COOKBOOK

Quick Installation Guide

Oracle 10*g* RAC Release 2 and Oracle Automatic Storage Management (ASM) on IBM @server System p and i Running AIX 5L with SAN Storage

Version 2.0
May 2007

ORACLE / IBM
Joint Solutions Center

IBM – Thierry Plumeau
Oracle/IBM Joint Solutions Center

Oracle - Frederic Michiara
Oracle/IBM Joint Solutions Center

This document is based on our experiences.
This is not an official (Oracle or IBM) documentation.
This document will be constantly updated
and we're open to any add-on
or feedback from your own experiences,
on same or different storage solution !!!

Document history :

Version	Date	Update	Who	Validated by
1.0	**January 2006**	- Creation	Frederic Michiara Thierry Plumeau	Fabienne Lepetit Paul Bramy Alain Roy Marc Tu-Duy-Kiem
1.1	**April 2006**	- Update (Storage, Multi-path I/O with IBM RDAC and MPIO, HDS Hitachi and EMC Power Path configuration, racgvip script modification for virtual network) - Appendix D	Frederic Michiara Thierry Plumeau	Fabienne Lepetit Paul Bramy Alain Roy Marc Tu-Duy-Kiem
2.0	**March 2007**	- Update - Removed GPFS part and created a specific ***Quick Installation Guide*** **Oracle 10*g* RAC Release 2 and IBM GPFS on IBM eServ System p and i running AIX 5L with SAN Storage** - Separate ASM Home from DB Home - asmcmd, ftp and http, dbconsole - Multi-pathing / storage update - Architecture update - 10.2.0.3 Patchset installation …	Frederic Michiara Thierry Plumeau	Fabienne Lepetit Paul Bramy Alain Roy

Contributors :

- **Paul Bramy : Oracle Corp. (ORACLE/IBM Joint Solutions Center)**
- **Alain Roy : IBM France (ORACLE/IBM Joint Solutions Center)**
- **Fabienne Lepetit : Oracle France (ORACLE/IBM Joint Solutions Center)**
- **Marc Tu-Duy-Kiem : IBM France (ORACLE/IBM Joint Solutions Center)**
- **Stephan Navaro : IBM France**
- **Rebecca Balough : IBM ATS (US)**

Contact :

- **ORACLE / IBM Joint Solutions Center ➔ oraclibm@fr.ibm.com**

1 THE AIM OF THIS DOCUMENT

This document is written to provide help installing Oracle10*g* Real Application Clusters (10.2) release 2 on IBM System p and i servers with AIX 5L.

We will describe step by step the architecture Oracle CRS (Cluster Ready Service) on raw disks and database on ASM (Automatic Storage Management).
For the architecture using GPFS for CRS Data and Database, see the other cookbook named :

Quick Installation Guide
Oracle 10*g* RAC Release 2 and IBM GPFS on IBM eServer System p and i running AIX 5L with SAN Storage.

As ASM and GPFS are mostly used in today customers architectures, "Concurrent Raw devices using IBM HACMP" has not been covered in this 10gRAC Release 2 cookbook, please read our 10gRAC Release 1 cookbook if you need to implement it. When implementing 10gRAC Release 2 with HACMP installed on the System p servers, you'll have to check metalink note 404474.1 Status of Certification of Oracle Clusterware with HACMP 5.3 & 5.4. As for today March 27, 2007 ONLY HACMP 5.2 is supported/certified with 10gRAC Release 2.

With these two different implementations, you'll be able to install some other combinations.

Metalink (http://metalink.oracle.com/metalink/plsql/ml2_gui.startup)

Titre	Origine	Référence
Oracle® Database Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10*g* Release 2 (10.2) for AIX http://download-uk.oracle.com/docs/cd/B19306_01/install.102/b14201/toc.htm	http://tahiti.oracle.com/	B14201-04
Oracle Clusterware and Oracle Real Application Clusters Administration and Deployment Guide - 10*g* Release 2 (10.2)	http://tahiti.oracle.com/	B14197-03
Oracle Database Release Notes – 10*g* Release 2 (10.2) for AIX 5L Based Systems (64bits)	http://tahiti.oracle.com/	B19074-06
Oracle Universal Installer Concepts Guide Release 2.2	http://tahiti.oracle.com/	A96697-01
Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries	Oracle Metalink http://metalink.oracle.com/	282036.1
Status Of Certification of Oracle Clusterware with HACMP 5.3 & 5.4	Oracle Metalink http://metalink.oracle.com/	404474.1
Oracle ASM and Multi-Pathing Technologies	Oracle Metalink http://metalink.oracle.com/	294869.1
Oracle® Database 2 Day + Real Application Clusters Guide 10*g* Release 2 (10.2)	http://tahiti.oracle.com/	B28759-01

The information contained in this paper resulted from :

- Oracle and IBM documentations
- Workshop experiences done in the Oracle/IBM Joint Solutions Center
- Benchmarks and POC implementations for customers performed by PSSC Montpellier
- This documentation is a joint effort from Oracle and IBM specialists.

Please also refer to Oracle online documentation for more information :

- http://docs.oracle.com
- http://tahiti.oracle.com
- Oracle RAC home page : http://www.oracle.com/database/rac_home.html
- For HACMP Documentation refer to : http://www1.ibm.com/servers/eserver/pseries/library/hacmp_docs.html
- For GPFS Documentation refer to : http://www-1.ibm.com/servers/eserver/pseries/library/gpfs.html
- For more information : http://www-1.ibm.com/servers/eserver/pseries/software/sp/gpfs_faq.html

Your comments are important for us, and we thanks the ones who send us their feedback about previous release, and about how this document did help them in their implementation. We want our technical papers to be as helpful as possible.

Please send us your comments about this document to the Oracle/IBM Joint Solutions Center.

Use our email address :	**Or our phone number :**
oraclibm@fr.ibm.com	**+33 (0)4 67 34 67 49**

2 ABOUT ORACLE CLUSTERWARE AND REAL APPLICATION CLUSTER

Extract from : Oracle® Database 2 Day + Real Application Clusters Guide 10*g* Release 2 (10.2)
Part Number B28759-01
http://download-uk.oracle.com/docs/cd/B19306_01/rac.102/b28759/toc.htm

Oracle RAC extends Oracle Database so that you can store, update, and efficiently retrieve data from multiple instances on different servers at the same time. Oracle RAC provides the software that allows the servers to work together in what is called a cluster. The physical structures that make up the database must reside on shared storage that is accessible from all servers that are part of the cluster. Each server in the cluster runs the Oracle RAC software.

Single-instance Oracle databases have a one-to-one relationship between datafiles and the instance. Oracle RAC environments, however, have a one-to-many relationship between datafiles and instances. In Oracle RAC environments, multiple cluster database instances form a single database. The instances can be on different servers, referred to as hosts or nodes. The combined processing power of the multiple servers provides greater throughput and scalability than is available from a single server.

Each cluster database instance in an Oracle RAC cluster uses its own memory structures and background processes. Oracle RAC uses Cache Fusion to synchronize the data stored in the buffer cache of each cluster database instance. Cache Fusion moves current data blocks (which reside in memory) between database instances, rather than having one database instance write the data blocks to disk and requiring another database instance to reread the data blocks from disk. When a data block located in the buffer cache of one instance is required by another instance, Cache Fusion transfers the data block directly between the instances using the interconnect, enabling the Oracle RAC database to access and modify data as if the data resided in a single buffer cache.

The Oracle RAC infrastructure is also a key component for implementing the Oracle enterprise grid computing architecture. Having multiple cluster database instances accessing a single database prevents the server from being a single point of failure. Any packaged or custom application that ran well on a single-instance Oracle database will perform well on an Oracle RAC database without requiring code changes.

Using our cookbook will ease the implementation of RAC with ASM on System p running AIX5L, but for more details on RAC, You will learn more about the operation of the database server in a cluster as :

- Administering Database Instances and Cluster Databases
- Administering Oracle Clusterware Components
- Administering Backup and Recovery
- Managing Database Workload Using Services
- Monitoring and Troubleshooting
- Adding Nodes and Instances
- Managing Oracle Software and Applying Patches

And all that from **Oracle® Database 2 Day + Real Application Clusters Guide 10*g* Release 2 (10.2) (Part Number B28759-01.**

3 ABOUT ORACLE AUTOMATIC STORAGE MANAGEMENT

Extract from : Oracle® Database 2 Day + Real Application Clusters Guide 10*g* Release 2 (10.2)
Part Number B28759-01
http://download-uk.oracle.com/docs/cd/B19306_01/rac.102/b28759/toc.htm

With Oracle RAC, each instance must have access to the datafiles and recovery files for the Oracle RAC database. Using Automatic Storage Management (ASM) is an easy way to satisfy this requirement.

ASM is an integrated, high-performance database file system and disk manager. ASM is based on the principle that the database should manage storage instead of requiring an administrator to do it. ASM eliminates the need for you to directly manage potentially thousands of Oracle database files.

ASM groups the disks in your storage system into one or more disk groups. You manage a small set of disk groups and ASM automates the placement of the database files within those disk groups.

ASM provides the following benefits:

- **Striping** — ASM spreads data evenly across all disks in a disk group to optimize performance and utilization. This even distribution of database files eliminates the need for regular monitoring and I/O performance tuning.

- **Mirroring** — ASM can increase data availability by optionally mirroring any file. ASM mirrors at the file level, unlike operating system mirroring, which mirrors at the disk level. Mirroring means keeping redundant copies, or mirrored copies, of each extent of the file, to help avoid data loss caused by disk failures. The mirrored copy of each file extent is always kept on a different disk from the original copy. If a disk fails, ASM can continue to access affected files by accessing mirrored copies on the surviving disks in the disk group.

- **Online storage reconfiguration and dynamic rebalancing** — ASM permits you to add or remove disks from your disk storage system while the database is operating. When you add a disk, ASM automatically redistributes the data so that it is evenly spread across all disks in the disk group, including the new disk. The process of redistributing data so that it is also spread across the newly added disks is known as rebalancing. It is done in the background and with minimal impact to database performance.

- **Managed file creation and deletion** — ASM further reduces administration tasks by enabling files stored in ASM disk groups to be managed by Oracle Database. ASM automatically assigns file names when files are created, and automatically deletes files when they are no longer needed.

ASM is implemented as a special kind of Oracle instance, with its own System Global Area and background processes. The ASM instance is tightly integrated with the database instance. Every server running one or more database instances that use ASM for storage has an ASM instance. In an Oracle RAC environment, there is one

ASM instance for each node, and the ASM instances communicate with each other on a peer-to-peer basis. Only one ASM instance is required for each node regardless of the number of database instances on the node. Oracle recommends that you use ASM for your database file storage, instead of raw devices or the operating system file system. However, databases can have a mixture of ASM files and non-ASM files.

4 WHAT'S NEW WITH RAC IMPLEMENTATION ON AIX5L

With Oracle 10gRAC on IBM System p, i running AIX5L :

- **HACMP is not any more necessary as clusterware software since 10gRAC Release 1.**
- **Oracle provide with 10gRAC its own clusterware, named as Oracle Clusterware, or CRS (Oracle Cluster Ready Service).**
- **Oracle Clusterware can cohabitate with HACMP under some conditions.**
- **With 10gRAC, Oracle Clusterware is mandatory as clusterware, even so other vendors clusterware software are installed on the same systems. You MUST check that third party clusterware can cohabitate with Oracle clusterware.**

  Subject: **Using Oracle Clusterware with Vendor Clusterware FAQ**

  Doc ID: **Note:332257.1**

- **With 10gRAC Release 2 :**
  - **Oracle VIP (Virtual IP) are now configured at the end of the Clusterware installation.**
  - **ASM software can be installed in its own ORACLE_HOME directory which will be different than the database software ORACLE_HOME.**

5 WAYS TO IMPLEMENT 10G RAC ON AIX5L

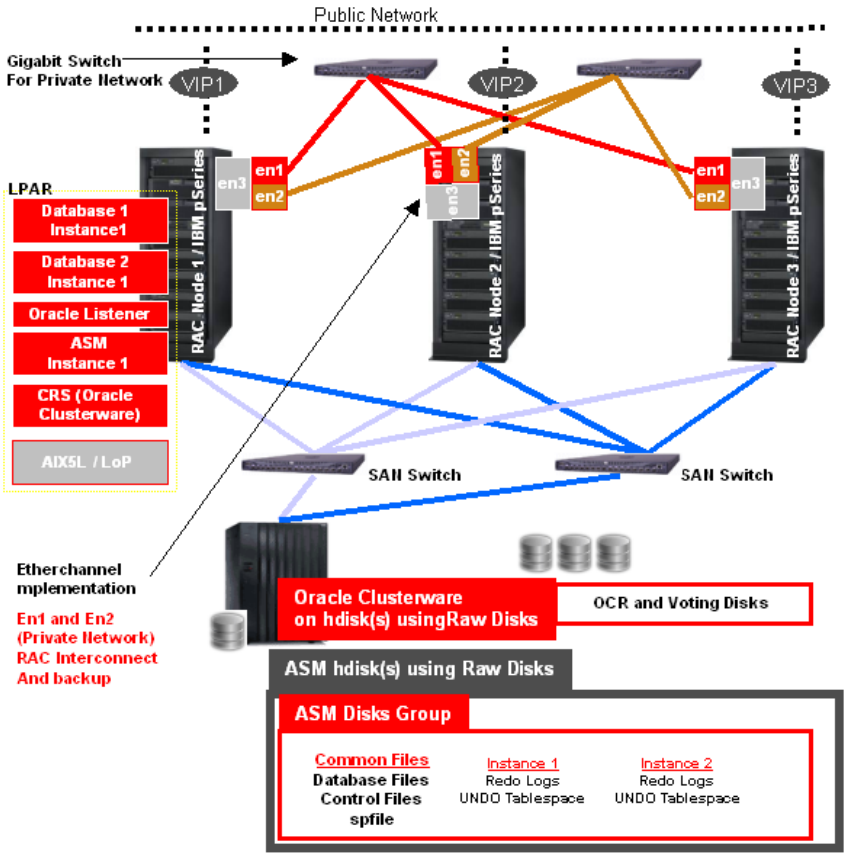
There is two type of infrastructure to implement RAC :

- **Over LAN (Local Area Network)** ➔ - 2 to n pSeries, 1 storage
  - 2 to n pSeries, 2 storages
- **Over MAN (Metropolitan Network)** ➔ - 2 to n pSeries, 2 storages with or without DWDM technology, depending of the distance between 2 sites.

5.1 10G RAC OVER LOCAL AREA NETWORK

5.1.1 ASM (Automated Storage Management)

10gRAC with OCR (Oracle Cluster Registry) disk(s), and Voting (Heartbeat) disk(s) on raw disks, and database on Oracle ASM (Automated Storage Management).



Oracle Storage Management Solution :

➔ No need for HACMP

➔ No need for GPFS

➔ CRS files (OCR and Voting) are placed on raw disks.

➔ Only Oracle databases files (datafiles, redo logs, archive logs, flash recovery area, …) are stored on the disks managed by Oracle ASM. No binaries.

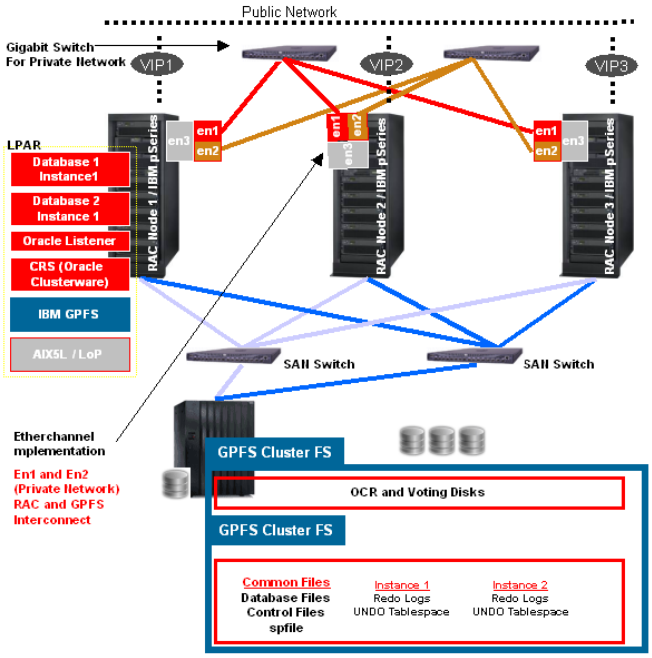
➔ ASM provided with oracle software

ORACLE CLUSTERWARE MANDATORY !!!

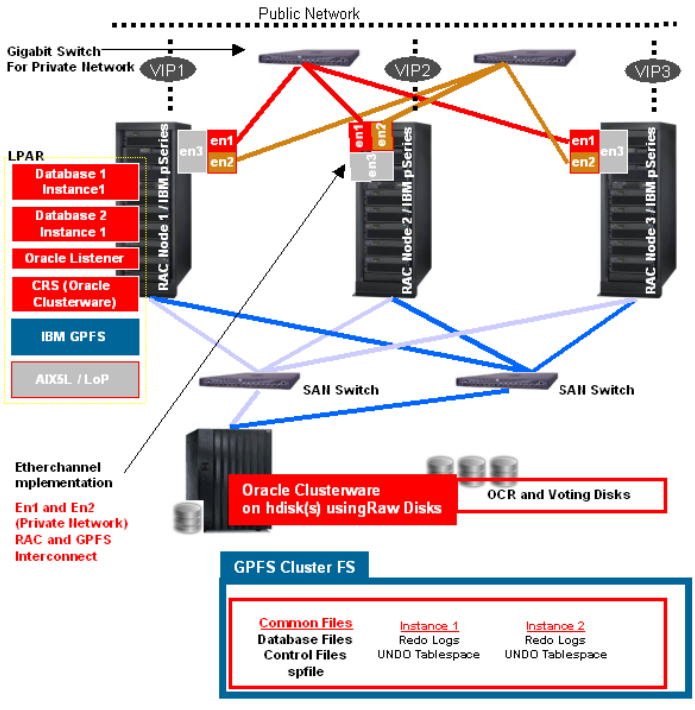
5.1.2 IBM GPFS (General Parallel File System)

10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s) and database on Cluster Files System with IBM GPFS (General Parallel File System).

Case 1 : OCR , voting disks, and Databases files on GPFS



Case 2 : OCR, voting disks on raw disks, and Databases files on GPFS



IBM Cluster Files System Solution :

➔ No need for HACMP

➔ GPFS required

➔ All files (Oracle database files, CRS files with OCR and Voting, any binaries, any files) can be stored on IBM GPFS.

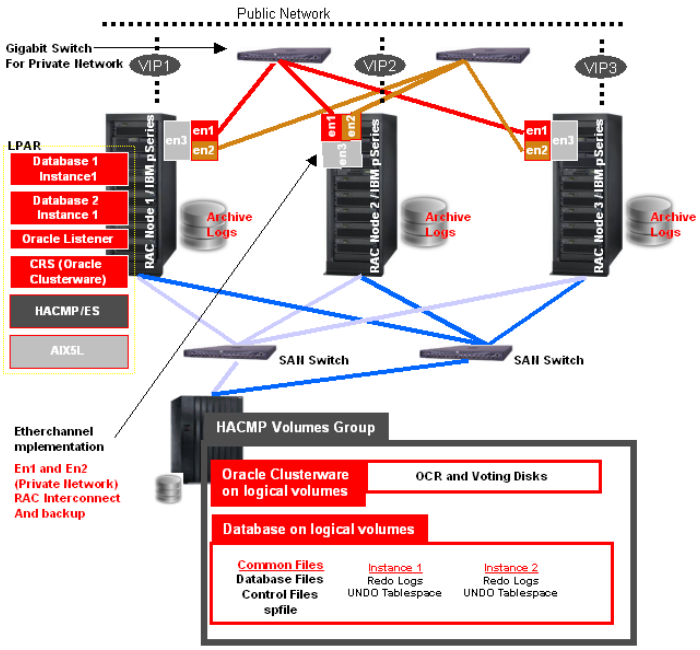
ORACLE CLUSTERWARE MANDATORY !!!

Implementation BEST practice

By experience,
It's best advice to store OCR and Voting disks on raw disks.

5.1.3 Concurrent Raw Devices with Volumes Groups (HACMP)

10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s), and database on concurrent raw devices with HACMP.



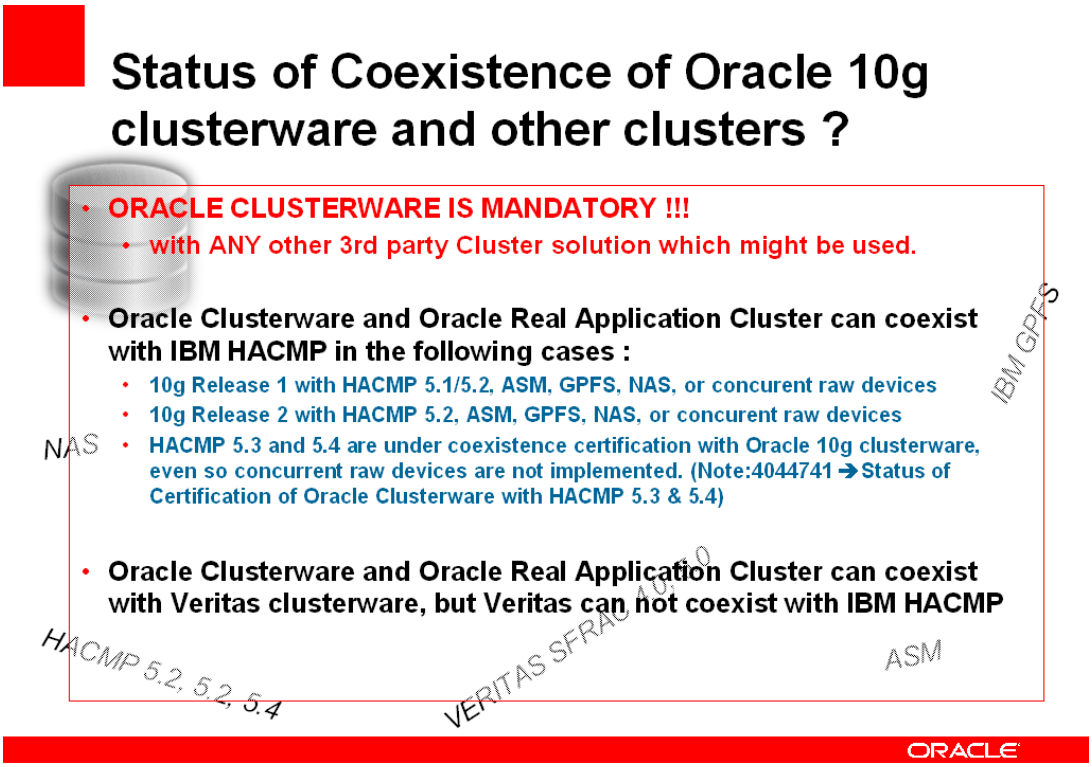
IBM HACMP Solution :

➔ **HACMP is ONLY mandatory to achieve concurrent acess (read/write) to shared raw devices declared in a Volume Group.**

➔ **HACMP can not replace Oracle Clusterware**, it's only used to make raw devices concurrent to all nodes.

➔ Only Oracle databases files (datafiles, redo logs, spfiles, CRS files with OCR and Voting) are stored on the concurrent raw devices. No archive logs, No binaries.

ORACLE CLUSTERWARE MANDATORY !!!

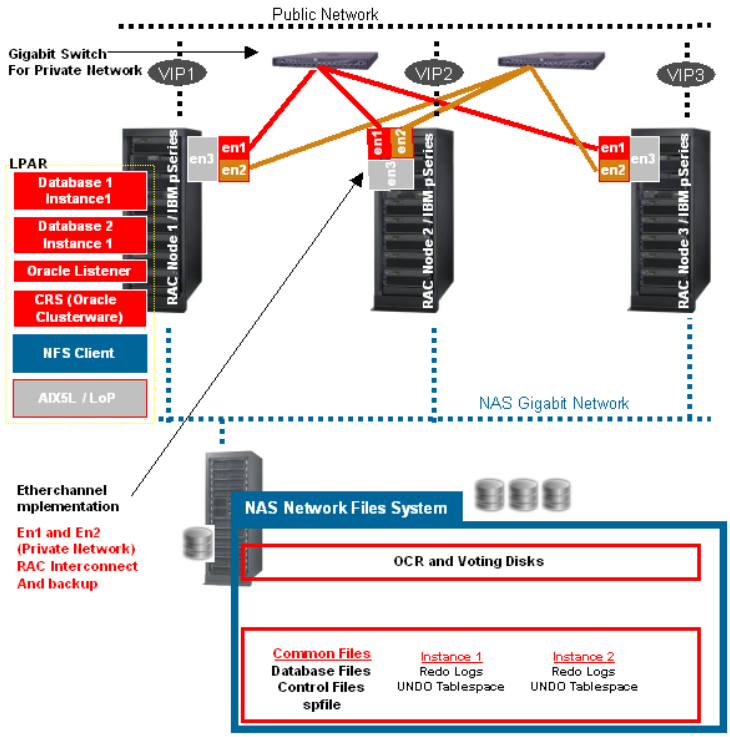


Subject: **Using Oracle Clusterware with Vendor Clusterware FAQ** Doc ID: **Note:332257.1**

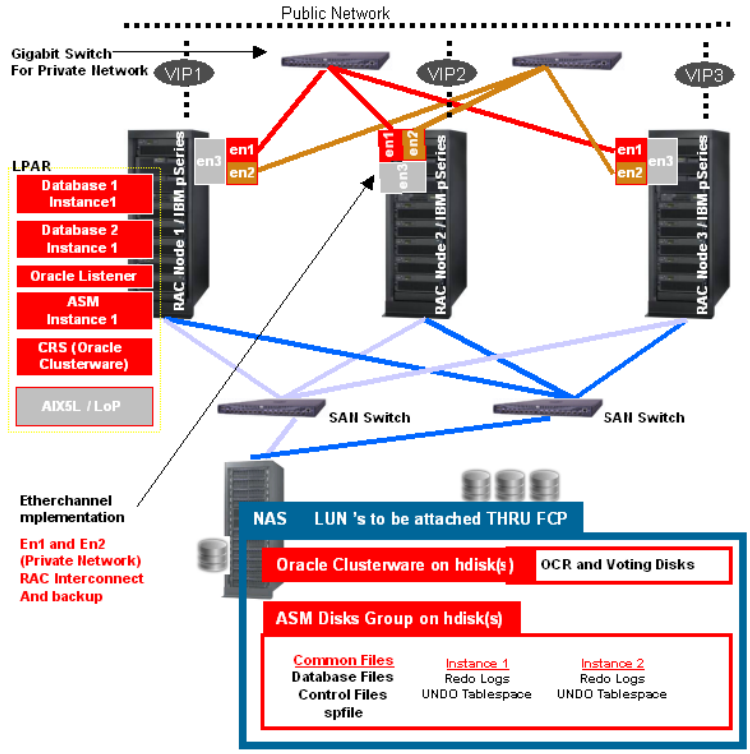
5.1.4 NAS (Network Appliance Filer)

10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s) and database on Network Appliance Filers in NFS mode.

Case 1 : NFS implementation



Case 2 : FCP implementation (as standard SAN)



!!! ONLY
With 10g RAC R2 !!!

Network Appliance Solution :

➔ No need for HACMP

➔ All files (Oracle database files, CRS files with OCR and Voting, any binaries, any files) can be stored on NAS filer.

➔ **Specific Cookbook from NetApp : Oracle Real Application Clusters 10g™ Release 2 with AIX 5L Version 5.3 and NetApp Storage Available on TR-3542**

ORACLE CLUSTERWARE MANDATORY !!!

Using IBM N Series Storage is same as NetApp NAS Storage.
IBM N Series are NetApp OEM.

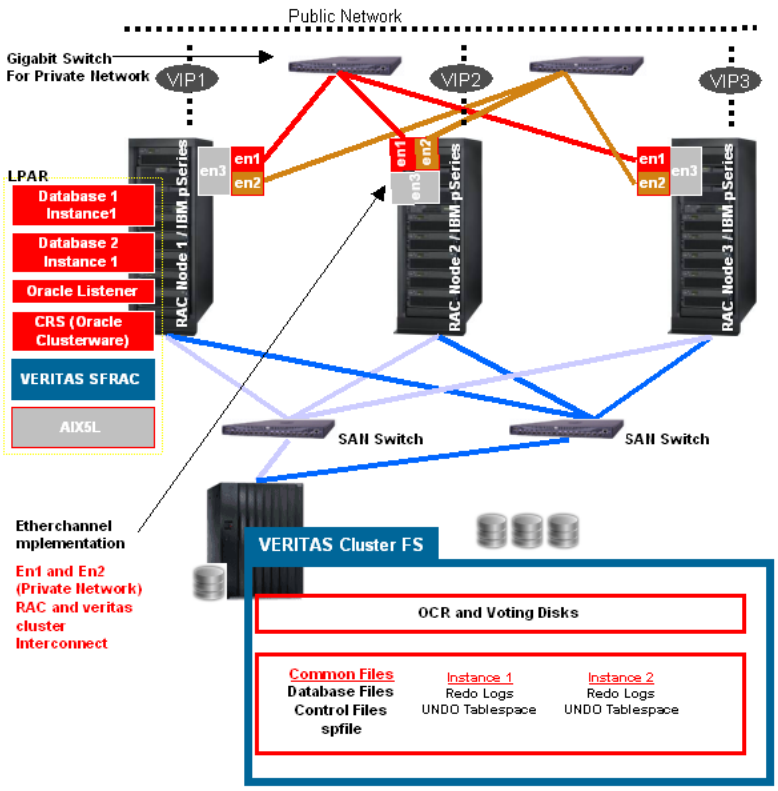
A NAS Storage can also be used as a traditional SAN, using FCP.

RAC on iSCSI with AIX5L is not supported as for today.

NetApp/Oracle RAC Supportability Matrix :
http://www.netapp.com/partners/oracle/oracle_RAC_certification_matrix.pdf

5.1.5 Veritas SFRAC

Veritas Storage Foudation for Oracle RAC, OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s) and database on the Veritas Cluster Files System …



VERITAS Cluster Files System Solution :

➔ **Veritas clusterware does not replace Oracle Clusterware**, it's only used for the cluster files system part.

➔ No need for HACMP

➔ Veritas SFRAC Does NOT cohabitate with HACMP

➔ All files (Oracle database files, CRS files with OCR and Voting, any binaries, any files) can be stored on Veritas cluster files system.

ORACLE CLUSTERWARE MANDATORY !!!

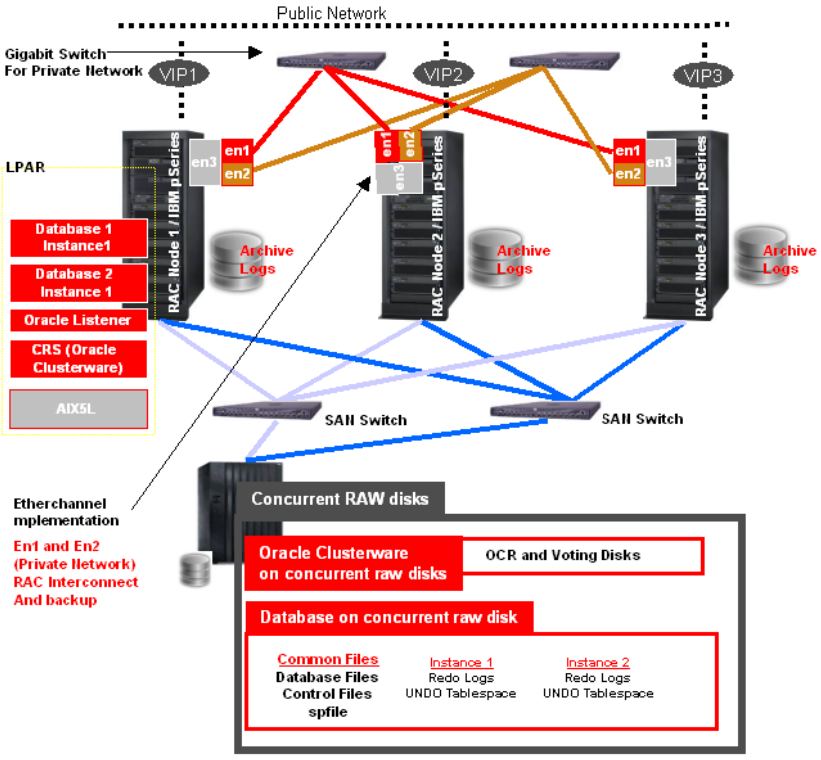
Subject: **Using Oracle Clusterware with Vendor Clusterware FAQ** Doc ID: **Note:332257.1**

Veritas Storage Foundation for RAC (SF-RAC) Supportability Matrix :
http://ftp.support.veritas.com/pub/support/products/DBE_Advanced_Cluster_for_Oracle_RAC/280186.pdf

➔ **Specific Cookbook from IBM :**
OracleOracle 10g RAC on AIX with Veritas Storage Foundation for Oracle RAC.
Available on IBM-WP100930

5.1.6 Concurrent Raw Devices without HACMP

10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s), and database on concurrent raw devices without HACMP, concurrent disks/raws will be prepares as done for ASM.



Concurrent raw devices Solution without HAMCP :

➔ **HACMP Is not necessary, even to provide concurrent raw devices.**

We'll have

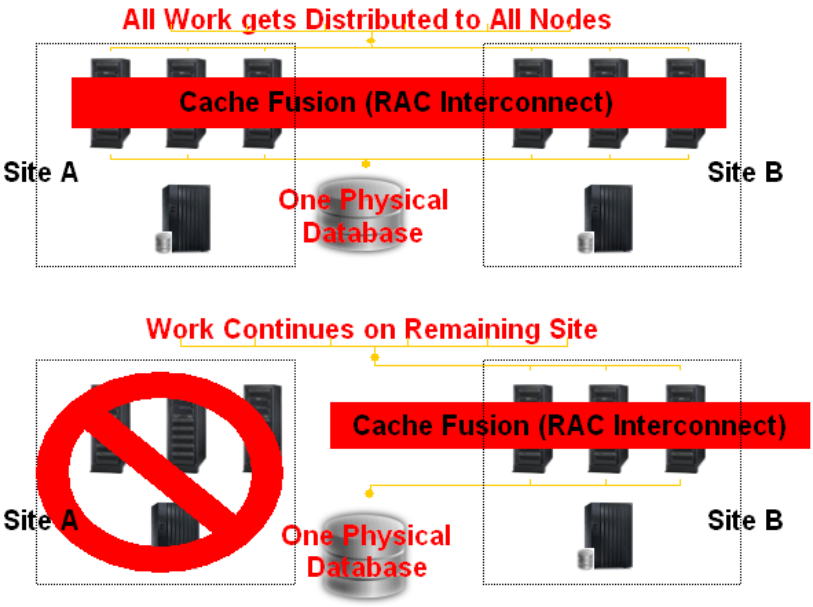
- **No Volume Group,**
- **No Logical volumes,**

ONLY concurrent raw disks using virtual devices.

➔ Only Oracle databases files (datafiles, redo logs, spfiles, CRS files with OCR and Voting) are stored on the concurrent raw devices. No archive logs, No binaries.

ORACLE CLUSTERWARE MANDATORY !!!

5.2 10G RAC over MAN (Metropolitan Area Network)



Architecture type known as :

- **Oracle RAC Extended Cluster**

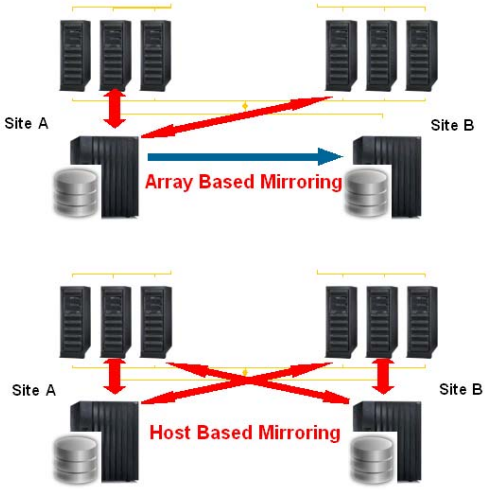
Or

- **Oracle RAC Stretched Cluster**

Customers are already in production with 10gRAC R2 using "Host based mirroring" or "Aray based mirroring" to maintain copies of data accros both storages.

