Data Management – exam of 08/06/2023 (A)

Problem 1

Let R(A,B,C) (with 26.000 pages) and Q(D,E,F) (with 10.000 pages) be two relations stored in a heap at processor P_0. We know that 200 free buffer frames are available at P_0 and that 100 tuples of each relation fit in one page. Also, we know that attribute C of R contains 500 values uniformly distributed in the tuples of the relation, and the same holds for attribute D of Q. Finally, we know that there are 10 processors $P_1, \ldots, P_{10}$ that we can use for processing queries, each with 80 free frames available. Consider the query that computes the equi-join between R and Q on the condition C = D, and compare the following two cases:

1.1 The query is executed at processor P_0.

1.2 The query is executed in parallel after sending the tuples of the two relations to processors $P_1, \ldots, P_{10}$.

For each of the above cases, illustrate the algorithm you would use and discuss its cost in terms of number of page accesses (case 1.1) and elapsed time (case 1.2)

Solution 1

3.1 The query is executed at processor P_0.

We cannot use a one-pass algorithm for computing the join, but, since $200 \times 199 > 26.000 + 10.000$, we can use a two-pass algorithm based on sorting. We know that we have two types of such algorithms, the simple sort-based join algorithm, and the sort-merge join algorithm and we know that the latter is more efficient, as far as the problem of "fragments with too many joining tuples" does not occur.

Let us compute the maximum number of tuples joining between the two relations, by computing the number of tuples in R having the same value in the joining attribute C as $26.000 \times 100/500 = 5.200$, and the number of tuples in Q having the same value in the joining attribute D as $10.000 \times 100/500 = 2.000$. This means that we need at least $2.000/100 = 20$ pages in the buffer to hold all the tuples of the maximum fragments of Q, and apply the one-pass join for such fragments. Now, in the case of sort-merge join algorithm, we have $26.000/200 = 130$ sublists for R and $10.000/200 = 50$ sublists for Q, and therefore, since $200 - (130 + 50) = 20$, and we need one buffer frame for the output, we do not have sufficient room in the buffer for holding the fragments of Q with tuples with the same value of D during the merging phase of the algorithm. On the other hand, it is immediate to verify that using the simple sort-based join algorithm, we have plenty of room for holding the fragments of Q with tuples with the same value of D during the merging phase and we conclude that we choose the simple sort-based join algorithm, with a cost of $5 \times (26.000 + 10.000) = 180.000$.

3.2 The query is executed in parallel after sending the tuples of the two relations to processors $P_1, \ldots, P_{10}$.

We assume that we have a good hash function on the values in the attributes C and D so that the hash-based partitioning distributes the tuples of the two relations uniformly. Using such hash function, we read the two relations R and Q at processor P_0 and we send to each of the 10 processors $P_1, \ldots, P_{10}$ 2.600 pages of R and 1.000 pages of Q. Let us now consider the situation at any of the 10 processor, say processor P_i. It is immediate to see that we cannot use a one-pass algorithm at P_i, but we can use a two-pass algorithm (e.g., based on sorting), because $2.600 + 1.000 < 80 \times 79$. Also, notice that at each processor, the number of tuples in R having the same value in the joining attribute C is $2.600 \times 100/50 = 5.200$, corresponding to 52 pages, and the number of tuples in Q having the same value in the joining attribute D is $1.000 \times 100/50 = 2000$, corresponding to 20 pages. This means that we need at least $2.000/100 = 20$ pages in the buffer to hold all the tuples of the maximum fragments of Q, and apply the one-pass join for such fragments. In the case of sort-merge join algorithm, we have $2.600/80 = 33$ sublists for R and $1.000/80 = 13$ sublists for Q after the first pass and therefore we have $80 - (33 + 13) = 34$ buffer frames available for holding all the 20 tuples of the fragments of Q with the same value for `D`, and apply the one-pass join for such fragments during the second pass (the "merging phase"). We conclude that we choose the sort-merge join algorithm, and, if we first store the two relations in the secondary storage when they arrive at processor P_i and then we execute the sort-merge join algorithm, the elapsed time will be $26.000 + 10.000$ (for reading the two relations at processor P_0) $+ 2.600 + 1.000$ (for storing the relation in the processor P_i) $+ 3 \times (2.600 + 1.000) = 50.400$ page accesses as elapsed time. We can actually do better by avoiding to store the two relations when they arrive at the processor, simply producing the corresponding sorted sublists, thus obtaining $26.000 + 10.000 + 2 \times (2.600 + 1.000) = 43.200$ page accesses as elapsed time.

