

# Oracle Real Application Cluster 12c inkl. ASM und ACFS

Sebastian Solbach  
Oracle Deutschland b.v. & Co.Kg  
München

## Schlüsselworte

12c, RAC, ASM, ACFS, Real Application Clusters, Automatic Storage Management, Grid Infrastruktur, ASM Cluster Filesystem

## Einleitung

Oracle 12.1.0.2 hat nicht nur bei der Datenbank viele Neuerungen eingeführt, auch im Bereich der darunterliegenden Grid Infrastruktur hat sich mit dem letzten Patch Set der Oracle Datenbank viel getan. Dabei gingen viele der Änderungen und nützlichen Tools leider im Trubel von InMemory, JSON und Multitenant unter. Dennoch liefert die Grid Infrastruktur viele Teile, um auch diese 12c Datenbanken in Zukunft hochverfügbar und ausfallsicher zu betreiben.

## 12c Real Application Clusters im Rahmen der MAA

Oracle 12c RAC ist ein Bestandteil der Maximum Availability Architektur(1) von Oracle.

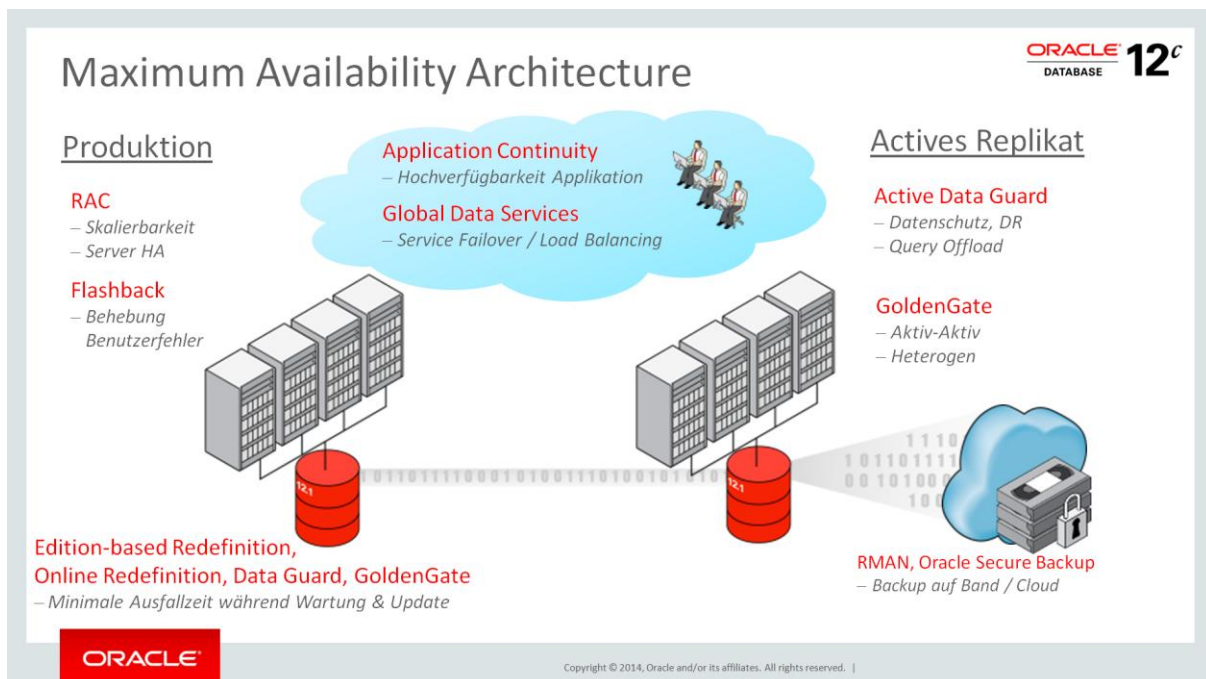


Abb.1: MAA Architektur Oracle

Die MAA Architektur wurde mit 12c dabei um 2 wichtige Technologien erweitert, die insbesondere Benutzer betreffen, die auf einen RAC zugreifen: Transaction Guard, Application Continuity und Global Data Services.

