1. What was the complex transformation u created

a. Normalizer transformation to create multiple rows out of single row
   manipulating the occurrences of the Key block of the row.

```
e.g. input  employee_id   dept    tasks
                112        HR      admin, interview, payroll


            output :
               112      HR   admin
               112      HR   interview
               112      HR   pay roll
```

 b. Used Informatica Metadata Exchange (MX) (Rep_Session_tbl_log, Rep_targ_tbls,
Rep_targ_Mapping, Rep_Src_Mapping etc)views to extract source, target, mapping, session
information's

c. Used parameter file (.txt file stored in the server file system) with input values for the Batch
            e.g.   [s_m_Map1] $$ACC_YEAR=2003 $$ACC_PERIOD=12
                     [s_m_Map2] $$ACC_YEAR=2003

2. How did u implement the Update Strategy multiple
                    Not sure : Is it ?

                    Target Update Override
                    By default, the Informatica Server updates targets based on
key values. However, you can override the default UPDATE statement for each
target in a mapping. You might want to update the target based on non-key
columns.
                    For a mapping without an Update Strategy transformation,
configure the session to mark source records as update.
                    If your mapping includes an Update Strategy transformation,
the Target Update option only affects source records marked as update. The
Informatica Server processes all records marked as insert, delete, or reject
normally. When you configure the session, mark source records as
data-driven. The Target Update Override only affects source rows marked as
update by the Update Strategy transformation.

Overriding the WHERE Clause
You can override the WHERE clause to include non-key columns. For example, you might want to
update records for employees named Mike Smith only. To do this, you edit the WHERE clause as
follows:

UPDATE T_SALES SET DATE_SHIPPED = :TU.DATE_SHIPPED,
TOTAL_SALES = :TU.TOTAL_SALES WHERE :TU.EMP_NAME = EMP_NAME and
EMP_NAME = 'MIKE SMITH'
Entering a Target Update Statement
Follow these instructions to create an update statement.
To enter a target update statement:

1. Double-click the title bar of a target instance.
2. Click Properties.
4. Click the arrow button in the Update Override field.
5. The SQL Editor displays.
5. Select Generate SQL.
   The default UPDATE statement appears.
6. Modify the update statement.
   You can override the WHERE clause to include non-key

columns.

7. Click OK.

3. **How did u handle error handling**
Use of filter to remove the existing records so only the on-existing will Flow through.

Basic error handling can be done using exp transformation to check for the known possible errors, if found give appropriate label to the rows, then pass them through router transformation and direct the rows to the error tables or the target.

Some of the generic checks which are common to most of the mapping e.g. check for zero, null, length etc can be defined as reusable expression transformation which then can used in different mappings.

4. **Strength and weakness of informatica**

1. Rows are processed one at a time, so calculation and checks among the rows is difficult, even though to some extent values can be stored in variables for further processing. This causes increase in the steps required to accomplish the job. E.g. first load the rows in a temporary table with some calculation, and then make this temp table the source for another tables for further row manipulations etc.

  Very good for load and extraction, however not so flexible to implement procedural programming logic.
  Every mapping needs a source and target, so sometimes Just end up using dummy source and target (dual tables) to build the logic.
 3. No Replace character function in 5.1 (6.0 has it)
 4. Ease of use, graphical representation of the ELT process, Maplets, and reusable exp transformation helps to standardize the ETL process across the organization.
5. Easy for knowledge transfer and maintenance.

Give an example of a complex transformation u created using Informatica

ETL Process in recent project and error handling

How Many Mapping u has in Ur recent Prj?

What is a repository

A Repository is a Meta data, which is stored in a centralized location, can be imported and exported to various systems stores information about all the components of informatica like information about sessions, mappings, transformation, mapplets. Session parameters, mapping parameters, local and global variable, reusable transformations

What are session Parameters , Difference between pre and post session parameters

Pre session parameters Ex:- like executing a stored procedure or a sql query dropping indexes to improve performance

Post Session Parameters Ex:- Recreating the indexes, EMAIL

Explain the ETL Architecture

Data Staging Area :- Temporary like also called operational data store one level down to data mart

Difference between PL/SQL and informatica. Why do u need informatica justify

Pl/SQl is free with oracle

Informatica is expensive

Easily u can load the data in a more sophisticated way using informatica

U can monitor the load process, Schedule Sessions etc in informatica

What are the Differences Between informatica 4.7 and 5.1 and 6.0

Router Transformation is available from 5.0 onwards

Debugging in Designer

Partition of session in session manager

In 6.0 complete heterogeneous targets one in oracle one in db2 into multiple targets

Data partitioning run as multiple sessions in informatica 6.0

Repository Server (New Component)

Workflow Manager (New Component)

Workflow Monitor (New Component)

Explain the process of moving ur data to production

1. Through Xml way of moving into production

2. Th scripts

System testing environment: - u will have already repository in test environment by xml way or th scripts,

Prepare the scripts for Tables, views with constraints, and indexes.

Change the session connections

Diff between lkp cache and unlkp cache

U can disable the option use lkp cache when u are creating a lkp tranfor
Performance issue if lkp is too large in size lkp cache also takes lot of space

What is diff between data warehousing and enterprise data warehousing

Widely used by whole enterprise data warehousing Total GE warehouse is called enterprise data warehouse Data warehouse build for all business in an enterprise is enterprise data warehousing and for a single business is data warehouse and module in a data warehouse is data mart

Ex:- GE capital , GE appliances, GE electrical each has their own data data warehouse

What was the challenging task u felt when u are working with informatica

What is sql override

Overriding a SQL in source qualifier or lookup for additional logic.

How do u fine tune a session

A mapping is running very slow how do u handle :- While creating a mapping in the development process I followed all the rules to fine tune like joiner transformation in the beginning reducing no of expressions In the beginning only I check how much time its taking to retrive At mapping level :- Source sql override (explain plan statistics)

Target we write to flat file then check the time againt to rdbms then u will know bottle neck is in the db or in the mapping Then we check network by calling network admin Then data cache index cache increase

How did u implement update strategy Complex Session scheduling in ur experience

Have u used a normalizer transformation where

No

Explain the difference between oracle normalization and data warehouse denormalization why do we need to denormalize in data warehousing

Reverse engineering is for the performance i.e. removing additional joins and decreasing the load on the OLTP system

How did u handle error handling

ETL-Row-Errors

Always I used have problems when I am retrieving data from flat files the data is not formatted arranged properly what I used to do is Using sql * loader I pulled flat file data into data They r temporarily tables then I used these temporary tables as source then applied transformations and loaded into target then deleted the temporary tables

Why did u choose this method?

Informatica taking lot of time to load data into source using sql * loader the job can be done in minutes

What are the different kinds of repots u have created in BO in the recent project

What are the different kinds of repots u have created in BO in the recent project Tabular, cross tab, Master detail, crosstab master detail, table multi master detail.

Can we have multiple conditions in a Lookup?
yes

Can we have multiple conditions in a Filter?
YES

How the flags are called in Update strategy?
IIF
0 - DD_INSERT , 1- DD_UPDATE , 2- DD_DELETE , 3- DD_REJECT

Is it possible to run the session other than Server manager? If so how?
YES USING PMCMD

What is the use of power plug?

For 3rd party connectors to sap, mainframe,Peoplesoft

What are all the versions of Power Mart/Power Center – 1.7/4.7/5.0/5.1 -??

Diff between lkp cache and unlkp cache

Look up cache Indicates whether the Lookup transformation caches lookup values during the session. When lookup caching is enabled, the Informatica Server queries the lookup table once, caches the values, and looks up values in the cache during the session. This can improve session performance. When you disable caching, each time a row passes into the transformation, the Informatica Server issues a select statement to the lookup table for lookup values.

When U are installing BO with what profile you will enter?

What is the difference between inner and outer join?

Inner is self join and outer join retrieves the information for the given condition and the information that dose not satisfy the condition

What is the syntax difference for using outer join in SQL Server / Oracle
In Oracle u use (+) symbol and in SQL Server u use *

How do you do requirement analysis?

This is a big question, we start with learning the existing system and interact with the end users for their requirements and then decide the steps that need to go in modeling, ETL and reporting

What is Exception?

Same as exception handler in Oracle

What is slowly changing Dimension?

