Deploying Oracle ASM with Oracle 10g RAC on AIX with IBM System Storage DS8000 and DS6000 Advanced Features

Betty Mason, IBM Systems and
Technology Group, Open Systems Lab
San Jose, California

Deploying Oracle ASM with Oracle 10g RAC on AIX® with IBM System Storage DS8000 and DS6000 Advanced Features

This document provides an overview of the IBM System Storage DS8000 and DS6000 series disk storage systems, along with a software architecture configuration of Oracle RAC with ASM on a multi-node cluster that utilizes the AIX operating system. Advanced duplication functions covered within this document include: disaster recovery, data migration, and data duplication. Use cases are covered explaining cold backup of a database using the FlashCopy® of Oracle 10g RAC database with Automatic Storage Management (ASM). Other functions covered are Oracle Hot backup Mode and Metro Copy (remote data mirror technology) with either the DS8000 or DS6000 storage system.

August 2007

Disclaimer

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing

IBM Corporation

North Castle Drive

Armonk, NY U.S.A. 10504-1785

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to: IBM World Trade Asia Corporation Licensing 2-31 Roppongi 3-chome, Minato-ku Tokyo 106-0032, Japan.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT,

MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

Trademarks

The following terms are trademarks of International Business Machines Corporation in the United States, other countries, or both:

IBM®

AIX®

DS8000™

DS6000™

eServer™

Enterprise Storage Server®

FlashCopy®

pSeries®

System p™

System Storage™

Oracle is trademark of Oracle Corporation in the United States, other countries, or both. Other company, product, or service names may be trademarks or service marks of others.

Table of Contents

Executive Summary

The IBM System Storage DS8000 and DS6000 Disk Storage systems are both high performance storage systems that offer expandable data open storage and data protective systems to correspond with specific business requirements.

Advanced Copy Services features available on the DS8000 include FlashCopy®, Remote Mirror and Copy, Metro Mirror, Global Copy, and Global Mirror.

A Subsystem Device Drive Path Control Module (SDDPCM) is available for IBM Storage DS8000 and DS6000 in multipath IO (MPIO) environments running on an IBM System p server.

Oracle 10g R2/Clusterware technology offers several features that work with both single database and RAC environments. Oracle Automatic Storage Management (ASM) technology is a feature that is designed to distribute information uniformly across all storage disk groups and has unique mirroring capabilities.

Usage cases of the Oracle 10g RAC data with ASM are applied to both cold backups and hot backups of Oracle Databases on ASM and DS8000 or DS6000.

Introduction

The IBM System Storage DS8000 and DS6000 enable enterprise customers to deploy Oracle Real Application Clusters (RAC) for their critical business needs. The DS8000 and DS6000 combined with RAC are designed to deliver robust, flexible, highly available, and cost-effective disk storage to support continuous operations for mission-critical transactions.

The architecture described in this paper documents the interoperability of AIX, IBM multi-path software SDDPCM, IBM System Storage DS8000 and DS6000 combined with Oracle10g R2 RAC features, Oracle Automatic Storage Management (ASM), and Oracle Clusterware. Oracle ASM was introduced in Oracle 10g and provides an integrated file system and volume manager for Oracle database files.

This paper documents a configuration of Oracle RAC with ASM on a multi-node cluster utilizing the AIX operating system and IBM System Storage DS8000 and DS6000.

IBM System Storage DS8000 and DS6000

The DS8000 series is a member of the IBM System Storage Disk Systems family and is a series of disk storage systems designed to provide high capacity and performance using IBM's POWER5™ technology. The DS8000 Series is a flexible and extendable disk storage subsystem designed to add and adapt to new technologies as they become available. New management tools provided with the DS8000 Series include the DS Storage Manager and the DS Command-Line Interface (CLI), and allow for the management and configuration of the DS8000 series as well as the DS6000 series. The DS8000 Series supports high availability environments and provide remote mirror and copy functions to support business continuity. The DS8000 can be configured as RAID-5, RAID-10, or a combination of both. RAID-5 offers excellent performance for many customer applications, while RAID-10 may offer better performance for selected applications.

