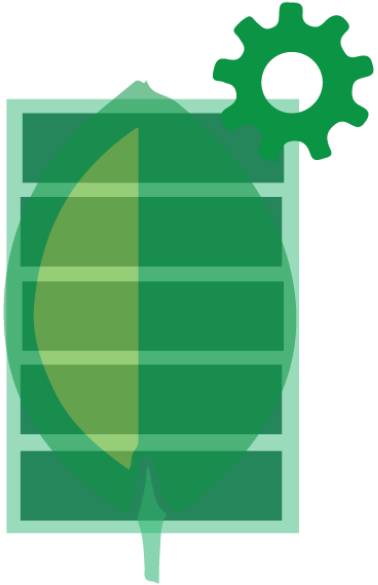


edureka!

MODULE-7
APPLICATION ENGINEERING AND MONGODB® TOOLS

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

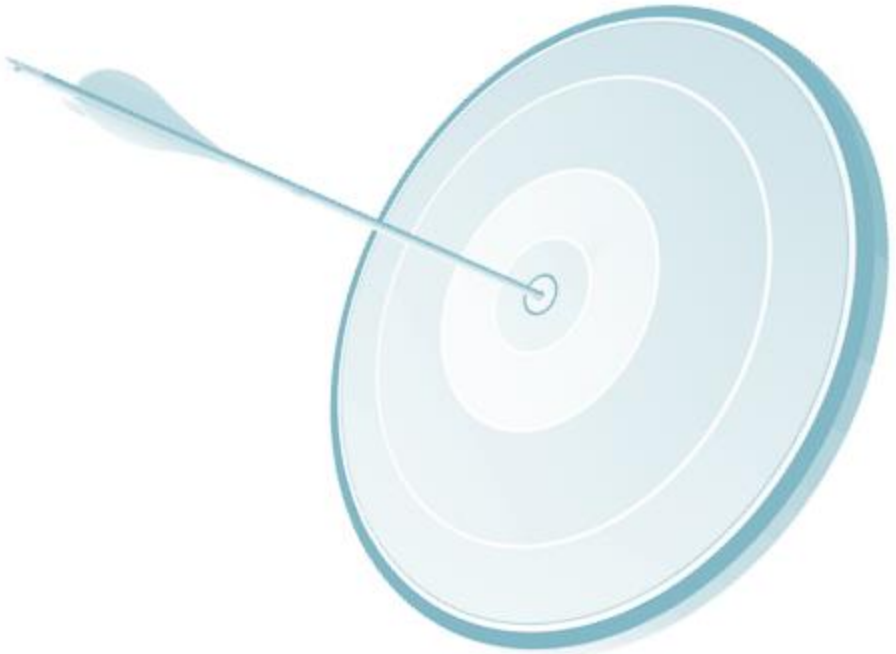
→ Use MongoDB® tools

→ Understand Jaspersoft integration with MongoDB ®

→ Know other popular use case of MongoDB ®

→ Know various API drivers available for MongoDB®

→ Run few query operators through http interface



Annie's Question

How can we create Index?

Annie's Answer

>db.COLLECTION_NAME.ensureIndex({KEY:1})

Annie's Question

Name the types of Index in MonogoDB®?

Annie's Answer

1.Single Field
2.Compound Index
3.Multikey Index
4.Geospatial Index
5.Text Indexes
6.Hashed Indexes

Annie's Question

What are the properties of Indexes in MonogoDB®?

Annie's Answer

1.TTL Indexes
2.Unique Indexes
3.Sparse Indexes

Annie's Question

What is TTL Index in MonogoDB®?

Annie's Answer

TTL indexes are special indexes that MongoDB can use to automatically remove documents from a collection after a certain amount of time. This is ideal for some types of information like machine generated event data, logs, and session information that only need to persist in a database for a limited amount of time.

Annie's Question

What is Unique Index in MonogoDB®?

Annie's Answer

A unique index causes MongoDB to reject all documents that contain a duplicate value for the indexed field.

Annie's Question

What is Sparse Indexes?

Annie's Answer

Sparse indexes only contain entries for documents that have the indexed field. If any document doesn't have the indexed field, that will be not be present in sparse index also.

Annie's Question

How many types of aggregation are there in MonogoDB®?

Annie's Answer

There are 3 types of aggregation in MongoDB they are:
Aggregation Pipeline
Map Reduce
Single Purpose Aggregation

Annie's Question

Can aggregation use Indexes to improve performance?

Annie's Answer

Yes, the aggregation pipeline can use indexes to improve its performance during some of its stages. In addition, the aggregation pipeline has an internal optimization phase.

Annie's Question

What is aggregation pipeline in MongoDB?

Annie's Answer

The MongoDB aggregation pipeline starts with the documents of a collection and streams, the documents from one pipeline operator to the next to process the documents.

Annie's Question

How aggregation pipeline works in MongoDB?

Annie's Answer

In MongoDB, the aggregate command operates on a single collection, logically passing the entire collection into the aggregation pipeline.

Annie's Question

What is the constraint with Map Reduce in MongoDB?

Annie's Answer

When returning the results of a map reduce operation inline, the result documents must be within the BSON Document Size Limit, which is currently 16 megabytes.

