

Disaster Recovery mit der Oracle Solaris Cluster Geographic Edition

**Hartmut Streppel
Oracle Deutschland B.V. & Co. KG
München**

Schlüsselworte

Business Continuity, K-Fall, Disaster Recovery, Cluster, Oracle Solaris Cluster Geographic Edition

Einleitung

Die meisten IT-Betreiber sind sich darüber im Klaren, dass Systeme in ihrem Rechenzentrum aufhören können, richtig und sinnvoll zu arbeiten. Um sich gegen solche Ausfälle abzusichern, werden Systeme redundant aufgebaut und betrieben, so dass einzelne Ausfälle nicht zu einem Stillstand komplexer Systeme führen können. Automatisierte Verfahren, z.B. Cluster, werden eingesetzt, um Fehler zu erkennen und automatisch auf Ersatzkomponenten oder – systeme umzuschalten.

Was aber, wenn die Ausfälle „Dinge“ betreffen, die nur einmal vorhanden sind

- ein kompletten Clusters
- Daten (d.h. Datenkorruption)
- Rechenzentrum (z.B. Überflutung legt alle Brandschutzabschnitte lahm)?

Dagegen hilft nur eine Disaster Recovery (DR) Lösung (K-Fall Absicherung), die in eine unternehmensweite Business Continuity (BC) Lösung integriert ist. Dabei sichert die DR-Lösung die IT-Infrastruktur so ab, dass selbst in einer Katastrophen-Situation der IT-Betrieb sehr schnell und ohne großen Datenverlust wieder aufgenommen werden kann, und somit die wichtigen Geschäftsprozesse weiter aufrecht erhalten werden können.

Im Folgenden soll eine DR-Lösung auf der Basis der Oracle Solaris Cluster Geographic Edition dargestellt werden.

Anforderungen an eine DR-Lösung

Die wesentlichste Anforderung an eine DR-Lösung ist eine möglichst weitgehende Unabhängigkeit der DR-Infrastruktur von der primär genutzten. Das klingt einfach, ist aber, wenn man den Gedanken der Unabhängigkeit zu Ende denkt, sehr schwer.

Der erste Punkt, der fast immer umgesetzt wird, ist eine Aufteilung kritischer Infrastrukturkomponenten auf mehrere Rechnerräume, die häufig auch noch geographisch weiter entfernt sind. Dies sichert gegen Katastrophen ab, die das RZ direkt betreffen. Der zweite wesentliche Punkt ist die normalerweise nur einfach vorhandene logische Kopie der Daten. Daten sind zwar mit RAID-Verfahren redundant gehalten, sollte aber der Inhalt zerstört werden, hilft auch das beste RAID-Level nicht. Hier hilft nur eine unabhängige Kopie. Um auch eine Zerstörung von Daten abzusichern, die z.B. durch fehlerhafte administrative Tätigkeiten erfolgt, sind Verfahren notwendig, die entweder eine Art „verzögerter“ Replikation durchführen, oder aber einen Sprung in die Vergangenheit ermöglichen, mit dem Fehler rückgängig gemacht werden können.

Weiterhin sollen keine Cluster über primäres und DR-RZ betrieben werden, um auszuschließen, dass ein fehlerhaft funktionierendes Cluster beide Infrastrukturen lahm legt. SAN und Netz sollten weitestgehend getrennt sein, um zu verhindern, dass fehlerhafte FC- oder Netz-Switches den Betrieb im DR-RZ beeinträchtigen können.

Hier wird es schon schwierig: Nutzt man die gleiche Technologie mit der gleichen Software in beiden Infrastrukturen, kann es bei einem Softwarefehler dazu führen, dass beide Standorte betroffen sind. Nein, das ist keine Theorie, sondern schlimme Erfahrung. Diese Überlegung kann man weiterführen über Betriebssysteme bis hin zur Anwendungssoftware. Und das Geld, das die NASA für die IT-Infrastruktur ihrer Raumfahrtprogramme ausgegeben hat, um jegliche Funktionalität dreifach zu implementieren, um in jeder Situation eine Mehrheitsentscheidung über das richtige Ergebnis herbeizuführen, mag kaum jemand ausgeben.