Problem 2
Consider the following schedule S:

$$B_1\ w_1(A)\ B_2\ w_2(D)\ B_3\ w_3(E)\ c_3\ B_4\ r_4(E)\ r_4(D)\ w_4(C)\ B_5\ r_5(C)\ c_4\ w_5(C)\ w_1(D)\ c_1\ w_2(A)\ c_2\ c_5$$

where the action B_i means "begin transaction T_i", the initial value of every item A, C, D, E is 100 and every write action increases the value of the element on which it operates by 100. Suppose that S is executed by PostgreSQL, and describe what happens when the scheduler analyzes each action (illustrating also which are the values read and written by all the "read" and "write" actions) in both the following two cases: (1) all the transactions are defined with the isolation level "read committed"; (2) all the transactions are defined with the isolation level "repeatable read".

Solution 2
We first deal with the isolation level "read committed". We remind the reader that such isolation level does not prevent the unrepeatable read anomaly (that, however, cannot occur in our case) nor the lost update anomaly (that actually occurs).

- $w_1(A)$: T_1 writes 200 on A in the local store.
- $w_2(D)$: T_2 writes 200 on D in the local store.
- $w_3(E)$: T_1 writes 200 on E in the local store.
- c_3: T_3 commits and the value 200 for E is written in the database.
- $r_4(E)$: T_4 reads 200.
- $r_4(D)$: T_1 reads 100.
- $w_4(C)$: T_4 writes 200 on C in the local store.
- $r_5(C)$: T_1 reads 100.
- c_4: T_4 commits and the value 200 for C is written in the database.
- $w_5(C)$: T_5 writes 300 on C in the local store (**a kind of lost update anomaly**).
- c_5: T_5 commits and the value 300 for C is written in the database.
- $w_1(D)$: not executed, because T_2 holds the write lock on D; so T_1 is suspended because it must wait for the end of T_2.
- $w_2(A)$: not executed because T_1 holds the write lock on A; a deadlock is recognized, T_2 is aborted and T_1 is resumed, writing the value 200 on D in the local store
- c_1: T_1 commits and the value 200 both for A and for D is written in the database.

- c_2: T_2 was aborted and now it rollbacks.

We now deal with the isolation level "repeatable read" (in bold the difference with respect to the previous case). We remind the reader that such isolation level prevents both the unrepeatable read anomaly (that, however, cannot occur in our case) and the lost update anomaly (that can occur).

- $w_1(A)$: T_1 writes 200 on A in the local store.
- $w_2(D)$: T_2 writes 200 on D in the local store.
- $w_3(E)$: T_1 writes 200 on E in the local store.
- c_3: T_3 commits and the value 200 for E is written in the database.
- $r_4(E)$: T_4 reads 200.
- $r_4(D)$: T_4 reads 100.
- $w_4(C)$: T_4 writes 200 on C in the local store.
- $r_5(C)$: T_5 reads 100.
- c_4: T_4 commits and the value 200 for C is written in the database.
- $w_5(C)$: T_5 **aborted with message: "ERROR: could not serialize access due to concurrent update"**.
- c_5: T_5 rollbacks.
- $w_1(D)$: not executed, because T_2 holds the write lock on D; so T_1 is suspended because it must wait for the end of T_2.
- $w_2(A)$: not executed because T_1 holds the write lock on A; a deadlock is recognized, T_2 is aborted and T_1 is resumed, writing the value 200 on D in the local store
- c_1: T_1 commits and the value 200 both for A and for D is written in the database.
- c_2: T_2 was aborted and now it rollbacks.

Problem 3

Answer the following two questions, providing a suitable motivation for each answer.

3.1 Give an example of three transactions, which obey 2PL and have the following properties: (i) there exists a schedule S on the three transactions such that when S is given in input to a 2PL scheduler, a deadlock occurs; (ii) for each pair of the three transactions and for any schedule S on such pair, no deadlock occurs when S is given in input to a 2PL scheduler.

