MODEL ANSWER

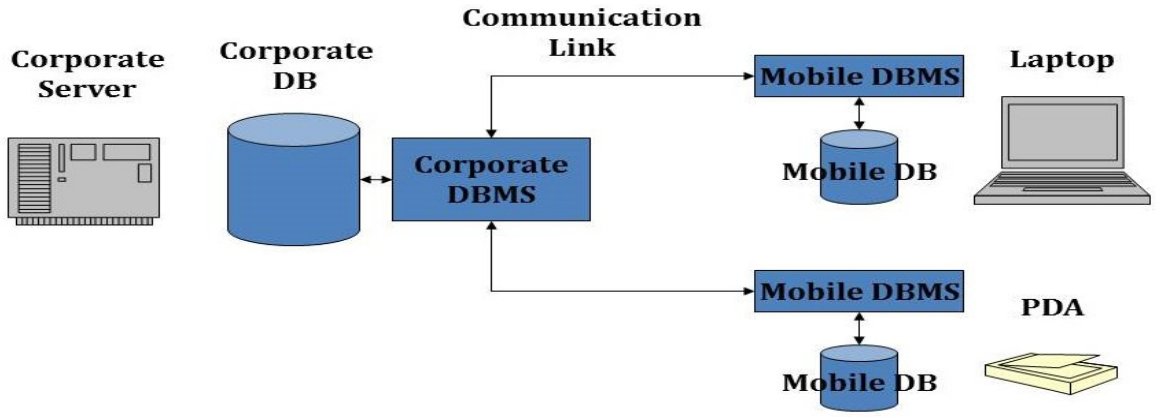
SUMMER– 17 EXAMINATION

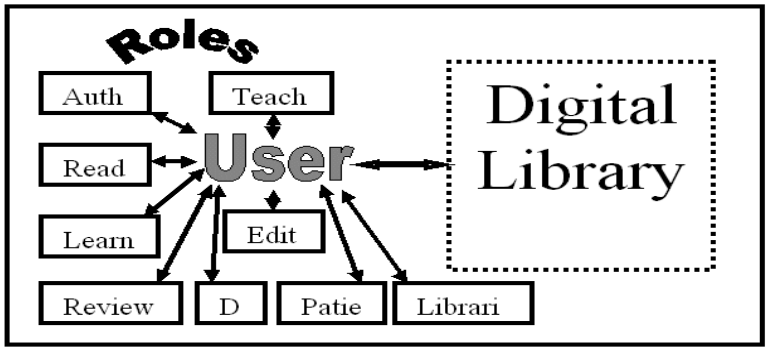
Subject Title: ADVANCED DATABASE MANAGEMENT **Subject Code:** 17631

Important Instructions to examiners:

1) The answers should be examined by key words and not as word-to-word as given in the model answer scheme.
2) The model answer and the answer written by candidate may vary but the examiner may try to assess the understanding level of the candidate.
3) The language errors such as grammatical, spelling errors should not be given more Importance (Not applicable for subject English and Communication Skills.
4) While assessing figures, examiner may give credit for principal components indicated in the figure. The figures drawn by candidate and model answer may vary. The examiner may give credit for any equivalent figure drawn.
5) Credits may be given step wise for numerical problems. In some cases, the assumed constant values may vary and there may be some difference in the candidate's answers and model answer.
6) In case of some questions credit may be given by judgement on part of examiner of relevant answer based on candidate's understanding.
7) For programming language papers, credit may be given to any other program based on equivalent concept.

Q. No.	Sub Q. N.	Answer	Marking Scheme
1.		**Attempt any FIVE :**	**5*4 = 20**
	(a)	**Define immediate update.** **(Explanation for 4 marks)**	**4**
	Ans:	Immediate update also called UNDO/REDO, is also another technique used to recover/support transaction failures that occur due to operating system, power, memory or machine failures. When a transaction runs, any of the updates or alterations made by the transaction are written directly in to the database. Both the original values and the new values are also recorded in the log file before changes are made to the database. On commit all changes made to the database are made permanent and the records in the log file are discarded. On rollback old values are restored in to the database using the old values stored in the log file. All the changes made by transactions to the database are discarded and this process is called "Un-doing". When the system restarts after a crash, all the database changes are made permanent for committed transactions. For uncommitted transactions, original values are restored using the values in the log file.	

	(b)	**Explain Architecture of mobile database.** **(Diagram 2 marks, explanation 2 marks)**	**2+2 = 4**
	Ans:	 **Architecture of Mobile Database** A mobile database is a database that can be connected to by a mobile computing device over a mobile network. A database that is portable and physically separate from the corporate database server. But Mobile Database is capable of communicating with that corporate database server from remote sites allowing the sharing of corporate database. Communicate with centralized database server through modes such as wireless or Internet access; replicate data on centralized database server and mobile device; Synchronize data on centralized database server and mobile device; capture data from various sources such as Internet. Manage/analyze data on the mobile device; create customized mobile applications.	
	(c)	**Explain how to access digital library.** **(Explanation 4 Marks)**	**4**
	Ans:	Ultimately, users will connect through a human-computer interface and interact with the digital library, though in some cases the digital library may be an embedded system that is seen only indirectly (e.g., through a word processor that allows one to search for a quotation). Most commonly, a digital library has an interface for users to search, browse, follow links, retrieve, and read documents.	



For example, the Computer Science Teaching Center, a digital library of courseware about computing, requests that all users, except those just browsing or searching, login to identify themselves. Then it knows and tailors the interface so that suitably authorized users can submit, review, or edit works. Further, CSTC encourages users to submit courseware they have developed that others might download, and supports their entering in suitable metadata as well as uploading their applets, demonstrations, laboratory exercises, interactive multimedia training resources, etc.

(d) **Explain any to multimedia database queries.** **2+2=4**
(each query 2 marks)

Ans:

Image Query (by keywords)

Police officer Rocky wants to examine pictures of "Big Spender".

Query: "Retrieve all images from the image library in which "Big Spender" appears."

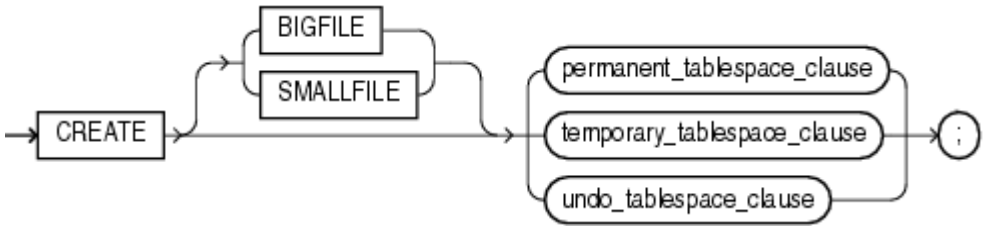
Video Query:

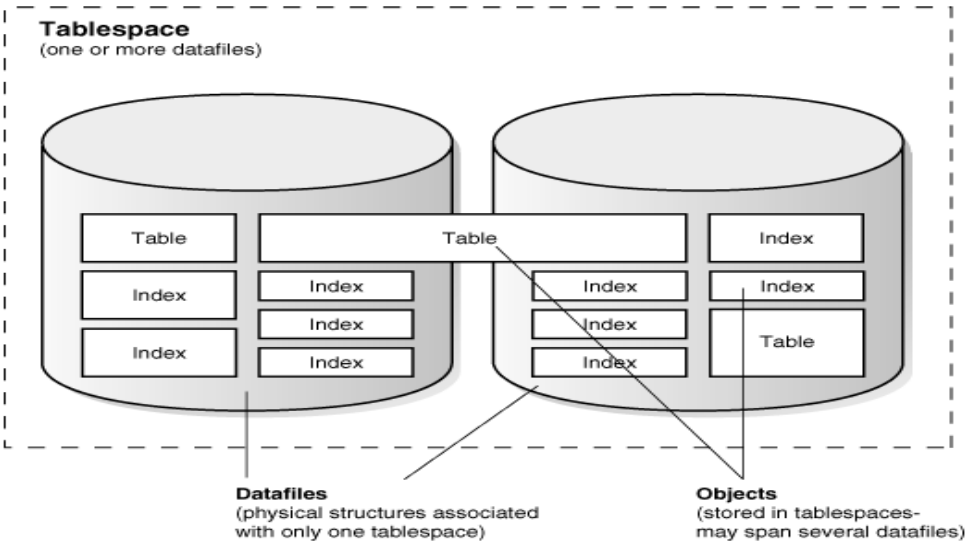
Police officer Rocky is examining a surveillance video of a particular person being fatally assaulted by an assailant. However, the assailant's face is occluded and image processing algorithms return very poor matches. Rocky thinks the assault was by someone known to the victim.

