

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»  
Інститут прикладного системного аналізу**

**ЛАБОРАТОРНА РОБОТА №3**

**з дисципліни «Високопродуктивні розподілені обчислювальні  
системи»**

**на тему: «Building MongoDB Environment»**

**Варіант №9**

**Виконали:**

Студенти групи КІ-31мп

Гордун М. В.

Заїка Б. Ю.

Решетник О. О.

**Перевірів:**

Кухарєв С. О.

Київ – 2023

## MongoDB sharding

docker-compose.yaml

```
version: '3'
services:

  ## Config Servers
  config01:
    image: mongo
    command: mongod --port 27017 --configsvr --replSet configserver
    volumes:
      - ./scripts:/scripts
  config02:
    image: mongo
    command: mongod --port 27017 --configsvr --replSet configserver
    volumes:
      - ./scripts:/scripts
  config03:
    image: mongo
    command: mongod --port 27017 --configsvr --replSet configserver
    volumes:
      - ./scripts:/scripts

  ## Shards
  shard01a:
    image: mongo
    command: mongod --port 27018 --shardsvr --replSet shard01
    volumes:
      - ./scripts:/scripts
  shard01b:
    image: mongo
    command: mongod --port 27018 --shardsvr --replSet shard01
    volumes:
      - ./scripts:/scripts
  shard01c:
    image: mongo
    command: mongod --port 27018 --shardsvr --replSet shard01
    volumes:
      - ./scripts:/scripts

  shard02a:
    image: mongo
    command: mongod --port 27019 --shardsvr --replSet shard02
```

```
volumes:
  - ./scripts:/scripts
shard02b:
  image: mongo
  command: mongod --port 27019 --shardsvr --replSet shard02
  volumes:
    - ./scripts:/scripts
shard02c:
  image: mongo
  command: mongod --port 27019 --shardsvr --replSet shard02
  volumes:
    - ./scripts:/scripts

shard03a:
  image: mongo
  command: mongod --port 27020 --shardsvr --replSet shard03
  volumes:
    - ./scripts:/scripts
shard03b:
  image: mongo
  command: mongod --port 27020 --shardsvr --replSet shard03
  volumes:
    - ./scripts:/scripts
shard03c:
  image: mongo
  command: mongod --port 27020 --shardsvr --replSet shard03
  volumes:
    - ./scripts:/scripts

shard04a:
  image: mongo
  command: mongod --port 27021 --shardsvr --replSet shard04
  volumes:
    - ./scripts:/scripts
shard04b:
  image: mongo
  command: mongod --port 27021 --shardsvr --replSet shard04
  volumes:
    - ./scripts:/scripts
shard04c:
  image: mongo
  command: mongod --port 27021 --shardsvr --replSet shard04
```

```

volumes:
  - ./scripts:/scripts

## Router
router:
  image: mongo
  command: mongos --port 27017 --configdb
configserver/config01:27017,config02:27017,config03:27017 --bind_ip_all
  ports:
    - "27017:27017"
  volumes:
    - ./scripts:/scripts
    - ./data:/data

```

## docker-compose up -d

```

C:\Users\1\VSpy\HL\Lab-3>docker-compose up -d
[+] Running 17/17
✔ Network lab-3_default Created 0.1s
✔ Container lab-3-shard04c-1 Started 1.4s
✔ Container lab-3-shard01c-1 Started 1.1s
✔ Container lab-3-shard02c-1 Started 1.4s
✔ Container lab-3-shard04a-1 Started 1.0s
✔ Container lab-3-config03-1 Started 1.3s
✔ Container lab-3-shard02a-1 Started 1.2s
✔ Container lab-3-shard02b-1 Started 1.3s
✔ Container lab-3-shard03b-1 Started 1.5s
✔ Container lab-3-shard01a-1 Started 1.0s
✔ Container lab-3-config02-1 Started 1.4s
✔ Container lab-3-config01-1 Started 1.5s
✔ Container lab-3-shard03c-1 Started 1.4s
✔ Container lab-3-shard04b-1 Started 1.3s
✔ Container lab-3-router-1 Started 1.4s
✔ Container lab-3-shard01b-1 Started 0.9s
✔ Container lab-3-shard03a-1 Started 1.0s

```

## Initializing shards

`docker-compose exec config01 sh -c "mongosh --host localhost --port 27017 /scripts/init-configserver.js"`

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec config01 sh -c "mongosh --host localhost --port 27017 /scripts/init-configserver.js"
Current Mongosh Log ID: 6585c2e42e8b06ed1bc0d4f
Connecting to:      mongodb://localhost:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:      2.1.1

