

Auch hochperformante Datenbank-Plattformen brauchen ein Backup. Dieser Artikel beschreibt die Überlegungen eines Telco-Unternehmens zum Backup-Konzept für eine Exadata Database Machine, die Problematiken und die Implementation mit RMAN und EMC Networker.

# Exadata-Backup mit EMC Networker

Konrad Häfeli, Trivadis AG

Das Projekt-Kickoff-Meeting im Dezember 2009 für die Installation und Inbetriebnahme der ersten Exadata Database Machine in der Schweiz hatte neben vielen anderen Punkten auch das Backup-Konzept auf der Agenda. Die guten Performance-Zahlen im Bereich „I/O“ sollten sich schon bald danach auch im Betrieb bestätigen. Von daher kann man davon ausgehen, dass sich dies auch auf die Backup/Recovery-Zeit auswirkt. Das optimale Backup-Konzept für Exadata wurde gesucht, und ein Vorschlag von Oracle in einem Whitepaper (siehe <http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503.pdf>) analysiert. Daraus ist folgende empfohlene Konfiguration ersichtlich:

- Backup-Media-Server involvieren
- Media-Server mit Infiniband anbinden
- Oracle Secure Backup als Media-Manager-Software

Mit diesem Konzept verspricht Oracle einen maximalen Durchsatz von rund 9 TB pro Stunde bei 14 LTO-4-Drives mit je rund 180MB/sec. Limitierend beim Backup-Durchsatz ist die Anzahl der Drives. Da das Backup nur die halbe Miete ist, sollte natürlich auch der Restore performant sein. Dabei ist die Anzahl der Disks pro Konfiguration entscheidend, sodass die 36 Disks bei einem Quarter-Rack den Restore-Durchsatz auf 6,5 TB pro Stunde limitieren. Ebenso wird empfohlen, alle Instanzen einzusetzen und einen RMAN-Channel pro Tape-Drive zu starten, in diesem Fall 14 Channels. So weit die Theorie.

Die Praxis im Projektleben: Die Backup-Verantwortung obliegt einer Tochterfirma, die eine bestehende Infrastruktur mit Software betreibt, an der die zu sichernden Systeme angehängt sind. Die Infiniband-Anbindung an einen Media-Server hätte neue Hardware bedeutet (physisch neben Exadata) und war in der Konfiguration nicht vorgesehen. Im Einsatz waren Server mit einem dedizierten 1-GB-Ethernet-Backup-Interface. Oracle Secure Backup war, obwohl eine kostengünstige und funktionelle Variante, kein Thema. Man hatte sich schon seit Jahren auf EMC Networker festgelegt. Folgende Problempunkte ergaben sich:

- Nur GB-Ethernet anstelle Infiniband-Anbindung
- EMC Networker anstelle Oracle Secure Backup
- Kein Bonding über mehrere Interfaces möglich

Das eingesetzte Exadata Quarter-Rack war eine V2 und noch nicht mit 10 GB-Ethernet-Karten ausgerüstet, sodass es auch keinen Vorteil gebracht hätte, wenn die Backup-Server wenigstens im Ethernet-Interface-Bereich aufgerüstet worden wären. Dies kommt dann erst mit dem Einsatz der Exadata X2-2 zum Tragen, doch das Unternehmen hat bereits zwei Half-Racks bestellt und wollte unbedingt vom größeren Durchsatz profitieren. Der limitierende Faktor bei der Konfiguration war klar das Netzwerk mit höchstens 120MB/sec, was etwa 400 GB pro Stunde entspricht. Dadurch war es auch nicht notwendig, ein Backup-Tuning zu machen, denn schon eine RMAN-Session liefert diesen Durchsatz.

