

Database Management System: Assignment 5

Total Marks : 20

August 14, 2023

Question 1

Identify the incorrect statement(s) about the 3-tier application architecture from the following.

Marks: 2 MCQ

- a) The Controller Layer acts as the middle layer between the front and the database.
- b) The Presentation Layer acts as the user interface between users and the database.
- c) The Business Logic Layer acts as the middle layer between the front-end and the database.
- d) The Data Access Layer acts as the backend of the application.

Answer: a)

Explanation: 3-tier Application Architecture consists of 3 sub-layers namely, Presentation Layer, Business Logic Layer, and Data Access Layer.

Frontend or Presentation Layer: user interface

- Forms, graphical user interfaces
- Many interfaces are Web-based or Mobile App

Middle Layer or Application / Business Logic Layer

- Functionality of the Application – links front and backend

Backend or Data Access Layer

- Persistent data, large in volume, needs efficient access

Refer to Module 22 slide 6

Hence, option a) is incorrect.

Question 2

Caching is used to improve web server performance. Which of the following sentences is not true about caching?

Marks: 2 MCQ

- a) Caching is used on the server side for JDBC connection pooling.
- b) Caching is used to increase the amount of data that needs to be sent over the network.
- c) Web proxy caching is used to store copies of frequently accessed Web objects.
- d) Caching is used on the server side for generated HTML.

Answer: b)

Explanation: Caching techniques are used to reduce the cost of serving pages by exploiting commonalities between requests

At the server site:

- Caching of JDBC connections between servlet requests
 - a.k.a. connection pooling

Caching results of database queries

- Cached results must be updated if the underlying database changes

Caching of generated HTML

- At the client's network
 - Caching of pages by Web proxy

Refer to Module 22 slide 16

Hence, option b) is false.

Question 3

A Redundant Arrays of Independent Disk (RAID) system has 8 disks and it uses block level striping to improve reliability. If we have to store a 16 KB file in it and the size of a disk block is 512 bytes. In which disk, the last file block will be stored?

Marks: 2 MCQ

- a) Disk 8
- b) Disk 7
- c) Disk 1
- d) Disk 0

Answer: a)

Explanation: In Block-level striping with n disks, block i of a file goes to disk $(i \% n) + 1$, so that requests for different blocks can run in parallel if the blocks reside on different disks.

In the above system no of disk = 8

Size of the file = 16 KB = 16×2^{10} bytes

One block size = 512 bytes = 2^9

No of blocks required = $2^{14} \div 2^9 = 2^5 = 32$

File blocks are numbered 0, 1, 2, ..., 31.

So, the last block will be 31.

Required block = $(31 \% 8) + 1 = 8$

Hence, option a) is correct.

Question 4

Which of the two alternatives would you choose if you need to support real-time queries that must be answered within a guaranteed short period of time?

Marks: 2 MSQ

- a) Magnetic Hard Disk
- b) Solid State Drive
- c) Redundant Arrays of Independent Disks
- d) Digital Video Disk

Answer: b), c)

Explanation: SSDs have much lower latency for data access, and higher data transfer bandwidth than magnetic disks.

Magnetic disks are vulnerable to failure, which could result in loss of data stored on the disk. We can reduce the likelihood of irretrievable data loss by retaining multiple copies of data. Mirroring reduces the probability of data loss greatly. More sophisticated methods based on redundant arrays of independent disks (RAID) offer further benefits.

By striping data across disks, these methods offer high throughput rates on large accesses; by introducing redundancy across disks, they improve reliability greatly.

So, for real-time queries that must be answered within a guaranteed short period of time, SSD or RAID can be preferred.

Hence, options b) and c) are correct.

Question 5

If a system contains 1,000 disk drives, and each of them has a 6,00,000-hour MTBF (Mean time between failure), how often a single drive failure will occur in that disk system?

Marks: 2 MCQ

- a) 25000 hours
- b) 14400 hours
- c) 600 hours
- d) 25 hours

Answer: c)

Explanation: $\text{MTBF (array)} = \text{MTBF (one disk)} / \text{Number of disks in the array}$

So, the answer is $600,000 / 1000 = 600$ hrs

Hence, option c) is correct.

Question 6

Suppose you have a 64-gigabyte disk pack with the following specifications: 32 double-sided platters, 1024 tracks per surface, and 512 sectors per track. What is the data size of one sector?

