Jonathan Roland

WEB 335 Introduction to NoSQL

Instructor: Prof. Krasso

11/27/2020

Discussion 7.1

1. **What are MongoDB pipelines?**

MongoDB Pipelines are similar to Linux pipelines in that they describe a chain of operations that contribute to a final output. In MongoDB, these operations are performed on the documents in a given collection. They expect a specific input and provide a processed document (or stream of documents) as an output.

One comparison would be a factory assembly line. The factory takes in a sheet of metal, the first station stamps it into a car panel, the second station trims the edges, the third station drills holes, the fourth station preps it for paint et cetera. This factory takes in a stream of metal sheets of an expected type and size, and outputs a stream of completed car body panels.

1. **How are pipelines used to perform complex operations?**

Each individual stage of a pipeline is called a data processing unit. Each data processing unit takes in, processes, and outputs a single document at a time. Each of these stages contains tunables, or parameters that we can set to perform actions like filtering, arithmetic, or aggregation. Sometimes we may even need to feed result data back through for another pass.

This is extremely similar to an effects chain for an electric guitar. The guitar’s signal enters the chain and is processed in serial order by the effects pedals. Each pedal has several knobs to control parameters for that type of effect. A simple chain might go like this:

Guitar->distortion pedal with distortion knob set to 75%->chorus pedal with depth at 50% and rate at 20%->reverb pedal with ‘special’ mode switch on and depth at 60%->amplifier with equalization settings dialed in.

This turns a simple clean signal from the guitar into a big, washed out sound with lots of harmonic overtones.

1. **What are MongoDB projections ($project)?**

Projections are a type of dot notation used to specify nested fields. Because documents most often contain nested hierarchies, these must be referred to by their hierarchical position to utilize them for searching and aggregation operations. When using an aggregation function, the $project parameter can be given a projection document to describe nested fields we wish to output.

1. **Provide an example of when and how you would use MongoDB projections ($project).**

Assume a collection named “people” and a document data structure like so:

{

“\_id”:12345,

“first\_name”: “Jonathan”,

“last\_name”: “Roland”,

“location”: {

“city”:”Knoxville”,

“state”:”Tennessee”,

“zip\_code”:37931

}

}

If we wanted to return the example documents first name, last name and state, we would use a projection to reach the locations nested “state” value:

db.people.aggregate([

{$match:{“\_id”:12345}},

{$project:{

\_id: 0,

first\_name: 1,

last\_name: 1,

state: “location.state”

}}

]).pretty();

This will return the following:

{

“first\_name”:”Jonathan”,

“last\_name”:”Roland”,

“state”:”Tennessee”

}

**Reference:**

Bradshaw, S., Brazil, E., & Chodorow, K. (2019). *MongoDB: The Definitive Guide: Powerful and Scalable Data Storage* (3rd ed.). O’Reilly Media.