Automating the creation of OLAP-cubes on the base of MS SQL AS for optimization of the infrastructure of Informational analytical system of the enterprise

Dzmitry Karzhounik, Natalia Lapitskaya

Abstract— Automating the creation of OLAP-cubes on the base of MS SQL AS in purpose to get around data access delays of regional informational analytical system caused by extension of raw data.

I. INTRODUCTION

In our days government facilities usually exploit information systems (IS) with architecture based on outdated technologies. Regional analytical information system (RAIS) was created in 2006 to provide reports exchange and collection of various statistical indicators from organizations in different areas. Such kind of data can be hierarchically organized but after long period of exploitation the growth of raw data in a data warehouse (DW) based on relational database (DB) model leads to noticeable delays in query execution. Existing client applications of the system are strongly connected to current DB model. To resolve that problem cost-effectively such situation creates a need to change already existing infrastructure to make use of that data without warehouse remodeling.

To update the infrastructure of informational exchange of the system new functionality must be added according to modern standards. Such standards can be shortly described by the concept of the "Analytical pyramid" (Figure 1).
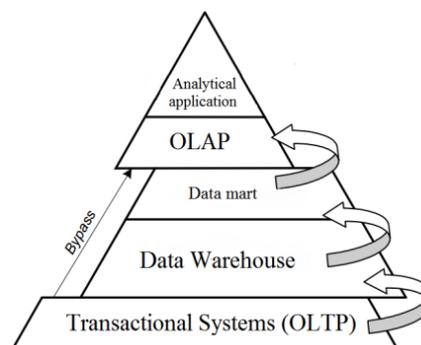

Figure 1. Analytical pyramid

II. ANALYTICAL SYSTEMS

RAIS provides information to the users by integrating data from different data sources. Though to view that data in a convenient way it should be processed and aggregated in a form of summary reports.

RAIS development to the next level means its update to OLAP-system. But the main goal is to create in the system means for enterprise resource planning or business process management.

D. Karzhounik, Faculty of Computer Systems and Networks, BSUIR, Minsk, Belarus, (e-mail: DimaKorzhovnik@yandex.com)
N. Lapitskaya, Faculty of Computer Systems and Networks, BSUIR, Minsk, Belarus, (e-mail: lapan@bsuir.by)

Such tasks are related to the area of Business intelligence (BI), so the software should be based on data marts using the technology of multidimensional OLAP-cubes. It's integration into the system will provide the required application performance in terms of growing volumes of data [1].

III. PRACTICAL RESULTS

At the moment RAIAS level according to analytical pyramid is DW and applications generate direct requests to it. If we add a level which includes Data marts and technology to ensure the functionality of OLAP-systems we will get the following schema of information interaction in the terms of Microsoft SQL server technology, which was selected cause it supports local cubes. (Figure 2).
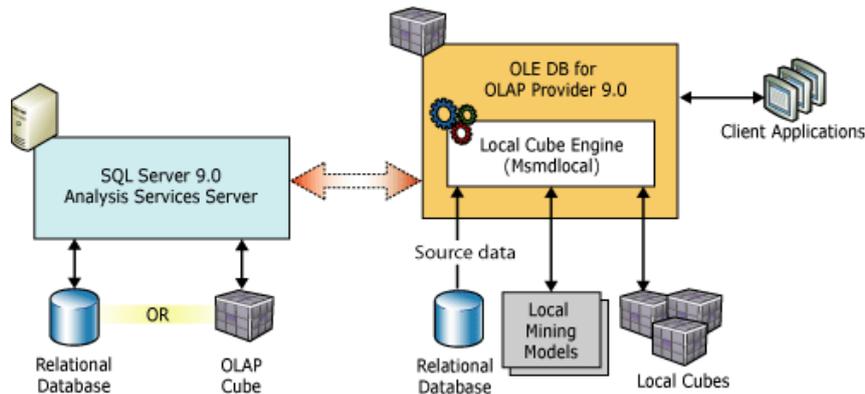

Figure 2. OLAP-systems interaction schema

Short explanation for Figure 2.The system collects reports with different, periodically varying statistical indicators from different branches of the organization in the region:

1. DW - Data Warehouse DBMS Oracle.
2. SI - statistical information, indicators of accounting, financial, etc. accounting
3. EDPS –local data bases of collected SI
4. IA - Interface array, the model describes how the lists are in SI DW.
5. SP IA SI – IA SI training system, designed to collect data from EDPS, convert them to a format IA and transfer the array to be loaded into DW.

