

BUT 3 Informatique

Université Sorbonne Paris-Nord

IUT de Villetaneuse

Transformation des données SQL en NoSQL : Une étude pratique avec Redis, MongoDB

Réalisé par :

Sidi ESSAADOUNI

Encadré par :

Christophe Cérin

Version 1.0

Table des matières

1	Introduction	3
2	Matériels et Méthodes	4
2.1	Outils utilisés	4
2.1.1	Architecture du projet	4
2.2	Stratégie de Migration de SQL vers NoSQL	5
2.2.1	Étape 1 :Regroupement des informations de Vols, Avions et Pilotes	6
2.2.2	Étape 2 :Intégration des Classes de Vols	7
2.2.3	Étape 3 :Gestion des Clients	7
2.2.4	Étape 4 :Représentation des Réservations	7
2.3	Conversion des données TXT en JSON	8
2.3.1	Insertion des documents dans Redis	8
2.3.2	Insertion des documents dans MongoDB	9
3	Résultats	10
3.1	Exemple de requêtes et de jointures	10
4	Discussion	11
4.1	Comparaison des Performances	11
4.2	Limites et Améliorations Possibles	11
5	Conclusion	12
A	Annexes	13
A.1	Configuration Docker pour Redis	13
A.2	Configuration Docker pour MongoDB	13
A.3	Configuration Docker pour Python	14
A.4	Création des Documents JSON	14
A.5	Insertion des documents dans Redis	16
A.5.1	Insertion des documents dans MongoDB	17

A.6 Profiling des Requêtes	18
A.7 Dépôt Git	18

Chapitre 1

Introduction

Dans le cadre de ce projet, nous avons entrepris la transformation d'une base de données relationnelle (SQL) classique en une architecture NoSQL, une démarche qui répond aux exigences modernes de flexibilité, d'évolutivité, et de gestion optimisée des données semi-structurées. Face à l'augmentation du volume de données et aux besoins croissants d'agilité dans les traitements, le modèle SQL a montré des limites, notamment en termes de performance et de rigidité. C'est pour cela que nous avons décidé de mettre en place une solution NoSQL qui permet d'adapter la structure des données aux besoins actuels.

Le choix d'une dénormalisation en format JSON, un processus permettant de simplifier et d'optimiser la manipulation des données, est au cœur de cette transformation. En effet, ce format est particulièrement bien adapté aux bases de données NoSQL, car il facilite l'agrégation des informations liées, élimine les jointures coûteuses, et permet une gestion simplifiée des données sans contrainte de schéma rigide.

Dans cette étude, nous avons utilisé deux bases NoSQL majeures : Redis et MongoDB. Redis, avec sa capacité de stockage en mémoire, offre des performances optimisées pour les lectures rapides et les données temporaires, tandis que MongoDB se distingue par sa persistance de données et sa capacité à gérer des structures JSON complexes. Ce rapport détaille les étapes de la migration, les choix techniques, et les avantages obtenus en termes de performance et de flexibilité, illustrant comment ces technologies peuvent répondre aux besoins métiers modernes.

Résumé

Ce rapport décrit le processus de migration d'une base de données SQL vers un environnement NoSQL, en adoptant une approche de dénormalisation au format JSON. En intégrant Redis et MongoDB, nous démontrons les avantages d'une architecture NoSQL pour le stockage et la manipulation de données semi-structurées, avec des gains en flexibilité et en performance. La combinaison de ces technologies permet d'adapter la gestion des données aux exigences actuelles, en simplifiant les traitements tout en assurant une évolutivité pour des volumes de données plus élevés.

Chapitre 2

Matériels et Méthodes

2.1 Outils utilisés

Voici la liste des outils utilisés pour le projet :

- Ubuntu 24.04: L'OS utilisé pour le développement mais non nécessaire suite à l'utilisation de Docker
- Docker / Docker Compose: pour orchestrer l'environnement de développement
- Redis: pour le stockage des données JSON en mémoire
- MongoDB: pour la persistance des documents JSON
- Python: pour le traitement des données et l'injection en base
- Poetry: pour gérer les dépendances
- LaTeX: pour la rédaction du rapport technique
- VSCode: pour le développement
- Git: pour le contrôle de version

Docker et Docker Compose

Docker est utilisé pour orchestrer l'environnement de développement. Le fichier `compose.yaml` définit les services, incluant Redis, MongoDB et Python (Poetry). Chacun a son propre Dockerfile (ex. `Dockerfile.redis`, `Dockerfile.mongo`). Cela garantit l'isolation des services, ce qui rend le projet facilement déployable sur différentes machines.

