

Technical Papers Database Questions: Database Part II

Examrace Placement Series prepares you for the toughest placement exams to top companies.

What is 5NF?

A Relation schema R is said to be 5NF if for every join dependency $\{ R_1, R_2, \dots R_n \}$ that holds R, one the following is true

$R_i = R$ for some i.

The join dependency is implied by the set of FD, over R in which the left side is key of R.

What is Domain-Key Normal Form?

A relation is said to be in DKNF if all constraints and dependencies that should hold on the the constraint can be enforced by simply enforcing the domain constraint and key constraint on the relation.

What are partial, alternate, artificial, compound and natural key?

Partial Key: It is a set of attributes that can uniquely identify weak entities and that are related to same owner entity. It is sometime called as Discriminator.

Alternate Key: All Candidate Keys excluding the Primary Key are known as Alternate Keys.

Artificial Key: If no obvious key, either stand alone or compound is available, then the last resort is to simply create a key, by assigning a unique number to each record or occurrence. Then this is known as developing an artificial key.

Compound Key: If no single data element uniquely identifies occurrences within a construct, then combining multiple elements to create a unique identifier for the construct is known as creating a compound key.

Natural Key: When one of the data elements stored within a construct is utilized as the primary key, then it is called the natural key.

What is indexing and what are the different kinds of indexing?

Indexing is a technique for determining how quickly specific data can be found.

Types:

- Binary search style indexing
- B-Tree indexing

- Inverted list indexing
- Memory resident table
- Table indexing

What is system catalog or catalog relation? How is better known as?

A RDBMS maintains a description of all the data that it contains, information about every relation and index that it contains. This information is stored in a collection of relations maintained by the system called metadata. It is also called data dictionary.

What is meant by query optimization?

The phase that identifies an efficient execution plan for evaluating a query that has the least estimated cost is referred to as query optimization.

What is join dependency and inclusion dependency?

Join Dependency: A Join dependency is generalization of Multivalued dependency. A JD $\{ R_1, R_2, \dots R_n \}$ is said to hold over a relation R if $R_1, R_2, R_3, \dots R_n$ is a lossless-join decomposition of R. There is no set of sound and complete inference rules for JD.

Inclusion Dependency: An Inclusion Dependency is a statement of the form that some columns of a relation are contained in other columns. A foreign key constraint is an example of inclusion dependency.

What is durability in DBMS?

Once the DBMS informs the user that a transaction has successfully completed, its effects should persist even if the system crashes before all its changes are reflected on disk. This property is called durability.

What do you mean by atomicity and aggregation?

Atomicity: Either all actions are carried out or none are. Users should not have to worry about the effect of incomplete transactions. DBMS ensures this by undoing the actions of incomplete transactions.

Aggregation: A concept which is used to model a relationship between a collection of entities and relationships. It is used when we need to express a relationship among relationships.

What is a Phantom Deadlock?

In distributed deadlock detection, the delay in propagating local information might cause the deadlock detection algorithms to identify deadlocks that do not really exist. Such situations are called phantom deadlocks and they lead to unnecessary aborts.

What is a checkpoint and When does it occur?

Visit examrace.com for free study material, doorseptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

A Checkpoint is like a snapshot of the DBMS state. By taking checkpoints, the DBMS can reduce the amount of work to be done during restart in the event of subsequent crashes.

What are the different phases of transaction?

Different phases are

Analysis phase

Redo Phase

Undo phase

What do you mean by flat file database?

It is a database in which there are no programs or user access languages. It has no cross-file capabilities but is user-friendly and provides user-interface management.

What is "transparent DBMS"

It is one, which keeps its Physical Structure hidden from user.

Brief theory of Network, Hierarchical schemas and their properties

Network schema uses a graph data structure to organize records example for such a database management system is CTCG while a hierarchical schema uses a tree data structure example for such a system is IMS.

What is a query?

A query with respect to DBMS relates to user commands that are used to interact with a data base. The query language can be classified into data definition language and data manipulation language.

What do you mean by Correlated subquery?

Subqueries, or nested queries, are used to bring back a set of rows to be used by the parent query. Depending on how the subquery is written, it can be executed once for the parent query or it can be executed once for each row returned by the parent query. If the subquery is executed for each row of the parent, this is called a correlated subquery.