Das Exadata-System zeichnet sich durch einen sehr hohen internen Datendurchsatz aus. Die Grundkonfiguration beinhaltet eine Flash Recovery Area (FRA), die neben „flashback logs“ auch „archived redologs“ und RMAN-Backup-Pieces aufnehmen kann. Mit einem zweistufigen Konzept können die Datenbank-Backups sehr schnell in die FRA gemacht und danach über das langsame externe Netzwerk auf Tape geschrieben werden, ohne die Datenbank zu belasten. Dies stellt eine Disaster-Recoverability sicher, die zwar keine kleinen Recovery Time Objectives (RTO) erfüllen kann, aber doch praktikabel für die Risikoabdeckung ist. Bei Daten bzw. Fileproblemen kann das Recovery sehr schnell ab dem letzten Backup von der Disk (FRA ASM-Diskgruppe) erfolgen. Folgende Backup-Komponenten sind zu berücksichtigen:

- Filesysteme (root-Filesystem, user-Filesystem, ORACLE\_BASE/ORACLE\_HOME)
- Incremental Level 0 Datenbank-Backup
- Archived redolog Backup
- Spfile und controlfile Backup
- Backup der Flash Recovery Area Backupfiles
- Housekeeping der Archived redologfiles
- Housekeeping der Backupfiles

Die Backups erfolgen grundsätzlich auf der Standby-Seite und dort über die Storage-Nodes auf die Virtual Tape Library (VTL) geschrieben. Jede Site der MAA-Konfiguration kann Standby-Systeme haben, sodass auch die Backup-Infrastruktur symmetrisch auf beiden Seiten aufgebaut wird. Für die Definition

der backupenden Knoten beziehungsweise Instanzen wird ein „Backup“-Service im Cluster erstellt, der diesen hochverfügbar zur Verfügung stellt. Die Backup-Umgebung verbindet über den „Service\_Namen“ auf die Target-Datenbank und stellt so den Ort des Backups sicher.

Tägliche Full-Backups (als Incremental level 0) werden auf den lokalen Storage (FRA) durchgeführt und danach die noch nicht auf Tape gesicherten Files der FRA auf die VTL geschrieben. In regelmässigen Intervallen (2 Stunden) werden auch die Archivefiles auf Disk gesichert und dann auf das Tape kopiert.

Das Housekeeping der in der FRA liegenden Files (Archivefiles, Backupfiles) wird über RMAN-Konfigurationen beziehungsweise FRA-Funktionalität erledigt (siehe [http://www.trivadis.com/uploads/tx\\_cabagdownloadarea/Konrad\\_Haefeli\\_Flash\\_Recovery\\_Area\\_in\\_derPraxis.pdf](http://www.trivadis.com/uploads/tx_cabagdownloadarea/Konrad_Haefeli_Flash_Recovery_Area_in_derPraxis.pdf)). Für die Sicherung des RMAN-Repositorys im Controlfile ist eine RMAN-Catalog-Datenbank in einer dritten Lokation im Einsatz.

