

Warum multidimensionale Datenbanken unersetzlich sind

**Whitepaper von Christian Wagner und
Harald Matzke**

Inhaltsverzeichnis

1	Warum multidimensionale Datenbanken unersetzlich sind.....	4
1.1	Der Siegeszug der Multidimensionalität	4
1.2	Die Merkmale von OLAP-Systemen	4
1.3	ROLAP, MOLAP und spaltenorientierte Datenbanken	5
1.4	Anforderungen der Anwender und Erfüllungsgrad der Alternativen.....	6
2	Fazit	7
3	Über die Autoren	8
4	Über Serviceware	8

1 Warum multidimensionale Datenbanken unersetzlich sind

1.1 Der Siegeszug der Multidimensionalität

In den 90er Jahren startete eine neue Datenbanktechnologie einen Siegeszug durch die Büros der Controlling-Abteilungen. Hunderte von inkonsistenten und fehlerhaften Excel- und Lotus 1-2-3-Tabellen wurden durch neue, multidimensionale und damit anwenderorientierte Datenbanktechnologien ersetzt. Diese haben erstmals die Möglichkeit geboten, die Analyse- und Planungsanforderungen der Fachabteilungen in einer zentralen Datenbank mit außergewöhnlichen Antwortzeiten und fast grenzenlosen Modellierungsmöglichkeiten zu erfüllen. Die in vielen Produkten übliche Anbindung an Microsoft Excel haben den Erfolg dieser Technologien beschleunigt. Als Synonym für diese Datenbanken wird oftmals der Begriff OLAP (Online Analytical Processing) verwendet.

1.2 Die Merkmale von OLAP-Systemen

OLAP-Systeme haben ihren Ursprung in der Welt des Finanzcontrollings und haben sich von dort auch in andere Bereiche wie dem Vertriebscontrolling, HR-Controlling oder Produktionscontrolling weiterverbreitet.

Kern der OLAP-Philosophie sind die **mehrdimensionalen Strukturen**. Datenmodelle im Controlling sind hochgradig strukturiert, unterliegen dabei aber ständigen Veränderungen. Berichtsstrukturen wie das Zeilenschema einer G&V und Bilanz sind in den rechtlichen Vorschriften des HGB oder IRFS vorgegeben. Die Konzern- oder Kostenstellenstruktur ergibt sich aus der jeweils aktuellen, sich aber immer wieder ändernden Organisationsstruktur eines Unternehmens ähnlich wie Struktur und Ausprägungen der Produkte, Märkte, Vertriebskanäle oder Geschäftssegmente. Analysen und Berichte vergleichen nicht nur Ist- mit Budgetdaten, sondern auch verschiedene Forecast-Versionen oder der Mittelfristplanung. Diese Ausprägungen ergeben sich aus den definierten Prozessen in der jeweiligen Controlling-Organisation.

Eine Dimension besteht aus Elementen, die mit anderen in einer Beziehung stehen. Die am häufigsten benutzte Beziehung ist die **Hierarchie**, beispielsweise die Aufteilung der Produkte in Produktsegmente, Produktfamilien, Produktgruppen und Einzelprodukten. Eine Dimension kann man sich auch als Baum vorstellen, wobei die Einzelprodukte die Blätter darstellen. Hierarchien können „ausbalanciert“ sein, d.h. auf jeder Ebene gibt es in jedem Zweig der Hierarchie Elemente, oder „unbalanciert“, d.h. in manchen Zweigen des Baumes gibt es mehr Ebenen als in anderen.

Die Pioniere der OLAP-Technologie gaben ihrer Datenbank von Anfang an einen Rechenkern mit. Dies bedeutet, dass die Datenbanken nicht nur Daten speichern und für den Analysefall wieder ausgeben können, sondern auch in der Datenbank automatisiert **Berechnungen** durchgeführt werden können. Dieses Merkmal war ein wesentlicher Faktor um die Tabellenkalkulationsblätter ablösen zu können. Einfache Berechnungen sind Aggregationen entlang der Hierarchie, wobei nicht nur reine Additionen, sondern auch Subtraktionen und andere mathematische Grundoperationen möglich sind. Eine Produktdimension wird typischerweise einfach aggregiert. Im Zeilenschema einer Deckungsbeitragsrechnung sind dagegen Additionen und Subtraktionen nötig. Mehrdimensionale Strukturen erfordern und ermöglichen aber auch Berechnungen über Dimensionen hinweg, quer durch den sich aufspannenden Datenraum. Führende OLAP-Datenbanken wie IBM Planning Analytics (TM1) enthalten heute eine Vielzahl von statistischen, mathematischen und betriebswirtschaftlichen Basisfunktionen wie man sie typischerweise auch in den Tabellenkalkulationen findet. Die Pflege der Datenmodelle ist so einfach, dass sie von Fachleuten der Fachabteilung durchgeführt werden kann.

