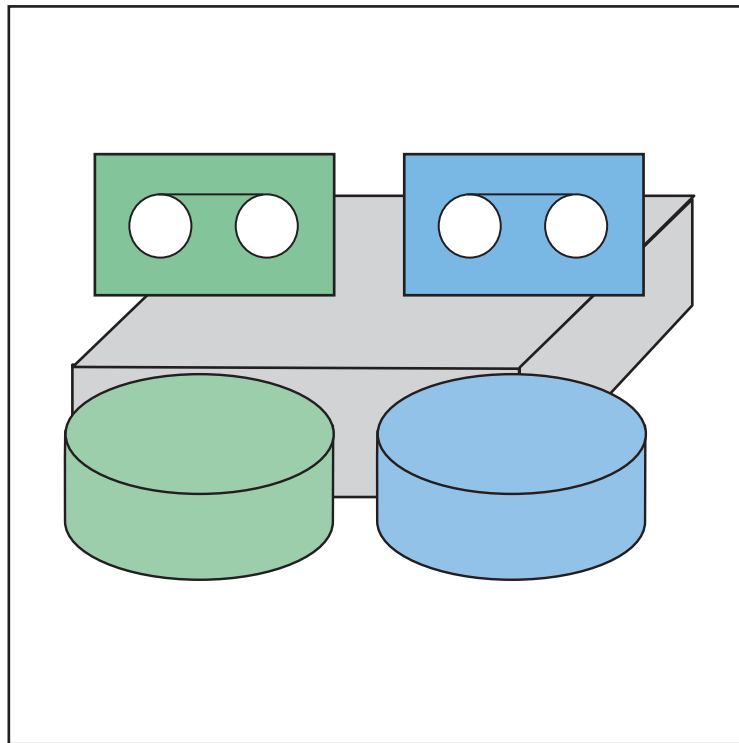


Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

von Dipl.-Inform. Matthias Egerland



Das Volumen der Daten, die innerhalb eines Unternehmens anfallen, wächst exponentiell. Wachstumsraten von 50% pro Jahr sind hier keine Seltenheit, auch bei unternehmenskritischen Daten. Hier spielt SAP eine zunehmend wichtige Rolle.

Unternehmen, die seit Jahren auf SAP setzen, sind genauso von der Frage nach

der Datensicherheit betroffen, wie Unternehmen, die die Einführung von SAP erst planen. Über mehrere Standorte verteilte Cluster-Lösungen können zwar den aktiven Datenstand sicherstellen, doch für logische Fehler ist die Verfügbarkeit von Datensicherungen essentiell. Da auch in einer Multi-Site-Umgebung der Ausfall eines beliebigen Standorts vom Backup-Konzept abgebildet werden muss, stellt sich gleich-

zeitig die Frage, wie Datensicherungen simultan auf mehrere Standorte verteilt werden können. Die hierfür relevanten sowohl organisatorischen als auch technologischen Überlegungen werden im Folgenden analysiert.

weiter auf nächster Seite

Schwerpunktthema

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

Fortsetzung von Seite 1



Dipl.-Inform. Matthias Egerland hat an der RWTH Aachen Informatik studiert und ist seit 2005 Mitarbeiter der ComConsult Beratung und Planung GmbH. Er ist Leiter des Competence Center Data Center und unterstützt die Competence Center IT-Sicherheit und Netze. Neben den Schwerpunkten Desktop-, Server- und Infrastruktur-Virtualisierung beschäftigt sich Herr Egerland insbesondere mit Speicherlösungen in virtualisierten Umgebungen. In Projekten erstellt er Konzepte und Ausschreibungen von IT-Infrastruktur-Lösungen gemäß UfAB in den Bereichen Lokale Netze (LAN) für mehrere tausend Teilnehmer, aktive Rechenzentrumskomponenten – insbesondere Server, Speicher, Netzwerk und Firewalls – sowie Virtuelle Informationstechnologie.

1. Datensicherung und -wiederherstellung im SAP-Umfeld

In diesem Kapitel werden grundlegende Verfahrensweisen zur Datensicherung und -wiederherstellung dargestellt. Nach den allgemeinen Aspekten und Datensicherungsverfahren inkrementeller, differentieller und block-basiert inkrementeller Backups wird auf das Konzept virtueller Tape Libraries (VTL) eingegangen. Nach der exemplarischen Darstellung des Aufbaus einer SAP-Umgebung werden die verschiedenen Software-Komponenten dargestellt, die für die Durchführung einer Datensicherung ineinander greifen. Da SAP auf Oracle-Datenbanken aufbaut, wird anschließend die Struktur näher beleuchtet, die Oracle zur Durchführung von Datensicherungen bereitstellt.

1.1 Allgemeines

Völlig unabhängig von den Redundanzen in der Speicherumgebung oder in den Teilkomponenten der Speicherumgebung sind Datensicherungen unabdingbar für die Erhaltung der Verfügbarkeit von Daten. Während Redundanzen im Wesentlichen vor Hardware-Fehlern schützen, bewahrt die Datensicherung vor dem Datenverlust bei logischen und operationellen Fehlern sowie Software-Schäden. Selbst beim Einsatz von Clusterlösungen auf Server-Speicherebene hat daher eine regelmäßige Datensicherung zu erfolgen. Zu sichern sind die Daten, die für den Fortbestand und die Aufrechterhaltung des produktiven Betriebes notwendig sind.

Eingesetzt werden die durchgeführten Datensicherungen für die Wiederherstellung von zerstörten, gelöschten oder aus sonstigen Gründen nicht mehr zugänglichen Daten.

Für ein genaues Begriffsverständnis werden die folgenden Definitionen festgehalten:

- Unter einer Sicherung oder Backup wird die Speicherung von aktiven Daten verstanden. Aktive Daten sind die Daten, die auf den Servern und/oder auf den Speichersystemen für den produktiven Betrieb der Systeme und Applikationen benötigt werden.
- Unter einer Wiederherstellung oder einem Recovery wird der Wiederherstellungsprozess von Daten verstanden, die zuvor gesichert worden sind. Wenn z.B. ein Verzeichnis gelöscht wurde, dann ist dieses Verzeichnis wiederherzustellen. Unter Recovery fällt auch die Wiederherstellung eines Serversystems mit den Betriebssystemdaten und den Applikationsdaten.

- Das Zurückholen von archivierten oder bereits ausgelagerten Daten wird als Restore bezeichnet.

Für die Konzipierung einer Datensicherung sind sowohl technische als auch organisatorische Aspekte zu beachten. Die technischen Aspekte dienen der Auswahl der einzusetzenden Technik, welche als Basis für eine Technologieauswahl die folgenden Punkte berücksichtigen sollte:

- Verfügbarkeit der Datensicherungslösung,
- Datenvolumen, welches zu sichern ist,
- Performance, die notwendig ist, um eine Sicherung innerhalb von definierten Zeiten durchführen zu können bzw. Daten wiederherstellen zu können (maximale Wiederherstellungszeit),

- Applikationen, die in Abhängigkeit der eingesetzten Sicherungssoftware entsprechende Unterstützung durch Agenten und/oder Optionen erfordern, und
- Sicherungsverfahren.

Ein wichtiges Ziel bei der Datensicherung ist die Minimierung der Zeit, die zwischen folgenden Punkten verstreicht:

- Recovery Point Objective (RPO): Zeitpunkt, von dem die letzten konsistenten und korrekten Daten vorliegen und auf den zurückgesetzt werden kann
- Recovery Time Objective (RTO): Zeitpunkt, von dem an die Informations- und Kommunikationsdienstleistungen wieder verfügbar sind mit dem Stand des letzten konsistenten und korrekten Sicherungspunktes

Der Zusammenhang zwischen RPO, Zeitpunkt des Katastropheneintritts, Wiederherstellungszeit und RTO ist in Abbildung 1 dargestellt.

Zwischen dem Eintritt der Katastrophe und dem RTO verstreicht die Wiederherstellungszeit. Im Katastrophenfall (zum Beispiel Ausfall eines Rechenzentrums) entspricht der Zeitpunkt des Eintritts der Katastrophe dem Konsistenzpunkt, wenn die Daten gespiegelt sind.

