
WHY OLAP – AN OVERVIEW OF ON-LINE ANALYTICAL PROCESSING

A DATABEACON.COM WHITE PAPER

On-Line Analytical Processing (OLAP) tools meet the need for interactive multidimensional reporting and analysis. They allow operational managers to perform trend, comparative, and time-based analysis by enabling exploration of pre-calculated and summarized data along multiple dimensions. Operational managers can explore data first at a summary level, then drill down through the data hierarchy to examine increasingly granular levels of detail.

This document provides a brief description of On-Line Analytical Processing, its history, and some of the main features that define it.

EXECUTIVE SUMMARY

The diversity and pace of today's business require complementary tools that support greater variability of use and dynamic interaction with the data to support operational managers as they explore and evaluate inter-relationships in the data. An operational manager doing trend analysis on a particular product line would require a mountain of static reports to accommodate his/her analysis needs, an approach that is not desired or sustainable.

On-Line Analytical Processing (OLAP) tools meet the need for interactive multidimensional reporting and analysis. They allow operational managers to perform trend, comparative, and time-based analysis by enabling exploration of pre-calculated and summarized data along multiple dimensions. Operational managers can explore data first at a summary level, then drill down through the data hierarchy to examine increasingly granular levels of detail.

This document provides a brief description of On-Line Analytical Processing, its history, and some of the main features that define it.

TABLE OF CONTENTS

SPECIAL TOOLS FOR SPECIAL NEEDS

1

ON-LINE ANALYTICAL PROCESSING

2

MULTIDIMENSIONAL

2

CONSISTENTLY FAST

3

INTUITIVE INTERFACE

4

FLEXIBLE ANALYSIS

6

SUMMARY

6

INTRODUCTION

Data is the life-blood of all organizations. It is constantly collected, manipulated, managed and explored by operational managers and other employees as they evaluate the health and operation of the business. They review past and present metrics such as cost and revenue to make decisions that will improve the performance and profitability of their organization.

The responsibility and challenge of collection and delivery of this data in a meaningful form traditionally belongs to the I.T. department. This is no small challenge as both data volume and the number of operational managers requesting access are growing exponentially. This rapid growth and an Internet-paced business environment are causing many I.T. purchase decisions to be driven by the "frontline" needs of the operational managers tasked with the delivery of revenue-enhancing or cost-saving initiatives. And in some cases, the owners of those initiatives are bypassing central I.T. entirely to achieve their objectives. Whether departmental or I.T. rooted, organizations are constantly looking for new methods of delivery and new methods to quickly distill exploding data into useful, actionable information.

SPECIAL TOOLS FOR SPECIAL NEEDS

From the beginning of data collection, we have had reporting tools - production, managed and ad hoc - that allow report authors and operational managers such as the VP Finance to access, navigate and explore relational data and quickly create reports with minimal

understanding of the underlying database language, connectivity and functionality. These tools evolved in capability and audience as we moved from database reporting, to Decision Support Systems (DSS), to Executive Information Systems (EIS) and now to Business Intelligence (BI) and Business Performance Management (BPM). The reporting solutions commonly offered in these toolsets invariably offer some snapshot of the data in a two-dimensional, static view. A powerful example is an Income Statement, a periodic static view of the business for review by the VP Finance and other executives.

The diversity and pace of today's business require complementary tools that support greater variability of use and dynamic interaction with the data to support operational managers as they explore and evaluate interrelationships in the data. An operational manager doing trend analysis on a particular product line would require a mountain of static reports to accommodate his/her analysis needs, an approach that is not desired or sustainable. As seen in Figure 1, operational managers need multiple degrees of freedom in their exploration and analysis, as made available in dynamic reports, and in particular, interactive multidimensional reports.

