DWMS: Data Warehouse Management System

Narendra Mohan
Software AG of Far East, Inc.

After a thorough investigation of end user requirement differences from a Database and a Data Warehouse, a strong case is made for specific Data Warehouse Management Systems (DWMS), in contrast to the trend of utilizing DBMS for Data Warehouse construction. The three type (Temporal, Nested Relational, Multidimensional) DWMS hypothesis, an original concept of Yoshioki Ishii, President, Software AG of Far East, Inc. is explained in detail. Customer solutions and related products are introduced.

From DBMS to DWMS

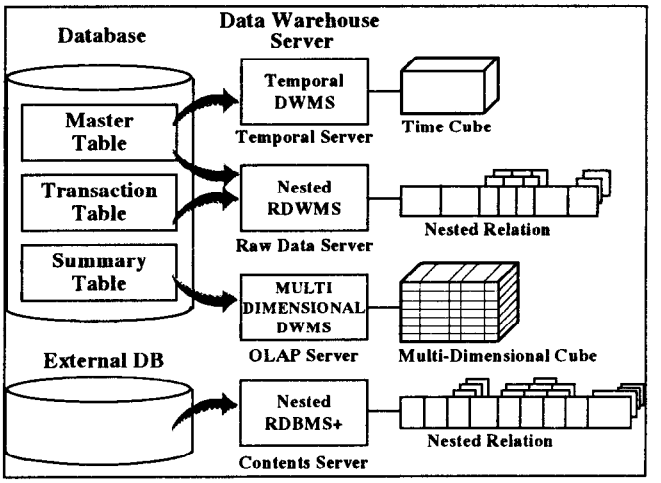
Data contained in a Database can be broadly classified into master type (fact) data, transaction type (event) data and summary type (aggregation) data. Attributes such as the number of fields, volume of updates, amount of data appends, and count of multiple occurrence fields differ vastly for each of the above. A Data Warehouse is basically constructed by systematic accumulation of the data that is originally stored in Database(s). For optimally utilizing the data warehousing concept, however; fundamental differences between Database and Data Warehouse such as current vs. historical data, large volume vs. very large volume data, mission critical vs. decision support application, etc. must be reviewed. Rather than adopting RDBMS for transaction based as well as information based applications, a clear distinction, DBMS for the former and DWMS for the latter, is recommended. In view of the distinct data types, an innovative DWMS design, deemed essential for achieving the intuitive comprehensibility and performance levels that end users aspire, is proposed.

Nested Relation over Normalization

Normalization methodology was intorduced for shrinking the overall database size, easing updates, and effecting program/data independence. Ironically, a Data Warehouse assumes expansive size, does not necessitate dynamic updates, and is used primarily for analytical rather than repetitive processes. Hence, normalization is not only futile but is in fact counterproductive towards factors like performance and ease of use. The numerous join processes generated between normally distributed tables causes slow and cryptic query systems (SQL based) with non-instinctive application logic for Data Warehouse users. The Nested Relational Model is presented as an ideal solution.

Proceedings of the 22nd VLDB Conference Mumbai(Bombay), India, 1996

Types of DWMS



While a DBMS is optimal when centralized, DWMS must be differentiated according to requirements. Corresponding to the database file types, three DWMS varieties, Temporal DWMS, Nested Relational DWMS and Multidimensional DWMS composed respectively of the Time Cube, Nested Relation, and Multi-Dimensional Cube as the fundamental units of data are introduced. For Data Warehouse applications, master type data (time variable inclusive) is managed in the Temporal DWMS, transaction type data (after undergoing normalization by merging with relevant master type data) in the Nested RDWMS, and summary type data in the Multidimensional DWMS.

Tools and Product Offerings

Customer solution packages with ADABAS (a high performance Nested Relational Model based database) as the Nested RDWMS, Essbase as the Multidimensional DWMS, SOAR as the Nested RDBMS+ Servers respectively, and DB-FRONT as an end user tool have been marketed with remarkable success in Japan. An original Temporal DWMS product is currently under development.