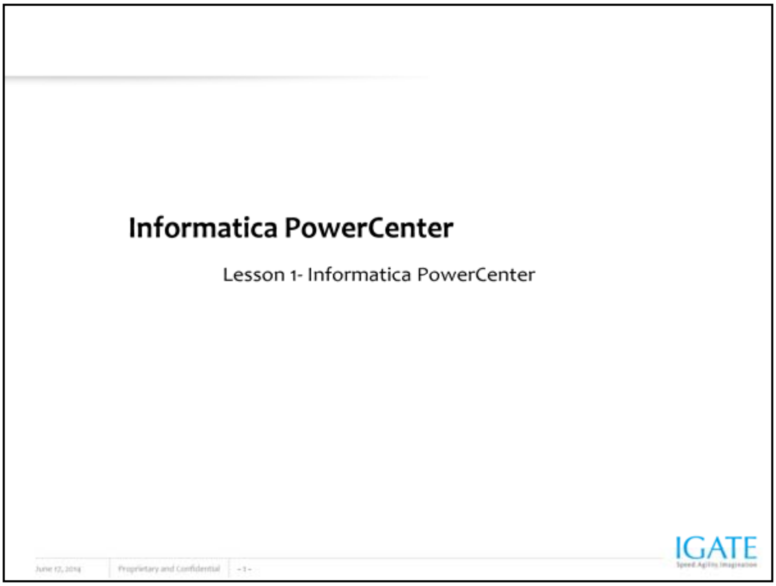
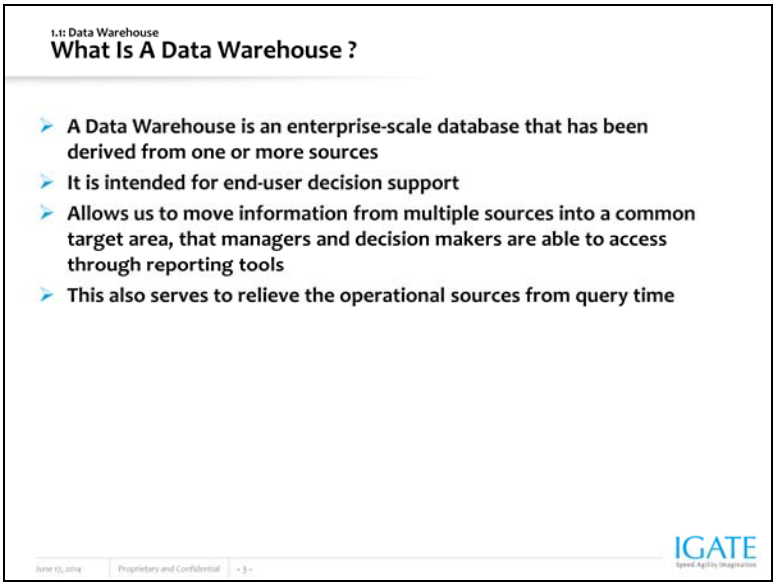

Informatica PowerCenter
Lesson 1- Informatica PowerCenter
IGATE

Lesson Objectives

- **In this lesson you will learn about:**
  - What is a Data Warehouse ?
  - Overview of Data Warehouse Model
  - ETL Process
  - Informatica PowerCenter Components
  - PowerCenter Domain
  - Integration Service
  - Repository Manager
  - User and Group Management
  - Folder



A data warehouse is a decision support database that is maintained separately from an organization’s operational databases. It usually resides on a dedicated server. This database is designed based on what kind of information a company is seeking (e.g., sales, marketing, healthcare membership and providers, etc.). A data warehouse is “the place” for top executives, managers, analysts, and other end-users to mine a rich source of company information. They can ask compelling business questions and find answers in their data, and make key and timely business decisions from their desktops using GUI based On-line Analytical Processing (OLAP) tools.

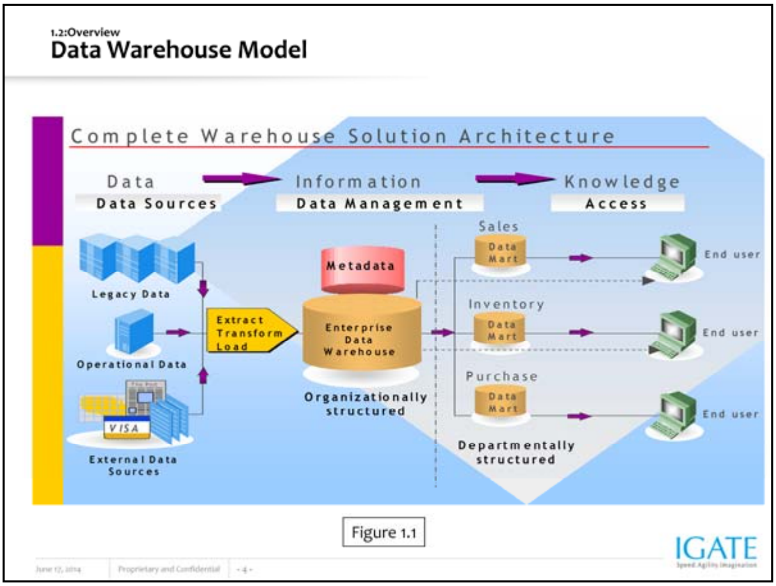
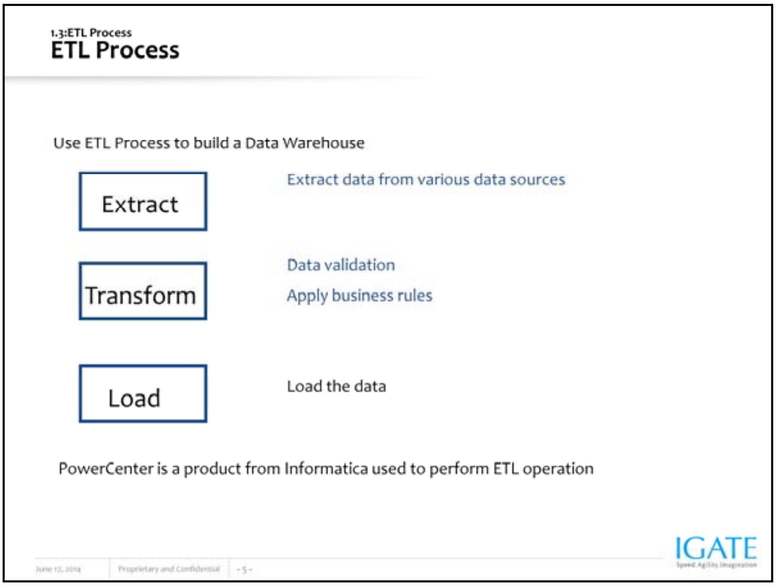


Figure 1.1

The above Figure 1.1. shows a simplified data warehouse model. The developer puts business rules for data transformation into the metadata repository. There is a transformation process that extracts data from the operational sources like database, legacy systems, and flat files. It transforms data according to the business rules, and loads the data into the data warehouse. From the data warehouse, atomic data flows to various departments for their customized usage. These departmental databases are called *data marts*.