Annie's Question

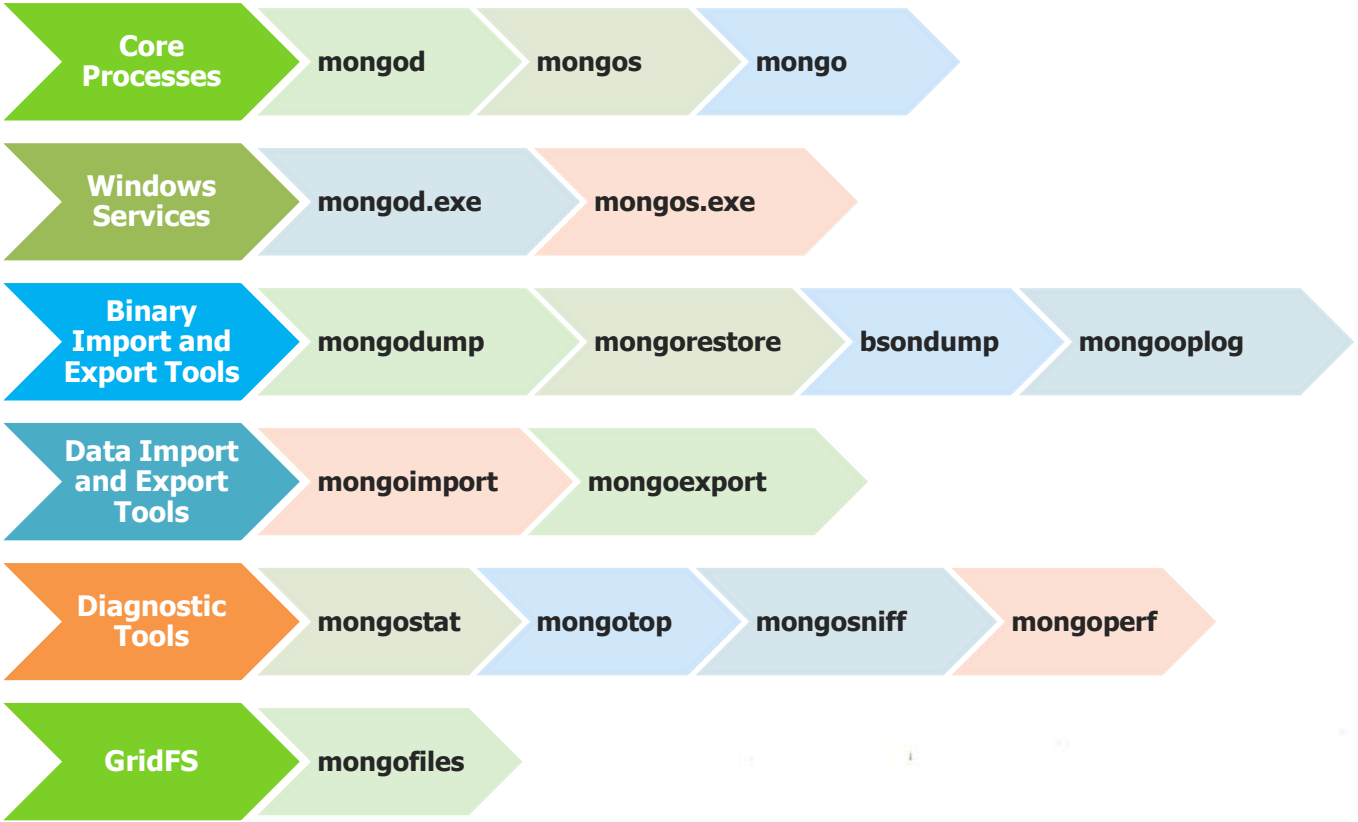
What are common examples of single purpose aggregations?

Annie's Answer

Count, Distinct and Group are common examples.

MongoDB Package Components (Tools)

edureka!



Configuration File Options

edureka!

→ Administrators and users can control mongod or mongos instances at runtime either directly from mongod's command line arguments or using a configuration file.

→ While both methods are functionally equivalent and all settings are similar, the configuration file method is preferable.

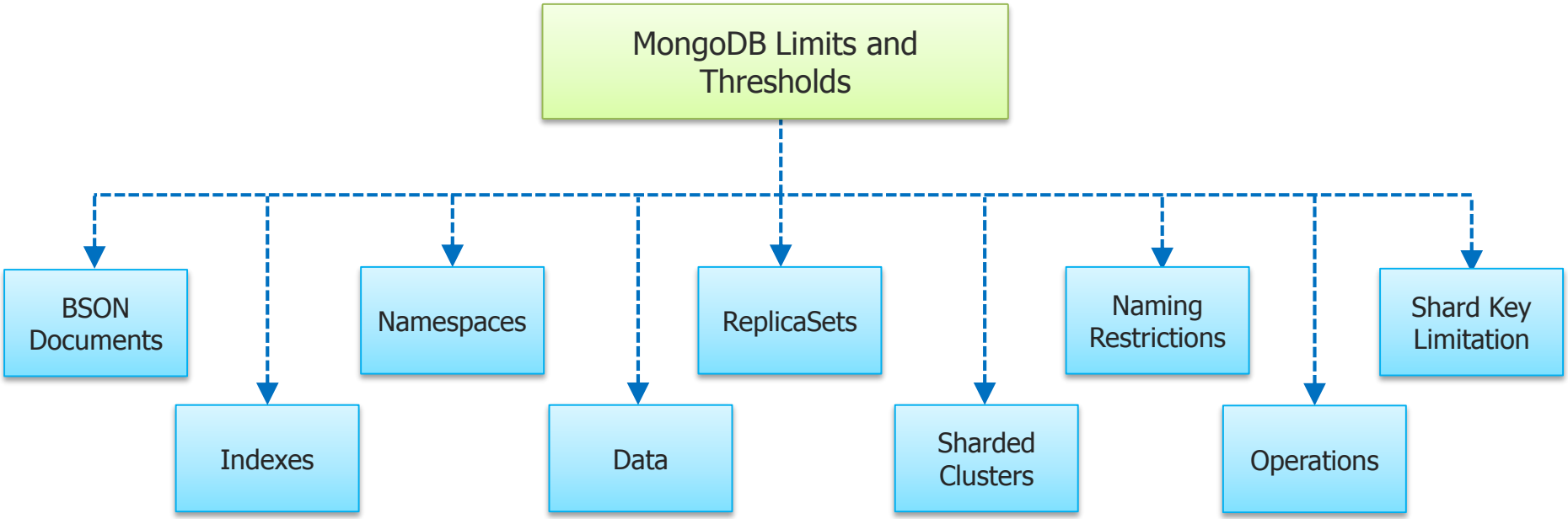
→ Ensure the configuration file uses ASCII encoding. mongod does not support configuration files with non-ASCII encoding, including UTF-8.