BI-solution module in this case means a software that automates the construction of server OLAP-cubes as data marts and OLAP-local cubes as a report file with the extension * .cub. [8][5]

Based on the requirements of the domain user module must obtain a three-dimensional structure of a cube star - model, cube will contain following aggregated information: timing, area and accountable organization, depending on enterprise filial, organizations, and relevant lists of indicators, which reduces amount of excess aggregated data in the reports and data marts.

The resulting solution involves the use of tools and programming language support software to generate local and server cubes based on the data from the repository, as well as provide a complete API for working with data in the cubes and allow the standard to test the correctness of their selection for automatic generation. Comparative analysis showed that for model creation and implementation of the described functionality suitable technology stack from Microsoft to use the .Net platform and C# [3].

Microsoft Analysis Services has OLAP server model supports for local cubes, and integrates with database MS SQL Server, which will provide connection to the DW as the data source. In general with the use of following technologies (Figure 3) helped us to develop required software module with advanced functionality and even provide means of testing the output data [9].
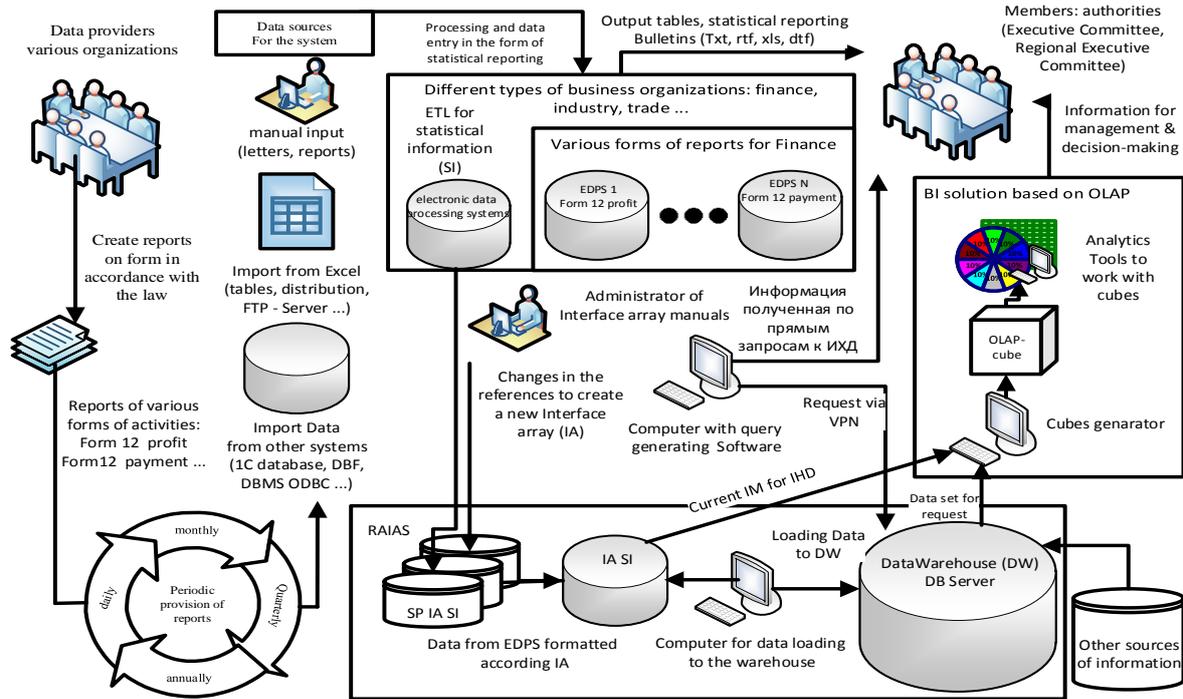
Figure 3. Infrastructure of information exchange system

NET Framework and C# gives following benefits [7]:

1. Modern data storage modes: xOLAP and Local cubes

2. C# + LINQ + Visual Studio + MS SQL Server. Good tool and means for software development

3. OLEDB (for Oracle DW) Connection to different data sources

4. AMO (Analysis Management Objects) for OLAP. Is used to automate the process of creating a cube, we need to do this programmatically [4]

5. ADOMD.NET provides API to work with cube data

6. MDX (Multidimensional Expressions) For testing aggregated data from the DW to cube

7. XML (standard form of data representation) For implementing special functionality

8. XML (standard form of data representation)

OLAP Server cube generation is divided into 7 steps:

1) Connect to the Analysis Services: Create a connection of Analysis Services server.

