

Oracle 12c: Neue ASM-Features

Markus Flechtner
Trivadis GmbH
Düsseldorf

Schlüsselworte

Oracle Database12c, Automatic Storage Management, ACFS, FlexASM, HA-NFS

Einleitung

Automatic Storage Management, kurz ASM, hat seit seiner Einführung mit Oracle 10g eine rasante Entwicklung genommen. Mit dem neuen Release 12c hat Oracle die Möglichkeiten von ASM noch erweitert. „Flex ASM“ bringt Hochverfügbarkeit und verbesserte Ressourcennutzung für ASM. Automatische Fehlerkorrektur, sog. „Disk Scrubbing“, erhöht die Datensicherheit. Verbesserungen in den Bereichen „Rebalance“ und „Resync“ ergänzen die Weiterentwicklung ab. Mit 12c erlaubt Oracle auch die Verwendung von ACFS für Datenbankdateien. Damit kann man für Datenbanken den Komfort eines Filesystems mit der ASM-Funktionalität verbinden. Weitere Änderungen bei ACFS sind „HA NFS“ sowie Verbesserungen in den Gebieten Replikation, Verschlüsselung und Auditing.

Erinnerung: Was macht ASM?

Bevor wir die neuen Features anschauen, hilft der Blick zurück: was macht ASM? ASM ist ein Volume Manager und Dateisystem für Datenbanken. Es verwaltet die Metadaten der Diskgruppen und Disks, d.h. die Aufteilung der vorhandenen Festplatten in Diskgruppen, den verfügbaren Speicherplatz der Diskgruppen und den Überblick über den ASM-Extents („extent map“).

ASM macht ein automatisches Striping über die vorhandenen Disks einer Diskgruppe und erlaubt über die Redundanzlevel „normal“ und „high“ ein Mirroring.

Im Hinblick auf die neuen Features sind zwei Aspekte wichtig:

1. Der Datenbank-I/O erfolgt direkt durch die Datenbank-Instanzen, nicht durch die ASM-Instanz. Die Datenbank-Instanzen erhalten die Metadaten von der ASM-Instanz und greifen dann direkt auf die ASM-Disks zu.
2. Fällt eine ASM-Instanz aus, so stürzen die Datenbank-Instanzen, die diese ASM-Instanz nutzen, ab.
- 3.

FlexASM

„FlexASM“ ist ein neues Feature für Oracle 12c-Cluster. Mit FlexASM ist keine lokale ASM-Instanz mehr erforderlich. Im Cluster laufen standardmäßig drei ASM-Instanzen. Gibt es eine lokale ASM-Instanz, so wird diese von den Datenbank-Instanzen genutzt. Ansonsten verbinden sie sich über das sog. „ASM-Net“ mit einer ASM-Instanz auf einem anderen Cluster-Knoten verbinden und erhalten von dort die notwendigen Meta-Informationen. Fällt eine ASM-Instanz aus, verbinden sich die Datenbank-Instanzen mit einer anderen im Cluster vorhandenen ASM-Instanz aus.

Mittels „ALTER SYSTEM RELOCATE CLIENT ..“ kann der Wechsel zu einer anderen ASM-Instanz auch erzwungen werden.

Durch die reduzierte Anzahl von ASM-Instanzen ist FlexASM insgesamt ressourcenschonender als die klassische ASM-Konfiguration.

Für FlexASM sind neue Cluster-Ressourcen erforderlich:

- ASM-Netzwerk
- ASM-Listener
- ADVM-Proxy

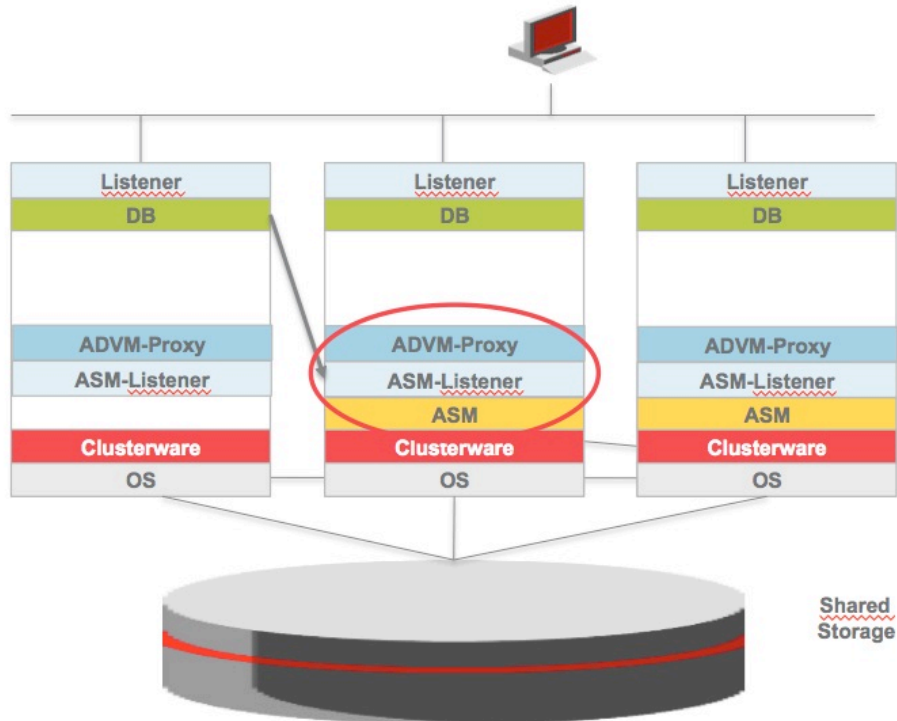


Abbildung 1: FlexASM: Architektur

Das ASM-Netzwerk ist dabei das Netzwerk, über das die Datenbank- und ASM-Instanzen miteinander kommunizieren, d.h. die Metadaten austauschen. Ein dediziertes Netzwerk wird empfohlen; es kann aber auch ein existierendes Netzwerk (z.B. der Cluster-Interconnect) verwendet werden.

Innerhalb des ASM-Netzwerkes gibt es die ASM-Listener, die – und das ist übliche Listener-Funktionalität – die Verbindungsanfragen der Datenbank-Instanzen entgegennehmen und an die ASM-Instanzen weiterreichen.

Der ADVM-Proxy ist notwendig, damit ACFS auf FlexASM zugreifen kann. Es sind eigene Instanzen, die mit SID „+APXn“ haben. Sie werden automatisch gestartet und müssen nicht konfiguriert werden.

```

grid@rac2node1:~/ [+ASM1] crsctl stat res -t
-----
Name                Target  State          Server          State details
-----
Local Resources
-----
ora.ASMNET1LSNR ASM.lsnr
      ONLINE  ONLINE        rac2node1      STABLE
      ONLINE  ONLINE        rac2node2      STABLE
      ONLINE  ONLINE        rac2node3      STABLE
..
ora.proxy_advm
      ONLINE  ONLINE        rac2node1      STABLE
      ONLINE  ONLINE        rac2node2      STABLE
      ONLINE  ONLINE        rac2node3      STABLE
..
ora.asm
      2      ONLINE  ONLINE        rac2node1      STABLE
      3      ONLINE  ONLINE        rac2node2      STABLE

```

Abbildung 2: neue Cluster-Ressourcen bei FlexASM

FlexASM wird mit dem bekannten Befehl „`srvctl`“ administriert:

```

srvctl modify asm [-listener <lsnr_name>] [-pwfile <password_file_path>]
      [-count {<number_of_instances>|ALL}] [-force]
srvctl remove asm [-proxy] [-force]
srvctl relocate asm -currentnode <current_node> [-targetnode <target_node>]
[-force]

```

Wir sehen, dass man die ASM-Kardinalität (Default: 3) ändern kann. Dabei ist zu beachten, dass bei einer Reduktion der Kardinalität die überzähligen Instanzen automatisch gestoppt werden, dass aber andersherum bei einer Erhöhung der Kardinalität zusätzliche Instanzen mit „`srvctl start asm`“ gestartet werden müssen.