→ To start mongod or mongos using a config file, use one of the following forms:

```
mongod --config /etc/mongodb.conf

mongod -f /etc/mongodb.conf

mongos --config /srv/mongodb/mongos.conf

mongos -f /srv/mongodb/mongos.conf
```

MongoDB Limits and Thresholds

edureka!



Connection String URI Format

edureka!

→ In the field of computer networking, a URI scheme is the top level of the Uniform Resource Identifier (URI) naming structure.

→ URI schemes are frequently and incorrectly referred to as "protocols", or specifically as URI protocols or URL protocols, since most were originally designed to be used with a particular protocol, and often have the same name.

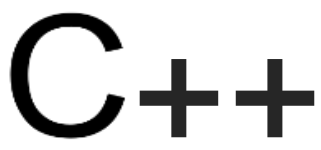
→ The http scheme, for instance, is generally used for interacting with web resources using HyperText Transfer Protocol. Today, URIs with that scheme are also used for other purposes, such as RDF resource identifiers and XML namespaces, that are not related to the protocol.

→ Internet standard STD 66 (also RFC 3986) defines the generic syntax to be used in all URI schemes. Every URI is defined as consisting of four parts, as follows:

```
<scheme name> : <hierarchical part> [ ? <query> ] [ # <fragment> ]
```

API and Drivers for MongoDB

edureka!



HTTP Interface

edureka!

→ To operate MongoDB through HTTP protocol, start your mongod database with --rest option

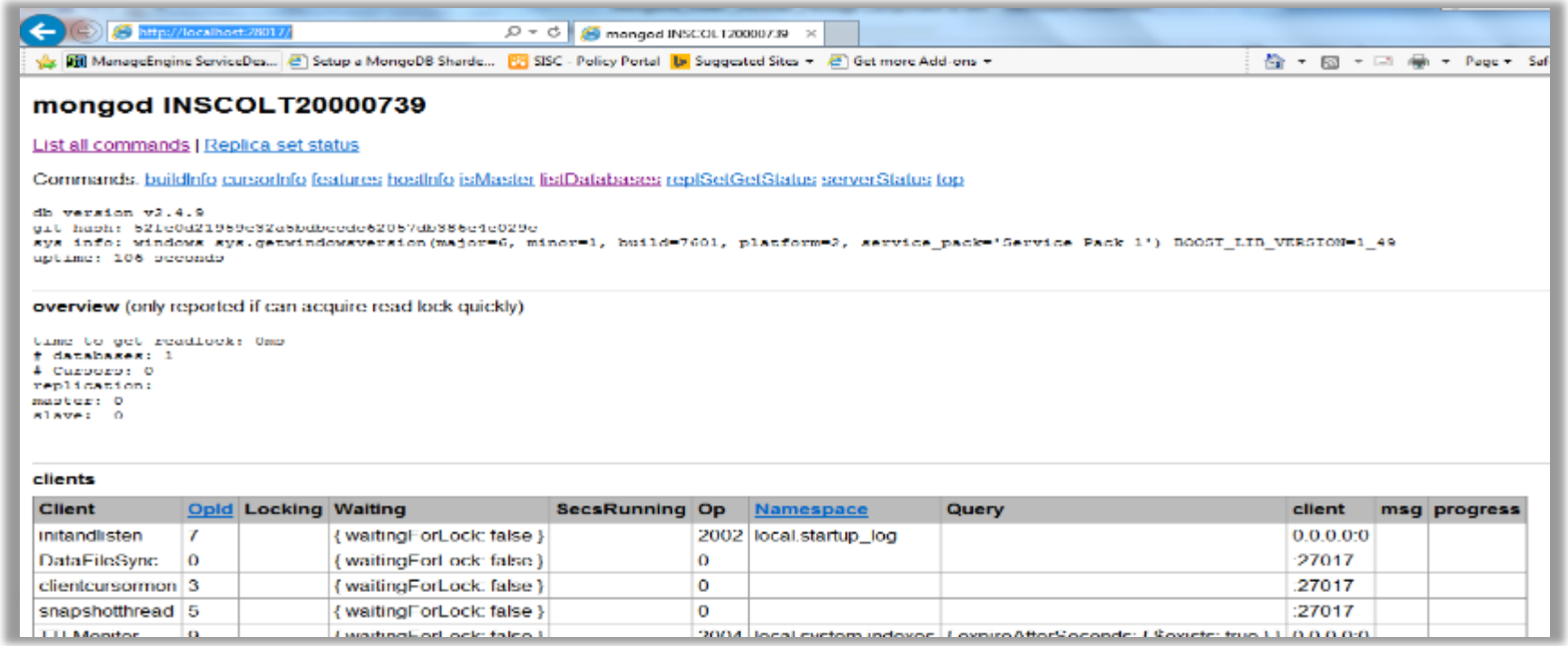
→ >mongod --rest

```
D:\mongodb\bin>mongod --rest
Mon Mar 24 22:22:26.305 [initandlisten] MongoDB starting : pid=7084 port=27017 d
bpath=\data\db\ 64-bit host=INSCOLT20000739
Mon Mar 24 22:22:26.307 [initandlisten] db version v2.4.9
Mon Mar 24 22:22:26.308 [initandlisten] git version: 52fe0d21959e32a5bdbecdc6205
7db386e4e029c
Mon Mar 24 22:22:26.308 [initandlisten] build info: windows sys.getwindowsversio
n(major=6, minor=1, build=7601, platform=2, service_pack='Service Pack 1') BOOST
_LIB_VERSION=1_49
Mon Mar 24 22:22:26.308 [initandlisten] allocator: system
Mon Mar 24 22:22:26.309 [initandlisten] options: { rest: true }
Mon Mar 24 22:22:26.339 [initandlisten] journal dir=\data\db\journal
Mon Mar 24 22:22:26.340 [initandlisten] recover : no journal files present, no r
ecovery needed
Mon Mar 24 22:22:26.457 [initandlisten] waiting for connections on port 27017
Mon Mar 24 22:22:26.457 [websvr] admin web console waiting for connections on po
rt 28017
```

HTTP Interface (Contd.)

edureka!

