



Big Data - Administration et infrastructure

## NoSQL, Apache Cassandra - Mise en oeuvre et administration

3 jours (21h00) | ★★★★★ 4/5 | NOSQL-CAS | Évaluation qualitative de fin de stage | Formation délivrée en présentiel ou distanciel

Formations Informatique > Big Data > Big Data - Administration et infrastructure

Contenu mis à jour le 13/10/2023. Document téléchargé le 23/05/2024.

### Objectifs de formation

A l'issue de cette formation, vous serez capable de :

- Installer et administrer des bases de données sous la solution NoSQL Apache Cassandra
- Décrire 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
- Définir le CQL (Cassandra Query Language)
- Créer une base de données et manipuler ses objets
- Expliquer la notion de grappe au sein de la base de données.

### Modalités, méthodes et moyens pédagogiques

Formation délivrée en présentiel ou distanciel\* (blended-learning, e-learning, classe virtuelle, présentiel à distance).

Le formateur alterne entre méthode\*\* démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation).

Variables suivant les formations, les moyens pédagogiques mis en oeuvre sont :

- Ordinateurs Mac ou PC (sauf pour certains cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel)
- Environnements de formation installés sur les postes de travail ou en ligne
- Supports de cours et exercices

En cas de formation intra sur site externe à M2i, le client s'assure et s'engage également à avoir toutes les ressources matérielles pédagogiques nécessaires (équipements informatiques...) au bon déroulement de l'action de formation visée conformément aux prérequis indiqués dans le programme de formation communiqué.

\* nous consulter pour la faisabilité en distanciel

\*\* ratio variable selon le cours suivi

## Prérequis

Avoir des connaissances de base de l'administration Linux ou Windows et d'une base de données relationnelle.

## Public concerné

Chefs de projets, administrateurs de bases de données, développeurs, architectes.

## Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence émarginée par demi-journée par les stagiaires et le formateur.

# Programme

## Découvrir l'architecture de NoSQL Apache Cassandra et ses apports par rapport aux autres solutions

- Introduction
- Historique, fonctionnalités de Cassandra, licence
- Format des données, "key-value", traitement de volumes importants, haute disponibilité, système réparti de base de données...

## Installer et configurer le SGBD NoSQL Apache Cassandra

- Installation et configuration
- Prérequis
- Plateformes supportées
- Etude du fichier de configuration : conf/cassandra.yaml
- Répertoire de travail, de stockage des données, gestion de la mémoire

## Exemples de travaux pratiques (à titre indicatif)

- Ateliers : démarrage d'un noeud et test de l'interface cliente "cqlsh"

## Appréhender le CQL (Cassandra Query Language)

- Commandes de base :
  - Connexion au système de base de données
  - Création de colonnes
  - Insertion
  - Modification
  - Recherche
- Le CQL : Cassandra Query Language
- Limitations du CQL

## Créer une base de données et manipuler ses objets

- Utilisation de Cassandra
- Création de bases et interrogation avec CQL
- Définition de la notion de consistance
- Eléments en jeu :

- Commit.log
- Memtable
- Quorum
- Comment écrire des requêtes ?
  - Approches

### **Exemples de travaux pratiques (à titre indicatif)**

- Ateliers : premiers pas avec une base de données Cassandra pré-chargée et mise à disposition sur l'infrastructure de travaux pratiques

## **Connaître la notion de grappe au sein de la base de données**

- Gestion de la grappe
- Principe et configuration des noeuds
- Notion de bootstrapping et de token
- Paramètres de démarrage des noeuds
- Réplication : topologie du réseau et "endpoint snitch"
- Stratégie de réplication
- Méthode d'ajout de noeuds et suppression
- Architecture de stockage mémoire et disque dur, gestion des tombstones, bloom filter

### **Exemple de travaux pratiques (à titre indicatif)**

- Atelier : mise en place d'une configuration de production (multi Data Centers, multi-racks)

## **Administrer et sécuriser un cluster Cassandra**

- Exploitation
- Gestion des noeuds Cassandra
- Sauvegardes, snapshots et export au format JSON
- Principe de cohérence, hinted\_handoff, digest request et read repair
- Sécurité
- Gestion des rôles et des autorisations sur une application standard

### **Exemples de travaux pratiques (à titre indicatif)**

- Ateliers : paramétrage, authentification et sécurisation de la base system\_auth

## **Support Hadoop et Spark**

- Principe de map/reduce
- Implémentation Hadoop et intégration Hadoop / Cassandra
- Support Spark : description rapide de l'architecture Spark

### **Exemples de travaux pratiques (à titre indicatif)**

- Ateliers :
  - Mise en oeuvre depuis Cassandra
  - Exécution d'application Spark s'appuyant sur une grappe Cassandra

## **Supervision et performances**

- Prometheus : apports et particularité de Prometheus pour la supervision Cassandra
- Supervision avec nodetool
- Principe des accès JMX, exports JMX vers des outils de supervision
- Performance : présentation de l'outil de test de performance Cassandra-stress

### **Exemples de travaux pratiques (à titre indicatif)**

- Ateliers :
  - Démonstration avec Prometheus et Grafana
  - Mise en place d'un plan de stress et paramétrage

Le contenu de ce programme peut faire l'objet d'adaptation selon les niveaux, prérequis et besoins des apprenants.

## **Modalités d'évaluation des acquis**

- Analyse des résultats individuels produits par chaque participant lors des travaux pratiques, correction au fur et à mesure du déroulé de la formation
- Réalisation d'un test complet sur un projet global de fin de formation, et analyse par le consultant formateur
- Un QCM individuel d'évaluation des acquis est proposé en fin de chaque session

## **Accessibilité de la formation**

Le groupe M2i s'engage pour faciliter l'accessibilité de ses formations. Les détails de l'accueil des personnes en situation de handicap sont consultables sur la page Accueil et Handicap.

## **Modalités et délais d'accès à la formation**

Les formations M2i sont disponibles selon les modalités proposées sur la page programme. Les inscriptions sont possibles jusqu'à 48 heures ouvrées avant le début de la formation. Dans le cas d'une formation financée par le CPF, ce délai est porté à 11 jours ouvrés.