

Allzeit bereit: Hochverfügbarkeit mit ORACLE 10g

A. Held / DOAG Regionaltreffen Rhein-Main

Inhalt

1	Übersicht Downtimekategorien und Oracle Technologien	2
2	Darstellung der Verfahrensweisen.....	4
2.1	Redundante Systeme.....	4
2.1.1	Oracle Failover Cluster (Cold Standby)	4
2.1.2	Oracle Standby-Datenbanken (Warm Standby)	7
2.1.3	Oracle Real Application Clusters	9
2.1.4	Geo Cluster.....	10
2.2	Online Technologien im Oracle-Umfeld.....	11
2.2.1	Online Reorganisation	11
2.2.2	Online Redefinition	12
2.2.3	Dynamic Reconfiguration	12
2.2.4	Automatic Storage Management.....	12
3	Schnelle Wiederherstellung im Fehlerfall	13
	Fast Recovery und RMAN	13
3.1.1	Flashback	13
4	Wiederanlaufzeiten der einzelnen Verfahren	15
5	Resümee	16

Abbildungen

Abbildung 1:	Downtimekategorien und korrelierende Oracle Technologien	3
Abbildung 2:	Schematische Darstellung eines Failover Clusters.....	5
Abbildung 3:	Schematische Darstellung einer Konfiguration mit Standby-Datenbanken	7
Abbildung 4:	Schematische Darstellung des Real Application Clusters	9

Tabellenübersicht

Tabelle 1:	Oracle Technologieübersicht inkl. geschätzter Wiederanlaufzeiten	14
Tabelle 2:	Oracle Verfügbarkeitstechnologien.....	15

Allzeit bereit: Hochverfügbarkeit mit ORACLE 10g

Die Auswahl und Implementierung einer geeigneten Verfügbarkeitsstrategie ist eine schwierige Aufgabe. Sie erfordert einen guten Überblick über mögliche Architekturen und einiges an Hintergrundwissen. Administratoren, Entwickler und Architekten mit diesem Know How sind noch immer rar und infolge dessen sehr gefragt! Zahlreiche Wege eröffnen die Möglichkeit, erhöhte Verfügbarkeitsanforderungen umzusetzen. Je nach Anforderung der Anwendung kann ein gänzlich anderer Weg sinnvoll sein. In diesem Artikel werden die wesentlichen Grundzüge jener Technologien dargestellt, die (Hoch-)Verfügbarkeit im Oracle-Umfeld gewährleisten.

Weitere Informationen, die dieses Papier ergänzen, finden Sie unter www.oracle-10g.de bzw. www.oracle-grid.de/verfuegbarkeit.

1 Übersicht Downtimekategorien und Oracle Technologien

Alle der in den nächsten Abschnitten dargestellten Oracle Technologien adressieren bestimmte Typen von Ausfallzeiten:

- Oracle Real Application Clusters sowie Failover Cluster sichern den Ausfall eines Rechners ab, der Datenbankdienste zur Verfügung stellt.
- Standby-Datenbanken, die in einem entfernten Rechenzentrum bereitgehalten werden, gewährleisten auch bei so genannten „Site Fehlern“ die Verfügbarkeit eines Datenbankservers. Der Ausfall eines Rechners, des Storage aber auch des gesamten Rechenzentrum durch ein Desaster wie Brand, Feuer oder Hochwasser kann damit abgesichert werden.
- Online Reorganisation und Dynamic Reconfiguration erlauben Wartungsarbeiten am laufenden Produktionssystem ohne Restart des Rechners oder der Datenbank Dienste.
- Flashback federt Benutzerfehler ab.
- Automatic Storage Management den Ausfall eines Festplatten Devices.
- Fast Recovery schließlich gewährleistet in Notfällen eine zumindest relativ schnelle Systemwiederherstellung.

Abbildung 1 auf Seite 3 zeigt die Technologien, die Oracle bietet, um den jeweiligen Ausfalltyp abzufangen und eine ununterbrochene Verfügbarkeit des Datenbank Servers zu gewährleisten.

