Q1 Lab 1 - Storage Management

10 Points

This quiz is meant to take only 20 minutes or so. While you have over 24 hours to finish at your convenience, don't feel that you need to write a lot of text. A paragraph or two for each question could be sufficient. Try to answer in a way that demonstrates your understanding of lab 1.

Q1.1

5 Points

Explain what happens in your implementation of SimpleDB from Lab 1 if multiple queries execute at the same time. Each query only performs a sequential scan of one table. All queries touch different tables. Discuss when data is read from disk and what goes on in the buffer pool. Do not worry about concurrency control nor failure recovery.

If multiple queries are executing at the same time, there will be two outcomes.

The first outcomes is that the main memory that are allocated is large enough for saving all the pages from different Queries' sequential scans. For example, if the main memory in buffer pool can hold 10 pages, and there are 5 queries doing sequential scan on 5 tables, and each tables only have 2 pages to store all its tuples. Then my code in lab1 will allow all operators to do the sequential scan separately and share the cache-memory in Buffer Pool manager.

The second outcomes are throwing DbException, because I had not written the eviction policy in lab1, so when the total number of pages are larger than what cache can hold in buffer pool, it will stop adding more pages in cache, and throw an DbException. If any queries had done the sequential scan before DbException are thrown. It will return the results, and the page will be stored in the cache, however, if the query has not done the sequential scan, then it will be stopped, because the DbException has been thrown out.

Q1.2

5 Points

Describe a scenario for multiple queries running in lab 1 that would benefit from caching in the buffer pool. Again assume that each query only does a sequential scan, but for this question queries may be accessing the same tables at the same time. In your example it might help to specify an example size for the buffer pool and relation(s) that will make the scenario clear.

If multiple queries are executing at the same time on the same table, there will be two outcomes as well, similar to Q1.1.

Let's say the size for the buffer pool that it can hold is 10 pages. Because all the query will be doing the sequential scan on the same table, then if the pages in the table needed is smaller than 10, so all of the pages can be stored in the cache. Then all the queries will be able to finish. Because my code in lab1 will return the page, if the page has already been stored in the main memory of buffer pool. So it don't need every query to store the page that is already in cache again.

In the other hand, If the table's tuples needs to be stored in more than 10 pages. With the code I wrote in lab1, it will throw an DbException and none of the queries can be finished.

Quiz 1

GRADED

STUDENT
Yuanfeng Li

TOTAL POINTS
10 / 10 pts

QUESTION 1