The end user will use a query tool to look at the metadata in the repository and the transformed data in the warehouse. The data in the data warehouse will be used to cater to dynamic and complex reporting requirements.

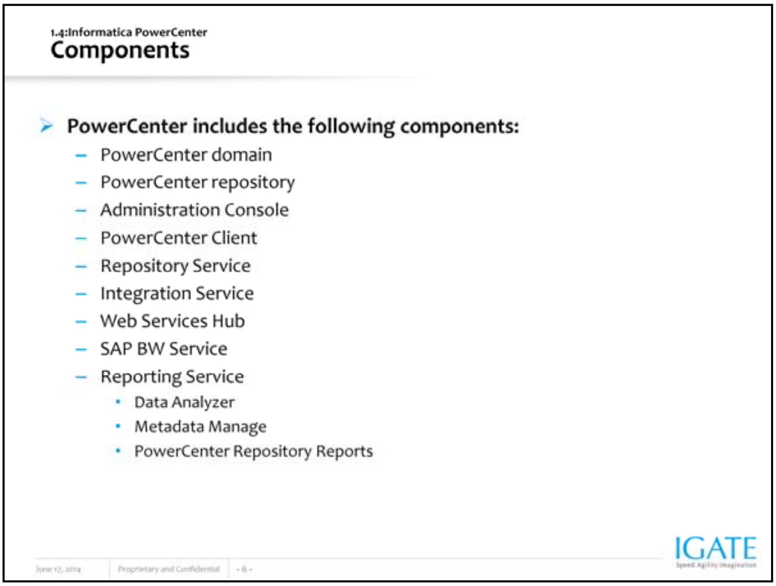
When Informatica's PowerCenter is fit into the above model, the developer can use PowerCenter Designer tool to develop mappings and other metadata for the Informatica repository.



To populate the data in the data warehouse, the steps are:

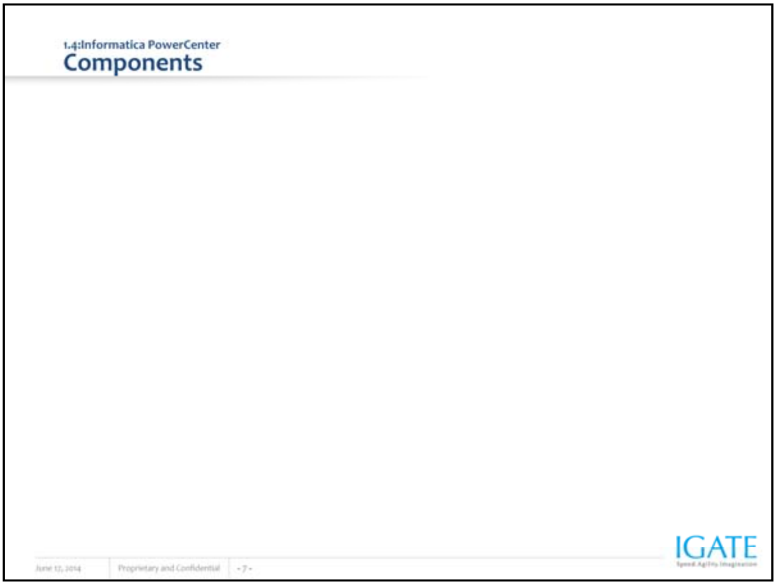
- **Extract -** Extracting appropriate data from existing operational database(s), flat files etc
- **Transform -** Cleansing or scrubbing the data, aggregating, denormalizing and filtering the data
- **Load -** Loading the data into the database

This data population process is also known as the *data transformation* process. Moving data from the operational databases to the data warehouse needs to be done via extraction tools. Operational data needs to be mapped to the target data warehouse. As a part of the data movement, data transformation is performed as specified by the meta data rules developed during the data modeling stage. **PowerCenter** is a product from Informatica used to perform ETL operation.



PowerCenter provides an environment that allows to load data into a centralized location, such as a datamart, data warehouse, or operational data store (ODS). Data can be extracted from multiple sources, transformed according to business logic built in the client application, and load the transformed data into file and relational targets. Informatica provides the following integrated components:

- **PowerCenter domain.** The Power Center domain is the primary unit for management and administration within PowerCenter. The Service Manager runs on a PowerCenter domain. The Service Manager supports the domain and the application services.
- **PowerCenter repository.** The PowerCenter repository resides in a relational database. The repository database tables contain the instructions required to extract, transform, and load data
- **Administration Console.** The Administration Console is a web application that you use to administer the PowerCenter domain and PowerCenter security
- **PowerCenter Client.** The PowerCenter Client is an application used to define sources and targets, build mappings and mapplets with the transformation logic, and create workflows to run the mapping logic. The PowerCenter Client connects to the repository through the Repository Service to modify repository metadata.
- **Repository Service.** The Repository Service accepts requests from the PowerCenter Client to create and modify repository metadata and accepts requests from the Integration Service for metadata when a workflow runs
- **Integration Service.** The Integration Service extracts data from sources and loads data to targets



• **Web Services Hub.** Web Services Hub is a gateway that exposes PowerCenter functionality to external clients through web services

• **SAP BW Service**. The SAP BW Service extracts data from and loads data to SAP BW. If you use the PowerExchange for SAP NetWeaver BW Option, you must create and enable a SAP BW Service in the PowerCenter domain

