

ESSBASE und die OBIEE 11g - Aufbruch zu "echten" OLAP-Analysen

Frank Tenzer
OPITZ CONSULTING Hamburg GmbH

Schlüsselworte:

ESSBASE, OBIEE 11g, OLAP, ROLAP

Einleitung

Die ORACLE BI-Suite 11g bietet zahlreiche neue Features, die "echte" OLAP-Analyse-Tätigkeit unterstützen. Außerdem gibt es in der Anbindung von ESSBASE Appliactions erhebliche Verbesserungen, die bisher für die OBI "verborgene" ESSBASE-Features wie UDA, Alias-Tables etc. nun endlich nutzbar machen. Der Vortrag verdeutlicht an zahlreichen praktischen Beispielen und einer Live-Demo die Einsatzmöglichkeiten und den Nutzen dieser neuen Features, die ESSBASE zur idealen MOLAP-Quelle für die OBI machen und performante, intuitiv anwendbare und mächtige OLAP-Analyse-Modelle in der OBI 11g ermöglichen. Die folgenden Ausführungen bieten einen Rahmen für die im Vortrag konkret vorgestellten Beispiele und sind als Einführung in das Thema, aber auch als Ergänzung zu den Vortragsfolien gedacht.

Was macht OLAP zu OLAP?

Die Abkürzung OLAP steht für Online Analytical Processing, frei übersetzt bedeutet das in etwa „direkte analytische Datenverarbeitung“. Der Begriff OLAP wurde vor allem durch Edgar Frank Codd im Zusammenhang mit den sogenannten 12 Codd'schen Regeln geprägt und etwa zwei Jahre später durch die von Pendse und Creeth geforderten FASMI-Kriterien auf eine weniger technische und stärker auf die Bedürfnisse der Anwender zugeschnittene Sicht erweitert. OLAP ist eine Methode aus dem Bereich der Analytischen Informations- oder Entscheidungsunterstützenden Systeme und ist charakterisiert durch Analyseverfahren, die im Gegensatz zum Standard-Reporting nicht die exakte Abfrage sowie die Form der Darstellung voraussetzen und damit ein reproduzierbares Ergebnis liefern, sondern einen eher kreativen, freien Umgang mit einer bereitgestellten Datenbasis erfordern, um eine zuvor gestellte Hypothese zu belegen. Eine Hypothese belegen - Man muss also das Ergebnis bereits im Voraus kennen? Nein, das verlangt OLAP nicht mal seinen kreativsten Analysten ab. Aber der OLAP-Analyst hat eine genaue Vorstellung von der Frage, die er an das System „stellen“ möchte. Der Umgang mit den Daten, die kreative „Bewegung“ innerhalb eines oft als Würfel (*Cube*) symbolisierten, multidimensionalen Datenraums führt letztlich zu einem Teilausschnitt aus dem OLAP-Cube (*Report-Slice*) führt, welcher die in der Frage enthaltene Hypothese entweder bestätigt oder widerlegt. Ein Beispiel: „Haben die Erweiterungen der Produktpalette in 2010 die geplante Umsatzsteigerung bewirkt?“ – Diese Frage eines Analysten enthält implizit die Hypothese, dass Maßnahmen erfolgreich waren. Wie kann OLAP dies nun belegen? Ein Vergleich zweier Schnitte aus dem Sales-Cube, entlang der Zeitdimension, des Ist- und Plan-Szenarios, der Kennzahl-Umsatz und jeweils einmal als „as was“ und als „as is“ Sicht auf die Produktdimension offenbaren, ob das Ziel erreicht wurde – und belegen damit die Hypothese. Einige der von Codd aufgestellten Regeln beeinflussen noch immer die grundlegenden analytischen Datenstrukturen hinter OLAP-Systemen, die durch Multidimensionalität, Hierarchien und Metainformationen geprägt sind. Das wohl bekannteste Modell hierfür ist das sogenannte Star-Schema. Die beschriebenen Arten von Bewegung im Datenraum, dem Cube, kennt man als „Slice“ (Schneiden), „Dice“ (Drehen), „Drill Down“ (In die

Tiefe gehen) etc. Dabei steht die Analyse-Tätigkeit weit im Vordergrund vor einem perfekt Vorbereiteten Ergebnis, wie wir es im Standard-Reporting-Bereich erwarten. Der Analyst arbeitet mit den Daten; besonders wichtig sind daher die Antwortzeiten von OLAP-Systemen: „Speed of thought analysis“ – So schnell, wie der Gedanke des Analysten die Idee liefert, zwei Dimensionen in der Pivot-Tabelle zu vertauschen oder per roll-up eine aggregierte Sicht zu gewinnen, ebenso schnell muss das System in der Lage sein, die Antwort zu liefern.

