



Présentation de la formation MongoDB, administration

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Noureddine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Le formateur

- **Noureddine DRISSI**
 - Expertise dans le domaine des bases de données
 - 15 années d'expérience sur tous les SGBDR du marché (Mysql, Oracle, SQL Server, Postgresql, DB2, MongoDB)
 - Issue de l'environnement bancaire
-  contact@valneo-xi.fr
- Mon profil sur alphorm.com : <http://www.alphorm.com/auteur/noureddine-drissi>
- Mes formations sur alphorm.com



MongoDB: Administration

alphorm.com™ ©

Connaissances préalables

- Connaissances du système d'exploitation Linux/UNIX ou Windows
- Connaissance de JavaScript (non indispensable)

A propos de ce cours

- Objectifs :
 - Comprendre le mouvement NoSQL;
 - installer MongoDB;
 - comprendre l'architecture de MongoDB;
 - travailler avec les objets dans MongoDB;
 - configurer la montée en charge;
 - optimiser les performances;
 - administrer MongoDB au quotidien.

Plan du cours

- Module 01 : Présentation du mouvement NoSQL
- Module 02 : Présentation de MongoDB
- Module 03 : Prise en main de MongoDB
- Module 04 : Travailler avec les documents
- Module 05 : Réplication et Sharding
- Module 06 : Performance et indexation
- Module 07 : Administration de MongoDB



Liens utiles

- La documentation MongoDB
 - <https://www.mongodb.org>
- Des livres : <http://www.mongodb.org/books> (en Anglais)
- Blog : <http://blog.mongodb.org/>
- Services commerciaux : <http://www.mongodb.com/>
- MongoDB Management Service (un service gratuit en cloud pour la surveillance et backup des déploiement de mongodb) :
<http://mms.mongodb.com/>



GO

MongoDB: Administration

alphorm.com™ ©

Alphorm



Présentation de la formation L'environnement

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

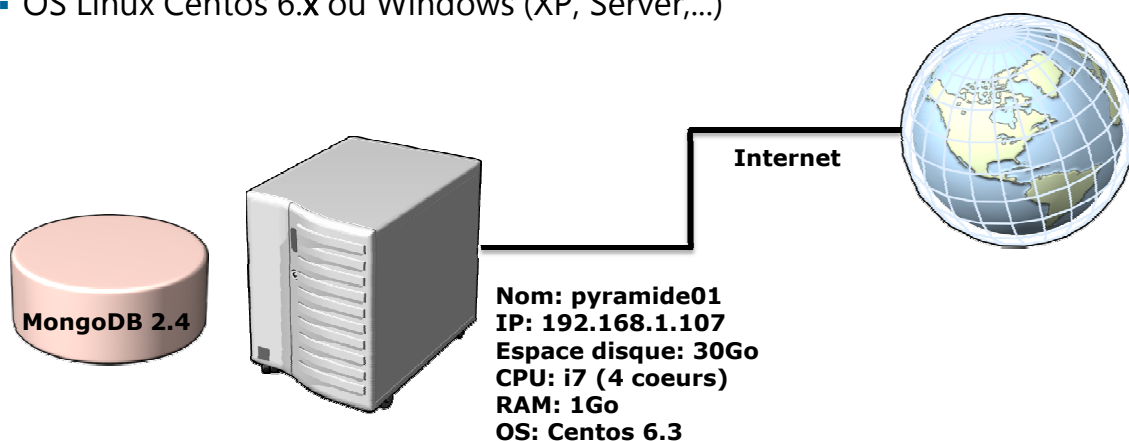
Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Environnement

- Prés-requis
 - Espace disque de 30G
 - OS Linux Centos 6.x ou Windows (XP, Server,...)



Présentation du mouvement NoSQL





Présentation du mouvement NoSQL

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©



Le mouvement NoSQL

- Historique
 - Le concept du NoSQL à une bonne décennie d'ancienneté
 - Utilisé pour la première fois en 1998
 - Développé par des start-up comme logiciel de stockage des données
 - Inauguration du NoSQL le 11 novembre 2009 à San Francisco
 - Début de normalisation d'un langage de manipulation des données UnQL (Unstructured Query Language) en 2011

MongoDB: Administration

alphorm.com™ ©



Le mouvement NoSQL

- Le NoSQL (Not Only SQL) c'est quoi ?
 - SGBD non relationnel issue du monde Web
 - Les données ne sont plus manipulées avec le SQL
 - L'unité de stockage n'est plus la table
 - La définition d'un schéma de données relationnel n'est plus nécessaire
 - Renonciation aux fonctionnalités des SGBDR (ACID)
 - Les principaux axes sont la haute disponibilité et le partitionnement des données
 - Permet de gérer de très grosses volumétries de données



Le mouvement NoSQL

- Catégories de bases NoSQL
 - Les bases clefs/valeurs
 - *Redis, Voldemort*
 - Les bases orientées documents
 - *MongoDB, CouchDB, Riak*
 - Les bases orientées colonnes
 - *Cassandra, BigTable, HBase*
 - Les bases orientées graphes
 - *Neo4j: en Java*
 - *FlockDB: utilisé par Twitter*

Le mouvement NoSQL

- Les bases clefs/valeurs
 - Principes du modèle
 - Associe un identifiant unique à chaque valeur dans la base de données
 - Une valeur peut être de type simple ou un objet sérialisé
 - Fournit quatre opérations basiques
 - création, lecture, modification, suppression
 - Toute l'intelligence dans la récupération des données se situe dans l'applicatif client

Le mouvement NoSQL

- Les bases orientée Colonnes
 - Principes du modèle
 - Stocke les informations sous forme de colonnes
 - Permet d'ajouter très facilement des informations
 - Limite les valeurs nulles





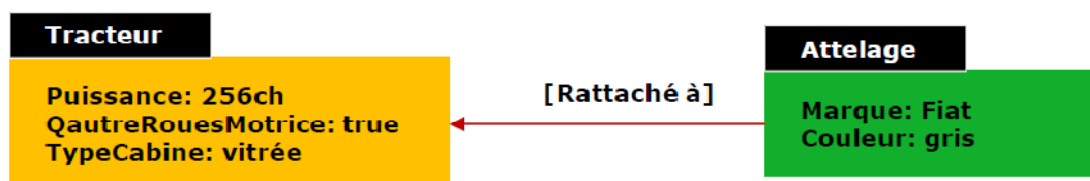
Le mouvement NoSQL

- Les bases orientée documents
 - Destinée à stocker des documents
 - Le contenu d'un document ne se limite pas à des attributs simples
 - Chaque document a sa propre structure dans une même collection



