

MongoDB Report

Step 1: Extract data from OSM

- We tried pushing the pbf files directly to the VM and then use that file to be uploaded to MongoDB, but we found that the JSON retrieved had to be formatted in order

Backup

- Shell Script for automated backup

```
GNU nano 7.2 mongo_db_backup.sh
BACKUP_DIR="/home/bitnami/mongodb_backups"
DATE=$(date +"%Y-%m-%d_%H-%M-%S")
DB_NAME="OSM"
CRED_FILE="/home/azureuser/bitnami_credentials"

# Extract credentials
MONGO_USER=$(grep "default username and password" $CRED_FILE | awk -F"'"' '{print $2}')
MONGO_PASS=$(grep "default username and password" $CRED_FILE | awk -F"'"' '{print $4}')

# Run mongodump
/opt/bitnami/mongodb/bin/mongodump --uri="mongodb://$MONGO_USER:$MONGO_PASS@localhost:27017/$DB_NAME?authSource=admin" --archive=$BACKUP_DIR

echo "Backup completed: $BACKUP_DIR/mongodb_backup_$DATE.gz"
```

[Wrote 14 lines]

^G Help	^O Write Out	^W Where Is	^K Cut	^J Execute	^C Location	^U Undo	^A Set Mark	^M To Bracket
^X Exit	^R Read File	^N Replace	^L Paste	^I Justify	^_ Go To Line	^E Redo	^D Copy	^Q Where Was

b. Scheduling backup script using crontab

```
GNU nano 7.2 -- bitnami@ssrawat2: ~ -- ssh -i ~/Desktop/Assignments/DBA/Unit 3/ssrawat2_key.pem azureuser@52.183.1163 /tmp/crontab.Mr0pcz/crontab *
# Edit this file to introduce tasks to be run by cron.
#
# Each task to run has to be defined through a single line
# indicating with different fields when the task will be run
# and what command to run for the task
#
# To define the time you can provide concrete values for
# minute (m), hour (h), day of month (dom), month (mon),
# and day of week (dow) or use '*' in these fields (for 'any').
#
# Notice that tasks will be started based on the cron's system
# daemon's notion of time and timezones.
#
# Output of the crontab jobs (including errors) is sent through
# email to the user the crontab file belongs to (unless redirected).
#
# For example, you can run a backup of all your user accounts
# at 5 a.m every week with:
# 0 5 * * 1 tar -zcf /var/backups/home.tgz /home/
#
# For more information see the manual pages of crontab(5) and cron(8)
#
# m h dom mon dow   command
10 * * * * /home/azureuser/mongodb_backup.sh >> /home/azureuser/mongodb_backup.log 2>&1|
```

