

# ORDIX news

Das IT-Magazin der ORDIX AG



## Das „i-Tüpfelchen“ in puncto Sicherheit

24 | Oracle 11gR2 Security - Welche Neuerungen bringt das Release R2?

6 | Die Qual der Wahl  
Oracle und XML:Datentyp XMLType

18 | Perl-Skripte konfigurieren  
Auslagerung der Skripte mit XML::Simple

10 | Pacemaker -  
Schrittmacher für Heartbeat?

36 | Resource Management mit Linux  
Control Groups - Prozesse in Gruppen



# Das Treffen der Oracle-Community

16. – 18. November 2010 in Nürnberg: **DOAG 2010** Konferenz + Ausstellung

## Prominente Keynote-Speaker



Jürgen Kunz  
Senior Vice President  
und verantwortlich für  
die ORACLE Deutschland  
B.V. & Co. KG  
**Software. Hardware.  
Complete.**



Tom Kyte  
Datenbank-Experte „Ask  
Tom“, Buchautor und  
Senior Technical Architect  
bei Oracle  
**What are we still doing  
wrong?**



James Gosling  
Urvater der Program-  
miersprache Java  
**Time to move on!**



Jean Reiczky  
Senior Vice President Oracle  
Customer Services for EMEA  
**Maximizing customers  
investments and success for  
the investment they have  
made in Oracle technology**



Prof. Hermann Maurer  
Institut für Informations-  
systeme und Computer  
Medien (IICM), TU Graz  
**„Es stehen uns noch viele  
Entwicklungen und Über-  
sorschungen ins Haus ...“**



Steven Feuerstein  
PL/SQL-Evangelist  
und Autor zahlreicher  
PL/SQL-Bücher  
**Coding Therapy for  
Software Developers**



Andrew Sutherland  
Senior Vice President  
und verantwortlich für  
Middleware Technology  
Oracle EMEA  
**InfraStructure & Cloud,  
The New Paradigm**



Chris Baker  
Senior Vice President Core Tech-  
nology Business Oracle EMEA,  
Middle East & Africa  
**InfraStructure & Cloud,  
The New Paradigm**  
(gemeinsam mit A. Sutherland)



Thomas Schröder  
Vice President  
Hardware Sales  
Oracle Northern Europe  
**Kundenvorteile durch  
den „integrated stack“**



Richard Mason  
Vice President MySQL  
Sales and Consulting  
Oracle Corp.  
**Oracle strategy and  
plans for MySQL**



Ed Burns  
Consulting Member Java-  
Server Faces Oracle Corp.  
**Oracle's Runtimes and  
Frameworks:  
A Roadmap for Deve-  
loper Productivity**



Brian Oliver  
Global Solutions Architect  
Oracle Coherence Engineering  
Oracle Fusion Middleware  
**The Science and Physics of  
Application Performance and  
Scalability**

TTT

Paderborn, September 2010

Die ARD möge es mir verzeihen, dass ich den Titel ihrer Sendung „Titel – Thesen – Temperature“ missbrauche. Das nicht ganz neue Spiel in der IT-Branche lautet aber ebenfalls TTT: Täglicher Telko Terror.

Nichts gegen sinnvolle Telefonkonferenzen, sie können meist kürzer und zielführender sein als deren Vorläufer SM (nein nicht Sado Maso, obwohl manchmal so anmutend). Ich rede von täglichen Status-Meetings in der Endlos- (und Ziellos-) Variante.

Aber woher rührt dieses Problem? Lassen Sie mich etwas ausholen. Vor mehr als 20 Jahren wünschte ich mir, dass es im Bereich der Entscheidungsträger jüngere Personen gäbe, die mit einem guten Technikverständnis auf Basis guter und knapper Informationen schnell plausible und nachvollziehbare Entscheidungen treffen. Heute wünsche ich mir, dass anstelle der McKinsey<sup>1</sup> gläubigen, jungen und völlig unerfahrenen Entscheidungsträger der eine oder andere kernige Mittfünfziger mit dementsprechender Vita auf dem richtigen Stuhl sitzen würde.

Was ist passiert? IT wurde im Laufe der Jahre immer teurer und teilweise wurden Unsummen für überflüssige Anschaffungen, gigantische Projekte, die in der Versenkung verschwanden, bevor sie jemals auf Anwender losgelassen werden konnten, verschwendet. Man denke nur an die Unsummen, die Y2K auf dem Gewissen hat.

Also galt es, die IT-Budgets einzufrieren oder besser auf ein Niveau deutlich vor 2000 zurückzuführen. Es mussten nun (ellenlange) Begründungen für Ausgaben und ROI-Dokumente erstellt werden. Die (Management-) Berater fühlten sich ebenfalls berufen und ihnen fiel das ein, womit Sie in den neunziger Jahren das produzierende Gewerbe „reformiert“ hatten: Das Factory-Prinzip. Nach dem Motto „Was in der Automobilbranche funktioniert hat, kann ja in der IT nicht so falsch sein“ wurden IT-Unternehmen und Abteilungen reihenweise umorganisiert. Wobei manche dieser Umorganisationen nicht mal annähernd die Lebensdauer einer belgischen Regierung erreicht.

Nein, ich möchte jetzt nicht die McKinsey<sup>1</sup> Berater als unfähig bezeichnen, letztendlich wäre ich vor knapp 30 Jahren ja auch beinahe einer geworden. Aber IT ist kein Fertigungsbetrieb, die meisten IT-Projekte lassen sich eher mit einer Manufaktur vergleichen. Die kann auch erfolgreich sein, aber auch dort würde ein Factory-Prinzip nicht funktionieren, selbst wenn man beim Einkauf bestimmter Teile auf „Factories“ zurückgreifen kann.

Was hilft? Im Laufe meiner nunmehr dreißigjährigen IT-Erfahrung hat sich eines herausgestellt: Wissen hilft. Ich erinnere mich gerne an Projekte und Situationen, wo Millionen für neue leistungsfähigere Hardware schon bereitgestellt waren und dann wenige tausend EURO für eine gut gemachte Performance-Analyse mit anschließend konsequent umgesetzten Maßnahmen ausreichten, um Performance und Anwender zufriedenzustellen.

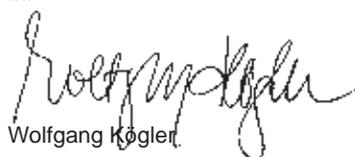
Und auch in der jüngeren Vergangenheit hat uns immer wieder bei der erfolgreichen Umsetzung von Kundenprojekten gerade die Aushebelung des Factory-Prinzips zum Erfolg verholfen.

Also nehmen Sie meinen Rat mit: Komplexe IT-Projekte brauchen Macher, Verantwortliche und Entscheider, die auf gutes Wissen und einen dementsprechenden Erfahrungsschatz zurückgreifen können. Solche Menschen treiben ein Projekt wesentlich effizienter voran als viele Linienverantwortliche, deren Interesse am Gelingen des einzelnen Projektes zu gering ist. Je umfangreicher das technische Wissen und der Erfahrungsschatz sind, desto erfolgreicher werden die Projekte sein.

Zum Aufbau Ihres Wissensschatzes tragen wir mit dieser Ausgabe in folgenden Themen bei: AIX7, Oracle und XML, DB2 und Perl, Open Source Cluster und Control Groups in Linux sowie Security und Oracle, die jüngsten Oracle RAC und Grid Features.

Die Tage werden kürzer, es ist die richtige Zeit, um wieder viel zu lesen. Ich wünsche Ihnen viel Spaß und wenn in der nächsten Telko mal wieder zu viel geredet wird, halten Sie es mit Dieter Nuhr: „Wenn man keine Ahnung hat, einfach mal Fr... halten.“

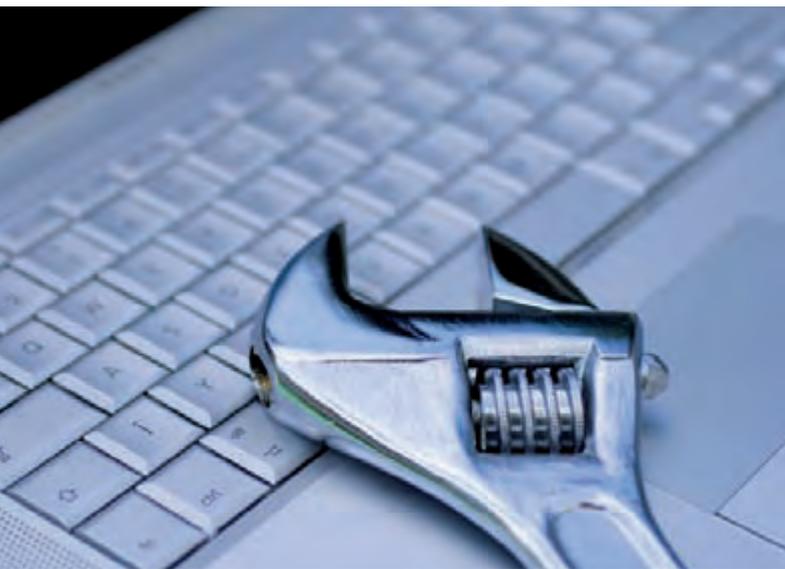
Ihr



Wolfgang Kögler



<sup>1</sup> Sie können den Namen McKinsey jederzeit durch ein anderes Beratungsunternehmen ersetzen (Accenture, PWC, ...).



18 | Perl-Skripte konfigurieren



24 | Oracle 11gR2 - Das „i-Tüpfelchen“ in puncto Sicherheit

## Java/JEE

- 33..... **Endlich dabei - die Criteria Query API in der JPA 2.0**  
Ein Einblick in die Criteria Query API, die nun Bestandteil der Java Persistence API 2.0 ist

## Betriebssysteme

- 30..... **IBM POWER 7 und AIX 7  
Leistung zur siebten Potenz**  
Dieser Artikel gibt einen Überblick über die Prozessortech-  
nologie der POWER7-Hardware und über zu erwartenden  
neuen Funktionen des Betriebssystems AIX7.

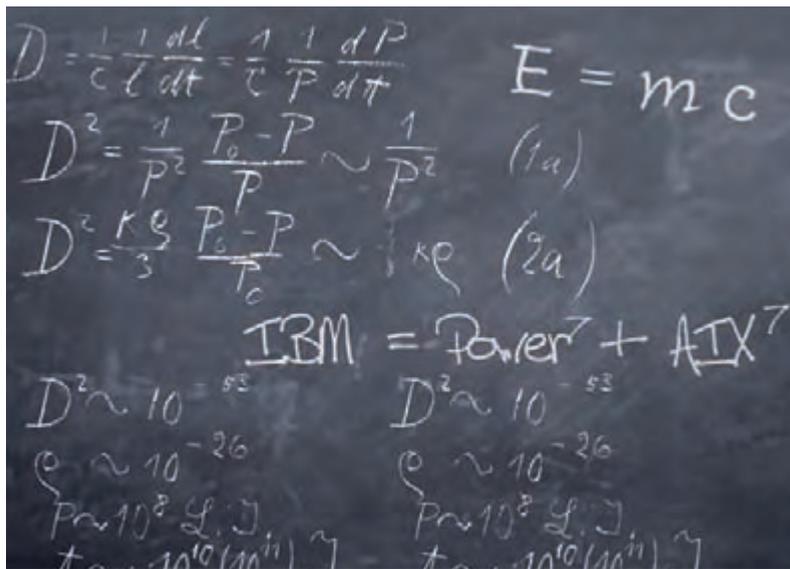
## Open Source

- 10..... **Open Source Cluster mit Pacemaker (Teil I):  
Pacemaker - Schrittmacher für Heartbeat?**  
Heartbeat hat sich zu einer beliebten Open Source  
Cluster-Lösung entwickelt. Diese neue Reihe beschäftigt  
sich mit vergangenen, aktuellen und zukünftigen Funkti-  
onen dieser Open Source-Lösung.
- 36..... **Resource Management mit Linux  
Control Groups - Linux legt Prozesse in Ketten**  
Control Groups sind seit der Version 2.6.24 ein Bestandteil  
des Standard-Linux-Kernels. Dieser Artikel erläutert, wie  
an Hand von Control Groups, Prozesse in Gruppen gefasst  
werden und deren Eigenschaften mit Hilfe Subsystemen  
manipuliert werden können.

## Datenbanken

- 6..... **Oracle und XML (Teil V):  
Die Qual der Wahl**  
Es werden drei verschiedene Möglichkeiten  
beleuchtet, wie XML-Daten als XMLType in  
einer Oracle Datenbank gespeichert wer-  
den können und welche Speicherungsform  
zur höchstmöglichen Performance führt.
- 18..... **Perl und DB2 im Verbund (Teil III):  
Perl-Skripte konfigurieren**  
Dieser Artikel erläutert wie man Perl-Skripte  
für die Verwendung mehrerer Datenbanken  
vorbereitet.
- 24..... **Oracle 11g Security (Teil III):  
Das „i-Tüpfelchen“ in puncto Sicherheit**  
Welche Neuerungen wurden mit dem  
Release R2 im Bereich der Sicherheit ein-  
geführt. Lohnt sich ein Umstieg?
- 40..... **Quantensprung von Oracle RAC-Version  
11gR1 auf 11gR2**  
Dieser Artikel beleuchtet die umfangreichen  
Änderungen der neuen RAC-Version und  
gibt einen Einblick über weitere Einsatz-  
möglichkeiten der Grid Infrastructure.

**ORDIX 2.0**  
1990 - 2010 | 20 Jahre bessere IT-Beratung



30 | Leistung zur siebten Potenz: Power7 und AIX7



36 | Control Groups - Linux legt Prozesse in Ketten

## Aktuell

- 14..... **DOAG Konferenz und Schulungstag 2010**  
ORDIX und coniatos präsentieren Ihr Know-how.
- 21..... **Larry Ratlos:**  
Das Problem mit den Zeitzonen.

## Standards

- 03..... **Editorial**
- 04..... **Inhalt**
- 05..... **Impressum**
- 22..... **Seminarübersicht:**  
September bis Februar 2011

## Impressum

### Herausgeber:

ORDIX AG Aktiengesellschaft für Softwareentwicklung, Beratung, Schulung und Systemintegration, Paderborn

**Redaktion:** Jens Pothmann, Evelyn Ernst

**V.i.S.d.P.:** Benedikt Georgi, Wolfgang Kögler

### Anschrift der Redaktion:

ORDIX AG  
Westernmauer 12 - 16  
33098 Paderborn  
Tel.: 05251 1063-0  
Fax: 0180 1673490

### Gestaltung/Layout:

Jens Pothmann

**Auflage:** 9.900

**Druck:** Media-Print Informationstechnologie GmbH, Paderborn

### Bildnachweis:

© istockphoto.com / Computer maintenance © sndr  
© istockphoto.com / Padlock going through chip © Photobvious  
© istockphoto.com / House Building © alexsl  
© sxc.hu / Drop of oil © pinzino  
© it-sa.de / Eingang Ost © it-sa

### Autoren dieser Ausgabe:

Dr. Uwe Bechthold, Marius Dorlöchter, Evelyn Ernst, Christian Fertsch, Kathrin Hammerschmidt, Kathleen Hock, Wolfgang Kögler, Norbert Munkel, Michael Skowasch

Die Zeitschrift ORDIX News wird von der ORDIX AG an ausgewählte Kunden verteilt und kann für 2,20 Euro bestellt werden. Sie können die Zusendung der ORDIX News jederzeit ohne Angabe von Gründen schriftlich (z. B. Brief, Fax, E-Mail) abbestellen. Die neueste Ausgabe wie auch ältere Ausgaben finden Sie im Archiv der ORDIX News im Internet unter: <http://www.ordix.de>. Schauen Sie mal rein!

Der Kontakt zu unseren Lesern ist uns sehr wichtig. Für Anregungen, Kritik und Anmerkungen zu den Themen, aber auch für interessante Ideen sind wir immer offen und dankbar. Wir freuen uns auf Ihr Feedback an [redaktion@ordix.de](mailto:redaktion@ordix.de).

### Copyright:

ORDIX AG. Alle Rechte, auch die der Übersetzung, des Nachdrucks und der Vervielfältigung der Artikel oder von Teilen daraus, bleiben uns vorbehalten. Kein Teil der Artikel darf ohne unsere schriftliche Genehmigung in irgendeiner Form reproduziert, insbesondere unter Verwendung elektronischer Systeme verarbeitet, verbreitet, vervielfältigt oder zu öffentlichen Wiedergaben benutzt werden.

### Haftung:

Eine Haftung für die Richtigkeit der Veröffentlichungen kann trotz sorgfältiger Prüfung durch die Redaktion vom Herausgeber nicht übernommen werden.

### Warenzeichen:

Einige der aufgeführten Bezeichnungen sind eingetragene Warenzeichen ihrer jeweiligen Inhaber. ORDIX® ist registrierte Marke der ORDIX AG.

# Die Qual der Wahl

Dieser Artikel richtet sich an Entwickler und Software-Architekten, die XML im Zusammenspiel mit Oracle-Datenbanken einsetzen möchten.

Es gibt drei verschiedene Möglichkeiten, XML-Daten als XMLType in der Oracle-Datenbank abzuspeichern. Aber wann sollte eigentlich welche Speicherungsform eingesetzt werden? Und wie erreiche ich die höchstmögliche Performance für meine Anwendung? Diese Fragen werden in diesem Artikel näher untersucht.

## Speicherungsformen

Für den Datentyp XMLType gibt es seit Oracle 11g drei verschiedene Speicherungsformen. Diese werden im Folgenden kurz vorgestellt:

- **Unstrukturierte Speicherung:**  
Bei der unstrukturierten Speicherung werden die XML-Daten intern als Text im CLOB abgelegt. Dazu wird zu jeder XMLType-Spalte intern eine CLOB-Spalte angelegt, die für den Benutzer allerdings nicht sichtbar ist. Die XMLType-Spalte enthält intern also nur eine Referenz auf diese CLOB-Spalte. Wurde beim Anlegen einer Spalte vom Datentyp XMLType kein XML-Schema angegeben, so wird automatisch ein CLOB erzeugt.
- **Objektrelationale Speicherung:**  
Diese Art der Speicherung ist nur möglich, wenn das passende XML-Schema mit Hilfe des Packages DBMS\_XMLSCHEMA in der Datenbank registriert wurde. In diesem Fall werden aufgrund der Informationen aus dem XML-Schema die entsprechenden objektrelationalen Tabellen oder Views erzeugt.
- **Binary XML:**  
Diese Speicherungsform gibt es seit Oracle 11g. Die XML-Dokumente werden in einem speziell entwickelten BinärfORMAT in der Datenbank abgespeichert. Binary XML bietet dazu einen flexiblen Schema-Support, so dass XML-Dokumente, die mehreren XML-Schemata genügen können, in der gleichen Spalte abgespeichert werden können.

Weitere Informationen über den Datentyp XMLType finden Sie in dem ORDIX News Artikeln der Ausgaben 1/2007 [1] und 2/2010 [3].

## Wann welche Speicherungsform einsetzen?

Vor der Speicherung von XML-Dokumenten sollte genau überlegt werden, welche Operationen auf den XML-Dokumenten in der Datenbank ausgeführt werden sollen. Es müssen also zunächst die Anwendungsfälle definiert werden. Erst dann kann geklärt werden, welche Speicherungsform verwendet werden soll. Die Wahl der Speicherungsform ist deshalb so wichtig, da sie die Performance der Anwendung stark beeinflusst.

Im ORDIX News Artikel „Performance Tuning“ der Ausgabe 2/2007 [2] wurde das Thema der Wahl der Speicherungsform bereits kurz beleuchtet. Allerdings gab es zu diesem Zeitpunkt die Speicherungsform Binary XML noch nicht. Aus diesem Grund greifen wir das Thema in diesem Artikel noch einmal auf.

## Performancetest

Anhand eines Performancetests soll aufgezeigt werden, wie stark die Speicherungsform die Performance beeinflusst. Dazu wurde zunächst für jede der Speicherungsformen jeweils eine Tabelle angelegt (siehe Abbildung 1). Für die objektrelationale Speicherung muss natürlich vor dem Anlegen der Tabelle das entsprechende XML-Schema an der Datenbank registriert werden. Nun führen wir nacheinander die

unten gelisteten Operationen durch und messen dabei jeweils die Verarbeitungsdauer. Bei den Select-Anweisungen wird das Statement jeweils zwei Mal ausgeführt und die Verarbeitungszeit des 2. Statements („warme Datenbank“) verwendet. Beim Insert werden die Statements durch eine Prozedur erzeugt. Durch den Aufruf der Prozedur mit den jeweiligen Tabellennamen als Aufrufparameter werden die einzelnen Tabellen gefüllt.

Folgende Operationen wurden ausgeführt:

- Einfügen von 10.000 XML-Dokumenten in die Tabellen (siehe Abbildung 2).
- Selektieren der kompletten XML-Dokumente aus der Datenbank (siehe Abbildung 3).
- Selektieren von Teilen der XML-Dokumente aus der Datenbank (siehe Abbildung 4).

Die Ergebnisse der Performancetests werden zum Vergleich in Abbildung 5 dargestellt. Es zeigt sich, dass die Wahl der Speicherungsform erhebliche Auswirkungen auf die Performance hat.

## Unstrukturierte Speicherung

Bei der Betrachtung der Ergebnisse fällt auf, dass das Einfügen von XML-Dokumenten in die Datenbank bei der unstrukturierten Speicherung am schnellsten geht. Auch das Auslesen der kompletten XML-Dokumente aus der Datenbank ist bei der unstrukturierten Speicherung am performantesten.

Dies lässt sich daraus ableiten, dass bei der strukturierten Speicherung das Dokument beim Insert analysiert und in mehrere Tabellen aufgesplittet werden muss. Bei Abfragen werden Dokumente wieder aus den einzelnen Tabellen zusammengesetzt. Diese Vorgänge können gerade bei sehr großen XML-Dokumenten viel Zeit in Anspruch nehmen. Soll also immer oder hauptsächlich auf das komplette XML-Dokument zugegriffen werden, so ist die unstrukturierte Speicherung zu empfehlen.

