Assignment-01

Subject:- Advanced Database Management System

Start Date: 15 September 2018

End Date: 20 September 2018

---

1. Why is the use of DBMS recommended? Explain by listing some of its major advantages.
2. What is the use of DROP command and what are the differences between DROP, TRUNCATE and DELETE commands?
3. Describe the different types of file organization? Explain using a sketch of each of them with their advantages and disadvantages. Mention the purpose of indexing. How this can be done by B+ tree? Explain.
4. Computer Sciences Department frequent fliers have been complaining to Dane County Airport officials about the poor organization at the airport. As a result, the officials decided that all information related to the airport should be organized using a DBMS, and you have been hired to design the database. Your first task is to organize the information about all the airplanes stationed and maintained at the airport. The relevant information is as follows:
   - Every airplane has a registration number, and each airplane is of a specific model.
   - The airport accommodates a number of airplane models, and each model is identified by a model number (e.g., DC-10) and has a capacity and a weight.
   - A number of technicians work at the airport. You need to store the name, SSN, address, phone number, and salary of each technician.
   - Each technician is an expert on one or more plane model(s), and his or her expertise may overlap with that of other technicians. This information about technicians must also be recorded.
   - Traffic controllers must have an annual medical examination. For each traffic controller, you must store the date of the most recent exam.
   - All airport employees (including technicians) belong to a union. You must store the union membership number of each employee. You can assume that each employee is uniquely identified by a social security number.
   - The airport has a number of tests that are used periodically to ensure that airplanes are still airworthy. Each test has a Federal Aviation Administration (FAA) test number, a name, and a maximum possible score.
   - The FAA requires the airport to keep track of each time a given airplane is tested by a given technician using a given test. For each testing event, the information needed is the date, the number of hours the technician spent doing the test, and the score the airplane received on the test.

a. Draw an ER diagram for the airport database. Be sure to indicate the various attributes of each entity and relationship set; also specify the key and participation constraints for each relationship set. Specify any necessary overlap and covering constraints as well.
b. The FAA passes a regulation that tests on a plane must be conducted by a technician who is an expert on that model. How would you express this constraint in the ER diagram? If you cannot express it, explain briefly.
c. Translate your ER diagram into a relational schema, and show the SQL statements needed to create the relations, using only key and null constraints. If your translation cannot capture any constraints in the ER diagram, explain why.
d. You can also modified the ER diagram to include the constraint that tests on a plane must be conducted by a technician who is an expert on that model. Can you modify the SQL statements defining the relations obtained by mapping the ER diagram to check this constraint?

5. Explain the difference between Hash indexes and B+-tree indexes. In particular, discuss how equality and range searches work, using an example.
6. If you were about to create an index on a relation, what considerations would guide your choice? Discuss:
   a. The choice of primary index.
   b. Clustered versus unclustered indexes.
   c. Hash versus tree indexes.
   d. The use of a sorted file rather than a tree-based index.
   e. Choice of search key for the index. What is a composite search key, and what considerations are made in choosing composite search keys? What are index-only plans, and what is the influence of potential index-only evaluation plans on the choice of search key for an index?
7. Both disks and main memory support direct access to any desired location (page). On average, main memory accesses are faster, of course. What is the other important difference between the two (from the perspective of the time required to access a desired page)?
8. Consider a disk with a sector size of 512 bytes, 2000 tracks per surface, 50 sectors per track, five double-sided platters, and average seek time of 10 msec.
   a. What is the capacity of a track in bytes? What is the capacity of each surface? What is the capacity of the disk.
   b. How many cylinders does the disk have?
   c. Give examples of valid block sizes. Is 256 bytes a valid block size? 2048? 51200?
   d. If the disk platters rotate at 5400 rpm (revolutions per minute), what is the maximum rotational delay? 5. If one track of data can be transferred per revolution, what is the transfer rate?
9. Explain what the buffer manager must do to process a read request for a page. What happens if the requested page is in the pool but not pinned?

10. Suppose that you have a sorted file and want to construct a dense primary B+ tree index on this file.
    a. One way to accomplish this task is to scan the file, record by record, inserting each one using the B+ tree insertion procedure. What performance and storage utilization problems are there with this approach?
    b. Explain how the bulk-loading algorithm described in the text improves upon this scheme.