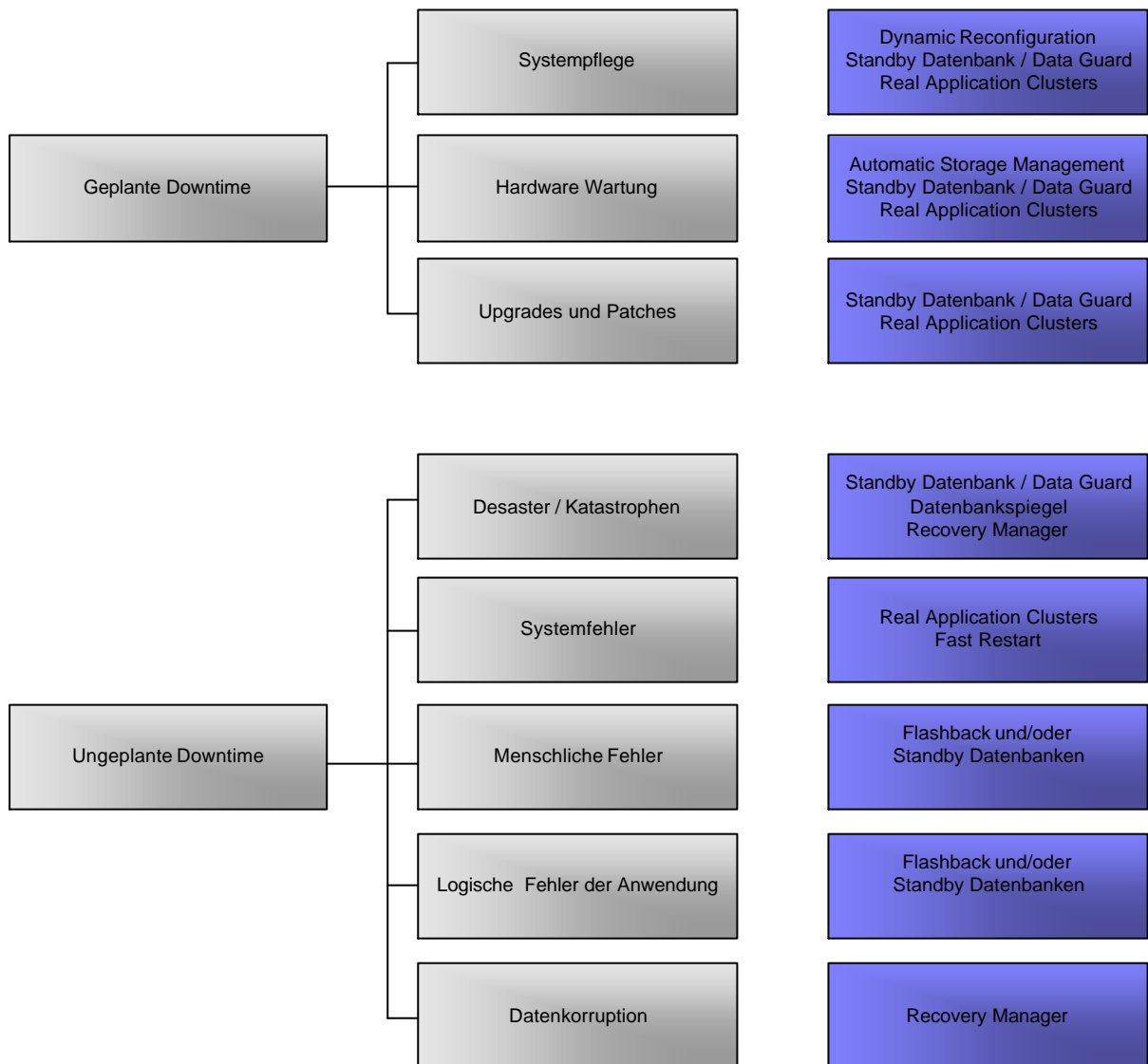


Abbildung 1: Downtimekategorien und korrelierende Oracle Technologien

2 Darstellung der Verfahrensweisen

Die folgende Darstellung der Verfahrensweisen gliedert sich in:

- Redundante Systeme
- Online Technologien im Oracle-Umfeld
- Schnelle Wiederherstellung

2.1 Redundante Systeme

Unter einem redundanten Rechnersystem ist ein Rechnerverbund zu verstehen, der im Fehlerfall eines Servers die Übernahme der Funktion durch einen Ersatzrechner gewährleistet. Fällt ein Datenbankserver aus, so übernimmt ein anderer dessen Aufgaben. Die Übernahmezeit kann dabei stark variieren: Oracle Real Application Cluster kann aufgrund seiner Aktiv/Aktiv-Architektur in bestimmten Konfigurationen in Sekundenschnelle übernehmen, während ein Cold Standby System in einem Failover Cluster durchaus auch zwanzig Minuten für die Übernahme benötigen kann. Auf den nächsten Seiten werden diese Architekturen und deren Übernahmezeiten genauer dargestellt.

Rechnerredundanz kann auf unterschiedlichen Wegen hergestellt werden. Die wichtigsten Architekturen für Oracle-Datenbankserver sind:

- Failover Cluster (Cold Standby)
- Standby-Datenbanken (Warm Standby)
- Oracle Real Application Clusters (Hot Standby)
- Geo Cluster
- Datenbankspiegel
- Replikationsverfahren

Diese werden in den folgenden Abschnitten genauer dargestellt.

2.1.1 Oracle Failover Cluster (Cold Standby)

Als *Cold Standby* bezeichnet man Systeme, deren Standby Rechner solange nicht im Betrieb ist, bis das Primärsystem ausfällt. Die Server sind redundant ausgelegt. Der Ersatzrechner ist dabei im Normalbetrieb deaktiviert, startet aber dann automatisch, sobald das Primärsystem nicht mehr verfügbar ist. Die Erkennung der Verfügbarkeit des Primärsystems erfolgt automatisch, meist über vorgefertigte Oracle Agents. In vielen Konfigurationen prüfen diese lediglich das Vorhandensein der Oracle Prozesse.

Jedoch ist dies nicht unbedingt aussagekräftig: Es gibt Fälle, in denen alle Prozesse laufen, aber dennoch nicht korrekt arbeiten. Einige Hersteller versuchen daher einen Connect zur Datenbank-Instanz, gelegentlich auch eine Transaktion wie ein Insert in eine für den Agenten vorgefertigte Tabelle oder die Abfrage der Systemzeit über ein SQL-Statement wie „select sysdate from dual“. Schlägt diese Aktion fehl, wird nach einer vorgefertigten Prozedur verfahren und der Failover eingeleitet.

Typischer Vertreter im Oracle Umfeld ist der *Failover Cluster*. In diesem ist entweder Rechner A oder Rechner B aktiv. Ressourcen wie Storage, logische IP-Adressen und Hostnamen, aber auch Dienste wie Oracle Instanzen und Oracle Listener werden in einem Failover Cluster in so genannten Ressourcengruppen verwaltet. Der jeweils aktive Rechner nutzt die Ressourcengruppen exklusiv. So kann zum Beispiel der Storage nur von maximal einem Rechner gemountet werden.

Ob der jeweils andere Rechner aktiv ist, wird in einem Failover Cluster über den Cluster Heartbeat überprüft. Beide Rechner sind hierfür mit einem privaten, exklusiv für den Cluster genutzten Netzwerk verbunden. Über dieses wird der Heartbeat gesandt, der dem jeweils anderen Rechner das Signal gibt, dass das eigene System funktionsfähig, also „am Leben“ ist. Verstummt der Heartbeat des Primärrechners, so geht der Standby Knoten davon aus, dass er den Betrieb übernehmen muss. Dazu übernimmt er die Ownership der Clusterressourcen. Er mountet den Storage, übernimmt logische IP-Adressen und Hostnamen, startet die Oracle Dienste und stellt damit die Betriebsfähigkeit des Datenbank Backend Systems wieder her. Je nach Ausfallursache und Größe der Datenbank dauert dies zwischen wenigen Minuten und etwa einer halben Stunde.