Please contact us at oraclibm@fr.ibm.com to discuss about such architecture.

RAC Extended Architecture : Array or Host based mirroring ?



Architecture Design with key points to solve considering Customer requirements.

Main key points :

- Distance between site, quality of existing network infrastructure (DWDM needed or Not ?)
- Array or Host based Mirroring
- Which Cluster files system ?
- Disaster scenarios and Split Brain Situation to cover
- Level of downtime supported
- Costs, etc ….

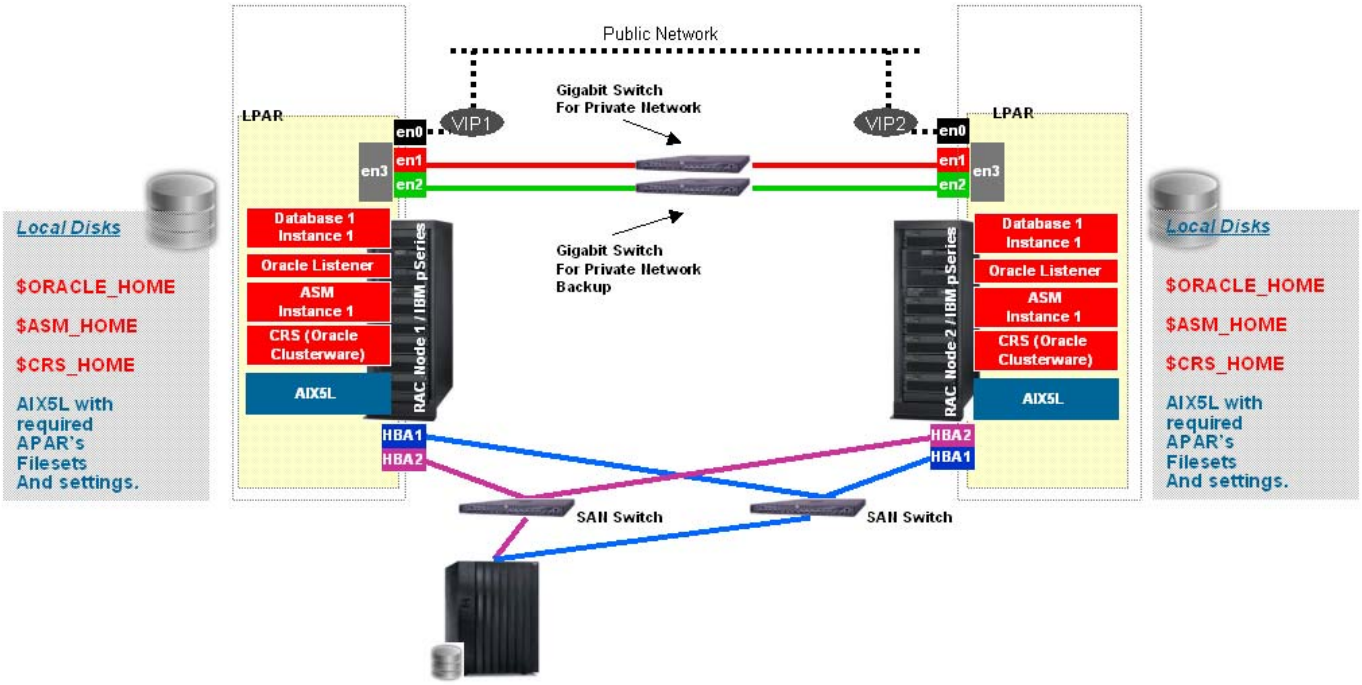
➔ **Oracle Real Application Clusters on Extended Distance Clusters(PDF)** November 2006
http://www.oracle.com/technology/products/database/clustering/pdf/ExtendedRAC10gR2.pdf

➔ **The Right Choice for Disaster Recovery: Data Guard, Stretch Clusters or Remote Mirroring**
Various solutions are available today to protect business-critical data, and enable enterprises to quickly restore their business operations in the event of outages or disasters. This paper discusses three such technologies - Oracle Data Guard in a Maximum Availability Architecture (MAA) configuration (i.e. Data Guard combined with Oracle Real Application Clusters (RAC)), Stretch Clusters based on RAC, and Remote Mirroring. It describes their capabilities using a DR solution assessment framework, and makes best practice recommendations regarding their applicability to various business situations. Presentation
http://www.oracle.com/technology/deploy/availability/pdf/DRChoices_TWP.pdf

6 HARDWARE ARCHITECTURE

For our infrastructure, we used a cluster which is composed of three partitions (IBM LPAR) on an IBM @server system p 570 using AIX 5L.

BUT in the real world, to achieve true high availability it's necessary to have at least two IBM Systems p / i servers as shown bellow :

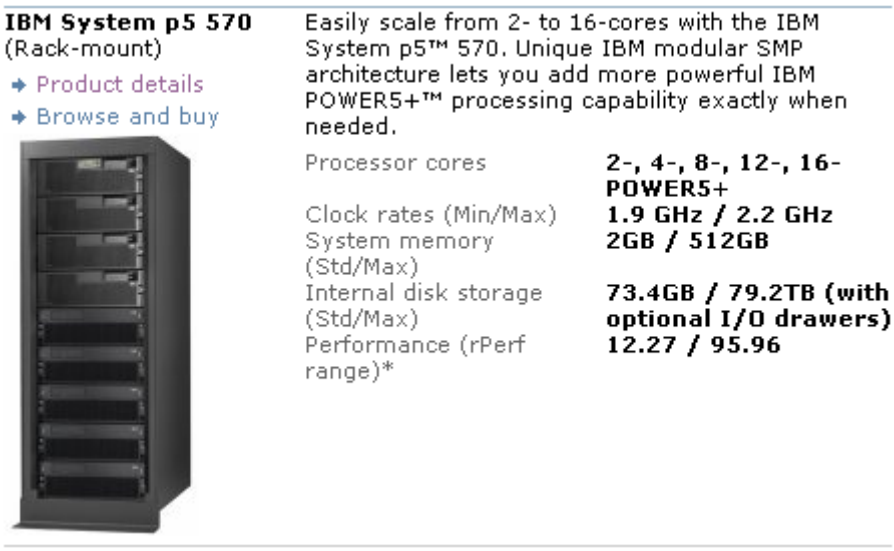


Each component of the infrastructure must be protected :

- Disk access path (2 HBA's and multi-pathing software)
- Interconnect network
- Public network
- …

6.1 IBM SYTEM P SERVERS

This is the IBM Sytem pServer we used for our installation …



http://www-03.ibm.com/servers/eserver/pseries/hardware/highend/

http://www-03.ibm.com/systems/p/

THEN you'll need 1 AIX5L LPAR on each server for real RAC implementation, with necessary memory and Power5 CPU assigned to each LPAR.

Commands to print the config for IBM System p on AIX5L :

```
{node1:root}/ -> prtconf
System Model: IBM,9117-570
Machine Serial Number: 10ED22C
Processor Type: PowerPC_POWER5
Number Of Processors: 2
Processor Clock Speed: 1654 MHz
CPU Type: 64-bit
Kernel Type: 64-bit
LPAR Info: 5 node1
Memory Size: 2048 MB
Good Memory Size: 2048 MB
Platform Firmware level: SF235_180
Firmware Version: IBM,SF235_180
Console Login: enable
Auto Restart: true
Full Core: false

...
```

...

Network Information

Host Name: node1
IP Address: 10.3.25.81
Sub Netmask: 255.255.255.0
Gateway: 10.3.25.254
Name Server:
Domain Name:

Paging Space Information

Total Paging Space: 1536MB
Percent Used: 16%

Volume Groups Information

```
==============================================================================
rootvg:
PV_NAME           PV STATE          TOTAL PPs   FREE PPs    FREE DISTRIBUTION
hdisk0            active            546         317         99..00..00..109..109
==============================================================================
```

INSTALLED RESOURCE LIST

The following resources are installed on the machine.
+/- = Added or deleted from Resource List.
* = Diagnostic support not available.

Model Architecture: chrp
Model Implementation: Multiple Processor, PCI bus

```
+ sys0                                                  System Object
+ sysplanar0                                            System Planar
* vio0                                                  Virtual I/O Bus
* vscsi0          U9117.570.10ED22C-V5-C5-T1                    Virtual SCSI Client Adapter
* hdisk0          U9117.570.10ED22C-V5-C5-T1-L810000000000      Virtual SCSI Disk Drive
* ent2            U9117.570.10ED22C-V5-C4-T1                    Virtual I/O Ethernet Adapter (l-lan)
* ent1            U9117.570.10ED22C-V5-C3-T1                    Virtual I/O Ethernet Adapter (l-lan)
* vsa0            U9117.570.10ED22C-V5-C0                       LPAR Virtual Serial Adapter
* vty0            U9117.570.10ED22C-V5-C0-L0                    Asynchronous Terminal
* ent0            U9117.570.10ED22C-V5-C2-T1                    Virtual I/O Ethernet Adapter (l-lan)
* pci1            U7879.001.DQD01JK-P1                          PCI Bus
* pci3            U7879.001.DQD01JK-P1                          PCI Bus
+ fcs0            U7879.001.DQD01JK-P1-C2-T1                    FC Adapter
* fcnet0          U7879.001.DQD01JK-P1-C2-T1                    Fibre Channel Network Protocol Device
* fscsi0          U7879.001.DQD01JK-P1-C2-T1                    FC SCSI I/O Controller Protocol Device
* dac0            U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404  3552  (500) Disk Array Controller
* dac3            U7879.001.DQD01JK-P1-C2-T1-W200300A0B80C5404  3552  (500) Disk Array Controller
+ L2cache0                                               L2 Cache
+ mem0                                                  Memory
+ proc0                                                 Processor
+ proc2                                                 Processor
+ hdisk1          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L0             3552  (500) Disk Array Device
+ hdisk2          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L1000000000000 3552  (500) Disk Array Device
+ hdisk3          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L2000000000000 3552  (500) Disk Array Device
+ hdisk4          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L3000000000000 3552  (500) Disk Array Device
+ hdisk5          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L4000000000000 3552  (500) Disk Array Device
+ hdisk6          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L5000000000000 3552  (500) Disk Array Device
+ hdisk7          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L6000000000000 3552  (500) Disk Array Device
+ hdisk8          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L7000000000000 3552  (500) Disk Array Device
+ hdisk9          U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L8000000000000 3552  (500) Disk Array Device
+ hdisk10         U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L9000000000000 3552  (500) Disk Array Device
+ hdisk11         U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LA000000000000 3552  (500) Disk Array Device
+ hdisk12         U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LB000000000000 3552  (500) Disk Array Device
+ hdisk13         U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LC000000000000 3552  (500) Disk Array Device
{node1:root}/ ->
```

Command to get information on the LPAR :

```
{node1:root}/ -> lparstat -i
Node Name                                  : node1
Partition Name                             : node1
Partition Number                           : 5
Type                                       : Shared-SMT
Mode                                       : Uncapped
Entitled Capacity                          : 0.20
Partition Group-ID                         : 32773
Shared Pool ID                             : 0
Online Virtual CPUs                        : 2
Maximum Virtual CPUs                       : 4
Minimum Virtual CPUs                       : 1
Online Memory                              : 2048 MB
Maximum Memory                             : 8192 MB
Minimum Memory                             : 1024 MB
Variable Capacity Weight                   : 128
Minimum Capacity                           : 0.10
Maximum Capacity                           : 2.00
Capacity Increment                         : 0.01
Maximum Physical CPUs in system            : 4
Active Physical CPUs in system             : 4
Active CPUs in Pool                        : 4
Unallocated Capacity                       : 0.00
Physical CPU Percentage                    : 10.00%
Unallocated Weight                         : 0
{node1:root}/ ->
```

6.2 OPERATING SYSTEM

Operating system must be installed the same way on each LPAR, with the same maintenance level, same APAR and FILESETS level.

AS Cluster,
RAC Nodes must have :

- **Same Processor Technology**
  - Ex : Power 5 (in our case)
  - From p505 to P595, including Blades Center with JS20/21
- **Same Operating System**
  - Ex : AIX5L 5.3
- **Same Maintenance Level**
  - Ex : TL5 SP3
- **Same APAR (s) / Filesets (s)**

➔ Check "PREPARING THE SYSTEM" chapter for Operating System requirements on AIX5L

- AIX5.1 is not supported
- **AIX5.2 / 5.3 are supported and certified**

↙ The IBM AIX clustering layer, HACMP filesets, MUST NOT be installed if you've chosen an implementation without HACMP. If this layer is implemented for other purpose, disks ressources necessary to install and run CRS data will have to be part of an HACMP volume group resource.

If you have previously installed HACMP, you must remove :

- HACMP filesets (cluster.es.*)
- rsct.hacmp.rte
- rsct.compat.basic.hacmp.rte
- rsct.compat.clients.hacmp.rte

If you did run a first installation of the Oracle Clusterware (CRS) with HACMP installed, ➔Check if /opt/ORCLcluster directory does exist and if so, remove it on all nodes.

THEN REBOOT ALL NODES …

6.3 NETWORK INFRASTRUCTURE

A private network (for instance a gigabit ethernet network, using a gigabit switch to link each cluster nodes) is designed only for Oracle interconnect use (cache fusion between instances). This dedicated network is mandatory.

Gigabit switch is mandatory for production implementation, even for only 2 nodes architecture.

(Cross-over cable can be used only for test purpose, and it's not supported by Oracle Support, please read RAC FAQ on http://metalink.oracle.com).

A second gigabit ethernet interconnect, with a different network mask, can be setup for security purposes or performance issues.

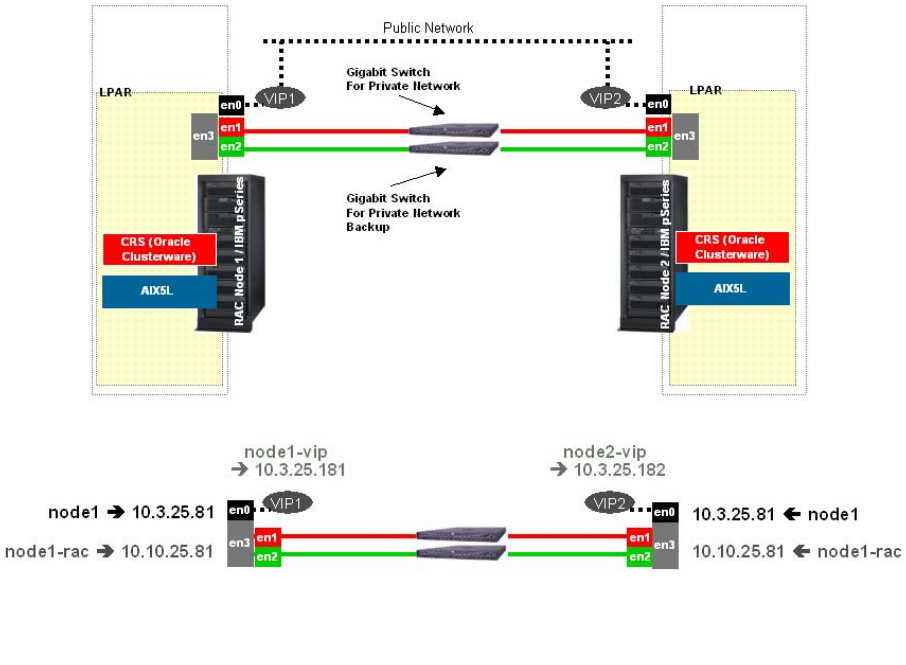
6.3.1 Standard Network Architecture

• Network cards for public network must have same name on each participating node in the RAC cluster.

• Network cards for Interconnect Network (Private) must have same Name on each participating Node in the RAC cluster.

1 virtual IP per node must be reserved, and not used on the network prior to Oracle clusterware installation. Don't set IP allias at AIX5L level, Oracle clusterware will take charge of it.

Network Layout for RAC/ASM Implementation Using AIX Etherchannel !!!



6.3.2 Network Architecture with AIX5L Etherchannel

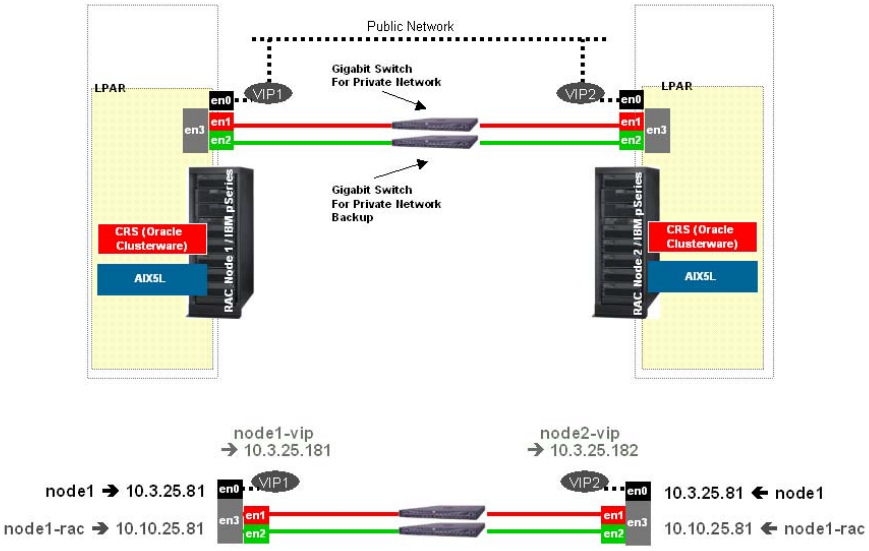
EtherChannel is supported to be used with RAC on AIX.

EtherChannel provide :

- network bandwith aggregation.
- failover.

This is implemented at AIX level and fully transparent to oracle RAC.

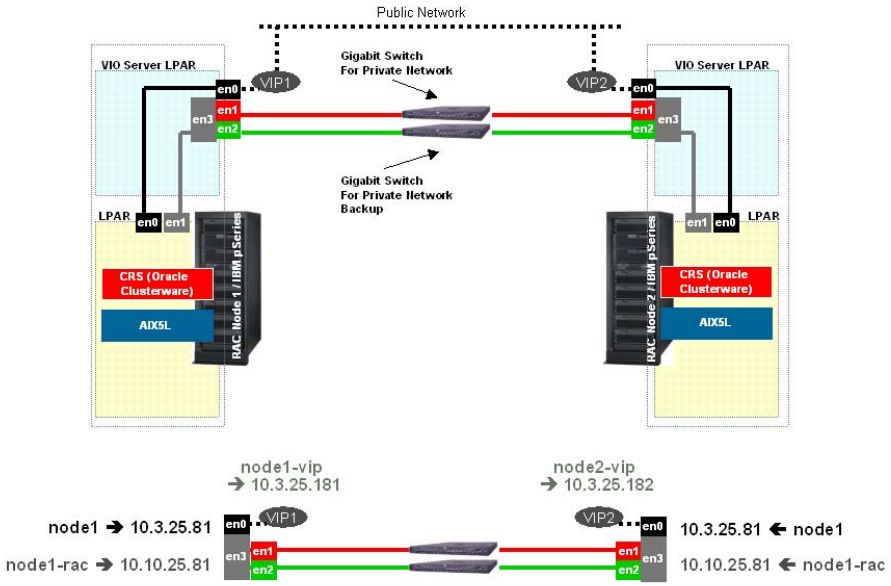
Network Layout for RAC/ASM Implementation Using AIX Etherchannel !!!



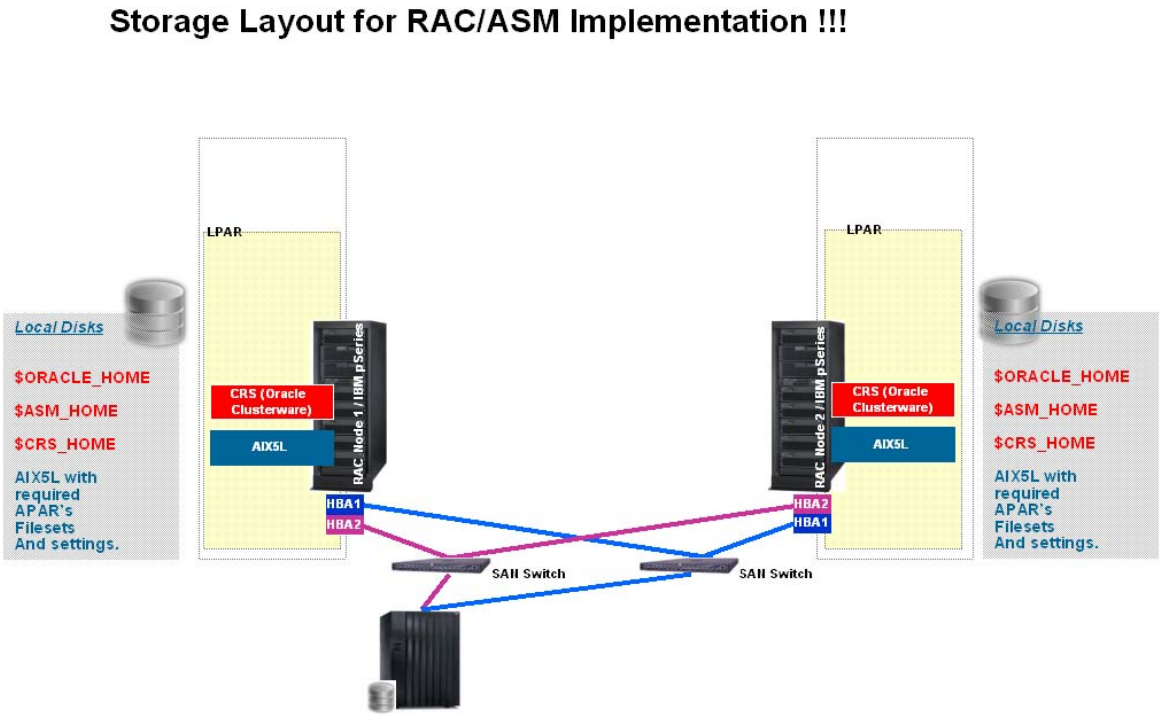
6.3.3 Network Architecture with AIX5L Virtual I/O Server

AIX Virtual I/O network feature is usable to implement RAC network when possible.

Network Layout for RAC/ASM Implementation Using AIX Etherchannel and VIO Server !!!



6.4 SAN Storage



When implementing RAC, you must be carrefull on the SAN storage to use. The SAN Storage must be capable thru it's drivers of read/write concurrency at same time from any member of the RAC cluster, which means that "reserve_policy" attribute from disks (hdisk, hdiskpower, dlmfdrv, etc …) discovered must be able to be set to "no_reserve" or "no_lock" values.

6.4.1 SAN and Virtual I/O Server

You can use virtual I/O disks for :

- AIX5L operating system
- Oracle clusterware ($CRS_HOME)
- RAC Software ($ORACLE_HOME)

BUT NOT to be used for :

- **OCR (Oracle Cluster Registry) disk,**
- **Voting (Heartbeat) disk,**

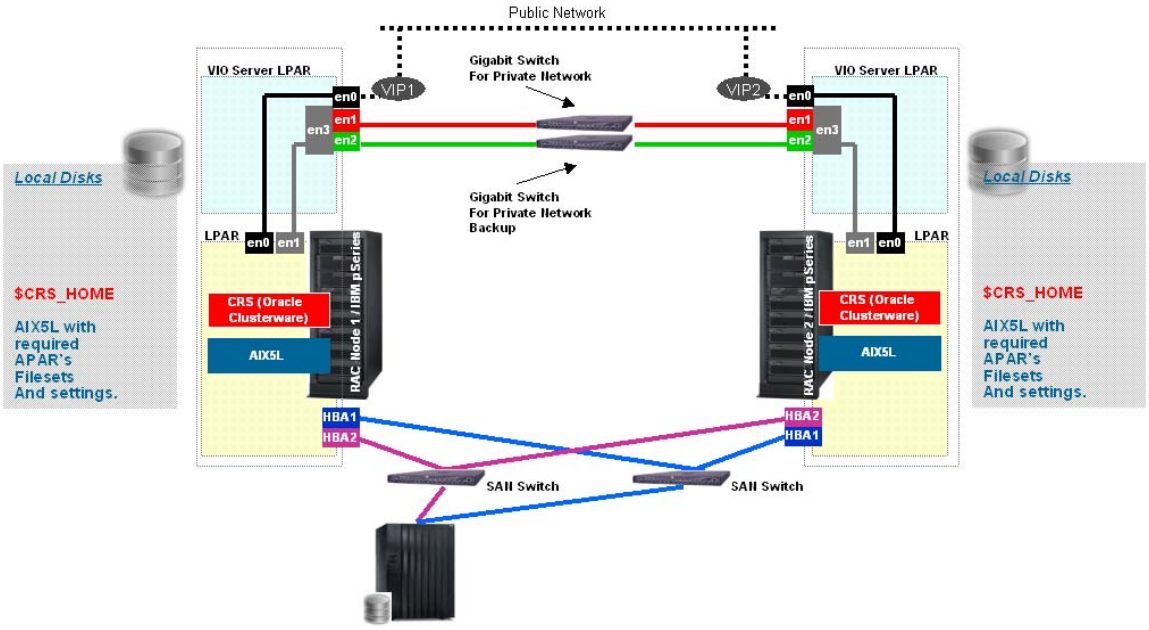
and

- **database files (datafiles, redo logs, archive logs, UNDO, TEMP, etc….)**

HBA Fiber Channel cards MUST NOT be shared thru Virtual I/O server, as for today (check certification status on Oracle Metalink).

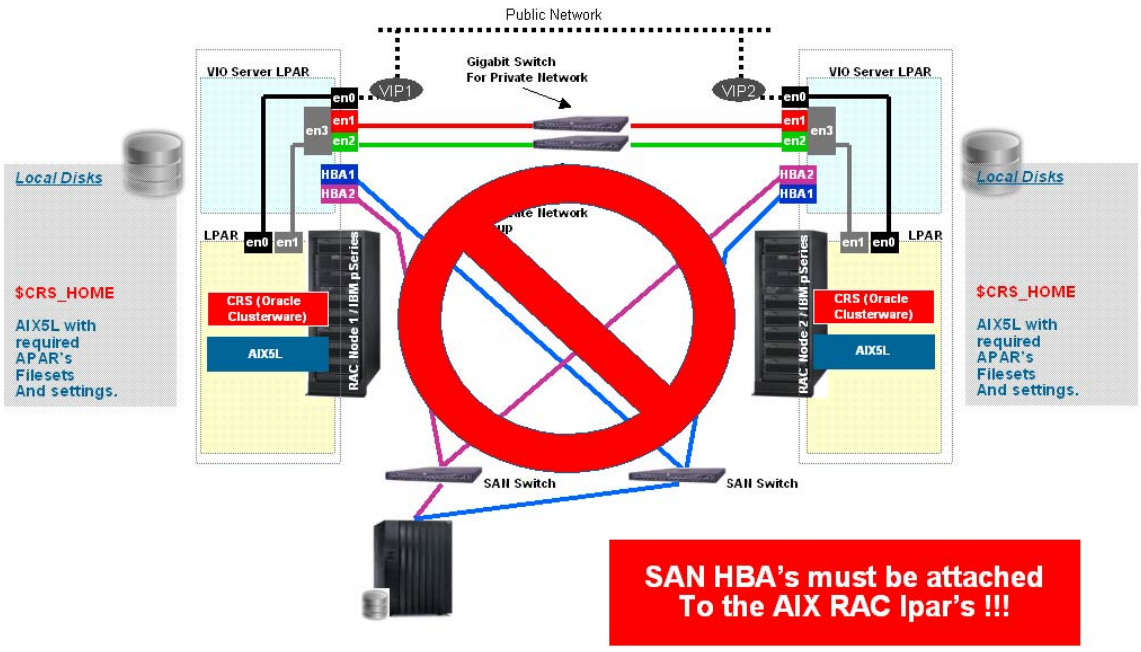
AIX Virtual I/O network feature is usable to implement RAC network when possible.

Supported Storage Layout for RAC/ASM Implementation Using AIX Etherchannel and VIO Server !!!



AIX Virtual I/O network feature is usable to implement RAC network when possible.

Non Supported Storage Layout for RAC/ASM Implementation Using AIX Etherchannel and VIO Server !!!



6.4.2 Multi-pathing and ASM

Please check Metalink note "Oracle ASM and Multi-Pathing Technologies" Doc ID: **Note:294869.1**

Note, that Oracle Corporation does not certify ASM against multipathing utilities. The MP utilities listed below are ones that known working solutions. As we do more testing, additional MP utilities will be listed here, thus, this document is an active document.

Multi-pathing allow SAN access failover, and load balancing accros SAN Fiber Channel attachements.

OS Platform	Multi-pathing tool	ASM Device Usage	Notes
AIX	**EMC PowerPath**	Use raw partitions thru the pseudo device /dev/rhdiskpowerx	
	IBM SDD (Vpath) **NOT SUPPORTED for RAC/ASM on AIX !!!**	/dev/vpathx	As of this writing, SDD-AIX is known to cause discovery and device handling problems for ASM, and thus is not viable solution. ASM needs to access disks/vpath thru non root user, which is not allowed by SDD as for today 27 March 2007. See SDDPCM section below for an alternative solution to SDD for AIX
	IBM SDDPCM	Use /dev/rhdiskx device	You must install SDDPCM filesets and enable SDDPCM.. SDDPCM cannot co-exist w/ SDD. SDDPCM only works with the following IBM storage components: DS8000,DS6000,Enterprise Storage Server (ESS) SDDPCM works also on top of IBM SVC (SAN Volume Controler), and on top of other supported storages like HDS, EMC, etc …
	IBM RDAC (Redundant Disk Array Controller)	Use /dev/rhdiskx device	RDAC is installed by default and must be used with IBM storage DS4000, and former FasTt series.
	Hitachi Dynamic Link Manager - HDLM	Use /dev/rdsk/cxtydz that's generated by HDLM Or /dev/dlmfdrvx	HDLM generates a scsi (cxtydzx) address where the controller is highest unused controller number. HDLM no longer requires HACMP. /dev/dlmfdrvx can be used out of HDLM Volume Group Or if using HDLM Volume Group, logical volumes must be created with "mklv" command using "-T O" options.
	Fujitsu ETERNUS GR Multipath Driver	Use /dev/rhdisk device	

6.4.3 IBM storage and multi-pathing

With IBM, please refer to IBM to confirm which IBM storage is supported with RAC, if not specified in our document.

IBM TotalStorage products for IBM Sytem p

IBM DS4000, DS6000 and DS8000 series are supported with 10gRAC.

IBM Storage DS300 and DS400 are not, and will not be supported with 10gRAC.

As for today March 27, 2007 IBM Storage DS3200 and DS3400 are not yet supported with 10gRAC.

IBM System Storage and TotalStorage products

➔ http://www-03.ibm.com/servers/storage/product/products_pseries.html

There are 2 cases when using IBM storage :

- IBM MPIO (Multi-Path I/O).
  MPIO driver is supported with IBM Total Storage ESS, DS6000 and DS8000 series only
  And with IBM SVC (SAN Volume Controler).

- IBM RDAC (Redundant Disk Array Controller) for IBM Total Storage DS4000.
  RDAC driver is supported with IBM Total Storage DS4000 series only, and former FasTt.

  ➔ You MUST use one or the other, depending on the storage used.

- **case 1: Lun's provided by the IBM storage with IBM MPIO installed as multi-pathing driver.**

  Disks (LUN's) will be seen as hdisk at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
…
```

- **case 2: Lun's provided by the IBM DS4000 storage with IBM RDAC installed as multi-pathing driver.**

  Disks (LUN's) will be seen as hdisk at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
…
```

6.4.3.1 IBM MPIO (Multi Path I/O) Setup Procedure

AIX Packages needed to install on all nodes :

```
devices.sddpcm.53.2.1.0.7.bff
devices.sddpcm.53.rte
devices.fcp.disk.ibm.mpio.rte
```

devices.fcp.disk.ibm.mpio.rte download page :
http://www-1.ibm.com/support/docview.wss?rs=540&context=ST52G7&dc=D410&q1=host+scripts&uid=ssg1S4000203&loc=en_US&cs=utf-8&lang=en

MPIO for AIX 5.3 download page :
http://www-1.ibm.com/support/docview.wss?uid=ssg1S4000201

On node 1, and node 2 …

Installing the filesets :

```
       smitty install

            Install and Update Software
                  Install Software

     * INPUT device / directory for software
[/mydir_with_my_filesets]

            SOFTWARE to install                                  []
                  Press F4

                  Select devices.fcp.disk.ibm.mpio


                                  Install Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                                  [Entry Fields]
* INPUT device / directory for software                .
* SOFTWARE to install
[devices.fcp.disk.ibm.> +
  PREVIEW only? (install operation will NOT occur)    no
+
  COMMIT software updates?                            yes
+
  SAVE replaced files?                                no
+
  AUTOMATICALLY install requisite software?           yes
+
  EXTEND file systems if space needed?                yes
+
  OVERWRITE same or newer versions?                   no
+
  VERIFY install and check file sizes?                no
+
  Include corresponding LANGUAGE filesets?            yes
+
  DETAILED output?                                    no
+
  Process multiple volumes?                           yes
+
  ACCEPT new license agreements?                      no
+
  Preview new LICENSE agreements?                     no
+
```

Check the insatllation succed and the installation summary message :

```
Installation Summary
--------------------
Name                        Level           Part        Event       Result
-------------------------------------------------------------------------------
devices.fcp.disk.ibm.mpio.r 1.0.0.0         USR         APPLY       SUCCESS


Install devices.sddpcm.53

Select :

  ¦   devices.sddpcm.53
ALL ¦
  ¦    + 2.1.0.0  IBM SDD PCM for AIX V53
¦
  ¦    + 2.1.0.7  IBM SDD PCM for AIX V53
¦


                                Install Software

Type or select values in entry fields.
Press Enter AFTER making all desired changes.


* INPUT device / directory for software                .
* SOFTWARE to install                                  [devices.sddpcm.53
      > +
  PREVIEW only? (install operation will NOT occur)     no
+
  COMMIT software updates?                             yes
+
  SAVE replaced files?                                 no
+
  AUTOMATICALLY install requisite software?            yes
+
  EXTEND file systems if space needed?                 yes
+
  OVERWRITE same or newer versions?                    no
+
  VERIFY install and check file sizes?                 no
+
  Include corresponding LANGUAGE filesets?             yes
+
  DETAILED output?                                     no
+
  Process multiple volumes?                            yes
+
  ACCEPT new license agreements?                       no
+
  Preview new LICENSE agreements?                      no
+


Installation Summary
--------------------
Name                        Level           Part        Event       Result
-------------------------------------------------------------------------------
devices.sddpcm.53.rte       2.1.0.0         USR         APPLY       SUCCESS
devices.sddpcm.53.rte       2.1.0.0         ROOT        APPLY       SUCCESS
devices.sddpcm.53.rte       2.1.0.7         USR         APPLY       SUCCESS
devices.sddpcm.53.rte       2.1.0.7         ROOT        APPLY       SUCCESS
devices.sddpcm.53.rte       2.1.0.7         USR         COMMIT      SUCCESS
devices.sddpcm.53.rte       2.1.0.7         ROOT        COMMIT      SUCCESS
```

Now you need to reboot all AIX nodes !!!

