



Implementing Oracle 10g RAC with ASM on AIX

R. Ballough
ballough@us.ibm.com
IBM Advanced Technical Support
December 10, 2005



Introduction..... 3

Automatic Storage Management (ASM) Overview 4

Considerations for choosing ASM..... 8

 Storage Requirements and Dependencies 8

 Administration Implications of ASM Implementation 10

Tuning Parameters and Settings for ASM 12

 AIX Parameters..... 12

 Oracle Parameters 13

Conclusion 15

Appendix A: A sample ASM install process using EMC PowerPath (Symmetrix) with
AIX 5.3 16

Reminders 19

Trademarks 19

References..... 19

Acknowledgements..... 20

Introduction

Prior to the release of Oracle Database Server 10g, database administrators had to choose between raw logical volumes or filesystems to house their data files. The decision to implement one or the other was always a compromise, as raw logical volumes offer the best performance, but filesystems are easier to administer.

The choice of filesystems offers many features for the dba. Since the unit of storage allocation is a filesystem, and the filesystem is owned by the Oracle user, the DBA can allocate additional data files as needed. Oracle's autoextend feature can also be used to increase the space allocation for data files when necessary, and free space can be easily seen at the system level to be used with capacity planning tools. When it comes to backups, the data copy can also take place at the system level, backing up only used data.

When raw logical volumes are used, on the other hand, a logical volume must be allocated for each database file. Adding or resizing a logical volume must be performed by the systems administrator. For backups, either RMAN must be used, or the entire raw logical volume must be written to the backup media, including unused sectors.

Although the management issues are more complex, raw logical volumes offer the best performance, though filesystem improvements to bypass the filesystem buffer cache such as AIX's Concurrent I/O (CIO) can offer response time near that of raw in some configurations.

In Oracle 10g, a new choice for data file storage has been introduced, called Automatic Storage Management (ASM). Built on raw device files, ASM offers the performance of raw logical volumes, but offers the configuration flexibility of filesystems to the DBA.

The purpose of this paper is to provide information about the requirements and considerations for ASM implementation for a Real Applications Cluster (RAC) environment using the AIX operating system. Although ASM may also be used in a single-instance environment, this document will only discuss ASM in the context of RAC.

Automatic Storage Management (ASM) Overview

With AIX, each LUN has a raw device file in the /dev directory such as /dev/rhdisk0. Traditionally, in a filesystem or raw environment, a volume group is created using one or more LUNs, and the volume group is subdivided into logical volumes, each of which will have its own raw device file, such as /dev/rlv0. For an ASM environment, this raw device file for either a LUN or a raw logical volume is assigned to the oracle user. An ASM instance manages these device files; in a RAC cluster, one ASM instance will be created per RAC node.

The raw device files given to ASM are organized into *ASM disk groups*. In a typical configuration, two disk groups are created, one for data and one for a recovery area, but if LUNs are of different sizes or performance characteristics, more disk groups should be configured, with only similar LUNs (similar in terms of size, performance, and raid characteristics) comprising each disk group. For each ASM diskgroup, a level of redundancy is defined, which may be *normal* (mirrored), *high* (3 mirrors), or *external* (no mirroring). Each disk group can also have a stripe width specified of 1MB or 128k. External mirroring with a stripe width of 1MB is typical in a fibre channel SAN environment.

When a file is created within ASM, it is automatically striped across all storage allocated to the disk groups, and each of these individual stripes is referred to as a *file extent*. When a database instance opens a file managed by ASM, the database instance queries the ASM instance for the map of extents for that file. Once that extent map is provided to the database instance, the database instance performs I/O to the device file locations indicated by the extent map – that is, the ASM instance provides the database with pointers to the data, and is not actually used to perform I/O. As a result, the ASM parameter file (spfile) does not need to be tuned for load, though SGA and parameter changes may be needed for the database instance(s) using ASM.

To create a data file in an ASM environment, a tablespace is simply created within the ASM instance, and no supporting filesystem or raw logical volume creation is necessary.

```
SQL> create tablespace index01 datafile '+DATADG1' size 1024M;
```

It is possible within ASM to create user-friendly names and directory structures using aliases, however, files created using aliases are not Oracle-managed files. The following statements will create the tablespace index01 in a directory labeled index:

```
SQL> alter diskgroup DATA add directory '+DATADG1/index';  
SQL> create tablespace index01 datafile '+DATADG1/index/index01' size  
1024M;
```



Rather than viewing data file information such as location and free space by executing operating system commands such as 'df', V\$ASM views within the ASM instances are queried.

