
Question 1
Consider the following database schema.

COURSES = {CNo, Title, Credits, Dept}
STUDENTS = {ID, Name, DOB, Dept}
TRANSCRIPTS = {ID, CNo, Semester, Year, Grade}

The primary key for COURSES is CourseNo and there is a primary B+ tree index on this attribute.
The primary key for STUDENTS is ID and there is a primary B+ tree index on this attributes.
The university has 40 departments.
COURSES has 1,000 tuples and 100 disk blocks.
STUDENTS has 20,000 tuples and 1,000 disk blocks.
TRANSCRIPTS has 400,000 tuples and 10,000 disk blocks.

Estimate the size of the answer (number of tuples) for each of the following queries:

Q1. (5%) \( \sigma_{Dept='Math'}(COURSES) \)
Q2. (5%) \( \sigma_{Dept='Math'}(STUDENTS) \)
Q3. (5%) \( STUDENTS \bowtie TRANSCRIPTS \)
Q4. (10%) \( (\sigma_{Dept='Math'}STUDENTS) \bowtie TRANSCRIPTS \)
Q5. (5%) \( \sigma_{Dept='Math'}(STUDENTS \bowtie TRANSCRIPTS) \)
Q6. (5%) \( COURSES \bowtie TRANSCRIPTS \)
Q7. (10%) \( (\sigma_{Dept='Math'}COURSES) \bowtie TRANSCRIPTS \)
Q8. (5%) \( \Pi_{ID,CNo,Title}(\sigma_{Dept='Math'}COURSES) \bowtie TRANSCRIPTS \)
Q9. (10%)

\[
\text{select *}
\text{from COURSES, STUDENTS, TRANSCRIPTS}
\text{where STUDENTS.ID = TRANSCRIPTS.ID and COURSES.CNo = TRANSCRIPTS.CNo}
\]

Q10. (40%) For this query, devise two execution plans and estimate their costs (that is, the number of block accesses). Do not include the cost of writing the final answer. Try to use more efficient execution plans.

\[
\text{select *}
\text{from COURSES, STUDENTS, TRANSCRIPTS}
\text{where STUDENTS.ID = TRANSCRIPTS.ID and COURSES.CNo = TRANSCRIPTS.CNo}
\text{and STUDENT.Dept = 'CS' and COURSES.Dept = 'Math'}
\]