

# **Erfahrungsbericht eines „Proof of Concepts“ 12TB RAC DWH → Destination: Exadata**

**Henrik Ruenger (+Andreas Groetz)  
bwin.party services (Austria) GmbH  
Wien**

## **Schlüsselworte**

Exadata, PoC, dwh, Erfahrungsbericht, Migration, Test, Storage, Batchläufe, Business Case, Transport

## **Einleitung**

Im Unternehmen befindet sich ein produktives Data Warehouse, welches kurz vor der Evaluierung der Exadata folgende Parameter aufwies:

- Größe: 12 Terrabyte
- Instanz: 2 RAC Nodes
- Version: 11.1.0.6
- Server: jeweils 1 exklusive Domain auf einer Oracle-Sun M9000
- Storage: EMC<sup>2</sup>-DMX

Aus nachfolgenden Gründen mussten Überlegungen angestellt werden, dieses System abzulösen:

- Weiterer Ausbau der Hardware technisch nicht möglich
- Weiterer Ausbau von Storage technisch nicht möglich
- Re-Evaluierung der Lizenzkosten
- Re-Evaluierung der Wartungskosten

Es sollte eine neue Umgebung evaluiert werden um all diese Punkte zu berücksichtigen und eine möglichst attraktive Lösung dafür zu finden. Die zu evaluierende Umgebung sollte ausserdem Scale-Up fähig sein, sprich nicht nur zum Kaufzeitpunkt richtig dimensioniert sein, sondern auch das weitere Wachstum des dwh's abdecken. Als zukunftsichere Alternative zur derzeitigen Infrastruktur hatte sich das Unternehmen dazu entschlossen ein Proof-of-concept mit der Exadata durchzuführen.

In diesem Vortrag werden sowohl alle theoretischen Vorüberlegungen sowie die direkt in Reading (ETC von Oracle in England) durchgeführten Tätigkeiten beschrieben. Abgerundet wird der Vortrag von den Ergebnissen, welche im praktischen Test erzielt werden konnten.

Zu beachten ist ausserdem, dass es zum Zeitpunkt der Evaluierung im europäischen Raum kaum bis gar keine Unterlagen (vor allen Dingen erfahrungsmäßiger Natur) gab und somit die gesamte Evaluierung von Grund auf selbst erarbeitet wurde.

Spannende Fragen die im Rahmen des Vortrags beantwortet werden sind u.A. die Folgenden:

- Welche Überlegungen muss man anstellen um den Betrieb einer Exadata gewährleisten zu können?
  - Stichworte: Verantwortlichkeiten, Stromverbrauch, Platzverbrauch
- Wie bekommt man das produktive dwh nach England?
  - Stichworte: Transport, Export, Größenbeschränkungen, Indexgrößen, Datenbankgröße
- Welche Tests müssen/sollen/können beim ETC stattfinden?
  - Stichworte: Business Objects (BO) Reports, kompletter produktiver Ladelauf, synthetische Tests
- Welche Vorteile / Nachteile zum Kauf der Exadata gibt es?
  - Stichworte: Business Case, Single-Point-of-Contact, organisatorische Vereinfachungen

## Erfolgskriterien, Prioritäten und High-Level Timeline für das Proof-of-Concept

### 1. Erfolgskriterien

- Es sollen mindestens 35% an Storage Einsparungen vorgewiesen werden
- Wenigstens 10 verschiedene Performance und 10 verschiedene operationale Tests sollten durchgeführt werden. Diese sollen zudem alle schneller laufen als auf dem produktiven System
- Statistiken und Indexe sollen innerhalb von 10 Stunden erstellt werden

### 2. Prioritäten

Test#	Test	Objective	Priorty
1	Laden von Testdaten	Ladezeiten	1
2	Messen der Abfrageperformance	Proof-of-performance	1
3	Tests der verschiedenen Exadata Kompressions-Techniken (anhand von produktiven Daten)	Kompressions-Vorteile erreichen	2
4	Standard DBA Aufgaben	SQL*Loader, Datafiles erstellen, Tables / Indexes erstellen, analyse statistics	3
5	Tests der Exadata cell features	Proof-of-resilience	4

