



IBM DB2 Cube Views

Eine Einführung

By Gary Robinson
IBM Corporation

Inhaltsverzeichnis:

2	<i>Einführung in OLAP</i>
2	<i>Die Grundlagen</i>
4	<i>Physische Architektur</i>
5	<i>MOLAP oder ROLAP</i>
8	<i>IBM DB2 Cube Views</i>
8	<i>Einführung in Cube Views</i>
11	<i>DB2 Cube Views intern</i>
11	<i>OLAP-Metadaten der DB2-Datenbank bekannt machen</i>
21	<i>Aggregationen und Materialized Query Tables</i>
27	<i>Web Services für OLAP</i>
28	<i>Ausgabe von Analyseergebnissen mit DB2 Office Connect Analytic Edition</i>
29	<i>DB2 Cube Views Partner</i>
32	<i>Zusammenfassung</i>

Einführung in OLAP

Dieses Dokument gibt Ihnen eine Kurzeinführung zu OLAP sowie eine detailliertere Beschreibung von IBM DB2 Cube Views, einer neuen Erweiterung zu IBM DB2 Universal Database für die optimale Unterstützung von OLAP-Abfragen und -Anwendungen.

Die Grundlagen

Ein gut entworfenes Data Warehouse muss den Benutzern schnelle, leicht zu benutzende und effiziente analytische Fähigkeiten zur Beantwortung einer großen Anzahl geschäftlicher Fragestellungen bieten. Das System muss in der Lage sein, den analytischen Anforderungen der Benutzer gerecht zu werden, aber dennoch einfach zu bedienen sein. Es sollte eine multidimensionale konzeptionelle Sicht auf die Daten liefern. Zusätzlich müssen alle Daten und abgeleitete Informationen für alle Geschäftsanwendungen verfügbar sein. Die IBM DB2 Cube Views Technologie in Verbindung mit geeigneten OLAP- oder anderen Abfragewerkzeugen helfen den Benutzern, diese Anforderungen zu erfüllen.

OLAP steht für On-Line Analytical Processing, ein Begriff der einen multidimensionalen Ansatz zur Entscheidungsunterstützung beschreibt. Während relationale Datenbanken eine Sammlung von Daten in Tabellen und Feldern sind, ist eine OLAP-Datenbank eine Sammlung von verwandten Daten, die in Würfeln gespeichert sind, strukturiert durch Dimensionen, Hierarchien und Member.



On-Line analytische Verarbeitung ermöglicht eine hoch komplizierte, multidimensionale Analyse von Daten, die für die Entscheidungsfindung benutzt werden können. Obwohl die Abgrenzung von OLAP zu anderen Entscheidungshilfen etwas vage ist, müssen OLAP-Produkte mindestens die folgenden minimalen Funktionen besitzen:

- *Roll-up*. Die Roll-up Funktion verbirgt die Dimensionsrangordnung entlang einer oder mehrerer Dimensionen, um die verbleibende(n) Dimension(en) in einer größeren Granularität anzuzeigen.
- *Drill-down*. In Kontrast dazu erlaubt die Drill-down Funktion dem Benutzer eine detailliertere Ansicht einer gegebenen Dimension.
- *Slice*. Hierbei wird eine Scheibe aus dem ursprünglichen Würfel entsprechend einem einzelnen Wert einer gegebenen Dimension angezeigt. Es wird keine Aggregation vorgenommen. Statt dessen kann sich der Benutzer auf die Einzelwerte fokussieren.
- *Dice*. Eine verwandte Operation ist das Dice. In diesem Fall wird ein Teilwürfel vom ursprünglichen Ergebnis definiert. In anderen Worten, durch das Spezifizieren von Wertbereichen von einer oder mehreren Dimensionen, kann der Benutzer ausgewählte Blöcke von summierten Daten anzeigen.
- *Pivot*. Pivot ist eine einfache aber effektive Operation, die dem OLAP-Benutzer erlaubt, Würfelwerte in einer natürlichen und intuitiven Form darzustellen.

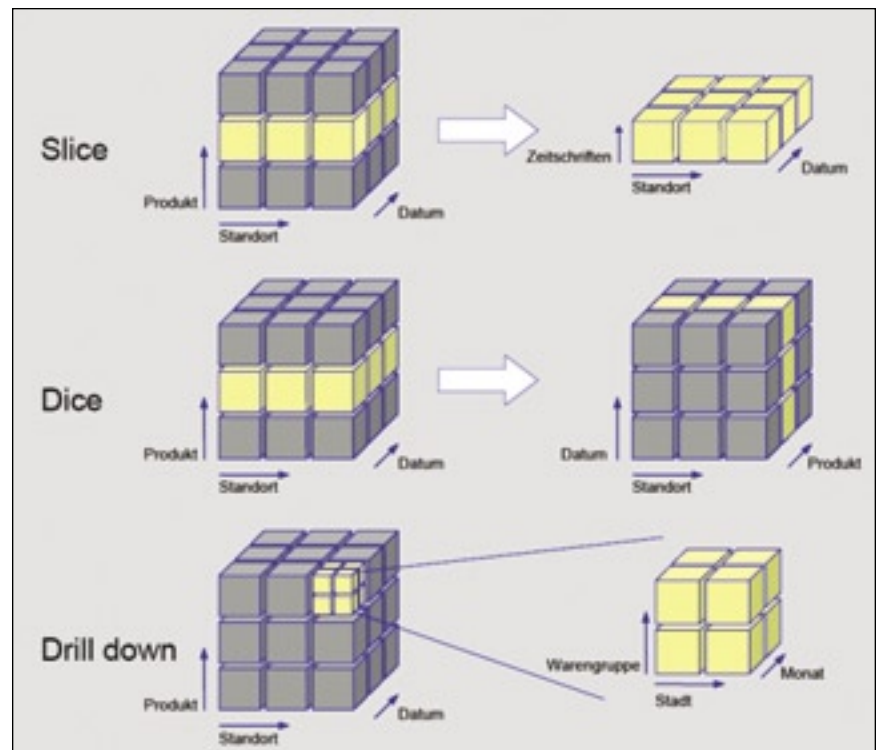


Abbildung 1: Typische OLAP-Operationen

Physische Architektur

Während OLAP-Produkte sehr einfach sein können, wie z.B. die Erweiterung einer Tabellenkalkulation, stehen größere Systeme fast ausnahmslos in Verbindung mit einem zugrundeliegenden Data Warehouse. Ein Data Warehouse ist eher über Themen als über atomare Transaktionen organisiert. Es repräsentiert aggregierte Informationen aus einer Vielfalt von Quellen. Es stellt eine Datensammlung über einen längeren Zeitraum hinweg dar, typischerweise Jahre, und ist normalerweise abgestimmt für Lesezugriffe.

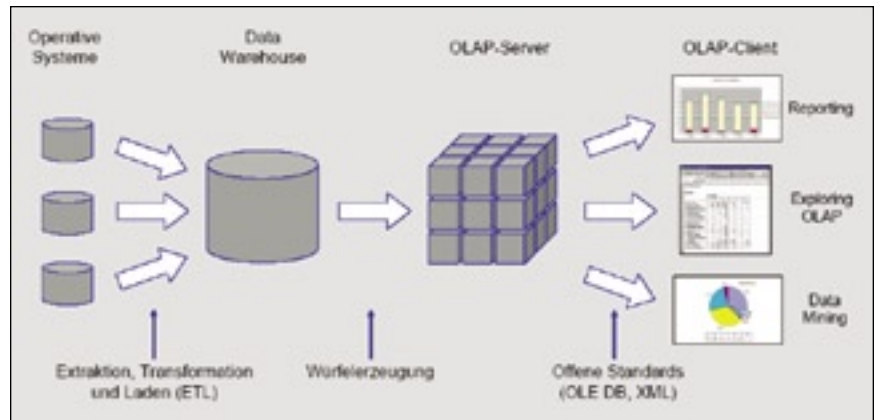


Abbildung 2: Dreischichtige Architektur von Data Warehouses

Im Allgemeinen können Data Warehouses als ein dreischichtiges Datenmodell gesehen werden. Wie in Abbildung 2 oben dargestellt, stellen die operativen Quellsysteme die erste Schicht dar. Aus ihnen werden die Informationen zuerst extrahiert, dann bereinigt, umgewandelt und in das Warehouse, welches die zweite Schicht darstellt, geladen. Dieser erste Schritt ist eine wichtige Aktivität, die bei den Anbietern von OLAP-Werkzeugen sehr stark in den Vordergrund gerückt ist. Oft befinden sich die Produktionsdaten in einer Ansammlung von verteilten, heterogenen Quellen und müssen sehr stark angepasst und bereinigt werden, bevor sie in das Warehouse integriert werden können.

MOLAP oder ROLAP

Eines der wichtigsten Elemente von OLAP-Umgebungen ist das Vertrauen in die multidimensionalen Datenwerte. Wie vorher schon erwähnt, repräsentieren Data Warehouses eher themenorientierte als transaktionsorientierte Sätze. Deshalb können aggregierte Werte innerhalb eines logischen Würfels als existent betrachtet werden, wobei der Benutzer den Würfel auf einer oder mehreren dimensional Ebenen indexieren kann. In der Nomenklatur von OLAP verursacht diese Art von konzeptioneller Repräsentation der Daten einen sogenannten *Datenwürfel*. Auf diesem würfelähnlichen Modell liegt der Fokus der dritten Schicht der Data Warehouse-Architektur, der Data Mart-Ebene.