c. Backup log as output file of backups

```
2024-11-30T20:10:01.943+0000 writing OSM.Kansas to archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:10:01.946+0000 writing OSM.Maine to archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:10:01.951+0000 writing OSM.kentucky to archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:10:01.953+0000 writing OSM.Texas to archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:10:04.922+0000 [##.....] OSM.Kansas 295169/2366351 (12.5%)
2024-11-30T20:10:04.922+0000 [#####.....] OSM.Texas 322278/418701 (77.0%)
2024-11-30T20:10:04.922+0000 [.....] OSM.kentucky 80288/3112803 (2.6%)
2024-11-30T20:10:04.922+0000 [##.....] OSM.Maine 144197/1631849 (8.8%)
2024-11-30T20:10:04.922+0000
2024-11-30T20:10:07.921+0000 [####.....] OSM.Kansas 419239/2366351 (17.7%)
2024-11-30T20:10:07.921+0000 [#####.....] OSM.Texas 397566/418701 (95.0%)
2024-11-30T20:10:07.921+0000 [##.....] OSM.kentucky 303545/3112803 (9.8%)
2024-11-30T20:10:07.921+0000 [####.....] OSM.Maine 274903/1631849 (16.8%)
2024-11-30T20:10:07.921+0000
2024-11-30T20:10:08.093+0000 [#####.....] OSM.Texas 418701/418701 (100.0%)
2024-11-30T20:10:10.724+0000 done dumping OSM.Texas (418701 documents)
2024-11-30T20:10:10.921+0000 [####.....] OSM.Kansas 468421/2366351 (19.8%)
2024-11-30T20:10:10.921+0000 [##.....] OSM.kentucky 492139/3112803 (15.8%)
2024-11-30T20:10:10.921+0000 [####.....] OSM.Maine 274903/1631849 (16.8%)
2024-11-30T20:10:10.921+0000
2024-11-30T20:10:13.922+0000 [####.....] OSM.Kansas 493432/2366351 (20.9%)
2024-11-30T20:10:13.922+0000 [##.....] OSM.kentucky 508181/3112803 (16.3%)
2024-11-30T20:10:13.922+0000 [####.....] OSM.Maine 339620/1631849 (20.8%)
2024-11-30T20:10:13.922+0000
2024-11-30T20:10:16.922+0000 [####.....] OSM.Kansas 518266/2366351 (21.9%)
2024-11-30T20:10:16.922+0000 [####.....] OSM.kentucky 538447/3112803 (17.3%)
2024-11-30T20:10:16.922+0000 [####.....] OSM.Maine 405311/1631849 (24.8%)
2024-11-30T20:10:16.922+0000
2024-11-30T20:10:19.921+0000 [####.....] OSM.Kansas 568541/2366351 (24.0%)
2024-11-30T20:10:19.921+0000 [####.....] OSM.kentucky 554343/3112803 (17.8%)
2024-11-30T20:10:19.921+0000 [####.....] OSM.Maine 405311/1631849 (24.8%)
2024-11-30T20:10:19.921+0000

^G Help      ^O Write Out  ^W Where Is   ^K Cut        ^T Execute    ^C Location   M-U Undo      M-A Set Mark  M-J To Bracket
^X Exit      ^R Read File  ^\ Replace    ^U Paste      ^_ Justify    ^/ Go To Line M-E Redo      M-6 Copy      ^Q Where Was

-- bitnami@ssrawat2: ~ -- ssh -i ~/Desktop/Assignments/DBA/Unit 3/ssrawat2_key.pem azureuser@52.163.1.163

GNU nano 7.2 backup.log
2024-11-30T20:12:49.922+0000 [#####.....] OSM.Maine 1597167/1631849 (97.9%)
2024-11-30T20:12:49.922+0000
2024-11-30T20:12:52.922+0000 [#####.....] OSM.Kansas 2169046/2366351 (91.7%)
2024-11-30T20:12:52.922+0000 [#####.....] OSM.kentucky 1993308/3112803 (64.0%)
2024-11-30T20:12:52.922+0000 [#####.....] OSM.Maine 1620255/1631849 (99.3%)
2024-11-30T20:12:52.922+0000
2024-11-30T20:12:53.848+0000 [#####.....] OSM.Maine 1631849/1631849 (100.0%)
2024-11-30T20:12:55.921+0000 [#####.....] OSM.Kansas 2283099/2366351 (93.1%)
2024-11-30T20:12:55.921+0000 [#####.....] OSM.kentucky 1993308/3112803 (64.0%)
2024-11-30T20:12:55.921+0000
2024-11-30T20:12:57.082+0000 done dumping OSM.Maine (1631849 documents)
2024-11-30T20:12:58.921+0000 [#####.....] OSM.Kansas 2258839/2366351 (95.5%)
2024-11-30T20:12:58.921+0000 [#####.....] OSM.kentucky 2061565/3112803 (66.2%)
2024-11-30T20:12:58.921+0000
2024-11-30T20:13:01.922+0000 [#####.....] OSM.Kansas 2283208/2366351 (96.5%)
2024-11-30T20:13:01.922+0000 [#####.....] OSM.kentucky 2123224/3112803 (68.2%)
2024-11-30T20:13:01.922+0000
2024-11-30T20:13:04.922+0000 [#####.....] OSM.Kansas 2354817/2366351 (99.5%)
2024-11-30T20:13:04.922+0000 [#####.....] OSM.kentucky 2154696/3112803 (69.2%)
2024-11-30T20:13:04.922+0000
2024-11-30T20:13:06.417+0000 [#####.....] OSM.Kansas 2366351/2366351 (100.0%)
2024-11-30T20:13:07.922+0000 [#####.....] OSM.kentucky 2222498/3112803 (71.4%)
2024-11-30T20:13:08.455+0000 done dumping OSM.Kansas (2366351 documents)
2024-11-30T20:13:10.921+0000 [#####.....] OSM.kentucky 2288517/3112803 (73.5%)
2024-11-30T20:13:13.922+0000 [#####.....] OSM.kentucky 2389588/3112803 (76.8%)
2024-11-30T20:13:16.921+0000 [#####.....] OSM.kentucky 2439220/3112803 (78.4%)
2024-11-30T20:13:19.921+0000 [#####.....] OSM.kentucky 2499639/3112803 (80.3%)
2024-11-30T20:13:22.921+0000 [#####.....] OSM.kentucky 2553793/3112803 (82.0%)
2024-11-30T20:13:25.921+0000 [#####.....] OSM.kentucky 2637097/3112803 (84.7%)
2024-11-30T20:13:28.921+0000 [#####.....] OSM.kentucky 2706018/3112803 (86.9%)
2024-11-30T20:13:31.922+0000 [#####.....] OSM.kentucky 2799091/3112803 (89.9%)
2024-11-30T20:13:34.922+0000 [#####.....] OSM.kentucky 2891897/3112803 (92.9%)
2024-11-30T20:13:37.921+0000 [#####.....] OSM.kentucky 2955825/3112803 (95.0%)
2024-11-30T20:13:40.921+0000 [#####.....] OSM.kentucky 3045494/3112803 (97.8%)
2024-11-30T20:13:43.922+0000 [#####.....] OSM.kentucky 3097766/3112803 (99.5%)
2024-11-30T20:13:44.138+0000 [#####.....] OSM.kentucky 3112803/3112803 (100.0%)
2024-11-30T20:13:44.514+0000 done dumping OSM.kentucky (3112803 documents)
Backup completed: /home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz

^G Help      ^O Write Out  ^W Where Is   ^K Cut        ^T Execute    ^C Location   M-U Undo      M-A Set Mark  M-J To Bracket
^X Exit      ^R Read File  ^\ Replace    ^U Paste      ^_ Justify    ^/ Go To Line M-E Redo      M-6 Copy      ^Q Where Was
```

