MongoDB Crash Course

A quick practical tour of MongoDB

Agenda

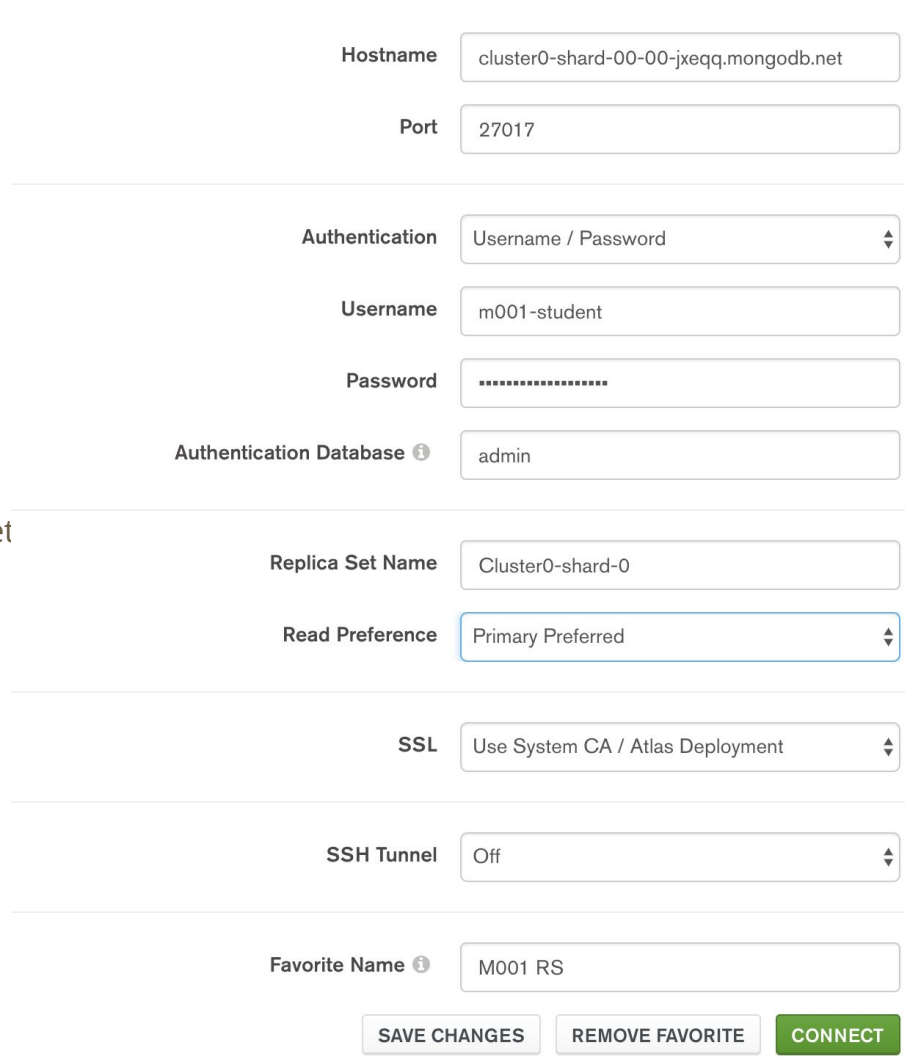
- MongoDB Overview
- Browsing data and schemata with Compass on Atlas
- Using Mongo Shell and starting a local single server
- CRUD Operations, Full Text Search, GeoSpatial Queries
- Read and Write Concerns
- The Aggregation Framework; Using Compass Aggregation Editor
- Data Models and Transactions (new in 4.0)
- Indexes and Optimization
- Replication and Sharding; Deployment Architectures
- Writing in Java for Mongo using Spring Boot: a Primer