Da Datenwürfel eine multidimensionale Interpretation der Daten vorschlagen, hat eine Anzahl von OLAP-Anbietern sich entschlossen, den Würfel als ein multidimensionales Feld zu modellieren. Diese MOLAP-Produkte (Multidimensionales OLAP) bieten schnelle Antwortzeiten auf alle OLAP-Anfragen, da direkt auf die Datenwürfel-Struktur zugegriffen wird, um Teilmengen von aggregierten Daten abzufragen. Leider skalieren MOLAP-Lösungen bei großen, vieldimensionalen Datenmodellen nicht immer effektiv (obwohl der MOLAP-Durchsatz in den letzten Jahren bedeutsam zugenommen hat). Das Problem ist, dass bei einer wachsenden Anzahl von Dimensionen die Daten in den Datenwürfeln zunehmend „*sparse*“ (dünn besiedelt) werden. Mit anderen Worten, viele der durch die Datenwürfel-Struktur vertretenen Attribut-Kombinationen enthalten keine aggregierten Daten. Daher kann eine so zustande gekommene MOLAP-Struktur eine enorme Anzahl von leeren Zellen enthalten, die einen unannehmbar grossen Speicherplatzbedarf bedingen. Kompressionstechniken können angewandt werden, um diese Probleme zu verringern, obwohl so teilweise der natürliche Index vernichtet wird, der MOLAP so attraktiv und performant macht.

Hochperformante OLAP-Technologien haben einige Alleinstellungsmerkmale, die aber für die Dimensionsanalyse nicht erforderlich sind. Gleiche Merkmale sind auch in relationalen Datenbanken vorhanden, die mit dem Design und der Technik erstellt wurden, welche Ralph Kimball in seiner wegweisenden Arbeit „The Data Warehouse Toolkit“ beschrieben hat. Doktor Kimball definiert Dimensionale- oder Sternschemata mit relationaler Datenbank-Technologie. Die Daten sind gegenüber der dritten Normalform aus ER-Modellen denormalisiert, um multidimensionale Geschäftsmodelle in den Daten zu reflektieren. Anstatt eines multidimensionalen Arrays, wird der Datenwürfel als eine Sammlung von bis zu 2^d relationalen Tabellen realisiert (wobei d die Anzahl der Dimensionen ist), von der jede eine eigenständige Ansicht beschreibt. Diese Art der multidimensionalen Datenanalyse bezeichnet man als ROLAP (Relationales OLAP). Da diese Ansichten konventionelle Datenbanktabellen sind, können sie mit traditionellen RDBMS-Techniken verarbeitet und



abgefragt werden (z.B. Indizes und Joins). Außerdem können sie in großen Data Warehouses effizienter sein, da nur jene Würfelzellen in den Tabellen gespeichert werden, die tatsächlich Daten enthalten. In seiner einfachsten und reinsten Form werden Maßeinheiten in einer Fakttable gespeichert und diese Fakttable wird über Hilfsschlüssel mit den Dimensionstabellen verbunden. Dieses Schema stellt sich normalerweise als eine grosse zentrale Fakttable umgeben von einer Anzahl kleinerer Dimensionstabellen dar. Dies kann man sich wie einen Stern (Star) vorstellen. Eine Variante des Star-Schemas ist das Snowflake-Schema (Schneeflocken-Schema), in welchem einige der Dimensionstabellen normalisiert wurden. Dies kann zu einem Schema führen, das wie eine Schneeflocke aussieht.

Welche Schema-Variante verwendet wird ist weniger wichtig als die Aufbereitung der Daten für den Benutzer, so dass dieser sie leicht und interaktiv für verschiedene geschäftliche Fragestellungen in verschiedenen Aggregationsstufen und aus verschiedenen Blickwinkeln analysieren kann. MOLAP und ROLAP kann man sozusagen als verwandt betrachten, und beide Techniken und Technologien werden häufig zusammen in einer Architektur eingesetzt, die auch als hybrides OLAP (HOLAP) bezeichnet wird. Diese Architektur erlaubt es den Benutzern, sowohl von der hervorragenden Verarbeitungsgeschwindigkeit von MOLAP als auch von der Skalierbarkeit relationaler Datenbanken zu profitieren. ROLAP (dimensionale Modellierung) und MOLAP sind die Grundsteine von Data Warehouse-Lösungen, denn sie bieten den Benutzern Freiheitsgrade in der Analyse von Geschäftsdaten, wie sie traditionelle Abfrage- und Berichts-Lösungen nicht bieten können. Ohne OLAP-Technologien in der einen oder anderen Form verschwenden Analysten oft eine Menge Zeit und Energie beim Zusammentragen der für eine geschäftliche Fragestellung relevanten Informationen aus einer Anzahl von verschiedenen Berichten und Systemen.



IBM DB2 Cube Views

Einführung in Cube Views

DB2 Cube Views ist ein OLAP-Beschleuniger für DB2. Die Entwicklung von DB2 Cube Views begann mit einer Serie von Strategiediskussionen in welchen wir den OLAP-Markt und unsere eigenen OLAP-Technologien betrachtet haben, um zu entscheiden, welche Anforderungen unsere Kunden an DB2 stellen, damit sie für OLAP-Lösungen eingesetzt werden kann. Unsere Analyse führte zu der Schlussfolgerung, dass die steigende Nachfrage nach OLAP-Lösungen sich fortsetzen wird und die OLAP-Lösungen selbst in Volumen und Umfang wachsen werden, während die Versorgung mit Informationen schneller als je zuvor geschehen muss. Wir erwarten einen wachsenden Markt für hoch skalierbare und sehr umfangreiche mehrdimensionale Modelle, die es Benutzern erlauben, ihre Geschäftsdatenmodelle ad hoc zu analysieren ohne an technische Beschränkungen zu stossen. Limitationen der Würfeltechnologie schränken derzeit den Bereich und die Tiefe von den in OLAP-Würfeln verfügbaren Informationen ein. Während das Datenvolumen wächst und der zunehmende Bedarf nach Echtzeitinformationen der Würfeltechnologie Grenzen auflegt, stellen die Benutzer immer anspruchsvollere Anforderungen an diese Systeme. Diese Anforderungen haben die OLAP-Anbieter schnell dazu gebracht, auf eine hybride Architektur zu konvergieren, die es OLAP-Designern ermöglicht, physische Würfel mit einer Teilmenge der Geschäftsdaten zu beladen, den Benutzern aber erlauben, transparent zusätzliche Dimensionen oder Stufen von Detaildaten zu analysieren, die in einem darunterliegenden Data Warehouse oder Data Mart gespeichert sind. Die OLAP-Lösung fasst dann beide Ansätze zusammen, den mehrdimensional abgespeicherten Würfel für hohe Verarbeitungs-geschwindigkeit und die relationale Datenbank mit grosser Skalierbarkeit und flexiblen Analysemöglichkeiten.



Die wachsende Nachfrage nach OLAP-Lösungen kann auf mehrere Faktoren zurückgeführt werden. OLAP wurde zuerst im Finanzbereich eingesetzt, wo Finanzanalytiker schnell den Vorteil sahen, statt einer Menge von Tabellenkalkulationen eine Technologie zu haben, die entworfen wurde, um ihren Anforderungen an eine schnellere Datenanalyse ihrer Finanzmodelle zu entsprechen. Viele OLAP-Verkäufe wurden abgeschlossen basierend auf der benötigten Zeit, die hochbezahlte Finanzanalytiker durch den Einsatz eines OLAP-Servers einsparen konnten, statt sie mit der Eingabe und Manipulation komplexer Tabellenkalkulationen zu verbringen. IT-Abteilungen, andere Benutzer und Führungskräfte sahen schnell Möglichkeiten, um OLAP-Technologie in vielen anderen geschäftlichen Bereichen anzuwenden, wo Analyse und „what if“ Modelle verlangt wurden. Als die OLAP-Technologie begann, sich vom Finanzbereich auf Vertrieb, Marketing und andere Geschäftsbereiche auszubreiten, forcierten die Benutzer diese Methode mit komplizierteren Geschäftsmodellen, tieferer Analyse und einer großen Auswahl von Schnittstellen. So kompliziert diese Anwendungen auch waren, es handelte sich typischerweise um Summeninformationen über eine beschränkte Anzahl von Dimensionen. Als diese Technologie sich vom Handel und der Industrie in andere Branchen, wie zum Beispiel Telekommunikationsunternehmen, hinein verbreitete, stieg die Nachfrage nach zunehmend ergiebigeren, flexibleren und skalierbaren Modellen an. In diesen Branchen stecken die wertvollsten Informationen oft in den Einzeldaten mit einer großen Anzahl von Attributen und sehr dynamischen Beziehungen. Das Laden von OLAP-Würfeln mit einer vordefinierten Untermenge von Attributen über relativ statische Dimensionen in Aggregationsstufen wertete diese Technik in vielen von diesen Fällen ab. Benutzer konnten keinen Kompromiss zwischen den Möglichkeiten und Grenzen von OLAP-Würfeln finden. Das heutige geschäftliche Umfeld macht diesen Kompromiss sogar noch schwieriger. Analytiker und Führungskräfte müssen ihr Geschäft zu



einem höheren Grad feinabstimmen als je zuvor, um schneller reagieren zu können und damit eine höhere Rentabilität und Wettbewerbsvorteile zu erzielen. Um den Wert jeder Transaktion und jeder Kundeninteraktion zu maximieren, muss das Auswertungswerkzeug diese Informationen benutzen, verstehen und darauf reagieren. Änderungen in der Gesetzgebung, wie zum Beispiel Sarbanes-Oxley, erhöhen den Druck auf Unternehmen noch weiter, denn finanzielle Transparenz und Unternehmensverantwortlichkeit verlangen ein sehr viel tieferes Verständnis von den internen und externen Geschäftstätigkeiten als je zuvor. Die Kombination der weit verbreiteten Adaption von OLAP mit der heutigen Volkswirtschaft, Unternehmen und gesetzgeberischen Anforderungen bringen jede OLAP-Anwendung bis an den Punkt, an dem Würfel nicht schnell genug mit ausreichenden Informationen auf der notwendigen Detailstufe geladen werden können, die erforderlich sind um konkurrenzfähig zu sein. So schnell auch Moores Gesetz die Hardware vorantreibt, reicht das nicht aus, um mit diesen Anforderungen Schritt zu halten.