Restore

a. Restore shell script

```
GNU nano 7.2 mongo_restore.sh
#!/bin/bash
# Variables (embedded in the script)
BACKUP_DIR="/home/bitnami/mongodb_backups"
DB_NAME="OSM"
CRED_FILE="/home/azureuser/bitnami_credentials"

# Extract credentials
MONGO_USER=$(grep "default username and password" $CRED_FILE | awk -F"'" '{print $2}')
MONGO_PASS=$(grep "default username and password" $CRED_FILE | awk -F"'" '{print $4}')

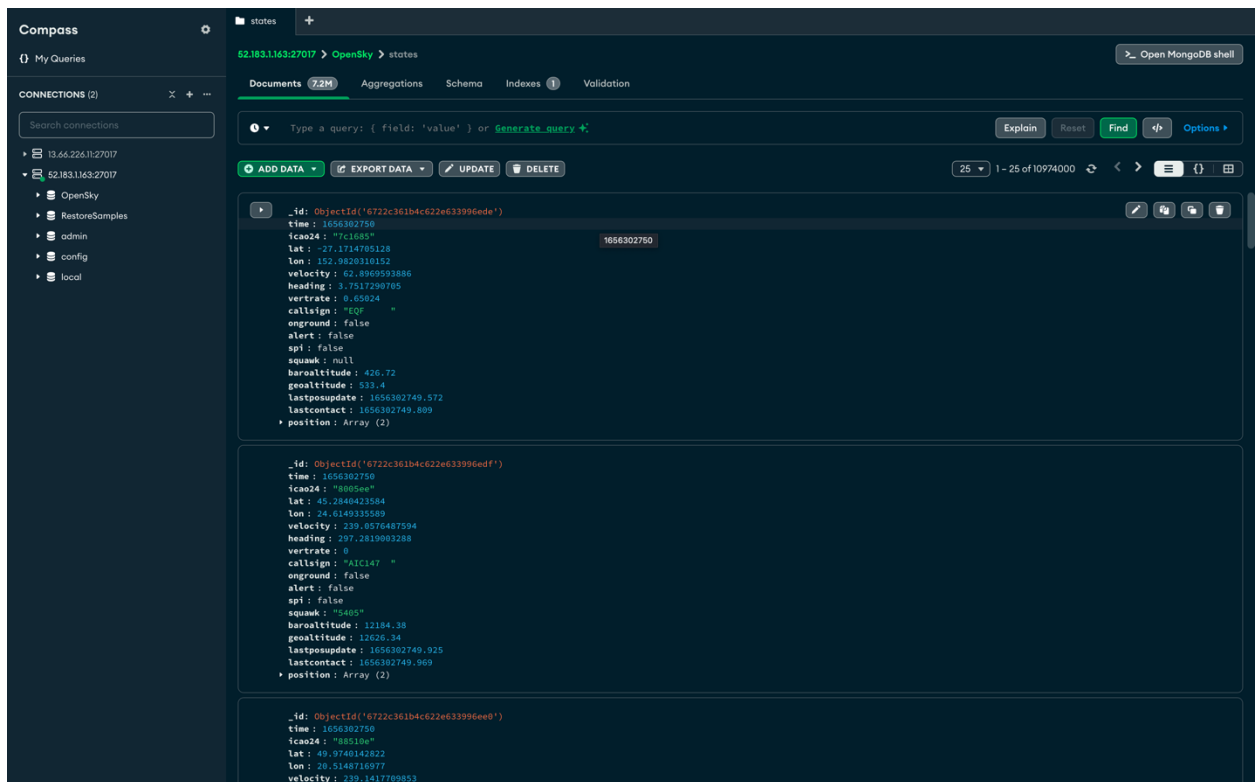
# Identify the latest backup file
LATEST_BACKUP=$(ls -t $BACKUP_DIR/mongodb_backup_*.gz | head -n 1)

# Run mongorestore
/opt/bitnami/mongodb/bin/mongorestore \
  --uri="mongodb://$MONGO_USER:$MONGO_PASS@localhost:27017/$DB_NAME?authSource=admin" \
  --archive=$LATEST_BACKUP \
  --gzip

echo "Restore completed from: $LATEST_BACKUP"
```

