

Projet sur les bases de données du Big Data :  
*Projet d'Approfondissement MONGODB ET D'UN MOTEUR  
NOSQL DE VOTRE CHOIX*  
*Au 25/02/2025*

## Préambule

Le travail ici demandé fait suite au cours « Les Bases de Données du Big Data ». Ces bases de données sont aussi souvent appelées Bases de Données NoSql. L'objectif est, à partir de d'un sujet, de mettre en œuvre de façon approfondie le SGBD Nosql MONGODB et un SGBD NOSQL de votre choix. Ce travail sera fait par **groupe de 4 étudiants au maximum (3 minimum)**. Ce TP permet, en plus de l'examen, d'évaluer le cours « Les Bases de Données du Big Data » (50% de la note). **Tous les livrables doivent être rendus le 15 avril 2025 au plus tard.** SOYEZ CREATIF. CELA SERA RECOMPENSE.

**FAIRE UN DOSSIER DE RENDU PAR MOTEUR NOSQL :** Un dossier MongoDB et un autre dossier du moteur NOSQL de votre choix avec une comparaison avec MongoDB

**Concernant l'autre moteur : Il doit être différent pour chaque groupe.**

## Table des matières

1. CHOIX DU SUJET	3
2. MCD MERISE	3
3. CONVERSION DU MCD MERISE EN DES OBJETS MONGODB ET DU MOTEUR NOSQL DE VOTRE CHOIX ET CLASSES JAVA	4
4. COMPLEMENTS SUR LE MOTEUR NOSQL MONGODB ET D'UN MOTEUR NOSQL DE VOTRE CHOIX	4
5. GENERATION AUTOMATIQUE DES DONNEES	4
6. CHARGEMENT DES DONNEES	5
7. MISES A JOUR DES DONNEES (INSERT, UPDATE, DELETE)	5
8. INTERROGATION ET MANIPULATION DES DONNEES	5
9. CONSTRUCTION D'UN CLUSTER MONGODB	5
10. LES RESULTATS DE L'ETUDE	6
11. REPARTITION DU TRAVAIL	6
12. OU S'INSCRIRE POUR CONSTITUER SON GROUPE DE 4	6
13. OU DEPOSER CES RESULTATS	7
14. EVALUATION DES RENDUS	7

# 1. Choix du sujet

Vous choisirez, si ce n'est déjà fait, une application appartenant à la liste suivante ou un sujet de votre choix :

- |  |                                      |
|--|--------------------------------------|
| 1) Agence de location de bateaux       | 12) Gestion d'un labo. de chercheurs |
| 2) Agence immobilière                  | 13) Cabinet de notaires              |
| 3) Agence bancaire                     | 14) Gestion de la bourse             |
| 4) Gestion d'une bibliothèque          | 15) Gestion d'un cabinet médical     |
| 5) Tournois de tennis                  | 16) Gestion d'une agence de voyages  |
| 6) Gestion d'une promotion d'étudiants | 17) Gestion de salons                |
| 7) Gestion des menus d'un restaurant   | 18) Gestion des anciens étudiants    |
| 8) Gestion d'un club sportif           | 19) Gestion d'une boîte de nuit      |
| 9) Tournois de trampoline              | 20) Jeux lignes                      |
| 10) Gestion de tweet                   | 2) profils des clients d'un téléco   |
| 11) Commerce en ligne                  | 22) autres sujets de votre choix     |

Attention : deux groupes ne peuvent pas avoir un même sujet.

## 2. MCD MERISE

Le résultat de cette phase doit être mis dans un fichier appelé 2specification\_analyse\_conception\_Nom1\_Nom2\_Nom3\_Nom4.docx

Et/ou 2specification\_analyse\_conception\_Nom1\_Nom2\_Nom3.pdf

Nom1, Nom2, Nom3 représentent les noms des membres d'un même groupe.

