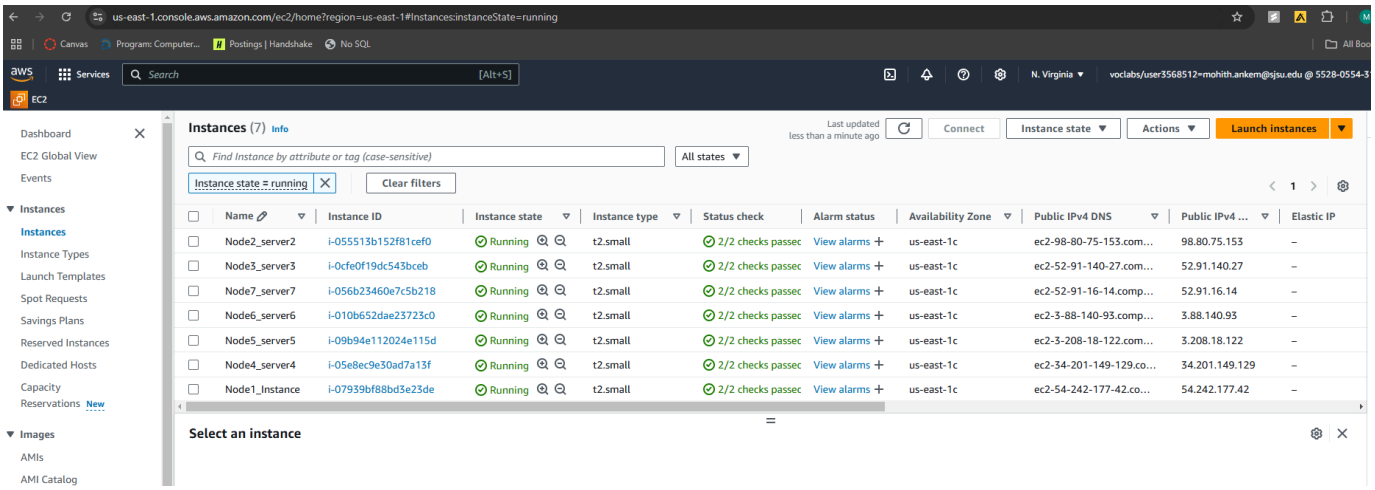
1. Set up nodes in AWS. Determine the number of nodes based on your deployment plan.

Ans:There are a total of 7 instances; 3 for config servers, 1 for mongos router and the other 3 for sharding.

Config Servers : Node1, Node2 and Node3

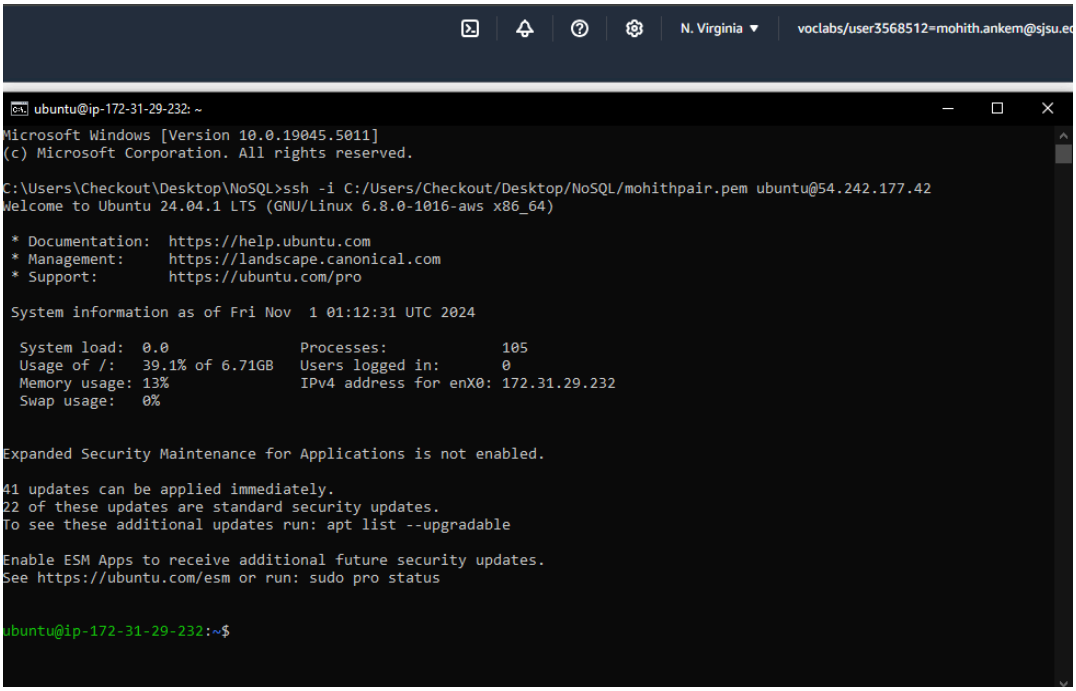
Mongos Server: Node4

Shard Servers: Node5, Node6 and Node7

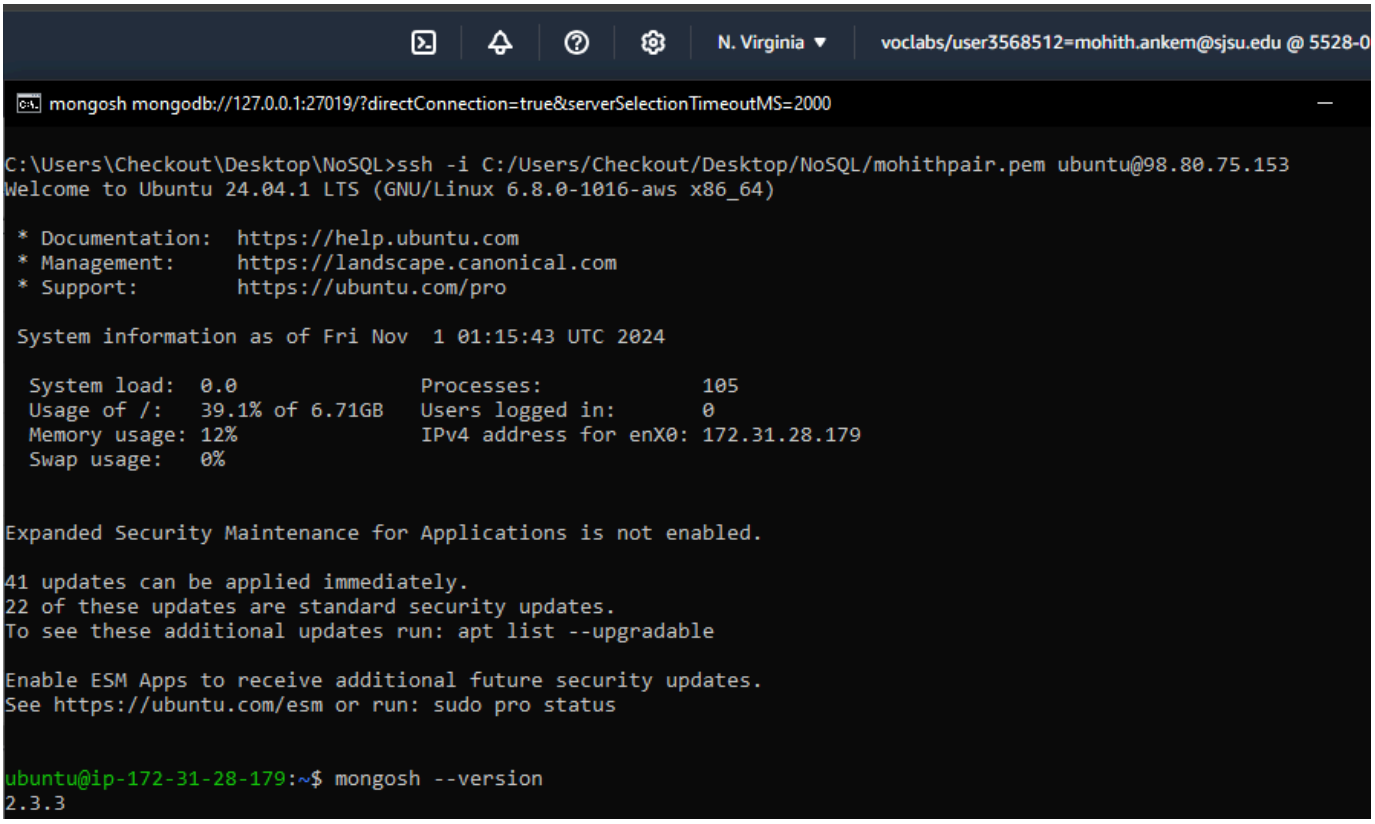


2. Access these instances (nodes) through SSH

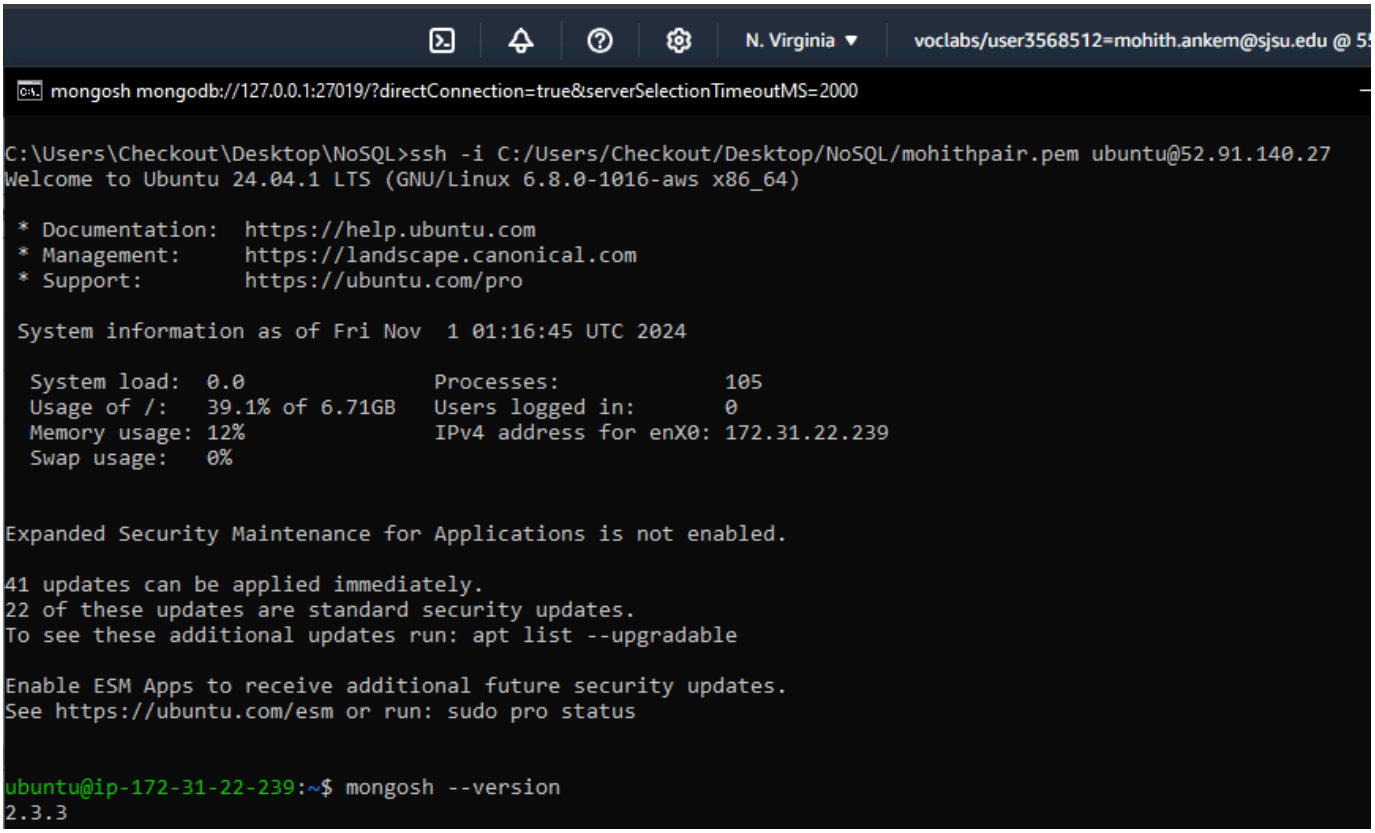
Instance 1:



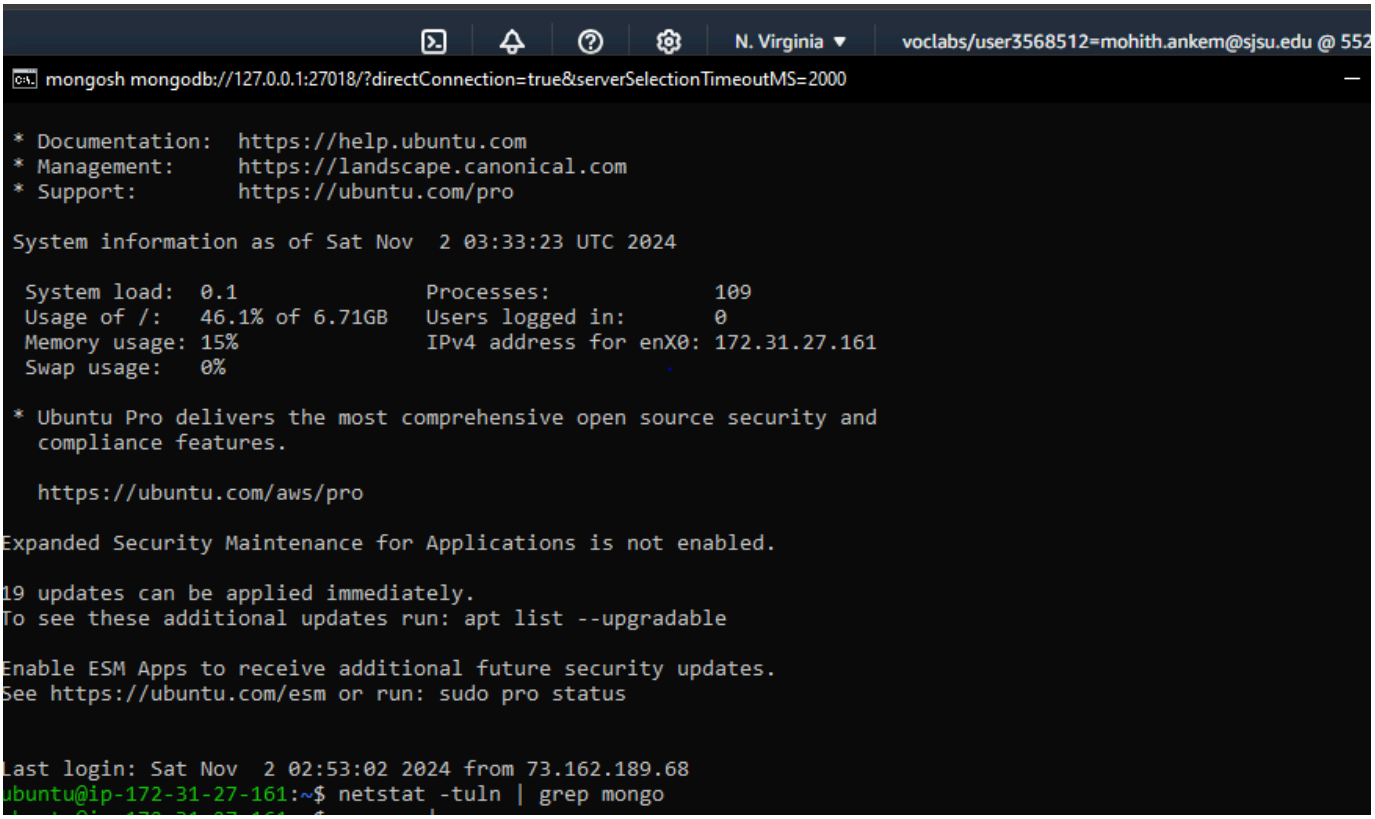
Instance 2:



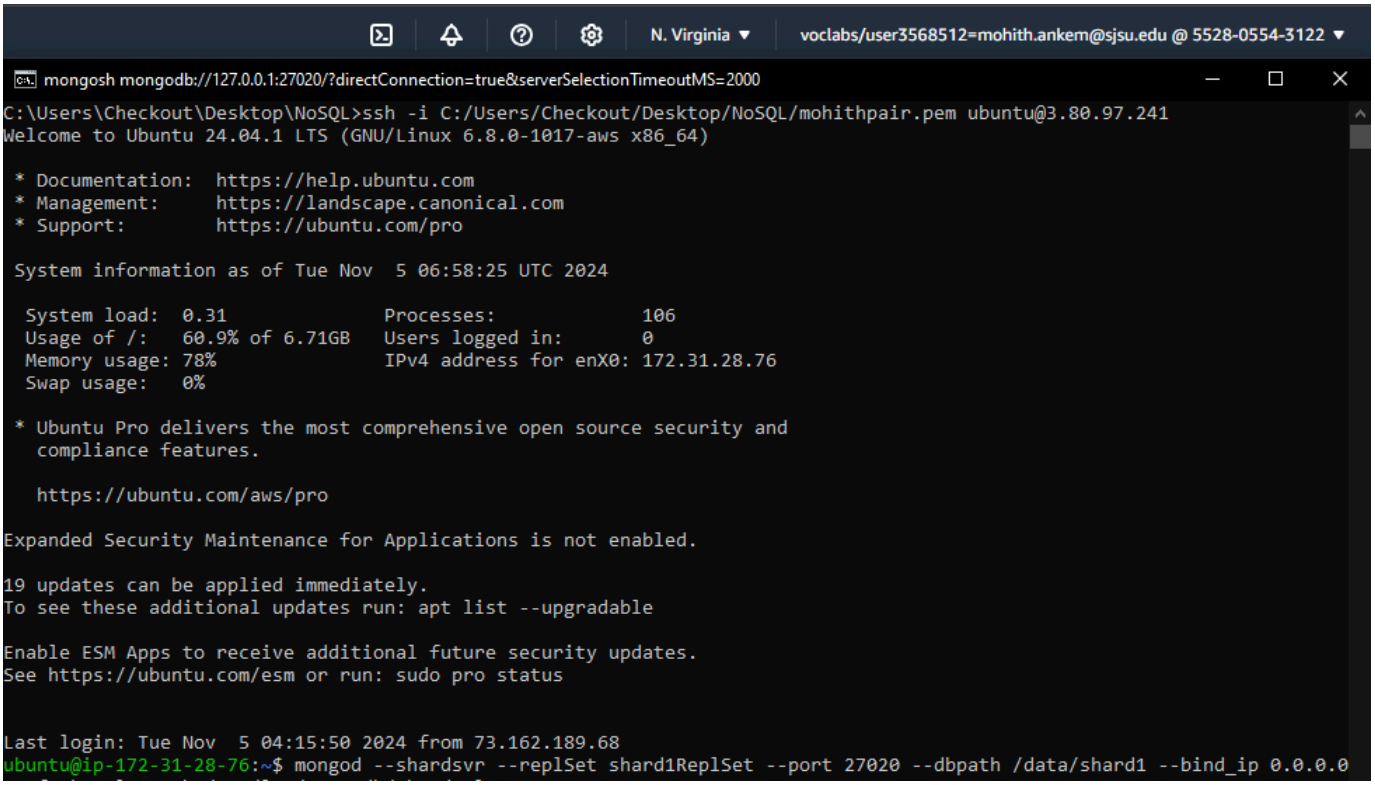
Instance 3:



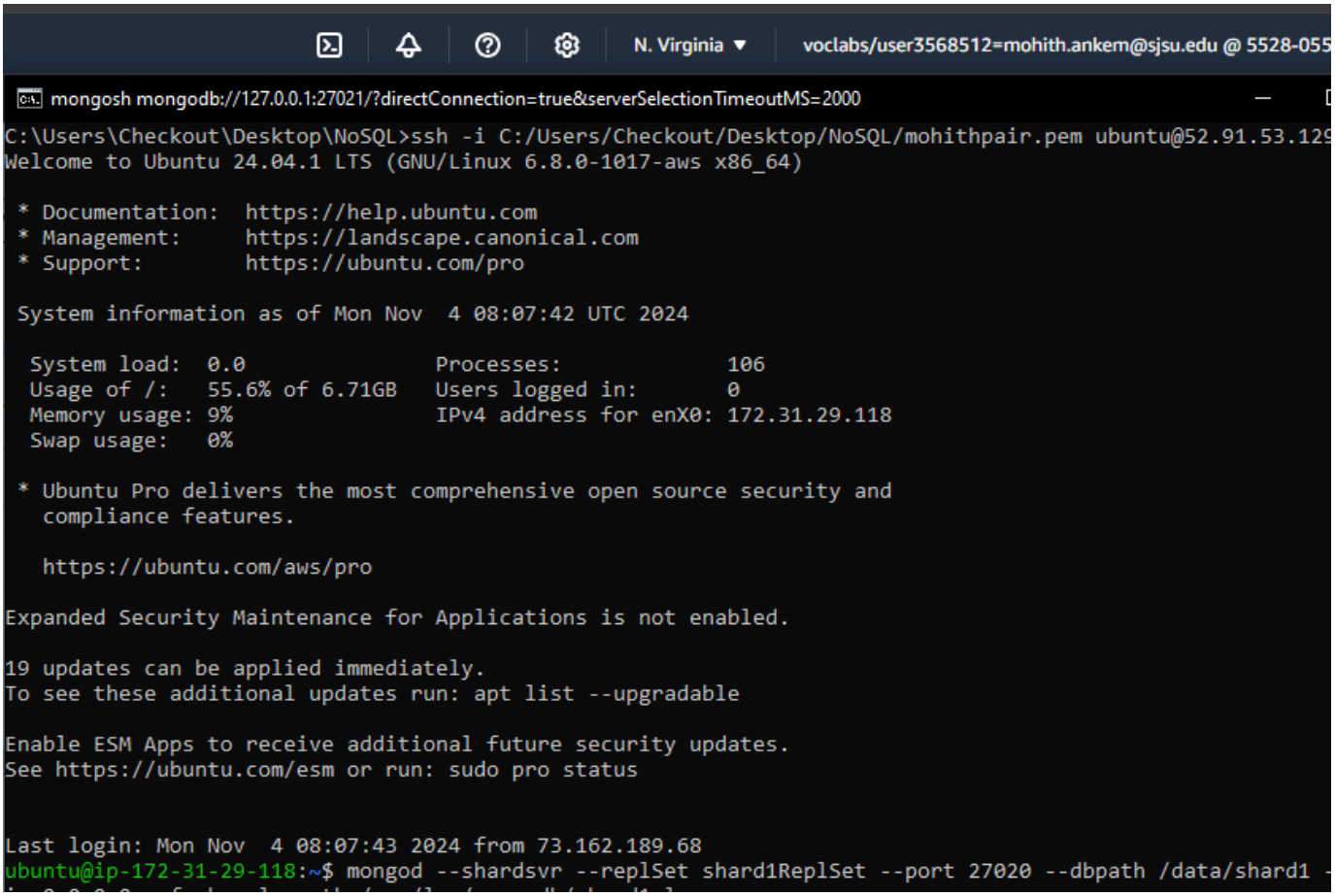
Instance 4:



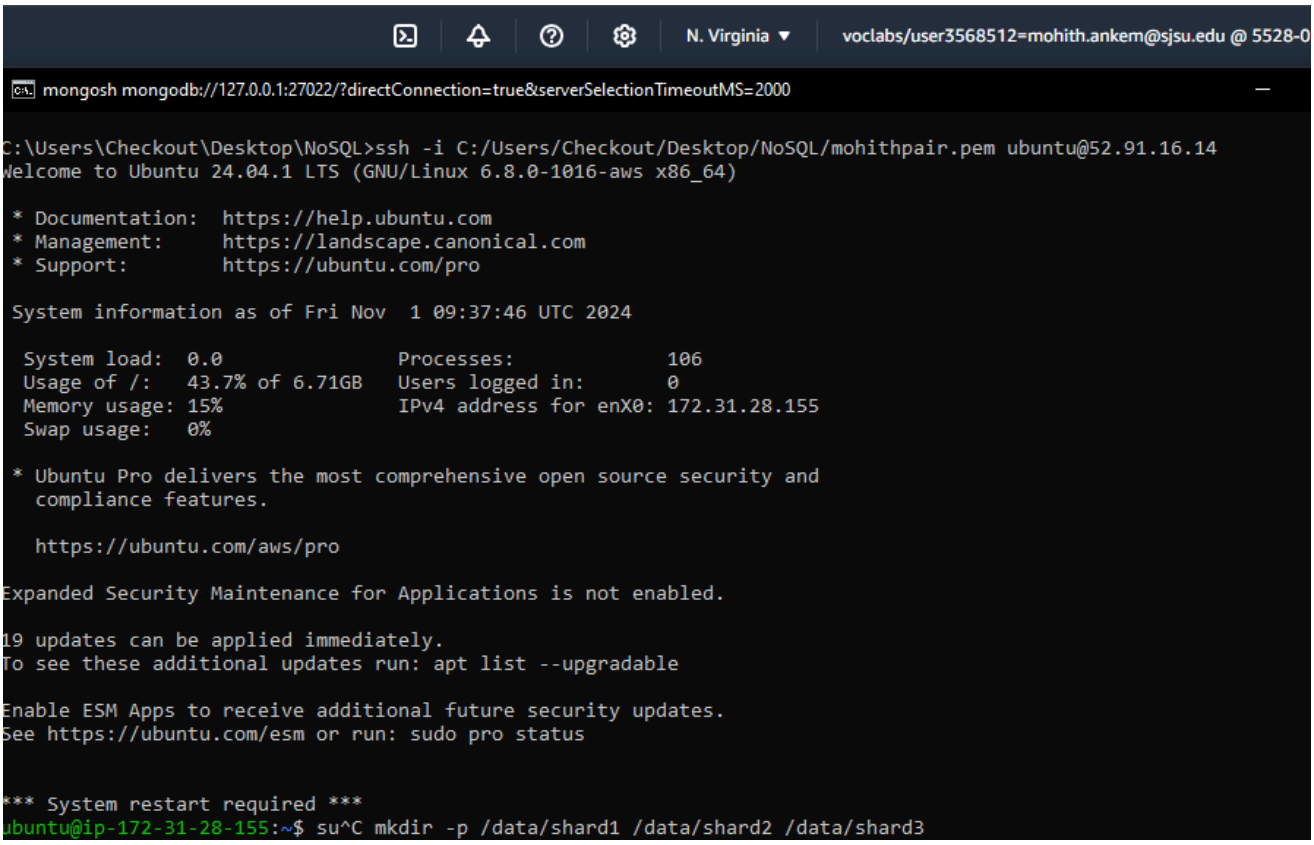
Instance 5:



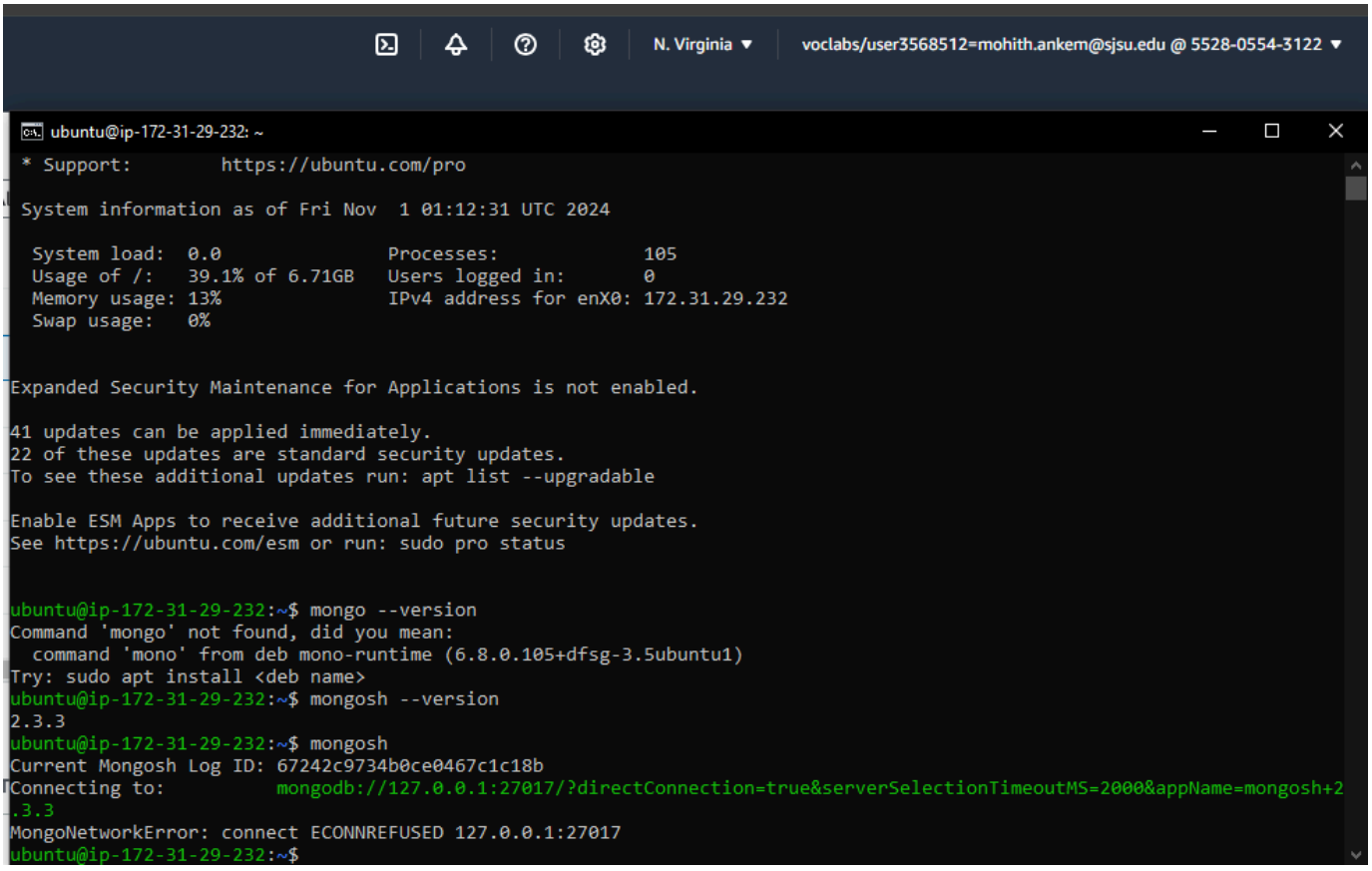
Instance 6:

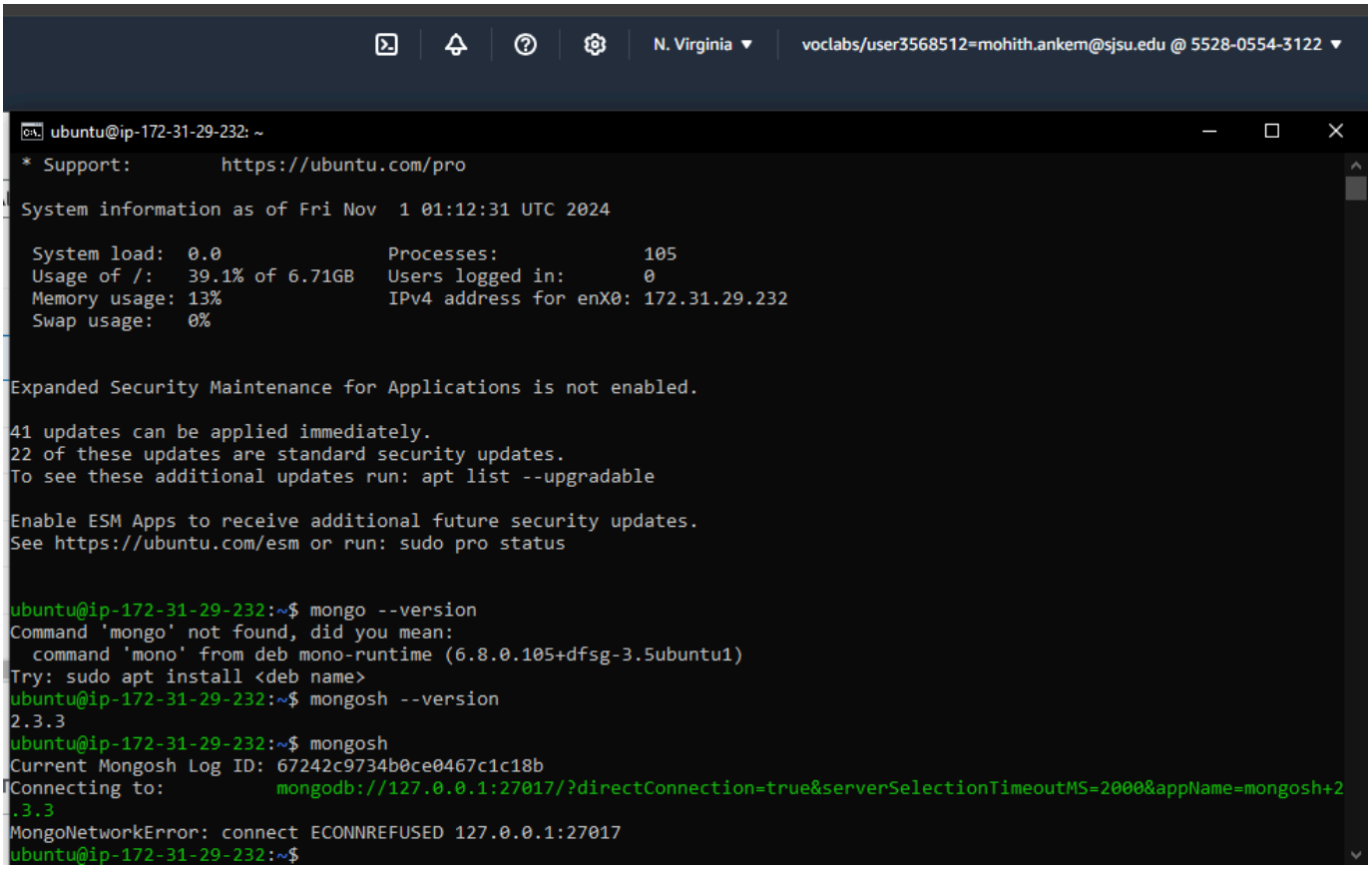


Instance 7:



3. Install MongoDB in each node (i.e. instance)





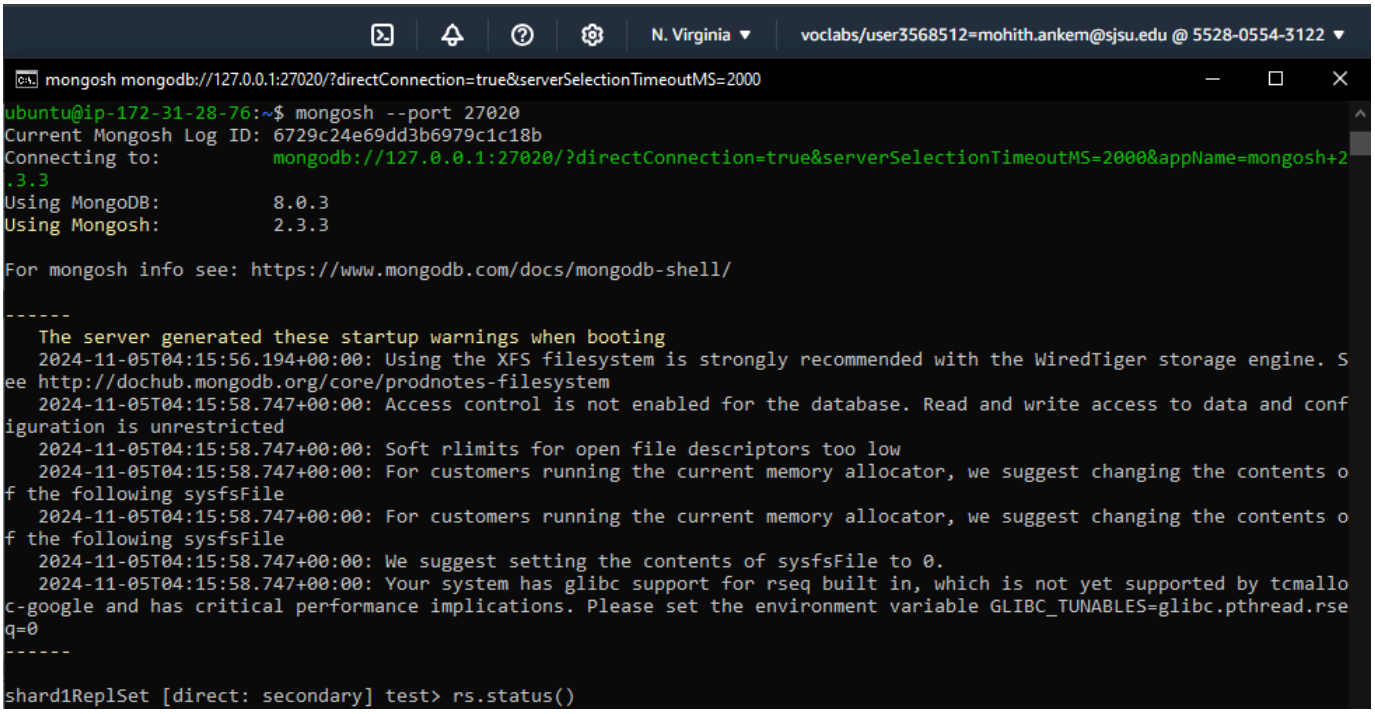
The Above two are proof that mongo is installed in the respective server.

