

Confluent Administrator for Apache Kafka

CODICE	DT0006
DURATA	4 gg
PREZZO	1.590,00 €
EXAM	

DESCRIZIONE

Obiettivo del corso è preparare alla certificazione "Confluent Certified Administrator for Apache Kafka", si partirà dall'installazione e configurazione della piattaforma Confluent. Si procederà poi ad utilizzare gli strumenti di gestione confluent per eseguire Kafka più facilmente, oltre che Archiviare ed elaborare i dati del flusso in ingresso, ottimizzare e gestire i cluster Kafka.

TARGET

Sistemisti

PREREQUISITI

- Esperienze sistemiche su sistemi Linux
- Consigliate conoscenze sistemiche di base

CONTENUTI

Modulo 1: Introduzione all'ecosistema software Kafka

- Il mondo Apache
- Concetti Base della Messagistica
- Principali modelli di Messagistica
- Protocolli
- Apache Kafka
- Campi di applicazione
- Kafka Connect
- Schema Registry
- REST Proxy
- Ksql
- Zookeeper
- Introduzione alle Kafka Java Client APIs

- Producer API
- Consumer API
- Streams API
- Connect API
- Installazione Ambiente
 - Installazione di Zookeeper
 - Installazione di Kafka
 - Installazione di altri moduli Confluent
 - Configurazione Java

Modulo 2: Kafka fundamentals

- Publish/Subscribe Messaging e Streaming
- Kafka's Command Line Tools
- Fundamentals
 - Apache Kafka architecture, design principles, and purposes
 - Distributed Systems - Scalability, Fault Tolerance, High Availability
 - Primary functions of: Producer, Consumer, Broker
 - Meaning of “immutable” log
 - Meaning of “committed”
 - Topics, Partitions
 - Essential services of Apache Zookeeper
 - Replication, Leaders, Followers
 - Kafka Messages, structure, make-up, metadata
 - Kafka Controller
 - Exactly Once Semantics
- Brokers & Cluster
- Il Flow completo di Kafka
- Approfondimenti
 - Disk-Based Retention
 - Scalabilità
 - Alte performance
 - Security
- Zookeeper

Modulo 3: Managing, configuring, and optimizing a cluster for performance

- Startup sequence; component dependencies
- How many partitions? Tradeoffs
- Scalability factors
- Sources and tools for monitoring; Display of metrics

- InSyncReplicas (ISR); Fully and Under replicated, and offline
- Consumer lag, Under/Over Consumption
- Broker failure, detection, and recovery
- Batching and its impacts/consequences
- Determining and solving data imbalance across brokers
- Impacts of average and maximum message sizes
- Quotas
- Tuning

Modulo 4: Kafka Avanzato

- Brokers and Zookeeper
- CPU, RAM, network, storage considerations
- Number of nodes
- Rack awareness
- Kafka Connect
 - Source and Sink Connectors
 - Scalability and High Availability
- Business Continuity / DR
- Data retention