### 3. High-Level Ablauf

Task
Mit der Umgebung vertraut machen (+Workshops)
DWH mit max. 1,5TB erstellen (Vorgabe von Oracle für proof-of-concepts im ETC)
Daten inklusive pl/sql packages laden, indexes erstellen und Statistiken berechnen
Komplette DWH Beladung
DML/DDL statements + Tuning
OLTP Tests und Exadata tuning + Ausfallszenarios (simuliert vom ETC Team)
Komplette DWH Beladung mit allen Tuning Tipps
Bereinigen der Umgebung

## Vorüberlegungen zum Betrieb einer Exadata im Unternehmen

Bevor überhaupt ein praktischer Test durchgeführt werden konnte haben wir uns damit auseinander gesetzt was eine Exadata für das Unternehmen bedeuten würde. Wie würde ein Betrieb einer Appliance aussehen und wie kann man theoretisch auf die Exadata migrieren?

Da der Kauf einer Exadata für uns nicht nur die Ablöse der bestehenden Infrastruktur für das dwh bedeutet hätte, sondern gleichzeitig auch eine Änderung für mehr oder weniger alle Datenbanken, wurden umfangreiche Überlegungen angestellt was alles zum Betrieb einer Exadata zu bedenken ist.

Dabei sind wir hergegangen und haben Antworten zu Fragen ausgearbeitet die uns ein klareres Bild der Exadata vermitteln sollten. Wir werden hier weniger auf die Beantwortung, als mehr auf die Fragestellungen eingehen welche wir ausgearbeitet haben um uns ein Bild darüber zu machen wie das Unternehmen mit einer Exadata in die Zukunft gehen könnte.

Bei dem Proof-of-concept sind verschiedene Schwerpunkte ausgearbeitet worden die jeweils noch einmal eine Priorität zugewiesen bekommen haben. Dabei bedeutet (A) = Priorität 1, (B) = Priorität 2, (C) = Priorität 3 und innerhalb der einzelnen Bereiche ist die Wichtigkeit vom ersten zum letzten Eintrag abfallend.

## 1. Migrationsaufwände

- (A) Migrationsplan
  - Welche Tools/Techniken kommen zur Migration in Frage?
  - Welche Datenbanken müssen/sollen/können überhaupt auf die Exadata migriert werden?
  - Welche Architekturänderungen müssen im Rahmen der Migration vorgenommen werden?
  - Muss das bestehende Backup-Konzept bei einem Umstieg auf die Exadata angegriffen werden?
  - Wird der ODI (Oracle Data Integrator) auf der Exadata unterstützt?
  - Inwieweit unterstützt uns Oracle bei der Installation/Konfiguration (Software, Anschlüsse, Hardware)?
- (A) Betrieb
  - Wie kann ein Betrieb gewährleistet werden (Aufteilung, Zuständigkeiten)?
  - Welche Schulungen sind notwendig?
  - Wie erfolgt das Monitoring?
  - Inwieweit vereinfacht sich der Betrieb (beim Patchen, HW-Austausch, Zusammenarbeit mit Oracle-MySupport)?
  - Ist Linux eine Hürde?
- (B) M9000 Shutdown
  - Wann läuft der Wartungsvertrag der alten Umgebung aus?
  - Welche Alternativen gibt es zur M9000 wenn nicht alle Umgebungen auf die Exadata migriert werden können?
  - Wie hoch ist der Zeitaufwand für die „Weg“-Migration von der M9000?
  - Wieviel Storage auf der EMC<sup>2</sup>-DMX wird pro Umgebung auf der M9000 frei?
  - Wie ist der Preis der M9000 im Vergleich zu anderen Angeboten und auch zur weiteren Exadata Aufrüstung?
  - Was an Hardware bleibt nach der Migration übrig?

