

Der Artikel zeigt die Konzepte und Erfahrungen eines Telco-Unternehmens, das Exadata als neue strategische Oracle-Datenbank-Server-Plattform gewählt hat. Im Detail geht es um die Umsetzung des Konsolidierungsvorhabens einer größeren Anzahl von Datenbanken auf die neue Umgebung.

Datenbank-Konsolidierung auf der Exdata Database Machine

Konrad Häfeli, Trivadis AG

Im Einzelnen waren bei dem großen Unternehmen in der Telco-Branche folgende Beweggründe ausschlaggebend:

- Geplanter Ersatz der Storage-Systeme
- Die bestehenden Standby-Systeme waren nicht gleichwertig den Primary-Systemen
- Die Standby-Systeme waren nicht mehr ausreichend
- Geplanter Einsatz von Oracle 11g R2
- die Bestehende Backup-Infrastruktur und -Abläufe waren zu komplex
- Der Energieverbrauch war heutigen Standards nicht mehr angemessen
- Mehrere Plattform-Konsolidierungen standen an
- Es war eine Lizenz-Kostensenkung erwünscht

Das Standard-Setup für Exadata von Oracle Advanced Customer Services (ACS) ist bekannt und die Migration einer Datenbank mit Oracle Golden Gate (Near Zero Downtime Migration) auf die Database Machine hat der Autor auf der DOAG 2011 Konferenz in einem Vortrag präsentiert (siehe http://www.trivadis.com/uploads/tx_cabag-downloadarea/Konrad_Haefeli_Oracle_GoldenGate_DB-Migration-auf-Exadata-mit-Fallback_NOV2010.pdf).

Ist-Situation

Das zu konsolidierende Mengengerüst besteht aus etwa 50 Datenbanken unterschiedlicher Größe (100 GB bis mehrere TB) und mit verschiedenen Versionen (hauptsächlich Oracle 10.2.0.4),

meist Single Instanz, aber auch RAC-Datenbanken. Die Infrastruktur wird applikationsbezogen von zwei Teams betrieben, die im Datenbank-Bereich eng zusammenarbeiten. Die Verfügbarkeit erfolgt meist über Failover-Cluster-Konfiguration, wobei ein Stretched Veritas Cluster über zwei nahegelegene (3 km) Datacenter im Einsatz ist. Die bestehende RAC-Konfiguration ist mittels Dataguard zu einem Maximum-Availability-Architecture-System (MAA) zusammengefasst.

Das Sizing der Exadata-Plattform wurde mittels Trivadis Configuration Manager (TVD-ConfMan) unterstützt. Über alle bestehenden Systeme wurden Metriken (KPIs) definiert und gemessen und danach mit den Performance-relevanten Zahlen auf der Exadata verglichen. Dabei stellte sich heraus, dass eine Half-Rack-Konfiguration ausreicht, um die bestehende Last zu bewältigen sowie für die geplanten neuen Datenbanken genügend Reserve bereitzustellen. Durch die Verfügbarkeitsanforderungen im Telco Bereich (24*365) war auch eine MAA-Konfiguration (je ein Half-Rack in zwei Datacenter 80 km getrennt) notwendig.

Konzeptüberlegungen und Bewertung

Durch das Zusammenfassen aller Datenbank-Infrastrukturen und deren Betrieb in einem Team (neue horizontale Organisation) können die applikationsbezogenen Unterteilungen der Datenbanken außer Betracht gelassen werden. Es stellt sich aber dennoch die Frage ob eine der beiden Extremvarianten überhaupt untersucht werden soll:

- 1:1-Migration, also für jede alte Datenbank eine neue auf der Exadata
- „All in one“-Datenbank, das heißt alle produktiven Datenbank-Schemata werden in eine Datenbank migriert / konsolidiert

Bei der ersten Variante werden keine Synergien und Möglichkeiten der verfügbaren RAC-Infrastruktur genutzt:

- Service-Hochverfügbarkeit
- Scaleout
- Reduktion der Ressourcen in den Bereichen „Memory“, „Background-Prozesse“ und „Anzahl Tablespace“

Die zweite Variante gibt Problemfelder im Bereich „Abhängigkeiten“ beziehungsweise bei der Abgrenzung der einzelnen bestehenden Applikationsschemas:

- Zum Teil sind viele globale Synonyme im Einsatz (noch mit gleichen Namen über die verschiedenen Applikations-Datenbanken hinweg)
- Überprivilegierte User
- Unterschiedliches Change-Management
- Klumpenrisiko bzw. Seiteneffekte untereinander

Es kristallisierte sich ein Kompromiss heraus, bei dem die Anzahl der Datenbanken (massiv) reduziert werden soll, dies aber in Anbetracht der Gemeinsamkeiten beziehungsweise Unterschiede. Folgende Definition wurde als generelle Grundlage für die Konsolidierung auf eine (Exadata) RAC-Infrastruktur gemacht:

1. Alle Applikationen verbinden sich nur noch über Datenbank-Service-Namen und haben keinen Bezug mehr zur physischen Datenbank beziehungsweise zum Datenbank-Namen.
2. Die Applikationen sind fähig (einerseits mittels Oracle-Net-Konfiguration, andererseits im Handling der Oracle-Verbindungen), einen kurzfristigen Verbindungsabbruch zu überstehen und einen neuen Verbindungsaufbau zu machen.

Die erste Anforderung lässt eine Abstraktion des Datenzugriffs zu, sodass es keine Rolle spielt, wo diese physisch liegen. Die zweite ist nötig, damit die Verschiebung der Datenbank-Services über die RAC-Knoten (und eventuell MAA-Knoten) hinweg keine Applikations-Abstürze hervorruft.

Auf den vier Knoten des Exadata-Half-Rack sollen die bestehenden Daten themenspezifisch auf möglichst wenigen Datenbanken konsolidiert werden. Die Zugriffe darauf sind über Datenbank-Service-Namen abstrahiert.

Alle Themen-Datenbanken haben gestartete Instanzen auf jedem Knoten, wobei die Applikationsservices einer Datenbank immer nur auf einem Knoten laufen und mittels Servicefailover hochverfügbar gehalten werden. Die Services werden so auf die Knoten verteilt, damit eine ausgeglichene Last auf dem System ist.

Vorteile: Durch die Zusammenfassung der Datenbanken können bestehende Datenbank-Links weggelassen werden. Ein spezifischer Datenbank-Service ist hochverfügbar, da ein Failover auf eine schon gestartete Instanz erfolgen kann. Eine mögliche Cache-Fusion-Problematik wird umgangen, da die jeweiligen Verbindungen auf dieselben Daten nicht über verschiedenen Instanzen hinweg gemacht werden (per default kein Scaleout, da die meisten Applikationen nicht „RAC-aware“ sind). Eine gute Lastverteilung ist auf der Granularität Datenbank-Service sehr einfach möglich.

Abbildung 1 zeigt zusammengefasst die redundanten Daten in den Exadata Storage Cells und darüber die vier Datenbank-Knoten mit der Oracle-Grid-

Infrastruktur (Cluster-Schicht), den ASM Instanzen (pro Server) und den verschiedenen thematisch konsolidierten Datenbanken (TH[a-d]). In der Grafik sind es vier Datenbanken, die jeweils eine Instanz (I[1-4]) pro Knoten haben und auf denen die Services (S[a-d]) der ganzen Datenbanken verteilt sind.

Mit dem Einsatz einer Exadata-Plattform ist die beim alten Konzept mittels „Stretched Cluster“ sichergestellte Disaster-Redundanz nicht gewährleistet. Aus diesem Grund wird beim neuen Konzept eine Maximum-Availability-Architecture (MAA) definiert. Diese stellt mittels Dataguard (Redologtransport und Apply) eine datenbankseitige Absicherung dar, welche im Fehlerfall die definierten Services hochfahren und bereitstellen kann. Ebenfalls ist mit Active Dataguard die Möglichkeit gegeben, „readonly“-Services auf den Standby-Datenbanken anzubieten, die das transaktionsorientierte (Primary-) System entlasten können.

Abbildung 2 stellt die beiden Datacenter dar, die im Wechsel oder sogar gemischt die Datenbanken im Primary- und Standby-Modus beherbergen. Eine

Überwachung kann optional aus einem dritten Datacenter gemacht werden, indem ein sogenannter „Dataguard-Observer“ konfiguriert wird, der den Failover der Primary-Site auf die Standby Site im Fehlerfall selbständig durchführt.