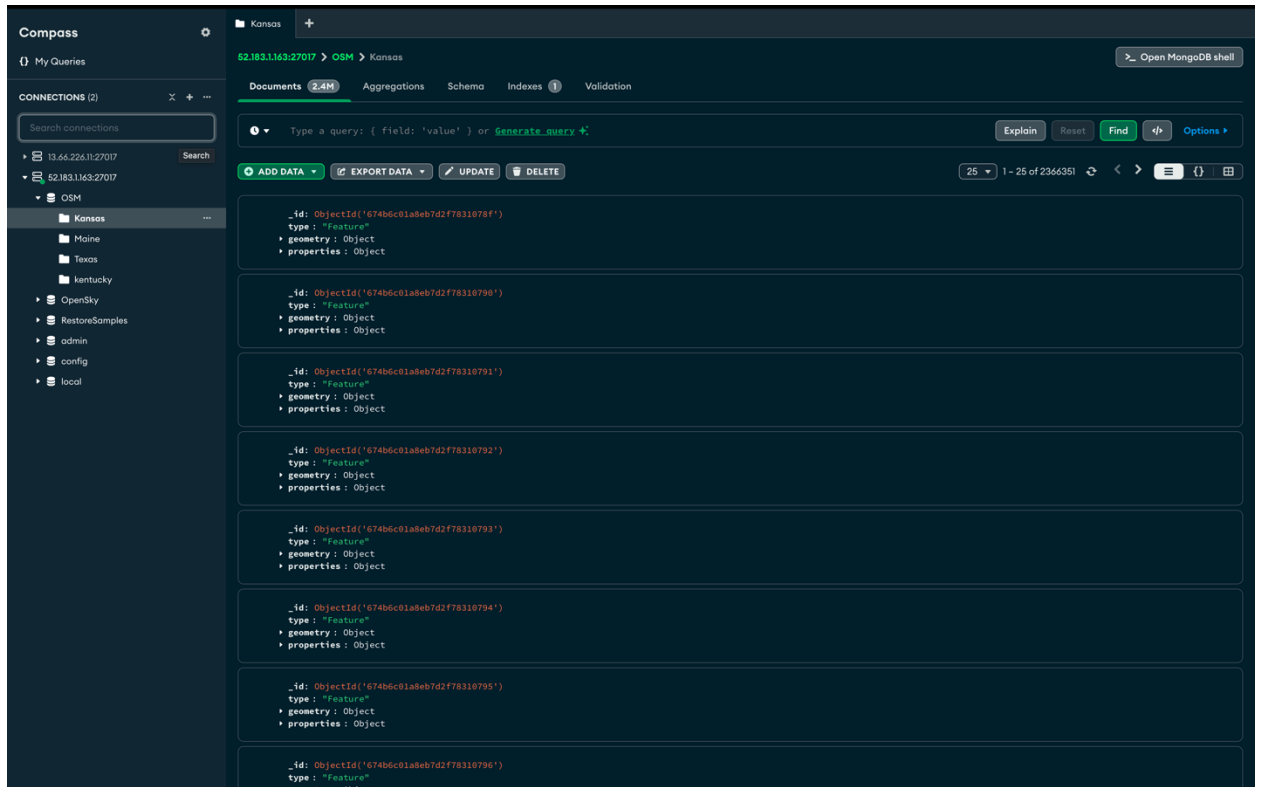
^G Help ^O Write Out ^W Where Is ^K Cut [Read 20 lines] ^T Execute ^C Location M-U Undo M-A Set Mark M-] To Bracket
^X Exit ^R Read File ^\ Replace ^U Paste ^J Justify ^_ Go To Line M-E Redo M-6 Copy ^Q Where Was

b. Deleted OSM Database from compass (Replicating any data loss)



