Group Project - Building a simple custom relational database (RDb) and its management system (RDbMS)

Project Title: 5408_RDbMS_Custom(RDb)
Team Size: 3 to 4

Project Objective:

By completing this project, a student will be able to describe concepts in modelling notation (e.g., Entity-Relation Diagrams or UML) and how they would be used. A student will be able to describe the most common designs for core database system components including the query optimizer, query executor, storage manager, access methods, and transaction processor. By completing this task, a student will be able to explain the techniques used for data replication, and allocation during the database design process

In this project, each group is required to create a light-weight relational database management system (e.g. metadata management, data structure design, data storing, retrieval, building database and logs, analysis, and security), with custom database structure for in memory operations using various data structures such as Arrays, Tree, Linked Lists etc., and custom file design for persistent storage.

Plagiarism Policy:

- This project is a group task. Collaboration of any type (outside the assigned group) amounts to a violation of the academic integrity policy and will be reported to the AIO.
- Content cannot be copied verbatim from any source(s). Please understand the concept and write in your own words. In addition, cite the actual source. Failing to do so will be considered as plagiarism and/or cheating.
- The Dalhousie Academic Integrity policy applies to all material submitted as part of this course. Please understand the policy, which is available at:
  https://www.dal.ca/dept/university_secretariat/academic-integrity.html

Timeline and Deliverables:

Phase 1 (Feasibility Study) - (2 %)

- Each group will work on the given problem and identify the required technologies, and list the new learning.
- Each group will record their initial meeting logs and in addition, publish a timeline in the form gantt chart
- Each group will submit their tentative design, and implementation plan as a 3-page written
- document.

Submission: One 3-page report from each group must be submitted on or before Jun 10, 2021 (11:59 pm)

Project Rubric:

Based on the discussion board rubric (McKinney, 2018)

	Excellent (20%)	Proficient (15%)	Marginal (10%)	Unacceptable (0%)
Completeness including Citation	All required tasks are completed	Submission highlights tasks completion. However, missed some tasks in between, which created a disconnection	Some tasks are completed, which are disjoint in nature.	Incorrect and irrelevant
Correctness	All parts of the given tasks are correct	Most of the given tasks are correct However, some portions need minor modifications	Most of the given tasks are incorrect. The submission requires major modifications.	Incorrect and unacceptable
Novelty	The submission contains novel contribution in key segments, which is a clear indication of application knowledge	The submission lacks novel contributions. There are some evidences of novelty, however, it is not significant	The submission does not contain novel contributions. However, there is an evidence of some effort	There is no novelty
Clarity	The written or graphical materials, and developed applications provide a clear picture of the concept, and highlights the clarity	The written or graphical materials, and developed applications do not show clear picture of the concept. There is room for improvement	The written or graphical materials, and developed applications fail to prove the clarity. Background knowledge is needed	Failed to prove the clarity. Need proper background knowledge to perform the tasks
Group Work	Evidence of group work, meeting logs, Coordination	Evidence of group work. However, missed meeting logs, room for improvement in Coordination	Missed meeting logs, failed to display group coordination	No group work done. Project is unacceptable

Project Requirements

In this project, you need to build a simple relational database (RDb), and its database management system (RDbMS). Your team should explore data structure concepts for creating the databases. In addition, your team should create a programming framework that can work as a RDbMS application layer. Your database should handle multi-user (2 users) requests. The database management system layer should provide a command-line interface (Graphical User Interface is not required), and perform various functionalities of relational database management system.

Expected functionalities of 5408_RDbMS_Custom(RDb)

You do not have to implement these in Phase 1; Only report is needed in Phase 1

- Create, manage databases with at least 2 users, and provide MFA (multi-factor authentication) for users. E.g. encrypted password, and security question
- Accepts user inputs in standard SQL format only from standard I/O console
- Process the queries and converts those into instructions for create/delete/update/alter/select operations etc.
- Select and implement data structures which will be used for in memory operations of the database. E.g. after query execution, the processed query will be added in a stack for execution. Once done, any permanent changes will be reflected to the custom RDb file in the persistent storage.

- Conceptualize and implement a custom file format for persistent storage (JSON/CSV or standard file formats are not accepted).

- Creation of various dynamic logs such as,
  - **General Logs**: query execution time, state of the database (e.g. how many tables are there with number of records at a given time)
  - **Event Logs:** changes in database, concurrent transactions, crash reports, etc.
  - **Query Logs:** capture user queries and timestamp of query submission

- Depending on the design - creation of Data dictionary, metadata, and lower limit/upper limit of datatypes etc.
- Options to specify constraints.

- Option to create SQL Dump (table structure and values)
- **Locking for concurrency control -**The RDBMS must follow ACID properties and implementation of two-phase locking

- **ERD** The RDBMS should create simple ERD from existing databases (data structure). The concept of ERD generation is identical to reverse engineering that is available on standard commercial DBMSs. However, the ERD does not have to be graphical. Information such as cardinality, entity names, relationships, and primary key, foreign key constrains should be captured on a text file as ERD.

Tasks for Phase 1 (Feasibility Study): Due Jun 10, 2021 @11:59 pm

- Meet your team (using Teams) and discuss the strategy;
- Capture screenshot of meetings; and write down the discussed agenda. (You do not need to keep video recordings)
- Select a programming framework and related technologies. Provide written explanation for your selection.
  - [**Note**: You need to use core Java programming. For in memory operation or buffers, you can only use standard data structure (e.g. arrays, linkedlists, tree etc.**).**
  - You **cannot** use JSON/CSV for capturing records.
  - You cannot use any external libraries or 3rd party libraries]
- Explore Data Structure(s) that is suitable for your relational database
- Document your plan
  - create flowchart highlighting the operations of the DBMS *5408_RDbMS_Custom(RDb)*
  - Write the assumptions
  - Add your initial algorithms or pseudocodes
  - Write initial Test cases for performing the unit tests
- The database should handle integer, real numbers, and text data

Task for Phase 2 (Go-Live and Project Closure): Due End of Term

- Implementation of the system
- Performing Unit Test and Validation Test

- Each group will submit a pre-recorded group video presentation of **maximum 1 hour** highlighting various components of the project, and the working model.
  - The details of the pre-recorded presentation will be posted on Teams.

- There will be a synchronous Q&A session with each group at the end of the project.
-
- Detailed documentation of meeting logs screenshots, use case, test cases, pseudocode, evidence of testing, team and individual contributions, limitations etc. in the form of a final report.
- Each group will submit a report (10 to 15pages) at the end of the final project Q&A session