    I guess you know the TYPE 1,2,3

What is slowly Growing Dimension?
The dimensions, which dose not have many changes for the warehouse, e.g. Region dimension dose not have many changes it may add a row in a year. You have a wizard in Informatica just go through

The Target mapping filters source rows based on user-defined comparisons, and then inserts only those found to be *new* to the target. Use the Target mapping to determine which source rows are new and to load them to an existing target table. In the Target mapping, all rows are current. Use the Target mapping to load a fact or dimension table, one in which existing data does not require updates.
For example, you have a site code dimension table that contains only a store name and a corresponding site code that you update only after your company opens a new store. Although listed stores might close; you want to keep the store code and name in the dimension for historical analysis. With the Target mapping, you can load new source rows to the site code dimension table without deleting historical sites.

What kind of Test plan? What kind of validation you do?
In Informatica we create some test SQL to compare the number or records and validate scripts if the data in the warehouse is loaded for the logic incorporated.

What is the usage of unconnected/connected look up?
We use a lookup for connecting to a table in the source or a target. There are 2 ways in which a lookup can be configured i.e. connected or unconnected

You can configure a connected Lookup transformation to receive input directly from the mapping pipeline, or you can configure an unconnected Lookup transformation to receive input from the result of an expression in another transformation

Connected Lookup	Unconnected Lookup
Receives input values directly from the pipeline.	Receives input values from the result of a :LKP expression in another transformation
You can use a dynamic or static cache.	You can use a static cache.
Cache includes all lookup columns used in the mapping (that is, lookup	Cache includes all lookup/output ports in

table columns included in the lookup condition and lookup table columns linked as output ports to other transformations).	the lookup condition and the lookup/return port.
Can return multiple columns from the same row or insert into the dynamic lookup cache.	The dynamic lookup cache. Designate one return port (R). Returns one column from each row
If there is no match for the lookup condition, the Informatica Server returns the default value for all output ports. If you configure dynamic caching, the Informatica Server inserts rows into the cache.	If there is no match for the lookup condition, the Informatica Server returns NULL.
Pass multiple output values to another transformation. Link lookup/output ports to another transformation.	Pass one output value to another transformation. The lookup/output/return port passes the value to the transformation calling: LKP expression.
Supports user-defined default values.	Does not support user-defined default values

What is complex mapping and explain?
We can say a mapping involving many lookups; Joiners and complex calculation may be called as a complex mapping

If you can't find what you are looking for in lookup table, how do you handle?

If u has data coming from diff. sources what transformation will u use in your designer?
Joiner Transformation

How many transformations are there?
15 types of transformations

Can u stop a session in a concurrent batch?
Nope

What happens when u uses the delete or update or reject or insert statement in Ur update strategy?

How many sessions u can run?

The sessions those are configured in server. Max you can is 50-100(not sure)

What are parameter variables and where do u use it?

Mapping Parameters

A mapping parameter represents a constant value that you can define before running a session. A mapping parameter retains the same value throughout the entire session.
When you use a mapping parameter, you declare and use the parameter in a mapping or mapplet. Then define the value of the parameter in a parameter file for the session. During the session, the Informatica Server evaluates all references to the parameter to that value.

Mapping Variables

Unlike a mapping parameter, a mapping variable represents a value that can change through the session. The Informatica Server saves the value of a mapping variable to the repository at the end of each successful session run and uses that value the next time you run the session.

Local variable: You can use local that you create within a mapping in any transformation expression. For example, if you use a complex tax calculation throughout a mapping, you might want to write the expression once and designate it as a. you thereby increase performance since the Informatica Server performs the calculation only once. Local are especially useful when used with stored procedure expressions to capture multiple return values.

System variable $$$SessStartTime these are constant through out the mapping and cannot be changed.

What are counters?

The performance details provide that help you understand the session and mapping efficiency. Each Source Qualifier, target definition, and individual transformation appears in the performance details, along with that display performance information about each transformation

Understanding Performance

All transformations have some basic that indicates the number of input rows, output rows, and error rows. Source Qualifiers, Normalizes, and targets have additional that indicates the efficiency of data moving into and out of buffers. You can use these to locate performance bottlenecks. Some transformations have specific to their functionality. For example, each Lookup transformation has an indicates the number of rows stored in the lookup cache. When you read performance details, the first column displays the transformation name as it appears in the mapping, the second column contains the name, and the third column holds the resulting number or efficiency percentage. When you

partition a source, the Informatica Server generates one set of for each partition. The following performance illustrate two partitions for an Expression transformation:

Transformation		Value
EXPTRANS [1]	Expression_input rows	8
	Expression_output rows	8
EXPTRANS [2]	Expression_input rows	16
	Expression_output rows	16

Note: When you partition a session, the number of aggregate or rank input rows may be different from the number of output rows from the previous transformation.

How do you set the size of block buffer?

How do you pass parameters to unconn. Stored proc?

Configure the expression to send any input parameters and capture any output parameters or return value You must know whether the parameters shown in the Expression Editor are input or output parameters. You insert variables or port names between the parentheses in the exact order that they appear in the stored procedure itself. The datatypes of the ports and variables must match those of the parameters passed to the stored procedure.

For example, when you click the stored procedure, something similar to the following appears:
:SP.GET_NAME_FROM_ID()

This particular stored procedure requires an integer value as an input parameter and returns a string value as an output parameter. How the output parameter or return value is captured depends on the number of output parameters and whether the return value needs to be captured.

If the stored procedure returns a single output parameter or a return value (but not both), you should use the reserved variable PROC_RESULT as the output variable. In the previous example, the expression would appear as:
:SP.GET_NAME_FROM_ID(inID, PROC_RESULT)

InID can be either an input port for the transformation or a variable in the transformation. The value of PROC_RESULT is applied to the output port for the expression.

If the stored procedure returns multiple output parameters, you must create variables for each output parameter. For example, if you created a port called varOUTPUT2 for the stored procedure expression, and a variable called varOUTPUT1, the expression would appears as:
:SP.GET_NAME_FROM_ID (inID, varOUTPUT1, PROC_RESULT)
The value of the second output port is applied to the output port for the expression, and the value of the first output port is applied to varOUTPUT1. The output parameters are returned in the order they are declared in the stored procedure itself.
With all these expressions, the datatypes for the ports and variables must match the datatypes for the input/output variables and return value.

How do u implement unconn. Stored proc. In a mapping?

Can u access a repository created in previous version of Informatica?
No

What happens if the info. Server doesn't find the session parameter in the parameter file?

How does run session from command line?
PMCMD command
What is diff. Things u can do using PMCMD?
Start, Stop and abort the session

What are different ports in Informatica?
    Input, Output, Variable, Return/Rank, Lookup and Master.

What is a Variable port? Why it is used?
    Variable port is used to store intermediate results.
    Variable ports can reference input ports and variable ports, but not output ports.

What are the different connectivity Informatica uses to connect to sources, targets and the repository?
    Power mart and Power Center uses
            Network Protocol
            Native Drivers
            ODBC
    The Server Manager and the Informatica Server use TCP/IP or IPX/SPX to communicate to each other.

Difference between Data cleansing and Data scrubbing?
    Data cleansing is a process of removing errors and resolving inconsistencies in source data before loading data into targets.
    Data scrubbing is a process of filtering, merging, decoding and translating the source data to create the validation data for data warehouse.

What are reusable transformations? Can we rollback this option?
Reusable transformations can be used in multiple transformations. Only one transformation can be used in reusable transformation. **You can rollback this option.**

Diff between Active and passive transformation?

Transf can be active or passive , activ tranf can change the no of rec passed th it, a passive tranf can never change the rec cnt**, Active trsnf that might change the rec cnt are** advan ext proc, aggrega, filter,joiner,normalizer, rank , updte strategy, source qualifier**, if u use powerconnect to access erp sources, erp source quail is also an active tranfor**
Passive tranf :- lookup, expression, external procedure, seq generator, stored procedure
U can connect only 1 active tranf to the same tranf or target can connect any no of pass tranf

What are Mapplet?
A mapplet is a reusable object that represents a set of transformations. It allows you to reuse transformation logic and can contain as much transformation as you need.

What are mapping parameters and variables?
 A mapping parameter represents a constants value that you can define before running a session. A mapping parameter retains the same value throughout the entire session.
Unlike a mapping parameter, a mapping variable represents a value that can change through the session.