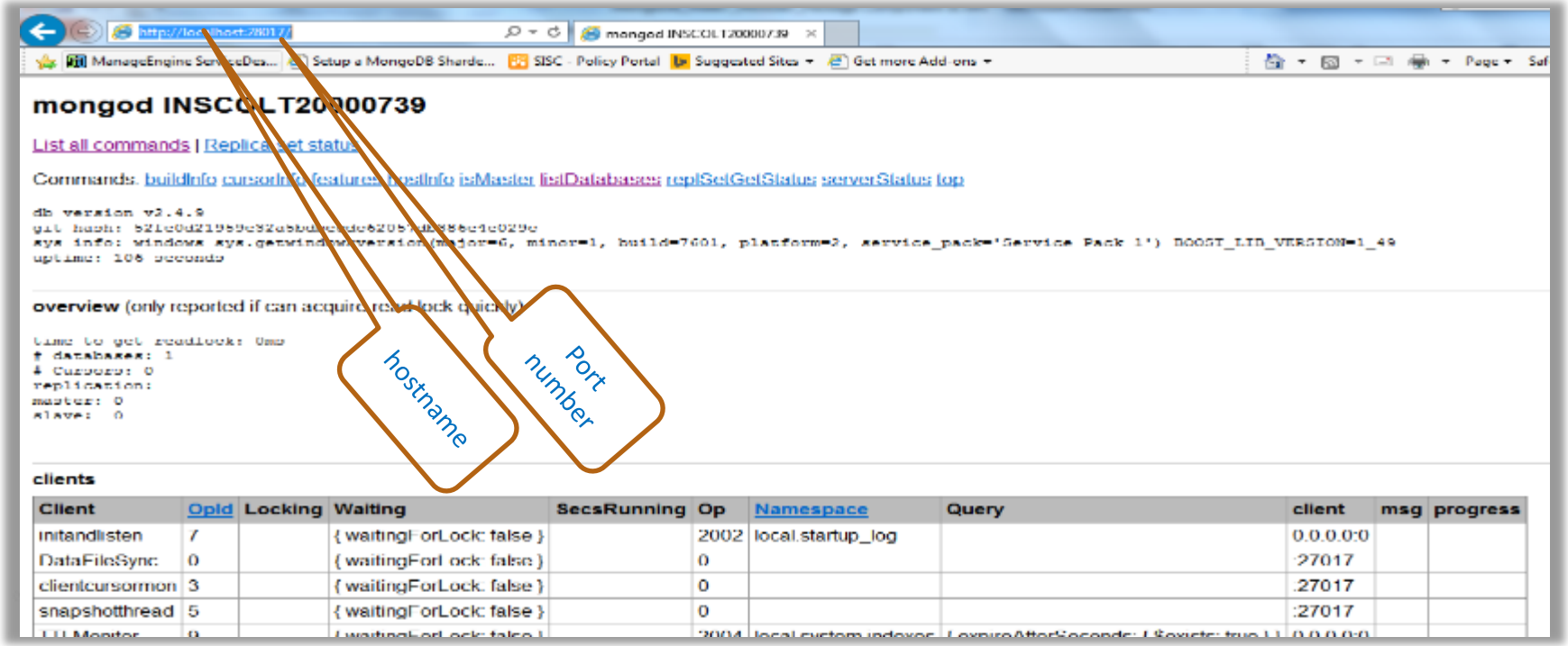
→ In browser use this URL http://localhost:28017/ (please node the trailing "/")



HTTP Interface (Contd.)

edureka!

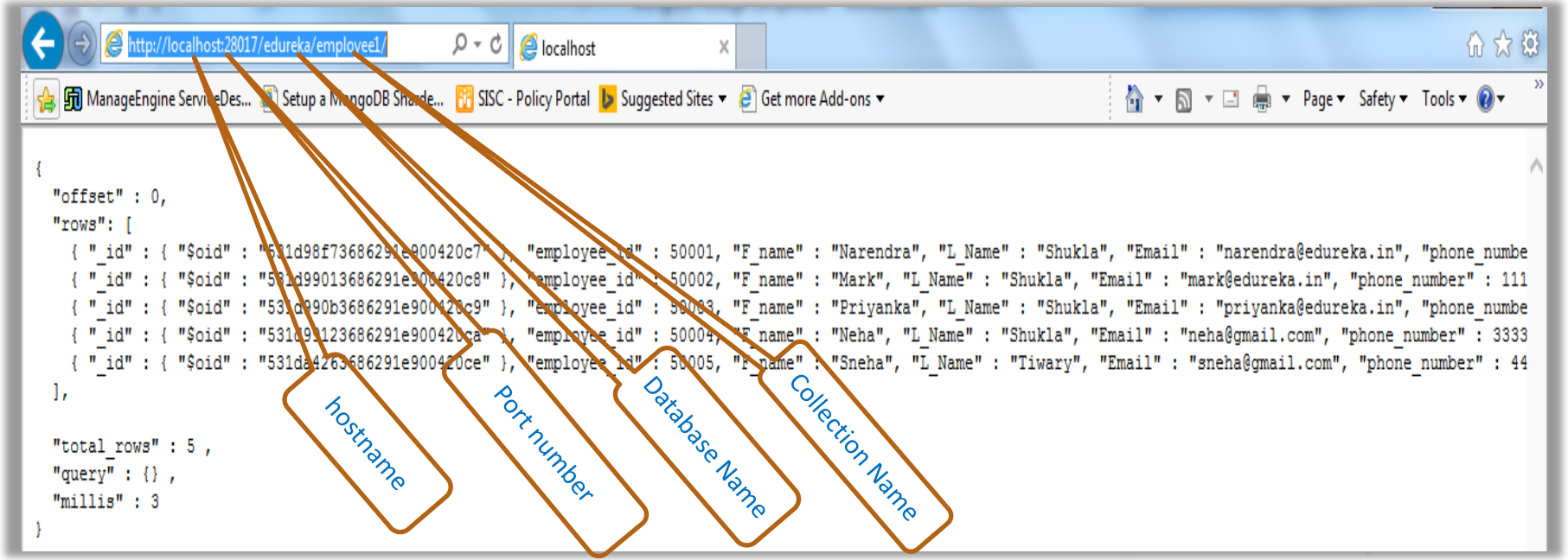
→ In browser use this URL "http://localhost:28017/"



HTTP Interface (Contd.)

edureka!

