

Kafka : ingestion de messages

Mettre en œuvre Kafka pour la distribution de messages

Détails

- Code : AE-KAF
- Durée : 3 jours (21 heures)

Public

- Chefs de projets
- Consultants
- Consultants informatiques
- Développeurs
- Professionnels de l'IT

Pré-requis

- Connaissance de l'écosystème hadoop
- Bases de programmation dans un langage objet

Objectifs

- Comprendre le fonctionnement de Kafka et les principes de communications inter applications
- Acquérir les bonnes pratiques de distribution de messages
- Savoir configurer Kafka pour intégrer les données de différents formats et de sources différentes
- Appréhender les différentes APIs de Kafka
- Mettre en œuvre KSQL

Programme

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

Acquérir les bonnes pratiques de distribution de messages

- Etude de la configuration du broker
- *Atelier : création d'une configuration multi-broker, démarrage de plusieurs noeuds*
- *Atelier : création d'un topic simple et mise en œuvre d'une chaîne de base. Visualisation des messages avec kafka-console-consumer*

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*

Appréhender les différentes APIs de Kafka

- Conception d'applications avec Kafka
- Principe de fonctionnement

- *Atelier : Développement de prototypes en Python, Java, Scala*
- Couplage avec SparkStreaming en mode batch, en mode continu
- Principe et architecture de Kafka Streams

Mettre en œuvre 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 œuvre d'une chaîne de traitement avec KSQL*

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*

Exploitation

- Mise en œuvre de kafka-tool
- Gestion des logs
- Détection de pannes : méthodes et outils
- Benchmarks

Modalités

- **Type d'action** :Acquisition des connaissances
- **Moyens de la formation** :Formation présentielle – 1 poste par stagiaire – 1 vidéo projecteur – Support de cours fourni à chaque stagiaire
- **Modalités pédagogiques** :Exposés – Cas pratiques – Synthèse
- **Validation** :Exercices de validation – Attestation de stages