

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

Jochen Hinderberger, Dialog Semiconductor GmbH

17.11.2010

# Dialog Semiconductor

## Führend in der Power Management Integration



- ▶ Weltweit führender Entwickler und Lieferant von kundenspezifischen Halbleiterlösungen im mixed signal Bereich
- ▶ Erfolgreich bei hochvolumigen und schnellwachsenden mobilen Geräten
  - Smartphones
  - 3G / HSDPA Handys
  - Mobile Mediengeräte
- ▶ Bereit für die nächste Generation hochintegrierter Geräte
  - eBook
  - Netbook
  - Digital picture frame
- ▶ Innovative Displaytreiber der nächsten Generation in Entwicklung
  - PM-OLED
  - e-ink™
  - Mirasol™
- ▶ Wachstumsstory und Zulieferer von weltweit führenden Herstellern
- ▶ Dezember 2009: Elektra Company of the Year & Fabless Semiconductor Supplier of the Year

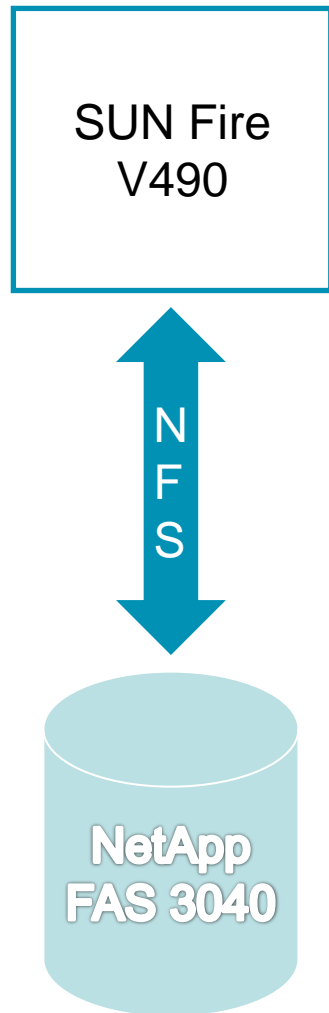




- Motivation und Planungsphase
- Test und Migration
- Fazit und Ausblick

- Motivation und Planungsphase
  - Ehemaliger DB Server
  - Erwartungen
  - Exadata und Alternativen
  - Planungsphase
- Test und Migration
- Fazit und Ausblick





- Vorherige Datenbank:
  - Single Server (Solaris 10 Container)
  - 6 Cores und 12GB RAM für Oracle
  - Wurde zuletzt auf Oracle 11.1 migriert
  - NFS weil geringster Admin Aufwand
  - Nachteil: 2x Protokoll overhead (Ethernet + NFS)
  - Verschiedene NetApp Volumes für oracle home, data, reco
- System war im Herbst 2009 an der Leistungsgrenze

# Was soll die neue HW leisten?



- Skalierbarkeit
- Hochverfügbarkeit
- Einfache Administration
- Guter Support
- Geringe Kosten
- Flexibilität durch eigenes Know-How

# Alles nur Exadata?



- Folgende Lösungen standen zur Auswahl:
  - Neuer, leistungsstarker single Server
  - RAC Eigenbau mit Oracle Advanced Support
  - Exadata Maschine
  - Outsourcing der Datenbank in externes Rechenzentrum
- Exadata konnte als einzige Lösung alle Anforderungen abdecken  
entsprechend hoch waren die Erwartungen an Geschwindigkeit und  
Zuverlässigkeit...



- Ursprünglich Exadata V1 im Gespräch
- Entscheidung fiel Ende September 2009 auf die Exadata V2
- Danach Preis- und Lizenzverhandlungen, internes approval
- Vertrag wurde Ende November unterschrieben
- Noch im Dezember mussten infrastrukturtechnische Fragen mit SUN und Oracle geklärt werden
- Auslieferung wurde für Mitte Januar geplant





- Motivation und Planungsphase
- Test und Migration
  - Auslieferung und Aufbau
  - Testphase und Hardwareprobleme
  - Neue (alte) Backupstrategie
  - Migration
  - Applikationsänderungen für Exadata
- Fazit und Ausblick



- Auslieferung und Aufbau von SUN peinlichst genau überwacht
- 1 Woche Tests durch SUN Techniker vor interner Übergabe an Oracle  
→ hoher Aufwand durch diverse Checklisten
- Mehrtägige Tests durch Oracle vor Übergabe an den Kunden
  - Einbinden ins Kundennetz
  - Installation einer vorläufigen Test DB
- Gemeinsame Installation der Dialog DB





