Der Artikel zeigt die Zusammenhänge zwischen virtualisierten Umgebungen, Hochverfügbarkeit und dem K-Fall „Absicherung".

Virtualisierung und Hochverfügbarkeit

Hartmut Streppel, ORACLE Deutschland B.V. & Co. KG

Zunächst ist es notwendig, Zusammenhänge und Modelle zu verstehen, um eine wirkliche Hochverfügbarkeit zu erreichen und folgende Fragen beantworten zu können: Welche Vorteile bieten virtualisierte Umgebungen? Können mit dem Einsatz von Virtualisierungstechnologien höhere Verfügbarkeiten erreicht werden? Wenn ja, wie? Hilft Virtualisierung beim Schutz gegen Katastrophen? Wo sind die Fallstricke? Wie kombiniert man Hochverfügbarkeits-Technologien mit Virtualisierung?

Virtualisierte Umgebungen

Virtualisierung ist eine alte Technologie. Schon vor mehr als dreißig Jahren waren virtuelle Maschinen auf Großrechnern im Einsatz. Leider wird heute Virtualisierung im Umfeld von Betriebssystemen fast immer mit solchen virtuellen Maschinen gleichgesetzt. In den letzten Jahren wurden auch neue Varianten wie Solaris-Zonen entwickelt, die einer Anwendung ebenfalls eine virtuelle Ablaufumgebung zur Verfügung stellen und ähnliche, aber auch weitergehende Vorteile bieten. Die Vor- und Nachteile dieser Technologien und ihre Einsatzfelder sind oft beschrieben und verglichen worden. Wenn im Folgenden von virtuellen Umgebungen die Rede ist, sind sowohl Zonen als auch virtuelle Maschinen gemeint.

Hochverfügbarkeit und Single Points of Failures

Hochverfügbarkeit (HA) wird erreicht, indem Systeme zuverlässiger gemacht und sogenannte „Single Points of Failures" eliminiert und beherrscht wer-

KeepTool mit neuer Version 10
Das handliche Werkzeug für Oracle™-Datenbanken

Zahlreiche neue Funktionen, z.B.
- Schnelle Textsuche quer über alle Tabellen im Schema.
- Praktische Tooltip-Hinweise im DataContent zeigen während der Dateneingabe Look-up-Daten zu Fremdschlüsselwerten sowie Kommentare und Datentyp-informationen zu den Spalten-überschriften an.
- Mehrstufige Pivot-Ansicht mit Visualisierung im DataContent.
Laden Sie die kostenlose Testversion unter www.keeptool.com herunter.
ORACLE Gold Partner
keeptool

den. Die deutsche Wikipedia definiert: „Unter einem Single Point of Failure (SPOF), zu Deutsch etwa „einzelne Stelle des Scheiterns“, versteht man einen Bestandteil eines technischen Systems, dessen Ausfall den Ausfall des gesamten Systems nach sich zieht.“ Kurz und vereinfacht gesagt ist alles, was nur einmal vorhanden ist, ein SPOF. Das kann ein einzelner Server sein, aber auch der einzige Anschluss ans öffentliche Stromnetz oder genau eine eingesetzte Software-Version. SPOFs werden durch den Einsatz von Redundanz, etwa mithilfe mehrerer Server, redundanter SANs etc., und auch durch den Einsatz von Cluster-Software, die Komponenten überwacht, um im Fehlerfall zu reagieren, eliminiert. Wenn ein System also keinen SPOF mehr besitzt, also alle SPOFs durch Redundanz und Cluster abgesichert sind, kann ein System nicht mehr ausfallen, solange nur einzelne Fehler auftreten.

