**目录**

[1. 概述 1](#_Toc465784763)

[2. 数据导出导入操作 1](#_Toc465784764)

[2.1数据导出 1](#_Toc465784765)

[2.1.1导出为json 1](#_Toc465784766)

[2.1.2导出为csv 1](#_Toc465784767)

[2.1.3带过滤条件导出 1](#_Toc465784768)

[2.2导入数据 1](#_Toc465784769)

[2.2.1导入json格式 1](#_Toc465784770)

[2.2.2导入csv格式 1](#_Toc465784771)

[3. 数据备份与恢复 1](#_Toc465784772)

[3.1数据备份 1](#_Toc465784773)

[3.1.1备份一个collection 1](#_Toc465784774)

[3.1.2备份一个数据库 1](#_Toc465784775)

[3.1.3无用户名密码备份 1](#_Toc465784776)

[3.1.4mongodump相关的问题 1](#_Toc465784777)

[3.2数据恢复 1](#_Toc465784778)

[3.2.1恢复一个数据库 1](#_Toc465784779)

[3.2.2恢复一个collection 1](#_Toc465784780)

[4. 再说点什么 1](#_Toc465784781)

[4.1导入导出备份恢复应该考虑的问题 1](#_Toc465784782)

[4.2json丢失数据精度 1](#_Toc465784783)

[4.3导入导出备份恢复有关的锁 1](#_Toc465784784)

[5. 参考资料 1](#_Toc465784785)

**jar cvfm MyWordCount.jar manifest.data \*.class**

# 概述

数据库一旦大动，数据库数据的导入、导出、备份、还原操作就显得尤为重要，本篇介绍MongoDB的数据导入、导出、本分、还原操作。

MongoDB对于数据的导出还提供了条件过滤等高级操作，非常好！下面所有命令测试的MongoDB版本是2.4.9。

# 数据导出导入操作

## 2.1数据导出

数据导出命令mongoexport，常用参数解析如下：

|  |  |  |
| --- | --- | --- |
| 参数选项 | 含义 | 其他 |
| -h | 主机ip |  |
| -port | 端口 |  |
| -u | 用户名 |  |
| -p | 密码 |  |
| -d | db |  |
| -c | collection |  |
| -f | 列，指定导出某些列 |  |
| -q | 过滤条件 |  |
| --csv | csv格式 |  |
| -o | 输出文件 |  |
| --forceTableScan | 不用快照，全表扫描 |  |
| --help | 显示帮助信息 |  |
| -v/--version | 显示版本信息 |  |
| --ssl | 所有链接使用ssl |  |
| --limit，--skip | 二者配合使用，导出数目的文档 | mongoexport -d test -c records --sort '{a: 1}' --limit 100 --out export.0.json  mongoexport -d test -c records --sort '{a: 1}' --limit 100 --skip 100 --out export.1.json  mongoexport -d test -c records --sort '{a: 1}' --limit 100 --skip 200 --out export.2.json  但是--limit，--skip会不会影响效率倒没有测试。 |
|  |  |  |

不同的版本也有不同的参数支持，具体视个版本而定。

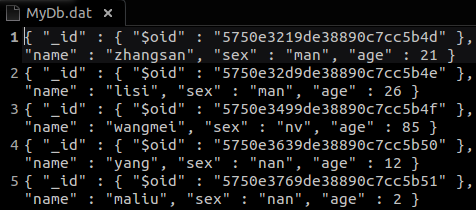
### 2.1.1导出为json

**lt@lt-vm:~/桌面$ mongoexport -d MyDb -c TestCreateTime -o MyDb.dat**

**connected to: 127.0.0.1**

**exported 5 records**

导出后的MyDb.dat文件时一行一行的json格式串，如下：



### 2.1.2导出为csv

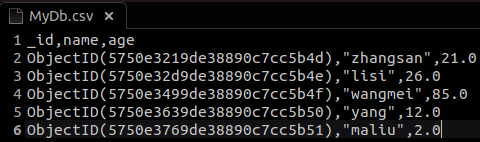
导出csv格式，使用的参数选项是--csv -f，注意-f之后必须指定所要导出的列名（列名必须存在）。

