

Oracle Snap Management Utility – Snap mir eine

Marco Schwab
Witt-Gruppe
Weiden i. d. Opf.

Schlüsselworte

Klonen, Snap Management Utility, ZFS, Testsysteme, Exadata

Die Migration unserer unternehmensweiten Datenbank von einer IBM P-Serie mit EMC Storage auf eine Exadata hat uns bei der Witt-Gruppe vor neue Herausforderungen beim Aufbau unserer Testsysteme gestellt. Bisher wurden Storage-Technologien der EMC verwendet, um alle Disks der Produktiv-Datenbank zu kopieren und damit eine Test-Datenbank zu starten. Durch den EMC Mechanismus „Space Efficient Clone“ wurde dafür nur ein Bruchteil des ursprünglichen Speicherplatzes benötigt.

Jetzt wird eine Lösung benötigt, die dies in Verbindung mit einer Exadata ermöglicht. Oracle bietet mit seinem „Snap Management Utility“ (SMU) in Verbindung mit einer ZFS Storage Appliance (ZFSA) die Möglichkeit, per Knopfdruck Backups und Klone einer Datenbank zu erzeugen.

Ziel dieses Papers ist es, dieses Tool vorzustellen und einen Überblick über Installation und Verwendung zu geben. Dazu wird zunächst die Umgebung der Witt-Gruppe und deren Anforderungen an Testsysteme vorgestellt. Anschließend wird die Installation des SMU beschrieben und kurz auf die Konfiguration der ZFS-Shares eingegangen. Zum Schluss werden die Möglichkeiten der Klon-Erstellung mit dem SMU genannt und an einem Beispiel näher erläutert.

Datenbank-Umgebung der Witt-Gruppe

Die Umgebung der Witt-Gruppe besteht aus 2 Exadata X6-2 Quarter-Racks mit je einer ZFSA im Rack integriert. Die beiden Systeme sind in unterschiedlichen Rechenzentren verteilt, um so eine maximale Ausfallsicherheit zu gewährleisten. Abbildung 1 zeigt dabei die Verteilung der Datenbanken.