Zu Projektbeginn war noch nicht klar, ob überhaupt eine Drittanbieter-Software wie EMC Networker-Client und das nötige Networker-Modul für Oracle auf der Exadata installiert werden darf. Dies war jedoch schnell geklärt und stellte kein Problem dar. Immer eine Frage bei der Implementierung von Backup-Konzepten ist die „Hoheit“ des Scheduling: Darf der DBA selbst die Backups auslösen, die er auf Tape macht, oder obliegt das dem Backup-Manager, der die Enterprise Backup Infrastruktur unterhält und auch noch eine Vielzahl von anderen Backup-Anforderungen berücksichtigen muss. In diesem Fall standen relativ großzügige Backup-Zeitfenster zur Verfügung, innerhalb derer die Backups eigenständig zu managen waren. Es wurde beschlossen, das Backup via Enterprise Manager Job-Scheduler zu erledigen, dies hatte den Vorteil der zentralen Verwaltung und der Alarmierung im Fehlerfall über die Grid-Control-Infrastruktur. Im Gegensatz dazu wurde für das Filesystem-Backup eine Backup-Gruppe im Networker Administrator eingerichtet, die ent-

```
CONFIGURE RETENTION POLICY TO RECOVERY WINDOW OF 21 DAYS;
CONFIGURE BACKUP OPTIMIZATION ON;
CONFIGURE DEFAULT DEVICE TYPE TO DISK;
CONFIGURE CONTROLFILE AUTOBACKUP ON;
CONFIGURE CONTROLFILE AUTOBACKUP FORMAT FOR DEVICE TYPE DISK TO
'+reco/thaa_site1/backupset/%d_%F';
CONFIGURE CONTROLFILE AUTOBACKUP FORMAT FOR DEVICE TYPE ,SBT_TAPE' TO
,%F';
CONFIGURE DEVICE TYPE ,SBT_TAPE' PARALLELISM 2 BACKUP TYPE TO BACKUP-
SET;
CONFIGURE DEVICE TYPE DISK PARALLELISM 4 BACKUP TYPE TO COMPRESSED
BACKUPSET;
CONFIGURE CHANNEL DEVICE TYPE ,SBT_TAPE' PARMS ,ENV=(NSR_
SERVER=backupsrv1,NSR_DATA_VOLUME_POOL=ExadataTape,NSR_SAVESET_
RETENTION=2Months,NSR_SAVESET_BROWSE=2Months)';
CONFIGURE ARCHIVELOG DELETION POLICY TO BACKED UP 1 TIMES TO ,SBT_
TAPE';
CONFIGURE SNAPSHOT CONTROLFILE NAME TO ,/u01/app/oracle/admin/thaa/
backup/snapcf_THAA1.f';
```

#### Listing 1

sprechend den Interval-Definitionen des Media Managers die Files via den Standard-Networker-Client sichert.

Für das Datenbank-Backup war als erstes natürlich die Tapeschnittstelle von Oracle zum Networker zu definieren und zu konfigurieren. Dazu war die Installation des Moduls für Oracle nötig. Durch das Erstellen eines Links im „ORACLE\_HOME/lib“ mit dem Namen „libobk“, der auf das „nmo“-Modul zeigt, war die Anbindung erledigt. Damit der Datenbank-Server die Berechtigung hat, auf den jeweiligen Backup-Server beziehungsweise vorgeschobenen Media-Server (Storage-Node) zu schreiben, ist noch die Aufnahme des DB-Servers auf dem Networker Administrationsserver als Backup-Client nötig. Durch die Angabe der Networker-Daten beim Konfigurieren des Tape-Channels ist die Anbindung abgeschlossen und kann getestet werden.

Damit der nicht zu unterschätzende Scriptaufwand gering gehalten werden konnte wurde das Backup Tool von Trivadis (siehe <http://www.trivadis.com/produkte/datenbank-tools/tvd-backuptm.html>) eingesetzt, das die jeweiligen Datenbank und Backup-spezifischen Scripts aus zentralen Konfigurationsfiles generiert, startet, Logfiles schreibt, diese auch auf Fehler prüft und den jeweiligen Status an den aufrufenden Scheduler zurückgibt. Listing 1 zeigt den Aufruf der mitgelieferten RMAN Script-Templates.

Es empfiehlt sich, die Konfiguration bei jedem Backup neu zu erstellen. Dies ist kein großer Aufwand, wenn einmal die Script-Umgebung steht, garantiert aber, dass nicht ein DBA-Kollege etwas anpasst, das dann im weiteren Verlauf der Backups die Funktionalität verändert.

Die Backup-Retention ist auf drei Wochen eingestellt, wird aber über den Networker-Pool (siehe die „ENV“-Parameter, die für den Device-Type „SBT\_TAPE mitgeliefert werden“) noch abgesichert, sodass keine Backups älter als zwei Monate auf den Tapes zu finden sind. Komprimierte Backups auf Disk ermöglichen bei vertretbarem CPU-Aufwand eine höhere Haltezeit der Backups auf der schnell zugreifbaren FRA und reduzieren die zu übertragende Datenmenge über das limitierte Backup-Netzwerk.

