**1. What are the differences between connected lookup and unconnected lookup?**

|  |  |
| --- | --- |
| **Connected vs Unconnected Lookups** | |
| **Connected Lookup** | **Unconnected Lookup** |
| 1. It receives input from the pipeline & participates in the data flow. | 1. It receives input from the result of an LKP. |
| 2. It can use both, dynamic and static cache. | 2. It can’t be dynamic. |
| 3. It can return more than one column value i.e. output port. | 3. It can return only one column value. |
| 4. It caches all lookup columns. | 4. It caches only the lookup output ports in the return port & lookup conditions. |
| 5. It supports user-defined default values. | 5. It doesn’t support user-defined default values. |

**2. What is Lookup transformation?**

* Lookup transformation is used to look up a source, source qualifier, or target in order to get relevant data.
* It is used to look up a ‘flat file’, ‘relational table’, ‘view’ or ‘synonym’.
* Lookup can be configured as Active or Passive as well as Connected or Unconnected transformation.
* When the mapping contains the lookup transformation, the integration service queries the lookup data and compares it with lookup input port values. One can use multiple lookup transformations in a mapping.
* The lookup transformation is created with the following type of ports:
  + Input port (I)
  + Output port (O)
  + Look up Ports (L)
  + Return Port (R)

**3. How many input parameters can exist in an unconnected lookup?**

Any number of input parameters can exist. For instance, you can provide input parameters like column 1, column 2, column 3, and so on. But the return value would only be one.

**4. Name the different lookup cache(s)?**

Informatica lookups can be cached or un-cached (no cache). Cached lookups can be either static or dynamic. A lookup cache can also be divided as persistent or non-persistent based on whether Informatica retains the cache even after completing session run or if it deletes it.

* Static cache
* Dynamic cache
* Persistent cache
* Shared cache
* Recache

**5. Is ‘sorter’ an active or passive transformation?**

It is an active transformation because it removes the duplicates from the key and consequently changes the number of rows.

**6. What are the various types of transformation?**

* Aggregator transformation
* Expression transformation
* Filter transformation
* Joiner transformation
* Lookup transformation
* Normalizer transformation
* Rank transformation
* Router transformation
* Sequence generator transformation
* Stored procedure transformation
* Sorter transformation
* Update strategy transformation
* XML source qualifier transformation

**7. What is the difference between active and passive transformation?**

Active Transformation:- An active transformation can perform any of the following actions:

* Change the number of rows that pass through the transformation: For instance, the Filter transformation is active because it removes rows that do not meet the filter condition.
* Change the transaction boundary: For e.g., the Transaction Control transformation is active because it defines a commit or roll back transaction based on an expression evaluated for each row.
* Change the row type: For e.g., the Update Strategy transformation is active because it flags rows for insert, delete, update, or reject.

Passive Transformation: A passive transformation is one which will satisfy all these conditions:

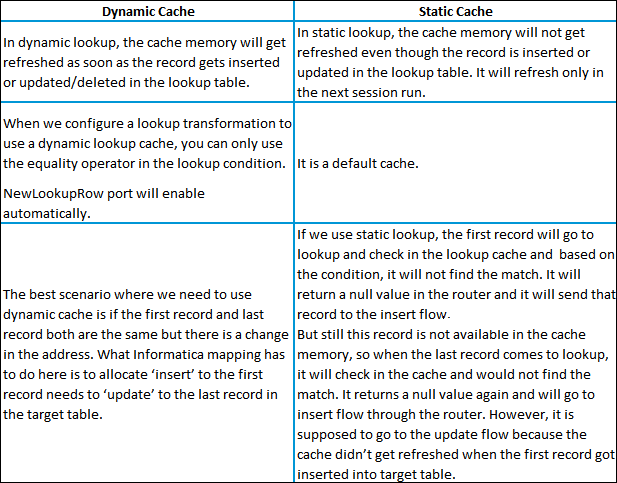
* Does not change the number of rows that pass through the transformation
* Maintains the transaction boundary
* Maintains the row type

