Release: 20230414

Topics we cover

Use the database logs - getLog

Data profiling to identify slow ops

Configure the data profiling level

Common causes of slow operations

Finding slow operations

The most common source of bad performance is a missing index

Not the only cause - bad client code is a close second cause

Two ways of finding inefficient queries:

1. Database log
2. Profiling a database*

* Profiling requires a feature that is not available on Atlas Shared Instances

Slow operations are usually caused by a missing or incorrect index.
Atlas and Cloud Manager now use the database log for profiling rather than using native profiling.

The getLog Command

getLog gets the last 1024 log entries

Returns a single document - logs are in the log array as strings

In MongoDB 4.4+, these strings are JSON - previously it was just text

```
> var loglines = db.adminCommand({getLog:"global"})
...
> printjson(loglines)

{
      "totalLinesWritten" : 2625918,
      "log" : [ ... ],
      "ok" : 1,
      "$clusterTime" : {
      "clusterTime" : Timestamp(1594293362, 5),
      "signature" : {
      "hash" : BinData(0,"3+McnOj...c3Qs="),
      "keyId" : NumberLong("683...8756")
      }
      },
      "operationTime" : Timestamp(1594293362, 5)
}
```

The MongoDB log file can show slow queries
From MongoDB 4.4, the log is JSON, so it can be written to a searchable collection.

Note: The getLog command does not return log lines on Atlas Shared Tiers currently.

Copying the log to a collection

In MongoDB 4.4+, all log output is now in JSON format

Copy the getLog output to a collection to search or aggregate the data

We may not always see the results in the beginning of our deployment because the slow operations only appear in the log if the duration exceeds the slow operation threshold. In the versions prior to 4.4, you can put it in a collection, but then only query using regular expressions would be possible as these would be just lines of text.

Note: The getLog command does not return log lines on Atlas Shared Tiers currently.

```
use temp
var x = db.adminCommand( { getLog: global } )
db.log.drop()

db.log.insertMany(x.log.map(l=>({l})))
db.log.find({l:/COLLSCAN/}).pretty()
```

The setProfilingLevel command

`db.setProfilingLevel(level, slowms)`

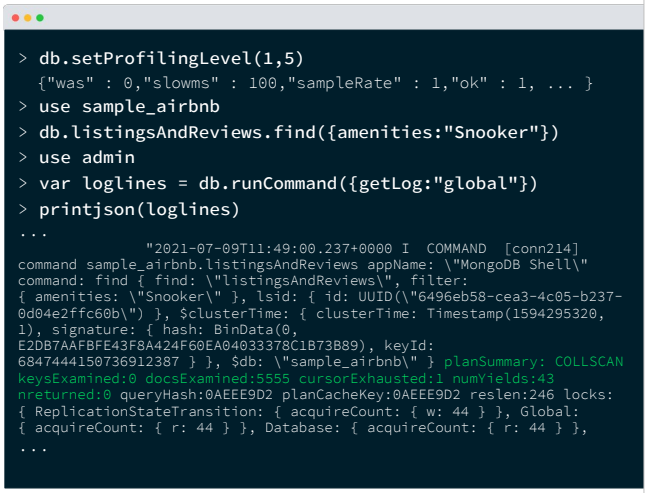
level: 0, 1, or 2

0 - Profiler is off

1 - Store only slow queries

2 - Store all queries

slowms: threshold in milliseconds for slow operations



The db profiler can also be used to find slow queries - slowms sets the threshold for 'slow'
samplerate is what proportion 0-1(1 is 100%) of slow queries to log
Note: The setProfilingLevel() command is not allowed on Atlas Shared Tiers.

Note: The queries are stored in a collection system.profile.

Recording slow operations

Use to capture longer term but adds extra writes - **Remember to turn this off after use as slows production**

Records in a capped collection (10MB max)

Visualization and aggregations are possible

```
> use sample_airbnb

> db.setProfilingLevel(1,5)

> db.listingsAndReviews.find({amenities:"Snooker"})

> db.system.profile.findOne()

{
      "op" : "query",
      "ns" : "sample_airbnb.listingsAndReviews",
      "keysExamined" : 0,
      "docsExamined" : 5555,
      "cursorExhausted" : true,

      ...

}

> db.setProfilingLevel(0,100)
```

Important: The db profiler should be turned off after used as it will degrade performance
Note: The setProfilingLevel() command would not be allowed on Atlas Shared Tiers.

Causes of slow operations

Missing indexes leading to lots of Disk I/O

Writes to cache outstripping disk write capability - Waiting for cache to flush

Locking

- You can see in logs what locks things take
- Most things don't block others - but a few admin things do

There are other causes of slow operations to investigate as well.

Causes of slow operations (cont.)

Excessive CPU usage

- Constantly logging in
- Document lock contention.
- Inappropriately large arrays
- Running code in the database

Nested Joins exploding complexity

There are other causes of slow operations to investigate

Quiz Time!

#1. Select three data items that can be found in Log entries for slow operations

A	The client IP address
B	The number of documents read
C	The number of documents in the database
D	The total server time taken
E	The time spent waiting for any locks

Answer in the next slide.

#1. Select the items found in Log entries for slow operations.

A	The client IP address	B	The number of documents read	C	The number of documents in the database
D	The total server time taken	E	The time spent waiting for any locks		

Answer: B, D, E