Umsetzung der Konsolidierung

Die Zusammenfassung der Datenbanken sollte nach folgenden Kriterien erfolgen:

- Datenzusammengehörigkeit
 - Gleiche Applikationsfamilie
 - Gemeinsame Daten (Ersatz von DB-Links)
- Explizite Datentrennung
 - Trennung wegen nötiger Unabhängigkeit
 - Klumpenrisiko
- Security
 - Sensitive Daten
 - Schema braucht hohe Privilegien
- Verfügbarkeit
 - Multi Instanz Datenbank
 - Single Instanz Datenbank
 - Dataguard
 - DG Protection Level

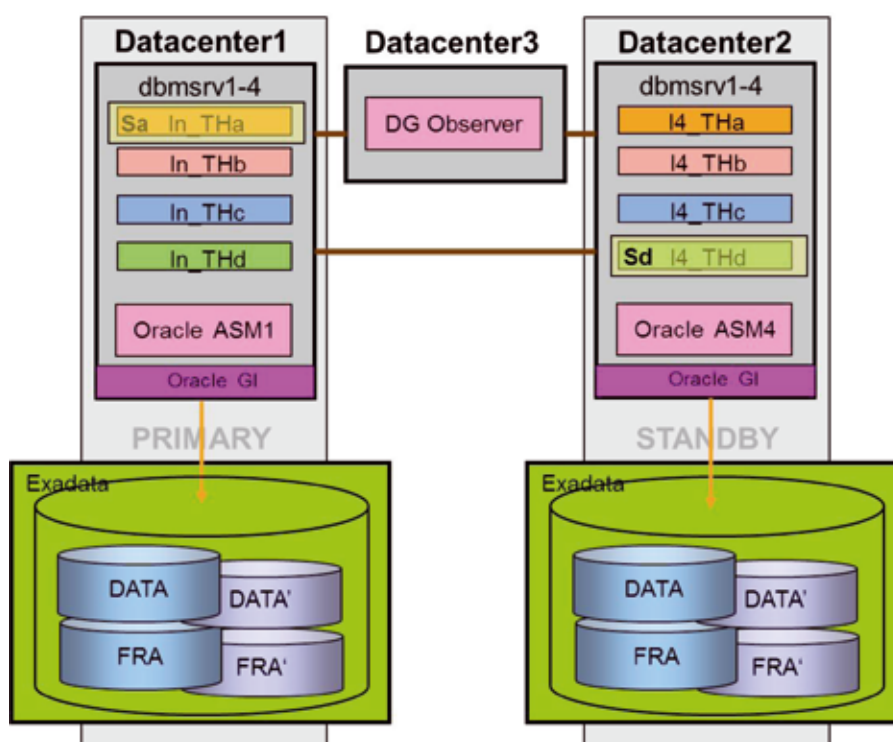


Abbildung 1: Konzeptgrafik Exadata Database-Konsolidierung auf Datenbank-Services

