

## Formation Hadoop : Administration

■ <b>Durée :</b>	2 jours (14 heures)
■ <b>Tarifs inter-entreprise :</b>	1 875,00 € HT (standard) 1 500,00 € HT (remisé)
■ <b>Public :</b>	Administrateurs systèmes
■ <b>Pré-requis :</b>	Connaissances en administration système, préférablement Java
■ <b>Objectifs :</b>	- Installer les services d'un nœud Hadoop - Assembler plusieurs nœuds Hadoop - Déployer une nouvelle application sur un cluster existant - Effectuer une restauration de données suite à une reprise sur incident
■ <b>Modalités pédagogiques, techniques et d'encadrement :</b>	<ul style="list-style-type: none"><li>• Formation synchrone en présentiel et distanciel.</li><li>• Méthodologie basée sur l'Active Learning : 75 % de pratique minimum.</li><li>• Un PC par participant en présentiel, possibilité de mettre à disposition en bureau à distance un PC et l'environnement adéquat.</li><li>• Un formateur expert.</li></ul>
■ <b>Modalités d'évaluation :</b>	<ul style="list-style-type: none"><li>• Définition des besoins et attentes des apprenants en amont de la formation.</li><li>• Auto-positionnement à l'entrée et la sortie de la formation.</li><li>• Suivi continu par les formateurs durant les ateliers pratiques.</li><li>• Évaluation à chaud de l'adéquation au besoin professionnel des apprenants le dernier jour de formation.</li></ul>
■ <b>Sanction :</b>	Attestation de fin de formation mentionnant le résultat des acquis
■ <b>Référence :</b>	BUS100612-F
■ <b>Note de satisfaction des participants:</b>	Pas de données disponibles
■ <b>Contacts :</b>	commercial@dawan.fr - 09 72 37 73 73

■ <b>Modalités d'accès :</b>	Possibilité de faire un devis en ligne (www.dawan.fr, moncompteformation.gouv.fr, maformation.fr, etc.) ou en appelant au standard.
■ <b>Délais d'accès :</b>	Variable selon le type de financement.
■ <b>Accessibilité :</b>	Si vous êtes en situation de handicap, nous sommes en mesure de vous accueillir, n'hésitez pas à nous contacter à referenthandicap@dawan.fr, nous étudierons ensemble vos besoins

## Préambule

Qu'est-ce que le Big Data ?  
La problématique de taille  
Position de Hadoop dans le paysage

## Administration de Hadoop

Présentation d'un nœud existant  
Organisation des services et étude du séquençement avec YARN

## Atelier : modifier la taille des blocs HDFS pour diminuer le nombre de Map/Reduce

## Mettre Hadoop en place

Relation entre la plateforme installée et les framework de développement  
Proposer de frameworks indépendants pour assurer la compatibilité : Spring Data

## Atelier : déployer une application d'accès à HBase au travers d'un mapping O/R Spring Data

## Travailler avec MapReduce

Déployer un programme Map/Reduce sur un cluster de nœuds Hadoop  
Recherche des logs  
Remonter les anomalies aux développeurs  
Proposer l'usage de file Kafka

## Atelier : utilisation de file d'entrée sortie pour un programme Map/Reduce

## Routage de données

Définition de routes logicielles

Mettre en place un cas de calcul où les données déclenchent les programmes

**Atelier : faire un routage de données depuis un répertoire HDFS vers une file Kafka qui est l'entrée d'un programme Map/Reduce**

### **Utilisation des vues**

Utilisation des vues Ambari

Visualisation de l'état des nœuds d'un cluster

Importer/exporter des fichiers de configuration

**Atelier : relancer une grappe de services, utilisation des vues YARN et Tez**

### **Gestion des droits**

Gestion des comptes utilisateurs

Gestion des droits de fichier sur un système de fichier distribué

Utilisation de certificat

**Atelier : configurer les services Knox et Ranger**