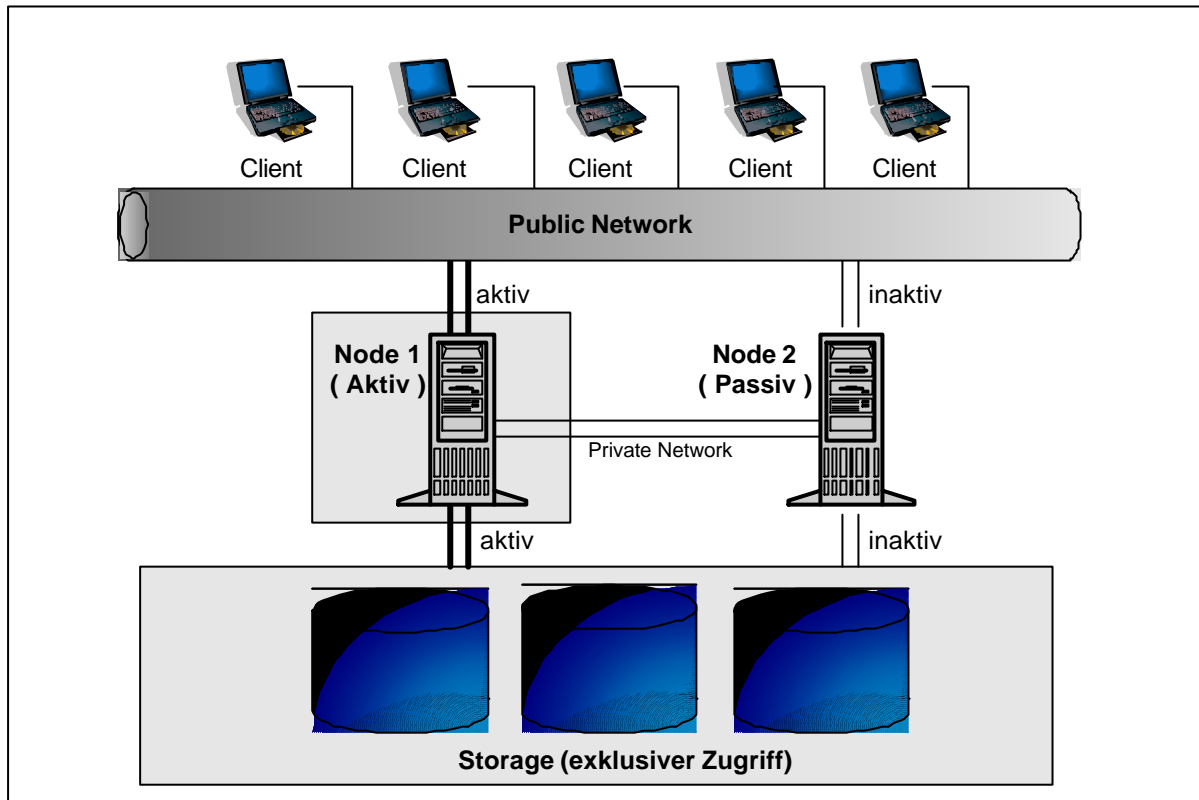


Abbildung 2: Schematische Darstellung eines Failover Clusters

Damit eine Übernahme der Funktionalitäten möglich ist, müssen beide Rechner einen gemeinsamen Zugriff auf einige Ressourcen haben. So müssen beide Rechner den Storage, auf dem die Datenbankdateien liegen, mounten können. Daher muss für den Failover Cluster ein externes Storage-System verwendet werden, an das mehrere Hosts angeschlossen werden können. Möglich ist dies bei Nutzung von *SCSI-Storage*, durch Einsatz eines *SAN (Storage Area Networks)*, *iSCSI (IP-basiertes SCSI)* oder *NAS (Network Attached Storage)*. Erfahrungsgemäß ist die Nutzung von *SCSI-Storage* und *SAN-Technologie* unbedingt vorzuziehen.

Die Verbindung zwischen den beiden Rechnern des Failover Cluster wird über das Private Network Interface realisiert. Zum Einsatz kommen hier häufig einfache Ethernet Netzwerkkarten. Für eine schnellere Verbindung werden oft auch proprietäre Lösungen des Hardwareherstellers verwendet. Das Private Network sollte unbedingt - wie andere Komponenten auch – redundant ausgelegt werden. Zwischen beiden Rechnern sollten mindestens zwei Netzwerkschnittstellen implementiert werden, so dass auch das Private Network keinen *SPOF (Single Point of Failure)* bildet und bei Ausfall einer Netzwerkkarte die jeweils andere die Aufgaben übernehmen kann.

Fallen doch einmal alle Verbindungen des Private Network aus, besteht die Gefahr, dass beide Knoten die Ressourcen übernehmen wollen. Gerade der konkurrierende Zugriff auf die

eigentlich exklusiv zu nutzenden Ressourcen wie den Storage kann zu erheblichen Datenkorruptionen führen. Aufgefangen wird dies in der Regel durch den Mechanismus eines Quorum Device bzw. einer Voting Disk.

Diese Verfahrensweise, bei der eine Festplattenpartition für die Kommunikation der Cluster-Rechner genutzt wird, ist im Grunde recht einfach: Auch wenn im Cluster über keines der Network Interfaces Kommunikation möglich ist, besteht immer noch Kontakt zum gemeinsam angeschlossenen Storage. Kann ein Rechnerknoten auch darauf nicht mehr zugreifen, so kann er die Datenbank sowieso nicht mehr mounten und ist betriebsunfähig. In diesem Fall stellt er auch keine Gefahr dar: Ein unkontrollierter konkurrierender Zugriff beider Knoten auf den exklusiv zu nutzenden Storage ist in diesem Fall ausgeschlossen. Können aber beide Knoten auf den Storage zugreifen, so können sie auch hierüber kommunizieren. Nur jener Knoten, der über die Voting Disk die Ownership erlangt, übernimmt die entsprechenden Clusterressourcen und stellt die Betriebsfähigkeit wieder her. Der andere Knoten fährt automatisch herunter.

Das Private Network wird exklusiv für die Kommunikation im Cluster verwendet. Keinesfalls sollte dieses mit Funktionalitäten des Public Network, das für den Client-Zugriff reserviert ist, belastet werden. Hält man sich nicht daran, kann es zu nicht unerheblichen Seiteneffekten kommen. So kann das Network Interface den Cluster Heartbeat bei starkem Netzwerk Traffic nicht korrekt oder nicht rechtzeitig weiterleiten. Timeout-Zeiten werden möglicherweise überschritten. Der Effekt: Der zweite Rechner im Cluster nimmt an, dass der Primärknoten ausgefallen ist, und versucht die Betriebsfunktionalität zu übernehmen.

Die Verbindung zu Clients wie zu den Applikationsservern erfolgt über das Public Network. Dieses bildet die Netzwerkschnittstelle zur Außenwelt. Auch hier sollten pro Rechner mindestens zwei Netzwerkkarten zum Einsatz kommen, um einen SPOF zu vermeiden. Diese werden in das Unternehmensnetz eingebunden und erlauben den Zugriff auf den Failover-Cluster von außen.

Cold Standby Systeme benötigen einen gesonderten Cluster Layer, der nicht von Oracle geliefert, sondern von Drittanbietern eingekauft werden muss. So bietet Sun den Sun Cluster für Solaris. Unter AIX kann HACMP, für HPUX der HP MC/ServiceGuard eingesetzt werden. Veritas bietet den Veritas Cluster für verschiedene Plattformen wie Sun Solaris, HPUX und Linux an.

Notwendige Komponenten eines Cold Standby Systems sind:

- Zwei Server mit mindestens je zwei Private und zwei Public Network Interfaces.
- An beide Server angeschlossener Storage (externes SCSI, SAN, iSCSI, NAS).
- Cluster Management Software wie Sun Cluster, Veritas Cluster, HACMP oder ServiceGuard.

Für Planung und Implementierung sind neben Datenbank und Netzwerk Know How auch Kenntnisse der Cluster Management Software nötig. Die Oracle Datenbank selbst wird ähnlich aufgesetzt wie ein normales Single Instance System und anschließend über Ressourcengruppen in den Cluster eingebunden.

