THE BUCHAREST UNIVERSITY OF ECONOMIC STUDIES

DATABASE SYSTEMS JOURNAL

ISSN: 2069 – 3230

Vol. IV, Issue 1/2013

LISTED IN

Cabell's Directories of Publishing Opportunities,
RePEc, EBSCO, DOAJ, Open J-Gate,
Index Copernicus, Google Scholar

BUSINESS INTELLIGENCE

ERP

DATA MINING

DATA WAREHOUSE

DATABASE

Contact
Calea Dorobanţilor, no. 15-17, room 2017, Bucharest, Romania
Web: http://dbjournal.ro
E-mail: editor@dbjournal.ro

Contents

ARH_Db_Tuner: The GUI tool to Monitor and Diagnose the SGA Parameters Automatically

Hitesh KUMAR SHARMA[1], Aditya SHASTRI[2], Ranjit BISWAS[3]
[1]Assistant Professor, University of Petroleum & Energy Studies
[2]Vice-Chancellor, Banasthali University, Rajasthan-304022, India
[3]Head and Professor,CSE Dept, Jamia Hamdard (Hamdard University)
hkshitesh@gmail.com, adityashastri@yahoo.com, ranjitbiswas@yahoo.com

Database administrators should be aware of resource usages to maintain database system performance. As database applications become more complex and diverse, managing database systems becomes too costly and prone to error. Autonomic database tuning becomes more important than ever. One of the major issues to address in regards to ORACLE database performance is the size of the database. The bulk of the information consists of a large number of records, contained in many tables, each ranging from thousands of rows. There are many factors that can have direct effects on the performance of the database. CPU, Memory, Network, Disk I/O are among other factors. In order to make a database up and run efficiently, each factor must be addresses carefully, and the best tuning strategy must be applied for optimum performance. ORACLE performance issues are complex, and for a DBA, there is a large number of values to monitor and examine in order to decide on best tuning strategy.

The aim of the work behind this paper was to design and implement a Database Performance Tuning Measurement Toolkit for ORACLE Database Servers working on MS-Windows platforms. This system is called ARH_Db_Tuner. This system is the advance version of ORACLE Performance Monitoring Toolkit [19] (OPMT for short). Some of the future aspects of OPMT have been implemented in ARH_Db_Tuner. It is used for testing, analysis and reporting of the database performance.*

*** ARH stands for Aditya-Ranjit-Hitesh (The first names of Authors)**

Keywords: *SGA, SGA Dynamic Parameters, Database Tuning, DBA, Automated Tuning.*

1 Introduction

Database vendors are becoming aware that the human cost of operating large database systems is growing dramatically. As the scope of relational database functions has expanded in recent years, the complexity of database systems has also grown. The added complexity and the increase in data size (now frequently into tens of terabytes) have increased the burden on database administrators. The combination of increased data volumes, larger systems, and increased function, has motivated the need for autonomic capability within database management systems in order to reduce cost of ownership and to enable databases to operate in environments with limited access to skilled administration personnel [1].

A new and interesting approach to this management problem is an *autonomic DBMS* that is able to automatically manage its resources to maintain acceptable performance in the face of changing conditions [2]. Other terms used in the literature for an autonomic DBMS are *self-tuning DBMSs* [3] and *no knobs operation* [4]. An autonomic DBMS must be able to perform the configuration tasks currently carried out by DBAs to initially set up a system. Typical configuration tasks include determining appropriate allocations for main memory areas such as the buffer pools and the sort heap, mapping database objects (user defined tables and indexes) to disk storage and mapping database objects to

buffer pools. Performing configuration tasks requires knowledge of the available system resources and of the characteristics of the workload.

Many system performance evaluation models have been proposed, varying in degree of accuracy, cost to build and run, and the reliance of the method on configuration-specific parameters. Many of these models are concerned with evaluating the appropriateness of a particular hardware configuration on the mix of programs that will run on the system. These methods, generally used for capacity planning, include queuing models, simulation, and monitors of contrived or real workloads. Detailed simulation models, such as IBM's CSS, FIVE [5], and ANCICSVS [6], QSIM [7], and APLOMB [8], are complicated, requiring large initial investments of time and money. These investments are not warranted by a single use; thus, these simulators are best suited for development by specific manufacturers and may not be publicly available. General packages, such as SCERT and CASE, are more useful as flexible evaluation tools [9].

On the other hand, and as businesses across all industries are getting more dependent on IT infrastructure, the problem of availability of IT systems becomes an increasingly strategic business concern. This is because the interdependencies of modern business are often linked through software. On one hand, employees, customers, and partners communicate and conduct commerce through networked systems, at a more intense level today than ever before. On the other hand, when one system in the network fails, it can break dependencies and impact business processes across the enterprise and beyond to customers and partners.

To avoid the ramifications of database downtime, many corporations have taken a renewed interest in database availability. For some, the goal is continuous availability, where a database server never fails, which would result in so-called unplanned downtime. Most companies do not need such a stringent level of database availability; they are satisfied with high availability, which allows for a small room of planned downtime allocated for database maintenance. Most strategies for continuous availability and high availability assume that a slight repair to hardware and/or software is all that is required for recovery.

High availability is often associated with fault-tolerant systems. The term *fault-tolerant* means a system can operate in the presence of hardware component failures A single component failure in a fault-tolerant system will not cause a system interruption because the alternate component will take over the task transparently. As the cost of components continues to drop, and the demand for system availability increases, many *non-fault-tolerant* systems have redundancy built-in at the subsystem level. As a result, many non-fault-tolerant systems can tolerate hardware faults— consequently, the line between a fault-tolerant system and a non-fault-tolerant system becomes increasingly blurred [10, 11, 12, 13]

Another key point in bringing database systems users' satisfaction is the system performance. Performance can be defined as the ability of a system to deliver the results based on the request of the users, while keeping them satisfied. The key point to remember is satisfaction. If current database key indicators reveal that the database is running optimally, but users are not satisfied with the response times, then there will be a need for tuning [14].

One of the biggest responsibilities of a DBA is to ensure that the database is tuned properly. The ORACLE RDBMS is usually tunable and allows the database to be monitored and adjusted to increase its performance. One should do performance tuning for the following reasons [15]:

- The speed of computing might be wasting valuable users' time (users waiting for response);

- Enable your system to keep-up with the speed business is conducted; and
- Optimize hardware usage to save money (companies are spending millions on hardware).

Tuning the database performance is not a simple task and it depends on the specific requirements, the operating system and the target hardware. There is no "one fits it all" approach. The goal is to avoid obvious slowdowns and balance the available resources (I/O bandwidth, memory and CPU) [16].

2. Database Performance Tuning

Every DBA has experienced a situation in which an application slows down after it has been in production for a while. But why this happens is not always evident. Perhaps the number of transactions issued has increased or maybe the volume of data has increased [17]. The overall performance of a system can be generally measured according to transaction response times, that is, the time it takes to complete a query or task, which is also the time the user must wait for the task to finish and possibly return results. Slow (long) response times translate into bad performance and frustrated users, whereas quick response times mean better performance and happy users. Database system may show different levels of performance (good or bad) at different times of the day, according to heavy or light user activity. If a user query takes a relatively long time to finish, based on previous tuning, it is an indication that system may have a performance problem that may be resolved, and this should be investigated.

Another way to know if there is a performance problem is by simply monitoring the system on a regular basis, and this is the theme of the work behind this paper.

3. The ARH_Db_Tuner System

The structure of the system include buffer cache, database blocks tuning, input/output tuning, operating system tuning, shared pool and others. It is out of the scope of this paper to include the flowcharts of these.

ARH_Db_Tuner consists of three functional parts, each is interfaced through a number of screens, the first part is dedicated for testing the system, and it is interfaced by Five screens that the user can move between by selecting the appropriate tab. Each screen is related to different category.

The second part is for displaying the various options to generate graph just on one click. There are for different charts available for some major SGA Parameter.

The Third functional part is the Auto/Manual Refreshing of Charts and Table data.

The user interface screens are available to perform system tests, these screens are used for:

Shared pool: The first screen "Shared Pool" is divided into three database blocks. These are: Library cache—in which three records were retrieved, these are: Gethitratio, Pins and Reloads. When the button “Calculate reloads–to–Pins Ratio” is pressed the system will display the value of “sum(reloads)/sum(pins)”. This value is analyzed and the result is displayed in a separate window as either "good" or "bad” alongside some recommendation that will appears if the analysis is bad. This is to inform the user about the necessary actions that must be taken to improve the system. In addition to the recommendation, some brief comments that are related to the monitored database system status will be displayed also as shown in figure 2 below. Shared pool reserved area—in which two records are retrieved; these are: "Request Misses" and "Request Failures". These values are displayed when the V$SHARED_POOL_RESERVED” button is pressed. Data dictionary cache—in which three values are retrieved; these are: "Parameter", "Gets" and "Getmisses". When the user clicks on “sum of GETMISSES to

sum of GETS" button, OPMT calculates the percentage of the sum of "GETMISSES to the sum of GETS" and checks if this value is less than 15 per cent, then it will show some analysis regarding this calculated value.

Redo Log buffer: The Third screen "Redo log buffer" is used for two system testing functionalities; these are:Session Wait. This is used to check the length of system waiting period in seconds. Pressing on "Ckeck Seconds in Wait" button, as shown in the upper part of figure 4 below, four pieces of information will be retrieved and displayed. These are SID "Session ID", the Event, the waiting period in Seconds, and State of the system. If the waiting period is relatively too long, the system will display "Bad" in the "Analysis" window in the left bottom corner of the screen. It also gives the necessary recommendations to deal with this problem and some comments too. System Status "Sysstat". This to check for the value of redo buffer allocation retries. Again, if it is relatively large the system will show the results of the analysis, recommendation and comments too.

Input/Output: The third user interface is the "Input/Output" screen and it is about scan statistics. This screen is mainly used for retrieving and displaying two values, namely: the name and the value of the scanned table. Again, if the value is undesirably high, the system will display "Bad" in the Analysis window and recommendations and comments will be given too. This screen is shown in figure 5 below.

Undo Segments. This is the forth screen which is used for checking the number of deleted values. The user can get the number of deleted values by clicking on the "Check if number of waits > 1% number of requests" button. The desirable value must be less than 1% of the total number of requests. If this is the case, the system will display "Good", otherwise, the user will get "Bad" result as described above. This is shown in figure 6 below.

The graphical representation of the statistics has been show in figure 7, 8, 9 and 10. All graphical representation windows are having a Auto Refresh buttons. On clicking on these buttons the graphs will refresh automatically after 5 seconds and display the real time statistics of the database.