Beim gleichzeitigen Auftreten von Fehlern spricht man in der Regel von Katastrophen oder dem K-Fall, der nur noch mit aufwändigen Disaster-Recovery-Mechanismen beherrscht werden kann. Eine Katastrophe ist also nicht nur die Überschwemmung, die den Rechnerraum überflutet, oder der Terrorangriff, der einen kompletten Campus verwüstet, sondern auch die Datenkorruption und die daraufhin entdeckte, nicht nutzbare Datensicherung. Die Übergänge sind fließend.

Cluster wie Oracle Solaris Cluster und Oracle Clusterware bestehen aus ausgefeilten Software-Komponenten, die eine extrem hohe Zuverlässigkeit garantieren und auch mit komplexen Fehlersituationen wie „split brain“ sicher umgehen können. Alternativ werden vor allem in zustandslosen Umgebungen häufig Lastverteiler und hoch redundante Umgebungen eingesetzt, bei denen Ausfälle einzelner Komponenten kaum bemerkt und ausgefallene transparent durch andere ersetzt werden. Aber Achtung: Bei der Verwendung von Lastverteilern ist das Problem der Hochverfügbarkeit nicht verschwunden, sondern nur auf die Lastverteiler und das Netzwerk verlagert worden.

Virtualisierung und Hochverfügbarkeit

Eine häufige Annahme besteht darin, dass beim Einsatz virtualisierter Umgebungen einfache Hochverfügbarkeit erreichbar sei, beispielsweise durch die Fähigkeit einer Kontroll-Instanz, in einer virtuellen Umgebung eine fehlerhafte virtuelle Maschine zu entdecken und diese auf einem anderen Server neu zu starten. Ein anderes Wundermittel wäre die Möglichkeit, eine laufende virtuelle Maschine auf einen anderen Server zu verschieben, um Hochverfügbarkeit zu erreichen. Alle

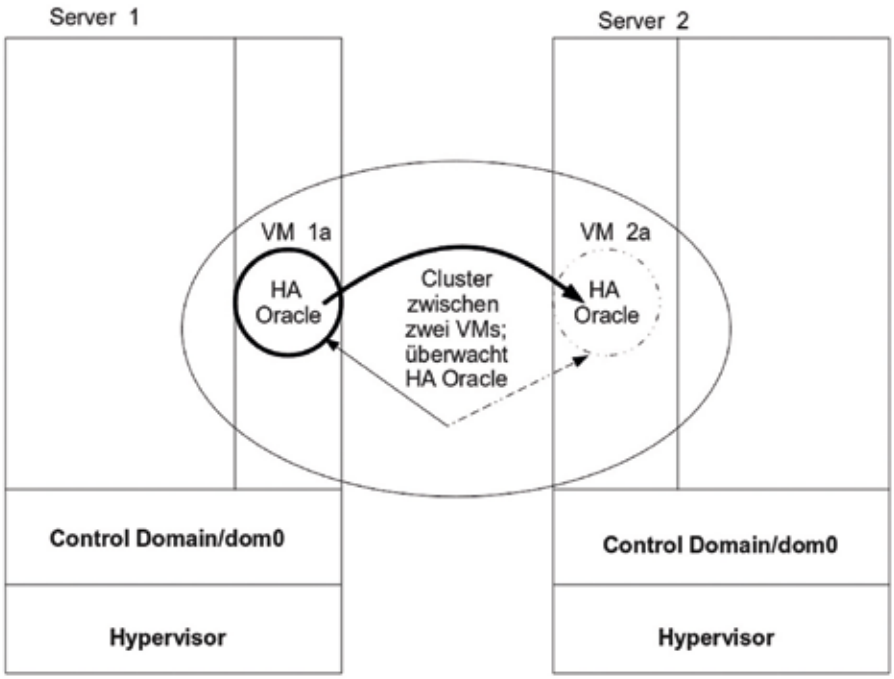

Abbildung 1: Virtuelle Maschinen als Clusterknoten: HA Oracle wird im Fehlerfall geschwenkt

