

Bases de Datos a Gran Escala

Abraham Trashorras Rivas

Práctica 4

Objetivo, Crear cluster MONGODB:

Vamos a crear un cluster con esta estructura:



E1:

Creamos 3 máquinas para el cluster a partir de `greiBDServerBase`. Aprovechamos las que ya teníamos creadas de la Práctica 3:

Nombre	Ip
Citus1	192.168.56.104
Citus2	192.168.56.103
Citus3	192.168.56.102

E2:

Primero creo la carpeta `mongoConf` en todas las máquinas y después ejecuto el comando:

```
nohup mongod --configsvr --replSet rsConfServer --port 27017 --dbpath  
/home/alumnogreibd/mongoConf --bind_ip localhost,192.168.56.IP > /dev/null &
```

Usamos `nohup` para que el proceso no se cierre al cerrar sesión.

Donde se encuentra IP se referencia a los últimos 3 dígitos que identifican cada máquina. De esta forma configuramos los nodos de MongoDB.

Ahora tenemos que usar `rs.initiate` para configurar el `replicaSet`. Desde una de las máquinas, en este caso `citius1`, ejecutamos:

```
rs.initiate({
  _id: "rsConfServer",
  configsvr: true,
  members: [
    { _id: 0, host: "192.168.56.104:27017" },
    { _id: 1, host: "192.168.56.103:27017" },
    { _id: 2, host: "192.168.56.102:27017" }
  ]
})
```

E3:

Vamos a crear los shards con la información disponible en la siguiente tabla:

Shard	Folder	Port
Shard1	/home/alumnogreibd/shard1	27018
Shard2	/home/alumnogreibd/shard2	27019
Shard3	/home/alumnogreibd/shard3	27020

Debemos hacer los siguientes pasos para cada shard N:

Primero, crear la carpeta en cada máquina:

```
mkdir shardN
```

Ahora tenemos que crear los shards en cada máquina de esta forma:

```
nohup mongod --shardsvr --replSet rsShardN --port PORTN --dbpath
/home/alumnogreibd/shardN --bind_ip localhost,192.168.56.IP > /dev/null &
```

Donde N es el número de shard e IP son los tres últimos números de IP de cada máquina.

Después nos vamos a conectar desde `citius1` a mongo para iniciar los shards:

```
mongo --host localhost --port PORTN
rs.initiate({
  _id: "rsShard1",
  members: [
    { _id: 0, host: "192.168.56.104:PORTN" },
    { _id: 1, host: "192.168.56.103:PORTN" },
    { _id: 2, host: "192.168.56.102:PORTN" }
  ]
})
```

E4:

En cada máquina debemos iniciar una instancia de mongos con el siguiente comando:

```
nohup mongos --port 27021 --configdb
```

```
rsConfServer/192.168.56.104:27017,192.168.56.103:27017,192.168.56.102:27017 --bind_ip localhost,192.168.56.IP
> /dev/null &
```

Ahora nos conectamos en citius1 al servidor y añadimos los shards

```
mongo --host localhost --port 27021
```

```
sh.addShard("rsShard1/192.168.56.104:27018,192.168.56.103:27018,192.168.56.102:27018")
sh.addShard("rsShard2/192.168.56.104:27019,192.168.56.103:27019,192.168.56.102:27019")
sh.addShard("rsShard3/192.168.56.104:27020,192.168.56.103:27020,192.168.56.102:27020")
```

E5:

Aun conectados al servidor, seleccionamos la base de datos correcta y cargamos los datos generados en la practica 3 con peliculasjson:

```
use bdge
sh.enableSharding("bdge")
```

```
mongoimport --host=localhost:27021 --db=bdge --collection peliculas --
file=/home/alumnogreibd/peliculas.json
```

E6:

Seguimos conectados al servidor, creamos un índice para los datos y añadimos la colección a los shards con dicho índice.

```
db.peliculas.createIndex({ "id": "hashed" })
sh.shardCollection("bdge.peliculas", { "id": "hashed" })
```

E7:

Seguimos conectados, podemos ya ejecutar consultas. Aquí tenemos la consulta para solucionar el ejercicio 7:

```
db.peliculas.find(
  {
    presupuesto: { $gt: 500000 }
  }
).count()
```

Resultado: 7705

E8:

Pauso citus2 y reejecuto la consulta. Sigue dando la misma respuesta!