The IBM System Storage DS6000 series provides high availability and high performance in a small, modular package. This series,

along with the DS8000 series, is intended to offer an enterprise-class continuum of storage systems with shared replication services and common management interfaces. The DS6000 series systems are well suited to help simplify storage infrastructure, support business continuity, and improve information lifecycle management.

The DS6800 is designed to provide medium and large businesses with a low-cost, enterprise-class storage solution. The DS6800 can help simplify data management, provide extensive data protection, recovery capabilities, and easy scalability for both mainframe and open system storage needs

Advanced Features – Copy Services

Copy Services are a collection of functions supporting disaster recovery, data migration, and data duplication operations. Copy Services run on the IBM System Storage DS8000 storage unit and support UNIX, Windows, and System z environments. Many design characteristics and advanced functions of the DS8000 contribute to the protection of data. The DS8000 has a number of advanced Copy Services functions that are part of the IBM System Storage Resiliency family. These functions are supported on the previous generation of storage systems, IBM TotalStorage Enterprise Storage® Server (ESS). Copy Services include the following types of functions:

FlashCopy, a point-in-time copy function

Remote mirror and copy functions which include:

- Metro Mirror
- Global Copy
- Global Mirror

Management of the Copy Services functions is done through the IBM System Storage command-line interface (DSCLI) and Web-based interface called the IBM System Storage DS Storage Manager. The DS Storage Manager allows set up and management of data-copy functions from any point where network access is available.

FLASHCOPY

The FlashCopy feature enables creation of full volume copies of data. When you set up a FlashCopy operation, a relationship is established between source and target volumes, and a bitmap of the source volume is created. Once this relationship and a bitmap are created, the target volume can be accessed as though all the data had been physically copied. While a relationship between the source and target volume exists, a background process copies from the source to the target volume.

REMOTE MIRROR AND COPY

The remote mirror and copy feature is a flexible data mirroring technology that allows replication between volumes on two or more disk storage systems. You can also use this feature for data backup and disaster recovery. The DS8000 storage units participate in remote mirror and copy solutions with the ESS Model 750, ESS Model 800, and DS6000 storage units. The Remote Mirror and Copy feature operates in the following modes:

METRO MIRROR

Metro Mirror provides real-time mirroring of logical volumes between two or more IBM storage units that can be located up to 300 km from each other. It is a synchronous copy solution where write operations are completed on both copies (local and remote site) before they are considered to be completed.

GLOBAL COPY

Global Copy copies data asynchronously and over longer distances than is possible with Metro Mirror. When operating in Global Copy mode, the source volume sends a periodic, incremental copy of updated tracks to the target volume instead of a constant stream of updates. This causes less impact to application writes for source volumes and less demand for bandwidth resources, while allowing a more flexible use of the available bandwidth.

GLOBAL MIRROR

Global Mirror provides a long-distance remote copy feature across two sites using asynchronous technology. Global Mirror operations provide the following benefits

- Support for virtually unlimited distance between the local and remote sites, with the distance typically limited only by the capabilities of the network and the channel extension technology.
- This ”unlimited” distance enables you to choose your remote site location based on business needs and enables site separation to add protection from localized disasters.
- A consistent and restartable copy of the data at the remote site.

Note: For more information about copy services, visit *http://www.redbooks.ibm.com/abstracts/sg246788.html*

SDDPCM

SDDPCM is a Subsystem Device Driver Path Control Module for IBM System Storage DS8000 and DS6000 in Multi-Path IO (MPIO) environments, running on an IBM System p server. SDDPCM manages the paths to provide High Availability (HA) and load balancing of storage I/O, automatic path-failover protection, concurrent download of licensed internal code and prevention of a single-point-failure caused by host bus adapter, fiber channel cable, or host-interface adapter on supported storage.

Note: For more information about copy services, visit http://www-1.ibm.com/support/docview.wss?uid=ssg1S7000303&aid=1

Oracle Technologies

Oracle10g R2/Clusterware

Oracle10g R2 provides features that can help ease complex Oracle administrative tasks. Automatic Storage Management (ASM), Oracle Clusterware or Cluster Ready Services (CRS), Enhanced RMAN features, Flashback technology, performance tuning enhancements, job scheduler, Data Pump and Automatic Work Load Repository are some of the key features.