→ In browser use this URL "http://localhost:28017/"



Package Component and API Hands On

Annie's Question

What is the maximum size of namespace, database and collection name?

Annie's Answer

Must be shorter than 123 bytes.

Annie's Question

What is the maximum length of names of indexes, including their namespace (i.e. database and collection name)?

Annie's Answer

Cannot be longer than 125 characters.

Annie's Question

What is the maximum number of fields that can be there in compound index?

Annie's Answer

31

Annie's Question

Can we use group operators in sharded cluster?

Annie's Answer

No, The group does not work with sharding. Use Map-Reduce instead.

Annie's Question

What is the maximum size of a shard key?

Annie's Answer

Shard key cannot exceed 512 bytes.

Annie's Question

Can we change a shard key after sharding the collection?

Annie's Answer

You cannot change a shard key after sharding the collection.

Annie's Question

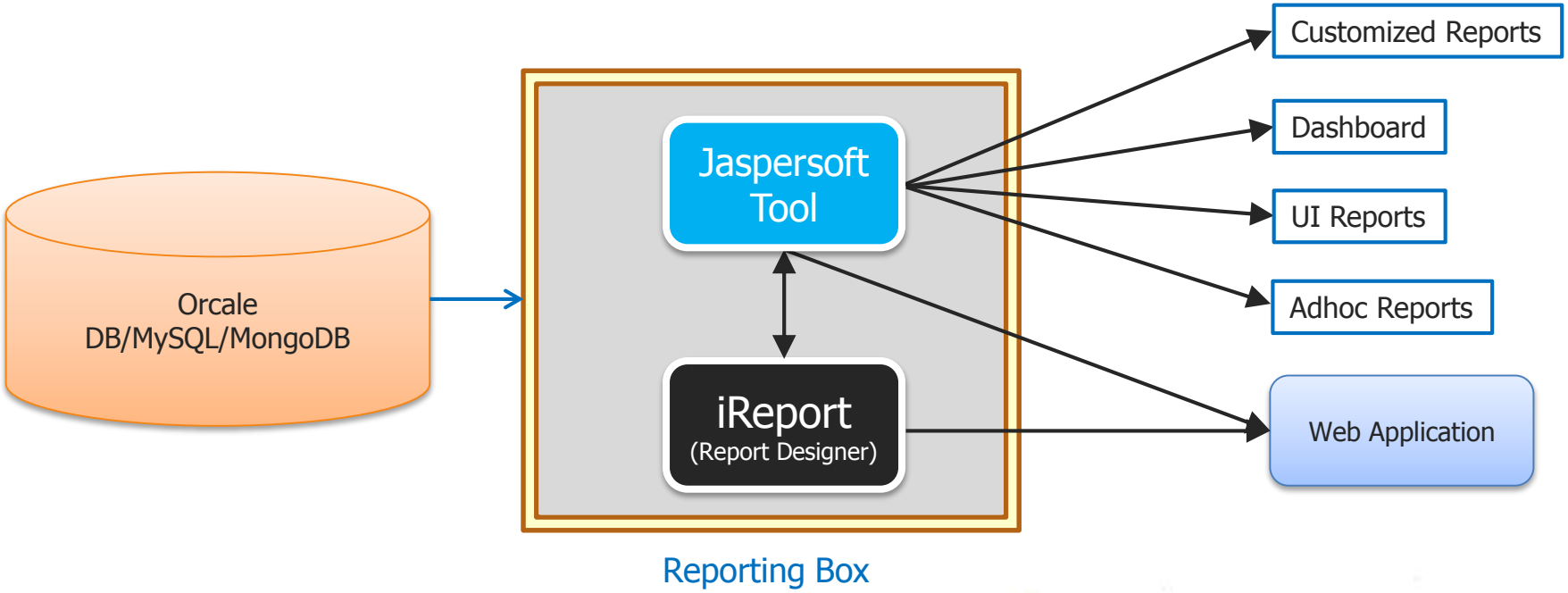
What is the maximum length of database name in MongoDB?

Annie's Answer

Database names cannot be empty and must have fewer than 64 characters.

