Session 3

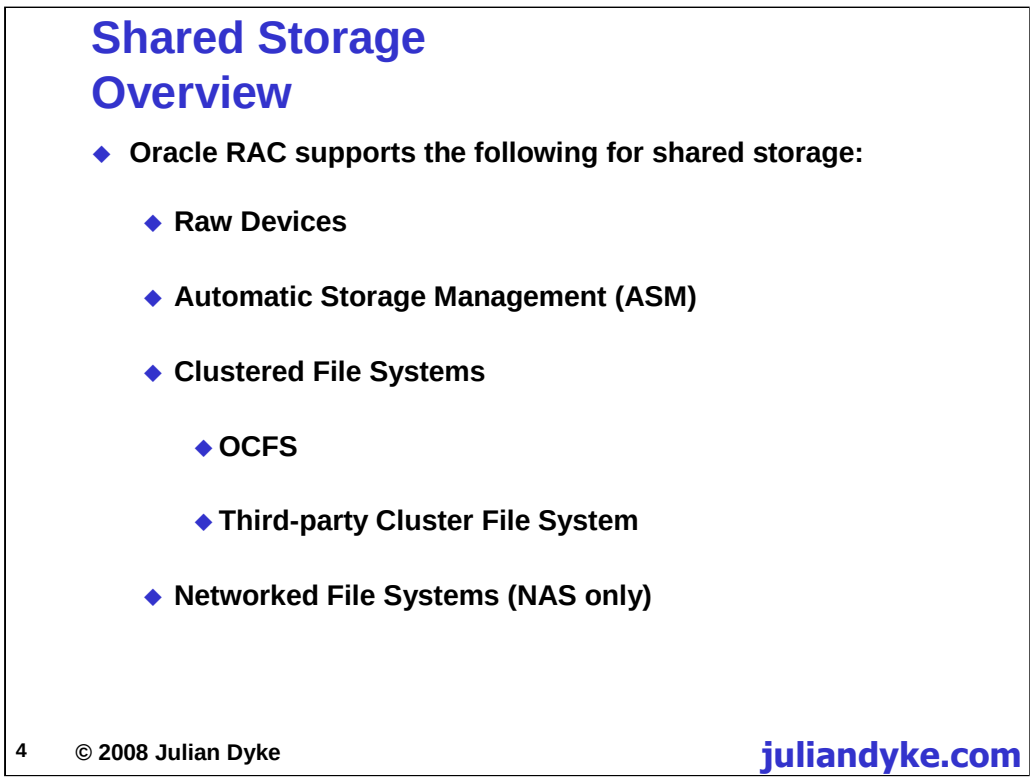
Storage

Advanced RAC
Auckland - May 2008

Agenda

- Shared Storage
- Automatic Storage Management

Shared Storage



Oracle supports four types of shared storage

•Raw Devices

•Automatic Storage Management (ASM)

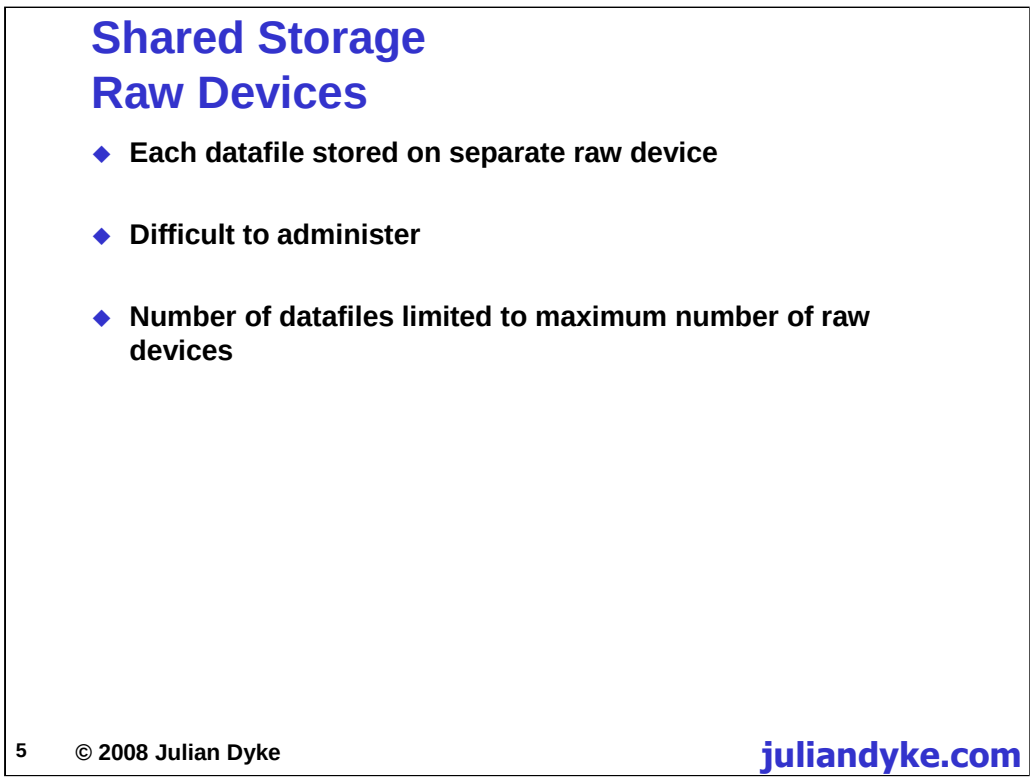
•Clustered File System (CFS)

•Networked File System (NFS on NAS)

Oracle strongly recommends the use of Automatic Storage Management (ASM), which is largely stable in single-instance environments and in RAC-environments implementing external redundancy. There are some outstanding question marks over ASM configurations using double or triple mirroring in early versions of Oracle 10.2 and I would therefore recommend it is only configured with external redundancy in 10.2.0.1 and 10.2.0.2

On Linux Oracle supports the Oracle Cluster File System (OCFS and OCFS2) which are both Open Source products. OCFS packages must be downloaded from the otn.oracle.com website (as opposed to oss.oracle.com) in order to be fully supported. A proprietary version of OCFS which has similar functionality to OCFS2 is supported on Windows A number of other third-party cluster file systems are available which have varying acceptance in the marketplace.

Oracle only supports limited versions of NFS running on Network Attached Storage (NAS). NAS-platforms must be certified as part of the OSCP program in order to be fully supported.



For many years, administrators have implemented databases on raw devices. This is because raw devices are generally faster than file systems (perhaps 10%). Raw devices do not use the file system cache when reading and writing blocks. This avoids the overhead of double buffering which often represents a significant cost in terms of CPU and memory.

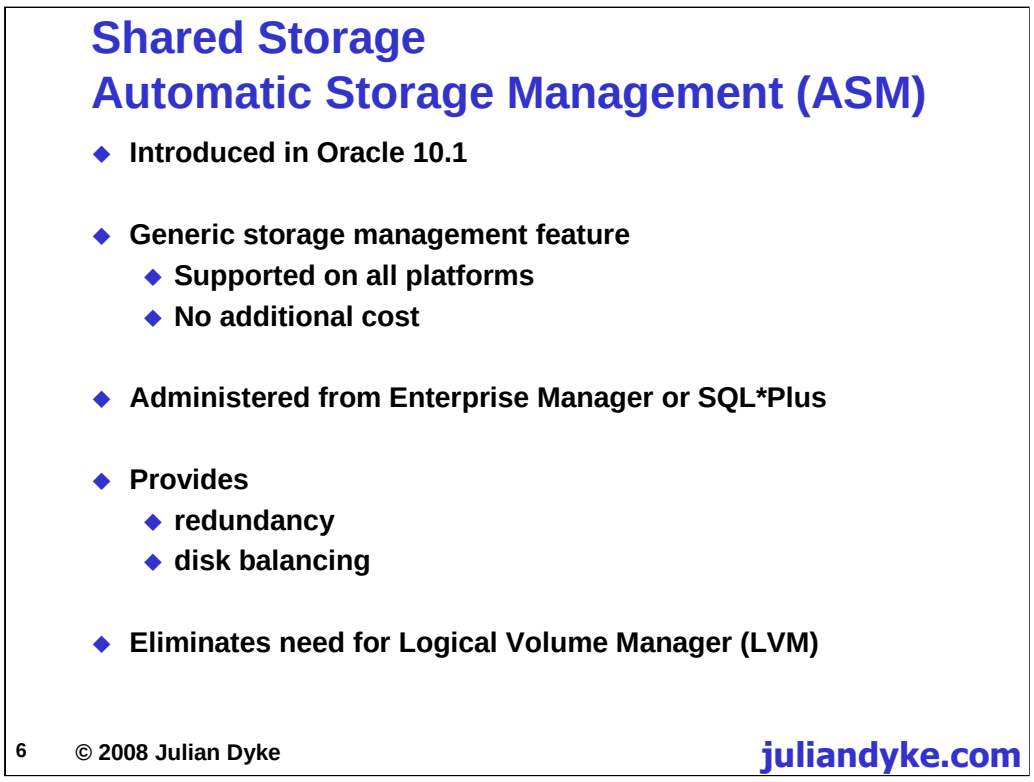
However, raw devices can be difficult to manage, particularly in a clustered environment. In an Oracle database each datafile must be placed in a separate raw device.

Raw devices were popular in early RAC versions (Oracle 9.0.1 and sometimes Oracle 9.2). It is very rare to see them in Oracle 10.1 and above as sites will normally deploy ASM. Because they present some significant manageability issues. users frequently deploy raw devices within the framework of a Logical Volume Manager (LVM).

Archived redo logs cannot be stored on raw devices. Therefore file systems are required these files. The file systems can be local file systems or clustered file systems located on central storage. If local file systems are used, then in a RAC environment they are normally exported using NFS to enable all other nodes in the cluster to access the archived redo logs.

In Red Hat Advanced Server 2.1, the maximum number of raw devices is 128. In Red Hat Enterprise Linux 3.0 and above, this limit is increase to 256 raw devices. This still may not be sufficient for some large databases.

In recent years, the performance advantage of raw devices has narrowed with the introduction of direct I/O. I am not aware of any new Oracle 10g RAC deployments on raw devices.

