

Szenarien zu Hochverfügbarkeit und Skalierung mit und ohne Oracle RAC

Alexander Scholz
its-people
Frankfurt am Main

Schlüsselworte

Hochverfügbarkeit ohne Ausfallzeiten mit Oracle RAC gewährleisten, Oracle RAC One Node mit Online Relocation als Alternative zu Oracle RAC, Oracle Data Guard als Hochverfügbarkeitslösung zwischen mehreren Standorten

Einleitung

Außer Oracle RAC gibt es noch Oracle RAC One Node und Oracle Data Guard, um Hochverfügbarkeitsanforderungen zu gewährleisten. Auch wenn diese Lösungen nicht den vollen Umfang eines Oracle RAC bieten, reichen sie für viele Kundenbedürfnisse aus und sind oftmals kostengünstiger. Nachfolgend wird die grundsätzliche Arbeitsweise von Oracle RAC, Oracle RAC One Node und Oracle Data Guard vorgestellt und deren Unterschiede aufgezeigt. Eine detaillierte Feature Liste oder Konfigurationsbeschreibung ist nicht Gegenstand dieser Beschreibung. Abschließend wird betrachtet, welches Produkt für welche Anforderungen am besten geeignet ist.

Oracle Real Application Cluster (RAC)

Wenn es um Hochverfügbarkeit und Skalierbarkeit geht, ist Oracle RAC das Produkt, das am häufigsten genannt wird. Oracle RAC kombiniert Hochverfügbarkeit und Skalierbarkeit und garantiert dadurch hochverfügbare Lösungen auch für sehr große und komplexe Systemumgebungen. Lizenztechnisch ist Oracle RAC eine Option der Enterprise Edition und setzt diese für die Lizenzierung voraus.

Eine Datenbankanwendung kann ohne Designänderungen in einem RAC eingesetzt werden und steht gleichzeitig auf mehreren physikalischen Servern, den sogenannten Knoten, zur Verfügung. Wenn ein Knoten ausfällt, werden alle Anfragen an die verbliebenen Knoten weitergeleitet und die Anwendung steht weiterhin zur Verfügung. Aus interner Sicht beschreibt ein Cluster einen Pool unabhängiger Server, die nach außen wie ein einziges System zu sehen sind.

Die Hochverfügbarkeit eines Clusters wird durch redundante Systeme sichergestellt. Die Redundanz erstreckt sich dabei nicht nur auf die Server, sondern auch auf den Plattenspeicher und die Netzwerkverbindungen. Die Oracle Instanz ist von der Oracle Datenbank getrennt. Unter der Oracle Instanz werden die Prozesse und Hauptspeicherstrukturen verstanden. Diese befinden sich auf den einzelnen Servern im RAC. Unter der Oracle Datenbank werden die physikalischen Speicherstrukturen, also die Datafiles, verstanden. Die Datenbank befindet sich in einem von allen Servern geteilten und gleichzeitig zugänglichen Plattenspeicher.

Die Instanzen und die Listener im RAC werden durch die Clusterware überwacht und bei Ausfall automatisch neu gestartet. Die Clusterware kommuniziert mit den einzelnen Knoten über den sogenannten Interconnect. Dies ist eine private Netzwerkverbindung, die den Knoten vorbehalten ist.

Die Clients kommunizieren mit dem RAC über eine öffentliche virtuelle Netzwerkverbindung, die bei Ausfall eines Knotens auf den nächst verfügbaren gelenkt wird. Ein Zugriff auf einzelne Server innerhalb des RAC von den Clients aus ist nicht notwendig. Für die Clients verhält sich der RAC stets wie ein einziges System.

Wenn die Performance des bestehenden RAC nicht mehr ausreicht, kann der RAC ohne Ausfall oder Unterbrechung beliebig skaliert werden. Es können weitere Server hinzugefügt werden und bestehende Anwendungen können beliebig Servern zugeordnet werden. Der RAC sorgt für ein automatisches Loadbalancing der Clientanfragen und lastet so die Knoten optimal aus. Auch SQL Abfragen im parallel Modus können auf verschiedene Knoten verteilt werden.

Die nachfolgende Abbildung zeigt schematisch die Funktionsweise eines RAC. Eine detailliertere Darstellung der einzelnen Prozesse und Funktionsweisen innerhalb des RAC ist im Rahmen dieses Vortrages nicht möglich.