Und was ist mit Data-Mining?

OLAP und Data-Mining werden immer wieder in einem Atemzug genannt, haben Berührungspunkte, sind aber auf keinen Fall identisch. Unter Data-Mining versteht man vielmehr eine Sammlung ganz bestimmter mathematisch-statistischer Analysen, die die Erkennung von Mustern in sehr großen Datenbeständen ermöglichen und somit ebenfalls Entscheidungsunterstützung liefern können. Zu den Aufgaben und damit zu den Analyse-Disziplinen im Data-Mining gehören Ausreißererkennung, Clusteranalyse, Klassifikation, Assoziations- und Regressionsanalyse. Aus den erkannten Mustern in den Analysierten Datensätze lassen sich Prognosen über die Gesamtheit der Datensätze ableiten, die Entscheidungen unterstützen können. Prominente Vertreter sind z.B. Scoring-Verfahren wie die der Schufa (Clustering) oder die Warenkorb- (Assoziations-) Analyse – „n% der Kunden, die eine Zahnbürste kauften, kauften auch Zahnpasta“ – Die Entscheidung, die aus diesem Muster erwachsen kann ist, beide Produkte im selben Regal zu platzieren, um dem Kunden den „Paar-Kauf“ möglichst einfach zu machen und damit das potentielle Umsatz-Volumen möglichst gut auszunutzen.

Was verbindet nun OLAP und Data-Mining? Die Technologie – Beide nutzen die oben angesprochene Art von analytischen Datenstrukturen. Darüber hinaus kann OLAP ein Weg sein, Data-Mining-Analysen zu entwickeln.

ESSBASE – ein (M)OLAP Werkzeugkasten für Profis...und Einsteiger

OLAP scheint fest verbunden mit dem Begriff der Multidimensionalität von Daten – Wozu denn nun noch MOLAP, also multidimensional OLAP? Unter MOLAP versteht man eine Technologie, die ein analytisches Datenmodell und die enthaltenen Daten in einer optimierten Array-Struktur speichert, die über eine speziell für die jeweilige (Produkt- oder Hersteller-spezifische) Struktur entwickelte Engine verwaltet wird, zugegriffen und bearbeitet werden kann - Im Gegensatz zum relationalen (R-) OLAP, welches das Datenmodell in einer nahezu beliebigen relationalen Datenbankstruktur implementiert auf die mit standard-Abfragesprachen wie SQL zugegriffen wird. ESSBASE ist ein MOLAP-DBMS, ein multidimensionales, auf OLAP-Strukturen spezialisiertes Datenbank-Management-System. Damit liegt auf der Hand: ESSBASE ist eine ideale Basis für OLAP-Aktivitäten. Was bietet ESSBASE aber neben diesem offensichtlichen „Match“? In erster Linie sind hier sicherlich zu nennen:

- Effiziente, Leistungsfähige und skalierende Storage-Technologie, die auch große analytische Modelle mit vielen Dimensionen hoher Kardinalität umgehen kann
- Die extrem Leistungsstarke und flexible Calculation-Engine, die mit mehreren hundert Funktionen und der Möglichkeit benutzerdefinierbarer Funktionen alle Anforderungen von Aggregation, Berechnung und Allokation der Daten in den Cubes ermöglicht
- Eine intuitive und fachlich motivierte Modellierungs-technologie, die eine schnelle, agile und aufwandsgünstige Modellierung und Implementierung der Analyse-Modelle - der Cubes bzw. sogenannten Applications - in enger Kopplung mit den Fachanwendern ermöglicht („rapid application development“)
- Flexible, leistungsstarke und schnelle Integrationsmechanismen für unterschiedlichste Arten von Datenquellen von relationalen Datenbanken über Flatfiles bis hin zu den allseits beliebten Excel-Spreadsheets mit den beinhalteten erprobten und komplexen Planungs- und Controlling-Modellen.

