

# 1. 环境准备

本项目需要完成如下的环境配置和数据准备

## 1.1. 本项目需要以下环境

- Docker
- Java 8
- nodejs
- npm
- Maven
- [mongosh](#)
- [MongoDB Database Tools](#)
- [Mongo Compass](#)
- [Studio 3T](#)
- [Prometheus 2.30.3](#)
- [Grafana 8.4.2](#)

## 1.2. 数据准备

- 将 db-generation 解压后放到 \${PROJECT}/env 下
- 将 db-generation 重命名为 data

# 2. 部署方法

## 2.1. 包结构

本项目的包结构如下

```
└─ ${PROJECT}
    ├─ frontend 前端代码
    ├─ backend 后端代码
    └─ env 脚本及配置文件
        ├─ data 数据文件位置
        ├─ grafana-dashboard Grafana 面板位置
        ├─ hdfs HDFS配置文件
        ├─ mongodb MongoDB配置文件
        └─ scripts 脚本文件
```

## 2.2. 脚本介绍

- /env/script 目录
  - setup.sh、setup.bat：初始化并启动容器，包括数据加载等，用于第一次启动
  - shutdown.sh、shutdown.bat：停止容器的运行，并删除容器
  - start.sh、start.bat：开始容器的运行（第一次启动请使用 [setup.sh](#)）
  - stop.sh、stop.bat：停止容器的运行

## 2.3. 部署流程

### 2.3.1. 数据生成

使用脚本：/env/data-generate/genTable\_mongoDB10G.py

```
# 进入到相应目录，然后运行 genTable_mongoDB10G.py
python3 genTable_mongoDB10G.py
```

### 2.3.2. 初始化服务

首次启动：使用脚本 /script/setup.sh

```
# 进入到 /script 目录，然后运行 setup.sh
./setup.sh
```

setup.sh 脚本说明：

```

#!/bin/bash

# hdfs
docker-compose -f ./hdfs/docker-compose.yml up -d
sleep 10
# 创建目录, 将 /data/articles/ 上传至 hdfs
docker exec -it namenode hadoop fs -mkdir /articles/
docker exec -it namenode hadoop fs -put /data/articles/ /articles/

# redis
docker pull redis:7.0.0
docker run -itd --name redis -p 6379:6379 redis
docker network connect hdfs_hadoopnet redis --ip 172.21.0.6

# mongodb
# config servers
docker-compose -f ./mongodb/docker-compose-configsvr.yml up -d
mongosh --host localhost --port 40011 ./mongodb/setup-configsvr.js
# shard servers
docker-compose -f ./mongodb/docker-compose-shardsvr.yml up -d
mongosh --host localhost --port 40021 ./mongodb/setup-shardsvr-rs1.js
mongosh --host localhost --port 40031 ./mongodb/setup-shardsvr-rs2.js
# mongos
docker-compose -f ./mongodb/docker-compose-mongos.yml up -d
sleep 5
mongosh --host localhost --port 40002 ./mongodb/setup-mongos.js
docker network connect hdfs_hadoopnet mongos --ip 172.21.0.7

# load data
# 使用 mongoimport 直接导入生成的数据
mongoimport --host localhost --port 40002 -d ddbms -c user --file ../data/user.dat
mongoimport --host localhost --port 40002 -d ddbms -c article --file ../data/article.dat
mongoimport --host localhost --port 40002 -d ddbms -c read --file ../data/read.dat
# 使用 js 脚本生成 be-read 和 rank 表
mongosh --host localhost --port 40002 ../mongodb/generate-beread.js
mongosh --host localhost --port 40002 ../mongodb/generate-rank.js

```

### 2.3.3. 启动前端

进入 /frontend 文件夹, 执行如下命令本地编译:

```
npm install
```

执行如下命令启动前端

```
npm run dev
```

然后可以在localhost:8080访问到前端

## 2.3.4. 查看监控

以 MacOS 为例

### 2.3.4.1. Prometheus

1. 下载Prometheus的二进制包到本地，解压后进入对应文件夹：

```
tar xvfz prometheus-*.tar.gz  
cd prometheus-*
```

