

Datenbank als RAC oder in einer virtualisierten Umgebung?

Autor: Jochen Kutscheruk, merlin.zwo InfoDesign GmbH & Co. KG

Sowohl Oracle Datenbanken in virtualisierten Umgebungen als auch Hochverfügbarkeitslösungen auf Basis der Oracle Grid Infrastructure erfreuen sich zu Recht einer beachtlichen Nachfrage. Dabei wird immer wieder die Frage gestellt, welche der beiden Lösungen denn die Bessere wäre. Darauf möchte ich in diesem Artikel eingehen.

Die Begriffe Real Application Cluster (RAC) und Grid Infrastructure werde ich dabei nicht sauber trennen. RAC als Hochverfügbarkeits-Technologie für die Datenbank hat inzwischen einen gewissen Bekanntheitsgrad erreicht, während man von dem Umfang der Übermenge „Grid Infrastructure“ meistens noch keine klare Vorstellung hat. Es ist jedoch nicht Ziel dieses Artikels, die Feinheiten dieses Unterschieds zu erläutern.

Eines vorweg: eine Oracle Real Application Cluster (RAC) Datenbank zu virtualisieren und damit scheinbar von den Vorteilen beider Lösungen auf einmal zu profitieren, halte ich persönlich für keine gute Idee.

Oracle RAC ist nur auf einer einzigen Virtualisierungsplattform – Oracle VM - von Oracle zertifiziert. Auf allen anderen (ESX, HyperV, ...) wird RAC tatsächlich nicht funktionieren. Dies hat technische Gründe im Timing – von einer virtualisierten Umgebung wird normalerweise keine bestimmte Antwortzeit garantiert. Dies genau ist jedoch für die Funktion der Grid Infrastructure essentiell notwendig.

Welches Ziel möchten Sie erreichen?

Das ist die alles entscheidende Frage. Wenn Ihre Antwort lautet: „Ich möchte für meine kleine, mäßig belastete Oracle Datenbank nicht extra einen ganzen Server anschaffen“, dann sollten Sie diese Datenbank tatsächlich einfach auf einem virtualisierten Server betreiben. Punkt.

Wenn Ihre Antwort jedoch lautet: „Ich möchte meine Oracle Datenbank **skalierbar** und **hochverfügbar** machen“, dann sind ein paar zusätzlich Überlegungen notwendig.

Das Ziel der Skalierbarkeit

Vor nicht allzu langer Zeit war ganz klar, dass man eine Datenbank nur mit der RAC-Technologie (fast) beliebig skalieren kann. Die verfügbaren Server im x86-Bereich waren mit maximal zwei oder vier Prozessoren ausgestattet, und diese hatten nur einen oder zwei Kerne. Machte also in Summe maximal 8 CPU-Kerne für einen Server. Ebenso war der maximal mögliche Hauptspeicher begrenzt.

Waren dennoch größere Server notwendig, so gab es nur erheblich teurere Lösungen von z.B. IBM, HP oder SUN. Im Vergleich hierzu war eine x86-basierte RAC-Lösung bei vergleichbarer Leistungsfähigkeit oftmals deutlich günstiger zu implementieren.

Inzwischen erhält man jedoch auch im x86-Bereich bezahlbare 4-Prozessor Maschinen mit 4, 6, 8 oder n-Core CPUs. Dies bedeutet, dass man für seinen Datenbankserver leicht 32 Prozessorkerne und 288GB Hauptspeicher zur Verfügung stellen kann. Das sollte für überwiegende Anzahl der Datenbanken, die im produktiven Einsatz sind, mehr als ausreichend sein. Ich selbst kenne nur wenige Datenbanken, die höhere Anforderungen stellen.

Sollten dennoch höhere Anforderungen gestellt werden, so benötigt man tatsächlich eine RAC-Lösung, um entsprechend skalieren zu können. Wobei ich keinesfalls eine „gesplittete“ Lösung mit einem oder mehreren Active DataGuards unterschlagen möchte. Aber wie gesagt: diese Anforderungen werden nur von ganz wenigen Anwendungen gestellt, sie sind daher nicht Bestandteil dieser Überlegungen.

Grundsätzlich kann heute auch ein x86 basierter Server ausreichend skaliert werden, so dass damit auch anspruchsvollere Anforderungen erfüllt werden können.