How many transformations have you used?
10- 12

What is Aggregate transformation?
 An aggregator transformation allows you to perform aggregate calculations, such as average and sums. The Aggregator transformation is unlike the Expression transformation, in that you can use the Aggregator transformation to perform calculations on groups.
What is Router Transformation? How is it different from Filter transformation?
 A Router transformation is similar to a Filter transformation because both transformations allow you to use a condition to test data. A Filter transformation tests data for one condition and drops the rows of data that do not meet the condition. However, a router transformation tests data for one or more conditions and gives you the option to route rows of data that do not meet any of the conditions to default output group.
What are connected and unconnected transformations?
Connected transformations are the transformation, which are in the data flow, whereas unconnected transformation will not be in the data flow. These are dealt in Lookup and Stored procedure transformations.
What is Normalizer transformation?

Normalizer transformation normalizes records from COBOL and relational sources allowing you to organize the data according to your needs. A normalizer transformation can appear anywhere in a data flow when you normalize a relational source.

How to use a sequence created in Oracle in Informatica?

By using stored procedure transformation.

What are source qualifier transformations?

The source qualifier represents the records that the Informatica Server reads when it runs a session.

What is a dimension?

A set of level properties that describe a specific aspect of a business, used for analyzing the factual measures of one or more cubes, which use that dimension. Ex: geography, time, customer and product.

What is a XML source qualifier?

The XML source qualifier represents the data elements that the Informatica server reads when it runs a session with XML sources.

What is a cube?

A set of related factual measures, aggregates, and dimensions for a specific dimensional analysis problem. Ex: regional product sales.

What is Load Manager Process?

Load manager is the primary Informatica server process. It performs the following tasks:
a. Manages sessions and batch scheduling.
b. Locks the sessions and reads properties.
c. Reads parameter files.
d. Expands the server and session variables and parameters.
e. Verifies permissions and privileges.
f. Validates sources and targets code pages.
g. Creates session log files.
h. Creates Data Transformation Manager (DTM) process, which executes the session.

Define a session?

A session is a set of instructions that describes how and when to move data from source to targets

What are pmcmd commands?

pmcmd is a command line program to communicate with the Informatica server. This does not replace the server manager, since there are many tasks that you can perform only with server Manager.

What are cache and their types in Informatica?

The Informatica server creates **index** and **data** cache for aggregator, Rank, joiner and Lookup transformations in a mapping. The Informatica server stores key values in the index cache and output values in the data cache.

What are minimum and maximum values for index and data cache?

Index cache: Min: 12MB   Max: 24 MB

Data cache: Min: 12MB   Max: 24MB

What is default block buffer size?
   64K
What is default LM shared memory size?
   2MB
What is an incremental aggregation?

   In Incremental aggregation, you apply captured changes in the source to aggregate
   calculations in a session. If the source changes only incrementally and you can capture
   changes, you can configure the session to process only those changes. This allows the
   Informatica server to update your target incrementally, rather than forcing it to process the
   entire source and recalculate the same calculation each time you run the session.

What is Reject loading?
 During a session, the Informatica server creates a reject file for each target instance in the mapping. If
the writer or the target rejects data, the Informatica server writes the rejected row into reject file.
   The reject file and session log contain information that helps you determine the cause of the
   reject. You can correct reject files and load them to relational targets using the Informatica
   reject load utility. The reject loader also creates another reject file for the data that the writer or
   target reject during the reject loading.

How many servers' instances can you register to a repository?

   What was the size of your warehouse?
   What is Star Schema?
   What is the difference between inner and outer join
   How big was your fact table?
   How did you handle performance issues?
   If you have data coming in from multiple sources, just walk thru the process of loading it into
   the target?
   How will u convert rows into columns or columns into rows?
   Do you have any experience in OOP
   What is the syntax difference for using outer join in SQL Server / Oracle
   What is a repository?
The place where u store the metadata is called a repository

What is the difference between PowerCenter and PowerMart ?

Power Center and Power mart call the instructions to get data,change it ,where to write the
information Data Mart

PowerCenter	PowerMart
Can create Local Repository which could be upgraded to a Global Repository.	Feature is not available

Can Connect to Enterprise Data (SAP/R3, SIEBEL, VANTIVE, IBM MQ Series……	Feature is not available
Ability to register multiple servers, share metadata across repositories, and partition data	All features except distributed metadata, multiple registered servers, and data partitioning

What are the max numbers of sessions that you could configure in Informatica Server?

You can set three parameters in the Informatica Server configuration that control how the Load Manager allocates memory to sessions:

- **MaxSessions.** The maximum session's parameter indicates the maximum number of session slots available to the Load Manager at one time for running or repeating sessions. For example, if you select the default MaxSessions of 10, the Load Manager allocates 10 session slots. This parameter helps you control the number of sessions the Informatica Server can run simultaneously.
- **LMSharedMemory.** The Load Manager shared memory parameter is set in conjunction with the maximum sessions parameter to ensure that the Load Manager has enough memory for each session. The Load Manager requires approximately 200,000 bytes of shared memory for each session slot. The default setting is 2,000,000 bytes. For each increase of 10 sessions in the MaxSessions setting, you need to increase LMSharedMemory by 2,000,000 bytes.
- **KeepRepeatSessionInShm.** The Keep Repeating Sessions in Shared Memory option determines how the Load Manager behaves when there are no available slots for sessions to run. By default, this option is disabled, and the Load Manager swaps repeating sessions out of session slots to run new session requests. If all session slots are filled with running sessions, the Load Manager places new session requests in a waiting state until it can open a session slot. If you enable this option, the Load Manager retains repeating sessions in shared memory, and fails to run new session requests.

What is unconnected transformation?   What is the advance of using Unconnected transformation?

- Unconnected transformation is not part of the data flow
- The Informatica Server queries the Lookup table based on the logic of the expression calling the lookup.
- Unconnected Lookup Transformation to check the record whether it exists in the target database or not.

What are the steps involved in the migration from older version to newer version of Informatica Server?

Have you worked in conjunction with Oracle DBA(s) and UNIX administrators?

Have you worked with Oracle front-end tools?

What are the main features of Oracle 8i with context to datawarehouse?

Are you comfortable with Business Objects?

What kind of analysis you did in the existing Data warehouse environment?

What problem you faced and what were your recommendations (based on your last project)?

How many dimension tables and fact tables were there in your datawarehouse?

If you can't find what you are looking for in lookup table, how do you handle?

What is Star schema / Snow Flake schema?

What are Target Types on the Server?
Target Types are File, Relational and ERP
What are Target Options on the Servers?
Target Options for File Target type are FTP File, Loader and MQ
There are no target options for ERP target type
Target Options for Relational are Insert, Update (as Update), Update (as Insert), Update (else Insert), Delete,and   TruncateTable.
How do you identify existing rows of data in the target table using lookup transformation?
Can identify existing rows of data using **unconnected** lookup transformation.

What are Aggregate transformation?
Aggregator transformation allows you to perform aggregate calculations, such as averages and sums.

What are various types of Aggregation?
Various types of aggregation are SUM, AVG, COUNT, MAX, MIN, FIRST, LAST, MEDIAN, PERCENTILE, STDDEV, and VARIANCE.

What are Dimensions and various types of Dimensions?
A set of level properties that describe a specific aspect of a business, used for analyzing the factual measures of one or more cubes which use that dimension.

Egs. Geography, time, customer and product.

What are 2 modes of data movement in Informatica Server?
The data movement mode depends on whether Informatica Server should process single byte or multi-byte character data. This mode selection can affect the enforcement of code page relationships and code page validation in the Informatica Client and Server.

a) Unicode – IS allows 2 bytes for each character and uses additional byte for each non-ascii character (such as Japanese characters)
b) ASCII – IS holds all data in a single byte

The IS data movement mode can be changed in the Informatica Server configuration parameters. This comes into effect once you restart the Informatica Server.

What is Code Page Compatibility?

Compatibility between code pages is used for accurate data movement when the Informatica Sever runs in the Unicode data movement mode. If the code pages are identical, then there will not be any data loss. One code page can be a subset or superset of another. For accurate data movement, the target code page must be a superset of the source code page.

