

Oracle ASM Cluster File System (ACFS) und ASM Dynamic Volume Manager (ADVM)

Sven Morling
OPITZ CONSULTING Schweiz GmbH
Thalwil, Zürich

Schlüsselworte:

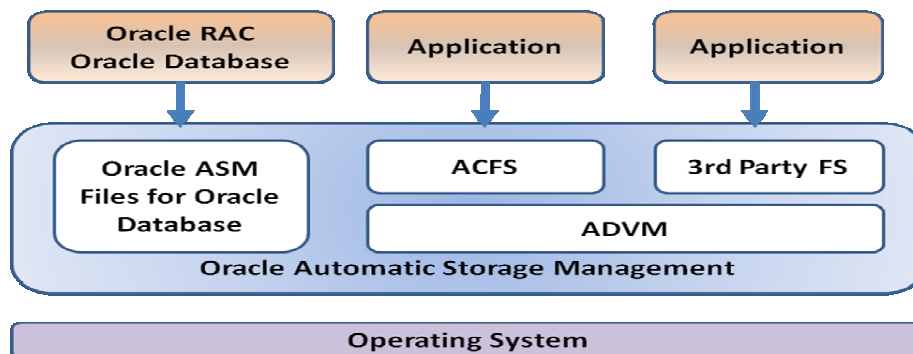
Automatic Storage Management, Cluster File System, ASM, Logical Volume Management, ADVM

Einleitung

Die Geschichte von ASM beginnt mit der Oracle Datenbank 10g Release 1. Damals war ASM eine Blackbox. Der Widerstand der DBA's war groß. Keiner wollte dieser Blackbox trauen. Oracle hat sich dieser Problematik in der Version 10g Release 2 angenommen. Es kam „asmcmd“. Die Taschenlampe für die Blackbox. Damit konnte man auch endlich in das Mysterium ASM hineinschauen. Somit gewann ASM immer mehr Anhänger. Damit wurden aber auch die Ansprüche an ASM größer. Mit Version 11g Release 1 erweiterte Oracle die Funktionalität von ASM. Es kamen Features wie Preferred Mirror Read, Fast Mirror Resync und noch weitere kleine nette Features. Aber mit Version 11g Release 2 bringt Oracle nun einen ECHTEN Volume Manager auf den Markt. OK, ASM war ja schon ein Quasi-Volume Manager. Aber nur für Datenbankfiles. Aber nun, mit ACFS und ADVM ist es möglich verschiedenste Dateien im „ASM“ abzulegen.

Konzept von ACFS und ADVM

Oracle Automatic Storage Management Cluster File System (ACFS) ist für den Betrieb auf vielen Plattformen vorbereitet. Dabei bedient sich ACFS komplett der Features von ASM. Eingeschlossen sind hier Mirroring, Striping sowie Preferred Mirror Read und auch Fast Mirror Resync. Es wird empfohlen ACFS für alle Dateien zu benutzen, die nicht im ASM direkt abgelegt werden können. Zusätzlich ist es nicht supportet, Dateien, zum Beispiel Datafiles, in einem ACFS abzulegen, die man direkt im ASM ablegen kann. Weiterhin ist es nicht supportet, Dateien von Grid Infrastructure, sowie Voting Files und Oracle Cluster Registry Dateien im ACFS abzulegen.



Installation

Seit der Version 11g Release 2 ist das ASM nicht mehr im RDBMS Home enthalten und muss nun aus der Grid Infrastructure Installation gestartet werden. Wenn man dann die ASM Instanz hat, kann man eine Diskgroup mit den notwendigen Kompatibilitätseinstellungen erstellen:

```
[oracle@test] # asmca -silent -createDiskGroup \
                -diskString "ORCL:*" -diskGroupName ACFS \
                -diskList ORCL:D01,ORCL:D02,ORCL:D03,ORCL:D04 \
                -redundancy normal -compatible.asm 11.2 \
                -compatible.rdbms 11.2 -compatible.advm 11.2
```

```
SQL> select a.name, b.name, b.value
       2 from v$asm_diskgroup a, v$asm_attribute b
       3 where a.group_number=b.group_number
       4    and b.name like 'compa%'
       5    and a.name = 'ACFS';
```

NAME	NAME	VALUE
ACFS	compatible.asm	11.2.0.0.0
ACFS	compatible.rdbms	11.2.0.0.0
ACFS	compatible.advm	11.2.0.0.0

Damit ist die Diskgroup nun in der Lage ACFS Volumes aufzunehmen. Dafür gibt es dann auch von Oracle einen neuen ASM File Typ namens „ASMVOL“:

```
SQL> select distinct type from v$asm_file;
```

```
TYPE
-----
...
ASMVOL
...
```

