**Related to DDL, DML**

Why can't we write ddl statements directly in PL/SQL block, for example when i write

CREATE OR REPLACE PROCEDURE test IS

BEGIN

truncate table table\_name; // error

END test;

/

But,

CREATE OR REPLACE PROCEDURE test IS

BEGIN

execute immediate 'truncate table table\_name'; // works fine

END test;

/

Why second one executed successfully ?

Only dynamic SQL can execute the following types of statements within PL/SQL program units:

* Data definition language (DDL) statements such as CREATE, DROP, GRANT, and REVOKE

A TRUNCATE operation is DDL.

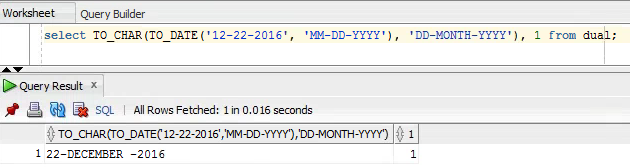
When using EXECUTE IMMEDIATE, remember that any DDL operations you execute will implicitly COMMIT the current transaction.

**Name of the column after firing an SQL statement**

select column\_name, num\_distinct, num\_nulls, histogram

from user\_tab\_col\_statistics

where table\_name = 'T';

****

**11. What is the difference between STOP and ABORT options in Workflow Monitor?**

On issuing the STOP command on the session task, the integration service stops reading data from the source although it continues processing the data to targets. If the integration service cannot finish processing and committing data, we can issue the abort command.

ABORT command has a timeout period of 60 seconds. If the integration service cannot finish processing data within the timeout period, it kills the DTM process and terminates the session

**12. How can we store previous session logs?**

If you run the session in the time stamp mode then automatically session log out will not overwrite the current session log.

Go to Session Properties –> Config Object –> Log Options

Select the properties as follows:

Save session log by –> SessionRuns

Save session log for these runs –> Change the number that you want to save the number of log files (Default is 0)

If you want to save all of the log files created by every run, and then select the option Save session log for these runs –> Session TimeStamp

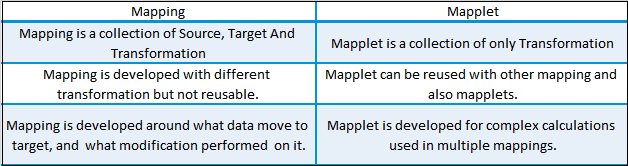
You can find these properties in the session/workflow Properties.

## ****16. What are data driven sessions?****

When you configure a session using update strategy, the session property data driven instructs Informatica server to use the instructions coded in mapping to flag the rows for insert, update, delete or reject. This is done by mentioning DD\_UPDATE or DD\_INSERT or DD\_DELETE in the update strategy transformation.

“Treat source rows as” property in session is set to “Data Driven” by default when using a update strategy transformation in a mapping.

## ****18. What is the difference between Mapping and Mapplet?****

[](http://www.edureka.co/informatica?utm_source=blog&utm_medium=image&utm_campaign=top-informatica-interview-questions-you-must-prepare-for-in-2016-14jan16)

## ****19. How can we delete duplicate rows from flat files?****

We can make use of sorter transformation and select distinct option to delete the duplicate rows.

## ****24. State the differences between SQL Override and Lookup Override?****

* The role of SQL Override is to limit the number of incoming rows entering the mapping pipeline, whereas Lookup Override is used to limit the number of lookup rows to avoid the whole table scan by saving the lookup time  and the cache it uses.
* Lookup Override uses the “Order By” clause by default. SQL Override doesn’t use it and should be manually entered in the query if we require it
* SQL Override can provide any kind of ‘join’ by writing the queryLookup Override provides only Non-Equi joins.
* Lookup Override gives only one record even if it finds multiple records for a single condition  
  SQL Override doesn’t do that.

## ****26. What are the different ways to implement parallel processing in Informatica?****

We can implement parallel processing using various types of partition algorithms:

**Database partitioning:** The Integration Service queries the database system for table partition information. It reads partitioned data from the corresponding nodes in the database.

**Round-Robin Partitioning:** Using this partitioning algorithm, the Integration service distributes data evenly among all partitions. It makes sense to use round-robin partitioning when you need to distribute rows evenly and do not need to group data among partitions.