Schon in 11g konnten geplante Ausfälle mit Hilfe von Fast Application Notification (FAN) und Connection Pools abgedeckt werden. Dazu wurden die Connection Pools der Applikationen über unterschiedlichste Mechanismen (Advanced Queuing oder Oracle Notification Server, kurz ONS) benachrichtigt, wenn sich auf dem Cluster eine Änderung ergab. Hierzu musste der Connection Pool aber die entsprechenden Informationen erhalten. Dies erforderte unter anderem die Angabe der einzelnen RAC Knoten. Mit 12c ist die Funktion weiterhin vorhanden, die Integration wurde aber vereinfacht: So werden einerseits keine unterschiedlichen Mechanismen zur Benachrichtigung mehr benötigt, sondern nur noch der ONS, andererseits ist auch die Angabe der einzelnen Knoten überflüssig, da diese Information nun die Listener beim Verbindungsaufbau mit übermitteln.

Transaction Guard (2) und Application Continuity (3) ermöglichen Applikationen nun auch mit ungeplanten Ausfällen umzugehen. So bietet 12c zum ersten Mal die Möglichkeit für eine Applikation den Transaktionsstatus einer Transaktion nach einem Abbruch der Verbindung zur Datenbank festzustellen. Diese als Transaction Guard bezeichnete Technologie steht allen Clients zur Programmierung zur Verfügung und wurde bei JDBC Thin Treiber in eine fertige Lösung verpackt: Application Continuity kurz AC. AC erlaubt es, dass Applikationen ohne spezielle Programmierung nach dem Ausfall einer Verbindung zum RAC Server die Transaktionen wiederholen können. Implizit verwendet AC hierbei die von Transaction Guard bereitgestellte Funktionalität.

Global Data Services (GDS) hingegen bezeichnet die Möglichkeit, einen zentralen Listener für das Unternehmen bereitzustellen, der Benutzer auch bei verteilten und replizierten Datenbanken an die beste geeignete Instanz weiterleitet. GNS ist also praktisch wie ein Single Client Access Name (SCAN) Listener für alle Datenbanken, nicht nur für einen Cluster.

Viele weitere Änderungen mit 12c und mit dem letzten Patchset 12.1.0.2 erkennt man schon bei der Installation der Grid Infrastruktur (GI), daher wird der Installationsprozess der GI nun näher betrachtet.

### **12c Real Application Clusters anhand der Installation**

Die Vorbereitungen der Grid Infrastruktur (GI) Installation sind auch in 12c sehr umfangreich, da eine Clusterware tief in das Betriebssystem eingreift. Wie bisher auch gibt es neben der offiziellen Dokumentation auf <http://docs.oracle.com> auch die sogenannten RAC Starter Kits (My Oracle Support, kurz MOS, Note 810394.1) (4), die Best Practices für die Installation auflisten. Unter Linux fällt in diesem Zusammenhang auf, dass es für die 12c Datenbank nun ein neues RDBMS Preinstall Paket gibt (ehemals Oracle-Validated). Sind die Vorbereitungen getroffen startet man für die Installation den Oracle Universal Installer.

Wie bisher auch hat man nun die Wahl ob die GI als Cluster oder nur für Single Instanz Datenbanken verwendet wird. In diesem Zusammenhang sollte auf MOS Note 1584742.1 (5) „Support Impact of the Deprecation Announcement of Oracle Restart with Oracle Database 12c“ hingewiesen werden, in der der Umstand erläutert ist, warum in Zukunft Oracle Restart in dieser Art nicht mehr verfügbar sein wird.