Le mouvement NoSQL

- Les bases orientée Graphes
 - Le moins connus et assez complexe à mettre en oeuvre
 - Principes du modèle
 - Repose sur un modèle mathématique appelé Théorie des Graphes
 - Données modélisées sous format de noeud
 - Données reliées entre elles par des Arc nommés





Le mouvement NoSQL

- Autres types de bases NoSQL
 - Base de données hiérarchique
 - Base de données objet



Le mouvement NoSQL

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



Le mouvement NoSQL

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



Le mouvement NoSQL

- Avantages de MongoDB
 - Répond aux besoins de performances
 - Garantit la scalabilité horizontale (réplication et sharding)
 - Nombreuses fonctionnalités (count, group by, order by, SUM, MIN, etc.)
 - Supporte l'indexation pour optimiser les performances



Le mouvement NoSQL

- La présence sur le marché de MongoDB
 - Utilisé par de grands comptes
 - MTV
 - Disney
 - Doodle
 - CERN (Organisation Européenne pour la Recherche Nucléaire)
 - ...



Le mouvement NoSQL

- Analogie avec le SQL
 - NoSQL (Base) = SQL (Base)
 - NoSQL (Collection) = SQL (Table)
 - NoSQL (Document) = SQL (Enregistrement)
 - 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



Le mouvement NoSQL

- Analogie avec le SQL
 - NoSQL (Base) = SQL (Base)
 - NoSQL (Collection) = SQL (Table)
 - NoSQL (Document) = SQL (Enregistrement)
 - 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



Le mouvement NoSQL

- Les principaux produits NoSQL
 - MongoDB
 - Accumulo
 - Berkeley DB (Oracle)
 - BigTable (Google)
 - Cassandra (Facebook, Twitter, Digg)
 - CouchDB
 - SimpleDB (Amazon.com)
 - ...



Le mouvement NoSQL

- Cas d'utilisation
 - Sites web populaires
 - Répartition de charge (load balancing)
 - Utilisation de matériel peu coûteux
 - Optimisation et administration simplifiés



Le mouvement NoSQL

- Comparaison NoSQL / SGBDR
 - SGBDR
 - Les points forts :
 - Beaucoup de fonctionnalités et de règles pour garantir des bases cohérentes et complètes : mécanisme de verrous pour la gestion des accès concurrentiels, respect des propriétés ACID (Atomicité, Cohérence, Isolation, Durabilité).
 - Utilisation des outils spécialisés comme le modèle entité-relation qui est le Modèle Conceptuel de Données (MCD) utilisé pour modéliser la structure de la base.
 - Les points faibles :
 - Performances qui déclinent avec l'augmentation du volume de données
 - Problématique de distribution des données d'une base.



Le mouvement NoSQL

- Comparaison NoSQL / SGBDR
 - NoSQL
 - Les points forts :
 - Bons temps de réponse malgré de très gros volumes de données
 - Facilement distribuable
 - Plus flexible en cas de panne (disponibilité partielle voire totale)



Le mouvement NoSQL

- Comparaison NoSQL / SGBDR
 - NoSQL
 - Les points faibles :
 - Moins de propriétés garantissant un état cohérent de la base. Conformément au théorème de Brewer (aussi appelé théorème du CAP), seules deux des trois propriétés suivantes peuvent être respectées par un SGBD NoSQL : Cohérence, Disponibilité et Résistance au morcellement.
 - Pas de mécanismes de jointures, le côté client doit pallier à ce problème.
 - Il est rare qu'un SGBD de type NoSQL implémente un mécanisme de verrous pour garantir la cohérence pendant des accès concurrents.



Prise en main de MongoDB



MongoDB: Administration

alphorm.com™©

Alphorm.com



Prise en main de MongoDB

Installation sous Linux

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Installation sous Linux

- Installation sous Linux
 - Deux solutions
 - Passer par le gestionnaire de paquets (yum sous Centos)
 - Pour les distributions Debian, Ubuntu
 - <http://downloads-distro.mongodb.org/repo/>
 - Télécharger le code source directement (plus pénible)

Installation sous Linux

- Configuration de yum
 - Pour un OS 64 bits
 - Ajouter le le fichier /etc/yum.repos.d/10gen.repo
 - Ajouter les entrées suivantes

```
[10gen]
name=10gen Repository
baseurl=http://downloads-distro.mongodb.org/repo/redhat/os/x86_64
gpgcheck=0
enabled=1
```

Installation sous Linux

▪ Configuration de yum

- Pour un OS 32 bits
 - Ajouter le le fichier /etc/yum.repos.d/10gen.repo
 - Ajouter les entrées suivantes

```
[10gen]
name=10gen Repository
baseurl=http://downloads-distro.mongodb.org/repo/redhat/os/i686
gpgcheck=0
enabled=1
```

Installation sous Linux

▪ Installation via yum

- yum install mongo-10gen mongo-10gen-server



Installation de MongoDB sous Windows



MongoDB: Administration

alphorm.com™©

Alphorm.com



Prise en main de MongoDB

Installation sous Windows

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Installation sous Windows

- Les étapes
 - Télécharger l'archive à partir de <http://www.mongodb.org/downloads>
 - Décompresser l'archive
 - Création de l'arborescence de stockage de l'instance par défaut
 - Configuration du path
 - Création du service d'arrêt/démarrage de Windows
 - Démarrage de MongoDB



Instance MongoDB





Prise en main de MongoDB Instance MongoDB

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Instance MongoDB

- Caractéristiques d'une instance MongoDB
 - Un port d'écoute (par défaut 27017)
 - Un processus serveur
 - Un répertoire racine de stockage
 - Un fichier de log
 - Un fichier de configuration: *mongod.conf*