Commands

to know AIX WWPN

```
{node1:root}/ -> pcmpath query wwpn
Adapter Name    PortWWN
fscsi0          10000000C935A7E7
fscsi1          10000000C93A1BF3
```

Commands

to check disks :

```
{node1:root}/ -> lsdev -Cc disk -t 2107

hdisk2  Available 0A-08-02 IBM MPIO FC 2107
hdisk3  Available 0A-08-02 IBM MPIO FC 2107
hdisk4  Available 0A-08-02 IBM MPIO FC 2107
hdisk5  Available 0A-08-02 IBM MPIO FC 2107
hdisk6  Available 0A-08-02 IBM MPIO FC 2107
hdisk7  Available 0A-08-02 IBM MPIO FC 2107
hdisk8  Available 0A-08-02 IBM MPIO FC 2107
hdisk9  Available 0A-08-02 IBM MPIO FC 2107
hdisk10 Available 0A-08-02 IBM MPIO FC 2107
hdisk11 Available 0A-08-02 IBM MPIO FC 2107
```

Commands

to check disks :

```
{node1:root}/ -> pcmpath query device

DEV#:   2  DEVICE NAME: hdisk2  TYPE: 2107900  ALGORITHM:  Load Balance
SERIAL: 75271812000
==========================================================================
Path#      Adapter/Path Name          State     Mode     Select     Errors
    0           fscsi0/path0          CLOSE   NORMAL          0          0
    1           fscsi1/path1          CLOSE   NORMAL          0          0

DEV#:   3  DEVICE NAME: hdisk3  TYPE: 2107900  ALGORITHM:  Load Balance
SERIAL: 75271812001
==========================================================================
Path#      Adapter/Path Name          State     Mode     Select     Errors
    0           fscsi0/path0          CLOSE   NORMAL          0          0
    1           fscsi1/path1          CLOSE   NORMAL          0          0


DEV#:   4  DEVICE NAME: hdisk4  TYPE: 2107900  ALGORITHM:  Load Balance
SERIAL: 75271812002
==========================================================================
Path#      Adapter/Path Name          State     Mode     Select     Errors
    0           fscsi0/path0          CLOSE   NORMAL          0          0
    1           fscsi1/path1          CLOSE   NORMAL          0          0
```

6.4.3.2 IBM AIX RDAC (FCP.ARRAY filesets) Setup Procedure

This ONLY apply to use of of DS4000 storage series. NOT to DS6000, DS8000 and ES800.

RDAC is installed by default on AIX5L.
Each node must have 2 HBA cards, for multi-pathing. With ONLY 1 HBA per node, it will works but path to SAN will not be protected. THEN in production, 2 HBA per node must be used.

All AIX hosts in your storage subsystem must have the RDAC multipath driver installed.
In a single server environment, AIX allows load sharing (also called *load balancing*). You can set the load balancing parameter to yes. In case of heavy workload on one path the driver will move other LUNs to the controller with less workload and, if the workload reduces back to the preferred controller. Problem that can occur is disk thrashing. That means that the driver moves the LUN back and forth from one controller to the other. As a result the controller is more occupied by moving disks around than servicing I/O. The recommendation is to NOT load balance on an AIX system. The performance increase is minimal (or performance could actually get worse).

RDAC (fcp.array filesets) for AIX support round-robin load-balancing

Setting the attributes of the RDAC driver for AIX
The AIX RDAC driver files are not included on the DS4000 installation CD.
Either install them from the AIX Operating Systems CD, if the correct version is included, or download them from the following Web site: http://techsupport.services.ibm.com/server/fixes
or http://www-304.ibm.com/jct01004c/systems/support/

Commands

to check that necessary filesets are present for RDAC …

```
{node1:root}/ -> lslpp -L devices.fcp.disk.array.rte
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  devices.fcp.disk.array.rte
                             5.3.0.52    A     F    FC SCSI RAIDiant Array Device
                                                    Support Software


State codes:
 A -- Applied.
 B -- Broken.
 C -- Committed.
 E -- EFIX Locked.
 O -- Obsolete.  (partially migrated to newer version)
 ? -- Inconsistent State...Run lppchk -v.

Type codes:
 F -- Installp Fileset
 P -- Product
 C -- Component
 T -- Feature
 R -- RPM Package
{node1:root}/ ->

{node1:root}/ -> lslpp -L devices.common.IBM.fc.rte
  Fileset                      Level  State  Type  Description (Uninstaller)
  ----------------------------------------------------------------------------
  devices.common.IBM.fc.rte
                             5.3.0.50    C     F    Common IBM FC Software

State codes:
 A -- Applied.
 B -- Broken.
 C -- Committed.
 E -- EFIX Locked.
 O -- Obsolete.  (partially migrated to newer version)
 ? -- Inconsistent State...Run lppchk -v.

Type codes:
 F -- Installp Fileset
 P -- Product
 C -- Component
 T -- Feature
 R -- RPM Package
{node1:root}/ ->
```

Commands

to check the RDAC configuration and HBA path to hdisk …

On node1 …

```
{node1:root}/ -> fget_config -v -A

---dar0---

User array name = 'DS4000_JSC'
dac0 ACTIVE dac3 ACTIVE

Disk      DAC   LUN Logical Drive
hdisk1    dac0    0 G8_spfile
hdisk2    dac3    1 G8_OCR1
hdisk3    dac0    2 G8_OCR2
hdisk4    dac3    3 G8_Vote1
hdisk5    dac0    4 G8_Vote2
hdisk6    dac3    5 G8_Vote3
hdisk7    dac0    6 G8_Data1
hdisk8    dac3    7 G8_Data2
hdisk9    dac0    8 G8_Data3
hdisk10   dac3    9 G8_Data4
hdisk11   dac0   10 G8_Data5
hdisk12   dac3   11 G8_Data6
hdisk13   dac0   12 G8_tie
{node1:root}/ ->
```

On node2 …

```
{node2:root}/ -> fget_config -v -A

---dar0---

User array name = 'DS4000_JSC'
dac0 ACTIVE dac3 ACTIVE

Disk      DAC   LUN Logical Drive
hdisk0    dac0    0 G8_spfile
hdisk1    dac3    1 G8_OCR1
hdisk2    dac0    2 G8_OCR2
hdisk3    dac3    3 G8_Vote1
hdisk4    dac0    4 G8_Vote2
hdisk5    dac3    5 G8_Vote3
hdisk6    dac0    6 G8_Data1
hdisk7    dac3    7 G8_Data2
hdisk8    dac0    8 G8_Data3
hdisk9    dac3    9 G8_Data4
hdisk10   dac0   10 G8_Data5
hdisk12   dac3   11 G8_Data6
hdisk13   dac0   12 G8_tie
{node2:root}/ ->
```

Commands

to check the RDAC configuration and HBA path to hdisk for one specific "dar" …

```
{node1:root}/ -> fget_config -l dar0
dac0 ACTIVE dac3 ACTIVE
hdisk1    dac0
hdisk2    dac3
hdisk3    dac0
hdisk4    dac3
hdisk5    dac0
hdisk6    dac3
hdisk7    dac0
hdisk8    dac3
hdisk9    dac0
hdisk10   dac3
hdisk11   dac0
hdisk12   dac3
hdisk13   dac0
{node1:root}/ ->
```

Fcs0 is one of the HBA.

Commands

To see the HBA fiber channel statistics :

```
{node1:root}/ -> fcstat fcs0

FIBRE CHANNEL STATISTICS REPORT: fcs0

Device Type: FC Adapter (df1080f9)
Serial Number: 1F41709923
Option ROM Version: 02E01871
Firmware Version: H1D1.81X1
World Wide Node Name: 0x20000000C93F8E29
World Wide Port Name: 0x10000000C93F8E29

FC-4 TYPES:
  Supported:
0x0000012000000000000000000000000000000000000000000000000000000000
  Active:
0x0000010000000000000000000000000000000000000000000000000000000000
Class of Service: 3
Port Speed (supported): 2 GBIT
Port Speed (running):   2 GBIT
Port FC ID: 0x650B00
Port Type: Fabric

Seconds Since Last Reset: 2795

        Transmit Statistics     Receive Statistics
        -------------------     ------------------
Frames: 41615                   96207
Words:  1537024                 12497408

LIP Count: 0
NOS Count: 0
Error Frames:  0
Dumped Frames: 0
Link Failure Count: 269
Loss of Sync Count: 469
Loss of Signal: 466
Primitive Seq Protocol Error Count: 0
Invalid Tx Word Count: 51
Invalid CRC Count: 0

IP over FC Adapter Driver Information
  No DMA Resource Count: 0
  No Adapter Elements Count: 0

FC SCSI Adapter Driver Information
  No DMA Resource Count: 0
  No Adapter Elements Count: 0
  No Command Resource Count: 0

IP over FC Traffic Statistics
  Input Requests:   0
  Output Requests:  0
  Control Requests: 0
  Input Bytes:  0
  Output Bytes: 0

FC SCSI Traffic Statistics
  Input Requests:   24721
  Output Requests:  8204
  Control Requests: 252
  Input Bytes:  46814436
  Output Bytes: 4207616
{node1:root}/ ->
```

6.4.4 EMC storage and multi-pathing

With EMC, please refer to Hitachi to see which EMC storage is supported with RAC.

There are 2 cases when using EMC storage :

- **case 1: Lun's provided by the EMC storage with IBM MPIO installed as multi-pathing driver.**

  Disks (LUN's) will be seen as hdisk at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
…
```

Then for disks to be used for ASM, and on all nodes :

1. Install MPIO on all nodes, attach the LUN to each node, dicover LUN's with "cfgmgr".
2. Identify hdisk names on each nodes, for a given LUN ID.
3. remove PVID from hdisk and change the reserve policy to no reserve using :

   chdev -l hdisk… -a pv=clear

   chdev -l hdisk… -a reserve_policy=no_reserve
4. set ownerchip to oracle:dba to the /dev/rhdisk…
5. set read/write permissions to 660 to the /dev/rhdisk…
6. access the disk thru /dev/rhdisk… for ASM diskgroup configuration

- **case 2 : Lun's provided by the EMC storage with EMC PowerPath installed as multi-pathing driver.**

  Disks (LUN's) will be seen as hdiskpower at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
hdiskpower0          00ced22cf79098ff                    rootvg          active
hdiskpower1          none                                None
hdiskpower2          none                                None
hdiskpower3          none                                None
hdiskpower4          none                                None
…
```

Then for disks to be used for ASM, and on all nodes :

1. Install PowerPath on all nodes, attach the LUN to each node, dicover LUN's with "cfgmgr".
2. Identify hdiskpower names on each nodes, for a given LUN ID.
3. remove PVID from hdiskpower and change the reserve policy to no reserve using :

   chdev -l hdiskpower… -a pv=clear

   chdev -l hdiskpower… -a reserve_lock=no
4. set ownerchip to oracle:dba to the /dev/rhdiskpower…
5. set read/write permissions to 660 to the /dev/rhdiskpower…
6. access the disk thru /dev/rhdiskpower… for ASM diskgroup configuration

6.4.4.1 *EMC PowerPath Setup Procedure*

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

On node 1, and node 2 …

1. **Install EMC ODM drivers and necessary filesets …**
   - 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
   {node1:root}/ -> rmdev –dl hdiskX

3. run /usr/lpp/EMC/Symmetrix/bin/emc_cfgmgr to detect devices

4. **Install PowerPath version 4.3.0 minimum using smit install**

5. register PowerPath
   {node1:root}/ -> emcpreg –install

6. initialize PowerPath devices
   {node1:root}/ -> powermt config

7. **verify that all PowerPath devices are named consistently across all cluster nodes**
   {node1:root}/ -> /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.

   Identify two small luns to be used for OCR and voting

   if the hdiskpowerX names for the OCR and VOTE devices are different, create a new device for each of these functions as follows:
   {node1:root}/ -> mknod /dev/ocr c <major # of OCR LUN> <minor # of OCR LUN>
   {node1:root}/ -> 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"
   {node1:root}/ -> chdev -l hdiskpowerX -a reserve_lock=no

9. Verify the attribute is set
   {node1:root}/ -> lsattr –El hdiskpowerX

10. Set permissions on all hdiskpower drives to be used for ASM, voting, or the OCR as follows :
    {node1:root}/ -> chown oracle:dba /dev/rhdiskpowerX
    {node1:root}/ -> chmod 660 /dev/rhdiskpowerX

The Oracle Installer will change these permissions and ownership as necessary during the CRS install process.

6.4.5 HITACHI storage and multi-pathing

With Hitachi, please refer to Hitachi to see which HDS storage is supported with RAC.

There are 3 cases when using Hitachi HDS storage :

- **case 1: Lun's provided by the HDS storage with IBM MPIO installed as multi-pathing driver.**

  Disks (LUN's) will be seen as hdisk at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
…
```

Then for disks to be used for ASM, and on all nodes :

1. Install MPIO on all nodes, attach the LUN to each node, dicover LUN's with "cfgmgr".
2. Identify hdisk names on each nodes, for a given LUN ID.
3. remove PVID from hdisk and change the reserve policy to no reserve using :

   chdev -l hdisk… -a pv=clear

   chdev -l hdisk… -a reserve_policy=no_reserve
4. set ownerchip to oracle:dba to the /dev/rhdisk
5. set read/write permissions to 660 to the /dev/rhdisk
6. access the disk thru /dev/rhdisk for ASM diskgroup configuration

- **case 2 : Lun's provided by the HDS storage with HDLM installed as multi-pathing driver.**

  Disks (LUN's) will be seen as dlmfdrv at AIX level using lspv command.

On node 1 …

```
{node1:root}/ -> lspv
dlmfdrv1          00ced22cf79098ff                  rootvg          active
dlmfdrv1          none                              None
dlmfdrv2          none                              None
dlmfdrv3          none                              None
dlmfdrv4          none                              None
…
```

Then for disks to be used for ASM, and on all nodes :

1. Install HDLM on all nodes, attach the LUN to each node, dicover LUN's with "cfgmgr".
2. Identify dlmfdrv names on each nodes, for a given LUN ID.
3. remove PVID from hdisk and change the reserve policy to no reserve using :

   chdev -l dlmfdrv… -a pv=clear

   chdev -l dlmfdrv… -a dlmrsvlevel=no_reserve
4. set ownerchip to oracle:dba to the /dev/rdlmfdrv…
5. set read/write permissions to 660 to the /dev/rdlmfdrv…
6. access the disk thru /dev/rdlmfdrv… for ASM diskgroup configuration

- **Case 3 : Lun's provided by the HDS storage with HDLM as multi-pathing driver.**

  Disks will be seen as dlmfdrv at AIX level using lspv command, and part of a HDLM VG (volume groups).

On node 1 …

```
{node1:root}/ -> lspv
dlmfdrv0          00ced22cf79098ff                    rootvg          active
dlmfdrv1          none                                vg_asm          active
dlmfdrv2          none                                vg_asm          active
dlmfdrv3          none                                vg_asm          active
dlmfdrv4          none                                vg_asm          active
dlmfdrv5          none                                vg_asm          active
…

or

{node1:root}/ -> lspv
dlmfdrv0          00ced22cf79098ff                    rootvg          active
dlmfdrv1          none                                vg_ocr_disk1    active
dlmfdrv2          none                                vg_voting_disk1 active
dlmfdrv3          none                                vg_asm_disk1    active
dlmfdrv4          none                                vg_asm_disk2    active
dlmfdrv5          none                                vg_asm_disk3    active
…
```

Then for disks to be used for ASM, and on all nodes :

1. Install HDLM on all nodes, attach the LUN to each node, dicover LUN's with "cfgmgr".
2. Turn off reserve locking on all nodes ➔ dlnkmgr set –rsv on 0 –s
3. Create VGs (Volume Groups) and LVs (Logical Volumes)

   2 options :

1/ Create a VG with all dlmfdrv, and create LV's out of it

On node 1 …

```
1) create the volume groups

 dlmmkvg -y vg_asm -B -s 256 -V 101 dlmfdrv1 dlmfdrv2
dlmfdrv3 dlmfdrv4 dlmfdrv5

2) enable the volume groups

dlmvaryonvg vg_asm

3) Create the logical volumes

  mklv -y lv_ocr_disk1 -T O -w n -s n -r n vg_asm 2
  mklv -y lv_voting_disk1 -T O -w n -s n -r n vg_asm 2
  mklv -y lv_asm_disk1 -T O -w n -s n -r n vg_asm 2
  mklv -y lv_asm_disk2 -T O -w n -s n -r n vg_asm 2
  mklv -y lv_asm_disk3 -T O -w n -s n -r n vg_asm 2
  …

4) Change permissions, owner, group of the raw devices
    B10811-05 Oracle DB Installation Guide p2-70

  chown oracle:dba /dev/rlv_ocr_disk1
  chown oracle:dba /dev/rlv_voting_disk1
  chown oracle:dba /dev/rlv_asm_disk1
  chown oracle:dba /dev/rlv_asm_disk2
  chown oracle:dba /dev/rlv_asm_disk3
  …
  chmod 660 /dev/rlv_ocr_disk1
  chmod 660 /dev/rlv_voting_disk1
  chmod 660 /dev/rlv_asm_disk1
  chmod 660 /dev/rlv_asm_disk2
  chmod 660 /dev/rlv_asm_disk3
  …
```

2/ Create on VG for one dlmfdrv, and one LV for one VG.

On node 1 …

```
1) create the volume groups

 dlmmkvg -y vg_ocr1 -B -s 256 -V 101 dlmfdrv1
 dlmmkvg -y vg_vot1 -B -s 256 -V 101 dlmfdrv2
 dlmmkvg -y vg_asm1 -B -s 256 -V 101 dlmfdrv3
 dlmmkvg -y vg_asm2 -B -s 256 -V 101 dlmfdrv4
 dlmmkvg -y vg_asm3 -B -s 256 -V 101 dlmfdrv5
  …

2) enable the volume groups

 dlmvaryonvg vg_ocr1
 dlmvaryonvg vg_vot1
 dlmvaryonvg vg_asm1
 dlmvaryonvg vg_asm2
 dlmvaryonvg vg_asm3
  …

3) Create the logical volumes

  mklv -y lv_ocr_disk1 -T O -w n -s n -r n vg_ocr1 2
  mklv -y lv_voting_disk1 -T O -w n -s n -r n vg_vot1
2
  mklv -y lv_asm_disk1 -T O -w n -s n -r n vg_asm1 2
  mklv -y lv_asm_disk2 -T O -w n -s n -r n vg_asm2 2
  mklv -y lv_asm_disk3 -T O -w n -s n -r n vg_asm3 2
  …

4) Change permissions, owner, group of the raw
devices
    B10811-05 Oracle DB Installation Guide p2-70

  chown oracle:dba /dev/rlv_ocr_disk1
  chown oracle:dba /dev/rlv_voting_disk1
  chown oracle:dba /dev/rlv_asm_disk1
  chown oracle:dba /dev/rlv_asm_disk2
```

```
5) disable the volume groups

dlmvaryoffvg vg_asm
```

On node 2 …

```
6) Identify which dlmfdrv correspond to dlmfdrv1 from
node1.

 dlmfdrv1 on node1 ➔ dlmfdrv0 on node2

7) import the volume groups on node2

This will copy the vg/lv configuration that was made on
node1

 dlmimportvg -V 101 -y vg_asm dlmfdrv0

7) enable the volume groups on node2

 dlmvaryonvg vg_asm

8) this will ensure the vg will not get varyon'd at boot

 dlmchvg -a n vg_asm
```

On node 1 …

```
9) enable the volume groups on node1

 dlmvaryonvg vg_asm
```

Check for document :

[Hitachi Dynamic Link Manager (HDLM) for IBM AIX Systems User's Guide](http://sys-admin.net/uploads/SAN/hldm_admin_guide.pdf) :
http://sys-admin.net/uploads/SAN/hldm_admin_guide.pdf

```
  chown oracle:dba /dev/rlv_asm_disk3
  …
  chmod 660 /dev/rlv_ocr_disk1
  chmod 660 /dev/rlv_voting_disk1
  chmod 660 /dev/rlv_asm_disk1
  chmod 660 /dev/rlv_asm_disk2
  chmod 660 /dev/rlv_asm_disk3
  …

5) disable the volume groups

 dlmvaryoffvg vg_ocr1
 dlmvaryoffvg vg_vot1
 dlmvaryoffvg vg_asm1
 dlmvaryoffvg vg_asm2
 dlmvaryoffvg vg_asm3
  …
```

On node 2 …

```
6) Identify which dlmfdrv correspond to dlmfdrv1
from node1, and so on with other dlmfdrv

 dlmfdrv1 on node1 ➔ dlmfdrv0 on node2
 dlmfdrv2 on node1 ➔ dlmfdrv1 on node2
 dlmfdrv3 on node1 ➔ dlmfdrv2 on node2
 dlmfdrv4 on node1 ➔ dlmfdrv3 on node2
 dlmfdrv5 on node1 ➔ dlmfdrv4 on node2
  …

7) import the volume groups on node2

This will copy the vg/lv configuration that was made
on node1

 dlmimportvg -V 101 -y vg_ocr1 dlmfdrv0
 dlmimportvg -V 101 -y vg_vot1 dlmfdrv1
 dlmimportvg -V 101 -y vg_asm1 dlmfdrv2
 dlmimportvg -V 101 -y vg_asm2 dlmfdrv3
 dlmimportvg -V 101 -y vg_asm3 dlmfdrv4
  …


7) enable the volume groups on node2

 dlmvaryonvg vg_ocr1
 dlmvaryonvg vg_vot1
 dlmvaryonvg vg_asm1
 dlmvaryonvg vg_asm2
 dlmvaryonvg vg_asm3
  …

8) this will ensure the vg will not get varyon'd at
boot

 dlmchvg -a n vg_ocr1
 dlmchvg -a n vg_vot1
 dlmchvg -a n vg_asm1
 dlmchvg -a n vg_asm2
 dlmchvg -a n vg_asm3
 …
```

On node 1 …

```
9) enable the volume groups on node1

 dlmvaryonvg vg_ocr1
 dlmvaryonvg vg_vot1
 dlmvaryonvg vg_asm1
 dlmvaryonvg vg_asm2
 dlmvaryonvg vg_asm3
  …
```

4. remove any PVID from dlmfdrv…
5. access the disk thru /dev/rlv_asm_disk… for ASM diskgroup configuration

6.4.6 Others, StorageTek, HP EVA storage and multi-pathing

For most of the storage solutions, please contact the providing company for supported configuration, as read/write concurrent acces from all RAC nodes must be possible to implement a RAC solution. That means possibility to setup the disk reserve_policy to no_reserve or equivalent.

7 SPECIFIC CONSIDERATIONS FOR RAC/ASM SETUP WITH HACMP INSTALLED

Oracle Clusterware does not requires HACMP to work, but some customers may need to have HACMP installed on the RAC node cluster to protect third party products, or ressources. Oracle clusterware could replace HACMP for most operations as cold failover for 10g and 9i single database, applications servers or any applications.

Please check following documents for details on
http://www.oracle.com/technology/products/database/clustering/index.html :

- Comparing Oracle Real Applicaton Clusters to Failover Clusters for Oracle Database(PDF) December 2006
- Workload Management with Oracle Real Application Clusters (FAN, FCF, Load Balancing) (PDF) May 2005
- Using Oracle Clusterware to Protect 3rd Party Applications (PDF) May 2005
- Using Oracle Clusterware to Protect a Single Instance Oracle Database (PDF) November 2006
- Using Oracle Clusterware to protect Oracle Application Server (PDF) November 2005
- How to Build End to End Recovery and Workload Balancing for Your Applications 10g Release 1(PDF) Dec 2004
- Oracle Database 10*g* Services (PDF) Nov 2003
- Oracle Database 10*g* Release 2 Best Practices: Optimizing Availability During Unplanned Outages Using Oracle Clusterware and RAC ***New!***

HOWEVER, if customer still need to have HACMP, Oracle Clusterware can cohabitate with HACMP. ONLY HACMP 5.1 and 5.2 are supported to work on same servers as Oracle Clusterware 10g R2.

1. HACMP must not take-over/failover the Oracle Clusterware ressources (VIP, database, etc …)

2. HACMP VIP must not be configured on IP from Public node name used by RAC (hostname), or Oracle Clusterware VIP

3. With 10g, it's not necessary to declare the RAC interconnect in HACMP

4. It's not mandatory to declare hdisks used for ASM in HACMP as logical volumes (LV) from Volume Groups (VG). In this case follow the cookbook to prepare the disks for OCR, Voting and ASM disks.

5. If the choice is to declare hdisks used by ASM in HACMP Volume Groups, THEN you'll have to prepare the disks for OCR, Voting, ASM spfile and ASM disks as describe in official Oracle document available on http://tahiti.oracle.com

   Oracle® Database Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10*g* Release 2 (10.2) for AIX
   Part Number B14201-04
   http://download-uk.oracle.com/docs/cd/B19306_01/install.102/b14201/toc.htm

Check for the following chapters :

3.3 Configuring Storage for Oracle Clusterware Files on Raw Devices
The following subsections describe how to configure Oracle Clusterware files on raw partitions.

Configuring Raw Logical Volumes for Oracle Clusterware
Creating a Volume Group for Oracle Clusterware
Configuring Raw Logical Volumes in the New Oracle Clusterware Volume Group
Importing the Volume Group on the Other Cluster Nodes
Activating the Volume Group in Concurrent Mode on All Cluster Nodes

For OCR/voting disks, do create a volume group (VG), and create logical volumes (lv) with names as /dev/Ocr_Disk1, /dev/Ocr_Disk2, /dev/Voting_Disk1, etc …
And don't forget to remove the reserve policy on all hdisks …

3.6 Configuring Database File Storage on Raw Devices
The following subsections describe how to configure raw partitions for database files.

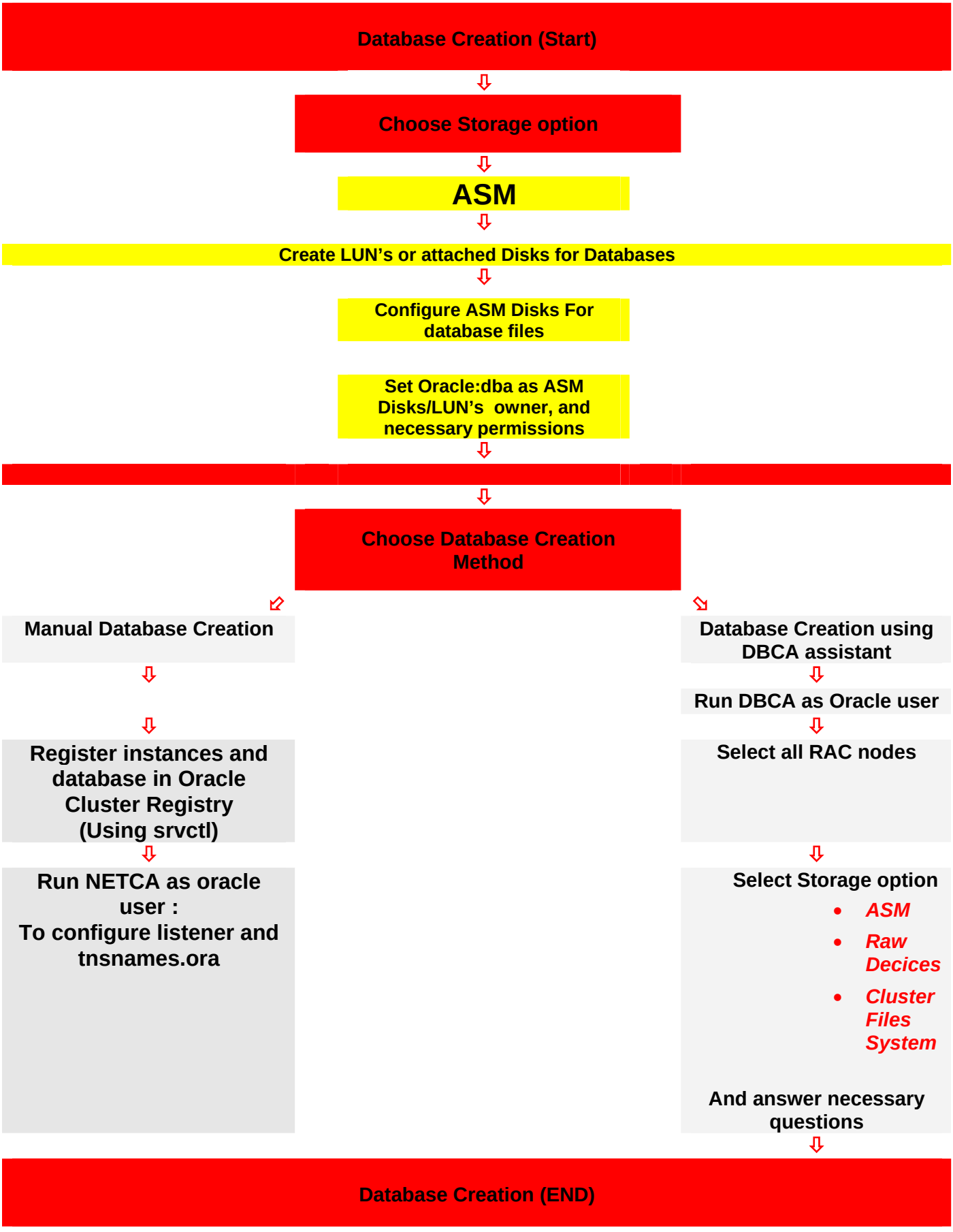
Configuring Raw Logical Volumes for Database File Storage
Creating a Volume Group for Database Files
Creating Database File Raw Logical Volumes in the New Volume Group
Importing the Database File Volume Group on the Other Cluster Nodes
Activating the Database File Volume Group in Concurrent Mode on All Cluster Nodes

For ASM disks, do create a volume group (VG), and create logical volumes (lv) with names as /dev/ASM_Disk1, /dev/ASM_Disk2, /dev/ASM_Disk3, etc …
And don't forget to remove the reserve policy on all hdisks …

8 INSTALLATION STEPS

8.1 RAC INSTALLATION STEPS

8.2 DATABASE CREATION STEPS



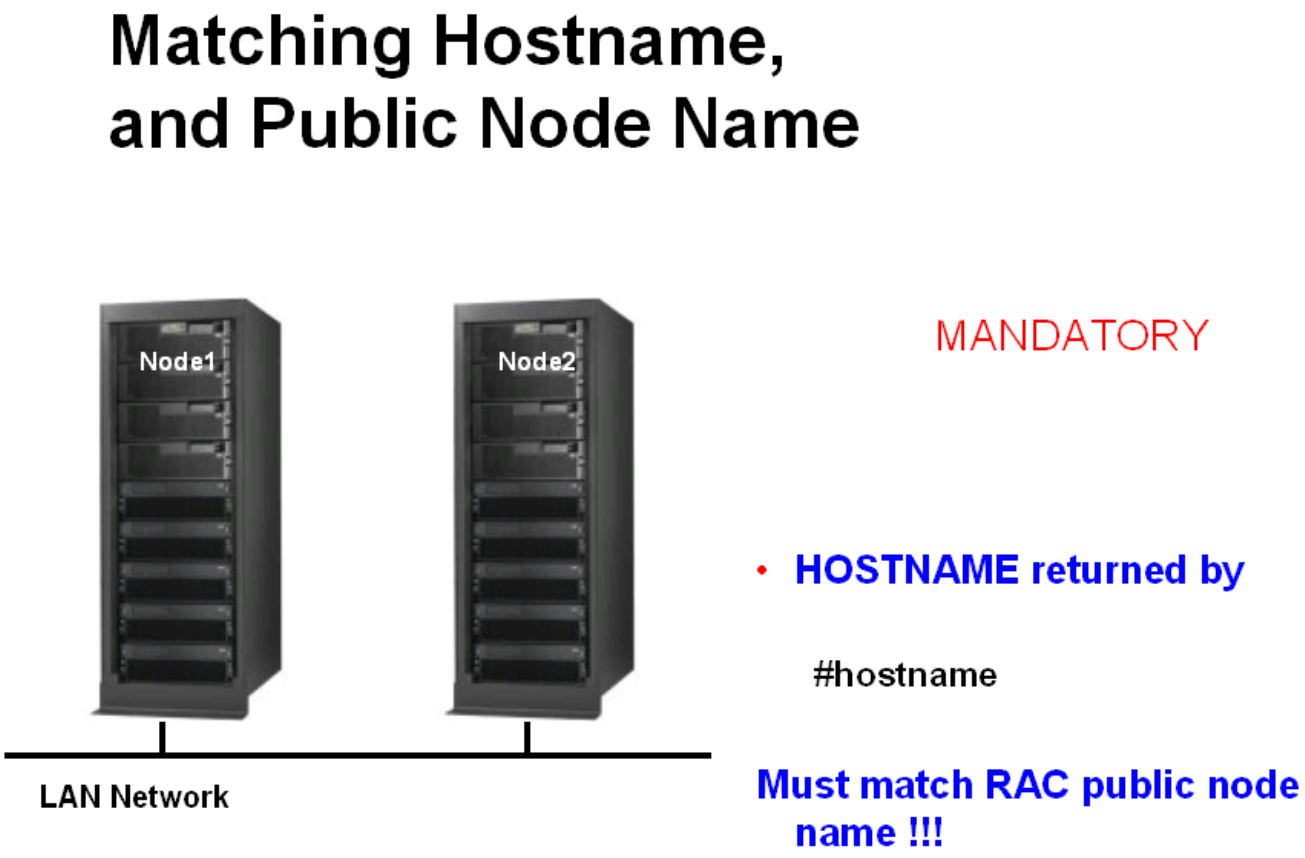
9 CHECK LIST TO USE AND FOLLOW

This is the list of operations you should do, before moving to Oracle Installation steps :

Operations		Done on each node : Yes/No ?					
		Node 1	Node 2	Node 3	Node 4	Node 5	...
1	• Check the Hardware Requirements						
2	• Check the Network Requirements						
3	• Check the Software Requirements						
4	• Create Required UNIX Groups and User						
5	• Configure Kernel Parameters and Shell Limits						
6	• Identify Required Software Directories						
7	• Identify or Create an Oracle Base Directory						
8	• Create the CRS Home Directory						
9	• Choose a Storage Option for Oracle CRS, Database, and Recovery Files						
11	• Create Directories for Oracle CRS, Database, or Recovery Files						
12	• Configure Disks for Automatic Storage Management ➔ **And ACCESS Disk Storage concurently from all nodes participating to the RAC cluster**						
13	• Synchronize the System Time on Cluster Nodes						
14	• Stop Existing Oracle Processes						
15	• Configure the oracle User's Environment						

10 PREPARING THE SYSTEM

10.1 Matching Server Hostname and RAC Public Node Name



Before setting the default hostname of each server, make sure to match it with the RAC Public node !!!

From node1 :

{**node1**:root}/ -> **hostname**
node1
{node1:root}/ ->

From node2 :

{**node2**:root}/ -> **hostname**
node2
{node2:root}/ ->

Default Server Hostname		RAC Public Node	
en0			
Node Name	IP	Node Name	IP
node1	10.3.25.81	node1	10.3.25.11
node2	10. 3.25.82	node2	10. 3.25.12

10.2 HARDWARE REQUIREMENTS

Requirements to meet

on ALL nodes !!!

- **RAM >= 512 MB minimum**
  ➔. Command to check the physical memory :
  lsattr –El sys0 –a realmem

- **Internal disk >= 12 GB for the oracle code (CRS_HOME, ASM_HOME, ORACLE_HOME)**

- **Paging space = 2 x RAM,** with a minimum of 400 MB and a maximum of 2 GB.
  ➔ To check the paging space configured :
  lsps –a

- **Temporary Disk Space :** The Oracle Universal Installer requires up to 400 MB of free space in the /tmp directory.
  ➔ To check the free temporary space available:
  df –k /tmp
  ➔ You can use an other filesystem instead of /tmp.
  Set the TEMP environment variable (used by Oracle)
  and the TMPDIR environment variable to the new location.

  For example :
  export TEMP=/new_tmp
  export TMPDIR=/new_tmp
  export TMP=/new_tmp

10.3 AIX SOFTWARE REQUIREMENTS

To have the latest information please refer to Metalink Note 282036.1. on http://metalink.oracle.com, this document include last update. Check also certification status on metalink.oracle.com, or otn.oracle.com

AIX release supported with Oracle 10g RAC R2 as for March 27, 2007.

- AIX 5L version 5.3, Maintenance Level 02 or later
- AIX 5L version 5.2, Maintenance Level 04 or later

To determine which version of AIX is installed, enter the following command :