Ce fichier doit contenir :

- **Une page de garde** : avec le titre du projet et les noms des membres
- **La description du sujet** : ce que doit faire cette application, décrire textuellement ces structures et ses services
- **Le dictionnaire de données MERISE**. Pour chaque entité décrire chacune des propriétés : Titre / description / format des données / type / Identifiant / contraintes
- **La description textuelle des associations** : Décrire textuellement les associations entre entités
- **La définition du Modèle Entité-Association MERISE** (en utilisant le logiciel Poweramc de SYBASE/SAP ou manuellement). Vous devez vous limiter à 10 entités maximum et 5 minimum. Vous devez ici prendre en compte les contraintes identifiées lors de la description du dictionnaire de données. Exemple de liens d'association pour deux entités A et B ayant une liaison 1 : N ou N-M (exemple UN PILOTE ASSURE 0, 1 ou plusieurs VOL, un VOL est assuré par 1 et 1 PILOTE au plus)

**Attention :** Le schéma conceptuel obtenu ici n'est qu'un indicatif.

### **3. Conversion du MCD MERISE en des objets MONGODB et du moteur nosql de votre choix et classes java**

- **Spécification des modèles de documents à mettre dans chaque collection**

Les objets doivent contenir si utile entre autres des tableaux, des records, etc.

Il ne s'agit de reproduire une base de données relationnelle. L'esprit MONGODB et D'UN MOTEUR NOSQL DE VOTRE CHOIX doit primer.

Vous devez organiser les objets en fonction des impératifs du moteur nosql cible (clé/valeur, clé/document, clé/colonnes, graphe).

- **Spécification des classes et des méthodes JAVA**

- . Méthodes CRUD (insertion, modification, suppression)

- . Indexes secondaires en fonction de chacun e vos moteurs nosql

- . Méthodes applicatives de consultation (jointure, groupement, tri, traitement en masse de documents) en fonction de chacun e vos moteurs nosql.

### **4. Compléments sur le moteur NoSql MONGODB et D'UN MOTEUR NOSQL DE VOTRE CHOIX**

Merci de développer les points ci-dessous théoriquement et pratiquement si possible :

- Modèles de données supportés
- Réévaluer la procédure d'installation du moteur et des utilitaires
- Architecture du moteur NoSql (avec des schémas expliqués)
- Méthode de partitionnement (avec des schémas expliqués)
- Méthode de réplication (avec des schémas expliqués)
- Montée en charge (avec des schémas expliqués)
- Gestion du ou des caches mémoire (avec des schémas expliqués)

### **5. Génération automatique des données**

Pour chacun de vos moteurs (MongoDB et le le moteur nosql de votre choix), vous devez générer automatiquement des objets (plusieurs milliers si possible).

Le site <https://mockaroo.com/> (ou un autre) permet de générer des données gratuitement et en grand nombre, dans divers formats, dont JSON, CSV, .... Il sait générer des noms, des dates, des valeurs numériques dans certains intervalles etc.

D'autres solutions de solutions de génération de données existent. Vous pouvez les utiliser à la place.

## 6. Chargement des données

Vous devez ici procéder au chargement des objets générés plus haut, dans votre BD Nosql. Trois solutions possibles :

- Ecrire une classe Java par container (collection, table, ...) avec l'api fournit par le sgbd nosql
- Utiliser l'interface ligne de commandes du sgbd nosql
- Utiliser un utilitaire de chargement de données

Vous avez le droit de choisir une des solutions ou de mixer.

## 7. Mises à jour des données (insert, update, delete)

Afin d'effectuer les mises à jour vous devez :

- Définir une classe Java par container (collection, table, ...) qui permet d'insérer, modifier ou supprimer des enregistrements (le faire pour un enregistrement et aussi pour plusieurs)
- Effectuer des tests