This is done using the install_mongo.bat script shown below:

```
@echo off
setlocal

:: Set the path to your private key
set "KEY_PATH=C:\Users\Checkout\Desktop\NoSQL\mohithpair.pem"

:: List of AWS instance IP addresses
set "instances=98.80.75.153 52.91.140.27 52.91.16.14 3.88.140.93 3.208.18.122 34.201.149.129 54.242.177.42"

:: MongoDB installation commands
set "install_commands=curl -fsSL https://www.mongodb.org/static/pgp/server-8.0.asc | sudo gpg -o /usr/share/keyrings/mongod

:: Loop through each instance and install MongoDB
for %%i in (%instances%) do (
    echo Connecting to %%i
    ssh -i "%KEY_PATH%" -o "StrictHostKeyChecking=no" ubuntu@%%i "%install_commands%"
    echo Finished installing MongoDB on %%i
)

endlocal
echo All done!
pause
```

Using the respective port numbers all the servers are connected to mongos using the command

Mongosh –port 27019

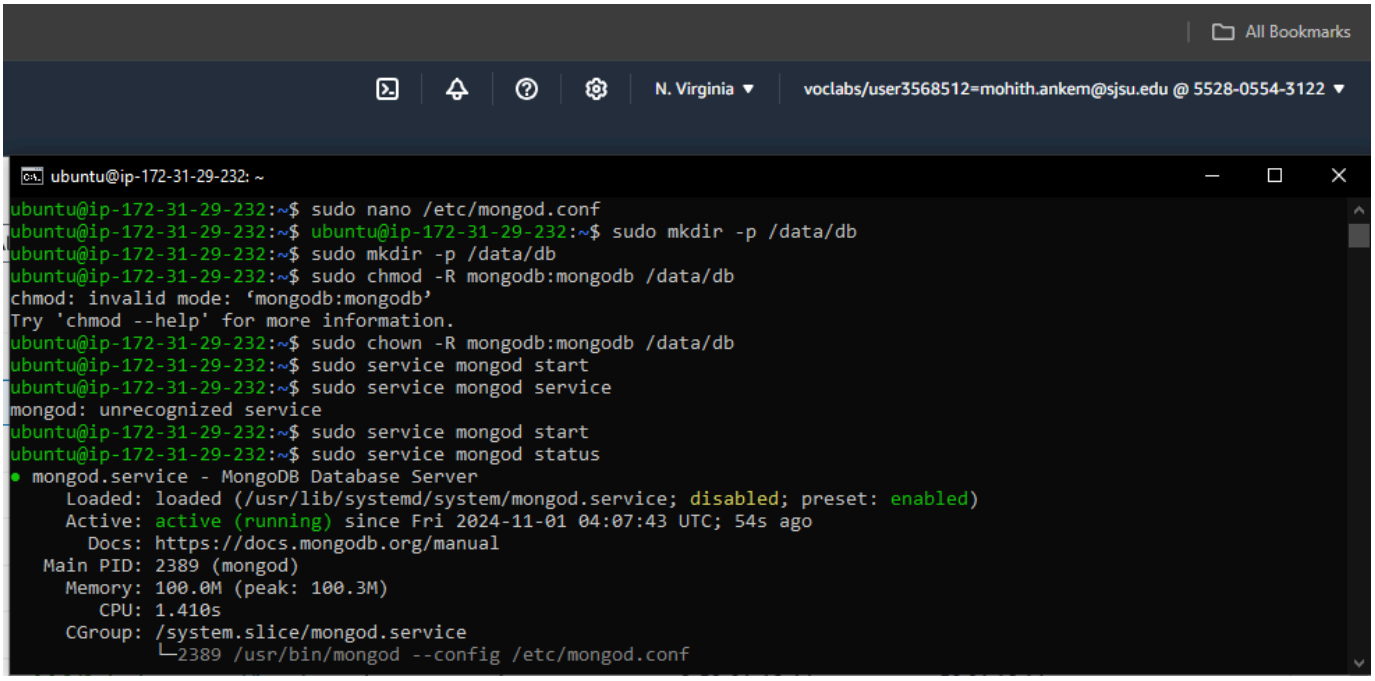


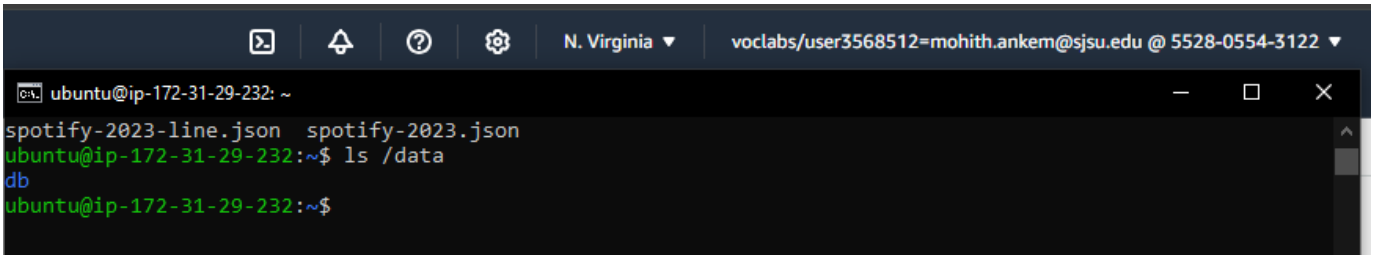
4. Create a directory to store database in each node

The Commands that are used to make directory in the servers to store database in each node are as follows:

- Sudo mkdir -p /data/db
- Sudo chown -R mongodb:mongodb /data/db

The Screenshot is attached below to make the directories which are done in all servers.



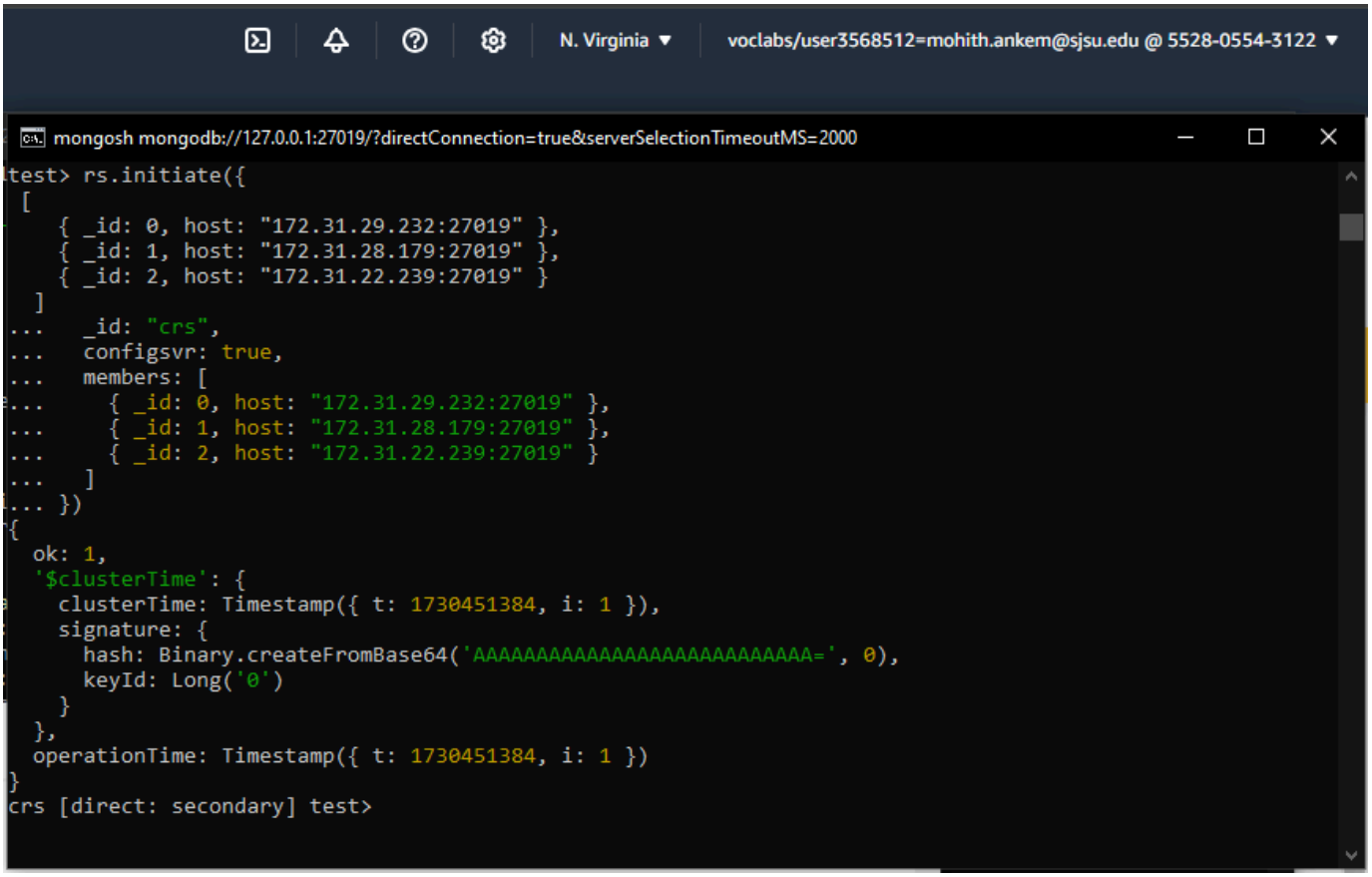


5. Specify Public and Private IP Addresses of AWS instances used in your solution.

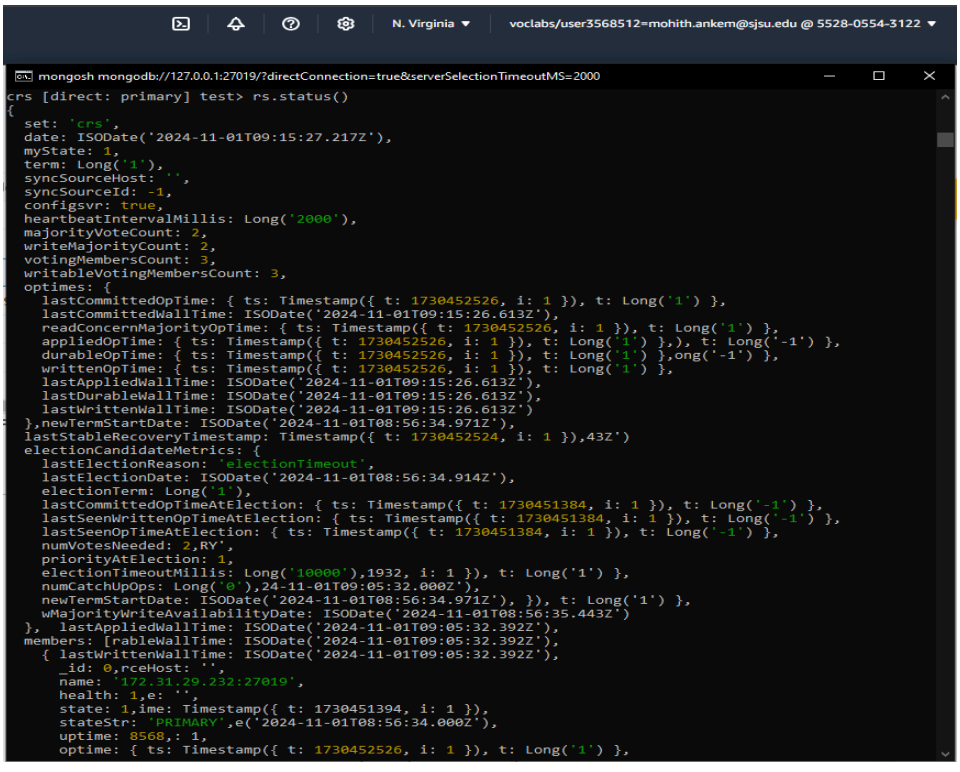
Instance Name	Public IP	Private IP
Node1_Instance	54.197.4.176	172.31.29.232
Node2_server2	3.87.107.88	172.31.28.179
Node3_server3	18.212.98.236	172.31.22.239
Node4_server4	23.22.61.21	172.31.27.161
Node5_server5	3.80.97.241	172.31.28.76
Node6_server6	3.95.243.227	172.31.29.118
Node7_server7	54.226.141.42	172.31.28.155

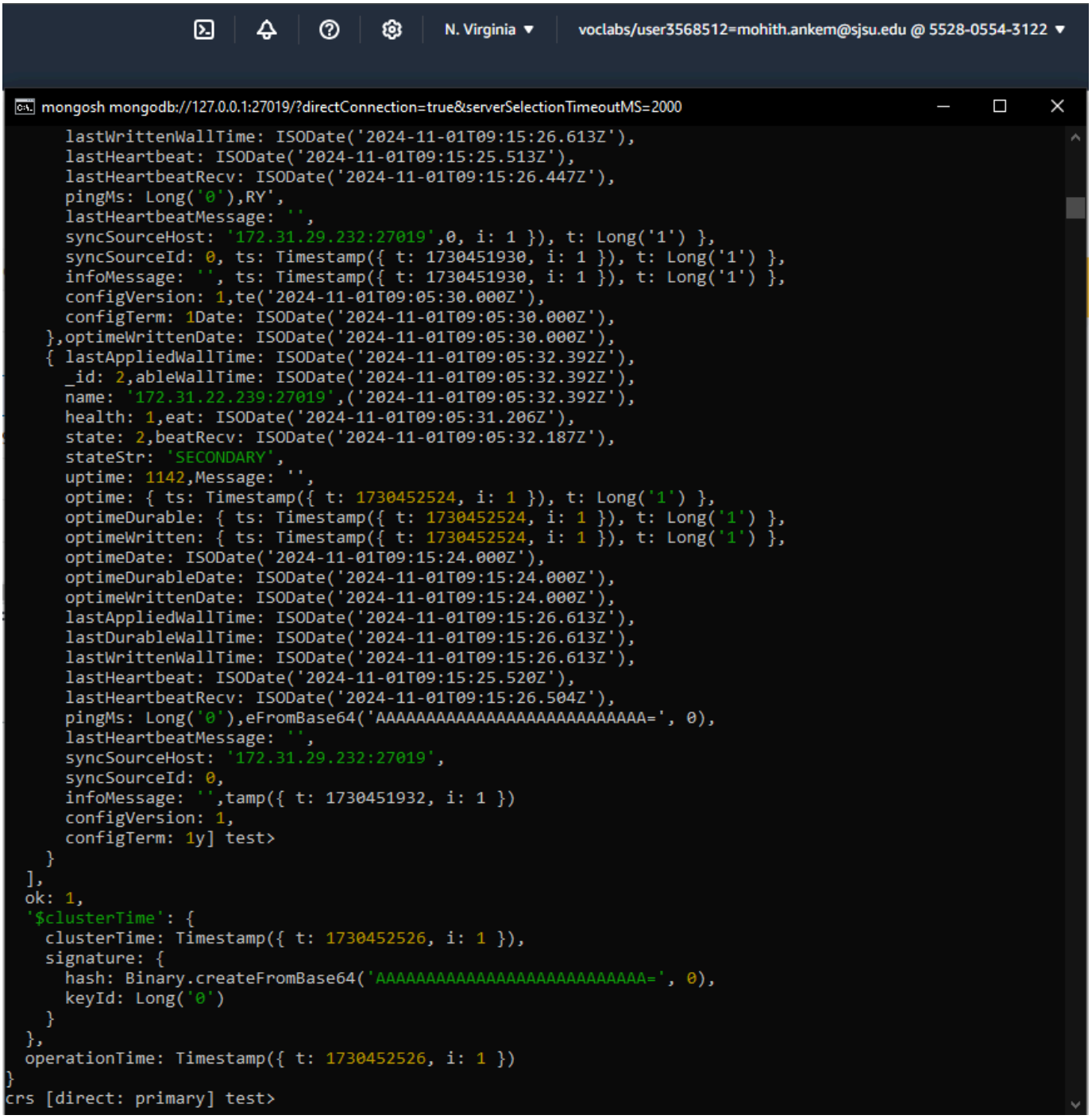
6. Set up and launch three config servers in a replica set.

I have used the rs.initiate() command to initialize the replica set for the config servers.



rs.status():

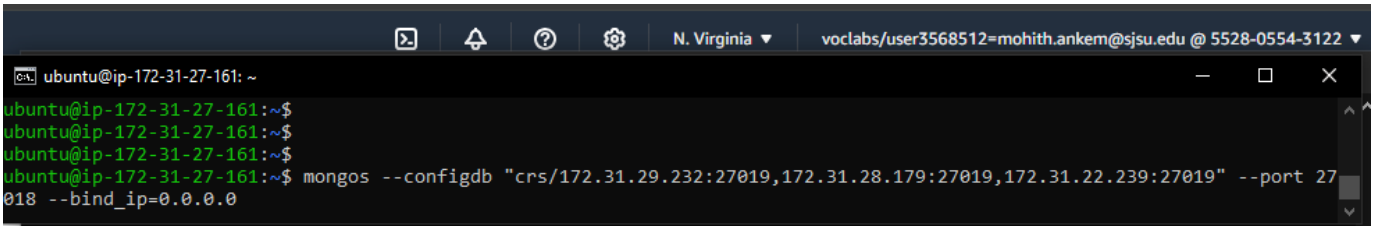




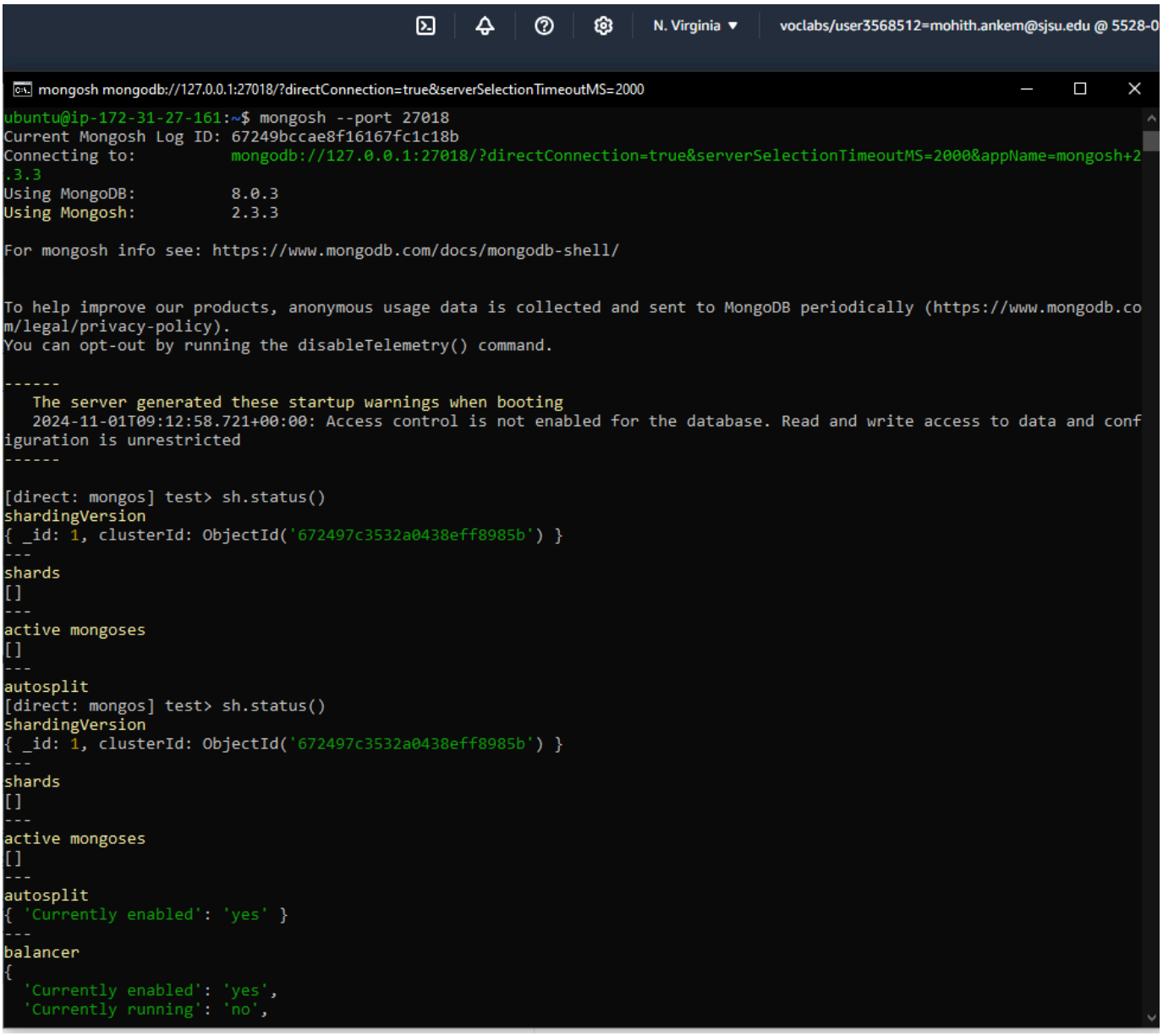
7. Connect mongos to each config server.

The command to run mongos server is: **mongos –config /etc/mongos.conf**

Using this command I modified the sharding configuration which will setup mongos server to each config server.



Using the above command in the command prompt , I kept the server running and opened another terminal to access the monos server and sh.status() after executing the command is attached below.



8. Set up and launch each of the three shards. Make sure to include the result of sh.status() before adding the shards, which will be done in the next task.

On all of the shard servers we run the following commands:

sudo mkdir -p /data/shard1 /data/shard2 /data/shard3

sudo chown ubuntu:ubuntu /data/shard1 /data/shard2 /data/shard3 /var/log/mongodb

mongod --shardsvr --replSet shard1ReplSet --port 27020 --dbpath /data/shard1 --bind_ip 0.0.0.0 --fork --logpath /var/log/mongodb/shard1.log

mongod --shardsvr --replSet shard2ReplSet --port 27021 --dbpath /data/shard2 --bind_ip 0.0.0.0 --fork --logpath /var/log/mongodb/shard2.log

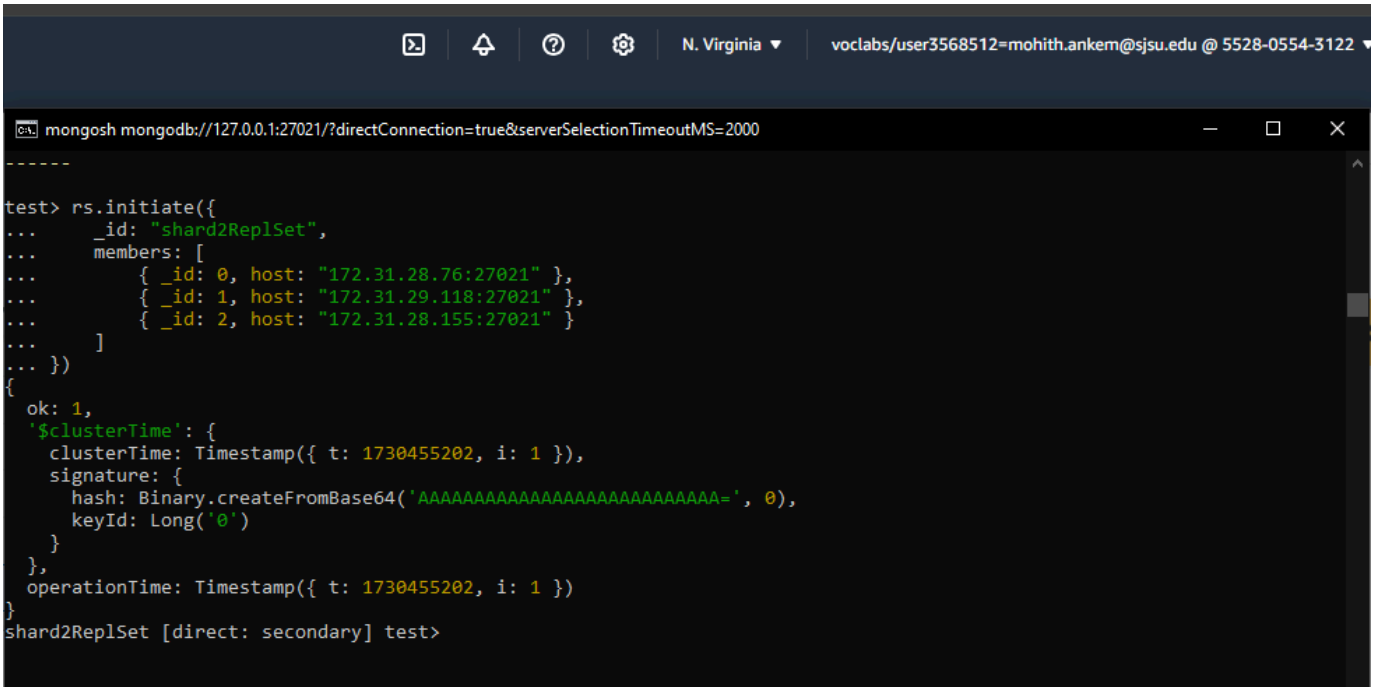
mongod --shardsvr --replSet shard3ReplSet --port 27022 --dbpath /data/shard3 --bind_ip 0.0.0.0 --fork --logpath /var/log/mongodb/shard3.log

After running these 3 commands on all three shard servers replica sets are created for each shard.Then we connect to mongosh to see the status.

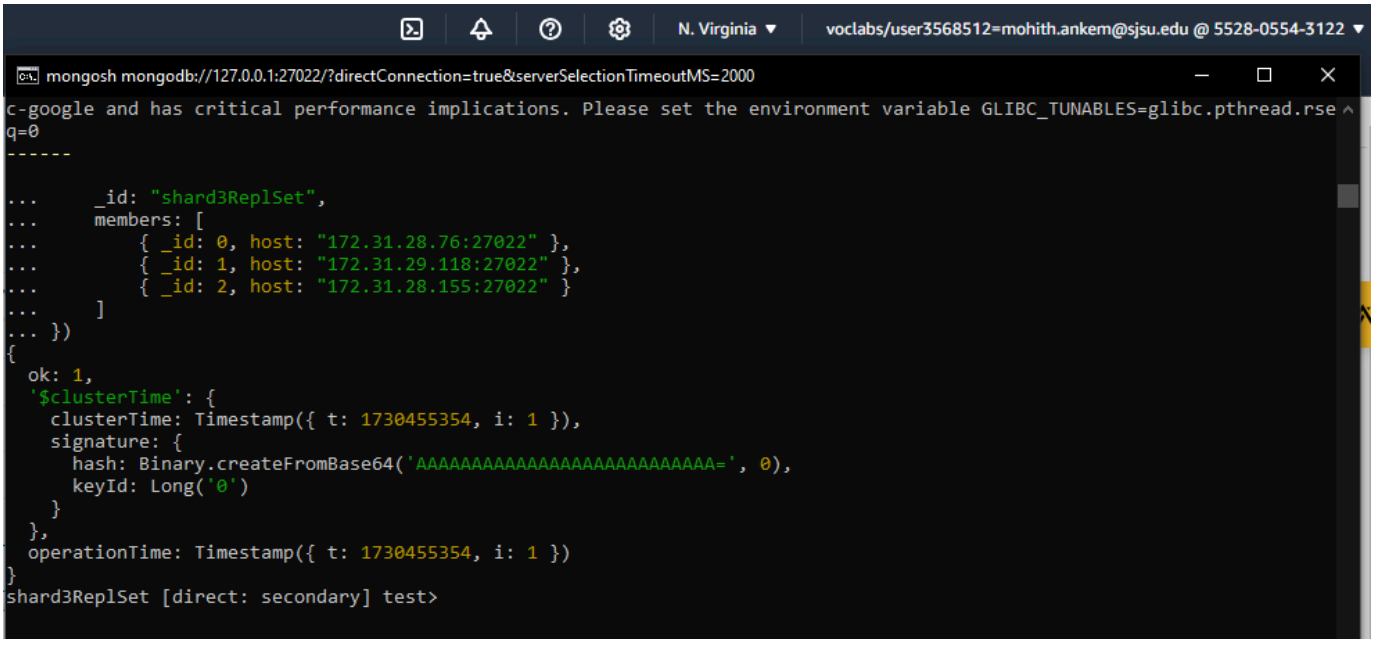
Shard 1:

```
rs.initiate({

    _id: "shard1ReplSet",

    members: [

        { _id: 0, host: "172.31.28.36:27020" },

        { _id: 1, host: "172.31.29.118:27020" },

        { _id: 2, host: "172.31.28.155:27020" }

    ]

})
```

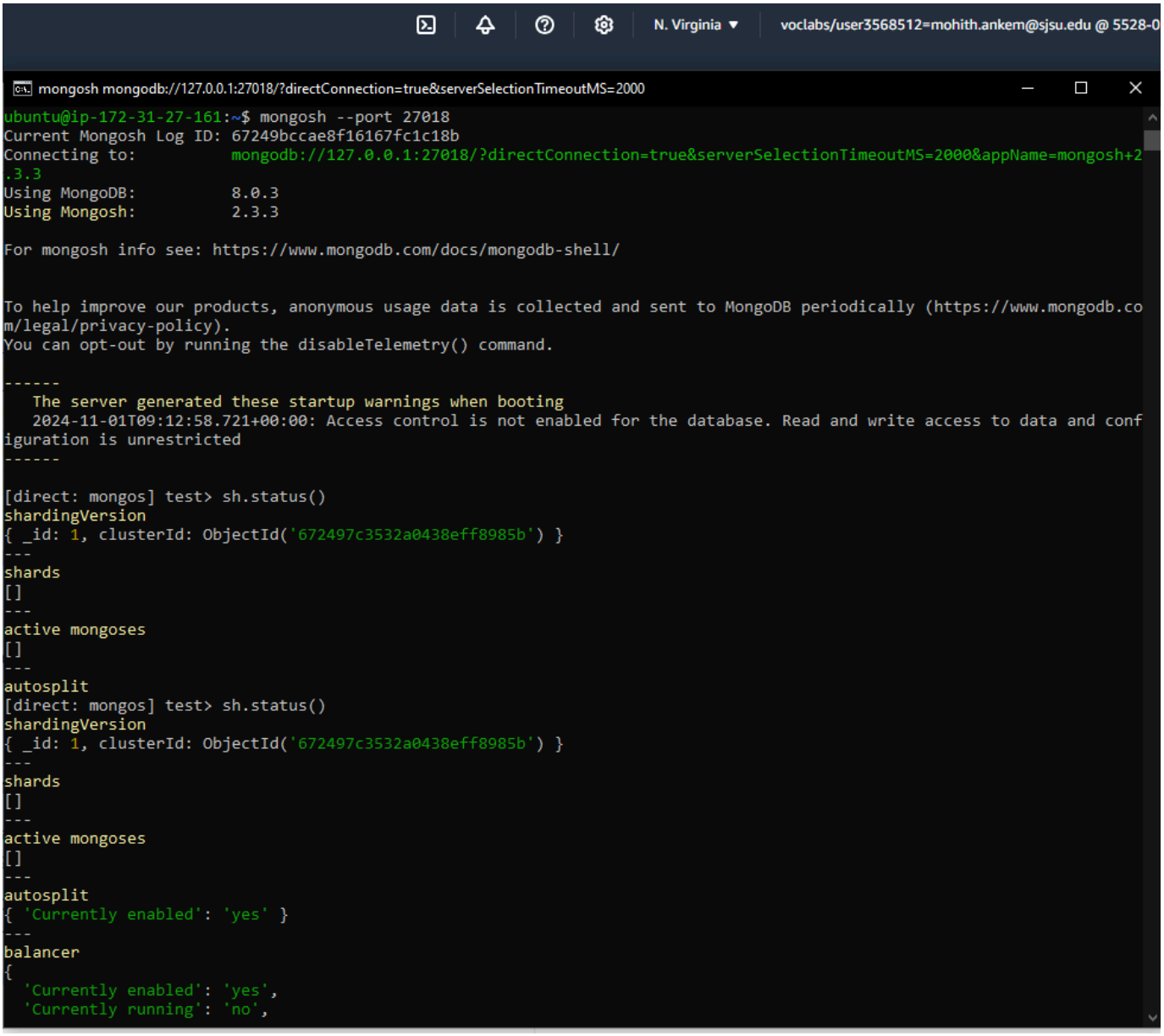
Shard2-Initiate:



Shard 3-Initiate:

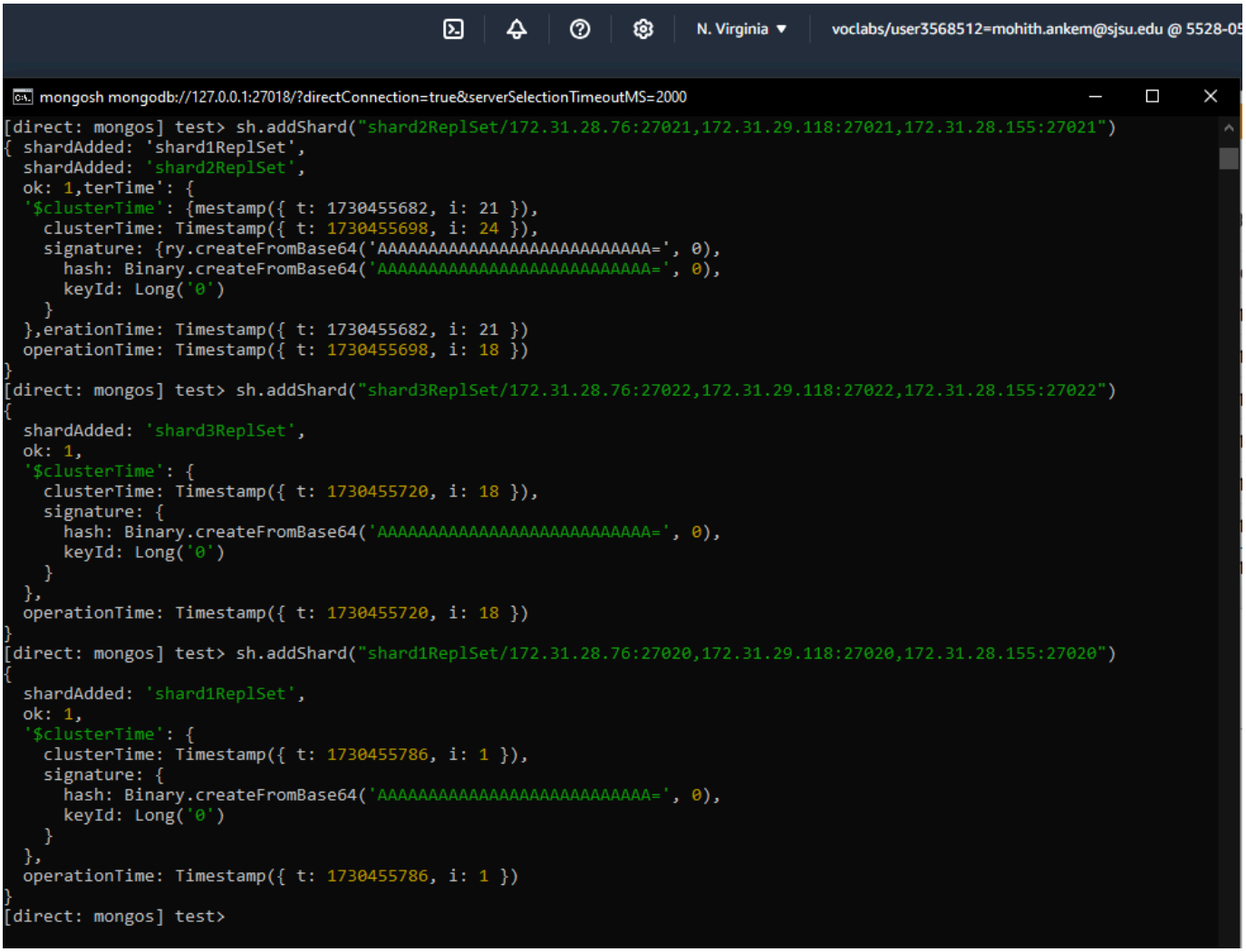


sh.status() before adding shards:

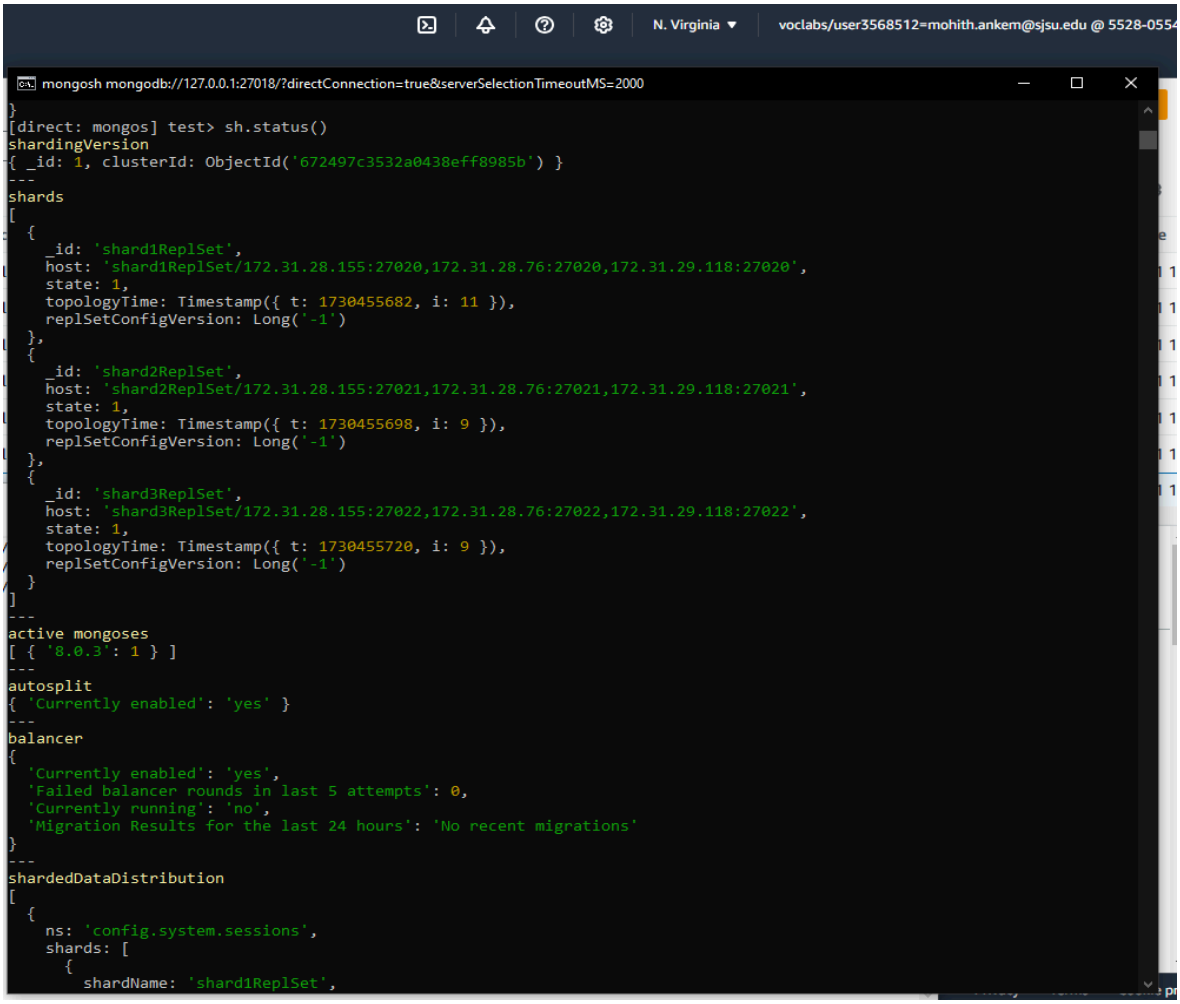


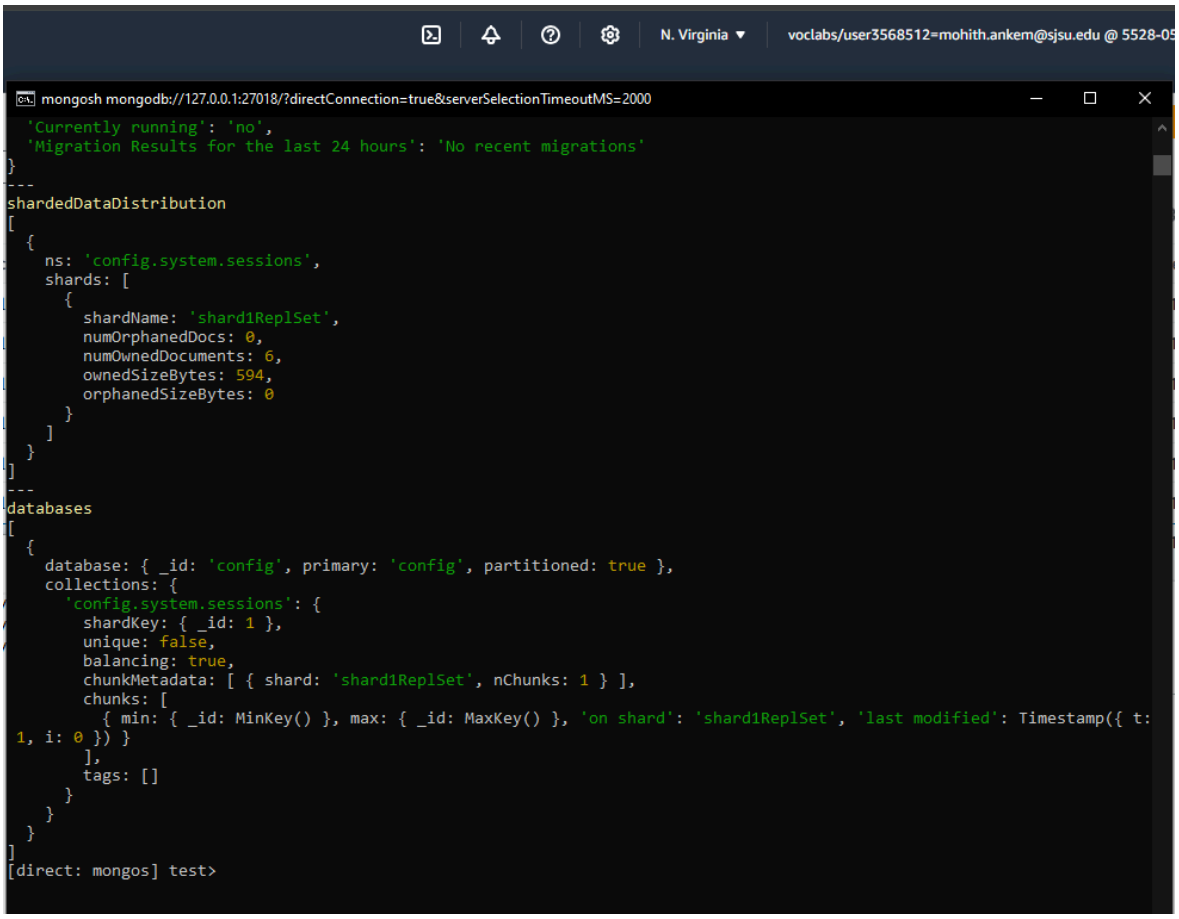
9. Add the shards and ensure you include the result of sh.status() after adding them.

sh.addShard() Command is used to add the shard servers.



sh.status() after adding the shards:


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
}
[direct: mongos] test> sh.status()
shardingVersion
{ _id: 1, clusterId: ObjectId('672497c3532a0438eff8985b') }
---
shards
[
{
_id: 'shard1ReplSet',
host: 'shard1ReplSet/172.31.28.155:27020,172.31.28.76:27020,172.31.29.118:27020',
state: 1,
topologyTime: Timestamp({ t: 1730455682, i: 11 }),
replSetConfigVersion: Long('-1')
},
{
_id: 'shard2ReplSet',
host: 'shard2ReplSet/172.31.28.155:27021,172.31.28.76:27021,172.31.29.118:27021',
state: 1,
topologyTime: Timestamp({ t: 1730455698, i: 9 }),
replSetConfigVersion: Long('-1')
},
{
_id: 'shard3ReplSet',
host: 'shard3ReplSet/172.31.28.155:27022,172.31.28.76:27022,172.31.29.118:27022',
state: 1,
topologyTime: Timestamp({ t: 1730455720, i: 9 }),
replSetConfigVersion: Long('-1')
}
]
---
active mongoses
[{ '8.0.3': 1 }]
---
autosplit
{ 'Currently enabled': 'yes' }
---
balancer
{
'Currently enabled': 'yes',
'Failed balancer rounds in last 5 attempts': 0,
'Currently running': 'no',
'Migration Results for the last 24 hours': 'No recent migrations'
}
---
shardedDataDistribution
[
{
ns: 'config.system.sessions',
shards: [
{
shardName: 'shard1ReplSet',


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-05
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
'Currently running': 'no',
'Migration Results for the last 24 hours': 'No recent migrations'
}
---
shardedDataDistribution
[
{
ns: 'config.system.sessions',
shards: [
{
shardName: 'shard1ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 6,
ownedSizeBytes: 594,
orphanedSizeBytes: 0
}
]
}
]
---
databases
[
{
database: { _id: 'config', primary: 'config', partitioned: true },
collections: {
'config.system.sessions': {
shardKey: { _id: 1 },
unique: false,
balancing: true,
chunkMetadata: [{ shard: 'shard1ReplSet', nChunks: 1 }],
chunks: [
{ min: { _id: MinKey() }, max: { _id: MaxKey() }, 'on shard': 'shard1ReplSet', 'last modified': Timestamp({ t:
1, i: 0 }) }
],
tags: []
}
}
}
]
[direct: mongos] test>

10. Enable the shards and explain the nature of the shard key (ascending, random, or location-based) as well as the sharding strategy (range-based or hash-based) used in your deployment.

A. Enabling Shards:

We enabled sharding on the database testdb with the command:

db.runCommand({ enableSharding: "testdb" });

After enabling sharding on testdb, we shard the grading collection with the specified shard key using the command:

sh.shardCollection("testdb.grading", { student_id: "hashed" });

B. Shard Key:

The chosen shard key for the grading collection is **student_id**, with a **hashed sharding strategy.**

1. Nature of Shard Key:

Hashed keys are random in nature. They work by taking the value of the specified field (student_id) and hashing it to distribute documents across shards uniformly.

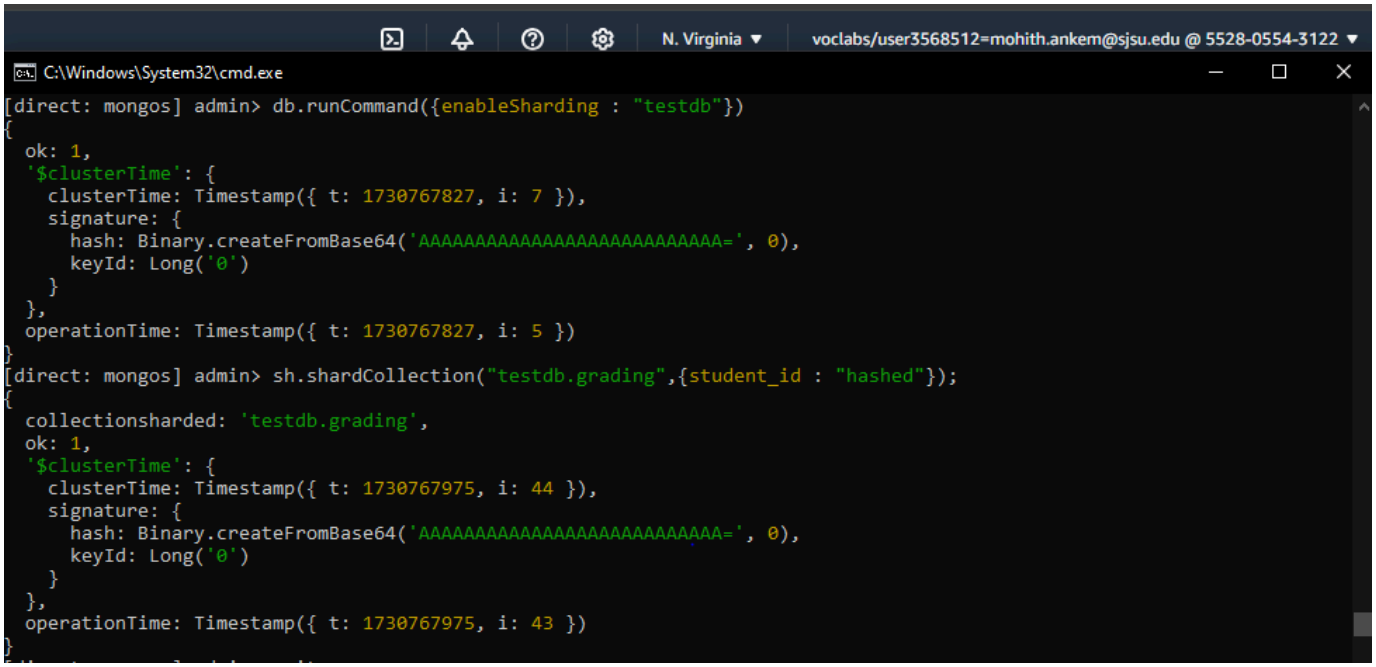
C. Sharding Strategy:

We used a hash-based sharding strategy by specifying **{ "student_id": "hashed" }** as the shard key.

Reason for Choosing Hash-Based Sharding:

In this deployment, hash-based sharding is suitable because it provides consistent and balanced data distribution. Since student_id has a high cardinality (many unique values), hashing each value distributes the documents randomly across the shards.

This approach ensures that each shard receives an approximately equal amount of data, helping to balance the load and optimize query performance.



11. Populate the cluster with data using a public dataset. Explain your collection and include the code to populate the data, along with the result of sh.status() after the data is populated. Specify the URL for the dataset. (Refer to Task 12 to choose an appropriate dataset for executing the given queries. You are allowed to clean and reduce the public dataset of your choice to populate a reasonable amount of data to be distributed across the shards. You may determine what constitutes a reasonable amount.) You are not permitted to use zips.json provided in the prior assignment. If the public dataset you choose is not large enough, you are allowed to synthesize additional data to scale it.

- Dataset:

  The dataset used is based on a modified version of a public dataset, grades.json, which contains student grading information. Each document in the collection represents the scores of a student across different types of assessments (exam, quiz, and homework) for a specific class.

  Dataset URL: The dataset can be found on Kaggle website

  https://www.kaggle.com/datasets/shrashtisinghal/mongo-db-datsets?select=grades.json

  In the given link we can find grades.json in the set of mongoDB Datasets.

- Collection Structure:

  The collection, grading, includes the following fields:

  _id: A unique identifier for each document.

  student_id: An integer representing the student ID.

  scores: An array of subdocuments where each entry includes:

  type: The type of assessment (e.g., "exam", "quiz", "homework").

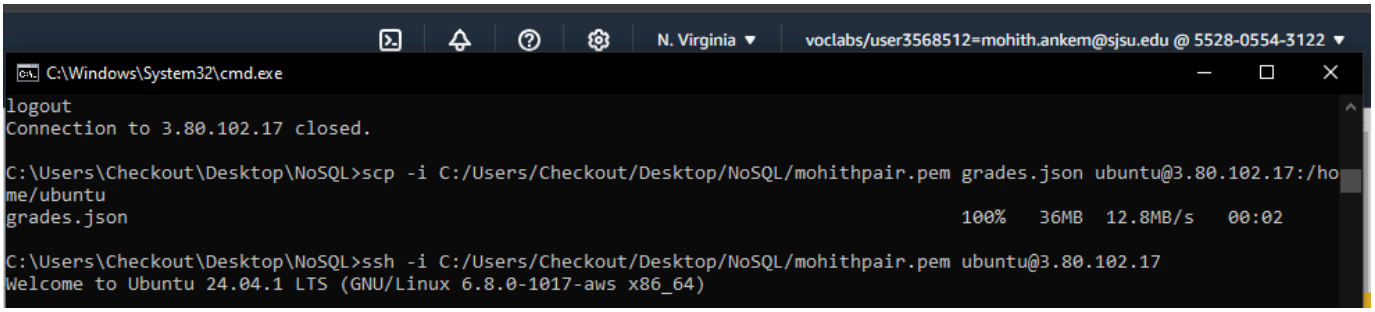
  score: The score received on that assessment.

  class_id: An integer representing the class or course ID.

  This data structure allows us to query based on student_id, class_id, or within the array field scores using operators like $elemMatch and aggregation.

- Shard Key and Sharding Strategy:

  We used student_id as the shard key with a hashed sharding strategy, ensuring a balanced distribution of documents across shards. The hashed strategy helps achieve even distribution, as the student_id values have sufficient cardinality (uniqueness).

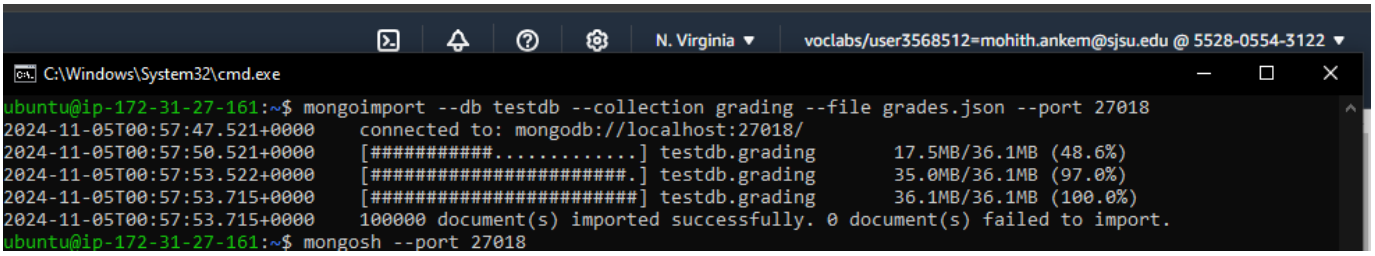


Using the

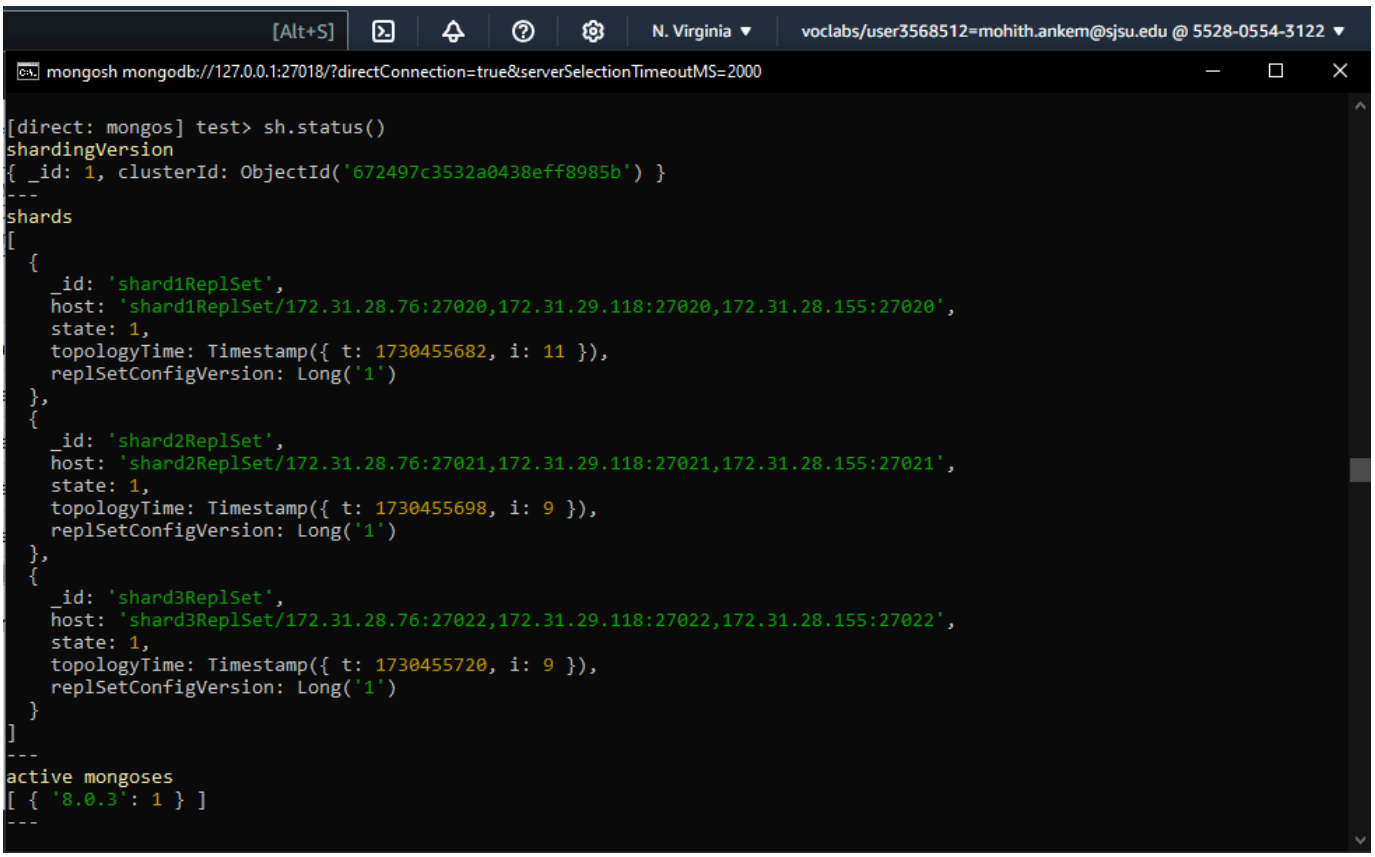
scp -i C:/Users/Desktop/NoSQL/mohithpair.pem grades.json ubuntu@publicIP:/home/ubuntu

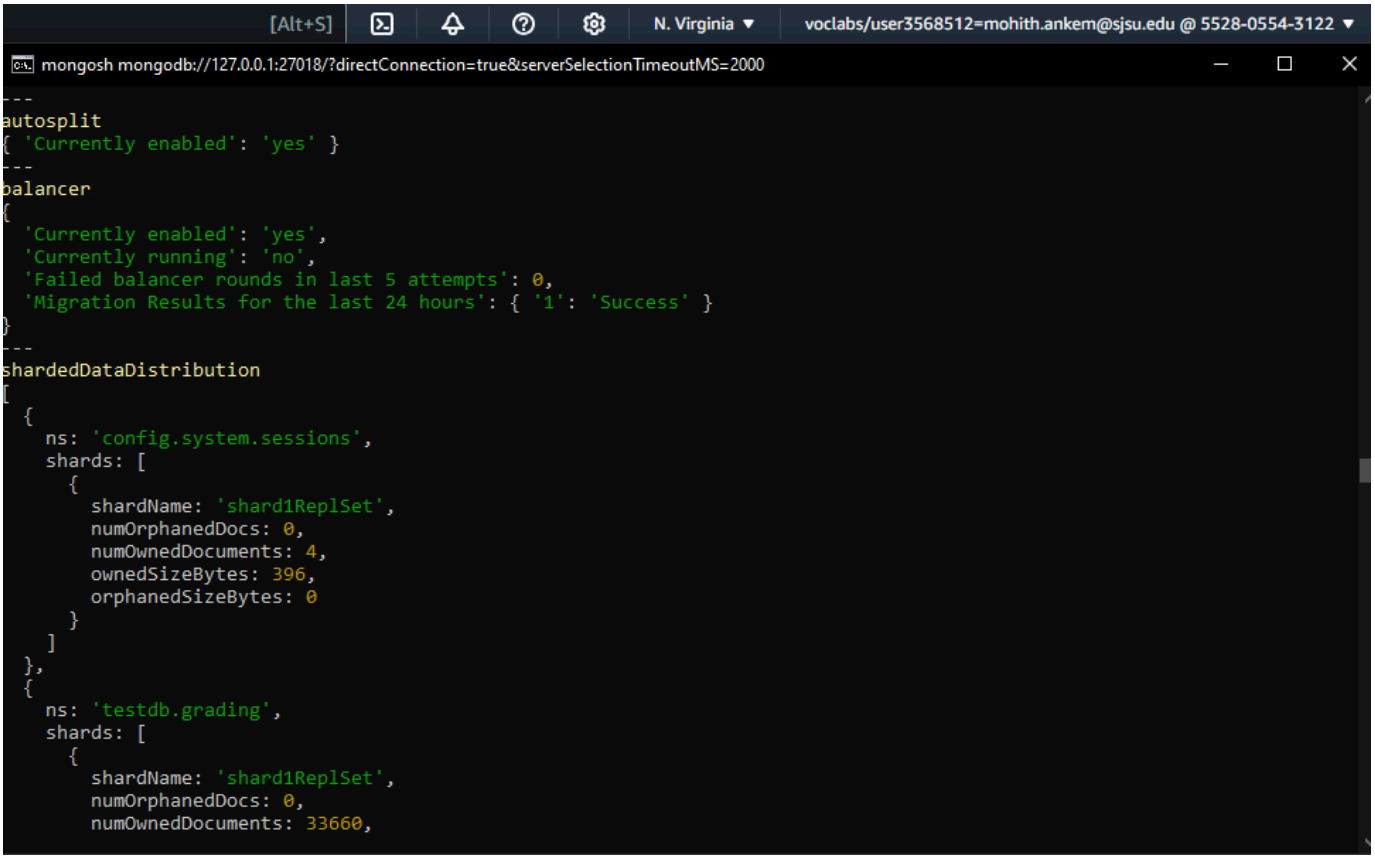
We uploaded the file from the local machine to the server.

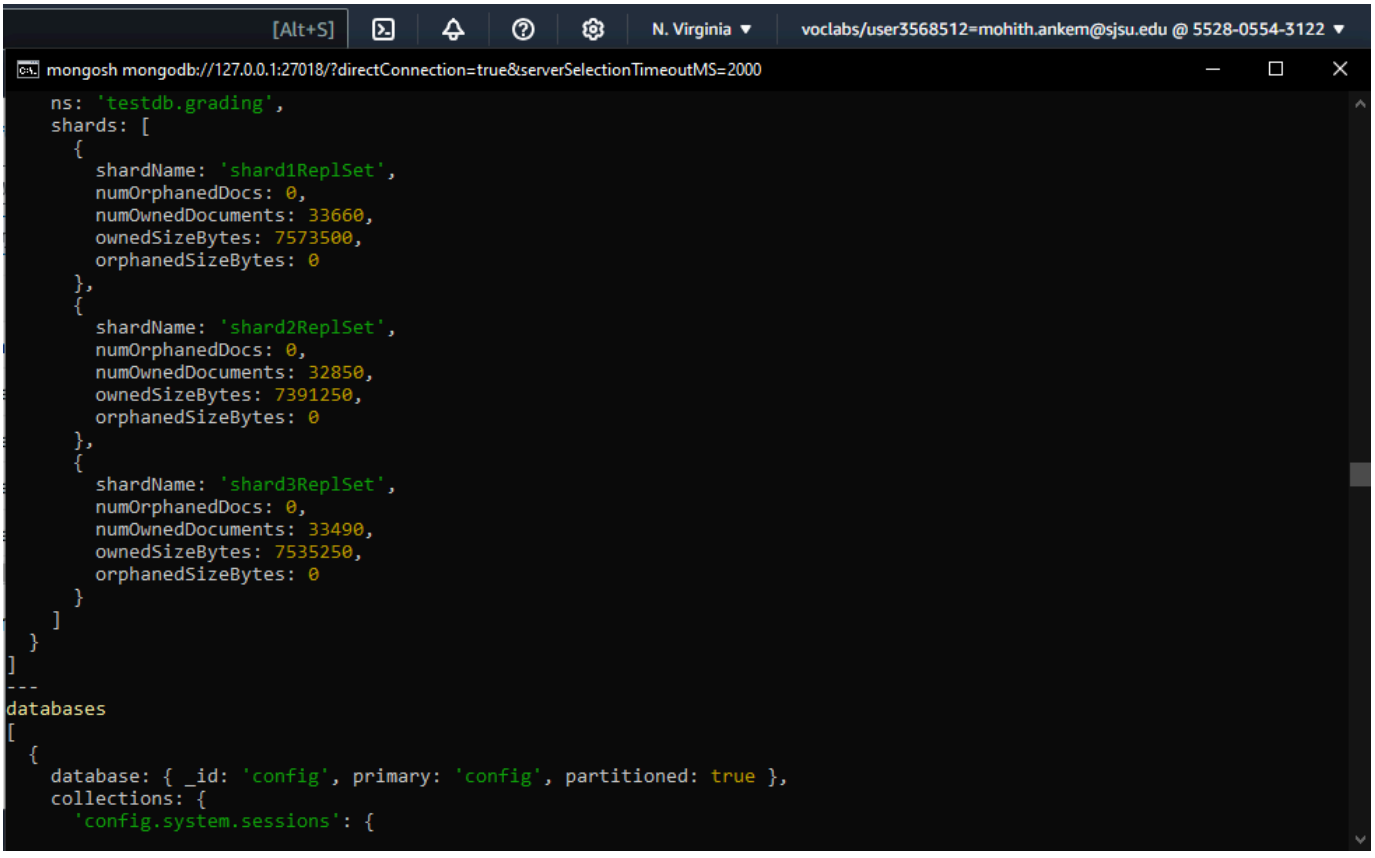
Then after enabling the shard database and sharding strategy as explained in the above question we import the dataset into mongos using **mongoimport** command.