Nach dieser Auswahl steht die Entscheidung für einen Flex Cluster oder einen Standard Cluster an. Ein Standard Cluster ist ein Cluster, wie er auch zu 11g Zeiten schon bekannt war. Ein Flex Cluster vereint sogenannte Leaf Clusterknoten - Knoten die keinen Zugriff auf das Storage Subsystem haben, mit den Hub Knoten, die die Funktion herkömmlicher Datenbankclusterknoten übernehmen. Die Idee hinter einem Flex Cluster ist, Applikationsserver unter die Überwachung der Clusterware zu stellen, damit Abhängigkeiten zu einer Datenbank definiert werden können. Da diese Applikationen keinen Zugriff auf das shared Storage brauchen, platziert man diese im Flex Cluster auf den sogenannten Leaf Knoten. Dass die Clusterware in 12c ideal geeignet ist auch generische Applikationen zu überwachen,

erkennt man schon an dem neuen Typ des „generic\_application“. So können Applikationen, die über eine Prozess ID (PID) Datei zur Statusüberwachung verfügen, mit einem Befehl unter die Überwachung der Clusterware gestellt werden:

```
$ crsctl add resource sambal -type generic_application -attr
"START_PROGRAM='/etc/init.d/smb start',
STOP_PROGRAM='/etc/init.d/smb stop', CLEAN_PROGRAM='/etc/init.d/smb
stop', PID_FILES='/var/run/smbd.pid,/var/run/nmbd.pid'"
```

Einige Bildschirme im Installer später werden die Informationen für den SCAN Listener abgefragt. Details zum SCAN findet man im Whitepaper „Oracle RAC – Overview of SCAN“ auf OTN. Neu mit 12c ist die Unterstützung mehrerer SCANS in unterschiedlichen publik Subnetzen. Allerdings kann dies nicht direkt bei der Installation angegeben werden und wird manuell nachträglich konfiguriert. Neuerdings kann für das Public Netzwerk auch IPv6 verwendet werden und wie bisher auch gibt es die Möglichkeit die Namensauflösung im Cluster an den Grid Naming Service kurz GNS zu delegieren und somit die Anpassungen nicht mehr zwingend im eigenen DNS Server durchzuführen. Allerdings benötigt der GNS in 12c weder Dynamische Netzwerke (DHCP bzw. Stateless Address Configuration) noch die Zone Delegation im DNS, die man bisher brauchte. Allerdings benötigt ein Flex Cluster zwingend diesen GNS, der letztendlich aber dann nur noch eine zusätzliche IP Adresse ist. Ein dedizierter GNS Cluster kann dabei auch die GNS Aufgaben für mehrere Cluster übernehmen.

### **Flex ASM und andere ASM Neuerungen**

Unabhängig von Flex Cluster ist Flex ASM, auch wenn der Name anderes vermuten lässt. Hinter Flex ASM verbirgt sich hauptsächlich die Funktionalität, dass die einzelnen Datenbankinstanzen eines Clusters nicht mehr abhängig von der lokalen ASM Instanz sind. Konkret bedeutet dies, dass der Ausfall einer ASM Instanz keine Auswirkungen mehr für die Datenbanken hat, die sich nun einfach auf eine andere ASM Instanz verbinden können. Sollten allerdings noch 11gR2 Datenbanken im selben Cluster betrieben werden, so sind diese weiterhin von der lokalen ASM Instanz abhängig. Im Unterschied zu 12.1.0.1 hat der letzte Patchset bei Flex ASM übrigens eine Minimum Kardinalität von 2.

Auch sonst bietet ASM 12c einiges mehr: Von 511 Diskgruppen über einen Befehl ASM Platten umzubenennen bis hin zur Speicherung von OCR Backup und Passwordfile direkt in ASM.

