Introduction to Hadoop Development

Description

Cette formation à la certification des développeurs Apache Hadoop vous aidera à vous faire une idée détaillée de Big Data et de Hadoop. Certains des sujets abordés sont une introduction à l'écosystème Hadoop, la compréhension des HDFS et de MapReduce, y compris l'abstraction de MapReduce. Apprenez à installer, mettre en œuvre les différents composants de Hadoop comme Pig, Hive, Flume, Sqoop et YARN.

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 de logiciels, analystes, professionnels de la BI, de l'ETL et de data warehouse
- Les développeurs, architectes et testeurs de Big Data

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

© Copyright ITTA – Tous les contenus sont protégés par le droit d'auteur.

- Planification des tâches à l'aide d'Oozie

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

© Copyright ITTA – Tous les contenus sont protégés par le droit d’auteur.