The V\$ASM_DISKGROUP view stores information about each disk group and can be used to query space available in the disk group.

```
SQL> select group_number, name, state, type, total_mb, free_mb
from v$asm_diskgroup;
```

GROUP_NUMBER	NAME	STATE	TYPE	TOTAL_MB	FREE_MB
1	DATADG1	MOUNTED	EXTERN	168478	137383

The V\$ASM_DISK view breaks this information down by each device assigned to ASM:

```
SQL> select path, state, total_mb, free_mb from v$asm_disk;
```

PATH	STATE	TOTAL_MB	FREE_MB
/dev/rhdisk10	NORMAL	168478	137383

This view also includes performance statistics similar to system iostat commands:

```
SQL> select path, reads, read_time, bytes_read, writes, write_time,
bytes_written from v$asm_disk;
```

PATH	READS	READ_TIME	BYTES_READ	WRITES	WRITE_TIME	BYTES_WRITTEN
/dev/rhdisk10	170	4.42	2682880	1558	7.27	6381568

The V\$ASM_FILE view shows the detail at the file level, but lists it by file number, rather than a file name, so queries to this table should be joined with the V\$ASM_ALIAS view to show the file names:

```
SQL> select a.file_number, a.name, f.block_size, f.blocks, f.bytes,
f.space, f.type from v$asm_alias a, v$asm_file f
2 where a.file_number=f.file_number;
```

FILE_NUMBER	NAME	BLOCK_SIZE	BLOCKS	BYTES	SPACE	TYPE
256	Current.256.574691443	16384	431	7061504	8388608	CONTROLFILE
257	Current.257.574691443	16384	431	7061504	8388608	CONTROLFILE
261	UNDOTBS1.261.574696463	8192	1048577	8589942784	8592031744	DATAFILE
262	SYSTEM.262.574696449	8192	131073	1073750016	1075838976	DATAFILE



Starting with release 10.2.0.1, Oracle also provides a utility called ASMCMD to provide a unix-like command line environment for querying information about ASM storage. The following commands are available: *cd, du, find, help, ls, lsct, lsdg, mkalias, mkdir, pwd, rm, and rmalias.*

```
$ export ORACLE_SID=+ASM1
$ asmcmd
ASMCMD> cd +DATADG1/TEST/datafile
ASMCMD> ls -al
Type          Redund Striped Time          Sys Name
DATAFILE UNPROT COARSE NOV 28 21:00:00 Y none=>SOE.264.575538945
DATAFILE UNPROT COARSE NOV 28 21:00:00 Y none=>SYSAUX.270.574696517
DATAFILE UNPROT COARSE NOV 28 21:00:00 Y none=>SYSTEM.262.574696449
DATAFILE UNPROT COARSE NOV 28 21:00:00 Y none=>UNDOTBS1.261.574696463
DATAFILE UNPROT COARSE NOV 28 21:00:00 Y none=>USERS.266.574696529
ASMCMD> du
Used_MB      Mirror_used_MB
10429        10429
ASMCMD> lsdg
State  Type  Rebal  Unbal  Sector  Block  AU  Total_MB
Free_MB Req_mir_free_MB Usable_file_MB Offline_disks Name
MOUNTED EXTERN N      N      512    4096  1048576 168478
137383          0      137383          0 DATADG1/
```

Databases using ASM can be seen either using the ASMCMD command 'lsct', or by querying the V\$ASM_CLIENT view from the ASM instance:

```
ASMCMD> lsct
DB_Name  Status          Software_Version  Compatible_version
Instance_Name
test     CONNECTED      10.2.0.1.0       10.2.0.1.0 test1

SQL> select * from v$asm_client;

GROUP_NUMBER INSTANCE_NAME DB_NAME STATUS SOFTWARE_VERSION COMPATIBLE_VERSION
-----
1          test1 test CONNECTED 10.2.0.1.0 10.2.0.1.0
```

If a drive is added to or removed from an ASM diskgroup, ASM will automatically redistribute file extents evenly across the disks in the diskgroup. When the change occurs, the rebalance process (RBAL) detects the change, and messages one or more ARBx processes to perform the actual rebalance. The number of ARBx processes is determined from the parameter ASM_POWER_LIMIT. By default, this parameter is set to 1, making rebalancing a low impact operation. For some environments, even one rebalancing process may not be desirable behavior during peak workloads. If ASM_POWER_LIMIT is set to 0, no rebalancing will occur when a disk is added. If this setting is chosen, care must be taken to manually increase ASM_POWER_LIMIT when rebalancing is deemed acceptable.



