

ZDLRA: Die neue Backup-Appliance reduziert Datenverlust und Belastung

Attila Mester, Frank Schneede und Sebastian Solbach, ORACLE Deutschland B.V. & Co. KG



Die „Zero Data Loss Recovery Appliance“ genannte Maschine ist neues Mitglied in der Reihe der Engineered Systems und verändert die Datensicherung von Oracle-Datenbanken – nicht nur in größeren Rechenzentren. Sie wurde bereits auf der Oracle OpenWorld 2013 angekündigt und ist seit dem 1. Oktober 2014 generell verfügbar. Dieser Artikel stellt die Funktionsweise und die wichtigsten Eigenschaften dieser Appliance vor.

Eines der wichtigsten Ziele bei der Entwicklung der Maschine war, dass die Wiederherstellung der gesicherten Datenbanken auch beim Ausfall des kompletten Produktsystems ohne Datenverlust möglich ist. Dies lässt sich nur durch ein Continuous-Data-Protection-System (CDP) erreichen, das alle Änderungen der Datenbank in Real Time sichert.

Heutige Backup-Lösungen haben das Problem, die Datenbank während des Backups sehr stark zu belasten. Dies ist mithilfe der neuen Backup-Strategie der ZDLRA erheblich verbessert. Durch die neue Architektur kann diese mit einem System Hunderte von Oracle-Datenbanken im Rechenzentrum effizient sichern.

Traditionelle Sicherungssysteme

Die meisten Kunden nutzen für die Sicherung ihrer Oracle-Umgebungen den von Oracle bereitgestellten und empfohlenen Recovery Manager (RMAN), der die Datenbanken auf Disk, Band oder auf einer Kombination der beiden sichert. Eine gängige Methode ist die periodische (beispielsweise wöchentliche) Vollsicherung mit nachfolgenden inkrementellen Sicherungen. Das war bis jetzt ein guter Kompromiss zwischen der Belastung der Da-

tenbank während des Backups und der Geschwindigkeit beim Restore, da nur eine überschaubare Anzahl von Sicherungsdaten zurückgespielt und verarbeitet werden mussten. Bei herkömmlichen inkrementellen Backups dauert jedoch ein Restore deutlich länger als das Restore einer einzigen Vollsicherung, da nach dem eigentlichen Restore die inkrementellen Backups in die zurückgesicherten Daten integriert werden müssen.

Eine andere moderne Methode ist das Verfahren der „incrementally updated backup strategy“. Bei dieser werden zuerst ein Image-Backup und danach nur noch inkrementelle Backups erstellt. Das Image wird dann mit jedem inkrementellen Backup gesichert, womit dann ein neues Full Image für ein schnelles Restore zur Verfügung steht. Nachteil ist, dass die Datenbank mit dem Prozess des ständig wiederkehrenden Recovery stark belastet ist und eine längere Vorhaltung von Backup-Daten, das sogenannte „Recovery Point Objective“(RPO), nur umständlich realisierbar ist.

Darüber hinaus gibt es Systeme, die mit Deduplizierung der Daten auf Storage-Ebene Plattenplatz sparen sollen. Dies funktioniert jedoch für Backup-Daten

der Oracle-Datenbank nicht optimal und erzwingt häufig auch die Verwendung eines Full Backups, um von der Deduplikationsfunktion optimal zu profitieren. Selbst wenn die Daten mithilfe von RMAN auf solchen Systemen gespeichert sind, ist eine Belastung des primären Systems erheblich. Hinzu kommt, dass auch die Validierung der Backup-Daten auf Wiederherstellbarkeit auf Storage-Ebene fehlt und damit auch ein Mechanismus, um korrupten Blöcken im Backup vorzubeugen, was bei Block-basierter Deduplikation nicht nur das eigene Backup betreffen kann, sondern auch Backups anderer Datenbanken.