**lt@lt-vm:~/桌面$ mongoexport --csv -f \_id,name,age -d MyDb -c TestCreateTime -o MyDb.csv**

**connected to: 127.0.0.1**

**exported 5 records**

导出后的数据格式如下：



### 2.1.3带过滤条件导出

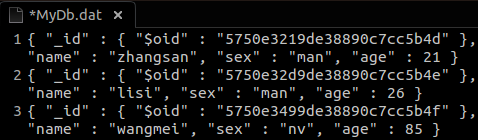
带过滤条件的导出是我们经常使用的一个功能，如下操作，导出age大于等于20的记录：

**lt@lt-vm:~/桌面$ mongoexport -d MyDb -c TestCreateTime -q '{age:{"$gte":20}}' -o MyDb.dat**

**connected to: 127.0.0.1**

**exported 3 records**

导出的记录信息如下：



## 2.2导入数据

导出数据的工具是mongoimport，常用参数选项与mongoexport相同，不同参数如下：

|  |  |  |
| --- | --- | --- |
| 参数选项 | 含义 | 其他 |
| -h | 主机ip |  |
| -port | 端口 |  |
| -u | 用户名 |  |
| -p | 密码 |  |
| -d | db |  |
| -c | collection |  |
| -f | 列 |  |
| --help | 显示帮助信息 |  |
| -v/--version | 显示版本信息 |  |
| --ssl | 所有链接使用ssl |  |
| --dbpath | 直接访问mongod数据库文件，不用连接mongod服务，该操作需要锁定参数指定的数据目录，所以如果mongod正在访问该目录，那么将无法操作。 | 该选项在>=3.0的版本中已经去掉了，不建议使用这个选项。 |
| --directoryperdb | 与dbpath配合，每个db指定一个目录 |  |
| -type | 导入文件的格式，json、csv、tsv，默认json |  |
| --file | 要导入的文件 |  |
| --drop | 首先删除collection |  |
| --upsert | 如果已经存在，则插入或者更新对象 |  |
| --headerline | csv/tsv文件第一行是头 |  |
| … … | … … | … … |

### 2.2.1导入json格式

**lt@lt-vm:~/桌面$ mongoimport -d MyDb -c TestCreateTime MyDb.dat**

**connected to: 127.0.0.1**

**Fri Jun 3 16:56:06.537 imported 5 objects**

### 2.2.2导入csv格式

**lt@lt-vm:~/桌面$ mongoimport -d MyDb -c TestCreateTime --type csv --headerline --file MyDb.csv**

**connected to: 127.0.0.1**

**Fri Jun 3 16:57:39.321 imported 5 objects**

**csv格式很好****，主流数据库都支持该类格式的迁移。**

# 数据备份与恢复

## 3.1数据备份

MongoDB的数据备份命令是mongodump，参数与mongoexport等类似。mongodump可以备份数据库级别或表级别的内容。mongodump工具的参数选项如下：

|  |  |  |
| --- | --- | --- |
| 参数选项 | 含义 | 其他 |
| --help | 查看帮助信息 |  |
| -v，--version | 查看版本信息 |  |
| -h | 主机ip |  |
| --port | 端口号 |  |
| --ssl | 所有的连接使用ssl |  |
| -u | 用户名 |  |
| -p | 密码 |  |
| -d | db，数据库信息 |  |
| -c | collection，文档信息 |  |
| -o | 输出位置 |  |
| -q | json格式的过滤条件 |  |
| --repair | 尝试恢复crashed的数据库 |  |
| --forceTableScan | 不使用快照，强制扫描全表 |  |
| --dbpath | 直接访问mongod数据库文件，不用连接mongod服务，该操作需要锁定参数指定的数据目录，所以如果mongod正在访问该目录，那么将无法操作。 | 该选项在>=3.0的版本中已经去掉了，不建议使用这个选项。 |

