

Oracle Database 11g Automatic Storage Management New Features Overview

*An Oracle Technical White Paper
June 2007*

NOTE:

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

**ORACLE DATABASE 11G
AUTOMATIC STORAGE MANAGEMENT
NEW FEATURES OVERVIEW**

NOTE:	1
ORACLE DATABASE 11G AUTOMATIC STORAGE MANAGEMENT NEW FEATURES OVERVIEW	2
INTRODUCTION	3
AUTOMATIC STORAGE MANAGEMENT OVERVIEW	3
AUTOMATIC STORAGE MANAGEMENT NEW FEATURES	4
IMPROVED SCALABILITY AND PERFORMANCE	5
FAST MIRROR RESYNC	5
VLDB SUPPORT	6
<i>Variable size extents</i>	6
MULTIPLE ALLOCATION UNIT SIZE	7
FASTER REBALANCE WITH RESTRICTED MOUNT OPTION.....	7
PREFERRED MIRROR READ.....	8
ASM ROLLING UPGRADES AND PATCHING	8
TABLESPACE LEVEL MIGRATION WIZARD IN ENTERPRISE MANAGER	9
MANAGEABILITY ENHANCEMENTS	10
DISK GROUP COMPATIBILITY ATTRIBUTES	10
NEW ASMCMD COMMANDS	10
<i>ASMCMD cp</i>	10
<i>ASMCMD disk group metadata backup and restore</i>	11
<i>ASMCMD lsdk</i>	12
<i>ASMCMD remap</i>	12
SYSASM ROLE	12
MOUNT/DROP FORCE DISK GROUP	13
CONCLUSION	14

INTRODUCTION

“With the ASM feature, DBAs won’t have to worry about optimizing disk I/O. ASM eliminates hot spots by evenly distributing data providing the optimal bandwidth that an end user or an application needs” *Arvind Gidwani, Senior Manager IT, Qualcomm*

Automatic Storage Management (ASM) in Oracle Database 11g is an evolution in file system and volume management functionality for Oracle database files. ASM further enhances automation and simplicity in storage management that is critical to the success of the Oracle grid architecture. ASM also improves file system scalability and performance, manageability, and database availability for single-instance Oracle databases as well as for Oracle Real Application Clusters (Oracle RAC) environments.

This white paper is targeted at a technical audience mainly comprising:

- Database, system and storage administrators
- Architects
- Consultants
- System engineers
- Technical managers

Automatic Storage Management Overview

ASM is a feature of Oracle Database 11g that provides an integrated cluster file system and volume management capabilities at no additional cost. ASM lowers the Oracle database storage total cost of ownership and increases storage utilization while improving performance and availability over traditional file system and volume management solutions. With ASM, a fraction of the time is needed to manage your database storage environment and datafiles.

ASM is easier to manage than conventional file systems, and it has the performance of raw volumes and is tightly integrated with the Oracle Database forming the foundation of a storage grid for Oracle. Additionally, ASM eliminates the need for 3rd-party volume managers and file systems for managing the Oracle database files.