A correlated subquery can be easily identified if it contains any references to the parent subquery columns in its WHERE clause. Columns from the subquery cannot be referenced anywhere else in the parent query. The following example demonstrates a non-correlated subquery.

Eg Select * From CUST Where 10/03/1990 IN (Select ODATE From ORDER Where CUST.CNUM = ORDER.CNUM)

What are the primitive operations common to all record management systems?

Addition, deletion and modification.

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

Name the buffer in which all the commands that are typed in are stored 'Edit' Buffer

What are the unary operations in Relational Algebra? PROJECTION and SELECTION.

Are the resulting relations of PRODUCT and JOIN operation the same? No.

PRODUCT: Concatenation of every row in one relation with every row in another.

JOIN: Concatenation of rows from one relation and related rows from another.

What is RDBMS KERNEL?

Two important pieces of RDBMS architecture are the kernel, which is the software, and the data dictionary, which consists of the system-level data structures used by the kernel to manage the database

You might think of an RDBMS as an operating system (or set of subsystems), designed specifically for controlling data access; its primary functions are storing, retrieving, and securing data. An RDBMS maintains its own list of authorized users and their associated privileges; manages memory caches and paging; controls locking for concurrent resource usage; dispatches and schedules user requests; and manages space usage within its table-space structures

Name the sub-systems of a RDBMS: I/O, Security, Language Processing, Process Control, Storage Management, Logging and Recovery, Distribution Control, Transaction Control, Memory Management, Lock Management

Which part of the RDBMS takes care of the data dictionary? How Data dictionary is a set of tables and database objects that is stored in a special area of the database and maintained exclusively by the kernel.

What is the job of the information stored in data-dictionary?

The information in the data dictionary validates the existence of the objects, provides access to them, and maps the actual physical storage location.

Not only RDBMS takes care of locating data it also

determines an optimal access path to store or retrieve the data

How do you communicate with an RDBMS?

You communicate with an RDBMS using Structured Query Language (SQL)

Define SQL and state the differences between SQL and other conventional programming Languages

SQL is a nonprocedural language that is designed specifically for data access operations on normalized relational database structures. The primary difference between SQL and other conventional programming languages is that SQL statements specify what data operations should be performed rather than how to perform them.

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

Name the three major set of files on disk that compose a database in Oracle

There are three major sets of files on disk that compose a database. All the files are binary. These are

- Database files
- Control files
- Redo logs

The most important of these are the database files where the actual data resides. The control files and the redo logs support the functioning of the architecture itself.

All three sets of files must be present, open, and available to Oracle for any data on the database to be useable. Without these files, you cannot access the database, and the database administrator might have to recover some or all of the database using a backup, if there is one.

What is an Oracle Instance?

The Oracle system processes, also known as Oracle background processes, provide functions for the user processes functions that would otherwise be done by the user processes themselves

Oracle database-wide system memory is known as the SGA, the system global area or shared global area. The data and control structures in the SGA are shareable, and all the Oracle background processes and user processes can use them.

The combination of the SGA and the Oracle background processes is known as an Oracle instance

What are the four Oracle system processes that must always be up and running for the database to be useable

The four Oracle system processes that must always be up and running for the database to be useable include DBWR (Database Writer), LGWR (Log Writer), SMON (System Monitor), and PMON (Process Monitor).

What are database files, control files and log files. How many of these files should a database have at least? Why?

Database Files

The database files hold the actual data and are typically the largest in size. Depending on their sizes, the tables (and other objects) for all the user accounts can go in one database file but that's not an ideal situation because it does not make the database structure very flexible for controlling access to storage for different users, putting the database on different disk drives, or backing up and restoring just part of the database.

Visit examrace.com for free study material, doorseptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

You must have at least one database file but usually, more than one files are used. In terms of accessing and using the data in the tables and other objects, the number (or location) of the files is immaterial.

The database files are fixed in size and never grow bigger than the size at which they were created

Control Files

The control files and redo logs support the rest of the architecture. Any database must have at least one control file, although you typically have more than one to guard against loss. The control file records the name of the database, the date and time it was created, the location of the database and redo logs, and the synchronization information to ensure that all three sets of files are always in step. Every time you add a new database or redo log file to the database, the information is recorded in the control files.

