

# Data Warehousing als Herausforderung

## Mehrdimensionale Analysen machen Marktentwicklungen verständlich

**Applix**

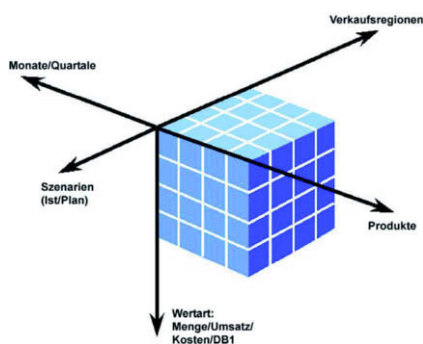


Rund 75 Prozent der Unternehmen räumen laut einer Studie von PricewaterhouseCoopers signifikante Probleme bei Datenstruktur und Inhalten ein. Weniger als 50 Prozent vertrauen ihren eigenen Daten.

Aufgrund der Heterogenität und Inkonsistenz lassen sich nur 20 Prozent aller verfügbaren Datenbestände zur Entscheidungsfindung heranziehen. Data-Warehousing bietet einen erfolgversprechenden Ansatz, diese Informationen in eine Zwischenschicht zu übertragen und zu vereinheitlichen, sodass die Daten anschließend für individuelle Analysen bereitstehen.

Die Anwendungssoftware der operativen Bereiche eines Unternehmens (Auftragsverwaltung, Finanzen, Warenwirtschaft, Distribution etc.) generiert nicht nur die zugehörigen Angaben über Kunden, Artikel oder Preise, sondern sie liefert automatisch auch das Ausgangsmaterial für strategische Entscheidungen. Die Datenbestände liegen allerdings häufig in unterschiedlichen Datenhaltungs-Systemen vor, sodass ein zeitnaher und gezielter Zugriff auf entscheidungsrelevante Angaben nicht möglich ist. Die Transaktionsebene gestattet außerdem nur eine beschränkte Sicht, da die Angaben in der Regel zeitlich begrenzt nur Wochen oder maximal wenige Monate zurückreichen. An diesem Punkt setzen Strategien für Data-Warehousing an. Die Herausforderung besteht darin, die Daten aus dem operativen Umfeld in einer Warehouse-Datenbank zusammenzuführen und für die Analyse zu vereinheitlichen.

Rund 95 Prozent der zentralen Warehouse-Daten sind in relationalen Datenbanksystemen (RDBMS) archiviert. Diese Systeme verwalten selbst Terabyte-große Datenmengen. Je umfangreicher die Berechnungen bei der Abfrage der archivierten Daten sind, desto länger warten Benutzer auf die Beantwortung ihrer Fragestellungen. Angaben über den Verkauf eines Produkts in einem Geschäftsfeld an einem bestimmten Tag lassen sich einfach darin wiederfinden. In der Regel reicht ein SQL-Statement, um diese Einträge herauszufiltern. Wenn aber die in einer Region erzielten Ergebnisse einer Produktgruppe für einen bestimmten Zeitraum abgefragt werden, sind komplexe Berechnungen notwendig, die relationalen Systemen Performance-Probleme bereiten.



Beispiel eines multidimensionalen Datenmodells mit 5 Dimensionen.

Die Reaktionszeit der Warehouse-Datenbanken wird gesenkt und nicht selten das gesamte System temporär lahmgelegt.

### OLAP-Datenbanken zur dynamischen Datenanalyse

Einen Lösungsansatz bieten zusätzliche Datenbanken für Online Analytical Processing (OLAP), die es mit zwei unterschiedlichen Ausprägungen für dynamische und statische Wertzuweisung gibt. Beim statischen Analysemodell werden alle theoretisch möglichen Werte vorausgerechnet und als sogenannte Hyper-Cubes auf der Festplatte hinterlegt. Dies kann bei komplexen Modellen Stunden oder sogar Tage dauern. Man findet diese Architekturen vor allem bei konstanten Analyse- und Reporting-Aufgaben sowie bei abfrageintensiven Anwendungen mit hohem Detaillierungsgrad. Auch dort, wo keine zeitnahen Auswertungen gefordert sind, kann der Einsatz sinnvoll sein. Bei der von Applix mit iTM1 realisierten dynamischen Variante - Multidimensional OLAP (MOLAP) oder Real-Time OLAP genannt - werden die Analyseergebnisse erst zum Bedarfszeitpunkt errechnet (Calculation-on-Demand). Will man zum Beispiel die Verkäufe einer Produktgruppe ermitteln, werden die Einzelwerte aller Produkte von der Software automatisch zusammengefaßt. Will ein Mitarbeiter die Entwicklung einer Region betrachten, werden die Daten der Geschäftsstellen konsolidiert. Sind aktuelle Quartalswerte gefordert, konsolidiert die Software zu diesem Zeitpunkt nur die Tages- und Monatsergebnisse.