Redis et Redis CLI

Redis est utilisé comme base NoSQL pour le stockage des données JSON en mémoire. Les fichiers de connexion et de requêtes (`connectionRedis.py` et `requests_redis.py`) gèrent l'interface entre l'API et Redis. Redis CLI est utilisé pour des tests rapides, et un fichier `.env.redis` gère la configuration sécurisée des variables d'environnement. MongoDB est utilisé pour la persistance des documents JSON. Il est intégré via PyMongo, avec des fichiers de connexion et de requêtes spécifiques (`connectionMongo.py` et `requests_json.py`). Les données semi-structurées sont ainsi manipulées efficacement. Le fichier `Dockerfile.mongo` assure la configuration de MongoDB dans un conteneur.

Python et Poetry

Python est le langage principal pour le traitement des données et l'injection en base. Poetry est utilisé pour gérer les dépendances, avec un fichier `pyproject.toml` pour créer un environnement cohérent.

2.1.1 Architecture du projet

La figure 2.1 représentant l'architecture du projet avec Docker et bases NoSQL. Cette architecture est basé sur un principe de séparation des responsabilités en séparant en différentes couches :

- **API** : Cette couche communique avec la couche Services. Elle traite les requêtes HTTP (Get, Post, Delete) et envoie les réponses en JSON
- **DAL** : Cette couche contient les fonctions de manipulation des données. Elle communique avec les bases de données via des connexions.
- **Services** : Cette couche contient les fonctions de traitement des données. Elle communique avec les bases de données via des connexions.
- **Config** : Cette couche contient les paramètres de configuration. Elle communique avec les bases de données via des connexions.

Cette architecture est bénéfique car elle permet d'avoir un code propre, modulable et réutilisable. En effet, chaque couche peut être modifiée indépendamment, ce qui facilite la maintenance et le développement de l'application. Concernant les performances, la séparation des couches permet de réduire le temps de chargement des données, ce qui améliore la performance de l'application.

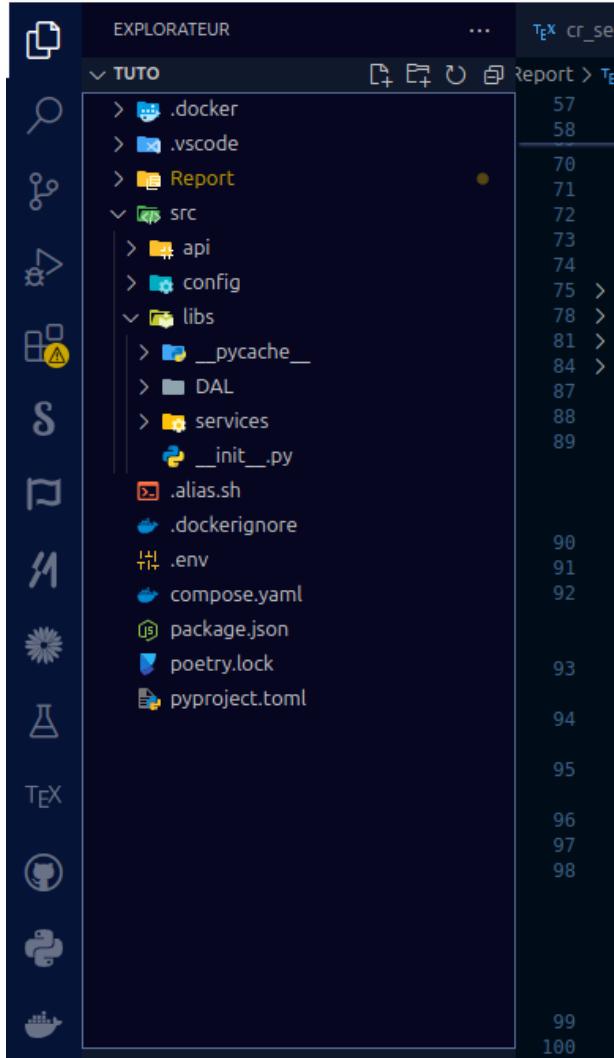


FIGURE 2.1 – Architecture du projet avec Docker et bases NoSQL

2.2 Stratégie de Migration de SQL vers NoSQL

La base de données SQL originale se composait de plusieurs tables interconnectées qui suivaient un modèle relationnel strict. Cependant, ce modèle, bien que robuste pour certains cas d'utilisation, a montré ses limites dans la gestion des données dynamiques et évolutives. Cela nous a conduit à prendre la décision de passer à une architecture NoSQL, plus adaptée aux besoins actuels.

