

ETLLabs.com

# INFORMATICA INTERVIEW QUESTIONS, 25 Scenarios/Solutions

Informatica Interview Questions [ Version 1.1 ]

Compiled by - mahender, uma  
2/1/2013

## Informatica Scenarios

### Scenario1:

We have a target source table containing 3 columns: Col1, Col2 and Col3. There is only 1 row in the table as follows:

Col1	Col2	Col3
a	b	c

There is target table contain only 1 column Col. Design a mapping so that the target table contains 3 rows as follows:

Col
a
b
c

### Solution:

Not using a Normalizer transformation:

Create 3 expression transformations exp\_1, exp\_2 and exp\_3 with 1 port each. Connect col1 from Source Qualifier to port in exp\_1. Connect col2 from Source Qualifier to port in exp\_2. Connect col3 from source qualifier to port in exp\_3. Make 3 instances of the target. Connect port from exp\_1 to target\_1. Connect port from exp\_2 to target\_2 and connect port from exp\_3 to target\_3.

### Scenario 2:

There is a source table that contains duplicate rows. Designs a mapping to load all the unique rows in 1 target while all the duplicate rows (only 1 occurrence) in another target.

### Solution:

Bring all the columns from source qualifier to an Aggregator transformation. Check group by on the key column. Create a new output port COUNT\_COL in aggregator transformation and write an expression COUNT (KEY\_COLUMN). Make a router transformation with 2 GROUPS: Dup and Non-Dup. Check the router conditions COUNT\_COL > 1 in Dup group while COUNT\_COL = 1 in Non-dup group. Load these 2 groups in different targets.

### Scenario 3:

There is a source table containing 2 columns Col1 and Col2 with data as follows:

Col1	Col2
a	l
b	p
a	m
a	n
b	q

x      y

Design a mapping to load a target table with following values from the above mentioned source:

Col1	Col2
a	l, m, n
b	p, q
x	y

## Solution:

Use a sorter transformation after the source qualifier to sort the values with col1 as key. Build an expression transformation with following ports (order of ports should also be the same):

1. COL1\_PREV: It will be a variable type port. Expression should contain a variable example: VAL
2. COL1: It will be Input/output port from Sorter transformation
3. COL2: It will be input port from sorter transformation
4. VAL: It will be a variable type port. Expression should contain Col1
5. CONCATENATED\_VALUE: It will be a variable type port. Expression should be decode (Col1,Col1\_prev,Concatenated\_value|'|'||Col2,Col1)
6. CONCATENATED\_FINAL: It will be an output port connecting the value of CONCATENATED\_VALUE.

After expression, build a Aggregator Transformation. Bring ports Col1 and CONCATENATED\_FINAL into aggregator. Group by Col1. Don't give any expression. This effectively will return the last row from each group.

Connect the ports Col1 and CONCATENATED\_FINAL from aggregator to the target table.

<UMA> this can be achieved by using a database stored procedure also, but it might end up parsing the statement every time. So it would be always good to go with the above mentioned solution.

## Scenario 4:

Design an Informatica mapping to load first half records to 1 target while other half records to a separate target.

## Solution:

You will have to assign a row number with each record. To achieve this, either use Oracle's PSUDO column ROWNUM in Source Qualifier query or use NEXTVAL port of a Sequence generator. Let's name this column as ROWNUMBER.

From Source Qualifier, create 2 pipelines:

First Pipeline:

Carry first port Col1 from SQ transformation into an aggregator transformation. Create a new output port "tot\_rec" and give the expression as COUNT(Col1). Do not group by any port. This will give us the total number of records in Source Table. Carry this port tot\_rec to an Expression Transformation. Add another port DUMMY in expression transformation with default value 1.

Second Pipeline:

from SQ transformation, carry all the ports(including an additional port rownumber generated by rownum or sequence generator) to an Expression Transformation. Add another port DUMMY in expression transformation with default value 1.

Join these 2 pipelines with a Joiner Transformation on common port DUMMY. carry all the source table ports and 2 additional ports tot\_rec and rownumber to a router transformation. Add 2 groups in Router : FIRST\_HALF and SECOND\_HALF. Give condition rownumber<=tot\_rec/2 in FIRST\_HALF. Give condition rownumber>tot\_rec/2 in SECOND\_HALF. Connect the 2 groups to 2 different targets.