## 2. Anforderungen

- (A) Sizing wie CPU, RAM, Storage ...
  - Welche Ausstattung wird bei der Exadata benötigt?
  - Wie viele Exadata Maschinen werden benötigt?
  - CPU Convert Matrix
  - Datenbank Wachstumsplan
  - Wie funktioniert das Nachrüsten der Exadata?
  - Wie sind die Skalierungsmöglichkeiten (Anzahl der RAC-Nodes)?
  - Sollen die Exadata Datacenter-redundant gebaut werden?
  - Soll eine Stresstest/QA-Umgebung auf der Exadata gebaut werden?
- (A) Analyse der bestehenden Ressourcen (Verfügbarkeiten, Wachstumsraten)
  - Welche Wachstumsraten hat das Data-Warehouse?
  - Zu welchen Zeitpunkten laufen Wartungsverträge aus?
  - Welche Servertypen/Anzahl gibt es im bestehenden Einsatz?
  - Welche Netzwerkgeschwindigkeiten herrschen vor?
- (A) Forecast / Analyse von Zukunftsbedarf (Einteilung dwh & Non-dwh)

- Storage Ausblick auf 3-5 Jahre?
- Welche fachlichen Abhängigkeiten gibt es?
- CPU Ausblick auf 3-5 Jahre?
- Welche Puffergröße bezüglich neuer Datenbanken muss bedacht werden?
- Datenklassifizierung bezüglich Kompression?
- Welche Vor/Nachteile bringt eine 2-Architekturstrategie (x86/Sparc)?
- Netzwerk Durchsatz, Ausblick auf 3-5 Jahre?
- DB-Version (Zertifizierungsmatrix)
- HA-Betrachtung
- (B) Datacenter Kapazitäten
  - Stromverbrauch der Exadata?
  - Platzverbrauch der Exadata?
  - Welche Klimakapazität ist noch verfügbar?
  - In welches Datacenter soll die Exadata gestellt werden?
  - Wie viel KW sind noch verfügbar?
  - Wie sieht es mit der Notwendigkeit eines Datacenter übergreifenden Setups aus?
  - Notfallplan bei Datacenter Überkapazität?
  - Anzahl der Netzwerkports?
- (C) Testsysteme
  - Kann man einen Testbereich auf der Exadata separiert betreiben?
  - Wie sieht die Duplizierung von der Produktion auf Test aus?
  - Wie kann ein mit der Exadata vergleichbares System gebaut werden?
- (C) Identifikation von Datenbanken die (nicht) auf die Exadata migriert werden sollen
  - Warum sollte eine Umgebung auf die Exadata migriert werden?
  - Kriterien die erfüllt sein müssen um eine Datenbank auf die Exadata zu bringen?

### **Trip nach Reading (ETC)**

Die Bedingung für das proof-of-concept im ETC war, dass die mitgebrachten Daten 1.5 TB nicht übersteigen dürfen. Dafür mussten einige Überlegungen angestellt werden um das dwh von seiner ursprünglichen Größe von 12 TB auf 1.5 TB zu reduzieren (die 1.5TB Maximalgröße sind eine Vorgabe von Oracle für proof-of-concepts im ETC).

Um das proof-of-concept nicht zu gefährden wurde folgender Ablaufplan aufgestellt:

- Schritt 1: proof-of-concept Dokumente werden zu Oracle geschickt damit eine Reservierung der Testumgebung im ETC vorgenommen werden kann
- Schritt 2: Skripterstellung damit das dwh auf 1.5 TB verkleinert wird (Daten eines Jahres ohne Indexe)
- Schritt 3: detaillierter Testplan muss bereits 3 Wochen vor der Anreise ins ETC stehen (Aufbau und Rekalibrierungen nehmen 50% der Zeit in Reading ein)
- Schritt 4: interne Performance Tests um Vergleichszahlen im Produktivsystem auf der M9000 zu haben. Eine Baseline um Verbesserungspotentiale im ETC testen zu können
- Schritt 5: Export der Produktivdaten zu Flatfiles
- Schritt 6: Durchführung der Tests im ETC
- Schritt 7: eine Woche später Präsentation der proof-of-concept Ergebnisse vor allen Stakeholdern