Fig. 1. Login Window

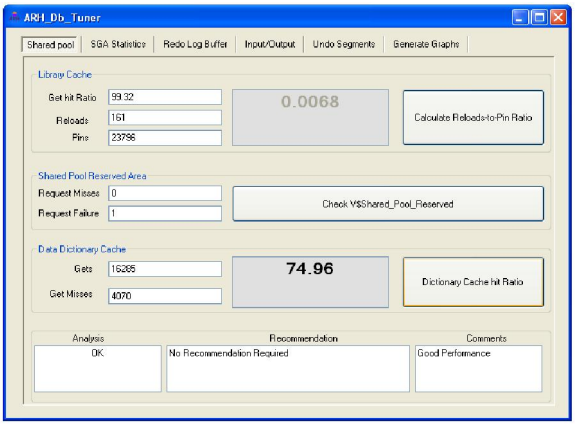

Fig. 2. Shared Pool Statistics Window

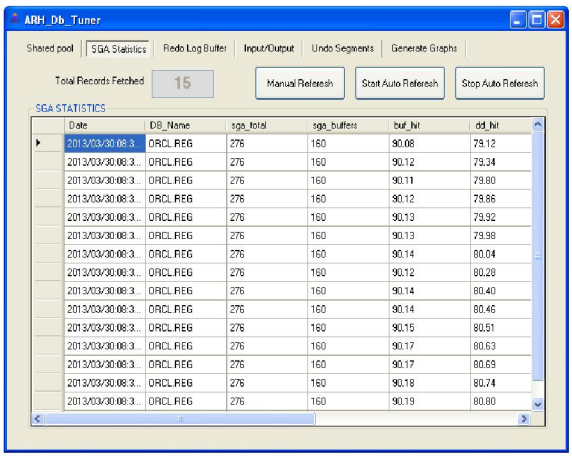

Fig. 3. SGA Statistics Window

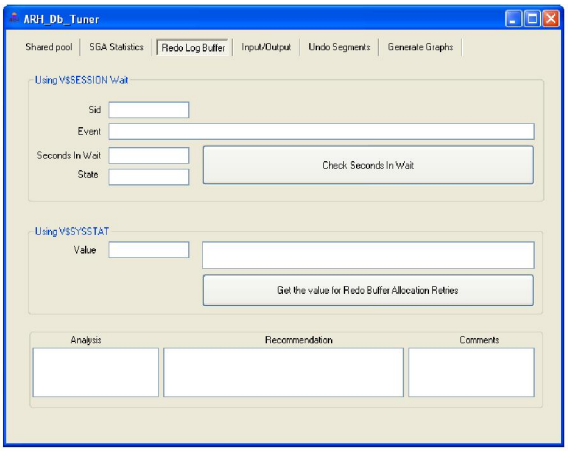

Fig. 4. Redo Log Buffer Statistics Window

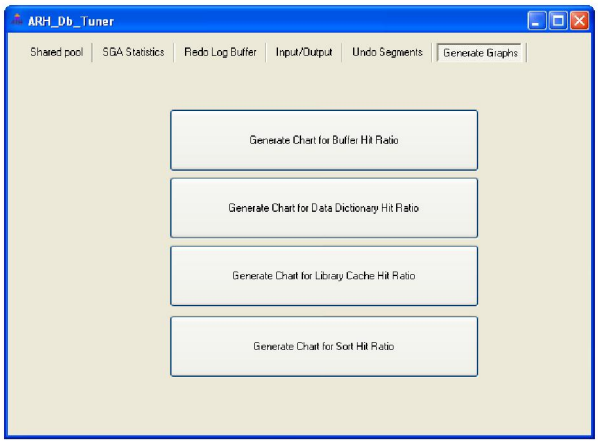

Fig. 7. Graph Generation Window

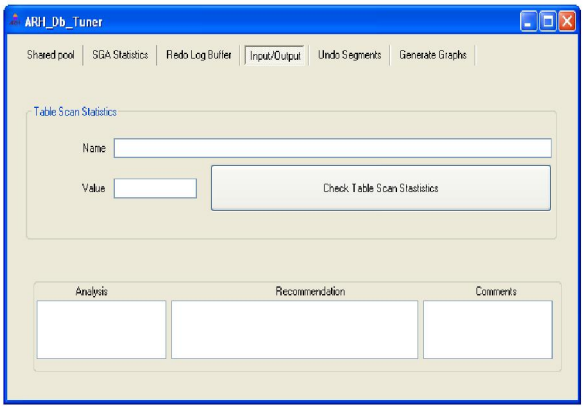

Fig. 5. I/O Statistics Window

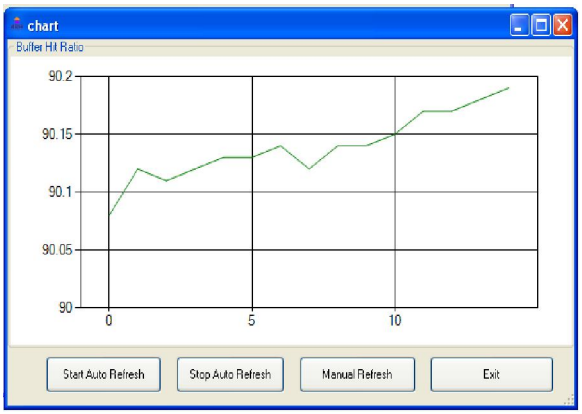

Fig. 8. Buffer Hit Ratio Graph Window

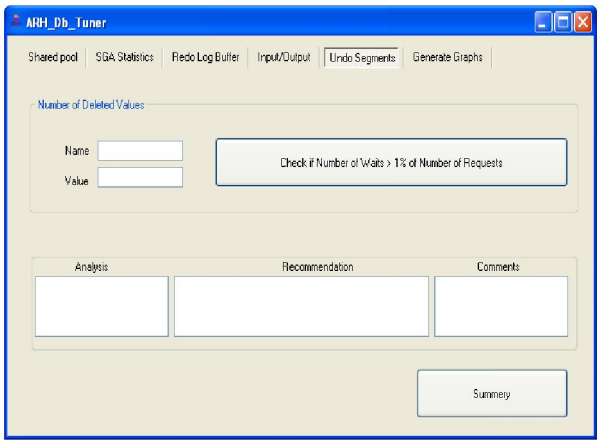

Fig. 6. Undo Segment Statistics Window

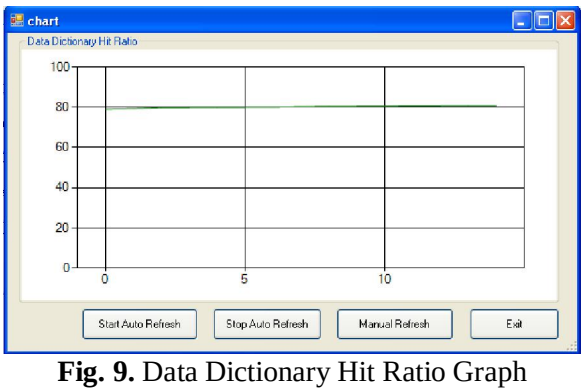

Fig. 9. Data Dictionary Hit Ratio Graph Window

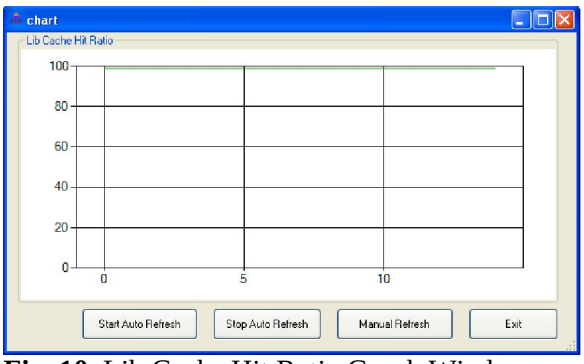

Fig. 10. Lib Cache Hit Ratio Graph Window

4. Conclusion and Future Work

Tuning the database can become quite complex, but modern databases offers the administrator an unparalleled ability to control the PGA and SGA. Until old databases evolve into a completely self-tuning architecture, the DBA was responsible for adjusting the dynamic configuration of the system RAM. Automated SGA adjustment scripts can be used to allow the DBA to grow and shrink the SGA regions. Manual tuning cost more for an organization but it is one of the major needs for an organization to attract the customer. So we have proposed a solution to fulfill the need of an organization in the shape of this Automation Framework. This framework will not take any cost and it will give faster result compare to manual tuning.

ARH_Db_Tuner is a database performance monitoring and tuning system that was developed to be used with ORACLE Database system run on Microsoft Windows operating systems. The main objective of speeding up database systems performance tuning processes and simplifying the DBA duties in a reliable and flexible manner was achieved. The system has been tested and proved efficient and reliable. ARH_Db_Tuner provides a flexible structure that can be further developed with minor changes. It provides friendly interfaces that can be easily used by the DBA and developers to monitor database performance. All test information and produced history files for each performance test are stored and can displayed and examined at any time.

The system design is flexible and can be easily expanded. Future expansions and changes may include:

- Make the system multi-lingual.
- Expand it to work on other database management systems like MS-SQL Server, DB2.
- Make it to run on other operating systems, like UNIX and Linux.

References

[1] Lightstone, S. *et al.*, "Toward Autonomic Computing with DB2 Universal Database", *SIGMOD Record*, Vol. 31, No. 3, September 2002.

[2] Xu, X., Martin, P. and Powley, W., "Configuring Buffer Pools in DB2 UDB", IBM Canada Ltd., the National Science and Engineering Research Council (NSERC) and Communication and Information Technology Ontario (CITO), 2002.

[3] Chaudhuri, S. (ed). Special Issue on, "Self-tuning Databases and Application Tuning", IEEE Data Engineering, *Bulletin 22(2)*, June 1999.

[4] Bernstein, P. *et al.*, "The Asilomar Report on Database Research", ACM SIGMOD Record 27(4), December 1998, pp. 74 - 80.

[5] Nguyen, H. C., Ockene, A., Revell, R., and Skwish, W. J., "The role of detailed simulation in capacity planning". IBM Syst. J. 19, 1 (1980), 81-101.

[6] Seaman, P. H., "Modeling considerations for predicting performance of CICS/VS systems", IBM Syst. J. 19, 1 (1980), 68-80.

[7] Foster, D. V., Mcgehearty, P. F., Sauer, C. H., and Waggoner, C. N., "A language for analysis of queuing models", *Proceedings of the 5th Annual Pittsburgh Modeling and Simulation Conference* (Univ. of Pittsburgh, Pittsburgh, Pa., Apr. 24-26). 1974, pp. 381-386.

[8] Reiser, M., and Sauer, C. H., "Queuing network models: Methods of solution and their program implementation", *Current Trends in Programming Methodology*. Vol. 3, Software Modeling and Its Impact on Performance, K. M. Chandy and R. T. Yeb, Eds. Prentice-Hall, Englewood Cliffs, N. J., 1978, pp. 115-167.

[9] Borovits, I., and Neumann, S., "Computer Systems Performance Evaluation", D.C. Heath and Co., Lexington, Mass., 1979.

[10] Enrique Vargas, "High Availability Fundamentals", Sun BluePrints™ OnLine, November 2000,
http://www.sun.com/blueprints

[11] Harry Singh, "Distributed Fault-Tolerant/High-Availability Systems", Trillium Digital Systems, a division of Intel Corporation, 12100 Wilshire Boulevard, Suite 1800 Los Angeles, CA 90025-7118 U.S.A. Document Number 8761019.12.
[12] David McKinley, "High availability system. High availability system platforms", Dedicated Systems Magazine - 2000 Q4 (http://www.dedicated-systems.com)
[13] Sasidhar Pendyala, "Oracle's Technologies for High Availability", Oracle Software India Ltd., India Development Centre.
[14] James Koopmann, "Database Performance and some Christmas Cheer", an article in the Database Journal, January 2, 2003.
[15] Frank Naudé, "Oracle Monitoring and Performance Tuning", http://www.orafaq.com/faqdbapf.htm
[16] Michael Marxmeier, "Database Performance Tuning", http://www.hp-eloquence.com/support/misc/dbtuning.html
[17] Sharma H., Shastri A., Biswas R. **"**Architecture of Automated Database Tuning Using SGA Parameters", *Database System Journal*, Romania, 2012
[18] Sharma H., Shastri A., Biswas R. "A Framework for Automated Database TuningUsing Dynamic SGA Parameters and Basic Operating System Utilities", *Database System Journal*, Romania, 2013
[19] Mihyar Hesson, "Database performance Issues"
[20] PROGRESS SOFTWARE, Progress Software Professional Services, http://www.progress.com/za/services/index.ssp
[21] Ralph Kimball,http://www.informatik.uni-trier.de/~ley/db/indices/atree/k/Kimball:Ralph.html
[22] Information Builders, "OLAP–Online Analytical Processing–Tools", http://www.informationbuilders.com/olap-online-analytical-processing-tools.html

About Authors:

Hitesh KUMAR SHARMA, The author is an Assistant Professor in University of Petroleum & Energy Studies, Dehradun. He has published 8 research papers in National Journals and 5 research paper in International Journal. Currently He is pursuing his Ph.D. in the area of database tuning**.**

Aditya SHASTRI, Ph.D. MIT, Published about 200 research papers in international journals on Graph Theory with applications in Communication, Computer Graphics and Parallel Processing ,Vice Chancellor, Director, BanasthaliUniversity, Banasthali, INDIA

Ranjit BISWAS, Head and Professor Jamia Hamdard (Hamdard University) Published about 100 research papers in International journals/bulletins.

Reverse Engineering in Data Integration Software

Vlad DIACONITA
The Bucharest Academy of Economic Studies
diaconita.vlad@ie.ase.ro

Integrated applications are complex solutions that help build better consolidated and standardized systems from existing (usually transactional) systems.
Integrated applications are complex solutions, whose complexity are determined by the economic processes they implement, the amount of data employed (millions of records grouped in hundreds of tables, databases, hundreds of GB) and the number of users [11].
Oracle, once mainly known for his database and e-business solutions has been constantly expanding its product portfolio, providing solutions for SOA, BPA, Warehousing, Big Data and Cloud Computing. In this article I will review the facilities and the power of using a dedicated integration tool in an environment with multiple data sources and a target data mart.
***Keywords**: ODI, reverse engineering, SOA, data mart*

1 Data integration software

There are many software packages that can aid to system integration. One of them is the Oracle Data Integrator (ODI) that has its roots in the acquisition of Sunopsis by Oracle in 2006. The primary use of ODI is to move and transform data from one place to another so there are many other Oracle products that are benefiting and thus creating the need of using this tool in many organizations.
Like shown in [7] data Warehouses and data marts (a subset of the data warehouse that is usually oriented to a specific department) are data-intensive systems that are used for analytical tasks in businesses such as analyzing sales/profits statistics, cost/benefit relation statistics, customer preferences statistics, etc. The idea of a data warehouse is to extract data from operational databases and to store them separately. The justification for this approach is that OLAP largely deals with condensed data, thus does not depend on the latest updates by transactions. Furthermore, OLAP requires only read-access to the data, so the separation of the data for OLAP from OLTP allows time-consuming transaction management to be dispensed with.
The need for data warehouse usually comes from the fact that the operational systems cannot be overload with the additional queries required by the business intelligence needs. Integrating data from multiple, usually heterogeneous sources, is a problem addressed by ODI.
Linking SOA with ODI and Data Warehousing (DW) can be beneficial. From a business perspective the key word in dealing with SOA is definitely flexibility. Companies must be able to keep pace with the rapid changing conditions of the business environment. In the same time the trend in IT architectures leads toward an integrated model by building business processes that span multiple operational systems and by enabling interoperability between legacy systems and newly developed systems [3].
As shown in [4] SOA has the merit to introduce a new kind of technological "democracy" where the application systems are considered a federation per se, thus opening the doors to a new kind of logical distributed computing approach where the technological platforms are downgraded to the implementation or physical level.
Most SOA efforts have centered on transaction systems, but data warehousing can benefit from SOA with the ability to join various actions (services) from different areas of the DW to create

composite applications or common services. The services that are part of a data warehouse such data extraction, transformation, loading, querying or updating should be part of the SOA from the start. This should make more comprehensive business intelligence possible, and could assist in the development of fully integrated SOA [2]. The usage of SOA makes data location of little importance to its users so using it in conjunction with ODI processes and transformations makes lot of sense. SOA can enable an abstract layer that makes data available inside an enterprise using homogeneous services.

ODI can also have a role in normalizing data definition so different applications reference the same data. It can also have a role in processing Big Data by delegating and distributing processing.

2. ETL and ELT

As an alternative to ETL, ODI proposes ELT which extracts data from the source, loads it into the target and processes there by using SQL the needed transformations.

This approach exploits database optimizers as opposed to transformation that is performed in-flight or requiring a separate intermediary. The ELT approach directly impacts performance and has proven to make data loading fast, efficient and incredibly reliable [9].

An ELT has the ability to manage a staging area, generate code and execute instructions on target systems but also on source systems, such systems that are being managed by any DBMS.

The components of ODI architecture are the repository, the studio, the graphical interface of the software, the agents and the console. In SOA, repositories are used to manage services and support service discovery at runtime. Usually, there is a process needs a data access service (DAS) to execute a query usually against a database. DAS are variations of the ordinary service concept; they are more data-intensive and designed to expose data as a service [5]. In an ELT solution the data is not stored in the repository; it is moved directly to the target. The repository is usually stored in a schema of an existing database and is composed of the master repository for the sensitive data and the work repository for the data needed by the developers. In a production environment an execution repository is also present. It stores only the operational metadata. The exchange of data can be done through versioning or by importing or exporting XML files. There are two types of agents: the standalone agent that can be installed on any platform and the JEE agent that runs on a Weblogic server. Usually the JEE agents has the role to distribute execution requests and balance load across different other agents which are usually standalone agents. A strategy using only standalone agents is also possible.

The key elements of ODI are Execution Contexts, Knowledge Modules, Models, Interfaces, Packages and Load Plans.

For assuring independence from the physical location of the data, logical schemas can be used, at execution this are translated into physical ones so the maintenance of the connection parameters, location of the databases, and schema names is entirely independent of the code itself.

Metadata can be imported using knowledge modules from applications, where objects usually are representation of the data or from databases using models. After it's imported, metadata can be enhanced in ODI (for example by adding constraints). Also new metadata can be created.

Another key element in ODI is the interfaces where the transformations are built. An interface contains among others description, mappings and flows.

Packages put together elements such as interfaces, variables and procedures. They are compiled into scenarios which execution can be organized with load plans.

3. Reverse-engineering the model metadata

Like shown in [1], to use ODI, first we declare a new data server in the ODI Physical Architecture and then a reference to a Physical Schema located on that server that holds the business data. We can also construct a work schema for every physical schema to store temporary data.

A Physical Schema definition needs to be associated with a Logical Schema name which will be exclusively used to build ODI models. The models are abstracted, independent of the data source but seem homogeneous to the developer. Usually they are built by reverse-engineering the structural data, that coming from databases, flat files, XML files or different ERP systems. Different other elements, such as structural integrity data can be added to the captured metadata.

Like shown in [6], the **Reverse-Engineering Knowledge Modules** role is to perform customized reverse engineering for a model. It connects to the application or metadata provider then transforming and writing the resulting metadata into Oracle Data Integrator's repository. The metadata is written temporarily into the SNP_REV_xx tables. The RKM then calls the Oracle Data Integrator API to read from these tables and write to Oracle Data Integrator's metadata tables of the work repository in incremental update mode (figure 1).

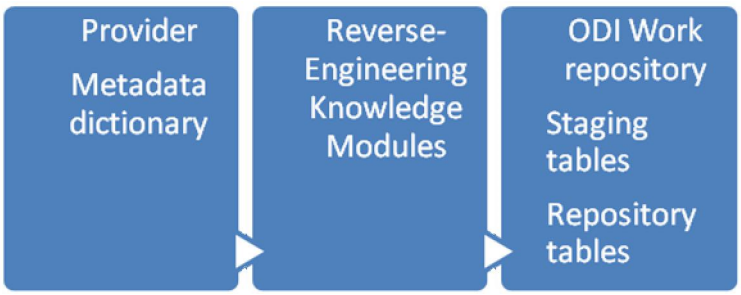

Fig. 1. RKM

As shown in [7], Oracle Data Integrator implements five different types of KMs. Each of them covers one phase in the transformation process from source to target. The three most important types of modules are the *integration knowledge module* (**IKM**), the *loading knowledge module* (**LKM**), and the *check knowledge module* (**CKM**). As explained in [1] when a Knowledge Module has been previously imported into the parent project and applied to the interface target and the interface is subsequently executed, it is the steps within the **IKM** that determine the what, how, and when data is moved into the target data store. **LKMs** load data into the staging area from other servers. If the source data is in the same server as the staging area then **LKM** is not needed. **CKMs** are used to check and enforce data integrity through testing conformance to constraints and references, either statically on data tables on source or target systems, or dynamically during the process of a data flow defined in an ODI interface.