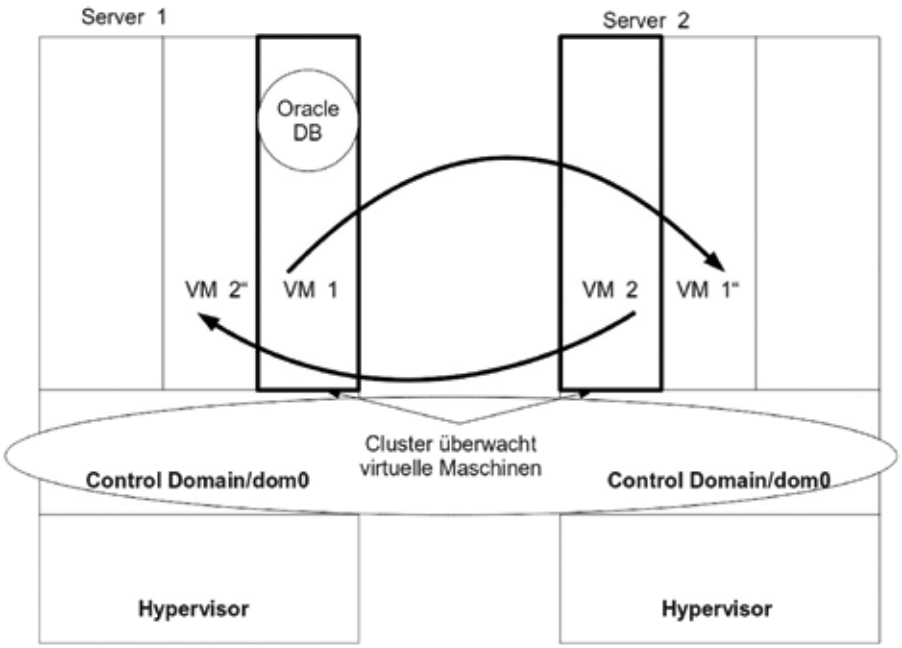

Abbildung 2: Cluster nutzt virtuelle Maschinen als Ressourcen: VM1 und VM2 können geschwenkt werden

diese Punkte sind valide, müssen aber präzisiert und erweitert werden, um ihre Bedeutung, vor allem aber ihre Lücken zu verstehen.

Virtualisierung, Cluster und zwei Hochverfügbarkeitsmodelle

Beim Einsatz von Clustern gibt es zwei grundsätzliche Modelle zur Erreichung von Hochverfügbarkeit mit virtualisierten Umgebungen:

- Eine virtualisierte Umgebung stellt hochverfügbare Dienste zur bereit
- Eine virtualisierte Umgebung ist selbst hochverfügbar (sogenannte „Black Box")

Im ersten Fall bieten die Cluster Dienste hochverfügbar an, etwa eine Datenbank oder ein ERP-System. Die virtuellen Umgebungen sind Clusterknoten. Auf diesen laufen die normalen Cluster-Mechanismen ab, um Komponenten und Dienste zu starten, zu stoppen, zu überwachen und im Fehlerfall einzugreifen. Dieses Modell wird beispielsweise implementiert mit Oracle Solaris Cluster in einer virtuellen Gastmaschine unter Oracle VM Server for SPARC oder auch mit Oracle Clusterware in einer virtuellen Gastmaschine unter Oracle VM Server for x86 (siehe Abbildung 1).

Das zweite Modell betrachtet virtuelle Umgebungen als hochverfügbare Ressourcen. Ein Cluster startet, stoppt und überwacht diese und ist in der Lage, im Fehlerfall eine solche komplette Umgebung auf einem anderen Clusterknoten neu zu starten (siehe Abbildung 2).