Die Antwort der OLAP-Anbieter auf diese Herausforderung ist, den Umfang von OLAP-Lösungen mit einer hybriden Architektur zu erweitern. Eine schnelle Übersicht über den OLAP-Markt wird schnell bestätigen, dass die führenden OLAP-Anbieter bereits hybride Lösungen anbieten. Wenn diese Architektur auch das Versprechen hält, das Beste von Hochgeschwindigkeitswürfel-Strukturen und skalierbarer relationaler Speicherung zu sein, bringt sie auch neue Herausforderungen mit sich. Die OLAP-Technologie muss nahtlos die Lücke von Würfeln zu relationalen Tabellen überbrücken, und die IT Abteilung muss dann optimierte Lösungen auf Basis beider Technologien entwickeln und einsetzen. Die Entwicklung von hybridem OLAP wurde bis jetzt von den OLAP-Anbietern mit wenig Unterstützung der Anbieter relationaler Datenbanken vorangetrieben. Mit der Einführung von DB2 Cube Views ist DB2 die erste relationale Datenbank, welche die Wichtigkeit von OLAP anerkennt und insbesondere der Notwendigkeit zur Unterstützung und Optimierung für OLAP in der relationalen Datenbank Rechnung trägt und diese mit anderen OLAP-Technologien integriert.

**DB2 Cube Views intern**

Über eine Vielzahl von Versionen hinweg hat DB2 eine Reihe von Eigenschaften und Funktionen zur Verfügung gestellt, die Business Intelligence (BI) und OLAP unterstützen. Dazu gehören unter anderem: analytische Funktionen, wie zum Beispiel „ranking“ und „correlation“, Verwaltung von aggregierten Daten durch Materialized Query Tables (MQTs), hervorragende Gruppierungsfunktionen (grouping sets), „roll up“ und „cube by“ sowie die „Multidimensional Clustering“ (MDC) Index-Technologie. Dies gab IBM eine sehr starke Basis für die Entwicklung von DB2 Cube Views. IBM hat drei Schlüsselbereiche identifiziert, die benötigt werden, um die zusätzlichen Werkzeuge und Technologien zu entwickeln, die in DB2 Cube Views enthalten sind. Diese sind:

- OLAP Modelle in der DB2-Datenbank abbilden und bekannt machen,
- Entwicklung eines „Advisors“ zur Erstellung von Aggregationstabellen
- Auslieferung einer modernen OLAP-Schnittstelle für Anwendungsentwickler.

Dies sind die wichtigsten Komponenten von DB2 Cube Views Version 8.1, das jetzt als Zusatzprodukt zu DB2 Universal Database verfügbar ist.

OLAP-Metadaten der DB2-Datenbank bekannt machen

DB2 speichert seit langer Zeit multidimensionale Datenmodelle. Jede Version von DB2 lieferte eine bessere Unterstützung für multidimensionale (Star und Snowflake) Schemata und Analysefunktionen, aber diese waren tief in der DB2-Engine eingebettet. Wenn man die DB2-Katalogtabellen betrachtet oder DB2-Daten abfragt, findet sich keine Spur von Dimensionalität, Rangordnungen, Kennzahlen oder Attributen. Tatsächlich ist es sehr schwierig herauszufinden, ob ein Schema überhaupt multidimensional ist, es sei denn, Sie sind in der Lage, das unterliegende Schema zu untersuchen und seine Beziehungen sichtbar zu machen.

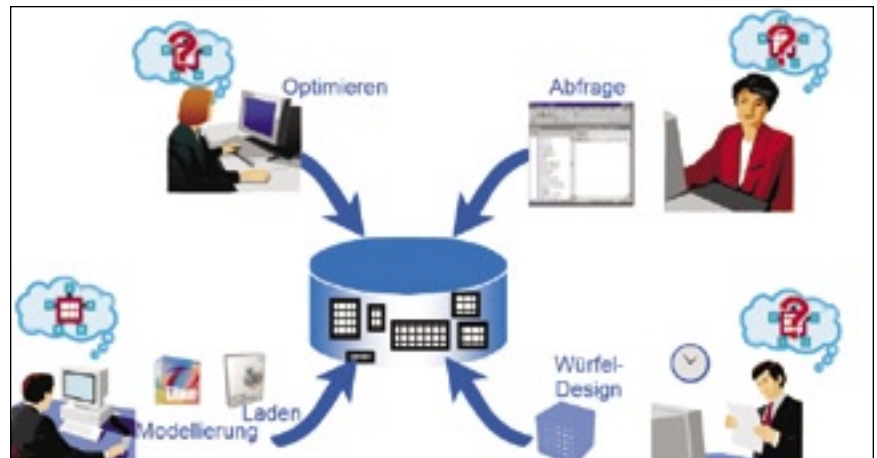


Abbildung 3: Data Warehouse Umgebung

In einer typischen, oder vielleicht idealisierten Umgebung wie sie in Abbildung 3 gezeigt wird, beginnt ein Designer, ein multidimensionales Schema zu definieren, welches den Bedürfnissen einer Gruppe von Endbenutzern angepasst ist. Er benutzt ein Modellierungswerkzeug zur Definition des Schemas, zum Markieren von Dimensionen, Fakten, Attributen und anderen Aspekten des Modells. Durch Betätigung der Entertaste wird das Schema automatisch in DB2 angelegt. Um dieses Schema mit Daten zu füllen, benutzt unser Designer ein ETL-Werkzeug. Dieses Werkzeug sollte auch multidimensionale Modelle bearbeiten können, das heisst, es sollte die Tabellen und Spalten verstehen, welche die Dimensionen und Fakten beschreiben, sowie korrekte Schlüssel erzeugen und Änderungen in den Dimensionen handhaben können. Wenn unser Designer die ETL-Prozesse ausführt, werden Daten in DB2 übertragen und das Schema ist abfragebereit. Leider gehen beim Anliefern des Schemas eine Menge an Informationen verloren. Die Modellierung und das ETL-Werkzeug waren nicht in der Lage, ihre Kenntnisse der multidimensionalen Struktur des Schemas nach DB2 zu transferieren. Dies führt dazu, dass die Abfragewerkzeuge der Endanwender keine ausreichenden Informationen von DB2 bekommen, um den Benutzern das multidimensionale Modell präsentieren zu können, das

der Designer angelegt hat. Die einzigen in DB2 verfügbaren Informationen sind die über die in DB2 gespeicherten Tabellen und Spalten. Es gibt keine Informationen über die höheren Stufen der multidimensionalen Organisation, die der Designer so sorgfältig erstellt hat. Um dieses Defizit auszugleichen, zwingt die Mehrheit der Abfragewerkzeuge die Benutzer dazu, zuerst alle multidimensionalen Abhängigkeiten wieder herauszufinden. Diese Informationen werden dann erneut eingegeben und als werkzeugspezifische Metadaten gespeichert. Die Endanwender können jetzt mit Ihrer Arbeit fortfahren. Aber was geschieht, wenn Sie sich für den Einsatz eines anderen Werkzeugs entscheiden, um einen OLAP-Würfel aus dem multidimensionalen Data Mart zu bauen, oder nur den Datenbank-Administrator (DBA) bitten, das Schema zu optimieren? Die Antwort ist, der gesamte Prozess fängt von vorne an, und eine weitere Insellösung von multidimensionalen Metadaten wird befüllt, wie in Abbildung 4 dargestellt. Dieser Prozess ist fehleranfällig, zeitaufwendig und unnötig.

