

<http://www.doag.org/home/aktuelle-news/article/oracle-zero-data-loss-recovery-appliance-ein-neues-engineered-system.html>

# Oracle Zero Data Loss Recovery Appliance – ein neues Engineered System

Sebastian Solbach, ORACLE Deutschland B.V. & Co. KG

**Das neueste Mitglied in der Reihe der Engineered Systems, die Zero Data Loss Recovery Appliance (ZDLRA), ist seit dem 1. Oktober 2014 auf dem Markt. Der Vorteil der neuen Backup-Maschine: Die Datensicherung erfolgt in Real-Time als „Virtual Full Backup“. Dadurch sollen die Backups platzsparend und die Restores deutlich schneller erfolgen. DOAG Online geht näher auf Funktionsweise und Steuerung des Systems ein.**

Eines der wichtigsten Ziele bei der Entwicklung der Maschine war die



Wiederherstellung der gesicherten Datenbanken ohne Datenverlust auch bei einem Ausfall des kompletten Produktivsystems. Dies lässt sich durch eine Real-Time-Sicherung aller Änderungen der Datenbank erreichen. Aktuelle Backuplösungen belasten jedoch häufig während des Backups sehr stark die Datenbank. Mit Hilfe der neuen Backupkonzeption der ZDLRA werden genau diese Themen adressiert, sodass ein ZDLRA-System durch die neue Architektur hunderte von Oracle-Datenbanken im Rechenzentrum effizient sichern kann.

## **Traditionelle Sicherungssysteme und deren Nachteile**

Die meisten Kunden nutzen heutzutage für die Sicherung ihrer Oracle-Umgebungen den von Oracle bereitgestellten und empfohlenen „Recovery Manager“ (RMAN), mit dem Datenbanken auf geeigneten Medien gesichert werden. Eine übliche Methode ist die periodische (z. B. wöchentliche) Vollsicherung in Zeiten geringer Belastung des Produktionssystems mit nachfolgenden inkrementellen Sicherungen. Dadurch wird die Belastung der Datenbank möglichst gering gehalten und die Wiederherstellungszeit minimiert, da nur eine überschaubare Anzahl von Sicherungsdaten zurückgespielt und verarbeitet werden muss. Eine Wiederherstellung aus einem inkrementellen Backup dauert

allerdings deutlich länger als das Restore einer einzigen Vollsicherung, da nach dem eigentlichen Restore die inkrementellen Backups nachgefahren werden müssen. Eine andere, moderne Methode ist das Verfahren der „incrementally updated backup strategy“. In diesem Verfahren werden nach einem initialen „Image Backup“ nur noch inkrementelle Backups gemacht. Das Image wird nach jedem inkrementellen Backup recovered, wodurch ein neues „Full Image“ für ein schnelles Restore zur Verfügung steht. Der Nachteil dieser Methode ist, dass die Datenbank mit dem Prozess des ständig wiederkehrenden Recovery stark belastet wird, und dass eine längere Vorhaltung von Backup-Daten, das sogenannte „Recovery Point Objective“ (RPO), nur umständlich realisierbar ist. Andere Systeme sollen mit Deduplikation der Daten auf Storage-Ebene Plattenplatz sparen. Eine optimale Deduplikation ist jedoch häufig nur mittels Full Backup zu erreichen. Selbst wenn die Daten mit Hilfe von RMAN auf solche Systeme gespeichert werden, ist eine Belastung des primären Systems erheblich. Hinzu kommt, dass die Validierung der Backup-Daten auf Wiederherstellbarkeit auf Storage-Ebene fehlt und damit auch ein Mechanismus, korrupten Blöcken im Backup vorzubeugen. Das kann bei blockbasierender Deduplikation nicht nur das eigene Backup betreffen, sondern auch Backups anderer Datenbanken. Schließlich können alle Sicherungssysteme die Datenbanken nur soweit wiederherstellen, wie es die mit der letzten Sicherung archivierten Redo-Logs ermöglichen, wenn die Online-Redo-Logs der Primärsysteme nicht mehr vorhanden sind. Mit „Data Guard“ wird die verlustfreie Absicherung zwar erreicht, für eine große Anzahl von Datenbanken ist dies aber eine recht teure Lösung und erhöht zudem den Administrationsaufwand. Grundsätzlich ist eine „Disaster Site“ kein Ersatz für ein echtes Datensicherungssystem. Die hier vorgestellte ZDLRA räumt mit diesen Nachteilen auf, da sie gerade diese Defizite adressiert.