Cluster Synchronization Services

ASM is designed to work with both single database and RAC environments, ASM requires that Cluster Synchronization Services (CSS) be installed and started prior to ASM becoming available. In a single instance database environment, CSS maintains synchronization between the ASM and database instances. Oracle's Cluster Ready Services (CRS) includes the CSS component which is automatically installed on each node that runs ASM and is started when each node boots.

ASM and the shared disk storage components such as disks and disk groups are inherently monitored by CSS, providing cluster management and node monitoring management. Mounted disk groups and ASM itself are registered with CSS upon startup thereby keeping disk group metadata in sync across all RAC nodes. CSS thereby dynamically registers any new disk groups that are created and broadcast to the other cluster nodes.

Activities in both the database and ASM instances are synchronized using inter-node communication and the health of those ASM instances is verified by CSS. Structural changes that require synchronization such as adding or deleting a disk initiate the inter-node messages. Efficient synchronization of both ASM and the database are thereby accomplished with the same integrated lock management infrastructure.

Automatic Storage Management (ASM)

Automatic Storage Management (ASM) is a new feature in Oracle Database 10g. ASM provides a solution for storage management challenges. ASM provides an integrated storage management interface which maintains consistent volumes across all servers and storage platforms. ASM virtualizes storage into disk groups. ASM distributes data evenly across storage resources within disk groups to optimize performance and utilization. ASM provides three mirroring options for protection against disk failure as follows:

Mirror option	Mirror Description
external	defers redundancy to storage device
normal	2 way mirroring
high	3 way mirroring

SM uses a unique mirroring algorithm. ASM does not mirror disks, but rather it mirrors extents. As a result, a hot spare disk is not required. Only spare capacity is required within the disk group.

If a disk fails, ASM automatically reconstructs the contents of the failed disk on the surviving disks. Reading the mirrored contents from the surviving disks the I/O hit from a disk failure is spread across several disks, rather than on the single disk that mirrors the failed drive. When ASM allocates a primary extent of a file to one disk in a disk group, it allocates a mirror copy of that extent to another disk in the disk group. Primary extents on a given disk will have their respective mirror extents on one of several partner disks in the disk group. Each disk in a disk group has the same ratio of primary and mirror extents.

Advantages of ASM technology

- Reduction of administration tasks
- Reduction of volume management tasks
- Automation of I/O tuning for workloads
- Simplification of addition and removal of disks
- Reduction in downtime as there is no file system interface for file management

- Automatic balancing of load across disks in ASM disk group with database running following change to storage capacity

Automatic Storage Management Disk Groups

To reduce the complexity of managing ASM disk groups, Oracle recommends that generally no more than two disk groups be maintained and managed per RAC cluster or single ASM instance.

- Database Area: Active database files, control files, online redo logs and change tracking files used in incremental backups are stored in this area.
- Flash Recovery Area: Recovery related files multiplexed copies of current control files and redo logs, archive logs, backup sets and flashback log files

Defining a failure group within ASM is an additional level of security against disk failure. A failure group is a set of disks in a disk group sharing a common resource. ASM ensures that a primary extent and its mirror copy never reside in the same failure group. If you define failure groups for your disk group, ASM can tolerate the simultaneous failure of multiple disks in a single failure group. This allows ASM to mirror across disk on separate storage units to protect against the failure of an entire storage unit. Since it is tightly integrated with the database, ASM is able to take advantage of database features to recover from I/O failures.

Use Cases with Advanced Features

USE CASE 1 – COLD BACKUP USING FLASHCOPY OF ORACLE 10G RAC DATABASE WITH ASM

Historically a cold backup of a database is the desired method to ensure the reliable recovery of a database. A cold backup requires all database processes to be stopped; the database is unavailable to the users until the entire database is backed up. Utilizing the flash copy features of the DS8000 and DS6000 provide for the completion of a cold backup without the wait of conventional backup methods.