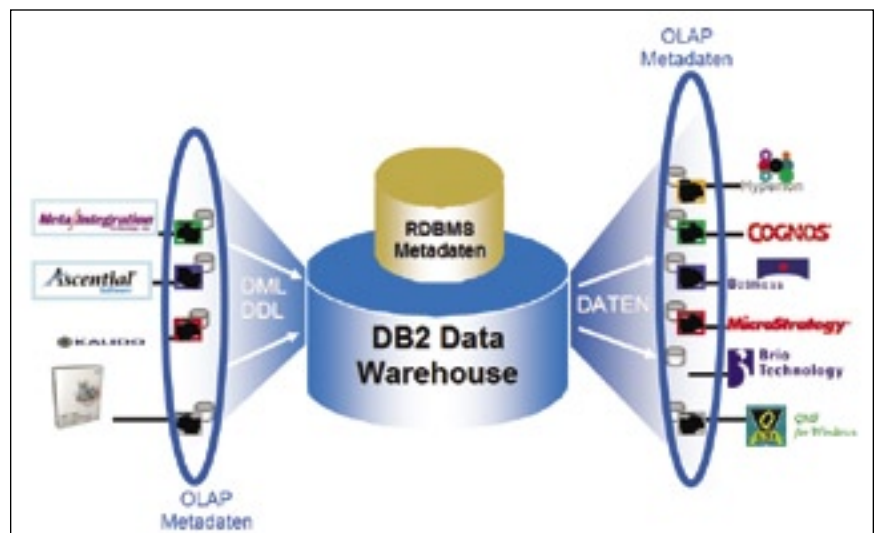


Abbildung 4: Austausch von OLAP-Metadaten im Data Warehouse

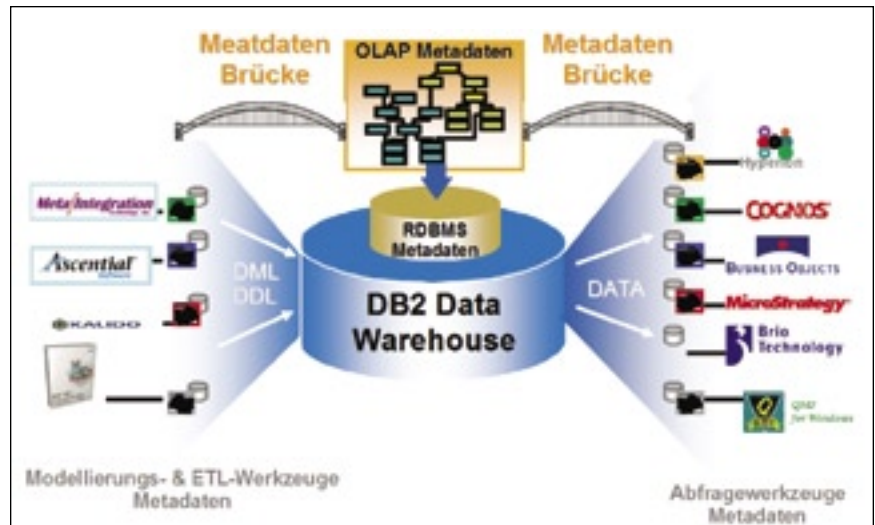


Abbildung 5: DB2 im Zentrum der OLAP-Metadaten

Hier kommt DB2 Cube Views ins Spiel. Wenn wir sagen, wir machen OLAP der DB2-Datenbank bekannt, denken wir zuerst an eine Reihe von Erweiterungen in den DB2-Katalogen die Metadaten über die multidimensionalen Abhängigkeiten speichern, die das Schema bilden. Mit DB2 Cube Views können Sie jetzt in DB2 abfragen, welche Informationen es über ein multidimensionales Schema gibt, und bekommen eine vollständige Beschreibung von Dimensionen, Attributen, Kennzahlen, Rangordnungen und mehr zurück. Als wir DB2 Cube Views entworfen haben, fanden wir heraus, dass es zwei Wege gab: Entweder konnten wir eine weitere Insellösung für Metadaten schaffen, oder wir konnten Vorteile aus der Tatsache ziehen, dass DB2 genau richtig im Zentrum des potentiellen Metadatenflusses liegt, nämlich zwischen den Modellierungs- und ETL-Werkzeugen und den Abfragewerkzeugen, wie in Abbildung 5 dargestellt. Indem wir das „Back End“ mit dem „Front End“ verbanden, konnten wir DB2 zu dem Kanal machen, durch den alle multidimensionalen Metadaten fließen. Dadurch konnte der gesamte Prozess der nutzbaren OLAP-Lösungen rationalisiert werden.



Um alles in die richtigen Wege zu leiten, suchten wir das Gespräch mit unseren Partnern. Glücklicherweise sprachen wir mit einer Anzahl unserer besten Partner über unser Vorhaben, noch bevor DB2 Cube Views als offizielles Projekt genehmigt wurde. Die Rückmeldung, die wir von diesen Partnern bekamen war, dass der „End to End“ Metadatenfluss durch DB2 sehr viel Sinn macht, aber nur dann wenn das Ganze einfach aufzusetzen ist. Eine komplizierte Schnittstelle, die ein tiefes Verständnis von Katalogtabellen und der SQL Sprache voraussetzte, würde nicht funktionieren. Basierend auf dieser Rückmeldung, entwickelten wir eine einfache Schnittstelle über eine DB2 Stored Procedure. Diese Schnittstelle basiert vollständig auf XML. Wenn DB2 über ein multidimensionales Schema unterrichtet werden soll, ruft man die Stored Procedure auf und gibt ihr eine XML-Beschreibung des Schemas als Parameter mit. Wenn Sie von DB2 Informationen zu einem Schema benötigen, rufen sie dieselbe Stored Procedure auf und erhalten so ein XML-Dokument, welches Ihr Schema beschreibt. Zusätzlich zur Bereitstellung einer einfachen Schnittstelle für unsere Partner, schirmt Sie die Stored Procedure von den unterliegenden Tabellen ab und liefert zusätzliche Informationen in Form einer automatischen Plausibilitätsprüfung der ankommenden XML-Dokumente. Mit der richtigen Schnittstelle waren wir in der Lage, mit unseren Partnern gemeinsam einen „End to End“ Datenfluss von multidimensionalen Metadaten durch DB2 zu realisieren. Der Code, den Partner entwickeln, um DB2 Cube Views anzubinden, ist als „Bridge“ bekannt. Wir waren sehr froh, DB2 Cube Views bereits mit einer Anzahl von fertigen „Bridges“ von führenden Partnern auf den Markt bringen zu können.

Eine andere Tatsache, die uns in der Zusammenarbeit mit unseren Partnern aufgefallen ist, war, dass es keinen einheitlich akzeptierten Weg gab, um multidimensionale Metadaten auszutauschen. Einige von Ihnen werden nun denken „Was ist mit CWM (Common Warehouse Metamodel)“?. Wir betrachteten CWM sehr gründlich und diskutierten es mit einer Anzahl von Partnern und Kunden. Wir hätten gerne einen Metadaten-Standard verwendet, aber waren der Meinung, dass CWM, so wie es im Moment definiert ist, nicht ausreicht, um einen sinnvollen Austausch von



aussagefähigen Informationen entsprechend den Anforderungen unserer Partner zu gewährleisten. Unsere Metadaten müssten durch zwei CWM Pakete befördert werden und dennoch hätten wir eigene Erweiterungen hinzufügen müssen, um einen sinnvollen Austausch zu gewährleisten. Partner hätten dann eine DB2 Cube Views spezifische CWM-Bridge entwickeln müssen, und die schwierige Aufgabe, Informationen aus zwei verschiedenen Paketen konsolidiert zurückzubringen. Leider scheint dies der allgemeine heutige Status von CWM zu sein. Wir wollen den CWM-Standard nicht herabsetzen, er ist die einzige Bestrebung der Industrie, die das Bedürfnis nach Metadatenaustausch adressiert. IBM ist ein aktiver Teilnehmer in den Standard-Ausschüssen für CWM, aber in diesem Fall traf es nicht ganz unsere Bedürfnisse und erschwerte die Aufgaben für unsere Partner. Einer von unseren aktuellen Partnern (Meta Integration Technology AG oder MITI), der sich auf Metadatenaustausch spezialisiert hat, machte es möglich, DB2 Cube Views Metadaten mit ihrem Werkzeug (Meta Integration Modell Bridge) in das CWM-Format umzuwandeln. Wenn Sie CWM konforme Metadaten aus DB2 Cube Views benötigen, hat MITI eine Lösung für Sie.

Zusätzlich zu dem Problem des Metadatenaustauschs fanden wir auch, dass jeder der Partner eine leicht differenzierte Betrachtungsweise zu OLAP-Metadaten hat und wie unterschiedliche Perspektiven von multidimensionalen Modellen in verschiedenem Ausmaß wirksam einzusetzen sind. Hyperion Essbase und IBM DB2 OLAP Server haben sehr strenge multidimensionale Strukturen, die durch die Dimensionen und verbundene Hierarchien definiert sind. Diese Strukturen liefern sehr schnelle „slice and dice“- und „drill“-Fähigkeiten entlang definierter Pfade. MicroStrategy, einer unserer führenden Partner, hat einen unterschiedlichen Denkansatz, der mehr durch



die Attribute als durch die Dimensionen definiert ist. Diese aufgelockerte Struktur erlaubt es Benutzern, überall effektiver zu „drillen“ und irgendein Attribut mit irgendeinem anderen Attribut zu verbinden. Diese verschiedenen Ansätze bringen unterschiedliche Anforderungen an die Metadaten mit sich. Wir hatten zu entscheiden, wie wir einen effektiven Metadatenaustausch zu einem breiten Spektrum von Partneranwendungen ermöglichen wollten. Wie es in solchen Situationen häufig der Fall ist, hätte der kleinste gemeinsame Nenner wenig Wert gehabt. Daher benötigten wir einen anderen Ansatz. Unsere Entwicklungsteam kam auf die einfache, aber hocheffektive Idee, die multidimensionalen Objekte zu schichten und zu verkapseln. Dies ermöglicht unseren Partnern, von diesen Informationen so viele abzurufen, wie sie benötigen und diese mit Ihren eigenen Mitteln zu interpretieren. Ein Nebeneffekt dieses Ansatzes war unsere Erkenntnis, dass diese Methode des Metadatenaustausches wesentlich effektiver und wertvoller für einen viel größeren Bereich von BI-Tools ist. Reportingwerkzeuge können feststellen, welche Attribute verfügbar sind und mit welchen Entitäten (Dimensionen) sie verbunden sind. Data Mining-Werkzeuge können einen Vorsprung in der Attribut-Entdeckung bekommen und dann die Ergebnisse mit den ermittelten Attributen kommunizieren. Geografische Analysewerkzeuge (Spatial Tools) können auch Nutzen aus neuen Attributen ziehen, die schnell für Reporting-, OLAP- und sogar Mining-Werkzeuge verfügbar gemacht werden können. Durch die Zusammenarbeit mit einer steigenden Anzahl von BI-Partnern, erwarten wir, dass die DB2 Cube Views Metdaten zunehmend eine wichtige Rolle spielen werden, um den nahtlosen Austausch von Informationen zwischen einer großen Anzahl von analytischen Werkzeugen und Technologien zu erleichtern.

