

Multidimensional Basics

Introduction to Multidimensional Basics

Incessant progress and dynamics in the economy opened the door to a lucrative opportunity for vendors of various technologies that handle an organization's day-to-day transactional operations. Old paper-based systems for tracking operational results are not only grueling, but it is prone to much alteration, rendering the results less reliable. To top it all, preservation of data in this form has many drawbacks, since amount of data becomes extremely cumbersome and hard to maintain.

Consequently, as a business' size and market broadens, so does the mounting volume of data required to be handled. Hence, several database applications and operational software have emerged, brandishing capabilities for validation, storage, retrieval and other kinds of transactional data maintenance.

Information processing is the backbone of a business organization because it is where company-oriented decisions are based. For that reason, good information processing came to be the vital factor for the survival of any business organization. However, good information comes in various states. The need to rely and analyze not only the past and current but also the projected future data is very essential in arriving at a good decision. Just as evaluation of operational performance is a key to maintaining a smoothly running business, so is the ability to make predictions on the possible movements and trends in the industry which normally involves formulation of assumptions, backed by use of derived descriptions, application of certain algorithms and using other information apart from those provided by the actual transactional results. From here, the limiting factors of operational software surfaced, because operational software is unable to provide anything other than historical data. This triggered the birth of several analysis applications in support of analysis-oriented processing.

How then can businesses take advantage of the benefits derived from utilization of these technologies? The system put in place by most companies allows a user to pull, extract, or 'query' data based on particular information requirements in support of some ad-hoc analysis directed toward a specific objective. Not only do business organizations put these systems into use for their regular operating activities, but they tend to heavily rely on the information provided by these applications for reporting and analysis purposes.

Typically, SQL databases and Spreadsheet technologies are being employed by most users in analysis and reporting activities. However, technology developers and users have found some flaws and encountered certain difficulties or limitations in using either of the two applications as they do not fully support the analytical requirements of end-users.

For example, the use of [spreadsheets](#) is most convenient for the majority of end users, as most people are familiar and have been exposed to this type of application. But although users can move comfortably in a spreadsheet environment, performing an analysis of data

in this format can prove to be a very challenging task. Often, users need to refer to multiple numbers of worksheets, particularly when dealing with large volumes of data. Apart from that, performing a complex calculation can also be a major problem, as spreadsheets have only limited calculation functions and features. Spreadsheets are able to handle cell-based formulas, but a user is still required to redefine the formula for the same intersections of data when working on multiple worksheets. Also, spreadsheets also do not fully support hierarchies and multiple hierarchies. Additionally, spreadsheets remain inflexible and do not present an easy and quick method of restructuring views. In instances where a user wants to generate a different view of the data, he will be compelled to redo everything, including redefining formulas and necessary look up tables.

On the other hand, a SQL Database application uses a structured query language to pull specific data from a database source. Unfortunately, many users find it difficult to handle this type of application. As most end users are non-technical people, query language is unnatural for them as it is complex and technically sophisticated. To address this problem, accessible query templates and easy-to-use query tools have been provided. In spite of this, many are still indisposed to use this application because there are still situations where the predefined query templates and tools are not in sync with a user's specific needs. Oftentimes, when dealing with this type of application, users are inclined to be highly dependent on technical support, which suppresses the ability of information to be readily available in a timely manner.

The Concept of Multidimensionality

The problems with analyzing and reporting data from operating systems through SQL highlight the importance of the concept of multidimensionality. In very few words, multidimensionality is the ability to manipulate the view of data in various ways without disrupting the integrity of the data. It is extremely important that users have the ability to customize and change their views of data to suit their specific needs and enable users to perform any type of analysis within the database.

Data or information handled by a small organization varies not only in volume but also in complexity as compared to a medium- to large-sized organization. Large companies would most likely, for instance, have a wider array of products, broader market and consumers, many outlets in different locations, several suppliers, a larger employee count and large companies handle more denominations and measures. Also since there are different users of information, whom may come from different levels within the organizational structure, they can differ not only in the needed views of data but also have varying reporting requirements and need different reporting schemas.

