

Data Warehouse – Opportunity for Local Authorities

Stanislava Šimonová

Institute of System Engineering and Informatics,
Faculty of Economics and Administration, University of Pardubice

Abstract

Local government exists in a very complex organizational environment, which has been subjected to an ever increasing pace of changes. This situation has generated a huge impact on the local government's reaction towards two major elements in its administration – decision making and technology approaching. The data warehouse technology offers a solution to this problem. But there are various obstacles that inhibit the government from adopting this technology.

Key words

local government, data warehouse, decision making processes, technology approach

1. Introduction

There is a difference between commercial organizations and public organizations in their considerations of approaching information and communication technology at all, at particular data warehouse technology. It depends on motivations and the obstacles of introducing or implanting such a technology. Unlike the public organizations, the main motive in commercial organizations is to make profits [3]. By collecting customers' data, analyze it and make better decisions that increase the business wealth. While in the public organization, profits and business wealth have no meaning at all. It is the only local government approaching this technology, the motive is to providing the citizens with better services, which means, better education, better workforce, better roads, healthy population, cleaner environment, safer society etc. It was realized that better services requires the management to accelerate their decision making process so they can respond quickly to the on going changes. One key to this accelerated decision making is having right information at the right time and easily accessible by the right people, which with data warehouse technology can provide.

2. Data warehouse technology

Operational systems usually have strict performance requirements, predictable workloads, small units of work, and high utilization. By contrast, decision support systems typically have varying performance requirements, unpredictable workloads, large units of work and erratic utilization. Decision support data usually needs to be collected from a variety of operational systems and kept in data store of its own on a separate platform. That separate data store is a data warehouse. A data warehouse is a collection of operational / transaction data specifically structured for querying, analysis, and reporting. The data warehouse is a direct reflection of the various function rules of the enterprise, not just of a specific function unit, as they apply to strategic decision-support information; it is the historical store of strategic information, with the history relating to either the data or its relationships or both [1] [2]. The data warehouse is also the collection point for the integrated, subject-oriented strategic information that is handled during the data acquisition process; it is the source of stable data regardless of how the processes may change.

It is the source of information that is subsequently delivered to the data marts. The data marts in question may be used for exploration, data mining, managed queries, or online analytical

processing [4]. The data mart is a smaller, more focused data warehouse. In many cases find it useful to create data marts for specific business units that have equally specific data analysis needs. Although the larger data warehouse could support those needs, the enormous bulk of data contained within a typical data warehouse could reduce the efficiency of a consistently focused data analysis effort. The data mart established for a specific online analytical effort is both the target of delivery and the direct source of data accessed by end users associated with that data mart. In contrast to the larger data warehouse, the data mart is a reflection of the business rules of a specific function unit, not the enterprise, as they apply to strategic decision-support information; the business rules reflected by the data mart need to be consistent with the enterprise rules, but are commonly tailored to the unique business capabilities addressed by the data mart. The data mart obtains its data from a relatively stable, cleansed, and integrated source – the organizational data warehouse; it does not need to cleanse or integrate the incoming data. The data mart is a set of tables designed for direct access by users who need to analyze data according to a set of predefined parameters; this characteristic requires a structure that supports easy, intuitive analysis across parameters and hierarchies within those parameters. It is a set of tables that is designed for aggregation; the fact table may contain detailed data, but most of the queries will view this data by the function constraints that form the dimensions. The data mart is typically not a data source for traditional statistical analyses needed for exploration or data mining. Omission of this requirement permits the data to be collected along predefined, known relationships.

3. Data Warehouse for Public Administration

An obvious premise of users' access to regional data warehouse should be the access via a network, most often through Internet (Fig. 1).

Regarding the demands for data within public administration, we can recognise the following requirements that an architecture of regional data warehouse should meet: first - regional data warehouse should work with data extracted from operational applications of public administrations, i.e. without any direct access to the operational databases; second - the data must be transformed and purified, i.e. potential redundancies must be removed, the data must be consistent, aggregated data will be created etc [3].

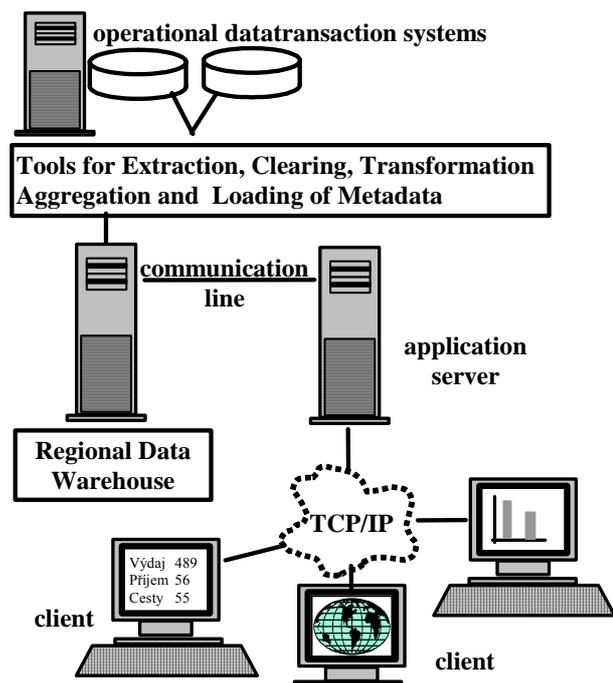


Fig.1: Regional data warehouse

The architecture of independent data marts works with various levels of both detailed data and aggregated data formed on the basis of extraction, purification, and transformation of operational data, which is positively reflected in their consistency (Fig. 2).

