

# Database Management Systems

Lecture 13

Problems

I

Let R and S be 2 relations. R has 10,000 records; a page can hold 10 R records. S has 2,000 records; a page can hold 10 S records.

1. 52 buffer pages are available. Compute the cost of:

```
SELECT *  
FROM R INNER JOIN S ON R.a = S.b
```

using *page-oriented nested loops join* and *block nested loops join*; S is the outer relation.

R – 10,000 records; a page can hold 10 R records => 1,000 pages

S – 2,000 records; a page can hold 10 S records => 200 pages

- page-oriented nested loops join
  - $200 + 200 * 1000 = 200,200$  I/Os
- block nested loops join
  - block size: 50 =>  $\left\lceil \frac{200}{50} \right\rceil = 4$  S blocks
  - $200 + 4 * 1000 = 4,200$  I/Os

Let R and S be 2 relations. R has 10,000 records; a page can hold 10 R records. S has 2,000 records; a page can hold 10 S records.

2. Compute the cost of sorting R using *external merge sort* with 200 buffer pages.

- $2 * 1000 * 2 = 4,000$  I/Os
- $2 * N * \left( \left\lceil \log_{B-1} \left\lceil \frac{N}{B} \right\rceil \right\rceil + 1 \right)$  I/Os

Let R and S be 2 relations. R has 10,000 records; a page can hold 10 R records. S has 2,000 records; a page can hold 10 S records.

3. R is stored at București, S is stored at Cluj-Napoca. Compute the cost of:

```
SELECT *  
FROM R INNER JOIN S ON R.a = S.b
```

using *simple nested loops join (tuple-oriented)* in Cluj-Napoca, without caching; S is the outer relation.

- $t_d$  time to R / W a page from / to disk
- $t_s$  time to ship a page
  - $200t_d + 2000 * 1000 (t_d + t_s) = 200t_d + 2,000,000 (t_d + t_s)$   
 $= 2,000,200t_d + 2,000,000t_s$

4. Encode the data *de gustibus non disputandum* using the secret encryption key *metallica* and the table of codes below. Write the last 5 characters in the result.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	-
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

- see lecture 5

||



1. T1 and T2 are 2 concurrent transactions, both active at time  $t$ . Choose the correct answer(s):
- a. The following execution describes a *write read* conflict: At time  $t$ , T2 is reading a data object previously written by T1.
  - b. The following execution describes a *write read* conflict: At time  $t$ , T2 is writing a data object previously read by T1.
  - c. The following execution describes a *read write* conflict: At time  $t$ , T2 is reading a data object previously written by T1.
  - d. The following execution describes a *read write* conflict: At time  $t$ , T2 is writing a data object previously read by T1.
  - e. none of the above answers is correct.

2. A schedule S:

- a. is conflict serializable if and only if its precedence graph has exactly one cycle.
- b. is conflict serializable if and only if its precedence graph is acyclic.
- c. is conflict serializable if and only if its precedence graph has exactly two cycles.
- d. is conflict serializable if and only if its precedence graph has exactly three cycles.
- e. none of the above answers is correct.

3. In SQL Server, under the READ UNCOMMITTED isolation level:
- a. S locks must be acquired to perform read operations.
  - b. read operations are performed without acquiring S locks.
  - c. X locks must be acquired to perform write operations.
  - d. write operations are performed without acquiring X locks.
  - e. none of the above answers is correct.

4. In horizontal fragmentation:

- a. the reconstruction operator is the natural join.
- b. the union of the horizontal fragments must be equal to the original relation.
- c. fragmentation is performed with projection operators.
- d. fragmentation is performed with selection predicates.
- e. none of the above answers is correct.

5. I is an index with search key  $\langle C1, C2, C3, C4 \rangle$ .

a. If I is a hash index, I matches condition  $C1 > 10 \text{ AND } C2 > 7$ .

b. If I is a hash index, I matches condition  $C1 = 10 \text{ AND } C2 = 7 \text{ AND } C3 = 1 \text{ AND } C4 = 5$ .

c. If I is a B+ tree index, I matches condition  $C1 = 10 \text{ AND } C2 = 7$ .

d. If I is a B+ tree index, I matches condition  $C2 = 7 \text{ AND } C3 = 9$ .

e. none of the above answers is correct.

6. Let R be a relation with P pages. The cost of sorting R using *simple two-way merge sort* (i.e., with 3 pages in the buffer pool) is:

- a.  $\pi^P$
- b.  $2P(\lceil \log_4 P \rceil + 1)$
- c.  $2P(\lceil \log_2 P \rceil + 1)$
- d.  $2P(\lceil \log_3 P \rceil + 1)$
- e. none of the above answers is correct.

