



International Journal of Engineering, Business and Enterprise Applications (IJEBEA)

(Open Access, Double Blind Peer-reviewed, Refereed and Indexed Journal)

www.iasir.net

The Opportunities and Limitations of Using On-Line Analytical Processing Cubes for Business Intelligence Reporting in Banking

O Ulas¹, M Ozturan²

¹M.A.Graduate, Department of Management Information Systems, Bogazici University, Istanbul, Turkey.

²Professor, Department of Management Information Systems, Bogazici University, Istanbul, Turkey.

Abstract: The purpose of this study was to demonstrate opportunities and limitations of using on-line analytical processing (OLAP) cubes for business intelligence reporting in banking from a technological point of view. Triangulation technique on the basis of methods and data sources was adopted to enhance the validity and reliability of the study and crosscheck the findings from a particular method. First, related literature was surveyed. Secondly, a comparative experiment through development of an OLAP cube was carried out to determine the opportunities and limitations of OLAP cubes against relational model and re-evaluate the arguments in theory. Besides, structured interviews were conducted with banking information technology experts to confirm the findings of the experiment. The results indicated that OLAP was more appropriate than relational model due to greater query performance and simplicity in conducting complex queries with specific purposes. However, it was less flexible than the relational model, which could serve further needs and easily develop structured reports.

Keywords: Business intelligence, BI, On-line analytical processing, OLAP, Business intelligence reporting, Management information reporting; MI reporting, banking.

I. Introduction

The widespread use of information technology generates tremendous amounts of data within an organisation which contain information that is invaluable to the organisation's decision makers [1]. These data need to be transformed into consumable information for business organisations to be able to compete in the market. Business intelligence (BI) is a type of decision support system (DSS) and refers this transformation. Useful information can be derived from data available inside and outside the organisation through BI technologies to promote effective decision-making. Data warehousing, data mining and on-line analytical processing (OLAP) are key BI technologies. Data are stored and managed in data warehouse (DW), while OLAP and data mining convert the data into consumable information. OLAP cube services are considered to promote fast and multidimensional analysis of information on business variables.

There are significant changes in banking industry in recent years. Globalization, technological innovations, domestic or international mergers and acquisition, diversification and ongoing regulatory changes force banking industry to create new business strategies for competition, success or even survival in the market. Banks have to be more efficient in making decisions, improve operational performance, maximize productivity and profitability, increase workforce performance and reduce risks and operational costs; therefore, they have to invest more in BI technology. The main benefits of BI systems bring to any bank are their abilities to provide a deep understanding of past and current operations, and forecast future events [2]. Reference [3] states that BI systems for banking and finance are applications that provide a helping hand to banking people to let them work with information the way they want, hence business intelligence has gained high acceptance in this sector [4].

In this study, the role of using OLAP cubes in improving management information (MI) reporting in a banking organisation was examined from a technology perspective to gain a more reliable understanding in this scope. The study was based on existing BI infrastructure and needs of one of the leading commercial banks of Turkey whose name will remain undisclosed upon their request and be referred as "the Bank" within the text.

Main research question, which guided this study, is "What are the opportunities and limitations of using OLAP cubes for BI reporting in banking?"

II. Background

BI systems are the tools that combine data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information, provide actionable information delivered at the right time, at the right location, and in the right form to planners and decision makers [5]. BI process is considered as a cycle of "identify intelligence needs", "gather data", "process data", "analyze information" and "act based on the analysis" phases which occurs repeatedly until the decision-makers find appropriate answers. The outputs of a process serve as input for a new iteration of an ongoing BI process [6].

Lowering costs and expenses, reducing risks, increasing revenue, improving productivity, operational performance, profitability and customer satisfaction are denoted in literature as business benefits of BI [1], [7].

In theory, widely considered models for database design are dimensional and relational models. In dimensional database model data are presented in dimensional structure whereas the relational model stores the data in a normalized form. A school of thought suggests that relational environment is shaped by the corporate or enterprise data model while dimensional model is shaped by the end-user requirements. According to this approach, granular data in the relational model are not only used to service unknown future needs for information, but also to meet existing needs of existing user groups as in dimensional model [8]. Flexibility and performance issues are described as most important difference between the two models in theory. It is considered that relational model is highly flexible but is not optimized for performance while dimensional model is good at in performance but not at flexibility [8]. On the other hand, another group of authors consider that conducting the basic operations such as complex data aggregation, comparison and consolidation is difficult in relational method [9], [10]. Some authors suggests that dimensional model and DM are not scalable; those are departmental and not enterprise solutions and only appropriate when there is a predictable usage pattern [8] while the others refute these assumptions.

