

# Hadoop : l'écosystème

## INFORMATIONS GÉNÉRALES

**Type de formation :** Formation continue

**Éligible au CPF :** Non

**Domaine :** IA, Big Data et Bases de données

**Action collective :** Non

**Filière :** Big Data

**Rubrique :** NoSQL et Hadoop

**Code de formation :** BD018

## € Tarifs

**Prix public :** 720 €

### Tarif & financement :

Nous vous accompagnons pour trouver la meilleure solution de financement parmi les suivantes :

- **Le plan de développement des compétences** de votre entreprise : rapprochez-vous de votre service RH.
- **Le dispositif FNE-Formation.**
- **L'OPCO** (opérateurs de compétences) de votre entreprise.
- **Pôle Emploi** sous réserve de l'acceptation de votre dossier par votre conseiller Pôle Emploi.
- **CPF -MonCompteFormation**

Contactez nous pour plus d'information

## PRÉSENTATION

### Objectifs & compétences

Chefs de projets, développeurs, et toute personne souhaitant comprendre les mécanismes Hadoop et le rôle de chaque composant.

### Public visé

Chefs de projets, développeurs, et toute personne souhaitant comprendre les mécanismes Hadoop et le rôle de chaque composant.

### Pré-requis

Connaissances générales des systèmes d'information.

## 📍 Lieux & Horaires

**Campus :** Ensemble des sites

**Durée :** 7 heures

### Délai d'accès :

Jusqu'à 8 jours avant le début de la formation

**Distanciel possible :** Oui

## PROGRAMME

### Introduction

Rappels sur NoSQL.

Le théorème CAP.

Historique du projet hadoop

Les fonctionnalités : stockage, outils d'extraction, de conversion, ETL, analyse, ...

Exemples de cas d'utilisation sur des grands projets.

Les principaux composants : HDFS pour le stockage et YARN pour les calculs.

Les distributions et leurs caractéristiques (HortonWorks, Cloudera, MapR, GreenPlum, Apache, ...) 2.

### L'architecture

Terminologie : NameNode, DataNode, ResourceManager

Rôle et interactions des différents composants

**Présentation des outils d'infrastructure :** ambari, avro, zookeeper; de gestion des données : pig, oozie, tez, falcon, pentaho, sqoop, flume; d'interfaçage avec les applications GIS; de restitution et requêtage : webhdfs, hive, hawq, impala, drill, stinger, tajo, mahout, lucene, elasticSearch, Kibana

**Les architectures connexes :** spark, cassandra

### Exemples interactifs

Démonstrations sur une architecture Hadoop multi-nœuds.

Mise à disposition d'un environnement pour des exemples de calcul

Travaux pratiques : Recherches dans des données complexes non structurées.

### Applications

Cas d'usages de hadoop.

Calculs distribués sur des clusters hadoop

## 📅 Prochaines sessions

Cliquez sur la date choisie pour vous inscrire :

### ■ 25 / 11 / 2024

📍 : Ensemble des sites

✓ : Distanciel possible

🕒 : 7 heures

📅 : 1 jours

### ■ 02 / 09 / 2025

📍 : Ensemble des sites

✓ : Distanciel possible

🕒 : 7 heures

📅 : 1 jours

## MODALITÉS

**Modalités**

**Modalités** : en présentiel, distanciel ou mixte – Horaires de 9H à 12H30 et de 14H à 17H30 soit 7H – Intra et Inter entreprise

**Pédagogie** : essentiellement participative et ludique, centrée sur l'expérience, l'immersion et la mise en pratique. Alternance d'apports théoriques et d'outils pratiques.

**Ressources techniques et pédagogiques** : Support de formation au format PDF ou PPT Ordinateur, vidéoprojecteur, Tableau blanc, Visioconférence : Cisco Webex / Teams / Zoom

**Pendant la formation** : mises en situation, autodiagnosics, travail individuel ou en sous-groupe sur des cas réels

**Méthode**

**Fin de formation** : entretien individuel

**Satisfaction des participants** : questionnaire de satisfaction réalisé en fin de formation

**Assiduité** : certificat de réalisation (validation des acquis)