ASM manages data spread over multiple volumes. When FlashCopy is used for multiple volumes, they volumes must be at a consistent level. Consistency groups create a consistent Point-in-Time copy across multiple volumes, and even across multiple DS8000 storage systems, thus managing the consistency of *dependent writes*.

When a FlashCopy operation is invoked, it takes only a few seconds to complete the process of establishing the FlashCopy consistency group pairings and creating the necessary control bitmaps. Thereafter, you have access to a Point-in-Time Copy of the source ASM disk group. As soon as the pairs have been established, you can read and write to both the source and target volumes. After creating the bitmap, a background process begins to copy the real data from the source to the target volumes. If you access the source or the target volumes during the background copy, FlashCopy manages these I/O requests, and facilitates both reading from and writing to both the source and target copies. When all the data has been copied to the target, the FlashCopy relationship is ended.

With the Freeze FlashCopy Consistency Group option, the System Storage unit holds off I/O activity to a volume for a time period by putting the source volume in a *queue full* state. Therefore, a time slot can be created during which dependent write updates do not occur, and FlashCopy uses that time slot to obtain a consistent point-in-time copy of the related volumes. I/O activity resumes when all FlashCopy relationships are established.

To utilize the FlashCopy process for a cold backup of an Oracle database on ASM and DS8000 or DS6000, shutdown the database and execute the FlashCopy relationship. While the relationship between the source and target volume exists, a background process copies the tracks from the source to the target volume. Restart the database and resume normal operations.

Note: For more information about Creating a FlashCopy relationship, visit: http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp

USE CASE 2 – ORACLE HOT BACKUP

A hot backup of an Oracle database provides the convenience of making a backup while the database is up and running thus eliminating the inconvenience of shutting the database down that a cold backup requires. While it is possible for users to continue I/O on the database when it is in hot backup mode the writes to the data files are held until the database is taken out of hot backup mode.

To utilize the FlashCopy process for a hot backup of an Oracle database on ASM and DS8000 or DS6000, alter the database to backup mode and execute the FlashCopy relationship. With hot backup it is imperative to execute the FlashCopy relationship with consistency groups. While the relationship between the source and target volume exists, a background process copies the tracks from the source to the target volume. Alter the database and end the backup mode once the FlashCopy relationship has been made.

Note: For more information about Creating a FlashCopy relationship, visit: *http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp*

USE CASE 3 – REMOTE HOT BACKUP WITH METRO MIRROR

The Remote Mirror and Copy is a flexible data mirroring technology that allows replication between volumes on two or more disk storage systems. DS8000 Storage Units participate in Remote

Mirror and Copy solutions with the ESS Model 750, ESS Model 800, and DS6000 Storage Units. Metro Mirror provides real-time mirroring of logical volumes between two DS8000s that can be located up to 300 km from each other. It is a synchronous copy solution where write operations are completed on both copies (local and remote site) before they are considered to be complete.

Remote hot backup uses the same FlashCopy mechanism with the link between source and target disks managed over a connection between Storage units that may be separated over a greater distance such as a mirror site located in another location. The same rules apply Consistency groups are required and a fibre connection between the storage units is required. The database is put into hot backup mode and the Metro Mirror process is executed. Once the relationship is executed the database is then brought out of hot backup mode and continues normal operation

