

# **OLAP Council White Paper**

## **Introduction**

The purpose of the paper that follows is to define On-Line Analytical Processing (OLAP), who uses it and why, and to review the key features required for OLAP software as referenced in the OLAP Council benchmark specification.

## **What Is OLAP?**

During the last ten years, a significant percentage of corporate data has migrated to relational databases. Relational databases have been used heavily in the areas of operations and control, with a particular emphasis on transaction processing (for example, manufacturing process control, brokerage trading). To be successful in this arena, relational database vendors place a premium on the highly efficient execution of a large number of small transactions and near fault tolerant availability of data.

More recently, relational database vendors have also sold their databases as tools for building Data Warehouses. A Data Warehouse stores tactical information that answers "who?" and "what?" questions about past events. A typical query submitted to a Data Warehouse is: "What was the total revenue for the eastern region in the third quarter?"

It is important to distinguish the capabilities of a Data Warehouse from those of an OLAP (On-Line Analytical Processing) system. In contrast to a Data Warehouse, which is usually based on relational technology, OLAP uses a multidimensional view of aggregate data to provide quick access to strategic information for further analysis.

OLAP enables analysts, managers, and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information. OLAP transforms raw data so that it reflects the real dimensionality of the enterprise as understood by the user.

While OLAP systems have the ability to answer "who?" and "what?" questions, it is their ability to answer "what if?" and "why?" that sets them apart from Data Warehouses. OLAP enables decision-making about future actions. A typical OLAP calculation is more complex than simply summing data, for example: "What would be the effect on soft drink costs to distributors if syrup prices went up by \$.10/gallon and transportation costs went down by \$.05/mile?"

OLAP and Data Warehouses are complementary. A Data Warehouse stores and manages data. OLAP transforms Data Warehouse data into strategic information. OLAP ranges from basic navigation and browsing (often known as "slice and dice"), to calculations, to more serious analyses such as time series and complex modeling. As decision-makers exercise more advanced OLAP capabilities, they move from data access to information to knowledge.

### **Who Uses OLAP and Why?**

OLAP applications span a variety of organizational functions. Finance departments use OLAP for applications such as budgeting, activity-based costing (allocations), financial performance analysis, and financial modeling. Sales analysis and forecasting are two of the OLAP applications found in sales departments. Among other applications, marketing departments use OLAP for market research analysis, sales forecasting, promotions analysis, customer analysis, and market/customer segmentation. Typical manufacturing OLAP applications include production planning and defect analysis.

Important to all of the above applications is the ability to provide managers with the information they need to make effective decisions about an organization's strategic directions. The key indicator of a successful OLAP application is its ability to provide information as needed, i.e., its ability to provide "just-in-time" information for effective decision-making. This requires more than a base level of detailed data.

Just-in-time information is computed data that usually reflects complex relationships and is often calculated on the fly. Analyzing and modeling complex relationships are practical only if response times are consistently short. In addition, because the nature of data relationships may not be known in advance, the data model must be flexible. A truly flexible data model ensures that OLAP systems can respond to changing business requirements as needed for effective decision making.

Although OLAP applications are found in widely divergent functional areas, they all require the following key features:

- Multidimensional views of data
- Calculation-intensive capabilities
- Time intelligence

### **Multidimensional Views**

Multidimensional views are inherently representative of an actual business model. Rarely is a business model limited to fewer than three dimensions. Managers typically look at financial data by scenario (for example, actual vs. budget), organization, line items, and

time; and at sales data by product, geography, channel, and time.

A multidimensional view of data provides more than the ability to "slice and dice"; it provides the foundation for analytical processing through flexible access to information. Database design should not prejudice which operations can be performed on a dimension or how rapidly those operations are performed. Managers must be able to analyze data across any dimension, at any level of aggregation, with equal functionality and ease. OLAP software should support these views of data in a natural and responsive fashion, insulating users of the information from complex query syntax. After all, managers should not have to understand complex table layouts, elaborate table joins, and summary tables.

Whether a request is for the weekly sales of a product across all geographical areas or the year-to-date sales in a city across all products, an OLAP system must have consistent response times. Managers should not be penalized for the complexity of their queries in either the effort required to form a query or the amount of time required to receive an answer.

The OLAP Council APB-1 performance benchmark tests a server's ability to provide a multidimensional view of data by requiring queries of varying complexity and scope. Basic aggregation is performed on some of the dimensions (product, customer, and channel). More complex calculations are performed on other dimensions. The measure dimension computes ratios and averages. Variances are computed along the scenario dimension. A complex model based on historical performance is used to compute the forecast scenario. Consistently quick response times to these kinds of queries are key to establishing a server's ability to provide multidimensional views of information.