Cluster in virtuellen Umgebungen

Einige Beispiele illustrieren die Nutzung virtueller Umgebungen als Clusterknoten (siehe Abbildungen 1 und 3):

- *Oracle VM Server*
  Oracle Clusterware läuft in virtuellen Gastsystemen auf mehreren Servern unter Oracle VM Server und überwacht dort Oracle Real Application Clusters oder andere Anwendungen
- *Oracle VM Server for SPARC*
  Oracle Solaris Cluster läuft in virtuellen Gastsystemen auf mehreren SPARC-T-Klasse-Servern unter Oracle VM Server for SPARC und überwacht dort beliebige Anwendungen
- *Oracle Solaris Zonen*
  Oracle Solaris Cluster läuft als virtuelles Zonencluster in nicht-globalen Zonen mehrerer Oracle-Solaris-Instanzen und überwacht dort beliebige Anwendungen

Der wesentliche Vorteil dieses Ansatzes ist, dass sowohl die Konfiguration als auch der Betrieb solcher Cluster mehr oder weniger identisch ist zum Einsatz in nicht-virtualisierten Umgebungen. Unterschiedlich sind nur die Namen der nun virtuellen Devices. Das virtuelle Zonencluster vereinfacht den Betrieb noch weiter, da wesentliche Komponenten wie Quorum und Cluster Interconnect allein vom globalen Cluster verwaltet werden. Anwendungen werden mit den Standard-Mechanismen ins Cluster integriert. Für Betreiber mit suboptimalem Change Management ist die Anforderung, die Umgebungen der Clusterknoten konsistent zu halten, manchmal eine Herausforderung (siehe Abbildung 3).

Probleme mit Hypervisoren

Cluster sind aus guten Gründen sensibel bezüglich der Systemzeit und der Zuverlässigkeit der Cluster-Interconnects. In virtuellen Maschinen, nicht aber in Solaris-Zonen, ist in der Regel kein direkter Zugriff von der Gastmaschine auf die Hardware, in der die Zeit verwaltet wird, möglich. Auch gibt es keinerlei Mechanismen, um die Zuverlässigkeit von virtuellen Netzen zu beeinflussen. Deshalb können bei Clustern in virtuellen Maschinen grundsätzlich Probleme auftauchen, die bis zum Ausfall von Clusterknoten führen, weil zum Beispiel wegen Überlastung der Kontroll-Instanzen oder des Hypervisors Clockticks oder Heartbeat-Pakete verlorengehen. Nur eine perfekte Integration von Cluster- und Hypervisor-Software, die nur gewährleistet ist, wenn beide aus einer Hand kommen, ist der Garant für ein reibungsloses Funktionieren des Clusters. Die Oracle-VM-Server-Produkte sind mit Oracles Cluster-Produkten bestens integriert.