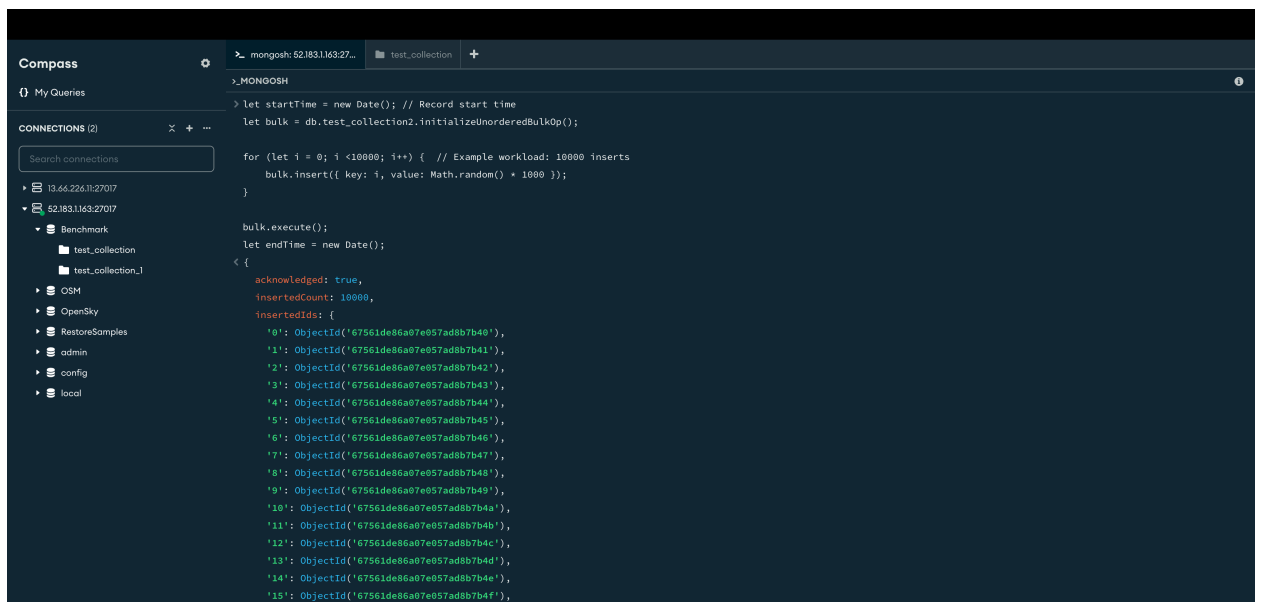
c. Restore script execution and output

```
bitnami@ssrawat2:~$ ./mongo_restore.sh
2024-11-30T20:54:06.423+0000 The --db and --collection flags are deprecated for this use-case; please use --nsInclude instead, i.e. with -
--nsInclude=${DATABASE}.${COLLECTION}
2024-11-30T20:54:06.455+0000 preparing collections to restore from
2024-11-30T20:54:06.473+0000 reading metadata for OSM.Texas from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01
.gz'
2024-11-30T20:54:06.473+0000 reading metadata for OSM.kentucky from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-
01.gz'
2024-11-30T20:54:06.473+0000 reading metadata for OSM.Maine from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01
.gz'
2024-11-30T20:54:06.473+0000 reading metadata for OSM.Kansas from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-0
1.gz'
2024-11-30T20:54:06.518+0000 restoring OSM.Kansas from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:54:07.587+0000 restoring OSM.Texas from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:54:09.423+0000 OSM.Kansas 30.2MB
2024-11-30T20:54:09.423+0000 OSM.Texas 16.0MB
2024-11-30T20:54:09.423+0000 restoring OSM.Maine from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:54:10.523+0000 restoring OSM.kentucky from archive '/home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz'
2024-11-30T20:54:12.321+0000 OSM.Kansas 32.0MB
2024-11-30T20:54:12.423+0000 OSM.Texas 48.0MB
2024-11-30T20:54:12.423+0000 OSM.Maine 16.0MB
2024-11-30T20:54:12.423+0000 OSM.kentucky 1.97MB
2024-11-30T20:54:12.423+0000
2024-11-30T20:54:15.423+0000 OSM.Kansas 43.2MB
2024-11-30T20:54:15.423+0000 OSM.Texas 48.0MB
2024-11-30T20:54:15.424+0000 OSM.Maine 32.0MB
2024-11-30T20:54:15.424+0000 OSM.kentucky 32.0MB
2024-11-30T20:54:15.424+0000
2024-11-30T20:54:18.423+0000 OSM.Kansas 64.0MB
2024-11-30T20:54:18.423+0000 OSM.Texas 60.6MB
2024-11-30T20:54:18.423+0000 OSM.Maine 48.0MB
2024-11-30T20:54:18.423+0000 OSM.kentucky 48.0MB
2024-11-30T20:54:18.423+0000
2024-11-30T20:54:21.425+0000 OSM.Kansas 70.6MB
2024-11-30T20:54:21.425+0000 OSM.Texas 80.0MB
2024-11-30T20:54:21.425+0000 OSM.Maine 48.0MB
2024-11-30T20:54:21.425+0000 OSM.kentucky 80.0MB
2024-11-30T20:54:21.426+0000
2024-11-30T20:54:24.423+0000 OSM.Kansas 96.0MB
2024-11-30T20:56:39.211+0000 OSM.kentucky 1.57GB
2024-11-30T20:56:39.211+0000 finished restoring OSM.kentucky (3112803 documents, 0 failures)
2024-11-30T20:56:39.212+0000 no indexes to restore for collection OSM.Maine
2024-11-30T20:56:39.212+0000 no indexes to restore for collection OSM.Kansas
2024-11-30T20:56:39.212+0000 restoring indexes for collection OSM.Texas from metadata
2024-11-30T20:56:39.222+0000 index: &idx.IndexDocument{Options:primitive.M{"2dsphereIndexVersion":3, "name":"position_2dsphere", "v":2}, K
ey:primitive.D{primitive.E{Key:"position", Value:"2dsphere"}}, PartialFilterExpression:primitive.D(nil)}
2024-11-30T20:56:39.222+0000 no indexes to restore for collection OSM.kentucky
2024-11-30T20:56:41.474+0000 7529704 document(s) restored successfully. 0 document(s) failed to restore.
Restore completed from: /home/bitnami/mongodb_backups/mongodb_backup_2024-11-30_20-10-01.gz
```



