

Business Intelligence In-Memory

Annett Thurm-Meyer ORACLE Deutschland B.V. & Co. KG

Business-Anforderungen werden komplexer, Zeitfenster kürzer, Ressourcen knapper. Der Schlüssel zu erfolgreichem Business Intelligence ist die zeitnahe Gewinnung von Einblicken und Erkenntnissen, die entscheidungsrelevante Informationen liefern.

Performance ist neben den funktionalen Innovationen die Top-Anforderung im Umfeld von Business Intelligence. Während komplexe Berichte und Analysen vor geraumer Zeit noch Stunden, wenn nicht Tage in Anspruch genommen haben, ist heute die Verarbeitung derselben Berichte und Analysen In-Memory, also im Hauptspeicher, in kurzer Zeit möglich. Damit kann Business Intelligence auch als Entscheidungsunterstützung für operative Prozesse, ja sogar near Realtime genutzt werden. Der Artikel zeigt, wie Business Intelligence mit In-Memory-Technologie zusammenarbeitet.

Wie Business Intelligence In-Memory funktioniert

Die für Business Intelligence genutzte In-Memory-Funktionalität bezieht sich bei Oracle hauptsächlich auf ein Datenbank-Management-System, das auch den Arbeitsspeicher eines Servers als Datenspeicher nutzt. Herkömmliche Systeme halten die Daten auf Festplatten. Dabei kostet der ständige Lesezugriff auf die Festplatte Zeit, die auch durch schnellere Platten nicht beliebig beschleunigt werden kann.

Die Zugriffsgeschwindigkeiten und Algorithmen für den Zugriff auf Daten im Arbeitsspeicher sind deutlich einfacher und damit performanter. Durch die gesunkenen Preise für Arbeitsspeicher ist es heute sinnvoll möglich, Daten im Arbeitsspeicher zu halten. Reicht ein Arbeitsspeicher nicht aus, können herkömmliche Technologien wie Grid Computing dazu verwendet werden, die Datenhaltung auf die Arbeitsspeicher mehrerer Server zu verteilen.

Die traditionelle relationale Oracle-Datenbank hat die Daten in Zeilen gespeichert. Jeder neue Datenbank-Eintrag wird als weitere Zeile abgespeichert. Jede Zeile ist dann wiederum in mehreren Spalten gegliedert, die jeweils Attribute beinhalten.

Diese Art der Datenhaltung ist perfekt für die Speicherung von operationalen Daten (OLTP) geeignet. Neu ist die Datenhaltung in Spalten („Columns“). Sie speichert jedes Attribut einer Transaktion oder eines Datenbank-Eintrags in einer separaten Spaltenstruktur.

Das Spaltenformat ist ideal für Analysen, weil es ein schnelleres Auslesen der Daten erlaubt, wenn nur einige Daten und nicht ganze Zeilen aus einem großen Datenbestand ausgelesen werden müssen. Logisch ist der Umkehrschluss, dass sich eine spaltenbasierte Haltung von Daten nicht für DML („einfügen“, „update“ oder „löschen“) eignet. Die In-Memory-Option der Oracle-Datenbank ermöglicht, das Beste aus beiden Welten zu nutzen und damit Oracle Business Intelligence optional zu unterstützen (siehe Abbildung 1).

Sobald eine In-Memory-Speicherung für Tabellen, Materialized Views oder auch nur ausgewählte Spalten aktiviert ist, erzeugt

die Oracle-Datenbank automatisch eine In-Memory-Kopie der Daten und hält diese synchronisiert.

Diese In-Memory-Daten stehen dann für hoch performante Analysen und Berichte zur Verfügung. Dabei ist es nicht erforderlich, die analytische Applikation anzupassen. Der Optimizer der Datenbank entscheidet, aus welchem Speicher die Daten schnellstmöglich geliefert werden können, um die Anfrage des BI-Servers zu bedienen.