Es gibt allerdings auch noch einige weitere Gründe, die unstrukturierte Speicherung zu verwenden: Die unstrukturierte Speicherung muss zwingend verwendet werden, wenn das Original XML-Dokument (mit allen Whitespaces) beim Auslesen wieder hergestellt werden muss. Außerdem ist die Speicherung in einem CLOB dann vorteilhaft, wenn die Struktur der XML-Dokumente stark variiert oder wenn große

```
CREATE TABLE auto_text(
  autoid NUMBER,
  herstellerliste XMLTYPE
);

CREATE TABLE auto_objektrelational(
  autoid NUMBER,
  herstellerliste XMLTYPE
)
XMLTYPE COLUMN herstellerliste
XMLSCHEMA "http://www.ordix.de/herstellerliste.xsd"
ELEMENT "Auto";

CREATE TABLE auto_binary(
  autoid NUMBER,
  herstellerliste XMLTYPE
)
XMLTYPE COLUMN herstellerliste STORE AS BINARY XML;
```

Abb. 1: Anlegen der Tabellen für den Performancetest.

```
create or replace procedure fill_table(p_table_name in varchar2)
is
  j number:= 0;
begin
  FOR i in 1 .. 10000 LOOP
    j := i+1;
    execute immediate 'insert into '||p_table_name||' values
('||i||', XMLTYPE('<?xml version="1.0" encoding="UTF-8"?>
<Auto xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://www.ordix.de/herstellerliste.xsd">
  <Herstellerliste>
  <Hersteller id="||i||">
  <Name>Motorenwerke' ||i||'</Name>
  </Hersteller>
  <Hersteller id="||j||">
  <Name>Best-Blech" ||j||'</Name>
  </Hersteller>
  </Herstellerliste>
  </Auto>''))';
    commit;
  end loop;
end;
/
```

Abb. 2: Anlegen der Prozedur zum Füllen der Tabellen.

Datenmengen zu speichern sind. Veränderungen in der Struktur sind für eine strukturierte Speicherung nicht hinnehmbar: Jedes gespeicherte Dokument muss zum registrierten XML-Schema passen.

## Objektrelationale Speicherung

Anhand der Ergebnisse des Performancetests lässt sich auch erkennen, dass beim Selektieren von Teilen des XML-Dokuments, die objektrelationale Speicherungsform mit am schnellsten ist. Somit sollte auf diese Spei-

```
select herstellerliste from auto_text;
select herstellerliste from auto_objektrational;
select herstellerliste from auto_binary;
```

Abb. 3: Selektieren des kompletten XML-Dokuments.

```
select count(*) from (
select
  extractvalue(value(h), ,Hersteller/Name/text()`) as herstel-
  lername,
  extractvalue(value(h), ,Hersteller/@id`) as herstellerid
  from auto_text, table(xmlsequence(extract(herstellerliste,
  ,/Auto/Herstellerliste/Hersteller`))) (+) h);

select count(*) from (
select
  extractvalue(value(h), ,Hersteller/Name/text()`) as herstel-
  lername,
  extractvalue(value(h), ,Hersteller/@id`) as herstellerid
  from auto_objektrational, table(xmlsequence(extract(herste-
  llerliste, ,/Auto/Herstellerliste/Hersteller`))) (+) h);

select count(*) from (
select
  extractvalue(value(h), ,Hersteller/Name/text()`) as herstel-
  lername,
  extractvalue(value(h), ,Hersteller/@id`) as herstellerid
  from auto_binary, table(xmlsequence(extract(herstellerliste,
  ,/Auto/Herstellerliste/Hersteller`))) (+) h);
```

Abb. 4: Selektieren eines Teils des XML-Dokuments.

	Unstrukturierte Speicherung	Objektrationale Speicherung	Binary XML
Insert	00:00:35.62	00:00:50.42	00:00:39.90
Select des gesamten Dokuments	00:00:24.15	00:00:25.53	00:00:26.53
Select auf Teile des XML-Dokuments	00:00:04.34	00:00:00.28	00:00:00.26

Abb. 5: Ergebnisse des Performatetests.

cherungsform zurückgegriffen werden, wenn viele Zugriffe auf Teile der einzelnen XML-Dokumente erfolgen müssen. Diese Speicherungsform ist allerdings nur dann möglich, wenn die abgespeicherten XML-Dokumente alle dem gleichen XML-Schema genügen. Beim Insert in die Datenbank werden die XML-Dokumente aufgesplittet und in Tabellen und Objekten abgespeichert. Somit muss bei Abfragen oder Updates auf Teile der XML-Dokumente nicht mehr das gesamte Dokument gelesen bzw. geschrieben werden. Es findet ein Query Rewrite statt, bei dem die Abfrage mit dem XPath-Ausdruck in eine relationale

Abfrage umgewandelt wird. Dadurch muss nicht das gesamte XML-Dokument gelesen werden, sondern es wird direkt auf die relationalen Tabellen zugegriffen in denen die Daten in der Datenbank abgespeichert wurden.

Diese Speicherungsform sollte nicht eingesetzt werden, wenn viele Inserts erfolgen und nur selten Teile der XML-Dokumente abgefragt werden, sondern häufig das gesamte XML-Dokument.

### Binary XML

Neben der objektrationalen Speicherung erzielt auch die Speicherung als Binary XML sehr gute Ergebnisse bei unserem Performatetest zum Abfragen einzelner Fragmente aus den XML-Dokumenten. Aber auch die Verarbeitungszeiten bei Einfüge-Operationen waren kaum schlechter als bei der unstrukturierten Speicherung. Diese Speicherungsform kann also dann eingesetzt werden, wenn viele Inserts in die Datenbank gemacht werden und wenn häufig Abfragen auf Teile der XML-Dokumente erfolgen.

Binary XML ist aus mehreren Gründen sehr performant. Da die XML-Dokumente im speziellen Binärformat in der Datenbank abgelegt werden, ist die Größe der zu übertragenen Daten viel kleiner. Das Parsen und Serialisieren der XML-Daten beim Austausch der Daten zwischen den einzelnen Applikationsschichten bzw. vor dem Speichern in der Datenbank fällt zudem weg. Bei der Verwendung von Binary XML als Speichermethode für XML-Dokumente wird der Netzwerk-Overhead verringert, da für die Übertragung das kompakte, interne Format verwendet wird und nicht das traditionelle serialisierte Textformat. Des Weiteren liefert Oracle einen neuen Parser und einen neuen XML-Schema Validator mit. Weder der Parser noch der Validator nutzen einen DOM-Baum zum Parsen bzw. Validieren der Daten. Mehr Informationen dazu finden sich im dem ORDIX News Artikel der Ausgabe 2/2010 [3].

Ein Vorteil gegenüber der objektrationalen Speicherung ist, dass die abgespeicherten XML-Dokumente nicht einem einzelnen XML-Schema genügen müssen. Es können verschiedene Encodings beim Anlegen der Tabelle definiert werden. Somit kann festgelegt werden, ob die XML-Dokumente einem bestimmten Schema, einem von mehreren Schemata oder gar keinem Schema genügen müssen.

Müssen viele Abfragen auf Teile der XML-Dokumente durchgeführt werden, wenn die XML-Dokumente keinem einheitlichen XML-Schema genügen, so ist Binary XML die Speicherungsform der Wahl. Dies gilt auch dann wenn viele Abfragen auf Teile der XML-Dokumente durchgeführt werden müssen und zusätzlich viele Inserts in die Tabelle gemacht werden müssen.

### Fazit

Die Wahl der Speicherungsform hat erhebliche Auswirkungen auf die Performance. Somit sollte beim Design der Datenbank genau darüber nachgedacht werden, welche Art der Speicherung verwendet werden soll. Wenn man sich nicht sicher ist, welche Form am sinnvollsten ist, empfiehlt es sich einen Performancetest durchzuführen, denn die Wahl der falschen Speicherungsform kann durch kein Tuning mehr ausgeglichen werden.



*Kathrin  
Hammerschmidt  
(info@ordix.de).*

### Glossar

<b>DOM</b>	Document Object Model. DOM ist ein API für den Zugriff auf XML-Dokumente. Das API erlaubt, dynamisch die Struktur, das Layout und den Inhalt von XML-Dokumenten zu ändern.
<b>XML</b>	Extensible Markup Language (XML). XML ist eine so genannte META-Sprache zur Beschreibung von Dokumenten. Ein Vorteil von XML ist der vereinfachte Austausch von Daten, da XML-Formate in einer strengen Grammatik definiert werden können und so die Implementierung von zuverlässigen Schnittstellen erlauben.
<b>XML-Schema</b>	Ein XML-Schema beschreibt die Struktur von XML-Dokumenten und erlaubt ihre inhaltliche Überprüfung.
<b>XMLType</b>	XMLType ist ein Oracle Datentyp, durch den SQL-Operationen auf XML-Inhalte genauso ermöglicht werden wie XML-Operationen auf SQL-Inhalte.
<b>XPath</b>	XPath stellt Funktionen und Ausdrücke zur Verfügung, um Knoten innerhalb von XML-Dokumenten zu lokalisieren. Mit XPath können auch Ausdrücke ausgewertet und Berechnungen durchgeführt werden. Es dient als Grundlage für XSLT, XPointer und Xquery.

### Links

- ▶ [1] Artikel: ORDIX News Artikel „XMLType“:  
[http://www.ordix.de/ORDIXNews/1\\_2007/Datenbanken/xmltype.html](http://www.ordix.de/ORDIXNews/1_2007/Datenbanken/xmltype.html)
- ▶ [2] Artikel: ORDIX News Artikel „Performance Tuning“:  
[http://www.ordix.de/ORDIXNews/2\\_2007/Datenbanken/oracle\\_xml\\_performance\\_tuning.html](http://www.ordix.de/ORDIXNews/2_2007/Datenbanken/oracle_xml_performance_tuning.html)
- ▶ [3] Artikel: ORDIX News Artikel „Verwendung von Binary XML“:  
[http://www.ordix.de/ORDIXNews/2\\_2010/Datenbanken/oracle\\_xml.html](http://www.ordix.de/ORDIXNews/2_2010/Datenbanken/oracle_xml.html)
- ▶ [4] Seminarempfehlung „Oracle und XML“:  
<http://training.ordix.de/siteengine/action/load/kategorie/Programmierung/nr/271/index.html>

## Seminarempfehlung: Oracle und XML

▶ **Informationen/Online-Anmeldung:** <http://training.ordix.de/siteengine/action/load/kategorie/Programmierung/nr/271/index.html>

In diesem Seminar werden Sie mit den grundlegenden Eigenschaften und potentiellen Einsatzgebieten von XML unter Oracle vertraut gemacht. Neben den verschiedenen Methoden der Speicherung von XML-Objekten in einer Oracle Datenbank lernen Sie die Möglichkeiten zur Erzeugung und Verarbeitung von XML-Dokumenten kennen.

### Seminarinhalte

- Überblick über die Oracle- und XML-Komponenten
- XML Support innerhalb der Datenbank
- Erzeugen und Auswerten von XML
- Datentyp XMLType
- Einsatz von XML-Schema
- Performance Tuning
- XML-Dokumente generieren und speichern
- 11g Neuheiten

### Termine

17.11. - 19.11.2010 in Wiesbaden

**Seminar-ID:** P-XML-03

**Dauer:** 3 Tage

**Preis pro Teilnehmer:** 1.190,00 € (zzgl. MwSt.)

**Frühbucherpreis:** 1.071,00 € (zzgl. MwSt.)

Wir führen unsere Seminare auch jederzeit an einem geeigneten Ort Ihrer Wahl durch und bringen, wenn nötig, auch das entsprechende Equipment mit. Informieren Sie sich am besten im Internet über unsere Kundenseminare und die mobilen Schulungen: <http://training.ordix.de>.

# Pacemaker - Schrittmacher für Heartbeat?

Dieser Artikel richtet sich an Administratoren, sowie Entscheider und Berater, die sich mit neuen Funktionalitäten der aktuellen Heartbeat-Version vertraut machen wollen.

Häufig haben wir in der ORDIX News bereits über die bei uns und unseren Kunden sehr beliebte Open Source Cluster-Lösung Heartbeat berichtet. In den letzten Jahren wurden zahlreiche Funktionen implementiert und verbessert, so dass sich Heartbeat zu einer Lösung entwickelt hat, die sich vor kommerziellen Cluster-Lösungen nicht verstecken muss. Mit diesem Artikel möchten wir eine Reihe beginnen, die sich mit vergangenen, aktuellen und zukünftigen Funktionen beschäftigt.

## Cluster Manager - Wächter der Verfügbarkeit

Sollen Dienste hochverfügbar angeboten werden, um den Ausfall einer einzelnen Hard- oder Software-Komponente nicht zum Problem für das Unternehmen werden zu lassen, werden häufig zwei oder mehr Systeme zu einem Cluster (Bündel) zusammengefasst. Von außen betrachtet stellt dieser immer einen logischen Server dar, mit gleichem Betriebssystem und Softwarestand. Empfehlenswerterweise besitzen diese Systeme auch die gleiche Hardware-Ausstattung.

Der Cluster Manager selbst ist eine Software für die Verwaltung eines Cluster. Er ist für alle Vorgänge innerhalb des Cluster, wie z. B. die Übernahme von Diensten beim Ausfall eines im Cluster beteiligten Servers oder einer Software-Komponente zuständig. Je nach eingesetzter Cluster-Lösung wird der Cluster Manager häufig auch als Clusterware oder Cluster Framework bezeichnet.

Um einen Ausfall eines Cluster-Knotens zu erkennen, senden die Cluster-Knoten Heartbeat-Pakete über redundante Netzwerkleitungen. Dadurch wird die Erreichbarkeit des Partners überprüft. Um nicht vollkommen abhängig von Netzwerkkabeln und dem TCP/IP-Protokoll zu sein findet bei vielen Cluster-Lösungen auch eine Kommunikation über gemeinsam ge-

nutzte Festplatten statt. Diese werden je nach eingesetzter Lösung als Quorum-, Voting-, oder Coordinatordisks bezeichnet.

## Heartbeat 1 - Einer für alles

Die Entwicklung von Heartbeat [1] in der ersten Version begann 1997/1998 durch Alan Robertson. Der Grundgedanke hierbei war die Überwachung des Partnerknotens und Übernahme der Ressourcen im Fehlerfall. Unterstützt wurden hierbei Cluster mit zwei Knoten. Die Ressourcen konnten Active/Passive oder Active/Active konfiguriert werden. Hat der Cluster eine Ressource auf einem Knoten erfolgreich gestartet, hat dieser nie wieder geprüft ob die Ressource noch läuft. Das Überwachen der Ressourcen musste mit anderen Werkzeugen durchgeführt werden, beliebte Lösungen waren hier `mon` [2] und `monit` [3]. Auf Quorumdisks wurde bei Heartbeat 1 verzichtet, hier wurde eine Lösung namens STONITH implementiert, welche bei einem Ausfall aller Heartbeat-Leitungen den (vermeintlich) ausgefallenen Cluster-Knoten mit einer schaltbaren Steckdosenleiste stromlos schaltet. Die Konfiguration fand mit Hilfe von drei Konfigurationsdateien im Verzeichnis `/etc/ha.d` statt.

- **authkeys:**  
In der Datei `authkeys` kann ein geheimer Schlüssel hinterlegt werden, mit dem eine

gegenseitige Authentifizierung der Cluster-Knoten stattfindet.

- haresources:**  
 In der Datei `haresources` werden die konfigurierten Cluster-Dienste eingetragen und der bevorzugte (primäre) Knoten auf dem die Ressource laufen soll.
- ha.cf:**  
 In der Datei `ha.cf` sind alle notwendigen Einstellungen des Cluster durchzuführen, wie z. B. die Namen der beteiligten Knoten, verwendete Netzwerkkarten für den Heartbeat und die eingesetzten STONITH-Geräte.

Wurde in einer der Konfigurationsdateien eine Änderung durchgeführt, mussten die geänderten Dateien manuell auf den Partnerknoten kopiert werden. Um dem Cluster die Änderungen bekannt zu machen, mussten die Dienste neu gestartet oder geladen werden. In der Prozessliste tauchte der heartbeat-Prozess auf, der sowohl für alle Steuerungsaufgaben, wie auch die Kommunikation der beteiligten Knoten mit heartbeat-Paketen zuständig war. Die Kommunikationsschicht wird hier auch als Consensus Cluster Membership (CCM)-Schicht bezeichnet. Abbildung 1 zeigt den Aufbau des Cluster Manager bei verschiedenen Heartbeat-Versionen.

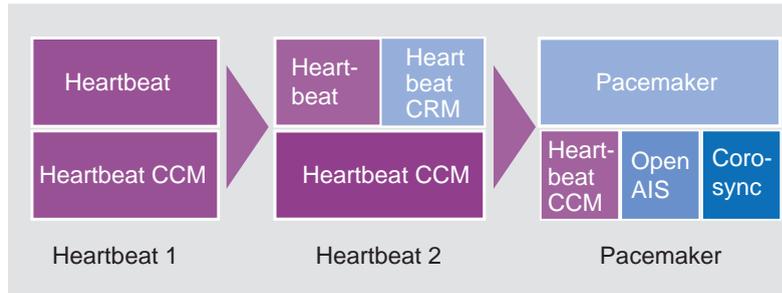


Abb. 1: Aufteilung der Kernkomponenten der unterschiedlichen Versionen.

```
<primitive id="resource_IP2" class="ocf" type="IPAddr"
provider="heartbeat">
  <instance_attributes>
    <attributes>
      <nvpair name="ip" value="192.168.14.102"/>
    </attributes>
  </instance_attributes>
</primitive>
```

Abb. 2: Konfiguration einer IP-Adresse mit XML.

### Heartbeat 2 - Viele Dienste für komplexe Cluster

Mitte des Jahres 2005 wurde Heartbeat in der Version 2 veröffentlicht und brachte viele neue Funktionalitäten mit, die in der Version 1 vermisst wurden. Ein Cluster konnte nun aus mehr als zwei Knoten bestehen, die Entwickler haben Cluster mit bis zu 16 Knoten getestet. Heartbeat fühlte sich nun auch für die Überwachung der Ressourcen zuständig, so dass über frei einstellbare Zeitintervalle ein Monitoring stattfinden konnte. Bei Fehlern kann der Cluster mit einem Neustart der Ressourcen oder einem Failover auf einen anderen Knoten reagieren. Die Ressourcen werden bei Heartbeat 2 Clustern jetzt nicht mehr in der Datei `/etc/ha.d/haresources` durchgeführt. Sie ist einer Konfiguration aus einer XML-Syntax gewichen. Änderungen an Ressourcen werden jetzt immer Online durchgeführt und die Konfiguration wird auf alle Knoten automatisch abgeglichen. Veränderungen der Dateien `authkeys` und `ha.cf` bedürfen weiterhin einem manuellen Kopieren mit anschließendem Reload der Dienste auf den Knoten. Aber immer mehr Funktionalitäten, wie z. B. die Konfiguration der

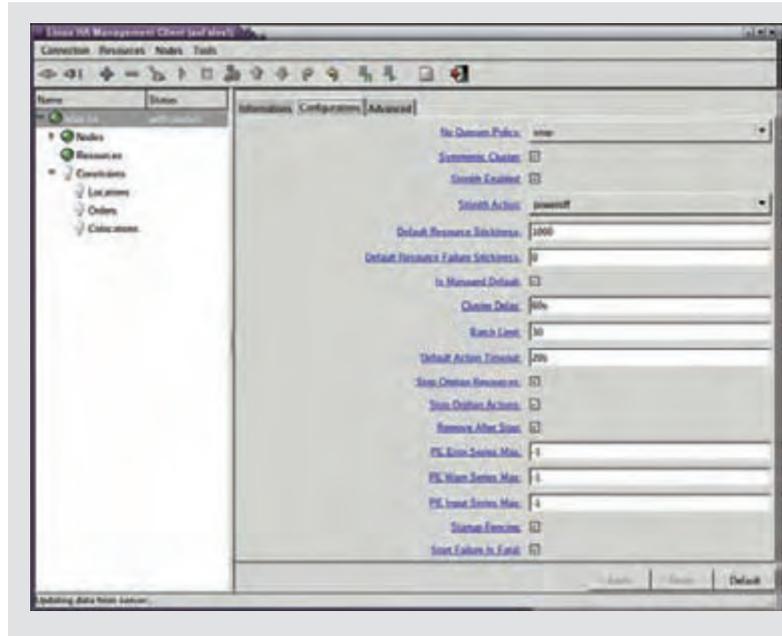


Abb. 3: hb\_gui zur grafischen Konfiguration des Cluster.

STONITH-Geräte wurde nach und nach aus der Datei `ha.cf` in die XML-Konfiguration verlagert. Abbildung 2 zeigt, wie nun eine IP-Adresse als Ressource konfiguriert werden muss, die dann anschließend mit dem Befehl  `cibadmin` in den Cluster eingelesen wird.

```

crm(live)# help

This is the CRM command line interface program.

Available commands:

    cib                manage shadow CIBs
    resource           resources management
    node               nodes management
    options            user preferences
    configure          CRM cluster configuration
    ra                 resource agents information center
    status             show cluster status
    quit,bye,exit     exit the program
    help              show help
    end,cd,up         go back one level

crm(live)#

```

Abb. 4: CRM Shell zur Konfiguration von Cluster und Ressourcen.

Um die Komplexität von Heartbeat 2 Clustern zu bewältigen wurden die Aufgaben auf verschiedene Prozesse aufgeteilt. Der CRM-Prozess (Cluster Ressource Manager) bildet den Hauptprozess. Folgende Aufgaben und Rollen existieren innerhalb eines Heartbeat 2 Cluster.

- Cluster Resource Manager (CRM) - Übernimmt die eigentliche Cluster-Verantwortung, läuft auf allen Knoten als Anlaufstelle für alle Ereignisse und anderen Dienste.
- Local Resource Manager (LRM) - Ist für die lokalen Ressourcen des Knoten zuständig und kennt deren Status.
- Cluster Information Base (CIB) – XML-basierte Konfiguration im Arbeitsspeicher der Cluster-Knoten, die automatisch synchronisiert wird.
- Cluster Consensus Membership (CCM) – Protokollschicht, welche die Kommunikation (heartbeat) der Cluster-Knoten untereinander gewährleistet.
- Policy Engine (PE) - Prüft bei Änderung der Cluster-Konfiguration welche Schritte unternommen werden müssen, um das angestrebte Ziel zu erreichen
- Transition Engine (TE) - Führt die Schritte, wie z. B. das Starten/Stoppen von Ressourcen aus, um das durch die Policy Engine (PE) errechnete Ziel zu erreichen.
- Designated Coordinator (DC) - Die Rolle des DC wird durch die Eigenschaften der PE und TE dargestellt und existiert nur auf einem der beteiligten Cluster-Knoten.

Für die Konfiguration von Ressourcen, die Abhängigkeiten und das Monitoring innerhalb des Cluster haben die Entwickler eine GUI mitgeliefert, mit der die Konfiguration durchgeführt werden kann. Abbildung 3 zeigt die GUI, welche mit dem Befehl `hb_gui` gestartet werden kann.

