

Big Data - L'écosystème centralisé de streaming avec Apache Kafka

Diffuser des données en continu

 Présentiel ou en classe à distance



3 jours (21 h)

Prix inter : 2.460,00 € HT

Réf.: BD554



Idéal en
Distanciel

Apache Kafka joue un rôle central dans les architectures de streaming actuelles. Dans l'univers du Big Data en constante mutation, Kafka s'est imposé comme un élément clé.

Cette révolution a remodelé les métiers et compétences exigées, demandant désormais une maîtrise approfondie de Kafka pour traiter les flux de données. Cette formation est idéale pour comprendre en profondeur son fonctionnement, les principes de communication inter-applications et les bonnes pratiques de distribution de messages.

Vous découvrirez comment configurer Kafka pour intégrer des données de différents formats et sources, maîtriser ses différentes APIs et mettre en oeuvre KSQL. De plus, vous acquerez les compétences pour travailler avec Kafka en toute sécurité.

A qui s'adresse cette formation ?



Pour qui

- Architectes, chefs de projet, développeurs, équipes DevOps...



Prérequis

- Posséder des notions d'architectures applicatives en entreprise et des connaissances des protocoles réseaux
- Avoir de bonnes connaissances en Java ou d'un autre langage
- Des connaissances en Spring ou Node.js sont recommandées

Programme

1 - Comprendre le fonctionnement de Kafka et les principes de communications inter applications

- Le projet Kafka : historique, fonctionnalités, principe de fonctionnement
- Présentation de l'architecture et du rôle de chaque composant : broker, producer, consumer
- Liaison éventuelle avec Zookeeper. Impacts

2 - Acquérir les bonnes pratiques de distribution de messages

- Étude de la configuration du broker
Ateliers :

Création d'une configuration multi-broker, démarrage de plusieurs noeuds

Création d'un topic simple et mise en oeuvre d'une chaîne de base. Visualisation des messages avec kafka-console-consumer

3 - Savoir configurer Kafka pour intégrer les données de différents formats et de sources différentes

- Kafka Connect : présentation des fonctionnalités : intégration de données d'origines multiples, modes de fonctionnement (standalone ou distribué)
 - Types de connecteurs
- Atelier :

Configuration de connecteurs, ingestion de données, création d'une chaîne de transformation

4 - Appréhender les différentes APIs de Kafka

- Conception d'applications avec Kafka
 - Principe de fonctionnement
 - Couplage avec SparkStreaming en mode batch, en mode continu
 - Principe et architecture de Kafka Streams
- Atelier :

Développement de prototypes en Python, Java, Scala

5 - Mettre en oeuvre KSQL

- Conception d'application avec KSQL
 - La sur-couche KSQL
 - Présentation de ksqldb
 - Création de flux en KSQL
 - Les ajouts SQL pour permettre le traitement en flux
 - Micro-batches
 - Introduction au water-marking
- Atelier :

Mise en oeuvre d'une chaîne de traitement avec KSQL

6 - Travailler en sécurité avec Kafka

- Intégrité et confidentialité : chiffrement SSL et authentification SASL
 - Sécurisation de l'infrastructure : Zookeeper, proxy Rest
 - Disponibilité : la réplication
 - Facteur de réplication
 - Partitions
- Atelier :

Tests de haute disponibilité dans une configuration multi-noeuds

7 - Exploitation

- Mise en oeuvre de Kafka-tool
- Gestion des logs
- Détection de pannes : méthodes et outils
- Benchmarks



Les objectifs de la formation

- Être capable de comprendre le fonctionnement de Kafka et les principes de communications inter applications
- Pouvoir acquérir les bonnes pratiques de distribution de messages
- Savoir configurer Kafka pour intégrer les données de différents formats et de sources différentes
- Être capable d'appréhender les différentes APIs de Kafka
- Apprendre à mettre en oeuvre KSQL
- Comprendre comment travailler en sécurité avec Kafka



Evaluation

- Cette formation fait l'objet d'une évaluation formative.



Les points forts de la formation

- Une formation très opérationnelle durant laquelle s'alternent les phases d'apports théoriques, d'échanges, de partage d'expériences.
- A l'issue de la formation les participants seront capable de mettre en oeuvre l'outil d'Apache Kafka pour utiliser un système de messagerie distribuée à haut débit.
- Les participants réalisent sur un cas concret une chaîne de traitement Streaming.
- 96% des participants à cette formation se sont déclarés satisfaits ou très satisfaits au cours des 12 derniers mois.



Dates et villes 2026 - Référence BD554



Dernières places disponibles



Session garantie

A distance

du 21 janv. au 23 janv.

du 26 août au 28 août

du 4 mai au 6 mai

du 2 nov. au 4 nov.

Paris

du 21 janv. au 23 janv.

du 26 août au 28 août

du 4 mai au 6 mai

du 2 nov. au 4 nov.