E9:

Ahora pauso tanto citius2 como citius3 y reejecuto la consulta. Se queda y devuelve:

```

mongos> db.peliculas.find( { presupuesto:{$gt:500000} } ).count()
uncaught exception: Error: count failed: {
  "ok" : 0,
  "errmsg" : "failed on: rsShard3 :: caused by :: Could not find host matching read preference { mode: \"primary\" } for set rsShard3",
  "code" : 133,
  "codeName" : "FailedToSatisfyReadPreference",
  "operationTime" : Timestamp(1703273758, 1),
  "$clusterTime" : {
    "clusterTime" : Timestamp(1703273758, 1),
    "signature" : {
      "hash" : BinData(0,"AAAAAAAAAAAAAAAAAAAAAAAAAAAA="),
      "keyId" : NumberLong(0)
    }
  }
} :
_getErrorWithCode@src/mongo/shell/utils.js:25:13
DBQuery.prototype.count@src/mongo/shell/query.js:386:11
@(shell):1:1

```

E10:

Despauo las maquinas y vuelvo a ejecutarlo. Me devuelve la solución correcta de 7702.

E11:

```

db.peliculas.updateMany(
  {},
  [
    {
      "$set": {
        "fecha_emision": {
          "$dateFromString": {
            "dateString": "$fecha_emision"
          }
        }
      }
    }
  ]
);

```

```

mongos> db.peliculas.updateMany(
... {},
... [
... {
... "$set": {
... "fecha_emision": {
... "$dateFromString": {
... dateString: "$fecha_emision"
... }
... }
... }
... }
... ]
... );
{ "acknowledged" : true, "matchedCount" : 69062, "modifiedCount" : 68934 }

```

E12:

Consulta creada:

```
db.peliculas.find(
  {
    "generos.nombre": "Science Fiction",
    ingresos: { $gt: 800000000 }
  },
  {
    titulo: 1,
    presupuesto: 1,
    _id: 0
  }
).pretty()
```

Resultado:

```
{
  "titulo" : "Transformers: Revenge of the Fallen",
  "presupuesto" : 150000000
}
{ "titulo" : "Avatar", "presupuesto" : 237000000 }
{ "titulo" : "Inception", "presupuesto" : 160000000 }
{ "titulo" : "Transformers: Dark of the Moon", "presupuesto" : 195000000 }
{ "titulo" : "Star Wars: The Force Awakens", "presupuesto" : 245000000 }
{ "titulo" : "Captain America: Civil War", "presupuesto" : 250000000 }
{ "titulo" : "Guardians of the Galaxy Vol. 2", "presupuesto" : 200000000 }
{ "titulo" : "Independence Day", "presupuesto" : 75000000 }
{ "titulo" : "Transformers: Age of Extinction", "presupuesto" : 210000000 }
{ "titulo" : "The Hunger Games: Catching Fire", "presupuesto" : 130000000 }
{ "titulo" : "Jurassic Park", "presupuesto" : 63000000 }
{ "titulo" : "The Avengers", "presupuesto" : 220000000 }
{ "titulo" : "Iron Man 3", "presupuesto" : 200000000 }
{ "titulo" : "Avengers: Age of Ultron", "presupuesto" : 280000000 }
{ "titulo" : "Jurassic World", "presupuesto" : 150000000 }
{ "titulo" : "Rogue One: A Star Wars Story", "presupuesto" : 200000000 }
{
  "titulo" : "Star Wars: Episode III - Revenge of the Sith",
  "presupuesto" : 113000000
}
{
  "titulo" : "Star Wars: Episode I - The Phantom Menace",
  "presupuesto" : 115000000
}
```

E13:

Consulta creada:

```
db.peliculas.aggregate([
  {
    $match: {
      "reparto.persona.nombre": "Penélope Cruz",
      "generos.nombre": "Action"
    }
  },
  {
    $project: {
      titulo: 1,