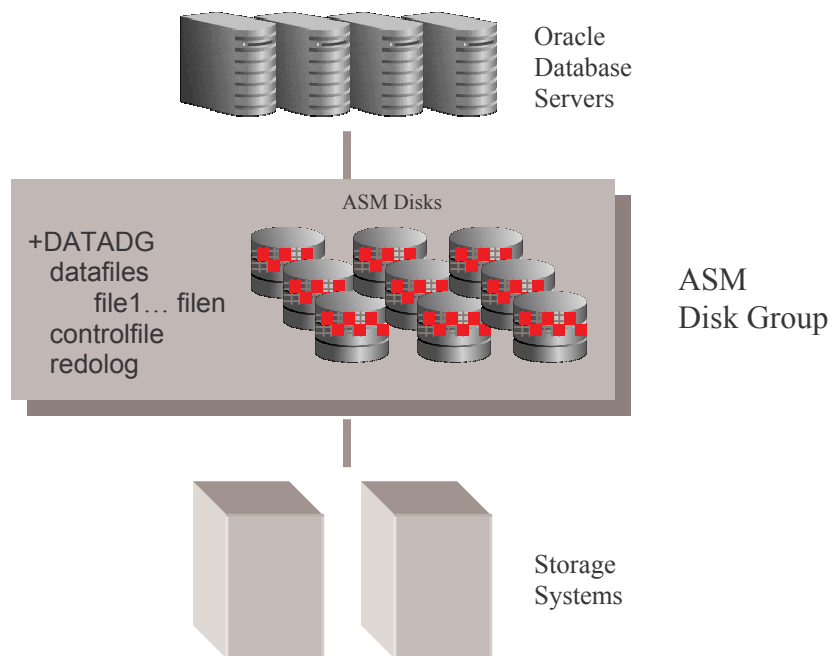


Figure 1

ASM brings significant key values to Oracle Database platforms. ASM improves manageability by simplifying storage provisioning, storage array migration, and storage consolidation. ASM provides flexible easy to manage interfaces including the SQL*Plus, Oracle Enterprise Manager GUIs and a UNIX-like command line interfaces. ASM provides sustained best in class performance because of its innovative rebalancing feature that distributes data evenly across all storage resources, providing an even distribution of IO and optimal performance. ASM scales to multi-TB databases (VLDB) efficiently without compromising functionality or performance.

ASM is built to maximize database availability. ASM provides self-healing automatic mirror reconstruction and resynchronization, rolling upgrades and patching. ASM also supports dynamic and on-line storage reconfigurations in both single instance and Oracle RAC database configurations. ASM customers realize significant cost savings and achieve lower total cost of ownership because of features such as just-in-time provisioning, and clustered pool of storage that is ideal for database consolidation. ASM provides all of this without an additional license or licensing fees.

ASM is the preferred file system and volume manager for Oracle database files for the following reasons:

- Simplifies and automates storage management
- Increases storage utilization, uptime, and agility
- Delivers predictable performance and availability service level agreements

Automatic Storage Management New Features

“ASM solves the majority of the performance and manageability issues of using low-cost storage to run Oracle on Linux and allows you to build very large, high-performance systems”*Grant McAlister*
Principal Database Engineer
Amazon.com

ASM in Oracle Database 11g provides major functionality enhancements in the following areas:

- **Scalability and performance**
 - Improved scalability and performance for VLDB
 - Fast mirror resynchronization
 - Preferred mirror read in a clusters
 - Support for larger allocation units (AU)
 - Fast rebalance
- **Rolling upgrade and patching**
- **Manageability enhancements**
 - Table level migration wizard in EM
 - Disk group compatibility attributes
 - New ASMCMD commands

- New SYSASM privilege separate from the SYSDBA privilege
- More flexible FORCE options to MOUNT and DROP disks group

IMPROVED SCALABILITY AND PERFORMANCE

Fast Mirror Resync

The ability to resync mirrors is not only key to the success of our extended cluster, but it also allows for proactive preventative maintenance resulting in greater uptime.

ASM Fast Mirror Resync significantly reduces the time required to recover from transient disk failures in normal and high redundancy disk groups. Proactive maintenance of ASM disks and/or transient ASM disk failures can both highly benefit from this feature. When a disk goes offline following a transient failure, ASM tracks the extents that are modified during the outage. When the transient failure is repaired, ASM can quickly resynchronize only the ASM mirror extents that have been updated during the outage. A transient failure assumes that the contents of the mirrored disks are intact and the reasons for the disk unavailability are because of failures such as a loose cable, a power failure or a recycle, a host bus adapter, or a disk controller.

The new DISK_REPAIR_TIME attribute defines a time window in which the transient failure can be repaired and the mirror disk can be brought back on-line. ASM only resynchronizes the changed blocks (extents), which results in much faster recovery time. This can mean a difference between minutes to hours or days depending on the amount of data on the disks being brought on-line.

