

ARCHITECTURE ON-LINE ANALYTICAL PROCESSING

DIANA ELENA CODREANU, DENISA PARPANDEL, IONELA POPA
"CONSTANTIN BRANCOVEANU" UNIVERSITY OF PITEȘTI
FMMAE

2/A BASCOVULUI STREET, PITEȘTI, ROMANIA
codreanudia@yahoo.com, parpandeldenisa@yahoo.com, popaionela80@yahoo.com

Abstract:

OLAP systems are part of decision support systems and will assist analysts and managers, those who are responsible for the smooth running of an organization and improving organizational performance, giving them quick access to data stored in the system by reflecting the multidimensional aspect data.

OLAP tools provide users with a fast response even if the query request is made on a large volume of data. Basically, OLAP tools provide the ability to transform huge volumes of data that exist in the organization into useful information to support decision-making process.

Key words: OLAP tools, analytical processing, cubes, data warehouses, database

JEL classification: D89

1. Introduction

OLAP has become fast enough foundation solutions including intelligent and can speak for business performance management, planning, allocation resources, budget allocation, making forecasts, financial reports, findings and reports data warehouses.

Specialists consider that an OLAP system is an information server that allows quick access to data (atomic data and derived data) and facilities calculations [4].

Because OLAP tools work with multidimensional data models, perform complex analytical queries, ad hoc and has a high processing speed, some scholars have suggested that they rather be called FASMI (Fast Analysis of Shared Multidimensional Information) [5].

Thus, Nigel and Richard Creeth Pendle, from the idea that "*18 or even 12 rules are too much*" wanted to define the characteristics of OLAP applications in a simple way and thus propose FASMI test, five key words [3]:

- *Key word "speed"* - users have earlier access to data at any time (basically the smooth running of a company is assured of making correct decisions to be taken quickly based on data that is stored in a warehouse data, the speed of access becomes a necessity). If simple analysis response time is one second, and for more complex analysis, response time may be longer than 20 seconds;

- *key word "analysis"* - OLAP system should provide the user to perform various tests with the help of mathematical methods or statistical. Virtually, should provide the opportunity definition of ad-hoc calculations making reports in any form without the use of professional programming languages.

- *key word "shared"* - allowing data to be accessed by many concurrent users (this means that OLAP applications by means of specific mechanisms to ensure confidentiality);

- *key word "multi-dimensional"* - the backbone of OLAP applications is the multidimensional database, part of the data warehouse;

- *key word "information"* - all necessary data at a time can be accessed regardless of source or volume.

2. OLAP Architecture

OLAP systems have a structured architecture based on three essential components:

- *Database* - the data source used for OLAP analysis. As database can use a relational database to ensure our multidimensional storage facilities, a multidimensional database, a data warehouse, etc.
- *OLAP server* - the one that manages multidimensional data structure and at the same time a link between the database and OLAP customer.
- *OLAP customer* - are those that provide data mining applications but also supports the generation of results (graphs, reports, etc.).

The figure below shows the architecture of OLAP systems, which vary depending on how data storage and processing of their type, but generally on how one can identify three levels of data: the data sources, OLAP server and the presentation of data or interface user.

