ASSIGNMENT – B4

TITLE :-

Map reduce operation in Mongo DB

PROBLEM STATEMENT :-

Write an example of map reduce using Mongo DB

THEORY

→ Map - Reduce is a data preprocessing algorithm for condensing large volumes of data into useful aggregated units
→ It can solve some problems that are too complex to express using the aggregation framework
→ It is fairly slow and should not be used for real time data analysis
→ Map reduce can be easily parallelized across multiple servers

Map Phase :-

Maps an operation onto every document in a collection
The operation could be either 'do nothing' or 'emit' these keys with X values.

Keys are grouped & lists of emitted values are created for each key.

Reduce phase :-

Takes list of values and reduces it to a single element
This element is returned to the shuffle step until each key has a list containing a single value the result

Syntax :-

```
db.collection.mapReduce(
    function() { emit(key, value); };
    function(key, value) {
        return reducedFunction
    },
    {
        out : collection,
        query : document,
        sort : document,
        limit : number
    }
)
```

→ map is a JS function that maps a value with a key & emits key value pair

→ reduce is a JS function that groups all documents having same key

→ out specific location location of query result.

→ query specifies optional selection criteria for selecting documents

→ sort specifies optional sort criteria

→ limit specifies optional max number of documents to be returned.

CONCLUSION :-

We have successfully understood & implemented mapReduce() in MongoDB.