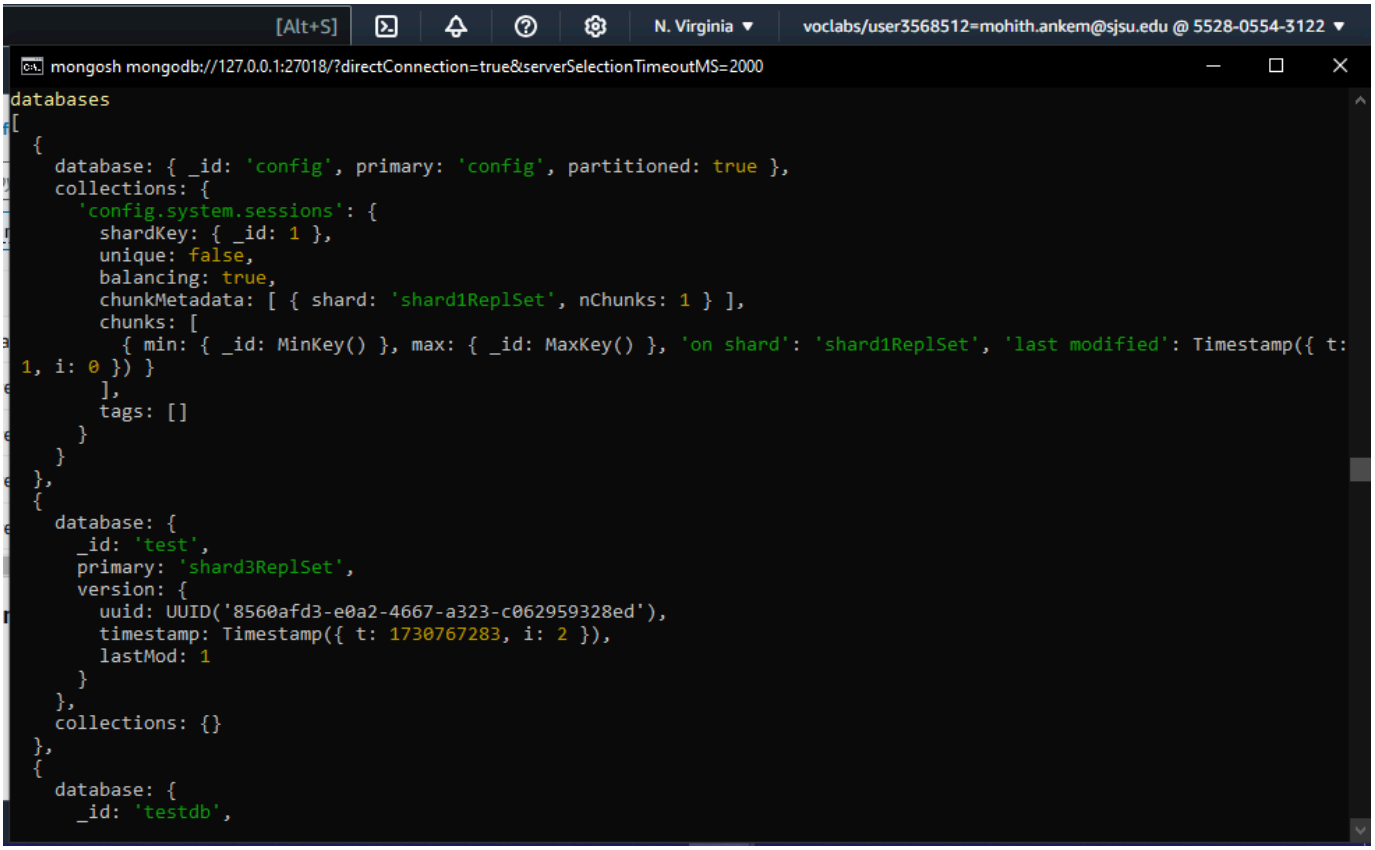


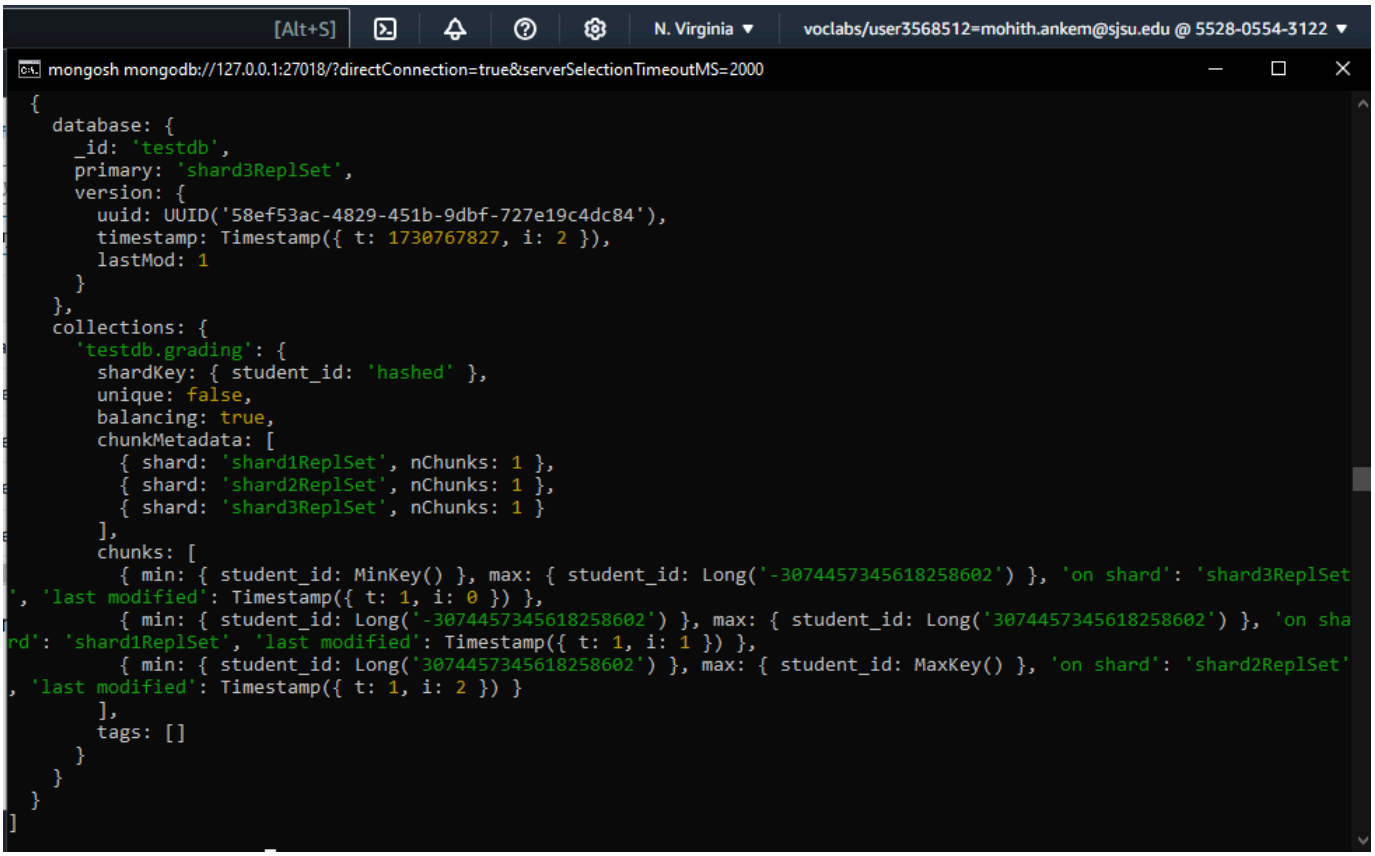
sh.status() after data population:




[Alt+S]
N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
---
autosplit
{ 'Currently enabled': 'yes' }
---
balancer
{
'Currently enabled': 'yes',
'Currently running': 'no',
'Failed balancer rounds in last 5 attempts': 0,
'Migration Results for the last 24 hours': { '1': 'Success' }
}
---
shardedDataDistribution
[
{
ns: 'config.system.sessions',
shards: [
{
shardName: 'shard1ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 4,
ownedSizeBytes: 396,
orphanedSizeBytes: 0
}
]
},
{
ns: 'testdb.grading',
shards: [
{
shardName: 'shard1ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 33660,


[Alt+S]
N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
ns: 'testdb.grading',
shards: [
{
shardName: 'shard1ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 33660,
ownedSizeBytes: 7573500,
orphanedSizeBytes: 0
},
{
shardName: 'shard2ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 32850,
ownedSizeBytes: 7391250,
orphanedSizeBytes: 0
},
{
shardName: 'shard3ReplSet',
numOrphanedDocs: 0,
numOwnedDocuments: 33490,
ownedSizeBytes: 7535250,
orphanedSizeBytes: 0
}
]
}
]
---
databases
[
{
database: { _id: 'config', primary: 'config', partitioned: true },
collections: {
'config.system.sessions': {



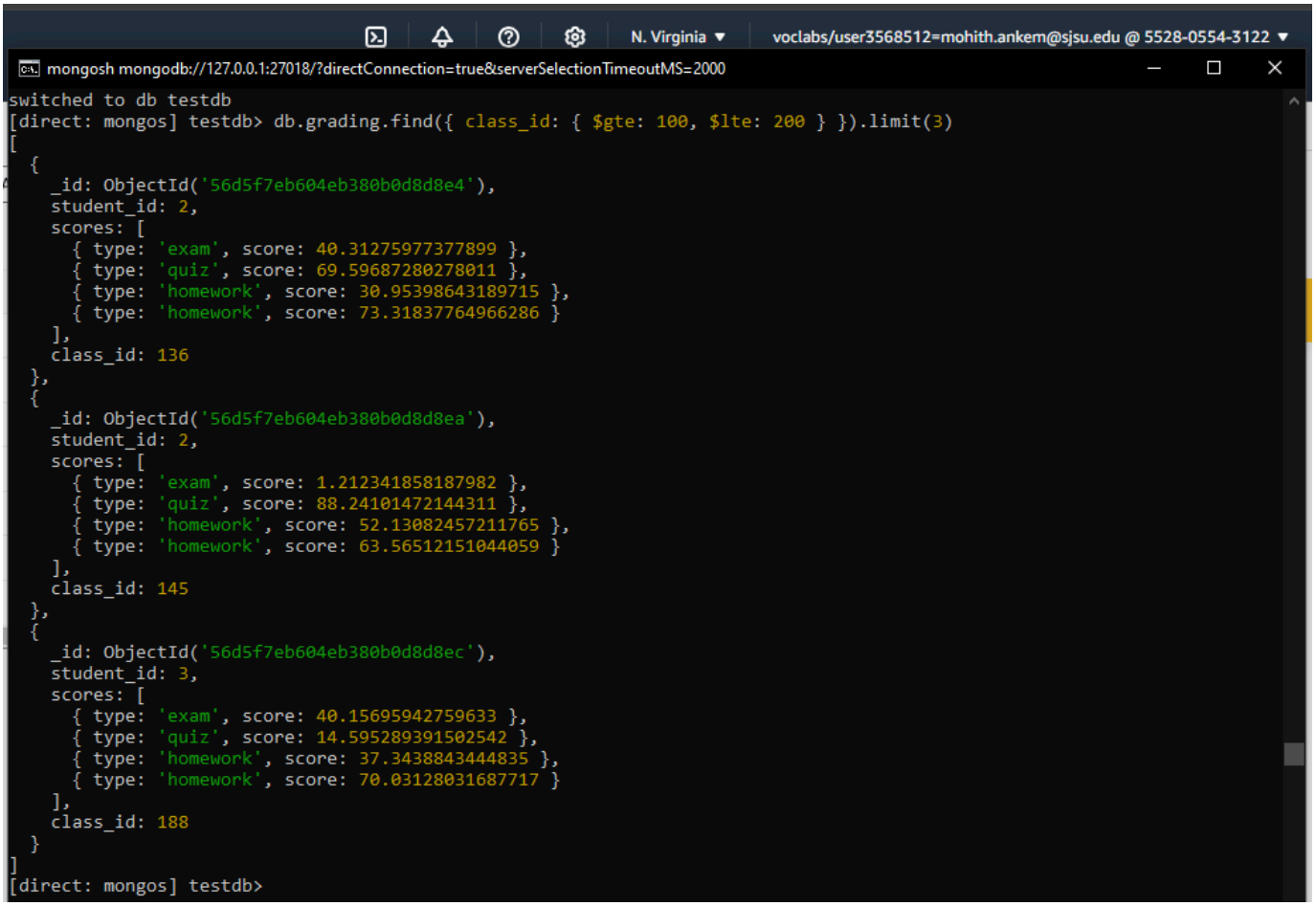


Since we used hashing strategy the documents are split into shard1ReplSet has 33660 documents and 2 chunks, shard2ReplSet has 32850 documents and 1 chunk, whereas shard3ReplSet has 33490 documents and 1 chunks.

12. Generate the following queries for the populated data. For each query, show its execution time and also show which shard served the query.
    1. **A range query to find documents in a given range.**

db.grading.find({ student_id: { $gte: 100, $lte: 200 } }).explain("executionStats");

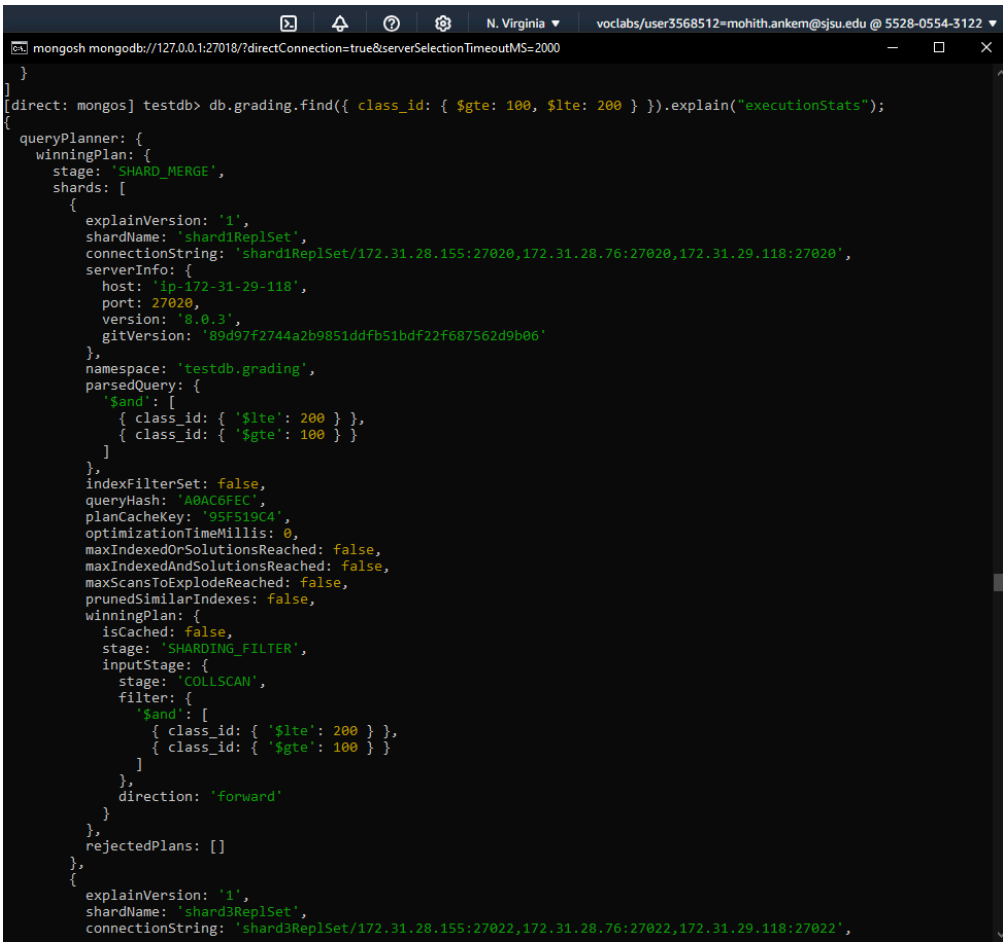
The range query i have used retrieves documents that has student_id in between 100 and 200. Output limited to 3.

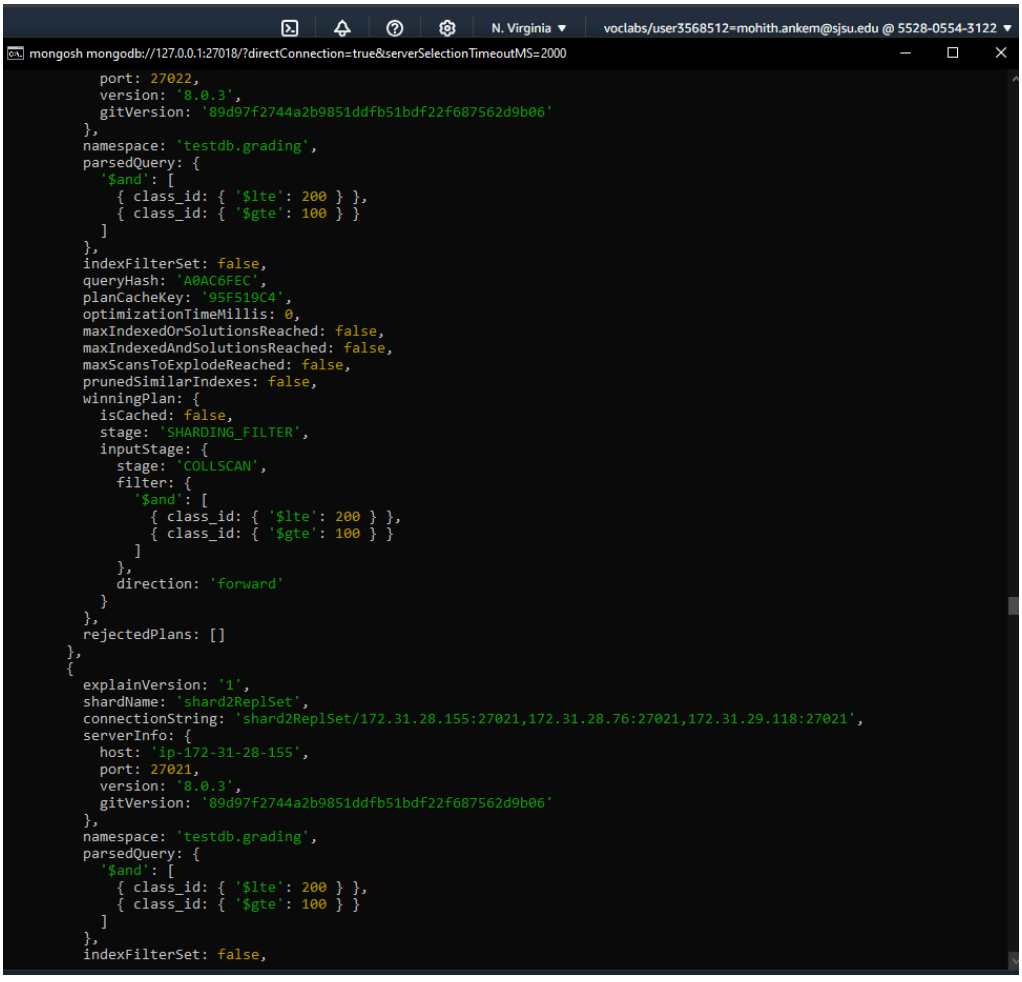


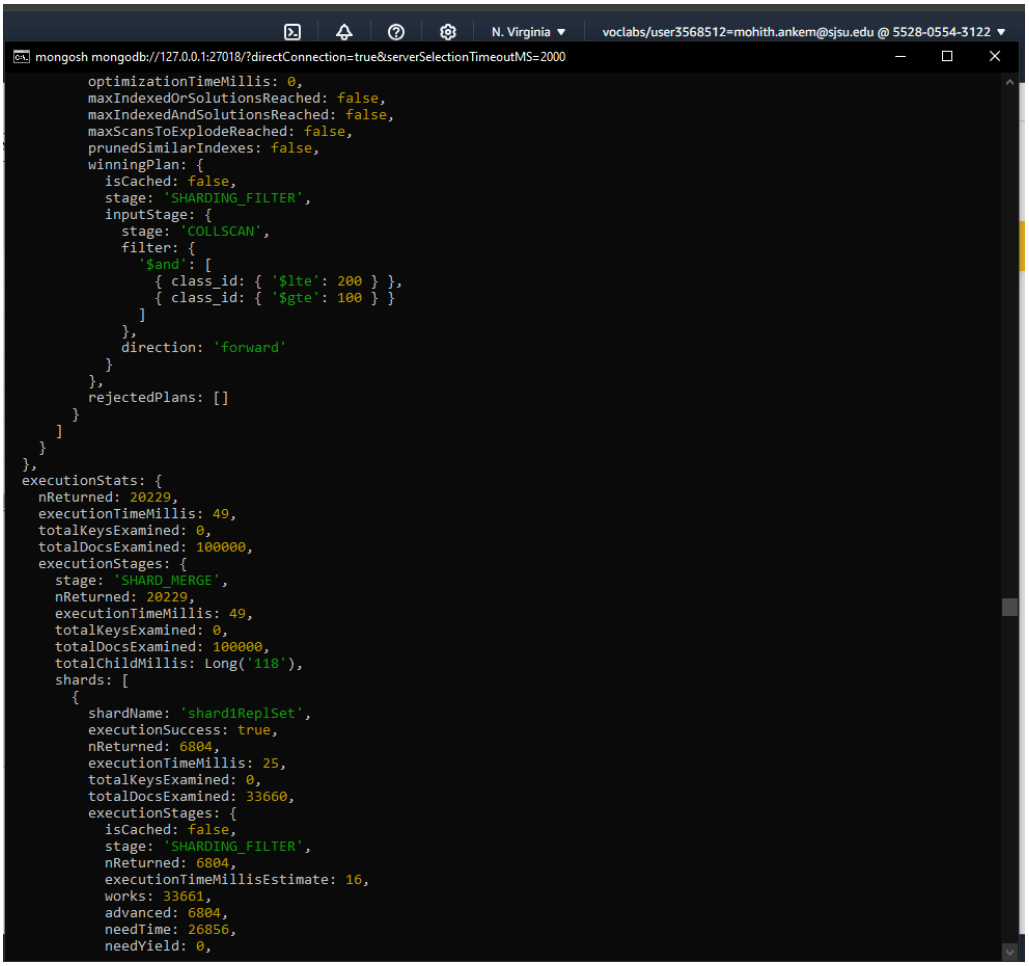
Execution stats after running the query:

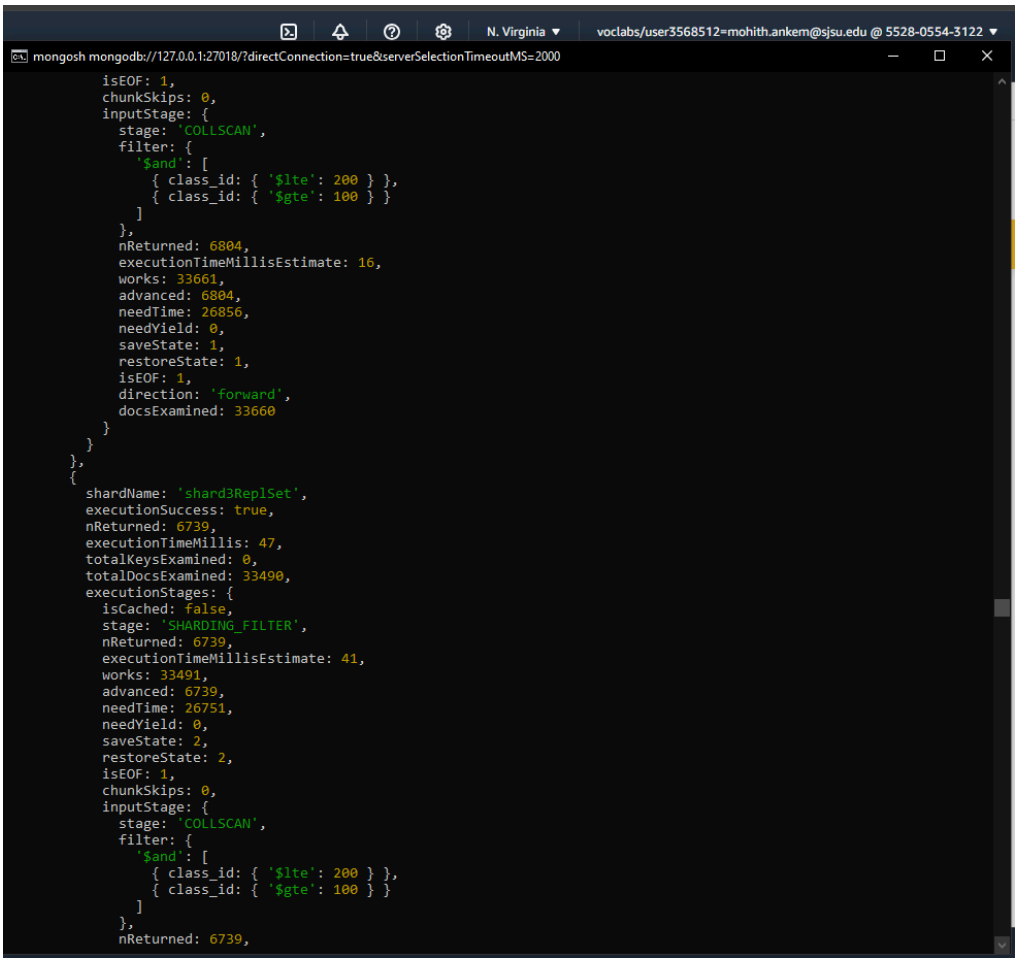
For this Range query since we used hashed strategy all 3 shards served the query.

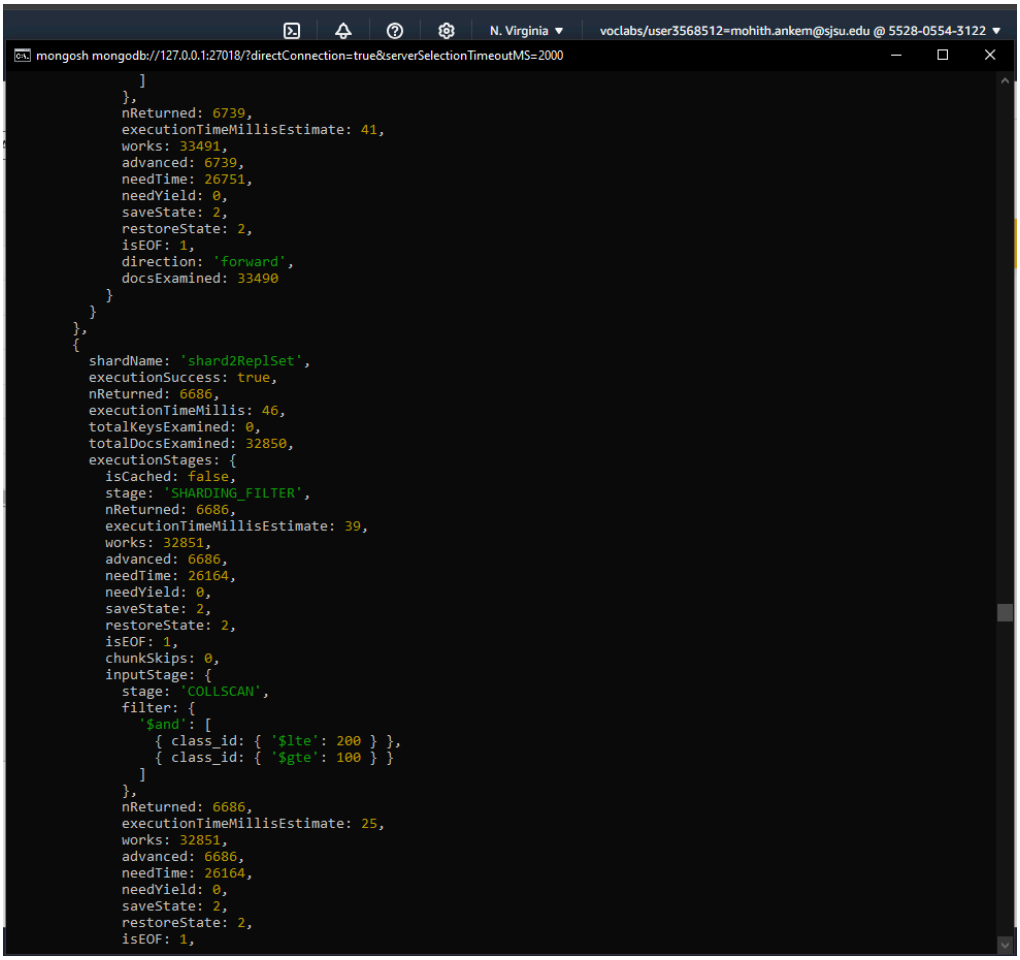
Shard1 has returned around 6884 docs and shard2 has returned around 6686 docs and shard3 has returned around 6739 docs . The total Execution time it took to run the query is **49ms.**

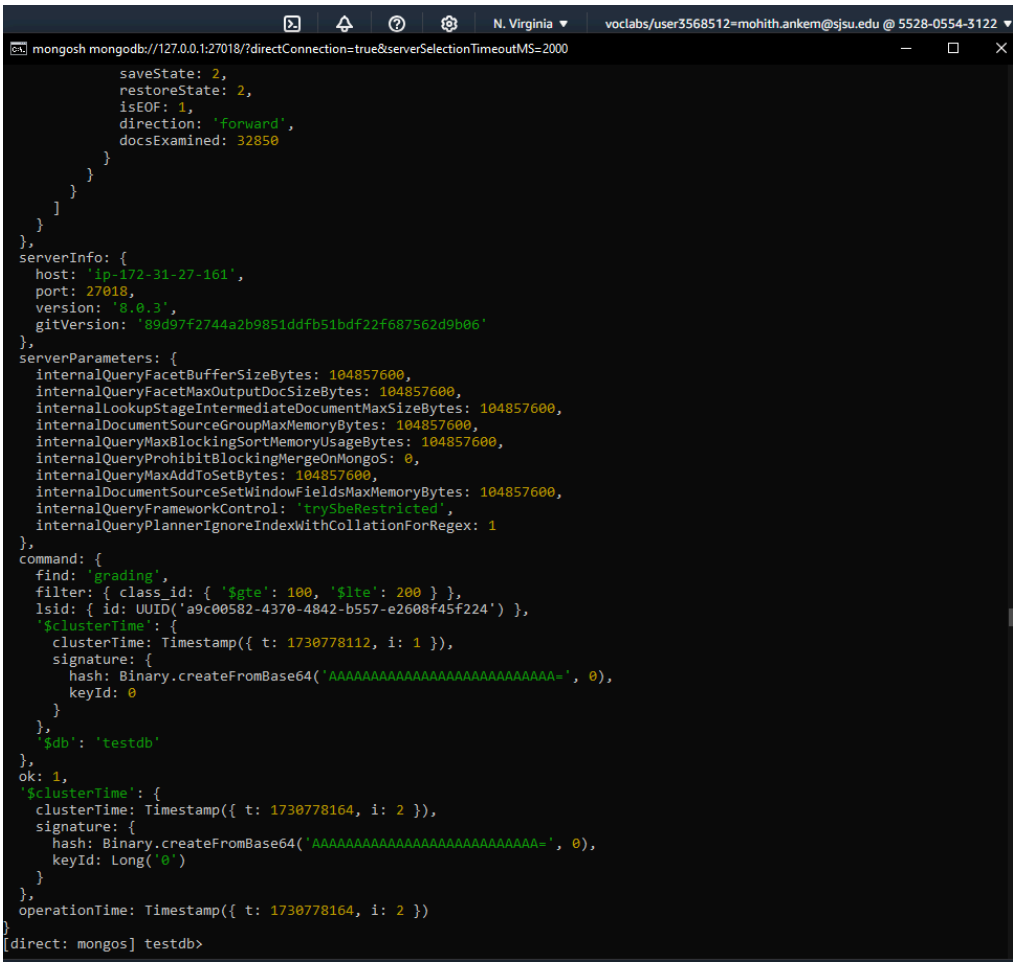

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
[direct: mongos] testdb> db.grading.find({ class_id: { $gte: 100, $lte: 200 } }).explain("executionStats");
queryPlanner: {
winningPlan: {
stage: 'SHARD_MERGE',
shards: [
explainVersion: '1',
shardName: 'shard1ReplSet',
connectionString: 'shard1ReplSet/172.31.28.155:27020,172.31.28.76:27020,172.31.29.118:27020',
serverInfo: {
host: 'ip-172-31-29-118',
port: 27020,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
namespace: 'testdb.grading',
parsedQuery: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
indexFilterSet: false,
queryHash: 'A0AC6FEC',
planCacheKey: '95F519C4',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'SHARDING_FILTER',
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
direction: 'forward'
rejectedPlans: []
explainVersion: '1',
shardName: 'shard3ReplSet',
connectionString: 'shard3ReplSet/172.31.28.155:27022,172.31.28.76:27022,172.31.29.118:27022',