2) Create a Database: Create a database in Analysis Services server and save it.

3) Create a Data Source: Add a data source (our DW) to the database and set its connection string.

4) Create a Data Source View:

  a) Create a Data Set.

  b) Add Fact tables in Data Set.

  c) Add Dimension tables in Dataset and Relation between them.

  d) Create a Data Source View based on the created Data Set.

5) Create the Dimension, Attribute, Hierarchy, and Member Property Objects:

  a) Add Dimension to the Database.

  b) Add Dimension Attributes.

  c) Set Attribute usage and source

6) Create the Cube, Measure Group, Measure, and Partition Objects:

  a) Add Cube to the Database and set Cube source to the Data Source View.

  b) Add Measure Group to the Cube.

  c) Add Measure to Group and set Measure source.

7) Generate Cube:
   a) Add Dimension to the Cube.
   b) Use Regular Relationship Between Dimension and Fact Table Measure Group.
   c) Link Table Key in Dimension Table with Table Key in Fact Table Measure Group.
   d) Save Cube and all objects to the Analysis Services

After creating the cube, it can be processed manually from Microsoft Analysis Services [6], or programmatically so that logic could be included into business logic layer of the software added into the system infrastructure.

IV. CONCLUSION

It has been found that the OLAP-based solution can be implemented in several steps. First AMO is used to create multi-dimensional structures on the basis of MS SQL analytics server service. On the second step a client application can be developed to provide remote access to the data stored and implement any specific functionality required by the users. On the third step the software module based on ActiveX Data Objects Multi-Dimensional (ADO MD) can be used to create other software application integrated with .NET Framework, which would enhance system infrastructure.

The following tasks were completed in the project:

1) A software method of constructing a multidimensional data model, OLAP-cube, in order to optimize data storage and increase the speed of processing user requests, also increasing the flexibility of the system by excluding redundant data in local marts.

2) Implementation of a software module with the support of AMO.NET ADOMD.NET technologies that allow to connect to the server relational and analytical databases, ensuring the creation of these databases on the server, and provide an ability for automate testing correct data structures.

3) Feature of creating local cubes of data, those can be used as summary reports, gives a convenient way to work with the data from the repository without a need to connect to the database server, just using MS Excel instead.

Members of analytics team have reduced time spent on reporting, due to the developed software module that uses data marts of OLAP cubes, and as a result requests for information are faster. The analysis and reporting procedures has become much easier due to use of local files in custom tools such as Pivot Table.

REFERENCES

[1]   Isaev, DV - Analytical information systems [electronic resource] [Accessed on 11/01/2015] Available from: http://www.hse.ru/data/192/270/1238/ИсаевДВ%20-%20Аналитические%20ИС.pdf
[2]   Comparison of OLAP Servers [electronic resource] [Accessed on 08/02/2015] Available from: https://en.wikipedia.org/wiki/Comparison_of_OLAP_Servers
[3]   Using MS SQL Server Analysis Services 2008 for data warehousing [electronic resource] [Accessed on 22/02/2015] Available from: http://www.intuit.ru/studies/courses/568/424/info
[4]   Object Model AMO [electronic resource] [Accessed on 11/03/2015] Available from: https://technet.microsoft.com/en-us/library/ms345083(v=sql.110).aspx
[5]   Supports Local OLAP cubes model [electronic resource] [Accessed on 25/03/2015] Available from: https://technet.microsoft.com/en-us/library/bb522640(v=sql.110).aspx
[6]   An example of a server programmatically create OLAP-cube [electronic resource] [Accessed on 11/04/2015] Available from: http://avinashkt.blogspot.com/2008/04/olap-cube-generation-programmatically.html
[7]   Essence server core OLAP MS SQL Server Analysis Services [electronic resource] [Accessed on 24/04/2015] Available from: https://technet.microsoft.com/en-us/library/ms174776(v=sql.110).aspx
[8]   An example of the software create a local OLAP-cube [electronic resource] [Accessed on 11/05/2015] Available from: http://www.wiktorzychla.com/2007/11/complete-olap-infrastructure-without.html
[9]   Unit testing OLAP-cubes SQL Server using C # [electronic resource] [Accessed on 23/05/2015] Available from: http://msdn.microsoft.com/ru-ru/magazine/dn519927.aspx