Hochverfügbarkeit und K-Fall-Absicherung mit Engineered Systems

Hartmut Streppel und Andris Perkons, Oracle Deutschland B.V. & Co. KG

Die Absicherung von Geschäftsprozessen gegen Unterbrechungen durch Katastrophen aller Art ist notwendiger denn je. Solche Business-Continuity-Projekte sind sehr komplex, da sie immer komplette Geschäftsabläufe betreffen und nicht nur den IT-Anteil.

Der Artikel betrachtet allerdings nur einen Ausschnitt der IT: die Oracle Engineered Systems, vor allem SPARC SuperCluster, und ihre Einbettung in K-Fall-Absicherungs-Konfigurationen. Im Rahmen eines internen Proof of Concept wird untersucht, wie komplette und komplexe Anwendungs-Stacks, von der Datenbank über Dateisystem-Daten bis zum Anwendungs-Layer, zwischen Engineered Systems in unterschiedlichen Rechenzentren einfach, schnell und sicher geschwenkt werden können.

Es ist interessant, Hochverfügbarkeits-Deployments in unterschiedlichen Ländern zu betrachten. Es scheint vor allem in Deutschland, aber auch in den nordischen Ländern, eine Präferenz zu geben, unternehmenskritische Systeme über getrennte Rechenzentren hinweg zu betreiben und zu clustern. In Großbritannien und den USA scheint das nach der kleinen internen Recherche der Autoren nicht der Fall zu sein.

Solche, vor allem in Deutschland anzutreffenden Campus- beziehungsweise Metro-Konfigurationen über zwei entfernte Rechenzentren sollen den Ausfall eines kompletten Rechenzentrums absichern. Cluster, die zusammen mit Daten-Spiegelung oder -Replikation über diese Rechenzentren hinweg konfiguriert sind, sollen Anwendungsschwenks zuverlässig auslösen und vollautomatisch durchführen, ohne dass ein Administrator eingreifen muss. Dies funktioniert, in vielen Fällen nachgewiesen, etwa bei plötzlichen Stromausfällen recht gut – natürlich in Abhängigkeit von der Qualität des eingesetzten Produkts und der Korrektheit der Konfiguration. Ob solche Cluster auch komplexere und zum Beispiel schleichende Katastrophen wie Brände in den Infrastruktur-Komponenten beherrschen, ist eine interessante Frage. Was solche „2-RZ-Deployments“ sicherlich nicht absichern, sind komplexe Fehler in den Single Points of Failures (SPOF):

- Daten, obwohl physikalisch mehrfach vorhanden, existieren logisch nur einmal
- Ein defektes Cluster könnte nicht mehr booten, etwa nach einer fehlerhaften Patch-Aktion
- In der Anwendungssoftware kann ein fehlerhaftes Update zu einer Situation führen, in der die Anwendung nicht mehr fehlerfrei läuft

Für all diese Fälle gibt es historische Beispiele, bei deren Betrachtung offensichtlich wird, dass auch dafür Vorsorge betrieben werden muss.

Oracle Engineered Systems

Engineered Systems (ES) von Oracle wie Exadata und SPARC SuperCluster sind hochintegrierte Systeme, die Server, Storage und Netz-Infrastruktur in einem Rack enthalten. Die Hochverfügbarkeit wird erreicht durch Redundanz innerhalb eines Systems oder einer Gruppe zusammengeschalteter ES, kombiniert mit der Nutzung von Software, die diese Redundanzen im Fehlerfall nutzt. Dies ist vor allem Cluster-Software wie Oracle Grid Infrastructure (Clusterware) und Oracle Solaris Cluster. Der Aufbau von „stretched“ oder Campus-Clustern über RZ-Grenzen hinweg, mit dem häufig RZ-Ausfälle abgesichert werden sollen, ist mit ES nicht vorgesehen und nicht unterstützt.

Ein RZ-Ausfall wird als Katastrophe (Disaster) angesehen und mit den dafür zur Verfügung stehenden Mitteln abgesichert. Für die Datenbank steht hierzu natürlich Oracle Data Guard zur Verfügung. Es sorgt für eine kontinuierliche Übertragung aller anfallenden Redo-Informationen zum System, auf dem eine Schatten-Datenbank betrieben wird und wendet die Redo-Information auf diese an. Abhängig von Parametern wie Distanz zwischen den Rechenzentren, akzeptablem Datenverlust im K-Fall, Wiederanlaufzeiten etc. kann Data Guard entweder für eine synchrone oder asynchrone Datenübertragung konfiguriert werden.