On-Line Analytical Processing (OLAP) tools meet the need for interactive multidimensional reporting and analysis. They allow operational managers to perform trend, comparative, and time-based analysis by enabling exploration of pre-calculated and summarized data along multiple dimensions. Operational managers can

THE DIVERSITY AND PACE OF TODAY'S BUSINESS REQUIRE COMPLEMENTARY TOOLS THAT SUPPORT GREATER VARIABILITY OF USE AND DYNAMIC INTERACTION WITH THE DATA TO SUPPORT OPERATIONAL MANAGERS AS THEY EXPLORE AND EVALUATE INTERRELATIONSHIPS IN THE DATA

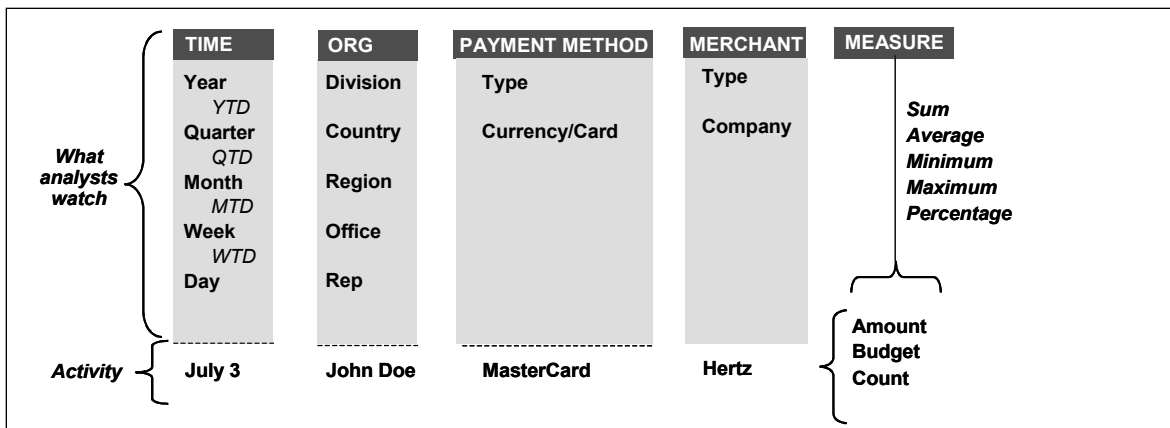


Figure 1: While activity data is usually captured at a detailed level, analysts most frequently work dimensionally at some level of summarization.

THE KEY TO OLAP IS MULTIDIMENSIONALITY, PROVIDING OPERATIONAL MANAGERS WITH THE ABILITY TO INTERACT WITH AND VIEW INFORMATION FROM A NUMBER OF PERSPECTIVES AND AGGREGATIONS TO MAKE CRITICAL BUSINESS DECISIONS WITH MINIMAL COMPUTING EXPERTISE

explore data first at a summary level, then drill down through the data hierarchy to examine increasingly granular levels of detail.

ON-LINE ANALYTICAL PROCESSING

In 1962 Ken Iverson first addressed multidimensional reporting in A Programming Language (APL), which was later adopted and implemented by IBM. It recognized the dimensionality of data, and offered techniques on how to process and display it. Further, it targeted operational managers; in particular, business analysts. The analyst community accepted and adopted OLAP technology because multidimensionality is a concept it could instinctively grasp. Operational managers tend to think in metrics, dimensions and hierarchies, as opposed to abstract fields, records and tables. OLAP enabled these operational managers to set up their own analyses, deploy their own solutions, often with the blessing of the overworked and understaffed I.T. department.

In 1993, E.F. Codd authored a seminal paper on multidimensional analysis, and in so doing, coined the term OLAP for On-Line Analytical Processing. He also defined 12 rules for OLAP, and set the basis for many of the products that followed, either as distinct solutions or as part of more complex platforms. These tools, in one way or another, offer a powerful calculation engine that creates/reads a multidimensional file or cube to enable operational managers to analyze data across multiple dimensions.

Databeacon is an example of an OLAP product, albeit different in that it is the first product purposely architected for Web Reporting and Data Analysis over the Internet via no-touch smart client deployment. Its innovative Client OLAP Reporting Environment (CORE) and smart client report viewers offer unique advantages in several areas including scalability, integration and deployment. From Codd's list of twelve rules, the following four most clearly differentiate OLAP from production, managed and ad hoc reporting.

1. Multidimensional

There is a multidimensional data structure and view that allows operational managers to analyze numerical values from different perspectives, e.g. product, time, and geography.

2. Consistently fast

To ensure fast, predictable query times, OLAP

vendors pre-aggregate data. This is done using pre-aggregated relational tables, or through a highly compressed multidimensional file known as a cube.