Die organisatorischen Aspekte einer Konzipierung berücksichtigen den Standort der Datensicherungsmedien und deren Lagerung bzw. Auslagerung. Ebenso ist eine Sicherungsstrategie notwendig, die besonders im Falle der Wiederherstellung einen ausschlaggebenden Zeitfaktor darstellen kann. So ist eine inkremen-

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

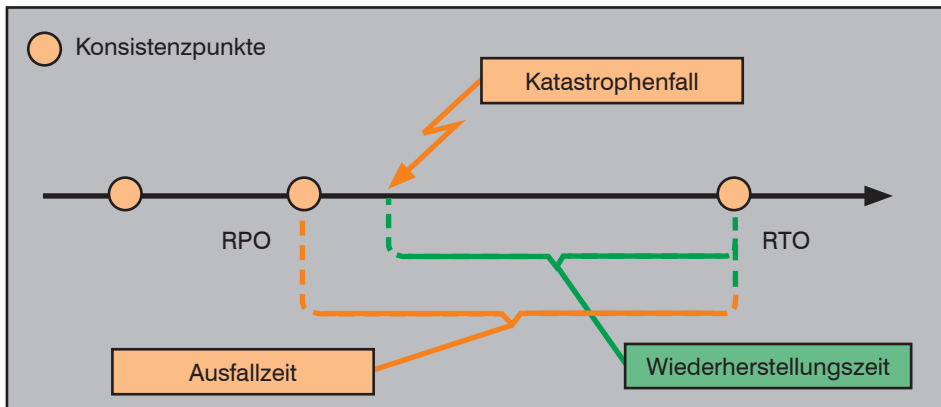


Abbildung 1: K-Fall-Vorsorge durch Datensicherung

telles Sicherung bei der Wiederherstellung organisatorisch zeitaufwendiger als eine differenzielle oder volle Sicherung. Ein weiterer Aspekt ist die Aufteilung der Verantwortungsbereiche nach Sicherungsdurchführung und Wiederherstellung bzw. Restore.

Grundsätzlich sind die Anforderungen an die Durchführung eines zeitgemäßen Backups in den letzten Jahren gestiegen. Die Erstellung von Duplikaten von Sicherungen und eine sichere Auslagerung werden im Kontext eines Disaster-Recovery-Konzeptes immer wichtiger. Im Folgenden werden Lösungen und Techniken für die Datensicherung dargestellt.

1.2 Sicherungsverfahren

Von marktüblicher Datensicherungs-Software werden mindestens die nachfolgend aufgeführten Sicherungsverfahren unterstützt:

- Bei einer Vollsicherung werden mit jeder Ausführung alle Daten im vollen Umfang gesichert.
- Bei einer inkrementellen Sicherung werden die hinzugefügten oder geänderten Daten seit der letzten inkrementellen Sicherung in die Sicherung aufgenommen, wie aus Abbildung 2 hervorgeht. In Verbindung mit einer initialen Vollsicherung kann so über mehrere Jobs und in der Regel auch Medien eine komplette Sicherung aufgebaut werden. Zu beachten ist, dass hierbei längere Recovery-Zeiten aufgrund der höheren Anzahl von Jobs und Medien, die zu diesem Zweck eingesetzt werden, einzuplanen sind.
- Bei einer differenziellen Sicherung werden die hinzugefügten oder geänderten Dateien seit der letzten Vollsicherung in die Sicherung aufgenommen, wie in Abbildung 3 dargestellt. In Verbindung mit einer Vollsicherung kann hier mit-

tels zwei Jobs eine Komplettsicherung aufgebaut werden. Dies hat zur Folge, dass die differenziellen Sicherungen von Tag zu Tag größer und damit zeitlich länger werden. Für die Wiederherstellungszeit ist dies jedoch positiv, da hier nur zwei Jobs zu berücksichtigen sind.

1.2.1 Block Level Incremental Backups (BLIB)

Die Möglichkeit, inkrementelle Datensicherungen auf Blockebene durchzuführen, wird als Block Level Incremental Backup (BLIB) bezeichnet. Hierbei werden zunächst nur alle belegten Blöcke einer zu sichernden Partition übertragen. Bei jedem inkrementellen Sicherungsvorgang werden anschließend nur die Blöcke gesichert, die sich gegenüber dem letzten Vorgang geändert haben. Dies hat i.W. drei Vorteile:

1. Das Volumen der zu sichernden Daten kann auf Blockebene deutlich kleiner ausfallen als auf Dateiebene, da nur die veränderten Daten einer Datei (blockbasiert) gesichert werden.

2. Da die Datensicherung selbst stets aus

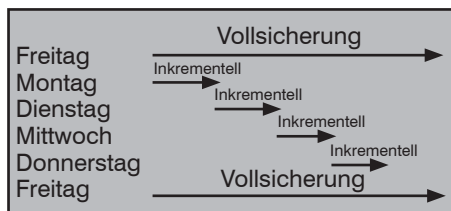


Abbildung 2: Inkrementelle Sicherung

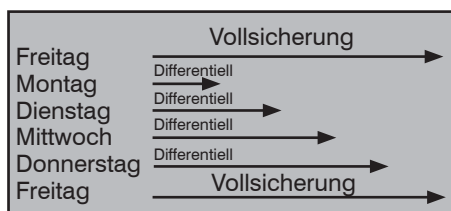


Abbildung 3: Differenzielle Sicherung

einem blockbasierten Abbild der Festplatten-Partition besteht, kann dieses Abbild z.B. per iSCSI an einem neuen Server gemounted werden. Dies gilt auch für die inkrementellen Backups im Zusammenhang mit den vorangegangenen Teilsicherungen und der letzten Vollsicherung. Damit ist ein System in Disaster-Recovery-Fällen nahezu ohne Zeitverlust wieder lauffähig. Dieser Vorgang wird typischerweise „Instant Availability Mapping“ genannt.

3. Einzelne Dateien können aus dem blockbasierten Gesamtabbild restauriert werden, ohne die gesamte Datensicherung zunächst vollständig zurückspielen zu müssen. Nach dem Mounten des blockbasierten Abbilds wird hierfür auf die zu restaurierenden Daten dateibasiert zugegriffen.

Diese Vorteile sprechen für einen BLIB-Ansatz zur Datensicherung. Allerdings wird dieser nicht von jeder Backup-Software und für alle zu sichernden Umgebungen unterstützt, da er Snapshot-basiert arbeitet und zur Erzeugung von applikationskonsistenten Snapshots ein entsprechender Agent erforderlich ist.

1.2.2 Synthetische Backups

Anstelle in regelmäßigen Abständen den kompletten Datenbestand für eine aktuelle Vollsicherung über das Netzwerk zu transportieren, können neue Full-Backups auch auf Basis der letzten Vollsicherung und einer Reihe von inkrementellen Sicherungen neu berechnet werden. Allerdings kann der Ressourcenaufwand für eine solche Konsolidierung von Datensicherungen beträchtlich sein und bei unzureichender Auslegung der hierfür verantwortlichen Server-Hardware die Backup-Zeit um ein Vielfaches übersteigen. Außerdem zeigt die Praxis, dass es dennoch ratsam ist, in regelmäßigen Abständen eine neue Vollsicherung zu erzeugen. Damit wird verhindert, dass sich aufgrund von Übertragungs- oder Software-Fehlern Inkonsistenzen in die synthetischen Backups einschleichen, die erst im Recovery-Fall offensichtlich werden.

1.3 Virtuelle Tape Library (VTL)

Eine virtuelle Tape Library ist ein Festplattenspeichersystem, das sich mittels einer entsprechenden Software gegenüber den Servern als physische Bandbibliothek darstellt. Hierbei wird der Bandlaufwerkstyp eines speziellen Herstellers so perfekt emuliert, dass aus Sicht der zu sichernden Server und der Backup-Software kein Unterschied zwischen einer physischen Bandbibliothek und der virtuellen Tape Library erkannt werden kann.