Superset -A code page is a superset of another code page when it contains the character encoded in the other code page. It also contains additional characters not contained in the other code page.

Subset - A code page is a subset of another code page when all characters in the code page are encoded in the other code page.

What is Code Page used for?

Code Page is used to identify characters that might be in different languages. If you are importing Japanese data into mapping, then u must select the Japanese code page for the source data.

What is Router transformation?

Router transformation allows you to use a condition to test data. It is similar to filter transformation. It allows the testing to be done on one or more conditions.

What is Load Manager?

The load Manager is the Primary informatica Server Process. It performs the following tasks -
- Manages session and batch scheduling.
- Locks the session and read session properties.
- Reads the parameter file.
- Expand the server and session variables and parameters.
- Verify permissions and privileges.
- Validate source and target code pages.
- Create the session log file.
- Create the Data Transformation Manager which executes the session.

What is Data Transformation Manager?

After the load manager performs validations for the session, it creates the DTM process. The DTM process is the second process associated with the session run. The primary purpose of the DTM process is to create and manage threads that carry out the session tasks.

The DTM allocates process memory for the session and divide it into buffers. This is also known as buffer memory. It creates the main thread, which is called the master thread. The master thread creates and manages all other threads.

If we partition a session, the DTM creates a set of threads for each partition to allow concurrent processing. When Informatica server writes messages to the session log it includes thread type and thread ID. Following are the types of threads that DTM creates:

- **MASTER THREAD** - Main thread of the DTM process. Creates and manages all other threads.
- **MAPPING THREAD** - One Thread to Each Session. Fetches Session and Mapping Information.
- **Pre and Post Session Thread** - One Thread Each To Perform Pre And Post Session Operations.
- **READER THREAD** - One Thread for Each Partition for Each Source Pipeline.
- **WRITER THREAD** - One Thread for Each Partition If Target Exist In The Source pipeline Write To The Target.
- **TRANSFORMATION THREAD** - One or More Transformation Thread For Each Partition.

WHAT IS SESSION AND BATCHES?

SESSION -A Session Is a set of instructions that tells the Informatica Server How and when to move data from sources to targets. After creating the session, we can use either the server manager or the command line program pmcmd to start or stop the session.

BATCHES - It provides a way to Group Sessions For Either Serial Or Parallel Execution By The Informatica Server. There Are Two Types Of Batches :
1. **SEQUENTIAL** - Run Session One after the Other.
2. **CONCURRENT** - Run Session At The Same Time.

What is a source qualifier?
It represents all data queried from the source.

Why we use lookup transformations?
Lookup Transformations can access data from relational tables that are not sources in mapping. With Lookup transformation, we can accomplish the following tasks:
a) **Get a related value** - Get the Employee Name from the Employee table based on the Employee ID
b) **Perform Calculation**
c) **Update slowly changing dimension tables -** We can use **unconnected lookup transformation** to determine whether the records already exist in the target or not.

1. Difference between Power Center and Mart?
2. What are the new features of Power Center 5.0?
3. How to run a session, which contains mapplet?

4. Differentiate between Load Manager and DTM?

Business Objects:
1. What is Universe ?
2. Define Loops and how to resolve them ?
3. What were the errors which you encountered most while running BO reports ?
4. How to schedule BO Reports ?
5. What are the methods available in Document class of the Web-I SDK ?
6. What is the objective behind Partitioning in Oracle ?

CODE PAGE OVERVIEW

A code page contains the encoding to specify characters in a set of one or more languages. An encoding is the assignment of a number to a character in the character set. You use code pages to identify data that might be in different languages. For example, if you are importing Japanese data into a mapping, you must select a Japanese code page for the source data.

To change the language to English and require the system to use the Latin1 code page, in UNIX, execute the following command. **setenv LANG en_US.iso88591**

If You Are Using PowerCenter

With PowerCenter, you receive all product functionality, including the ability to register multiple servers, share metadata across repositories, and partition data.
A PowerCenter license lets you create a single repository that you can configure as a global repository, the core component of a data warehouse.
When this guide mentions a PowerCenter Server, it is referring to an Informatica Server with a PowerCenter license.

If You Are Using PowerMart

This version of PowerMart includes all features except distributed metadata, multiple registered servers, and data partitioning. Also, the various options available with PowerCenter (such as PowerCenter Integration Server for BW, PowerConnect for IBM DB2, PowerConnect for IBM MQSeries, PowerConnect for SAP R/3, PowerConnect for Siebel, and PowerConnect for PeopleSoft) are not available with PowerMart.
When this guide mentions a PowerMart Server, it is referring to an Informatica Server with a PowerMart license.

Normalizer Transformation:

Normalization protects the data and makes the database more flexible by eliminating redundancy and inconsistent dependencies.

Used mainly with COBOL sources.
With relational sources, used to create multiple rows from a single row of data.

Use a single Normalizer T to handle multiple levels of denormalization in the same record. For example, a single record might contain two different detail record sets. Rather than using two Normalizer T to handle the two different detail record sets, you handle both normalizations in the same transformation.

Pivoting can be done. Changing columns to rows and vice versa.

A Normalizer column id is created for OCCURS, Pivoting used to identify the Columns.

When we are importing COBOL and Normalizer T automatically created and **we can't modify the field definitions in the Normalizer T**. We have to modify only in the Source Analyzer.

Setting	Description
Reset	If selected, the Informatica Server resets the generated key value after the session finishes to its original value.
Restart	If selected, the Informatica Server restarts the generated key values from 1 every time the session runs.
Tracing level	Determines the amount of information about this transformation that the Informatica Server writes to the session log when it runs the session. You can override this tracing level when you configure a session.

Can I use my existing PowerMart 4.0/PowerCenter 1.0 mappings that contain COBOL sources?

In PowerMart 4.0 and PowerCenter 1.0, the Designer did not support REDEFINE statements in COBOL sources or copybooks. If your COBOL sources included REDEFINE statements in older mappings and you implemented workarounds, you do not need to edit your existing mappings or sessions. However, if you want to improve session performance, you can either create a new mapping or edit the existing mapping by reimporting the COBOL sources and modifying any filter transformations or target as needed. You can then create one session to load the data from the COBOL source.

Cannot I edit the ports in my Normalizer transformation when using a relational source?

When you create ports manually, you must do so on the Normalizer tab in the transformation, not the Ports tab.

<u>Source Qualifier Transformation:</u>
When you add a relational or a flat file source definition to a mapping, you need to connect it to a Source Qualifier transformation. **The Source Qualifier represents the records that the Informatica Server reads when it runs a session**.

- To join data originating from the same DB.
- Filter records in the Source itself.
- To specify an outer join instead of a default inner join.
- To specify sorter ports.
- To select distinct values from the source.
- To create a custom query to issue a special select statement for the Informatica server to read source data. For example, we might use a custom query to perform aggregate calculations or execute a stored procedure.

Don't alter the data types in the Source Qualifier.
If we have connected multiple SQ to multiple targets, we can designate the order in which the targets are loaded.
If we connect one SQ to multiple targets, we can enable constraint-based loading in a session to have the IS load data based on target table primary and foreign key relationships.
The IS reads only the columns in SQ that are connected to other transformations.
Informatica Server

Informatica server moves data from source to target based on mapping and session metadata stored in a repository. A mapping is a set of source and target definitions linked by transformation objects that define the rules for data transformation. A session is a set of instructions that describes how and when to move data from source to targets.

Session Process

The Informatica server uses both process memory and system shared memory to perform these tasks. It runs as a daemon on UNIX and as a service on Windows NT/2000.

The load manager process. Start the session, creates the DTM process, and sends post-session email when the session completed.

The DTM process (Data Transformation Manager). Creates threads to initialize the session, read, write, and transform data, and handle pre- and post-session operations.

Partitioning Data

The Informatica server can achieve high performance by partitioning source data and performing the extract, transformation, and load for each partition in parallel.

Configure the session to partition source data.
Install the Informatica server on a machine with multiple CPUs.

When you partition a session, configure partitioning based on source qualifiers in a mapping.

For relational source, the Informatica server creates multiple database connections to a single source and extracts a separate range of data for each connection. For XML or file sources, the Informatica server reads multiple files concurrently.

