Name: Shubham Chemate
Roll No: 31118

PICT, PUNE

PBSL Lab

Assignment 03 (Group B)

Title: MongoDB: map-reduces operations

Problem-statement:

Implement map-reduce operations with suitable examples in mongoDB

Objectives:

1) To understand the concept of map-reduce
2) To apply map-reduce as a data-processing paradigm for condensing large volumes of data into useful aggregate methods.

Outcomes:

1) Learn about concept of map-reduce in mongoDB.
2) The application & execution of map-reduce to condense large volume of data into aggregate result.

Software requirements:

1) mongoDB updated to latest version
2) mongosh updated to latest version
3) Operating system - windows 10 (64 bit)

Hardware requirements

1) Intel core-i5 8th gen 4-core processor
2) 8GB RAM, 512 GB SSD storage

Theory Related concepts:

PICT, PUNE

Map-reduce in mongoDB:

- map-reduce is a data-processing paradigm, for condensing large amount of data into useful aggregate results.
- to perform map-reduce operations, mongoDB provides the map-reduce database command.

map-reduce command:

Syntax:

```
db.collection.mapReduce(
    function() {emit(key,val);}   //map function
    function(key,values)
      {return reducefunction},
    {
      out: collection,
      query: document,
      sort: document,
      limit: number
    }
)
```

- the map reduce operation first queries the collection, then maps the result documents to emit key-value pairs, which is then reduced based on the keys that have multiple value.
- In the above syntax:
  1) map function: It is javascript function that maps a value with key & emits a key-value pair.
  2) reduce function: It is JS function that reduces or groups all documents having

PICT, PUNE

same key

3) out: It specifies the location of the map-reduce query result.

4) query: It specifies the optional selection criteria for selecting documents.

5) sort: It specifies the optional sort criteria

6) limit: It specifies the optional maximum number of documents to be returned

- In mongoDB, the map-reduce operation can write results to a collection or return the result inline.
- If the map-reduce output is written in a collection (using 'out' parameter), then we can view the result using find operator.

syntax: db.map_reduce_output.find()

where

'map_reduce_output' is the name of the collection to which map-reduce operation written the result.

- The map-reduce queries can be used to construct large complex aggregation queries. The use of custom JS function makes use of map-reduce which is very flexible & powerful.

Test cases

The code files & output screenshots are attached separately.

PICT, PUNE

Conclusion:

In this assignment, I learned the concept of map-reduce & implemented it to condense large volume of data into aggregate results in a collection.