To reverse engineer, after defining the topology, we create the model like shown in figure 2.

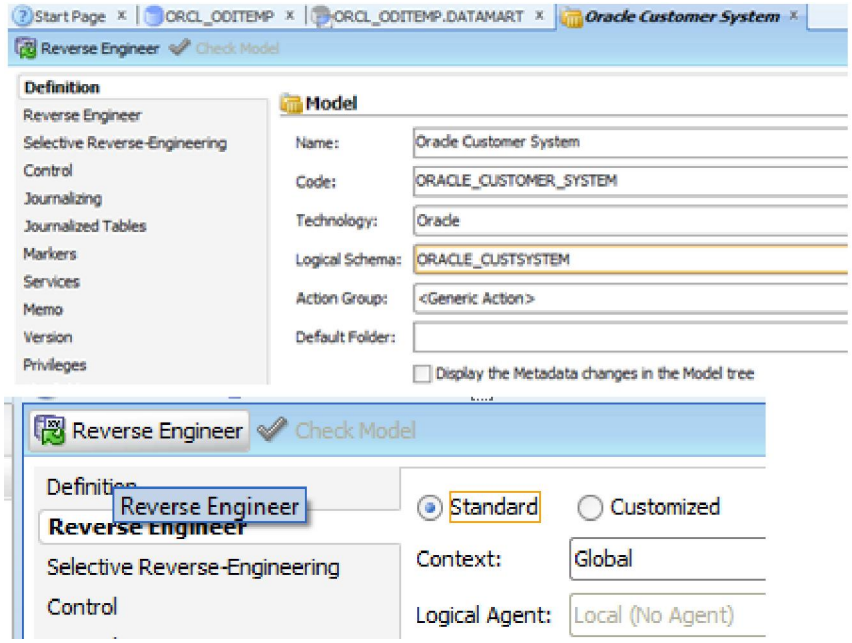

Fig. 2. Create the model

After configured the ODI representations of our data objects an interface can be built to move and transform the data into the data mart, like shown in figure 3.We can use the Automatic Mapping and also do manual mappings and transformations using SQL code and apply some timestamps useful for audit purposes. The transformed data will be loaded from an Oracle source into a target Oracle data mart.

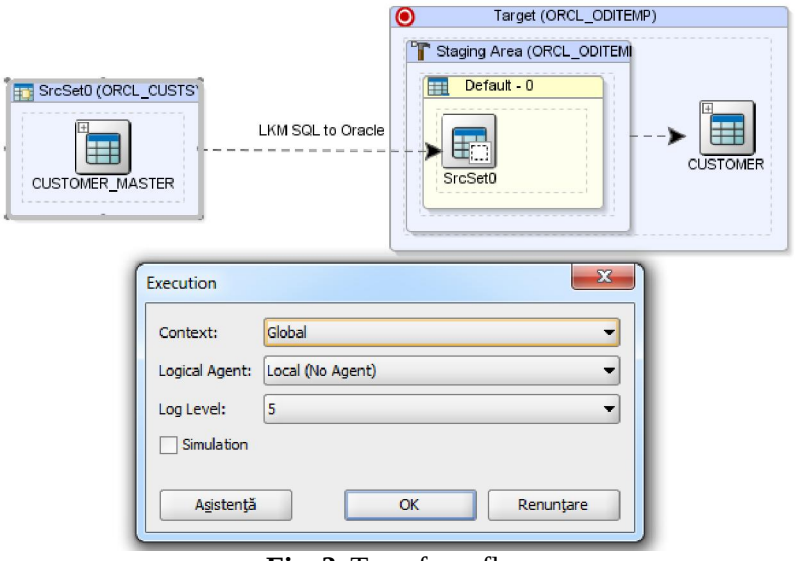

Fig. 3. Transform flow

We can add some additional complexity by introducing joins and lookups, heterogeneous data sources and data aggregation. And we can do this with multiple source databases usually using different JDBC connectors (figure 4).

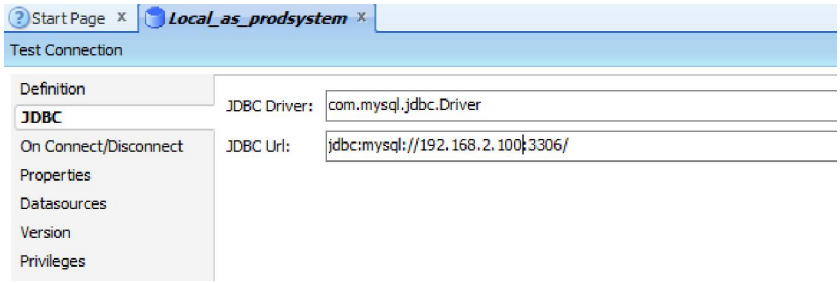

Fig. 4. MySql JDBC Driver

In the example shown in figure 5 we join three sources (actually we have 2 instances of the same data source table) and link them to the target data mart.

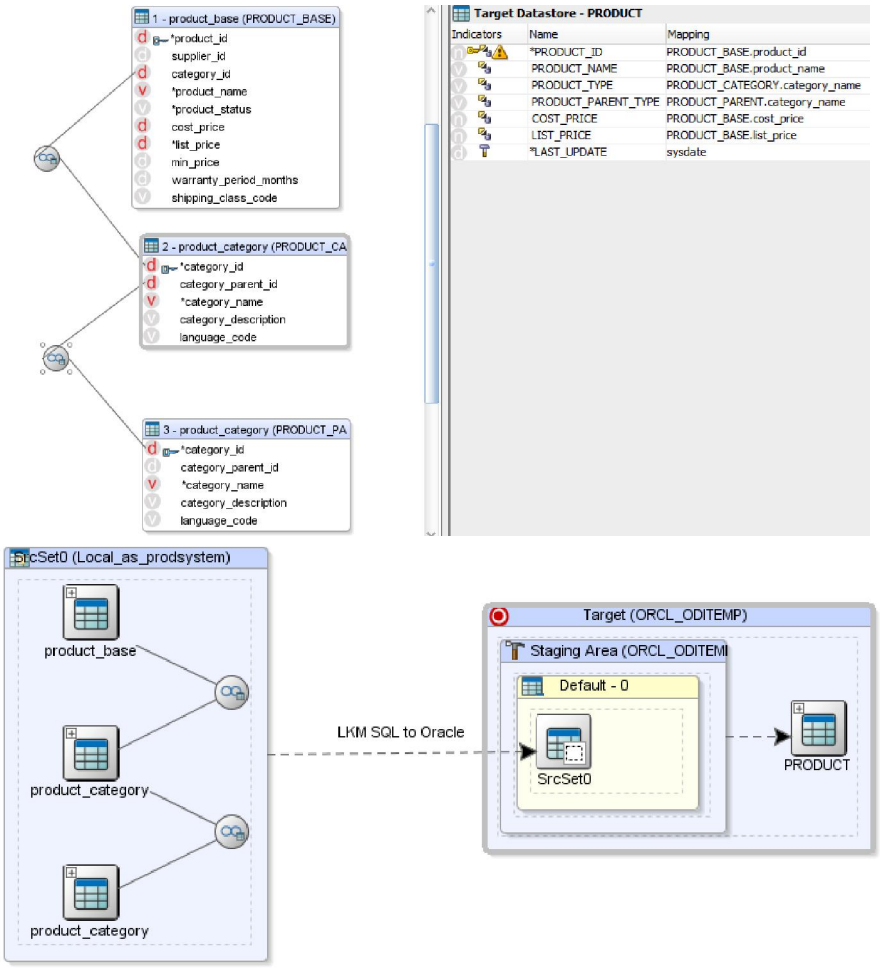

Fig. 5. Join transform flow

Like shown in figure 6, an interface can be built to move data using JDBC from a third-party DBMS to an Oracle data mart.

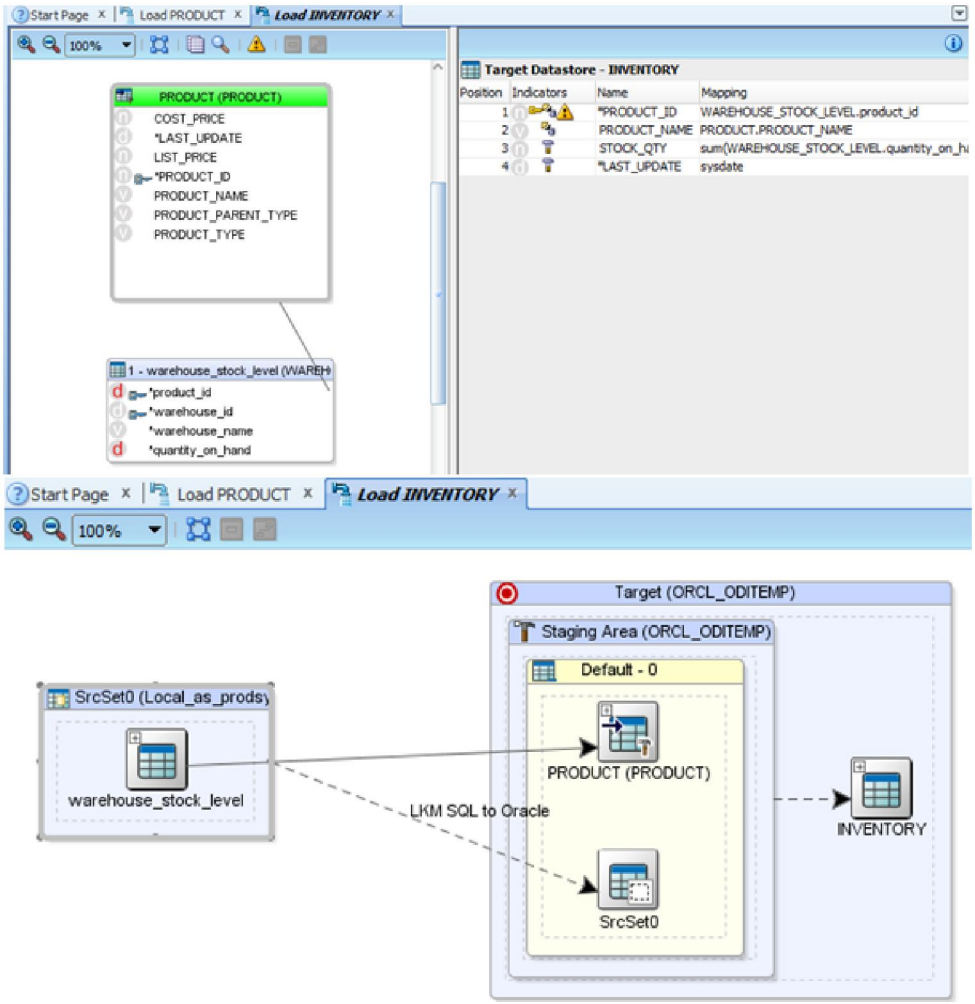

Fig. 6. Lookup transform flow

As shown in [10] eXtensible Markup Language (XML) is a platform-independent format for representing data and was designed as a standard for information exchange over the Internet. XML enables easy exchange of information, which allows interoperability between applications due to data encapsulation with metadata. To use XML inside ODI a JDBC driver is required which is available out-of-the-box (figure 7).

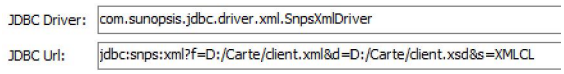

Fig. 7. XML JDBC

Let's look at the following XML file:

```
<CLIENT>
  <ID_CLIENT>10</ID_CLIENT>
  <FIRST_NAME>Jan</ FIRST_NAME >
  <LAST_NAME>Roberts</ LAST_NAME >
  <CREDIT>600</ CREDIT >

<EMAIL>Ishwarya.Roberts@LAPWING.COM</E
MAIL >
  <DATE_OF_BIRTH>21-MAR-44</
DATE_OF_BIRTH >
  <CIVIL_STATUS>single</ CIVIL_STATUS
>
  <SEX>F</SEX>
  <INCOME >G: 130,000 - 149,999</
INCOME >
 </CLIENT>
```

This would reverse-engineer to a table called CLIENT that has the following columns: ID_CLIENT, FIRST_NAME, LAST_NAME etc. The mapping is shown in figure 8.

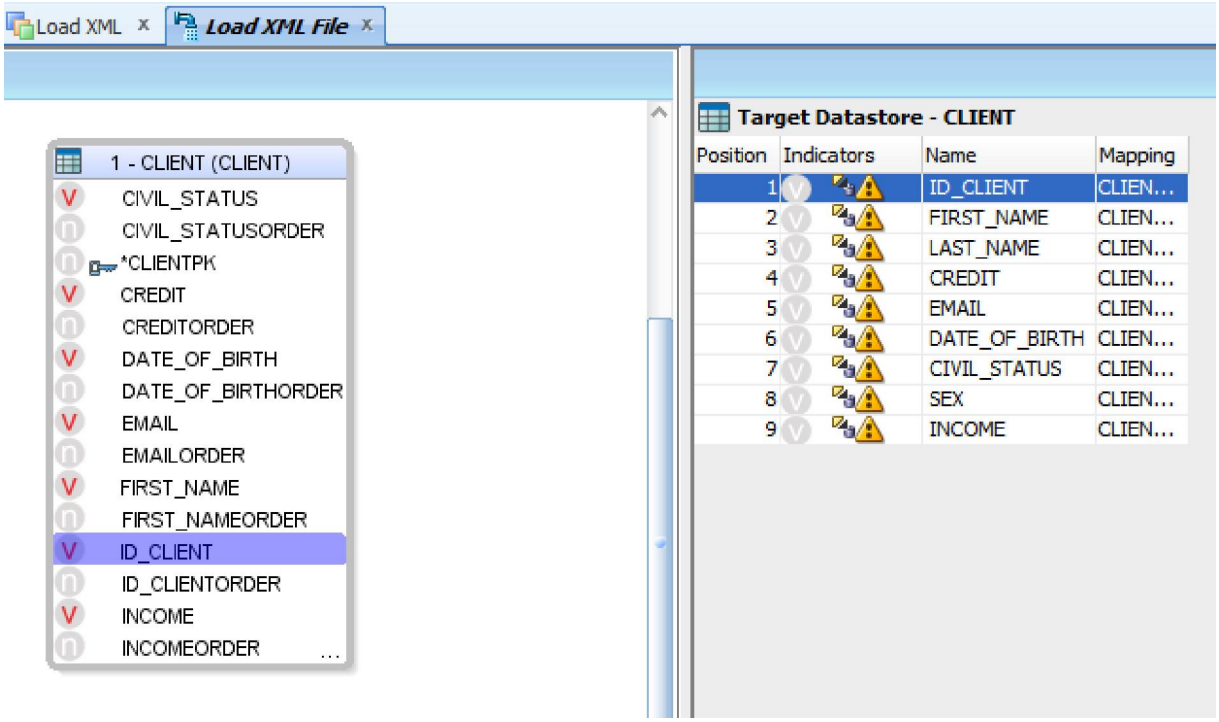

Fig. 8. XML-Relational Mapping

Conclusions

Data integration is very often a necessity in bigger projects. Using an integrated tool can be much more powerful and useful in projects that imply using data from heterogeneous sources, targets, and applications. Such products provide great aid in integrating databases, ERPs, CRMs, B2B systems, flat files, XML data, LDAP, JDBC or ODBC. It also can help in cutting hardware costs through improved utilization and high-performance data integration. Using external services for data integration and by deploying data services and transformation services that can be integrated within an SOA infrastructure Also, SOA business processes can assign large data operations to Oracle Data Integrator by using web services.