MongoDB: Administration

alphorm.com™ ©

Instance MongoDB

- Arrêt / Démarrage d'une instance MongoDB
 - Sous Linux
 - Démarrage
 - `/etc/init.d/mongod start` (démarrage de mongoDB via script)
 - `mongod --dbpath <racine_instance>` (démarrage de mongoDB)
 - `mongod -f <fichier.conf>`
 - Arrêt
 - `/etc/init.d/mongod stop` (permet d'arrêter mongoDB)
 - `/etc/init.d/mongod restart` (permet le redémarrage de mongoDB)
 - Commande `db.shutdownServer()` à partir du client mongo et la base admin

Instance MongoDB

- La racine de l'instance
 - Sous Windows
 - Via le gestionnaire de services Windows
 - En mode ligne de commande
 - `net start MongoDB` (démarrage du service)
 - `net stop MongoDB` (arrêt du service)



Les outils et commandes MongoDB



MongoDB: Administration

alphorm.com™©



Alphorm.com

Prise en main de MongoDB

Les outils et commandes MongoDB

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Noureddine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

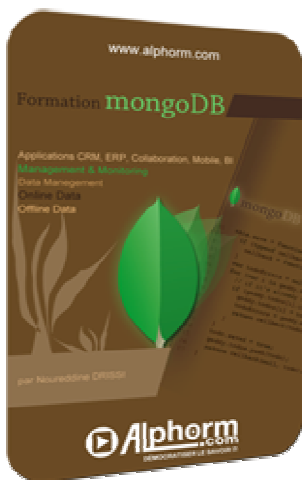
Les outils et commandes MongoDB

- Inventaire
 - **mongod** (le moteur de base)
 - **mongo** (le shell javascript)
 - **mongos** (le contrôleur de Sharding)
 - Les outils d'import/export
 - **mongoimport**, **mongoexport**, **mongodump**, **mongorestore**, **bsondump**
 - **mongofiles** (l'utilitaire GridFS)
 - **mongostat** (visualisation des stats d'une instance mongoDB)
 - **mongosniff** (le tcpdump de mongo)
 - **mongotop**, **mongoperf**



Bases de données MongoDB





Prise en main de MongoDB

Les bases de données

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

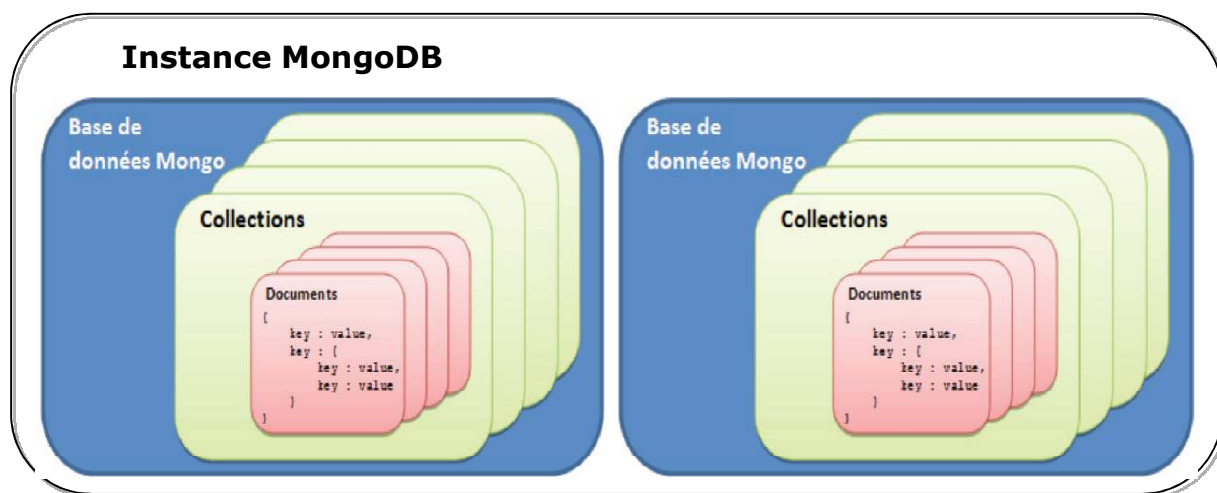
Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Les bases de données

■ Architecture



MongoDB: Administration

alphorm.com™ ©

Les bases de données

- Création / Suppression
 - Création de bases de données
 - Automatique
 - # mongo <db_name>
 - > use <db_name>
 - Suppression de bases de données
 - # mongo
 - > use <db_name>;
 - > db.runCommand({dropDatabase: 1});
 - > db.dropDatabase();

Les bases de données

- Stockage des données physiques d'une base
 - Limite en environnement 32 bits
 - Ne permet pas de stocker plus de 2.5Gb de données
 - Déconseillé pour de la production
 - Le répertoire de stockage utilise la variable --dbPath
 - En production
 - Utiliser la version 64 bits
 - Le répertoire de stockage utilise la variable --dbPath

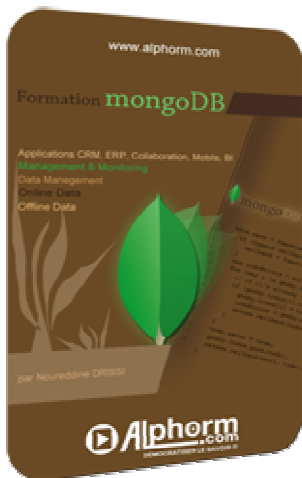
Les bases de données

- Stockage des données physiques d'une base
 - Caractéristiques
 - Pour chaque base
 - initialisation de deux fichiers
 - <db_name>.0 et <db_name>.1
 - La taille de chaque fichier est le double de la taille du précédent (64, 128, 256, etc.)
 - Taille maximale d'un fichier = 2Gb
 - Le fichier <nom_fichier>.ns stocke les namespaces de la base (16 Mb par défaut)
 - modifiable via l'option -nssize, avec pour taille maximum 2Gb



Les espaces de noms





Prise en main de MongoDB

Les espaces de noms (namespace)

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Les espaces de noms

- Caractéristiques
 - Chaque collection et chaque indexe compte comme un espace de nom
 - Concrètement: concaténation du nom de la base et celui de la collection
 - Limites
 - Par défaut: 24000 espaces de noms
 - Possibilité d'augmenter cette taille avec le paramètre **-nssize**
 - Possibilité de redimensionner une base existante
 1. On redémarre mongod avec l'option **-nssize**
 2. On exécute la commande **db.repairDatabase()**