Marks: 2 MCQ

- a) 64 Kilobyte
- b) 32 Kilobyte
- c) 4 Kilobyte
- d) 2 Kilobyte

Answer: d)

Explanation: Capacity of the disk (V_D) = capacity of one surface (V_S) * number of platters (N_P) * number of sides per platter

Capacity of each surface (V_S) = capacity of track (V_T) * number of tracks per surface (N_T)

Capacity of a track (V_T) = data size of a sector (V_{Sec}) * number of sectors per track (N_S)

Therefore, data size of a sector (V_{Sec}) = $\frac{V_T}{N_S} = \frac{V_S}{N_T * N_S}$

$$\begin{aligned} &= \frac{V_D}{N_T * N_S * N_P * \text{number of sides per platter}} \\ &= \frac{64 * 2^{30}}{1024 * 512 * 32 * 2} = \frac{2^6 * 2^{30}}{2^{10+9+5+1}} = \frac{2^{36}}{2^{25}} = 2 \text{ Kilobyte} \end{aligned}$$

Hence, option d) is correct.

Question 7

Consider the following string of reference:

33, 25, 81, 42, 33, 25, 50, 33, 25, 81, 42, 50

Find the number of replacements (where an existing value is replaced by a new value because the buffer is full) incurred using the least recently used (LRU) buffer replacement algorithm with 3 empty buffer frames.

Marks: 2 MCQ

- a) 7
- b) 8
- c) 10
- d) 12

Answer: a)

Explanation:

Buffer			Comments
33			33 is added to an empty buffer
33	25		25 is added to an empty buffer
33	25	81	81 is added to an empty buffer
42	25	81	33 is the least recently used. Hence, replaced by 42
42	33	81	25 is the least recently used. Hence, replaced by 33
42	33	25	81 is the least recently used. Hence, replaced by 25
50	33	25	42 is the least recently used. Hence, replaced by 50
50	33	25	No replacement, 33 present
50	33	25	No replacement, 25 present
81	33	25	50 is replaced, as 33, 25 are recently used
81	42	25	33 is replaced, as it is the least recently used element
81	42	50	25 is replaced

Hence, altogether 7 replacements have been done.

So, option a) is the answer.

Question 8

Consider the following schema:

`Movie(movie_id, title, language, genre, director, year, production_company, reg_company)`

`production_company → movie_id, title, language, genre, director, year`

`reg_company → production_company`

If you perform the schema refinement technique on `Movie`, which of the following statements will be true?

Marks: 2 MCQ

- a) Table `Movie` do not need any refinement.
- b) It is better to decompose in -
`Movie(movie_id, title, language, genre, director, year)`
`Production(production_company, reg_company)`
- c) It is better to decompose in -
`Movie(movie_id, title, language, genre)`
`Direction(production_company, director, year)`
`Production(production_company, reg_company)`
- d) It is better to decompose in -
`Movie(movie_id, title, language, genre, director, year, production_company)`
`Production(production_company, reg_company)`

Answer: d)

Explanation: There is redundancy in table `Movie`. So, it must be decomposed into BCNF. In option d) both are in BCNF. Decomposition is lossless and dependency-preserving. Hence, option d) is the answer.

Question 9

Consider the file organization for the relation **Customer**

Customer		
cust_Name	contact_No	Pointer
Smith	116642	1
Pitter	319981	3
Black	824045	NULL
Cherry	524327	4
Kiti	736686	2

contact.No is the primary key field. The field **Pointer** in this file denotes the serial number of the next record in the file. The serial number for the first record is 0 and the **Pointer** field of the last record is set to **NULL**. After insertion of the record {"Thomas", 249137}, the file is as follows.

Customer		
cust_Name	contact_No	Pointer
Smith	116642	5
Pitter	319981	3
Black	824045	NULL
Cherry	524327	4
Kiti	736686	2
Thomas	249137	1

Which type of file organization is used to store the relation **Customer**?

Marks: 2 MCQ

- a) Heap File Organization
- b) Sequential File Organization
- c) Hash File Organization
- d) Multitable Clustering File Organization

Answer: b)

Explanation: Records of **Customer** are stored in the sequential order of **contact_No**, based on the value of the search key of each record.

So, it is a Sequential File Organization.

Hence, option b) is correct.

Question 10

Which of the following statement(s) is (are) correct?

Marks: 2 MSQ

- a) The **disk controller** acts as an interface between the computer system and the disk drive hardware.
- b) When a sector is found to be **bad**, the disk controller remaps the logical sector to a different physical sector.
- c) **Mean time to failure (MTTF)** is the maximum time, for which a disk can run continuously without any failure.
- d) When a huge number of disks are connected by a high-speed network to a number of servers, it is called **Network Area Storage**.

Answer: a), b)

Explanation: The **disk controller** acts as an interface between the computer system and the disk drive hardware. And when a **sector** is found to be **bad**, the disk controller remaps the logical sector to a different physical sector.

So, these two statements are right.

Mean time to failure (MTTF) is the **average time**, when a disk may run continuously without any failure.

When a huge number of disks are connected by a high-speed network to a number of servers, is called **Storage Area Networks**.

Hence, the statements given in options a) and b) are correct.