7. Consider the query:

SELECT \*

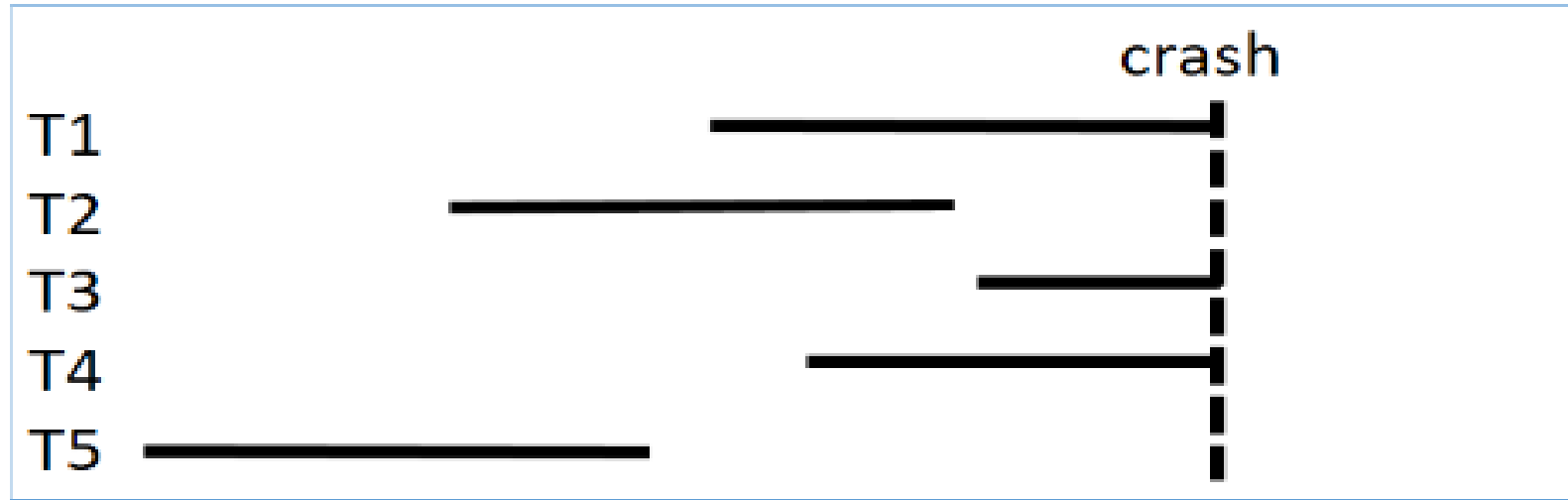
FROM R1, R2, R3

WHERE p1 AND p2 AND p3

The conditions tested by the predicates in the WHERE clause are statistically independent. The cardinality of a relation R is denoted by  $|R|$ . The reduction factor associated with predicate p is denoted by  $RF(p)$ . The cardinality of the query's result set can be estimated by:

- a.  $\frac{|R1| * |R2| * |R3|}{RF(p1) + RF(p2) + RF(p3)}$
- b.  $|R1| * |R2| * |R3| * RF(p1) * RF(p2) * RF(p3)$
- c.  $RF(p1) * RF(p2) * RF(p3) - (|R1| + |R2| + |R3|)$
- d.  $|R1| + |R2| + |R3| + RF(p1) + RF(p2) + RF(p3)$
- e. none of the above answers is correct.

8. Consider the execution below. When the system comes back up after the crash, it must ensure that:



- a. T1, T3, T4 are durable; T2 and T5 are undone.
- b. T1, T3, T4 are undone; T2 and T5 are durable.
- c. T1 is undone only if T2 and T4 are also undone.
- d. T2 is durable only if T5 is undone.
- e. none of the above answers is correct.



9. In data replication:

- a. *primary site replication* is an asynchronous replication technique.
- b. *primary site replication* is a synchronous replication technique.
- c. *read-any write-all* is a synchronous replication technique.
- d. *read-any write-all* is an asynchronous replication technique.
- e. none of the above answers is correct.

10. A database access request contains:

- a. the requesting user.
- b. the criminal record of the requesting user.
- c. the operation the user wants to perform.
- d. the requested object.
- e. none of the above answers is correct.

11. Consider schedule S below over transactions T1, T2, T3, T4 (all transactions commit):

T1	T2	T3	T4
W(A)			
			R(C)
	R(B)		
		W(D)	
	R(A)		
R(D)			
			W(B)
R(C)			

- a. S is conflict serializable.
- b. S is not conflict serializable.
- c. (R(T4, C), R(T1, C)) belongs to the conflict relation of S.
- d. (W(T1, A), R(T2, A)) belongs to the conflict relation of S.
- e. none of the above answers is correct.