Vorteile eines Oracle Failover Clusters:

- Übernahme der Funktionalität bei Ausfall eines Servers ohne administrativen Eingriff
-

Nachteile eines Oracle Failover Clusters:

- Erhöhter Zeitaufwand (im Vergleich zu Warm und Hot Standby) für den Wiederanlauf
- Bei Ausfall eines Rechenzentrums (zum Beispiel durch einen Brand) fällt der gesamte Cluster aus.
- Redundante Hardware kann im Normalbetrieb nicht genutzt werden.

2.1.2 Oracle Standby-Datenbanken (Warm Standby)

Auch in einem Warm Standby System sind die Datenbankserver redundant ausgelegt. Anders als beim Cold Standby ist hier jedoch der zweite Rechner bereits betriebsbereit und kann daher bei einem Ausfall des Primärserver die Funktionalität schneller übernehmen. Typische Vertreter im Oracle Umfeld sind replizierte und Standby-Datenbanken.

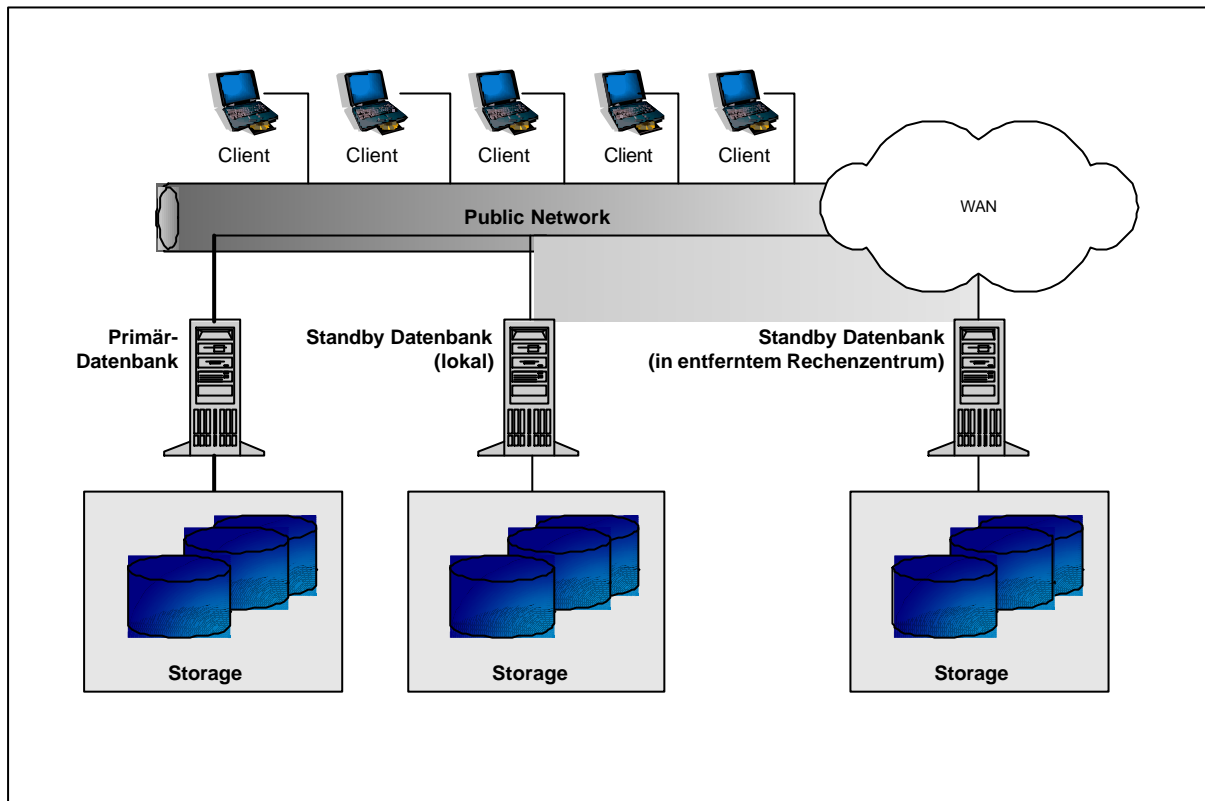


Abbildung 3: Schematische Darstellung einer Konfiguration mit Standby-Datenbanken

In einer Standby-Datenbank wird eine transaktionskonsistente Kopie der Primärdatenbank auf einem zweiten vollständigen System gehalten. Primär- und Standby System haben dabei in der Regel ihren eigenen privaten Storage. Initial wird die Primärdatenbank einmalig über ein Backup auf das Standby System übertragen. Anschließend werden alle Transaktionen des Primärsystems auf das Standby System repliziert. Die Replikation kann in Echtzeit oder auch zeitverzögert erfolgen.

Zeitverzögerte Verarbeitung wird häufig zum Abfangen von Benutzerfehlern eingesetzt. Wird in einer Standby-Datenbank eine Zeitverzögerung von 30 Minuten implementiert, so werden auch Benutzerfehler erst nach diesem Zeitraum appliziert. Löscht ein Benutzer „versehentlich“ Daten, so sind diese Daten auf dem Standby-System noch 30 Minuten verfügbar. Erst dann wird die Transaktion auch auf dem Standby-Rechner ausgeführt. So können die Daten im Bedarfsfall recht schnell aus dem Standby-System rekonstruiert werden. Allerdings hat diese Zeitverzögerung auch einen Nachteil: Fällt das Primärsystem aus, so müssen die Transaktionen der letzten 30 Minuten auf dem Standby System nachgefahren werden, bevor dieses die Betriebsfähigkeit wiederherstellen kann. Die hierfür notwendige Zeit ist für manche Anwendungen jedoch nicht akzeptabel. Der Ausfall eines Systems ist nicht nur Flughäfen und Krankenhäuser, sondern auch für das Kassensystem eines Warenhauses an einem Samstag kurz vor Weihnachten eine teure Angelegenheit.

Ab Oracle 10g ist es nicht mehr notwendig, ein solches Delay einzusetzen, um Benutzerfehler abfangen zu können. Vielmehr kann hier die Oracle Flashback Technologie eingesetzt werden. Diese erlaubt ein vorübergehendes Zurücksetzen einer Datenbank in

einen früheren konsistenten Zustand. Flashback lässt sich auch auf einer Standby-Datenbank ausführen, um von hier aus die Daten über einen Datenbank Link zurück zur Primärdatenbank zu übertragen. Abschließend kann die Standby-Datenbank durch Wiederholung aller Transaktionen des Primär-Systems wieder auf den aktuellen Zustand der Primärdatenbank vorgerollt werden.

Standby-Datenbanken können auch als Disaster Recovery Strategie eingesetzt werden. Man kann Standby-Datenbanken lokal oder auch in entfernten Rechenzentren betreiben. Beides lässt sich auch innerhalb einer Standby Konfiguration miteinander verbinden: Eine Standby-Datenbank kann beispielsweise lokal, eine zweite entfernt betrieben werden. Dabei sind sehr große Distanzen – auch über Kontinente hinweg – möglich. Bei Ausfall eines der Rechenzentren kann der Betrieb weiterhin aufrechterhalten werden. Die Daten können auf diesem Wege über eine große Distanz redundant gespeichert werden.

