mongoDB
FOR GIANT IDEAS

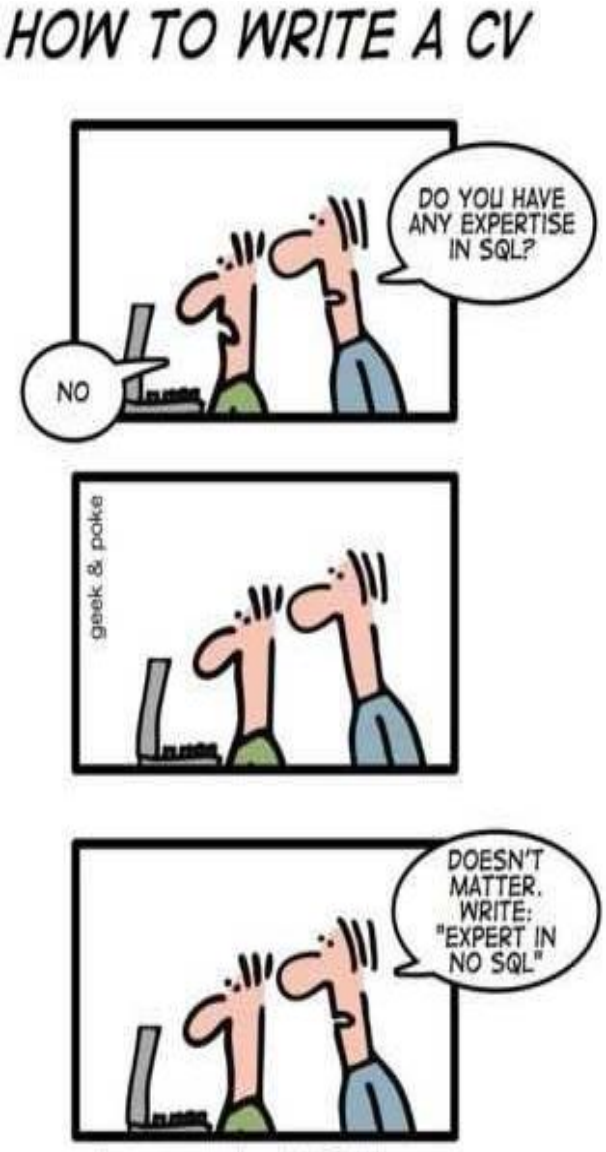
mongoDB
{ name: mongo, type: DB }

Cours de Bases de données NOSQL (MongoDB)

madaniabdellah@gmail.com

Plan du cours

- Rappels :
  - JSON
  - Bases de Données Relationnelles
- Introduction : Mouvement NoSQL
- MongoDB
  - **Caractéristiques de MongoDB**
  - Requêtes avec MongoDB
  - MongoDB et Java
  - MongoDB et PHP
  - Réplication et reprise sur panne dans MongoDB
  - Partitionnement dans MongoDB
  - TextSearch



Caractéristiques de MongoDB

madaniabdellah@gmail.com

MongoDB

MongoDB (humongous = énorme)

- Sponsorisé par 10gen depuis 2007
- Base de données orientée document
- Open-source
- Développé en C++
- Données stockées sous forme BSON
- Absence de tables
- Utilise la notion de « schemaless »
- Fournit un shell javascript pour l'accès aux données et l'administration

MongoDB

- S'adapte bien aux applications Web :
  - cherche à répondre aux besoins de performance en temps réel
  - garantit la scalabilité horizontale (réplication et *sharding*)
  - offre des nombreuses fonctionnalités que l'on trouve dans le monde relationnel (count, groupBy, etc.) mais aussi le support de la recherche **full-text** ou **MapReduce**
  - supporte l'indexation pour l'optimisation des recherches
  - nombreuses fonctionnalités (count, group by, order by, SUM, MIN, etc.)

