CSL7090 Assignment 2

Saurabh Shashikant Burewar (B18CSE050)

MongoDB and SQL

MongoDB setup

- We are running MongoDB on our system and not Atlas.
- Mongo server runs in the background in the system and all the operations are performed on the mongo shell.

1. Create database

The "use" command creates a new database and switches to it.

```
> show dbs
admin 0.000GB
config 0.000GB
local 0.000GB
> use LibraryDB
switched to db LibraryDB
```

We can create a new collection using - db.createCollection("Books").

2. Add documents

All entries are added manually in the database using db.Books.insert().

```
ob.Books.insert(("ISBN:"978013250884", "Accession Not:"108230900011", "Title":"Cleam Code: A Hundbook of Agile Software Craftsmanship", "Author":"Robert C. Mortin", "Publisher":"Prentice Hall, 2009", "Edition ":"Illustrated", "Year of Publication":"2009", "Category":"DBS", "Total_number_of_pages":"All? "Price":"55.40")

ob.Books.insert(("ISBN:"978013457704", "Accession_Mo":"201120180005", "Title":"Refactoring: Improving the Design of Existing Code", "Author":"Nartin Fowler", "Publisher":"Addison-Wesley Professional, 2018", "Edition":"2", "Year of Publication":"2018", "Category":"Jown", "Total_number_of_pages":"1448", "Price":"75.40")

ob.Books.insert(("ISBN:"978013433136", "Accession_Mo":"108012019090014", "Title":"The C. Book, Featuring the AUSI C Standard", "Author":"Nichael Francis Banaham", "Publisher":"Addison-Wesley Publishing Company", Writee:"120.90"))

ob.Books.insert(("ISBN:"978133051120", "Accession_Not:"1080420160022", "Title":"Tython for Everybody: Exploring Data Using Python 3", "Author":"Charles R. Severance", "Publisher":"Charles R. Severance", "Edition":"1", "Year of publication":"2016", "Category:"Python", "Total_number_of_pages":"231", "Price":"79.00"))

ob.Books.insert(("ISBN:"978140935690", "Accession_Not:"128020130016", "Title":"Learning Python: Powerful Object-Oriented Programming", "Author":"Mark Lutz", "Publisher":"O'Reilly Media", "Edition":"5", "Year of Publication":"231)

ob.Books.insert(("ISBN:"978140935690"), "Accession_Not:"128020130000", "Title":"Learn to Program with C", "Author":"Noel Kalicharan", "Publisher":"Apress", "Edition":"2", "Year_of_Publication":"2015", "Category:"19thon", "Total_number_of_pages":"108.90")

ob.Books.insert(("ISBN:"978180431311", "Accession_Not:"131220150007", "Title":"Learn to Program with C", "Author":"Noel Kalicharan", "Publisher":"Apress", "Edition":"2", "Year_of_Publication":"2015", "Category:"19thon", "Total_number_of_pages":"704.9", "Price":"20.99")

writeResult(("Ininserted": 1))

ob.Books.insert(("ISBN:"97808365745", "Accession_Not:"0112010000
```

3. Create a replica set (important when writing script)

Next, we create a replica set which is needed for us to watch for changes in the database. To do this, we first make some changes to the mongo configuration file to enable replication and restart the server to incorporate the changes. Then, we open the mongo shell and start replication using *rs.initiate()*. This returns a log of successful start.

MySQL setup

- MySQL installed on system
- Create database using shell

mysql -u "root" -p"<password>" -e "CREATE DATABASE librarydb"