Any rebalance activity can be seen by querying the V\$ASM_OPERATION view.

```
SQL> select * from V$ASM_OPERATION;  
no rows selected
```

Considerations for choosing ASM

With Oracle 10g RAC, Oracle introduced its own clusterware, Cluster-Ready Services (CRS), which replaces the heartbeat and cluster membership functions normally provided by HACMP. HACMP is not necessary in a 10g RAC implementation unless the use of traditional raw logical volumes is desired, requiring HACMP to manage the concurrent LVM environment. If a clustered filesystem is desired, GPFS or Veritas software must be purchased. However, in an ASM environment, no additional software is needed, which can result in cost savings over filesystem or raw environments.

In addition to software requirements for ASM, customers should examine both storage requirements and dependencies, and administration implications of ASM implementation.

Storage Requirements and Dependencies

Typically with products and features of the database, Oracle publishes certification information at <http://metalink.oracle.com>. With ASM, however, Oracle has not provided information specific to ASM requirements.

There are several storage requirements which must be met in order to utilize ASM in a RAC environment. First, as in any RAC environment, the device must obviously support physical connectivity to multiple systems – such as dual-initiated SCSI, fibre channel, or network attached (this paper will only address fibre channel attachment). Second, the raw logical device to be used by ASM must support ownership by the ‘oracle’ user rather than the default of ‘root’. And lastly, the device must support the capability to remove any type of locking mechanism that would prevent true sharing by multiple hosts.

Typically in a fibre channel environment, multipathing software is used to facilitate use of multiple fibre channel paths to disk, either for redundancy or bandwidth. It is important to note that, according to the document Oracle Database 10g Release 2 Automatic Storage Management Overview and Technical Best Practices¹, “Oracle Corporation does not certify or qualify these multipathing tools”. In my experience, this function has been performed by the disk vendor, who typically asks for an RPQ to support an ASM configuration.

Below is a list of storage subsystems known to have been implemented with ASM on AIX. The information provided is not intended to be a comprehensive list of all storage

¹ Oracle Database 10g Release 2 Automatic Storage Management Overview and Technical Best Practices by Nitin Vergurlekar, Oracle Corporation. A link to this document is included in the References section of this paper.

products which may be used with AIX, nor is it a substitute for vendor install guides and recommendations, but should instead be used as a starting point for obtaining information to architect an appropriate storage solution for an ASM environment.

IBM eServer storage products:

DS8000, DS6800 and ESS storage models have been successfully implemented using AIX MPIO for multipathing. Using SDDPCM is highly recommended. SDD multipathing software does not allow non-root ownership of devices, and may not be used at this time.

- In order to turn off device locking, all disk devices used for ASM, OCR, or voting must set `reserve_policy=no_reserve`
`# chdev -l hdisk# -a reserve_policy=no_reserve`
- Verify that this is set correctly on all devices:
`# lsattr -El hdisk# | grep reserve`

<code>reserve_policy</code>	<code>no_reserve</code>	Reserve Policy	True
-----------------------------	-------------------------	----------------	------

DS4000 storage models use RDAC for path failover capability, and this is also known to work with ASM, but the parameter in this case to turn off device locking is `reserve_lock=no`

- `# chdev -l hdisk# -a reserve_lock=no`
- Verify that this is set correctly on all devices:
`# lsattr -El hdisk# | grep reserve`

<code>reserve_lock</code>	<code>no</code>	Reserve device on open	True
---------------------------	-----------------	------------------------	------

EMC Storage products:

Symmetrix and Clariion systems using PowerPath version 4.3 and above are known to work with ASM, and installation instructions can be found in the [EMC Networked Storage Topology Guide](#).

- In order to turn off device locking, all `hdiskpower` devices used for ASM, OCR, or voting must set `reserve_lock=no`
- `# chdev -l hdiskpower# -a reserve_lock=no`
- Verify that this is set correctly on all devices:
`# lsattr -El hdiskpower# | grep reserve`

<code>reserve_lock</code>	<code>no</code>	Reserve device on open	True
---------------------------	-----------------	------------------------	------

HDS 99xx Storage:

Hitachi storage can be implemented with ASM in two different ways. The first option is to use native AIX MPIO, in which case the implementation is built on raw devices, as with ESS storage. As with MPIO used in conjunction with ESS, the parameter `reserve_policy=no_reserve` must be set on all devices (see instructions for DS8000, DS6800 and ESS above).