Cluster überwachen virtuelle Umgebungen

Wenn, wie als Modell 2 beschrieben, Cluster virtuelle Umgebungen über-

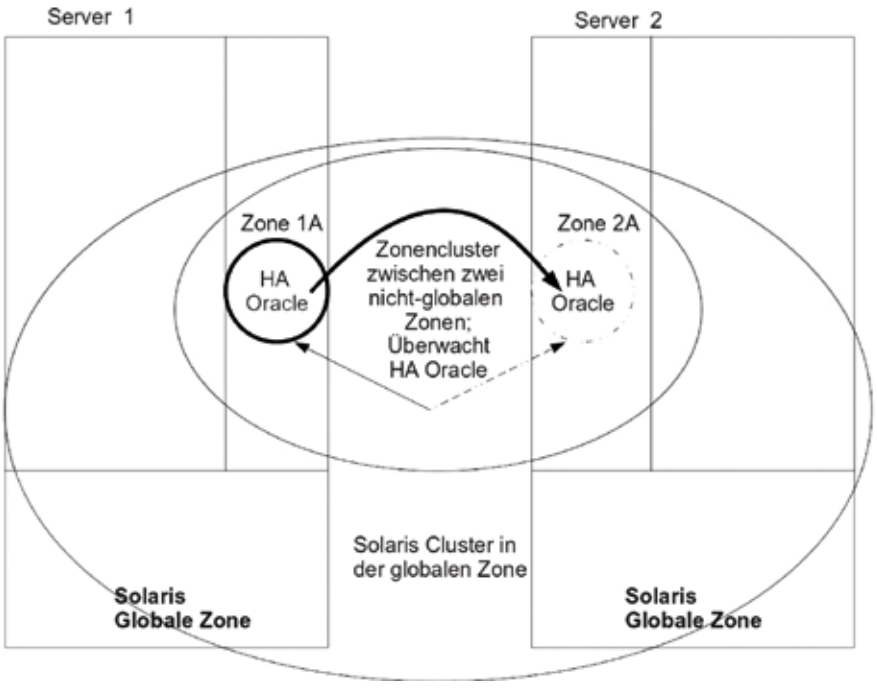

Abbildung 3: Solaris-Zonen als virtuelle Clusterknoten: HA Oracle wird im Fehlerfall geschwenkt

wachen, sind die Cluster-Komponenten Bestandteil der übergeordneten Instanz, also der Dom0, der Control Domain oder der globalen Zone (siehe auch Abbildungen 2 und 4). Da diese Komponenten direkten Zugriff auf die Hardware-Ressourcen haben, können die aufgeführten Probleme nicht auftreten: Heartbeats, Cluster-Interconnects und die Systemzeit sind zuverlässig verwaltet. Als hochverfügbare Ressourcen lassen sich nun komplette virtuelle Umgebungen wie virtuelle Maschinen oder Solaris-Container konfigurieren. Eine virtuelle Maschine ist nun kein Clusterknoten, sondern eine hochverfügbare Ressource. Innerhalb dieser werden vom Cluster keinerlei Mechanismen bereitgestellt, mit denen Dienste hochverfügbar gemacht werden können. Es müssen stattdessen betriebssystemseitige Mechanismen wie Solaris Service Management Facility (SMF) verwendet werden, um Anwendungen zu kontrollieren. Beispiele für solche hochverfügbaren Umgebungen sind:

- *Oracle VM Server for SPARC*
  Oracle Solaris Cluster läuft in der Control Domain. Gast-VMs werden mithilfe des HA-Agenten „Oracle VM Server" überwacht und können im Fehlerfall auf einem anderen Server neu gestartet werden.
- *Oracle VM Server for X86*
  Der Oracle VM Agent (unter anderem der Poolmaster) läuft in der Dom0 und überwacht virtuelle Gastmaschinen, die ebenfalls im Fehlerfall auf einem anderen Server neu gestartet werden können.
- *Oracle Solaris Zonen*
  Oracle Solaris Cluster läuft in der globalen Zone und überwacht mithilfe des „HA for Solaris Zones"-Agenten nicht-globale Solaris-Zonen. Diese können im Fehlerfall auf einem anderen Server in einer anderen globalen Solaris-Zone neu gestartet werden. Dieses Modell ist auch als „Flying Container" bekannt.

Der wesentliche Vorteil dieses Ansatzes ist, dass an der nun hochverfügbaren virtuellen Umgebung nichts geändert werden muss. Sie ist quasi „clusterfrei". Außerdem – das wird häufig als wesentliches Kriterium gesehen – ist nur eine Umgebung zu verwalten, die dann ja zwischen Clusterknoten geschwenkt werden kann. Der wesentliche Nachteil ist, dass die virtuelle Umgebung jetzt selbst zum SPOF geworden ist. Fällt sie zum Beispiel durch eine Fehlkonfiguration aus, stehen die Anwendungen, die innerhalb der virtuellen Umgebung konfiguriert wurden, nicht mehr zur Verfügung (siehe Abbildung 4).