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

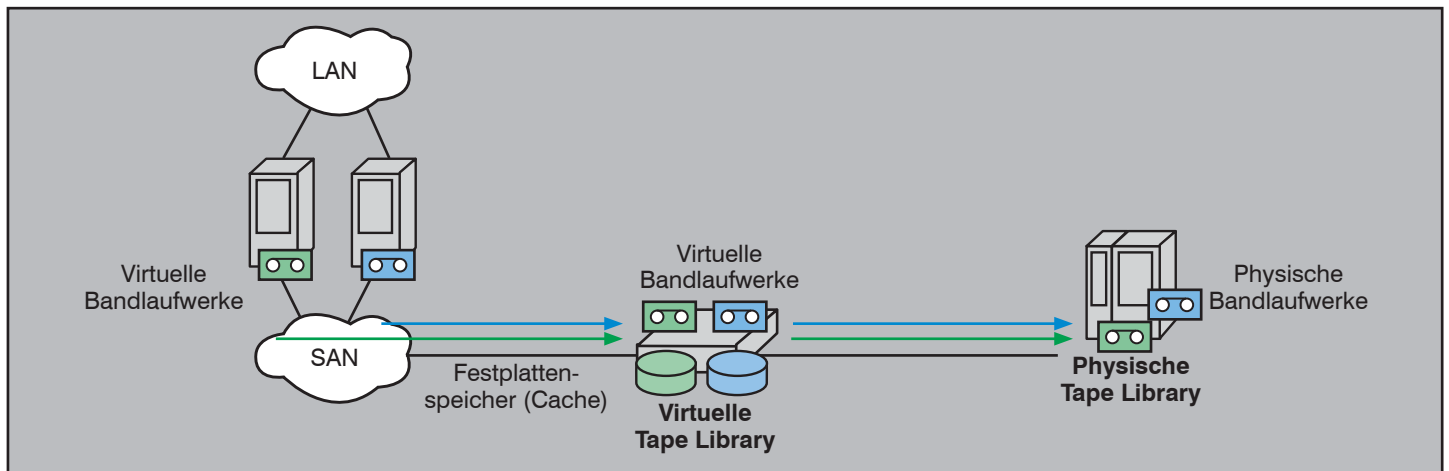


Abbildung 4: Anbindung Virtuelle Tape Library

Virtuelle Tape Libraries nutzen den Geschwindigkeitsvorteil durch die Sicherung auf Festplatten im ersten Schritt des gesamten Datensicherungsprozesses aus. Dieser Geschwindigkeitsvorteil besteht nicht nur in den höheren Schreib- und Leserraten auf das virtuelle Band. Insbesondere das Verteilen von Backup-Datenströmen und der anschließende Zugriff auf verteilte vorliegende Datensicherungen werden durch eine VTL massiv beschleunigt, da sämtliche Start- / Stoppvorgänge und das Spulen eines physischen Bandes entfallen.

Die VTLs können über das SAN angebunden werden und den Serversystemen über dieses SAN virtuelle Bandlaufwerke zur Verfügung stellen. Diese Laufwerke emulieren unterschiedliche Bandtechnologien. Der Zugriff erfolgt durch das Serversystem wie auf ein physikalisches Bandlaufwerk. Tatsächlich jedoch beinhaltet die VTL keine Bandlaufwerke, sondern über RAID-Technologie abgesicherte Festplatten in einem Verbund. Eine Software stellt jeweils Teile dieses Festplattenverbundes für die Serversysteme als Bandlaufwerk dar. Die Sicherung erfolgt somit tatsächlich auf Festplatten mit dem damit verbundenen Performancevorteil gegenüber physikalischen Bandlaufwerken. Der Start-Stopp-Betrieb entfällt bei dieser Technik, ebenso wie die mechanischen Beanspruchungen von Bändern. Mittels dieser Technik kann jedem Serversystem, das über das SAN angebunden ist, von zentraler Stelle aus ein Bandlaufwerk zur Verfügung gestellt werden.

Bei Einsatz von VTL wird die gleichzeitige Durchführung von Datensicherungsjobs von mehreren Serversystemen unterstützt. Unterschiedliche Bandtechnologien für die Jobs können emuliert werden.

Optional können entweder durch einen Algorithmus der VTL oder mittels einer Regel der Backup-Software die auf den Festplatten vorgehaltenen Daten auf tatsächliche physikalische Bandlaufwerke ausgelagert werden, die entweder direkt an die VTL oder über das SAN angebunden sind. Diese physikalischen Tape Libraries besitzen wiederum mehrere physikalische Bandlaufwerke und ausreichend Platz für die Aufnahme von Bandmedien. Die auf dem Markt zu diesem Zweck allgemein eingesetzten Tape Libraries sind in Bezug auf die Anzahl der Laufwerke und die Anzahl der Medienslots äußerst skalierbar (siehe auch Abbildung 4).

Mittels VTL ist nicht nur eine schnellere Sicherung realisierbar, sondern auch eine schnellere Wiederherstellung, solange die Daten noch auf den Festplatten innerhalb der virtuellen Tape Libraries vorgehalten werden. Ansonsten werden die Daten vom Bandmedium in die virtuelle Tape Library zurückgeholt und mittels der Emulation auf das Serversystem zurückgespielt.

Die Daten, die auf einem zentralen Speichersystem abgelegt sind, können über die Serversysteme und das SAN ggf. unter Einsatz von separaten HBAs für die Datensicherung auf die Festplatten der virtuellen Library und von dort auf die physikalische Tape-Library gesichert werden.

Zusammengefasst bestehen die Vorteile einer VTL in den folgenden Aspekten:

- Sie kann das Backup beschleunigen weil die Daten nicht auf Band, sondern auf ein schnelles Disk-System geschrieben werden.
- Besonders bei der Wiederherstellung von vielen kleinen Dateien bietet eine VTL einen Geschwindigkeitsvorteil gegenüber Bandbibliotheken.

genüber Bandbibliotheken.

- Eine VTL bietet erhebliche Geschwindigkeitsvorteile bei auf mehrere Laufwerke gestreuten Lesevorgängen.
- Der Datenfluss unterliegt nicht den Kontinuitätsanforderungen eines physischen Bandlaufwerks.
- Physische Bandmedien können eingespart werden oder werden weniger beansprucht.
- Backups können unabhängig von den Zeitfenstern für physische Laufwerke geschrieben werden.
- VTLs können nahtlos in die vorhandene Infrastruktur integriert werden. Da sie eine Bandbibliothek emulieren, sind keine Änderungen an der vorhandenen Backup-Struktur erforderlich. Die Server verwenden wie bisher ihren Backup-Client.
- Zugewinn an Flexibilität: zusätzliche (virtuelle) Laufwerke können „auf Knopfdruck“ als Backup-Ziele definiert werden.

Dem gegenüber stehen die folgenden Nachteile einer VTL:

- Zusätzliche Investitionskosten
- Zusätzlicher Betriebsaufwand
- Erhöhte Komplexität in der SAN Infrastruktur
- Mangelnde Transparenz der VTL-Funktionalitäten gegenüber der Backup-Software
- Beispiel: Datensicherungen, die von der VTL dupliziert werden, sind der Backup-Software nicht bekannt.

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

- Zu hohe Transparenz der VTL gegenüber der Backup-Software
- Die Backup-Software kann zwischen virtuellen und physischen Medien nicht unterscheiden. Besteht der Bedarf zur Auslagerung des Backups auf ein physisches Band, muss ein solches Laufwerk besonders kenntlich gemacht werden.
- Einfachere Transportierbarkeit von physischen Bändern, z.B. zur Archivierung entfällt
- Geringere Schutzmöglichkeiten der Daten im Vergleich zu ausgelagerten Bändern

1.4 Aufbau einer SAP-Umgebung

In einer SAP-Umgebung sind in der Regel die Applikations-Server im Sinne einer Multi-Tiering-Architektur von den Datenbanken getrennt. D.h., dass ein Server die SAP-Applikation beinhaltet, während ein anderer Server die von der Applikation genutzte Datenbank betreibt.

Da SAP zu den unternehmenskritischen Diensten gehört, wird dieser Dienst aus Verfügbarkeitsgründen in der Regel über zwei Rechenzentren verteilt. Die geclusterten SAP-Applikations-Server verwenden die "Enqueue-Replikation" zur zentralen Verwaltung von logischen Sperren. Auf der Oracle-Seite sorgt der Real Application Cluster (RAC) für die synchrone Replikation der Daten. Abbildung 5 verdeutlicht diesen Aufbau.

1.5 Datensicherungen im SAP-Umfeld

1.5.1 Backup-Software

Im SAP-Umfeld besteht häufig die Anforderung, Datensicherung direkt aus der eigenen Oberfläche heraus auslösen zu können. Dies wird über die Einbettung von entsprechenden Skripten in Schaltflächen oder Menüeinträge des SAP-GUI realisiert. Ein solches Skript generiert eine Kommandozeile für die SAP Backup & Recovery Tools (BRTOOLS). Die BRTOOLS wiederum erzeugen Kommandos für den Oracle Recovery Manager (RMAN). RMAN ist in der Lage, per „Block Change Tracking“ nur die veränderten Blöcke der Datenbank lesen zu müssen. Dies entspricht dem weiter oben beschriebenen Block Level Incremental Ansatz.