Basics: CRUD Operations

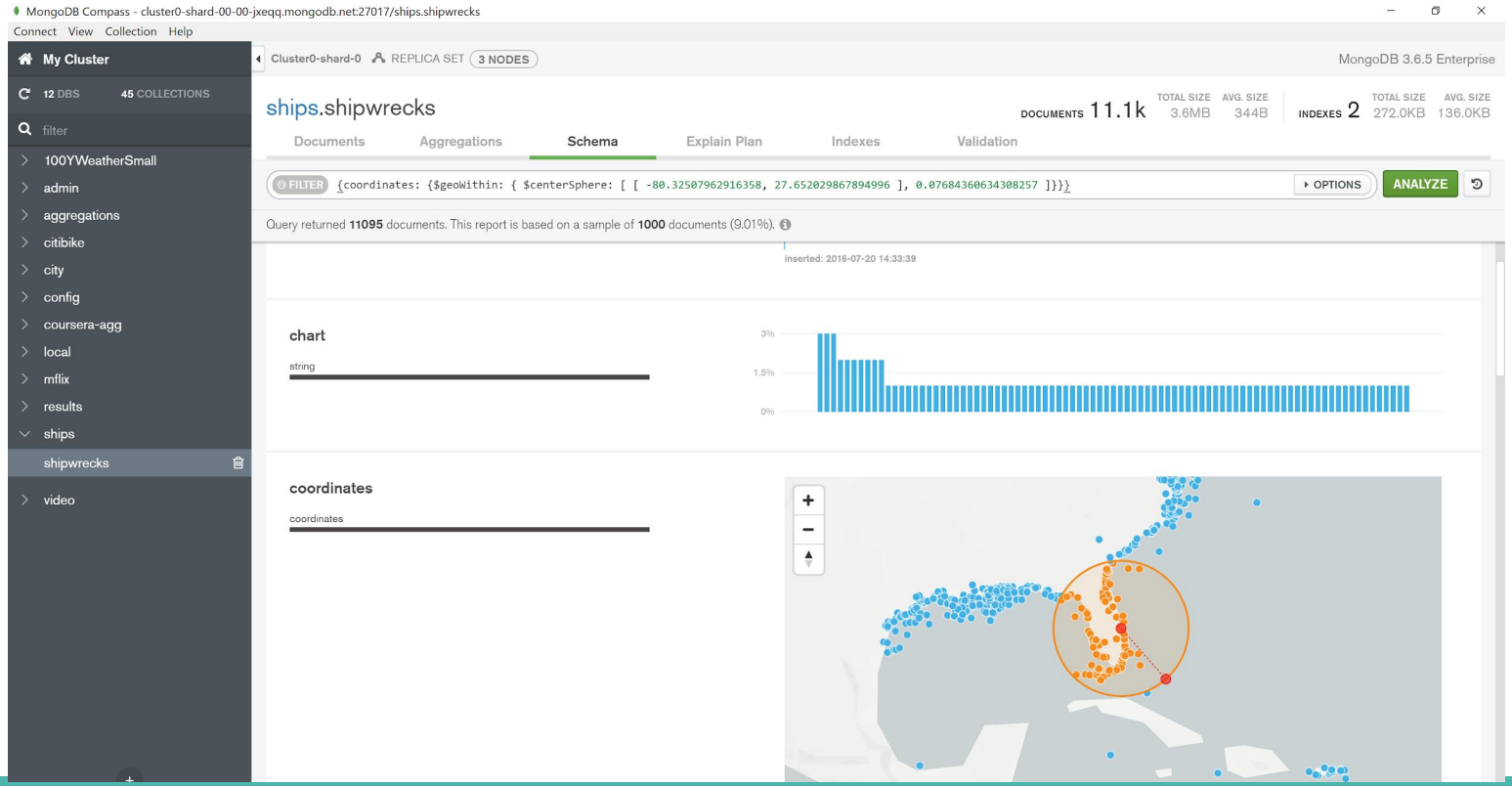
Connecting to the sample data sets

Using Compass (NOT Community):

Hostname: cluster0-shard-00-00-jxeqq.mongodb.net
Username: m001-student
Password: m001-mongodb-basics
Replica Set Name: Cluster0-shard-0
Read Preference: Primary Preferred



Schema Browsing



Start server and Mongo Shell

https://docs.mongodb.com/manual/mongo/

- Installation
- **The mongo Shell**
  - Configure the mongo Shell
  - Access the mongo Shell Help
  - Write Scripts for the mongo Shell
  - Data Types in the mongo Shell
  - mongo Shell Quick Reference

Overview of CRUD operations

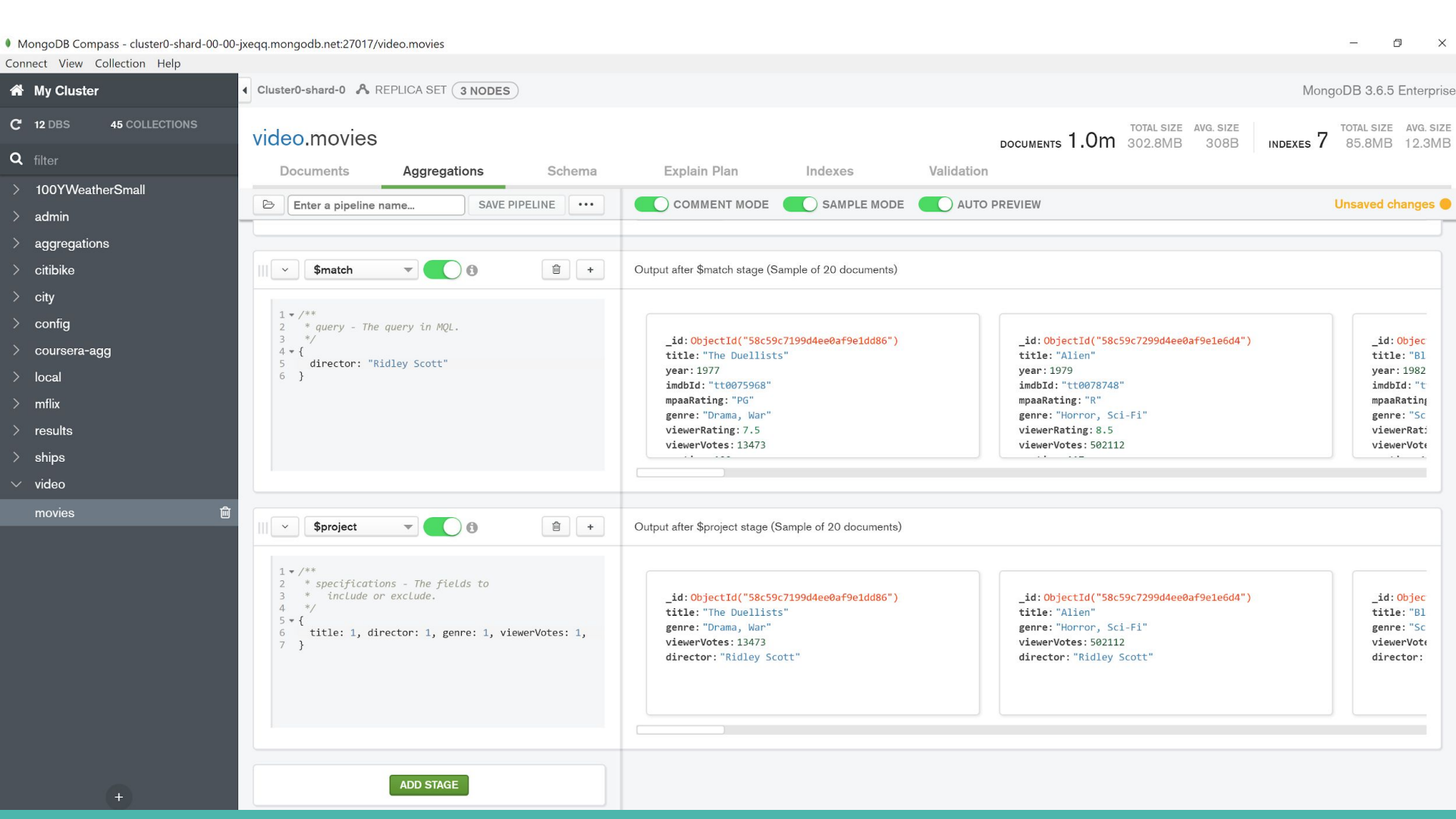
- **MongoDB CRUD Operations**
  - Insert Documents
  - Query Documents
  - Update Documents
  - Delete Documents
  - Bulk Write Operations
  - Retryable Writes
  - SQL to MongoDB Mapping Chart
  - Text Search
  - Geospatial Queries
  - Read Isolation (Read Concern)
  - Write Acknowledgement (Write Concern)
  - MongoDB CRUD Concepts

Aggregation Pipeline

Overview of the Aggregation Pipeline

Aggregation Pipeline

- Aggregation Pipeline Optimization
- Aggregation Pipeline Limits
- Aggregation Pipeline and Sharded Collections
- Example with ZIP Code Data
- Example with User Preference Data


