOU WILL LEARN** ● Creation of Kafka producers, consumers, and brokers

using command line. ● End-to-end implementation of Kafka messaging system with Java in Eclipse. ● Perform installation and creation of a Storm Cluster and execute Storm Management commands. ● Implement Spouts, Bolts and a Topology in Storm for Transaction alert application system. ● Perform the implementation of a microservice using Spark in Scala IDE. ● Learn about the various approaches of integrating Kafka and Spark. ● Perform integration of Kafka and Storm using Java in the Eclipse IDE.

WHO THIS BOOK IS FOR This book is intended for Software Developers, Data Scientists, and Big Data Architects who want to build software systems to process data streams in real time. To understand the concepts in this book, knowledge of any programming language such as Java, Python, etc. is needed.

TABLE OF CONTENTS

1. Introduction to Kafka
2. Installing Kafka
3. Kafka Messaging
4. Kafka Producers
5. Kafka Consumers
6. Introduction to Storm
7. Installation and Configuration
8. Spouts and Bolts
9. Introduction to Spark
10. Spark Streaming
11. Kafka Integration with Storm
12. Kafka Integration with Spark

Encyclopedia of Big Data

Technologies Simon and Schuster

Today's network is about agility, automation, and continuous improvement. In *Kafka Up and Running for Network DevOps*, we will be on a journey to learn and set up the hugely popular Apache Kafka data messaging system. Kafka is unique in its principle to treat network data as a continuous flow of information that can adapt to the ever-changing business requirements. Whether you need a system to aggregate log messages, collect metrics, or something else, Kafka can be the reliable, highly redundant system you want. We will begin by learning about the core concepts of Kafka, followed by detailed steps of setting up a Kafka system in a lab environment. For the production environment, we will take advantage of the various public cloud provider offerings. Next, we will set up our Kafka cluster in Amazon Managed Kafka Service to host our Kafka cluster in the AWS cloud. We will also learn about AWS Kinesis, Azure Event Hub, and Google Cloud Pub/Sub. Finally, the book will illustrate several use cases of how to integrate Kafka with our network from

data enhancement, monitoring, to an event-driven architecture. The *Network DevOps Series* is a series of books targeted for the next generation of Network Engineers who wants to take advantage of the powerful tools and projects in modern software development and the open-source communities.

[Kafka Streams in Action, Second Edition](#)
Packt Publishing Ltd

Every enterprise application creates data, whether it consists of log messages, metrics, user activity, outgoing messages, or something else. Moving all of this data is just as important as the data itself. This book's updated second edition shows application architects, developers, and production engineers new to the Kafka open source streaming platform how to handle real-time data feeds. Additional chapters cover Kafka's AdminClient API, new security features, and tooling changes. Engineers from Confluent and LinkedIn 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. You'll examine: How publish-subscribe messaging fits in the big data ecosystem Kafka producers and consumers for writing and reading messages Patterns and use-case requirements to ensure reliable data delivery Best practices for building data pipelines and applications with Kafka How to perform monitoring, tuning, and maintenance tasks with Kafka in production The most critical metrics among Kafka's operational measurements Kafka's delivery capabilities for stream processing systems *Architecting Data Intensive Applications* Vintage

In diesem shortcut geht es um Apache Kafka, den verteilten, partitionierenden und replizierenden Service für Datenströme. Kapitel 1 stellt die Konzepte vor, mithilfe derer Apache Kafka seine Funktionen zur Verfügung stellt, und erläutert deren effektive Performance. Im zweiten Kapitel geht es um die Abfrage und Verwaltung aktueller und historischer Definitionen von Datenstrukturen mithilfe

von Schema Registry. In diesem Zusammenhang werden Avro-Schemas erläutert und es wird auf die Vorteile einer zentralen Schemaverwaltung eingegangen. Kapitel 3 widmet sich dem Kafka-REST-Proxy, der eine zusätzliche Möglichkeit bietet, über HTTP/REST mit einem Kafka-Cluster zu interagieren. Neben den Funktionen dieses Proxys demonstrieren die Autoren anhand einer Beispielanwendung den Umgang mit dem REST-API.