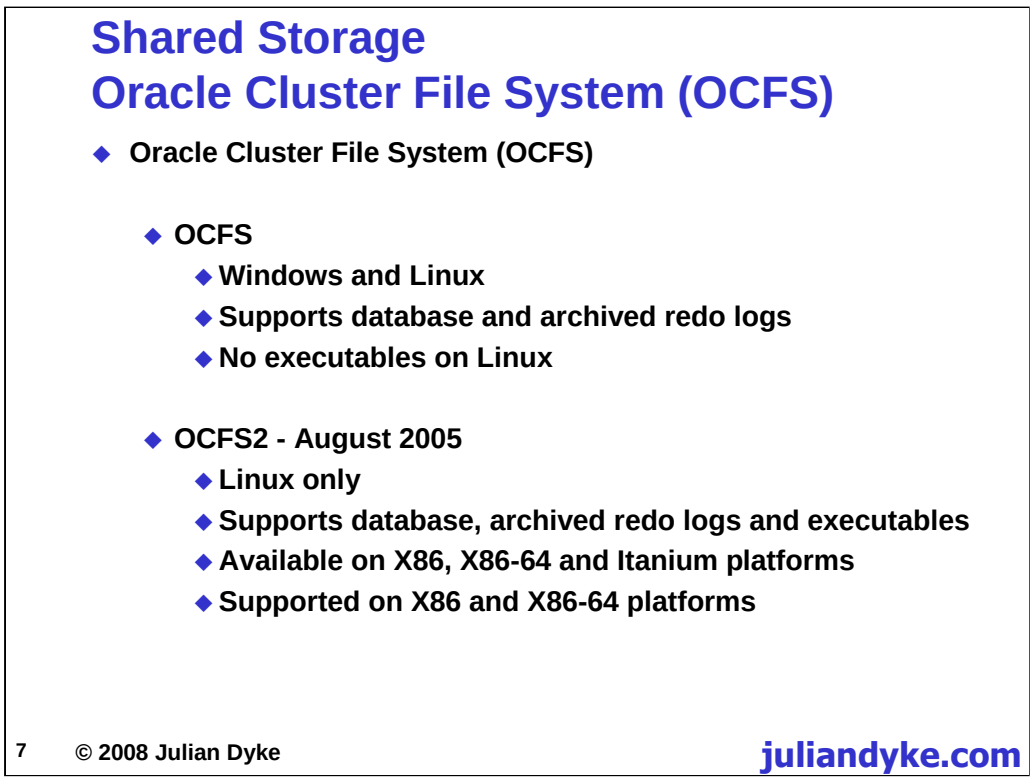


ASM was introduced in Oracle 10.1. In a RAC environment ASM provides a shared storage subsystem. Some users still use an intermediate layer such as Veritas. However, it is possible to deploy ASM directly on the LUNs presented by the SAN thus eliminating the need for a Logical Volume Manager (LVM). Where the database shares storage with other applications, an LVM may still be necessary.

Unlike OCFS which is only available on Linux and Windows, ASM is a generic storage management feature which is available on all Oracle 10g-supported platforms.

In addition to providing all LVM functionality for Oracle databases, ASM also provides redundancy and disk balancing capabilities. ASM file systems can be configured to provide double or triple mirroring. However, the majority of users specify external mirroring in which case mirroring is performed by the underlying storage subsystem e.g. SAN or NAS. This is generally more efficient in terms of resource consumption and avoids diverting expensive CPU cycles from activities directly related to th e database on the Oracle-licensed servers.

In Oracle 10.2 and above, the contents of an ASM file system are more transparent due to the introduction of the ASMCMD command line utility and also an FTP interface.

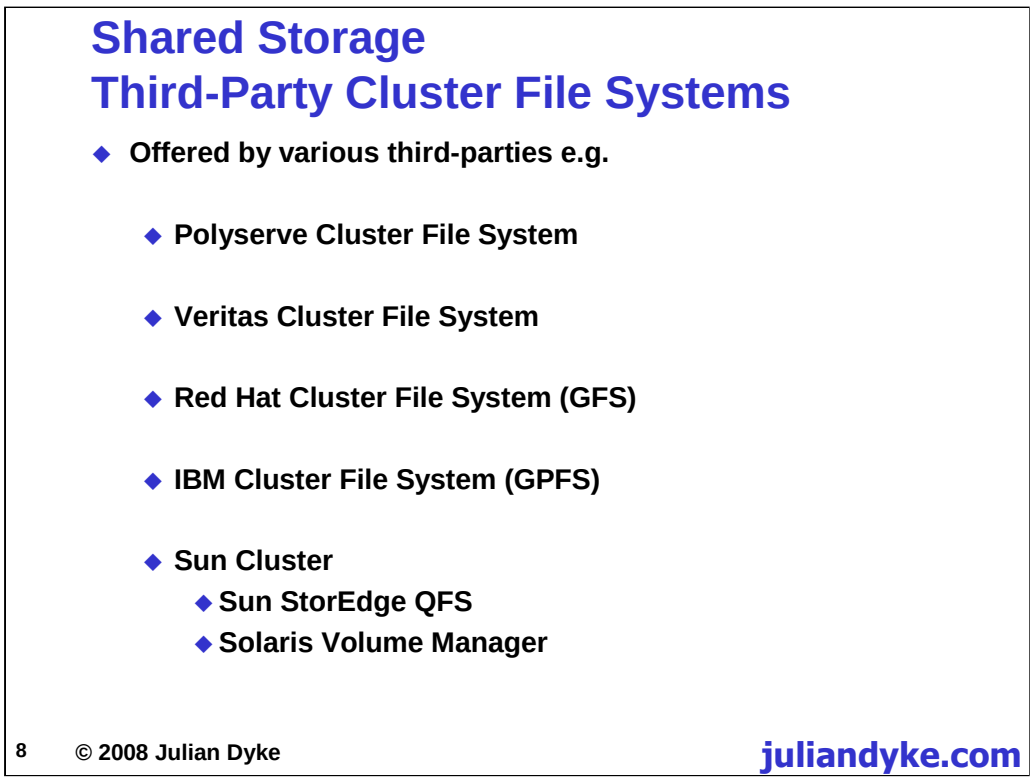


OCFS is an Open Source cluster file system developed by Oracle for the Linux platform. It is also available for Windows, though this version is not Open Source.

OCFS Version 1 is supported on Red Hat Advanced Server 2.1 and Red Hat Enterprise Linux 3.0. It is not supported on Red Hat Enterprise Linux 4.0. OCFS Version 1 supports database files and archived redo logs, but does not support Oracle binaries. Therefore it is not possible to create a shared Oracle home on OCFS Version 1.

OCFS Version 2 is supported on Red Hat Enterprise Linux 4.0. It is not, however, supported on Red Hat Advanced Server 2.1 and Red Hat Enterprise Linux 3.0. OCFS Version 2 supports database files, archived redo logs and Oracle binaries. It is therefore possible to create a shared Oracle home on OCFS Version 2.

However, there are some question marks over the long term future of OCFS2. Oracle strongly recommends ASM which is a generic product available on all Oracle-supported platforms. On the other hand, OCFS2 is only available on Linux. In addition, OCFS2 was available for download from August 2005. However, support was not announced for the X86 version on Red Hat until February 2006. Support was not announced for the X86-64 version until May 2006. This does not appear to be an overwhelming endorsement of the technology. I strongly recommend therefore, that you avoid using OCFS2 until its future becomes more definite.



Various third-parties offer cluster file systems at additional cost. Because of the additional cost, most RAC users choose to deploy ASM or OCFS.

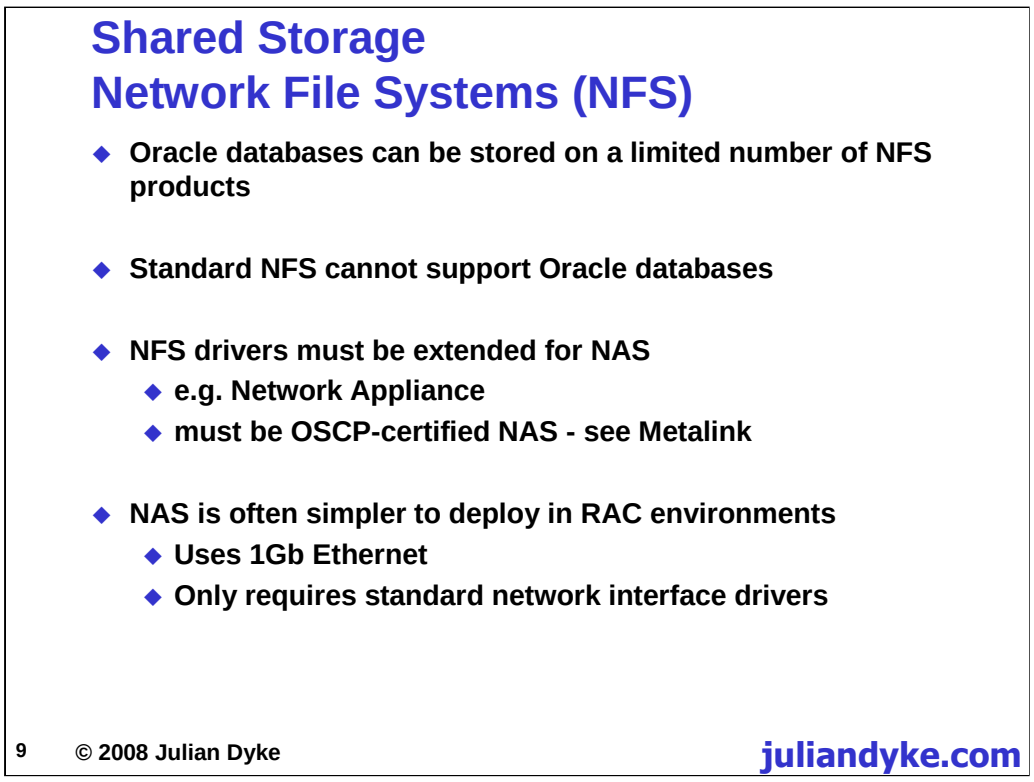
Polyserve products provide an additional clustering layer between RAC and the operating system. Whilst Polyserve is undoubtedly better than the standard Oracle RAC offerings, it is not always cost effective for smaller sites as Polyserve targets sites with 16 or more nodes.

An Oracle RAC version of the Veritas Cluster File System was available in Oracle 9i. However, few RAC sites have adopted this. The marketing of this product for Oracle 10g has been very understated which suggests that, perhaps it is a product which should be avoided at present.

Red Hat Cluster File system is an additional cost CFS for Red Hat environments. It does not offer significantly different functionality to OCFS2 and has seen little uptake amongst RAC users.

A few RAC users have experimented with the IBM Cluster File System (GPFS), but again, few have deployed it into production.

Finally, RAC is supported on Sun Cluster. RAC can be deployed on either Sun StorEdge QFS or Solaris Volume Manager. Both options require Sun Cluster licences. There does appear to be some uptake of these products amongst RAC users. They are, of course, platform-specific and therefore can only be used on Sun hardware.



Oracle databases can be stored on a limited number of NFS products.

