BDA Lab-1 Report

Shaan Subbaiah B C - 1BM18CS096 - 6B - May 23, 2021

---

1. Cassandra commands

```
cqlsh> create keyspace students with replication = { 'class': 'SimpleStrategy',
'replication_factor': 1 };


cqlsh> describe keyspaces;

students  system_auth         system_schema  system_views
system    system_distributed  system_traces  system_virtual_schema


cqlsh> use students;
cqlsh:students> create table student_info( rollNo int primary key, name text, joinDate
timestamp, lastExamPerc double );


cqlsh:students> describe tables

student_info


cqlsh:students> describe table student
student_info  students.


cqlsh:students> describe table student_info

CREATE TABLE students.student_info (
    rollno int PRIMARY KEY,
    joindate timestamp,
    lastexamperc double,
    name text
) WITH additional_write_policy = '99p'
    AND bloom_filter_fp_chance = 0.01
    AND caching = {'keys': 'ALL', 'rows_per_partition': 'NONE'}
    AND cdc = false
    AND comment = ''
    AND compaction = {'class':
'org.apache.cassandra.db.compaction.SizeTieredCompactionStrategy', 'max_threshold':
'32', 'min_threshold': '4'}
    AND compression = {'chunk_length_in_kb': '16', 'class':
'org.apache.cassandra.io.compress.LZ4Compressor'}
```

```
   AND crc_check_chance = 1.0
   AND default_time_to_live = 0
   AND extensions = {}
   AND gc_grace_seconds = 864000
   AND max_index_interval = 2048
   AND memtable_flush_period_in_ms = 0
   AND min_index_interval = 128
   AND read_repair = 'BLOCKING'
   AND speculative_retry = '99p';


cqlsh:students> begin batch insert into student_info(rollno, joindate, lastexamperc,
name) values (1, '2021-05-23', 90.0, 'Adam') insert into student_info(rollno, joindate,
lastexamperc, name) values (2, '2021-05-22', 97.7, 'Eve') apply batch;


cqlsh:students> select * from student_info;

rollno | joindate                         | lastexamperc | name
--------+----------------------------------+--------------+------
     1 | 2021-05-22 18:30:00.000000+0000 |           90 | Adam
     2 | 2021-05-21 18:30:00.000000+0000 |         97.7 |  Eve

(2 rows)


cqlsh:students> update student_info set name = 'Micheal' where rollno = 1;


cqlsh:students> select * from student_info where rollno in (1,2);

rollno | joindate                         | lastexamperc | name
--------+----------------------------------+--------------+---------
     1 | 2021-05-22 18:30:00.000000+0000 |           90 | Micheal
     2 | 2021-05-21 18:30:00.000000+0000 |         97.7 |     Eve

(2 rows)


cqlsh:students> create index on student_info(lastexamperc);


cqlsh:students> select rollno, name from student_info limit 2;

rollno | name
--------+---------
     1 | Micheal
     2 |     Eve
```

```
(2 rows)


cqlsh:students> create index on student_info(name);


cqlsh:students> update student_info set name='Eve2', lastexamperc=100.0 where rollno=2;


cqlsh:students> select * from student_info;

 rollno | joindate                         | lastexamperc | name
--------+----------------------------------+--------------+---------
      1 | 2021-05-22 18:30:00.000000+0000 |           90 | Micheal
      2 | 2021-05-21 18:30:00.000000+0000 |          100 |    Eve2

(2 rows)


cqlsh:students> delete lastexamperc from student_info where rollno=2;


cqlsh:students> select * from student_info;

 rollno | joindate                         | lastexamperc | name
--------+----------------------------------+--------------+---------
      1 | 2021-05-22 18:30:00.000000+0000 |           90 | Micheal
      2 | 2021-05-21 18:30:00.000000+0000 |         null |    Eve2

(2 rows)


cqlsh:students> delete from student_info where rollno=2;


cqlsh:students> select * from student_info;

 rollno | joindate                         | lastexamperc | name
--------+----------------------------------+--------------+---------
      1 | 2021-05-22 18:30:00.000000+0000 |           90 | Micheal

(1 rows)
```