### MOLAP für schnelle Kalkulations-Logik und -Methodik

Im Unterschied zur zweidimensionalen Tabellendarstellung relationaler Datenbanksysteme basieren multidimensionale OLAP-Datenbanken auf einer Repräsentation aggregierter Werte in Form beliebig vieler Dimensionen, die in Form eines Würfels angeordnet sind. Sie übernehmen nur aggregierte Werte der Warehouse-

Datenbank, da Anwender für Analyse- und Reporting-Aufgaben, beispielsweise interaktive „Was wäre wenn“-Szenarien, ohne die Ergebnisse der Einzeltransaktionen auskommen. Das Einrichten dieser Zwischenebene macht einen Großteil der Zugriffe auf die zentrale Warehouse-Datenbank überflüssig, da Anwender Analysen direkt auf der MOLAP-Datenbank ausführen. Die Leistungsfähigkeit zeigt sich besonders dann, wenn zum Beispiel die Ergebnisse von Produktgruppen pro Region zu ermitteln und die Angaben mit den Vorjahresdaten zu vergleichen sind. Die schnelle Kalkulations-Logik und -Methodik der Hauptspeicherorientierten Architektur von Applix iTM1 erlaubt Daten sofort hochzukonsolidieren, um zum Beispiel zu ermitteln, wie sich die Planungsergebnisse auf ein Cost-Center oder den Geschäftsfeldplan auswirken.

Theoretisch ist die multidimensionale Datenbank auf einen Umfang von zwei Milliarden Elementen pro Dimension limitiert; jeder Würfel enthält maximal bis zu 16 Dimensionen. Applix iTM1 arbeitet mit einer Multi-Cube-Architektur. Damit sind selbst komplexe Applikationen für das Risk-Management oder die Portfolio-Analyse in Banken oder für den Logistik-Bereich im Versandhandel umsetzbar, die aus dreißig und mehr Würfeln bestehen, die über Regeln miteinander verbunden sind. Eine Limitierung gibt es seitens des adressierbaren Hauptspeichers nur theoretisch. Dessen Wert liegt bei der heute üblichen 32-Bit-Architektur bei vier Gigabyte. Mit der 64-Bit-Version von iTM1 ist der adressierbare Speicher praktisch unbegrenzt.

Allerdings sind in der Praxis nahezu alle Würfel dünn besiedelt, da zum Beispiel nicht jedes Produkt, in jeder Region, von jedem Vertriebsmann, zu jedem Zeitpunkt verkauft wird. Die Kapazität der Würfel ist folglich nie voll ausgeschöpft. Da die Anzahl der theoretisch möglichen, gegenüber der Summe der tatsächlich belegten Zellen zwischen 0,3 und 5 Prozent liegt,



## Über Applix

Durch die Verknüpfung von CRM und Business Intelligence versorgt Applix die Geschäftsführung und die Fachabteilungen mit den e-Business-Lösungen und dem Know-how, die für eine erfolgreiche Unternehmensentwicklung notwendig sind.

Die CRM-Lösung Applix iEnterprise umfaßt Module für Vertrieb, Marketing, Kundenservice, HelpDesk, Call-Center sowie CRM-Analyse. Applix iTM1 dient zur Planung, Analyse und Budgetierung in allen Unternehmensbereichen.

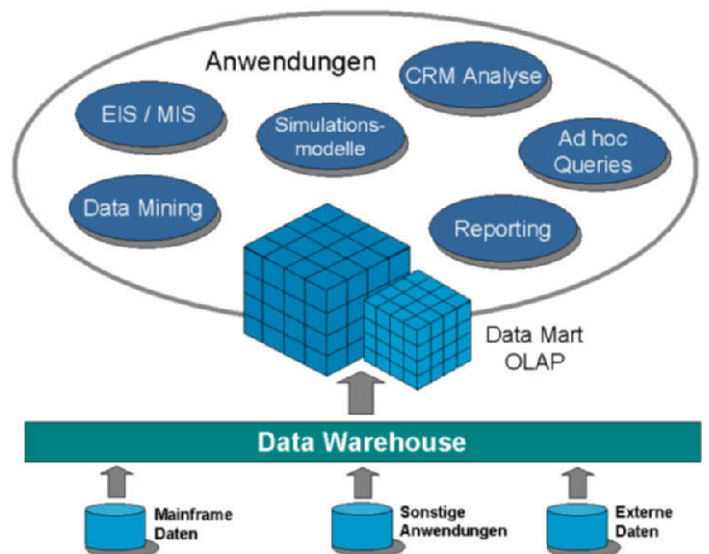
sind hierdurch selbst sehr komplexe Modelle weitgehend ohne Beschränkung umsetzbar.

