RDBMS2NOSQL: A Migration Model From RDBMS to NoSql

Yasser Aderghal. Author[1]

1. *E-mail any correspondence to: yasser.aderghal@gmail.com*

Abstract

The need to store large amount of data has increased last decades while we entering new chapter of Web, a web that require to store as much as possible of data so that enterprises could compete each other, as a result, legacy database systems needed to be upgraded to fill the deficit.

At the present, new systems emerged into the market and one of them is mongodb which of our concern in this project, as we are trying to propose a model that can transform data from a relational database management system (mysql in our case) to document-oriented database system (mongodb). Fortunately , our model passed tests on different databases and ready for use but it drops in performance while processing large files.

Keywords: Mysql; Mongodb; Migration; Django; Python

Introduction

Last decades , technologies faced a tremendous advancements as well as the growth of businesses . And as well known , database is the backbone of any enterprise that trying to compete on high level , therefore, database also should be upgraded . Most of the time enterprise used to just upgrade their legacy systems to newer and more secure systems , but recently as we generate big data and software companies trying to develop different type of databases .We are facing huge demand on data migration mostly from one type of database to another type . And that ,lead use to take a close look of how these data migration programs or services (if we called that) and how can we benefit from these services .

The report is arranged as follow: First section we gave a brief definitions and background on databases followed by example of programs that we are trying to mimic. Additionally , a section explaining how we tackled the problem and the dependencies. Finally , an example of the transformation is given and a brief conclusion.

Background

Databases it's types

In simple words, data can be facts related to any object in consideration. For example, your name, age, height, weight, etc. are some data related to you. A picture, image, file, pdf, etc. can also be considered data.

A database is a systematic collection of data. They support electronic storage and manipulation of data. Databases make data management easy. There exist different type of database the most known and traditional one is Relational database , this type of database defines database relationships in the form of tables. The name comes from the mathematical notion of "relation." which is known now for relational algebra. Also exist graph-oriented database which are based on graph theory to store and manipulate data . Document-oriented database is also another type of databases usually called noSql , the data is kept in document collections, usually using the XML, JSON, BSON formats. One record can store as much data as you want, in any data type (or types) you prefer.

Database migration

Database migration is the process of migrating a data from one or more source database to one or more target database using a database service or programs . When a migration is finished, the dataset in the source databases resides fully, though possibly restructured, in the target databases. Clients that accessed the source databases are then switched over to the target databases, and the source databases are turned down.

Mysql & Mongodb

MySQL is an open-source relational database management system (RDBMS).Its name is a combination of "My", the name of co-founder Michael Widenius's daughter ,and "SQL", the abbreviation for Structured Query Language.

Founded in 2007, MongoDB Inc. was a new approach to databases design. MongoDB created a way to store

the "humongous" amount of data necessary for scalable use-cases, and the "Mongo" in the name is short for humongous. MongoDB's design is in response to a need for nimble, information-rich database performance. It stores data as MongoDB documents, which is the touchstone of its design.

Related work

Mongify is a utility written in the ruby language and is used to migrate databases from SQL to mongodb. Mongify utility migrates databases while not caring about primary keys and foreign keys as in case of an RDBMS. It supports data migration from MySQL, SQLite and other relational databases . [1]

Studio 3T is another utility that can let use migrate , it already has a SQL Import to MongoDB feature, which allows you to import records from a single table into a single collection.

Methodology

Dependencies

This project is developed on linux environment, and written in the python language . Libraries used here are as follow :

```
django==3.2.5
sql-metadata==2.3.0
gunicorn==20.1.0
django-cors-headers==3.11.0
```

Software used for deployment:

```
Docker==20.10.12
nginx==1.19.0
```

(**NOTE**) : It is preferably to run this application on docker, it runs on port 1300 . otherwise , if it ran directly,you must change debug mode to true.

(**NOTE**) : Readme file shows how to run this application. The source code can be downloaded [2].

Model

Despite the complexity of the sql documents , we realized that there exist patterns. For every command in the sql file, a comment above it ends with (− and a new line "\n"), therefore , the program split the file based on that which results in a list of all commands and configurations in the file .

Secondly , the app select elements of that list which it had "INSERT" or "FOREIGN KEY". Afterward , and by using a library which facilitate the use of regular expression , we took attributes and values of "INSERT" command , and combine the two using `zip()` to return a dictionary object. As for "FOREIGN KEY" , it's relatively easy to extract our concerned information from it by using " ` " a character for splitting and store them in list of lists.

Finally, the app looks for table name , foreign key , referenced table and referenced key for every list of "FOREIGN KEYS" and then join the tables if it's needed.

Example

This subsection expose what we have tried to explain previously. If we consider the two tables below (tables 1 2), this application let you choose to join them (left join or\and right join) as the prototype (Listing 1) show, or returning simple json files without references of other tables, (Listing 2) in this case.

code	numero	nom
0	22	Etudiant 0
1	21	Etudiant 1
2	27	Etudiant 2
3	14	Etudiant 3
4	37	Etudiant 4

Table 1: Etudiant table.

numero	libelle
0	Module 0
1	Module 1
2	Module 2
3	Module 3
4	Module 4

Table 2: Filiere table.

```json
[{
    "numero": 0,
    "libelle": "Module 0",
    "etudiant": [
        {
            "code": 29,
            "numero": 0,
            "nom": "Etudiant 29"
        },
        {
            "code": 61,
            "numero": 0,
            "nom": "Etudiant 61"
        },
        {
            "code": 78,
            "numero": 0,
            "nom": "Etudiant 78"
        }
    ]
}]
```

Listing 1: filiere.json

```
[{
    "code": 0,
    "numero": 22,
    "nom": "Etudiant 0"
}]
```

Listing 2: etudiant.json

Conclusion

In this report , we presented and explained the purpose of our project. An introduction and a brief background on database with it's types and database migration is given at the beginning of the report , as we continued explaining how we attacked the problem along with an example of the transformation from two tables of which were given by the professor to json file format .

After testing on various databases (in size to be precise) , the results validate that data processing algorithms need to be improved or changed completely, as we noticed drops in performance while the program trying to process large files. Additionally, the program can only process dumped files out-of phpMyAdmin .

References

1. Ran. From mysql to nosql! how to migrate your mysql data to mongodb using mongify utility. Unixmen 2022 Feb. Available from: https://www.unixmen.com/migrate-mysql-to-mongodb-using-mongify-utility-linux/

2. Aderghal Y. Yasseraderghal/MnMmigrating: MnMmigrating is a program for converting SQL files to JSON. GitHub. Available from: https://github.com/YasserAderghal/MnMmigrating.git