MongoDB Compass - cluster0-shard-00-00-jxeqq.mongodb.net:27017/video.movies
Connect View Collection Help
My Cluster
12 DBS 45 COLLECTIONS
filter
100YWeatherSmall
admin
aggregations
citibike
city
config
coursera-agg
local
mflix
results
ships
video
movies
Cluster0-shard-0 REPLICA SET 3 NODES
MongoDB 3.6.5 Enterprise
video.movies
DOCUMENTS 1.0m TOTAL SIZE 302.8MB AVG. SIZE 308B
INDEXES 7 TOTAL SIZE 85.8MB AVG. SIZE 12.3MB
Documents Aggregations Schema Explain Plan Indexes Validation
Enter a pipeline name...
SAVE PIPELINE
COMMENT MODE SAMPLE MODE AUTO PREVIEW
Unsaved changes
$match
/**
* query - The query in MQL.
*/
{
director: "Ridley Scott"
}
Output after $match stage (Sample of 20 documents)
_id: ObjectId("58c59c7199d4ee0af9e1dd86")
title: "The Duellists"
year: 1977
imdbId: "tt0075968"
mpaaRating: "PG"
genre: "Drama, War"
viewerRating: 7.5
viewerVotes: 13473
_id: ObjectId("58c59c7299d4ee0af9e1e6d4")
title: "Alien"
year: 1979
imdbId: "tt0078748"
mpaaRating: "R"
genre: "Horror, Sci-Fi"
viewerRating: 8.5
viewerVotes: 502112
$project
/**
* specifications - The fields to
* include or exclude.
*/
{
title: 1, director: 1, genre: 1, viewerVotes: 1,
}
Output after $project stage (Sample of 20 documents)
_id: ObjectId("58c59c7199d4ee0af9e1dd86")
title: "The Duellists"
genre: "Drama, War"
viewerVotes: 13473
director: "Ridley Scott"
_id: ObjectId("58c59c7299d4ee0af9e1e6d4")
title: "Alien"
genre: "Horror, Sci-Fi"
viewerVotes: 502112
director: "Ridley Scott"
ADD STAGE

Aggregation reference and examples

- Aggregation Reference
  - Aggregation Pipeline Quick Reference
  - Aggregation Commands
  - Aggregation Commands Comparison
  - Variables in Aggregation Expressions
  - **SQL to Aggregation Mapping Chart**

SQL Terms, Functions, and Concepts	MongoDB Aggregation Operators
WHERE	`$match`
GROUP BY	`$group`
HAVING	`$match`
SELECT	`$project`
ORDER BY	`$sort`
LIMIT	`$limit`
SUM()	`$sum`
COUNT()	`$sum`
join	`$lookup`

New in version 3.2.

Data Models and Transactions

Thinking in Documents

Data Models

- Data Modeling Introduction
- Schema Validation
- Data Modeling Concepts
    - Data Model Design
    - Operational Factors and Data Models

https://www.mongodb.com/blog/post/thinking-documents-part-1
https://www.mongodb.com/blog/post/thinking-documents-part-2

Data Model Examples and Patterns

Data Models

- Data Modeling Introduction
- Schema Validation
- Data Modeling Concepts
- Data Model Examples and Patterns
    - Model Relationships Between Documents
        - Model One-to-One Relationships with Embedded Documents
        - Model One-to-Many Relationships with Embedded Documents
        - Model One-to-Many Relationships with Document References

Using transactions in multi-document models

https://docs.mongodb.com/manual/core/transactions/

- Retriable transaction
- Retriable commit
- Read/Write concern at transaction level

```
void commitWithRetry(ClientSession clientSession) {
    while (true) {
        try {
            clientSession.commitTransaction();
            System.out.println("Transaction committed");
            break;
        } catch (MongoException e) {
            // can retry commit
            if (e.hasErrorLabel(MongoException.UNKNOWN_TRANSACTION_COMMIT_RESULT_LABEL)) {
                System.out.println("UnknownTransactionCommitResult, retrying commit operat
                continue;
            } else {
                System.out.println("Exception during commit ...");
                throw e;
            }
        }
    }
}
```

Indexes

Type of indexes

Indexes

- Single Field Indexes
- Compound Indexes
- Multikey Indexes
- Text Indexes
- Hashed Indexes

https://examples.javacodegeeks.com/software-development/mongodb/mongodb-full-text-search-tutorial/

Index mechanics

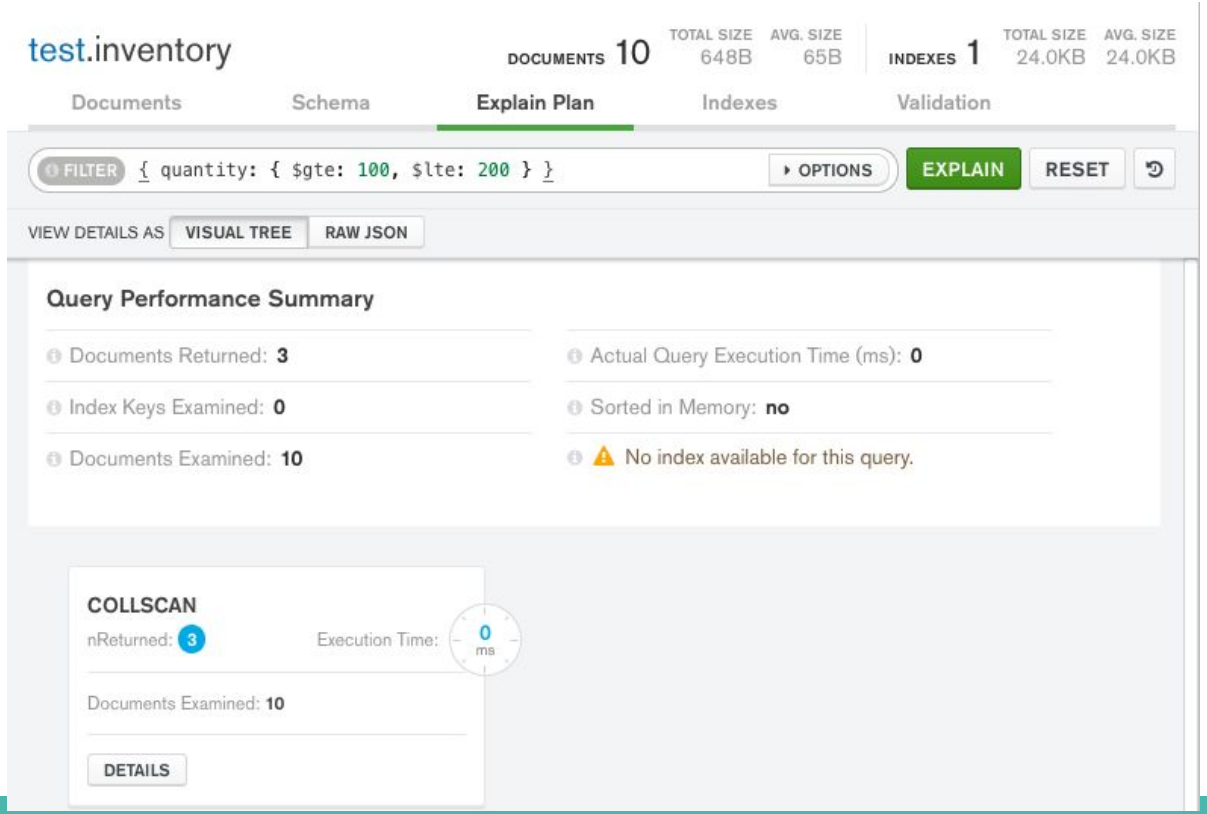
- Index Properties
- Index Build Operations on a Populated Collection
- Index Intersection
- Manage Indexes
- Measure Index Use
- Indexing Strategies
- Indexing Reference

