

1. Record Nr.	UNINA9910808446903321
Autore	Lopez Estrada Raul Eduardo
Titolo	Apache Kafka quick start guide : leverage Apache Kafka 2.0 to simplify real-time data processing for distributed applications // Raul Estrada
Pubbl/distr/stampa	Birmingham : , : Packt, , 2018
Edizione	[1st edition]
Descrizione fisica	1 online resource (186 pages)
Disciplina	005.713769
Soggetti	Electronic data processing Telecommunication - Message processing
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Sommario/riassunto	<p>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 ?y. 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 asex, 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</p>

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

- How to validate data with Kafka
- Add information to existing data
- How to 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.