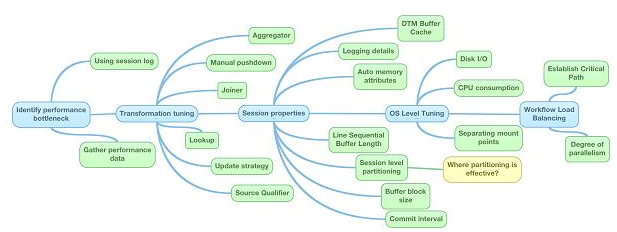
**Hash Auto-Keys Partitioning:** The Powercenter Server uses a hash function to group rows of data among partitions. When the hash auto-key partition is used, the Integration Service uses all grouped or sorted ports as a compound partition key. You can use hash auto-keys partitioning at or before Rank, Sorter, and unsorted Aggregator transformations to ensure that rows are grouped properly before they enter these transformations.

**Hash User-Keys Partitioning:** Here, the Integration Service uses a hash function to group rows of data among partitions based on a user-defined partition key. You can individually choose the ports that define the partition key.

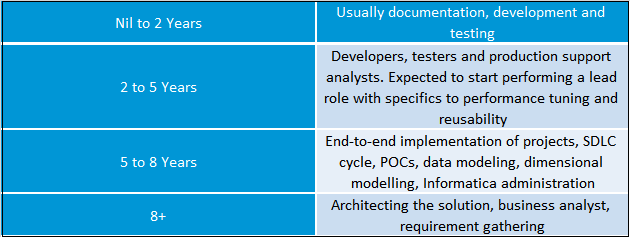
**Key Range Partitioning:** With this type of partitioning, you can specify one or more ports to form a compound partition key for a source or target. The Integration Service then passes data to each partition depending on the ranges you specify for each port.

**Pass-through Partitioning:** In this type of partitioning, the Integration Service passes all rows from one partition point to the next partition point without redistributing them.

## ****27. What are the different levels at which performance improvement can be performed in Informatica?****



## ****29. What are the different types of profiles in Informatica?****



### **QUESTION #1**

**What is a control task?**

A control task is used to alter the normal processing of a workflow by stopping, aborting, or failing a workflow or worklet.

### **QUESTION #2**

**What is a pipeline partition and how does provide a session with higher performance?**

Within a mapping, a session can break apart different source qualifier to target pipelines into their own reader/transformation/writer thread(s). This allows the Integration Service to run the partition in parallel with other pipeline partitions in the same mapping. The parallelism creates a higher performing session.

### **QUESTION #3**

**What is the maximum number of partitions that can be defined for in a single pipeline?**

You can define up to 64 partitions at any partition point in a pipeline.

### **QUESTION #4**

**Pipeline partitions is designed to increase performance, however list one of it’s disadvantages?**

Increasing the number of partitions increases the load on the node. If the node does not contain enough CPU bandwidth, you can overload the system.

### **QUESTION #5**

**What is a dynamic session partition?**

A dynamic session partition is where the Integration Service scales the number of session partitions at runtime. The number of partitions is based on a number of factors including number of nodes in a grid or source database partitions.

### **QUESTION #6**

**List three dynamic partitioning configurations that cause a session to run with one partition**

1. You set dynamic partitioning to the number of nodes in the grid, and the session does not run on a grid.

2. You create a user-defined SQL statement or a user-defined source filter.

3. You use dynamic partitioning with an Application Source Qualifier.

### **QUESTION #7**

**What is pushdown optimization?**

Pushdown optimization is a feature within Informatica PowerCenter that allows us to push the transformation logic in a mapping into SQL queries that are executed by the database. If not all mapping transformation logic can be translated into SQL, then the Integration Service will process what is left.

### **QUESTION #8**

**List the different types of pushdown optimization that can be configured?**

1. Source-side pushdown optimization – The Integration Service pushes as much transformation logic as possible to the source database.

2. Target-side pushdown optimization – The Integration Service pushes as much transformation logic as possible to the target database.

3. Full pushdown optimization – The Integration Service attempts to push all transformation logic to the target database. If the Integration Service cannot push all transformation logic to the database, it performs both source-side and target-side pushdown optimization.

