

NoSQL - Apache Cassandra, mise en oeuvre et administration

Installer et administrer des bases de données sous la solution NoSQL Apache Cassandra

 Présentiel ou en classe à distance



3 jours (21 h)

Prix inter : 2.050,00 € HT
Forfait intra : 8.350,00 € HT

Réf.: BD513



Idéal en
Distanciel

Système de gestion de base de données NoSQL, Cassandra a été conçu pour gérer des quantités massives de données réparties sur plusieurs serveurs tout en assurant une disponibilité maximale des données qu'elle héberge. Du fait de sa structure décentralisée, Cassandra offre nativement une tolérance aux pannes individuelles. Enfin, supportant l'intégration avec Hadoop, Cassandra pourrait se révéler une plate-forme d'avenir pour les projets Big Data. Les participants à cette formation acquerront, en 3 jours, les compétences nécessaires pour mettre en oeuvre une solution de données avec Cassandra.

A qui s'adresse cette formation ?



Pour qui

- Chefs de projet, administrateurs de bases de données, développeurs, architectes



Prérequis

- Avoir des connaissances de base de l'administration de Linux ou Windows et d'une base de données relationnelles
- **Disposez-vous des connaissances nécessaires pour suivre cette formation ? Testez-vous !**

Programme

1 - Introduction

- Historique, fonctionnalités de Cassandra, licence
- Format des données, "Key-value", traitement de volumes importants
- Haute disponibilité, système réparti de bases de données, ...

2 - Installation et configuration

- prérequis
- Plates-formes supportées
- Étude du fichier de configuration : conf/cassandra.yaml
- Répertoire de travail, de stockage des données, gestion de la mémoire
- Démarrage d'un noeud et test de l'interface cliente cqlsh

3 - CQL

- Commandes de base : connexion au système de base de données
- Création de colonnes, insertion, modification recherche
- Le CQL : Cassandra Query Language
- Exécution de scripts
- Comment écrire des requêtes ?
- Approches
- Limitations du CQL

4 - Gestion de la grappe

- Principe
- Configuration des noeuds
- Notion de bootstrapping et de token
- Paramètres de démarrage des noeuds
- Réplication : topologie du réseau et EndpointSnitch
- Stratégie de réplication
- Ajout de noeuds, suppression
- Architecture de stockage mémoire et disque dur, gestion des tombstones, bloom-filter
- Travaux pratiques : mise en place d'une configuration de production (multi-datacenters, multi-racks)

5 - Supervision

- Prometheus : installation, lancement et découverte
- Supervision avec nodetool
- Principe des accès JMX
- Exports JMX vers des outils de supervision
- Démonstration avec Prometheus et Grafana

6 - Exploitation

- Gestion des noeuds Cassandra
- Sauvegardes, snapshots et export au format JSON
- Principe de cohérence, hinted_handoff, digest request et read repair

7 - Support Hadoop

- Principe de map/reduce
- Implémentation Hadoop et implémentation Hadoop/Cassandra

8 - Support Spark

- Description rapide de l'architecture Spark
- Mise en oeuvre depuis Cassandra
- Exécution d'application Spark s'appuyant sur une grappe Cassandra

9 - Performance

- Outil de test de performance Cassandra-stress
- Mise en place d'un plan de stress et paramétrage

10 - Sécurité

- Paramétrage, authentification, sécurisation de la base system_auth
- Gestion des rôles et des autorisations



Les objectifs de la formation

- Découvrir l'architecture de NoSQL Apache Cassandra et ses apports par rapport aux autres solutions
- Installer et configurer le SGBD NoSQL Apache Cassandra
- Administrer et sécuriser un cluster Cassandra
- Appréhender le CQL (Cassandra Query Language)
- Créer une base de données et manipuler ses objets
- Connaître la notion de grappe au sein de la base de données



Evaluation

- Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.



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 avec Cassandra.
- Des échanges avec les animateurs sur les meilleures pratiques pour garantir la disponibilité des bases et leurs performances.



Dates et villes 2026 - Référence BD513



Dernières places disponibles



Session garantie

Paris

du 12 janv. au 14 janv.

du 9 mars au 11 mars

du 4 mai au 6 mai

du 6 juil. au 8 juil.

du 21 sept. au 23 sept.

du 28 sept. au 30 sept.

du 7 déc. au 9 déc.

A distance

du 12 janv. au 14 janv.

du 9 mars au 11 mars

du 4 mai au 6 mai

du 6 juil. au 8 juil.

du 21 sept. au 23 sept.

du 28 sept. au 30 sept.

du 7 déc. au 9 déc.

Toulouse

du 12 janv. au 14 janv.

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Marseille

du 12 janv. au 14 janv.

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Aix-en-Provence

du 12 janv. au 14 janv.

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Lille

du 12 janv. au 14 janv.

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Nantes

du 9 mars au 11 mars

du 6 juil. au 8 juil.

du 7 déc. au 9 déc.

Rennes

du 9 mars au 11 mars

du 6 juil. au 8 juil.

du 7 déc. au 9 déc.

Lyon

du 9 mars au 11 mars

du 6 juil. au 8 juil.

du 7 déc. au 9 déc.

Rouen

du 9 mars au 11 mars

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Sophia Antipolis

du 9 mars au 11 mars

du 4 mai au 6 mai

du 21 sept. au 23 sept.

Bordeaux

du 9 mars au 11 mars

du 6 juil. au 8 juil.

du 7 déc. au 9 déc.

Strasbourg

du 9 mars au 11 mars

du 4 mai au 6 mai

du 21 sept. au 23 sept.