

# No Rocket Science: Migration & Betrieb der ersten deutschen Exadata V2

**Jochen Hinderberger  
Dialog Semiconductor  
Kirchheim/Teck-Nabern**

## **Schlüsselworte:**

Exadata V2, Migration, Flash, Compression, Hochverfügbarkeit

## **Einleitung**

Als weltweit führender Entwickler und Lieferant von kundenspezifischen Halbleiterlösungen für die Sektoren Mobilfunk, Automobilindustrie und industrielle Anwendungen fallen bei Dialog Semiconductor weltweit Daten an unterschiedlichsten Produktionsstandorten an.

Für viele unternehmensweite Anwendungen ist eine große, zentrale Oracle Datenbank die Grundlage bei Dialog Semiconductor.

## **Motivation und Planungsphase**

Im Sommer 2009 begann die Planung zur Ablösung des bestehenden Systems: Eine Single Server Instanz mit Oracle 11.1 auf einer Sun Fire 490 (6 Cores / 12GB RAM für Oracle) in Verbindung mit einem NetApp FAS 3040 Storage. Dieses System war zum damaligen Zeitpunkt an seiner Leistungsgrenze, auch die gestiegenen Anforderungen bezüglich der Verfügbarkeit konnten nicht mehr erfüllt werden.

## **Erwartungen**

Die wichtigsten Ziele für das neue System waren: Skalierbarkeit, Hochverfügbarkeit und guter Support bei möglichst geringen Kosten. Als Alternative wurde über ein Managed Hosting der Datenbank, ein RAC-Eigenbau sowie eine Servererweiterung nachgedacht.

Aufgrund der genannten Punkte fiel die Auswahl Ende September kurz nach deren Release auf die Exadata V2. Die Erwartungen waren, dass durch die Gesamtlösung nicht nur ein korrektes Zusammenspiel der einzelnen Komponenten des RACs gewährleistet ist - durch den Wegfall des NetApp Filers sollten auch Kosten eingespart sowie der I/O Durchsatz erhöht werden.

Als positiver Nebeneffekt verbleibt bei dieser Lösung, im Gegensatz zum Managed Hosting, das vorhandene Datenbankwissen in der Firma und kann weiter vertieft werden.

## **Die Auslieferung**

Vor der Auslieferung wurden peinlichst genau diverse Checklisten bezüglich den hohen Exadata Anforderungen (Netzwerkanschlüsse, verschiedene I/P Bereiche, Umgebungsbedingungen im Serverraum, Transportwege wie z.B. Tür- und Aufzugshöhen,...) besprochen.

Am 18. Januar 2010 stand das V2 Quarter Rack an dem geplanten Platz im Rechenzentrum.

Nach dem Aufbau durch SUN-Techniker, Akklimatisierung und mehrtägigem Check aller einzelnen Komponenten wurde das System Oracle intern an Mitglieder des Installationsteams und der STU Gruppe (Strategische Technische Unterstützung) übergeben. Diese installierten eine Testdatenbank und verifizierten alle Funktionen anhand fester Prüfprotokolle. Nach der Löschung dieser temporären Datenbank wurde die Exadata nach Kundenvorgaben installiert.

Zwei Wochen nach Anlieferung wurde die Exadata V2 an Dialog Semiconductor übergeben.

### **Anfängliche Probleme**

Während der vorbereiteten eigenen Testphase wurden probeweise Schematas erstellt, Daten geladen, Anwendungen verbunden und Exadatafunktionen wie die Verwendung des Storage Indexes, Compression,... ausprobiert. Während dieser Tests, welche immer noch vor Ort von der STU Gruppe unterstützt wurden, fielen immer wieder einzelne Flashkarten in den Zellen aus und mussten durch SUN Techniker getauscht werden. Durch diese Hardware- (und auch vereinzelt Software-) Probleme konnte der Oracle Support und dessen Antwortzeiten ausführlich auf die Probe gestellt werden. Da diese Probleme in der Übergangsphase von SUN zu Oracle auftraten, waren die Supportprozesse noch nicht eingespielt und auch das Exadata System oft noch nicht bekannt. Bei den ersten Hardwarefehlern war es daher nicht möglich die vereinbarten SLAs einzuhalten, da Ersatzteile für die Flashkarten schlicht nicht vorhanden waren. Uns wurde von der STU Gruppe ein Fallmanager gestellt, welcher alle Supportanfragen koordinierte und zusammen mit uns Lösungen erarbeitete.