Metro Cluster

Oracle Solaris Cluster ist Oracles Hochverfügbarkeitstechnologie. Sie überwacht Hardware und Software innerhalb eines Clusters und ist in der Lage, Aktionen auszulösen, die dafür sorgen, dass hochverfügbare Dienste nach einem Fehler ohne oder mit minimaler Unterbrechung zur Verfügung stehen.

Eine solche Clusterlösung kann auch in sog. Metro Clustern verwendet werden, d.h. Infrastrukturen, die sich über mehrere RZs, die auch weiter entfernt sein können, erstrecken. Solche Metro – oder auch „stretched“ Cluster – sichern allerdings nicht gegen Katastrophen wie oben genannt ab.

Eine weitestgehende Unabhängigkeit ist nur gegeben, wenn Cluster sich gerade nicht über mehrere Standorte erstrecken.

Die Oracle Solaris Cluster Geographic Edition

Die Oracle Solaris Cluster Geographic Edition (OSC Geo) bietet genau die Unabhängigkeit, die für eine richtige DR-Lösung gefordert ist. Die Verbindung zwischen Ausweich- und Produktions-RZ basiert nur auf IP über ein – in der Regel - „öffentliches“ Netz. Über dieses wird eine Art Heartbeat betrieben, der allerdings, da es sich nicht um einen traditionellen Cluster Interconnect und in der Regel größere Entfernungen handelt, mit relativ hohen Timeout Werten überwacht wird.

Ist eins der Cluster nicht mehr erreichbar, wird eine Alarmierung ausgelöst, die den Betrieb veranlassen, zu untersuchen, was für ein Problem aufgetreten ist. Es ist eine bei DR-Experten allgemein anerkannte Best Practice, in einem solchen Problemfall NICHT automatisch einen Schwenk einer u.U. sehr komplexen Anwendung auslösen zu lassen. Stattdessen muss, nachdem das die Alarmierung auslösende Problem analysiert wurde, ein Entscheidungsprozess ablaufen, an dessen Ende eine eindeutige Entscheidung steht, hochverfügbare, unternehmenskritische Anwendungen entweder im Ausweich-RZ zu starten, oder aber die Behebung des ursächlichen Problems zu betreiben.

Interessanterweise benötigt man für eine solche Entscheidung auch eine Mehrheit aus 3 Stimmen, ähnlich wie ein Quorum-Algorithmus im Cluster, bei dem auch – allerdings automatisiert - die Mehrheit aus 3 Stimmen gesucht wird. Nehmen wir an, dass ein RZ nicht mehr erreichbar ist, auch nicht über Satellitentelefon; in diesem Fall ist eine Entscheidung einer zweiten Person im Ausweich-RZ nicht ausreichend, denn die Person im nicht mehr erreichbaren RZ könnte ebenfalls der Meinung sein, sie könne eine Entscheidung treffen. Lösung: es muss eine definierte dritte Person die Entscheidung mit einer der beiden anderen treffen.

Partnerships and Protection Groups

Mit Hilfe der Oracle Solaris Cluster Geographic Edition werden Partnerschaften zwischen Clustern etabliert. Dabei kann ein Cluster Partnerschaften mit mehreren anderen Clustern haben, d.h. es sind 1:N und N:1 Lösungen implementierbar.