2. 修改Prometheus的配置文件prometheus.yml文件如下：

```
global:  
  scrape_interval: 15s  
  
scrape_configs:  
  - job_name: "prometheus"  
    static_configs:  
      - targets: ["localhost:9090"]  
  - job_name: "dbms"  
    static_configs:  
      - targets: ["localhost:8088"]
```

3. 启动 Prometheus：

```
./prometheus --config.file=prometheus.yml
```

4. 访问 <http://localhost:9090>，进入Status下的Target界面查看联通情况

## 2.3.4.2. Grafana

1. 下载Grafana的二进制包到本地，解压后进入对应文件夹：

```
tar -zxvf grafana-*.tar.gz  
cd grafana-*
```

2. 启动Grafana

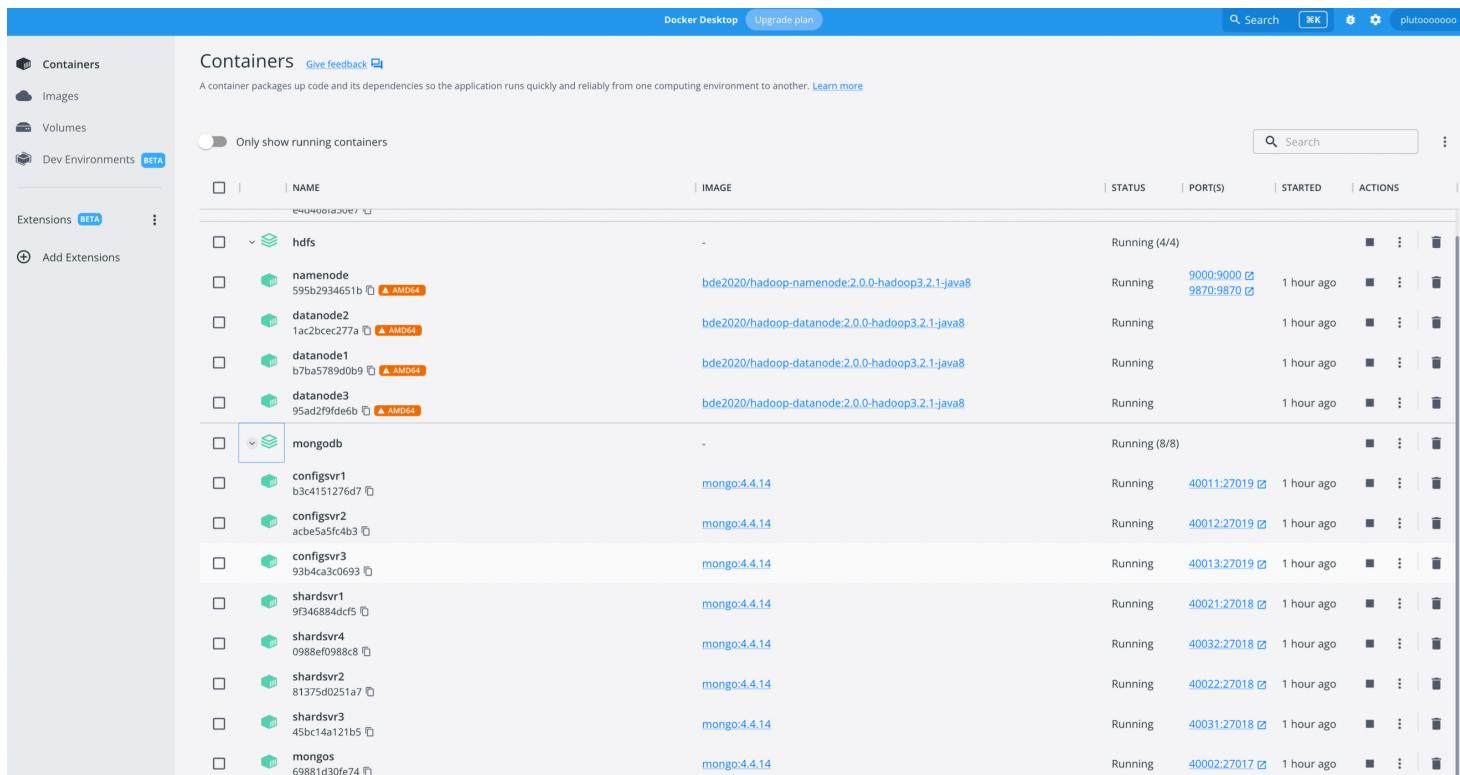
```
./bin/grafana-server web
```

3. 访问 <http://localhost:3000>，进入Grafana
4. 配置对应的Data Source为刚刚安装好的Prometheus的位置
5. 在Dashboard界面，Import面板的配置 /env/grafana-dashboard/DBMS Dashboard.json，然后选择对应的job和instance后即可访问

## 2.4. 环境查看说明

### 2.4.1. Mongo

首先在 Docker Desktop 中 进入 mongos 容器：



The screenshot shows the Docker Desktop interface with the following details:

NAME	IMAGE	STATUS	PORT(S)	STARTED	ACTIONS
namenode	bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8	Running	9000:9000, 9870:9870	1 hour ago	[Actions]
datanode2	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	Running		1 hour ago	[Actions]
datanode1	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	Running		1 hour ago	[Actions]
datanode3	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	Running		1 hour ago	[Actions]
configsvr1	mongo:4.4.14	Running	40011:27019	1 hour ago	[Actions]
configsvr2	mongo:4.4.14	Running	40012:27019	1 hour ago	[Actions]
configsvr3	mongo:4.4.14	Running	40013:27019	1 hour ago	[Actions]
shardsvr1	mongo:4.4.14	Running	40021:27018	1 hour ago	[Actions]
shardsvr4	mongo:4.4.14	Running	40032:27018	1 hour ago	[Actions]
shardsvr2	mongo:4.4.14	Running	40022:27018	1 hour ago	[Actions]
shardsvr3	mongo:4.4.14	Running	40031:27018	1 hour ago	[Actions]
mongos	mongo:4.4.14	Running	40002:27017	1 hour ago	[Actions]