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:09:14.615+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2023-12-22T17:09:22.055+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-22T17:09:22.055+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to 'never'
2023-12-22T17:09:22.055+00:00: vm.max_map_count is too low
-----
Loading file: /scripts/init-configserver.js
MongoServerError: Rejecting initiate with a set name that differs from command line set name, initiate set name: rs-config-server, command line set name: configserver
```

`docker-compose exec shard01a sh -c "mongosh --host localhost --port 27018 /scripts/init-shard01.js"`

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec shard01a sh -c "mongosh --host localhost --port 27018 /scripts/init-shard01.js"
Current Mongosh Log ID: 6585c30aa5bb6bd2874ace4
Connecting to:      mongodb://localhost:27018/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:      2.1.1

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:09:14.528+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2023-12-22T17:09:22.030+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-22T17:09:22.031+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to 'never'
2023-12-22T17:09:22.031+00:00: vm.max_map_count is too low
-----
Loading file: /scripts/init-shard01.js
MongoServerError: Rejecting initiate with a set name that differs from command line set name, initiate set name: rs-shard-01, command line set name: shard01
```

`docker-compose exec shard02a sh -c "mongosh --host localhost --port 27019 /scripts/init-shard02.js"`

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec shard02a sh -c "mongosh --host localhost --port 27019 /scripts/init-shard02.js"
Current Mongosh Log ID: 6585c34178de3baa55cf7a86
Connecting to:      mongodb://localhost:27019/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:      2.1.1

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:09:14.539+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2023-12-22T17:09:22.097+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-22T17:09:22.098+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to 'never'
2023-12-22T17:09:22.098+00:00: vm.max_map_count is too low
-----
Loading file: /scripts/init-shard02.js
MongoServerError: Rejecting initiate with a set name that differs from command line set name, initiate set name: rs-shard-02, command line set name: shard02
```

docker-compose exec shard03a sh -c "mongosh --host localhost --port 27020 /scripts/init-shard03.js"

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec shard03a sh -c "mongosh --host localhost --port 27020 /scripts/init-shard03.js"
Current Mongosh Log ID: 6585c399e2ee7648b6eceb8f
Connecting to:      mongodb://localhost:27020/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:      2.1.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:09:14.573+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2023-12-22T17:09:22.081+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-22T17:09:22.081+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to 'never'
2023-12-22T17:09:22.081+00:00: vm.max_map_count is too low
-----

Loading file: /scripts/init-shard03.js
MongoServerError: Rejecting initiate with a set name that differs from command line set name, initiate set name: rs-shard-03, command line set name: shard03
```

docker-compose exec shard04a sh -c "mongosh --host localhost --port 27021 /scripts/init-shard04.js"

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec shard04a sh -c "mongosh --host localhost --port 27021 /scripts/init-shard04.js"
Current Mongosh Log ID: 6585c3bf587a3f1a56325821
Connecting to:      mongodb://localhost:27021/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:      2.1.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:09:14.559+00:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2023-12-22T17:09:22.085+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2023-12-22T17:09:22.086+00:00: /sys/kernel/mm/transparent_hugepage/enabled is 'always'. We suggest setting it to 'never'
2023-12-22T17:09:22.086+00:00: vm.max_map_count is too low
-----

Loading file: /scripts/init-shard04.js
```

## Initializing router

docker-compose exec router sh -c "mongosh --host localhost --port 27017 /scripts/init-router.js"

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec router sh -c "mongosh --host localhost --port 27017 /scripts/init-router.js"
Current Mongosh Log ID: 6585c54e41d317d103dba189
Connecting to:      mongodb://localhost:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:       2.1.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

-----
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2023-12-22T17:19:23.513+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

Loading file: /scripts/init-router.js
```

## Importing data

`docker-compose exec router sh -c "mongosh < /scripts/enable-sharding.js"`

```
C:\Users\1\VSpy\HL\Lab-3>docker-compose exec router sh -c "mongosh < /scripts/enable-sharding.js"
Current Mongosh Log ID: 6585c5a421f1b88793a18e3b
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.1.1
Using MongoDB:      7.0.4
Using Mongosh:       2.1.1

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

-----
The server generated these startup warnings when booting
2023-12-22T17:19:23.513+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----

[direct: mongos] test> sh.enableSharding('taxi')
{
  ok: 1,
  '$clusterTime': {
    clusterTime: Timestamp({ t: 1703265701, i: 8 }),
    signature: {
      hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
      keyId: Long('0')
    }
  },
  operationTime: Timestamp({ t: 1703265701, i: 2 })
}
[direct: mongos] test> db.rides.createIndex( { 'driver': 'hashed' }, { name: 'driver' } )
driver
[direct: mongos] test> sh.shardCollection('taxi.rides', { 'driver': 'hashed' })
{
  collectionsSharded: 'taxi.rides',
  ok: 1,
  '$clusterTime': {
    clusterTime: Timestamp({ t: 1703265702, i: 51 }),
    signature: {
      hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAAAAAAA=', 0),
      keyId: Long('0')
    }
  },
  operationTime: Timestamp({ t: 1703265702, i: 51 })
}
[direct: mongos] test> sh.shardCollection('taxi.rides', { 'driver': 'hashed' })
```