Schließlich können alle Sicherungssysteme die Datenbanken nur soweit wiederherstellen, wie die mit der letzten Sicherung gesicherten, archivierten Redo-Logs es ermöglichen, wenn die Online-Redo-Logs der Primärsysteme nicht mehr vorhanden sind. Mit Data Guard wird diese verlustfreie Absicherung zwar erreicht, aber für eine große Anzahl von Datenbanken ist dies eine recht teure Lösung, die auch den Administrationsaufwand erhöht. Grundsätzlich ist eine Disaster Site auch kein Ersatz für ein echtes Datensicherungssystem.

Funktionsweise der ZDLRA im Rechenzentrum

Auch die Recovery Appliance nutzt RMAN, um die Daten der zu sichernden Datenbanken auf die ZDLRA zu übertragen. Das erfolgt über ein ZDLRA-spezifisches SBT-Interface im RMAN. Weitere Änderungen sind auf der Client-Seite nicht notwendig. Die entsprechenden SBT Libraries gibt es dabei schon bereits seit der Version 10.2.0.5 für eine Datenbank, aber selbstverständlich auch für alle neueren Versionen.

Es sei erwähnt, dass die Recovery Appliance auch auf bereits vorhandene Platten-basierte Backups der Datenbank zurückgreifen kann. Dieser „Poll“-Mechanismus transferiert die Backup-Daten, wie diese schon in der Fast Recovery Area (FRA) enthalten sind, zur Recovery Appliance. Er lässt sich auch in regelmäßigen Abständen konfigurieren.

Die Methode, mit der Backups von der Datenbank auf die Recovery Appliance erstellt werden, heißt „Delta Push“. Dabei werden initial ein Level-0-Full-Backup erzeugt und danach nur noch differenziell-inkrementelle Backups. Diese „Incremental-Forever“-Strategie beeinträchtigt die Leistungsfähigkeit der Datenbank während des Backups nur minimal. Damit alle Änderungen der Datenbank zeitnah von der Recovery Appliance gespeichert werden können, ist das Delta-Push-Verfahren optional für den Real-Time-Redo-Transport konfiguriert. Der zurzeit asynchrone Redo-Transport kommt direkt aus dem Redo-Log-Buffer der Datenbank und erlaubt somit im Falle eines totalen Ausfalls des primären Systems, das System bis zur beinahe letzten Transaktion zu recovern. Daher kommt auch der Name dieser Appliance: „Zero Data Loss“.

Diese Funktion ist auch von Data-Guard-Umgebungen bekannt. Genau wie dort lässt sich über eine „Far Sync Instance“ ein echter Zero-Data-Loss-Modus realisieren. Im Gegensatz zu Data Guard unterstützt die Recovery Appliance aber auch die Redo-Informationen von Betriebssystem-Plattformen mit unterschiedlichen Endian-Strukturen. Die Redo-Streams werden in der Appliance archiviert, was das Primärsystem ebenfalls entlastet, da beim inkrementellen Backup nicht mehr unbedingt die archivierten Redo-Logs nochmals gesichert werden müssen.

Die empfangenen Backup-Daten sind in der Recovery Appliance im sogenannten „Delta Store“ gespeichert. Dieser ist das Herz der Recovery Appliance, in dem er für alle Operationen zuständig ist, die für die Aufbereitung und Speicherung der Backup-Daten notwendig sind. Nachdem die Appliance die Backup-Daten empfangen hat, validiert und speichert sie diese in Pools in komprimierten Blöcken. Die De-Duplizierung findet nicht in der Appliance, sondern bereits auf der Client-Seite durch RMAN statt, weil ein inkrementelles Backup die redundanten Blöcke beim Lesen bereits eliminiert. Nach jedem inkrementellen Backup wird aus den gespeicherten Blöcken in der Recovery Appliance eine virtuelle Full-Repräsentation des Backups erzeugt, was auch so in dem Recovery-Katalog, der sich ebenfalls in der Appliance befindet, eingetragen ist.