Redo Logs

Any database must have at least two redo logs. These are the journals for the database; the redo logs record all changes to the user objects or system objects. If any type of failure occurs, the changes recorded in the redo logs can be used to bring the database to a consistent state without losing any committed transactions. In the case of non-data loss failure, Oracle can apply the information in the redo logs automatically without intervention from the DBA.

The redo log files are fixed in size and never grow dynamically from the size at which they were created.

What is ROWID?

The ROWID is a unique database-wide physical address for every row on every table. Once assigned (when the row is first inserted into the database), it never changes until the row is deleted or the table is dropped.

The ROWID consists of the following three components, the combination of which uniquely identifies the physical storage location of the row.

Oracle database file number, which contains the block with the rows

Oracle block address, which contains the row

The row within the block (because each block can hold many rows)

The ROWID is used internally in indexes as a quick means of retrieving rows with a particular key value. Application developers also use it in SQL statements as a quick way to access a row once they know the ROWID

What is Oracle Block? Can two Oracle Blocks have the same address?

Oracle “formats” the database files into a number of Oracle blocks when they are first created making it easier for the RDBMS software to manage the files and easier to read data into the memory areas.

The block size should be a multiple of the operating system block size. Regardless of the block size, the entire block is not available for holding data; Oracle takes up some space to manage the contents of the block. This block header has a minimum size, but it can grow.

These Oracle blocks are the smallest unit of storage. Increasing the Oracle block size can improve performance, but it should be done only when the database is first created.

Each Oracle block is numbered sequentially for each database file starting at 1. Two blocks can have the same block address if they are in different database files.

What is database Trigger?

A database trigger is a PL/SQL block that can be defined to automatically execute for insert, update, and delete statements against a table. The trigger can be defined to execute once for the entire statement or once for every row that is inserted, updated, or deleted. For any one table, there are twelve events for which you can define database triggers. A database trigger can call database procedures that are also written in PL/SQL.

Name two utilities that Oracle provides, which are used for backup and recovery.

Along with the RDBMS software, Oracle provides two utilities that you can use to back up and restore the database. These utilities are Export and Import.

The Export utility dumps the definitions and data for the specified part of the database to an operating system binary file. The Import utility reads the file produced by an export, recreates the definitions of objects, and inserts the data.

If Export and Import are used as a means of backing up and recovering the database, all the changes made to the database cannot be recovered since the export was performed. The best you can do is recover the database to the time when the export was last performed.

What are stored-procedures? And what are the advantages of using them.

Stored procedures are database objects that perform a user defined operation. A stored procedure can have a set of compound SQL statements. A stored procedure executes the SQL commands and returns the result to the client. Stored procedures are used to reduce network traffic.

How are exceptions handled in PL/SQL? Give some of the internal exceptions' names.

PL/SQL exception handling is a mechanism for dealing with run-time errors encountered during procedure execution. Use of this mechanism enables execution to continue if the error is not severe enough to cause procedure termination.

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

The exception handler must be defined within a subprogram specification. Errors cause the program to raise an exception with a transfer of control to the exception-handler block. After the exception handler executes, control returns to the block in which the handler was defined. If there are no more executable statements in the block, control returns to the caller.

User-Defined Exceptions

PL/SQL enables the user to define exception handlers in the declarations area of subprogram specifications. User accomplishes this by naming an exception as in the following example:

```
ot_failure EXCEPTION.
```

In this case, the exception name is `ot_failure`. Code associated with this handler is written in the `EXCEPTION` specification area as follows:

```
EXCEPTION
```

```
when OT_FAILURE then
```

```
out_status_code: = g_out_status_code.
```

```
out_msg: = g_out_msg.
```

The following is an example of a subprogram exception:

```
EXCEPTION
```

```
when NO_DATA_FOUND then
```

```
g_out_status_code: = 'FAIL'
```

```
RAISE ot_failure.
```

Within this exception is the `RAISE` statement that transfers control back to the `ot_failure` exception handler. This technique of raising the exception is used to invoke all user-defined exceptions.

System-Defined Exceptions

Exceptions internal to PL/SQL are raised automatically upon error. `NO_DATA_FOUND` is a system-defined exception. Table below gives a complete list of internal exceptions.

