---

1) How do you define Teradata? Give some of the primary characteristics of the same.

Teradata is basically an RDMS which is used to drive the Datamart, Datawarehouse, OLAP, OLTP, as well as DSS Appliances of the company. Some of the primary characteristics of Teradata are given below.

- Is capable of running on Single-nodes, as well as multi-nodes.
- Parallelism is built into the system.
- Very much compatible with the standards of ANSI.
- Tends to act in the same way as a server.
- It is an Open System that basically executes for UNIX MR-RAS, Suse Linux ETC, WIN2K, etc.

2) What are the newly developed features of Teradata?

Some of the newly developed features of Teradata are: -

- Automated temporal analytics
- Extension in the compression capabilities which allows flexible compression of data about 20 times more data than the previous version.
- Customer associated innovation like tetradata viewpoint.

3) Highlight a few of the important components of Teradata.

Some of the important components of Teradata are: -

- Bynet
- Access Module Processor (AMP)
- Parsing Engine (PE)
- Virtual Disk (vDisk)
- Virtual Storage System (VSS)

4) Mention the procedure via which, we can run Teradata jobs in a UNIX environment.

All you have to do is perform execution in UNIX in the way as mentioned below.

$Sh > BTEQ < [Script Path] > [Logfile Path]

or

$Sh > BTEQ < [Script Path] TEE [Logfile Path]

---

5) In Teradata, how do we Generate Sequence?

In Teradata, we Generate Sequence by making use of Identity Column

6) During the Display time, how is the sequence generated by Teradata?

All you have to do is use CSUM.

7) A certain load is being imposed on the table and that too, every hour. The traffic in the morning is relatively low, and that of the night is very high. As per this situation, which is the most advisable utility and how is that utility supposed to be loaded?

The most suggestible utility here has to be Tpump. By making use of packet size decreasing or increasing, the traffic can be easily handled.

8) If Fast Load Script fails and only the error tables are made available to you, then how will you restart?

There are basically two ways of restarting in this case.

- Making the old file to run – Make sure that you do not completely drop the error tables. Instead, try to rectify the errors that are present in the script or the file and then execute again.
- Running a new file – In this process, the script is executed simply using end loading and beginning statements. This will help in removing the lock that has been put up on the target table and might also remove the given record from the fast-log table. Once this is done, you are free to run the whole script once again.

9) Mention a few of the ETL tools that come under Teradata.

Some of the ETL tools which are commonly used in Teradata are DataStage, Informatica, SSIS, etc.

10) Highlight a few of the advantages that ETL tools have over TD.

Some of the advantages that ETL tools have over TD are: -

- Multiple heterogeneous destinations, as well as sources can be operated.
- Debugging process is much easier with the help of ETL tools owing to full-fledged GUI support.
- Components of ETL tools can be easily reused, and as a result, if there is an update to the main server, then all the corresponding applications connected to the server are updated automatically.
- De-pivoting and pivoting can be easily done using ETL tools.

11) What is the meaning of Caching in Teradata?

Caching is considered as an added advantage of using Teradata as it primarily works with the source which stays in the same order i.e. does not change on a frequent basis. At times, Cache is usually shared amongst applications.

12) How can we check the version of Teradata that we are using currently?

Just give the command .SHOW VERSION.

13) Give a justifiable reason why Multi-load supports NUSI instead of USI.

The index sub-table row happens to be on the same Amp in the same way as the data row in NUSI. Thus, each Amp is operated separately and in a parallel manner.

14) How is MLOAD Client System restarted after execution?

The script has to be submitted manually so that it can easily load the data from the checkpoint that comes last.

15) How is MLOAD Teradata Server restarted after execution?

The process is basically carried out from the last known checkpoint, and once the data has been carried out after execution of MLOAD script, the server is restarted.

16) What is meant by a node?

A node basically is termed as an assortment of components of hardware and software. Usually a server is referred to as a node.

17) Let us say there is a file that consists of 100 records out of which we need to skip the first and the last 20 records. What will the code snippet?

We need to use BTEQ Utility in order to do this task. Skip 20, as well as Repeat 60 will be used in the script.

18) Explain PDE.

PDE basically stands for Parallel Data Extension. PDE basically happens to be an interface layer of software present above the operation system and gives the database a chance to operate in a parallel milieu.

19) What is TPD?

TPD basically stands for Trusted Parallel Database, and it basically works under PDE. Teradata happens to be a database that primarily works under PDE. This is the reason why Teradata is usually referred to as Trusted Parallel or Pure Parallel database.

---

20) What is meant by a Channel Driver?

A channel driver is software that acts as a medium of communication between PEs and all the applications that are running on channels which are attached to the clients.

21) What is meant by Teradata Gateway?

Just like channel driver, Teradata Gateway acts as a medium of communication between the Parse Engine and applications that are attached to network clients. Only one Gateway is assigned per node.

22) What is meant by a Virtual Disk?

Virtual Disk is basically a compilation of a whole array of cylinders which are physical disks. It is sometimes referred to as disk Array.

23) Explain the meaning of Amp?

Amp basically stands for Access Module Processor and happens to be a processor working virtually and is basically used for managing a single portion of the database. This particular portion of database cannot be shared by any other Amp. Thus, this form of architecture is commonly referred to as shared-nothing architecture.

