Data Warehousing Solutions for Reporting Problems

Juha Kontio
Turku Polytechnic, Finland

INTRODUCTION

Reporting is one of the basic processes in all organizations. It provides information for planning and decision making and, on the other hand, information for analyzing the correctness of the decisions made at the beginning of the process. Reporting is based on the data that the operational information systems contain. Reports can be produced directly from these operational databases, but an operational database is not organized in a way that naturally supports analysis. An alternative way is to organize the data in such a way that supports analysis easily. Typically, this method leads to the introduction of a data warehouse.

In the summer of 2002, a multiple case study research was launched in six Finnish organizations (see Table 1). The researchers studied the databases of these organizations and identified the trends in database exploitation. One of the main ideas was to study the diffusion of database innovations. In practice this meant that the researchers described the present database architecture and identified the future plans and present problems. The data for this research was mainly collected with semistructured interviews, and altogether, 54 interviews were arranged.

The research processed data of 44 different information systems. Most (40%) of the analyzed information systems were online transaction processing systems, such as order-entry systems. The second largest category (30%) comprised information systems relating to decision support and reporting. Only one pilot data warehouse was among these systems, but on the other hand, customized reporting systems were used, for example, in SOK, SSP, and OPTI. Reporting was commonly recognized as an area where interviewees were not satisfied and were hoping for improvements.

This article focuses on describing the reporting problems that the organizations are facing and explains how they can exploit a data warehouse to overcome these problems.

BACKGROUND

The term *data warehouse* was first introduced as a subject-oriented, integrated, nonvolatile, and time-variant collection of data in support of management's decisions (Inmon, 1992). A simpler definition says that a data warehouse is a store of enterprise data designed to facilitate management decision making (Kroenke, 2004). A data warehouse differs from traditional databases in many ways. Its structure is different than that of traditional databases, and different functionalities are required (Elmasri & Navathe, 2000). The aim is to integrate all corporate information into one repository, where the information is easily accessed, queried, ana-

Table 1. The case organizations

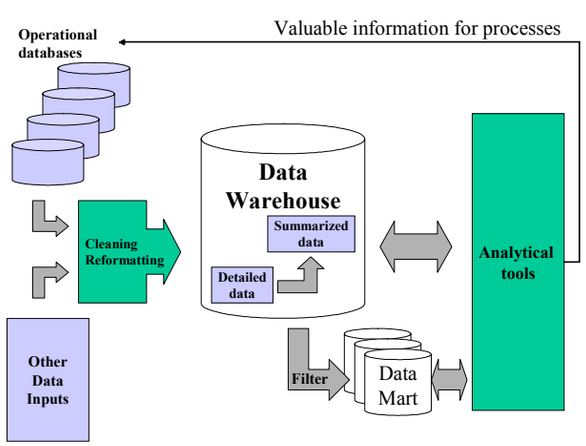
Organization/ Abbreviation Used in This Research	Line of Business	Private/ Public	Turnover 2002 *(Millions €)*	Employees 2002
SOK Corporation/SOK	Co-operative society (main businesses: food and groceries, hardware)	Private	2998	4645
Salon Seudun Puhelin, Ltd./SSP	Telecommunication	Private	28	121
Statistics Finland/STAT	National statistics	Public	52	1074
State Provincial Office of Western Finland/WEST	Regional administrative authority	Public		350
TS-Group, Ltd.,/TS	Printing services and communications	Private	69.7	2052 (Consolidated corporation)
Optiroc, Ltd./OPTI	Building materials	Private	149	388

lyzed, and used as a basis for the reports (Begg & Connolly, 2002). A data warehouse provides decision support to organizations with the help of analytical databases and Online Analytical Processing (OLAP) tools (Gorla, 2003). A data warehouse (see Figure 1) receives data from the operational databases on a regular basis, and new data is added to the existing data. The warehouse contains both detailed aggregated data and summarized data to speed up the queries. It is typically organized in smaller units called *data marts,* which support the specific analysis needs of a department or business unit (Bonifati, Cattaneo, Ceri, Fuggetta, & Paraboschi, 2001).

In the case organizations, the idea of the data warehouse has been discussed, but so far no data warehouses exist, although in one case, a data warehouse pilot is in use. The rationale for these discussions is that at the moment, the reporting and the analyzing possibilities are not serving the organizations very well. Actually, the interviewees identified many problems in reporting.

In the SOK Corporation, the interviewees complained that information is distributed in numerous information systems; thus, building a comprehensive view of the information is difficult. Another problem is in financial reporting. A financial report taken from different information systems gives different results, though they should be equal. A reason for this inequality is that the data is not harmonized and processed similarly. In the restaurant business of SOK Corporation, an essential piece of information is the sales figures of the products. It should be able to analyze which, where, and how many products have been bought. In the whole SOK Corporation, analyzing different customers and their behavior in detail is, at the moment, impossible. The interviewees also mentioned that a common database containing all products of the co-operative society might help in reporting, but defining a common classification of the products will be a demanding task. Centralization of the data is also one topic that has been discussed in SOK Corporation, which has been justified with improvements in reporting.

Figure 1. Data warehousing



In Salon Seudun Puhelin, Ltd., the interviewees mentioned that the major information system is somehow used inconsistently. Therefore, the data is not consistent and influences the reporting. This company has developed its own reporting application with Microsoft Access, but the program is not capable of managing files over 1 GB, which reduces the possibilities of using the system. According to the interviewees, this limit prevents, for example, the follow-up of daily sales. Another problem concerning the reporting system is that users are incapable of defining their own reports when their needs change. Analyzing customer data is also difficult, because collecting all customer data together is a very burdensome task. Therefore, Salon Seudun Puhelin, Ltd., has also discussed a data warehouse solution for three reasons: a) to get rid of the size limits, b) to provide a system for users where they can easily define new reports, and c) to gain more versatile analysis possibilities.

The State Provincial Office of Western Finland is a joint regional administrative authority of seven ministries. One of their yearly responsibilities is to evaluate the basic service in their region. In practice, this responsibility means that they gather and analyze a large amount of data. The first problem is that they have not used a special data management tool. The lack of an adequate tool for data management makes it difficult to do any time-series analysis, which many of the interviewees hoped for. Another problem is that the results should be easily distributed in forms of different reports, but at the moment, this is not the case.

In TS-Group, Ltd., a data warehouse pilot has been implemented. The pilot enables versatile reporting and, as the interviewees mentioned, this opportunity should not be lost. However, some reporting problems still exist. For example, the distribution and the format of the reports should be solved. In the department of financial management, the reporting system does not support the latest operating systems and, therefore, only some computers are capable of using the system.

In Optiroc, Ltd., reports are generated directly from the operational databases that are not designed for reporting purposes. One problem attendant on this design is that the reports run slowly. In principle, users should also be able to create their reports by themselves, but in reality, only a few of them are able to. Maybe this is one reason that the interviewees presented very critical comments on reporting. The interviewees mentioned as well that the implementation of a data warehouse system is strongly supported and is seen as a solution for prob-