Das Löschen der Archivefiles (Housekeeping) wird automatisch von den Files gemacht, die schon einmal auf Tape gesichert wurden, aber nur dann, wenn der Disk-Platz benötigt wird. Das zeigt die View „v\$flash\_recovery\_area\_usage“. Die Spalte „PERCENT\_SPACE\_RECLAIMABLE“ steht für Platz, der zwar zum Beispiel mit Archivefiles noch belegt ist, aber bei Bedarf überschrieben werden kann (siehe Listing 2).

Für das Setzen der Flash Recovery Area werden folgende DB Parameter benötigt:

```
alter system set db_recovery_file_dest_size=3000G
scope=spfile;
alter system set db_recovery_file_dest='+RECO' scope=spfile;
```

Folgender Parameter konfiguriert die Archivelog-Destination auf die FRA:

```
alter system set log_archive_dest_1='USE_DB_RECOVERY_FILE_DEST';
```

Diese RMAN Kommando-Sequenzen werden in den Scripts aufgerufen:

- Full Backup
- backup incremental level 0 device type disk database filesper set 10
- Archlog Backup
- backup device type disk archivelog all not backed up 1 times filesper set 20
- Flash recovery area Backup
- Backup device type 'SBT\_TAPE' recovery area

Der Einsatz der Exadata als „preconfigured“ und „balanced“ System kommt vielfach schon beim Backup mit den Rahmenbedingungen der bestehenden Datencenter-Infrastruktur in Konflikt und kann demzufolge nicht immer mit der empfohlenen Variante implementiert werden. Vor allem bei einem Quarter-Rack, wo sich die Datenmenge noch in überschaubaren Größen hält, lässt sich auch ein Kompromiss mittels einer bestehenden Backup-Lösung – wie hier mit EMC Networker – umsetzen.

```
SQL> select * from v$flash_recovery_area_usage
```

FILE_TYPE	PERCENT_SPACE_USED	PERCENT_SPACE_RECLAIMABLE	NUMBER_OF_FILES
CONTROL FILE	0	0	0
REDO LOG	.31	0	13
ARCHIVED LOG	20.55	12.1	1416
BACKUP PIECE	44.2	24.02	125
IMAGE COPY	0	0	0
FLASHBACK LOG	9.08	3.23	331

Listing 2

Konrad Häfeli  
Trivadis AG

konrad.haefeli@trivadis.com



**ORACLE** PARTNER

## Technology for success

Wir unterstützen Ihren Erfolg mit der Konzeption und dem Aufbau Ihrer Datenbankanwendungen, sowie deren technischem Support.

Wir von der Krug & Partner GmbH bestehen aus einem hoch motivierten Team von Oraclespezialisten. Regelmäßige Schulungen halten unsere Fachkräfte immer auf dem aktuellsten Stand der Technik.

- **Datenbanken**  
Datenbank-Installation & -Konfiguration  
Administration & Wartung  
Backup/Recovery  
Health-Check
- **Projekte**  
Gewerblicher Rechtsschutz (IP)  
Automotive Medien- & Energiewirtschaft
- **Application Server Systeme**  
JDeveloper Forms & Reports  
PL/SQL (CMSDK, OEM, Disco...) u.v.m.
- **Lizenzierung**  
Beratung & Analyse des Lizenzstatus  
Lizenzierung

Krug & Partner GmbH ◦ Treitschkestr. 3 ◦ D-69115 Heidelberg  
Telefon: +49 (0) 62 21/60 79 0 ◦ Telefax: +49 (0) 62 21/60 79 60

E-Mail: [info@krug-und-partner.de](mailto:info@krug-und-partner.de)  
[www.krug-und-partner.de](http://www.krug-und-partner.de)



for-success