

Turn in this assignment electronically using the digital dropbox on Blackboard. You may prepare your answers using a word processor or you may hand-write them, but you must submit a PDF file to the digital dropbox (i.e. hand-written solutions need to be scanned in and converted to PDF and word processing documents need to be converted to PDF).

Name your file as described below. Points will be deducted for improperly formatted filenames:

<youronyen>-hw9.pdf

<youronyen> should be your UNC Onyen. This will be used to email your grade to you. The next character is a dash (-), NOT an underscore, followed by "hw9" for homework 9, and ".pdf".

### **Problem 1 (adapted from Ramakrishnan)**

Consider a distributed database that has the following two relations:

Employees (eid: integer, did: integer, ename: varchar, salary: real)  
Departments (did: integer, mgrid: integer, location: varchar)

The Employees relation is stored in New York and the Departments relation is stored in Los Angeles. The Employees table is 1000 pages and the Departments table is 50 pages. For each query below, describe the evaluation plan and give its cost in terms of D (the cost to read a page) and S (the cost to ship a page).

- Find the ename of the employee with the highest salary. Assume the query is run at Los Angeles and the "ship to one site" approach is used. Note that no join is needed for this query.
- Generate a list of all the managers, their department id number, and a list of all the employees that work for that department (i.e. join Employees and Departments). Assume the query is run at Los Angeles and the "fetch as needed" approach is used with a nested loop join.
- Generate a list of all the managers, their department id number, and a list of all the employees that work for that department (i.e. join Employees and Departments). Assume the query is run at Los Angeles and the "ship to one site" approach is used with a nested loop join.

### **Problem 2 (from Ramakrishnan)**

Suppose that the two-phase commit protocol with presumed abort is used. Explain how the system recovers from failure and deals with a particular transaction T in each of the following cases.

- A subordinate site for T fails before receiving a prepare message.
- A subordinate site for T fails after receiving a prepare message and force-writing an abort log record, but before responding to the prepare message.

### **Problem 3**

Consider the following query on the demp, trucks, drives schema discussed in class.

```
select demp.eid, trucks.tid, trucks.color
from demp, trucks, drives
where trucks.tid = drives.tid AND demp.eid = drives.eid;
```

Re-write the FIRST LINE of the query to provide an optimizer hint that demp and drives should be joined before trucks and drives.