



Hadoop Apache pour les développeurs (certification cloudera)

Déscription :

Le BigData n'est plus réservé aux géants du Web, c'est maintenant l'affaire de tous. On estime que 80% des données d'une entreprise sont non-structurées et inexploitées... et le volume augmente de plus en plus rapidement. De plus, les infrastructures modernes nous permettent de conserver de plus en plus de données à moindre coût : rien n'est à jeter, tout peut être utile.

A la suite de la formation, les stagiaires auront la possibilité de passer l'examen « Cloudera Certified Developer ». Ils deviendront alors des experts certifiés Cloudera dans leur entreprise.

Objectifs

- Pourquoi choisir Hadoop
- Hadoop: les concepts de base
- Rédiger un programme « MapReduce »
- L'écosystème Hadoop
- Intégrer Hadoop dans un « Workflow »
- Les algorithmes « MapReduce » communs
- Travailler avec Apache Hive et Pig
- Débugger des programmes « MapReduce »
- Le développement « MapReduce » avancé
- L'utilisation de Graph avec Hadoop
- Le nouvel API Hadoop

Publics

Cette formation s'adresse aux développeurs qui développeront, maintiendront et optimiseront des tâches Hadoop.

Durée

4 jours

Pré-requis

Bonnes connaissances du développement Java

Programme de cette formation

Les technologies Hadoop

- Comment fonctionne « Hadoop Distributed File System » et « MapReduce »
- Comment développer des applications "MapReduce"

•

MapReduce

- Comment réaliser des tests unitaires d'applications MapReduce
- Comment utiliser des combinateurs MapReduce, des partitionners et des caches distribués
- Les meilleures pratiques pour développer et déboguer des applications MapReduce
- Comment mettre en œuvre l'entrée et la sortie des données dans les applications MapReduce
- Les algorithmes pour les principales tâches "MapReduce"
- Comment relier des ensembles de données dans MapReduce
- Comment Hadoop s'intègre dans le Data Center

Utilisation des algorithmes d'apprentissage automatique de Mahout

- Comment Hive et Pig peuvent être utilisés pour développer rapidement des applications
- Comment créer des workflow important avec Oozie