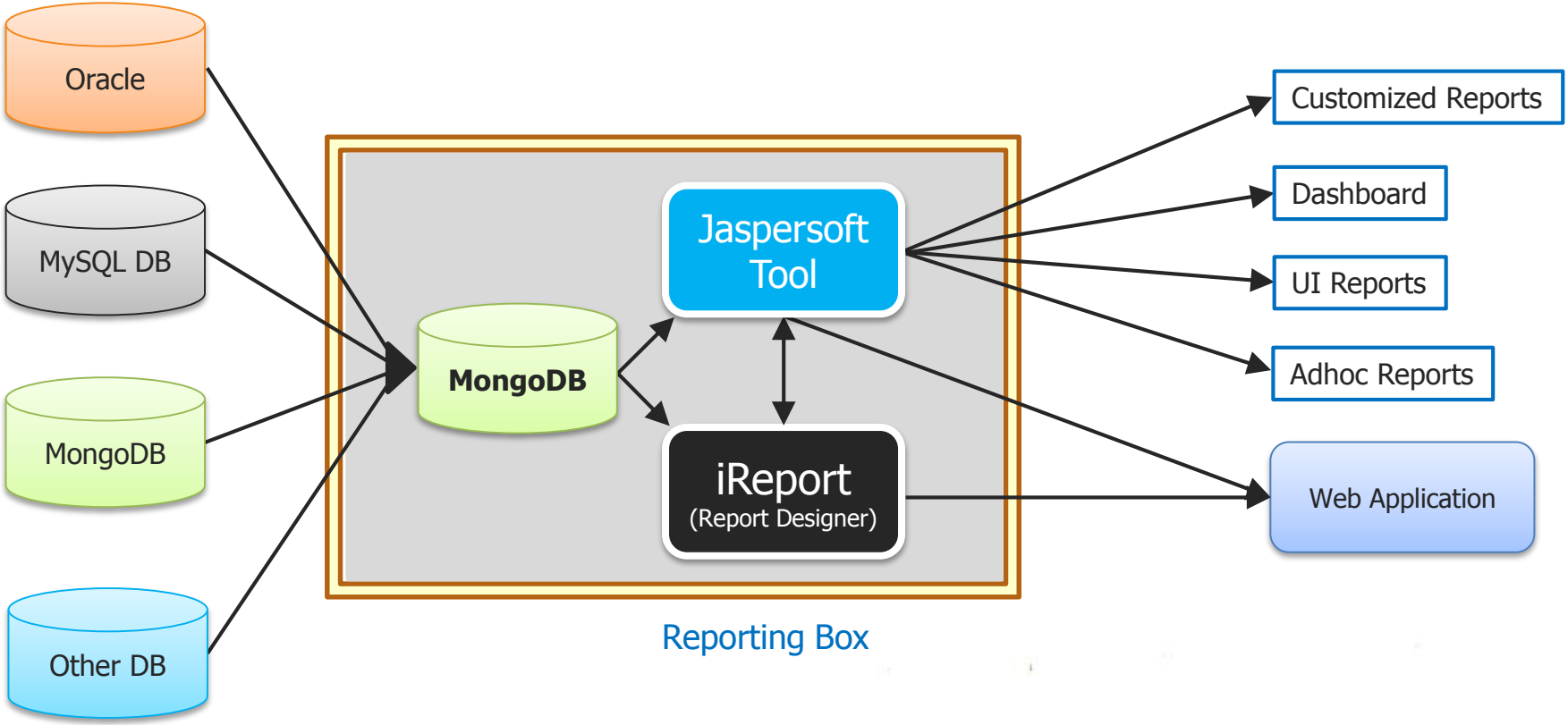
Integration of Analytics and Reporting (Scenario 1)

edureka!



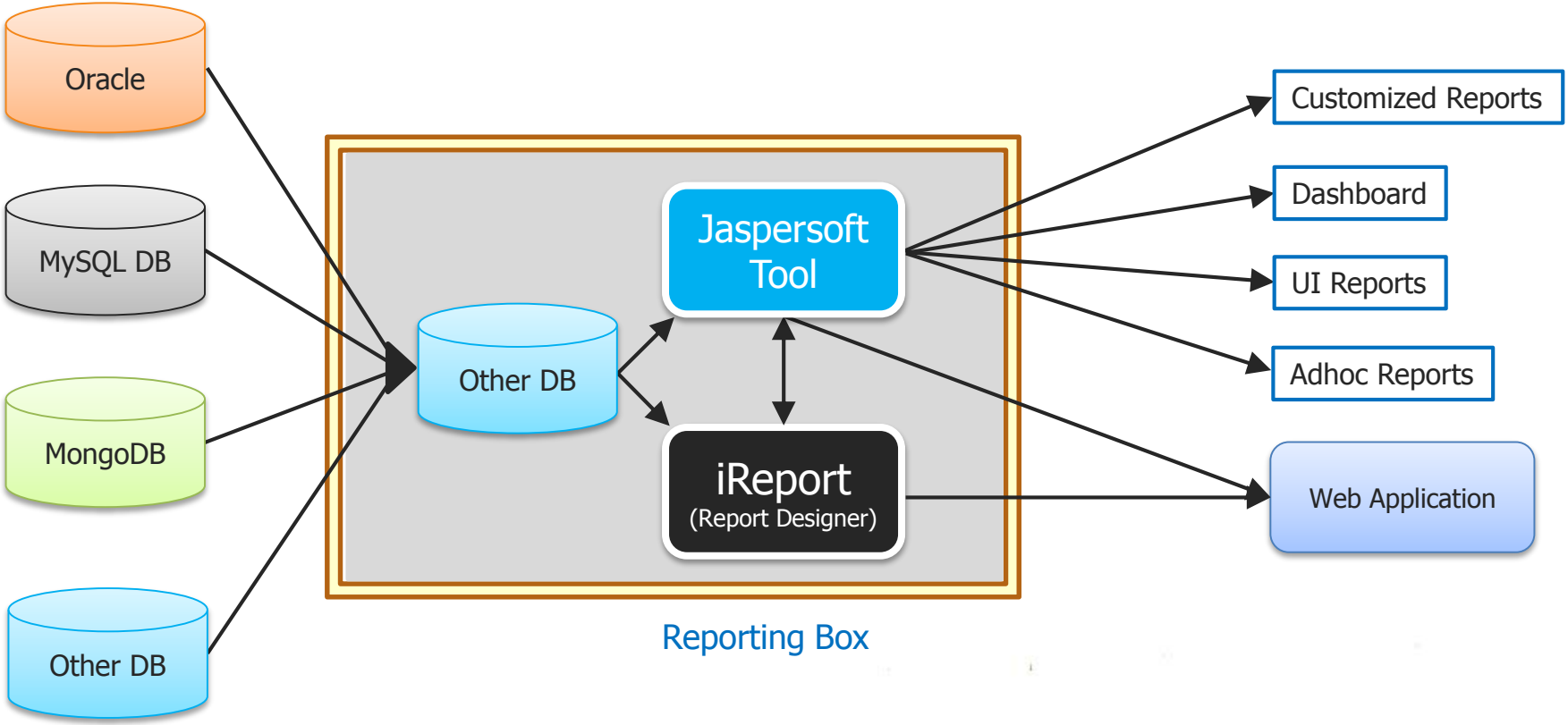
Integration of Analytics and Reporting (Scenario 2)

edureka!

