

# MongoDB

*Thomas Boch - Infusion XIX, 24 mai 2013*

# MongoDB, base de données orientée documents



mongoDB

```
{name: "mongo", type: "DB"}
```



- ▶ Dans la mouvance NoSQL (*Not only SQL*)
- ▶ Développé par 10gen - licence AGPL
- ▶ schemaless (schéma implicite)
- ▶ pas de jointures
- ▶ on travaille sur des documents (<4MB)
  - ▶ objets JSON
- ▶ Vocabulaire

SGBDR	MongoDB
database	database
table	collection
tuple	document

# Création d'une collection et ingestion de documents

```
> use test
> db.movies.insert({title: "2001, a space odyssey", year: 1969
                    director: {first_name: "Stanley", last_name: "Kubrick"}})

> db.movies.find()
{ "_id" : ObjectId("519dbc09d388164273591067"), "title" : "2001, a space
odyssey", "year": 1969, "director" : { "first_name" : "Stanley",
"last_name" : "Kubrick" } }

> db.movies.insert({title: "Un prophète", year: 2009, director: {first_name:
"Jacques", last_name: "Audiard"}, cast: [{first_name: "Tahar", last_name:
"Rahim"}, {first_name: "Niels", last_name: "Arestrup"}], tags: ["drama",
"crime"]})
```



# Requêtage (1/2)

- ▶ Nombre de documents dans la collection :

```
> db.movies.count()  
7
```

- ▶ Films sortis en 1979 :

```
> db.movies.find({year: 1979})  
{ "_id" : ObjectId("519dc267d38816427359106d"), "title" : "La vie de Brian", "year" : 1979, ... }  
{ "_id" : ObjectId("519dc3f9d38816427359106f"), "title" : "Apocalypse now", "year" : 1979, ... }
```

- ▶ Films sortis avant 1960 :

```
> db.movies.find({year: {$lt: 1960}})  
{ "_id" : ObjectId("519dc37ed38816427359106e"), "title" : "12 angry men", "year" : 1957, ... }  
{ "_id" : ObjectId("519dc60fd388164273591070"), "title" : "Citizen Kane", "year" : 1941, ... }  
  
> db.movies.find({$where: 'this.year < 1960'})
```

# Requêtage (2/2) ■

- ▶ Films réalisés par Kubrick :

```
> db.movies.find({"director.last_name": "Kubrick"}, {title: true, _id: false})
{ "title" : "2001, a space odyssey" }
```

- ▶ Films taggés comme **drama** ou **comedy** :

```
> db.movies.find({tags: {$in: ["drama", "comedy"]}})
...
```

- ▶ Films ayant 3 acteurs :

```
> db.movies.find({cast: {$size: 3}}, {title: true})
{ "_id" : ObjectId("519dd502d388164273591072"), "title" : "LA Confidential" }
```



# Mais aussi ...

- ▶ sharding et partitionnement horizontal
- ▶ index géographique (GeoJSON)
- ▶ agrégation et map/reduce

# Drivers

- ▶ Disponible dans la plupart des langages : C, C++, C#, Erlang, Java, Javascript, Perl, PHP, Python, Ruby, Scala
- ▶ PyMongo :

```
from pymongo import MongoClient

mongoConn = MongoClient()
mongoDb = mongoConn['test']
moviesCollection = mongoDb['movies']

moviesCollection.insert({"title": "Blade runner"})
```



# Utilisation de MongoDB comme cache pour Sesame

"Messier 1"



```
{  
  ra: 83.633,  
  dec: 22.014,  
  ...  
}
```

# Méthodologie

- ▶ Récupération des logs Sesame sur 6 mois
- ▶ Extraction des noms d'objets résolus par Sesame
- ▶ Ingestion des identifiants SIMBAD et positions associées
- ▶ Hypothèse (simplificatrice ?) : *la casse et les blancs ne sont pas significatifs*
  - ▶ " Messier 1 " == "Messier1" == "mEsSiEr 1"
- ▶ Soumission de la liste d'objets à MongoDB



```
{
  "id": "cosmosj095932.33+014305.1",
  "position": {
    "ra": 149.8847666,
    "dec": 1.7180694
  },
  "otype": "Galaxy",
  "magnitudes": [
    {
      "B": 23.922
    },
    {
      "V": 23.279
    },
    {
      "K": 20.028
    }
  ]
}
```

# Résultats

- ▶ 1.6M requêtes traitées en 3000 secondes (533 requêtes/seconde)
- ▶ 81% des noms d'objets sont résolus
  - ▶ correspond à 83% des requêtes
- ▶ exemple de noms d'objets non résolus
  - ▶ 11 Per --> \* 11 Per
  - ▶ 1ES 1320+084 --> 1ES 1320+08.4
  - ▶ 2dFS 0332 --> 2dFS 332



# Conclusions

- ▶ 83% des requêtes Sesame résolues peuvent l'être sans interrogation de SIMBAD
  - ▶ réduit la dépendance aux pannes externes
  - ▶ réduit la charge induite sur SIMBAD
- ▶ MongoDB peut servir de cache d'objets

# Liens

- ▶ Tutoriel interactif:

<http://try.mongodb.org/>

- ▶ Getting started:

<http://docs.mongodb.org/manual/tutorial/getting-started/>

- ▶ Common use cases:

<http://docs.mongodb.org/manual/use-cases/>