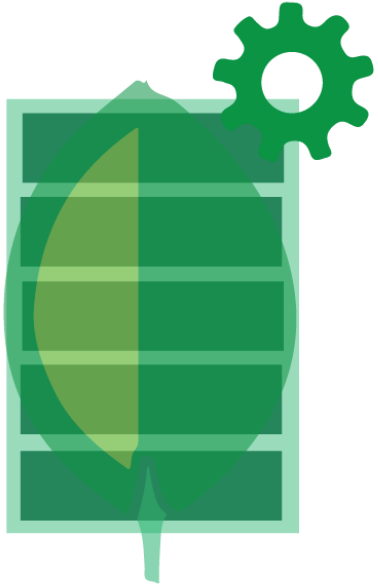


edureka!

MODULE-4
MONGODB® ADMINISTRATION

Course Topics

edureka!

- Module 1
  - Design Goals, Architecture and Installation
- Module 2
  - CRUD Operations
- Module 3
  - Schema Design and Data Modelling
- Module 4
  - **Administration**
- Module 5
  - Scalability and Availability
- Module 6
  - Indexing and Aggregation Framework
- Module 7
  - Application Engineering and MongoDB Tools
- Module 8
  - Project, Additional Concepts and Case Studies

Objectives

edureka!

At the end of this module, you will be able to

→ Take database backup and restore MongoDB®

→ Export and import data from/to a MongoDB® instance

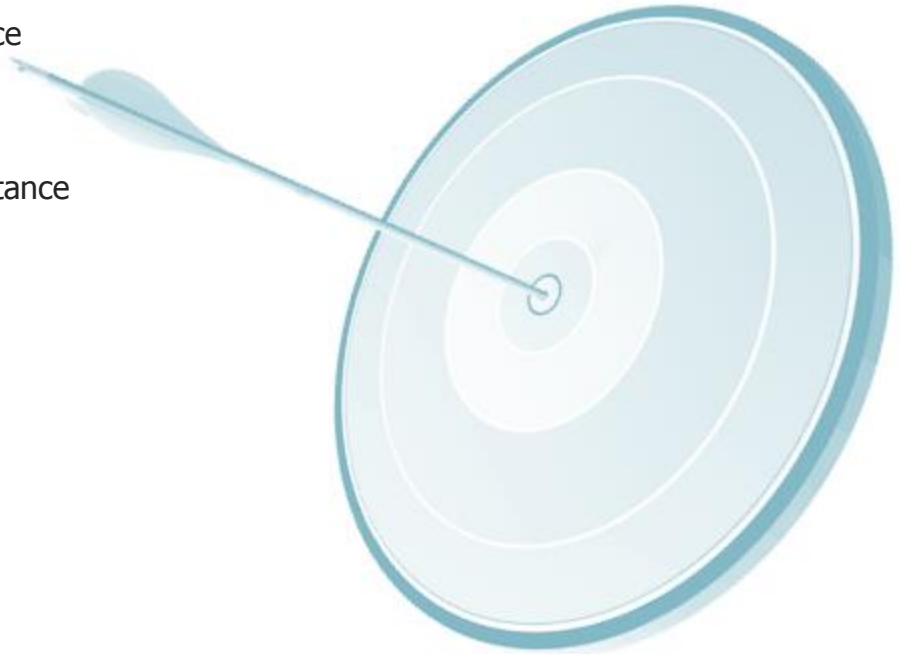
→ Check server status, db status

→ Monitor various resource utilization of a mongod instance

→ Understand various optimization strategies

→ Create capped collection

→ Use TTL in MongoDB®



Annie's Question

What is GridFS?

Annie's Answer

GridFS is a specification for storing and retrieving files that exceed the BSON-document size limit of 16MB.

Annie's Question

What is nested depth for BSON documents?

Annie's Answer

MongoDB® supports no more than 100 levels of nesting for BSON documents.

Annie's Question

What is the max size of namespace file?

Annie's Answer

Namespace files can be no larger than 2047 megabytes. By default namespace files are 16 megabytes. You can configure the size using the nssize option.

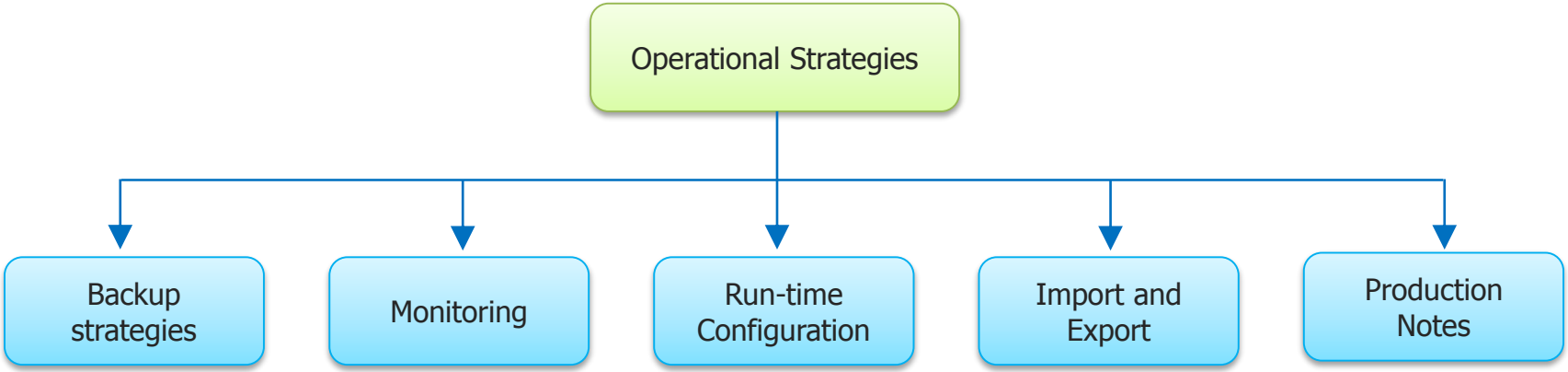
Annie's Question

How many collections GridFS stores files?

