Process managers are responsible for maintaining the flow of data both into and out of the data warehouse. There are three different types of process managers:

- Load manager
- Warehouse manager
- Query manager

Data Warehouse Load Manager

Load manager performs the operations required to extract and load the data into the database. The size and complexity of a load manager varies between specific solutions from one data warehouse to another.

Load Manager Architecture

The load manager does performs the following functions:

- Extract data from the source system.
- Fast load the extracted data into temporary data store.
- Perform simple transformations into structure similar to the one in the data warehouse.
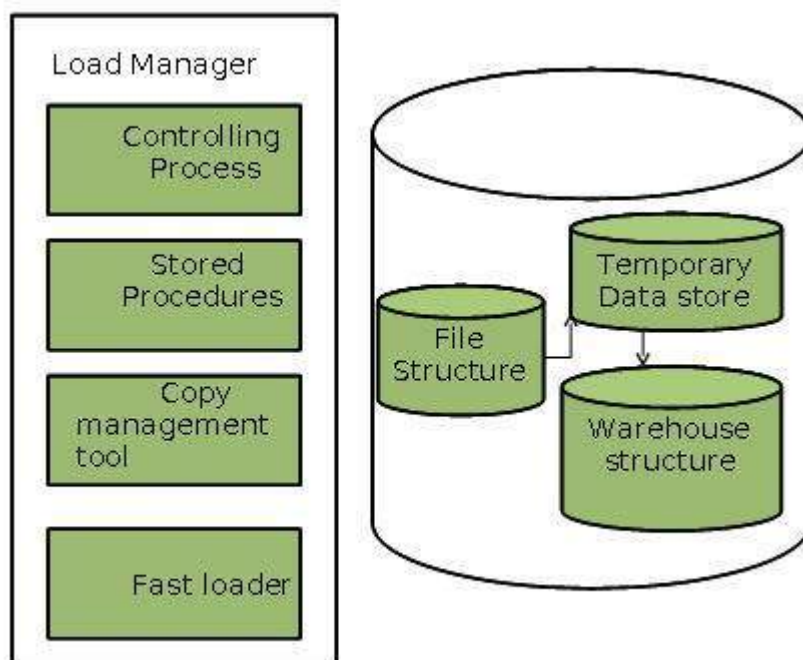
Extract Data from Source

The data is extracted from the operational databases or the external information providers. Gateways are the application programs that are used to extract data. It is supported by underlying DBMS and allows the client program to generate SQL to be executed at a server. Open Database Connection *ODBC* and Java Database Connection *JDBC* are examples of gateway.

Fast Load

- In order to minimize the total load window, the data needs to be loaded into the warehouse in the fastest possible time.
- Transformations affect the speed of data processing.

- It is more effective to load the data into a relational database prior to applying transformations and checks.

- Gateway technology is not suitable, since they are inefficient when large data volumes are involved.

Simple Transformations

While loading, it may be required to perform simple transformations. After completing simple transformations, we can do complex checks. Suppose we are loading the EPOS sales transaction, we need to perform the following checks:

- Strip out all the columns that are not required within the warehouse.

- Convert all the values to required data types.

Warehouse Manager

The warehouse manager is responsible for the warehouse management process. It consists of a third-party system software, C programs, and shell scripts. The size and complexity of a warehouse manager varies between specific solutions.

Warehouse Manager Architecture

A warehouse manager includes the following:

- The controlling process

- Stored procedures or C with SQL

- Backup/Recovery tool

- SQL scripts
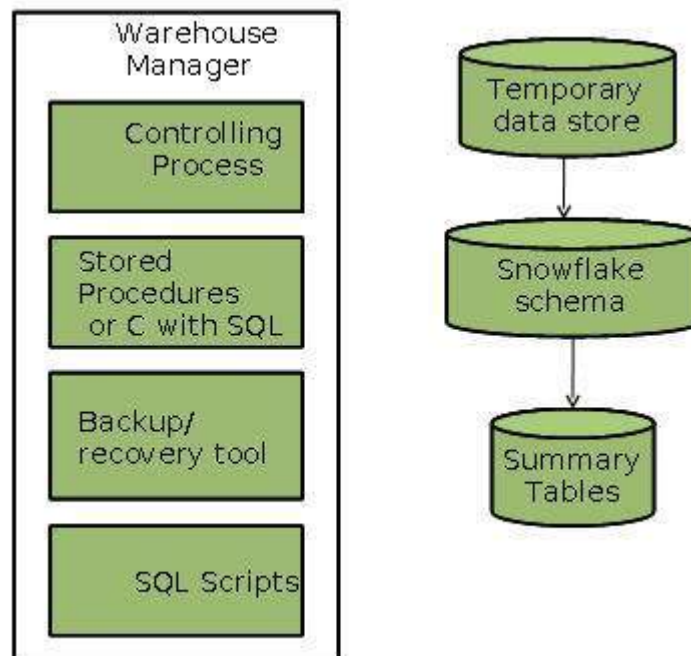
Functions of Warehouse Manager

A warehouse manager performs the following functions:

- Analyzes the data to perform consistency and referential integrity checks.