- Erweiterungen des rein dimensions-, hierarchie- und zahlengetriebenen Modells eines OLAP-Cubes um Attribut-Dimensionen, user-defined-attributes (UDA), die z.B. der Clustering dienen, alias-tables und text-based-measures ermöglichen aussagekräftige Analysemodelle, die die Sprache der Anwender sprechen.

Diese Faktoren machen ESSBASE zu einer perfekten Wahl im Bereich der OLAP-Middleware. Das klassische Analyse-Frontend bildet dabei Excel im Zusammenspiel mit dem ESSBASE-spezifischen Add-In zur Abfrage und auch Bearbeitung der Daten in den Cubes. Erweiterte Funktionalitäten im OLAP-Bereich bietet der zusätzliche SmartView-Add-In mit seinen Member-Selection-Funktionen, definierbaren Reports-Slices als Basis für wiederkehrende Analyse-Aufgaben und dem Point-of-View-(POV-) Manager, der eine effiziente Navigationsmöglichkeit über die Dimensionen bietet, die nicht im Pivot-Spreadsheet enthalten sind. Eine leistungsstarke Toolbox – aber nur wenig integrativ. Es droht das Leben in der „Spreadsheet-Hölle“, viele dutzend Excel-Sheets, die sich in unterschiedlichsten Bearbeitungsständen vervielfacht in Mailboxen und Dateisystemen wiederfinden, erhöhen den Aufwand und das Fehlerpotential – geht das nicht noch besser?

OBI 11g – ein integratives Frontend für den Analysten...und alle anderen

Die ORACLE BI-Suite 11g bietet eine vollständige und hochintegrative Middleware-Plattform für alle Arten von analytischen Fragestellungen. Von der Analysten-Tätigkeit, der Adhoc-Analyse auf den zentral über den BI-Server angebotenen Metadaten-Katalogen und den darunter integrierten Datenbeständen über die flexible, dynamische und interaktive Entscheidungsunterstützung durch gemeinsame und individuelle Dashboards bis hin zum automatisch in verschiedensten Formaten und auf eine Reihe unterschiedlichster Endgeräte publizierten, „pixel-perfekt“ gestalteten Standard-Reports – jede Anforderung wird bedient. Eine klar strukturierte, skalierende und integrative Schichten-Architektur ermöglicht dabei die Durchführung aller Aufgaben über den Webbrowser- sei es am Arbeitsplatz oder mit dem iPad im Taxi. Dabei glänzt die aktuelle Version der Suite mit OLAP-relevanten Features wie hierarchiebasierten Zeilen- und Spalten-Darstellungen, die echtes drill-down und roll-up ermöglichen, dem korrekten Umgang mit unbalancierten Hierarchien, unterschiedlichsten grafischen Darstellungsformen und vielem mehr. Zu den Quell-Systemen gehören dabei neben relationalen Star-Schemata auch ESSBASE als echter MOLAP-Provider. Die aktuelle Version der OBI kann die ESSBASE-Strukturen native ansprechen, ohne den Umweg über Standards wie MDX zu gehen. Das Resultat sind zum Einen performante Analysen auch auf umfangreichen und komplexen Cubes, zum Anderen aber auch der Gewinn an Flexibilität und Unterstützung im Analyse-Frontend, da z.B. auch die ESSBASE-seitig angebotenen Funktionen zur Aggregation, Berechnung oder Allokation für den Analysten nutzbar werden. Diese können bei Bedarf in Analysen verwendet werden, ohne dass der Cube angepasst werden muss, was wiederum die freie Analysetätigkeit unterbrechen würde. Nicht zu unterschätzen ist aber auch die Fähigkeit der aktuellen OBI, die von den ESSBASE-Applications angebotenen Metadaten zu Dimensionen und ihren Hierarchien vollständig nutzbar zu machen und direkt in den Business- und den Präsentation-Layer des OBI-Repositorys zu übernehmen. Mit wenigen Mausklicks promoten Administratoren exakt die mit den Fachabteilungen entwickelten und abgestimmten neuen oder angepassten Cubes direkt in das BI-Frontend. Keine zusätzliche Metadaten-Definition ist erforderlich. Somit geben die OBI 11g und ESSBASE offenbar ein perfektes Team, wenn es um OLAP-Analysen geht.