### Pacemaker - Schrittmacher für Heartbeat

Ab dem Jahr 2007 wurde die Hauptkomponente von Heartbeat 2 der Cluster Ressource Manager herausgelöst und so modifiziert das auf der Cluster Consensus Membership-Schicht auch andere Protokolle als das heartbeat-Protokoll unterstützt werden. Der Name für die neue Komponente war seitdem Pacemaker [4] (Schrittmacher). Aktuell werden von Pacemaker zwei weitere Verfahren zur Kommunikation namens `OpenAIS` und `Corosync` unterstützt. Auf die Konfiguration werden wir in einer kommenden ORDIX News eingehen. Aktuell wird sehr intensiv an Pacemaker weiterentwickelt, während es um den ursprünglichen Teil des Cluster Resource Manager von Heartbeat 2 etwas ruhiger geworden ist. Auch bei Pacemaker findet die Konfiguration weiterhin mit Hilfe von XML statt, wobei sich gegenüber Heartbeat 2 die Syntax etwas verändert hat. Auch die GUI, die Pacemaker mitbringt, wurde optisch gegenüber der Heartbeat 2 GUI etwas abgeändert.

### CRM Shell - Pacemaker spricht mit mir

Um unter Heartbeat 2 Ressourcen zu konfigurieren gab es für den Administrator bislang zwei Möglichkeiten: Entweder er schrieb XML-Schnipsel in Textdateien und importierte diese in den Cluster, oder er nahm die GUI. Dieses hatte drei Nachteile:

- 1) Um die GUI zu starten wird ein X-Server benötigt.
- 2) Nur wenige Administratoren werden sich mit XML auseinander setzen wollen, um eine Ressource zu konfigurieren oder zu verändern.
- 3) Die GUI hatte gerade am Anfang einige Bugs und deckte nicht alle Möglichkeiten der Konfiguration ab.

Mit Pacemaker wird nun eine neue Konfigurationsmethode mitgeliefert, die CRM Shell. Mit dem Befehl `crm` landet man in einer interaktiven Shell (siehe Abbildung 4), in der durch Befehle (`cib`, `resource`, `node`, etc.) wei-

tere Unterpunkte interaktiv aufgerufen werden können. Änderungen der Konfiguration können jetzt durch diese Unterpunkte leicht durchgeführt werden. Natürlich ist auch eine Möglichkeit vorgesehen, mit der per Skript die Konfiguration durchgeführt werden kann und keine manuelle Auswahl notwendig ist. Mit Hilfe der Tab-Taste können, wie in der Bash, Befehle vervollständigt werden. Bei Fehlern in der Konfiguration werden diese farbig gekennzeichnet. Konfigurationsänderungen müssen nicht immer auf die aktive Konfiguration angewendet werden, sondern können an einer `shadow`-Konfiguration getestet werden.

### Fazit und Ausblick

Heartbeat 2 setzte schon viele Wünsche der Administratoren in die Praxis um und stellte gleichzeitig die Weichen für einen zukunftssicheren modularen Aufbau der beliebten Open-Source Cluster-Lösung. Mit Pacemaker wurde eine Unterstützung für alternative Protokolle zur Prüfung der Verfügbarkeit von Knoten implementiert. Mit der CRM Shell bekommt der Administrator zudem, abseits von GUI und XML, ein leistungsstarkes Konfigurationswerkzeug in die Hand. In den kommenden ORDIX News werden wir detailliert auf die Features, die Pacemaker liefert, eingehen.

### Glossar

<b>Heartbeat</b>	Name einer Open Source Cluster-Lösung.
<b>heartbeat</b>	Kommunikation zwischen den Cluster-Knoten zur Prüfung ob der andere Knoten noch erreichbar ist.
<b>Quorumdisk</b>	Komponente eines Cluster zur Wahrung der Datenintegrität bei Ausfall der Netzwerkleitungen.
<b>STONITH</b>	„Shoot the other node in the head“, Möglichkeit bei Heartbeat den Partnerknoten bei Ausfall aller Heartbeat-Leitungen zuverlässig zu eliminieren.

### Links

- ▶ [1] Projektseite von Heartbeat Version 1 und 2: <http://www.linux-ha.org>
- ▶ [2] Projektseite mon: [https://mon.wiki.kernel.org/index.php/Main\\_Page](https://mon.wiki.kernel.org/index.php/Main_Page)
- ▶ [3] Projektseite monit: <http://mmonit.com/monit/>
- ▶ [4] Projektseite Pacemaker: <http://www.clusterlabs.org/>



Christian Fertsch  
([info@ordix.de](mailto:info@ordix.de)).

## Seminarempfehlung: Linux Hochverfügbarkeits-Cluster

▶ **Informationen/Online-Anmeldung:** <http://www.ordix.de/trainingsshop/siteengine/action/load/kategorie/Betriebssysteme/nr/363/index.html>

In diesem Seminar lernen Sie die Grundlagen der Hochverfügbarkeit und der Cluster-Technologie am Beispiel der Heartbeat/Pacemaker Hochverfügbarkeitslösung kennen. Shared Storage unter Einsatz von DRBD wird in diesem Workshop genauso thematisiert, wie der Cluster Interconnect oder die Problematik des SplitBrain. Neben Failover-Techniken wird das Loadbalancing in Active/Active Clustern betrachtet.

### Seminarinhalte

- Grundlagen der Hochverfügbarkeit
- Grundlagen der Cluster-Technologie
- Distributed Replicated Block Device (DRBD)
- Verwaltung der Konfiguration mit drbdlinks und csync2
- Heartbeat, OpenAIS und Corosync
- Heartbeat Version 3
- Fencing mit STONITH
- Erweiterung und Upgrade eines Cluster
- Linux Virtual Server (LVS)
- Die Theorie wird durch praktische Übungen am System vertieft

### Termine

- 27.09. - 01.10.2010 in Wiesbaden
- 08.11. - 12.11.2010 in Wiesbaden

**Seminar-ID:** BS-09

**Dauer:** 5 Tage

**Preis pro Teilnehmer:** 1.890,00 € (zzgl. MwSt.)

**Frühbuchepreis:** 1.701,00 € (zzgl. MwSt.)

Wir führen unsere Seminare auch jederzeit an einem geeigneten Ort Ihrer Wahl durch und bringen, wenn nötig, auch das entsprechende Equipment mit. Informieren Sie sich am besten im Internet über unsere Kundenseminare und die mobilen Schulungen: <http://training.ordix.de>.



DOAG Konferenz und Schulungstag 2010

2010  
**DOAG**  
Konferenz + Ausstellung

## ORDIX und coniatos präsentieren ihr Know-how

In diesem Jahr trifft sich die Oracle Community vom 16. bis 19. November 2010 im CongressCenter Nürnberg Ost. In weit mehr als 400 Fachvorträgen bietet die DOAG Konferenz + Ausstellung 2010 wie gewohnt Neuigkeiten zu Oracle und praxisnahe Erfahrungsberichte. ORDIX ist in diesem Jahr mit neun hochkarätigen Vorträgen und einer Schulung dabei. Darüber hinaus präsentiert unser Tochterunternehmen, die coniatos AG aus Wiesbaden, ihr Know-how mit zwei Vorträgen.

### Hochverfügbarkeit mit Data Guard – Möglichkeiten und Grenzen

Täglich wird der Datenbankadministrator mit neuen Anforderungen konfrontiert. Vor allem soll alles hochverfügbar und ausfallsicher sein. Doch was ist eigentlich Hochverfügbarkeit? Welche Möglichkeiten bietet Oracle zu dem Thema? Wo ist die Abgrenzung zwischen Data Guard und RAC? Was können wir mit Data Guard leisten? Anhand von Beispielen aus der täglichen Praxis geht der Referent Andreas Kother in seinem Vortrag auf diese Fragen ein und gibt die entsprechenden Antworten.

### Partitioning – Technik und Anwendungsbeispiele

Partitioning gibt es bei Oracle seit der Version 8.0. In vielen Projekten werden unterschiedliche Ansätze angewendet (Range/Hash/List). Im Vortrag werden die möglichen Techniken - auch unter Einbeziehung der Version 11g - dargestellt und Anwendungsbeispiele gegeben. Anhand einer Demo zeigt Klaus Reimers, darüber hinaus die vielen Möglichkeiten der Partitionierung im Zusammenhang.

### PL/SQL-Tuning

Wie die Laufzeit von PL/SQL-Applikationen deutlich beschleunigt werden kann, erläutert der Referent Markus Fiegler in seinem Vortrag. Dabei werden unter anderem Funktionen, wie PL/SQL Inlining, PL/SQL Function Result Cache oder die XML-Speicherungsform Binary XML beleuchtet, die zur Performance-Steigerung von PL/SQL-Applikationen führen. Darüber hinaus liefert Markus Fiegler einige Tipps und Tricks aus der Praxis bezogen auf Performance-Verbesserungen von PL/SQL-Programmen.

### Eine kurze Geschichte der Zeit Datums- und Zeitverwaltung in Oracle

Mit diesem Vortrag stellt Martin Hoermann die zeitbezogenen Oracle Datentypen DATE, TIMESTAMP, TIMESTAMP WITH [LOCAL] TIME ZONE, INTERVAL und die dazugehörigen Funktionen vor. Nach einer Einführung in die einfache Zeitarithmetik behandelt der Vortrag die interne Verwaltung der Daten, die Darstellung von Zeiten sowie die Handhabung und Konvertierung zwischen verschiedenen

Zeitzonen. Darüber hinaus geht der Referent auf Besonderheiten verschiedener Kalender, Zeit- und Datumsumstellungen ein. Eine Betrachtung unterschiedlicher Fehlermeldungen im Zusammenhang mit Datumswerten rundet den Vortrag ab.

### Endlich dabei – die Criteria-QueryAPI in der JPA 2.0

Viele Entwickler haben bisher die Criteria Query API – wie sie in Hibernate zur Verfügung steht – in der Java Persistence API (JPA) vermisst. Nun ist diese API in der JPA 2.0 integriert. Sie bietet ein objektorientiertes Verfahren zum dynamischen Aufbau von Abfragen. Sie ist damit kein Ersatz für die mächtige Java Persistence Query Language (JPQL), sondern vielmehr eine sinnvolle Ergänzung, wenn es um dynamische Zusammenstellung von Abfragen in einer Anwendung geht. Der Referent Oliver Kaluza beleuchtet in seinem Vortrag die Funktionalitäten der Criteria-Query API und stellt die Vor- und Nachteile gegenüber der JPQL dar. Zudem gibt er Empfehlungen, wann der Einsatz der beiden APIs sinnvoll ist.

### Kontext und Dependency Management in Java EE - Einführung in JSR-299

Der JSR-299 ist ein Standard für Kontext und Dependency Injection in JEE6-Umgebungen. Kontext und Dependency Management Mechanismen erlauben eine effiziente Kopplung verschiedenster Schichten einer JEE 6 Applikation. Insbesondere Zugriffe von Presentation (JSF/JSP) auf Business Tier (EJB) lassen sich erheblich vereinfachen. Eine erste Implementierung des Standards ist JBoss Weld. Björn Konrad zeigt anhand von JBoss Weld, wie sich der JSR-299 in gängigen Java Enterprise Applikationen einsetzen lässt.

### Be dynamic! RichClient-Funktionalitäten per Google Web Toolkit

In modernen Web-Anwendungen wird die Nachfrage nach RichClient-Funktionalitäten zur Umsetzung spezieller Anforderungen immer größer. Diese Funktionalitäten sollen dem Nutzer jedoch ohne die Installation von Browser Plugins zur Verfügung stehen. Das Google Webtoolkit (GWT) ermöglicht die Erstellung von dynamischen Web-Applikationen und kann durch eine Vielzahl von Erweiterungen spezielle Anforderungen besser abdecken als andere Frameworks. Dennoch ist gerade bei unterschiedlichen Anforderungspro-

## Die Vorträge zum Thema Oracle im Überblick



### Hochverfügbarkeit mit Data Guard

Dienstag, 16.11.2010  
12:00 – 12:45 Uhr, Raum 19  
Referent: Andreas Kother,  
ORDIX AG



### Partitioning – Technik und Anwendungsbeispiele

Mittwoch, 17.11.2010  
10:00 – 10:45 Uhr, Raum 1  
Referent: Klaus Reimers,  
ORDIX AG



### PL/SQL-Tuning

Mittwoch, 17.11.2010  
10:00 – 10:45 Uhr, Raum 6  
Referent: Markus Fiegler,  
ORDIX AG



### Datums- und Zeitverwaltung in Oracle

Donnerstag, 18.11.2010  
10:00 – 10:45 Uhr, Raum 1  
Referent: Martin Hoermann,  
ORDIX AG

## Die Vorträge zum Thema Java im Überblick



### Endlich dabei – die Criteria-QueryAPI in der JPA 2.0

Dienstag, 16.11.2010  
16:00 – 16:45 Uhr, Raum 17  
Referent: Oliver Kaluza,  
ORDIX AG



### Kontext und Dependency Management in Java EE - Einführung in JSR-299

Mittwoch, 17.11.2010  
13:00 – 13:45 Uhr, Raum 17  
Referent: Björn Konrad,  
ORDIX AG



### RichClient- Funktionalitäten per Google Web Toolkit

Donnerstag, 18.11.2010  
15:00 – 15:45 Uhr, Raum 17  
Referent: Julian Gärtner,  
ORDIX AG



### Versionierung von Webservice-Schnittstellen

Mittwoch, 17.11.2010  
16:00 – 16:45 Uhr, Raum 17  
Referent: Christian Wiesing,  
ORDIX AG

### Weitere Vorträge im Überblick



#### MySQL Backup- und Recovery-Strategien

Dienstag, 16.11.2010  
13:00 – 13:45 Uhr, Raum 4  
Referent: Matthias Jung,  
ORDIX AG



#### Datenbank- und Betriebsoptimierung dank Agilem Projektmanagement

Mittwoch, 17.11.2010  
12:00 – 12:45 Uhr, Raum 4  
Referent: Lars Eisenblatt,  
coniatos AG



#### Strategic Fit - Strategischer Wertbeitrag von Enterprise Architectures

Donnerstag, 18.11.2010  
13:00 – 13:45 Uhr, Raum 8  
Referent: Matthias Heint,  
coniatos AG

filen die Kombination und Integration mit anderen Frameworks von Vorteil. Der Referent Julian Gärtner zeigt anhand eines Beispiels aus der Praxis die erfolgreiche Integration einer RichClient-Applikation basierend auf dem Google Webtoolkit in einer klassischen Web-Anwendung.

#### Versionierung von Webservice-Schnittstellen

Webservice-Schnittstellen, die einmal produktiv genommen wurden, können in der Regel nicht mehr geändert werden, ohne gleichzeitig auch die zugreifenden Systeme anzupassen. Um dennoch neue Funktionalitäten zur Verfügung stellen zu können, ist eine Versionierung der Schnittstellen notwendig. Christian Wiesing stellt in seinem Vortrag anhand von praktischen Erfahrungen und konkreten Beispielen ein Konzept zur Versionierung von Webservice-Schnittstellen vor und geht auf mögliche Probleme und Fallstricke ein.

#### MySQL Backup- und Recovery-Strategien

Eine der wichtigsten Aufgaben eines Datenbankadministrators ist es, ein zuverlässiges und effizientes Sicherungskonzept zu erstellen.

Matthias Jung, stellt mit seinem Vortrag die Standardwerkzeuge und Funktionen von MySQL vor, mit deren Hilfe sich auch anspruchsvolle Backup-Anforderungen realisieren lassen. Er gibt Antworten auf die von Kunden oft gestellten Fragen:

- Wie lassen sich MySQL-Datenbanken online ohne störenden Einfluss auf den Betrieb der Datenbank sichern?
- Wie lässt sich ein Point-in-Time-Recovery realisieren?
- Lohnt sich der Einsatz von Replikationsmechanismen im Rahmen eines Back-up-Konzeptes?
- Wie kann die vorhandene technische Infrastruktur (Storage, Betriebssysteme) sinnvoll im Rahmen dieses Konzeptes eingebunden werden?

#### Flink und flexibel - Datenbank- und Betriebsoptimierung dank Agilem Projektmanagement

In einem IT-Service-Unternehmen wird der Rollout und der Betrieb von Datenbanken ständig vor neue Herausforderungen gestellt. Ressourcen-Mangel und enge Zeitraster stellen große Risikofaktoren für die zeitgerechte Umsetzungen dar. Hier gilt es, schnell und flexibel auf kurzfristig auftretende Anforderungen reagieren zu können. Agile Verfahren im Projektmanagement können diesen Forderungen Rechnung tragen. Lars Eisenblatt, Vorstand der coniatos AG, liefert in seinem Vortrag einen Erfahrungsbericht aus einem Optimierungsprojekt für einen umfangreichen Datenbankbetrieb und zeigt auf, wie agiles Projektmanagement (SCRUM) auch in Projekten im IT-Betrieb erfolgreich angewendet werden kann. Darüber hinaus wird die Interaktion mit dem klassischen Projektcontrolling beleuchtet. Abgerundet wird der Vortrag durch die Vorstellung interessanter Lessons Learned.

#### Strategic Fit - Strategischer Wertbeitrag von Enterprise Architectures

Erfolgreiche IT-Unternehmensarchitekturen werden aus der Geschäftsstrategie des Unternehmens abgeleitet. Auf diese Weise wird das entscheidende Fundament für eine langfristig erfolgreiche Modellierung von IT-Landschaften gebildet. In der Praxis stellt sich die Frage, wie der strategische Wertbeitrag von IT-Unternehmensarchitekturen ermittelt werden kann. Matthias Heint, coniatos AG,

stellt die Vorgehensweisen mit dem Einsatz wichtiger Kennzahlen zur Bewertung der IT-Strategiekonformität vor und zeigt, wie die erfolgreiche Umsetzung einer IT-Strategie sichergestellt werden kann.

### DOAG Schulungstag

Direkt im Anschluss an die DOAG Konferenz und Ausstellung lädt Sie die ORDIX AG am 19. November 2010 von 9 bis 16 Uhr zu einer eintägigen Schulung ebenfalls in das CCN Ost ein. Der Referent, Klaus Reimers, stellt in seinem Seminar das Thema „AWR und ASH Analyse“ vor.

In diesem Seminar lernen Sie den Umgang mit den seit Oracle 10 zur Verfügung gestellten Daten des Active Workload Repository (AWR) und der Active Session History (ASH). Sie lernen, historische Daten zu analysieren und zu interpretieren. Des Weiteren werden die automatisierten Schnittstellen des AWR zur Performance-Analyse und dem Tuning vorgestellt. Weitere Informationen und Anmeldung über die Konferenzseiten der DOAG [5].

**Wir freuen uns auf Sie bei der DOAG 2010 in Nürnberg!**

#### Links

- ▶ [1] Rückblick DOAG Konferenz 2009:  
[http://www.ordix.de/ORDIXNews/1\\_2010/Aktuell/doag\\_konferenz.html](http://www.ordix.de/ORDIXNews/1_2010/Aktuell/doag_konferenz.html)
- ▶ [2] Oracle Dienstleistungen der ORDIX AG:  
<http://www.ordix.de/Portfolio/Technologien/Datenbanken/oracle.html>
- ▶ [3] Oracle Seminare:  
[http://training.ordix.de/index\\_oracle.htm](http://training.ordix.de/index_oracle.htm)
- ▶ [4] Java Seminare:  
<http://training.ordix.de/siteengine/action/load/kategorie/Java-JEE/index.html>
- ▶ [5] DOAG Webseite:  
<http://www.doag.org/konferenz/doag/2010/>
- ▶ [6] coniatos AG Webseite:  
<http://www.coniatos.de>



## DOAG-Schulungstag: AWR und ASH Analyse

#### Inhalte:

- AWR Überblick / ASH Überblick
- AWR Prozesse, Konfiguration, Snapshots
- AWR Datenmodell
- Interpretation der Standardreports AWR und ASH
- Beispielhafte Analyse anhand von praxisorientierten Fragestellungen mit AWR (Fragestellungen können von den Teilnehmern eingebracht werden)
- Prozedurale Schnittstelle von AWR und ADDM
- SQL Tuning Advisor

#### Zeit/Ort:

19. November 2010  
09:00 - 16.00 Uhr  
CongressCenter Ost, Nürnberg

**Referent:** Klaus Reimers

#### Preis:

450 € zuzügl. MwSt.

Melden Sie sich gleich an unter:

<http://www.doag.org/konferenz/doag/2010/>

Weitere Seminarangebote finden Sie im ORDIX-Trainingsshop unter: <http://training.ordix.de>.



Perl und DB2 im Verbund (Teil III)

## Perl-Skripte konfigurieren

Dieser Artikel richtet sich an DB2-Administratoren, die Perl als Skriptsprache verwenden sowie alle Perl-Programmierer, die die Konfigurierbarkeit ihrer Skripte verbessern wollen.

Ein Skript ist geschrieben, die Verbindung zur Datenbank funktioniert und ein wenig Ordnung haben wir dank `SQL::Library` auch geschaffen. In diesem Artikel werden wir unsere Skripte für die Nutzen bei einer größeren Anzahl von Datenbanken vorbereiten, indem wir die Konfiguration der Datenbank-Verbindungen mit Hilfe von `XML::Simple` auslagern.

### Das Problem

In den bisherigen Artikeln dieser Reihe haben wir die Datenbank-Verbindungen in den Skripten „hart“ codiert. Werden diese Skripte in weiteren Datenbanken eingesetzt, muss man sich Gedanken über die Konfigurierbarkeit machen. Es stehen uns verschiedene Ansätze zur Verfügung, die im Laufe dieses Artikels erläutert werden.

### @ARGV: Die „faule“ Variante

Das Array `@ARGV` enthält immer sämtliche Argumente der Kommandozeilen (siehe Abbildung 1).

Wir könnten daher einfach mit `$ARGV[0]` auf z. B. den Datenbanknamen zugreifen, mit `$ARGV[1]` auf den Benutzer und mit `$ARGV[2]` auf das Passwort. Dies hat allerdings folgende Nachteile:

- Die Reihenfolge der Argumente ist festgelegt und muss dem Anwender bekannt sein.
- Das Passwort ist potentiell in der Prozessliste für jedermann sichtbar.
- Es entsteht zuviel Programmierarbeit.

### Getopt: Etwas mehr Kontrolle

Bei der Verwendung von Optionsschaltern mit `Getopt::Std` oder `Getopt::Long` ersparen wir dem User zumindest das Einhalten einer bestimmten Reihenfolge und können z. B. eine Online-Hilfe einbauen (siehe Abbildung 2). Die übrigen Probleme bleiben:

- Alle Connect-Daten für alle Datenbanken müssen bekannt sein.
- Das Passwort ist potentiell in der Prozessliste für jedermann sichtbar.
- Es entsteht noch mehr Tipparbeit.

### Externe CSV-Datei: Wider das Vergessen

Eine externe CSV-Datei ist eine bewährte und einfache Möglichkeit um Konfigurationsdaten zu externalisieren. Ein unbestreitbarer Vorteil liegt in der Verfügbarkeit ohne die zusätzliche Installation eines speziellen Perl-Moduls.

Im Beispiel (siehe Abbildung 3) übergeben wir nur noch den Datenbank-Alias. Als nachteilig erweist sich, dass die Konfigurationsdatei nicht selbstdokumentierend ist.