24) What does Amp contain and what are all the operations that it performs?

Amp basically consists of a Database Manager Subsystem and is capable of performing the operations mentioned below.

- Performing DML
- Performing DDL
- Implementing Aggregations and Joins.
- Releasing and applying locks, etc.

25) What is meant by a Parsing Engine?

PE happens to be a kind Vproc. Its primary function is to take SQL requests and deliver responses in SQL. It consists of a wide array of software components that are used to break SQL into various steps and then send those steps to AMPs.

26) What do you mean by parsing?

Parsing is a process concerned with analysis of symbols of string that are either in computer language or in natural language.

---

27) What are the functions of a Parser?

A Parser: -

- Checks semantics errors
- Checks syntactical errors
- Checks object existence

28) What is meant by a dispatcher?

Dispatcher takes a whole collection of requests and then keeps them stored in a queue. The same queue is being kept throughout the process in order to deliver multiple sets of responses.

29) How many sessions of MAX is PE capable of handling at a particular time?

PE can handle a total of 120 sessions at a particular point of time.

30) Explain BYNET.

BYNET basically serves as a medium of communication between the components. It is primarily responsible for sending messages and also responsible for performing merging, as well as sorting operations.

31) What is meant by a Clique?

A Clique is basically known to be an assortment of nodes that is being shared amongst common disk drives. Presence of Clique is immensely important since it helps in avoiding node failures.

32) What happens when a node suffers a downfall?

Whenever there is a downfall in the performance level of a node, all the corresponding Vprocs immediately migrate to a new node from the fail node in order to get all the data back from common drives.

33) List out all forms of LOCKS that are available in Teradata.

There are basically four types of LOCKS that fall under Teradata. These are: -

- Read Lock
- Access Lock
- Exclusive Lock
- Write Lock

----

34) What is the particular designated level at which a LOCK is liable to be applied in Teradata?

- Table Level – All the rows that are present inside a table will certainly be locked.
- Database Level Lock – All the objects that are present inside the database will be locked.
- Row Hash Level Lock – Only those rows will be locked which are corresponding to the particular row.

35) In the Primary Index, what is the score of AMPs that are actively involved?

Only one AMP is actively involved in a Primary Index.

36) In Teradata, what is the significance of UPSERT command?

UPSERT basically stands for Update Else Insert. This option is available only in Teradata.

37) Highlight the advantages of PPI(Partition Primary Index).

PPI is basically used for Range-based or Category-based data storage purposes. When it comes to Range queries, there is no need of Full table scan utilization as it straightaway moves to the consequent partition thus skipping all the other partitions.

38) Give the sizes of SMALLINT, BYTEINT and INTEGER.

SMALLINT – 2 Bytes – 16 Bites -> -32768 to 32767

BYTEINT – 1 Bytes – 8 Bits -> -128 to 127

INTEGER – 4 Bytes – 32 Bits -> -2,147,483,648 to 2,147,483,647

39) What is meant by a Least Cost Plan?

A Least Cost Plan basically executes in less time across the shortest path.

40) Highlight the points of differences between the database and user in Teradata.

- A database is basically passive, whereas a user is active.
- A database primarily stores all the objects of database, whereas a user can store any object whether that is a macro, table, view, etc.
- Database does not has password while the user has to enter password.

41) Highlight the differences between Primary Key and Primary Index.

- Primary index is quite mandatory, whereas Primary Key is optional.
- Primary Index has a limit of 64 tables/columns, whereas Primary Key does not have any limit.
- Primary Index allows duplicates and nulls, whereas Primary Key doesn't.
- Primary Index is a physical mechanism, whereas Primary Key is purely logical mechanism.

---

42) Explain how spool space is used.

Spool space in Teradata is basically used for running queries. Out of the total space that is available in Teradata, 20% of the space is basically allocated to spool space.

43) Highlight the need for Performance Tuning.

Performance tuning in Teradata is basically done to identify all the bottlenecks and then resolve them.

44) Comment whether bottleneck is an error or not.

Technically, bottleneck is not a form of error, but it certainly causes a certain amount of delay in the system.

45) How can bottlenecks be identified?

There are basically four ways of identifying a bottleneck. These are: -

- Teradata Visual Explain
- Explain Request Modifier
- Teradata Manager
- Performance Monitor

46) What is meant by a Highest Cost Plan?

As per Highest Cost Plan, the time taken to execute the process is more, and it takes the longest path available.

47) Highlight all the modes that are present under Confidence Level.

Low, No, High and Join are the four modes that are present under Confidence Level.

48) Name the five phases that come under MultiLoad Utility.

Preliminary Phase, DML Phase, Data Acquisition Phase, Application Phase and End Phase.

49) Highlight the limitations of TPUMP Utility.

Following are the limitations of TPUMP utility: -

- We cannot use SELECT statement.
- Data Files cannot be concatenated.
- Aggregate and Exponential operators are not supported.
- Arithmetic functions cannot be supported.

---

50) In BTEQ, how are the session-mode parameters being set?

.set session transaction BTET -> Teradata transaction mode

.set session transaction ANSI -> ANSI mode

These commands will work only when they are entered before logging into the session.