

## Tipps & Tricks zum Aufbau einer Unternehmenscloud mit CC13c

**Stefan Seck, Borys Neselovskyi**  
**OPITZ CONSULTING Deutschland GmbH**

### Schlüsselworte

Oracle Cloud Control, Self Service Portal, Charge Pack, Cloud, Exadata, VLAN Tagging, ZFS Storage Appliance, Pluggable Database

### Einleitung

Seit mehreren Jahren bietet die Firma Oracle seinen Kunden eine breite Palette von Lösungen. Die berühmte Oracle Datenbank ist fast eine Legende und sehr beliebt und verbreitet. Die Middleware Lösungen wie Oracle Forms oder SOA Suite finden auch sehr große Akzeptanz und werden in der Regel mit der Datenbank zusammen betrieben. In den letzten Jahren hat Oracle einige Server (Begriff: *Oracle Engineered Systems*) auf die Markt gebracht, die für das Betreiben von Oracle Software Lösungen prädestiniert sind.

Eine *Exadata* Maschine ist für den Betrieb von unternehmenskritischen, hoch frequentierten Datenbanken entwickelt. Die beste Hardware Ausstattung in der Kombination mit speziellen Software Lösungen wie *Smart Scan* machen eine Exadata zu der besten Datenbank-Plattform.

Etwas kleinere *Oracle Database Appliance* kann sowohl die Datenbanken als auch Middleware Produkte hosten. Diese Maschine ist für den Betrieb bei mittelständischen Unternehmen geeignet.

Die ZFS Storage Appliance ist eine sehr mächtige Speicherlösung, die einen performanten Speicher zur Verfügung stellt und für die Speicherung von Datenbank – Sicherungen bestens geeignet ist. Die Unterstützung von *Snapshot* Technologie sowie von modernen Replikationsmechanismen machen *ZFS Storage Appliance* zu den besten Speicherlösungen auf der Markt.

Viele Kunden betreiben erfolgreich heute Oracle Produkte in eigenen Rechenzentren. Es gibt aber auch viele Unternehmen, die gern das Betreiben von Infrastrukturen auslagern und sich nur auf den Kerngeschäft konzentrieren wollen. So können Firmen die Personal- und Infrastrukturkosten sparen und diese in Betriebskosten umwandeln (CAPEX zu OPEX).

Auch die Marktführer haben diese Tendenz erkannt und bieten Cloud Service an, wobei die gesamte Firmeninfrastruktur oder einige Teilbereiche ausgelagert werden können. Die Firmen wie Oracle, Amazon und Microsoft bieten zahlreiche Cloud Lösungen in Bereichen:

- IaaS (Infrastructure as a Service)
- PaaS (Platform as a Service)
- DaaS (Database as a Service)
- SaaS (Software as a Service) an.

Dabei handelt es sich um so genannte Public Cloud: physikalische Server stehen in Rechenzentren des Cloud Anbieters. Der Zugriff auf die Cloud Dienste erfolgt durch die zahlreiche Tools, die die Cloud Betreiber zur Verfügung stellen.

Einige Probleme und offene Fragen machen den Umzug in die Public Cloud für viele Unternehmen allerdings noch sehr schwierig bis fast unmöglich. Fragen wie "Sind sensible Firmendaten in der Cloud sicher?", "Wie funktioniert das Zusammenspiel zwischen Cloud- und On-Premise-

Komponenten?" oder "Wer ist für die Performance und Sicherheit in hybriden Umgebungen verantwortlich?" können momentan nur sehr schwer und unvollständig beantwortet werden.

Viele Firmen würden gerne die Cloud Angeboten nutzen, die Daten aber lieber in einem eigenen Rechenzentrum aufbewahren.

Genau für diese Anforderungen kommt das Modell „private Cloud“ ins Spiel: Die benötigte Infrastruktur wird im Kundenrechenzentrum aufgebaut und von Partner betreut. Der Kunde überlässt dem Partnerunternehmen den Betrieb der Umgebung und konzentriert sich auf den Kerngeschäft. Der Kunde hat aber die Hoheit über eigene Daten.

