Data Warehouse Supported Customer Relationship Management in Electric Energy Distribution

IMRE PETKOVIC[1], PERE TUMBAS[2]
[1]Politechnical Engineering College, Subotica peti@vts.su.ac.yu
[2]Faculty of Economics, Subotica ptumbas@yunord.net

ABSTRACT

In the environment where the demand of customers is prevailing and the competition is merciless and global the ideal of new trends in sales management represents the concept of strategic sales management. The change of the organization from that founded on the product to that turned to the customer is possible only by CRM (Customer Relationship Management). It is a public secret that CRM is the most wanted business term on the Internet and if the organization wants its web site to be visited, it is necessary to place this abbreviation somewhere. This paper deals with the characteristics of customer relationship management in electric energy distribution companies depending on which the software system supporting the management and the background data warehouse can be built. The paper pays special attention to the creation of meta-data (meta-data bases) of data warehouses in the given field of business.

Characteristics of CRM

Defining CRM is different in sales management, but often it is described as a process having four substantial steps (Peppers & Rogers):

- preliminary identification of needs and customers' wishes,
- differentiating customers to the needs and importance of realizing organization targets,
- interaction with customers from the view of the precise identification of their needs and
- shaping products, i.e. services that are offered to some groups of customers, on the basis of "learned" by the help of interaction.

CRM is not a technology, but it requires a technology. If we want to implement such a technology that will help the organization to turn to the customer, then the solution must supply the following three services:

- marketing is managed by knowledge,
- service of support to customers with the history of their activities (buying, interest, habit, and so on), and
- automated process of selling.

Therefore, the most trivial interpretation of the CRM notion means that it is a strategy focusing on better familiarizing with the needs and behavior of the customer in order to build a stronger connection with him and realize the relation that is the key of success of a contemporary business organization, and which is supported by contemporary computer technology. The differences between the old and new approach to the function of selling are seeing on the Table 10-1.

Table 10-1. The Old and the New Approaches

Old approach	New approach
Win new customers	Keep existing customers
Receive orders	Become preferred provider
Managing all customers in the identical way	Managing every single customer in his/her own way because of maximal long-range profitability
Sell to anyone	Concentrate on the customers who can provide high profit

There are technological components of CRM, but to think about it, primarily, as a technological term is a big mistake. It is more useful to understand CRM as a process for receiving much information about customers, sales, successfulness and responsibility of marketing, market trends and so on. The idea of CRM is to help business in creating the value and satisfying the customer using technology and human resources, and it can be attained by:

– providing better services to customers,
– developing more efficient call services,
– more effective selling of products,
– giving help to selling staff to conclude business quicker,
– simplifying the process of marketing and sales,
– winning new customers and
– increasing the customers' income.

However, simple buying software and its installation can do all this. For CRM to be really useful, the organization must first makes decisions what kinds of information about customers wants and what it wants to do with the information. Then, it must take care about the different ways of receiving information from customers, where and how these data are stored and used. One organization, for instance, can be in interaction with customers by e-mail campaigns, web sites, call centers, mobile selling staff, marketing, advertisements, and so on. A solid CRM connects all these forms of mutual interaction. There are some data that a CRM project should collect:

– reactions to campaigns,
– data on delivery and satisfying time limits,
– data on selling and buying,
– data about customers,
– registering web data,
– filing services and
– data on web selling and so on.

In the last years there was a sudden need to think in a new way about doing business that should induce the organizations to change radically the way of business and the relations with customers. The increase of market transparency and globalization caused the need of keeping off the customers as the most important question of contemporary organizations. The necessary conditions of such an orientation forced the need for CRM software, after all, for the part relating to sales management, e.g. the adequate customer relationship management. The automating operative processes (operative CRM), the analytical solutions (analytical CRM) and the electronic business with providers and customers (collaborative CRM), the collected information are necessary for understanding customers' behavior, adequate relationship management and fruitful strategic organization management. Analyzing the characteristics of CRM software (Siebel, People Soft, Oracle, SAP, J.D. Edwards), some mutual and the most important uses of electronic CRM solutions can be cited:

– more effective segmentation of customers and target group optimization,
– better anticipation of market development,
– profitability analysis of every individual customer,
– constant increase of customers' loyalty,
– quality increase of services for customers,
– better possibility of sales, using mobile equipments,
– intelligent product configuration, using the Internet technology
– shorter selling cycle and bigger profitability of sales process,
– synchronization of information from many different sources, and
– faster reaction to market changes.