When the Informatica Server loads relational data, it loads relational data; it creates multiple database connections to the target and loads partitions of data concurrently. When the Informatica server loads data to file targets, it creates a separate file for each partition.

Update strategy

When you design the data warehouse, you need to decide what type of information to store in targets. As part of your target table design, you need to determine whether to maintain all the historic data or just the most recent changes.

In PowerMart and Power Center, you set your update strategy at two different levels:

Within a session. When you configure a session, you can instruct the server to either treat all records in the same way (for example, treat all records as inserts), or use instructions coded into the session mapping to flag records for different database operations.

Within a mapping. Within a mapping, you use the Update Strategy transformation to flag records for insert, delete, update, or reject.

Follow these steps to define an update strategy:

To control how records are flagged for insert, update, delete, or reject within a mapping, add an Update Strategy transformation to the mapping. Update Strategy transformations are essential if you want to flag records destined for the same target for different database operations, or if you want to reject records.

Define how to flag records when you configure a session. You can flag all records for insert, delete, or update, or select the Data Driven option, where the Informatica Server follows instructions code into Update Strategy transformations within the session mapping.

Define insert, update, and delete options for each target when you configure a session. On a target-by-target basis, you can allow or disallow inserts and deletes, and you can choose three different ways to handle updates.

When you configure a session, you have several options for handling specific database operations, including updates.

Specifying an operation for all rows

Setting	Description
Insert	Treat all records as inserts. If inserting the record violates a primary or foreign key constraint in database, the server reject the records
Delete	Treat all records as deletes. For each record, if the server finds a corresponding record in the target table, the server deletes it.
Update	Treat all records as updates. For each existing records, server updates the record.
Data Driven	The informatica server follows instructions code into update strategy transformations within the session mapping to determine how to flag records for insert, delete, update or reject.

Specifying operations for individual Target Tables.

Insert. Select this option to insert a row into a target table.
Delete. Select this option to delete a record from a table.
Update. You have three different options.
    Update as update. Update each record flagged for update if exists in the target table.
    Update as insert. Insert each record flagged for update.
    Update else insert. Update the record if it exist. Otherwise, insert it.
Truncate Table. Select this option to truncated the target table before loading data.

Flagging Records within a Mapping.

Following are database operation

Operation	Constant	Numeric Value
Insert	DD_INSERT	0
Update	DD_UPDATE	1
Delete	DD_DELETE	2
Reject	DD_REJECT	3

The server treats any other value as an insert.

Update strategy transformation is frequently the first transformation in a mapping, before data reaches a target table. You can use the Update strategy transformation to determine how to flag that record. Later, when you configure a session based on this transformation, you can determine what to do with records flagged for insert, delete, or update.

Forwarding Rejected Rows

You can configure the Update strategy transformation to either pass rejected rows to the next transformation or drop them. By default, the informatica server forwards rejected rows to the next transformation.

Update Strategy Expression.

Frequently, the update strategy expression uses the IIF or DECODES function from the transformation language to test each record to see if it meets a particular condition. On the condition you can assign a numeric code to flag it for a particular database operation.

Aggregator and update strategy transformation

Position the Aggregator before the Update Strategy Transformation. In this case, you perform the aggregate calculation, and then used the update strategy transformation to flag records that contain the results of this calculation for insert, delete, or update.

Position the aggregator after the update strategy transformation. Here, you flag records for insert, delete, update, or reject before you perform the aggregate calculation. How you flag a particular record determines how the aggregator transformation treats any values in that record used in the calculation.

This behavior has changed since PowerMart 3.5. In version 3.5, the informatica server performed all aggregate calculations before it flagged records for insert, update, delete, or reject through a data driven expression

What are Target Types on the Server?
 Target Types are File, Relational and ERP

What are Target Options on the Servers?

 Target Options for File Target type are FTP File, Loader and MQ
There are no target options for ERP target type
Target Options for Relational are Insert, Update (as Update), Update (as Insert), Update (else Insert), Delete,   and   Truncate   Table.

How do you identify existing rows of data in the target table using lookup transformation?
Can identify existing rows of data using **unconnected** lookup transformation.

What are Aggregate transformation?
Aggregator transformation allows you to perform aggregate calculations, such as averages and sums.

What are various types of Aggregation?
Various types of aggregation are SUM, AVG, COUNT, MAX, MIN, FIRST, LAST, MEDIAN, PERCENTILE, STDDEV, and VARIANCE.

What are Dimensions and various types of Dimensions?
A set of level properties that describe a specific aspect of a business, used for analyzing the factual measures of one or more cubes which use that dimension. Egs. Geography, time, customer and product.

What are 2 modes of data movement in Informatica Server?
The data movement mode depends on whether Informatica Server should process single byte or multi-byte character data. This mode selection can affect the enforcement of code page relationships and code page validation in the Informatica Client and Server.

  c) Unicode – IS allows 2 bytes for each character and uses additional byte for each non-ascii character (such as Japanese characters)
  d) ASCII – IS holds all data in a single byte

The IS data movement mode can be changed in the Informatica Server configuration parameters. This comes into effect once you restart the Informatica Server.

What is Code Page Compatibility?
Compatibility between code pages is used for accurate data movement when the Informatica Sever runs in the Unicode data movement mode. If the code pages are identical, then there will not be any data loss. One code page can be a subset or superset of another. For accurate data movement, the target code page must be a superset of the source code page.

Superset - *A code page is a superset of another code page when it contains the character encoded in the other code page. It also contains additional characters not contained in the other code page.*

Subset - A code page is a subset of another code page when all characters in the code page are encoded in the other code page.

What is Code Page used for?
Code Page is used to identify characters that might be in different languages. If you are importing Japanese data into mapping, then u must select the Japanese code page for the source data.

What is Router transformation?

Router transformation allows you to use a condition to test data. It is similar to filter transformation. It allows the testing to be done on one or more conditions.

What is Load Manager ?

Ans. The load Manager is the Primary informatica Server Process. It Performs the following tasks -
- Manages session and batch scheduling.
- Locks the session and read session properties.
- Reads the parameter file.
- Expand the server and session variables and parameters.
- Verify permissions and privileges.
- Validate source and target code pages.
- Create the session log file.
- Create the Data Transformation Manager which execute the session.

12. What is Data Transformation Manager?

Ans. After the load manager performs validations for the session, it creates the DTM process. The DTM process is the second process associated with the session run. The primary purpose of the DTM process is to create and manage threads that carry out the session tasks.

The DTM allocates process memory for the session and divide it into buffers. This is also known as buffer memory. It creates the main thread, which is called the master thread. The master thread creates and manages all other threads.

If we partition a session, the DTM creates a set of threads for each partition to allow concurrent processing.. When Informatica server writes messages to the session log it includes thread type and thread ID. Following are the types of threads that DTM creates:

- **MASTER THREAD** - Main thread of the DTM process. Creates and manages all other threads.
- **MAPPING THREAD** - One Thread to Each Session. Fetches Session and Mapping Information.
- **Pre And Post Session Thread** - One Thread Each To Perform Pre And Post Session Operations.
- **READER THREAD** - One Thread for Each Partition for Each Source Pipeline.
- **WRITER THREAD** - One Thread for Each Partition If Target Exist In The Source pipeline Write To The Target.
- **TRANSFORMATION THREAD** - One or More Transformation Thread For Each Partition.

13. WHAT IS SESSION AND BATCHES?

Ans. **SESSION** - A Session Is A set of instructions that tells the Informatica Server How And When To Move Data From Sources To Targets. After creating the session, we can use either the server manager or the command line program pmcmd to start or stop the session.

BATCHES - It Provides A Way to Group Sessions For Either Serial Or Parallel Execution By The Informatica Server. There Are Two Types Of Batches :

3. **SEQUENTIAL** - Run Session One after the Other.

4. **CONCURRENT** - Run Session At The Same Time.

14. What is a source qualifier?
 It represents all data queried from the source.

15. Why we use lookup transformations?
 Lookup Transformations can access data from relational tables that are not sources in mapping. With Lookup transformation, we can accomplish the following tasks:
d) **Get a related value** - Get the Employee Name from the Employee table based on the Employee ID
e) **Perform Calculation**
f) **Update slowly changing dimension tables -** We can use **unconnected lookup transformation** to determine whether the records already exist in the target or not.

