Formation NoSQL - Architecture et solutions

Informations

Durée : 2 jours (14h.)

Tarif* : Nous consulter

Réf : NOS2

Niveau : Facile

intra

*tarif valable jusqu'au 31/12/2024

Prochaines sessions

Contactez-nous pour connaitre nos futures sessions.

Pré-requis

- Connaître les bases des architectures techniques

Objectifs

- Identifier les différences des BDD SQL et des BDD NoSQL
- Évaluer les avantages et les inconvénients inhérents aux technologies NoSQL
- Analyser les principales solutions du monde NoSQL
- Identifier les champs d'application des BDD NoSQL

Programme

Le NoSQL

Le NoSQL : définition, historique
Les motivations : extensibilité (scalability), facilité de développement
Les données structurées, semis et non structurées : la provenance, la typologie
La comparaison des systèmes de gestion des bases de données relationnelles (SGBDR) et du NoSQL
Le NoSQL, le Big Data et les architectures Cloud : les principes d'architecture communs et divergents
Le positionnement du NoSQL au sein du Big Analytics : de l'ère de la transaction à l'ère de l'interaction
Le NoSQL et les outils de stockage et d'analyse du Big Data
La classification des analyses

Le Relationnel et le NoSQL

Des qualités ACID aux qualités BASE
Resource Description Framework (RDF)
Le théorème CAP (cohérence, disponibilité, tolérance au partitionnement)
Les différents niveaux de cohérence
SGBDR : avantages et limites
L'évolution vers le distribué : extensibilité verticale et horizontale
Les différentes approches de gestion de bases de données
Les bases hiérarchiques, modèle relationnel, bases objets, bases XML, NoSQL
Le compromis du NoSQL

Les fonctionnalités techniques communes des bases NoSQL

La structure souple des données : la conception du schéma, l'agrégation, la dénormalisation, la duplication
L'architecture distribuée : les principes, le shared-nothing
L'équilibrage de charge
Les méthodes de distribution et de duplication des données
La disponibilité et la cohérence différée : gossip, timestamps, vector clock, règle de majorité, arbre de Merkle
Les méthodes d'amélioration des performances : caches en lecture, en écriture, MVCC
L'architecture in-memory (exemple : ParStream)
Introduction à Hadoop et Map Reduce
L'écosystème Hadoop : différences avec les SGBDR, relations avec le NoSQL
L'usage d'Hadoop dans l'entreprise

Le requêtage

Un essai de classification du NoSQL

Les classements : les différents critères
Les bases de données clé-valeur : le modèle de données, l'extensibilité, réplication, la clé-valeur ordonnée
Les bases de données colonnes : le modèle de données, la hiérarchie (familles, super colonnes), l'extensibilité
Les bases de données document : le modèle de données, l'extensibilité
BDD graphes : le modèle de données, le requêtage, l'extensibilité, les jointures et les graphes, la comparaison avec les SGBDR
L'utilisation des bases graphes, piles logicielles
Les bases de données stream et Complex Event Processing (CEP). Détection de fraude chez PayPal
La base en mémoire. L'avenir : la mémoire non volatile
Les bases XML

Quelques caractéristiques des bases NoSQL existantes

Les bases clé-valeur : Amazon Dynamo, Redis, Riak, Voldemort, Memcached
Les bases colonnes. Google BigTable : principes et API. Hbase : qualités, structure au sein d'Hadoop
Hypertable. Cassandra : propriétés, avantages et faiblesses, requêtage, quelques références
Les bases documents : CouchDB, Couchbase. MongoDB : principes, avantages, limites, cas d'usage
Les bases graphes : Neo4J, AllegroGraph, uRika, une comparaison
Les bases en mémoire : les ancêtres - SolidDB, TimesTen
Les bases actuelles : HANA, SciDB

Vers le NewSQL

Motivation : le relationnel et le NoSQL
Comparaison entre le NewSQL, le NoSQL et les SGBDR
Leader actuel MarkLogic : propriétés, architecture, moteur de recherche, interface, intégration avec Hadoop, cas d'usage
VoltDB : principes et exemple de cas d'usage
Les autres bases : Google Spanner, NuoDB, uCIRRUS, MemSQL, Clustrix

Les infrastructures matérielles pour les solutions NoSQL

Les critères de sélection : performance, capacité, extensibilité
Des principes similaires aux architectures cloud
Les choix matériels : serveurs, stockage disque, les ensembles serveur + stockage, le réseau.
Les serveurs à grande mémoire

Aller ou non vers le NoSQL ?

La migration de ses données vers le NoSQL : MongoDB, bases graphes
Les outils d'intégration de l'écosystème Hadoop
Les usages de bases de données NoSQL
Les critères de différenciation entre projets SGBDR et NoSQL
Les critères de comparaisons entre les différents types de bases NoSQL et les bases relationnelles