Le schéma suivant illustre la structure de la base SQL avant migration :

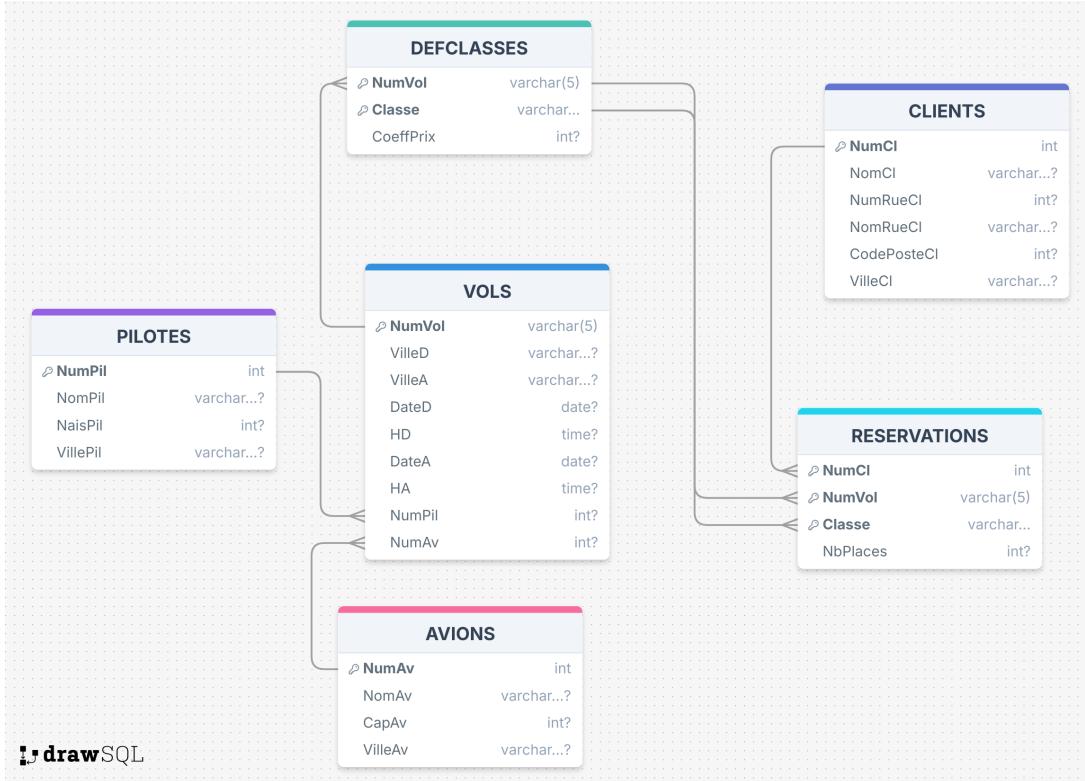


FIGURE 2.2 – Schéma de la base SQL

2.2.1 Étape 1 :Regroupement des informations de Vols, Avions et Pilotes

Dans le modèle SQL, pour récupérer toutes les informations d'un vol, il était nécessaire de joindre les tables 'VOLS', 'AVIONS' et 'PILOTES', car chaque table stockait les données d'une entité spécifique. Par exemple, la table 'VOLS' contenait les informations sur le vol, mais il fallait ensuite faire une jointure avec 'AVIONS' pour connaître les détails de l'avion, et avec 'PILOTES' pour obtenir les informations du pilote.

En NoSQL, nous avons choisi de regrouper toutes ces informations dans un seul document au sein de la collection 'Vols'. Cette approche permet de conserver toutes les données pertinentes pour un vol dans un seul document JSON.

```
Vol = {
  _id : 'V001',
  VilleD : 'Paris',
  VilleA : 'New York',
  DateD : '2024-12-01',
  HD : '08:00',
  DateA : '2024-12-01',
  HA : '14:00',
  Avion : { ... },
  Pilote : { ... }
}
```

Ce choix a été motivé par plusieurs raisons :

- **Élimination des jointures** : En SQL, récupérer ces informations nécessitait de multiples jointures. En NoSQL, en regroupant ces informations dans un seul document, nous réduisons considérablement le coût des requêtes.
- **Accès optimisé aux données** : Toutes les données relatives à un vol, un avion et un pilote étant dans un seul document, cela permet un accès plus rapide et plus direct.
- **Regroupement logique** : Les informations sur les vols, les avions et les pilotes sont souvent interrogées ensemble, il est donc logique de les stocker ensemble.

2.2.2 Étape 2 :Intégration des Classes de Vols

Dans la base SQL, les différentes classes de service ('Economy', 'Business') étaient stockées dans une table séparée appelée 'DEFCLASSES', et elles étaient liées à la table 'VOLS' par le numéro de vol. Cela nécessitait une jointure supplémentaire pour récupérer les informations des classes disponibles pour chaque vol.

Dans le modèle NoSQL, nous avons décidé d'inclure les classes disponibles directement dans une liste au sein du document 'Vol'. Chaque élément de cette liste contient la classe ('Economy', 'Business') ainsi que son coefficient de prix.

```
/* Classes disponibles pour le vol */
Classes : [
  {
    Classe : 'Economy',
    CoeffPrix : 1.0
  },
  {
    Classe : 'Business',
    CoeffPrix : 1.5
  }
]
```

Ce choix s'explique par les avantages suivants :

- **Simplification des requêtes** : En incluant les informations des classes directement dans le document 'Vol', nous évitons d'avoir à effectuer une jointure supplémentaire pour les récupérer.
- **Centralisation des données du vol** : Toutes les informations pertinentes concernant un vol (y compris les classes de service) sont désormais disponibles en une seule requête.
- **Flexibilité du modèle NoSQL** : Le format JSON nous permet de structurer ces informations sous forme de liste, ce qui correspond bien à la nature dynamique des classes de vol, et permet d'ajouter facilement de nouvelles classes si nécessaire.