MongoDB: Administration

alphorm.com™ ©



Travailler avec les documents



MongoDB: Administration

alphorm.com™©



Alphorm.com

Travailler avec les documents

Les collections

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Noureddine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Les collections

- Une Collection c'est quoi ?
 - Equivalent de la table en relationnel
 - Deux modes de création
 - Automatiquement lors d'une insertion de document (tuple)
 - En exécutant la commande `db.createCollection('<nom_collection>');`

Les collections

- Les opérations sur les Collections
 - Créer
 - > `db.createCollection('gens');`
 - Lister
 - > `show collections;`
 - > `db.getCollections();`
 - Insérer un document
 - > `db.<collection_name>.insert( { var1: "valeur", var2: "valeur", var3: "valeur", } );`
 - Supprimer
 - > `db.<nom_collection>.drop();`



Les documents



MongoDB: Administration

alphorm.com™©

Alphorm.com



Travailles avec les documents

Les documents

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

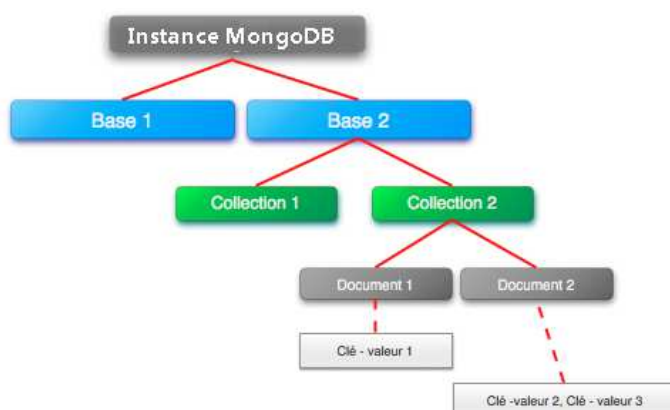
alphorm.com™©

Les documents

- Un document c'est quoi ?
 - Equivalent d'un enregistrement (tuple) en relationnel
 - Deux modes d'insertion
 - En utilisant la commande **db.<nom_collection>.insert()** du schell
 - En utilisant la commande **db.<nom_collection>.save ()** du schell
 - En utilisant un driver: PHP, Ruby, Java, Python
 - Chaque document crée contient le champ « **_id** »
 - Généré automatiquement
 - Type primary key
 - Indexé

Les documents

- Rappel de la hiérarchie des objets



Les documents

- Le format JSON/BSON
 - MongoDB stocke les documents au format BSON (Binary JSON)
 - BSON est la représentation binaire des objets JSON (JavaScript Object Notation)

Les documents

- Le format JSON c'est quoi ?
 - JSON (JavaScript Object Notation)
 - Permet de représenter de l'information structurée
 - Format de données textuelle
 - Utilise une notation JavaScript
 - Décrit par la RFC 4627
 - Un document JSON, ne comprend que deux éléments structurels :
 - des ensembles de paires nom / valeur ;
 - des listes ordonnées de valeurs.

Les documents

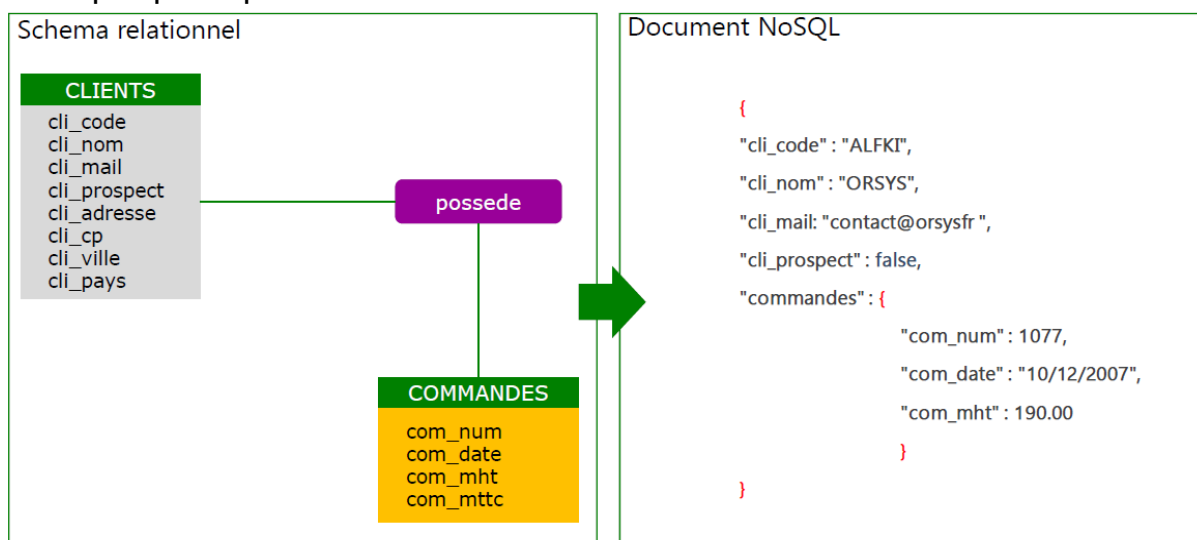
- Le format JSON

- Exemple

```
{
  "pospect" : "ALFKI",
  "nom" : "ORSYS",
  "adresse" : {
    "numero" : 77,
    "codepostal" : "31000",
    "ville" : "Toulouse"
  }
}
```