Es ist möglich, bis zu neun Standby-Datenbanken an einer Primärdatenbank zu betreiben. So bietet es sich in manchen Umgebungen an, zwei Standby-Datenbanken zu verwenden. Dies gilt insbesondere bei Verwendung spezieller Protection Modes, die garantieren, dass kein Datenverlust entsteht.

Komponenten für den Einsatz von Standby-Datenbanken sind:

- Ein Primärsystem und ein Standby System (jeweils mit eigenem Storage)
- Eine Netzwerkverbindung zwischen Primär- und Standby System

Vorteile einer Standby-Datenbank:

- Bei Verwendung von Oracle Data Guard: Übernahme der Funktionalität bei Ausfall eines Servers ohne administrativen Eingriff
- Bei Ausfall eines Rechenzentrums kann unter Verwendung einer entfernten Standby-Datenbank der Betrieb weiter aufrechterhalten werden
- Disaster Schutz durch Einsatz eines Remote Rechenzentrums möglich
- Schutz vor Benutzerfehlern (abhängig von der Verzögerungsdauer bzw. bei Einsatz von Flashback Database auf der Standby-Seite)

Nutzung der Standby Database für das Reporting

Nachteile einer Standby-Datenbank:

- Übernahme der Betriebsfähigkeit durch Standby benötigt etwas mehr Zeit als bei Hot Standby
- Erhöhte Netzlast durch Übertragung der Redo Log Informationen zwischen Primär- und Standby System. Eventuell sollte zusätzliche Bandbreite zur Verfügung gestellt werden.

2.1.3 Oracle Real Application Clusters

Bei Nutzung des Oracle Real Application Clusters sind im Normalbetrieb zwei oder mehrere Rechnerknoten im Cluster aktiviert. Diese können auch für verteiltes Rechnen im Oracle-Grid genutzt werden. Der Vorteil: Fällt ein Rechner aus, so können Clients unmittelbar und ohne Wiederanlaufzeit auf einen verbleibenden Rechnerknoten konnektieren. Langlaufende Operationen können zudem auch über den gesamten Cluster parallelisiert werden. So ist neben der erhöhten Verfügbarkeit auch das Thema Skalierung im Oracle 10g Grid eine der Stärken des Real Application Cluster. Redundante Hardware wird nicht nur im Fehlerfall eingesetzt, sondern kann – anders als bei Failover Clustern - im Normalbetrieb genutzt werden.

Ähnlich wie in einem Oracle Failover Cluster wird auch hier ein Private Network für die Clusterkommunikation sowie ein davon separiertes Public Network benötigt. Allerdings wird im Oracle Real Application Cluster konkurrierend auf die Datenbank Dateien eines Shared Storages zugegriffen. Dies erfordert zusätzlichen Kommunikationsaufwand zwischen den Knoten, der zur Abstimmung von Sperrungen und veränderten Block Images nötig ist. In Real Application Cluster Umgebungen ist es daher wichtig, einen möglichst schnellen privaten Cluster Interconnect zu nutzen.

Mindestens Gigabit Ethernet sollte eingesetzt werden. Besser noch sind spezielle, oft proprietäre Lösungen mit höherer Bandbreite. Beispiele sind Memory Channel (alphabasierte HP Cluster), Myrinet (Linux Systeme), Scalable Coherent Interconnect (SUN), Veritas LLT (verschiedene Plattformen) oder HP Hyper Fabric HMP.

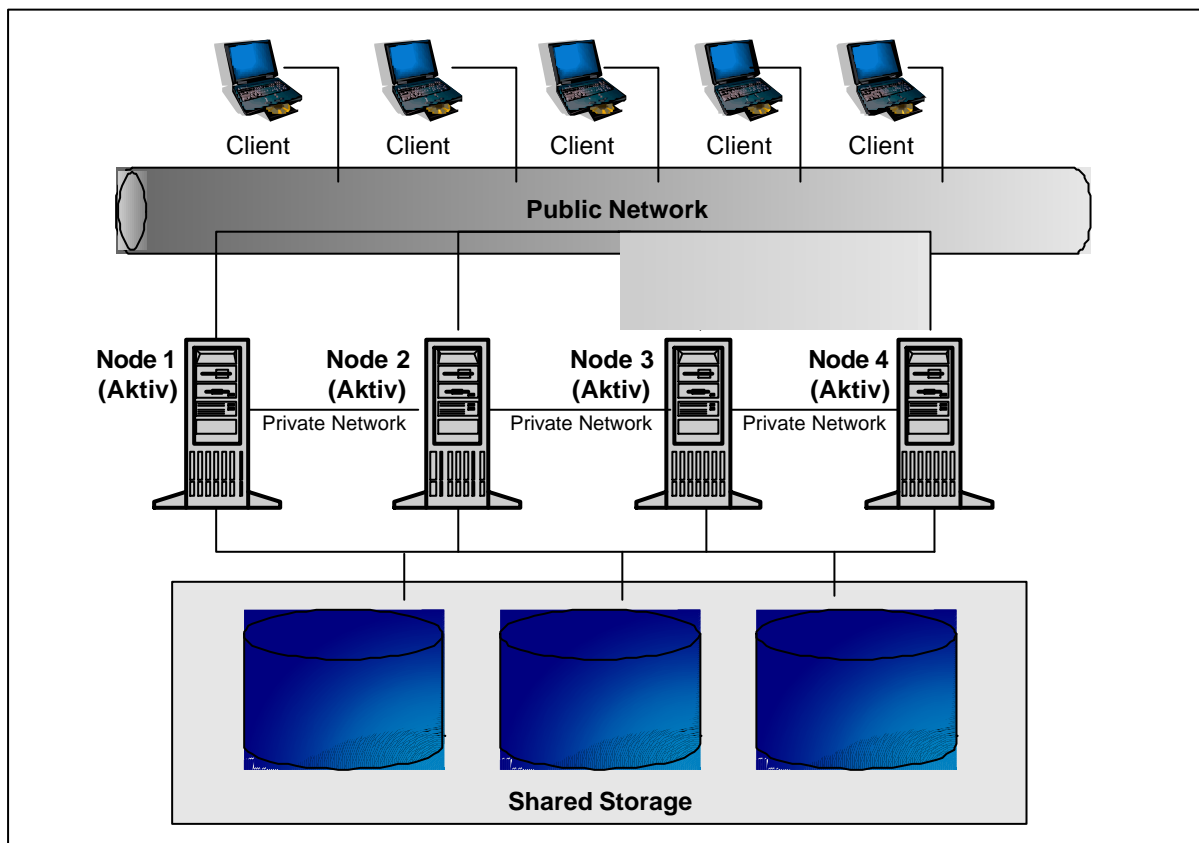


Abbildung 4: Schematische Darstellung des Real Application Clusters

Notwendige Komponenten für den Einsatz eines Oracle Real Clusters sind:

- Shared Storage für alle beteiligten Cluster Nodes (konkurrierender Zugriff)
- Verwendung von Cluster Filesystem, ASM oder Raw Devices auf dem Shared Storage
- Für jeden Cluster Node mindestens je zwei Private und zwei Public Network Interfaces
- Oracle Cluster Ready Services oder Cluster Management Software eines Drittherstellers.