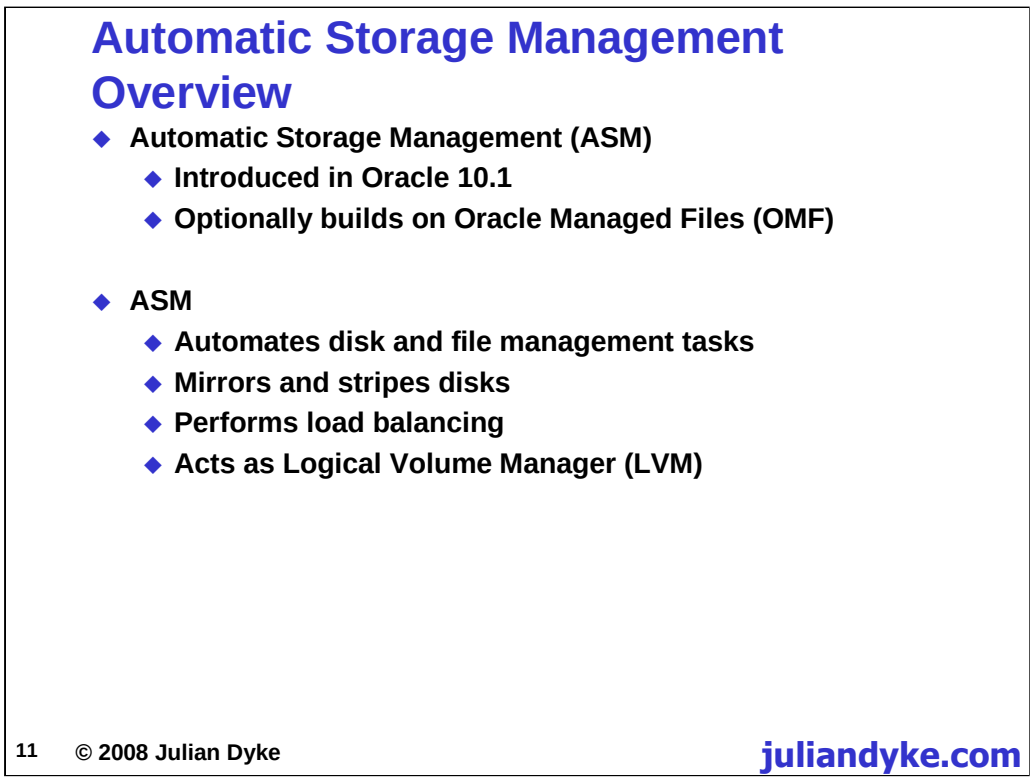
Oracle RAC databases are not supported on standard NFS. However, some extended NFS drivers are supported on Network Attached Storage (NAS). If you wish to use NAS for shared storage then this must be on the OSCP certified list. See Metalink for details.

NAS is often simpler to deploy in a RAC environment. Usually a dedicated 1Gb Ethernet is used to connect the cluster nodes to the NAS. However Network Appliance storage also supports iSCSI and fibre channel.

Remember that NFS and iSCSI are more resource intensive on the cluster nodes as additional CPU is required to build the network messages required to support I/O. It is possible, though much more expensive to purchase network cards which perform additional processing on-board, thus alleviating pressure on the CPUs. This may be a cost-effective option, particular for users with Enterprise Edition licenses.

See Metalink Note 359515.1- Mount options for Oracle files when used with NAS devices

Automatic Storage Management



In all database systems disk I/O is much slower the memory access and is often the single biggest bottleneck

Automatic Storage Management (ASM) was introduced in Oracle 10.1. It optionally uses a feature called Oracle Managed Files (OMF) which was introduced in Oracle 9.0.1. OMF automatically assigns names to datafiles and also creates and drops datafiles thereby eliminating some management effort.

ASM automates disk and file management tasks. It performs disk striping and mirroring. It also performs load balancing.

ASM acts as Oracle's own Logical Volume Manager (LVM). It therefore may not be necessary to deploy a third-party LVM.

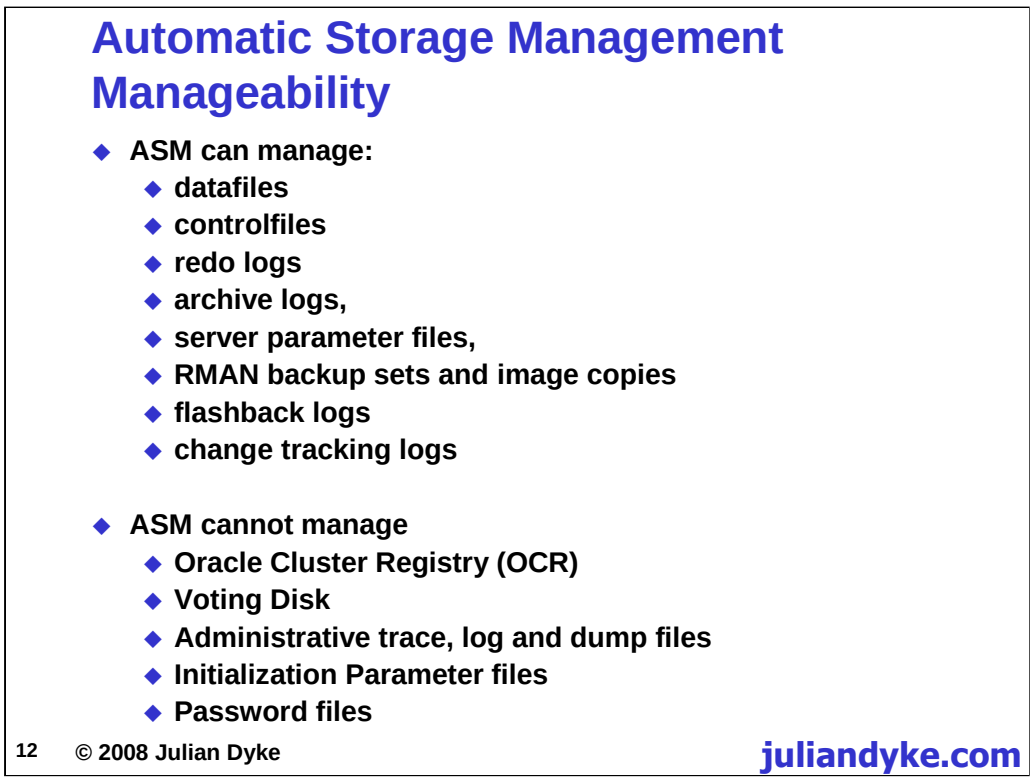
You can combine operating system, OMF and ASM files in same database, although this is rarely done in practice. It may, however, be useful to support complex migrations.

The main benefits of ASM are reductions in deployment costs and management overheads.

Cannot use operating system commands or utilities to access ASM files

Must use RMAN to copy ASM files

Logical concepts including segments, extents and tablespaces unchanged when using ASM



ASM can manage a number of different file types including

- datafiles
- controlfiles
- redo logs
- archived redo logs
- server parameter files (SPFILE)
- RMAN backup sets
- RMAN image copies
- flashback logs
- change tracking logs

However ASM cannot manage a number of files that are essential in a RAC environment including

- Oracle Clusterware Registry (OCR)
- Voting Disks
- Administrative files (trace, log and dump files)
- Initialization Parameter files (PFILE) and Password Files (where required)

Most ASM users assign raw devices to the OCR and Voting Disks and use local file systems for the adminstrative files.

Automatic Storage Management Advantages

- Reduces deployment costs
  - ASM is free
  - May not require third-party LVM
  - Oracle DBA can perform most disk management tasks
  - Do not need expert in file systems, RAID or logical volumes
- Reduced management overhead
  - Only maintain a few disk groups instead of maintaining a large number of datafiles
  - Easier to manage large databases
  - Logical concepts including segments, extents and tablespaces unchanged in ASM
- Integrated with Oracle Managed Files (OMF)
  - Tablespaces can be managed without specifying filenames

The advantages of using Automatic Storage Management include reduced deployment costs. ASM is free with both Standard Edition and Enterprise Edition licences. In RAC configurations ASM can often be used to replace third-party LVMs such as Veritas. Once ASM has been configured, the Oracle DBA can perform most disk management tasks avoiding the need to use a systems administrator to create file systems, RAID devices or logical volumes. It will still be necessary, however, to have a storage administrator who can create additional LUNs or expand existing meta-LUNs as the size of the database increases.

ASM can reduce the management overhead as it is only necessary to maintain a few disk groups instead of maintaining a large number of datafiles. This is particularly an advantage over using raw devices for shared database storage. It is easier to manage large databases.

Logical concepts including segments, extents and tablespaces are unchanged in ASM

ASM is integrated with Oracle Managed Files (OMF) which means that tablespaces can be managed without specifying filenames.

Automatic Storage Management Advantages

- Performs mirroring and striping
  - increases reliability and performance
  - mirroring is applied at file level (as opposed to disk level) giving more control Automatically balances I/O load in parallel across all available disk drives
  - eliminates hotspots
  - maximizes performance
- When new disks are added ASM automatically moves data around to balance I/O load among disks
  - ASM load-balances file activity by uniformly distributing file extents across all disks in a disk group
- Maintains redundant copies of data to provide fault tolerance
  - Can optionally build ASM storage system on vendor-supplied storage mechanisms

ASM can perform both striping and mirroring better than third party LVMs because ASM understands Oracle file types and uses appropriate strategy for each type. In ASM data is managed by specifying desired reliability and performance characteristics for classes of data

Automatic Storage Management Advantages

- Available for all supported platforms in Oracle 10.1 and above
  - Generic code
  - Non-proprietary
  - OCFS2 only available on Linux
- Included with Oracle database software
  - OCFS2 must be separately downloaded
- Does not require separate licence
  - Third-party storage managers usually additional cost
- Tested by Standard Edition RAC users
  - ASM is mandatory in this configuration
  - High-level of deployment in Oracle 10.1 and above
- In Oracle 10.2 and above
  - Can support multiple versions of database software
  - Increases upgrade options

ASM is available on all supported platforms in Oracle 10.1 and above. It is generic code (i.e. not platform specific). The only Oracle-supplied alternative is OCFS2 which is only available on Linux.

ASM is included with the standard database software distribution. OCFS2 must be downloaded separately.

ASM does not require a separate licence. Other third-party storage managers such as Red Hat GFS, IBM GPFS, Veritas and Sun StorEdge all require additional licences.

The free RAC option for Standard Edition users on condition that ASM was used for shared storage was a stroke of genius leading to a higher rate of adoption of ASM during the first release (10.1) than is normally seen with other new features. This means that ASM has been tested more thoroughly than many other features would have been at the same stage in their development. By definition however it has only been tested on smaller systems that can operate under the limitations of the Standard Edition licence.

In Oracle 10.2 and above, ASM can support multiple versions of the database software. This increases the upgrade options. For example, you can upgrade the ASM version first, while the databases stay on the old version. Then you can upgrade all or some of the databases on another occasion.

Automatic Storage Management Disadvantages

- Cannot use operating system commands or utilities to access ASM files
  - Must use RMAN to copy ASM files
  - Impacts organizations with separate storage teams
- Requires separate ASM instance on each node
  - Additional CPU and memory
  - Additional background processes / daemons
- Cannot store:
  - Oracle Clusterware files (OCR and Voting Disk)
  - Oracle executables / shared Oracle home
  - Administrative files
    - Trace or log files
    - Parameter files
    - Password files

ASM also has a number of disadvantages.