### XML::Simple: Die elegante Lösung

XML hat gegenüber CSV den unschlagbaren Vorteil, bei entsprechend sinnvoller Wahl der Tags selbstdokumentierend zu sein. Zudem ist es erheblich einfacher, eine XML-Datenstruktur nachträglich zu erweitern, da der Zugriff auf die Daten standardisiert ist.

**XML::Simple** macht seinem Namen alle Ehre und liefert eine wirklich einfache Schnittstelle zur Verarbeitung von XML-Daten.

Um DTDs und Ähnliches müssen wir uns nicht kümmern. Dies mag einigen auch als Nachteil erscheinen. Für unsere Zwecke reichen die Fähigkeiten von **XML::Simple** aber bei Weitem aus.

### Installation und Aufruf

Nach der Installation (z. B. mit `perl-MCPAN-eshell`) bzw. nach den Anweisungen der README-Datei kann das Modul mit `use XML::Simple` eingebunden werden.

### Das Prinzip

**XML::Simple** kann XML-Dateien sowohl lesen als auch schreiben. Beim Lesen werden die Werte aus der XML-Datei einer Hash-Referenz zugewiesen. Beim Schreiben wird der Inhalt eines Hash in eine XML-Struktur umgewandelt.

Die beiden wesentlichen zur Verfügung gestellten Methoden hierfür sind `XMLin()` und `XMLout()`.

### Die Konfigurationsdatei

Es ist Ihnen überlassen, wie Sie die Konfigurationsdatei gestalten und welche zusätzlichen

```
my $database_name=$ARGV[0];
my $database_user=$ARGV[1];
my $database_pwd=$ARGV[2];
```

Abb. 1: Argumente über @ARGV einlesen. Schnell gemacht aber unzuverlässig.

```
use Getopt::Std;
our ($opt_d, $opt_u, $opt_p);
getopts('d:u:p:') or die "usage: $0 -d database -u user -p
passwd\n";
my ($dbname, $dbuser, $dbpass) = ($opt_d, $opt_u, $opt_p);
```

Abb. 2: Verwendung von Getopt. Fordert aber ein gutes Gedächtnis.

```
# dbconfig.csv
sample,db2inst1,1234
test,db2inst2,admin

# get_csv.pl
my $infile = 'dbconfig.csv';
my $dbname = lc( $ARGV[0] ); # Datenbankname wird übergeben
my %config; # enthält später die Konfiguration

# Konfiguration einlesen:
open( CONFIG, "<$infile" ) or die "$!\n";
while (<CONFIG>) {
    my ( $dbname, $dbuser, $dbpass ) = split( /,/, $_ );
    chomp( $dbpass );
    $dbname = lc($dbname);
    $config{$dbname}{dbuser} = $dbuser;
    $config{$dbname}{dbpass} = $dbpass;
}
my $dbuser = $config{$dbname}{dbuser};
my $dbpass = $config{$dbname}{dbpass};
```

Abb. 3: Der Klassiker. Geht immer. Allerdings wenig flexibel.

```
<config logdir="/tmp" logfile="$0.log">
  <database name="sample">
    <user>db2inst1</user>
    <pass>1234</pass>
    <description>Beispiel-Datenbank für DB2</description>
    <description>generiert mit db2sample</description>
  </database>
  <database name="test">
    <user>db2inst2</user>
    <pass>supergeheim</pass>
    <description>Test-Datenbank</description>
    <description>Development</description>
    <description>Buchhaltungssoftware</description>
  </database>
</config>
```

Abb. 4: Eine Beispiel-Konfiguration für 2 Datenbanken.

```
$config = {
  'database' => {
    'test' => {
      'pass' => 'supergeheim',
      'user' => 'db2inst2',
      'description' =>
        [ 'Test-Datenbank', 'Development', 'Buchhaltungssoftware' ]
    },
    'sample' => {
      'pass' => '1234',
      'user' => 'db2inst1',
      'description' =>
        [ 'Beispiel-Datenbank für DB2', 'generiert mit db2sample' ]
    }
  },
  'logdir' => '/tmp',
  'logfile' => '$0.log'
};
```

Abb. 5: Die von `XMLin()` generierte Datenstruktur.

```
use XML::Simple;
my $dbname = $ARGV[0];
my ($dbuser, $dbpass);
my $config = XMLin('dbconfig.xml') or die "$!\n";
if ( defined $config->{database}{$dbname} ) {
  $dbuser = $config->{database}{$dbname}{user};
  $dbpass = $config->{database}{$dbname}{pass};
}
else {
  print "Datenbank $dbname nicht konfiguriert\n"; exit 1;
}
```

Abb. 6: Einfacher und direkter Zugriff.

## Links

- ▶ [1] ORDIX News Artikelreihe „DB2 und Perl“:  
<http://www.ordix.de/ORDIXNews/artikelreihen.html#DB2Perl>
- ▶ [2] Dokumentation zu „XML::Simple“: <http://bit.ly/d8irkw>
- ▶ [3] Seminarempfehlungen „IBM DB2 für Linux/Unix/Windows Administration“:  
<http://training.ordix.de/siteengine/action/load/kategorie/Datenbanken/nr/279/index.html>

Informationen Sie als sinnvoll erachten (siehe Abbildung 4). Eine spätere Erweiterung ist jederzeit möglich. So könnten Sie eine zentrale Konfigurationsdatei für sämtliche Skripte in Ihrer RZ-Umgebung erstellen.

## Konfiguration einlesen

Abbildung 5 zeigt die von `XMLin()` generierte Datenstruktur. Die Elemente des Hash sind entsprechend den XML-Tags in der Konfigurationsdatei benannt. Abbildung 6 zeigt den einfachen und übersichtlichen Zugriff auf die Daten.

Mit `XML::Simple` und `SQL::Library` haben wir zwei einfach zu bedienende Technologien vorgestellt, mit denen wir den Aufwand für Pflege und Anpassung unserer Datenbank-Skripte optimieren können. In der nächsten Folge der Reihe werden wir uns zumindest zeitweise von DB2 entfernen und den Fokus noch stärker auf Perl-Technologien verlagern. Das zentrale Thema wird die Erstellung eigener Nagios-Plugins unter Perl sein.



Norbert Munkel  
([info@ordix.de](mailto:info@ordix.de)).

## Seminarempfehlung: Perl Programmierung Grundlagen

- ▶ **Informationen/Online-Anmeldung:** <http://training.ordix.de/siteengine/action/load/kategorie/Programmierung/nr/189/index.html>

In diesem Seminar werden Sie mit der Programmiersprache Perl vertraut gemacht. Neben der Einführung in die verschiedenen Daten- und Kontrollstrukturen werden auch die für Systemadministratoren wichtigen Themen "reguläre Ausdrücke" und "Zugriff auf Dateien" behandelt. Anhand von Beispielen wird der Einsatz von Perl im CGI- und Datenbank-Umfeld gezeigt.

### Seminarinhalte

- Variablentypen unter Perl
- Ablaufsteuerung von Programmen
- Reguläre Ausdrücke
- Arbeiten mit Zeichenketten (String-Manipulation)
- Verwendung/Erstellung von Funktionen und Modulen
- Zugriff auf Dateien und Verzeichnisse
- Arbeiten mit der Online-Hilfe von Perl
- Einsatz von Perl im CGI-Umfeld und im Datenbank-Umfeld
- Vertiefung der Theorie durch praktische Übungen und Beispiele

### Termine

04.10. - 08.10.2010 in Wiesbaden

**Seminar-ID:** P-PERL-01

**Dauer:** 5 Tage

**Preis pro Teilnehmer:** 1.690,00 € (zzgl. MwSt.)

**Frühbucherpreis:** 1.521,00 € (zzgl. MwSt.)

Wir führen unsere Seminare auch jederzeit an einem geeigneten Ort Ihrer Wahl durch und bringen, wenn nötig, auch das entsprechende Equipment mit. Informieren Sie sich am besten im Internet über unsere Kundenseminare und die mobilen Schulungen: <http://training.ordix.de>.

Larry Ratlos

# Larry hat Probleme mit den Zeitzonen



Larry hat heute ein Problem mit einer Oracle-Datenbank. Nachdem ihm beim letzten Problem mit der Select-Box so viele Leser weiter helfen konnten, ist Larry zuversichtlich, das ihm auch diesmal die Leser der ORDIX News zur Seite stehen werden.

## TIMESTAMP (ohne Timezone)

Larry hat eine Oracle-Datenbank in der Zeiten in einer Spalte vom Datentyp TIMESTAMP (ohne Timezone) abgespeichert werden (siehe Abbildungen 1 und 2). Die Zeiten sollen zukünftig in UTC abgespeichert werden.

Larry möchte nun Einträge in der lokalen Zeit (z. B. Europe/Berlin) anzeigen und von der lokalen Zeit die UTC-Zeit abspeichern. Er hat eine zeitlang in der Dokumentation gesucht, aber nichts passendes gefunden.

## Können Sie Larry helfen?

Schicken Sie Larry Ihre richtige Lösung, damit er die korrekte Anzeigen in seiner Datenbank erhält. Auf Ihren Lösungsvorschlag freut sich Larry **bis zum 18. Oktober 2010 an kniffel@ordix.de**.

Sollten Sie das Zeitzonen-Problem nicht lösen können, dann besuchen Sie zusammen mit Larry den Vortrag von Martin Hoermann auf der DOAG Konferenz (s. Seite 14) zum Thema: „Eine kurze Geschichte der Oracle Zeit“.

## Larry bedankt sich bei allen Einsendern für die PHP-Codes für sein Select-Box-Problem

Viele unserer Leser konnten Larry mit seinem Select-Box Problem helfen. Larry bedankt sich bei Frau Göb, Herrn Wolf, Herrn Linten und Herrn Riemer, die diesmal die schnellsten Einsender waren.

Den PHP-Code, der zur richtigen Anzeige in der Select-Box führt, finden Sie im rechten Fenster.

```
CREATE TABLE zeiten( utc_zeit TIMESTAMP );
utc_zeit => Europe/Berlin
Europe/Berlin => utc_zeit
```

Abb. 1: Code aus der Oracle-Datenbank.

```
SQL> SELECT sessiontimezone, dbtimezone FROM dual;
SESSIONTIMEZONE      DBTIME
-----
+02:00                +00:00
CREATE TABLE zeiten( utc_zeit TIMESTAMP );
```

Abb. 2: Aktuelle Einstellung der Datenbank.

## Lösung aus der Ausgabe 2/2010

Die problematische Stelle ist in Zeile 6 des Skriptes zu finden:

```
- $db_array[$j][] = explode(",|", $lines[$j]);
+ $db_array[$j] = explode(",|", $lines[$j]);
```

Somit lautet der vollständige PHP-Code in korrekter Form:

```
<?php
$lines = file("../db_name.txt");
for ($j = 0; $j < count($lines); $j++) {
    $db_array[$j] = explode(",|", $lines[$j]);
}
echo ,
<label for="db_server">Datenbank</label>
<select name="db_server">`;

for ($i = 0; $i < count($db_array); $i++) {
    echo ,<option name="`. $db_array[$i][0]. `"'>`. $db_array[$i][1]. `</op-
tion>`;
}
echo ,</select>`;
?>
```

Dieser ergibt den erwarteten HTML-Output:

```
<label for="db_select">Datenbank:</label>
<select name="db_select"><option name="prod_db_1">Produktionsdatenbank 1
</option><option name="entw_db_1">Entwicklungsdatenbank 1
</option><option name="entw_db_1">Entwicklungsdatenbank 2
</option></select>
```

September 2010 - Februar 2011

	Dauer	Preis	Sept.		Oktober			November			Dezember			Januar			Februar										
			KW 37	KW 38	KW 39	KW 40	KW 41	KW 42	KW 43	KW 44	KW 45	KW 46	KW 47	KW 48	KW 49	KW 50	KW 51	KW 52	KW 1	KW 2	KW 3	KW 4	KW 5	KW 6	KW 7	KW 8	KW 9
<b>Datenbanken</b>																											
Datenbank-Hochverfügbarkeitslösungen für Entscheider	1 Tag	550,00 €																									
Datenbank-Modellierung	2 Tage	890,00 €																									
Oracle SQL	5 Tage	1.890,00 €																									
Oracle SQL für Experten	3 Tage	1.290,00 €																									
Oracle Datenbankprogrammierung mit PL/SQL Grundlagen	5 Tage	1.890,00 €																									
Oracle Datenbankprogrammierung mit PL/SQL Aufbau	5 Tage	1.890,00 €																									
Objektorientierung in Oracle	5 Tage	1.990,00 €																									
Oracle Datenbankadministration Grundlagen	5 Tage	1.990,00 €																									
Oracle Datenbankadministration Aufbau	5 Tage	1.990,00 €																									
Manuelles Oracle Backup und Recovery	5 Tage	1.990,00 €																									
Oracle Tuning und Monitoring	5 Tage	1.990,00 €																									
Oracle AWR und ASH Analyse und Interpretation	3 Tage	1.290,00 €																									
Oracle Troubleshooting Workshop	5 Tage	1.990,00 €																									
Oracle 11gR2 RAC und Grid Infrastructure	5 Tage	1.990,00 €																									
Oracle 11g Neuheiten	5 Tage	1.990,00 €																									
Oracle Security	4 Tage	1.590,00 €																									
Oracle Data Guard	4 Tage	1.590,00 €																									
Oracle Backup und Recovery mit RMAN	4 Tage	1.590,00 €																									
Oracle Grid Control	3 Tage	1.290,00 €																									
Oracle Streams Advanced Queuing	3 Tage	1.290,00 €																									
Oracle Replikation	3 Tage	1.290,00 €																									
Oracle Migration und Patching	3 Tage	1.290,00 €																									
Oracle Capacity Planning (nur auf Anfrage)	3 Tage	1.290,00 €																									
IBM Informix SQL	5 Tage	1.790,00 €																									
IBM Informix Dynamic Server Administration	5 Tage	1.990,00 €																									
IBM Informix Dynamic Server Tuning und Monitoring	5 Tage	1.990,00 €																									
IBM Informix Dynamic Server Backup und Recovery	3 Tage	1.290,00 €																									
IBM Informix Dynamic Server HV-Technologien unter Unix	4 Tage	1.590,00 €																									
IBM DB2 für Linux/Unix/Windows SQL Grundlagen	5 Tage	1.890,00 €																									
IBM DB2 für Linux/Unix/Windows Administration	5 Tage	1.990,00 €																									
IBM DB2 für Linux/Unix/Windows Monitoring und Tuning	3 Tage	1.290,00 €																									
MySQL Administration	3 Tage	1.190,00 €																									
Microsoft SQL Server Administration	5 Tage	1.790,00 €																									
<b>Programmierung</b>																											
PHP Programmierung Grundlagen	5 Tage	1.690,00 €																									
PHP Programmierung Aufbau	3 Tage	1.190,00 €																									
Perl Programmierung Grundlagen	5 Tage	1.690,00 €																									
Perl Programmierung Aufbau	5 Tage	1.690,00 €																									
Shell, Awk und Sed	5 Tage	1.690,00 €																									
AWK Intensiv-Workshop	3 Tage	1.190,00 €																									
Einführung in XML	3 Tage	1.190,00 €																									
XML Programmierung unter Java mit DOM und SAX	2 Tage	890,00 €																									
Oracle und XML	3 Tage	1.190,00 €																									
<b>Web- und Applikations-Server</b>																											
Apache Web-Server Installation und Administration	3 Tage	1.190,00 €																									
Tomcat Konfiguration und Administration	3 Tage	1.190,00 €																									
WebSphere Application Server Installation/Administration	3 Tage	1.390,00 €																									
Administration und Konfiguration für JBoss	3 Tage	1.190,00 €																									

Online-Anmeldung und stets aktuelle Seminarinhalte und Termine unter: <http://training.ordix.de>

	Dauer	Preis	Sept.		Oktober			November			Dezember			Januar		Februar											
			KW 37	KW 38	KW 39	KW 40	KW 41	KW 42	KW 43	KW 44	KW 45	KW 46	KW 47	KW 48	KW 49	KW 50	KW 51	KW 52	KW 1	KW 2	KW 3	KW 4	KW 5	KW 6	KW 7	KW 8	KW 9
<b>Java-JEE</b>																											
Einführung in die objektorientierte Programmierung	3 Tage	1.190,00 €																									
Java Programmierung Grundlagen	5 Tage	1.690,00 €																									
Java Programmierung Aufbau	5 Tage	1.690,00 €																									
Java GUI Entwicklung mit Swing	5 Tage	1.690,00 €																									
JEE für Entscheider	1 Tag	590,00 €																									
Einführung in JEE	3 Tage	1.290,00 €																									
JSP und Servlet Programmierung	5 Tage	1.590,00 €																									
EJB Programmierung	5 Tage	1.590,00 €																									
Web-Anwendungen mit JavaServer Faces (JSF)	5 Tage	1.590,00 €																									
Entwickeln mit dem Spring-Framework	3 Tage	1.190,00 €																									
Java Web Services	3 Tage	1.190,00 €																									
Hibernate und die Java Persistence API	5 Tage	1.690,00 €																									
Java Performance Tuning	3 Tage	1.290,00 €																									
<b>Betriebssysteme</b>																											
Unix/Linux Grundlagen für Einsteiger	5 Tage	1.690,00 €																									
Linux Systemadministration	5 Tage	1.690,00 €																									
Server-Virtualisierung mit XEN (nur noch auf Anfrage)	3 Tage	1.190,00 €																									
Linux Hochverfügbarkeits-Cluster	3 Tage	1.290,00 €																									
OpenLDAP - Praxiseinsatz im Netzwerk	4 Tage	1.490,00 €																									
Solaris Systemadministration Grundlagen	5 Tage	1.990,00 €																									
Solaris Systemadministration Aufbau	5 Tage	1.990,00 €																									
Solaris kompakt für erfahrene Unix-Administratoren	5 Tage	1.990,00 €																									
Solaris 10 für erfahrene Systemadministratoren	5 Tage	1.990,00 €																									
Solaris Containers	3 Tage	1.290,00 €																									
Zettabyte Filesystem (ZFS) Workshop	2 Tage	890,00 €																									
IBM AIX Systemadministration Grundlagen	5 Tage	1.990,00 €																									
IBM AIX Installation, Backup und Recovery	3 Tage	1.290,00 €																									
IPv6 für Entscheider	1 Tag	590,00 €																									
IPv6 für Systemadministratoren	3 Tage	1.190,00 €																									
<b>Systemmanagement</b>																											
Systemüberwachung mit Nagios Grundlagen	3 Tage	1.190,00 €																									
Systemüberwachung mit Nagios Aufbau	2 Tage	890,00 €																									
<b>Projektmanagement</b>																											
IT-Projektmanagement	5 Tage	1.990,00 €																									
Systemisches Projektmanagement	4 Tage	1.850,00 €																									
Krisenmanagement in Projekten	2 Tage	1.100,00 €																									
IT-Projektcontrolling	3 Tage	1.290,00 €																									
<b>IT-Management</b>																											
IT-Architekturen	3 Tage	1.650,00 €																									
E-Business	2 Tage	1.100,00 €																									
Systemische Strategieentwicklung	3 Tage	1.650,00 €																									
IT-Management	5 Tage	1.990,00 €																									
IT-Risikomanagement	3 Tage	1.650,00 €																									
IT-Prozessmanagement	3 Tage	1.650,00 €																									
Konfliktmanagement	2 Tage	1.100,00 €																									

■ Seminar in Wiesbaden

■ Seminar in Bielefeld (IPS)

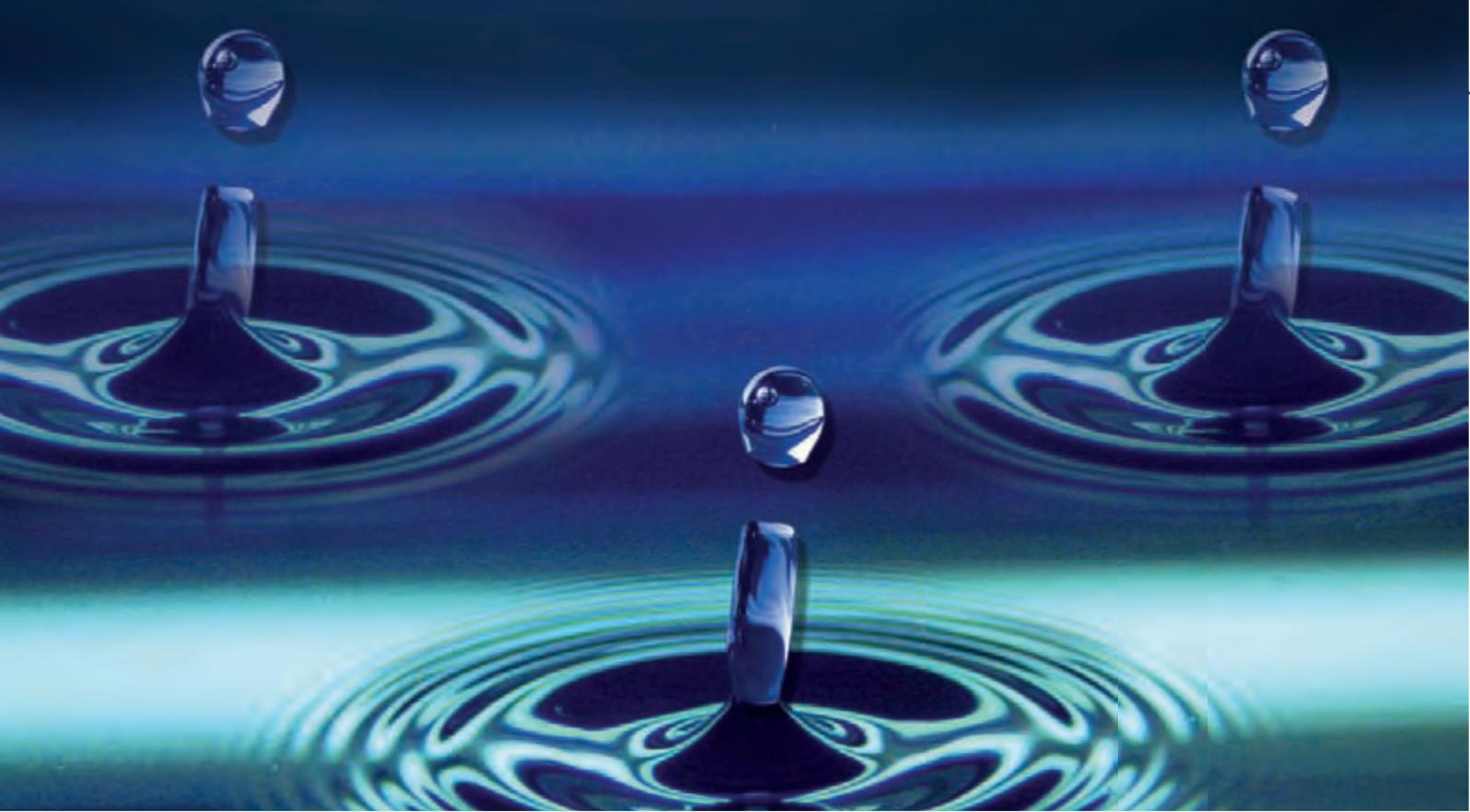
**Informationen und Anmeldung**

Für Informationen und Fragen zu individuell zugeschnittenen Seminaren, Ausbildungsreihen oder Inhouse-Schulungen stehen wir Ihnen gerne zur Verfügung. Auf Wunsch senden wir Ihnen auch unser komplettes Seminarprogramm zu.

