CSCI 725 : NoSQL and NewSQL Data Systems

Assignment 2: MongoDB

Data Modeling:

The database I have chosen is MongoDB. It stores data in the document form. It can have Documents and Arrays inside Documents (Embedded Documents, Arrays). It has JSON like Document schema. Each document has a unique identifier ("_id") which is like an object. Just like Table in MySQL, it has Collections which have multiple documents stored in it.
For the data in this assignment,

- I have created a **schema "ITEM".**
- It has **2 collections namely : "Users" and "Products"**. "Users" store all the information related to a User. "Products" store all the information related to a product.
- **"Users"** Collection has documents and each document has following fields:
    1. **_id** : Unique identifier. It stores username in String form.
    2. **Password**: It stores password of user
    3. **firstName**: It stores first name of user
    4. **lastName**: It stores last name of user
    5. **orders**: It is an array of Documents. A user can place multiple orders. Each Document inside this array has field:
        a. Orderid: It has the id of order. Each order will have a unique different id.
        b. Date: date when order is placed.
        c. cart: It is an array of Documents. Each Document has fields : "product" and "quantity".
        Each cart will have a list of documents storing different products and quantities. An order will consist of multiple products and the quantity of each product.
    6. **productrating**: It is an array of Documents. Each user can give rating to multiple products. So this array consists of all the products the user has rated and the rating given. Each Document inside this array has field:
        a. product: stores product id.
        b. rating: Stored rating given by user.
- **"Products"** Collection has documents and each document has following fields:
    1. **_id**: Unique identifier. It stores product id.
    2. **name**: It stores the name of a product.
    3. **description**: It stores description of a product.
    4. **price**: It stores the cost of a product.
    5. **initialStock**: It stores the initial stock quantity of the product.
    6. **review**: It is an array of Documents. A product can have review by multiple users. Each user can give review to multiple different products. A user cannot give

review to same product again (this case is handled in code). It stores information related to review given by a user for a product. Each document in it has fields :

a. user: username of user
b. text: Review text given by user.
c. rating: rating given by user
d. date: Date of the given review.

I am using property of MongoDB to store embedded arrays and documents inside documents to store all the related data together. So Array of Order Documents have all the items contained in that item, their quantity and date of order. It is all stored together. Since User places an order, it is stored in the “Users” Collection.
User also gives rating so storing all the products and their ratings given by user in “productrating” array in “Users”. This way stores all the information related to User together in Users Collection. This helps to efficiently find the average rating for a product given by a user as we dont need to scan the entire data.

Array of Review Documents have review details given by multiple users for a product. So we can see all the reviews and rating for a product given by multiple users, by just matching the productID. Thus, it doesn’t need to scan the entire data. So, reviews and all information related to products are stored in “Products”. It provides data integrity as everything is stored at one place.