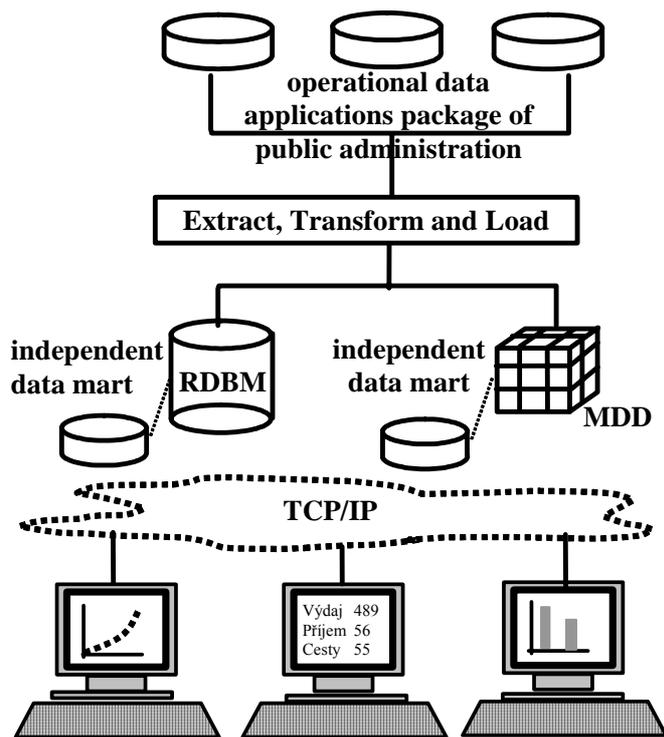


Fig. 2: Architecture of Regional Data Warehouse

Data marts react to actual information needs, quick responses being obtainable within a framework of a defined set of enquiries thanks to the dimensional approach to data modelling.

The complex and sophisticated variant of construction of regional data warehouse would consist in the architecture of enterprise data warehouse. However, this variant is exacting technologically, organizationally as well as economically, i.e. a variant requiring considerable experience in the field of data warehouse technology.

An acceptable architecture of regional data warehouse can be seen in a certain combination of

both ways of solution, i.e. the architecture of independent data marts without constructing the database of central metadata but with application of coordination across the data marts within public administration inclusive of unification of professional terminologies of the individual areas of interest.

3.1 Needs and Barriers

Local authorities are complicated and highly diverse organizations in terms of the administrative structure, governance services, participants, and size. Local authorities are expected to prognosticate and identify citizen's needs, react to complex social, demographics, economics, political and environmental problems, respond to the changes in citizens' expectations, respond to changes in business and national economy, provide effective and efficient set of services, and moreover, to subject to the central government constraints. Major changes in the public organization structure resulted to a certainly diversity. This diversity, instability and complexity of the local government environment, has a huge impact on the implementation and the usability of the technology on which e-Government relies on. This situation generates inconstancy in all the technological and manual systems that run the local authority, increases the disparity of the data stored in those systems and inhibit any effort of integrating those systems.

In the case of data warehouse technology, very little understanding obtained among the current and the potential users. The users in local authorities are divided to various categories, decision makers, officers as well as citizens. Now, no consider citizens as potential customers of regional data sources. It's found that the majority of decision makers have very modest knowledge on the data warehouse technology or what this technology can offer them. The officers, whose in charge of using this technology on a daily bases and producing reports that reflect the organization wide view to the decision maker, either have no idea at all or a very high expectation, which reflects non-acquaintance in both operational system the use and the

data warehouse. One of the major reasons for that are the employment policy applied in local authorities. And other reasons related to the financial, communication and social issues.

However, the regional “informational” situation is not mature enough for an enterprise data warehouse (demanding sophisticated construction and technology, and high financial investments) to be constructed; nevertheless, this is not inevitably necessary for the above-specified obtaining of high-quality outputs. The success of data warehouse is built not on specific technology or platform. The workforce are not equipped or qualified enough to be able to determine how to meet their business requirements, then implementing such technology will do to vain.

Conclusion

A data warehouse is a subject-oriented collection of data for support of management's decision needs. Implementing data warehouse technology to support decision-making policy in local authorities has many obstacles, both in the sophisticated technology (lack of model) and in the area of public administration (obstacles financial, structural, educational). But data warehouse technology is the right technology for support of decision making - where the right information at the right time and easily accessible is by the right people.

References

- [1] Humpries, M. Dara warehousing. Praha: Computer Press, 2002, 257 s.
- [2] Lacko, L. Datové sklady, anlyza OLAP as dolování dat. Brno: Computer Press, 2003, 286 s.
- [3] Šimonová, S. Regional Data Warehouse. Pardubice: Univerzita Pardubice, 2002, 132 s.
- [4] Marakas, G., M. Modern Data Warehousing, Mining, and Visualization. Prentice Hall, 2003, 274 s.

Kontaktní adresa:

Ing. Stanislava Šimonová, Ph.D.
Institute of System Engineering and Informatics, Faculty of Economics and Administration,
University of Pardubice, Studentská 84, 53210 Pardubice
tel.: 466 036 009
e-mail: Stanislava.Simonova@upce.cz

Recenzovala:

Ing. Jitka Komárková, Ph.D., FES, ÚSII, UPa