MongoDB to MySQL script

- 1. Libraries used
 - Pymongo
 - Mysql-connector-python
- 2. Get MongoDB data and add to MySQL
 - We create a client to access mongodb server. We use standard functions to access the collections and get all the documents in the collections.
 - We create a connection to MySQL database and create a table with all the required columns like ISBN, Accession No., etc.
 - We parse all the documents in MongoDB and insert them into the SQL table.
 - We run a select query to print all the entries in the SQL table.
 - This Script parses all the documents in the collection and adds them in the MySQL table, so if the run is interrupted, we won't lose our progress.

```
SS.D:LITIAth year/Sem VIIINSEVASsignments/22 python mongotosql.py

cilysers/Saurabh/AppoRata/Local/Programs/Python/Python/Slibisite-packages/requests\_init__py:89: RequestsDependencyMarning: urllib3 (1.26.7) or chardet (3.0.4) doesn't match a supported "carnot find "mokoks" collection in the database...

('O788133757884', '018208908011', 'clean Code: A Handbook of Agile Software Craftsmanship', 'Robert C. Martin', 'Prentice Hall, 2009', 'illustrated', '2009', 'DRMS', '431', '49.99')

('0788133757844', '03112888085', '9featoring: Improving the Design of Existing Code', 'Martin Fowler', 'ddison-Wesley Professional, 2018', '2', '2018', '348', '55.49')

('078813475744', '201128080805', 'Python for Everybody: Exploring Data Using Python 3', 'Charles R. Severance', 'Charles Severance', '1', '2015', 'Python', '231', '79.00')

('078134091510922', 'Python', '21, '191', 'c', '310', '79.00')

('078134091510921', '120620130016', 'Learning Python: Powerful Object-Oriented Programming', Mark Lutz', '0' Melly Media', '5', '2015', 'Python', '1648', '29.99')

('07803185706354', '01112008034', 'Database Management System (ORMS)s A Practical Approach', 'Bython', '21, '21, '2', '2018', '1048', '29.99')

('07803185706373', '011120080034', 'Python Programming: An Introduction to Computer Science', 'John N. Lelle', 'Franklin, Beedle', 'a', '2004', 'Python', '517', '50.99')

('07803185706371', '03102800034', 'Python Brogramming: An Introduction to Computer Science', 'John N. Lelle', 'Franklin, Beedle', 'a', '2004', 'Python', '517', '50.99')

('07803185706171', '0310280010033', 'Brogramming: An Introduction to Computer Science', 'John N. Lelle', 'Franklin, Beedle', 'a', '2004', 'Python', '517', '50.99')

('07803185706171', '0310280010003', 'Python', '51004', '1241 McGraw Hill', '3', '2011', '1088', '780', '38.99')

('07803185706170', '0310280010003', 'Programming: An Introduction to the Python Language', 'Mark Summerfield', 'Addison-Wesley', 'illustrated 1', '2010', 'Python', '630', '92.99')
```

3. Listening to MongoDB and updating MySQL

The script keeps running to listen to any changes in MongoDB using db.collections.watch().

```
PS D:\IIIJ\4th year\Sem VIII\SDE\Assignments\2> python mongotosql.py
C:\Users\Saurabh\AppData\Local\Programs\Python\Python38\lib\site-packages\requests\_init__.py:89: RequestsDependencyWarning: urllib3 (1.26.7) or chardet (3.0.4) doesn't match a supported version!
warningswarn("urllib3 ({}) or chardet ({}}) doesn't match a supported "
Cannot find "Books" collection in the database...
('978013258084', '0180820090011', 'Clean Code: A Handbook of Agile Software Craftsmanship', 'Robert C. Martin', 'Prentice Hall, 2009', 'illustrated', '2009', 'DEWS', '431', '49.99')
('9780134787704', '201120180005', 'Refactoring: Improving the Design of Existing Code', 'Martin Fowler', 'Addison-Wesley Professional, 2018', '2', '2018', 'Jav a', '448', '55.49')
('978020154336', '010219990014', 'The C Book, Featuring the ANSI C Standard', 'Michael Francis Banahan', 'Addison-Wesley Publishing Company', '2', '1991', 'C', '310', '26.99')
('9781530051120', '090420160022', 'Python for Everybody: Exploring Data Using Python 3', 'Charles R. Severance', 'Charles Severance', '1', '2016', 'Python', '2 31', '79.00')
('9781484213711', '151220150007', 'Learning Python: Powerful Object-Oriented Programming', 'Mark Lutz', "O'Reilly Media", '5', '2013', 'Python', '1,648', '20.9 '9')
('9781887902991', '011220040034', 'Database Management System (DBMS): A Practical Approach', 'Rajiv Chopra', 'S. Chand Limited', '5', '2010', 'DBMS', '704', '2 0.99')
('9781887902991', '011220040034', 'Python Programming: An Introduction to Computer Science', 'John M. Zelle', 'Franklin, Beedle', '4', '2004', 'Python', '517', '50.99')
('9780321680563', '011120100013', 'Programming in Python 3: A Complete Introduction to the Python Language', 'Mark Summerfield', 'Addison-Wesley', 'illustrated '1', '20780332586017', '031020010003', 'Effective Java', 'Joshua Bloch', 'Pearson India', '3', '2001', 'Java', '264', '43.86')
Listening to changes in MongoD8 database....
```