When we talk about a multidimensional database, the foremost consideration would be the dimension, which happens to be the basic structure of a multidimensional application. Each dimension has a structure, a set of members for each category, and a child/parent

relationship between members. Fields of data may be organized into logical categories such as Time, Geography, Customers, Products and Measures, to name a few examples of dimensions. Some dimensions are arranged following a certain hierarchy structure. Thus, in a geographic dimension, local branches may be consolidated into regions, areas and countries. In such case where a hierarchical dimension is present, we will regard the entire hierarchy as one single dimension.

Normally, a user may want to pull only a particular set of data. For example, a sales manager would want to infer how well their sales operations are doing, and he wants to break the information down by products, by location, by customer and by season. Every time we mention the word “by” we are talking about adding another dimension to the analysis. In this case, the sales manager needs only to pull the data from specific dimensions, only those which are necessary for his analysis requirements.

We may represent a set of data with predefined dimensions as a cube which closely resembles a Rubix cube. Each face of the cube represents a particular dimension; the dimensions being coexistent with each other at every point and, at the same time, are also independent from each other. Similar to twisting a Rubix cube, a user can also manipulate a multidimensional cube in several ways to view data. The key element here is that regardless of the way the data is viewed or displayed, the same intersections of data will still show the same measure, meaning the actual data is preserved.

In some multidimensional database applications such as an OLAP product, formulas are defined inside the cube, or set of dimensions. Formulas assign an equation to be performed on a specific portion of a cube, which is much faster and more effective than defining a formula in a spreadsheet application based on a specific cell. By applying the formulas directly to the data in the cube, the multidimensional database application eliminates the redundancy of redefining formulas in Excel, for example. And, heedless of the way the data is being viewed or organized, the formula will still apply to the same portion of the cube.

One technique for manipulating the data view is Rotation or more commonly known as Slicing and dicing. Slicing and dicing is a function that enables users to see the different faces of a cube in a single view by displaying them in a flat spreadsheet-like format. The reason for employing a spreadsheet-like view of data is that a computer screen is only limited to providing a flat layout of data. This permits users to customize the output of data by moving the dimensions from rows to columns depending on a specific layout preferred by user.

Although data is presented in a two-dimensional or flat layout, it doesn't necessarily imply that only two dimensions can be displayed at once. Actually, it is very possible that there are more dimensions present in the data view. Some dimensions may simply be nested inside another dimension. How is this possible? To explain it more clearly, nesting dimensions is merely combining the intersections of dimensions to show more than two dimensions of information in a single view. This is done by nesting the dimensions either in rows, columns or both.

For example a user may want to see comparatively the Sales, Cost and Profit results of different branches for the past three years. Here we are dealing with the dimensions Year, Accounts (Aggregation of members Sales, Cost and Profit), Branch, and Measure. Rather than presenting the needed information using three different reports, the user can nest one dimension under another. For example Branch may be nested under Accounts dimension along the rows and display year along the columns.

		2005	2006	2007
Sales	Branch 1	345,798.00	354,875.00	349,212.00
	Branch 2	597,545.00	645,231.00	591,311.00
	Branch 3	248,846.00	281,346.00	346,441.00
Cost	Branch 1	122,554.00	124,112.00	122,995.00
	Branch 2	264,465.00	284,643.00	264,111.00
	Branch 3	97,883.00	102,315.00	135,422.00
Profit	Branch 1	223,244.00	230,763.00	226,217.00
	Branch 2	333,080.00	360,588.00	327,200.00
	Branch 3	150,963.00	179,031.00	211,019.00

What happens is that Accounts and Branch dimension will be displayed simultaneously on the rows. This enables the user to directly compare the information for the past three years by looking at a single data only. This is an optimized view for analysis purposes.

In this short introduction, you can start to grasp the significant role that multidimensionality plays in the business world.

Want to know more about how a multi-dimensional solution can help transform your data into information? Do you want to create an application that can impact how you run your business? To hear more about a solution that's right for your company, contact Vince Kelly, President, Kelly Solution Providers, Inc., (215) 340-2481 or vze2gmct@verizon.net.