**TERADATA BASICS NOTES**

[**http://www.testsahoy.com/it-certification/mock-tests-for-teradata**](http://www.testsahoy.com/it-certification/mock-tests-for-teradata)

[**https://atozmcqs.blogspot.com/2014/02/teradata-multiple-choice-questions-and.html**](https://atozmcqs.blogspot.com/2014/02/teradata-multiple-choice-questions-and.html)

/\*CREATE TABLE STUDENT\_161259(STUDENT\_ID INTEGER, STUDENT\_NAME VARCHAR(30));\*/

/\*SEL \* FROM STUDENT\_161259;\*/

/\*HELP TABLE STUDENT\_161259;\*/

SHOW TABLE STUDENT\_161259;

-------------------------------------------------------------

set---> will not allow duplicate row

multiset---> will allow

-----------------------------------------------------------------

PRIMARY INDEX ( STUDENT\_ID );

the above means it is NON UNIQUE PRIMARY INDEX

-------------------------------------------------

protection mechanisms

FALLBACK – TO PREVENT FROM AMP FAILURE, INSTEAD OF COPYING THE DATA FROM ALL AMPS IT USES FALLBACK.

JOURNAL

-----------------------------------------------------

**EXPLAIN** **SEL** \* **FROM** STUDENT\_161259**;**

1) First, **we lock** **TD\_BIM\_FR\_TRNG\_DB.STUDENT\_161259** for **read** on a

reserved **RowHash** to prevent global deadlock***(TO PREVENT LOCKING FROM OTHER USERS)…(TALKS ABT PSUEDO TABLE)***

2) Next, **we lock** **TD\_BIM\_FR\_TRNG\_DB.STUDENT\_161259** for **read**.

3) We do an all-AMPs **RETRIEVE** step from ***(FOR SEC.INDEX 🡪 TWO AMPS OPN &***

***FOR PRI.INDEX SINGLE AMPS OPN)***

**TD\_BIM\_FR\_TRNG\_DB.STUDENT\_161259** by way of an **all-rows scan** with

no residual conditions ***(SINCE WE USED NO WHERE CLAUSE)*** into **Spool 1** (group\_amps), which is **built**

**locally** on the AMPs. The size of **Spool 1** is estimated with low

confidence to be **80 rows** (2,960 bytes). The estimated time for

this step is 0.02 seconds.

4) Finally, we send out an END TRANSACTION step to all AMPs involved

in processing the request.

-> The contents of **Spool 1** are sent back to the user as the result of

statement 1. The total estimated time is 0.02 seconds.

**3 LEVELS OF CONFIDENCE**

**HIGH** --- QUERY FREQUENTLY USED AND HAVE ALL THE STATISTICAL INFO WITH WHICH WE WORKED (abt acces,abt data, abt nodes)

**LOW**----EXECUTED JUST 2-3 TIMES, DOESNOT HAVE ALL THE STATISTICAL INFO

**NO**---- 1ST TIME EXECUTION, NO STATISTICAL INFO PRESENT

BUCKET 🡪COLLECTION OF AMP, GIVES AMP DETAILS ALONG WITH ROW HASH VALUES

1ST HASH MAP A

PE (MATCHED VALUE OF AMP)

EMP

DEPT

EMP 5

DEPT 1

DEPT

DEPT

DEPT

LOGO. 🡪 BUCKET🡪 CHECKS FOR ROWHASH VALUE🡪PE STAGE

OCB🡪COLLECTION OF ORACLE BLOCK

VIRTUAL PROCESSOR🡪 FASTEST WAY OF PROCESSING

SMP OR MPP 🡪 EITHER SINGLE TO SINGLE PONT OR BROADCASTING

**3 DIFFNT. FAILURES**

1. IF NODE FAILS 🡪 CLIQUE(COLLECTION OF NODES) METHOD IS USED, HAS HOT STANDBY NODE (HSN-EMPTY NODE)

IF NODE 2 FAILS, THE NODE WILL 1ST GO TO HSN, TAKES DATA FROM THE AMPS(CHECKS FOR AVAILABLE NODES),DATA GETS DISTRIBUTED TO AVAILABLE NODES AND THEN IT RELEASES THE MEMORY AS NODE 2 IS FAILED AND THEN DISCARDS THE NODE