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
port: 27022,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
namespace: 'testdb.grading',
parsedQuery: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
indexFilterSet: false,
queryHash: 'A0AC6FEC',
planCacheKey: '95F519C4',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'SHARDING_FILTER',
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
direction: 'forward'
rejectedPlans: []
explainVersion: '1',
shardName: 'shard2ReplSet',
connectionString: 'shard2ReplSet/172.31.28.155:27021,172.31.28.76:27021,172.31.29.118:27021',
serverInfo: {
host: 'ip-172-31-28-155',
port: 27021,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
namespace: 'testdb.grading',
parsedQuery: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
indexFilterSet: false,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'SHARDING_FILTER',
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
]
},
direction: 'forward'
}
},
rejectedPlans: []
}
]
}
},
executionStats: {
nReturned: 20229,
executionTimeMillis: 49,
totalKeysExamined: 0,
totalDocsExamined: 100000,
executionStages: {
stage: 'SHARD_MERGE',
nReturned: 20229,
executionTimeMillis: 49,
totalKeysExamined: 0,
totalDocsExamined: 100000,
totalChildMillis: Long('118'),
shards: [
{
shardName: 'shard1ReplSet',
executionSuccess: true,
nReturned: 6804,
executionTimeMillis: 25,
totalKeysExamined: 0,
totalDocsExamined: 33660,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 6804,
executionTimeMillisEstimate: 16,
works: 33661,
advanced: 6804,
needTime: 26856,
needYield: 0,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
]
},
nReturned: 6804,
executionTimeMillisEstimate: 16,
works: 33661,
advanced: 6804,
needTime: 26856,
needYield: 0,
saveState: 1,
restoreState: 1,
isEOF: 1,
direction: 'forward',
docsExamined: 33660
}
}
},
{
shardName: 'shard3ReplSet',
executionSuccess: true,
nReturned: 6739,
executionTimeMillis: 47,
totalKeysExamined: 0,
totalDocsExamined: 33490,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 6739,
executionTimeMillisEstimate: 41,
works: 33491,
advanced: 6739,
needTime: 26751,
needYield: 0,
saveState: 2,
restoreState: 2,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
]
},
nReturned: 6739,

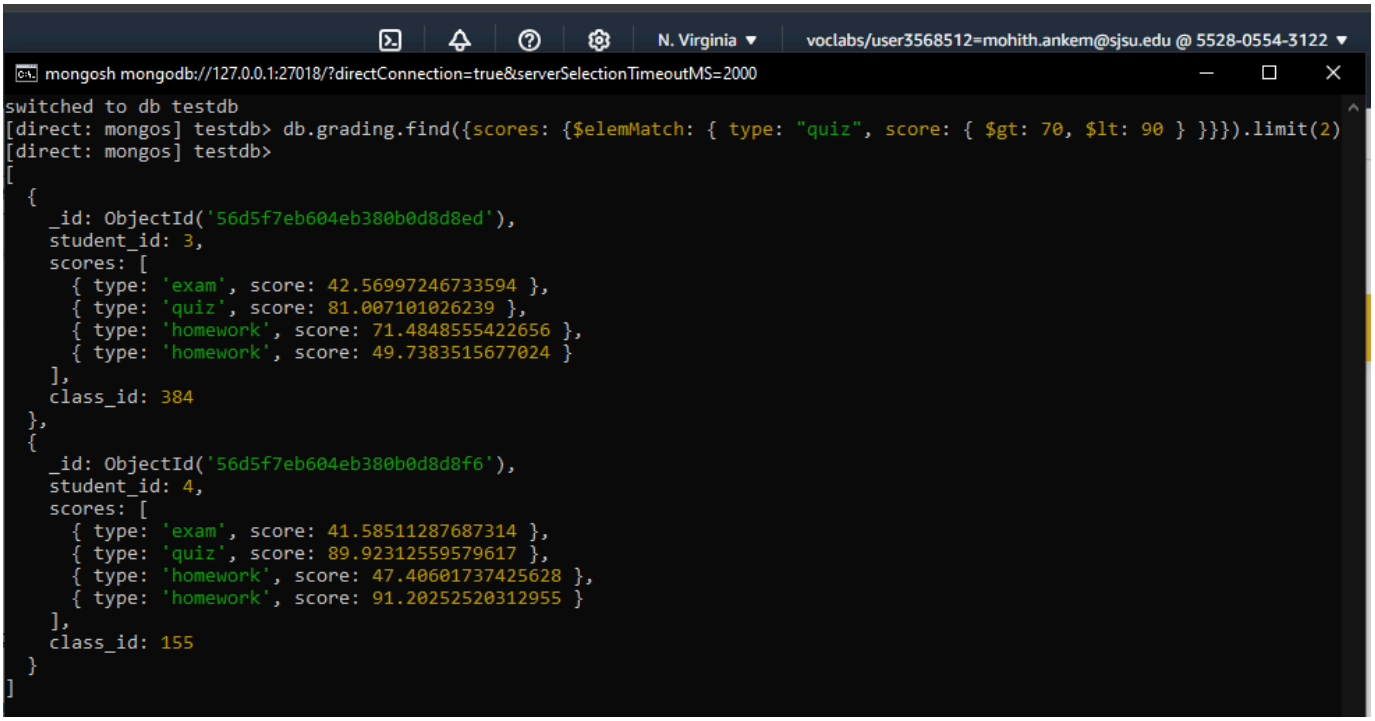

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
]
},
nReturned: 6739,
executionTimeMillisEstimate: 41,
works: 33491,
advanced: 6739,
needTime: 26751,
needYield: 0,
saveState: 2,
restoreState: 2,
isEOF: 1,
direction: 'forward',
docsExamined: 33490
}
}
},
{
shardName: 'shard2ReplSet',
executionSuccess: true,
nReturned: 6686,
executionTimeMillis: 46,
totalKeysExamined: 0,
totalDocsExamined: 32850,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 6686,
executionTimeMillisEstimate: 39,
works: 32851,
advanced: 6686,
needTime: 26164,
needYield: 0,
saveState: 2,
restoreState: 2,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
'$and': [
{ class_id: { '$lte': 200 } },
{ class_id: { '$gte': 100 } }
]
},
nReturned: 6686,
executionTimeMillisEstimate: 25,
works: 32851,
advanced: 6686,
needTime: 26164,
needYield: 0,
saveState: 2,
restoreState: 2,
isEOF: 1,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
saveState: 2,
restoreState: 2,
isEOF: 1,
direction: 'forward',
docsExamined: 32850
}
}
}
]
}
},
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
},
serverParameters: {
internalQueryFacetBufferSizeBytes: 104857600,
internalQueryFacetMaxOutputDocSizeBytes: 104857600,
internalLookupStageIntermediateDocumentMaxSizeBytes: 104857600,
internalDocumentSourceGroupMaxMemoryBytes: 104857600,
internalQueryMaxBlockingSortMemoryUsageBytes: 104857600,
internalQueryProhibitBlockingMergeOnMongoS: 0,
internalQueryMaxAddToSetBytes: 104857600,
internalDocumentSourceSetWindowFieldsMaxMemoryBytes: 104857600,
internalQueryFrameworkControl: 'trySbeRestricted',
internalQueryPlannerIgnoreIndexWithCollationForRegex: 1
},
command: {
find: 'grading',
filter: { class_id: { '$gte': 100, '$lte': 200 } },
lsid: { id: UUID('a9c00582-4370-4842-b557-e2608f45f224') },
'$clusterTime': {
clusterTime: Timestamp({ t: 1730778112, i: 1 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: 0
}
},
'$db': 'testdb'
},
ok: 1,
'$clusterTime': {
clusterTime: Timestamp({ t: 1730778164, i: 2 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: Long('0')
}
},
operationTime: Timestamp({ t: 1730778164, i: 2 })
}
[direct: mongos] testdb>

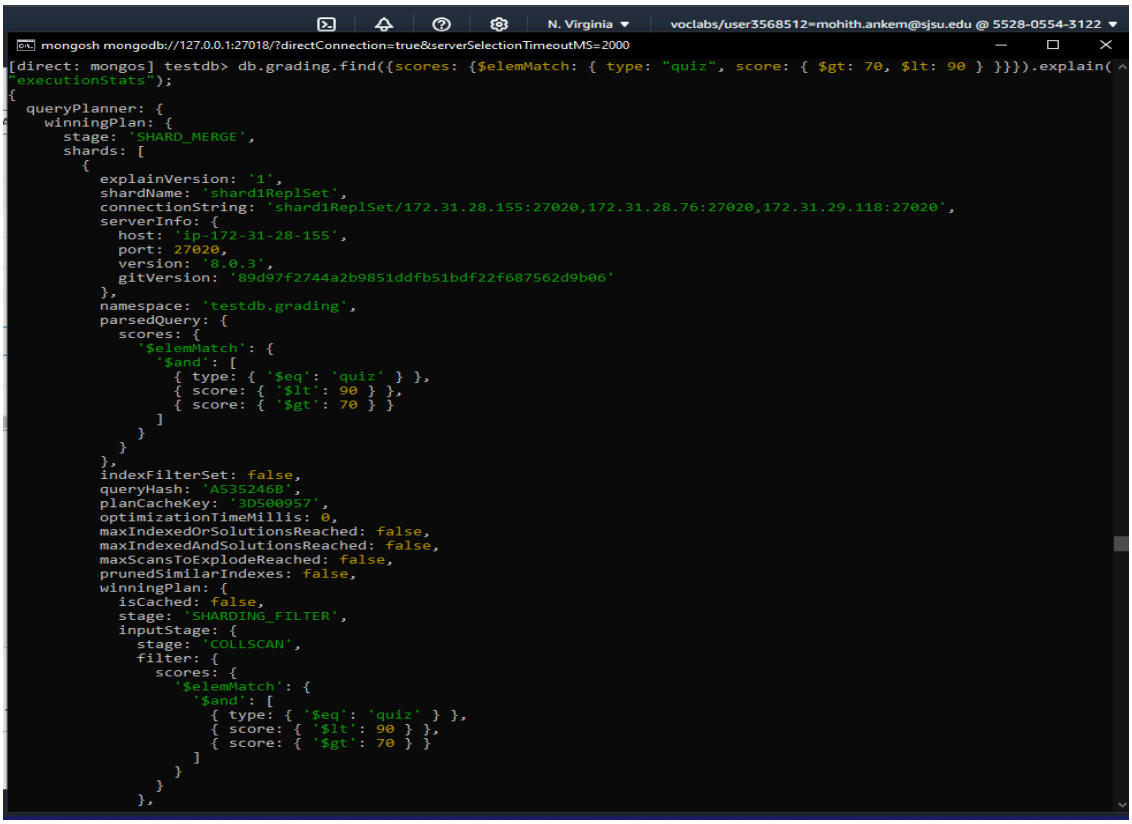
2. A query involving $elemMatch involving at least two conditions.

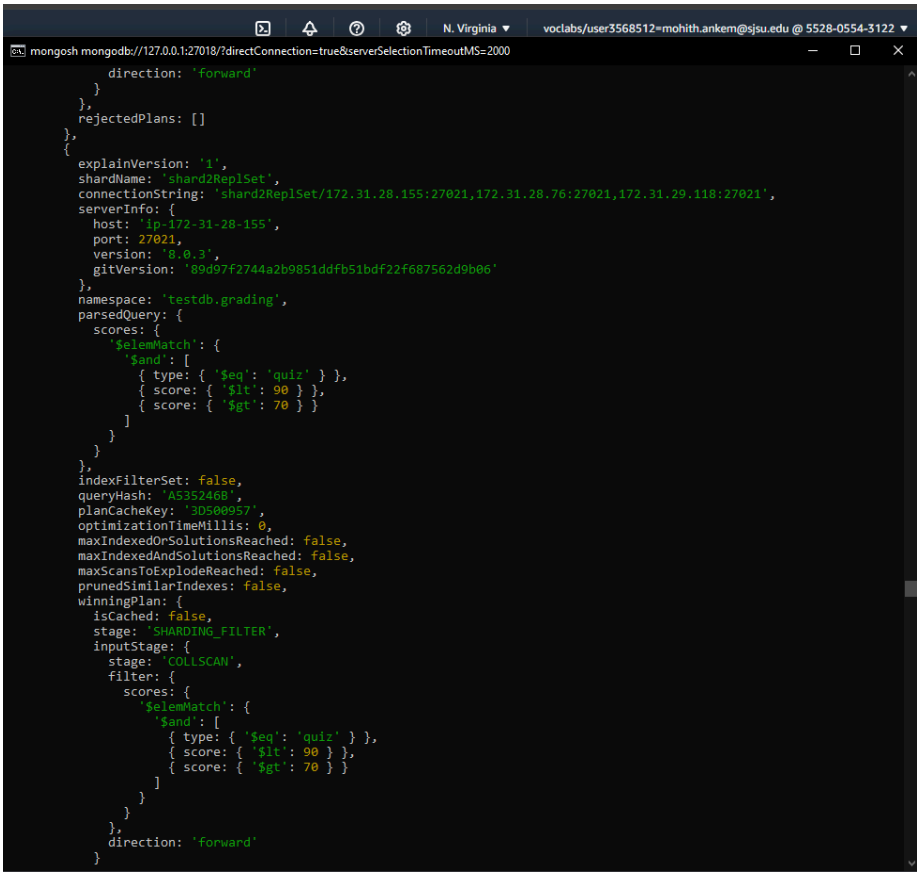
The $elemMatch condition applies to individual elements within the scores array. Specifically, it filters for documents where there is at least one score entry in the array with:

- type equal to "quiz".
- score greater than 70 and less than 90.

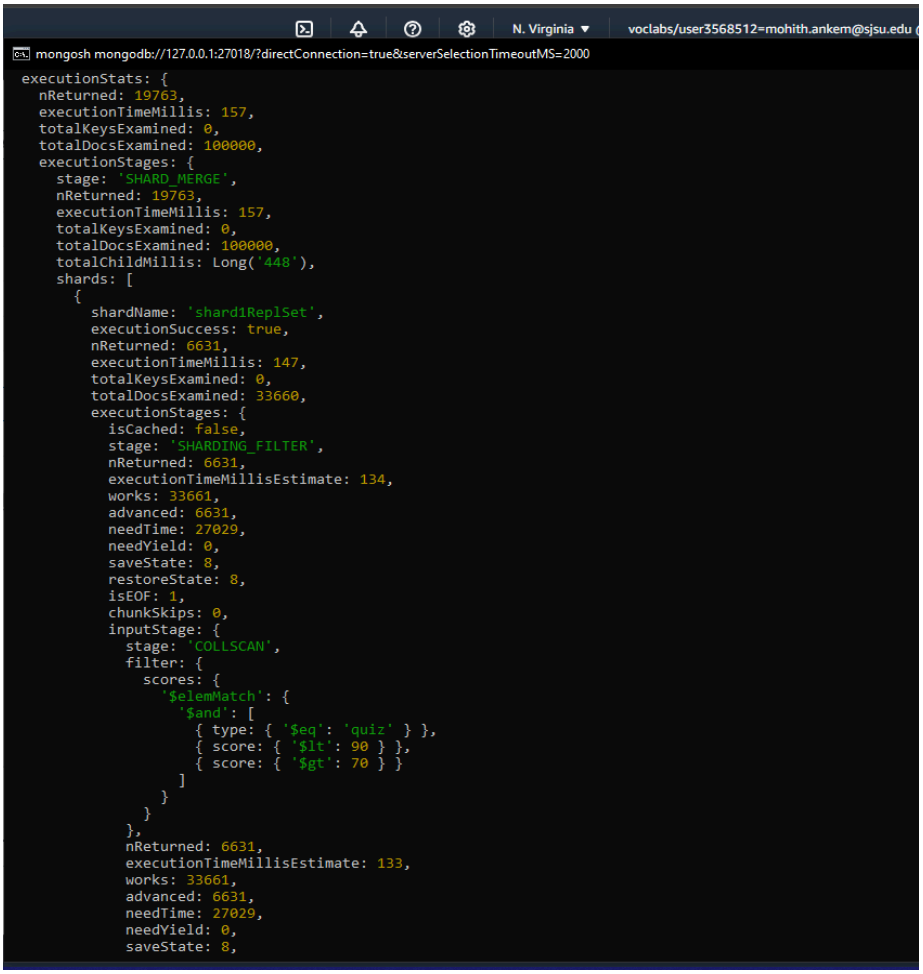


Execution stats:

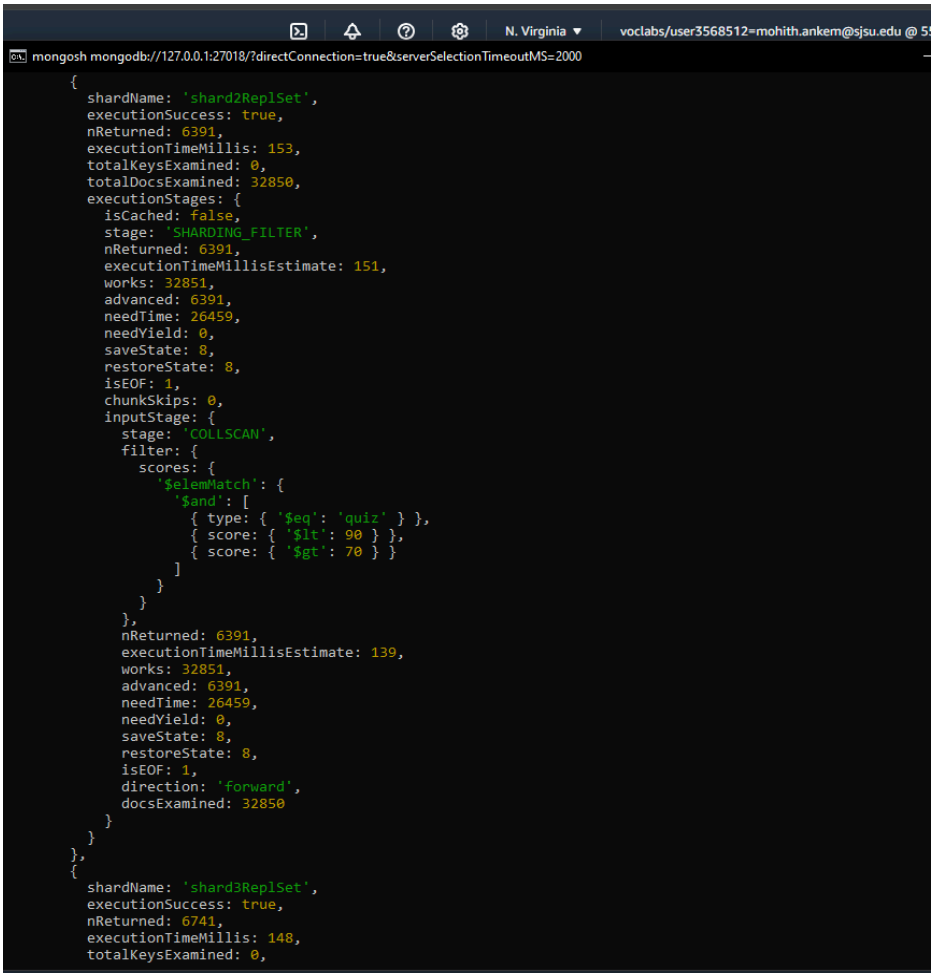



N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
direction: 'forward'
rejectedPlans: []
explainVersion: '1',
shardName: 'shard2ReplSet',
connectionString: 'shard2ReplSet/172.31.28.155:27021,172.31.28.76:27021,172.31.29.118:27021',
serverInfo: {
host: 'ip-172-31-28-155',
port: 27021,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
namespace: 'testdb.grading',
parsedQuery: {
scores: {
'$elemMatch': {
'$and': [
{ type: { '$eq': 'quiz' } },
{ score: { '$lt': 90 } },
{ score: { '$gt': 70 } }
indexFilterSet: false,
queryHash: 'A535246B',
planCacheKey: '3D500957',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'SHARDING_FILTER',
inputStage: {
stage: 'COLLSCAN',
filter: {
scores: {
'$elemMatch': {
'$and': [
{ type: { '$eq': 'quiz' } },
{ score: { '$lt': 90 } },
{ score: { '$gt': 70 } }
direction: 'forward'

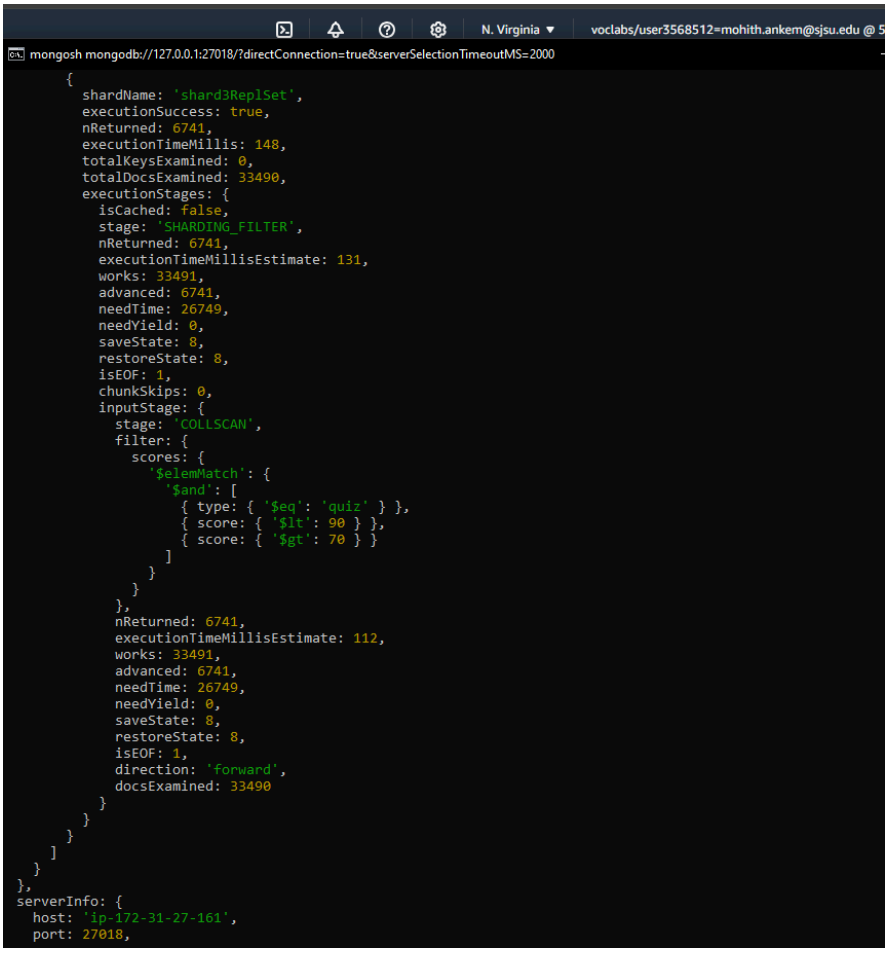

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
explainVersion: '1',
shardName: 'shard3ReplSet',
connectionString: 'shard3ReplSet/172.31.28.155:27022,172.31.28.76:27022,172.31.29.118:27022',
serverInfo: {
host: 'ip-172-31-28-155',
port: 27022,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
namespace: 'testdb.grading',
parsedQuery: {
scores: {
'$elemMatch': {
'$and': [
{ type: { '$eq': 'quiz' } },
{ score: { '$lt': 90 } },
{ score: { '$gt': 70 } }
indexFilterSet: false,
queryHash: 'A535246B',
planCacheKey: '3D500957',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'SHARDING_FILTER',
inputStage: {
stage: 'COLLSCAN',
filter: {
scores: {
'$elemMatch': {
'$and': [
{ type: { '$eq': 'quiz' } },
{ score: { '$lt': 90 } },
{ score: { '$gt': 70 } }
direction: 'forward'
rejectedPlans: []

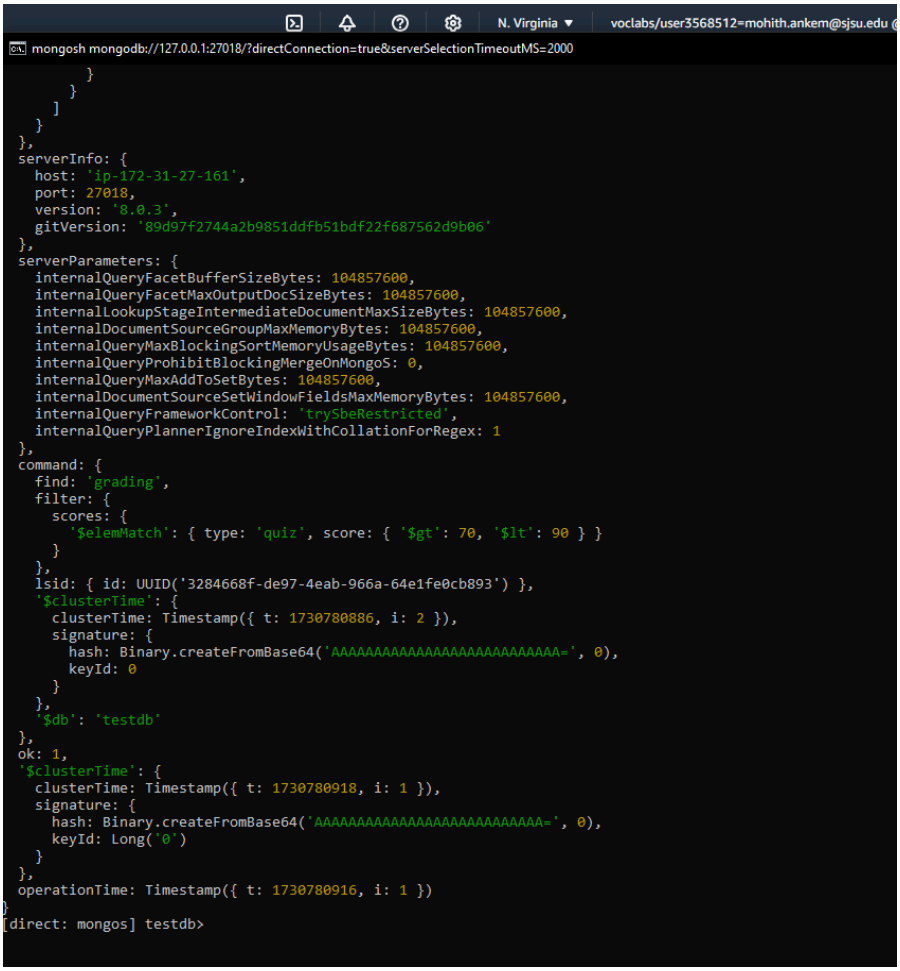



```
executionStats: {
  nReturned: 19763,
  executionTimeMillis: 157,
  totalKeysExamined: 0,
  totalDocsExamined: 100000,
  executionStages: {
    stage: 'SHARD_MERGE',
    nReturned: 19763,
    executionTimeMillis: 157,
    totalKeysExamined: 0,
    totalDocsExamined: 100000,
    totalChildMillis: Long('448'),
    shards: [
      {
        shardName: 'shard1ReplSet',
        executionSuccess: true,
        nReturned: 6631,
        executionTimeMillis: 147,
        totalKeysExamined: 0,
        totalDocsExamined: 33660,
        executionStages: {
          isCached: false,
          stage: 'SHARDING_FILTER',
          nReturned: 6631,
          executionTimeMillisEstimate: 134,
          works: 33661,
          advanced: 6631,
          needTime: 27029,
          needYield: 0,
          saveState: 8,
          restoreState: 8,
          isEOF: 1,
          chunkSkips: 0,
          inputStage: {
            stage: 'COLLSCAN',
            filter: {
              scores: {
                '$elemMatch': {
                  '$and': [
                    { type: { '$eq': 'quiz' } },
                    { score: { '$lt': 90 } },
                    { score: { '$gt': 70 } }
                  ]
                }
              }
            },
            nReturned: 6631,
            executionTimeMillisEstimate: 133,
            works: 33661,
            advanced: 6631,
            needTime: 27029,
            needYield: 0,
            saveState: 8,
```




```
      {
        shardName: 'shard2ReplSet',
        executionSuccess: true,
        nReturned: 6391,
        executionTimeMillis: 153,
        totalKeysExamined: 0,
        totalDocsExamined: 32850,
        executionStages: {
          isCached: false,
          stage: 'SHARDING_FILTER',
          nReturned: 6391,
          executionTimeMillisEstimate: 151,
          works: 32851,
          advanced: 6391,
          needTime: 26459,
          needYield: 0,
          saveState: 8,
          restoreState: 8,
          isEOF: 1,
          chunkSkips: 0,
          inputStage: {
            stage: 'COLLSCAN',
            filter: {
              scores: {
                '$elemMatch': {
                  '$and': [
                    { type: { '$eq': 'quiz' } },
                    { score: { '$lt': 90 } },
                    { score: { '$gt': 70 } }
                  ]
                }
              }
            },
            nReturned: 6391,
            executionTimeMillisEstimate: 139,
            works: 32851,
            advanced: 6391,
            needTime: 26459,
            needYield: 0,
            saveState: 8,
            restoreState: 8,
            isEOF: 1,
            direction: 'forward',
            docsExamined: 32850
          }
        }
      },
      {
        shardName: 'shard3ReplSet',
        executionSuccess: true,
        nReturned: 6741,
        executionTimeMillis: 148,
        totalKeysExamined: 0,
```


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
{
shardName: 'shard3ReplSet',
executionSuccess: true,
nReturned: 6741,
executionTimeMillis: 148,
totalKeysExamined: 0,
totalDocsExamined: 33490,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 6741,
executionTimeMillisEstimate: 131,
works: 33491,
advanced: 6741,
needTime: 26749,
needYield: 0,
saveState: 8,
restoreState: 8,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
scores: {
'$elemMatch': {
'$and': [
{ type: { '$eq': 'quiz' } },
{ score: { '$lt': 90 } },
{ score: { '$gt': 70 } }
]
}
}
},
nReturned: 6741,
executionTimeMillisEstimate: 112,
works: 33491,
advanced: 6741,
needTime: 26749,
needYield: 0,
saveState: 8,
restoreState: 8,
isEOF: 1,
direction: 'forward',
docsExamined: 33490
}
}
}
]
}
},
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,

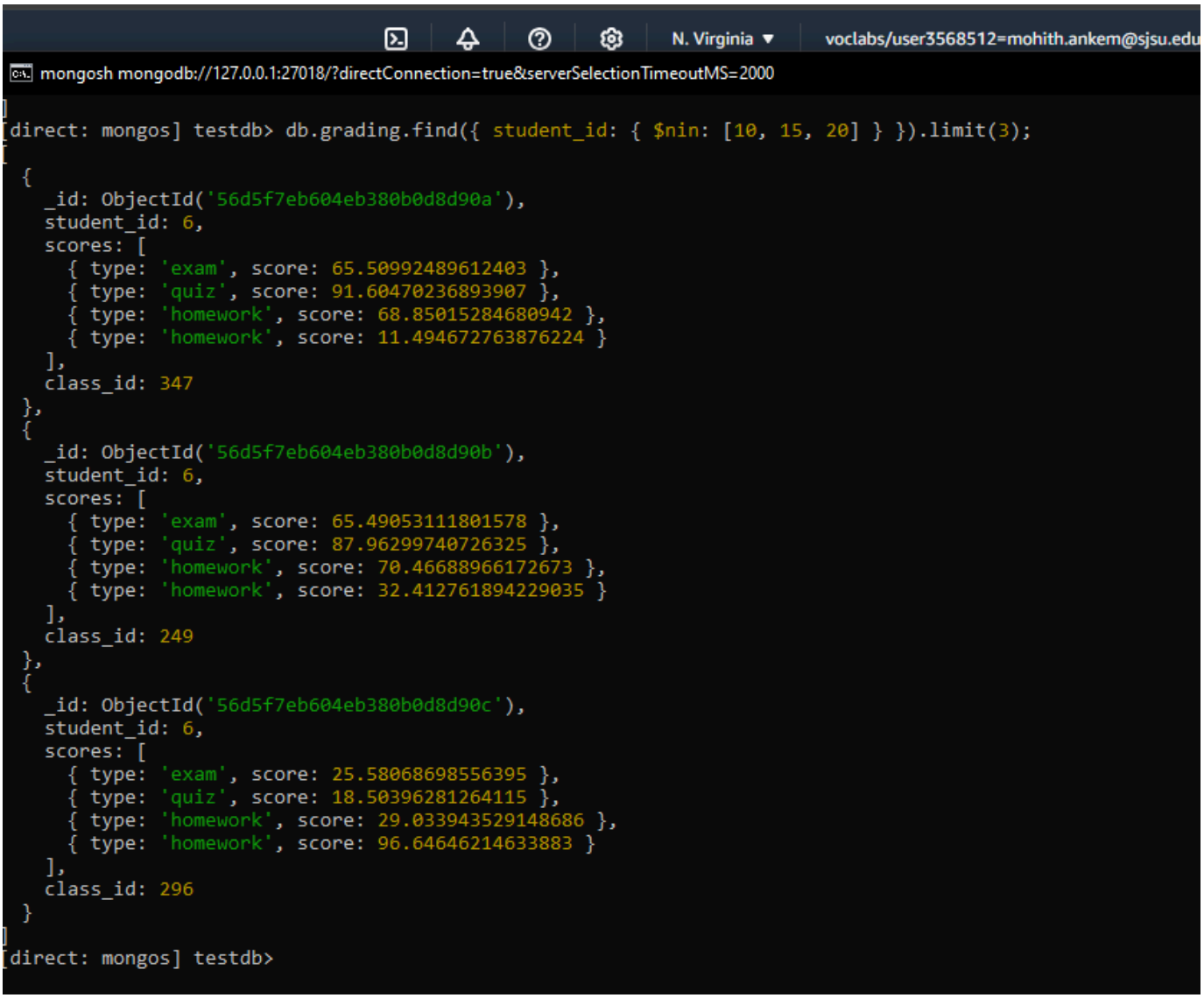

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
}
}
]
}
},
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
},
serverParameters: {
internalQueryFacetBufferSizeBytes: 104857600,
internalQueryFacetMaxOutputDocSizeBytes: 104857600,
internalLookupStageIntermediateDocumentMaxSizeBytes: 104857600,
internalDocumentSourceGroupMaxMemoryBytes: 104857600,
internalQueryMaxBlockingSortMemoryUsageBytes: 104857600,
internalQueryProhibitBlockingMergeOnMongoS: 0,
internalQueryMaxAddToSetBytes: 104857600,
internalDocumentSourceSetWindowFieldsMaxMemoryBytes: 104857600,
internalQueryFrameworkControl: 'trySbeRestricted',
internalQueryPlannerIgnoreIndexWithCollationForRegex: 1
},
command: {
find: 'grading',
filter: {
scores: {
'$elemMatch': { type: 'quiz', score: { '$gt': 70, '$lt': 90 } }
}
},
lsid: { id: UUID('3284668f-de97-4eab-966a-64e1fe0cb893') },
'$clusterTime': {
clusterTime: Timestamp({ t: 1730780886, i: 2 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: 0
}
},
'$db': 'testdb'
},
ok: 1,
'$clusterTime': {
clusterTime: Timestamp({ t: 1730780918, i: 1 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: Long('0')
}
},
operationTime: Timestamp({ t: 1730780916, i: 1 })
}
[direct: mongos] testdb>

It took a total of 157 ms to run the query and in this case also all 3 shards have returned the results with shard1 returned 6631, shard2 returned 6391 and shard3 returned 6741 docs.

