MongoDB Operations Rapid Start Training

MongoDB Operations Rapid Start Training

Release 3.4

MongoDB, Inc.

Jun 05, 2017

Contents

1 Introduction

1.1 Warm Up

Introductions

- Who am I?
- My role at MongoDB
- My background and prior experience

Note:

- Tell the students about yourself:
  - Your role
  - Prior experience

Getting to Know You

- Who are you?
- What role do you play in your organization?
- What is your background?
- Do you have prior experience with MongoDB?

Note:

- Ask students to go around the room and introduce themselves.
- Make sure the names match the roster of attendees.
- Ask about what roles the students play in their organization and note on attendance sheet.
- Ask what software stacks students are using.
  - With MongoDB and in general.
  - Note this information as well.

MongoDB Experience

- Who has never used MongoDB?
- Who has some experience?
- Who has worked with production MongoDB deployments?
- Who is more of a developer?
- Who is more of an operations person?

1.2 MongoDB - The Company

10gen

- MongoDB was initially created in 2008 as part of a hosted application stack.
- The company was originally called 10gen.
- As part of their overarching plan to create the 10gen platform, the company built a database.
- Suddenly everybody said: "I like that! Give me that database!"

Origin of MongoDB

- 10gen became a database company.
- In 2013, the company rebranded as MongoDB, Inc.
- The founders have other startups to their credit: DoubleClick, ShopWiki, Gilt.
- The motivation for the database came from observing the following pattern with application development.
  - The user base grows.
  - The associated body of data grows.
  - Eventually the application outgrows the database.
  - Meeting performance requirements becomes difficult.

1.3 MongoDB Overview

Learning Objectives

Upon completing this module students should understand:

- MongoDB vs. relational databases and key/value stores
- Vertical vs. horizontal scaling
- The role of MongoDB in the development stack
- The structure of documents in MongoDB
- Array fields
- Embedded documents
- Fundamentals of BSON

MongoDB is a Document Database

Documents are associative arrays like:

- Python dictionaries
- Ruby hashes
- PHP arrays
- JSON objects

An Example MongoDB Document

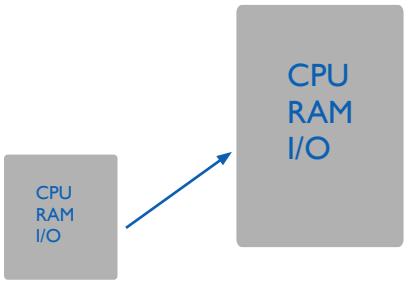
A MongoDB document expressed using JSON syntax.

```
{
  "_id" : "/apple-reports-second-quarter-revenue",
  "headline" : "Apple Reported Second Quarter Revenue Today",
  "date" : ISODate("2015-03-24T22:35:21.908Z"),
  "author" : {
    "name" : "Bob Walker",
    "title" : "Lead Business Editor"
  },
  "copy" : "Apple beat Wall St expectations by reporting ...",
  "tags" : [
    "AAPL", "Earnings", "Cupertino"
  ],
  "comments" : [
    { "name" : "Frank", "comment" : "Great Story" },
    { "name" : "Wendy", "comment" : "When can I buy an Apple Watch?" }
  ]
}
```

Note:

- How would you represent this document in a relational database? How many tables, how many queries per page load?
- What are the pros/cons to this design? (hint: 1 million comments)
- Where relational databases store rows, MongoDB stores documents.
- Documents are hierarchical data structures.
- This is a fundamental departure from relational databases where rows are flat.

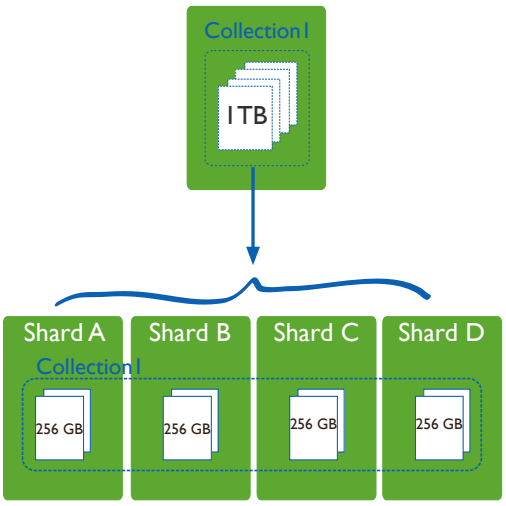
Vertical Scaling



Note: Another difference is in terms of scalability. With an RDBMS:

- If you need to support a larger workload, you buy a bigger machine.
- The problem is, machines are not priced linearly.
- The largest machines cost much more than commodity hardware.
- If your application is successful, you may find you simply cannot buy a large enough a machine to support your workload.

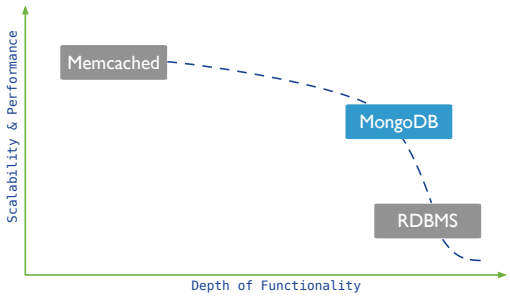
Scaling with MongoDB



Note:

- MongoDB is designed to be horizontally scalable (linear).
- MongoDB scales by enabling you to shard your data.
- When you need more performance, you just buy another machine and add it to your cluster.
- MongoDB is highly performant on commodity hardware.

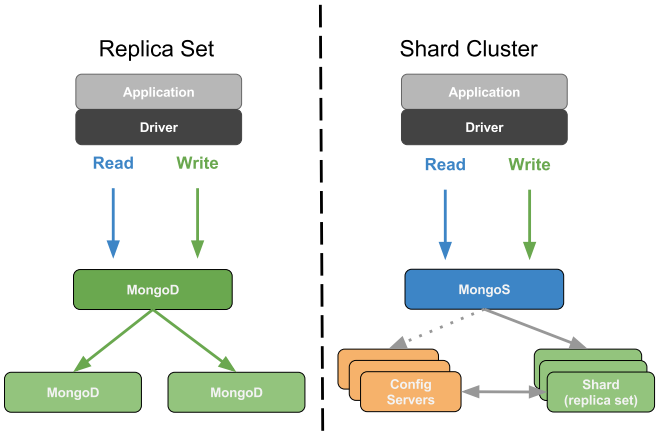
Database Landscape



Note:

- We've plotted each technology by scalability and functionality.
- At the top left, are key/value stores like memcached.
- These scale well, but lack features that make developers productive.
- At the far right we have traditional RDBMS technologies.
- These are full featured, but will not scale easily.
- Joins and transactions are difficult to run in parallel.
- MongoDB has nearly as much scalability as key-value stores.
- Gives up only the features that prevent scaling.
- We have compensating features that mitigate the impact of that design decision.

MongoDB Deployment Models



Note:

- MongoDB supports high availability through automated failover.
- Do not use a single-server deployment in production.
- Typical deployments use replica sets of 3 or more nodes.
  - The primary node will accept all writes, and possibly all reads.
  - Each secondary will replicate from another node.
  - If the primary fails, a secondary will automatically step up.
  - Replica sets provide redundancy and high availability.
- In production, you typically build a fully sharded cluster:
  - Your data is distributed across several shards.
  - The shards are themselves replica sets.
  - This provides high availability and redundancy at scale.

1.4 MongoDB Stores Documents

Learning Objectives

Upon completing this module, students should understand:

- JSON
- BSON basics
- That documents are organized into collections

JSON

- JavaScript Object Notation
- Objects are associative arrays.
- They are composed of key-value pairs.

A Simple JSON Object

```
{
    "firstname" : "Thomas",
    "lastname" : "Smith",
    "age" : 29
}
```

JSON Keys and Values

- Keys must be strings.
- Values may be any of the following:
  - string (e.g., “Thomas”)
  - number (e.g., 29, 3.7)
  - true / false
  - null
  - array (e.g., [88.5, 91.3, 67.1])
  - object
- More detail at json.org[1].

Example Field Values

```
{
  "headline" : "Apple Reported Second Quarter Revenue Today",
  "date" : ISODate("2015-03-24T22:35:21.908Z"),
  "views" : 1234,
  "author" : {
    "name" : "Bob Walker",
    "title" : "Lead Business Editor"
  },
  "tags" : [
    "AAPL",
    23,
    { "name" : "city", "value" : "Cupertino" },
    { "name" : "stockPrice", "value": NumberDecimal("143.51")},
    [ "Electronics", "Computers" ]
  ]
}
```

[1] http://json.org/

BSON

- MongoDB stores data as Binary JSON (BSON).
- MongoDB drivers send and receive data in this format.
- They map BSON to native data structures.
- BSON provides support for all JSON data types and several others.
- BSON was designed to be lightweight, traversable and efficient.
- See bsonspec.org[2].

Note: E.g., a BSON object will be mapped to a dictionary in Python.

BSON Hello World

```
// JSON
{ "hello" : "world" }

// BSON
"\x16\x00\x00\x00\x02hello\x00
 \x06\x00\x00\x00world\x00\x00"
```

Note:

- \x16\x00\x00\x00 (document size)
- \x02 = string (data type of field value)
- hello\x00 (key/field name, \x00 is null and delimits the end of the name)
- \x06\x00\x00\x00 (size of field value including end null)
- world\x00 (field value)
- \x00 (end of the document)

A More Complex BSON Example

```
// JSON
{ "BSON" : [ "awesome", 5.05, 1986 ] }

// BSON
"\x3b\x00\x00\x00\x04BSON\x00\x26\x00
 \x00\x00\x020\x00\x08\x00\x00
 \x00awesome\x00\x011\x00\x33\x33\x33\x33\x33\x33
 \x14\x40\x102\x00\xc2\x07\x00\x00
 \x00\x00"
```

[2] http://bsonspec.org/#/specification

Documents, Collections, and Databases

- Documents are stored in collections.
- Collections are contained in a database.
- Example:
  - Database: products
  - Collections: books, movies, music
- Each database-collection combination defines a namespace.
  - products.books
  - products.movies
  - products.music

The `_id` Field

- All documents must have an `_id` field.
- The `_id` is immutable.
- If no `_id` is specified when a document is inserted, MongoDB will add the `_id` field.
- MongoDB assigns a unique ObjectId as the value of `_id`.
- Most drivers will actually create the ObjectId if no `_id` is specified.
- The `_id` field is unique to a collection (namespace).

1.5 MongoDB Data Types

Learning Objectives

By the end of this module, students should understand:

- What data types MongoDB supports
- Special consideration for some BSON types

What is BSON?

BSON is a binary serialization of JSON, used to store documents and make remote procedure calls in MongoDB. For more in-depth coverage of BSON, specifically refer to bsonspec.org[3]

Note: All official MongoDB drivers map BSON to native types and data structures

---

BSON types

MongoDB supports a wide range of BSON types. Each data type has a corresponding number and string alias that can be used with the `$type` operator to query documents by BSON type.

Double 1 “double”

String 2 “string”

Object 3 “object”

Array 4 “array”

Binary data 5 “binData”

ObjectId 7 “objectId”

Boolean 8 “bool”

Date 9 “date”

Null 10 “null”

BSON types continued

Regular Expression 11 “regex”

JavaScript 13 “javascript”

JavaScript (w/ scope) 15 “javascriptWithScope”

32-bit integer 16 “int”

Timestamp 17 “timestamp”

64-bit integer 18 “long”

Decimal128 19 “decimal”

Min key -1 “minKey”

Max key 127 “maxKey”

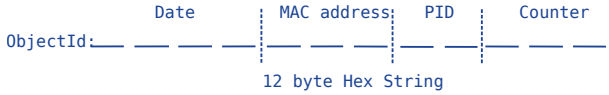
Note: Take the opportunity to show students how to query using `$type` operator:

```
use datatypes
db.sample.insertMany([
  {a: 1, b: "hello"},
  {a: 1.2, b: { c: "goodbye"}}
])
```

[3] http://bsonspec.org/

```
db.sample.find({b: {$type: 'string'}})
db.sample.find({b: {$type: 3}})
```

ObjectId




```
> ObjectId()
ObjectId("58dc309ce3f39998099d6275")
```

Note:

- An ObjectId is a 12-byte value.
- The first 4 bytes are a datetime reflecting when the ObjectID was created.
- The next 3 bytes are the MAC address of the server.
- Then a 2-byte process ID
- Finally, 3 bytes that are monotonically increasing for each new ObjectId created within a collection.

Timestamps

BSON has a special timestamp type for *internal* MongoDB use and is **not** associated with the regular Date type.

Date

BSON Date is a 64-bit integer that represents the number of milliseconds since the Unix epoch (Jan 1, 1970). This results in a representable date range of about 290 million years into the past and future.

- Official BSON spec refers to the BSON Date type as UTC datetime
- Signed data type. Negative values represent dates before 1970.

```
var today = ISODate() // using the ISODate constructor
```

Decimal

In MongoDB 3.4, support was added for 128-bit decimals.

- The **decimal** BSON type uses the decimal128 decimal-based floating-point numbering format.
- This supports 34 significant digits and an exponent range of **-6143** to **+6144**.
- Intended for applications that handle monetary and scientific data that requires exact precision.

How to use Decimal

For specific information about how your preferred driver supports decimal128, click here[4].

In the Mongo shell, we use the *NumberDecimal()* constructor.

- Can be created with a string argument or a double
- Stored in the database as *NumberDecimal("999.4999")*

```
> NumberDecimal("999.4999")
NumberDecimal("999.4999")
> NumberDecimal(999.4999)
NumberDecimal("999.4999")
```

Note: Using a double as the argument can lead to loss of precision. A string argument is preferred.

```
MongoDB Enterprise > NumberDecimal("999.4999999999999")
NumberDecimal("999.4999999999999")
MongoDB Enterprise > NumberDecimal(999.4999999999999)
NumberDecimal("999.500000000000")
```

Decimal Considerations

- If upgrading an existing database to use **decimal128**, it is recommended a new field be added to reflect the new type. The old field may be deleted after verifying consistency
- If any fields contain **decimal128** data, they will not be compatible with previous versions of MongoDB. There is no support for downgrading datafiles containing decimals
- **decimal** types are not strictly equal to their **double** representations, so use the **NumberDecimal** constructor in queries.

Note: Show the following example:

```
MongoDB Enterprise > use foo
MongoDB Enterprise > db.numbers.insertMany([
  { "_id" : 1, "val" : NumberDecimal( "9.99" ), "description" : "Decimal" },
  { "_id" : 2, "val" : 9.99, "description" : "Double" },
  { "_id" : 3, "val" : 10, "description" : "Double" },
  { "_id" : 4, "val" : NumberLong(10), "description" : "Long" },
  { "_id" : 5, "val" : NumberDecimal( "10.0" ), "description" : "Decimal" }
])
```

[4] https://docs.mongodb.com/ecosystem/drivers/

```
MongoDB Enterprise > db.numbers.find().sort({"val": 1})
{ "_id" : 1, "val" : NumberDecimal("9.99"), "description" : "Decimal" }
{ "_id" : 2, "val" : 9.99, "description" : "Double" }
{ "_id" : 3, "val" : 10, "description" : "Double" }
{ "_id" : 4, "val" : NumberLong(10), "description" : "Long" }
{ "_id" : 5, "val" : NumberDecimal("10.0"), "description" : "Decimal" }

MongoDB Enterprise > db.numbers.find({"val": NumberDecimal("10")}).sort({"val": 1})
{ "_id" : 3, "val" : 10, "description" : "Double" }
{ "_id" : 4, "val" : NumberLong(10), "description" : "Long" }
{ "_id" : 5, "val" : NumberDecimal("10.0"), "description" : "Decimal" }

MongoDB Enterprise > db.numbers.find({"val": NumberDecimal("9.99")}).sort({"val": 1})
{ "_id" : 1, "val" : NumberDecimal("9.99"), "description" : "Decimal" }

MongoDB Enterprise > db.numbers.find({"val": 9.99}).sort({"val": 1})
{ "_id" : 2, "val" : 9.99, "description" : "Double" }
```

1.6 Lab: Installing and Configuring MongoDB

Learning Objectives

Upon completing this exercise students should understand:

- How MongoDB is distributed
- How to install MongoDB
- Configuration steps for setting up a simple MongoDB deployment
- How to run MongoDB
- How to run the Mongo shell

Production Releases

64-bit production releases of MongoDB are available for the following platforms.

- Windows
- OSX
- Linux
- Solaris

Installing MongoDB

- Visit https://docs.mongodb.com/manual/installation/.
- Please install the Enterprise version of MongoDB.
- Click on the appropriate link, such as "Install on Windows" or "Install on OS X" and follow the instructions.
- Versions:
  - Even-numbered builds are production releases, e.g., 2.4.x, 2.6.x.
  - Odd-numbers indicate development releases, e.g., 2.5.x, 2.7.x.

Linux Setup

```
PATH=$PATH:<path to mongodb>/bin

sudo mkdir -p /data/db

sudo chmod -R 744 /data/db

sudo chown -R `whoami` /data/db
```

Note:

- You might want to add the MongoDB bin directory to your path, e.g.
- Once installed, create the MongoDB data directory.
- Make sure you have write permission on this directory.

If you are using Koding these are a few instructions you can follow:

- Download MongoDB tarbal and setup the environment

```
wget http://downloads.mongodb.org/linux/mongodb-linux-x86_64-ubuntu1204-3.2.
↪1.tgz
tar xzvf mongodb-linux-x86_64-ubuntu1204-3.2.1.tgz
cd mongodb-linux-x86_64-ubuntu1204-3.2.1/bin
export PATH=`pwd`:$PATH
```

Install on Windows

- Download and run the .msi Windows installer from mongodb.org/downloads.
- By default, binaries will be placed in the following directory.

```
C:\Program Files\MongoDB\Server\<VERSION>\bin
```

- It is helpful to add the location of the MongoDB binaries to your path.
- To do this, from "System Properties" select "Advanced" then "Environment Variables"

Note: Can also install Windows as a service, but not recommended since we need multiple mongod processes for future exercises

---

Create a Data Directory on Windows

- Ensure there is a directory for your MongoDB data files.
- The default location is `\data\db`.
- Create a data directory with a command such as the following.

```
md \data\db
```

Note: Optionally, talk about the –dbpath variable and specifying a different location for the data files

---

Launch a `mongod`

Explore the `mongod` command.

```
<path to mongodb>/bin/mongod --help
```

Launch a `mongod` with the MMAPv1 storage engine:

```
<path to mongodb>/bin/mongod --storageEngine mmapv1
```

Alternatively, launch with the WiredTiger storage engine (default).

```
<path to mongodb>/bin/mongod
```

Specify an alternate path for data files using the `--dbpath` option. (Make sure the directory already exists.) E.g.,

```
<path to mongodb>/bin/mongod --dbpath /test/mongodb/data/wt
```

Note:

- Please verify that all students have successfully installed MongoDB.
- Please verify that all can successfully launch a mongod.

The MMAPv1 Data Directory

```
ls /data/db
```

- The mongod.lock file
  - This prevents multiple mongods from using the same data directory simultaneously.
  - Each MongoDB database directory has one .lock.
  - The lock file contains the process id of the mongod that is using the directory.
- Data files
  - The names of the files correspond to available databases.
  - A single database may have multiple files.

Note: Files for a single database increase in size as follows:

- sample.0 is 64 MB
- sample.1 is 128 MB
- sample.2 is 256 MB, etc.
- This continues until sample.5, which is 2 GB
- All subsequent data files are also 2 GB.

The WiredTiger Data Directory

```
ls /data/db
```

- The mongod.lock file
  - Used in the same way as MMAPv1.
- Data files
  - Each collection and index stored in its own file.
  - Will fail to start if MMAPv1 files found

Import Exercise Data

```
unzip usb_drive.zip

cd usb_drive

mongoimport -d sample -c tweets twitter.json

mongoimport -d sample -c zips zips.json

mongoimport -d sample -c grades grades.json

cd dump

mongorestore -d sample city

mongorestore -d sample digg
```

Note: If there is an error importing data directly from a USB drive, please copy the sampledata.zip file to your local computer first.

Note: For local installs

- Import the data provided on the USB drive into the *sample* database.

For Koding environment

- Download *sample* data from:

```
wget https://www.dropbox.com/s/54xsjwq59zoq1fe/sample.tgz
```

Launch a Mongo Shell

Open another command shell. Then type the following to start the Mongo shell.

```
mongo
```

Display available commands.

```
help
```

Note: On Koding environment do the following:

- Create a new *Terminal* and rename it to **Client**

Explore Databases

Display available databases.

```
show dbs
```

To use a particular database we can type the following.

```
use <database_name>

db
```

Note:

- This assigns the variable `db` to a connection object for the selected database.
- We can display the name of the database we are currently using by evaluating `db` in the mongo shell.
- Highlight the power of the Mongo shell here.
- It is a fully programmable JavaScript environment.
  - To demonstrate this you can use the following code block

```
for(i=0;i<10;i++){ print("this is line "+i)}
```

Exploring Collections

```
show collections

db.<COLLECTION>.help()

db.<COLLECTION>.find()
```

Note:

- Show the collections available in this database.
- Show methods on the collection with parameters and a brief explanation.
- Finally, we can query for the documents in a collection.

