Introduction to Database Systems – Midterm

Spring 2021, Duration: 100 minutes

1. Explain the following terms:
   (1) ACID (Please also explain what each word means) (4%)
   (2) Plan & Scan (4%)
   (3) Phantom Read (4%)
   (4) Cascading Delete (4%)

2. Consider the following steps, called write-ahead-logging (WAL), for writing a record into disk in a DBMS:
   (1) Write log to log buffer;
   (2) Write data to buffer;
   (3) Flush log buffer to disk;
   (4) Flush data buffer to disk.

   Is it valid to switch steps (3) and (4)? Explain your reason. (4%)

3. Why does a DBMS usually manage the block caches by itself rather than relying on the memory management function (e.g., paging system) provided by the OS? (5%)

4. What benefit does Strict 2 Phase Locking (S2PL) provide? What disadvantages result? (5%)

5. Consider the table below.

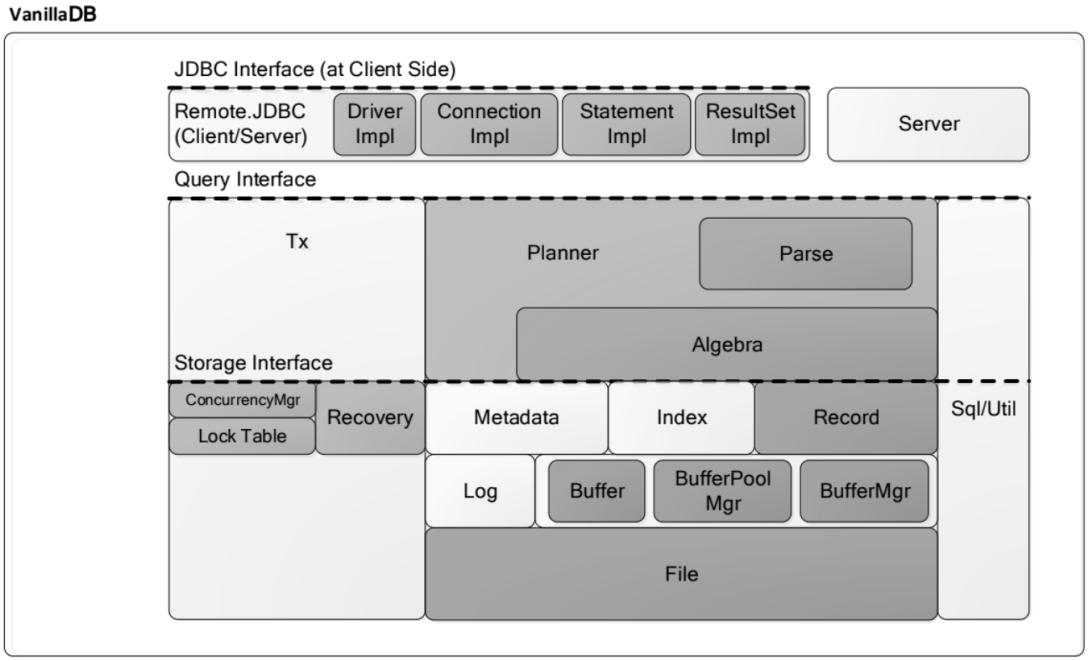
Student ID	Hometown	Zip	Course ID, Score
75312	Taichung City	400	(S5302, 89), (S5381, 98)
75534	Kaohsiung City	800	(S5302, 88)
75302	Hsinchu City	300	(S5345, 90), (S3581, 84)

   Please decompose the table in order to make it agree with 3rd normal form while preserving the following functional dependency:

   (1) Student ID → { Hometown, (Course ID, Score) }
   (2) Hometown → Zip

   Note that you should draw your answer in the form of tables along with records. (10%)

6. Given a general DBMS architecture below:



Among the shaded components, which are thread-safe? (thread-safe: safe for sharing between threads) (6%)

7. Consider the following code in VanillaCore.

```
RandomAccessFile f = new RandomAccessFile(file, "rws");
```

What does "s" stand for in "rws"? What is the effect? (5%)

8. Given the following schedule:

Transaction 1	Transaction 2
Read A	
Write A	
	Read A
	Read B
Write B	
	Write B
	Commit
Commit	

