

RAC News in Oracle 11g Release 2

Dierk Lenz
Herrmann & Lenz Services GmbH
Burscheid

Schlüsselworte:

Real Application Clusters, RAC, Database, 11g Release 2, New Features, Grid Infrastructure

Einleitung

Seit dem 1. September 2009 – also etwas mehr als einem Jahr – gibt es nun die neue Version der Oracle Datenbanksoftware inklusive der Grid Infrastructure. Wurden im letzten Jahr noch Erfahrungen aus dem Beta-Test präsentiert, so können in diesem Jahr bereits Erfahrungen aus diversen Installationen wiedergegeben werden.

Verfügbarkeit

Die erste freigegebene Plattform war bekanntlich Linux (in den Varianten 32 Bit und 64 Bit). Bis zum Jahresende folgten dann weitere Unix-Portierungen (HP-UX Itanium, Solaris, AIX). Windows-Anhänger mussten bis zum April dieses Jahres warten, bekamen dafür aber gleich auch die ersten Oracle-Produkte für die neuen Betriebssysteme Windows Server 2008 Release 2 und Windows 7.

Für Linux ist soeben das erste Patchset 11.2.0.2 erschienen, das für die anderen Plattformen bis zum Jahresende noch erscheinen soll. Interessant ist, dass dieses Release für das erste Quartal 2011 auch für HP OpenVMS Itanium angekündigt ist, nachdem es 11g Release 1 hierfür nicht gegeben hat.

Eine RAC-Installation besteht aus der Grid Infrastructure und der Database. Bemerkenswert ist, dass es für Windows zwar eine 32 Bit Database, aber keine Grid Infrastructure gibt. Ein Windows-RAC ist also nur mit 64 Bit-Windows möglich. (Hier spielt sicherlich die Microsoft-Strategie eine Rolle, Windows Server 2008 Release 2 nur als 64 Bit-Version zu veröffentlichen; ältere Windows Server-Versionen gibt es aber immer noch mit 32 Bit.)

Installation

Die Clusterware und ASM sind ab 11.2 Bestandteil der Grid Infrastructure, die vor der Database Software installiert werden muss. Neu ist ebenfalls, dass die Grid Infrastructure auch auf Einzelknoten installiert werden kann. In diesem Fall werden Oracle Restart und ASM installiert. Restart sorgt für den automatischen Start von Komponenten wie Listener, ASM und DB-Instanzen und bringt ein bisschen „Cluster-Feeling“ mit: Die Definition und Administration erfolgt wie beim Cluster mit dem Werkzeug srvctl.

Ganz neu ist die Vorgehensweise bei Patchsets, also erstmalig bei Version 11.2.0.2. Das Patchset ist nun ein vollständig neues Installationspaket. Somit muss für eine Neuinstallation nicht mehr langwierig ein Basisrelease und das passende Patchset installiert werden; man kann sofort mit der gewünschten Version beginnen.

Für bestehende Installationen kann man sich aussuchen, ob man den Upgrade „in place“ oder „out of place“ durchführt. Oracle selbst rät ganz entschieden zum out-of-place-Upgrade. Das hat sicher den Vorteil, ein wesentlich saubereres Oracle-Home zu hinterlassen. Allerdings braucht man während des Upgrades den doppelten Platz für die Software-Installation. Für die Grid Infrastructure ist der out-of-place-Upgrade obligatorisch.

ASM Cluster File System

Mit dem ASM Cluster File System ist es Oracle gelungen, dem Anspruch „Alle Daten werden in Oracle gespeichert“ ein wenig näher zu kommen. Es geht zwar hier nicht um die Speicherung in einer Datenbank, sondern ASM wurde um entsprechende Treiber und Funktionen erweitert.

Resultat dieser Funktionalität ist, dass innerhalb einer ASM-Diskgruppe ein Volume angelegt werden kann, das mit einem ACFS-Dateisystem versehen und auf jedem Clusterknoten gemountet werden kann. ACFS ist ein schnelles, Cluster-fähiges Dateisystem, das auch mit großen Volumes sehr gut umgehen kann.

