Jonathan Roland

WEB 335 Introduction to NoSQL

Instructor: Prof. Krasso

11/19/2020

Discussion 6.1

1. **What are indexes and how does MongoDB select them?**

Indexes are lists of a deliberately pre-gathered type of data (by a specified key) in a MongoDB collection. They allow for faster queries of the specified data in order to avoid scanning the entire collection to find a match. While this greatly improves query speed, it will slow write operations on the same data since the change must be reflected in any indexes that reference it.

Indexes are only deliberately created by the user by using the createIndex() function on a collection. Once created, any query on the indexed key will utilize the index rather than perform a collection scan. This does not require any adjustment of syntax to perform.

1. **What are compound indexes and when should you use them?**

Standard indexes only gather a single set of values based on a common key. We often need more than that for our queries, such as gathering the first and last name of all users with a particular job title. Compound indexes allow us to nominate multiple sets of values to be stored together in order to speed up multi-value queries. We can achieve this very simply by giving the createIndex() function multiple field keys.

While this does speed up multi-value queries, it should be noted that some forethought is required when creating a compound index. If dealing with large volumes of data, create your compound index to be ordered in the same way as you would query it (which field is first, starting at min vs starting at max). This will help to stave off memory limit issues that can be caused by sorting a return in memory. By structuring the compound index to match a common query case, the index will stay updated in the same order the query will search it, reducing computational burden.

1. **What are capped collections?**

Capped collections are sequentially written collections with a fixed size. They do not accept deletions and reject updates that would increase the size of the collection. When a capped collection hits its maximum size, any new insertion causes the deletion of the oldest document to make room for the newest document. This collection type has limited uses, such as a running log or other expiring serial storage case. Capped collections must be created deliberately by passing the “capped” and “size” parameters into the createCollection() function. Capped must be set to true and size specified in bytes. Once created, the capped collection cannot be modified and can only be dropped.

Reference:

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