### **QUESTION #9**

**What databases are we able to configure pushdown optimization?**

IBM DB2

Microsoft SQL Server

Netezza

Oracle

Sybase ASE

Teradata

Databases that use ODBC drivers

### **QUESTION #10**

**List several transformations that work with pushdown optimization to push logic to the database?**

Aggregator, Expression, Filter, Joiner, Lookup, Router, Sequence Generator, Sorter, Source Qualifier, Target, Union, Update Strategy

### **QUESTION #11**

**What is real-time processing?**

Data sources such as JMS, WebSphere MQ, TIBCO, webMethods, MSMQ, SQP, and webservices can publish data in real-time. These real-time sources can be leveraged by Informatica PowerCenter to process data on-demand. A session can be specifically configured for real-time processing.

### **QUESTION #12**

**What types of real-time data can be processed with Informatica PowerCenter**

1. Messages and message queues. Examples include WebSphere MQ, JMS, MSMQ, SAP, TIBCO, and webMethods sources.

2. Web service messages. Example includes receiving a message from a web service client through the Web Services Hub.

3. Change data from PowerExchange change data capture sources.

### **QUESTION #13**

**What is a real-time processing terminating condition?**

A real-time processing terminating condition determines when the Integration Service stops reading messages from a real-time source and ends the session.

### **QUESTION #14**

**List three real-time processing terminating conditions?**

1. Idle time – Time Integration Service waits to receive messages before it stops reading from the source.

2. Message count – Number of messages the Integration Service reads from a real-time source before it stops reading from the source.

3. Reader time limit – Amount of time in seconds that the Integration Service reads source messages from the real-time source before it stops reading from the source

### **QUESTION #15**

**What is real-time processing message recovery?**

Real-time processing message recovery allows the Integration Service to recover unprocessed messages from a failed session. Recovery files, tables, queues, or topics are used to recover the source messages or IDs. Recovery mode can be used to recover these unprocessed messaged.

### **QUESTION #16**

**What factors play a part in determining a commit point?**

1. Commit interval

2. Commit interval type

3. Size of the buffer blocks

### **QUESTION #17**

**List all configurable commit types?**

1. Target-based commit – Data committed based on the number of target rows and the key constraints on the target table.

2. Source-based commit – Data committed based on the number of source rows.

3. User-defined commit – Data committed based on transactions defined in the mapping properties.

### **QUESTION #18**

**What performance concerns should you be aware of when logging error rows?**

Session performance may decrease when logging row errors because the Integration Service processes one row at a time instead of a block of rows at once.

### **QUESTION #19**

**What functionality is provided by the Integration Service when error logging is enabled?**

Error logging builds a cumulative set of error records in a error log file or error table created by the Integration Service.

### **QUESTION #20**

**What is the difference between stopping and aborting a workflow session task?**

A **stop command** tells the Integration Service to stop reading session data, but will continue writing and committing data to targets.

A **abort command** works exactly like the stop command, however it will tell the Integration to stop processing and committing data to targets after 60 seconds. If all processes are not complete after this time out period, the session gets terminated.

### **QUESTION #21**

**What is a concurrent workflow?**

A concurrent workflow is a workflow that can run as multiple instances concurrently. Concurrent workflows can be configured in one of two ways:

1. Allow concurrent workflows with the same instance name.

2. Configure unique workflow instances to run concurrently.

### **QUESTION #22**

**What is Informatica PowerCenter grid processing and its benefits?**

Grid processing is a feature of PowerCenter that enables workflows and sessions to be run across multiple domain nodes. PowerCenter grid’s parallel processing provides increased performance and scalability.

### **QUESTION #23**

**List the types of parameters and variables that can be defined within a parameter file?**

Service variables, service process variables, workflow and worklet variables, session parameters, and mapping parameters and variables.

### **QUESTION #24**

**With PowerCenter, what two locations can one specify a parameter file?**

1. Within the session task.

2. Within the workflow.

### **QUESTION #25**

**How are mapplet parameters and variables defined withing a parameter file different?**

Mapplet parameters and variables are different because they must be preceded with the mapplet name they were definee within. For example, a parameter by name of MyParameter, defined within mapplet MyMapplet, would be set to a value of 10 in a related parameter file by using syntax: MyMapplet.MyParameter=10.