**ORDIX AG**  
 Westermauer 12 - 16  
 33098 Paderborn  
 Tel.: 05251 1063-0

**ORDIX AG**  
 Kreuzberger Ring 13  
 65205 Wiesbaden  
 Tel.: 0611 77840-00

Preise gelten pro Seminar pro Teilnehmer in Euro zzgl. ges. MwSt. Inhousepreise auf Anfrage.



Oracle 11g Security (Teil III)

# Oracle 11gR2: Das „i-Tüpfelchen“ in puncto Sicherheit

Dieser Artikel richtet sich an Datenbankadministratoren, Sicherheitsbeauftragte und Datenbankentwickler mit Verantwortung im Security Bereich.

Im letzten Teil unserer Artikelreihe „Oracle 11g Security“ beschäftigen wir uns mit den Neuerungen, die Oracle im zweiten Release seiner aktuellen Datenbankversion in puncto Sicherheit eingeführt hat. Dieser Artikel zeigt, ob das Oracle 11g Release 2 hier etwas wesentlich Neues bringt und ob sich der Umstieg auf 11gR2 aufgrund interessanter und nützlicher Security-Funktionen wirklich lohnt.

## Audit-Daten verwalten: Package DBMS\_AUDIT\_MGMT

Was ursprünglich für Oracle Audit Vault – dem zentralen „Data Warehouse-System für Audit-Daten“ – entwickelt und eingeführt (AV 10.2.3) wurde, hat nun auch in Form des neuen Systempakets DBMS\_AUDIT\_MGMT im Standard-Lieferumfang der Datenbank Einzug gehalten.

Das Package DBMS\_AUDIT\_MGMT hält Funktionen zur Verwaltung von Inhalten der verschiedenen Audit-Pfade auf den Oracle Datenbanken bereit.

Folgende Audit-Pfade, gesteuert über den Parameter `AUDIT_TRAIL`, werden hierbei von den Prozeduren des neuen Pakets unterstützt:

- `AUDIT_TRAIL` steht auf „DB“ oder „DB,EXTENDED“  
→ Standard-Auditing wird in Tabelle `SYS.AUD$` protokolliert.  
→ Fine Grained Auditing wird in Tabelle `SYS.FGA_LOG$` protokolliert.
- `AUDIT_TRAIL` steht auf „OS“, „XML“ oder „XML,EXTENDED“  
→ Auditing erfolgt in das Betriebssystem als `*.aud` oder `*.xml` Dateien.

Mit Einführung von Oracle 11g hat sich der Default des Standard-Auditing dahingehend verändert, dass der Parameter `AUDIT_TRAIL` jetzt auf DB statt auf NONE steht und zudem viele Aktionen zwangsweise per Standard-Auditing in die `AUD$` protokolliert werden. Gerade vor diesem Hintergrund sind die Funktionen des neuen Paketes eine willkommene

Hilfe, von denen die Wesentlichen nachfolgend vorgestellt werden.

### Package DBMS\_AUDIT\_MGMT – Verschieben der Audit-Tabellen

Wegen möglicher Speicherengpässe auf dem SYSTEM Tablespace und Audit-Daten, die nicht zum Funktionieren der Datenbank nötig sind, gehören die Audit-Tabellen streng genommen nicht in diesen Tablespace. Mit dem Package DBMS\_AUDIT\_MGMT bietet Oracle – nun endlich – einen offiziell unterstützten Weg zum Verschieben der Audit-Tabelle(n) aus dem SYSTEM Tablespace heraus an (siehe Abbildung 1).

Mit dem Aufruf der Prozedur `SET_AUDIT_TRAIL_LOCATION` wird sowohl die Tabelle `AUD$` als auch die für das „Fine Grained Auditing“ verwendete Tabelle `FGA_LOG$` in den Tablespace `AUDIT_AUX` verschoben. Über die Verwendung alternativer Konstanten (`AUDIT_TRAIL_AUD_STD`, `AUDIT_TRAIL_FGA_STD`) ist es auch möglich, jeweils nur eine der genannten Audit-Tabellen zu verschieben.

### Package DBMS\_AUDIT\_MGMT – Automatisches Löschen der Audit-Daten

Wurden bisher eigenentwickelte Skripte zur manuellen Bereinigung von Datensätzen aus den Audit-Tabellen (`AUD$`, `FGA_LOG$`) und von Audit-Dateien aus dem Filesystem verwendet, bietet das `DBMS_AUDIT_MGMT`-Paket nun komfortable Prozeduren an, mit denen Audit-Informationen

- manuell  
(Prozedur `CLEAN_AUDIT_TRAIL`)

oder

- automatisiert im Rahmen eines Batch-Jobs

aus den einzelnen Audit-Pfaden gelöscht werden können.

Der Scheduler Batch-Job (siehe `DBA_SCHEDULER_JOBS`) wird hierbei über den Aufruf der Prozedur `CREATE_PURGE_JOB` und der Angabe eines gewünschten Zeitintervalls eingestellt (siehe Abbildung 2).

Das Löschen der Inhalte des jeweiligen Audit-Pfades selbst erfolgt entweder komplett

```
BEGIN
  DBMS_AUDIT_MGMT.SET_AUDIT_TRAIL_LOCATION(
    audit_trail_type => DBMS_AUDIT_MGMT.AUDIT_TRAIL_DB_STD,
    audit_trail_location_value => 'AUDIT_AUX');
END;
/
```

Abb. 1: Aufruf zum Verschieben der Audit-Tabellen.

```
BEGIN
  DBMS_AUDIT_MGMT.CREATE_PURGE_JOB(
    audit_trail_type          => DBMS_AUDIT_MGMT.AUDIT_TRAIL_
ALL,
    audit_trail_purge_interval => 24 /* Stunden */,
    audit_trail_purge_name     => 'PURGE_ALL_AUDIT_TRAILS',
    use_last_arch_timestamp    => TRUE);
END;
/
```

Abb. 2: Einrichten eines Scheduler Jobs zum täglichen Löschen aller Audit-Pfade.

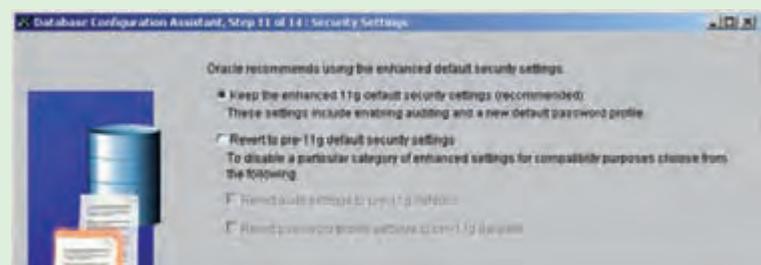


Abb. 3: Ehemalige Abfrage im DBCA 11g R1, ob die Enhanced default security settings eingeschaltet werden sollen.

oder in Abhängigkeit eines im Vorfeld gesetzten Archivierungsdatums („Last Archive Timestamp“).

Es wurden außerdem eine Reihe neuer `DBMS_AUDIT_MGMT_*` Views eingeführt, um die Einstellungen und Aktionen auf den Audit Trails überwachen zu können.

### Package DBMS\_AUDIT\_MGMT – Rotation für Audit-Dateien

Mit Hilfe der Prozedur `SET_AUDIT_TRAIL_PROPERTY` kann eine Rotation der Audit-Dateien (`*.aud`, `*.xml`) im OS nach maximaler Dateigröße (KB) und/oder -alter (Tage) aufgesetzt werden. Beim Erreichen einer der genannten Grenzwerte wird automatisch eine neue Datei für Audit-Daten auf dem OS eröff-

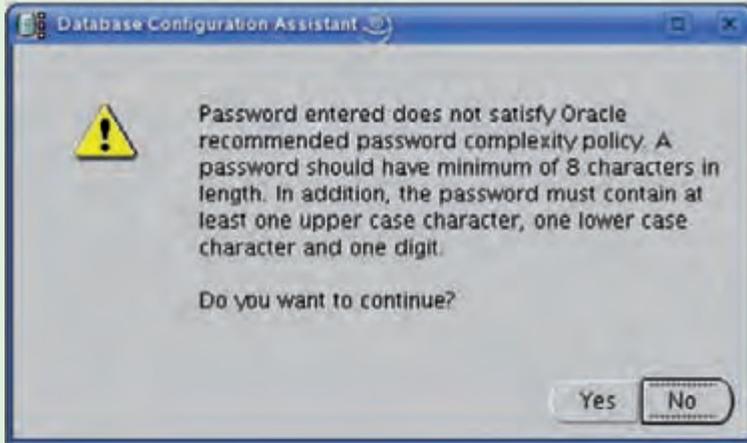


Abb. 4: Password-Check im DBCA 11gR2.

```
$ cat get_uptime.sh
#!/bin/bash
/usr/bin/uptime

SQL> CREATE TABLE uptime (data varchar2(255))
  ORGANIZATION EXTERNAL
  (
    TYPE oracle_loader
    DEFAULT DIRECTORY data_dir
    ACCESS PARAMETERS
    (
      RECORDS DELIMITED BY NEWLINE
      PREPROCESSOR exec_dir:'get_uptime.sh'
      FIELDS TERMINATED BY ',' LDRTRIM
    )
    location ('get_uptime.sh')
  );

SQL> SELECT * FROM uptime;

DATA
-----
2:30am up 11 days 7:49, 6 users, load average: 0.01, 0.06, 0.03
```

Abb. 5: Anlegen einer externen Tabelle mit PREPROCESSOR-Klausel.

net. Die Größe der OS und XML Audit-Dateien kann man somit leicht unter Kontrolle halten.

Es sei noch darauf hingewiesen, dass das Package DBMS\_AUDIT\_MGMT bereits als Patch für 10.2.0.3, 10.2.0.4 verfügbar und im Patchset 11.1.0.7 enthalten ist.

Und noch ein Hinweis in Sachen eigener Sicherheit: Achten Sie darauf, wem Sie das EXECUTE-Recht auf dieses Paket und somit das Recht, aus der Datenbank heraus Audit-Inhalte löschen zu dürfen, vergeben!

## Security Neuheiten im DBCA 11gR2

Noch in 11g Release 1 fragte der DBCA sowohl beim Anlegen als auch beim nachträglichen Ändern der Datenbank, ob die **Enhanced Default Security Settings** eingeschaltet werden sollen (Abbildung 3).

Wie schon im ersten Teil unserer Artikelreihe [1] geschildert, haben diese Sicherheitseinstellungen – sofern sie denn im DBCA ausgewählt wurden – Anpassungen des Standard-Benutzerprofils sowie ein standardmäßiges Auditing bestimmter Operationen zur Folge.

Oracle bietet mit 11g Release 2 diese Wahl überhaupt nicht mehr an; die neuen Sicherheitseinstellungen sind hier somit gesetzt und können im aktuellen DBCA auch nicht mehr abgewählt werden. Ein Verhalten, das grundsätzlich auch bei per Skript erstellten Datenbanken gilt.

Der Weg an den neuen Sicherheitseinstellungen vorbei, führt hier nur noch über die nachträgliche Ausführung der beiden Skripte `undopwd.sql` und `undoaud.sql`, die sich im Verzeichnis `$ORACLE_HOME/rdbms/admin` befinden.

Eine weitere Änderung im DBCA 11gR2 betrifft die Kennworteingabe der Systembenutzer SYS, SYSTEM, SYSMAN und DBSNMP. Die gewählten Passwörter werden nun auf ausreichende Komplexität überprüft (siehe Abbildung 4).

Ist diese nicht gegeben erscheint eine Warnung, die allerdings mit der Option „YES“ ignoriert werden kann.

Auch wenn man beide Neuerungen wie eben aufgezeigt übergehen kann, sind sie aus Security-Sicht der richtige Ansatz und sollten genutzt werden!

## EXECUTE- Privileg auf Verzeichnisobjekte

Seit 11g Release 2 besteht die Möglichkeit, ASCII-Dateien für die Nutzung von External Tables vorab einem „Preprocessing“ zu unterziehen.

Über den neuen Parameter `PREPROCESSOR` des Zugriffstreibers `ORACLE_LOADER` kann ein (Shell-)Programm in die externe Tabellendefinition eingebunden werden, um somit Daten, die nicht im entsprechenden Ladeformat

vorliegen, aufzubereiten oder diese direkt als Eingabe zu verwenden.

Folgende Einsatzmöglichkeiten sind hierbei denkbar:

- Entpacken der Datei (unzip)
- Aufbereitung des Dateiinhaltes mittels eigener Skripte (Perl, awk, etc.)
- Weitergabe der Skriptausgabe an eine externe Tabelle (z. B. Auflistung Verzeichnisinhalt, etc.); siehe hierzu Beispiel in Abbildung 5.
- Übertragen der Datei (scp, ftp)
- ...

Voraussetzung ist, dass sich das auszuführende Preprozessor-Programm in einem Verzeichnis befindet, welches der Datenbank über einen logischen Platzhalter (= Directory) bekannt gemacht wurde:

```
CREATE DIRECTORY exec_dir AS '/.../
scripts';
```

Der Benutzer, der die in diesem Verzeichnis enthaltenen Programme ausführen soll, benötigt zudem das Privileg **EXECUTE**:

```
GRANT EXECUTE ON DIRECTORY exec_dir TO
<user>;
```

### Sicherheitsaspekte bei Verwendung des EXECUTE-Privilegs

Für den Fall, dass unterschiedliche Benutzer auf unterschiedliche Programme zum Preprocessing zugreifen, sollte aus Gründen der Sicherheit mit separaten Verzeichnissen gearbeitet werden.

Die Verwendung separater Verzeichnisse ist aus Security-Sicht grundsätzlich auch für Programme, Benutzerdaten und Daten, die der Zugriffstreiber erzeugt, zu empfehlen.

Das Privileg, Dateien zu schreiben (**WRITE**) und Dateien auszuführen (**EXECUTE**) sollte nicht gleichzeitig auf einem Verzeichnis existieren.

Neben der Erstellung und Änderung von Verzeichnisobjekten, sollten nun auch Zugriffe auf die Programme zum Preprocessing mit ins Auditing aufgenommen werden:

```
AUDIT EXECUTE ON DIRECTORY exec_dir;
```

### ASM-Administration nur noch mit SYSASM-Privileg!

Bereits mit Einführung des SYSASM-Privilegs im ersten Release von Oracle 11g erfolgte eine logische Trennung zwischen Datenbank- und Storage-/System-Administration; eine durchaus sinnvolle Trennung, da beide Administrationsbereiche in der Praxis oft an unterschiedliche Personengruppen vergeben sind.

Jede Aktion, die unter der „deprecated“ Anmeldung mit SYSDBA an der Instanz 11g R1 ASM durchgeführt wird, führt zu einer Warnung im alert.log. Obwohl SYSASM und SYSDBA in 11g R1 noch gleichwertig zur Verfügung stehen, sollte die Ablösung durch SYSASM schon hier praktiziert werden.

Release 2 setzt diese Trennung konsequent fort, indem die Anmeldung als SYSDBA an einer ASM-Instanz zwar noch möglich ist, bis dato bekannte Kommandos zur ASM-Administration allerdings zum Fehler „ORA-01031: insufficient privileges“ führen.

Somit erfordert die Administration der ASM-Instanz ab 11g Release 2 die Anmeldung unter Verwendung des SYSASM-Privilegs!

Auch in 11g Release 2 ist es noch optional, mit einer eigenen OS-Gruppe (OSASM) bei der Authentifizierung über das Betriebssystem zu arbeiten. Hier gewährleistet dann nur noch die Mitgliedschaft in der OSASM (**asmadmin**) Gruppe die lokale Anmeldung mit **sqlplus / as sysasm** an der ASM-Instanz.

### TDE zur Datenverschlüsselung: Umgang mit Master Key und Wallet

Mit der unter Oracle 10.2 eingeführten und unter Oracle 11g stark erweiterten transparenten Datenverschlüsselung (TDE) geht es darum, wie sensible Daten auf physikalischer Ebene – also Daten in Datenbankdateien als auch Daten im RMAN Backup und im Data Pump Export – in verschlüsselter Form abgelegt werden können.

Bei diesem Verfahren spielt der sogenannte „Master Key“ eine zentrale Rolle. Dieser befindet sich außerhalb der Datenbank in einer kleinen Datei (**ewallet.p12**). Diese Datei wird als „Wallet“ (Brieftasche) bezeichnet und ist mit einem Passwort geschützt.

Um auf die verschlüsselten Daten zugreifen zu können, muss dieses Wallet nach jedem Start der Datenbank unter Angabe des Pass-

**Glossar**

<b>ASM</b>	Automatic Storage Management. ASM ist ein im Oracle Kernel integrierter Volume Manager. Er ist ab Oracle 10g verfügbar und arbeitet nach dem SAME-Prinzip (stripe and mirror everything). ASM kann als zentraler Shared Storage für RAC-Systeme verwendet werden.
<b>DBCA</b>	Database Configuration Assistant. Grafischer Assistent zur Erstellung und Konfiguration von Oracle Datenbanken und anderen Oracle Komponenten.
<b>Externe Tabellen</b>	Tabellen, die in der Datenbank definiert, deren Daten aber außerhalb gespeichert sind.
<b>Fine Grained Auditing</b>	Fein granularere Audit Option, die eine Überwachung des Datenzugriffs basierend auf dem Inhalt erlaubt.
<b>Hardware Security Modul</b>	Spezialisierte Hardware im Netzwerk, die sowohl zur Generierung und Ablage von Schlüsseln als auch zur Beschleunigung des Ver- und Entschlüsselns eingesetzt werden kann.
<b>PCI-DSS</b>	Payment Card Industry Data Security Standard; Regelwerk für Kreditkartendaten-verarbeitende Unternehmen.
<b>Standard-Auditing</b>	Über AUDIT-Befehle werden die drei Bereiche Objekte, Kommandos und Privilegien überwacht.
<b>TDE</b>	Transparent Data Encryption (deutsch: Transparente Datenverschlüsselung).

wortes manuell geöffnet werden. Erst dann kann sich die Datenbank den Master Key holen, um die Tabellen-/Tablespace-Schlüssel zu entschlüsseln, was wiederum die Voraussetzung für die anschließende Entschlüsselung der Daten selbst darstellt.

Möchte man den Datenbankstart hingegen vollautomatisch ablaufen lassen, kann das Wallet mit automatischer Anmeldung („Auto Open (Login) Wallet“) aufgesetzt werden. Hierdurch wird eine zweite Datei namens `cwallet.sso` angelegt, in der das Wallet-Passwort enthalten ist. Dieses wird fortan von der Datenbank genutzt, um ohne Angabe des Passworts das sonst verschlüsselte Wallet abfragen zu können.

**TDE 11gR2 - Local Auto-Login Wallet**

Konsequenz aus dem „Auto Open Wallet“ ist aber auch, dass im Falle eines Umzuges des Wallet inkl. `sso`-Datei, dieses geöffnet bleibt und so auf einem entfernten Rechner für Entschlüsselungen ge- bzw. missbraucht werden kann.

Um diese Sicherheitslücke zu schließen, kann seit Oracle 11gR2 mit einem lokalen Auto-Login gearbeitet werden.

Local Auto-Login Wallets besitzen nämlich die Eigenschaft, dass sie immer nur auf der Maschine, auf der sie erzeugt wurden, geöffnet bleiben. Beim Umzug bzw. Kopieren auf einen anderen Rechner, muss das Local Auto-Login Wallet wieder einmalig und lokal mit dem Wallet-Passwort geöffnet werden.

Ein Local Auto-Login Wallet wird mit dem Kommandozeilen-Tool `orapki` angelegt bzw. geändert:

```
# orapki wallet create -wallet <wallet_location>
-auto_login_local [-pwd <password>]
```

**TDE 11gR2 – Unified Master Key**

Neben der transparenten Verschlüsselung von Spalten können seit 11g Release 1 auch ganze Tablespaces verschlüsselt werden. Da die Verschlüsselung auf Tablespace-Ebene sämtliche Einschränkungen der Spaltenverschlüsselung aufhebt, ist sie dieser grundsätzlich vorzuziehen.

Auch die Verschlüsselung von Tablespaces arbeitet mit einem Master Key, der die einzelnen Tablespace Encryption Keys im Data Dictionary verschlüsselt und zuvor mit dem Befehl

```
alter system set encryption key
identified by "<password>";
```

im Wallet anzulegen ist.

Dieser Befehl legt bis 11.1.0.7 zwei unterschiedliche Schlüssel an, einen Master Key für die Spalten- und einen Master Key für die Tablespace-Verschlüsselung. Nur derjenige für die Spaltenverschlüsselung konnte neu generiert (Rekey) werden.

In Release 2 legt der gleiche Befehl nur noch einen „Unified Master Key“ im Wallet an, der sowohl für die TDE Spalten- als auch Tablespace-Verschlüsselung verwendet wird und zudem mit der angegebenen Syntax jederzeit gewechselt (Rekey) werden kann. Eine Eigenschaft, die Richtlinien wie z. B. PCI-DSS entgegenkommt, die einen periodischen Schlüsselwechsel verlangen.

**TDE 11gR2 und Hardware Security Module (HSM)**

Wem die Konfiguration mit Software Wallets noch zu unsicher ist oder die Schlüsselverwaltung in verteilten Umgebungen (z. B. mit

Data Guard und RAC) erleichtern möchte, kann Hardware Security Module (= Hardware Wallets) zur sichereren Generierung und Ablage von TDE Masterkeys benutzen.

Die HSM-basierte Verschlüsselung wird seit Oracle 11g unterstützt, unterscheidet sich hier allerdings je nach installiertem Patchlevel:

#### 11.1.0.6:

HSM-basierte Spaltenverschlüsselung wird unterstützt; Master Key für Tablespace-Verschlüsselung muss im Software Wallet abgelegt werden.

#### 11.1.0.7:

TDE unterstützt auch die HSM-basierte Tablespace-Verschlüsselung; allerdings nicht für den nachträglichen Einsatz (kein Rekey).

#### 11gR2:

Ein Unified Master Encryption Key im HSM für TDE Spalten und Tablespace-Verschlüsselung; vorhergehende Restriktionen bzgl. TDE und HSM sind aufgehoben.

