TERADATA:

Lab08trg1 is a schema where it is a container and created automatically

CONCEPTUAL,INTERNAL EXTERNAL-SCHEMA LEVEL

Internal level-physically stored as 0 and 1

Memory management for variables will be continuous to each other in c program as like schema for tables obj structure

Max number will be 38 digits in DB

Every oracle block has header in which it has owner, no of tables, no of rows, obj present

Teradata has bucket in which it has amp number where it generates row hash values

Schema is a collection related obj structure of tables under user

External level-Select statement for all columns selected or particular columns or end user view

Key in TD:

Key which uniquely identifies tuples with no null-candidate key and td has more than one candidate key

Composite primary key-combines 2 primary key (unique with no null)

Foreign key-referential integrity

One to many (primary and foreign key)

Many to many (M:N or N:M)

Normalization:

2nf avoids partial dependencies

3nf avoids data redundancy and transitive dependency

Transitive dependency-indirect relationship a->b->c

Intra key dependency problem-if any of 2 candidate key depends on each other which rarely happens

BYSCODE FORM is very rare (extended 3nf)

TD:

TD ARCHIECTURE-basic unit of td is node and architect is logical

Node-memory,system,server,files….

PE-virtual processor and takes query and compile it and generates 32 bit row hash value and 32 bit unique total is 64 bit

Amp-virtual processor and it has 4 amp by default

Bynet-message parsing layer and it has default BYNET 0 AND 1 and receives from amp and gives it to pe

Components of PE- dispatcher, optimizer, session handler, parser (execution plan)

Parser-hard/soft parsing

* Parser: The Parser checks for the syntax, if true forward the query to Session Handler.
* Session Handler: it does all the security checks, such as checking of logging credentials and whether the user has permission to execute the query or not.(default 120 sessions)
* Optimizer: It finds out the best possible and optimized plan to execute the query.
* Dispatcher: The Dispatcher forwards the query to the AMPs.(it creates locks)

Oracle doesn’t have pipelining parallelism by default ->main drawback of oracle where it prefers bulk load.

In TD it creates 120 sessions for faster execution of large volume of data.

Our Td is MPP(Broadcast)

MPP- broadcast, multicast, pt to pt

Hash map algorithm-looks 1st bucket for amp number

Bucket-collection of amp number(amp details)

Amp number stores unique row id

Default table created – set

First column will be NUPI (default)

Failures:

1)node failure- it is resolved by CLIQUE mechanism(collection of nodes where failure node will be moved to HSN NODE)

2)VDISK failure –it is resolved by raid mechanism

3)amp failure –it is resolved by fallback mechanism

Transient Journal:

Concurrently writing to a page which is maintained by server and has transaction details of data in Teradata

Set will not allow duplicate rows and multiset can have duplicate rows.

Sample-will not allow duplicates

Sample with replacement-allows duplicates

1)vproc

Pe and amp

2)clique

Set of nodes

3)foreign key features

Identify table realitionship

4)primary key

Set of columns uniquely identify row

5)obj of pi

Ensures enough amp involved to reduce query processing time

6)what happens if pe failed

All channels will migrate to another node in clique

7)architecture of td

Shared nothing

8) features of td architecture

We require amp access and

9)when a user create small temp table in spool space which table type used

Both volatile and derived tables

10)adv or features of views

It provides additional authority or security for the table

11)minus operator equivalent to

Except

12)full outer join

Matching and unmatching rows from both table

13)adv of upi

It enforces uniqueness on primary key column

14)which column is affected by distinct key

For both columns(all columns will be affected,distinct is written only once after select statement)

15)pe

Vproc

16)what will u consider for oltp applications?

Few of many tables can be accessed

17)working in td mode

[Td sql assistant(default creates set table)-teradata mode

Bteq(default creates multiset table)-ansi mode]

Td sql assistant

18)Insert(write lock),select(read lock) ,update(write lock), there is a failure in update, what happens?

All locks will be released

19)purpose of merge

Can operate on identity col and only single row statements are supported

20) If we give alter(used only for permanent table) table table\_name, which referential integrity option is processed

With check option

21)which journal can be defined for table

Permanent local after journal

22)what is compulsory whether to calculate summary results while create summary table

Volatility of col to summarise

23)how to see if the sys is performed in optimal way…

Statistics collect on all indexes

24)if u want to ensure if a person is working in current consistent data what lock to be consider for select statement

Read lock

25)in explain query ,we have locked ds name.table name for access means..