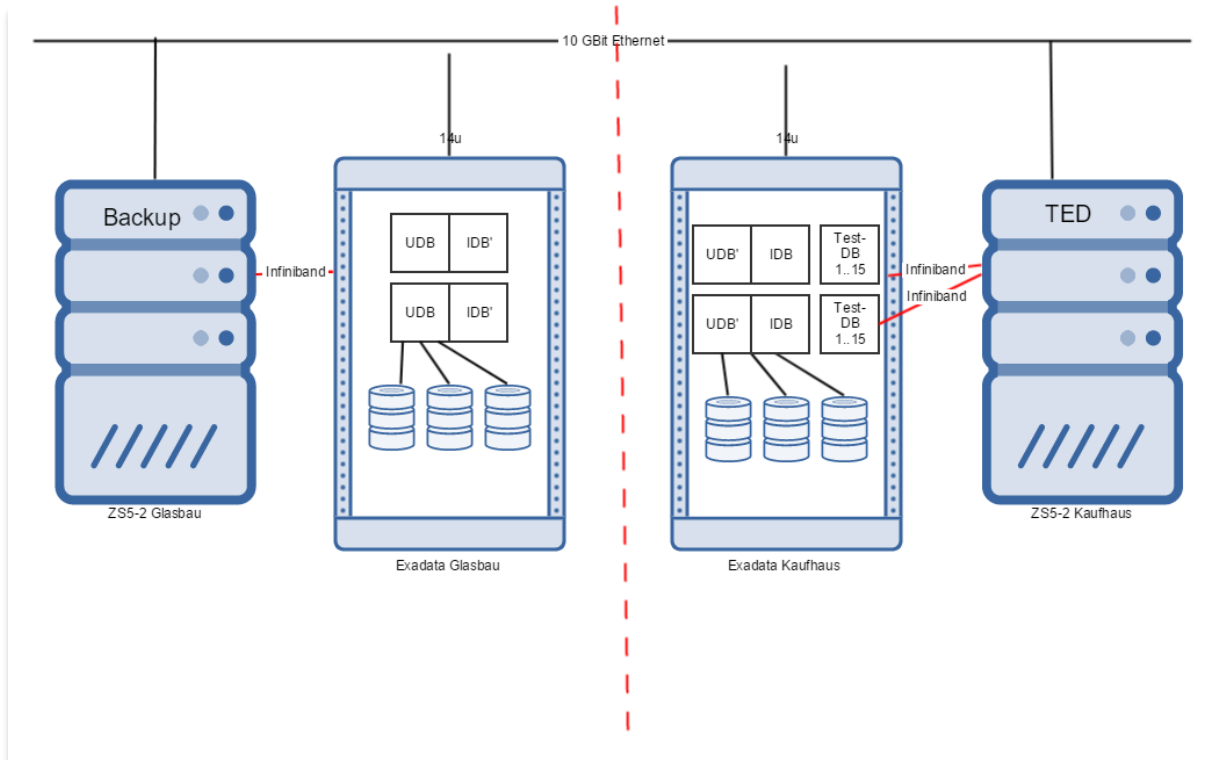


Abbildung 1: DB-Landschaft Witt-Gruppe

Das Produktiv-System UDB läuft auf dem System EXA03 als RAC-Datenbank und besitzt eine Dataguard Standby-Datenbank auf dem System EXA04 im anderen Rechenzentrum. Eine weitere Dataguard Standby-DB dient dabei als Masterklon für die Erstellung der Testsysteme. Ihre Daten liegen jedoch nicht auf den Storage-Zellen der Exadata, sondern auf Exporten der ZFSA. Die RAC-Datenbank IDB ist ein Voll-Klon der UDB und befindet sich auf dem System EXA04. Sie dient als praxisnahe Integrations-Datenbank und wird daher über RMAN-Duplicate von der UDB erstellt. Daneben befinden sich auf EXA04 noch bis zu 15 Testsysteme, deren Daten auf geklonten Exporten der ZFSA liegen und durch das SMU erstellt werden.

Anforderungen an die Testsysteme

Innerhalb der Witt-Gruppe haben wir folgende Anforderungen an unsere Testsysteme:

- Die Möglichkeit bis zu 15 unabhängige Klone nahezu gleichzeitig zu erstellen
- Zeit zum Erstellen eines Klons (ohne Individualisierung) unter einer Stunde
- Möglichst geringer Storage-Bedarf (im Verhältnis zur Änderungsrate)
- Erstellung muss über Skripte automatisierbar sein

All diese Anforderungen werden durch das SMU erfüllt. Das Tool kümmert sich um den kompletten Klonprozess und macht es möglich eine Vielzahl an Klone gleichzeitig zur Verfügung zu stellen. Die Erstellung eines Klon ist je nach Konfiguration der Datenbanken in wenigen Minuten möglich. Außerdem bietet die ZFSA ähnliche Techniken wie EMC, wodurch ein Klon, je nach Änderungsrate der Daten, einen Bruchteil der Größe des Originals benötigt. Das integrierte „Command Line Interface“ (CLI) ermöglicht es, problemlos per Skript auf das Tool zuzugreifen.

Installation unter Linux

Das Oracle SMU wird auf einem eigenen Server unter Oracle Linux 7.3 installiert. Oracle liefert für jede Version ein RPM-Package aus, welches mittels Paket-Manager installiert wird. Lediglich eine aktuelle JRE wird benötigt. Da das Tool als Root-User installiert und gestartet werden muss, wird kein Oracle-User benötigt.