### 3.1.1备份一个collection

备份在当前目录：

**lt@lt-vm:~/桌面$ mongodump -d MyDb -c TestCreateTime**

**connected to: 127.0.0.1**

**Fri Jun 3 17:03:05.597 DATABASE: MyDb to dump/MyDb**

**Fri Jun 3 17:03:05.609 MyDb.TestCreateTime to dump/MyDb/TestCreateTime.bson**

**Fri Jun 3 17:03:05.610 5 objects**

**Fri Jun 3 17:03:05.610 Metadata for MyDb.TestCreateTime to dump/MyDb/TestCreateTime.metadata.json**

如此命令会在当前目录生成一个dump目录，结构如下：



### 3.1.2备份一个数据库

备份到指定目录：

**lt@lt-vm:~/桌面$ ls**

**MyDb.csv MyDb.dat**

**lt@lt-vm:~/桌面$ mongodump -d MyDb -o /home/lt/桌面/dump**

**connected to: 127.0.0.1**

**Fri Jun 3 17:07:23.407 DATABASE: MyDb to /home/lt/桌面/dump/MyDb**

**Fri Jun 3 17:07:23.408 MyDb.system.indexes to /home/lt/桌面/dump/MyDb/system.indexes.bson**

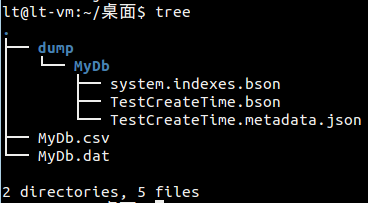
**Fri Jun 3 17:07:23.408 1 objects**

**Fri Jun 3 17:07:23.408 MyDb.TestCreateTime to /home/lt/桌面/dump/MyDb/TestCreateTime.bson**

**Fri Jun 3 17:07:23.408 5 objects**

**Fri Jun 3 17:07:23.408 Metadata for MyDb.TestCreateTime to /home/lt/桌面/dump/MyDb/TestCreateTime.metadata.json**

目录结构如下：



### 3.1.3无用户名密码备份

没有用户名和密码，可以忽略-u、-p选项。导出本机数据库忽略-h、-p选项。默认端口忽略--port选项。导出所有数据库忽略-d选项，备份所有数据库：

**lt@lt-vm:~/桌面$ mongodump -o ./alldb/**

**connected to: 127.0.0.1**

**Fri Jun 3 17:54:34.702 all dbs**

**Fri Jun 3 17:54:34.702 DATABASE: MyDb to ./alldb/MyDb**

**Fri Jun 3 17:54:34.703 MyDb.system.indexes to ./alldb/MyDb/system.indexes.bson**

**Fri Jun 3 17:54:34.703 1 objects**

**Fri Jun 3 17:54:34.703 MyDb.TestCreateTime to ./alldb/MyDb/TestCreateTime.bson**

**Fri Jun 3 17:54:34.707 5 objects**

**Fri Jun 3 17:54:34.707 Metadata for MyDb.TestCreateTime to ./alldb/MyDb/TestCreateTime.metadata.json**

### 3.1.4mongodump相关的问题

（1）mongodump**所备份的数据不包含索引**，因此当你备份完导入到新的数据库之后，需要重新创建索引；

（2）mongodump对于[2.2, +]之后的数据格式与[-, 2. 2)之前的格式不兼容，也就是如果你使用mongodump 2.2之后的版本，是无法备份2.2版本之前的数据的。最好使用同等版本的工具进行操作；

