

• **DURÉE**

3 jours / 21h

APPRÉCIATION

Évaluation qualitative de fin de stage

MODALITÉS

ET MOYENS

PÉDAGOGIQUES

Démonstrations, cas pratiques, synthèse et évaluation des acquis

► **Objectifs**

pédagogiques

- Pouvoir définir l'architecture d'une solution Hadoop
- Installer et configurer un cluster Hadoop
- Configurer et personnaliser un environnement Hadoop pour optimiser la disponibilité et le débit des données

► **Public concerné**

Chef de projets, Administrateur Système et toute personne souhaitant mettre en œuvre un système Hadoop

► **Prérequis**

Connaissance systèmes Unix/Linux

Big Data

HADOOP - ARCHITECTURE & ADMINISTRATION

2 030 € HT

Inter-Entreprises
Intra-entreprises sur devis

Hadoop est la principale plateforme de Big Data. Elle assure le stockage et le traitement d'immenses volumes de données. Cette formation vous permettra de comprendre son architecture et d'acquérir les compétences nécessaires pour installer, configurer et administrer un cluster Hadoop.

Introduction

- Présentation générale d'Hadoop
- Exemples d'utilisations d'Hadoop
- Principales distributions d'Hadoop
- Planification

Architecture et composants de la plateforme Hadoop

- Approche conceptuelle d'Hadoop
- HDFS : NameNode / DataNode
- Paradigme MapReduce et YARN
- Les technologies émergentes

Installation et configuration

- Installation d'une plateforme Hadoop
- Définir l'architecture
- Installation d'Hadoop / HDFS
- Installation des principaux composants: Hive, Pig, etc.
- Mise en œuvre: installation d'Hadoop et de quelques composants
- Configuration de l'environnement
- Étude des fichiers de configuration
- Modes de configuration: ligne de commandes, API Rest, serveur http intégré, API natives
- Configuration des services: HDFS, YARN, Logs, etc.
- Configuration des droits d'accès

Administration

- Gestion des ressources
- Affecter les quotas
- Gérer les planificateurs
- Administration de HDFS
- Gestion des démons Hadoop
- Ajout / suppression des nœuds
- Gestion des données dans HDFS
- Administration de MapReduce
- Gestion des tâches MapReduce
- Gestion des nœuds de calcul
- Planification des stratégies de sécurité et de récupération des données

Travaux pratiques

Installation, configuration et administration pas à pas d'un cluster Hadoop