## 8. Interrogation et manipulation des données

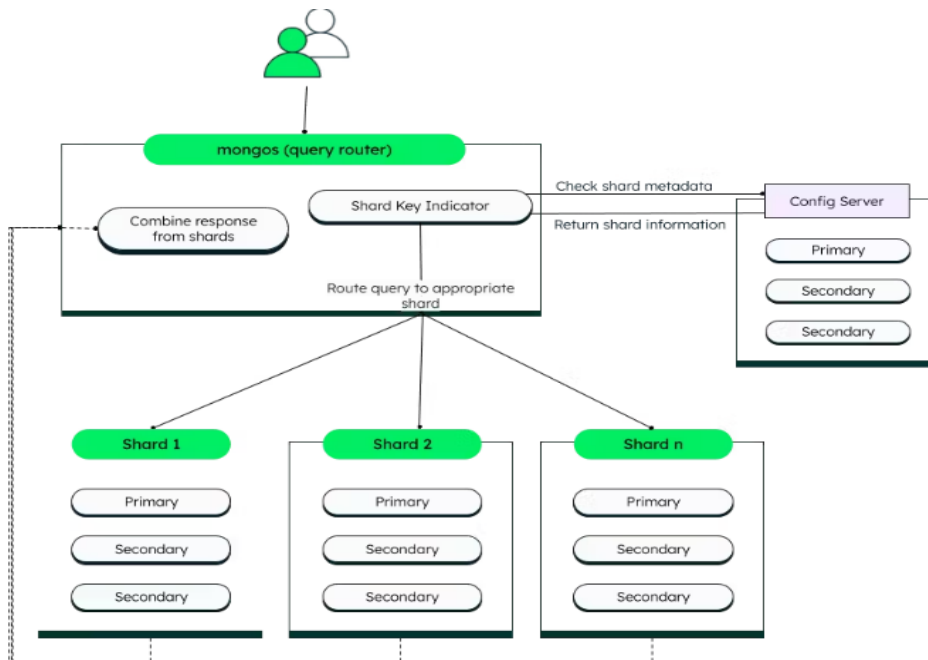
Définir dans une classe Java pour chacun de vos containers (collection, table, ...) 5 à 6 méthodes permettant de manipuler les objets (documents, lignes, valeurs, ...) qu'ils contiennent. Concernant MongoDB :

- Ecrire des méthodes (5 à 6) permettant le traitement des documents par étape (pipeline d'agrégations) avec y compris des fonctions complexes tels que les groupements, jointures, la recherche hiérarchique, ...
- Effectuer des tests.

Faire la même chose avec le deuxième moteur.

## 9. Construction d'un cluster MongoDB

Chaque groupe devra construire un cluster MongoDB avec au moins trois shard et un serveur de configuration. Voir l'architecture ci-dessus. N'hésitez pas à dokkeriser.



Voici quelques points supplémentaires à explorer :

- Sauvegarde et restauration – supervision de la base de données
- La sécurité d’une installation MongoDB (utilisateurs, droits, ...)

## 10. Les résultats de l’étude

Pour chaque moteur (MONGODB et le moteur NOSQL de votre choix) dans des dossiers séparés :

- Un vidéo (contexte, démo) obligatoire de 10 minutes environ présentant pas à pas travail
- Rapport PDF contenant les résultats des chapitres 1, 2, 3, 4 et 5. A mettre dans un word nommé : 2mcd\_conversionMCD\_Generation\_Objets \_Nom1\_Nom2\_Nom3.docx
- Le dossier des objets générées
- Les scripts de chargement des objets
- Les sources java contenant les résultats des chapitres 6, 7 et 8
- Dossier de comparaison MONGODB et du moteur NOSQL de votre choix
- Les scripts, le rapport, la construction par étape du cluster MONGODB

Tout des être bien organisé pour faciliter la lecture.

## 11. Répartition du travail

- 4 Membres par groupe (activités par membre) : Répartir équitablement le travail entre les membres d’un groupe
- Vous devez lors de la restitution identifier ce que chaque membre a fait

## 12. Ou s’inscrire pour constituer son groupe de 4

Dans le OneDrive, voici le lien :

[Voir mes mails](#)

## **13. Ou déposer ces résultats**

J'aviserais.

## **14. Evaluation des rendus**

Je vais m'appuyer sur la vidéo et les autres rendus pour corriger.