3.2 Try to generalize the above example to the case of n transactions. More precisely, try to come up with an example of n transactions, which obey 2PL and have the following properties: (i) There exists a schedule S on the n transactions such that when S is given in input to a 2PL scheduler, a deadlock occurs. (ii) For each $n-1$ transactions of the n chosen transactions and for any schedule S on such $n-1$ transactions, no deadlock occurs when S is given in input to a 2PL scheduler.

Solution 3

3.1 Here is the example:
$T_1 = w_1(x_1)\ w_1(x_2)$,
$T_2 = w_2(x_2)\ w_2(x_3)$,
$T_3 = w_3(x_3)\ w_3(x_1)$.
Consider now the schedule S: $w_1(x_1)\ w_2(x_2)\ w_3(x_3)\ w_1(x_2)\ w_2(x_3)\ w_3(x_1)$. It is easy to see that:

- when S is given in input to a 2PL scheduler, a deadlock occurs;
- for every pair of transactions T_1, T_2, T_3 and for every schedule S' on such pair, no deadlock occurs when S' is given in input to a 2PL scheduler.

3.2 The generalization of the above example to n transactions is immediate:
$T_1 = w_1(x_1)\ w_1(x_2)$,
$T_2 = w_2(x_2)\ w_2(x_3)$,
$\ldots$
$T_{n-1} = w_{n-1}(x_{n-1})\ w_{n-1}(x_n)$,
$T_n = w_n(x_n)\ w_n(x_1)$.
Consider now the schedule S: $w_1(x_1)\ w_2(x_2) \ldots w_{n-1}(x_{n-1})\ w_n(x_n)\ w_1(x_2)\ w_2(x_3) \ldots\ w_{n-1}(x_n)\ w_n(x_1)$. It is easy to see that:

- when S is given in input to a 2PL scheduler, a deadlock occurs;
- for every pair of transactions $T_1, T_2, T_3, \ldots, T_n$ and for every schedule S' on such pair, no deadlock occurs when S' is given in input to a 2PL scheduler.

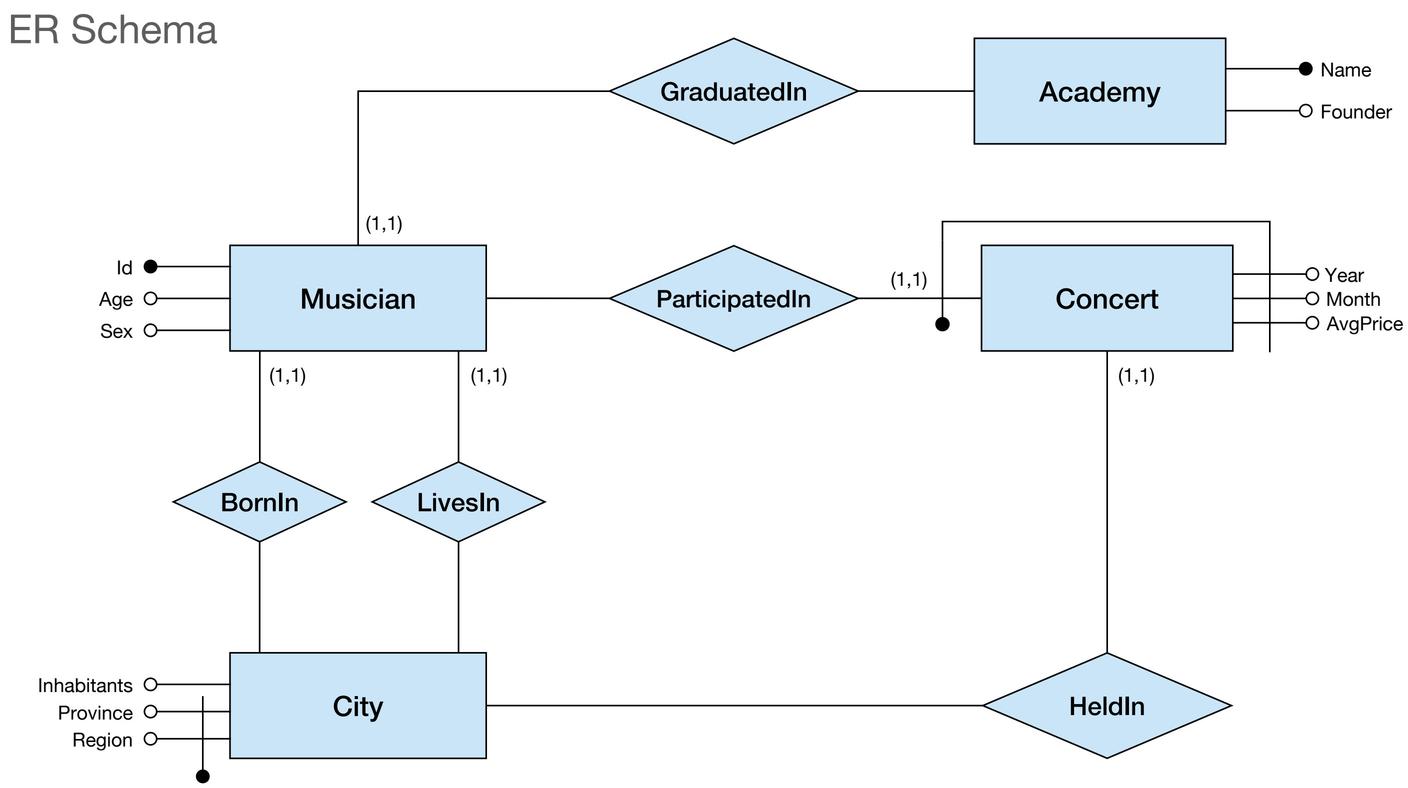
Problem 4
Let `Person(id,age)` be a relation with 500 pages stored in a file sorted on ⟨`id,age`⟩, `City(name,region)` a relation with 200 pages sorted on ⟨`name,region`⟩ and `Visited(id,age,name,region)` a relation with 20.000 pages sorted on ⟨`id,age,name,region`⟩. With the goal of knowing, for each person, the cities that (s)he has not visited, we want to compute the set difference between the cartesian product of `Person` and `City` and the relation `Visited`. Tell which is the best algorithm that a query engine with 206 free buffer frames should use for this task, indicating also the cost of executing such algorithm.