Lab Test Environment

Oracle RAC Servers

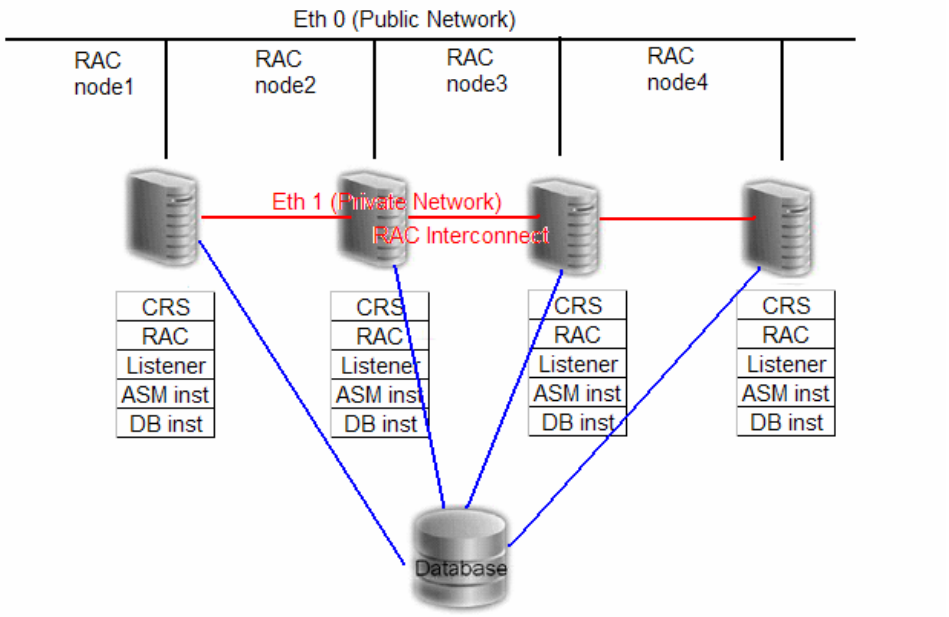
Four IBM eServer™ pSeries® p630 Servers with 4 CPUs each are configured as Oracle RAC servers.

Features	
Machine Model	IBM eServer pSeries p630
CPU	4
RAM	4GB
Internal Disk	4 x 9 GB
I/O adapter	2 x LP9002
OS version	AIX 5.3 ML3
Oracle version	Oracle 10.2.0.2 DB, Oracle Clusterware, Oracle ASM

Storage Hardware

Features		
Storage	IBM DS8000 2107-900	IBM DS6000 1750
Fibre Switch	IBM 2109	IBM 2109
Network	Ethernet switch	Ethernet switch
Storage devid	IBM.2107-7502301	IBM.1750-13AB73A

Topology of 4-node Oracle 10g R2 RAC with ASM



Configuration by Example

Storage Side Configuration using DSCLI

Storage DSCLI commands are the same for DS8000 and DS6000. Use the following steps for a DS6000 identifier, IBM.1750-13AB73A, to set up the storage configuration:

Set up storage pool (extent pool)

Perform the following steps to create the extent pool:

1. Create an array.

```
dscli> mkarray -dev IBM.1750-13AB73A -raidtype 5 -
arsite IBM.1750-13AB73A/S1
```

2. Create one fixed block from one array.

```
dscli> mkrank -dev IBM.1750-13AB73A -array A0 -stgtype
fb
```

3. Create a fixed block storage type extent pool.

```
dscli> mkextpool –rankgrp 0 –stgtype fb ora_RAC
```

4. Assign an unassigned rank to a extent pool.

```
dscli> chrank -extpool p0 r0
```

5. Display a list of array sites and status information.

```
dscli> lsarraysite

arsite       DA Pair      dkcap (10^9B) State   Array
=============================================================
IBM.1750-13AB73A/S1 IBM.1750-13AB73A/0     146.0 Assigned  IBM.1750-
13AB73A/A0
```

6. Display a list of defined ranks in a storage image and status information.

```
dscli> lsrank

ID      Group State datastate Array  RAIDtype extpoolID
stgtype
==========================================================================
=================

IBM.1750-13AB73A/R0   0 Normal Normal  IBM.1750-13AB73A/A0 5
IBM.1750-13AB73A/P0 fb
```

List the extent pool.

```
dscli> lsextpool

Name    ID  stgtype rankgrp status availstor (2^30B)
%allocated available reserved numvols
=============================================================

ora_RAC  IBM.1750-13AB73A/P0 fb   0 below   48     81
48    0   28
```

Make volume group and LUNs

Perform the following steps to create volume group and LUNs:

1. Create a volume group in a storage image.

```
dscli> mkvolgrp -type scsimask  Aix_oracle
 creates volumegroup named Aix_oracle and assigns an
identifier in this case the identifier is v12
```