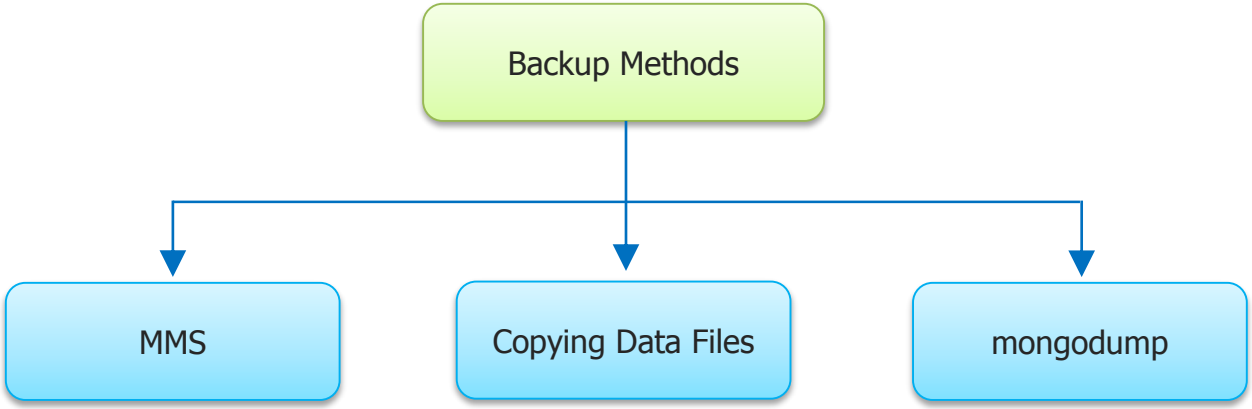
Annie's Answer

GridFS actually stores your files in two collections
chunks stores the binary chunks.
For details files stores the file's metadata.

Operational Strategies



Backup Methods



Backup Methods (MMS)

edureka!

→ The MongoDB® Management Service supports backup and restore for MongoDB deployments.

→ MMS continually backs up MongoDB® replica sets and sharded systems by reading the oplog data from your MongoDB® cluster.

→ MMS Backup offers point in time recovery of MongoDB replica sets and a consistent snapshot of sharded systems.

→ MMS achieves point in time recovery by storing oplog data so that it can create a restore for any moment in time in the last 24 hours for a particular replica set.

→ Sharded cluster snapshots are difficult to achieve with other MongoDB backup methods.

Backup Methods (Copying Data Files)

edureka!

→ We can create a backup by copying MongoDB®'s underlying data files.

→ To get a correct snapshot of a running mongod process, we must have journaling enabled.

→ This method can be used only if the volume where MongoDB stores data files supports point in time snapshots.

→ To get a consistent snapshot of a sharded system, we must disable the balancer.

→ Since copying multiple files is not an atomic operation, we must stop all writes to the mongod before copying the files.

→ We can use these snapshots to create backups of a MongoDB® system at an exact moment in time.

→ If your storage system does not support snapshots, we can copy the files directly.

Backup Methods (mongodump)

edureka!

→ The mongodump tool reads data from a MongoDB® database and creates high fidelity BSON files.

→ mongodump and mongorestore can operate against a running mongod process, and can manipulate the underlying data files directly.

→ When connected to a MongoDB® instance, mongodump can adversely affect mongod performance. If your data is larger than system memory, the queries will push the working set out of memory.

→ The mongorestore tool can populate a MongoDB® database with the data from these BSON files .

→ By default, mongodump does not capture the contents of the local database.

→ We can use mongodump directly to the data files even though mongod is not running.

Backup Methods (mongodump) Contd.

edureka!

→ If applications modify data while mongodump is creating a backup, mongodump will compete for resources with those applications.