Für Anwendungen, die beispielsweise in Anwendungsdomains auf einem SPARC SuperCluster laufen und die zusätzlich Daten-Replikation benötigen, heißt die technische Lösung „Oracle Solaris Cluster (OSC) Geographic Edition“. Diese überwacht Cluster und Daten-Replikation und stellt Schwenk-Mechanismen zur Verfügung, mit denen Cluster-Ressourcen-Gruppen zusammen mit ihrer Daten-Replikation in ein entferntes Cluster geschwenkt werden können. Ein Schwenk vom primären RZ ins Ausweich-RZ kann manuell nach entsprechender Entscheidungsfindung gemäß den K-Fall-Prozessen initiiert werden. Dies ist die in der Business-Continuity-Community empfohlene Methode. Eine Automatisierung ist ebenfalls möglich. Allerdings wird im Allgemeinen wegen der Schwierigkeiten, eine Katastrophe algorithmisch sicher zu diagnostizieren, von einer Automatisierung abgeraten.

Die Trennung von lokaler Hochverfügbarkeit und Disaster Recovery über

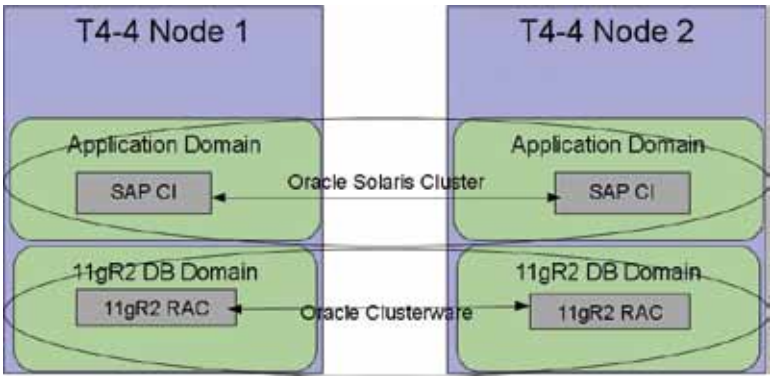


Abbildung 1: SSC-Setup mit RAC und HA SAP

RZ-Grenzen hinweg entspricht der Maximum Availability Architecture (MAA), die Oracles Blaupause für diese Themen ist (siehe http://www.oracle.com/technetwork/database/features/availability/maa-096107.html).

Ein Proof of Concept mit SPARC SuperCluster T4-4

Um ein Gefühl für die Komplexität oder besser für die Einfachheit einer solchen Konfiguration zu bekommen, führten wir einen kleinen Proof of Concept durch. Das beispielhafte Setup bestand aus zwei SPARC SuperCluster (SSC) in getrennten Rechenzentren mit je zwei Datenbank- und je zwei Anwendungs-Domains. Solche High-End-Umgebungen stehen selten als Testumgebung zur Verfügung, daher wurden sie mit zwei kleinen, T4-basierten Systemen nachgebaut. Da ein SSC aus Oracle-Standard-Komponenten besteht, war die Testumgebung funktional fast 100-prozentig identisch mit einer echten Umgebung.

Als Test-Anwendung diente eine SAP-Zentral-Instanz in der Anwendungs-Domain. Ein zusätzlicher SAP Application Server wurde für diesen kurzen Test auf einem externen Server betrieben. Die Hochverfügbarkeit der Datenbank war gewährleistet durch die Verwendung von Real Application Clusters (RAC) mit der dazugehörigen Clusterware und für die SAP-Zentral-Instanz durch Oracle Solaris Cluster (OSC) und den neuen SAP-NetWeaver-Agenten.

Hochverfügbarkeit innerhalb eines SSC

Alle typischen Einzelfehler, dazu gehören auch Ausfälle einer Domain oder sogar eines kompletten Servers, werden entweder durch die redundante Hardware, durch RAC/Clusterware oder durch OSC-Agenten abgefedert. Typische Service-Unterbrechungszeiten, wenn sie überhaupt vom Endbenutzer (beim Test war das eine interaktive „sapgui"-Session) bemerkt werden, bewegen sich hierbei im niedrigen zweistelligen Sekundenbereich. Viele Fehler sind, weil sie von noch tieferen Schichten der Architektur abgesichert sind, für den Anwender vollständig transparent. Die Tests erfolgten ohne beziehungsweise mit nur geringer Last; unter Volllast können diese Zeiten sicherlich auch etwas länger sein.