References

[1] Peter C. Boyd-Bowman, Christophe Dupupet, Denis Gray, David Hecksel, Julien Testut, Bernard Wheeler, *Getting Started with Oracle Data Integrator 11g: A Hands-On Tutorial*, May 2012

[2]http://www.information-management.com/news/1065308-1.html

[3] Alexandra Florea, Anca Andreescu, Vlad Diaconita, Adina Uta, "Approaches Regarding Business Logic Modeling in Service Oriented Architecture", *Informatica Economică* vol. 15, no. 3/2011

[4] Cătălin Strîmbei, "Smart Data Web Services", *Informatica Economică* vol. 16, no. 4/2012

[5] M. Turner, D. Budgen, P. Brereton, "Turning software into a service", *Computer*, 36 (2003), pp. 38–44

[6]http://docs.oracle.com/cd/E15586_01/integrate.1111/e12645/intro.htm

[7] Uli Bethke, *Developing a Knowledge Module in Oracle Data Integrator*, 2009

[8] Jane Zhao, Hui Ma, "ASM-based design of data warehouses and on-line analytical processing systems", *Journal of Systems and Software*, Volume 79, Issue 5, May 2006, Pages 613–629

[9]http://www.oracle.com/technetwork/middleware/data-integrator/learnmore/odi-for-soa-wp-1555852.pdf

[10] Iuliana Botha, "Managing XML Data to optimize Performance into Object-

Relational Databases", *Database Systems Journal* vol. II, no. 2/2011

[11] Vlad Diaconita, "Hybrid Solution for Integrated Trading", *Informatica Economică* vol. 14, no. 2/2010

Vlad DIACONIŢA is a lecturer at the Department of Economic Informatics and Cybernetics, within the Faculty of Economic Cybernetics, Statistics from the Bucharest Academy of Economic Studies. He has graduated the faculty at which he is now teaching in 2005 and since 2010 holds a PhD in the field of Cybernetics and Statistics. He is the co-author of 3 books in the domain of economic informatics, and of more than 25 papers in journals and conference proceedings. He participated as team member in 3 research projects that have been financed from national research programs. He is a member of the IEEE Computer and INFOREC professional associations. Domains of competence: Database systems, Data warehouses, OLAP and Business Intelligence, Integrated Systems, SOA.

Introducing Virtual Law offices in the Existing Judiciary

Rizul KHANNA[1], Rahul BHAGAT[2]
[1]Dept. of ECE, [2]Dept. of ECE, NIT, Jalandhar
Punjab, India
rijulkhanna.ec.09@nitj.ac.in, rahulbhagat.cs.10@nitj.ac.in

With the rise of omnipresent foundation of computing resources over the past years, every IT Setup is expanding their horizons in the Cloud services and related technologies. Cloud provides dynamically scalable virtualized computing resources as a service over the Internet and this key characteristic differentiates it from traditional computing paradigm. It is the application of Cloud and mobile computing technologies for improving communication between lawyers, their clients, and any other person involved. This framework basically digitalizes the existing judicial file system to form an e-library using Clouds infrastructure and using the internet services like GPRS or GSM / CDMA or 3G/4G etc for information retrieval from this e-library after the request has received proper authorization and authentication of the regulatory body. Data is received from both Web and mobile based applications so that each person can access the judicial data anytime and anywhere. For the realization of this system, a web interface is created with e-library serving as the main database and with user-friendly interface to do the above data acquisition and analysis which ultimately gives pace to the slow process of case management.

Keywords: *Applications, Cloud Computing, Data Interaction, Internet, E-library Regulatory body, Security*

1 Introduction

The past few years have seen a great and an increasing thrust in computing technologies .In today's life users prefer to use web services for access to their data and applications through an ISP rather than manually visiting the service provider. Today, people use communication devices such as mobile and PC's for more than conversation purpose, they are also heavily demanded in various other areas like accessing Web services, video streaming through Internet etc. For example: Banking system is digitized, a user can access all of its banking services such as withdrawing, transferring money by simply accessing websites through their PC or mobile anywhere anytime in the world. Particularly in web services, this is especially true with the emergence of digital phones and Tablets using WAP (Wireless Application protocol) in order to provide cell phone access to special websites [1]. Web services, besides providing an easy access to various services, bring with them a great deal of Transparency. Cloud computing is a 'metaphor' for internet. Cloud computing is basically storing of our data on someone else's Server, so rather than using our own hard– disk or maintaining our own Server within our office, we will use a third party server [2]. There is no need to install or store anything on local computers .This frequently takes the form of web-based tools or applications that users can access and use through a web browser as if they were programs installed locally on their own computers. The nice thing about Cloud computing is that it is Cost effective, you don't have to own the server hardware. It gives you flexibility because you can access your data anywhere as long as you have an internet connection .It makes you more Nimble. Cloud computing allows you to compete on a larger scale with other larger firms that have these resources to maintain these servers and purchase these software and pay licensing fee over and over again. So by using Cloud

resources, one can pay only a subscription fee every month instead of purchasing the software.

A. Some of the Services offered by Cloud

1) Infrastructure as a Service (IaaS): Cloud consumers directly use IT infrastructures provided in the IaaS Cloud. Virtualization is extensively used in IaaS Cloud in order to integrate/decompose physical resources in an ad-hoc manner to meet growing or shrinking resource demand from Cloud consumers. An example of IaaS is Amazon's EC2 which allows users to rent virtual computers on which to run their own computer applications [3].

2) Software as a Service (SaaS): Cloud consumers release their applications on a hosting environment, which can be accessed through networks from various clients by application users to achieve economies of scale and optimization in terms of speed, security, availability, disaster recovery, and maintenance .It can be accessed by the customers on pay per use basis. Examples of SaaS include SalesForce.com, Google Mail, Google Docs etc.

3) Platform as a Service (PaaS): It is a development platform supporting the full *Software Lifecycle* which allows Cloud consumers to develop Cloud applications (e.g. SaaS) directly on the PaaS Cloud. Google App Engine is its famous known type.

4) Data as a Service (DaaS): The delivery of virtualized storage on demand becomes a separate Cloud service - data storage service called DaaS, could be seen as a special type IaaS. DaaS allows consumers to pay for what they are actually using rather than for the entire database. Some DaaS offerings provide table-style abstractions that are designed to scale out to store and retrieve a huge amount of data within a very compressed timeframe, often too large, too expensive or too slow for most RDBMS to cope with.

B. Types of Cloud Models

1) Private Cloud: Computing architecture of this Cloud is dedicated to the customer and is not shared with other organizations and managed by the organization or a third party regardless whether it is located premise or off premise.

2) Public Cloud: The public Cloud is used by the general public Cloud consumers and the Cloud service provider has the full ownership of the public Cloud with its own policy, value, and profit, costing, and charging model. The customer has no visibility over the location of the Cloud computing infrastructure. The computing infrastructure is shared between organizations. Many popular Cloud services are public Clouds including Amazon EC2, S3, Google App Engine, Force.com, etc.

3) Community Cloud: Several organizations jointly construct and share the same Cloud infrastructure as well as policies, requirements, values, and concerns. The Cloud community forms into a degree of economic scalability and democratic equilibrium. For example, all the government agencies in a city can share the same Cloud.

4) Hybrid Cloud: The Cloud infrastructure is a combination of two or more Clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability. Organizations use the hybrid Cloud model in order to optimize their resources to increase their core competencies by margining out peripheral business functions onto the Cloud while controlling core activities on-premise through private Cloud.

2. Problem Statement

Till date, the above mentioned services of Cloud computing had emerged in fields of Hospitality management system, Machine monitoring system etc. Apart from these, there exists a system that lacks digitalization. That is the judicial system.

This system, mainly in developing countries, still involves a large scale paper or file work, administered manually and not much of transparency is provided to the end user. **How do we use Cloud resources for above mentioned aspects of judicial system?** [4].

The Solution: Remote monitoring by specialized Lawyer through communicating devices such as mobiles and Tablet PC's. Whole of courts documents will be stored on servers provided by the Cloud known as e- library (IaaS) and anyone who tries to access these files will use a software (SaaS) that will be designed dedicatedly for this purpose. The person accessing these files could be a judge, a lawyer or a client. There will be a regulatory body that will provide different access controls to different users of the e-library. Following is the prototype of above framework in a sequential manner:

(i). Courts existing library, which is based on contemporary file system, is transformed into e-library using the infrastructure provided by the Cloud which will form a Digitalized database. Proper indexing of these files should be maintained according to some meaningful criteria. This criteria could be the year in which the case was filed, type of case (i.e criminal, civil case etc.), or the extent up to which it has been resolved [5].

(ii) A dedicated software service (SaaS) will be provided by any web hosting firm such as Amazon Web Services (AWS) to various users of e-library, to access the documents in it. Users will use a web browser available on any type of communicating device or PDA (Personal Digital Assistant) such as Tablet PC's or Mobile phones etc. to access this data at any time and at anyplace with a greater speed through a strong internet connection [6].

(iii) A Regulatory body will be placed in between the end user and e-library that will check the access control available on the database to different type of users i.e. lawyers, Clients or judges [7]. It will decide different privileges such as Reading, Writing and Updating on the database to different users trying to access the e-library [8].

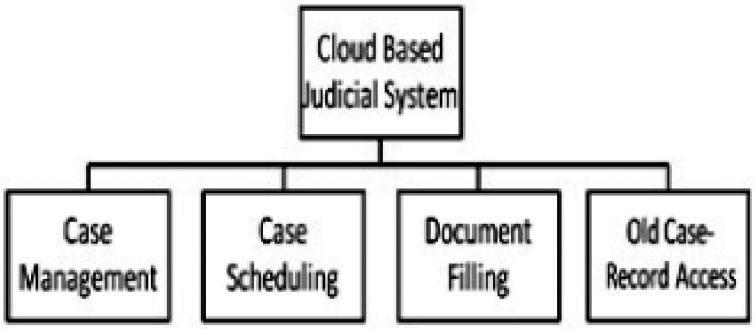

Fig.1. Prototype of the system

3. Proposed Solution

1) e-library: First of all we need to develop a database that will form the platform for the services we want to give to the end user [7]. This will be provided by using the IaaS service offered by the Cloud. We will either develop a Cloud of our own or use the services of some other third party public Cloud by paying a subscription fees every month or year. What we will do initially is that we will take all the existing file system records available with the courts and different lawyers about different type of cases and put the whole file system on Cloud's infrastructure. This might take long time maybe a year or so. The infrastructure we could use would be of something like that of Amazon S3 (Simple Storage Service) [9].This will develop the basic database for the various services we will provide to the users. This database will be called the e-library. Since, this

library uses the servers provided by the Cloud it will be highly flexible in terms of the amount of data it can store.

2) SaaS for end users: After we are through with our initial database development part, now we have to look upon how to access this database through the internet? For this purpose we will develop a dedicated software that will have a simple and interactive Graphical User Interface (GUI) that will be available to all the end users through the internet [10]. All of the users will have to register themselves before accessing the e-library and they will be provided unique Id's that will differ from user to user, depending on the type of user who is demanding the access. The main purpose of this software will be to fetch the data stored in e-library and display it on the user end through internet in a web browser i.e. Reading the particulars of any case from the library. Another purpose it needs to serve will be to write into the Cloud that may include updating and adding the data in the library, but this authority could not be provided to each and every user considering security issues. These services would depend upon the amount of authority available to the user who is accessing the library. One such SaaS service available is Law fusion direct software that provides hosted case management service for Lawyers (those who want to join it at will, moreover it is not for the whole judicial system).

3) Regulatory Body: Now we have both the e-library and a software service available to access it. But now the big issue arises as to how to keep track and maintain the different type of users that are trying to access the library for different purpose. Even more important part that needs to be dealt with is, unauthorized access by any user to the library. For this purpose, a regulatory body will be interleaved between the e-library and the end user. The main functions of the regulatory body will be as stated below:

*3.1) User Authentication :*When the user makes an access call for the library it will be first passed through the regulatory body ,which will check which type of user it is i.e. whether it is a case client, lawyer or a judge appointed to that case and then, it will provide the proper access control of library to the user [11].

3.2) Access Control of e-library: Once the type of user is identified it is to be decided as to which user should be provided with what amount of rights. Following will be the access rights depending on the type of users:

3.2.1) Client: A client will only have the right to Read the case that he/she has lodged or the petition filed. He will be able to see the progress of the case or case summary but will not have the authority to amend the data present in library [12].

3.2.2) Lawyer: If the user is a lawyer he/she will not only have the right to Read but also the right to write i.e. file a petition on behalf of some client and case scheduling and case management. But he/she will not have the authorization to change the contents of the library [13].

3.2.3) Judges: If the user trying to access the library will be a judge then he/she will be provided with full access control to the e-library. He will have READING, WRITING and UPDATING rights to the data available through the library. UPDATING means that judge will have the authority to amend the data present in the library. Being the supreme authority this type of user will even have the control of closing a case for access to other type of users.

4. Basic Block and Information Flow Diagrams

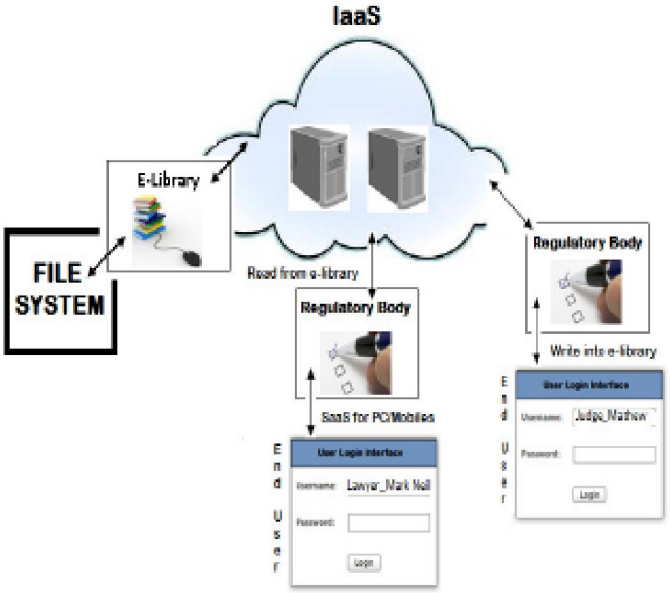

Fig.2. Information flow diagram

The above presented framework has some main components shown in figure 2.

(i). The e-library which is formed using the data present in the existing File System into the infrastructure available in the Cloud.

(ii). This library will form the main database for the web and mobile applications that want to retrieve any information from it.

(iii). A regulatory body will also have be placed between the Library and End users that is necessary from security point of view as it will give access controls to different type of user who are trying to access library for various interests of their own. Every information flow from library to the user must be authenticated from this regulatory body only then the information would be transmitted to the end user, where we have the SAAS application to access the data.

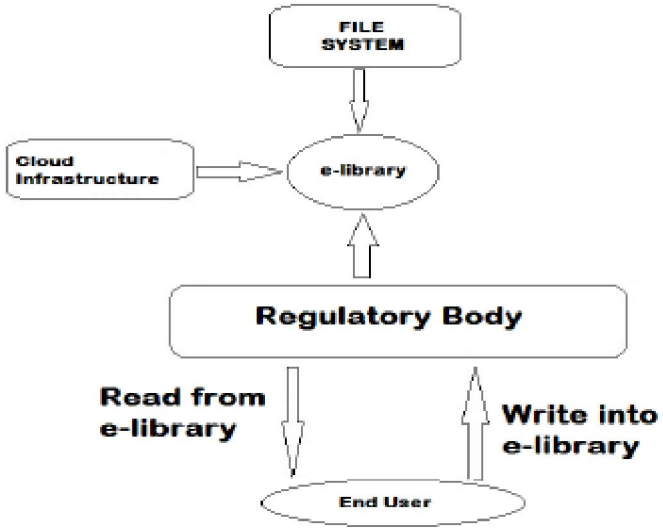

Fig.3. Block Diagram

5. Issues to be considered

A. Database provisioning of the e-library. One of the basic issue that arises while trying to implement the above scenario is that, initial storage of data in e-library or in other words we can say that, Digitalization

of data from File system into Servers of Cloud is a very time consuming and cumbersome process as there will exist millions of records that will have to be digitalize during this process to create the strong database, which the most basic requirement for this scenario to work [14].

