

Hadoop - HBase mise en oeuvre et administration

Installer et configurer la base de données HBase

 Présentiel ou en classe à distance



2 jours (14 h)

Prix inter : 1.600,00 € HT

Réf.: BD514

Dans un contexte où les organisations manipulent des volumes massifs de données, **Apache HBase** s'impose comme une solution incontournable pour le **stockage distribué et scalable**. Conçue pour les environnements **Big Data basés sur Hadoop**, cette formation permet de comprendre en profondeur l'**architecture HBase**, son fonctionnement et ses apports en termes de performance, de haute disponibilité et de gestion de données massives.

Les participants découvrent comment **installer, configurer et administrer HBase en mode distribué**, depuis la mise en place de clusters jusqu'à l'optimisation des requêtes. L'apprentissage couvre également le rôle essentiel de **Zookeeper**, la gestion de la consistance, ainsi que l'exploitation des outils natifs comme **HBase Shell, hbck ou region splitter**.

A qui s'adresse cette formation ?



Pour qui

- Chefs de projet, administrateurs
- Toute personne souhaitant stocker des données avec Hbase



Prérequis

- Connaissance des principes de base Hadoop et des bases de données

Programme

1 - Introduction

- Rappels rapides sur l'écosystème Hadoop
- Les fonctionnalités du Hadoop Common, HDFS, YARN, Spark, MapReduce.
- Présentation des principaux modules : Hadoop Common, HDFS, YARN, Spark, MapReduce
- Présentation HBase
- Historique
- Lien avec HDFS

2 - Architecture et fonctionnement de HBase

- Architecture et fonctionnement de HBase : table, région, ligne, famille de colonnes, cellules, espace de nommage, ...
- Fonctionnalités : failover automatique, sharding, interface avec des jobs MapReduce
- Rôles des composants : HBase Master Node, Region Master, clients HBase
- Haute disponibilité et cohérence des données. Présentation du rôle de Zookeeper

Atelier
Définir une architecture HBase adaptée à des contraintes d'utilisation

3 - Les apports d'HBase pour le stockage distribué

- Comprendre le format des données dans HBase
- Comparaison avec d'autres bases de type clé/valeur

- Découverte des interfaces disponibles pour l'exploitation des données
- Panorama des outils HBase : hbase pe et hbase lrt (tests de performance) hbase shell (exploitation et gestion)

Atelier

Manipuler une base HBase avec hbase shell

4 - Installation

- Choix des packages
- Installation et configuration en mode distribué
- Installation en mode distribué
- Intégration avec HDFS dans un cluster Hadoop
- Test de connexion avec hbase shell

Atelier

Déployer une grappe de serveurs HBase en mode distribué

Effectuer des interrogations via le serveur HTTP intégré

5 - Configuration

- Fonctionnement en mode distribué : organisation des démons (HMaster, HRegionServer, Zookeeper)
- Gestion de la cohérence et de la consistance des données
- Suivi et contrôle

Atelier

Utiliser les outils d'exploitation (hbck, hfile...)

Mettre en oeuvre des splits sur des tables réparties avec regionsplitter



Les objectifs de la formation

- Comprendre le fonctionnement de Hadoop
- Comprendre l'architecture de HBase et ses principaux composants
- Identifier les apports d'HBase en termes de stockage distribué des données
- Mettre en place l'installation du système
- Savoir mettre en place une configuration distribuée



Evaluation

- Cette formation fait l'objet d'une évaluation formative.



Les points forts de la formation

- Une pédagogie basée sur l'alternance de phases théoriques et de mises en pratique qui permet aux participants d'acquérir une première expérience concrète du stockage de données en environnement HBase
- Des échanges avec les animateurs sur les meilleures pratiques pour garantir la disponibilité des bases et leurs performances



Dates et villes 2026 - Référence BD514



Dernières places disponibles



Session garantie

A distance

du 5 mars au 6 mars

du 18 juin au 19 juin

du 15 oct. au 16 oct.

du 3 déc. au 4 déc.

Paris

du 5 mars au 6 mars

du 18 juin au 19 juin

du 15 oct. au 16 oct.

du 3 déc. au 4 déc.