## Sample questions for File Orgs and Transaction Analysis

**Q1**

## Database management systems allow multiple users to access the same data concurrently. It is generally recognised that there are three types of situation in which a pair of concurrent transactions may interfere with one another. Produce an analysis of two of these situations and show clearly how a database management system can overcome these problems utilising 2 Phase Locking.

Consistency:

Isolation:

Properties of Transaction:

1. Using 2 Phase Locking can lead to deadlock. Describe (possibly using a timing diagram). What is meant by deadlock in this context?

*By using 2 phase locking can lead to a deadlock, this is done when 2 process’ are locking down a variable which is trying to be read by the other process. By doing so, both process’ will be locked out and will need to rollback after a set timer.*

|  |  |  |
| --- | --- | --- |
| *Time* | *V1* | *V2* |
| *T1* |  | *Read(Lock(x))* |
| *T2* | *Read(Lock(n))* | *Start(timer)* |
| *T3* | *Start(timer)* | *Read(n)* |
| *T4* | *Read(x)* | *Wait* |
| *T5* | *Wait* | *timeout* |
| *T6* | *Timeout* | *Release(x)* |
| *T7* | *Release(n)* | *Rollback* |
| *T8* | *Rollback* |  |

**Q2**

1. Hash files contain records which are stored at addresses calculated using a hash function on an attribute value (normally the key). Explain what is meant by the term ‘synonym’ in this context.

**(b)** Using the following ‘key’ values:

23,45,54,33,67,77,29,27,93

and the following hash algorithm

divide by 11 and take the remainder as the generated address

to place the ‘records’ into the addresses (0-10) generated by the hash function.

Use diagrams (different sets for i, ii and iii) to demonstrate

i) Dealing with synonyms using Open Addressing.

ii) Dealing with synonyms using Unchained Overflow

iii) Dealing with synonyms using Chained Overflow

**(20 marks)**

**Q3**

**(a)** A technique for maximising the degree of concurrency in a system is that of Serialisation. Serialisability can be achieved in a number of ways. One of the most widely used ways incorporates the ‘Locking’ of data items. A transaction must claim a lock on a data item before performing a read or write operation. Describe the two types of lock that can be ‘placed’ on a data item and explain the difference between them.

(6 marks)

**(b)** Describe the locking process that takes place when a transaction needs to access a data item.

(5 marks)

**(c)** An Indexed Sequential File is a sequentially sorted data file, with the index on the field (attribute) on which the file has been sequentially ordered. It normally has a storage area for the actual data file, a separate index (or indexes) and an overflow area.

The data file can be organised into ‘buckets’ and so it is not necessary to have one index entry for every record in the file, but an index entry for every ‘bucket’.

Insert the following key values into an indexed sequential file that has 3 records in each bucket.

This file will have 1 data file, 1 index and 1 overflow area.

Leave 1 record ‘slot’ free in each bucket.

S02, S04, S05, S07, S09, S10, S14, S16, S17, S19

Document the changes to the Index that occur as you add each record to the data file. (You may find it useful to make the ‘box’ into which you place the index value large enough to put a line through previous values as new records are added to the data file)

After setting up this file, add the following records:

S01, S08, S11, S15, S12, S20, S18.

Remember to use the overflow area where necessary and document the changes to the index that occur as you add each record to the data file.

(14 marks)

**Q4**

**(a)** There are two approaches used by DBMS’s when placing records onto secondary storage:

The DBMS uses the underlying Operating System’s file organisations and access methods.

The DBMS does not use the underlying Operating System’s file organisations and access methods. The internal organisation of the files that make up the physical database are managed by the DBMS itself. It simply uses the primitive disk management operations provided by the Operating System to retrieve and store sectors on a disk.

State the advantages and disadvantages of both these approaches.

**(6 marks)**

**(b)** Connolly and Begg define a transaction as:

“An action, or series of actions, carried out by a single user or application program, which reads or updates the contents of the database”

Describe the two main states in which a transaction can exist.

**(5 marks)**

**(d)** An Indexed Sequential File is a sequentially sorted data file, with the index on the field that the file is sequentially ordered on. It normally has a storage area for the actual data file, a separate index (or indexes) and an overflow area.

The data file can be organised into ‘buckets’ and so it is not necessary to have one index entry for every record in the file, but an index entry for every ‘bucket’.

For large data files, the corresponding index file would be very large. To overcome this problem we can implement a multilevel index.

Insert the following key values into a data file that has 3 records in each bucket.

Leave 1 record ‘slot’ free in each bucket.

Set up a 2 level index to be used to access these records. (Use 2 ‘records’ in each index)

S04,S05, S07, S09, S11, S12, S14, S16.

(You can ignore the overflow area for this example).

Document the changes to the Indexes that occur as you add each record to the Data file. (You may find it useful to make the ‘box’ into which you place the index value large enough to put a line through previous values as new records are added to the data file)

(14 marks)

**Other possible parts of questions:**

The use of locks alone does not guarantee serialisability, thus we have to implement two-phase locking. Two-phase locking states that each transaction can be divided into two phases. Describe each of these two phases.

(**4 marks)**

The consistency and reliability aspects of transactions are due to the ‘ACIDity’ properties of transactions.

Describe and discuss each of these properties.

(**12 marks)**

## Database management systems allow multiple users to access the same data concurrently. It is generally recognised that there are three types of situation in which a pair of concurrent transactions may interfere with one another. Produce an analysis of two of these situations and show clearly how a database management system can overcome these problems utilising 2 Phase Locking.

## (10 marks)

Using 2 Phase Locking can lead to deadlock. Describe (possibly using a timing diagram). What is meant by deadlock in this context?

**(5 marks)**

## Database management systems allow multiple users to access the same data concurrently. It is generally recognised that there are three types of situation in which a pair of concurrent transactions may interfere with one another. Produce an analysis of one of these situations and show clearly how a database management system can overcome this problem utilising 2 Phase Locking.

## (10 marks)