Data Model for CRM Supporting

The philosophy of the data warehouse is very different than the philosophy of classical databases that are regularly used today for recording current (and updated) data. Instead of immediate current data, Data Warehouse contains a complete history of value changes of recorded characteristics on the basis of which different predictions can be formulated. Therefore, data warehouses can be efficiently applied in the preparation and strategic and tactic business decision-making. There are many methods for data warehouse design, and also many data are

available both at physical and abstraction levels, the most famous are relational and multidimensional models.

The need for data warehouses is formulated at the end of the '70s when the need for decision support systems appeared. Such systems existed at the end of the '60s when they were called Executive Information Systems. In the '80s, besides their continual development, their English name was also changed: MIS – Management Information Systems. Very soon, it was evident that the system based on the classical databases cannot satisfy the expectations in the field of decision support from objective structural limitations of databases. DSS (Decision Support Systems) applications could process efficiently only the current data. There was a possibility for processing filed data, but it was not very effective. Databases and classical information systems were planned for quick and efficient data memorizing (store housing), data processing and generating answers to the short SQL queries. DSS applications asked for not only the current state but also all the previous states for several months and years back. It could be done by the increase of memory capacities with an avoidable increase of time for processing. For analysis of the huge volume of data and discovering data trends it was necessary to provide big computers, but the results were halfway ones. Then the researches began to find out the appropriate structures for memorizing a huge volume of data, along with the most often used aggregate value of these data, all with a view of making easier the application of statistical data for data analyses. The classical OLTP processing (OLTP – On line Transaction Processing) could not provide such requirements. The type of processing that could successfully do the tasks expected by DSS systems was called OLAP processing (OLAP – On Line Analytical Processing). OLAP processing relied on the newly constructed system for data management called: Data Warehouse (DW).

Data Warehouse by Data Models

There are some differences in conceiving the notion of the data warehouse if we listen to the pioneers, the founders of these fields. R. Kimball talks about the data warehouse: "The conglomeration of an organization's data warehouse staging and presentation areas, where operational data are specifically structured for query and analysis performance and ease-of-use." [3] . W. H. Inmon's definition is cited even more often: "A data warehouse is a subject oriented, integrated, nonvolatile, and time variant collection of data in support of management's decisions." [2] . R. Kimball's definition is more general, and Inmon's one is more accepted and it determines the field of application of the data warehouse. The epithet subjectively oriented is related to the fact the data warehouse includes the date of only narrow thematic fields, not the complete business. The integral character of the data warehouse covers the fact that data are collected from different operational systems and they are memorized after their standardization. The nonvolatile characteristic states the unchangeability of data in the data warehouse. Data change in operative systems, namely, they don't cause data change in the data warehouse, but only the widening and integrating of new data with the old ones. The time characteristic reflects the fact that the data in the data warehouse are stored with the associated timestamp, e.g. the time mark of data origin. In later analyses, this will enable the analysis on the time basis. Data Warehouses are loaded from the operative systems with the data necessary for strategic decision-making. Their size can be measured in gigabytes (GB) and terabytes (TB). Fewer data warehouses of the local type supporting decisions of the smaller (limited) importance are called Data Marts. To some opinions, the data warehouse that helps the functioning of CRM applications also represents only data marts because it covers only one (the most important to many opinions) function of the enterprise. Data marts use only one part of the data warehouse and they can be of different types

depending on the purpose of using. The most important data marts types are [1] :

– *OLAP data mart* is realized by the software for database management of a multidimensional organization that uses a data cube, worked out on the basis of a star like scheme; they are used for in advanced defined and often used queries.
– *Exploration warehouse* is a typical data mart for researching and ad hoc processing, applying special processing in strong servers and by an optimized processing of ad hoc query gives answers for a short time. It is used for hypothesis development.
– *Data mining warehouse* is used for controlling and proving hypotheses constructed by the help of the exploration warehouse. Special techniques of data mining are used.