OLAP-Datenbanken decken den gesamten Controlling Zyklus von Reporting, Analyse, Planung, Forecasting, Simulation und Szenario Rechnung ab. Damit brauchen die Endanwender Lese- und Schreibrechte auf Basis eines oftmals detaillierten Zugriffsrechtekonzepts.

1.3 ROLAP, MOLAP und spaltenorientierte Datenbanken

Die ersten und heute noch marktprägenden Systeme gehören in die Kategorie der **MOLAP**-Systeme. MOLAP-Systeme sind Produkte, die ihre Daten in echten mehrdimensionalen Strukturen abspeichern. Aus IT-Sicht sind das proprietäre Strukturen. Moderne Systeme wie IBM Planning Analytics bieten aber alle wesentlichen Funktionen sich nahtlos in einen IT-Betrieb zu integrieren.

Der Wunsch der IT, unternehmenskritische Daten in bekannten und gewohnten Strukturen zu haben, führte zur Entwicklung der **relationalen OLAP**-Datenbanken (ROLAP). Bekannte Vertreter sind Microstrategy und SAP BW. Die Daten werden in einem Sternschema oder Schneeflockenschema modelliert. Dabei gibt es typischerweise eine zentrale Faktentabelle, die alle Daten auf unterster Ebene enthält, sowie die über Schlüssel verknüpften Dimensionstabellen, in denen dann auch die Hierarchien abgebildet werden. Die relationale Datenbanktechnologie hat gewaltige Technologiesprünge gemacht, die auch den ROLAP-Datenbanken sehr geholfen haben.

Spätestens mit dem Marketing-Hype um SAP HANA generieren die spaltenorientierten Datenbanken einen zusätzlichen Konkurrenten zu MOLAP und zeilenorientiertem ROLAP. Auch HANA arbeitet mit Tabellen als Technologiebasis. Aber die Abfragen laufen hier anders als in einer klassisch relationalen Datenbank. Man arbeitet oft mit einer großen Tabelle, in der auf Spalten gefiltert wird.

Bei vielen IT-Strategen und IT-Architekten wird MOLAP deshalb als veraltete Technologie gesehen, die perspektivisch ausgemustert werden kann.

In allen IT-Organisationen ist das Knowhow über tabellarische Datenbanktechnologien, ob zeilen- oder spaltenorientiert, weit verbreitet. Oft wird zur Standardisierung eine Konsolidierung aller Applikationen auf eine solche technologische Plattform angestrebt. Dabei besteht oft die Ansicht, dass tabellenbasierte Technologien die Anforderungen der Fachabteilungen genauso gut unterstützen können wie die MOLAP-Systeme.

Stellt man die Anforderungen der Fachabteilung in den Mittelpunkt, dann ist dies aus unserer Sicht eine grobe Fehleinschätzung. Entsprechende Ablöseversuche gehen in der Regel schief.

Die Ursache dafür liegt in den dem Controlling zugrundeliegenden Datenstrukturen und Prozessabläufen. Wir möchten im Folgenden auf die Anforderungen und die Eignung der Technologien für die spezifische Fragestellung näher eingehen. Wir betrachten nur analytische und explizit nicht transaktional orientierte Fragestellungen. Für solche sind MOLAP Systeme nicht geeignet.

1.4 Anforderungen der Anwender und Erfüllungsgrad der Alternativen

Planung, Analyse und Self-Service Reporting sind heute auf allen Managementebenen im Einsatz. Einfache Bedienung, vorstrukturierte Daten und Antwortzeiten im Bereich von Sekunden und Zehntelsekunden sind dafür zwingend nötig. Je höher der Managementlevel, desto wichtiger werden Antwortzeiten und vordefinierte Berichte, in der Regel sogenannte Dashboards. Auf der Ebene der Entscheidungsvorbereiter stehen die Möglichkeiten der Adhoc-Analyse und des Self-service Reporting im Vordergrund. Beide Benutzergruppen müssen aber auf der gleichen Datenbasis mit den gleichen Strukturen auf dem berühmten „single source of truth“ arbeiten.

Einfache mehrdimensionale Strukturen können mittlerweile problemlos in zeilen- oder spaltenorientierten Datenbanksystemen umgesetzt werden. Spaltenorientierte Systeme haben dabei häufig sehr große Performance-Vorteile, da Standardanalysen häufig Filter auf Dimensionen nutzen, die in der spaltenorientierten Technologie bessere Antwortzeiten ermöglicht.

