

# Kafka : Ingestion de message

Référence : PYCB041

Durée : 2 jours

Certification : **Aucune**

## CONNAISSANCES PREALABLES

- 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 PREPAREE

Aucune

## METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

## FORMATEUR

Consultant-Formateur expert Bigdata

## METHODE D'EVALUATION DES ACQUIS

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

## CONTENU DU COURS

### 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

### 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

### Multi-broker

- Etude de la configuration du broker

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

### 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

### La réplication

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

### **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

### **Sécurisation**

- Chiffrement SSL et authentification SASL

- Sécurisation de l'infrastructure : Zookeeper, proxy Rest

### **Exploitation**

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