Hochverfügbarkeit über RZ-Grenzen

Doppelfehler oder komplexe Einzelfehler können grundsätzlich nicht durch Cluster abgesichert werden. So sind beispielsweise beim Ausfall mehrerer Storage-Zellen oder auch beim – sehr unwahrscheinlichen – Ausfall eines kompletten Racks die hochverfügbaren Anwendungen nicht mehr verfügbar. Auch das Cluster kann hier nicht mehr eingreifen, da auch die redundanten Komponenten mit ausgefallen sind. Um solche Fehlersituationen zu überstehen, wird noch mehr Redundanz benötigt. Diese muss vollständig unabhängig von der primären Umgebung sein, damit sichergestellt ist, dass lokale Fehler nicht in das Ausweich-RZ ausstrahlen. So sind etwa erweiterte SANs oder auch gekoppelte (IP-)Netze, die über RZ-Grenzen hinweg konfiguriert werden, anfällig gegenüber komplexen Problemen, die von zentralen Switches, IPs und FCs, ausgelöst werden. Deshalb gilt aus Sicht der K-Fall-Absicherung die Regel: Je weniger Kopplung, desto besser.

Die ausschließliche Verwendung von IP-Kommunikation zur Daten-Replikation ist deshalb eine wesentliche Anforderung für eine unabhängige Umgebung im Ausweich-RZ. Sowohl Oracle DataGuard (ODG) als auch Oracle Solaris Cluster (OSC) Geographic Edition nutzen nur IP-basierte Kommunikation. Auch die meisten Storage-basierten Replikations-Technologien, wie ZFS Storage Appliance (ZFSSA), arbeiten auf diese Weise.

Die getestete Konfiguration nutzt ODG und OSC Geographic Edition, um ein schnelles und sicheres Umschalten von der primären auf die Ausweich-Konfiguration durchzuführen. Die OSC Geographic Edition ist nicht nur in der Lage, Solaris-Cluster-Res-

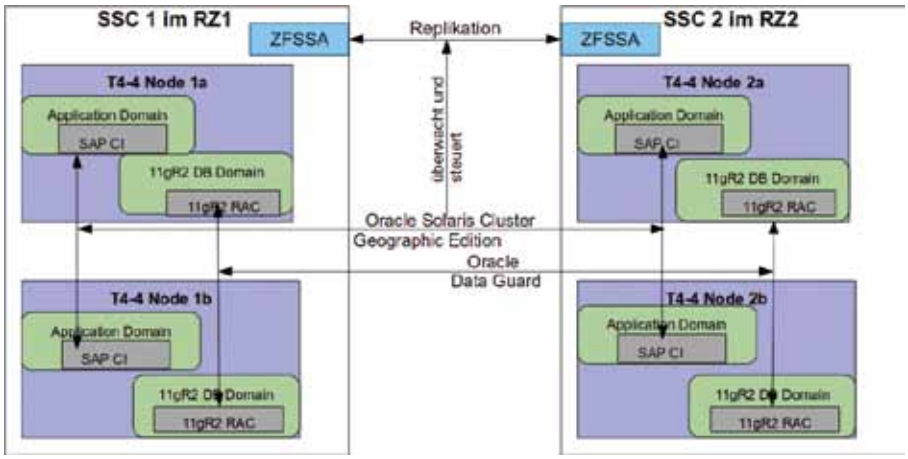


Abbildung 2: Disaster-Recovery-Lösung zwischen zwei SSCs

sourcen-Gruppen zwischen Clustern (also zwischen Clustern in verschiedenen RZs) zu schwenken, sondern auch dazu, die Replikation einer ZFSSA zu kontrollieren und zu schwenken. In diesem Test konnte die Integration der ZFSSA aus Zeitgründen allerdings nicht getestet werden.

Orchestrierung des Switchover zwischen Rechenzentren

Der Traum des Administrationsteams – und auch der Geschäftsleitung – ist es natürlich, dass selbst nach einem katastrophalen Ereignis wie einem schweren Brand in einem RZ alle geschäftskritischen IT-Systeme mit minimaler Unterbrechung und vor allem ohne Datenverlust im Ausweich-RZ weiter betrieben werden können. Um dies zu erreichen, sollte allein die Installation von ein wenig Software und im Ernstfall dann das Drücken eines großen roten Knopfs ausreichen. So eine Lösung, die eine heterogene Hard- und Software-Landschaft überwacht und dann im K-Fall schwenkt, gibt es heute nicht.

In unserem POC kombiniert ein einfaches Shell-Skript die beiden Mechanismen ODG und OSC Geographic Edition. Da die SAP-Zentral-Instanz einen automatischen Reconnect an die DB durchführt, können beide Komponenten, die Datenbank in den Datenbank-Domains und die SAP-Zentral-Instanz in den Anwendungsdomains, parallel geschwenkt werden. Um ein Gefühl für die Dauer der Service-Unterbrechungen zu bekommen, wurden einige Messungen durchgeführt. Alle Schwenks dauerten weniger als eine Minute, allerdings wieder mit der Einschränkung, dass ohne Last auf der Datenbank getestet wurde.