Schon unbalancierte Baumstrukturen bringen die tabellarischen Technologien an ihre Grenzen. Betrachtet man im Finanzcontrolling eine G&V-Struktur, dann ist diese im höchsten Maße unbalanciert. Absteigend vom Gewinn bis zu den einzelnen Kostenarten finden wir sehr schnell 7-8 Hierarchiestufen, während das Zinsergebnis häufig mit 3-4 Stufen auskommt. Mit dem Parent-Child Ansatz können diese Anforderungen zwar behoben werden, aber mehrere unbalancierte Strukturen in einer Abfrage führen zu einer extrem hohen und exponentiell steigenden Anzahl an Lookups in den Datenbanken. Selbst auf der schnellsten Hardware und im Hauptspeicher gerechnet, werden diese Datenbanken langsamer als vergleichbare in-memory MOLAP Systeme. Je komplexer die Hierarchiestrukturen, desto größer werden die Performanceunterschiede. Jeder, der versucht hat, so etwas mit externen Hierarchien in SAP abzubilden, weiß wie schnell die Performance in die Knie geht.

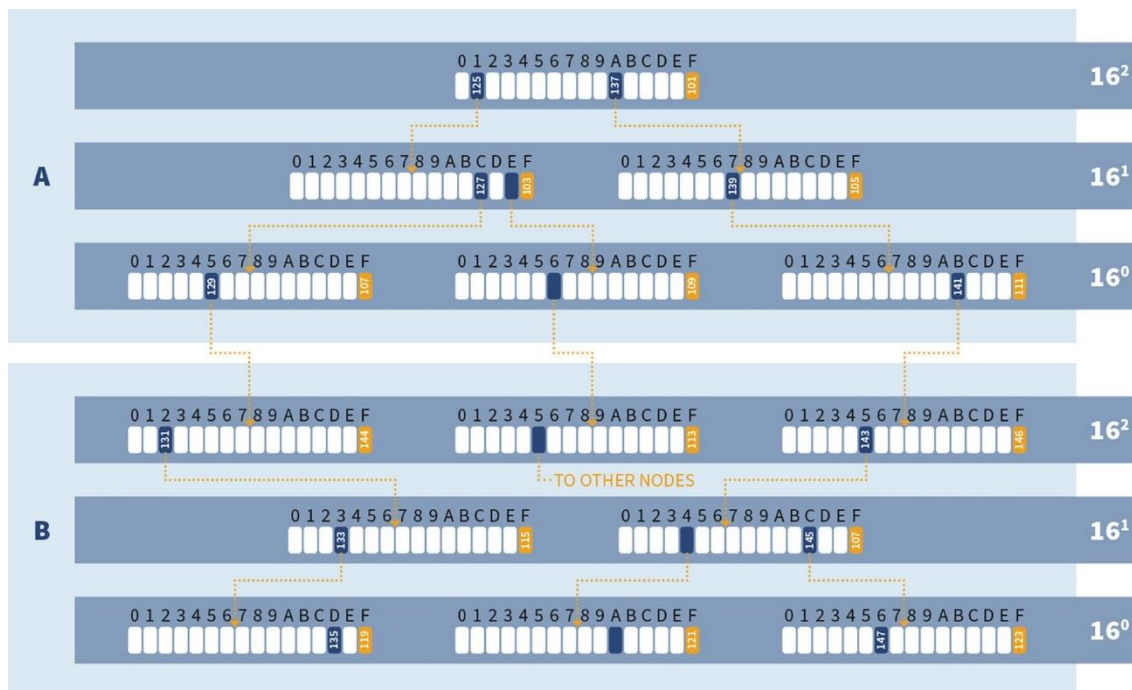


Abbildung: Die in effizienten MOLAP-Systemen verwendete mehrdimensional gestapelte Baumstruktur als Basis für die effiziente Verarbeitung hierarchisch- multidimensionaler Strukturen und Berechnungen – Auszug aus Patentschrift TM1

Dimensionsübergreifende Berechnungen sind eine weitere, sehr große Herausforderung für die tabellarischen Technologien. In einer Segmentrechnung werden beispielsweise Gemeinkosten der Logistik auf der Basis von Produkten und Vertriebskanälen auf die Segmente verteilt. Im tabellarischen Modell muss dabei ein IT-Spezialist diese Anforderungen durch einen komplexen SQL-Prozess umsetzen. Dabei ist voraussichtlich der immense Aufwand, um dies in einem tabellarischen System umzusetzen und zu pflegen, das Hauptproblem. Wenn zusätzlich die Anforderung besteht, dies Real-Time in einem Planungsprozess umzusetzen, dann kulminiert das in einem Performance- und Wartbarkeitsproblem.

