



# MongoDB, le NoSQL #1






# Historique de la solution

Une histoire simple mais originale dans le monde du NoSQL

## Historique de la solution




10Gen, MongoDB Inc.

L'histoire de MongoDB ne commence pas en Californie avec un grand nom du Web mais à NYC.

10Gen, en 2007, lance le développement d'une service de PAAS (Plateform As A Service), avec des attentes similaires à Google Cloud Platform ou Window Azure.  
Finalement en 2009, 10Gen introduit sur le marché la base de données Open Source MongoDB.

Plus tard 10Gen change de nom pour devenir MondoDB Inc.

## Historique de la solution



### Première release stable et adoption massive

La première release considéré industriellement viable est la version 1.4 sorti en mars 2010.

La dernière release stable (3.6.1) date du 26 décembre 2017 et une release candidate de la 3.6.2 a d'ores et déjà été publié.

MongoDB, développé en C, C++ et javascript, est actuellement la 5ème base de données la plus populaire d'après DbEngine et la première NoSQL.



# Intérêt de la solution

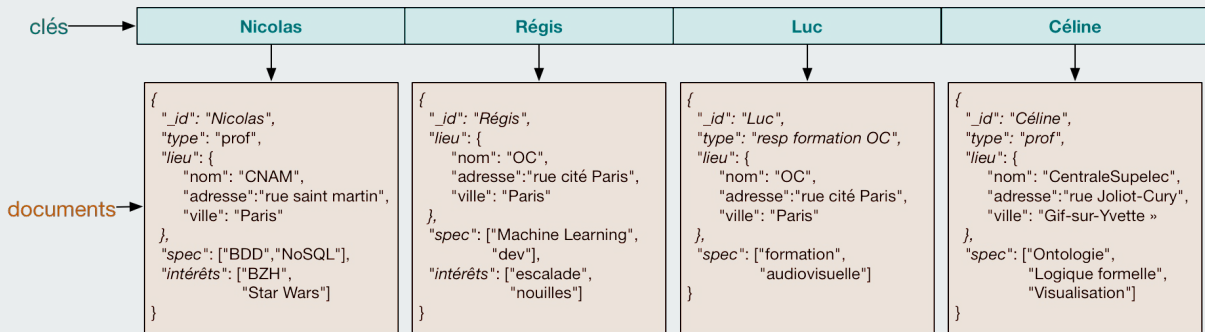
MongoDB (de l'anglais humongous qui peut être traduit par « énorme ») est un système de gestion de base de données orientée documents, répartitionnable sur un nombre quelconque d'ordinateurs et ne nécessitant pas de schéma prédéfini des données.

# Intérêt de la solution

## Le concept de base orienté Document

Stockage orienté lignes

id	type	lieu	spec	intérêts
Nicolas	prof	CNAM	BDD, NoSQL	BZH, Star Wars
Régis		OC	Machine Learning, Dev	escalade, nouilles chinoises
Luc	resp formation OC	OC	formation, audiovisuel	
Céline	prof	CentraleSupélec	Ontologie, logique formelle, visualisation	



## Intérêt de la solution



### Stockage de data dans MongoDB

MongoDB stocke les données au format BSON (JSON Binaire). Cette structure lui permet de sauvegarder dans son format de document des structures de données simples et des tableaux associatifs.

MongoDB n'a pas de schéma prédéterminé. Les données prennent la forme de documents enregistrés eux-mêmes dans des collections, une collection contenant un nombre quelconque de documents. Les collections sont comparables aux tables, et les documents aux enregistrements des bases de données relationnelles.

Contrairement aux bases de données relationnelles, les champs d'un enregistrement sont libres et peuvent être différents d'un enregistrement à un autre au sein d'une même collection.

## Intérêt de la solution



### Stockage de fichier dans MongoDB

Le système GridFS dans MongoDB permet de stocker directement des fichiers bénéficiant de l'avantage de répartition de de tolérance au panne propre à MongoDB.

