

Kann die Exadata Database Machine so wichtig sein und den geschäftlichen Erfolg eines Unternehmens beeinflussen?

Business-Faktor Exadata im Data Warehouse

Alfred Schlaucher, ORACLE Deutschland B.V. & Co. KG

Es gibt Unternehmen, die Data-Warehouse-Systeme strategisch nutzen. Sie sehen einen Zusammenhang zwischen wirtschaftlichem Erfolg und dem Einsatz ihres Data-Warehouse-Systems. Wenn es dann im Data-Warehouse-System nicht mehr „rund“ läuft, so hat das auch spürbare Auswirkungen auf den Geschäftserfolg. In solchen Krisen-Situationen ist es nicht schwer, diese Unternehmen vom Nutzen einer extrem schnellen Exadata-Maschine zu überzeugen.

Um die Nutzenpotenziale einer Exadata-Maschine im Data Warehouse zu erkennen, sollte man sich zunächst die besonderen Herausforderungen ansehen. Ein Data Warehouse ist heute eines der wichtigsten IT-Systeme. Sein Erfolg führte in den letzten fünfzehn Jahren zu einer sehr intensiven Nutzung und zu einer engen Verflechtung mit anderen IT-Systemen. Außerdem verwendet man es zunehmend operativ, was zu mehr Detailtiefe der Daten führt. Das Ergebnis sind riesige Datenmengen, die aufgrund der besonderen Anforderungen an Data-Warehouse-Systeme immer wieder in aufwändigen Batch-Läufen umgewälzt werden.

Betrachtet man die Exadata-Einsätze der letzten Jahre etwas genauer, so sind es Batch-Läufe, die als Hauptmotivation für den Einsatz gegolten haben. Es waren Auswertungen und Berechnungen mit großen Datenbeständen, die entweder nicht rechtzeitig fertig wurden oder aufgrund der langen Laufzeit nur begrenzt häufig und wohl dosiert stattfinden konnten. Diese Engpässe waren schließlich für jedermann und auch im Management sichtbar.

Unternehmen haben gelernt, aus Informationen Geld zu machen, etwa durch Aktien-Orders, Kontostände über Bonus-Punkte, weltweite Währungsbewegungen oder einfach nur Statistik-Informationen als Grundlage für Investitionsentscheidungen. Detailgenaue Auswertungen lassen sich verkaufen; man kann sie zur Steuerung von Kampagnen verwenden oder sie bilden die Grundlage für Ratings, die bares Geld wert sein können. Der Ausfall einer verkaufbaren Auswertung oder stunden- und tagelange Laufzeiten sind durch entgangene Einnahmen beziehungsweise eingeschränkte Kampagnenhäufigkeit monetär messbar.

Noch mehr „Blech“ im Rechenzentrum

Die genannten Data-Warehouse-Anwendungen benötigten unter ihrer alten Hardware Laufzeiten von Stunden und sogar Tagen. Um die Laufzeiten auf ein spürbar sinnvolles Maß zu reduzieren, wäre sehr viel mehr traditionelle Hardware nötig, wobei es nicht um günstige Standard-Hardware, sondern um ausfallsichere Hochleistungssysteme

geht, die die Budgets extrem belasten.

Exadata-Kunden hatten vor ihrer Entscheidung genau diese Überlegungen vor Augen: Kann durch den bezahlbaren Einsatz extremer Performance (in Form von Exadata) ein deutlich erhöhter Nutzen in den Ablaufprozessen rund um das Data Warehouse gewonnen werden? Waren rechnende Strategen an dem Auswahlprozess beteiligt, so fiel die Entscheidung meist zugunsten von Exadata aus. In einem Fall hat allein die Verdoppelung der in einem Jahr durchführbaren Kunden-Kampagnen die Investitionskosten in eine Exadata-Maschine bereits nach neun Monaten komplett ausgeglichen. Anmerkung: Die genannten Einsatzbeispiele sind anonym. Kontakte zu den jeweiligen Firmen sind jedoch über den Autor möglich.

Zentrale Dienstleistungen (SAN) werden unbezahlbar

Gerade große Data-Warehouse-Systeme kann man heute immer weniger mit zentral verwaltetem Storage und



Abbildung 1: Die Durchführung von Kampagnen kann aufgrund der vielen Ablaufschritte sehr zeitaufwändig sein und mehrere Wochen dauern. Man muss Fehler und damit Zeit für Wiederholungen mit einkalkulieren.

