

Lab #8

👤 Выполнил:	Ю Юрлов Павел ИАТ22-1м
@ FA E-mail:	221759@edu.fa.ru
📎 Telegram:	t.me/paulyurlov

Задачи лабораторной:

1. Создайте коллекцию документов для обработки её с помощью MapReduce.
2. Наполните коллекцию документами.
3. Произведите обработку коллекции с использованием модели распределенных вычислений MapReduce.

Ход выполнения работы:

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2...
nosql_curse> db.selling.mapReduce(map, reduce, {out: {inline: 1}})
DeprecationWarning: Collection.mapReduce() is deprecated. Use an aggregation instead.
See https://docs.mongodb.com/manual/core/map-reduce for details.
{
  results: [
    {
      _id: { name: 'Nexus One', year: 2013, month: 0, day: 21 },
      value: { count: 2 }
    },
    {
      _id: { name: 'iPhone 4', year: 2013, month: 0, day: 20 },
      value: { count: 1 }
    },
    {
      _id: { name: 'Nexus One', year: 2013, month: 0, day: 22 },
      value: { count: 1 }
    },
    {
      _id: { name: 'iPhone 4', year: 2013, month: 0, day: 21 },
      value: { count: 3 }
    },
    {
      _id: { name: 'Nexus One', year: 2013, month: 0, day: 20 },
```

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
nosql_curse> db.sellingResult.find()
[
  {
    _id: { name: 'Nexus One', year: 2013, month: 0, day: 21 },
    value: { count: 2 }
  },
  {
    _id: { name: 'iPhone 4', year: 2013, month: 0, day: 20 },
    value: { count: 1 }
  },
  {
    _id: { name: 'Nexus One', year: 2013, month: 0, day: 22 },
    value: { count: 1 }
  },
  {
    _id: { name: 'iPhone 4', year: 2013, month: 0, day: 21 },
    value: { count: 3 }
  },
  {
    _id: { name: 'Nexus One', year: 2013, month: 0, day: 20 },
    value: { count: 3 }
  }
]
nosql_curse> 
```

1. Создайте коллекцию документов для обработки её с помощью MapReduce.

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2...
nosql_curse> db.createCollection('students')
{ ok: 1 }
nosql_curse> 
```

2. Наполните коллекцию документами.

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTime...
nosql_curse> db.students.insertMany([
... {"name": "Midhu", "subject": "science", "marks": 68 },
... {"name": "Midhu", "subject": "maths", "marks": 98 },
... {"name": "Midhu", "subject": "sports", "marks": 77 },
... {"name": "Akhil", "subject": "science", "marks": 67 },
... {"name": "Akhil", "subject": "maths", "marks": 87 },
... {"name": "Akhil", "subject": "sports", "marks": 89 },
... {"name": "Anish", "subject": "science", "marks": 67 },
... {"name": "Anish", "subject": "maths", "marks": 78 },
... {"name": "Anish", "subject": "sports", "marks": 90 }
... ])
{
  acknowledged: true,
  insertedIds: {
    '0': ObjectId("63c00b49cb3be1635eb0f701"),
    '1': ObjectId("63c00b49cb3be1635eb0f702"),
    '2': ObjectId("63c00b49cb3be1635eb0f703"),
    '3': ObjectId("63c00b49cb3be1635eb0f704"),
    '4': ObjectId("63c00b49cb3be1635eb0f705"),
    '5': ObjectId("63c00b49cb3be1635eb0f706"),
    '6': ObjectId("63c00b49cb3be1635eb0f707"),
    '7': ObjectId("63c00b49cb3be1635eb0f708"),
    '8': ObjectId("63c00b49cb3be1635eb0f709")
  }
}
nosql_curse> 
```

3. Произведите обработку коллекции с использованием модели распределенных вычислений MapReduce.

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeout...
}
nosql_curse> var map = function() {emit(this.name,this.marks)};

nosql_curse> var reduce = function(name,marks) {return Array.sum(marks)};

nosql_curse> db.students.mapReduce(
...   map,
...   reduce,
...   { out: "totals" }
... );
{ result: 'totals', ok: 1 }
nosql_curse> db.students.find({})
```

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeout...
  name: 'Anish',
  subject: 'sports',
  marks: 90
}
]
nosql_curse> db.totals.find({})
[
  { _id: 'Akhil', value: 243 },
  { _id: 'Midhu', value: 243 },
  { _id: 'Anish', value: 235 }
]
nosql_curse> []
```

Теоретические вопросы

1. Что означает термин MapReduce?

MapReduce - модель распределённых вычислений, используемая для параллельных вычислений над очень большими наборами данных

2. Из каких шагов состоит работа MapReduce?

Map-шаг и Reduce-шаг

3. Какими преимуществами обладает MapReduce по сравнению с обычными вычислениями?

Высокая производительность за счет распределения вычислений по кластерам

4. Опишите работу с MapReduce в MongoDB.

Для работы с MapReduce в MongoDB используется команда MapReduce со следующим синтаксисом:

```
db.runCommand(
{
  mapReduce: <string>,
  map: <string or JavaScript>,
  reduce: <string or JavaScript>,
  finalize: <string or JavaScript>,
  out: <output>,
  query: <document>,
  sort: <document>,
  limit: <number>,
  scope: <document>,
  jsMode: <boolean>,
  verbose: <boolean>,
  bypassDocumentValidation: <boolean>,
  collation: <document>,
  writeConcern: <document>,
  comment: <any>
```

```
}  
)
```