Contrasting a Data Warehouse with an OLTP System
Figure 1–1 illustrates some of the key differences between a data warehouse's model and an OLTP system's.

Figure 1–1 Contrasting OLTP and Data Warehousing Environments

OLTP		Data Warehouse
Complex data structures (3NF databases)		Multidimensional data structures
Few	Indexes	Many
Many	Joins	Some
Normalized DBMS	Duplicated Data	Denormalized DBMS
Rare	Derived data and Aggregates	Common

One major difference between the types of system is that data warehouses are not usually in third-normal form.

Data warehouses and OLTP systems have vastly different requirements. Here are some examples of the notable differences between typical data warehouses and OLTP systems:

Workload

>   Data warehouses are designed to accommodate ad hoc queries. The workload of a data warehouse may not be completely understood in advance, and the data warehouse is optimized to perform well for a wide variety of possible query operations.
>   OLTP systems support only predefined operations. The application may be specifically tuned or designed to support only these operations.

Data Modifications

>   The data in a data warehouse is updated on a regular basis by the ETT process (often, every night or every week) using bulk data-modification techniques. The end users of a data warehouse do not directly update the data warehouse. In an OLTP system, end users routinely issue individual data-modification statements in the database. The OLTP database is always up-to-date, and reflects the current state of each business transaction.

Schema Design

>   Data warehouses often use denormalized or partially denormalized schemas (such as a star schema) to optimize query performance. OLTP systems often use fully normalized schemas to optimize update/insert/delete performance, and guarantee data consistency.

Typical Operations

>   A typical data warehouse query may scan thousands or millions of rows. For example, "Find the total sales for all customers last month." A typical OLTP operation may access only a handful of records. For example, "Retrieve the current order for a given customer."

Historical Data

>   Data warehouses usually store many months or years of historical data. This is to support historical analysis of business data. OLTP systems usually store only a few weeks' or months' worth of data. The OLTP system only stores as much historical data as is necessary to successfully meet the current transactional requirements.

Typical Data Warehouse Architectures

As you might expect, data warehouses and their architectures can vary depending upon the specifics of each organization's situation. Figure 1–2 shows the most basic architecture for a data warehouse. In it, a data warehouse is fed from one or more source systems, and end users directly access the data warehouse.

Figure 1–2 Typical Architecture for a Data Warehouse

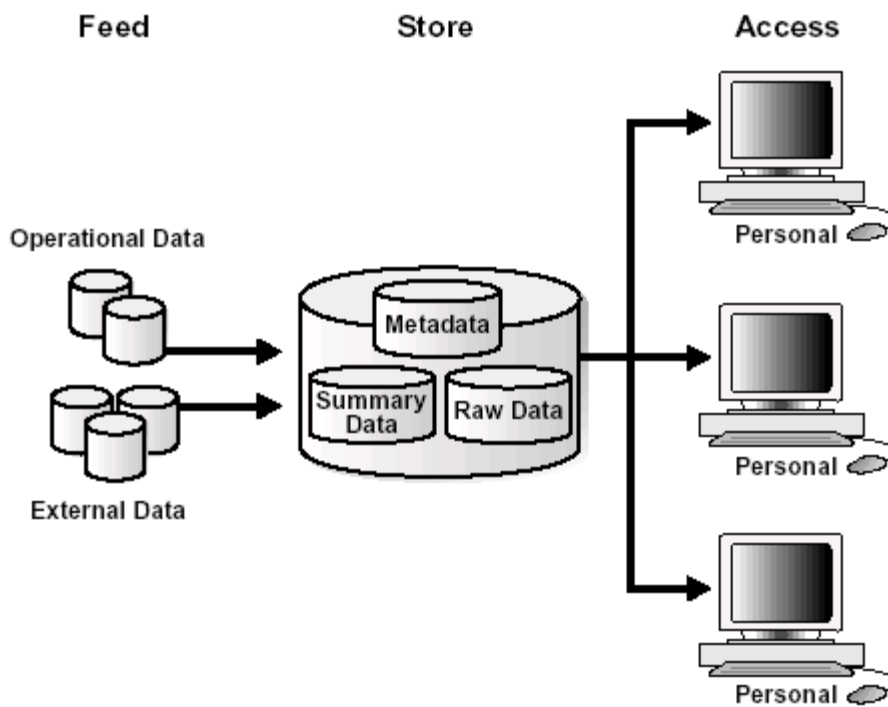Figure 1–2 Typical Architecture for a Data Warehouse

Figure 1–3 illustrates a more complex data warehouse environment. In addition to a central database, there is a staging system used to cleanse and integrate data, as well as multiple data marts, which are systems designed for a particular line of business.

Figure 1–3 Typical Architecture for a Complex Data Warehouse
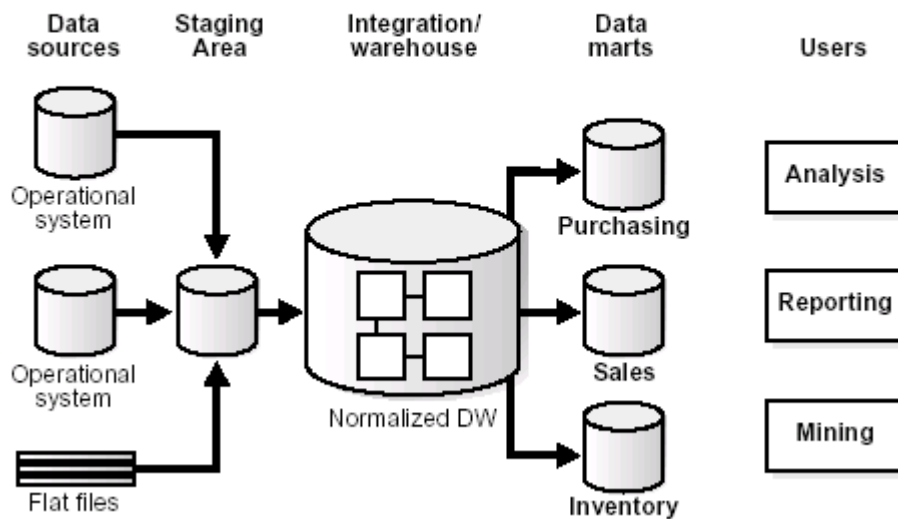
What are confirmed dimensions
Confirmed dimensions are linked to multiple fact tables
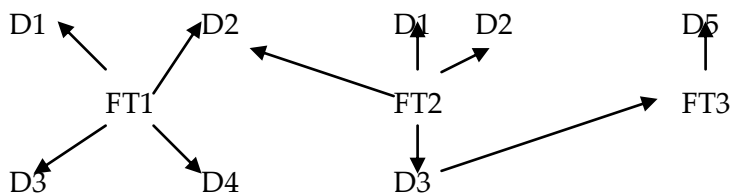
Data modeling tool :- Erwin by computer associates

Reverse engineering :- Modifying existing tables
Forwarding :- creating tables (u create the sql scripts and execute)

Explain the Difference between OLAP, MOLAP,ROLAP ,HOLAP
   OLTP (online transaction processing system) stores day to day transactions
DW stores historical Data. OLTP data is in normalized form, multiple tables joins where u run a query, complex joins takes a lot of time
DW is a denormalized one or more fact tables each connected to multiple dimension tables (Star Schema)
Snow Flake Schema:- When u have multiple lookup tables, u can go for snowflake schema (Partial Normalized) when u have large volume of data in dimension tables
Ex:- When u have a country table to be linked to both customers, resot u can create a lkp country table so when u make any changes to ctry table u don't need to change in multiple dimension tables

ROLAP:- Star, Snow Flake Schema
MOLAP:-Multidimensional data is stored in hirarechical way (Customer-> Region->Country->Continent) Data is stored in the form of cubes . MOLAP is very expensive and complex design
When u need results very very quickly response time is more in MOLAP
HOLAP also called as Hybrid OLAP combination of both ROLAP and MOLAP
What is ZABO(Zero based administration)

What is AIS (APPlication Integration Software)

Name Some Cleansing Tools
Informatica can be used as a cleansing tool, Ab initio, Data stage

What is Slicing and dicing

Explain linking universes

To link the universes u must have exported to where the kernel universe at lease once otherwise the designer does not allow the link

Data Warehousing

What is Data warehousing
A DW is a DB used for query, analysis and reporting . By definition DW is a subject oriented, intergrated, non volatile and time variant
Subject Oriented:- Represents a subject areas like sales, Mktg
Integrated: Data Colleted from multiple source systems integrated into a user readable unique format
Ex:- male, female ,0,1, M,F, T, F
Non Volatile :- Dw stores historical data
Time Variant :- Stores data timewise like weekly,monthly,quarterly, yearly

Oracle
Query Optimization (Performance tuning SQL)
Oracle currently (as of Oracle9i Release 2) provides 2 different optimization approaches. These are Cost Based Optimization (CBO) and Rule Based Optimization (RBO). The CBO bases optimization choices on pre-gathered table and index statistics while the RBO makes it's decisions based on a set of 'best practice' rules and does not rely on any statistical information. CBO's reliance on statistics makes it vastly more flexible than the RBO since as long as up to date statistics are maintained it will accurately reflect real data volumes. The RBO is Oracle's legacy optimizer and is to be disported in Oracle10i.