Aus Perspektive des Datenbank-Benutzers ist ein solches Backup nicht von einem nicht-virtuellen Full Backup zu unterscheiden und auch in den Recovery-Zeiten mit einem Full Backup vergleichbar. Der Restore-Prozess von einem Full Backup ist wesentlich schneller als bei einer Full/Incremental-Strategie. Man gewinnt also Geschwindigkeit beim Restore, obwohl nur der Plattenplatz für inkrementelle Backups beansprucht wird. Damit man sicher sein kann, dass die virtuellen Backups immer wiederherstellbar sind, werden im Hintergrund periodisch automatische Validierungs-Jobs ausgeführt. Alle diese Prozesse werden von der

Appliance autonom mit der Rechenkraft der Appliance ausgeführt. Die produktiven Datenbanken sind somit von Datensicherungsaktivitäten weitgehend entlastet (siehe Abbildung 1).

Der Recovery-Prozess ist ebenfalls denkbar einfach. Wenn eine Datenbank oder Teile davon wiederhergestellt werden sollen, verwendet RMAN den Katalog, der auf der Recovery Appliance installiert ist. Die Zeitangabe für ein PIT-Recovery ist im Rahmen der spezifizierten Datenvorhaltung (Retention) beliebig möglich. Die Recovery Appliance liefert dann das nächstgelegene Full Backup, die notwendigen Archive-Logs und die empfangenen Redo-Streams, die hier den Online-Redo-Logs der Datenbank entsprechen, sollten diese auf dem Primärsystem nicht mehr verfügbar sein. Somit ist ein „Zero Data Loss Recovery“ der Datenbank möglich. Es sei hier erwähnt, dass nicht alle Datenbanken zwingend mit „Redo Transport“ gesichert sein müssen. In diesem Fall sollten bei der RMAN-Sicherung der Datenbank dann auch archivierte Redo-Logs mit gesichert werden, wie bei herkömmlichen Backups. Die virtuellen Full Backups stehen in diesem Fall aber auch für ein Recovery zur Verfügung.

Die Recovery Appliance kann im Rechenzentrum sinnvoll ergänzt werden. Wenn weitere Backup-Kopien erwünscht sind, kann eine entfernte Appliance alle gesicherten Daten aus der ersten Appliance durch Replikation zeitnah empfangen. Die zwei Backup-Kopien werden sowohl in der ersten als auch in der zweiten