Les documents

- Exemple pratique d'utilisation



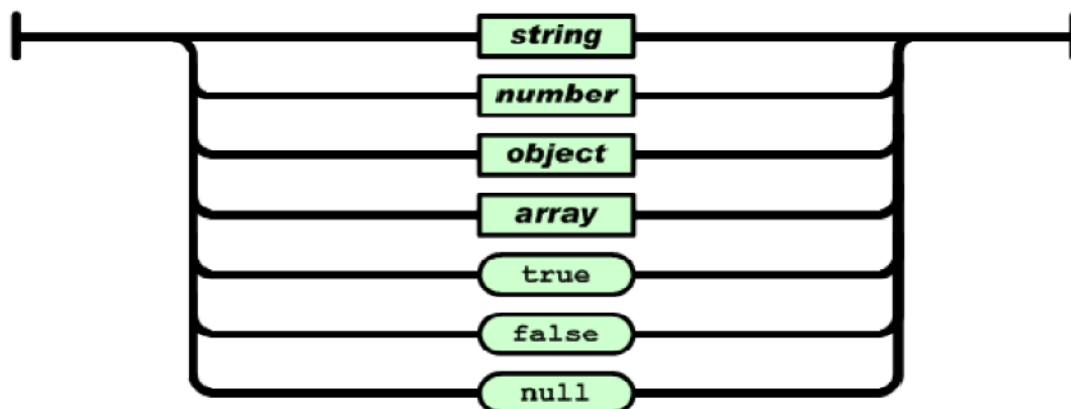
Les documents

- Avantage du format JSON
 - Format abstrait pour une représentation simplifiée dans les différents langages
 - Indépendant du langage de programmation
 - Simple et complet pour la représentation des objets
 - Utilise des types de données connus et simples à décrire
 - Une bonne lisibilité de la syntaxe
 - Temps de traitement très proche de celui des fichiers XML
 - Moins volumineux en terme de stockage
 - Pour JavaScript un document JSON représente un objet

Les documents

- Types disponibles en JSON
 - issus de json.org

value





Les documents

- Le format BSON c'est quoi ?
 - BSON est la représentation binaire des objets JSON
 - Ajoute des améliorations et des capacités supplémentaires
 - expressions régulières
 - Bytes Arrays
 - Date et TimeStamps
 - stockage de blocs de code ou de fonctions JavaScript



Les documents

- Les types de données
 - Une base MongoDB est constituée de Collection(s)
 - Une collection est constituée de document(s) regroupant des paires clef/valeur dont la clef est une string et la valeur de type :
 - des objets ;
 - des tableaux ;
 - des valeurs génériques de type tableau, objet, booléen, nombre, chaîne ou null.



Les opérations CRUD



MongoDB: Administration

alphorm.com™©



Alphorm.com

Réplication et Sharding

La réplication

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

La réplication

- Présentation
 - Deux fonctionnalités importantes
 - Depuis la version 1.6
 - Les Replica Set
 - Le Sharding

La réplication

- Qu'est-ce que le Replica Set ?
 - Permet d'assurer la redondance et la haute disponibilité
 - Permet aux données d'être dupliquées de manière transparente
 - Utilise la notion de groupe de serveur appelé set
 - Chaque set possède
 - **un** noeud principal (primaire) : lecture/écriture
 - **n** serveurs de backup (secondaire) : lecture uniquement
 - **Concrètement:** les Replica Set permettent de répliquer les données entre des instances **MongoDb**

La réplication

- Qu'est-ce que le Replica Set ?
 - Permet d'assurer la redondance et la haute disponibilité
 - Permet aux données d'être dupliquées de manière transparente
 - Utilise la notion de groupe de serveur appelé set
 - Chaque set possède
 - **un** noeud principal (primaire) : lecture/écriture
 - **n** serveurs de backup (secondaire) : lecture uniquement
 - **Concrètement:** les Replica Set permettent de répliquer les données entre des instances **MongoDb**

La réplication

- Mise en place
 - Réplication
 - Démarrage d'un serveur dans un Replica Set
 - **mongod**
 - **--replSet** (indique à quel Replica Set appartient l'instance)
 - **--shardsvr** (active le partitionnement horizontal)
 - ...
 - Le serveur maître génère un journal de transaction
 - Consommé par les différentes bases esclaves

La réplication

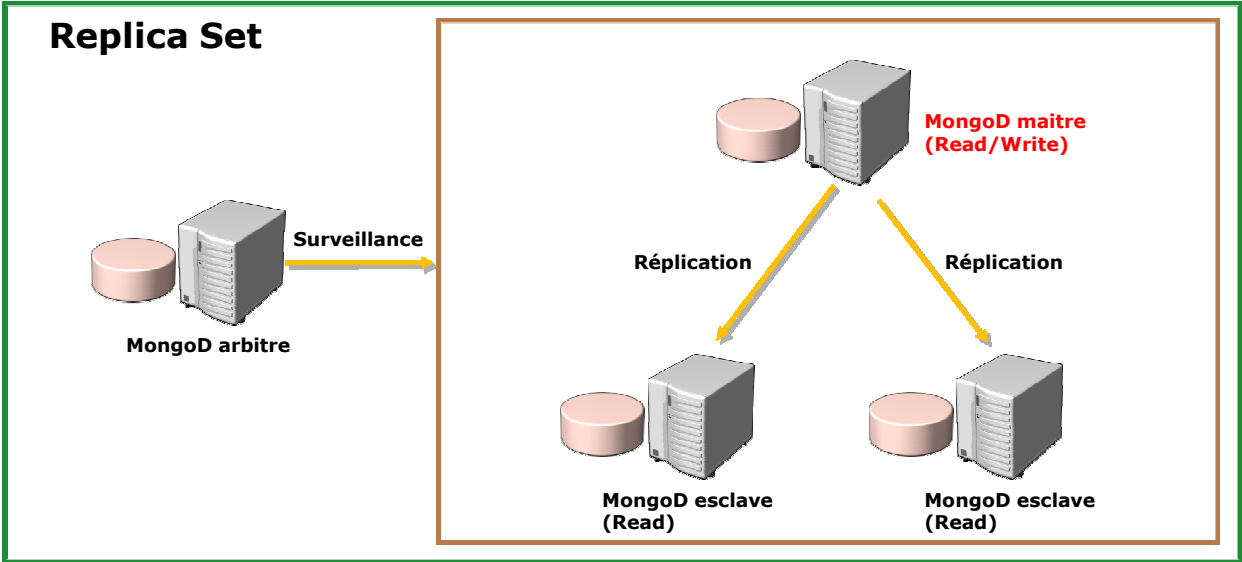
- Mise en place
 - Réplication
 - Ajout d'un noeud dans un Replica Set
 - Sur le serveur maître
 - PRIMARY> `rs.add("host:port")`
 - Sortir un noeud d'un Replica Set
 - Sur le serveur maître
 - PRIMARY> `rs.remove("host:port")`

La réplication

- Mise en place
 - Réplication
 - Les principales commandes d'exploitation
 - `rs.initiate(cfg)`; (permet d'initialiser la configuration d'un Replica Set)
 - `db.isMaster()`; (permet de vérifier qui est le maître)
 - `rs.status()`; (permet de vérifier l'état du Replica Set)
 - `rs.slaveOk()`; (permet d'activer les lecture sur les serveurs esclaves)
 - `rs.syncFrom("host:port")`; (permet de forcer la synchronisation)
 - ...