Benchmarking

For Scaling Factor -10000



```

  '9982': ObjectId('67561de96a07e057ad8ba23e'),
  '9983': ObjectId('67561de96a07e057ad8ba23f'),
  '9984': ObjectId('67561de96a07e057ad8ba240'),
  '9985': ObjectId('67561de96a07e057ad8ba241'),
  '9986': ObjectId('67561de96a07e057ad8ba242'),
  '9987': ObjectId('67561de96a07e057ad8ba243'),
  '9988': ObjectId('67561de96a07e057ad8ba244'),
  '9989': ObjectId('67561de96a07e057ad8ba245'),
  '9990': ObjectId('67561de96a07e057ad8ba246'),
  '9991': ObjectId('67561de96a07e057ad8ba247'),
  '9992': ObjectId('67561de96a07e057ad8ba248'),
  '9993': ObjectId('67561de96a07e057ad8ba249'),
  '9994': ObjectId('67561de96a07e057ad8ba24a'),
  '9995': ObjectId('67561de96a07e057ad8ba24b'),
  '9996': ObjectId('67561de96a07e057ad8ba24c'),
  '9997': ObjectId('67561de96a07e057ad8ba24d'),
  '9998': ObjectId('67561de96a07e057ad8ba24e'),
  '9999': ObjectId('67561de96a07e057ad8ba24f')
},
matchedCount: 0,
modifiedCount: 0,
deletedCount: 0,
upsertedCount: 0,
upsertedIds: {}
}
> let elapsedTime = (endTime - startTime) / 1000; // Convert to seconds
print("Elapsed Time (seconds):", elapsedTime);
< Elapsed Time (seconds):
< 12.211
> let tps = 10000 / elapsedTime; // TPS calculation
print("Transactions Per Second (TPS):", tps);
< Transactions Per Second (TPS):
< 818.9337482597058
Benchmark>
```

TPS came out to be **818.93**

Monitoring Screenshots