1. VDISK FAILS🡪RAID MECH. FOR PREVENTION (RAID 5 IS USED)..parity bit is used
2. AMP FAILS🡪 FALLBACK IS USED (4->1,1->2,2->3,3->4)

**RECENTLY USED RAID 10**

RAID 1🡪S/M PERFORMS MIRRORING

RAID 1&5 🡪 TERADATA SUPPORTS

RAID 5🡪 ATTACH A PARITY BIT IN THE CHANGED ROW

INBUILT INFRA.

JOURNAL🡪 TO HAVE AN ENTRY OF DML DATA TO JOURNAL, DURING ANY NEED THE SYSTEM MATCHES ALL THE RECORDS BY ITSELF

Transient journal…concurrently writing into a page.. maintained by td server

**1st Nov**

Index

**PRIMARY** 🡪 UPI & NUPI (PRI.CAN HAVE NULL VALUES,LOGICAL)

ONLY 1 PRI.INDEX WILL BE PRESENT IN THE TABLE(UPI OR NUPI)

BY DEFAULT NUPI WILL BE ASSIGNED

UPI INDEX WILL STORE DATA IN EVEN DISTRIBUTION

NUPI WILL NOT EVENLY DISTRIBUTE THE DATA IT DITRIBUTES BASED ON INDEXES

BOTH THE INDEX WILL GO FOR SINGLE AMP Opn.

**SECONDARY** 🡪 USI (2 AMP OPN) (LESS VALUE RANGE,LIMITED OPTIONS) & NUSI (2 OR ALL AMP OPN, WHEN NO INDEX IS SET) (WIDE RANGE VALUE, WIDE OPTIONS)

**BOTH HAVE INDEX SUBTABLE**

**ALL AMP METHOD IS BETTER BECAUSE WE PROVIDE PARALLELISM**

TD KEEPS 1ST COL AS INDEXED COLM.

IF A TABLE IS DROPPED PRI.INDEX IS ALSO DROPPED

CAN ALSO USE DROP INDEX INDEX\_NAME

PE🡪ROWS

BYNET

ROWID’S (AMP)

IF QUERY IS SEARCHED ON BASIS OF USI OR NUSI IT TAKES

**FOR NUSI**

PE

BYNET

SUBTABLE

BYNET

32 SEC.INDEXES CAN BE CREATED

**USI ARE ROW HASH DISTRIBUTED**

NUSI ARE UNEVEN SEARCHING, IT TALKS ABT ROWIDS

ALL AMP RETRIVAL IS DONE IN NUSI

PARTITION PRI.INDEX

2 TYPES 🡪RANGE & CASE

**UTILITIES 🡪 ADDITIONAL OPTIONS FOR THE FILE TO TABLE EXCHANGE**

*TD utilities*

BTEQ 🡪 EXPORT & IMPORT (ONLY FOR SMALLER OR MIN. AMOUNT OF DATA) **(LOADS FROM TABLE TO FILE)**

FAST EXPORT **(LOADS FROM TABLE TO FILE) (.fs)**

MUTI LOAD (.ml)

FILE TO TABLE

FAST LOAD(.fl)

TPUMP

IF THE TABLE IS EMPTY **FAST LOAD** WILL **INSERT** (INSERTION ONLY), WORKS FOR EMPTY TABLES ONLY, ONLY FOR ONE TABLE MAX. **(.fl)**

IN **MULTI LOAD** BOTH **UPDATION** AND **INSERTION** CAN BE DONE (UPSERT) **(.ml)**

**MULTI LOAD** WORKS FOR EMPTY & NON EMPTY TABLES **(5 DIFFNT OPNS CAN BE PERFORMED)**

**TPUMP** IS SIMILAR TO MULTI LOAD OPN….IT TAKES RECENT CHANGES INTO THE TABLE

**Max of 15 jobs can be done at a time in fast and multi load**

**TO PERFORM FASTLOAD AND MULTI LOAD**

IF THE TGT TABLE HAS SEC.INDEX, JOIN INDEX, FOR.KEY CONST, TRIGGERS.