If the dimensional models are presented within a RDBMS, then these dimensionally modelled tables are referred to as star schemas. If dimensional models stored in an OLAP database, they're commonly referred as cubes.

OLAP is commonly defined as "a category of software technology that enables analyst and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user" [11] OLAP achieves the multidimensional functionality by using the cube structure which is specially designed to retrieve, analyze and report data efficiently on the OLAP database platform. An OLAP cube consists of numeric facts called measures, which are categorized by dimensions.

There are three major types of OLAP tools that differ with respect to their functionality and architecture: Multidimensional OLAP (MOLAP), Relational OLAP (ROLAP), Hybrid OLAP (HOLAP). MOLAP systems are dedicated on-line analytical processing implementations whereas ROLAP systems depend on relational databases [9]. HOLAP model attempts to incorporate the best features of MOLAP and ROLAP into a single architecture. However, there is not an agreement between the authors on the most effective type of OLAP architectures.

The followings are considered as the differences between relational reporting and OLAP analysis in theory [12]:

- OLAP provides interaction with data while relational reporting can provide limited interactivity.
- OLAP cubes are very good at providing very complex calculations whereas it is difficult in relational reporting.
- Relational reporting tools are better at providing report formatting than OLAP tools.
- OLAP provides very fast query performance when it is compared with relational model.
- Relational reporting is better in handling textual information than OLAP.
- OLAP may provide less maintenance since all aggregations are automatically provided within the cube. Relational reporting can be maintenance intensive if numerous "aggregation tables" are necessary.
- Ad hoc query creation is very easy with OLAP tools when it is compared with relational reporting.
- Conducting trend relationship between several terms of data is easier in OLAP systems with respect to the relational ones.

OLAP cubes are considered as specifically developed tools for the queries in large databases with many relations involved, such as those in banks and optimized for short query response times. OLAP systems are utilized to discover trends and to analyze critical factors and to keep the banking managers informed about the business situation. Since OLAP cubes are precalculated, which results in fast query performance that is critical for executive use, OLAP is well suited to handle complicated organisational roll-ups, as well as complex calculations, including inter-row manipulations. With these features, OLAP naturally fits for banking reporting [9].

III. Methodology

The goal of this study was to compare OLAP reporting with traditional techniques. Due to considerably broad scope of the study and in order to enhance the validity and reliability, triangulation technique was adopted on the basis of methods and data sources, which provides a check of findings from a particular method [13]. The type of the study and approach to research issue involved several directions and methods. Both quantitative and qualitative techniques were combined during the study and data were gathered from several sources.

First, a conceptual analysis through the review of theoretical background and literature was performed, which composed the theoretical and qualitative part of the study. A conceptual framework was established to define, describe, classify and organize existing information related to subject of the study before quantification and measurements.

Secondly, a comparative experiment through design and development of an OLAP cube was carried out for experimental verification of theoretical predictions by quantitative evidences. It was aimed to demonstrate the opportunities and constraints of an OLAP cube implementation empirically by testing its performance and to observe OLAP queries in the context of functional properties. Current BI infrastructure, management reporting

system and relevant data of the Bank were investigated and identified prior to this experimental design. After development of the cube, several performance tests were conducted. It was aimed to measure, examine and compare certain aspects of relational and OLAP reporting on existing data structure of the Bank, to obtain objective evidence about the opportunities and limitations of using OLAP cubes in banking reporting. Lastly, the opinions of banking IT experts with BI experience (on an average of around 5 years) were investigated through structured interviews, which were conducted to confirm the findings of the experimental cube implementation and to examine qualitative aspects of OLAP cube utilization in reporting. During these interviews 11 participants answered 30 questions, which consist of certain statements derived from the results of performance tests and certain theoretical assumptions on functional properties of the two reporting system,. The participants were also asked to submit additional comments at the end of each group of questions, which might clarify their responses.

The main approach of this study was deductive since it mainly intended to test theoretical assumptions and provided empirical evidences on opportunities and limitations of using OLAP cube in banking; however it involved inductive approach at some points of the project. Certain new patterns, new impacts of OLAP cube implementation in managerial reporting could be detected or observed as consequences of experimental analysis.

Similarly, both descriptive and explanatory research methods were chosen considering the objective to obtain first hand data from the performance tests and investigate the opinions of the experts on the findings of these tests. This study was descriptive since it reviewed the basic facts of BI, DW and OLAP technologies along with theoretical discussions in this field and tried to provide a clear view and detailed picture on usage of such kind of applications through investigation of common opinion of users and experts. In the meantime, it was an explanatory research due to the fact that it tested the relational and OLAP reporting processes and provided evidences through a comparative experimental analysis whether they supported or refuted the several arguments in theory; enriched the explanations on BI and OLAP applications in literature.

