ETL development – Pre requisites.

1. Setup Source and Target database.
2. Creation of ODBC connections.
   1. Creating source ODBC connection.
   2. Creating target ODBC connection.

ETL development process.

1. Creation of source metadata.
2. Creation of target metadata.
3. Design mapping without business rules.
4. Creating session for each mapping.
   1. Create reader connection (source).
   2. Create writer connection (Target).
   3. Create workflow.
   4. Run workflow.
   5. Monitoring ETL process. (view state).

What is ETL?

ETL stands for Extract-Transform-Load. ETL is the process of extracting the data from different source (Operational databases) systems, integrating the data and transforming the data into a homogeneous format and loading into the target warehouse database. Simply the overall process of ETL (Extraction, Transformation and Loading) is called Data Acquisition.

**Extraction** : Extraction is the process of reading the data from source databases into staging areas.

**Transformation** : Transformation is the process of converting the source data into required warehouse format.

**Loading** : Loading is the process of writing converted data from staging area into target warehouse systems.

What are the GUI based ETL tools?
1. Informatica.
2. DataStage.
3. Data Junction.
5. Cognos Decision Stream.

What are the programmatic based ETL tools?
1. PI/Sql.
2. SAS Base.
3. Tera ACCESS.
4. Tera Data Utilities.
   1. BTEQ.
   2. Fast Load.
   4. T(Trickle) Pump.

An Informatica PowerCenter is a GUI based ETL (Extract, Transform,Load) tool from Informatica Corporation which was founded in Redwood city, Los Angeles (1993). An Informatica PowerCenter is a client server process which allows you to design, rub, monitor and administer ETL applications known as ‘Mapping’. ETL.

ETL – Client – Server process.

A client is a graphical user interface which provides a design time environment or development environment where ETL applications are designed.

ETL applications.

ETL application logically defined Extraction, Transformation and Loading. An ETL application is designed with the following components.

Source (Extraction).

Business Rules (Transformation).

Targets (Loading).

What is Repository?

A repository is the brain of an ETL system that stores metadata which is required to perform ETL and ETL applications.

What is ETL Server?

It is an engine that provides a run time environment that executes ETL applications and it performs extraction, transformation and loading.

PowerCenter Integration Service (PRIS).
Informatica PowerCenter Integration service is an ETL engine or ETL server that performs Extraction, Transformation and Loading.

- It is a runtime environment where ETL objects are executed.
- Integration service reads metadata such as sessions, mappings from repository through repository service.
- It creates metadata such as log file, which is saved in Repository through Repository service.

**Setup source and target database.**

To setup target database, navigate to Start | Programs | Oracle | Application development | SQL plus.

- Logon to Oracle with the following user credentials.
  - username: system
  - password: Admin12345.
  - Host string: ORCL.

Click on OK button to login. Now we have to create target user in Oracle database, in this Informatica tutorial we are going to create user called ‘Batch10am’. Follow the steps given below to create ‘Batch10am’ user.

```
Creating new Database user
Enter user-name: system
Enter password:

Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - 64bit Production
With the Partitioning, OLAP, Data Mining and Real Application Testing options

SQL> Create user Batch10am identified by Target;
User created.

SQL> Grant DBA TO Batch10am;
Grant succeeded.

SQL> show user;
USER is "SYSTEM"

SQL> Conn Batch10am/ Target;
Connected.

SQL> Show user;
USER is "BATCH10AM"
```

As shown above, we have created new user in Oracle database. Now we have to create a table for the user ‘Batch10am’. Learn how to create a sample table in Oracle database.
Creating ODBC connection

What is ODBC?

ODBC means Open Database Connectivity. IT is a middle-ware or an interface that connects to various databases.

ODBC connection for Source and Target can be made from System DSN tab, navigate to Start | Settings | Control panel | Administrative tools | Data sources (ODBC) | System DSN | Click on add for new connection.

How to start Informatica PowerCenter 10.1.

Before going to ETL development process, first we have to start Informatica PowerCenter 10.1.0. To start navigate to Start | Settings | Control Panel | Administrative Tools | Services | Select Informatica 10.1.0 | Start (right click).

We can also start Informatica PowerCenter 10.1.0 service from Start | Programs | Informatica 10.1.0 | Server | Start Informatica services.

How to create folder in Informatica PowerCenter?

– Learn how to create folder in Informatica PowerCenter.

ETL development process.

A source metadata is created using Source analyzer tool from PowerCenter designer client component.

What is a Session?

A session is a task that executes a Mapping by sending instructions to server what to do and server will follow session instructions. Session can be created using task developer tool from Informatica PowerCenter Workflow manager client component.
### Informatica Introduction
- Informatica Tutorial - Home
- What is Informatica?
- Informatica - Architecture.
- Informatica - What is PowerCenter Integration Service?
- Informatica - PowerCenter Repository Service
- Informatica - PowerCenter 10.0.1 Installation Step by Step
- Informatica - Creating Integration Service
- Informatica - Data Transformation Manager (DTM)
- Informatica - DTM Threads
- Informatica - Load Balancing
- Informatica Server 9.0 installation Step by Step.
- How to add Repository and Connect

### Informatica Transformations
- Introduction to Informatica transformations
- Aggregator Transformation
- Filter Transformation
- Lookup Transformation
- Rank Transformation
- Router Transformation
- Joiner Transformation
- Expression Transformation
- Normalizer Transformation
- SQL Transformation
- Source Qualifier Transformation
- Sequence Transformation
- Sorter Transformation
- Stored Transformation
- Transaction Control Transformation
- Union Transformation
- Update Strategy Transformation
- XML Generator Transformation
- XML Parser Transformation