You must use RMAN to copy ASM files - there are no command line utilities to copy ASM files though you can use the get and put commands in ASM.

A separate ASM instance is required. This uses additional CPU and memory for the SGA, background processes and, if configured, the executables and libraries in any separate ASM home.

You cannot store Oracle Clusterware files (the Oracle Cluster Registry (OCR) and CSS Voting Disk) in ASM. In Oracle 10.2 you cannot store the Oracle executables or shared Oracle home in ASM. Finally you cannot store administrative files such as trace or log files, parameter files and password files in ASM.

Automatic Storage Management Components

- ASM has three components
  - ASM instance
  - ASM disk groups
  - ASM files

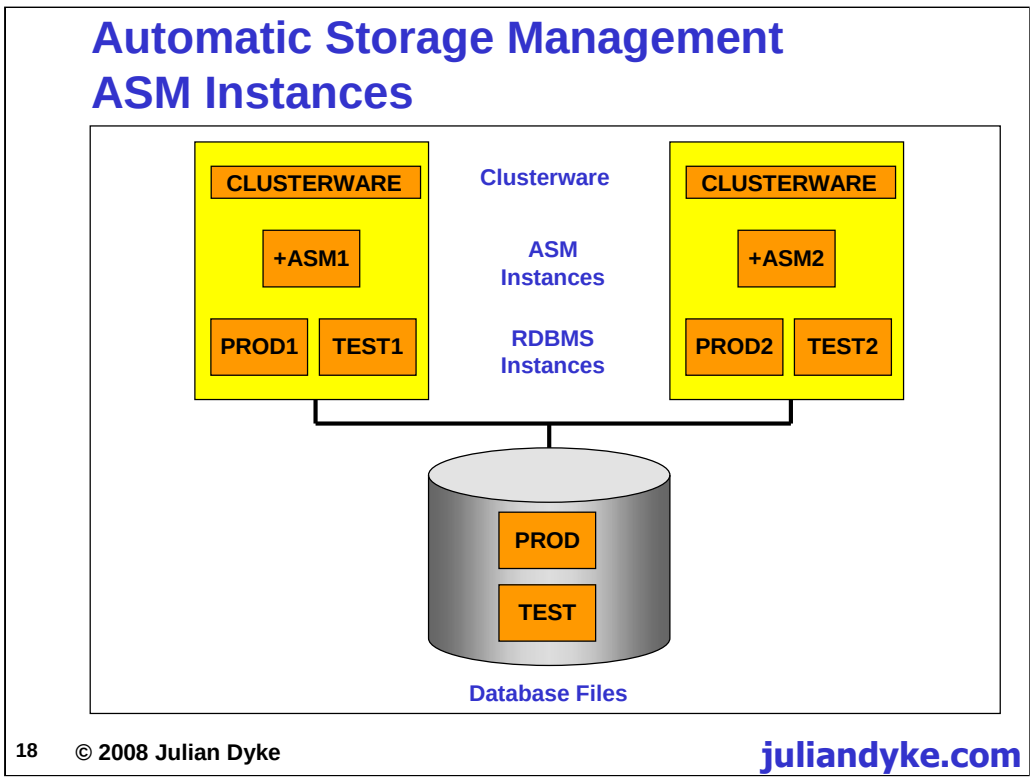
ASM has three components

• ASM instance

• ASM disk groups

• ASM files

These components will be discussed over the next few slides.



This slide shows the relationship between Clusterware, ASM instances and RDBMS instances.

On a cluster running ASM, clusterware will run on each node. Clusterware includes the following daemons or services:

- CRS - in Unix crsd
- CSS - in Unix ocssd
- EVM - in Unix evmd
- PROCD in Unix oprocd

There will normally be one ASM instance on each node. These will be called +ASM1, +ASM2 etc. Each ASM instance has a complete set of background processes.

There will be one RDBMS instance for each database in the cluster. In the above slide these are called PROD1, PROD2 for the PROD database and TEST1, TEST2 for the TEST database. Each RDBMS instance has a complete set of background processes.

ASM Instance Overview

- ASM Instance
  - Similar to small database instance
  - Helps database access ASM files
  - Must be started in order for database to access ASM files
  - Does not mount any Oracle database files
  - Does not have data dictionary
  - Only maintains ASM file metadata
  - Oracle database
    - contacts ASM instance for information about ASM datafiles
    - accesses ASM files directly

ASM has its own instance and background processes. An ASM instance must be running on each node in the cluster for the database to be able to use the ASM file system.

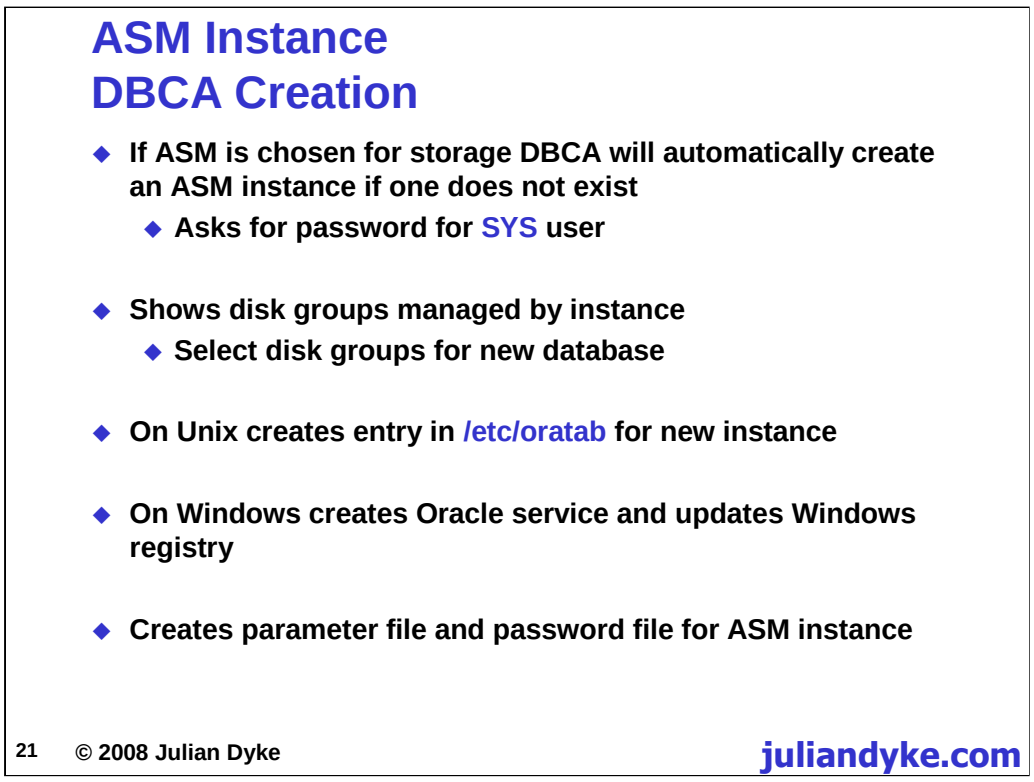
An ASM instance is similar to a small database instance. It helps the database to access the ASM files.

ASM Instance Overview

- ASM Instance
  - Manages and protects disk groups
  - All disk group management commands must be issued from within ASM instance
    - Database has no direct connection to ASM disk groups
  - Managed by administrator using
    - Operating system authentication with SYSDBA or SYSOPER privilege
    - Password file
  - Created using DBCA (recommended) or manually

ASM instances can be created using the DBCA or manually. Although manual creation is trivial, I recommend using the DBCA as this automatically updates the Oracle Cluster Registry (OCR) for each node in the cluster.

An ASM instance normally requires approximately 100MB of disk space and a 64MB SGA.



During ASM instance configuration, the DBCA performs the following actions:

• If ASM is chosen for storage DBCA will automatically create an ASM instance if one does not exist. A password will be requested for the SYS user.

• The DBCA shows the disk groups that are currently managed by the instance and allows you to create disk groups for the new databases.

• During installation on Linux an entry is written to /etc/oratab for the new instance. For example

```
+ASM1:/u01/app/oracle/product/10.2.0/db_1:N
```

• During installation on Windows an Oracle service is created and the Windows registry is updated.

• A parameter file and password file are also created for the ASM instance.

ASM Instance Privileges

- ASM instances can only be managed by users with SYSDBA or SYSOPER privilege:
  - Only SYSDBA can
    - create ASM instances
    - CREATE DISKGROUP
    - ADD DISK / DROP DISK
    - ALTER DISKGROUP RESIZE
  - Both SYSDBA and SYSOPER can:
    - STARTUP/SHUTDOWN ASM instances
    - ALTER DISKGROUP CHECK
    - ALTER DISKGROUP MOUNT/DISMOUNT/REMOUNT
    - ALTER DISKGROUP OFFLINE
    - access ASM instance dynamic performance views

22 juliandyke.com

ASM instances can only be managed by users with the SYSDBA or SYSOPER privilege.

You must have SYSDBA privileges to

• create ASM instances

• CREATE DISKGROUP

• ADD DISK / DROP DISK

• ALTER DISKGROUP RESIZE

Both SYSDBA and SYSOPER can:

• STARTUP / SHUTDOWN ASM instances

• ALTER DISKGROUP CHECK

• ALTER DISKGROUP MOUNT / DISMOUNT / REMOUNT

• ALTER DISKGROUP OFFLINE

• access ASM instance dynamic performance views

ASM Instance Parameter Files

- Each ASM instance must have a parameter file
  - Can be text or SPFILE
- Only requires a handful of initialization parameters
- For example:

```
INSTANCE_TYPE = ASM
DB_UNIQUE_NAME = +ASM
ASM_POWER_LIMIT = 1
ASM_DISKSTRING = ‘/dev/sdc1’,’/dev/sdd1’
ASM_DISKGROUPS = DATA, RECOVERY
LARGE_POOL_SIZE = 16M
```


Each ASM instance must have a parameter file. This can be a text file (PFILE) or a binary file (SPFILE)

Only a handful of initialization parameters are required.

For example:

```
INSTANCE_TYPE = ASM
DB_UNIQUE_NAME = +ASM
ASM_POWER_LIMIT = 1
ASM_DISKSTRING = ‘/dev/sdc1’,’/dev/sdd1’
ASM_DISKGROUPS = DATA, RECOVERY
LARGE_POOL_SIZE = 16M
```

ASM Instance
INSTANCE_TYPE parameter

- Introduced in Oracle 10.1
- Specifies type of instance
- Can be
  - RDBMS (default)
  - ASM
- If INSTANCE_TYPE parameter is set to ASM
  - ASM instance will be started
  - All other ASM parameters will have default values

The INSTANCE_TYPE parameter was introduced in Oracle 10.1. This parameter defaults to RDBMS in which case a database instance will be started. If this parameter is set to ASM, then an ASM instance will be started instead.

If the INSTANCE_TYPE parameter is set to ASM, all other ASM parameters will be given appropriate default values which can optionally be overwritten in the parameter file.