Now, we open a mongo shell to make changes in our mongo database. We insert two documents in the database.

```
rs0:PRIMARY> db.Books.insert(("ISBN": "9781484213711", "Accession_No": "151220150045", "Title": "Learn to Program with C", "Author": "Noel Kalicharan", "Publisher": "Apress", "Edition": "2", "Year_of_Publication": "2015", "Category": "C", "Total_number_of_pages": "312", "Price": "45.99" }\
MriteResult(C'innserted": 1)
rs0:PRIMARY> db.Books.insert(("ISBN": "9781449355692", "Accession_No": "120620130054", "Title": "Learning Python: Powerful Object-Oriented Programming", "Author": "Mark Lutz", "Publisher": "O'Reilly Media",
"Edition": "5", "Vear_of_Publication": "2013", "Category": "Python", "Total_number_of_pages": "1,648", "Price": "20.99" }\)
MriteResult(("Innserted": 1))
```

When a change is detected, the watch function returns an object with all details of the change. We read if the change was insert type and get the key (id) of the document added. Now, we can use the function collections.find_one({id}) to get the document that was added. Once we have the document, we use the same method we used above to add this new entry into the SQL table.

Wrapper Script

1. Prepare database

- First, we need to prepare the database so we can write and test our script.
- A python script is used to create two tables of Books and Readers and fill them with 5
 entries each. The Books table has the Accession_no as the primary key which makes the
 Accession_no in Readers a foreign key.

```
PS D:\IITJ\4th year\Sem VIII\SDE\Assignments\2> python prepareCassandra.py

Row(accession_no='090420160022', author='Charles R. Severance', isbn='9781530051120', publisher='Charles Severance', title='Python for Everybo
a Using Python 3')

Row(accession_no='010219990014', author='Michael Francis Banahan', isbn='9780201544336', publisher='Addison-Wesley Publishing Company', title=
uring the ANSI C Standard')
ming')

Row(accession_no='201120180005', author='Martin Fowler', isbn='9780134757704', publisher='Addison-Wesley Professional, 2018', title='Refactori
Design of Existing Code')

Row(accession_no='010820090011', author='Robert C. Martin', isbn='9780132350884', publisher='Prentice Hall, 2009', title='Clean Code: A Handbo
are Craftsmanship')

Row(reader_id='23645677', accession_no='120620130016', issue_date='13/04/2021', return_date='23/06/2021')

Row(reader_id='67345635', accession_no='1201120180005', issue_date='14/04/2021', return_date='01/12/2021')

Row(reader_id='12349870', accession_no='090420160022', issue_date='13/07/2021', return_date='11/10/2021')

Row(reader_id='56723456', accession_no='090420160022', issue_date='13/07/2021', return_date='11/10/2021')

Row(reader_id='56723456', accession_no='090420160022', issue_date='13/07/2021', return_date='11/10/2021')
```

2. Script

- The script keeps running in the background and keeps track of the number of entries in the table Books and also keeps a record of primary keys.
- It checks for count every five seconds and if the count decreases, it means a row has been deleted. It checks which row is deleted using the keys and deletes the corresponding row in Readers. Then, it updates our initial count and keys record with these new values. Now, it again starts checking for counts.
- This way referential integrity is maintained since the Readers will never have an entry which is not in Books.

```
PS D:\IITJ\4th year\Sem VIII\SDE\Assignments\2> python refint.py
Initial count - 5
Count now - 5
```

