Database Systems

A.Y. 2021/2022

Master Degree in Computer ScienceFirst written exam 01/07/2022 from 9:30 to 14:00

“Computer management”	
1. 2. 3. 4. 5. 6. 7. 8. 9. 10 11	We want to design a database containing information about the software installed on the computers of a company. The system must keep track of the available packages. The name and description are known for each package available. For each package, there are different versions available, which are identified by a couple of numbers: major release and minor release (for example ver. 3.5). In addition to the number, for each version, we know its size on disk and release date. Multiple versions may be available for the same package, but not with the same release date. Furthermore, there is a dependency relationship between packages: if package A depends on package B it means that A needs the presence of B. This means that not all the packages can be installed on all the computers. Furthermore, some specific versions of a package may have additional dependencies on other packages. Available versions of the packages can be installed on computers. For each installation of a version of a package on a computer, the date on which it was made and any notes are recorded. For each package, only one version can be installed on a single computer.

analyze these specifications, filtering the ambiguities present and then grouping them homogeneously. Represent the specifications with an ER diagram. Indicate the strategy followed in the conceptual modeling phase. Complete the schema documentation with any constraints not expressed by the conceptual ER schema. *(7 points)*

2) Consider the conceptual scheme defined in the previous exercise. Suppose that the following operations are carried out on this data:
Op1: Verify the possibility of installing a package, knowing the version, on a given PC (160 times a day)
Op2: Install a package (150 times a day)
Op3: Search for computers on which a given package is installed, sorted by version (100 times a week)
Op4: Display of installed packages, including the number of PCs on which they are installed (15 times a day)
Op5: Removal of a computer (10 times a day)

Taking into account that there are 600 computers, define the table of volumes and accesses for the defined conceptual schema, then design the logical schema of an object-relational database. *(6 points)*

3) Consider the following schema:
CUSTOMER (CF, name, city)
BANK (name, number_of_employees)
BRANCH (code, bank, city) with referential integrity constraint between the bank attribute and the BANK table
WITHDRAWAL (code, customer, agency, account, amount_withdrawn, date) with a referential integrity constraint between the customer attribute and the CUSTOMER table, between the agency attribute and the AGENCY table, between the account attribute and the ACCOUNT table
ACCOUNT (code, customer, agency, balance, opening_date) with a referential integrity constraint between the customer attribute and the CUSTOMER table, between the agency attribute and the AGENCY table
Define
• a trigger that, in the case of two withdrawals greater than € 1000 made on the same day, sends a warning to the Bank of Italy (suppose an existing procedure).
• a materialized view that provides, for each Agency, the number of daily withdrawals and the total withdrawn, in the year 2022. This (materialized) view must be kept updated by means of triggers. *(5 points)*

4) Given query 3 of exercise 1, show a physical query plan for this query for a DBMS that allows primary heap data structures and secondary dense B+trees, assuming there are no indexes. For each intermediate partial result, report the expected size and for each step, the expected cost. Hint: start from the result of the algebraic optimization. *(6 points)*

5) Analyze the following schedule of a set of transactions, and clarify whether it constitutes a view-serializable schedule or not, whether it constitutes a conflict-serializable schedule or not, and whether or not it can be used to generate a schedule in 2PL. Motivate all three responses.
S: r1 (x), w2 (x), r3 (x), r1 (y), r4 (z), w2 (y), r1 (v), w3 (v), r4 (v), w4 (y), w5 (y), w5 (z) *(5 points)*

6) Describe the 2PC protocol. *(4 points)*