• **Reporting Service.** The Reporting Service runs the Data Analyzer web application. Data Analyzer provides a framework for creating and running custom reports and dashboards. You can use Data Analyzer to run the metadata reports provided with PowerCenter, including the PowerCenter Repository Reports and Data Profiling Reports. Data Analyzer stores the data source schemas and report metadata in the Data Analyzer repository

   - **Data Analyzer.** Data Analyzer provides a framework to perform business analytics on corporate data. With Data Analyzer, you can extract, filter, format, and analyze corporate information from data stored in a data warehouse, operational data store, or other data storage models
   - **Metadata Manager Service.** The Metadata Manager Service runs the Metadata Manager web application. You can use Metadata Manager to browse and analyze metadata from disparate metadata repositories. Metadata Manager helps you understand and manage how information and processes are derived, how they are related, and how they are used. Metadata Manager stores information about the metadata to be analyzed in the Metadata Manager repository
   - **PowerCenter Repository Reports.** PowerCenter Repository Reports are a set of prepackaged Data Analyzer reports and dashboards to help you analyze and manage PowerCenter metadata

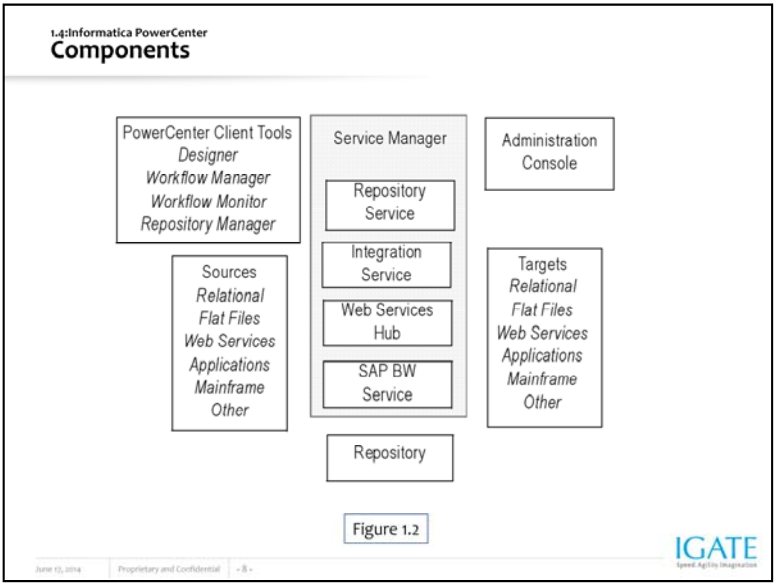
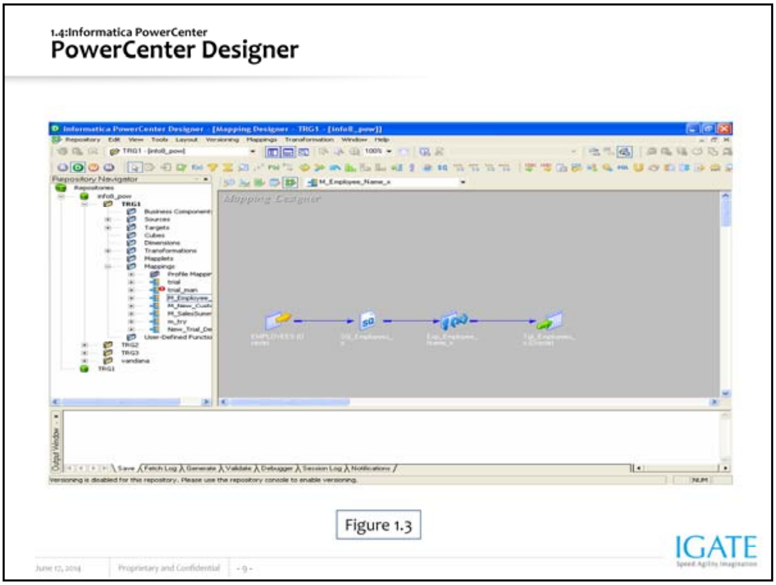


Figure 1.2

<u>Sources</u>
PowerCenter accesses the following sources:
Relational: Oracle, Sybase ASE, Informix, IBM DB2, Microsoft SQL Server, and Teradata.
File: Fixed and delimited flat file, COBOL file, XML file, and web log.
Application: Business sources such as Hyperion Essbase, WebSphere MQ, IBM DB2 OLAP Server, JMS, Microsoft Message Queue, PeopleSoft, SAP NetWeaver, SAS, Siebel, TIBCO, and webMethods.
Mainframe: Mainframe databases such as Adabas, Datacom, IBM DB2 OS/390, IBM DB2 OS/400, IDMS, IDMS-X, IMS, and VSAM.
Other: Microsoft Excel, Microsoft Access, and external web services.

<u>Targets</u>
PowerCenter can load data into the following targets:
Relational: Oracle, Sybase ASE, Sybase IQ, Informix, IBM DB2, Microsoft SQL Server, and Teradata.
File: Fixed and delimited flat file and XML.
Application: Business sources such as Hyperion Essbase, WebSphere MQ, IBM DB2 OLAP Server, JMS, Microsoft Message Queue, mySAP, PeopleSoft EPM, SAP BW, SAS, Siebel, TIBCO, and webMethods.
Mainframe: Mainframe databases such as IBM DB2 for z/OS, IMS, and VSAM.
Other: Microsoft Access and external web services.