Abbildung 6 veranschaulicht, wo DB2 Cube Views Metadaten gespeichert sind und mit welchen DB2-Objekten sie kommunizieren. Lassen Sie uns das Würfel-Modell näher betrachten. Stellen Sie sich das Würfel-Modell als einen kompletten Satz von Metadaten vor, die ein multidimensionales Schema beschreiben. Innerhalb des Würfel-Modelles starten wir mit der niedrigsten Stufe von Objekten, den Kennzahlen und Attributen. Dies können einfach Pointer zu einer Spalte in einer Faktentabelle oder einer Spalte in einer Dimensionstabelle sein. Kennzahlen haben drei Komponenten: ein Name, einen Ausdruck und ein Summierungsskript. Mit diesen drei Komponenten es ist möglich, eine große Anzahl von komplexen Kennzahlen zu beschreiben, inklusive Kennzahlen, die auf anderen Kennzahlen basieren, abgeleiteten Kennzahlen aus der Benutzung von mathematischen SQL-Funktionen, verdichteten Kennzahlen, die unter Benutzung einer Vielfalt von Operatoren und Funktionen über verschiedene Dimensionen erstellt werden. Gekoppelt mit den analytischen Funktionen von DB2 sind DB2 Cube Views Kennzahlen sehr leistungsstark. Auch Attribute können unter Benutzung ähnlicher Fähigkeiten abgeleitet und berechnet werden. Auf dieser niedrigen Stufe beschreiben DB2 Cube Views Metadaten auch Joins. Im Falle eines einfachen Star-Schemas ist dies enthalten im Join zwischen der Fact- und Dimensionstabelle.

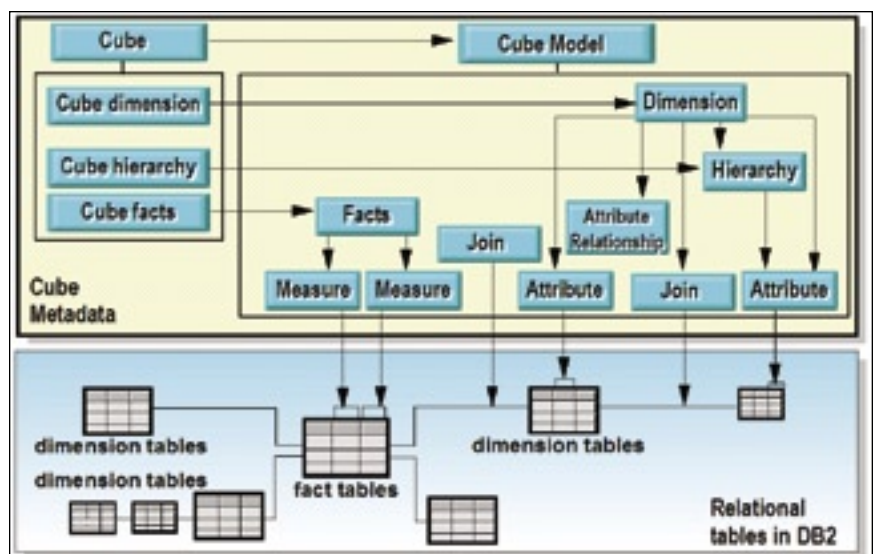


Abbildung 6: OLAP-Metadaten in DB2 Cube Views

In einem komplexeren Snowflake-Schema muss der Join zwischen Dimensionstabellen und „outtrigger“ Tabellen modelliert werden. Auf der nächsten Stufe der DB2 Cube Views Metadaten finden wir die Attribut-Beziehungen. Eine Attribut-Beziehung verkapselt Informationen über die Beziehung zwischen zwei Attributen. Die Information enthält den Typ der Beziehung: beschreibend oder bezogen, die Kardinalität der Beziehung und in einigen Fällen die Tatsache, dass eine funktionelle Abhängigkeit zwischen den Attributen existiert. Auf der nächsten Stufe der Metadaten finden wir Hierarchien. DB2 Cube Views erlaubt Ihnen, eine grosse Anzahl von Hierarchietypen einschließlich: „balanced, unbalanced, ragged und network“ zu modellieren. Zusätzlich können Sie entweder „Standard“ oder „Rekursiv“ Einsatz für „ragged“ Hierarchien spezifizieren. Eine beliebige Anzahl von alternativen Hierarchien können für jede Dimension definiert werden. Die nächste Stufe der Metadaten ist die Dimension selbst, die verwandte Sätze von Attributen und Rangordnungen in einen Zusammenhang bringt. Diese geringe Menge von geschichteten Metadatenobjekten gibt uns ausreichende Möglichkeiten und Aussagekraft, um ein multidimensionales Schema im Detail zu beschreiben, Damit können wir eine große Anzahl von Partnern effektiv unterstützen.

Die linke Seite von Abbildung 6 zeigt den DB2 Cube Views Würfel. Dies ist ein vollkommen virtueller Würfel, der in keiner Weise materialisiert wird. Würfel in DB2 Cube Views haben zwei Verwendungszwecke. Wie Sie sich vorstellen können, sind Würfel relativ einfache Teilmengen eines vollständigen multidimensionalen Schemas, das von den anspruchsloseren Reportingwerkzeugen am unteren Ende des OLAP-Spektrums abgefragt werden kann. Ein Würfel ist eine Anzahl von Pointern, die auf das Würfel-Modell zeigen. Diese Pointer wählen Kennzahlen, Dimensionen und eine Hierarchie pro Dimension aus. Da Würfel nicht materialisiert sind, kann man sich vorstellen, viele Würfel aufzusetzen, die eindeutige Sichten auf die Daten für bestimmte Benutzer oder Benutzergruppen zur Verfügung stellen. Der zweite Verwendungszweck von Würfeln ist der Anteil der Informationen, den wir benutzen, um das multidimensionale Schema aus den Informationen zu optimieren. Wir werden dies später noch im Detail betrachten.

An diesem Punkt sollten wir erwähnen, dass DB2 Cube Views Metadaten in Erweiterungen der DB2-Kataloge gespeichert sind. Sie ermöglichen den „End to End“ Datenfluss von multidimensionalen Informationen zur Unterstützung von OLAP und Business Intelligence-Lösungen. DB2 Cube Views sollte nicht mit einem Metadaten Repository verwechselt werden. Es bietet keine „Check in“ und „Check out“ Fähigkeiten, Versionskontrolle, Abstammungshinweise, Impact-Analyse und andere Eigenschaften, die ein allgemeines Metadatenrepository unterstützen muss. Wenn Sie nach dieser Funktionalität suchen, dann bieten unsere Partner Ascential und MITI ausgezeichnete Lösungen an. Da beide Partner Bridges zu DB2 Cube Views anbieten, können multidimensionale Metadaten zwischen diesen Repositories und DB2 ausgetauscht werden.

DB2 Cube Views enthält eine Administrationsschnittstelle, das sogenannte OLAP-Center (s. Abbildung 7). Von dieser Schnittstelle aus können Administratoren und Designer multidimensionale Metadaten von DB2 Cube Views importieren und exportieren, definieren, anschauen und aktualisieren sowie den Optimization Advisor starten.

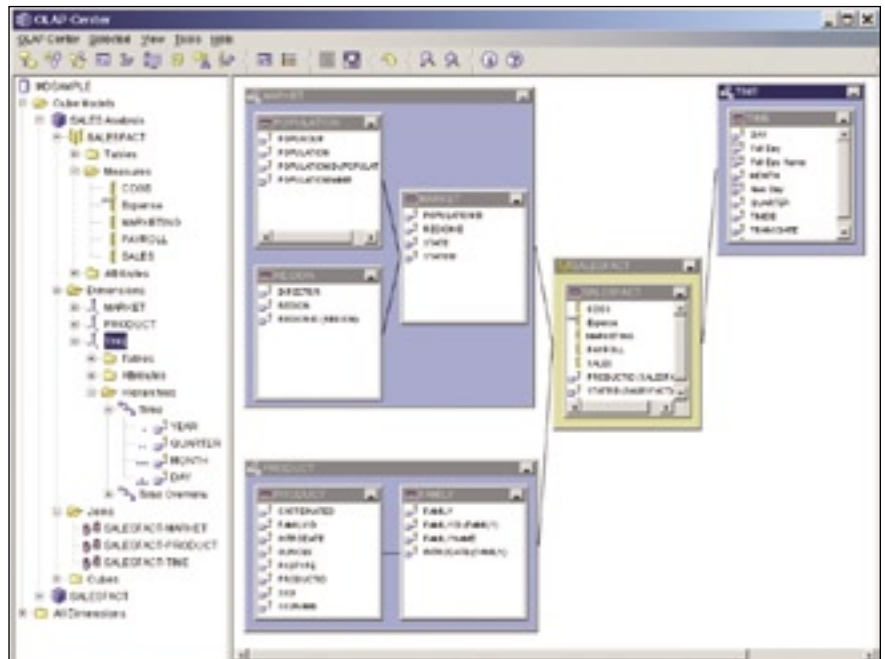


Abbildung 7: Die Administrationsschnittstelle von DB2 Cube Views: Das OLAP-Center



Jetzt „kennt“ DB2 die OLAP-Hierarchien und ist ein aktiver Teilnehmer am „End to End“ Datenfluss der multidimensionalen Metadaten. Die vorhandenen Informationen können nun benutzt werden, um die Business Intelligence-Lösung zu optimieren.

