

CPS216 Data-Intensive Computing Systems - Fall 2010

Written Assignment 2

- Total points = 100. Due date: Tuesday, Nov. 30, 2010 (5.00 PM).
 - Submission: In class, or email solutions in pdf or plain text to gang@cs.duke.edu. You can also drop off the solutions at Gang's office: N303A North Building.
 - Do not forget to indicate your name on your submission.
 - State all assumptions. For questions where descriptive solutions are required, you will be graded both on the correctness and clarity of your reasoning.
 - Email questions to gang@cs.duke.edu and shivnath@cs.duke.edu.
-

Question 1

Points 15

Pig has an ILLUSTRATE operator/command that Chris Olston covered in his talk. You can find more details on the ILLUSTRATE operator in: (a) Page 307 of Tom White's book, and (b) the paper that describes the algorithm used by the ILLUSTRATE operator available at: <http://research.yahoo.com/pub/2807>

Consider the following Pig Latin script taken from:
http://developer.yahoo.com/blogs/hadoop/posts/2010/01/comparing_pig_latin_and_sql_fo/

```
Users           = load 'users' as (name, age, ipaddr);
Clicks          = load 'clicks' as (user, url, value);
ValuableClicks = filter Clicks by value > 0;
UserClicks     = join Users by name, ValuableClicks by user;
Geoinfo        = load 'geoinfo' as (ipaddr, dma);
UserGeo        = join UserClicks by ipaddr, Geoinfo by ipaddr;
ByDMA          = group UserGeo by dma;
ValuableClicksPerDMA = foreach ByDMA generate group, COUNT(UserGeo);
store ValuableClicksPerDMA into 'ValuableClicksPerDMA';
```

- (a) What are the important challenges faced by the ILLUSTRATE operator for a Pig Latin script like the one above?
- (b) Give one complete example of the output that the ILLUSTRATE operator may produce for the above Pig Latin script when we invoke ILLUSTRATE ValuableClicksPerDMA. Present your answer like Figure 2 in the paper available at: <http://research.yahoo.com/pub/2807>. You can also use the format used on Page 308 of Tom White's book.

Question 2

Points 15

Consider the following SQL query over tables $R(a, b, c)$, $S(a, d)$, and $T(b, e)$, and $U(c, d)$.

```
Q1: Select R.a
From   R, S, T, U
Where  R.a = S.a and R.b = T.b and R.c = U.c and S.d = U.d and R.a < 20
```

[1.] Explain how the Selinger algorithm that we learned in class will optimize Query Q1. Draw the lattice. Don't forget to take interesting orders into account.

Now consider the following SQL query over the same tables:

```
Q2: Select DISTINCT R.a
     From   R, S, T, U
     Where  R.a = S.a and R.b = T.b and R.c = U.c and S.d = U.d and R.a < 20
```

[2.] How will you extend the Selinger algorithm to optimize Query Q2?

Finally, consider the following SQL query over the same tables:

```
Q3: Select DISTINCT T.e
     From   R, S, T, U
     Where  R.a = S.a and R.b = T.b and R.c = U.c and S.d = U.d and R.a < 20
```

[3.] How will you extend the Selinger algorithm to optimize Query Q3?

Question 3

Points 15

The following information is available about relations R and S:

- Relation R is clustered and the blocks of R are laid out contiguously on disk. $B(R) = 1250$ and $T(R) = 12,500$.
 - Relation S is clustered and the blocks of S are laid out contiguously on disk. $B(S) = 1000$ and $T(S) = 10000$.
 - $M = 101$ blocks.
- a. For this question, assume that our cost model is the same as the one we have been using in class, namely, the total number of blocks read or written, excluding the writes for the final output. Compute the number of buckets and the cost for the most efficient Hybrid Hash Join of relations R and S.
- b. Suppose everything in the question remains the same except now $M=51$. Compute the number of buckets and the cost for the most efficient Hybrid Hash Join of relations R and S.

Question 4

Points 15

Suppose a database system reboots after a crash and finds that both *A* and *B* on disk have the value 10. The log is found to be:

```
<START T1>
<START T2>
<T1, A, 5>
<COMMIT T1>
<T2, B, 5>
<T2, A, 15>
[CRASH]
```

1. If the system is using UNDO logging, then give the initial state of the database before T1 and T2 began executing (i.e., what were the initial values of A and B on the disk?).
2. If the system is using UNDO logging, then what will be the final state of the database after recovery (i.e., what will be the values of A and B on the disk after the recovery process has finished?).

3. If the system is using REDO logging, then give the initial state of the database before T1 and T2 began executing (i.e., what were the initial values of A and B on the disk?).
4. If the system is using REDO logging, then what will be the final state of the database after recovery (i.e., what will be the values of A and B on the disk after the recovery process has finished?).

Question 5

Points 15

Assume that a database system using UNDO/REDO logging and nonquiescent checkpointing crashes with the log records on disk given below. Record $\langle T, X, v, w \rangle$ means that transaction T changed the value of database element X; its former value was v, and its new value is w.

```

<START T1>
<T1, X, 14, 28>
<T1, Y, 15, 5>
<START T2>
<T2, Z, 20, 10>
<COMMIT T1>
<START CHKPT (T2)>
<T2, W, 4, 7>
<START T3>
<END CHKPT>
<T3, X, 28, 17>
<COMMIT T2>

```

1. What are all of the possible values on disk for each of the database elements W, X, Y and Z?
2. Which, if any, transactions will need to be redone in the recovery process?
3. How would your answers to parts (1) and (2) change if $\langle \text{END CHKPT} \rangle$ were not present in the log?

Question 6

Points 15

Consider the following transaction log from the start of the run of a database system that is capable of doing UNDO/REDO logging with checkpointing:

```

1) <START T1>
2) <T1, A, 45, 10>
3) <START T2>
4) <T2, B, 5, 15>
5) <T2, C, 35, 10>
6) <T1, D, 15, 5>
7) <COMMIT T1>
8) <START T3>
9) <T3, A, 10, 15>
10) <START CHKPT (T2, T3)>
11) <T2, D, 5, 20>
12) <COMMIT T2>
13) <END CHKPT>
14) <START T4>
15) <T4, D, 20, 30>
16) <T3, C, 10, 15>
17) <COMMIT T3>
18) <COMMIT T4>

```

Assume the log entries are in the format <Tid, Variable, Old value, New value>. What are the values of the data items A, B, C, and D on disk after recovery:

1. If the system crashes just before line 6 is written to disk?
2. If the system crashes just before line 10 is written to disk?
3. If the system crashes just before line 12 is written to disk?
4. If the system crashes just before line 13 is written to disk?
5. If the system crashes just before line 16 is written to disk?
6. If the system crashes just before line 18 is written to disk?

Question 7

Points 10

1. Explain briefly in your own words the advantages and disadvantages of HBase over a Relational Database Management System (RDBMS).
2. We talked about UNDO, REDO, and UNDO/REDO logging in class. What type of logging does HBase use? Explain briefly.