University of Helsinki, Department of Computer Science
Database Management, separate / renewal exam 18.1.2005 / H.Laine

Write the name of the course, the date of exam, your name, signature, and date of birth on each separate answer paper.

1. Explain briefly the structure of a typical file page. (12)

2. Define the concept dynamic hashing. Describe the principle of either extendible or linear hashing. (12)

3. Let's consider the tables `Order(orderID, dateOrdered, deliveryAddress, customerNo, etc)` and `Customer(customerNo, name, city, billingAddress, etc)`. Table Order has 30000 rows. Average size of the rows is 400 bytes. Table Customer has 1000 rows with the average size of 200 bytes. Page size in both files is 4KB. What is the most efficient way to execute the query

    ```
    select orderID, name, billingAddress, city, deliveryAddress
    from Order, Curtomer
    where Order.customerNo=customer.customerNo;
    ```
    You may use at most about 200 pages of buffer / main memory space for the execution (assumptions to simplify the calculations may be done). Both tables are implemented as piles. They both have a secondary index on their primary key. Table Order has also an index based on column customerNo. All indexes are implemented by hashing. They have so few overflow pages that they need not be considered in calculations. Motivate your answer. How many disk accesses are needed? (14)

4. a) What information is stored in a database transaction log?
    b) How is the log involved in committing a transaction (commit protocol)?
    c) How is the log used in database recovery? (12)