2. Cassandra Employee Database

```
cqlsh> create keyspace employee_info with
replication={'class':'SimpleStrategy','replication_factor':1};


cqlsh> use employee_info;
cqlsh:employee_info> create table employee_details(emp_id int, emp_name text,
designation text, doj timestamp, salary double, dept_name text, primary
key(emp_id,salary));


cqlsh:employee_info> describe table employee_details;

CREATE TABLE employee_info.employee_details (
    emp_id int,
    salary double,
    dept_name text,
    designation text,
    doj timestamp,
    emp_name text,
    PRIMARY KEY (emp_id, salary)
) WITH CLUSTERING ORDER BY (salary ASC)
    AND additional_write_policy = '99p'
    AND bloom_filter_fp_chance = 0.01
    AND caching = {'keys': 'ALL', 'rows_per_partition': 'NONE'}
    AND cdc = false
    AND comment = ''
    AND compaction = {'class':
'org.apache.cassandra.db.compaction.SizeTieredCompactionStrategy', 'max_threshold':
'32', 'min_threshold': '4'}
    AND compression = {'chunk_length_in_kb': '16', 'class':
'org.apache.cassandra.io.compress.LZ4Compressor'}
    AND crc_check_chance = 1.0
    AND default_time_to_live = 0
    AND extensions = {}
    AND gc_grace_seconds = 864000
    AND max_index_interval = 2048
    AND memtable_flush_period_in_ms = 0
    AND min_index_interval = 128
    AND read_repair = 'BLOCKING'
    AND speculative_retry = '99p';


cqlsh:employee_info> begin batch insert into
employee_details(emp_id,emp_name,designation,doj,salary,dept_name) values
(100,'tanya','manager','2020-09-11',30000,'testing') insert into
employee_details(emp_id,emp_name,designation,doj,salary,dept_name) values
(111,'sriram','associate','2020-06-11',25000,'development') insert into
```

```
employee_details(emp_id,emp_name,designation,doj,salary,dept_name) values
(121,'shiva','manager','2020-01-03',35000,'hr') apply batch;


cqlsh:employee_info> select * from employee_details;

emp_id | salary | dept_name   | designation | doj                             |
emp_name
--------+--------+-------------+-------------+---------------------------------+-------
---
   111 |  25000 | development |   associate | 2020-06-10 18:30:00.000000+0000 |
sriram
   121 |  35000 |          hr |     manager | 2020-01-02 18:30:00.000000+0000 |
shiva
   100 |  30000 |     testing |     manager | 2020-09-10 18:30:00.000000+0000 |
tanya

(3 rows)


cqlsh:employee_info> update employee_details set emp_name='shaan' where emp_id = 121
and salary=35000;


cqlsh:employee_info> select * from employee_details;

emp_id | salary | dept_name   | designation | doj                             |
emp_name
--------+--------+-------------+-------------+---------------------------------+-------
---
   111 |  25000 | development |   associate | 2020-06-10 18:30:00.000000+0000 |
sriram
   121 |  35000 |          hr |     manager | 2020-01-02 18:30:00.000000+0000 |
shaan
   100 |  30000 |     testing |     manager | 2020-09-10 18:30:00.000000+0000 |
tanya

(3 rows)


cqlsh:employee_info> alter table employee_details add project text;


cqlsh:employee_info> update employee_details set project='chat app' where emp_id=111
and salary=25000;
cqlsh:employee_info> update employee_details set project='campusx' where emp_id=121 and
salary=35000;
cqlsh:employee_info> update employee_details set project='canteen app' where emp_id=100
and salary=30000;
```

```
cqlsh:employee_info> select * from employee_details;

 emp_id | salary | dept_name   | designation | doj                             |
emp_name | project
--------+--------+-------------+-------------+---------------------------------+-------
---+-------------
    111 |  25000 | development |   associate | 2020-06-10 18:30:00.000000+0000 |
sriram |    chat app
    121 |  35000 |          hr |     manager | 2020-01-02 18:30:00.000000+0000 |
shaan |     campusx
    100 |  30000 |     testing |     manager | 2020-09-10 18:30:00.000000+0000 |
tanya | canteen app

(3 rows)


cqlsh:employee_info> insert into
employee_details(emp_id,emp_name,designation,doj,salary,dept_name)
values(113,'sam','manager','2020-09-09',30000,'testing') using ttl 30;


cqlsh:employee_info> select ttl(emp_name) from employee_details where emp_id=113 and
salary=30000;

 ttl(emp_name)
---------------
            22

(1 rows)


cqlsh:employee_info> paging off;
Disabled Query paging.


cqlsh:employee_info>  select * from employee_details where emp_id in (111,121,100)
order by salary;

 emp_id | salary | dept_name   | designation | doj                             |
emp_name | project
--------+--------+-------------+-------------+---------------------------------+-------
---+-------------
    111 |  25000 | development |   associate | 2020-06-10 18:30:00.000000+0000 |
sriram |    chat app
    100 |  30000 |     testing |     manager | 2020-09-10 18:30:00.000000+0000 |
tanya | canteen app
```

```
   121 |  35000 |          hr |     manager | 2020-01-02 18:30:00.000000+0000 |
shaan |     campusx

(3 rows)
```