```

```

        titulo_original: 1,
        _id: 0
    }
},
{ $sort: { fecha_emision: 1 } }
])

```

Resultado:

```

{ "titulo" : "Bandidas", "titulo_original" : "Bandidas" }
{ "titulo" : "G-Force", "titulo_original" : "G-Force" }
{ "titulo" : "Sahara", "titulo_original" : "Sahara" }
{ "titulo" : "Grimsby", "titulo_original" : "Grimsby" }
{ "titulo" : "The Hi-Lo Country", "titulo_original" : "The Hi-Lo Country" }
{ "titulo" : "Pirates of the Caribbean: On Stranger Tides", "titulo_original" : "Pirates of th
e Caribbean: On Stranger Tides" }

```

E14:

Consulta creada:

```

db.peliculas.aggregate([
  { $unwind: '$generos' },
  {
    $group: {
      _id: '$generos.nombre',
      peliculas: { $sum: 1 },
      ingresos_totales: { $sum: '$ingresos' },
      presupuesto_total: { $sum: '$presupuesto' }
    }
  },
  {
    $project: {
      genero: '$_id',
      peliculas: 1,
      ingresos_totales: 1,
      presupuesto_total: 1,
      beneficios: { $subtract: ['$ingresos_totales', '$presupuesto_total'] }
    }
  },
  { $sort: { beneficios: -1 } }
])

```

Resultado:

```

{ "_id" : "Adventure", "peliculas" : 3490, "ingresos_totales" : NumberLong("199978669360"), "presupuesto_total" : NumberLong("65706926364"), "genero" : "Adventure", "beneficios" : NumberLong("134271742996") }
{ "_id" : "Action", "peliculas" : 6592, "ingresos_totales" : NumberLong("201388050019"), "presupuesto_total" : NumberLong("77709151433"), "genero" : "Action", "beneficios" : NumberLong("123678898586") }
{ "_id" : "Comedy", "peliculas" : 13176, "ingresos_totales" : NumberLong("166862896741"), "presupuesto_total" : NumberLong("60136127186"), "genero" : "Comedy", "beneficios" : NumberLong("106726769555") }
{ "_id" : "Drama", "peliculas" : 20244, "ingresos_totales" : NumberLong("168083769347"), "presupuesto_total" : NumberLong("70050619008"), "genero" : "Drama", "beneficios" : NumberLong("90753150339") }
{ "_id" : "Thriller", "peliculas" : 7619, "ingresos_totales" : NumberLong("129749706040"), "presupuesto_total" : NumberLong("55510980272"), "genero" : "Thriller", "beneficios" : NumberLong("74238725768") }
{ "_id" : "Family", "peliculas" : 2767, "ingresos_totales" : NumberLong("10709957659"), "presupuesto_total" : NumberLong("33519899995"), "genero" : "Family", "beneficios" : NumberLong("73580857664") }
{ "_id" : "Fantasy", "peliculas" : 2309, "ingresos_totales" : NumberLong("103938276500"), "presupuesto_total" : NumberLong("34278683311"), "genero" : "Fantasy", "beneficios" : NumberLong("69659593189") }
{ "_id" : "Science Fiction", "peliculas" : 3044, "ingresos_totales" : NumberLong("97847960421"), "presupuesto_total" : NumberLong("35364867470"), "genero" : "Science Fiction", "beneficios" : NumberLong("62483092951") }
{ "_id" : "Animation", "peliculas" : 1931, "ingresos_totales" : NumberLong("67433244964"), "presupuesto_total" : NumberLong("19957486079"), "genero" : "Animation", "beneficios" : NumberLong("47475758885") }
{ "_id" : "Romance", "peliculas" : 6730, "ingresos_totales" : NumberLong("73473647770"), "presupuesto_total" : NumberLong("26037454973"), "genero" : "Romance", "beneficios" : NumberLong("47436192797") }
{ "_id" : "Crime", "peliculas" : 4304, "ingresos_totales" : NumberLong("63386623657"), "presupuesto_total" : NumberLong("27601759972"), "genero" : "Crime", "beneficios" : NumberLong("35784863685") }
{ "_id" : "Mystery", "peliculas" : 2464, "ingresos_totales" : NumberLong("34754614989"), "presupuesto_total" : NumberLong("14425146930"), "genero" : "Mystery", "beneficios" : NumberLong("20329468059") }
{ "_id" : "Horror", "peliculas" : 4671, "ingresos_totales" : NumberLong("30837094673"), "presupuesto_total" : NumberLong("11823585606"), "genero" : "Horror", "beneficios" : NumberLong("19013509067") }
{ "_id" : "Music", "peliculas" : 1597, "ingresos_totales" : NumberLong("13370292367"), "presupuesto_total" : NumberLong("4466931321"), "genero" : "Music", "beneficios" : NumberLong("8903361046") }
{ "_id" : "War", "peliculas" : 1322, "ingresos_totales" : NumberLong("15910458263"), "presupuesto_total" : NumberLong("7805981358"), "genero" : "War", "beneficios" : NumberLong("8104476905") }
{ "_id" : "History", "peliculas" : 1398, "ingresos_totales" : NumberLong("14902198420"), "presupuesto_total" : NumberLong("8914261694"), "genero" : "History", "beneficios" : NumberLong("5987936726") }
{ "_id" : "Western", "peliculas" : 1042, "ingresos_totales" : NumberLong("5122498884"), "presupuesto_total" : NumberLong("3178439427"), "genero" : "Western", "beneficios" : NumberLong("1944059457") }
{ "_id" : "Documentary", "peliculas" : 3930, "ingresos_totales" : 1449528578, "presupuesto_total" : 466941285, "genero" : "Documentary", "beneficios" : 982587293 }
{ "_id" : "Foreign", "peliculas" : 1619, "ingresos_totales" : 291633311, "presupuesto_total" : 401004833, "genero" : "Foreign", "beneficios" : -100371522 }
{ "_id" : "TV Movie", "peliculas" : 766, "ingresos_totales" : 42000000, "presupuesto_total" : 179449000, "genero" : "TV Movie", "beneficios" : -137449000 }

