

NoSQL, Apache Cassandra, mise en œuvre et administration

Référence : PYCB010

Durée : 3 jours (21 heures)

Certification : Aucune

Connaissances préalables

- Connaissances générales sur les bases de données

Profil des stagiaires

- Chefs de projet
- Gestionnaires de bases de données

Objectifs

- Connaître les apports de Cassandra, savoir l'installer et le configurer

Certification préparée

- Aucune

Méthodes pédagogiques

- Mise à disposition d'un poste de travail par participant
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions

Formateur

- Consultant-Formateur expert Bigdata

Méthodes d'évaluation des acquis

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation des compétences acquises envoyée au stagiaire
- Attestation de fin de stage adressée avec la facture

Contenu du cours

1. Introduction

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

2. Installation et configuration

- Prérequis. Plate-formes supportées. Etude 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

4. Gestion de la grappe

- Principe. Préparation du premier noeud : adresse d'écoute
- Configuration de nouveaux noeuds. Notion de bootstrapping et de token
- Paramètres listen_address et rpc_address
- Réplication : topologie du réseau et EndpointSnitch. Stratégie de réplication. Ajout de noeuds, suppression
- Cassandra dans un cloud. Mise en oeuvre avec OpenStack

5. Supervision

- OpsCenter : installation, lancement. Utilisation de base
- Supervision avec nodetool cfstats, ou export JMX vers des outils de supervision comme Nagios

6. Exploitation

- Sauvegardes. Import/export au format JSON

7. Support Hadoop

- Principe de MapReduce. Implémentation Hadoop. Mise en oeuvre depuis Cassandra

8. Support Spark

- Description rapide de l'architecture spark. Mise en oeuvre depuis Cassandra
- Execution de travaux Spark s'appuyant sur une grappe Cassandra

Notre référent handicap se tient à votre disposition au [01.71.19.70.30](tel:0171197030) ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.