CSE 322 Midterm II

Olcay Taner YILDIZ

I. QUESTION (21 POINTS)

Consider the following relations about the suppliers-parts-projects database. The significance of an shipment row is that the specified supplier supplies the specified part to the specified project in the specified quantity.

Suppliers(sid:integer, sname:string, status:integer, city:string)

Parts(pid:integer, pname:string, color:string, weight:integer, city:string)

Projects(prjid:integer, prjname:string, city:string)

Shipment(sid:integer, pid:integer, prjid:integer, quantity:integer)

Write the following queries in SQL.

1) Get supplier numbers for suppliers who supply project 1.
2) Get full details for parts supplied by a supplier in London.
3) Get part numbers for parts supplied by a supplier in London to a project in London.
4) Get total number of projects supplied by supplier 1.
5) Get project numbers for projects supplied entirely by supplier 1.
6) Get part numbers for parts supplied to all projects in London.
7) Get all cities in which at least one supplier, part, or project is located.

II. QUESTION (19 POINTS)

Give Java codes for the following operations:

1) Establish a connection to url 'jdbc:mysql:www.cse322.com:3083' with userID 'root' password 'root'. Do not forget exception handling codes (Use try-catch structure).
2) Prepare an SQL statement (Use prepareStatement) to the following query: 'SELECT sid FROM Suppliers WHERE status = ? AND city = ?'. Get the status and city values from user.
3) Execute the query above to get a ResultSet. Write the results to the screen.

III. QUESTION (24 POINTS)

In a simple file organization

- The number of data pages when records are packed onto pages with no space is 1024 (B).
- The number of records per page is 100 (R).
- The average time to read and write disk page is 10 milliseconds (D).
- The fanout is 32 (F).

Given the following two queries, 'SELECT sname FROM Suppliers WHERE sid = 34' and 'SELECT sname FROM Suppliers WHERE sid < 10', what is the total cost to process these two queries for

1) Heap file
2) File of supplier records sorted on sid
3) Clustered B+ tree file with search key sid (Assume %100 occupancy)
4) Heap file with an unclustered hash index on sid

Explain briefly how you have calculated the total cost for each case.

IV. QUESTION (15 POINTS)

1) Sort the following memory types in increasing order of speed: cache, tape, main memory, magnetic disk, cpu.
2) Consider a disk with a sector size of 512 bytes, 2000 tracks per surface, 50 sectors per track, five double-sided platters. What is the capacity of a track in bytes? What is the capacity of the disk?
3) Give an advantage of fixed-length record format over variable-length record format.

V. QUESTION (21 POINTS)

Suppose that a page can contain at most four data values and that all data values are integers. Using only B+ trees of order 2, give examples of each of the following:

1) A B+ tree whose height changes from 2 to 3 when the value 25 is inserted. Show your structure before and after the insertion.
2) A B+ tree in which the deletion of the value 25 leads to a redistribution. Show your structure before and after the deletion.
3) A B+ tree in which the deletion of the value 25 causes a merge of two nodes but without altering the height of the tree. Show your structure before and after the deletion.