Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

Topics: SQL Queries Optimization, Transaction Management, Hash Indexes

Instructions:

- Complete Section 1 and submit
  1. *4-FirstLastName-Lab.txt (spool file for Part A)*
  2. *4--FirstLastName-Lab_T1.txt (spool file for transaction 1 for Part B)*
  3. *4--FirstLastName-Lab_T2.txt (spool file for transaction 2 for Part B)*
  4. *4-FirstLastName-Lab.pdf (explanation of spool files for Part A and B)*

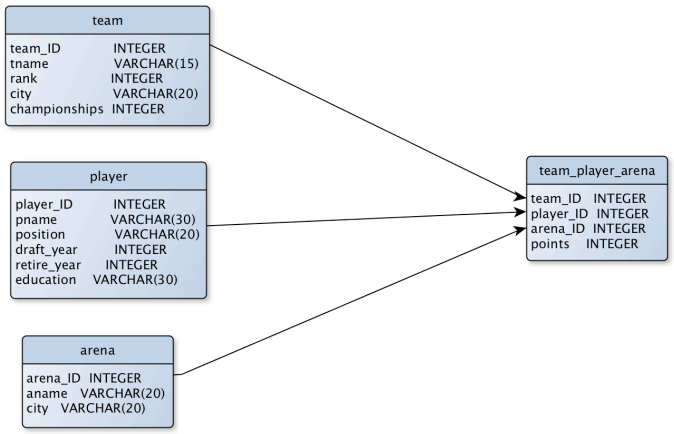
- Complete Section 2 and submit *4-FirstLastName.pdf*

Section 1:

This section covers the practical implementation of a database schema. (50 points)

- Login in to the Linux Oracle server and access the database using *sqlplus* (refer to Oracle Linux Server Instructions)
- Display the SQL commands using: SET ECHO ON
- Create your homework submission log file using: SPOOL *4-FirstLastName-Lab.txt*
- To terminate log file use: SPOOL OFF command

Using this schema that you have already created answer the questions that follow:



- The *team* table has a *team_id* primary key
- The *player* table has a *player_id* primary key
- The *arena* table has an *arena_id* primary key
- The *team_play_arena* table has a (*team_id, player_id, arena_id*) composite primary key
- The *team_play_arena* table has a *team_id* foreign key column that references the *team* table
- The *team_play_arena* table has a *player_id* foreign key column that references the *player* table

Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

- The *team_play_arena* table has a *arena_id* foreign key column that references the *arena* table

Part A: Query Optimization

- To see execution plans for queries use: SET AUTOTRACE TRACEONLY EXPLAIN

1. Run the following two queries:

Query 1:
```
select p.pname from player p, team_player_arena tpa;
```

Query 2:
```
select p.pname from player p, team_player_arena tpa
 where p.player_id = tpa.player_id;
```

(10 points) Looking at the explain plan for each query, which one is less efficient and why?

2. Run the following two queries:

Query 1:
```
select a.aname, a.city
from   arena a, team_player_arena tpa, team t
where  tpa.team_id = t.team_id
and    tpa.arena_id = a.arena_id;
```

Query 2:
```
select distinct a.aname, a.city
from   arena a, team_player_arena tpa, team t
where  tpa.team_id = t.team_id
and    tpa.arena_id = a.arena_id;
```

(10 points) Looking at the explain plan for each query, which query costs less and why?

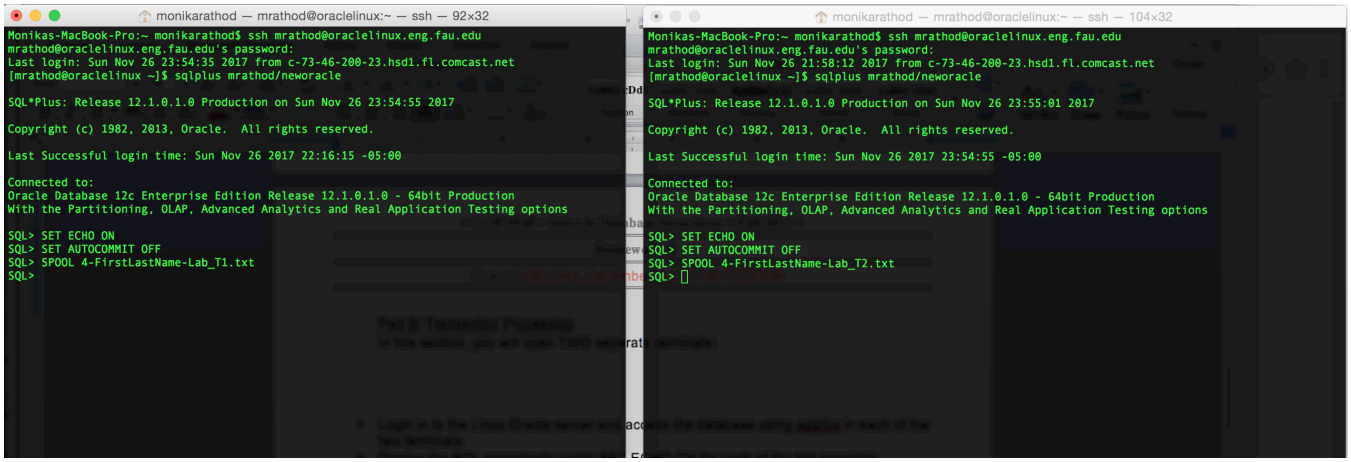
- To turn off explain plan for queries: SET AUTOTRACE OFF

Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

Part B: Transaction Processing

In this section, you will open TWO separate terminals:



- Login in to the Linux Oracle server and access the database using *sqlplus* in each of the two terminals
- Display the SQL commands using: SET ECHO ON for each of the two terminals
- To enable manual commit for changes: SET AUTOCOMMIT OFF for each of the two terminals
- Create transaction1 log file using in 1st terminal: SPOOL *4-FirstLastName-Lab_T1.txt*
- Create transaction 2 log file using in 2nd terminal: SPOOL *4-FirstLastName-Lab_T2.txt*

1. Execute the following statements and observe:

	Terminal 1	Terminal 2
1	insert into team values (7, 'Lakers' , 3, 'LA', 0);	
2		insert into team values (7, 'Knicks' , 8, 'NY', 0);
3	commit;	
4		commit;

a. (5 points) What happens after step 2 and why?
b. (5 points) What happens after step 3 and why?

Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

2. Execute the following statements and observe:

	Terminal 1	Terminal 2
1	insert into team values(8, 'Knicks' , 8 , 'NY' , 0);	
2		insert into team values (9, 'Heat', 12 , 'MI', 0);
3	select * from team;	
4		select * from team;
5	commit;	
6		select * from team;
7		rollback;
8		select * from team;
9	select * from team;	
10	insert into team values (9, 'Celtics', 10 , 'BN', 0);	
11		insert into arena values (8, 'Pepsi Center' , 'Denver');
12	insert into arena values(9, 'Oracle Arena', 'Oakland');	
13		insert into team values (10, 'Pelicans', 15, 'NO', 0);
14	rollback;	
15	select * from team;	
16	select * from arena;	
17		select * from team;
18		select * from arena;
19		commit;
20		select * from team;
21		select * from arena;

a. (10 points) What happens during steps 1 to 4 and why?
b. (5 points) What happens at step 6 and why?
c. (5 points) What happens after step 8 and why?
d. (5 points) What happens after step 13 and why?
e. (5 points) What happens after step 14 and why?
g. (5 points) What happens after step 18 and why?
f. (5 points) What happens after step 19 and why?

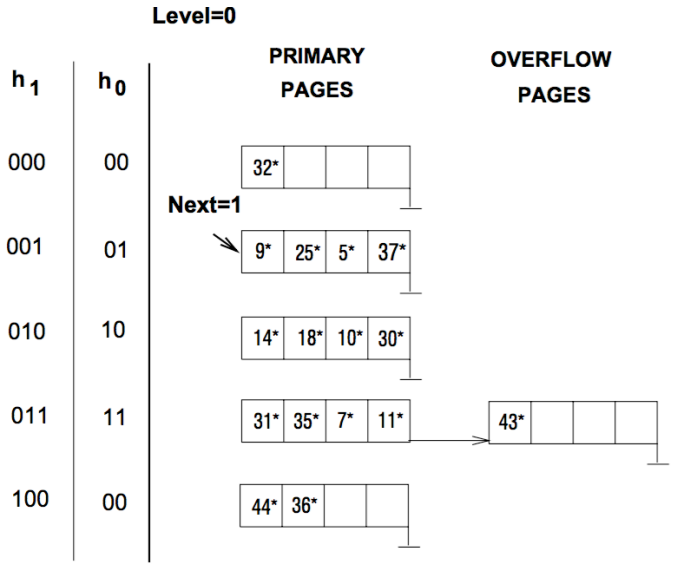
Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

Section 2:

This section covers query optimization and hash indexes (50 points).

1. Using the following Hash index, answer the questions that follow:



Show the resulting hash index after

a. (5 points) Inserting data entry 22*
b. (5 points) Inserting data entry 17*

Note: Use the results from the previous question as a starting point for the next question.

2. Using the following schema:

Suppliers(sid, sname, address)
Parts(pid, pname, color)
Catalog(pid, sid, cost, rating)

consider the following query:

```
SELECT s.sname, c.cost
FROM suppliers s, parts p, catalog c
WHERE p.pid = c.pid AND c.sid = s.sid
AND c.rating < 5 AND c.cost > 15 AND p.name = 'screw driver'
```

a) (10 pts) Draw a query tree using selection, projection and cross product (σπ×).
b) (10 pts) Draw a query tree that optimizes the SQL query using pushing selections.

Homework 4

Due on Saturday, December 2, 2017 at 11:59 PM

3. Using two relations R and S, with the join condition $R_i = S_j$, and given the following information:

- M pages in R with pR tuples per page
- N pages in S with pS tuples per page

- R contains:
  - 2000 pages
  - 150 tuples per page
- S contains:
  - 1000 pages
  - 90 tuples per page
- Buffer size is 102 (inclusive of 2 additional buffers)

a) (10 points) Compute the I/O cost for R and S using a block nested loop join.
b) (10 points) Compute the I/O cost for R and S using a hash join.