Mittels `asmcmd` erzeugt man jetzt das Volume, welches hinterher im `/dev`-Pfad erreichbar ist.

```
ASMCMD> volcreate -G ACFS -s 2g acfstest
```

```
[oracle@test ~]$ file /dev/asm/acfstest-233
/dev/asm/acfstest-233: block special (252/119297)
[root@test asm]# mkfs -t acfs /dev/asm/acfstest-233
mkfs.acfs: version = 11.2.0.1.0.0
mkfs.acfs: on-disk version = 39.0
mkfs.acfs: volume = /dev/asm/acfstest-233
mkfs.acfs: volume size = 2147483648
mkfs.acfs: Format complete.
[root@test asm]# echo "/dev/asm/acfstest-233 /mnt acfs defaults
1 2" >> /etc/fstab
[root@test asm]# mount /mnt
[root@test asm]# df -k /mnt
Filesystem 1K-blocks Used Available Use% Mounted on
/dev/asm/acfstest-233 2097152 41272 2055880 2% /mnt
```

Damit ist jetzt das Filesystem verfügbar. Allerdings ist Oracle hier noch etwas nachlässig, was die Persistenz bei ACFS nach einem Reboot angeht. Beziehungsweise, da es sich ja um eine von root zu administrierende Komponente handelt, wurde ACFS nicht in den normalen Oracle Restart-Stack mit eingebaut. Das bedeutet, daß man die ACFS Komponenten noch in den Systemstart einbinden muss. Dies sollte zwingend vor einem Reboot passieren, da man sonst versucht das Filesystem durch den „/etc/fstab“-Eintrag beim Systemstart zu mounten und damit fehlschlägt.

```
[root@test ~]# cat /etc/init.d/acfsload
#!/bin/sh

# chkconfig: 2345 30 21
# description: Load Oracle ACFS drivers at system boot
/u01/app/oracle/product/11.2.0/grid/bin/acfsload start -s
[root@test ~]# chkconfig --add acfsload
[root@test ~]# chkconfig --list acfsload
acfsload          0:off  1:off  2:on   3:on   4:on   5:on   6:off
```

Somit ist die Einrichtung eines ACFS Volumes jetzt komplett und benutzbar.

Anwendung

Oracle bindet den ACFS Treiber direkt in die Systembibliotheken ein. Das bedeutet, daß man mkfs, chfs, mount und ähnliche Linux Systemprogramme benutzen kann. Auf Windows heißen die Programme parallel dazu acfschkdsk, acfsformat, acfsmountvol und acfsdismount. Das bedeutet aber auch, daß man durchaus andere Filesysteme als das von Oracle jetzt mitgeliefert ACFS benutzen kann.

Die Administration der Volumes kann sowohl aus dem asmcmd als auch ueber die acfs* Utilities im Betriebssystem gemacht werden. Im asmcmd stehen hierfür diverse neue Befehle zur Verfügung:

```
ASMCMDB> help
...
    volcreate, voldelete, voldisable, volenable, volinfo
    volresize, volset, volstat
...
```

Im Betriebssystem hat man ebenfalls einige Utilities um ACFS oder auch direkt ADVM zu administrieren beziehungsweise, Informationen über Volumes und Mountpoints zu erhalten. Die beiden wichtigsten hierbei sind:

```
[root@test ~]# acfsutil

Usage: acfsutil [-h] command ... # (Version 11.2.0.1.0.0)

    -h - help

Command Subcmd Arguments
-----
...
info fs                                - Display all ACFS file systems
...
info id          ACFS file identifier and mountpoint.
rmfs             <device>                - Remove unmounted ACFS file sys
```

```

    size          [-|+]nnn[K|M|G|T|P] [-d <device>] <path> - Resize file
system
registry        [-l]                               - Display registry contents
...
tune            AcfsMaxOpenFiles                    - Maximum number of open files
tune                                                    (Windows only)
snap create <snap_name> <mountpoint> - create a file system snapshot
snap delete <snap_name> <mountpoint> - delete a file system snapshot

```

und

```
[root@test ~]# advmutil
```

```
Usage: advmutil [-h] command ... # (Version 0.0.0.0.0.0)
```

```
-h - help
```

```

Command Arguments
-----
help            - display this message
tune            - list all tunables
...
volinfo <device name> - display ADVM volume
information
volinfo          using a one entry per line
format
...

```

Snapshots

Snapshots bieten die Möglichkeit eine On-Disk-Sicherung der Daten. Dazu wird Im Moment der Erzeugung des Snapshots nur gering Speicher verbraucht. Oracle greift hier auf die Copy-On-Write Technik zurück. Das bedeutet, das erst bei Veränderung des Blocks bzw der Datei eine Kopie erstellt wird.