- Creates indexes, business views, partition views against the base data.

- Generates new aggregations and updates the existing aggregations.

- Generates normalizations.

- Transforms and merges the source data of the temporary store into the published data warehouse.

- Backs up the data in the data warehouse.

- Archives the data that has reached the end of its captured life.

Note: A warehouse Manager analyzes query profiles to determine whether the index and aggregations are appropriate.
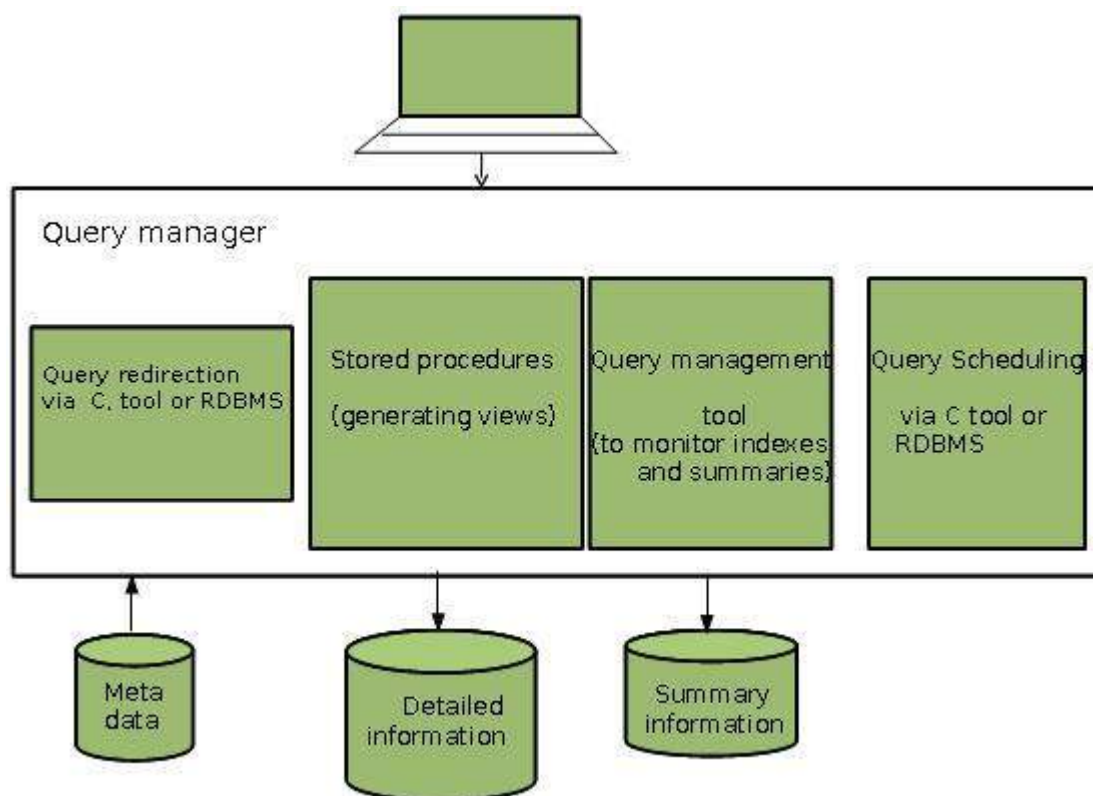
Query Manager

The query manager is responsible for directing the queries to suitable tables. By directing the queries to appropriate tables, it speeds up the query request and response process. In addition, the query manager is responsible for scheduling the execution of the queries posted by the user.

Query Manager Architecture

A query manager includes the following components:

- Query redirection via C tool or RDBMS

- Stored procedures

- Query management tool

- Query scheduling via C tool or RDBMS

- Query scheduling via third-party software

Functions of Query Manager

- It presents the data to the user in a form they understand.

- It schedules the execution of the queries posted by the end-user.

- It stores query profiles to allow the warehouse manager to determine which indexes and aggregations are appropriate