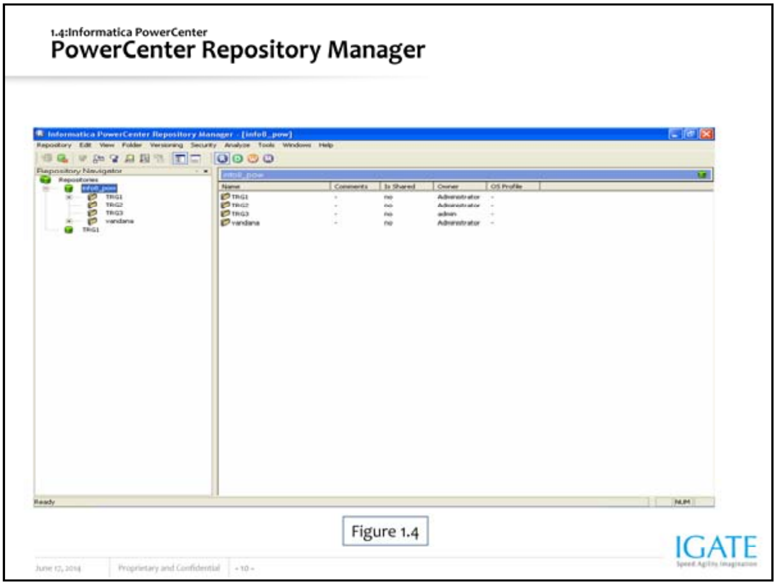


The PowerCenter Designer has five tools to build mappings and mapplets so as to specify how to move and transform data between sources and targets.
The five tools are as follows:

- Source Analyzer
- Target designer
- Transformation Developer
- Mapplet Designer
- Mapping Designer

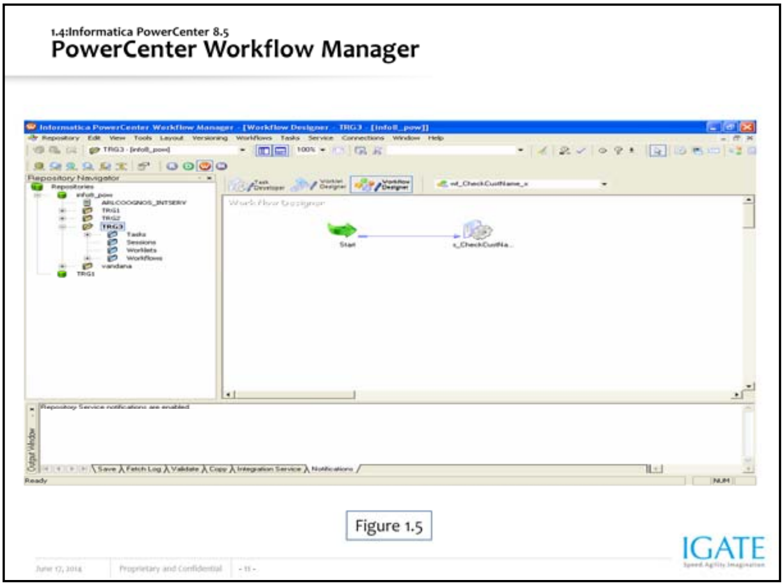
The Designer helps create source definitions, target definitions, and transformations to build the mappings.

The Designer allows to work with multiple tools at one time and to work in multiple folders and repositories at the same time. It also includes windows so as to view folders, repository objects, and Tasks.



The Repository Manager client tool allows to navigate through multiple folders and repositories and perform basic repository Tasks. The Repository Manager can be used to:

- Add a repository
- Remove a repository
- Connect to a repository
- Search for repository objects, etc.



The Workflow Manager is used to define a set of instructions called a *Workflow* to execute mappings built in the Designer. Generally, a Workflow contains a session and any other Task to be performed when a session is executed.

1.4:Informatica PowerCenter

PowerCenter Workflow Monitor

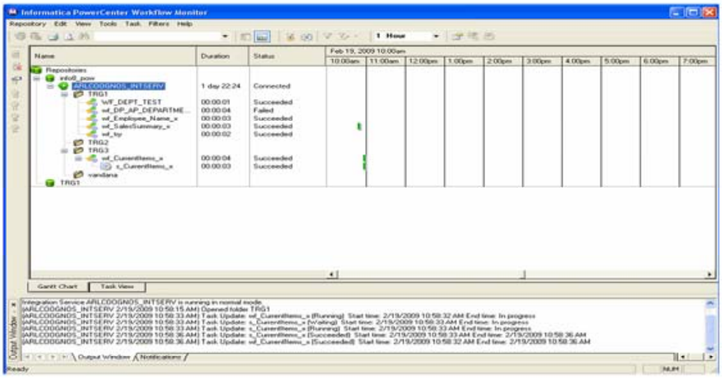


Figure 1.6

The Workflow Monitor is a tool that allows to monitor data loading process.

1.4:Informatica PowerCenter

PowerCenter Administration Console

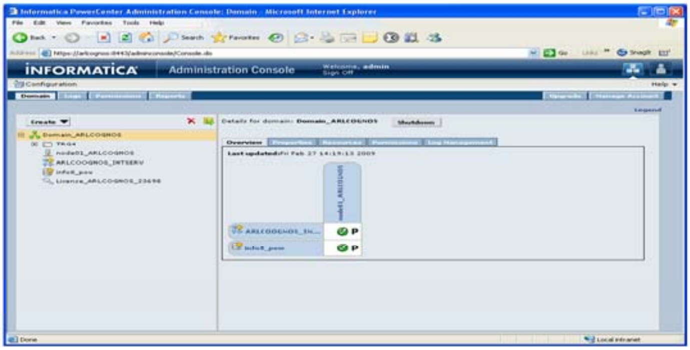


Figure 1.7

The PowerCenter Administration Console is the administration tool you use to administer the PowerCenter domain and PowerCenter security. Use the Administration Console to perform the following tasks:

Domain administrative tasks. Manage logs, domain objects, user permissions, and domain reports. Domain objects include services, nodes, grids, folders, and licenses.

Security administrative tasks. Manage users, groups, roles, and privileges.

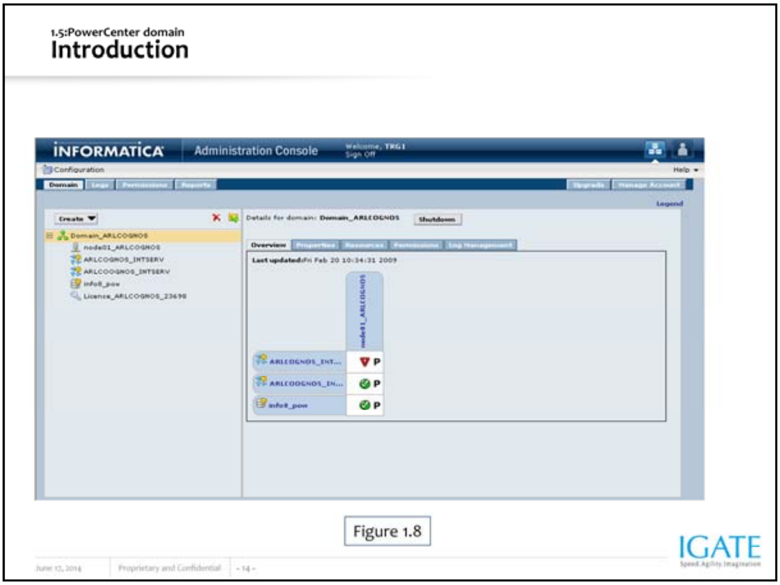


Figure 1.8

A domain is the primary unit for management and administration of services in PowerCenter. PowerCenter provides the PowerCenter domain to support the administration of the PowerCenter services.