IV. Implementation and Performance Tests

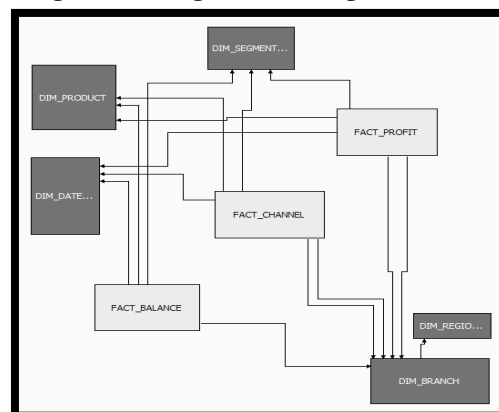
Main steps of the implementation phase of the study can be summarized as:

- Identification of relevant data of the Bank being used for management reporting
- Preparation of the cube development environment
- Design and development of the cube's source database
- Preparation of the mapping documents
- Design and development of extract, transform and load (ETL) processes
- Loading data into the cube DM
- Design and development of the cube
- Preparation of the static reports

The current MI reporting system of the Bank running on the corporate DW was based on scheduled reports, which involved some critical measures of the corporation in a highly aggregated form. This system was analyzed and the information currently reported within the system was extracted from the scheduled static reports. The reported fields were listed and then grouped under certain sets of dimensions and measures in order to provide a reference model for the cube design. It was not appropriate to develop such kind of OLAP cube on the DW environment. Therefore, it was decided that the cube's source database should be on another and more convenient environment rather than the DW. For this reason a new DM model was developed consisting of related data sets.

Next stage was the preparation of the ETL jobs, which would feed the cube DM. The data mapping documents were prepared which described the relation between the DW and the cube DM. Then, ETL jobs were prepared and were run multiple times and the DM was filled with the termly data. About 250 GB of data were transferred into the DM, which filled fifty-one partitions at the initial load phase. For the continuity of the data transfer, the job was scheduled to a scheduler.

Figure 1: A High-Level Design of the Cube



The OLAP cube was built using MS SQL Server 2005: Analysis Services (SSAS) with the model shown in the Figure 1. The dimensional model at the relational tables was represented multidimensionally within this cube.

A separate partition for each month in the measure groups had to be created within the cube. That was because there were a great amount of data in the fact tables and maintenance process of the termly data would be painful during each update in the source. The cube was developed with MOLAP functionality while the query response times were the most prior concern of the study.

Following the completion of the design phase, the cube was deployed to the production server and all the dimensions and measure partitions were fully processed. In order to keep the cube up-to-date, a job was scheduled to run daily to process the dimensions and current partitions of the measures. For monthly process, another job was created as well; however, it wasn't scheduled due to the manual interference requirements.

By the development of the cube, in the final BI architecture of the Bank, there were source systems at the bottom level. Then the data were loaded into the DW. Internal and external applications also existed within the system and the cube took place at external application layer. Finally the data from the DW and those applications were presented within the reporting layer.

Afterwards, two performance tests were applied for the relational and multidimensional environments. First test was related to the evaluation of report development processes for the systems. For this purpose, a set of eight reports with distinct characteristics were chosen from the current reporting system, which differed in report complexity (query cost), number of joined tables, row count of the queried tables, row count of the result set, number of queried terms. To prevent potential performance advantage of the powerful DW of the Bank, chosen reports were redeveloped in the same formats on the new relational DM. MS SQL Server 2005: Reporting Services (SSRS) were used to prepare the static reports on the OLAP cube.

The second test was related to the proportion of the “additional” progresses needed to accomplish reporting by the two environments. There were a number of sub-tasks, which would ease observing the total time necessary for each of these additional progresses. Table 1 demonstrates these tasks with definitions and environment dependent necessities. The sub-tasks in relational and cube environments were classified as ‘Necessary’, ‘Optional’ or ‘Completely Unnecessary’. “Necessary” was a step that had to be fulfilled for the implementation whereas “Optional” was a step without which implementation could also be realized. However, in order to see the results immediately or to obey the development rules of the Bank, these optional steps might still be needed. The tests in this section that measure duration were mainly related to necessary human effort required to accomplish the tasks. Therefore, the calculations were mostly made on the basis of MH (Man/Hour).

