Department of Computer Science

Prof. Dr. Nikolaus Augsten
Jakob-Haringer-Str. 2
5020 Salzburg, Austria
Telefon: +43 662 8044 6347
E-Mail: nikolaus.augsten@plus.ac.at

Paris Lodron Universität Salzburg

Non-Standard Database Systems — Exam
Summer Semester 2022/2023 — 12.07.2023

Name: ______________________ **Student ID:** ____________________

Hints

- Check whether you received all pages of the exam (9 pages).
- Write your name or your student ID on each sheet of the exam and hand in all pages.
- All answers are expected to be written on the exam sheet.
- Number any additional pages that you use for longer answers and reference them clearly visible on the corresponding exercise sheet.
- Use only pencils that are permanent. No red pencil.
- Use the notation and techniques discussed during the lecture.
- Exercises with more than one solution are not graded.
- You may use one A4 sheet with your personal notes (on both sides, hand written or printed).
- Exam duration: 90 minutes

Signature ______________________

Grading

Filled by the examiner

Exercise	1	2	3	4	5	6	7	8	Summe
Total Points	2	2	2	2	2	2	2	2	16
Points Reached									

Exercise 1 - *Throughput and Response Time.* 2 Points

A database system must process transactions t_1, t_2, t_3, and t_4. Assume that shorter transactions are given priority over longer transactions, i.e., the waiting queue is ordered by execution time, and no transactions are aborted.

Fill in the start time (when the transaction enters the system) and the execution time for each transaction such that the avg. response time is x seconds and the throughput is $x\frac{1}{sec}$ for some value of x that you can choose.

Transaction	Start Time	Execution Time
t_1		
t_2		
t_3		
t_4		

Exercise 2 - *Parallel System Architecture.* 2 Points

Mark the following statements as true (**T**) or false (**F**).

1. The shared memory and the shared disk architecture are identical. ☐
2. NUMA allows processors to access the memory of other processors through a high-speed network. ☐
3. Shared memory scales to larger number of processors than shared nothing architectures. ☐
4. The top level of hierarchical architectures is shared nothing. ☐

Exercise 3 - *Horizontal Partitioning.* 2 Points

Relation $r[A]$ in Table 1 should be horizontally partitioned onto four disks, D_i, $0 \leq i \leq 3$. Partition the tuples on attribute A using

1. round-robin (with processing order left-to-right in Table 1), and
2. hash partitioning with hash function $h(a) := a \bmod 4$.

Further, answer the following questions:

3. What is the downside of hash partitioning compared to round-robin?
4. For which type of queries is hash partitioning preferable over round-robin?

A	30	72	54	46	66	34	42	60	10	22	84	96

Table 1: Relation $r[A]$.

Exercise 4 - *2-Phase-Commit (2PC).* 2 Points

Transaction T is initiated at site S_i with coordinator C_i, $1 \leq i \leq n$, and is executed at sites S_k, $1 \leq k \leq n$. Discuss the situations when a message between coordinator C_i and site S_k gets lost in

1. Phase 1 of the protocol, and
2. Phase 2 of the protocol.

Exercise 5 - *Persistent Messaging.* 2 Points

Consider a sender S that sends a message to receiver R using the persistent messaging protocol. Table 2 shows the initial entries in the relations *messages_to_send* of the sender and *received_messages* of the receiver. Newer events have larger time stamps.

messages_to_send

number	message	time	ack
1	$Q \leftarrow Q + 9$	2	received
3	$A \leftarrow A + 3$	3	received
7	$Q \leftarrow Q + 3$	5	
8	$B \leftarrow B - 9$	7	received
9	$C \leftarrow C - 6$	8	

received_messages

number	message	time	ack
7	$Q \leftarrow Q + 3$	5	sent
8	$B \leftarrow B - 9$	7	sent
9	$C \leftarrow C - 6$	8	sent

Table 2: Relations *messages_to_send* at sender S and *received_messages* at receiver R.

1. Assume that S receives the acknowledgement for the message with number 9. Compute the value of T_{OLD}.
2. Show the relation *received_messages* after receiver R has received and processed the value of T_{OLD}.

Exercise 6 - *Deadlock Handling*. 2 Points

Construct a scenario with two transactions T_1 and T_2 that results in an unnecessary rollback using 2-phase locking.

Exercise 7 - *Vector clocks.* 2 Points

Assume a single data item Q that is replicated on sites S_1, S_2, and S_3. A site S_i can do (i) a local write on Q, $W(Q)$, which changes the value of the local copy Q_i, or (ii) copy the value from a different site S_j, $C(S_j)$, $j \neq i$, which copies the value of Q_j to Q_i. All vectors are initialized with the zero vector.

1. Show the vector clocks resulting from the schedule in Figure 1.
2. Identify the first operation that requires reconciliation.

Note: Local reads, which will typically precede a local write in a real schedule, are not relevant for conflict detection and omitted from the schedule.

S_1	S_2	S_3
	$W(Q)$	
	$W(Q)$	
$C(Q_2)$		
$W(Q)$		
		$C(Q_2)$
$C(Q_3)$		

Figure 1: Schedule on replicated data item Q.

Exercise 8 - *Parallel Join.* 2 Points

Given a system with 6 processing nodes p_i, $1 \leq i \leq 6$, and two relations $R[A] = [4, 7, 13, 14, 16, 24, 25, 36, 44, 55, 57, 62, 68, 72, 78, 81, 85, 92]$ and $S[A] = [7, 34]$ (each number is an attribute value of A and forms a unary tuple). The tuples of the two relations are round-robin partitioned on the processing nodes p_i.

Find an appropriate parallel join technique for the following query and count:

1. the number of tuples that must be transferred over the network between any pair of nodes;
2. the number of tuple pairs from R and S that must be joined per processing node.

```
SELECT * FROM R, S
WHERE R.A <= S.A;
```