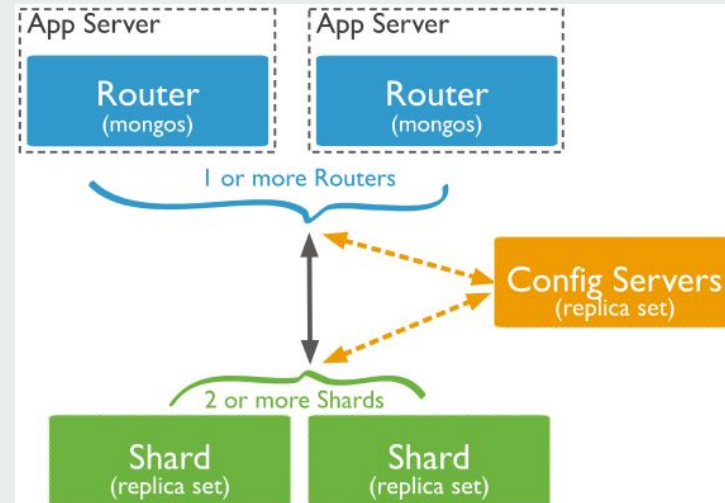
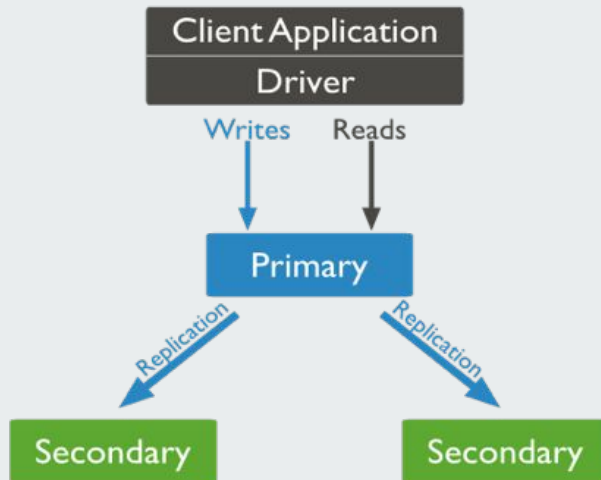
Cette particularité en fait un choix particulièrement pertinent pour les système de gestions électroniques de document.



## Intérêt de la solution

### Déploiement de mongoDB

Suivant la méthode utilisée pour déployer mongoDB, la base supportera une scalabilité et un comportement d'écriture différent.

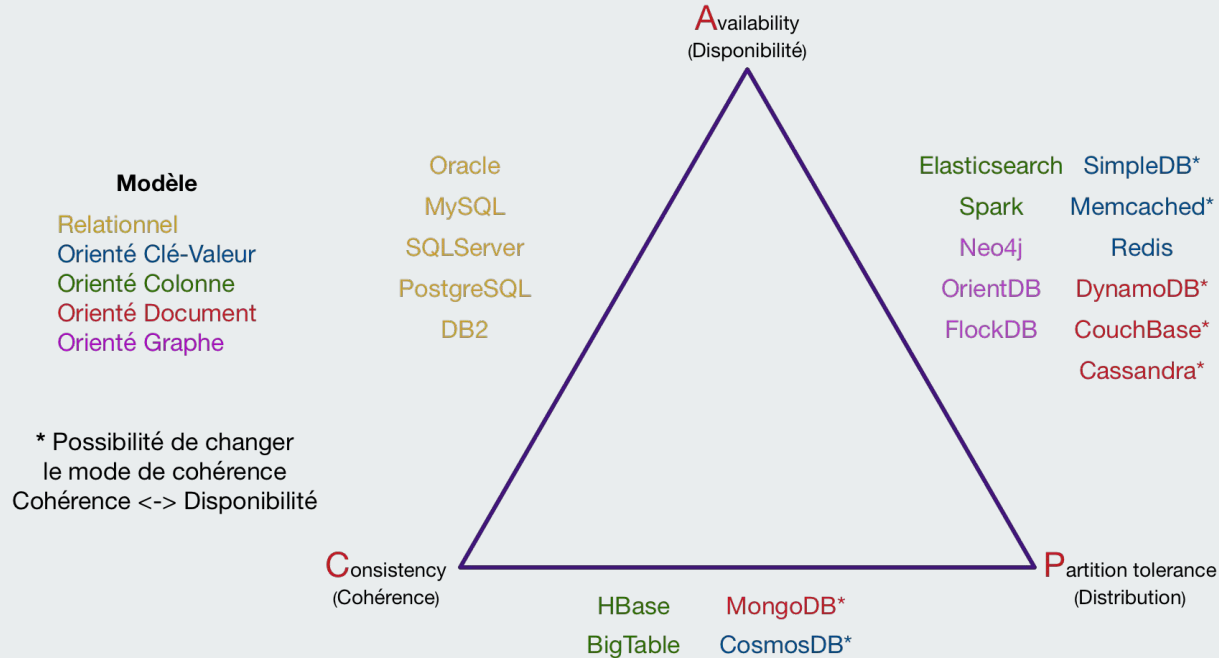
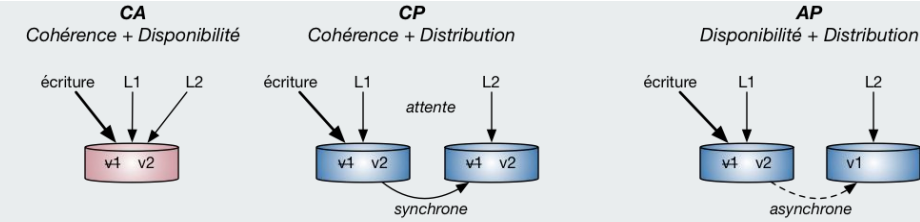