3. **A query involving $in, $nin, or $all**

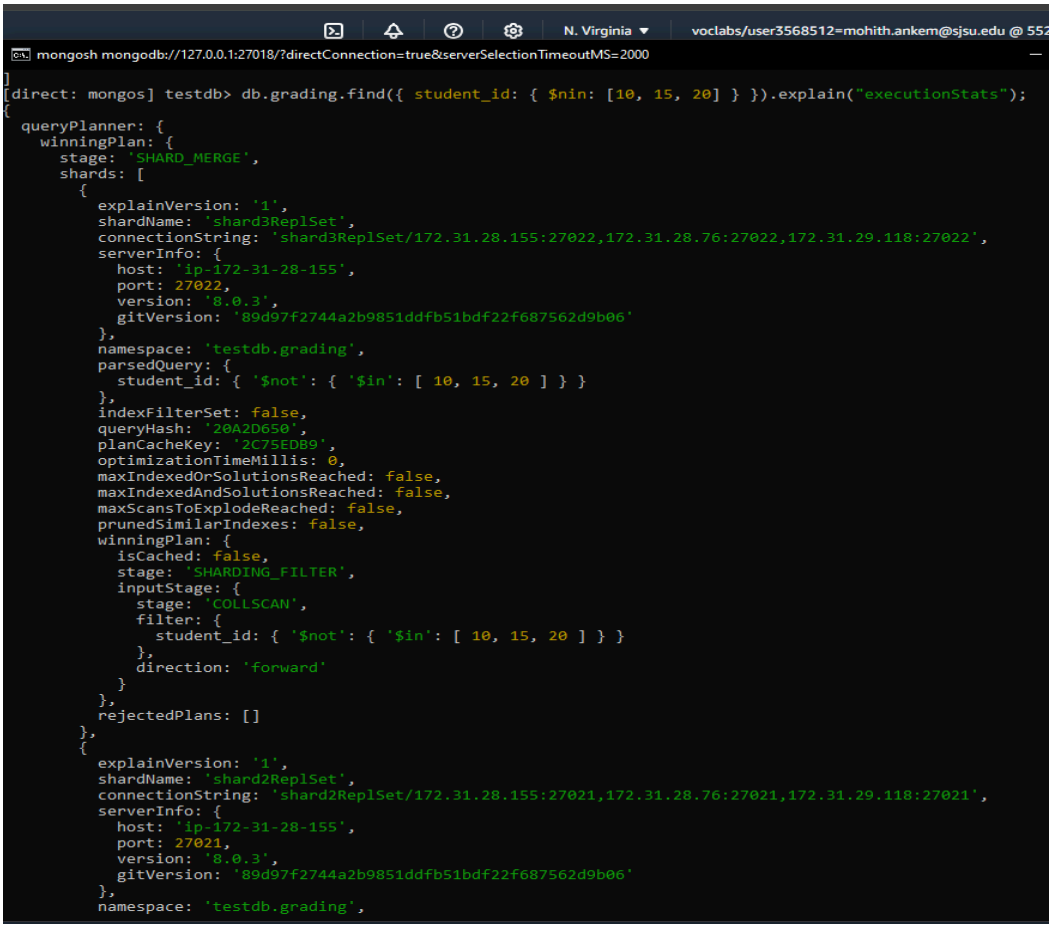
db.grading.find({ student_id: { $nin: [10, 15, 20] } }).limit(3);

This query will retrieve documents where student_id is *not* 10, 15, or 20. By using $nin, MongoDB will filter out any documents where student_id is one of the values in the array [10, 15, 20].

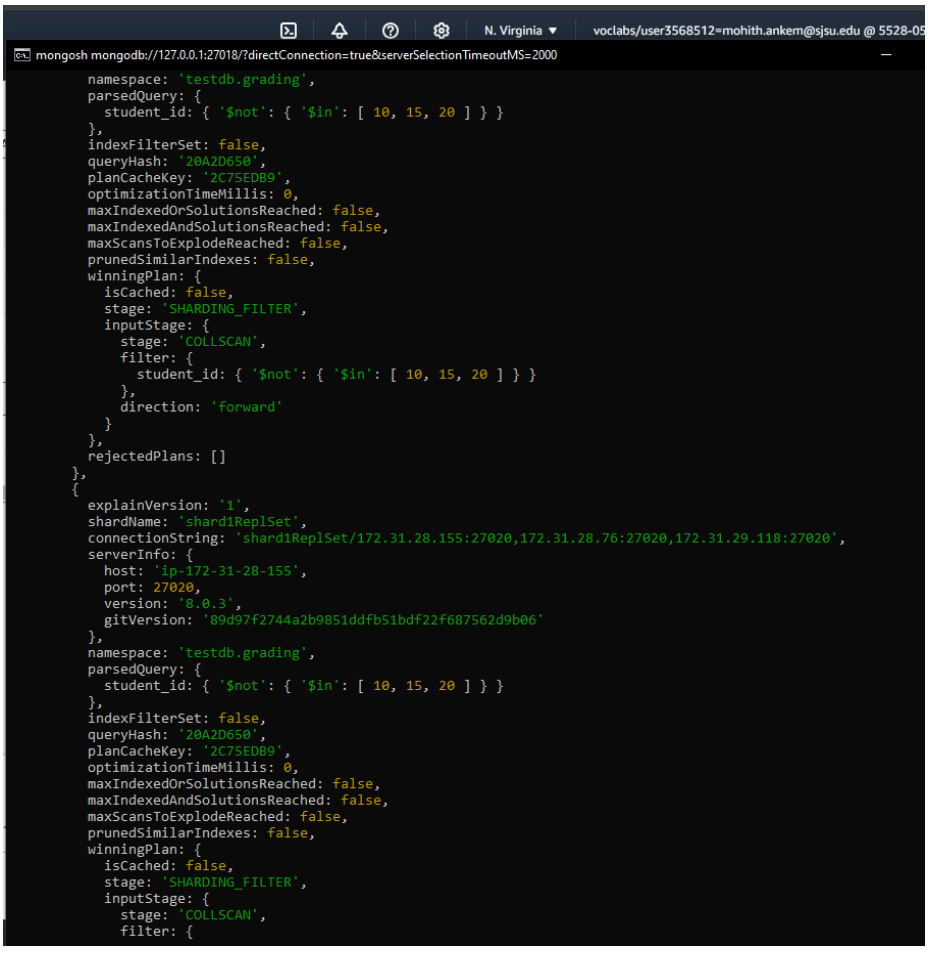


Execution Stats:

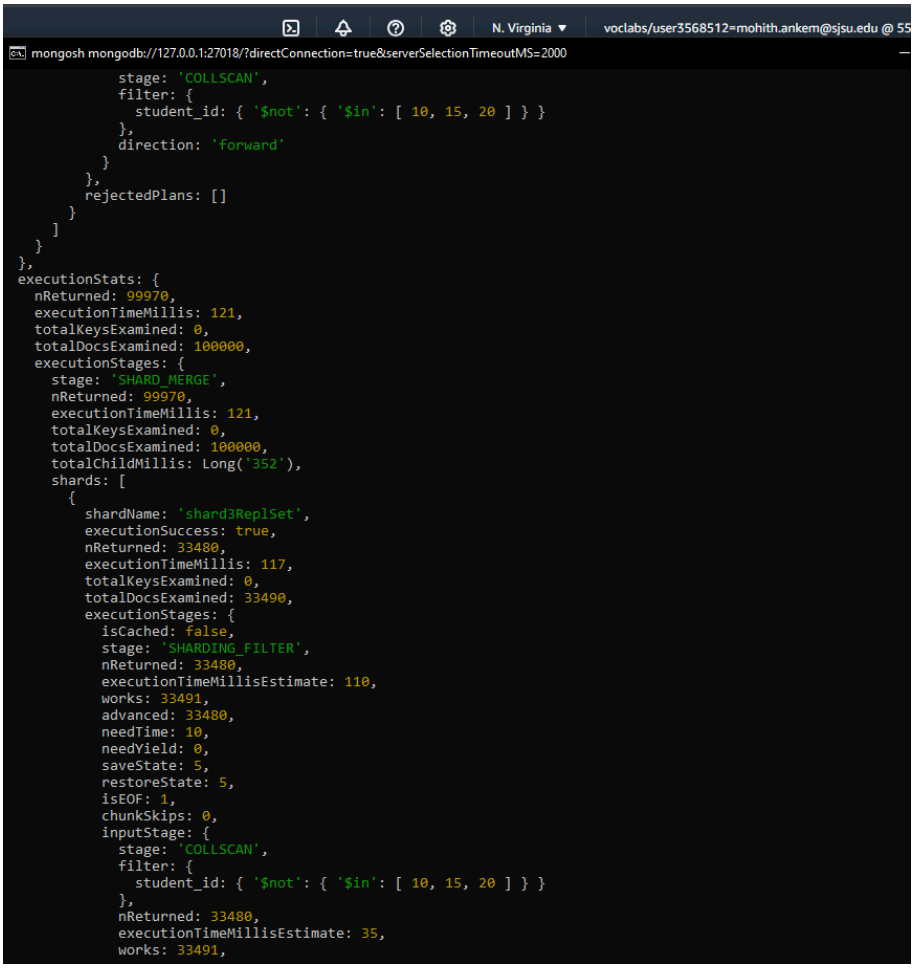
It took a total of 121 ms to run the query and in this case also all 3 shards have returned the results with shard1 returned 33660, shard2 returned 33830 and shard3 returned 33480 docs.

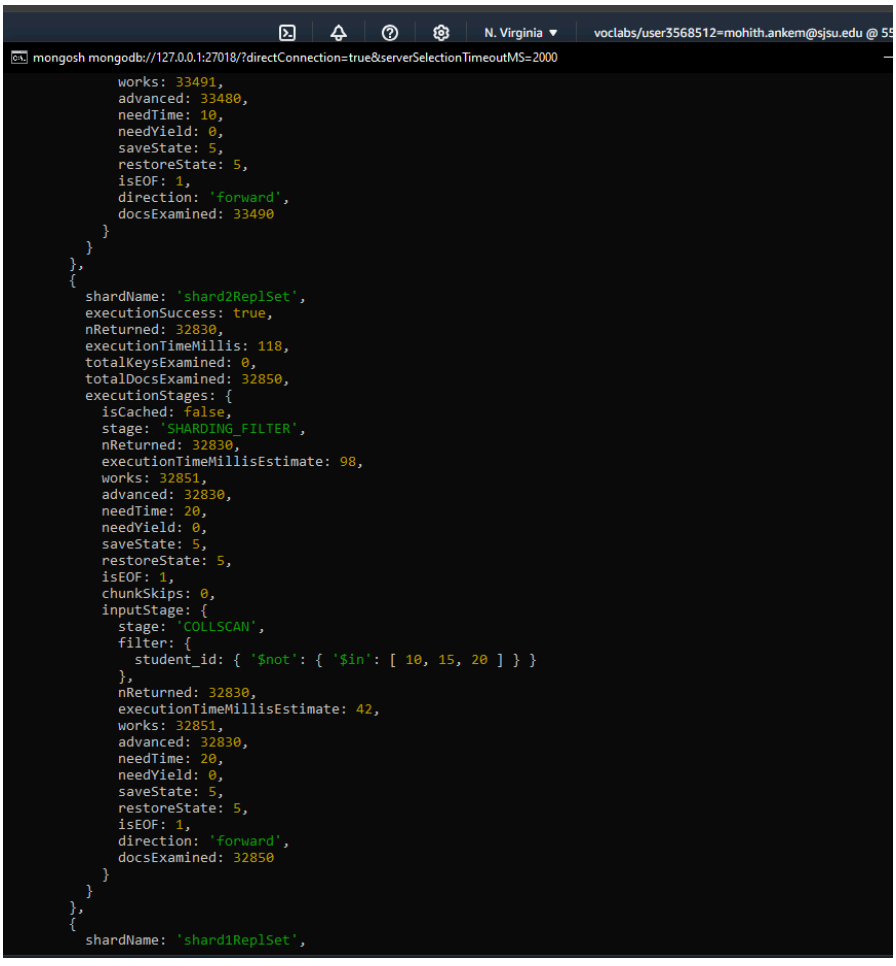


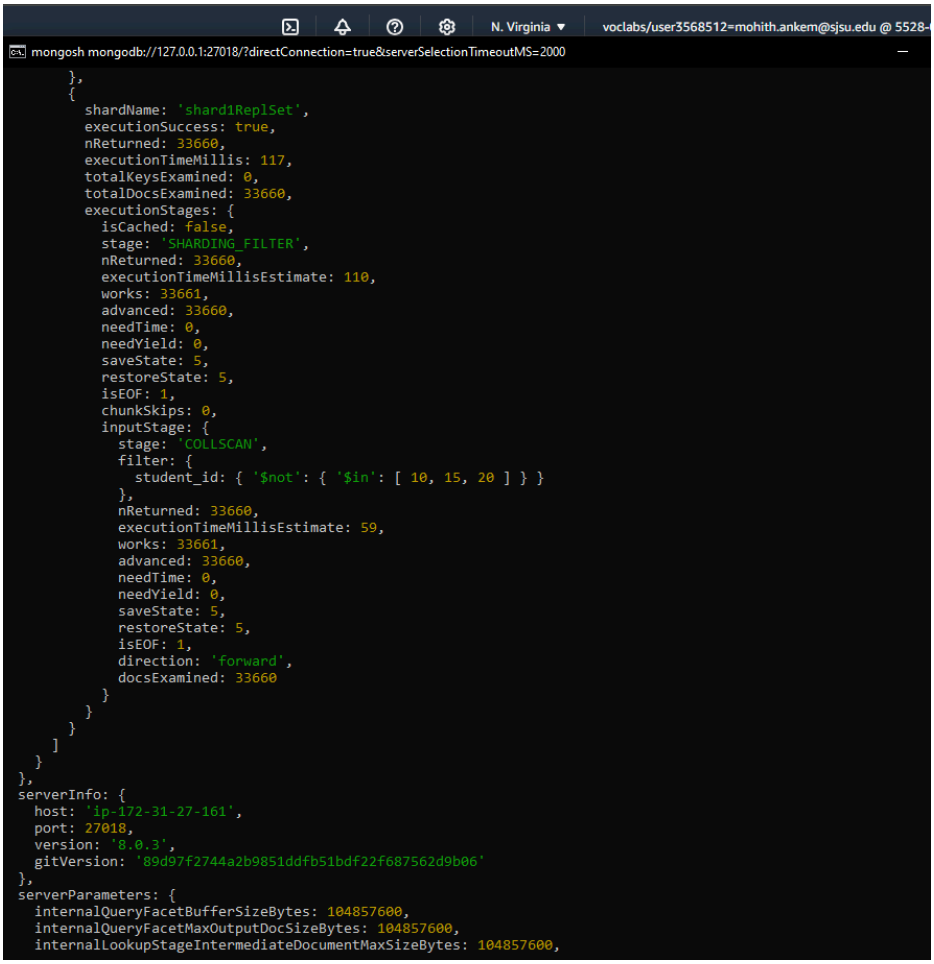

```
]
[direct: mongos] testdb> db.grading.find({ student_id: { $nin: [10, 15, 20] } }).explain("executionStats");
{
  queryPlanner: {
    winningPlan: {
      stage: 'SHARD_MERGE',
      shards: [
        {
          explainVersion: '1',
          shardName: 'shard3ReplSet',
          connectionString: 'shard3ReplSet/172.31.28.155:27022,172.31.28.76:27022,172.31.29.118:27022',
          serverInfo: {
            host: 'ip-172-31-28-155',
            port: 27022,
            version: '8.0.3',
            gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
          },
          namespace: 'testdb.grading',
          parsedQuery: {
            student_id: { '$not': { '$in': [ 10, 15, 20 ] } }
          },
          indexFilterSet: false,
          queryHash: '20A2D650',
          planCacheKey: '2C75EDB9',
          optimizationTimeMillis: 0,
          maxIndexedOrSolutionsReached: false,
          maxIndexedAndSolutionsReached: false,
          maxScansToExplodeReached: false,
          prunedSimilarIndexes: false,
          winningPlan: {
            isCached: false,
            stage: 'SHARDING_FILTER',
            inputStage: {
              stage: 'COLLSCAN',
              filter: {
                student_id: { '$not': { '$in': [ 10, 15, 20 ] } }
              },
              direction: 'forward'
            }
          },
          rejectedPlans: []
        },
        {
          explainVersion: '1',
          shardName: 'shard2ReplSet',
          connectionString: 'shard2ReplSet/172.31.28.155:27021,172.31.28.76:27021,172.31.29.118:27021',
          serverInfo: {
            host: 'ip-172-31-28-155',
            port: 27021,
            version: '8.0.3',
            gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
          },
          namespace: 'testdb.grading',
```

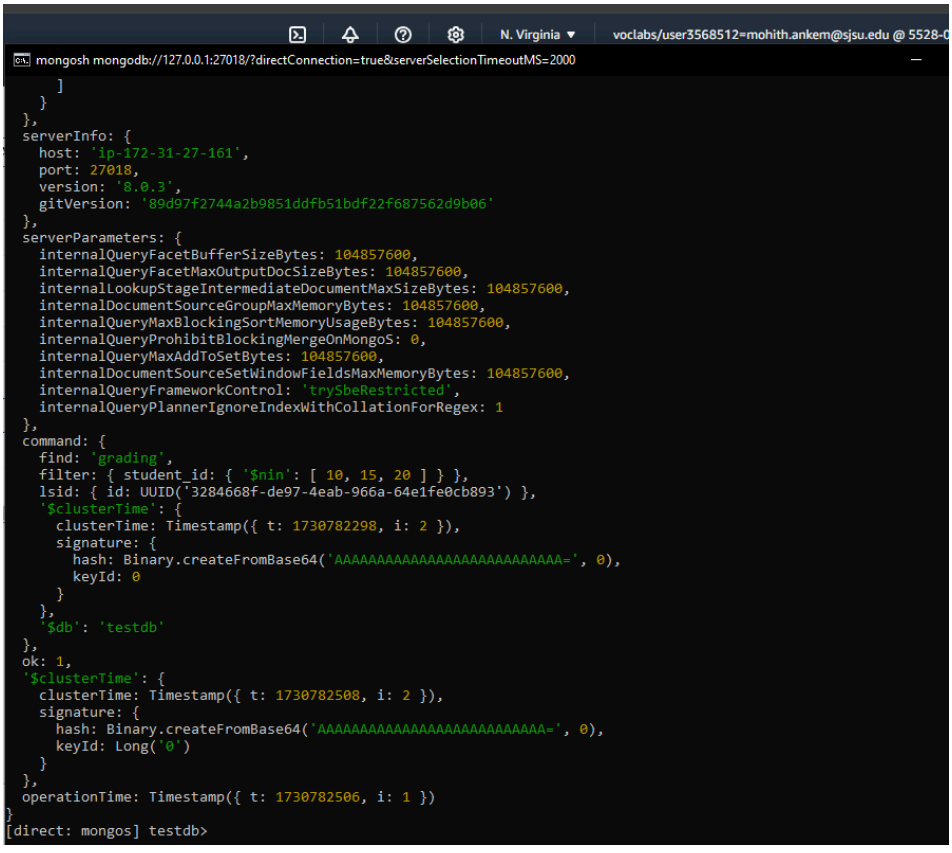



```
          namespace: 'testdb.grading',
          parsedQuery: {
            student_id: { '$not': { '$in': [ 10, 15, 20 ] } }
          },
          indexFilterSet: false,
          queryHash: '20A2D650',
          planCacheKey: '2C75EDB9',
          optimizationTimeMillis: 0,
          maxIndexedOrSolutionsReached: false,
          maxIndexedAndSolutionsReached: false,
          maxScansToExplodeReached: false,
          prunedSimilarIndexes: false,
          winningPlan: {
            isCached: false,
            stage: 'SHARDING_FILTER',
            inputStage: {
              stage: 'COLLSCAN',
              filter: {
                student_id: { '$not': { '$in': [ 10, 15, 20 ] } }
              },
              direction: 'forward'
            }
          },
          rejectedPlans: []
        },
        {
          explainVersion: '1',
          shardName: 'shard1ReplSet',
          connectionString: 'shard1ReplSet/172.31.28.155:27020,172.31.28.76:27020,172.31.29.118:27020',
          serverInfo: {
            host: 'ip-172-31-28-155',
            port: 27020,
            version: '8.0.3',
            gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
          },
          namespace: 'testdb.grading',
          parsedQuery: {
            student_id: { '$not': { '$in': [ 10, 15, 20 ] } }
          },
          indexFilterSet: false,
          queryHash: '20A2D650',
          planCacheKey: '2C75EDB9',
          optimizationTimeMillis: 0,
          maxIndexedOrSolutionsReached: false,
          maxIndexedAndSolutionsReached: false,
          maxScansToExplodeReached: false,
          prunedSimilarIndexes: false,
          winningPlan: {
            isCached: false,
            stage: 'SHARDING_FILTER',
            inputStage: {
              stage: 'COLLSCAN',
              filter: {
```


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 55
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
stage: 'COLLSCAN',
filter: {
student_id: { '$not': { '$in': [10, 15, 20] } }
},
direction: 'forward'
}
},
rejectedPlans: []
}
]
}
},
executionStats: {
nReturned: 99970,
executionTimeMillis: 121,
totalKeysExamined: 0,
totalDocsExamined: 100000,
executionStages: {
stage: 'SHARD_MERGE',
nReturned: 99970,
executionTimeMillis: 121,
totalKeysExamined: 0,
totalDocsExamined: 100000,
totalChildMillis: Long('352'),
shards: [
{
shardName: 'shard3ReplSet',
executionSuccess: true,
nReturned: 33480,
executionTimeMillis: 117,
totalKeysExamined: 0,
totalDocsExamined: 33490,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 33480,
executionTimeMillisEstimate: 110,
works: 33491,
advanced: 33480,
needTime: 10,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
student_id: { '$not': { '$in': [10, 15, 20] } }
},
nReturned: 33480,
executionTimeMillisEstimate: 35,
works: 33491,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 55
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
works: 33491,
advanced: 33480,
needTime: 10,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
direction: 'forward',
docsExamined: 33490
}
}
},
{
shardName: 'shard2ReplSet',
executionSuccess: true,
nReturned: 32830,
executionTimeMillis: 118,
totalKeysExamined: 0,
totalDocsExamined: 32850,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 32830,
executionTimeMillisEstimate: 98,
works: 32851,
advanced: 32830,
needTime: 20,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
student_id: { '$not': { '$in': [10, 15, 20] } }
},
nReturned: 32830,
executionTimeMillisEstimate: 42,
works: 32851,
advanced: 32830,
needTime: 20,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
direction: 'forward',
docsExamined: 32850
}
}
},
{
shardName: 'shard1ReplSet',


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
},
{
shardName: 'shard1ReplSet',
executionSuccess: true,
nReturned: 33660,
executionTimeMillis: 117,
totalKeysExamined: 0,
totalDocsExamined: 33660,
executionStages: {
isCached: false,
stage: 'SHARDING_FILTER',
nReturned: 33660,
executionTimeMillisEstimate: 110,
works: 33661,
advanced: 33660,
needTime: 0,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
filter: {
student_id: { '$not': { '$in': [10, 15, 20] } }
},
nReturned: 33660,
executionTimeMillisEstimate: 59,
works: 33661,
advanced: 33660,
needTime: 0,
needYield: 0,
saveState: 5,
restoreState: 5,
isEOF: 1,
direction: 'forward',
docsExamined: 33660
}
}
}
]
}
},
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
},
serverParameters: {
internalQueryFacetBufferSizeBytes: 104857600,
internalQueryFacetMaxOutputDocSizeBytes: 104857600,
internalLookupStageIntermediateDocumentMaxSizeBytes: 104857600,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
]
}
},
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
},
serverParameters: {
internalQueryFacetBufferSizeBytes: 104857600,
internalQueryFacetMaxOutputDocSizeBytes: 104857600,
internalLookupStageIntermediateDocumentMaxSizeBytes: 104857600,
internalDocumentSourceGroupMaxMemoryBytes: 104857600,
internalQueryMaxBlockingSortMemoryUsageBytes: 104857600,
internalQueryProhibitBlockingMergeOnMongoS: 0,
internalQueryMaxAddToSetBytes: 104857600,
internalDocumentSourceSetWindowFieldsMaxMemoryBytes: 104857600,
internalQueryFrameworkControl: 'trySbeRestricted',
internalQueryPlannerIgnoreIndexWithCollationForRegex: 1
},
command: {
find: 'grading',
filter: { student_id: { '$nin': [10, 15, 20] } },
lsid: { id: UUID('3284668f-de97-4eab-966a-64e1fe0cb893') },
'$clusterTime': {
clusterTime: Timestamp({ t: 1730782298, i: 2 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: 0
}
},
'$db': 'testdb'
},
ok: 1,
'$clusterTime': {
clusterTime: Timestamp({ t: 1730782508, i: 2 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: Long('0')
}
},
operationTime: Timestamp({ t: 1730782506, i: 1 })
}
[direct: mongos] testdb>

4. **A query involving aggregate()**

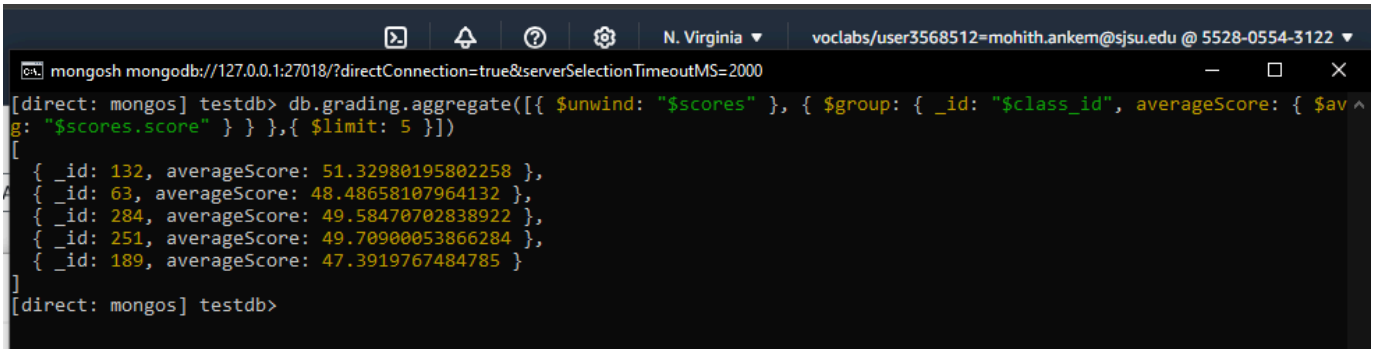
```
db.grading.aggregate([ { $unwind: "$scores" },

  { $group: { _id: "$class_id", averageScore: { $avg: "$scores.score" } } }

]).explain("executionStats");
```

$unwind:

- { $unwind: "$scores" } takes each document's scores array and "unwinds" it. This means that for each element in the scores array, it creates a separate document.
- If a document has multiple scores, each score type (like "exam", "quiz", or "homework") will become its own document, allowing us to handle each score individually.

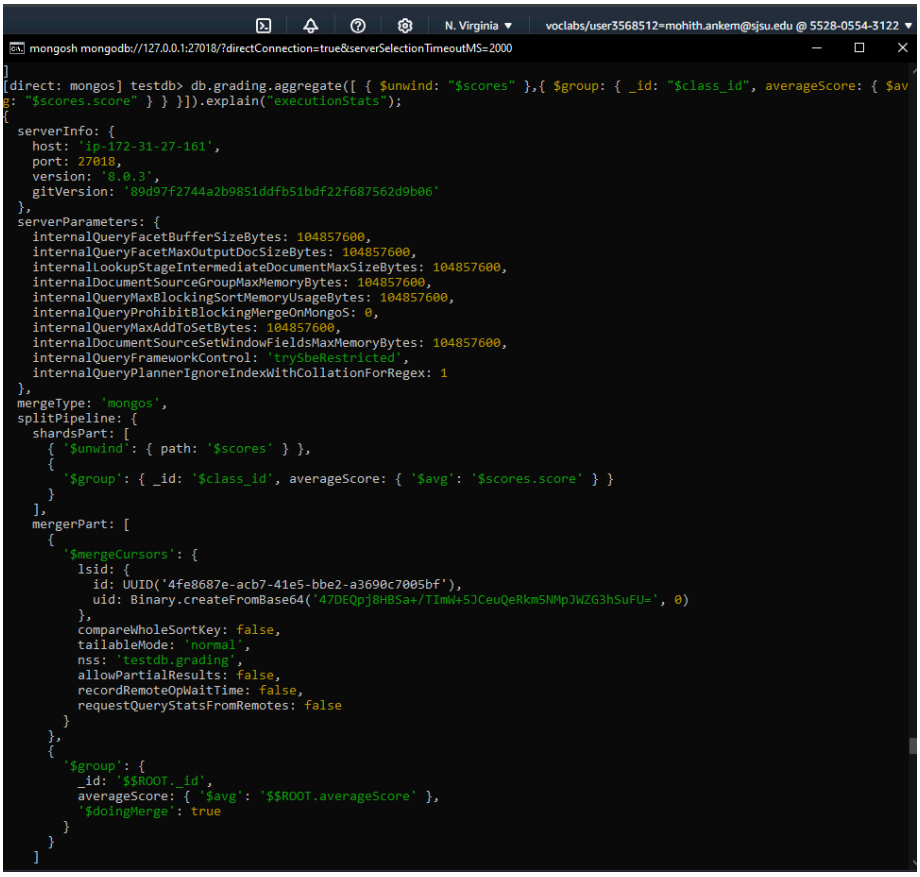
$group:

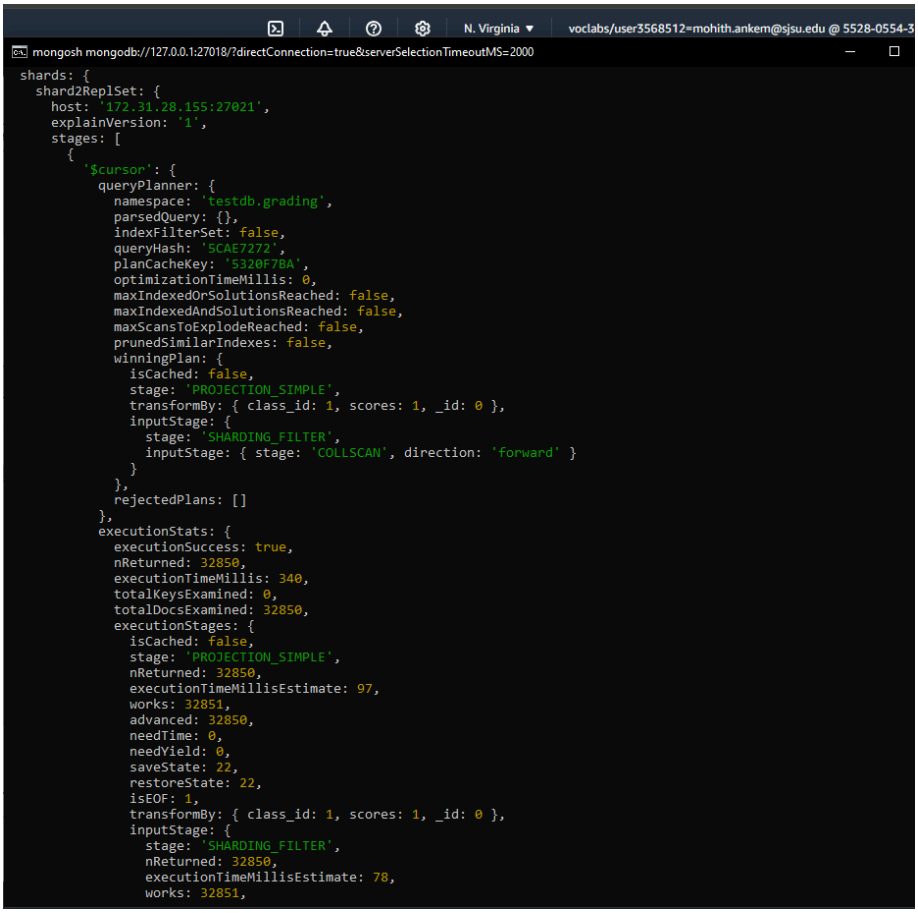
- { $group: { _id: "$class_id", averageScore: { $avg: "$scores.score" } } } groups the unwound documents by class_id and calculates the average of the score field for each class.
- The result is one document per class_id containing the calculated averageScore.