### **Durchgeführte Tests und Fokusbereiche**

Folgende Schwerpunktbereiche wurden ausgewählt:

- Produktive Tests
- Synthetische Tests
- Hardware Tests

Der Ablauf der Tests stellte sich wie folgt dar:

1. Erster Tag im ETC (Ziel: Klon der Produktion auf der Exadata nachstellen, Vorbereitungen abgeschlossen für eine dwh Beladung am zweiten Tag)
  - Daten wurden bereits präventiv ins DBFS geladen
  - Datenbank wurde erstellt
  - Vertraut machen mit der Umgebung
  - Schematas erstellen (DDL)
  - Daten aus externen Tabellen laden
  - Indexe erstellen
2. Zweiter Tag im ETC (Ziel: Vorbereitungen für Business Objects und dwh erledigt)
  - Rest der Strukturen erstellen
  - Cell storage Struktur Session zur Erklärung
  - Infrastruktur (Views und Prozeduren) werden für BO-Tests geladen
  - Erstellung der Indexes abgeschlossen
  - Synthetische Tests mit verschiedenen Komprimierungsstufen werden durchgeführt
  - Vorbereitungen für manuelle BO-Tests werden abgeschlossen
3. Dritter Tag im ETC (Ziel: BO-Reports + Ergebnisse, Tuning)
  - Statistiken berechnen
  - OS-Management/Cell-Management/Monitoring Know-How Session
  - BO-Reports werden durchgeführt
  - Synthetische Tests werden durchgeführt
  - DBA Tasks werden durchgeführt
4. Viertes Tag im ETC (Ziel: funktionierende BO-Reports, dwh Beladung, Tuning)
  - Erste Ergebnisse Kompression und Performance liegen vor
  - Dwh Beladung wird durchgeführt
  - Komprimierte Tabellen für BO-Lauf
  - Diverse dwh Beladungen mit komprimierten Tabellen
  - Synthetische Tests werden durchgeführt
5. Fünfter Tag im ETC (Ziel: Tuning Läufe, Crash-Tests, Cleanup, Vorbereiten der Reports)
  - BO-Reports Longrunner
  - Dwh Beladung wird durchgeführt
  - Cleanup des Systems
  - Datensammlung für Berichterstattung

## Ergebnisse

1. Probleme
  - BO-Reports (100 Langläufer) mit ausgeschalteten Indexes
    - 98% sind durchgelaufen, 2% bedürfen detailliertes Tuning
  - Estimate statistics lief stundenlang
    - 3600 (~10 Jahre) tägliche Partitionen lagen auf einer Tabelle und diese wurde seriell berechnet
  - Monitoring mit Database-Control auf der exadata
    - Das ETC hat uns ein selbstgeschriebenes Tool zur Verfügung gestellt welches wesentlich sinnvoller zum Überwachen der gesamten Exadata ist
  - Für differenziertere Tests wäre mehr Zeit notwendig gewesen

- Einen produktiven Klon mit individuellen Start-Routinen zu bauen dauert alleine schon ca. 2.5 Tage

## 2. Learnings

- Zugriff auf komprimierte Tabellen ist gleich schnell (mit Vorteil beim Verbrauch des Speicherplatzes)
- Zwei Tage werden für den Aufbau der Produktivumgebung gebraucht (mit minimalen Konfigurationsänderungen)
- “Out-of-the-box” Datenbank kann komplett ohne Tuning benutzt werden bei sehr guten Testergebnissen
- Alle 8 Nodes des RAC wurden ohne Probleme hinzugeschaltet
- Storage Cells wurden hinzugefügt (bei 4.9 TB hat das Rebalancing 3.5 Stunden gedauert)
- Es waren während der gesamten Woche keine administrativen Arbeiten (patches, bugs) notwendig