PL/SQL internal exceptions.

Exception Name Oracle Error

`CURSOR_ALREADY_OPEN` ORA-06511

`DUP_VAL_ON_INDEX` ORA-00001

`INVALID_CURSOR` ORA-01001

`INVALID_NUMBER` ORA-01722

LOGIN_DENIED ORA-01017

NO_DATA_FOUND ORA-01403

NOT_LOGGED_ON ORA-01012

PROGRAM_ERROR ORA-06501

STORAGE_ERROR ORA-06500

TIMEOUT_ON_RESOURCE ORA-00051

TOO_MANY_ROWS ORA-01422

TRANSACTION_BACKED_OUT ORA-00061

VALUE_ERROR ORA-06502

ZERO_DIVIDE ORA-01476

In addition to this list of exceptions, there is a catch-all exception named OTHERS that traps all errors for which specific error handling has not been established.

Does PL/SQL support “overloading” Explain

The concept of overloading in PL/SQL relates to the idea that you can define procedures and functions with the same name. PL/SQL does not look only at the referenced name, however, to resolve a procedure or function call. The count and data types of formal parameters are also considered.

PL/SQL also attempts to resolve any procedure or function calls in locally defined packages before looking at globally defined packages or internal functions. To further ensure calling the proper procedure, you can use the dot notation. Prefacing a procedure or function name with the package name fully qualifies any procedure or function reference.

Tables derived from the ERD

1. Are totally unnormalised
2. Are always in 1NF
3. Can be further denormalised
4. May have multi-valued attributes Answer: (b) Are always in 1NF Spurious tuples may occur due to
 - a. Bad normalization
 - b. Theta joins
 - c. Updating tables from join

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

Which is correct?

- a. i & ii
- b. ii & iii
- c. i & iii
- d. ii & iii

Answer: (a) i & iii because theta joins are joins made on keys that are not primary keys.

A B C is a set of attributes. The functional dependency is as follows

$AB \rightarrow B$

$AC \rightarrow C$

$C \rightarrow B$

- 1. is in 1NF
- 2. is in 2NF
- 3. is in 3NF
- 4. is in BCNF

Answer: Correct answer is A. It is in 1NF since $(AC) + = \{ A, B, C \}$ hence AC is the primary key. Since $C \rightarrow B$ is a FD given, where neither C is a Key nor B is a prime attribute, this it is not in 3NF. Further B is not functionally dependent on key AC thus it is not in 2NF. Thus the given FDs is in 1NF.

In mapping of ERD to DFD

- 1. entities in ERD should correspond to an existing entity/store in DFD
- 2. entity in DFD is converted to attributes of an entity in ERD
- 3. relations in ERD has 1 to 1 correspondence to processes in DFD
- 4. relationships in ERD has 1 to 1 correspondence to flows in DFD

Answer: (a) entities in ERD should correspond to an existing entity/store in DFD

A dominant entity is the entity

- 1. on the N side in a 1: N relationship
- 2. on the 1 side in a 1: N relationship
- 3. on either side in a 1: 1 relationship

4. nothing to do with 1: 1 or 1: N relationship

Answer: (b) on the 1 side in a 1: N relationship

Select 'NORTH' CUSTOMER From CUST_DTLS Where REGION = 'N' Order By

CUSTOMER Union Select 'EAST' CUSTOMER From CUST_DTLS Where REGION = 'E' Order By CUSTOMER

The above is

1. Not an error
2. Error-the string in single quotes 'NORTH' and 'SOUTH'
3. Error-the string should be in double quotes
4. Error-ORDER BY clause

Answer: (d) Error-the ORDER BY clause. Since ORDER BY clause cannot be used in UNIONS

What is Storage Manager?

It is a program module that provides the interface between the low-level data stored in database, application programs and queries submitted to the system.

What is Buffer Manager?

It is a program module, which is responsible for fetching data from disk storage into main memory and deciding what data to be cache in memory.

What is Transaction Manager?

It is a program module, which ensures that database, remains in a consistent state despite system failures and concurrent transaction execution proceeds without conflicting.

What is File Manager?

It is a program module, which manages the allocation of space on disk storage and data structure used to represent information stored on a disk.

What is Authorization and Integrity manager?

It is the program module, which tests for the satisfaction of integrity constraint and checks the authority of user to access data.