Query: "Find all video segments in which the victim of the assault appears."

By examining the answer of the above query, Rocky hopes to find other people who have previously interacted with the victim.

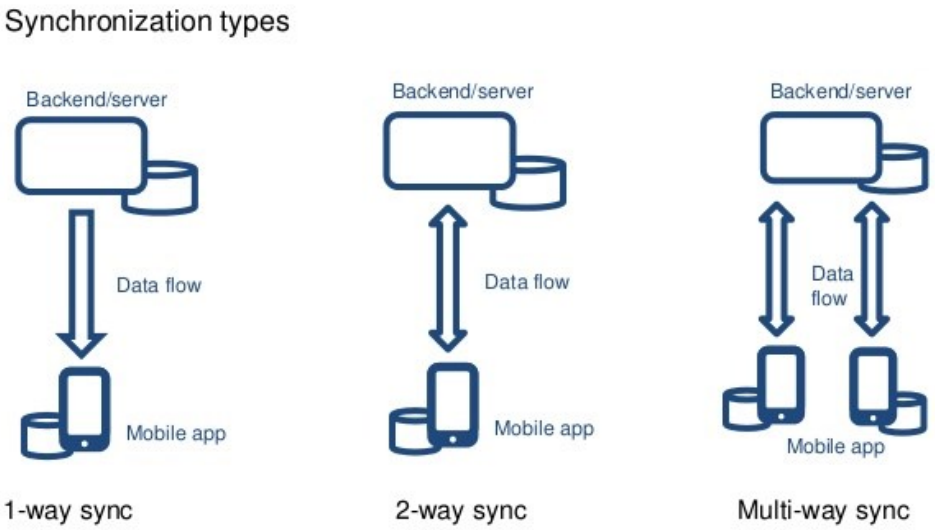
	(e)	**Describe data warehousing.** **(Description 4 marks)**	**4**
	Ans:	A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process. **Subject-Oriented**: A data warehouse can be used to analyze a particular subject area. For example, "sales" can be a particular subject. **Integrated:** A data warehouse integrates data from multiple data sources. For example, source A and source B may have different ways of identifying a product, but in a data warehouse, there will be only a single way of identifying a product. **Time-Variant:** Historical data is kept in a data warehouse. For example, one can retrieve data from 3 months, 6 months, 12 months, or even older data from a data warehouse. This contrasts with a transactions system, where often only the most recent data is kept. For example, a transaction system may hold the most recent address of a customer, where a data warehouse can hold all addresses associated with a customer. **Non-volatile:** Once data is in the data warehouse, it will not change. So, historical data in a data warehouse should never be altered. **Nonupdateable** Data in the data warehouse are loaded and refreshed from operational systems, but cannot be updated by end users. A data warehouse is not just a consolidation of all the operational databases in an organization. Because of its focus on business intelligence, external data, and time-variant data (not just current status), a data warehouse is a unique kind of database. Data warehousing is the process whereby organizations create and maintain data warehouses and extract meaning and inform decision making from their informational assets through these data warehouses.	
	(f)	**Explain how to create a table space.** **(Explanation 2 marks, Syntax 1 mark, example 1 mark)**	**2+1+1= 4**
	Ans:	Use the CREATE TABLESPACE statement to create a **tablespace**, which is an allocation of space in the database that can contain schema objects. A **permanent tablespace** contains persistent schema objects. Objects in permanent tablespaces are stored in **datafiles**. You must have the CREATE TABLESPACE system privilege. To create the SYSAUX tablespace, you must have the SYSDBA system privilege. Before you can create a tablespace, you must create a database to contain it, and the database must be open.	

		Syntax **create_tablespace: =**  **Example** CREATE BIGFILE TABLESPACE bigtbs_01 DATAFILE 'bigtbs_f1.dat' SIZE 20M AUTOEXTEND ON; creates a bigfile tablespace bigtbs_01 with a datafile bigtbs_f1.dat of 10 MB	
	(g)	**State meaning of redundancy in Web-database.** **(Explanation 4 marks)**	**4**
	Ans:	**Data redundancy definition** Data redundancy in database means that some data fields are repeated in the database. This data repetition may occur either if a field is repeated in two or more tables or if the field is repeated within the table. Data can appear multiple times in a database for a variety of reasons. For example, an online shopping may have the same customer's name appearing several times if that customer has bought several different products at different dates. **Disadvantages Of Data Redundancy** 1. Increases the size of the web database unnecessarily. 2. Causes data inconsistency. 3. Decreases efficiency of web database. 4. May cause data corruption. 5. Slow access of web service 6. Web database conflict to retrieve data 7. Errors/confusion in SQL queries	

2.		**Attempt any FOUR :**	**4*4=16**
	(a)	**Explain use of table space.** **(Definition 1 mark, Diagram 1 mark, explanation 2 marks)**	**4**
	Ans:	**Definition**: Oracle stores data logically in **tablespaces** and physically in **datafiles** associated with the corresponding tablespace.  Databases, tablespaces, and datafiles are closely related, but they have important differences: • An Oracle database consists of one or more logical storage units called tablespaces, which collectively store all of the database's data. • Each tablespace in an Oracle database consists of one or more files called datafiles, which are physical structures that conform to the operating system in which Oracle is running. • A database's data is collectively stored in the datafiles that constitute each tablespace of the database. For example, the simplest Oracle database would have one tablespace and one datafile. Another database can have three tablespaces, each consisting of two datafiles (for a total of six datafiles). Allocate More Space for a Database You can enlarge a database in three ways: • Add a datafile to a tablespace • Add a new tablespace • Increase the size of a datafile	

	(b)	**Explain causes of failure in database recovery.** **(each point 1 mark) (any four)**	**1*4=4**
	Ans:	**Database Hardware Failure** A database is just a combination of software and data. Frequently, when a database fails to boot, it's due to hardware failure. The RAID array, server or hard drive has failed, resulting in the inability to access the database. If your database is stored on a RAID array and only one drive in the array has failed, it may be possible for IT professionals to perform a hot swap and replace the failed drive in a RAID 1 configuration or higher**.** **File Corruption** Databases may fail at the file level, which means one or more files in the database have become damaged, causing corruption. Corrupted files represent logical damage to the database and hard drive. **File System Damage** Sometimes, operating system files will become damaged or corrupted if a server or computer is powered down incorrectly, experiences a power surge, or something happens to interrupt the process while data is being written to the files. **Network Failure** When your system uses networks such as local area networks and phone lines to connect client workstations to database servers, or to connect several database servers to form a distributed database system, network failures such as aborted phone connections or network communication software failures can interrupt the normal operation of a database system. **Process Failure** A process failure is a failure in a user, server, or background process of a database instance such as an abnormal disconnect or process termination. When a process failure occurs, the failed subordinate process cannot continue work, although the other processes of the database instance can continue. **User Error** A database administrator can do little to prevent user errors such as accidentally dropping a table. Usually, user error can be reduced by increased training on database and application principles.	