Table level access

26)check constraint

table level constraint is applied

27)there is a select statement, with subquery as inner query returns null values with connecting condition NOT IN in outer query

The first table returns all non-null rows

28)we created a table with col1,2,3,col is date and we partitioned the date interval from 2002 to 2012 ith interval 1 year,no range

rows are moved from dropped partition to no range partition

29)there is table of some gb of active dataware house environment approx. 30% of data is updated on nite on bteq ,a table remains avail for 24\*7 for update from user access &update

Implement daily online archive

30)which query has shortest access path….

Store key option

31)u have a table 1,2 is date,3 is varchar,again range by partition form 2001 to 2005 interval 5 day,query as select col from table\_name where date=15\* apr,2004 which parttion will be considered for the above query

Single partition

32)how to create a empty table form existing table

Create table t1 as another\_table\_name with no date

33)there is query,

select city,id

from orders

sample with replacement

when city =ny then 2,4

when state =au then 5,7

else 50

order by 2

Assume 10 rows from each city is order by table then how many distinct city will be there….

->5

**Soln:**

each city has 10 rows

for city ny,

(2+4)\*10=60 rows, where 6 rows will be generated for 10 values in the city ny

For state au,

(5+7)=12 rows

Distinct city=60/12=5 rows

33)macro with separator, write a simple query….

Create macro macro\_name as echo

34)which of the join doesn’t require all amp operation

Nested

35)modify priary index ref integrity option in avail with alter table option

false

36)diff b/w active and traditional data ware house

Data freshness(avail only in active DWH)

37)with check option and with no check option. what are the benefits

During insertion

38)u have 8 node sys with two 4 node clique. we have to insert another 2 node clique in hard stand by node.what is use of hsn

During node failure the hsn receives migrating vproc from the service level of sys

39)how many amp in cluster provides max security against sys failure

4 amps

40)which ref integrity characteristics will inaccurately affects performance

Need for the existence of ref table

41)best feature of active DWH

Low cost of disk to data ratio

42)how will you avoid full table scan

Use of where clause in PI in the query

43)define transaction

Logical unit of work

44)select city,state,sample id

from table

sample with replacement

when state=wi then 4

else 3

order by 3

How many distinct values are then. Provided 10 rows for all…..

->2

**Soln:**

For wi we have same id and else state have id so it will return 2 rows

45)SELECT \* FROM TABLENAME WHERE COL=1

SELECT \* FROM TABLENAME WHERE COL=2

What kind of parallelism exists?

Multi statement

46)which col type supported with compression

integer

47)imagine table is empty

select avg(sal) from tablename

No records found

48)a new work load management is established without user and apllication,there is dba what are the information to collect

TD architecture

49)i want delete data faster…

Multiload delete

50)limitation of tpump

Arithmetic fn are supported

51)when multiple values of col are compressed ,where will be compressed value stored..

Table header

52)if u give help session cmd it display current ongoing session

True

53)what is true about fastload…

No journaling

54) what is true about multiload…

Upsert fn is supported

55)choose appropriate label rows per NUPI value

Set table with many rows with each NUPI has much worse insert performance than Multiset table with many rows with each NUPI

56)how will find space requirement for NUSI

Index value size

57)u writing a statement exp query spool.. which is duplicated on all amps. what type of join is it?

Product join

58)there is user, creating ref integrity on joining two tables in the view. What is true about statement…

If the data is requested by referencing table, td will automatically lock and take data from referenced table without joining

59) Feature of NUSI

Index row and table row on the same amp

60)u have composite primary key,how will u minimize the row distribution for composite primary key?

Defining NUPI col will minimize it(same indexed value on the same amp)

61)why we have to choose UPI than NUPI

To improve the performace of query using ranges

62)what option will give the sys availability

Hsn,transient journal

63)there is active 24\*7 app users active altering and batch processing.u have a task to select and update on partitioned table for month. which is minimal tech for ………

Using insert, select to copy the data into temp table & archive the temp table and using sql delete statement we can remove the last month data.

64)can we create view with aggregation fn or derived table with aggregation fn. what are the benefits for view vs derived table

Aggregation will always gets refreshed with updated data

65)purpose of denormalization on application coding?

It generally makes new application difficult for implementation

66)which statement describes locking level of Full Table Scan

It access each data only once

67)index level lock is not supported by TD

True

68)which is not correct about derived table

They may be defined with DDL

69)to\_\_\_\_\_\_ NUPI is better than UPI---->allow access with non unique value

70)which action can’t be done by users with no perm space allocation perform on sys

Index compression cannot be executed