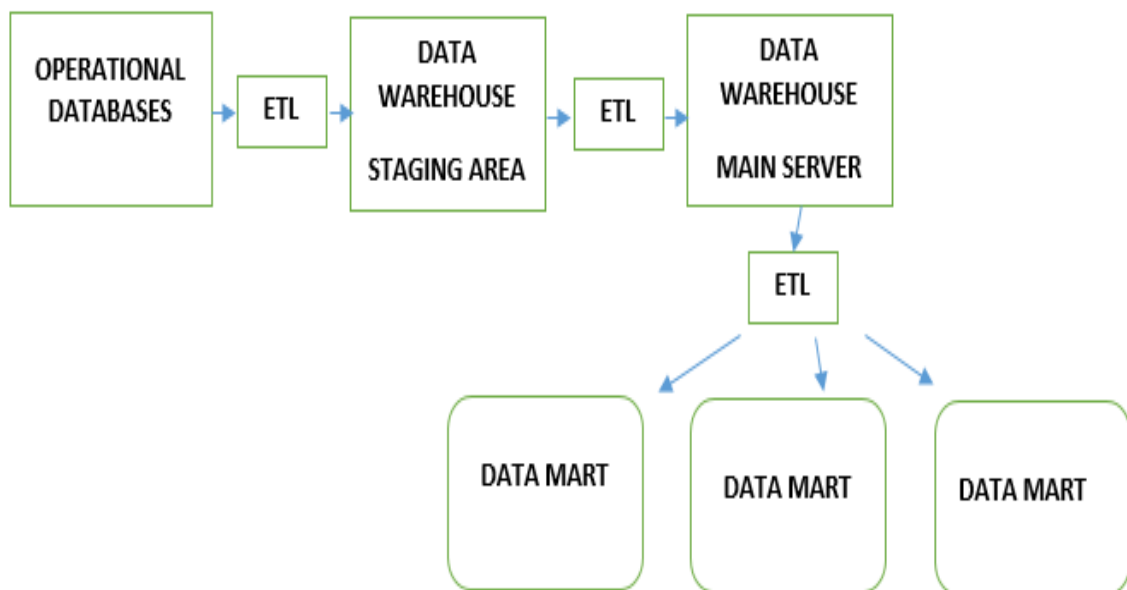# Title: Architecture of a Data warehouse

The Architectural Blueprint

To understand the innumerable Data Warehousing concepts, get accustomed to its terminology, and solve problems by uncovering the various opportunities they present, it is important to know the architectural model of a Data warehouse.

There are several different architectural models of Data Warehouses which have been designed on the basis of the specific requirements of a business.

A generalized model is as follows: As data is transferred from an organization's operational databases to a staging area, from there it is finally moved into a data warehouse and is set into conformed data marts. The copying of data is carried out by means of an ETL technology where data is extracted, transformed, and loaded. This is also represented with a schematic diagram below.

Systems in operation

Most businesses find their corporate data assets fragmented across disparate application systems which are running on various technical platforms in multiple geographical locations. This heterogeneity in data structure does not support good decision making as there is monotony which leads to the loss of data quality.

As a current trend for businesses, integration of operational data from various organizations has led to the development of mutually co existent business partners. For the same, sharing of

consolidated historical data among such business partners can improve their business prospects and profits.

The databases which are operational in an organization generally deal with a relational data view with a primary focus of data entry and do not support the consolidation of data, the generalization of data, and analytics.

Data Warehousing is the solution for suchbusiness requirements wherein data is consolidated and integrated from the various operational databases of an organizationwhich runs on several technical platforms across different physical locations.

ETL technology

ETL stands for Extract, Transform, and Load which are important operations of the architectural model of Data Warehousing. Transfer of all kinds of consolidated data is possible through ETL technology. Data is moved from one component of the model to another, all of which are accessible by decision makers.

Apart from the transfer of data which involves extraction and loading, ETL is also responsible fortransforming of inconsistent data, cleansing and filtering of data. Owing to such critical importance, ETL scheduling is critical as a single failure would disturb the entire process.

Next in line, A Staging Area Component

Utilizing the ETL technology, once data from source databases is copied, it is moved into a temporary location called aData warehouse staging area. The primary reason for the existence of a staging area is to ensure that all needed data is consolidatedbefore it can be integrated into the main components of a Data Warehouse.

In an active business, there exist many limitations in the hardware, network resource as well as differences in business cycles and data processing cycles which makes it a challenge to extract all the data from the databases simultaneously.

Let's tackle this with a very practical example, if you were a business which deals in sales, it wouldn't be convenient for you to extract data of sales on a very frequent basisas the data is meant for end-of-the-month evaluation. Similarly, extracting data is also affected by time zones which change greatly with geographical location.

The Data Warehouse Server

From the staging area by means of ETL, the data is then integratedwith the various internal and external operational databases of the organization which operate across the globe. This leads to a humongous collection of detailed data.

For example, the data of every sale ever recorded by a business would be convoluted which enables it to bestatistically analyzed very efficiently. With such abundance of data, the organization's reviewers would not access the Data Warehouse server directly. They access only the various front-end OLAP tools that analyze subject-oriented data and represent it as Data Marts.

Front-end Data marts

With assistance from the ETL technology, operations of transferring data from the warehouse to a data mart is done. Extracted data is represented on one or several Data Marts which enables it to be accessed by the organizations reviewers. The Data Marts often showcase a multi-dimensional view of extracted data with the help of front-end Data Warehousing OLAP Tools will be used to visualize the analyzed data or information.

A Data Mart resembles an excel spreadsheet. It contains several tables, columns, and rows, each representing a subject of the organization's Data Warehouse. Let's take the example of a business, their transaction data mart would contain several tables of their client's transactions from the previous/current year. Similarly, a data mart which focuses on the customers would contain data listed in several columns and rows of their customer's information like names, phone numbers, and addresses. For a Sales Data Mart, only data related to products sold and additional purchases would exist.