ASM Instance
DB_UNIQUE_NAME parameter

- Only required for clustered ASM instances
- Specifies unique name for group of ASM instances within cluster or node
- Default value is +ASM
- Change this parameter if more than one ASM instance running on same node
- Default instance names are
  - +ASM1
  - +ASM2
  - etc

The DB_UNIQUE_NAME parameter is only required for clustered ASM instances. It specifies a unique name for the group of ASM instances within the cluster or node. The default value is +ASM.

You will need to change this parameter if there is more than one ASM instance running on the same node.

By default the instance names are +ASM1 on the first node, +ASM2 on the second node and so on.

ASM Instance
ASM_POWER_LIMIT parameter

- Indicates maximum speed to be used by ASM instance during disk rebalance operations
- When individual disks are added/removed ASM moves around an amount of data equal to the storage added or deleted from the disk group
  - Required to evenly redistribute datafiles and balance I/O load across disks
- Default is 1
- Range is 1 (slowest) to 11 faster
- Rebalance speed can also be specified in POWER clause of REBALANCE command

```
ALTER DISKGROUP dgroup1 REBALANCE POWER 4;
```


The ASM_POWER_LIMIT parameter indicates the amount of resources to be used by the ASM instance during disk rebalance operations.

When individual disks are added/removed ASM moves around an amount of data equal to the storage added or deleted from the disk group. This activity is required to evenly redistribute datafiles and balance I/O load across disks.

The default value for this parameter is 1; the range is from 1 (slowest) to 11 (fastest).

The rebalance speed can also be specified in the POWER clause of the ALTER DISKGROUP REBALANCE command. For example:

```
ALTER DISKGROUP dgroup1 REBALANCE POWER 4;
```

ASM Instance
ASM_DISKSTRING parameter

- Specifies disk location for Oracle to use during disk discovery process
- When new disk added to disk group ASM instance discovers new disk by searching in directories listed in ASM_DISKSTRING parameter
- Default value is NULL
- Can use asterisk as a wildcard
- Example:

```
ASM_DISKSTRING = '/dev/sdc1','/dev/sdd1*';
```


The ASM_DISKSTRING parameter specifies the disk locations which Oracle should use during the disk discovery process.

When a new disk is added to the disk group, the ASM instance discovers that disk by searching in directories listed in the ASM_DISKSTRING parameter.

The default value is NULL.

You can use an asterisk as a wildcard value.

For example:

```
ASM_DISKSTRING = '/dev/sdc1','/dev/sdd1*';
```

ASM Instance
ASM_DISKGROUPS parameter

- Specifies names of disk groups that should be automatically mounted at instance startup
- Default value is NULL
- If text parameter file used then names of disk groups must be added manually
- If SPFILE used then Oracle will update SPFILE when disk groups are created or dropped

The ASM_DISKGROUPS parameter specifies the names of disk groups that should be automatically mounted at instance startup.

The default value is NULL.

If a text parameter file is used then the names of disk groups must be added manually.

If a server parameter file is used, then Oracle will update the SPFILE when the disk groups are created or dropped.

This is a good argument for using a server parameter file if you use the CREATE DISKGROUP and DROP DISKGROUP commands to maintain the disk groups. However, if you use the DBCA to maintain the disk groups, it will update the text parameter files with any changes.

ASM Instance Background Processes

- ASM instance has several background processes including
  - SMON, PMON, LGWR, DBWR, MMAN, PSP0,
  - LMON, LCK0, LMD0, LMSn, DIAG
- Two new background processes
  - RBAL – ASM Rebalance Master
    - ASM Rebalance Master
    - In charge of coordinating disk activity
  - ARBn – ASM Rebalance Slaves (ARB0, ARB1 etc)
    - ARBn Background Process performs rebalancing by moving data extents between disks

29 © 2008 Julian Dyke juliandyke.com

The ASM instance has several background processes including:

• SMON, PMON, LGWR, DBW0, CKPT and MMAN

In a RAC configuration, the ASM instance also has a full set of RAC background processes including:

• LMON, LMD0, LMS0, LCK0 and DIAG

There are also some ASM specific background processes including:

• RBAL - ASM Rebalance Master which coordinates disk activity

• ARBn - ASM Rebalance Processes which perform rebalancing by moving data extents around. Up to 11 ARBn processes can exist which are numbered ARB0 to ARB9 and ARBA.

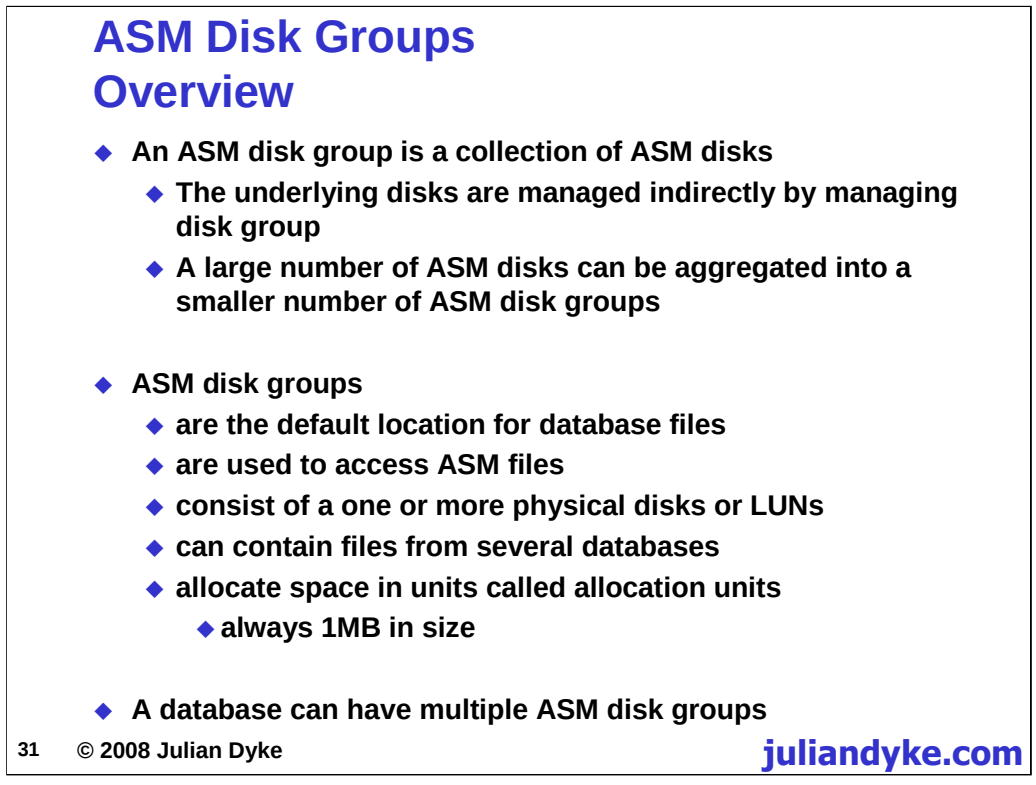
ASM Instance
RDBMS Instances

- Any Oracle database instance using ASM has two new ASM-related background processes
  - RBAL background process
    - Performs global opens of the disks in the ASM disk groups
  - ASMB background process
    - connects as foreground process into ASM instance
    - acts as link between ASM instance and database instance
    - communicates information such as
      - datafile creation
      - datafile deletion
      - updating statistics
      - performing instance health checks

Any Oracle database instance using ASM has two new ASM-related background processes

The RBAL background process performs global opens of the disks in the ASM disk groups

The ASMB background process connects as a foreground process in the ASM instance. It acts as a link between the ASM instance and the database instance. The ASMB background process communicates information such as datafile creation and deletion, statistics updates. It also performs instance health checks.



An ASM disk group is a collection of ASM disks. The underlying disks are managed indirectly by managing disk group. A large number of ASM disks can be aggregated into a smaller number of ASM disk groups

ASM disk groups are the default location for database files. They are used to access ASM files. An ASM disk group consists of a one or more physical disks or LUNs. It can contain files from several databases

Space is allocated in an ASM disk group in units called allocation units. These are always 1mb in size.

A database can have multiple ASM disk groups. An ASM disk group can contain multiple databases.

ASM Disk Groups
Striping

- ASM uses striping to provide optimal I/O performance
- Two types of striping dependent on database file type
  - Coarse striping
    - Stripe size is 1MB
    - All database files except control files, online redo logs and flashback logs
  - Fine striping
    - Used to reduce file latency for files needing faster access
    - Stripe size is 128KB
    - Used for control files, online redo logs and flashback logs

ASM uses striping to provide optimal I/O performance

There are two types of striping dependent on database file type

• Coarse striping - used for all database files except control files, online redo logs and flashback logs. The stripe size is 1MB

• Fine striping - used for control files, online redo logs and flashback logs only. Fine striping is used to reduce file latency for files needing faster access. The stripe size is 128KB.

In order to maximize the benefits of ASM striping, you should use disks of same type and performance capacity in a disk group

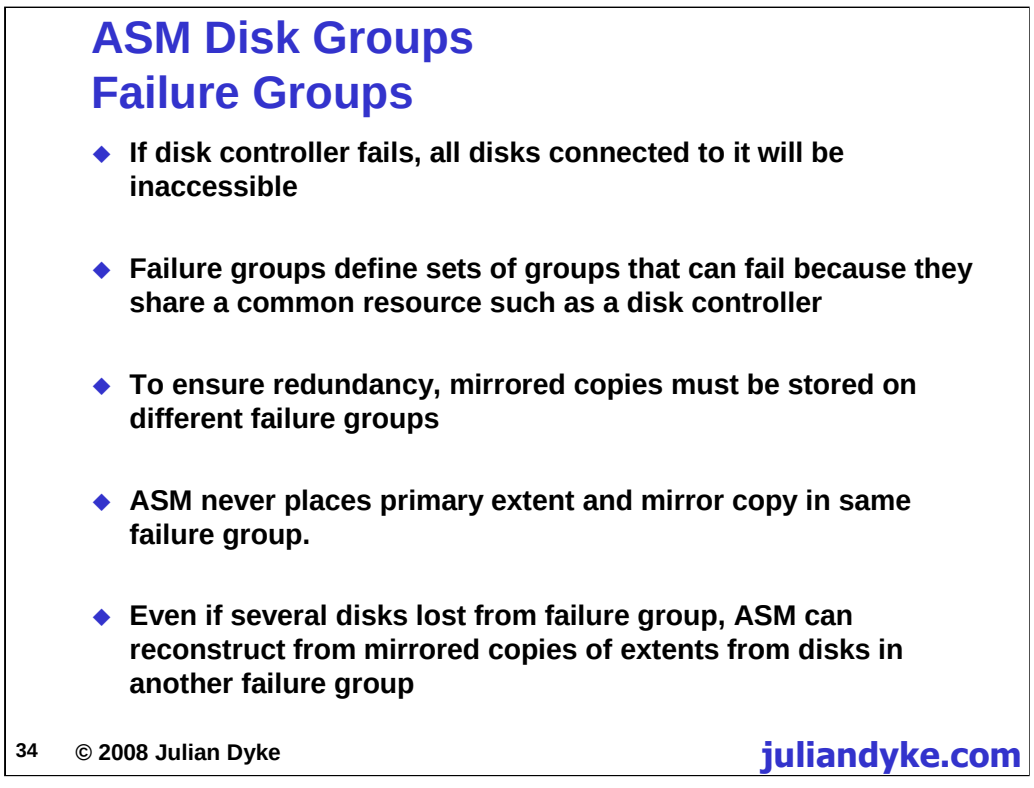
ASM Disk Groups
Mirroring

- Provides data redundancy
- If disk is lost, mirror can continue operations without data loss
- ASM mirroring is different from operating system mirroring
  - Operating system mirrors entire disks
  - ASM mirrors extents
- When any disk in disk group fails, ASM reconstructs failed disk using mirrored extents from other disks in disk group
- Mirroring extents reduces I/O required to construct mirror after disk failure because extents are spread across several disks

Mirroring provides data redundancy. If a disk fails, the mirror can continue operations without data lost.