**8. Name the output files created by Informatica server during session running.**

* Informatica server log: Informatica server (on UNIX) creates a log for all status and error messages (default name: pm.server.log). It also creates an error log for error messages. These files will be created in the Informaticahome directory.
* Session log file: Informatica server creates session log files for each session. It writes information about sessions into log files such as initialization process, creation of SQL commands for reader and writer threads, errors encountered and load summary. The amount of detail in the session log file depends on the tracing level that you set.
* Session detail file: This file contains load statistics for each target in mapping. Session detail includes information such as table name, number of rows written or rejected. You can view this file by double clicking on the session in the monitor window.
* Performance detail file: This file contains session performance details which tells you where performance can be improved. To generate this file, select the performance detail option in the session property sheet.
* Reject file: This file contains the rows of data that the writer does not write to targets.
* Control file: Informatica server creates a control file and a target file when you run a session that uses the external loader. The control file contains the information about the target flat file such as data format and loading instructions for the external loader.
* Post session email: Post session email allows you to automatically communicate information about a session run to designated recipients. You can create two different messages. One if the session completed successfully and another if the session fails.
* Indicator file: If you use the flat file as a target, you can configure the Informatica server to create an indicator file. For each target row, the indicator file contains a number to indicate whether the row was marked for insert, update, delete or reject.
* Output file: If a session writes to a target file, the Informatica server creates the target file based on file properties entered in the session property sheet.
* Cache files: When the Informatica server creates a memory cache, it also creates cache files. For the following circumstances, Informatica server creates index and data cache files.

**9. How do you differentiate dynamic cache from static cache?**

The differences are shown in the table below:

[](https://www.edureka.co/informatica)

**10. What are the types of groups in router transformation?**

* Input group
* Output group
* Default group

**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

[[](https://www.edureka.co/informatica)](https://www.edureka.co/informatica" \t "_blank)

**[Informatica Training & Certification](https://www.edureka.co/informatica" \t "_blank)**

* *[Instructor-led Sessions](https://www.edureka.co/informatica" \t "_blank)*
* *[Real-life Case Studies](https://www.edureka.co/informatica" \t "_blank)*
* *[Assignments](https://www.edureka.co/informatica" \t "_blank)*
* *[Lifetime Access](https://www.edureka.co/informatica" \t "_blank)*

[Explore Curriculum](https://www.edureka.co/informatica" \t "_blank)

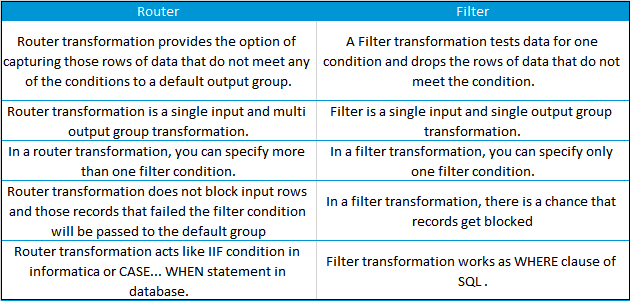
Save session log for these runs –> Change the number that you want to save the number of log files (Default is 0)

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.

**13. What are the similarities and differences between ROUTER and FILTER?**

The differences are:

[](https://www.edureka.co/informatica)

Advantages of Router transformation over Filter transformation:

* Better Performance; because in mapping, the Router transformation Informatica server processes the input data only once instead of as many times, as you have conditions in Filter transformation.
* Less complexity; because we use only one Router transformation instead of multiple Filter transformations.
* Router transformation is more efficient than Filter transformation.

For E.g.:

Imagine we have 3 departments in source and want to send these records into 3 tables. To achieve this, we require only one Router transformation. In case we want to get same result with Filter transformation then we require at least 3 Filter transformations.

Similarity:

A Router and Filter transformation are almost same because both transformations allow you to use a condition to test data.