Is this schedule serializable? Explain why if your answer was "no". (5%)

Is this schedule recoverable? Explain why if your answer was "no". (5%)

9. Given a table named “sailors” as shown in the following table.

s-id	*s-name*	*rating*	*age*
22	Dustin	7	45.0
29	Brutus	10	33.0
31	Lubber	8	55.5
32	Andy	5	25.5
58	Rusty	1	35.0
71	Zorba	10	16.0
75	Horatio	4	35.0
84	Art	13	25.5
93	Bob	4	63.5

(1) Show the result of the following query. (4%)

```
SELECT  s.rating, AVG(s.age) AS avgage
FROM    sailors s
GROUP BY    s.rating
```

(2) Show the result of the following query. (5%)

```
SELECT  s.rating, MIN(s.age) AS minage
FROM    sailors s
WHERE   s.age >= 18
GROUP BY    s.rating
HAVING  COUNT(*) > 1
```

10. Consider the following sequence of actions, listed in the order they are submitted to the DBMS.

T1 : R(X) , T2 : W(X) , T2 : W(Y) , T3 : W(Y) , T1 : W(Y) ,

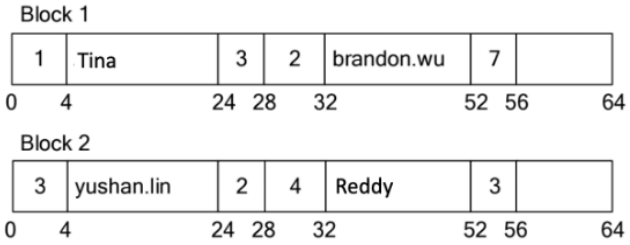
T1 : Commit , T2 : Commit , T3 : Commit

Assume that the DBMS use S2PL as its concurrency control mechanism. Please show the order of these actions **actually performed** in the system. (10%)

11. Suppose each data block has size 64 bytes. Given the following table:

Field Name	Type	Max Size (in bytes)
s_id	int	4
s_name	varchar(10)	20 (2 bytes/char)
grade	int	4

A DBMS stores unspanned, fixed-length records in a block and gets a block layout as follows (Note that we ignore INUSE flags here):



Draw alternative block layouts if

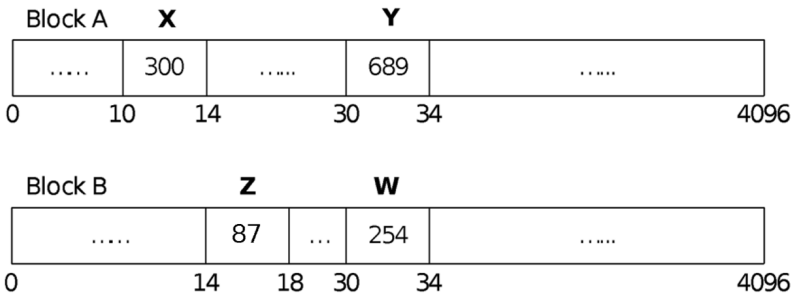
(1) Unspanned, variable-length records are stored. (5%)

(2) Spanned, variable-length records are stored. We assume that values are also spanned as well (5%)

12. Suppose a DBMS is shut down due to power outage. Consider the following log records on the disk:

```
<Tx.11, Start>
<Tx.12, Start>
<Tx.11, Block B, offset 14, 11, 15>
<Tx.12, Block A, offset 10, 200, 300>
<Tx.13, Start>
<Tx.11, Commit>
<Tx.13, Block A, offset 30, 17, 689>
<Tx.12, Block B, offset 14, 15, 87>
<Checkpoint, 12, 13>
<Tx.14, Start>
<Tx.12, Block A, offset 10, 300, 999>
<Tx.12, Block B, offset 30, 254, 361>
<Tx.12, Commit>
<Tx.14, Block B, offset 14, 87, 52>
<Tx.13, Block A, offset 10, 999, 700>
<Tx.13, Abort>
<Tx.15, Start>
<Tx.15, Block B, offset 30, 361, 11>
<Tx.15, Commit>
```

and blocks on the disk:



Use the undo-redo algorithm to recover the system state, and write down the values of X, Y, Z, and W after recovery? (10%)