Bei einer virtualisierten Lösung kann es jedoch – obwohl der Virtualisierungsserver selbst ausreichend mit CPU, Hauptspeicher und Festplatte ausgestattet ist - Einschränkungen geben:

- Die Anzahl der CPUs, die dem virtualisierten Datenbankserver zur Verfügung gestellt werden können, kann - je nach verwendeter Virtualisierungssoftware - unter Umständen begrenzt sein.
- Zusätzlich kann der Hauptspeicher, welcher dem virtuellen Server zugewiesen werden kann, begrenzt sein.

Ebenfalls ein Problem in einer virtualisierten Umgebung kann die Festplatte sein.

Grundsätzlich gibt es zwei Möglichkeiten:

- Die (virtuellen) Festplatten werden als Dateien im Dateisystem der Virtualisierungsumgebung abgebildet. Dies bedingt im Normalfall eine vergleichsweise „schlechte“ I/O-Performance – auf jeden Fall schlechter als bei einem nativen Festplattenzugriff.
- Von der Virtualisierungsumgebung werden physische Festplatten der zentralen Storage direkt an den virtualisierten Server durchgereicht. Wenn zusätzlich der Festplattenzugriff über exklusive Festplattenkanäle (z.B. einen dedizierten FC- oder iSCSI-Adapter) direkt aus dem virtualisierten Server erfolgen kann, erreicht man annähernd die gleiche Festplattenperformance wie auf echter Hardware.

Bei einer RAC-Lösung mit direktem physischen Zugriff auf die zentrale Storage können zumindest leichte I/O-Performancevorteile erwartet werden. Allerdings gibt es derart viele Möglichkeiten, eine Storage schlecht zu konfigurieren, dass ich dies nicht als generellen Vorteil einer RAC-Lösung werten möchte. Ein gut konfigurierter „virtueller“ Festplattenzugriff kann um Klassen schneller sein als eine schlecht konfigurierte Storage für einen RAC.

Andererseits sind durch eine RAC-Lösung auch Performanceeinbußen zu erwarten. Dabei kann es sich um eine vernachlässigbare Größenordnung handeln; bei nicht RAC-geeigneter Software kann der „Global Cache“ aber durchaus einen Flaschenhals und eine Performancebremse darstellen.

Bis auf wenige klare Ausnahmefälle kann man bei der Frage der Skalierbarkeit nicht automatisch davon ausgehen, dass eine RAC-Lösung besser sein muss als eine virtualisierte Lösung.

Das Ziel der Hochverfügbarkeit

Sowohl die RAC- als auch die Virtualisierungs-Variante können im Bereich „Hochverfügbarkeit“ punkten und sind dabei einer „konventionellen“ Single Server Lösung eindeutig überlegen.

Bei der RAC-Variante ist der Ausfall eines Servers quasi schon von Hause aus zulässig. Aber auch in einer virtualisierten Umgebung werden beim Ausfall eines Virtualisierungsservers die darauf befindlichen virtuellen Server auf einem der verbleibenden Server im Verbund neu gestartet. Das geschieht zwar nicht „unterbrechungsfrei“, kann aber in der Praxis vollkommen ausreichend sein.

Bei einer (geplanten) Hardwarewartung sind beide Lösungen theoretisch gleichauf. In einer virtualisierten Umgebung werden dabei die auf dem zu wartenden Server befindlichen virtuellen Server auf andere, im gleichen Verbund befindliche Server „migriert“. Dies geschieht im laufenden Betrieb und ohne dass es eine Anwendung oder ein Anwender bemerken würde.

In einer RAC-Umgebung wird der zu wartende Server heruntergefahren, die anderen Server übernehmen dessen Aufgaben, die bestehenden Datenbankverbindungen werden auf die verbleibenden Server „umgezogen“. Theoretisch sollte das zum gleichen Ergebnis führen wie in der virtualisierten Welt, leider zeigt jedoch die Praxis, dass fast keine Anwendung wirklich für den Einsatz in einer RAC-Umgebung angepasst ist. Der Failover der Verbindung klappt noch problemlos, wenn sich jedoch die Anwendung z.B. in Packagevariablen Einstellungen gespeichert hat, dann sind diese nach dem Failover weg, die Anwendung fällt auf die Nase und muss neu gestartet werden. Das ist eindeutig nicht „hochverfügbar“ (auch wenn es nicht um einen Fehler des RAC handelt).

Insofern hat also bei der Hardwarewartung in der Praxis die virtualisierte Lösung die Nase vorne – die Anwendung behält hier die Datenbankverbindung und muss nicht angepasst werden.