### **QUESTION #26**

**What is an SQL Transformation in Informatica?**

The SQL transformation is is active, passive, and connected. It allows for runtime SQL processing. It allows data to be retrieved, inserted, updated, and deleted midstream in a mapping pipeline. SQL transformations have two modes, script mode and query mode. Script mode allows for external located script files to be called to execute SQL. Query mode allows for SQL to be placed within the transformation’s editor to execute SQL logic.

### **QUESTION #27**

**What is dynamic lookup cache?**

Dynamic lookup cache is cache that has been built from the first lookup request. Each subsequent row that passes through the lookup will query the cache. As these rows are processed or inserted into the lookup’s target table, the lookup cache is also updated dynamically.

### **QUESTION #28**

**What is a Unstructured Data transformation?**

The Unstructured Data transformation is active, passive, and connected. It leverages the Data Transformation application to transform unstructured, semi-structured, and structured file formats such as messaging formats, HTML pages, PDF documents, ACORD, HIPAA, HL7, EDI-X12, EDIFACT, and SWIFT. Once data has been transformed by Data Transformation, it can be returned in a mapping pipeline and further transformed and/or loaded to an appropriate target.

### **SUMMARY**

Well I hope these **25+ Advanced Informatica interview questions** have been helpful and can assist you in landing your next big data integration gig!

Feel free to post any potential interview questions of your own or a question I can help you answer.

Remember to checkout my posts below for even more Informatica and Data Warehouse Interview question assistance…

### **QUESTION #24**

***What is the purpose of the INITCAP function?***

The INITCAP function capitalizes the first letter in each word of a string and converts all other letters to lowercase.

EX: INITCAP(IN\_DATA)

|  |  |
| --- | --- |
| **IN\_DATA** | **RETURN VALUE** |
| informatica interview questions | Informatica Interview Questions |

### **QUESTION #34**

***What is the recommended order for Optimizing Informatica PowerCenter performance tuning bottlenecks?***

1. Target
2. Source
3. Mapping
4. Transformation
5. Session
6. Grid Deployments
7. PowerCenter Components
8. System

## QUESTION #4

**What is a star schema?**

A star schema is a simplistic two-dimensional, star shaped data model. Its useful design connects one or more fact tables with a number of dimension tables via primary/foreign key relationships to enable strategic decision making.

## QUESTION #5

**What is a dimension table?**

A dimension table contains the descriptive attributes of a business process event that enable filtering and grouping of facts. Fact tables can be found in both star schemas and OLAP cubes where there are relationships between dimension and fact tables. Dimension table attributes describe the “who, what, when, where, why, and how” of the business process being modeled.

## QUESTION #6

**What is a fact table?**

Fact tables contain business process event measurements that are almost always numeric. A fact table’s grain is defined through this physical observable business event. Fact tables can be found in both star schemas and OLAP cubes where there are relationships between dimension and fact tables. There are several categories of fact tables including: Additive, Semi-Additive, and Non-Additive.

## QUESTION #7

**Numeric measures in a fact table fall into these three categories?**

1. Additive

2. Semi-Additive

3. Non-Additive

## QUESTION #8

**What is an Additive Fact?**

An additive fact is a measure in a fact table that can be fully summed across any of the dimensions associated with it.

## QUESTION #9

**What is a Semi-Additive Fact?**

A semi-additive fact is a measure in a fact table that can be summed across some dimensions associated with it, but not all.

## QUESTION #10

**What is a Non-Additive Fact?**

A non-additive fact is a measure in a fact table that cannot be summed across any of the dimensions associated with it. A good example of this are ratios.

## QUESTION #11

**What are the three fundamental types of fact tables?**

1. Transaction

2. Periodic Snapshot

3. Accumulating Snapshot

## QUESTION #12

**What is a transaction fact table?**

A transaction fact table represents an event that occurred at an instantaneous point in time.

## QUESTION #13

**What is a periodic snapshot fact table?**

A periodic snapshot fact table captures cumulative business performance on regular, predictable time intervals.

## QUESTION #14

**What is a accumulating snapshot fact table?**

Accumulating snapshot fact tables represent processes that have a definite beginning and end together with a standard set of intermediate process steps.

## QUESTION #15

**What is an Enterprise Data Warehouse Bus Matrix?**

An enterprise data warehouse bus matrix is a document of the data warehouse bus architecture.