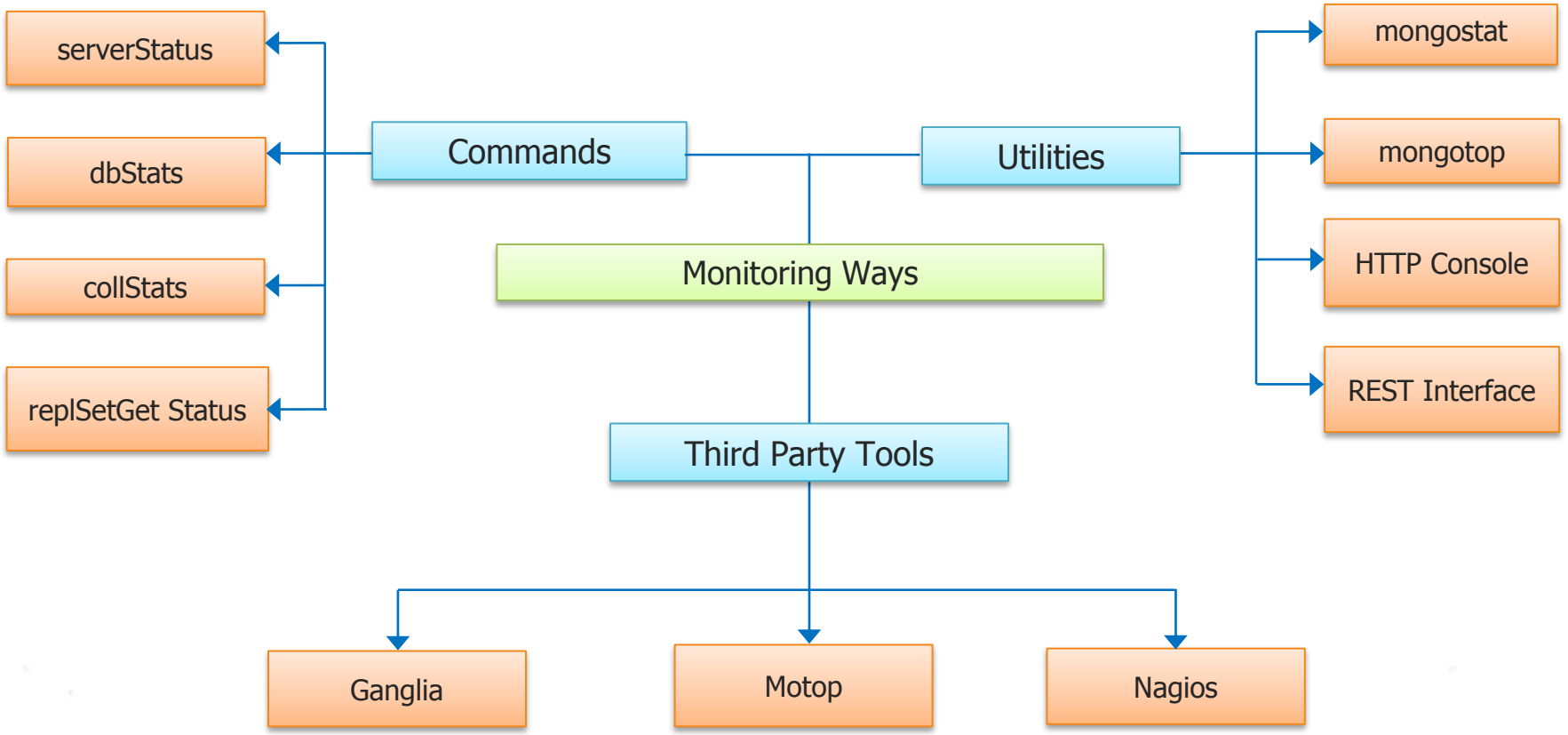
→ These tools are simple and efficient for backing up small MongoDB deployments, but are not ideal for capturing backups of larger systems.

→ mongodump only captures the documents in the database. The resulting backup is space efficient, but mongorestore must rebuild the indexes after restoring data.

→ For replica sets, mongodump also supports a point in time feature with the --oplog option. Applications may continue modifying data while mongodump captures the output.

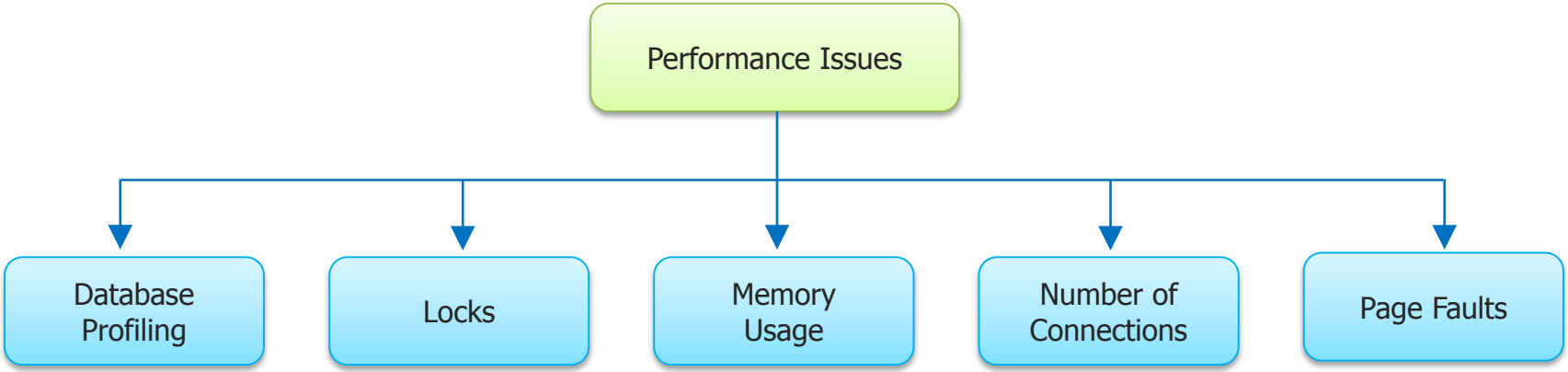
Monitoring for MongoDB®

edureka!



Diagnosing Performance Issues

edureka!



Annie's Question

When does page fault occurs in MongoDB®?

Annie's Answer

Page faults will occur if you're attempting to access part of a memory-mapped file that isn't in memory.

Annie's Question

Where can I find information about a mongod process that stopped running unexpectedly?