Indexing strategies

Indexes

- Indexing Strategies
  - Create Indexes to Support Your Queries
  - Use Indexes to Sort Query Results
  - Ensure Indexes Fit in RAM
  - **Create Queries that Ensure Selectivity**

Using Indexes and Explain Plan in Compass

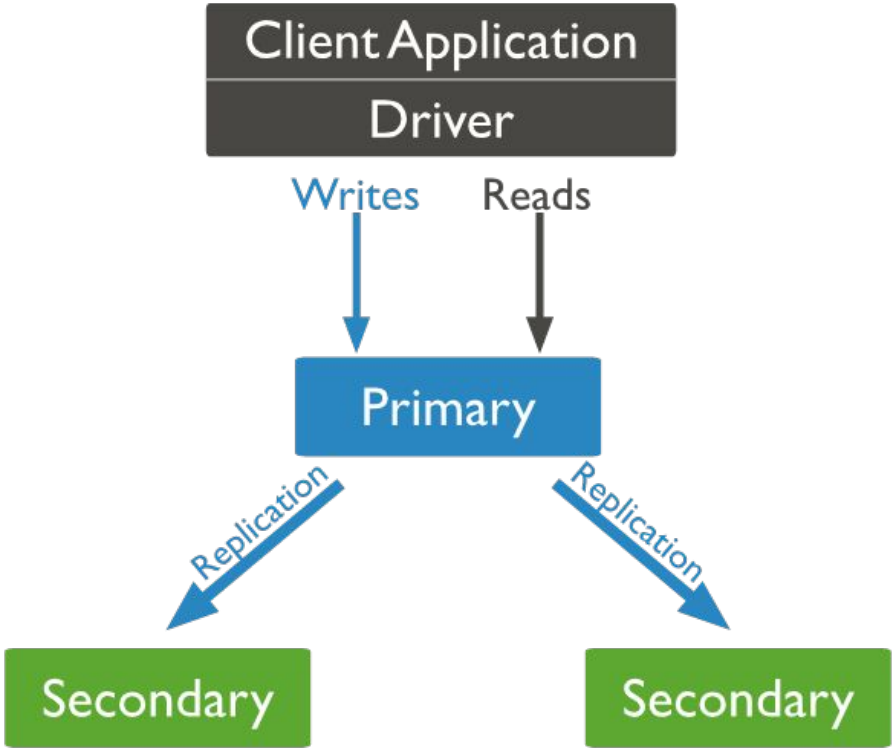


Replication and Sharding

Replica set members

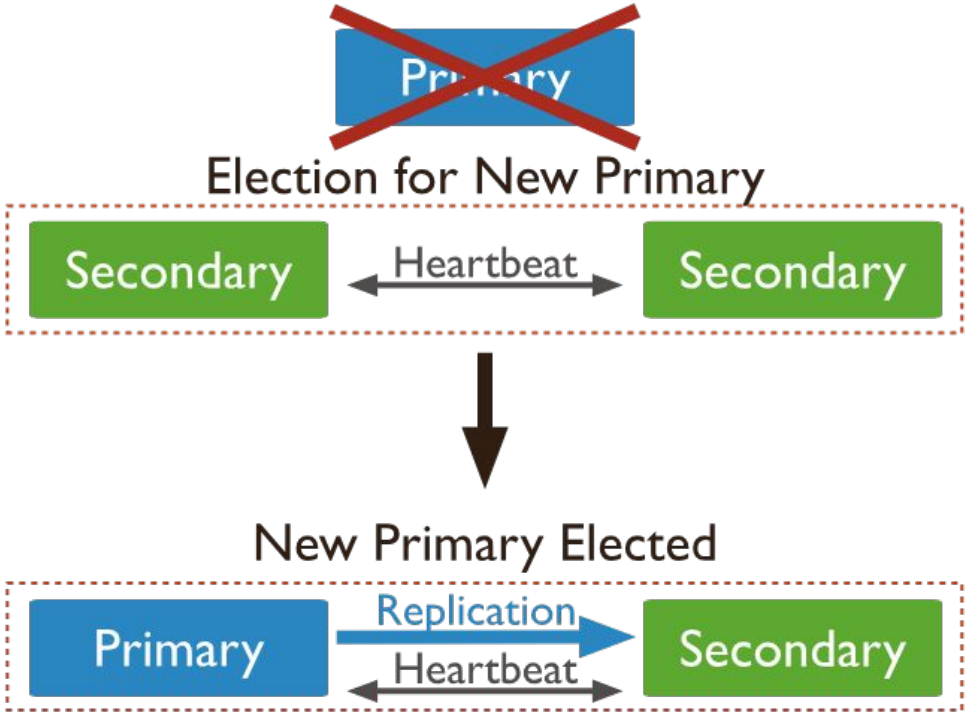
Replication

- **Replica Set Members**
  - Replica Set Primary
  - Replica Set Secondary Members
  - Replica Set Arbiter