The second option is to use HDLM, version 5.4.2 or above, which requires creation and use of raw logical volumes instead of LUNs for ASM disks. One non-concurrent raw logical volume should be created per LUN at the size of the entire LUN. Creating the logical volumes using the 'mklv -T 0' flag is recommended.

Turning off reserve locking in this environment is performed using the HDLM command

```
# dlkmgr set -rsv on 0 -s
```

Additional installation instructions can be found in the document [Hitachi Dynamic Link Manager for AIX Distribution Set Release Notes](#).

Administration Implications of ASM Implementation

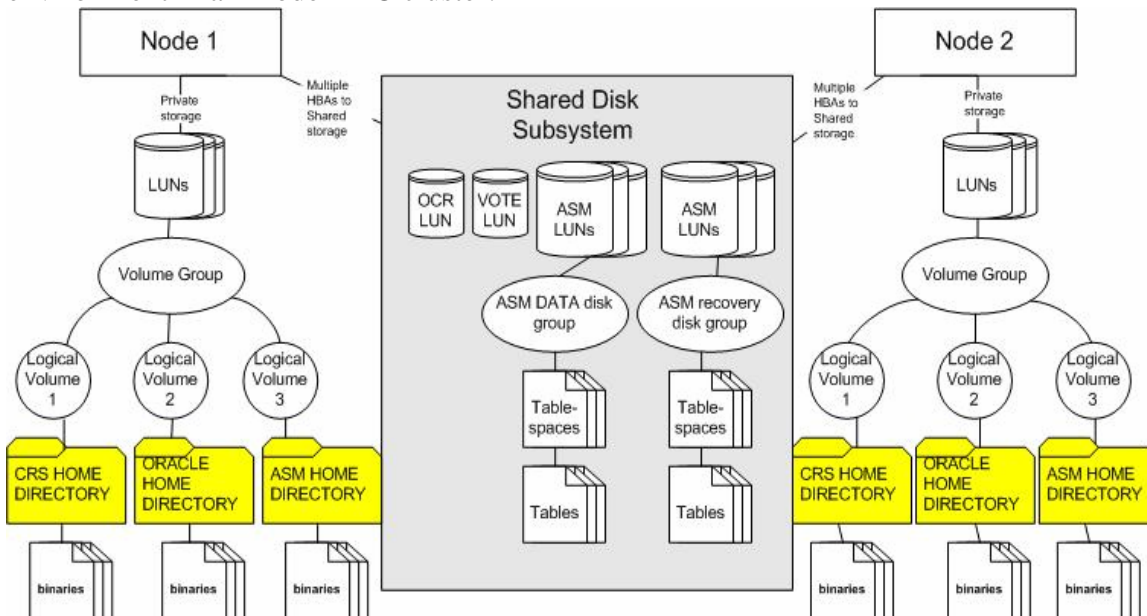
As ASM is built on raw devices, it is essential that systems administration processes and procedures be able to work with a raw environment. This includes using RMAN for backups, and the ability to change any scripts or processes involving use of filesystem commands (such as utilization scripts using the 'df' command). Any procedures for moving data between various hosts will need to include rebuilding the ASM environment on the target host, and any file transfer will no longer be able to be done at the os level (for example, with a 'cp', 'mv' or 'rcp' command). Disk-based copy technologies such as PPRC or SRDF can still be used with ASM, but the secondary copy must be connected to a second node to be used; it cannot be used concurrently with the primary copy on the primary server.

Moving files between locations, such as from one ASM diskgroup to another (required if the level of redundancy needs to be changed, for example) requires using the DBMS_FILE_TRANSFER utility or RMAN restore. In addition to copying ASM files to other ASM files, DBMS_FILE_TRANSFER can also copy ASM files to OS files (and vice versa) to migrate data to or from ASM.