Der Backup-Job wird letztlich an die Dritthersteller Backup-Software übertragen und von dort auf die Ziel-Laufwerke geschrieben. Die Dritthersteller-Software übernimmt dabei die Steuerung und Verwaltung der Bandlaufwerke und -roboter.

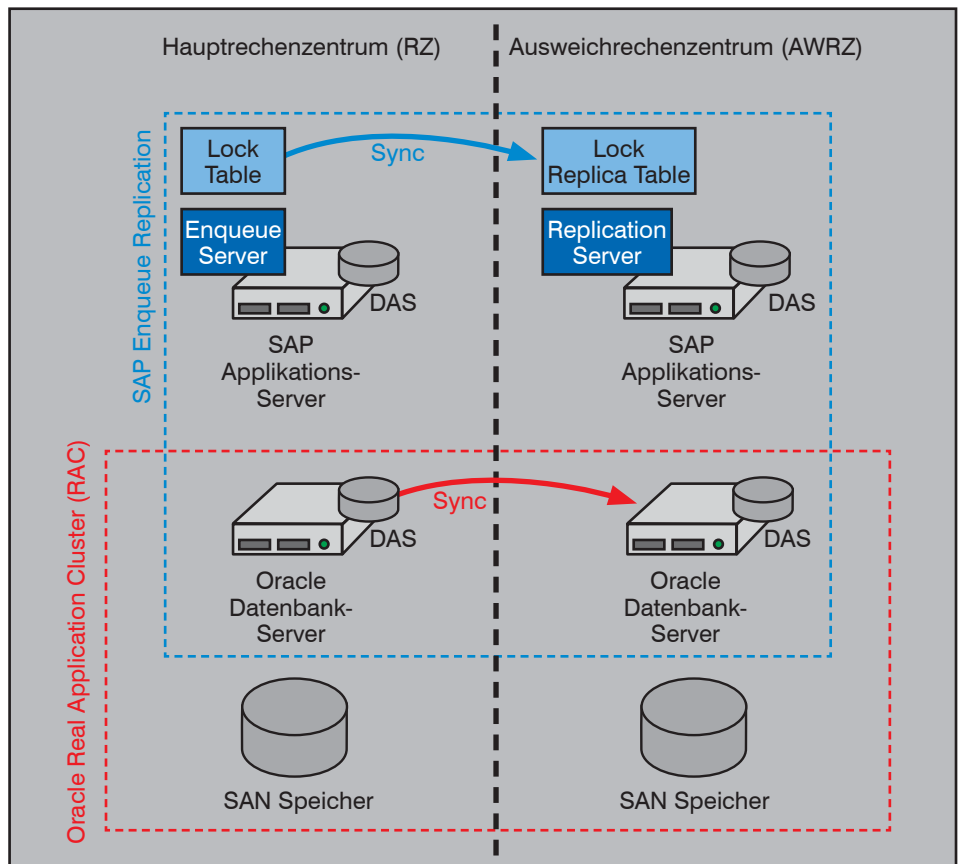


Abbildung 5: Redundante SAP-Umgebungen durch Oracle Real Application Cluster (RAC) und SAP Enqueue Replication

Abbildung 6 zeigt dieses Ineinandergreifen der unterschiedlichen Backup-Produkte im SAP-Umfeld.

Problematisch wird diese Konstellation, wenn insbesondere RMAN und die Dritthersteller-Software nur begrenzt miteinander harmonieren. Ist z.B. nicht klar geregelt, wer die „führende Rolle“ übernimmt, kann es zu Inkonsistenzen in den jeweiligen Backup-Katalogen kommen oder es muss mit entsprechenden Aufwänden für Konsistenz gesorgt werden.

Um veraltete Backups beispielsweise zu löschen, muss in der Dritthersteller-Software eine „Retention Period“ von „Unendlich“ eingetragen werden. Unter RMAN bzw. über die BRTOOLS wird dann die Lebenszeit der Backups definiert. Durch einen täglichen „Delete Obsolete“-Lauf von RMAN werden die überflüssigen Backups schließlich gelöscht.

1.5.2 Backup von Oracle-Datenbanken

SAP setzt als Datenbanksystem Produkte des Herstellers Oracle ein. Insofern sind für Datensicherungen im SAP-Umfeld die Mechanismen für das Backup von Oracle Datenbanken zu betrachten.

Der gültige Ausgangszustand einer Oracle Datenbank enthält alle Einträge, Relationen und Schlüssel in konsistenter Form. Alle Änderungen und Ergänzungen, die im

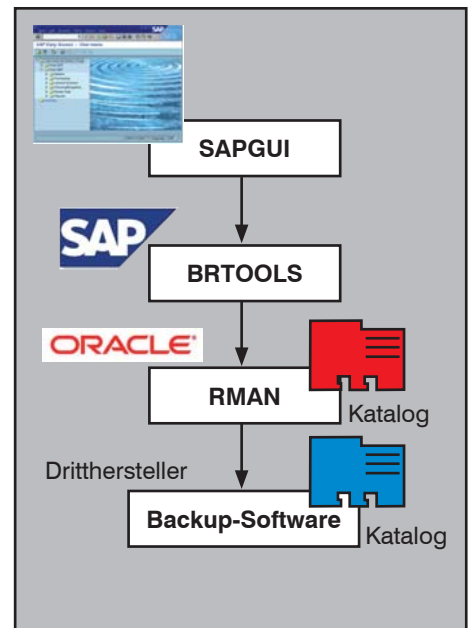


Abbildung 6: Ineinandergreifen unterschiedlicher Backup-Software im SAP-Umfeld

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

Laufe der Zeit an der Datenbank durchgeführt werden, werden in so genannte „Online Redo Logs“ gespeichert. Wird ein Fehler oder eine Inkonsistenz in der Datenbank entdeckt, kann sie mit Hilfe der Redo-Logs schrittweise wieder rekonstruiert werden. Um den Speicherplatzbedarf zu reduzieren, werden die Redo-Logs in zyklischer Abfolge überschrieben. Dabei wird in der Regel eine zeitliche Abfolge definiert.

Wird beispielsweise alle 6 Stunden ein neues Online Redo Log geschrieben und nach 4 Online Redo Logs das erste wieder überschrieben, steht stets Restaurierungsmaterial für die letzten 24 Stunden zur Verfügung. Dabei macht jedes Online Redo Log im Schnitt ein Viertel des gesamten Änderungsvolumens dieses Zeitraums aus.

Die Datensicherung von Oracle-Datenbanken betrifft insofern 2 unterschiedliche Aspekte. Zum einen ist dies das Backup der gesamten Datenbank, das zur Reduktion der zu übertragenen Gesamtmenge blockweise inkrementell oder differentiell erfolgen kann. Dabei lautet die Bezeichnung für differentielle Backups unter RMAN „incremental level 0“.

Zum Anderen sind es die Online Redo Logs, die archiviert werden sollten, um möglichst schnell fehlerhafte Änderungen korrigieren zu können. Das Sichern der Online Redo Logs wird bei SAP über den ARCHIVELOG-Modus im SAPDBA-Menü eingestellt.

Abbildung 7 veranschaulicht diese Vorgehensweise.

Durch diese kombinierten Datensicherungsverfahren - Vollsicherung der Datenbank einerseits, Archivierung der Redo Logs andererseits - können folgende Aspekte in der Backup-Strategie berücksichtigt werden:

- Häufiges Archivieren der - vom Datenvolumen her relativ kleinen - Redo Logs, um beliebige Zwischenstände der an der Datenbank vorgenommenen Änderungsvorgänge rekonstruieren zu können.
- Selteneres Sichern der vollständigen Datenbank - gegebenenfalls kombiniert mit inkrementellen oder differentiellen Zwischenschritten.

2. Synchrone Backups an unterschiedlichen Standorten

In der Praxis besteht häufig die Anforderung, Datensicherungen nicht nur an einem Standort vorzuhalten, sondern auch auf einen zweiten Standort zu schreiben.

