50606 Mina Almas

29. Automatic Generation of Data Warehouse

Abstract

A Data Warehouse is not an individual repository product. Rather, it is an overall process, for building decision support systems and a knowledge-based applications architecture and environment that supports both everyday tactical decision making and long-term business strategizing. Data Warehouse applications provide the business community with access to accurate combines information from various internal and external sources. Building and maintaining Data Warehouses (DW) is one of the hottest topics in today's IT industry. Even though data warehousing has been around for more than 23 years, the total cost of owning a DW has been unaffordable for small to midsize companies.

Derivation of Data Warehouse structure from the metadata of operational databases is a tedious and time consuming task. Hence, a successful data warehouse designer requires knowledge and training in specific design techniques combined with practical experience. This research demonstrates that availability of computer readable meta-data description mapping of ERP/operational database structure helps to automate generation of data warehouse. The goal of my research is to accomplish the automatic generation of data warehouse from ERP/operational databases. The proposed application uses SQL Server2005 classes (SMO and AMO) on DOT.NET 2005 framework. The research shows that availability of computer readable meta-data description of mapping of operational database structure helps to automate generation of data warehouse.

My proposed Method is based on 7 main steps. Firstly all computer readable meta-data store in SQL tables.

Secondly the operational database mapped with the computer readable meta-data description and mapped transactions metadata is stored in SQL table.

Thirdly SMO Application Engine, designed in Dot Net, will identify the dimensions and measures from operational database's readable computer metadata which is stored in 5 different SQL tables.

Fourth, when the dimensions and measures have been identified and the dimensional model has been built, it's time to move on to the physical implementation. AMO will perform this task by using AMO application engine. Then we will be able to create, modify, and delete objects such as cubes, dimensions, mining structures, and Analysis Services databases.

The implementation of proposed application is presented using a case study of an organization's ERP/operational database.