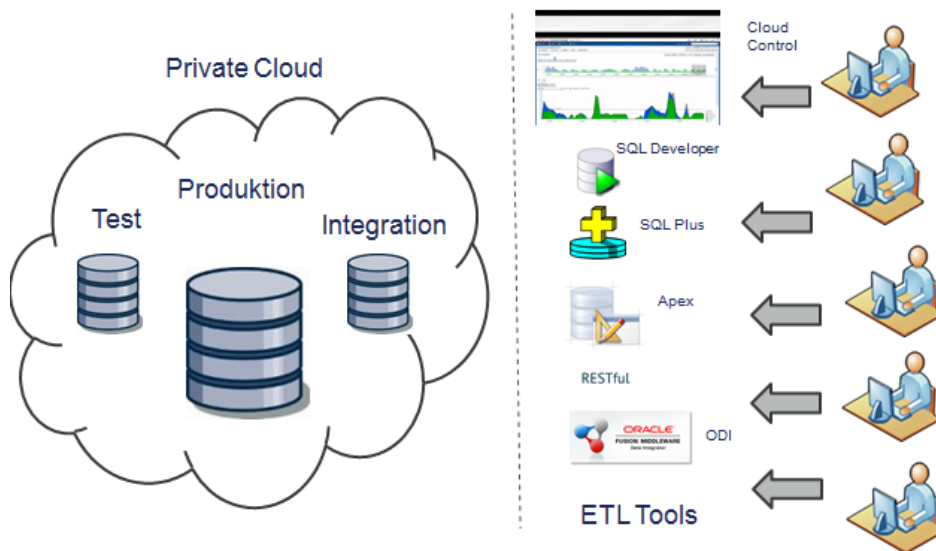


Figure 1: Cloud Nutzung

Die Abbildung zeigt, dass Cloud Nutzer via zahlreiche Tools auf Datenbanken zugreifen. Die Wartung der Infrastruktur übernimmt der Cloud Anbieter.

## Architektur

### Topologie

Für Oracle-lästige Umgebungen ist es durchaus sinnvoll, die private Cloud mit der Hilfe von *Oracle Engineered Systems* zu realisieren. Sehr wichtig ist dabei, die Kriterien festzulegen, die für den Betrieb maßgeblich sind. Dazu zählen die Service-Level-Agreement (SLA) Vereinbarungen, die Antwortzeiten einer Anwendung bzw. einer Datenbank oder minimalen Durchsatz von Daten.

Für 24 x 7 Umgebungen könnte nur ein Rechenzentrum nicht ausreichen, für den Katastrophenfall muss ein Konzept mit einem Ausfallrechenzentrum evaluiert werden.

Die hochkritischen und performance-intensive Datenbanken sollen auf einer Exadata betrieben werden. Durch die exzellente Hardware Ausstattung, sehr performanten Storage und durch die Infiniband Anbindung zählt Exadata zu den besten und leistungsfähigen Maschinen auf dem Markt.

Darüber hinaus verfügt eine Exadata über einige spezielle Software, die beste Performance und Stabilität bei dem Betrieb von Oracle Datenbanken gewährleistet:

- Exadata Smart Scan (Smart I/O):
  - o Unterstützung für JSON und XML Analytics Vorgänge
  - o Storage Indizes
- Hybrid Columnar Compression (HCC):
  - o Reduzierung der Storage Nutzung durch Komprimierung
- Smart Flash Cache
  - o Abfragen werden schneller ausgeführt
  - o Voraussetzung: RDBMS 12.1.0.2
- Database Filesystem (DBFS): Dateien (z.B. Word oder Excel Dokumente) können von einem System (z.B. von einem FTP Server) direkt in die Datenbank geladen werden
- In-Memory Database: gesamte Datenbank wird im Speicher betrieben
- DB Snapshots (Pluggable Database)
  - o Smart Scans und Smart Flash Cache sind unterstützt
- Exadata Events in AWR Reports

Kleinere Datenbanken mit mäßigen IO-Verhalten können auf einer *Oracle Database Appliance* (ODA) betrieben werden. ODA X5 unterstützt alle Enterprise Edition Eigenschaften. Die Datenbanken können auf einer ODA als RAC, RAC One Node oder Single Instanzen aufgebaut werden. Außerdem können auf der ODA Virtual Plattform Oracle Middleware Produkte installiert und betrieben werden.

Die neue ODA X6 ist etwas kleinere Distribution, die für die Standard Edition Datenbanken geeignet ist.