What is dynamic caching
One Important feature of Informatica 6.2 is dynamic lookup cache
dynamic means.......for example if you are using lookup transformation and lookup cache for a particular table so this lookup cache will read the table at a time and will keep in the memory in the form of lookup cache, but this is generally under assumption you the lookup table is not getting changed during that time

let's if the lookup table values getting changed during that time this lookup cache will not hold most recent values While creating aggregator transformation have u enabled sorted input option what exactly it does Explain the process so instead of standard lookup cache you use dynamic cache in that case

What is the size of your repository ?
10 GB
How many source tables u have
30 to 40 (Target database 60 GB)
How many mapping u have done
20

How many sessions
Batches :- 4-5 nested
Sessions :- 20-25
Individual :- depencies I put in sequential (ex:- first one session, next session ,next a batch)

How do u sechedule sessions
On Demand (Manual)
Time based
Event based

How do u handle session failure
I run from recovery mode, PMCMD run (if error occurs session stops, wherever it is stopped it continues)

Dashboards:- GE From one prj how much revenue generated , how well u did a process (Graphical representation) Bo Reporting (Use graphs)

What does sorted input option in aggregator transformation does
Improves session performance but the data has to be passed into aggregator transformation must be sorted before (if use sorted input option and data is not sorted before session fails)

When u pass presorted data ex:- 101 if
4 records whenever it finds new no 102 it performs the calculation

When the sorted input is not selected the Informatica server performs aggregate calculations as it reads the Informatica server stores data for each group until it reads the entire source to ensure all aggregate calculations are accurate
When sorted input option is selected Informatica assumes all data is sorted by group as it reads rows for a group performs aggregate calculations as it reads

How did u do increment load

What is Click Stream In Data Warehousing

Click stream is basically web based data warehousing analysis basically e-web intelligence

When to index a particular column on what percentage
Generally if you are creating ordinary b-tree index.........if you know that that columns will be mostly reference in select statement
 if your select statement is retrieving more than 30 % rows then...it's better not to use index
index will be used out of the table when you are retrieving smaller no. of rows
if you retrieving more no. of rows. you can read the table directly rather than using the index(Full table scan)

STAR QUERY transformer generally used to improve the query performance in generally star schema models

The star transformation is a cost-based query transformation aimed at executing star queries efficiently. Whereas the star optimization works well for schemas with a small number of dimensions and dense fact tables, the star transformation may be considered as an alternative if any of the following holds true:

In order to get the best possible performance for star queries, it is important to
follow some basic guidelines:

A bitmap index should be built on each of the foreign-key columns of the fact
table(s).

The initialization parameter STAR_TRANSFORMATION_ENABLED should be
set to TRUE. This enables an important optimizer feature for star-queries; it is
set to FALSE by default for backwards-compatibility.

n The cost-based optimizer should be used. [This does not apply solely to star
schemas: all data warehouses should always use the cost-based optimizer].

If u have data coming from diff. Sources what transformation will u use in your designer?
> Joiner Transformation

What is a reusable transf.. What is a mapplet . Explain diff. Bet them

Reusable tranformation:- if u want to create tranfor that perform common tasks such as avg sal in a dept

Mapplet:- Is a reusuable object that represents a set of transformations

How **many transformations are there**.
> Standard and reusable
> Standard:- Aggregator, expression, external procedure, advanced external procedure, filter, joiner, lookup, normalizer, sequence, source qualifer, stored procedure, update strategy

Can u stop a session in a concurrent batch?
> If the session u want to stop is part of a batch u must stop the batch to stop the session. If the session is part of a batch that is nested in a series of batches u must stop the outermost batch to stop the session

What happens when u use the delete or update or reject or insert statement in ur update strategy?
> Inserts:- treats all records as inserts , while inserting if the record violates primary, foreign key or foreign key in the database it rejects the record

What are variables and where do u use it?
> Use variable to simplify complex expressions, temporarily store date, store values from prior rows, compare values, capture multiple return values from a stored procedure, store the results of an unconnected lookup

What is conn. And unconn. Lookup?

A connected lookup transformation is part of the mapping flow,. With connected lookup values u can have multiple return values, Support Default values

Unconnected lookup transformation exists separate from the data flow in the mapping

U write an expr using :lkp reference qualifier to call the lookup within another transformation

Uses:- testing the results of a lookup in a transformation

Filtering records based on the lookup results

Calling the same lookup multiple times in one mapping

To update slowly changing dimensions, Does not Support Default values

Can u access a repository created in previous version of Informatica?

What are the diff.ports available?

Input Port :- Which receives data (Target)

Output Port:- Which provides data (Source)

Input/Output Port:- Which pass data through them (Mapplets)

Variable Port :- Used to store components of expressions

What happens if the info. Server doesn't find the session parameter in the parameter file?

How do run session from command line?

pmcmd

What are diff. Things u can do using PMCMD ?

U can stop or abort a session

What is pmrep command

Have u created parallel sessions How do u create parallel sessions ?

U can improve performance by creating a concurrent batch to run several sessions in parallel on one informatic server, if u have several independent sessions using separate sources and separate mapping to populate diff targets u can place them in a concurrent batch and run them at the same time, if u have a complex mapping with multiple sources u can separate the mapping into several simpler mappings with separate sources. Similarly if u has session performing a minimal no of transformations on large amounts of data like moving flat files to staging area, u can separate the session into multiple sessions and run them concurrently in a batch cutting the total run time dramatically

What is dynamic insert

What are diff types of cache diff between persistent and static cache

Static cache :- Read only will be deleted after use

Persistent cache :- Reused saved on the server

What are session parameters ? How do u set them ?

When do u use mapping parameters ? (In which transformations)

What is a parameter When and where do u them when does the value will be created
Desing time, run time . If u don't create parameter what will happen

How do you do error handling in Informatica ?
   Error handling is very primitive.
   Log files can be generated which contain error details and code.
   The error code can be checked from troubleshooting guide and corrective action   taken.The log
   file can be increased by giving appropriate tracing level in the session properties. Also we can
   give that one Session can stop after 1,2 or n number of errors.
How do you implement configuration management in Informatica?
   There are several methods to do this .Some of them are :-
Taking a back up of the repository as a binary file and treat it as a configurable item.
Implement Folder Versioning utility in Informatica.
Scenario :-. A mapping contains Source Table S_Time (Start_Year, End_Year)
   Target Table Time_Dim (Date, Day, Month, Year, Quarter)
   Stored procedure transformation : A procedure has two input parameters I_Start_Year,
   I_End_Year and output parameter as O_Date, Day , Month, Year, Quarter. If this session is
   running, how many rows will be available in the target and why ?.
   Only one row the last date of the End_Year.
   All the subsequent rows are overriding the previous rows.

What is the difference between lookup cache and lookup index.
   Look up Cache contains Index cache and data cache
   Index cache:Contains columns used in condition
   Data cache: :Contains other output columns than the condition columns.
 Discuss two approaches for updation of target table in informatica and   how they are different.

Update strategy transformation: We can write our own code .It is flexible.
Normal insert / update /delete (with proper variation of the update option) :
   It can be configured in the Session properties.
   Any change in the row will cause an update. Inflexible.