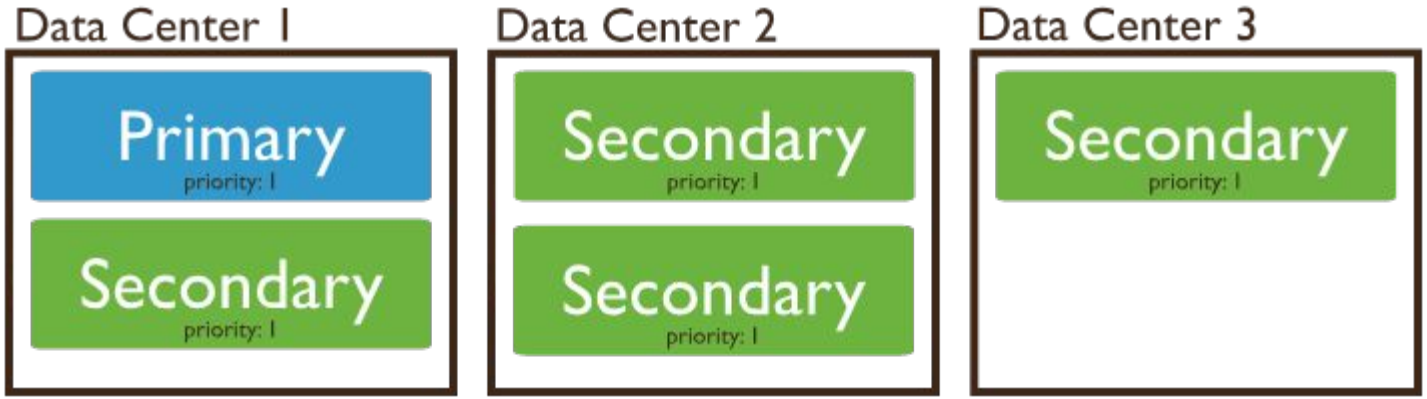
Replica Set High Availability

- Replica Set High Availability
  - **Replica Set Elections**
  - Rollbacks During Replica Set Failover



Replica Set Deployment Architectures

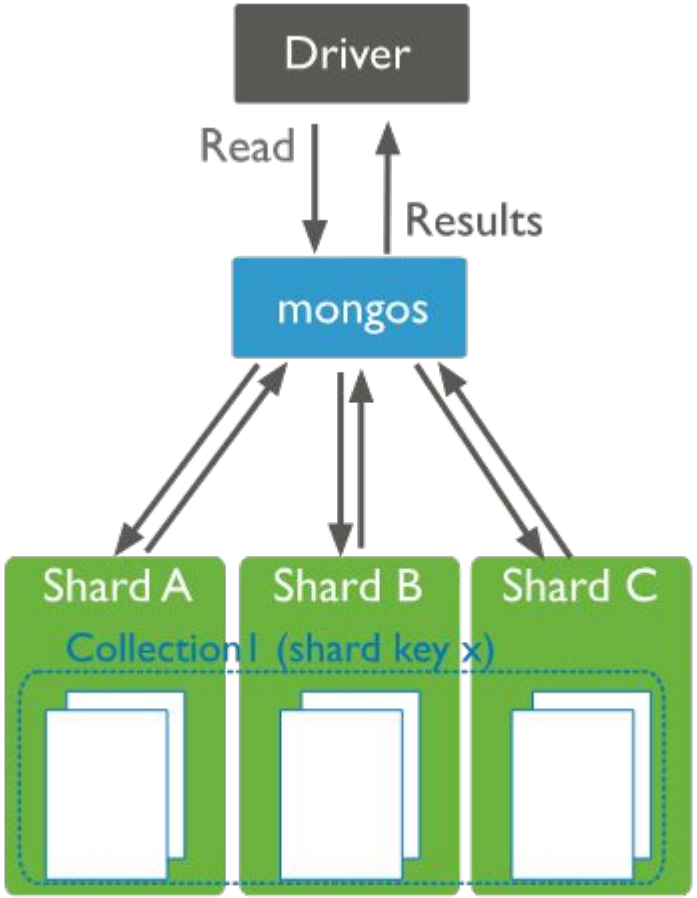
- Replica Set Deployment Architectures
  - **Three Member Replica Sets**
  - Replica Sets Distributed Across Two or More Data Centers



Sharding

Sharding

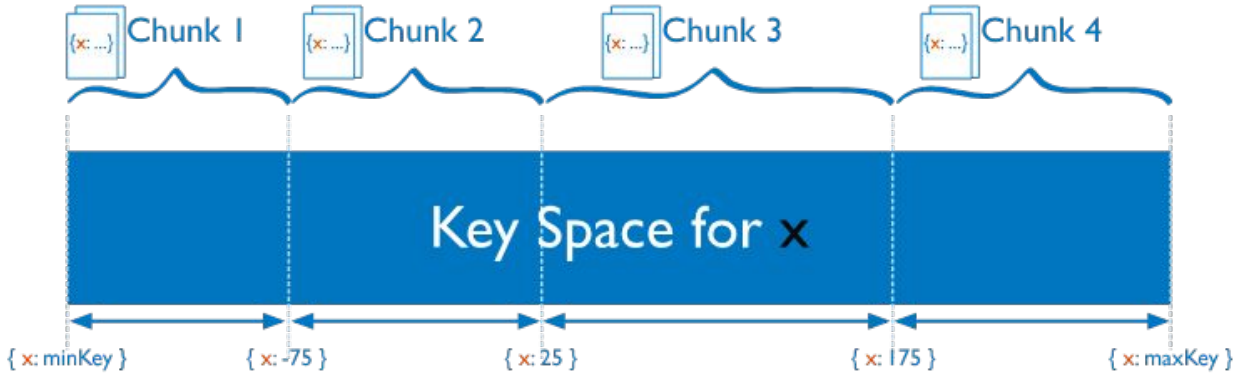
- Sharded Cluster Components
  - **Shards**
  - Config Servers (metadata)
  - Router (mongos)



