Formation Hadoop Cloudera : administration

N° ACTIVITÉ : 11 92 18558 92 | TÉLÉPHONE : 01 85 77 07 07 | E-MAIL : inscription@hubformation.com

Référence	CB032
Durée	3 jours (21h)
Tarif	2 160 €HT
Repas	69 €HT(en option)

Objectifs

| Connaître les principes du framework Hadoop et savoir l'installer et le configurer.
| Maitriser la configuration et la gestion des services avec Cloudera Manager.

Public

| Chefs de projet, administrateurs et toute personne souhaitant mettre en oeuvre un système distribué avec Hadoop.

Prérequis

| Connaissance des commandes des systèmes unix/linux.

Programme de la formation

Introduction
| Les fonctionnalités du framework Hadoop.
| Les différentes versions.
| Distributions : Apache, Cloudera, Hortonworks, EMR, MapR, DSE.
| Spécificités de chaque distribution.
| Les apports de la distribution Cloudera.
| Architecture et principe de fonctionnement.
| Terminologie : NameNode, DataNode, ResourceManager, NodeManager.
| Rôle des différents composants.
| Le projet et les modules : Hadoop Common, HDFS, YARN, Spark, MapReduce
| Oozie, Pig, Hive, HBase, ...

Les outils Hadoop
| Infrastrucure/Mise en oeuvre : Avro, Ambari, Zookeeper, Pig, Tez, Oozie, Falcon, Pentaho
| Vue d'ensemble
| Gestion des données.
| Exemple de sqoop.
| Restitution : webhdfs, hive, Hawq, Mahout,ElasticSearch ..
| Outils complémentaires:
| Spark, SparkQL, SparkMLib, Storm, BigTop, Zebra
| de développement : Cascading, Scalding, Flink/Pachyderm
| d'analyse : RHadoop, Hama, Chukwa, kafka

Installation et configuration
| Présentation de Cloudera Manager
| Trois modes d'installation : local, pseudo-distribué, distribué
| Première installation.
| Mise en oeuvre avec un seul noeud Hadoop.
| Présentation de Cloudera Manager.
| Configuration de l'environnement,étude des fichiers de configuration :
| core-site.xml, hdfs-site.xml, mapred-site.xml, yarn-site.xml et capacity-scheduler.xml
| Création des users pour les daemons hdfs et yarn,droits d'accès sur les exécutables et répertoires.
| Lancement des services.

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)
du 22 au 24 octobre 2025

PARIS
du 22 au 24 octobre 2025

VOIR TOUTES LES DATES

HUB Formation - 14 avenue de l'Opéra - 75001 Paris
01 85 77 07 07 - www.hubformation.com - SIREN 525 287 785

| Démarrage des composants : hdfs, hadoop-daemon, yarn-daemon, etc ..
| Gestion de la grappe, différentes méthodes :
| ligne de commandes, API Rest, serveur http intégré, APIS natives
| Exemples en ligne de commandes avec hdfs, yarn, mapred
| Présentation des fonctions offertes par le serveur http
| Travaux pratiques :
| Organisation et configuration d'une grappe hadoop avec Cloudera Manager
| Traitement de données.
| Requêtage SQL avec Impala.

Administration Hadoop
| Outils complémentaires à yarn et hdfs : jConsole, jconsole yarn
| Exemples sur le suivi de charges, l'analyse des journaux.
| Principe de gestion des noeuds, accès JMX.
| Travaux pratiques :
| mise en oeuvre d'un client JMX
| Administration HDFS :
| présentation des outils de stockage des fichiers, fsck, dfsadmin
| Mise en oeuvre sur des exemples simples de récupération de fichiers
| Gestion centralisée de caches avec Cacheadmin

Haute disponibilité
| Mise en place de la haute disponibilité sur une distribution Cloudera. Atelier : passage d'un système HDFS en mode HA Explication/démonstration d'une fédération de cluster Hadoop

Sécurité
| Mécanismes de sécurité et mise en oeuvre pratique :
| Activation de la sécurité avec Kerberos dans core-site.xml, et dans hdfs-site.xml pour les NameNode et DataNode. Sécurisation de yarn avec la mise en oeuvre d'un proxy et d'un Linux Container Executor.

Exploitation
| Installation d'une grappe Hadoop. Lancement des services.
| Principe de la supervision des éléments par le NodeManager.
| Présentation de Ganglia,Kibana
| Travaux pratiques :
| Visualisation des alertes en cas d'indisponibilité d'un noeud.
| Configuration des logs avec log4j.

Méthode pédagogique

Les travaux pratiques sont réalisés sur une distribution Hadoop Cloudera.

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

HUB Formation - 14 avenue de l'Opéra - 75001 Paris
01 85 77 07 07 - www.hubformation.com - SIREN 525 287 785

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.