DAWAN

Plan de la formation

www.dawan.fr

Formation Programmer Hadoop en Java : Approfondissement

Durée :	2 jours (14 heures)
Tarifs inter-entreprise :	2 075,00 CHF HT (standard) 1 660,00 CHF HT (remisé)
Public :	Développeurs Java, Administrateurs (DBA ou systèmes)
Pré-requis :	Avoir suivi Programmer Hadoop en Java Initiation ou notions équivalentes
Objectifs :	Charger et récupérer des données de HBase et du système de fichiers distribué Hadoop (HDFS) - User Defined Functions de Hive et Pig
Modalités pédagogiques, techniques et d'encadrement :	• Formation synchrone en présentiel et distanciel. • Méthodologie basée sur l'Active Learning : 75 % de pratique minimum. • Un PC par participant en présentiel, possibilité de mettre à disposition en bureau à distance un PC et l'environnement adéquat. • Un formateur expert.
Modalités d'évaluation :	• Définition des besoins et attentes des apprenants en amont de la formation. • Auto-positionnement à l'entrée et la sortie de la formation. • Suivi continu par les formateurs durant les ateliers pratiques. • Évaluation à chaud de l'adéquation au besoin professionnel des apprenants le dernier jour de formation.
Sanction :	Attestation de fin de formation mentionnant le résultat des acquis
Référence :	BUS102000-F
Note de satisfaction des participants:	Pas de données disponibles

Contacts :	commercial@dawan.fr - 09 72 37 73 73
Modalités d'accès :	Possibilité de faire un devis en ligne (www.dawan.fr, moncompteformation.gouv.fr, maformation.fr, etc.) ou en appelant au standard.
Délais d'accès :	Variable selon le type de financement.
Accessibilité :	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

Bien-fondé des données distribuées

Optimiser les performances du débit des données
Utiliser la redondance pour récupérer les données

Interfacer avec le système de fichiers distribué Hadoop

Analyser la structure et l'organisation du HDFS
Charger des données brutes et récupérer le résultat
Lire et écrire des données avec un programme
Manipuler les types SequenceFile d'Hadoop
Partager des données de référence avec DistributedCache

Structurer les données avec HBase

Passer du stockage structuré au stockage non structuré
Appliquer les principes NoSQL avec une application de modèle à la lecture, se connecter à HBase à partir des tâches MapReduce, comparer HBase avec d'autres types de magasins de données NoSQL

Exploiter la puissance de SQL avec Hive

Structurer bases de données, les tables, les vues et les partitions
Intégrer des travaux MapReduce avec des requêtes Hive
Lancer des requêtes avec HiveQL
Accéder aux servers Hive via IDBC, ajouter des fonctionnalités à HiveQL avec les fonctions définies par l'utilisateur

Tester et déboguer le code Hadoop

Enregistrer des événements importants à auditer et à déboguer
Valider les spécifications avec MRUnit
Déboguer en mode local

Déployer, surveiller et affiner les performances

Déployer la solution sur un cluster de production
Uutiliser des outils d'administration pour optimiser les performances
Surveiller l'exécution des tâches via les interfaces utilisateur web