Nikit Gokhe

Class – Comp D1

Roll No. 224024

GR No. 21810522

**ASSIGNMENT NO.09**

**Aim**:

Implement Map reduces operation with suitable example using MongoDB.

**Objective**:

To understand and implement map reduce.

**Theory**:

Collection: The data goes through the following phases

Input Splits:

An input to a MapReduce job is divided into fixed-size pieces called input splits Input split is a chunk of the input that is consumed by a single map

Mapping

This is the very first phase in the execution of map-reduce program. In this phase data in each split is passed to a mapping function to produce output values. In our example, a job of mapping phase is to count a number of occurrences of each word from input splits (more details about input-split is given below) and prepare a list in the form of <word, frequency>

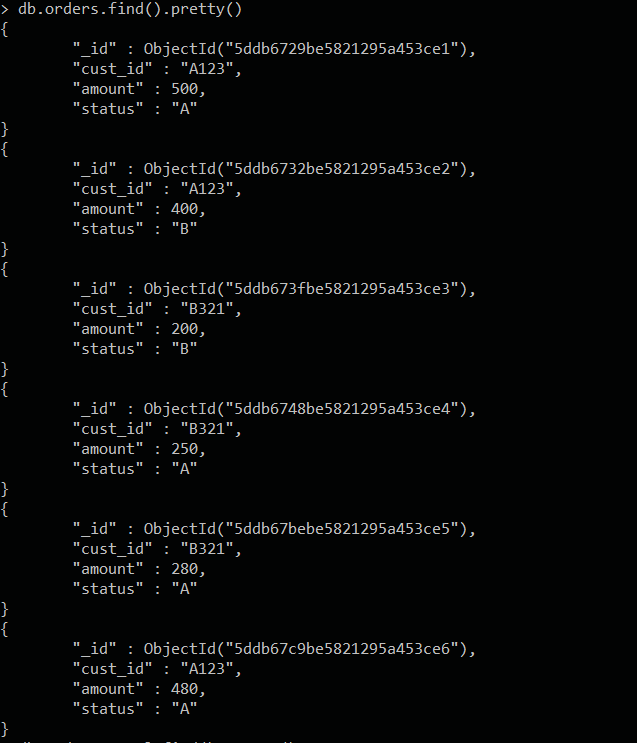
Shuffling

This phase consumes the output of Mapping phase. Its task is to consolidate the relevant records from Mapping phase output. In our example, the same words are clubbed together along with their respective frequency.

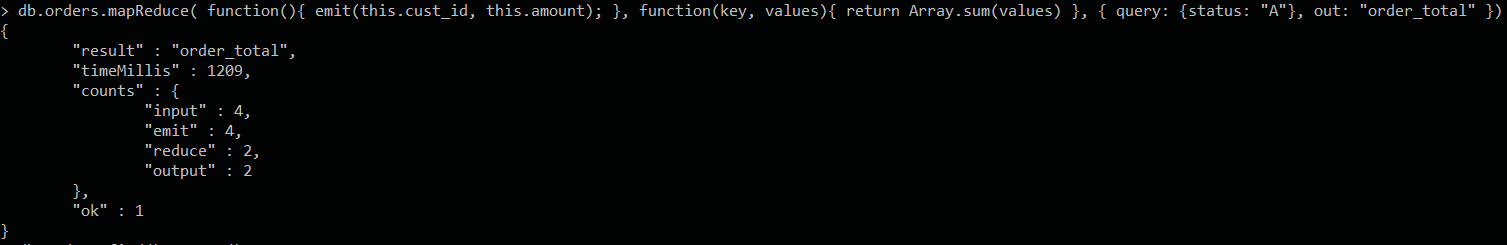
Reducing

In this phase, output values from the Shuffling phase are aggregated. This phase combines values from Shuffling phase and returns a single output value. In short, this phase summarizes the complete dataset.

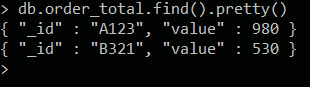
**Input:**



**Implementing map reduce**:



**Output**:



**Conclusion :**

Successfully implemented map reduce on an example schema using MongoDB.