UE23C642A– Advanced Data Structures **05 April 2024**

Assignment : Skip List & Log Structured Merge Tree (LSMT)

Guidelines:

a) Code can be developed in **PYTHON/Java/C/C++ (Open Source compiler IDEs)**

b) ***Assignment will have to be carried out in teams of size TWO.***

c) Submission (Code, Readme files, Test Data etc , Snapshot of results) will have to be done, on, or before deadline, in to the Google Drive shared folder

d) Summary report of the assignment will have to be handed over in *hard copy after the Demo*

e) Approx 4 -5 W***eeks*** of time will be available before submission. Actual dates will be broadcast. Hence look out !!

f) Follow fair code of ethics and , **develop your version** of code. You can discuss/consult with anyone, but write your version of the code. Plagiarism will get you zero marks !!

g) You will be called upon to Demo the assignment, to match with submission data you have provided in the Google Drive /Hard Copy .

Problem Definition, Data Generation, Testing and Logging Stats

Problem:

Implement a Key-Value Store, that can scale upto 10 Million records, with
a) Randomised Skip List b) Log Structured Merge Trees (LSMT)

Reading Material:

a) Building a Log-Structured Merge Tree Database - Part 1
b) A survey of LSM-Tree based Indexes, Data Systems and KV-stores
c) 26 Top Key-Value Databases Compared (Apr 2024)

DataSet:

- Pick any publicly available Key-value Datasets (e.g Aadhaar, Credit Card, Twitter , Kaggle etc)

Solution Strategy:

- Bulid and Populate Skip List
- Build and Populate LSMT key-value Store
- Enable a standard (Construct, Free, Insert, Find, Delete, Range Query) Interface for both

Demo and Reports:

- Your Demo should enable Interface Ops from Simple UI or Command line
- Over all Timing Metrics for Read/ Write and Rang Queries , should be shown
- Overall Storage for both Skip List and LSMT should be shown
- Produce a 4-5 Report on your Design, Challenges, Limits , Timings etc

Last para of your report should contain your observations on the Learning Outcomes of this project