Aggregationen und Materialized Query Tables

Die effektive Verwaltung von Aggregationen ist eine Schlüsselfunktion aller OLAP-Lösungen. Die Fähigkeit zur Analyse summierter Informationen hinunter bis zur Detailinformation über mehrere Stufen und Dimensionen ist eine Anforderung, die OLAP-Benutzer an eine solche Lösung stellen. Um diese Anforderung zu erfüllen, versuchen OLAP-Lösungen die Benutzeranforderungen gegen die benötigte Zeit und den benötigten Speicherplatz abzuwägen, um die notwendigen Aggregationen aufzubauen, zu laden und zu speichern. Die richtige Mischung herauszufinden, ist nicht einfach, weil die Anzahl der OLAP-Datensätze schnell explodiert. Jedes zusätzliche Member, das einem OLAP-Modell hinzugefügt wird, hat eine Sprengwirkung auf den potentiell erforderlichen Zeit- und Speicherbedarf für das Laden, die Berechnung und die Speicherung der Aggregationen. Scheinbar einfache und selbstverständlich scheinende Anforderungen können sehr schnell die Grenzen der OLAP-Technologie sprengen. Viele der heutigen OLAP-Server benutzen Würfel-Speichertechniken, die dieses Problem in gewissem Maße adressieren, aber keine dieser Lösungen, so elegant sie auch ist, befriedigt die Erfordernisse der gegenwärtigen Benutzer vollständig. Als die beste OLAP-Architektur stellt sich zunehmend ein würfelbasierendes Zwischenspeichern der Daten heraus, das den Benutzern jedoch zusätzlich ein transparentes „Drillen“ zu Detailstufen und zusätzlichen Attributen erlaubt, die in der darunterliegenden relationalen Datenbank gespeichert sind. Diese hybride Architektur erlaubt OLAP-Lösungen, die sowohl OLAP-Würfel-Strukturen als auch relationale Datenbanken umfassen. Diese Architektur unterstützt die Benutzeranforderungen an die Breite und die

Tiefe von Analysen und erlaubt Administratoren, einige Daten im Würfel und die anderen Daten in der relationalen Datenbank zu speichern. Um diese Architektur optimal arbeiten zu lassen, muss die darunterliegende relationale Speicherung für OLAP optimiert sein, sonst werden die Benutzer ein nicht akzeptables Antwortzeitverhalten feststellen, wenn die Datenanalyse von den vorberechneten Aggregationen, die im Zwischenspeicher stehen, zu den Aggregationen umschaltet, die ad hoc in der relationalen Datenbank berechnet werden. Um diese Herausforderung anzugehen, muss die relationale Datenbank auch Aggregationen speichern. DB2 hat eine grosse Anzahl von OLAP-Unterstützungsfunktionen und sehr gute Aggregationsmechanismen. Bis zum Erscheinen von DB2 Cube Views war jedoch die Entscheidung darüber, welche Aggregationen zu bilden waren, um diese Fähigkeiten zu nutzen, eine extreme Herausforderung an den Datenbank-Administrator.

Es gibt theoretisch zwei Entscheidungen beim Verwalten von Aggregationen: blindlings alle möglichen Summen bilden, die potentiell von Interesse sein könnten, oder Summen nur bedarfsgerecht bilden, wenn der Benutzer sie abfragt. Die erste Methode ist nicht praktikabel, da nicht genug Zeit und Speicherplatz zur Verfügung steht, um jede mögliche Aggregation zu berechnen, zu speichern und damit verfügbar zu machen. Die andere Methode gibt den Benutzern nicht die benötigten hoch-interaktiven Analysemöglichkeiten, weil die Antwortzeiten sehr lange und unberechenbar werden können. Die Lösung ist, nur einige Aggregationen aufzubauen. Damit stellt sich aber die Frage: Welche Aggregationen sind aufzubauen? Hier kommt der DB2 Cube Views „Optimization Advisor“ ins Spiel.

Die Kernfrage des angesprochenen Problems ist: Wie minimiert man die Zeit- und den Speicherplatzbedarf für den Aufbau von Aggregationen, ohne im Voraus zu wissen, welche Aggregationen der Benutzer abfragen will. In DB2 Cube Views gibt es eine Anzahl von Techniken, die Ihnen helfen, diese Gleichung zu optimieren.

1. Eingaben der Datenbank-Administratoren

Der DBA wählt eines oder mehrere analytische Muster aus einer vordefinierten Liste aus, die der „Opimization Advisor“ anbietet. Diese Muster weisen darauf hin, wo innerhalb des gesamten Modells die meisten Benutzeraktivitäten erwartet werden. Wenn der DBA erwartet, dass „Extrahieren“ die am meisten gewählte Operation ist, dann wissen wir, dass ganz bestimmte Schnitte durch den Datenbestand abgerufen werden. Somit können die Aggregationen in Richtung dieser Scheiben optimiert werden. Wenn der DBA erwartet, dass ein „drill down“ die am meisten gewählte Operation ist, dann erwarten wir, dass die Datenabfrage auf einer hohen Stufe innerhalb des Modells anfängt und dann weiter nach unten drillt. Die Berechnung sehr hoher Aggregationsstufen ist teuer, aber die Breite des Modells nimmt ab, je höher man in der Rangordnungen nach oben klettert, daher können wenige gut platzierte Aggregationen nahe oder direkt am Anfang eines Modells enorm helfen. Dies bedeutet nicht, dass wir mit der Aggregation am Anfang starten und uns dann herunterarbeiten, denn es kann effizienter sein, Aggregationen eine oder zwei Ebenen tiefer im Modell aufzubauen, auf die DB2 dann bei der ad hoc Berechnung von Werten auf höheren Aggregationsstufen zurückgreifen kann. Um zu bestimmen, welche Aggregate erstellt werden sollten, wird ebenfalls das multidimensionale Modell analysiert.

2. Das multidimensionale Modell

Die Form des multidimensionalen Modells ist von besonderer Bedeutung, wenn wir erwägen, wo Aggregate angelegt werden müssen. Insbesondere muß die Form der Dimensionen auf jeder Stufe berücksichtigt werden. Wenn wir ein Modell haben, das sich stufenweise über Detailstufen entfaltet, können wir Aggregationen aufbauen, die regelgerecht durch das Modell schneiden. Die meisten Modelle sind jedoch nicht so geordnet, daher lohnt es sich nach Regionen innerhalb der Dimensionen Ausschau zu halten, in denen viele „Children“ in einen „Parent“ münden. Dann werden Aggregationen in jenen Teilen des Modells gebildet, wo die Datenbank eine grosse Anzahl von Zeilen auslesen müsste, um eine Summe zu erhalten. Dies ist offensichtlich eine gute Investition. Das Szenario wird noch komplexer, wenn die Daten selbst nicht einheitlich sind, daher berücksichtigt DB2 Cube Views auch bestimmte Muster innerhalb der Daten.

3. Die Daten

Multidimensionale Datenbestände sind üblicherweise sehr spärlich und ungleichmäßig gefüllt. Jedes internationale Unternehmen hat Produkte, die auf individuelle Märkte zugeschnitten sind. In einem OLAP-Vertriebsmodell für ein solches Unternehmen, wird der Schnittmenge von europäischen Produkten und amerikanischen Kunden sehr spärlich gefüllt sein. Viele Firmen haben einen uneinheitlichen Verlauf ihrer Umsätze, entsprechend der Zyklen innerhalb der Branchen, die sie beliefern. Wenn wir uns um spärlich und uneinheitlich gefüllte Zellen nicht kümmern, bilden wir Aggregationen, die basierend auf der vorherigen Analyse perfekt und vernünftig aussehen, aber dann zu sehr „dünn besiedelten“ Datenregionen innerhalb des Würfels führen. Um dies zu vermeiden, entnehmen wir mit den neuen DB2 Stichproben-Funktionen eine Stichprobe der Daten. Dies hilft uns, zu verstehen, wo die dichten gefüllten Zellen sind und wo deshalb die I/O-Kosten hoch wären, wenn die Aggregationen ad hoc gebildet würden.

Basierend auf der Analyse von diesen und anderen Faktoren wird der DB2 Cube Views Advisor Aggregationen empfehlen, die entsprechend dem verfügbaren Speicherplatz und der erforderlichen Zeit für die Analyse den besten Nutzen bringen.

Die Aggregationen stehen jetzt in DB2 gespeichert physisch zur Verfügung. In der Vergangenheit gab es verschiedene Werkzeuge, die diese Aggregationen benutzt haben. Wenn man sich für eine passende Menge von Aggregationen entschieden hat und mit ETL-Verfahren entsprechende Aggregationstabellen in Übereinstimmung mit den in den Werkzeugen spezifizierten Regeln definiert hat, dann konnte man diese dem Werkzeug deklarieren, so dass es sie, wann immer möglich, ausnutzt. Dieser Ansatz war jedoch nicht sehr erfolgreich, wenn man von wohlverstandenen und relativ statischen Anwendungsschemata absieht, die manuell durch Anwendungsspezialisten optimiert und durch vordefinierte ETL-Prozeduren unterstützt wurden. Für die Mehrheit der multidimensionalen Schemata gab es keine passende Aggregationsunterstützung und es wurden häufig Aggregationen gebildet, welche die Berechnung bestimmter Berichte oder Kennzahlen in Berichten unterstützten und beschleunigten. Wenn sich die Anforderungen änderten oder neue Berichte hinzukamen, musste man neue zusätzliche Aggregationen hinzufügen. Im Laufe der Zeit wird es für den DBA unmöglich, zu verfolgen, ob Aggregationen noch in Benutzung sind und wie oft sie benutzt werden. Schon nach kurzer Zeit existiert eine grosse Menge von Aggregationstabellen, eine Menge Redundanz und Duplikate, Verwaltungsaufwand und Prüfungen, sowie ein wachsendes Zeit- und Speicherproblem. Hinzu kommt der Aufwand für die Wartung und das Betreiben der ETL-Prozesse.

Der Modell-basierende Ansatz von DB2 Cube Views löst diese Designprobleme und DB2 Materialized Query Tables (MQTs) lösen die ETL- und Verwaltungsprobleme. DB2 MQTs sind sehr nützliche Datenbankobjekte, die DB2 für Sie verwaltet. Die Aktualisierung einer MQT erfolgt über ein einfaches SQL-Statement und DB2 kümmert sich um der Rest. Außerdem bietet DB2 flexible Möglichkeiten zur Aktualisierung von MQTs, wenn Änderungen in den zugrundeliegenden Daten auftreten: Sie können automatisch oder kontrolliert über eine Batch-Prozedur aktualisiert werden. Auch kann der Administrator entscheiden, ob bei einer Aktualisierung die ganze Tabelle ersetzt werden soll oder ob DB2 auf Basis einer Analyse der Datenänderungen selbst bestimmen soll, welche aggregierten Werte neu zu berechnen sind. Die Kombination aus dem modellbasierten Design von Aggregationen mit DB2 Cube Views und den Aggregationsmöglichkeiten mit DB2 Materialized Query Tables ermöglicht es Administratoren auf sehr effektive Art und Weise, DB2 für OLAP-Lösungen zu optimieren.