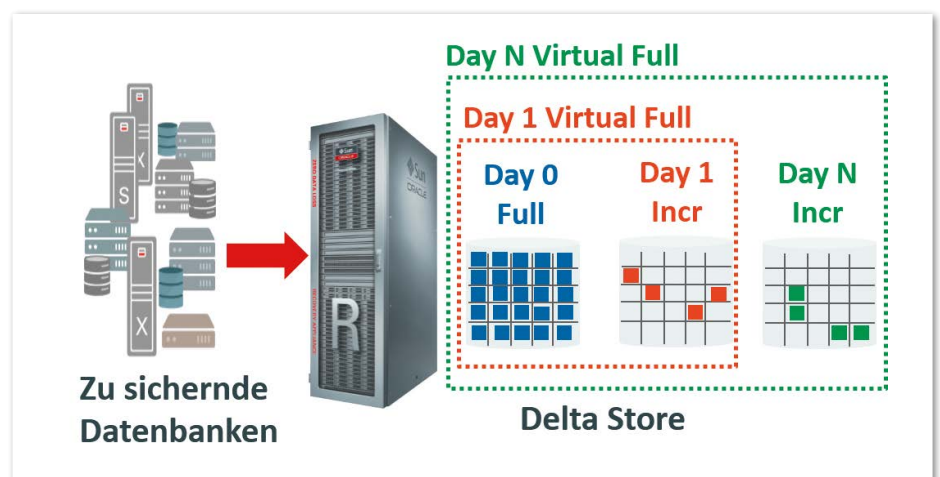


Abbildung 1: Delta Store mit virtuellen Full Backups

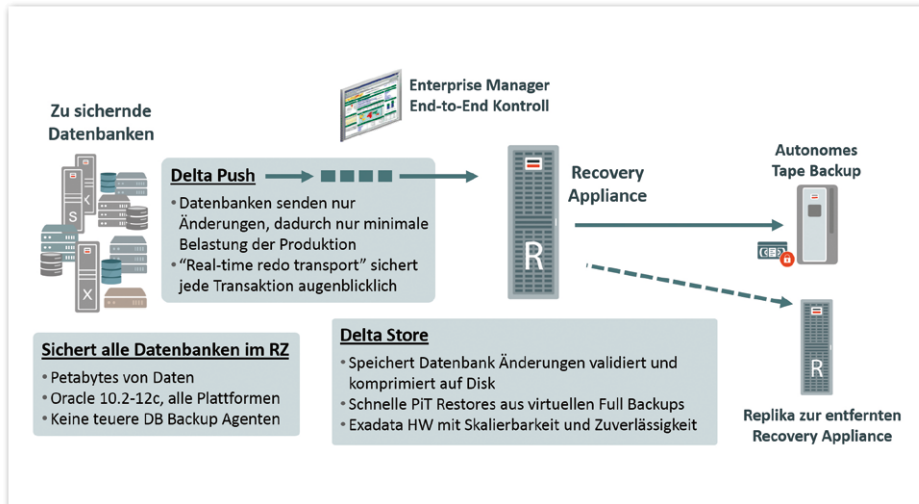


Abbildung 2: ZDLRA im Rechenzentrum

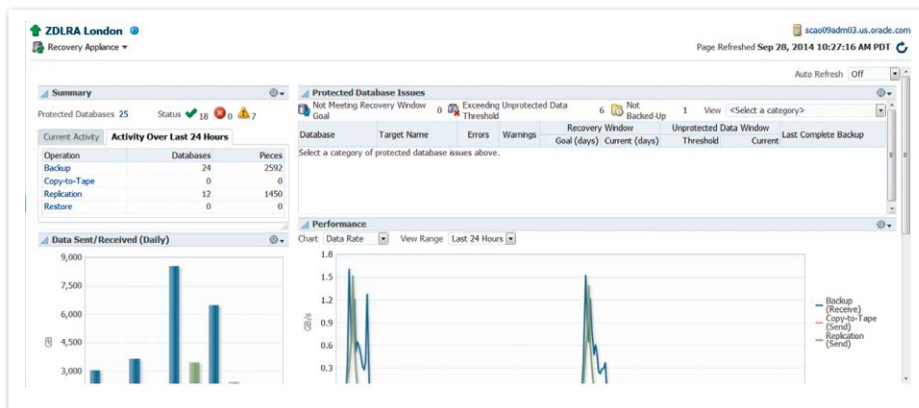


Abbildung 3: Grafische Benutzeroberfläche der ZDLRA im OEM CC12c

Recovery Appliance katalogisiert und stehen somit transparent für RMAN-Restores zur Verfügung (siehe Abbildung 2).

Die Recovery Appliance ist standardmäßig mit dem Backup-Tool „Oracle Secure Backup“ vorkonfiguriert, um Backup-Kopien auf Bandlaufwerken zu erstellen. Die Erstellung der bis zu vier Bandkopien wird durch die Appliance autonom, ohne Belastung der zu sichernden Primärsysteme erledigt. Die Bandkopien können unterschiedliche Retention-Zeiten haben. Damit entsteht eine kostengünstige Archivierung.