## QUESTION #17

**What is a centipede fact table?**

A fact table where a modeler has decided to normalize the fact table istead of snowflaking the dimension table(s) (which is also discouraged).

## QUESTION #18

**What is a stakeholder matrix?**

A stakeholder matrix is similar to a enterprise bus matrix. However, it replaces common dimensions with stakeholder such as merchandising, marketing, store operations, logistics, and finance. It is a guide to help understand which business departments are interested in each business process.

## QUESTION #19

**Do conformed dimensions work for an agile development environment?**

While some argue that the organizational agreement necessary to build conformed dimensions in organization does not fit well with an agile environment, arguments can be made to the contrary.

One might argue that conformed dimension allow tables to be build once and replicated rather than re-creating and re-inventing the wheel all over the organization. Development will actually begin to speed up greatly once common dimensions are established.

## QUESTION #20

**What is a conformed facts?**

They are measures that can be re-used across multiple dimensional models. Key performance indicators (KPI) such as profit, revenue, and costs are all example of possible conformed facts.

agreement of what a single fact table represents. A business process’s lowest level of captured data can be referred to as atomic grain.

## QUESTION #22

**What is a type 0 slowly changing dimension?**

A type 0 slowly changing dimension is a dimension table with attributes values that never change. Attributes are considered and may be labeled “original”.

## QUESTION #23

**What is a type 1 slowly changing dimension?**

A type 1 slowly changing dimension is a dimension table with current value attributes. In other works, old attribute values are overwritten by the most recent record values.

## QUESTION #24

**What is a type 2 slowly changing dimension?**

A type 2 slowly changing dimension updates dimension attribute values by inserting a new row in the dimension and inactivating its prior record.

## QUESTION #25

**What is a type 3 slowly changing dimension?**

A type 3 slowly changing dimension, a new row is not inserted with a change, however a prior attribute column is added and updated with the old attribute value within the same record. This allows business to understand a previous value for an attribute as well as the current one, all within the same record. Reports would then align with the current value attribute, but business could easily report on the previous value if needed.

## QUESTION #26

**What is a junk dimension?**

A junk dimension, also referred to as an indicator or transaction profile dimension is a separate dimension table that contains flags and indicators which have been removed from a fact table. By placing these flags and indicators into their own table, we are able to clean up the complex transactional fact table.

## QUESTION #27

**What is a audit dimension?**

While loading certain fact tables, like a shipment invoice line transaction fact, metadata is generated that can be useful to not only IT, but to business. ETL processing characteristics like data quality indicators and others can be stored in this audit dimension. Business can then possibly leverage it by helping them understand anomalous values, versions, and gain a general confidence in reported numbers.

## QUESTION #28

**Define Kimball’s four-step dimensional design process?**

1. Select the business process.

2. Declare the grain.

3. Identify dimensions.

4. Identify the facts.

46. What is the target load order?  
  
      U specify the target loadorder based on source qualifiers in a maping.If u have the multiple   
      source qualifiers connected to the multiple  targets,U can designatethe order in which informatica  
      server loads data into the targets.

## 47.What is the default join that source qualifier provides?      Inner equi join.

51.What is the default source option for update stratgey transformation?  
  
        Data driven.  
  
52. What is Datadriven?  
  
          The informatica server follows instructions coded into update strategy transformations with  
        in the session maping determine how to flag records for insert,update,,delete  or reject  
        If u do not choose data driven  option setting,the informatica server ignores all update strategy  
       transformations in the mapping.  
  
53.What r the options in the target session of  update strategy transsformatioin?  
  
      Insert  
      Delete  
      Update  
      Update as update  
      Update as insert  
      Update esle insert  
      Truncate table

59. What r two types of processes that informatica runs the session?  
  
  Load manager Process: Starts the session, creates the DTM process, and sends post-session email when the session completes.  
 The DTM process. Creates threads to initialize the session, read, write, and transform data, and handle pre- and post-session operations.