## Scenario 5:

A source table contains emp\_name and salary columns. Develop an Informatica mapping to load all records with 5th highest salary into the target table.

### Solution:

The mapping will contain following transformations after the Source Qualifier Transformation:

1. Sorter : It will contain 2 ports - emp\_name and salary. The property 'Direction' will be selected as 'Descending' on key 'Salary'
2. Expression transformation: It will 6 ports as follows -
  - a> emp\_name : It will be an I/O port directly connected from previous sorter transformation
  - b> salary\_prev : It will be a variable type port. Give any variable name e.g val in its Expression column
  - c> salary : It will be an I/O port directly connected from previous transformation
  - d> val : It will be a variable port. The expression column of this port will contain 'salary'
  - e> rank: It will be a variable type port. The expression column will contain decode (salary,salary\_prev,rank,rank+1)
  - f> rank\_o : It will be an output port containing the value of 'rank'.
3. Filter Transformation : It will have 2 I/O ports emp\_name and salary with a filter condition rank\_o = 5

The ports emp\_name and salary from Filter Transformation will be connected to target

## Scenario 6:

Let's say I have more than have record in source table and I have 3 destination table A,B,C. I have to insert first 1 to 10 records in A then 11 to 20 in B and 21 to 30 in C.  
Then again from 31 to 40 in A, 41 to 50 in B and 51 to 60 in C.....So on up to last record.

### Solution:

Generate sequence number using informatica, add filter or router transformations and define the conditions accordingly...

Define group condition as follows under router groups....

Group1 = mod(seq\_number,30) >= 1 and mod(seq\_number,30) <= 10

Group2 = mod(seq\_number,30) >= 11 and mod(seq\_number,30) <= 20

Group3 = (mod(seq\_number,30) >=21 and mod(seq\_number,30) <= 29 ) or mod(seq\_number,30) = 0

Connect Group1 to A, Group2 to B and Group3 to C

## Scenario 7:

Validation rules for connecting transformations in Informatica?

### Solution:

Some validation rules:

- 1-You can only link ports with compatible datatypes.
- 2-You cannot connect an active transformation and a passive transformation to the same downstream transformation.
- 3-You cannot connect more than one active transformation to the same downstream transformation or transformation input group. only way to do it using joiner with sorted ports.

## Scenario 8:

Source is a flat file and want to load unique and duplicate records separately into two separate targets; right??

### Solutions:

Here comes the solution -

SRC - SQ\_SRC - SRT - EXP - RTR - TGT

Try with the above and add logic of Sorter to identify duplicates, Use expression to mark the duplicates and finally Router to route to different targets.

## Scenario 9:

Input file

-----

10  
10  
10  
20  
20  
30

output file

-----

1  
2  
3  
1  
2  
1

scenario-it will count the no of records for example in this above case first 10 is there so it will count 1,den again 10 is there so it will be 2, when 20 comes it will be 1 again.

## Solution:

First import source, then use a sorter transformation. sort it by your column, then use an expression. In expression make this column, like this

1. column\_num(coming from sorter)
2. current\_num= check if column\_num=previous\_num,then add (first\_value +1),else 1
3. first\_value=current\_num.
4. previous\_num(new column)= column\_num

Pass current\_num to target.

## Scenario 10:

Input file

-----

10  
10  
10  
20  
20  
30

output file

-----

1  
2  
3  
4  
5  
6

## Solution:

<UMA> Sequence Generator can be used

## Scenario 11:

input file

-----

10  
10  
10

```

20
20
30
output file
----->
1
1
1
2
2
3

```

## Solution:

Sort => Expr (%10)=>Target

## Scenario 12:

There are 2 tables(input table)

table aa table bb

```

-----
id name id name
-- ----
101 ramesh 106 harish
102 shyam 103 hari
103 ---- 104 ram
104 ----

```

output file

```

-----
id name
-- ----
101 ramesh
102 shyam
103 hari
104 ram

```

## Solution:

One SQ => Exclude NULL value trx => Filter 106

Table bb is master and aa is detail table; If I do Master outer join, It will give common records from Master and additional records from detail – So output will be 101, 102, 103, 104

Use sorter and direct to target

**SQL: SELECT \* FROM BB LEFT OUTER JOIN AA ON (BB.ID = AA.ID)**

## Scenario 13:

```
table aa(input file)
```

```
-----  
id name
```

```
-- ----
```

```
10 aa
```

```
10 bb
```

```
10 cc
```

```
20 aa
```

```
20 bb
```

```
30 aa
```

```
Output
```

```
----
```

```
id name1 name2 name3
```

```
-----
```

```
10 aa bb cc
```

```
20 aa bb --
```

```
30 aa -- --
```