Admin Commands

- There are also a number of admin commands at our disposal.
- The following will shut down the mongod we are connected to through the Mongo shell.
- You can also just kill with Ctrl-C in the shell window from which you launched the mongod.

```
db.adminCommand( { shutdown : 1 } )
```

- Confirm that the mongod process has indeed stopped.
- Once you have, please restart it.

2 CRUD

2.1 Creating and Deleting Documents

Learning Objectives

Upon completing this module students should understand:

- How to insert documents into MongoDB collections.
- `_id` fields:
- How to delete documents from a collection
- How to remove a collection from a database
- How to remove a database from a MongoDB deployment

Creating New Documents

- Create documents using `insertOne()` and `insertMany()`.
- For example:

```
// Specify the collection name
db.<COLLECTION>.insertOne( { "name" : "Mongo" } )

// For example
db.people.insertOne( { "name" : "Mongo" } )
```

Example: Inserting a Document

Experiment with the following commands.

```
use sample

db.movies.insertOne( { "title" : "Jaws" } )

db.movies.find()
```

Note:

- Make sure the students are performing the operations along with you.
- Some students will have trouble starting things up, so be helpful at this stage.

Implicit `_id` Assignment

- We did not specify an `_id` in the document we inserted.
- If you do not assign one, MongoDB will create one automatically.
- The value will be of type ObjectId.

Example: Assigning _ids

Experiment with the following commands.

```
db.movies.insertOne( { "_id" : "Jaws", "year" : 1975 } )
db.movies.find()
```

Note:

- Note that you can assign an _id to be of almost any type.
- It does not need to be an ObjectId.

Inserts will fail if...

- There is already a document in the collection with that `_id`.
- You try to assign an array to the `_id`.
- The argument is not a well-formed document.

Example: Inserts will fail if...

```
// fails because _id can't have an array value
db.movies.insertOne( { "_id" : [ "Star Wars",
                                 "The Empire Strikes Back",
                                 "Return of the Jedi" ] } )

// succeeds
db.movies.insertOne( { "_id" : "Star Wars" } )

// fails because of duplicate id
db.movies.insertOne( { "_id" : "Star Wars" } )

// malformed document
db.movies.insertOne( { "Star Wars" } )
```

Note:

- The following will fail because it attempts to use an array as an `_id`.

```
db.movies.insertOne( { "_id" : [ "Star Wars",
                                 "The Empire Strikes Back",
                                 "Return of the Jedi" ] } )
```

- The second insert with `_id : "Star Wars"` will fail because there is already a document with `_id` of "Star Wars" in the collection.
- The following will fail because it is a malformed document (i.e. no field name, just a value).

```
db.movies.insertOne( { "Star Wars" } )
```

`insertMany()`

- You may bulk insert using an array of documents.
- Use `insertMany()` instead of `insertOne()`

Note:

- In the case of an ordered bulk operation, every operation will be executed in the order they are added to the bulk operation.
- In the case of an unordered bulk operation however there is no guarantee what order the operations are executed.
- With an unordered bulk operation, the operations in the list may be reordered to increase performance.

Ordered `insertMany()`

- For ordered inserts MongoDB will stop processing inserts upon encountering an error.
- Meaning that only inserts occurring before an error will complete.
- The default setting for `db.<COLLECTION>.insertMany` is an ordered insert.
- See the next exercise for an example.

Example: Ordered `insertMany()`

Experiment with the following operation.

```
db.movies.insertMany( [ { "_id" : "Batman", "year" : 1989 },
                        { "_id" : "Home Alone", "year" : 1990 },
                        { "_id" : "Ghostbusters", "year" : 1984 },
                        { "_id" : "Ghostbusters", "year" : 1984 } ] )
db.movies.find()
```

Note:

- This example has a duplicate key error.
- Only the first 3 documents will be inserted.

Unordered `insertMany()`

- Pass `{ ordered : false }` to `insertMany()` to perform unordered inserts.
- If any given insert fails, MongoDB will still attempt all of the others.
- The inserts may be executed in a different order than you specified.
- The next exercise is very similar to the previous one.
- However, we are using `{ ordered : false }`.
- One insert will fail, but all the rest will succeed.

Example: Unordered `insertMany()`

Experiment with the following insert.

```
db.movies.insertMany( [ { "_id" : "Jaws", "year" : 1975 },
                        { "_id" : "Titanic", "year" : 1997 },
                        { "_id" : "The Lion King", "year" : 1994 } ],
                      { ordered : false } )
db.movies.find()
```

The Shell is a JavaScript Interpreter

- Sometimes it is convenient to create test data using a little JavaScript.
- The mongo shell is a fully-functional JavaScript interpreter. You may:
  - Define functions
  - Use loops
  - Assign variables
  - Perform inserts

Exercise: Creating Data in the Shell

Experiment with the following commands.

```
for (i=1; i<=10000; i++) {
    db.stuff.insert( { "a" : i } )
}

db.stuff.find()
```

Deleting Documents

You may delete documents from a MongoDB deployment in several ways.

- Use `deleteOne()` and `deleteMany()` to delete documents matching a specific set of conditions.
- Drop an entire collection.
- Drop a database.

Using `deleteOne()`

- Delete a document from a collection using `deleteOne()`
- This command has one required parameter, a query document.
- The first document in the collection matching the query document will be deleted.

Using `deleteMany()`

- Delete multiple documents from a collection using `deleteMany()`.
- This command has one required parameter, a query document.
- All documents in the collection matching the query document will be deleted.
- Pass an empty document to delete all documents.

Example: Deleting Documents

Experiment with removing documents. Do a `find()` after each `deleteMany()` command below.

```
for (i=1; i<=20; i++) { db.testcol.insertOne( { _id : i, a : i } ) }

db.testcol.deleteMany( { a : 1 } )  // Delete the first document

// $lt is a query operator that enables us to select documents that
// are less than some value. More on operators soon.
db.testcol.deleteMany( { a : { $lt : 5 } } )  // Remove three more

db.testcol.deleteOne( { a : { $lt : 10 } } )  // Remove one more

db.testcol.deleteMany()  // Error: requires a query document.

db.testcol.deleteMany( { } )  // All documents removed
```

Dropping a Collection

- You can drop an entire collection with `db.<COLLECTION>.drop()`
- The collection and all documents will be deleted.
- It will also remove any metadata associated with that collection.
- Indexes are one type of metadata removed.
- **All collection and indexes files are removed and space allocated reclaimed.**
  - Wired Tiger only!
- More on meta data later.

Note: Mention that `drop()` is more performant than `deleteMany()`.

Example: Dropping a Collection

```
db.colToBeDropped.insertOne( { a : 1 } )
show collections  // Shows the colToBeDropped collection

db.colToBeDropped.drop()
show collections  // collection is gone
```

Dropping a Database

- You can drop an entire database with `db.dropDatabase()`
- This drops the database on which the method is called.
- It also deletes the associated data files from disk, freeing disk space.
- Beware that in the mongo shell, this does not change database context.

Example: Dropping a Database

```
use tempDB
db.testcol1.insertOne( { a : 1 } )
db.testcol2.insertOne( { a : 1 } )

show dbs  // Here they are
show collections  // Shows the two collections

db.dropDatabase()
show collections  // No collections
show dbs  // The db is gone

use sample  // take us back to the sample db
```

2.2 Reading Documents

Learning Objectives

Upon completing this module students should understand:

- The query-by-example paradigm of MongoDB
- How to query on array elements
- How to query embedded documents using dot notation
- How the mongo shell and drivers use cursors
- Projections
- Cursor methods: `.count()`, `.sort()`, `.skip()`, `.limit()`

The `find()` Method

- This is the fundamental method by which we read data from MongoDB.
- We have already used it in its basic form.
- `find()` returns a cursor that enables us to iterate through all documents matching a query.
- We will discuss cursors later.

Query by Example

- To query MongoDB, specify a document containing the key / value pairs you want to match
- You need only specify values for fields you care about.
- Other fields will not be used to exclude documents.
- The result set will include all documents in a collection that match.

Example: Querying by Example

Experiment with the following sequence of commands.

```
db.movies.drop()
db.movies.insertMany( [
    { "title" : "Jaws", "year" : 1975, "imdb_rating" : 8.1 },
    { "title" : "Batman", "year" : 1989, "imdb_rating" : 7.6 }
] )
db.movies.find()

db.movies.find( { "year" : 1975 } )

// Multiple Batman movies from different years, find the correct one
db.movies.find( { "year" : 1989, "title" : "Batman" } )
```

Note: Matching Rules:

- Any field specified in the query must be in each document returned.
- Values for returned documents must match the conditions specified in the query document.

- If multiple fields are specified, all must be present in each document returned.
- Think of it as a logical AND for all fields.

Querying Arrays

- In MongoDB you may query array fields.
- Specify a single value you expect to find in that array in desired documents.
- Alternatively, you may specify an entire array in the query document.
- As we will see later, there are also several operators that enhance our ability to query array fields.

Note: Students might find it helpful to think of an array field as having multiple values – one for each of its elements.

Example: Querying Arrays

```
db.movies.drop()
db.movies.insertMany(
 [{ "title" : "Batman", "category" : [ "action", "adventure" ] },
  { "title" : "Godzilla", "category" : [ "action", "adventure", "sci-fi" ] },
  { "title" : "Home Alone", "category" : [ "family", "comedy" ] }
 ])

// Match documents where "category" contains the value specified
db.movies.find( { "category" : "action" } )

// Match documents where "category" equals the value specified
db.movies.find( { "category" : [ "action", "sci-fi" ] } )  // no documents

// only the second document
db.movies.find( { "category" : [ "action", "adventure", "sci-fi" ] } )
```

Note: Later, we'll see operators that will allow us to do things like match all documents where an array field contains any of a set of values.

Querying with Dot Notation

- Dot notation is used to query on fields in embedded documents.
- The syntax is:

```
"field1.field2" : value
```

- Put quotes around the field name when using dot notation.

Example: Querying with Dot Notation

```
db.movies.insertMany(
    [ {
          "title" : "Avatar",
          "box_office" : { "gross" : 760,
                           "budget" : 237,
                           "opening_weekend" : 77
                         }
      },
      {
          "title" : "E.T.",
          "box_office" : { "gross" : 349,
                           "budget" : 10.5,
                           "opening_weekend" : 14
                         }
       }
     ] )

db.movies.find( { "box_office" : { "gross" : 760 } } ) // no values

// dot notation
db.movies.find( { "box_office.gross" : 760 } )  // expected value
```

Example: Arrays and Dot Notation

```
db.movies.insertMany( [
    { "title" : "E.T.",
      "filming_locations" :
         [ { "city" : "Culver City", "state" : "CA", "country" : "USA" },
           { "city" : "Los Angeles", "state" : "CA", "country" : "USA" },
           { "city" : "Cresecent City", "state" : "CA", "country" : "USA" }
         ] },
    { "title": "Star Wars",
      "filming_locations" :
         [ { "city" : "Ajim", "state" : "Jerba", "country" : "Tunisia" },
           { "city" : "Yuma", "state" : "AZ", "country" : "USA" }
         ] } ] )

 db.movies.find( { "filming_locations.country" : "USA" } ) // two documents
```

Note:

- This query finds documents where:
  - There is a `filming_locations` field.
  - The `filming_locations` field contains one or more embedded documents.
  - At least one embedded document has a field `country`.
  - The field `country` has the specified value (“USA”).
- In this collection, `filming_locations` is actually an array field.
- The embedded documents we are matching are held within these arrays.

Projections

- You may choose to have only certain fields appear in result documents.
- This is called projection.
- You specify a projection by passing a second parameter to `find()`.

Projection: Example (Setup)

```
db.movies.insertOne(
{
   "title" : "Forrest Gump",
   "category" : [ "drama", "romance" ],
   "imdb_rating" : 8.8,
   "filming_locations" : [
     { "city" : "Savannah", "state" : "GA", "country" : "USA" },
     { "city" : "Monument Valley", "state" : "UT", "country" : "USA" },
     { "city" : "Los Anegeles", "state" : "CA", "country" : "USA" }
   ],
   "box_office" : {
     "gross" : 557,
     "opening_weekend" : 24,
     "budget" : 55
   }
 })
```

Projection: Example

```
db.movies.findOne( { "title" : "Forrest Gump" },
                   { "title" : 1, "imdb_rating" : 1 } )
{
   "_id" : ObjectId("5515942d31117f52a5122353"),
   "title" : "Forrest Gump",
   "imdb_rating" : 8.8
}
```

Projection Documents

- Include fields with `fieldName: 1`.
  - Any field not named will be excluded
  - except _id, which must be explicitly excluded.
- Exclude fields with `fieldName: 0`.
  - Any field not named will be included.

Example: Projections

```
for (i=1; i<=20; i++) {
    db.movies.insertOne(
        { "_id" : i, "title" : i,
          "imdb_rating" : i, "box_office" : i } )
}
db.movies.find()
// no "box_office"
db.movies.find( { "_id" : 3 }, { "title" : 1, "imdb_rating" : 1 } )
// no "imdb_rating"
db.movies.find( { "_id" : { $gte : 10 } }, { "imdb_rating" : 0 } )
// just "title"
db.movies.find( { "_id" : 4 }, { "_id" : 0, "title" : 1 } )
// just "imdb_rating", "box_office"
db.movies.find( { "_id" : 5 }, { _id : 0, "title" : 0 } )
// Can't mix inclusion/exclusion except _id
db.movies.find( { "_id" : 6 }, { "title" : 1, "imdb_rating" : 0 } )
```

The last `find()` fails because MongoDB cannot determine how to handle unnamed fields such as `box_office`.

Cursors

- When you use `find()`, MongoDB returns a cursor.
- A cursor is a pointer to the result set
- You can get iterate through documents in the result using `next()`.
- By default, the mongo shell will iterate through 20 documents at a time.

Example: Introducing Cursors

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    db.testcol.insertOne( { a : Math.floor( Math.random() * 100 + 1 ),
                            b : Math.floor( Math.random() * 100 + 1 ) } )
}
db.testcol.find()

it
it
```

Note:

- With the `find()` above, the shell iterates over the first 20 documents.
- `it` causes the shell to iterate over the next 20 documents.
- Can continue issuing `it` commands until all documents are seen.

Example: Cursor Objects in the Mongo Shell

```
// Assigns the cursor returned by find() to a variable x
var x = db.testcol.find()

// Displays the first document in the result set.
x.next()

// True because there are more documents in the result set.
x.hasNext()

// Assigns the next document in the result set to the variable y.
y = x.next()

// Return value is the value of the a field of this document.
y.a

// Displaying a cursor prints the next 20 documents in the result set.
x
```

Cursor Methods

- `count()`: Returns the number of documents in the result set.
- `limit()`: Limits the result set to the number of documents specified.
- `skip()`: Skips the number of documents specified.

Example: Using `count()`

```
db.testcol.drop()
for (i=1; i<=100; i++) { db.testcol.insertOne( { a : i } ) }

// all 100
db.testcol.count()

// just 41 docs
db.testcol.count( { a : { $lt : 42 } } )

// Another way of writing the same query
db.testcol.find( { a : { $lt : 42 } } ).count(  )
```

Note:

- You may pass a query document like you would to `find()`.
- `count()` will count only the documents matching the query.
- Will return the number of documents in the collection if you do not specify a query document.
- The last query in the above achieves the same result because it operates on the cursor returned by `find()`.

Example: Using `sort()`

```
db.testcol.drop()
for (i=1; i<=20; i++) {
    db.testcol.insertOne( { a : Math.floor( Math.random() * 10 + 1 ),
                            b : Math.floor( Math.random() * 10 + 1 ) } )
}

db.testcol.find()

// sort descending; use 1 for ascending
db.testcol.find().sort( { a : -1 } )

// sort by b, then a
db.testcol.find().sort( { b : 1, a : 1 } )

// $natural order is just the order on disk.
db.testcol.find().sort( { $natural : 1 } )
```

Note:

- Sort can be executed on a cursor until the point where the first document is actually read.
- If you never delete any documents or change their size, this will be the same order in which you inserted them.
- Sorting two or more fields breaks the convention of javascript objects that key / value pairs are unordered.
- In some drivers you may need to take special care with this.
- For example, in Python, you would usually query with a dictionary.
- But dictionaries are unordered in Python, so you would use an array of tuples instead.

The `skip()` Method

- Skips the specified number of documents in the result set.
- The returned cursor will begin at the first document beyond the number specified.
- Regardless of the order in which you specify `skip()` and `sort()` on a cursor, `sort()` happens first.

The `limit()` Method

- Limits the number of documents in a result set to the first `k`.
- Specify `k` as the argument to `limit()`
- Regardless of the order in which you specify `limit()`, `skip()`, and `sort()` on a cursor, `sort()` happens first.
- Helps reduce resources consumed by queries.

The `distinct()` Method

- Returns all values for a field found in a collection.
- Only works on one field at a time.
- Input is a string (not a document)

Example: Using `distinct()`

```
db.movie_reviews.drop()
db.movie_reviews.insertMany( [
    { "title" : "Jaws", "rating" : 5 },
    { "title" : "Home Alone", "rating" : 1 },
    { "title" : "Jaws", "rating" : 7 },
    { "title" : "Jaws", "rating" : 4 },
    { "title" : "Jaws", "rating" : 8 } ] )
db.movie_reviews.distinct( "title" )
```

Note: Returns

```
{
    "values" : [ "Jaws", "Home Alone" ],
    "stats" : { ... },
    "ok" : 1
}
```

2.3 Query Operators

Learning Objectives

Upon completing this module students should understand the following types of MongoDB query operators:

- Comparison operators
- Logical operators
- Element query operators
- Operators on arrays

Comparison Query Operators

- `$lt`: Exists and is less than
- `$lte`: Exists and is less than or equal to
- `$gt`: Exists and is greater than
- `$gte`: Exists and is greater than or equal to
- `$ne`: Does not exist or does but is not equal to
- `$in`: Exists and is in a set
- `$nin`: Does not exist or is not in a set

Example (Setup)

```
// insert sample data
db.movies.insertMany( [
  {
    "title" : "Batman",
    "category" : [ "action", "adventure" ],
    "imdb_rating" : 7.6,
    "budget" : 35
  },
  {
    "title" : "Godzilla",
    "category" : [ "action",
    "adventure", "sci-fi" ],
    "imdb_rating" : 6.6
  },
  {
    "title" : "Home Alone",
    "category" : [ "family", "comedy" ],
    "imdb_rating" : 7.4
  }
] )
```

Example: Comparison Operators

```
db.movies.find()

db.movies.find( { "imdb_rating" : { $gte : 7 } } )

db.movies.find( { "category" : { $ne : "family" } } )

db.movies.find( { "title" : { $in : [ "Batman", "Godzilla" ] } } )

db.movies.find( { "title" : { $nin : [ "Batman", "Godzilla" ] } } )
```

Logical Query Operators

- `$or`: Match either of two or more values
- `$not`: Used with other operators
- `$nor`: Match neither of two or more values
- `$and`: Match both of two or more values
  - This is the default behavior for queries specifying more than one condition.
  - Use `$and` if you need to include the same operator more than once in a query.

Example: Logical Operators

```
db.movies.find( { $or : [
   { "category" : "sci-fi" }, { "imdb_rating" : { $gte : 7 } }
 ] } )

// more complex $or, really good sci-fi movie or medicore family movie
db.movies.find( { $or : [
   { "category" : "sci-fi", "imdb_rating" : { $gte : 8 } },
   { "category" : "family", "imdb_rating" : { $gte : 7 } }
 ] } )

// find bad movies
db.movies.find( { "imdb_rating" : { $not : { $gt : 7 } } } )
```

Note:

- `db.movies.find( { "imdb_rating" : { $not : { $gt : 7 } } } )` also returns everything that doesnt have an "imdb_rating"

Example: Logical Operators

```
// find movies within an imdb_rating range
db.movies.find( { "imdb_rating" : { $gt : 5 , $lte : 7 } } )  // and is implicit

// queries can be nested, why are there no results?
db.movies.find( { $and : [
   { $or : [
     { "category" : "sci-fi", "imdb_rating" : { $gte : 8 } },
     { "category" : "family", "imdb_rating" : { $gte : 7 } }
   ] } ,
   { $or : [
     { "category" : "action", "imdb_rating" : { $gte : 6 } }
   ] }
 ] } )
```

Element Query Operators

- `$exists`: Select documents based on the existence of a particular field.
- `$type`: Select documents based on their type.
- See BSON types[5] for reference on types.

Example: Element Operators

```
db.movies.find( { "budget" : { $exists : true } } )

// type 1 is Double
db.movies.find( { "budget" : { $type : 1 } } )