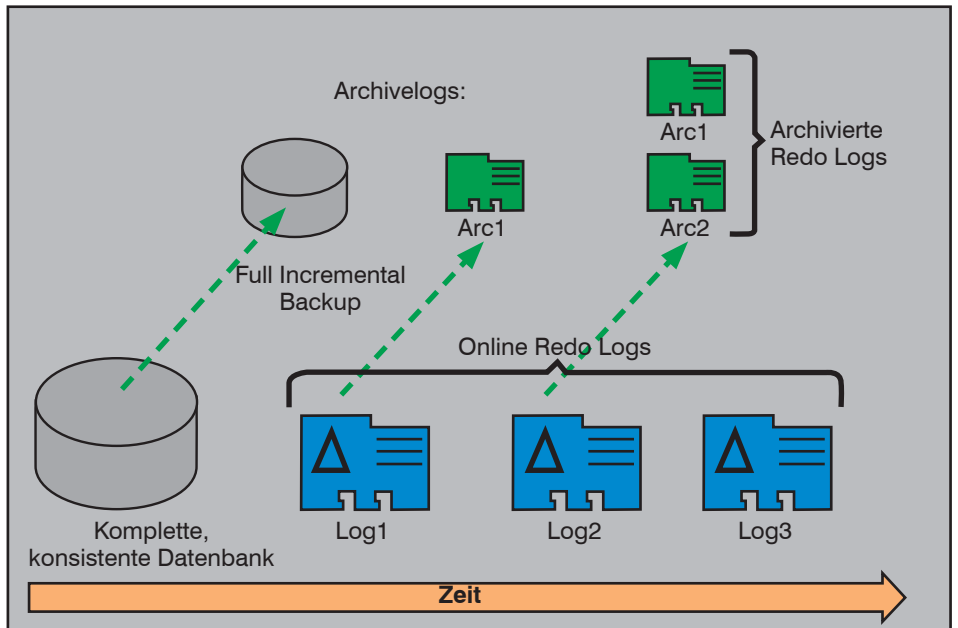


Abbildung 7: Datensicherung von Oracle-Datenbanken

Dabei gebietet es ein Disaster Recovery Konzept, dass diese Datensicherungen möglichst zeitgleich und möglichst identisch an beiden Standorten vorliegen. Im Folgenden werden mögliche Lösungsvarianten für das Verteilen von Datensicherungen auf verschiedene Backup-Ziele aufgezeigt und bewertet.

2.1 Variante 1: Parallele Sicherungsjobs

Die einfachste Methode, eine Datensicherung an unterschiedliche Backup-Ziele zu übertragen, besteht darin, auf dem zu sichernden Server zwei parallele Backup-Jobs zu starten. Diese lesen allerdings die Daten unabhängig voneinander, so dass es zu doppelten Zugriffen auf die - gegebenenfalls im SAN platzierten - Platten-

speicher des Servers kommt. Außerdem verbrauchen die simultanen Backup-Jobs 2x die Systemressourcen auf dem zu sichernden Server. Des Weiteren werden die Backup-Daten 2x über die Serverschnittstelle über das SAN in Richtung der Backup-Ziele geschrieben. Da die Backup-Jobs letztlich unabhängig voneinander sind, werden sie auch separat inventarisiert und besitzen unterschiedliche Zeitstempel.

Dem gegenüber besteht ein Vorteil in der Einfachheit dieser Lösung: Bei manchen am Markt befindlichen Backup-Produkten ist dies außerdem die einzige Möglichkeit, Datensicherungen auf zwei Backup-Ziele zu verteilen.

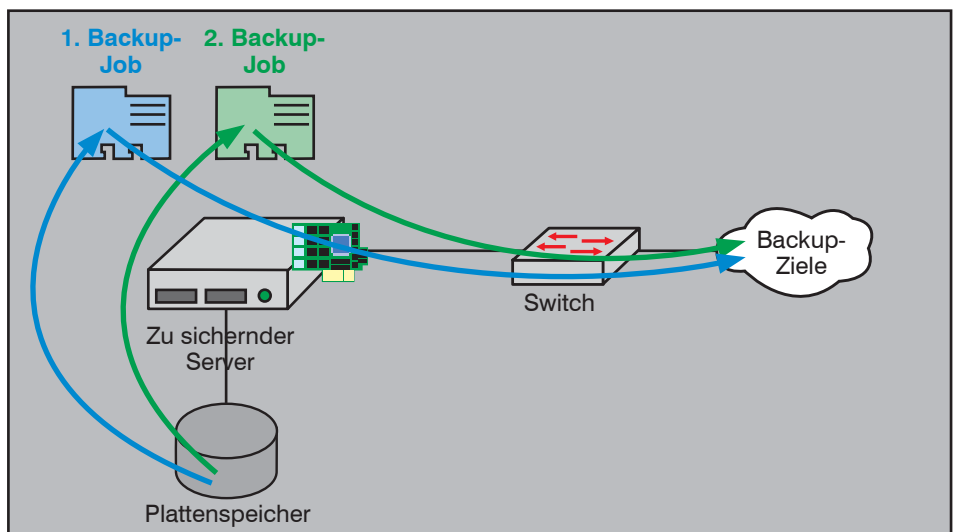


Abbildung 8: Replikation von Datensicherungen durch parallele Backup-Jobs

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

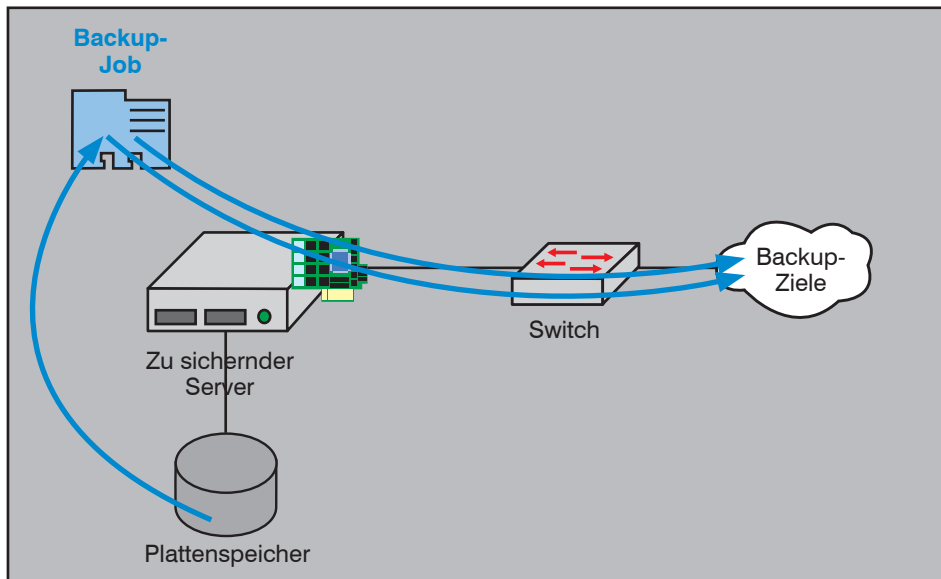


Abbildung 9: Replikation von Datensicherungen durch Adressierung mehrerer Backup-Ziele

Abbildung 8 veranschaulicht diesen Ansatz.

2.2 Variante 2: Adressierung mehrerer Backup-Ziele

Einige Backup-Produkte gestatten es, mehrere Backup-Ziele von ein und demselben Backup-Job aus zu adressieren. Dies führt zu dem Vorteil, dass die zu sichernden Daten zur 1x von Platte gelesen werden müssen und nur ein Backup-Job das zu sichernde System belastet. Da die Backups den gleichen Ursprung haben, besitzen sie außerdem den gleichen Zeitstempel.

In Richtung der Backup-Ziele muss die Datensicherung jedoch weiterhin mehrfach transportiert werden, was zu einer entsprechenden Belastung der Server-Schnittstellen und des Speichernetzes in dieser Richtung führt. Der Backup-Job wird erst dann abgeschlossen, wenn auch das langsamere der beiden Backup-Ziele fertig ist (typischerweise das an einem Remote-Standort).

Abbildung 9 veranschaulicht diesen Ansatz.

2.3 Variante 3: Verteilung des Datenstroms

Soll der zu sichernde Server noch weiter entlastet werden von den zu verteilenden Datensicherungen, kann der Medien-Server – der die Backup-Ziele anbindet – zur Verteilung des Datenstroms genutzt werden. Dieser Ansatz ist mit dem im IP-Netz bekannten Multicasting vergleichbar. Ein einzelner Datenstrom wird von einer nachgelagerten Komponente in Richtung mehrerer Ziele verteilt.

Der Medien-Server kann einerseits im IP-Netz stehen, was allerdings eine LAN-seitige Datensicherung zur Basis hätte. Steht hierfür kein vom Produktivnetz getrenntes Backup-Netz zur Verfügung, ist dieser Ansatz unbedingt zu vermeiden.

