



NoSQL : Apache Cassandra, mise en oeuvre et administration

CB010

Durée: 3 jours

2 080 €

5 au 7 février
3 au 5 juin

16 au 18 septembre
9 au 11 décembre

Public :

Chefs de projet, Administrateurs de bases de données, Développeurs, Architectes...

Objectifs :

A l'issue de la formation, le stagiaire sera capable d'installer et d'administrer des bases de données sous la solution NoSQL Apache Cassandra

Connaissances préalables nécessaires :

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

Objectifs pédagogiques :

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
Connaitre la notion de grappe au sein de la base de données

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.

Atelier : démarrage d'un noeud et test de l'interface cliente `cqlsh`.



Phirio

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.

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

Connaitre la notion de grappe au sein de la base de données

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.
Méthode d'ajout de noeuds et suppression.
Architecture de stockage mémoire et disque dur, gestion des tombstones, bloom-filter

Atelier : mise en place d'une configuration de production (multi-datcenters, 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

Atelier : paramétrage, authentification et sécurisation de la base system_auth.

Gestion des rôles et des autorisations sur une application standard.

Support Hadoop et Spark

Principe de map/reduce. Implémentation Hadoop et intégration Hadoop/Cassandra.
Support Spark :
Description rapide de l'architecture spark.

Atelier : Mise en oeuvre depuis Cassandra. Execution d'application Spark s'appuyant sur une grappe Cassandra.



Phirio

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.

Atelier : démonstration avec Prométheus et Grafana.

Performance :
Présentation de l'outil de test de performance Cassandra-stress

Atelier : mise en place d'un plan de stress et paramétrage.