Vorteile eines Oracle Real Application Clusters:

- Übernahme der Funktionalität bei Ausfall eines Servers ohne administrativen Eingriff
- Kein Zeitaufwand für Wiederanlauf, Client Reconnect direkt ohne Zeitversatz auf den Ersatzknoten
- Redundante Hardware kann auch während des Normalbetriebs genutzt werden.
- Parallelisierung über alle Clusterknoten hinweg realisierbar
- Dynamische Zuordnung von Services möglich

Nachteile eines Oracle Real Application Clusters:

- Bei Ausfall eines Rechenzentrums (zum Beispiel durch einen Brand) fällt der gesamte Cluster aus. Dieser Fehlerfall kann bei Bedarf durch Einsatz einer zusätzlichen Standby-Datenbank in Kombination mit Oracle Real Application Clusters abgefangen werden.

2.1.4 Geo Cluster

Einige Drittanbieter wie Veritas oder Libelle bieten erweiterte Möglichkeiten der Disaster Absicherung. Geo Cluster erlauben Clustering ohne geographische Grenzen. Je nach Architektur können die Cluster Nodes in Rechenzentren stehen, die von einigen wenigen bis zu mehr als hundert Kilometer entfernt sind. Manche Hersteller bieten auch Lösungen, die Spiegel- und Replikationsmechanismen über Kontinente hinweg unterstützen. Die Architekturen der einzelnen Hersteller sind ebenso unterschiedlich wie die Implementierung.

Verfahren wie Hardware Replikation und Datenbankspiegel schützen ebenfalls vor logischen und physikalischen Ausfällen. Kernkonzept ist die Spiegelung über große Distanzen. Das Ergebnis ist von der Funktionalität her ähnlich einer Standby-Datenbank. Der Layer zur Realisierung liegt jedoch unterhalb der Datenbankmanagementschicht: Hardware Spiegel und Spiegel auf Filesystemebene sind die häufigsten Vertreter dieser Architektur.

Replikation kann auch innerhalb der Datenbank implementiert werden. Snapshot und Multimaster Replikation sowie Oracle Streams bieten hier weitere Möglichkeiten, erhöhte Verfügbarkeit umzusetzen.

Voraussetzungen zur Implementierung von Geo Clustern und Datenbank Spiegeln:

- Abhängig vom Anbieter, sehr unterschiedliche Implementierungen

Vorteile:

- Datensicherheit und Verfügbarkeit auch bei Disastern

Nachteile:

- Kosten für Lizenzen der Drittanbieter-Software
- Oftmals erhöhte Bandbreite nötig

Cluster-Technologien allein schützen also nicht vor Katastrophen wie Feuer, Wasser und Blitz. Soll ein System auch vor solchen Gefahren geschützt werden, so müssen die Cluster- und RAID-Systeme räumlich voneinander getrennt werden. Hierzu sollte ein räumlicher Abstand von mindestens mehreren hundert Metern eingehalten werden. Zudem sollte das Backup-System nicht im gleichen Gebäude bzw. Brandabschnitt stehen, so dass im Katastrophenfall mindestens dieses für die Wiederaufbau eines Ersatzsystems zur Verfügung steht.

Die räumliche Trennung ist für den Cluster noch relativ einfach realisierbar, die Mehrkosten halten sich in Grenzen. Die räumliche Trennung eines RAID-Systemes jedoch bedeutet einen erheblichen technischen Aufwand und verursacht enorme Mehrkosten. Zur Realisierung wird in der Regel ein weiteres gleiches RAID-System in einem anderen Gebäude, eine exklusiv für die Datenspiegelung genutzte Glasfaserverbindung so wie spezielle Controller und Software benötigt. Kosten hierfür gehen schnell in die Hunderttausende.

2.2 Online Technologien im Oracle-Umfeld

Für die Umsetzung eines hochverfügbaren Systems ist es nicht nur erforderlich, ungeplante Ausfallzeiten zu reduzieren. Auch Wartungsarbeiten, die ein Wartungsfenster benötigen, reduzieren die Verfügbarkeit des Systems. Oracle bietet für solche Wartungsarbeiten inzwischen zahlreiche Workarounds, die ein paralleles Arbeiten der Anwender erlauben. Hierzu zählen insbesondere *Online Reorganisation*, *Online Redefinition* und *Dynamic Reconfiguration*.

2.2.1 Online Reorganisation

Sind Daten in der Datenbank stark fragmentiert, wird unter Umständen eine Reorganisation notwendig. Mussten früher Daten exportiert und wieder importiert werden, so kann dies heute ohne Downtime online, also im laufenden Betrieb geschehen. Trotzdem sollte man natürlich nicht Zeiten der Hochlast für eine Reorganisation wählen. Schließlich benötigen auch Online Vorgänge einige Ressourcen.

Eine Tabellenspalte umzubenennen, eine Spalte hinzuzufügen oder zu entfernen, Speicherplatz Klauseln zu verändern oder ein Datensegment von einem Tablespace in einen anderen zu verschieben, war früher eine aufwändige Angelegenheit. Online Redefinition erlaubt diese Arbeiten im laufenden Betrieb. Das klingt nach einer Selbstverständlichkeit. Doch bedenkt man, dass zum Umbenennen einer Tabellenspalte in älteren Oracle-Version eine Temporärtabelle angelegt, die Originaltabelle gelöscht, mit dem geänderten Spaltennamen erzeugt und abschließend mit den Daten aus der Temporärtabelle befüllt werden musste, so ist Online-Redefinition ein enormer Fortschritt im Hinblick auf Verfügbarkeit.

Voraussetzungen:

- Keine

Vorteile des Einsatzes von Online Reorganisation:

- Keine Downtimes
- Einfache Handhabung

Nachteile des Einsatzes von Online Reorganisation:

- Keine

2.2.2 Online Redefinition

Unter Online Redefinition ist die Änderung der Struktur eines Datenbankschemas im laufenden Betrieb zu verstehen. Hierzu zählt das Anfügen oder Entfernen von Tabellenspalten, aber auch Partitionierung von Tabellen oder die Umwandlung einer Tabelle in eine *IOT (Index Organized Table)*.

