

HADOOP : L'ÉCOSYSTÈME

IBD07

1 jour (7h) 720,00 €^{HT}

Objectifs

Disposer d'une vision claire des différents éléments de l'écosystème Hadoop, projet Open Source initié par Google, destiné à faciliter le stockage, l'exploitation et l'analyse de très grands volumes d'information hébergés sur plusieurs machines en réseaux • Être à l'aise avec les concepts et termes propres à Hadoop et comprendre les rôles et les interactions des différents composants d'Hadoop au travers d'exemples significatifs

Profil Stagiaire(s)

chef de projets, développeur, et toute personne souhaitant comprendre les mécanismes Hadoop et le rôle de chaque composant dans un projet Big Data

Pré-requis

Connaissances générales des systèmes d'information

Profil Animateur(s)

Consultant-Formateur expert Big Data

best

PROGRAMME

Introduction

Rappels sur NoSQL

Le théorème CAP

Historique du projet hadoop

Les fonctionnalités : stockage, outils d'extraction, de conversion, ETL, analyse, ...

Exemples de cas d'utilisation sur des grands projets

Les principaux composants : HDFS pour le stockage et YARN pour les calculs

Les distributions et leurs caractéristiques : HortonWorks, Cloudera, MapR, GreenPlum,

Apache, ...

Architecture

Terminologie : NameNode, DataNode, ResourceManager

Rôle et interactions des différents composants

Présentation des outils :

. d'infrastructure (ambari, avro, zookeeper)

. de gestion des données (pig, oozie, falcon, pentaho, sqoop, flume)

. d'interfaçage avec les applications GIS

. de restitution et requêtage : webhdfs, hive, hawq, impala, drill, stinger, tajo, mahout, lucene,

elasticSearch, Kibana

Les architectures connexes : spark, cassandra

Exemples interactifs

Démonstrations sur une architecture Hadoop multi-noeuds.

Mise à disposition d'un environnement pour des exemples de calcul

Travaux pratiques : Recherches dans des données complexes non structurées

Applications

Cas d'usages de Hadoop

Infrastructures hyperconvergées avec les appliances Hadoop

Calculs distribués sur des clusters Hadoop

Contactez-nous

Conseiller formation

09 88 66 10 00

inscriptions@demofr

Nos sessions

28 mars 2022 : Paris / A Distance

3 oct. 2022 : Paris / A Distance

HADOOP : L'ÉCOSYSTÈME

IBD07

1 jour (7h) 720,00 €^{HT}

Méthode pédagogique

Remise d'une documentation pédagogique papier ou numérique pendant le stage
La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions
Les sessions "A distance" sont réalisées avec l'outil de visioconférence Teams, permettant au formateur d'adapter sa pédagogie.

Moyens pédagogiques et techniques de mise en œuvre

Nos formateurs DEMOS sont recrutés conformément aux spécifications mentionnées pour chaque formation. Ce sont des professionnels en activité et/ou des experts dans leur domaine.

Ils utilisent des méthodes et outils appropriés aux formations qu'ils dispensent et adaptent leur pédagogie au public accueilli.

Par ailleurs, nos centres de formation DEMOS sont tous équipés :

- Salles de formation lumineuses, spacieuses, design, ergonomiques, mobiles et équipées d'écrans plasma, de PC portables/fixes si nécessaire à la formation suivie.
- Accès au wifi haut débit sur l'ensemble des lieux
- Espace de co-learning et webcorners
- Cafés, collations et rafraîchissement sont offerts

Et tous sont accessibles aux personnes à mobilité réduite

Le dispositif de suivi et d'évaluation

Pour assurer un suivi individuel, Demos a mis en place 2 types d'évaluation :

- Des évaluations des acquis en cours et en fin de formation

Elles peuvent être faites de différente manière selon le contenu de la formation suivie :

Quiz, exercice pratique, étude de cas, jeu de rôles, mise en situation, soutenance devant un jury pour les formations à finalité certifiante.

- Une évaluation de la satisfaction de chaque stagiaire est réalisée en ligne. Cette évaluation est complétée par l'appréciation du formateur à l'issue de chaque session.