Auch in einem Fibre Channel SAN kann der Medien-Server zur Verteilung eines Backup-Stroms herangezogen werden. In diesem Fall kommen so genannte „Fibre Channel Transport Services“ auf dem Medien-Server zum Einsatz, die den Datenstrom aufsplitten und auf mehrere Ziele verteilen. Zu beachten ist in diesem Szenario, ob die Backup-Software auf ihrem Medien-Server diese Dienste anbietet und ob das Betriebssystem auf dem Medien-Server diese unterstützt.

Abbildung 10 veranschaulicht diesen Ansatz.

2.4 Verteilung der Datensicherungen über VTLs

Eine VTL kann nicht nur eingesetzt werden, um den Datensicherungsvorgang zu beschleunigen. Sie kann auch ein Hilfsmittel zur Verteilung der Datensicherungen auf mehrere Backup-Ziele darstellen. Dabei stehen einerseits Bordmittel der VTL zur Verfügung, die ein Backup auf eine zweite VTL replizieren können und dabei gegebenenfalls von VTL-spezifischen Deduplizierungsmechanismen profitieren können. Andererseits ist eine solche über die VTLs ausgeführte Kopie der Datensicherung für die Backup-Software intransparent, so dass sie nicht in ihrer Katalogdatei geführt werden kann und für einen Restore-Vorgang manuell eingebunden werden müsste.

Insofern ist ein Replizierungsverfahren vorzuziehen, das als Speicherplattform zwei VTLs nutzt, jedoch über die zentrale Datensicherungs-Software gesteuert wird. Über so genannte „Backup-Policies“ kann bei einigen Backup-Produkten definiert werden, wie mit einer Datensicherung weiter verfahren werden soll. So kann z.B. eine nachgelagerte Replikation einer Datensicherung auf einen oder mehrere andere Datenträger vorgenommen werden. Dabei behält die Backup-Datei nicht nur den gleichen Inhalt, sondern auch den gleichen Zeitstempel. Da auch die replizierten Backup-Dateien in der Katalogdatei der Backup-Software geführt werden, stehen sie für eine unmittelbare Restaurierung zur Verfügung.

Abbildung 11 veranschaulicht diesen Ansatz.

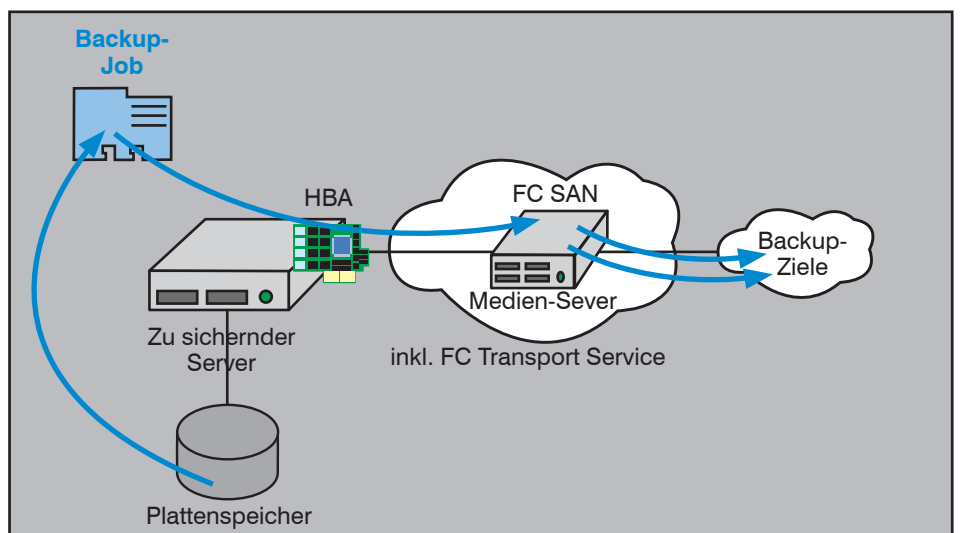


Abbildung 10: Replikation von Datensicherungen durch Nutzung eines Fibre Channel Transport Services

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

3. Backup-Szenarien mittels VTL

Im Folgenden werden drei aus den in Kapitel 2 dargestellten Varianten abgeleitete Szenarien einer Standort-übergreifenden Replizierung von Backups mittels VTLs diskutiert.

3.1 Szenario 1: Sequentiell auf VTL

In diesem ersten Szenario werden die Daten von dem zu sichernden Server zunächst auf die VTL des eigenen Standorts geschrieben. In einem nachgelagerten Prozess repliziert die Backup-Software (nicht die VTL) diese Datensicherung in das Ausweichrechenzentrum. In beiden Rechenzentren können die Datensicherungen optional noch auf physische Bänder ausgelagert werden, wenn dies z.B. zur Archivierung erforderlich ist. Dieses Szenario basiert auf der in Abschnitt 2.4 dargestellten Variante.

Anforderungen

- Die Backup-Software muss in der Lage sein, die Datensicherungen über die VTLs zu duplizieren, um diese Datensicherungen für den Wiederherstellungsfall zentral verwalten können.

Vorteile

- Nach Abschluss des Sicherungsvorgangs liegen an beiden RZ-Standorten identische Datensicherungen mit gleichem Zeitstempel vor.
- Durch die Nutzung einer VTL als Ziel für die zu sichernden Server werden die Leistungsanforderungen an das Netzwerk auf dieser Strecke entschärft.
- Das für die Durchführung der Datensicherung erforderliche Zeitfenster hängt von der Geschwindigkeit der lokalen VTL ab und liegt in der Regel unter dem eines physischen Bandlaufwerks.

Nachteile

- Da die Datensicherungen in einem nachgelagerten Prozess über die VTLs auf den Remote-Standort dupliziert werden, liegen die zwei identischen Datensicherungen nicht simultan an den beiden Standorten vor.

Abbildung 12 veranschaulicht dieses Szenario.

3.2 Szenario 2: Parallel auf VTL

Szenario 2 basiert auf einer Datensicherungs-Software, die in der Lage ist, innerhalb eines Backup-Jobs zwei verschiedene Backup Ziele gleichzeitig zu adressieren. Einerseits ist dies die VTL des eigenen Rechenzentrums, andererseits wird die VTL des Ausweichrechen-