- Einarbeitung in diverse Exadata Tools und Features gemeinsam mit Oracle
  - Überwachung einzelner HW-Komponenten
  - „neue“ RAC Funktionalitäten
  - Testweise connects verschiedener Applikationen
  - Start / stop des gesamten Systems oder einzelnen Komponenten mit Failover
- Tests mit Storage Index, HCC, Flashspeicher,...
  - Neue Explainpläne
  - Tests diverser regelmäßiger Jobs auf Originaldaten
  - Verhalten unter maximaler Last

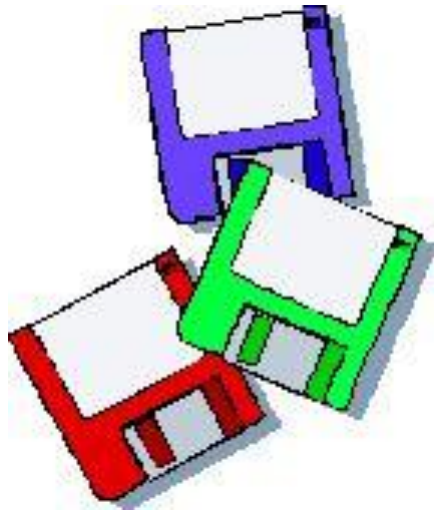
→ Nach wenigen Tagen Probleme mit Flashkarten!



- Probleme:
  - Defekt einer Karte wird nicht durch System erkannt
  - Karte wird weiterhin „normal“ verwendet → deutlicher Performanceeinbruch
  - Manuelle Deaktivierung des gesamten Flashcaches einer Zelle notwendig
- Eigentlicher Defekt nicht Flashspeicher sondern „super cap“
- Karten mehrmals ausgetauscht, andauernde Ausfälle
  - Nach Eskalation zu Günther Stürner:
    - Oracle Fallmanager vor Ort
    - Kompletter Austausch aller Flashkarten
    - SW Update um diverse andere Probleme (unerwarteter Reboot,...) zu lösen
- Fehlersuche und Bereitstellung neu designter Flashkarten dauerte Wochen



# Welche Backupstrategie?



- Früher:  
NetApp Snapshots auf Filesystemebene
- Jetzt:  
RMAN Sicherung mittels Legato Networker auf VTL
- Definition der Backupstrategie und Vorhaltezeiten  
(wöchentliche Fullbackups, tägliche Inkrements)
- Diverse Backuptests (Parallelisierung, Kompression)
- Recovery einer korrupten DB und einzelner DB files



- Komplette Migration sollte gescriptet und nahezu automatisiert sein  
→ Tests pro App. möglich, genaue Ausfallzeiten vorhersagbar
  - Keine generelle Auszeit der gesamten DB und aller Systeme:
    - Sukzessive Migration kleinerer Applikationen (max. 20 GB Daten)
    - Eigene Migrationsstrategie für große Datenmengen erforderlich
- Kein „big bang“ sondern sanfte Migration
- Dadurch keine Probleme mit zu vielen Redologs



- Migration „kleiner Applikationen“ mit vier Scripten:
    - Export des kompletten Schemas + Daten in ein dump file
    - Scp auf einen Exadata Knoten
    - Erstellen des Users und eines Tablespaces pro Applikation (Konsolidierung)
    - Import des Schemas mit Remapping der Tablespaces
  - Jede Migration wurde mindestens einmal getestet
  - Genaue Abschätzung der benötigten Downtime
- Migration dauerte nur jeweils wenige Minuten
- Aktionen für die Mittagspause



- Migration großer Datenmengen:
  - Export der Metadaten und erstellen der Schematas auf der Exadata
  - Neustrukturierung verschiedener partitionierter Tabellen
  - Kopieren aller „alten“ Partitionen mittels DB Links und bulk inserts  
→ 2.5 TB Daten im Vorraus migriert
  - Letzter Schritt: Inkrementelle Synchronisation innerhalb einer angekündigten 2h Downtime mittels DB Links, Applikationen währenddessen umgestellt
- Vorteile von DB Links:
  - Kein Platzbedarf auf filesystem Ebene
  - Kein Transfer der Indizes (werden von der Exadata online neu aufgebaut)
  - Automatischer Reorg aller betroffenen Tabellen → Platzgewinn