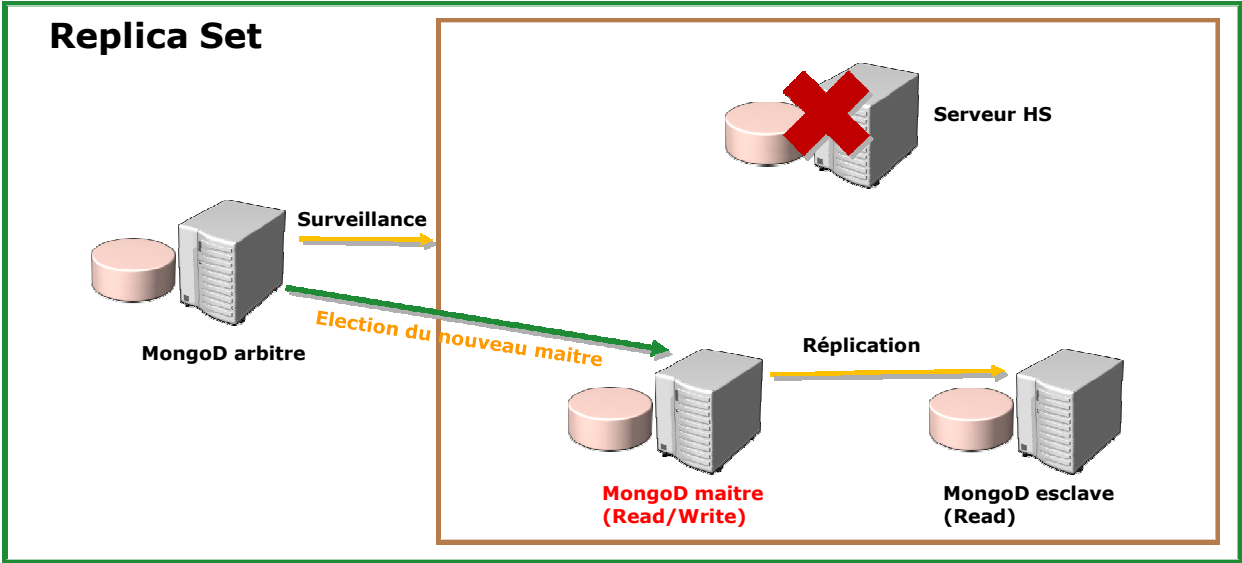
La réplication

- Exemple de Replica Set



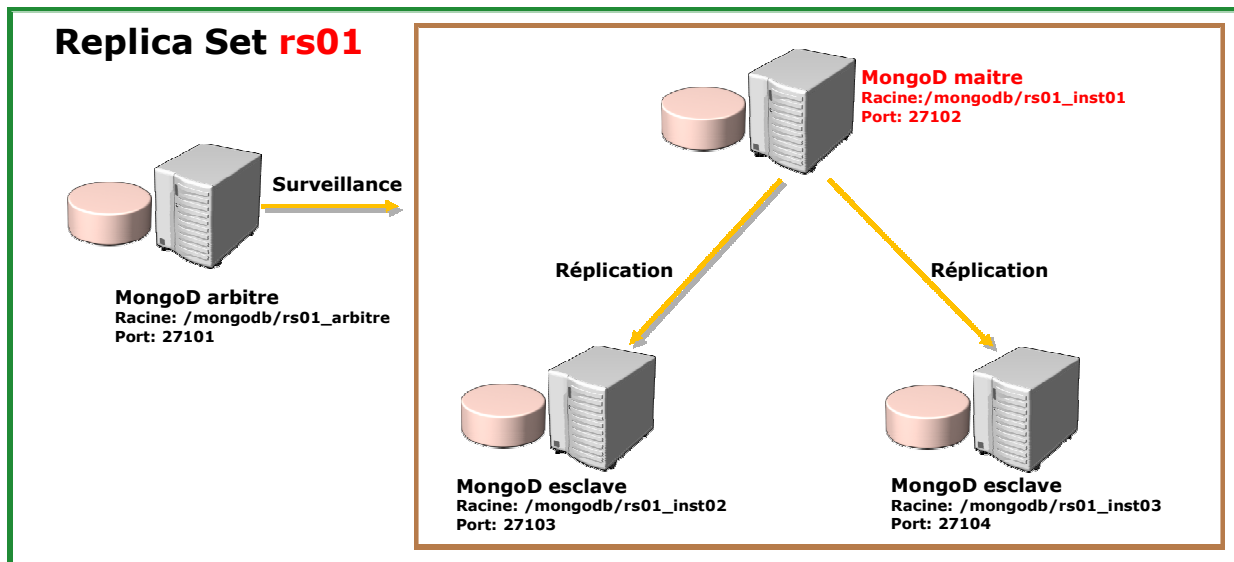
La réplication

- En cas de défaillance



La réplication

■ TP



MongoDB: Administration

alphorm.com™©



Le sharding



MongoDB: Administration

alphorm.com™©



Réplication et Sharding Le sharding

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Le sharding

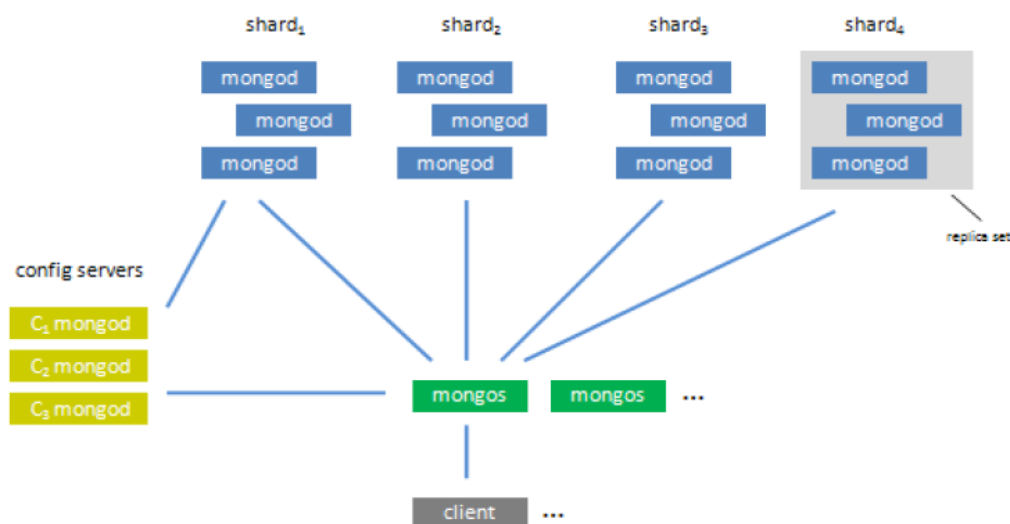
- Qu'est-ce que le Sharding ou partitionnement ?
 - Pré-requis: mise en place des Replica Set
 - Permet une scalabilité horizontale
 - Distribution du stockage des données sur différentes instances MongoDB
 - Chaque machine stock un sous ensemble des données (chunk)
 - Utilise une clé de Sharding pour répartir le stockage

MongoDB: Administration

alphorm.com™ ©

Le sharding

▪ Architecture du sharding



Le sharding

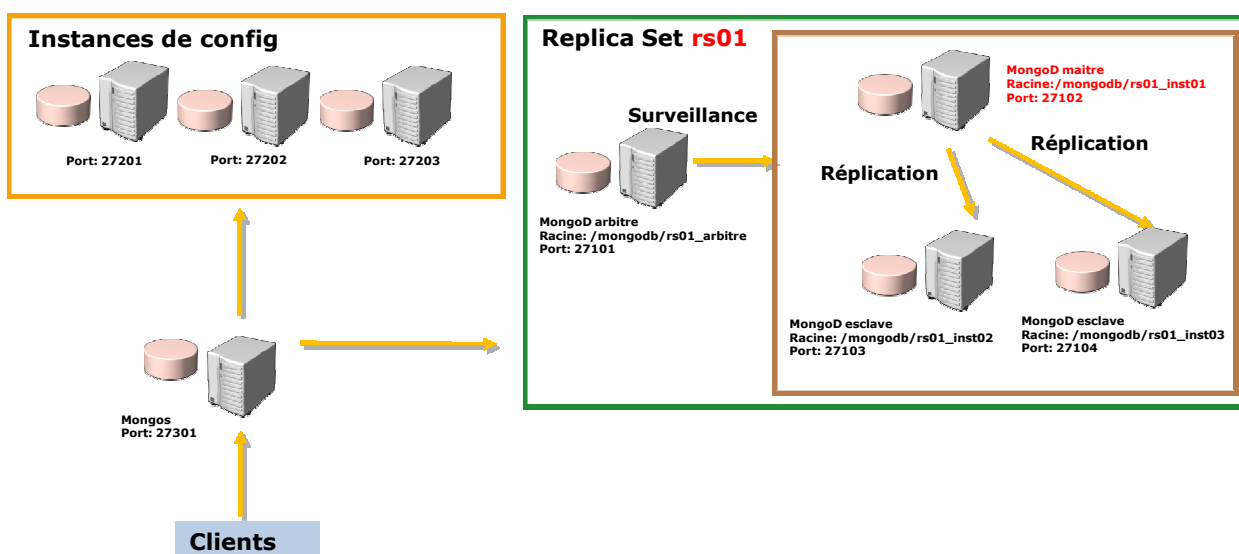
- Les étapes de mise en oeuvre
 - Mise en place du Replica Set
 - Mise en place du sharding
 - Les Shard servers : ou instances mongod
 - Les Config servers : Serveur de configuration qui va stocker les metadata du shard
 - Le mongos : Sert de routeur vers les différents shards
 - Test de l'infrastructure

Le sharding

- Les principales commandes
 - Ajouter des noeuds au sharding
 - `db.runCommand({addshard : "<nom_Replica_Set/host1:port,host2:port,hostN:port"});`
 - Activer le sharding pour une base de données
 - `db.runCommand({enablesharding : "<nom_base>"});`
 - Définir une clef de partitionnement pour le sharding
 - `db.runCommand({shardcollection : "<namespace>", key:{"<nom_champ>":1}});`
