

UNIVERSITY OF THE WEST INDIES
Department of Computing
COMP6720 – Advanced Database Management Systems - 2021
Lecturer: Dr. Kevin Miller

Problem Definition

The performance of big database systems is very important. When designing an application, caution must be taken in how the database system is designed, as oftentimes, the bottleneck of the application is the database. Hence, this assignment will help you to make the right decisions when designing and creating database systems.

Algorithm analysis is the branch of computer science, that deals with the calculation of the amount of time and the storage requirement that will be needed for an algorithm to perform optimally. This technique involves counting the number of steps (Time Complexity) and also counting the number of storing locations (Space Complexity) that will be used by the algorithm. The benefit of this technique is that we can figure out the performance of algorithms irrespective of the specification of the machine that the algorithm will run on.

Similarly with algorithm analysis, we try to figure out the performance of our queries and we do this by counting the number of records accessed. As a result of this technique, and by using relational algebraic expressions, we can figure out how to arrange these expressions so that our queries will be much more efficient.

In this assignment, students will install and configure a relational database (MySQL) and a NoSQL database (MongoDB) on their system. In these two database systems, the students will create the same databases. Using any computer scripting language, create a script to use an appropriate library to populate your databases. Prepare a technical document, with the requirements below.

PART 1

The requirements are stated below:

- a).** In the relational database, create three (3) tables with at least one shared foreign key among all tables (at least 1,000 records should be related).
- b).** In the NoSQL database, create three collections with similar set up as in **a).**
- c).** Insert 10,000 records in each table in your relational database.
- d).** Insert 10, 000 documents in your NoSQL database.
- e).** Create two queries in relational algebra that requires the joining of all tables using two (2) strategies. By counting the number of rows accessed, calculate the performance of these queries when running on a relational database.
- f).** Develop your own strategy for analyzing the performance of the same query in **e)** using your NoSQL database (*This technique should not depend on your machine*).

PART 2

The student will perform a number of CRUD operations and measure the time taken by each database management system. This data should be tabulated with the support of appropriate graphs to show the comparisons.

- a).** Select and compare the retrieval of 1,000 records, 5,000 records and 10,000 records.
- b).** Delete and compare the time taken to delete of 1,000 records, 5,000 records and 10,000 records.
- c).** Update and compare the time taken to update 1,000 records, 5,000 records and 10,000 records.
- d).** Compare with and without the use of indexes, and at least one different storage engine.
(eg. of storage engine: MyISAM).
- e).** Discuss the following using your results:
 - How does MySQL compares to MongoDB as it relates to :
 - a. Inserting Records
 - b. Selected Records
 - c. Deleting Records
 - d. Updating Records
 - Propose reasons for your observations.
- f).** Discuss the advantages and disadvantages of using a relational database vs a NoSQL database.
Your discussion should include a specific situation in which you would use one over the other.

The document and the script should be uploaded as a zipped file.

Due Date: October 27, 2021 @ 10:00pm

NB: A submission container will be created at the appropriate time.

**** THE END ****