COMARCH
Business Intelligence

Comarch Data Warehouse Manager

9.6

INTRODUCTION TO DATA WAREHOUSE MANAGEMENT

Companies, as a result of their operations generate huge amounts of data. Collecting this data from many data sources and its integration in one place called data warehouse in order to be able to prepare analyses and reports is a complicated process, but it is necessary, if company would like to take decisions based on knowledge. Periodical loading of data from all systems and additional files to data warehouse is the main subject of this integration. Members of the board and operational managers are usually the main group to use data warehouse through reporting applications and tools. Taking into consideration quantity of data, and its complexity, properly build data warehouse is an excellent solution that helps to search for new trends, regularities, or logical patterns. Comarch DWM (Data Warehouse Manager) is a modern tool, which enables to build, and manage data warehouses.
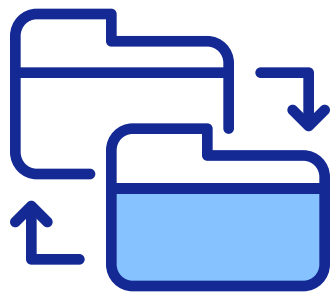
TRANSFORMATION AND LOADING (ETL)

Comarch DWM platform helps to gather data in order in one place, which is named data warehouse. Data warehouses enable to manage data efficiently in the most optimal way. Data before being loaded to data warehouse shall be extracted from source systems, properly transformed, cleansed, integrated, and finally loaded into final base. ETL is the process responsible for these operations.
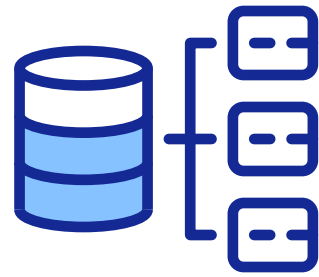

EXTRACT
Extraction of the data from source systems

TRANSFORM
Transofrmation and integration with final model
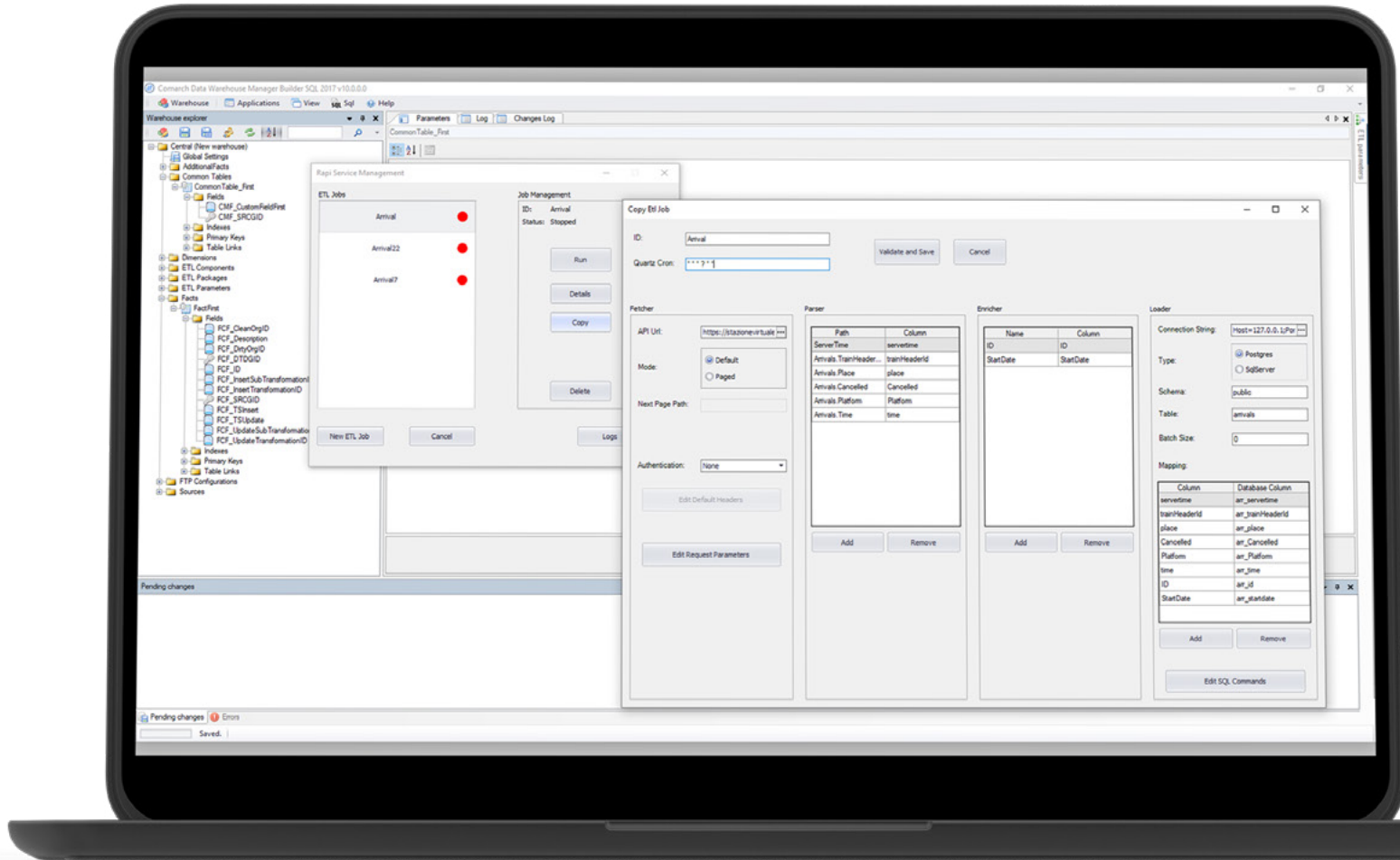
LOAD
Loading of the data into the data warehouse