1.6:Integration Service

Introduction

- The Integration Service moves data from sources to targets based on workflow and mapping metadata stored in a repository.
- When a workflow starts, the Integration Service retrieves mapping, workflow, and session metadata from the repository.
- It extracts data from the mapping sources and stores the data in memory while it applies the transformation rules configured in the mapping.

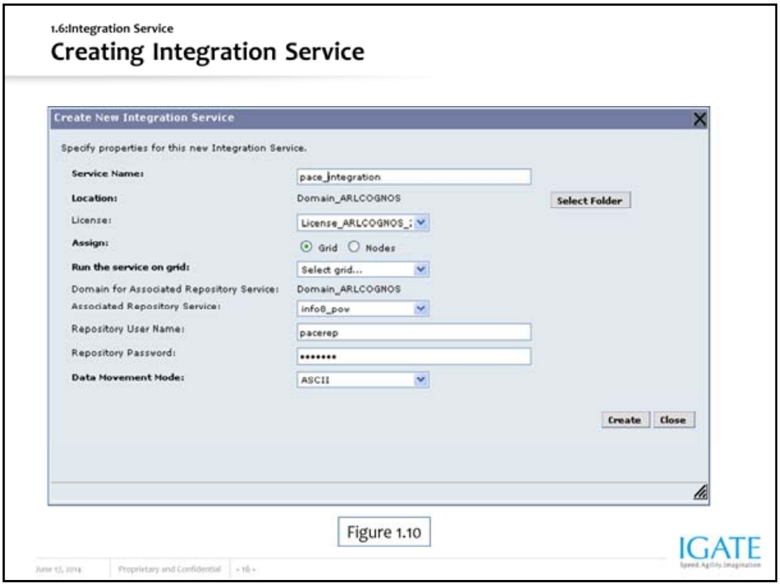


Figure 1.10

Steps to Create Integration Service:

• In the Administration Console, click ***Create > Integration Service.***

The Create New Integration Service dialog box appears.

• Enter values for the following Integration Service options ***Service Name, Location, License, Assign, Run the Service on Grid, Primary Node, Backup Nodes, Domain for Associated Repository Service, Associated Repository Service, Repository User Name, Repository Password, Security Domain, Data Movement Mode***

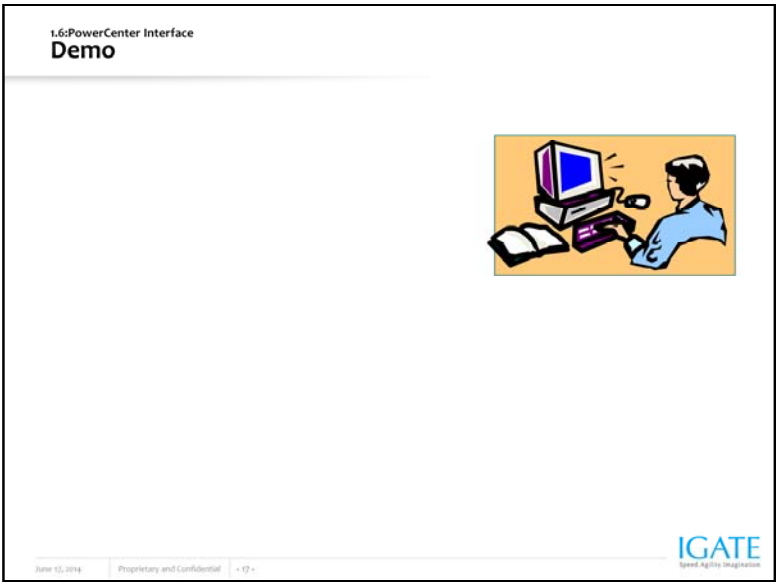
• Click OK.

If you do not specify an associated repository, the following message appears:

No Repository Service is associated with this Integration Service. Select an associated Repository Service to view a list of available codepages.

You cannot enable the Integration Service until you assign the code page for each Integration Service process node.

• Click OK.



PowerCenter Graphical User Interface

1.7:Repository Manager

Creating Repository Service

- Open PowerCenter Administration Console and then click on Domain->Create from the drop down list select Repository Service

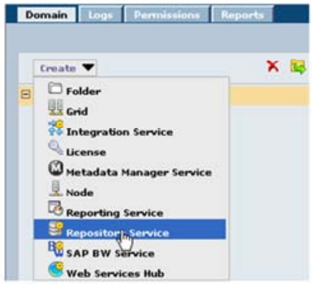


Figure 1.12

1.7:Repository Manager

Creating Repository Service

- Enter the properties of the Repository Service

Figure 1.13

Following are the steps to create a repository:

1. In the Navigator of the Administration Console, select the ***folder*** where you want to create the Repository Service.

   Note: If you do not select a folder, you can move the Repository Service into a folder after you create it.

2. Click ***Create > Repository Service.***

   The Create New Repository Service dialog box appears.

3. Enter values for the following Repository Service options:

Service Name, Location, License, Node, Database Type, CodePage, ConnectString, DBUser, DBPassword, TableSpaceName, Creation Mode

4. If you create a Repository Service for a repository with existing content and the repository existed in a

different PowerCenter domain, verify that users and groups with privileges for the Repository Service exist in the current domain.

