

- 1) Difference between open studio to enterprise edition
 - ➔ Cron tab , manual code movable
 - ➔ Joblets, tac, cdc, test cases, dynamic schema, code collaboration, log server(job fails automatically captured),project preferences ,AMC, code reusability
- 2) Migrating data from one database to another what are the issues we are facing
 - ➔ Data type issues
 - ➔ Pseudo columns
- 3) Components u r using to store the data in databases?
 - ➔ Toracleoutput component
- 4) What about toraclerow component?
 - ➔ We can execute the sql statements like update and insert data like that
 - ➔ Single and multiple queries we can execute
- 5) Testing approach does u have in your project?
 - ➔ Put some limits while extracting data
 - ➔ While loading into the target first we will check with the tlogrow weather data is coming or not
- 6) Bulkcomponents
 - ➔ Before inserting the data row by row. It create one temp file and load it into the target
 - ➔ Limitation: while inserting first time, table should be empty
- 7) Scenarios: latest file we need to fetch
 - ➔ Tfilelist➔ order by date asc/desc
- 8) Tmap and tjoin difference?
 - ➔Tmap: 1 as main & remaining "n" i/p as lookup || tjoin: 1 as main,1as lookup
 - N o/p || 1 as main & 1 as inner join reject
 - left outer join &inner join || inner join
 - first, all, unique matches || unique
- 9) If job is running more than 30 minutes we need to stop the job how you will do that?
- 10) Incremental loading?
 - ➔Insert record based on the last execution.(generated key)
- 11) Tjava,tjavarow,tjavaflex
 - Tjava-code executes only once.(own java code)
 - Tjavarow- code executes for each row
 - Tjavaflex-combination of both
- 12) How to avoid duplicate rows in talend?
 - ANS: tUniqueRow
 - ➔ If it is in database level we can write query as distinct option
- 13) Input having ab, bc, cd, ef
 - How can i show ab,
 - bc,
 - cd,
 - ef
 - ans: normaliser.
- 14) 5 independent sub jobs in and job got failed at 3rd job do 1 & 2nd jobs values gets inserted to database and how to overcome the error and insert values to db
 - ➔ I will do toraclecommit component at postjob once all the jobs got executed.
 - ➔ If this scenario we will put a checkpoints on onsubjob ok. We can execute those job with those check points
- 15) Input abcxzy@gmail.com how can i show output only "abcxyz".

- ➔ Left(mailed ,index(mail9d,@))
- 16) have you used right join in tjoin -- Only Inner join
 - ➔ by change the main as look up and look up as main we can implement rightouter join
- 17) How can you get data using ftp in talend? (Deloitte)
 - ➔ Ftpget component
- 18) What is the flow transforming data from one database to another? (Deloitte)
- 19) How you build the jobs and deploy them and how you schedule (Automate) the jobs? (Deloitte)?
 - ➔ Nexus(compiled code),GIT(generate code)
 - ➔ Generated-deployment-run
 - ➔ Triggers(normal trigger,cron trigger,file trigger)
- 20) What are triggers in pl/sql? (Virtusa)

Triggers are stored programs , which is fired when a DML statements executes.
- 21) How to get Second highest salary from a table? (Virtusa)
 - ➔ Select sal from (select sal,dense_rank over (order by sal desc) r) where r=2;
 - ➔ SELECT MAX(TaxRate) from Sales.SalesTaxRate where TaxRate Not in(SELECT MAX(TaxRate) from Sales.SalesTaxRate)
 - ➔ SELECT TOP 1 TaxRate FROM (SELECT TOP 2 TaxRate FROM Sales.SalesTaxRate ORDER BY TaxRate DESC) AS MyTable ORDER BY TaxRate ASC;
- 22) What is the transformation components used in talend? (Virtusa)
 - ➔ tmap, tAggregateRow,taggregatesortrow
- 23) 12. A --> B --> C, A is an input and B is an output which should pass into C o/p file in single mapping using TMap component. (Virtusa)
 - ➔ We add the tjava component in between the two output components
- 24) Have you used the Garbage collectors(GC) in Talend? (Virtusa)
 - ➔ usually we will get this type of errors when we use huge components like tsortrow,taggregaterow,tmap(sort on disk and use temp data option is we need to choose)
 - ➔ increase jvm parameters(by default it will be 256mto1024m)
- 25) What is the unique option in tJoin which is not in tMap? (Virtusa)
 - ➔ We can collect the rejected output data in tjoin. Where as in tmap we won't observe this option
 - ➔ Scenario :---