Betrachtet man etwa einen kompletten SAP-Stack, bestimmen bei Entwurf einer HA-/DR-Architektur mit Engineered Systems einige Details, wie schnell und reibungslos der Schwenk in ein Ausfall-Rechenzentrum durchführbar ist:

- Wechselt die Zentral-Instanz den Host-Namen beziehungsweise die IP-Adresse? In diesem Fall muss im Rahmen der Orchestrierung auch die Namens-Auflösung berücksichtigt werden. Schlimmstenfalls müssen die Applikations-Instanzen durchgestartet werden.
- Was passiert mit den NFS-Verzeichnissen für „/sapmnt“? Ist der NFS-Server über RZ-Grenzen hinweg verfügbar oder sind zusätzliche Tätigkeiten notwendig?

Fazit

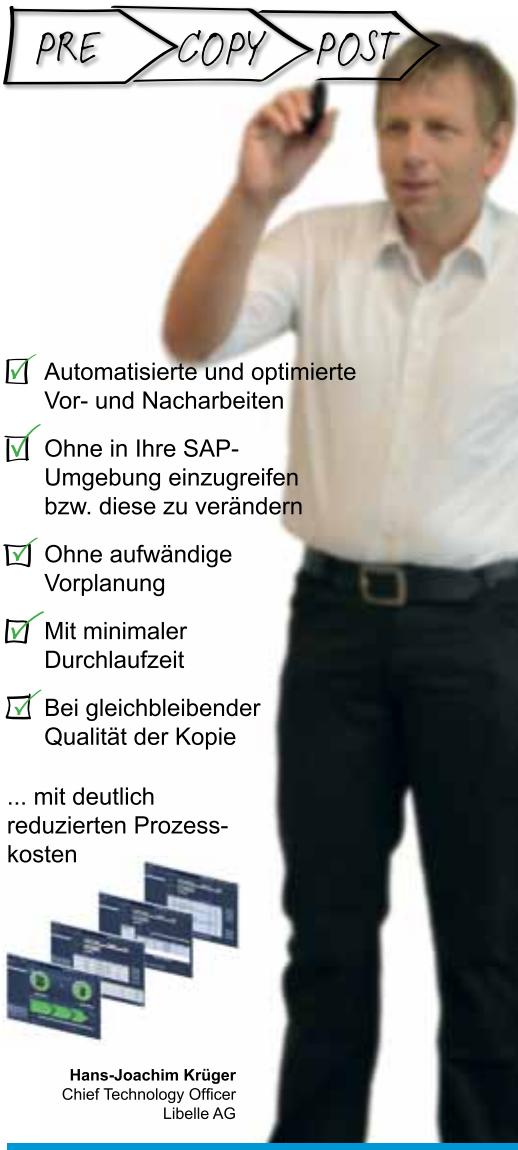
Die Tests haben gezeigt, dass mit Standard-Technologien wie Oracle Data-Guard und Oracle Solaris Cluster Geographic Edition komplette und komplexe Anwendungs-Stacks, die innerhalb eines Engineered Systems laufen, sicher und schnell auf ein anderes Engineered System in einem Ausweich-RZ geschwenkt werden können. Die Umschaltzeiten sind so kurz, dass die durch sie bedingten Service-Unterbrechungszeiten kein Grund dafür sein dürften, im Notfall für bestimmte administrative Aufgaben an Produktions-Systemen keinen Schwenk eines kompletten Stacks durchzuführen. Für die Weiterführung unternehmenskritischer Aufgaben im K-Fall sind diese Technologien erste Wahl.

Andris Perkons
andris.perkons@oracle.com

Hartmut Streppel
hartmut.streppel@oracle.com

Automatisierte SAP-Systemkopien auf Knopfdruck

Libelle SystemCopy



- Automatisierte und optimierte Vor- und Nacharbeiten
- Ohne in Ihre SAP-Umgebung einzugreifen bzw. diese zu verändern
- Ohne aufwändige Vorplanung
- Mit minimaler Durchlaufzeit
- Bei gleichbleibender Qualität der Kopie

... mit deutlich reduzierten Prozess-kosten

Hans-Joachim Krüger
Chief Technology Officer
Libelle AG

Erfahren Sie mehr:
www.Libelle.com/systemcopy

ORACLE Gold Partner

Libelle

Libelle AG
Gewerbestr. 42 • 70565 Stuttgart, Germany
T +49 711 / 78335-0 • F +49 711 / 78335-148
www.Libelle.com • sales@libelle.com