ACFS kann z.B. genutzt werden, um eine Database-Installation zentral abzulegen und für alle Knoten gemeinsam zu nutzen. Zu einer solchen Vorgehensweise ist zwar aufgrund von Verfügbarkeitsüberlegungen bei Upgrades generell abzuraten; allerdings wird ein zentrales Oracle-Home z.B. bei SAP-Installationen vorausgesetzt.

Für die zentrale Speicherung von Skripten, Ladedaten usw. hat sich ACFS allerdings sehr bewährt. Bei vielen Installationen werden z.B. Ladeprozesse grundsätzlich vom gleichen Knoten gestartet. Nur im Fehlerfall wird ein anderer Knoten hierzu verwendet. Oft wird erst dann eine Synchronisation der Skripte, Konfigurationsdaten usw. durchgeführt. Hat man jedoch alle wichtigen Komponenten zentral abgelegt, so entfällt das Synchronisieren.

Einziges Wermutstropfen bei ACFS ist, dass dieses Feature noch nicht auf allen Plattformen verfügbar ist. Anfangs konnte man lediglich Oracle Enterprise Linux, Red Hat Enterprise Linux und Windows Server 2003 verwenden; für 11.2.0.2 sind weitere Plattformen angekündigt: SLES10, Solaris und AIX.

Single Client Access Name (SCAN)

RAC-Installationen erfreuen sich wachsender Beliebtheit – Netzwerk- und IO-Infrastrukturen verbessern sich ständig, so dass mit überschaubaren Investitionen Hardware für ein Cluster mit guter Performance bereitgestellt werden kann. In früheren Versionen führte ein Punkt bei der Anbindung des RAC-Systems zu Verwirrung: Der TNSNAMES-Eintrag für Clients musste immer eine Adressliste mit allen virtuellen IP-Adressen (VIPs) enthalten. Hier ein Beispiel:

```
RAC.HL.DE =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP) (HOST = 192.168.233.113) (PORT = 1521))
      (ADDRESS = (PROTOCOL = TCP) (HOST = 192.168.233.114) (PORT = 1521))
      (ADDRESS = (PROTOCOL = TCP) (HOST = 192.168.233.115) (PORT = 1521))
      (ADDRESS = (PROTOCOL = TCP) (HOST = 192.168.233.116) (PORT = 1521))
    )
    (CONNECT_DATA =
      (SERVICE_NAME = rac.hl.de)
    )
  )
```

Das hat mehrere gravierende Nachteile:

1. Die tnsnames.ora-Datei muss auf jeden Fall geändert werden. Da die Dateien oft auf hunderten Clients verteilt sind, kann das zu sehr aufwendigen Aktionen führen.
2. Nur der Oracle-eigene Thin-JDBC-Treiber kann Oracle-TNS-Syntax für den Verbindungsaufbau interpretieren – andere JDBC-Treiber bleiben „außen vor“.
3. Ändert sich die RAC-Konfiguration, so müssen die TNSNAMES-Einträge nachgezogen werden.

Insbesondere beim letzten Punkt bestand für Oracle Handlungsbedarf, da ein neues 11gR2-RAC-Feature die automatische Provisionierung ist, d.h. eine RAC-Datenbank kann sich automatisch in einem Cluster ausdehnen.