ABOUT DWM TOOL

Comarch Data Warehouse Manager is an element of Comarch Business Intelligence Platform, which gathers all integration modules. Their purpose is to support and manage process of loading data into data warehouse, administration of data, its arrangement and describing. Application of particular modules depends on specificity of project, number of data sources and needs of company connected with analyses:

- Manage structure of data warehouse
- Manage process of loading and transformation
- Create logs for processes of loading, transformation, and informs customer's administrators
- Manage and process additional mappings, groupings, and keep dictionaries in additional data sources, if such requirements are defined

MAIN FUNCTIONALITIES

- Comarch DWM platform helps to generate a structure of data warehouse, dimensions and facts.
- It enables automatic generation of ETL packets, which load data to particular tables of data warehouse.
- All data concerning ETL process, and structure of data warehouse are being kept in meta tables. As a result of it, there is a possibility to generate reports which describe structure of bases and operating of ETL process.
- System enables management of data sources and information about source of data are being kept in destination table. It allows simple generation of reports regarding particular data sources.

The main advantages of DWM tool are:

- The main advantages of DWM tool are:
- Automatic generation of data warehouses
- Easy automation and control over the loading process
- Loading of data at various levels of detail
- Support of Slowly Changing Dimensions functionality (SCD)
- Integration with many different data sources: flat files, XLS files , MS SQL 2008, 2012, 2014, 2016 databases, Oracle databases, OLE DB, others
- Parallel loading to multiple tables at the same time
- Real time view of ETL processing

NEW DWM RELEASES AND ITS FUNCTIONALITIES

Our product is under continuous improvement process. In the latest versions the solution's performance has been improved. In addition, the following functionalities have been implemented:

- Enabling loading fact tables and dimensions from a json file and verification the correctness of the file structure,
- Ability to save changes to the data warehouse including comments to each change of the object,
- A new label "Changes Logs" has been added, which allows you to view all changes made on DW by selected categories,
- A separate application "Migrator" was added which allows the migration of META databases and the creation of migration scripts,
- It was allowed to group in table schemes in the facts and dimensions list,
- A new, universal mechanism for updating DW data structures has been developed,
- Additional modules supporting REST API and Kafka queue were developed, allowing for:
  - creating connections to the external Rest API and Kafka services (operating on Linux or Windows systems),
  - testing the connection and saving its settings,
  - previewing the connection data to the service, their modification and deletion the service from the list,
  - connection to selected, saved services,
  - an overview of the ETLJobs to the connected service,
  - ETLJob administration of the connected service - running, stopping, previewing, editing, deleting, copying, creating a new one,
  - previewing service logs.

DEVELOPMENT DIRECTIONS

Following the current trends in Business Intelligence area, the Comarch DWM tool is developing towards a distributed ETL system. It means, that our solution works with both commercial and open source data sources. Furthermore, the implementation of mechanisms such as Kafka support, allows for real-time data extraction for further processing them in data warehouse.

What is more, further development of the application will be continued based on the cloud ready approach. Comarch DWM as a network application will facilitate the maintenance of the system and enable quick response to changes and implementation requirements. At the same time, the computing power needed to perform individual loading tasks, can be distributed over many different environments.

For more information visit **www.bi.comarch.com**

COMARCH
ERP