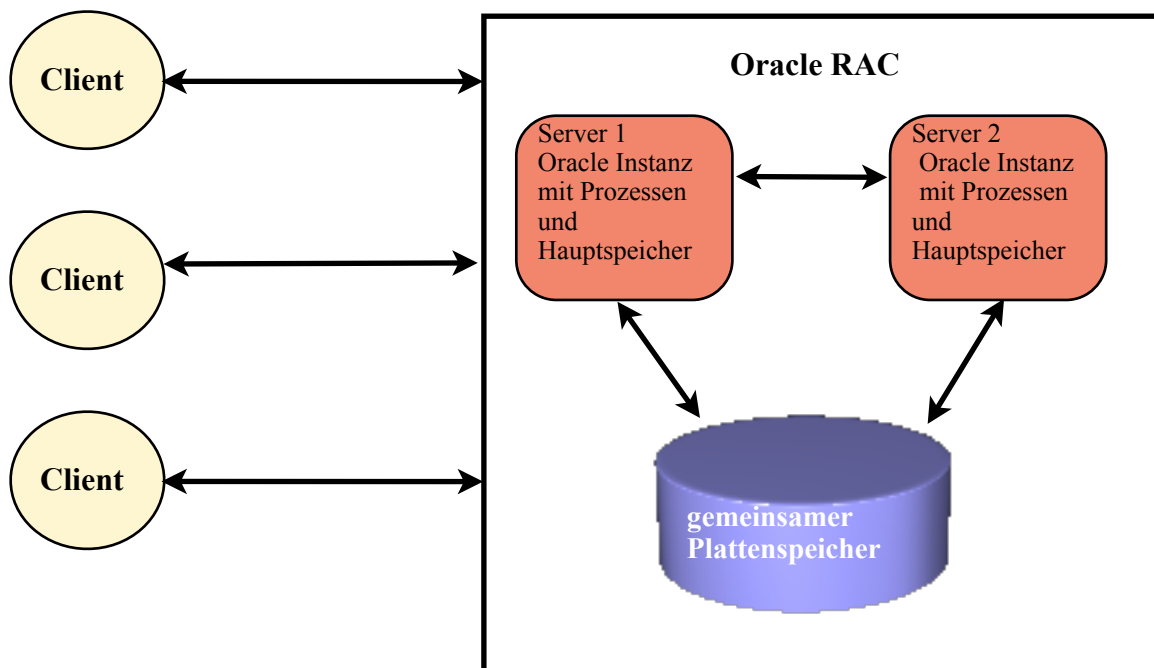


Abb1. schematischer Aufbau Oracle RAC

Oracle RAC One Node

Oracle RAC One Node kann mit einer hochverfügbaren Single Instanz Datenbank verglichen werden. Oracle RAC One Node ist ähnlich wie ein normaler RAC aufgebaut. Allerdings sind die einzelnen Knoten nicht gleichzeitig aktiv, sondern nur jeweils eine Instanz. Alle anderen Instanzen im RAC sind heruntergefahren und dienen als Standby Datenbanken.

Analog zum Oracle RAC nutzen alle Server einen gemeinsamen Plattenspeicher und werden über die Clusterware kontrolliert. Wenn ein Server ausfällt, wird über die Clusterware ein Standby Knoten hochgefahren. Alle Clients greifen dann automatisch auf die Standby Datenbank zu.

Dadurch dass immer nur ein Server verfügbar ist, ist die Datenbank bis zum hochfahren des Standby Systems nicht verfügbar. Oracle RAC One Node ist daher nur für Anwendungen geeignet, die eine kurze Ausfallzeit tolerieren können. Anders als bei reinen Single Instanz Lösungen erfolgt das Starten der Standby Datenbank automatisch und bedarf keiner Interaktion des Administrators. Die Ausfallzeit beschränkt sich daher auf die Zeit, die das Standby System zum Hochfahren benötigt.

Oracle RAC One Node ermöglicht das Online Relocation einer Datenbank. Das bedeutet, dass eine Datenbank online von einem Server auf den anderen verschoben werden kann. Bei anstehenden Wartungsarbeiten, wie zum Beispiel Betriebssystempatch, muss die Datenbank nicht heruntergefahren werden. Die Standby Datenbank wird hochgefahren und übernimmt nach und nach alle Sessions und alle laufenden Transaktionen der ursprünglichen Datenbank. Sobald die Standby Datenbank alles vollständig übernommen hat, wird die ursprüngliche Datenbank heruntergefahren und steht für Wartungsarbeiten bereit. In der Zeit bis alle Sessions übernommen wurden, funktionieren beide Knoten wie ein normales RAC als Aktiv-Aktiv Komponenten. Danach arbeiten die Server wieder als Aktiv-Passiv Cluster weiter.