Kafka on the Shore O'Reilly Media Discover the real power of Spring Framework 5.0 and learn to create powerful applications in its newest version Key Features Learn reactive programming by implementing a reactive application with Spring Webflux Create a robust and scalable messaging application with Spring messaging support Apply your knowledge to build three real-world projects in Spring Book Description With growing demands, organizations are looking for systems that are robust and scalable. Therefore, the Spring Framework has become the most popular framework for Java development. It not only simplifies software development but also improves developer

productivity. This book covers effective ways to develop robust applications in Java using Spring. The book has three parts, where each one covers the building of a comprehensive project in Java and Spring. In the first part, you will construct a CMS Portal using Spring's support for building REST APIs. You will also learn to integrate these APIs with AngularJS and later develop this application in a reactive fashion using Project Reactor, Spring WebFlux, and Spring Data. In the second part, you'll understand how to build a messaging application, which will consume the Twitter API and perform filtering and transformations. Here, you will also learn about server-sent events and explore Spring's support for Kotlin, which makes application development quick and efficient. In the last part, you will build a real microservice application using the most important techniques and patterns such as service discovery, circuit breakers, security, data streams, monitoring, and a lot more from this architectural style. By the end of the book, you will be confident about using Spring to build your applications. What you will learn Implement REST APIs with Spring REST

support Introduce the Spring Boot and understand how it makes creating robust applications extremely simple Understand how Spring Data helps us add persistence in MongoDB and SQL databases Introduce Reactive Programming and use this with Spring Webflux Implement a Reactive REST client and learn how it can create asynchronous applications Create a robust, scalable, and fault tolerant application with Spring Messaging Implement a websocket to add interactive behaviors in your applications Introduce the Spring Cloud projects Who this book is for If you're a developer starting out with Spring, then this book will help you learn about the new Spring 5.0 framework concepts followed by their implementation in Java and Kotlin. The book will also help experienced Spring developers gain insights into the new features added in Spring 5.0.

Kafka in Action Simon and Schuster The Encyclopedia of Big Data Technologies provides researchers, educators, students and industry professionals with a comprehensive authority over the most relevant Big Data Technology concepts. With over 300 articles written by

worldwide subject matter experts from both industry and academia, the encyclopedia covers topics such as big data storage systems, NoSQL database, cloud computing, distributed systems, data processing, data management, machine learning and social technologies, data science. Each peer-reviewed, highly structured entry provides the reader with basic terminology, subject overviews, key research results, application examples, future directions, cross references and a bibliography. The entries are expository and tutorial, making this reference a practical resource for students, academics, or professionals. In addition, the distinguished, international editorial board of the encyclopedia consists of well-respected scholars, each developing topics based upon their expertise.

Effective Kafka entwickler.Press The book will follow a step-by-step tutorial approach which will show the readers how to use Apache Kafka for messaging from scratch. Apache Kafka is for readers with software development experience, but no prior exposure to Apache Kafka or similar technologies is assumed. This book is also for enterprise application developers and

big data enthusiasts who have worked with other publisher-subscriber based systems and now want to explore Apache Kafka as a futuristic scalable solution. *Apache Kafka* "O'Reilly Media, Inc." Everything you need to implement stream processing on Apache Kafka? using Kafka Streams and the ksqldb event streaming database. This totally revised new edition of Kafka Streams in Action has been expanded to cover more of the Kafka platform used for building event-based applications. You'll also find full coverage of ksqldb, an event streaming database purpose-built for stream processing applications. In Kafka Streams in Action, Second Edition you'll learn how to: Design streaming applications in Kafka Streams with the KStream and the Processor API Integrate external systems with Kafka Connect Enforce data compatibility with Schema Registry Build applications that respond immediately to events in either Kafka Streams or ksqldb Craft materialized views over streams with ksqldb Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology The lightweight Kafka

Streams library provides exactly the power and simplicity you need for event-based applications, real-time event processing, and message handling in microservices. The KSQL database makes it a snap to create applications that respond immediately to events, such as real-time push and pull updates. About the book *Kafka Streams in Action, Second Edition* teaches you to implement stream processing within the Kafka platform. In this easy-to-follow book, you'll explore real-world examples to collect, transform, and aggregate data, work with multiple processors, and handle real-time events. You'll also dive into processing event data with KSQL. Practical to the very end, it finishes with testing and operational aspects, such as monitoring, debugging, and gives you the opportunity to explore a few end-to-end projects. About the reader Assumes experience with building Java applications, concepts like threading, serialization, and with distributed systems. No knowledge of Kafka or streaming applications required. About the author Bill Bejeck is a Confluent engineer and a Kafka Streams contributor with over 15 years of software development experience. Bill is

also a committer on the Apache Kafka project.