**BTEQ MODE (DEFAULT MODE) ,** EXTN .btq

*TYPES OF MODES*

DATA/RECORD MODE

INDIC DATA exacty like data mode, rep or indication for null values in this mode num nulls 🡪0 and charc nulls 🡪space

REPORT/FIELD human readable format, looks like tabular format with row and cols

DIF **DATA INTERCHANGE FORMAT**

**.IMPORT FILE=D:\LOGAVIGNESH\bteqexp.txt;**

**(OR)**

**.IMPORT VARTEXT ‘,’ FILE=D:\LOGAVIGNESH\bteqexp.txt;**

.logon 10.58.0.66/TD\_USER21,TD\_USER21;

DATABASE TD\_BIM\_FR\_TRNG\_DB;

.IMPORT VARTEXT ',' FILE=D:\LOGAVIGNESH\emp\_val.txt;

.REPEAT \*

USING(A VARCHAR(30),B VARCHAR(30),C VARCHAR(30),D VARCHAR(30))

INSERT INTO new\_emp\_161259 VALUES((cast(:a as integer)),:b,(cast(:c as integer)),(cast(:d as integer)));

.LOGOFF

**2ND NOV**

**FAST EXPORT – BY DEF. THE NO. OF EXPORT(SESSIONS) CREATED IS 4, WORKS ON DATA & INDIC DATA**

.LOGTABLE EMP\_LOG;

.LOGON 10.58.0.66/TD\_USER21,TD\_USER21;

DATABASE TD\_BIM\_FR\_TRNG\_DB;

.BEGIN EXPORT SESSIONS 10;

.EXPORT OUTFILE ‘D:\LOGAVIGNESH\TEMP1\_FS.txt’;

SELECT \* FROM emp\_161259;

.END EXPORT

.LOGOFF

**FASTLOAD WILL WORK IN 2 PHASES,** **ACQUISITION** (*DATA GETS LOADED FROM THE FILE TO TABLEIN UNSORTED MANNER, THEN MOVES TO APPLICATION PHASE)*

**AND APPLICATION** *(TAKES DATA FROM WT TO AMPS)*

S/M ITSELF WILL CREATE THE WT

CREATE ONE TABLE WITH EXTN AS .ET OR .UV

**MULTI LOAD HAVE 5 PHASES,**

**1)PRELIMINARY, (ALLOCATION OF MEMORY)**

**2)DML TRANSACTION , (USING DML)**

**3)ACQUISITION, (WORK TABLE)**

**4)** **APPLICATION, (APPLY THE CHANGES)**

**5)CLEAR STAGE(CLEAN UP)**

**COMMAND USED HERE IS MLOAD**

**Worktable** 🡪 WT

LOADING NEEDS AN INTERME.TABLE WHERE THE DATA COMES IN AN ORDER

**Error table** 🡪 ET🡪fast and multi load need 2 error tables

**2 types of error table, ET (ERROR TABLE) & UV (UNIQUE VIOLATION, DUPLICATE ENTRIES WILL BE CHECKED)**

**When loading there shouldn’t be any ET and UV tables**

FAST LOAD AND MULTI LOAD SHOULD NOT HAVE NUSI

**27th nov,18**

Database is a coll of tables.

Dbms is database management system

Rdbms… min 6 princi,

Pri key… uniq identifies each tuples(row)

Col…attributes

File…relation

Lab 08 trg# is the schema

Schema(logical way of organizing…coll. of multiple obj cretaed) is created automatically, which has related obj in it

**Conceptual, internal, external level of schema**

**Internal**….physically stored as 0 nd 1….here it s stored as str in c program(heterog coll. in one place)

Mem. Management is well organized in c program str

Max number can have 38 digits in database

Ocb…oracle block….evry ocb has a header….displays owner, obj present

Conceptual…syntax for the db

External…..select statements…end user point of view

Key,….which uniq identifies each tuple and not involve null is candidate key

More than candi. Key allowed

Composite pri. Key(combining 2 or more primary key)

If that is uniq nd not null is called composite pri. Key

Foreign key…referential intergrity(tab of one is related with tab of another)

One to many(relation b/w pri.key and for. Key)

Many to many(m:n or n:m)

Normalization…completely normal is 3nf…avoid redund….