Die Abbildung gibt einen Überblick über die eingesetzte Oracle Engineered Systems und die Verwendung:

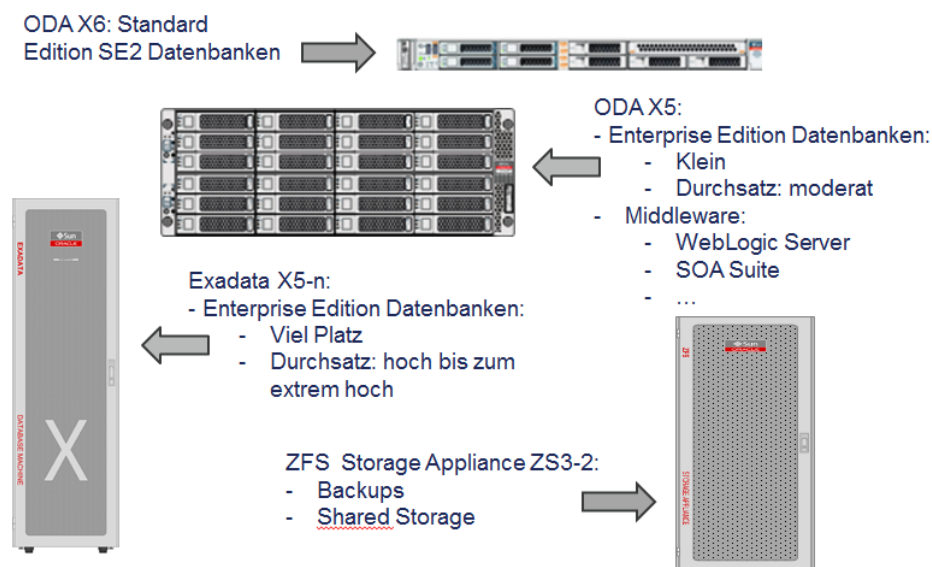


Figure 2: Engineered Systems: Überblick

Denkbares Architekturszenario für eine private Cloud: in zwei Rechenzentren sind Oracle Engineered Systems. Die Rechenzentren sind über performantes und leistungsfähiges Netzwerk verbunden.

Im Rechenzentrum 1 werden folgende Server gehostet:

System	Anzahl	Verwendung
Exadata X5-2 Quarter Rack	1	Primäre Datenbanken Produktion EE
ODA X5-2 (VP)	1	Primäre Datenbanken Produktion EE Oracle Middleware
ODA X6-2M	1	Datenbanken SE2
ZFS Storage Appliance ZS3-2	1	Backups Shared Storage

Im Rechenzentrum 2 werden folgende Maschinen aufgebaut:

System	Anzahl	Verwendung
Exadata X5-2 Half Rack	1	Standby Datenbanken Produktion EE Datenbanken Entwicklung EE Datenbanken Test EE
ODA X5-2 (VP)	1	Standby Datenbanken Produktion EE Oracle Middleware
ZFS Storage Appliance ZS3-2	1	Backups Shared Storage

Die nachfolgende Abbildung zeigt die Verteilung von Engineered Systems über zwei Rechenzentren:

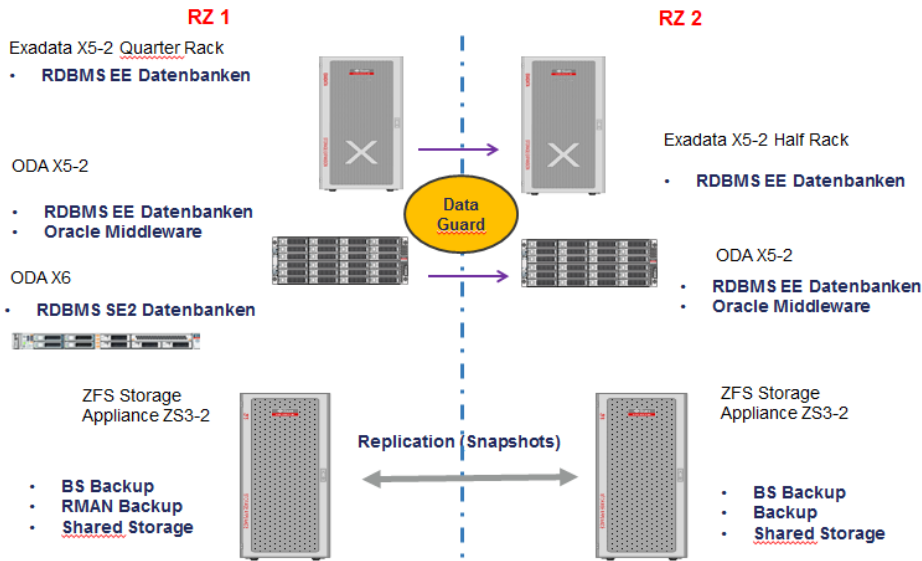


Figure 3: Topologie

### Datenbanken

Datenbanken werden mit der Option *Pluggable Database* installiert und in einem *Real Application Cluster (RAC)* betrieben. Mittels Oracle Data Guard werden diese in das zweite Rechenzentrum repliziert:

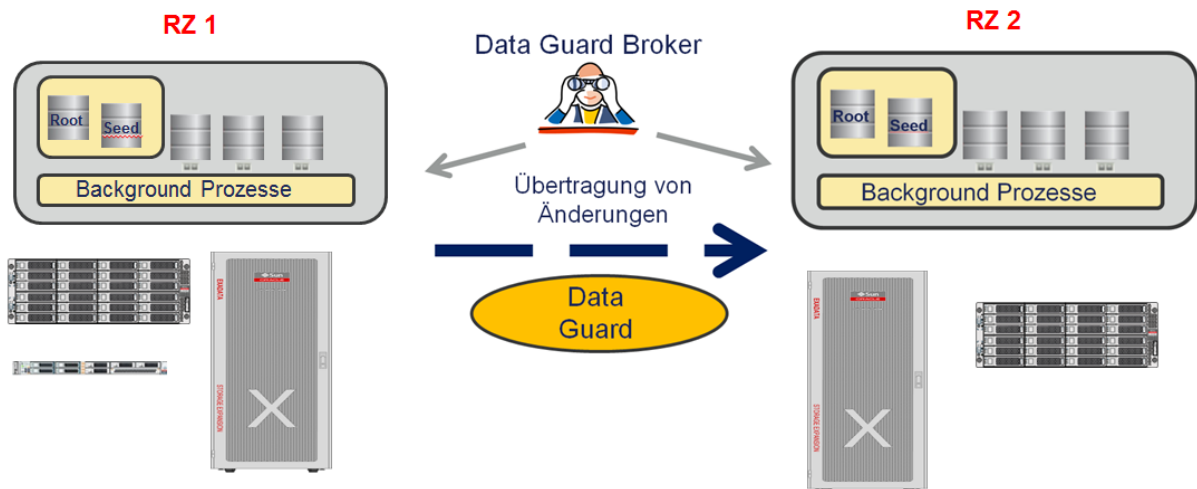


Figure 4: RDBMS Architektur

### Isolierung von Umgebungen

In einer privaten Cloud sollen unterschiedliche Umgebungen betrieben werden. Kunde möchte die Anwendungen in einer Testumgebung entwickeln, in der Abnahmeumgebung diese ausgiebig testen. Die Performance Tests sollen in einer Lasttestumgebung betrieben werden. Eine Referenzumgebung dient zum Nachstellen von Problemen im Betrieb und zur Erstellung von Bugfixes.

Am besten sollen unterschiedliche Umgebungen komplett isoliert voneinander installiert werden. Die produktive Umgebung soll den höchsten Sicherheitsstandards entsprechen. In einer Entwicklungsumgebung spielt die Sicherheit weniger wichtige Rolle.

Wie trenne ich unterschiedliche Umgebungen innerhalb einer Exadata oder ODA netzwerktechnisch?

ODA und Exadata können virtualisiert werden. Außerdem mit dem Technik *VLAN Tagging* können unterschiedliche Netzwerksegmente innerhalb einer Maschine konfiguriert werden. VLAN Tagging basiert auf der 802.1Q Technologie und ermöglicht, zusätzliche Netzwerke über das Public Netzwerk zu bauen. Dabei werden in Cluster Registry zusätzliche Netzwerke konfiguriert. Die Cluster Ressourcen (Datenbanken, Services, Listener, Virtuelle IP Adressen) werden einem virtuellen Netzwerk zugeordnet. So entstehen unterschiedliche virtuelle Netzwerke mit dazugehörigen Ressourcen, die eine Trennung von z.B. Test- und Produktionsumgebung ermöglicht.