Annie's Answer

You can check in the mongodb log message where you have set the path for log e.g.
sudo grep mongod /var/log/messages
sudo grep score /var/log/messages

Annie's Question

What tools are available for monitoring MongoDB®?

Annie's Answer

MMS, Command line Utilities, Third Party Utilities

Annie's Question

How do I calculate how much RAM I need for my application?

Annie's Answer

The amount of RAM you need depends on several factors, including but not limited to:
The relationship between database storage and working set.
The operating system's cache strategy for LRU
The impact of journaling
The number or rate of page faults and other MMS gauges to detect when you need more RAM

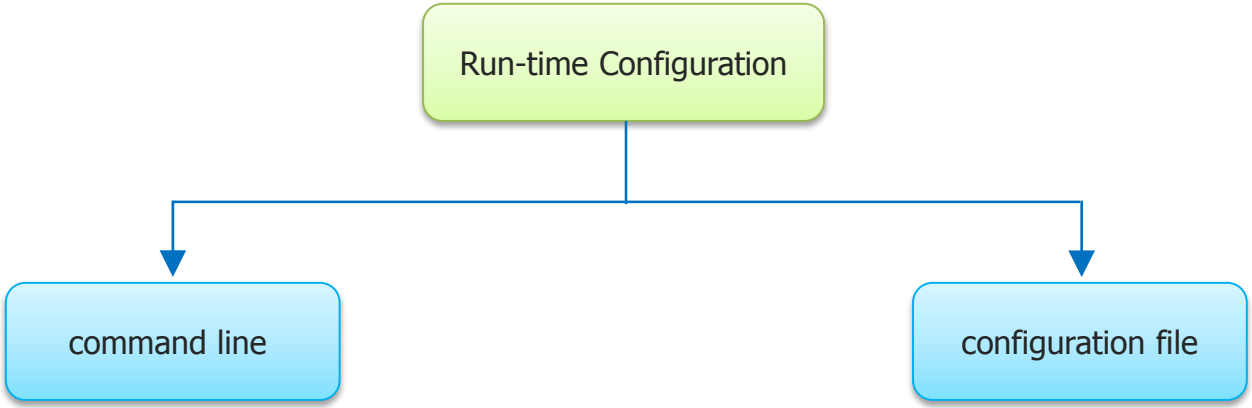
Annie's Question

Can mongodump take database backup even when mongod is not running?

Annie's Answer

Yes it can

Run-time Database Configuration



Import and Export MongoDB® Data

Annie's Question

Which kind of file can be imported and exported in MongoDB® ?

Annie's Answer

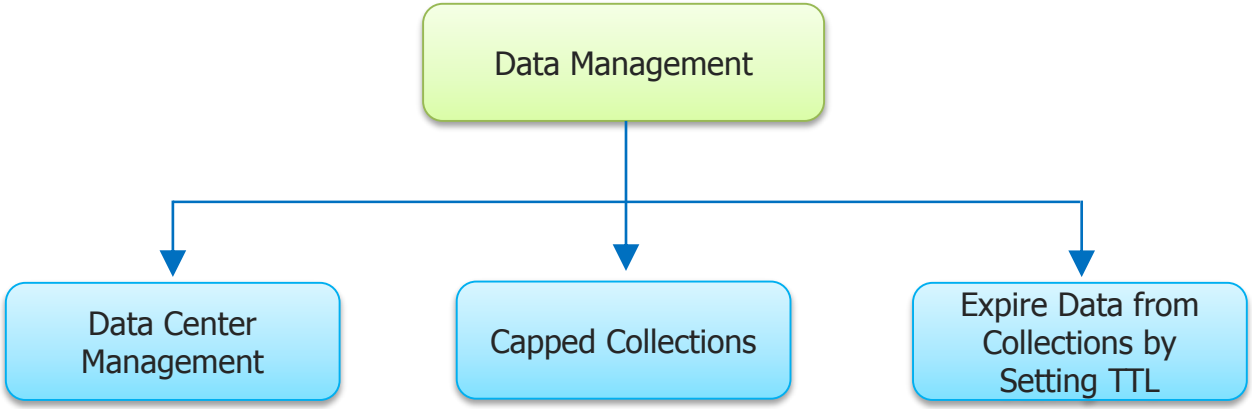
1. JSON
2. CSV
3. TSV

Import and Export MongoDB® Data

edureka!

DEMO

Data Management



Data Center Awareness

edureka!