B. Who will actually own the Cloud? Other important part is that, who will be providing the Cloud services for this purpose. As everything is confidential and a great deal of security issues will arise, the authorities access this services will have to decide a rightful Service provider. It will be a matter of trust for judicial system whether to go for a Private Cloud or a Public Cloud. Government can also setup their own Private Cloud which will assure the users of security of data to some extent [15].

C. Mobile Phone-Connectivity The access of information to users from the database available will be done both through web and mobile applications. Data transference basically is done by access the web i.e by using web services. GPRS / 3G / 4G / Wimax services must be subscribed at both ends for this type of data transfer [16].

D. Application for Computational and Work Analysis A Web application interface for automatic data transfer, and an Open source based application such as android application is made for mobiles which possesses functionality inclusively able to access data over the internet and save it in own end database, to graphically present the data for analysis with parameters of a particular case, and also to pictorially display of various inferences like crime scene photos, weapons used etc. and able to transfer results or call or messages to the app at other end to some other user. Following is a prototype of the Web interface that could help meet the requirements for users using web services:

Fig.4. Web Interface

Following could be a prototype of Android based or some other Open sources application interface that will help the mobile users to enjoy this service through mobile phones anywhere anytime:

Fig.5. Mobile Web Application

E. Estimation of data that needs to be stored in the e-library: Another major issues is to be taken care of is the estimation of data that is needed to be stored in the e-library as we are talking of a very huge amount of data the Cloud provider will have to be ready to give us that much of space on the Cloud and furthermore this will be a constantly increasing database that needs to be maintained and for this purpose estimation of data is very important before deployment of this idea.

6. Work to be carried

In order to set up a system that provides this information anytime and anywhere, there is a need to exploit latest software technologies to have SaaS and mobile application having features similar to that mentioned above.

A. Infrastructure System Required

1) SaaS Development: The personal computer and the Internet have been the two driving forces in today's businesses. The personal computer has made a lot of tasks easier and has streamlined a lot of operations within an organization while the Internet totally changed the business model of a lot of companies, because it was able to reach an audience far from what traditional media could ever hope for. Now, the Internet and personal computers are changing the landscape again by changing how software is deployed to businesses and users. The current trend is slowly moving toward Software as a Service or SaaS. Many companies are warming up to the idea of getting the same level of results but with lesser costs [17].

2) Web Server: Apache Tomcat can be used as a server for deploying the Java and web applications. The integrated environments which are used now-a-days have servers (Tomcat) already integrated in it.

3) *Algorithm for Efficient Data Searching:* We need an efficient algorithm for quick and accurate search requests from the e-library. A good search engine does not attempt to return the pages that best match the input query. A good search engine tries to answer the underlying question. If you become aware of this you'll understand why Google (and other search engines), use a complex algorithm to determine what results they should return [18].

4) Creating Mobile Application for displaying records in an organized way: Windows and Android *platforms* provide several options for you to save persistent application data [6]. The solution you choose depends on your specific needs and how much space your data requires.

7. Challenges

The implementation of above framework is bound to face following challenges:

A. Security Since the data to be stored and maintained in the library is very confidential therefore its one of the main challenges to maintain security of such a large amount of data [19].

B. Size of Data Centre Size of data centre is very large and it keeps on increasing as more numbers of cases are filed. So, it's difficult to estimate as to how much data needs to be stored in the Cloud's infrastructure.

C. Innovation To achieve a truly contiguous working, these solutions will require an unprecedented level of innovation to fully exploit the full spatial context of the surroundings from case scheduling to case management and everything else until the closing of case [20].

D. Acceptance The world wide acceptance relies on those who are in direct contact with them, and most of them will try to resist it, unless and until they are familiarized or gets acquainted with such smart technologies [21].

E. Time consumption The conversion of files into e-library takes a large amount of time which is not good to serve our purpose.

Conclusion

With the progressive trends of Cloud infrastructure, their compatibility with communication devices and more and more dependence on Data center technologies, prospects of automating case management of Judicial System has edified. With the emergence of Cloud computing as universal system of computing resources delivered as a service, implementing such a system will get a robust start. This system will try to save the response time and increase transparency. Also with arrival of WIMAX and incoming technology (4G), data rate and bandwidth limitations will be subsided enormously and further helps in establishment of this concept.

Acknowledgment

The first author would like to acknowledge the support of faculty (ECE) of NIT Jalandhar in the development of this framework mostly derived from the Bachelors of Technology final year project.

References

[1] Internet: www.it.iitb.ac.in/~sri/talks/wap01.ppt

[2] Internet: http://techterms.com/definition/cloudcomputing

[3] Internet: en.wikipedia.org/wiki/Cloud_computing

[4] Internet: http://chardjeffrey.wordpress.com/2013/01/22/practiceguidelines- for-cloud-computing-by-lawyers/

[5] Internet: http://westlawinsider.com/cyberlaw/the-lawyers-guide-tocloud- computing-part-1-of-4/

[6] Internet: http://www.youblawg.com/business-development-forlaw- firms/cloud-computing-in-financial-sector-and-legal-system

[7] Bikramjit Singh, Rizul Khanna, and Dheeraj Gujral, "Cloud computing: A need for a regulatory body," in *High Performance, Architecture and Grid Computing*, vol. 169 of *Communications in Computer and Information Science*, pp. 119-125. Springer Berlin Heidelberg, 2011.

[8] Sharon Bradley, "Ethics on the Wing: Examination of Opinions on Electronic Services and Cloud Computing", University of Georgia School of Law Library

[9] Internet: http://aws.amazon.com/s3/

[10] Internet: http://www.goclio.com/

[11] Internet: http://blogs.msdn.com/b/dachou/archive/2008/08/19/cloudcomputing- and-user-authentication.aspx

[12] Internet: http://chardjeffrey.wordpress.com/2013/01/22/practiceguidelines- for-cloud-computing-by-lawyers/

[13] Eugene Clark, Ian M. Sims," Flying into the hurricane: law, policy and practice implications of cloud computing," Int. J. of

Technology Policy and Law, 2012 Vol.1, No.2, pp.135 – 151
[14] Internet: http://judis.nic.in/
[15] James M. McCauley, Ethics Counsel, Virginia State Bar, "Cloud Computing — A Silver Lining or Ethical Thunderstorm for Lawyers?", Consultus Electronica, www.vsb.org/docs/valawyermagazine/vl0211_consultus.pdf
[16] Internet: http://en.wikipedia.org/wiki/List_of_mobile_ standards "List of`Mobile Phone Standards"
[17] Internet: http://www.saas-development.com/
[18] Internet: http://www.seomoz.org/blog/search-engine-algorithmbasics
[19] Internet: http://technet.microsoft.com/en-us/security/jj554736.aspx
[20] Jordana Hausman, "Who's Afraid of the Virtual Lawyers? The Role of Legal Ethics in the Growth and Regulation of Virtual Law Offices", Summer, 2012, Georgetown Journal of Legal Ethics
[21] Internet: http://www.netforlawyers.com/content/legal-ethicslawyers- attorneys-client-information-cloud-computing-0037

Database Security for an Integrated Solution to Automate Sales Processes in Banking

Alexandra Maria Ioana FLOREA
Economic Informatics and Cybernetics Department
The Bucharest University of Economic Studies
Bucharest, ROMANIA
alexandra.florea@ie.ase.ro

In order to maintain a competitive edge in a very active banking market the implementation of a web-based solution to standardize, optimize and manage the flow of sales / pre-sales and generating new leads is requested by a company. This article presents the realization of a development framework for software interoperability in the banking financial institutions and an integrated solution for achieving sales process automation in banking.
The paper focuses on presenting the requirements for security and confidentiality of stored data and also on presenting the identified techniques and procedures to implement these requirements.
Keywords: *interoperability, security, sales process automation, databases*

1 Introduction

One of the current trends is the increasing cooperation between companies throughout the product life cycle. This trend is related to specific business needs such as the need to reduce costs, flexibility and product innovation.
As enterprises have to face increasing pressure to be open to new collaborations, integration, interoperability and external business processes with business partners, conducted under a common business network become essential. From the point of view of business applications connecting to business partners, access to data stored on mainframes, integration applications installed on different operating systems are just a few examples of situations where interoperability becomes an absolute requirement.
In previous research we have conducted a case study to create a development framework for software interoperability in the financial/ banking institutions whose business is focused mainly on the retail area. This platform was created following the stages of a methodology we have proposed and presented in [1]. In each stage we followed the specific steps of data processes and software products integration, in order to minimize the influence of risk factors and to maximize interoperability and system performance.
The main steps followed in developing the framework are:

- identifying strategic objectives,
- investigation business process analysis and business process modeling,
- identification technologies and software platforms to be used,
- simulation and implementation models,
- execution and system monitoring,
- evaluation and optimization of the solution,

as shown in Figure 1.

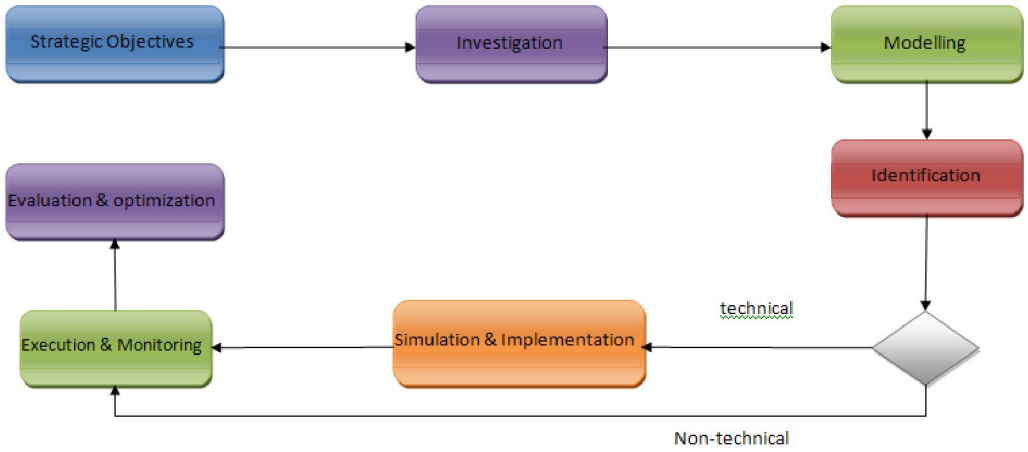

Fig 1. Proposed methodology for developing a BPM solution [1]

This article aims to summarize the developed framework, emphasizing requirements relating to the security database and the procedures and techniques used to achieve them.

2. Strategic objectives of the proposed framework

Retail banking has seen a significant expansion especially in the first part of the first decade of the XXI century this type of activity being relatively little promoted before. The main target was represented by corporate clients like banks and even those who gave greater importance to retail customers looked upon them rather as depositors, suppliers of financial resources, which were subsequently used for lending to corporate customers.

With the disappearance of the possibility for banks to place their money in government securities with attractive interest rates and the increased difficulty in attracting corporate clients, things began to change. More attention began to be given increasingly to the supply of credit to households. This shift in approach was helped significantly by the intense increase in household demand for loans. This demand has been in an upward trend until the end of the decade, as determined by the need and desire to purchase goods with long-term use or property without having to pay cash. This resulted in a significant diversification of supply from banks that have introduced products increasingly attractive to customers because of the increased competition. Clearly, increasing competition among banking players means more benefits for the population that can benefit from more favorable conditions, lower interest rates and fees, but also because banks are forced to make their activities more efficient in order to meet market conditions.

Receiving a loan depends on the outcome of an assessment called scoring. This assessment is made based on an algorithm that takes into account many parameters that have different values for each client. The parameters included in the scoring calculation are assessed as having different influences in determining the risk associated with a loan so that each receives its own weight. These weights vary over time depending on economic conditions in the medium.

The scoring algorithm is used to identify and size the risk associated with each customer, depending on its features [2]. Based on the obtained score, each client is assigned a risk category such as very low risk, low risk, medium risk, high risk and very high risk. The score is inversely proportional to the level of risk associated with a client. A higher score implies a reduced risk associated with allowing easier access to a loan or even higher amounts.

The criteria scale considered in calculating

credit score and its associated weights are not constant. The changes that occur over time can be determined by several factors such as: changes in the degree of risk a financial institution is willing to take it for each client segment, micro and macro-economic conditions change, adjustments or corrections credit scoring model.

Given the fact that sales of credit type banking products involves a great potential risk for companies that offer this type of product is essential to have a sound and prudent management of the credit granting process. This will reduce associated risks and thus minimize the risk of solvency, liquidity etc. In order for financial institutions to operate in optimal conditions certain conditions need to be fulfilled that a bank wants to achieve by introducing automation solutions for sales:

- making the credit granting process more efficient by streamlining steps to minimize the number of components
- improve communication and coordination between users;
- implementation of standards based on good customer service.

3. Implemented business processes

The proposed solution offered modules and capabilities for

- Customer management;
- Management of offered credit products;
- Sales areas management;
- A flexible and easy to use module for the definition and interpretation of financial indicators;
- A module for loan applications management that includes: recording financial information of customer data registration guarantees deposited automatically calculating and interpreting financial indicators, keeping a history of scoring and prescoring simulations made on loan applications, determine eligibility of loan applications;
- Lending Analysis module.

In addition to the expressed needs of the management of existing customers, there are some needs regarding lead management and sales opportunities and facilitate the qualification of potential customers (leads) to reduce the effort of agents to each client, resulting in increased number of potential new customers managed with the same effort and the same time. We want to achieve standardization, ease of communication, elimination of the redundant manual data processing and better control of the entire flow of generating prospects for management, so as to have consistent information in real time and can take corrective action to improve performance knowingly.

From the point of view of ensuring greater flexibility in the interaction with the system, there is the need to incorporate both mobile components and web-based interfaces.

Mobile components must operate both online (for synchronization of information) and offline, for situations where the user is out of coverage.

Mobile interface for agents (sales representatives) must provide functionality that would allow them to reduce the effort and the number of visits required for each potential customer. This includes the following modules:

- Data acquisition from the territory - allows the introduction of minimal information on an agent's new clients. This information can then be enhanced through the web-based interface.
- Prescoring - allows a quick calculation of the potential customer's qualification.
- Validating uncertain qualifications (online) - if we cannot obtain a decisive qualification for a potential customer based on the prescoring number, if there is connectivity to the communications network, the agent may request a loan officer validation or invalidation of that case,

automatically sending all information collected and may be notified if that potential client qualifies or not for credit.
- Products - interactive product catalog periodically synchronized with the central system. The agents can find information relevant to all products the organization offers.

Web-based interfaces will ensure access when users are not traveling. The solution will run on a server within the company and can be accessed online using a username and a password predefined by administrators from any computer in the organization. Each user will be defined as having a role in the system. Administrators can define access to the system for each role by enabling or limiting some functionality.

Agencies will have an interface within the online platform that can be used independently or as a complement coming from processed data on mobile devices. It is necessary to implement the functionality in the following areas:
- Potential customers: interface to view and edit information about potential customers.
- Visits processing interface: here the agent will see all of the latest visits after processing and generate requests to their loan officers on completing all the necessary information for a standard credit application.

Loan officers are designed to filter requests coming from the territory and to send them to central approval. Also they have a role in coordination of agents. Functionalities should be implemented in the following areas:
- View agents' requests: to allow viewing of all loan applications and related information coming from agents.
- Processing agents' claims: approval of this application by loan officers to lead to an automatic export of data from the new system in the current credit management system used by the organization.
- Processing customer visits: to allow supplement the information gathered through the mobile interface of the territory on current customers to complete visits according to procedures.