Nachdem das File aus MOS heruntergeladen wurde, kann es, wie in Listing 1 beschrieben, installiert werden.

```
cd /opt
unzip p25042289_130_Linux-x86-64.zip

cd Oracle_Snap_Management_Utility_Linux-1.3.2-3/
rpm -i oracle-smu-1.3.2-3.noarch.rpm
```

Listing 1: Installation des SMU

Für die Nutzung des „Beautiful User Interface“ (BUI) bzw. des „Command Line Interface“ (CLI) müssen anschließend noch die Ports freigegeben werden, falls die Firewall des Oracle Linux aktiv ist. Die Freigabe der beiden benötigten Ports ist in Listing 2 beschrieben.

```
firewall-cmd --zone=public --list-all

# BUI
firewall-cmd --permanent --zone=public --add-port=8443/tcp
# CLI
firewall-cmd --permanent --zone=public --add-port=8002/tcp

service firewalld restart
```

Listing 2: Freischalten der Firewall-Ports für BUI und CLI

Nun muss lediglich der Service als root gestartet werden. Damit bei einem unerwarteten Boot das Tool von selbst startet, wird der Service noch per „systemctl“ in die Bootsequenz eingereiht. Listing 3 beschreibt das Starten und das Aktivieren des Auto-Starts.

```
service smud start

systemctl enable smud
```

Listing 3: Starten des SMU

Konfiguration der ZFS-Shares

Damit SMU eine Datenbank klonen kann, müssen alle Daten (Datafiles, Archivelogs, Temp-Files und Controlfiles) auf Exporten der ZFSA liegen. Um für jeden File-Typ die optimale Performance zu erreichen, werden nach Empfehlung von Oracle (DocID 2087231.1) die in Tabelle 1 dargestellten Shares angelegt. Diese Einstellungen werden beim Klonen automatisch übernommen. Daher sollten auch die Anforderungen an die späteren Testsysteme bei der Konfiguration berücksichtigt werden.

Oracle Files	Record Size	Sync Write Bias	Read Cache	Compression	Share Name
Datafiles	32K	latency	all data and metadata	LZ4	data
Tempfiles	128K	latency	do not use cache devices	LZ4	temp
Archivelogs	1M	throughput	do not use cache devices	LZ4	archive
Redologs	128K	latency	do not use cache devices	off	reco
Controlfiles	8K	latency	all data and metadata	off	control

Tabelle 1: Share-Konfiguration für Testsysteme

Bei den Datafiles sind allgemein mehrere Empfehlungen für die Blocksize zu finden. Man sollte hier berücksichtigen, ob hauptsächlich Single-Block-Reads oder eher Full-Table-Scans ausgeführt werden und je nachdem eine kleine oder große Blocksize wählen. Für uns hat sich entgegen den Empfehlungen von Oracle eine Blocksize von 128K für die Datafiles bewährt.

Die erstellten Exporte müssen anschließend auf einem Host, in unserem Fall der Exadata, gemountet werden. Dazu empfiehlt Oracle im MOS (DocID 359515.1) folgende Mount-Optionen für Linux:

```
rw,bg,hard,nointr,rsize=32768,wsiz=32768,tcp,actimeo=0,vers=3,timeo=600
```

Listing 4: Mount-Optionen für Linux

Klon-Erstellung mit SMU

Für die Erstellung eines Klons wird immer eine Ausgangs-DB benötigt. Diese kann jedoch auf verschiedenen Wegen zur Verfügung gestellt werden. Die beiden Möglichkeiten sollen hier kurz erläutert werden:

- Klonen aus RMAN Backup

Hat man ein RMAN Image-Copy auf einem ZFS-Export liegen, kann dieses als Master-DB genutzt werden. Die Funktion „Import RMAN Backup Image“ erstellt daraus eine lauffähige DB, aus der Klone erstellt werden können. Das Image kann dann über inkrementelle Backups immer wieder erneuert werden. Leider hat diese Methode den Nachteil, dass alle Klone zerstört werden müssen bevor das Image erneuert werden kann. Es ist also nicht möglich unabhängige Klone zu erstellen, die zu unterschiedlichen Zeiten neu aufgebaut werden können.