Die interessanteste Änderung hat sich aber beim ASM Cluster Filesystem, welches auf Basis von ASM läuft, ergeben. So ist dieses seit September 2014 komplett kostenfrei auch mit den zusätzlichen Funktionalitäten von Snapshots, Replikation, Security, Tagging und Verschlüsselung. Allerdings erfordert der Snapshot einer Datenbank eine Enterprise Edition und steht der Standard Edition nicht zur Verfügung. Aber gerade mit den 1023 Read-Writeable Snapshots und dem Hochverfügbaren NFS Server auf ACFS Basis ist ACFS ein perfektes, platzsparendes Dateisystem für Testdatenbanken.

Für das automatische Anlegen, Klonen und Verwalten von Datenbanken auf ACFS gibt es seit kurzem auf OTN auch das gDBClone Skript, welches das Erstellen von Testdatenbanken auf Basis dieser platzsparenden Redirect-On-Write Snapshot Technologie erheblich vereinfacht. Die Oracle Database Appliance 12.1.2 von Oracle hat eine ähnliche Funktionalität ebenfalls auf Basis des 12c ACFS Filesystems umgesetzt.

Werden Platten für ASM bzw. ACFS zur Verfügung gestellt, wurde früher unter Linux ASMLIB verwendet. Allerdings schützte ASMLIB nicht davor, dass der Systemadministrator die Platten von ASM zerstörte, wenn er die Platten nicht richtig zuordnen konnte und formatierte. Dies soll mit dem neuen 12.1.0.2 ASM Filter Driver (kurz AFD) nun anders werden. AFD steht nicht nur unter Linux

zur Verfügung, sondern auch unter Solaris, AIX und Windows. AFD schützt dabei nicht nur die Platten vor versehentlicher Formatierung, sondern unterstützt implizit auch Thin Provisioned Storage.

### **Cluster Tools (CLUVFY, CHM, TFA, ORACHK)**

Bevor nun das Cluster Verification Utility (CLUVFY) mit seinen erweiterten 12c Checks die Systeme überprüft, ob die Installationsvoraussetzungen erfüllt sind, bietet der 12c GI Installer nun die Ausführung der root.sh Skripte direkt über den Installer an. Das vereinfacht die Nutzung erheblich, allerdings hat man etwas weniger Kontrolle über die Reihenfolge der root.sh Skriptausführung.

CLUVFY begegnet einem generell nun häufiger. So hat auch der komplett überarbeitete Database Configuration Assistance (DBCA) in 12c den CLUVFY integriert. Des weiteren bietet CLUVFY auch noch ein paar andere sehr sinnvolle Funktionen, die nicht unbedingt im Zusammenhang mit einer Installation stehen. So ist es möglich mit CLUVFY eine Baseline einer Installation zu ziehen und die Ergebnisse mit anderen Clustern zu vergleichen. So erkennt man auf Anhieb, wenn es zwischen den Knoten der unterschiedlichen Cluster größere Unterschiede im Setup geben sollte. Gerade sinnvoll, wenn ein funktionierender Cluster mit einem anderen verglichen werden soll. Allerdings ist dies nicht mit dem umfangreichen Change Management Pack des Enterprise Managers zu vergleichen.

Zwar nicht neu bei der Installation der GI ist das Anlegen des Grid Infrastruktur Management Repositories kurz GIMR. Dieses gab es auch schon in 11g, war dort aber noch eine Berkley Datenbank. Erst mit 12c wurden die Daten des Cluster Health Monitors (CHM) in eine Oracle Datenbank geschrieben. CHM bietet dabei für den Cluster ein paar sehr hilfreiche Funktionen, wie Memory Guard, Quality of Service Management, Trace File Analyzer (TFA) und Enterprise Manager Integration. Details hierzu finden Sie im Oracle DBA Community Tipp zum CHM. (6)