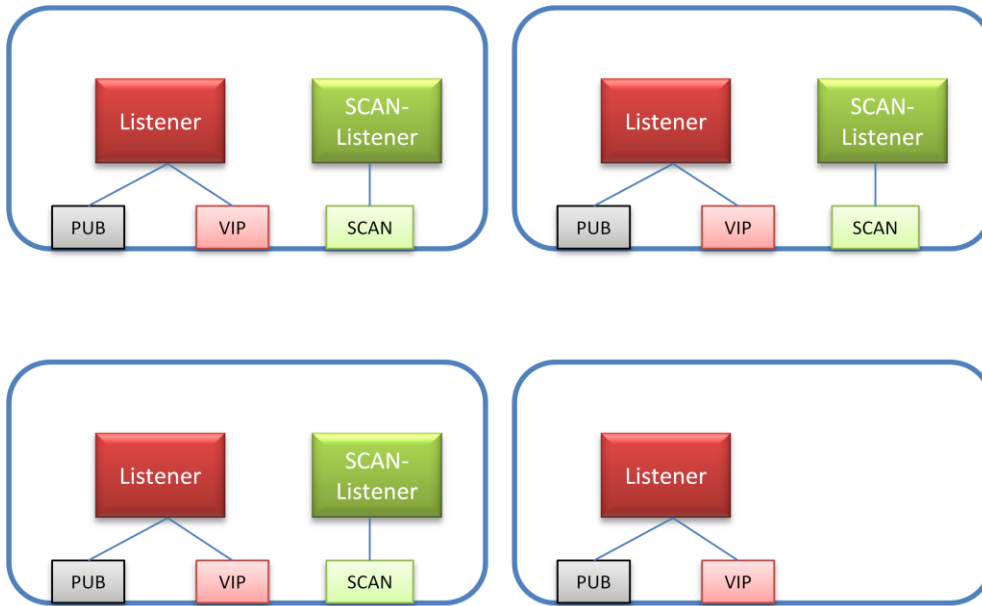
Die zentrale Frage lautet also: Wie bekommt man einen „normal“ aussehenden TNSNAMES-Eintrag für RAC-Umgebungen hin, der weiterhin alle relevanten Anforderungen erfüllt (Hochverfügbarkeit und Skalierbarkeit)?

Oracles Antwort lautet: SCAN! Der SCAN ist zunächst ein Name, der im DNS konfiguriert sein muss, und zwar mit drei IP-Adressen, die zyklisch bei der DNS-Abfrage zurückgegeben werden. Der SCAN wird für die TNSNAMES-Einträge genutzt, so dass dieser wie folgt aussieht:

```
RAC.HL.DE =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP) (HOST = rac-scan) (PORT = 1521))
    )
    (CONNECT_DATA =
      (SERVICE_NAME = rac.hl.de)
    )
  )
```

Die zum SCAN konfigurierten IP-Adressen werden von der Grid Infrastructure als virtuelle IP-Adressen für das Cluster eingerichtet. Um sie von den (weiterhin vorhandenen) Knoten-VIPs zu unterscheiden, werden diese auch als SCAN-VIPs bezeichnet. Im Unterschied zu den Knoten-Listnern, die auf der Public-IP und der VIP aufsetzen und somit ausschließlich auf dem „richtigen“ Knoten laufen, setzen die neu hinzugekommenen SCAN-Listener auf je einer SCAN-VIP auf und können zwischen den Knoten hin- und hergeschaltet werden.