// type 3 is Object (embedded document)
db.movies.find( { "budget" : { $type : 3 } } )
```

Array Query Operators

- `$all`: Array field must contain all values listed.
- `$size`: Array must have a particular size. E.g., `$size : 2` means 2 elements in the array
- `$elemMatch`: All conditions must be matched by at least one element in the array

Example: Array Operators

```
db.movies.find( { "category" : { $all : [ "sci-fi", "action" ] } } )

db.movies.find( { "category" : { $size : 3 } } )
```

Example: $elemMatch

```
db.movies.insertOne( {
    "title" : "Raiders of the Lost Ark",
    "filming_locations" : [
      { "city" : "Los Angeles", "state" : "CA", "country" : "USA" },
      { "city" : "Rome", "state" : "Lazio", "country" : "Italy" },
      { "city" : "Florence", "state" : "SC", "country" : "USA" }
    ] } )

// This query is incorrect, it won't return what we want
db.movies.find( {
    "filming_locations.city" : "Florence",
    "filming_locations.country" : "Italy"
  } )

// $elemMatch is needed, now there are no results, this is expected
db.movies.find( {
    "filming_locations" : {
      $elemMatch : {
```

[5] http://docs.mongodb.org/manual/reference/bson-types

```
        "city" : "Florence",
        "country" : "Italy"
      } } } )
```

Note:

- Comparing the last two queries demonstrates `$elemMatch`.

2.4 Updating Documents

Learning Objectives

Upon completing this module students should understand

- The `replaceOne()` method
- The `updateOne()` method
- The `updateMany()` method
- The required parameters for these methods
- Field update operators
- Array update operators
- The concept of an upsert and use cases.
- The `findOneAndReplace()` and `findOneAndUpdate()` methods

The `replaceOne()` Method

- Takes one document and replaces it with another
  - But leaves the _id unchanged
- Takes two parameters:
  - A matching document
  - A replacement document
- This is, in some sense, the simplest form of update

Note:

- By "simplest," we mean that it's simple conceptually – that replacing a document is a sort of basic idea of how an update happens.
- We will later see update methods that will involve only changing some fields.

First Parameter to `replaceOne()`

- Required parameters for `replaceOne()`
  - The query parameter:
    - Use the same syntax as with `find()`
    - Only the first document found is replaced
- `replaceOne()` cannot delete a document

Second Parameter to `replaceOne()`

- The second parameter is the replacement parameter:
  - The document to replace the original document
- The _id must stay the same
- You must replace the entire document
  - You cannot modify just one field
  - Except for the _id

Note:

- If they try to modify the _id, it will throw an error

Example: `replaceOne()`

```
db.movies.insertOne( { title: "Batman" } )
db.movies.find()
db.movies.replaceOne( { title : "Batman" }, { imdb_rating : 7.7 } )
db.movies.find()
db.movies.replaceOne( { imdb_rating: 7.7 },
                      { title: "Batman", imdb_rating: 7.7 } )
db.movies.find()
db.movies.replaceOne( { }, { title: "Batman" } )
db.movies.find()  // back in original state
db.movies.replaceOne( { }, { _id : ObjectId() } )
```

Note:

- Ask the students why the first replace killed the `title` field
- Ask why the final replace failed

The `updateOne()` Method

- Mutate one document in MongoDB using `updateOne()`
  - Affects only the _first_ document found
- Two parameters:
  - A query document
    - same syntax as with `find()`
  - Change document
    - Operators specify the fields and changes

`$set` and `$unset`

- Use to specify fields to update for `UpdateOne()`
- If the field already exists, using `$set` will change its value
  - If not, `$set` will create it, set to the new value
- Only specified fields will change
- Alternatively, remove a field using `$unset`

Example (Setup)

```
db.movies.insertMany( [
  {
    "title" : "Batman",
    "category" : [ "action", "adventure" ],
    "imdb_rating" : 7.6,
    "budget" : 35
  },
  {
    "title" : "Godzilla",
    "category" : [ "action",
    "adventure", "sci-fi" ],
    "imdb_rating" : 6.6
  },
  {
    "title" : "Home Alone",
    "category" : [ "family", "comedy" ],
    "imdb_rating" : 7.4
  }
] )
```

Example: `$set` and `$unset`

```
db.movies.updateOne( { "title" : "Batman" },
                     { $set : { "imdb_rating" : 7.7 } } )
db.movies.updateOne( { "title" : "Godzilla" },
                     { $set : { "budget" : 1 } } )
db.movies.updateOne( { "title" : "Home Alone" },
                     { $set : { "budget" : 15,
                                "imdb_rating" : 5.5 } } )
db.movies.updateOne( { "title" : "Home Alone" },
                     { $unset :  { "budget" : 1 } } )
db.movies.find()
```

Update Operators

- `$inc`: Increment a field's value by the specified amount.
- `$mul`: Multiply a field's value by the specified amount.
- `$rename`: Rename a field.
- `$set`: Update one or more fields (already discussed).
- `$unset` Delete a field (already discussed).
- **`$min`: Updates the field value to a specified value if the specified value** is less than the current value of the field
- **`$max`: Updates the field value to a specified value if the specified value** is greater than the current value of the field
- `$currentDate`: Set the value of a field to the current date or timestamp.

Example: Update Operators

```
db.movies.updateOne( { title: "Batman" }, { $inc: { "imdb_rating" : 2 } } )
db.movies.updateOne( { title: "Home Alone" }, { $inc: { "budget" : 5 } } )
db.movies.updateOne( { title: "Batman" }, { $mul: { "imdb_rating" : 4 } } )
db.movies.updateOne( { title: "Batman" },
                     { $rename: { budget: "estimated_budget" } } )
db.movies.updateOne( { title: "Home Alone" }, { $min: { budget: 5 } } )
db.movies.updateOne( { title: "Home Alone" },
      { $currentDate : { last_updated: { $type: "timestamp" } } } )
// increment movie rating by 1
db.movie_mentions.updateOne( { title: "Batman" },
      { $inc:  { "imdb_rating" : 1 } } )
```

The `updateMany()` Method

- Takes the same arguments as `updateOne`
- Updates all documents that match
  - `updateOne` stops after the first match
  - `updateMany` continues until it has matched all

Warning: Without an appropriate index, you may scan every document in the collection.

Example: `updateMany()`

```
// let's start tracking the number of sequals for each movie
db.movies.updateOne( { }, { $set : { "sequels" : 0 } } )
db.movies.find()
// we need updateMany to change all documents
db.movies.updateMany( { }, { $set : { "sequels" : 0 } } )
db.movies.find()
```

Array Element Updates by Index

- You can use dot notation to specify an array index
- You will update only that element
  - Other elements will not be affected

Example: Update Array Elements by Index

```
// add a sample document to track mentions per hour
db.movie_mentions.insertOne(
    { "title" : "E.T.",
      "day" : ISODate("2015-03-27T00:00:00.000Z"),
      "mentions_per_hour" : [ 0, 0, 0, 0, 0, 0, 0,
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
      0, 0 ]
    } )

// update all mentions for the fifth hour of the day
db.movie_mentions.updateOne(
    { "title" : "E.T." } ,
    { "$set" :  { "mentions_per_hour.5" : 2300 } } )
```

Note:

- Pattern for time series data
- Displaying charts is easy
  - Can change granularity to by the minute, hour, day, etc.

Array Operators

- `$push`: Appends an element to the end of the array.
- `$pushAll`: Appends multiple elements to the end of the array.
- `$pop`: Removes one element from the end of the array.
- `$pull`: Removes all elements in the array that match a specified value.
- `$pullAll`: Removes all elements in the array that match any of the specified values.
- `$addToSet`: Appends an element to the array if not already present.

Note:

- These operators may be applied to array fields.

Example: Array Operators

```
db.movies.updateOne(
    { "title" : "Batman" },
    { $push : { "category" : "superhero" } } )
db.movies.updateOne(
    { "title" : "Batman" },
    { $pushAll : { "category" : [ "villain", "comic-based" ] } } )
db.movies.updateOne(
    { "title" : "Batman" },
    { $pop : { "category" : 1 } } )
db.movies.updateOne(
    { "title" : "Batman" },
    { $pull : { "category" : "action" } } )
db.movies.updateOne(
    { "title" : "Batman" },
    { $pullAll : { "category" : [ "villain", "comic-based" ] } } )
```

Note:

- Pass $pop a value of -1 to remove the first element of an array and 1 to remove the last element in an array.

The Positional $ Operator

- $[6] is a positional operator that specifies an element in an array to update.
- It acts as a placeholder for the first element that matches the query document.
- `$` replaces the element in the specified position with the value given.
- Example:

```
    db.<COLLECTION>.updateOne(
        { <array> : value ... },
        { <update operator> : { "<array>.$" : value } }
    )
```

[6] http://docs.mongodb.org/manual/reference/operator/update/postional

Example: The Positional $ Operator

```
// the "action" category needs to be changed to "action-adventure"
db.movies.updateMany( { "category": "action",  },
                      { $set: { "category.$" : "action-adventure" } } )
```

Upserts

- If no document matches a write query:
  - By default, nothing happens
  - With `upsert: true`, inserts one new document
- Works for `updateOne()`, `updateMany()`, `replaceOne()`
- Syntax:

```
db.<COLLECTION>.updateOne( <query document>,
                           <update document>,
                           { upsert: true } )
```

Upsert Mechanics

- Will update if documents matching the query exist
- Will insert if no documents match
  - Creates a new document using equality conditions in the query document
  - Adds an `_id` if the query did not specify one
  - Performs the write on the new document
- `updateMany()` will only create one document
  - If none match, of course

Example: Upserts

```
db.movies.updateOne( { "title" : "Jaws" },
                     { $inc: { "budget" : 5 } },
                     { upsert: true } )

db.movies.updateMany( { "title" : "Jaws II" },
                      { $inc: { "budget" : 5 } },
                      { upsert: true } )

db.movies.replaceOne( { "title" : "E.T.", "category" : [ "scifi" ] },
                      { "title" : "E.T.", "category" : [ "scifi" ], "budget" : 1 },
                      { upsert: true } )
```

Note:

- Note that an updateMany works just like updateOne when no matching documents are found.
- First query updates the document with “title” = “Jaws” by incrementing “budget”

- Second query: 1) creates a new document, 2) assigns an _id, 3) sets "title" to "Jaws II" 4) performs the update
- Third query: 1) creates a new document, 2) sets "title" : "Jaws III", 3) Set budget to 1

`save()`

- The `db.<COLLECTION>.save()` method is syntactic sugar
  - Similar to `replaceOne()`, querying the `_id` field
  - Upsert if `_id` is not in the collection
- Syntax:

```
db.<COLLECTION>.save( <document> )
```

Example: `save()`

- If the document in the argument does not contain an `_id` field, then the `save()` method acts like `insertOne()` method
  - An ObjectId will be assigned to the `_id` field.
- If the document in the argument contains an `_id` field: then the `save()` method is equivalent to a replaceOne with the query argument on `_id` and the upsert option set to true

```
// insert
db.movies.save( { "title" : "Beverly Hills Cops", "imdb_rating" : 7.3 })

// update with { upsert: true }
db.movies.save( { "_id" : 1234, "title" : "Spider Man", "imdb_rating" : 7.3 })
```

Note:

- A lot of users prefer to use update/insert, to have more explicit control over the operation

---

Be careful with `save()`

Careful not to modify stale data when using `save()`. Example:

```
db.movies.drop()
db.movies.insertOne( { "title" : "Jaws", "imdb_rating" : 7.3 } )

db.movies.find( { "title" : "Jaws" } )

// store the complete document in the application
doc = db.movies.findOne( { "title" : "Jaws" } )

db.movies.updateOne( { "title" : "Jaws"  }, { $inc: { "imdb_rating" : 2 } } )
db.movies.find()

doc.imdb_rating = 7.4
```

```
db.movies.save(doc)  // just lost our incrementing of "imdb_rating"
db.movies.find()
```

`findOneAndUpdate()` and `findOneAndReplace()`

- Update (or replace) one document and return it
  - By default, the document is returned pre-write
- Can return the state before or after the update
- Makes a read plus a write atomic
- Can be used with upsert to insert a document

`findOneAndUpdate()` and `findOneAndReplace()` Options

- The following are optional fields for the options document
- `projection:` `<document>` - select the fields to see
- `sort:` `<document>` - sort to select the first document
- `maxTimeoutMS:` `<number>` - how long to wait
  - Returns an error, kills operation if exceeded
- `upsert:` `<boolean>` if true, performs an upsert

Example: `findOneAndUpdate()`

```
db.worker_queue.findOneAndUpdate(
    { state : "unprocessed" },
    { $set: { "worker_id" : 123, "state" : "processing" } },
    { upsert: true } )
```

`findOneAndDelete()`

- Not an update operation, but fits in with `findOneAnd` ...
- Returns the document and deletes it.
- Example:

```
db.foo.drop();
db.foo.insertMany( [ { a : 1 }, { a : 2 }, { a : 3 } ] );
db.foo.find();  // shows the documents.
db.foo.findOneAndDelete( { a : { $lte : 3 } } );
db.foo.find();
```

3 Indexes

3.1 Index Fundamentals

Learning Objectives

Upon completing this module students should understand:

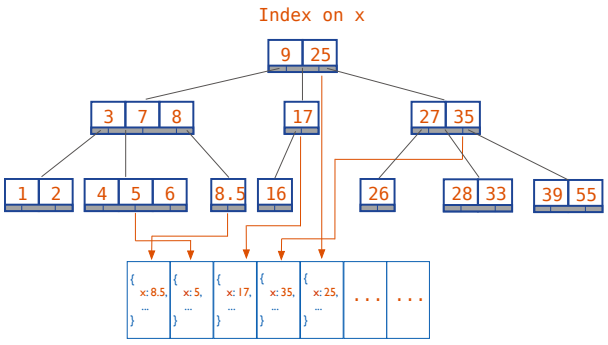
- The impact of indexing on read performance
- The impact of indexing on write performance
- How to choose effective indexes
- The utility of specific indexes for particular query patterns

---

Note:

- Ask how many people in the room are familiar with indexes in a relational database.
- If the class is already familiar with indexes, just explain that they work the same way in MongoDB.

Why Indexes?



Note:

- Without an index, in order to find all documents matching a query, MongoDB must scan every document in the collection.
- This is murder for read performance, and often write performance, too.
- If all your documents do not fit into memory, the system will page data in and out in order to scan the entire collection.

- An index enables MongoDB to locate exactly which documents match the query and where they are located on disk.
- MongoDB indexes are based on B-trees.

Types of Indexes

- Single-field indexes
- Compound indexes
- Multikey indexes
- Geospatial indexes
- Text indexes

Note:

- There are also hashed indexes and TTL indexes.
- We will discuss those elsewhere.

Exercise: Using `explain()`

Let's explore what MongoDB does for the following query by using `explain()`.

We are projecting only `user.name` so that the results are easy to read.

```
db.tweets.find( { "user.followers_count" : 1000 },
                { "_id" : 0, "user.name": 1 } )

db.tweets.find( { "user.followers_count" : 1000 } ).explain()
```

Note:

- Make sure the students are using the sample database.
- Review the structure of documents in the tweets collection by doing a find().
- We'll be looking at the user subdocument for documents in this collection.

Results of `explain()`

With the default `explain()` verbosity, you will see results similar to the following:

```
{
  "queryPlanner" : {
    "plannerVersion" : 1,
    "namespace" : "twitter.tweets",
    "indexFilterSet" : false,
    "parsedQuery" : {
      "user.followers_count" : {
        "$eq" : 1000
      }
    },
```

Results of `explain()` - Continued

```
  "winningPlan" : {
      "stage" : "COLLSCAN",
      "filter" : {
        "user.followers_count" : {
          "$eq" : 1000
        }
      },
      "direction" : "forward"
    },
    "rejectedPlans" : [ ]
  },
  ...
}
```

`explain()` Verbosity Can Be Adjusted

- **default:** determines the winning query plan but does not execute query
- **executionStats:** executes query and gathers statistics
- **allPlansExecution:** runs all candidate plans to completion and gathers statistics

---

Note:

- Default will be helpful if you're worried running the query could cause sever performance problems
- executionStats will be the most common verbosity level used
- allPlansExecution is for trying to determine WHY it is choosing the index it is (out of other candidates)

`explain("executionStats")`

```
> db.tweets.find( { "user.followers_count" : 1000 } )
  .explain("executionStats")
```

Now we have query statistics:

```
..
"executionStats" : {
  "executionSuccess" : true,
  "nReturned" : 8,
  "executionTimeMillis" : 107,
  "totalKeysExamined" : 0,
  "totalDocsExamined" : 51428,
  "executionStages" : {
    "stage" : "COLLSCAN",
    "filter" : {
      "user.followers_count" : {
        "$eq" : 1000
      }
    },
```

`explain("executionStats")` - Continued

```
    "nReturned" : 8,
    "executionTimeMillisEstimate" : 100,
    "works" : 51430,
    "advanced" : 8,
    "needTime" : 51421,
    "needFetch" : 0,
    "saveState" : 401,
    "restoreState" : 401,
    "isEOF" : 1,
    "invalidates" : 0,
    "direction" : "forward",
    "docsExamined" : 51428
  }
  ...
}
```

`explain("executionStats")` Output

- `nReturned` : number of documents returened by the query
- `totalDocsExamined` : number of documents touched during the query
- `totalKeysExamined` : number of index keys scanned
- A `totalKeysExamined` or `totalDocsExamined` value much higher than `nReturned` indicates we need a better index
- Based `.explain()` output, this query would benefit from a better index

---

Note:

- By documents "touched", we mean that they had to be in memory (either already there, or else loaded during the query)
- By "better" index, we mean one that matches the query more closely.

Other Operations

In addition to `find()`, we often want to use `explain()` to understand how other operations will be handled.

- `aggregate()`
- `count()`
- `group()`
- `update()`
- `remove()`
- `findAndModify()`
- `insert()`

Note:

- Has not yet been implemented for the new CRUD API.
  - No `updateOne()`, `replaceOne()`, `updateMany()`, `deleteOne()`, `deleteMany()`, `findOneAndUpdate()`, `findOneAndDelete()`, `findOneAndReplace()`, `insertMany()`

`db.<COLLECTION>.explain()`

`db.<COLLECTION>.explain()` returns an ExplainableCollection.

```
> var explainable = db.tweets.explain()
> explainable.find( { "user.followers_count" : 1000 } )
```

equivalent to

```
> db.tweets.explain().find( { "user.followers_count" : 1000 } )
```

also equivalent to

```
> db.tweets.find( { "user.followers_count" : 1000 } ).explain()
```

Note:

- This will get confusing for students, may want to spend extra time here with more examples

---

Using `explain()` for Write Operations

Simulate the number of writes that would have occurred and determine the index(es) used:

```
> db.tweets.explain("executionStats").remove( { "user.followers_count" : 1000 } )
```

```
> db.tweets.explain("executionStats").update( { "user.followers_count" : 1000 },
  { $set : { "large_following" : true } }, { multi: true } )
```

Note:

- Walk through the "nWouldModify" field in the output to show how many documents would have been updated

---

Single-Field Indexes

- Single-field indexes are based on a single field of the documents in a collection.
- The field may be a top-level field.
- You may also create an index on fields in embedded documents.

Creating an Index

The following creates a single-field index on `user.followers_count`.

```
db.tweets.createIndex( { "user.followers_count" : 1 } )
db.tweets.find( { "user.followers_count" : 1000 } ).explain()
```

`explain()` indicated there will be a substantial performance improvement in handling this type of query.

Note:

- nscannedObjects should now be a much smaller number, e.g., 8.
- Operations teams are accustomed to thinking about indexes.
- With MongoDB, developers need to be more involved in the creation and use of indexes.

Listing Indexes

List indexes for a collection:

```
db.tweets.getIndexes()
```

List index keys:

```
db.tweets.getIndexKeys()
```

Indexes and Read/Write Performance

- Indexes improve read performance for queries that are supported by the index.
- Inserts will be slower when there are indexes that MongoDB must also update.
- The speed of updates may be improved because MongoDB will not need to do a collection scan to find target documents.
- An index is modified any time a document:
  - Is inserted (applies to *all* indexes)
  - Is deleted (applies to *all* indexes)
  - Is updated in such a way that its indexed field changes

Note:

- For mmapv1, all indexes will be modified whenever the document moves on disk
  - i.e., When it outgrows its record space

Index Limitations

- You can have up to 64 indexes per collection.
- You should NEVER be anywhere close to that upper bound.
- Write performance will degrade to unusable at somewhere between 20-30.

Use Indexes with Care

- Every query should use an index.
- Every index should be used by a query.
- Any write that touches an indexed field will update every index that touches that field.
- Indexes require RAM.
- Be mindful about the choice of key.

Note:

- If your system has limited RAM, then using the index will force other data out of memory.
- When you need to access those documents, they will need to be paged in again.

Additional Index Options

- Sparse
- Unique
- Background

Sparse Indexes in MongoDB

- Sparse indexes only contain entries for documents that have the indexed field.