3. Cassandra Library Database

```
cqlsh> create keyspace library_info with replication =
{'class':'SimpleStrategy','replication_factor':1};


cqlsh> use library_info;


cqlsh:library_info> create table library_details(stud_id int,counter_value
               ... counter,stud_name text,book_name text,date_of_issue
timestamp,book_id
               ... int,primary key(stud_id,stud_name,book_name,date_of_issue,book_id));


cqlsh:library_info> update library_details set counter_value=counter_value+1
               ... where stud_id=111 and stud_name='sam' and book_name='ML' and
               ... date_of_issue='2020-11-09' and book_id=200;
cqlsh:library_info> update library_details set counter_value=counter_value+1
               ... where stud_id=112 and stud_name='shaan' and book_name='BDA' and
               ... date_of_issue='2020-01-01' and book_id=300;
cqlsh:library_info> update library_details set counter_value=counter_value+1
               ... where stud_id=113 and stud_name='ayman' and book_name='OOMD' and
               ... date_of_issue='2020-06-01' and book_id=400;


cqlsh:library_info>  select * from library_details;

stud_id | stud_name | book_name | date_of_issue                   | book_id |
counter_value
---------+-----------+-----------+---------------------------------+---------+---------
------
     111 |       sam |        ML | 2020-11-08 18:30:00.000000+0000 |     200 |
1
     113 |     ayman |      OOMD | 2020-05-31 18:30:00.000000+0000 |     400 |
1
     112 |     shaan |       BDA | 2019-12-31 18:30:00.000000+0000 |     300 |
1

(3 rows)


cqlsh:library_info> update library_details set counter_value=counter_value+1
               ... where stud_id=112 and stud_name='shaan' and book_name='BDA' and
               ... date_of_issue='2020-01-01' and book_id=300;

cqlsh:library_info>  select * from library_details where stud_id=112;

stud_id | stud_name | book_name | date_of_issue                   | book_id |
counter_value
```

```
---------+----------+----------+---------------------------------+---------+---------
------
    112 |     shaan |       BDA | 2019-12-31 18:30:00.000000+0000 |     300 |
2

(1 rows)


cqlsh:library_info>  copy
library_details(stud_id,stud_name,book_name,book_id,date_of_issue,counter_value) to
'library.csv' ;
Using 1 child processes

Starting copy of library_info.library_details with columns [stud_id, stud_name,
book_name, book_id, date_of_issue, counter_value].
cqlshlib.copyutil.ExportProcess.write_rows_to_csv(): writing row
cqlshlib.copyutil.ExportProcess.write_rows_to_csv(): writing row
cqlshlib.copyutil.ExportProcess.write_rows_to_csv(): writing row
Processed: 3 rows; Rate:      32 rows/s; Avg. rate:      32 rows/s
3 rows exported to 1 files in 0.123 seconds.


cqlsh:library_info> truncate library_details;


cqlsh:library_info>  copy
library_details(stud_id,stud_name,book_name,book_id,date_of_issue,counter_value) from
'library.csv' ;
Using 1 child processes

Starting copy of library_info.library_details with columns [stud_id, stud_name,
book_name, book_id, date_of_issue, counter_value].
Processed: 3 rows; Rate:       5 rows/s; Avg. rate:       8 rows/s
3 rows imported from 1 files in 0.394 seconds (0 skipped).


cqlsh:library_info> select * from library_details;

stud_id | stud_name | book_name | date_of_issue                   | book_id |
counter_value
---------+----------+----------+---------------------------------+---------+---------
------
    111 |       sam |        ML | 2020-11-08 18:30:00.000000+0000 |     200 |
1
    113 |     ayman |      OOMD | 2020-05-31 18:30:00.000000+0000 |     400 |
1
    112 |     shaan |       BDA | 2019-12-31 18:30:00.000000+0000 |     300 |
2
```

4. MongoDB Commands

1. Create a new collection
```
use Student
```

2. Insert a value
```json
db.Student.insert({
   "Name" : "XYZ",
   "RollNo:" : 1,
   "Age" : 21,
   "ContactNo" : "1234567890",
   "EmailId": "user1@lab.com"
})
```