The DISK_REPAIR_TIME attribute is set to 3.6 hours by default. You can use ALTER DISKGROUP <disk group name> SET ATTRIBUTE "DISK_REPAIR_TIME"=" 24H" to, as in this example, extend this value to a longer time period. After you repair the disk, run the SQL statement ALTER DISKGROUP DISK ONLINE. This statement brings a repaired disk group back online and starts the resynchronization process.

The V\$ASM_ATTRIBUTE view shows the current disk repair time attribute value. You can see the time remaining value in the REPAIR_TIME column of the V\$ASM_DISK or V\$ASMDISK_STAT views. You can also manually OFFLINE a disk or failure group, or force drop offline disks that cannot be repaired. In this case, a full resynchronization is necessary after replacing the failed disks and adding them to the disk group.

This feature applies to all cases where ASM mirroring is used and further improves database availability. Applications such as extended clusters, where you create two failure groups (one on each site), derive additional benefits because a network or connectivity outage can be recovered very quickly without requiring a full mirror resync. Proactive preventative maintenance is also now feasible because you can offline failure groups, perform maintenance, and then online the failure group to update your disk contents in relatively much shorter time without the need to take the database down.

VLDB Support

Variable size extents

The variable size extent feature in Oracle Database 11g enables support for much larger ASM files, reduces SGA memory requirements for very large databases, and improves performance for file create and open functions.

Allocation Unit (AU)

Selected at disk group creation time and may be 1,2,4,8,16,32,64 MB

Striping

Coarse Stripe size always = one AU
Fine Stripe size always = 128 KB

Variable size ASM file extents

Extent size = AU size up to 20,000 extents

Extent size = 8*AU up to 40,000 extents

Extent size = 64*AU beyond 40,001 extents
(Not shown on this graph)

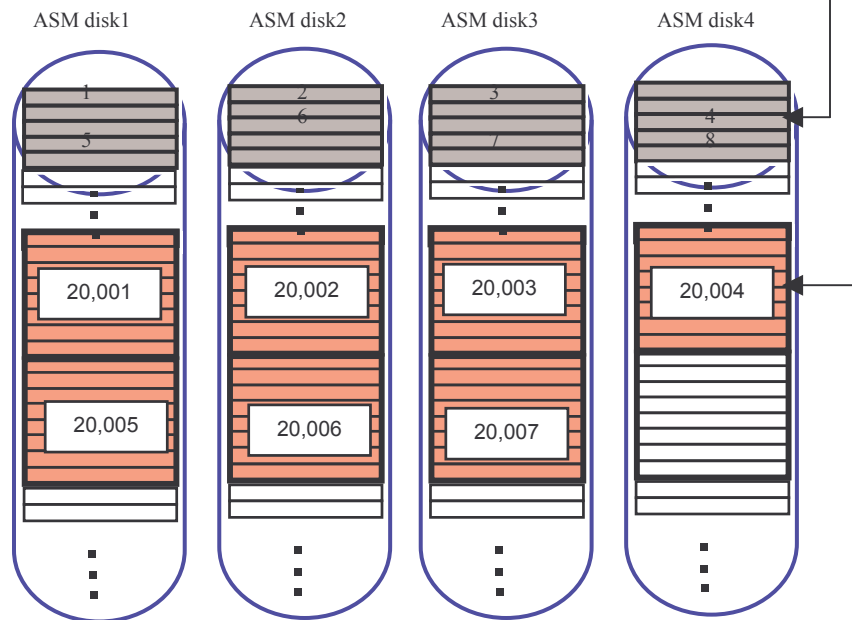


Figure 2

An ASM file can begin with 1 MB extents (assuming 1MB AU) and as the file size increases, the extent size also increases to 8MB and 64MB at a predefined number of extents. Therefore, the size of the extent map defining a file can be smaller by a factor of 8 and 64 depending on the size of the file. The initial extent size is equal to the allocation unit size and it increases by the 8 and 64 factor at predefined thresholds.