3. Intuitive interface

Skilled analysts and non-technical users alike can quickly manipulate and analyze data and generate reports without involving their organization's I.T. department.

4. Complex calculations

With multiple dimensions come more complex, cross-dimensional calculations. An analysis session might require the subtotal of sales for a particular state to be expressed as a percentage of the whole country, or for that singular product, or both! Further, this result may require presentation as part of a time-series analysis, i.e. current quarter versus last quarter versus a year ago.

MULTIDIMENSIONAL

The key to OLAP is multidimensionality, providing operational managers with the ability to interact with and view information from a number of perspectives and aggregations to make critical business decisions with minimal computing expertise. There are three basic dimensions - categories, time and measures.

Categories offer the distinct identifiers or classifications on which measures (values) are based. Bob, Ted and Melissa are all categories of Salespeople on whom we attribute and measure revenue. But the multidimensionality of categories means that their performance can be aggregated and viewed at a Sales Office level, a City level, a State level, a Country level, and a Business Region level. Alternately, you can start at a Region level and drill down from the summary to any lower level of detail, analyzing and comparing results as you go.

Time is a special form of category where the OLAP tool is aware that the value is an ordinal, and offers special decomposition and arithmetic rules to support the data type. A single date can be viewed individually, or separated as a day, month, quarter or year value. This OLAP feature is essential for trend and seasonality analysis. Advanced OLAP tools will further offer relative time support, permitting operational managers to compare as an example, "current" month against "previous" month, or Year-to-Date against Previous-Year-to-Date.

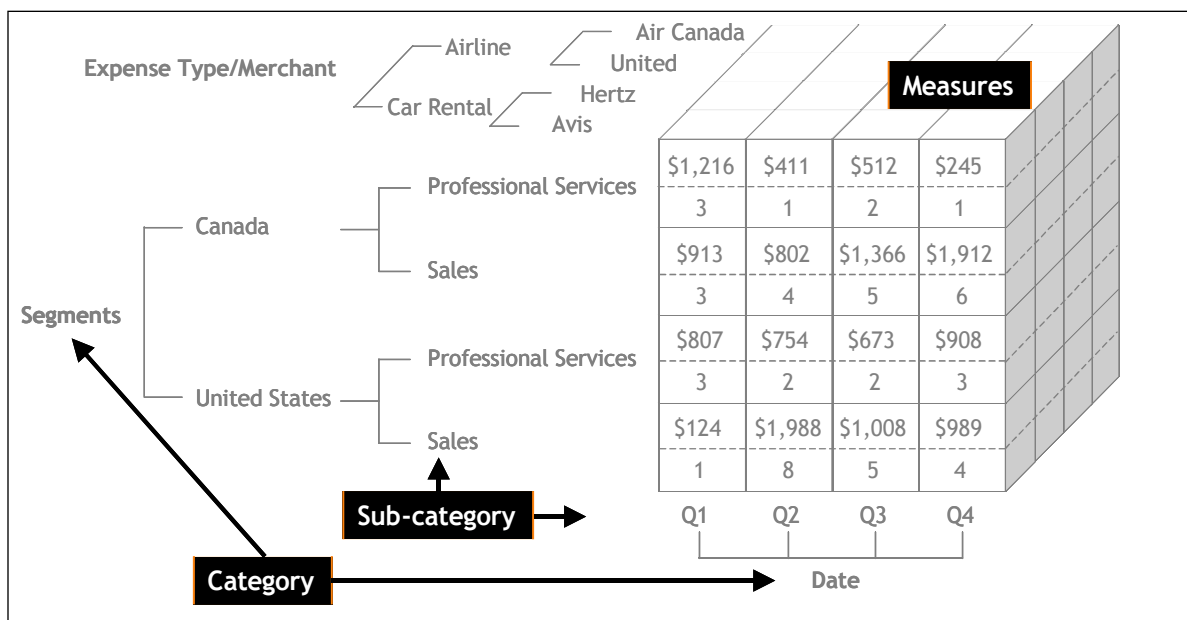


Figure 2: Multidimensional data and its aggregation as illustrated in a cube design

Measures refer to the numbers, the values of a particular category's performance. This can be revenue, costs, salary, discounts and more. Measures can also be constructions made from the raw data, values such as growth, variance and performance ratios.