Die komplette Logfile Struktur hat sich mit 12.1.0.2 geändert und die GI Logfiles sind nun ebenfalls komplett im Automatic Diagnostic Repository (ADR) integriert und über die Command Line des ADR (ADRCI) (7) abzurufen:

```
adrci> show alert
Choose the home from which to view the alert log:
1: diag/tnslsnr/bumucsvm6/asmnet1lsnr_asm
2: diag/tnslsnr/bumucsvm6/mgmtlsnr
3: diag/tnslsnr/bumucsvm6/listener
4: diag/tnslsnr/bumucsvm6/listener_scan1
5: diag/rdbms/_mgmtdb/-MGMTDB
6: diag/apx/+apx/+APX1
7: diag/asm/+asm/+ASM1
8: diag/crs/bumucsvm6/crs
```

Der mit 12.1.0.2 GI automatisch installierte Trace File Analyzer versteht sowohl diese neue Struktur der Logfiles, als auch die anderen wichtigen Logfile Dateien von einem RAC Cluster, wie CHM Daten, Betriebssystemlogs, ASM und Datenbank Logfiles. TFA kann automatisch im Fehlerfall oder auch manuell diese Daten sammeln, die danach an Oracle Support hochgeladen werden können. Dabei verringert TFA die hochzuladen Daten immens, außerdem sollte dadurch ein „Ping“-„Pong“ mit dem Oracle Support der Vergangenheit angehören. Näheres zum Umgang mit dem TFA finden Sie ebenfalls in einem Oracle DBA Community Tipp (8).

Neben diesen Tools sollte auch ORACHK nicht vergessen werden. ORACHK überprüft komplette Oracle Systeme (nicht nur GI und Datenbank, sondern auch Enterprise Manager Systeme) auf über 7000 Oracle Best Practices.

## 12c RAC Datenbanken

Auch mit 12c gibt es bei Clustern die Wahl zwischen sogenannten Policy Managed Datenbanken und Administrator Managed Datenbanken. Der Hauptunterschied besteht in der Verwaltung derselben, da im Gegensatz zu einer Admin Managed Datenbank, bei der der Administrator im Details festgelegt, wo etwas laufen soll, Policy Managed Datenbanken vom Cluster verwaltet werden und der Administrator lediglich die Wichtigkeit und Größe entsprechender DB Service festlegt.

12c Policy Managed Datenbanken bietet dabei eine Vielzahl an neuen Möglichkeiten, so dass der Cluster auch Tages zeitabhängig eine Rekonfiguration des Clusters durchführen kann. So kann tagsüber der OLTP Anwendung mehr Rechenressourcen gegeben werden, während Nachts die Auswertungen stärker bevorzugt werden. In diesem Zusammenhang gibt es nun auch die Funktion Server zu kategorisieren, damit bestimmte Server bevorzugt für Testinstanzen, andere für Produktivinstanzen angezogen werden.

Da dieses Art des Setups nun beliebig komplex Ausfallen kann, bietet 12c mit der „What-If“ Analyse in der Clusterware nun die Gelegenheit das Verhalten des Clusters zu prüfen. Hierzu wird den Clusterbefehlen „crsctl“ und „srvctl“ einfach die „eval“ Funktion mitgegeben:

```
$ srvctl relocate server -servers lnxcra12srv1 -serverpool mail -eval
```

```
$ crsctl eval relocate server lnxcra12srv1 -to ora.mail -f
```

Letztendlich ermöglichen Policy Managed Datenbanken auch die Verwendung des Quality of Service Management, so dass aktiv bei Beeinträchtigung eines SLAs der Cluster bestimmte Aktionen durchführen kann. Dies reicht von einfacher Veränderung der Resource Pläne innerhalb des Resource Managers der Datenbank über eine Anpassung des CPU\_COUNT Parameters bis hin zum Verschieben einzelner Server in andere Serverpools.

Zwar steht mit 12.1.0.2 auch Admin Managed Datenbanken das Monitoring von SLAs zur Verfügung, der Cluster greift aber nicht aktiv ein. Zumindest aber den Schutz vor einer Überlastung durch zu hohen Memory Verbrauch gibt es nun mit Memory Guard auch für diese Art der Datenbankadministration.