- Klonen aus einer Datenbank

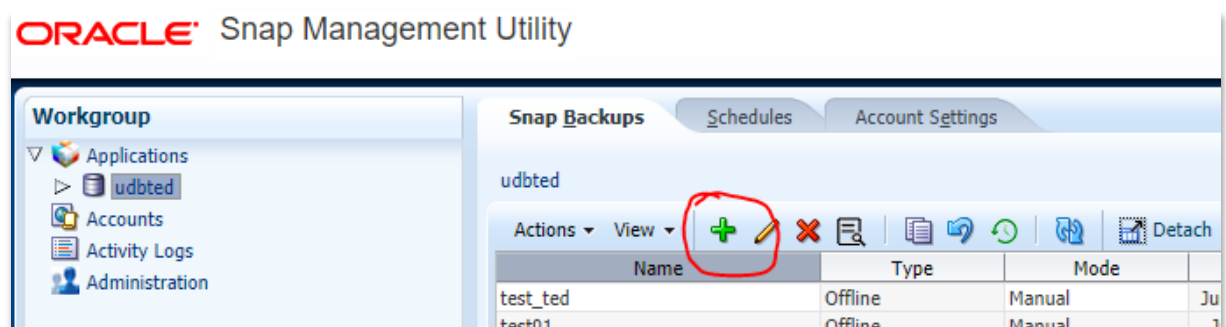
Alternativ kann man eine bestehende Datenbank, deren Files auf der ZFSA liegen, einlesen. Daraus kann man dann, genau wie beim RMAN-Image, Klone erzeugen. Da die Klon-Erstellung in der Regel online stattfindet, wird die Datenbank nicht beeinflusst und keine Downtime benötigt. Außerdem ist es möglich eine Data Guard Standby-Datenbank zu nutzen. Dies hat den Vorteil, dass die Master-Datenbank automatisch auf einem aktuellen Stand ist. Leider ist bei dieser Variante die Erstellung des für den Klon benötigten Backups nur offline möglich, weshalb die Standby-Datenbank eine kurze Downtime benötigt, die in der Regel nicht länger als 1 Minute dauert.

Beispiel-Erstellung einer Klon-DB

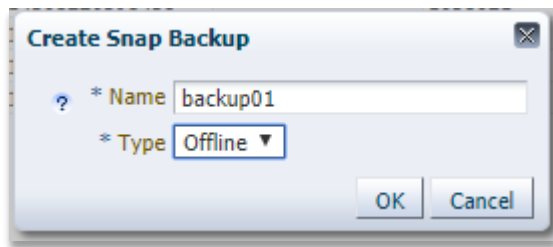
Da bei der Witt-Gruppe Testsysteme in oft sehr kurzen Intervallen erstellt werden müssen, ist die Variante mit einem RMAN Image zu umständlich. Das verwendete Verfahren zum Erstellen eines Klons mittels Data Guard Standby-Datenbank als Master-Klon wird nun beispielhaft beschrieben.

Zum Erstellen eines Klons sind nur zwei Schritte notwendig. Zuerst wird ein Backup einer im SMU hinterlegten Datenbank erzeugt. Aus diesem Backup können anschließend beliebig viele Klone erzeugt werden.

1. Erstellen eines Backups



Hat man einen Master-Klon unter Applications ausgewählt, kann man mit einem Klick auf das Plus-Symbol ein neues Backup erzeugen. Dabei öffnet sich folgender Dialog.



Hier wird ein Name für das Backup und die Art des Backups ausgewählt. Dabei kann grundsätzlich zwischen Online und Offline gewählt werden. Voraussetzung für den Online-Modus ist, dass sich die Datenbank im Archivelog-Modus befindet. Bei Standby-Datenbanken ist jedoch nur die Backup-Erstellung im Offline-Modus möglich.