## Solution:

Use Sorter => EXPR (val=in\_name; out=if (out=val

## Scenario 14:

```
table aa(input file)
```

```
-----  
id name
```

```
-- ----
```

```
10 a
```

```
10 b
```

```
10 c
```

```
20 d
```

```
20 e
```

```
output
```

```
-----
```

```
id name
```

```
-- ----
```

```
10 abc
```

```
20 de
```



## Scenario 15:

**In the below scenario how can I split the row into multiple depending on date range?**

The source rows are as

ID Value from\_date(mm/dd) To\_date(mm/dd)

1 \$10 1/2 1/3

2 \$5 1/5 1/8

3 \$20 1/9 1/11

The target should be

ID Value Date

1 \$10 1/2

1 \$10 1/3

2 \$5 1/5

2 \$5 1/6

2 \$5 1/7

2 \$5 1/8

3 \$20 1/9

3 \$20 1/10

3 \$20 1/11

What is the informatica solution?

## Solution:

Use a Normalizer transformation with 3 ports ID, Value, Date. Set the 'Occurs' property of Normalizer to 2 for Date port and 1 for ID and Value ports. Now normalizer will be created with 4 input ports ID, Value, Date1 and Date2 and there will be 3 output ports ID, Value, Date. Connect from\_date to Date1 and To\_Date to Date2, and connect the rest of the matching ports. Connect the normalizer to your target.

Use datediff function to calculate no of days between dates.

Use that difference as the number of iteration .than use add\_to\_date function to increment date till number of iteration and load it to target.

## Scenario 16:

**How can the following be achieved in 1 single Informatica Mapping.**

\* If the Header table record has error value(NULL) then those records and the corresponding child records in the SUBHEAD and DETAIL tables should also not be loaded into the target(TARGET1,TARGET 2 or TARGET3).

\* If the HEADER table record is valid, but the SUBHEAD or DETAIL table record has an error value (NULL) then the no data should be loaded into the target TARGET1,TARGET 2 or TARGET3.

\* If the HEADER table record is valid and the SUBHEAD or DETAIL table record also has valid records only then the

data should be loaded into the target TARGET1,TARGET 2 and TARGET3.

```
=====
HEADER
COL1 COL2 COL3 COL5 COL6
1 ABC NULL NULL CITY1
2 XYZ 456 TUBE CITY2
3 GTD 564 PIN CITY3

SUBHEAD
COL1 COL2 COL3 COL5 COL6
1 1001 VAL3 748 543
1 1002 VAL4 33 22
1 1003 VAL6 23 11
2 2001 AAP1 334 443
2 2002 AAP2 44 22
3 3001 RAD2 NULL 33
3 3002 RAD3 NULL 234
3 3003 RAD4 83 31

DETAIL
COL1 COL2 COL3 COL5 COL6
1 D001 TXX2 748 543
1 D002 TXX3 33 22
1 D003 TXX4 23 11
2 D001 PXX2 56 224
2 D002 PXX3 666 332
=====

TARGET1
2 XYZ 456 TUBE CITY2

TARGET2
2 2001 AAP1 334 443
2 2002 AAP2 44 22

TARGET3
2 D001 PXX2 56 224
2 D002 PXX3 666 332
```

## Solution:

I don't know. Let us know if you know this.

## Scenario 17:

If i had source like unique & duplicate records like 1,1,2,3,3,4 then i want load unique records in one target like 2,4 and i want load duplicate records like 1,1,3,3

## Solution:

Source => SQ => Aggregator => Joiner => Router => Target1,2

## Scenario 18:

I Have 100 Records in a relational table and i want to load the record in 3 targets , first records goes to target 1 and second to target 2 and third to target 3 and so on ,what are the tx used in this.

### Solution:

- 1) From source qualifier get the records to the Expression.
- 2) Use one Sequence generator in which set the max value as 3, enable cycle option. Connect it to the expression.
- 3) Then use router & create 2 groups, 1st group condition as Next value = 1 another as next value = 2 and default.
- 4) These should be connected to the 3 target tables.