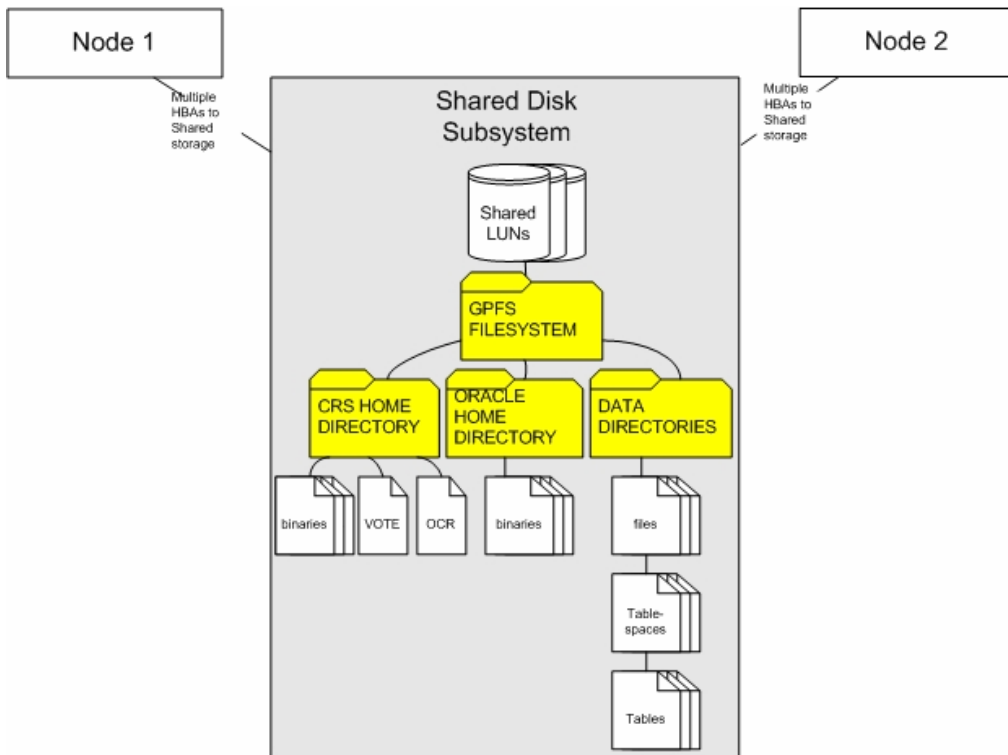
Any time a device is deleted using 'rmdev -dl <hdisk#>' and redetected using cfgmgr, the device ownership will revert to the default, resulting in disks which will not be useable by ASM. Raw device file ownership in /dev should be noted in case these permissions need to be set in the future.

ASM manages only storage for database files; it cannot manage storage for Oracle binaries, or for OCR and VOTE devices. If shared binaries are desired, a clustered filesystem such as GPFS must be used. Full LUNs should be configured for each of the OCR and VOTE devices; this is not a requirement in an HACMP or GPFS environment.

The following diagram represents a typical storage configuration for an ASM environment in a 2-node RAC cluster:



An environment using a clustered filesystem would not need to separate LUNs or even logical volumes for the OCR or Vote functions; in fact, the entire logical volume layer for GPFS is hidden and does not need to be defined. A GPFS clustered environment could have a structure similar to the one below:



Tuning Parameters and Settings for ASM

AIX Parameters

The best practices for AIX system setup for an ASM environment are virtually the same as when using raw logical volumes. The following initial settings are recommended for the operating system, although environments do vary and additional tuning may be required.

1. System attributes (found in 'smit system')

The system maintains a parameter, `maxuproc`, to identify the maximum number of user processes. This parameter should be increased from the default to a number which will exceed the number of oracle processes; typically a value of 4096 is sufficient.

2. Asynchronous I/O

AIX can use two types of asynchronous I/O, kernelized and threaded. All filesystem implementations used threaded asynchronous I/O, and require the configuration of the threaded asynchronous I/O subsystem through 'smit aio', typically increasing the number of I/O servers (maximum # of servers) and the size of the request queue (maximum # of requests). ASM, however, like raw, uses the kernelized asynchronous I/O subsystem, which does not require configuration, although in order to install the Oracle software, the 'STATE to be configured at system restart' for asynchronous I/O threads must be set to 'available' from the 'smit aio' menu.

For Oracle to take advantage of asynchronous I/O, the spfile must include the parameters 'disk_asynch_io=TRUE and filesystemio_options=asynch'.

3. /etc/security/limits

This file sets process resource limits for users. Historically, gating resources with this file has caused issues including inability to start CRS, inability to open database files, and inability to create files greater than 2GB in size. It is highly recommended to set all values other than core to unlimited (-1) for the oracle user.

