

# Exam: Advanced Database Systems (COMP90050\_2021\_SM1)

Started: Jun 18 at 15:00

## Quiz Instructions

[Links to technical support for the exam](https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/technical-support) [\\_ \(https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/technical-support\)](https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/technical-support)

**Total number of questions: 12**

**Total marks: 50**

### Authorised Materials:

This is an open book exam – the lecture slides, lecture recordings, textbooks, and study materials can be used. Note that there is a strict time limit for the exam, hence you should be mindful of the time spent using such resources.

While you are undertaking this exam you ***must not***:

- make use of any messaging or communications technology.
- act in any manner that could be regarded as providing assistance to another student who is undertaking this assessment, or will in the future be undertaking this assessment.

The work you submit ***must be based on your own knowledge and skills***, without assistance from any other person.

### Instructions to Students:

This exam begins at the time mentioned in 'Available from' in Canvas (lms.unimelb.edu.au). The exam must be completed and submitted by the mentioned 'Due time'. This exam has a strict time limit.

- **Once submitted, you will NOT be able to reopen the quiz to change your answers. This means that you will be able to click the "Submit" button ONLY once.**
- Answers should only contain simple text. You do NOT need to upload any image. DO NOT click anything in the editing toolbar, e.g., changing format, uploading media, adding equation, etc.
- **Questions have unequal marks**, and we encourage you to **attempt all questions**.

## Academic Integrity Declaration

By commencing and/or submitting this assessment I agree that I have read and understood the [University's policy on academic integrity](https://academicintegrity.unimelb.edu.au/#online-exams). [\\_ \(https://academicintegrity.unimelb.edu.au/#online-exams\)](https://academicintegrity.unimelb.edu.au/#online-exams)

I also agree that:

1. Unless paragraph 2 applies, the work I submit will be original and solely my own work (cheating);
2. I will not seek or receive any assistance from any other person (collusion) except where the work is for a designated collaborative task, in which case the individual contributions will be indicated; and,
3. I will not use any sources without proper acknowledgment or referencing (plagiarism).
4. Where the work I submit is a computer program or code, I will ensure that:
  - a. any code I have copied is clearly noted by identifying the source of that code at the start of the program or in a header file or, that comments inline identify the start and end of the copied code; and
  - b. any modifications to code sourced from elsewhere will be commented upon to show the nature of the modification.

**Question 1****2 pts**

Which of the following RAID settings has the equal number of write operations on average among all the disks?

- ☐ RAID 4 with 3 disks
- ☐ RAID 4 with 5 disks
- ☐ RAID 3 with 3 disks
- ☐ RAID 1 with 3 disks

**Question 2****2 pts**

In a distributed database with shared data, the following options show the desired properties in pairs. Which one of these pairs cannot be achieved together at the same time?

- ☐ Consistency and partition tolerance
- ☐ Availability and consistency
- ☐ Availability and partition tolerance

**Question 3****2 pts**

Weather service Australia has installed multiple sensors in Melbourne with a small computing device in them. Each computing devices store and manage the data of its own sensor. When the sensors record any new data, they connect with the other nearby sensors in an ad-hoc network, share those data, and then disconnect. Which of the following database architecture is the most suitable choice for this scenario?

- ☐ P2P databases
- ☐ Cloud storage
- ☐ Centralised database
- ☐ Distributed database

**Question 4****2 pts**

Which system is expected to be available more (e.g., expected to function for longer) than the other system? (i) System A with 5 devices on failfast (ii) System B with 5 devices as a supermodule

- ☐ System B
- ☐ System A
- ☐ Both system A and system B have the same availability

**Question 5****4 pts**

John needs to buy a new computer. He usually uses his computer to play online games, listen to music, watch movies, work on word documents for his studies, and browse social networks. He cannot decide which of the following options of

hardware will be suitable for him. Please explain the advantages and disadvantages of each option for him. Assume both options cost the same.

Option 1: 250 GB solid state drive, 8 MB cache

Option 2: 500 GB hard disk drive, 32 MB cache

### Question 6

4 pts

In a nested transaction, a sub-transaction can commit, but the actual commit will not take place until all its ancestor transactions commit. Then why is a nested transaction still useful?

### ~~Question 7~~

4 pts

In a database, the table 'student\_records' is very frequently accessed by many concurrent transactions and processes, for both reading and writing purposes. In the same database, the table 'classroom\_records' is rarely accessed. What type of semaphore should be used for each of these two database tables? Are these semaphores of the same type for both tables? Why or why not?

### Question 8

4 pts

Many systems employ a simple strategy to handle deadlocks - they just allow a transaction to wait for certain maximum time on a lock, and then force it to rollback. Why is this strategy used by most systems? What are the benefits of this strategy over the other possible strategies to handle deadlocks?

### Question 9

6 pts

Degree 2 isolation may have 'non-repeatable read' type of concurrency problem. What are the benefits of using Degree 2 isolation then? Can you give an example application scenario where a Degree 2 isolation can be useful over a strict Degree 3 isolation.

## Question 10

6 pts

The following transactions are issued in a system at the same time. Answer for both scenarios.

(i) Scenario 1: When the value of the variable some\_input is 3, which of the following transactions can run concurrently from the beginning till commit (that is, all operations and locks are compatible to run concurrently with another one) and which ones need to be delayed? Please give explanation for the delayed transactions.

(ii) Scenario 2: When the value of the variable some\_input is 1, which of the following transactions can run concurrently from the beginning till commit (that is, all operations and locks are compatible to run concurrently with another one) and which ones need to be delayed? Please give explanation for the delayed transactions.

	T2	T3
T1	Lock (IS,A)	Lock (IX,A)
Lock (S,A)	If(some_input == 3){	If(some_input == 3){
Read A	Lock(S,A)	Lock (X,A)
Unlock A	Read A	Write A
	}	}
	Unlock A	Unlock A

A compatibility matrix is as follows -

**Question 11****6 pts**

We have a simplified log at the time of a system crash. Assume that there is no log record before the checkpoint. The format of a log record is (LSN, Operation Details).

(00, begin checkpoint)

(05, end checkpoint)

(10, T1 write Page1)

(20, T2 write Page2)

(30, T1 write Page5)

(40, T2 commit)

(45, T2 end)

(50, T3 write Page5)

(60, T1 abort)

(70, CLR undo T1 LSN 30)

(80, CLR undo T1 LSN 10)

(90, T1 end)

CRASH

The system recovery consists of three phases: analysis, redo and undo. Please answer each of the following questions.

A. What information will be in the dirty page table after the analysis phase (write as a list of the format (Page id, LSN))?

B. If the pageLSN of Page5 stored in the database is found as **30**, then what will be the order of the LSNs to be redone in the Redo phase? Assume that all the necessary pages are in the dirty page table, all LSNs in the log are greater than or equal to the corresponding page's recLSN, and all LSNs in the log are greater than the corresponding page's pageLSN (except for Page5).

C. What is the order of the LSNs to undo in the Undo phase?

## Question 12

8 pts

A company needs to store some sensitive medical records on disk. It is very critical that no error occurs, and no data is lost while storing the records. They can use multiple disks for the storage purpose if necessary. What strategies can they use to minimize the chance of any error while writing the data on disk? What strategies can they use to detect whether any error occurred or not? How can they still use the data if an error is detected? You do not need to describe the detailed steps of the strategies.

Not saved

Submit Quiz