```
{node1:root}/ -> oslevel -s
5300-05-04 ➔ gives level of ML or TL and sub-level of
service pack, which means AIX5L TL05 SP4
```

If the operating system version is lower than the minimum required, upgrade your operating system to this level. AIX 5L maintenance packages are available from the following Web site :
http://www-912.ibm.com/eserver/support/fixes/

10.3.1 Filesets Requirements for 10gRAC R2 / ASM (NO HACMP)

AIX filesets required on ALL nodes for 10gRAC Release 2 implementation with ASM !!!

Check that the required filsets are installed on the system. (Note: If the PTF is not downloadable, customers should request an efix through AIX customer support.)	**AIX 5L version 5.2, Maintenance Level 04 or later**	**AIX 5L version 5.3, Maintenance Level 02 or later**
	Filesets • bos.adt.base • bos.adt.lib • bos.adt.libm • bos.perf.libperfstat • bos.perf.perfstat • bos.perf.proctools • rsct.basic.rte • rsct.compat.clients.rte • xlC.aix50.rte 7.0.0.4 or 8.xxx • xlC.rte 7.0.0.1 or 8.xxx **Specific Filesets** **For EMC Symmetrix :** • EMC.Symmetrix.aix.rte.5.2.0.1 **For EMC CLARiiON :** • EMC.CLAR11ON.fcp.rte.5.2.0.1	**Filesets** • bos.adt.base • bos.adt.lib • bos.adt.libm • bos.perf.libperfstat • bos.perf.perfstat • bos.perf.proctools • rsct.basic.rte • rsct.compat.clients.rte • xlC.aix50.rte 7.0.0.4 or 8.xxx • xlC.rte 7.0.0.1 or 8.xxx **Specific Filesets** **For EMC Symmetrix :** • EMC.Symmetrix.aix.rte.5.2.0.1 **For EMC CLARiiON :** • EMC.CLAR11ON.fcp.rte.5.2.0.1

Depending on the AIX Level that you intend to install, verify that the required filesets are installed on the system. The following procedure describes how to check these requirements.

To ensure that the system meets these requirements, follow these steps :

And Check that required filesets are all installed.

And that :
xlC.aix50.rte and xlC.rte are at minimum release of 7.0.0.4

```
{node1:root}/ -> lslpp -l bos.adt.base bos.adt.lib bos.adt.libm bos.perf.libperfstat
bos.perf.perfstat bos.perf.protocols rsct.basic.rte rsct.compat.clients.rte
xlC.aix50.rte xlC.rte

  Fileset                      Level  State      Description
  ----------------------------------------------------------------------------
Path: /usr/lib/objrepos
  bos.adt.base              5.3.0.51  APPLIED    Base Application Development
                                                 Toolkit
  bos.adt.lib               5.3.0.50  COMMITTED  Base Application Development
                                                 Libraries
  bos.adt.libm              5.3.0.40  COMMITTED  Base Application Development
                                                 Math Library
  bos.perf.libperfstat      5.3.0.50  COMMITTED  Performance Statistics Library
                                                 Interface
  bos.perf.perfstat         5.3.0.50  COMMITTED  Performance Statistics
                                                 Interface
  bos.perf.proctools        5.3.0.50  COMMITTED  Proc Filesystem Tools
  rsct.basic.rte             2.4.5.0  COMMITTED  RSCT Basic Function
  rsct.compat.clients.rte    2.4.5.0  COMMITTED  RSCT Event Management Client
                                                 Function
  xlC.aix50.rte              8.0.0.4  COMMITTED  C Set ++ Runtime for AIX 5.0
  xlC.rte                    8.0.0.4  COMMITTED  C Set ++ Runtime

Path: /etc/objrepos
  bos.perf.libperfstat       5.3.0.0  COMMITTED  Performance Statistics Library
                                                 Interface
  bos.perf.perfstat         5.3.0.50  COMMITTED  Performance Statistics

                                                 Interface
  rsct.basic.rte             2.4.5.0  COMMITTED  RSCT Basic Function
{node1:root}/ ->
```

If a fileset is not installed and committed, then install it. Refer to your operating system or software documentation for information about installing filesets.

10.3.2 APAR's Requirements for 10gRAC R2 / ASM (NO HACMP)

Check that the required APAR's are installed on the system.

AIX Patches (APAR) required on ALL nodes for 10gRAC R2 implementation with ASM !!!

(Note: If the PTF is not downloadable, customers should request an efix through AIX customer support.)

AIX 5L version 5.2, Maintenance Level 04 or later	AIX 5L version 5.3, Maintenance Level 02 or later
APAR's **Authorized Problem Analysis Reports (APARs) for AIX 5L v5.2 ML 4:** • **IY63133**: large percentage of CPU time spent in ldata_balance routine • **IY63366**: dlsym returns null even for valid symbol in AIX520 ML-4 • **IY64691**: chvg -b can cause corruption and crash • **IY64737**: AIO can hang in knotunlock • **IY65001**: mklvcopy on a striped lv is failing to update lvcb • **IY64978**: deadlock with concurrent renaming and unlinking under JFS • **IY65001** • **IY75901, if IY69518 was previously installed.** • **IY70029**: "CORRUPTION FROM SIMULTANEOUS CIO WRITES WITH O_DSYNC ON JFS2" If using the IBM Journal File System Version 2 (JFS2) for Oracle Database files. **JDK : (Not mandatory for the installation)** • **IY58350** Patch for SDK 1.3.1.16 (32-bit) • **IY63533** Patch for SDK 1.4.2.1 (64-bit) • **IY65305** Patch for SDK 1.4.2.2 (32-bit)	**APAR's** **For AIX 5L V5.3 Technology Level 5 (TL 5300-05) IBM AIX PTF for APAR IY89080 must be installed.** **In addition, Oracle customers should contact Oracle support to obtain the fix for Oracle Bug 5496862.** **TL 5300-05-05 and TL 5300-05-05 does include the APAR IY89080.** **Authorized Problem Analysis Reports (APARs) for AIX 5L v5.3 ML02:** • **IY68989**: "WRITE TO MMAPPED SPACE HANGS" • **IY68874**: An application that is run with mandatory large page data (LDR_CNTRL=LARGE_PAGE_DATA=M) may core-dump on the AIX 5.3 64-bit kernel in a LPAR environment. • **IY70031**: "CORRUPTION FROM SIMULTANEOUS CIO WRITES WITH O_DSYNC ON JFS2" If using the IBM Journal File System Version 2 (JFS2) for Oracle Database files. • **If using AIX 5L V5.3 Maintenance Level 3 (ML03) for any Oracle product, install the PTF for APAR IY76140**. **JDK : (Not mandatory for the installation)** • **IY58350** Patch for SDK 1.3.1.16 (32-bit) • **IY63533** Patch for SDK 1.4.2.1 (64-bit) • **IY65305** Patch for SDK 1.4.2.2 (32-bit)

To ensure that the system meets these requirements, follow these steps:

To determine whether an APAR is installed, enter a command similar to the following:

This is an example for AIX5L 5.3, with TL05 :

```
{node1:root}/ -> /usr/sbin/instfix -i -k "IY68989 IY68874 IY70031 IY76140 IY89080"
    All filesets for IY68989 were found.
    All filesets for IY68874 were found.
    All filesets for IY70031 were found.
    All filesets for IY76140 were found.
    All filesets for IY89080 were found.
{node1:root}/ ->
```

If an APAR is not installed, download it from the following Web site and install it:
http://www-912.ibm.com/eserver/support/fixes/

10.4 Oracle Software Requirements

Link to donwload code from http://otn.oracle.com/
http://www.oracle.com/technology/software/products/database/oracle10g/htdocs/10201aixsoft.html

Orale CD's needed for the RAC installation	Oracle Database 10*g* Release 2 (10.2.0.1.0) Enterprise/Standard Edition for AIX5L rdbms_aix5l64_database.cpio.gz (1,268,576,110 bytes) (cksum - 3772623559) Oracle Clusterware Release 2 (10.2.0.1.0) rdbms_aix5l64_cluster.cpio.gz (732,798,920 bytes) (cksum - 1766193337)
Clusterware and Database Patchset needed **Best is to implement last patchset for Oracle Clusterware and ASM software**	Patch Number 4547817 10.2.0.2 PATCH SET FOR ORACLE DATABASE SERVER 10.2.0.2 Delivered : 17-AUG-2006 Size : 1.3G Patch Number 5337014 10.2.0.3 PATCH SET FOR ORACLE DATABASE SERVER 10.2.0.3 Delivered : 22-FEB-2007 Size : 1.5G **Application of patchset to the database software depends if your application has been tested with it, it's your choice, your desision.**
Extra Clusterware Patches needed	CRS Bundle # 3 ONLY if 10.2.0.2 PATCH SET implemented CRS Bundle # 3 included in the 10.2.0.3 PATCH SET
Extra ASM Patches needed	ARU p5893614
Extra Database Patches needed	As needed for your project ….

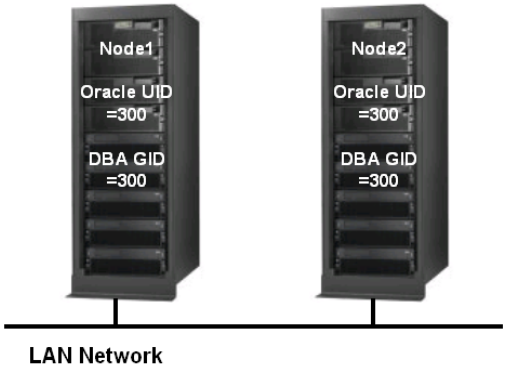
Directions to extract contents of .gz files :

1. Unzip the file: **gunzip** <filename>
2. Extract the file: **cpio -idcmv <** <filename>
3. Installation guides and general Oracle Database 10*g* documentation can be found here.
4. Review the certification matrix for this product here.

10.5 Users and Groups

This setup has to be done on all the nodes of the cluster.

Oracle User and DBA group



User and Group

MANDATORY

- **Oracle username must be same on each node**
- **Oracle user ID must be same on each node**
- **DBA oracle group should be same on each node, and set as primary group for oracle user.**

Be sure that all the groups and user numbers (500, 501 in our case) are identical thru the nodes.

To create the following groups :

smit group

dba ➔ Primary group for the oracle user.
oinstall ➔ The ora inventory group. This group is not mandatory. If it exists, it will be the group owner of the oracle code files. This group is a secondary group for oracle user.

To create the users :

smit user

oracle ➔ Owner of the database.

The **oracle** user must have **dba** as primary group, **oinstall** as secondary groups.

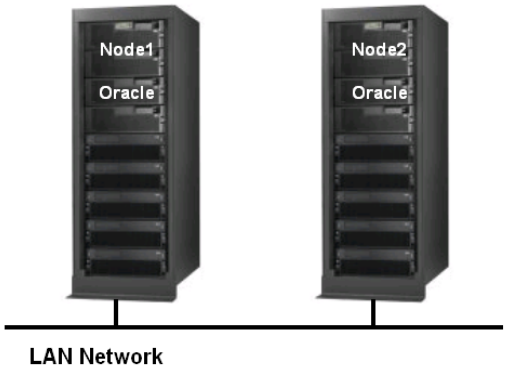
Verification: check if the file /etc/group contains lines such as : (the numbers could be different)

{**node1**:root}/ -> **cat /etc/group |grep oracle**
dba:!:300:oracle
oinstall:!:301:oracle

{**node1**:root}/ -> su - oracle
{node1:oracle}/oracle -> **id**
uid=300(oracle) gid=300(dba) groups=1(staff)
{node1:oracle}/oracle ->

{**node2**:root}/ -> su - oracle
{node2:oracle}/oracle -> **id**
uid=300(oracle) gid=300(dba) groups=1(staff)
{node2:oracle}/oracle ->

Setting Oracle User Password



Oracle Users

MANDATORY

- Oracle User Password must be set

  #passwd oracle

- Telnet connection to oracle user must be done to test successful connection, and changing password at first time if needed !!!

  As oracle #telnet node1
  #telnet node2

Set a password to oracle user, the same for all the nodes of the cluster, with the command :

{**node1**:root}/ -> **passwd oracle**
Changing password for "oracle"
oracle's New password:
Enter the new password again:
{node2:root}/ ->

{**node2**:root}/ -> **passwd oracle**
Changing password for "oracle"
oracle's New password:
Enter the new password again:
{node2:root}/ ->

And connect at least once to each node as oracle user, to validate the password, as first connexion will ask to change the password :

{**node1**:root}/ -> **telnet node1**
…….

{**node2**:root}/ -> **telnet node2**
…….

10.6 Configure Kernel Parameters and Shell Limits

Configuring Shell Limits, System Configuration, and Network Tuning Parameters (Extract from Oracle Documentation)

> **Note:**
>
> The parameter and shell limit values shown in this section **are minimum recommended values** only. For production database systems, Oracle recommends that you tune these values to optimize the performance of the system. Refer to your operating system documentation for more information about tuning kernel parameters.

Oracle recommends that you set shell limits, system configuration parameters, and network tuning parameters as described in this section on all cluster nodes, for Oracle and root users.

10.6.1 Configure Shell Limits

Verify that the shell limits shown in the following table are set to the values shown. The procedure following the table describes how to verify and set the values.

Shell Limit (As Shown in smit)	Recommended Value for oracle user	Recommended Value for root user
Soft FILE size	-1 (Unlimited)	-1 (Unlimited)
Soft CPU time	-1 (Unlimited) **Note:** This is the default value.	-1 (Unlimited) **Note:** This is the default value.
Soft DATA segment	-1 (Unlimited)	-1 (Unlimited)
Soft STACK size	-1 (Unlimited)	-1 (Unlimited)

To view the current value specified for these shell limits, and to change them if necessary,

follow these steps:

1. **Enter the following command:**

   # smit chuser

2. **In the** User NAME **field, enter the user name of the Oracle software owner, for example oracle.**
3. **Scroll down the list and verify that the value shown for the soft limits listed in the previous table is -1.**

   If necessary, edit the existing value.

4. **When you have finished making changes, press F10 to exit.**

OR for root and oracle user on each node, to check thru ulimit command :

```
{node1:root}/ -> ulimit -a
time(seconds)        unlimited
file(blocks)         unlimited
data(kbytes)         unlimited
stack(kbytes)        4194304
memory(kbytes)       unlimited
coredump(blocks)     unlimited
nofiles(descriptors) unlimited
{node1:root}/
```

```
{node1:root}/ -> su - oracle
{node1:oracle}/oracle -> ulimit -a
time(seconds)        unlimited
file(blocks)         unlimited
data(kbytes)         unlimited
stack(kbytes)        4194304
memory(kbytes)       unlimited
coredump(blocks)     unlimited
nofiles(descriptors) unlimited
{node1:oracle}/
```

10.6.2 Configure System Configuration Parameters

Verify that the maximum number of processes allowed per user is set to 2048 or greater :

> **Note:**
> For production systems, this value should be at least 128 plus the sum of the PROCESSES and PARALLEL_MAX_SERVERS initialization parameters for each database running on the system.

To check the value …

```
{node1:root}/ -> lsattr -El sys0 -a maxuproc
maxuproc 1024 Maximum number of PROCESSES allowed per user True
{node1:root}/ ->
```

To edit and modify the value …

```
Setting value to 10000

{node1:root}/ -> chdev -l sys0 -a maxuproc='10000'
{node1:root}/ ->

{node1:root}/ -> lsattr -El sys0 -a maxuproc
maxuproc 10000 Maximum number of PROCESSES allowed per user True
{node1:root}/ ->
```

OR

1. **Enter the following command:**

   ```
   # smit chgsys
   ```

2. **Verify that the value shown for Maximum number of PROCESSES allowed per user is greater than or equal to 2048.**

   If necessary, edit the existing value.

3. **When you have finished making changes, press F10 to exit.**

Verify that the lru_file_repage parameter is set to 0 :

The default value is "1", **but it is recommended to set this to "0"**. This setting hints to the VMM to only steal file pages (from the AIX file buffer cache) and leave the computational pages (from the SGA) alone.
By setting "lru_file_repage=0", AIX only frees file cache memory. This guarantees working storage stays in memory, and allows file cache to grow.

So in the past you might have set maxclient at 20% for database servers. Today you could set maxclient at 90% and "lru_file_repage=0". The exact setting will vary based on your application and amount of memory. Contact IBM Support if you need help determining the optimum setting.

This new lru_file_repage parameter is only available on AIX 5.2 ML04+ and AIX 5.3 ML01+

To check the value on each node ...

```
{node1:root}/ -> vmo -L lru_file_repage
NAME                      CUR    DEF    BOOT   MIN    MAX    UNIT           TYPE
     DEPENDENCIES
------------------------------------------------------------------------------
lru_file_repage           1      1      1      0      1      boolean          D
------------------------------------------------------------------------------
{node1:root}/ ->
```

Change the value to 0 on each node :

```
{node1:root}/ -> vmo -p -o lru_file_repage=0
Setting lru_file_repage to 0 in nextboot file
Setting lru_file_repage to 0
{node1:root}/ ->

{node1:root}/ -> vmo -L lru_file_repage
NAME                      CUR    DEF    BOOT   MIN    MAX    UNIT           TYPE
     DEPENDENCIES
------------------------------------------------------------------------------
lru_file_repage           0      1      0      0      1      boolean          D
------------------------------------------------------------------------------
{node1:root}/ ->
```

Typical vmo settings for Oracle :

- lru_file_repage=0 (default=1) (AIX 5.2 ML04 or later)
  - Forces file pages to be repaged before computational pages
- minperm%=5 (default 20)
  - Target for minimum % of physical memory to be used for file system cache
- maxperm%=15 (default 80)

Setting Asynchronous I/O :

To check the value on each node …

```
{node1:root}/ -> smitty aio
```

```
                                  Asynchronous I/O (Legacy)

Move cursor to desired item and press Enter.

  Change / Show Characteristics of Asynchronous I/O
  Remove Asynchronous I/O; Keep Definition
  Configure Defined Asynchronous I/O
  Generate Error Report
  Trace Asynchronous I/O
```

Select change / Show Characteristics of Asynchronous I/O

```
                        Change / Show Characteristics of Asynchronous I/O

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                      [Entry Fields]
  MINIMUM number of servers                          [150]
  MAXIMUM number of servers per cpu                  [300]
  Maximum number of REQUESTS                         [16384]
  Server PRIORITY                                    [39]
  STATE to be configured at system restart            available
  State of fast path                                  enable
```

Set STATE to “available”

OR

```
{node1:root}/ -> lsattr -El aio0
autoconfig available STATE to be configured at system restart True
fastpath   enable    State of fast path                       True
kprocprio  39        Server PRIORITY                          True
maxreqs    16384     Maximum number of REQUESTS               True
maxservers 300       MAXIMUM number of servers per cpu        True
minservers 150       MINIMUM number of servers                True
{node1:root}/ ->
```

10.6.3 Configure Network Tuning Parameters

Verify that the network tuning parameters shown in the following table are set to the values shown or higher values. The procedure following the table describes how to verify and set the values.

Network Tuning Parameter	Recommended Value on all nodes
ipqmaxlen	512
rfc1323	1
sb_max	1310720
tcp_recvspace	65536
tcp_sendspace	65536
udp_recvspace	655360 **Note:** The recommended value of this parameter is 10 times the value of the udp_sendspace parameter. The value must be less than the value of the sb_max parameter.
udp_sendspace	65536 **Note:** This value is suitable for a default database installation. For production databases, the minimum value for this parameter is 4 KB plus the value of the database DB_BLOCK_SIZE initialization parameter multiplied by the value of the DB_MULTIBLOCK_READ_COUNT initialization parameter: (DB_BLOCK_SIZE * DB_MULTIBLOCK_READ_COUNT) + 4 KB

To check values …

```
{node1:root}/ -> for i in ipqmaxlen rfc1323 sb_max tcp_recvspace
tcp_sendspace udp_recvspace udp_sendspace
<
> do
>       no -a |grep $i
> done
                ipqmaxlen = 512
                  rfc1323 = 1
                   sb_max = 1310720
            tcp_recvspace = 65535
            tcp_sendspace = 65535
            udp_recvspace = 655350
            udp_sendspace = 65535
{node1:root}/ ->

{node2:root}/ -> for i in ipqmaxlen rfc1323 sb_max tcp_recvspace
tcp_sendspace udp_recvspace udp_sendspace
<
> do
>       no -a |grep $i
> done
                ipqmaxlen = 512
                  rfc1323 = 1
                   sb_max = 1310720
            tcp_recvspace = 65535
            tcp_sendspace = 65535
            udp_recvspace = 655350
            udp_sendspace = 65535
{node2:root}/ ->
```

To change the current values to required ones, if necessary,

follow these steps :

1. **If you must change the value of any parameter, enter the following command to determine whether the system is running in compatibility mode:**

   # /usr/sbin/lsattr -E -l sys0 -a pre520tune

   If the system is running in compatibility mode, the output is similar to the following, showing that the value of the pre520tune attribute is enable:

   pre520tune enable Pre-520 tuning compatibility mode True

   By default, with AIX5L, compatibility mode is set to false !!!
   Change it to true ONLY if necessary !!!

   **** if you want to enable the compatibility mode, issue the following command :**
   # chdev -l sys0 -a pre520tune=enable

2. **If the system is running in compatibility mode,**

THEN

follow these steps to change the parameter values:

Enter commands similar to the following to change the value of each parameter:

```
# /usr/sbin/no –o parameter_name=value

For example:
# /usr/sbin/no –o udp_recvspace=655360
```

Add entries similar to the following to the /etc/rc.net file for each parameter that you changed in the previous step:

```
if [ -f /usr/sbin/no ] ; then
   /usr/sbin/no -o udp_sendspace=65536
   /usr/sbin/no -o udp_recvspace=655360
   /usr/sbin/no -o tcp_sendspace=65536
   /usr/sbin/no -o tcp_recvspace=65536
   /usr/sbin/no -o rfc1323=1
   /usr/sbin/no -o sb_max=1310720
   /usr/sbin/no -o ipqmaxlen=512
fi
```

By adding these lines to the /etc/rc.net file, the values persist when the system restarts.
THEN YOU NEED TO Modify and REBOOT all nodes !

ELSE

 # /usr/sbin/lsattr -E -l sys0 -a pre520tune

Will give the following :

```
pre520tune disable Pre-520 tuning
compatibility mode True
```

The system is not running in compatibility mode, enter commands similar to the following to change the parameter values:

Enter commands similar to the following to change the value of parameter ipqmaxlen:

```
ipqmaxlen parameter:
/usr/sbin/no -r -o ipqmaxlen=512
```

Enter commands similar to the following to change the value of each others parameters:

```
/usr/sbin/no -p –o udp_sendspace=65536
/usr/sbin/no -p –o udp_recvspace=655360
/usr/sbin/no -p –o tcp_sendspace=65536
/usr/sbin/no -p –o tcp_recvspace=65536
/usr/sbin/no -p –o rfc1323=1
/usr/sbin/no -p –o sb_max=1310720
```

Note: If you modify the ipqmaxlen parameter, you must restart the system.

These commands modify the /etc/tunables/nextboot file, causing the attribute values to persist when the system restarts.

Content of /etc/tunables/nextboot

For node1 :

{node1:root}/ -> **cat /etc/tunables/nextboot**

```
# IBM_PROLOG_BEGIN_TAG
# This is an automatically generated prolog.
#
# bos530 src/bos/usr/sbin/perf/tune/nextboot 1.1
#
# Licensed Materials - Property of IBM
#
# (C) COPYRIGHT International Business Machines Corp. 2002
# All Rights Reserved
#
# US Government Users Restricted Rights - Use, duplication or
# disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
#
# IBM_PROLOG_END_TAG

vmo:
        lru_file_repage = "0"
        minperm% = "5"
        maxperm% = "15"
        maxpin% = "80"
        maxclient% = "35"
        strict_maxclient = "0"
        minfree = "3000"
        maxfree = "4000"

no:
        sb_max = "1310720"
        udp_recvspace = "655350"
        rfc1323 = "1"
        ipqmaxlen = "512"
        udp_sendspace = "65535"
        tcp_sendspace = "65535"
        tcp_recvspace = "65535"
{node1:root}/ ->
```

10.7 NETWORK CONFIGURATION

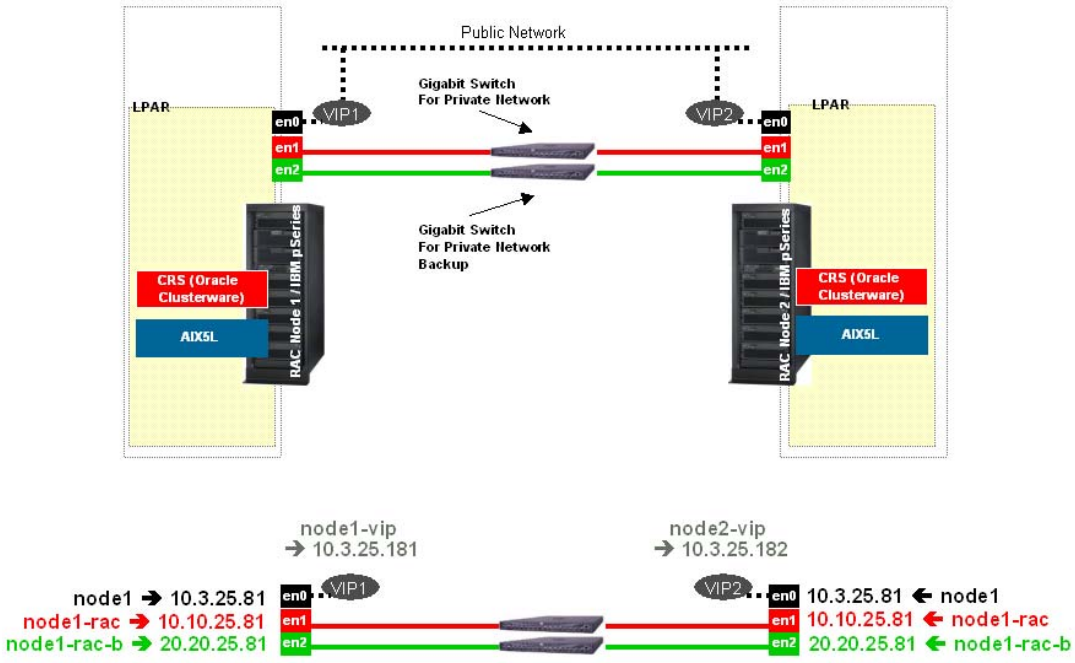
10.7.1 Possible network configuration layout of our Sytem p RAC cluster

Network architecture can be done with 2, or more nodes, but always with a Gigabit switch for the RAC interconnect.

- Network cards for public network must have same name on each participating node in the RAC cluster.
- Network cards for Interconnect network must have same name on each participating node in the RAC cluster.

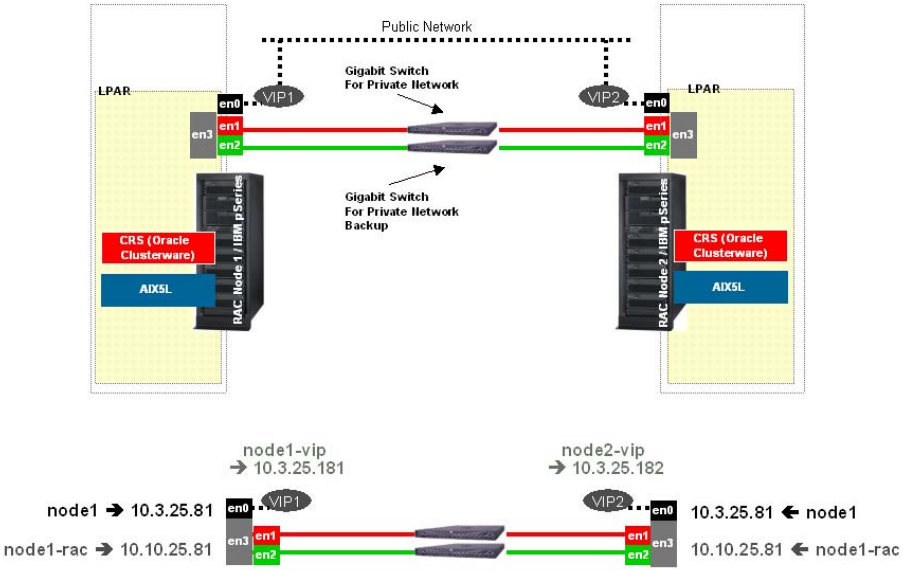
10.7.1.1 Network implementation with 2 Nodes

Minimum Network Layout for RAC/ASM Implementation !!!



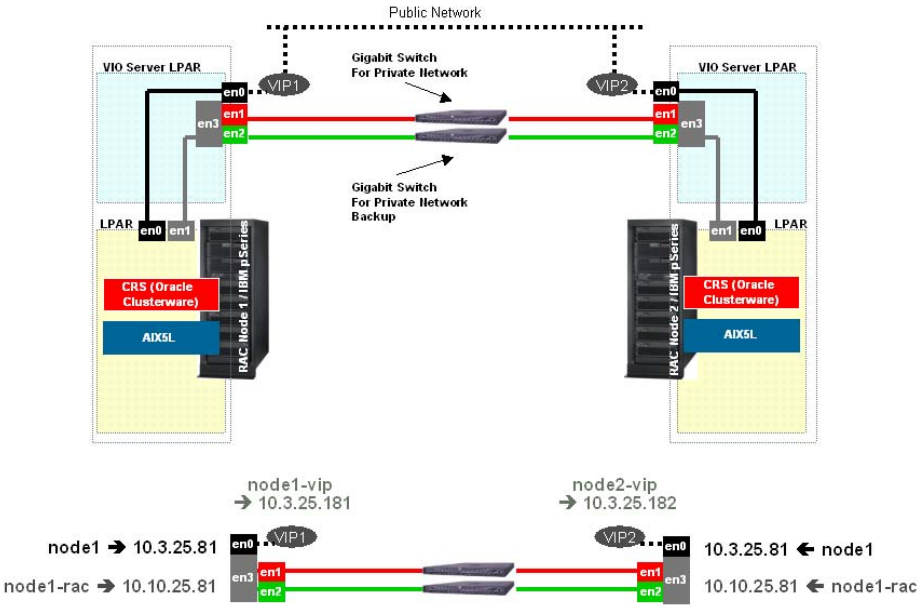
10.7.1.2 Network implementation with 2 Nodes, and Eterchannel

Network Layout for RAC/ASM Implementation Using AIX Etherchannel !!!



10.7.1.3 Network implementation with 2 Nodes, Eterchannel and VI/O Server

Network Layout for RAC/ASM Implementation Using AIX Etherchannel and VIO Server !!!



10.7.2 Configuration steps to follow

10.7.2.1 Nodes Name and IP Identification

Please make a table as follow to have a clear view of your RAC network architecture :

PUBLIC NODE NAME MUST BE NAME RETURNED BY "hostname" AIX COMMAND

Issue the AIX command "hostname" on each node to identify default node name :

```
{node1:root}/ ->hostname
node1

{node1:root}/ -> ping node1
PING node1: (10.3.25.81): 56 data bytes
64 bytes from 10.3.25.81: icmp_seq=0 ttl=255 time=0 ms
64 bytes from 10.3.25.81: icmp_seq=1 ttl=255 time=0 ms
64 bytes from 10.3.25.81: icmp_seq=2 ttl=255 time=0 ms
^C
----node1 PING Statistics----
3 packets transmitted, 3 packets received, 0% packet
loss
round-trip min/avg/max = 0/0/0 ms
{node1:root}/ ->
```

```
{node2:root}/ ->hostname
node2

{node2:root}/ -> ping node2
PING node2: (10.3.25.82): 56 data bytes
64 bytes from 10.3.25.82: icmp_seq=0 ttl=255 time=0 ms
64 bytes from 10.3.25.82: icmp_seq=1 ttl=255 time=0 ms
64 bytes from 10.3.25.82: icmp_seq=2 ttl=255 time=0 ms
^C
----node2 PING Statistics----
3 packets transmitted, 3 packets received, 0% packet
loss
round-trip min/avg/max = 0/0/0 ms
{node2:root}/ ->
```

Public		VIP (Virtual IP)		RAC Interconnect (Private Network)		RAC Interconnect backup (if no etherchannel implemented)	
en ?				en ?		en ?	
Node Name	IP	Node Name	IP	Node Name	IP	Node Name	IP
node1	**10.3.25.81**						
node2	**10. 3.25.82**						

Oracle clusterware VIP's IP adress and corresponding nodes names must not be used on the network prior to Oracle Clusterware installation. Don't make any AIX allias on the public network interface, the clusterware installation will do it. Just reserve 1 VIP and it's hostname per RAC node.

Public		VIP		RAC Interconnect (Private Network)		RAC Interconnect backup (if no etherchannel implemented)	
en ?				en ?		en ?	
Node Name	IP	Node Name	IP	Node Name	IP	Node Name	IP
node1	10.3.25.81	**node1-vip**	**10.3.25.181**				
node2	10. 3.25.82	**node2-vip**	**10. 3.25.182**				

Oracle Clusterware VIP's IP and corresponding nodes names can be declared in the DNS, or at minimum in the local hosts file.

10.7.2.2 Network Card Identification

As root, Issue the AIX command "ifconfig –l" to list network card on each node :

Result from node1 :

```
{node1:root}/ -> ifconfig -l
en0 en1 en2 lo0
{node1:root}/ ->
```

Result from node2 :

```
{node2:root}/ -> ifconfig -l
en0 en1 en2 lo0
{node2:root}/ ->
```

As root, Issue the following shell to get necessary information from network interfaces on each node :

Result from node1 :

```
{node1:root}/ -> for i in en0 en1 en2
do
  echo $i
  for attribut in netaddr netmask broadcast state
     do
        lsattr -El $i -a $attribut
     done
done
en0
netaddr 10.3.25.81 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
en1
netaddr 10.10.25.81 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
en2
netaddr 20.20.25.81 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
{node1:root}/ ->
```

Result from node2 :

```
{node2:root}/ -> for i in en0 en1 en2
do
  echo $i
  for attribut in netaddr netmask broadcast state
     do
        lsattr -El $i -a $attribut
     done
done
en0
netaddr 10.3.25.82 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
en1
netaddr 10.10.25.82 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
en2
netaddr 20.20.25.82 Internet Address True
netmask 255.255.255.0 Subnet Mask True
broadcast  Broadcast Address True
state up Current Interface Status True
{node2:root}/ ->
```

- **en0 is set to be the "Public Network Interface" on all nodes.**
- **en0 is set to be the "VIP Network Interface" on all nodes.**
- **en1 is set to be the "Private Network Interface", also named as "RAC Interconnect" on all nodes.**
- **en2 is set to be the "Private Network Interface Backup", also named as "RAC Interconnect Backup" on all nodes.**

Public		VIP		RAC Interconnect (Private Network)		RAC Interconnect backup (if no etherchannel implemented)	
en0				en1		en2	
Node Name	IP	Node Name	IP	Node Name	IP	Node Name	IP
node1	10.3.25.81	node1-vip	10.3.25.181				
node2	10. 3.25.82	node2-vip	10. 3.25.182				

As root, Issue the AIX command "ifconfig –a" to list network card on each node :

Result example from node1 :

```
{node1:root}/ ->ifconfig –a
en0:
flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.3.25.81 netmask 0xffffff00 broadcast 10.3.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en1:
flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.10.25.81 netmask 0xffffff00 broadcast 10.10.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en2:
flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 20.20.25.81 netmask 0xffffff00 broadcast 20.20.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
lo0: flags=e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT>
        inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
        inet6 ::1/0
         tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1
{node1:root}/ ->
```

To see details on the network interface :

```
{node1:root}/oracle -> lsattr -El en0
alias4                      IPv4 Alias including Subnet Mask           True
alias6                      IPv6 Alias including Prefix Length         True
arp           on            Address Resolution Protocol (ARP)          True
authority                   Authorized Users                           True
broadcast                   Broadcast Address                          True
mtu           1500          Maximum IP Packet Size for This Device     True
netaddr       10.3.25.81    Internet Address                           True
netaddr6                    IPv6 Internet Address                      True
netmask       255.255.255.0 Subnet Mask                                True
prefixlen                   Prefix Length for IPv6 Internet Address    True
remmtu        576           Maximum IP Packet Size for REMOTE Networks True
rfc1323                     Enable/Disable TCP RFC 1323 Window Scaling True
security      none          Security Level                             True
state         up            Current Interface Status                   True
tcp_mssdflt                 Set TCP Maximum Segment Size               True
tcp_nodelay                 Enable/Disable TCP_NODELAY Option          True
tcp_recvspace               Set Socket Buffer Space for Receiving      True
tcp_sendspace               Set Socket Buffer Space for Sending        True
{node1:root}/oracle ->
```

HEN, we will get the following table with our system :

Public		VIP		RAC Interconnect (Private Network)		RAC Interconnect backup (if no etherchannel implemented)	
en0				en1		en2	
Node Name	IP	Node Name	IP	Node Name	IP	Node Name	IP
node1	10.3.25.81	node1-vip	10.3.25.181	**node1-rac**	**10.10.25.81**	**node1-rac-b**	**20.20.25.181**
node2	10. 3.25.82	node2-vip	10. 3.25.182	**node1-rac**	**10.10.25.82**	**node1-rac-b**	**20.20.25.182**

10.7.2.3 Host file Setup

You should have the entries on each node for : /etc/hosts, /etc/hosts.equiv and on root/oracle home directory $HOME/.rhosts.

Update/check entries in hosts file on each node

node1:root-/> **pg /etc/hosts** ↵

```
# Public Network
10.3.25.81          node1
10.3.25.82          node2

# Virtual IP address
10.3.25.181         node1-vip
10.3.25.182         node2-vip

# Interconnect RAC
10.10.25.81         node1-rac
10.10.25.82         node2-rac

# Interconnect RAC Backup
10.10.25.81         node1-rac-b
10.10.25.82         node2-rac-b
```

NOTA :

RAC Interconnect should be protected :

- having 2 links, 1 active (en1 for our case), and 1 passive as failover backup (en2 for our case) configured in the CRS, thru oicfg command.

Or

- having 2 links, and AIX etherchanel implemented on top of the 2 links (en1 and en2 for our case).

10.7.2.4 Defining Default gateway on public interface

Using “route add”, do **set the default gateway on public network interface**, en0 in our case, and on all nodes :

MUST BE DONE on each node !!!

To establish a default gateway, type:

```
{node1:root}/ -> route add 0 10.3.25.254
{node2:root}/ -> route add 0 10.3.25.254
```

The value 0 or the default keyword for the Destination parameter means that any packets sent to destinations not previously defined and not on a directly connected network go through the default gateway. The 10.3.25.254 address is that of the gateway chosen to be the default.

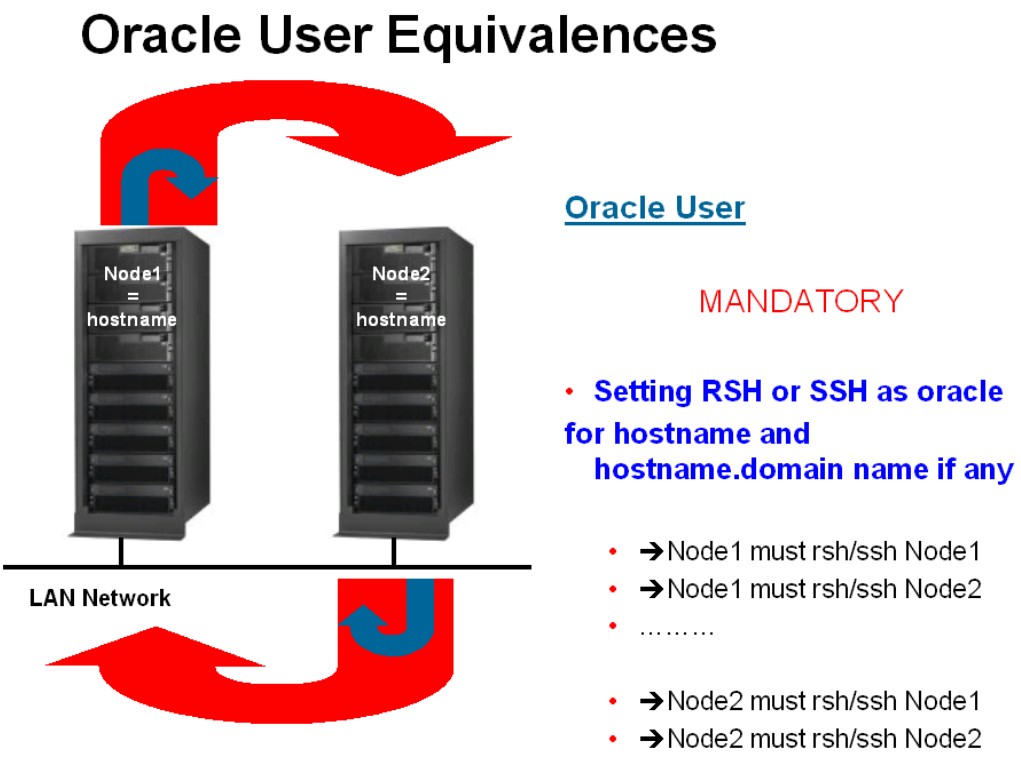
To check if default gateway is set :

Use “netstat -r” On both nodes.