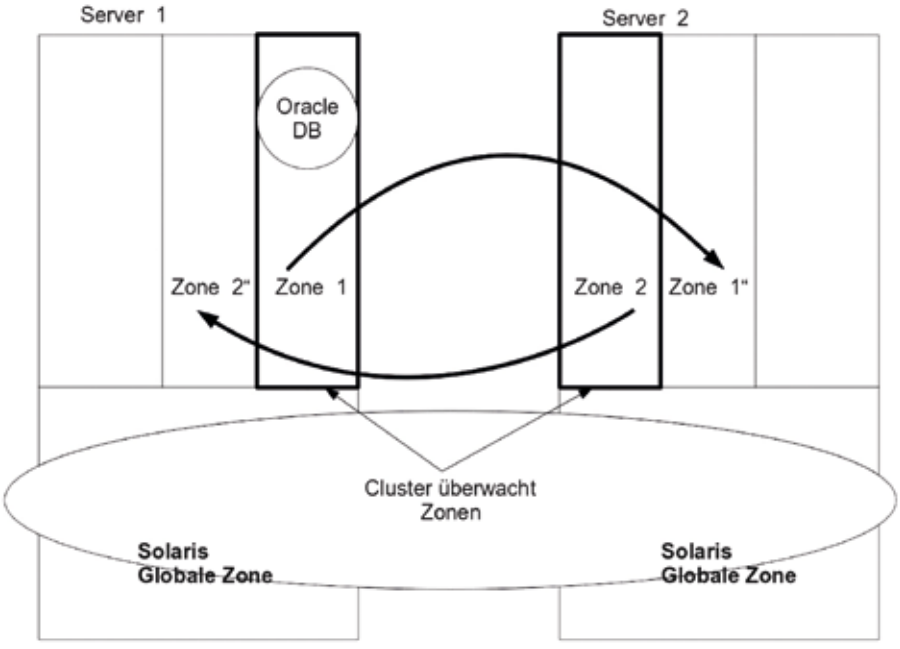

Abbildung 4: Cluster nutzt nicht-globale Zonen als Ressourcen: Zone1 und Zone2 können geschwenkt werden

Kombination verschiedener Virtualisierungstechnologien

Oracle Solaris Zonen sind unabhängig von der darunterliegenden Virtualisierungstechnologie. Ob die Oracle-Solaris-Instanz auf einer physischen Maschine, in einer virtuellen Maschine oder in einer Systems Domain eines M-Klasse-Servers läuft, macht für Zonen keinen Unterschied. Für Zonen-Cluster bedeutet das, dass auch sie in all diesen Kombinationen eingesetzt werden können.

Virtuelle Maschinen, die Clusterknoten sind, zusätzlich als Cluster-Ressourcen zu konfigurieren, die in kritischen Situationen geschwenkt werden können, scheint ein bestechender Gedanke, aber auch eine zu komplexe Realisierung, die heute in keinem dem Autor bekannten Produkt umgesetzt ist.

Live-Migration

Die Möglichkeit, virtuelle Maschinen im laufenden Betrieb auf einen anderen Server zu verschieben, wird häufig als das Mittel zur Verbesserung der Verfügbarkeit angesehen. Doch es gibt nur zwei Aspekte, unter denen Live-Migration dies tatsächlich liefert:

- Wartung der Server-Hardware, die nicht im laufenden Betrieb erfolgen kann
- Wartung des Hypervisors beziehungsweise der Dom0 oder der Control-Domain

In diesen beiden Fällen kann Live-Migration helfen, die „planned Downtime" eines Systems zu verringern, indem vor den Wartungsmaßnahmen das Gastsystem im laufenden Betrieb ohne Unterbrechung auf einen anderen Knoten verlagert wird. In allen anderen Fällen hilft Live-Migration nicht, vor allem bei:

- Patch-Operationen der virtuellen Maschine
- Problemen der virtuellen Maschine
- Anwendungsproblemen
- Fast allen Serverproblemen

Es steht entweder das Quellsystem nicht mehr zur Verfügung, etwa nach einem Hardware-Fehler, oder das virtuelle System hat selbst ein Problem, das durch eine Live-Migration nicht behoben wird.