### **What are the new features of Informatica 9.x in developer level?**

From a developer's perspective, some of the new features in Informatica 9.x are as follows:

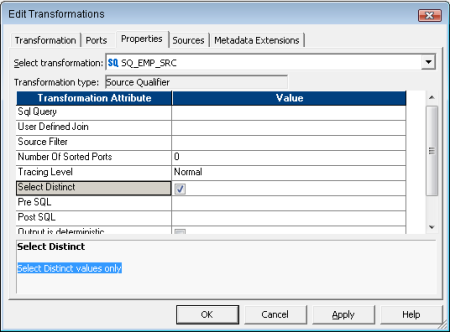
* Now Lookup can be configured as an active transformation - it can return multiple rows on successful match
* Now you can write SQL override on un-cached lookup also. Previously you could do it only on cached lookup
* You can control the size of your session log. In a real-time environment you can control the session log file size or time
* Database deadlock resilience feature - this will ensure that your session does not immediately fail if it encounters any database deadlock, it will now retry the operation again. You can configure number of retry attempts.

### **How to Delete duplicate row using Informatica**

#### **Scenario 1: Duplicate rows are present in relational database**

Suppose we have Duplicate records in Source System and we want to load only the unique records in the Target System eliminating the duplicate rows. What will be the approach?

Assuming that the source system is a **Relational Database**, to eliminate duplicate records, we can check the **Distinct** option of the **Source Qualifier** of the source table and load the target accordingly.



But what if the source is a flat file? Then how can we remove the duplicates from flat file source?

#### **Scenario 2: Deleting duplicate rows / selecting distinct rows for FLAT FILE sources**

Here since the source system is a **Flat File** you will not be able to select the distinct option in the source qualifier as it will be disabled due to flat file source table. Hence the next approach may be we use a **Sorter Transformation** and check the **Distinct** option. When we select the distinct option all the columns will the selected as keys, in ascending order by default.

#### **Deleting Duplicate Record Using Informatica Aggregator**

Other ways to handle duplicate records in source batch run is to use an **Aggregator Transformation** and using the **Group By** checkbox on the ports having duplicate occurring data. Here you can have the flexibility to select the *last or the first* of the duplicate column value records.

There is yet another option to ensure duplicate records are not inserted in the target. That is through **Dynamic lookup** cache. Using Dynamic Lookup Cache of the target table and associating the input ports with the lookup port and checking the Insert Else Update option will help to eliminate the duplicate records in source and hence loading unique records in the target

### **Loading Multiple Target Tables Based on Conditions**

#### **Scenario**

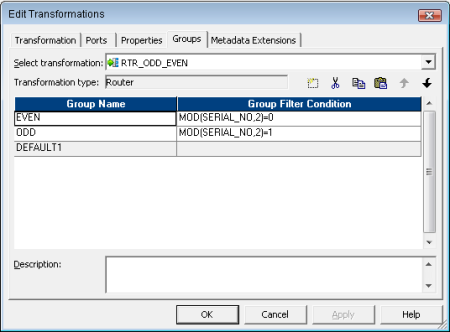
Suppose we have some serial numbers in a flat file source. We want to load the serial numbers in two target files one containing the EVEN serial numbers and the other file having the ODD ones.

#### **Answer**

After the Source Qualifier place a **Router Transformation**. Create two **Groups** namely **EVEN and ODD**, with filter conditions as:

MOD(SERIAL\_NO,2)=0 and MOD(SERIAL\_NO,2)=1

... respectively. Then output the two groups into two flat file targets.



### **Normalizer Related Questions**

#### **Scenario 1**

Suppose in our Source Table we have data as given below:

| **Student Name** | **Maths** | **Life Science** | **Physical Science** |
| --- | --- | --- | --- |
| Sam | 100 | 70 | 80 |
| John | 75 | 100 | 85 |
| Tom | 80 | 100 | 85 |

We want to load our Target Table as:

| **Student Name** | **Subject Name** | **Marks** |
| --- | --- | --- |
| Sam | Maths | 100 |
| Sam | Life Science | 70 |
| Sam | Physical Science | 80 |
| John | Maths | 75 |
| John | Life Science | 100 |
| John | Physical Science | 85 |
| Tom | Maths | 80 |
| Tom | Life Science | 100 |
| Tom | Physical Science | 85 |

Describe your approach.