# Intérêt de la solution

Pour quels besoins ?

# Intérêt de la solution / Pour quels besoins ?



## Intérêt de la solution / Pour quels besoins ?



Le NOSQL le plus proche d'un besoin RDBMS

Sa structure souple de document et sa souplesse de requêtage en font une des base NoSQL la plus proche des besoins RDBMS.

Doctrine (un ORM d'abstraction de base SQL en PHP) propose d'ailleurs d'être utiliser avec MongoDB proposant d'abstraire un peu plus le NoSQL aux développeurs.

MongoDB est donc un choix crédible pour les projets avec une désérialisation simple sans jointure (ou très peu) et pas de transaction.

## Intérêt de la solution / Pour quels besoins ?

### Exemple de cas d'utilisation d'une base orienté document

- gestion de contenu (bibliothèques numériques, collections de produits, dépôts de logiciels, collections multimédia, etc.),
- framework stockant des objets,
- collection d'événements complexes,
- gestion des historiques d'utilisateurs sur réseaux sociaux.

Les moteurs de blog, de micro-blogging, les GED, les annuaires sont de bonnes applications du MongoDB.

Les applications mobiles, la gestion de contenu, les analytics et l'IOT sont aussi des domaines où MongoDB sera à l'aise.



# Intérêt de la solution

Quels sont les équivalents ?

## Intérêt de la solution / Quels sont les équivalents ?

Les concurrents et équivalent

Les alternatives en base documentaire sont les même que pour cassandra

- CouchBase : avec un requêtage SQL like
- CouchDB : base documentaire XML
- Cassandra : base documentaire/colonne avec requêtage SQL like



**Couchbase**





# Intérêt de la solution

Quand l'utiliser et quand préférer autre chose ?



## Intérêt de la solution / Quand l'utiliser et quand préférer autre chose ?



### Les pors et les contres

MongoDB excelle en scalabilité et cache pour du traitement temps réel d'analytcs ou si le schéma n'est pas clairement défini.