**14. Why is sorter an active transformation?**

When the Sorter transformation is configured to treat output rows as distinct, it assigns all ports as part of the sort key. The integration service discards duplicate rows that were compared during the sort operation. The number of input rows will vary as compared to the output rows and hence it is an active transformation.

**15. When do you use SQL override in a lookup transformation?**

You should override the lookup query in the following circumstances:

1. Override the ORDER BY clause. Create the ORDER BY clause with fewer columns to increase performance. When you override the ORDER BY clause, you must suppress the generated ORDER BY clause with a comment notation.  
   Note: If you use pushdown optimization, you cannot override the ORDER BY clause or suppress the generated ORDER BY clause with a comment notation.
2. A lookup table name or column names contains a reserved word. If the table name or any column name in the lookup query contains a reserved word, you must ensure that they are enclosed in quotes.
3. Use parameters and variables. Use parameters and variables when you enter a lookup SQL override. Use any parameter or variable type that you can define in the parameter file. You can enter a parameter or variable within the SQL statement, or use a parameter or variable as the SQL query. For example, you can use a session parameter, $ParamMyLkpOverride, as the lookup SQL query, and set $ParamMyLkpOverride to the SQL statement in a parameter file. The designer cannot expand parameters and variables in the query override and does not validate it when you use a parameter or variable. The integration service expands the parameters and variables when you run the session.
4. A lookup column name contains a slash (/) character. When generating the default lookup query, the designer and integration service replace any slash character (/) in the lookup column name with an underscore character. To query lookup column names containing the slash character, override the default lookup query, replace the underscore characters with the slash character, and enclose the column name in double quotes.
5. Add a WHERE clause. Use a lookup SQL override to add a WHERE clause to the default SQL statement. You might want to use the WHERE clause to reduce the number of rows included in the cache. When you add a WHERE clause to a Lookup transformation using a dynamic cache, use a Filter transformation before the Lookup transformation to pass rows into the dynamic cache that match the WHERE clause.  
   Note: The session fails if you include large object ports in a WHERE clause.
6. Other. Use a lookup SQL override if you want to query lookup data from multiple lookups or if you want to modify the data queried from the lookup table before the Integration Service caches the lookup rows. For example, use TO\_CHAR to convert dates to strings.

**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.

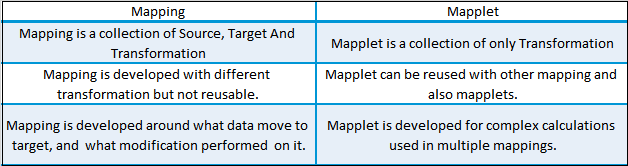
“Treat source rows as” property in session is set to “Data Driven” by default when using a update strategy transformation in a mapping.

**17. What are mapplets?**

•A Mapplet is a reusable object that we create in the Mapplet Designer.

•It contains a set of transformations and lets us reuse that transformation logic in multiple mappings.

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

[](https://www.edureka.co/informatica)

**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.

**20. What is the use of source qualifier?**

The source qualifier transformation is an active, connected transformation used to represent the rows that the integrations service reads when it runs a session. You need to connect the source qualifier transformation to the relational or flat file definition in a mapping. The source qualifier transformation converts the source data types to the Informatica native data types. So, you should not alter the data types of the ports in the source qualifier transformation.

The source qualifier transformation can be used to perform the following tasks:

**Data Warehousing and ETL Training**

**[INFORMATICA TRAINING & CERTIFICATION](https://www.edureka.co/informatica" \o "Informatica Training & Certification" \t "_blank)**

**[Informatica Training & Certification](https://www.edureka.co/informatica" \o "Informatica Training & Certification" \t "_blank)**

*[Reviews](https://www.edureka.co/informatica" \o "Informatica Training & Certification" \t "_blank)*

**[5](https://www.edureka.co/informatica" \o "Informatica Training & Certification" \t "_blank)**[(14149)](https://www.edureka.co/informatica" \o "Informatica Training & Certification" \t "_blank)