2.2.3 Étape 3 :Gestion des Clients

Dans le modèle SQL, la table 'CLIENTS' stockait des informations sur les clients, notamment leur nom, adresse, et autres coordonnées. Les champs d'adresse (numéro de rue, nom de rue, code postal, ville) étaient représentés par plusieurs colonnes.

Dans la structure NoSQL, nous avons décidé d'imbriquer les informations d'adresse dans un sous-document appelé 'Adresse' pour chaque document 'Client'.

```
Client = {
  _id : 789,
  NomCl : 'Alice Dupont',
  Adresse : { ... },
  Email : 'alice.dupont@example.com',
  Telephone : '0123456789'
}
```

Ce choix a été fait pour les raisons suivantes :

- **Regroupement logique des informations** : En imbriquant les informations d'adresse dans un sous-document, nous améliorons la cohérence et la gestion des données.
- **Flexibilité** : Si des champs supplémentaires liés à l'adresse sont nécessaires à l'avenir (comme le pays ou la région), ils peuvent être ajoutés sans avoir à modifier toute la structure.

2.2.4 Étape 4 :Représentation des Réservations

Dans la base SQL, la table 'RESERVATIONS' établissait des relations entre les clients et les vols via des clés étrangères, reliant également les classes de service. Dans le modèle NoSQL, nous avons conservé ce concept de relation

en utilisant des identifiants ('VolId' et 'ClientId') pour relier les documents 'Reservation' aux documents 'Vol' et 'Client'.

```
Reservation = {
    _id : 'R001',
    VolId : 'V001',
    ClientId : 789,
    NbPlaces : 2,
    Classe : 'Economy'
}
```

Nous avons fait ce choix pour les raisons suivantes :

- **Conservation des relations critiques** : Même dans un modèle NoSQL, certaines relations doivent être maintenues. Ici, nous utilisons des identifiants pour lier les réservations aux vols et aux clients sans dupliquer inutilement les données.
- **Économie de stockage** : En stockant uniquement les références aux documents 'Vol' et 'Client', nous évitons de dupliquer les informations des vols et des clients dans chaque réservation.

2.3 Conversion des données TXT en JSON

Le script A.4 montre la création de 3 documents json à partir de nos 6 tables, ces tables peuvent être intégrées facilement à Redis en ayant comme index le nom de la collection et indice (id) et pour mongo cela peut être représenté en trois collections chacune avec ses documents. et chaque document a un id unique.

2.3.1 Insertion des documents dans Redis

Pour insérer les documents dans Redis nous utiliserons le script A.5 qui charge les données JSON, les insère dans Redis et affiche un message de succès.

la figure 2.3 représente la sortie de la commande `redis-cli` qui affiche les données insérées dans Redis.