Weiterhin können ASM-Instanzen im Cluster verschoben werden. Der Befehl „`srvctl relocate asm`“ ist allerdings nur in der Online-Hilfe von `srvctl` dokumentiert, nicht aber in der Oracle-Dokumentation. Dabei ist es allerdings nicht erlaubt, die ASM-Instanz des OCR-Masters zu verschieben. In diesem Fall kommt die etwas kryptische Fehlermeldung „CRS-0223: Resource 'ora.asm 3 1' has placement error.“ bzw. die Meldung im `crsd.log` „CRS-5027: Cannot stop ASM instance on node "rac2node2" because it is currently being used by Cluster Ready Service“.

Bei der Konvertierung einer vorhandenen ASM-Installation zu FlexASM hilft der „ASM Configuration Assistant“ (`asmca`), der alle Schritte vorbereitet. Voraussetzung für FlexASM ist, dass alle Diskgruppen mit ASM-Kompatibilitätslevel 12.1 angelegt sind und dass SPFILE, Password-Datei und die Cluster Registry (OCR) in ASM abgelegt wurden.

Werden auf einem Cluster mit FlexASM auch Datenbanken der Version 11 betrieben, so muss auf jedem Knoten eine ASM-Instanz laufen (`cardinality=ALL`), denn Oracle 11g-Datenbankinstanzen erwarten eine lokale ASM-Instanz.

Weitere neue ASM-Features in Oracle 12c

Password-Dateien in ASM

Eine weitere Neuerung in Oracle 12c ist, dass die Password-Dateien in einer ASM-Diskgruppe abgelegt werden können. Früher mussten diese Dateien im Dateisystem abgelegt werden und z.B. bei der Änderung des SYS-Passwortes einer Datenbank mussten alle Password-Dateien einzeln geändert werden. Dies entfällt mit Oracle 12c.

```
grid@node1:~/[+ASM1] orapwd file='+GRID/ASM/orapwasm' asm=y
```

```
oracle@node1:~/[RACDB1] orapwd file='+GRID/RACDB/orapwRACDB'\  
dbuniqueName=RACDB
```

Disk-Scrubbing

Mittels „disk scrubbing“ kann ASM Diskgruppen automatisch auf Korruptionen prüfen. Dies geschieht automatisch bei einem „rebalance“; kann aber auch manuell ausgelöst werden:

```
SQL> ALTER DISKGROUP U01 SCRUB;  
SQL> ALTER DISKGROUP U01 SCRUB REPAIR;  
SQL> ALTER DISKGROUP U01 SCRUB DISK U01_0001;  
SQL> ALTER DISKGROUP U01 SCRUB  
2 FILE '+U01/ORCL/DATAFILE/users.269.806412493' REPAIR POWER LOW;
```

In Diskgruppen mit normaler oder hoher Redundanz können entdeckte Datenkorruptionen repariert werden.

Even-Read für Diskgruppen

Beim Lesen von einer Diskgruppe mit normaler oder hoher Redundanz wurde bis Oracle 11g immer die primäre Kopie eines Extents gelesen. Ausgenommen natürlich die Systeme mit ASM Preferred Mirror Read (Stretch Cluster), die so konfiguriert sind, dass jeweils von der lokalen Failure Group gelesen wird.

Mit Oracle 12c werden Lesezugriffe gleichmäßig auf alle Platten, die die jeweiligen Extents enthalten, verteilt. „Preferred Mirror Read“ funktioniert natürlich weiterhin.

OCR-Backup in ASM

Die Oracle-Clusterware sichert regelmäßig die Cluster Registry (OCR). Der Backup wird üblicherweise innerhalb des Oracle-Home-Verzeichnis auf dem sog. OCR-Master abgelegt. Mit Oracle 12c kann der Backup auch in einer ASM-Diskgruppe abgelegt werden

```
ocrconfig -backuploc +FRA
```

Leider kann dieser Backup nicht einfach mittels asmcmd in ein Dateisystem kopiert werden. Dazu ist der Umweg via „DBMS_FILE_TRANSFER“ erforderlich.