MongoDB

Parmi les caractéristiques la démarquant des BDR, on retiendra :

- Le **concept de « schéma » n'existe pas**. Chaque document est libre de suivre sa propre structure.
- **Jointures** : les **données** sont généralement **embarquées** dans le même « document ». Cette forme de stockage peut ainsi être vue comme des jointures déjà exécutées (même si la possibilité de *linkage* entre documents existe surtout pour modéliser des relations N-M)
- L'**atomicité** des transactions n'est garantie que sur **un seul document**
- La modification concurrente doit être gérée au niveau de l'application
- La **flexibilité**, la **performance** et la **scalabilité** se font au détriment de la capacité de gérer des transactions complexes
- **Terminologie :** Table = Collection, Ligne = Document, Index = Index, Jointure = Données embarquées

MongoDB vs SGBDR

RDBMS	MongoDB
Database	**Database**
Table	**Collection**
Tuple/Row	**Document**
column	**Field**
Join	**Join Embedded Documents**
Primary Key	**Primary Key (Default key _id provided by mongodb itself)**

- Dans les bases SQL chaque enregistrement de la table contient exactement les mêmes champs, seul le contenu varie
- Dans une collection MongoDB, les documents peuvent avoir des champs totalement différents

MongoDB

- Une installation de mongo peut contenir plusieurs *bases de données*
- Une entrée dans une base mongo est appelée un *document*.
- On stocke des *documents* dans des *collections*.
- Une *base* contient donc des *collections* de documents.
- Un document peut être vu comme un ensemble de couples clef / valeur.

MongoDB

- Exemples de documents très simples :

```
{
    'prenom': 'Christophe',
    'nom': 'Prieur‘
}
{
    'prenom': 'Toto‘,
    ‘dateNaissance’:ISODate("1968-10-20")
}
{
    'code': 'IO2',
    'intitulé': 'Internet et outils‘
}
```

- On utilise ici la syntaxe BSON, qui est utilisée par l’interpréteur mongo

MongoDB

- Rien n'oblige les documents d'une collection à posséder les mêmes champs, ni à ce que les champs de même nom aient des valeurs de même type (la cohérence de la base est laissée à la charge du programmeur).
- Une collection n'est qu'un ensemble de documents, sans contrainte pour ceux-ci.

MongoDB

- Les valeurs peuvent être simples ou complexes (tableaux, listes, dictionnaires, etc.)

```
{
   'code': 'IO2',
   'intitulé': 'Internet et outils',
   'semestre': 2,
    'outils': ['html', 'css', 'php', 'mongo', 'javascript']
}
```

MongoDB

- Chaque document est pourvu d'un identifiant supposé être unique
- Si on utilise le système d'affection d'identifiant par défaut, il sera bien unique
- C'est la clef _id (noter l'underscore).

MongoDB

- **Par défaut, si on ne spécifie pas de valeur explicite pour le champs _id, alors MongoDB en générera un de 24 caractères alpha-numériques.**
- **Par exemple: 56563b19f4c8c7c9597d7d95**
- **Par défaut, cet identifiant est composé:**
  - **D'un timestamp sur 4 octets (indiquant la date/heure de création).**
  - **D'un identifiant de machine sur 3 octets.**
  - **D'un identifiant de processus sur 2 octets.**
  - **D'un compteur sur 3 octets, démarrant à une valeur aléatoire.**
- **Cela veut dire que, pour tout document, on dispose d'ores et déjà d'une date de création sans avoir à créer de champs explicite.**

Installation de MongoDB

- Il faut connaître la version de votre Windows :

```
C:\>wmic os get osarchitecture
```

- Télécharger la version de votre Windows à partir de :

http://www.mongodb.org/downloads

- Installer MongoDB
- MongoDB nécessite un dossier pour stocker ses fichiers :

```
C:\> md data
C:\data>md db
```

- Démarrer le serveur MongoDB

```
C:\mongodb\bin>mongod
```

- Démarrer le shell MongoDB

```
C:\mongodb\bin>mongo
```