Die Multi-Cube-Architektur von Applix iTM1 speichert nur die Basiselemente; die übrigen Angaben werden dynamisch generiert, sodass das auf der Festplatte hinterlegte Datenvolumen relativ klein bleibt. Bei einer Großbank, deren größter Würfel 13 Dimensionen umfaßt, dauert die längste Konsolidierungsphase rund fünf Minuten. Beim Einsatz einer statischen OLAP-Plattform hätte die Lösung über zwei Terabyte Speicherplatz belegt und die Kalkulation des Würfels würde rund 48 Stunden dauern. Ein weiterer Applix-Kunde, ein Telefondienstleister, betreibt über ein WAN (Wide Area Network) eine sehr große interaktive Planungsanwendung, die auf einem Server mit Applix iTM1 basiert. Bis zu 100 Anwender aus den verschiedenen Niederlassungen arbeiten parallel mit Informationen der Analyse-Datenbank. Zu den Anforderungen bei der Projektrealisierung zählten vor allem Interaktion und Planungsflexibilität, verbunden mit massiver Schreibfunktionalität für eine derart hohe Zahl gleichzeitiger Benutzer. Schreiben bedeutet in diesem Zusammenhang die Planung von „Was-wäre-wenn“-Szenarien, also die interaktive Eingabe von Daten in das MOLAP-System.

### OLE/DB für OLAP als Auswahlkriterium

Werkzeuge zur Entwicklung der Datenmodelle sind Bestandteile multidimensionaler Datenbanken, ebenso wie Migrations-Tools für die Übernahme der Daten aus der Warehouse-Datenbank bzw. aus den operativen Systemen. Der Datenzugriff erfolgt sowohl über standardisierte,

offene ODBC-, ASCII- oder API-Schnittstellen, als auch über proprietäre Interfaces älterer Host-Systeme. Die Unterstützung von OLE/DB für OLAP sollte bei der Suche nach einer OLAP-Engine ein auswahlentscheidendes Kriterium sein. Dieser offene Standard stellt sicher, dass zum Beispiel bei einer Firmenübernahme oder der Eingliederung eines Unternehmenszweigs die dort installierten Systeme ohne Programmieringriffe mit der MOLAP-Datenbank kommunizieren können.



Einordnung von Online Analytical Processing (OLAP) innerhalb eines Data Warehouse.

Während einer solchen Integrationsphase kommt es darauf an, möglichst schnell umfassende und genaue Planungsangaben zu bekommen, die Aufschluß über Produkte, Kunden und Vertriebsergebnisse geben.

Anwender aus den unterschiedlichen Funktions- und Geschäftsbereichen eines Unternehmens greifen mit Abfrage-Tools auf die im multidimensionalen Würfel gesammelten Informationen zu.



Die Anforderungen der verschiedenen Benutzergruppen differieren: Vorstände oder Geschäftsführer legen zum Beispiel Wert auf grafische Front-ends und vorgefertigte Reports mit einfachen Bedieneroberflächen, während Knowledge-Worker oder Controller in der Regel mit Spreadsheets als Front-ends ihre Daten analysieren.

Deshalb legen diese vor allem Wert auf eine enge Integration von multidimensionalen Datenbanken und Spreadsheet-Programmen. Sofern die MOLAP-Engine OLE/DB für OLAP unterstützt, kann der Anwender unter den marktführenden Front-end-Analysewerkzeugen auswählen, die sich nahtlos in die Anwendungsumgebung integrieren. Solche Analyse-Werkzeuge stammen unter anderem von Anbietern wie Arcplan, Bissantz & Company, Cognos, Gentia, Knosys, Hyperion, Microsoft oder Seagate.

Inter-Cube-Kalkulationsregeln erlauben den Anwendern, Daten zwischen Würfeln zu transferieren und ebenso Funktionalitäten weiterzugeben. Das erhöht den Grad der Wiederverwertbarkeit, vereinfacht die Pflege und macht Modelle überschaubar.

Ein Beispiel für die Mehrfachnutzung bietet der Aufbau eines Währungskalkulationswürfels als einheitliche Basis für die Konvertierung von D-Mark in US-Dollar, Yen oder Britische Pfund bei unterschiedlichen Planungs- und Auswertungsprozessen. Ein Unternehmen kann diesen übergeordneten Cross-Currency-Calculator vom Vertriebsanalyse-, Service- oder Kostenplanungs-Würfel aus referenzieren und mitbenutzen. Andere Funktionen erlauben darüber hinaus, Teile eines Würfels oder Modells auf einzelne Geschäftsstellen zu replizieren. Die Planung erfolgt mit dem jeweiligen Modell vor Ort. Über Nacht werden die verschiedenen Datenbestände synchronisiert.