```

E15:

Consulta creada:

```

db.peliculas.aggregate([
  { $unwind: '$personal' },

```

```

{
  $match: {
    "personal.trabajo": "Director",
    "ingresos": { $gt: 0 }
  },
  {
    $group: {
      _id: '$personal.persona.nombre',
      peliculas: { $sum: 1 },
      ingresos_totales: { $sum: '$ingresos' },
      duracion_media: { $avg: '$duracion' }
    },
    {
      $project: {
        _id: 0,
        'director': '$_id',
        'peliculas': 1,
        'ingresos_totales': 1,
        'duracion_media': 1
      }
    },
    { $sort: { duracion_media: -1 } },
    { $limit: 10 }
  }
})

```

Resultado:

```

{ "peliculas" : 2, "ingresos_totales" : 37923936, "duracion_media" : 234, "director" : "Ronald F. Maxwell" }
{ "peliculas" : 1, "ingresos_totales" : 393816, "duracion_media" : 287, "director" : "Sergei Loban" }
{ "peliculas" : 3, "ingresos_totales" : 6613926, "duracion_media" : 199.33333333333334, "director" : "Anurag Kashyap" }
{ "peliculas" : 3, "ingresos_totales" : 37280000, "duracion_media" : 197.33333333333334, "director" : "Ashutosh Gowariker" }
{ "peliculas" : 1, "ingresos_totales" : 11000000, "duracion_media" : 193, "director" : "D.W. Griffith" }
{ "peliculas" : 2, "ingresos_totales" : 71063463, "duracion_media" : 193, "director" : "Joseph L. Mankiewicz" }
{ "peliculas" : 1, "ingresos_totales" : 100000000, "duracion_media" : 190, "director" : "Aditya Chopra" }
{ "peliculas" : 2, "ingresos_totales" : 93900000, "duracion_media" : 190, "director" : "Sooraj R. Barjatya" }
{ "peliculas" : 1, "ingresos_totales" : 1345903, "duracion_media" : 189, "director" : "Hector Babenco" }
{ "peliculas" : 2, "ingresos_totales" : 40385320, "duracion_media" : 185.5, "director" : "Yash Chopra" }

```