5. Click OK.

The created repository service will be seen in the folder. Then select the Repository Service in Navigator window and then the repository service can be enabled or disabled using ***Enable*** or ***Disable***

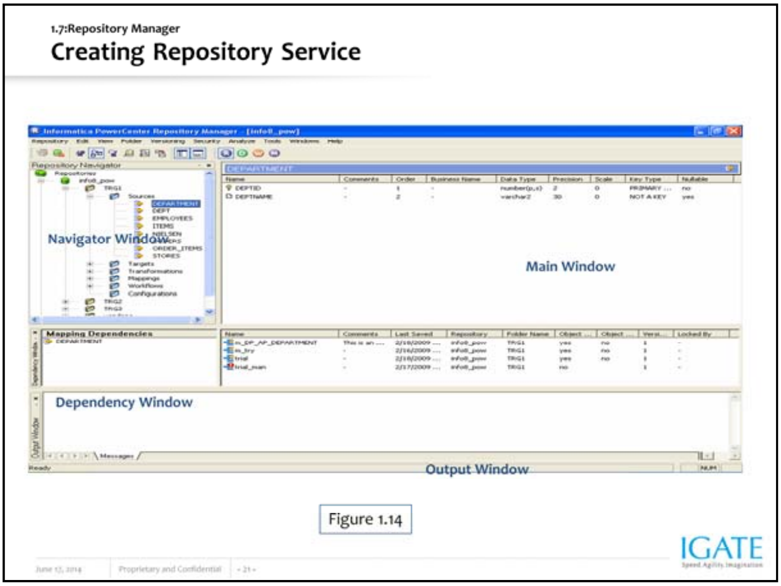


Figure 1.14

The Repository Manager allows to perform basic repository tasks. It has the following four windows:

- **Navigator Window -** Displays all the objects that are created in the Repository Manager, the Designer, and the Workflow Manager. It is organized first by repository, then by folder.
- **Main Window -** Provides properties of the object selected in the Navigator window. The columns in this window will change depending upon the object selected in the Navigator window.
- **Dependency Window -** Shows dependencies on sources, targets, mappings, and shortcuts for objects selected in either the Navigator or Main window
- **Output Window -** Provides the output of procedures executed within Repository Manager

Other Tasks that can be performed within the Repository Manager include:

- Create folders and edit folder properties
- Establish users, groups, privileges, and edit folder permissions
- Search metadata and perform dependency analyses
- View locks and unlock objects

Demo

Steps to connect to the repository using Repository Manager for the first time:

1. Launch the Repository Manager tool, choose Start | Programs | Informatica PowerCenter | Client | Repository Manager.
3. Click on the ***Add Repository*** menu option.
4. Enter the following information:
   - Repository name
   - Username
5. Choose ***Repository* | *Connect*** or right click on the repository name and choose ***Connect***.
6. The ***Connect to Repository*** dialog box appears.
7. Enter the repository username and password.
8. Click ***More***.
9. The Connect to Repository dialog box expands.
10. Enter the host name of the machine the Repository Server is running on and the port number the Repository Server uses for connections.
11. Click ***Connect***.

The Repository Server now opens a connection to the database, and a new icon representing the repository appears in the Repository Manager. Folders within the repository now appear underneath the icon for that repository.

Steps to connect to a repository that has been accessed before:

1. Launch the Repository Manager tool, choose Start | Programs | Informatica PowerCenter | Informatica PowerCenter Client | Repository Manager.
2. Select the icon for the repository and click the ***Connect*** button on the toolbar.
3. Alternatively, the repository icon can be selected, choose Repository | Connect.
4. Enter the repository username and password.
5. Click ***Connect***.

Follow the same steps to connect to the repository using PowerCenter Designer, Workflow Manager, and Workflow Monitor.

1.8:User and Group Management

Introduction

- To access the services and objects in the PowerCenter domain and to use the PowerCenter applications, you must have a user account.
- The tasks you can perform depend on the type of user account you have.
- User account management in PowerCenter involves the following key components:
  - Users. You can set up different types of user accounts in the PowerCenter domain.

1.8:User and Group Management

Introduction

- Groups. You can set up groups of users and assign different roles, privileges, and permissions to each group.
- Privileges and Roles. Privileges determine the actions that users can perform in PowerCenter applications.
- A role is a collection of privileges that you can assign to users and groups.
- Authentication. When a user logs in to PowerCenter, the Service Manager authenticates the user account in the PowerCenter domain and verifies that the user can use the PowerCenter applications.

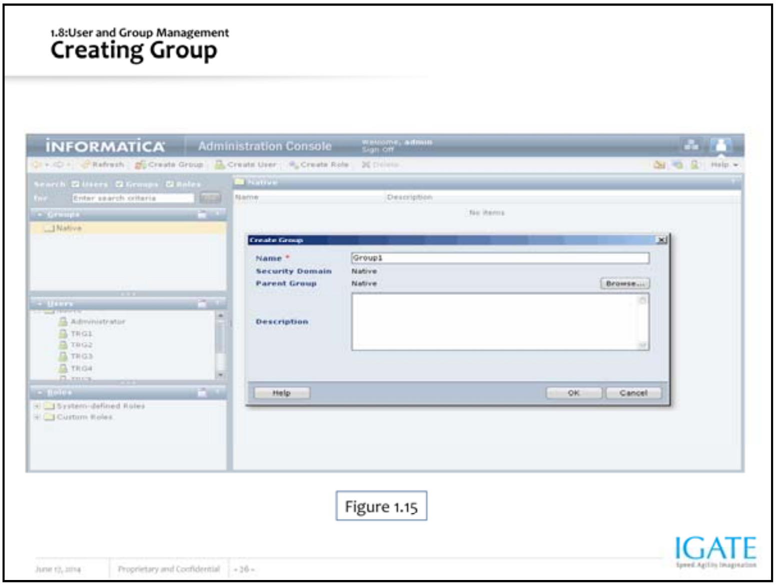


Figure 1.15

The steps for creating a group are as follows:

1. In the Administration Console Screen, Select ***Configure Security***
2. Select ***Create Group***
3. Enter Name and Description
4. ***Browse*** and select the Group in case of creating sub grouping