Bei einer Softwarewartung (Update der Datenbanksoftware) lassen sich keine wirklichen Unterschiede zwischen beiden Varianten feststellen. Bei der RAC-Lösung muss man normalerweise jedoch nicht nur die Datenbanksoftware aktualisieren, sondern zusätzlich noch die Grid Infrastructure Installation. Dieses Update der Grid Infrastructure kann jedoch im laufenden Betrieb „reihum“ und ohne Unterbrechung erfolgen – zumindest ab Oracle Version 11.2.

Beim Update der Datenbank selbst sind die gleichen Schritte durchzuführen wie bei einer Single Instance. Ein funktionierendes „Rolling Update“ habe ich – obwohl es inzwischen möglich sein sollte (Stichwort „Editionen“) – noch nicht gesehen.

Bleibt noch die Frage, wie viel Hochverfügbarkeit denn tatsächlich benötigt wird. Muss die Datenbank wirklich absolut unterbrechungsfrei zur Verfügung stehen, so bleibt tatsächlich nur die RAC-Lösung. Dazu müssen jedoch vorab penibel sämtliche Ausfallszenarien durchgetestet und auch die Anwendung an die RAC-Umgebung (Stichwort: FAN-Notification) angepasst werden. Dies ist ein nicht zu unterschätzender Aufwand, eventuell lässt sich die notwendige Anpassung der Software, insbesondere bei „Standardsoftware“, auch gar nicht erreichen.

Oftmals ist jedoch eine wirklich unterbrechungsfreie Lösung tatsächlich nicht notwendig. Speziell wenn der zusätzliche Aufwand und die Risiken einer Grid Infrastructure Installation mit in Betracht gezogen werden. Die zusätzliche Software kann wiederum selbst Fehler beinhalten, die gerade zum Ausfall der eigentlich hochverfügbaren Datenbank führt. Man handelt sich also eine weitere mögliche Fehlerquelle ein, die den scheinbaren Vorteil der „unterbrechungsfreien Hochverfügbarkeit“ wieder zunichte machen kann. Dies ist insbesondere in Umgebungen zu beobachten, welche nicht mit der notwendigen Erfahrung und den erforderlichen umfangreichen Failover-Tests implementiert wurden.

Bei dem Wunsch nach Hochverfügbarkeit stellt sich immer die Frage, wie tragisch denn eine zwei-, drei- oder fünfminütige Unterbrechung der Verfügbarkeit der Datenbank objektiv betrachtet wirklich wäre – insbesondere vor dem Hintergrund, dass dies ja nur extrem selten geschieht. Oftmals ist dies tatsächlich kein wirkliches Problem.

Ein Missverständnis gilt es eventuell noch auszuräumen: der Begriff „Hochverfügbarkeit“ bezieht sich nicht auf die Sicherheit der Daten – also Schutz vor physischem Festplattencrash oder logischer Zerstörung. Vor diesem Problem können beide Varianten von sich aus nicht schützen. Dies gelingt nur entweder mit einer passenden Sicherheits- und Sicherungsstrategie oder aber mit DataGuard.

Zusammenfassung

Tatsächlich gibt es keine einfache Antwort, ob eine RAC- oder eine virtualisierte Lösung die bessere Variante ist. „It depends“ – leider gilt dies auch hier.

Nur in wenigen Ausnahmefällen kann die Antwort klar zu Gunsten einer RAC-Lösung ausfallen:

- Die Datenbank muss tatsächlich unterbrechungsfrei verfügbar sein, die Software kann angepasst werden
- Die Datenbank lastet die verfügbare Hardware bereits massiv aus – sowohl was CPU als auch was I/O betrifft. Eine bereits ausgelastete Hardware ist kein guter Kandidat für eine Virtualisierungslösung.
- Die virtualisierte Umgebung kann nicht die notwendige Performance liefern, eventuell weil bereits andere Server mit hoher Last auf den gleichen Servern virtualisiert wurden.

Oft zu beobachten ist auch, dass Firmen in einer Art „Virtualisierungs-Überschwang“ alle verfügbaren Server virtualisieren. Dadurch werden jedoch letztlich sowohl die Virtualisierungsserver als auch die Storage überlastet.

Es empfiehlt sich tatsächlich, vorab etwas Zeit in Beratung und auch Tests zu investieren, um die wirklich beste Lösung für die konkreten Anforderungen zu finden. Die zu beachtenden Punkte konnten hier nur angerissen werden.

Kontakt:

Jochen Kutscheruk
merlin.zwo InfoDesign GmbH & Co. KG
jochen.kutscheruk@merlin-zwo.de