

Sicherung von Oracle Datenbanken mit der ZFS Backup Appliance

Attila Mester
Oracle Deutschland
Potsdam

Schlüsselworte

Datenbanksicherung, Exadata, RMAN, ZFS Storage Array, ZFS Backup Appliance

Einleitung

Die auf dem ZFS Storage Array aufbauende Oracle ZFS Backup Appliance (ZBA) wird zunehmend für die Datenbanksicherung der Engineered Systems wie Exadata verwendet, kann aber auch bei herkömmlichen Datenbankarchitekturen Verwendung finden. Basierend auf realisierten Kundenprojekten und Proof of Concepts werden solche Architekturen und Konfigurationen gezeigt mit dem die sehr hohen Durchsatzwerte unter einer optimalen Storage Auslastung erreicht werden können. Richtig gewählte RMAN Strategien können die vorhandene physikalische Kapazität der Storage Systeme optimal nutzen, somit kann man auf die Verwendung von externen, teuren deduplizierenden Systeme in den meisten Fällen verzichten. Die in der ZBA vorhandenen Storage Services wie Snapshot, Cloning und Replikation erlauben die erweiterte Nutzung der ZBAs über reine Datensicherungsaufgaben hinweg und ermöglichen den einfachen Aufbau von Test-, Entwicklungs- und Migrationsdatenbanken. Im Vortrag werden auch mögliche Integrationsbeispiele der ZFS Backup Appliance in bestehende Datensicherungsarchitekturen mit NetBackup, NetWorker und TSM gezeigt.

Überblick der ZFS Storage Arrays

Die ZFSSAs sind Teil des Oracle Storage Portfolios und sind unter den NAS Storages zu finden. Wir befinden uns gerade beim Übergang von der 3. zur 4. Generation dieser Geräte. Die Modelle 7120, 7320 und 7420 werden durch die neuen ZS3-2 und ZS3-4 ersetzt. Sie sind ca. 2-mal schneller und noch besser ausbaubar da sie über mehr PCIe Slots verfügen als die vorgänger Version und auch mehr Disk Shelves aufnehmen können. Die neuen Disk Shelves tragen die Bezeichnung DE2-24C oder DE2-24P und können Write Flash SSDs, schnelle 900 GB Platten oder hochkapazitive 4 TB Platten aufnehmen. Diese neuen Shelves können nicht mit den Shelves der älteren ZFSSA Systeme gemischt werden.

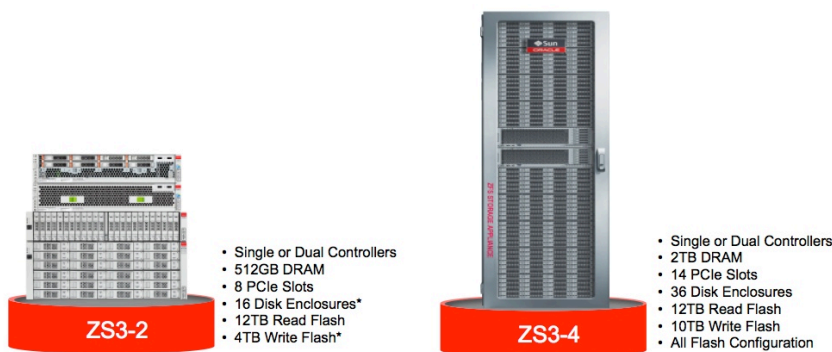


Abb. 1: Die neue ZS-3 Generation

Die ZFSSAs sind sehr zuverlässige Storage Systeme. Die Integrität der gespeicherten Daten stellt die selbst-heilende ZFS File System Architektur und die komplette Validierung des gesamten I/O Pfades sicher. Die Redundanz der Daten kann als zweifach oder dreifach Spiegel erhöht werden, und für RAID ähnliche Strukturen stehen RAIDZ-1 bis 3 mit einfacher, doppelter oder dreifacher Parität. Auf System Ebene sorgt das batterie lose Design für erhöhte Datensicherheit und die redundant vorhandenen aktiv-aktiv Controller schützen auch vor dem kompletten Ausfall eines NAS Kopfes. Die Storage Services Snapshot und Lokal Replikation sind kostenlos nutzbar und stehen innerhalb eines Systems zur Verfügung, während die Remote Replikation kostenpflichtig ist und erlaubt das asynchrone Kopieren der Daten über eine IP Strecke zwischen mehreren ZFSSAs. Die Remote Replikation ist für unstrukturierte Daten empfohlen, für Oracle Datenbank Daten sollte man Data Guard einsetzen weil die Log-Shipping Architektur dafür die meisten Vorteile bietet.