This feature is completely automatic for newly created files (or the extended part of files) once compatible.asm and compatible.rdbms have been advanced to 11.1.

ASM VLDB support is extended to:

- 140 Peta Byte (PB) in External Redundancy (no ASM mirroring)
- 42PB in Normal Redundancy (2-way ASM mirroring)
- 15PB in High Redundancy (3-way ASM mirroring)
- 40 ExaBytes of storage
- 4PB for each ASM disk
- 1 million files per disk group
- 10,000 ASM disks
- 63 disk groups per ASM instance

Please note that Oracle database limits file sizes to 128TB with big file tablespaces and 32k block size which is considerably lower than ASM file size limits.

Multiple Allocation Unit size

You can now set the ASM Allocation Unit (AU) size to be 1/2/4/8/16/32/64MB at disk group creation time. An allocation unit is the minimum segment of a disk that can be assigned to a file. Each extent is an integral number of allocation units. Each file contains an integral number of extents. Our research has shown that significant performance can be gained for some workloads and storage system types deployed by using larger AU sizes.

The default AU size is set to 1MB for ASM to match the MAXIO default for Oracle Database IO buffer size. The larger AU sizes are likely to benefit large sequential reads and writes. It may be beneficial to increase the MAXIO buffer size to 4MB and create the ASM disk group with a 4MB allocation unit for optimal performance for applications performing large sequential IO.

Please note that the FINE-grained striping remains 128KB as before, but the COARSE-grained stripe is equal to the AU size selected.

Faster Rebalance with Restricted Mount Option

A disk group mounted in RESTRICTED mode is mounted exclusively by only one node. Clients of ASM on that node cannot access that disk group while it is mounted in RESTRICTED mode. The RESTRICTED mode enables you to perform all maintenance tasks on a disk group in the ASM instance without any external interaction. Rebalance operations performed while the diskgroup is in RESTRICTED mode eliminate the lock/unlock extent maps messaging between ASM instances in an Oracle RAC environment, thus improving overall rebalance throughput.

At the end of the maintenance cycle you have to explicitly dismount the disk group and remount it in normal mode.

The ALTER DISKROUP *diskgroupname* MOUNT command is extended to allow for ASM to mount the diskgroup in restricted mode. Here is an example of the sequence of commands:


```
ALTER DISKGROUP data DISMOUNT;  
ALTER DISKGROUP data MOUNT RESTRICT;Maintenance task: You can add/drop  
online/offline disks, drop files, create directories, etc.ALTER DISKGROUP data DISMOUNT;  
ALTER DISKGROUP data MOUNT;
```

When you use the RESTRICTED option to startup an ASM instance, all of the disk groups defined in the ASM_DISKGROUPS parameter are mounted in RESTRICTED mode.

Preferred Mirror Read

When ASM is managing redundancy, you can configure an ASM instance on a node to read from a preferred mirror copy (i.e. a preferred read failure group). The default behavior is to always read from the primary copy.

This feature is beneficial when you have an extended Oracle RAC cluster, where the nodes and the failure groups are separated by a long distance to enable disaster recovery. In this case, the Oracle RAC nodes on each site can be configured to read from their local storage mirror copies instead of going through a network with potentially high latencies.

Set the ASM_PREFERRED_READ_FAILURE_GROUPS initialization parameter to specify a list of failure group names as preferred read disks. This parameter is a multi-valued parameter and should contain a string with a list of *<disk group name>.<failure group name>* separated by a comma. The new column PREFERRED_READ has been added to the V\$ASM_DISK and V\$ASM_DISK_IOSTAT views. If the disk group that the disk is in belongs to a preferred read failure group, then the value of this column is Y.

To set up preferred mirror read between sites A and B, the failure group name SITEA or SITEB should be preceded by the disk group name:

```
ASM_PREFERRED_READ_FAILURE_GROUP=DATA.SITEA or
```

```
ASM_PREFERRED_READ_FAILURE_GROUP=DATA.SITEB
```

ASM ROLLING UPGRADES AND PATCHING

The rolling upgrade feature is a huge win and significantly increases our database availability and management flexibility for patching and upgrades.