files:

file: 1- -customer file:-customer name, customer id, country id

file: 2-- country:-country id , country name

matching record in one output and non matching records in another output
- 26) How do you find out the DataType error while you build a job with no time? (Virtusa)
 - ➔ Bottom of the tool we can see the Code tab.
- 27) Difference b/w ETL & ELT? (Accenture)
 - ➔ETL: data is extracted, done transformations & loaded into target.

→ELT: where data is extracted, then loaded into a staging table in the database, transforming it in the database and then loading it into the target .

28) Real time scenario on DeNormalization concept? (Accenture)

29) Features of TMap? (Accenture)

→Store on disk to avoid heap memory issues

→Parallel option(lookup)

30) Difference b/w tAggregateRow and tAggregateSortedRow? (GenPact)

Taggregaterow→accepts i/p even it is not stored and perform sum ,avg.

Standard & union can also be performed.

TaggregateSortedrow→accepts i/p only when sorted.

No.of rows count to be specified.

31) About getting files from one server to another server (tFTP)? (GenPact)

32) tSystem, tLogCatcher component? (GenPact)

tSystem→Calls other system processing commands already there and run tat in other job.

tLogCatcher→catches error based on the trigger of Java exception, tDie or tWarn.

33) Multitasking concept? (GenPact)

Run subjobs in parallel (performance tuning).

34) Job got error after transforming data then how i need to resume the job from the error row? (GenPact)

35)About RowNum Analytical Function in Talend? (GenPact)

→Analytical funtions are quite useful in SQL and help us avoid extra coding..

→Based on the id it generates the rownum

35) Auto commit concept in talend? (GenPact)

auto commit function commits each SQL statement as a single transaction immediately after the statement is executed

36) How do you handle errors (tLogCatcher component)? (Accenture)

→tlogcatcher

37) About tParallelise compound? (GenPact)

run the specific jobs on parallel ...tparallelize(parallel, synchronize options)

38) Roles And Responsibilities

39) Factors considering before designing the job

Test case,memory management,dynamic sql syntax,parallise poptions

40) Project explanation

41) How the code will be moved to the QC

42) QC knowlegde about talend

43) How code move to production

44) How to execute the job using build job

Run batch file in cmd prompt or scheduler in windows

.sh file and cron tab in linux.

45) Components frequently used in the project

46) Diff b/w buffer and hash

Buffer: 1 bufero/p with same metadata stored in memory.

Hash: "n" multiple hash output with diff data.

47) Advantages of buffer

the buffer can be read by parent jobs to send data back up to the calling parent job