Exalytics – die Business Intelligence In-Memory Machine

Oracle hat das Thema „Business Intelligence In-Memory“ mit extrem leistungsstarker Hardware zu einem Engineered System gebündelt, der Exalytics-Maschine. Dies hat neben der besseren Performance auch noch niedrigere Total Costs of Ownership zur Folge, da das komplette System zusammen gewartet und gepatcht wird. Daneben wird auch noch das operative Risiko

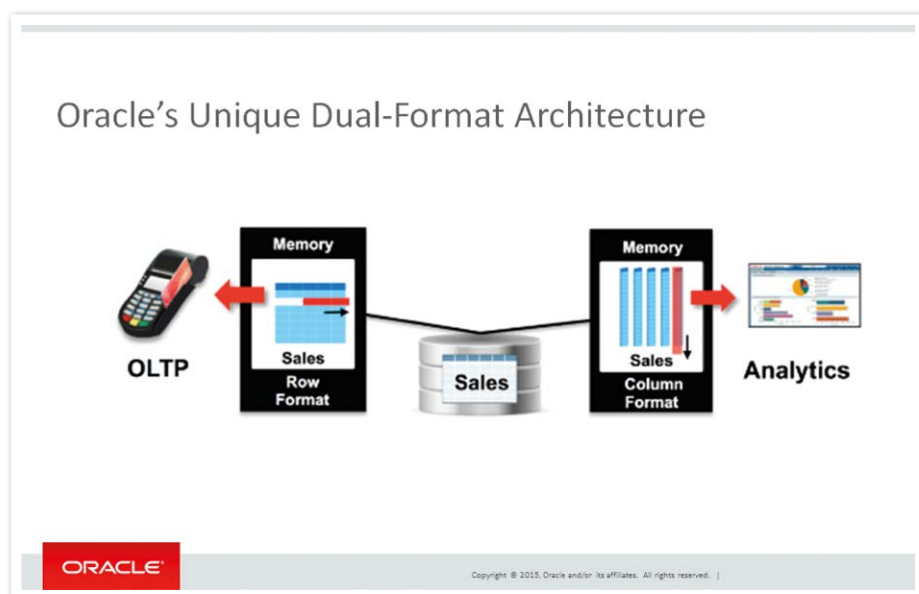


Abbildung 1: Dual-Format Architecture

deutlich gesenkt, da alle Komponenten aufeinander abgestimmt sind.

Auf diesem für die analytischen Anforderungen optimierten System kann nun neben der eigentlichen BI-Software (Business Intelligence Suite Foundation Edition) auch die Datenbank 12c mit der Database-In-Memory-Option installiert sein. Außerdem kann Oracle auch für die Enterprise-Performance-Management-Software verwendet werden. Die Exalytics kann somit nicht nur alle analytischen, sondern auch alle Anforderungen rund um das Thema „Enterprise Performance Management“ bedienen. Dazu zählen Themen wie „Ertragsmanagement“, „Umsatzmanagement“, „Nachfrage-Forecast“, „Inventarmanagement“, „Preisoptimierung“, „Profitabilitätsmanagement“ und „rollierende Forecasts“. Die Exalytics-Maschine kann wahlweise mit weiteren Engineered Systems kombiniert werden. Exadata ist für die Haltung von großen Datenbanken und Exalogic für den Betrieb der Middleware optimiert. Es kann Designgründe dafür geben, dass eine Kombination dieser verschiedenen Engineered Systems bedeutende Vorteile für das Unternehmen bringt.

Die Nutzen-Perspektive für den Anwender

Für den Fachanwender ändert sich mit Exalytics nichts an der gewohnten Arbeitsumgebung (identische Software), das BI-System ist jedoch deutlich schneller als zuvor. Die Benutzer-Akzeptanz verbessert sich stark, denn eine hohe Analyse-Performance ist das wichtigste Akzeptanz-Kriterium eines BI-Systems, wenn Informationsrelevanz und Datenqualität als gegeben gelten. Geschwindigkeit hat allerdings per se noch keinen Wert. Konkreter geschäftlicher Nutzen aufgrund der Performance entsteht aber bei folgenden Anforderungen aus der BI-Praxis:

- Bisher dauerte eine Ad-hoc-Analyse eine halbe Stunde und länger. Nun benötigt sie nur noch einen Bruchteil der Zeit, es sind also zum einen schneller Ergebnisse da, zum anderen besteht für den Fachanwender nun die Möglichkeit, tiefergehend in die Analyse einzusteigen und so neue, zusätzliche Erkenntnisse zu gewinnen.
- Bislang war die Analyse-Verknüpfung so komplex, dass selbst nach einem Tag

Laufzeit die Berechnung noch nicht beendet war. Der Prozess musste abgebrochen werden. Alternativ war diese Analyse schrittweise manuell durchzuführen, indem Einzel-Ergebnisse ermittelt und miteinander verbunden wurden. Wenn eine solche komplexe Analyse mit der Exalytics nun zuverlässig in beispielsweise zwei Stunden ohne jeglichen manuellen Aufwand durchläuft, hat das einen hohen Geschäftswert. Mit der Exalytics sind Analysen durchführbar, die vorher aufgrund ihrer Komplexität nicht umzusetzen waren.

- Das BI-System skaliert aufgrund der Hardware-Ausstattung der Exalytics sehr gut und kann deutlich mehr internen Benutzern zur Verfügung gestellt werden, was die Transparenz im Unternehmen verbessert.
- Erst durch die sekundenschnelle Analyse-Antwort werden komplett neue Geschäftsmodelle im Internet abbildbar, zum Beispiel indem in einem Webshop eine Online-Analyse dem Kunden direkte Berechnungsergebnisse bezogen auf seine aktuelle Auswahl zeigt: Entweder ist diese Analyse sofort da (Realtime) oder das ganze Angebot ergibt keinen Sinn, da selbst fünf Sekunden Wartezeit im eCommerce-Umfeld aus Kundensicht unakzeptabel sind.
- BI-Angebote in der Cloud nehmen zu, flankiert von klar formulierten Service Level Agreements (SLA). BI-Cloud-Anbieter können sich mit der Exalytics Wettbewerbsvorteile verschaffen, da sie sehr schnelle Performance zusagen können.
- Im Big-Data-Kontext sind Realtime-Filterungen der eingehenden Massendaten wichtig, um schnell die relevanten Daten zu identifizieren und diese direkt weiter zu analysieren. Hier ist höchste Performance im Datenstrom ein Muss.

Die Nutzen-Perspektive für die IT

In Data-Warehouse-Projekten ist die Optimierung der Analyse-Performance ein Dauerthema: Aus Anwendersicht stehen die Analyse-Ergebnisse nie schnell genug zur Verfügung. Exalytics bietet in vielen Fällen einen erheblichen Performance-Schub – ohne oder mit nur minimalem personellen Aufwand. Exalytics ist somit

für das IT-Projektmanagement interessant, wenn die hundert Personentage, die für die DWH-Performance-Optimierung (zusätzliche Materialized Views, Test von OLAP-Würfeln etc.) eingeplant waren, nun entfallen können. Bestehende Oracle-BI-Kunden haben zusätzlich den Vorteil, dass sie mit ihrer bestehenden Oracle-BI-Umgebung nahtlos auf die Exalytics-Plattform umziehen können. Das offerierte „All-in-one-patching“ zum Projektstart der Exalytics und die quartalsweisen Updates bedeuten eine Risiko- und Aufwands-Minimierung im BI-Betrieb.

Monitoring und Überwachung der Exalytics-Maschine mit allen Server-Komponenten erfolgen mittels Cloud Control (von der Hardware bis zum Dashboard). Hier zeigt sich der Vorteil eines optimierten, ausbalancierten Gesamtsystems, die Komplexität wird wesentlich leichter beherrschbar; zwanzig bis vierzig Prozent Aufwandsersparnis gegenüber traditioneller Vorgehensweise sind nachweisbar.