Die Ursache der Probleme bei den Flashkarten wurde erkannt und behoben, woraufhin alle weltweit verbauten und fehlerhaften Karten ausgetauscht wurden.

Als positiver Nebeneffekt konnten wir während dieser Zeit im laufenden Testbetrieb den Austausch von HW-Komponenten durchspielen, DB Knoten neu starten, usw... Die einzelnen redundanten Systeme, vor allem ASM, haben wie erwartet reagiert: Die Datenbank war stets verfügbar.

Die verbliebenen Softwareprobleme wurden durch einen Patch des gesamten Softwarestacks von Oracle gelöst (OS, Zellensoftware, DB Software, Firmware auf Switches,...). Anschliessend wurde das System von unserer Seite für die Migration freigegeben.

### **Migration und Backuptests**

Die Migration der verschiedenen Applikationen wurde Schema für Schema getestet und vorgenommen. Es kamen zwei verschiedene Techniken zum Einsatz, beide jedoch stets gescriptet und komplett automatisiert:

1) Für „kleinere Applikationen“: Ein bis drei Schematas ohne weitere Abhängigkeiten und maximal ca. 20GB Daten:

- Export des kompletten Schemas inklusive Daten vom alten System in ein dump file
- Kopieren des dump files auf einen Exadata Knoten
- Erstellen des Users und Tablespaces (Konsolidierung auf einen Tablespace pro Schema)
- Import des Schemas mit Remapping der Tablespaces, falls notwendig

→ Diese 4 Schritte wurden komplett gescriptet und mindestens einmal getestet. Somit war eine reproduzierbare Migration gewährleistet und die dafür benötigte Zeit konnte sehr genau abgeschätzt werden.

2) Alle anderen Schematas, z.B. mit partitionierten Tabellen, komplexen Abhängigkeiten und großen Datenmengen:

- Export der Metadaten der betroffenen Schematas vom alten System
- Erstellen der benötigten Tablespaces und User auf der Exadata. Hierbei wurden teilweise angewachsene Tabellen neu partitioniert, bzw. nach ausführlichen Tests die Partitionsstrategie neu festgelegt (s.u.)
- Import der restlichen Metadaten
- Kopieren der Daten mittels Datenbanklinks und bulk inserts:  
Im ersten Schritt wurden innerhalb einer Woche die Daten aller historischen Partitionen (keine Änderungen des Datenbestandes) vom alten System auf die Exadata kopiert. Dabei wurde die Last auf dem (noch produktiven) alten System gering gehalten und es konnten ca. 2.5TB Daten migriert werden.  
Im zweiten Schritt wurden alle Daten der restlichen Tabellen innerhalb einer downtime von ca. 2h der betroffenen Applikationen ebenfalls mittels Datenbanklink migriert. Die dazu benötigten Scripte wurden selbst mittels SQL Statements automatisiert erstellt.

Nach dem Test der Migration und vor der Inbetriebnahme der ersten Applikation auf der Exadata wurde das Backup und Recovery mit dem zu diesem Zeitpunkt vorhandenen Datenbestand durchgeführt. Die Exadata wird mittels eines Legato Backupclients und RMAN gesichert. Es erfolgt eine wöchentliche Vollsicherung mit täglichen Inkrementen.

Es wurden verschiedene Recoveryscenarien durchgespielt, beispielsweise das Löschen von Datenbankfiles direkt im ASM mit nachfolgendem Recovery.

Daraufhin konnten die ausführlich getestete Migration nach und nach durchgeführt werden. Für die „kleineren Applikationen“ betrug die erforderliche downtime nur einige Minuten und konnte je nach Nutzerkreis teilweise kurz in der Mittagspause durchgeführt werden.