	(c)	**Describe data synchronization in mobile database.** **(Diagram 1 mark, Explanation 3 marks)**	**4**
	Ans:	**Data synchronization** is the process of establishing consistency among data from a source to a target data storage and vice versa and the continuous harmonization of the data over time	



Two-way synchronization of partial

Between mobile-device and personal area computer

For example, whenever the list of contacts and personal information manager data is modified at any of them, it is made consistent after synchronization

Server-alerted synchronization

The server alerts the client the data modification or additions

The client synchronizes the modified or new data by pull request

For example, alerting new e-mail and the device pulling that

One-way server-initiated

Server initiates synchronization of any new modification since communication of last modification

Sends modified data copies to the client

When a new email arrives at a server, it initiates the

		synchronization as and when the device connects to the server and pushes the mail **Client initiated refresh synchronization** The client initiates synchronization with the server for refreshing its existing data copies For refreshing the configuration parameters saved at the server for it	
	(d)	**Explain ways of Mandatory Access Control.** **(Each point 1 mark)**	**1*4=4**
	Ans:	Mandatory Access Control, or MAC, relies on labels that correspond to the sensitivity levels of information for clients and objects. MAC policy compares the sensitivity label at which the user is working to the sensitivity label of the object being accessed and refuses access unless certain MAC checks are passed. **Top Secret (T):** Shall be applied to information, the unauthorized disclosure of which reasonably could be expected to cause exceptionally grave damage to the national security. **Secret (S):** Shall be applied to information, the unauthorized disclosure of which reasonably could be expected to cause exceptionally serious damage to the national security. **Confidential I:** Shall be applied to information, the unauthorized disclosure of which reasonably could be expected to cause exceptionally damage to the national security. **Unclassified (U): no security restriction** Classification is the security level given to information based on policy, We will use the same classification levels used for clearances to assign classifications. Clearance and classification go together, in that, a user's clearance is a limit to the access of information based on the information's classification. Classification relation is null < U < C < S < T. Each security level is said to dominate itself and all others below it in this hierarchy	

	(e)	**List any four key benefits of three-tire applications.** **(Each point 1 mark)**	**1*4=4**
	Ans:	Three-Tier Architecture provides the following benefits. **Scalability** Each tier can scale horizontally. For example, you can load-balance the Presentation tier among three servers to satisfy more Web requests without adding servers to the Application and Data tiers. **Performance** Because the Presentation tier can cache requests, network utilization is minimized, and the load is reduced on the Application and Data tiers. If needed, you can load-balance any tier. **Availability** If the Application tier server is down and caching is sufficient, the Presentation tier can process Web requests using the cache. **Independency** Managing data is independent from the physical storage. As each tier is independent it is possible to use different sets of developers	
	(f)	**Explain SGA and PGA.** **(SGA 2 marks, PGA 2 marks)**	**2*2=4**
	Ans:	**1. System Global Area (SGA)** The System Global Area (SGA) is a group of shared memory structures, known as SGA components, that contain data and control information for one Oracle Database instance. The SGA is shared by all server and background processes. Examples of data stored in the SGA include cached data blocks and shared SQL areas. **Database buffer cache** Before data stored in the database can be queried or modified, it must be read from a disk and stored in the buffer cache. All user processes connected to the database share access to the buffer cache. For optimal performance, the buffer cache should be large enough to avoid frequent disk I/O operations. **Shared pool** The shared pool caches information that is shared among users: • SQL statements that can be reused • Information from the data dictionary such as user account data, table and index	

descriptions, and privileges
- Stored procedures, which are executable code that is stored in the database

Redo log buffer

This buffer improves performance by caching redo information until it can be written to the physical online redo log files stored on disk. Redo information and online redo log files are discussed in "About Online Redo Log Files".

2. Program Global Area (PGA)

A **Program Global Area (PGA)** is a memory region that contains data and control information for a server process. It is nonshared memory created by Oracle Database when a server process is started.

Access to the PGA is exclusive to the server process. There is one PGA for each server process. Background processes also allocate their own PGAs. The total PGA memory allocated for all background and server processes attached to an Oracle Database instance is referred to as the **total instance PGA memory**, and the collection of all individual PGAs is referred to as the **total instance PGA**, or just **instance PGA**.

The amount of PGA memory used and the contents of the PGA depend on whether the instance is running in dedicated server or shared server mode.

The PGA is used to process SQL statements and to hold logon and other session information. A large part of the PGA is dedicated to **SQL work areas**, which are working memory areas for sorts and other SQL operations.

Session Memory Session memory **is the memory allocated to hold a session's variables (logon information) and other information related to the session. For a shared server, the session memory is shared and not private**

The private SQL area

Contains data such as bind variable values, query execution state information, and query execution work areas. Each session that issues a SQL statement has a private SQL area. Each user that submits the same SQL statement has his or her own private SQL area that uses a single shared SQL area. Thus, many private SQL areas can be associated with the same shared SQL area.

3.		**Attempt any FOUR :**	**4*4=16**
	(a)	**Explain object Privilege.** **(Definition 1 mark, privileges 3 marks)**	**1+3=4**
	Ans:	An object privilege is the right to perform a particular action on an object or to access another user's object. Objects include tables, views, materialized views, indexes, synonyms, sequences, cache groups, replication schemes and PL/SQL functions, procedures and packages. **Object privileges**	

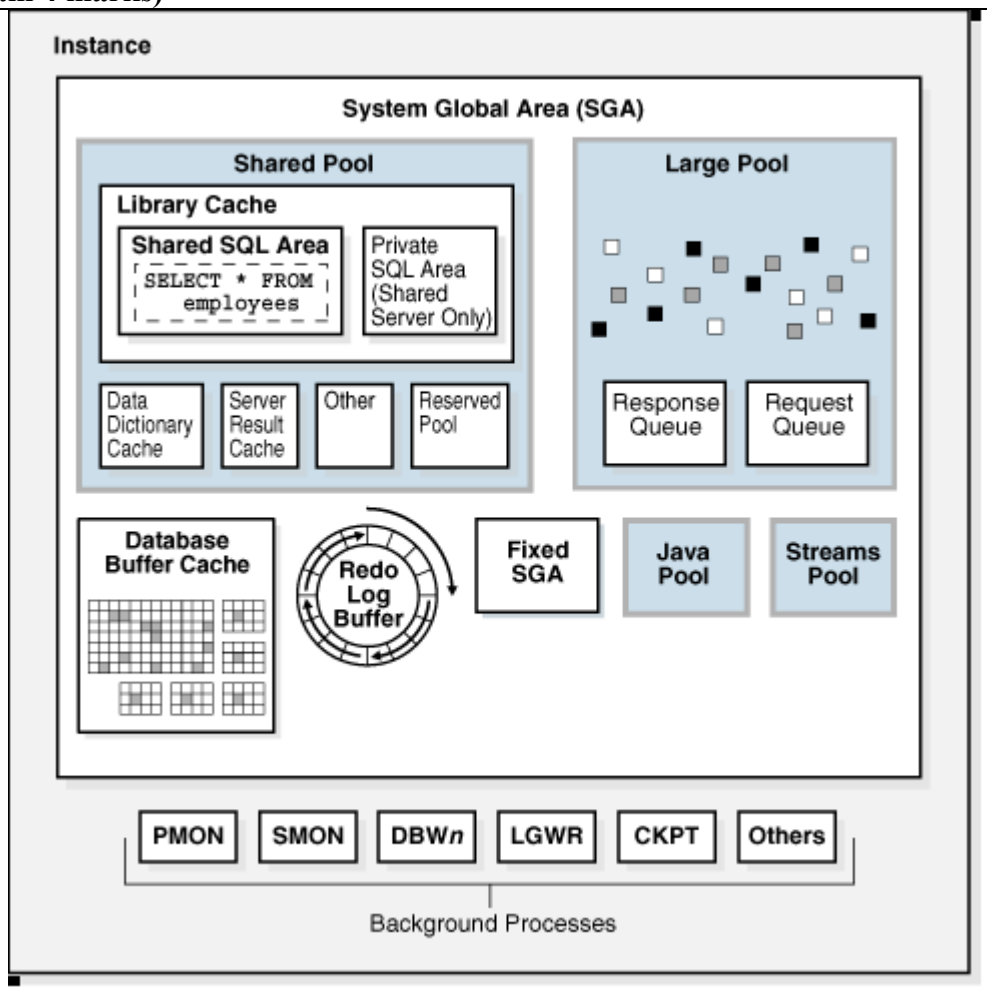
Privilege	Object type	Description
DELETE	Table	Enables a user to delete from a table.
EXECUTE	PL/SQL package, procedure or function	Enables a user to execute a PL/SQL package, procedure or function directly.
FLUSH	Cache group	Enables a user to flush a cache group.
INDEX	Table or materialized view	Enables a user to create an index on a table or materialized view.
INSERT	Table or synonym	Enables a user to insert into a table or into the table through a synonym.
UNLOAD	Cache group	Enables a user to unload a cache group.
UPDATE	Table	Enables a user to update a table.