4. Network Tunables (seen with `no -a`)

The size of the tcp send buffer (`tcp_sendspace`) dictates how much data the application can send, and the tcp receive buffer (`tcp_recvspace`) indicates how much data can be received before being read by the receiving application. Values which are too small risk having application waits due to insufficient buffer space for tcp requests. The default values for AIX are too small for most two or three tier application environments. At a minimum, the following values are suggested:

- `tcp_sendspace = 65536`

- `tcp_recvspace = 65536`

These values may be insufficient for some environments, particularly where network backups are implemented. In order to allow any `sendspace` or `recvspace` buffers to be greater than 65536, the parameter 'rfc1323=1' must also be set.

When UDP over gigabit Ethernet is used as the cluster interconnect, the following parameters should also be set, as indicated in the Oracle Database Administrator's Reference for UNIX Systems:

- `udp_sendspace = db_block_size * db_file_multiblock_read_count + 4k`

- `udp_recvspace = 10*(udp_sendspace)`

- `rfc1323 = 1`

5. Virtual memory manager parameters (seen with `vmo -a`)

By default, AIX allocates up to 80% of real memory for use for filesystem buffer cache.

When a filesystem is not used, such as with ASM, or if filesystem buffer cache is bypassed with options such as concurrent or direct I/O, it is advisable to decrease the amount of memory which can be allocated to filesystem buffer cache. The value `minperm%` dictates the minimum percentage of real memory that is allocated to filesystem buffer cache, `maxperm%` dictates the maximum percentage of real memory that can be allocated to filesystem buffer cache for JFS filesystems, and `maxclient%` dictates the maximum percentage of real memory that can be allocated to filesystem buffer cache for JFS2 filesystems. We typically set these values as follows in an ASM or raw logical volume environment:

- `minperm% = 5`
- `maxclient% = 15`
- `maxperm% = 15`

With AIX5.2ML4 and above, we instead typically use the following values:

- `lru_file_repage=0`
- `minperm% = 5`
- `maxclient% = 80`
- `maxperm% = 80`

The `lru_file_repage` parameter was introduced in AIX5.2ML4, and changes the behavior of the least recently used daemon (`lrud`) such that as long as the number of file pages in use (`numperm` or `numclient`) is greater than `minperm`, only file pages, not computational pages, will be paged out. When `lru_file_repage=0`, the values for `maxperm%` and `maxclient%` can be left at the default value of 80%, which allows any available memory up to 80% which is not in use to be used for filesystem buffer cache.

6. **Paging space (smit chps)**

The default allocation of paging space on AIX is 512MB, which is too small to support Oracle installations. Typically a reasonable rule of thumb is 1x real memory, though larger memory configurations will not need this much memory.

Oracle Parameters

The following parameters are suggested in the document [Oracle Database 10g Release 2 Automatic Storage Management Overview and Technical Best Practices](#):

Parameters to be included in the spfile for databases using ASM:

1. Increase **Processes by 16**
2. Increase **Large_Pool by 600k**
3. Increase **Shared Pool by (1M per 100GB of usable space) + 2M** (assumes external redundancy is chosen)

Parameters to be included in the spfile of the ASM instances:

1. **ASM_POWER_LIMIT=1**

The Best Practices guide suggests this value to make ASM rebalance operations a low priority; however, this can potentially cause rebalance operations to occur during peak volume times. This parameter can also be set to 0 to prevent accidental rebalance



operations from occurring during peak loads, which may be preferable for many environments. This parameter can be raised specifically when rebalancing is desired.

2. **Processes=25 + 15n**, where “n” is the number of databases which will use ASM.



Conclusion

ASM is a viable option for data storage in an Oracle RAC environment on AIX. Its benefits include decreased software cost, raw I/O performance, and ease of database administration. However, systems management practices and storage level implications should be carefully examined before making a decision to implement ASM.