The task of building Data warehouse can be done on the basis of two basic approaches: top-down and bottom-up. The CIF (Corporate Information Factory) architecture is used for the realization of the top-down approach [1] , and the bottom-up approach results in Bus architecture, also called a multidimensional architecture [6] . In the CIF architecture the data warehouse takes the central place that is in fact the database with relations in 3NF including only elementary (non-aggregate) data. On the basis of it, different data marts, being of multidimensional architecture, can be generated and provided by data. The bus architecture is a set of coordinated data marts (super marts) that can be built according to the necessity. To this approach of building architecture, the central data warehouse doesn't exist because it would be an unnecessary redundancy – data repeating. The data of information systems e.g. classical data bases can be described at three application levels, i.e. they can build three kinds of data models [10] : (1) conceptual (level), (2) logical (level) and (3) physical (level) models. The same levels, the same data models can be defined for data warehouses, too [11] . For the Bus architecture, the multidimensional model (MDM) is used as a conceptual data model, and Object Definition Language (ODL) can be applied for specifying the CIF architecture in building the conceptual model or Entity Relationship Diagrams (ERD), for data warehouse and MDM for data marts (see Figure 10-1).

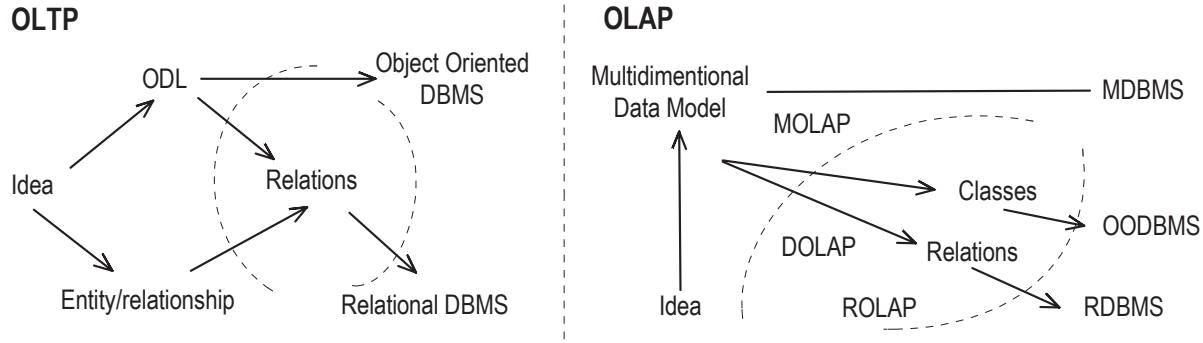

Figure 10-1. The OLTP and OLAP architecture levels

MDM is sometimes called a data-cube model and this name is applied both to three dimensional and n-dimensional data sets. In the element (cell) of the cube are found numerical data or the set of data (measures). The borders of the n-dimensional cube represent the dimensions on which numerical data can be grouped according to aggregation functions. The dimensions are sometimes organized in the hierarchy, for instance, time dimensions: a year, half a year, a quarter, a month, a week or a day. The attribute is an element of the dimensional hierarchy, for instance a month, and a tag is one concrete variable of the observed attribute, for instance April. The detailed description of the cube (granularity, to Kimball: grain) is determined by

the element (cell) of the cube, and it is the smallest "unit" of the numerical data, that can be got if all dimensions are used for addressing. If there is more than one numerical data in the cell, their granularity must be the same. The conceptual model has the operators that are used in generating different reports, and they realize the most often used functions: roll up, drill down or roll down, pivoting, selection, filtering, slicing and dicing. For the documentation of MDM models, graphical representation usually is the ERD model used. Namely, the MDM models differentiate only in the structure from the classical ERD models [13] , namely:

– The structure of the MDM model is simple: the entity of facts is in the center of the structure and the entities of dimensions are connected to it (star scheme) or to the hierarchy of dimensions (snowflake scheme).
– There is only one type of the link 1:N in the MDM model, where the end is "more" on the side of the entity of facts for every link, and from that side the link is obligatory, and the end to the dimension is of the type "one", from that side the link is optional: so the links in the MDM model are not marked.

The schemes constructed by ERD divide into star schemes and snowflake schemes, as the above cited. The difference is that for the star schemes it is not possible to define the hierarchy of dimensions ([9] , see Figure 10-2) while for the snowflake schemes the hierarchy of dimensions are regularly presented, if it is wanted and needed. The hierarchies of dimensions with the snowflake schemes enable to generate reports for every defined hierarchical level of any hierarchical dimension.