48) Diff b/w tmap and tjoin

49) Explain diff join models and matches

Right,left,full,inner,cross & first,all,unique

50) Advantages of enabling the store temp data in tmap

10gb source→Sometimes we don't have enough memory in RAM ,use store on disk

- To overcome memory issues
- 51) Is it possible to set the value dynamically on hard disk?
- 52) Exception handling in talend
tlogcatcher
- 53) What is the job server(TAC):- scheduling jobs
- 54) How do job details or value available in AMC in TAC
- 55) How do AMC giving the statistical information about job
- 56) Scenario :-
Server and talend studio is diff
Files available in server, shouldn't download the file
- 57) how to get the files apart from FTP
winscp, filezilla
- 58) Diff b/w Integrity constraint and unique constraint
- 59) What is surrogate key?
Columns that declared as primary key instead of real key.
Speed increases
Sequence in oracle.
- 60) What is index?
Index is a small copy of db used to speed up the retrieve of data.
Clustered-row data in order.
Non-clustered-random.
- 61) Without using any extra components (tlogrow, tmap) is it possible to view the data from any input file?
→ If it table, we can use data from database(sql, mssql)
→ Drag and drop the input file right click>data viewer, there we can view the file data
- 62) Can we use code reusability in Talend?
→ Yes, we can use code in any number of times with the help of joblets
- 63) Can we run multiple sub jobs parallel in Talend?
→ Yes, we can run the sub jobs parallel with the option in the
job>extra>multithreading
→ In case if u want to run the specific jobs on parallel we can use
tparallelize(parallel, synchronize options) component
- 64) Difference between on subjob ok and on component ok?
→ On subjob ok: if subjob complete success then only next job will trigger same as on component ok means when that component executes then only next components will run.
→ Tfilelist(100 files)-> input->tbufferoutput
a) After completion of copying the data from file to buffer I want to copy that file from that file location to another file
b) If we put them on subjob ok to tfilecopy . it copies only the last file that copies
c) But if we mention on component ok it copies the all files once completion of the copy
- 65) Multiple tables, same schema/metadata. Single output file what u can do?
→ Append option is the in tfileoutputdelimited.
- 66) Input file; only name Mahesh records we need. What we can do?
→ Tfilterrow-> ename='mahesh'

- 67) Inputfile,only Mahesh records and filter output deptno desc ?
- ➔ Tfilterrow-> ename='mahesh'
 - ➔ Tsortrow-> deptno asc/desc
- 68) Bharath.txt file is there, how can we know that where that file is there?
- ➔ Tfilelist->tjava
 - ➔ In java:ctrl+space tfilelist.currentfilepath
- 69) In joiner condition if we use want to display matched records and un matched records in single output then..
- ➔ In tmap, + symbol on rightside top right, click on it. It will ask two options ie new output and create join table from. Select create output table from output1. Then I will come out as same main flow
- 70) Jvm arguments will use ram memory or heap memory?
- ➔ Jvm arguments will use only ram memory where as in tmap if u want use temp data it will ask temp data path, it will use heap memory
- 71) Can context variables can able to pass from child job to parent job?
- ➔ Yes, in child job tcontextdump component can be connected to tbufferoutput
 - ➔ Tbufferinput to context load
- 72) Can we pass context variables from parent job to child job?
- ➔ Yes, in chiljob>component>transmit whole context
- 73) How can we use context variables without declaring at repository?
- ➔ Yes, we can declare it by using tcontextload
 - ➔ Declare all context variables in a input file. Read the input file as filedelimited and drag and drop it in job connect it with tcontext load
 - ➔ Main drawback is, if u add any new context variable, we need to run again that job where as if you declare In repository we can easily drag and drop.
- 74) How to resume job execution from same where job get failed due to unexpected error in Talend?
- 75) What are the components you have used in Talend projects?
- ➔ Tmap,tschema compliance check, buffer input n output etc..
- 76) Explain about toracleconnection component?
- ➔ Establish connection
 - ➔ We can re use that connection throughout the job
- 77) Is it possible to use the single database connection which is defined in parent job at child job? If yes how you what we need to do?
- ➔ Yes, we can declare the single database connection and use the same connection in child table also.
- 78) If we receive data in integer column and we want to load it into sting column. how u can do this conversion?
- ➔ Integer(col_name).toString(method) ----integer to output string
 - ➔ String.valueOf(col_name) ---- string to output integer
 - ➔ By using tconverttype component we can convert/change datatypes
- 79) Suppose if there are 10 files in one folder and wants to get current file name how u will get this?
- ➔ Tfilelist>tfileinputcomponent>tmap