Credit brokerage companies that are partners with the organization are designed to find and process loan applications in the field. They are both responsible for finding new potential customers and with filtering them so as to provide bank applications with greater chance of success. Implemented functionalities must include both those for agents and for loan officers in the input and processing demands of the market. These requests will arrive, as in the case of loan officers, in the central lending system.

Web-based interface for the management team should ensure functionality in the following areas:
- Tracking Agents: similar credit facilities officers, those in central can analyze in detail the macro or national or regional activity of agents
- Tracking the activity of credit officers: to monitor the performance of processing claims submitted by agents or tracking visits activity to existing clients of the loan officer.
- Other specific management reporting process.

The online platform must have a centralized notification module that can be used by all users of this platform (agents, loan officers, the management team). The system's aim is sending official notifications on processes and procedures of the organization with respect to the activity of generating leads.

In order to develop these components we'll apply solutions based on data integration, portal, business processes and services.

4. Functional requirements

Customer Management – the solution must allow the registration of three types of clients: individuals, businesses and

freelancers. This level will retain only general data on customers and for all customers we must be able to register an unlimited number of addresses and contacts, as well as the preferred means of contact.

Customers are assigned a sales area, and each client has a sales agent responsible for its management. This may be the responsible agent who handled the registration of client (owner) or another agent who was sent to the client. For each client the history of interactions with him must be visible.

Between different categories of customers we should be able to define relations like:

- Individual - individual: family / life partner / acquaintance;
- Individual - legal: associate / administrator / employer / employee;
- Legal - legal: branch / agency.

Also, for each customer record we must be able to attach files in various formats. These files will be stored in the organization's portal, portal developed Sharepoint2010.

Management of offered lending products - The system must provide the possibility of introducing, by responsible, new product packages that the bank puts on the market and validation before they can be used by agents. Once placed, the products are subject to approval by the manager of the "credit products." department. If the product is approved for sale he enters the flow and can be used in the loan applications. Otherwise, it can be rejected leading to removal from the credit package offer. In this case the desired storage method if archiving the product for future reference.

For additional information about loan products the sales documentation are is used. This area allows attachment of useful files for agents and their management depending on the term of validity associated. The small number of additional documents linked with each loan product determines that integration with the portal to store these documents is not necessary.

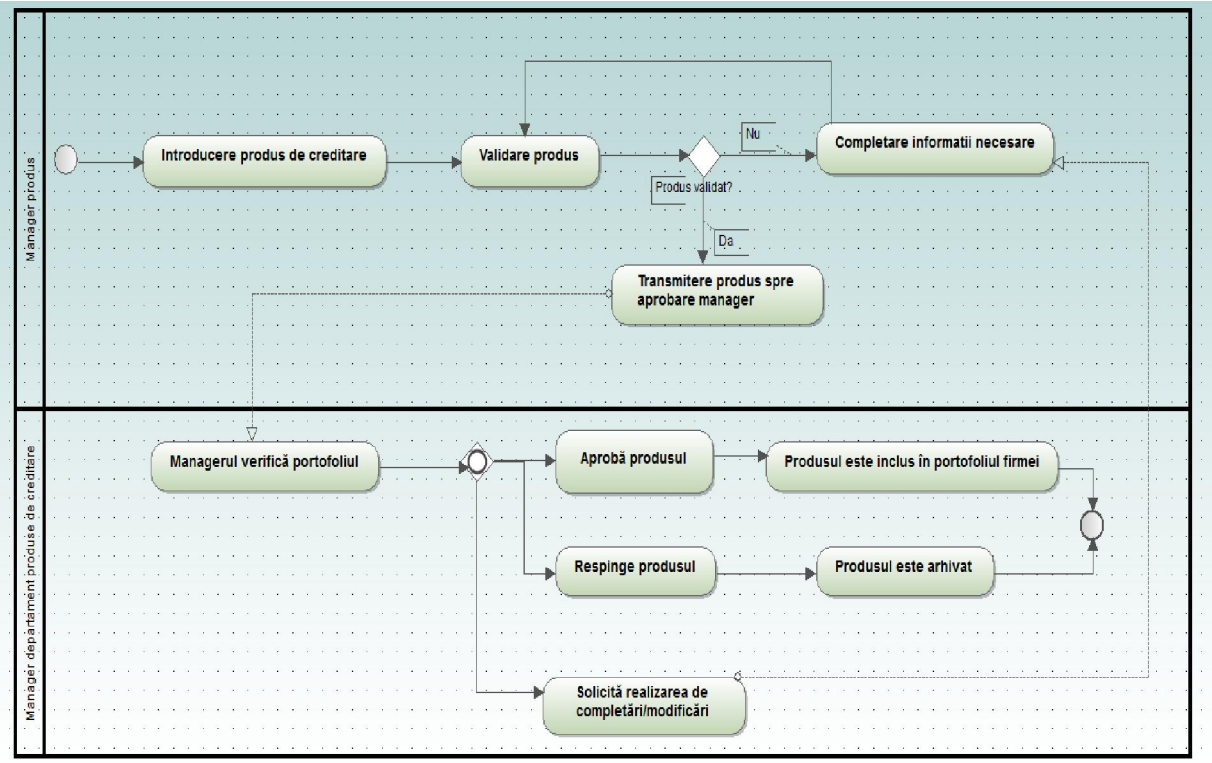

Fig.2 Lending products management process

Administration of loan applications – the solution allows the recording of the client's financial information, the set of information that must be stored being determined by the type of client. Records must allow the addition of parameters so that it can contain all the details of a legal organization, respectively all relevant data in financial terms for individuals. This information will be stored at the credit application level and not at the customer registration level as they are variable in time. Also, following the registration of financial information, we can register information regarding the offered securities. Collaterals are divided into categories/sub-categories, each collateral introduced requires an associated value and the type and amount of collateral may be used to determine the outcome of prescoring/scoring. All supporting documents relating to guarantees will be stored in the portal.

Simulations for prescoring and scoring can be made and the eligibility of a credit application will be determined. Eligibility is determined by the result of the last executed simulation. The solution will implement the company's prescoring and scoring algorithms, algorithms that can be modified according to market requirements. Eligible loan applications are submitted after pre-scoring to the credit bureau to determine their eligibility after scoring performance.

After scoring algorithm calculation, calculating and interpreting financial ratios associates should be done automatically. It is possible that a credit application considered eligible when calculating the prescoring will be rejected after the analysis and scoring calculation. A history of all prescoring/scoring simulations conducted on loan applications will be saved.

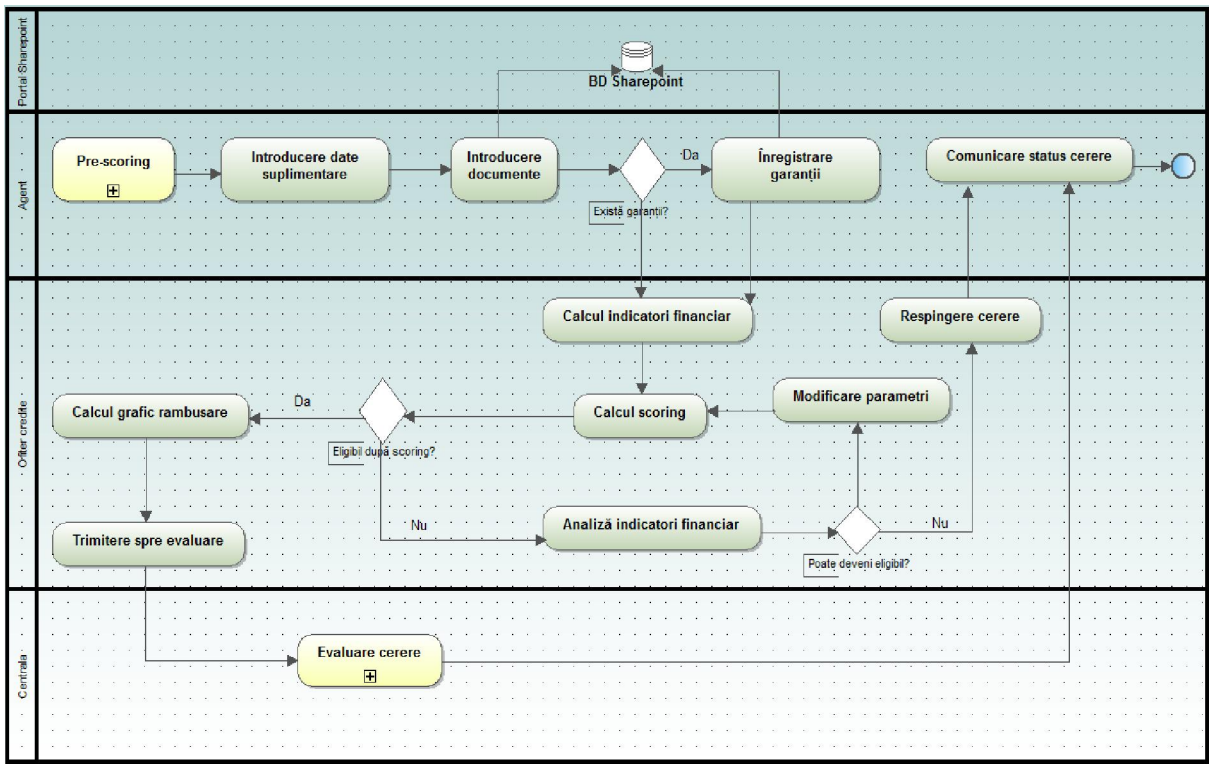

Fig. 3 Scoring process

Managing sales areas - sales areas will be the main spatial organization of business units. Sales areas overlap the main regions of the country, namely Moldova, Muntenia, Dobrogea, Oltenia, Banat, Crişana, Transylvania and Maramures. For each area there is a manager (sales manager) who supervises the work of agents in the area. Each sales area must be divided into sales territories and for each

territory there is a responsible sales agent. *Lending Analysis module* will lead to synthetic reporting situations that will help improve operations and performance. Statements may be obtained on the performance of agents such as number of meetings with customers or the number of requests made processed. It will also be able to evaluate the process of selling on a territorial basis by creating reports at territory or area-wide sale level.

5. Requirement regarding data privacy and security.

These requirements mainly deal with: how to access the system, security, data requirement (confidentiality, effectiveness), performance, flexibility and networking.

After investigating the particularities of the lending activity the requirements were divided into two categories: efficiency and protection of stored data and privacy and security.

For the first category the main requirement include the need to maximize efficiency by eliminating redundant work and delays, by facilitating communication, ensuring data integrity and by automating repetitive tasks and manual using appropriate technology. The system must incorporate on-line editing of data and contextual default values for certain fields to reduce the risk of errors and it must ensure the validity of data legally introduced based on legal provisions on the issue of communications. Also the system must be built to natively support the use of appropriate technologies to avoid the risk of data loss. The system must ensure backup and archiving processes, programs at various dates and times, repetitive or not.

The backup process will be done on the network drive, and the system must employ appropriate technology for recovering data from backup files.

In regards to the second category of requirements we have identified that the system must prevent unauthorized access to confidential data, the identification of persons in all its functions: data entry and data changes, data transmission and storage, access to records, the creation of children. The application must ensure that information required reconfiguration flexibility, access control and data consolidation.

The application must provide appropriate access to users, ensuring the protection of personal data with authentication and authorization mechanisms and centralized user management.

The system must meet the minimum safety requirements for the processing of personal data provided in the Order of the Ombudsman no. 52/2002.

The application must allow for a certain group of users (in Back Office) access is exclusive to a particular type of request from the integrated solution, this portion can be set as inaccessible to other groups.

6. Solving requirements regarding security and privacy

To access the application a Web interface based on Windows credentials will be used. In order to gain access, the user will have to connect with a registered account in the dedicated domain. Checking credentials takes place when logging on to the workstation, so once logged in a user can access the application via the web interface, without the need to introduce other identification data. Once you open the application, each user can view data, but not all customers will be able to make changes. They will be allowed to process only customers they are responsible for. Interface accessing will be made using the Internet Explorer browser, which is the only one supported by Microsoft Dynamic CRM 4. If you want to use other browsers (Google Chrome, Mozilla Firefox) will use Internet Explorer add-ons.

System security policies are implemented using Active Directory (AD). Running AD servers are called domain controllers. Such a controller authenticates and authorizes all

users and workstations in a Windows domain network type. Through Active Directory all suitable users are identified as part of the application it is stated and implicit right to work with it. When a user logs on to a workstation that is included in a domain, AD verify password entered to determine whether the user has the role of administrator or regular user.

User passwords must meet a set of security policies, namely:

- Have a size of at least 8 characters and contain at least one letter pattern, a digit and a special character;
- Password should be changed after not more than 31 days but no sooner than three days after the last change;
- It will store a history of passwords used so the same password cannot be reused until after 12 changes.

Users need to be grouped into roles and data access will be restricted based on these roles. Also there must be a possibility to share data so that users can get access to objects that do not belong to them, in order to accomplish various tasks in cooperation. When an employee leaves the institution, the user rights of the respective will be deleted.

Back-up strategy is dictated by the standard CRM policies and back-up implies a full week and a daily incremental backup. The latter achieves storing transaction logs at the end of the day. There will be two back-up media, one locally, in the IT department and one outside the department, in a disaster recovery center.

The backup process will be done on the network drive, and the system must employ appropriate technology for recovering data from backup files.

The archive will be on optical storage media (DVDs) or network drive.

For DVDs, they will automatically receive an ID generated by the system. DVDs archives will contain cases (data sheet in .pdf format) and not native portions of the database. When searching for a case that was previously archived on DVD, the system must provide the ID of DVD on which it is located.

If the volume of data to be archived is higher than the capacity of a DVD, the system will burn, in turn, all DVDs needed. Each DVD burner will have its own ID.

Archiving process must be transparent for the supervisor.

The system must leave it to the supervisor if you delete or not the archived information from the database

However, the system will allow the restoration of records in the database using records made up to that point.

To ensure the availability and security of operations, there will be regular monitoring of all systems and servers that contain confidential and secure information and checks of the available capacity. Servers must be protected and tested prior to any modification.

All information regarding access and attempts to access the system will be stored in order to identify unauthorized access. There will also be recorded and retained for a predetermined period all security incidents, in order to achieve access control and audit in order to investigate possible errors.

Data requirements, deal with issues such as consistency, confidentiality and efficiency. The solution offered should enable managed information assurance. It will be taken to ensure the uniqueness of information entered into the system by using tools for identifying and eliminating duplicates.

Performance is measured by the number of users that can use the solution and the time required to extract the information. The solution should allow for the simultaneous introduction of data by at least 1200 users.

The flexibility of the solution translates into the possibility of further development and optimization at both the hardware and software level. The system should permit better technological solution chosen to be built after deployment in a cost effective

manner. In order to upgrade the software level, the solution uses Microsoft Dynamics CRM 4 in native mode for transferring customizations to the latest version recently appeared on the market, Microsoft Dynamics CRM 2011. Document management portal is developed using the latest available version of Microsoft SharePoint. In terms of hardware, the system can be scaled vertically by adding processors and RAM on existing servers.

Interconnection with other systems must assume that the solution will allow data transfer and communication between interface developed in CRM and the SharePoint portal of the organization. In addition the solution must enable data export in Excel format. The solution integrates natively with Microsoft Exchange and Active Directory, and Microsoft Dynamics CRM integrates natively with Outlook and other Microsoft Office applications suite.

Conclusions

The solution is mainly dedicated lending management, achieving this integration of two elements commonly used in current organizations: a solution for customer management and business process that involves (Microsoft Dynamics CRM 4.0) and an online portal for document management. Security of information from the project refers mainly, but is not restricted, to six main components: confidentiality, authentication, authorization, integrity and non-repudiation logging.