Die Administration der Snapshots gestaltet sich sehr einfach:

```

[root@test ctest]# echo "ORACLE ACFS Snapshot Test" >
/u01/app/oracle/acfsmounts/ctest/snap.test
[root@test ctest]# acfsutil snap create 1 /u01/app/oracle/acfsmounts/ctest
acfsutil snap create: Snapshot operation is complete.
[root@gartentrainer ctest]# acfsutil info fs \
/u01/app/oracle/acfsmounts/ctest
/u01/app/oracle/acfsmounts/ctest
ACFS Version: 11.2.0.1.0.0
flags:          MountPoint,Available
mount time:     Mon Sep  6 15:26:06 2010
volumes:       1
total size:     1073741824
total free:     979058688
primary volume: /dev/asm/btest-330
label:
flags:          Primary,Available,ADVM
on-disk version: 39.0

```

```

allocation unit:      4096
major, minor:       252, 168962
size:               1073741824
free:               979058688
ADVM diskgroup      FRA
ADVM resize increment: 268435456
ADVM redundancy:    unprotected
ADVM stripe columns: 4
ADVM stripe width:  131072
number of snapshots: 1
snapshot space usage: 32768
[root@test ctest]# ls /u01/app/oracle/acfsmounts/ctest/.ACFS/
.fileid/ repl/     snaps/
[root@test ctest]# ls /u01/app/oracle/acfsmounts/ctest/.ACFS/snaps/1/
lost+found oradata snap.test
[root@test ctest]# cat /u01/app/oracle/acfsmounts/ctest/snap.test
ORACLE ACFS Snapshot Test
[root@test ctest]# rm -f /u01/app/oracle/acfsmounts/ctest/snap.test
[root@test ctest]# file /u01/app/oracle/acfsmounts/ctest/snap.test
/u01/app/oracle/acfsmounts/ctest/snap.test: ERROR: cannot open
`/u01/app/oracle/acfsmounts/ctest/snap.test' (No such file or directory)
[root@test ctest]# cp
/u01/app/oracle/acfsmounts/ctest/.ACFS/snaps/1/snap.test
/u01/app/oracle/acfsmounts/ctest/snap.test
[root@test ctest]# cat /u01/app/oracle/acfsmounts/ctest/snap.test
ORACLE ACFS Snapshot Test
[root@test ctest]# acfsutil snap delete 1 /u01/app/oracle/acfsmounts/ctest
acfsutil snap delete: Snapshot operation is complete.
[root@test ctest]# ls /u01/app/oracle/acfsmounts/ctest/.ACFS/snaps/*
ls: /u01/app/oracle/acfsmounts/ctest/.ACFS/snaps/*: No such file or
directory

```

New Features 11.2.0.2.0

Das Patchset 11.2.0.2.0 bringt noch zusätzliche Neuerungen die ACFS betreffen. Hier kommen Sicherheits- und Verschlüsselungserweiterungen, Replikation und Tagging hinzu. Zur Sicherheitserweiterung ist zu sagen, das es sich um eine Erweiterung der Zugriffsbeschränkung des Betriebssystems handelt. Die Verschlüsselung soll den Zugriff auf die Daten bei Verlust oder Diebstahl sperren. Tagging bietet die Möglichkeit, einzelne Dateien mit einem zusätzlichen „Tag“ zu versehen. Wenn man dann mehrere Dateien mit dem selben „Tag“ versieht, ist es so möglich Dateien zu gruppieren. Dies kann man zum Beispiel für die Replikation verwenden, wenn man nicht das ganze Filesystem replizieren möchte.

Desweiteren wurde der Support für AIX und Solaris angekündigt, sowie das benutzen von Snapshots auf diesen beiden Betriebssystemen.

Fazit

Die Verbreiterung der Oracle Produktpalette schreitet unaufhörlich weiter. Konzentrieren wir uns hierbei mal nur auf den Markt rund um die Datenbank. Erst kam die Veröffentlichung eines eigenen Linux. Nachfolgend wurde mit HP eine DB-Hardware-Appliance auf den Markt gebracht. Diese wurde mit dem Kauf von SUN erneuert und die Hardware wurde durch SUN-eigene ausgetauscht. Inzwischen haben wir Exadata V2. Und nun kommt mit Oracle 11g Release 2 auch noch ein echter Logical Volume Manager an den Start. Da müssen man sich fragen, wie weit wird es noch gehen. Allerdings muß man auch zugeben, das es doch seinen Charme hat, alles aus einer Hand zu haben. Damit werden Supportanfragen schneller und besser abgewickelt. Die Informationsquellen werden damit auch übersichtlicher. Und wenn dann jetzt noch der Preis und die Stabilität stimmt!

Kontaktadresse:

Sven Morling

OPITZ CONSULTING Schweiz GmbH
Seestrasse 97
CH-8800 Thalwil

Telefon: +41 44 721 11 22
E-Mail: sven.morling@opitz-consulting.com
Internet: www.opitz-consulting.com