Ein eingespieltes Middleware-Team unterstützt „echtes“ OLAP

Wie sieht es nun „auf den zweiten Blick“ mit der Unterstützung von OLAP-Anforderungen aus, wenn wir das Gespann aus OBI 11g & ESSBASE betrachten? Entfernen wir uns einmal von einer wohlwollenden Bewertung anhand der vielen genannten Vorteile und ziehen uns auf eine neutrale Bewertung anhand der FASMI-Regeln von Pendse und Creeth zurück:

1. *„Fast: Abfragen sollen durchschnittlich fünf Sekunden dauern dürfen. Dabei sollen einfache Abfragen nicht länger als eine Sekunde und nur wenige, komplexere Abfragen bis zu 20 Sekunden Verarbeitungszeit beanspruchen“*: Die genannten Zahlenwerte sind sicherlich nicht mehr als aktuell repräsentativ anzusehen, sicherlich belegbar ist aber, dass die Kombination aus den leistungsstarken storage-engines (BSO, ASO), der calculation-engine und einem effektiven caching-Konzept und dem verbesserten ESSBASE-Integration in der OBI 11g und dem damit verbundenen nativen Zugriff auf die Cubes deutlich in der Lage sind „speed of thought analysis“ zu ermöglichen.
2. *„Analysis: Ein OLAP-System soll jegliche benötigte Logik bewältigen können. Dabei soll die Definition einer komplexeren Analyseabfrage durch den Anwender mit wenig Programmieraufwand zu realisieren sein“*: Wer einmal mit der OBI beziehungsweise dem beinhalteten „Answers“-Nachfolger „Analytics“ gearbeitet hat, wird bestätigen, dass der Programmieraufwand durch den Anwender meist gegen Null geht. Lediglich extrem komplexe Ausnahme-Situationen erfordern die Verwendung von ESSBASE-Funktionsaufrufen, die Masse der Anforderungen lassen sich im Stil der Verwendung einer Excel-Pivot-Tabelle realisieren.
3. *„Shared: Ein OLAP-System soll für den Mehrbenutzerbetrieb ausgelegt sein. Dies bedingt eine Verfügbarkeit geeigneter Zugriffsschutzmechanismen“*: Die OBI-Architektur stellt erwiesenermaßen eine skalierbare, mehrbenutzerfähige Middleware-Plattform dar. ESSBASE bietet zusätzlich Multithreading- und Locking-Mechanismen an, um einen performanten Mehrbenutzerbetrieb zu ermöglichen. Zugriffsschutz-Mechanismen bestehen ebenfalls in beiden Systemen, wobei ESSBASE-Cubes mit einer fachlich orientierten, Filter-basierten und intuitiv anwendbaren Row-Level-Security ein Höchstmaß an individuellen Security-Anforderungen unterschiedlichster Unternehmensbereiche erfüllen kann.
4. *„Multidimensional: Als Hauptkriterium fordern Pendse und Creeth eine mehrdimensionale Strukturierung der Daten mit voller Unterstützung der Dimensionshierarchien“*: ESSBASE ist ein rein auf MOLAP-technologie basierendes DBMS. Dadurch ist implizit nur eine rein multidimensionale und streng hierarchiebasierte Modellierung möglich.
5. *„Information: Bei der Analyse sollen einem Anwender alle benötigten Daten transparent zur Verfügung stehen. Eine Analyse darf nicht durch Beschränkungen des OLAP-Systems beeinflusst werden“*: Durch die neuen Features der OBI 11g in Bezug auf ESSBASE stehen sämtliche relevanten Informationen im Analyse-Frontend zur Verfügung: Von den Dimensions- und Hierarchie-Metadaten des ESSBASE-Outline bis hin zu sämtlichen Erweiterungen des Datenmodells durch Attribut-Dimensionen, UDAs oder text-based-measures.

Soweit zur Theorie. Das dieses eingespielte Team auch in der Praxis überzeugt offenbart sich an konkreten Beispielen, wie sie in der Live-Demo zum Vortrag enthalten sind. Ohne Frage: Das Duo OBI 11g & ESSBASE stellt ein echtes OLAP-System dar und bietet so die Chance, vom puren Reporting zu Entscheidungsunterstützenden Analyseformen aufzubrechen.

Kontaktadresse:

Frank Tenzer

OPITZ CONSULTING Hamburg GmbH

Butendeichsweg 2

D-21129 Hamburg

Telefon: +49 (0) 40-741122-1360

Fax: +49 (0) 40-741122-4360

E-Mail Frank.Tenzer@opitz-consulting.com

Internet: www.opitz-consulting.com