Question 1 [10 Pts.]

Discuss the advantages and disadvantages of using (a) an unordered file, (b) an ordered file, and (c) a static hash file with fixed number of buckets and chaining employed. Which operations can be performed efficiently on each of these primary file organizations, and which operations are expensive?

Question 2 [20 Pts.]

A data file, consisting of 5632 disk blocks, requires sorting using the external sort-merge algorithm. The system provides a main-memory buffer that can accommodate 64 blocks at a time.

1. Determine the number of passes required in the merge phase of the external sort merge algorithm, considering the limitations of the buffer space. [10 points]

2. Estimate the total number of disk block accesses involved in the overall process. This estimation should include all reads and writes during both the initial run creation, sort, and the merge phase. Explain your work. [10 points]

Question 3 [10 Pts.]

What are the differences among primary, secondary (e.g., a Hash index or a B+ Tree index), and clustering indexes? How do these differences affect the ways in which these indexes are implemented? Which of the indexes are dense, and which are not?

Question 4 [50 Pts.]

For each of the following statements about Nested-Loop Joins, determine whether they are True or False and explain WHY. [10 points each]

a. In a simple nested loop join where both tables fit entirely in memory, the choice of which table to use as the inner/outer table significantly affects I/O costs.

b. In a simple nested loop join where only one of the tables fits entirely in memory, it is beneficial to use that table as the inner table.

c. If neither table fits entirely in memory, I/O costs would be lower if we process both tables on a per-block basis rather than a per-tuple basis.

d. For a block nested loop join, in the worst case (indexes are absent), each block in the inner table has to be read once for each tuple in the outer table.

e. In an index nested loop join where only one of the tables has an index on the join attribute, the choice of inner/outer tables should be made based on the table sizes.

Question 5 [10 Pts.]

Extend the Sort-Merge Join algorithm to implement the RIGHT OUTER-JOIN opera tion. Explain your work.