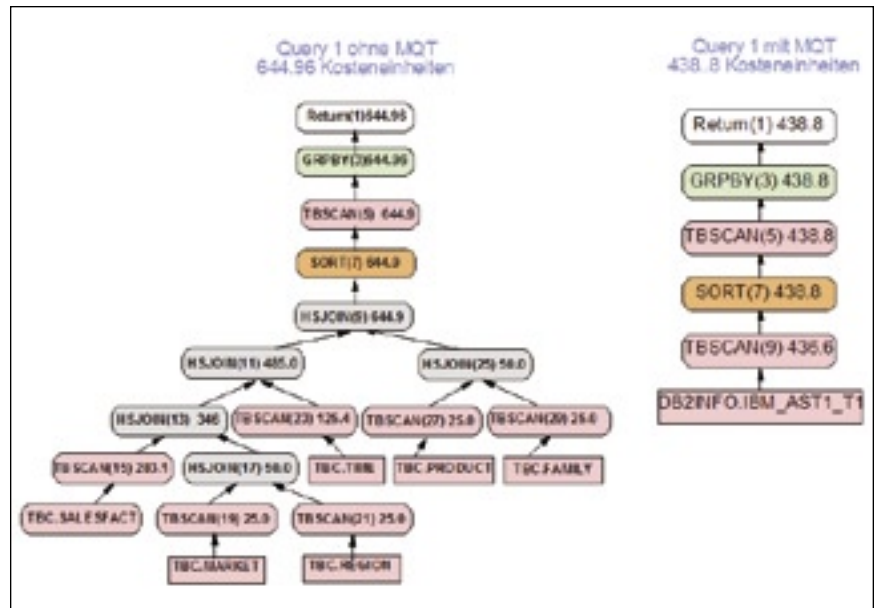


Abbildung 8: Query-Abarbeitungsplan (EXPLAIN) ohne und mit DB2 Materialized Query Tables

Ein weiterer Vorteil der Benutzung von MQTs ergibt sich aus der Tatsache, dass anwendungsspezifische Aggregationstabellen nur jeweils der Anwendung zur Verfügung stehen, für die sie erstellt wurden. Andere Anwendungen, die dieselbe Aggregation verwenden könnten, fehlt der Zugriff auf diese Tabelle, so dass deren Nutzen stark eingeschränkt ist. Da MQTs normale DB2-Objekte sind und damit auch dem DB2-Optimierer bekannt sind, stehen so gebildete Aggregationen für alle Anwendungen und Werkzeuge zur Verfügung und keines von ihnen muss sich isoliert darum kümmern. DB2 schreibt auf Basis einer kostenbasierenden Analyse alle eingehenden SQL-Abfragen automatisch um (Query Rewrite), so dass alle zur Verfügung stehenden Aggregationen bzw. MQTs in optimaler Weise zur Beantwortung der Query herangezogen werden. Das hat zur Folge, dass alle Werkzeuge und Anwendungen die verfügbaren MQTs automatisch und für den Benutzer völlig transparent ausnutzen, auch wenn sie von deren Existenz gar nichts wissen. Abbildung 6 zeigt den DB2-Abarbeitungsbaum für eine SQL-Abfrage ohne und mit Verwendung von MQTs. DB2 ist sehr hoch entwickelt in der Benutzung von MQTs. Es schreibt nicht nur Abfragen um, wenn die abgefragten Daten genau einer existierenden Aggregation entsprechen sondern verwendet wann immer möglich auch teilweise Übereinstimmungen. Dies wird in Abbildung 9 veranschaulicht.

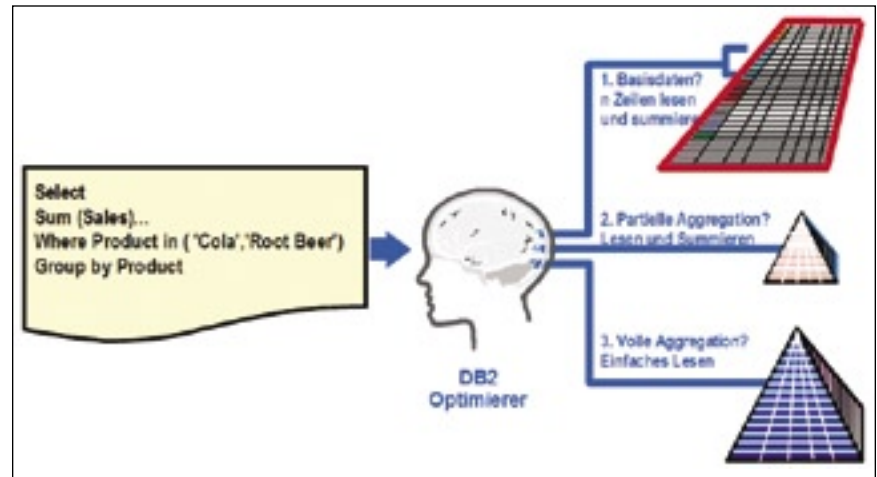


Abbildung 9: Nutzung von DB2 Materialized Query Tables durch den DB2 Optimierer

Web Services für OLAP

Das dritte Ziel von DB2 Cube Views war es, OLAP-Web Services für Anwendungsentwickler bereitzustellen. Viele Organisationen setzen Web Services als neues Modell für die Anwendungsentwicklung ein. Die serviceorientierte Architektur von Web Services und die begleitenden Standards und Infrastruktur werden sehr schnell die bevorzugte Architektur für die Entwicklung neuer Anwendungen werden. Viele der neuen Anwendungen, die auf diesem Weg entwickelt werden, benötigen Zugriff auf multidimensionale Informationen. Ein Beispiel könnte eine Inventarverwaltungs-Anwendung sein, die zur Bestimmung der richtigen Nachbestellmenge dynamische Abfragen auf historische Informationen benötigt. Ein anderes eine neue Service-Anwendung, die Call Center Mitarbeiter mit Informationen über Kunden und kürzlich geleistetem Produktservice versorgt. Dies sind keine typischen OLAP-Endbenutzeranwendungen sondern Anwendungen, die durch die Einbettung von analytischen Informationen in einen Geschäftsprozess einen Mehrwert zur Verfügung stellen. Die Spezialisten, die solche Anwendungen entwickeln, haben ein umfassendes Wissen über die Geschäftsprozesse, die Sie bearbeiten.

Sie sind jedoch keine OLAP-Fachleute, müssen aber im Stande sein, sehr schnell und leicht analytische Informationen in ihre Anwendungen einzubetten, ohne Experten auf diesem Gebiet zu werden. Insbesondere sollten wir nicht erwarten, dass diese Anwendungsentwickler eine komplexe OLAP-Abfragesprache beherrschen müssen, bevor sie für eine Anwendung ein paar Informations-Zellen abrufen oder in einen Algorithmus einbetten können.

Im Hinblick auf diesen Entwickler-Bedarf haben wir eine Technologievorschau von Web Services für OLAP zur Verfügung gestellt, die auf der DB2 Cube Views Infrastruktur aufbaut. Diese Web Services sind schnell und einfach zu benutzen und basieren auf XML-Technologien, so dass sie hervorragend für Entwickler von Web-Anwendungen geeignet sind, die schnellen Zugriff auf analytische Informationen benötigen, ohne OLAP-Experten zu sein.

Ausgabe von Analyseergebnissen mit DB2 Office Connect Analytic Edition

DB2 Cube Views beinhaltet ein „Add-In“ für Excel-Spreadsheets, das interaktiv „Slice“, „Dice“ und „Drill“ direkt gegen DB2-Daten und DB2 Cube Views unterstützt. Das DB2 Office Connect Analytic Edition Add-In liest die DB2 Cube Views Metadaten und erlaubt dem Benutzer direkt, einen Würfel zu wählen, den er analysieren will. Basierend auf den Benutzerauswahlen erzeugt DB2 Office Connect eine SQL-Abfrage, die dynamisch eine Excel Pivot-Tabelle befüllt. Wenn der Benutzer durch den Würfel navigiert, holt sich DB2 Office Connect die entsprechenden Daten aus DB2. Viele dieser Queries fragen aggregierte Daten ab, und natürlich schreibt der DB2-Optimierer diese Abfragen um, damit vorhandene MQTs ausgenutzt werden können. Das Ergebnis ist Hochgeschwindigkeits-„Slice“, „Dice“ und „Drill“ direkt gegen DB2. DB2 Office Connect Analytic Edition bietet unseren Kunden eine benutzerfreundliche und effektive Methode, einfache Analysedaten aus DB2 direkt an Microsoft Excel-Benutzer zu verteilen. Da DB2 Office Connect nicht pro Benutzer lizenziert wird, können Sie ohne zusätzliche Lizenzkosten Analyseergebnisse einer größeren Anzahl von Benutzern als je zuvor zugänglich und nutzbar machen.



	A	B	C	D	E	F	G
2	PRODUCT	PI					
4			TIME->YEAR	TIME->QUARTER			
5			2000				Grand Total
6	MARKET-HP	Data	Qtr1	Qtr2	Qtr3	Qtr4	
7	Central	COGS	13671.38	14401.46	14545.57	13671.71	56289.12
8		SALES	31268.83	33028.87	33554.36	31438.02	129300.08
9	East	COGS	9069.56	9576.1	10056.71	9190.8	37913.17
10		SALES	20573.89	22432.99	22959.09	21334.19	87299.16
11	South	COGS	5221.39	5442	5790.94	5425.77	21880.1
12		SALES	12199.08	12802	13324.89	12642.41	50970.18
13	West	COGS	14850	15713.43	16908.7	15383.66	62855.82
14		SALES	32890.59	34377.9	36436.22	33837.5	137542.21
15	Total COGS		42882.32	45131.29	47181.52	43999.65	179200.82
16	Total SALES		96982.39	102431.73	106273.33	99252.11	404939.56

Abbildung 10: Analyse von DB2 Daten mit Microsoft Excel mit DB2 Office Connect Analytics Edition

DB2 Cube Views Partner

DB2 Cube Views ist eine Plattform für OLAP, die durch die Optimierung zahlreicher Technologien OLAP-Abfragen beschleunigt. Hauptbestandteil unserer Bemühungen ist der nahtlose Fluss von multidimensionalen Metadaten durch DB2, vom Design-Werkzeug bis zur Endbenutzeranwendung. Zur Realisierung dieser Vision arbeiteten wir mit führenden Business Intelligence- und Anwendungsentwicklungs-Partnern zusammen, die wie wir DB2 für die führende Datenbank-Plattform für OLAP halten.