While categories, time and measures are the core elements of an OLAP design, how they are defined and presented to the operational manager will vary from product to product. Some tools ask that you select your categories from the database metadata layer, and others require you to define a brand new layer of metadata. One of the more common methods used by vendors is to define the multidimensional structure of how data is to be presented, and then use that definition against the data to build multidimensional data cubes. Figure 2 offers a "cube" illustration of this file. The technique enables solutions that can summarize large volumes of data into small-footprint cubes.

Databeacon is an example of a product that has a "Publisher" to define the architecture, and an "Insight" viewer (in addition to two other more entry-level viewers in Databeacon Smart Client Standard and Databeacon Smart Client Professional) to analyze and report on the data content. Extending beyond similar tools, Databeacon can both design and build cubes on

an ad hoc basis under program control, and can create highly compressed analytical reports that are easily deployed for fast distributed OLAP processing at an operational manager's PC.

CONSISTENTLY FAST

OLAP tools will invariably deliver fast response time as compared to the cycle of generating reports and spreadsheets, however the response that an operational manager will see will vary by product, and by load - both data and users.

When dealing with large volumes of data, OLAP vendors will pre-aggregate the data to ensure fast, predictable query times. This can be done using pre-aggregated relational tables, or through a compressed multidimensional cube. But "predictable" response time does not mean "consistent" response time. Operational managers will not necessarily understand the complexities that go into resolving a multi-table SQL statement, but they will definitely not be happy with varying response times as they analyze and explore the data. It is a leading cause of "shelfware" - software no one is interested in using anymore.

A multidimensional solution offers a more consistent response, although even its response time will degrade if the solution is deployed in a client/server architecture and there are a

OLAP TOOLS WILL INVARIABLY DELIVER FAST RESPONSE TIME AS COMPARED TO THE CYCLE OF GENERATING REPORTS AND SPREADSHEETS, HOWEVER THE RESPONSE THAT AN OPERATIONAL MANAGER WILL SEE WILL VARY BY PRODUCT, AND BY LOAD - BOTH DATA AND USERS

THE POWER OF OLAP IS IN THE OPERATIONAL MANAGER'S ABILITY TO INTERACT WITH THE PRESENTATION OF THE DATA, ITERATIVELY CHANGING THE DIMENSIONS, MEASURES, GRANULARITY, NESTING AND GROUPINGS OF DATA AS IMPORTANT BUSINESS INSIGHTS GET DISCOVERED

large number of users accessing an overworked server. Newer solutions such as Databeacon employ a smart client architecture to deliver the scalability and fast response time required of Web-based no-touch deployments.

INTUITIVE INTERFACE

While multidimensionality is at the heart of OLAP, the power is in the operational manager's ability to interact with the presentation of the data, iteratively changing the dimensions, measures, granularity, nesting and groupings of data as important business insights get discovered. OLAP delivers the untold degrees of freedom you as an operational manager need to quickly and accurately explore business data.

Let's look at this through an example - the Sales example that ships with Databeacon Open Client Standard. The example shows several performance measures for a variety of products, sold at different stores in cities in two states. It also offers information on whether the transaction was paid for in cash or with a credit card.

Figure 3 illustrates a report showing quarterly sales performance of freezers at the Home Store. The reports shows Quantity sold, Sales revenue and Cost of Goods sold, three measures that can be used to evaluate sales performance. But this is a static view, offering a

data perspective that could be used next quarter to see if there is any change. Valuable, but limited. With OLAP, and in particular

Databeacon, operational managers can expand the view, and get more insight out of the data. Complex calculations are performed as the cube is built, or by the individual operational manager as they explore the multidimensional data pertinent to their need. In figure 4, calculations are used to create two new performance measures, Profit and Average Sell Price, allowing the operational manager to more quickly compare and understand business performance.

These measures can also be viewed at any of the dimensional levels defined for any of the categories in the cube. In figure 5, we see the Date category showing measures for the Date total, the year 2000, Q1 2000, January and February 2000, and several days in February. While this is impressive to see the aggregation applied for all measures (real and calculated) at each level of the Date dimension, it is more so when you realize that Date is a Time category type, and that the input data was simply a single date column. The OLAP tool read the date field, automatically decomposed it into year, quarter, month and day levels, and then aggregated the measures for each.