**[TALEND FOR DATA INTEGRATION AND BIG DATA](https://www.edureka.co/talend-for-big-data" \o "Talend for Data Integration and Big data" \t "_blank)**

**[Talend for Data Integration and Big data](https://www.edureka.co/talend-for-big-data" \o "Talend for Data Integration and Big data" \t "_blank)**

*[Reviews](https://www.edureka.co/talend-for-big-data" \o "Talend for Data Integration and Big data" \t "_blank)*

**[5](https://www.edureka.co/talend-for-big-data" \o "Talend for Data Integration and Big data" \t "_blank)**[(2071)](https://www.edureka.co/talend-for-big-data" \o "Talend for Data Integration and Big data" \t "_blank)

**[DATA WAREHOUSING AND BI CERTIFICATION TRAINING](https://www.edureka.co/data-warehousing-and-bi" \o "Data Warehousing and BI Certification Training" \t "_blank)**

**[Data Warehousing and BI Certification Training](https://www.edureka.co/data-warehousing-and-bi" \o "Data Warehousing and BI Certification Training" \t "_blank)**

*[Reviews](https://www.edureka.co/data-warehousing-and-bi" \o "Data Warehousing and BI Certification Training" \t "_blank)*

**[5](https://www.edureka.co/data-warehousing-and-bi" \o "Data Warehousing and BI Certification Training" \t "_blank)**[(4024)](https://www.edureka.co/data-warehousing-and-bi" \o "Data Warehousing and BI Certification Training" \t "_blank)

**[DATA WAREHOUSING CERTIFICATION TRAINING](https://www.edureka.co/datawarehousing-self-paced" \o "Data Warehousing Certification Training" \t "_blank)**

**[Data Warehousing Certification Training](https://www.edureka.co/datawarehousing-self-paced" \o "Data Warehousing Certification Training" \t "_blank)**

*[Reviews](https://www.edureka.co/datawarehousing-self-paced" \o "Data Warehousing Certification Training" \t "_blank)*

**[4](https://www.edureka.co/datawarehousing-self-paced" \o "Data Warehousing Certification Training" \t "_blank)**[(3743)](https://www.edureka.co/datawarehousing-self-paced" \o "Data Warehousing Certification Training" \t "_blank)

* Joins: You can join two or more tables from the same source database. By default, the sources are joined based on the primary key-foreign key relationships. This can be changed by explicitly specifying the join condition in the “user-defined join” property.
* Filter rows: You can filter the rows from the source database. The integration service adds a WHERE clause to the default query.
* Sorting input: You can sort the source data by specifying the number for sorted ports. The integration service adds an ORDER BY clause to the default SQL query
* Distinct rows: You can get distinct rows from the source by choosing the “Select Distinct” property. The integration service adds a SELECT DISTINCT statement to the default SQL query.
* Custom SQL Query: You can write your own SQL query to do calculations.

**21. What are the different ways to filter rows using Informatica transformations?**

* Source Qualifier
* Joiner
* Filter
* Router

**22. What are the different transformations where you can use a SQL override?**

* Source Qualifier
* Lookup
* Target

**23. Why is it that in some cases, SQL override is used?**

The Source Qualifier provides the SQL Query option to override the default query. You can enter any SQL statement supported by your source database. You might enter your own SELECT statement, or have the database perform aggregate calculations, or call a stored procedure or stored function to read the data and perform some tasks.

**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.

If you want to get hands-on learning on Informatica, you can also check out the tutorial given below. In this tutorial, you will learn about Informatica Architecture, Domain & Nodes in Informatica, and other related concepts.

**25. What is parallel processing in Informatica?**

After optimizing the session to its fullest, we can further improve performance by exploiting under utilized hardware power. This refers to parallel processing and we can achieve this in Informatica Powercenter using Partitioning Sessions.