Die Band-Bibliotheken sind über Fiber Channel mit der Appliance direkt verbunden. Die Möglichkeit zur Verwendung einer bestehenden Tape-Backup-Infrastruktur wie NetBackup besteht ebenfalls, der fremde Media-Server muss hier jedoch außerhalb der Appliance installiert sein. Auf der Appliance

selber wird dann lediglich der Backup Agent installiert.

Die Steuerung der gesamten Oracle-Datenbank-Backup-Umgebung geschieht über den Oracle Enterprise Manager Cloud Control 12c (OEM CC), der durch ein spezielles ZDLRA-Plug-in ergänzt ist. Der OEM CC ist die einfach zu bedienende grafische Oberfläche, die für die Steuerung und Überwachung aller Engineered Systems strategisch wichtig ist. Das Management der Backup-Umgebung über den OEM CC umfasst die Parametrisierung der RMAN-Backups der zu sichernden Datenbanken, die Konfiguration der Recovery Appliance selbst und die Verwaltung der Replikas sowie der Bandkopien, insofern diese mit OSB erstellt wurden. Es stehen auch umfangreiche Monitoring- und Reporting-Funktionen zur Verfügung, um die gesamte Infrastruktur verwalten zu können (siehe Abbildung 3).

Die Recovery Appliance ermöglicht den Aufbau einer privaten Backup-Cloud für die gesamte Oracle-Datenbank-Infrastruktur eines Rechenzentrums, um somit „Backup as a Service“ anzubieten. Um das Backup von Hunderten von Oracle-Datenbanken effizient zu verwalten, werden Protection Policies gebildet und die Datenbanken diesen zugeordnet. Zu den typischen Attributen der Protection Policies gehören die verschiedenen Recovery-Zeitfenster (RTO) für Disk beziehungsweise Tape sowie Einstellungen für die Replikationen. Die als Replika genutzten Recovery Appliances können eigene Policies enthalten, da diese in den jeweiligen Appliances als Metadaten abgespeichert sind. Rollenaufteilungen für mehrere Administratoren erlauben zudem die Bildung von Management-Inseln, bei denen die jeweiligen Administratoren nur die ihnen zugewiesenen Datenbank-Backups sehen und verwalten können (siehe Abbildung 4).

Konfigurationen und Skalierbarkeit, Kapazität und Performance

Die Recovery-Appliance-Hardware entspricht einer Exadata mit High Capacity Disks. Die kleinste bestellbare Einheit besteht aus zwei Compute- und drei Storage-Servern. Diese sind über redundante InfiniBand-Switches miteinander verbunden. Die Netzwerk-Verbindungen nach außen bestehen aus acht 10-GbE-Ports pro Rack, über diese Anschlüsse werden die Backups, Recoveries und Replikationen zu anderen Appliances durchgeführt. Die Storage-Server enthalten jeweils zwölf mal vier TB Disks. Die Basis-Einheit mit drei Storage-Servern hat eine nutzbare Kapazität von 37 TB. Die Platten der Storage-Server sind – genau wie bei Exadata – mit dem Automatic Storage Management gespiegelt. Dabei werden die Backup-Daten einfach gespiegelt abgespeichert, die Meta- und Katalog-Daten sogar mit dreifacher Spiegelung. Vom Basis Rack auf ein Full Rack kommt man durch schrittweisen, bedarfsgesteuerten Ausbau durch einzelne Storage-Server (siehe Abbildung 5).

Ein Recovery-Appliance-Full-Rack kann bis zu vierzehn Storage-Server aufnehmen, das entspricht einer nutzbaren Kapazität von 224 TB. Die ersten drei

Storage-Server verfügen über weniger nutzbare Kapazität für Backup-Daten als die Erweiterungen, da dort die zusätzlichen Metadaten und der RMAN-Recovery-Katalog untergebracht sind. Die erreichbare tatsächliche Backup- oder Recovery-Performance (sogenannte „Ingest“) liegt bei etwa zwölf TB pro Stunde für ein voll ausgebautes Rack. Wenn höhere Kapazitäten und Performance erforderlich sind, können insgesamt 18 Racks miteinander zu einer Einheit zusammengeschaltet werden. Dabei erhöhen sich Kapazität und Durchsatz linear.