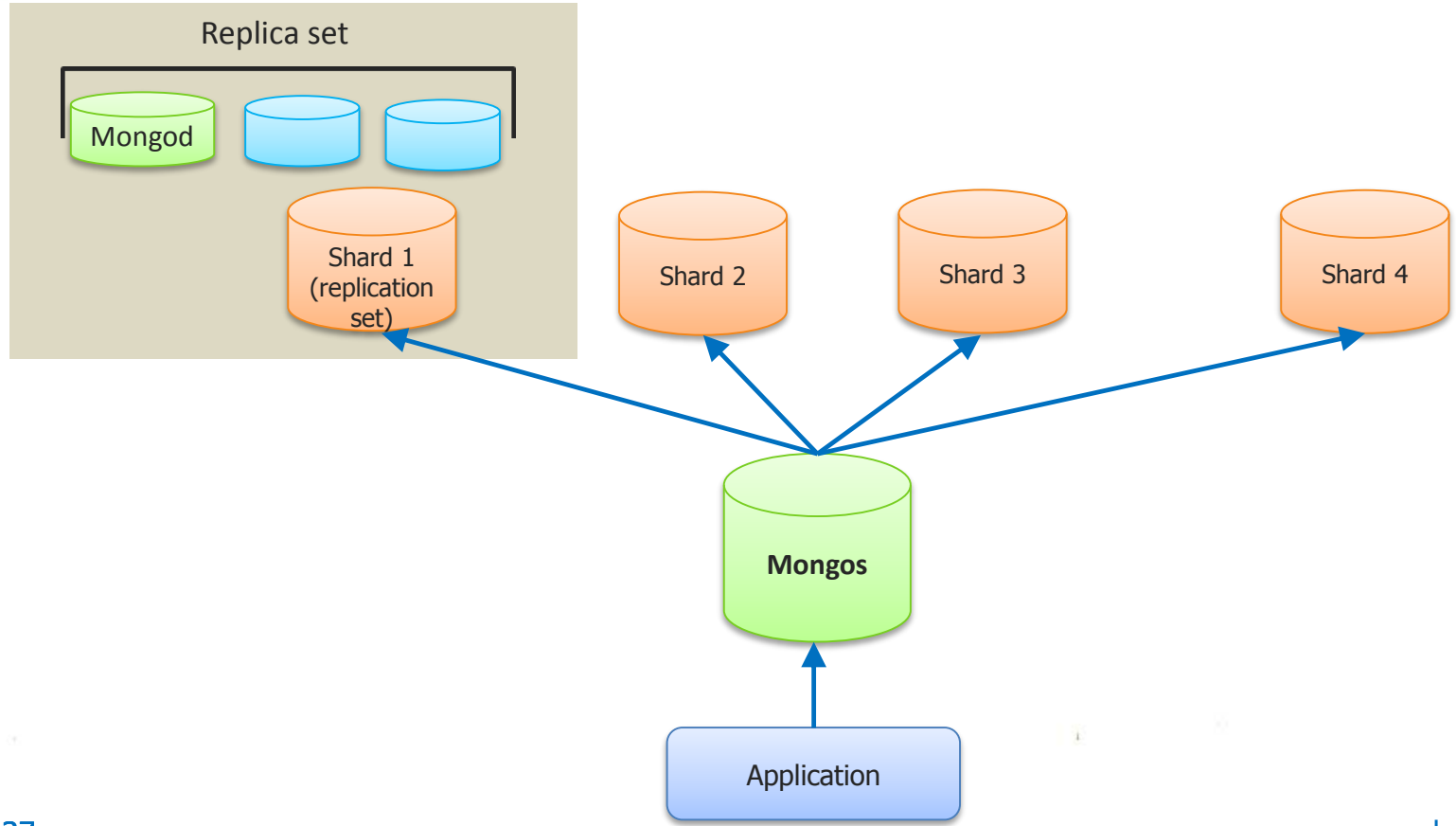
→ MongoDB® can be configured in various way for production environment.

→ It depend on kind of requirement and way of setting up environment.

→ In next slides I have shown few scenarios.