Year 2000 Quarterly Freezer Sales								
Date	+	Q1 2000	+	Q2 2000	+	Q3 2000	+	Q4 2000
Quantity		46		28		22		44
Sales		\$ 16,581.42		\$ 9,959.56		\$ 7,868.94		\$ 15,917.88
Cost of Goods		\$ 10,943.72		\$ 6,573.32		\$ 5,193.53		\$ 10,505.81

Figure 3: Databeacon report shows quarterly sales of freezers at the Home Store in the year 2000

Year 2000 Quarterly Freezer Sales								
Date	+	Q1 2000	+	Q2 2000	+	Q3 2000	+	Q4 2000
Quantity		46		28		22		44
Sales		\$ 16,581.42		\$ 9,959.56		\$ 7,868.94		\$ 15,917.88
Cost of Goods		\$ 10,943.72		\$ 6,573.32		\$ 5,193.53		\$ 10,505.81
Profit		\$ 5,637.70		\$ 3,386.24		\$ 2,675.41		\$ 5,412.07
Average Sell Price		\$ 360.47		\$ 355.70		\$ 357.68		\$ 361.77

Figure 4: Calculations create new measures like Profit and Average Selling Price to better illustrate and compare business performance results

		Freezer Sales						
Date	Quantity	Date	2000	Q1 2000	Jan 2000	Feb 2000	Feb 01 2000	Feb 06 2000
	Quantity		229	140	46	4	17	2
	Sales		\$ 81,950.33	\$ 50,327.80	\$ 16,581.42	\$ 1,407.08	\$ 6,135.09	\$ 723.54
	Cost of Goods		\$ 54,087.26	\$ 33,216.38	\$ 10,943.72	\$ 928.68	\$ 4,049.15	\$ 477.54
	Profit		\$ 27,863.07	\$ 17,111.42	\$ 5,637.70	\$ 478.40	\$ 2,085.94	\$ 246.00
	Average Sell Price		\$ 357.86	\$ 359.48	\$ 360.47	\$ 351.77	\$ 360.89	\$ 361.77

Figure 5: Date columns are automatically dimensioned by Databeacon Collaboration Edition into year, quarter, month and day levels with measures aggregated to each level

Figure 5 is but one view or report on the data. If we looked at all levels of granularity for the Date category in the Sales example, there would be 2 years, 8 quarters, 24 months and 730 days. This represents 765 different aggregations for each measure that are automatically presented to the operational manager for analysis. If we look further at the combinations and permutations of these aggregations, it means that the operational manager has many millions of ways to report on timed performance results.

Figure 6 illustrates one such combination, a yearly quarter over quarter performance comparison. In this report, sales of freezers are shown for all stores, not just the Home Stores, trebling the number of possible displays. Imagine the combinations if we also relate the stores to the 6 cities in 2 states, and the 5 different payment methods. And that's just for freezer sales! Freezers are but one category in a Products dimension that has stoves, small appliances, and more. The number of combinations in which the operational managers can view their business data is almost endless.

The power of OLAP goes beyond the ability to interrelate particular data segments. OLAP reporting tools allow operational managers to choose alternate chart formats, sort, segment, nest and more. Figure 7 shows a bar chart visualization that might be used to accelerate the transfer of knowledge, particularly when the report is published to other operational managers.

OLAP reporting tools allow non-I.T. people such as operational managers to easily explore large quantities of data in perspectives that make sense to their decision needs. The intuitiveness of the interface combined with the breadth of capabilities is what gives operational managers the freedom they need to do their job effectively.