Netzwerken über Data-Service-Center bedienen. Sind die Umlage-Kosten für OLTP-Systeme noch überschaubar, sprengen sie bei mehrstelligen Terabyte-Speicher-Anforderungen die Budgets. Wenn der weitaus größte Anteil der Warehouse-Kosten in die zentrale Storage-Bereitstellung fließt, dann läuft etwas falsch. Solche Kosten steigen kontinuierlich, weil die Datenmengen in Data-Warehouse-Systemen immer schneller wachsen. Skalierung und Datenwachstum ist dann kein technisches, sondern ein massives finanzielles Thema. Für einen großen deutschen Finanzdienstleister war genau das der wichtigste Grund für die Einführung von Exadata. Seit Jahren hatte er die Speicherkosten für seine Data-Warehouse-Umgebung im Blick. Als die Marke von 75 Prozent anteiliger Kosten überschritten war, wurde Exadata eingeführt. Die Speicherkosten sanken danach um zwei Drittel.

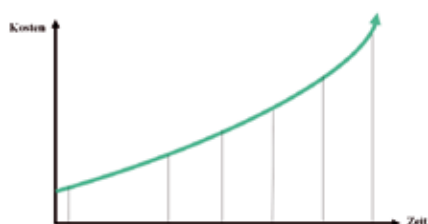


Abbildung 2: Kosten für Speicher im Data Warehouse wachsen nicht linear, sondern exponentiell

Hinzu kommt die fehlende Flexibilität von zentralen Betreibern: Braucht man im Data Warehouse mehr Speicherkapazität, ist diese zentral zu beantragen und genehmigen. Sind dann zur Bedienung der Anforderungen neue Platten erforderlich, benötigt man einen Beschaffungsantrag und hofft, dass nicht ein aufwändiger Ausschreibungsprozess anfällt. Ist in der Data-Warehouse-Datenbank ein neuer Patch von einer bestimmten Betriebssystem-Version abhängig, so muss erst auf ein Betriebssystem-Upgrade gewartet werden. Sind dann noch andere Nicht-Datenbank-Anwendungen auf den Servern und dem Storage, so muss man noch auf die Verträglichkeit des BS-Upgrade und dieser Fremdanwendungen warten. Das bremst jede Agilität aus.

Bleiben noch die technischen Unzulänglichkeiten: OLTP-Systeme sind idealerweise Prozess-orientiert, Warehouse-Systeme Daten-orientiert. Im Ergebnis sind Daten-orientierte und auf hohe Datenmengen ausgerichtete Anwendungen wie das Data Warehouse durch das Storage-System mit extrem hohen IO-Durchsatzraten und kräftigen Datenleitungen zu bedienen. Der zentrale Rechenzentrumsbetrieb unterstützt jedoch primär OLTP-Systeme, die in der Regel Datenzugriffe mit überschaubaren Datenmengen und oft nur auf einzelne Sätze erfordern. Dieser Umstand ist als Performance-Killer für Data-Warehouse-Systeme bekannt. Pauschal kann man daher fordern: Für das Data Warehouse privaten Storage und ein privates Netzwerk bereitstellen. Eine Exadata-Maschine passt aus folgenden Gründen ideal für den Data-Warehouse-Betrieb:

- Hardware und Software sind über eine Stelle administrierbar. Das kann auch der Datenbank-Administrator erledigen. Dies senkt Betriebskosten, und IT-technische Abläufe werden schneller und einfacher. Das war einer der Hauptgründe für eine deutsche Direktbank, die Database Machine einzuführen. Bei diesem Beispiel ist gut erkennbar, dass es nicht immer die sehr großen Firmen mit großen IT-Abteilungen sind, die sich mit Exadata beschäftigen. Das „All-In-One-Konzept“ der Database Machine scheint gerade für Unternehmen mit mittelgroßen IT-Abteilungen interessant zu sein, die den administrativen Aufwand zum Betreiben der IT gering halten wollen.

Newsticker

SAP-Anwendungen für Oracle Exadata Database Machine zertifiziert

Mit der Zertifizierung können nun auf den Oracle Exadata Database Machines X2-2 und X2-8 SAP-Applikationen betrieben werden, die auf SAP NetWeaver 7.x für den Einsatz mit Oracle 11g R2 zertifiziert sind. Dazu zählen SAP ERP und SAP Business Warehouse (BW). SAP-Kunden, die Rund-um-die-Uhr-Verfügbarkeit ihrer Anwendungen benötigen, profitieren zudem von der Maximum Availability Architecture. Exadata ist besonders für SAP-Anwender interessant, die ihre Hardware-Plattform erneuern wollen.