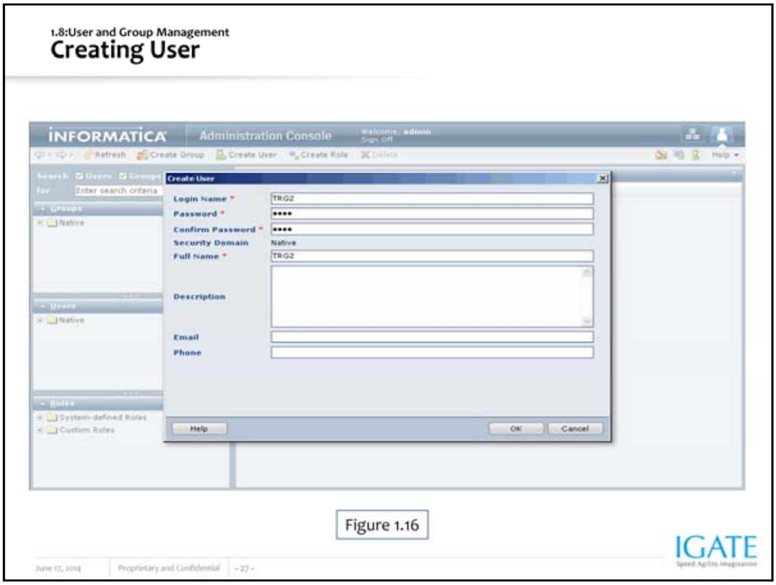


Figure 1.16

The steps for creating a user are as follows:

1. In the Administration Console Screen, Select ***Configure Security***
2. Select ***Create User***
3. Enter ***Name, Password, Full Name, Description, Email, Phone***
4. Click OK

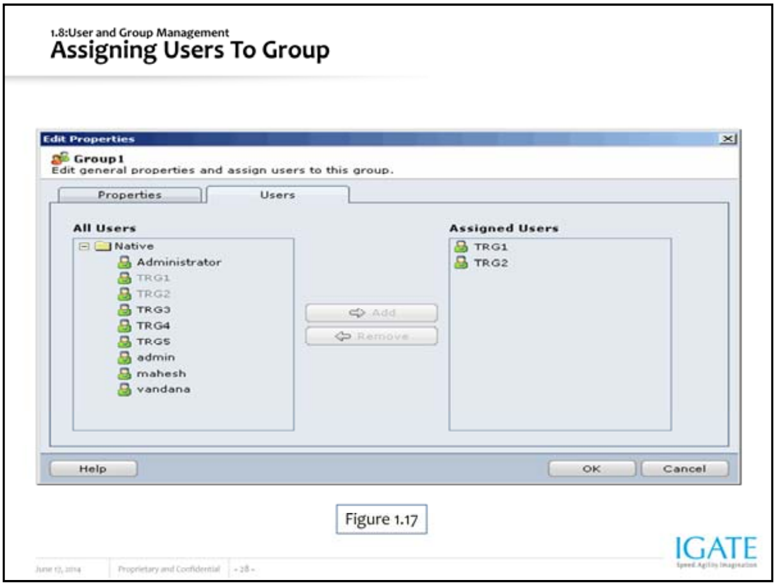


Figure 1.17

The steps for Assigning the Users to the Group are as follows:
• Select the Group from Navigator Window
• ***Overview|Edit |Users***
• Add the users to the group

1.8:User and Group Management

Assigning Roles And Privileges

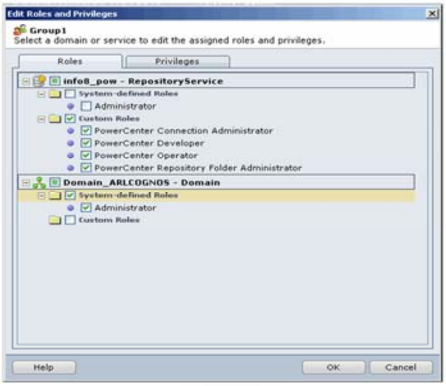


Figure 1.18

June 17, 2014 | Proprietary and Confidential | - 29 -

The steps for Assigning the Roles to the Group are as follows:

1. Select Group from Navigator Window
2. ***Overview|Edit |Roles***
3. Select the Appropriate Roles from:
   - ***System-Defined Roles***
   - ***Custom Roles***
4. Press OK

The steps for Assigning the Privileges to the Group are as follows:

1. Select Group from Navigator Window
2. ***Overview|Edit |Privileges***
3. Select the appropriate privileges from:
   - ***Tools:*** Includes the privilege to log in to the Administration Console.
   - ***Security Administration:*** Includes privileges to manage users, groups, roles, and privileges.
   - ***Domain Administration:*** Includes privileges to manage the domain, folders, nodes, grids, licenses, and application services.
4. Press OK

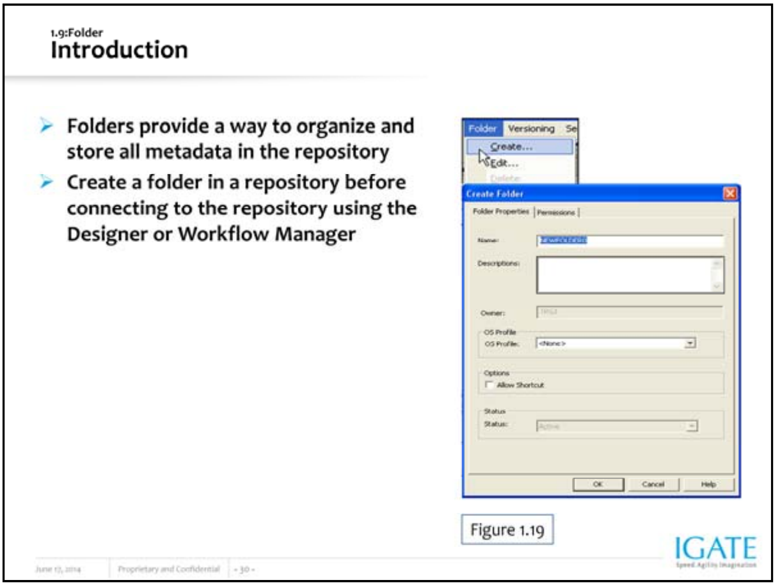


Figure 1.19

Folders provide a way to organize and store all metadata in the repository, including mappings, schemas, and sessions. Folders are designed to be flexible, to logically organize the repository. Each folder has a set of configurable properties that helps to define how users access the folder. For example, create a folder that allows all repository users to see objects within the folder, but not to edit them. Or, create a folder that allows users to share objects within the folder.