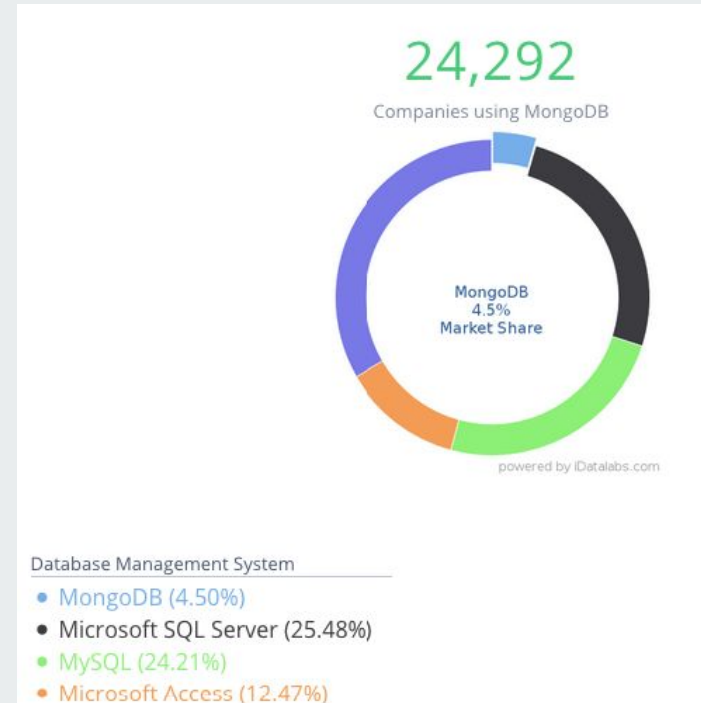
#### POURS

Stockage de document non structuré, bonnes performance sur les inesrtions

#### CONTRES

La base n'est pas ACID compliant, pas de fonctions ou de procédures stockées, non prévu pour de la données transactionnelles.

# Qui l'utilise ?




## Qui l'utilise ?



# Analyse d'un cas client



## Analyse d'un cas client



### Shutterfly, le challenge

Shutterfly est une plateforme de stockage et d'échange d'images, elle stocke environ 6 milliards de photos avec une charge jusqu'à 10 000 transactions par secondes.

Historiquement, Shutterfly tournait sur le SGBD Oracle.

Les premières recherches d'alternatives à Oracle ont conduit à tester MySql et Postgersql (2 RDBMS open source), puis à comparer plusieurs solutions NoSql :

- Cassandra
- CouchDB
- BerkeleyDB
- En bien entendu MongoDB

## Analyse d'un cas client



### Shutterfly, l'application

MongoDB n'est pas venu remplacer le système RDBMS existant sur l'ensemble des données de Shutterfly.

C'est principalement autours de la régie pub, du stockage des photos et de leur métadatas que MongoDB a été utilisé.

La facturation et la gestion des utilisateurs, nécessitant une consistance critique, a été conservé dans Oracle.

## Analyse d'un cas client

### Shutterfly, l'application

GridFS for large  
object storage

Flexible data model  
for similar, but  
different objects

Geo spatial indexing  
for location based  
searches

```
{ camera: "Nikon d4",  
  location: [ -122.418333, 37.775 ]  
}
```

```
{ camera: "Canon 5d mkII",  
  people: [ "Jim", "Carol" ],  
  taken_on: ISODate("2012-03-07T18:32:35.002Z")  
}
```

```
{ origin: "facebook.com/photos/xwdf23fsdf",  
  license: "Creative Commons CC0",  
  size: {  
    dimensions: [ 124, 52 ],  
    units: "pixels"  
  }  
}
```

Horizontal scalability  
for large data sets



```
db.archives.  
  find({ "country": "Egypt" });
```

Indexing and rich  
query API for easy  
searching and sorting

Flexible data model  
for similar, but  
different objects

```
{ type: "Artefact",  
  medium: "Ceramic",  
  country: "Egypt",  
  year: "3000 BC"  
}
```