ASM mirroring is different from operating system mirroring. The operating system mirrors entire disks whereas ASM mirrors extents.

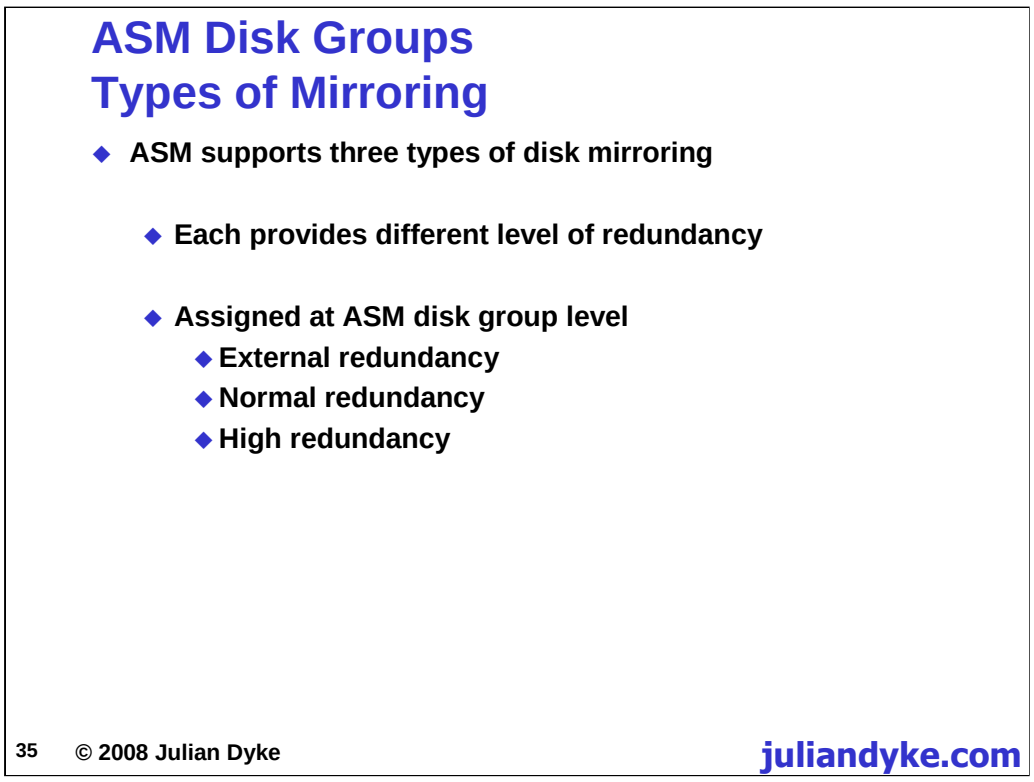
When any disk in disk group fails, ASM reconstructs failed disk using mirrored extents from other disks in disk group. Mirroring extents reduces I/O required to construct mirror after disk failure because extents are spread across several disks



If a disk controller fails, all disk connected to it will be inaccessible.

Failure groups define sets of groups that can fail because they share a common resource such as a disk controller. To ensure redundancy, mirrored copies must be stored on different failure groups

If you have configured normal or high redundancy then ASM never places the primary extent and the mirror copy of that extent in same failure group. Even if several disks lost from failure group, ASM can reconstruct from mirrored copies of extents from disks in another failure group



ASM supports three types of mirroring, each providing a different level of redundancy. A level of redundancy can be assigned to each ASM disk group.

The following redundancy levels are available:

• External redundancy

• Normal redundancy

• High redundancy

ASM Disk Groups
External Redundancy

- External redundancy
  - No mirroring
  - Using operating system storage array protection
  - No failure groups
  - Most commonly deployed redundancy option
  - Recommended for most RAC users including those with SAN and NAS

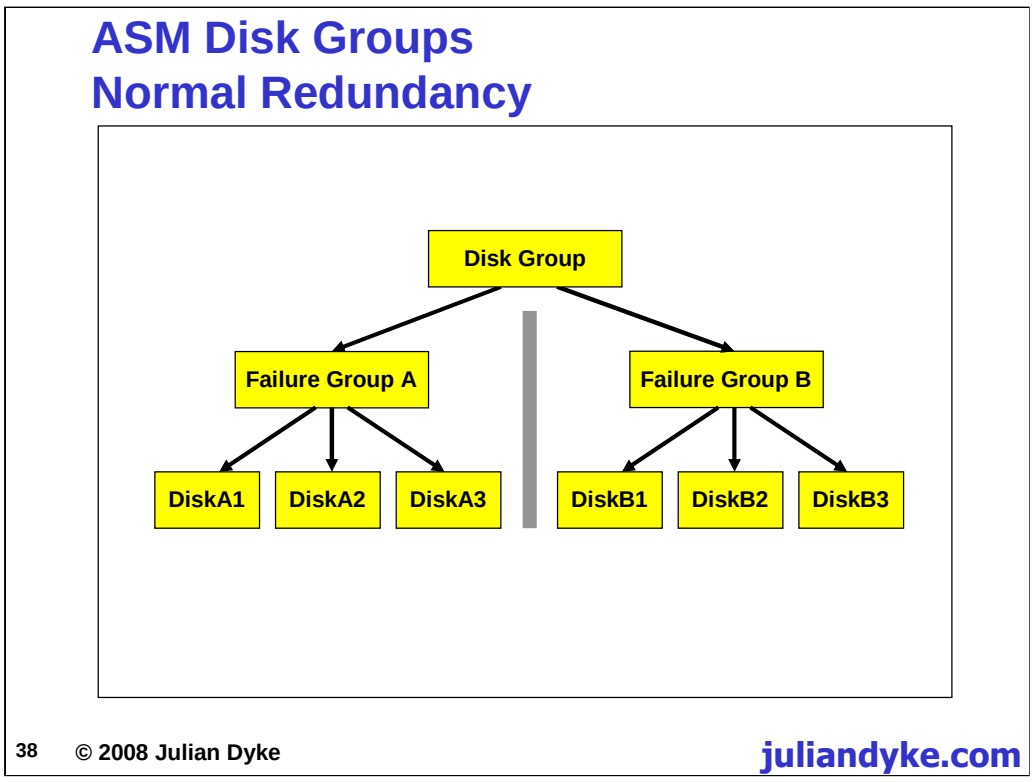
Most RAC users specify External Redundancy for ASM Disk Groups. I recommend this because, by definition a RAC database must have shared storage which is normally either SAN or NAS. These storage types generally support RAID configurations e.g. RAID1+0 or RAID5. In my opinion it is better to rely on the technology implemented by the storage vendor over the past 10-15 years than to rely on the technology released by Oracle in the past 24 months.

ASM Disk Groups
Normal Redundancy

- Normal redundancy
  - Two-way mirroring
  - Must have at least two failure groups
  - Each extent written to two disk groups
  - One primary extent and one mirrored extent
  - Mandatory with host-based mirroring for extended clusters
  - Rarely encountered in the field

Normal redundancy specifies that ASM should maintain one mirrored copy of every extent. You must have at least two failure groups. Each extent will be written to two disk groups. There is one primary extent and one mirrored extent.

Normal redundancy is mandatory if you are implementing host-based mirroring for extended clusters. Otherwise it is rarely encountered in the field as the vast majority of customers use external redundancy.



This slide shows how normal redundancy is configured. Failure groups are a implicit concept; it is not necessary to create them explicitly.

When normal redundancy is configured,. each disk should be allocated to a failure group. The failure groups should exist on physically redundant hardware. This can be JBODs, SAN LUNs or NAS file system provided all physical components in the storage are redundant between the failure groups.

Normal redundancy is used by some major RAC users to provide low cost storage using JBOD arrays. This can be much more cost effective than using SAN or NAS storage for large data volumes.

ASM Disk Groups
High Redundancy

- High redundancy
  - Three-way mirroring
  - Must have at least three failure groups
  - Each extent written to three disk groups
  - One primary extent and two mirrored extents
  - No known deployments at present (December 2006)

Normal redundancy specifies that ASM should maintain two mirrored copies of every extent. You must have at least three failure groups. Each extent will be written to three disk groups. There is one primary extent and two mirrored extents.

At the time of writing (December 2006) I am not aware of any deployments of ASM using high redundancy.

ASM Disk Groups
Disk Group Creation

- ASM disk groups can be created
  - using Enterprise Manager
  - in Database Configuration Assistant (DBCA)
  - in SQL*Plus using CREATE DISKGROUP command

```
CREATE DISKGROUP diskgroup1
NORMAL REDUNDANCY
FAILGROUP controller1 DISK
    '/dev/sda1',
    '/dev/sda2',
FAILGROUP controller2 DISK
    '/dev/sdb2',
    '/dev/sdb2';
```

```
CREATE DISKGROUP diskgroup1
HIGH REDUNDANCY
FAILGROUP controller1 DISK
    '/dev/sda1',
    '/dev/sda2',
FAILGROUP controller2 DISK
    '/dev/sdb1',
    '/dev/sdb2',
FAILGROUP controller3 DISK
    '/dev/sdc1',
    '/dev/sdc2';
```


I recommend using DBCA to create disk groups. However, you can also use SQL*Plus commands as shown in the above slide.

For NORMAL REDUNDANCY two failure groups must be specified using the FAILGROUP clause; for HIGH REDUNDANCY, three failure groups must be defined using the FAILGROUP clause.

If no FAILGROUP is specified then each disk will be in its own failure group. If no NAME is specified then Oracle will assign a system-generated name.

ASM will automatically mount newly created disk group. If you are using an SPFILE for the ASM instance then the name will be added to the SPFILE automatically; if you are using a PFILE then it will be necessary to add the disk group name manually.