- Das Datenmengen-Problem erledigt sich fast von selbst: Ein Full-Rack mit schnellen SAS-Platten speichert unkomprimiert 28 TB. Bei dem Einsatz von HybridColumnarCompression schaffen auch SAS-Platten weit über 100 TB an Benutzerdaten. Damit ist in den meisten Fällen die Forderung nach planbarer Skalierung erfüllt. Ein Telco-Unternehmen in der Türkei konnte sein 120 TB großes Warehouse komplett in einem mit SAS-Platten bestückten Full Rack speichern. Ein deutsches Unternehmen aus der gleichen Branche reduzierte den belegten Platz von zuvor 33 TB auf etwa 8 TB auf der Exadata-Maschine.
- Die Exadata-Maschine liefert über ihre technischen Mittel wie Smart Scan und die Verwendung von Infiniband einen idealen IO-Durchsatz für das Data Warehouse. 25 GB pro

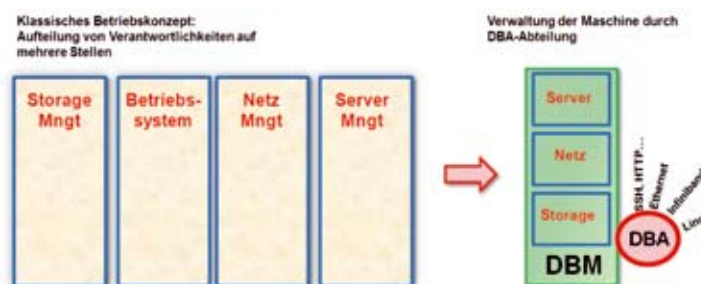


Abbildung 3: Bis zu 70 Prozent weniger Betriebskosten für das Data Warehouse bei einer deutschen Direktbank

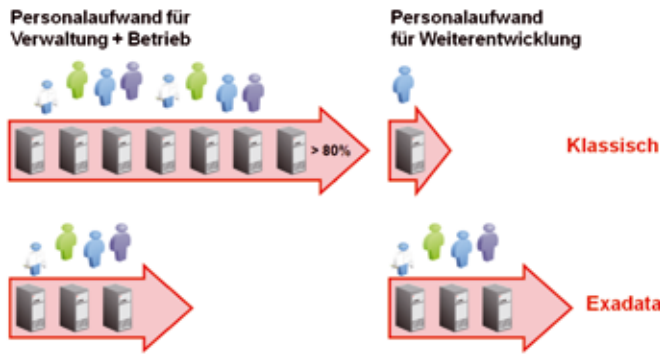


Abbildung 4: Mitarbeiterressourcen für echte Neu- und Weiterentwicklung gewinnen; Anwendungsstau abbauen



Abbildung 5: Schlanke, integrierte Data-Warehouse-Architektur für kürzere Latenzzeiten

Sekunde können Mengenbewegende Batchläufe ideal ausnutzen. Bei typischen Tabellenkonstruktionen mit Partitioning und dem Wunsch, sehr stark zu parallelisieren, spielt die Database Machine ihre eigentlichen Stärken hervorragend aus.

Auch für die Neu- und Weiterentwicklung von Warehouse-Systemen liefert die Database Machine besondere Vorteile. Die Entwicklungszeiten werden kürzer und die Ergebnisse qualitativ besser.

In Data-Warehouse-Umgebungen ist das A und O die Effizienz, mit der Daten über längere Wegstrecken hinweg geprüft und transformiert werden. Entwickler müssen im Data Warehouse besonders häufig auch mit großen Datenmengen testen, um den späteren Anforderungen in der Produktion möglichst nahe zu kommen. Die hohe Performance beschleunigt diesen Prozess. Tests können unter Echtbedingungen mit größeren Datenmengen stattfinden. Es entstehen auch Erfahrungswerte über potentielle Laufzeiten. Dadurch

kann man bereits in Testphasen sehr viel mehr Fehler finden, die später in der Produktion nicht mehr auftreten.

Diese Erfahrung machte ein deutscher IT-Dienstleister, der eine halbe Database Machine für Entwicklungsprojekte nutzt. Die Software-Entwickler äußern sich extrem positiv über diese Bedingungen. Das Konzept, Exadata auch in der Entwicklung einzusetzen, hat sich bewährt, sodass man sogar an einen Ausbau des Systems denkt.

Freischaufeln und Zeit für Neues gewinnen

Gleich bei mehreren Exadata-Einsätzen tritt ein besonders interessanter Effekt auf. Die Effektivität scheint bei manchen, vor allem größeren Data-Warehouse-Systemen mit der Zeit verlorengegangen zu sein. Über Jahre hat man die Systeme weiterentwickelt und bei neuen Benutzeranforderungen immer neue Funktionalität entwickelt. Oft genug wurden dabei Grundsätze beiseitegelassen. Anstatt über bestehende In-

formationen Synergien zu schaffen, hat man immer wieder komplett neue ETL-Strecken implementiert. Redundanzen und unnötige Komplexität schlichen sich ein. Gleichzeitig wurden mehr und mehr Vorsysteme angeschlossen. Die Datenmengen wuchsen rasant. Solche Systeme lassen sich dann nur noch mit Mehraufwand verwalten. Ein Mehraufwand, der unnötig Weiterentwicklungs-Ressourcen bindet. Dieser Sog kann bis zum Stillstand führen und verschlingt gewaltige Summen, ohne dass man genau erkennen kann wofür. Der Autor hat in seiner Praxis einige solcher Systeme vor allem im zweistelligen TB-Bereich beobachten können.