Die ZFSSAs werden vorwiegend in den folgenden Gebieten am erfolgreichsten eingesetzt:

- Allgemeiner NAS Storage als Daten Filer über die gängigen Protokolle
- als VMware Storage
- Primäre Lokation von Oracle Datenbanken
- Target Device für Oracle Datenbank Backups

In diesem Vortrag konzentrieren wir uns auf das Einsatzgebiet Datensicherung für Oracle Datenbanken. Um die Einsetzbarkeit für diesen Zweck einfach zu gestalten, wurden von Oracle die ZFS Backup Appliance genannten Geräte vorgestellt.

ZFS Backup Appliance

Die ZBA ist im Grunde ein ZFSSA welches mit den notwendigen Schnittstellen bereits ausgestattet ist und im Werk fertig verkabelt in einem Rack montiert wird. Ihre Konfiguration wurde von Oracle Engineering für RMAN Backups getestet und validiert. Somit hat der Kunde eine hohe Planungssicherheit und die Inbetriebnahme ist entsprechend kurz.

Die momentanen Backup Appliances kommen entweder mit den High-Performance oder mit den High-Capacity Disk Shelves und basieren auf dem ZFSSA 7420. Sie verfügen immer über redundante Controller, InfiniBand und 10GbE Anschlüsse. Sie verfügen in der Basisausführung über 4 HP oder 2 HC Disk Shelves. Sie sind natürlich mit weiteren Shelves aufrüstbar. Der Datendurchsatz erreicht beim Backup bis zu 20 TB/h und beim Restore 11 TB/h.

Die neuen ZBAs welche in Kürze angekündigt werden basieren auf dem neuen ZS3-4 System, haben ebenfalls zwei Controller und 4 High-Capacity Shelves, ihr Durchsatz wurde auf 25 TB/h Backup, 15 TB/h Restore erhöht. Später werden auch ZS3-2 basierende ZBAs bestellbar sein welche preislich günstiger liegen werden, jedoch weniger Ausbaubarkeit bieten.

ZFSSA für Oracle DB Backup

Die Sicherung der Oracle Datenbank auf ein angeschlossenes ZFSSA ist eine getestete und von der Oracle "Maximum Availability Architecture" Gruppe empfohlene Methode, insbesondere für die physikalische Sicherung der Exadata Maschinen. Dabei werden die Storage Volumes des ZFSSAs über das Infiniband Netzwerk der Exadata Maschine gemountet. Das verwendete Protokoll dNFS und die serielle Hochgeschwindigkeitsübertragungstechnologie Infiniband erlaubt sehr hohe Datendurchsätze bei ausreichender Parallelisierung der RMAN Sicherungsströme. Die typische HW Konfiguration für einen Oracle Datenbank Backup ist jedoch eine kombinierte Disk Backup und Tape Backup Architektur. Die Konfiguration hier ist in Verbindung mit einer Exadata Maschine erläutert. Die Sicherung der Exadata Maschinen soll nach den Oracle empfohlenen Methoden mit dem

Recovery Manager (RMAN) primär direkt auf Disk erfolgen. Für die Datensicherung mit langer Vorhaltezeit ist eine Tape Sicherung immer noch die erste Wahl und sollte als Ergänzung zu einer RMAN Sicherung auf Disk mit in Betracht gezogen werden.

Ein Weiterkopieren der Sicherungsdaten auf Tape mittels einer Datensicherungssoftware ist für längere Verfügbarkeit der Sicherungsdaten und zweite Kopien mittels verschiedenen Architekturen möglich und oft gewünscht.

