

FORMATION HADOOP-DÉVELOPPER DES APPLICATIONS BIG DATA

Ce cours vous permettra de développer des applications vous permettant de traiter des données distribuées en mode batch. Vous collecterez, stockerez et traiterez des données de formats hétérogènes avec Apache Hadoop afin de mettre en place des chaînes de traitement intégrées à votre Système d'Information

PROGRAMME

1/ Le Big Data

- Définition du périmètre du Big Data
- Le rôle du projet Hadoop
- Les concepts de base des projets Big Data
- Présentation du Cloud Computing
- Différence entre Cloud Computing privé et public
- Les architectures Big Data à base du projet Hadoop

2/ Collecte de données et application de Map Reduce

- Analyse des flux de données dans l'entreprise
- Données structurées et non-structurées
- Les principes de l'analyse sémantique des données d'entreprise
- Graphe des tâches à base de MapReduce
- La granularité de cohérence des données
- Transfert de données depuis un système de persistance dans Hadoop
- Transfert de données d'un Cloud dans Hadoop

3/ Le stockage des données avec HBase

- Plusieurs types de base de données XML
- Patterns d'usages et application au Cloud
- Application de Hadoop database au sein d'un workflow
- Utilisation des projets Hive/Pig
- Utilisation du projet HCatalog
- L'API Java HBase

4/ Le stockage des données sur HDFS

- Patterns d'usages et application au Cloud
- Architecture et installation d'un système HDFS, journal, NameNode, DataNode
- Opérations, commandes et gestion des commandes
- L'API HDFS Java
- Analyse de données avec Apache Pig
- Le langage Pig Latin. Utiliser Apache Pig avec Java
- Requêtage avec Apache Hive
- Réplication de données. Partage de données sur une architecture HDFS

PRIX (INTER-ENTREPRISE) : 2710 euros

28h jours

←→ HAD01

Vous souhaitez organiser cette formation dans vos locaux ?

Demandez Houilly au
01 84 25 05 10

OBJECTIFS

- Construire un programme à base de Map Reduce
- Intégrer Hadoop HBase dans un workflow d'entreprise
- Travailler avec Apache Hive et Pig depuis HDFS
- Utiliser un graphe de tâches avec Hadoop

PRE-REQUIS

- Bonne expérience en développement Java
- Connaissances en architecture Web

PUBLIC CONCERNE

- Concepteurs
- Développeurs

DATES INTER-ENTREPRISES

Délais d'entrée : sans

[A distance](#)

04/01/2022 , 19/04/2022 , 23/08/2022 , 22/11/2022 , Paris

04/01/2022 , 19/04/2022 , 23/08/2022 , 22/11/2022 ,

Si vous souhaitez organiser cette formation a une autre date contactez-nous.

NOUS CONTACTER

Openska
21 rue Louise Weiss
75013 Paris

Tel : 01 84 25 05 10
Tel : 01 84 17 44 76
www.openska.com



5/ Spring Data Hadoop

- Introduction à Spring et Spring Data
- Le namespace Hadoop pour Spring
- Utiliser Spring pour simplifier la configuration Hadoop
- Configuration du cache distribué
- Définition des Jobs et dépendance entre Jobs
- Intégration des outils (Pig, Hive...)

ELEMENTS COMPLÉMENTAIRES À TRANSMETTRE IMPÉRATIVEMENT À VOTRE OPCO POUR VOTRE DEMANDE DE PRISE EN CHARGE

HORAIRES

- 9h30-13h
- 14h-17h30

ACCÈS HANDICAP

Oui, Consultez notre référent interne.

MODALITÉS D'ÉVALUATION

L'évaluation des acquis se fait tout au long de la formation au travers de multiples exercices et mise en situation

MODALITÉS DE SUIVI DE L'EXÉCUTION

- Contrôle systématique des présences par demi-journée
- Attestation de présence remis à chaque participant

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la

pédagogie, et ce pour chaque formation qu'ils animent. Ils ont majoritairement cinq à dix années d'expérience dans leur domaine de compétences et ont une grande pratique en entreprise

MOYENS PÉDAGOGIQUES

Les supports pédagogiques sont imprimés par nos soins et transmis à chacun des stagiaires. Les supports sont aussi remis au format électronique aux participants. Ce support est projeté via un vidéo projecteur afin d'animer la formation. Des exercices d'application ou études de cas sont prévus afin de valider les acquis des stagiaires.

MOYENS D'ACCOMPAGNEMENT

8 personnes maximum, formation animée par un formateur expert sur le sujet.
Mise en situation : Pédagogie active et participative.
Apport théorique et méthodologique.
Etude de cas. Atelier pratique.

NATURE DES TRAVAUX DEMANDÉS AU STAGIAIRE

Un exercice est réalisé par le participant à la fin de chaque chapitre. Le participant dispose de 20 min avant de passer à la correction avec le formateur.

MODALITÉS TECHNIQUES EN CAS DE PROBLÈMES

Les connexions et installations d'outil peuvent être testées en amont de la formation avec le formateur et les stagiaires ou le donneur d'ordre. Ce test permet d'éviter tout accident technique lié aux outils de connexion à distance.

Le formateur prend la main sur le poste du participant en cas de difficulté durant la formation.



ORGANISME DE FORMATION RÉFÉRENCÉ SOUS LE NUMÉRO : 117 555 432 75