

Apache Kafka

Référence : **PYCB041**

Durée : **3 jours (21 heures)**

Certification : **Aucune**

Connaissances préalables

- Connaissance de l'écosystème hadoop et bases de programmation dans un langage objet (java ou scala ou python)

Profil des stagiaires

- Chefs de projet, développeurs souhaitant mettre en oeuvre kafka pour la distribution de messages.

Objectifs

- Comprendre le fonctionnement de kafka
- Acquérir les bonnes pratiques de distribution de messages
- Savoir configurer kafka pour intégrer les données de différents formats et de sources différentes

Certification préparée

- Aucune

Méthodes pédagogiques

- Mise à disposition d'un poste de travail par participant
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions

Formateur

- Consultant-Formateur expert Bigdata

Méthodes d'évaluation des acquis

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation des compétences acquises envoyée au stagiaire
- Attestation de fin de stage adressée avec la facture

Contenu du cours

1. Introduction

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

2. Mise en oeuvre

- Préconisations d'installation et prérequis
- Travaux pratiques: installation et lancement de zookeeper et du kafka-server, Création d'un topic simple, Mise en oeuvre d'une chaîne de base
- Visualisation des messages avec kafka-console-consumer

3. Multi-broker

- Etude de la configuration du broker
- Travaux pratiques : création d'une configuration multi-broker, démarrage de plusieurs noeuds

4. Applicatif

- Conception d'applications avec Kafka. Principe de fonctionnement
- Travaux pratiques : Développement de prototypes en Python, Java, Scala, Couplage avec SparkStreaming en mode batch, en mode continu, Principe et architecture de Kafka Streams, Exemples avec KSQL

5. La réplication

- Facteur de réplication
- Partitions
- Travaux pratiques : tests de haute disponibilité dans une configuration multi-noeuds

6. Kafka Connect

- Présentation des fonctionnalités : intégration de données d'origines multiples, modes de fonctionnement (standalone ou distribué)
- Types de connecteurs
- Travaux pratiques : configuration de connecteurs, ingestion de données, création d'une chaîne de transformation

7. Sécurisation

- Chiffrement SSL et authentification SASL
- Sécurisation de l'infrastructure : Zookeeper, proxy Rest

8. Exploitation

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

Notre référent handicap se tient à votre disposition au [01.71.19.70.30](tel:0171197030) ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.