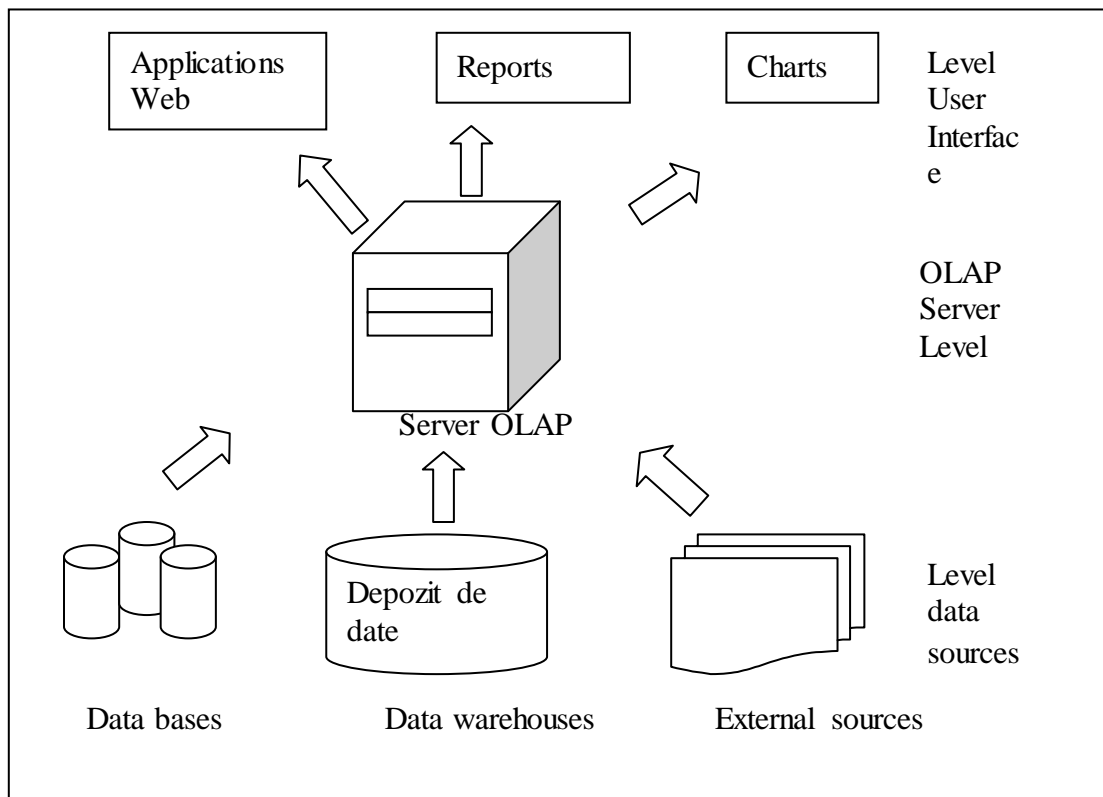


Figure 1: Architecture of OLAP systems

There are several options in OLAP data could be stored and processed. Thus, depending on the method of organizing and storing data, there may be three options [1]:

- *Client Files* - data is stored locally on a client computer as files are organized, on which operations can be applied to analyze the processing and transformation. This organization of data has some drawbacks of which we can enumerate: the amount that can be processed is indulged reduced time to processing information is quite high, the data shows a poor security, lack of advanced multidimensional analysis.

- *Relational databases* - this arrangement is used when the data comes as a relational DBMS and data warehouse is a repository be implemented virtually or using a relational model.

- *Databases multidimensional* - in this case, the data are organized into a data warehouse on a dedicated server, which is called multidimensional server. In this

situation we have a data warehouse with multidimensional objects. On multidimensional data can be applied by the server multidimensional operations. The data are drawn from various sources (relational databases, files), transformed and loaded into tables of facts and dimensions and units on different levels, pre-processed and prepared for analysis [1]. The most appropriate way of organizing because it delivers high capacity data storage, ETL process requires you exist for the processing and loading data but also the implementation of the operations performed at the server optimized for multidimensional analysis.

As regards data processing, we have all three options, namely:

➤ *The core of SQL* - is not the most appropriate options to perform complex multidimensional calculations because SQL language does not have facilities to perform multidimensional calculations directly and to have the same results as when using multi-dimensional functions and operations, be made several stages.

➤ *Multidimensional client engine* - some operations can be performed locally if the user has a relatively powerful system. If this alternative, the user must have advanced knowledge because he will have to build and apply analysis functions.

➤ *Multidimensional server engine* - we can say that this is the best option for achieving multidimensional operations. Dedicated server so the user performing the multidimensional and beyond final this task. Dedicated server provides access resources and applications competitor analysis is performed in real time.

OLAP tools enable users to store data in both relational databases and multidimensional databases. If we consider how to store data in databases, OLAP systems can be classified as:

- *ROLAP systems*

- *MOLAP systems*

- *HOLAP systems*

2.1. Systems ROLAP (Relational OLAP) - OLAP systems that allow for appropriate data is analyzed as a relational database (data warehouse). Basic data storage and data is aggregated in a relational database, and deterioration is done Working on a separate copy of data (such as data warehouse or data center), so based on transactional data . Storing multidimensional data in data warehouses or data center will be run using star or snow scheme. It aims to ensure quick access and flexibility for handling multidimensional.

Where there are many variables, there is a risk that data can not be stored in one table when facts and facts prove that the table is not an effective solution. Upgrading a single large fact tables proved to be an ineffective solution. Thus, the table actually used for more complex applications is divided into groups of variables, taking into account the degree of dispersion, setting the appropriate data sources and sizes. If we can meet complex OLAP applications and facts tables 10-20. .

In 1984 he launched the first product was launched ROLAP, and was named metaphor used producing marketing applications for consumer goods companies. Subsequently they were launched a series of products, of which we can mention: AXSYS - the company launched Advantage, Beacon of Company Products.

With ROLAP can exploit and sometimes large amounts of data base management systems offer specific facilities ROLAP data. Also, as the advantage we might remember that the data are stored in relational databases and can be accessed using SQL reporting tools (tools not be OLAP tools). As a comparison between ROLAP and MOLAP ROLAP can say that the systems are much slower. Basically it is a disadvantage of ROLAP systems response time to information requests, which sometimes can be quite high when complex queries are performed with multiple data sources.

