

# Backupverfahren aus der Praxis - optimiert und intelligent

**Stirnemann Roland**  
**Trivadis AG**  
**CH-8152 Zürich**

## **Schlüsselworte**

Risikominimierung von Datenverlust, effiziente Ressourcennutzung, bedarfsgerechte Sicherungen, Backup Optimierung, Transparenz mit Backup Reports, Volumenreduktion

## **Einleitung**

IT-Ausfälle und Datenverlust sind für Unternehmen ein Albtraum. Häufig sind Software- oder Netzwerkprobleme, aber auch menschliche Fehler mögliche Ursachen. Die Kosten können - je nach Kritikalität und Dauer - in die Millionenhöhe gehen. Entsprechende Risiken gilt es zu minimieren durch ein verlässliches Backup & Recovery der Datenbanken. Trotz angemessener Sicherheit soll aber die Effizienz nicht verschlechtert werden, da mehr Sicherheit oft in Verbindung mit mehr Daten und Ressourcen gesehen wird und somit letztendlich Mehrkosten verursacht. Es stellt sich also die Frage: Wie können Risiken minimiert werden ohne die Effizienz zu vernachlässigen? Ein optimales Backupverfahren ist erforderlich, welches so wenig sichert wie möglich und so viel wie nötig. Wie das in der Praxis umgesetzt werden kann wird der DOAG Präsentation aufgezeigt.

## **Aktuelle Backup Situation in Unternehmen**

Die Sicherung von Datenbanken wird in heutigen Unternehmen oft sehr pauschal durchgeführt und generiert deshalb unverhältnismässige Betriebskosten. Alle sind sich einig, dass es ein Backup braucht, jedoch ist es in der Praxis häufig schwierig Argumente zu finden, um Investitionen in Verbesserungen/Erneuerungen tätigen zu können. Erst im Fehlerfall werden Entscheider und Business wachgerüttelt und sind sich der Wichtigkeit einer zuverlässigen und performanten Sicherung und Wiederherstellung bewusst. Die nachfolgende Zusammenstellung nennt aktuelle Problemzonen aus der Praxis, denen es zu begegnen gilt:

**Backup Performance:** Jede Sicherung generiert spürbare Systemlast, primär I/O Operationen, und Datenvolumen im entsprechenden Speichersystem. Das Volumen ist jedoch nur eine Seite der Medaille, denn die Last durch das Lesen der Daten ist in Bezug auf die Datenbankperformance und letztendlich für die Applikation entscheidend. In der Praxis ist die Backuplast zudem häufig schlecht verteilt.

**Backupsteuerung:** Die Steuerung und somit die Verteilung der Backup Jobs ist bei mehreren hundert Datenbanken nicht ganz einfach. Häufig fehlt der Überblick was wann startet und so kommt es zu Lastspitzen aufgrund zu vieler paralleler Backup Operationen.

**Backupverwaltung:** Generell finden viele Aufgaben direkt auf jedem einzelnen Datenbankserver statt und kosten den Administrator Zeit und Nerven. Auch die Konfiguration von unterschiedlichen Backup Anforderungen (z.B. seitens SLA) ist in grossen Umgebungen eine grosse Herausforderung, um den Überblick jederzeit zu behalten.

**Überwachung und Reporting:** In irgendeiner Form existiert meistens eine Backupüberwachung. Doch was kann diese bzw. was kann sie nicht? Häufig basiert die Überwachung auf Mailnachrichten, welche im Fehlerfall versendet werden. Übergreifende ad-hoc Auswertungen zum Backup Volumen, zu Backup Laufzeiten, zu laufenden und fehlgeschlagenen Sicherungen müssen manuell erstellt werden und sind somit zeitaufwendig. Umfassende Transparenz auf Knopfdruck ist vielerorts ein Wunschdenken.

**Wiederherstellung:** Weil ein Backup seit Jahren funktioniert ist dies keine Garantie, dass auch die Wiederherstellung einwandfrei klappt. Die Erfahrung zeigt, dass solche Vorgänge länger dauern als angenommen oder nicht so reibungslos funktionieren wie gedacht. Dies führt zu längeren Ausfallzeiten im Fehlerfall und letztendlich mehr Kosten.

### **Anwendungszwecke einer Sicherung**

Bevor die Frage geklärt wird, wie eine optimale Sicherung aussehen könnte, sollte sich ein Unternehmen im Klaren sein wozu überhaupt die Backups verwendet werden. Natürlich kommt den meisten Leuten der Disaster Restore als erster Gedanken in den Sinn. Jedoch kann ein Backup in der Praxis für weit mehr verwendet werden, wodurch je nach Anwendungsfall unterschiedliche Anforderungen bestehen. Insbesondere für den Disaster Fall glauben teilweise Administratoren, dass eine Hochverfügbarkeitslösung wie Data Guard ein Backup ersetzt. Diese Aussage ist natürlich nicht allgemeingültig. Ein weiterer Anwendungszweck für das Backup ist die Wiederherstellung der Datenbank nach logischen Korruptionen. Die Datenbank ist eigentlich noch funktionsfähig, jedoch ist der Inhalt (Daten) fehlerhaft. Das Klonen von Datenbanken kann ebenfalls auf RMAN Sicherungen basieren.