## Scenario 19:

There are three columns empid, salmonth, sal contains the values 101,jan,1000  
101 feb 1000 ...

like twelve rows are there then my required out put is like contains 13 columns empid jan feb march ..... dec and the values are 101 1000, 1000, 1000 etc

Make 13 columns and add expression transformation and as out and put the conditions like MAX(if(month='jan',sal)), same for other months.

## Scenario 20:

I have a source either file or db table

Eno ename sal dept

101 sri 100 1  
102 seeta 200 2  
103 lax 300 3  
104 ravam 76 1  
105 soorp 120 2

Want to run a session 3 times.

First time: it should populate dept 1

Second time: dept 2 only

Third time: dept 3 only

How can we do this?

### Solution:

Not sure how to do it.

## Scenario 21:

If I have a source as below:

EmployeeId, FamilyId, Contribution

1,A,10000

2,A,20000

3,A,10000

4,B,20

5,B,20

---

And my desired target is as below:

EmployeeId,Contribution(In Percent)

1,25%

2,50%

3,25%

4,50%

5,50%

---

**Explanation:** The contribution field in target is the individual Employee's share in a family's contribution. Say if total family contribution is 40000 then if A has contributed 10000 then target should have a value of 25%.

---

Can you please suggest me an approach to solve the specified problem?

## Solution:

Here goes the sql override

```
SELECT A.empid(B.contribution/A.BB)*100 AS CONTRIBUT FROM (SELECT empid SUM(contribution) OVER (PARTITION BY familyid) AS BB  
FROM table1) A (SELECT empid contribution FROM table1) B WHERE A.empid B.Empid
```

## Scenario 22:

Is it more Advantageous to use the Pre and Post SQL properties in the Workflow Designer task properties or in the Mapping Designer's? I am copying data from production to a staging table. As part of the process, I need to drop the indexes and triggers before the move, and recreate them after the move. What is the benefit (if any) of using the pre and post sql property in the WorkFlow rather than the Mapping?

## Solution:

It's good to go with Pre-SQL, Session, & Post SQL options with some modifications: (Never validated)

## Scenario 23:

I have a session which has a truncate target table option enabled. When the session fails for some reason the data is truncated in the table. How can I avoid truncating a target table in case of session failure?

### Solution:

#### Step - 1:

In the Pre-SQL:

Statement sequence would be --

1. Write savepoint statement like "SAVEPOINT STARTLOAD"
2. DELETE STATEMENT like "DELETE FROM TABLE TABLE\_NAME" .....

#### Step - 2:

At session level:

1. Uncheck the "Truncate table option"
2. Increase the commit interval to max value that informatica allows i.e, 2,147,483,647 as the commit interval considering that during that load you will have source records coming less than 2,147,483,647 record count.
3. Enable the option "**Rollback Transactions on Errors**" This will help to rollback the operations till the savepoint set in the Pre-SQL section.

#### Step -3:

In Post-SQL: Target Definition -

Write a Commit statement. "COMMIT". : This will help to commit the operations done on the target table after savepoint set in the Pre-SQL.

## Scenario 24:

How to pass one mapping variable / parameter to another with in the same workflow?

### Solution:

To pass a mapping variable or parameter value from one session to another in a workflow, do the following:

Create two consecutive sessions (session1 and session2) in a Workflow.

Create Workflow variable(s) in the Workflow.

In Session1, go to Edit >Components > Post-Session On Success Variable Assignment, assign values from mapping variables/parameter to workflow variables.

In Session2 Edit > Components > Pre-session Variable Assignment, assign values from workflow variables to mapping variables/parameter.

OR

from PowerCenter 8.6, there is an option to share the mapping variable to multiple sessions in the same workflow using the pre\_session\_variable\_assignment option and then create a workflow variable.

## Scenario 25:

In Informatica, what is the benefit apart from Performance of using more than one INTEGRATION SERVICES in a Domain

### Solution:

Load balancing and failover

## INFORMATICA INTERVIEW QUESTIONS

Please contact on [uma@etllabs.com](mailto:uma@etllabs.com) , [mahender@etllabs.com](mailto:mahender@etllabs.com) , [sadiq@etllabs.com](mailto:sadiq@etllabs.com) if you want answer for specific (It should be new , not present in PDF) Informatica Interview Questions. We do our jobs as well so we may take time to respond. It might take 10-15 days to respond to you. You can either use contact us form as well <http://etllabs.com/contact-us/>.