### *VLAN Tagging Konfiguration*

Die Konfiguration erfolgt nur auf den Datenbankservern. Speichereinheiten (Storage Cells) dürfen nicht geändert werden.

Die Netzwerke müssen in Vorfeld geplant werden. Die bestehende Infrastruktur im Rechenzentrum soll bereits existieren. Die externe 1G/10G Ethernet Switches sollen bereits für VLAN Tagging konfiguriert und an die Exadata Datenbankserver (DBM) via Ethernet Kabel eingebunden werden.

Die Netzwerkschicht sollen Bonding unterstützen. Genaue Anweisungen für die Switch Konfiguration sind auf jedem Exadata Datenbank Server in der Datei `/usr/share/doc/kernel-doc-<Version>/Documentation/networking/bonding.txt` gespeichert.

Die die Netzwerkschnittstelle `bondeth0` ist für VLAN Tagging verantwortlich und ist bereits während der Installation konfiguriert. Die zusätzlichen virtuellen Netzwerke können über das Interface `bondeth0` eingebunden werden. Dafür sollen folgende Schritte unter der Benutzerkennung `root` durchgeführt werden:

- Das Linux Paket `vconfig` muss auf Datenbankservern installiert werden:

Prüfen: `rpm -qa | grep vconfig`  
Ausgabe (Beispiel): `vconfig-1.9-8.1.el6.x86_64`

- Kernel Module für VLAN Tagging sollen bereits installiert werden (Clusterware Version 11.2.0.2 oder höher):

Prüfen: `find /usr/src/kernels/^uname -r*/net -name 8021q`  
Ausgabe (Beispiel): `/usr/src/kernels/2.6.39-400.248.3.el6uek.x86_64/net/8021q`

- Prüfen, ob VLAN Schnittstellen existieren (in diesem Beispiel sind keine Interfaces konfiguriert):

Prüfen: `cat /proc/net/vlan/config`  
Ausgabe: `cat: /proc/net/vlan/config: No such file or directory`

- Für die reibungslose VLAN Funktionalität muss das Paket `initscripts` (Version 8.45.38-2.0.1 und höher) installiert werden

Prüfen: `rpm -qa | grep initscripts`

Ausgabe (Beispiel): `initscripts-9.03.46-1.0.4.el6_6.1.x86_64`

- Die zusätzliche VLAN Netzwerk kann mit der Hilfe des Shell Skriptes `dbm-vlan-config.sh` konfiguriert werden. Das Skript ist von Oracle bereitgestellt und kann von der MOS Note 1423676.1 heruntergeladen werden. Im nachfolgenden Beispiel wird ein Netzwerk mit der ID 1064, Netzwerkadresse, Netmask und Gateway konfiguriert. Alle Konfigurationsdateien werden automatisch erzeugt. Dabei entsteht ein Netzwerkinterface: `bondeth0.1064`. Die Schnittstelle wird sofort aktiviert und ist rebootfest:

```
./dbm-vlan-config.sh add -v 1064 -i bondeth0 -ip 10.12.36.175 -n 255.255.255.0 -g 10.12.36.1
```

Mit dem Kommando „`./dbm-vlan-config.sh show`“ kann man die konfigurierte Konfiguration anzeigen lassen:

```
./dbm-vlan-config.sh show
```

Ausgabe:

```
[Info] VLAN (8021q) support is enabled in system
bondeth0.1064 | 1064 | bondeth0 | 10.12.36.175 | 255.255.255.0 | up
```

In den nächsten Schritten binden wir das Netzwerk 1064 in Cluster. Diese Aufgaben sollen vom Eigentümer der Clusterinstallation (in der Regel User *grid*) durchgeführt werden

- Die Cluster Umgebung muss geladen werden:

```
. oraenv
ORACLE_SID = [ASM1] ? +ASM1
The Oracle base has been set to /u01/app/oracle
```

- Mit dem Kommando `oifcfg` wird ein neues Public Netzwerk auf Basis von VLAN 1064 konfiguriert werden (Dabei wird eine neue IP Adresse festgelegt: in diesem Beispiel: 10.12.36.180):

```
oifcfg setif -global bondeth0.1064/10.12.36.180:public
```