3. Insert multiple values at once
```json
var MyStudents = [
   {
       "Name" : "ABC",
       "RollNo:" : 3,
       "Age" : 22,
       "ContactNo" : "2234567890",
       "EmailId": "user2@lab.com"
   },
   {
       "Name" : "DEF",
       "RollNo:" : 5,
       "Age" : 21,
       "ContactNo" : "3234567890",
       "EmailId" : "user3@lab.com"
   },
   {
       "Name" : "GHI",
       "RollNo:" : 7,
       "Age" : 20,
       "ContactNo" : "4234567890",
       "EmailId" : "user4@lab.com"
   },
   {
       "Name" : "JKL",
       "RollNo:" : 10,
       "Age" : 18,
       "ContactNo" : "5234567890",
       "EmailId" : "user5@lab.com"
   },
```

```
]

db.Student.insert(MyStudents);
```

4. Print all current values
```json
db.getCollection('Student').find({}).forEach(printjson)
```
```json
{
    "_id" : ObjectId("606ad5a6e581cc0b904470a5"),
    "Name" : "XYZ",
    "RollNo:" : 1,
    "Age" : 21,
    "ContactNo" : "1234567890",
    "EmailId" : "user1@lab.com"
}
{
    "_id" : ObjectId("606ad60fe581cc0b904470a6"),
    "Name" : "ABC",
    "RollNo:" : 3,
    "Age" : 22,
    "ContactNo" : "2234567890",
    "EmailId" : "user2@lab.com"
}
{
    "_id" : ObjectId("606ad60fe581cc0b904470a7"),
    "Name" : "DEF",
    "RollNo:" : 5,
    "Age" : 21,
    "ContactNo" : "3234567890",
    "EmailId" : "user3@lab.com"
}
{
    "_id" : ObjectId("606ad60fe581cc0b904470a8"),
    "Name" : "GHI",
    "RollNo:" : 7,
    "Age" : 20,
    "ContactNo" : "4234567890",
    "EmailId" : "user4@lab.com"
}
{
    "_id" : ObjectId("606ad60fe581cc0b904470a9"),
    "Name" : "JKL",
    "RollNo:" : 10,
    "Age" : 18,
    "ContactNo" : "5234567890",
    "EmailId" : "user5@lab.com"
```

```
}
```
5. Update RollNo of a student
```json
db.Student.update(
{"RollNo:" : 10},
{$set: { "EmailId" : "modified@lab.com"}});
```
```json
db.getCollection('Student').find({"RollNo:":10}).forEach(printjson)
```
```json
{
    "_id" : ObjectId("606ad60fe581cc0b904470a9"),
    "Name" : "JKL",
    "RollNo:" : 10,
    "Age" : 18,
    "ContactNo" : "5234567890",
    "EmailId" : "modified@lab.com"
}
```
6. Update Name of a student
```json
db.Student.update(
{"Name" : "XYZ"},
{$set: { "Name" : "EcksWhyZee"}});
```
```json
db.getCollection('Student').find({"Name" : "EcksWhyZee"}).forEach(printjson)
```
```json
{
    "_id" : ObjectId("606ad5a6e581cc0b904470a5"),
    "Name" : "EcksWhyZee",
    "RollNo:" : 1,
    "Age" : 21,
    "ContactNo" : "1234567890",
    "EmailId" : "user1@lab.com"
}
```
7. Export to json
```
mongoexport --db testdb --collection Student --out
C:\Users\shaan\Desktop\Exported\Student.json
```
```json
{"_id":{"$oid":"606ad5a6e581cc0b904470a5"},"Name":"EcksWhyZee","RollNo:":1.0,"Age":21.0
,"ContactNo":"1234567890","EmailId":"user1@lab.com"}
```

{"_id":{"$oid":"606ad60fe581cc0b904470a6"},"Name":"ABC","RollNo:":3.0,"Age":22.0,"Conta
ctNo":"2234567890","EmailId":"user2@lab.com"}

{"_id":{"$oid":"606ad60fe581cc0b904470a7"},"Name":"DEF","RollNo:":5.0,"Age":21.0,"Conta
ctNo":"3234567890","EmailId":"user3@lab.com"}

{"_id":{"$oid":"606ad60fe581cc0b904470a8"},"Name":"GHI","RollNo:":7.0,"Age":20.0,"Conta
ctNo":"4234567890","EmailId":"user4@lab.com"}

{"_id":{"$oid":"606ad60fe581cc0b904470a9"},"Name":"JKL","RollNo:":10.0,"Age":18.0,"Cont
actNo":"5234567890","EmailId":"modified@lab.com"}

8. Drop Student

```json
db.getCollection('Student').drop()
```

9. Import from exported file

```
mongoimport --db testdb --collection Student
C:\Users\shaan\Desktop\Exported\Student.json
```