### **Möglichkeiten zur Backup Optimierung mit Oracle Bord Mitteln**

Oracle bietet von Haus aus schon seit längerem diverse Funktionalitäten um Backup und Restore Vorgänge zu optimieren. Inkrementelle Sicherungen helfen bei der Reduktion des Backupvolumens (Output) und belasten somit den Backup Storage weniger. Die gesamte Datenbank wird nicht mehr täglich gesichert, sondern nur alle paar Tage. In der Oracle Enterprise Edition steht zusätzlich Block Change Tracking zur Verfügung. Damit lässt sich auch das Inputvolumen, also die Leseoperationen, massiv verringern. Weitere Optimierungen beim Backupoutput sind mittels Komprimierung möglich, die direkt in RMAN eingebaut ist. Jedoch ist beim Aktivieren der Komprimierung darauf zu achten, dass je nach verwendetem Algorithmus die Lizenz für die Advanced Compression Option zu erwerben ist. Für die zentrale Steuerung, Konfiguration, Überwachung und Verwaltung der Backups bietet sich Oracle Enterprise Manager an. Daneben gibt es auf dem Markt diverse Tools zur zentralen Steuerung, jedoch sind diese bei der Verwaltung der Backups häufig limitiert und dem DBA eher fremd. Letztendlich ist bei zentralisierten Lösungen die Verfügbarkeit sehr wichtig, denn bei den meisten Lösungen muss der zentrale Management Server verfügbar sein, damit die Agenten die Sicherungsaufgaben ausführen. Ein längerer Stillstand des Management Servers kann zum Datenbankstillstand führen (z.B. Archiver Stuck) und stellt ein Single-Point-of-Failure dar. Im Bereich der Restore- und Volumen-Optimierung bietet Oracle die Zero Data Loss Recovery Appliance (ZDLRA) an. Dieses Engineered System verringert das Backupvolumen, indem lediglich ein initiales Komplettbackup erforderlich ist. Danach reichen ausschliesslich inkrementelle Sicherungen aus. Basierend auf der initialen Komplettsicherung und den nachfolgenden inkrementellen Sicherungen kann sich die Appliance über die logische Verknüpfung der Datenblöcke auf jeden Zeitpunkt innerhalb der Aufbewahrungsdauer einen komplettes incremental level 0 Backup zusammenbauen.

### **Die intelligente und zentrale Backup Steuerung**

Trotz all den hilfreichen technischen Funktionen wurde bisher eine Komponente zur Optimierung nicht betrachtet, nämlich die Steuerung der Backupvorgänge. Fragen Sie sich einmal: Wann tanken Sie Ihr Auto? Jeden Montag exakt um sieben Uhr oder vielleicht doch erst wenn der Tankinhalt zur Neige geht. Wohl eher das Letztere oder mit anderen Worten Sie tanken bedarfsgerecht, wenn es nötig wird. Genauso verhält sich eine moderne Backupsteuerung, welche auf starre Sicherungsintervalle verzichtet und stattdessen anhand von Richtlinien entscheidet wann ein Backup erfolgen muss. Für diese auf Richtlinien basierende Entscheidung braucht die zentrale Steuerung einen intelligenten Algorithmus. Die Abbildung 1 zeigt grafisch die Architektur einer möglichen und bewährten Implementation, welche spürbar die Anzahl Backupvorgänge und somit die Gesamtlast reduziert.