## 3. Dwh Beladungsergebnisse

Die folgende Grafik stellt, von links nach rechts, die einzelnen Level dar welche innerhalb der dwh Beladung während der Nacht abgearbeitet werden. Dabei ist zu erkennen das jeder senkrechte Bereich auf der Exadata (mit wenigen Longrunnern die abgebrochen werden mussten) nur zwischen 5-20% der Zeit einnimmt welche es noch auf der Produktivumgebung benötigt hatte:

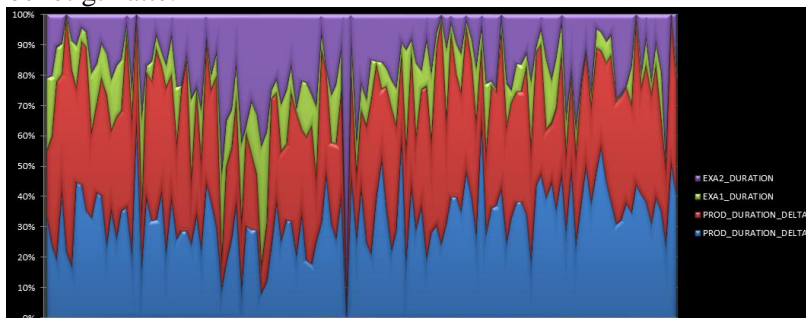


Abbildung 1: DWH-Batchlauf (einzelne Steps)

## 4. Synthetische Tests – Ergebnisse der Kompression

Im folgenden Bild wird dargestellt wie sich der Platzverbrauch in der Produktivumgebung (orange bzw. auf der linken Seite in den Graphen) und auf der Exadata darstellt. Dabei ist zu erkennen dass in der Exadata immer eine massive Platzeinsparung erwirkt wird

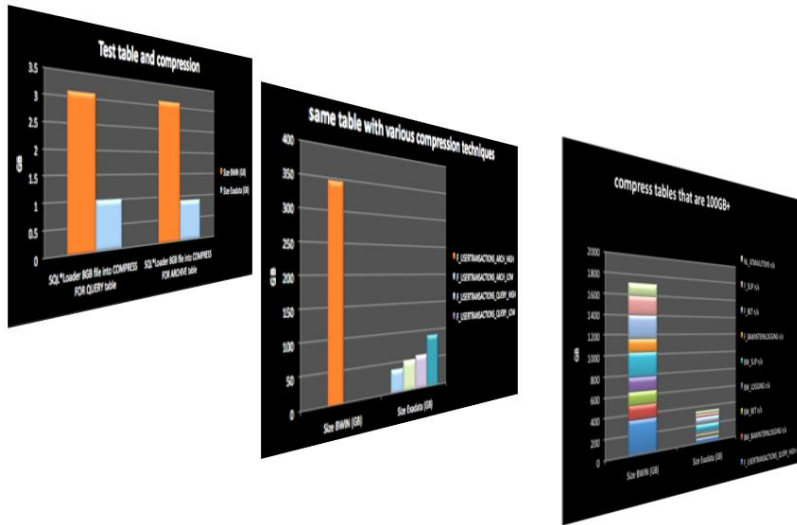


Abbildung 2: Kompressionsergebnisse

- Forecast des benötigten Speicherplatzes für die nächsten Jahre  
Die beiden linken Linien beziehen sich auf den benötigten Speicherplatz für die nächsten Jahre (dabei wird ein Puffer, in blau, eingerechnet um etwaige neue Projekte mit aufnehmen zu können) sowie in grün der prognostizierte Speicherplatz der auf der Exadata benötigt würde (durch Kompression und Verzicht auf Indexe)