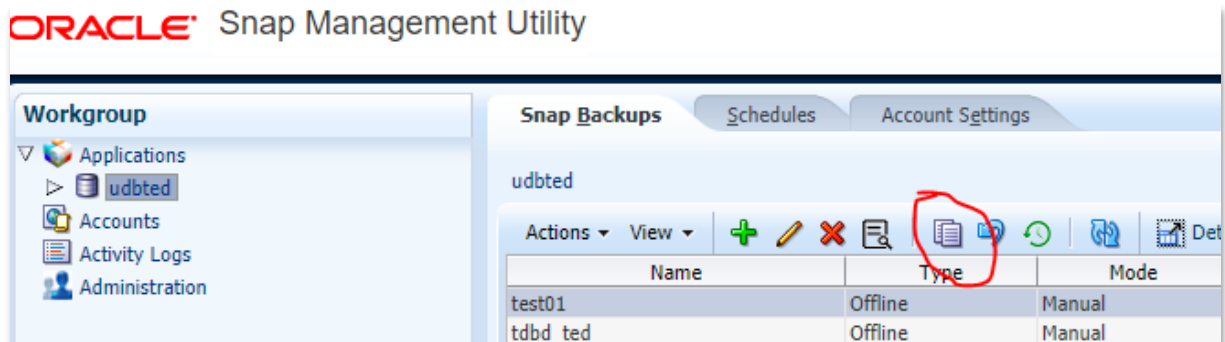
Technisch wird beim Erstellen eines Backups ein Snapshot der für die Datenbank relevanten Shares auf der ZFSA erzeugt. Daraus können später neue Shares erzeugt werden.

Das Log zeigt deutlich, welche Schritte das SMU zur Erstellung eines Backups ausführt.

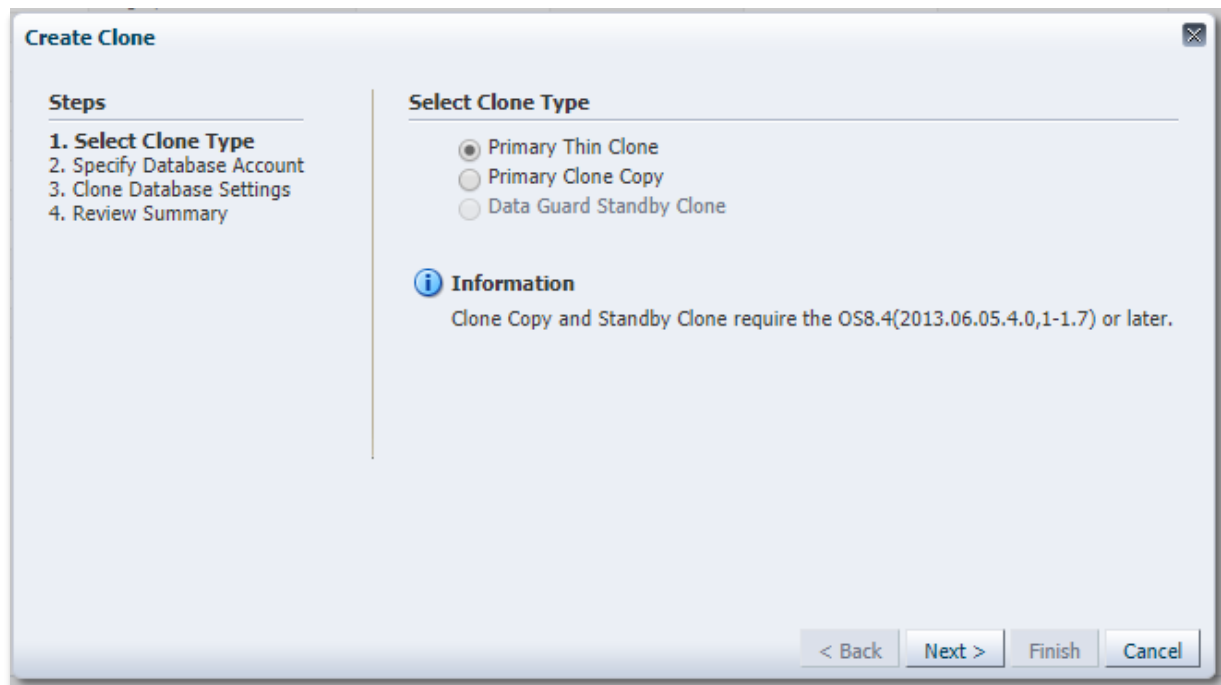
```
Backup all file shares
Look up shares mounted on database host: OK
Remote share list
zfs02h01:/export/ted/data
zfs02h01:/export/ted/temp
zfs02h01:/export/ted/archive
zfs02h01:/export/ted/reco
zfs02h01:/export/ted/control
Look up storage shares: OK
Storage share list
/export/ted/data
/export/ted/temp
/export/ted/archive
/export/ted/reco
/export/ted/control
Perform pre-backup operations
Check database status: OK
Shutdown database: OK
Create snapshots
Snapshot shares on storage head 0: OK
Perform post-backup operations
Put database back to previous status
Start database in MOUNTED: OK
Start managed recovery: OK
```