### **Complex Calculations**

The real test of an OLAP database is its ability to perform complex calculations. OLAP databases must be able to do more than simple aggregation. While aggregation along a hierarchy is important, there is more to analysis than simple data roll-ups. Examples of more complex calculations include share calculations (percentage of total) and allocations (which use hierarchies from a top-down perspective).

Key performance indicators often require involved algebraic equations. Sales forecasting uses trend algorithms such as moving averages and percentage growth. Analyzing the sales and promotions of a given company and its competitors requires modeling complex relationships among the players. The real world is complicated -- the ability to model complex relationships is key in analytical processing applications.

OLAP software must provide a rich tool kit of powerful yet succinct computational methods. To make developers more efficient and business users more self-sufficient, the

vehicle for implementing computational methods should be clear and non-procedural. If the method for creating the desired calculations is not clear, development time and/or usage will suffer. If the calculation method is procedural, changes to the system cannot be done in a timely manner, effectively eliminating access to just-in-time information.

Whereas transaction processing systems are judged on their ability to collect and manage data, analytical processing systems are judged on their ability to create information from data. The OLAP Council APB-1 performance benchmark contains a representative selection of calculations, both simple and complex. An example of a simple calculation contained in the performance benchmark is the calculation of margin (sales minus costs). The computation of the forecast is the most complex calculation contained in the current version of the performance benchmark. Historical data is used to project the future and aggregate data is used to estimate input data. Other more complex calculations, such as allocations and trend analysis, are also often found in OLAP systems.

### **Time Intelligence**

Time is an integral component of almost any analytical application. Time is a unique dimension because it is sequential in character (January always comes before February). True OLAP systems understand the sequential nature of time. Business performance is almost always judged over time, for example, this month vs. last month, this month vs. the same month last year.

The time hierarchy is not always used in the same manner as other hierarchies. For example, a manager might ask to see the sales for May or the sales for the first five months of 1995. The same manager might also ask to see the sales for blue shirts but would never ask to see the sales for the first five shirts. Concepts such as year-to-date and period over period comparisons must be easily defined in an OLAP system.

In addition, OLAP systems must understand the concept of balances over time. For example, if a company sold 10 shirts in January, five shirts in February, and 10 shirts in March, then the total balance sold for the quarter would be 25 shirts. If, on the other hand, a company had a head count of 10 employees in January, only five employees in February, and 10 employees again in March, what was the company's employee head count for the quarter? Most companies would use an average balance. In the case of cash, most companies use an ending balance.

The OLAP Council APB-1 performance benchmark contains examples of how time is used in OLAP applications. Smoothed sales are computed as a three month moving average. Inventory is aggregated as an ending balance. The forecast calculation uses this year's vs. last year's knowledge, year-to-date knowledge, and annualization factors.

## OLAP Benefits

Successful OLAP applications increase the productivity of business managers, developers, and whole organizations. The inherent flexibility of OLAP systems means business users of OLAP applications can become more self-sufficient. Managers are no longer dependent on IT to make schema changes, to create joins, or worse. Perhaps more importantly, OLAP enables managers to model problems that would be impossible using less flexible systems with lengthy and inconsistent response times. More control and timely access to strategic information equal more effective decision-making.

IT developers also benefit from using the right OLAP software. Although it is possible to build an OLAP system using software designed for transaction processing or data collection, it is certainly not a very efficient use of developer time. By using software specifically designed for OLAP, developers can deliver applications to business users faster, providing better service. Faster delivery of applications also reduces the applications backlog.

OLAP reduces the applications backlog still further by making business users self-sufficient enough to build their own models. However, unlike standalone departmental applications running on PC networks, OLAP applications are dependent on Data Warehouses and transaction processing systems to refresh their source level data. As a result, IT gains more self-sufficient users without relinquishing control over the integrity of the data.

IT also realizes more efficient operations through OLAP. By using software designed for OLAP, IT reduces the query drag and network traffic on transaction systems or the Data Warehouse.

Lastly, by providing the ability to model real business problems and a more efficient use of people resources, OLAP enables the organization as a whole to respond more quickly to market demands. Market responsiveness, in turn, often yields improved revenue and profitability.

*Sarah Forsman is vice president of marketing of Symmetry Corporation, an On-Line Analytical Processing (OLAP) consulting firm in San Rafael, CA. She has 15 years of experience marketing and selling both OLAP and relational software and services.*