### Fazit

Abgesehen von „echten“ Neuheiten wie z. B. das Paket DBMS\_AUDIT\_MGMT zur Verwaltung von Audit-Daten beschränken sich die

### Links

- ▶ [1] ORDIX News Artikel: „Secure by Default“ in Oracle 11g – Die sichere Datenbank out-of-the-box?“ [http://www.ordix.de/ORDIXNews/1\\_2010/Datenbanken/secure\\_by\\_default.html.html](http://www.ordix.de/ORDIXNews/1_2010/Datenbanken/secure_by_default.html.html)
- ▶ [2] Buch „HOWTO Secure and Audit Oracle 10g and 11g“, Autor: Ron Ben Natan, ISBN-10: 1420084127, ISBN-13: 978-1420084122
- ▶ [3] Oracle Database Advanced Security Administrator's Guide 11g Release 2
- ▶ [4] Oracle Database PL/SQL Packages and Types Reference 11g Release 2
- ▶ [5] Oracle DBA Community: <http://www.oracle.com/global/de/community/dbadmin/index.html>

Neuerungen von 11gR2 in puncto Sicherheit größtenteils auf Vervollständigungen und Ergänzungen; eben das „i-Tüpfelchen“ für die in 11g Release 1 eingeführten Security-Funktionalitäten.

Rein aus Sicht der Security ist ein Umstieg auf das zweite Release also nicht unbedingt notwendig; auch wenn dessen Ergänzungen durchweg sinnvoll und nützlich sind.



Kathleen Hock  
([info@ordix.de](mailto:info@ordix.de)).

## Seminarempfehlung: Oracle Security

▶ **Informationen/Online-Anmeldung:** <http://training.ordix.de/siteengine/action/load/kategorie/Datenbanken/nr/837/index.html>

In diesem Seminar werden potenzielle Sicherheitsschwachstellen aufgezeigt und Lösungsansätze erarbeitet. Somit können Sie anschließend Daten vor unberechtigtem Zugriff schützen und in der Datenbank sowie im Netz verschlüsseln. Außerdem können Sie dafür sorgen, dass Applikationen je nach Berechtigung Daten individuell präsentieren.

### Seminarinhalte

- Auditing
- Datenzugriffskontrolle
- Identifizierung und Authentifizierung
- Verschlüsselung
- Absicherung der Datenübertragung
- Überprüfen von Sicherheitslücken
- Database Security in Oracle 11g
- Vertiefung der Theorie durch praktische Übungen und Beispiele

### Termine

18.10. - 21.10.2010 in Wiesbaden  
06.12. - 09.12.2010 in Wiesbaden

**Seminar-ID:** DB-ORA-33A

**Dauer:** 4 Tage

**Preis pro Teilnehmer:** 1.590,00 € (zzgl. MwSt.)

**Frühbucherpreis:** 1.431,00 € (zzgl. MwSt.)

Wir führen unsere Seminare auch jederzeit an einem geeigneten Ort Ihrer Wahl durch und bringen, wenn nötig, auch das entsprechende Equipment mit. Informieren Sie sich am besten im Internet über unsere Kundenseminare und die mobilen Schulungen: <http://training.ordix.de>.

$$D^2 = \frac{1}{P^2} \frac{P_0 - P}{P} \sim \frac{1}{P^2} \quad (1a)$$

$$D^2 = \frac{K_0}{3} \frac{P_0 - P}{P} \sim K_0 \quad (2a)$$

$$\text{IBM} = \text{Power}^7 + \text{AIX}^7$$

$$D^2 \sim 10^{-53}$$

$$D^2 \sim 10^{-53}$$

$$e \sim 10^{-26}$$

$$e \sim 10^{-26}$$

$$P \sim 10^8 \text{ G.J.}$$

$$P \sim 10^8 \text{ G.J.}$$

IBM POWER 7 und AIX 7 - Neuheiten

## Leistung zur siebten Potenz: POWER 7 und AIX 7

Der Artikel richtet sich an Berater, Systemadministratoren und Entscheider, die sich für technologische Neuerungen im Umfeld von IBM und Unix interessieren.

IBM hat vor einigen Wochen die neue POWER 7 Hardware auf den Markt gebracht. Das Betriebssystem AIX 7 folgt in den nächsten Wochen. Wie schon AIX 6 wird die neue Version über ein open beta-Programm den Marktteilnehmern vorgestellt und ist für den Herbst angekündigt. Dieser Artikel gibt einen Überblick über die zu erwartenden Neuerungen.

### AIX Roadmap

Die IBM Version AIX 5.3 wird ab Anfang 2011 nicht mehr vertrieben und ein Jahr später nur noch gegen Aufpreis gewartet. IBM setzt hiermit die Kunden unter Druck, auf AIX 6.1 zu migrieren, welches erst ab Anfang 2013 innerhalb eines Jahres aus dem Markt geholt wird. Momentan ist das Betriebssystem in drei Editionen verfügbar: Standard, Enterprise und Express Edition.

Die unterschiedlichen Editionen haben verschiedene Leistungsumfänge, besonders hinsichtlich ihrer Virtualisierungsmöglichkeiten. Sie werden preislich abgestuft nach Server-Größe angeboten. Eine AIX-Version der Enterprise Edition beinhaltet dann beispielsweise neben AIX 6.1 auch den WPAR Manager V2.1 und die Systems Director Enterprise Edition. Ein AIX der Express Edition hingegen ist be-

schränkt auf maximal 4 Cores und 8 GB Hauptspeicher pro Partition.

### Automatisierte Systemkonfiguration

Eine neue Eigenschaft von AIX 6.1 TL04 ist der AIX Runtime Expert (ARTEX). ARTEX liefert eine Methode, die es ermöglicht von einem zentralen Punkt Konfigurationsinformationen (Profil) eines AIX-Systems (Parameter zu NFS, RAS, Sicherheit, Tuning, ...) in Form von xml-Dateien auf ein anderes System zu verteilen und dort wirksam werden zu lassen.

Die Profile können manuell erstellt, oder von einem existierenden, konfigurierten System abgezogen und verteilt werden. So unterstützen sie eine weitere Automatisierung der Systembereitstellung. Vorgefertigte Profile können dann eingesetzt werden, um das System nach Vorgaben des Workload

(Datenbank, Webserver, Applicationserver, SAP) automatisch zu tunen.

### Erweiterte Unterstützung von Workload Partitions

Das AIX-Pendant zu den Solaris Zonen "Workload Partitions" ist um einige Funktionen erweitert worden. Es ist nun möglich, Teile einer rootvg (die Dateisysteme /, /var, /tmp, /admin, /home) auf einer dedizierten Platte auszulagern (WPAR rootvg). Die WPAR unabhängigen Dateisysteme /usr, /opt und /proc werden weiterhin von der globalen WPAR in die WPAR-rootvg per NFS eingehängt.

Weiterhin wird die Lauffähigkeit des AIX 5.2 unter AIX 7 WPARS auf POWER 7 unterstützt. Dabei werden Umgebungen konsolidiert, in denen das alte Betriebssystem nicht migriert werden kann. AIX 7 ist binärkompatibel zu AIX 6, AIX 5-Applikationen, auch unter 32-bit, müssen nicht rekompiliert werden. Des Weiteren unterstützt AIX 7 Plattformen bis zurück zu POWER 4 und ist direkt von AIX 5 aus migrierbar.

### Die neue Hardware

IBM hat sich bei der Entwicklung der neuen Prozessoren von der Jagd nach immer höheren Taktfrequenzen verabschiedet und setzt auf mehr Prozessorkerne und Threads pro CPU und schlägt damit die Richtung anderer Hersteller ein. Zur Zeit stehen die Modelle IBM-POWER 750, 755, 770 und 780 zur Verfügung. Sie bieten im Vergleich zu den jeweiligen Vorgängermodellen eine um den Faktor 4 erhöhte Anzahl an Cores und die Möglichkeit per SMT4 vier Threads pro Core laufen zu lassen. Die Anzahl der Threads verachtfacht sich somit. Für das zukünftige Flaggschiff des Nachfolgers der p595-Server, muss AIX imstande sein 1024 Threads zu bedienen.

Eine CPU Version mit nur 6 Cores ist auch erhältlich, sie bietet erhöhte Taktraten. Die Performance eines Cores konnte im Vergleich zur POWER 6 Plattform zwar nur um etwa 20 – 25 % gesteigert werden. Die Leistungsfähigkeit der gesamten CPU steigert sich dennoch durch die erhöhte Anzahl an Cores.

### Das Ende allen Hauptspeichermangels?

IBM bietet zwei neue Eigenschaften innerhalb der POWER-Architektur, um dem allgegenwärtigen Mangel an Hauptspeicher zu begegnen. Zum einen ist dies die Möglichkeit

durch „Active Memory Sharing“ (AMS) einem Speicherpool mehreren Partitionen zur Verfügung zu stellen, ähnlich der Bereitstellung eines Prozessor-Pools für viele Partitionen. Dies kündigte IBM zwar schon 2006 an, erhältlich ist diese Eigenschaft erst seit April 2009 für POWER 6. Die zweite Eigenschaft ist „ActiveMemoryExpansion“ (AME), hierbei wird der Hauptspeicher wie bei einem „Zip“-Archiv gepackt. Beide Eigenschaften sind kombinierbar.

### Speicher für alle – und zwar geteilt

Die Idee zu AMS ergibt sich als logische Folge des Konzepts der geteilten Prozessor-Ressourcen. Nicht eine Partition erhält eine fest zugewiesene Menge an Hauptspeicher, sondern mehrere Partitionen bedienen sich je nach Anforderung aus einem, durch den Hypervisor, kontrollierten Pool. Dabei benötigt ein gering ausgelasteter Server wenig Speicher und gibt diesen dem Pool zurück, während ein zweiter Server sich die Ressourcen aus dem Pool in Hochlastphasen herausziehen kann. Partitionen mit fest zugewiesenem Speicher und solche mit Benutzung eines Speicherpools können auf einem System dabei gleichzeitig existieren.

Die Zuweisung des Speichers funktioniert vollautomatisch und dynamisch. Diese Zuweisung muss nicht durch DLPAR-Operationen per Hand eingeleitet werden. Die Verteilung des Hauptspeichers auf die Partitionen lässt sich somit flexibler gestalten, wobei sich nicht jeder Workload auf einem System für den Einsatz von AMS eignet.

Workloads, die von AMS profitieren sind solche, mit wechselnden und „sanften“ Schwankungen in der Speicherauslastung, bzw. Applikationen, die mit Latenzzeiten durch auftretendes Paging zurecht kommen oder häufig Filecaching nutzen.

Schlechte Kandidaten für AMS sind Systeme, die einen Hochleistungszugriff auf den Speicher benötigen, plötzliche und starke Schwankungen in der Speicherausnutzung aufweisen oder Speicher nutzen, der nicht gepaged werden darf (Pinned Memory). Vor der Nutzung ist die Workload zu analysieren und über Benchmarking vor und nach der Einführung von AMS auf Leistungsverluste zu testen.

### AME als Speicherpacker

Eine weitere Eigenschaft ist die Active Memory Expansion der POWER 7 Architektur, welche mit AIX 6.1 nutzbar ist.

**Glossar**

<b>DLPAR</b>	Dynamic Logical Partitioning. Möglichkeit seit POWER 5, Hardware-Ressourcen einer Partition online zuweisen zu können.
<b>Hypervisor</b>	Hardware-Baustein, zur Steuerung der POWER-Rechensysteme.
<b>rootvg</b>	Die Systemfestplatten von AIX.
<b>Thread</b>	Anzahl der Programmabläufe, die gleichzeitig durch einen Prozessor bearbeitet werden.
<b>WPAR</b>	Workload Partition, eine von der „Global Partiton“ betriebene Software-Partition von AIX.

**Links**

- ▶ [1] ORDIX News Artikel „IBM POWER5: Logische Partitionierung & Virtualisierung“:  
[http://www.ordix.de/ORDIXNews/3\\_2007/Betriebssysteme/ibm\\_power5.html](http://www.ordix.de/ORDIXNews/3_2007/Betriebssysteme/ibm_power5.html)
- ▶ [2] ORDIX News Artikel „POWER6 und AIX 6.1 - IBMs neues Traumpaar?“:  
[http://www.ordix.de/ORDIXNews/4\\_2007/Betriebssysteme/POWER6\\_AIX6.html](http://www.ordix.de/ORDIXNews/4_2007/Betriebssysteme/POWER6_AIX6.html)
- ▶ [4] Seminarempfehlungen:  
IBM AIX Systemadministration Grundlagen  
<http://training.ordix.de/siteengine/action/load/kategorie/Betriebssysteme/nr/554/index.html>  
Analyse von komplexen AIX Performance Problemen und Schritte zur Optimierung:  
<http://training.ordix.de/siteengine/action/load/kategorie/Betriebssysteme/nr/997/index.html>  
Analyse von komplexen AIX Installation, Backup und Recovery:  
<http://training.ordix.de/siteengine/action/load/kategorie/Betriebssysteme/nr/853/index.html>
- ▶ [5] IBM POWER und Storage Symposium, Wiesbaden, 10. – 12.05.2010  
<http://www.ibm.com>

der dann tatsächlich für die Kompressionsrate eingestellt wird, wird nach Maßgabe der freien Prozessor-Ressourcen gesetzt. Natürlich gibt es auch hier Einschränkungen: Die Daten im Hauptspeicher müssen komprimierbar und die Partition darf nicht am Limit der Prozessorlast sein. AME erscheint IBM als so nützlich, dass es zusätzlich lizenziert werden muss.

**Fibre Channel Adapter – neu virtualisiert**

Eine sehr nützliche Neuerung, die allerdings kein AIX 7/POWER 7 voraussetzt, verbirgt sich hinter dem sperrigen Namen N-Port ID Virtualization (NPIV). Beim bisherigen Weg der Virtualisierung von Platten im SAN werden die Geräte als „hdisks“ auf dem VIOS bereitgestellt und von dort aus einzeln auf die Virtual Server SCSI Adapter verteilt. Dies hat zur Folge, das auf einem VIOS hunderte von Platten hängen können, was deren Administration erheblich kompliziert.

Mit NPIV kann eine FC-Adapterkarte 32 virtualisierte WWN-Adressen erhalten, d. h. bis zu 32 Partitionen bekommen einen virtualisierten FC-Adapter mit eigener WWN. Das Zoning geschieht dann direkt mit den virtuellen WWNs, so dass die Verwaltung der Platten vom VIOS auf die einzelnen Partitionen verlagert wird.

**POWER wird grün**

Die neue Prozessor-Generation bietet einige Energiespareigenschaften. So kann, wenn es die Umgebungsbedingungen zulassen, der Prozessortakt kurzfristig erhöht werden, damit mehr Rechenleistung bereitsteht.

Umgekehrt kann bei Minderauslastung die Prozessortaktung herabgesetzt werden bis zum Punkt, wo einzelne Cores abgeschaltet werden. IBM verspricht hier eine Vervierfachung der Rechenleistung bei gleichzeitiger Erhöhung der Energieaufnahme um 14%.

**Fazit**

Als weitreichendste Neuerungen können sicherlich das verbesserte Speichermanagement mittels AME und AMS, die Virtualisierung von FC-Adapttern und die erhöhte Energieeffizienz der Plattform genannt werden. Sie flexibilisieren den Einsatz von Hardware-Ressourcen und erhöhen deren Effizienz.

Nach dem Quantensprung der Einführung von POWER 5-Systemen baut IBM auch in Zukunft auf die durchgehende Virtualisierung von IT-Landschaften.

Sie nutzt „überschüssige“ Prozessorkapazität um Seiten aus dem Hauptspeicher zu lesen und komprimiert wieder zurückzuschreiben. Mit AME lässt sich unter sehr günstigen Umständen ein Kompressionsfaktor von 10 erreichen, allerdings ist dies durch die dann exzessive Prozessorauslastung nicht praktikabel. Mit Kompressionsfaktoren um die 1,5 z. B. erzeugt man in vielen Einsatzbereichen so gut wie keine zusätzliche Prozessorlast durch die Kompression.

Dazwischen gibt es einen Bereich, der durch den Administrator ausgelotet werden muss. Eine Partition, der man beispielsweise 12 GB zuweisen möchte, wird so eingerichtet, dass sie 8 GB physikalischen Speicher bekommt und ein Kompressionsfaktor von 1,5 besitzt. Sie besitzt dann 12 GB „komprimierten“ Speicher. Diese Eigenschaft ist partitionsweise für das AIX-Betriebssystem einsetzbar.

Interessant ist, das durch ein bei AIX 6.1 TL04 SP02 mitgeliefertes Planungs-Tool (amepat) abgeschätzt werden kann, wieviel zusätzliche Prozessorlast auf einer gegebenen Partition durch verschiedene Werte für die Kompression hervorgerufen werden würden. Der Wert,



Dr. Uwe Bechthold  
([info@ordix.de](mailto:info@ordix.de)).

## Nützliche Neuheit in der JPA 2.0

# Endlich dabei – die Criteria Query API in der JPA 2.0

Viele Entwickler haben bisher die Criteria Query API – wie sie in Hibernate zur Verfügung steht – in der Java Persistence API (JPA) vermisst. Nun ist diese API in der JPA 2.0 integriert. Sie bietet ein objektorientiertes Verfahren zum dynamischen Aufbau von Abfragen. Damit soll sie nicht die mächtige JPQL (Java Persistence Query Language) ersetzen, sondern eine sinnvolle Ergänzung sein, wenn es um die dynamische Zusammenstellung von Abfragen in einer Anwendung geht.

Dieser Artikel richtet sich an erfahrene Java-Entwickler, die bereits mit JPA gearbeitet haben.

## Einführung

Viele Entwickler kennen die Criteria Query API von Hibernate und haben sich gefragt, wann diese endlich in den Standard der JPA übernommen wird. Mit der Version JPA 2.0 ist dies nun geschehen. Die Criteria Query API ist aber nicht als Ersatz für die Java Persistence Query Language (JPQL), sondern als Ergänzung für bestimmte Anwendungsfälle gedacht. Man kann mit ihr sehr gut Queries in objektorientierter Weise zur Laufzeit zusammensetzen und dabei die Vollständigkeit und die Typkorrektheit weitestgehend sicherstellen. Geeignet ist diese API z. B. für die Auswahl von verschiedenen Filtern, für die Einschränkung der Ergebnismenge von Projekten, Mitarbeitern, Aufträgen etc.

Fangen wir mit einem einfachen Beispiel an. Eine Query mit JPQL sieht wie folgt aus:

```
SELECT w FROM worker w
WHERE w.name = 'Oliver'
```

Wie diese Abfrage mit der Criteria Query API dann aussieht, zeigt die Abbildung 1. Wie man sofort feststellt, ist das Schreiben von Queries mit der Criteria Query API nicht unbedingt kürzer, aber dies ist auch nicht die Absicht. Mit dem Befehl `cb.createQuery(...)` erzeugen wir eine Criteria Query für den Ergebnistyp `Worker`. Mit `cq.from(...)` erzeugen wir eine Projektionsvariable (Root) und mit `cq.select(...)` legen wir die Selektion fest. Mit `cq.where(...)` wird dann die Bedingung beschrieben. Es tauchen also mit `from`, `select` und `where` alle Elemente auf, wie man sie von einer normalen JPQL-Query kennt. Der Aufbau ist also der gleiche, man benutzt nur eine andere Schreibweise.

## Metamodell-Klassen

Diese Schreibweise stellt leider immer noch nicht die vollständige Typsicherheit und die Tatsache sicher, ob das verwendete Attribut der Entity existiert. Wir arbeiten hier immer noch mit Strings, bei denen es zu Schreibfehlern kommen kann. So kommt es beispielsweise erst zu einem Laufzeitfehler, wenn es kein Attribut `name` in der Entity `Worker` gibt, oder das Attribut vom falschen Typ ist. An dieser Stelle kommen die Metamodell-Klassen ins Spiel. Möchte man diese einsetzen, sollte zu jeder Entity-Klasse eine Metamodell-Klasse existieren. Diese kann man theoretisch manuell erstellen. Viel sinnvoller ist aber die automatische Generierung mit Hilfe von Annotations-Prozessoren, wie z. B. dem `HibernateMetamodel` Generator.

In unserem Beispiel (siehe Abbildung 2) haben wir eine Entity-Klasse `Worker`. Die dazugehörige Metamodell-Klasse sehen Sie ebenfalls in Abbildung 2.

Diese Klasse können wir jetzt einfach in unserer Criteria Query benutzen und sind so vollkommen typsicher. Fast alle Fehler können so schon zur Kompilierungszeit festgestellt werden:

```
cq.where(cb.equal(w.get
(Worker_.name), "Oliver"));
```

## Pfad-Ausdrücke

Aus der JPQL kennen wir bereits folgende Pfad-Ausdrücke:

```
CriteriaBuilder cb = entityManager.getCriteriaBuilder();
CriteriaQuery<Worker> cq = cb.createQuery(Worker.class);
Root<Worker> w = cq.from(Worker.class);
cq.select(w);
cq.where(cb.equal(w.get("name"), "Oliver"));
TypedQuery<Worker> q = entityManager.createQuery(cq);
List<Worker> result = q.getResultList();
```

Abb. 1: Eine Query-Abfrage mit Criteria Query API.

```
@Entity
public class Worker {
    @ID
    private Integer id;
    private String name;
    private Integer age;
    @ManyToOne
    private Department department;
    @OneToMany
    List<Project> myProjects;
    ...
}
@StaticMetamodel(Worker.class)
public class Worker_ {
    public static volatile SingularAttribute< Worker, Integer> id;
    public static volatile SingularAttribute< Worker, String> name;
    public static volatile SingularAttribute< Worker, Integer> age;
    public static volatile SingularAttribute< Worker, Department > department;
    public static volatile ListAttribute< Worker, Project> myProjects;
```

Abb. 2: Die Entity Klasse Worker mit der dazugehörigen Metamodell-Klasse.

**Direkte Benutzung:**

```
cq.where(cb.equal(w.get(Worker_.department)
    .get(Department_.location), "Paderborn"));
```

**Mit separater Definition:**

```
Root<Worker> w = cq.from(Worker.class);
Path<Department> dep = w.get(Worker_.department);
Path<String> loc = dep.get(Department_.location);
cq.select(w).where(cb.equal(loc, "Paderborn"));
TypedQuery<Worker> q = entityManager.createQuery(cq);
List<Worker> result = q.getResultList();
```

Abb. 3: Einbau des Pfades in die where-Methode und die separate Abfrage.

```
CriteriaQuery<String> cq = cb.createQuery(String.class);
Root<Worker> w = cq.from(Worker.class);
Path<String> loc = w.get(Worker_.department).get(Department_.
location);
cq.select(loc).distinct(true);
cq.where(cb.lessThan(w.get(Worker_.age), 18));
TypedQuery<String> q = entityManager.createQuery(cq);
List<String> result = q.getResultList();
```

Abb. 4: Auslieferung aller Standorte mit Altersdefinition.

```
Root<Worker> w = cq.from(Worker.class);
Join<Worker, Project> p = w.join(Worker_.myProjects, JoinType.
Inner);
cq.select(w).distinct(true);
cq.where(cb.equal(p.get(Project_.status), "delayed"));
TypedQuery<Worker> q = entityManager.createQuery(cq);
List<Worker> result = q.getResultList();
```

Abb. 5: Abfragedefinition mit Hilfe der Join-Ausdrücke.

```
SELECT w FROM worker w
WHERE w.department.location =
'Paderborn'
```

Dieser Pfad selektiert alle Arbeiter, deren Abteilung sich am Standort Paderborn befindet. Auch dies ist natürlich mit der Criteria Query API möglich. Der Pfad (`Path`) kann z. B. direkt in der `where`-Methode eingebaut oder separat definiert werden. Damit ist der `Path` an verschiedenen Stellen nutzbar wie z. B. in der `where`- oder auch der `select`-Methode, um den Rückgabewert zu definieren (siehe Abbildung 3).