Appendix A: A sample ASM install process using EMC PowerPath (Symmetrix) with AIX 5.3

Basic System Setup

1. Install AIX 5.3 + latest maintenance level, and check metalink note 282036.1 for any additional system prerequisites for Oracle
2. Verify the following filesets are installed, or install if not present:
 - bos.adt.base
 - bos.adt.lib
 - bos.adt.libm
 - bos.adt.syscall
 - bos.perf.libperfstat
 - bos.perf.perfstat
 - bos.perf.proctools
 - bos.perf.gtools
 - rsct.basic
 - rsct.basic.compat
3. Create dba and oinstall groups with the same GID across all cluster nodes
4. Create oracle user with the same UID across all cluster nodes, primary group dba
5. set date and timezone (smit system)
6. start xntpd (smit xntpd)
7. implement tuning parameters from the Tuning Parameters and Settings for ASM section of this document

Configure Network Settings & Services

1. Set up tcpip on the en0 adapter


```
# smitty tcpip
- Minimum configuration and startup for en0  ** public network **
  - rac1:          10.1.1.101
  - rac2:          10.1.1.102
  - rac3:          10.1.1.103
- Minimum configuration and startup for en1  ** RAC Interconnect **
  - rac1-en1:     10.1.10.101
  - rac2-en1:     10.1.10.102
  - rac3-en1:     10.1.10.103
- Minimum configuration and startup for en2  ** RAC Interconnect **
  - rac1-en2:     10.1.11.101
  - rac2-en2:     10.1.11.102
  - rac3-en2:     10.1.11.103
```
2. Update /etc/hosts with all IP/DNS entries
3. Create entries in /etc/hosts.equiv for the oracle user


```
rac1 oracle
rac2 oracle
rac3 oracle
rac1-en1 oracle
rac2-en1 oracle
rac3-en1 oracle
rac1-en2 oracle
```

```
rac2-en2    oracle
rac3-en2    oracle
```

Logical Volumes & Filesystems

1. Increase filesystem sizes:
 - / = 256 MB
 - /tmp = > 500 MB free
 - /var = 512 MB

2. Make filesystems for Oracle SW (\$ORACLE_HOME), ASM (\$ORACLE_ASM_HOME) and CRS (\$ORA_CRG_HOME),
 - \$ORACLE_HOME, eg /opt/oracle/product/10.1.0, should be ~ 5-6 GB
 - \$ORA_CRG_HOME, eg /crs/oracle/product/10.1.0, should be ~ 2 GB
 - mount filesystems after creation
 - change ownerships & permissions, example:
 - chown -R oracle:oinstall /opt/oracle
 - chown -R 775 /opt/oracle
 - mkdir -p /crs/oracle/product/10.1.0
 - chown -R oracle:oinstall /crs/oracle
 - chmod -R 755 /crs/oracle

3. Add \$ORA_CRG_HOME/bin to root's PATH

POWERPATH installation

See *PowerPath for AIX version 4.3 Installation & Administration Guide, P/N 300-001-683 for details*

1. Install EMC ODM drivers
 - 5.2.0.1 from ftp://ftp.emc.com/pub/elab/aix/ODM_DEFINITIONS/EMC.AIX.5.2.0.1.tar.Z
 - install using smit install
2. remove any existing devices attached to the EMC
 - # rmdev -dl hdiskX
3. run /usr/lpp/EMC/Symmetrix/bin/emc_cfgmgr to detect devices
4. Install PowerPath version 4.3.0 using smit install
5. register PowerPath
 - # emcprep -install
6. initialize PowerPath devices
 - # powermt config
7. verify that all PowerPath devices are named consistently across all cluster nodes
 - # /usr/lpp/EMC/Symmetrix/bin/inq.aix64 | grep hdiskpower
 - compare results. Consistent naming is not required for ASM devices, but LUNs used for the OCR and VOTE functions must have the same device names on all rac systems.
 - if the names for the OCR and VOTE devices are different, create a new device for each of these functions as follows:
 - # mknod /dev/ocr c <major # of OCR LUN> <minor # of OCR LUN>
 - # mknod /dev/vote c <major # of VOTE LUN> <minor # of VOTE LUN>
 Major and minor numbers can be seen using the command 'ls -al /dev/hdiskpower*'
8. On all hdiskpower devices to be used by Oracle for ASM, voting, or the OCR, the reserve_lock attribute must be set to "no"
 - # chdev -l hdiskpowerX -a reserve_lock=no

9. Verify the attribute is set
lsattr -El hdiskpowerX
 10. Identify two small luns to be used for OCR and voting
 11. Set permissions on all hdiskpower drives to be used for ASM, voting, or the OCR as follows:
chown oracle:dba /dev/rhdiskpowerX
chmod 660 /dev/rhdiskpowerX
- The Oracle Installer will change these permissions and ownership as necessary during the CRS install process.