Durch diese Online Relocation Funktionalität ist auch eine Skalierung in eingeschränktem Umfang möglich. Solange die Anwendung auf einem Server mit besserer Hardware noch performant läuft, kann die Datenbank auf einen anderen Server verschoben werden und die alte Hardware kann ausgetauscht werden.

Das nachfolgende Bild verdeutlicht die Funktionsweise des Oracle RAC One Node.

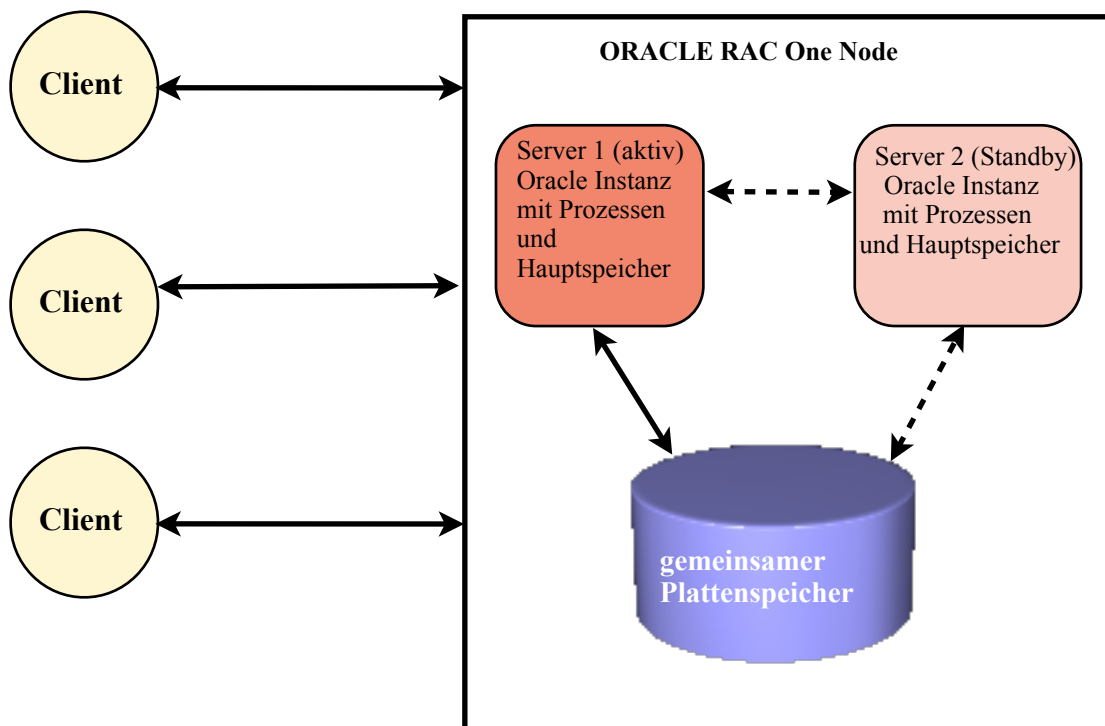


Abb2. schematischer Aufbau Oracle RAC One Node

Oracle Data Guard

Data Guard wird zum Einrichten, Verwalten und Überwachen von synchronisierten Standby Datenbanken verwendet. Es können bis zu 30 Standby Datenbanken eingerichtet werden. Data Guard kann für Hochverfügbarkeit und Disaster Recovery verwendet werden. Auf das Standby System kann bei einem Ausfall des Hauptsystems automatisch gewechselt werden. Fällt das Hauptsystem aus, steht die Standby Datenbank als Hauptsystem zur Verfügung. Im Unterschied zu RAC One Node befinden sich die Daten nicht in einem gemeinsamen Plattenbereich, sondern jeder Server hat seinen eigenen Plattenbereich. Standby Datenbanken mittels Data Guard können auch zwischen weit entfernten Standorten eingerichtet werden.

Data Guard sammelt die Redo Informationen aus dem Primärsystem ein, überträgt sie auf die Standby Datenbank und verarbeitet die Redo Informationen auf der Standby Datenbank. Die Übertragung der Redo Informationen kann synchron oder asynchron erfolgen. Bei einer synchronen Übertragung wartet das Primärsystem bei einem Commit auf die Bestätigung der Standby Datenbank, dass die Redo Informationen übertragen wurden. Dadurch ist sichergestellt, dass keine Redo Information verloren geht. Je nach Netzwerkverbindung kann es hier zu Performanceeinbußen kommen, da nach jedem Commit auf die Bestätigung gewartet wird. Bei der asynchronen Übertragung wartet das Primärsystem nicht auf die Bestätigung der Standby Datenbank, sondern fährt mit der weiteren Verarbeitung fort. Performanceeinbußen entstehen hierbei nicht, allerdings ist die Vollständigkeit der Redo Informationen nicht gewährleistet. Es kann passieren, dass Redo Informationen verloren gehen und damit eine synchrone Standby Datenbank nicht mehr gewährleistet ist.