Man kann dieses Beispiel auch mit nur einem Schritt, ohne den Zwischenschritt mit `dep` durchführen. Die Pfad-Angabe kann, wie bereits beschrieben, auch im `select`-Teil benutzt werden. In Abbildung 4 wird ein Beispiel aufgezeigt, indem alle Standorte geliefert werden, an denen Arbeiter unter 18 Jahren arbeiten. Nebenbei zeigt es noch die Verwendung des `distinct`-Befehl.

**Join-Ausdruck**

Wie bei SQL und JPQL gibt es natürlich auch bei der Criteria Query API Joins, bei denen ein `Join`-Typ angegeben werden kann. Gibt man keinen Typ an, so handelt es sich um einen Inner-Join. Das Beispiel in Abbildung 5 zeigt, wie man mit Hilfe der `Join`-Ausdrücke eine Abfrage definiert, die alle Arbeiter liefert, die in Projekten tätig sind, die sich verzögern. Auch bei den `Join`-Ausdrücken gibt es wie bei den `Path`-Ausdrücken die Möglichkeit, diese zu verketten.

**Parameter**

Die bisherigen Beispiele haben stets feste Werte bei den Abfragen. In der Praxis möchte man diese in der Regel parametrisieren. Dies erfolgt in der Criteria Query API mit Hilfe von Parameter-Expressions. Im folgenden Beispiel (siehe Abbildung 6) führen wir auch gleich die `Predicates` mit ein.

Dies sind separat definierte Bedingungen, die man im `where`-Befehl benutzen kann. Das Beispiel soll alle Mitarbeiter liefern, die ein bestimmtes Alter haben und an einem bestimmten Ort arbeiten.

In den beiden Variablen `location` und `age` stehen dann die eigentlichen Werte, die übergeben wurden.

## In-Befehl

Häufig möchte man ein Attribut bei mehreren Werten abfragen. Dazu dient der `in`-Befehl. Das Beispiel in Abbildung 7 ermittelt alle Arbeiter, die in Paderborn, Wiesbaden oder Köln arbeiten.

Die Schreibweise ist erst einmal etwas ungewöhnlich und bei vielen Werten auch unpraktisch. Daher gibt es eine Kurzform, die allerdings nicht so viele Optionen ermöglicht:

```
cq.where(loc.in("Paderborn",
"Wiesbaden", "Köln"));
```

## Construct-Befehl

Die Ergebniswerte aus der Query können dazu genutzt werden, um eine neue Klasse zu instanziiieren. Dies kann mit der Criteria Query API direkt in der Abfrage geschehen. Dazu benötigt man den `construct`-Befehl. Für das Beispiel in Abbildung 8 haben wir eine Klasse `WorkerDetails`, die nur Namen und Alter enthält. Ferner hat die Klasse einen Konstruktor mit diesen beiden Parametern als Eingabe.

## Fazit

Die Criteria Query API steht der JPQL in ihrem Funktionsumfang in nichts nach. Durch ihre doch recht ausschweifende Schreibweise gegenüber der JPQL, bietet sie sich nicht für jeden Zweck als Alternative an. Für das dynamische Zusammensetzen von Queries zur Laufzeit ist sie aber erste Wahl. Auch wer auf Typsicherheit steht, ist – im Zusammenhang mit den Metamodell-Klassen – mit der Criteria Query API gut bedient. In Summe kann man also von einer guten und sinnvollen Ergänzung des JPA-Standards sprechen

Übrigens würde ich mich sehr freuen, wenn Sie meinen Vortrag zu diesem Thema auf der diesjährigen DOAG Konferenz besuchen würden.



Oliver Kalutza  
([info@ordix.de](mailto:info@ordix.de)).

```
CriteriaQuery<Worker> cq = cb.createQuery(Worker.class);
Root<Worker> w = cq.from(Worker.class);
Path<String> l = w.get(Worker_.department).get(Department_.
location);

ParameterExpression<String> pLoc = cp.parameter(String.class, "loc");
Predicate criterial = cb.equal(l, pLoc);
ParameterExpression<Integer> pAge = cp.parameter(Integer.class, "age");
Predicate criteria2 = cb.equal(w.get(Worker_.age), pAge);

cq.select(Worker).distinct(true);
cq.where(cb.and(criterial, criteria2));
TypedQuery<Worker> q = entityManager.createQuery(cq);
q.setParameter("loc", location);
q.setParameter("age", age);
List<Worker> result = q.getResultList();
```

Abb. 6: Beispiel von ParameterExpressions.

```
Root<Worker> w = cq.from(Worker.class);
Path<String> loc =w.get(Worker_.department).get(Department_.
location);
cq.select(w)
cq.where(cb.in(loc).value("Paderborn").value("Wiesbaden").
value("Köln"));
```

Abb. 7: Abfrage mit dem `in`-Befehl.

```
CriteriaQuery<WorkerDetails> cq =
cb.createQuery(WorkerDetails.class);
Root<Worker> w = cq.from(Worker.class);
Path<String> name = w.get(Worker_.name);
Path<Integer> age = w.get(Worker_.age);
cq.select(cb.construct(WorkerDetails.class, name, age));
TypedQuery<WorkerDetails> q = entityManager.createQuery(cq);
List<WorkerDetails > result = q.getResultList();
```

Abb. 8: Abfrage mit dem `construct`-Befehl.

## Glossar

- API** Application Programming Interface (API) - Sammlung von Klassen, welche eine Schnittstelle zu einer Hardware oder Anwendung definiert.
- Hibernate** Open Source Persistence Framework für Java. Das Framework ermöglicht es, den Zustand eines Objekts in einer relationalen Datenbank zu speichern und aus entsprechenden Datensätzen wiederum Objekte zu erzeugen. Dies bezeichnet man auch als Object Relational Mapping (OR-Mapping oder ORM).
- JPA** Java Persistence API - ist Standard für die Persistenz-Abbildung im Java EE- und Java SE-Bereich. Ein wichtiges Merkmal der Java Persistence API ist, dass sie für beliebige POJOs verwendet werden kann und nicht auf EJBs beschränkt ist. Somit steht sie auch den Entwicklern zur Verfügung, die keine EJB-Komponenten in ihren Applikationen benötigen.
- JPQL** Die Java Persistence Query Language (JPQL) ist die Abfragesprache der Enterprise JavaBeans Komponentenarchitektur. Sie ist angelehnt an SQL, fokussiert aber Objekte. Die Java Persistence Query Language ist im Vergleich zur EJB QL erheblich erweitert und teilweise neu entwickelt worden.

## Link

- ▶ [1] Spezifikation von der JPA 2.0:  
<http://www.jpcc.org/en/jsr/detail?id=317>



Resource Management mit Linux

# Control Groups - Linux legt Prozesse in Ketten

**Dieser Artikel richtet sich an Linux-Administratoren, die in größeren Umgebungen Hardware-Ressourcen auf einem Server aufteilen wollen.**

**Jeder kennt die guten, alten Nice-Level um Prozesse unter Linux zu priorisieren, aber wie können Prozesse auf eine bestimmte CPU gepinnt werden? Wie grenzt man die Speichernutzung von Prozessgruppen ein? Welche Möglichkeit gibt es den Zugriff von Prozessen auf bestimmte Geräte zu beschränken? All diesen und noch mehr Fragen können Sie nach diesem Artikel mit nur einer Antwort begegnen: Control Groups. Um was es sich dabei genau handelt und wie Sie diese Prozesse im täglichen Betrieb nutzen können, werden wir in diesem Artikel erläutern.**

## Einführung

Der Name Control Groups oder kurz cgroups ist auf deren Aufbau zurückzuführen. Die Prozesse auf einem Server sind in hierarchisch angeordneten Gruppen sortiert. Diese Gruppen können über verschiedene Subsysteme kontrolliert werden. Die Subsysteme bieten in erster Linie die Möglichkeit den cgroups Hardware-Ressourcen zuzuweisen bzw. Beschränkungen aufzuerlegen. Dies geht über die bis jetzt bekannten Möglichkeiten der Priorisierung und Hardware-Zuweisung von bzw. zu Prozessen hinaus.

Da die cgroups seit Version 2.6.24 ein Bestandteil des Standard-Linux-Kernels sind, können sie unter den meisten aktuellen Distributionen genutzt werden.

## Konzept und Aufbau

Der Kernel stellt ein Interface, in Form von virtuellen File-Systemen, zur Verfügung mit dem die verschiedenen Control Groups in Hierarchien aufgebaut werden können. Zum Verständnis des Konzepts sind die folgenden drei Elemente wichtig:

- Eine cgroup ist eine Gruppe von Prozessen denen verschiedene Parameter für eine oder mehrere Subsysteme zugewiesen sind.
- Ein Subsystem ist ein Modul bei dem es sich normalerweise um einen Resource Controller handelt. Die unterschiedlichen Subsysteme bieten verschiedene Möglichkeiten, um Prozesse auf cgroup-Basis zu beeinflussen.

- Das Resource Management ist in eine oder mehrere Hierarchien aufgebaut. Eine Hierarchie besteht aus cgroups, die in einer Baumstruktur aufgebaut sind und einem oder mehreren Subsystemen, die der Struktur zugewiesen sind. Jede Hierarchie ist in dem System als virtuelles File-System sichtbar.

In Abbildung 1 wird der Zusammenhang von cgroups, Subsystemen und Hierarchien grafisch dargestellt.

### Ein Controller für alle Lebenslagen

Es stehen Subsysteme für die unterschiedlichen Aufgaben zur Verfügung. Sie bieten zum Beispiel die Möglichkeit eine Gruppe von Prozessen nur auf bestimmten CPU Cores zu berechtigen oder einen Core für eine Prozessgruppe zu reservieren. Die Tabelle in Abbildung 2 bietet eine Übersicht über die wichtigsten bereits implementierten Subsysteme. Welche Subsysteme auf dem eigenen Server zur Verfügung stehen, ob sie einer Hierarchie zugeordnet sind und wie viele cgroups in den Hierarchien existieren, kann mit dem Befehl `cat /proc/cgroups` abgefragt werden.

Die Konfiguration der Gruppen erfolgt über die von den Subsystemen bereitgestellten Dateien. Um beispielsweise eine Gruppe von Prozessen auf die CPU 3 zu pinnen, muss in der Datei `cpuset.cpus` eine 2 eingetragen werden. Welche Dateien von den einzelnen Subsystemen zur Verfügung gestellt werden und welche Funktion sie haben, kann der Kernel-Dokumentation [1] entnommen werden.

### Implementierung – Zu Fuß

Control Groups können einfach mit Betriebssystemmitteln erstellt werden. Folgende Schritte sind notwendig, um eine Hierarchie zu erstellen, die es ermöglicht die Priorisierung auf CPU-Ebene und die Memory-Nutzung von Prozessen zu beeinflussen. Als erstes muss das virtuelle File-System, das die Grundlage für die Hierarchie bildet, eingebunden werden.

```
mkdir -p /cgroup/hierarchie_1
mount -t cgroup -o cpu,memory hierarchie_1 \
/cgroup/hierarchie_1
```

Durch diese Befehle wird eine Hierarchie bestehend aus den beiden Subsystemen `cpu` und `memory`, die eine Control Group enthält, angelegt. Dieses virtuelle File-System enthält die Dateien, über die die Control Group konfiguriert

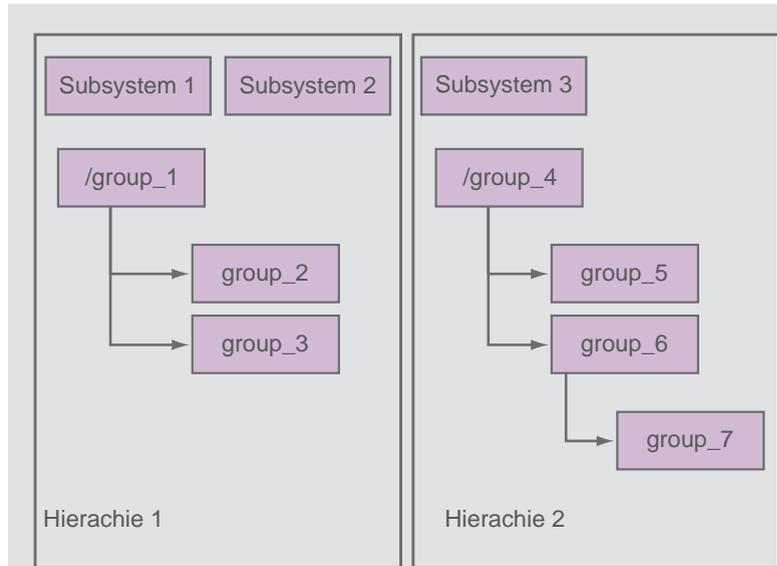


Abb. 1: Aufbau und Konzept von cgroups.

Subsystem	Beschreibung
cpuset	Zuweisen von CPUs und Memory Nodes zu Control Groups.
cpu	Beeinflussen des CPU-Schedulers. Beispielsweise kann das Prioritätsverhältnis der einzelnen cgroups untereinander geregelt werden.
cpuacct	Generierung von automatischen Reports der CPU-Nutzung der Prozesse in einer Gruppe.
memory	Setzen von Memory-Limits für eine Gruppe von Prozessen, Auswertungen zur Memory-Nutzung von Prozessen in der Gruppe, Einstellen der „swapiness“ für eine Gruppe.
devices	Erlauben bzw. verbieten von Zugriffen von Tasks in einer Gruppe auf Devices.
freezer	Suspend und Resume von Prozessen in einer Gruppe
net_cls	Markieren von Netzwerkpaketen mit einer classid zu einer Gruppe. Diese kann mit dem tc ausgewertet und interpretiert werden.

Abb. 2: Auswahl von Subsystemen.

werden kann. In der Datei `tasks` sind standardmäßig alle Prozesse des Systems enthalten. Weitere Control Groups in dieser Hierarchie können folgendermaßen angelegt werden:

```
cd /cgroup/hierarchie_1
mkdir apache oracle
```

Durch das einfache Anlegen der beiden Ordner werden die Gruppen `apache` und `oracle`

```

mount {
    cpu = /cgroup/hierarchie_1;
    memory = /cgroup/hierarchie_1;
}

group apache {
    cpu {
        cpu.shares = 1024;
    }
}

group oracle {
    cpu {
        cpu.shares = 2048;
    }
}

```

Abb. 3: Auszug aus einer cgconfig.conf.

<user>:<process>	<subsystem>	<destination>
@oracle	cpu,memory	oracle
*:httpd	cpu,memory	apache
* *		default

Abb. 4: Auszug aus einer cgrules.conf.

erstellt. Diese enthalten die gleichen Dateien wie der Ordner /cgroup/hierarchie\_1 nur, dass die beiden tasks-Dateien leer sind. Soll nun ein Prozess von der oberen Control Group z. B. in die Gruppe apache verschoben werden, reicht das Kommando:

```
echo <PID> > /cgroup/hierarchie_1/apache/tasks
```

Alle Prozesse, die nun von dem verschobenen Prozess gestartet werden, sind nun automatisch in der Gruppe apache angelegt. In welchen Control Groups sich ein Prozess befindet, kann in der Datei /proc/<PID>/cgroup nachgeschlagen werden. Um zu erfahren, welche Gruppen zur Zeit auf dem Rechner aktiv sind, reicht die Eingabe von `ps -o cgroup`. Die Gewichtung von Prozessen in einer Gruppe, im Bezug auf die anderen Gruppen, wird über die Datei `cpu.shares` bestimmt. Der Wert in dieser Datei muss immer im Zusammenhang mit den Werten in den Dateien der anderen Gruppen gesehen werden. Bei einem Default-Wert von 1024 bedeutet eine Verdoppelung auf 2048, dass die Prozesse in dieser Gruppe eine doppelt so hohe Priorität haben die CPU zu nutzen wie die Prozesse in den anderen Gruppen.

```
echo 2048 > /cgroup/hierarchie_1/oracle/cpu.shares
```

Genau wie in diesen Beispielen lassen sich die anderen Subsysteme und deren Parame-

ter bedienen. Mit den Control Groups lässt sich somit ein umfangreiches Ressourcen Management unter Linux realisieren.

### Implementierung – libcgroup

Das Problem an der gerade vorgenommenen Konfiguration ist, dass sie nur bis zum nächsten Neustart hält und auf Dauer recht kompliziert zu implementieren ist. Damit nicht jeder seine eigenen Skripte zur Verwaltung der Control Groups entwickeln muss und um die Administration zu vereinfachen wurde die Bibliothek libcgroup entwickelt. Sie vereinfacht mit einigen Tools und Konfigurationsdateien die Administration von Gruppen und ermöglicht somit den Einsatz in einer produktiven Umgebung.

Die Bibliothek libcgroup bringt zwei Daemons mit, die über jeweils eine Datei in ihrem Verhalten beeinflusst werden können. Der cgconfig-Service ermöglicht es persistente Gruppen zu konfigurieren, die bei jedem Systemstart automatisch angelegt und mit den konfigurierten Parametern versehen werden. Dazu liest der Daemon die Datei /etc/cgconfig.conf ein und wertet sie aus. In Abbildung 3 ist ein Auszug aus einer Beispielkonfiguration abgebildet.

Der zweite Daemon heißt cgregd und dient dazu mit Hilfe der Regeln in der /etc/cgrules.conf Prozesse in die gewünschte Control Group zu verschieben. Diese Regeln basieren auf dem Namen und dem Owner eines Prozesses (siehe Abbildung 4).

Über die beiden Daemons hinaus bietet die Library noch einige Tools mit denen cgroups zur Laufzeit administriert werden können. Die folgende Auflistung ist eine kurze Beschreibung der nützlichsten Werkzeuge:

- **cgcreate**  
Erstellen von Control Groups zur Laufzeit.
- **cgset**  
Setzen von Parametern eines Subsystems in einer cgroup.
- **cgexec**  
Starten eines Prozesses in einer bestimmten Gruppe. Dies ignoriert die Einstellungen in der cgrules.conf.
- **cgclassify**  
Verschieben von Prozessen in eine andere cgroup.
- **cglear**  
Löschen aller Control Groups.

Um einen Service direkt in einer bestimmten Gruppe zu starten, ohne ihn in der `cgrules.conf` zu hinterlegen, ist es möglich in seiner `/etc/sysconfig`-Datei die Variable `CGROUP_DAEMON` zu setzen.

## Fazit und Blick in die Zukunft

Da die Control Groups seit der Version 2.6.24 fester Bestandteil des Linux-Kernels sind, ist es in aktuellen Distributionen möglich die Technik direkt über das Kernel-Interface zu nutzen. Mit der Bibliothek `libcgroup` steht zudem eine Tool-Sammlung zur Verfügung, mit der sich die Gruppen „administratorenfreundlich“ anwenden lassen. In Red Hat Enterprise Linux 6 werden Control Groups, genau wie in diesem Artikel beschrieben, inklusive `libcgroup` nutzbar sein [2]. Somit stehen auch in den zukünftigen Enterprise Distributionen alle Türen offen, um Ressourcen auf Linux komfortabel zu managen. Die Entwicklung der `cgroups` steht jedoch noch nicht an ihrem Ende. Besonders im Bereich der

## Links

- ▶ [1] Kernel Dokumentation:  
<http://www.kernel.org/doc/Documentation/cgroups/>
- ▶ [2] RHEL 6 Documentation:  
[http://www.redhat.com/docs/en-US/Red\\_Hat\\_Enterprise\\_Linux/6-Beta/pdf/Resource\\_Management\\_Guide/Red\\_Hat\\_Enterprise\\_Linux-6-Resource\\_Management\\_Guide-en-US.pdf](http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6-Beta/pdf/Resource_Management_Guide/Red_Hat_Enterprise_Linux-6-Resource_Management_Guide-en-US.pdf)

Subsysteme wird sich in Zukunft noch einiges bewegen. Bereits angekündigt ist z. B. das Subsystem `blkio`, mit dem sich der Zugriff auf Block-Devices reglementieren lässt. So wird es möglich sein, Prozessen unterschiedliche Gewichtungen beim Schreiben oder Lesen auf Devices zu zuweisen. Im Allgemeinen kann den Control Groups nur eine gute Zukunft vorausgesagt werden, da sie Linux um eine mächtige Technologie erweitern.



Marius Dorlöchter  
([info@ordix.de](mailto:info@ordix.de))

## Seminarempfehlung: Linux Systemadministration

▶ **Informationen/Online-Anmeldung:** <http://www.ordix.de/trainingsshop/siteengine/action/load/kategorie/Betriebssysteme/nr/85/index.html>

In diesem Seminar werden Sie mit der Installation, Konfiguration und Administration eines Linux-Systems vertraut gemacht. Neben den allgemeinen Grundlagen der Administration von Unix-Systemen lernen Sie die Besonderheiten und Erweiterungen im Linux-Umfeld kennen. Der Umgang mit Geräte-dateien und dem Logical Volume Manager (LVM) wird ebenso behandelt wie das Startverhalten oder die Benutzerverwaltung.

### Seminarinhalte

- Installation
- Systemstart und Runlevel-Konzept
- Softwaremanagement mit RPM
- Der Kernel: Funktionsweise, Module, Installation, Konfiguration
- Bootmanager GRUB
- Druckereinrichtung mit CUPS
- Netzwerkkonfiguration
- Systemmeldungen
- Benutzerverwaltung
- Festplattenkonfiguration und Dateisysteme (ext2/3)
- Einbinden externer Verzeichnisse (NFS)
- Prozessverwaltung
- Zeitgesteuerte Jobverwaltung (cron, at)
- Diagnose-Tools
- Software RAID (Multi Devices)
- Gerätedateien: `udev`
- Logical Volume Management (LVM) und Software-RAID
- Swapping/Paging
- Diagnose und Troubleshooting
- Tipps und Tricks aus der Praxis
- Vertiefung der Theorie durch praktische Übungen und Beispiele

### Termine

15.11. - 19.11.2010 in Wiesbaden

**Seminar-ID:** BS-02

**Dauer:** 5 Tage

**Preis pro Teilnehmer:** 1.690,00 € (zzgl. MwSt.)

**Frühbucherpreis:** 1.521,00 € (zzgl. MwSt.)

Wir führen unsere Seminare auch jederzeit an einem geeigneten Ort Ihrer Wahl durch und bringen, wenn nötig, auch das entsprechende Equipment mit. Informieren Sie sich am besten im Internet über unsere Kundenseminare und die mobilen Schulungen: <http://training.ordix.de>.