#### **Answer**

Here to convert the Rows to Columns we have to use the **Normalizer Transformation** followed by an Expression Transformation to Decode the column taken into consideration. For more details on how the mapping is performed please visit [Working with Normalizer](https://dwbi.org/etl/informatica/147-using-informatica-normalizer-transformation)

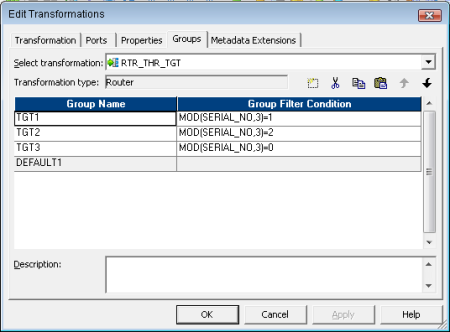
#### **Scenario 2**

Suppose we have a source table and we want to load three target tables based on source rows such that first row moves to first target table, second row in second target table, third row in third target table, fourth row again in first target table so on and so forth. Describe your approach.

#### **Answer**

We can clearly understand that we need a **Router transformation** to route or filter source data to the three target tables. Now the question is what will be the filter conditions. First of all we need an **Expression Transformation** where we have all the source table columns and along with that we have another i/o port say seq\_num, which is gets sequence numbers for each source row from the port **NextVal**of a **Sequence Generator start value 0 and increment by 1**. Now the filter condition for the three router groups will be:

* MOD(SEQ\_NUM,3)=1 connected to 1st target table
* MOD(SEQ\_NUM,3)=2 connected to 2nd target table
* MOD(SEQ\_NUM,3)=0 connected to 3rd target table



#### **How can we implement Aggregation operation without using an Aggregator Transformation in Informatica?**

#### **Answer**

We will use the very basic concept of the **Expression Transformation** that at a time we can access the previous row data as well as the currently processed data in an expression transformation. What we need is simple Sorter, Expression and Filter transformation to achieve aggregation at Informatica level.

**Q27.** How does Joiner transformation treat NULL value matching.

**Ans.** The Joiner transformation **does not match null values**

**NOTE:** Only equality operator is available in joiner join condition.

**Q33.** Suppose we have a source table populating two target tables. We connect the NEXTVAL port of the Sequence Generator to the surrogate keys of both the target tables.

Will the Surrogate keys in both the target tables be same? If not how can we flow the same sequence values in both of them.

**Ans.** When we connect the **NEXTVAL** output port of the **Sequence Generator** directly to the surrogate key columns of the target tables, the **Sequence number will not be the same**.

A block of sequence numbers is sent to one target tables surrogate key column. The second targets receives a block of sequence numbers from the Sequence Generator transformation only after the first target table receives the block of sequence numbers.

**Q34.** Suppose we have 100 records coming from the source. Now for a target column population we used a Sequence generator.

Suppose the Current Value is 0 and End Value of Sequence generator is set to 80. What will happen?

**Ans.** **End Value** is the maximum value the Sequence Generator will generate. After it reaches the End value the session fails with the following error message:

TT\_11009 Sequence Generator Transformation: Overflow error.

**Q39.** How does Aggregator Transformation handle NULL values?

**Ans.** By default, the aggregator transformation treats null values as NULL in aggregate functions. But we can specify to treat null values in aggregate functions as NULL or zero.

**Q40.** What is Incremental Aggregation?

**Ans.** We can enable the session option, Incremental Aggregation for a session that includes an Aggregator Transformation. When the Integration Service performs incremental aggregation, it actually passes changed source data through the mapping and uses the historical cache data to perform aggregate calculations incrementally.

**Q45.** Suppose we do not group by on any ports of the aggregator what will be the output.

**Ans.** If we do not group values, the Integration Service will return **only the last row** for the input rows.

**Q46.** What is the expected value if the column in an aggregator transform is neither a group by nor an aggregate expression?

**Ans.** Integration Service produces one row for each group based on the group by ports. The columns which are neither part of the key nor aggregate expression will return the corresponding value of last record of the group received. However, if we specify particularly the FIRST function, the Integration Service then returns the value of the specified first row of the group. So default is the **LAST** function.

**Q47.** Give one example for each of Conditional Aggregation, Non-Aggregate expression and Nested Aggregation.

**Ans.**

Use conditional clauses in the aggregate expression to reduce the number of rows used in the aggregation. The conditional clause can be any clause that evaluates to TRUE or FALSE.