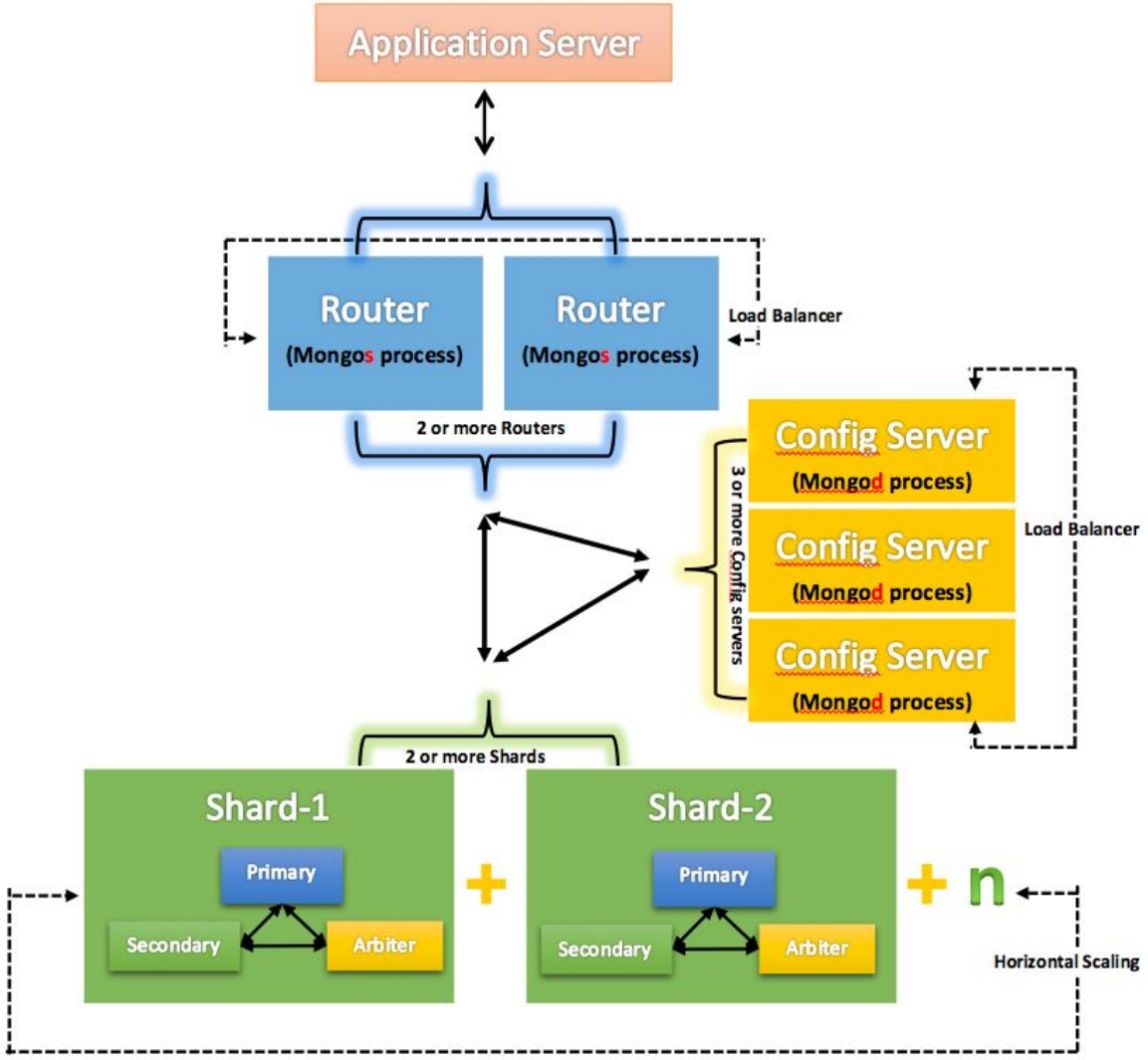
Shard Keys

Sharding

- Sharded Cluster Components
- Shard Keys
- Hashed Sharding
- **Ranged Sharding**

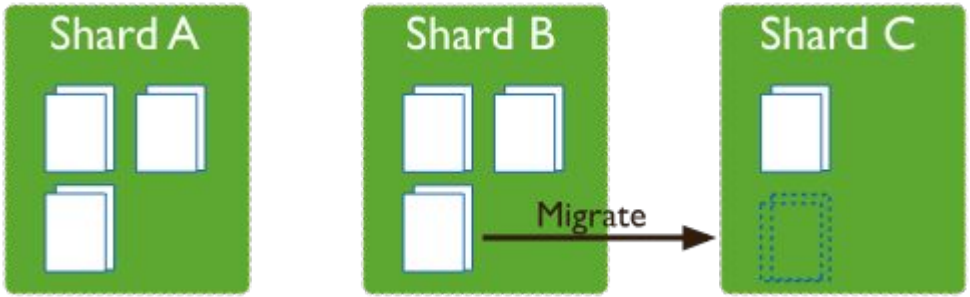


A full architecture with replica-set shards

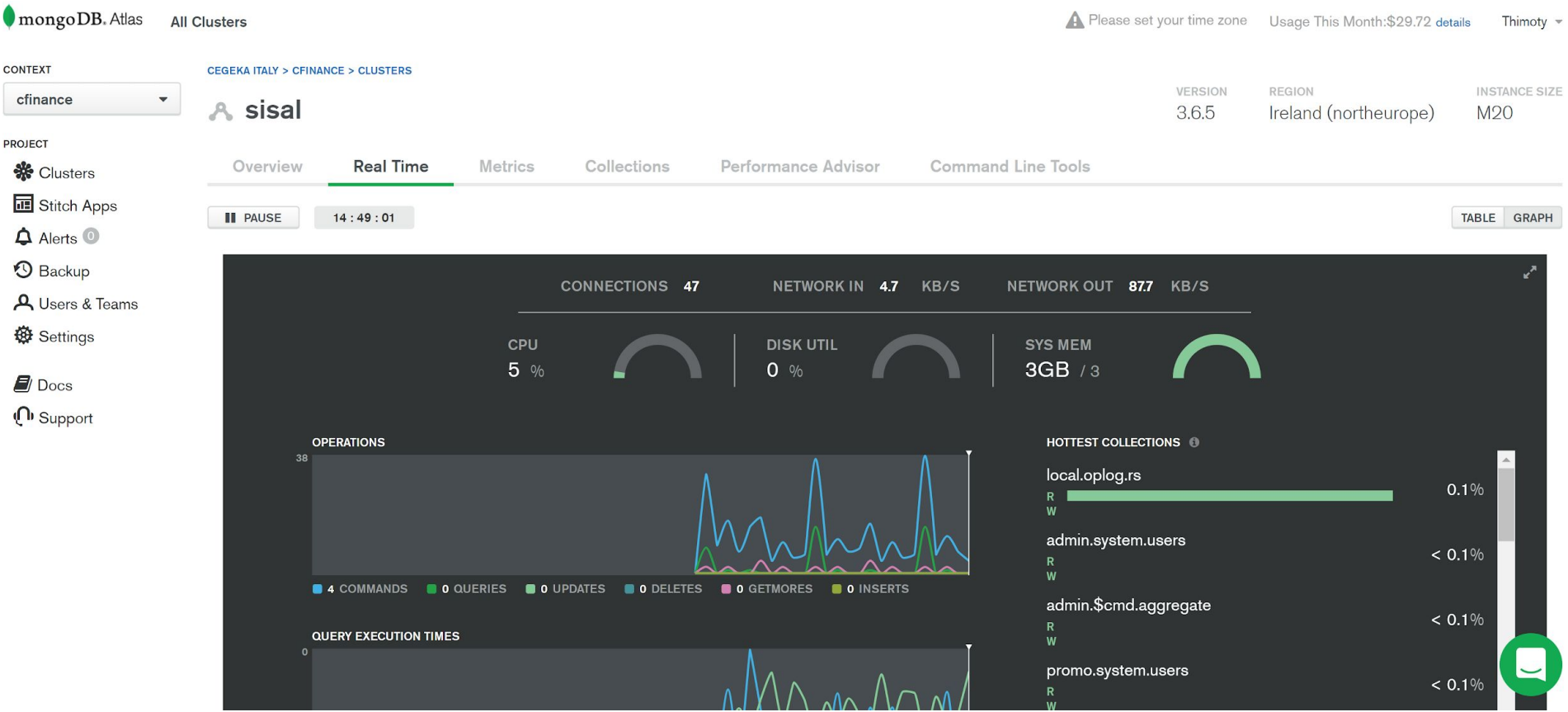


Sharded Cluster Balancer

- Data Partitioning with Chunks
- **Balancer**
  - Manage Sharded Cluster Balancer
  - Migrate Chunks in a Sharded Cluster