Listing 5: Auszug aus Log des Backupvorgangs

2. Erstellen des Klons aus dem Backup



Um einen Klon aus dem erstellten Backup zu erzeugen wird das Backup markiert und der Menüpunkt „Create Clone“ angeklickt. Anschließend können in einem Dialog die Eigenschaften des Klons festgelegt werden.



Klone können auf mehrere Arten erstellt werden. Beim Erstellen eines „Thin Clones“ werden auf der ZFSA Shares erstellt, die lediglich Zeiger auf die Originaldaten enthalten. Erst wenn die Originaldaten geändert werden oder neue Daten auf den Shares geschrieben werden, wird der Zeiger aufgelöst und die Daten physikalisch geschrieben. Dadurch benötigt ein Klon nur einen Bruchteil der Speicherkapazität des Originals, je nach Änderungsrate der Daten. Dieser Mechanismus nennt sich „copy on write“. Bei einer „Clone Copy“ werden die Shares auf der ZFSA physikalisch kopiert. Der Klon braucht so exakt die gleiche Kapazität wie das Original.

Create Clone

Steps

1. Select Clone Type
- 2. Specify Database Account**
3. Clone Database Settings
4. Review Summary

Specify Database Account - Thin Clone

* Host Account:

? * Application Name:

? * Database Name:

? * SID:

? * Database Unique Name:

* Listener Port:

X * Database Configuration:

? * Database Home:

User:

* Password:

* Confirm Password:

< Back Next > Finish Cancel

Als nächstes werden die grundlegenden Einstellungen für den Klon festgelegt. Dazu zählen der Host, auf dem der Klon erstellt wird, sowie einige datenbankspezifische Parameter. Hervorzuheben ist hier die Datenbank Konfiguration. Bei der Erstellung kann gewählt werden, ob die Datenbank als „Single Instance“, als „RAC“, oder als „RAC One Node“ erstellt werden soll.

Create Clone

Steps

1. Select Clone Type
2. Specify Database Account
- 3. Clone Database Settings**
4. Review Summary

Clone Database Settings - Thin Clone

* System Global Area Size(int|K|M|G):