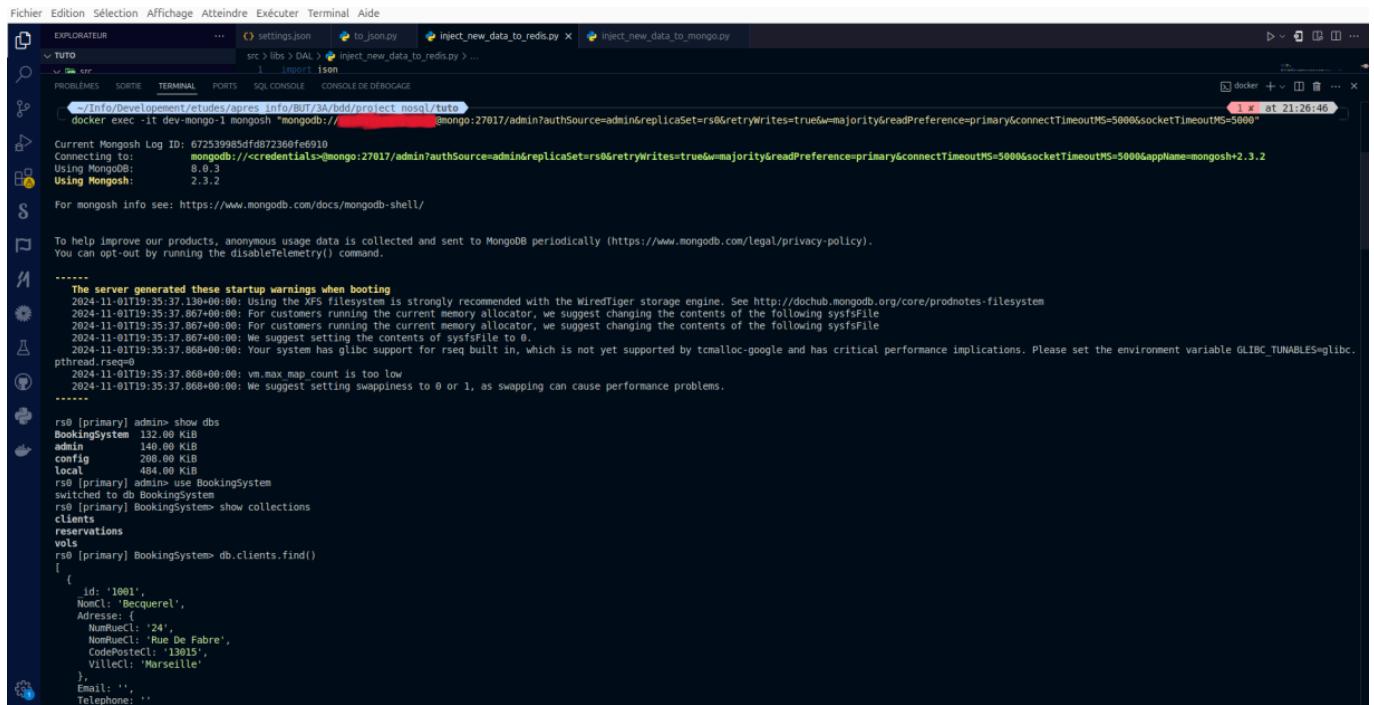
```
Fichier Edition Sélection Affichage Atteindre Exécuter Terminal Aide
FICHIERER ... settings.json to_json.py inject_new_data_to_redis.py inject_new_data_to_mongo.py
src > libs > DAL > inject_new_data_to_redis.py ...
PROBLÈMES SORTIE TERMINAL PORTS SQL CONSOLE CONSOLE DE DÉBOGAGE
` 1 insert JSON
2023-05-15T21:33:43+00:00 [INFO] [Redis] INFO 2023-05-15T21:33:43+00:00 [INFO] [Redis] 127.0.0.1:6379> keys *
0K
1) "vol:V603"
2) "vol:V211"
3) "reservation:R025"
4) "vol:V220"
5) "vol:V210"
6) "vol:V628"
7) "vol:V889"
8) "vol:V622"
9) "vol:V688"
10) "vol:V151"
11) "reservation:R043"
12) "vol:V518"
13) "client:1038"
14) "vol:V927"
15) "vol:V690"
16) "vol:V614"
17) "client:1001"
18) "vol:V209"
19) "vol:V623"
20) "vol:V909"
21) "vol:V805"
22) "client:1027"
23) "client:1025"
24) "vol:V606"
25) "client:1028"
26) "vol:V214"
27) "reservation:R047"
28) "vol:V613"
29) "vol:V626"
30) "vol:V606"
31) "vol:V180"
32) "vol:V157"
33) "vol:V153"
34) "vol:V511"
35) "reservation:R004"
36) "reservation:R013"
37) "vol:V621"
38) "vol:V213"
39) "vol:V635"
40) "vol:V886"
41) "vol:V627"
42) "reservation:R017"
43) "reservation:R030"
44) "vol:V157"
45) "vol:V602"
```

FIGURE 2.3 – Insertion des documents dans Redis

2.3.2 Insertion des documents dans MongoDB

Pour insérer les documents dans MongoDB nous utiliserons le script A.5.1 qui charge les données JSON, les insère dans MongoDB et affiche un message de succès.

la figure 2.4 représente la sortie de la commande mongo qui affiche les données insérées dans MongoDB.



The screenshot shows a terminal window with several tabs at the top: "Fichier", "Édition", "Sélection", "Affichage", "Atteindre", "Exécuter", "Terminal", and "Aide". The "TERMINAL" tab is active, displaying the mongo command output. The command "mongo:27017/admin?authSource=admin&replicaSet=rs0&retryWrites=true&w=majority&readPreference=primary&connectTimeoutMS=5000&socketTimeoutMS=5000" is run, followed by "Using Mongosh: 2.0.3" and "Using MongoDB: 2.0.3". The mongo shell then lists the contents of the "clients" collection, showing a single document with the ID "1001".

```
Current Mongosh Log ID: 672539985fd872360fe010
Connecting to: mongo:27017/admin?authSource=admin&replicaSet=rs0&retryWrites=true&w=majority&readPreference=primary&connectTimeoutMS=5000&socketTimeoutMS=5000
Using Mongosh: 2.0.3
Using MongoDB: 2.0.3

For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2024-11-01T19:35:37.130+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2024-11-01T19:35:37.867+00:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2024-11-01T19:35:37.867+00:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2024-11-01T19:35:37.867+00:00: We suggest setting the contents of sysfsFile to 0.
2024-11-01T19:35:37.868+00:00: Your system has glibc support for rseq built in, which is not yet supported by tcmalloc-google and has critical performance implications. Please set the environment variable GLIBC_TUNABLES=glibc.
pthread_rseq=0
2024-11-01T19:35:37.868+00:00: vm.max_map_count is too low
2024-11-01T19:35:37.868+00:00: We suggest setting swappiness to 0 or 1, as swapping can cause performance problems.
-----
rs0 [primary] admin> show dbs
BookingSystem   132.00 KB
admin           149.00 KB
config          208.00 KB
local           484.00 KB
rs0 [primary] admin> use BookingSystem
switched to db BookingSystem
rs0 [primary] BookingSystem> show collections
clients
reservations
vols
rs0 [primary] BookingSystem> db.clients.find()
{
  "_id": "1001",
  "NomCle": "Becquerel",
  "AdresseCle": "24",
  "Num RueCle": "24",
  "Nom RueCle": "Rue De Fabre",
  "Code PostalCle": "13015",
  "VilleCle": "Marseille"
},
Email: ...
Telephone: ...
```

