



ORGANISME DE FORMATION AUX TECHNOLOGIES ET METIERS DE L'INFORMATIQUE

Formation Hadoop - Cloudera pour architectes et administrateurs

Mise en oeuvre d'une solution de stockage réparti

N° ACTIVITÉ : 11 92 18558 92

TÉLÉPHONE : 01 85 77 07 07

E-MAIL : inscription@hubformation.com

S'imposant peu à peu comme la solution standard pour les projets Big Data, Hadoop est un framework composé de divers outils permettant de traiter des volumes de données considérables. Base de tout projet de Big Data, le stockage des données revêt une dimension stratégique puisque de son organisation initiale et de sa qualité dépendront la pertinence et l'efficacité de leur exploitation. A l'issue de cette formation, les architectes et administrateurs en charge de la mise en oeuvre et de la gestion d'une solution Big Data seront à même d'utiliser les différents outils fournis par la distribution Cloudera de Hadoop pour mener à bien leur mission.

Objectifs

- | Savoir élaborer une architecture Hadoop
- | Identifier comment Hadoop Distributed File System (HDFS) et Yarn/MapReduce fonctionnent
- | Être en mesure d'exploiter et gérer un cluster Hadoop
- | Apprendre à configurer Hadoop et ses outils
- | Pouvoir mettre en oeuvre un cluster Hadoop
- | Identifier comment gérer la maintenance des noeuds dans un cluster
- | Savoir déplacer et gérer les données dans un cluster
- | Être capable d'administrer un cluster Hadoop

Public

| Administrateurs systèmes et personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement

Prérequis

- | La plate-forme reposant sur une architecture Linux, il est demandé aux participants de savoir naviguer et modifier des fichiers dans cet environnement
- | Une connaissance préliminaire d'Hadoop n'est pas exigée

Programme de la formation

INTRODUCTION

- | Présentation générale d'Hadoop
- | Exemples d'utilisation dans différents secteurs
- | Historique et chiffres clés : Quand parle-t-on de Big Data ?
- | Les possibilités d'implantation de l'infrastructure et les coûts associés

L'ÉCOSYSTÈME D'HADOOP

- | Le système de fichier HDFS
- | Le paradigme MapReduce et l'utilisation à travers YARN
- | Le garant des informations : Zookeeper
- | Le système de stockage de la donnée basé sur le socle : HBase
- | La présentation de l'architecture globale : où installer les services ?
- | La définition du dimensionnement pour un cluster Hadoop
- | Les particularités de MapR

Référence	BIG503
Durée	3 jours (21h)
Tarif	2 080 €HT

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)

- du 17 au 19 juillet 2024
- du 14 au 16 octobre 2024

PARIS

- du 17 au 19 juillet 2024
- du 14 au 16 octobre 2024

[VOIR TOUTES LES DATES](#)

INSTALLATION DU SOCLE D'HADOOP

| Installation et configuration manuelle pas-à-pas sur plusieurs machines de : HDFS, YARN, Zookeeper, HBase

MANIPULATION DES DONNÉES DANS UN CLUSTER HADOOP

| Architecture, objectif et fonctionnement de : Hue, Hive, Pig, Oozie, Sqoop, Flume, Spark

| Mise en pratique sur les différents outils

EXPLOITATION D'UN CLUSTER HADOOP

| Architecture, objectif et fonctionnement du manager : MapR Control System

| Mise en pratique

CONCEPTS D'ARCHITECTURE COMPLEXE

| Architecture en rack : les avantages et fonctionnement et paramétrage du rack awareness

| Scalabilité et performances : le Namenode Federation

| High Availability des serveurs maîtres

| Utilisation des snapshots : sauvegarde et restauration

Méthode pédagogique

Chaque participant travaille sur un poste informatique qui lui est dédié. Un support de cours lui est remis soit en début soit en fin de cours. La théorie est complétée par des cas pratiques ou exercices corrigés et discutés avec le formateur. Le formateur projette une présentation pour animer la formation et reste disponible pour répondre à toutes les questions.

Méthode d'évaluation

Tout au long de la formation, les exercices et mises en situation permettent de valider et contrôler les acquis du stagiaire. En fin de formation, le stagiaire complète un QCM d'auto-évaluation.

Suivre cette formation à distance

Voici les prérequis techniques pour pouvoir suivre le cours à distance :

| Un ordinateur avec webcam, micro, haut-parleur et un navigateur (de préférence Chrome ou Firefox). Un casque n'est pas nécessaire suivant l'environnement.

| Une connexion Internet de type ADSL ou supérieure. Attention, une connexion Internet ne permettant pas, par exemple, de recevoir la télévision par Internet, ne sera pas suffisante, cela engendrera des déconnexions intempestives du stagiaire et dérangera toute la classe.

| Privilégier une connexion filaire plutôt que le Wifi.

| Avoir accès au poste depuis lequel vous suivrez le cours à distance au moins 2 jours avant la formation pour effectuer les tests de connexion préalables.

| Votre numéro de téléphone portable (pour l'envoi du mot de passe d'accès aux supports de cours et pour une messagerie instantanée autre que celle intégrée à la classe virtuelle).

| Selon la formation, une configuration spécifique de votre machine peut être attendue, merci de nous contacter.

| Pour les formations incluant le passage d'une certification la dernière journée, un voucher vous est fourni pour passer l'examen en ligne.

| Pour les formations logiciel (Adobe, Microsoft Office...), il est nécessaire d'avoir le logiciel installé sur votre machine, nous ne fournissons pas de licence ou de version test.

| Horaires identiques au présentiel.

Accessibilité

Les sessions de formation se déroulent sur des sites différents selon les villes ou les dates, merci de nous contacter pour vérifier l'accessibilité aux personnes à mobilité réduite.

Pour tout besoin spécifique (vue, audition...), veuillez nous contacter au 01 85 77 07 07.