Als Sizing-Richtlinie gilt, dass die Recovery Appliance für zwei Wochen Backup-Datenvorhaltung etwa ein- bis zweimal die Kapazität der zu sichernden Datenbanken haben sollte. Natürlich spielen für die genauere Berechnung die tägliche Änderungsrate, das Redo-Log-Aufkommen und die Komprimierbarkeit der Datenbank ebenfalls eine Rolle.

Der Preis einer Recovery Appliance setzt sich aus dem Hardware- und dem Software-Preis zusammen. Die Hardware-Kosten entsprechen etwa denen der Exadata-Maschinen. Obwohl die Funktionen der Recovery Appliance durch die installierten Oracle-RAC-Datenbanken realisiert sind, fallen keine zusätzlichen Datenbank-Lizenzen an; nur von der Kapazität abhängige, spezielle ZDLRA-Software-Lizenzen werden berechnet.

Fazit

Oracle hat bereits mit Data Guard eine Zero-Data-Loss-Absicherung und mit der ZFS Backup Appliance ein performantes Backup-Target für RMAN-Sicherungen. Beide Lösungen kommen aber hauptsächlich als einzelne Lösungen für die besonders wichtigen Datenbanken vor. In Umgebungen, in denen für alle Datenbanken ein Sicherheitskonzept gesucht oder für viele Datenbanken Zero Data Loss Recovery gewünscht wird und bis jetzt Komplexität und Kosten von Data-Guard-Architekturen die Implementierung verhindert haben, ist ZDLRA eine Alternative. Backup-Konsolidierung oder Insourcing von Oracle-Datenbanken in großen Rechenzentren ist sicherlich auch ein passendes Anwendungsfeld, dort spielen die Skalierbarkeit und die einfache Administration mit Mandantenfähigkeit eine wichtige Rolle.

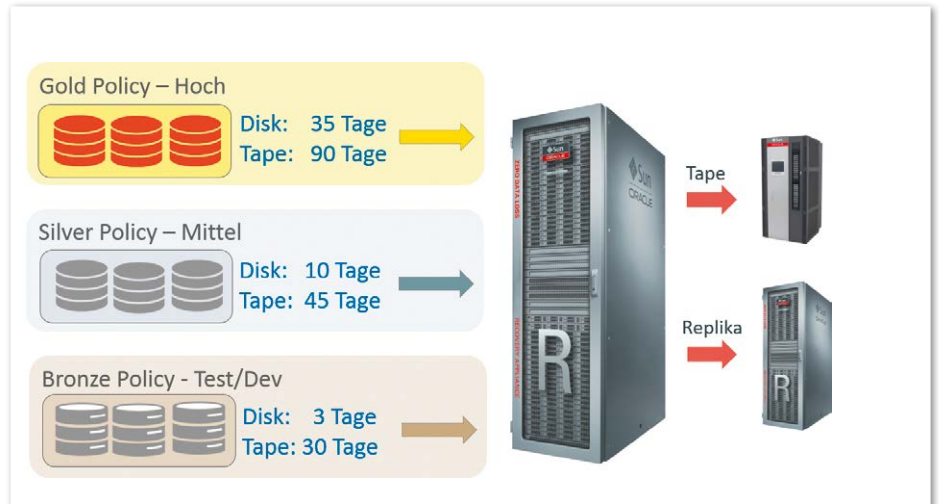


Abbildung 4: Beispiel für Protection Policies



Abbildung 5: Base Rack und Kapazitätserweiterung



Attila Mester
attila.m.mester@oracle.com



Frank Schneede
frank.schneede@oracle.com



Sebastian Solbach
sebastian.solbach@oracle.com