2nf… avoid partial dependency…if any col totally depends on primary key it is totally dep…partially on pri. Key and along with some other col. It is partially depnt

Transitive dependency…2) if a key col. Does not Depends on any of the col …1) a to b , b to c

Boys code normal form….. very rare…it is the extension of 3NF

If any of 2 candidate key depends on each other it is called intra key dependency….it happens rarely

Mostly all tables are in 2nf

**Archi**

PE..parsing engine…is a virtual processor..takes quesry form user…compile…generates 32 bit row hash value..and 32 bit totally 64 bit..

Basic unit of td is node

Node is…mem…op system..server…

All archi are logical

Amp ….is virt. Processor…

Bynet…form pe it comes…2 bynet layer…bynet 0 nd 1

Bynet…msg passing layer..where pe gives req. and it gets passed via bynet…receives from amp and give to pe

By def bynet has…4 amp opn

Amp is connected to vdisk

Compo present in pe… Dispatcher,optimizer, session handler, parser

Parser…either hard/soft parsing….if plan is there it goes for soft parsing….

Opti…optimizes the code..best way is found

Sess.handler…120 sessions are created for single pe…. pipelining and parallelism

Drawback of oracle…pipelining and parallelism…block by block movememt in oracle

In td..by def…in a single pe it creates 120 session max..for a single slct statement

Dispatcher…. Will avoid dead lock situations….

Mpp..pt to pt..broadcast..multicast…usually broadcasted

Hashmap..finds a bucket…which gives amp no.

Bucket ..col. of amp details.. from amp row id are takes..

Generally set tales are created….NUPI is made by def… where no duplicates is allowed

Set doesn’t allows duplicates…multi set allows

1 What are vproc…..Pe and amp

2 def of cliques…. Collection of nodes

3 features of foreign key? Identify tab relation

4 pri key… set of col. That uniq identify a row

5 obj of primary index…. To ensure enough amp are invloved ,To improve query performance

6 what happens if pe fails…. All channel pe will migrate to another node…

7 archi feature… shared nothing

8 we req amp access ….PE communicates thro bynet & AMP associates from vdisk

9 which table is used when a user want to create small temp table but only spool space is there….ans is both volatile and derived

10 advan of views…. Provides an additional security for the table…

11 minus is equal to…. except

12 Advan of usi…both 1 and 2..enforces uniqness on a pri key col & improves speed

13 which col is affected by distinct key wrd….it gets appli. For both cols(distinct is written only once, that too infront only, immediately after slct statemement, combi. Shud not be same)

14 parsing is typ of…vproc

15 wat will u consi. For oltp app….few of many tab must be accessed

16 working in td mode … betq(;) ansi mode.. by def multiset, td sql assistant(execute)…. no termination ; …by def set tables

Insert(write), slct(read), update(write) is given,,there is a fail in update, then? Then all locks are released after failure

17 purpose of merge……can operate on tab with identity col and only single row statements are suprtd in merge

18 if u give alter tab table name(for perm tab) which integrity option is present …with check option

19 which journal can be defined for tab….perm. local after journal

20 wat is the compulsion to consider while creating summary table or cal. Summary res on demand….Volatility of col. to summarize

21 how 2 see if the sys has performed in a optimal way…..statistics collect on all indexes

22 if u want to ensure a person is wrking in current consistent data wat lock is to be wrked out for slct query ……read access

22 wat is true abt alter cmd…

Check constraint…..a tab level check const is applied

23 exp query we have locked dn name.tab name for access means……table level access

24 there is sclt statement with sub q where sal not null....