Neue ACFS-Features in Oracle 12c

ACFS ist ein auf ASM basierendes Dateisystem, das es seit Oracle 11g gibt. Auch hier gibt es zahlreiche Neuerungen.

Das Produkt „Oracle Cluster File System“ (OCFS) beruht auf den Komponenten „ACFS“ (ASM Cluster File System) und „ADVM“ (ASM Dynamic Volume Manager).

OCFS-Lizensierung

Mit Oracle 12c wurde die OCFS-Lizensierung des Oracle geändert. Die „Grundfunktionalität“ ist jetzt kostenlos. Für die Nutzung der erweiterten Funktionalität, d.h. Snapshots, Replikation, Tagging, Security und Auditing ist eine gesonderte OCFS-Lizenz erforderlich.

ACFS für alle Datenbank-Dateien

Bis Oracle 11.2 galt, dass Datenbank-Dateien (Datendateien, Redolog-Dateien, Control-Dateien etc.) nicht in einem ACFS-Dateisystem abgelegt werden durften. Dies ist jetzt erlaubt.

Zusammen mit ACFS-Snapshots (lizenzpflichtig!) ist dies eine einfache Möglichkeit, Test-Datenbanken bereitzustellen.

High-Availability-NFS mit Oracle Grid Infrastructure

Bei HA-NFS führt Oracle mehrere existierende Funktionalitäten der Grid Infrastructure zusammen und ermöglicht ein hochverfügbares Network File System:

- ACFS
- Virtuelle IP-Adressen (VIP)
- Überwachung von Cluster-Ressourcen mittels Clusterware

Nach den NFS- und DNS-seitigen Vorbereitungen sind für HA-NFS folgende Aktivitäten erforderlich:

1. Anlegen der VIP und des exportierten Verzeichnisses als Cluster-Resource

```
root# srvctl add havip -id demovip -address demohanfs.trivadis.com
root# srvctl add exportfs -id demovip \
    -path /u00/app/oracle/acfs_acfsvol -name demohanfs
```

2. Starten der virtuellen IP

```
root# srvctl start havip -id demovip
```

Neuerungen bei den ACFS-Snapshots

Zwei angenehme Neuerungen gibt es bei den ACFS-Snapshots. Zum einen können Read-Only-Snapshots jetzt in Read-Write-Snapshots umgewandelt werden:

Von Read-Only nach Read-Write:

```
acfsutil snap convert -w ACFS2_SNAP1 /acfs2
```

Von Read-Write nach Read-Only:

```
acfsutil snap convert -r ACFS2_SNAP1 /acfs2
```

Weiterhin können Snapshots als Grundlage für weitere Snapshots dienen („snap on snap“):

```
grid@node1: acfsutil snap create -r -p ACFS2_SNAP1
grid@node1: acfsutil snap info ACFS2_SNAP2 /acfs2
snapshot name:          ACFS2_SNAP2
```

RO snapshot or RW snapshot: RW
parent name: ACFS2_SNAP1
snapshot creation time: Wed Jul 24 10:08:28 2013

Erweiterte Unterstützung für ACFS-Funktionalität

Seit Beginn der Verfügbarkeit von ACFS mit Oracle 11.2 wurde die Funktionalität und die Unterstützung der verschiedenen Plattform stetig erweitert. Mit Oracle 12c ist dieser Prozess fast abgeschlossen: Snapshots, Tagging und Replikation stehen für alle Plattformen (Linux, Windows, Solaris und AIX) zur Verfügung. Die Security-Features und Verschlüsselung sind nur für AIX (noch) nicht verfügbar.

Kontaktadresse:

Markus Flechtner
Trivadis GmbH
Werdener Straße 4
D-40227 Düsseldorf

Telefon: +49 (0) 211 58 66 64 70
Fax: +49 (0) 211 58 66 64 71
E-Mail: Markus.Flechtner@trivadis.com
Internet: www.trivadis.com