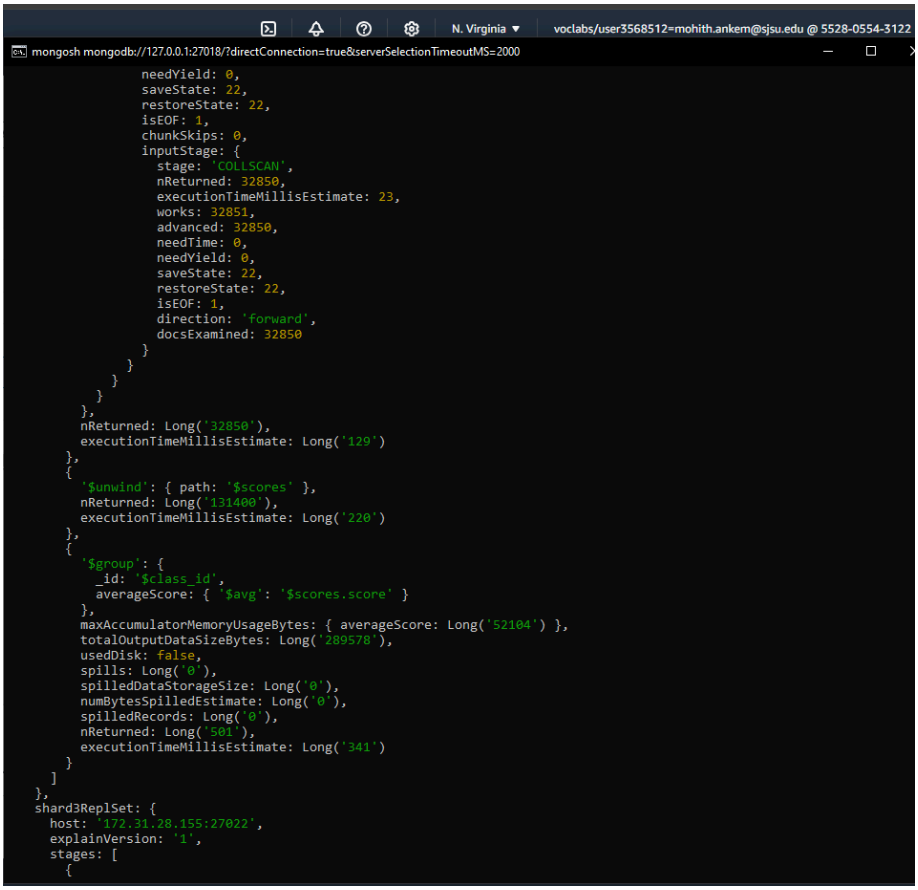


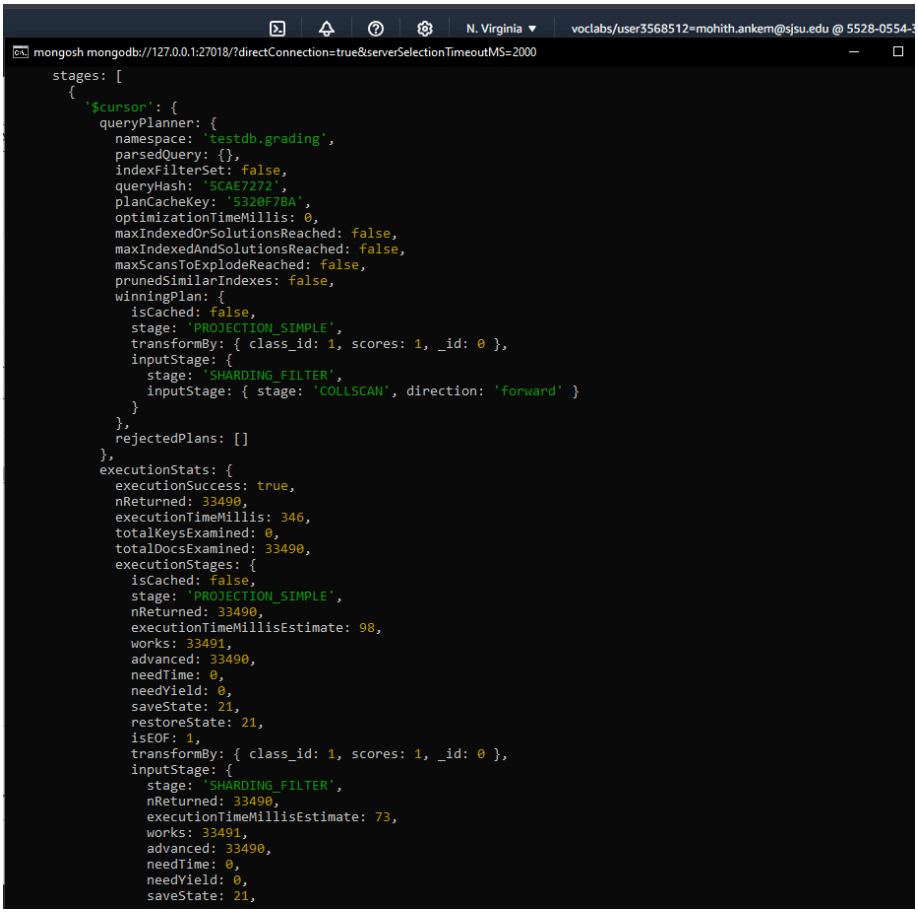
Execution Stats:

It took a total of 346 ms to run the query and in this case also all 3 shards have returned the results with shard1 returned 33660, shard2 returned 33850 and shard3 returned 33490 docs.

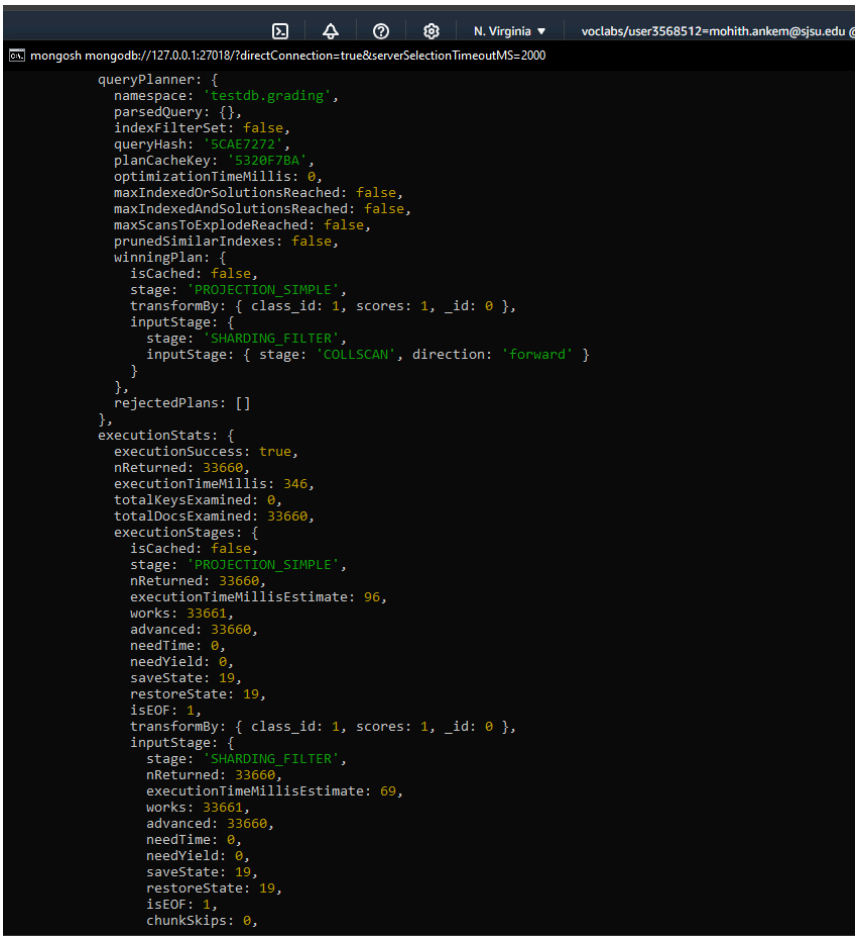

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
[direct: mongos] testdb> db.grading.aggregate([{ $unwind: "$scores" },{ $group: { _id: "$class_id", averageScore: { $av
g: "$scores.score" } } }]).explain("executionStats");
{
serverInfo: {
host: 'ip-172-31-27-161',
port: 27018,
version: '8.0.3',
gitVersion: '89d97f2744a2b9851ddfb51bdf22f687562d9b06'
},
serverParameters: {
internalQueryFacetBufferSizeBytes: 104857600,
internalQueryFacetMaxOutputDocSizeBytes: 104857600,
internalLookupStageIntermediateDocumentMaxSizeBytes: 104857600,
internalDocumentSourceGroupMaxMemoryBytes: 104857600,
internalQueryMaxBlockingSortMemoryUsageBytes: 104857600,
internalQueryProhibitBlockingMergeOnMongoS: 0,
internalQueryMaxAddToSetBytes: 104857600,
internalDocumentSourceSetWindowFieldsMaxMemoryBytes: 104857600,
internalQueryFrameworkControl: 'trySbeRestricted',
internalQueryPlannerIgnoreIndexWithCollationForRegex: 1
},
mergeType: 'mongos',
splitPipeline: {
shardsPart: [
{ '$unwind': { path: '$scores' } },
{
'$group': { _id: '$class_id', averageScore: { '$avg': '$scores.score' } }
}
],
mergerPart: [
{
'$mergeCursors': {
lsid: {
id: UUID('4fe8687e-acb7-41e5-bbe2-a3690c7005bf'),
uid: Binary.createFromBase64('47DEQpj8HBSa+/TImW+5JCeuQeRkm5NMpJWZG3hSuFU=', 0)
},
compareWholeSortKey: false,
tailableMode: 'normal',
nss: 'testdb.grading',
allowPartialResults: false,
recordRemoteOpWaitTime: false,
requestQueryStatsFromRemotes: false
}
},
{
'$group': {
_id: '$$ROOT._id',
averageScore: { '$avg': '$$ROOT.averageScore' },
'$doingMerge': true
}
}
]

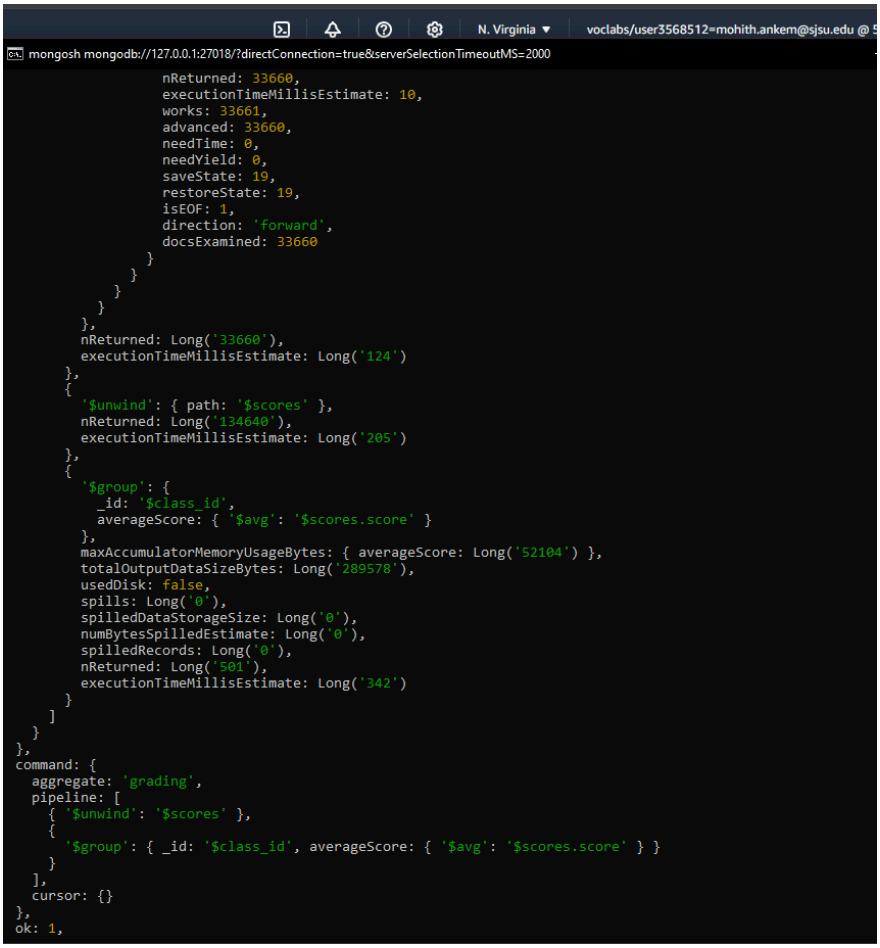

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
shards: {
shard2ReplSet: {
host: '172.31.28.155:27021',
explainVersion: '1',
stages: [
{
'$cursor': {
queryPlanner: {
namespace: 'testdb.grading',
parsedQuery: {},
indexFilterSet: false,
queryHash: '5CAE7272',
planCacheKey: '5320F7BA',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
inputStage: { stage: 'COLLSCAN', direction: 'forward' }
}
},
rejectedPlans: []
},
executionStats: {
executionSuccess: true,
nReturned: 32850,
executionTimeMillis: 340,
totalKeysExamined: 0,
totalDocsExamined: 32850,
executionStages: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
nReturned: 32850,
executionTimeMillisEstimate: 97,
works: 32851,
advanced: 32850,
needTime: 0,
needYield: 0,
saveState: 22,
restoreState: 22,
isEOF: 1,
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
nReturned: 32850,
executionTimeMillisEstimate: 78,
works: 32851,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
needYield: 0,
saveState: 22,
restoreState: 22,
isEOF: 1,
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
nReturned: 32850,
executionTimeMillisEstimate: 23,
works: 32851,
advanced: 32850,
needTime: 0,
needYield: 0,
saveState: 22,
restoreState: 22,
isEOF: 1,
direction: 'forward',
docsExamined: 32850
nReturned: Long('32850'),
executionTimeMillisEstimate: Long('129')
'$unwind': { path: '$scores' },
nReturned: Long('131400'),
executionTimeMillisEstimate: Long('220')
'$group': {
_id: '$class_id',
averageScore: { '$avg': '$scores.score' }
maxAccumulatorMemoryUsageBytes: { averageScore: Long('52104') },
totalOutputDataSizeBytes: Long('289578'),
usedDisk: false,
spills: Long('0'),
spilledDataStorageSize: Long('0'),
numBytesSpilledEstimate: Long('0'),
spilledRecords: Long('0'),
nReturned: Long('501'),
executionTimeMillisEstimate: Long('341')
shard3ReplSet: {
host: '172.31.28.155:27022',
explainVersion: '1',
stages: [


N. Virginia
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
stages: [
'$cursor': {
queryPlanner: {
namespace: 'testdb.grading',
parsedQuery: {},
indexFilterSet: false,
queryHash: '5CAE7272',
planCacheKey: '5320F7BA',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
inputStage: { stage: 'COLLSCAN', direction: 'forward' }
rejectedPlans: []
executionStats: {
executionSuccess: true,
nReturned: 33490,
executionTimeMillis: 346,
totalKeysExamined: 0,
totalDocsExamined: 33490,
executionStages: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
nReturned: 33490,
executionTimeMillisEstimate: 98,
works: 33491,
advanced: 33490,
needTime: 0,
needYield: 0,
saveState: 21,
restoreState: 21,
isEOF: 1,
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
nReturned: 33490,
executionTimeMillisEstimate: 73,
works: 33491,
advanced: 33490,
needTime: 0,
needYield: 0,
saveState: 21,


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
chunkSkips: 0,
inputStage: {
stage: 'COLLSCAN',
nReturned: 33490,
executionTimeMillisEstimate: 7,
works: 33491,
advanced: 33490,
needTime: 0,
needYield: 0,
saveState: 21,
restoreState: 21,
isEOF: 1,
direction: 'forward',
docsExamined: 33490
nReturned: Long('33490'),
executionTimeMillisEstimate: Long('150')
'$unwind': { path: '$scores' },
nReturned: Long('133960'),
executionTimeMillisEstimate: Long('257')
'$group': {
_id: '$class_id',
averageScore: { '$avg': '$scores.score' }
maxAccumulatorMemoryUsageBytes: { averageScore: Long('52104') },
totalOutputDataSizeBytes: Long('289578'),
usedDisk: false,
spills: Long('0'),
spilledDataStorageSize: Long('0'),
numBytesSpilledEstimate: Long('0'),
spilledRecords: Long('0'),
nReturned: Long('501'),
executionTimeMillisEstimate: Long('341')
shard1ReplSet: {
host: '172.31.28.155:27020',
explainVersion: '1',
stages: [
'$cursor': {
queryPlanner: {
namespace: 'testdb.grading',
parsedQuery: {},


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
queryPlanner: {
namespace: 'testdb.grading',
parsedQuery: {},
indexFilterSet: false,
queryHash: '5CAE7272',
planCacheKey: '5320F7BA',
optimizationTimeMillis: 0,
maxIndexedOrSolutionsReached: false,
maxIndexedAndSolutionsReached: false,
maxScansToExplodeReached: false,
prunedSimilarIndexes: false,
winningPlan: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
inputStage: { stage: 'COLLSCAN', direction: 'forward' }
rejectedPlans: []
executionStats: {
executionSuccess: true,
nReturned: 33660,
executionTimeMillis: 346,
totalKeysExamined: 0,
totalDocsExamined: 33660,
executionStages: {
isCached: false,
stage: 'PROJECTION_SIMPLE',
nReturned: 33660,
executionTimeMillisEstimate: 96,
works: 33661,
advanced: 33660,
needTime: 0,
needYield: 0,
saveState: 19,
restoreState: 19,
isEOF: 1,
transformBy: { class_id: 1, scores: 1, _id: 0 },
inputStage: {
stage: 'SHARDING_FILTER',
nReturned: 33660,
executionTimeMillisEstimate: 69,
works: 33661,
advanced: 33660,
needTime: 0,
needYield: 0,
saveState: 19,
restoreState: 19,
isEOF: 1,
chunkSkips: 0,



5. **A update**

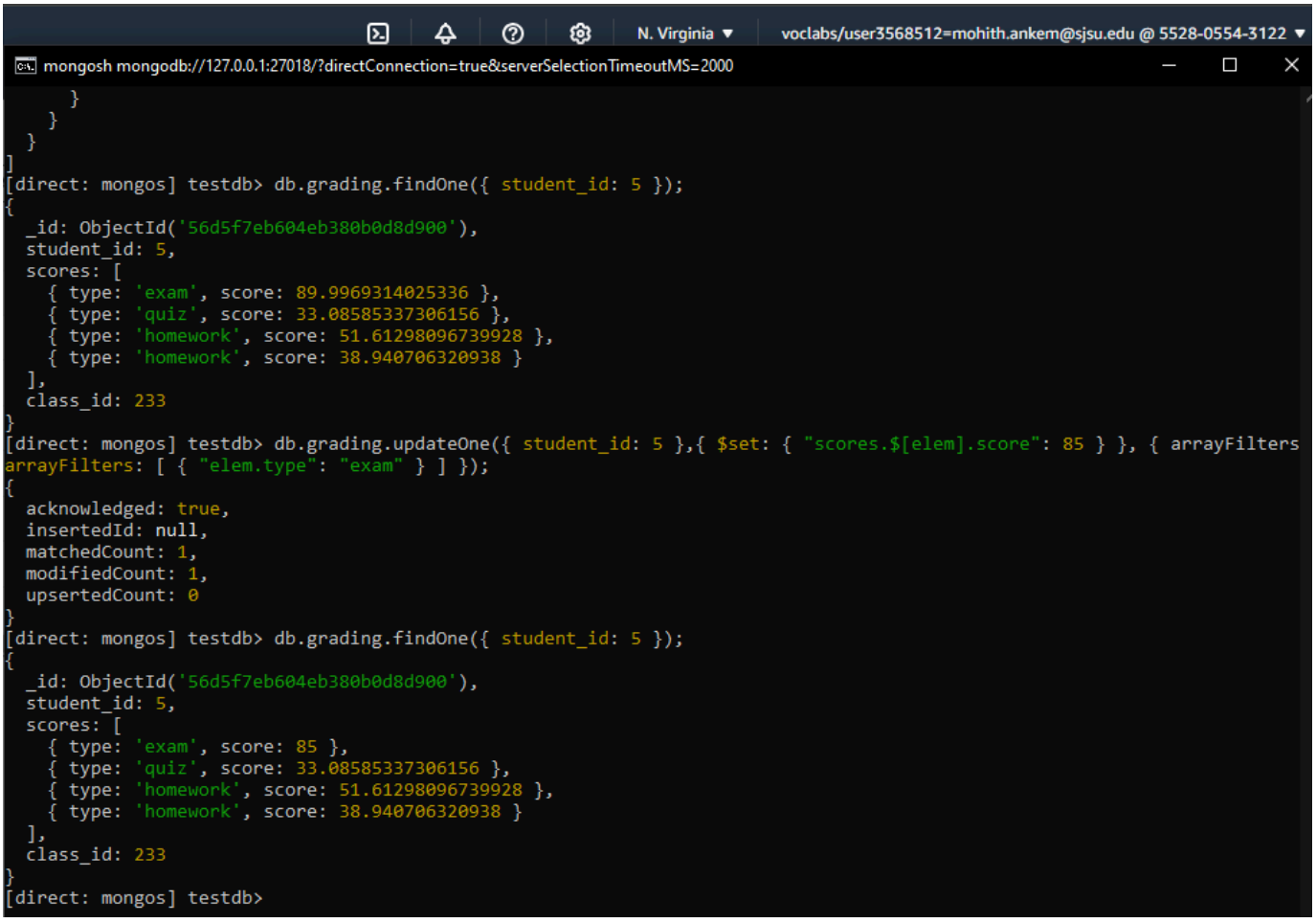
db.grading.updateOne({ student_id: 5 }, { $set: { "scores.$[elem].score": 85 } },

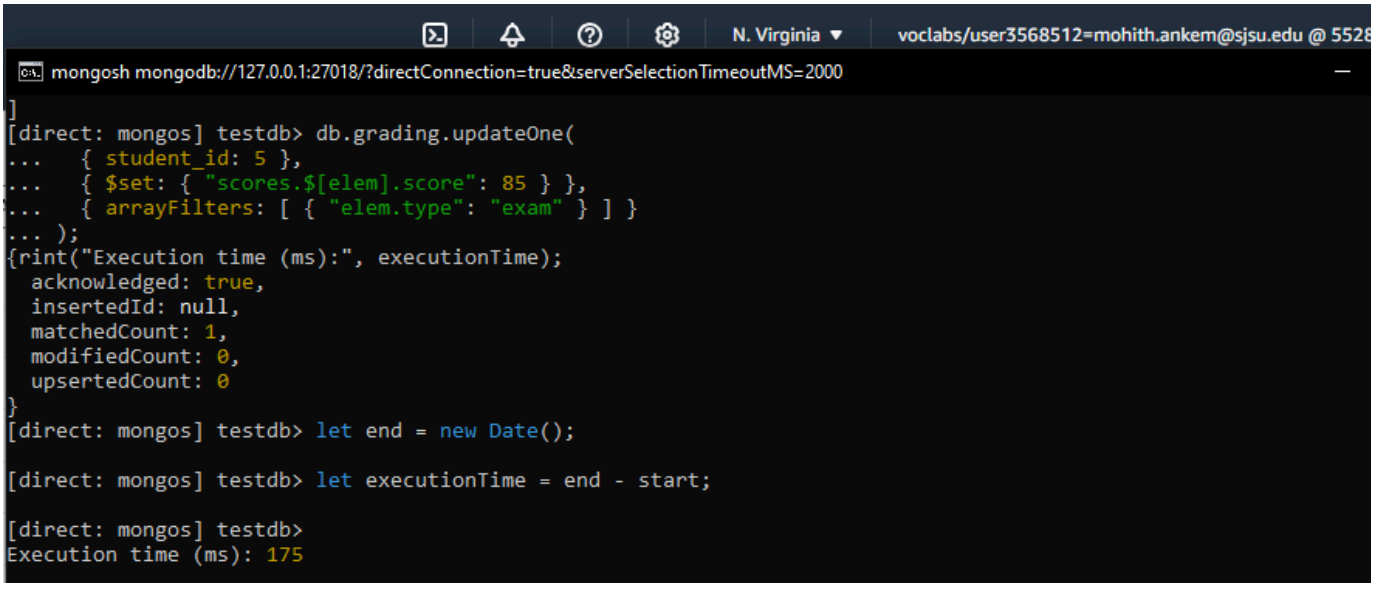
{ arrayFilters: [{ "elem.type": "exam" }] });

This command updates the score field to 85 for the score entry with "type": "exam" in the scores array for the document with student_id 5.

I have used commands to update the records and attached the screenshots before and after running the command for updating the record.

- It took 175 ms for the command to execute and as the explainStats command doesn't work for update query the 3 shards will serve the query.

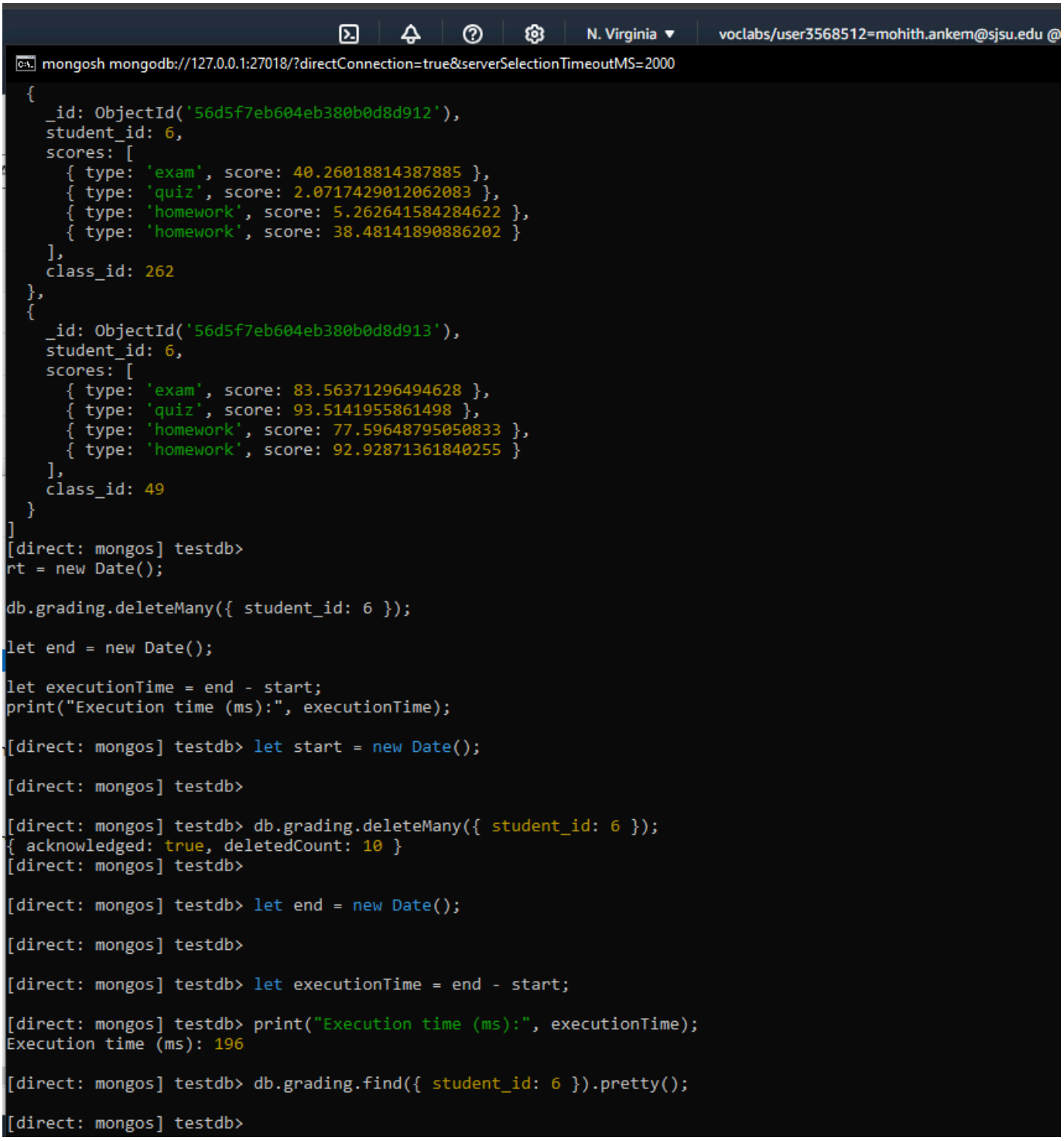

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-3122
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
}
}
}
]
[direct: mongos] testdb> db.grading.findOne({ student_id: 5 });
{
_id: ObjectId('56d5f7eb604eb380b0d8d900'),
student_id: 5,
scores: [
{ type: 'exam', score: 89.9969314025336 },
{ type: 'quiz', score: 33.08585337306156 },
{ type: 'homework', score: 51.61298096739928 },
{ type: 'homework', score: 38.940706320938 }
],
class_id: 233
}
[direct: mongos] testdb> db.grading.updateOne({ student_id: 5 },{ $set: { "scores.$[elem].score": 85 } }, { arrayFilters
arrayFilters: [{ "elem.type": "exam" }] });
{
acknowledged: true,
insertedId: null,
matchedCount: 1,
modifiedCount: 1,
upsertedCount: 0
}
[direct: mongos] testdb> db.grading.findOne({ student_id: 5 });
{
_id: ObjectId('56d5f7eb604eb380b0d8d900'),
student_id: 5,
scores: [
{ type: 'exam', score: 85 },
{ type: 'quiz', score: 33.08585337306156 },
{ type: 'homework', score: 51.61298096739928 },
{ type: 'homework', score: 38.940706320938 }
],
class_id: 233
}
[direct: mongos] testdb>


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528
mongosh mongodb://127.0.0.1:27018/?directConnection=true&serverSelectionTimeoutMS=2000
]
[direct: mongos] testdb> db.grading.updateOne(
... { student_id: 5 },
... { $set: { "scores.$[elem].score": 85 } },
... { arrayFilters: [{ "elem.type": "exam" }] }
...);
{rint("Execution time (ms):", executionTime);
acknowledged: true,
insertedId: null,
matchedCount: 1,
modifiedCount: 0,
upsertedCount: 0
}
[direct: mongos] testdb> let end = new Date();
[direct: mongos] testdb> let executionTime = end - start;
[direct: mongos] testdb>
Execution time (ms): 175

6. A delete

For Delete i have used **deleteMany()** Command to delete all the records of docs that have student_id as ‘6’ and have shown before and after deleting the records.