```
{node1:root}/ -> netstat -r
Routing tables
Destination      Gateway           Flags   Refs     Use  If   Exp  Groups

Route Tree for Protocol Family 2 (Internet):
default          10.3.25.254       UG        1     73348 en0      -      -   =>
default          9.212.131.254     UG        0       435 en0      -      -
10.3.25.0        node1             UHSb      0         0 en0      -      -   =>
10.3.25/24       node1             U        11   6113590 en0      -      -
node1            loopback          UGHS     37   1034401 lo0      -      -
node1-vip        loopback          UGHS      8     80831 lo0      -      -
10.3.25.255      node1             UHSb      0         4 en0      -      -
10.10.25.0       node1-rac         UHSb      0         0 en1      -      -   =>
10.10.25/24      node1-rac         U        25    350557 en1      -      -
node1-rac        loopback          UGHS     16       481 lo0      -      -
10.10.25.255     node1-rac         UHSb      0         4 en1      -      -
20.20.25.0       node1-rac-b       UHSb      0         0 en2      -      -   =>
20.20.25/24      node1-rac-b       U        16    176379 en2      -      -
node1-rac-b      loopback          UGHS      5       392 lo0      -      -
20.20.25.255     node1-rac-b       UHSb      0         4 en2      -      -
127/8            loopback          U        49    187105 lo0      -      -

Route Tree for Protocol Family 24 (Internet v6):
::1              ::1               UH        0        32 lo0      -      -
{node1:root}/ ->
```

10.7.2.5 User Equivalence Setup

Before installing Oracle Real Application clusters, you must configure user equivalence for the **oracle** user on all cluster nodes.



You have two type of User Equivalence implementation :

- RSH (Remote shell)
- SSH (Secured Shell)

> **When SSH is not available, the Installer uses the `rsh` and `rcp` commands instead of `ssh` and `scp`. .**

You have to choose one or the other, but don't implement both at the same time.

➔ Usually, customers will implement SSH.
AND if SSH is started and used, do configure SSH

On the public node name returned by the AIX command "hostname" :

node1 must ssh or rsh to node1, as oracle user.
node1 must ssh or rsh to node2, as oracle user.
node2 must ssh or rsh to node1, as oracle user.
node2 must ssh or rsh to node2, as oracle user.

10.7.2.5.1 RSH implementation

Set up user equivalence for the oracle and root account, to enable rsh, rcp, rlogin commands.

You should have the entries on each node for : /etc/hosts, /etc/hosts.equiv and on root/oracle home directory $HOME/.rhosts.

/etc/hosts.equiv

Update/check entries in hosts.equiv file on each node

```
node1:root-/> pg /etc/hosts.equiv ↵

node1        root
node2        root
node1        oracle
node2        oracle
```

$HOME/.rhosts

Update/check entries in .rhosts file on each node for root user :

```
node1:root-/> su –  root
node1:root-/ > cd
node1:root-/> pg$HOME/.rhosts
node1        root
node2        root
node1        oracle
node2        oracle
```

Update/check entries in .rhosts file on each node for oracle user :

```
node1:root-/> su – oracle
node1:oracle-/home/oracle>  cd
node1:oracle-/home/oracle>pg  $HOME/.rhosts
node1        oracle
node2        oracle
```

Note : It is possible, but not advised because of security reasons, to put a "+" in hosts.equiv and .rhosts files.

Test if the user equivalence is correctly set up (node2 is the secondary cluster machine).

You are logged on node1 as root :

```
node1:root-/>  rsh node2 (=> no password)
node2:root-/>  rcp /tmp/toto node1:/tmp/toto
node2:root-/> su - oracle
node2 :oracle-/home/oracle> rsh node1 date

Mon Apr 23 17:26:27 DFT 2007
node2 :oracle-/home/oracle>
```

10.7.2.5.2 SSH implementation

Before you install and use Oracle Real Application clusters, you must configure secure shell (SSH) for the **oracle** user on all cluster nodes. Oracle Universal Installer uses the **ssh** and **scp** commands during installation to run remote commands on and copy files to the other cluster nodes. You must configure SSH so that these commands do not prompt for a password.

Note:

This section describes how to configure OpenSSH version 3. **If SSH is not available, then Oracle Universal Installer attempts to use rsh and rcp instead**.
To determine if SSH is running, enter the following command:

```
$ ps -ef | grep sshd
```

If SSH is running, then the response to this command is process ID numbers. To find out more about SSH, enter the following command:

```
$ man ssh
```

Configuring SSH on Cluster Member Nodes

To configure SSH, you must first create RSA and DSA keys on each cluster node, and then copy the keys from all cluster node members into an authorized keys file on each node. To do this task, complete the following steps:

Create RSA and DSA keys on each node: Complete the following steps on each node:

1. Log in as the **oracle** user.
2. If necessary, create the **.ssh** directory in the **oracle** user's home directory and set the correct permissions on it:
   ```
   $ mkdir ~/.ssh
   $ chmod 700 ~/.ssh
   ```
3. Enter the following commands to generate an RSA key for version 2 of the SSH protocol:
   ```
   $ /usr/bin/ssh-keygen -t rsa
   ```
4. At the prompts:
   Accept the default location for the key file.
   Enter and confirm a pass phrase that is different from the **oracle** user's password.
   This command writes the public key to the ~/.ssh/id_rsa.pub file and the private key to the ~/.ssh/id_rsa file. Never distribute the private key to anyone.

   Enter the following commands to generate a DSA key for version 2 of the SSH protocol:
   ```
   $ /usr/bin/ssh-keygen -t dsa
   ```
5. At the prompts:
   Accept the default location for the key file
   Enter and confirm a pass phrase that is different from the **oracle** user's password
   This command writes the public key to the ~/.ssh/id_dsa.pub file and the private key to the ~/.ssh/id_dsa file. Never distribute the private key to anyone.

Add keys to an authorized key file: Complete the following steps:

1. On the local node, determine if you have an authorized key file (**~/.ssh/authorized_keys**). If the authorized key file already exists, then proceed to step 2. Otherwise, enter the following commands:

   $ touch ~/.ssh/authorized_keys
   $ cd ~/.ssh
   $ ls

   You should see the **id_dsa.pub** and **id_rsa.pub** keys that you have created.

2. Using SSH, copy the contents of the **~/.ssh/id_rsa.pub** and **~/.ssh/id_dsa.pub** files to the file **~/.ssh/authorized_keys**, and provide the Oracle user password as prompted. This process is illustrated in the following syntax example with a two-node cluster, with nodes node1 and node2, where the Oracle user path is **/home/oracle:**

   ```
   [oracle@node1 .ssh]$ ssh node1 cat /home/oracle/.ssh/id_rsa.pub >> authorized_keys
   oracle@node1's password:
   [oracle@node1 .ssh]$ ssh node1 cat /home/oracle/.ssh/id_dsa.pub >> authorized_keys
   [oracle@node1 .ssh$ ssh node2 cat /home/oracle/.ssh/id_rsa.pub >> authorized_keys
   oracle@node2's password:
   [oracle@node1 .ssh$ ssh node2 cat /home/oracle/.ssh/id_dsa.pub >>authorized_keys
   oracle@node2's password:
   ```

3. Use SCP (Secure Copy) or SFTP (Secure FTP) to copy the **authorized_keys** file to the Oracle user .ssh directory on a remote node. The following example is with SCP, on a node called node2, where the Oracle user path is **/home/oracle:**
   [oracle@node1 .ssh]scp authorized_keys node2:/home/oracle/.ssh/

4. Repeat step 2 and 3 for each cluster node member. When you have added keys from each cluster node member to the authorized_keys file on the last node you want to have as a cluster node member, then use SCP to copy the complete authorized_keys file back to each cluster node member

   > **Note:**
   >
   > The Oracle user's **/.ssh/authorized_keys** file on every node must contain the contents from all of the **/.ssh/id_rsa.pub** and **/.ssh/id_dsa.pub** files that you generated on all cluster nodes.

5. Change the permissions on the Oracle user's **/.ssh/authorized_keys** file on all cluster nodes:

   $ chmod 600 ~/.ssh/authorized_keys

Note:

Repeat this process for each node in the cluster !!!

At this point, if you use ssh to log in to or run a command on another node, you are prompted for the pass phrase that you specified when you created the DSA key.

Enabling SSH User Equivalency on Cluster Member Nodes

To enable Oracle Universal Installer to use the ssh and scp commands without being prompted for a pass phrase,

follow these steps:

1. **On the system where you want to run Oracle Universal Installer, log in as the oracle user.**

2. **Enter the following commands:**
   ```
   $ exec /usr/bin/ssh-agent $SHELL
   $ /usr/bin/ssh-add
   ```

3. **At the prompts, enter the pass phrase for each key that you generated.**

If you have configured SSH correctly, then you can now use the **ssh** or **scp** commands without being prompted for a password or a pass phrase.

4. **If you are on a remote terminal**, and the local node has only one visual (which is typical), then use the following syntax to set the DISPLAY environment variable:

   Bourne, Korn, and Bash shells
   ```
   $ export DISPLAY=hostname:0
   ```

   C shell:
   ```
   $ setenv DISPLAY 0
   ```

For example, if you are using the Bash shell, and if your hostname is node1, then enter the following command:
```
$ export DISPLAY=node1:0
```

5. **To test the SSH configuration**, enter the following commands from the same terminal session, testing the configuration of each cluster node, where nodename1, nodename2, and so on, are the names of nodes in the cluster:
   ```
   $ ssh nodename1 date
   $ ssh nodename2 date
   ```

These commands should display the date set on each node.
If any node prompts for a password or pass phrase, then verify that the **~/.ssh/authorized_keys** file on that node contains the correct public keys.
If you are using a remote client to connect to the local node, and you see a message similar to "Warning: No xauth data; using fake authentication data for X11 forwarding," then this means that your authorized keys file is configured correctly, but your ssh configuration has X11 forwarding enabled. To correct this, proceed to step 6.

Note:
The first time you use SSH to connect to a node from a particular system, you may see a message similar to the following:

```
The authenticity of host 'node1 (140.87.152.153)' can't be established.
RSA key fingerprint is 7z:ez:e7:f6:f4:f2:4f:8f:9z:79:85:62:20:90:92:z9.
Are you sure you want to continue connecting (yes/no)?
```

Enter **yes** at the prompt to continue. You should not see this message again when you connect from this system to that node.
If you see any other messages or text, apart from the date, then the installation can fail. Make any changes required to ensure that only the date is displayed when you enter these commands.
You should ensure that any parts of login scripts that generate any output, or ask any questions, are modified so that they act only when the shell is an interactive shell.

6. **To ensure that X11 forwarding will not cause the installation to fail, create a**

user-level SSH client configuration file for the Oracle software owner user, as follows:

a. Using any text editor, edit or create the **~oracle/.ssh/config** file.

b. Make sure that the ForwardX11 attribute is set to no. For example:
Host *
ForwardX11 no

7. **You must run Oracle Universal Installer from this session or remember to repeat steps 2 and 3 before you start Oracle Universal Installer from a different terminal session.**

Preventing Oracle Clusterware Installation Errors Caused by stty Commands

During an Oracle Clusterware installation, Oracle Universal Installer uses SSH (if available) to run commands and copy files to the other nodes. During the installation, hidden files on the system (for example, .bashrc or .cshrc) will cause installation errors if they contain stty commands.

To avoid this problem, you must modify these files to suppress all output on STDERR, as in the following examples:

- **Bourne, Bash, or Korn shell:**

```
if [ -t 0 ]; then
   stty intr ^C
fi
```

- **C shell:**

```
test -t 0
if ($status == 0) then
   stty intr ^C
endif
```

10.8 ORACLE ENVIRONMENT SETUP

Oracle environment : **$HOME/.profile** file in Oracle's home directory

To be done on each node.

```
export ORACLE_BASE=/oracle
export AIXTHREAD_SCOPE=S        (S for system-wide thread scope)
export TEMP=/tmp
export TMP=/tmp
export TMPDIR=/tmp
umask 022
```

Notes:

On AIX, when using multithreaded applications or LAN-free, especially when running on machines with multiple CPUs, we strongly recommend setting AIXTHREADSCOPE=S in the environment before starting the application, for better performance and more solid scheduling.

For example:

EXPORT AIXTHREADSCOPE=S

Setting AIXTHREAD_SCOPE=S means that user threads created with default attributes will be placed into system-wide contention scope. If a user thread is created with system-wide contention scope, it is bound to a kernel thread and it is scheduled by the kernel. The underlying kernel thread is not shared with any other user thread.

AIXTHREAD_SCOPE (AIX 4.3.1 and later)

Purpose:
Controls contention scope. P signifies process-based contention scope (M:N). S signifies system-based contention scope (1:1).

Values:
Possible Values: P or S. The default is P.

Display:
echo $AIXTHREAD_SCOPE (this is turned on internally, so the initial default value will not be seen with the echo command)

Change:
AIXTHREAD_SCOPE={P|S}export AIXTHREAD_SCOPE Change takes effect immediately in this shell. Change is effective until logging out of this shell. Permanent change is made by adding the AIXTHREAD_SCOPE={P|S} command to the `/etc/environment` file.

Diagnosis:
If fewer threads are being dispatched than expected, then system scope should be tried.

10.9 LOCAL DISK FOR ORACLE CODE (ORACLE CLUSTERWARE AND RAC SOFTWARE)

The oracle code can be located on an internal disk and propagated on the other machines of the cluster. The Oracle Universal Installer manage the cluster-wide installation, that is done only once. Regular file systems are used for Oracle code.

NOTA : You can also use virtual I/O disks for :

- AIX5L operating system
- Oracle clusterware ($CRS_HOME)
- RAC Software ($ORACLE_HOME)

On both nodes, create the file system for Oracle code. This file system of 6 GB, is generally located on an internal disk.

To list the internal disks :

On node 1 …

```
node1:root-/>  lsdev –Ccdisk | grep SCSI
hdisk0  Available        Virtual SCSI Disk Drive
```

On node 2 …

Create a volume group called oraclevg :

On node 1 …

```
node1:root-/>  mkvg -f -y'oraclevg' -s'32' hdisk0
```

On node 2 …

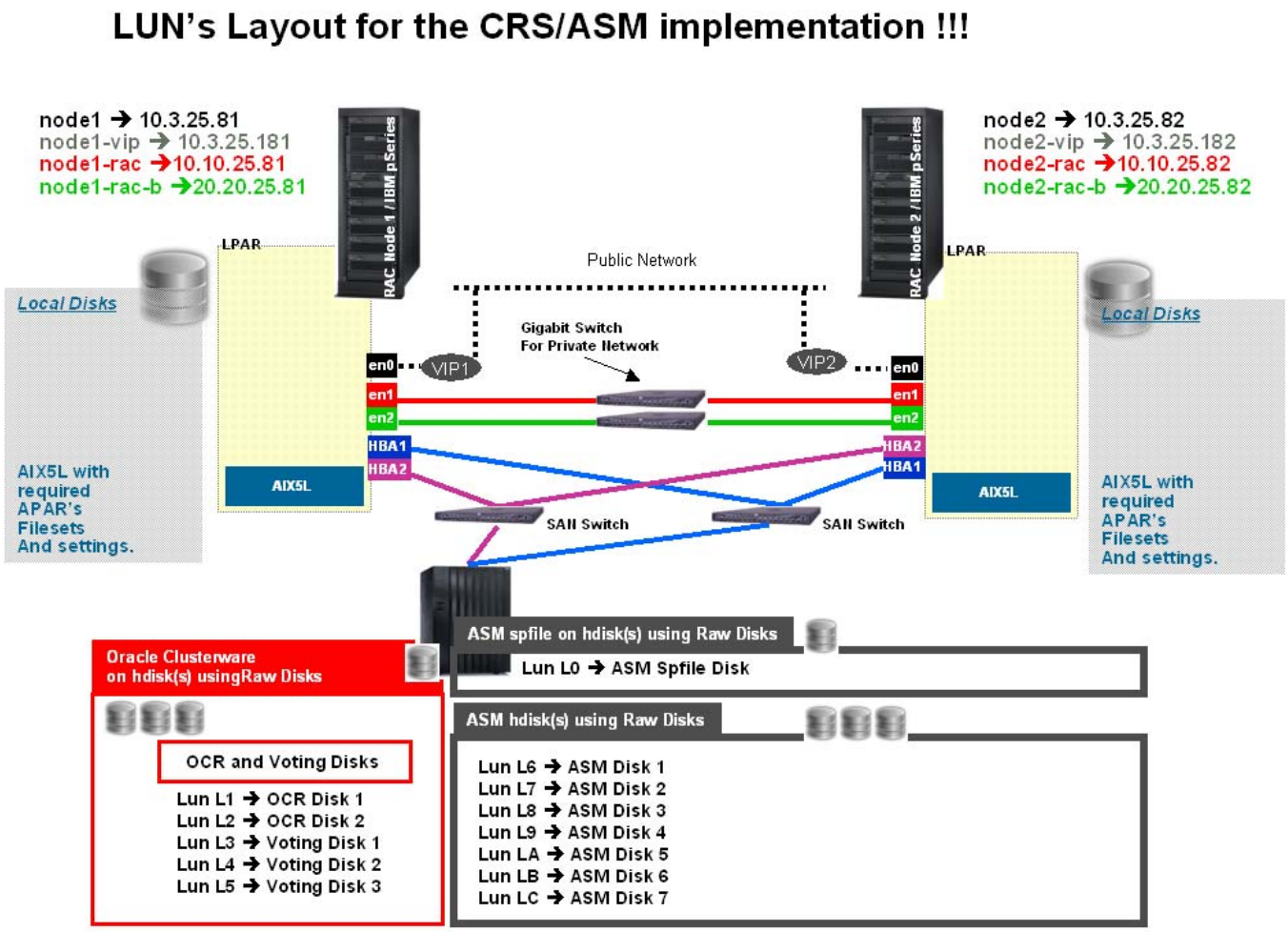
Create a 12GB file system /oracle in the previous volume group (large file enabled) :

On node 1 …

```
node1:root-/> crfs -v jfs2 -a bf=true -g'oraclevg' -a size='12582912' -m'/oracle' -A'yes' -
p'rw' -t'no' -a nbpi='8192' -a ag='64'

node1:root-/>  mount /oracle
node1:root-/>  chown oracle:dba /oracle
```

On node 2 …

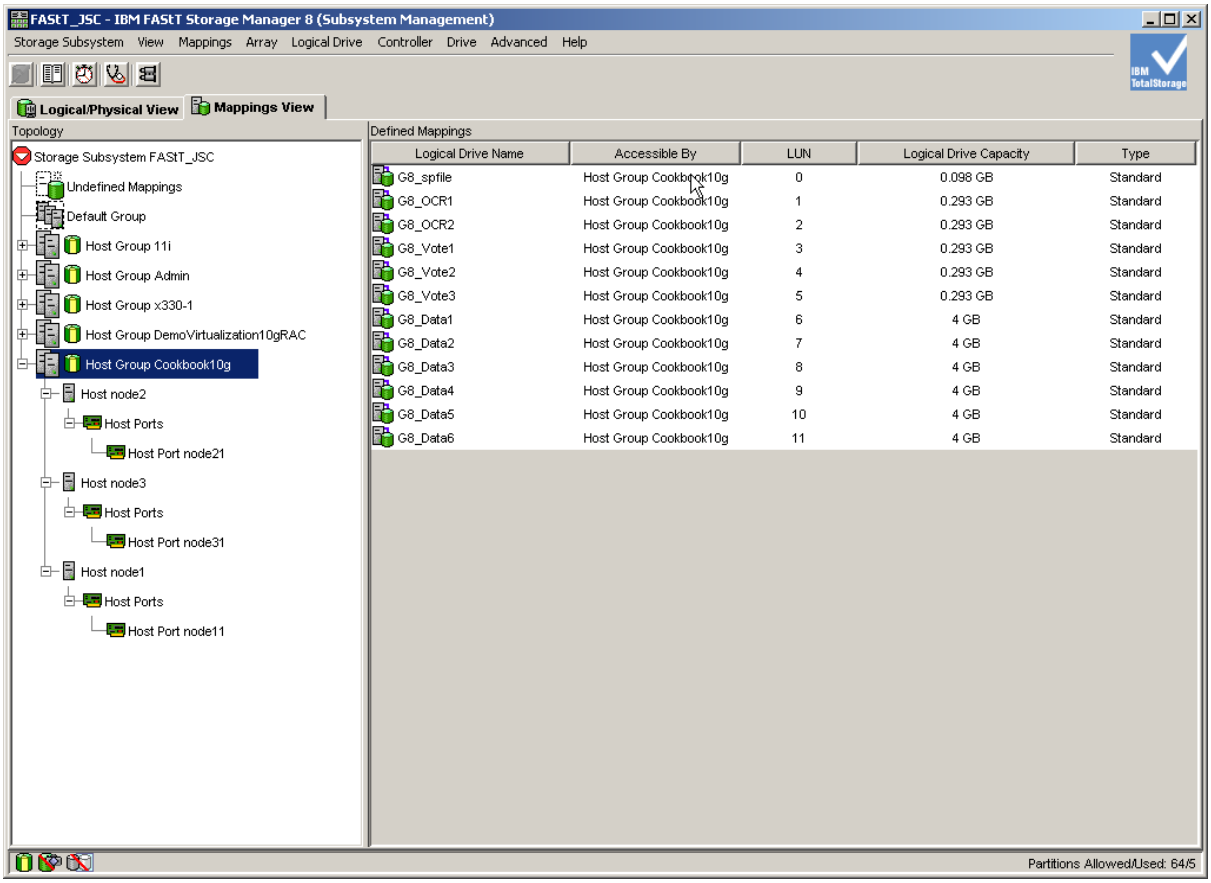
10.10 ASM Implementation



10.10.1 LUN's creation

Using the storage administration console, you have to create :

- **LUN's for OCR and Voting Disks**
- LUN's for disks to be used with ASM, for DiskGroup Creation or give physical disks to ASM



The following screen shows the LUN mapping for nodes used in our cluster. The LUN for OCR disks and Voting disks have ids 2 to 6. These ID's will help us on to identify which hdisk will be used.

Disks	LUN's ID Number	LUN's Size
ASM spfile	L**0**	100 MB
OCR1	L**1**	300 MB
OCR2	L**2**	300 MB
Voting1	L**3**	300 MB
Voting2	L**4**	300 MB
Voting3	L**5**	300 MB
Disk 1 for ASM	L**6**	4 GB
Disk 2 for ASM	L**7**	4 GB
Disk 3 for ASM	L**8**	4 GB
Disk 4 for ASM	L**9**	4 GB
Disk 5 for ASM	L**A (meaning Lun 10)**	4 GB
Disk 6 for ASM	L**B (meaning Lun 11)**	4 GB
Disk 7 for ASM	L**C (meaning Lun 12)**	4 GB

10.10.2 Register LUN's at AIX level

Before registration of LUN's :

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
{node1:root}/ ->
```

On node 2 …

As root on each node, update the ODM repository using the following command : "**cfgmgr**"

You need to register and identify LUN's at AIX level, and LUN's will be mapped to hdisk and registered in the AIX ODM.

After registration of LUN's in ODM thru cfgmgr command :

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
hdisk11         none                                None
hdisk12         none                                None
hdisk13         none                                None
{node1:root}/ ->
```

On node 2 …

When hdisk is not marked rootvg, but None, it's imortant to check that it's not used at all.

Use the following command to read hdisk header :

On node 1 … for hdisk2

```
{node1:root}/ ->lquerypv -h /dev/rhdisk2
00000000   00000000 00000000 00000000 00000000  |................|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
{node1:root}/ ->
```

On node 2 …

If all lines are ONLY full of "0"
THEN the hdisk is free to be used.

If it's not the case, check first that which LUN is mapped to the hdisk (next pages), and if it's the LUN you should use, you must then "dd" (zeroing) the /dev/rhdisk2 rdisk

On next step, we will need to identify which hdisk is mapped to which LUN in the ODM repository.

10.10.3 Preparing Raw Disks for CRS/Voting disks

We know the LUN's available for OCR and Voting disks are L3 to L7.

Disks	LUN's ID Number
OCR 1	L1
OCR 2	L2
Voting 1	L3
Voting 2	L4
Voting 3	L5

Identify the disks available for OCR and Voting disks, on each node, knowing the LUN's numbers.

Knowing the LUN's number to use, we know need to identify the corresponding hdisks on each node of the cluster as detailed in the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node …. Corresponding hdisk
OCR 1	L1			
OCR 2	L2			
Voting 1	L3			
Voting 2	L4			
Voting 3	L5			

There are two methods to identify the corresponding hdisks :

- **Identify LUN ID assign to hdisk, using "lscfg –l hdisk?" command**
- **Identify hdisks by assigning momently a PVID to each hdisk not having one**

We strongly recommend not using PVID to identify hdisk.

If some hdisks are already used, setting a PVID to a used hdisk can corrupt the the hdisk header, and have generate issues as loosing data stored on the hdisk.

Get the list of hdisk available on each node :

On node 1 …

```
{node1:root}/ -> lsdev -Ccdisk
hdisk0  Available          Virtual SCSI Disk Drive
hdisk1  Available 03-08-02 3552     (500) Disk Array Device
hdisk2  Available 03-08-02 3552     (500) Disk Array Device
hdisk3  Available 03-08-02 3552     (500) Disk Array Device
hdisk4  Available 03-08-02 3552     (500) Disk Array Device
hdisk5  Available 03-08-02 3552     (500) Disk Array Device
……
……
{node1:root}/ ->
```

On node 2 …

```
{node1:root}/ -> rsh node2 lsdev -Ccdisk
hdisk0  Available 04-08-02 3552     (500) Disk Array Device
hdisk1  Available 04-08-02 3552     (500) Disk Array Device
hdisk2  Available 04-08-02 3552     (500) Disk Array Device
hdisk3  Available 04-08-02 3552     (500) Disk Array Device
hdisk4  Available 04-08-02 3552     (500) Disk Array Device
hdisk5  Available 04-08-02 3552     (500) Disk Array Device
……
……
{node1:root}/ ->
```

Get the List of the hdisks on node1 :

List of available hdisks on node 1 :

rootvg is hdisk0 !!!

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
hdisk11         none                                None
hdisk12         none                                None
hdisk13         none                                None
{node1:root}/ ->
```

➔ **NO PVID are assigned apart for the rootvg hdisk**

Using lscfg command, try to identify the hdisks in the list generated by lspv on node1 :

Identify LUN ID assign to hdisk, using "lscfg –vl hdisk?" command

On node 1 …

```
{node1:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
 do
 lscfg -vl hdisk$i
 done
  hdisk0    U9117.570.10ED22C-V5-C5-T1-L810000000000  Virtual SCSI Disk Drive
  hdisk1    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L0  3552     (500) Disk Array Device
  hdisk2    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L1000000000000  3552     (500) Disk Array Device
  hdisk3    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L2000000000000  3552     (500) Disk Array Device
  hdisk4    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L3000000000000  3552     (500) Disk Array Device
  hdisk5    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L4000000000000  3552     (500) Disk Array Device
  hdisk6    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L5000000000000  3552     (500) Disk Array Device
  hdisk7    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L6000000000000  3552     (500) Disk Array Device
  hdisk8    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L7000000000000  3552     (500) Disk Array Device
  hdisk9    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L8000000000000  3552     (500) Disk Array Device
  hdisk10   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L9000000000000  3552     (500) Disk Array Device
  hdisk11   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LA000000000000  3552     (500) Disk Array Device
  hdisk12   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LB000000000000  3552     (500) Disk Array Device
  hdisk13   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LC000000000000  3552     (500) Disk Array Device
{node1:root}/ ->
```

THEN, We get the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node …. Corresponding hdisk
OCR 1	L1	hdisk2		hdisk?
OCR 2	L2	hdisk3		
Voting 1	L3	hdisk4		hdisk?
Voting 2	L4	hdisk5		
Voting 3	L5	hdisk6		

➔ **No need to assign PVID when using this method.**

Get the List of the hdisks on node2

List of available hdisks on node 2 :

rootvg is hdisk11, whichis not same hdisk as on node 1, which means that it could be the same for each hdisk. On both nodes, same hdisk names might not be attached to same LUN.

On node 2 …

```
{node2:root}/ -> lspv
hdisk0          none                None
hdisk1          none                None
hdisk2          none                None
hdisk3          none                None
hdisk4          none                None
hdisk5          none                None
hdisk6          none                None
hdisk11         00ced22c826884d3                    rootvg          active
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
hdisk12         none                                None
hdisk13         none                                None
{node2:root}/ ->
```

➔ **NO PVID are assigned apart for the rootvg hdisk**

Using lscfg command, try to identify the hdisks in the list generated by lspv on node2 :

Be carefull hdisk2 on node1 is not necessary hdisk2 on node2.

Identify LUN ID assign to hdisk, using "lscfg –vl hdisk?" command

On node 1 …

```
{node2:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
 do
 lscfg -vl hdisk$i
 done
  hdisk0    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L0  3552     (500) Disk Array Device
  hdisk1    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L1000000000000  3552     (500) Disk Array Device
  hdisk2    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L2000000000000  3552     (500) Disk Array Device
  hdisk3    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L3000000000000  3552     (500) Disk Array Device
  hdisk4    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L4000000000000  3552     (500) Disk Array Device
  hdisk5    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L5000000000000  3552     (500) Disk Array Device
  hdisk6    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L6000000000000  3552     (500) Disk Array Device
  hdisk7    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L7000000000000  3552     (500) Disk Array Device
  hdisk8    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L8000000000000  3552     (500) Disk Array Device
  hdisk9    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L9000000000000  3552     (500) Disk Array Device
  hdisk10   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LA000000000000  3552     (500) Disk Array Device
  hdisk11   U9117.570.10ED22C-V1-C5-T1-L810000000000  Virtual SCSI Disk Drive
  hdisk12   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LB000000000000  3552     (500) Disk Array Device
  hdisk13   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LC000000000000  3552     (500) Disk Array Device
{node2:root}/ ->
```

THEN, We get the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node …. Corresponding hdisk
OCR 1	L1	hdisk2	hdisk1	hdisk?
OCR 2	L2	hdisk3	hdisk2	hdisk?
Voting 1	L3	hdisk4	hdisk3	hdisk?
Voting 2	L4	hdisk5	hdisk4	hdisk?
Voting 3	L5	hdisk6	hdisk5	hdisk?