Die Standby Datenbank kann als physikalische oder logische Standby Datenbank implementiert werden. Bei der physikalischen Standby Datenbank wird mittels Redo Apply eine exakte Replik der Primärdatenbank erstellt. Alle Speicherstrukturen stimmen mit der primären Datenbank überein. Bei der logischen Standby Datenbank werden mittels SQL Apply alle Daten aus der Primärdatenbank kopiert. Die physikalische Organisation und Datenstruktur kann aber von der primären Datenbank abweichen. Der Vorteil der logischen anstelle der physikalischen Implementierung ist, dass auf der Standby Datenbank zusätzliche Datenbankobjekte erstellt werden können, die auf der primären Datenbank nicht vorhanden sind. Auf diese Weise kann die Standby Datenbank für zusätzliche Auswertungen verwendet werden und entlastet damit das primäre System.

Mit Hilfe der Active Data Guard Option kann die Standby Datenbank für nur lesende Zugriffe zeitgleich mit der Synchronisierung benutzt werden. Diese Active Option setzt die Enterprise Edition voraus und ist zusätzlich lizenzierungspflichtig.

Für Testzwecke kann ein Snapshot der Standby Datenbank für Lese- und Schreiboperationen erstellt werden. Die Redo Informationen werden weiterhin eingesammelt, aber nicht in die Standby Datenbank eingespielt. Nachdem der Test beendet ist, wird der Snapshot wieder in eine synchronisierte Standby Datenbank konvertiert und die bisher im Primärsystem gemachten Änderungen werden in das Standby System übernommen.

Die nachfolgende Abbildung zeigt schematisch die Arbeitsweise von Data Guard.

Oracle Data Guard

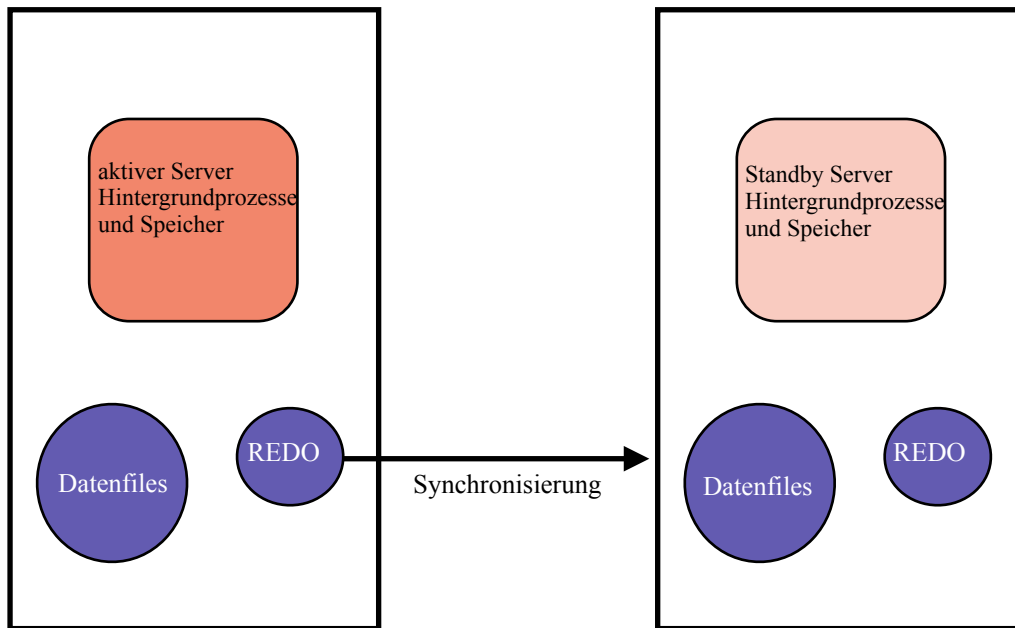


Abb3. schematischer Aufbau Oracle Data Guard

Lösungen von Hochverfügbarkeit und Skalierbarkeit außerhalb von Oracle

Neben den oben vorgestellten Lösungen gibt es auch noch weitere Möglichkeiten um eine hochverfügbare Lösung zu realisieren.