Zusammenfassend ergibt sich folgendes Bild für die im Controlling anzutreffenden Datenstrukturen:

Anforderung	MOLAP	Zeilenorientiertes ROLAP	Spaltenorientiertes ROLAP
Abbildung von mehrdimensionalen Strukturen	X X X	X X	X X
Abbildung von ausbalancierten Bäumen	X X X	X X	X X
Abbildung von unbalancierten Bäumen	X X X	X	X
Antwortzeiten in kleinen Strukturen	X X X	X X	X X X
Antwortzeiten in mittleren Strukturen	X X X	X	X X
Antwortzeiten in sehr großen Strukturen	X X	X	X
einfache Berechnungen	X X X	X X	X X
komplexe Berechnungen innerhalb einer Dimension	X X X	X	X
dimensionsübergreifende Berechnungen	X X X	X	X
Berechtigungskonzepte auf Dateninhalte	X X X	X	X
Dateneingabe auch auf aggregierten Ebenen	X X X	X	X

2 Fazit

ROLAP-Systeme und insbesondere spaltenorientierte Datenbanksysteme wie SAP HANA können in einfachen mehrdimensionalen Datenräumen die Anforderungen problemlos erfüllen. Die wirklichen Herausforderungen, insbesondere im Finanzcontrolling, sind aber nur durch erheblichen Einsatz von Rechenleistung, gleichzeitigem Programmieraufwand von SQL-Spezialisten und auch dann nur eingeschränkt umsetzbar.

MOLAP schließt weiterhin die Technologielücke, die tabellarische Datenbanken aufweisen. Diese bilden die transaktionalen Prozesse und filterorientierte Fragestellungen sehr effizient ab. Im analytischen Umfeld und bei einer Kombination aus Dimensionalität, Hierarchien und Berechnungen sind weiterhin MOLAP Datenbanken gefragt. Sie werden Bestandteil der IT-Landkarte bleiben und weiterwachsen, wenn man die Abbildung der Anforderungen der Controller und des Managements sicherstellen möchte. Die Einführung von SAP HANA mag die Anwendungsbereiche etwas verschieben, ändert diesen Sachverhalt aber nicht grundsätzlich. Viele Unternehmen machen derzeit genau diese Erfahrung.

3 Über die Autoren

Christian Wagner ist verantwortlich für das TM1 Ökosystem in der DACH-Region sowie Initiator und Veranstalter des jährlichen Planning Analytics TM1 Anwendertags. Er bringt 19 Jahre Berufserfahrung in der Implementierung von OLAP-Systemen für Planung, Analyse und Reporting mit, davon 15 Jahre mit TM1. Christian Wagner ist derzeit leitender Anwendungsspezialist bei IBM mit den Beratungsschwerpunkten Machbarkeitsstudien, Bau prototypischer Datenmodelle, Präsentationen, Trainings. Außerdem ist er zentraler Kontaktpunkt zum Produktmanagement in den USA. Er studierte Wirtschaftsingenieurwesen an der TU München.

Harald Matzke hat Betriebswirtschaft studiert und zunächst im Controlling der Hewlett-Packard GmbH seine berufliche Laufbahn begonnen. Dort war er im Jahr 1986 erstmals mit dem Aufbau und der Durchführung der Kostenstellenplanung auf der Basis von Lotus 1-2-3 beauftragt. Im Jahr 1990 wechselte er intern in die IT, um die Bereiche Executive Informationssysteme und Decision Support Systeme für die Hewlett-Packard GmbH aufzubauen. Im Jahr 1993 gründete er mit Thomas Volz gemeinsam die Firma cubus. Von Anfang an lag der Schwerpunkt von cubus auf betriebswirtschaftlichen Planungs- und Berichtssystemen. Neben der konzeptionellen Beratung und der Implementierung von Performance Management und Business Intelligence Systemen sind eigene Softwaresysteme auf der Basis etablierter Technologien von Oracle, IBM und Microsoft die Kernelemente des Produktangebots. Seit Juni 2019 gehört cubus zu Serviceware SE.

4 Über Serviceware

Serviceware ist ein führender Anbieter von Softwarelösungen zur Digitalisierung und Automatisierung von Serviceprozessen, mit denen Unternehmen ihre Servicequalität steigern und ihre Servicekosten effizient managen können. Die einzigartige integrierte und modulare ESM-Plattform besteht aus den proprietären Softwarelösungen helpLine (Service Management), anafee (Financial Management), Careware (Field und Customer Service Management), SABIO (Knowledge-Management) und cubus (Corporate Performance Management / Business Intelligence).

Serviceware hat mehr als 800 Kunden aus verschiedensten Branchen, darunter 9 DAX-Unternehmen sowie 4 der 7 größten deutschen Unternehmen. Der Hauptsitz des Unternehmens befindet sich in Bad Camberg, Deutschland. Serviceware beschäftigt mehr als 400 Mitarbeiter.

Weitere Informationen erhalten Sie unter www.serviceware-se.com

Kontakt



cubus - A Serviceware Company

Bahnhofstraße 29

71083 Herrenberg

Tel. +49 70 32 9451-0

Fax +49 70 32 9451-30

E-Mail: info@cubus.eu

Web: www.cubus.eu