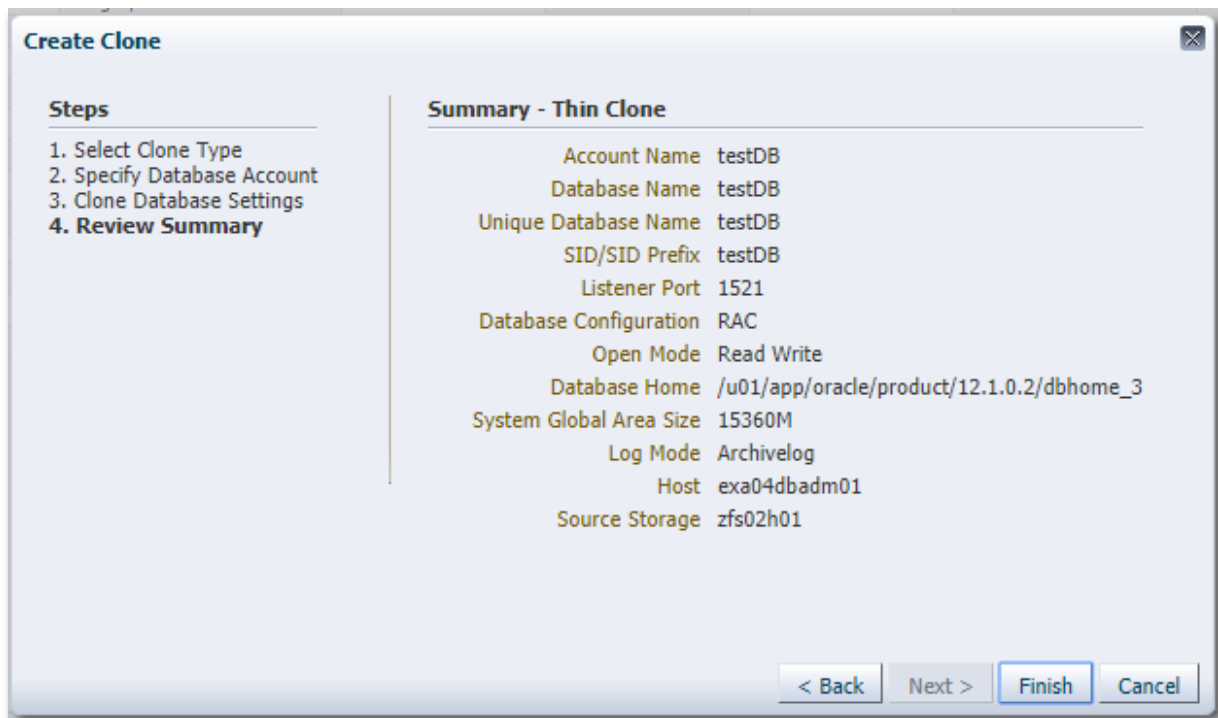
Open Mode:

Log Mode:

i **Information**
The initialization parameters for the clone database.

< Back Next > Finish Cancel

Danach können noch Einstellungen wie SGA-Größe gewählt werden und ob der Klon nach der Erstellung im Archivelog-Modus sein soll. Außerdem kann der Zustand der Datenbank mit dem „Open Mode“ festgelegt werden.



Nach der Zusammenfassung wird mit einem Klick auf den Button „Finish“ die Erstellung gestartet. Auch hier ist im Log deutlich zu sehen, wie Oracle beim Erstellen des Klons vorgeht.

Clone all file shares

Look up storage shares: OK

Storage share list

/export/ted/temp

/export/ted/data

/export/ted/archive

/export/ted/reco

/export/ted/control

Clone shares on head 0: OK

Clone share list

/export/smu-clone-1507055070917-0

/export/smu-clone-1507055070917-1

/export/smu-clone-1507055070917-2

/export/smu-clone-1507055070917-3

/export/smu-clone-1507055070917-4

Configure target database

Mount the clone share(s): OK

Configure and start up clone database

Validate clone uid: OK

Create necessary database directories: OK

Create the password file

Get database running status: OK

Get database running status: OK

Add the database to the cluster

Get database running status: OK

Database details

Database is not running.

```
Create pfile for local instance: OK
Create controlfile: OK
Validate and start up database: OK
Clean up dummy files and recreate / recompile necessary component(s): OK
Perform post-script and start up database on all nodes
  Changing clone DBID: OK
  Create shared spfile: OK
```

Listing 6: Auszug aus Log des Klonvorgangs

Oracle verwaltet fast alle Schritte, die bei der Erstellung eines funktionierenden Klons nötig sind. Nutzt man jedoch dNFS, so muss man sich selbst um die nötigen Einträge in der „oranfstab“ kümmern. Diese werden zum aktuellen Stand noch nicht durch das SMU erzeugt.