What are stand-alone procedures?

Procedures that are not part of a package are known as stand-alone because they independently defined. A good example of a stand-alone procedure is one written in a SQL * Forms application. These types of procedures are not available for reference from other Oracle tools. Another

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

limitation of stand-alone procedures is that they are compiled at run time, which slows execution.

What are cursors give different types of cursors.

PL/SQL uses cursors for all database information accesses statements. The language supports the use two types of cursors

Implicit

Explicit

What is cold backup and hot backup (in case of Oracle)?

Cold Backup:

It is copying the three sets of files (database files, redo logs, and control file) when the instance is shut down. This is a straight file copy, usually from the disk directly to tape. You must shut down the instance to guarantee a consistent copy.

If a cold backup is performed, the only option available in the event of data file loss is restoring all the files from the latest backup. All work performed on the database since the last backup is lost.

Hot Backup:

Some sites (such as worldwide airline reservations systems) cannot shut down the database while making a backup copy of the files. The cold backup is not an available option.

So different means of backing up database must be used the hot backup. Issue a SQL command to indicate to Oracle, on a tablespace-by-tablespace basis, that the files of the tablespace are to be backed up. The users can continue to make full use of the files, including making changes to the data. Once the user has indicated that he/she wants to back up the tablespace files, he/she can use the operating system to copy those files to the desired backup destination.

The database must be running in ARCHIVELOG mode for the hot backup option.

If a data loss failure does occur, the lost database files can be restored using the hot backup and the online and offline redo logs created since the backup was done. The database is restored to the most consistent state without any loss of committed transactions.

What are Armstrong rules? How do we say that they are complete and/or sound

The well-known inference rules for FDs

Reflexive rule:

If Y is subset or equal to X then X Y.

Augmentation rule:

If $X \rightarrow Y$ then $XZ \rightarrow YZ$.

Transitive rule:

If $\{ X \rightarrow Y, Y \rightarrow Z \}$ then $X \rightarrow Z$.

Decomposition rule:

If $X \rightarrow YZ$ then $X \rightarrow Y$.

Union or Additive rule:

If $\{ X \rightarrow Y, X \rightarrow Z \}$ then $X \rightarrow YZ$.

Pseudo Transitive rule:

If $\{ X \rightarrow Y, WY \rightarrow Z \}$ then $WX \rightarrow Z$.

Of these the first three are known as Armstrong Rules. They are sound because it is enough if a set of FDs satisfy these three. They are called complete because using these three rules we can generate the rest all inference rules.

How can you find the minimal key of relational schema?

Minimal key is one which can identify each tuple of the given relation schema uniquely. For finding the minimal key it is required to find the closure that is the set of all attributes that are dependent on any given set of attributes under the given set of functional dependency.

Algo. I Determining X^+ , closure for X, given set of FDs F

Set $X^+ = X$

Set Old $X^+ = X^+$

For each FD $Y \rightarrow Z$ in F and if Y belongs to X^+ then add Z to X^+

Repeat steps 2 and 3 until Old $X^+ = X^+$

Algo. II Determining minimal K for relation schema R, given set of FDs F

Set K to R that is make K a set of all attributes in R

For each attribute A in K

Compute $(KA)^+$ with respect to F

If $(KA)^+ = R$ then set $K = (KA)^+$

What do you understand by dependency preservation?

Given a relation R and a set of FDs F, dependency preservation states that the closure of the union of the projection of F on each decomposed relation R_i is equal to the closure of F. Ie,

$((PR_1(F)) \cup \dots \cup (PR_n(F)))^+ = F^+$

Visit examrace.com for free study material, doorsteptutor.com for questions with detailed explanations, and "Examrace" YouTube channel for free videos lectures

if decomposition is not dependency preserving, then some dependency is lost in the decomposition.

What is meant by Proactive, Retroactive and Simultaneous Update.

Proactive Update: The updates that are applied to database before it becomes effective in real world.

Retroactive Update: The updates that are applied to database after it becomes effective in real world.

Simultaneous Update: The updates that are applied to database at the same time when it becomes effective in real world.

What are the different types of JOIN operations? **Equi Join:** This is the most common type of join which involves only equality comparisons. The disadvantage in this type of join is that there