 How do you handle performance issues in Informatica? Where can you monitor **the performance?**
There are several aspects to the performance handling .Some of them are:-
   Source tuning
   Target tuning
   Repository tuning
   Session performance tuning

Incremental Change identification in source side.
Software, hardware (Use multiple servers) and network tuning.
Bulk Loading
Use the appropriate transformation.
To monitor this
Set performance detail criteria
Enable performance monitoring
Monitor session at runtime &/ or Check the performance monitor file.

What is a suggested method for validating fields / marking them with errors?
One of the successful methods is to create an expression object, which contains variables. One variable per port that is to be checked. Set the error "flag" for that field, then at the bottom of the expression trap each of the error fields. From this port you can choose to set flags based on each individual error which occurred, or feed them out as a combination of concatenated field names – to be inserted in to the database as an error row in an error tracking table.

Where is the cache (lookup, index) created and how can you see it.
The cache is created in the server. Some default memory is allocated for it. Once that memory is exceeded than these files can be seen in the Cache directory in the Sever, not before that.
When do you use SQL override in Look up Transformation.

Use SQl override when you have more than one look up table To use where condition to reduce records in cache.

Explain how "constraint based load ordering" works?.
- Constraint based load ordering in PowerMart / PowerCenter works like this: it controls the order in which the target tables are committed to a relational database. It is of no use when sending information to a flat file. To construct the proper constraint order: links between the TARGET tables in Informatica need to be constructed. Simply turning on "constraint based load ordering" has no effect on the operation itself. Informatica does NOT read constraints from the database when this switch is turned on. Again, to take advantage of this switch, you must construct primary / foreign key relationships in the TARGET TABLES in the designer of Informatica. Creating primary / foreign key relationships is difficult - you are only allowed to link a single port (field) to a single table as a primary / foreign key.

What is the difference between Power mart and Power Centre.
Power Center - has all the functionality .
distributed metadata(repository).
global repository and can register multiple Informatica servers. One can share metadata across repositories.

Can connect to Varied sources like Peoplesoft,SAP etc.
Has bridges which can transport meta data from opther tools (like Erwin)
Cost around 200K US $.

Power Mart – Subset of Power centre.
One repository and can register only one Informatica server.
Cannot connect to Varied sources like Peoplesoft,SAP etc
Cost around 50K US $.

What is the difference between Oracle Sequence and Informatica Sequence and which is better.
Oracle sequence can be used in a Pl/Sql stored procedure, which in turn can be used with stored procedure transformation of Informatica.
Informatica sequence is generated through sequence generator transformation of Informatica.
It depends upon the user needs but Oracle sequence provides greater control.

How do you execute a set of Sql commands before running as session and after completion of session in Informatica? Explain.
Sql commands can be put in stored procedures.
Two Unconnected Stored procedure Transformations are created pointing to respective procedures one pre session ,other post session.
When the Session is run these two procedures are executed before the session and after the session.

How can you utilize COM components in Informatica.
By writing C+,VB,VC++ code in External Stored Procedure Transformation

What is an indicator file and how it can be used.
Indicator file is used for Event Based Scheduling when you don't know when the Source Data is availaible., A shell command ,script or a batch file creates and send this indicator file to the directory local to the Informatica Server.Server waits for the indicator file to appear before running the session.

What persistent cache? When it should be used.
When Lookup cache is saved in Look up Transformation It is called persistent cache.
The first time session runs it is saved on the disk and utilized in subsequent running of the Session.
It is used when the look up table is Static i.e does'nt change frequently

What is Incremental Aggregation and how it should be used .
If the source changes only incrementally and you can capture changes, you can configure the session to process only those changes. This allows the Informatica Server to update your target incrementally, rather than forcing it to process the entire source and recalculate the same calculations each time you run the session. Therefore, only use incremental aggregation if:
Your mapping includes an aggregate function.
The source changes only incrementally.
You can capture incremental changes. You might do this by filtering source data by timestamp.
Before implementing incremental aggregation, consider the following issues:
Whether it is appropriate for the session

What to do before enabling incremental aggregation
When to reinitialize the aggregate caches

Scenario :-Informatica Server and Client are in different machines. You run a session from the server manager by specifying the source and target databases. It displays an error. You are confident that everything is correct. Then why it is displaying the error?

The connect strings for source and target databases are not configured on the Workstation conatining the server though they may be on the client m/c.

Informatica

Duration : 1 Hr.   Max. Marks : 100

Where exactly the sources and target information stored ?   (2)
Informatica Repository Tables
What is the difference between power mart and power centre. Elaborate. (2)
Power center is a global repository
What are variable ports and list two situations when they can be used?   (2)

What are the parts of Informatica Server?   (2)
How does the server recognise the source and target databases.
    Elaborate on this.   (2)
List the transformation used for the following: (10)
    Heterogeneous Sources
    Homogeneous Sources
    Find the 5 highest paid employees within a dept.
    Create a Summary table
    Generate surrogate keys
**What is the difference between sequential batch and concurrent batch and which is recommended
    and why ?**   (2)
Designer is used for _____   (1)
Repository Manager is used for _____   (1)
Server Manager is used for _____   (1)
Server is used for _____   (1)
A session S_MAP1 is in Repository A. While running the session error message has displayed
    ' server hot-ws270 is connect to Repository B '. What does it mean ?   (2)
How do you do error handling in Informatica ?   (2)
How do you implement scheduling in Informatica ?   (2)
What is the meaning of upgradation of repository ?   (2)
How can you run a session without using server manager ?   (2)
What is indicator file and where it is used ?   (2)
What are pre and post session stored procedures ? Write a suitable example. (2)
Consider two cases :

1. Power Center Server and Client on the same machine
2. Power Center Sever and Client on the different machines

What is the basic difference in these two setups and which is recommended? (2)

Informatica Server and Client are in different machines. You run a session from the server manager by specifying the source and target databases. It displays an error. You are confident that everything is correct. Then why it is displaying the error?   (2)

When you connect to repository for the first time it asks you for user name & password of repository and database both.

But subsequent times it asks only repository password. Why ?   (2)

What is the difference between normal and bulk loading.

Which one is recommended?   (2)

What is a test load ?   (2)

What is an incremental aggregation and when it should be implemented? (2)

How can you use an Oracle sequences in Informatica? You have an Informatica sequence generator transformation also. Which one is better to use?   (2)

What is the difference between a shortcut of an object and copy of an object?

Compare them.   (2)

What is mapplet and a reusable transformation?   (2)

How do you implement configuration management in Informatica?   **(3)**

What are Business Components in Informatica?   (2)

Dimension Object created in Oracle can be imported in Designer (T/ F)   (1)

Cubes contain measures (T / F)   (1)

COM components can be used in Informatica (T / F)   (1)

Lookup is an Active Transformation (T/F)   (1)

What is the advantage of persistent cache? When it should be used.   **(1)**

When will you use SQL override in a lookup transformation?   (2)

Two different admin users created for repository are _____ and_____ (1)

Two Default User groups created in the repository are ____ and _____   (1)

A mapping contains

Source Table S_Time (Start_Year, End_Year)

Target Table Tim_Dim (Date, Day, Month, Year, Quarter)

Stored procedure transformation : A procedure has two input parameters I_Start_Year, I_End_Year and output parameter as O_Date, Day , Month, Year, Quarter. If this session is running, how many rows will be available in the target and why ?   (5)

39. Two Sources S1, S2 containing measures M1,M2,M3, 4 Dimensions D1,D2,D3,D4,
1 Fact F1 containing measures M1, M2,M3 and Dimension Surrogate keys K1,K2,K3,K4

(a) Write a SQL statement to populate Fact table F1

(b) Design a mapping in Informatica for loading of Fact table F1.   (5)

40. What is the difference between connected lookup and unconnected lookup.   (2)

41. What is the difference between datacahe and lindex cahe.   (2)

42. When should one create a lookup transformation?   (2)

43. How do you handle performance issues in Informatica.Where can you monitor the performance ?   **(3)**

44. List and Discuss two approaches for updation of target table in informatica and how they are different.   (3)
45. You have created a lookup tansformation for a certain condition which if true returns multiple rows .When you go to the target and see only one row has come and not all. Why is it so and how it can be corrected.   (2)
46. Where are the log files generally stored. Can you change the path of the file. What can the path be?   (2)
47. Where is the cache (lookup, index) created and how can you see it.   (2)