	(b)	**Explain advantages and limitations of web database.** **(2 advantages-2 marks, 2 limitations-2marks)**	**2+2=4**
	Ans:	**Advantages:** **1. Save Money** One of the advantages of online database software is that it can save your business money. When you don't need to buy a software program for your business, this could result in a major savings overall. In most cases, businesses pay for a software program and then pay for a licensing fee for each computer that uses it. Using an online database may prove cheaper, depending on the number of computers you use. You also don't need to invest in servers to store the data at your business. **2. Flexible Use** Another benefit of using an online database program is that it allows your business to be flexible. You only pay for the amount of storage that you use. You need not worry about purchasing servers as you go or eliminating them when they are no longer needed. If your business grows or shrinks, you do not need to be concerned about the costs of database	

		management software or servers. **1. Technical Support** Another advantage of using a Web-based database program is that you can shift the technical support burden to someone else. Paying a company for access to an online database includes technical support. If the database has problems, you simply contact the company and the staff handles it. You don't need to pay for an information technology professional for this purpose. If you already have an IT department, your employees can focus on other things. **Limitations:** **1. Reliability** The internet is currently an unreliable and slow communication medium-when a request is carried across the internet, there is no real guarantee of delivery. **2. Security** Security is of great concern for an organization that makes its database accessible on the web. User authentication and secure data transmission are critical because of the large number of potentially anonymous users. **3. Cost** Cost of maintenance is very expensive.	
	(c)	**State use of transaction logs.** **(Each point 1 mark)**	**1*4=4**
	Ans:	• **Individual transaction recovery** If an application issues a ROLLBACK statement, or if the Database Engine detects an error such as the loss of communication with a client, the log records are used to roll back the modifications made by an incomplete transaction. • **Recovery of all incomplete transactions when Server is started** If a server fails, the databases may be left in a state where some modifications were never written from the buffer cache to the data files, and there may be some modifications from incomplete transactions in the data files. When an instance of SQL Server is started, it runs a recovery of each database. Every modification recorded in the log which may not have been written to the data files is rolled forward. Every incomplete transaction found in the transaction log is then rolled back to make sure the integrity of the database is preserved. • **Rolling a restored database, file, filegroup, or page forward to the point of failure**	

		After a hardware loss or disk failure affecting the database files, you can restore the database to the point of failure. You first restore the last full database backup and the last differential database backup, and then restore the subsequent sequence of the transaction log backups to the point of failure.+ As you restore each log backup, the Database Engine reapplies all the modifications recorded in the log to roll forward all the transactions. When the last log backup is restored, the Database Engine then uses the log information to roll back all transactions that were not complete at that point. • **Supporting transactional replication** The Log Reader Agent monitors the transaction log of each database configured for transactional replication and copies the transactions marked for replication from the transaction log into the distribution database	
	(d)	**List four potential benefits of digital library.** **(Each point 1 mark)**	**1*4=4**
	Ans:	**No physical boundary**. The user of a digital library need not to go to the library physically; people from all over the world can gain access to the same information, as long as an Internet connection is available. **Round the clock availability** A major advantage of digital libraries is that people can gain access 24/7 to the information. **Multiple access**. The same resources can be used simultaneously by a number of institutions and patrons. This may not be the case for copyrighted material: a library may have a license for "lending out" only one copy at a time; this is achieved with a system of digital rights management where a resource can become inaccessible after expiration of the lending period or after the lender chooses to make it inaccessible (equivalent to returning the resource). **Information retrieval**. The user is able to use any search term (word, phrase, title, name, subject) to search the entire collection. Digital libraries can provide very user-friendly interfaces, giving click able access to its resources. **Preservation and conservation**. Digitization is not a long-term preservation solution for physical collections, but does succeed in providing access copies for materials that would otherwise fall to degradation from repeated use. Digitized collections and born-digital objects pose many preservation and conservation concerns that analog materials do not. Please see the following "Problems" section of this page for examples. **Space**. Whereas traditional libraries are limited by storage space, digital libraries have the potential to store much more information, simply because digital information requires very little physical space to contain them and media storage technologies are more affordable than ever before.	

	(e)	**Explain how to create a DBA user.** **(Explanation 3 marks, Example 1 mark)**	**3+1=4**
	Ans:	**Creating a User** Once connected as SYSTEM, simply issue the CREATE USER command to generate a new account. CREATE USER books_admin IDENTIFIED BY MyPassword; Here we're simply creating a books_admin account that is IDENTIFIED or authenticated by the specified password. **The Grant Statement** With our new books_admin account created, we can now begin adding privileges to the account using the GRANT statement. GRANT is a very powerful statement with many possible options, but the core functionality is to manage the privileges of both users and roles throughout the database. Grant all on "book_admin" with grant option; **Example** create user bob; identified by password; grant connect to bob; grant all privileges to bob;	

	(f)	**Draw Oracle instance architecture.** **(Diagram 4 marks)**	**4**
	Ans:		
4.		**Attempt any FOUR :**	**4*4=16**
	(a)	**Explain types of oracle database structure.** **(Each point 2 marks)**	**2+2=4**
	Ans:	The database has **logical structures** and **physical structures**. Because the physical and logical structures are separate, the physical storage of data can be managed without affecting the access to logical storage structures. **1. Physical Database Structures** physical database structures of an Oracle database, including datafiles, redo log files, and control files. **Datafiles** Every Oracle database has one or more physical **datafiles**. The datafiles contain all the database data. The data of logical database structures, such as tables and indexes, is	

physically stored in the datafiles allocated for a database.

Control Files

Every Oracle database has a **control file**. A control file contains entries that specify the physical structure of the database. For example, it contains the following information:

Database name
Names and locations of datafiles and redo log files
Time stamp of database creation

Redo Log Files

Every Oracle database has a set of two or more **redo log files**. The set of redo log files is collectively known as the redo log for the database. A redo log is made up of redo entries (also called **redo records**).

The primary function of the redo log is to record all changes made to data. If a failure prevents modified data from being permanently written to the datafiles, then the changes can be obtained from the redo log, so work is never lost.

Archive Log Files

You can enable automatic archiving of the redo log. Oracle automatically archives log files when the database is in ARCHIVELOG mode.

Parameter Files

Parameter files contain a list of configuration parameters for that instance and database.

Backup Files

To restore a file is to replace it with a backup file. Typically, you restore a file when a media failure or user error has damaged or deleted the original file.

2. Logical Database Structures

The logical storage structures, including data blocks, extents, and segments, enable Oracle to have fine-grained control of disk space use.