### **Funktionsweise der ZDLRA im Rechenzentrum**

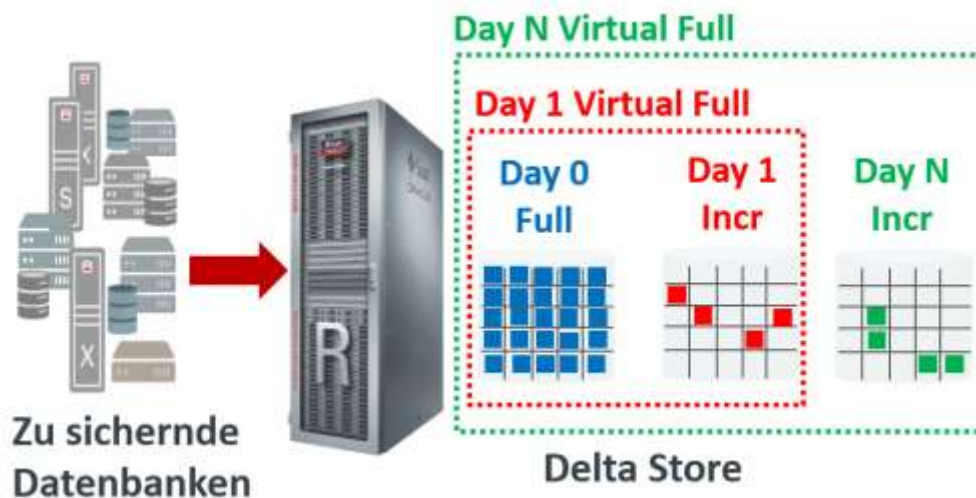
Die „Recovery Appliance“ nutzt ebenfalls RMAN, um die Daten der zu sichernden Datenbanken auf die ZDLRA zu übertragen. Das geschieht über ein ZDLRA spezifisches SBT Interface im RMAN, es werden Datenbanken ab Version 10.2.0.5 unterstützt.

Die Recovery Appliance kann auch über einen Poll-Mechanismus auf bereits vorhandene plattenbasierte Backups der Datenbank zurückgreifen, die z. B. in der „Fast Recovery Area“ (FRA) enthalten sind, und diese in konfigurierbaren Abständen regelmäßig zur Recovery Appliance übertragen.

Die Methode, wie Backups von der Datenbank auf der Recovery Appliance erstellt werden, heißt „Delta Push“. Dabei wird initial ein „Level 0 Full Backup“ erzeugt, danach nur noch differenziell-inkrementelle Backups, wodurch die Leistungsfähigkeit der Datenbank während des Backups nur minimal beeinträchtigt wird. Damit alle Änderungen der Datenbank zeitnah von der Recovery Appliance gespeichert werden können, kann das Delta Push-Verfahren für den aus „Data Guard“-Umgebungen bekannten „Real-Time Redo-Transport“ konfiguriert werden. Der zurzeit asynchrone „Redo-Transport“ kommt direkt aus dem „Redo-Log-Buffer“ der Datenbank und erlaubt 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. Wie in „12c Data Guard“-Umgebungen lässt sich ein echter „Zero Data Loss“-Modus realisieren, indem man eine „Far Sync Instance“ einsetzt. Im Gegensatz zu Data Guard unterstützt die Recovery Appliance auch Redo-Informationen von Betriebssystemen mit unterschiedlicher Endianness. Die Redo-Streams werden in der Appliance archiviert, was das Primärsystem ebenfalls entlastet, da beim inkrementellen Backup die archivierten Redo-Logs nicht mehr unbedingt nochmals gesichert werden müssen.

Die empfangenen Backup-Daten werden in der Recovery Appliance im sogenannten „Delta Store“ gespeichert. Der Delta Store ist das Herz der Recovery Appliance und ist für alle Operationen zuständig, 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 komprimierten Blöcken. Die Deduplikation der Backup-Daten findet nicht in der Appliance, sondern bereits durch RMAN statt, weil ein inkrementelles Backup redundante Blöcke schon beim Lesen eliminiert. Nach jedem inkrementellen Backup wird aus den gespeicherten Blöcken in der Recovery Appliance ein sogenanntes „Virtual Full Backup“ erzeugt, welches in den Recovery-Katalog der Appliance eingetragen wird. Aus der Perspektive des Datenbankbenutzers ist ein solches Backup nicht von einem nicht-virtuellen Full Backup unterscheidbar und auch in den Recovery-Zeiten mit einem Full Backup vergleichbar. Man gewinnt 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 Validierungsjobs ausgeführt. All diese Prozesse werden von der Appliance autonom ausgeführt und belasten die produktiven Datenbanken nicht.