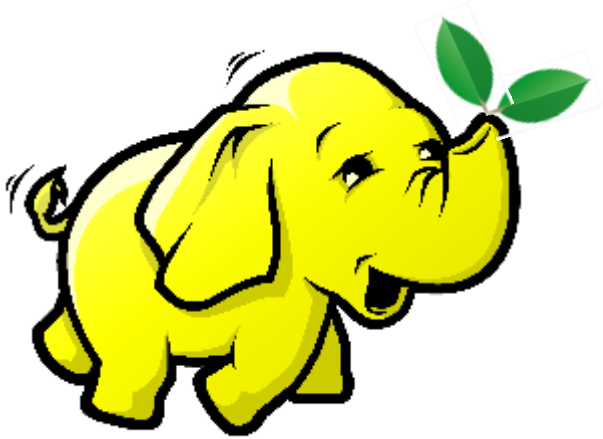


Integration of Analytics and Reporting (Scenario 3)

edureka!



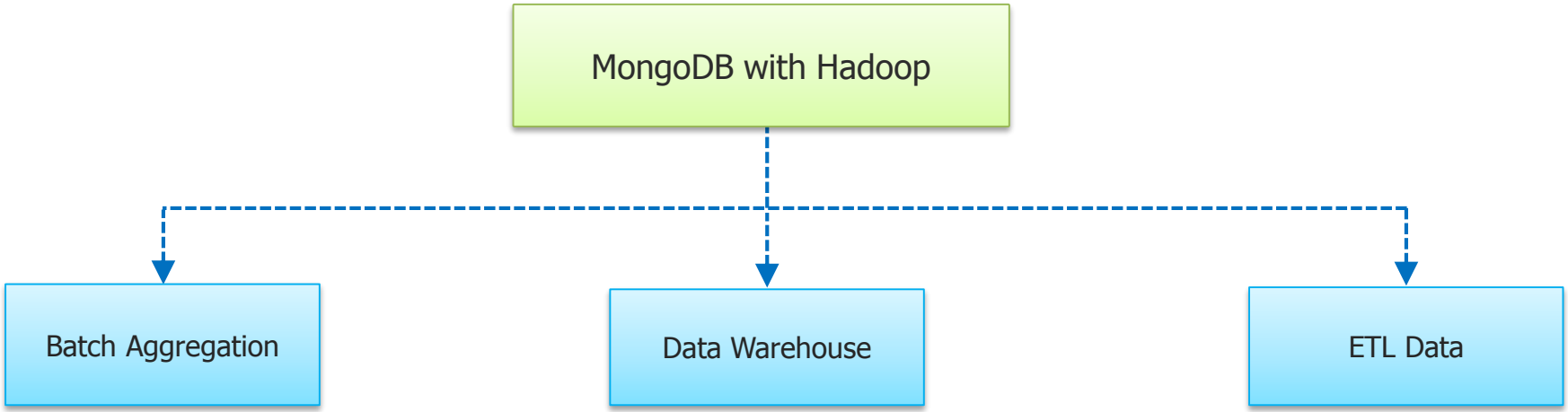
MongoDB with Hadoop



- There are some examples where MongoDB is used as the "operational" real-time data store and Hadoop is used for offline batch data processing and analysis.
- MongoDB & Hadoop can provide a high-level description and can fit together in a typical Big Data stack.

MongoDB with Hadoop (Contd.)

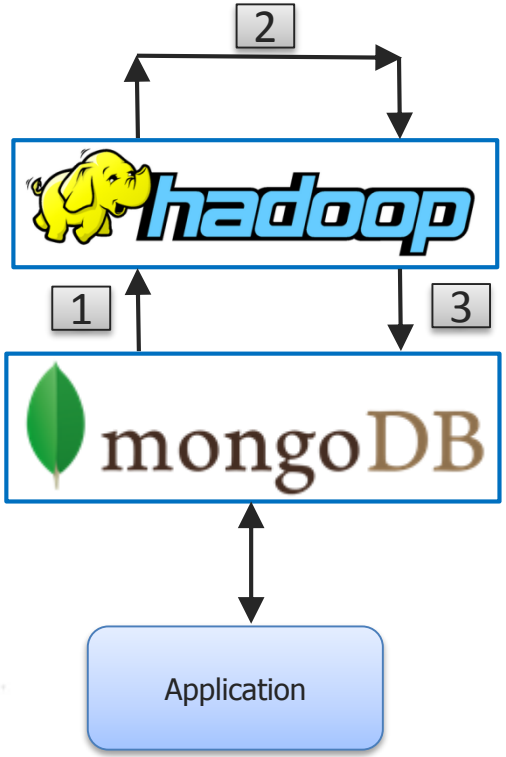
edureka!