In an Oracle RAC environment, the ASM rolling upgrade feature enables you to independently upgrade or patch ASM nodes in a cluster, thus providing greater flexibility and uptime. This feature is available starting from Oracle Database 11g forward.

Before patching or upgrading ASM, you must place the ASM cluster in a 'rolling migration' mode. This enables an ASM instance to operate in a multi-version environment. Rolling migration mode is expected to be a short-term state. You should have the appropriate software installed on all nodes of the cluster prior to entering rolling migration mode. Oracle Clusterware must be upgraded to the appropriate version on all nodes prior to upgrading ASM. All normal database file

operations are allowed, but most disk group configuration changes are prevented while in rolling migration mode.

```
ALTER SYSTEM START ROLLING MIGRATION TO number;
```

Now you can shut down each ASM instance individually and perform a software upgrade on that instance. The upgraded ASM instance can then rejoin the cluster after the upgrade because it is in rolling migration mode. After all the nodes are successfully upgraded, you can end the rolling migration mode to return to full functionality normal operation.

```
ALTER SYSTEM STOP ROLLING MIGRATION;
```

You can also use the same procedure to roll back (downgrade) nodes if you encounter problems during the migration. You cannot enter rolling migration mode if there are rebalancing operations occurring anywhere in the cluster. You must wait until the rebalance completes before initiating a rolling migration.

New instances that join the cluster during a migration, switch to a rolling migration state immediately upon startup. You can use the following SQL function to identify the state of a clustered ASM environment:

```
select SYS_CONTEXT('sys_cluster_properties', 'cluster_state') from dual;
```

show output form this command

Tablespace level Migration Wizard in Enterprise Manager

ASMMigrate is an Oracle Enterprise Manager (EM) feature that migrates a database residing on file system or raw devices to one residing on Automatic Storage Management (ASM) on the same host.

ASMMigrate will now support tablespace level migration in addition to the whole database migration introduced in Oracle Database 10g Release 2.

In Oracle Database 10g Release 2, we introduced new EM features to extend ASM to be configured on other nodes of the cluster if ASM was configured on only a subset of the cluster hosts. In addition, we provided the capability to extend ASM transparently when we extend the RAC database to a new node that ASM is not already configured on.

Starting with Oracle Database 11g, if you have a single instance ASM on one node, we can convert this to a "clustered" ASM that runs on all or specified list of the nodes in the cluster.

MANAGEABILITY ENHANCEMENTS

Disk Group Compatibility Attributes

Two new disk group compatibility attributes are introduced in Oracle Database 11g ASM. The new *compatible.asm* and *compatible.rdbms* disk group compatibility attributes determine the minimum version of ASM and database instances that can connect to an ASM disk group.

If you upgrade your ASM and database version from 10.1 to 11.1, you have the option to keep the ASM disk compatibility attribute at 10.1. In this case, your ASM and database 11.1 will function but not be able to take advantage of all the 11.1 new features yet. This is valuable since this allows you to upgrade your software and make sure you are happy with the release before committing. You can advance the disk group compatibility attribute from 10.1 to 11.1 to enable all the features in 11.1. You can revert back to the previous release (10.1) if you have not advanced the disk group compatibility attributes. Care must be taken to advance the attributes since this action is irreversible.

You can set the disk group compatibility by using either the CREATE DISKGROUP or ALTER DISKGROUP commands.

```
CREATE DISKGROUP DATA NORMAL REDUNDANCY
DISK '/dev/raw/raw1','/dev/raw/raw2'
ATTRIBUTE 'compatible.asm'='11.1' compatible.rdbms='11.1'
```

The following features will be enabled when you advance asm and rdbms disk group compatibility attributes:

Advancing compatible.asm or compatible.rdbms attributes from 10.1 to 11.1 will enable the following key features:

- Preferred mirror read
- Variable size extents
- Fast mirror resync

New ASMCMD commands

We have introduced three new functionalities to the ASMCMD utility to improve node recovery after a crash, repair bad blocks on a disk, copy files and simplify the listing of ASM disks in a disk group.