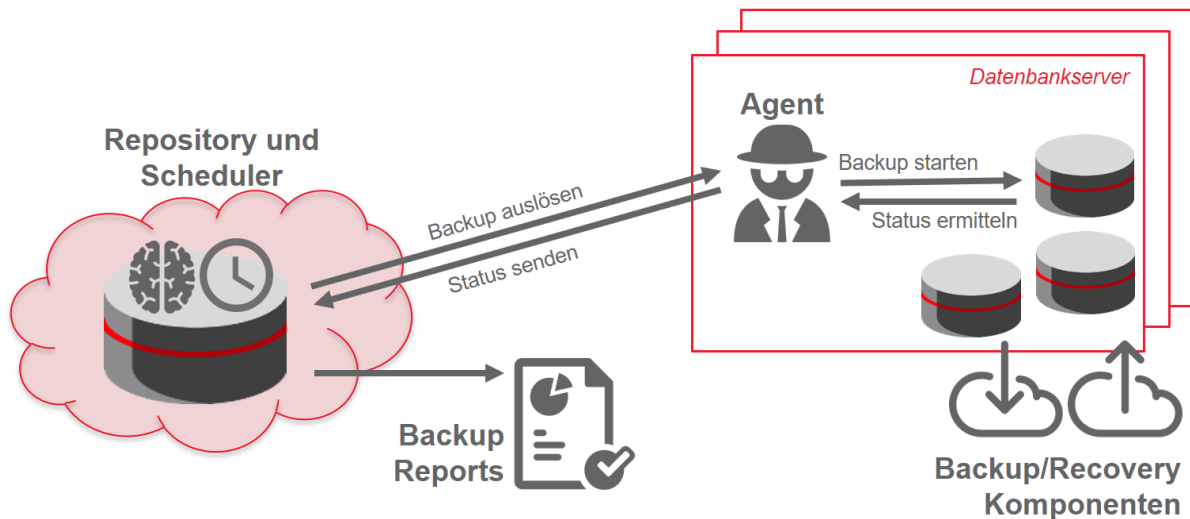


Abb. 1: Architektur einer intelligenten Backup Steuerung

Auf jedem Datenbankserver läuft ein Agent, welcher kontinuierlich die Situation zur System- und Datenbankanlastung ermittelt und an das zentrale Repository sendet. Darunter fallen beispielsweise der CPU Verbrauch, die I/O Belastung, das Transaktionsvolumen, der Füllgrad der Archive Log Destination, die Rolle der Datenbank (Primary oder Standby), etc. Im Repository selbst läuft in kurzen Abständen ein Algorithmus in Form eines Datenbankjobs, welcher aufgrund der aktuellen Situation und den definierten Richtlinien für eine bestimmte Datenbank entscheidet, ob ein Backup erfolgen soll oder nicht. Entsprechende Backup Jobs werden erstellt und vom Agenten abgeholt und ausgeführt. Abschliessend meldet der Agent den Status über das Backup an das Repository zurück und der Kreislauf schliesst sich. Jede Datenbank ist einer bestimmten Richtlinie zugewiesen und somit gilt es initial zu überlegen wie die Datenbanken kategorisiert werden sollen (z.B. Produktion, Test, DWH). Mit anderen Worten widerspiegelt eine Richtlinie letztendlich die Anforderungen an eine bestimmte Kategorie. Nachfolgend werden nicht abschliessend ein paar Anforderungen genannt, welche in einer Richtlinie enthalten sein könnten: Zeitfenster für die Sicherung, Transaktionsvolumen zwischen zwei Sicherungen, Füllgrad der Archive Log Destination, Maximale CPU & I/O Last, Recovery Point Objective (RPO). Eine intelligente Backupsteuerung entscheidet aber nicht nur aufgrund der Anforderungen einer Richtlinie, sondern plant die Sicherungen innerhalb des konfigurierten Zeitfensters optimal ein. Dank dieser Verteilung der Backupoperationen werden Performance Engpässe vermieden und die Backup bedingte Last ist ausgeglichener ohne grosse Schwankungen.

Neben dem Herzstück der intelligenten Steuerung sollten bei der Implementation der Architektur in Abbildung 1 weitere Punkte beachtet werden. Der Agent muss in der Lage sein bei einem temporären Ausfall der zentralen Steuerung eigenständig weiter zu sichern, indem sich der Agent lokal die Sicherungsintervalle der letzten Tage merkt und gleich weiterfährt. Dadurch wird einem Datenbankstillstand durch einen Archiver Stuck vorgebeugt und dem Risiko eines Single-Point-of-Failure entgegengewirkt. Der zentrale Ansatz bietet insbesondere im Bereich der Backup Verwaltung riesige Vorteile indem der DBA mit einem einzelnen Kommando beliebige Server/Datenbanken in den Wartungsmodus setzen kann. Auch im Sinne der Transparenz lässt die Lösung ihre Muskeln spielen, denn sämtliche Informationen rund um das Backup sind im Repository zentral gespeichert und können für Auswertungszwecke weiterverwendet werden.

**Fazit**

Die Praxis und die Erfahrung zeigt, dass die Nutzung von Oracle Funktionalität bereits einiges an Optimierung ermöglicht. Doch erst die bedarfsgerechte und intelligente Steuerung der Backupoperationen bringt die entscheidenden Mehrwerte im Bereich Effizienz, Risikominimierung und Transparenz zum Vorschein.

**Kontaktadresse:**

Roland Stirnimann  
Trivadis AG  
Sägereistrasse 29  
CH-8152 Glattbrugg

Telefon: +41 58 459 55 55  
E-Mail [roland.stirnimann@trivadis.com](mailto:roland.stirnimann@trivadis.com)  
Internet: [www.trivadis.com](http://www.trivadis.com)