➔ **No need to assign PVID when using this method.**

Setup reserve_policy on ocr and voting hdisks, on each node :

Example for one hdisk :

Issue the command "lsattr –E –l hdisk3" to vizualize all attributes for hdisk3

```
{node1:root}/ -> lsattr -El hdisk3
PR_key_value    none                                Persistant Reserve Key Value           True
cache_method    fast_write                          Write Caching method                   False
ieee_volname    600A0B80000C54180000026245EE3C96 IEEE Unique volume name                   False
lun_id          0x0002000000000000                  Logical Unit Number                    False
max_transfer    0x100000                            Maximum TRANSFER Size                  True
prefetch_mult   1                                   Multiple of blocks to prefetch on read False
pvid            none                                Physical volume identifier             False
q_type          simple                              Queuing Type                           False
queue_depth     10                                  Queue Depth                            True
raid_level      0                                   RAID Level                             False
reassign_to     120                                 Reassign Timeout value                 True
reserve_policy single_path                          Reserve Policy                         True
rw_timeout      30                                  Read/Write Timeout value               True
scsi_id         0x690600                            SCSI ID                                False
size            300                                 Size in Mbytes                         False
write_cache     yes                                 Write Caching enabled                  False
{node1:root}/ ->
```

Or only lsattr –E –l hdisk3 | grep reserve

```
{node1:root}/ -> lsattr -El hdisk3 | grep reserve
reserve_policy single_path                          Reserve Policy                         True
{node1:root}/ ->
```

- **On IBM storage (ESS, FasTt, DSXXXX) : Change the "reserve_policy" attribute to "no_reserve"**

  chdev -l hdisk? -a reserve_policy=no_reserve

- **On EMC storage : Change the "reserve_lock" attribute to "no"**

  chdev -l hdisk? -a reserve_lock=no

- **On HDS storage with HDLM driver, and no disks in Volume Group : Change the "dlmrsvlevel" attribute to "no_reserve"**

  chdev -l dlmfdrv? -a dlmrsvlevel=no_reserve

Change the "reserve_policy" attributes for each disks dedicated to ASM, on each nodes of the cluster :

In our case, we have an IBM storage !!!

On node 1 ...

```
{node1:root}/ -> for i in 2 3 4 5 6
 do
 chdev -l hdisk$i  -a  reserve_policy=no_reserve
 done
changed
changed
changed
changed
changed
{node1:root}/ ->
```

On node 2 ...

```
{node2:root}/ -> for i in 1 2 3 4 5
 do
 chdev -l hdisk$i  -a  reserve_policy=no_reserve
 done
changed
changed
changed
changed
changed
{node2:root}/ ->
```

Example for one hdisk on node1 :

Issue the command "lsattr –E –l hdisk3 |grep reserve" to vizualize modified attributes for hdisk3

```
{node1:root}/ -> lsattr -El hdisk3 | grep reserve
reserve_policy no_reserve                           Reserve Policy                         True
{node1:root}/ ->
```

As described before, disks might have different names from one node to another for example hdisk2 on node1 might be hdisk3 on node2, etc…

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	/dev/ocr_disk1	hdisk2			hdisk1		
OCR 2	L2	/dev/ocr_disk2	hdisk3			hdisk2		
Voting 1	L3	/dev/voting_disk1	hdisk4			hdisk3		
Voting 2	L4	/dev/voting_disk2	hdisk5			hdisk4		
Voting 3	L5	/dev/voting_disk3	hdisk6			hdisk5		

Identify minor and major number for each hdisk, on each node …

On node1 :

```
{node1:root}/ -> for i in 2 3 4 5 6
 do
 ls -la /dev/*hdisk$i
 done
brw-------   1 root     system       20,   6 Mar 07 10:31 /dev/hdisk2
crw-------   1 root     system       20,   6 Mar 07 10:31 /dev/rhdisk2
brw-------   1 root     system       20,   7 Mar 07 10:31 /dev/hdisk3
crw-------   1 root     system       20,   7 Mar 07 10:31 /dev/rhdisk3
brw-------   1 root     system       20,   8 Mar 07 10:31 /dev/hdisk4
crw-------   1 root     system       20,   8 Mar 07 10:31 /dev/rhdisk4
brw-------   1 root     system       20,   9 Mar 07 10:31 /dev/hdisk5
crw-------   1 root     system       20,   9 Mar 07 10:31 /dev/rhdisk5
brw-------   1 root     system       20,  10 Mar 07 10:31 /dev/hdisk6
crw-------   1 root     system       20,  10 Mar 07 10:31 /dev/rhdisk6
{node1:root}/ ->
```

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	/dev/ocr_disk1	hdisk2	**20**	**6**	hdisk1		
OCR 2	L2	/dev/ocr_disk2	hdisk3	**20**	**7**	hdisk2		
Voting 1	L3	/dev/voting_disk1	hdisk4	**20**	**8**	hdisk3		
Voting 2	L4	/dev/voting_disk2	hdisk5	**20**	**9**	hdisk4		
Voting 3	L5	/dev/voting_disk3	hdisk6	**20**	**10**	hdisk5		

Identify minor and major number for each hdisk, on each node …

On node2 :

```
{node2:root}/ -> for i in 1 2 3 4 5
 do
 ls -la /de> do
 ls -la /dev/*hdisk$i
 done
brw-------   1 root     system       20,   6 Mar 07 10:32 /dev/hdisk1
crw-------   1 root     system       20,   6 Mar 07 10:32 /dev/rhdisk1
brw-------   1 root     system       20,   7 Mar 07 10:32 /dev/hdisk2
crw-------   1 root     system       20,   7 Mar 07 10:32 /dev/rhdisk2
brw-------   1 root     system       20,   8 Mar 07 10:32 /dev/hdisk3
crw-------   1 root     system       20,   8 Mar 07 10:32 /dev/rhdisk3
brw-------   1 root     system       20,   9 Mar 07 10:32 /dev/hdisk4
crw-------   1 root     system       20,   9 Mar 07 10:32 /dev/rhdisk4
brw-------   1 root     system       20,  10 Mar 07 10:32 /dev/hdisk5
crw-------   1 root     system       20,  10 Mar 07 10:32 /dev/rhdisk5
{node2:root}/ ->
```

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	/dev/ocr_disk1	hdisk2	20	6	hdisk1	**20**	**6**
OCR 2	L2	/dev/ocr_disk2	hdisk3	20	7	hdisk2	**20**	**7**
Voting 1	L3	/dev/voting_disk1	hdisk4	20	8	hdisk3	**20**	**8**
Voting 2	L4	/dev/voting_disk2	hdisk5	20	9	hdisk4	**20**	**9**
Voting 3	L5	/dev/voting_disk3	hdisk6	20	10	hdisk5	**20**	**10**

As disks have different names from one node to another for example L1 correspond to hdisk2 on node1, and hdisk1 on node2, etc…

THEN from the table, for node1 we need to create virtual devices :

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	**/dev/ocr_disk1**	hdisk2	**20**	**6**	hdisk1	20	6
OCR 2	L2	**/dev/ocr_disk2**	hdisk3	**20**	**7**	hdisk2	20	7
Voting 1	L3	**/dev/voting_disk1**	hdisk4	**20**	**8**	hdisk3	20	8
Voting 2	L4	**/dev/voting_disk2**	hdisk5	**20**	**9**	hdisk4	20	9
Voting 3	L5	**/dev/voting_disk3**	hdisk6	**20**	**10**	hdisk5	20	10

To create same virtual devices on each node called :

/dev/ocr_disk1
/dev/ocr_disk2
/dev/voting_disk1
/dev/voting_disk2
/dev/voting_disk3

we need to use major and minor number of hdisks which will make the link between the virtual devices and the hdisks.

Using the command : mknod Device_Name c MajNum MinNum

For first node, as root user

```
node1:root-/> mknod /dev/ocr_disk1 c 20 6
node1:root-/> mknod /dev/ocr_disk2 c 20 7
node1:root-/> mknod /dev/voting_disk1 c 20 8
node1:root-/> mknod /dev/voting_disk2 c 20 9
node1:root-/> mknod /dev/voting_disk3 c 20 10
```

THEN set ownership of the created virtual devices to oracle:dba

```
node1:root-/> chown oracle:dba /dev/ocr_disk1
node1:root-/> chown oracle:dba /dev/ocr_disk2
node1:root-/> chown oracle:dba /dev/voting_disk1
node1:root-/> chown oracle:dba /dev/voting_disk2
node1:root-/> chown oracle:dba /dev/voting_disk3
```

THEN set read/write permissions of the created virtual devices to 660

```
node1:root-/> chmod 660 /dev/ocr_disk1
node1:root-/> chmod 660 /dev/ocr_disk2
node1:root-/> chmod 660 /dev/voting_disk1
node1:root-/> chmod 660 /dev/voting_disk2
node1:root-/> chmod 660 /dev/voting_disk3
```

Checking the modifications

After Oracle clusterware installation,

"crw-rw-r–" will be changed to **"crw-r--r–"** for the OCR and voting virtual devices.

And **"oracle dba"** ownership will be changed to **"root dba"** for the OCR virtual devices.

```
{node1:root}/ -> ls -la /dev/* | grep "20,  6"
brw-------   1 root     system       20,  6 Mar 07 10:31 /dev/hdisk2
crw-rw-r--   1 root     dba          20,  6 Mar 12 15:03 /dev/ocr_disk1
crw-------   1 root     system       20,  6 Mar 07 10:31 /dev/rhdisk2
{node1:root}/ -> ls -la /dev/* | grep "20,  7"
brw-------   1 root     system       20,  7 Mar 07 10:31 /dev/hdisk3
crw-rw-r--   1 root     dba          20,  7 Mar 12 15:03 /dev/ocr_disk2
crw-------   1 root     system       20,  7 Mar 07 10:31 /dev/rhdisk3
{node1:root}/ -> ls -la /dev/* | grep "20,  8"
brw-------   1 root     system       20,  8 Mar 07 10:31 /dev/hdisk4
crw-------   1 root     system       20,  8 Mar 07 10:31 /dev/rhdisk4
crw-rw-r--   1 oracle   dba          20,  8 Apr 03 14:24 /dev/voting_disk1
{node1:root}/ -> ls -la /dev/* | grep "20,  9"
brw-------   1 root     system       20,  9 Mar 07 10:31 /dev/hdisk5
crw-------   1 root     system       20,  9 Mar 07 10:31 /dev/rhdisk5
crw-rw-r--   1 oracle   dba          20,  9 Apr 03 14:24 /dev/voting_disk2
{node1:root}/ -> ls -la /dev/* | grep "20, 10"
brw-------   1 root     system       20, 10 Mar 07 10:31 /dev/hdisk6
crw-------   1 root     system       20, 10 Mar 07 10:31 /dev/rhdisk6
crw-rw-r--   1 oracle   dba          20, 10 Apr 03 14:24 /dev/voting_disk3
```

From the table, for node2 :

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	**/dev/ocr_disk1**	hdisk2	20	6	hdisk1	**20**	**6**
OCR 2	L2	**/dev/ocr_disk2**	hdisk3	20	7	hdisk2	**20**	**7**
Voting 1	L3	**/dev/voting_disk1**	hdisk4	20	8	hdisk3	**20**	**8**
Voting 2	L4	**/dev/voting_disk2**	hdisk5	20	9	hdisk4	**20**	**9**
Voting 3	L5	**/dev/voting_disk3**	hdisk6	20	10	hdisk5	**20**	**10**

By chance Major and minor number are the same on both nodes, for corresponding hdisks, but it could be different.

To create same virtual devices on each node called :

/dev/ocr_disk1
/dev/ocr_disk2
/dev/voting_disk1
/dev/voting_disk2
/dev/voting_disk3

we need to use major and minor number of hdisks which will make the link between the virtual devices and the hdisks.

Using the command : mknod Device_Name c MajNum MinNum

For first node, as root user

```
node2:root-/> mknod /dev/ocr_disk1 c 20 6
node2:root-/> mknod /dev/ocr_disk2 c 20 7
node2:root-/> mknod /dev/voting_disk1 c 20 8
node2:root-/> mknod /dev/voting_disk2 c 20 9
node2:root-/> mknod /dev/voting_disk3 c 20 10
```

THEN set ownership of the created virtual devices to oracle:dba

```
node2:root-/> chown oracle:dba /dev/ocr_disk1
node2:root-/> chown oracle:dba /dev/ocr_disk2
node2:root-/> chown oracle:dba /dev/voting_disk1
node2:root-/> chown oracle:dba /dev/voting_disk2
node2:root-/> chown oracle:dba /dev/voting_disk3
```

THEN set read/write permissions of the created virtual devices to 660

```
node2:root-/> chmod 660 /dev/ocr_disk1
node2:root-/> chmod 660 /dev/ocr_disk2
node2:root-/> chmod 660 /dev/voting_disk1
node2:root-/> chmod 660 /dev/voting_disk2
node2:root-/> chmod 660 /dev/voting_disk3
```

Checking the modifications

After Oracle clusterware installation,

"crw-rw-r–" will be changed to **"crw-r--r–"** for the OCR and voting virtual devices.

And **"oracle dba"** ownership will be changed to **"root dba"** for the OCR virtual devices.

```
{node2:root}/ -> ls -la /dev/* | grep "20,  6"
brw-------   1 root     system       20,  6 Mar 07 10:31 /dev/hdisk1
crw-rw-r--   1 root     dba          20,  6 Mar 12 15:03 /dev/ocr_disk1
crw-------   1 root     system       20,  6 Mar 07 10:31 /dev/rhdisk1
{node2:root}/ -> ls -la /dev/* | grep "20,  7"
brw-------   1 root     system       20,  7 Mar 07 10:31 /dev/hdisk2
crw-rw-r--   1 root     dba          20,  7 Mar 12 15:03 /dev/ocr_disk2
crw-------   1 root     system       20,  7 Mar 07 10:31 /dev/rhdisk2
{node2:root}/ -> ls -la /dev/* | grep "20,  8"
brw-------   1 root     system       20,  8 Mar 07 10:31 /dev/hdisk3
crw-------   1 root     system       20,  8 Mar 07 10:31 /dev/rhdisk3
crw-rw-r--   1 oracle   dba          20,  8 Apr 03 14:24 /dev/voting_disk1
{node2:root}/ -> ls -la /dev/* | grep "20,  9"
brw-------   1 root     system       20,  9 Mar 07 10:31 /dev/hdisk4
crw-------   1 root     system       20,  9 Mar 07 10:31 /dev/rhdisk4
crw-rw-r--   1 oracle   dba          20,  9 Apr 03 14:24 /dev/voting_disk2
{node2:root}/ -> ls -la /dev/* | grep "20, 10"
brw-------   1 root     system       20, 10 Mar 07 10:31 /dev/hdisk5
crw-------   1 root     system       20, 10 Mar 07 10:31 /dev/rhdisk5
crw-rw-r--   1 oracle   dba          20, 10 Apr 03 14:24 /dev/voting_disk3
```

Now, we did format the virtual devices :

Format (Zeroing) and Verify that you can read on the disks from each node :

On node 1 ...

```
node1:root-/> for i in 1 2
do
dd if=/dev/zero of=/dev/ocr_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.

node1:root-/> for i in 1 2 3
do
dd if=/dev/zero of=/dev/voting_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
```

Verify devices concurrent read/write access by running at the same time dd command from each node :

At the same time :

- **on node1**
- **on node2**

On node 1 ...

```
node1:root-/> for i in 1 2
do
dd if=/dev/zero of=/dev/ocr_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
```

On node 2 ...

```
node2:root-/> for i in 1 2
do
dd if=/dev/zero of=/dev/ocr_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
```

Same for voting disks ...

10.10.4 Preparing Disks for ASM

We know the LUN's to use for the ASM disks :

Disks	LUN's Number
ASM Spfile Disk	L**0**
Disk 1 for ASM	L**6**
Disk 2 for ASM	L**7**
Disk 3 for ASM	L**8**
Disk 4 for ASM	L**9**
Disk 5 for ASM	L**A (meaning Lun 10)**
Disk 6 for ASM	L**B (meaning Lun 11)**
Disk 7 for ASM	L**C (meaning Lun 12)**
…	…

dentify the disks available for ASM and ASM spfile disks, on each node, knowing the LUN's numbers.

Disk 8 for ASM

Disks	LUN's Number	Node 1	Node 2	Node ….
ASM Spfile Disk	L**0**			
Disk 1 for ASM	L**6**			
Disk 2 for ASM	L**7**			
Disk 3 for ASM	L**8**			
Disk 4 for ASM	L**9**			
Disk 5 for ASM	L**A**			
Disk 6 for ASM	L**B**			
Disk 7 for ASM	L**C**			
…	…	…	…	…

There are two methods to identify the corresponding hdisks :

- **Identify LUN ID assign to hdisk, using "lscfg –l hdisk?" command**
- **Identify hdisks by assigning momently a PVID to each hdisk not having one**

We strongly recommend not using PVID to identify hdisk.

If some hdisks are already used, setting a PVID to a used hdisk can corrupt the the hdisk header, and have generate issues as loosing data stored on the hdis

Get the List of the hdisks on node1 :

List of available hdisks on node 1 :

rootvg is hdisk0 !!!

On node 1 …

```
{node1:root}/ -> lspv
hdisk0          00ced22cf79098ff                    rootvg          active
hdisk1          none                                None
hdisk2          none                                None
hdisk3          none                                None
hdisk4          none                                None
hdisk5          none                                None
hdisk6          none                                None
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
hdisk11         none                                None
hdisk12         none                                None
hdisk13         none                                None
{node1:root}/ ->
```

➔ **NO PVID are assigned apart for the rootvg hdisk**

Using lscfg command, try to identify the hdisks in the list generated by lspv on node1 :

Identify LUN ID assign to hdisk, using "lscfg –vl hdisk?" command

On node 1 …

```
{node1:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
 do
 lscfg -vl hdisk$i
 done
  hdisk0    U9117.570.10ED22C-V5-C5-T1-L810000000000  Virtual SCSI Disk Drive
  hdisk1    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L0  3552     (500) Disk Array Device
  hdisk2    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L1000000000000  3552     (500) Disk Array Device
  hdisk3    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L2000000000000  3552     (500) Disk Array Device
  hdisk4    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L3000000000000  3552     (500) Disk Array Device
  hdisk5    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L4000000000000  3552     (500) Disk Array Device
  hdisk6    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L5000000000000  3552     (500) Disk Array Device
  hdisk7    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L6000000000000  3552     (500) Disk Array Device
  hdisk8    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L7000000000000  3552     (500) Disk Array Device
  hdisk9    U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L8000000000000  3552     (500) Disk Array Device
  hdisk10   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-L9000000000000  3552     (500) Disk Array Device
  hdisk11   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LA000000000000  3552     (500) Disk Array Device
  hdisk12   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LB000000000000  3552     (500) Disk Array Device
  hdisk13   U7879.001.DQD01JK-P1-C2-T1-W200200A0B80C5404-LC000000000000  3552     (500) Disk Array Device
{node1:root}/ ->
```

THEN

We get the following table :

Disks	LUN's Number	Node 1	Node 2	Node ….
ASM Spfile Disk	L**0**	**hdisk1**		hdisk?
Disk 1 for ASM	L**6**	**hdisk7**		hdisk?
Disk 2 for ASM	L**7**	**hdisk8**		…
Disk 3 for ASM	L**8**	**hdisk9**		
Disk 4 for ASM	L**9**	**hdisk10**		
Disk 5 for ASM	L**A**	**hdisk11**		
Disk 6 for ASM	L**B**	**hdisk12**		
Disk 7 for ASM	L**C**	**hdisk13**		
…	…	…	…	…

➔ **No need to assign PVID when using this method.**

Get the List of the hdisks on node2

List of available hdisks on node 2 :

rootvg is hdisk11, whichis not same hdisk as on node 1, which means that it could be the same for each hdisk. On both nodes, same hdisk names might not be attached to same LUN.

On node 2 …

```
{node2:root}/ -> lspv
hdisk0          none                None
hdisk1          none                None
hdisk2          none                None
hdisk3          none                None
hdisk4          none                None
hdisk5          none                None
hdisk6          none                None
hdisk11         00ced22c826884d3                    rootvg          active
hdisk7          none                                None
hdisk8          none                                None
hdisk9          none                                None
hdisk10         none                                None
hdisk12         none                                None
hdisk13         none                                None
{node2:root}/ ->
```

➔ **NO PVID are assigned apart for the rootvg hdisk**

Using lscfg command, try to identify the hdisks in the list generated by lspv on node2 :

Be carefull hdisk7 on node1 is not necessary hdisk7 on node2.

Identify LUN ID assign to hdisk, using "lscfg –vl hdisk?" command

On node 1 …

```
{node2:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
 do
 lscfg -vl hdisk$i
 done
  hdisk0    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L0  3552     (500) Disk Array Device
  hdisk1    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L1000000000000  3552     (500) Disk Array Device
  hdisk2    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L2000000000000  3552     (500) Disk Array Device
  hdisk3    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L3000000000000  3552     (500) Disk Array Device
  hdisk4    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L4000000000000  3552     (500) Disk Array Device
  hdisk5    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L5000000000000  3552     (500) Disk Array Device
  hdisk6    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L6000000000000  3552     (500) Disk Array Device
  hdisk7    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L7000000000000  3552     (500) Disk Array Device
  hdisk8    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L8000000000000  3552     (500) Disk Array Device
  hdisk9    U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-L9000000000000  3552     (500) Disk Array Device
  hdisk10   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LA000000000000  3552     (500) Disk Array Device
  hdisk11   U9117.570.10ED22C-V1-C5-T1-L810000000000  Virtual SCSI Disk Drive
  hdisk12   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LB000000000000  3552     (500) Disk Array Device
  hdisk13   U7879.001.DQD01JK-P1-C6-T1-W200200A0B80C5404-LC000000000000  3552     (500) Disk Array Device
{node2:root}/ ->
```

THEN

We get the following table :

Disks	LUN's Number	Node 1	Node 2	Node ….
ASM Spfile Disk	L**0**	hdisk1	**hdisk0**	hdisk?
Disk 1 for ASM	L**6**	hdisk7	**hdisk6**	hdisk?
Disk 2 for ASM	L**7**	hdisk8	**hdisk7**	…
Disk 3 for ASM	L**8**	hdisk9	**hdisk8**	
Disk 4 for ASM	L**9**	hdisk10	**hdisk9**	
Disk 5 for ASM	L**A**	hdisk11	**hdisk10**	
Disk 6 for ASM	L**B**	hdisk12	**hdisk12**	
Disk 7 for ASM	L**C**	hdisk13	**hdisk13**	
…	…	…	…	…

➔ **No need to assign PVID when using this method.**

Setup reserve_policy on ocr and voting hdisks, on each node :

Example for one hdisk :

Issue the command "lsattr –E –l hdisk7" to vizualize all attributes for hdisk7

```
{node1:root}/ -> lsattr -El hdisk7
PR_key_value    none                                Persistant Reserve Key Value           True
cache_method    fast_write                          Write Caching method                   False
ieee_volname    600A0B80000C54180000026345EE3CCC IEEE Unique volume name                   False
lun_id          0x0006000000000000                  Logical Unit Number                    False
max_transfer    0x100000                            Maximum TRANSFER Size                  True
prefetch_mult   1                                   Multiple of blocks to prefetch on read False
pvid            none                                Physical volume identifier             False
q_type          simple                              Queuing Type                           False
queue_depth     10                                  Queue Depth                            True
raid_level      0                                   RAID Level                             False
reassign_to     120                                 Reassign Timeout value                 True
reserve_policy  no_reserve                          Reserve Policy                         True
rw_timeout      30                                  Read/Write Timeout value               True
scsi_id         0x690600                            SCSI ID                                False
size            4096                                Size in Mbytes                         False
write_cache     yes                                 Write Caching enabled                  False
{node1:root}/ ->
```

Or only lsattr –E –l hdisk3 | grep reserve

```
{node1:root}/ -> lsattr -El hdisk7 | grep reserve
reserve_policy single_path                          Reserve Policy                         True
{node1:root}/ ->
```

- **On IBM storage (ESS, FasTt, DSXXXX) : Change the "reserve_policy" attribute to "no_reserve"**

  chdev -l hdisk? -a reserve_policy=no_reserve

- **On EMC storage : Change the "reserve_lock" attribute to "no"**

  chdev -l hdisk? -a reserve_lock=no

- **On HDS storage with HDLM driver, and no disks in Volume Group : Change the "dlmrsvlevel" attribute to "no_reserve"**

  chdev -l dlmfdrv? -a dlmrsvlevel=no_reserve

Change the "reserve_policy" attributes for each disks dedicated to ASM, on each nodes of the cluster :

In our case, we have an IBM storage !!!

On node 1 …

```
{node1:root}/ -> for i in 1 7 8 9 10 11 12 13
 do
 chdev -l hdisk$i  -a  reserve_policy=no_reserve
 done
changed
changed
changed
changed
changed
changed
changed
changed
{node1:root}/ ->
```

On node 2 …

```
{node2:root}/ -> for i in 0 6 7 8 9 10 12 13
 do
 chdev -l hdisk$i  -a  reserve_policy=no_reserve
 done
changed
changed
changed
changed
changed
changed
changed
changed
{node2:root}/ ->
```

Example for one hdisk on node1 :

Issue the command "lsattr –E –l hdisk3 |grep reserve" to vizualize modified attributes for hdisk3

```
{node1:root}/ -> lsattr -El hdisk7 | grep reserve
reserve_policy no_reserve                       Reserve Policy                  True
{node1:root}/ ->
```

As described before, disks might have different names from one node to another for example hdisk7 on node1 might be hdisk7 on node2, etc…

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
ASM Spfile Disk	L0	/dev/asmspf_disk	hdisk1			hdisk0		
Disk 1 for ASM	L6	/dev/rhdisk?	hdisk7			hdisk6		
Disk 2 for ASM	L7	/dev/rhdisk?	hdisk8			hdisk7		
Disk 3 for ASM	L8	/dev/rhdisk?	hdisk9			hdisk8		
Disk 4 for ASM	L9	/dev/rhdisk?	hdisk10			hdisk9		
Disk 5 for ASM	LA	/dev/rhdisk?	hdisk11			hdisk10		
Disk 6 for ASM	LB	/dev/rhdisk?	hdisk12			hdisk12		
Disk 7 for ASM	LC	/dev/rhdisk?	hdisk13			hdisk13		

We need to get major and minor number for each hdisk of node1 :

To obtain minor and major numbers of each hdisk, on node1 …

We need to issue the command :

ls –la /dev/hdisk?

On node1 :

```
{node1:root}/ -> for i in 1 7 8 9 10 11 12 13
 do
 ls -la /dev/*hdisk$i
 done
brw-------   1 root     system       20,  5 Mar 12 12:18 /dev/hdisk1
crw-------   1 root     system       20,  5 Mar 12 12:18 /dev/rhdisk1
brw-------   1 root     system       20, 11 Mar 07 10:31 /dev/hdisk7
crw-------   1 root     system       20, 11 Mar 07 10:31 /dev/rhdisk7
brw-------   1 root     system       20, 12 Mar 07 10:31 /dev/hdisk8
crw-------   1 root     system       20, 12 Mar 07 10:31 /dev/rhdisk8
brw-------   1 root     system       20, 13 Mar 07 10:31 /dev/hdisk9
crw-------   1 root     system       20, 13 Mar 07 10:31 /dev/rhdisk9
brw-------   1 root     system       20, 14 Mar 07 10:31 /dev/hdisk10
crw-------   1 root     system       20, 14 Mar 07 10:31 /dev/rhdisk10
brw-------   1 root     system       20, 15 Mar 07 10:31 /dev/hdisk11
crw-------   1 root     system       20, 15 Mar 07 10:31 /dev/rhdisk11
brw-------   1 root     system       20, 16 Mar 07 10:31 /dev/hdisk12
crw-------   1 root     system       20, 16 Mar 07 10:31 /dev/rhdisk12
brw-------   1 root     system       20, 17 Mar 27 16:13 /dev/hdisk13
crw-------   1 root     system       20, 17 Mar 27 16:13 /dev/rhdisk13
{node1:root}/ ->
```

To fill in the following table, taking result from command for hdisk1, we get "20, 5" which give us 20 as major number, and 5 as minor number.

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
ASM Spfile Disk	L0	/dev/asmspf_disk	hdisk1	20	5	hdisk0		
Disk 1 for ASM	L6	/dev/rhdisk?	hdisk7	20	11	hdisk6		
Disk 2 for ASM	L7	/dev/rhdisk?	hdisk8	20	12	hdisk7		
Disk 3 for ASM	L8	/dev/rhdisk?	hdisk9	20	13	hdisk8		
Disk 4 for ASM	L9	/dev/rhdisk?	hdisk10	20	14	hdisk9		
Disk 5 for ASM	LA	/dev/rhdisk?	hdisk11	20	15	hdisk10		
Disk 6 for ASM	LB	/dev/rhdisk?	hdisk12	20	16	hdisk12		
Disk 7 for ASM	LC	/dev/rhdisk?	hdisk13	20	17	hdisk13		

We need to get major and minor number for each hdisk of node2 :

To obtain minor and major numbers of each hdisk, on node2 …

We need to issue the command :

ls –la /dev/hdisk?

On node2 :

```
{node2:root}/ -> for i in 0 6 7 8 9 10 12 13
 do
 ls -la /dev/*hdisk$i
 done
brw-------   1 root     system       20,  5 Mar 12 12:17 /dev/hdisk0
crw-------   1 root     system       20,  5 Mar 12 12:17 /dev/rhdisk0
brw-------   1 root     system       20, 11 Mar 07 10:32 /dev/hdisk6
crw-------   1 root     system       20, 11 Mar 07 10:32 /dev/rhdisk6
brw-------   1 root     system       20, 12 Mar 07 10:32 /dev/hdisk7
crw-------   1 root     system       20, 12 Mar 07 10:32 /dev/rhdisk7
brw-------   1 root     system       20, 13 Mar 07 10:32 /dev/hdisk8
crw-------   1 root     system       20, 13 Mar 07 10:32 /dev/rhdisk8
brw-------   1 root     system       20, 14 Mar 07 10:32 /dev/hdisk9
crw-------   1 root     system       20, 14 Mar 12 13:55 /dev/rhdisk9
brw-------   1 root     system       20, 15 Mar 07 10:32 /dev/hdisk10
crw-------   1 root     system       20, 15 Mar 07 10:32 /dev/rhdisk10
brw-------   1 root     system       20, 16 Mar 07 10:32 /dev/hdisk12
crw-------   1 root     system       20, 16 Mar 07 10:32 /dev/rhdisk12
brw-------   1 root     system       20, 17 Mar 27 16:13 /dev/hdisk13
crw-------   1 root     system       20, 17 Mar 27 16:13 /dev/rhdisk13
{node2:root}/ ->
```

To fill in the following table, taking result from command for hdisk0, we get "20, 5" which give us 20 as major number, and 5 as minor number.

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
ASM Spfile Disk	L**0**	/dev/asmspf_disk	hdisk1	20	5	hdisk0	**20**	**5**
Disk 1 for ASM	L**6**	/dev/rhdisk?	hdisk7	20	11	hdisk6	**20**	**11**
Disk 2 for ASM	L**7**	/dev/rhdisk?	hdisk8	20	12	hdisk7	**20**	**12**
Disk 3 for ASM	L**8**	/dev/rhdisk?	hdisk9	20	13	hdisk8	**20**	**13**
Disk 4 for ASM	L**9**	/dev/rhdisk?	hdisk10	20	14	hdisk9	**20**	**14**
Disk 5 for ASM	L**A**	/dev/rhdisk?	hdisk11	20	15	hdisk10	**20**	**15**
Disk 6 for ASM	L**B**	/dev/rhdisk?	hdisk12	20	16	hdisk12	**20**	**16**
Disk 7 for ASM	L**C**	/dev/rhdisk?	hdisk13	20	17	hdisk13	**20**	**17**

As disks have different names from one node to another for example L0 correspond to hdisk1 on node1, and hdisk0 on node2, etc…

THEN from the following table :

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
ASM Spfile Disk	L0	/dev/asmspf_disk	hdisk1	20	5	hdisk0	20	5
Disk 1 for ASM	L6	/dev/rhdisk?	hdisk7	20	11	hdisk6	20	11
Disk 2 for ASM	L7	/dev/rhdisk?	hdisk8	20	12	hdisk7	20	12
Disk 3 for ASM	L8	/dev/rhdisk?	hdisk9	20	13	hdisk8	20	13
Disk 4 for ASM	L9	/dev/rhdisk?	hdisk10	20	14	hdisk9	20	14
Disk 5 for ASM	LA	/dev/rhdisk?	hdisk11	20	15	hdisk10	20	15
Disk 6 for ASM	LB	/dev/rhdisk?	hdisk12	20	16	hdisk12	20	16
Disk 7 for ASM	LC	/dev/rhdisk?	hdisk13	20	17	hdisk13	20	17

We need to create a virtual device for the asm spfile disk.

To create same virtual devices on each node called :

/dev/asmspf_disk

we need to use major and minor number of hdisks which will make the link between the virtual devices and the hdisks.

```
Using the command : mknod Device_Name c MajNum MinNum

For first node, as root user

node1:root-/> mknod /dev/asmspf_disk c 20 5

For Second node, as root user

Node2:root-/> mknod /dev/asmspf_disk c 20 5
```

THEN set ownership of the created virtual devices to oracle:dba

```
For first node, as root user
node1:root-/> chown oracle:dba /dev/asmspf_disk

For Second node, as root user
node2:root-/> chown oracle:dba /dev/asmspf_disk
```

THEN set read/write permissions of the created virtual devices to 660

```
For first node, as root user
node1:root-/> chmod 660 /dev/asmspf_disk

For Second node, as root user
node2:root-/> chmod 660 /dev/asmspf_disk
```

Checking the modifications

After Oracle clusterware installation,

```
For first node, as root user
{node1:root}/ -> ls -la /dev/* | grep "20,  5"
crw-rw----   1 oracle   dba          20,  5 Mar 12 15:06 /dev/asmspf_disk
brw-------   1 root     system       20,  5 Mar 12 12:18 /dev/hdisk1
crw-------   1 root     system       20,  5 Mar 12 12:18 /dev/rhdisk1
{node1:root}/ ->

For Second node, as root user
{node2:root}/ -> ls -la /dev/* | grep "20,  5"
crw-rw----   1 oracle   dba          20,  5 Mar 12 15:06 /dev/asmspf_disk
brw-------   1 root     system       20,  5 Mar 12 12:17 /dev/hdisk0
crw-------   1 root     system       20,  5 Mar 12 12:17 /dev/rhdisk0
{node2:root}/ ->
```

Format (Zeroing) device from node1 :

```
node1:root-/> dd if=/dev/zero of=/dev/ocr_disk$i bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
```

NOW, we need to create a finalize preparation of the ASM disks.

2 options are possibles :

- **1 - Setting ONLY the oracle:dba awnerschip, and 660 permissions to the /dev/rhdisk? of each node**

Or

- **Creating Virtual Devices like /dev/ASM_Disk?, if wanted for humans/administrator conveniences.**

With option 1, we'll have the following table :

We'll not use Major and Minor numbers.

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk
ASM Spfile Disk	L0	/dev/asmspf_disk	hdisk1	hdisk0
Disk 1 for ASM	L6	/dev/rhdisk7 on 1 --- /dev/rhdisk6 on 2	hdisk7	hdisk6
Disk 2 for ASM	L7	/dev/rhdisk8 on 1 --- /dev/rhdisk7 on 2	hdisk8	hdisk7
Disk 3 for ASM	L8	/dev/rhdisk9 on 1 --- /dev/rhdisk8 on 2	hdisk9	hdisk8
Disk 4 for ASM	L9	/dev/rhdisk10 on 1 --- /dev/rhdisk9 on 2	hdisk10	hdisk9
Disk 5 for ASM	LA	/dev/rhdisk11 on 1 --- /dev/rhdisk10 on 2	hdisk11	hdisk10
Disk 6 for ASM	LB	/dev/rhdisk12 on 1 --- /dev/rhdisk12 on 2	hdisk12	hdisk12
Disk 7 for ASM	LC	/dev/rhdisk13 on 1 --- /dev/rhdisk13 on 2	hdisk13	hdisk13

We just need to set oracle:dba ownerchip to /dev/rhdisk? mapped to LUN for ASM.
And to set 660 read/write permissions to these rhdisk? :

THEN set ownership of the created virtual devices to oracle:dba

For first node, as root user

```
node1:root-/> for i in 7 8 9 10 11 12 13
>do
>chown oracle:dba /dev/rhdisk$i
>done
node1:root-/>
```

For Second node, as root user

```
node2:root-/> for i in 6 7 8 9 10 12 13
>do
>chown oracle:dba /dev/rhdisk$i
>done
node2:root-/>
```

THEN set read/write permissions of the created virtual devices to 660

For first node, as root user

```
node1:root-/> for i in 7 8 9 10 11 12 13
>do
>chmod 660 /dev/rhdisk$i
>done
node1:root-/>
```

For Second node, as root user

```
node2:root-/> for i in 6 7 8 9 10 12 13
>do
>chmod 660 /dev/rhdisk$i
>done
node2:root-/>
```

Checking the modifications

Check also on second node ...

