FORMATION · BIG DATA

Hadoop – Déployer du Big Data

Durée de la formation : 4 jours

Tarif : 2400 €
Réf : BI20
Délai d'accès : 2 à 4 semaines

Prérequis

Connaissances en administration système, préalablement Java.

Public concerné

Administrateurs systèmes

Objectifs pédagogiques

Comprendre le Big Data et ses enjeux Savoir déployer Hadoop et son écosystème
Comprendre HDFS, MapReduce
Structurer les données ave HBase Écrire des requêtes avec HiveQl Lancer une analyse avec Pig

Objectifs opérationnels

Avoir acquis les connaissances sur le Big Data et Hadoop
Savoir mettre en place Hadoop Connaître les intérêts de MapReduce Travailler en autonomie

Méthodes mobilisées et modalités d'évaluation

L'évaluation se fait en fin de parcours sous forme de quizz et d'étude de cas

HADOOP

Hadoop, un framework révolutionnaire dans le domaine du Big Data, offre une solution robuste pour le stockage distribué et le traitement parallèle des données massives, nécessitant une expertise pour en maîtriser pleinement les capacités et les meilleures pratiques associées.

De la théorie mais beaucoup de pratiques vous permettront de maitriser l'usage d'Hadoop.

Inscription

Contactez-nous au 01 86 95 71 60 ou à l'adresse formation@adbi.fr pour connaitre les dates de notre prochaine session

Modes de participation

Nos formations sont proposées sous 2 formats :

- Chez nous prêt de place d'Italie
- Chez vous dans vos locaux

Accessibilité

Nos salles de formation sont accessibles aux personnes à mobilité réduite. N'hésitez à nous informer si vous avez des besoins particuliers et s'il y a des besoins d'aménagements spécifiques lors de votre inscription.

Attestation et certification

Une attestation de fin de formation sera remise à tous les participants à l'issue de leur parcours

Qualiopi processus certifié
RÉPUBLIQUE FRANÇAISE

Programme :

• Introduction au Big Data

- Qu'est-ce que le Big Data ?
- Source des données : l'homme, la machine
- La problématique de taille
- Position de Hadoop dans le paysage

• Introduction à Hadoop

- L'origine du projet
- Le système de fichiers HDFS
- Comprendre l'algorithme MapReduce
- L'environnement d'Hadoop : HBase, Zookeeper, Hive, Pig…
- L'APIYARN

• Mettre Hadoop en place : HDFS

- Du mode autonome au mode complètement distribué en cluster
- Pré-requis, distributions Hadoop
- Cluster Hadoop : NameNode, ResourceManager, DataNode,NodeManager
- Les fichiers de configuration
- Opérations de base sur le duster HDFS : formatage, démarrage, arrêt

Atelier pratique : installer Hadoop sur 2 nœuds, formater et manipuler

• Travailler avec MapReduce

- L'intérêt de MapReduce
- Mappers, reducers, parallélisme et indépendance des traitements
- Entrées, sorties
- Soumission d'un job à Hadoop

Atelier pratique : exécuter une tâche via MapReduce, avec sortie dans HDFS