While developing the application and the security measures, not only the possible external threats were taken into consideration, but also the potential vulnerability from the inside, one of the most neglected aspects when dealing with security issues.

References

[1] A. Florea, "Business Process Management Solutions Performance Tuning and Configuration", *Database Systems Journal*, Vol II, Issue 3/2011, September 2011, ISSN 2069-3230, pp 23-32

[2] X.Xu, C. Zhou, Z. Wang, "Credit scoring algorithm based on link analysis ranking with support vector machine", *Expert Systems with Applications*, Vol 36, Issue 2, March 2009, ISSN 1468-0394, pp 225-2632

[3] A. Lane, *A Database Security Management White Paper: Securing the Information Business Relies On*, 2012, http://www.technologyexecutivesclub.com/Articles/security/databasesecurity.php

[4] A. Bara, A.Florea, I.Botha, S. Oprea, "Solutions for CRM systems integration in organizational systems", *The 1st WSEAS International Conference on CLOUD COMPUTING (CLC '12)*, Vienna, Austria, November 10-12, 2012

Alexandra Maria Ioana FLOREA (born 1984) has graduated from the Faculty of Cybernetics, Statistics and Economic Informatics of the Academy of Economic Studies in 2007 and also from the Faculty of Marketing in 2008. Since then she is a PhD candidate, studying to obtain her PhD in the field of economic informatics.

At present she is pre-assistant lecturer at the Academy of Economic Science from Bucharest, Economic Informatics Department and her fields of interest include integrated information systems, information system analysis and design methodologies and database management systems.

Decision Support System in National Power Companies. A Practical Example (Part I)

Adela BÂRA, Iuliana BOTHA, Ion LUNGU, Simona-Vasilica OPREA
Academy of Economic Studies, Bucharest
Transelectrica Company, Bucharest
bara.adela@ie.ase.ro, iuliana.botha@ie.ase.ro, ion.lungu@ie.ase.ro,
simona.oprea@transelectrica.ro

The paper presents the developing stages of the decision support prototype in which the data warehouse and the presentation level are built and validated. The paper also extends the results published in the 12th international conference on Informatics in Economy (IE 2013) proceedings and will presents the major steps for developing the data warehouse that integrates the sources from the Wind Power Plants (WPP) from the national parks and also the interface modules that allow managers to analyze data at a central level.
Keywords: Data Warehouse, Business Intelligence, Analytics, Decision Support System

1 Introduction

The matter of developing decision support systems has now become a trend in the national companies. The large volumes of heterogeneous data gathered from different types of measuring devices and different formatting sources represent a challenge for developing integration solutions for extracting, transforming and loading data into a central data warehouse from where data can be analyzed and presented through special analytic tools like Business Intelligence (BI). This scenario is the base for our case study and also for the research project in the National Grid Company that has the objective to develop a prototype for decision support system for analyzing the renewable resources like wind power. Also, these aspects have been presented in a limited manner in the paper [1].

The system's architecture consists in four levels: data, model, presentation and telecommunications. The architecture has been presented in previous papers like [2], [3]. The aim of the present paper is to present the developing stage of the prototype in which the data warehouse and the presentation level are built. We'll describe some of the major steps for developing the data warehouse that integrates the sources from the Wind Power Plants (WPP) from the national parks and also some interface modules that allow managers to analyze data at a central level.

2. Considerations regarding data warehouses

The solution adopted for the informatics system prototype that we are going to develop is based on data warehouse implementation using a multidimensional model (MOLAP). The data warehouse built at this stage will allow metrics aggregation and also cubes processing and storage.

Usually the data warehouse architecture varies depending on the specific situation of each organization. For a basic architecture (figure 1), data is loaded from one or more sources, and the users directly access the data warehouse [5].

The architecture of a data warehouse has three main components, as shown in the figure below:

- Data warehouse itself and its management system;
- Data acquisition system from OLTP systems and other sources;
- Data analysis and presentation system.

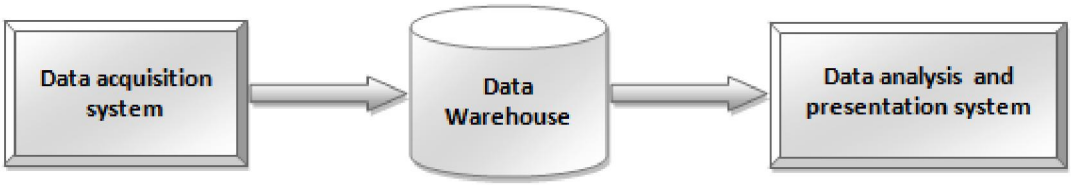

Fig. 1 – Main components of a data warehouse
Source: Authors

A more complex architecture uses integration and cleaning systems and also multiple data mart systems designed for each organization department, as indicated in [5].

As shown in figure 2, such complex architecture is divided into four different levels to achieve the data, as follows: sources, transformations, collections of data, analysis.

The data sources level collects heterogeneous data from various internal or external operational systems of the organization. It uses a process of data integration using a separate module called data source module.

The data transformations level uses a process of extracting, processing (cleaning) and data loading (ETL - Extract, Transform, Load), which requires data processing in terms of: integrity, precision, accuracy, format.

The data collections level contains processed data, aggregated and loaded into multidimensional structures on different levels prepared for being used in analysis. At this level we can design multiple data mart systems for enterprise departments.

The data analysis level involves extracting data from data warehouse and using business intelligence technologies for the analysis and interpretation of the information provided. At this level are used: OLAP products for analysis (so that data are presented graphically, in spreadsheets, embedded in portals etc.), features for data retrieval using more criteria, data mining products, complex and dynamic reports.

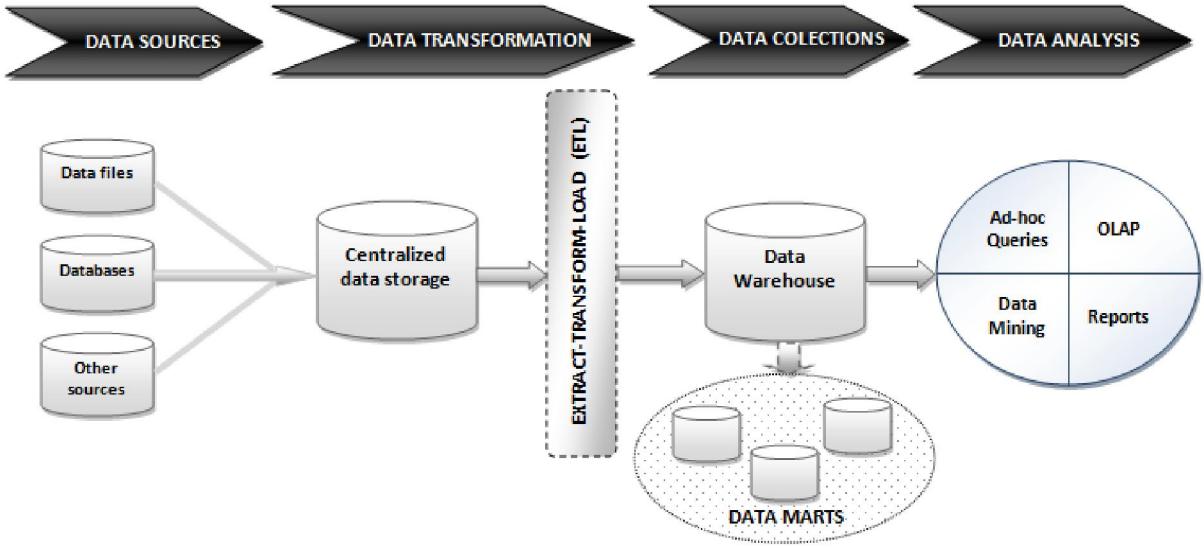

Fig. 2 – Complete architecture of a data warehouse system
Source: Authors

From the above architecture we can obtain three functional modules of data warehouse development: operational, central, and strategic [5].

The operational module is represented by organizational data which are usually stored in different formats at different locations. These data may come from

applications or operational systems from companies or from external sources. Regardless of their source, the data must be collected and brought into a consistent format to be used. This process of data transformation is the basis for building a consistent data warehouse. Data transformation is a complex process of extracting, filtering, cleaning, compounding, validating and loading data (ETL).

The central data warehouse module is represented by software products, servers on which they run, chosen implementation for the data warehouse. This module can be implemented in two versions: 1) implementation of a distributed system, where data is stored in decentralized, independent units (Independent Data Marts), each containing data relevant to a particular aspect of organization operations; 2) implementation of a unique, centralized data source, which can be accessed by users from all departments of the institution.

The strategic module (*business module*) offers the final value of a data warehouse which is due to the advantages offered by the various processes of decision making and analysis. By using different ways to access data and use processing technologies available, users can get information that will help in the process of establishing company strategy. In this way, data are prepared for interpretation and analysis using specific instruments (graphics, presentations, dynamic reports, web browsers, data visualization etc.).

3. Developing the data warehouse

As we shown in [1], to implement the model we have chosen Oracle Warehouse Builder 11g development environment (documentation available at [4]) and we have followed specific steps: creating the source module, designing destination module, making mappings, creating and validating objects.

Developing the data source module means actually importing metadata definitions from external sources of data (figure 3). All data sources need to be mapped to the data warehouses objects which are created when designing the destination module.

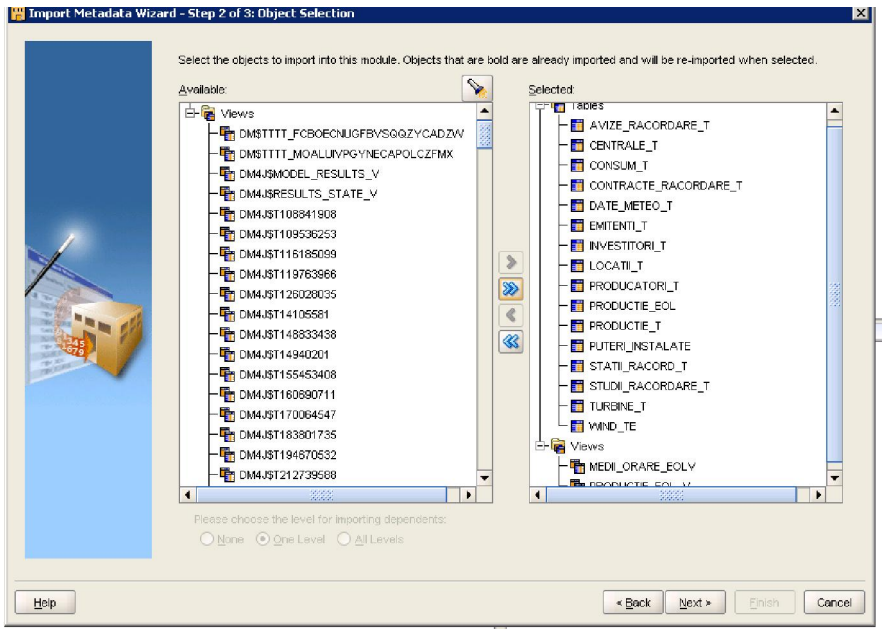

Fig. 3 – Creating the data source module
Source: [1]

Depending on analysis requirements of the decision makers, we have identified the dimension tables, which are the main structural entities that will be part of the organization and selection of multidimensional data. The dimensions are

organized on several levels, so as to permit typical OLAP operations: rotation, sections, navigation in detail (drill down) and aggregation (drill up / roll up).
For example, DIM_CENTRALA dimension contains attributes related to wind turbines and wind power stations and is based on a virtual table created on relational tables CENTRALE_T, STATII_RACORD_T and TURBINE_T:

```
CREATE OR REPLACE VIEW
dim_centrale AS
SELECT s.id_statie_racord,
s.denumire denumire_statie,
s.id_locatie, c.id_centrala,
c.denumire denumire_centrala,
c.id_investitor,
c.putere_instalata PI_Centrala,
c.nr_contract, t.serie_turbina,
t.puterea_instalata PI_turbina,
t.data_punere_fct,
t.id_producator, t.imagine
FROM turbine_t t, centrale_t c,
statii_racord_t s
WHERE t.id_centrala=c.id_centrala
AND
c.id_statie_racord=c.id_statie_rac
ord;
```

In this case, the hierarchical levels in DIM_CENTRALA dimension are: L_STATIE - L_CENTRALA - L_TURBINA.

Central tables containing measures attributes (metrics) and foreign keys to dimension tables are called data cubes. They usually contain numeric data that can be aggregated and analyzed based on existing levels in dimensions [5].
Particularly important for subsequent analysis and data presentation is CUB_PRODUCTIE cube that contains metrics related to energy produced by wind turbines. This cube is based on a virtual table created on tables PRODUCTIE_T and TURBINE_T:

```
CREATE OR REPLACE VIEW
fact_productie AS
SELECT t.serie_turbina,
t.data_punere_fct, t.data_revizie,
p.data, p.viteza, p.directie,
p.energie_produsa,
p.stare_turbina, t.id_loc,
t.id_producator
FROM productie_t p, turbine_t t
WHERE
p.serie_turbina=t.serie_turbina;
```

The design of data warehouse schema is made using joins between dimension and fact tables within cubes. In our case, a snowflake data warehouse schema model was built (figure 4), because there are joins also between dimensions [1].

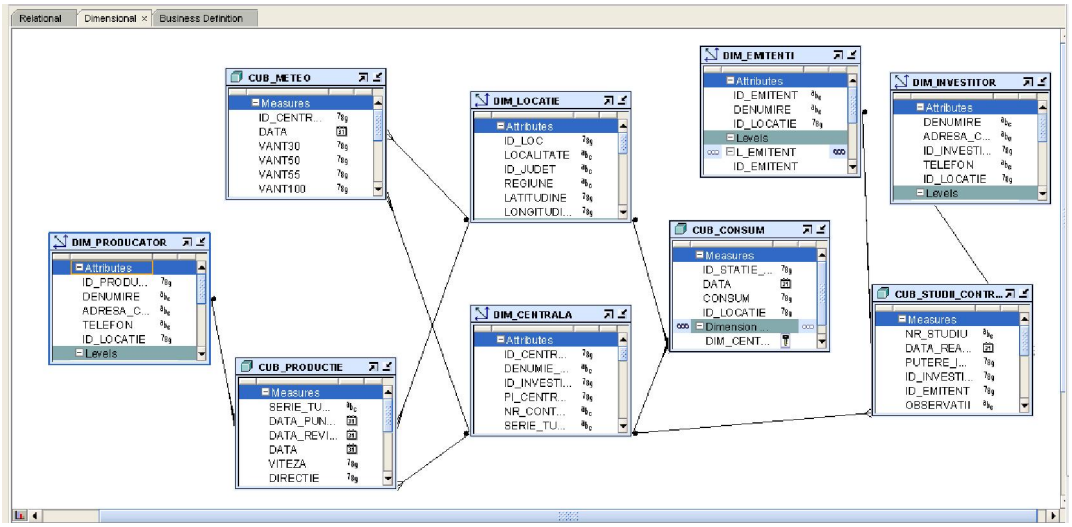

Fig. 4 – Snowflake schema of the data warehouse
Source: [1]

All objects of the data warehouse are stored in database, for each object being created a set of specific PL/SQL procedures and functions for creating attributes, hierarchical levels of the

dimensions, measures and joins within cubes.

4. Developing the presentation level

The presentation level consists in two major components: interface modules for local Wind Power Plants and interface modules for central management (executives and operative department from central level). Each module has two components: one for current activities like production, consumption, resource allocation, short term prediction, and another component for data analysis based on OLAP model for historical and predictive reporting.

The modules for current management are developed using JDeveloper environment and the modules for analysis are developed using Oracle Business Intelligence Suite.