 - Afficher les informations sur le sharding
 - `db.printShardingStatus();`
 - `db.<collection>.stats();`

Le sharding

▪ TP





Performance et indexation



MongoDB: Administration

alphorm.com™©



Alphorm.com

Performance et indexation

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Performance et indexation

- Les indexes sous MongoDB
 - Pourquoi faire ?
 - Améliorer les temps d'exécution des requêtes
 - Les différents types d'indexes
 - Index sur un seul champ
 - Index unique
 - Index composé, Sparse et Hashed
 - Indexation par défaut
 - Le champ `_id` est indexé par défaut (impossible de le supprimer)

Performance et indexation

- Index sur un seul champ
 - Deux modes
 - Ordre ascendant (1 pour ASC)
 - Ordre descendant (-1 pour DESC)
 - Poser un index (ascendant et descendant)
 - `db.<collection>.ensureIndex({<nom_champ>:1});`
 - `db.<collection>.ensureIndex({<nom_champ>:-1});`

Performance et indexation

- Index composés
 - Deux modes
 - Ordre ascendant (1 pour ASC)
 - Ordre descendant (-1 pour DESC)
 - Poser un index (ascendant et descendant)
 - `db.<collection>.ensureIndex({<nom_champ1>:1, <nom_champ2>:1});`
 - `db.<collection>.ensureIndex({<nom_champ1>:-1, <nom_champ2>:-1});`

Performance et indexation

- Les index unique
 - Deux modes
 - Ordre ascendant (1 pour ASC)
 - Ordre descendant (-1 pour DESC)
 - Poser un index unique (ascendant et descendant)
 - `db.<collection>.ensureIndex({<nom_champ1>:1, {unique: true}});`
 - `db.<collection>.ensureIndex({<nom_champ1>:-1, {unique: true}});`