Eine Performance-Maschine wie Exadata kann auf einen Schlag so viele Ressourcen liefern, dass die IT-Abteilung wieder Raum für eine echte Weiterentwicklung beziehungsweise sogar für ein komplettes Re-Design des Systems gewinnt. Diese Performance kann allerdings schlechtes Design nicht korrigieren. Es werden jedoch Berichte in kürzerer Zeit fertig oder sind plötzlich vollständig; andere Berichte können überhaupt zum ersten Mal geliefert werden, während bis dahin die Generierungsjobs immer abgebrochen waren; es entfällt das permanente Kontrollieren der Ressourcen und Log-Files etc.; Mitarbeiter gewinnen Zeit, die sie für eine echte Mehrwertleistung einbringen können und nicht zum Kaschieren alter Sünden.

Nach der Einführung der allerersten Database Machine in Deutschland konnten die Administratoren eine erstaunliche Feststellung treffen. In dem Unternehmen gibt es Mitarbeiter, die direkt mit SQL auf Daten in der Datenbank zugreifen, also ohne SQL-generierende Werkzeuge. Nachdem die Anwender merkten, dass die Antworten in viel kürzerer Zeit an den Bildschirm geliefert wurden (bis dahin war man an Antwortzeiten in der Länge einer Mittagspause gewöhnt), formulierten sie mit Exadata zunehmend komplexere und deutlich mehr Abfragen. Bis dahin hatten sich die Anwender aufgrund der Ressourcen-Engpässe der alten Umgebung selbst beschränkt. Welchen zusätzlichen Nutzen sie aus den Daten ziehen, können sie erst jetzt nach der

Einführung der Database Machine erfahren.

Ein letzter Pluspunkt soll noch erwähnt werden: Gerade große Unternehmen betreiben auch heute noch eine verteilte Data-Warehouse-Landschaft mit zentralem Data Warehouse, aber vielen dezentralen Data Marts. Das dadurch erzwungene ständige Kopieren zu den Data Marts kostet Zeit und Ressourcen. Andererseits traut man sich aufgrund der befürchteten Performance-Verluste nicht, die Data Marts mit auf der zentralen Warehouse-Plattform zu betreiben. Exadata macht dies möglich. Sowohl das zentrale Data Warehouse als auch viele Data Marts (und damit alle System-Schichten) kann man auf einer Maschine betreiben.

Fazit

Noch einmal zusammengefasst eine Liste der relevanten Business-Faktoren:

- Aufgrund der extremen Performance von Exadata sind mehr und zusätz-

liche geschäftliche Aktivitäten möglich. Vermehrtes Geschäft bringt mehr Umsatz.

- Exadata bremst die Kostenschraube in dem besonders speicherhungrigen Data-Warehouse-Segment. Hier sollten alle Faktoren bei der Kalkulation berücksichtigt werden, also auch Personalkosten und ermittelbare Kosten für Reibungsverluste in den Verwaltungsprozessen.
- Aufgrund des hohen Ressourcen-Angebots schafft Exadata Freiräume bei Planung, Ausbau und Erneuerung des Data-Warehouse-Systems.
- Das „All-In-One-Konzept“ von Exadata liefert Agilitätsvorteile gegenüber einer trägen, zentralen Rechenzentrumsverwaltung. Schnelleres Reagieren in Richtung Fachanwender wird mit einem schlanken Betrieb möglich.
- Exadata führt zu schnellerem Entwickeln mit qualitativ besseren Ergebnissen.
- Endanwender können ihre Erwartungen und Informationserforder-

nisse besser verwirklichen und formulieren mehr und komplexere Abfragen.

- Exadata erlaubt aufgrund des extrem hohen Ressourcen-Angebots integrierte und schlanke Data-Warehouse-Architekturen, bei denen alle System-Schichten auf einer Plattform betrieben werden können.

Alfred Schlaucher
ORACLE Deutschland B.V. & Co. KG
alfred.schlaucher@oracle.com



Natürlich können Sie auch nach Amerika rudern ...

... aber warum sich das Leben unnötig schwer machen? Wir sagen: Am besten erreicht man sein Ziel direkt und komfortabel – das gilt für Atlantiküberquerungen genauso wie für Datenbankentwicklung und -administration. Allen Unternehmen, die mit Oracle™ Datenbanken arbeiten, bietet KeepTool mit Hora ein mächtiges Werkzeug: intuitiv, zuverlässig und universell einsetzbar; unterstützt durch kostenlosen und schnellen Support.

Ohne Umwege – direkt mit KeepTool.

www.keeptool.com

keeptool

ORACLE Gold Partner