Tablespaces

A database is divided into logical storage units called **tablespaces**, which group related logical structures together. For example, tablespaces commonly group together all application

		objects to simplify some administrative operations. Online and Offline Tablespaces A tablespace can be **online** (accessible) or **offline** (not accessible). A tablespace is generally online, so that users can access the information in the tablespace. However, sometimes a tablespace is taken offline to make a portion of the database unavailable while allowing normal access to the remainder of the database. This makes many administrative tasks easier to perform. **Oracle Data Blocks** At the finest level of granularity, Oracle database data is stored in **data blocks**. One data block corresponds to a specific number of bytes of physical database space on disk. The standard block size is specified by the DB_BLOCK_SIZE initialization parameter. In addition, you can specify up to five other block sizes. A database uses and allocates free database space in Oracle data blocks. **Extents** The next level of logical database space is an **extent**. An extent is a specific number of contiguous data blocks, obtained in a single allocation, used to store a specific type of information. **Segments** Above extents, the level of logical database storage is a **segment**. A segment is a set of extents allocated for a certain logical structure.	
	(b)	**List characteristics of mobile computing.** **(Each point ½ mark)**	**½ * 8=4**
	Ans:	1. Portability 2. Social Interactivity 3. Connectivity 4. Context Sensitivity 5. Individual 6. Small Size	

		7. Wireless Communication 8. Speed	
	(c)	**Explain how to Revoke the assigned privilege.** **(Revoke Explanation 2 mark, syntax 1 mark, example 1 mark)**	**2+1+1= 4**
	Ans:	Use the REVOKE statement to: • Revoke system privileges from users and roles • Revoke roles from users and roles • Revoke object privileges for a particular object from users and roles To revoke a **system privilege,** you must have been granted the privilege with the ADMIN OPTION. To revoke a **role**, you must have been granted the role with the ADMIN OPTION. You can revoke any role if you have the GRANT ANY ROLE system privilege. **Syntax** REVOKE privilege_name ON object_name FROM {user_name |PUBLIC |role_name} **For Example:** REVOKE SELECT ON employee FROM user1; This command will REVOKE a SELECT privilege on employee table from user1.When you REVOKE SELECT privilege on a table from a user, the user will not be able to SELECT data from that table anymore.	
	(d)	**Explain catastrophic failures.** **(Definition 1 mark, Explanation 2 marks, example 1 mark)**	**1+2+1= 4**
	Ans:	**Catastrophic Failure** **Definition -** "A **catastrophic failure** is a sudden and total failure from which recovery is impossible. Catastrophic failures often lead to cascading systems failure." Catastrophic failure is a complete, sudden, often unexpected breakdown in a machine,	

		electronic system, computer or network. Such a breakdown may occur as a result of a hardware event such as a disk drive crash, memory chip failure or surge on the power line. Catastrophic failure can also be caused by software conflicts or malware. Sometimes a single component in a critical location fails, resulting in downtime for the entire system **For example,** catastrophic failure can be observed in steam turbine rotor failure, which can occur due to peak stress on the rotor; stress concentration increases up to a point at which it is excessive, leading ultimately to the failure of the disc. In firearms, catastrophic failure usually refers to a rupture or disintegration of the barrel or receiver of the gun when firing it.	
	(e)	**State the need of backup plan.** **(Each point 1 mark)**	**1*4=4**
	Ans:	**1. Preparing for Problems** Numerous daily hazards can cause system failures. As you plan your database backup and recovery strategy, be sure to consider all of these various threats to database integrity and availability. **2. Image Copy Backups** A fundamental component of a database backup and recovery plan is creating backup copies of data. When an error occurs that damages the integrity of the database, a backup copy of the data can be used as the basis to recover or restore the database. However, the full story on backing up a database is not quite that simple. **3. Document Your Backup Strategy** Once your backup strategy has been established and implemented, the backup system can run for a long time without any DBA intervention required. Such automation is a mixed blessing, though. Over time, things can be forgotten and the DBA staff can change, both of which can cause confusion during a hectic database recovery. For this reason it is imperative that the backup and recovery strategy, implementation, and procedures be thoroughly tested and documented by the DBA. **4. Database Recovery** When problems impact the database, the DBA can use the image copy backups and the database log to recover the database. Whatever the cause of the problem, the DBA must be able to recover data quickly so that the business can continue to operate. When data is unavailable, your company may be losing thousands or even millions of dollars. Recognizing the need for a database recovery is quite different from actually performing a recovery in a speedy and proper fashion. **Database recovery can be a very complex task** that is prone to errors and difficult to manage.	

	(f)	**Describe Log Writer (LGWR).** **(Explanation 4 marks)**	**4**

	Ans:	**Log Writer Process (LGWR)**
		
		The **log writer process (LGWR)** is responsible for redo log buffer management—writing the redo log buffer to a redo log file on disk. LGWR writes all redo entries that have been copied into the buffer since the last time it wrote.
		
		The redo log buffer is a circular buffer. When LGWR writes redo entries from the redo log buffer to a redo log file, server processes can then copy new entries over the entries in the redo log buffer that have been written to disk. LGWR normally writes fast enough to ensure that space is always available in the buffer for new entries, even when access to the redo log is heavy.
		
		LGWR writes one contiguous portion of the buffer to disk. LGWR writes:
		
		• A commit record when a user process commits a transaction
		• Redo log buffers
		  o Every three seconds
		  o When the redo log buffer is one-third full
		  o When a DBW*n* process writes modified buffers to disk, if necessary
		
		GWR writes synchronously to the active mirrored group of redo log files. If one of the files in the group is damaged or unavailable, LGWR continues writing to other files in the group and logs an error in the LGWR trace file and in the system alert log. If all files in a group are damaged, or the group is unavailable because it has not been archived, LGWR cannot continue to function.
		
		When a user issues a COMMIT statement, LGWR puts a commit record in the redo log buffer and writes it to disk immediately, along with the transaction's redo entries. The corresponding changes to data blocks are deferred until it is more efficient to write them. This is called a **fast commit** mechanism. The atomic write of the redo entry containing the transaction's commit record is the single event that determines the transaction has committed. Oracle returns a success code to the committing transaction, although the data buffers have not yet been written to disk.
		
		When a user commits a transaction, the transaction is assigned a **system change number (SCN)**, which Oracle records along with the transaction's redo entries in the redo log. SCNs are recorded in the redo log so that recovery operations can be synchronized in Real Application Clusters and distributed databases.	

5.		**Attempt any TWO :**	**2*8=16**
	(a)	**List types of threat issues for database security & explain one in detail.** **(List 2 marks, explanation 4 marks, example 2 marks)**	**2+4+2= 8**
	Ans:	**List** 1. Excessive Privilege Abuse: 2. Legitimate Privilege Abuse 3. Privilege Elevation 4. Platform Vulnerabilities 5. SQL Injection 6. Weak Audit Trails 7. Denial of Service 8. Database Communication Protocol Vulnerabilities 9. Backup Data Exposure **SQL Injection** SQL injection is a code injection technique that might destroy your database. SQL injection is one of the most common web hacking techniques. SQL injection is the placement of malicious code in SQL statements, via web page input. **SQL in Web Pages** SQL injection usually occurs when you ask a user for input, like their username/userid, and instead of a name/id, the user give you an SQL statement that you will unknowingly run on your database. Look at the following example which creates a SELECT statement by adding a variable (txtUserId) to a select string. The variable is fetched from user input (getRequestString): **Example** txtUserId = getRequestString("UserId");	

txtSQL = "SELECT * FROM Users WHERE UserId = " + txtUserId;

SQL Injection Based on 1=1 is Always True

Look at the example above again. The original purpose of the code was to create an SQL statement to select a user, with a given user id.

If there is nothing to prevent a user from entering "wrong" input, the user can enter some "smart" input like this:

UserId:

105 OR 1=1

Then, the SQL statement will look like this:

SELECT * FROM Users WHERE UserId = 105 OR 1=1;

The SQL above is valid and will return ALL rows from the "Users" table, since OR 1=1 is always TRUE.