FIGURE 2.4 – Insertion des documents dans MongoDB

Chapitre 3

Résultats

Les résultats de la migration sont présentés sous forme de documents JSON pour chaque entité, illustrant les avantages de la dénormalisation et les structures simplifiées obtenues. On peut visualiser les résultats en effectuant des requêtes vers la base de données NoSQL et donc au JSON car le format JSON est utilisé dans redis et mongo en tant que document.

Pour illustrer les avantages de la dénormalisation, nous avons utilisé des exemples de requêtes et de jointures pour montrer comment les données peuvent être manipulées et agrégées de manière plus efficace. Ces exemples sont présentés dans la section suivante.

3.1 Exemple de requêtes et de jointures

En ayant convertit nos tables SQL en JSON, nous pouvons normaliser les requêtes et les jointures pour manipuler les données de manière plus efficace que les données viennent d'un fichier JSON, récupéré par redis ou par mongo.

Chapitre 4

Discussion

Cette section analyse comment les résultats de la migration répondent aux objectifs fixés et aux besoins d'optimisation des performances. Nous avons mesuré les performances de chaque base de données à travers plusieurs requêtes (profiling détaillé en annexe A.6).

4.1 Comparaison des Performances

Les résultats montrent que Redis offre des temps de réponse plus rapides pour les opérations de lecture, tandis que MongoDB est plus adapté pour les requêtes complexes grâce à son système de stockage JSON natif.

4.2 Limites et Améliorations Possibles

Bien que NoSQL ait permis une flexibilité accrue, certains cas d'utilisation nécessitent encore une réflexion pour l'optimisation des écritures massives.

Chapitre 5

Conclusion

La migration de SQL vers NoSQL, en utilisant Redis et MongoDB, a permis d'optimiser l'accès aux données tout en offrant une flexibilité accrue pour gérer des données semi-structurées. Les performances, analysées dans la section Discussion, montrent les avantages de chaque base de données selon les types de requêtes, ce qui confirme leur complémentarité pour des applications modernes.

Annexe A

Annexes

A.1 Configuration Docker pour Redis

```
redis:
  build:
    context: ./docker/redis/
    dockerfile: Dockerfile.redis
  ports:
    - "${REDIS_PORT}:${REDIS_PORT}"
  volumes:
    - ./docker-data/redis/data:/data
    - ./docker-data/redis/logs:/var/log/redis
  restart: always
  env_file:
    - ./docker/redis/.env.redis
  environment:
    - REDIS_ARGS=--requirepass ${REDIS_PASSWORD}
  healthcheck:
    test: ["CMD", "redis-cli", "-a", "${REDIS_PASSWORD}", "ping"]
    interval: 30s
    timeout: 10s
    retries: 5
```

A.2 Configuration Docker pour MongoDB

```
mongo:
  build:
    context: ./docker/mongo/
    dockerfile: Dockerfile.mongo
  ports:
    - 27017:27017 # Expose le port 27017
    - 27018:27018 # Expose le port 27018
    - 27019:27019 # Expose le port 27019
  restart: always # Redémarre toujours en cas de crash
  volumes:
    - ./docker-data/mongo/data:/data/db # Monte le volume de données MongoDB
  env_file:
    - ./docker/mongo/.env.mongo # Fichier d'environnement pour MongoDB
  healthcheck: # Vérifie l'état de santé de MongoDB
    test: echo 'db.runCommand("ping").ok' | mongosh mongo:27017/test --quiet
    interval: 30s # Intervalle entre les vérifications
    timeout: 10s # Délai d'expiration de la commande
    retries: 5 # Nombre de tentatives avant de considérer le service comme défaillant
```

```

start_period: 20s # Temps avant le démarrage des vérifications
networks:
  dev:
    aliases:
      - localhost # Alias pour le service MongoDB dans le réseau Docker