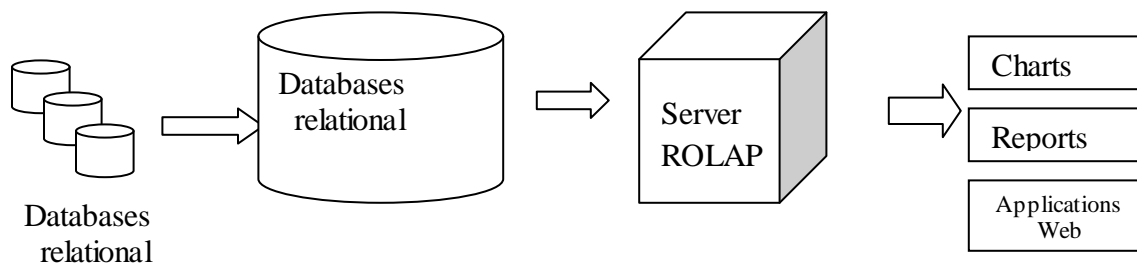


Figure 2: Architecture ROLAP

2.2 Systems MOLAP (Multidimensional OLAP) - usually used when we have less amount of initial data and solutions are seen as classic multidimensional analysis. Stores both data base and aggregated data in a multidimensional database, called the Cube, which are used as effective tools for operations analysis and to perform complex calculations. Basically, MOLAP mainframe systems provide users with a multidimensional view of data.

MOLAP systems have focused on optimizing the flexibility and storage techniques and the concept of transaction [2].

MOLAP systems are much faster in terms of data aggregation and in terms of queries, however, generates large volumes of data hedge. Response time the query is improved because of precalculate aggregations of such data and responses to queries are prepared before launching the application.

Analyzing disk space and response time performance of complex queries, we can say that MOLAP cubes are best.

As we mention advantages of MOLAP systems [2]:

- Relational tables are not suitable for multidimensional data;
- Multidimensional Arrays allow storage of multidimensional data efficiently;
- SQL language is not suitable for multidimensional operations.

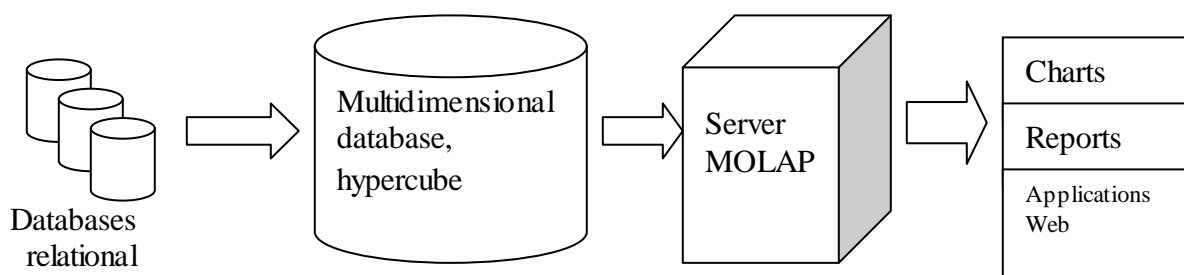


Figure 3: Architecture MOLAP

2.3 HOLAP (HYBRID OLAP) - can be called a compromise between the first two versions, which attempts to combine the advantages of MOLAP and ROLAP, to ensure users the best possible solution in terms of performance. If HOLAP basic data are stored in relational database while the data are aggregated in the cube HOLAP saved.

We can also say there are not too advanced in terms of scalability, but not in terms of speed. However, they both provide them a level and within acceptable parameters.

3. Conclusions

We can say that OLAP technology, multi-dimensional modeling of the backbone. "Multidimensional modeling helps us to present the results of economic activities in close conjunction with one or more factors that have participated in their training.

Multidimensional databases are optimized data structures used to exploit the data stored in data warehouses and OLAP for analysis.

The data stored in OLAP technology for data storage is used effectively in on-line data analysis, ensuring a rapid response to complex queries.

Multidimensional OLAP model with specific aggregation techniques provide for the organization of large data sets to enable a quick and easy interpretation. OLAP data analysts provide flexibility and speed of work necessary to substantiate decisions in real time.

REFERENCES

1. Lungu I., Bâra A., *Sisteme informatice executive*, Editura Ase, București, 2007
2. Munteanu M., *Inițiere în tehnologia OLAP: teorie și practică*, Editura Ase, București, 2004
3. Pense N., *What is OLAP? An analysis of what the often misused OLAP term is supposed to mean.* (<http://www.bi-verdict.com/fasmi.htm>)
4. Velicanu M., Munteanu M., *O caracterizare a sistemelor OLAP actuale*, Revista Informatică Economică, nr. 3 (19)/2001
5. www.marketwach.ro/articol/816/O_abordare_practică_a_bazelor_de_date_OLAP