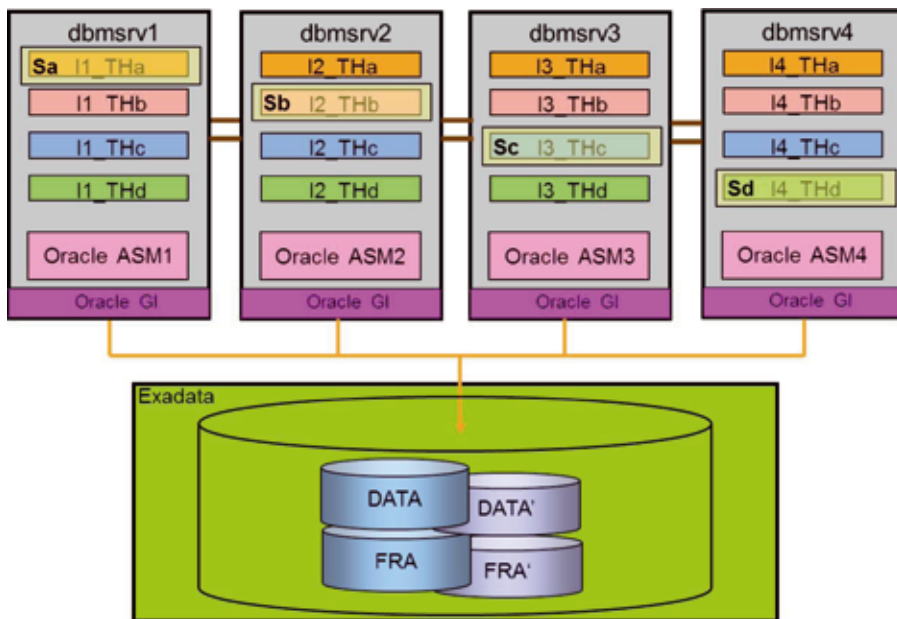


Abbildung 2: Konzeptgrafik Exadata MAA-Konfiguration

- RDBMS Version
 - Major/Minor Releases
 - Patch/Patchset

Für die Umsetzung der Konsolidierung musste zuerst ein Katalog aufgestellt werden, der die Kriterien der Zusammenfassung der Datenbanken auflistet und pro Datenbank bewertet. Diese zuerst als einfache Matrix gedachte Liste, stellte sich dann aber doch als recht aufwändig heraus, weil die verschiedenen Kriterien sich zum Teil ausschlossen, andererseits die verschiedensten Applikations-, Plattform-, Service-, und andere Manager dazu befragt werden mussten.

Die Auswertung der gesammelten Daten, wobei nicht für alle Projekte alle Informationen beschafft werden konnten, ermöglichte folgende Aufteilung und Priorisierung:

- Implementation aller neuen Datenbank-Ressource-Anfragen auf der neuen Plattform
- Migration von ein paar unkritischen Datenbanken als „quick wins“, für die Lieferung von nötigen Erfahrungswerten in der Migration und Betrieb
- Identifikation der zeitlich kritischen Datenbanken und Behandlung mit Projektleiterfokus, damit auch diese termingerecht konsolidiert sind

Erfahrungen

Der hohe Innovations-Rhythmus in der Unternehmung machte es notwendig, die Konsolidierungsplattform sehr schnell umzusetzen. Folgende Aufgaben mussten deshalb zuerst angegangen werden:

- Namenskonzepte für Datenbanken, Instanzen und Services
- Entwicklung von Templatescripts für die Erstellung von standardisierten Datenbanken
- Entwicklung von Scripts für die Konfiguration der Datenbanken und Services im Cluster
- Konzept und Implementation des Ressourcen-Managements (Exadata IORM, DBRM)
- Implementation des Backup-Konzepts
- Implementation des Monitoring-Konzepts

Die Angst der involvierten Applikations-Verantwortlichen und Betreiber vor dem Zusammenlegen ihrer Daten in einer Datenbank war um einiges größer als erwartet. Der Widerstand gegen die Konsolidierung hatte zu einem großen Teil damit zu tun, dass Mehraufwand durch Abgabe von bestehenden Privilegien befürchtet wurde und Reduktion der Verfügbarkeit durch An-

hängigkeiten von weniger stabilen Systemen. Durch ein moderates Auslegen des Konsolidierungsdrucks, indem die heikelsten Datenbanken 1:1 migriert wurden, konnten die Anwender von der guten Funktionalität der Plattform überzeugt werden und sind so eventuell in einer späteren Phase bereit, das Datenbankschema mit anderen zusammen zu legen.

Erst der produktive Betrieb hat die Nützlichkeit der MAA-Konfiguration wirklich aufgezeigt. Durch Switchover von einem Datacenter auf das andere waren große System-Patches ohne Risiko möglich. Auch bei den aufgetretenen technischen Fehlern (Disk-Ausfall, Motherboard-Austausch), die nicht immer ohne Probleme gelöst werden konnten, war das zweite System von Nutzen. Durch die Übernahme der Services war der nötige Hardware-Austausch ohne Last durchführbar.

Wichtig beim Betrieb einer Exadata-Plattform ist eine enge Zusammenarbeit mit dem Oracle Support, der sich speziell bei Exadata-Service-Requests sehr engagiert und zeitnahe Lösungen bringt. Denn eines ist klar, die Plattform, Software wie Hardware, braucht ein intensives Lifecycle-Management und dazu gehört auch das Beheben von unvorhergesehenen Problemen.

Die Kostenfrage wird natürlich bei so einer Konsolidierung immer wieder in den Vordergrund gestellt. Die Erfahrungen im Bereich „Lizenzen“ waren sehr gut, wobei da natürlich auch noch gute Verhandlungen mit Oracle als Early-Adopter der Exadata-Technologie geführt werden konnten.

Konrad Häfeli
Trivadis AG
konrad.haefeli@trivadis.com