# Quantensprung von Oracle RAC-Version 11gR1 auf 11gR2

Der Artikel richtet sich an Administratoren, die Oracle, RAC 11gR2 aufsetzen und betreiben möchten.

Die Änderungen im Oracle Real Application Cluster (RAC) beim Versionswechsel von 10gR2 auf 11gR1 waren überschaubar. Mit der Version 11gR2 hat Oracle nun allerdings gravierende Erweiterungen eingeführt. Dieser Artikel berichtet über diesen Riesenschritt von der Version 11gR1 auf 11gR2. Ebenso erhalten Sie einen Einblick über weitere Einsatzmöglichkeiten der neuen Grid Infrastructure.

## Grid Infrastructure und RAC

Die Installation der Clusterware wird ab der Version 11gR2 durch die Installation der Grid Infrastructure ersetzt. Die Grid Infrastructure ist ein Paket aus Clusterware, Automatic Storage Management (ASM) und Listener.

Mit der Installationsoption: „Grid Infrastructure für einen Cluster installieren und konfigurieren“ wird die Grid Infrastructure auf allen beteiligten Server installiert. Die Grid Infrastructure und die Datenbank-Software sollten jeweils unter einen eigenen OS-Benutzer installiert werden. Somit trennt Oracle die Infrastructure von der Datenbank-Software.

Die Oracle Cluster Registry (OCR)- und Voting-Disk liegen ab 11gR2 in einer Disk-Gruppe im ASM oder in einem zertifizierten Cluster File System. Neu ist die Oracle Local Registry (OLR), in der lokale Cluster-Informationen gespeichert werden.

Das **Spfile** der ASM-Instanzen wird nun in einer Disk-Gruppe im ASM platziert.

Auch ältere Datenbankversionen können die Grid Infrastructure verwenden. Hier gibt es allerdings diverse Probleme, siehe Metalink-Note: 948456.1.

Für Windows-32bit-Systeme wird die Grid Infrastructure und somit der RAC nicht mehr angeboten.

## Grid Infrastructure und Oracle RAC One Node

Ein neues Produkt in 11gR2 ist der RAC One Node. Dieser ist eine Art Aktiv/Passiv-Cluster von Oracle. Die Datenbank läuft als Single-Instanz innerhalb der Grid Infrastructure nur auf einem Knoten.

Einsatzzwecke:

- Failover der Datenbank, beim Ausfall des aktiven Server auf einen anderen Server der Grid Infrastructure. Der Anwender kann nach einem kurzen Ausfall mit der Datenbank weiterarbeiten.
- Für Wartungsarbeiten und zur Lastverteilung können die Datenbanken auf einen anderen Server der Grid Infrastructure mit dem Werkzeug `omotion` verlegt werden.

Für die Verwendung von RAC One Node wird die Installation der Grid Infrastructure und Datenbanksoftware wie beim RAC durchgeführt.

RAC One Node ist eine Preisooption der Enterprise Edition. Ein Spare-Server innerhalb der Grid Infrastructure kann lizenzfrei betrieben werden, allerdings darf eine Datenbank nur 10 Tage produktiv pro Jahr auf diesem Spare-Server liegen.

Mit dem Werkzeug `racone2rac` kann eine RAC One Node-Datenbank nach RAC konvertiert werden.

### Grid Infrastructure für Einzelknoten (Oracle Restart)

Die Grid Infrastructure kann auch auf Einzelknoten installiert werden und zwar mit der Installationsoption: „Grid Infrastructure für einen Standalone Server installieren und konfigurieren“. Single-Instanzen und Listener werden auf dem Einzelknoten über die Clusterware, nach einem Neustart des Servers, automatisch neu gestartet. Zudem versucht die Clusterware nach Fehlern die Single-Instanz und/oder den Listener neu zu starten. Mit dem Werkzeug `srvctl` müssen die Single-Instanzen/Datenbanken administriert werden.

### Cluster Time Synchronisation Service (CTSS)

Die Grid Infrastructure liefert einen eigenen Time Server mit. Soll das Network Time Protocol (NTP) nicht verwendet werden, so muss NTP vor der Installation auf den Cluster-Knoten gestoppt und deaktiviert werden. Der Cluster Time Synchronisation Service (CTSS) sorgt, statt NTP, für synchrone Zeiten auf den Cluster-Knoten. Während der Installation der Grid Infrastructure installiert und konfiguriert sich der Time Server automatisch.

### Neuer Installer

Mit der Oracle Version 11gR2 wird ein komplett neuer Installer zur Installation der Grid Infrastructure und auch der Datenbank-Software mitgeliefert (siehe Abbildung 1).

Mit dem Befehl `runInstaller` wird der Installer, wie gehabt, gestartet. Während der Installation erfolgt die Herstellung der Benutzeräquivalenz (SSH-Verbindung zwischen Cluster-Knoten) über den Installer. Alle Prüfungen die vor Version 11gR2 mit dem Cluster Verification Utility (`cluvfy`) noch manuell durchgeführt werden mussten, werden nun komplett aus dem Installer heraus ausgeführt. Probleme während dieser Prüfung können über einen Fix-It-Job bereinigt werden. Der Installer bietet auch die Möglichkeit direkt ein Response-File zur Silent-Installation zu erstellen.

Bei der Installation der Datenbank-Software ist nun keine Custom-Installation mehr möglich. Es sind nur noch bestimmte Optionen wählbar.

### Deinstallation Grid Infrastructure

Die Deinstallation der Grid Infrastructure und somit der Clusterware ist nun einfacher

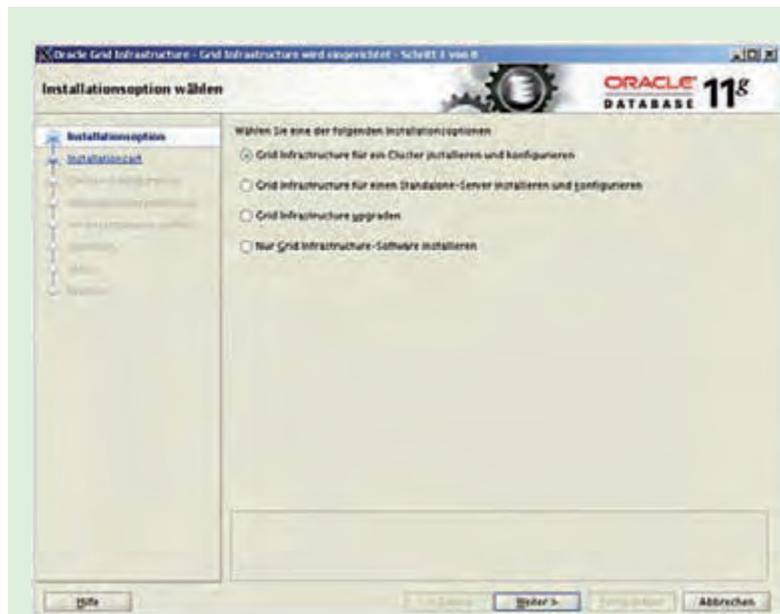


Abb. 1: Installer Grid Infrastructure.

```
RACDB =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP) (HOST = suse_cluster (PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = racdb)
    )
  )
)
```

The text 'suse\_cluster' in the HOST field is circled in red, and a red arrow points from the word 'SCAN' above it to the circle.

Abb. 2.: tnsnames.ora mit SCAN.

durchzuführen. Folgende Schritte müssen für die Deinstallation durchgeführt werden:

- Als Benutzer root:  
\$GRID\_Home/crs/install/rootcrs.pl –verbose –deconfig –force  
(auf allen Knoten)
- Auf dem letzten Knoten:  
Als Benutzer root: \$GRID\_Home/crs/install/rootcrs.pl –verbose –deconfig –force –lastnode  
(vom letzten Knoten aus wird die OCR- und Voting-Disk gelöscht)
- Als Benutzer grid (Installationsbenutzer der Grid Infrastructure):  
\$GRID\_Home/deinstall/deinstall

```
EZConnect:
sqlplus system/manager@suse_cluster:1521/racdb

JDBC:
jdbc:oracle:thin:@suse_cluster:1521/racdb
```

Abb. 3: Verbindungsaufbau mit EZConnect und JDBC.

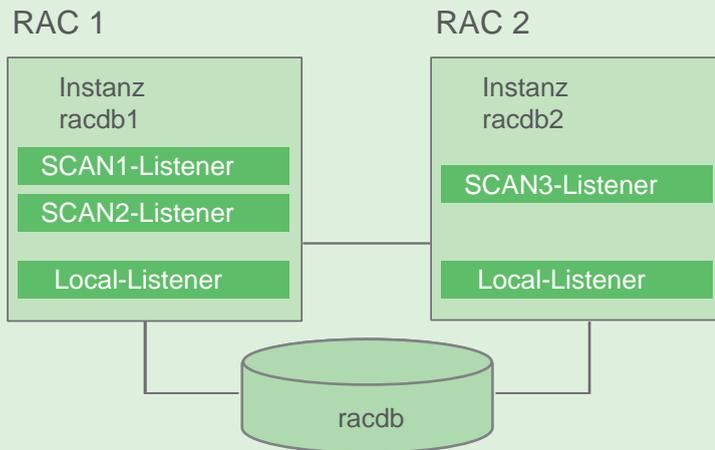


Abb.4: SCAN Listener auf einem Zwei-Knoten RAC.

senvergabe im RAC. Die einzig statische IP-Adresse ist die GNS-VIP-IP. Dazu wird vorab eine GNS-Subdomain im Public-Netzwerk und DHCP konfiguriert. Die GNS-VIP-IP „läuft“ auf einem Knoten im Cluster. Während der Installation der Grid Infrastructure kann der Cluster mit oder ohne GNS konfiguriert werden. Eine manuelle Konfiguration der VIP-Adressen und des Single Client Access Namen entfällt. Empfehlenswert ist die Konfiguration mit GNS für große Cluster, weil die Netzwerkkonfiguration beim Löschen und Anlegen neuer Knoten im Cluster entfällt.

### Single Client Access Name (SCAN)

Mit dem Single Client Access Namen (SCAN) kann nun mit nur einem Namen auf die Datenbank zugegriffen werden (siehe Abbildung 2). Ein Verbindungsaufbau zur RAC-Datenbank kann nun auch einfach mit EZConnect oder JDBC über den zentralen und serverunabhängigen SCAN erfolgen (siehe Abbildung 3). Der SCAN sorgt für ein Load Balancing und Failover der Clientverbindungen zur RAC-Datenbank.

Der Single Client Access Name (SCAN) wird während der Installation der Grid Infrastructure konfiguriert.

Wird GNS nicht verwendet, so müssen im DNS-Server dem SCAN drei IP-Adressen zugeordnet werden (Vorgabe von Oracle). Oracle konfiguriert daraufhin drei SCAN-Listener. Die IP-Adressen und somit die SCAN-Listener werden vom Client nach dem Round-Robin-Prinzip verwendet. Ein Ping auf den SCAN liefert also eine der drei im DNS-Server konfigurierte IP-Adressen zurück. In einem Zwei-Knoten-Cluster befinden sich zwei SCAN-Listener auf dem ersten und der dritte SCAN-Listener auf dem zweiten Knoten (siehe Abbildung 4). Der SCAN-Listener reicht die Client-Anmeldung an einem Local-Listener weiter. Diese Weiterverbindung zum Local-Listener erfolgt nach bestimmten konfigurierten Bedingungen:

- Server-seitiges Load-Balancing oder
- feste Zuordnung der Anmeldung auf die gleiche RAC-Instanz

SCAN-Listener sind eigene Ressourcen in der Clusterware. Fällt ein Knoten aus, so verlegt die Clusterware die SCAN-Ressource/n auf einen verfügbaren Knoten im RAC. Die SCAN-Listener können aber auch manuell mit dem Werkzeug `srvctl` verlegt werden.

### Neues Werkzeug für die Verwaltung von ASM

Mit der Version 11gR2 können durch den DBCA keine ASM-Instanzen/Plattengruppen verwaltet werden. Oracle stellt dafür das neue grafische Programm ASM Configuration Assistant (ASMCA) zur Verfügung.

### Grid Plug And Play (GPnP)

Mit Grid Plug and Play (GPnP) vereinfacht Oracle das Hinzufügen und Löschen von Knoten im RAC. Für GPnP wird ein zentrales Profile verwendet, das auf jedem Knoten im RAC liegt. Wird bei der Installation von Grid Infrastructure Grid Naming Service (GNS) konfiguriert, entfällt sogar die Netzwerkkonfiguration, weil GNS im Cluster für eine dynamische IP-Adressenvergabe sorgt.

### Grid Naming Service (GNS)

In allen RAC-Versionen waren die IP-Adressen im Cluster statisch. Grid Naming Service (GNS) sorgt für eine dynamische IP-Adres-

Ohne die Verwendung eines DNS-Servers, lässt sich die Grid Infrastructure somit nicht installieren. Es sei denn, man schafft einen Workaround und ersetzt die `nslookup`-Datei auf OS-Ebene durch ein entsprechendes Shell-Skript (nur für Testzwecke). Oracle überprüft nämlich während der Installation der Grid Infrastructure mit dem Werkzeug `nslookup`, ob von einem DNS-Server IP-Adressen zurückgeliefert werden.

Es können auch weiterhin die VIP-Adressen (virtuelle IP-Adressen) für eine Client-Verbindung zur RAC-Datenbank verwendet werden.

### Cluster Datenbank Konfiguration

Die Datenbank im Cluster kann mit unterschiedlichen Konfigurationsoptionen erstellt werden:

Admin-verwaltet

- Erstellen der Datenbank auf definierten Servern
- Administration erfolgt, wie bisher, manuell

Policy-verwaltet

- Dynamische Instanzverwaltung
- Erstellen der Datenbank auf Server-Pool Ressourcen

### Neue Befehle und Ressourcen

Der Statusbefehl `crs_stat -t` [1] ist in 11gR2 deprecated. Der neue Statusbefehl `crsctl status resource -t` oder abgekürzt `crsctl stat res -t` bietet einen guten Überblick über die Cluster-Ressourcen an. Der Befehl `crs_stat -t` funktioniert zwar noch, allerdings zeigt dieser nur noch die Ressource-Datenbank und nicht mehr die Instanzen des Clusters an. Ab Version 11gR2 können nun alle Ressourcen von einem Server aus mit `crsctl stop cluster -all` gestoppt werden.

Es gibt einige neue Ressourcen in RAC 11gR2. Disk-Gruppen sind nun auch Ressourcen, ebenfalls gibt es `network`, `oc4j` und die `SCAN`-Listener.

### Fazit

Mit dem Versionssprung von RAC 11gR1 auf 11gR2 hat Oracle sehr viele Erweiterungen eingebaut, so dass man schon von einer neuen Version sprechen kann. Die Version 11gR2 ist schon längere Zeit auf dem Markt, doch

### Glossar

<b>ASM</b>	Automatic Storage Management (ASM) ist ein im Oracle Kernel integrierter Volume Manager. Er ist ab Oracle 10g verfügbar und arbeitet nach dem SAME-Prinzip (stripe and mirror everything). ASM kann als zentraler Storage für RAC-Systeme verwendet werden.
<b>DBCA</b>	Database Configuration Assistant (DBCA). Grafischer Assistent zur Erstellung und Konfiguration von Oracle Datenbanken und anderen Oracle Komponenten.
<b>EZConnect</b>	Einfache Verbindungsmethode zur Datenbank, ohne Verwendung der <code>tnsnames.ora</code> .
<b>Grid Infrastructure</b>	Paket aus Clusterware, ASM und Listener.
<b>OCR</b>	Oracle Cluster Registry Disk. Die OCR Disk ist eine Datei, in der die Oracle Clusterware die gesamte Cluster-Konfiguration abspeichert.
<b>Oracle Restart</b>	Grid Infrastructure für Einzelknoten.
<b>RAC One Node</b>	Aktiv/Passiv-Cluster von Oracle.
<b>SCAN</b>	Single Client Access Name. Der SCAN ermöglicht die Verbindung mit nur einem „Namen“ auf die RAC-DB.
<b>SRVCTL</b>	Server Control Utility-Tool zur Administration von RAC.
<b>Voting-Disk</b>	In der Voting-Disk trägt sich der Master-Server ein, der beim Verlust des Interconnects „überlebt“.

### Links

- ▶ [1] Metalink-Note 948456.1: <http://metalink.oracle.com>
- ▶ [2] Seminarempfehlung „Oracle 11gR2“: <http://training.ordix.de/siteengine/action/load/kategorie/Datenbanken/nr/125/index.html>

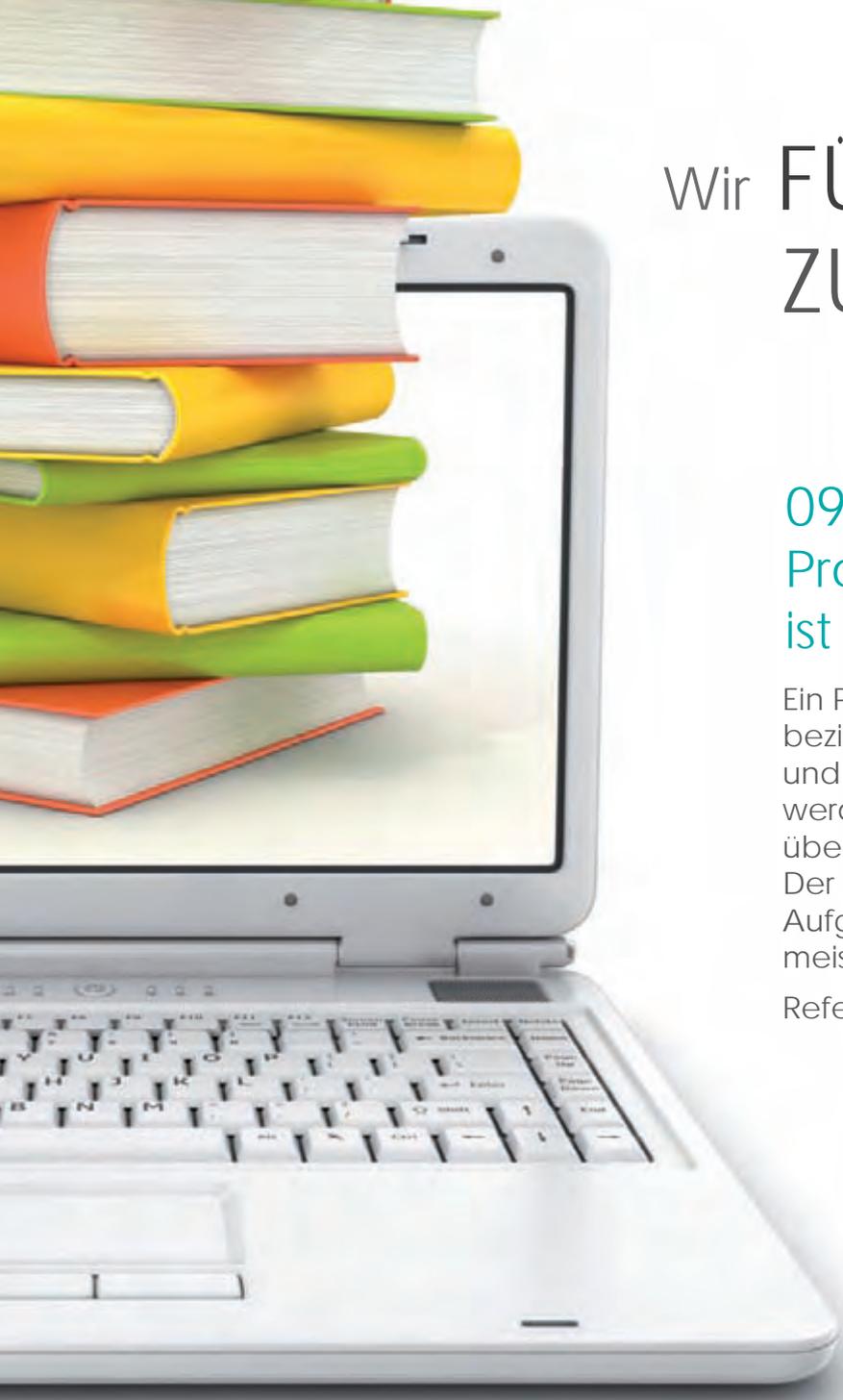
aufgrund der vielen Änderungen gehen erst jetzt vermehrt produktive RAC 11gR2-Umgebungen an den Start. Eine Herausforderung stellt die Migration älterer RAC-Versionen auf die neue Version 11gR2 dar.

In unserem Seminar „Oracle 11gR2: Real Application Cluster (RAC) & Grid Infrastructure“ [2] erhalten Sie einen tiefen Einblick in die Oracle Grid Infrastructure und Oracle RAC Architektur 11gR2. Neben der Installation der Oracle Grid Infrastructure, Oracle RAC Software und Erstellung einer Cluster-Datenbank werden Sie mit der Administration und Fehlerbehebung sowie der Konfiguration des Load-Balancing und Failover und dem Thema Backup- und Disaster Recovery im RAC 11gR2 vertraut gemacht.

Zukünftig werden wir Sie in weiteren Artikeln über die Neuheiten des Real Application Cluster auf dem Laufenden halten.



Michael Skowasch  
([info@ordix.de](mailto:info@ordix.de)).



# Wir **FÜHREN** Interessierte **ZUSAMMEN!**

09.11.2010  
**Projektmanagement  
ist Management**

Ein Projektmanager ist ein Manager - beziehungsweise sollte es sein. Um so größer und komplexer ein Projekt ist, desto wichtiger werden die Managementfähigkeiten gegenüber den fachinhaltlichen Kompetenzen. Der Referent beschreibt die verschiedenen Aufgaben und Rollen, die ein Projektmanager meistern muss.

Referent: Benedikt Georgi, ORDIX AG

Der „Treffpunkt RheinMainIT“ wird quartalsweise in unserem Trainingszentrum in Wiesbaden durchgeführt und bietet Ihnen ein abwechslungsreiches Vortragsprogramm zu aktuellen IT-Themen.

Der Vortrag beginnt um 18.00 Uhr, im Anschluss rundet ein Informationsaustausch bei einem kleinen Imbiss die Veranstaltung ab.

Treffpunkt  
**RheinMain IT**

**Kostenfreie  
Veranstaltungsreihe!**

Informationen und Anmeldung unter [www.RheinMainIT.de](http://www.RheinMainIT.de)