

Introduction à NoSQL

Référence : PYCB002

Durée : 1 jour

Certification : **Aucune**

CONNAISSANCES PREALABLES

- Connaissance générale des systèmes d'informations et des bases de données.

PROFIL DES STAGIAIRES

- Experts en bases de données, chefs de projet et toute personne souhaitant comprendre le fonctionnement et les apports des bases NoSQL.

OBJECTIFS

- Connaître les caractéristiques techniques des bases de données NoSQL, les différentes solutions disponibles.
- Identifier les critères de choix.

CERTIFICATION PREPAREE

Aucune

METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

FORMATEUR

Consultant-Formateur expert Bigdata

METHODE D'EVALUATION DES ACQUIS

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

CONTENU DU COURS

Introduction

- Origine des bases de données, les notions de transaction les SGBD, la standardisation SQL
- L'arrivée de nouveaux besoins : volumes importants liés aux technologies et aux nouveaux usages, traitements optimisés de flux de données au fil de l'eau
- Développement des techniques sur différents aspects : stockage, indexation/recherche, calcul
- Définition ETL : Extract Transform Load

Caractéristiques NoSQL

- Structure de données proches des utilisateurs, développeurs : sérialisation, tables de hachage, JSON
- Priorité au traitement du côté client
- Protocoles d'accès aux données, interfaces depuis les langages classiques
- Données structurées et non structurées, documents, images

- Stockage réparti : réplication, sharding, gossip protocol, hachage, ..
- Parallélisation des traitements : implémentation de MapReduce
- Cohérence des données et gestion des accès concurrents : "eventual consistency" et multi-version concurrency control

Principaux acteurs

- Les solutions NoSQL et leurs choix techniques : CouchDB, MongoDB, Cassandra, HBase (Hadoop), ElasticSearch, ..
- Démonstrations avec Cassandra et couchDB
- Critères de choix

Mise en oeuvre

- Points à vérifier : méthode d'utilisation des données, format de stockage JSON, XML, choix de la clé, notion de clé composite, ...

- Aspects matériels, besoins en mémoire, disques, répartition, ..
- Import des données : outils et méthodes selon les moteurs NoSQL