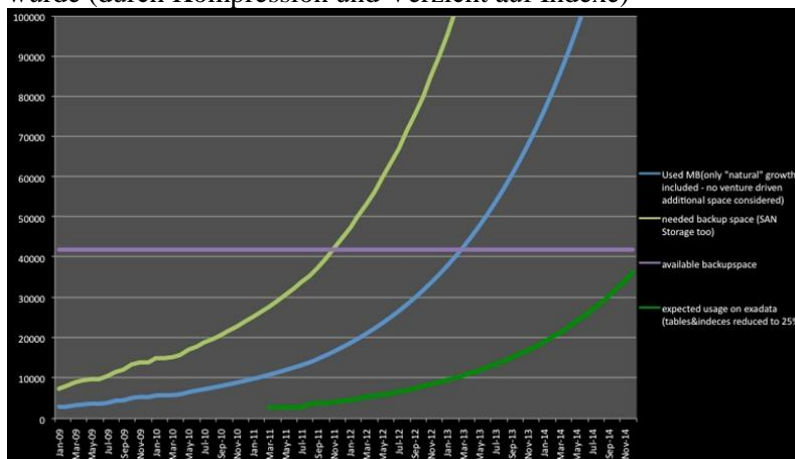


Abbildung 3: prognostizierter Platzbedarf des DWH

- Synthetische Tests – Ergebnisse der Performance  
Auch hier ist in orange die produktive Umgebung abgebildet und jeweils rechts daneben die Ergebnisse mit exakt denselben Daten auf der Exadata. Dabei ist ein massiver Geschwindigkeitsgewinn zu erkennen. Dieser betrug in jedem Fall immer mehr als 60%.

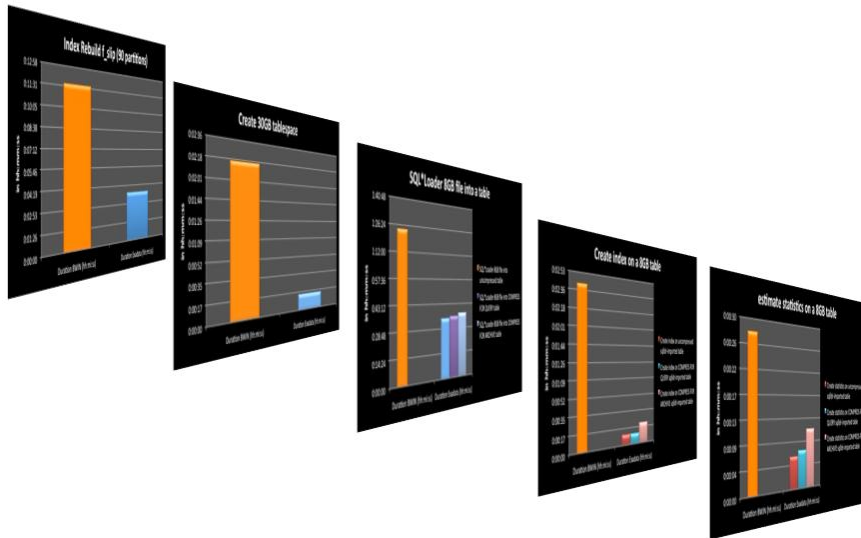


Abbildung 4: Performanceergebnisse

### Fazit

In allen Belangen konnte die Exadata überzeugen. Als Abschluss soll uns folgende Auflistung der Möglichkeiten einen Überblick über die Schwerpunktbereiche ermöglichen, die uns vor allem interessiert hatten. Ausserdem wird dadurch der Effekt der Exadata in den einzelnen Bereichen klarer.