```
db.<COLLECTION>.createIndex(
  { field_name : 1 },
  { sparse : true } )
```

Note:

- Partial indexes are now preferred to sparse.
- You can create the functional equivalent of a sparce index with `{ field : { $exists : true } }` for your partialFilterExpression.

Defining Unique Indexes

- Enforce a unique constraint on the index
  - On a per-collection basis
- Can't insert documents with a duplicate value for the field
  - Or update to a duplicate value
- No duplicate values may exist prior to defining the index

```
db.<COLLECTION>.createIndex(
  { field_name : 1 },
  { unique : true } )
```

Note:

- Unique indexes do not work well with multikey indexes
- The following would produce a collision and return an error

```
db.test.insertMany([
  {a: [1, 2, 3]},
  {a: [2, 4, 6]}
])
db.test.createIndex({a: 1}, {unique: true})
```

- Unique indexes do not evaluate subdocument contents
- The following would **not** produce an error

```
db.test.insertMany([
  {a: {b: 2, c: 3}},
  {a: {c: 3, b: 2}}
])
db.test.createIndex({a: 1}, {unique: true})
```

Building Indexes in the Background

- Building indexes in foreground is a blocking operation.
- Background index creation is non-blocking, however, takes longer to build.
- Initially larger, or less compact, than an index built in the foreground.

```
db.<COLLECTION>.createIndex(
  { field_name : 1 },
  { background : true } )
```

3.2 Compound Indexes

Learning Objectives

Upon completing this module students should understand:

- What a compound index is.
- How compound indexes are created.
- The importance of considering field order when creating compound indexes.
- How to efficiently handle queries involving some combination of equality matches, ranges, and sorting.
- Some limitations on compound indexes.

Introduction to Compound Indexes

- It is common to create indexes based on more than one field.
- These are called `compound indexes`.
- You may use up to 31 fields in a compound index.
- You may not use hashed index fields.

The Order of Fields Matters

Specifically we want to consider how the index will be used for:

- Equality tests, e.g.,

```
db.movies.find( { "budget" : 7, "imdb_rating" : 8 } )
```

- Range queries, e.g.,

```
db.movies.find( { "budget" : 10, "imdb_rating" : { $lt : 9 } } )
```

- Sorting, e.g.,

```
db.movies.find( { "budget" : 10, "imdb_rating" : 6 }
               ).sort( { "imdb_rating" : -1 } )
```

Note:

- The order in which the fields are specified is of critical importance.
- It is especially important to consider query patterns that require two or more of these operations.

Designing Compound Indexes

- Let's look at some guiding principles for building compound indexes.
- These will generally produce a good if not optimal index.
- You can optimize after a little experimentation.
- We will explore this in the context of a running example.

Example: A Simple Message Board

Requirements:

- Find all messages in a specified timestamp range.
- Select for whether the messages are anonymous or not.
- Sort by rating from highest to lowest.

Load the Data

```
a = [ { "timestamp" : 1, "username" : "anonymous", "rating" : 3 },
      { "timestamp" : 2, "username" : "anonymous", "rating" : 5 },
      { "timestamp" : 3, "username" : "sam", "rating" : 1 },
      { "timestamp" : 4, "username" : "anonymous", "rating" : 2 },
      { "timestamp" : 5, "username" : "martha", "rating" : 5 } ]
db.messages.insertMany(a)
```

Start with a Simple Index

Start by building an index on { timestamp : 1 }

```
db.messages.createIndex( { timestamp : 1 }, { name : "myindex" } )
```

Now let's query for messages with timestamp in the range 2 through 4 inclusive.

```
db.messages.find( { timestamp : { $gte : 2, $lte : 4 } } ).explain("executionStats")
```

Analysis:

- Explain plan shows good performance, i.e. `totalKeysExamined` = `n`.
- However, this does not satisfy our query.
- Need to query again with `{username: "anonymous"}` as part of the query.

Query Adding `username`

Let's add the `user` field to our query.

```
db.messages.find( { timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous" } ).explain("executionStats")
```

`totalKeysExamined` > `n`.

Include `username` in Our Index

```
db.messages.dropIndex( "myindex" );
db.messages.createIndex( { timestamp : 1, username : 1 },
                         { name : "myindex" } )
db.messages.find( { timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous" } ).explain("executionStats")
```

`totalKeysExamined` is still > `n`. Why?

`totalKeysExamined` > `n`

timestamp	username
1	"anonymous"
2	"anonymous"
3	"sam"
4	"anonymous"
5	"martha"

Note:

- The index we have created stores the range values before the equality values.
- The documents with timestamp values 2, 3, and 4 were found first.
- Then the associated anonymous values had to be evaluated.

A Different Compound Index

Drop the index and build a new one with user.

```
db.messages.dropIndex( "myindex" );
db.messages.createIndex( { username : 1 , timestamp : 1 },
                         { name : "myindex" } )

db.messages.find( { timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous" } ).explain("executionStats")
```

`totalKeysExamined` is 2. `n` is 2.

`totalKeysExamined == n`

username	timestamp
“anonymous”	1
“anonymous”	2
“anonymous”	4
“sam”	2
“martha”	5

Note:

- This illustrates why.
- There is a fundamental difference in the way the index is structured.
- This supports a more efficient treatment of our query.

Let Selectivity Drive Field Order

- Order fields in a compound index from most selective to least selective.
- Usually, this means equality fields before range fields.
- When dealing with multiple equality values, start with the most selective.
- If a common range query is more selective instead (rare), specify the range component first.

Adding in the Sort

Finally, let’s add the sort and run the query

```
db.messages.find( {
                    timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous"
                  } ).sort( { rating : -1 } ).explain("executionStats");
```

- Note that the `winningPlan` includes a `SORT` stage
- This means that MongoDB had to perform a sort in memory
- In memory sorts on can degrade performance significantly
  - Especially if used frequently
  - In-memory sorts that use > 32 MB will abort

In-Memory Sorts

Let's modify the index again to allow the database to sort for us.

```
db.messages.dropIndex( "myindex" );
db.messages.createIndex( { username : 1 , timestamp : 1, rating : 1 },
                         { name : "myindex" } );
db.messages.find( {
                    timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous"
                  } ).sort( { rating : -1 } ).explain("executionStats");
```

- The explain plan remains unchanged, because the sort field comes after the range fields.
- The index does not store entries in order by rating.
- Note that this requires us to consider a tradeoff.

Avoiding an In-Memory Sort

Rebuild the index as follows.

```
db.messages.dropIndex( "myindex" );
db.messages.createIndex( { username : 1, rating : 1, timestamp : 1 },
                         { name : "myindex" } );
db.messages.find( {
                    timestamp : { $gte : 2, $lte : 4 },
                    username : "anonymous"
                  } ).sort( { rating : -1 } ).explain("executionStats");
```

- We no longer have an in-memory sort, but need to examine more keys.
- `totalKeysExamined` is 3 and and `n` is 2.
- This is the best we can do in this situation and this is fine.
- However, if `totalKeysExamined` is much larger than `n`, this might not be the best index.

No need for `stage : SORT`

username	rating	timestamp
"anonymous"	2	4
"anonymous"	3	1
"anonymous"	5	2
"sam"	1	2
"martha"	5	5

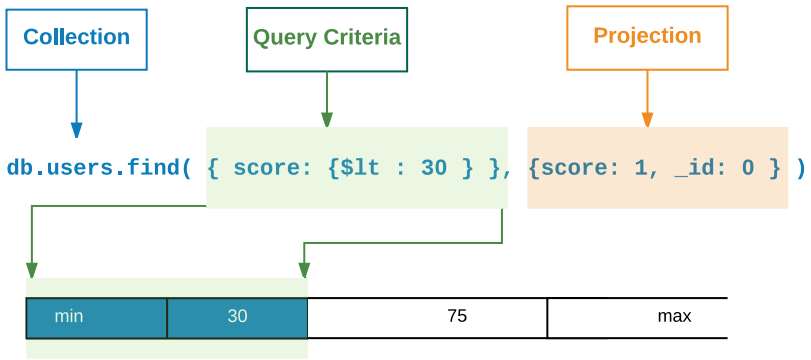
Note:

- general index illustration
- we can see that the returned results are already sorted.
- no need to perform an in-memory sort

General Rules of Thumb

- Equality before range
- Equality before sorting
- Sorting before range

Covered Queries



- When a query and projection include only the indexed fields, MongoDB will return results directly from the index.
- There is no need to scan any documents or bring documents into memory.
- These covered queries can be very efficient.

Exercise: Covered Queries

```
db.testcol.drop()
for (i=1; i<=20; i++) {
  db.testcol.insertOne({ "_id" : i, "title" : i, "name" : i,
                         "rating" : i, "budget" : i })
};
db.testcol.createIndex( { "title" : 1, "name" : 1, "rating" : 1 } )

// Not covered because _id is  present.
db.testcol.find( { "title" : 3 },
                 { "title" : 1, "name" : 1, "rating" : 1 }
                 ).explain("executionStats")

// Not covered because other fields may exist in matching docs.
db.testcol.find( { "title" : 3 },
                 { "_id" : 0, "budget" : 0 } ).explain("executionStats")

// Covered query!
db.testcol.find( { "title" : 3 },
                 { "_id" : 0, "title" : 1, "name" : 1, "rating" : 1 }
                 ).explain("executionStats")
```

3.3 Multikey Indexes

Learning Objectives

Upon completing this module, students should understand:

- What a multikey index is
- When MongoDB will use a multikey index to satisfy a query
- How multikey indexes work
- How multikey indexes handle sorting
- Some limitations on multikey indexes

Introduction to Multikey Indexes

- A multikey index is an index on an array.
- An index entry is created on each value found in the array.
- Multikey indexes can support primitives, documents, or sub-arrays.
- There is nothing special that you need to do to create a multikey index.
- You create them using `createIndex()` just as you would with an ordinary single-field index.
- If there is an array as a value for an indexed field, the index will be multikey on that field.

Example: Array of Numbers

```
db.race_results.drop()
db.race_results.createIndex( { "lap_times" : 1 } )
a = [ { "lap_times" : [ 3, 5, 2, 8 ] },
      { "lap_times" : [ 1, 6, 4, 2 ] },
      { "lap_times" : [ 6, 3, 3, 8 ] } ]
db.race_results.insertMany( a )

// Used the index
db.race_results.find( { lap_times : 1 } ).explain()

// One document found.
// Index not used, because it is naive to position.
db.race_results.find( { "lap_times.2" : 3 } ).explain()
```

Exercise: Array of Documents, Part 1

Create a collection and add an index on the `comments.rating` field:

```
db.blog.drop()
b = [ { "comments" : [
          { "name" : "Bob", "rating" : 1 },
          { "name" : "Frank", "rating" : 5.3 },
          { "name" : "Susan", "rating" : 3 } ] },
      { "comments" : [
          { name : "Megan", "rating" : 1 } ] },
      { "comments" : [
          { "name" : "Luke", "rating" : 1.4 },
          { "name" : "Matt", "rating" : 5 },
          { "name" : "Sue", "rating" : 7 } ] }]
db.blog.insertMany(b)

db.blog.createIndex( { "comments" : 1 } )
// vs
db.blog.createIndex( { "comments.rating" : 1 } )

// for this query
db.blog.find( { "comments.rating" : 5 })
```

Note:

- Note: JSON is a dictionary and doesn't guarantee order, indexing the top level array (comments array) won't work

---

Exercise: Array of Documents, Part 2

For each of the three queries below:

- How many documents will be returned?
- Will it use our multi-key index? Why or why not?
- If a query will not use the index, which index will it use?

```
db.blog.find( { "comments" : { "name" : "Bob", "rating" : 1 } } )
db.blog.find( { "comments" : { "rating" : 1 } } )
db.blog.find( { "comments.rating" : 1 } )
```

Note:

```
// Never do this, won't give you the results expected
// JSON is a dictionary, and won't preserve ordering, second query will return no␣
↪results

db.blog.find( { "comments" : { "name" : "Bob", "rating" : 1 } } )
db.blog.find( { "comments" : { "rating" : 1 } } )
```

Exercise: Array of Arrays, Part 1

Add some documents and create an index simulating a player in a game moving on an X,Y grid.

```
db.player.drop()
db.player.createIndex( { "last_moves" : 1 } )
c = [ { "last_moves" : [ [ 1, 2 ], [ 2, 3 ], [ 3, 4] ] },
      { "last_moves" : [ [ 3, 4 ], [ 4, 5 ] ] },
      { "last_moves" : [ [ 4, 5 ], [ 5, 6 ] ] },
      { "last_moves" : [ [ 3, 4 ] ] },
      { "last_moves" : [ [ 4, 5 ] ] } ]
db.player.insertMany(c)
db.player.find()
```

Exercise: Array of Arrays, Part 2

For each of the queries below:

- How many documents will be returned?
- Does the query use the multi-key index? Why or why not?
- If the query does not use the index, what is an index it could use?

```
db.player.find( { "last_moves" : [ 3, 4 ] } )
db.player.find( { "last_moves" : 3 } )
db.player.find( { "last_moves.1" : [ 4, 5 ] } )
db.player.find( { "last_moves.2" : [ 2, 3 ] } )
```

Note:

```
// 3 documents
db.player.find( { "last_moves" : [ 3, 4 ] } )
// Uses the multi-key index
db.player.find( { "last_moves" : [ 3, 4 ] } ).explain()

// No documents
db.player.find( { "last_moves" : 3 } )

// Does not use the multi-key index, because it is naive to position.
db.player.find( { "last_moves.1" : [ 4, 5 ] } )
db.player.find( { "last_moves.2" : [ 2, 3 ] } )
```

How Multikey Indexes Work

- Each array element is given one entry in the index.
- So an array with 17 elements will have 17 entries – one for each element.
- Multikey indexes can take up much more space than standard indexes.

Multikey Indexes and Sorting

- If you sort using a multikey index:
  - A document will appear at the first position where a value would place the document.
  - It will not appear multiple times.
- This applies to array values generally.
- It is not a specific property of multikey indexes.

Exercise: Multikey Indexes and Sorting

```
db.testcol.drop()
a = [ { x : [ 1, 11 ] }, { x : [ 2, 10 ] }, { x : [ 3 ] },
      { x : [ 4 ] }, { x : [ 5 ] } ]
db.testcol.insert(a)

db.testcol.createIndex( { x : 1 } )

// x : [ 1, 11 ] array comes first. It contains the lowest value.
db.testcol.find().sort( { x : 1 } )

// x : [ 1, 11 ] array still comes first. Contains the highest value.
db.testcol.find().sort( { x : -1 } )
```

Note:

```
// x : [ 1, 11 ] array comes first, because it contains the lowest value
db.testcol.find().sort( { x : 1 } )

// x : [ 1, 11 ] array still comes first, because it contains the highest value
db.testcol.find().sort( { x : -1 } )
```

Limitations on Multikey Indexes

- You cannot create a compound index using more than one array-valued field.
- This is because of the combinatorics.
- For a compound index on two array-valued fields you would end up with N * M entries for one document.
- You cannot have a hashed multikey index.
- You cannot have a shard key use a multikey index.
- We discuss shard keys in another module.
- The index on the `_id` field cannot become a multikey index.

Example: Multikey Indexes on Multiple Fields

```
db.testcol.drop()
db.testcol.createIndex( { x : 1, y : 1 } )

// no problems yet
db.testcol.insertOne( { _id : 1, x : 1, y : 1 } )

// still OK
db.testcol.insertOne( { _id : 2, x : [ 1, 2 ], y : 1 } )

// still OK
db.testcol.insertOne( { _id : 3, x : 1, y : [ 1, 2 ] } )

// Won't work
db.testcol.insertOne( { _id : 4, x : [ 1, 2 ], y : [ 1, 2 ] } )
```

3.4 Text Indexes

Learning Objectives

Upon completing this module, students should understand:

- The purpose of a text index
- How to create text indexes
- How to search using text indexes
- How to rank search results by relevance score

What is a Text Index?

- A text index is based on the tokens (words, etc.) used in string fields.
- MongoDB supports text search for a number of languages.
- Text indexes drop language-specific stop words (e.g. in English “the”, “an”, “a”, “and”, etc.).
- Text indexes use simple, language-specific suffix stemming (e.g., “running” to “run”).

Creating a Text Index

You create a text index a little bit differently than you create a standard index.

```
db.<COLLECTION>.createIndex( { <field name> : "text" } )
```

Exercise: Creating a Text Index

Create a text index on the “dialog” field of the montyPython collection.

```
db.montyPython.createIndex( { dialog : "text" } )
```

Creating a Text Index with Weighted Fields

- Default weight of 1 per indexed field.
- Weight is relative to other weights in text index.

```
db.<COLLECTION>.createIndex(
 { "title" : "text", "keywords": "text", "author" : "text" },
 { "weights" : {
   "title" : 10,
   "keywords" : 5
 }})
```

- Term match in “title” field has 10 times (i.e. 10:1) the impact as a term match in the “author” field.

Creating a Text Index with Weighted Fields

- The default weight is 1 for each indexed field.
- The weight is relative to other weights in a text index.

```
db.<COLLECTION>.createIndex(
    { "title" : "text", "keywords": "text", "author" : "text" },
    { "weights" : {
        "title" : 10,
        "keywords" : 5
    }})
```

- Term match in “title” field has 10 times (i.e. 10:1) the impact as a term match in the “author” field.

Text Indexes are Similar to Multikey Indexes

- Continuing our example, you can treat the `dialog` field as a multikey index.
- A multikey index with each of the words in `dialog` as values.
- You can query the field using the `$text` operator.

Exercise: Inserting Texts

Let's add some documents to our montyPython collection.

```
db.montyPython.insertMany( [
{ _id : 1,
  dialog : "What is the air-speed velocity of an unladen swallow?" },
{ _id : 2,
  dialog : "What do you mean? An African or a European swallow?" },
{ _id : 3,
  dialog : "Huh? I... I don't know that." },
{ _id : 45,
  dialog : "You're using coconuts!" },
{ _id : 55,
  dialog : "What? A swallow carrying a coconut?" } ] )
```

Querying a Text Index

Next, let's query the collection. The syntax is:

```
db.<COLLECTION>.find( { $text : { $search : "query terms go here" } } )
```

Exercise: Querying a Text Index

Using the text index, find all documents in the montyPython collection with the word "swallow" in it.

```
// Returns 3 documents.
db.montyPython.find( { $text : { $search : "swallow" } } )
```

Exercise: Querying Using Two Words

- Find all documents in the montyPython collection with either the word 'coconut' or 'swallow'.
- By default MongoDB ORs query terms together.
- E.g., if you query on two words, results include documents using either word.

```
// Finds 4 documents, 3 of which contain only one of the two words.
db.montyPython.find( { $text : { $search : "coconut swallow" } } )
```

Search for a Phrase

- To match an exact phrase, include search terms in quotes (escaped).
- The following query selects documents containing the phrase “European swallow”:

```
db.montyPython.find( { $text: { $search: "\"European swallow\"" } } )
```

Text Search Score

- The search algorithm assigns a relevance score to each search result.
- The score is generated by a vector ranking algorithm.
- The documents can be sorted by that score.

```
db.<COLLECTION>.find(
    { $text : { $search : "swallow coconut"} },
    { textScore: {$meta : "textScore" } }
).sort(
        { textScore: { $meta: "textScore" } }
) )
```

3.5 Lab: Using `explain()`

Exercise: explain(“executionStats”)

Drop all indexes from previous exercises:

```
mongo performance
> db.sensor_readings.dropIndexes()
```

Create an index for the “active” field:

```
db.sensor_readings.createIndex({ "active" : 1 } )
```

How many index entries and documents are examined for the following query? How many results are returned?