Ebenso ist es möglich, Anwendungen auf ein Notebook zu replizieren, unterwegs zu benutzen und die Modelle nach Rückkehr mit dem zentralen Server zu synchronisieren.

Den Vorstellungen der Anwender, vorzeigbare Planungs- und Analysemodelle kurzfristig umzusetzen, kommt das Angebot vorkonfigurierter Lösungen entgegen.



Nutzung von OLAP-Funktionalität innerhalb des Customer Relationship Management ermöglicht effiziente Analyse von Kundendaten.



**Applix GmbH**  
 Boschetsrieder Str. 67  
 D-81379 München  
 Tel: +49 89 7485 890  
 Fax: +49 89 7485 8920  
 e-Mail: info@applix.de  
 www.applix.de

Applix (Schweiz) AG  
 Hagenholzstr. 81  
 CH-8050 Zürich  
 Tel: +41 43 299 6000  
 Fax: +41 43 299 6009  
 e-Mail: info@applix.ch  
 www.applix.ch

Beispiele hierfür sind Lösungen für den Vertriebsbereich beziehungsweise CRM (Customer Relationship Management) in Kombination mit einem MOLAP-Server. Die Daten aus dem CRM-System werden automatisiert in die multidimensionale Datenbank eingelesen und stehen dort zur freien Analyse zur Verfügung. Vordefinierte Reports beschleunigen den Aufbau eines Vertriebs-Informationssystems. Andere MOLAP-Lösungen adressieren Finanzdienstleister. Banken nutzen eine Vielzahl von Marktdaten, die beispielsweise von Anbietern wie Reuters oder Bloomberg stammen. Die Händler benutzen zwar unterschiedliche Informationsquellen, doch oft haben die Produkte ähnliche Inhalte, sodass Zusatzkosten durch Redundanzen anfallen.

In Großbanken liegen die jährlichen Ausgaben für Marktdaten nicht selten zwischen 10 und 50 Millionen Dollar.

#### **Aufspüren von Einsparungspotentialen**

Banken-Mitarbeiter transferieren die Marktdaten der verschiedenen Anbieter in ein Analyse-Modell, das ihnen einen Überblick darüber verschafft, welche Details in welchen Paketen enthalten sind und welche Bestandteile wie oft genutzt werden. Mit den Auswertungen ist es relativ einfach nachzuweisen, dass beispielsweise in den vergangenen Wochen oder Monaten ein bestimmtes Angebot kaum oder gar nicht abgerufen wurde.

Dies bedeutet erhebliches Einsparungspotential durch Abschalten des betreffenden Informationsdienstes. Vorgefertigte Berichte geben darüber hinaus Auskunft, ob das Angebot eines Informationsdienstleisters vergleichbar mit dem eines eventuell preisgünstigeren Mitbewerbers ist. Die analytische Datenbank übernimmt die Vergleichsangaben und identifiziert etwaige Einsparungspotentiale.

Andere Lösungsanbieter konzentrieren sich zum Beispiel auf vordefinierte Analysemodelle für den Versandhandel oder den Medienbereich.

Aus operativen Systemen kommende Daten weisen einen hohen Detaillierungsgrad auf. Sie werden im Data-Warehouse konsolidiert, methodisch vereinheitlicht und hinterlegt, um das Herausfiltern der erforderlichen Informationen zu erleichtern. OLAP-Werkzeuge bieten letztlich die Möglichkeit, in dieser Informations-Pyramide beliebig zu navigieren und auf die passenden Daten zu jeweiligen Bedarfszeitpunkt zuzugreifen. Flexibilität bei der Analyse und Modellierungsgeschwindigkeit sprechen für MOLAP- anstelle von ROLAP-Lösungen. Die MOLAP-Engine residiert auf einem separaten Server, wobei etwa 80 Prozent der Installationen unter NT laufen und die übrigen 20 Prozent unter Unix. Statt mit vorgefertigten Antwortschemata zu arbeiten, kann der Anwender seine Reports frei definieren. Fällt ihm beispielsweise auf, dass die Umsätze einer Produktgruppe in einem Quartal unter der Norm liegen, kann er sofort mit Drill-downs von Ebene zu Ebene tiefer in die Details gehen und sich die Daten der Einzelprodukte zeigen lassen. So könnte sich z.B. herausstellen, dass der Umsatzrückgang auf starke Preisreduktionen eines Produktes infolge eines Sonderangebotes zurückzuführen ist.

© Copyright 2001 Applix GmbH, München.  
 Alle genannten Marken und Produktnamen sind Warenzeichen oder eingetragene Warenzeichen der jeweiligen Markeninhaber.