Kafka Streams in Action "O'Reilly Media, Inc."

What separates the traditional enterprise from the likes of Amazon, Netflix, and Etsy? Those companies have refined the art of cloud native development to maintain their competitive edge and stay well ahead of the competition. This practical guide shows Java/JVM developers how to build better software, faster, using Spring Boot, Spring Cloud, and Cloud Foundry. Many organizations have already waded into cloud computing, test-driven development, microservices, and continuous integration and delivery. Authors Josh Long and Kenny Bastani fully immerse you in the tools and methodologies that will help you transform your legacy application into one that is genuinely cloud native. In four sections, this book takes you through: The Basics: learn the motivations behind cloud native thinking; configure and test a Spring Boot application; and move your legacy application to the cloud Web Services: build HTTP and RESTful services with Spring; route requests in your distributed

system; and build edge services closer to the data Data Integration: manage your data with Spring Data, and integrate distributed services with Spring's support for event-driven, messaging-centric architectures Production: make your system observable; use service brokers to connect stateful services; and understand the big ideas behind continuous delivery *Big Data SMACK* Packt Pub Limited Introduce your children to the whimsical existential surrealism of everyone's favourite German modernist writer! For real, though - Kafka's writing is surprisingly well suited for young children, as this short illustrated collection perfectly, well, illustrates. [Real-Time Streaming with Apache Kafka, Spark, and Storm](#) Apress The software architecture landscape has evolved dramatically over the past decade. Microservices have displaced monoliths. Data and applications are increasingly becoming distributed and decentralised. But composing disparate systems is a hard problem. More recently, software practitioners have been rapidly converging on event-driven architecture as a sustainable way of dealing with

complexity - integrating systems without increasing their coupling. In *Effective Kafka*, Emil Koutanov explores the fundamentals of Event-Driven Architecture - using Apache Kafka - the world's most popular and supported open-source event streaming platform. You'll learn:

- The fundamentals of event-driven architecture and event streaming platforms-
- The background and rationale behind Apache Kafka, its numerous potential uses and applications-
- The architecture and core concepts - the underlying software components, partitioning and parallelism, load-balancing, record ordering and consistency modes-
- Installation of Kafka and related tooling - using standalone deployments, clusters, and containerised deployments with Docker-
- Using CLI tools to interact with and administer Kafka classes, as well as publishing data and browsing topics-
- Using third-party web-based tools for monitoring a cluster and gaining insights into the event streams-
- Building stream processing applications in Java 11 using off-the-shelf client libraries-
- Patterns and best-practice for organising the application architecture, with emphasis on maintainability and testability

of the resulting code- The numerous gotchas that lurk in Kafka's client and broker configuration, and how to counter them- Theoretical background on distributed and concurrent computing, exploring factors affecting their liveness and safety- Best-practices for running multi-tenanted clusters across diverse engineering teams, how teams collaborate to build complex systems at scale and equitably share the cluster with the aid of quotas- Operational aspects of running Kafka clusters at scale, performance tuning and methods for optimising network and storage utilisation- All aspects of Kafka security -including network segregation, encryption, certificates, authentication and authorization. The coverage is progressively delivered and carefully aimed at giving you a journey-like experience into becoming proficient with Apache Kafka and Event-Driven Architecture. The goal is to get you designing and building applications. And by the conclusion of this book, you will be a confident practitioner and a Kafka evangelist within your organisation - wielding the knowledge necessary to

teach others.