2. Create an open systems fixed block volume in a storage image.

```
dscli> mkfbvol -extpool p0 -cap 1 –volgrp v12 -name
oracle_#h 0210-0213
1GB luns for CRS and vote disks (4 total)
dscli> mkfbvol -extpool p0 -cap 10 –volgrp v12 -name
oracle_#h 0410-0419
10g luns for ASM disks (10 total)
```

3. List the resulting volume group and its members.

```
dscli> showvolgrp v12
Name Aix_oracle
ID  IBM.1750-13AB73A/V12
Type SCSI Mask
```

```
     Vols 0210 0211 0212 0213 0410 0411 0412 0413 0414 0415 0416
     0417 0418 0419
```

Make host connection and LUN assignments to nodes

Each node has two HBAs. Therefore, there will be two host connections per node. Perform the following steps to make the host connections:

1. Make an I/O port and host connect configuration.

```
dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c92ce4fd -profile
"IBM pSeries - AIX"  -volgrp v12    node1_h0

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c92ce46e -profile
"IBM pSeries - AIX"  -volgrp v12  node1_h1

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c92ceb85 -profile
"IBM pSeries - AIX"  -volgrp v12  node2_h0

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c92ce23c -profile
"IBM pSeries - AIX"  -volgrp v12   node2_h1

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c930e637 -profile
"IBM pSeries - AIX"  -volgrp v12  node3_h0

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c930e59b -profile
"IBM pSeries - AIX"  -volgrp v12  node3_h1

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c930e45a -profile
"IBM pSeries - AIX"  -volgrp v12  node4_h0

dscli> mkhostconnect  -dev IBM.1750-13AB73A -wwname 10000000c930e698 -profile
"IBM pSeries - AIX"  -volgrp v12  node4 h1
```

2. Display a list of host connections.

```
dscli> lshostconnect

Date/Time: October 31, 2006 5:20:34 PM PST IBM DSCLI Version: 5.0.6.229
DS: IBM.1750-13AB73A
Name         ID           WWPN        HostType     Profile       portgrp
volgrpID       ESSI
==========================================================================
==================node1_h0  IBM.1750-13AB73A/0002 10000000C92CE4FD
IBM pSeries - AIX    0 IBM.1750-13AB73A/V12 all

node1_h1  IBM.1750-13AB73A/0003 10000000C92CE46E    IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all

node2_h0  IBM.1750-13AB73A/0004 10000000C92CEB85    IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all

node2_h1  IBM.1750-13AB73A/0005 10000000C92CE23C    IBM pSeries - AIX
```

```
0 IBM.1750-13AB73A/V12 all

node3_h0  IBM.1750-13AB73A/0006 10000000C930E637     IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all

node3_h1  IBM.1750-13AB73A/0007 10000000C930E59B    IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all

node4_h0  IBM.1750-13AB73A/0008 10000000C930E45A    IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all

node4_h1  IBM.1750-13AB73A/0009 10000000C930E698     IBM pSeries - AIX
0 IBM.1750-13AB73A/V12 all
```

Note: For information about DSCLI command, visit: http://www.redbooks.ibm.com/redbooks/SG246452

3. **Storage allocation on hosts**

Use Emulex LP9002 HBAs with the following parameter settings:

- Setting up the "Fast I/O failure" supports faster failover to the alternate path.
- Dynamic tracking logic is called when the adapter driver receives an indication from the switch that there has been a link event involving a remote storage device port

These features should be set on all fscsi controllers in an AIX host as follows:

1. Change the characteristics of a device

```
$ chdev -l fscsi0 -a fc_err_recov=fast_fail

$ chdev -l fscsi0 -a dyntrk=yes
```

2. Display attribute characteristics

```
      $ lsattr –El fscsi0
attach       switch       How this adapter is CONNECTED      False
dyntrk       yes          Dynamic Tracking of FC Devices    True
fc_err_recov fast_fail    FC Fabric Event Error RECOVERY Policy   True
scsi_id      0xa1900      Adapter SCSI ID            False
sw_fc_class  3      FC Class for Fabric              True
```