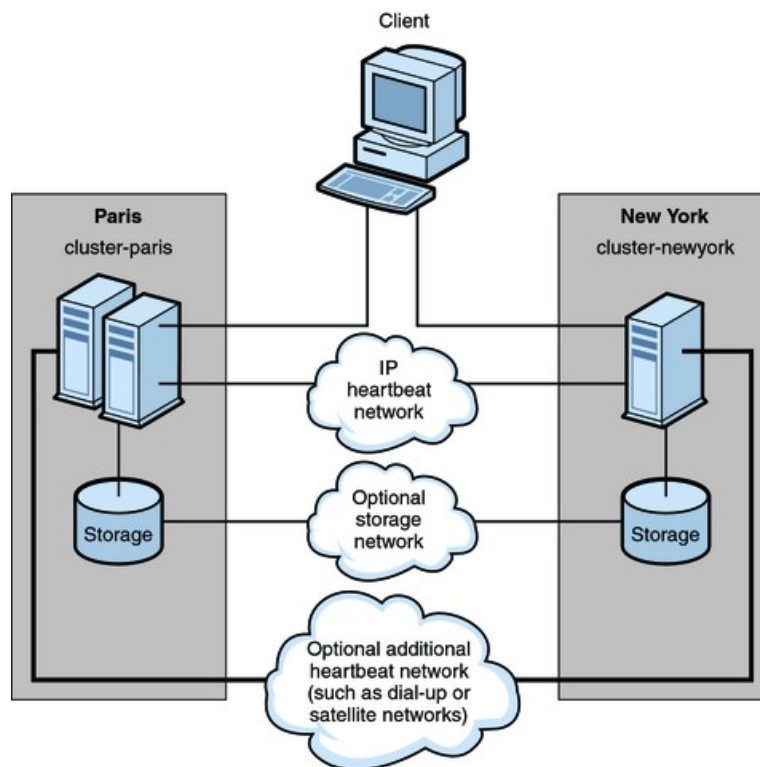


Abb. 1: Oracle Solaris Cluster Geographic Edition Konfiguration

Die wesentlichste Einheit, die von der OSC Geo verwaltet wird, ist die „protection group“ (PG). In ihr werden zwei Elemente zusammengefasst: eine Oracle Solaris Cluster Ressource Gruppe (RG) und eine Data Replication Device Group (DG), bzw. eine Replikationskomponente (RC), wenn eine ZFS Storage Appliance verwendet wird..

Die DG/RC ist hierbei verantwortlich für die Replikation der Datenbereiche, die die hochverfügbare Anwendung in der RG im Ausweich-RZ benötigt, um dort neu mit den aktuellen Daten gestartet zu werden. Als Replikationstechnologien werden derzeit unterstützt;

- ZFSSA Replikation
- Oracle Data Guard
- MySQL Replikation
- Sun Availability Suite
- EMC SRDF
- HDS TrueCopy/Universal Replicator (nur OSC Geo 3.x).

Dies sind 3 sehr unterschiedliche Replikationstechnologien: Host basierte Block-Replikation, Storage basierte Block-Replikation und applikations basierte Replikation.

Switchover und Takeover

Es gibt 2 Verfahren eine Protection Gruppe zwischen Clustern zu migrieren. Achtung: Live Migration

ist vor allem wg. der geforderten Unabhängigkeit der Umgebungen keine sinnvolle Lösung für eine DR-Umgebung. Switchover ist die freundliche Übernahme einer PG. Hier werden die zur PG gehörende Replikations-Gruppe und die Cluster Ressource Gruppe zunächst gestoppt und dann auf der Zielumgebung in umgekehrter Reihenfolge wieder gestartet. Dabei wird auch die Replikationsrichtung geändert. Am Ende des Switchover Prozesses sind beide Umgebungen weiterhin intakt, die Replikation funktioniert in die korrekte Richtung und die Anwendung ist nach kurzer Unterbrechung wieder aktiv.

Im K-Fall ist die alte produktive Umgebung in der Regel nicht mehr erreichbar, ein Switchover also nicht möglich. Stattdessen wird ein Takeover durchgeführt, der sowohl die Replikations-Gruppe als auch die Ressource Gruppe auf dem Ziel-Cluster startet. Dabei wird auch die Replikationsrichtung umkonfiguriert, aber in der Regel auf Grund der ausgefallenen ehemaligen primären Umgebung, nicht aktiviert. Es ist offensichtlich, dass bei diesem Verfahren die neue und die alte Umgebung auseinander laufen. D.h. es ist notwendig, die alte primäre Seite wieder auf den aktuellen Stand zu bringen, nachdem sie wieder aktiviert wurde.

Um zu verdeutlichen, dass der Schwenk einer PG (hier die `oracle-pg`) einfach ausgelöst werden kann, soll das notwendige Kommando dargestellt werden. Hinter der `oracle-pg` kann sich eine komplette HA Oracle Ressource Gruppe verbergen plus einer dazugehörigen Oracle Data Guard Konfiguration. Alle Komponenten werden mit diesem einen Befehl umgezogen:

```
geopg switchover -m <ziel cluster> oracle-pg
```

Das entsprechende takeover Kommando, das ausgeführt wird, wenn das produktive Cluster nicht mehr erreichbar ist, sieht so aus:

```
geopg takeover oracle-pg
```

Da ein Takeover nur in einer echten K-Fall-Situation in Frage kommt, kann das Kommando nur auf dem neuen Zielcluster ausgeführt werden. Damit ist das Ziel implizit Teil des Kommandos.

