

Create Database in Java - Project Design

Insert: (InsertData.java)

1. Program accepts a text file with no headers, in the following format:
(STB|TITLE|PROVIDER|DATE|REV|VIEW_TIME)
2. Program divides data into chunks of 1 million records and processes them. (If input file has more than 1 million records)
3. Records are stored in a text file
4. A corresponding text file is created which stores a unique hash for each record. This helps create unique values by STB, TITLE and DATE. If there is data collision, the new input record will not be entered into the main table. Instead it will be entered into a log table (log.txt).
5. Once the data is inserted into the main table, the log file is scanned against the main table to override the data. (Works like a Database commit).
6. Insert and data overriding is done in the same function call

Observation: I tested the code on Eclipse IDE in a PC with 8 GB RAM. The average time to insert and override an 85MB file with roughly 2 million records takes up to 22 seconds. In the attached submission, I have included the sample data you had provided (attached sample.txt).

Queries: (Queries.java)

1. Search by Title name
2. Search by date
3. Search all, ordered by name and date
4. Analysis report - calculates the total View Time, total Revenue and number of records for a query

Datastore structure:

1. A text file based storage is used for this code
2. Main table: A separate file for each date
3. Hash table: A separate file for each date to store hash for each unique record
4. Log table: Temporarily stores log records which needs to be updated in the main table
5. Temp table: Used to write data while updating records (only while committing data from log table) and then writing back to main table

Assumptions:

1. Input is in a text format with pipe delimited data with no headers
2. Input files maintains the data columns in the following order (STB|TITLE|PROVIDER|DATE|REV|VIEW_TIME)
3. Absolute path for the input text file will be provided (eg. C:\Users\aakashy\Desktop\sampl e.txt)
4. You can add add data one text file at a time.