```

A.3 Configuration Docker pour Python

```

poetry:
  build:
    context: ./docker/poetry/
    dockerfile: Dockerfile.poetry
    stdin_open: true
    tty: true
    volumes:
      - .:/workspace
    working_dir: /workspace
  environment:
    - POETRY_VIRTUALENVS_IN_PROJECT=true
  command: [ "/bin/sh" ]

```

A.4 Crédit des Documents JSON

```

import json
import os

# Table de correspondance des colonnes
tableCorespondance = {
  "AVIONS.txt": ["NumAv", "NomAv", "CapAv", "VilleAv"],
  "CLIENTS.txt": ["NumCl", "NomCl", "NumRueCl", "NomRueCl", "CodePosteCl", "VilleCl"],
  "DEFCLASSES.txt": ["NumVol", "Classe", "CoefPrix"],
  "PILOTES.txt": ["NumPil", "NomPil", "NaisPil", "VillePil"],
  "RESERVATIONS.txt": ["NumCl", "NumVol", "Classe", "NbPlaces"],
  "VOLS.txt": ["NumVol", "VilleD", "VilleA", "DateD", "HD", "DateA", "HA", "NumPil", "NumAv"],
}

# Lire les fichiers .txt et les stocker dans un dictionnaire
dictAllJson = {}
for fileName in os.listdir("src/libs/db/txt"):
  if fileName.endswith(".txt"):
    file_path = os.path.join("src/libs/txt", fileName)
    fields = tableCorespondance[fileName]

    # Vérifier si la première colonne est une clé unique
    if fileName in ["DEFCLASSES.txt", "RESERVATIONS.txt"]:
      # Stocker les données dans une liste pour ces fichiers
      dictAllJson[fileName] = []
      with open(file_path, 'r') as fh:
        for line in fh:
          description = list(line.strip().split("\t"))
          if len(description) < len(fields):
            # Gérer les lignes avec des colonnes manquantes
            continue
          data_entry = {}
          for i, categorie in enumerate(fields):
            data_entry[categorie] = description[i]
          dictAllJson[fileName].append(data_entry)
    else:
      # Utiliser un dictionnaire avec la première colonne comme clé

```

```

dictAllJson[file_name] = {}
with open(file_path, 'r') as fh:
    for line in fh:
        description = list(line.strip().split("\t"))
        if len(description) < len(fields):
            # Gérer les lignes avec des colonnes manquantes
            continue
        key = description[0]
        data_entry = {}
        for i, categorie in enumerate(fields[1:], start=1):
            data_entry[categorie] = description[i]
        dictAllJson[file_name][key] = data_entry

# Construire la liste des vols
vols_list = []
for vol_id, vol_data in dictAllJson["VOLS.txt"].items():
    vol_dict = {"_id": vol_id}

    # Copier les champs du vol
    vol_fields_mapping = {
        "VilleD": "VilleD",
        "VilleA": "VilleA",
        "DateD": "DateD",
        "HD": "HD",
        "DateA": "DateA",
        "HA": "HA",
    }
    for src_field, dest_field in vol_fields_mapping.items():
        vol_dict[dest_field] = vol_data.get(src_field, "")

    # Ajouter l'avion
    num_av = vol_data.get("NumAv", "")
    avion_data = dictAllJson["AVIONS.txt"].get(num_av, {})
    vol_dict["Avion"] = {
        "NumAv": num_av,
        "NomAv": avion_data.get("NomAv", ""),
        "CapAv": avion_data.get("CapAv", ""),
        "VilleAv": avion_data.get("VilleAv", ""),
    }

    # Ajouter le pilote
    num_pil = vol_data.get("NumPil", "")
    pilote_data = dictAllJson["PILOTES.txt"].get(num_pil, {})
    vol_dict["Pilote"] = {
        "NumPil": num_pil,
        "NomPil": pilote_data.get("NomPil", ""),
        "NaisPil": pilote_data.get("NaisPil", ""),
        "VillePil": pilote_data.get("VillePil", ""),
    }

    # Ajouter les classes
    vol_dict["Classes"] = []
    for classe in dictAllJson["DEFCLASSES.txt"]:
        if classe.get("NumVol") == vol_id:
            vol_dict["Classes"].append({
                "Classe": classe.get("Classe", ""),
                "CoeffPrix": classe.get("CoeffPrix", ""),
            })

    vols_list.append(vol_dict)

# Construire la liste des clients
clients_list = []