Voraussetzungen:

- Genaue Prüfung der möglichen Seiteneffekte durch die Schema-Änderung

Vorteile des Einsatzes von Online Reorganisation:

- Keine Downtimes
- Einfache Handhabung

Nachteile des Einsatzes von Online Reorganisation:

- Keine

2.2.3 Dynamic Reconfiguration

Bereits mit Oracle 9i wurde die dynamische *System Global Area (SGA)*, also ein dynamisch veränderbarer Aufbau des von der Oracle Instanz genutzten Arbeitsspeichers eingeführt. In älteren Versionen war es notwendig, die Instanz nach der Anpassung entsprechender Werte im Parameter File der Datenbank neu zu starten. Auch wenn die Downtimes hierfür meist nur wenige Minuten dauerten, so bietet Dynamic Reconfiguration durch Änderungen der SGA im laufenden Betrieb eine höhere Verfügbarkeit.

Voraussetzungen:

- Verwendung dynamischer Speicherparameter

Vorteile des Einsatzes von Dynamic Reconfiguration:

- Änderungen der Datenbank Instanz im laufenden Betrieb ohne Restart der Instanz möglich.

Nachteile des Einsatzes von Dynamic Reconfiguration:

- Keine

2.2.4 Automatic Storage Management

Mit Oracle 10g wurde Automatic Storage Management (ASM) eingeführt. Es bietet Funktionalitäten wie Disk Striping und Mirroring. Das ist im Grunde nichts Neues. Jedoch musste bisher entweder Software von Drittanbietern (Logical Volume Manager) oder aber intelligente - und damit teure - Hardware eingesetzt werden. Mit ASM bietet Oracle einen Built-In Spiegel auf Hard Disk Basis, der in Version 11 sicher ein Standard sein wird.

Voraussetzungen zur Implementierung von ASM:

- Verwendung einer ASM Instanz

Vorteile des Einsatzes von ASM:

- Mirror und Striping ohne externe Hilfsmittel innerhalb Oracle DBMS möglich.

Nachteile des Einsatzes von ASM:

Hardware RAID ist noch immer performanter.

3 Schnelle Wiederherstellung im Fehlerfall

Fast Recovery und RMAN

Ist der Ernstfall doch einmal eingetreten, wird im schlechtesten Fall die Wiederherstellung der Datenbank aus einem Backup nötig. Hier kommt der Recovery Manager zum Einsatz.

Unter Fast Recovery ist die schnellstmögliche Wiederherstellung eines Systems zu verstehen. Diese kann – je nach Fehlertyp – über ein Flashback Database, ein disk-basierendes oder ein tape-basierendes Restore mit Recovery.

Oracle 10g bietet mit Automatic Disk Based Backup eine wesentlich schnellere Wiederherstellungsmöglichkeit, als sie bisherige Verfahren bieten konnten. Dabei wird ein Speicherbereich, die Flash Recovery Area, genutzt, die günstigere Zugriffsmöglichkeiten als Tapes bietet. Die im Flash Recovery Area gespeicherten Backups können automatisiert auf Band archiviert werden. So liegen die neuesten Backups in Form von Files auf Disk vor, ältere Backups können – wenn auch mit mehr Zeitaufwand – von Tape wiederhergestellt werden.

Die Wiederherstellungsmöglichkeiten aus Automated Disk Based Backups decken alle Level ab: Angefangen von Block Level Media Recovery über einzelne Data Files bis hin zur gesamten Datenbank wird jede manuelle oder in *Oracle Recovery Manager (RMAN)* zur Verfügung gestellte Restore Option unterstützt.

Der Recovery verhielt sich bei seiner Einführung – wie dies häufig bei Einführung neuer Optionen der Fall ist – recht unzuverlässig. Vielen Oracle Anwendern blieb er zudem mit seiner teilweise recht kryptischen Syntax und den damals noch recht zahlreichen Bugs eher suspekt. Dieses Image ist sicherlich der Grund dafür, dass RMAN heute nicht häufiger eingesetzt wird.

RMAN gestattet inkrementelle Backups (Verringerung der Backup Fenster), schnelle Wiederherstellung defekter Datenbankbereiche, Verwendung von Disk Based Backup und Recovery sowie Reparatur von Block Corruptions im laufenden Betrieb und vieles mehr.

Oracle 10g arbeitet mit einem Change Tracking Verfahren. Hierbei wird ein Bitmap mitgeführt, das jene Blöcke kennzeichnet, die seit dem letzten Backup geändert wurden. Dies ermöglicht mit RMAN ein um etwa Faktor 20 schnelleres inkrementelles Backup. Aber auch das Recovery ist wesentlich schneller, die Bedienung über die graphischen Werkzeuge des Oracle Enterprise Manager *DBControl* und *Grid Control* zudem einfacher.

Voraussetzungen zur Implementierung:

- Keine

Vorteile des Einsatzes von RMAN:

- Zahlreiche Erweiterungen in Oracle 10g, die Backup und Recovery Time enorm verkürzen können

Nachteile des Einsatzes von RMAN:

- Bei Einsatz eines Repository: Erhöhte Komplexität sowie eigene Backup Strategie für die Repository Datenbank

3.1.1 Flashback

Oracle Flashback bietet hervorragende Wiederherstellungsmöglichkeiten bei Benutzerfehlern. Tabellen, die gelöscht wurden, können in Sekundenschnelle wiederhergestellt werden. Wurden Daten versehentlich geändert, so lassen sich diese auf Tabellenebene auf einen bestimmten Zeitpunkt zurücksetzen. Auch die gesamte Datenbank

kann auf einen früheren Zeitpunkt zurückgerollt werden. Dazu musste früher ein Backup eingespielt und anschließend „Point in Time“ reconvert werden. Heute genügt ein einfacher SQL Befehl. Möglich wird dies durch spezielle Flashback Logs.

Flashback lässt sich ebenfalls mit Standby-Datenbanken kombinieren. So kann im Fehlerfall die Standby-Datenbank zurückgesetzt werden, um anschließend den alten Datenbestand zurück in das Produktionssystem zu übertragen. Dieses kann durchgehend in Betrieb bleiben. Für manche Umgebungen wie z.B. ein Auskunftssystem (Fahrplanabfrage, Telefonauskunft, Katalogdaten und ähnliches) kann auf diese Weise ohne Downtime die Wiederherstellung versehentlich gelöschter oder veränderter Daten erfolgen. Aber auch für andere Systeme bildet Flashback eine schnelle Wiederherstellungsmöglichkeit in Situationen, in denen früher das enorm zeitaufwändige Restore der gesamten Datenbank mit anschließendem Point-in-Time-Recovery nötig war.

Voraussetzungen zur Implementierung von Flashback:

- Ausreichend großer Undo Tablespace für die Wiederherstellung auf Tabellenebene
- Archive Log Mode und Flashback Mode für die Wiederherstellung auf Datenbankebene

Vorteile des Einsatzes von Flashback:

- Sehr schnelle Wiederherstellung von Daten bei logischen Fehlern

Nachteile des Einsatzes von Flashback:

- Höherer Verbrauch an Festplattenspeicher für Undo Informationen und Flash Recovery Logs abhängig von der exakten Konfiguration

Eine Gesamtübersicht der Fähigkeiten einzelner Oracle Technologien in Bezug auf mögliche Fehlertypen und Ausfallursachen zeigt Tabelle 1.