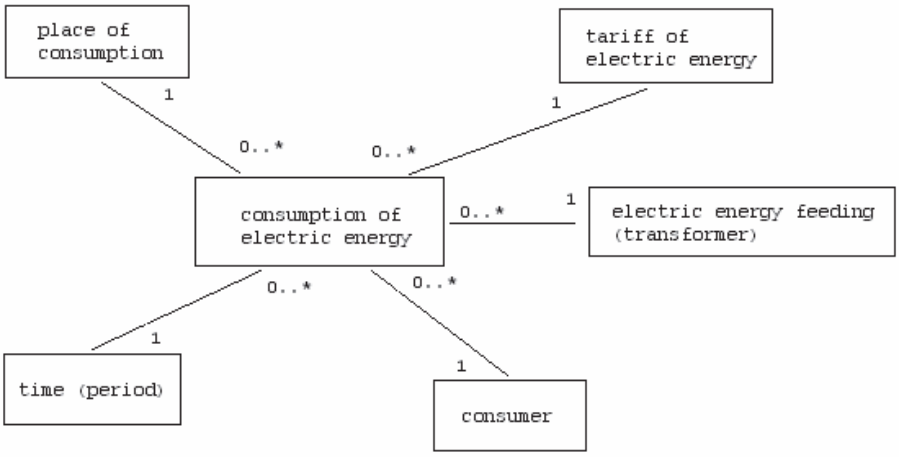

Figure 10-2. Star scheme model

The number of different reports is rather large, equal to the result of the following variables: the number of numerical data in the basic cell, the number of dimensions, and the total number of hierarchical level for all dimensions. In the data warehouses the process of data loading results different hierarchical level in every dimension that is not counted as a rule. In the case of a great number of hierarchical levels the process of loading can considerably slow down; not speaking about the necessary memory space and the time for answering to the queries. Ralph Kimball [5] suggests the following process for modeling data warehouses:

– choice of a business process from the business model,
– determining the level of granularity,
– choice of an attribute for dimensions, determining the hierarchy of dimensions,
– determining numerical data (measures) in the elementary cells.

The Starter Model for Supporting CRM

The starter model for supporting CRM enables a successful realization of customer business management, and it is independent of the enterprise and business fields ([1] Figure 10-3). The starter model represents a base for building concrete CRM solutions (applications). Contacts with consumers are established through different channels: from direct personal contacts by e-mails, but most often by post. The company and the customer can initiate the contact, that can be realized for many reasons: for information, coordinating data, offer, activities of the campaign, changing messages, and so on. For winning the sympathy of customers, contacts must be of personal character: customers are addressed by their names, not surnames. The campaigns are mostly of mass character, today: they are applied for the categories of consumers, but personification is preferred. All this is done, of course, for the purpose of increasing efficiency and successfulness of the campaign.

