

# ETL and Informatica



# Objectives

- **At the end of this module, the learners will be able to:**
  - Define ETL Terms
  - Explain ETL process flow
  - Outline Informatica overview
  - List Informatica key features
  - Distinguish server side and client side components

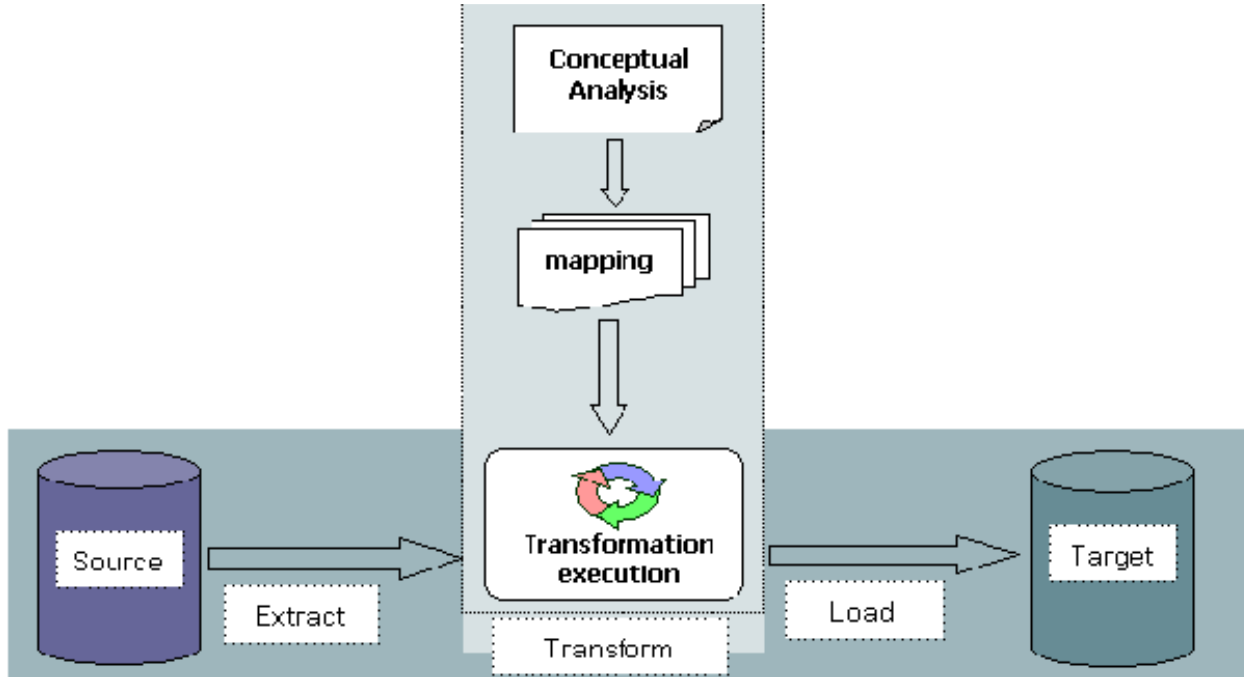
# ETL Terms

- Source Systems
- Mapping
- Metadata
- Staging area
- Data Cleansing
- Transformation
- Target System

# Pick and Answer

- What is the necessity of a staging area?
- What is meant by Mapping?
- Define Data Cleansing
- Define Transformation

# ETL Process Flow



# Informatica: An Overview

- Provides a single enterprise data integration platform
- Can integrate data to provide business users holistic access to enterprise data
- Scale and respond to business needs for information
- Simplify design, collaboration, and re-use to reduce developers' time to results

# Informatica: Key Features

- Enterprise-level data integration
- Security
- Visual Interface
- Developer productivity
- Compatibility

# Informatica Components

- Integration Service
- Repository Service
  - Repository Server Administration Console
- Web Services Hub
  - Repository Web Designer
- SAP BW service
  - Workflow Manager
  - Workflow Designer

## • Server components

## • Client components



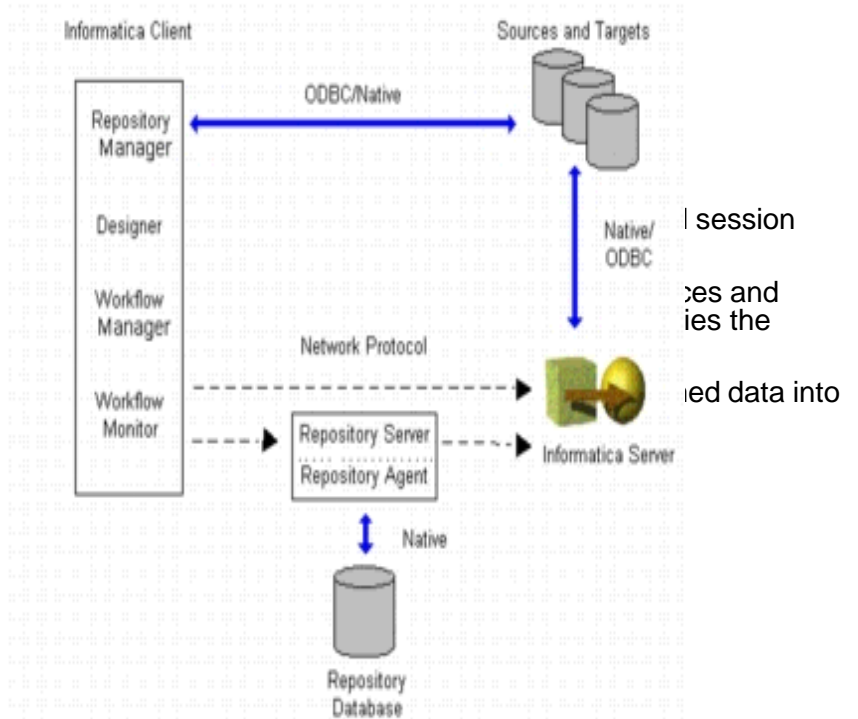
# PowerCenter Repository

- Stores the metadata
- Repository Manager creates the metadata tables in the database
- Tasks in the Power Center Client application such as creating users, analyzing sources, developing mappings or Mapplets, or creating sessions creates metadata
- Integration Service reads metadata created in the Client application when a session runs
- Global and local repositories can be created to share metadata
- Exchange of Metadata with other BI Tools
- MX Views are relational views which allow easy access to metadata information using queries written in SQL

# PowerCenter Client Tools

- Repository Manager
  - Use the Repository Manager to assign permissions to users and groups and manage folders
  - Manage user and group permissions
  - Perform folder functions
  - View metadata
- Designer
  - Add source and target definitions to the repository
  - Create mappings that contain data transformation instructions
- Workflow Manager & Workflow Monitor
  - Create, schedule, execute and monitor workflows

# Overview - Integration Service



# Overview – Sources/ Target

- Relational
- File
- Application
- Mainframe
- Other sources

# Quick Recap

- ETL stands for extraction, transformation and loading
- Informatica provides a single enterprise data integration platform
- With Informatica, you can achieve
  - Enterprise level data integration
  - Scalability
  - Developer productivity
- Client tools of Informatica includes
- Designer
  - Workflow Manager
  - Workflow Monitor
  - Repository Manager
  - Repository Administration Console

# Summary

In this module, we learnt to:

- Define ETL Terms
- Explain ETL process flow
- Outline Informatica overview
- List Informatica key features
- Distinguish server side and client side components



Thank You