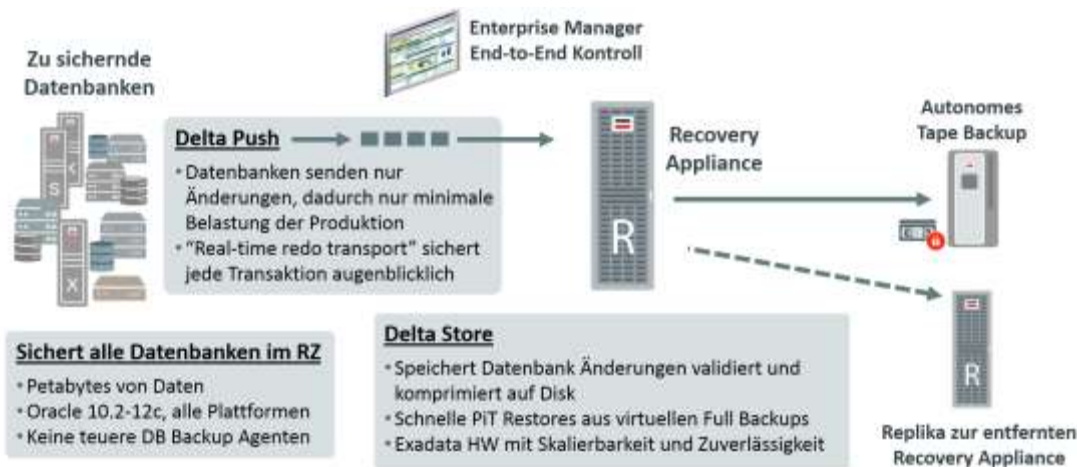


Der Recovery-Prozess ist ebenfalls einfach. Wenn eine Datenbank oder Teile davon recovert werden sollen, verwendet RMAN den Katalog, der auf der Recovery Appliance installiert ist. Die Zeitangabe für eine „Point In Time“ (PIT)-Recovery ist im Rahmen der spezifizierten Datenvorhaltung (Retention) beliebig 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 eine „Zero Data Loss Recovery“ der Datenbank möglich. Nicht alle Datenbanken müssen zwingend mit Redo-Transport gesichert werden - in diesem Fall sollten bei der RMAN-Sicherung der Datenbank 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.

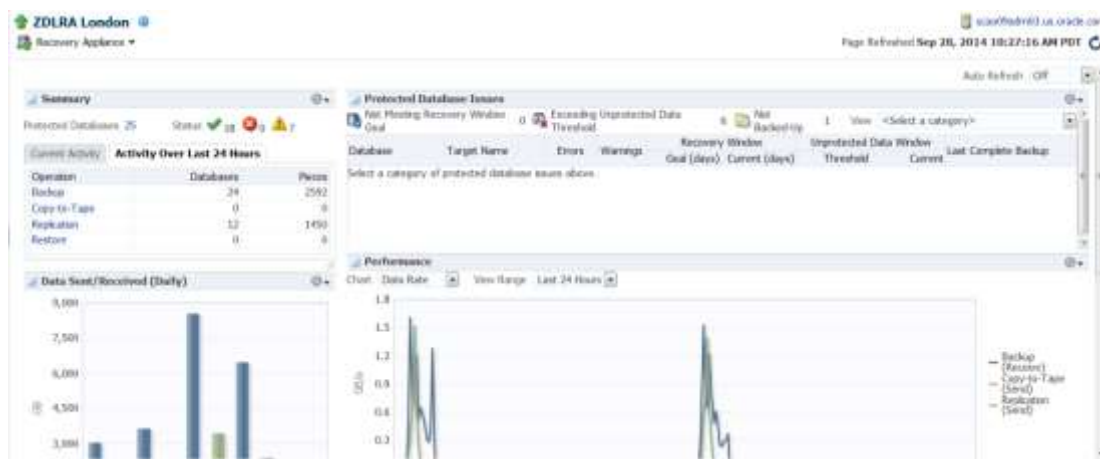
Die Recovery Appliance kann im Rechenzentrum sinnvoll ergänzt werden. Wenn weitere Backup-Kopien gewünscht werden, kann eine entfernte Appliance alle gesicherten Daten aus der ersten Appliance durch Replikation zeitnah empfangen. Die zwei Backup-Kopien werden sowohl in der ersten wie auch in der zweiten Recovery Appliance katalogisiert und stehen transparent für RMAN Restores zur Verfügung.