Folder Attributes

- FOLDER OWNER - User who serves as focal point to folder permissions
- PERMISSIONS - Rights to read, write, and/or execute objects in a folder
- SHARED - Property that allows users to make shortcuts in a folder
- SHORTCUT - A dynamic link to an object stored in a shared folder
- VERSIONS - Folder iterations that indicate development stages

A folder can be designated as shared by checking '*Allow Shortcuts*' in the Folder Properties. When a folder is shared, users can create shortcuts to objects in the folder, thereby sharing the metadata in the folder. Once a folder is designated as shared, it cannot be reverted back to make it non-shared. If a folder is not shared, users will only be able to make copies of objects in the folder.

Folders can be versioned, allowing to keep copies at different stages of development. Versioning provides the flexibility to:

- Revert back to earlier copy of an object
- Re-create an object if one is accidentally deleted
- Save a functioning copy while developing and testing

Folders are created and maintained by the Repository Manager. In the Repository Manager, the users accessing folders, the objects and versions within each folder can be seen.

In the Designer, use folders to store sources, transformations, mapplets, targets, and mappings.

In the Workflow Manager, use folders to store Workflows, Tasks, and sessions. When creating a Workflow, include any session or Task in the folder.

Note: Create a folder in a repository before connecting to the repository using the Designer or Workflow Manager.

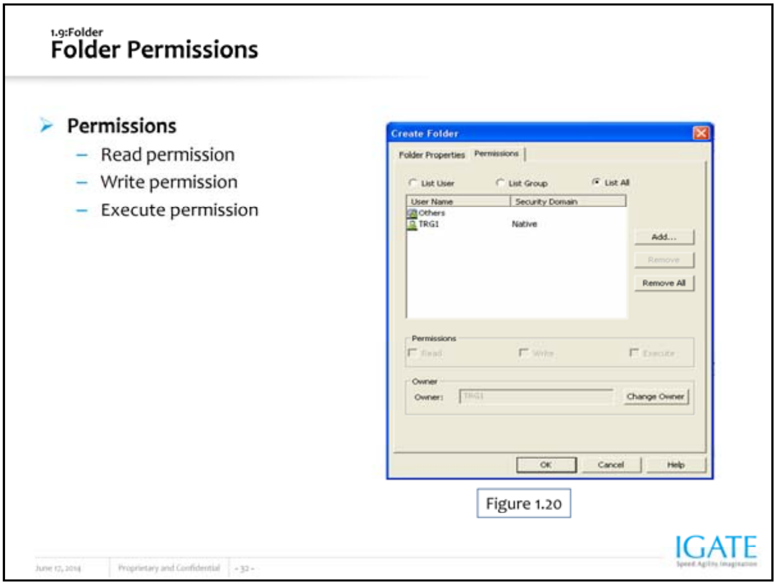


Figure 1.20

Permissions
Allows repository users to perform Tasks within a folder. To perform Tasks within a folder having the privilege to perform the Task in the repository as well as the applicable folder permission is a must. The available folder permissions are:

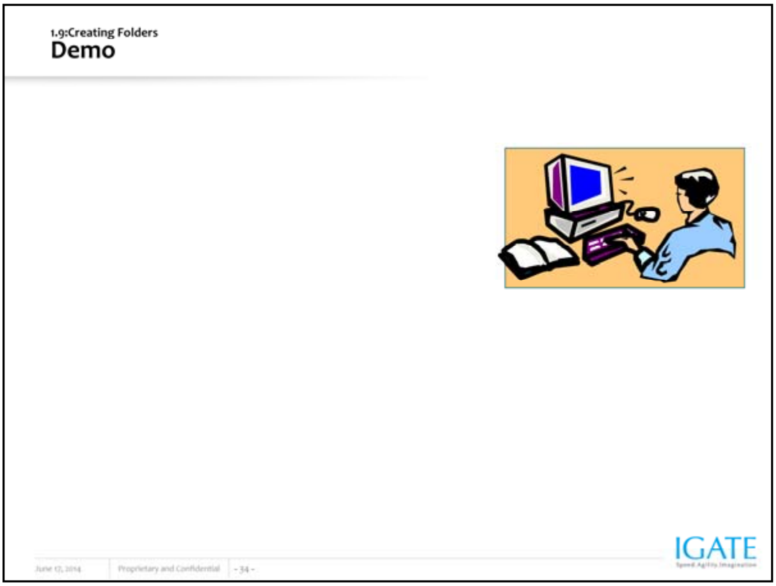
- **Read permission -** Allows to view the folder as well as objects in the folder
- **Write permission -** Allows to create or edit objects in the folder
- **Execute permission -** Allows to execute or schedule a Workflow in the folder

Note: Permissions work in conjunction with privileges. Privileges are actions that a user performs in PowerCenter applications. A user with the privilege to perform certain actions can require permissions to perform the action on a particular object.

1.9:Folder

Folder Permissions

- **Folder permissions can be granted on three levels of security.**
  - List User: User of the folder.
  - List Group: Users of a particular group.
  - List All: All the user belonging to repository services.



Steps to create a Folder:

1. In the Repository Manager, connect to the repository.
2. Choose ***Folder*** | ***Create***.
3. Choose ***Folder Properties***
4. Enter the following information:
    - Name – Folder name, Required.
    - Description - Description of the folder, Optional.
    - Owner - Owner of the folder. Any user in the repository can be the folder owner, Required.
    - Allow Shortcut - If selected, makes the folder shared, Optional

5. Choose ***Permissions*** Tab
    - Select the Radio Button from ***List User, List Group, List All***
    - Then select Permission check boxes accordingly

Summary

- **After completing this lesson you now:**
  - Understand the ETL Process
  - Know the role of PowerCenter ETL tool in a Data warehouse model
  - Know the various components of PowerCenter
  - Know your work environment in terms of a Repository, Users, and Folders

Review Question

- Question 1: The three levels of folder permissions are ______, ______ and ______
- Question 2: Informatica Client, and Informatica Server connect to the repository through the __________
- Question 3: One can start working in PowerCenter without having an assigned folder.
  - True
  - False
- Question 4: The first step in setting up repository security is to establish groups and users.
  - True
  - False

June 17, 2014 | Proprietary and Confidential | - 36 -

IGATE