Tabelle 1: Oracle Technologieübersicht inkl. geschätzter Wiederanlaufzeiten

Downtime Kategorie	Downtime Ursachen	Oracle Technologie	Wiederanlaufzeit
Geplant und ungeplant	Wartungsarbeiten Rechnerausfall	Cold Standby mit Failover Cluster	1-20 Minuten
Geplant und ungeplant	Wartungsarbeiten Benutzerfehler Komponentenausfall Rechnerausfall Disaster mit Rechnerverlust	Warm Standby mit Standby- Datenbanken	1-10 Minuten, bei zeitverzögerter Übertragung auch länger
Geplant und ungeplant	Wartungsarbeiten Rechnerausfall	Hot Standby mit Oracle Real Application Clusters	Wenige Sekunden bis zu 2 Minuten
Geplant und ungeplant	Wartungsarbeiten Benutzerfehler Komponentenausfall Rechnerausfall Disaster mit Rechnerverlust	GeoCluster und Datenbankspiegel	1-30 Minuten
Ungeplant	Ausfall einer Festplatte	Automatic Storage Management	Keine Ausfallzeit (Hot Swap)
Ungeplant	Medienfehler Disaster	Fast Recovery / Fast Restart	Abhängig von Ausfallgrund und wiederherzustellender Datenmenge
Ungeplant	Benutzerfehler	Flashback	Keine Ausfallzeit
Ungeplant	Datenkorruptionen Datenverlust Rechnerverlust	Recovery Manager	Abhängig vom Ausfallgrund
Geplant	Rekonfiguration	Dynamic Reconfiguration	Keine
Geplant	Reorganisation	Online Reorganisation	Keine

4 Wiederanlaufzeiten der einzelnen Verfahren

Je nach Fehlerfall und -typ unterscheiden sich die Wiederanlaufzeiten zum Teil enorm. Doch selbst innerhalb eines Fehlertyps variiert der Zeitraum, der für die Wiederherstellung der Verfügbarkeit eines Oracle Datenbankservers benötigt wird. So benötigt ein Standby System für den Failover eine Initialisierung durch einen Administrator und muss möglicherweise zunächst noch Redo Informationen verarbeiten, bevor die Datenbank Instanz auf dem Standby System gestartet werden kann.

Bei einem Cold Standby System müssen Ressourcen Gruppen den Cluster Knoten wechseln. Entsprechende Disk Devices müssen gemountet werden, bevor die Oracle Datenbank angesprochen werden kann.

Eine Hot Standby Lösung dagegen gestattet einen enorm schnellen Failover: Da weitere Instanzen bereits geöffnet und betriebsbereit sind, muss nur ein Client-Failover erfolgen, der insbesondere bei Anwendungen, die cluster aware sind, in weniger als einer Minute erfolgen kann. Als cluster aware werden Anwendungen bezeichnet, die über Schnittstellen mit dem Cluster Manager kommunizieren bzw. auf clusterspezifische Fehlermeldungen reagieren.

Andere Technologien wiederum vermeiden Downtimes. Dynamic Reconfiguration und Online Reorganisation gestatten Wartungsarbeiten an der Datenbank im laufenden Betrieb.

Tabelle 2: Oracle Verfügbarkeitstechnologien

	Schutz vor logischen Fehlern	Schutz vor Datenverlust	Disaster Schutz	Rechnerverfügbarkeit	Volle Ressourcennutzung
Failover Cluster	Nein	Nein	Siehe 1	Ja	Nein
Standby-Datenbanken	Nein	Ja	Siehe 2	Ja	Teilweise3
Real Application Clusters	Nein	Nein	Siehe 4	Siehe 5	Ja
Geo Cluster	Nein	Ja	Ja	Ja	Siehe 6
Datenbankspiegel	Siehe 7	Ja	Ja	Ja	Siehe 8
Dynamic Reorganisation	Nein	Nein	Nein	Nein	Ja
Dynamic Reconfiguration	Nein	Nein	Nein	Nein	Ja
Flashback	Ja	Ja	Nein	Ja	-
Recovery Manager	Ja	Ja	Ja	Nein	-
Automatic Storage Management	Nein	Siehe 9	Nein	Ja	-
Infrastrukturmaßnahmen	Nein	Ja	Ja	Ja	-
Hardware Redundanz	Nein	Siehe 10	Nein	Siehe 11	-

Anmerkungen zu „Tabelle 1: Oracle Verfügbarkeitstechnologien“

- 1 Bei Distanz zwischen Rechnerknoten: Ja
- 2 Bei Verwendung von Remote Standby DB: Ja
- 3 Reporting auf Standby DB möglich
- 4 Nein (Ausnahme: entfernte Cluster Nodes)
- 5 Ja, auch bei Hardwareausfall: Unterbrechungsfreier Betrieb möglich
- 6 Abhängig vom Hersteller
- 7 Bei zeitversetzter Spiegelung: Ja
- 8 Abhängig vom Hersteller
- 9 Auf Hardware-Level: Ja
- 10 Bei Verwendung von RAID 1 oder 5: Ja
- 11 Bei Hardwareausfall: Unterbrechungsfreier Betrieb möglich

5 Resümee

Jede der von Oracle angebotenen Technologien adressiert bestimmte Ausfallkategorien. Die meisten Technologien lassen sich miteinander kombinieren. So kann in einem Real Application Cluster eine Standby-Datenbank verwendet werden, in der wiederum Oracle Flashback eingesetzt wird. Eine solche Kombination deckt nahezu alle Fehlertypen ab: Der Ausfall eines Rechners wird durch einen zweiten Clusterknoten des Real Application Cluster aufgefangen, Benutzerfehler werden durch Flashback abgedeckt. Fällt ein komplettes Rechenzentrum aus, so steht die Standby-Datenbank des Remote Rechenzentrums zur Verfügung.

Ob tatsächlich alle diese Maßnahmen notwendig sind, ist im Rahmen der Bedarfsanalyse zu bestimmen. Nicht immer ist der finanzielle und technische Aufwand gerechtfertigt. Oft genügt auch eine recht einfache Struktur, um die zuvor ermittelten Anforderungen abzudecken.

Oracle 10g Hochverfügbarkeit

Die ausfallsichere Datenbank mit RAC, Data Guard und Flashback

A. Held, Addison-Wesley 2004/ Edition Oracle

ISBN: 3-8273-2163-8

ca. 650 Seiten - 1 CD, 1-farbig

€ 59,95 [D]*

Weitere Informationen sind unter <http://www.oracle-grid.de/verfuegbarkeit> erhältlich.