#### 2.4.1.1. 查看数据

在 Terminal 中 运行如下命令可以查看导入的数据

```
# 进入 mongodb
```

```
mongo
```

```
# 查看 db
```

```
show dbs
```

```
# 使用 ddbms 并查询 user 表
```

```
use ddbms
```

```
db.user.find()
```

```
Docker Desktop Upgrade plan Search ⓘ
Containers Images Volumes Dev Environments Extensions Add Extensions
Logs Inspect Terminal Stats
mongos mongo:4.4.14 RUNNING
MongoDB shell version v4.4.14
connecting to: mongodb://127.0.0.1:27017/?compressors=disabled&gssapiServiceName=mongodb
Implicit session: session { "id" : UUID("e677e08b-46b2-468d-8551-7916390ba565") }
MongoDB server version: 4.4.14
Welcome to the MongoDB shell.
For interactive help, type "help".
For more comprehensive documentation, see
    https://docs.mongodb.com
Questions? Try the MongoDB Developer Community Forums
    https://community.mongodb.com
The server generated these startup warnings when booting:
2022-11-27T14:33:46.984+00:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
...
mongo> show dbs
switched to db ddbms
...
{
  "_id": ObjectId("63b3756a393677ab59706568"),
  "timestamp": "1506328859011",
  "id": "u11",
  "uid": "11",
  "name": "user11",
  "gender": "male",
  "email": "email11",
  "phone": "phone11",
  "dept": "dept8",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag10",
  "obtainedCredits": 26
}
{
  "_id": ObjectId("63b3756a393677ab59706569"),
  "timestamp": "1506328859012",
  "id": "u12",
  "uid": "12",
  "name": "user12",
  "gender": "male",
  "email": "email12",
  "phone": "phone12",
  "dept": "dept13",
  "grade": "grade2",
  "language": "en",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag11",
  "obtainedCredits": 5
}
{
  "_id": ObjectId("63b3756a393677ab5970656a"),
  "timestamp": "1506328859014",
  "id": "u14",
  "uid": "14",
  "name": "user14",
  "gender": "male",
  "email": "email14",
  "phone": "phone14",
  "dept": "dept12",
  "grade": "grade3",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag40",
  "obtainedCredits": 66
}
{
  "_id": ObjectId("63b3756a393677ab5970656b"),
  "timestamp": "1506328859013",
  "id": "u13",
  "uid": "13",
  "name": "user13",
  "gender": "male",
  "email": "email13",
  "phone": "phone13",
  "dept": "dept4",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag19",
  "obtainedCredits": 21
}
{
  "_id": ObjectId("63b3756a393677ab5970656c"),
  "timestamp": "1506328859015",
  "id": "u16",
  "uid": "16",
  "name": "user16",
  "gender": "female",
  "email": "email16",
  "phone": "phone16",
  "dept": "dept4",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag35",
  "obtainedCredits": 85
}
{
  "_id": ObjectId("63b3756a393677ab5970656e"),
  "timestamp": "1506328859018",
  "id": "u18",
  "uid": "18",
  "name": "user18",
  "gender": "male",
  "email": "email18",
  "phone": "phone18",
  "dept": "dept15",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag49",
  "obtainedCredits": 16
}
{
  "_id": ObjectId("63b3756a393677ab5970656f"),
  "timestamp": "1506328859019",
  "id": "u19",
  "uid": "19",
  "name": "user19",
  "gender": "male",
  "email": "email19",
  "phone": "phone19",
  "dept": "dept8",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag31",
  "obtainedCredits": 10
}
{
  "_id": ObjectId("63b3756a393677ab59706570"),
  "timestamp": "1506328859020",
  "id": "u20",
  "uid": "20",
  "name": "user20",
  "gender": "male",
  "email": "email20",
  "phone": "phone20",
  "dept": "dept17",
  "grade": "grade2",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role2",
  "preferTags": "tag32",
  "obtainedCredits": 69
}
{
  "_id": ObjectId("63b3756a393677ab59706571"),
  "timestamp": "1506328859021",
  "id": "u23",
  "uid": "23",
  "name": "user23",
  "gender": "female",