- ➔ Add column, ctrl+space tfilelist currentfile name
 - ➔ FYI, it displays the file names as how many records passed through the main job
- 80) While loading 15 records from source to target but unfortunately, due to network issue you have loaded 10 records only. Now, I want to load from 11th record. How is it possible?
- ➔ Tmap>lookup>join condition as pk
 - ➔ Isnull->insert
- 81) I have a folder which is having csv files with same delimiter/metadata, I want to load all the csv files target tables. How is it possible? Which components you can use?
- ➔ Tfilelist,tfileinputdelimited,tbufferoutput,tbufferinput,tfileoutputdelimited
 - ➔ Tfilelist,tfileinputdelimited,tfileoutputdelimited(append option select)
- 82) How many types of context variables in talend?
- ➔ Global: repository>context>add context variable
 - ➔ Local: job >component+context add
 - ➔ Flat file,tcontextload
- 83) What is the difference between delete and truncate?
- ➔ Where clause,particular records deleted, rollback option,DML command-delete
 - ➔ Where clause absent, truncate deletes all the data, rollback wont possible, DDL command-truncate
 - ➔ In delete triggers are fire., where as in truncate triggers can't be fire.
- 84) What is the difference between union and union all?
- ➔ Union- unique values only accepted and also sorting those values
 - a) Whenever we are going to use set operator there must be same datatype
 - ➔ Union all-unique values along with duplicates
- 85) How to find out the fifth highest salary?
- ➔ Select * from (select deptno,ename,sal,dense_rank() over(order by sal desc) r from emp) where r=5;
- 86) In my inputfile 200 records is there, how can we read first 50 records?
- ➔ Inputfile- limit 50
 - ➔ Tfileinput,tsamplerow(1..50)
- 87) What is the main difference between tJava,tJavarrow,tJavaflex ?
- ➔ The main difference between between tjava and tJavarrow is , tjava executes only once whereas tJavarrow iterates by each record
 - ➔ TJavarrow will have edit schema whereas tjava won't be there
 - ➔ Tjavaflex is the combination of tjava+tJavarrow
- 88) What u will do when u face Heap memory issue?
- ➔ Run console>advance settings>use specific JVM arguments(increase the values)
 - ➔ Some components like tsortrow(on sort disk at advance settings),tmap having (use temp data)
- 89) 100 records is there in input file. Split the output records into 10 records in each file?
- ➔ Inputfile,outputfile(advance settings>splits the output 10 records option)
 - ➔ Inputfile,tmap(sequence),tsortrow,outputtable

- 90) Input file has some data; in my output file I need output like first 100 records in one file, next 200 records in another file and so on?
- ➔ Multiple Tfileinputs(limit 100, header1 and limit 200 header 101 like so on)
 - ➔ Inputfile-tsamplerow(1..100,101..300,301..600 etc)
 - ➔ Inputfile-tmap(sequence number at variable declaration)-add multiple outputs
add one column in each output()
- 91) Which module you used in your project?
- 92) Difference between open studio and enterprise edition?
Code collaboration, dynamic schema, data viewer, code reusability.
- 93) Data warehousing concepts.
- 94) What are the issues you faced while transforming data from one data base to another data base?
- Data type error
 - Pseudo column
- 95) Components used to store output?
tFileoutputdelimited, toracleoutput
- 96) Toracle row
Executes the query.
- 97) Sql concepts
- 98) Testing process you have to do while developing the job?
- 99) What are the components you are used in your project?
Tmap,bulk components,log components,uniqrow
- 100) Difference between tmap and tjoin?
- 101) Difference between delete and truncate?
- ➔truncate—delete all the rows
 - ➔delete—delete rows using where clause.
- 102) On batch size, commit differences?
- ➔committing the rows in db within each batch rows together.
 - ➔insertion of bulk number of rows.
- 103) Routines?
- ➔ reusable custom java code
- 104) Tparallelize?
- 105) Tmap having two outputs. We need to run one after another based on key values?
- 106) Joins(fuller)
- ➔right,left join-tunite-tunique
- 107) Data modelling
- 108) Writing procedures, functions and Unix shell scripting
- 109) What type of sources you're getting?
Flat files,excel,xml,json,table
- 110) Tsendmail component in master job?
- ➔allows to send email directly using SMTP(success/error)
- 111) Heap memory issues
Jvm in job level
- 112) Log components?
Tlogcatcher,tdie,twarn,tassert
- 113) Functions vs stored procedures
- 114) Context vs global variables
- 115) Out of memory

- 116) 100 record in a file →target I need 1-95?
- 117) Nvl vs nvl2?
- 118) xyz@gmail.com, I need gmail.com as o/p?
- 119) How will you deploy ur job?
- 120) What are best practices you have done?
- 121) 2 cluster data load into into db?