# Änderungen für Exadata?

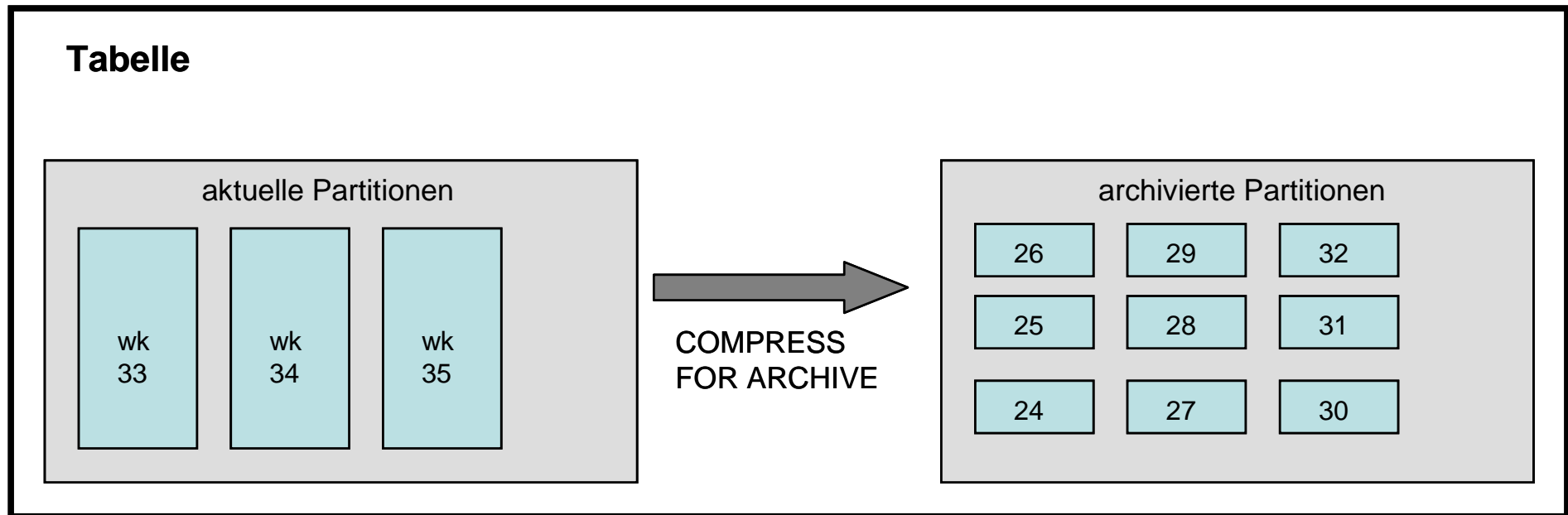


- Prinzipiell keine Änderungen der Applikationen notwendig  
→ Exadata ist eine „normale“ Oracle 11.2 Datenbank
- Alle früheren Tuningmaßnahmen wie Hints entfernt
- Höhere Parallelisierung regelmäßiger Aggregationsjobs möglich
- Längere Vorhaltezeiten verschiedener Chip-Testdaten möglich, dadurch:
  - Änderungen von nachgelagerten Arbeitsabläufen
  - Auswertungen über längere Zeiträume möglich
  - Neue Anforderungen an die DB

# Änderungen für Exadata?



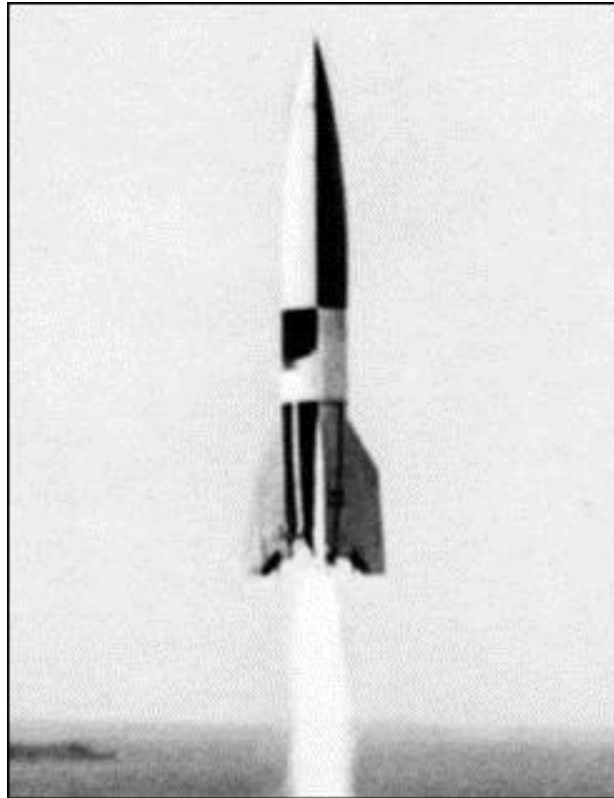
Partitionsstrategie mancher Tabellen wurde geändert:



→ Massive Platzersparnis durch Hybrid Columnar Compression (HCC)

→ I/O Durchsatz und somit Auswertegeschwindigkeit wird erhöht

# Produktiver Start unserer V2



25.03.2010:

Produktiver Launch unserer Exadata V2

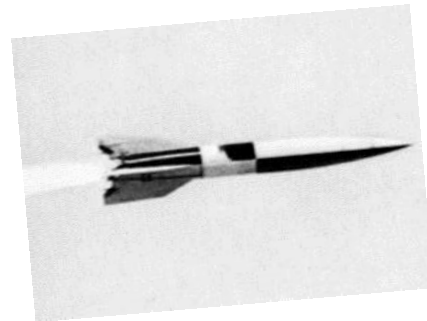
- Motivation und Planungsphase
- Test und Migration
- Fazit und Ausblick
  - Verbesserungen im Produktivbetrieb
  - CPU und I/O Durchsatz: Versprechen eingehalten?
  - Ausblick



# War die Entscheidung richtig?



- Seit Ende März ohne nennenswerte Probleme im Einsatz



- Administrativer Aufwand sehr gering, aber:
  - Gewisse Komplexität des Systems
  - Lernkurve beachten
  - Internes Know-How bezüglich Exadata und RAC aufgebaut
- Bereits drei neue Applikationen auf der Datenbank
- Genügend Raum für Erweiterungen
- Support hat sich bereits verbessert, aber noch Luft nach oben

# Wirklich alles schneller...?



- Daten upload und post processing:
  - Bulk inserts von jeweils ca. 10.000 rows
  - Automatische Aggregation über rechenintensive Trigger in Schleifen (ca. 1000 Ausführungen pro Ladejob auf jeweils 25.000 Zeilen)
  - Vorher 60 min → jetzt 6 min
- Aggregationsjobs / Datamining:
  - Sehr rechenintensive PL/SQL Prozeduren
  - Viele Daten (>20.000 Zeilen pro Select von einer partitionierten Tabelle mit ca. 100.000.000.000 Zeilen)
  - Vorher ca. 2.5 h → jetzt 15-20 min

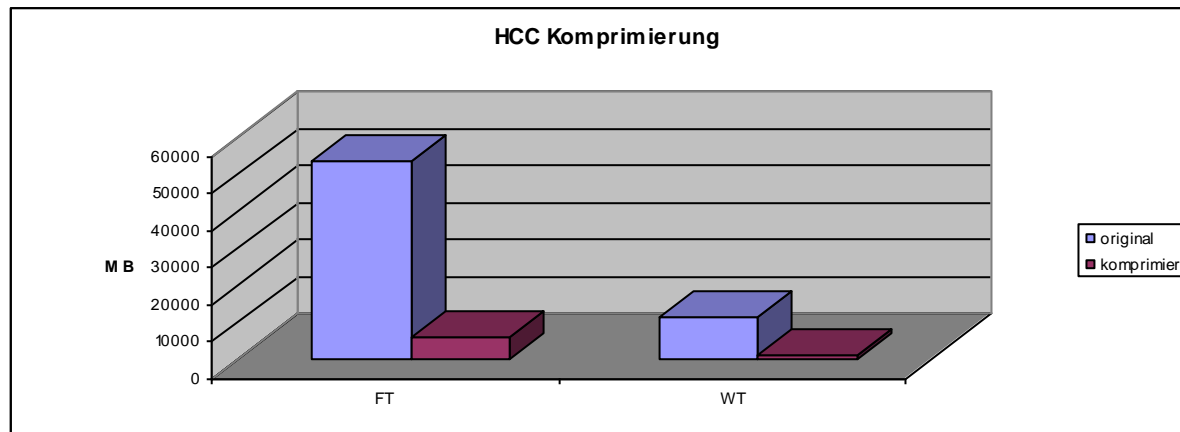


# ...und einfacher...?



- Explain Plans (fast) immer optimal  
→ keine Hints mehr notwendig
- Storage Index funktioniert wie erwartet
- Automatisches parallelisieren – CPU und I/O skaliert  
→ Exadata Komponenten sind aufeinander abgestimmt
- 11.2 Feature: „SQL Überwachung“ im EM sehr hilfreiches Tool

- Extreme Kompression durch HCC (Faktor 10):



- Messbarer Performancezuwachs durch Index Kompression:
  - Geringerer I/O zwischen Zellen und DB-Knoten
  - Keine spürbare CPU Mehrbelastung
- CPU Lastverteilung Richtung Zellen (auch durch sql offload)



- Neues DWH-Projekt:

- In Implementierungsphase
- Firmenweites Management Reporting
- Basierend auf Oracle Business Suite
- SAP → OWB → Exadata



- DBFS Idee:

- Evaluierung des DBFS für das Repository eines Dokumentenverwaltungssystems
- Vorteile wären:
  - Performance und Kompression
  - Einfacher Zugriff mittels SQL auf Dateien aus anderen DB-Anwendungen



## Fragen und Antworten





**Vielen Dank  
für Ihre Aufmerksamkeit**

Jochen Hinderberger  
Dialog Semiconductor  
[jochen.hinderberger@diasemi.com](mailto:jochen.hinderberger@diasemi.com)  
[www.dialog-semiconductor.com](http://www.dialog-semiconductor.com)