Mit 12.1.0.2 hat auch InMemory in die Datenbank Einzug gehalten. Oracle RAC unterstützt natürlich auch diese neue Funktion, allerdings gibt es 2 unterschiedliche Modi den speziellen Memory Bereich für Objekte zu verwenden: DISTRIBUTED und DUPLICATE. DISTRIBUTED bedeutet, dass die Objekte automatisch auf die Memory Bereiche aller RAC Knoten aufgeteilt werden. Damit hat man zwar den Vorteil mehr Memory für InMemory zur Verfügung zu haben, der Ausfall eines Knotens hätte aber zur Folge, dass auf die Platte zurückgegriffen werden muss, da die Blöcke nicht mehr im Memory sind. Dem hingegen sorgt DUPLICATE dafür, dass ein Objekt in allen Knoten im RAC vorhanden ist und somit Hochverfügbarkeit gegeben ist. DUPLICATE steht allerdings nur unter Engineered Systemen (Exadata, Supercluster) zur Verfügung.

## Fazit

Auch 12c Grid Infrastruktur ist die Basis Ihre Datenbank hochverfügbar zu betreiben. Dabei stecken viele Neuerungen eher in der Umgebung der Grid Infrastruktur, wie im Bereich Logging, Fehlerverbeugung, Best Practices und Administration als direkt im RAC selber. Dennoch sind insbesondere die kostenfreie Verfügbarkeit von ACFS mit seinen platzsparenden Snapshots für Datenbanken auch schon einen Blick Wert, denn die 12c GI bietet dies auch für aktuelle 11.2.0.4 Datenbanken.

## Links

1) Maximum Availability Architecture

<http://www.oracle.com/goto/maa>

2) White Paper: Transaction Guard with Oracle Database 12c

<http://www.oracle.com/technetwork/database/database-cloud/private/transaction-guard-wp-12c-1966209.pdf>

3) White Paper: Application Continuity with Oracle Database 12c

<http://www.oracle.com/technetwork/database/options/clustering/application-continuity-wp-12c-1966213.pdf>

4) RAC Starter Kits

<https://support.oracle.com/epmos/faces/DocumentDisplay?id=810394.1>

5) Support Impact of the Deprecation Announcement of Oracle Restart with Oracle Database 12c

<https://support.oracle.com/epmos/faces/DocumentDisplay?id=1584742.1>

6) Cluster Health Informationen im Enterprise Manager

[https://apex.oracle.com/pls/apex/GERMAN\\_COMMUNITIES.SHOW\\_TIPP?P\\_ID=3201](https://apex.oracle.com/pls/apex/GERMAN_COMMUNITIES.SHOW_TIPP?P_ID=3201)

7) ADRCI: Automatic Diagnostic Repository Command Interpreter

<http://www.oracle.com/webfolder/technetwork/de/community/dbadmin/tips/adrci/index.html>

8) Diagnose Daten einfach sammeln mit dem Oracle Trace File Analyzer

[https://apex.oracle.com/pls/apex/GERMAN\\_COMMUNITIES.SHOW\\_TIPP?P\\_ID=2881](https://apex.oracle.com/pls/apex/GERMAN_COMMUNITIES.SHOW_TIPP?P_ID=2881)

9) ORAchK - Health Checks for the Oracle Stack (Doc ID 1268927.2)

<https://support.oracle.com/epmos/faces/DocumentDisplay?id=1268927.2>

## Kontaktadresse:

Sebastian Solbach

Oracle Deutschland b.v. & Co. Kg.

Riesstrasse 25

D-80992 München

Telefon: +49 (0) 711-72840 239

E-Mail [sebastian.solbach@oracle.com](mailto:sebastian.solbach@oracle.com)

Internet: [www.oracle.de](http://www.oracle.de)