SUM( SALARY, JOB = CLERK )

Use non-aggregate expressions in group by ports to modify or replace groups.

IIF( PRODUCT = Brown Bread, Bread, PRODUCT )

The expression can also include one aggregate function within another aggregate function, such as:

MAX( COUNT( PRODUCT ))

**Q50.** What is a RANK port and RANKINDEX?

**Ans.** Rank port is an input/output port use to specify the column for which we want to rank the source values. By default Informatica creates an output port RANKINDEX for each Rank transformation. It stores the ranking position for each row in a group.

**Q55.** How does Rank transformation handle string values?

**Ans.** Rank transformation can return the strings at the top or the bottom of a session sort order. When the Integration Service runs in Unicode mode, it sorts character data in the session using the selected sort order associated with the Code Page of IS which may be French, German, etc. When the Integration Service runs in ASCII mode, it ignores this setting and uses a binary sort order to sort character data.

**Q58.** How does Sorter handle Case Sensitive sorting?

**Ans.** The Case Sensitive property determines whether the Integration Service considers case when sorting data. When we enable the Case Sensitive property, the Integration Service sorts uppercase characters higher than lowercase characters.

**Q59.** How does Sorter handle NULL values?

**Ans.** We can configure the way the Sorter transformation treats null values. Enable the property Null Treated Low if we want to treat null values as lower than any other value when it performs the sort operation. Disable this option if we want the Integration Service to treat null values as higher than any other value.

**Q60.** How does a Sorter Cache works?

**Ans.** The Integration Service passes all incoming data into the Sorter Cache before Sorter transformation performs the sort operation.

**Q61.** What is a Union Transformation?

**Ans.** The Union transformation is an Active, Connected non-blocking multiple input group transformation use to merge data from multiple pipelines or sources into one pipeline branch. Similar to the UNION ALL SQL statement, the Union transformation does not remove duplicate rows.

**Q65.** What is the difference between Reusable transformation and Mapplet?

**Ans.** Any Informatica Transformation created in the in the Transformation Developer or a non-reusable promoted to reusable transformation from the mapping designer which can be used in multiple mappings is known as Reusable Transformation. When we add a reusable transformation to a mapping, we actually add an instance of the transformation. Since the instance of a reusable transformation is a pointer to that transformation, when we change the transformation in the Transformation Developer, its instances reflect these changes.

A Mapplet is a reusable object created in the Mapplet Designer which contains a **set of transformations**and lets us reuse the transformation logic in multiple mappings. A Mapplet can contain as many transformations as we need. Like a reusable transformation when we use a mapplet in a mapping, we use an instance of the mapplet and any change made to the mapplet is inherited by all instances of the mapplet.

**Q66.** What are the transformations that are not supported in Mapplet?

**Ans.** Normalizer, Cobol sources, XML sources, XML Source Qualifier transformations, Target definitions, Pre- and post- session Stored Procedures, Other Mapplets.

**Q69.** Can we copy a session to new folder or new repository?

**Ans.** Yes we can copy session to new folder or repository provided the corresponding Mapping is already in there.

**Q70.** What type of join does Lookup support?

**Ans.** Lookup is just similar like SQL LEFT OUTER JOIN.

**16. What does role playing dimension mean?**

The dimensions that are utilized for playing diversified roles while remaining in the same database domain are called role playing dimensions.

**17. How can repository reports be accessed without SQL or other transformations?**

Ans:Repositoryreports are established by metadata reporter. There is no need of SQL or other transformation since it is a web app.

**18. What are the types of metadata that stores in repository?**

The types of metadata includes Source definition, Target definition, Mappings, Mapplet, Transformations.

**19. Explain the code page compatibility?**

When data moves from one code page to another provided that both code pages have the same character sets then data loss cannot occur. All the characteristics of source page must be available in the target page. Moreover if all the characters of source page are not present in the target page then it would be a subset and data loss will definitely occur during transformation due the fact the two code pages are not compatible.

**20. How can you validate all mappings in the repository simultaneously?**

All the mappings cannot be validated simultaneously because each time only one mapping can be validated.

**40.What is predefined event?**

It is a file-watch event. It waits for a specific file to arrive at a specific location.

**41. How can you define user defied event?**

User defined event can be described as a flow of tasks in the workflow. Events can be created and then raised as need arises.