ASMCMD cp

The ASMCMD cp option allows you to copy files between ASM disk groups and OS file systems and between two ASM servers. The following file copy permutations are supported:

ASM disk group → OS file system

OS file system → ASM disk group

ASM disk group → Another ASM disk group on the same server

ASM disk group → ASM disk group on a remote server

This example shows an ASM file being copied to a file system directory /backups:

```
asmcmd cp +DG1/VDB.CTF1 /backups
ASMCMD> cp +DG1/vdb.ctf1 /backups/vdb.ctf1
copying file(s)...
copying file(s)...
```

```
copying file(s)...  
file, /backups/vdb.cfb1, copy committed.
```

ASMCMD disk group metadata backup and restore

ASMCMD is extended to include ASM disk group metadata backup and restore functionality. This provides the ability to recreate a pre-existing ASM disk group with the same disk paths, disk names, failure groups, attributes, templates, and alias directory structure. Currently if an ASM disk group is lost, then it is possible to restore the lost files using RMAN but you have to manually recreate the ASM disk group and any required user directories/templates. There is no need to backup and restore ASM metadata itself.

ASM metadata backup and restore works in two modes. In backup mode it gathers information about existing disks and failure group configurations, attributes, templates, and alias directory structures and dumps it to a backup file. In restore mode, it reads the previously generated file to reconstruct the disk group and its metadata. You have the possibility to control the behavior in restore mode to do a full, nodg, or newdg restore. The 'full' mode restores the disk group exactly as it was at the time of backup. The 'nodg' mode restores the attributes, templates, and alias directory structure specified in the backup file to an existing disk group. The 'newdg' mode allows the user to override the disk group name, disk, and failure group specifications as part of a disk group creation, but retains the attribute, template, and alias directory structure from the backup. This example describes how to backup ASM metadata using the md_backup command, and how to restore them using the md_restore command.

```
ASMCMD> md_backup -b backup_file -g data  
Disk group to be backed up: DATA#  
ASMCMD>
```

The above statement specifies the -b option and the -g option of the command. This is to define the name of the generated file containing the backup information as well as the disk group that needs to be backed up.

Before you can restore the database files it contained, you have to restore the disk group itself. Let's assume that a storage array failure lost all of the disks in disk group DATA. In this case, you initiate the disk group recreation as well as restoring its metadata using the md_restore command. Here, you specify the name of the backup file generated at step one, as well as the name of the disk group you want to restore, and also the type of restore you want to do. Here a full restore of the disk group is done because it no longer exists.

```
ASMCMD> md_restore -b jfv_backup_file -t full -g data  
Disk group to be restored: DATA#  
ASMCMDAMBR-09358, Option -t newdg specified without any override options.  
Current Diskgroup being restored: DATA  
Diskgroup DATA created!  
User Alias directory +DATA/jfv  
created!  
ASMCMD>
```

Once the disk group is recreated, you can restore its database files using RMAN for example.

ASMCMD lsdisk

The `lsdisk` command lists ASM disk information. This command can run in two modes: connected and non-connected. In connected mode, ASMCMD uses the `V$` and `GV$` views to retrieve disk information. In non-connected mode, ASMCMD scans disk headers to retrieve disk information, using an ASM disk string to restrict the discovery set. The connected mode is always attempted, first. This is very useful for system or storage administrators who want a consolidated list of disks with ownerships.

```
ASMCMD> lsdisk -kspt '/dev/sdb6'
```

ASMCMD remap

When a normal read from an ASM disk group fails with an I/O error ASM satisfies the write from a mirrored extent in Normal and High redundancy mode. ASM attempts to remap the extent where the read failed by relocating it to another allocation unit on the same disk with the contents from the good mirror. If the remap succeeds on the same disk, the old allocation unit is marked as unusable. This process happens automatically only on blocks that are read. It is possible that some blocks and extents on an ASM disk group are seldom read; e.g. the secondary extents. The ASMCMD `remap` invokes bad block remapping for the specified blocks. One can use the ASMCMD `remap` command if the storage array returns an error on a physical block, then the ASMCMD `repair` can initiate a read on that block to trigger the repair.

```
ASMCMD> asmcmd remap DATA DATA_00015000-7500
```

SYSASM role

This feature introduces a new SYSASM role that is specifically intended for performing ASM administration tasks. Using the SYSASM role instead of the SYSDBA role separates ASM administration from database administration.