A tour in MongoDB Atlas


mongoDB. Atlas
All Clusters
Please set your time zone
Usage This Month:$29.72 details
Thimoty
CONTEXT
cfinance
PROJECT
Clusters
Stitch Apps
Alerts 0
Backup
Users & Teams
Settings
Docs
Support
CEGEKA ITALY > CFINANCE > CLUSTERS
sisal
VERSION
3.6.5
REGION
Ireland (northeurope)
INSTANCE SIZE
M20
Overview
Real Time
Metrics
Collections
Performance Advisor
Command Line Tools
PAUSE
14 : 49 : 01
TABLE
GRAPH
CONNECTIONS 47
NETWORK IN 4.7 KB/S
NETWORK OUT 87.7 KB/S
CPU
5 %
DISK UTIL
0 %
SYS MEM
3GB / 3
OPERATIONS
38
4 COMMANDS
0 QUERIES
0 UPDATES
0 DELETES
0 GETMORES
0 INSERTS
QUERY EXECUTION TIMES
0
HOTTEST COLLECTIONS
local.oplog.rs
R
W
0.1%
admin.system.users
R
W
< 0.1%
admin.$cmd.aggregate
R
W
< 0.1%
promo.system.users
R
W
< 0.1%

CLUSTERS > CREATE NEW CLUSTER

Create New Cluster

Global Cluster Configuration

Global Writes Enabled
AWS, 6 Zones

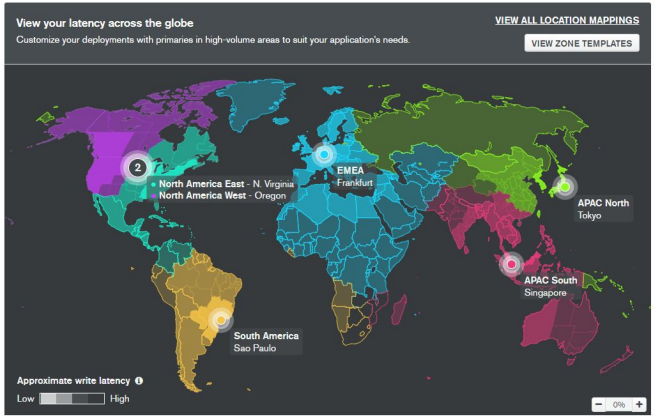
Enable Global Writes (M30 and up) ON

Low-latency reads and writes from anywhere in the world. View documentation

- Define multiple zones within a single cluster made up of one or more cloud regions.
- After deployment, we will show you how associate data with the nearest zone using location attributes.

Choose the cloud provider for all of your zones' regions

aws | Google Cloud Platform | Azure



Zone configuration summary

CONFIGURE LOCAL READS IN ALL ZONES

North America East (N. Virginia)

+ Add a Zone

Please note: Global Write Clusters require configuring shard keys after deployment. You can finish configuring your cluster after it has deployed.

$4.27/hour

Pay-as-you-go! *You will be billed hourly and can terminate your cluster anytime. Excludes variable data transfer, backup, and taxes.*