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 erledigt. Die Bandkopien können unterschiedliche Retentionszeiten haben, damit kann eine kostengünstige Archivierung erreicht werden. Die Bandbibliotheken werden über Fiber Channel mit der Appliance direkt verbunden. Die Möglichkeit zur Verwendung einer bestehenden Tape-Backup-Infrastruktur,

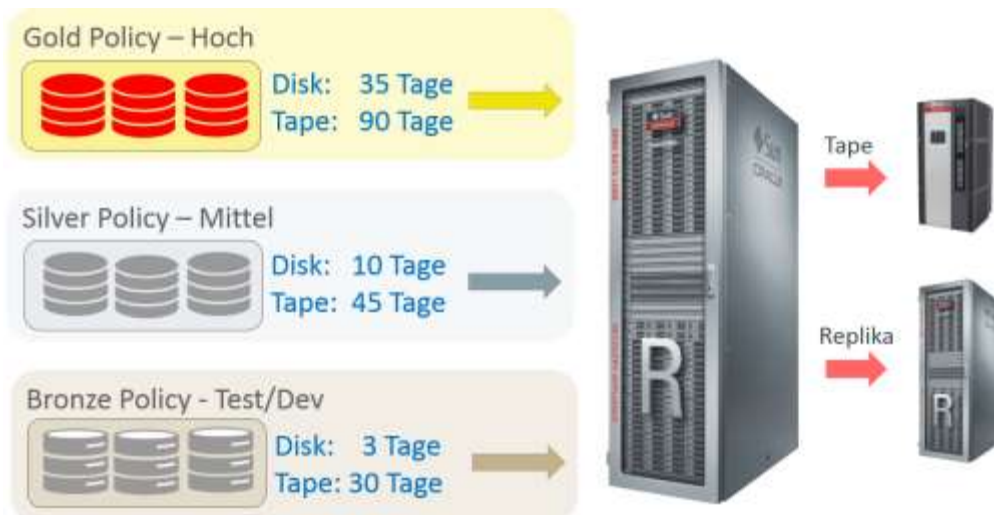
wie z. B. NetBackup, besteht ebenfalls, der fremde Media Server muss hier jedoch außerhalb der Appliance installiert werden. Auf der Appliance selbst wird dann lediglich der Backup-Agent installiert.



Die Steuerung der gesamten Oracle-Datenbank-Backup-Umgebung geschieht über den „Oracle Enterprise Manager Cloud Control 12c“ (OEM CC), der mit einem ZDLRA-Plugin ergänzt wird. Der OEM CC ist die grafische Oberfläche, die für die Steuerung und Überwachung aller Engineered Systems empfohlen wird. 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 mit OSB erstellten Bandkopien. Es stehen umfangreiche Monitoring- und Reporting-Funktionen zur Verfügung, um die gesamte Infrastruktur zu verwalten.



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 bzw. Tape und 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, wo die jeweiligen Administratoren nur die ihnen zugewiesenen Datenbank-Backups sehen und verwalten können.



### 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 Servern und drei Storage Servern, die über redundante InfiniBand Switches verbunden sind. Die Netzwerkverbindungen ins Kundennetzwerk bestehen aus acht 10 GbE Ports pro Rack, für Backup und ggf. Replikation zu anderen Appliances. Die Storage Server enthalten jeweils 12 x 4 TB Disks, die Basiseinheit mit 3 Storage Servern hat eine nutzbare Kapazität von 37 TB. Auf den Platten kommt, wie bei Exadata, ASM zum Einsatz, wobei Backup-Daten mit normaler, Metadaten und Kataloge mit hoher Redundanz abgelegt werden. Die Erweiterung des Basis Rack bis zum Full Rack erfolgt bedarfsgesteuert durch einzelne Storage Server.



Ein ZDLRA-Full-Rack kann bis zu 14 Storage Server aufnehmen, das entspricht einer nutzbaren Kapazität von 224 TB. Die erreichbare tatsächliche Backup- oder Recovery Performance liegt bei ca. 12 TB/h für ein voll ausgebautes Rack. Ein Ausbau ist bis hin zu insgesamt 18 verbundenen Racks möglich, wobei sich die Kapazität und der Durchsatz linear erhöhen.

Als Sizing Richtlinie gilt, dass die Recovery Appliance für 2 Wochen Backup-Datenvorhaltung ca. 1- bis 2-mal 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.

### Lizensierung und Preisgestaltung

Der Preis einer Recovery Appliance setzt sich aus dem der Hardware und Software zusammen. Die Hardware-Preise liegen unter denen der Exadata-Maschinen. Obwohl die Funktionen der Recovery Appliance durch installierte Oracle RAC Datenbanken realisiert sind, fallen keine zusätzlichen Datenbank-Lizenzen an. Nur von der Kapazität abhängige, spezielle ZDLRA-Software-Lizenzen werden als Software-Preis gerechnet.