```
db.sensor_readings.find(
                          { "active": false, "_id": { $gte: 99, $lte: 1000 } }
                       ).explain("executionStats")
```

3.6 Lab: Finding and Addressing Slow Operations

Set Up

- In this exercise let's bring up a `mongo` shell with the following instructions

```
mongo --shell localhost/performance performance.js
```

In the shell that launches execute the following method

```
performance.init()
```

Exercise: Determine Indexes Needed

- In a mongo shell run `performance.b()`. This will run in an infinite loop printing some output as it runs various statements against the server.
- Now imagine we have detected a performance problem and suspect there is a slow operation running.
- Find the slow operation and terminate it. Every slow operation is assumed to run for 100ms or more.
- In order to do this, open a second window (or tab) and run a second instance of the mongo shell.
- What indexes can we introduce to make the slow queries more efficient? Disregard the index created in the previous exercises.

Note:

- Look at the logs to find queries over 100ms
- { "active": 1 }
- { "str": 1, "x": 1 }

---

4 Replica Sets

4.1 Introduction to Replica Sets

Learning Objectives

Upon completing this module, students should understand:

- Striking the right balance between cost and redundancy
- The many scenarios replication addresses and why
- How to avoid downtime and data loss using replication

Use Cases for Replication

- High Availability
- Disaster Recovery
- Functional Segregation

High Availability (HA)

- Data still available following:
  - Equipment failure (e.g. server, network switch)
  - Datacenter failure
- This is achieved through automatic failover.

Note: If we lose a server and MongoDB is correctly configured:

- Our database system can still service reads and writes, but by default not during failover period when the election takes place and there is no primary.
- Without manual intervention as long as there is still a majority of nodes available.

Disaster Recovery (DR)

- We can duplicate data across:
  - Multiple database servers
  - Storage backends
  - Datacenters
- Can restore data from another node following:
  - Hardware failure
  - Service interruption

Functional Segregation

There are opportunities to exploit the topology of a replica set:

- Based on physical location (e.g. rack or datacenter location)
- For analytics, reporting, data discovery, system tasks, etc.
- For backups

Note:

- You may direct particular queries to specific nodes (which may have different indexes or hardware) to increase overall performance.
- Backup data from secondaries to avoid performance penalties on the primary, especially when using tools like `mongodump` which are I/O intensive and evict the working set from memory (significant when data size is larger than RAM and disks are slow).
- Dedicate secondaries for other purposes such as analytics jobs.

Large Replica Sets

Functional segregation can be further exploited by using large replica sets.

- 50 node replica set limit with a maximum of 7 voting members
- Useful for deployments with a large number of data centers or offices
- Read only workloads can position secondaries in data centers around the world (closer to application servers)

Note:

- Sample use case: bank reference data distributed to 20+ data centers around the world, then consumed by the local application server

---

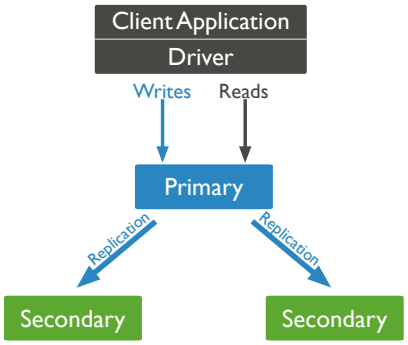
Replication is Not Designed for Scaling

- Can be used for scaling reads, but generally not recommended.
- Drawbacks include:
  - Eventual consistency
  - Not scaling writes
  - Potential system overload when secondaries are unavailable
- Consider sharding for scaling reads and writes.

Note:

- Eventual consistency: This is often tricky to handle as most developers are used to the strong consistency that most databases (and MongoDB) have by default. It also raises the question of how stale the data can be, and what to do when it crosses the threshold (e.g. fall back to reading from the primary). As of MongoDB 3.4, queries can be executed with **readConcern: linearizable** to ensure non-stale reads under certain circumstances. More on this will be covered later.
- Potential system overload: For example, consider a 3 data node replica set using secondaries to scale reads. Each node is serving reads at 70% of its capacity, which is a reasonable utilization rate. What happens if one of the secondaries fail or is intentionally taken down for maintenance (e.g. upgrading the OS or MongoDB)? Even if the load splits evenly between the 2 remaining nodes, they will be at `70+(70/2)= 105%` capacity.

Replica Sets



Note:

- MongoDB implements replication in the form of replica sets. Don't use the term master-slave as that is what we had before replica sets. It still exists for some corner cases (e.g. > 12 replicas) but should otherwise be avoided.
- A replica set consists of one or more `mongod` servers. Maximum 50 nodes in total and up to 7 with votes.
- There is at most one `mongod` that is "primary" at any one time (though there are edge cases/bugs when there is more than one).
- There are usually two or more other `mongod` instances that are secondaries.
- Secondaries may become primary if there is a failover event of some kind.
- Failover is automatic when correctly configured and a majority of nodes remain.

- The secondaries elect a new primary automatically. A primary may also voluntarily step down, like when it can no longer reach the majority of nodes to avoid a potential split brain scenario.

Primary Server

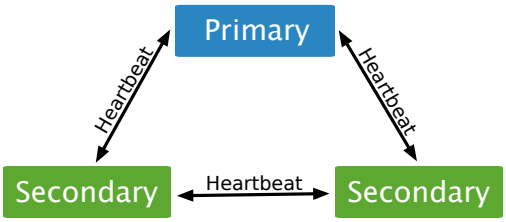
- Clients send writes to the primary only.
- MongoDB, Inc. maintains client drivers in many programming languages like Java, C#, Javascript, Python, Ruby, and PHP.
- MongoDB drivers are replica set aware.

Note: If the primary for a replica set changes from one node to another, the driver will automatically route writes to the correct `mongod`.

Secondaries

- A secondary replicates operations from another node in the replica set.
- Secondaries usually replicate from the primary.
- Secondaries may also replicate from other secondaries. This is called replication chaining.
- A secondary may become primary as a result of a failover scenario.

Heartbeats



Note:

- The members of a replica set use heartbeats to determine if they can reach every other node.
- The heartbeats are sent every two seconds.
- If a node is unreachable, this may indicate server failure, a network partition, or simply too slow to respond. The heartbeat will timeout and retried several times before the state is updated.

The Oplog

- The operations log, or oplog, is a special capped collection that is the basis for replication.
- The oplog maintains one entry for each document affected by every write operation.
- Secondaries copy operations from the oplog of their sync source.

Note: Remind students that capped collections are collections that do not indefinitely expand. Once their maximum size is reached they were roll back to the beginning (oldest insert) and start overwriting data.

Initial Sync

- Occurs when a new server is added to a replica set, or we erase the underlying data of an existing server (–dbpath)
- All existing collections except the *local* collection are copied
- As of MongoDB >= 3.4, all indexes are built while data is copied
- As of MongoDB >= 3.4, initial sync is more resilient to intermittent network failure/degradation

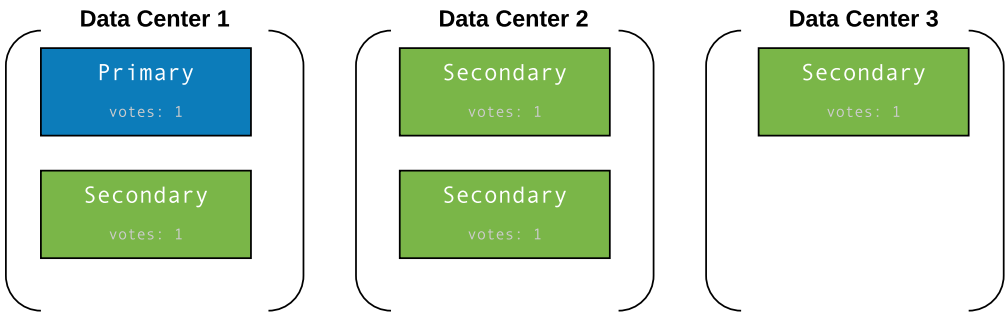
4.2 Elections in Replica Sets

Learning Objectives

Upon completing this module students should understand:

- That elections enable automated failover in replica sets
- How votes are distributed to members
- What prompts an election
- How a new primary is selected

Members and Votes

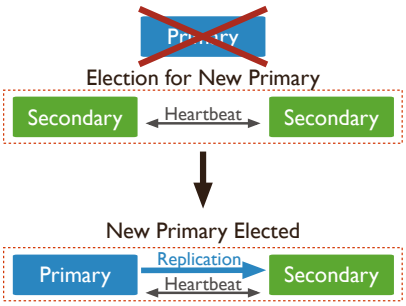


Note:

- In order for writes to occur, one member of a replica set must be primary.
- In the event the current primary becomes unavailable, the remaining members elect a new primary.
- Voting members of replica set each get one vote.

- Up to seven members may be voting members.
- This enables MongoDB to ensure elections happen quickly, but enables distribution of votes to different data centers.
- In order to be elected primary a server must have a true majority of votes.
- A member must have greater than 50% of the votes in order to be elected primary.

Calling Elections



Note:

- MongoDB uses a consensus protocol to determine when an election is required.
- Essentially, an election will occur if there is no primary.
- Upon initiation of a new replica set the members will elect a primary.
- If a primary steps down the set will hold an election.
- A secondary will call for an election if it does not receive a response to a heartbeat sent to the primary after waiting for 10 seconds.
- If other members agree that the primary is not available, an election will be held.

Selecting a New Primary

- Depends on which replication protocol version is in use
- PV0
  - Priority
  - Optime
  - Connections
- PV1
  - Optime
  - Connections

Priority

- PV0 factors priority into voting.
- The higher its priority, the more likely a member is to become primary.
- The default is 1.
- Servers with a priority of 0 will never become primary.
- Priority values are floating point numbers 0 - 1000 inclusive.

Note:

- Priority is a configuration parameter for replica set members.
- Use priority to determine where writes will be directed by default.
- And where writes will be directed in case of failover.
- Generally all identical nodes in a datacenter should have the same priority to avoid unnecessary failovers. For example, when a higher priority node rejoins the replica set after a maintenance or failure event, it will trigger a failover (during which by default there will be no reads and writes) even though it is unnecessary.
- More on this in a later module.
- PV1 does not factor priority into elections. However, after the replica set is stable, a secondary with higher priority will call for a new election to make itself the new primary. This increases the chance of there always being a primary, at the cost of potentially more elections.

Optime

- Optime: Operation time, which is the timestamp of the last operation the member applied from the oplog.
- To be elected primary, a member must have the most recent optime.
- Only optimes of visible members are compared.

Connections

- Must be able to connect to a majority of the members in the replica set.
- Majority refers to the total number of votes.
- Not the total number of members.

Note: To be elected primary, a replica set member must be able to connect to a majority of the members in the replica set.

When will a primary step down?

- After receiving the `replSetStepDown` or `rs.stepDown()` command.
- If a secondary is eligible for election and has a higher priority.
- If it cannot contact a majority of the members of the replica set.

replSetStepDown Behavior

- Primary will attempt to terminate long running operations before stepping down.
- Primary will wait for electable secondary to catch up before stepping down.
- "secondaryCatchUpPeriodSecs" can be specified to limit the amount of time the primary will wait for a secondary to catch up before the primary steps down.

Note:

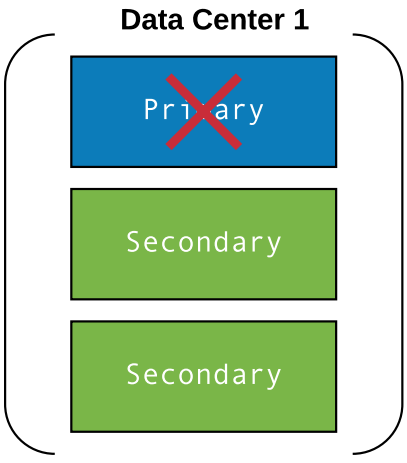
- Ask the class what the tradeoffs could be in setting secondaryCatchUpPeriodSecs to a very short amount of time (rollbacks could occur or operations not replicated)

Exercise: Elections in Failover Scenarios

- We have learned about electing a primary in replica sets.
- Let's look at some scenarios in which failover might be necessary.

Scenario A: 3 Data Nodes in 1 DC

Which secondary will become the new primary?

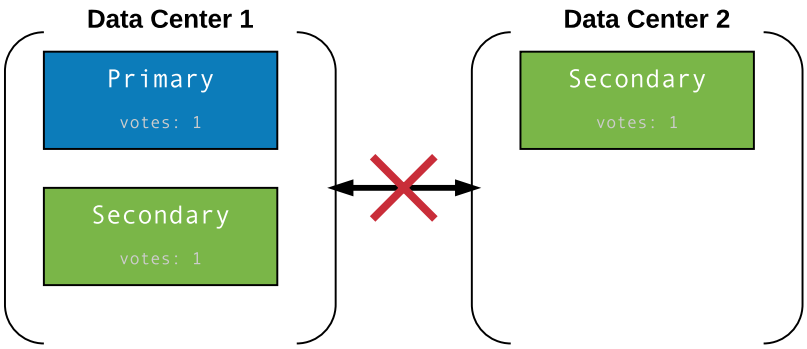


Note:

- It depends on the priorities of the secondaries.
- And on the optime.

Scenario B: 3 Data Nodes in 2 DCs

Which member will become primary following this type of network partition?

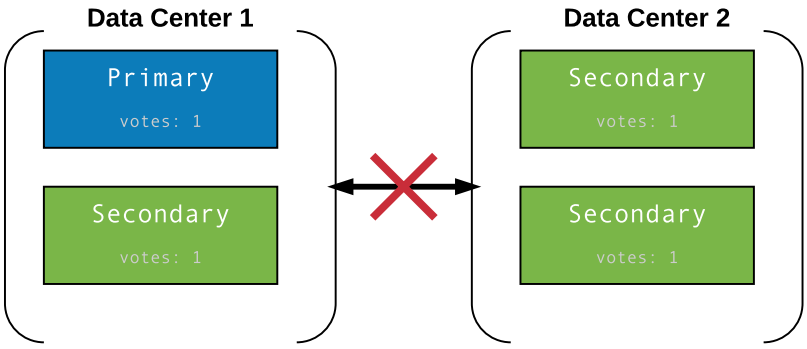


Note:

- The current primary is likely to remain primary.
- It probably has the highest priority.
- If DC2 fails, we still have a primary.
- If DC1 fails, we won't have a primary automatically. The remaining node in DC2 needs to be manually promoted by reconfiguring the replica set.

Scenario C: 4 Data Nodes in 2 DCs

What happens following this network partition?



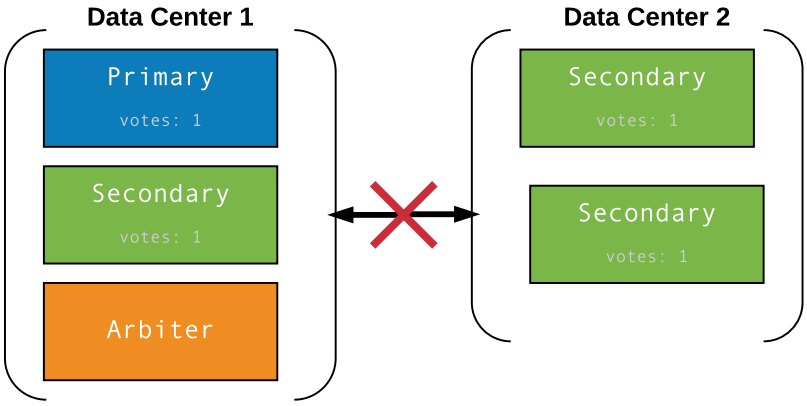
Note:

- We enter a state with no primary.
- Each side of the network partition has only 2 votes (not a majority).
- All the servers assume secondary status.
- This is avoidable.
- One solution is to add another member to the replica set.
- If another data node can not be provisioned, MongoDB has a special alternative called an arbiter that requires minimal resources.

- An arbiter is a `mongod` instance without data and performs only heartbeats, votes, and vetoes.

Scenario D: 5 Nodes in 2 DCs

The following is similar to Scenario C, but with the addition of an arbiter in Data Center 1. What happens here?

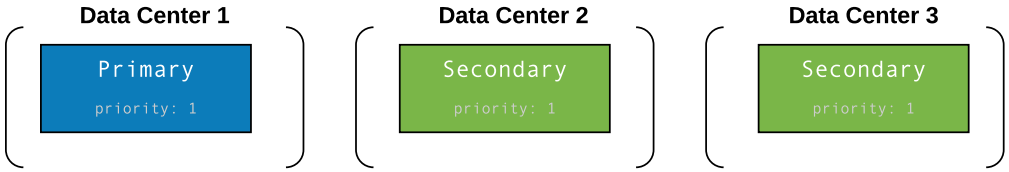


Note:

- The current primary is likely to remain primary.
- The arbiter helps ensure that the primary can reach a majority of the replica set.

Scenario E: 3 Data Nodes in 3 DCs

- What happens here if any one of the nodes/DCs fail?
- What about recovery time?

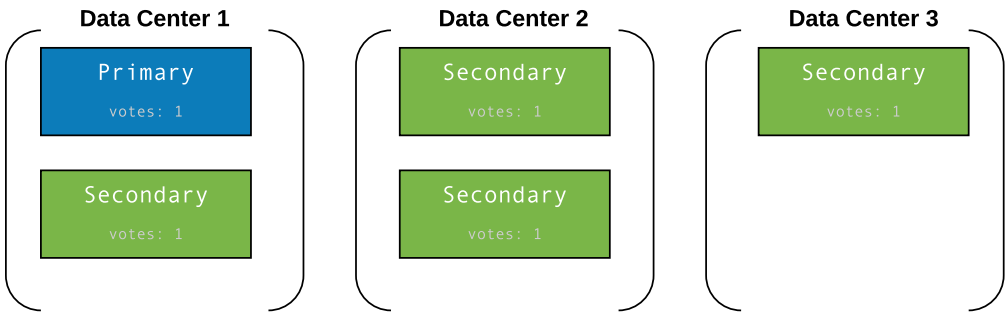


Note:

- The intent is to explain the advantage of deploying to 3 DCs - it's the minimum number of DCs in order for MongoDB to automatically failover if any one DC fails. This is generally what we recommend to customers in our consult and health check reports, though many continue to use 2 DCs due to costs and legacy reasons.
- To have automated failover in the event of single DC level failure, there must be at least 3 DCs. Otherwise the DC with the minority of nodes must be manually reconfigured.
- One of the data nodes can be replaced by an arbiter to reduce costs.

Scenario F: 5 Data Nodes in 3 DCs

What happens here if any one of the nodes/DCs fail? What about recovery time?



Note:

- Adds another data node to each "main" DC to reduce typically slow and costly cross DC network traffic if an initial sync or similar recovery is needed, as the recovering node can pull from a local replica instead.
- Depending on the data sizes, operational budget, and requirements, this can be overkill.
- The data node in DC3 can be replaced by an arbiter to reduce costs.

4.3 Replica Set Roles and Configuration

Learning Objectives

Upon completing this module students should understand:

- The use of priority to preference certain members or datacenters as primaries.
- Hidden members.
- The use of hidden secondaries for data analytics and other purposes (when secondary reads are used).
- The use of slaveDelay to protect against operator error.

Example: A Five-Member Replica Set Configuration

- For this example application, there are two datacenters.
- We name the hosts accordingly: `dc1-1`, `dc1-2`, `dc2-1`, etc.
  - This is just a clarifying convention for this example.
  - MongoDB does not care about host names except to establish connections.
- The nodes in this replica set have a variety of roles in this application.

Configuration

```
conf = {              // 5 data-bearing nodes
  _id: "mySet",
  members: [
    { _id : 0, host : "dc1-1.example.net:27017", priority : 5 },
    { _id : 1, host : "dc1-2.example.net:27017", priority : 5 },
    { _id : 2, host : "dc2-1.example.net:27017" },
    { _id : 3, host : "dc1-3.example.net:27017", hidden : true },
    { _id : 4, host : "dc2-2.example.net:27017", hidden : true,
      slaveDelay: 7200 }
  ]
}
```

Principal Data Center

```
{ _id : 0, host : "dc1-1.example.net", priority : 5 },
{ _id : 1, host : "dc1-2.example.net", priority : 5 },
```

Note:

- The objective with the priority settings for these two nodes is to prefer to DC1 for writes.
- The highest priority member that is up to date will be elected primary.
- Up to date means the member's copy of the oplog is within 10 seconds of the primary.
- If a member with higher priority than the primary is a secondary because it is is not up to date, but eventually catches up, it will force an election and win.

Data Center 2

```
{ _id : 2, host : "dc2-1.example.net:27017" },
```

Note:

- Priority is not specified, so it is at the default of 1.
- `dc2-1` could become primary, but only if both `dc1-1` and `dc1-2` are down.
- If there is a network partition and clients can only reach DC2, we can manually failover to `dc2-1`.

What about `dc1-3` and `dc2-2`?

```
// Both are hidden.
// Clients will not distribute reads to hidden members.
// We use hidden members for dedicated tasks.
{ _id : 3, host : "dc1-3.example.net:27017", hidden : true },
{ _id : 4, host : "dc2-2.example.net:27017", hidden : true,
  slaveDelay: 7200 }
```

Note:

- Will replicate writes normally.
- We would use this node to pull reports, run analytics, etc.
- We can do so without paying a performance penalty in the application for either reads or writes.

What about `dc2-2`?