For first node, as root user

```
{node1:root}/ -> ls -la /dev/rhdisk? | grep oracle
crw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/rhdisk7
crw-rw----   1 oracle   dba          20, 12 Mar 07 10:31 /dev/rhdisk8
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk9
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk10
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk11
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk12
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk13
{node1:root}/ ->
```

With option 2, we'll have the following table :

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
ASM Spfile Disk	L0	/dev/asmspf_disk	Hdisk1	20	5	hdisk0	20	5
Disk 1 for ASM	L6	/dev/ASM_Disk1	Hdisk7	20	11	hdisk6	20	11
Disk 2 for ASM	L7	/dev/ASM_Disk2	Hdisk8	20	12	hdisk7	20	12
Disk 3 for ASM	L8	/dev/ASM_Disk3	Hdisk9	20	13	hdisk8	20	13
Disk 4 for ASM	L9	/dev/ASM_Disk4	Hdisk10	20	14	hdisk9	20	14
Disk 5 for ASM	LA	/dev/ASM_Disk5	Hdisk11	20	15	hdisk10	20	15
Disk 6 for ASM	LB	/dev/ASM_Disk6	Hdisk12	20	16	hdisk12	20	16
Disk 7 for ASM	LC	/dev/ASM_Disk7	Hdisk13	20	17	hdisk13	20	17

We need to create a virtual device for the asm disks.

By chance Major and minor number are the same on both nodes, for corresponding hdisks, but it could be different.

To create same virtual devices on each node called :

/dev/asmspf_disk

we need to use major and minor number of hdisks which will make the link between the virtual devices and the hdisks.

Using the command : mknod Device_Name c MajNum MinNum

For first node, as root user

```
node1:root-/> mknod /dev/ASM_Disk1 c 20 11
node1:root-/> mknod /dev/ASM_Disk2 c 20 12
node1:root-/> mknod /dev/ASM_Disk3 c 20 13
node1:root-/> mknod /dev/ASM_Disk4 c 20 14
.....
```

For Second node, as root user

```
node2:root-/> mknod /dev/ASM_Disk1 c 20 11
node2:root-/> mknod /dev/ASM_Disk2 c 20 12
node2:root-/> mknod /dev/ASM_Disk3 c 20 13
node2:root-/> mknod /dev/ASM_Disk4 c 20 14
.....
```

THEN set ownership of the created virtual devices to oracle:dba

For first node, as root user

```
node1:root-/> for i in 1 2 3 4 5 6 7
>do
>chown oracle:dba /dev/ASM_Disk$i
>done
node1:root-/>
```

For Second node, as root user

```
node2:root-/> for i in 1 2 3 4 5 6 7
>do
>chown oracle:dba /dev/ASM_Disk$i
>done
node2:root-/>
```

THEN set read/write permissions of the created virtual devices to 660

For first node, as root user

```
node1:root-/> for i in 1 2 3 4 5 6 7
>do
>chmod 660 /dev/ASM_Disk$i
>done
node1:root-/>
```

For Second node, as root user

```
node2:root-/> for i in 1 2 3 4 5 6 7
>do
>chmod 660 /dev/ASM_Disk$i
>done
node2:root-/>
```

Checking the modifications

For first node, as root user

```
{node1:root}/ -> ls -la /dev/* | grep "20,  11"
brw-------   1 root     system       20,  11 Mar 07 10:31 /dev/hdisk7
crw-------   1 root     system       20,  11 Mar 07 10:31 /dev/rhdisk7
crw-rw-r--   1 oracle   dba          20,  11 Apr 03 14:24 /dev/ASM_Disk1
```

Check also for all disks, and on second node ...

Now, we need to format the virtual devices, or rhdisk. In both option, zeroing rhdisk is sufficient :

Format (Zeroing) and Verify that you can read on the disks from each node :

On node 1 ...

```
node1:root-/> for i in 7 8 9 10 11 12 13
>do
>dd if=/dev/zero of=/dev/rhdisk$i bs=8192 count=25000 &
>done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
...

node:root-/> for i in 6 7 8 9 10 12 13
```

Verify devices concurrent read/write access by running at the same time dd command from each node :

At the same time :

- **on node1**

On node 1 ...

```
node1:root-/> for i in 7 8 9 10 11 12 13
>do
>dd if=/dev/zero of=/dev/rhdisk$i bs=8192 count=25000 &
>done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
...
```

- **on node2**

On node 2 ...

```
node2:root-/> for i in 6 7 8 9 10 12 13
>do
>dd if=/dev/zero of=/dev/rhdisk$i bs=8192 count=25000 &
>done
25000+0 records in.
25000+0 records out.
25000+0 records in.
25000+0 records out.
...
```

10.10.5 Removing assigned PVID on hdisk

From following table, make sure that no PVID are assigned to hdisks from each node mapped to LUN's.

Disks	LUN's ID Number	Device Name seen on node1 (1) and node2 (2)	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR 1	L1	/dev/ocr_disk1	hdisk2	20	6	hdisk1	20	6
OCR 2	L2	/dev/ocr_disk2	hdisk3	20	7	hdisk2	20	7
Voting 1	L3	/dev/voting_disk1	hdisk4	20	8	hdisk3	20	8
Voting 2	L4	/dev/voting_disk2	hdisk5	20	9	hdisk4	20	9
Voting 3	L5	/dev/voting_disk3	hdisk6	20	10	hdisk5	20	10
ASM Spfile Disk	L0	/dev/asmspf_disk	hdisk1	20	5	hdisk0	20	5
Disk 1 for ASM	L6	/dev/rhdisk?	hdisk7	20	11	hdisk6	20	11
Disk 2 for ASM	L7	/dev/rhdisk?	hdisk8	20	12	hdisk7	20	12
Disk 3 for ASM	L8	/dev/rhdisk?	hdisk9	20	13	hdisk8	20	13
Disk 4 for ASM	L9	/dev/rhdisk?	hdisk10	20	14	hdisk9	20	14
Disk 5 for ASM	LA	/dev/rhdisk?	hdisk11	20	15	hdisk10	20	15
Disk 6 for ASM	LB	/dev/rhdisk?	hdisk12	20	16	hdisk12	20	16
Disk 7 for ASM	LC	/dev/rhdisk?	hdisk13	20	17	hdisk13	20	17

To remove PVID from hdisk, we will use the chdev command :

PVID must be removed from hdisk on each node.

IMPORTANT !!!!! Don't remove PVID to hdisk which are not yours !!!!

Using the command : chdev -l hdisk? -a pv=clear

For first node, as root user

```
node1:root-/> for i in 2 3 4 5 6 1 7 8 9 10 11 12 13
>do
>chdev -l hdisk$i -a pv=clear
>done
hdisk2 changed
hdisk3 changed
hdisk4 changed
hdisk5 changed
...
```

For Second node, as root user

```
Node2:root-/> for i in 1 2 3 4 5 6 0 7 8 9 10 12 13
>do
>chdev -l hdisk$i -a pv=clear
>done
hdisk2 changed
hdisk3 changed
hdisk4 changed
hdisk5 changed
...
```

Check whith lspv command as root on each node, if PVID are still assigned or not !!!

10.10.6 Checking for reserve_policy, and PVID settings.

Checking that no PVID are assigned to hdisks, and no single_path reserve policy are set :

As root user, using the command : `lsattr -El hdisk1`

```
{node1:root}/ -> lsattr -El hdisk1
PR_key_value   none                              Persistant Reserve Key Value              True
cache_method   fast_write                        Write Caching method                      False
ieee_volname   600A0B80000C54030000022D45F4DF5F  IEEE Unique volume name                   False
lun_id         0x0000000000000000                Logical Unit Number                       False
max_transfer   0x100000                          Maximum TRANSFER Size                     True
prefetch_mult  1                                 Multiple of blocks to prefetch on read    False
pvid           none                              Physical volume identifier                False
q_type         simple                            Queuing Type                              False
queue_depth    10                                Queue Depth                               True
raid_level     0                                 RAID Level                                False
reassign_to    120                               Reassign Timeout value                    True
reserve_policy no_reserve                        Reserve Policy                            True
rw_timeout     30                                Read/Write Timeout value                  True
scsi_id        0x690600                          SCSI ID                                   False
size           100                               Size in Mbytes                            False
write_cache    yes                               Write Caching enabled                     False
{node1:root}/ ->
```

To check all hdisks in one shot, use following shell script :

```
for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
do
lsattr -El hdisk$i | grep reserve_policy | awk '{print $1,$2 }'| read rp1 rp2
lsattr -El hdisk$i | grep pvid | awk '{print $1,$2 }'| read pv1 pv2
lsattr -El hdisk$i | grep lun_id | awk '{print $1,$2 }'| read li1 li2
if [ "$li1" != "" ]
then
echo hdisk$i' -> '$li1' = '$li2' / '$rp1' = '$rp2' / '$pv1' = '$pv2
fi
done
```

For first node, as root user running shell script

```
{node1:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
do
 lsattr -El hdisk$i | grep reserve_policy | awk '{print $1,$2 }'| read rp1 rp2
 lsattr -El hdisk$i | grep pvid | awk '{print $1,$2 }'| read pv1 pv2
 lsattr -El hdisk$i | grep lun_id | awk '{print $1,$2 }'| read li1 li2
 if [ "$li1" != "" ]
  then
   echo hdisk$i' -> '$li1' = '$li2' / '$rp1' = '$rp2' / '$pv1' = '$pv2
  fi
done
hdisk1 -> lun_id = 0x0000000000000000 / reserve_policy = no_reserve / pvid = none
hdisk2 -> lun_id = 0x0001000000000000 / reserve_policy = no_reserve / pvid = none
hdisk3 -> lun_id = 0x0002000000000000 / reserve_policy = no_reserve / pvid = none
hdisk4 -> lun_id = 0x0003000000000000 / reserve_policy = no_reserve / pvid = none
hdisk5 -> lun_id = 0x0004000000000000 / reserve_policy = no_reserve / pvid = none
hdisk6 -> lun_id = 0x0005000000000000 / reserve_policy = no_reserve / pvid = none
hdisk7 -> lun_id = 0x0006000000000000 / reserve_policy = no_reserve / pvid = none
hdisk8 -> lun_id = 0x0007000000000000 / reserve_policy = no_reserve / pvid = none
hdisk9 -> lun_id = 0x0008000000000000 / reserve_policy = no_reserve / pvid = none
hdisk10 -> lun_id = 0x0009000000000000 / reserve_policy = no_reserve / pvid = none
hdisk11 -> lun_id = 0x000a000000000000 / reserve_policy = no_reserve / pvid = none
hdisk12 -> lun_id = 0x000b000000000000 / reserve_policy = no_reserve / pvid = none
hdisk13 -> lun_id = 0x000c000000000000 / reserve_policy = no_reserve / pvid = none
{node1:root}/ ->
```

For Second node, as root user, running same shell script :

```
{node2:root}/ -> for i in 0 1 2 3 4 5 6 7 8 9 10 11 12 13
do
 lsattr -El hdisk$i | grep reserve_policy | awk '{print $1,$2 }'| read rp1 rp2
 lsattr -El hdisk$i | grep pvid | awk '{print $1,$2 }'| read pv1 pv2
 lsattr -El hdisk$i | grep lun_id | awk '{print $1,$2 }'| read li1 li2
 if [ "$li1" != "" ]
  then
   echo hdisk$i' -> '$li1' = '$li2' / '$rp1' = '$rp2' / '$pv1' = '$pv2
 fi
done
hdisk0 -> lun_id = 0x0000000000000000 / reserve_policy = no_reserve / pvid = none
hdisk1 -> lun_id = 0x0001000000000000 / reserve_policy = no_reserve / pvid = none
hdisk2 -> lun_id = 0x0002000000000000 / reserve_policy = no_reserve / pvid = none
hdisk3 -> lun_id = 0x0003000000000000 / reserve_policy = no_reserve / pvid = none
hdisk4 -> lun_id = 0x0004000000000000 / reserve_policy = no_reserve / pvid = none
hdisk5 -> lun_id = 0x0005000000000000 / reserve_policy = no_reserve / pvid = none
hdisk6 -> lun_id = 0x0006000000000000 / reserve_policy = no_reserve / pvid = none
hdisk7 -> lun_id = 0x0007000000000000 / reserve_policy = no_reserve / pvid = none
hdisk8 -> lun_id = 0x0008000000000000 / reserve_policy = no_reserve / pvid = none
hdisk9 -> lun_id = 0x0009000000000000 / reserve_policy = no_reserve / pvid = none
hdisk10 -> lun_id = 0x000a000000000000 / reserve_policy = no_reserve / pvid = none
hdisk12 -> lun_id = 0x000b000000000000 / reserve_policy = no_reserve / pvid = none
hdisk13 -> lun_id = 0x000c000000000000 / reserve_policy = no_reserve / pvid = none
{node2:root}/ ->
```

10.10.7 Recommandations, hints and tips for OCR / Voting disks

!!! IN ANY case, DON'T assign PVID to OCR / Voting disks when Oracle clusterware has been installed, and in test or production !!!

Assigning a PVID will erase the hdisk header !!!!, and with the risk to loose content.

AFTER CRS Installation :

How to identify hdisks used as OCR, and votings disks :

All hdisks prepared for OCR and voting disks have dba group assigned :

For OCR :

```
{node1:root}/ -> ls -la /dev/ocr*disk*
crw-r-----   1 root     dba           20,  6 Mar 12 15:03 /dev/ocr_disk1
crw-r-----   1 root     dba           20,  7 Mar 12 15:03 /dev/ocr_disk2
{node1:root}/ ->
```

Then

```
{node1:root}/ -> ls -la /dev/* |grep "20,  6"
brw-------   1 root     system        20,  6 Mar 07 10:31 /dev/hdisk2
crw-r-----   1 root     dba           20,  6 Mar 12 15:03 /dev/ocr_disk1
crw-------   1 root     system        20,  6 Mar 07 10:31 /dev/rhdisk2
{node1:root}/ ->
```

And using AIX command :

Example with OCR and corresponding rhdisk 2 on node1 :

```
{node1:oracle}/oracle/crs/bin ->lquerypv -h /dev/rhdisk2|grep 'z{|}'
00000010   C2BA0000 00001000 00012BFF 7A7B7C7D  |.........+.z{|}|
{node1:root}/ ->
{node1:oracle}/oracle/crs/bin ->lquerypv -h /dev/rhdisk2|grep '00820000
FFC00000 00000000 00000000'
00000000   00820000 FFC00000 00000000 00000000  |................|
{node1:root}/ ->
```

OR

```
{node1:root}/ -> lquerypv -h /dev/rhdisk2
00000000   00820000 FFC00000 00000000 00000000  |................|
00000010   C2BA0000 00001000 00012BFF 7A7B7C7D  |.........+.z{|}|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
{node1:root}/ ->
```

AFTER CRS Installation :

How to identify hdisks used as OCR, and votings disks :

For Voting :

```
{node1:root}/ -> ls -la /dev/vot*_disk*
crw-r--r--   1 oracle   dba          20,  8 Apr 05 14:12 /dev/voting_disk1
crw-r--r--   1 oracle   dba          20,  9 Apr 05 14:12 /dev/voting_disk2
crw-r--r--   1 oracle   dba          20, 10 Apr 05 14:12 /dev/voting_disk3
{node1:root}/ ->
```

Then

```
{node1:root}/ -> ls -la /dev/* |grep "20,  8"
brw-------   1 root     system       20,  8 Mar 07 10:31 /dev/hdisk4
crw-------   1 root     system       20,  8 Mar 07 10:31 /dev/rhdisk4
crw-r--r--   1 oracle   dba          20,  8 Apr 05 14:13 /dev/voting_disk1
{node1:root}/ ->
```

And using AIX command :

Example with OCR and corresponding rhdisk 2 on node1 :

```
{node1:oracle}/oracle/crs/bin ->lquerypv -h /dev/rhdisk4|grep 'z{|}'
00000010   A4120000 00000200 00095FFF 7A7B7C7D  |..........z{|}|
{node1:root}/ ->
{node1:oracle}/oracle/crs/bin ->lquerypv -h /dev/rhdisk4|grep '00220000
FFC00000 00000000 00000000'
00000000   00220000 FFC00000 00000000 00000000  |."..............|
{node1:root}/ ->
```

OR

```
{node1:root}/ -> lquerypv -h /dev/rhdisk4
00000000   00220000 FFC00000 00000000 00000000  |."..............|
00000010   A4120000 00000200 00095FFF 7A7B7C7D  |..........z{|}|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
```

How to free hdisk at AIX level, when hdisk are not anymore used for OCR or voting disk, or need to be reset for a failed CRS installation ?

You must reset the hdisk header having the OCR or Voting disk stamp :

```
{node1:root}/ ->dd if=/dev/zero of=/dev/rhdisk2 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
```

THEN query on the hdisk header will return nothing more than all lines full of '0':

```
{node1:root}/ -> lquerypv -h /dev/rhdisk2
00000000   00000000 00000000 00000000 00000000  |................|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
{node1:root}/ ->
```

10.10.8 Recommandations, hints and tips for ASM disks

!!! IN ANY case, DON'T assign PVID to OCR / Voting disks when Oracle clusterware has been installed, and in test or production !!!

Assigning a PVID will erase the hdisk header !!!!, and with the risk to loose content.

AFTER ASM Installation, and ASM diskgroup creation :

How to identify hdisks used by ASM.

All hdisks prepared for ASM are owned by oracle user, and group dba :

```
{node1:root}/ -> ls -la /dev/rhdisk* | oracle
brw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/hdisk7
crw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/rhdisk7
brw-rw----   1 oracle   dba          20, 12 Mar 07 10:31 /dev/hdisk8
crw-rw----   1 oracle   dba          20, 12 Mar 07 10:31 /dev/rhdisk8
brw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/hdisk9
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/rhdisk9
brw-rw----   1 oracle   dba          20, 14 Mar 07 10:31 /dev/hdisk10
crw-rw----   1 oracle   dba          20, 14 Mar 07 10:31 /dev/rhdisk10
brw-rw----   1 oracle   dba          20, 15 Mar 07 10:31 /dev/hdisk11
crw-rw----   1 oracle   dba          20, 15 Mar 07 10:31 /dev/rhdisk11
brw-rw----   1 oracle   dba          20, 16 Mar 07 10:31 /dev/hdisk12
crw-rw----   1 oracle   dba          20, 16 Mar 07 10:31 /dev/rhdisk12
brw-rw----   1 oracle   dba          20, 17 Mar 27 16:13 /dev/hdisk13
crw-rw----   1 oracle   dba          20, 17 Mar 27 16:13 /dev/rhdisk13
{node1:root}/ ->
```

Or if using virtual devices :

```
{node1:root}/ -> ls -la /dev/ASM*Disk* | grep oracle
crw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/ASM_Disk1
crw-rw----   1 oracle   dba          20, 12 Mar 07 10:31 /dev/ASM_Disk2
crw-rw----   1 oracle   dba          20, 13 Mar 07 10:31 /dev/ASM_Disk3
crw-rw----   1 oracle   dba          20, 14 Mar 07 10:31 /dev/ASM_Disk4
crw-rw----   1 oracle   dba          20, 15 Mar 07 10:31 /dev/ASM_Disk5
crw-rw----   1 oracle   dba          20, 16 Mar 07 10:31 /dev/ASM_Disk6
crw-rw----   1 oracle   dba          20, 17 Mar 27 16:13 /dev/ASM_Disk7
{node1:root}/ ->
```

THEN, for example with /dev/ASM_Disk1, we use major/minor number to identify the equivalent rhdisk

```
{node1:root}/ -> ls -la /dev/* | "20, 11"
brw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/hdisk7
crw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/rhdisk7
crw-rw----   1 oracle   dba          20, 11 Mar 07 10:31 /dev/ASM_Disk1
{node1:root}/ ->
```

How to identify hdisks used by ASM.

ORCLDISK standing for oracle ASM disk

ASMDB_GROUP standing for ASM Disk Group used for the ASMDB Database we have created in our example (the one you will create later …)

And using AIX command :

Example with ASM_Disk1 and corresponding rhdisk7 on node1 :

```
{node1:oracle}/oracle/crs/bin ->lquerypv -h /dev/rhdisk2|grep ORCLDISK
00000020   4F52434C 4449534B 00000000 00000000  |ORCLDISK........|
{node1:root}/ ->
```

OR

```
{node1:root}/ -> lquerypv -h /dev/rhdisk2
00000000   00820101 00000000 80000001 D12A3D5B  |.............@|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   4F52434C 4449534B 00000000 00000000  |ORCLDISK........|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   0A100000 00010203 41534D44 425F4752  |........ASMDB_GR|
00000050   4F55505F 30303031 00000000 00000000  |OUP_0001........|
00000060   00000000 00000000 41534D44 425F4752  |........ASMDB_GR|
00000070   4F555000 00000000 00000000 00000000  |OUP.............|
00000080   00000000 00000000 41534D44 425F4752  |........ASMDB_GR|
00000090   4F55505F 30303031 00000000 00000000  |OUP_0001........|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 01F5874B ED6CE000  |...........K.l..|
000000D0   01F588CA 150BA800 02001000 00100000  |................|
000000E0   0001BC80 00001400 00000002 00000001  |................|
000000F0   00000002 00000002 00000000 00000000  |................|
{node1:root}/ ->
```

NON used ASM disks with query on the hdisk header will return nothing more than all lines full of '0':

```
{node1:root}/ -> lquerypv -h /dev/rhdisk2
00000000   00000000 00000000 00000000 00000000  |................|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
{node1:root}/ ->
```

ORCLDISK standing for oracle ASM disk

ASMDB_FLASH RECOVERY standing for ASM Disk Group used for the ASMDB Database Flash Recovery Area we have created in our example (the one you will create later ...)

Example of output with an other rhdisk :

THEN query on the hdisk header will return :

```
{node1:root}/ ->lquerypv -h /dev/rhdisk8
00000000   00820101 00000000 80000000 DEC8B940  |..............@|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   4F52434C 4449534B 00000000 00000000  |ORCLDISK........|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   0A100000 00000103 41534D44 425F464C  |........ASMDB_FL|
00000050   41534852 45434F56 4552595F 30303030  |ASHRECOVERY_0000|
00000060   00000000 00000000 41534D44 425F464C  |........ASMDB_FL|
00000070   41534852 45434F56 45525900 00000000  |ASHRECOVERY.....|
00000080   00000000 00000000 41534D44 425F464C  |........ASMDB_FL|
00000090   41534852 45434F56 4552595F 30303030  |ASHRECOVERY_0000|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 01F5874E 4E47F800  |...........NNG..|
000000D0   01F588CA 14FA8C00 02001000 00100000  |................|
000000E0   0001BC80 00001400 00000002 00000001  |................|
000000F0   00000002 00000002 00000000 00000000  |................|
{node1:root}/ ->
```

How to free hdisk at AIX level, when hdisk are not anymore used for OCR or voting disk, or need to be reset for a failed CRS installation ?

You must reset the hdisk header having ASM disk stamp :

```
{node1:root}/ ->dd if=/dev/zero of=/dev/rhdisk2 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
```

THEN query on the hdisk header will return nothing more than all lines full of '0':

```
{node1:root}/ -> lquerypv -h /dev/rhdisk2
00000000   00000000 00000000 00000000 00000000  |................|
00000010   00000000 00000000 00000000 00000000  |................|
00000020   00000000 00000000 00000000 00000000  |................|
00000030   00000000 00000000 00000000 00000000  |................|
00000040   00000000 00000000 00000000 00000000  |................|
00000050   00000000 00000000 00000000 00000000  |................|
00000060   00000000 00000000 00000000 00000000  |................|
00000070   00000000 00000000 00000000 00000000  |................|
00000080   00000000 00000000 00000000 00000000  |................|
00000090   00000000 00000000 00000000 00000000  |................|
000000A0   00000000 00000000 00000000 00000000  |................|
000000B0   00000000 00000000 00000000 00000000  |................|
000000C0   00000000 00000000 00000000 00000000  |................|
000000D0   00000000 00000000 00000000 00000000  |................|
000000E0   00000000 00000000 00000000 00000000  |................|
000000F0   00000000 00000000 00000000 00000000  |................|
{node1:root}/ ->
```

Assigning PVID on the hdisk used by ASM will give same result, on the disk header !!!!

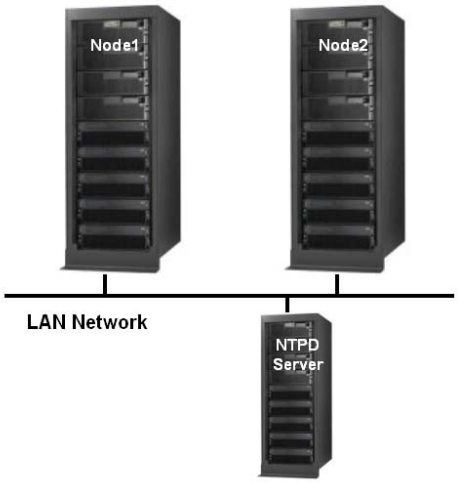
11 SYNCHRONIZE THE SYSTEM TIME ON CLUSTER NODES

Time Synchronisation

MANDATORY

To ensure that RAC operates efficiently, you must synchronize the system time on all cluster nodes.

Oracle recommends that you use xntpd for this purpose. xntpd is a complete implementation of the Network Time Protocol (NTP) version 3 standard and is more accurate than timed.



MANDATORY

- **Each node participating to the RAC cluster must have :**
  - Same date
  - Same time
- **NTPD (Time) server should be implemented !!!**

=> To ensure date/time are always synchronized

To configure xntpd, follow these steps on each cluster node :

1 / Enter the following command to create required files, if ecessary:

```
# touch /etc/ntp.drift /etc/ntp.trace /etc/ntp.conf
```

2 / Using any text editor, edit the /etc/ntp.conf file:

```
# vi /etc/ntp.conf
```

3 / Add entries similar to the following to the file:

```
server ip_address1
server ip_address2
server ip_address3
```

```
# Sample NTP Configuration file

# Specify the IP Addresses of three clock server systems.

timeserver1    10.3.25.101
timeserver2    10.3.25.102
timeserver3    10.3.25.103

# Most of the routers are broadcasting NTP time information. If your
# router is broadcasting, then the following line enables xntpd
# to listen for broadcasts.

broadcastclient

# Write clock drift parameters to a file. This enables the system
# clock to quickly sychronize to the true time on restart.

driftfile /etc/ntp.drift
tracefile /etc/ntp.trace
```

4 / To start xntpd, follow these steps:

A - Enter the following command: # **/usr/bin/smitty xntpd**

B - Choose **Start Using the xntpd Subsystem**, then choose **BOTH**.

```
                                        xntpd Subsystem

Move cursor to desired item and press Enter.

  Start Using the xntpd Subsystem
  Stop Using the xntpd Subsystem
```

```
                                  Start Using the xntpd Subsystem

Move cursor to desired item and press Enter.

  NOW
  Next System RESTART
  BOTH
```

12 IMPORTANT TIPS FOR ORACLE SOFTWARE AND PATCHS INSTALLATION INSTRUCTIONS

12.1 10G INSTALLATION ON AIX 5.3, FAILED WITH CHECKING OPERATING SYSTEM VERSION MUST BE 5200

Doc ID: Note:293750.1
Subject: 10g Installation on Aix 5.3, Failed with Checking operating system version must be 5200 Failed
Type: PROBLEM
Status: MODERATED
Content Type: TEXT/X-HTML
Creation Date: 13-DEC-2004
Last Revision Date: 02-MAR-2005

@ (AuthWiz 2.0) Created from SR 4219914.995.
@ Click here to edit in wizard.

This document is being delivered to you via Oracle Support's Rapid Visibility (RaV) process, and therefore has not been subject to an independent technical review.

The information in this document applies to:

Oracle Server - Enterprise Edition - Version: 10.1.0.2 to 10.1.0.2
AIX 5.3 Based Systems (64-bit)

Symptoms

10g Installation on Aix 5.3, Failed with Checking operating system version must be 5200 Failed
Installation Log show following details :

Using paramFile: /u06/OraDb10g/Disk1/install/oraparam.ini
Checking installer requirements...
Checking operating system version: must be 5200 Failed <<

Cause

This issue is same as the following :

Oracle bug fixes for AIX 5L v5.3 interoperability:
When running the Oracle Universal Installer (OUI) the following message or similar may appear:
"OUI-18001: The operating system “AIX Version 5300.0x” is not supported."

Fix

Workaround is to run the OUI as follows:

./ runInstaller -ignoreSysPrereqs

This parameter tell the installer to not stop because encountering an OS version greater than expected.

References : Note 282036.1 - Minimum software versions and patches required to Support Oracle Products on IBM pSeries.

12.2 10G RECOMMENDED STEPS BEFORE INSTALLATION AND APPLYING ANY PATCH SET ON AIX

Description

The patch set instructions for Installation or Patch Sets on AIX platforms do not include instructions to run "slibclean" before installing. This can lead to write errors and / or strange other errors during the application of the Patch Set or during upgrade / operation of a database after the Patch Set has been applied.

The recommended steps before installation and applying any Patch Set on AIX are:

1. Shutdown all instances which use the target ORACLE_HOME (*being sure to exit the SQLPLUS session used to shut down each instance*).
2. Stop all listeners started from the target ORACLE_HOME
3. Stop all client application code / daemons which use the target ORACLE_HOME
4. **Run the AIX command "/usr/sbin/slibclean" as "root" on all nodes to clean all unreferenced libraries from memory**.
5. Follow the install steps for the Patch Set

Possible Symptoms

- **Write Errors during Patch Set installation**

  <Note:169705.1> describes some of the "write" errors which can occur during application of the Patch Set if slibclean is not run.

Explanation

When AIX loads a shared library into memory the image of that library is kept in memory even if no process is using the library. If the on-disk copy of the library is altered then applications using that library still use the in-memory copy and not the updated disk copy. This is normal expected behaviour on AIX.

In the case of applying an Oracle Patch Set then shutting down all the instances , listeners and applications still leaves shared libraries in memory (eg: libjox9.a stays in memory). Application of the Patch Set updates the disk copy, but subsequent startup of an instance uses the in-memory library images (if they are still present). Hence the version banner can show the old release, and upgrade steps may fail as the instance is running an unsupported combination of libraries.

Running "slibclean" before starting the upgrade flushes libraries which are not currently in use from memory.

13 CLUSTER VERIFICATION UTILITY

13.1 UNDERSTANDING AND USING CLUSTER VERIFICATION UTILITY

NEW in 10gRAC R2 !!!!

Cluster Verification Utility (CVU) is a tool that performs system checks. CVU commands assist you with confirming that your system is properly configured for :

- Oracle Clusterware
- Oracle Real Application Clusters installation.

Introduction to Installing and Configuring Oracle Clusterware and Oracle Real Application Clusters
http://download-uk.oracle.com/docs/cd/B19306_01/install.102/b14201/intro.htm#i1026198

Oracle Clusterware and Oracle Real Application Clusters Pre-Installation Procedures
http://download-uk.oracle.com/docs/cd/B19306_01/install.102/b14201/part2.htm

13.2 USING CVU TO DETERMINE IF INSTALLATION PREREQUISITES ARE COMPLETE

On Both nodes, using oracle clusterware Disk1 as root user, DO run

…/clusterware/Disk1/rootpre/rootpre.sh

➔ CVU is using libraries installed by rootpre.sh script to run properly.

On node1

```
root@node1:/export/apps/oracle/clusterware/Disk1/rootpre
# ./rootpre.sh
./rootpre.sh output will be logged in /tmp/rootpre.out_06-01-
31.12:30:29
Kernel extension /etc/pw-syscall.64bit_kernel is loaded.
Unloading the existing extension: /etc/pw-syscall.64bit_kernel....

 Oracle Kernel Extension Loader for AIX
      Copyright (c) 1998,1999 Oracle Corporation


 Unconfigured the kernel extension successfully
 Unloaded the kernel extension successfully
Saving the original files in /etc/ora_save_06-01-31.12:30:29....
Copying new kernel extension to /etc....
Loading the kernel extension from /etc

 Oracle Kernel Extension Loader for AIX
      Copyright (c) 1998,1999 Oracle Corporation


 Successfully loaded /etc/pw-syscall.64bit_kernel with kmid:
0x4285700
 Successfully configured /etc/pw-syscall.64bit_kernel with kmid:
0x4285700
The kernel extension was successfuly loaded.

Configuring Asynchronous I/O....
Asynchronous I/O is already defined

Configuring POSIX Asynchronous I/O....
Posix Asynchronous I/O is already defined

Checking if group services should be configured....
Nothing to configure.
root@node1:/export/apps/oracle/clusterware/Disk1/rootpre
#
```

On node2

```
root@node2:/export/apps/oracle/clusterware/Disk1/rootpre
# ./rootpre.sh
./rootpre.sh output will be logged in /tmp/rootpre.out_06-01-31.12:31:00
Kernel extension /etc/pw-syscall.64bit_kernel is loaded.
Unloading the existing extension: /etc/pw-syscall.64bit_kernel....

 Oracle Kernel Extension Loader for AIX
      Copyright (c) 1998,1999 Oracle Corporation


 Unconfigured the kernel extension successfully
 Unloaded the kernel extension successfully
Saving the original files in /etc/ora_save_06-01-31.12:31:00....
Copying new kernel extension to /etc....
Loading the kernel extension from /etc

 Oracle Kernel Extension Loader for AIX
      Copyright (c) 1998,1999 Oracle Corporation


 Successfully loaded /etc/pw-syscall.64bit_kernel with kmid: 0x4285700
 Successfully configured /etc/pw-syscall.64bit_kernel with kmid:
0x4285700
The kernel extension was successfuly loaded.

Configuring Asynchronous I/O....
Asynchronous I/O is already defined

Configuring POSIX Asynchronous I/O....
Posix Asynchronous I/O is already defined

Checking if group services should be configured....
Nothing to configure.
root@node2:/export/apps/oracle/clusterware/Disk1/rootpre
#
```

IMPORTANT Extract from :

Oracle® Database Release Notes
10*g* Release 2 (10.2) for AIX 5L Based Systems (64-Bit)
Part Number B19074-03

➔ http://download-uk.oracle.com/docs/cd/B19306_01/relnotes.102/b19074/toc.htm

Third Party Clusterware

- If your deployment environment does not use HACMP, ignore the HACMP version and patches errors reported by Cluster Verification Utility (CVU). On AIX 5L version 5.2, the expected patch for HACMP v5.2 is IY60759. On AIX 5L version 5.3, the expected patches for HACMP v5.2 are IY60759, IY61034, IY61770, and IY62191.

- If your deployment environment does not use GPFS, ignore the GPFS version and patches errors reported by Cluster Verification Utility (CVU). On AIX 5L version 5.2 and version 5.3, the expected patches for GPFS 2.3.0.3 are IY63969, IY69911, and IY70276.

Check Kernel Parameter Settings

CVU does not check kernel parameter settings.
➔This issue is tracked with Oracle bug 4565046.

Missing Patch Error Message

➔When CVU finds a missing patch, it reports a xxxx patch is unknown error. This should be read as xxxx patch is missing.
This issue is tracked with Oracle bug 4566437.

Verify GPFS is Installed

Use the following commands to check for GPFS :

cluvfy stage -pre cfs -n node_list -s storageID_list [-verbose]
cluvfy stage -post cfs -n node_list -f file_system [-verbose]

This issue is tracked with Oracle bug 456039.

Oracle Cluster Verification Utility

Cluster Verification Utility (CVU) is a utility that is distributed with Oracle Clusterware 10*g*. It was developed to assist in the installation and configuration of Oracle Clusterware as well as Oracle Real Application Clusters 10*g* (RAC). The wide domain of deployment of CVU ranges from initial hardware setup through fully operational cluster for RAC deployment and covers all the intermediate stages of installation and configuration of various components. Cluster Verification Utility (CVU) download for all supported RAC 10*g* platforms will be posted as they become available.