Bewertung

Es ist offensichtlich, dass der wesentliche Wert der OSC Geographic Edition nicht in der Implementierung komplexer neuer Funktionalität liegt, sondern in der Zusammenfassung administrativer Schritte, die aufeinander aufbauen, in der richtigen Reihenfolge ausgeführt werden müssen, in der Regel aber sehr unterschiedliche Expertise benötigen. Diese Schritte sauber zu dokumentieren, alle möglichen Fehlerfälle zu besprechen und dann die Administratoren entsprechend zu trainieren, so dass im K-Fall auch das gesamte Wissen zur Verfügung steht, ist ein nicht zu unterschätzender Aufwand, der beim Einsatz eines Produkts wie der Oracle Solaris Cluster Geographic Edition zum großen Teile nicht notwendig ist.

Weitere Anwendungsmöglichkeiten

Vielleicht noch wichtiger als Ausweichmechanismus für hochverfügbare Anwendungen im K-Fall zu sein, ist die Möglichkeit, die OSC Geographic Edition als Vehikel zu verwenden, komplexere Software Updates durchzuführen. Da OSC Geo ein sog. lose gekoppeltes System ist, können in einer OSC Geo Partnerschaft durchaus Cluster und OSC Geo unterschiedlicher Release Stände vorhanden sein. Ähnliches gilt für Oracle Data Guard, bei dessen Einsatz es nicht erforderlich ist, dass die beteiligten Datenbanken in der identischen Version betrieben werden.

D.h. es gibt mit der Einführung einer DR-Infrastruktur die Möglichkeit, auch sog. „major“ Upgrades,

sowohl auf der Ebene von Oracle Solaris und Oracle Solaris Cluster, aber auch der Oracle Datenbank durchzuführen. Dabei wird zunächst die K-Fall Infrastruktur auf den neuesten Stand gebracht. Dies kann u.U. auch eine vollständige Neuinstallation bedeuten, da z.B. kein Upgrade von Solaris 10 auf Solaris 11 möglich ist. Danach wird das neue lokale Cluster und darauf OSC Geo aufgesetzt. OSC Geo ist nun in der Lage, auf der K-Fall Infrastruktur mit einem neuen Release betrieben zu werden, aber eine Partnerschaft mit dem alten Release in der Produktions Infrastruktur zu bilden. In dieser OSC Geo Partnerschaft ist es nun möglich, mit einem „geopg switchover“ Kommando eine OSC Ressource Gruppe aus dem Produktionscluster ins K-Fall-Cluster, d.h. vom alten Release ins neue zu migrieren. Und das Ganze mit minimaler Unterbrechung.

Ein solcher Anwendungsfall ist ausführlich unter [3] beschrieben.

Zusammenfassung

Eine DR-Lösung auf der Basis der Oracle Solaris Cluster Geographic Edition ist eine hochverfügbare Infrastruktur, die

- weitestgehende Unabhängigkeit der DR-Umgebung von der primären
- Überwachung der lokalen Cluster
- einfache, automatisierte Migration von OSC Ressource-Gruppen
- Verwaltung der zur Anwendung gehörenden Replikation

mit einfacher Administration bietet. Damit steht eine im K-Fall leicht und zuverlässig zu bedienende Lösung zur Verfügung, die einen komplexen Schwenkvorgang inklusive der Umkehrung der Replikationsrichtung schnell und sicher auf Knopfdruck durchführt.

Mehr Informationen

- Oracle Solaris Cluster Geographic Edition Overview - http://docs.oracle.com/cd/E29086_01/html/E29352
- Oracle Solaris and Oracle Solaris Cluster: Extending Oracle Solaris for Business Continuity - <http://www.oracle.com/technetwork/server-storage/solaris-cluster/documentation/solaris-cluster-businesscontinuity-168285.pdf>
- [3] How to Upgrade to Oracle Solaris Cluster 4.0 - <http://www.oracle.com/technetwork/articles/servers-storage-admin/upgrade-cluster4-0-1635866.html>

Kontaktadresse:

Hartmut Streppel
Oracle Deutschland B.V. & Co. KG
Riesstr. 25
D-80992 München

Telefon: +49 (0) 89-1430 2588
Fax: +49 (0) 89-1430 1150
E-Mail: Hartmut.Streppel@oracle.com
Internet: www.oracle.de