3. Display information about devices in the Device Configuration database

```
      $ lsdev  |grep hdisk

hdisk0   Available 1S-08-00-8,0 16 Bit LVD SCSI Disk Drive (internal disk)
```

```
hdisk1   Available 1S-08-00-9,0 16 Bit LVD SCSI Disk Drive (internal disk)
hdisk2   Defined  1H-08-01   IBM MPIO FC 2107
hdisk3   Defined  1H-08-01   IBM MPIO FC 2107
hdisk4   Defined  1H-08-01   IBM MPIO FC 2107
hdisk5   Defined  1H-08-01   IBM MPIO FC 2107
hdisk6   Defined  1H-08-01   IBM MPIO FC 2107
```

4. Display the name, location, and description of each device found in the current configuration

```
$ lscfg –v l hdisk2
$ lscfg -vl hdisk2 |grep Serial

Serial Number...............75023012100
```

5. List major and minor numbers for each host

```
$ ls -l hdisk2
brw-------  1 root   system    25, 7 Jul 20
16:57 hdisk2
```

Use SDDPCM to manage the fiber connections

SDDPCM provides commands to display the status of adapters used to access managed devices, to display the status of the devices that the device driver manages, or to map supported storage MPIO devices or paths to a supported storage device location. This section includes descriptions of the command.

```
Usage: pcmpath query adapter [n]
    pcmpath query device [n/-d <device_model>]
    pcmpath query adaptstats [n]
    pcmpath query devstats [n/-d <device_model>]
    pcmpath query portmap
    pcmpath query essmap
    pcmpath query wwpn
    pcmpath set adapter <n> online/offline
    pcmpath set device <n> path <m> online/offline
    pcmpath set device <n>/(<n> <n>) algorithm rr/fo/lb
    pcmpath set device <n>/(<n> <n>) hc_interval <t>
    pcmpath set device <n>/(<n> <n>) hc_mode
nonactive/enabled/failed
    pcmpath open device <n> path <m>
    pcmpath disable/enable ports <location> ess <essid>
      single port     = R1-Bx-Hy-Zz
      all ports on card  = R1-Bx-Hy
      all ports on bay   = R1-Bx
      refer portmap output for the location string and ESS
serial number
```

1. Display the SDDPCM path information.

```
       $ pcmpath query device

DEV#:  2 DEVICE NAME: hdisk2 TYPE: 2107900 ALGORITHM: Load
```

```
Balance

SERIAL: 7502301 2100
============================================================================
=========
Path#    Adapter/Path Name      State    Mode    Select    Errors
  0       fscsi0/path0        OPEN   NORMAL    171652      0
  1       fscsi0/path1        OPEN   NORMAL    171502      0
  2       fscsi1/path2        OPEN   NORMAL    171440      0
  3       fscsi1/path3        OPEN   NORMAL    171543      0
```

Note: For more information about SDDPCM, visit: http://www-1.ibm.com/support/docview.wss?uid=ssg1S7000303&aid=1

Create Cluster disks using DS8000/DS6000 Volumes

1. Create the shared volume using DS8000/DS6000.
2. Assign these volumes for the CRS OCR and vote disks.
3. Create the major and minor numbers and the character device of those disks.

```
$ chdev -l hdisk2 -a reserve_policy=no_reserve
$ chdev -l hdisk3 -a reserve_policy=no_reserve
$ mknod /dev/ocr_disk c 25 7
$ mknod /dev/vote_disk c 25 9
```

The following table illustrates how the disk major and minor numbers were applied on all 4 nodes:

Storage		Node1		Node2		Node3		Node4	
LUN ID		Major	Minor	Major	Minor	Major	Minor	Major	Minor
2100	ocr_disk	25	7	43	12	43	0	25	1
2101	vote_disk	25	9	43	11	43	10	25	7
4100	asm_disk1	25	2	43	0	43	3	25	2
4101	asm_disk2	25	13	43	10	43	13	25	11

4. Set proper permission and access level on those shared volumes.

```
$ chown oracle:dba ocr_disk
$ chown oracle:dba vote_disk
$ chmod 660 ocr_disk
```

```
$ chmod 660 vote_disk
```

Note: The disk names and their permissions should be consistent across all the RAC nodes.

