

Bases de données NoSQL, enjeux et solutions comment choisir la solution NoSQL adaptée à vos besoins

Séminaire de 2 jours - 14h

Réf : NSQ - Prix 2024 : 2 090€ HT

Les bases de données NoSQL n'offrent pas de langage de requête aussi riche que le SQL. Elles répondent d'abord à des contraintes de volumétrie et à un manque de structuration des données. Ce séminaire présente les différents types de bases NoSQL, leurs architectures, leurs utilisations ainsi que les produits du marché.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Identifier les différences entre SGBD SQL et SGBD NoSQL

Évaluer les avantages et les inconvénients inhérents aux technologies NoSQL

Analyser les principales solutions du monde NoSQL pour chaque modèle de données

Identifier les champs d'application des SGBD NoSQL en opérationnel et en analytique

Comprendre les différentes architectures, les modèles de données, les implémentations techniques

LE PROGRAMME

dernière mise à jour : 06/2021

1) Introduction au NoSQL

- L'historique du mouvement NoSQL.
- Les différentes approches de gestion de SGBD à travers le temps : hiérarchiques, relationnelles, objets, XML, NoSQL.
- Les cinq « V » du big data : Volume, Variété, Vitesse, Véracité, Validité.
- Données non structurées : activité web, échange de documents, réseaux sociaux, open data, IoT.
- Les grands acteurs à l'origine du mouvement NoSQL et du big data analytique : Google et Amazon.
- Vue synoptique des différents types de moteurs NoSQL du point de vue du modèle de données.
- Le NoSQL, le big data et les architectures cloud : principes d'architecture communs et divergents.
- Les modes de distribution : avec maître et décentralisé.
- Transactions distribuées, failover, points de sauvegarde, parallélisation des requêtes, équilibrage de charge.
- Le positionnement du NoSQL au sein du big data analytique : de l'ère de la transaction à l'ère de l'interaction.

Réflexion collective : Pourquoi le NoSQL ? Et pourquoi son succès ? Les besoins, l'évolution des architectures, la distribution et l'élasticité, le commodity hardware, quelques scénarios d'utilisation.

2) Le relationnel et le NoSQL.

- Les bases de données relationnelles : leurs forces et leurs limites.
- Structuration forte des données (schéma explicite) versus structure souple (schéma implicite) et la modélisation Agile.
- Des qualités ACID aux qualités BASE.

PARTICIPANTS

Direction informatique et fonctionnelle. Responsables informatique, chefs de projet, architectes, développeurs.

PRÉREQUIS

Connaissances de base des architectures techniques et du management SI. Connaissances en bases de données.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques... Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

- Théorème CAP (cohérence, disponibilité, tolérance au partitionnement).
- Les différents niveaux de cohérence.
- Le langage SQL, la performance des jointures. L'accès par la clé en NoSQL.
- L'évolution vers le distribué : extensibilité verticale et horizontale.
- Comprendre le NoSQL par le modèle de l'agrégat et de la centralité de la donnée.
- le NewSQL, un redesign des moteurs relationnels pour la distribution. Étude de CockroachDB.

Réflexion collective : Le modèle de l'agrégat versus le modèle relationnel : comment choisir ?
Comment gérer l'interopérabilité ?

3) Les mondes du NoSQL

- Monde du NoSQL à travers ses choix techniques et différentes bases NoSQL libres (du moins structuré au plus structuré).
- L'architecture distribuée : principes, le shared-nothing.
- Disponibilité et cohérence différée : gossip, timestamps, la règle de majorité, l'arbre de Merkle.
- Les patterns et les modèles. Comment modéliser et travailler efficacement en NoSQL.
- Les bases orientées clé-valeur et en mémoire : Redis, Riak, Aerospike.
- Les bases orientées documents : le format JSON. Couchbase Server, MongoDB.
- Les bases orientées colonnes distribuées pour le big data opérationnel : Hbase, Cassandra, ScyllaDB...
- Les moteurs orientés graphes : Neo4j, OrientDB...
- Les moteurs de recherche JSON : Elasticsearch, SOLR.
- Les bases de données de séries de temps : InfluxDB, KDB, Prometheus.

Démonstration : Démonstrations techniques, du point de vue de développement, de la mise en œuvre et de l'administration, des principaux moteurs NoSQL libres.

4) Choisir et mettre en place

- À quels usages correspondent les bases NoSQL ?
- Comment aborder la migration ?
- Comment développer efficacement avec des bases NoSQL ?
- Quels outils de supervision et comment les choisir ?
- Quelle est la complexité administrative et la courbe d'apprentissage ?
- Cas d'utilisation dans des entreprises existantes.
- Gérer les interactions avec les bases de données relationnelles.
- Implémenter des stratégies NoSQL avec des moteurs relationnels. L'exemple de PostgreSQL et ses extensions.
- Implémenter le NoSQL dans les clouds publics. Les pratiques et les offres de bases de données en tant que service.

Réflexion collective : Quel est l'intérêt de déployer des moteurs NoSQL dans le contexte de chacun et quel moteur NoSQL choisir ?

5) NoSQL et Big Data

- Big data analytique : l'écosystème Hadoop.
- Stockage et traitements. Les différentes formes de stockage dans HDFS : SequenceFile, Apache Parquet.
- Fonctions et usages : moteurs de recherche, outils de suggestion commerciale, détecteurs d'intrusion...
- Différents types de traitements : MapReduce, graphe orienté acyclique, flux, machine learning, graphes distribués...
- Fonctionnalités, outils et algorithmes : moteurs de recherche, Google Search, l'algorithme PageRank.
- L'outil intégré : Apache Spark.
- Connexion avec les moteurs opérationnels : ETL, Apache Sqoop.

Démonstration : Démonstrations de l'usage d'une plate-forme intégrée de big data analytique comme Apache Spark.

LES DATES

CLASSE À DISTANCE
2024 : 06 juin, 30 sept., 03 déc.

PARIS
2024 : 30 mai, 23 sept., 26 nov.