- o Performance:

Problem (Produktion)	Lösung (Exadata)
Ein Neustart der dwh Beladung beeinflusst das Business direkt	Neustart Möglichkeit, Verzögerungen bei Vorkomplexen in der Datenlieferung können ausgeglichen werden
Nur wenig Wartungsfenster sind möglich	Statistiken werden in kürzester Zeit berechnet, Rollouts gehen schneller, wenig individuelles Tuning notwendig
Administrative Arbeiten benötigen Ressourcen und Zeit	Automatisches Archivieren und Optimierungen beim Zugriff auf Daten, sowie ein Flashback Table werden schneller erledigt
Exklusiver Zugriff während der dwh Beladung wird benötigt	Zugriff während der dwh Beladung beeinflusst den Batch-Lauf in keinster Weise
Rollouts benötigen viel Zeit	Dramatisch reduzierte Zeiten für das Erstellen von Tabellen, Indexe, Statistiken
Complexes Datenbank-Design	Weniger Indexe werden benötigt, auch wesentlich komplexere Abfragen sind möglich
Desaster Recovery ist durch den benötigten Platzbedarf problematisch	Weniger Platz wird benötigt, dadurch verkürzen sich Recovery Zeiten dramatisch

- o Storage:



<b>Problem (Produktion)</b>	<b>Lösung (Exadata)</b>
Vorhandensein der QA ist teuer	Durch "out-of-the-box" Kompression wird weniger Platz benötigt
Das dwh wird teuer je mehr Daten geladen werden	Massive Kompression „out-of-the-box“
Im Disaster Fall wird eine extra Umgebung mit gleichem Platzbedarf benötigt	Eine kleine Umgebung ist permanent reserviert und kann trotzdem für "query-offloading" benutzt werden
Zurzeit ist die QA Umgebung nicht 100% baugleich	Eine 100% baugleiche Umgebung kann mehrmals eingerichtet werden, und das bei minimalem Zeitaufwand

### **Kritische Betrachtung**

- Eine Woche im ETC in Reading ist definitiv zu kurz um alle Möglichkeiten der Exadata auszuprobieren. Dazu kommt noch das wenn man seine eigenen Daten mitbringt, diese auf die bereits im Einsatz befindliche dwh Umgebung angepasst sind (Indexe, Hints, pl/sql). Um diese zu ändern und auf die neue Umgebung anzupassen sind umfangreiche Tests oder zumindest Codeänderungen erforderlich.
- Verantwortlichkeiten für den Betrieb einer Exadata müssen komplett neu überdacht werden. Oracle schlägt vor dass viele Dinge aus einer Hand (des DBA) bedient werden können, dafür müssen aber diese sich viel neues Wissen über Storage und Serveradministration aneignen. Des Weiteren ist die bei vielen Firmen etablierte Gewalten (Verantwortlichkeiten) Trennung mit der Exadata zu überdenken da eine klassische Aufteilung in Storage, Server, Netzwerk, Datenbank wenig Sinn macht und die Exadata auch als Appliance betrieben werden muss

### **Referenzen**

- Einsatzkriterien und Erfolgsfaktoren Oracle Database Machine/Exadata, Alfred Schlaucher, DOAG Konferenz 2010 (im Konferenzband)
- <http://www.beyenetwork.de/view/10194>
  - Für wen rechnet sich der Oracle Exadata Storage Server, Jaqueline Bloemen, April 2009
- Performance-Reports der in&out ag in der Schweiz
  - [http://www.inout.ch/files/pdf/news/exadata\\_fachbericht.pdf](http://www.inout.ch/files/pdf/news/exadata_fachbericht.pdf)
  - [http://www.inout.ch/files/pdf/download/inout\\_orabench\\_db\\_benchn\\_and\\_perf\\_metrics.pdf](http://www.inout.ch/files/pdf/download/inout_orabench_db_benchn_and_perf_metrics.pdf)

### **Kontaktadresse:**

Henrik Ruenger  
bwin.party services (Austria) GmbH  
Marxergasse 1b  
A-1030 Wien

Telefon: +43 (0) 664-853 4536  
E-Mail: [henrik.ruenger@bwinparty.com](mailto:henrik.ruenger@bwinparty.com)  
Internet: [www.bwinparty.com](http://www.bwinparty.com)