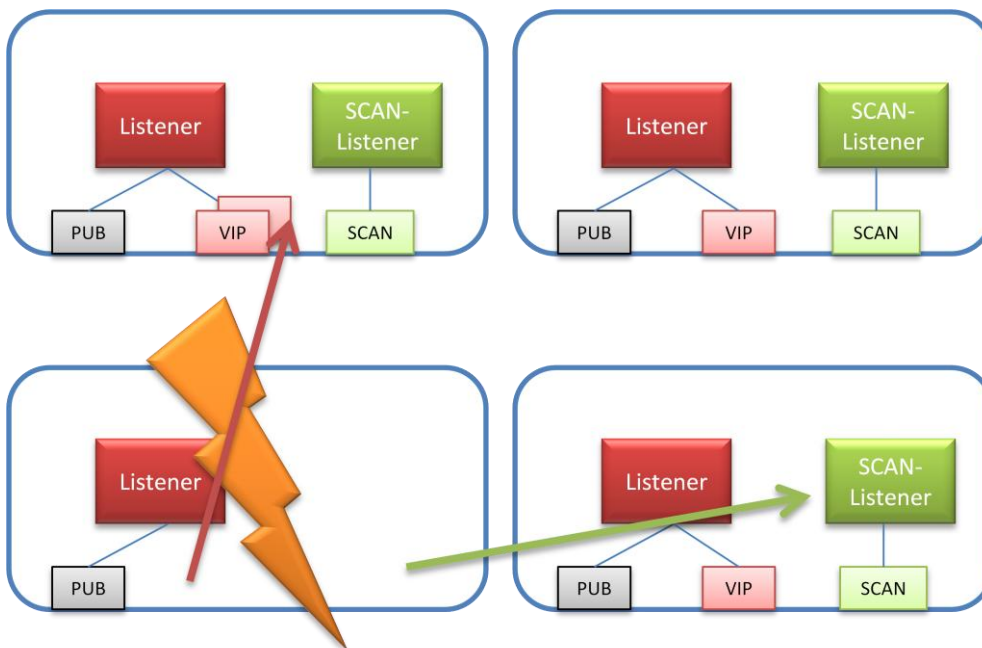
In einem Vier-Knoten-Cluster stellen sich die Listener insgesamt wie folgt dar:



Unabhängig von der Zahl der Knoten gibt es drei SCAN-VIPs und drei dazugehörige SCAN-Listener (LISTENER_SCAN1, LISTENER_SCAN2, LISTENER_SCAN3). Die SCAN-Listener werden automatisch gestartet und folgen der SCAN-VIP.

Damit die SCAN-Listener richtig funktionieren, werden sie bei den RAC-Instanzen als Remote Listener konfiguriert. Die Knoten-Listener werden für jede Instanz als lokaler Listener konfiguriert. Somit registrieren die Instanzen alle vorhandenen Datenbank-Services bei den SCAN-Listnern, die damit in der Lage sind, für das ganze Cluster passende Serverprozesse zu starten.

Wichtig ist nun das Verhalten beim Herunterfahren oder Ausfall eines Knotens:



Die Knoten-VIP wird von der Clusterware auf einen beliebigen anderen Knoten verschoben – ohne hierfür einen Listener zu starten. Die SCAN-VIP wird jedoch zusammen mit dem SCAN-Listener auf einen anderen Knoten verschoben, so dass nach kurzer Zeit alle SCAN-Listener wieder zur Verfügung stehen.

Oracle Net Konfiguration

Sobald man mit einer Grid Infrastructure arbeitet (ob im Cluster oder auf einem Einzelknoten), werden alle Listener aus dem Grid-Home gestartet und nicht mehr aus dem Database-Home. Das hat zur Folge, dass die Datei tnsnames.ora in diesen beiden Oracle-Homes synchron gehalten werden muss.

Schlussbemerkung

Nachdem es bei 11g Release 1 für RAC kaum Unterschiede zur Vorgängerversion gab, hat Oracle mit 11g Release 2 ordentlich „Gas gegeben“. Dass es bei größeren Umstellungen – wie der Einführung der Grid Infrastructure – zu einigen Stolperfallen bezüglich älterer Versionen kommt, war fast zu erwarten. Da das Ziel für ein RAC-System aber immer eine möglichst homogene Landschaft ist und die Änderungen viele Vorteile bringen, ist der Sprung ins kalte Wasser aber unter dem Strich empfehlenswert. Wie (fast) immer muss man sich mit den Neuerungen auseinandersetzen – um SCANs kommt man mit 11g Release 2 RAC nicht herum.

Kontaktadresse:

Dierk Lenz

Herrmann & Lenz Services GmbH
Höhestraße 37
51399 Burscheid

Telefon: +49 2174 6712 0
Fax: +49 2174 6712 12
E-Mail: dierk.lenz@hl-services.de
Internet: www.hl-services.de