Again using Databeacon as an example of current OLAP technology, here is a summary of the capabilities that support the Web Reporting and Data Analysis needs of operational managers

Again using Databeacon Collaboration Edition as an example of current OLAP technology,

OLAP REPORTING TOOLS ALLOW NON-I.T. PEOPLE SUCH AS OPERATIONAL MANAGERS TO EASILY EXPLORE LARGE QUANTITIES OF DATA IN PERSPECTIVES THAT MAKE SENSE TO THEIR DECISION NEEDS

		Freezer Sales at All Stores			
Date	Quantity	2000	Q1 2000	2001	Q1 2001
All Sto	Quantity	139	15	45	32
Corner Mart	Sales	\$ 49,066.03	\$ 5,396.55	\$ 16,069.65	\$ 11,496.64
	Cost of Goods	\$ 31,892.85	\$ 3,507.75	\$ 10,445.25	\$ 7,472.80
	Profit	\$ 17,173.18	\$ 1,888.80	\$ 5,624.40	\$ 4,023.84
	Average Sell Price	\$ 352.99	\$ 359.77	\$ 357.10	\$ 359.27

Figure 6: OLAP tools provide time series analysis, allowing operational managers to study year over year performance results and make better decisions

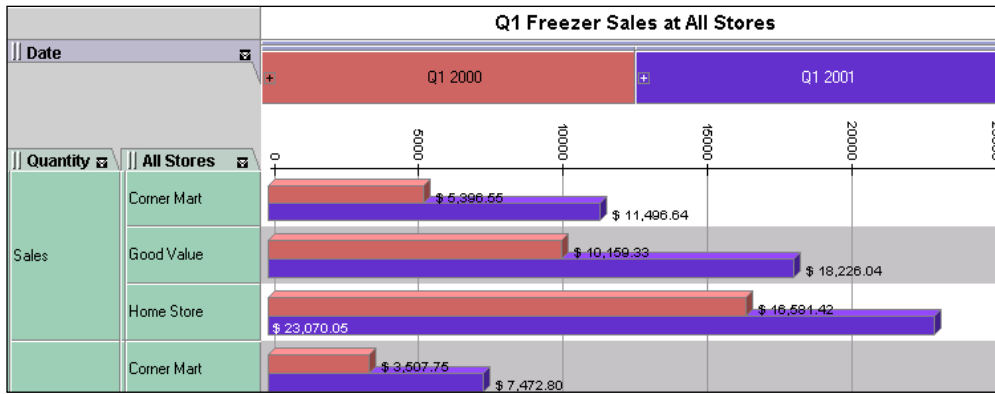


Figure 7: Charts and graphs allow operational managers to deploy data visualizations that accelerate understanding and better, quicker decisions

here is a summary of the capabilities that support the Web Reporting and Data Analysis needs of operational managers.

- * Access data from a number of sources
- * Look at data at different levels of granularity, such as Division / Department / Person or Year /Quarter / Month / Day
- * At each level of granularity, see the metric of performance shown as a value, count, total or average - all at once or one at a time.
- * Look at data through different visualizations, from a cross-tabular chart to pie, bar and line charts
- * View aggregated measures for dynamically chosen groups of categories, so as an example, the combined unit production for north-central plants in Detroit, Columbus and Chicago
- * See measures expressed as entered, or as a percentage of row, column or overall total
- * Filter the data that is presented, using individual and group selection, as well as unions and intersections
- * Sort, and limit the depth of data presented based on the sort.

OLAP GET INSIGHT OUT

There is an alternative to OLAP. It involves linking multiple reports as a way to deliver drill down, a key OLAP capability. However,

providing linked reports that can drill into one another assumes that a report designer can predict an operational manager's analysis process. It's not possible! OLAP functionality empowers operational managers to go wherever the data takes them.

At the core of OLAP reporting and analysis is this simple fact: you don't know what you don't know. If you did know, you could request just the relevant data and subsequently look at it using a parameterized report or pivot table viewer. Instead, you need the power of OLAP to deliver broad high level information so that you can drill down and through the data to discover what is important - that which you didn't know when you started.

As OLAP products such as Databeacon evolve, rich functionality is reaching more and more users, delivering here-to-fore unparalleled freedom and flexibility. Multidimensional Web Reporting and Data Analysis now services operational managers both inside and external to the organization. As Web-based explorable reports continue to gain greater popularity, multidimensional OLAP data delivery via no-touch smart client deployment will become commonplace.

AS OLAP PRODUCTS SUCH AS DATABEACON COLLABORATION EDITION EVOLVE, RICH FUNCTIONALITY IS REACHING MORE AND MORE USERS, DELIVERING HERE-TO-FORE UNPARALLELED FREEDOM AND FLEXIBILITY