The Informatica Powercenter Partitioning Option increases the performance of the Powercenter through parallel data processing. The Partitioning option will let you split the large data set into smaller subsets which can be processed in parallel to get a better session performance.

**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.

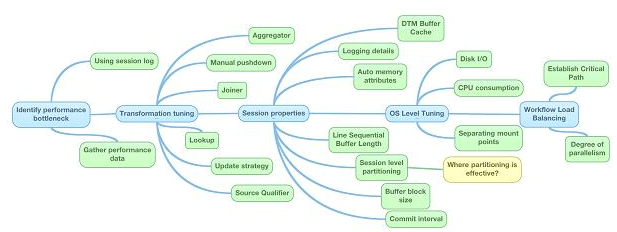
**Hash Auto-Keys Partitioning:** The Powercenter Server uses a hash function to group rows of data among partitions. When the hash auto-key partition is used, the Integration Service uses all grouped or sorted ports as a compound partition key. You can use hash auto-keys partitioning at or before Rank, 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?**



[Image Source](http://www.etltechblog.com/2012/05/tuning-sessions-for-better-performance.html)

**28. Mention a few design and development best practices for Informatica.**

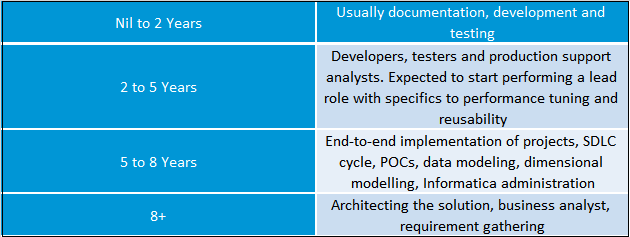
Mapping design tips:Standards – sticking to consistent standards is beneficial in the long run. This includes naming conventions, descriptions, environment settings, parameter files, documentation, among others.

* Reusability – in order to react quickly to potential changes, use Informatica components like mapplets, worklets, and reusable transformations.
* Scalability – when designing and developing mappings, it is a good practice to keep volumes in mind. This is caching, queries, partitioning, initial vs incremental loads.
* Simplicity – it is recommended to create multiple mappings instead of few complex ones. Use Staging Area and try to keep the processing logic as clear and simple as possible.
* Modularity – use the modular design technique (common error handling, reprocessing).

Mapping development best practices

* Source Qualifier – use shortcuts, extract only the necessary data, limit read of columns and rows on source. Try to use the default query options (User Defined Join, Filter) instead of using SQL Query override which may impact database resources and make unable to use partitioning and push-down.
* Expressions – use local variables to limit the amount of redundant calculations, avoid datatype conversions, reduce invoking external scripts (coding outside of Informatica), provide comments, use operators (||, +, /) instead of functions. Keep in mind that numeric operations are generally faster than string operations.
* Filter – use the Filter transformation as close to the source as possible. If multiple filters need to be applied, usually it’s more efficient to replace them with Router.
* Aggregator – use sorted input, also use as early (close to the source) as possible and filter the data before aggregating.
* Joiner – try to join the data in Source Qualifier wherever possible, and avoid outer joins. It is good practice to use a source with fewer rows, such as a Master source.
* Lookup – relational lookup should only return ports that meet the condition. Call Unconnected Lookup in expression (IIF). Replace large lookup tables with joins whenever possible. Review the database objects and add indexes to database columns when possible. Use Cache Calculator in session to eliminate paging in lookup cache.

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



**30. Explain shared cache and re cache.**

To answer this question, it is essential to understand persistence cache. If we are performing lookup on a table, it looks up all the data brings it inside the data cache. However, at the end of each session, the Informatica server deletes all the cache files. If you configure the lookup as a persistent cache, the server saves the lookup under an anonymous name. Shared cache allows you to use this cache in other mappings by directing it to an existing cache.

After a while, data in a table becomes old or redundant. In a scenario where new data enters the table, re cache ensures that the data is refreshed and updated in the existing and new cache.