Solution 4
An obvious and acceptable solution would be to compute the whole cartesian product sorted on the basis of the same sorting criteria as `Visited`, store it in a temporary relation and then compute in one pass the set difference between such temporary relation and the relation `Visited`. However, it is worth looking for a more efficient method, that does not require to materialize the cartesian product. For this purpose, we can notice that, since `Person(id,age)` is sorted on ⟨`id,age`⟩ and `City(name,region)` is sorted on ⟨`name,region`⟩, we can easily compute the cartesian product of `Person` and `City` in such a way that the resulting relation is sorted on ⟨`id,age,name,region`⟩: it is sufficient to combine the tuples of the two relation by respecting the order of each of them. Also, notice that the whole relation `City` fits in the buffer. This means that if we read `City` in 200 buffer frames and then we read `Person` one page at a time using the 201th buffer frame, we can produce the cartesian product sorted on ⟨`id,age,name,region`⟩ one page at a time using the 202th buffer fame. In turn this means that, based on the sorting on ⟨`id,age,name,region`⟩ of both the cartesian

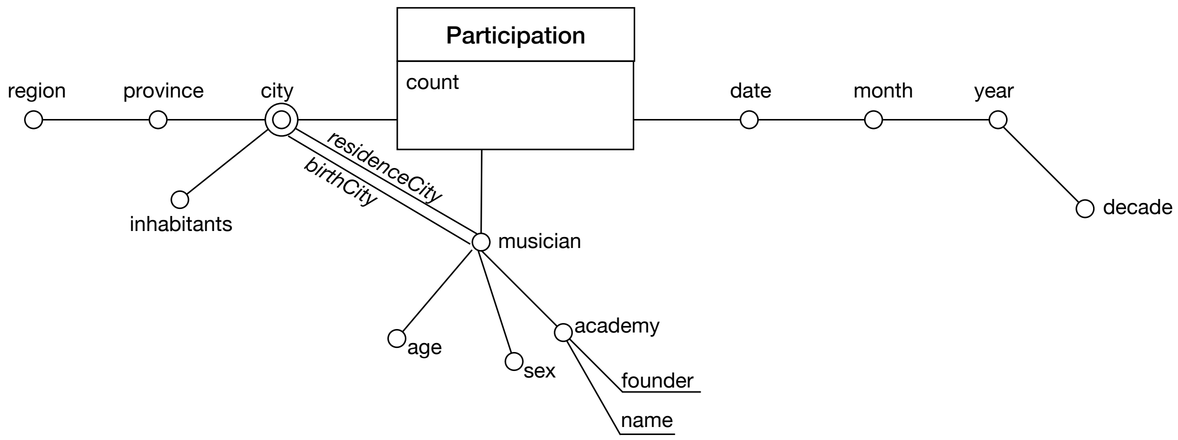
product of `Person` and `City` and the relation `Visited`, we can compute on the fly their difference by reading the relation `Visited` one page at a time in the 203th buffer frame and using the 204th buffer frame as the output frame. Since we access each page of all the relations only once, the whole algorithm is one-pass, and its cost is 200 + 500 + 20.000 = 20.700 page accesses.