Hauptmerkmale von In-Memory Business Intelligence

Die Exalytics-In-Memory-Technologie erlaubt Terabytes von Daten direkt im Arbeitsspeicher, um sie dort schnell zu berechnen und zu analysieren. Darauf basierend wird In-Memory Business Intelligence dann durch In-Memory für Exalytics und In-Memory Essbase bereitgestellt. Diese Datenmanagement-Systeme können in unterschiedlichen Ausprägungen, je nach Datenvolumen und -nutzung, durch vier unterschiedliche Techniken umgesetzt werden. Im Ergebnis erwarten Arbeitsgruppen, Abteilungen oder ganze Unternehmen Business Intelligence mit extremer Performance, um ihre Anforderungen bestmöglich zu unterstützen:

- *In-Memory Datenreplikation*
Viele unternehmensweite Business-Intelligence-Implementierungen können inklusive vordefinierten Oracle-Business-Intelligence-Applikationen komplett In-Memory gehalten werden. Dazu werden die 12c-Datenbank mit der In-Memory-Option direkt auf der Exalytics installiert und alle Daten des Data Warehouse im Speicher vorgehalten. Dieser Mechanismus erlaubt die In-Memory-Verarbeitung für das gesamte Spektrum aller Business-Intelligence-Analysen und weist den höchsten Grad der Flexibilität auf.

Oracle Exalytics X Series

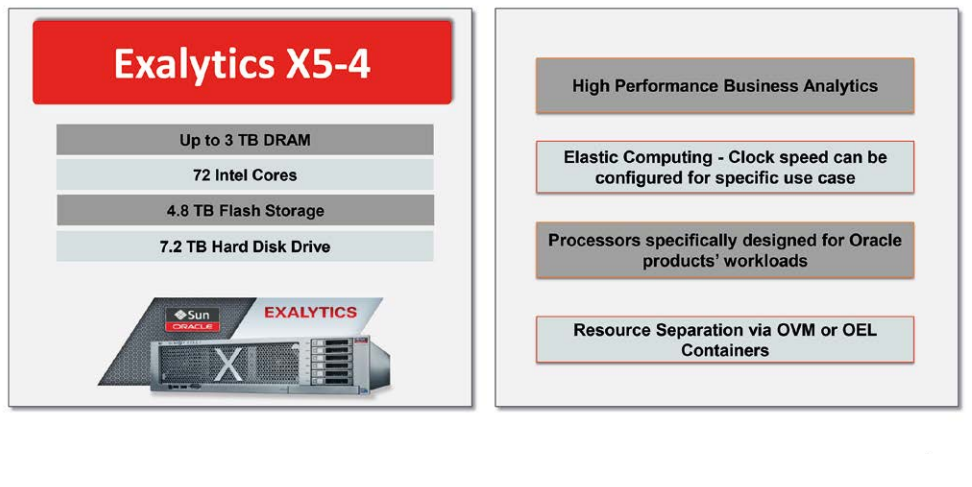


Abbildung 2: Oracle Exalytics X5-4 Maschine

- In-Memory Adaptive Data Marts**
 Die meisten Business-Intelligence-Applikationen haben Auslastungsmuster, die auf die Sammlung von spezifischen, besonders relevanten Daten aus dem Unternehmens-Data-Warehouse fokussieren. In diesem Fall ist es sinnvoll, aus diesen relevanten Daten einen Data Mart In-Memory zu definieren. Diese Vorgehensweise kann Performance-Steigerungen um den Faktor zwanzig realisieren.
- In-Memory Intelligent Result Cache**
 Exalytics Result Cache ist ein vollständig wiederverwendbarer In-Memory Cache, der mit Ergebnissen von bereits getätigten logischen Abfragen quasi automatisch befüllt wird. Jedes Abfrage-Ergebnis wird im Ergebnis-Cache gehalten und bedient ähnliche Abfragen, auch wenn nur eine Untermenge der gesamten Datenmenge benötigt wird. Zur Umsetzung einer optimalen Abfragebeschleunigung stellt Exalytics Werkzeuge zur Verfügung, um die Nutzung von Tabellen zu analysieren, die Notwendigkeit zur Vorab-Befüllung von Ergebnis-Caches zu identifizieren und automatisch auszuführen. Gerade die Vorab-Befüllung von Ergebnis-Caches („Cache Seeding“) erlaubt optimales Antwortverhalten für Abfragen zur Laufzeit.
- In-Memory Cubes**
 Essbase liefert mit seiner In-Memory-Optimierung für Exalytics eine weitere Opti-

