

- *Comment définir la taille d'un cluster et comment le déployer*
- *Comment déployer un cluster pour la première fois*
- *Comment configurer Hadoop et les Framework associés*
- *Comment gérer en permanence la maintenance des nœuds dans le cluster*
- *Comment équilibrer et optimiser la performance d'un cluster*
- *Comment déplacer et gérer les données dans le cluster*
- *Comment intégrer des statuts et des états de santé dans l'outil de monitoring existant*
- *Comment ajouter et enlever des nœuds de données*
- *Comment mettre en place une haute solution de disponibilité*
- *Les meilleures pratiques pour déployer un cluster Hadoop*

3**Prix : 2?100 € € / HT****PRÉ-REQUIS**

- Cette formation Hadoop – Cloudera pour architectes et administrateurs nécessite de savoir naviguer et modifier des fichiers dans Linux
- Une connaissance préliminaire d'Hadoop n'est pas exigée

MODALITÉS ET DÉLAIS D'ACCÈS**ATTESTATION OBTENUE****EFFECTIF DE LA FORMATION****CERTIFICATION****MODALITÉ PÉDAGOGIQUE**

Cours dispensé en mode présentiel avec une alternance d'apports théoriques et méthodologiques, et de mises en situations pratiques

PROCHAINES SESSIONS

Nous consulter.
Nous consulter.
Nous consulter.

OUTILS PÉDAGOGIQUES**MODALITÉS D'ÉVALUATION****MODALITÉS DE FINANCEMENT****MODALITÉS ET DÉLAIS D'ACCÈS****OBJECTIFS PÉDAGOGIQUES****ACCESSIBILITÉ****LES POINTS FORTS DE LA FORMATION****PROGRAMMES DE HADOOP – CLOUDERA POUR ARCHITECTES ET ADMINISTRATEURS****FONDEMENT, PLANIFICATION ET INSTALLATION**

- Introduction à la plateforme de données Hadoop 2.0
- Stockage Hadoop : Architecture HDFS
- Prérequis pour l'installation
- Gestion HDP : Ambari
- Ambari et la ligne de commande
- Hadoop Operating System (YARN) & MapReduce

CONFIGURATION ET GESTION DE DONNEES

- Service de configuration
- Configuration HDFS
- Configuration de Hadoop Operating System (YARN) & MapReduce
- Configuration HBase
- Configuration ZooKeeper
- Configuration Schedulers
- Intégrité des données
- Extract-Load-Transform (ELT) Data Movement
- Copie de données entre les clusters

MANAGEMENT DES DONNÉES / HORTONWORKS DATA PLATFORM (HDP) 2.0 OPERATIONS

- HDFS service Web
- Apache Hive Data Warehouse
- Transfert de données avec Sqoop
- Mouvement de données log avec Flume
- HDFS NFS Gateway
- Workflow management : Oozie
- Management du cycle de vie des données avec Falcon
- Monitoring HDP 2.0 Services
- Mise en service ou hors service des nœuds et services

HORTONWORKS DATA PLATFORM (HDP) 2.0

- Rack Awareness and Topology
- NameNode Federation Architecture
- NameNode High-Availability (HA) Architecture
- Backup & Recovery
- Security