



Big Data avec CASSANDRA

VOUS ÊTES

Chef de projet, gestionnaire de bases de données.

VOS OBJECTIFS

Connaître les apports de Cassandra.
Savoir l'installer et le configurer.
Maîtriser le CQL.
Maîtriser et sécuriser un cluster Cassandra.
Effectuer des calculs distribués avec Spark.

VOTRE FORMATION



DURÉE : 3 jours
21 heures



PROCHAINE SESSION :
Du 11 au 13 mai 2020



LIEU : En distanciel



PRIX : 2 172 €
net de taxes

PRÉ-REQUIS

Connaissances générales sur les bases de données.

MODALITÉS

La formation est accessible à distance en classe virtuelle : accès à l'infrastructure de travaux pratiques, machines physiques, outils pédagogiques, échanges avec le formateur

De 4 à 12 participants

Financement éligible au FNE Formation pour tout salarié d'entreprise en activité partielle

VOTRE CONTACT :



Andrea FALLOURD
Conseillère en formation
06 74 51 44 97
afalourd@itescia.fr

ITESCIA - Campus de Pontoise
8 rue Pierre de Coubertin
95300 PONTOISE

www.itescia.fr



VOTRE PROGRAMME

Introduction

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

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.
Démarrage d'un nœud et test de l'interface cliente Cqlsh.

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

Gestion de la grappe

Principe. Configuration des nœuds.
Notion de Bootstrapping et de Token.
Paramètres de démarrage des nœuds
Réplication : topologie du réseau et EndpointSnitch. Stratégie de réplication .
Ajout de nœuds, 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)

Supervision

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

Exploitation

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

Support Hadoop

Principe de MapReduce. Implémentation Hadoop et intégration Hadoop / Cassandra.

Support Spark

Description rapide de l'architecture Spark. Mise en œuvre depuis Cassandra.
Exécution de travaux Spark s'appuyant sur une grappe Cassandra.

Performance

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

Sécurité

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

Rejoignez nos réseaux sociaux