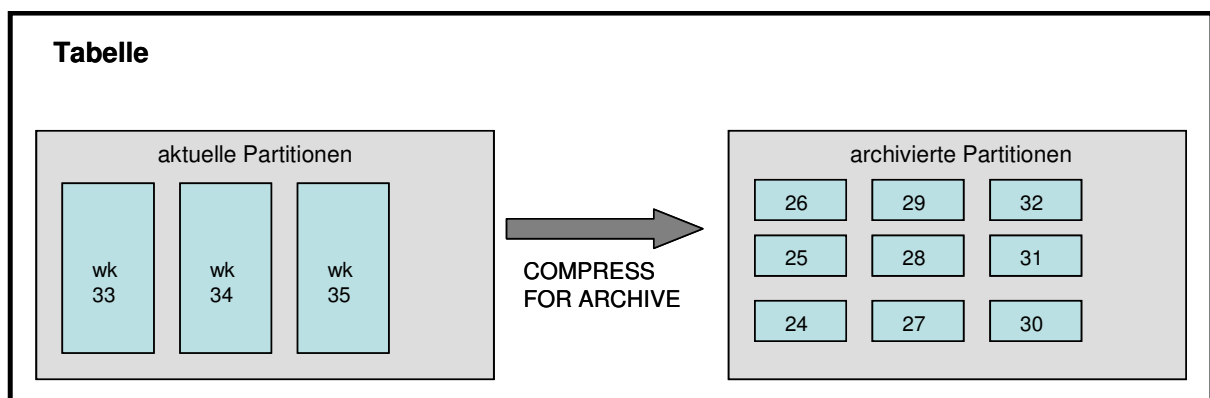
### Applikationsänderungen für Exadata

Prinzipiell müssen keine Änderungen an Applikationen durchgeführt werden, eine Exadata verhält sich wie jede andere Oracle Datenbank.

Im Zuge der Migration wurden von uns aber alle vorhandenen Hints in Prozeduren, Funktionen und Triggern zunächst entfernt und konnten bis auf eine sehr spezielle Ausnahme weggelassen werden.

Nach ausführlichen Tests wurde unsere Partitionsstrategie bei mehreren Tabellen geändert:

Aktuelle Partitionen auf Wochenbasis bleiben unverändert, ältere Partitionen ohne Änderungen werden durch einen wöchentlichen Job mittels Hybrid Columnar Compression (HCC) komprimiert und in einen Archivtablespace verschoben. Dadurch wird massiv Platz eingespart und der I/O Durchsatz erhöht, da die Daten vom Storage bis zur SQL Auswertung in den Knoten komprimiert bleiben.



## **Produktivbetrieb und Fazit**

Seit Ende April ist die Dialog Exadata V2 nach nur 3 Monaten als erste Installation in Deutschland ohne Probleme produktiv im Einsatz.

Drei weitere Applikationen wurden in der Zwischenzeit hinzugefügt oder sind geplant.

Mit der Exadata wurden alle erwarteten Ziele erreicht und zusätzlich durch den internen Storage die Netzwerklast verringert. Durch Verwendung der Komprimierung steht nun noch mehr Platz zur Verfügung als ursprünglich erwartet und zwar ohne Reduzierung des I/Os. Der administrative Aufwand des Gesamtsystems ist äusserst gering.

## **Fakten**

- PL/SQL basierte Datenaggregationsjobs von sehr großen Datenmengen:  
2.5h → 15-20min
- Datenuploads mittels bulk inserts und nachfolgenden Aggregationen mittels PL/SQL Trigger:  
60min → 6min
- Kompression von historischen Partitionen mittels HCC um durchschnittlich mindestens 85%

## **Ausblick**

Weitere Applikationen sollen nach Möglichkeit die Vorteile des bestehenden Exadatasystems nutzen:

Data-Warehouse und BI: Ein neues Projekt, welches derzeit in der Implementierungsphase ist, nutzt die Exadata für ein firmenweites Management-Reporting basierend auf Oracle Business Suite. Hierzu werden Daten aus SAP über den Oracle Warehousebuilder in die Exadata importiert und entsprechende Cubes berechnet.

DBFS: Eine weitere Idee ist das DBFS für das Repository unseres Dokumentenverwaltungssystems zu verwenden um auch an dieser Stelle von der Performance und der Kompression zu profitieren. Zusätzlich ist dadurch ein einfacher Zugriff mittels SQL auf die Dateien möglich, was die Anbindung an existierende, datenbankgestützte Systeme, erleichtern würde.

Kontaktadresse:

Jochen Hinderberger  
Dialog Semiconductor  
Neue Strasse 95  
D-73230 Kirchheim/Teck-Nabern

Telefon: +49 (0) 7021 805-303  
Fax: +49 (0) 7021 805-100  
Mobile +49 (0) 172 7661 949  
E-Mail [jochen.hinderberger@diasemi.com](mailto:jochen.hinderberger@diasemi.com)  
Internet: [www.dialog-semiconductor.com](http://www.dialog-semiconductor.com)