```
{ _id : 4, host : "dc2-2.example.net:27017", hidden : true,
  slaveDelay : 7200 }
```

Note:

- `slaveDelay` permits us to specify a time delay (in seconds) for replication.
- In this case it is 7200 seconds or 2 hours.
- slaveDelay allows us to use a node as a short term protection against operator error:
  - Fat fingering – for example, accidentally dropping a collection in production.
  - Other examples include bugs in an application that result in corrupted data.
  - Not recommended. Use proper backups instead as there is no optimal delay value. E.g. 2 hours might be too long or too short depending on the situation.

4.4 The Oplog: Statement Based Replication

Learning Objectives

Upon completing this module students should understand:

- Binary vs. statement-based replication.
- How the oplog is used to support replication.
- How operations in MongoDB are translated into operations written to the oplog.
- Why oplog operations are idempotent.
- That the oplog is a capped collection and the implications this holds for syncing members.

Binary Replication

- MongoDB replication is statement based.
- Contrast that with binary replication.
- With binary replication we would keep track of:
  - The data files
  - The offsets
  - How many bytes were written for each change
- In short, we would keep track of actual bytes and very specific locations.
- We would simply replicate these changes across secondaries.

Tradeoffs

- The good thing is that figuring out where to write, etc. is very efficient.
- But we must have a byte-for-byte match of our data files on the primary and secondaries.
- The problem is that this couples our replica set members in ways that are inflexible.
- Binary replication may also replicate disk corruption.

Note:

- Some deployments might need to run different versions of MongoDB on different nodes.
- Different versions of MongoDB might write to different file offsets.
- We might need to run a compaction or repair on a secondary.
- In many cases we want to do these types of maintenance tasks independently of other nodes.

Statement-Based Replication

- Statement-based replication facilitates greater independence among members of a replica set.
- MongoDB stores a statement for every operation in a capped collection called the `oplog`.
- Secondaries do not simply apply exactly the operation that was issued on the primary.

Example

Suppose the following command is issued and it deletes 100 documents:

```
db.foo.deleteMany({ age : 30 })
```

This will be represented in the oplog with records such as the following:

```
{ "ts" : Timestamp(1407159845, 5), "h" : NumberLong("-70461248769192690 8"),
  "v" : 2, "op" : "d", "ns" : "bar.foo", "b" : true, "o" : { "_id" : 65 } }
{ "ts" : Timestamp(1407159845, 1), "h" : NumberLong("601412634522501 9794"),
  "v" : 2, "op" : "d", "ns" : "bar.foo", "b" : true, "o" : { "_id" : 333 } }
{ "ts" : Timestamp(1407159845, 4), "h" : NumberLong("817879176423846 5439"),
  "v" : 2, "op" : "d", "ns" : "bar.foo", "b" : true, "o" : { "_id" : 447 } }
{ "ts" : Timestamp(1407159845, 3), "h" : NumberLong("-170739100170552 8381"),
  "v" : 2, "op" : "d", "ns" : "bar.foo", "b" : true, "o" : { "_id" : 1033 } }
{ "ts" : Timestamp(1407159845, 2), "h" : NumberLong("-681429739244240 6598"),
  "v" : 2, "op" : "d", "ns" : "bar.foo", "b" : true, "o" : { "_id" : 9971 } }
```

Replication Based on the Oplog

- One statement per document affected by each write: insert, update, or delete.
- Provides a level of abstraction that enables independence among the members of a replica set:
  - With regard to MongoDB version.
  - In terms of how data is stored on disk.
  - Freedom to do maintenance without the need to bring the entire set down.

Note:

- Can do maintenance without bringing the set down because statement-based replication does not depend on all nodes running the same version of MongoDB or other restrictions that may be imposed by binary replication.
- In the next exercise, we will see that the oplog is designed so that each statement is idempotent.
- This feature has several benefits for independent operation of nodes in replica sets.

Operations in the Oplog are Idempotent

- Each operation in the oplog is idempotent.
- Whether applied once or multiple times it produces the same result.
- Necessary if you want to be able to copy data while simultaneously accepting writes.

Note: We need to be able to copy while accepting writes when:

- Doing an initial sync for a new replica set member.
- When a member rejoins a replica set after a network partition a member might end up writing operations it had already received prior to the partition.

The Oplog Window

- Oplogs are capped collections.
- Capped collections are fixed-size.
- They guarantee preservation of insertion order.
- They support high-throughput operations.
- Like circular buffers, once a collection fills its allocated space:
  - It makes room for new documents.
  - By overwriting the oldest documents in the collection.

Sizing the Oplog

- The oplog should be sized to account for latency among members.
- The default size oplog is usually sufficient.
- But you want to make sure that your oplog is large enough:
  - So that the oplog window is large enough to support replication
  - To give you a large enough history for any diagnostics you might wish to run.

4.5 Write Concern

Learning Objectives

Upon completing this module students should understand:

- How and when rollback occurs in MongoDB.
- The tradeoffs between durability and performance.
- Write concern as a means of ensuring durability in MongoDB.
- The different levels of write concern.

What happens to the write?

- A write is sent to a primary.
- The primary acknowledges the write to the client.
- The primary then becomes unavailable before a secondary can replicate the write

Answer

- Another member might be elected primary.
- It will not have the last write that occurred before the previous primary became unavailable.
- When the previous primary becomes available again:
  - It will note it has writes that were not replicated.
  - It will put these writes into a `rollback file`.
  - A human will need to determine what to do with this data.
- This is default behavior in MongoDB and can be controlled using `write concern`.

Balancing Durability with Performance

- The previous scenario is a specific instance of a common distributed systems problem.
- For some applications it might be acceptable for writes to be rolled back.
- Other applications may have varying requirements with regard to durability.
- Tunable write concern:
  - Make critical operations persist to an entire MongoDB deployment.
  - Specify replication to fewer nodes for less important operations.

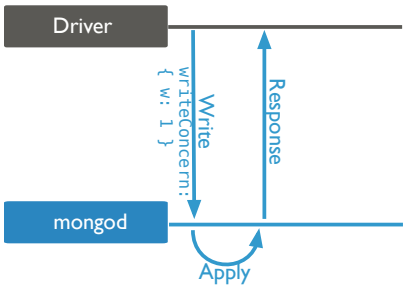
Note:

- MongoDB provides tunable write concern to better address the specific needs of applications.
- Clients may adjust write concern to ensure that the most important operations persist successfully to an entire MongoDB deployment.
- For other less critical operations, clients can adjust write concern to ensure faster performance.

Defining Write Concern

- MongoDB acknowledges its writes
- Write concern determines when that acknowledgment occurs
  - How many servers
  - Whether on disk or not
- Clients may define the write concern per write operation, if necessary.
- Standardize on specific levels of write concerns for different classes of writes.
- In the discussion that follows we will look at increasingly strict levels of write concern.

Write Concern: `{ w : 1 }`



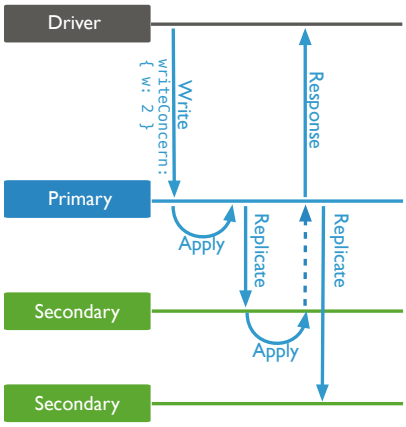
Note:

- We refer to this write concern as "Acknowledged".
- This is the default.
- The primary sends an acknowledgement back to the client that it received the write operation (in RAM).
- Allows clients to catch network, duplicate key, and other write errors.

Example: `{ w : 1 }`

```
db.edges.insertOne( { from : "tom185", to : "mary_p" },
                    { writeConcern : { w : 1 } } )
```

Write Concern: `{ w : 2 }`



Note:

- Called "Replica Acknowledged"
- Ensures the primary completed the write.
- Ensures at least one secondary replicated the write.

Example: `{ w : 2 }`

```
db.customer.updateOne( { user : "mary_p" },
                       { $push : { shoppingCart:
                         { _id : 335443, name : "Brew-a-cup",
                         price : 45.79 } } },
                       { writeConcern : { w : 2 } } )
```

Other Write Concerns

- `w` can use any integer for write concern.
- Acknowledgment guarantees the write has propagated to the specified number of voting members.
  - E.g., `{ w : 3 }`, `{ w : 4}`, etc.
- `j : true` ensures writes are also written to disk on the *primary* before being acknowledged
- When using PV1 (replication protocol version 1), writeConcernMajorityJournalDefault[7] is on by default for versions >= 3.4
  - so `w : majority` implies `j : true`

Write Concern: `{ w : "majority" }`

- Ensures the primary completed the write (in RAM).
  - By default, also on disk
- Ensures write operations have propagated to a majority of the **voting** members.
- Avoids hard coding assumptions about the size of your replica set into your application.
- Using majority trades off performance for durability.
- It is suitable for critical writes and to avoid rollbacks.

Example: `{ w : "majority" }`

```
db.products.updateOne({ _id : 335443 },
                       { $inc : { inStock : -1 } },
                       { writeConcern : { w : "majority" }})
```

[7] http://docs.mongodb.org/manual/reference/replica-configuration/#rsconf.writeConcernMajorityJournalDefault

Quiz: Which write concern?

Suppose you have a replica set with 7 data nodes. Your application has critical inserts for which you do not want rollbacks to happen. Secondaries may be taken down from to time for maintenance, leaving you with a potential 4 server replica set. Which write concern is best suited for these critical inserts?

- { w : 1 }
- { w : 2 }
- { w : 3 }
- { w : 4 }
- { w : "majority" }

Note: Answer: { w : "majority"}. This is the same as 4 for a 7 member replica set.

Further Reading

See Write Concern Reference[8] for more details on write concern configurations, including setting timeouts and identifying specific replica set members that must acknowledge writes (i.e. tag sets[9]).

[8] http://docs.mongodb.org/manual/reference/write-concern
[9] http://docs.mongodb.org/manual/tutorial/configure-replica-set-tag-sets/#replica-set-configuration-tag-sets

4.6 Read Preference

What is Read Preference?

- Read preference allows you to specify the nodes in a replica set to read from.
- Clients only read from the primary by default.
- There are some situations in which a client may want to read from:
  - Any secondary
  - A specific secondary
  - A specific type of secondary
- Only read from a secondary if you can tolerate possibly stale data, as not all writes might have replicated.

Use Cases

- Running systems operations without affecting the front-end application.
- Providing local reads for geographically distributed applications.
- Maintaining availability during a failover.

Note:

- If you have application servers in multiple data centers, you may consider having a geographically distributed replica set[10] and using a read preference of `nearest`.
- This allows the client to read from the lowest-latency members.
- Use `primaryPreferred` if you want an application to read from the primary under normal circumstances, but to allow possibly stale reads from secondaries during failover.

Not for Scaling

- In general, do *not* read from secondaries to provide extra capacity for reads.
- Sharding[11] increases read and write capacity by distributing operations across a group of machines.
- Sharding is a better strategy for adding capacity.

[10] http://docs.mongodb.org/manual/core/replica-set-geographical-distribution
[11] http://docs.mongodb.org/manual/sharding

Read Preference Modes

MongoDB drivers support the following read preferences. Note that `hidden` nodes will never be read from when connected via the replica set.

- **primary**: Default. All operations read from the primary.
- **primaryPreferred**: Read from the primary but if it is unavailable, read from secondary members.
- **secondary**: All operations read from the secondary members of the replica set.
- **secondaryPreferred**: Read from secondary members but if no secondaries are available, read from the primary.
- **nearest**: Read from member of the replica set with the least network latency, regardless of the member's type.

Tag Sets

- There is also the option to used tag sets.
- You may tag nodes such that queries that contain the tag will be routed to one of the servers with that tag.
- This can be useful for running reports, say for a particular data center or nodes with different hardware (e.g. hard disks vs SSDs).

For example, in the mongo shell:

```
conf = rs.conf()
conf.members[0].tags = { dc : "east", use : "production" }
conf.members[1].tags = { dc : "east", use : "reporting" }
conf.members[2].tags = { use : "production" }
rs.reconfig(conf)
```

4.7 Lab: Setting up a Replica Set

Overview

- In this exercise we will setup a 3 data node replica set on a single machine.
- In production, each node should be run on a dedicated host:
  - To avoid any potential resource contention
  - To provide isolation against server failure

Create Data Directories

Since we will be running all nodes on a single machine, make sure each has its own data directory.

On Linux or Mac OS, run the following in the terminal to create the 3 directories `~/data/rs1`, `~/data/rs2`, and `~/data/rs3`:

```
mkdir -p ~/data/rs{1,2,3}
```

On Windows, run the following command instead in Command Prompt or PowerShell:

```
md c:\data\rs1 c:\data\rs2 c:\data\rs3
```

Launch Each Member

Now start 3 instances of `mongod` in the foreground so that it is easier to observe and shutdown.

On Linux or Mac OS, run *each* of the following commands in its *own terminal window*:

```
mongod --replSet myReplSet --dbpath ~/data/rs1 --port 27017 --oplogSize 200
mongod --replSet myReplSet --dbpath ~/data/rs2 --port 27018 --oplogSize 200
mongod --replSet myReplSet --dbpath ~/data/rs3 --port 27019 --oplogSize 200
```

On Windows, run *each* of the following commands in its *own Command Prompt or PowerShell window*:

```
mongod --replSet myReplSet --dbpath c:\data\rs1 --port 27017 --oplogSize 200
mongod --replSet myReplSet --dbpath c:\data\rs2 --port 27018 --oplogSize 200
mongod --replSet myReplSet --dbpath c:\data\rs3 --port 27019 --oplogSize 200
```

Status

- At this point, we have 3 `mongod` instances running.
- They were all launched with the same `replSet` parameter of "myReplSet".
- Despite this, the members are not aware of each other yet.
- This is fine for now.

Note:

- In production, each member would run on a different machine and use service scripts. For example on Linux, modify `/etc/mongod.conf` accordingly and run:

  ```
  sudo service mongod start
  ```

- To simplify this exercise, we run all members on a single machine.
- The same configuration process is used for this deployment as for one that is distributed across multiple machines.

Connect to a MongoDB Instance

- Connect to the one of the MongoDB instances with the mongo shell.
- To do so run the following command in the terminal, Command Prompt, or PowerShell:

```
mongo    // connect to the default port 27017
```

Configure the Replica Set

```
rs.initiate()
// wait a few seconds
rs.add    ('<HOSTNAME>:27018')
rs.addArb('<HOSTNAME>:27019')

// Keep running rs.status() until there's a primary and 2 secondaries
rs.status()
```

Note:

- `rs.initiate()` will use the FQDN. If we simply use `localhost` when adding the data node and arbiter, MongoDB will refuse to mix the two and return an error.

Problems That May Occur When Initializing the Replica Set

- bindIp parameter is incorrectly set
- Replica set configuration may need to be explicitly specified to use a different hostname:

```
> conf = {
    _id: "<REPLICA-SET-NAME>",
    members: [
        { _id : 0, host : "<HOSTNAME>:27017"},
        { _id : 1, host : "<HOSTNAME>:27018"},
        { _id : 2, host : "<HOSTNAME>:27019",
          "arbiterOnly" : true},
      ]
    }
> rs.initiate(conf)
```

Write to the Primary

While still connected to the primary (port 27017) with mongo shell, insert a simple test document:

```
db.testcol.insert({ a: 1 })
db.testcol.count()

exit    // Or Ctrl-d
```

Read from a Secondary

Connect to one of the secondaries. E.g.:

```
mongo --port 27018
```

Read from the secondary

```
rs.slaveOk()
db.testcol.find()
```

Review the Oplog

```
use local
db.oplog.rs.find()
```

Changing Replica Set Configuration

To change the replica set configuration, first connect to the primary via mongo shell:

```
mongo --port <PRIMARY_PORT>   # e.g. 27017
```

Let's raise the priority of one of the secondaries. Assuming it is the 2nd node (e.g. on port 27018):

```
cfg = rs.conf()
cfg["members"][1]["priority"] = 10
rs.reconfig(cfg)
```

Note:

- Note that `cfg["members"][1]["priority"] = 10` does not actually change the priority.
- `rs.reconfig(cfg)` does.

Verifying Configuration Change

You will see errors like the following, which are expected:

```
2014-10-07T17:01:34.610+0100 DBClientCursor::init call() failed
2014-10-07T17:01:34.613+0100 trying reconnect to 127.0.0.1:27017 (127.0.0.1) failed
2014-10-07T17:01:34.617+0100 reconnect 127.0.0.1:27017 (127.0.0.1) ok
reconnected to server after rs command (which is normal)
```

Verify that the replica set configuration is now as expected:

```
rs.conf()
```

The secondary will now become a primary. Check by running:

```
rs.status()
```

Further Reading

- Replica Configuration[12]
- Replica States[13]

[12] http://docs.mongodb.org/manual/reference/replica-configuration/
[13] http://docs.mongodb.org/manual/reference/replica-states/

5 Sharding

5.1 Introduction to Sharding

Learning Objectives

Upon completing this module, students should understand:

- What problems sharding solves
- When sharding is appropriate
- The importance of the shard key and how to choose a good one
- Why sharding increases the need for redundancy

Contrast with Replication

- In an earlier module, we discussed Replication.
- This should never be confused with sharding.
- Replication is about high availability and durability.
  - Taking your data and constantly copying it
  - Being ready to have another machine step in to field requests.

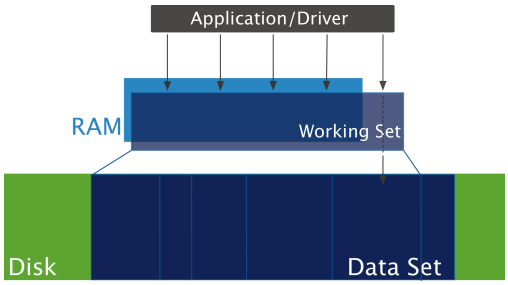
Sharding is Concerned with Scale

- What happens when a system is unable to handle the application load?
- It is time to consider scaling.
- There are 2 types of scaling we want to consider:
  - Vertical scaling
  - Horizontal scaling

Vertical Scaling

- Adding more RAM, faster disks, etc.
- When is this the solution?
- First, consider a concept called the `working set`.

The Working Set



Note:

- The working set for a MongoDB database is the portion of your data that clients access most often.
- Your working set should stay in memory, otherwise random disk operations will hurt performance.
- For the fastest processing, ensure that your indexes fit entirely in RAM so that the system can avoid reading the index from disk.
- In some cases, only recently indexed values must be in RAM.

Limitations of Vertical Scaling

- There is a limit to how much RAM one machine can support.
- There are other bottlenecks such as I/O, disk access and network.
- Cost may limit our ability to scale up.
- There may be requirements to have a large working set that no single machine could possible support.
- This is when it is time to scale horizontally.

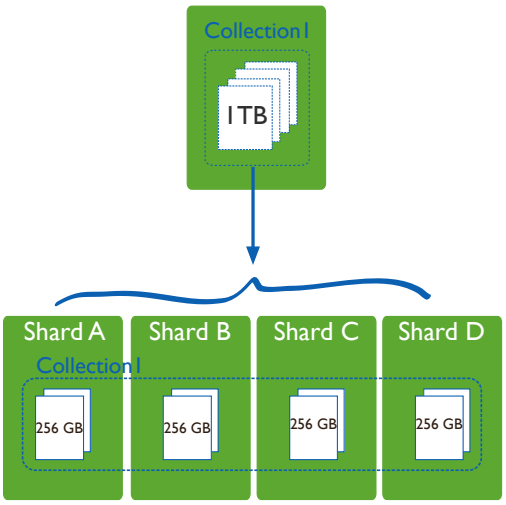
Sharding Overview

- MongoDB enables you to scale horizontally through sharding.
- Sharding is about adding more capacity to your system.
- MongoDB's sharding solution is designed to perform well on commodity hardware.
- The details of sharding are abstracted away from applications.
- Queries are performed the same way as if sending operations to a single server.
- Connections work the same by default.

When to Shard

- If you have more data than one machine can hold on its drives
- If your application is write heavy and you are experiencing too much latency.
- If your working set outgrows the memory you can allocate to a single machine.

Dividing Up Your Dataset



Note:

- When you shard a collection it is distributed across several servers.
- Each mongod manages a subset of the data.
- When you perform a read or write operation it will go to a router that will then direct the query to the appropriate server.
- Depending on the operation and how your cluster is configured you may need to touch only one server to complete the operation.

Sharding Concepts

To understanding how sharding works in MongoDB, we need to understand:

- Shard Keys
- Chunks

Shard Key

- You must define a shard key for a sharded collection.
- Based on one or more fields (like an index)
- Shard key defines a space of values
- Think of the key space like points on a line
- A key range is a segment of that line

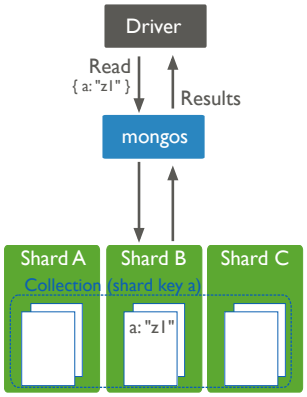
Shard Key Ranges

- A collection is partitioned based on shard key ranges.
- The shard key determines where documents are located in the cluster.
- It is used to route operations to the appropriate shard.
- For reads and writes
- Once a collection is sharded, you cannot change a shard key.

Note:

- For reads and updates, the shard key determines which shard holds the document and will handle the operation.
- When you insert a document, the shard key determines which server you will write to.

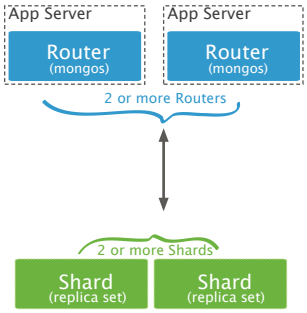
Targeted Query Using Shard Key



Chunks

- MongoDB partitions data into `chunks` based on shard key ranges.
- This is bookkeeping metadata.
- MongoDB attempts to keep the amount of data balanced across shards.
- This is achieved by migrating chunks from one shard to another as needed.
- There is nothing in a document that indicates its chunk.
- The document does not need to be updated if its assigned chunk changes.

Sharded Cluster Architecture



Note:

- This figure illustrates one possible architecture for a sharded cluster.
- Each shard is a self-contained replica set.
- Each replica set holds a partition of the data.
- As many new shards could be added to this sharded cluster as scale requires.
- At this point our cluster can handle a load approaching twice that of a single replica set using the same hardware.
- As mentioned, read/write operations go through a router.
- The server that routes requests is the mongos.

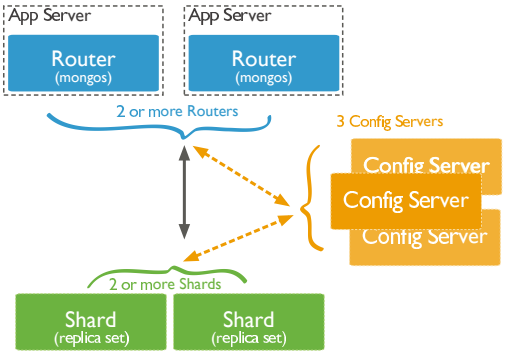
Mongos

- A mongos is responsible for accepting requests and returning results to an application driver.
- In a sharded cluster, nearly all operations go through a mongos.
- A sharded cluster can have as many mongos routers as required.
- It is typical for each application server to have one mongos.
- Always use more than one mongos to avoid a single point of failure.

Note:

- A mongos is typically deployed on an application server.
- There should be one mongos per app server.
- Scale with your app server.
- Very little latency between the application and the router.

Config Servers



Note:

- The previous diagram was incomplete; it was missing config servers.
- Use three config servers in production.
- In MongoDB 3.2, support for config server replica set (CSRS) was introduced.
- In MongoDB 3.4, mirrored (SCCC) config servers were deprecated. Config servers must now be set up in CSRS.
- In MongoDB 3.4, the balancing process was moved from the `mongos` server and is the responsibility of the **primary** config server
- These hold only metadata about the sharded collections.
  - Where your mongos servers are
  - Any hosts that are not currently available
  - What collections you have
  - How your collections are partitioned across the cluster
- Mongos processes use them to retrieve the state of the cluster.

- You can access cluster metadata from a mongos by looking at the `config` db.

Config Server Hardware Requirements

- Quality network interfaces
- A small amount of disk space (typically a few GB)
- A small amount of RAM (typically a few GB)
- The larger the sharded cluster, the greater the config server hardware requirements.

Possible Imbalance?

- Depending on how you configure sharding, data can become unbalanced on your sharded cluster.
  - Some shards might receive more inserts than others.
  - Some shards might have documents that grow more than those in other shards.
- This may result in too much load on a single shard.
  - Reads and writes
  - Disk activity
- This would defeat the purpose of sharding.

Balancing Shards

- If a chunk grows too large MongoDB will split it into two chunks.
- The MongoDB balancer keeps chunks distributed across shards in equal numbers.
- However, a balanced sharded cluster depends on a good shard key.

With a Good Shard Key

You might easily see that:

- Reads hit only 1 or 2 shards per query.
- Writes are distributed across all servers.
- Your disk usage is evenly distributed across shards.
- Things stay this way as you scale.

With a Bad Shard Key

You might see that:

- Your reads hit every shard.
- Your writes are concentrated on one shard.
- Most of your data is on just a few shards.
- Adding more shards to the cluster will not help.

Choosing a Shard Key

Generally, you want a shard key:

- That has high cardinality
- That is used in the majority of read queries
- For which the values read and write operations use are randomly distributed
- For which the majority of reads are routed to a particular server

More Specifically

- Your shard key should be consistent with your query patterns.
- If reads usually find only one document, you only need good cardinality.
- If reads retrieve many documents:
  - Your shard key supports locality
  - Matching documents will reside on the same shard

Cardinality

- A good shard key will have high cardinality.
- A relatively small number of documents should have the same shard key.
- Otherwise operations become isolated to the same server.
- Because documents with the same shard key reside on the same shard.
- Adding more servers will not help.
- Hashing will not help.

Non-Monotonic

- A good shard key will generate new values non-monotonically.
- Datetimes, counters, and ObjectIds make bad shard keys.
- Monotonic shard keys cause all inserts to happen on the same shard.
- Hashing will solve this problem.
- However, doing range queries with a hashed shard key will perform a scatter-gather query across the cluster.

Note:

- Documents will eventually move as chunks are balanced.
- But in the meantime one server gets hammered while others are idle.
- And moving chunks has its own performance costs.

Shards Should be Replica Sets

- As the number of shards increases, the number of servers in your deployment increases.
- This increases the probability that one server will fail on any given day.
- With redundancy built into each shard you can mitigate this risk.

5.2 Balancing Shards

Learning Objectives

Upon completing this module students should understand:

- Chunks and the balancer
- The status of chunks in a newly sharded collection
- How chunk splits automatically occur
- Advantages of pre-splitting chunks
- How the balancer works

Chunks and the Balancer

- Chunks are groups of documents.
- The shard key determines which chunk a document will be contained in.
- Chunks can be split when they grow too large.
- The balancer decides where chunks go.
- It handles migrations of chunks from one server to another.

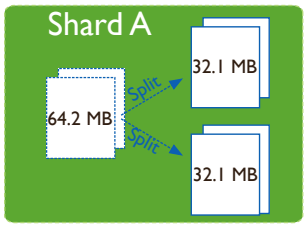
Chunks in a Newly Sharded Collection

- The range of a chunk is defined by the shard key values of the documents the chunk contains.
- When a collection is sharded it starts with just one chunk.
- The first chunk for a collection will have the range:

```
{ $minKey : 1 } to { $maxKey : 1 }
```

- All shard key values from the smallest possible to the largest fall in this chunk's range.

Chunk Splits



Note:

- When a chunk grows larger than the chunk size it will be split in half.
- The default chunk size is 64MB.
- A chunk can only be split between two values of a shard key.
- If every document on a chunk has the same shard key value, it cannot be split.
- This is why the shard key's cardinality is important
- Chunk splitting is just a bookkeeping entry in the metadata.
- No data bearing documents are altered.

Pre-Splitting Chunks

- You may pre-split data before loading data into a sharded cluster.
- Pre-splitting is useful if:
  - You plan to do a large data import early on
  - You expect a heavy initial server load and want to ensure writes are distributed

---

Note:

- A large data import will take time to split and balance without pre-splitting.

Start of a Balancing Round

- A balancing round is initiated by the balancer process on the primary config server.
- This happens when the difference in the number of chunks between two shards becomes to large.
- Specifically, the difference between the shard with the most chunks and the shard with the fewest.
- A balancing round starts when the imbalance reaches:
  - 2 when the cluster has < 20 chunks
  - 4 when the cluster has 20-79 chunks
  - 8 when the cluster has 80+ chunks

Balancing is Resource Intensive

- Chunk migration requires copying all the data in the chunk from one shard to another.
- Each individual shard can be involved in one migration at a time. Parallel migrations can occur for each shard migration pair (source + destination).
- The amount of possible parallel chunk migrations for *n* shards is *n*/2 rounded down.
- MongoDB creates splits only after an insert operation.
- For these reasons, it is possible to define a balancing window to ensure the balancer will only run during scheduled times.

Note: As of MongoDB 3.4, deployments using WiredTiger will see increased balancing speed. The default value **secondaryThrottle** is **false**, so the balancer will not wait for replication before proceeding to the next chunk during migration. This is the default now because MongoDB now has a dedicated balancing migration thread

---

Chunk Migration Steps

1. The balancer process sends the `moveChunk` command to the source shard.
2. The source shard continues to process reads/writes for that chunk during the migration.
3. The destination shard requests documents in the chunk and begins receiving copies.
4. After receiving all documents, the destination shard receives any changes to the chunk.
5. Then the destination shard tells the config db that it has the chunk.
6. The destination shard will now handle all reads/writes.
7. The source shard deletes its copy of the chunk.

Concluding a Balancing Round

- Each chunk will move:
  - From the shard with the most chunks
  - To the shard with the fewest
- A balancing round ends when all shards differ by at most one chunk.

5.3 Shard Zones

Learning Objectives

Upon completing this module students should understand:

- The purpose for shard zones
- Advantages of using shard zones
- Potential drawbacks of shard zones

Note: MongoDB 3.4 introduced Zones, which supersedes tag-aware sharding available in earlier versions.

Zones - Overview

- Shard zones allow you to "tie" data to one or more shards.
- A shard zone describes a range of shard key values.
- If a chunk is in the shard tag range, it will live on a shard with that tag.
- Shard tag ranges cannot overlap. In the case we try to define overlapping ranges an error will occur during creation.

Example: DateTime

- Documents older than one year need to be kept, but are rarely used.
- You set a part of the shard key as the ISODate of document creation.
- Add shards to the LTS zone.
- These shards can be on cheaper, slower machines.
- Invest in high-performance servers for more frequently accessed data.

Example: Location

- You are required to keep certain data in its home country.
- You include the country in the shard tag.
- Maintain data centers within each country that house the appropriate shards.
- Meets the country requirement but allows all servers to be part of the same system.
- As documents age and pass into a new zone range, the balancer will migrate them automatically.

Example: Premium Tier

- You have customers who want to pay for a "premium" tier.
- The shard key permits you to distinguish one customer's documents from all others.
- Tag the document ranges for each customer so that their documents will be located on shards of the appropriate tier (zone).
- Shards tagged as premium tier run on high performance servers.
- Other shards run on commodity hardware.
- See Manage Shard Zone[14]

Note:

- As customers move from one tier to another it will be necessary to execute commands that either add a given customer's shard key range to the premium tag or remove that range from those tagged as "premium".
- During balancing rounds, if the balancer detects that any chunks are not on the correct shards per configured tags, the balancer migrates chunks in tagged ranges to shards associated with those tags.
- After re-configuring tags with a shard key range, and associating it with a shard or shards, the cluster may take some time to balance the data among the shards.
- See: Tiered Hardware for varying SLA or SLO[15].

---

[14] http://docs.mongodb.org/manual/tutorial/manage-shard-zone/
[15] https://docs.mongodb.com/manual/tutorial/sharding-tiered-hardware-for-varying-slas/

Zones - Caveats

- Because tagged chunks will only be on certain servers, if you tag more than those servers can handle, you'll have a problem.
  - You're not only worrying about your overall server load, you're worrying about server load for each of your tags.
- Your chunks will evenly distribute themselves across the available zones. You cannot control things more fine grained than your tags.

6 Security

Security Introduction **(page 112)** An introduction to security options in MongoDB

Authorization **(page 115)** Authorization in MongoDB

Authentication **(page 123)** Authentication in MongoDB

Auditing **(page 125)** Auditing in MongoDB

Encryption **(page 127)** Encryption at rest in MongoDB

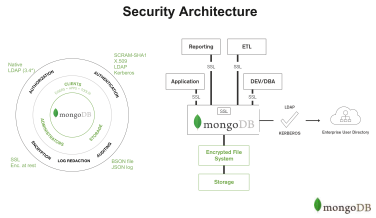
6.1 Security Introduction

Learning Objectives

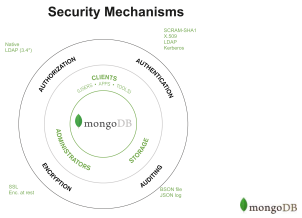
Upon completing this module students should understand:

- The high-level overview of security in MongoDB
- Security options for MongoDB
  - Authentication
  - Authorization
  - Transport Encryption
  - Enterprise only features

A High Level Overview



Security Mechanisms



Note:

- MongoDB provides numerous security features, to include:
  - Authentication

        * SCRAM-SHA-1
        * x.509 Certificate Authentication
    - Authorization
        * Role-Based Access Control
    - Transport Encryption
    - Enterprise Only Features
        * Kerberos Authentication
        * LDAP Proxy Authentication
        * Encryption at Rest
        * Auditing
    - Network Exposure Settings
- You should only run MongoDB in a trusted environment.
- You should run MongoDB from a non-root user.
- You are welcome to use any features you desire, or none.
- All security is off by default. This will change, and versions >= 3.6 will be restricted to localhost by default

---

Authentication Options

- Community
    - Challenge/response authentication using SCRAM-SHA-1 (username & password)
    - X.509 Authentication (using X.509 Certificates)
- Enterprise
    - Kerberos
    - LDAP

Note:

- Although there is a SCRAM-SHA-2 algorithm that addressed some vulnerabilities in SCRAM-SHA-1, it would not benefit MongoDB
- By cracking either algorithm an attacker would have to have access to the `db.users` collection and associated metadata

---

Authorization via MongoDB

- Predefined roles
- Custom roles
- LDAP authorization (MongoDB Enterprise)
  - Query LDAP server for groups to which a user belongs.
  - Distinguished names (DN) are mapped to roles on the `admin` database.
  - Requires external authentication (x.509, LDAP, or Kerberos).

Transport Encryption

- TLS/SSL
  - May use certificates signed by a certificate authority or self-signed.
- FIPS (MongoDB Enterprise)

Note:

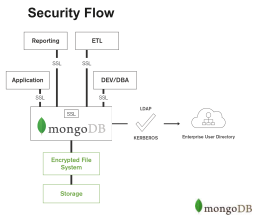
- FIPS (Federal Information Processing Standard)
- Users should use a certificate from a certificate authority in order for clients to verify the server's identity

---

Network Exposure Options

- `bindIp` limits the ip addresses the server listens on.
- Using a non-standard port can provide a layer of obscurity.
- MongoDB should still be run only in a trusted environment.

Security Flow



Note:

- A robust and secure applicate architecture can be created by using the provided security features and following this checklist[16]
- The overall strength of the security protocols in place is only as strong as the weakest link.
  - Use authentication

[16] https://docs.mongodb.com/manual/administration/security-checklist/

  - Use role-based authorization; limit users to what they need
  - Transmit data using TLS/SSL
  - Encrypt data at rest
  - Use valid certificates signed by a trusted certificate authority
  - Ensure all `mongod` and `mongos` servers are configured properly

6.2 Authorization

Learning Objectives

Upon completing this module, students should be able to:

- Outline MongoDB's authorization model
- List authorization resources
- Describe actions users can take in relation to resources
- Create roles
- Create privileges
- Outline MongoDB built-in roles
- Grant roles to users
- Explain LDAP authorization

Authorization vs Authentication

Authorization and Authentication are generally confused and misinterpreted concepts:

- Authorization defines the rules by which users can interact with a given system:
  - Which operations can they perform
  - Over which resources
- Authentication is the mechanism by which users identify and are granted access to a system:
  - Validation of credentials and identities
  - Controls access to the system and operational interfaces

Authorization Basics

- MongoDB enforces a role-based authorization model.
- A user is granted roles that determine the user's access to database resources and operations.

The model determines:

- Which roles are granted to users
- Which privileges are associated with roles
- Which actions can be performed over different resources

Note:

- You can bring up the following questions:
  - What are privileges?
  - What kind of resources can be found on a typical database?
- Have some open discussion about what defines an action.
- Also you can take the opportunity to give examples of different roles in a company and how they are organized in terms of procedures and resources.

What is a resource?

- Databases?
- Collections?
- Documents?
- Users?
- Nodes?
- Shard?
- Replica Set?

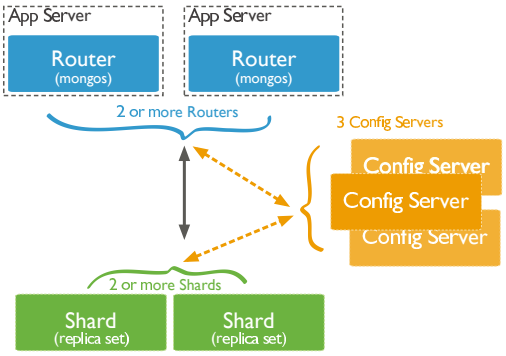
Note:

- A resource is a database, collection, set of collections, or the cluster.
- If the resource is the cluster, the affiliated actions affect the state of the system rather than a specific database or collection.

Authorization Resources

- Databases
- Collections
- Cluster

Cluster Resources



Note:

- Given the distributed nature of our database, MongoDB includes the cluster resource in the authorization module.
- Replica sets and shards comprise the cluster domain.

Types of Actions

Given a resource, we can consider the available actions:

- Query and write actions
- Database management actions
- Deployment management actions
- Replication actions
- Sharding actions
- Server administration actions
- Diagnostic actions
- Internal actions

Note:

- Actions are the operations that one can perform on database resources.
- The actions above are grouped by purpose.
- This organization is logical, not operational.

- Here we can ask the students which common operations they are familiar with while operating with a database and how those translate to MongoDB operations.

Specific Actions of Each Type

Query / Write	Database Mgmt	Deployment Mgmt
find	enableProfiler	planCacheRead
insert	createIndex	storageDetails
remove	createCollection	authSchemaUpgrade
update	changeOwnPassword	killop
	...	...

See the complete list of actions[17] in the MongoDB documentation.

Note: These are just a few examples of the list of actions available. The full list is available in MongoDB docs: https://docs.mongodb.org/v3.0/reference/privilege-actions/#privilege-actions

---

Authorization Privileges

A privilege defines a pairing between a resource as a set of permitted actions.

Resource:

```
{"db": "yourdb", "collection": "mycollection"}
```

Action: `find`

Privilege:

```
{
  resource: {"db": "yourdb", "collection": "mycollection"},
  actions: ["find"]
}
```

Note:

- We want to explain that we can set a privilege that enables multiple actions on a given resource.
- Also important to highlight that we can set *loose* resources like `all databases` or `all collections`

```
{
  resource: {"db": "", "collection": ""},
  actions: ["find", "insert"]
}
```

---

[17] https://docs.mongodb.com/manual/reference/privilege-actions/

Authorization Roles

MongoDB grants access to data through a role-based authorization system:

- Built-in roles: pre-canned roles that cover the most common sets of privileges users may require
- User-defined roles: if there is a specific set of privileges not covered by the existing built-in roles you are able to create your own roles

Built-in Roles

Database Admin	Cluster Admin	All Databases
dbAdmin	clusterAdmin	readAnyDatabase
dbOwner	clusterManager	readWriteAnyDatabase
userAdmin	clusterMonitor	userAdminAnyDatabase
	hostManager	dbAdminAnyDatabase

Database User	Backup & Restore
read	backup
readWrite	restore

Superuser	Internal
root	__system

Note: Built-in roles have been created given the generic users that interact with a database and their respective tasks.

- Database user roles: should be granted to application-side users;
- Database administrators: roles conceived for system administrator, DBAs and security officers
- Cluster Administrator roles: mostly for system administrators and DBAs; individuals that will deal with the overall administration of deployments
- Backup and Restore: for applications that perform only backup and restore operations (for example: Cloud and Ops Manager)
- All Database Roles: for global administrators of a deployment. If you want to avoid granting the same role for every single database
- Superuser: root level operations. Generally the first user that you create on any give system should probably have a root role and then add other specific users.
- Internal: it's documented, it's public but don't mention it too much. This a backdoor that only the cluster members (other replica set members, or a mongos) should have access to. Do not assign this role to user objects representing applications or human administrators.

Built-in Roles

To grant roles while creating an user:

```
use admin
db.createUser(
  {
    user: "myUser",
    pwd: "$up3r$3cr7",
    roles: [
      {role: "readAnyDatabase", db: "admin"},
      {role: "dbOwner", db: "superdb"},
      {role: "readWrite", db: "yourdb"}
    ]
  }
)
```

Built-in Roles

To grant roles to existing user:

```
use admin
db.grantRolesToUser(
  "reportsUser",
  [
    { role: "read", db: "accounts" }
  ]
)
```

Note: `grantRolesToUser` also allows to specify a `writeConcern` to ensure the durability of the operation, as any of the remaining authz methods.