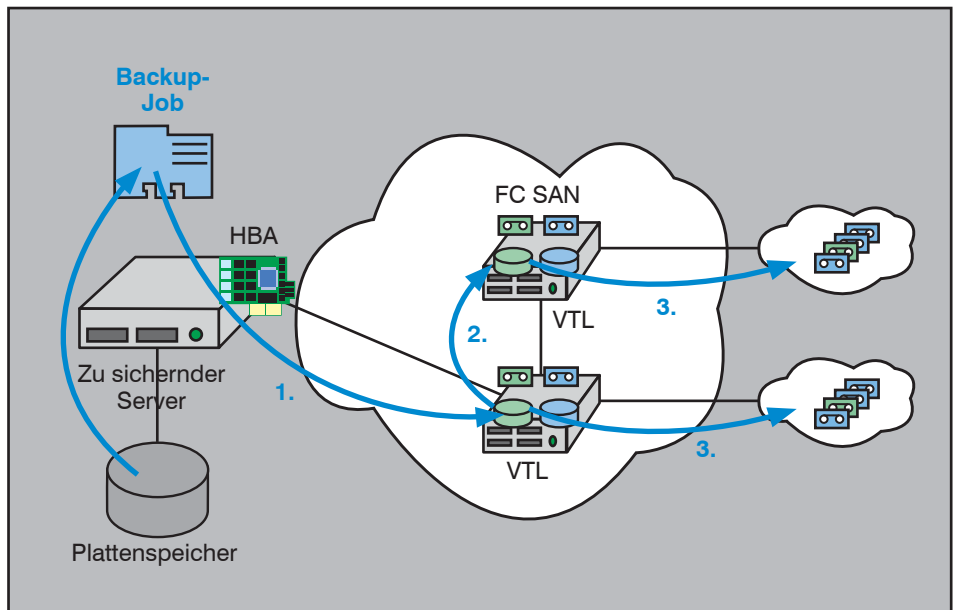


Abbildung 11: Replizierung von Datensicherungen über VTLs

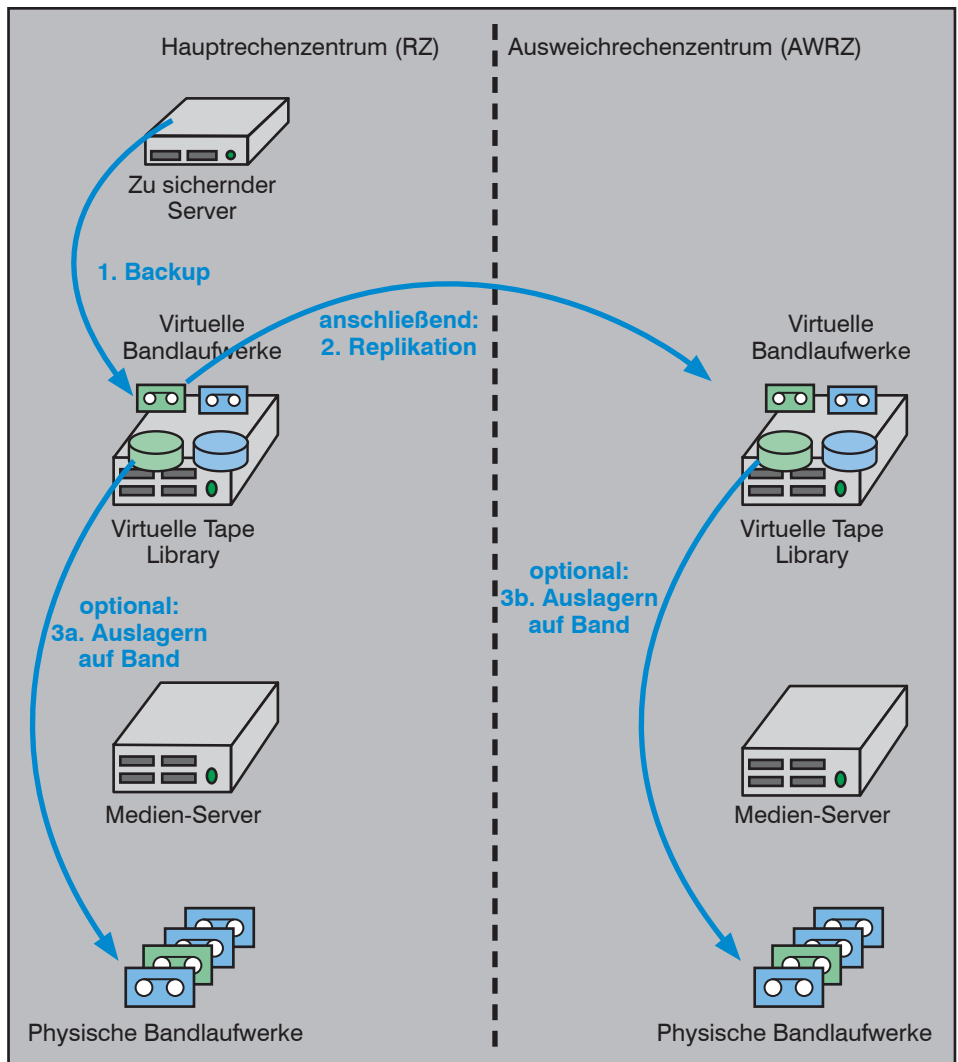


Abbildung 12: Szenario 1, sequentielle Datensicherung mittels VTLs

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

zentrums adressiert. Auch hier können bei Bedarf die virtuellen Bänder in einem nachgelagerten Prozess über einen Medien-Server auf physische Bandlaufwerke am gleichen Standort ausgelagert werden. Dieses Szenario basiert auf der in Abschnitt 2.2 dargestellten Variante.

Anforderungen

- Die Netzverbindung zwischen den RZ-Standorten muss eine der Server- und Laufwerksgeschwindigkeit angepasste Datenübertragungsrate gewährleisten.
- Die Netzwerkschnittstellen der Server müssen für die simultanen Datensicherungen die erforderlichen Netzwerk-IOs sowie eine der Laufwerksgeschwindigkeit angepasste Datenübertragungsrate gewährleisten.

Vorteile

- Durch die Nutzung einer VTL als Ziel für die zu sichernden Server besteht keine Anforderung hinsichtlich einer gleichmäßigen Datenübertragungsrate an das Netzwerk auf dieser Strecke.
- Da die simultanen Datensicherungen ihren Ursprung im gleichen Backup-Job haben, besitzen sie den gleichen Zeitstempel.

Nachteile

- Durch die simultane Durchführung der Datensicherungen werden die Anforderungen an Netzwerk-IOs und Datenübertragungsrate auf Seiten der Server-Schnittstelle verdoppelt.
- Die Größe des erforderlichen Backup-Fensters hängt von der Dauer des langsamsten der beiden simultanen Datensicherungen ab. Dies wird aufgrund der Weitverkehrsübertragung in der Regel die Sicherung in den Remote-Standort sein.

Abbildung 13 veranschaulicht dieses Szenario.

3.3 Szenario 3: Kombination aus Szenario 1 und 2

Szenario 3 kombiniert die Vorteile der Szenarien 1 und 2, um insbesondere im SAP-Umfeld den Charakter von Oracle-Datenbanken für eine möglichst effiziente Form der Datensicherung zu nutzen. Zunächst wird die vollständige Datenbank gemäß Variante 1 in regelmäßigen Abständen als Full-Backup auf die lokale VTL gesichert und von dort in einem nachgelagerten Prozess in das Ausweichrechenzentrum repliziert.

Dem gegenüber werden die Archive-

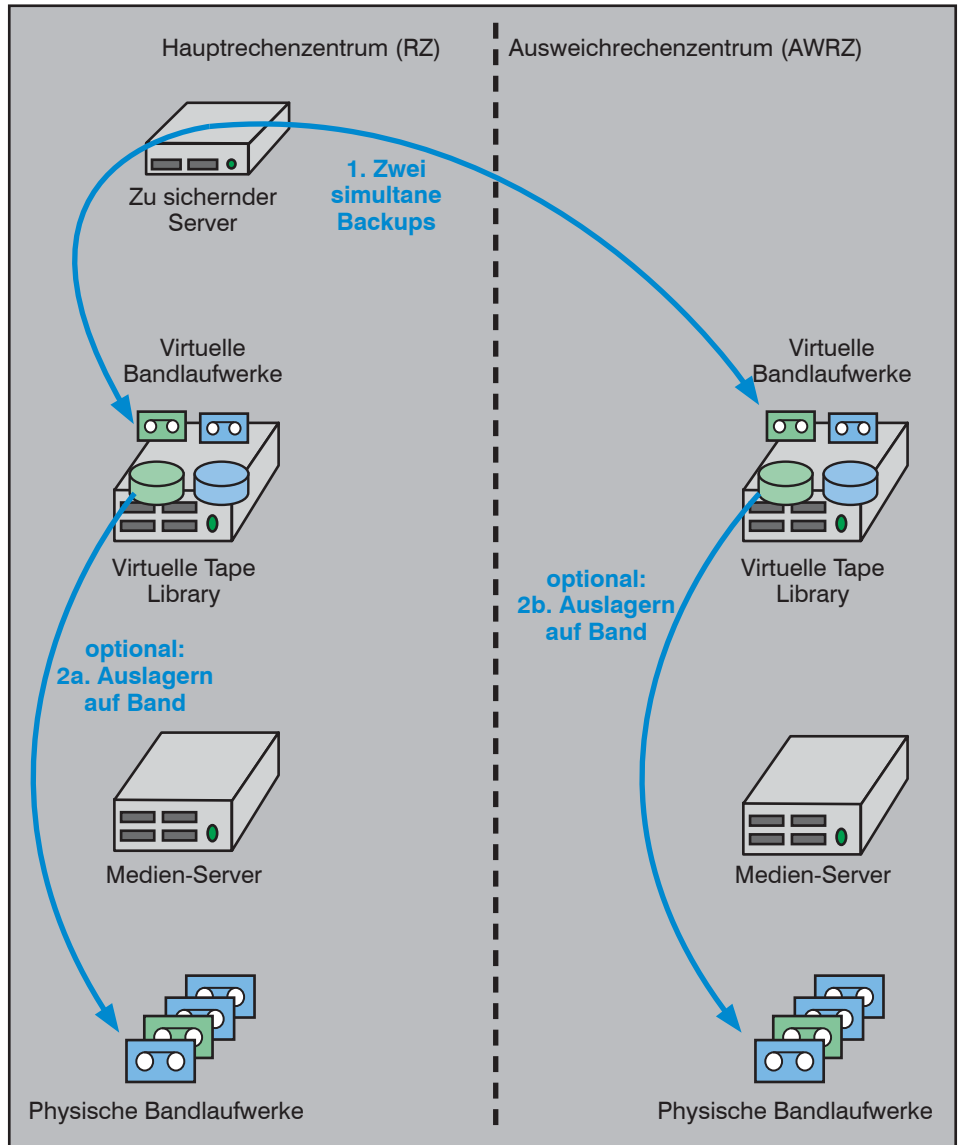


Abbildung 13: Szenario 2, parallele Datensicherung mittels VTLs

Logs gemäß Variante 2 auf die beiden VTLs simultan geschrieben. Da die Archive-Logs von der Größe her deutlich kleiner sind als die Datenbank, ist dieses Verfahren sehr effizient. Gleichzeitig haben die Archive-Logs eine sehr hohe Bedeutung, was die Verfügbarkeit der Daten angeht, da mit ihrer Hilfe sämtliche Zwischenstände seit der letzten Datenbanksicherung rekonstruiert werden können.

