

Mis à jour le 11/04/2024

S'inscrire

# Formation Cassandra

3 jours (21 heures)

## Présentation

**Cassandra** est une base de données distribuée NoSQL connue pour ses capacités à savoir gérer sans défaillance de très grosse quantité de données structurées, notamment utilisée à travers Spark dans les nouvelles architectures SMACK (Spark / Mesos / Akka / Cassandra / Kafka).

Ce système de gestion de base de données open-source est **très scalable** et bénéficie d'une architecture peer to peer pour atteindre un faible risque d'échec.

De plus, cet outil bénéficie de grandes possibilités d'analytics : recherche intégrée Solr, batch analytics avec Hadoop ou encore sa méthode d'analytics en temps réel.

Cette formation Cassandra vous donnera les connaissances et l'expérimentation pratique pour mettre en œuvre et administrer Apache Cassandra au quotidien. Vous découvrirez comment modéliser vos données, effectuer vos tâches de monitoring et optimiser l'administration de Cassandra à travers l'apprentissage des meilleures pratiques.

Comme toutes nos formations, celle-ci vous présentera la dernière version stable en date (**Cassandra 5**).

## Objectifs

- Comprendre l'architecture et le fonctionnement du SGBD NoSQL Apache Cassandra
- Installer et configurer un cluster Cassandra
- Créer une base de données et manipuler ses objets
- Découvrir les principes de l'optimisation et du développement
- Se familiariser avec les outils d'administration et monitoring Cassandra

## Public visé

- Développeurs
- Lead Developer
- Architectes
- Administrateurs systèmes

## Pré-requis

Connaissances de base d'un système Unix (connaissance de l'utilisation du terminal) ainsi qu'avec le langage Java.

## Programme de la formation Cassandra

### Introduction à Cassandra & NOSQL (Jour 1 - Matin)

- Architecture : d'hier à aujourd'hui "BIG DATA" ('SQL' => 'NOSQL')
- Polyglot Persistence\*. Typical architecture (hier et aujourd'hui)
- Introduction à Cassandra (Concepts, ACID => CAP)
- Cassandra Architecture (Comprendre son fonctionnement pour bien l'exploiter: commit log, data stockage, services, flux)
- Répartition & Distribution
- Consistency & Réplication (Hint, Repair, Tombstones)

### Data modeling avec Cassandra : Become a super modeler ! (Jour 1 - Après-midi)

- Le keyspace
- Big Table : Clé / Valeur, la clé est dans la key :)
- Le Relationnel en NOSQL => La dé-normalisation
- Don't be afraid of writes !
- Wide Row (Clustering columns)
- Counters
- Do you need a Transaction?
- Options des tables (compactions, gc graces, ttl ...)
- Best practises : ce à quoi il faut penser (row size, partitions, clustering or not clustering columns)
- TP (Use-cases & Discussions)

### Monitoring : quoi monitorer ? (Jour 2 - Matin)

- Throughput, read and write requests
- Latency, read and write latency
- Disk usage, disk space on each node
- Garbage collection frequency and duration
- Errors and overruns, especially unavailable exceptions which indicate failed requests due to unavailability of nodes in the cluster
- Outils (nodetool, JMX, Datastax OpsCenter (DSE) ...)

### Cassandra Tuning (Jour 2 - Après-midi)

- Cassandra pré-requis (hardware)
- cassandra.yaml (propriétés)
- Propriétés majeures (tuning)
- JVM options (tuning)
- Use-cases & Discussions

## Best Practices (Jour 3 - Matin)

- Java Driver (description, configuration, instanciation, etc.)
- Writes : best practises
- Reads : best practises
- Gestion des erreurs (exception)
- TP
- Use-cases & Discussions

## Administration (Jour 3 - Après-midi)

- Object Mapping
- Tests d'intégration (Cassandra embedded)
- TP
- Administration (overview) d'un cluster Cassandra : possibilité, les choses à faire, les choses à ne pas faire
- Retours d'expérience, pour aller plus loin (associer d'autres technologies ?)

## Pour aller plus loin

## Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

## Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

## Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

## Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

## Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

## Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.