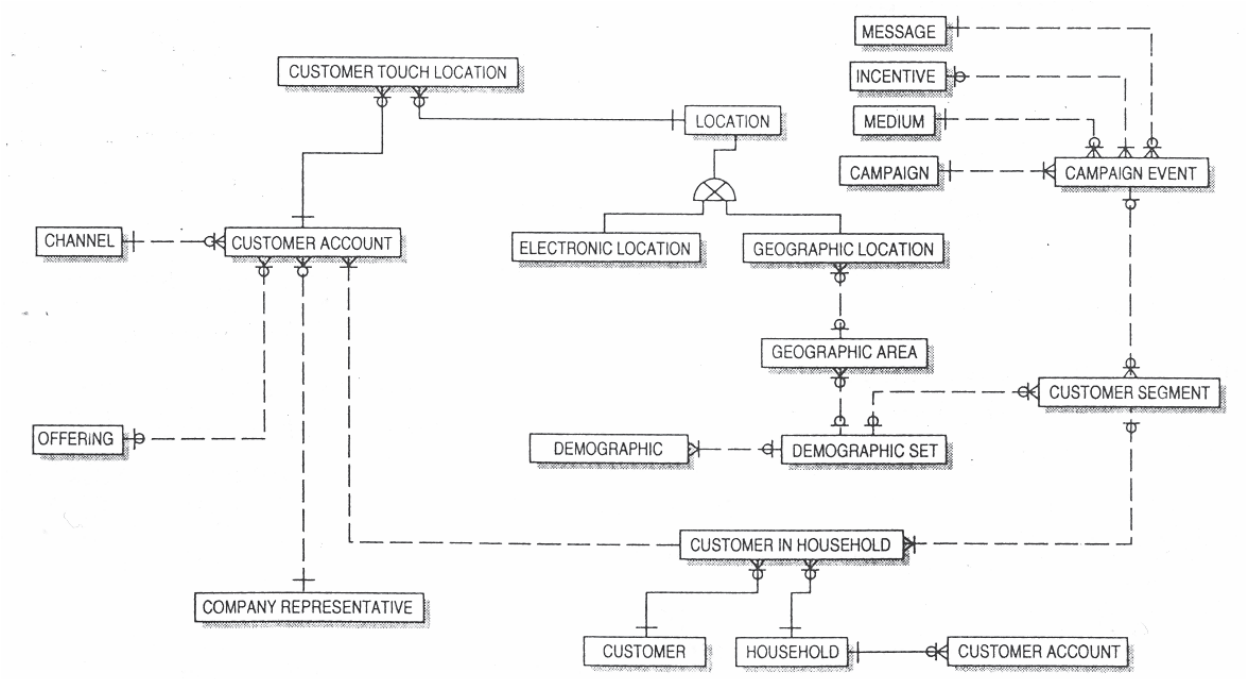

Figure 10-3. The starter model for supporting CRM

Metadata of the Data Warehouses

Simply to say, metadata are the data about data: they describe the components and characteristics of data using in an application, e.g. where and in what form (system of coding) they are stored, which connections are there between these data, where they originate from and what role they have in business operations.

In the '70s and '80s the term data dictionary was used, now we use the term of the metadata dictionary or repository[2]. In essence, today's metadata have really outgrown the notion of the former data dictionaries that described only the users' data. Contemporary meta databases include the following three kinds of data [12] .

- data describing business (the data of business logic or the data of logic of business)
- data describing data transformation (from the operation systems in DW), and
- administrative metadata (data on the possibility of navigation in data).

The data for describing business logic correspond mostly to the data in data dictionaries of the '70s and '80s. While the data dictionaries

[2] The comment of the Chief Editor.

describe the elements of data models: entities, attributes and connections between the entities, data for describing operations besides the previously cited, presented business rules, as well as the description of processing dictated by business operations. The competence for data capture and data maintenance belongs to the user of the applications.

Data for describing data transformation include the knowledge, i.e. the documentation about data and algorithms for data uniformity in the process of periodical DW uploading. Besides the description of DW structure, the source place where the data originate from and the destination of DW data, transformation data include the conditions and timing of transferring, as well as the processes of transformation. In the case of the CIF architecture, there is also a description of data mart uploading (conditions, timing and operations). If some software tools realize the ELT process, metadata are taken from the documentation.

The administrative data are more detailed than the description of navigation possibilities in database. The query paths indicating the navigation in database describe only the physical characteristics of the data seating themselves on external memory media, and the administrative metadata. This information keeps the time needed for uploading data into DW, the time for generating answers to queries and for generating reports with all the used patterns by the users. These data reflect the efficiency of uploading and functioning DW, e.g. applications, and on the basis of them we can uncover the data that are rarely used by queries.

The degree of metadata maintenance (updating) determines the quality and usefulness of CRM applications. The software for realizing meta databases implements the active or passive metadata management. The active metadata management can be done by software solutions that are the integral part of DW. With the active metadata management every type of data is registered in the meta-base before getting into DW, and from that moment it is automatic (thanks to the integral DW solution and metadata management). The advantage of this approach is an instantaneous promptness from the standpoint of data, their transformation and processing. The passive metadata management represents a software solution, independent of DW (CRM applications). It is projected and realized especially, and so it can be better in the field of business logic data than the finished integral solutions. However, it has irrefutable fault, and it is the maintenance. With these solutions, the user updates data (it is not automatic) and very often it happens that after the beginning enthusiasm in the maintenance of meta database updating, that activity is neglected.