Performance et indexation

- Plusieurs types d'indexes
 - Index BTREE
 - Index Sparse
 - Index Haché
 - Géospatial
 - Text

Performance et indexation

- Création des index en arrière plan
 - Par défaut construction de l'index en premier plan
 - Possibilité de construire un index en arrière plan
 - `db.collection.ensureIndex( { a: 1 }, { background: true } )`

Performance et indexation

- Visualisation des index
 - Index d'une collection
 - `db.<collection>.getIndexes()`
 - Index d'une base de données
 - `db.system.indexes.find()`

Performance et indexation

- Reconstruction d'index
 - Utilisation de la méthode `db.collection.reIndex()`
- Suppression d'index
 - > `db.<collection>.dropIndex('<nom_index>')`

Performance et indexation

- Monitoring des opérations sur les index (création/reconstruction)
 - Progression: `db.currentOp()`
 - Interruption: `db.killOp()`

Performance et indexation

- Visualisation des plans d'exécution
 - Utilisation de la méthode: `cursor.explain()`
 - `db.<collection>.find().explain()`
 - Forcer l'utilisation d'un index avec `hint()`
 - Utilisation de la méthode: `cursor.hint()`
 - `db.<collection>.find().hint({ <nom_champ>: 1 }).explain()`



Administration de MongoDB



MongoDB: Administration

alphorm.com™©

Alphorm



Administration de MongoDB

Sauvegarde et restauration

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Sauvegarde et restauration

- Sauvegarde à froid (le plus simple)
 - Les étapes
 - Arrêter le service mongod
 - Copier tous les fichiers de --dbpath
 - <nom_base>.<numero_fichier>
 - <nom_base>.ns
 - Démarrer le service mongod

Sauvegarde et restauration

- Export MongoDB
 - Commande **mongodump**
 - Permet d'exporter (collection, base, instance)
 - Contient autant de répertoire que de base
 - Contenant chacun un fichier BSON par collection, plus un fichier BSON d'index.

Sauvegarde et restauration

- Export MongoDB
 - **mongodump**
 - Les principales options
 - --db
 - --collection
 - --out (précise le répertoire d'export)
 - --dbpath (précise la racine à exporter)

Sauvegarde et restauration

- Export MongoDB
 - Commande **mongoexport**
 - Permet d'exporter
 - Une base
 - Une collection
 - Des champs
 - Deux formats d'exports possible
 - JSON
 - CSV

Sauvegarde et restauration

- Export MongoDB
 - `mongoexport`
 - Les principales options
 - `--db`
 - `--collection`
 - `--csv`
 - `--fields`

Sauvegarde et restauration

- Import MongoDB
 - Commande `mongoimport`
 - Permet d'importer
 - Une base
 - Une collection
 - Des champs
 - Deux formats d'imports possible
 - JSON
 - CSV

Sauvegarde et restauration

- Import MongoDB
 - Commande `mongorestore`
 - Permet de restaurer
 - Pré-requis
 - Droper la base cible
 - Les principales options
 - `--db`
 - `--collection`
 - `--dbpath`



La sécurité sous MongoDB





Administration de MongoDB

Map Reduce

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™ ©

Map Reduce

- Caractéristiques
 - Se compose de deux fonctions :
 - **map** qui passe à reduce une clé et des données à traiter
 - **reduce** qui, pour chaque clé donnée par map, va opérer une réduction sur les données en provenance de map selon une logique qu'il faut écrire.
- Fonctionnement
 - La fonction **map** va émettre (emit) des paires clé/valeur sur lesquelles reduce va travailler.
 - **map** traverse tous les documents d'une collection MongoDB et va pousser des infos à **reduce** selon une clé définie
 - Permet de paralléliser les traitements
 - Stocke le résultat dans une collection

MongoDB: Administration

alphorm.com™ ©



La journalisation



MongoDB: Administration

alphorm.com™©



Alphorm.com

Administration de MongoDB

La journalisation

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

La journalisation

- Caractéristiques
 - Permet d'assurer la fonctionnalité de durabilité
 - Utilise les write ahead logging (journal des transactions)
 - En version 2.0: la journalisation est activée par défaut en environnement 64-bits
 - Ne pas désactiver le mode journal en production
 - Au démarrage les opérations d'écritures sont écrites en mémoire et dans le journal des transactions avant l'application dans les fichiers data

La journalisation

- Activation du journal
 - Démarrage de mongodb avec l'option **--journal**
 - Si pas de fichier journal, mongodb pré-alloue un nouveau fichier de 1Go
 - Pas de connection possible pendant cette opération

La journalisation

- Recouvrement
 - Après un crash ou un arrêt intempestif
 - Au démarrage mongodb rejoue automatiquement les fichiers journaux
 - Si le démarrage échoue
 - Utiliser la commande `repairDatabase()`



Les principaux paramètres de
configuration d'une instance





Administration de MongoDB

Les principaux paramètres

Site : <http://www.alphorm.com>
Blog : <http://www.alphorm.com/blog>
Forum : <http://www.alphorm.com/forum>

Nouredine DRISSI
Expert consultant bases de données
Contact : contact@valneo-xi.fr

MongoDB: Administration

alphorm.com™©

Les principaux paramètres

- Renseignés dans le fichier **mongod.conf**
 - **dbpath** emplacement des données d'une instance
 - **logpath** emplacement des logs
 - **logappend** continu à écrire des anciens logs ou créer un nouveau fichier
 - **nojournal = true** désactive ou pas la journalisation
 - **norealloc = true** désactive la pré-allocation des fichiers data
 - **nssize = <size>** taille du fichier d'espace de nom
 - **port = 27017** port utilisé par une instance
 - **auth = true** active l'authentification sous MongoDB
 - **fork = true** démarrage de mongod en tâche de fond
 - **objcheck = true** inspection des documents avant insertion (utf-8 , type ect ..)
 - **nohttpinterface = true** active ou désactive l'interface web de monitoring (Defaults to localhost:27018).
 - ...

MongoDB: Administration

alphorm.com™©



Les interfaces graphiques d'administration