Es gibt eine kleine Klasse von technischen Problemen, bei denen eine proaktive Verlagerung von virtuellen Maschinen sinnvoll sein kann, etwa bei sich häufenden ECC-Fehlern, die auf einen baldigen Ausfall von Hauptspeichermodulen hindeuten. Hier besteht noch einiges Entwicklungspotenzial.

K-Fall-Absicherung

Ohne in die Tiefen der Anforderungen für eine echte Disaster-Recovery-Lösung zu gehen, wird noch ein wesentlicher Aspekt bei der Verwendung virtueller Umgebung in K-Fall-Lösungen diskutiert. Eine Reihe von Angeboten, virtuelle Umgebungen für den Katastrophenfall abzusichern, beruht auf der Replikation kompletter virtueller Maschinen in ein entferntes Rechenzentrum und dort vorgehaltener Server-Ressourcen, die diese Repliken dann produktiv betreiben können.

Dieses Verfahren, das übrigens auch für Solaris-Zonen möglich ist, hat einen wesentlichen Nachteil: Eine virtuelle Maschine ist ein SPOF. Wenn sie ein Problem hat, beispielsweise durch Fehlkonfiguration, administrative Fehler oder fehlerhafte Patches, lässt sich auch die Kopie nicht mehr betreiben. Ein Rechenzentrums-Betreiber muss selbst entscheiden, welche Katastrophe häufiger vorkommt: Der Ausfall des Rechenzentrums durch Überflutung oder Brand, dann hilft eine „1:1“-Kopie der virtuellen Umgebung, oder ein nicht mehr bootbares Betriebssystem, dann ist auch die Kopie nicht mehr zu verwenden. Eine Bedrohungs-Analyse hilft zu entscheiden, welche Methoden für die K-Fall-Absicherung sinnvoll sind. Die bekannten Regeln für eine Disaster-Recovery-Lösung gelten weiter. Die alternative Umgebung sollte folgende Bedingungen erfüllen:

- So unabhängig wie möglich sein (eine synchron gehaltene Kopie des Originals ist abhängig)
- So synchron wie notwendig sein

Welche Alternative gibt es zur Replikation kompletter virtueller Umgebungen? Vor allem die Vorhaltung komplett getrennter virtueller Maschinen oder Zonen, die nicht durch Fehler in den primären Umgebungen beeinträchtigt werden können. Für die Daten stehen dann die in Oracles Maximum Availability Architecture (MAA) beschriebenen Technologien zur Verfügung. An erster Stelle steht Oracle Data Guard, wenn es um die Oracle-Datenbank geht, das eine unabhängige Schatten-Datenbank verwaltet, wenn notwendig im synchronen Modus, und zusätzlich Mechanismen bereitstellt, um etwaige Daten-Inkonsistenzen durch Zurücksetzen auf einen alten, konsistenten Stand zurückzufahren.

Fazit

Virtuelle Umgebungen sind nicht per se hochverfügbar. Sie können aber mit Cluster-Technologien hochverfügbar gemacht werden. Dabei stehen zwei grundsätzlich verschiedene Methoden zur Wahl, deren Verständnis notwendig ist, um zu einer für den Betrieb passenden Lösung zu kommen:

- die Verwendung virtueller Umgebungen als Clusterknoten
- die Verwendung virtueller Umgebungen als Cluster-Ressourcen

Oracles Technologien wie Oracle VM Server, Oracle Clusterware, Oracle Solaris Cluster und Oracle Solaris Cluster Geographic Edition bieten die grundlegenden Möglichkeiten, diese Methoden sicher zu implementieren.

Hartmut Streppel
hartmut.streppel@oracle.com

Automatisierte SAP-Systemkopien auf Knopfdruck

Libelle SystemCopy

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

... mit deutlich reduzierten Prozesskosten

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