```

```

for client_id, client_data in dictAllJson["CLIENTS.txt"].items():
    client_dict = {
        "_id": client_id,
        "NomCl": client_data.get("NomCl", ""),
        "Adresse": {
            "NumRueCl": client_data.get("NumRueCl", ""),
            "NomRueCl": client_data.get("NomRueCl", ""),
            "CodePosteCl": client_data.get("CodePosteCl", ""),
            "VilleCl": client_data.get("VilleCl", ""),
        },
        "Email": "",          # Champs supplémentaires à remplir si nécessaire
        "Telephone": "",    # Champs supplémentaires à remplir si nécessaire
    }
    clients_list.append(client_dict)

# Construire la liste des réservations
reservations_list = []
reservation_id_counter = 1
for reservation_data in dictAllJson["RESERVATIONS.txt"]:
    reservation_dict = {
        "_id": f"R{str(reservation_id_counter).zfill(3)}",
        "VolId": reservation_data.get("NumVol", ""),
        "ClientId": reservation_data.get("NumCl", ""),
        "NbPlaces": reservation_data.get("NbPlaces", ""),
        "Classe": reservation_data.get("Classe", ""),
    }
    reservations_list.append(reservation_dict)
    reservation_id_counter += 1

# Écrire les données dans des fichiers JSON séparés
with open("src/libs/db/json/vols.json", "w") as f:
    json.dump(vols_list, f, indent=2, ensure_ascii=False)

with open("src/libs/db/json/clients.json", "w") as f:
    json.dump(clients_list, f, indent=2, ensure_ascii=False)

with open("src/libs/db/json/reservations.json", "w") as f:
    json.dump(reservations_list, f, indent=2, ensure_ascii=False)

print("Données extraites avec succès.")

```

A.5 Insertion des documents dans Redis

```

import json
import sys, os

# Définir le chemin
sys.path.append(os.path.abspath(os.path.join(os.path.dirname(__file__), '..', '..')))

# Définir le chemin vers le dossier contenant les fichiers JSON
SRC = os.path.abspath(os.path.join(os.path.dirname(__file__), '..', '..'))
DATA_DIR = os.path.join(SRC, 'libs', 'db', 'json')

from config.connectionRedis import connect_redis as connect

# Charger les données JSON
with open(os.path.join(DATA_DIR, 'vols.json'), 'r', encoding='utf-8') as f: vols = json.load(f)
with open(os.path.join(DATA_DIR, 'clients.json'), 'r', encoding='utf-8') as f:
    clients = json.load(f)
with open(os.path.join(DATA_DIR, 'reservations.json'), 'r', encoding='utf-8') as f:
    reservations = json.load(f)

```

```

r = connect()

def clear_redis_database():
    r.flushdb()

clear_redis_database()

# Insérer les vols dans Redis
for vol in vols:
    vol_id = vol['_id']
    # Stocker le vol sous la clé 'vol:{vol_id}'
    r.set(f'vol:{vol_id}', json.dumps(vol))

# Insérer les clients dans Redis
for client in clients:
    client_id = client['_id']
    # Stocker le client sous la clé 'client:{client_id}'
    r.set(f'client:{client_id}', json.dumps(client))

# Insérer les réservations dans Redis
for reservation in reservations:
    reservation_id = reservation['_id']
    # Stocker la réservation sous la clé 'reservation:{reservation_id}'
    r.set(f'reservation:{reservation_id}', json.dumps(reservation))

print("Données insérées avec succès dans Redis.")

```

A.5.1 Insertion des documents dans MongoDB

```

import sys, os, json

sys.path.append(os.path.abspath(os.path.join(os.path.dirname(__file__), '..', '..')))

from config.connectionMongo import connect_mongodb as connect

# Définir le chemin vers le dossier contenant les fichiers JSON
SRC = os.path.abspath(os.path.join(os.path.dirname(__file__), '..', '..'))
DATA_DIR = os.path.join(SRC, 'libs', 'DAL', 'db', 'json')

# Charger les données JSON
with open(os.path.join(DATA_DIR, 'vols.json'), 'r', encoding='utf-8') as f: vols = json.load(f)
with open(os.path.join(DATA_DIR, 'clients.json'), 'r', encoding='utf-8') as f: clients = json.load(f)
with open(os.path.join(DATA_DIR, 'reservations.json'), 'r', encoding='utf-8') as f: reservations = json.load(f)

# Se connecter à MongoDB
client = connect()

db = client.get_database('BookingSystem')

# Créer ou obtenir les collections
vols_collection = db.get_collection('vols')
clients_collection = db.get_collection('clients')
reservations_collection = db.get_collection('reservations')

# Effacer les collections existantes pour éviter les doublons lors de ré-exécutions
vols_collection.delete_many({})
clients_collection.delete_many({})
reservations_collection.delete_many({})

# Insérer les documents JSON dans chaque collection respective
vols_collection.insert_many(vols)

```

```
clients_collection.insert_many(clients)
reservations_collection.insert_many(reservations)

print("Les données ont été insérées avec succès dans MongoDB.")
```

A.6 Profiling des Requêtes

Détails de la performance des requêtes dans chaque base

A.7 Dépôt Git

Vous trouverez le dépôt Git de ce projet sur GitHub : <https://github.com/>.

Pour pouvoir tester le projet vous pourrez suivre les étapes mises dans le fichier README.md.