  "email": "email23",
  "phone": "phone23",
  "dept": "dept19",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag18",
  "obtainedCredits": 8
}
{
  "_id": ObjectId("63b3756a393677ab59706574"),
  "timestamp": "1506328859024",
  "id": "u24",
  "uid": "24",
  "name": "user24",
  "gender": "male",
  "email": "email24",
  "phone": "phone24",
  "dept": "dept7",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag20",
  "obtainedCredits": 13
}
{
  "_id": ObjectId("63b3756a393677ab59706575"),
  "timestamp": "1506328859025",
  "id": "u25",
  "uid": "5",
  "name": "user5",
  "gender": "male",
  "email": "email15",
  "phone": "phone5",
  "dept": "dept10",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag5",
  "obtainedCredits": 90
}
{
  "_id": ObjectId("63b3756a393677ab59706576"),
  "timestamp": "1506328859017",
  "id": "u17",
  "uid": "17",
  "name": "user17",
  "gender": "male",
  "email": "email11",
  "phone": "phone17",
  "dept": "dept15",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag31",
  "obtainedCredits": 82
}
{
  "_id": ObjectId("63b3756a393677ab59706577"),
  "timestamp": "1506328859003",
  "id": "u4",
  "uid": "4",
  "name": "user4",
  "gender": "male",
  "email": "email14",
  "phone": "phone14",
  "dept": "dept6",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag11",
  "obtainedCredits": 11
}
{
  "_id": ObjectId("63b3756a393677ab59706578"),
  "timestamp": "1506328859003",
  "id": "u3",
  "uid": "3",
  "name": "user3",
  "gender": "male",
  "email": "email13",
  "phone": "phone13",
  "dept": "dept2",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role2",
  "preferTags": "tag34",
  "obtainedCredits": 60
}
{
  "_id": ObjectId("63b3756a393677ab59706579"),
  "timestamp": "1506328859026",
  "id": "u2",
  "uid": "2",
  "name": "user2",
  "gender": "male",
  "email": "email12",
  "phone": "phone2",
  "dept": "dept13",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag17",
  "obtainedCredits": 77
}
{
  "_id": ObjectId("63b3756a393677ab59706577"),
  "timestamp": "1506328859028",
  "id": "u28",
  "uid": "28",
  "name": "user28",
  "gender": "male",
  "email": "email128",
  "phone": "phone28",
  "dept": "dept3",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role0",
  "preferTags": "tag20",
  "obtainedCredits": 49
}
{
  "_id": ObjectId("63b3756a393677ab59706584"),
  "timestamp": "1506328859030",
  "id": "u32",
  "uid": "32",
  "name": "user32",
  "gender": "male",
  "email": "email32",
  "phone": "phone32",
  "dept": "dept17",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role1",
  "preferTags": "tag35",
  "obtainedCredits": 88
}
{
  "_id": ObjectId("63b3756a393677ab59706585"),
  "timestamp": "1506328859033",
  "id": "u33",
  "uid": "33",
  "name": "user33",
  "gender": "male",
  "email": "email33",
  "phone": "phone33",
  "dept": "dept14",
  "grade": "gradel",
  "language": "en",
  "region": "Hong Kong",
  "role": "role0",
  "preferTags": "tag17",
  "obtainedCredits": 23
}
{
  "_id": ObjectId("63b3756a393677ab59706586"),
  "timestamp": "1506328859034",
  "id": "u34",
  "uid": "34",
  "name": "user34",
  "gender": "male",
  "email": "email34",
  "phone": "phone34",
  "dept": "dept11",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role0",
  "preferTags": "tag28",
  "obtainedCredits": 24
}
{
  "_id": ObjectId("63b3756a393677ab59706587"),
  "timestamp": "1506328859036",
  "id": "u36",
  "uid": "36",
  "name": "user36",
  "gender": "female",
  "email": "email136",
  "phone": "phone36",
  "dept": "dept0",
  "grade": "gradel",
  "language": "zh",
  "region": "Hong Kong",
  "role": "role0",
  "preferTags": "tag28",
  "obtainedCredits": 24
}
Type "it" for more
manan>
```

