

# Formation Hadoop Cloudera, Développement et administration moderne

## Présentation

Cette formation avancée permet aux participants de concevoir, optimiser et administrer des environnements Hadoop modernes basés sur Cloudera. Elle aborde l'architecture distribuée, l'optimisation des performances, la sécurisation, ainsi que les bonnes pratiques d'administration et de développement dans un environnement Big Data d'entreprise.

Durée : 21,00 heures (3 jours)

Tarif INTRA : [Nous consulter](#)

## Objectifs de la formation

- Concevoir des architectures Hadoop distribuées performantes
- Optimiser et administrer un cluster Cloudera moderne
- Évaluer et sécuriser les environnements Hadoop en production
- Déployer et exploiter des workflows Big Data avancés

## Prérequis

- Maîtrise des systèmes distribués
- Connaissances solides en Linux
- Expérience en gestion ou développement de systèmes Big Data

## Public

Architectes data, administrateurs systèmes, ingénieurs Big Data, experts techniques impliqués dans des projets Hadoop en entreprise.

## Programme de la formation

### Jour 1 – Architecture Hadoop avancée et écosystème Cloudera

#### Session du matin :

- Architecture Hadoop moderne : NameNode, DataNodes, YARN



- Présentation avancée de l'écosystème Cloudera (CDP)

**Session de l'après-midi :**

- Modèles de stockage et formats optimisés (Parquet, ORC)
- Exploration des services HDFS et YARN

**TP / Exercice :**

- Déploiement et configuration initiale d'un cluster Cloudera avec paramétrage avancé

**Points clés & takeaways :**

- Vision approfondie des composants Hadoop
- Compréhension des services clés de Cloudera

**Jour 2 – Administration, optimisation et sécurité**

**Session du matin :**

- Administration avancée via Cloudera Manager
- Optimisation des performances (tuning, réplication, I/O)

**Session de l'après-midi :**

- Sécurité : Kerberos, ACLs, chiffrement
- Supervision, logs, audits et monitoring avancé

**TP / Exercice :**

- Mise en place d'une stratégie de sécurité complète et optimisation d'un cluster existant

**Points clés & takeaways :**

- Maîtrise des outils d'administration Cloudera
- Gestion avancée de la sécurité et des performances

## Jour 3 – Développement et exploitation Big Data avancée

### Session du matin :

- Développement Big Data avec Hive, Impala et Spark
- Optimisation des jobs et des workflows

### Session de l'après-midi :

- Automatisation des pipelines Big Data
- Mise en production et bonnes pratiques CI/CD

### TP / Exercice :

- Conception et exécution d'un pipeline Big Data complet : ingestion, traitement et restitution

### Points clés & takeaways :

- Capacité à concevoir un pipeline complet sous Cloudera
- Maîtrise de l'intégration et de l'exploitation avancée

## Organisation

### Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes. Riches de leur expérience sur le sujet, ils sauront accompagner vos collaborateurs dans leur montée en compétences.

### Moyens pédagogiques et techniques

- Apports des connaissances communes.
- Mises en situation sur le thème de la formation et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.
- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.

- Pour les formations en présentiel dans les locaux mis à disposition, les apprenants sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

## **Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation**

En amont de la formation :

- Recueil des besoins des apprenants afin de disposer des informations essentielles au bon déroulé de la formation (profil, niveau, attentes particulières...).
- Auto-positionnement des apprenants afin de mesurer le niveau de départ.

Tout au long de la formation :

- Évaluation continue des acquis avec des questions orales, des exercices, des QCM, des cas pratiques ou mises en situation...

A la fin de la formation :

- Auto-positionnement des apprenants afin de mesurer l'acquisition des compétences.
- Evaluation par le formateur des compétences acquises par les apprenants.
- Questionnaire de satisfaction à chaud afin de recueillir la satisfaction des apprenants à l'issue de la formation.
- Questionnaire de satisfaction à froid afin d'évaluer les apports ancrés de la formation et leurs mises en application au quotidien.

## **Accessibilité**

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.