Does the example above look dangerous? What if the "Users" table contains names and passwords?

The SQL statement above is much the same as this:

SELECT UserId, Name, Password FROM Users WHERE UserId = 105 or 1=1;

A hacker might get access to all the user names and passwords in a database, by simply inserting 105 OR 1=1 into the input field.

SQL Injection Based on ""="" is Always True

Here is an example of a user login on a web site:

Username:

John Doe

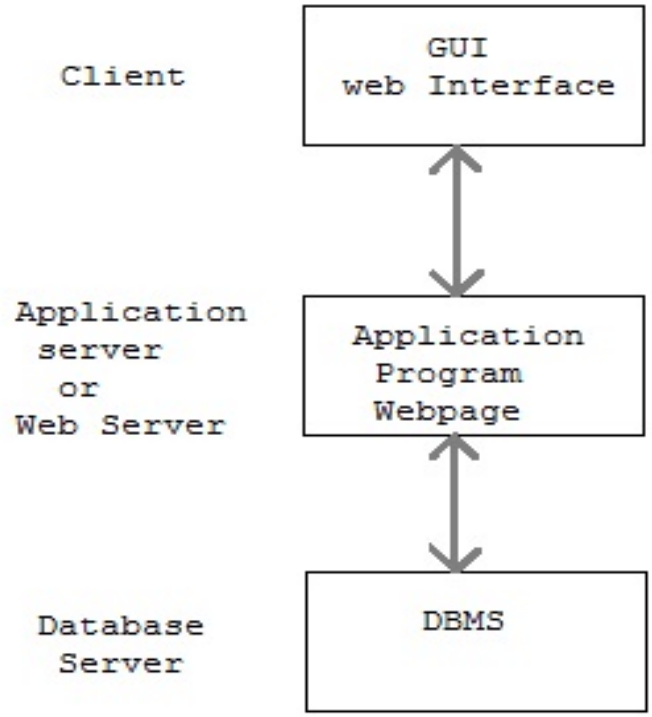
Password:

myPass

Example

uName = getRequestString("username");

		uPass = getRequestString("userpassword"); sql = 'SELECT * FROM Users WHERE Name ="' + uName + '" AND Pass ="' + uPass + '"' **Result** SELECT * FROM Users WHERE Name ="John Doe" AND Pass ="myPass" A hacker might get access to user names and passwords in a database by simply inserting " OR ""=" into the user name or password text box: User Name: " or ""=" Password: " or ""=" The code at the server will create a valid SQL statement like this: **Result** SELECT * FROM Users WHERE Name ="" or ""="" AND Pass ="" or ""="" The SQL above is valid and will return all rows from the "Users" table, since OR ""="" is always TRUE.	
	(b)	**Explain partitioning in client-server architecture.** **(Diagram 2 mark, Explanation 6 mark)**	**2+6=8**
	Ans:	A **partition** is a division of a logical database or its constituent elements into distinct independent parts. Database partitioning is normally done for manageability, performance or availability reasons, or for load balancing	



Client Server Architecture

Three-tier Client / Server database architecture is commonly used architecture for web applications. Intermediate layer called Application server or Web Server stores the web connectivity software and the business logic(constraints) part of application used to access the right amount of data from the database server. This layer acts like medium for sending partially processed data between the database server and the client

Three tier with an application server This architecture allows the main body of an application to run on a shared host rather than in the user system interface client environment. The application server shares business logic, computations and a data retrieval engine. In this architecture applications are more scalable and installation costs are less on a single server than maintaining each on a desktop client.

3-tier architecture provides

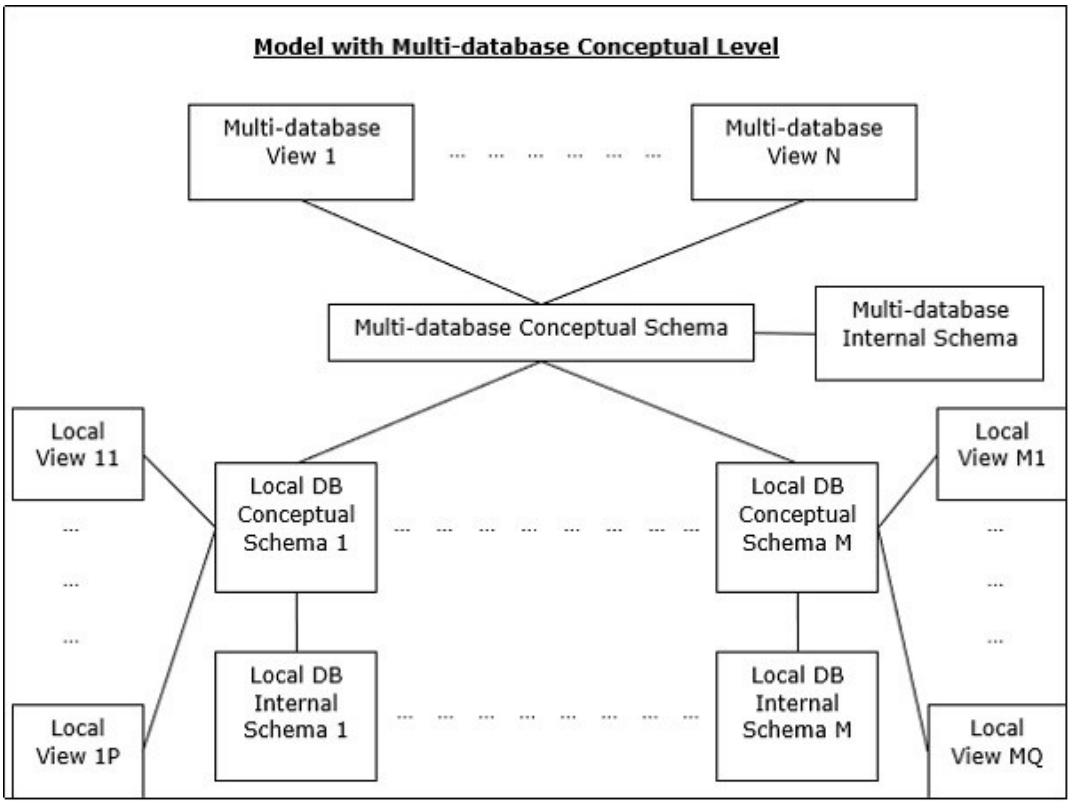
- A greater degree of flexibility
- Increased security, as security can be defined for each service, and at each level
- Increased performance, as tasks are shared between servers

