

Formation Kafka

Durée

3 jours (21 heures)

Présentation

Dans cette formation pratique, vous apprendrez à construire une application qui peut publier des données sur un cluster Apache Kafka et s'abonner à celles-ci. Vous apprendrez à utiliser Kafka pour manipuler les pipelines de données, à maîtriser ses concepts, à utiliser les composants principaux et l'architecture de base de Kafka et à appeler ses API. Outre le Core Kafka, le Kafka Connect et le Kafka Streams, la formation aborde également la plateforme Confluent, comme le Schema Registry et le REST Proxy.

Comme pour toutes nos formations, celle-ci vous présentera la toute dernière version de Kafka (à la date de rédaction de l'article : [Kafka 2.6.0](#) sortie en [août 2020](#)).

Objectifs

- Comprendre l'architecture de Kafka et son mode de fonctionnement dans un contexte multi-clients
- Publier et récupérer des données vers un cluster Kafka
- Utiliser les outils et les API permettant l'interaction avec un cluster Kafka
- Sécuriser les données à destination et en provenance d'un cluster Kafka
- Créer vos propres Consumers et Producers
- Implémenter un consumer mutli-thread
- Utiliser un proxy REST
- Stocker des données Avro dans Kafka avec le Schema Registry
- Ingestion de données avec Kafka Connect

Public visé

Développeurs, Data Scientists, Architectes, Administrateurs système, DevOps.

Pré-requis

- Connaissances de base d'un système Unix
- Connaissance dans un langage de développement moderne (Java, Python, Scala)

Pour aller plus loin

Programme de la formation Kafka

Introduction

- Historique et complexité des systèmes
- Le traitement en temps réel : une importance business
- Kafka : Une plateforme de données de flux

Les fondamentaux

- Un aperçu de Kafka et efficacité
- Producers
- Brokers
- Consumers
- Utilisation de ZooKeeper

Installation & Administration

- Installation et options avancés
- Considération matérielle (Hardware)
- Administrer Kafka

Architecture

- Fichiers journaux de Kafka
- Réplication et fiabilité
- Chemins d'écriture et de lecture de Kafka
- Partitions, Consumer Groups et Scalability

Développer avec Kafka

- Utiliser Maven pour la gestion de projet
- Accéder par programmation à Kafka
- Écrire un Producer in Java
- Utiliser REST API pour écrire un Producer
- Écrire un Consumer en Java
- Utiliser REST API pour écrire un Consumer

Développement avancé avec Kafka

- Créer un Consumer Multi-Thread
- Gestion des Offsets
- Consumer Rebalancing (rééquilibrage)
- Gestion manuelle des Offsets (Commits)
- Partitionnement des données
- Message Durability : gestion de la durée des messages

Schema Management

- Introduction à Avro
- Avro Schemas
- Utilisation du Schema Registry

Kafka Connect : Data Movement

- Cas d'utilisation de Kafka Connect
- Les bases de Kafka Connect
- Modes : Standalone and Distributed
- Configuration du mode distribué
- Tracking Offsets
- Connector Configuration
- Comparaison d'utilisation avec des options avancées

Kafka Streams

- Cas d'utilisation de Kafka Streams
- Kafka Streams : les fondamentaux Fundamentals
- Étude d'une application et usage de Kafka Streams

Tuning & Supervision avancé (Module Optionnel) - 2 jours

- Log Compaction
- Configuration Settings
 - Hardware and Capacity Planning
 - Traitement des Oversized Messages
- Monitoring & Alerte
 - ZooKeeper and OS-Level Monitoring
 - Key Kafka Metrics
- Cluster Administration
 - Key Cluster Administration Tasks
 - Replicating Clusters avec MirrorMaker

- Kafka Security
 - SSL for Encryption and Authentication
 - SASL for Authentication
 - Authorization
 - Migration vers Secure Cluste

MODULE COMPLÉMENTAIRE EN ANGLAIS SUR DEMANDE (+2 JOURS) - KAFKA DEVELOPER

- Training language : English
- Course level : Beginner to intermediate

Théorie : 60% Pratique : 40%

Audience :? Software Developers? Data Engineers? Architects

Getting Started

- Kafka Overview
- Kafka Basic Concepts
- Zookeeper Overview

Developing with Kafka

- First Kafka Producer (Java/Python)
- Inside Kafka Producer
- Advanced Kafka Producer (Java/Python)
- First Kafka Consumer (Java/Python)
- Inside Kafka Consumer
- Advanced Kafka Consumer (Java/Python)

Replications and Reliability

- Basic Replication Concepts
- Recovery and Failures
- Exactly Once Semantics (EOS)
- Controller Broker
- Inside Kafka Broker

Kafka Topic Management

- Topic Basic Operations
- Log Management
- Designing Topics
- Kafka Security

Kafka Connect

- Kafka Connect Concepts
- Types of Connectors
- Kafka Connect Implementation
- Standalone and Distributed Modes
- Configuring the Connectors

Kafka Streams

- Kafka Streams Concepts
- First Kafka Stream (Java/Python)
- Stateless KStreams
- Stateful KStreams

MODULE COMPLÉMENTAIRE EN ANGLAIS SUR DEMANDE (+2 JOURS) - KAFKA OPERATIONS

- Training language : English
- Course level : Beginner to intermediate

Théorie : 60% Pratique : 40%

Audience :

- Data Engineers
- Architects
- System Administrators
- DevOps

Getting Started

- Kafka Overview
- Kafka Basic Concepts
- Zookeeper Overview

Replications and Reliability

- Basic Replication Concepts
- Recovery and Failures
- Exactly Once Semantics (EOS)
- Controller Broker
- Inside Kafka Broker
- Inside Kafka Producer
- Inside Kafka Consumer

Log Retention and Compaction

- Retention
- Compaction

Managing Kafka Cluster

- Install and Configure Kafka
- Monitor Kafka
- Cluster Management
- Scaling Kafka Cluster

Kafka Security

- SSL for Encryption and Authentication
- SASL for Authentication
- Securing ZooKeeper
- Securing Kafka Broker

Hardware and Settings

- Cluster Sizing
- Configuration Settings
- Capacity Planning

Monitoring

- Kafka Broker Metrics
- Client Monitoring
- Monitoring Tools

Architecture

- Kafka Connect Overview
- Kafka Streams Overview
- Reference Architectures
- Available Kafka Distributions

Formation Spark Streaming

Formation Spark Tuning Avancé

Formation Spark et Machine Learning

Formation Cassandra

Formation MongoDB

Formation PostgreSQL

Sociétés concernées

Cette formation s'adresse aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiple permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.