Table 1: Development Tasks

	Relational		OLAP	
	Necessary	Optional	Necessary	Optional
Alter database logical model		x		x
Alter database physical model	x		x	
Modify ETL job design	x		x	
Deploy ETL jobs		x		x
Re-run ETL jobs and fill the new table/row		x	x	
Alter cube model			x	
Deploy cube			x	
Process cube			x	
Modify report template	x		x	
Process report		x		

V. Results and Discussion

The quantitative results from empirical phases of the study can be summarized as follows:

- Most costly operations occur while reading the data from the database and then while aggregating them hierarchically into several dimensions.
- Comparison of report development processes of relational and OLAP environments indicate that:
 - OLAP has a big advantage in terms of report processing duration when compared to the relational system.
 - Response times to for the necessary development tasks given in Table 1 come out to be 31.63 MH for OLAP where it is only 0.6 MH for relational environment.
 - There is a limitation at the relational side since a query can only compose aggregations up to a maximum level and therefore may not fulfil further needs.
 - Though it may be expected OLAP cube will have a smaller size since it kept only the summary data at the source DM, the total amount of storage necessary to maintain the cube is almost equal to the size of the DM itself.
 - Due to the full process requirement, it can be said that OLAP model is less flexible than the relational model.
 - Scalability is a limitation for both OLAP and relational systems while dealing with such large data sets.

- Since OLAP environment does not need to run beforehand (since they can be taken quickly from the cube), when the cube is full processed, all the reports become available for all the previous terms contradictory to relational system.

The qualitative results obtained through interviews done with the IT experts in the banking sector can be listed as follows:

- A great majority of the respondents agree that the OLAP systems have much greater reporting performance compared to the relational systems.
- According to most experts, aggregating, comparing and consolidating data within the relational system is harder than conducting those operations in an OLAP system. All of them support the idea considering that OLAP also performs better in conducting trend analysis.
- Common opinion among the participants is that the maintenance of an OLAP system is harder than the relational system. According to more than half of them, necessary effort to update the cube is greater than relational system because of its long process durations. This disadvantage will prevent OLAP from being up to date with the DW.
- The participants confirm that during the implementation of a change request to both systems, OLAP takes a longer time because of its full process need. On the other hand, except the process duration of the cube, the response times of both systems to the development needs are very close to each other.
- The response times of the relational and the multidimensional environments to the new development needs that have been calculated within the experiment are also confirmed by most of the respondents.
- The great majority of the participants agree that it is more flexible to create ad-hoc reports through OLAP. On the other hand, according to the common opinion, development of structured reports through relational reporting tools is easier than OLAP reporting tools.
- Most of the respondents strongly agree that fast query response times and several features of OLAP, which promote the presentation of the information from several perspectives, provide interactivity with the data.
- The respondents agree that OLAP is more appropriate than relational model for managerial reporting in banking. On the other hand, it is easier to serve further needs through the relational DM.
- Some of the experts refute the proposition stating that the total cost of ownership of OLAP for MI reporting is greater than the relational system. They say that modification and processing costs of OLAP become worthy when the report durations in the relational side are considered. Waiting duration for each set of queries will be much higher than OLAP, especially when trying to run them for the previous terms. Also, they add that, by using cubes, some very complex reports will start to be taken, which have not been possible within the current system.
- A great majority of the participants agree that OLAP cubes are not proper tools for rapidly changing environments; however, according to common opinion among them, OLAP is more appropriate than relational model for managerial reporting in banking since the several perspectives of data are needed to be reported in a much aggregated way.
-

VI. Conclusion

As stated in the beginning, main goal of the study was to demonstrate the opportunities and limitations of using OLAP cubes for BI reporting in banking, which are considered to promote fast and multidimensional analysis of information on the business variables in this field. Several research methods were used and data were collected from multiple sources to produce more reliable and valid conclusions. A comparative experiment was carried out for empirical verification of theoretical predictions. Additionally, interviews were conducted to confirm the findings from experiment

In order to evaluate, to what extent do the opportunities and limitations of OLAP utilization in banking differ from theory, findings from the theoretical and empirical parts of the study with the contributions of the results from the interviews can be stated as follows:

A group of authors suggest that relational environment is shaped by the corporate or enterprise data model while dimensional model is shaped by the end-user requirements. According to this approach, granular data in the relational model is also used to service unknown future needs for information, not only to meet existing needs of existing user groups as in dimensional model. The results of this study show that the theoretical assumptions are true as dimensional model specializes in particular points and serves only a specific purpose, however, reporting flexibility of the dimensional model is much better, thus, it should be well decided which model must be used in accordance with the purpose.