```
db.grantRolesToUser(
  "reportsUser",
  [
    { role: "read", db: "accounts" }
  ],
  { w: "majority" , wtimeout: 4000 }
)
```

User-defined Roles

- If no suitable built-in role exists, we can can create a role.
- Define:
  - Role name
  - Set of privileges
  - List of inherit roles (optional)

```
use admin
db.createRole({
```

```
  role: "insertAndFindOnlyMyDB",
  privileges: [
    {resource: { db: "myDB", collection: "" }, actions: ["insert", "find"]}
  ],
  roles: []})
```

Role Privileges

To check the privileges of any particular role we can get that information using the `getRole` method:

```
db.getRole("insertAndFindOnlyMyDB", {showPrivileges: true})
```

Note: There are many other authorization and user management commands and options that you should get your students acquainted with. All of those can be found in the security reference[18]

The output of this slide command is should be similar to the following:

```
{
"role": "insertAndFindOnlyMyDB",
"db": "admin",
"isBuiltin": false,
"roles": [ ],
"inheritedRoles": [ ],
"privileges": [
  {
    "resource": {
      "db": "myDB",
      "collection": ""
    },
    "actions": [
      "find",
      "insert"
    ]
  }
],
"inheritedPrivileges": [
  {
    "resource": {
      "db": "myDB",
      "collection": ""
    },
    "actions": [
      "find",
      "insert"
    ]
  }
 ]
}
```

---

[18] https://docs.mongodb.org/manual/reference/security/

LDAP Authorization

As of MongoDB 3.4, MongoDB supports *authorization* with LDAP.

How it works:

1. User authenticates via an external mechanism

```
$ mongo --username alice \
        --password secret \
        --authenticationMechanism PLAIN \
        --authenticationDatabase '$external'
```

LDAP Authorization (cont'd)

2. Username is tranformed into LDAP query

```
[
  {
    match: "(.+)@ENGINEERING",
    substitution: "cn={0},ou=engineering,dc=example,dc=com"
  }, {
    match: "(.+)@DBA",
    substitution:"cn={0},ou=dba,dc=example,dc=com"
  }
]
```

LDAP Authorization (cont'd)

3. MongoDB queries the LDAP server
    - A single entity's attributes are treated as the user's roles
    - Multiple entitiy's distinguished names are treated as the user's roles

Mongoldap

`mongoldap` can be used to test configurations between MongoDB and an LDAP server

```
$ mongoldap -f mongod.conf \
            --user "uid=alice,ou=Users,dc=example,dc=com" \
            --password secret
```

6.3 Authentication

Learning Objectives

Upon completing this module, you should understand:

- Authentication mechanisms
- External authentication
- Native authentication
- Internal node authentication
- Configuration of authentication mechanisms

Authentication

- Authentication is concerned with:
  - Validating identities
  - Managing certificates / credentials
  - Allowing accounts to connect and perform authorized operations
- MongoDB provides native authentication and supports X509 certificates, LDAP, and Kerberos as well.

Authentication Mechanisms

MongoDB supports a number of authentication mechanisms:

- SCRAM-SHA-1 (default >= 3.0)
- MONGODB-CR (legacy)
- X509 Certificates
- LDAP (MongoDB Enterprise)
- Kerberos (MongoDB Enterprise)

Note:

- Native: SCRAM-SHA-1 and MongoDB-CR are native mechanisms in the sense that they are fully managed by MongoDB instances.
- External: LDAP and Kerberos are external authentication mechanisms and are only available with MongoDB Enterprise.
- X509 can also be considered native in terms of management but they rely on certificates generated by 3rd parties and only enforced by MongoDB.

Internal Authentication

For internal authentication purposes (mechanism used by replica sets and sharded clusters) MongoDB relies on:

- Keyfiles
  - Shared password file used by replica set members
  - Hexadecimal value of 6 to 1024 chars length
- X509 Certificates

Simple Authentication Configuration

To get started we just need to make sure we are launching our mongod instances with the `--auth` parameter.

```
mongod --dbpath /data/db --auth
```

For any connections to be established to this mongod instance, the system will require a username and password.

```
mongo -u user -p
↪
↪                                                    MongoDB shell version: 3.2.
↪5
Enter password:
```

Note:

- Using the `--auth` parameter will only cause mongod to enable authentication.
- You need to create users separately.
- the shell
- You can take the opportunity to ask:
  - Q: What happens if we just launch a mongod without having any users created?
  - A: Nothing happens, we just can't access the instance.
- cover localhost exception
- create a first user with the following, avoid using `root` as the role

```
db.createUser({user:'admin',pwd:'pwd',roles:[{role:'userAdminAnyDatabase',db:'admin'}
↪]})
```

- run the following to see who is authenticated:

```
db.runCommand({connectionStatus:1})
```

6.4 Auditing

Learning Objectives

Upon completing this module, you should be able to:

- Outline the auditing capabilities of MongoDB
- Enable auditing
- Summarize auditing configuration options

Auditing

- MongoDB Enterprise includes an auditing capability for mongod and mongos instances.
- The auditing facility allows administrators and users to track system activity
- Important for deployments with multiple users and applications.

Audit Events

Once enabled, the auditing system can record the following operations:

- Schema
- Replica set and sharded cluster
- Authentication and authorization
- CRUD operations (DML, off by default)

Auditing Configuration

The following are command-line parameters to `mongod`/`mongos` used to configure auditing.

Enable auditing with `--auditDestination`.

- `--auditDestination`: where to write the audit log
  - syslog
  - console
  - file
- `--auditPath`: audit log path in case we define "file" as the destination

Auditing Configuration (cont'd)

- `--auditFormat`: the output format of the emitted event messages
  - BSON
  - JSON
- `--auditFilter`: an expression that will filter the types of events the system records

By default we only audit DDL operations but we can also enable DML (requires auditAuthorizationSuccess set to true)

Note:

- Explain what DML and DDL operations are:
  - DML means data manipulation language (inserts, updates, removes, grant user role ...)
  - DDL means data definition language (create collection, index, drop database ...)
- Q: Why do we not enable DML by default?
- A: Due to the performance impact of logging all write operations
- Q: In what circumstances might we want to enable DML?
- A: On highly sensitive namespaces or for given set of users.

Auditing Message

The audit facility will launch a message every time an auditable event occurs:

```
{
  atype: <String>,
  ts : { "$date": <timestamp> },
  local: { ip: <String>, port: <int> },
  remote: { ip: <String>, port: <int> },
  users : [ { user: <String>, db: <String> }, ... ],
  roles: [ { role: <String>, db: <String> }, ... ],
  param: <document>,
  result: <int>
}
```

Auditing Configuration

If we want to configure our audit system to generate a *JSON* file we would need express the following command:

```
mongod --auditDestination file --auditPath /some/dir/audit.log --auditFormat JSON
```

If we want to capture events from a particular user *myUser*:

```
mongod --auditDestination syslog --auditFilter '{"users.user": "myUser"}'
```

To enable DML we need to set a specific parameter:

```
mongod --auditDestination console --setParameter auditAuthorizationSuccess=true
```

Note:

- We can define filters on any particular field of the audit message
- These will work as regular MongoDB query filter expressions, but not all operators will apply.
- Be creative and ask students to set different filters based on roles or incoming connections.

6.5 Encryption

Learning Objectives

Upon completing this module, students should understand:

- The encryption capabilities of MongoDB
- Network encryption
- Native encryption
- Third party integrations

Encryption

MongoDB offers two levels of encryption

- Transport layer
- Encryption at rest (MongoDB Enterprise >=3.2)

Note:

- important to note to students that encryption at rest is an enterprise version feature

---

Network Encryption

- MongoDB enables TLS/SSL for transport layer encryption of traffic between nodes in a cluster.
- Three different network architecture options are available:
  - Encryption of application traffic connections
  - Full encryption of all connections
  - Mixed encryption between nodes

Note:

- mixed encryption means that we can have nodes in a replica set that communicate with some nodes not encrypted and others encrypted

---

Native Encryption

MongoDB Enterprise comes with a encrypted storage engine.

- Native encryption supported by WiredTiger
- Encrypts data at rest
  - AES256-CBC: 256-bit Advanced Encryption Standard in Cipher Block Chaining mode (default)
    - symmetric key (same key to encrypt and decrypt)
  - AES256-GCM: 256-bit Advanced Encryption Standard in Galois/Counter Mode
  - FIPS is also available
- Enables integration with key management tools

Encryption and Replication

- Encryption is not part of replication:
  - Data is not natively encrypted on the wire
    - Requires transport encryption to ensure secured transmission
  - Encryption keys are not replicated
    - Each node should have their own individual keys

Note:

- Important to raise awareness to this point
- Many students might get the impression that configuring encryption in one of the nodes would be enough when that's not the case
- Wire data needs to encrypted through TLS/SSL configuration
- Encrypted Storage Engine only provides encryption on data at rest
- We should use different encryption keys for different nodes.

Third Party Integration

- Key Management Interoperability Protocol (KMIP)
  - Integrates with Vormetric Data Security Manager (DSM) and SafeNet KeySecure
- Storage Encryption
  - Linux Unified Key Setup (LUKS)
  - IBM Guardium Data Encryption
  - Vormetric Data Security Platform
    - Also enables Application Level Encryption on per-field or per-document
  - Bitlocker Drive Encryption

Note:

- MongoDB offers some integration options for Key Management and Storage Encryption
- Key managers are recommended for good security practices like key expiration and rotation
- Key managers are important if we want to be complaint with HIPAA, PCI-DSS, and FERPA certifications

---

7 Reporting Tools and Diagnostics

Performance Troubleshooting **(page 130)** An introduction to reporting and diagnostic tools for MongoDB

7.1 Performance Troubleshooting

Learning Objectives

Upon completing this module students should understand basic performance troubleshooting techniques and tools including:

- `mongostat`
- `mongotop`
- `db.setProfilingLevel()`
- `db.currentOp()`
- `db.<COLLECTION>.stats()`
- `db.serverStatus()`

`mongostat` and `mongotop`

- `mongostat` samples a server every second.
  - See current ops, pagefaults, network traffic, etc.
  - Does not give a view into historic performance; use Ops Manager for that.
- `mongotop` looks at the time spent on reads/writes in each collection.

Exercise: `mongostat` (setup)

In one window, perform the following commands.

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    arr = [];
    for (j=1; j<=1000; j++) {
        doc = { _id: (1000 * (i-1) + j), a: i, b: j, c: (1000 * (i-1)+ j) };
        arr.push(doc)
    };
    db.testcol.insertMany(arr);
    var x = db.testcol.find( { b : 255 } );
    x.next();
    var x = db.testcol.find( { _id : 1000 * (i-1) + 255 } );
    x.next();
    var x = "asdf";
    db.testcol.updateOne( { a : i, b : 255 }, { $set : { d : x.pad(1000) } });
    print(i)
}
```

Exercise: `mongostat` (run)

- In another window/tab, run mongostat.
- You will see:
  - Inserts
  - Queries
  - Updates

Exercise: `mongostat` (create index)

- In a third window, create an index when you see things slowing down:

```
db.testcol.createIndex( { a : 1, b : 1 } )
```

- Look at mongostat.
- Notice that things are going significantly faster.
- Then, let's drop that and build another index.

```
db.testcol.dropIndexes()
db.testcol.createIndex( { b : 1, a : 1 } )
```

Exercise: `mongotop`

Perform the following then, in another window, run `mongotop`.

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    arr = [];
    for (j=1; j<=1000; j++) {
        doc = {_id: (1000*(i-1)+j), a: i, b: j, c: (1000*(i-1)+j)};
        arr.push(doc)
    };
    db.testcol.insertMany(arr);
    var x = db.testcol.find( {b: 255} ); x.next();
    var x = db.testcol.find( {_id: 1000*(i-1)+255} ); x.next();
    var x = "asdf";
    db.testcol.updateOne( {a: i, b: 255}, {$set: {d: x.pad(1000)}});
    print(i)
}
```

Note: Direct the students to the fact that you can see the activity on the server for reads/writes/total.

`db.currentOp()`

- currentOp is a tool that asks what the db is doing at the moment.
- currentOp is useful for finding long-running processes.
- Fields of interest:
  - microsecs_running
  - op
  - query
  - lock
  - waitingForLock

Exercise: `db.currentOp()`

Do the following then, connect with a separate shell, and repeatedly run `db.currentOp()`.

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    arr = [];
    for (j=1; j<=1000; j++) {
        doc = {_id: (1000*(i-1)+j), a: i, b: j, c: (1000*(i-1)+j)};
        arr.push(doc)
    };
    db.testcol.insertMany(arr);
    var x = db.testcol.find( {b: 255} ); x.next();
    var x = db.testcol.find( {_id: 1000*(i-1)+255 }); x.next();
    var x = "asdf";
    db.testcol.updateOne( {a: i, b: 255}, {$set: {d: x.pad(1000)}});
    print(i)
}
```

Note: Point out to students that the running time gets longer & longer, on average.

`db.<COLLECTION>.stats()`

- Used to view the current stats for a collection.
- Everything is in bytes; use the multiplier parameter to view in KB, MB, etc
- You can also use `db.stats()` to do this at scope of the entire database

Exercise: Using Collection Stats

Look at the output of the following:

```
db.testcol.drop()
db.testcol.insertOne( { a : 1 } )
db.testcol.stats()
var x = "asdf"
db.testcol2.insertOne( { a : x.pad(10000000) } )
db.testcol2.stats()
db.stats()
```

The Profiler

- Off by default.
- To reset, `db.setProfilingLevel(0)`
- At setting 1, it captures "slow" queries.
- You may define what "slow" is.
- Default is 100ms: `db.setProfilingLevel(1)`
- E.g., to capture 20 ms: `db.setProfilingLevel(1, 20)`

The Profiler (continued)

- If the profiler level is 2, it captures all queries.
  - This will severely impact performance.
  - Turns all reads into writes.
- Always turn the profiler off when done (set level to 0)
- Creates `db.system.profile` collection

Exercise: Exploring the Profiler

Perform the following, then look in your db.system.profile.

```
db.setProfilingLevel(0)
db.testcol.drop()
db.system.profile.drop()
db.setProfilingLevel(2)
db.testcol.insertOne( { a : 1 } )
db.testcol.find()
var x = "asdf"
db.testcol.insertOne( { a : x.pad(10000000) } )  // ~10 MB
db.setProfilingLevel(0)
db.system.profile.find().pretty()
```

Note:

- Mention to the students what the fields mean.
- Things to keep in mind:

- op can be command, query, or update
- ns is sometimes the db.<COLLECTION> namespace
  - but sometimes db.$cmd for commands
- key updates refers to index keys
- ts (timestamp) is useful for some queries if problems cluster.

`db.serverStatus()`

- Takes a snapshot of server status.
- By taking diffs, you can see system trends.
- Most of the data that MMS gets is from here.

Exercise: Using `db.serverStatus()`

- Open up two windows. In the first, type:

```
db.testcol.drop()
var x = "asdf"
for (i=0; i<=10000000; i++) {
    db.testcol.insertOne( { a : x.pad(100000) } )
}
```

- In the second window, type periodically:

```
var x = db.serverStatus(); x.metrics.document
```

Analyzing Profiler Data

- Enable the profiler at default settings.
- Run for 5 seconds.
- Slow operations are captured.
- The issue is there is not a proper index on the message field.
- You will see how fast documents are getting inserted.
- It will be slow b/c the documents are big.

Performance Improvement Techniques

- Appropriate write concerns
- Bulk operations
- Good schema design
- Good Shard Key choice
- Good indexes

Performance Tips: Write Concern

- Increasing the write concern increases data safety.
- This will have an impact on performance, however.
- This is especially true when there are network issues.
- You will want to balance business needs against speed.

Bulk Operations

- Using bulk operations (including `insertMany` and `updateMany`) can improve performance, especially when using write concern greater than 1.
- These enable the server to amortize acknowledgement.
- Can be done with both `insertMany` and `updateMany` .

Exercise: Comparing `insertMany` with `mongostat`

Let's spin up a 3-member replica set:

```
mkdir -p /data/replset/{1,2,3}
mongod --logpath /data/replset/1/mongod.log \
       --dbpath /data/replset/1 --replSet mySet --port 27017 --fork
mongod --logpath /data/replset/2/mongod.log \
       --dbpath /data/replset/2 --replSet mySet --port 27018 --fork
mongod --logpath /data/replset/3/mongod.log \
       --dbpath /data/replset/3 --replSet mySet --port 27019 --fork

echo "conf = {_id: 'mySet', members: [{_id: 0, host: 'localhost:27017'}, \
     {_id: 1, host: 'localhost:27018'}, {_id: 2, host: 'localhost:27019'}]}; \
     rs.initiate(conf)" | mongo
```

mongostat, insertOne with {w: 1}

Perform the following, with writeConcern : 1 and insertOne():

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    for (j=1; j<=1000; j++) {
        db.testcol.insertOne( { _id : (1000 * (i-1) + j),
                                a : i, b : j, c : (1000 * (i-1)+ j) },
                              { writeConcern : { w : 1 } } );
    };
    print(i);
}
```

Run mongostat and see how fast that happens.

Multiple insertOne s with {w: 3}

Increase the write concern to 3 (safer but slower):

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    for (j=1; j<=1000; j++) {
        db.testcol.insertOne(
            { _id: (1000 * (i-1) + j), a: i, b: j, c: (1000 * (i-1)+ j)},
            { writeConcern: { w: 3 } }
        );
    };
    print(i);
}
```

Again, run mongostat.

mongostat, insertMany with {w: 3}

- Finally, let's use insertMany to our advantage:
- Note that writeConcern is still { w: 3 }

```
db.testcol.drop()
for (i=1; i<=10000; i++) {
    arr = []
    for (j=1; j<=1000; j++) {
        arr.push(
            { _id: (1000 * (i-1) + j), a: i, b: j, c: (1000 * (i-1)+ j) }
        );
    };
    db.testcol.insertMany( arr, { writeConcern : { w : 3 } } );
    print(i);
}
```

Schema Design

- The structure of documents affects performance.
- Optimize for your application's read/write patterns.
- We want as few requests to the database as possible to perform a given application task.
- See the data modeling section for more information.

Shard Key Considerations

- Choose a shard key that distributes load across your cluster.
- Create a shard key such that only a small number of documents will have the same value.
- Create a shard key that has a high degree of randomness.
- Your shard key should enable a mongos to target a single shard for a given query.

Indexes and Performance

- Reads and writes that don't use an index will cripple performance.
- In compound indexes, order matters:
  - Sort on a field that comes before any range used in the index.
  - You can't skip fields; they must be used in order.
  - Revisit the indexing section for more detail.

8 Backup and Recovery

Backup and Recovery **(page 138)** An overview of backup options for MongoDB

8.1 Backup and Recovery

Disasters Do Happen

Human Disasters

Terminology: RPO vs. RTO

- **Recovery Point Objective (RPO)**: How much data can you afford to lose?
- **Recovery Time Objective (RTO)**: How long can you afford to be off-line?

Terminology: DR vs. HA

- **Disaster Recovery** (DR)
- **High Availability** (HA)
- Distinct business requirements
- Technical solutions may converge

Quiz

- Q: What's the hardest thing about backups?
- A: Restoring them!
- **Regularly test that restoration works!**

Backup Options

- Document Level
  - Logical
  - `mongodump`, `mongorestore`
- File system level
  - Physical
  - Copy files
  - Volume/disk snapshots

Document Level: `mongodump`

- Dumps collection to BSON files
- Mirrors your structure
- Can be run live or in offline mode
- Does not include indexes (rebuilt during restore)
- `--dbpath` for direct file access
- `--oplog` to record oplog while backing up
- `--query/filter` selective dump

`mongodump`

```
$ mongodump --help
Export MongoDB data to BSON files.

options:
  --help                  produce help message
  -v [ --verbose ]        be more verbose (include multiple times for
                          more verbosity e.g. -vvvvv)
  --version               print the program's version and exit
  -h [ --host ] arg       mongo host to connect to ( /s1,s2 for
  --port arg              server port. Can also use --host hostname
  -u [ --username ] arg   username
  -p [ --password ] arg   password
  --dbpath arg            directly access mongod database files in path
  -d [ --db ] arg         database to use
  -c [ --collection ] arg collection to use (some commands)
  -o [ --out ] arg (=dump)output directory or "-" for stdout
  -q [ --query ] arg      json query
  --oplog                 Use oplog for point-in-time snapshotting
```

File System Level

- **Must use journaling!**
- Copy `/data/db` files
- Or snapshot volume (e.g., LVM, SAN, EBS)
- *Seriously, always use journaling!*

Ensure Consistency

Flush RAM to disk and stop accepting writes:

- `db.fsyncLock()`
- Copy/Snapshot
- `db.fsyncUnlock()`

File System Backups: Pros and Cons

- Entire database
- Backup files will be large
- Fastest way to create a backup
- Fastest way to restore a backup

Document Level: `mongorestore`

- `mongorestore`
- `--oplogReplay` replay oplog to point-in-time

File System Restores

- All database files
- Selected databases or collections
- Replay Oplog

Backup Sharded Cluster

1. Stop Balancer (and wait) or no balancing window
2. Stop one config server (data R/O)
3. Backup Data (shards, config)
4. Restart config server
5. Resume Balancer

Restore Sharded Cluster

1. Dissimilar # shards to restore to
2. Different shard keys?
3. Selective restores
4. Consolidate shards
5. Changing addresses of config/shards

Tips and Tricks

- `mongodump/mongorestore`
  - `--oplog[Replay]`
  - `--objcheck/--repair`
  - `--dbpath`
  - `--query/--filter`
- `bsondump`
  - inspect data at console
- LVM snapshot time/space tradeoff
  - Multi-EBS (RAID) backup
  - clean up snapshots

Find out more
mongodb.com | mongodb.org
university.mongodb.com

Having trouble?
File a JIRA ticket:
jira.mongodb.org

Follow us on twitter
@MongoDBInc
@MongoDB