SERVICEWARE

Why multi-dimensional databases are irreplaceable

White Paper by Christian Wagner and Harald Matzke

Table of Contents

1 Why multi-dimensional databases are irreplaceable

1.1   The triumphant rise of multi-dimensionality

During the 1990s a new database technology emerged and took the offices of the finance and controlling departments by storm. Hundreds of inconsistent and erroneous Excel and Lotus 1-2-3 spreadsheets were replaced by new, multi-dimensional – and thus user-orientated – database technologies. For the first time, the analytics and planning requirements of the business units could be met by a single central database with stunning response times and virtually unlimited modelling capabilities. The connection to Microsoft Excel which is part of many products has boosted the success of these technologies further. The term OLAP (Online Analytical Processing) is commonly used as a synonym for these databases.

1.2   The characteristics of OLAP systems

OLAP systems were originally used in finance departments and spread into other areas, such as sales management, HR controlling and production management.

Multi-dimensional structures are the heart of the OLAP philosophy. While financial data models are highly structured, they are also subject to constant change. Reporting structures, such as the row schema of a P&L statement or balance sheet, are regulated by US GAAP or IFRS. The group or cost center structure is based on the current, but continuously changing organizational structure of a company, quite similar to the structure and characteristics of products, markets, sales channels or business segments. Analyses and reports do not only compare actuals vs. budget but also different versions of forecasts or of medium-term plans. The way in which items are analyzed depends on the processes defined by the organization.

A dimension consists of elements that are related to others. The most commonly used relationship is the **hierarchy,** for example the breakdown of products into segments, families, groups and individual products. A dimension can be envisaged as a tree, with the individual products representing the leaves. Hierarchies may be "well-balanced", i.e. at every level there are elements in every branch of the hierarchy, or "unbalanced", i.e. in some branches of the tree there are more levels than in others.

The pioneers of OLAP technology provided their databases with a calculation engine right from the start. This means that the databases not only store data and can output it for analysis but that **calculations** can be performed automatically within the database. This feature has been a decisive factor in replacing spreadsheets. Simple calculations are aggregations along the hierarchy and include additions, but also subtractions and other basic mathematical operations. Typically, a product dimension is simply aggregated. However, when calculating a contribution margin the row schema has to include additions and subtractions. Multi-dimensional structures also require and enable calculations across dimensions throughout the emerging data space. Today, leading OLAP databases, such as IBM Planning Analytics (TM1), provide a large number of basic statistical, mathematical and business functions, that can typically be found in spreadsheets as well. Maintaining the data models is so easy that the analysts in the business units can do it themselves.

OLAP databases cover the entire controlling cycle including reporting, analysis, planning, forecasting, simulation and scenario calculation. Thus, end users need read and write access based on what is often a detailed authentication and authorization concept.

1.3   ROLAP, MOLAP and column-oriented databases

The first systems, which are still dominating the market today, belonged to the **MOLAP** category. MOLAP systems are products that store data in real multi-dimensional structures. From an IT perspective, these are proprietary structures. However, modern systems such as IBM Planning Analytics offer all necessary features to seamlessly integrate with existing IT operations.

The desire of IT to have business-critical data in known and familiar structures led to the development of **relational OLAP databases** (ROLAP). Well-known products include Microstrategy and SAP BW. The data is modelled using a star or snowflake schema. There usually is a central fact table that contains all low level data as well as the dimension tables. These are linked by keys and also contain the hierarchies. Since relational database technology has advanced dramatically, ROLAP databases have benefited a lot as well.

Ever since the marketing hype about SAP HANA, column-orientated databases emerged as another competitor to MOLAP and the row-oriented ROLAP. HANA also relies on tables as a technological basis but in contrast to classical relational databases, queries are executed differently. Users often work with a large table that is filtered by columns.

Many IT strategists and IT architects regard MOLAP therefore as an outdated technology, that can be discarded in the long run.

In all IT organizations, the knowledge about tabular database technologies – may they be row- or column-based – is widespread. For standardization purposes, a consolidation of all applications on such a technological platform is a common goal. Particularly as many believe that tabular technologies are able to support the needs of the business departments just as well as MOLAP systems.

But when focusing on the requirements of the business, this is – in our opinion – a gross misjudgement. Attempts to replace MOLAP systems usually fail.