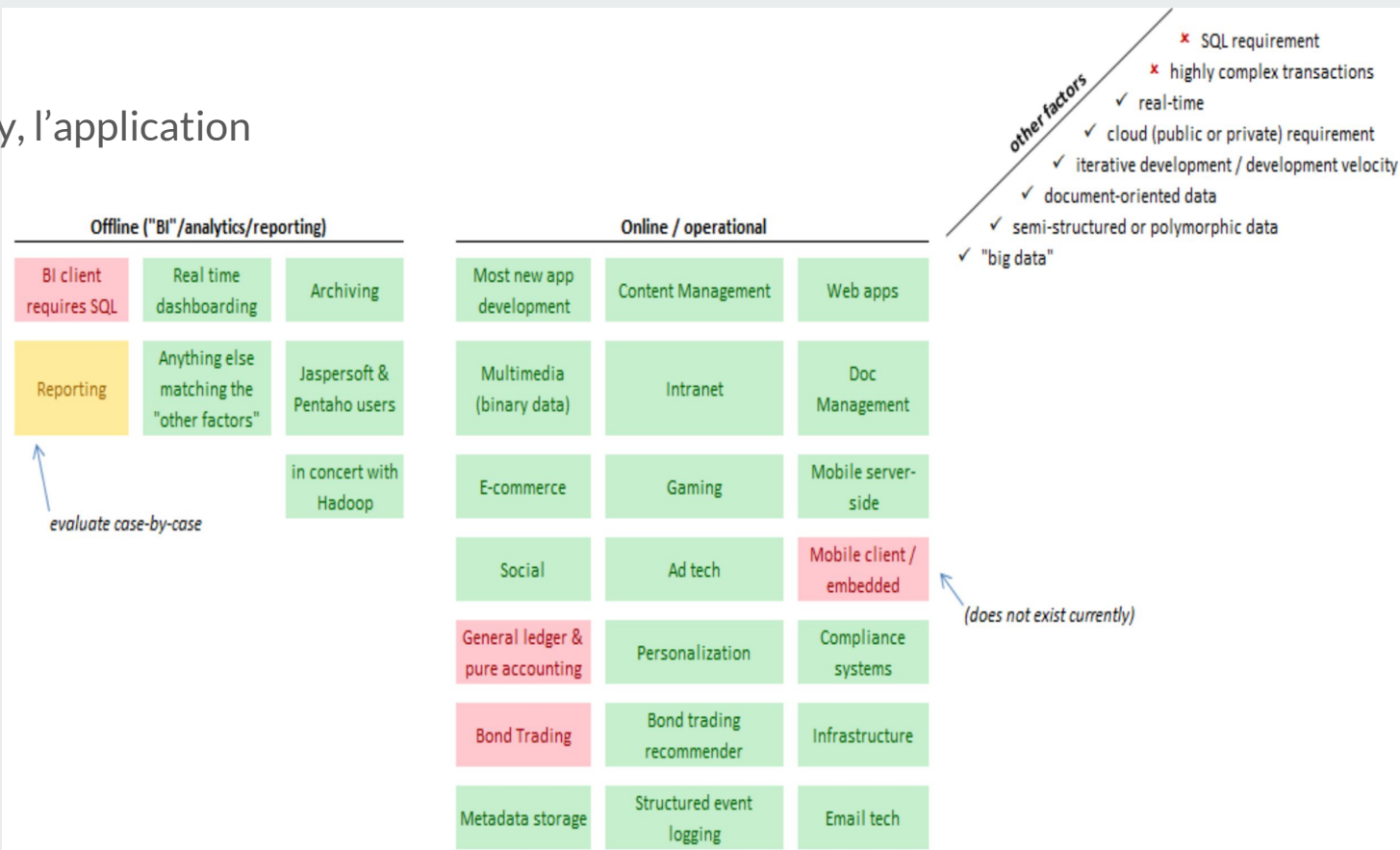


```
{ ISBN: "00e8da9b",  
  type: "Book",  
  country: "Egypt",  
  title: "Ancient Egypt"  
}
```



## Analyse d'un cas client

### Shutterfly, l'application







# Point sur les offres cloud

## Point sur les offres cloud

Directement disponible sur [mongodb.com](https://www.mongodb.com/cloud/atlas)

The screenshot shows the MongoDB Atlas cloud provider selection interface. It includes a 'Choose a Cloud Provider' section with options for AWS, Google Cloud Platform, and Microsoft Azure. Below this is a 'Choose A Region' dropdown menu set to 'eu-west-2 (London)'. The main section displays four instance types: M2, M5, M10, and M20. Each instance type shows its RAM and storage specifications. A 'Custom Storage Capacity' slider is at the bottom, with a green arrow pointing to 320GB. A cost estimate box on the right shows '\$0.80/hr Per 3 Node Replica Set'.

Instance Type	RAM	Storage
M2	Shared RAM	2 GB storage
M5	Shared RAM	5 GB storage
M10	2 GB RAM	10 GB storage
M20	4 GB RAM	20 GB storage

Custom Storage Capacity: 320 GB

Cost Estimate: \$0.80/hr Per 3 Node Replica Set



# Bon appétit