Oracle 10g RAC installation

1. Add the following to the oracle user's .profile:
ORACLE_BASE=<oracle base directory>; export ORACLE_BASE
ORA_CRS_HOME=<ora crs home>; export ORA_CRS_HOME
AIXTHREAD_SCOPE=S; export AIXTHREAD_SCOPE
umask 022
2. Run the Oracle installer to install CRS
\$ export LOG=/tmp/orainstall.log
\$ export ORACLE_HOME=/crs/oracle/product/10.1.0
Load the CRS install cd
run rootpre.sh on ALL nodes
\$ runInstaller -ignoreSysPrereqs
3. Check crs install for the correct number of nodes and interfaces
[rac1]crs/oracle/product/10.1.0/bin> # ./olsnodes -n
rac1 1
rac2 2
rac3 3
[rac1]crs/oracle/product/10.1.0/bin> # ./oifcfg getif
en0 10.1.3.0 global public
en1 10.1.30.0 global cluster_interconnect
en2 10.1.31.0 global cluster_interconnect
4. Install Oracle Binaries
\$export ORACLE_HOME=/home/oracle/product/10.1.0
\$ cd <10g DVD directory, Disk1>
\$ runInstaller -ignoreSysPrereqs
5. Install latest 10g patchset
6. Install additional Oracle patches listed in the PowerPath for AIX installation guide.
7. For Clariion systems, refer to the "Requirements for Oracle 10g RAC with ASM on AIX 5L" document from EMC to set the miscount settings appropriately.
8. Run DBCA to set up ASM instances and create database
Create Data and Recovery disk groups, each with external redundancy
use /dev/rhdisk* as the disk discovery path
Choose option to Use Oracle-Managed Files



Reminders

Copyright 2005 IBM Corporation. All Rights Reserved.

Neither this documentation nor any part of it may be copied or reproduced in any form or by any means or translated into another language, without the prior consent of the IBM Corporation. The information in this paper is provided by IBM on an "AS IS" basis. IBM makes no warranties or representations with respect to the content hereof and specifically disclaim any implied warranties of merchantability or fitness for any particular purpose. IBM assumes no responsibility for any errors that may appear in this document. The information contained in this document is subject to change without any notice. IBM reserves the right to make any such changes without obligation to notify any person of such revision or changes. IBM makes no commitment to keep the information contained herein up to date.

Version 1.0, published January 9, 2006

Trademarks

- IBM, AIX, and pSeries are trademarks or registered trademarks of the International Business Machines Corporation.
- Oracle, Oracle8, Oracle8i, Oracle9i, Oracle10g are trademarks or registered trademarks of Oracle Corporation.
- EMC, PowerPath are trademarks or registered trademarks of EMC.
- HDS, HDLM are trademarks or registered trademarks of Hitachi.
- UNIX is a registered trademark in the United States and other countries exclusively through X/Open Company Limited.
- All other products or company names are used for identification purposes only, and may be trademarks of their respective owners.

References

Oracle Database Release Notes 10g Release 1 (10.1.0.2.0) for AIX-Based Systems

<http://www.oracle.com/technology/documentation/database10g.html>

Oracle Database Installation Guide 10g Release 1 (10.1) for UNIX Systems

http://download-west.oracle.com/docs/pdf/B10811_05.pdf

Oracle Database Administrator's Reference 10g Release 1 (10.1) for UNIX Systems

http://download-west.oracle.com/docs/pdf/B10812_06.pdf

Oracle Database 10g Release 2 Automatic Storage Management Overview and Technical Best Practices

http://www.oracle.com/technology/products/database/asm/pdf/asm_10gr2_bptwp_sept05.pdf

Using Oracle 10g's Automatic Storage Management with EMC Storage Technology

http://www.emc.com/techlib/H1144-Best_Prac_Oracle_10gs_EMC_ldv.pdf

© 2005, IBM Advanced Technical Support Techdocs

Version 1/9/2006

<http://w3.ibm.com/support/Techdocs>



Metalink note #293819.1 “Placement of voting and OCR disk files in 10gRAC”
PowerPath for AIX version 4.3 Installation & Administration Guide, P/N 300-001-683

Acknowledgements

Thanks to the following people for their assistance with details for this paper:

- Phil Heintz, Software Solutions Manager, Hitachi Data Systems
- Gregory Verstraeten, IBM Advanced Technical Support
- Dale Martin, IBM Advanced Technical Support