Problem 5 (only for students enrolled in an A.Y. before 2021/22 who do **not** do the project)
Consider a database B about high schools, where: (i) for each student we are interested in the id, the age, the sex, the city of birth, the family (s)he is living with, and the school where (s)he has been enrolled in the various years, with the decades of the years; (ii) for each school we are interested in the type (classical, scientific, technical, etc.) and the city where the school is located; (iii) for each city we are interested in the province, the region and the number of inhabitants; (iv) for each family we are interested in the level of the income and the number of members. We want to build a data warehouse on all the above data for various analyses on student enrollments in the last 30 years. You are asked to show (1) the ER schema of the database B, (2) the DFM schema of the data warehouse, (3) the corresponding star schema (optionally, with the queries used to populate the star schema tables), and (4) based on such tables, the SQL query that computes the number of students, for each province of the school where they are enrolled, for each region of the city of residence of the students, and for each level of the income of the family of the students.

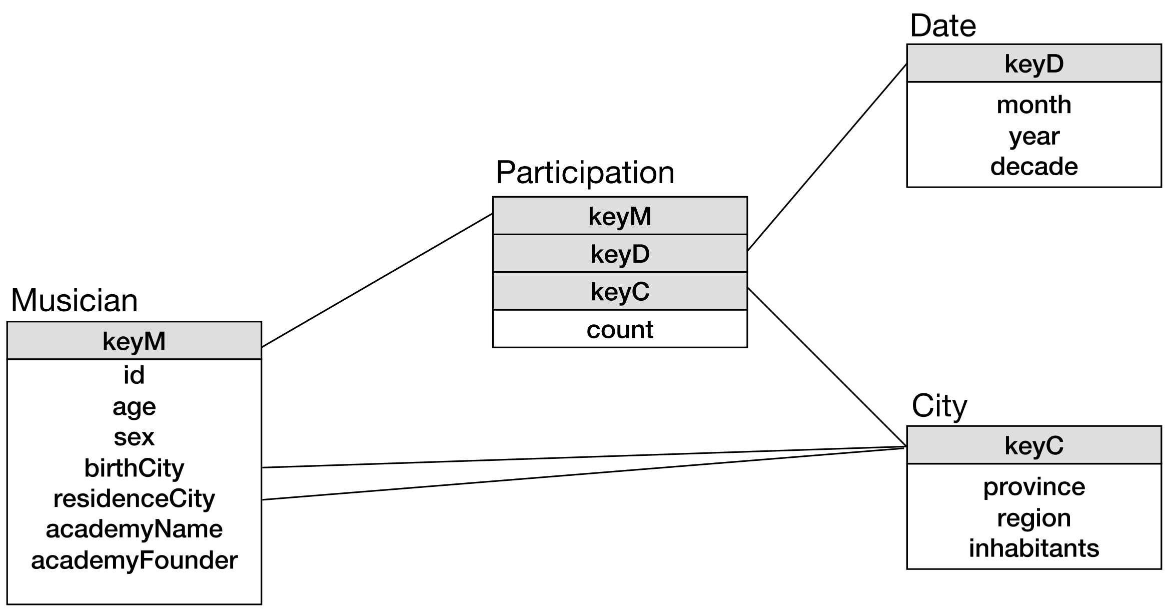
Solution 4



DFM Schema



Star Schema



SQL query

Number of musicians who participated in concerts, for each region of the city of birth of the musician, and for each music academy where the musician graduated

```
SELECT c.region, m.academyName, count(*) as noOfMusicians
FROM Participation p JOIN Musician m ON m.id =  p.keyM
          JOIN City c ON m.birthCity = c.keyC
GROUP BY c.region, m.academyName
```