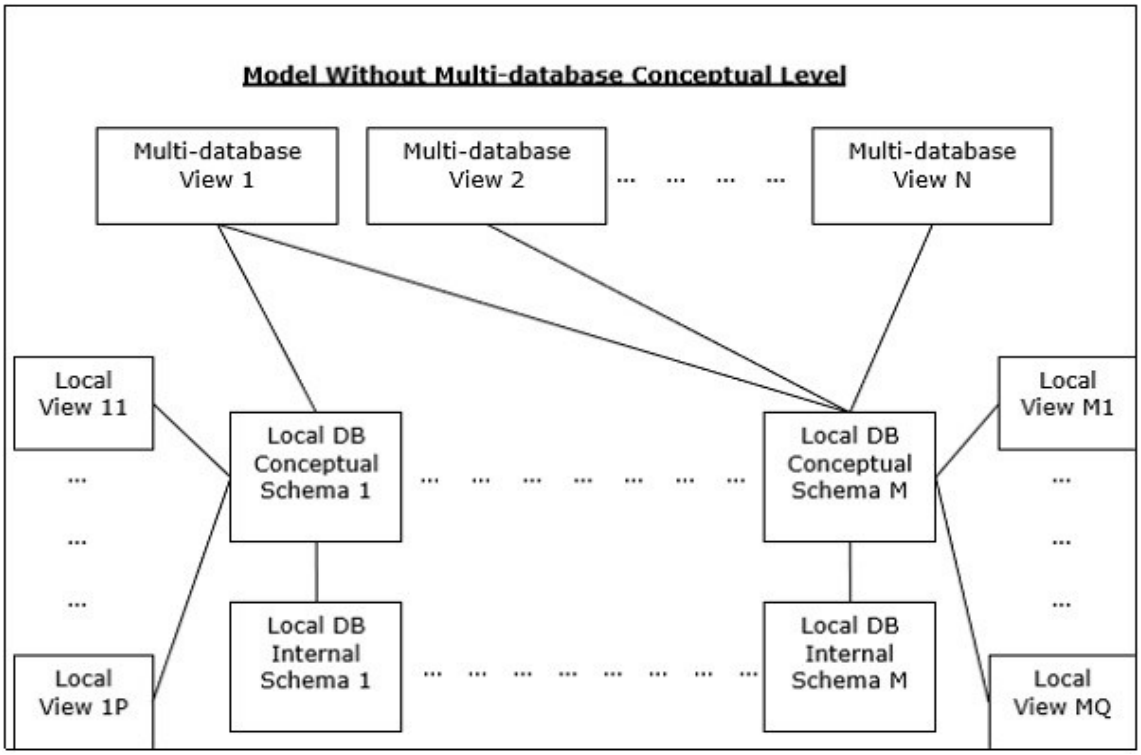
Benefits of the Oracle client/server architecture in a distributed processing environment

include the following:

- Client applications are not responsible for performing any data processing. Client applications can concentrate on requesting input from users, requesting desired data from the server, and then analyzing and presenting this data using the display capabilities of the client workstation or the terminal (for example, using graphics or spreadsheets).
- Client applications can be designed with no dependence on the physical location of the data. If the data is moved or distributed to other database servers, the application continues to function with little or no modification.
- Oracle exploits the multitasking and shared-memory facilities of its underlying operating system. As a result, it delivers the highest possible degree of concurrency, data integrity, and performance to its client applications.
- Client workstations or terminals can be optimized for the presentation of data (for example, by providing graphics and mouse support) and the server can be optimized for the processing and storage of data (for example, by having large amounts of memory and disk space).
- If necessary, Oracle can be *scaled*. As your system grows, you can add multiple servers to distribute the database processing load throughout the network (*horizontally scaled*). Alternatively, you can replace Oracle on a less powerful computer, such as a microcomputer, with Oracle running on a minicomputer or mainframe, to take advantage of a larger system's performance (*vertically scaled*). In either case, all data and applications are maintained with little or no modification, since Oracle is portable between systems.
- In networked environments, shared data is stored on the servers, rather than on all computers in the system. This makes it easier and more efficient to manage concurrent access.
- In networked environments, inexpensive, low-end client workstations can be used to access the remote data of the server effectively.
- In networked environments, client applications submit database requests to the server using SQL statements. Once received, the SQL statement is processed by the server, and the results are returned to the client application. Network traffic is kept to a minimum because only the requests and the results are shipped over the network.

	(c)	**Explain Multi-Database system. Explain recovery procedure in it.** **(Multi database 4 marks, Recovery procedure 4 marks)**	**4+4=8**
	Ans:	**1. Multi - DBMS Architectures** This is an integrated database system formed by a collection of two or more autonomous database systems. Multi-DBMS can be expressed through six levels of schemas − • **Multi-database View Level** − Depicts multiple user views comprising of subsets of the integrated distributed database. • **Multi-database Conceptual Level** − Depicts integrated multi-database that comprises of global logical multi-database structure definitions. • **Multi-database Internal Level** − Depicts the data distribution across different sites and multi-database to local data mapping. • **Local database View Level** − Depicts public view of local data. • **Local database Conceptual Level** − Depicts local data organization at each site. • **Local database Internal Level** − Depicts physical data organization at each site. There are two design alternatives for multi-DBMS − • Model with multi-database conceptual level. • Model without multi-database conceptual level.	





1. Recovery Procedure

After any application or system failure, there are two possible approaches to database recovery:

1. There is no need for recoverability, and all databases can be re-created from scratch. Although these applications may still need transaction protection for other reasons, recovery usually consists of removing the Berkeley DB environment home directory and all files it contains, and then restarting the application. Such an application may use the DB_TXN_NOT_DURABLE flag to avoid writing log records.

2. It is necessary to recover information after system or application failure. In this case, recovery processing must be performed on any database environments that were active at the time of the failure. Recovery processing involves running the db_recover utility or calling the DB_ENV->open() method with the DB_RECOVER or DB_RECOVER_FATAL flags.

During recovery processing, all database changes made by aborted or unfinished transactions are undone, and all database changes made by committed transactions are redone, as necessary. Database applications must not be restarted until recovery completes. After recovery finishes, the environment is properly initialized so that applications may be restarted.

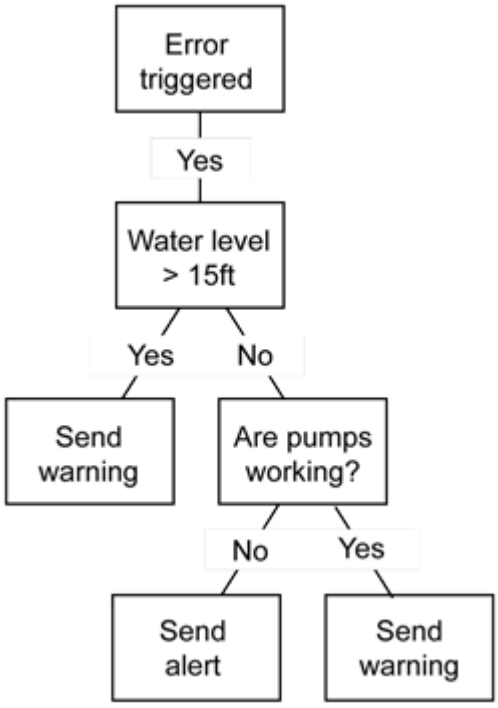
If performing recovery, there are two types of recovery

processing: *normal* and *catastrophic*. Which you choose depends on the source for the database and log files you are using to recover.

If up-to-the-minute database and log files are accessible on a stable filesystem, normal recovery is sufficient. Run the db recover utility or call the DB_ENV->open () method specifying the DB_RECOVER flag. However, the normal recovery case **never** includes recovery using hot backups of the database environment. For example, you cannot perform a hot backup of databases and log files, restore the backup and then run normal recovery — you must always run catastrophic recovery when using hot backups.
If the database or log files have been destroyed or corrupted, or normal recovery fails, catastrophic recovery is required.
For example, catastrophic failure includes the case where the disk drive on which the database or log files are stored has been physically destroyed, or when the underlying filesystem is corrupted and the operating system's normal filesystem checking procedures cannot bring that filesystem to a consistent state.

This is often difficult to detect, and a common sign of the need for catastrophic recovery is when normal Berkeley DB recovery procedures fail, or when checksum errors are displayed during normal database procedures.

6.		Attempt any TWO :	2*8=16
	(a)	**Explain any two data mining techniques.** **(Each technique 4 marks)**	4+4=8
	Ans:	**1. Prediction** Prediction is a wide topic and runs from predicting the failure of components or machinery, to identifying fraud and even the prediction of company profits. Used in combination with the other data mining techniques, prediction involves analyzing trends, classification, pattern matching, and relation. By analyzing past events or instances, you can make a prediction about an event. Using the credit card authorization, for example, you might combine decision tree analysis of individual past transactions with classification and historical pattern matches to identify whether a transaction is fraudulent. Making a match between the purchase of flights to the US and transactions in the US, it is likely that the transaction is valid. **2. Decision trees** Related to most of the other techniques (primarily classification and prediction), the decision tree can be used either as a part of the selection criteria, or to support the use and selection of specific data within the overall structure. Within the decision tree, you start with a simple question that has two (or sometimes more) answers. Each answer leads to a further question to help classify or identify the data so that it can be categorized, or so that a prediction can be made based on each answer.	