Interface level for the national operator

In this step we developed the analyses reports and dashboards for the national OTS to analyze the renewable production. The application consists in two components: one for current activities and another one for decision support made by a set of analytical reports.

Data from the wind power plants will be presented in different formats for the executives, like: reports, dashboards, pivot tables, charts, and interactive maps.

In the following sections we'll present only some components of the BI modules that allow managers to analyze the evolution of wind generation, consumption and allocation in different periods of time in the selected zones and parks. The interfaces consists of a set of dashboards and reports that presents data in a graphical and intuitive way and available through mobile devices.

Another facility that we've implemented is a geographical representation for an intuitive analysis of renewable resources in different areas and wind parks.

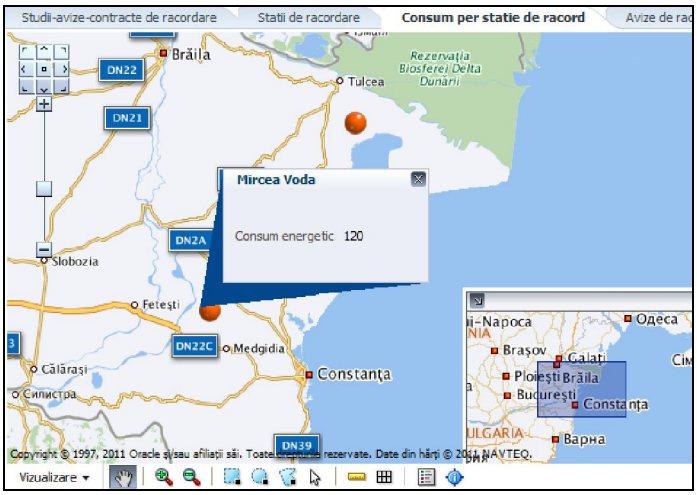

Fig. 5 - Interactive maps for analyzing WPP production
Source: [1]

The map in figure 4 visually depicts the energy consumption and production for each connected station. When selecting a geographic area and then the station marked the specific indicator is presented in detail.

Figure 5 shows an interactive map for the current studies and notices correlated with the actual contracts between the national OTS and the investors.

Fig. 6 – The studies and notices correlated with the actual contracts
Source: Authors

The results are presented through an OLAP report with navigation facilities between hierarchies (based on 1:*n* relationship): drill-down, roll-up, slice/dice.

The figure 7 presents the connection stations distributed over the regions having marked the production and the consumption.

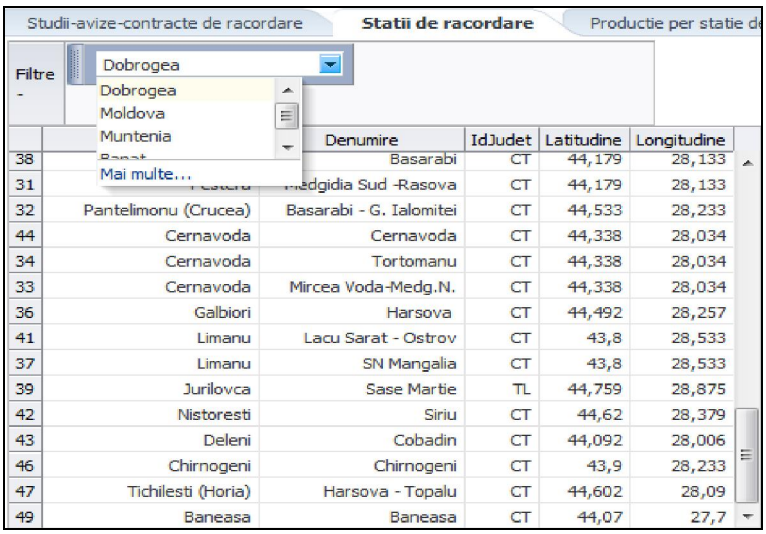

Fig. 7 – Analytical report with the connection stations over the geographic regions
Source: Authors

Analytical reports at central level

Based on the data warehouse we've developed a set of analytical reports for executives from the national level in order to observe and analyze the wind power plants connected in the system.

For developing we used the Oracle Business Intelligence Discoverer environment that offers different facilities for multidimensional analyses and graphical elements based on the virtual data warehouse.

In this step, by developing the analytical reports for executives, we can validate the data warehouse functionalities and we can establish new requirements and facilities

that will be implementing in the next phase.

At this stage, by making useful reports for the executives, the data warehouse functionality is validate, and new requirements and features may be implemented in the next phase of the prototype. We have built several sets of reports for different types of activities pursued in the analysis:

- Reports for grid connection, studies in progress and contracts in progress for WPPs;
- Reports for wind generation and production over time and in different regions;
- Reports for analyzing the consumption in periods of time and different measuring points;
- Reports for analyzing the functioning of national power over different periods of time;
- Reports for wind predictions correlated with wind generation in different WPPs.

For example, in the figure 8 the predictions for wind generation for the current month are represented in the same chart with the actual production and one can observe that the differences recorded are small. The predictions are made by using data mining techniques described in [3].

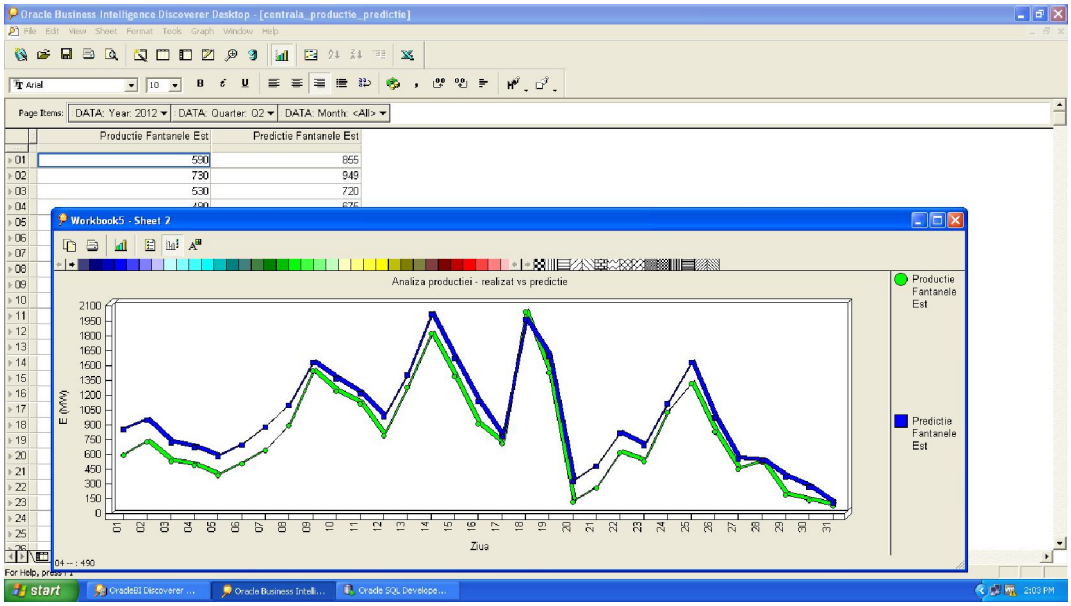

Fig. 8 - Wind generation predictions vs. actual production
Source: [1]

Reports for current production at national level over time

The production is observed and analyzed at national level in correlation and over regions and meteorological conditions. The analyses are realized for different types of turbines and producers due to noticeable differences.

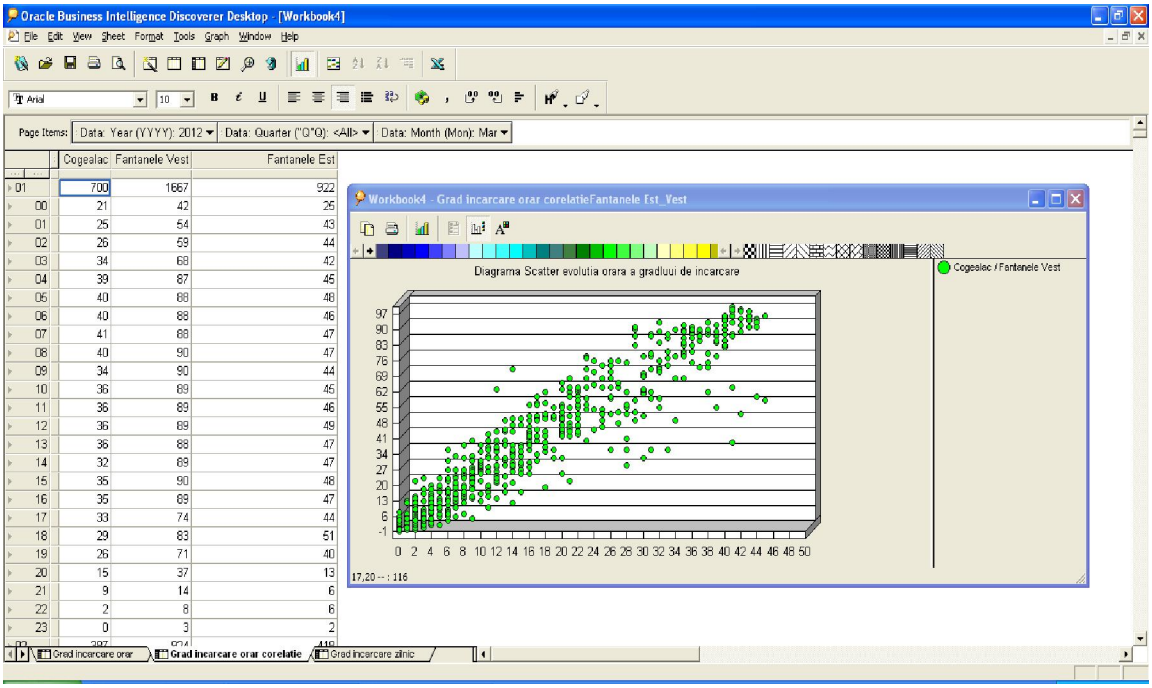

Fig. 9 – The correlation between different WPPs
Source: Authors

The report presented in figure 9 allows users to analyze the correlation between hourly productions from WPPs at national level by selecting the quarter, month and day as initial parameters. The reports are flexible, dynamic and allow the user to select the WPP for which the Scatter diagram is presented.

Conclusions and future work

The current research was conducted within a project won through national competition. In order to achieve rapid validation of data warehouse functionalities, the 2011 phase of the research project has used virtual objects and a ROLAP model (Relational OLAP). Thus, we designed the star / snowflake schema of data warehouse using Oracle BI Discoverer.

But the solution adopted for the prototype implementation is based on a real data warehouse which use multidimensional model (MOLAP), implemented using Oracle Warehouse Builder 11g. In this way the system performance is enhanced, and are allowed in an advanced manner aggregation of measures, storage and processing cubes, proper data validation and analysis.

The analytical module validates the data warehouse functionalities and offers an intuitive and interactive interface for executive and high-level managers to analyze the major components of WPPs: production, consumption and resource allocation.

Acknowledgment

This paper presents some results of the research project PN II, TE Program, Code 332: “Informatics Solutions for decision making support in the uncertain and unpredictable environments in order to integrate them within a Grid network”, financed within the framework of People research program.

References

[1] Adela Bâra, Iuliana Botha, Ion Lungu, Simona-Vasilica Oprea, “Prototype for a Decision Support System in National Power Companies”, *Proceedings of the 12th International Conference on Informatics in Economy (IE 2013)*, 2013, pp. 381-385, ISSN: 2284-7472

[2] Iuliana Botha, Adela Bâra, Simona-Vasilica Oprea, Ticiano Costa-Jordão, "Integrating XML Technology with Object-Relational Databases into Decision Support Systems", *Database Systems Journal* [Online], vol. III, no. 1/2012, pp. 11-20, ISSN: 2069–3230. Available: http://dbjournal.ro/archive/7/7_2.pdf

[3] Simona-Vasilica Oprea, Iuliana Botha, Adela Bâra, Anda Velicanu, "Prototype of a Decision Support System for analyzing and forecasting the Wind Energy Production in Romania", *Proceedings of the 1st International Conference on Information Technology and Computer Networks (ITCN '12), Vienna, Austria, Recent advances in computer engineering Series, Latest trends in Information Technology*, 2012, pp. 123-129, ISSN: 1790-5190, ISBN: 978-1-61804-134-0

[4] Oracle Warehouse Builder User's Guide, 11g Release 1, B31278-06, Available: http://docs.oracle.com/cd/B28359_01/owb.111/b31278.pdf

[5] Ion Lungu (coord.), Adela Bâra, Constanţa Bodea, Iuliana Botha, Vlad Diaconiţa, Alexandra Florea, Anda Velicanu, "Tratat de baze de date. Vol I. Baze de date. Organizare, proiectare şi implementare", Bucharest, ASE Publishing House, 2011, ISBN 978-606-505-472-1, ISBN vol. 978-606-505-481-3, pp. 402-408.

Adela BÂRA is a Lecturer at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Economic Cybernetics in 2002, holds a PhD diploma in Economics from 2007. She is the author of 7 books in the domain of economic informatics, over 40 published scientific papers and articles (among which over 20 articles are indexed in international databases, ISI proceedings, SCOPUS and 10 of them are ISI indexed). She participated as team member in 3 research projects and has gained as project manager one research contract, financed from national research programs. She is a member of INFOREC professional association. From May 2009, she is the director of the Oracle Excellence Centre in the university, responsible for the implementation of the Oracle Academy Initiative program. Domains of competence: Database systems, Data warehouses, OLAP and Business Intelligence, Executive Information Systems, Decision Support Systems, Data Mining.

Iuliana BOTHA is an Assistant Lecturer at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Cybernetics, Statistics and Economic Informatics in 2006 and the Databases for Business Support master program organized by the Academy of Economic Studies of Bucharest in 2008. Currently, she is a PhD student in the field of Economic Informatics at the Academy of Economic Studies. She is co-author of 5 books, 12 published articles (2 articles ISI indexed and the other 10 included in international databases), 18 scientific papers published in conferences proceedings (among which 6 paper ISI indexed). She participated as team member in 4 research projects that have been financed from national research programs. From 2007 she is the scientific secretary of the master program *Databases for Business Support* and she is also a member of INFOREC professional association. Her scientific fields of interest include: Databases, Database Management Systems, Design of Economic Information Systems, Grid Computing, e-Learning Technologies.

Ion LUNGU is a Professor at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. He has graduated the Faculty of Economic Cybernetics in 1974, holds a PhD diploma in Economics from 1983 and, starting with 1999 is a PhD coordinator in the field of Economic Informatics. He is the author of 22 books in the domain of economic informatics, 57 published articles (among which 2 articles ISI indexed) and 39 scientific papers published in conferences proceedings (among which 5 papers ISI indexed and 15 included in international databases). He participated (as director or as team member) in more than 20 research projects that have been financed from national research programs. He is a CNCSIS expert evaluator and member of the scientific board for the ISI indexed journal Economic Computation and Economic Cybernetics Studies and Research. He is also a member of INFOREC professional association and honorific member of Economic Independence academic association. In 2005 he founded the master program Databases for Business Support (classic and online), who's manager he is. His fields of interest include: Databases, Design of Economic Information Systems, Database Management Systems, Decision Support Systems, Executive Information Systems.

Simona Vasilica OPREA is a Senior Engineer and Project Manager at Transelectrica National Power Grid Company. She has graduated the Polytechnic University in 2001, holds a Master Diploma in Infrastructure Management Program, Yokohama National University, Japan in 2007 and a PhD diploma from 2009. She is the author of over 20 articles, from which 3 ISI Web of Science indexed and 2 included in SCOPUS international database. Domains of competence: wind farm, investment opportunity analysis, studies of prognosis, stationary and dynamic regimes, short circuit calculations.