In Oracle Database 11g, the OS group for SYSASM and SYSDBA is the same, and the default installation group for SYSASM is `dba` but they can be different. In future releases, separate OS groups will have to be created to separate the SYSDBA and SYSASM privileges. SYSDBA will only be able to query views and manipulate ASM files. SYSASM will be required to manage disk group configurations.

In this release, you can connect to an ASM instance as a member of the OSASM, which defaults to the `dba` group.

```
SQL> CONNECT / AS SYSASM
```

You also have the possibility to use the combination of `CREATE USER` and `GRANT SYSASM SQL` statements from an ASM instance to create a new SYSASM or SYSDBA user. This is possible as long as the name of the user is an

existing OS user name. These commands update the password file for the local ASM instance, and do not need the instance to be up and running. Similarly, you can revoke the SYSASM role from a user using the REVOKE command, and you can drop a user from the password file using the DROP USER command.

```
SQL> CREATE USER ossysasmusername IDENTIFIED by passwd;SQL> GRANT SYSASM TO ossysasmusername;  
SQL> CONNECT ossysasmusername / passwd AS SYSASM;  
SQL> DROP USER ossysasmusername;
```

Note: With Oracle Database 11g Release 1, if you log in to an ASM instance as SYSDBA, warnings are written in the corresponding alert.log file when you issue commands that alter disk group configuration.

Mount/Drop FORCE disk group

With Oracle Database 11g, ASM will fail to mount a disk group if there are any missing disks or failure groups during mount, unless the new FORCE option is used when mounting the incomplete disk group. This allows you to correct configuration errors like ASM_DISKSTRING set incorrectly, or connectivity issues before trying the mount again without incurring unnecessary rebalance operations.

```
ALTER DISKGROUP data MOUNT FORCE|[NOFORCE];
```

Disk groups mounted with the FORCE option will have one or more disks offline if they were not available at the time of the mount. You must take corrective actions before DISK_REPAIR_TIME expires to restore those devices. Failing to online those devices would result in the disks being expelled from the disk group and the need for a rebalance operation to restore redundancy for all the files in the disk group.

Mount with NOFORCE is the default mode In the NOFORCE mode, all of the disks that belong to a disk group must be accessible for the mount to succeed.

Drop disk group FORCE marks the headers of disks belonging to a disk group that cannot be mounted by the ASM instance as FORMER. However, the ASM instance first determines whether the disk group is being used by any other ASM instance using the same storage subsystem. If it is being used, and if the disk group is in the same cluster, or on the same node, then the statement fails. If the disk group is in a different cluster, then the system further checks to determine whether the disk group is mounted by any instance in the other cluster. If it is mounted elsewhere, then the statement fails. Please exercise caution when using the force option. DROP DISKGROUP FORCE is much safer than using dd to overwrite disk headers.

```
DROP DISKGROUP data FORCE INCLUDING CONTENTS;
```

CONCLUSION

The new ASM features in Oracle Database 11g further extend the power and simplicity of managing storage in an Oracle grid environment. ASM is a proven technology with over one thousand customers worldwide. ASM eliminates the need

for third party file systems and volume managers for the database files minimizing cost, improving storage utilization and consolidation that empowers reliable low cost computing across single instance Oracle Database and Oracle RAC environments.



White Paper Title
[Month] 2007
Author: Ara Shakian
Contributing Authors:

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Copyright © 2007, Oracle. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.