- Das neue Netzwerk soll jetzt unter der Benutzerkennung *root* gestartet werden:

```
srvctl add network -netnum 1064 -subnet 10.12.36.180/255.255.255.0/bondeth0.1064 -nettype static
```

- Die VIP Adresse hinzufügen (als *root*):

```
srvctl add vip -node exa01db01 -netnum 1064 -address exav106401-
vip/255.255.255.0/bondeth0.1064
```

- VIP Adresse starten und prüfen(als User *grid*):

```
srvctl start vip -i exav106401-vip
srvctl status vip -n exa01db01
```

```
VIP exav106401-vip is enabled
VIP exav106401-vip is running on node: exa01db01
```

- Jetzt kann ein Datenbank Listener konfiguriert und gestartet werden (als grid):

```
srvctl add listener -listener LIST_1064 -p 1521 -netnum 1064 -s
srvctl start listener -listener LIST_1064
```

- Jetzt soll die Datenbank für die Nutzung des Listeners LIST\_1064 konfiguriert werden. Dafür soll der Datenbank Parameter local\_listener angepasst werden:

```
sqlplus / as sysdba
```

```
SQL> alter system set
local_listener='(DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP)(HOST=
exav106401-vip)(PORT=1521))))' sid='orcl1';
```

Folgende Dokumentation ist sehr hilfreich:

- MOS Note 1423676.1 „Enabling 802.1Q VLAN Tagging in Exadata Database Machine over client networks “
- MOS Note 972500.1 „How to Modify SCAN Setting or SCAN Listener Port after Installation “
- Oracle Dokumentation:  
[http://docs.oracle.com/cd/E50790\\_01/doc/doc.121/e51950/network.htm#DBMIN20989](http://docs.oracle.com/cd/E50790_01/doc/doc.121/e51950/network.htm#DBMIN20989)

### *Backups via ZFS Storage Appliance*

Die ZFS Storage Appliance (ZA) ist eine ideale Plattform für die Speicherung von Datenbank Backups. Durch die direkte Anbindung via Infiniband an Exadata sind die Durchsatzraten sehr gut, was bei hoch frequentierten Datenbanken absolut notwendig ist. Die ZFS Storage Appliance unterstützt die Hybrid Columnar Compression (HCC): damit komprimierte Daten können auch für die Datenbank Clonen genutzt werden, ohne diese auspacken zu müssen.

Die Anbindung ZA an Exadata erfolgt über das Protokoll *Direct NFS* (dNFS) einzusetzen.

Um die Datenbanken auf einer ZA sichern zu können, müssen einige Vorarbeiten geleistet werden:

- Storage Profile:  
*Mirrored Storage Profile* wird in der Regel in der Exadata Umgebung konfiguriert. Es ist erforderlich für die OLTP Datenbanken mit dem hohen Durchsatz und für die inkrementelle Sicherungen.  
Für traditionelle RMAN Backups können singly parity oder double parity Profile in Betracht gezogen werden.
- Netzwerk: Oracle empfiehlt den Einsatz von Infiniband. Die active-active Multipath Gruppen sollen konfiguriert werden.



- Protokoll: Oracle Direct NFS soll verwendet werden
- Konfiguration ZFS *Projects* und *Shares*: Projekte werden benutzt für die logische Trennung von Backups. Man kann ein Projekt für eine Gruppe von Datenbanken oder ein Projekt pro Datenbank erstellen. In Projekten können verschiedene Shares (Speicherbereiche) konfiguriert werden
- Anbindung von ZFS Shares an eine Exadata: das Protokoll NFSv3 oder NFSv4 sind unterstützt.

- Mit dem Kommando `dcli` kann man auf allen Datenbankservern notwendigen Dienste (`rpcbind`, `nfslock` und `nfs`) aktivieren:

```
dcli -l root -g /home/oracle/dbs_group chmod 644 /etc/hosts.allow
dcli -l root -g /home/oracle/dbs_group chmod 644 /etc/hosts.deny
dcli -l root -g /home/oracle/dbs_group chkconfig rpcbind on
dcli -l root -g /home/oracle/dbs_group service rpcbind start
dcli -l root -g /home/oracle/dbs_group chkconfig nfslock on
dcli -l root -g /home/oracle/dbs_group service nfslock start
dcli -l root -g /home/oracle/dbs_group chkconfig nfs on
dcli -l root -g /home/oracle/dbs_group service nfs start
```