Anstelle des Oracle Clusters kann auch ein Betriebssystemcluster installiert werden. Als Datenbank kann auf jedem Knoten eine Standard Oracle Datenbank installiert werden. Da das Betriebssystem geclustert ist, wird bei Ausfall eines Knotens auf den anderen umgeschaltet. Da dieses Clustering nicht über die Oracle Prozesse überwacht wird, sind mit dieser Möglichkeit einige Nachteile verbunden. Zum einen funktioniert das Clustering nur, wenn das Betriebssystem ausfällt bzw. wenn ein Knoten nicht mehr verfügbar ist. Wenn der Knoten noch verfügbar ist, aber die Oracle Datenbank nicht mehr, wird nicht geclustert und die Datenbankanwendung steht. Auf dem Standby Knoten darf entweder keine Datenbank hochgefahren sein, oder diese muss einen anderen Namen haben. Im ersten Fall muss die Clusterlösung die Datenbank auf dem zweiten Knoten starten und dafür sorgen, dass auf dem ersten Knoten die Datenbank nicht mehr gestartet wird. Ansonsten existieren zwei Datenbanken unter dem gleichen Namen, was zu Konflikten führt. Im zweiten Fall müssen auf den Clients die Datenbanken für jeden Knoten extra in die `tnsnames.ora` eingetragen werden und die Anwendungslogik ist für die Auswahl des richtigen Knotens verantwortlich. Die Datensynchronisierung zwischen den beiden Knoten erfolgt nicht automatisch, sondern muss manuell aufgebaut werden.

Diesen Nachteilen steht nur der Vorteil der günstigeren Lizenzen im Vergleich zu den Oracle RAC Lösungen gegenüber. Da diese Lösung eigene Überwachung und Synchronisierungsmechanismen erwartet, ist sie nur für kleinere Anwendungen geeignet und ersetzt keine vollwertige Hochverfügbarkeitslösung. Sie ist nur als Ergänzung einer ansonsten Standalone stehenden Anwendung empfehlenswert, um mögliche Ausfallzeiten zu verkürzen.

Vergleich und Bewertung der Alternativen

Für eine Bewertung der individuell besten Möglichkeit ist zunächst zu klären ob bei einer Hochverfügbarkeitslösung ein kurzer Ausfall von einigen wenigen Minuten bis zum Start der Standby Datenbank akzeptabel ist oder nicht. Wenn ein auch noch so kurzer Ausfall der Anwendung nicht möglich ist, sollte Oracle RAC eingesetzt werden.

Wenn die Anwendung eine kurze Ausfallzeit zulässt muss geklärt werden, ob ein gemeinsamer Plattenbereich genutzt werden kann, oder ob für jeden Knoten ein separater Plattenbereich verwendet wird. Ist ein gemeinsamer Plattenbereich verfügbar, ist Oracle RAC One die bessere Alternative, da Datenbanken im laufenden Betrieb von einem Knoten zum anderen umgeschaltet werden können. Ein wartungsbedingter Ausfall wird dadurch vermieden.

Oracle Data Guard sollte eingesetzt werden, wenn die Anwendung eine kurze Ausfallzeit toleriert und kein gemeinsamer Speicher zur Verfügung steht. Da mit Data Guard auch weiter entfernte Datenbanken synchronisiert werden können, bietet sich hier eine Alternative um eine Hochverfügbarkeitslösung auf mehrere Standorte zu verteilen.

Eine Hochverfügbarkeitslösung mit eigenen Bordmitteln ist nur zur Ergänzung einer Standalone Lösung empfehlenswert. Das bedeutet, dass die Anwendung keine Hochverfügbarkeitslösung erfordert, die bestehende Verfügbarkeit aber durch den Einsatz eines Betriebssystemcluster verbessert wird.

Hinsichtlich der Skalierbarkeit ist der Einsatz eines Oracle RAC dann empfehlenswert, wenn die Anwendung aus Performancegründen auf mehreren Servern verteilt laufen muss. Im Oracle RAC kann jederzeit ein weiterer Knoten hinzugefügt werden, der dann Teile der Anwendungslogik übernehmen kann.

Solange die Performanceeinbußen durch Verwendung besserer Hardware ausgeglichen werden können und die Anwendung nicht zwingend auf mehreren Servern laufen muss, kann Oracle RAC One Node oder Dataguard eingesetzt werden.

Kontaktadresse:

Alexander Scholz
its-people
Lyoner Straße 44-48
D-60528 Frankfurt am Main

Telefon: +49 (0) 69-2475210-0
Fax: +49 (0) 69-2475210-21
E-Mail: alexander.scholz@its-people.de
Internet: www.its-people.de