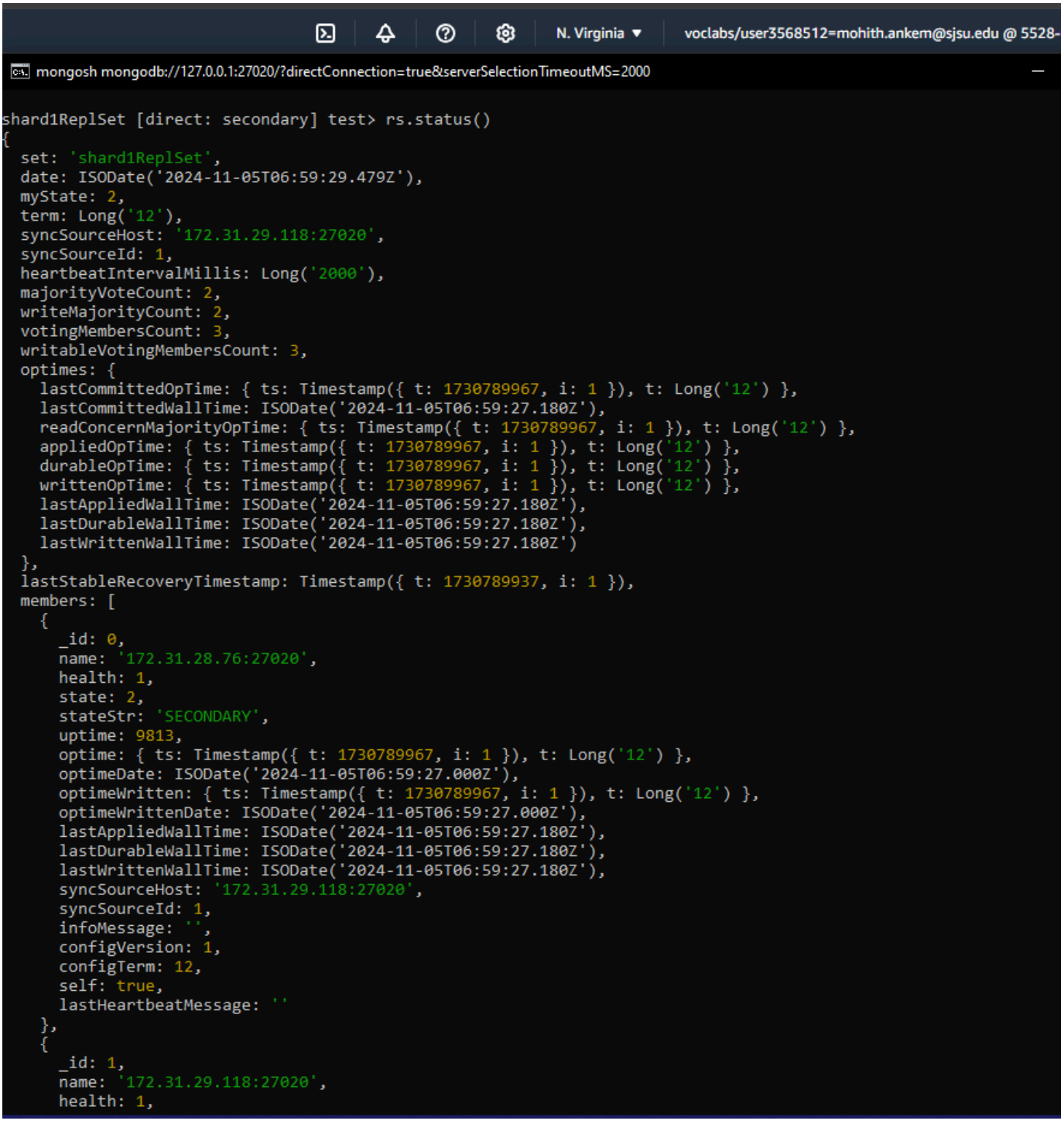
db.grading.deleteMany({ student_id: 6 });

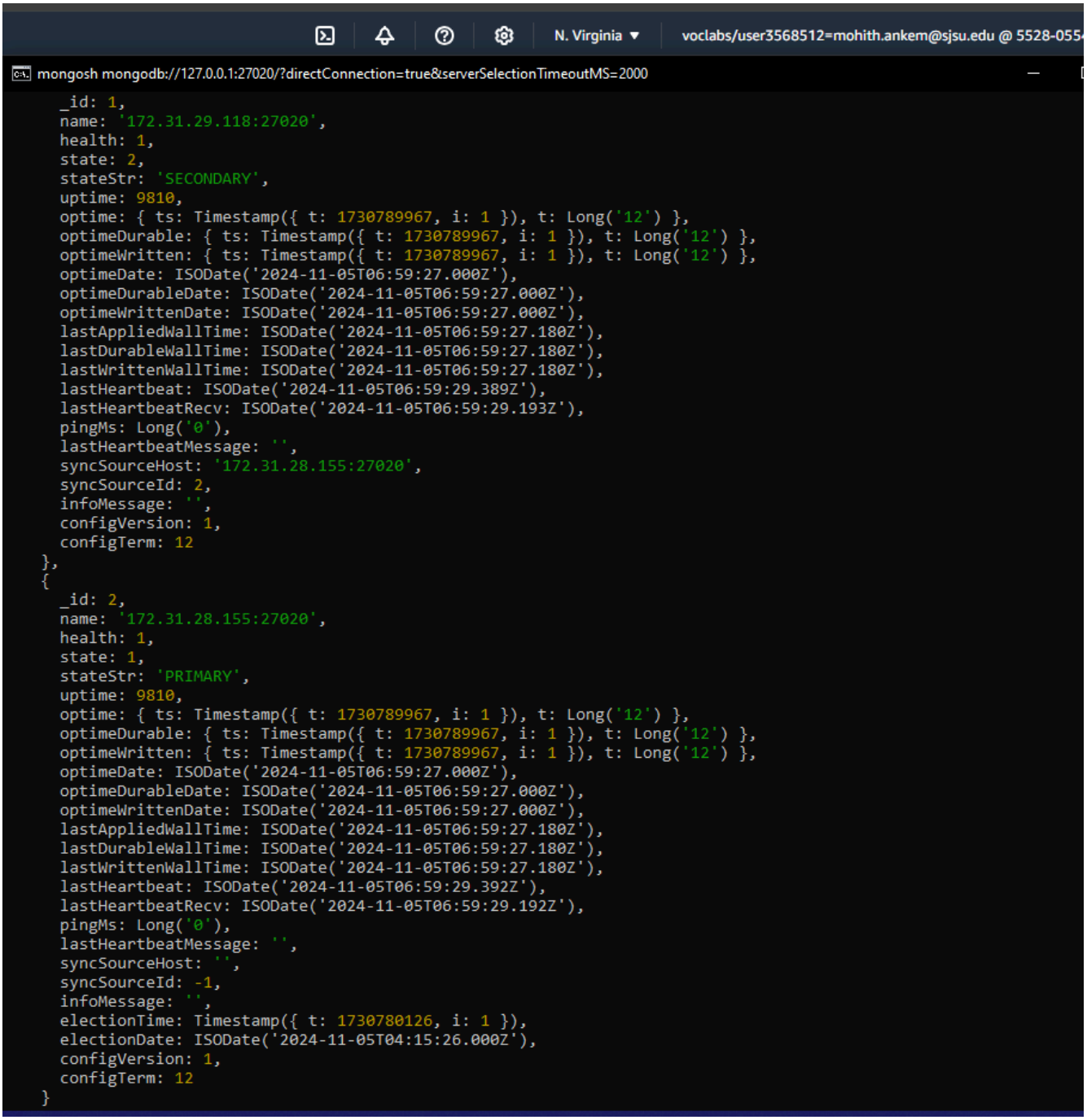


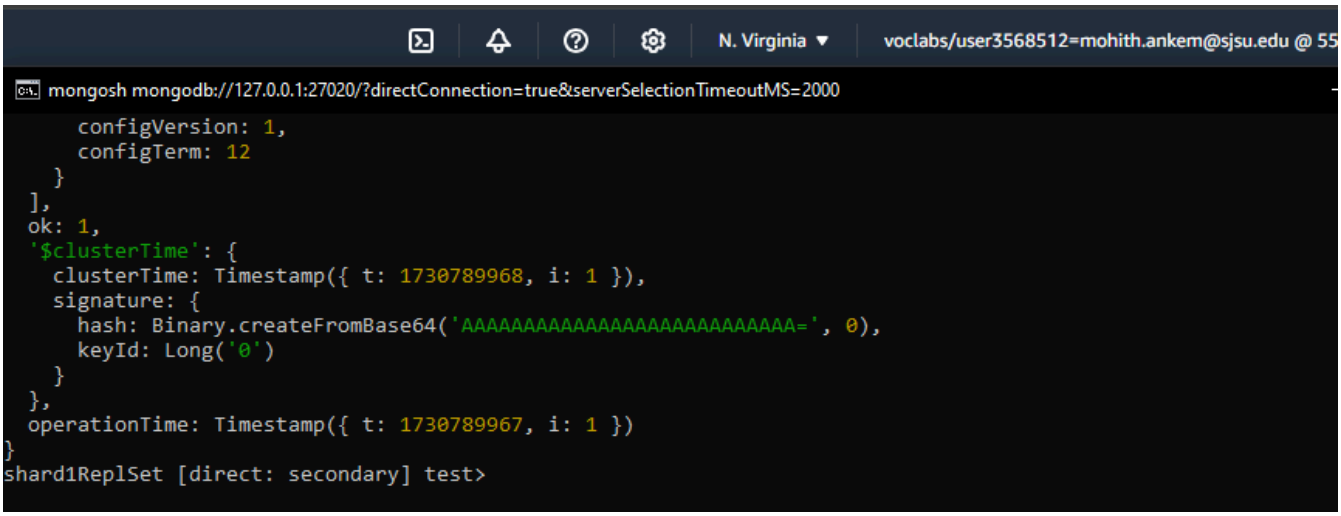
Deleted all the records that has student_id as ‘6’ and it took 196 ms to execute the query in which all 3 shards took part in this execution

13. Show shards are replicated using rs.status()

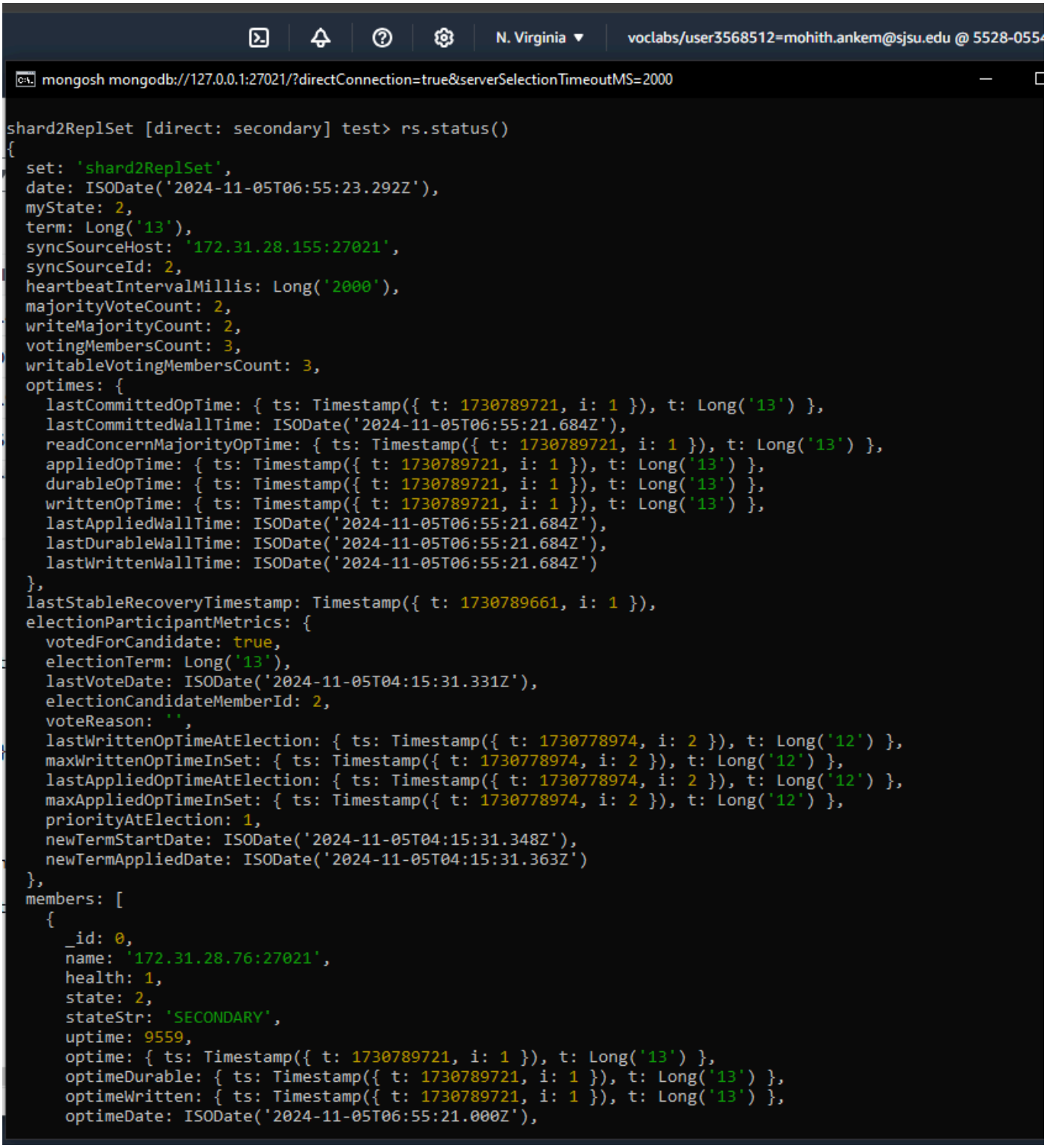
rs.status() for Shard1:

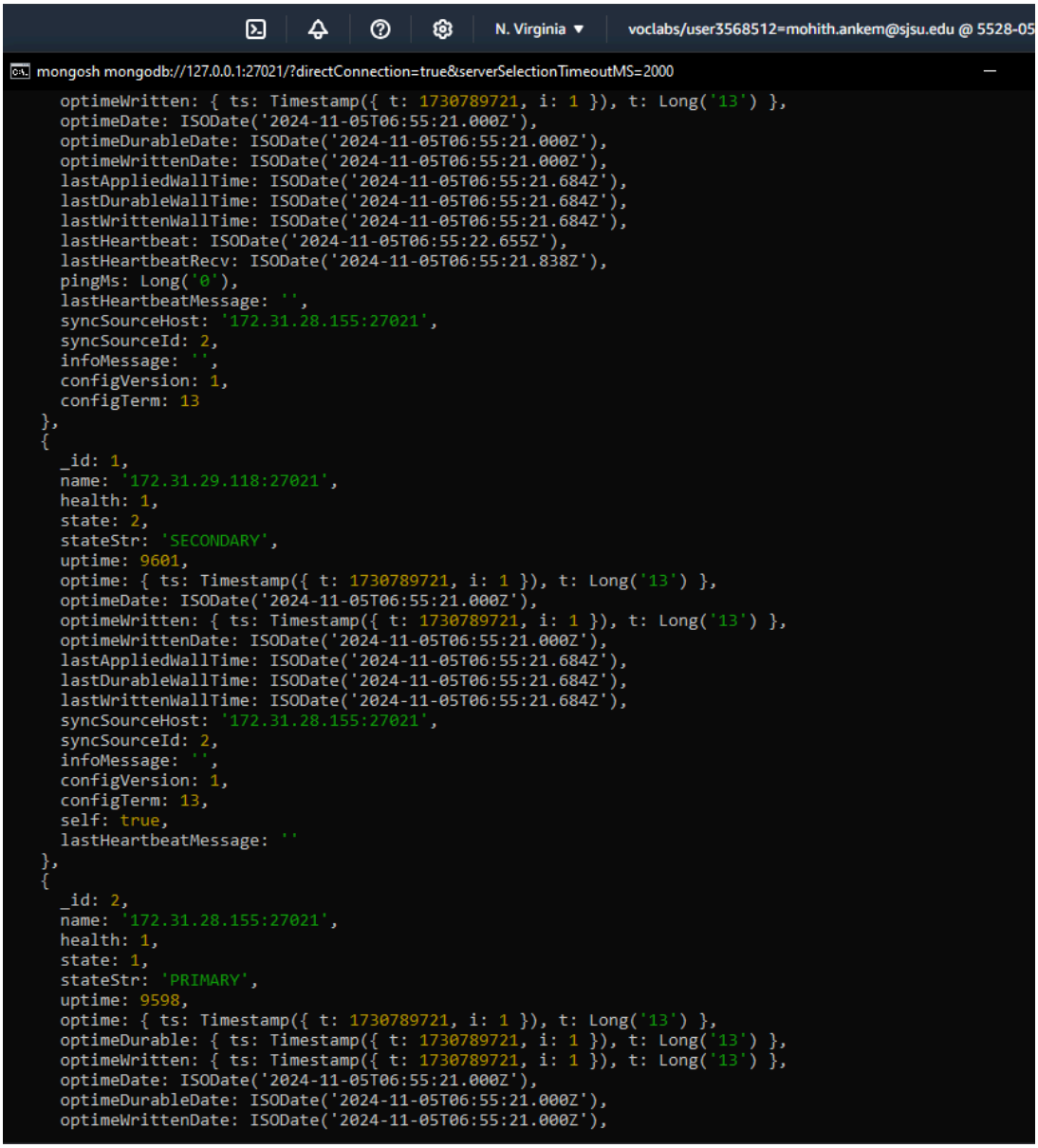



N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-055
mongosh mongodb://127.0.0.1:27020/?directConnection=true&serverSelectionTimeoutMS=2000
_id: 1,
name: '172.31.29.118:27020',
health: 1,
state: 2,
stateStr: 'SECONDARY',
uptime: 9810,
optime: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeDurable: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeWritten: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeDate: ISODate('2024-11-05T06:59:27.000Z'),
optimeDurableDate: ISODate('2024-11-05T06:59:27.000Z'),
optimeWrittenDate: ISODate('2024-11-05T06:59:27.000Z'),
lastAppliedWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastDurableWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastWrittenWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastHeartbeat: ISODate('2024-11-05T06:59:29.389Z'),
lastHeartbeatRecv: ISODate('2024-11-05T06:59:29.193Z'),
pingMs: Long('0'),
lastHeartbeatMessage: '',
syncSourceHost: '172.31.28.155:27020',
syncSourceId: 2,
infoMessage: '',
configVersion: 1,
configTerm: 12
},
{
_id: 2,
name: '172.31.28.155:27020',
health: 1,
state: 1,
stateStr: 'PRIMARY',
uptime: 9810,
optime: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeDurable: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeWritten: { ts: Timestamp({ t: 1730789967, i: 1 }), t: Long('12') },
optimeDate: ISODate('2024-11-05T06:59:27.000Z'),
optimeDurableDate: ISODate('2024-11-05T06:59:27.000Z'),
optimeWrittenDate: ISODate('2024-11-05T06:59:27.000Z'),
lastAppliedWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastDurableWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastWrittenWallTime: ISODate('2024-11-05T06:59:27.180Z'),
lastHeartbeat: ISODate('2024-11-05T06:59:29.392Z'),
lastHeartbeatRecv: ISODate('2024-11-05T06:59:29.192Z'),
pingMs: Long('0'),
lastHeartbeatMessage: '',
syncSourceHost: '',
syncSourceId: -1,
infoMessage: '',
electionTime: Timestamp({ t: 1730780126, i: 1 }),
electionDate: ISODate('2024-11-05T04:15:26.000Z'),
configVersion: 1,
configTerm: 12
}

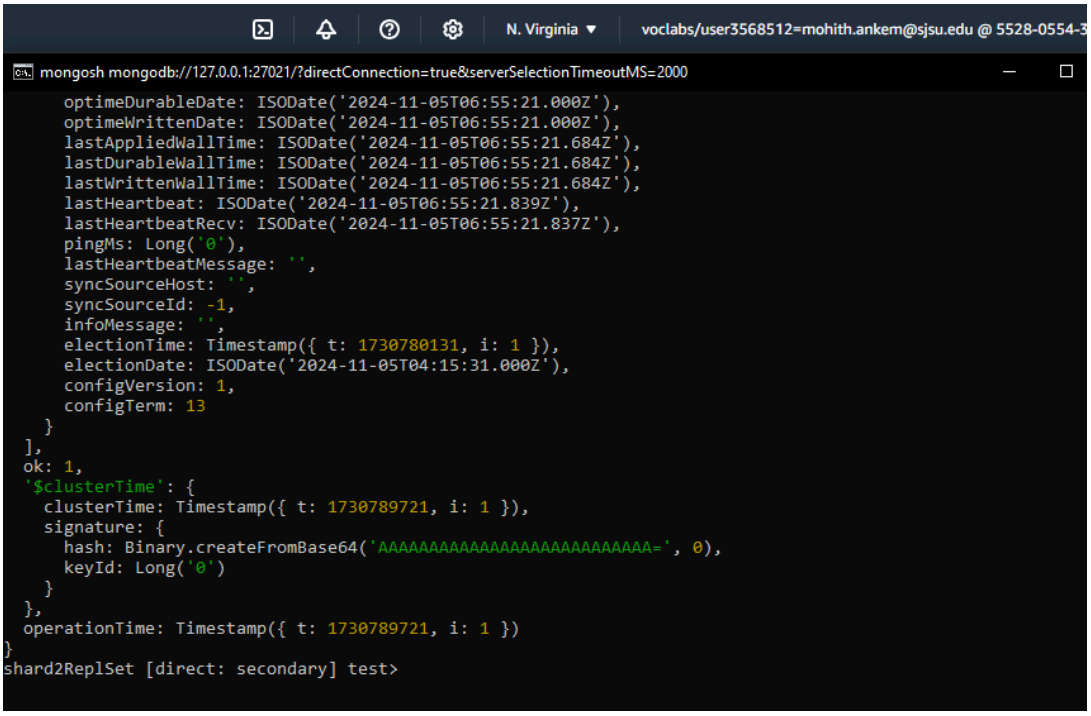

N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 55
mongosh mongodb://127.0.0.1:27020/?directConnection=true&serverSelectionTimeoutMS=2000
configVersion: 1,
configTerm: 12
}
],
ok: 1,
'$clusterTime': {
clusterTime: Timestamp({ t: 1730789968, i: 1 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: Long('0')
}
},
operationTime: Timestamp({ t: 1730789967, i: 1 })
}
shard1ReplSet [direct: secondary] test>

rs.status() for Shrad2:



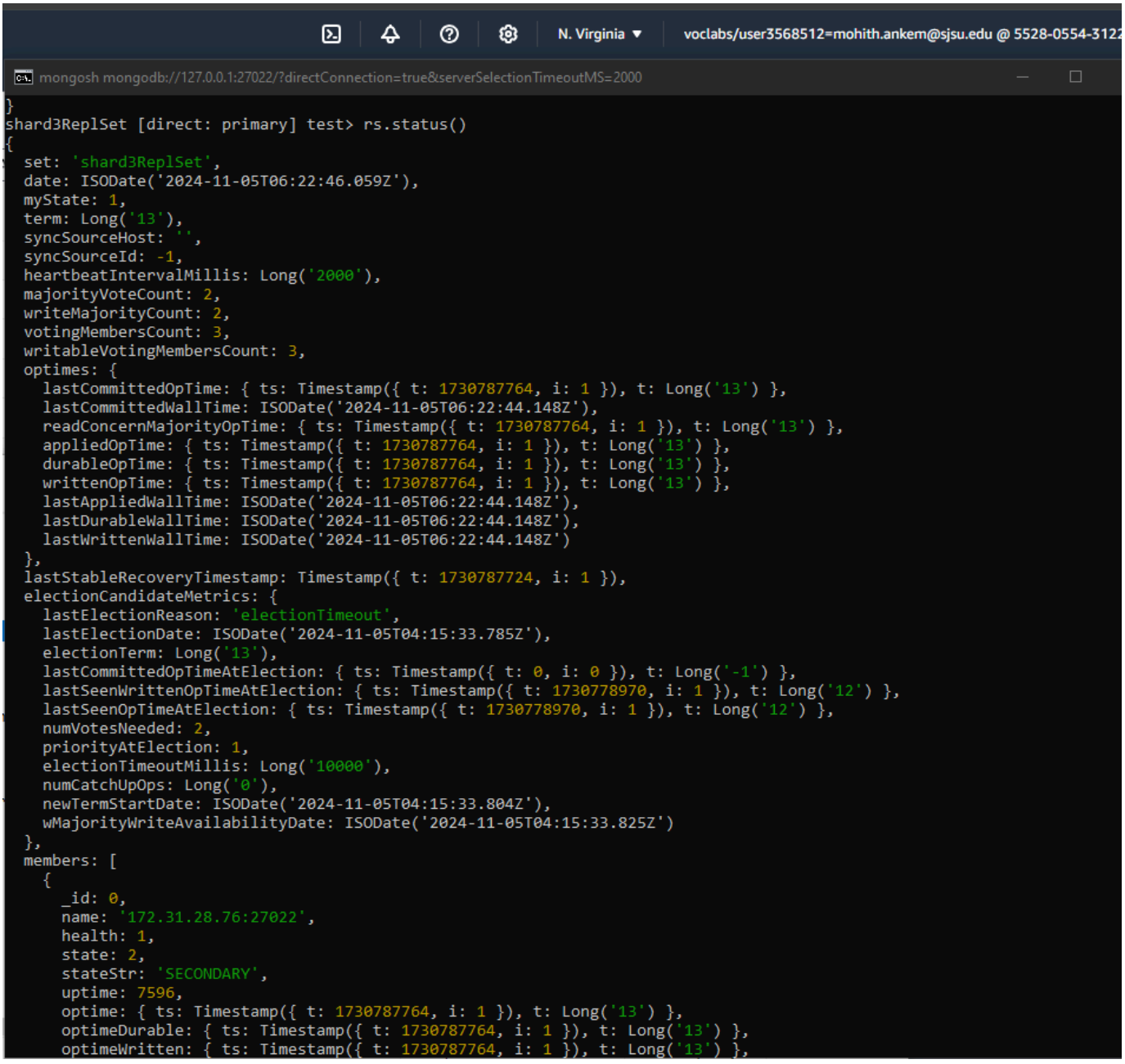


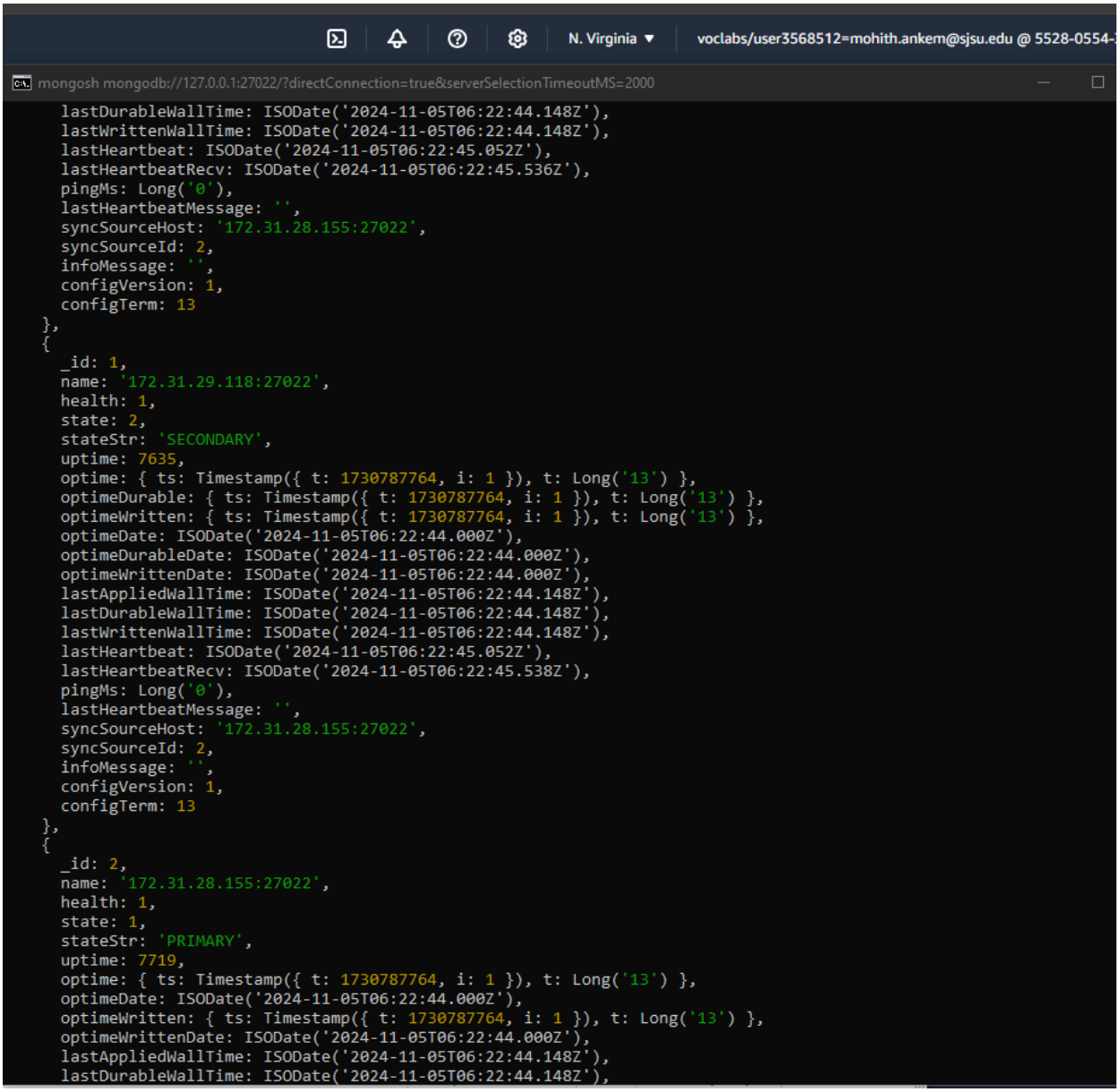

```
    optimeWritten: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeDurableDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeWrittenDate: ISODate('2024-11-05T06:55:21.000Z'),
    lastAppliedWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastDurableWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastWrittenWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastHeartbeat: ISODate('2024-11-05T06:55:22.655Z'),
    lastHeartbeatRecv: ISODate('2024-11-05T06:55:21.838Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: '',
    syncSourceHost: '172.31.28.155:27021',
    syncSourceId: 2,
    infoMessage: '',
    configVersion: 1,
    configTerm: 13
  },
  {
    _id: 1,
    name: '172.31.29.118:27021',
    health: 1,
    state: 2,
    stateStr: 'SECONDARY',
    uptime: 9601,
    optime: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeWritten: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeWrittenDate: ISODate('2024-11-05T06:55:21.000Z'),
    lastAppliedWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastDurableWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastWrittenWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    syncSourceHost: '172.31.28.155:27021',
    syncSourceId: 2,
    infoMessage: '',
    configVersion: 1,
    configTerm: 13,
    self: true,
    lastHeartbeatMessage: ''
  },
  {
    _id: 2,
    name: '172.31.28.155:27021',
    health: 1,
    state: 1,
    stateStr: 'PRIMARY',
    uptime: 9598,
    optime: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeDurable: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeWritten: { ts: Timestamp({ t: 1730789721, i: 1 }), t: Long('13') },
    optimeDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeDurableDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeWrittenDate: ISODate('2024-11-05T06:55:21.000Z'),
```

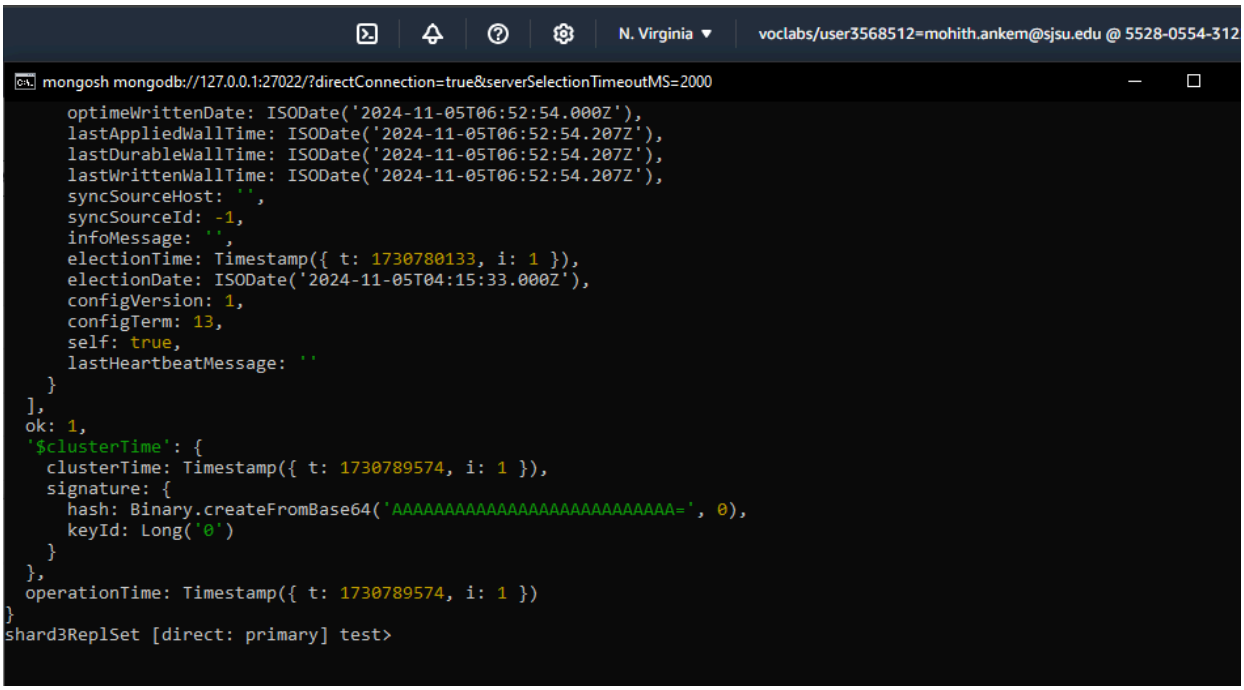



```
    optimeDurableDate: ISODate('2024-11-05T06:55:21.000Z'),
    optimeWrittenDate: ISODate('2024-11-05T06:55:21.000Z'),
    lastAppliedWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastDurableWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastWrittenWallTime: ISODate('2024-11-05T06:55:21.684Z'),
    lastHeartbeat: ISODate('2024-11-05T06:55:21.839Z'),
    lastHeartbeatRecv: ISODate('2024-11-05T06:55:21.837Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: '',
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: '',
    electionTime: Timestamp({ t: 1730780131, i: 1 }),
    electionDate: ISODate('2024-11-05T04:15:31.000Z'),
    configVersion: 1,
    configTerm: 13
  }
],
ok: 1,
'$clusterTime': {
  clusterTime: Timestamp({ t: 1730789721, i: 1 }),
  signature: {
    hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
    keyId: Long('0')
  }
},
operationTime: Timestamp({ t: 1730789721, i: 1 })
}
shard2ReplSet [direct: secondary] test>
```

rs.status() for shard3:




N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-
mongosh mongodb://127.0.0.1:27022/?directConnection=true&serverSelectionTimeoutMS=2000
lastDurableWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastWrittenWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastHeartbeat: ISODate('2024-11-05T06:22:45.052Z'),
lastHeartbeatRecv: ISODate('2024-11-05T06:22:45.536Z'),
pingMs: Long('0'),
lastHeartbeatMessage: '',
syncSourceHost: '172.31.28.155:27022',
syncSourceId: 2,
infoMessage: '',
configVersion: 1,
configTerm: 13
},
{
_id: 1,
name: '172.31.29.118:27022',
health: 1,
state: 2,
stateStr: 'SECONDARY',
uptime: 7635,
optime: { ts: Timestamp({ t: 1730787764, i: 1 }), t: Long('13') },
optimeDurable: { ts: Timestamp({ t: 1730787764, i: 1 }), t: Long('13') },
optimeWritten: { ts: Timestamp({ t: 1730787764, i: 1 }), t: Long('13') },
optimeDate: ISODate('2024-11-05T06:22:44.000Z'),
optimeDurableDate: ISODate('2024-11-05T06:22:44.000Z'),
optimeWrittenDate: ISODate('2024-11-05T06:22:44.000Z'),
lastAppliedWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastDurableWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastWrittenWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastHeartbeat: ISODate('2024-11-05T06:22:45.052Z'),
lastHeartbeatRecv: ISODate('2024-11-05T06:22:45.538Z'),
pingMs: Long('0'),
lastHeartbeatMessage: '',
syncSourceHost: '172.31.28.155:27022',
syncSourceId: 2,
infoMessage: '',
configVersion: 1,
configTerm: 13
},
{
_id: 2,
name: '172.31.28.155:27022',
health: 1,
state: 1,
stateStr: 'PRIMARY',
uptime: 7719,
optime: { ts: Timestamp({ t: 1730787764, i: 1 }), t: Long('13') },
optimeDate: ISODate('2024-11-05T06:22:44.000Z'),
optimeWritten: { ts: Timestamp({ t: 1730787764, i: 1 }), t: Long('13') },
optimeWrittenDate: ISODate('2024-11-05T06:22:44.000Z'),
lastAppliedWallTime: ISODate('2024-11-05T06:22:44.148Z'),
lastDurableWallTime: ISODate('2024-11-05T06:22:44.148Z'),


N. Virginia
voclabs/user3568512=mohith.ankem@sjsu.edu @ 5528-0554-312
mongosh mongodb://127.0.0.1:27022/?directConnection=true&serverSelectionTimeoutMS=2000
optimeWrittenDate: ISODate('2024-11-05T06:52:54.000Z'),
lastAppliedWallTime: ISODate('2024-11-05T06:52:54.207Z'),
lastDurableWallTime: ISODate('2024-11-05T06:52:54.207Z'),
lastWrittenWallTime: ISODate('2024-11-05T06:52:54.207Z'),
syncSourceHost: '',
syncSourceId: -1,
infoMessage: '',
electionTime: Timestamp({ t: 1730780133, i: 1 }),
electionDate: ISODate('2024-11-05T04:15:33.000Z'),
configVersion: 1,
configTerm: 13,
self: true,
lastHeartbeatMessage: ''
}
],
ok: 1,
'$clusterTime': {
clusterTime: Timestamp({ t: 1730789574, i: 1 }),
signature: {
hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
keyId: Long('0')
}
},
operationTime: Timestamp({ t: 1730789574, i: 1 })
}
shard3ReplSet [direct: primary] test>

14. List of hosts and describe what are deployed in each host. For example,

    node0:port# config server PRIMARY

    node1:port# config server SECONDARY

    …..

Ans:

node:port	**Description**
Node1_Instance1 : 27019	Config Server Primary
Node2_Server2 : 27019	Config Server Secondary
Node3_Server3 :27019	Config Server Secondary
Node4_Server4 : 27018 (mongos router)	Mongos Server
Node5_Server5 : 27020	Shard Server 1
Node6_Server6 : 27021	Shard Server 2
Node7_Server7 : 27022	Shard Server 3