Chapter 4  Data Management: Concepts and SAP Solutions

This chapter analyzes the possibilities for data management based on data warehouse solutions. First, general concepts and methods for setting up a data warehouse are discussed. Possible applications of the SAP Business Information Warehouse are then introduced.

4.1  The Data Warehouse Concept

4.1.1  Principles of the Data Warehouse Concept

Decisions in companies require the knowledge of current and historical data. These accumulate in all areas of the company as well as in the company environment and refer to sales, products and parts, work centers, consumption activities and payment transactions as well as to the behavior of customers and vendors, for example.

Data processing usually makes a distinction between transaction systems (administration and material requirements planning systems) and planning and control systems (e.g., data warehouse systems) (Meier et al. 2003).

The central task of transaction systems is the fast processing of activities and/or transactions which generate small datasets (e.g., recording of entries). Since the underlying databases — nowadays usually relational systems — are continuously updated, for example, when new records are inserted, old records are updated, corrected or deleted, the literature also speaks of OLTP (On-Line Transaction Processing) systems.

A disadvantage regarding decision support is the fact that historical data is no longer available for analyses because data records are being overwritten. In
addition, the preparation of the stored data is in many cases made difficult due to
the fact that in the areas of a company, for example production and materials
management, different transaction systems and thus several databases exist, which
in most cases have grown over the years and cover specific company
requirements. Also, OLTP systems can only inadequately meet many
requirements, for example, for a quick, flexible, multidimensional analysis of data
(Codd 1993). This is where the data warehouse concept applies with respect to a
decision support.

By using a data warehouse (for the specific requirements regarding such solutions,
refer to Inmon 1996) the goal is pursued to collect data from different OLTP
systems and to make it available for analyses as factor constant over time. Modern
data warehouse solutions provide tools both for the extraction of the data from the
OLTP systems and for the analysis of the data contained in a data warehouse (see
section 4.2).

A special feature of a data warehouse is the multidimensional data management
based on the On-Line Analytical Processing (OLAP) concept (Codd 1993). Here
the main idea is the representation of data in data cubes which are composed of
key figures and characteristics. While key figures, such as sales or profit, quantify
business correlations, characteristics are used to analyze these from different
points of view, for example by period.

The representation in form of such data cubes now provides several possibilities
for data analysis. This includes:

- **Drill Down**: Switch to a deeper consolidation level (detailed analysis).
- **Drill Up**: Switch to a higher consolidation level (more abstract analysis)
- **Slicing**: If data of a cube with at least three dimensions is to be displayed in a
table, the data must be projected to one level.
- **Dicing**: This method only deals with certain instances of the data cube
attributes, that is, smaller data cubes.

This way, certain analyses can be carried out more easily than with relational
database systems, which reveal shortcomings in particular with regard to the
changeability of the data view as well as the representation of different detail
levels.

The multidimensional data management in data warehouse solutions can be
implemented with the help of two different procedures (Bellatreche et al. 2001):

- **ROLAP** (Relational On-Line Analytical Processing) systems are based on a
conventional relational database (storage of the data in two-dimensional
tables). By successively arranging different instances of a table n times (e.g.,
52 week sales of a product), a virtual data cube is created. Reading a large
number of instances of a table in the case of a data analysis is time-consuming.