API

- 3. Language and Libraries used
 - Python
 - Flask
 - OpenCV

4. Server

- We are using Flask to create our simple API. So, we create a Flask application and create a route '/face_detect' with POST and GET methods.
- It takes files from clients and calls our face detection function (discussed below) to detect faces. Then, it returns the result to the client.
- We have also defined a simple HTML page which is displayed at the client side.
- Then, we run the server and keep it running.

```
PS D:\IITJ\4th year\Sem VIII\SDE\Assignments\2> python server.py

* Serving Flask app 'server' (lazy loading)

* Environment: production

WARNING: This is a development server. Do not use it in a production deployment.

Use a production WSGI server instead.

* Debug mode: on

* Restarting with stat

* Debugger is active!

* Debugger PIN: 491-177-929

* Running on all addresses.

WARNING: This is a development server. Do not use it in a production deployment.

* Running on http://192.168.29.135:5001/ (Press CTRL+C to quit)
```

5. Client

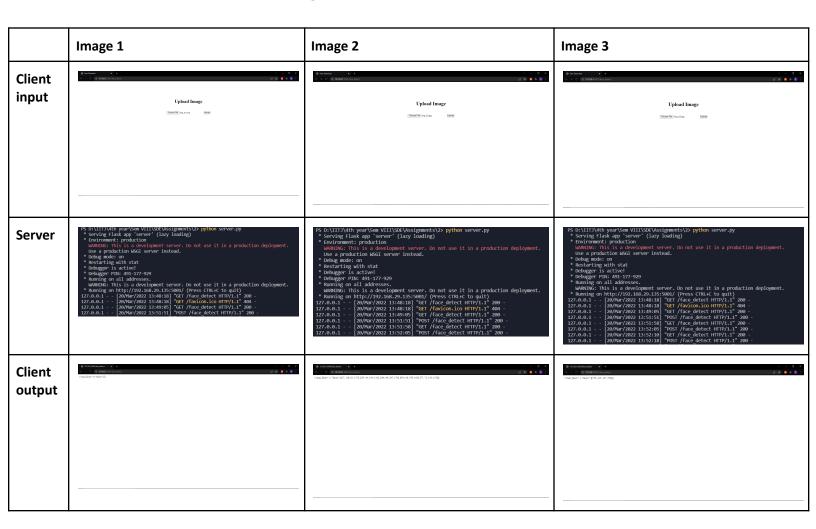
- We define a localhost for the client and once the server is running, the url will show our simple HTML page.
- Then, we can upload our image for detection. The client sends a request to the server.
- The client receives a response from the server which is then displayed in the browser window.



6. Face detection

Our face detection utility works using OpenCV. We use the Haar Cascade classifier to detect faces and return the number of faces detected and the location of those faces. This detector is not completely correct which is why it is only able to detect one face in image 3.

7. Results (zoom in to see images)



References

- MongoDB and SQL
 - https://www.youtube.com/watch?v=FwMwO8pXfq0
 - https://www.bogotobogo.com/DevOps/MongoDB/MongoDB_Replication_Replica_Set.p
 hp
 - https://onecompiler.com/questions/3vchw5r82/unable-to-use-mongodb-change-stream
 s-mongoerror-the-changestream-stage-is-only-supported-on-replica-sets
 - https://www.w3schools.com/python/python_mongodb_getstarted.asp

https://www.w3schools.com/python_mysql_getstarted.asp

Wrapper Script -

- https://www.youtube.com/watch?v=hJxlkHafYsQ
- https://docs.datastax.com/en/developer/python-driver/3.15/getting_started/
- https://cassandra.apache.org/doc/latest/cassandra/getting_started/drivers.html
- https://towardsdatascience.com/getting-started-with-apache-cassandra-and-python-81 e00ccf17c9

• Flask and OpenCV -

- https://flask.palletsprojects.com/en/2.0.x/quickstart/
- https://www.freecodecamp.org/news/how-to-build-a-web-application-using-flask-and-deploy-it-to-the-cloud-3551c985e492/
- https://blog.stoplight.io/python-rest-api
- o https://docs.opencv.org/3.4/db/d28/tutorial cascade classifier.html