The reason for this are the data structures and processes on which financial controlling is based. In the following paragraphs we would like to take a closer look at the requirements and the suitability of the technologies in regard of this particular question. We are only considering analytical and explicitly non-transactional issues. MOLAP systems are not suitable for the latter.

1.4   User requirements and degree of fulfilment provided by the alternatives

Planning, analysis and self-service reporting are nowadays used at all management levels. Therefore, ease of use, pre-structured data and response times in the range of seconds and tenths of a second are crucial. The importance of response times and predefined reports, usually in form of dashboards, increases with every step up the hierarchical ladder. In departments that prepare decisions, the focus is on ad hoc analysis and self-service reporting. However, both user populations need to be able to work with the same data, the same structures and the famous "single source of truth".

Today, simple, multi-dimensional structures can be easily implemented in row- or column-oriented database systems. Column-oriented systems tend to offer substantial performance advantages, since standard analytics often use filters on dimensions, that enable better response times in column-oriented technology.

It requires only a couple of unbalanced tree structures to push tabular technologies to their limits. Looking at a P&L structure in financial controlling it becomes obvious that it is highly unbalanced. Drilling-down from profit to

the individual cost types, 7-8 hierarchy levels are not uncommon, while the net interest income only requires 3-4 levels. This issue can be solved with the parent-child approach, but several unbalanced structures in the same query lead to an extremely high and exponentially increasing number of lookups in the databases. Even on the fastest hardware and when calculated in-memory, these databases are slower than comparable in-memory MOLAP systems. The more complex the hierarchical structures, the bigger the performance differences. If you have ever tried to create this structure using external hierarchies in SAP you know how quickly performance can go through the floor.
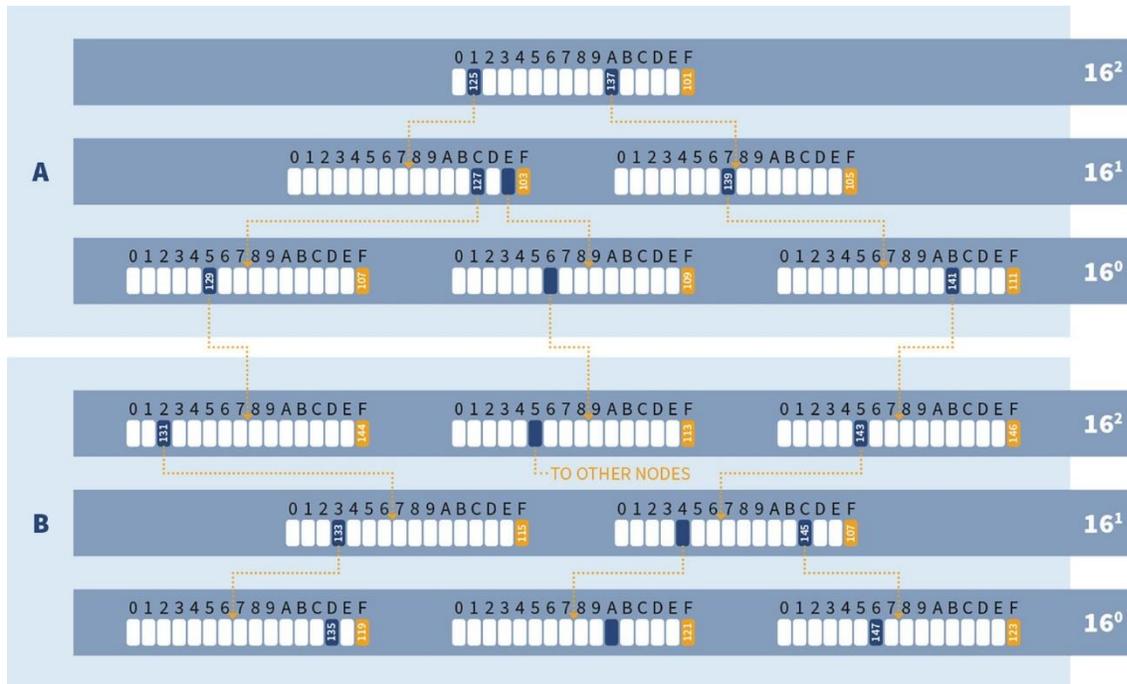
Figure: The multi-dimensional stacked tree structure used in efficient MOLAP systems as a basis for the efficient processing of hierarchical/multi-dimensional structures and calculations – excerpt from Patent Specification TM1.