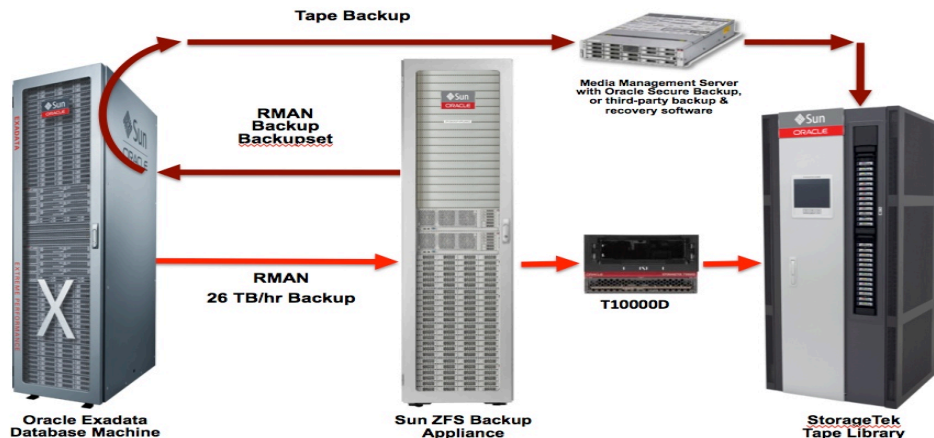


Abb. 2: Datenbankbackup auf ZFSSA mit Kopien auf Tape

Um die Oracle Sicherungsdaten von der ZFSSA oder auch direkt aus der Exadata Maschine auf Bänder schreiben zu können, muss eine Datensicherungsumgebung wie NetWorker, NetBackup, TSM, Oracle Secure Backup, etc. vorhanden sein. Die Datensicherungssoftware stellt das Media Management der Tapelibrary und die RMAN Schnittstelle für die Oracle Datenbank zur Verfügung. Für das Migrieren der Sicherungsdaten vom ZFSSA auf Tape bieten sich grundsätzlich zwei Möglichkeiten, die Methode mit RMAN oder durch eine NDMP Steuerung durch die Backup Software.

Bei der RMAN Methode werden die geschriebenen Backups wieder durch RMAN (backup backupset, oder backup image) auf Tape kopiert. Diese Methode benötigt geringfügige CPU Ressourcen und I/O auf der Datenbank Seite, aber sie ist die von Oracle empfohlene Methode weil dadurch alle Sicherungskopien durch RMAN katalogisiert werden und somit ohne grössere Vorbereitung direkt für Restores verwendet werden können. RMAN sendet die Sicherungsdaten über das Netzwerk (Infiniband oder 10GbE) zum Mediaserver der Backupsoftware welcher die Daten auf die angeschlossenen Tapelibraries schreibt. Eine direkte Fibre Channel Verbindung der Tapelaufwerke zur Exadata Maschine ist nicht zulässig, nur zusätzliche Software kann in der Exadata installiert werden, jedoch keine HW Teile wie FC HBAs.

Bei der NDMP Methode werden die Bandlaufwerke über FC direkt an das ZFSSA angeschlossen. Die ZFSSA Volumes werden mit NDMP Steuerung der Backupsoftware vom ZFSSA auf die Tapes kopiert. Diese Methode verursacht keine Belastung auf den Datenbank Servern, hat jedoch den Nachteil, dass RMAN die so erstellten Sicherungen nicht katalogisiert und auch nicht direkt für Datenbank Recovery nutzen kann. Bei Bedarf müssen die NDMP Sicherungskopien zuerst auf das ZFSSA zurückkopiert werden, erst dann kann RMAN mit dem Recovern der Datenbank beginnen. Die Geschwindigkeit der NDMP Sicherungen ist jedoch auch hoch und liegt im Bereich von ca. 8 TB/h, vorausgesetzt es stehen genügend Tape Ressourcen zur Verfügung.

Man kann die Verfügbarkeit der Sicherungsdaten in einem entfernten Rechenzentrum auch gewährleisten wenn Replikas der gewünschten ZFSSA Bereiche auf einem zweiten ZFSBA erstellt werden. Dieser Prozess ist eine Funktion des ZFSSA und geschieht asynchron über eine beliebige IP Strecke (1GbE oder 10 GbE). Im Falle eines Restores im entfernten RZ muss die Replika der dortigen RMAN Umgebung jedoch bekannt gemacht werden.

Die hier beschriebene Sicherungsarchitektur hat folgende wesentliche Vorteile:

- Sehr hohe Backup- und Restore Geschwindigkeit direkt von Disk
- Lange Backup Vorhaltezeiten und zweite Backup Kopie durch Tapelibrary Anbindung gegeben
- Einfaches Restore mit wenig Admin Overhead von Disk und auch vom Band durch RMAN gesteuerte Backups und Backup Kopien
- Möglichkeit der direkten Verwendung von HCC komprimierten RMAN Images für Datenbankkopien, da HCC komprimierte Datenbanken nur auf Oracle Storage laufen können

ZFSSA Sizing für Performance und Kapazität

Folgende Punkte sollte man hinsichtlich Performance der ZFSSA berücksichtigen:

- Der maximal erreichbare Durchsatz liegt bei 26 TB/h Backup und 15 TB/h Restore wenn InfiniBand Verbindungen verwendet werden
- Für normale Backup und Restore Aufgaben braucht man weder Read- noch Write Flash Speicher, die 4 TB SAS Disks reichen
- Pro Shelf liegt der Durchsatz der neuen Systeme bei 1300 MB/s, volle System Performance kommt also erst bei mehreren Shelves zustande
- Die Shelves werden mit RAIDZ (3 Data, 1 Parity) konfiguriert, die nutzbare Kapazität liegt somit bei ca. 64 TiB, mit 24 x 4 TB Disks gerechnet

RMAN Strategien für Disk Backup

Die Datenbanksicherung auf Disk erfolgt native mit dem RMAN Tool, eine externe Backup Software ist dafür nicht notwendig. Die Steuerung kann mit dem RMAN CLI oder dem Oracle Enterprise Manager verwaltet werden.