Anforderungen

- Die zu sichernden Daten müssen wie im Fall von SAP-/Oracle-Datenbanken unterscheidbar sein in „vollständige Datenbank“ und „Archive-Logs“, um unterschiedliche Backup-Strategien definieren zu können.

- Alternativ müsste die Backup-Software

unterschiedliche Varianten für „Full-Backups“ und „Incremental Backups“ unterstützen.

Vorteile

- Für die großen Datenvolumen der Full-Backups wird die asynchrone Replikation eingesetzt.
- Für die kleinen Datenvolumen der Archive-Logs wird die synchrone Replikation eingesetzt.

Nachteile

- Die Backup-Strategie wird komplexer, da unterschiedliche Varianten in Abhängigkeit vom Datensicherungstyp verwendet werden.

Abbildung 14 veranschaulicht dieses Szenario.

Storage: Backup von SAP-Umgebungen mit virtuellen Tape Libraries

4. Fazit

Da neben den allgemeinen Daten auch das Wachstum unternehmenskritischer Daten exponentiell zunimmt, kommt dem Datensicherungskonzept im Rahmen eines Disaster Recovery Plans eine sehr hohe Bedeutung zu. Immer mehr Unternehmen setzen auf SAP in Kombination mit Oracle Datenbanken als Kern für ihre „Business Critical Applications“.

Insofern ist es von enormer Bedeutung, frühzeitig - idealerweise schon in der Planungsphase - die richtige Kombination aus organisatorischen sowie technischen Maßnahmen zu wählen, um im Rahmen einer Backup-Strategie die hohen Verfügbarkeitsanforderungen an diese Daten zu gewährleisten.

Hier bietet die Kombination einer zeitlich nachgelagerten Verteilung der Vollsicherung einer Datenbank mit der parallelen Sicherung der Archive-Logs mit den Veränderungen an dieser Datenbank ein hohes Effizienzpotential. Ist gleichzeitig das Ziel der Datensicherungen eine virtuelle Tape Library, wird jede einzelne Datensicherung nicht nur beschleunigt, sondern insbesondere flexibilisiert.

Da für die Datensicherungen letztlich auf Bordmittel von SAP und Oracle in Form der BRTOOLS und RMAN zurückgegriffen wird (gesteuert von einer Dritthersteller-Software), sollten die folgenden Ansätze verfolgt werden:

- Regelmäßige Vollsicherungen der Datenbank, optimiert durch „Block Change Tracking“
- Regelmäßige Sicherung der Änderungen in Form der Archive-Logs

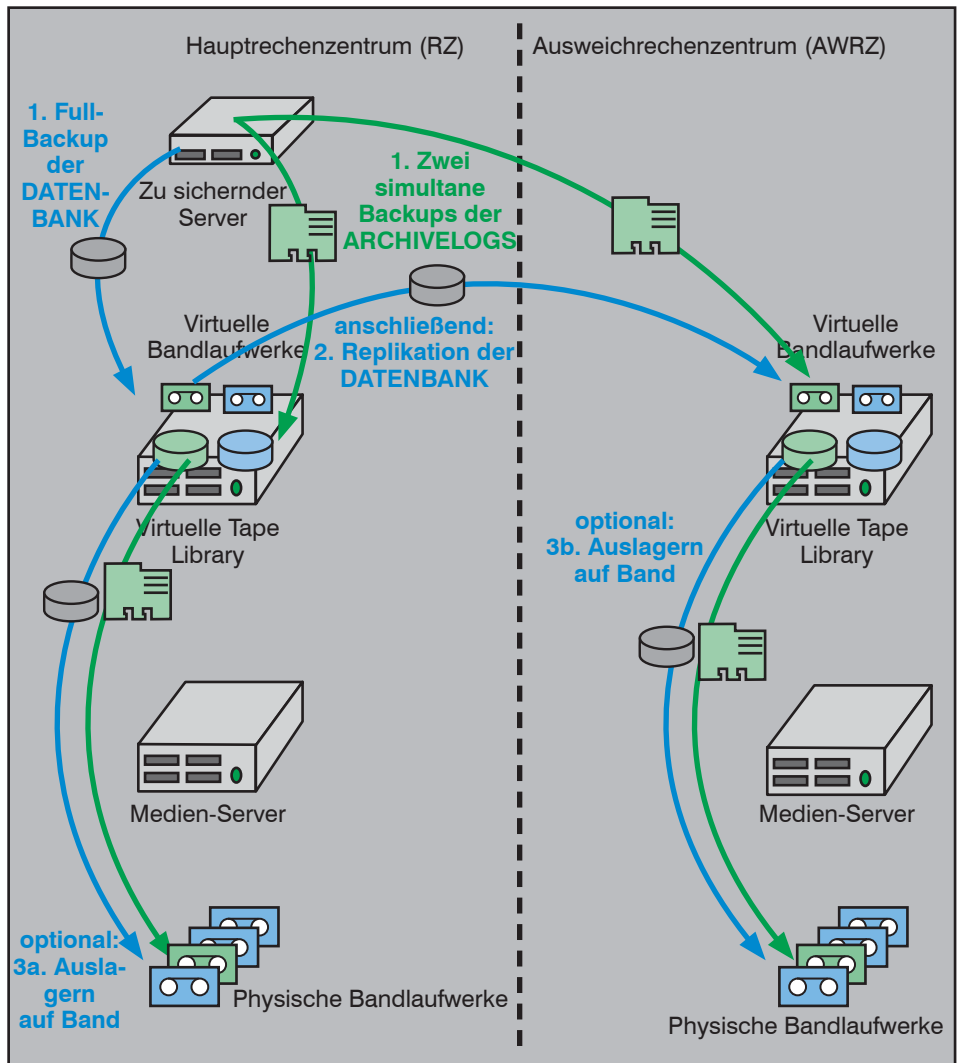


Abbildung 14: Szenario 3, Full-Backup per sequentieller Datensicherung, ARCHIVELOGS per paralleler Datensicherung

Die Häufigkeiten für diese beiden Backup-Aspekte hängen letztlich einerseits vom Datenvolumen, andererseits vom zur Verfügung stehenden Backup-Zeitfenster sowie den Anforderungen der jeweiligen Fachbereiche ab.

Jetzt Leser werden

Der Netzwerk Insider

Der Netzwerk Insider erscheint 12 Mal im Jahr im PDF-Format und informiert Sie per eMail über die Hintergründe aktueller Netzwerk-Technologien. Jeden Monat werden zwei Themen gewählt, über die in ausführlicher Form topaktuelle Insider-Informationen gegeben werden. Der Netzwerk-Insider vertritt die Sichtweise von Technologie-Anwendern und bewertet Produkte und Technologien im Sinne der wirtschaftlichen und erfolgreichen Umsetzbarkeit in der täglichen Praxis. Durch seine strenge wirtschaftliche Unabhängigkeit (keine Hersteller-Anzeigen) kann er es sich leisten, Schwachstellen und Nachteile offen anzusprechen. Der Netzwerk-Insider ist bekannt für seine kritische, herstellereutrale und fundierte Technologie-Bewertung.

Hier können Sie sich zum Netzwerk Insider kostenlos und ohne jede Verpflichtung registrieren lassen: <http://www.comconsult-akademie.de/de/Registrierung.php>