Cross-dimensional calculations are another huge challenge for tabular technologies. In segment accounting, for example, overhead costs of logistics are allocated to the segments based on products and distribution channels. In the tabular model, an IT specialist needs to implement these requirements using a complex SQL process. The immense effort needed for implementation and maintenance can be expected to be the main issue here. And if everything is supposed to work in real-time to support a planning process, a severe performance and maintainability problem will undoubtably ensue.

In summary, the following data structures are common for financial and controlling processes:

Requirement	MOLAP	Row-oriented ROLAP	Column-oriented ROLAP
Multi-dimensional structures	X X X	X X	X X
Balanced tree structures	X X X	X X	X X
Unbalanced tree structures	X X X	X	X
Response times in small structures	X X X	X X	X X X
Response times in medium-sized structures	X X X	X	X X
Response times in very large structures	X X	X	X
Simple calculations	X X X	X X	X X
Complex calculations within a dimension	X X X	X	X
Cross-dimensional calculations	X X X	X	X
Authorization concepts for data content	X X X	X	X
Data input also at aggregated levels	X X X	X	X

2 Conclusion

ROLAP systems and especially column-oriented database systems like SAP HANA can easily meet the requirements in simple multidimensional data rooms. However, the real challenges, especially in financial controlling, can only be met by considerable use of computing power and additional coding efforts by SQL specialists but even then only to a limited extent.

MOLAP continues to close the technology gap intrinsic to tabular databases. While these execute transactional processes and filter-related questions with great efficiency, MOLAP databases are still in popular demand where analytics and a combination of dimensionality, hierarchies and calculations are concerned. MOLAP will remain part of the IT landscape and continue to grow to ensure that the requirements of controllers and management are met. The introduction of SAP HANA may shift the areas of application to some extent but without fundamentally changing the issue. Many companies are currently experiencing exactly this.

3 About the Authors

Christian Wagner is responsible for the TM1 ecosystem in the DACH region as well as initiator and organizer of the annual Planning Analytics TM1 User Day. He brings 19 years of professional experience in the implementation of OLAP systems for planning, analysis and reporting, 15 years of which he spent with TM1. Christian Wagner is currently a leading application specialist at IBM with a consulting focus on feasibility studies, building prototypical data models, presentations and training. He is also the central point of contact for product management in the USA. He studied industrial engineering and management at the TU Munich.

Harald Matzke studied business administration and started his professional career in the controlling department of Hewlett-Packard GmbH. In 1986 he was first assigned the task of setting up and implementing cost center planning on the basis of Lotus 1-2-3. In 1990, he moved internally to IT to set up the Executive Information Systems and Decision Support Systems divisions for Hewlett-Packard GmbH. In 1993 he founded the company cubus together with Thomas Volz. From the beginning, cubus focused on business planning and reporting systems. In addition to conceptual consulting and the implementation of performance management and business intelligence systems, the core elements of the product range are proprietary software systems based on established technologies from Oracle, IBM and Microsoft. Since June 2019 cubus belongs to Serviceware SE.

4 About Serviceware

Serviceware is one of the leading providers of software solutions for the digitization and automation of service processes, enabling companies to increase service quality and efficiently manage service costs. The unique, integrated, and modular Enterprise Service Management (ESM) platform consists of the proprietary software solutions Serviceware Processes (helpLine),Serviceware Financial (anafee), Serviceware Resources (Careware), Serviceware Knowledge (SABIO), and Serviceware Performance (cubus). Serviceware has more than 800 customers from various business sectors; among those are 9 DAX companies as well as 4 of the 7 biggest corporations in Germany. Serviceware's headquarter is located in Bad Camberg, Germany and the company has over 400 employees. For more information visit www.serviceware-se.com

Contact

cubus - A Serviceware Company

Bahnhofstraße 29

71083 Herrenberg

Phone +49 70 32 9451-0

Fax +49 70 32 9451-30

E-Mail: info@cubus.eu

Web: www.cubus.eu