Jonathan Disla

Assignment 7.1

Web 335

Prof. Krasso

MongoDB can analyze data natively using the aggregation framework. The aggregation framework has many different stages. The stages in the aggregation framework are organized like a pipeline. According to the text, we take input from a MongoDB collection and pass the documents from that collection through one or more stages, each of which performs a different operation on its inputs (Shannon, Brazil, Chodorow, 161). The stages are what transform the data which can then be projected or visualized. The documents can also be manipulated within each stage.

Within the different stages such as $match, $project, and $limit which reduce the amount of data through each stage, we then project or output the data to a person or collection (Terpko). First however, we must use the aggregate function. Here is an example:

db.companies.aggregate([

{$match:{…}},

{$project: {...}}

])

In the example above, we tell mongo the aggregated data we want matches a certain parameter and then project the results. The $project stage reads and passes less data to the next stage thus requiring less RAM or processing resources. Within the $project stage, the fields specified can be existing fields from the input documents or newly computed fields (MongoDB Documentation). This stage tells Mongo the fields required and only those fields will be optimized.

As an example we can we can use the default value of one to retrieve all of the datat matching a criteria. I will write an example of a jobs collection where we will match the imported jobs and project the time the jobs take to process.

db.jobs.aggregate([

{$match:{“type”: “import”}},

{$project: {cluster:1, type:1, seconds:1, \_id:0}}

])

Reference List

Bradshaw Shannon, Braxil Eoin & Kristina Chodorow. “MongoDB The Definitive Guide Third Edition.” O’Reilly Media, Inc. December 2019.

mongoDB Documentation. “$project (aggregation)”. MongoDB.com, Accessed 26 April 2021.

<https://docs.mongodb.com/manual/reference/operator/aggregation/project/>

Terpko Jason. “Aggregation Pipeline”. Percona YouTube, 15 November 2017

https://www.youtube.com/watch?v=F81Qu6Xbhq0&ab\_channel=Percona