**Informatica Tutorial**

**(**[**https://www.javatpoint.com/informatica**](https://www.javatpoint.com/informatica)**)**

**INTRODUCTION: --**

1. Informatica can handle a large volume of data.
2. The Informatica tool provides a complete data integration solution and data management system.
3. In this tutorial, you will learn how Informatica performs various activities .
4. such as data profiling, data cleansing, transforming, and scheduling the workflows from source to target.

**What is Informatica?**

1. Informatica is introduced as a software development company in the market.
2. It provides a complete data integration solution and data management system.
3. It launched multiple products that mainly focused on data integration.

Informatica is used to extracting required data form operation all systems and transforms the same data on its server and load it to the data warehouse.

Informatica is also introduced as a data integration tool. This tool is based on the ETL architecture. It provides data integration software and services for different industries, businesses, government organizations, as well as telecommunication, health care, insurance, and financial services.

**For example**: -- we can connect with more than one server database. Here we connect to both Oracle and Microsoft SQL Server databases. And it also integrates the data into another system.

**Data Extraction**

The process of reading and extracting the data from multiple source systems into the Informatica server is called the data extraction. Informatica data can extract or read different methods such as SQL Server, Oracle, and many more.

**Data Transformation**

Data transformation is a process of converting the data into the required format. Data transformation supports the following activities, such as:

* **Data Merging:** It integrates the data from multiple sources.
* **Data Cleansing:** It cleans the data from unwanted or unnecessary information.
* **Data Aggregation:** It aggregates the data using the aggregate function such as Sum(), min(), max(), count(), etc.
* **Data Scrubbing:** It is used to derive the new data.

**Data Loading**

Data loading is used to inserting the data into a target system. There are two types of data loading, such as:

Ezoic

1. **Initial load or Full load:** It is the first step; it adds or inserts the data into an empty target table.
2. **Delta load or Incremental load or Daily load:** This step takes place after the initial load, and it is only used to load new records or changed the old records.

Informatica is an easy to use ETL tool, and it has a simple visual primary interface. You drag and drop the different objects and design process flow for data extraction transformation and load.

These designed process flow diagrams are called the mappings. Informatica has a property to communicate with all major data sources such as RDBMS, mainframe, Flat files, XML, SAP, VSM, and many more.

**Informatica Key Metrics**

Some main points that reflect the growth in demand for the Informatica certification, such as:

* 2015 income was $1.06 billion, more than the combined income of Abinitio, DataStage, SSIS, and other ETL tools.
* 7-year Annual CAGR is 30 percent.
* Partners are above 450.
* Major SI, ISV, OEM, and on-demand leaders.
* Customers are above 5,000.
* Customers in 82 countries and a direct presence in 28 countries.
* Number 1 in customer loyalty rankings from 7 years.

According to all the above facts, it confirms that Informatica is in the huge demand in the world. 9.6.0 is the latest version of Informatica PowerCenter, which is available in the market. Below are the following different editions for the PowerCenter:

* Standard Edition
* Advanced Edition
* Premium Edition

**Why We Need Informatica?**

* To perform some operations on the data at the backend in a data system, then we need the Informatica.
* To modify, cleaning up the data based on some set of rules, we need the Informatica.
* By using the Informatica, it is accessible to the loading of bulk data from one system to another.
* It provides a broad set of features such as integration of the data from multiple unstructured, semi-structured or structured systems, operations at row level on data, and scheduling operation of the data operation.
* It also supports the features of metadata, so it keeps preserved the information of the process and data operations.

**Informatica Architecture**

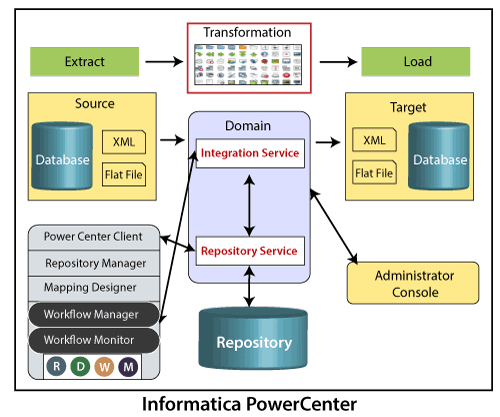
Informatica architecture is **service-oriented architecture (SOA**). A service-oriented architecture is defined as a **group of services that communicate with each other.** It means a simple data transfer during this communication, or it can be two or more services that coordinate the same activity.

The Informatica development depends upon the **component-based development techniques.**

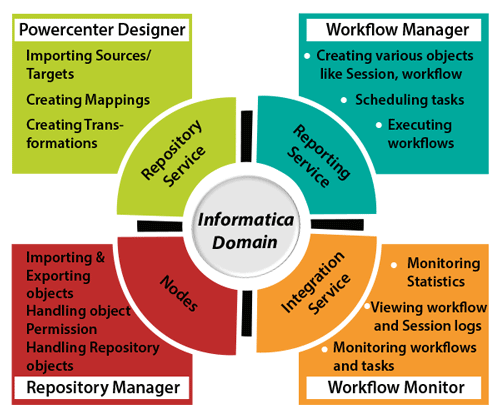
This technique uses the predefined components and functional units with their functionalities to get the result.

PowerCenter is based on the component-based development methodologies. To build a dataflow from the source to target, it used different components, and this process is called transformation.

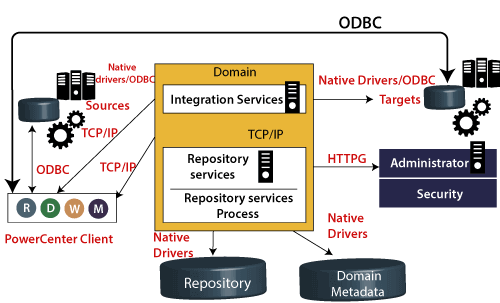
Informatica ETL tool has the below services and components, such as:



1. **Repository Service:** It is responsible for maintaining Informatica metadata and provides access to the same to other services.
2. **Integration Service:** This service helps in the movement of data from sources to the targets.
3. **Reporting Service:** This service generates the reports.
4. **Nodes:** This is a computing platform to execute the above services.
5. **Informatica Designer:** It creates the mappings between source and target.
6. **Workflow Manager:** It is used to create workflows or other tasks and their execution.
7. **Workflow Monitor:** It is used to monitor the execution of workflows.
8. **Repository Manager:** It is used to manage the objects in the repository.



Informatica Domain



* The Informatica domain is the fundamental administrative unit.
* The Informatica domain consists of nodes and services. These nodes and services are categorized into folders or sub-folders based on administration requirements and design architecture.
* The Console web page of the Informatica administrator creates a domain that looks like a folder. Inside this folder, we can create a node with the services.s
* In the Informatica domain, a node is a logical representation of the machine. All the services and processes run inside the domain in the Informatica. Multiple nodes can be present in a single domain. A gateway node receives the request from the clients and guides them to their respective services.
* The domain provides two types of services, such as:
  1. **Service Manager:** It manages domain operations such as logging, authentication, and authorization. It runs the application services on the nodes and leads users and groups.
  2. **Application Services:** It represents the server-specific services such as repository services, reporting services, and integration services. The application service can run on different nodes based on configuration.