Decision trees are often used with classification systems to attribute type information, and with predictive systems, where different predictions might be based on past historical experience that helps drive the structure of the decision tree and the output.

(b) **Explain how to access data from java program.**
(explanation 4 marks, program 4 marks) **4+4=8**

Ans:

Creating JDBC Application

There are following six steps involved in building a JDBC application −

- **Import the packages:** Requires that you include the packages containing the JDBC classes needed for database programming. Most often, using *import java.sql.** will suffice.
- **Register the JDBC driver:** Requires that you initialize a driver so you can open a communication channel with the database.
- **Open a connection:** Requires using the *DriverManager.getConnection()* method to create a Connection object, which represents a physical connection with the database.
- **Execute a query:** Requires using an object of type Statement for building and

submitting an SQL statement to the database.

- **Extract data from result set:** Requires that you use the appropriate *ResultSet.getXXX()* method to retrieve the data from the result set.
- **Clean up the environment:** Requires explicitly closing all database resources versus relying on the JVM's garbage collection.

Sample Code

```
//STEP 1. Import required packages
import java.sql.*;
public class FirstExample {
  // JDBC driver name and database URL
  static final String JDBC_DRIVER = "com.mysql.jdbc.Driver";
  static final String DB_URL = "jdbc:mysql://localhost/EMP";
  //  Database credentials
  static final String USER = "username";
  static final String PASS = "password";
    public static void main(String[] args) {
  Connection conn = null;
  Statement stmt = null;
  try{
    //STEP 2: Register JDBC driver
    Class.forName("com.mysql.jdbc.Driver");
    //STEP 3: Open a connection
    System.out.println("Connecting to database...");
    conn = DriverManager.getConnection(DB_URL,USER,PASS);
```

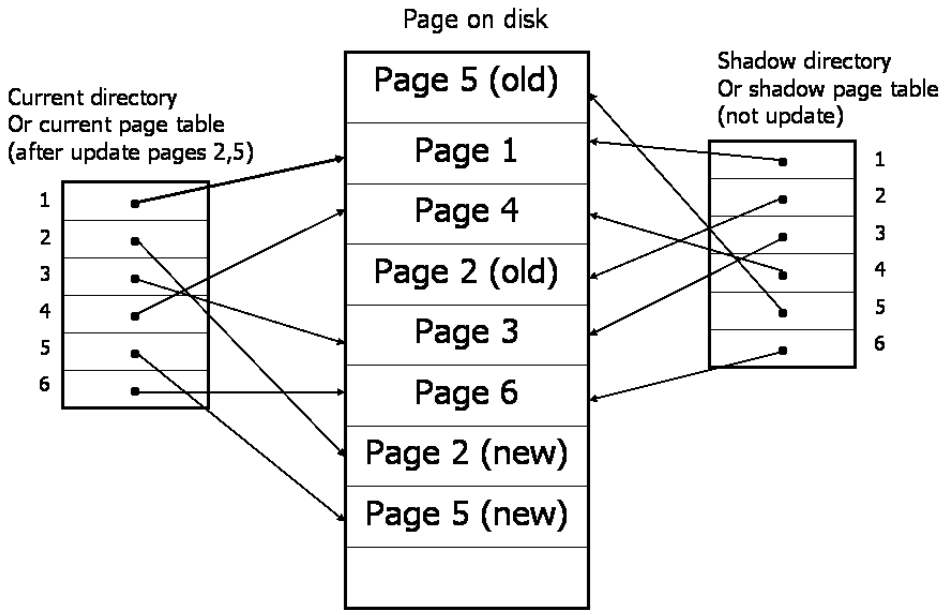
```
//STEP 4: Execute a query
System.out.println("Creating statement...");
stmt = conn.createStatement();
String sql;
sql = "SELECT id, first, last, age FROM Employees";
ResultSet rs = stmt.executeQuery(sql);
//STEP 5: Extract data from result set
while(rs.next()){
  //Retrieve by column name
  int id  = rs.getInt("id");
  int age = rs.getInt("age");
  String first = rs.getString("first");
  String last = rs.getString("last");
  //Display values
  System.out.print("ID: " + id);
  System.out.print(", Age: " + age);
  System.out.print(", First: " + first);
  System.out.println(", Last: " + last);
}
//STEP 6: Clean-up environment
rs.close();
stmt.close();
conn.close();
}catch(SQLException se){
```

```
    //Handle errors for JDBC
    se.printStackTrace();
  }catch(Exception e){
    //Handle errors for Class.forName
    e.printStackTrace();
  }finally{
    //finally block used to close resources
    try{
      if(stmt!=null)
        stmt.close();
    }catch(SQLException se2){
    }// nothing we can do
    try{
      if(conn!=null)
        conn.close();
    }catch(SQLException se){
      se.printStackTrace();
    }//end finally try
  }//end try
  System.out.println("Goodbye!");
}//end main
}//end FirstExample
```

Now let us compile the above example as follows −

```
C:\>javac FirstExample.java
```

		C:\> When you run FirstExample, it produces the following result − C:\>java FirstExample Connecting to database... Creating statement... ID: 100, Age: 18, First: Zara, Last: Ali ID: 101, Age: 25, First: Mahnaz, Last: Fatma ID: 102, Age: 30, First: Zaid, Last: Khan ID: 103, Age: 28, First: Sumit, Last: Mittal C:\>	
	(c)	**State any four importance of database backup. Explain shadow paging.** **(Importance 4 marks, shadow paging 4 marks)**	**4+4=8**
	Ans:	1. **Computer crashes** – always happen when you least need it, and can lead to data loss. 2. **Virus Infection** – aggressive malicious viruses can corrupt files and disable computers. 3. **Hard drive failure** - hard drives have a finite lifetime and can fail suddenly and without warning. The sudden death of a hard drive can cause the painful loss of months or years of irreplaceable files and the timing can be catastrophic – if this happens close to a work or college deadline it could be a nightmare scenario. 4. **Physical computer damage** – your files are more at risk if you use a laptop. Light and portable comes at the price of reduced durability. Laptops are sensitive and are easily damaged if dropped or have drinks spilled over them. 5. **Theft** – computers are sought after by thieves and cannot always be kept secure whilst travelling. **Shadow Paging** In computer science, **shadow paging** is a technique for providing atomicity and durability (two of the ACID properties) in database systems. An alternative to log-based crash-recovery techniques is **shadow paging**. Shadow paging is a copy-on-write technique for avoiding in-place updates of pages. Instead, when a page is to be modified, a **shadow page** is allocated. Since the shadow page has no references (from other pages on disk), it can be modified liberally, without concern for consistency constraints, etc. When the page is ready to become durable, all pages that referred to the original are updated to refer to the new replacement page instead. Because the page is "activated" only when it is ready, it is atomic.	



Shadow paging considers

- The database is partitioned into fixed-length blocks referred to as PAGES.
- Page table has n entries – one for each database page.
- Each contain pointer to a page on disk (1 to 1st page on database and so on…).

The idea is to maintain 2 pages tables during the life of transaction.

- The current page table
- The shadow page table

When transaction starts, both page tables are identical

- The shadow page table is never changed over the duration of the transaction.
- The current page table may be changed when a transaction performs a write operation.
- All input and output operations use the current page table to locate database pages on disk.

Advantages

- No Overhead for writing log records.

- No Undo / No Redo algorithm.
- Recovery is faster.

Disadvantages

- Data gets fragmented or scattered.
- After every transaction completion database pages containing old version of modified data need to be garbage collected.

Hard to extend algorithm to allow transaction to run concurrently.