Möglichkeiten zur Integration bestehen in mehreren verschiedenen Bereichen:

- Design, Modellierung und ETL-Werkzeuge bringen multidimensionale Metadaten über eine Bridge in DB2
- Metadatenmanagement-Werkzeuge können multidimensionale Metadaten entpacken und als Teil eines größeren Metadaten-Verwaltungsprozesses mit DB2 Cube Views austauschen
- OLAP-Anbieter können multidimensionale Informationen aus DB2 Cube Views extrahieren, um ihre eigene Verwaltung von OLAP-Metadaten zu betreiben
- Reportingwerkzeuge liefern OLAP-Analysen direkt aus DB2, wobei die DB2 Cube Views Metadaten ihnen helfen, die verfügbaren Dimensionen, Kennzahlen, Attribute und Hierarchien zu verstehen
- Applikationen, die ein in DB2 gespeichertes multidimensionales Schema benutzen, können Vorteile aus den von DB2 Cube Views entworfenen Materialized Query Tables ziehen, wenn DB2 Cube Views Metadaten definiert sind und der „Optimization Advisor“ gestartet ist.

Als DB2 Cube Views auf den Markt kam, lieferten eine Anzahl von Partnern bereits Bridges die eine oder mehrere dieser Möglichkeiten nutzten. Der folgende Abschnitt gibt Ihnen einen Überblick über die von jedem Partner verfügbaren Möglichkeiten:

Ascential

Ascential hat zu MetaStage eine Metabrokerbridge für den Austausch von Metadaten mit DB2 und DB2 Cube Views verfügbar gemacht.

Brio

Brio hat seine Business Intelligence Suite und seine neue Metrik Builder-Anwendung mit Bridges für DB2 Cube Views verfügbar gemacht.

Business Objects

Business Objects hat DB2 Cube Views über seine universelle Metadatabridge integriert.

Cognos

Cognos entwickelte den DB2 multidimensionalen Metadaten Assistent, um ihre Business Intelligence-Plattform mit DB2 Cube Views zu integrieren.

Hyperion

IBM und Hyperion haben zusammengearbeitet, um die Integration Server Bridge zu entwickeln . Diese Bridge ermöglicht den bidirektionalen Austausch von Metadaten zwischen DB2 Cube Views und dem Essbase/OLAP Integration Server.

Meta Integration Technologies AG (MITI)

MITI-Tools ermöglichen den Metadaten austausch über einen breiten Bereich von Werkzeugen und Anwendungen hinweg. Als Spezialist für den Austausch von Metadaten mit Modellierungs- und Designwerkzeugen, hat MITI einen fast lückenlosen Satz von Bridges für Modellierungswerkzeuge aller führenden Anbieter erstellt. In den vergangenen Jahren hat MITI eine wachsende Anzahl von Bridges zur Unterstützung von BI-Werkzeugen und -Anwendungen entwickelt.

MicroStrategy

MicroStrategy hat den DB2 Cube Views Import Assistenten entwickelt, um ein MicroStrategy Projekt aus den DB2 Cube Views Metadaten zu bauen.

Rocket Software

Rocket Software entwickelte eine OLAP-Abfrageschnittstelle für QMF for Windows. Diese Schnittstelle erlaubt QMF-Benutzern ein interaktives „Slice“, „Dice“ und „Drill“ auf DB2-Daten.

Die Liste von DB2 Cube Views Partnern wächst stetig. Besuchen Sie die DB2 Cube Views Webseite auf ibm.com/Data/db2/db2md, um mehr über die DB2 Cube Views Partner zu erfahren.



Zusammenfassung

Mit DB2 Cube Views hat IBM DB2 besser als je zuvor mit der BI- und OLAP-Welt integriert. Mit BI-Metadaten, die als Teil des DB2-Katalogs verwaltet werden, kennt DB2 jetzt Dimensionen, Hierarchien, Attribute, Kennzahlen und mehr, die alle multidimensionalen Konstruktionen umfassen, die für die Implementierung von BI-Lösungen erforderlich sind. Der DB2 Cube Views Optimization Advisor verwendet ausgefeilte analytische Techniken zum Entwurf von optimalen Aggregationen, um hochperformante BI-Lösungen zu unterstützen und eine optimale Nutzung von DB2s Summierungsmöglichkeiten, erstklassige Gruppenfunktionen und Materialized Query Tables zu gewährleisten.

DB2 Office Connect Analytic Edition gibt Ihnen die Möglichkeit, Analysen aus DB2 an einen grösseren Benutzerkreis als je zuvor zu einem niederen Preis zu liefern. Web Services stellen Analyseinformationen extrem schnell neuen analytischen Anwendungen basierend auf der serviceorientierter Architektur zur Verfügung. Die grosse Partnergemeinde, die DB2 Cube Views mit Metadaten Bridges unterstützt, ermöglichen es Ihnen, integrierte und optimierte Business Intelligence-Lösungen sehr schnell zu implementieren und ressourcensparend zu verwalten. Diese BI-Lösungen können von der Modellierung über ETL-Werkzeuge und DB2 bis hin zu den OLAP-, Abfrage- und Reporting-Werkzeugen Ihrer Wahl alle erforderlichen Komponenten umspannen.



DB2 Cube Views macht aus DB2 die führende Datenbank-Plattform für BI und OLAP. DB2 ist die einzige Datenbank, die die erforderlichen Leistungsmerkmale, Funktionen, Werkzeuge sowie die Integration von Metadaten und Partneranwendungen aufweist, um optimierte BI- und OLAP-Lösungen für Unternehmen bereitstellen zu können sowie Lösungen, die schnelle und zeitnahe Analysen in der Breite und Tiefe zur Verfügung stellen, wie sie heutige Führungskräfte und Analysten verlangen.

In der gegenwärtiger Volkswirtschaft ist Business Intelligence der Schlüssel, um einen tieferen Einblick in die Geschäftsdaten zu bekommen und damit Wettbewerbsvorteile realisieren zu können. Um dies effektiv zu tun, benötigen sie eine integrierte Lösung, die leicht eingesetzt und verwaltet werden kann. Sie benötigen eine Lösung, die die Benutzung von Systemressourcen und Datenbank-Einrichtungen optimiert, um hochperformant und bedarfsgerecht Analyseergebnisse bereitstellen zu können. Und vor allem benötigen Sie eine Lösung, die Ihnen schnell einen Return on Investment bringt. Mit anderen Worten, Sie benötigen DB2 Cube Views.

Für mehr Informationen

Um mehr über DB2 Cube Views zu erfahren, kontaktieren Sie bitte Ihren IBM Vertriebsbeauftragten, IBM Business Partner oder besuchen Sie unsere Website: ibm.com/software/data/db2/db2md

**IBM Deutschland GmbH**

D-70548 Stuttgart
Homepage: ibm.com/de

IBM Österreich

Obere Donaustraße 95
A-1020 Wien
Homepage: ibm.com/at

IBM Schweiz

Bändliweg 21, Postfach
CH-8010 Zürich
Homepage: ibm.com/ch

Die IBM Homepage finden Sie im Internet unter:
ibm.com

Vertragsbedingungen und Preise erhalten Sie bei den IBM Geschäftsstellen und den IBM Geschäftspartnern. Die Produktinformationen geben den derzeitigen Stand wieder. Alle Angaben ohne Gewähr. Gegenstand und Umfang der Leistungen bestimmen sich ausschließlich nach den jeweiligen Verträgen.

IBM, das IBM Logo, das @server Logo, DB2 Cube Views, DB2, DB2 Universal Database, Informix, Lotus und WebSphere sind eingetragene Marken der International Business Machines Corporation in den Vereinigten Staaten und in anderen Ländern.

Microsoft ist eine Marke der Microsoft Corporation in den Vereinigten Staaten und/oder anderen Ländern.

Marken anderer Unternehmen/Hersteller werden anerkannt.

Hinweise in dieser Veröffentlichung auf IBM Produkte, Programme oder Services bedeuten nicht, dass IBM diese in allen Ländern bereitstellt, in denen das Unternehmen tätig ist. Hinweise auf IBM Produkte, Programme oder Services bedeuten nicht, dass nur IBM Produkte, Programme und Services verwendet werden können. Sie können jedes andere Produkt mit identischer Spezifikation verwenden.

IBM Produkte sind fabrikneu hergestellt. Sie können neben neuen auch wieder verwendete Teile enthalten.

Bei IBM heißt Dienst am Kunden auch Dienst an unserer Umwelt. Wir nehmen Ihre IBM Altgeräte zurück und stellen deren umweltfreundliche Entsorgung zum Selbstkostenpreis sicher.

Die vorliegende Veröffentlichung dient ausschließlich der allgemeinen Information. Änderungen ohne Hinweis vorbehalten.

Die neuesten Informationen über IBM Produkte und Services erhalten Sie bei Ihrem lokalen IBM Händler.

Bei abgebildeten Geräten kann es sich um Entwicklungsmodelle handeln.

© Copyright IBM Corporation 2004.

Alle Rechte vorbehalten.