To create a disk group using Enterprise Manager. go to the Disk Group Administration page and specify the redundancy level, disk group name and a list of disks

ASM Disk Groups
Managing Disk Groups

- To add a disk to an existing disk group use ALTER DISKGROUP ADD DISK e.g.

```
ALTER DISKGROUP diskgroup1 ADD DISK '/dev/sda3' NAME disk3;
```

- To drop a disk from an existing disk group use ALTER DISKGROUP DROP DISK e.g.

```
ALTER DISKGROUP diskgroup1 DROP DISK disk3;
```

- To cancel the dropping of disks from an existing disk group use ALTER DISKGROUP e.g.

```
ALTER DISKGROUP diskgroup1 UNDROP DISKS;
```

- To drop a disk group use the DROP DISKGROUP command e.g.

```
DROP DISKGROUP diskgroup1 INCLUDING CONTENTS;
```

41 juliandyke.com

Note that you cannot use the UNDROP clause if the FORCE clause was specified in the ALTER DISKSGROUP DROP DISK command e.g.

ALTER DISKGROUP diskgroup1 DROP DISK disk3 FORCE;

ASM Disk Groups
Rebalancing Disk Groups

- ASM rebalances disk groups automatically and dynamically when a disk is added or removed from a disk group
- Disk groups can be rebalanced manually:

```
ALTER DISKGROUP diskgroup1 REBALANCE POWER 5;
```

- POWER clause
  - specifies level of parallelization for REBALANCE command
  - controls speed of rebalance operation
  - overrides ASM_POWER_LIMIT parameter
  - defaults to ASM_POWER_LIMIT if not specified
- In Oracle 10.2 and above, when several disks are added or dropped rebalancing operations can be consolidated into a single operation

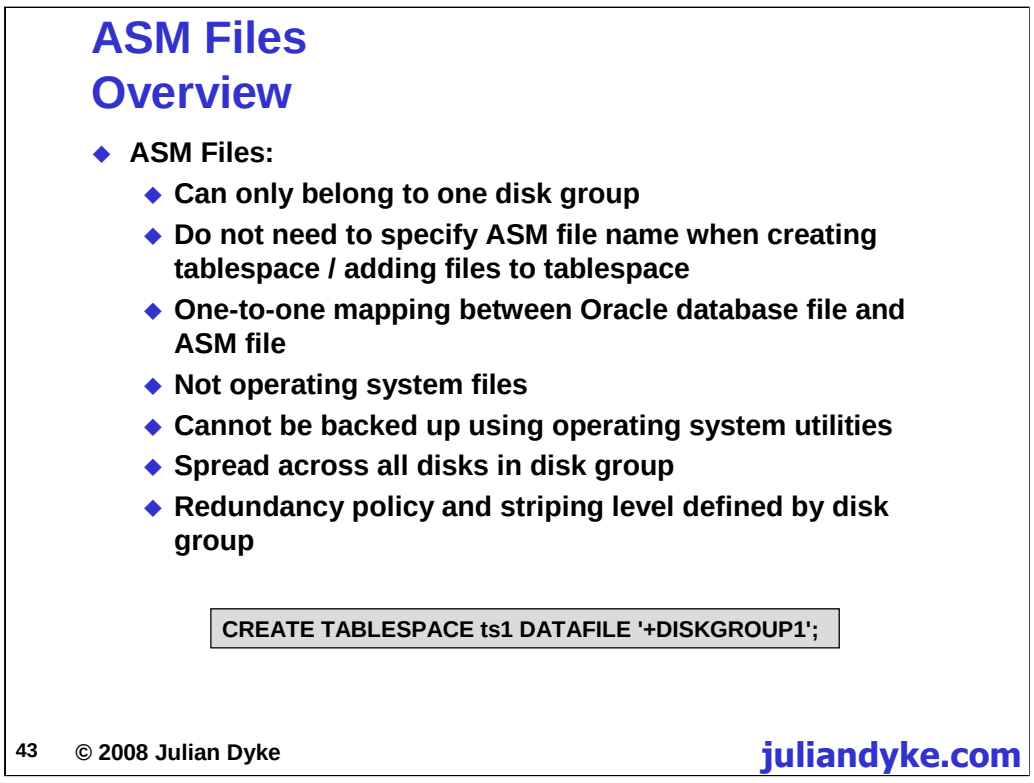
ASM rebalances disk groups automatically and dynamically when a disk is added or removed from a disk group

Disk groups can be also rebalanced manually. For example:

```
ALTER DISKGROUP diskgroup1 REBALANCE POWER 5;
```

The POWER clause specifies level of parallelization for REBALANCE command. It controls the speed of the rebalance operation. The POWER clause overrides the ASM_POWER_LIMIT parameter if specified; otherwise it defaults to the value of the ASM_POWER_LIMIT parameter.

In Oracle 10.2 and above, when several disks are added or dropped rebalancing operations can be consolidated into a single operation. This was a particular weakness in the original Oracle 10.1 implementation as there was no way to stop the rebalance operation starting after a disk group DDL statement was executed. Therefore it was not possible to group together a set of related disk group DDL statements.



ASM files can only belong to one disk group. You do not need to specify a filename for the ASM file when creating a tablespace or adding files to that tablespace. For example:

```
CREATE TABLESPACE ts1 DATAFILE '+DISKGROUP1';
```

There is a one-to-one mapping between Oracle database files and ASM files.

ASM files are not operating system files and cannot be backed up using operating system utilities; the ASM disk group must be backed up using RMAN.

ASM files can be spread across all disks in their disk group. The redundancy policy and striping level defined by disk group.

ASM Files Overview

- ASM files:
  - have OMF filenames
  - filenames are automatically generated
  - will be automatically deleted when tablespace is deleted
  - can also have user-specified alias
    - not considered an OMF file
    - will not be automatically deleted
  - cannot be seen at operating system level
  - can be seen by RMAN and other Oracle tools
  - can be seen in V$DATAFILE and V$LOGFILE
- ASM filenames stored in:
  - Controlfile
  - RMAN recovery catalog

You cannot use ASM file system for

- trace files
- audit files
- alert logs
- backup files
- export files
- tar files
- core files

ASM Files
Naming Conventions

- There are 4 naming conventions
  - Fully qualified ASM filenames
  - Numeric ASM filenames
  - Alias ASM filenames
  - Incomplete ASM filenames
- Oracle automatically generates fully qualified names for all ASM files
- User friendly alias ASM Filenames can be created for system generated filenames

45 © 2008 Julian Dyke juliandyke.com

The ASM filename usage rules are as follows:

For referring to an existing file use:
• fully qualified filename
• numeric filename
• alias filename (without template)

For creating a single file use:
• alias filename (with or without template)
• incomplete filename (with or without template)

For creating multiple files use:
• incomplete filename (with or without template)

Automatic Storage Management Dynamic Performance Views

- **The following dynamic performance views support ASM:**

View Name	Introduced
V$ASM_ALIAS	10.1
V$ASM_CLIENT	10.1
V$ASM_DISK	10.1
V$ASM_DISKGROUP	10.1
V$ASM_DISKGROUP_STAT	10.2
V$ASM_DISK_STAT	10.2
V$ASM_FILE	10.1
V$ASM_OPERATION	10.1
V$ASM_TEMPLATE	10.1

Most of the dynamic performance views were introduced in Oracle 10.1 with the exception of V$ASM_DISKGROUP_STAT and V$ASM_DISK_STAT which were introduced in Oracle 10.2.

The difference between V$ASM_DISKGROUP and V$ASM_DISKGROUP_STAT is that the latter dynamic performance view does not attempt to perform disk discovery when it is queried. The difference between V$ASM_DISK and V$ASM_DISK_STAT is similar.

For each instance-specific V$ view there is an equivalent cluster-wide GV$ view with a similar name e.g. GV$ASM_ALIAS.

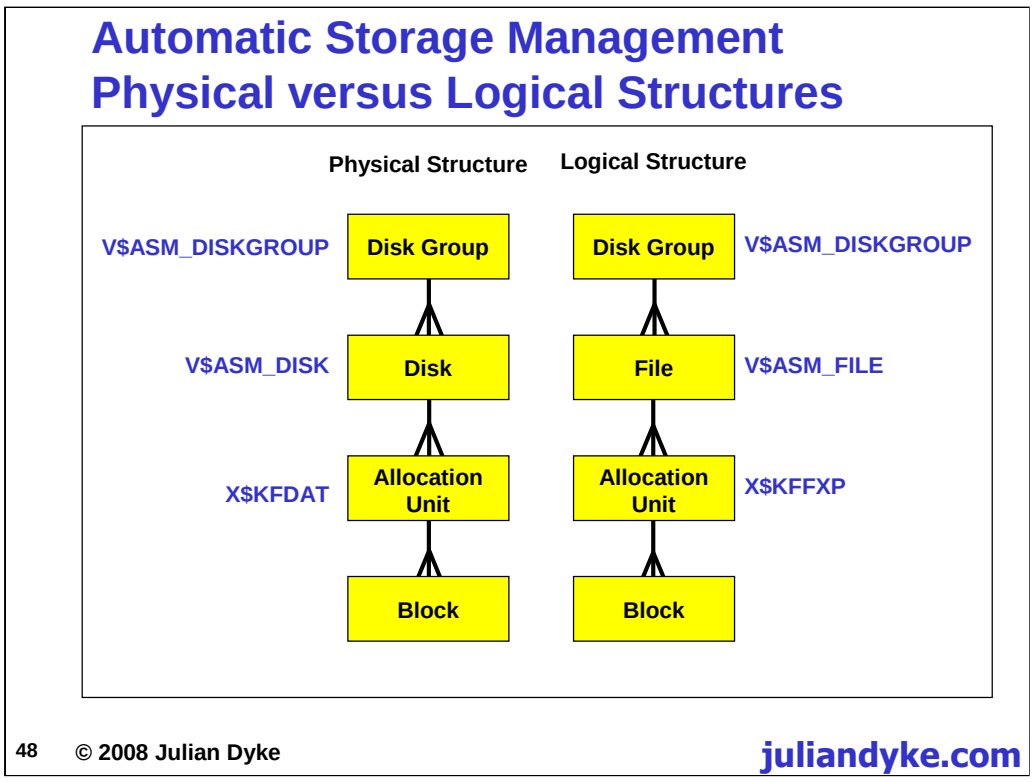
Automatic Storage Management Fixed Views

- **ASM is supported by the following fixed views**

View Name	Purpose
X$KFALS	Aliases
X$KFDSK	Disks
X$KFDSK_STAT	Disks (no discovery)
X$KFDAT	Disk Extents
X$KFGRP	Disk Groups
X$KFGRP_STAT	Disk Groups (no discovery)
X$KFGMG	Operations
X$KFKID	Disk performance
X$KFNCL	Clients
X$KFTMTA	Templates
X$KFFIL	Files
X$KFFXP	File Extents

ASM is supported by a set of fixed views including:

X$KFALS	Aliases
X$KFDSK	Disks
X$KFDSK_STAT	Disks (no discovery)
X$KFDAT	Disk Extents
X$KFGRP	Disk Groups
X$KFGRP_STAT	Disk Groups (no discovery)
X$KFGMG	Operations
X$KFKID	Disk Performance
X$KFNCL	Clients
X$KFTMTA	Templates
X$KFFIL	Files
X$KFFXP	File Extents

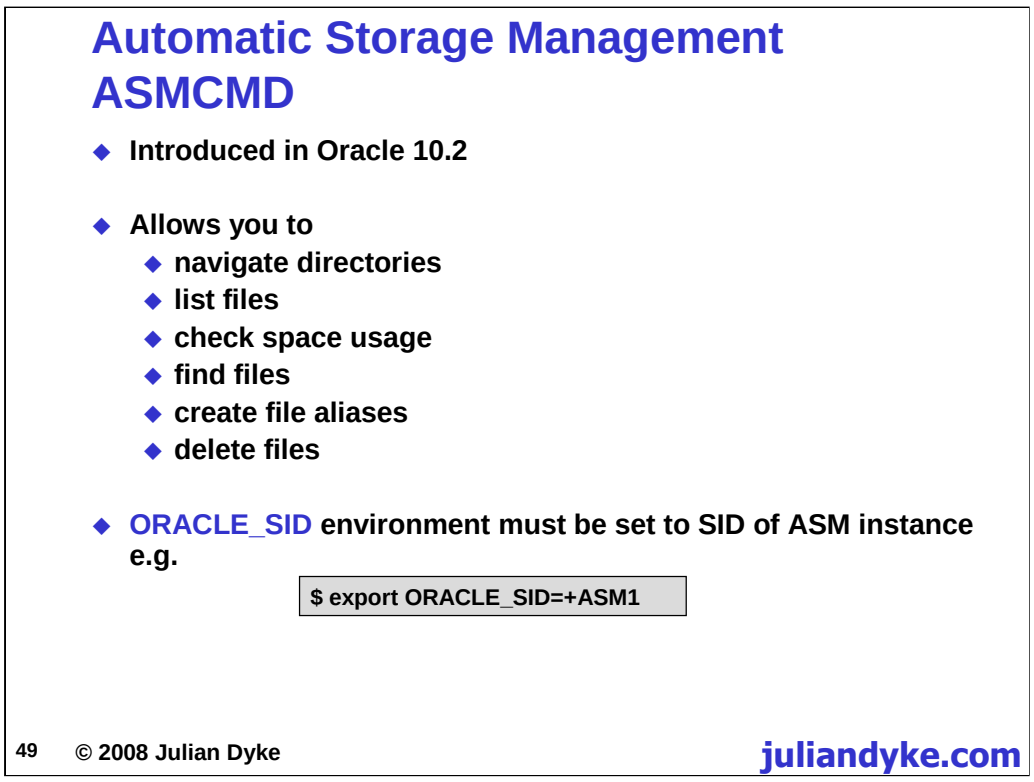


This slide summarizes the relationships between disk groups, disks, files, allocation units and blocks.

Each database can store data in one or more diskgroups. Similarly each diskgroup can contain data for one or more databases.

A disk group contains one or more physical disks. These are reported in V$ASM_DISK. Space within each physical disk is divided into allocation units. These are analogous to extents in a database segment. Allocation units are not reported in any V$ dynamic performance views. They are, however, reported in the X$KFDAT fixed view.

A disk group also contains zero or more files. These are reported in V$ASM_FILE. Each file uses one or more allocation units to store its contents. The allocation units are not necessarily contiguous. Allocation units are not reported in any V$ dynamic performance views. They are, however, reported in the X$KFFXP fixed view.



When ASM was introduced in Oracle 10.1, there was no supported way of viewing or modifying the contents of the file system. In Oracle 10.2 and above, the asmcmd utility allows you to inspect the files within the file system and to make very limited modifications.

In order to use asmcmd, you must first set your ORACLE_SID environment variable to the ASM instance name. For example:

export ORACLE_SID=+ASM1

asmcmd establishes a bequeath connection to the instance specified by $ORACLE_SID. The user must be a member of the SYSDBA group as this privilege is set implicitly by asmcmd during connection to the instance. When the asmcmd utility starts successfully, it will display the ASMCMD> prompt.

Automatic Storage Management ASMCMD

- **ASMCMD has a Unix-like command syntax including**
  - **cd** change directory
  - **du** disk usage
  - **find** find file
  - **help** show help
  - **ls** list files
  - **lsct** list clients
  - **lsdg** list diskgroups
  - **mkalias** make alias
  - **mkdir** make directory
  - **pwd** print working directory
  - **rm** remove file
  - **rmalias** remove alias

The syntax of the asmcmd utility is:

asmcmd [-p] command

If you specify the -p option the current directory will be displayed in the command prompt. For example:

ASMCMD [+DATAFILE/ORCL/CONTROLFILE] >

Automatic Storage Management FTP Support

- In Oracle 10.2 and above files in ASM file systems can be accessed using FTP
  - Requires XML DB installation
  - Created by default during DBCA installation
- Port numbers must be assigned manually. For example:

FTP	7001
HTTP/WebDAV	8001

- Port numbers can be assigned in
  - Enterprise Manager XDB Configuration Page
  - SQL*Plus
- For example, run the following SQL*Plus script as SYSDBA

```
SQL> @$ORACLE_HOME/rdbms/admin/catxdbdbca.sql 7001 8001
```

51 juliandyke.com

In Oracle 10.2 and above files in ASM file systems can be accessed using FTP. In order to use FTP, XML DB must have been installed. XML DB is created as part of the DBCA installation using the General Template.

Port numbers must be assigned manually for FTP and HTTP/WebDAV. For example you might specify port 7001 for FTP and port 8001 for HTTP/WebDAV.

The port numbers can be assigned in the Enterprise Manager XDB Configuration Page.

Alternatively you can assign the port numbers using SQL*Plus. For example:

```
SQL> @$ORACLE_HOME/rdbms/admin/catxdbdbca.sql 7001 8001
```

Automatic Storage Management FTP Support

- All ASM files are stored in /sys/asm
- Use any ftp client. For example:

```
$ ftp -u                                  # -u suppresses Kerberos authentication
ftp> open localhost 7001                  # use FTP port
ftp> user system
331 pass required for SYSTEM
Password: <oracle>
230 SYSTEM logged in
ftp> cd /sys/asm/diskgroup1/test/controlfile
ftp> ls                                   # list files in directory
-rw-r--r--  1 SYS    oracle 7258112 SEP 08 18:30 Current.305.566832105
ftp> bin                                  # switch to binary for file transfers
ftp> get Current.305.566832105    # can also put file
ftp> bye
```


All ASM files are externalized in /sys/asm. You can use any ftp client to access the ASM file system.

The slide shows an example FTP session. Note that the FTP port must be specified when opening the host. For example

ftp> open localhost 7001

You can navigate around the ASM file system as you would with any other directory-based file system.

You can use the pwd command to print the working directory, the cd command to change working directory and the ls command to inspect files.

Finally you can use the get and put commands to extract files from the ASM file system and to load files into the ASM file system respectively.

Note that it is very difficult to envisage any circumstances where it would ever be necessary to load files into an ASM file system from an external source using the FTP put command.

Automatic Storage Management Utilities

- **In Oracle 10.2 and above the kfed utility can be used to inspect and edit the contents of ASM blocks**

```
[oracle@server3 ~]$ $ORACLE_HOME/bin/kfed -h
as/mlib         ASM Library [asmlib='lib']
aun/um          AU number to examine or update [AUNUM=number]
aus/z           Allocation Unit size in bytes [AUSZ=number]
blkn/um         Block number to examine or update [BLKNUM=number]
blks/z          Metadata block size in bytes [BLKSZ=number]
ch/ksum         Update checksum before each write [CHKSUM=YES/NO]
cn/t            Count of AUs to process [CNT=number]
d/ev            ASM device to examine or update [DEV=string]
o/p             KFED operation type
                [OP=READ/WRITE/MERGE/NEW/FORM/FIND/STRUCT]
p/rovnm         Name for provisioning purposes [PROVNM=string]
s/eek           AU number to seek to [SEEK=number]
te/xt           File name for translated block text [TEXT=string]
ty/pe           ASM metadata block type number [TYPE=number]
```

- **This utility should only be used under the guidance of Oracle Support**

Automatic Storage Management Utilities

- In Oracle 10.2 and above the kfod utility can be used to test discovery of ASM instances, disk groups and disks

```
[oracle@server3 ~]$ $ORACLE_HOME/bin/kfod -h
_asm_a/llow_only_raw_disks  KFOD allow only raw devices

[_asm_allow_only_raw_disks=TRUE/(FALSE)]
_asm_l/ibraries             ASM Libraries

[_asm_libraries='lib1','lib2',...]
_asms/id                    ASM Instance[_asmsid=sid]
a/sm_diskstring             ASM Diskstring
                            [asm_diskstring='discoverystring'...]
d/isks                      Disks to discover [disks=raw,asm,all]
g/roup                      Disks in diskgroup [group='diskgroup']
n/ohdr                      KFOD header suppression [nohdr=TRUE/(FALSE)]
o/p                         KFOD options type
                            [OP=DISKS/GROUPS/INSTS/VERSION/CLIENTS/ALL]
p/file                      ASM parameter file [pfile='parameterfile']
s/tatus                     Include disk header status [status=TRUE/(FALSE)]
v/erbose                    KFOD verbose mode [verbose=TRUE/(FALSE)]
```

- This utility should only be used under the guidance of Oracle Support

Automatic Storage Management Oracle 11g New Features

- New ASM features in Oracle 11g include:
  - ASM fast mirror resynchronization
    - Applies to normal and high redundancy only
    - In Oracle 10g after disk access failure disks were automatically dropped and then rebuilt entirely
    - In Oracle 11g only extents modified since disk access failure are rebuilt
  - ASM preferred mirror read
    - Mainly applicable for extended clusters
    - Allows preferred failure group to be specified
  - Variable extent sizes
    - Extent sizes increased automatically
    - Allows larger files to be stored in ASM file system

Automatic Storage Management Oracle 11g New Features

- New ASM features in Oracle 11g include:
  - SYSASM role
    - Optional replacement for SYSDBA role
    - Allows storage administrators to work with ASM without compromising database security
  - New ASM disk group attributes
    - au_size
    - compatible_rdbms
    - compatible_asm
    - disk_repair_time
    - template.<name>.redundancy
    - template.<name>.stripe

Automatic Storage Management Oracle 11g New Features

- New ASM features in Oracle 11g include:
  - Fast Rebalance
    - In RAC clusters, disk group can be mounted on single instance only
    - Rebalance can proceed without locking overhead
  - ASMCMD Extensions
    - cp - copy command
    - md_backup - metadata backup
    - md_restore - metadata restore
    - repair

The ASMCMD cp command allows files to be copied to and from an ASM file system.