- Ganz wichtig ist die NFS Mount Optionen richtig einstellen. Die MOS Note 359515.1 befasst sich mit dem Thema ausgiebig. Folgende NFS Optionen sind üblich:

```
rw,bg,hard,nointr,rsize=1048576,wsiz=1048576,tcp,vers=3,timeo=600
```

- Erstellen *oranfstab* Datei: im `ORACLE_HOME/dbs` Verzeichnis soll die Datei *oranfstab* erstellt werden. Hier werden für alle Datenbanken die NFS Pfade gespeichert. Das Beispiel der Konfiguration kann aus der MOS Note 2087231.1 entnommen werden
- Auswahl der Backup Strategie: empfohlen wird RMAN Incremental Backup Strategie.
  - Schnellere tägliche Backups mit weniger Bandbreitenverbrauch durch den Einsatz von Block Change Tracking
  - Backups verbrauchen weniger Platz
  - Optimal für die Bandsicherung
- Konfiguration des RMAN Backup Service für jede Datenbank: mit dem Kommando `srvctl` soll für jede Datenbank ein Service konfiguriert werden.
- Vorbereitung der Datenbank für die Sicherung:
  - Block Change Tracking aktivieren
  - Anzahl von RMAN Kanälen einstellen
  - RMAN Tuning:
    - Parameter Section Size: empfohlen: 100G
    - Parameter Fileperset: empfohlen: 3
  - RMAN Kanäle einstellen (Beispiel):

```
RMAN> configure channel 1 device type disk connect
sys/welcome1@orc11_bkup1 format '/zssa/orc11/nckup1/%U';
```

- RMAN Skript (Beispiel):

```
run
```

```
{
sql 'alter system set "_backup_disk_bufcnt"=64 scope=memory';
sql 'alter system set "_backup_disk_bufsz"=1048576 scope=memory';
sql 'alter system set "_backup_file_bufcnt"=64 scope=memory';
sql 'alter system set "_backup_file_bufsz"=1048576 scope=memory';
backup as backupset incremental level 0 section size 100g database tag
'FULLBACKUPSET_L0';
backup as backupset section size 100g archivelog all not backed up 2 times
tag
'ARCHIVELOG';
backup current controlfile tag 'CONTROLFILE';
}
```

#### Nützliche Links:

- Configuring a Single Oracle ZFS Storage Appliance into an InfiniBand Fabric with Multiple Oracle Exadata Machines: <http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/multiple-exadata-zfssa-121013-2080035.pdf>
- Guidelines When Using ZFS Storage in an Exadata Environment (Doc ID 2087231.1)
- Oracle ZFS Storage: FAQ: Exadata RMAN Backup with The Oracle ZFS Storage Appliance (Doc ID 1354980.1)
- Oracle ZFS Storage: FAQ: Exadata RMAN Backup with The Oracle ZFS Storage Appliance (Doc ID 1354980.1)
- Oracle ZFS Storage Appliance: FAQ - RMAN Backup From ODA to Oracle ZFS Storage Appliance (Doc ID 1665928.1)
- Oracle ZFS Storage Appliance Administration Guide: [https://docs.oracle.com/cd/E27998\\_01/html/E48433/shares.html#scrolltoc](https://docs.oracle.com/cd/E27998_01/html/E48433/shares.html#scrolltoc)

#### **Services (SSE und BNE)**

Iaas

Paas

MWaaS

DBaaS

#### **Aufbau Self Service Portal (SSE)**

- Charge Pack

#### **Demo**

Aufbau PDB

Aufbau Middleware

Demo Beschreibung

**Fazit**

**Kontaktadresse:**

Stefan Seck

OPITZ CONSULTING Deutschland GmbH

Lazarettstraße 15

45127 Essen

Telefon: +49 (0) 2261-6001 0

Mobil: +49 (0) 173-7279792

E-Mail stefan.seck@opitz-consulting.com

Internet: [www.opitz-consulting.com](http://www.opitz-consulting.com)

Borys Neselovskyi

OPITZ CONSULTING Deutschland GmbH

Kirchstraße 6

D-51647 Gummersbach

Telefon: +49 (0) 2261-6001 0

Fax: +49 (0) 173-7279029

E-Mail borys.neselovskyi@opitz-consulting.com

Internet: [www.opitz-consulting.com](http://www.opitz-consulting.com)