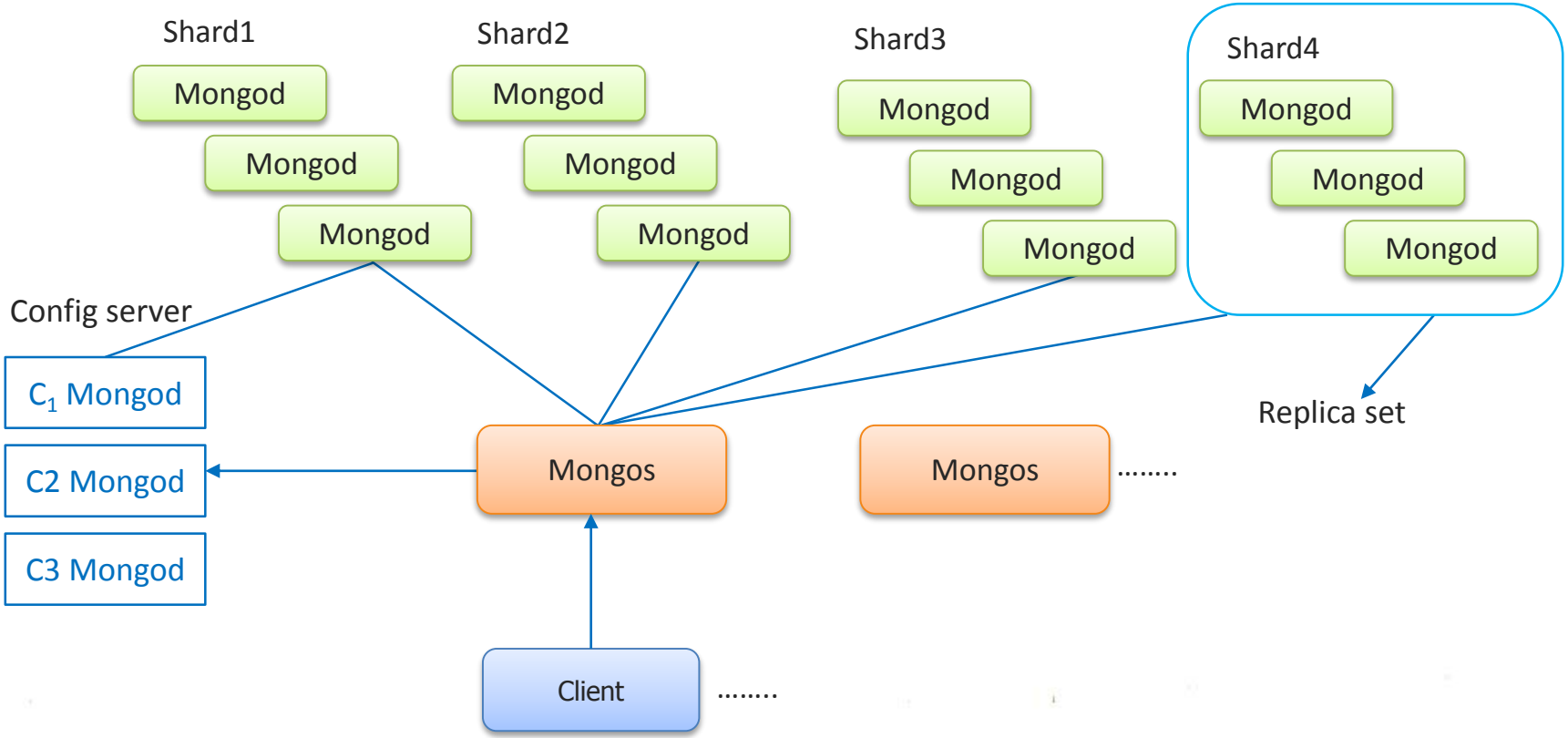
MongoDB® Production Architecture

edureka!



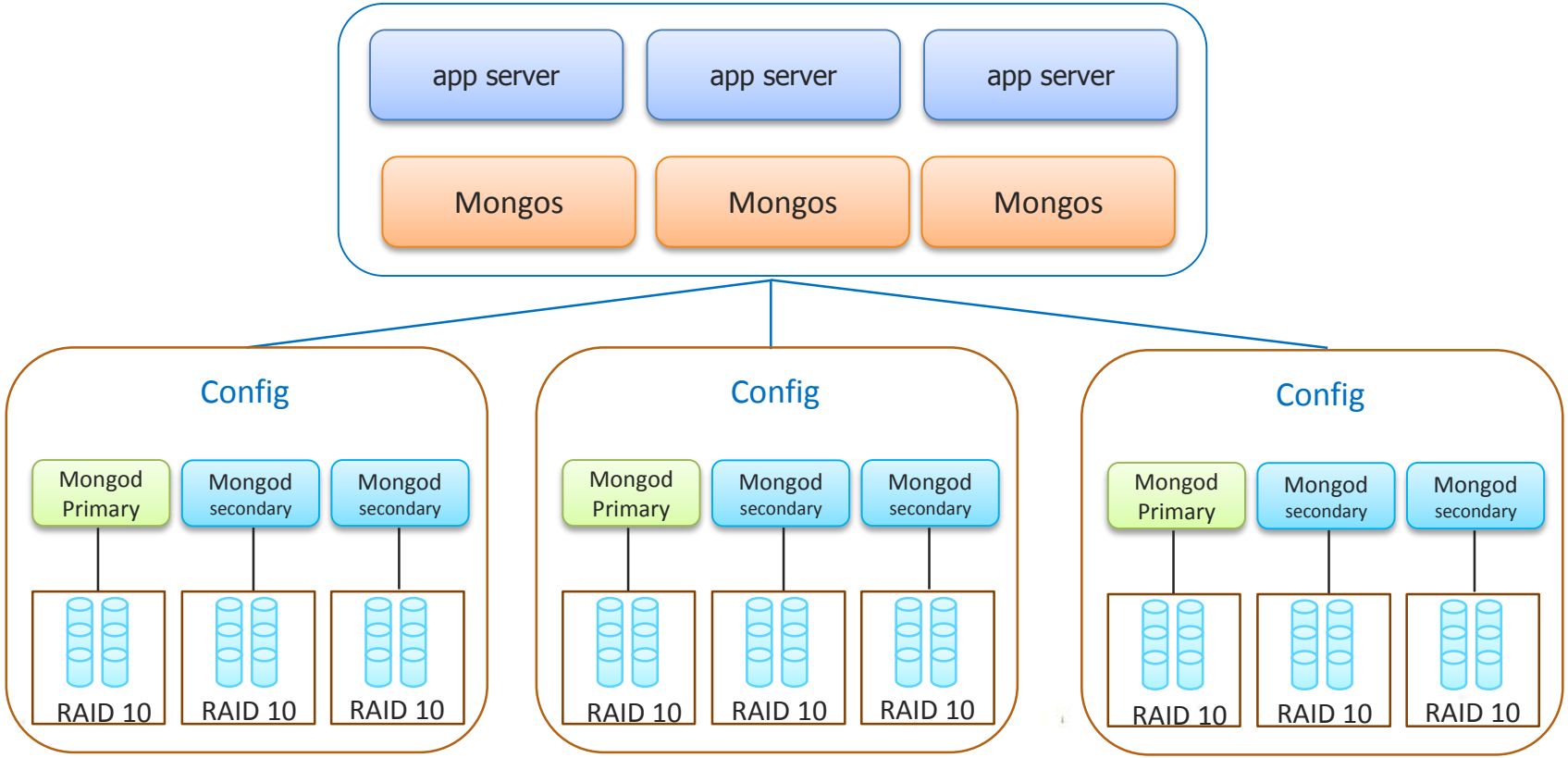
MongoDB® Production Architecture

edureka!