Flexibility and performance issues are described as most important differences between the two models in literature. It is considered that relational model is highly flexible but is not optimized for performance while dimensional model is good at in performance but not at flexibility. On the other hand, many authors of second group consider that conducting the basic operations such as complex data aggregation, comparison and consolidation is difficult in relational method. Experimental evidences provided by this study demonstrate that in order to operate these kinds

of applications, the relational model cannot provide sufficient performance. A relational model is a necessity to store the granule data and to serve the future needs but dimensional DMs should be built in order to develop BI applications effectively. Both architectures should be used together within a BI infrastructure.

Regarding functional features of OLAP, sufficient speed to support ad hoc analysis, unrestricted cross dimensional operations, intuitive data manipulation and unlimited dimensions and aggregation levels have been specified in theory as the opportunities of OLAP reporting. The findings of this study show that OLAP is actually appropriate for ad-hoc reporting and provides interactivity with the data through the mentioned features.

There is not an agreement between the authors on the most effective type of OLAP architectures, MOLAP, ROLAP and HOLAP. MOLAP by nature has a greater performance than ROLAP since it precalculates all the results and store them in the cube. MOLAP method has been also chosen for this study since the main goal has been to investigate whether OLAP improves the reporting performance. On the other hand, ROLAP has advantages like flexibility, scalability and easy maintenance since it doesn't require processing the cube beforehand. Opinions of the experts gathered through the interviews brought a new perspective to the OLAP concept by indicating that recently developed DW appliances may provide the performance of OLAP by using relational systems. By having such a high-end technology, MOLAP performance may be achieved when ROLAP model is used.

In theory, OLAP cubes are considered as specifically developed tools for the queries in large databases with many relations involved, such as those in banks and optimized for short query response times. OLAP systems are utilized to discover trends and to analyze critical factors and also well suited to handle complicated organisational roll-ups, as well as complex calculations. So, OLAP naturally fits for banking reporting. This study verifies this statements, however, it has been seen that handling the new development needs in rapidly changing environments can be painful when using cubes. Therefore, OLAP systems should be preferred to relational ones especially for the environments, which have static structures and reporting needs.

To conclude, the need to continue the studies on the methods toward more efficient reporting from both academic and business point of view is obvious due to the significant role of managerial reporting in banking in rapidly changing business environment.

VII. References

- [1] M. Ritacco and A. Carver, "The Business Value of Business Intelligence", Trends Quarterly, 2007. [Online]. Accessed: 21.03.2014. Available: <http://www.btquarterly.com/?mc=business-value-bi&page=bi-viewmarketplace>.
- [2] J. Knapik, "Business Intelligence in Retail Banking", Data Monitor, 2007. [Online]. Accessed: 27.04.2014. Available: <http://www.researchandmarkets.com/reports/586519>.
- [3] T. Vaidya, S. Ahirrao, K. Kadam, and K. Lugade, "Business Intelligence System for Banking and Finance", International Journal of Computer Science and Information Technologies, vol. 5(3), 2014, pp. 4336-4349.
- [4] G. Muhammad, J. Ibrahim, Z. Bhatti, and A. Waqas, "Business Intelligent as a Knowledge Management Tool in Providing Financial Consultancy Services", American Journal of Information Systems, vol. 2(2), 2014, pp. 26-32.
- [5] S. Negash, "Business Intelligence", Communications of the Association for Information Systems, vol. 13, 2004, pp. 177-195.
- [6] V. Pirttimäki, "Business Intelligence as a Managerial Tool in Large Finnish Companies", Ph.D. Thesis, 2007. Tampere University of Technology.
- [7] M. Robinson, "Business Intelligence Infrastructure", Information Management, 2002. [Online]. Accessed: 17.02.2014. Available: <http://www.information-management.com/specialreports/20020521/5211-1.html>
- [8] W.H. Inmon, Building the Data Warehouse, 4th ed., New York, USA: Wiley Publishing, 2005.
- [9] R. Kimball and M. Ross, The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd ed., New York, USA: John Wiley & Sons, 2002.
- [10] E. Thomsen, OLAP Solutions: Building Multidimensional Information Systems, 2nd ed., New York, USA: John Wiley & Sons.
- [11] OLAP Council, On-Line Analytical Processing, OLAP and OLAP Server Definitions, 1997. [Online]. Accessed: 19.05.2014. Available: <http://www.olapcouncil.org/research/glossaryly.htm#Defined>.
- [12] Business Intelligence Consulting, OLAP Analysis vs Relational Reporting, OLAP Business Solutions, n.d. [Online]. Accessed: 19.05.2014. Available: http://www.obs3.com/olap_vs_relational.shtml.
- [13] S. Greener, Business Research Methods, 1st ed., Copenhagen, Denmark: Ventus Publishing ApS, 2008.