（3）

## 3.2数据恢复

数据恢复命令是mongorestore，参数基本都类似。

### 3.2.1恢复一个数据库

**lt@lt-vm:~/桌面$ mongorestore -d MyDb ./dump/MyDb/**

**connected to: 127.0.0.1**

**Fri Jun 3 17:11:19.656 ./dump/MyDb/TestCreateTime.bson**

**Fri Jun 3 17:11:19.657 going into namespace [MyDb.TestCreateTime]**

**5 objects found**

**Fri Jun 3 17:11:19.664 Creating index: { key: { \_id: 1 }, ns: "MyDb.TestCreateTime", name: "\_id\_" }**

如果想要覆盖之前的数据，恢复可以加上--drop参数，如下：

**lt@lt-vm:~/桌面$ mongorestore -d MyDb --drop ./dump/MyDb/**

加上该参数就会在备份的时候先删除之前的数据然后进行恢复了。

不加任务选项恢复所有数据库：

**lt@lt-vm:~/桌面$ mongorestore alldb/**

**connected to: 127.0.0.1**

**Fri Jun 3 18:00:17.964 alldb/MyDb/TestCreateTime.bson**

**Fri Jun 3 18:00:17.964 going into namespace [MyDb.TestCreateTime]**

**5 objects found**

**Fri Jun 3 18:00:17.965 Creating index: { key: { \_id: 1 }, ns: "MyDb.TestCreateTime", name: "\_id\_" }**

### 3.2.2恢复一个collection

单个collection的恢复需要用mongoimport来恢复，如下：

**lt@lt-vm:~/桌面$ mongoimport -d MyDb -c TestCreateTime MyDb.dat**

# 再说点什么

## 4.1导入导出备份恢复应该考虑的问题

（1）如果这些操作对于线上有影响，适当的选择在业务访问不频繁的时间段；

（2）标记文件，便于我们确定所操作内容的时间点；

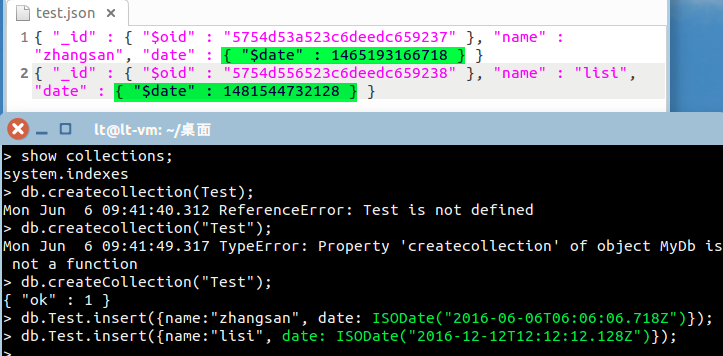
（3）确保数据的一致性状态，如果备份过程中数据不断更新的话，要确保数据的完整性和一致性；

（4）通过恢复和导入操作，来测试备份和导出的数据是可用的；

可以考虑使用copydb，clone，cloneCollection。

## 4.2json丢失数据精度

mongoimport和mongoexport并不能可靠的保存所有的富文本BSON数据类型。因为这两类工具所导出的格式基于json的，而json所支持的类型是bson的一个子集，比如bson中的Date和BinData类型，json中就没有，因此导入导出可能会丢失一些精度。但是测试发现并没有丢失，导出的json格式的数据不过是保留了json不能识别类型的类型信息而已：



不过还是建议，在存储的时候尽可能的避免使用json存在的类型，或者都采用字符串也行。

## 4.3导入导出备份恢复有关的锁

mongoexport貌似不会有锁的存在，经过测试，在向collection中写入数据的同时是可以进行mongoexport操作的，通常此时也是可以读mongo中的数据的，但是要考虑数据的一致性。

mongodump和其他几个暂未测试。

一个备份的策略尝试：将数据文件rsync到其他机器的目录中，然后对这个目录应用mongodump工具操作，不知道可行不？经过尝试不行，需要journal目录下一些记录的配合，而且必须是完整的数据文件才行，无法备份处理一个单独的数据文件。

# 参考资料

1. http://www.cnblogs.com/yangxia-test/p/3983271.html
2. http://chenzhou123520.iteye.com/blog/1641319
3. http://www.aspku.com/database/mongodb/44338.html
4. http://justcode.ikeepstudying.com/2016/01/mongodb-%E6%95%99%E7%A8%8B%E5%85%AD-mongodb%E7%AE%A1%E7%90%86%EF%BC%9A%E6%95%B0%E6%8D%AE%E5%AF%BC%E5%85%A5%E5%AF%BC%E5%87%BA%EF%BC%8C%E6%95%B0%E6%8D%AE%E5%A4%87%E4%BB%BD%E6%81%A2%E5%A4%8D%E5%8F%8A/
5. http://baike.baidu.com/link?url=3jPnPXZwRpBU9UlmQvF6Oqwl8FSegamM\_hOe3L\_v14odXkPWQDn4wtkU7hQbhPhbhI9Y1VTUAxhRxFdzvWx-6\_
6. http://www.cnblogs.com/yuechaotian/archive/2012/11/02/2751478.html
7. https://docs.mongodb.com/manual/reference/program/mongoexport/