25 we have slct statement where col name not in again col name from tab 2…………1st tab will ret. all the not null rows

26 we created a tab col1,2,3, col is date and we partition the date from 2002 to 2012 interval 1 yr…..rows are moved from the dropped partition from the no range partition

27 there is a tab of some gb of active environment aprrox 30% of data is being made updated on nite on bteq a tab remain avail for 24\*7 for update from user access & update …….implement the daily online archive

28 which query has shortest access path……… store key option

29 u hav a tab 1 int 2 date 3 varchar again range by partition is made from 2001 to 2005 with interval of 5 day, qry slct col from tab name where date=15th apr,2004, which partition will be there for the above qry……..ans is single partition

30 create tab tab name.. how to create empty tab with same tab definition…..create tab tab name as new tab name with data

31 select city id from order

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 in order by table then how many distinct city will be present ……….options are 1, 5, 32……. ans is 5

**solution**

* **ny is 2,4 so 2+4 =6 , for each city 10 rows, so 6\*10=60 combi for city**
* **for state 5,7 so 5+7=12**
* **therefore uniq id is 60/12=5**

32 macro with separator write a simple query…….create macro macro name as echo

33 which join method does not req. for all amp ….. nested

34 modify primary index ref. integrity option is avail with alter tab statement…false

35 diff b/w active and traditional data warehouse...... data freshness

36 with no check and with check option wat is the benefits…… during insertion

37 u have a 8 node Sys with 2 four node clique, now we should insert another 2 node clique in the HSN…….. during node fail the hsn receives the migrating vproc from the service level of the system

38 how many amps in the cluster provide max security against the system loss …..4

39 which ref. integrity(for. Key) characteristic will inaccurately affect the performance………. The need for the existence of ref. table

40 best feature of active dataware house……. Low cost of disk to the data ratio

41 how to avoid the full table scan….use of where clause in the query

42 define transaction…. Logical unit of work

43 slct city, state

Sample Id from table

Sample with rep.

When state=wi then 4

Else 3

Order by 3

How many distinct val(sample id) are there…… 2

Imagine 10 rows are the for all

44 slct \* from tab name where col=1

Slct \* frm tab name where col=2

Wat kind of parallelism will exist?

Ans is …… multi statement

45 wich col type is suprtd with compression?....integer

46 imagine a tab is empty

Slct avg(sal) form tab name Result?............. no recd

47 a new wrk load management practice is being estab. Without the application of user, there is a dba , which collect the information, what are the info to be collected?...... Teradata architecture

48 delete data faster… multi-load delete

49 limitations of TPUMP…..arithmetic funcs. are not suprtd

50 when multiple vals for a col is compressed, where it gets stored…… table header

51 help session displays current user

52 true abt fast loading……. No journaling

53 abt multi load…. Upsert is suprtd

54 choose the appropriate label abt rows per NUPI value…… set table with many rows for each nupi has much wrst insert perfm. Than a multi set tab with many rows for each nupi

55 how will u fine the space req. for NUSI?.... index value size

56 ur writing an explain qury spool… which is duplicated on all the amps wat type of join it is?...... Product join

57 there is user, creating a ref.Integrity on 2 tab joining the view, wat is true abt the statement……. If the data taken from referencing table then the td will lock and take the data directly from referencing table

58 wat is the feature of nusi…… index row and table row on same amp

59 suppose we have a composite pri. Key how to minimize row distribution? .......... defining nupi col will

60 why to choose upi instead of nupi? .... to improve query performance of the table

61 which option gives sys availability….. hsn & transient

62 active 24\*7 app uses active altering and batch processing for slct and update on a particular tab, which is a partition tab.

Ans is…… using insert slct to copy in to the temp tab & archive the temp table and sql del. we can remove the last month data

63 which is better, use of creating a view with aggregation func. Vs table with aggregation? ........ aggr. Is always getting refreshed with current data

64 effect of denormalization on application coding? .......generally makes new application more difficult to implement

65 which statement describes the locking level of full table scan (fts)….. access each data only once

66 index level is not suprtd in td? T or f… true

67 which is not crct abt derived tab? ……. They may be defined with DDL

68 to **\_\_\_\_\_** NUPI is better pri. Index than UPI? ……. Allow access with non uniq values

69 Which action cant be done with no perm space allocation perform on the system……. index compression