Project Phase 3

This document is created by Hasan Ertuğrul Çinar

In this project phase, you are going to apply indexing on your tables and analyze their effectiveness. You can check out how to apply indexing to a table by looking at recitation 8 users.sql

First, for each student in a group, you need to pick at least **two** tables to **join.** After choosing tables, populate them, where those tables need to have at least 1 million records. To populate your tables, you can choose any method you like, but it is advised to use some library similar to what is shown in recitation that will create some dummy data for you then you can insert those values through your populator application. After that, you need to write a query that will join these tables and add some filtering based on two different columns.

SELECT *

FROM table1

JOIN table2 ON table1.column1 = table2.column2

WHERE table2.columnx = ... and table1.columny = ...

LIMIT 1000000

Use the method shown in recitation to get your benchmark for execution time, then take clip the results as an image.

Then, apply indexing on two columns that you use as filters in your query. After adding indexes, run the same benchmark operation to record the new timing.

You must repeat this process two times, the same process of recording time difference between indexed and non-indexed versions of your tables. Each time, you must apply indexes on different columns.

In the end, it is expected that you have 2 different benchmarks for your queries on an indexed and non-indexed version, for different columns. Put the results on a report file for your group where each member shares their own benchmarks. Give brief descriptions about what you have done as well. A small paragraph would suffice.

Keep in mind that, MySQL uses a caching algorithm when a query is executed, so the next time It is executed, it will be a lot faster to retrieve the results, so make sure that you take your results in first try. If you want to change that, you may restart your application, you can disable caching for this step.

Next step of the project will be related to NoSQL where you will create some schemas for your given system and work on them through a python application.

Best

Hasan E. Cinar