`docker-compose exec router sh -c "mongoimport --host router --port 27017 --db taxi --collection rides --type csv --file data/data_small.csv --headerline --numInsertionWorkers=4 --batchSize=100"`

```
2023-12-22T17:29:22.147+0000 [#####] taxi.rides 1.00GB/1.03GB (97.8%)
2023-12-22T17:29:25.153+0000 [#####] taxi.rides 1.01GB/1.03GB (98.0%)
2023-12-22T17:29:28.148+0000 [#####] taxi.rides 1.01GB/1.03GB (98.1%)
2023-12-22T17:29:31.148+0000 [#####] taxi.rides 1.01GB/1.03GB (98.3%)
2023-12-22T17:29:34.148+0000 [#####] taxi.rides 1.01GB/1.03GB (98.4%)
2023-12-22T17:29:37.148+0000 [#####] taxi.rides 1.01GB/1.03GB (98.5%)
2023-12-22T17:29:40.148+0000 [#####] taxi.rides 1.01GB/1.03GB (98.7%)
2023-12-22T17:29:43.147+0000 [#####] taxi.rides 1.01GB/1.03GB (98.9%)
2023-12-22T17:29:46.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.0%)
2023-12-22T17:29:49.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.2%)
2023-12-22T17:29:52.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.3%)
2023-12-22T17:29:55.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.4%)
2023-12-22T17:29:58.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.6%)
2023-12-22T17:30:01.147+0000 [#####] taxi.rides 1.02GB/1.03GB (99.7%)
2023-12-22T17:30:04.148+0000 [#####] taxi.rides 1.02GB/1.03GB (99.9%)
2023-12-22T17:30:06.986+0000 [#####] taxi.rides 1.03GB/1.03GB (100.0%)
2023-12-22T17:30:06.992+0000 400000 document(s) imported successfully. 0 document(s) failed to import.
```

## Shard Distribution

```
Shard shard02 at shard02/shard02a:27019,shard02b:27019,shard02c:27019
{
  data: '448.83MiB',
  docs: 1032310,
  chunks: 2,
  'estimated data per chunk': '224.41MiB',
  'estimated docs per chunk': 516155
}
---
Shard shard03 at shard03/shard03a:27020,shard03b:27020,shard03c:27020
{
  data: '454.77MiB',
  docs: 1046086,
  chunks: 2,
  'estimated data per chunk': '227.38MiB',
  'estimated docs per chunk': 523043
}
---
Shard shard01 at shard01/shard01a:27018,shard01b:27018,shard01c:27018
{
  data: '424.09MiB',
  docs: 975492,
  chunks: 2,
  'estimated data per chunk': '212.04MiB',
  'estimated docs per chunk': 487746
}
---
Shard shard04 at shard04/shard04a:27021,shard04b:27021,shard04c:27021
{
  data: '411.35MiB',
  docs: 946112,
  chunks: 2,
  'estimated data per chunk': '205.67MiB',
  'estimated docs per chunk': 473056
}
```



## Totals

```
{
  data: '1.69GiB',
  docs: 4000000,
  chunks: 8,
  'Shard shard02': [
    '25.8 % data',
    '25.8 % docs in cluster',
    '455B avg obj size on shard'
  ],
  'Shard shard03': [
    '26.15 % data',
    '26.15 % docs in cluster',
    '455B avg obj size on shard'
  ],
  'Shard shard01': [
    '24.38 % data',
    '24.38 % docs in cluster',
    '455B avg obj size on shard'
  ],
  'Shard shard04': [
    '23.65 % data',
    '23.65 % docs in cluster',
    '455B avg obj size on shard'
  ]
}
```

## Querying data

docker-compose exec router sh -c "mongosh < /scripts/query-data.js"

```
[
  {
    _id: 'Driver took a longer route, claiming it was to avoid traffic, but it felt like a scam.',
    maxLength: 86
  },
  {
    _id: 'Taxi had a malfunctioning meter, and the fare was significantly higher than expected.',
    maxLength: 85
  },
  {
    _id: "Driver was talking on the phone loudly in a language passengers couldn't understand.",
    maxLength: 84
  },
  {
    _id: 'Driver was talking on the phone the entire time, ignoring my requests for silence.',
    maxLength: 82
  },
  {
    _id: 'The driver was talking loudly on a personal call, ignoring passenger discomfort.',
    maxLength: 80
  },
  {
    _id: "Driver refused to take a credit card, even though it's a listed payment option.",
    maxLength: 79
  },
  {
    _id: 'The taxi was dirty, and the seats were stained. Hygiene is a serious concern.',
    maxLength: 77
  },
  {
    _id: 'Cab had a strange rattling noise that persisted throughout the entire trip.',
    maxLength: 75
  },
  {
    _id: 'Taxi had a strange vibration that made it difficult to hold a conversation.',
    maxLength: 75
  },
  {
    _id: 'Cab had a strange odor; it was extremely unpleasant during the entire ride.',
    maxLength: 75
  }
]
```