Cancel | Create Cluster

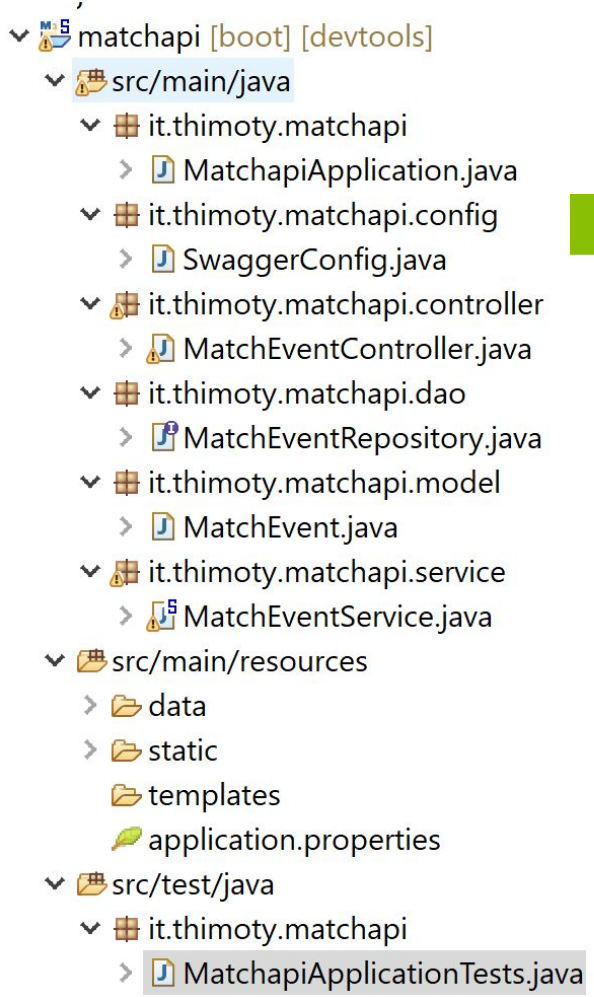
Programming using the Mongo Java Driver

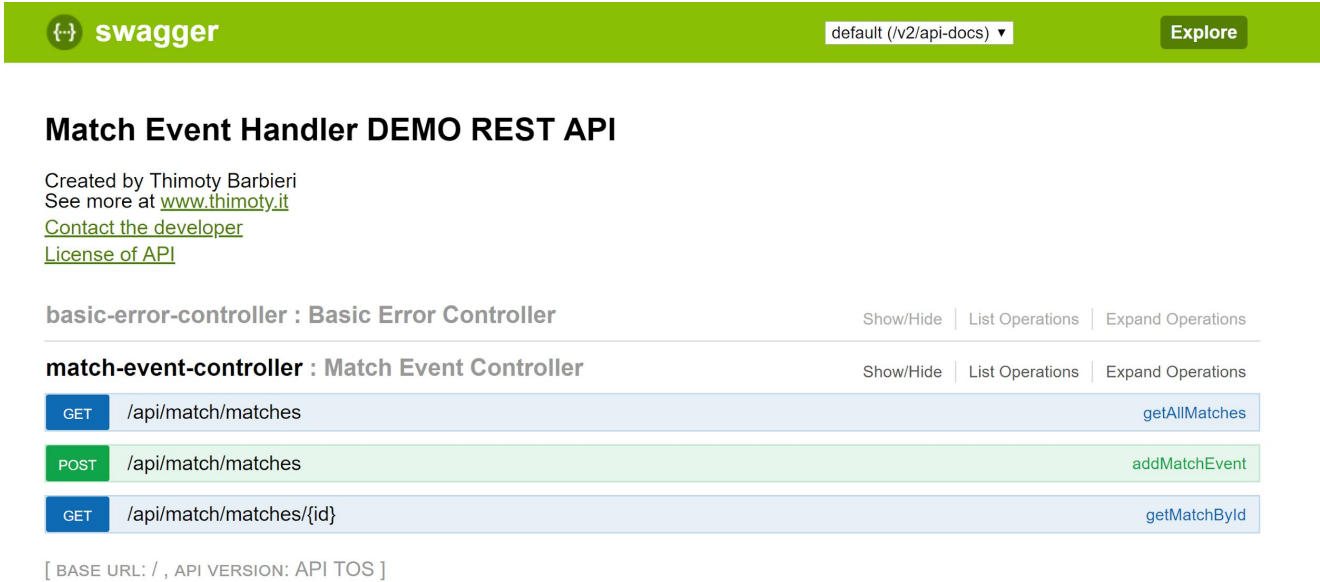
Cloning the repo with sample code

https://github.com/thimotyb/corsomongo

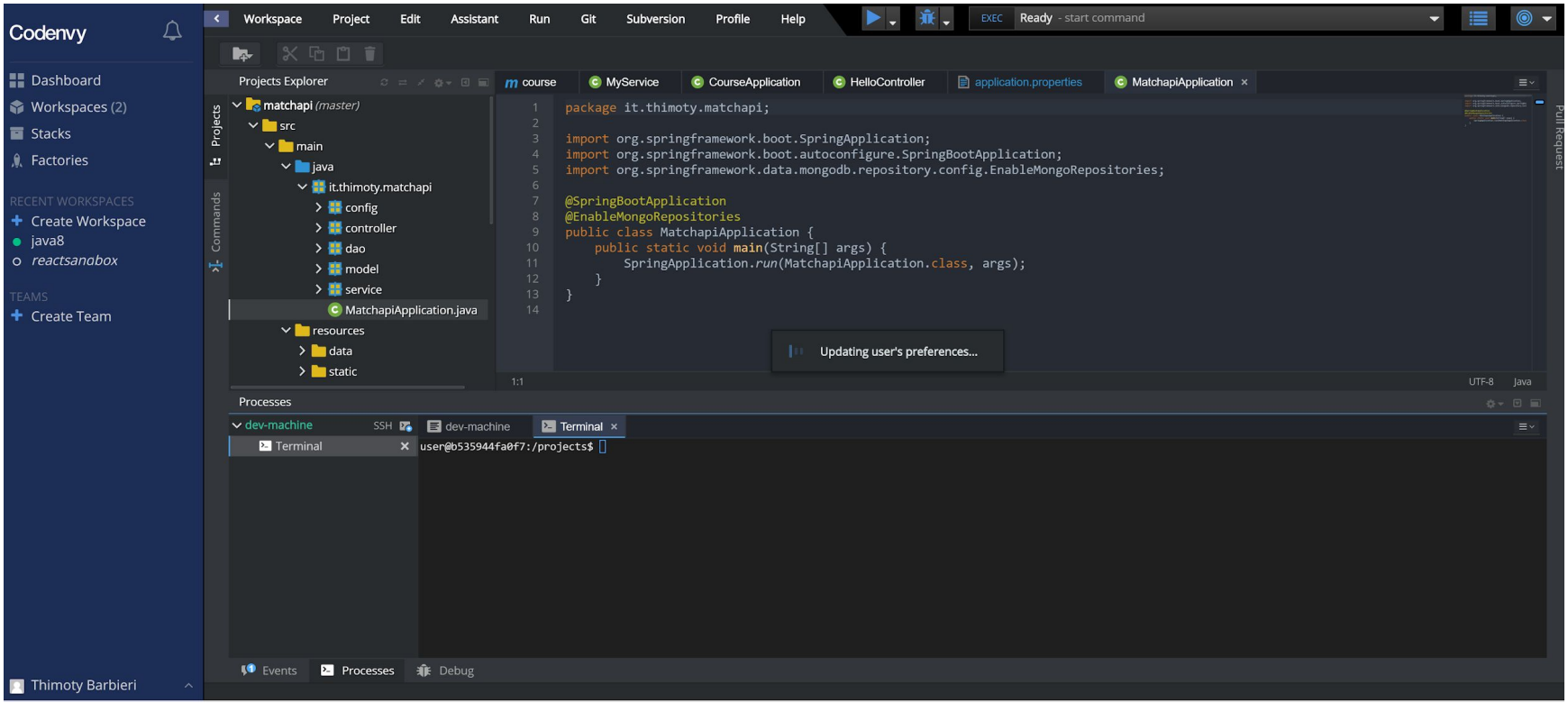
Then “Import maven project” in STS

Or: clone into your CodeEnvy.io environment
https://codenvy.io/dashboard/#/


matchapi [boot] [devtools]
src/main/java
it.thimoty.matchapi
MatchapiApplication.java
it.thimoty.matchapi.config
SwaggerConfig.java
it.thimoty.matchapi.controller
MatchEventController.java
it.thimoty.matchapi.dao
MatchEventRepository.java
it.thimoty.matchapi.model
MatchEvent.java
it.thimoty.matchapi.service
MatchEventService.java
src/main/resources
data
static
templates
application.properties
src/test/java
it.thimoty.matchapi
MatchapiApplicationTests.java


swagger
default (/v2/api-docs)
Explore
Match Event Handler DEMO REST API
Created by Thimoty Barbieri
See more at www.thimoty.it
Contact the developer
License of API
basic-error-controller : Basic Error Controller
Show/Hide
List Operations
Expand Operations
match-event-controller : Match Event Controller
Show/Hide
List Operations
Expand Operations
GET
/api/match/matches
getAllMatches
POST
/api/match/matches
addMatchEvent
GET
/api/match/matches/{id}
getMatchById
[BASE URL: / , API VERSION: API TOS]

CodeEnvy + Atlas



If the world was created by a programmer...

https://toggl.com/blog/world-created-programmer/