on zur Beschleunigung von Abfragen auf festgelegten Geschäftsbereichen (relevante Ausschnitte vom Gesamt-Datenmodell, die speziellen analytischen Anforderungen gerecht werden). Der Unterschied zu anderen Formen von In-Memory-Techniken ist der Umstand, dass hier auch zurückgeschrieben werden kann. Der BI Server für Exalytics baut diese Cubes auf, indem Daten über die semantische Schicht extrahiert werden. Notwendig und hilfreich sind diese multidimensionalen Cubes hauptsächlich für Advanced Scenario Modelling und „What If“-Analysen.

Es gibt also unterschiedliche Design-Möglichkeiten, Daten optimiert für Business Intelligence zur Verfügung zu stellen. Dabei unterstützen die Werkzeuge der Exalytics-Maschine und automatisieren Prozesse weitestgehend.

Beispiel für eine Konfiguration

Eine Exalytics X5-4 verfügt über die Fähigkeit, Business Intelligence durch In-Memory-Technologie zu beschleunigen und Daten in In-Memory Caches zu speichern. Außerdem kann eine Oracle Essbase als multidimensionale Datenbank gemäß ihrer Spezifikation beschleunigt werden, da diese ebenfalls für die Nutzung auf der Oracle Exalytics In-Memory Machine optimiert ist. Hinzu kommt die Fähigkeit, In-Memory zu cachen. Ferner kann auch die Pixel-perfekte Berichtskomponente BI Publisher auf der Exalytics-Maschine beschleunigt werden.

Die Exalytics Maschine unterstützt Business Intelligence mit 72 Cores, 2,6 GHz Intel Xeon Prozessor E7-8895 v3, speziell für Oracle hergestellt, mit bis zu 3 TB Speicher, 4,8 TB Solid-State Drive NVM Express (NVMe) Flash und 7,2 TB Festplatten-Laufwerken. Die Exalytics X5-4 liefert die perfekte Balance von Computing Power, Speicher sowie I/O Footprint (siehe Abbildung 2).

Zu der Maschine gehören im Kontext von Business Intelligence folgende Softwarepakete:

- Business Intelligence Suite Foundation wEdition**
 Dieses Softwarepaket liefert umfassende und vollständige Business-Intelligence-Unterstützung inklusive Unternehmensreporting, Dashboards, Ad-hoc-Analysen, Essbase (multi-dimensionaler OLAP Server), Scorecarding und Advanced Analytics auf einer integrierten Plattform.
- Database-12c-In-Memory-Option (optional)**
 Bei diesem Ansatz werden alle zu analysierenden Daten direkt auf der Exalytics im Speicher gehalten, um optimale Performance zu gewährleisten.

Fazit

Oracle Business Intelligence In-Memory ist im Kern die Fähigkeit des BI Servers, Daten aus einer spaltenbasierten Datenhaltung zu lesen. Wirkliche Effizienz-Steigerungen sind aber nur über die Nutzung eines Engineered Systems, der Exalytics-Maschine, sinnvoll zu realisieren, da hier diverse wichtige Mechanismen und Techniken zur Verfügung stehen, die, je nach Datenvolumen und Präferenz, die Daten optimal und weitestgehend automatisch aufbereiten und auch synchronisiert halten. Dieses Arbeitersparnis ist mit dem Einsatz von Business Intelligence und eines In-Memory-Datenbank-Management-Systems allein nicht zu realisieren.

Annett Thurm-Meyer
 annett-thurm-meyer@oracle.com