

Introduction au développement Hadoop

Description

La formation Hadoop "Introduction au développement Hadoop" s'adresse aux professionnels qui souhaitent comprendre et maîtriser les bases du Big Data. Dans un contexte où la donnée est devenue un atout stratégique, savoir traiter des volumes massifs d'informations est essentiel. Ce cours sur Hadoop propose une approche claire, progressive et adaptée pour entrer dans l'univers de Hadoop et de ses principaux composants.

Une formation Hadoop adaptée aux besoins du Big Data

Le programme couvre l'écosystème complet de Hadoop. Les participants découvrent le rôle du HDFS et de MapReduce, puis apprennent à manipuler des outils incontournables comme Pig, Hive, Flume, Sqoop et YARN. L'objectif est de donner une vision pratique et opérationnelle, pour comprendre comment mettre en place et gérer un environnement Hadoop performant.

Contenu du cours

Module 1 : Introduction à Big Data & l'écosystème Hadoop , Map Reduce et HDFS Module 2 : Installation et configuration de Hadoop Module 3 : Plongée profonde à Mapreduce Module 4 : Résolution de problèmes de graphiques Module 5 : Compréhension détaillée de Pig Module 6 : Compréhension détaillée de Hive Module 7 : (AVRO) Formats des données Module 8 : Introduction à l'architecture Hbase Module 9 : Mise en place du cluster et fonctionnement des Tâches Map Reduce Module 10 : Mapreduce avancé

Lab / Exercices

- Travailler avec le HDFS
- Programme d'écriture WordCount
- Rédaction d'un partitionneur personnalisé
- Mapreduce avec Combiner
- Carte Côté Rejoindre
- Réduire le nombre d'adhésions latérales
- Test unitaire Mapreduce
- Executer Mapreduce en mode "Local Job Runner"

Documentation

- Support de cours numérique inclus

Profils des participants

- Développeurs Big Data
- Analystes de données
- Architectes logiciels
- Professionnels BI et ETL
- Testeurs et intégrateurs Hadoop

Connaissances Préalables

- Pas de prérequis

Objectifs

- Apprendre l'architecture Hadoop
- Rédiger des programmes MapReduce et déployer des clusters Hadoop
- Développer des applications pour les grandes données en utilisant la technologie Hadoop
- Développer des programmes YARN sur la version 2.X de Hadoop
- Travaux sur l'analyse des données de grande taille à l'aide de Hive, Pig et YARN
- Intégrer MapReduce et HBase pour une utilisation et une indexation avancées
- Apprendre les principes fondamentaux du framework Spark et de son fonctionnement
- Comprendre RDD dans Apache Spark
- Apprendre les meilleures pratiques de développement
- Planification des tâches à l'aide d'Oozie

Description

Formation Hadoop

Niveau

Fondamental

Prix de l'inscription en Présentiel (CHF)

2300

Prix de l'inscription en Virtuel (CHF)

2150

Durée (Nombre de Jours)

3

Reference

HDP-01