Some Problems in CRM

Among other specificities of electric industry there is a fact that it is impossible to accumulate electric energy so it has to be distributed at the moment of receiving it. For functioning the electric power industry, it is necessary to provide the appropriate high-voltage, middle-voltage and low-tension grids. With a view of irreproachable functioning grids, it is needed to maintain them continuously, expand them according to the development of consumption and the increase of consumers, and there are unexpected damages that must be eliminated very soon. Timing and managing operations connected to the regular maintenance and damages elimination, coordination of maintenance groups and damages elimination (24-hour duty service) represent unavoidable widening of the standard CRM functions.

The energy distribution companies, realizing the function of supplying consumers with electrical energy, have the other specific characteristics in everyday work that widen the number of channels of CRM applications. Data gaining and remote control of the most vital elements of energy distribution grids (transformer stations of the voltage levels of 110/20 kV) and the remote control of electric meters and management of electric energy with the consumers has been an organized practice in the western countries for many years.

The remote control in Subotica was introduced 15 yeas ago in one transformer station in Backa Topola (34 km from Subotica) with the control center in Subotica. Nothing has happened in this field since then. The similar situation is with the remote control of electric meters and electric energy management with the consumers [7] . There is one own development, a hardware software solution [8] functioning since 1997 and including two distributive transformer stations of the voltage level of 20/0.4 kV and their consumers, about 200 households. There are not any prospects for further widening of remote reading now, although the functions realized in this system enable a fine analysis of the consumption function, as well as electric energy consumption management. There is a possibility for realizing some CFM functions, as the survey and memorizing the instant, 15 minutes, daily, weekly consumption, as well as the consumer balance of account and the category the consumer belongs to.

CRM is a complex philosophy including different scientific fields, starting from the psychology of behavior through shopping, marketing, sales till information and communication technologies. The successful realization of CRM functions is conditioned by an irreproachable functioning of all the above-cited (there are many of them, the most important of them were listed) factors of philosophy. Customer relationship management has the biggest, from the technical standpoint the most complex support by data warehouses that include the following groups of activities:

1. Initial and periodical data warehouse uploading, data extraction from the operational systems, data transformation to standardized form, and uploading of standardized data into data warehouse.
2. Generating the necessary reports.

On the basis of the previously presented facts we can conclude that the business of energy distribution companies is a pretty involved system and mostly covered by the appropriate software support – operative systems of the different nature. To many opinions, for the biggest problems in functioning data warehouses, only three letters: ETL can be accused. This abbreviation, according to Kimball [4] , can point to the existence of even 38 subsystems in functioning data warehouses.

References

[1] Imhoff, C. – Loftis, L. – Geiger J.: Building the Customer-Centric Enterprise: Data Warehouse Technics for Supporting Customer Relationship Management, Wiley Publishing, Inc, Indianapolis, 2003.

[2] Inmon, W.H.: Building the Data Warehouse - Third Edition, Wiley Publishing, Inc, Indianapolis, 2001.

[3] Kimball, R. – Ross, M.: The Data Warehouse Toolkit - Second Edition

[4] Kimball, R.: The 38 subsystems of ETL, White Paper, Intelligent Solutions, December, 2004.

[5] Kimball, R.: The Data Warehouse Lifecycle Toolkit, John Wiley & Sons, NewYork, 1998.

[6] Petković, I. – Doža, D, – Mesaroš L. – Ujhelji, R. – Krušedolac, V.: Elaborat uvodjenja sistema za daljinsko očitavanje brojila električne – ED Subotica, 1995.

[7] Petković, I. – Doža, D.: "Projekat realizacije sistema za daljinsko očitavanje električnih brojila i upravljanje tokovima električne energije", ED Subotica, ugovor br. 1996.

[8] Petković, I.: Applikacija DALJCIT – ED Subotica, 1996.

[9] Petković, I.: Adattárházak adatmodelljeinek tervezése Rational Rose fejlesztőeszközzel (Data Warehouse Data Model Planning with Rational Rose), Tudomány Napja Konferencia, Dunaújvárosi Főiskola, Dunaújváros, 2004.

[10] Raffai, M: Information System Development – Novadat, 1999.

[11] Sidló, Cs.: Összefoglaló az adattárházak témaköréből http://scs.web.elte.hu/Work/

[12] Todman C.: Designing a Data Warehouse – Hewlett-Packard Company, 2001.

[13] Todman, C.: Projektovanje skladišta podataka, HP-CET, Beograd, 2001.