

FORMATION HADOOP

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

PROGRAMME

1/ Étude des fichiers de configurations

- Gestion des utilisateurs pour les daemons hdfs et yarn
- Droits d'accès sur les exécutables et répertoires
- Architecture et gestion des services généraux Hadoop
- HDFS
- YARN
- MapReduce
- HBase

2/ Monitoring du cluster Hadoop

- Suivi de charges journaux (jConsole)
- Gestion des noeuds Accès JM
- Mise en oeuvre d'un client JMX
- Administration HDFS
- Stockage des fichiers, fsck, dfsadmin
- Gestion centralisée de caches avec Cacheadmin

3/ Sécurité

- Activation de la sécurité avec Kerberos dans core-site.xml et dans hdfs-site.xml pour les NameNode et DataNode.
- Gestion de la sécurité avec Apache Sentry

4/ Exploitation

- Supervision des éléments par le NodeManage
- Monitoring graphique avec Ambari, Kibana, Cloudera Manager
- Visualisation des alertes en cas d'indisponibilité d'un noeud
- Configuration des logs avec log4j

5/ HDFS

- Architecture
- Commandes SHELL

6/ Mapreduce

- Architecture MapReduce
- Exécuter un code MapReduce

PRIX (INTER-ENTREPRISE) : 1990 euros

21h jours

←→ HAD

Vous souhaitez organiser cette formation dans vos locaux ?

Demandez Houily au
01 84 25 05 10

OBJECTIFS

- Appréhender l'architecture d'un système Hadoop.
- Détailler les services principaux, leur configuration, la sécurité au sein d'un cluster mais aussi son exploitation.
- Passer en revue les différentes composantes logicielles pour manipuler les données big data (MapReduce, Pig, Hive Sqoop).

PRE-REQUIS

- La connaissance du SQL est un plus
-

PUBLIC CONCERNE

- Directeurs techniques
- Chefs de projets
- Architectes
- Consultants
- DBA
- Développeurs d'applications

DATES INTER-ENTREPRISES

Delais d'entree : sans

Paris et à distance

16/01/2023, 20/02/2023, 03/04/2023, 05/06/2023, 07/08/2023, 09/10/2023, 18/12/2023,

Si vous souhaitez organiser cette formation a une autre date contactez-nous.

NOUS CONTACTER

Openska
 21 rue Louise Weiss
 75013 Paris

Tel : 01 84 25 05 10
 Tel : 01 84 17 44 76
www.openska.com



7/ HUE

- Introduction
- Fonctionnalités et utilisation
- HBASE
- Architecture
- Commandes SHELL
- Création base de données, tables, familles
- Interrogation des données

8/ HIVE

- Architecture
- Méthodes d'accès à Hive
- HiveQL
- Création de bases de données, tables, vues
- Interrogation des données grâce au HiveQL
- Manipulations de fonctions définies par l'utilisateur (UDF)
- Partitionner vos données
- Archivage de vos données

9/ PIG

- Introduction
- Méthodes d'exécution
- Pig Latin
- Communication entre Pig et Hive

10/ SQOOP

- Introduction
- Pour quels usages ?
- Méthodes d'utilisation
- Import et export des données

11/ OOZIE

- Introduction
- Planification de workflows avec paramètres

ELEMENTS COMPLÉMENTAIRES À TRANSMETTRE IMPÉRATIVEMENT À VOTRE OPCO POUR VOTRE DEMANDE DE PRISE EN CHARGE

HORAIRES

- 9h30-13h
- 14h-17h30

ACCÈS HANDICAP

Oui, Consultez notre référent interne.

MODALITÉS D'ÉVALUATION

L'évaluation des acquis se fait tout au long de la formation au travers de multiples exercices et mise en situation

MODALITÉS DE SUIVI DE L'EXÉCUTION

- Contrôle systématique des présences par demi-journée
- Attestation de présence remis à chaque participant

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la

pédagogie, et ce pour chaque formation qu'ils animent. Ils ont majoritairement cinq à dix années d'expérience dans leur domaine de compétences et ont une grande pratique en entreprise

MOYENS PÉDAGOGIQUES

Les supports pédagogiques sont imprimés par nos soins et transmis à chacun des stagiaires. Les supports sont aussi remis au format électronique aux participants. Ce support est projeté via un vidéo projecteur afin d'animer la formation. Des exercices d'application ou études de cas sont prévus afin de valider les acquis des stagiaires.

MOYENS D'ACCOMPAGNEMENT

8 personnes maximum, formation animée par un formateur expert sur le sujet.
Mise en situation : Pédagogie active et participative.
Apport théorique et méthodologique.
Etude de cas. Atelier pratique.

NATURE DES TRAVAUX DEMANDÉS AU STAGIAIRE

Un exercice est réalisé par le participant à la fin de chaque chapitre. Le participant dispose de 20 min avant de passer à la correction avec le formateur.

MODALITÉS TECHNIQUES EN CAS DE PROBLÈMES

Les connexions et installations d'outil peuvent être testées en amont de la formation avec le formateur et les stagiaires ou le donneur d'ordre. Ce test permet d'éviter tout accident technique lié aux outils de connexion à distance.

Le formateur prend la main sur le poste du participant en cas de difficulté durant la formation.



ORGANISME DE FORMATION RÉFÉRENCÉ SOUS LE NUMÉRO : 117 555 432 75