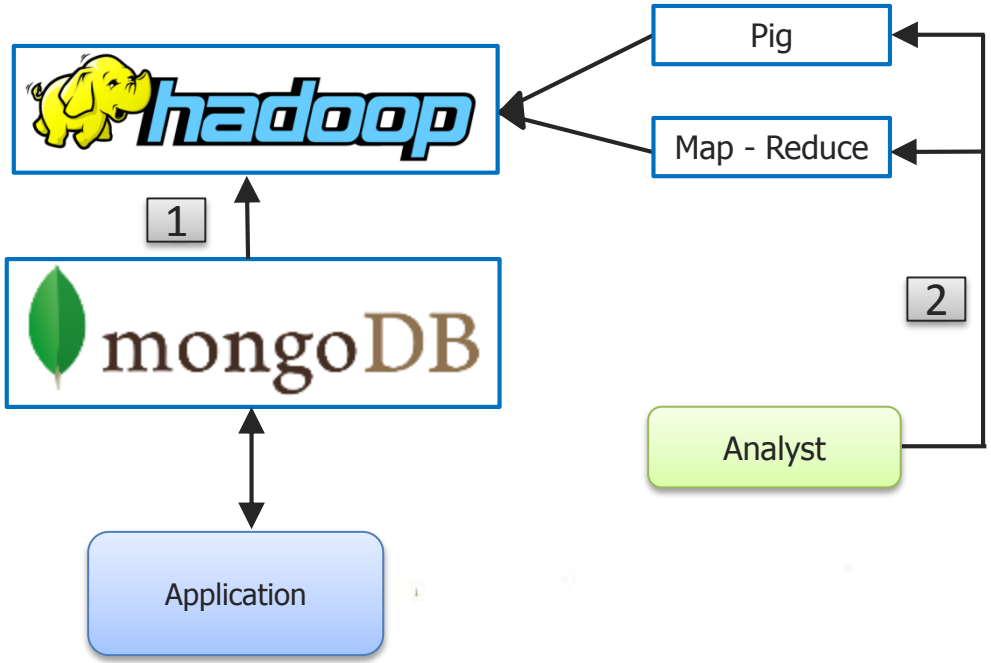
Batch Aggregation and Data Warehouse

edureka!

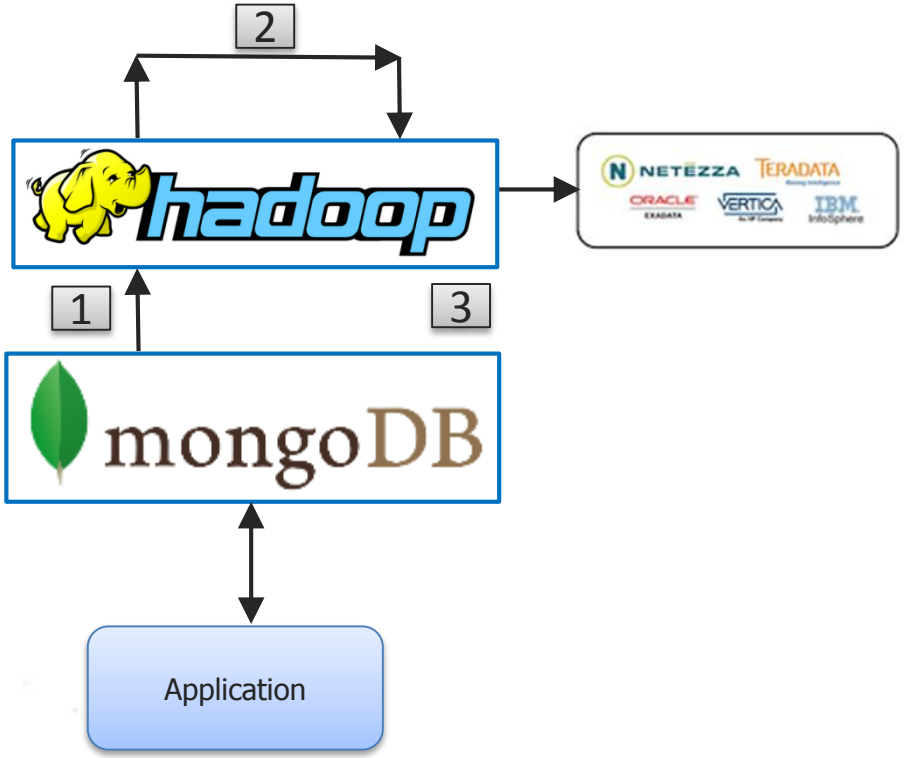
Batch Aggregation



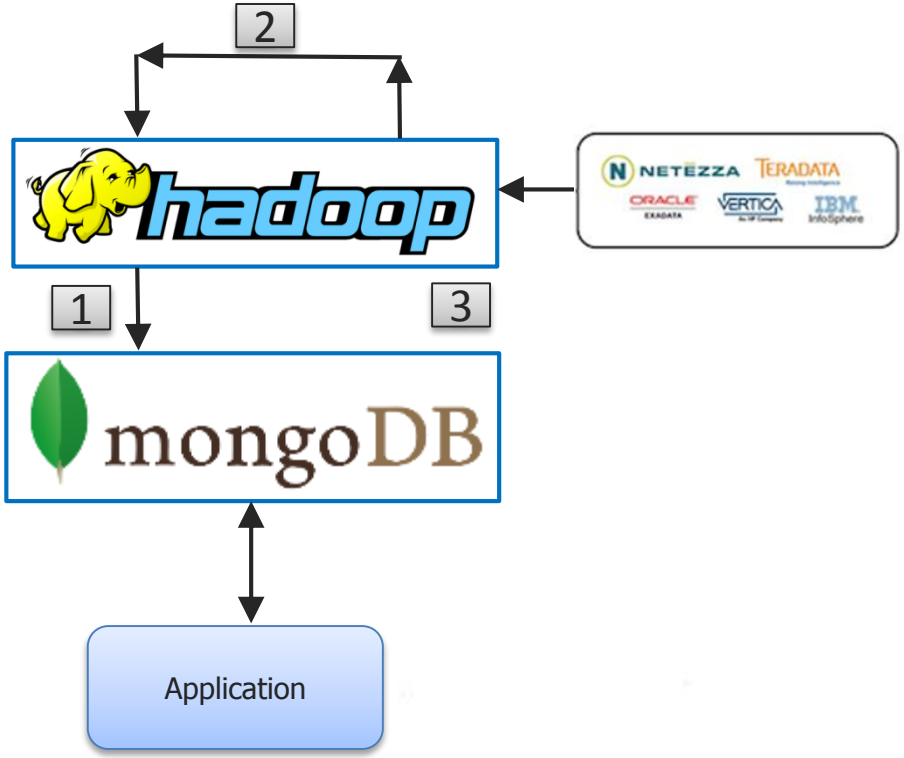
Data warehouse



ETL from MongoDB



ETL to MongoDB



Assignment

edureka!

→ Attempt the assignment using document present in LMS.

Pre-work

- Read Module 7 FAQ
- Attempt Module 7 Quiz
- Do R & D about Jasepersoft R

Agenda for Next Class

edureka!

- → Advance Security
- → MongoDB Methods
- → Course Project
- → Case Studies

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!