Die implementierbaren Backup Strategien umfassen Image- oder Backupset Sicherungen und die empfohlene Methode des "rolling Forward Incremental" Backups, wobei nach einer initialen Vollsicherung der Datenbank nur noch tägliche inkrementelle Sicherungen erfolgen. Diese werden anschliessend mit der vorherigen Vollsicherung zusammengeführt, wodurch eine neue Vollsicherung entsteht. So kann man die Vorteile der kurzen täglichen inkrementellen Sicherungen mit der Möglichkeit einer Datenbank Recovery aus einem vollen Sicherungsimage kombinieren.

Bei Verwendung von Image Backups werden Datenbankbereiche unverändert kopiert, somit kann die eingeschaltete Komprimierung auf dem ZFSSA erheblichen Speicherplatz sparen.

Durch die Snapshot Erstellung können die täglich erstellten Vollsicherungen über längere Zeit aufbewahrt werden und somit eine Versionisierung der Backups ermöglichen. Weil hier nur die geänderten Blöcke tatsächlich Plattenplatz benötigen, ist die Effektivität der Snapshots ähnlich hoch wie bei einer deduplizierenden Datenhaltung.

Verschlüsselte oder bereits komprimierte Datenbanken eignen sich nicht für die Daten Deduplizierung. Die gewünschten Einsparungen hinsichtlich Plattenplatz für Oracle Backups sowie

die Übertragungsgeschwindigkeiten können mit speziellen Deduplizierungslösungen wie z.B. Datadomain Systemen somit in den meiste Fällen nicht realisiert werden. Für die Sicherung von Exadata Systemen kann die DD-Boost Software zwar eingesetzt werden, aber Tests zeigten, dass die durch EMC prognostizierten Plattenplatz Einsparungen nicht erreicht werden könnten. Die DD-Boost Methode ist auf der Exadata von Oracle nicht zertifiziert. Oracle empfiehlt, Datenbanken mittels Infiniband oder 10 GbE auf ZFS Systeme zu sichern. Ein weiterer Vorteil liegt darin, dass die RMAN Backup Images für Test und Development auf dem ZFSSA weiterverwendet werden können. Durch die Verwendung von Snapshots und Volume Kopien kann der Aufbau von Test Datenbanken einfach und platzsparend erfolgen, was durch die optionale Nutzung des Oracle Snap Management Utilities (SMU) weiter vereinfacht werden kann.

ZFSSA in Datensicherungsumgebungen

NetWorker, NetBackup, TSM

- Empfohlen wird die initiale native RMAN Sicherung aus der Datenbank auf die ZFSSA
- RMAN Module auf dem Database Node kopiert RMAN Backupsets in das Media Management von NetWorker, NBU oder TSM und somit auf Tape
- RMAN liefert bei der Verwendung von „Weekly Full“ und „Daily Incremental“ Backupset Sicherungen mit eingeschalteter Komprimierung bessere Ergebnisse als deduplizierende Systeme
- Zu beachten ist bei TSM der fehlender IBM Support für Oracle Enterprise Linux und Solaris x86 für den TSM Oracle Modul und den TSM Server, wichtig vor allem bei Exadata Backup

Oracle Secure Backup

- OSB ist die Tape Backup Software von Oracle für Oracle und MySQL Datenbanken, Filesystemdaten und NAS Filer mit NDMP Steuerung
- Oracle supported für alle Datenbank Versionen und die gängigen Filesysteme
- Sehr günstige Lizenzierung, da nur die Anzahl der physikalischen Tapedrives relevant
- Zu beachten ist, dass Sicherung nur auf Tape, VTL oder Cloud Storage erfolgt

Kontaktadresse:

ORACLE Deutschland B.V. & Co. KG

Schiffbauergasse 14

D-14467 Potsdam

Telefon: +49 (0) 172-8125947

E-Mail attila.m.mester@oracle.com

Internet: www.oracle.com