5. During the CRS installation, select the external redundant disks to be used as the OCR and voting disks.

Prepare DS8000/DS6000 Volumes for ASM

Perform the following steps to create the ASM disks:

1. Create the shared volume using DS8000/DS6000.
2. Assign these volumes for the ASM disks.
3. Create a major and minor numbers and a character device of those disks.

```
$ chdev -l hdisk6 -a reserve_policy=no_reserve
$ chdev -l hdisk7 -a reserve_policy=no_reserve
$ mknod /dev/asm_disk1 c 25 2
$ mknod /dev/asm_disk2 c 25 13
```

4. Set proper permission and access level on those ASM disks.

```
$ chown oracle:dba /dev/asm_disk1
$ chown oracle:dba /dev/asm_disk2
$ chmod 660 /dev/asm_disk1
$ chmod 660 /dev/asm_disk2
```

Best Practice Recommendations

1. Use LUNs with similar capacity and performance within an ASM disk group.
2. Use multiple LUNs per ASM disk group to evenly distribute database reads and writes across members of the disk group.
3. Place datafiles and logfiles in same disk group.
4. Configure ASM disk groups to use external redundancy for DS8000 and DS6000.
5. Use IBM SDDPCM Multipathing device driver for high availability.
6. Enable dynamic tracking and fast failover if Emulex HBAs are used
7. Use disks with similar capacity and performance when adding space to a disk group.
8. Do not use more than two disk groups per RAC cluster or single ASM instance.

ASM simplifies database file management on IBM System Storage disk products. External redundancy allows the host to use CPU resources for data management and less CPU resources for storage management. These combined technologies enhance productivity for both database administrators and system administrators.

Conclusion

The use of Oracle ASM with IBM System Storage DS8000 and DS6000 products can help greatly simplify storage administration. ASM assumes the workload of balancing I/O within the disk group, making the addition and/or removal of physical storage possible without having to shutdown the Oracle database. ASM is designed to provide fault tolerance, allowing the configuration of duplicate or triplicate copies of any database files. ASM also takes advantage of existing DS8000 and DS6000 fault tolerance mechanisms for increased data protection.

References

1. Oracle Real Application Clusters Installation and Configuration Guide 10g Release 1 (10.1) for AIX-Based Systems, Apple Mac OS X, hp HP-UX, hp Tru64 UNIX, Linux, Solaris Operating System, and Windows Platforms Part No. B10766-08 http://download-east.oracle.com/docs/html/B10766_08/racinstl.htm
2. ASM 10gR2 New Features
   http://asm.us.oracle.com/pdf/ASM R2 New Features.pdf
   http://www.oracle.com/technology/products/database/asm/pdf/asm%20r2%20new%20features.pdf
3. ASM Best practices
   http://www.oracle.com/technology/products/database/asm/pdf/asm_10gr2_bptwp_sept05.pdf
   http://www.oracle.com/technology/products/database/asm/pdf/asm_10gr2_bestpractices%2005-07.pdf
4. Technical White Paper on Automatic Storage Management
   http://www.oracle.com/technology/products/manageability/database/pdf/asmwp.pdf
5. Taking the Guesswork out of db tuning
   http://www.oracle.com/technology/products/database/asm/pdf/take%20the%20guesswork%20out%20of%20db%20tuning%2001-06.pdf
6. Overview of Automatic Storage Management
   http://www.oracle.com/technology/products/manageability/database/pdf/asmov.pdf
7. IBM System Storage DS8000 Series: Copy Services in Open Environments
   http://www.redbooks.ibm.com/abstracts/sg246788.html
8. The IBM System Storage DS8000 Series: Concepts and Architecture
   http://www.redbooks.ibm.com/redbooks/SG246452

9. The IBM System Storage DS8000 Information Center http://publib.boulder.ibm.com/infocenter/dsichelp/ds8000ic/index.jsp
10. SDDPCM Users Manual http://www-1.ibm.com/support/docview.wss?uid=ssg1S7000303&aid=1
11. FINAL Release V1 COOKBOOK 10gRAC R2 – AIX5L – SAN Storage http://www.oracleracsig.org/pls/htmldb/RAC_SIG.download_my_file?p_file=1000520