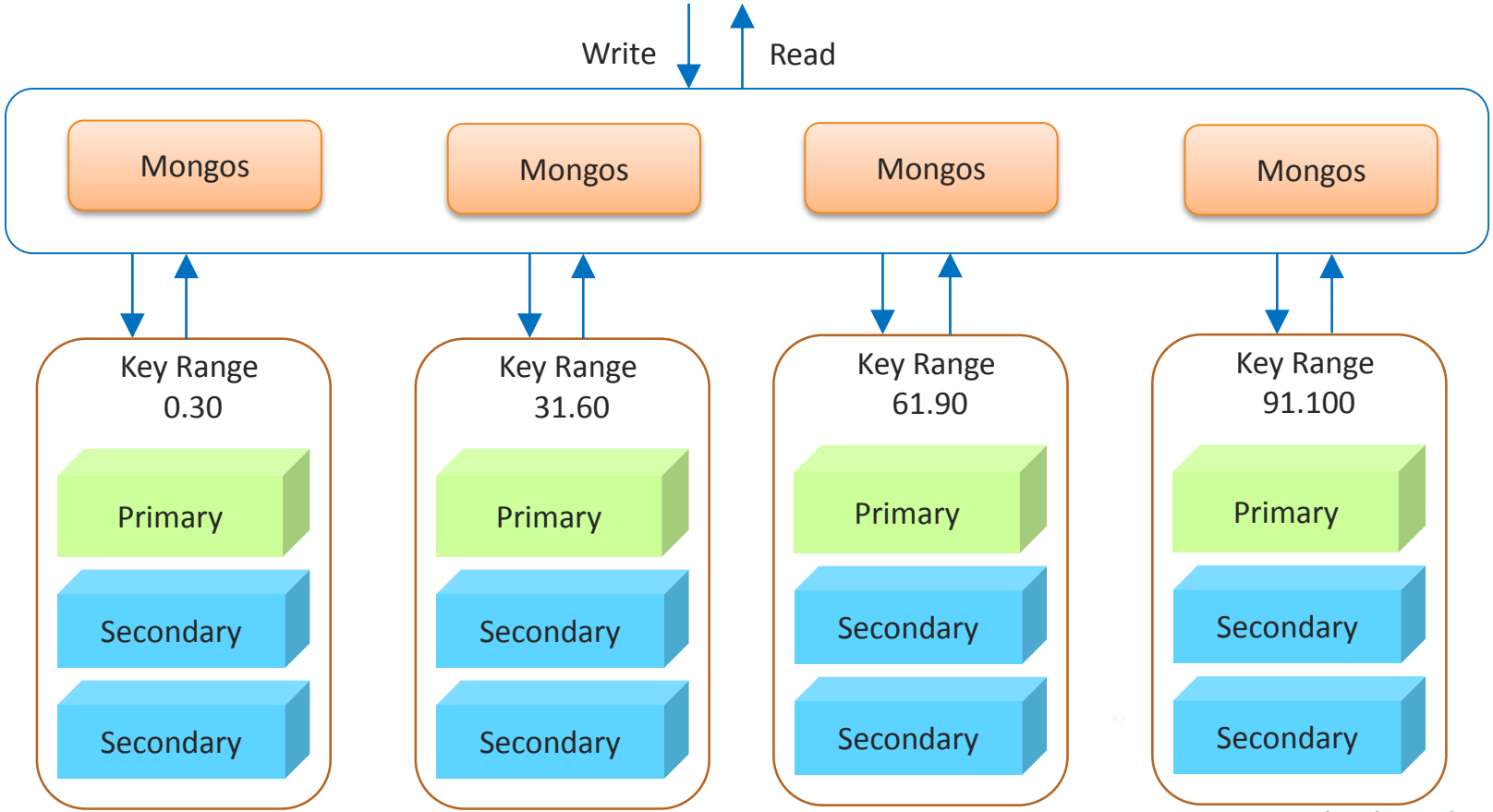
MongoDB® Production Architecture

edureka!



MongoDB® Production Architecture

edureka!



Optimization Strategies for MongoDB®

edureka!

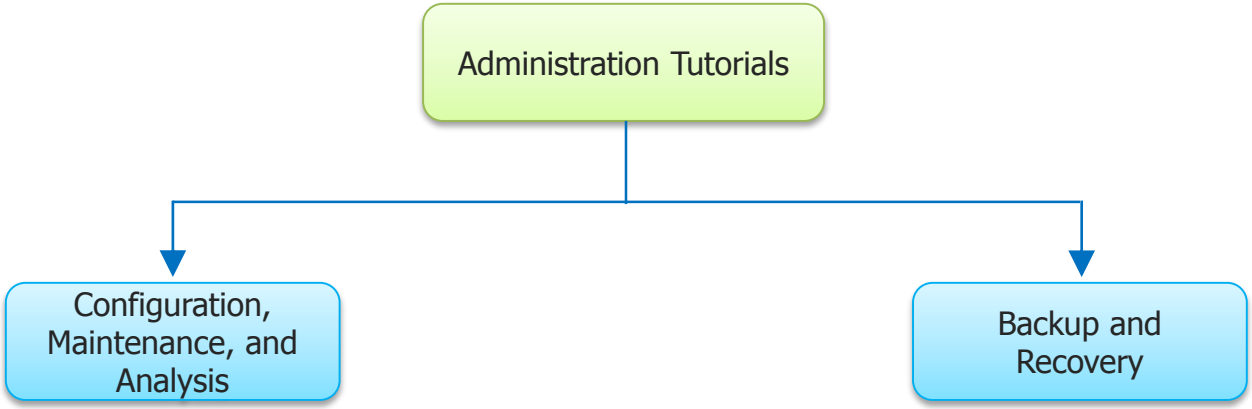
→ Evaluate Performance of Current Operations