[Transactional Machine Learning with Data Streams and AutoML](#) Packt Publishing Ltd
More and more data-driven companies are looking to adopt stream processing and streaming analytics. With this concise ebook, you'll learn best practices for designing a reliable architecture that supports this emerging big-data paradigm. Authors Ted Dunning and Ellen Friedman (Real World Hadoop) help you explore some of the best technologies to handle stream processing and analytics, with a focus on the upstream queuing or message-passing layer. To illustrate the effectiveness of these technologies, this book also includes specific use cases. Ideal for developers and non-technical people alike, this book describes:

- Key elements in good design for streaming analytics, focusing on the essential characteristics of the messaging layer
- New messaging technologies, including Apache Kafka and MapR Streams, with links to sample code
- Technology choices for streaming analytics: Apache Spark Streaming, Apache Flink, Apache Storm, and Apache Apex
- How stream-based architectures are helpful to support microservices

Specific

use cases such as fraud detection and geo-distributed data streams Ted Dunning is Chief Applications Architect at MapR Technologies, and active in the open source community. He currently serves as VP for Incubator at the Apache Foundation, as a champion and mentor for a large number of projects, and as committer and PMC member of the Apache ZooKeeper and Drill projects. Ted is on Twitter as @ted_dunning. Ellen Friedman, a committer for the Apache Drill and Apache Mahout projects, is a solutions consultant and well-known speaker and author, currently writing mainly about big data topics. With a PhD in Biochemistry, she has years of experience as a research scientist and has written about a variety of technical topics. Ellen is on Twitter as @Ellen_Friedman.

Apache Kafka 1.0 Cookbook Packt Publishing Ltd

Design and administer fast, reliable enterprise messaging systems with Apache Kafka About This Book Build efficient real-time streaming applications in Apache Kafka to process data streams of data Master the core Kafka APIs to set up Apache Kafka clusters and start writing

message producers and consumers A comprehensive guide to help you get a solid grasp of the Apache Kafka concepts in Apache Kafka with practical examples Who This Book Is For If you want to learn how to use Apache Kafka and the different tools in the Kafka ecosystem in the easiest possible manner, this book is for you. Some programming experience with Java is required to get the most out of this book What You Will Learn Learn the basics of Apache Kafka from scratch Use the basic building blocks of a streaming application Design effective streaming applications with Kafka using Spark, Storm &, and Heron Understand the importance of a low-latency, high-throughput, and fault-tolerant messaging system Make effective capacity planning while deploying your Kafka Application Understand and implement the best security practices In Detail Apache Kafka is a popular distributed streaming platform that acts as a messaging queue or an enterprise messaging system. It lets you publish and subscribe to a stream of records, and process them in a fault-tolerant way as they occur. This book is a comprehensive guide to designing and

architecting enterprise-grade streaming applications using Apache Kafka and other big data tools. It includes best practices for building such applications, and tackles some common challenges such as how to use Kafka efficiently and handle high data volumes with ease. This book first takes you through understanding the type messaging system and then provides a thorough introduction to Apache Kafka and its internal details. The second part of the book takes you through designing streaming application using various frameworks and tools such as Apache Spark, Apache Storm, and more. Once you grasp the basics, we will take you through more advanced concepts in Apache Kafka such as capacity planning and security. By the end of this book, you will have all the information you need to be comfortable with using Apache Kafka, and to design efficient streaming data applications with it. Style and approach A step-by-step, comprehensive guide filled with practical and real-world examples

Kafka Streams in Action, Second Edition "O'Reilly Media, Inc."
Learn how to integrate full-stack open source big data architecture and to choose

the correct technology—Scala/Spark, Mesos, Akka, Cassandra, and Kafka—in every layer. Big data architecture is becoming a requirement for many different enterprises. So far, however, the focus has largely been on collecting, aggregating, and crunching large data sets in a timely manner. In many cases now, organizations need more than one paradigm to perform efficient analyses. Big Data SMACK explains each of the full-stack technologies and, more importantly, how to best integrate them. It provides detailed coverage of the practical benefits

of these technologies and incorporates real-world examples in every situation. This book focuses on the problems and scenarios solved by the architecture, as well as the solutions provided by every technology. It covers the six main concepts of big data architecture and how to integrate, replace, and reinforce every layer: The language: Scala The engine: Spark (SQL, MLib, Streaming, GraphX) The container: Mesos, Docker The view: Akka The storage: Cassandra The message broker: Kafka What You Will Learn: Make big data architecture without using

complex Greek letter architectures Build a cheap but effective cluster infrastructure Make queries, reports, and graphs that business demands Manage and exploit unstructured and No-SQL data sources Use tools to monitor the performance of your architecture Integrate all technologies and decide which ones to replace and which ones to reinforce Who This Book Is For: Developers, data architects, and data scientists looking to integrate the most successful big data open stack architecture and to choose the correct technology in every layer