## 2.4.1.2. 查看数据分片

```
mongos> sh.status()
--- Sharding Status ---
sharding version: {
    "_id" : 1,
    "minCompatibleVersion" : 5,
    "currentVersion" : 6,
    "clusterId" : ObjectId("6383754ed5180dab2a303515")
}
shards:
{ "_id" : "shardsvr_rs1", "host" : "shardsvr_rs1/shardsvr1:27018,shardsvr2:27018", "state" : 1, "tags" : [ "DBMS1", "DBMS1-2" ] }
{ "_id" : "shardsvr_rs2", "host" : "shardsvr_rs2/shardsvr3:27018,shardsvr4:27018", "state" : 1, "tags" : [ "DBMS2", "DBMS1-2" ] }
active mongoses:
"4.4.14" : 1
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:
        No recent migrations
databases:
{ "_id" : "config", "primary" : "config", "partitioned" : true }
{ "_id" : "ddbsms", "primary" : "shardsvr_rs1", "partitioned" : true, "version" : { "uuid" : UUID("51ad99c4-def4-41d0-a11d-8e6eb6ed9e76"), "lastMod" : 1 } }
    ddbsms.article
        shard key: { "category" : 1, "aid" : 1 }
        unique: false
        balancing: true
        chunks:
            shardsvr_rs1 3
            shardsvr_rs2 2
{ "category" : { "$minKey" : 1 }, "aid" : { "$minKey" : 1 } } -->> { "category" : "science", "aid" : { "$minKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 0)
{ "category" : "science", "aid" : { "$minKey" : 1 } } -->> { "category" : "science", "aid" : { "$maxKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 1)
{ "category" : "science", "aid" : { "$maxKey" : 1 } } -->> { "category" : "technology", "aid" : { "$minKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 2)
{ "category" : "technology", "aid" : { "$minKey" : 1 } } -->> { "category" : "technology", "aid" : { "$maxKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 3)
{ "category" : "technology", "aid" : { "$maxKey" : 1 } } -->> { "category" : "science", "aid" : { "$maxKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 4)
tag: DBMS1-2 { "category" : "science", "aid" : { "$minKey" : 1 } } -->> { "category" : "science", "aid" : { "$maxKey" : 1 } }
tag: DBMS2 { "category" : "technology", "aid" : { "$minKey" : 1 } } -->> { "category" : "technology", "aid" : { "$maxKey" : 1 } }
    ddbsms.beread
        shard key: { "category" : 1, "_id" : 1 }
        unique: false
        balancing: true
        chunks:
            shardsvr_rs1 3
            shardsvr_rs2 2
{ "category" : { "$minKey" : 1 }, "id" : { "$minKey" : 1 } } -->> { "category" : "science", "id" : { "$minKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 0)
{ "category" : "science", "id" : { "$minKey" : 1 } } -->> { "category" : "science", "id" : { "$maxKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 1)
{ "category" : "science", "id" : { "$maxKey" : 1 } } -->> { "category" : "technology", "id" : { "$minKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 2)
{ "category" : "technology", "id" : { "$minKey" : 1 } } -->> { "category" : "technology", "id" : { "$maxKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 3)
{ "category" : "technology", "id" : { "$maxKey" : 1 } } -->> { "category" : "science", "id" : { "$maxKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 4)
tag: DBMS1-2 { "category" : "science", "id" : { "$minKey" : 1 } } -->> { "category" : "science", "id" : { "$maxKey" : 1 } }
tag: DBMS2 { "category" : "technology", "id" : { "$minKey" : 1 } } -->> { "category" : "technology", "id" : { "$maxKey" : 1 } }
    ddbsms.rank
        shard key: { "temporal_granularity" : 1, "_id" : 1 }
        unique: false
        balancing: true
        chunks:
            shardsvr_rs1 3
            shardsvr_rs2 4
{ "temporal_granularity" : { "$minKey" : 1 }, "id" : { "$minKey" : 1 } } -->> { "temporal_granularity" : "daily", "id" : { "$minKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 0)
{ "temporal_granularity" : "daily", "id" : { "$minKey" : 1 } } -->> { "temporal_granularity" : "daily", "id" : { "$maxKey" : 1 } } on : shardsvr_rs1 Timestamp(1, 1)
{ "temporal_granularity" : "daily", "id" : { "$maxKey" : 1 } } -->> { "temporal_granularity" : "monthly", "id" : { "$minKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 2)
{ "temporal_granularity" : "monthly", "id" : { "$minKey" : 1 } } -->> { "temporal_granularity" : "monthly", "id" : { "$maxKey" : 1 } } on : shardsvr_rs2 Timestamp(1, 3)
```

## 2.4.2. Hadoop

访问 localhost:9870 可以看到数据成功导入

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

### Browse Directory

/articles/articles

File	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	Action
..	-rw-r--r--	root	supergroup	12 KB	Nov 27 21:18	3	128 MB	.DS_Store	trash
article0	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article0	trash
article1	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article1	trash
article10	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article10	trash
article11	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article11	trash
article12	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article12	trash
article13	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article13	trash
article14	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article14	trash
article15	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article15	trash
article16	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article16	trash
article17	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article17	trash
article18	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article18	trash
article19	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article19	trash
article2	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article2	trash
article20	drwxr-xr-x	root	supergroup	0 B	Nov 27 21:18	0	0 B	article20	trash