→ Optimize Query Performance

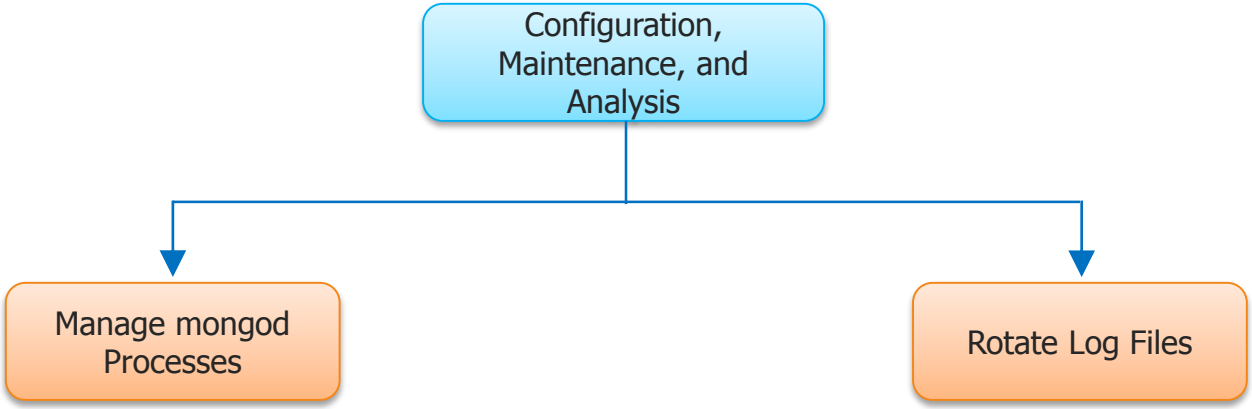
→ Design Notes

→ Use Capped Collections for Fast Writes and Reads

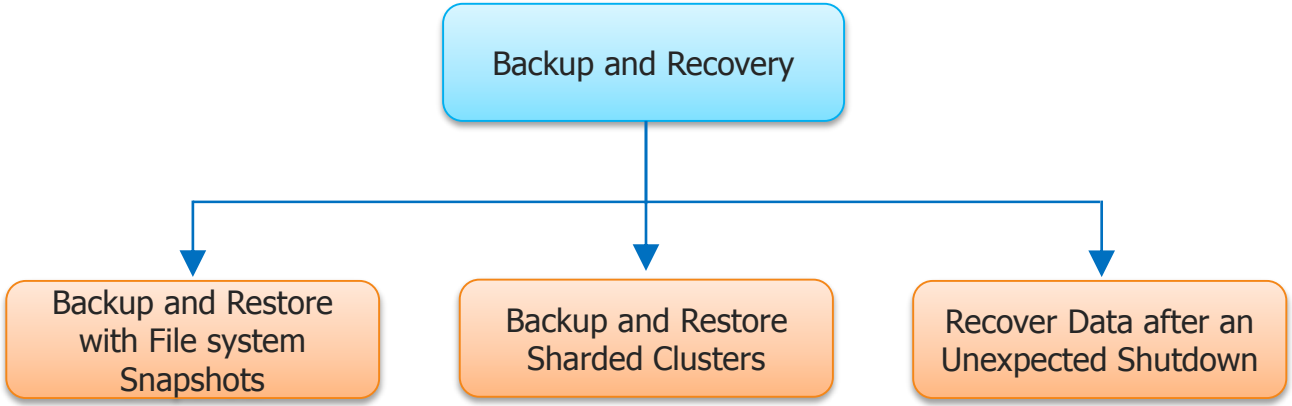
Administration Tutorials



Administration Tutorials



Administration Tutorials



Annie's Question

Which method you should prefer for Backup ?

Annie's Answer

It depends on scenario, but using mongodump is better option.

Annie's Question

What is Rotate Log Files?

Annie's Answer

By default MongoDB uses a standard approach that archives the current log file and starts a new one.
Using Log rotation MongoDB keeps on appending the same log file only.

Assignment

edureka!

- What is mongodump and mongorestore?
- Mention the ways to take backup for MongoDB® databases?
- How to import and export data in MongoDB®?
- What is the issue with taking backup with mongoexport?
- How to recover database with unexpected shutdown?
- What is rotate logfile in MongoDB®?
- What are the techniques there to monitor MongoDB database?
- What is sharding in MongoDB®?
- What is GridFS in MongoDB®?
- What is replication in MongoDB®?

Pre-work

- Read Module 4 FAQ
- Attempt Module 4 Quiz
- Complete assignment

Agenda for Next Class

edureka!

- → Replication Concepts
- → Replica Set Tutorials
- → Replication Reference
- → Sharding Introduction
- → Sharding Concepts
- → Sharded Cluster Tutorials
- → Sharding Reference

Survey

edureka!

Your feedback is important to us, be it a compliment, a suggestion or a complaint. It helps us to make the course better!

Please spare few seconds to take the survey after the webinar.

edureka!