Cluster Verification Utility Frequently Asked Questions (PDF)
http://www.oracle.com/technology/products/database/clustering/cvu/faq/cvu_faq.pdf
Download CVU for Oracle RAC 10*g* :
http://www.oracle.com/technology/products/database/clustering/cvu/cvu_download_homepage.html
Download CVU readme file for Oracle RAC 10*g* on AIX5L :
http://www.oracle.com/technology/products/database/clustering/cvu/readme/AIX_readme.pdf

- Subject: **Shared disk check with the Cluster Verification Utility** Doc ID: **Note:372358.1**

MAKE SURE You have "unzip" tool or symbolic link to unzip in /usr/bin on both node

On node1 Execute the following script As oracle user :

Setup and export your TMP, TEMP and TMPDIR variables

export TMP=/tmp With /tmp or other destination having enough free space
export TEMP=/tmp
export TMPDIR=/tmp

```
oracle@node1:/export/apps/oracle/clusterware/Disk1/cluvfy
# ls
cvupack.zip   jrepack.zip   runcluvfy.sh
```

./runcluvfy.sh stage -pre crsinst -n node1,node2 –verbose

If you want the result in a text file do the following :

./runcluvfy.sh stage -pre crsinst -n node1,node2 –verbose > /tmp/cluvfy_oracle.txt

And analyze the results :

At this stage, node connectivity is checked !!!

Performing pre-checks for cluster services setup

Checking node reachability...

```
Check: Node reachability from node "node1"
  Destination Node                  Reachable?
  ------------------------------------  ------------------------
  node1                             yes
  node2                             yes
Result: Node reachability check passed from node "node1".
```

➔ this is equivalent to hostname ping …

```
Checking user equivalence...

Check: User equivalence for user "oracle"
  Node Name                         Comment
  ------------------------------------  ------------------------
  node2                             passed
  node1                             passed
Result: User equivalence check passed for user "oracle".
```

➔ this is ssh or rsh tests for oracle user …

From node1 to node2
From node1 to node1
From node2 to node1
From node2 to node2

Checking administrative privileges...

```
Check: Existence of user "oracle"
  Node Name     User Exists               Comment
  ------------  ------------------------  ------------------------
  node1         yes                       passed
  node2         yes                       passed
Result: User existence check passed for "oracle".

Check: Existence of group "oinstall"
  Node Name     Status                    Group ID
  ------------  ------------------------  ------------------------
  node1         exists                    501
  node2         exists                    501
Result: Group existence check passed for "oinstall".

Check: Membership of user "oracle" in group "oinstall" [as Primary]
  Node Name         User Exists   Group Exists  User in Group  Primary
Comment
  ----------------  ------------  ------------  ------------  ------------  ------------
  tbas11b           yes           yes           yes           no            failed
  tbas11a           yes           yes           yes           no            failed
Result: Membership check for user "oracle" in group "oinstall" [as Primary] failed.
```

➔ this is user and group existence tests on node1

The following message is not a big issue :

Result: Group existence check failed for "oinstall".

➔ Just create oinstall group at AIX level, using smitty as root

Administrative privileges check failed

At this stage, node connectivity is checked !!!

```
Checking node connectivity...

Interface information for node "node1"
 Interface Name                 IP Address                     Subnet
 ------------------------------ ------------------------------ ---------------
 en0                            10.3.25.81                     10.3.25.0
 en1                            10.10.25.81                    10.10.25.0
 en2                            20.20.25.81                    20.20.25.0


Interface information for node "node2"
 Interface Name                 IP Address                     Subnet
 ------------------------------ ------------------------------ ---------------
 en0                            10.3.25.82                     10.3.25.0
 en1                            10.10.25.82                    10.10.25.0
 en2                            20.20.25.82                    20.20.25.0


Check: Node connectivity of subnet "10.3.25.0"
 Source                         Destination                    Connected?
 ------------------------------ ------------------------------ ---------------
 node1:en0                      node2:en0                      yes
Result: Node connectivity check passed for subnet "10.3.25.0" with node(s) node1,node2.

Check: Node connectivity of subnet "10.10.25.0"
 Source                         Destination                    Connected?
 ------------------------------ ------------------------------ ---------------
 node1:en1                      node2:en1                      yes
Result: Node connectivity check passed for subnet "10.10.25.0" with node(s) node1,node2.

Check: Node connectivity of subnet "20.20.25.0"
 Source                         Destination                    Connected?
 ------------------------------ ------------------------------ ---------------
 node1:en2                      node2:en2                      yes
Result: Node connectivity check passed for subnet "20.20.25.0" with node(s) node1,node2.


Suitable interfaces for VIP on subnet "20.20.25.0":
node1 en2:20.20.25.81
node2 en2:20.20.25.82

Suitable interfaces for the private interconnect on subnet "10.3.25.0":
node1 en0:10.3.25.81
node2 en0:10.3.25.82

Suitable interfaces for the private interconnect on subnet "10.10.25.0":
node1 en1:10.10.25.81
node2 en1:10.10.25.82

Result: Node connectivity check passed.
```

➔ this is the network interface tests on node1 and node2 …

➔ Node connectivity of subnet on node1 and node2, for each network interface …

➔ Testing suitable interfaces for VIP (public network), and private network …

Don't worry if you have messase as

ERROR:
Could not find a suitable set of interfaces for VIPs.

Result: Node connectivity check failed.

Just carry on (it will not affect the installation).

ERROR: Could not find a suitable set of interfaces for VIPs.

➔ **This is due to a CVU issue as explain bellow :**

Metalink Node ID 338924.1

CLUVFY Fails With Error: Could not find a suitable set of interfaces for VIPs

Per BUG:4437727, cluvfy makes an incorrect assumption based on RFC 1918 that any IP address that begins with any of the following octets is non-routable and hence may not be fit for being used as a **VIP**: 172.16.x.x192.168.x.x10.x.x.x However

At this stage, node system requirements for crs is checked !!!

Checking system requirements for 'crs'...

Check: Kernel version

Node Name	Available	Required	Comment
node1	AIX 5.3	AIX 5.2	passed
node2	AIX 5.3	AIX 5.2	passed

Result: Kernel version check passed.

➔ Kernel version test … Detecting AIX release

Check: System architecture

Node Name	Available	Required	Comment
node1	powerpc	powerpc	passed
node2	powerpc	powerpc	passed

Result: System architecture check passed.

➔System architecture test … Detecting Power processor

Check: Total memory

Node Name	Available	Required	Comment
node1	2GB (2097152KB)	512MB (524288KB)	passed
node2	2GB (2097152KB)	512MB (524288KB)	passed

Result: Total memory check passed.

➔Checking memory requirements …

Check: Swap space

Node Name	Available	Required	Comment
node1	1024GB (1073741824KB)	1GB (1048576KB)	passed
node2	1024GB (1073741824KB)	1GB (1048576KB)	passed

Result: Swap space check passed.

➔Checking swap space requirements …

Check: Free disk space in "/tmp" dir

Node Name	Available	Required	Comment
node1	400.35MB (409960KB)	400MB (409600KB)	passed
node2	400.35MB (409960KB)	400MB (409600KB)	passed

Result: Free disk space check passed.

➔Checking /tmp free space requirements …

Check: Free disk space in "/oracle" dir

Node Name	Available	Required	Comment
node1	4.63GB (4860044KB)	4GB (4194304KB)	passed
node2	4.63GB (4860044KB)	4GB (4194304KB)	passed

Result: Free disk space check passed.

➔Checking oracle user home directory free space requirements …

At this stage, node system requirements for crs is checked !!!

Checking system requirements for 'crs'... (Continued …)

Check: Package existence for "vacpp.cmp.core:7.0.0.2"

Node Name	Status	Comment
node1	vacpp.cmp.core:6.0.0.0	failed
node2	vacpp.cmp.core:6.0.0.0	failed

Result: Package existence check failed for "vacpp.cmp.core:7.0.0.2".

Check: Operating system patch for "IY65361 "

Node Name	Applied	Required	Comment
node1	unknown	IY65361	failed
node2	unknown	IY65361	failed

Result: Operating system patch check failed for "IY65361 ".

Check: Package existence for "vac.C:7.0.0.2"

Node Name	Status	Comment
node1	vac.C:7.0.0.2	passed
node2	vac.C:7.0.0.2	passed

Result: Package existence check passed for "vac.C:7.0.0.2".

Check ONLY for necessary requirements for ASM as explained in the cookbook (chapter 7).

CVU is testing existence of all prerequirements needed for all implementations as :

- RAC implementation
- ASM implementation
- GPFS implementation
- HACMP implementation (Concurrent raw devices or cohabitation between ASM or GPFS with HACMP)

```
Check: Package existence for "xlC.aix50.rte:7.0.0.4"
  Node Name                     Status                    Comment
  ----------------------------  ------------------------  ----------------
  node1                         xlC.aix50.rte:7.0.0.4     passed
  node2                         xlC.aix50.rte:7.0.0.4     passed
Result: Package existence check passed for "xlC.aix50.rte:7.0.0.4".

Check: Package existence for "xlC.rte:7.0.0.1"
  Node Name                     Status                    Comment
  ----------------------------  ------------------------  ----------------
  node1                         xlC.rte:7.0.0.1           passed
  node2                         xlC.rte:7.0.0.1           passed
Result: Package existence check passed for "xlC.rte:7.0.0.1".
```

xlC…… MUST be at minimum release 7.0.0.1

With Release 6…. The Oracle Clusterware will not start …

At this stage, node system requirements for crs is checked !!!

Checking system requirements for 'crs'... (Continued …)

```
Check: Package existence for "gpfs.base:2.3.0.3"
  Node Name                     Status                    Comment
  ----------------------------  ------------------------  ----------------
  node1                         gpfs.base:2.3.0.5         passed
  node2                         gpfs.base:2.3.0.5         passed
Result: Package existence check passed for "gpfs.base:2.3.0.3".

Check: Operating system patch for "IY63969"
  Node Name     Applied                 Required                Comment
  ------------  ----------------------  ----------------------  ----------
  node1         IY63969:gpfs.baseIY63969:gpfs.docs.dataIY63969:gpfs.msg.en_US  IY63969                 passed
  node2         IY63969:gpfs.baseIY63969:gpfs.docs.dataIY63969:gpfs.msg.en_US  IY63969                 passed
Result: Operating system patch check passed for "IY63969".

Check: Operating system patch for "IY69911"
  Node Name     Applied                 Required                Comment
  ------------  ----------------------  ----------------------  ----------
  node1         IY69911:gpfs.base       IY69911                 passed
  node2         IY69911:gpfs.base       IY69911                 passed
Result: Operating system patch check passed for "IY69911".

Check: Operating system patch for "IY70276"
  Node Name     Applied                 Required                Comment
  ------------  ----------------------  ----------------------  ----------
  node1         IY70276:gpfs.base       IY70276                 passed
  node2         IY70276:gpfs.base       IY70276                 passed
Result: Operating system patch check passed for "IY70276".

Check: Package existence for "cluster.license:5.2.0.0"
  Node Name                     Status                    Comment
  ----------------------------  ------------------------  ----------------
  node1                         missing                   failed
  node2                         missing                   failed
Result: Package existence check failed for "cluster.license:5.2.0.0".


Check: Operating system patch for "IY60759"
  Node Name     Applied                 Required                Comment
  ------------  ----------------------  ----------------------  ----------
  node1         unknown                 IY60759                 failed
  node2         unknown                 IY60759                 failed
Result: Operating system patch check failed for "IY60759".

Check: Operating system patch for "IY61034"
  Node Name     Applied                 Required                Comment
  ------------  ----------------------  ----------------------  ----------
  node1         IY61034:bos.mpIY61034:bos.mp64  IY61034                 passed
  node2         IY61034:bos.mpIY61034:bos.mp64  IY61034                 passed
Result: Operating system patch check passed for "IY61034".
```

Checking system requirements for 'crs'... (Continued …)

Check: Operating system patch for "IY61770"

Node Name	Applied	Required	Comment
node1	IY61770:rsct.basic.rteIY61770:rsct.core.errmIY61770:rsct.core.hostrm IY61770:rsct.core.rmcIY61770:rsct.core.sec IY61770:rsct.core.sensorrmIY61770:rsct.core.utils	IY61770	passed
Node2	IY61770:rsct.basic.rteIY61770:rsct.core.errmIY61770:rsct.core.hostrm IY61770:rsct.core.rmcIY61770:rsct.core.sec IY61770:rsct.core.sensorrmIY61770:rsct.core.utils	IY61770	passed

Result: Operating system patch check passed for "IY61770".

Check: Operating system patch for "IY62191"

Node Name	Applied	Required	Comment
node1	IY62191:bos.adt.profIY62191:bos.rte.libpthreads	IY62191	passed
node2	IY62191:bos.adt.profIY62191:bos.rte.libpthreads	IY62191	passed

Result: Operating system patch check passed for "IY62191".

Check: Package existence for "ElectricFence-2.2.2-1:2.2.2"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "ElectricFence-2.2.2-1:2.2.2".

Check: Package existence for "xlfrte:9.1"

Node Name	Status	Comment
node1	xlfrte:8.1.1.4	failed
node2	xlfrte:8.1.1.4	failed

Result: Package existence check failed for "xlfrte:9.1".

Check: Package existence for "gdb-6.0-1:6.0"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "gdb-6.0-1:6.0".

Check: Package existence for "make-3.80-1:3.80"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "make-3.80-1:3.80".

Check: Package existence for "freeware.gnu.tar.rte:1.13.0.0"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "freeware.gnu.tar.rte:1.13.0.0".

Check: Package existence for "Java14_64.sdk:1.4.2.1"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "Java14_64.sdk:1.4.2.1".

Check: Package existence for "Java131.rte.bin:1.3.1.16"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "Java131.rte.bin:1.3.1.16".

Checking system requirements for 'crs'... (Continued …)

Check: Package existence for "Java14.sdk:1.4.2.2"

Node Name	Status	Comment
node1	Java14.sdk:1.4.2.10	failed
node2	Java14.sdk:1.4.2.10	failed

Result: Package existence check failed for "Java14.sdk:1.4.2.2".

Check: Operating system patch for "IY65305"

Node Name	Applied	Required	Comment
node1	IY65305:Java14.sdk	IY65305	passed
node2	IY65305:Java14.sdk	IY65305	passed

Result: Operating system patch check passed for "IY65305".

Check: Operating system patch for "IY58350"

Node Name	Applied	Required	Comment
node1	unknown	IY58350	failed
node2	unknown	IY58350	failed

Result: Operating system patch check failed for "IY58350".

Check: Operating system patch for "IY63533"

Node Name	Applied	Required	Comment
node1	unknown	IY63533	failed
node2	unknown	IY63533	failed

Result: Operating system patch check failed for "IY63533".

Check: Package existence for "mqm.server.rte:5.3"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "mqm.server.rte:5.3".

Check: Package existence for "mqm.client.rte:5.3"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "mqm.client.rte:5.3".

Check: Package existence for "sna.rte:6.1.0.4"

Node Name	Status	Comment
node1	missing	failed
node2	missing	failed

Result: Package existence check failed for "sna.rte:6.1.0.4".

Check: Package existence for "bos.net.tcp.server"

Node Name	Status	Comment
node1	bos.net.tcp.server:5.3.0.30	passed
node2	bos.net.tcp.server:5.3.0.30	passed

Result: Package existence check passed for "bos.net.tcp.server".

Check: Operating system patch for "IY44599"

Node Name	Applied	Required	Comment
node1	unknown	IY44599	failed
node2	unknown	IY44599	failed

Result: Operating system patch check failed for "IY44599".

Checking system requirements for 'crs'... (Continued …)

Check: Operating system patch for "IY60930"

Node Name	Applied	Required	Comment
node1	IY60930:bos.mpIY60930:bos.mp64	IY60930	passed
node2	IY60930:bos.mpIY60930:bos.mp64	IY60930	passed

Result: Operating system patch check passed for "IY60930".

Check: Operating system patch for "IY58143"

Node Name	Applied	Required	Comment
node1	IY58143:X11.Dt.libIY58143:X11.base.rteIY58143:bos.acctIY58143:bos.adt.includeI Y58143:bos.adt.libmIY58143:bos.adt.profIY58143:bos.alt_disk_install.rteIY58143:bos.diag.com IY58143:bos.mp64IY58143:bos.mpIY58143:bos.net.ewlm.rteIY58143:bos.net.ipsec.keymgt IY58143:bos.net.ipsec.rteIY58143:bos.net.nfs.clientIY58143:bos.net.nfs.serverIY58143:bos.net.tcp.client IY58143:bos.net.tcp.serverIY58143:bos.net.tcp.smitIY58143:bos.perf.libperfstatIY58143:bos.perf.perfstat IY58143:bos.perf.toolsIY58143:bos.rte.archiveIY58143:bos.rte.bind_cmdsIY58143:bos.rte.boot IY58143:bos.rte.controlIY58143:bos.rte.filesystemIY58143:bos.rte.installIY58143:bos.rte.libc IY58143:bos.rte.lvmIY58143:bos.rte.manIY58143:bos.rte.methodsIY58143:bos.rte.security IY58143:bos.rte.serv_aidIY58143:bos.sysmgt.nim.clientIY58143:bos.sysmgt.quotaIY58143:bos.sysmgt.serv_aid IY58143:bos.sysmgt.sysbrIY58143:devices.chrp.base.rteIY58143:devices.chrp.pci.rteIY58143:devices.chrp.vdevice.rte IY58143:devices.common.IBM.atm.rteIY58143:devices.common.IBM.ethernet.rteIY58143:devices.common.IBM.fc.rte IY58143:devices.common.IBM.fda.diagIY58143:devices.common.IBM.mpio.rteIY58143:devices.fcp.disk.rte IY58143:devices.pci.00100f00.rteIY58143:devices.pci.14100401.diagIY58143:devices.pci.14103302.rte IY58143:devices.pci.14106602.rteIY58143:devices.pci.14106902.rteIY58143:devices.pci.14107802.rte IY58143:devices.pci.1410ff01.rteIY58143:devices.pci.22106474.rteIY58143:devices.pci.33103500.rte IY58143:devices.pci.4f111100.comIY58143:devices.pci.77101223.comIY58143:devices.pci.99172704.rte IY58143:devices.pci.c1110358.rteIY58143:devices.pci.df1000f7.comIY58143:devices.pci.df1000f7.diag IY58143:devices.pci.df1000fa.rteIY58143:devices.pci.e414a816.rteIY58143:devices.scsi.disk.rte IY58143:devices.vdevice.IBM.l-lan.rteIY58143:devices.vdevice.IBM.vscsi.rteIY58143:devices.vdevice.hvterm1.rte IY58143:devices.vtdev.scsi.rteIY58143:sysmgt.websm.appsIY58143:sysmgt.websm.framework IY58143:sysmgt.websm.rteIY58143:sysmgt.websm.webaccess	IY58143	passed
Node2	IY58143:X11.Dt.libIY58143:X11.base.rteIY58143:bos.acctIY58143:bos.adt.includeI Y58143:bos.adt.libmIY58143:bos.adt.profIY58143:bos.alt_disk_install.rteIY58143:bos.diag.com IY58143:bos.mp64IY58143:bos.mpIY58143:bos.net.ewlm.rteIY58143:bos.net.ipsec.keymgt IY58143:bos.net.ipsec.rteIY58143:bos.net.nfs.clientIY58143:bos.net.nfs.serverIY58143:bos.net.tcp.client IY58143:bos.net.tcp.serverIY58143:bos.net.tcp.smitIY58143:bos.perf.libperfstatIY58143:bos.perf.perfstat IY58143:bos.perf.toolsIY58143:bos.rte.archiveIY58143:bos.rte.bind_cmdsIY58143:bos.rte.boot IY58143:bos.rte.controlIY58143:bos.rte.filesystemIY58143:bos.rte.installIY58143:bos.rte.libc IY58143:bos.rte.lvmIY58143:bos.rte.manIY58143:bos.rte.methodsIY58143:bos.rte.security IY58143:bos.rte.serv_aidIY58143:bos.sysmgt.nim.clientIY58143:bos.sysmgt.quotaIY58143:bos.sysmgt.serv_aid IY58143:bos.sysmgt.sysbrIY58143:devices.chrp.base.rteIY58143:devices.chrp.pci.rteIY58143:devices.chrp.vdevice.rte IY58143:devices.common.IBM.atm.rteIY58143:devices.common.IBM.ethernet.rteIY58143:devices.common.IBM.fc.rte IY58143:devices.common.IBM.fda.diagIY58143:devices.common.IBM.mpio.rteIY58143:devices.fcp.disk.rte IY58143:devices.pci.00100f00.rteIY58143:devices.pci.14100401.diagIY58143:devices.pci.14103302.rte IY58143:devices.pci.14106602.rteIY58143:devices.pci.14106902.rteIY58143:devices.pci.14107802.rte IY58143:devices.pci.1410ff01.rteIY58143:devices.pci.22106474.rteIY58143:devices.pci.33103500.rte IY58143:devices.pci.4f111100.comIY58143:devices.pci.77101223.comIY58143:devices.pci.99172704.rte IY58143:devices.pci.c1110358.rteIY58143:devices.pci.df1000f7.comIY58143:devices.pci.df1000f7.diag IY58143:devices.pci.df1000fa.rteIY58143:devices.pci.e414a816.rteIY58143:devices.scsi.disk.rte IY58143:devices.vdevice.IBM.l-lan.rteIY58143:devices.vdevice.IBM.vscsi.rteIY58143:devices.vdevice.hvterm1.rte IY58143:devices.vtdev.scsi.rteIY58143:sysmgt.websm.appsIY58143:sysmgt.websm.framework IY58143:sysmgt.websm.rteIY58143:sysmgt.websm.webaccess	IY58143	passed

Result: Operating system patch check passed for "IY58143".

Check: Operating system patch for "IY66513"

Node Name	Applied	Required	Comment
node1	IY66513:bos.mpIY66513:bos.mp64	IY66513	passed
node2	IY66513:bos.mpIY66513:bos.mp64	IY66513	passed

Result: Operating system patch check passed for "IY66513".

Check: Operating system patch for "IY70159"

Node Name	Applied	Required	Comment
node1	IY70159:bos.mpIY70159:bos.mp64	IY70159	passed
node2	IY70159:bos.mpIY70159:bos.mp64	IY70159	passed

Result: Operating system patch check passed for "IY70159".

Checking system requirements for 'crs'... (Continued …)

Check: Operating system patch for "IY59386"

Node Name	Applied	Required	Comment
node1	IY59386:bos.rte.bind_cmds	IY59386	passed
node2	IY59386:bos.rte.bind_cmds	IY59386	passed

Result: Operating system patch check passed for "IY59386".

Check: Package existence for "bos.adt.base"

Node Name	Status	Comment
node1	bos.adt.base:5.3.0.30	passed
node2	bos.adt.base:5.3.0.30	passed

Result: Package existence check passed for "bos.adt.base".

Check: Package existence for "bos.adt.lib"

Node Name	Status	Comment
node1	bos.adt.lib:5.3.0.30	passed
node2	bos.adt.lib:5.3.0.30	passed

Result: Package existence check passed for "bos.adt.lib".

Check: Package existence for "bos.adt.libm"

Node Name	Status	Comment
node1	bos.adt.libm:5.3.0.30	passed
node2	bos.adt.libm:5.3.0.30	passed

Result: Package existence check passed for "bos.adt.libm".

Check: Package existence for "bos.perf.libperfstat"

Node Name	Status	Comment
node1	bos.perf.libperfstat:5.3.0.30	passed
node2	bos.perf.libperfstat:5.3.0.30	passed

Result: Package existence check passed for "bos.perf.libperfstat".

Check: Package existence for "bos.perf.perfstat"

Node Name	Status	Comment
node1	bos.perf.perfstat:5.3.0.30	passed
node2	bos.perf.perfstat:5.3.0.30	passed

Result: Package existence check passed for "bos.perf.perfstat".

Check: Package existence for "bos.perf.proctools"

Node Name	Status	Comment
node1	bos.perf.proctools:5.3.0.30	passed
node2	bos.perf.proctools:5.3.0.30	passed

Result: Package existence check passed for "bos.perf.proctools".

Check: Package existence for "rsct.basic.rte"

Node Name	Status	Comment
node1	rsct.basic.rte:2.4.3.0	passed
node2	rsct.basic.rte:2.4.3.0	passed

Result: Package existence check passed for "rsct.basic.rte".

Check: Package existence for "perl.rte:5.0005"

Node Name	Status	Comment
node1	perl.rte:5.8.2.30	passed
node1	perl.rte:5.8.2.30	passed

Result: Package existence check passed for "perl.rte:5.0005".

Check: Package existence for "perl.rte:5.6"

Node Name	Status	Comment
node1	perl.rte:5.8.2.30	passed
node2	perl.rte:5.8.2.30	passed

Result: Package existence check passed for "perl.rte:5.6".

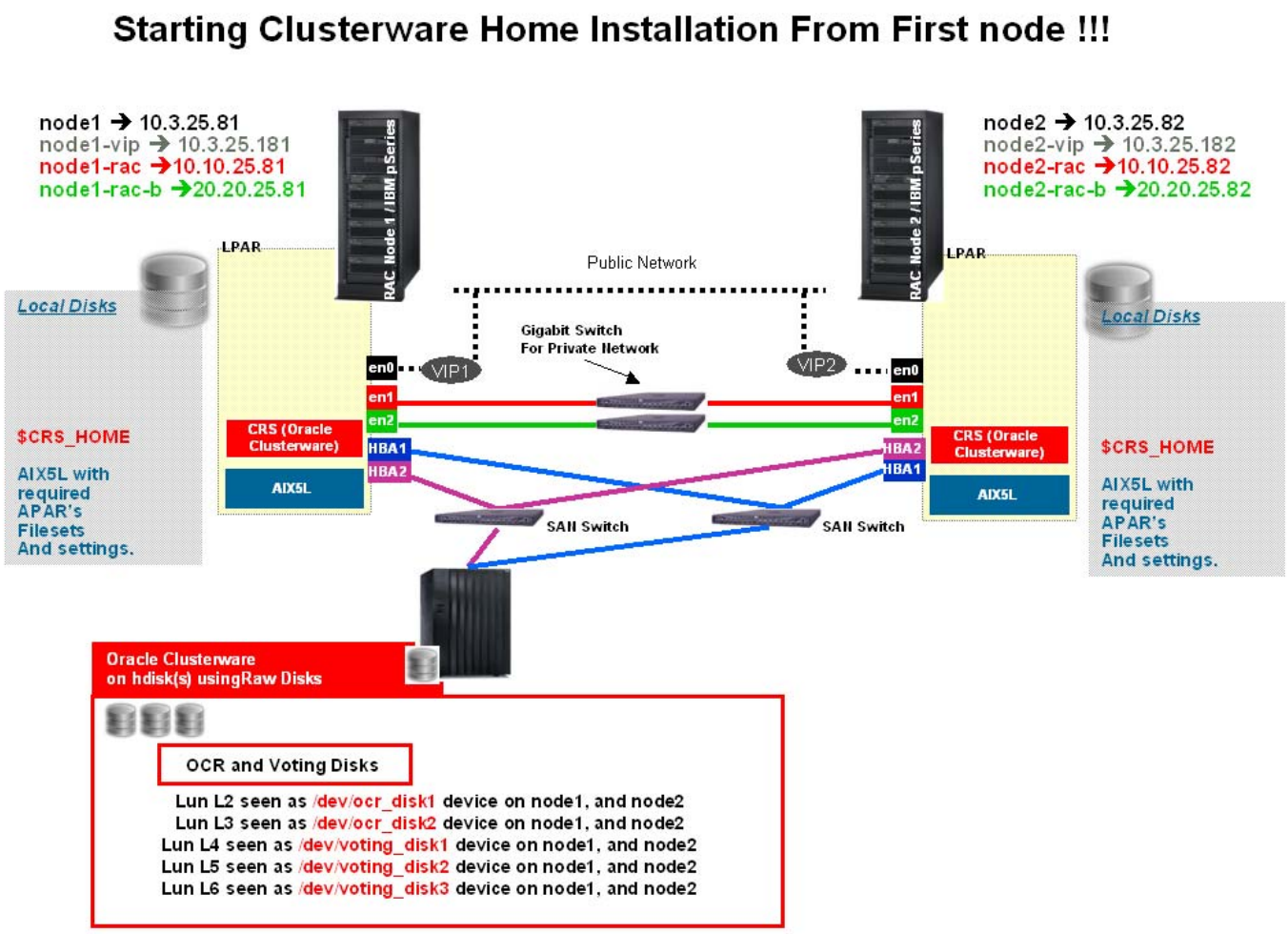
Checking system requirements for 'crs'... (Continued …)

```
Check: Package existence for "perl.rte:5.8"
  Node Name                     Status                        Comment
  ----------------------------  ----------------------------  ----------------
  node1                         perl.rte:5.8.2.30             passed
  node2                         perl.rte:5.8.2.30             passed
Result: Package existence check passed for "perl.rte:5.8".

Check: Package existence for "python-2.2-4:2.2"
  Node Name                     Status                        Comment
  ----------------------------  ----------------------------  ----------------
  node1                         missing                       failed
  node2                         missing                       failed
Result: Package existence check failed for "python-2.2-4:2.2".

Check: Package existence for "freeware.zip.rte:2.3"
  Node Name                     Status                        Comment
  ----------------------------  ----------------------------  ----------------
  node1                         missing                       failed
  node2                         missing                       failed
Result: Package existence check failed for "freeware.zip.rte:2.3".

Check: Package existence for "freeware.gcc.rte:3.3.2.0"
  Node Name                     Status                        Comment
  ----------------------------  ----------------------------  ----------------
  node1                         missing                       failed
  node2                         missing                       failed
Result: Package existence check failed for "freeware.gcc.rte:3.3.2.0".

Check: Group existence for "dba"
  Node Name     Status                  Comment
  ------------  ------------------------  ------------------------
  node1         exists                  passed
  node2         exists                  passed
Result: Group existence check passed for "dba".

Check: User existence for "nobody"
  Node Name     Status                  Comment
  ------------  ------------------------  ------------------------
  node1         exists                  passed
  node2         exists                  passed
Result: User existence check passed for "nobody".

System requirement failed for 'crs'
```

Pre-check for cluster services setup was unsuccessful on all the nodes.

Don't worry about the "Pre-check for cluster services setup was unsuccessful on all the nodes."
➔ this is a normal message as we don't want all APAR and FILESETS to be installed.

14 IMPLEMENTING ORACLE CLUSTERWARE (CRS)



14.1 ORACLE CLUSTERWARE INSTALLATION

Please read following Metalink note about CRS and 10gRAC :

Metalink Note ID 259301.1 - CRS and 10g Real Application Clusters

Oracle Cluster Ready Services installation is necessary and mandatory. This installation just have to be done only starting from one node. Once the first node is installed, Oracle OUI automatically starts the copy of the mandatory files on the others nodes, using **rcp** command. This step should not last long. **But in any case, don't think the OUI is stalled, and look at the network traffic before canceling the installation !**

As root user on each node, DO Create a symbolic link from /usr/sbin/lsattr to /etc/lsattr

ln -s /usr/sbin/lsattr /etc/lsattr

"/etc/lsattr" is used in vip check action

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node1:root}/oracle/products ->

{node2:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products ->
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

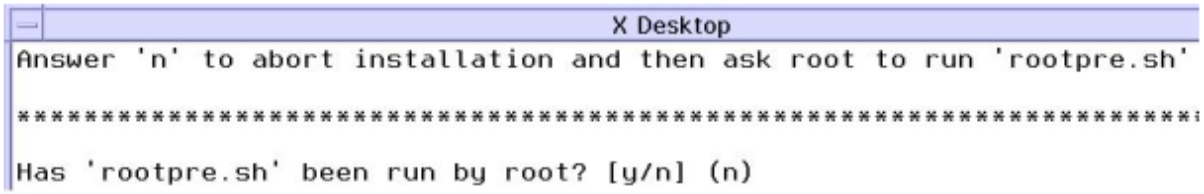
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

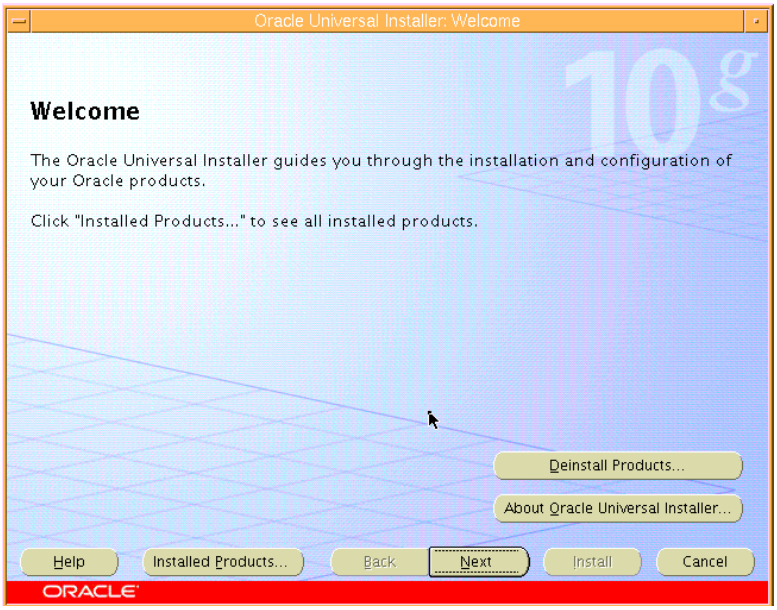
Or execute : ./runInstaller -ignoreSysPrereqs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
All filesets for 5300-04_AIX_ML were found.
All filesets for 5300-05_AIX_ML were found.

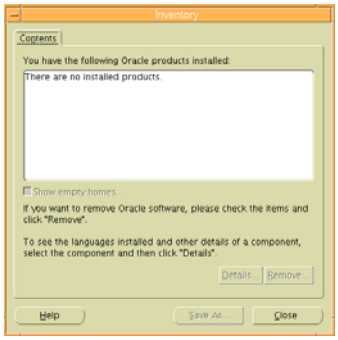


Make sure to execute rootpre.sh on each node before you click to the next step (If not done yet with CVU). ➔ make an NFS mount of the CRS Disk1 on other nodes, or remote copy files to other nodes, THEN run rootpre.sh on each node !!!



At the OUI Welcome screen

You can check if any oracle product is already installed, click on “Installed Products” :

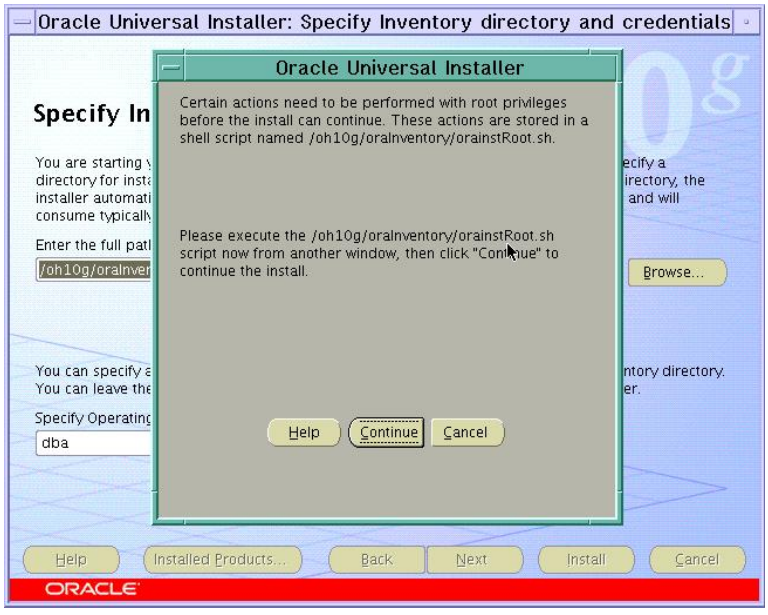


Just click Next ...

Specify where you want to create the inventory directory.
By default it will be created in the $ORACLE_BASE

Operating system group name should be set as dba

Then click Next ...

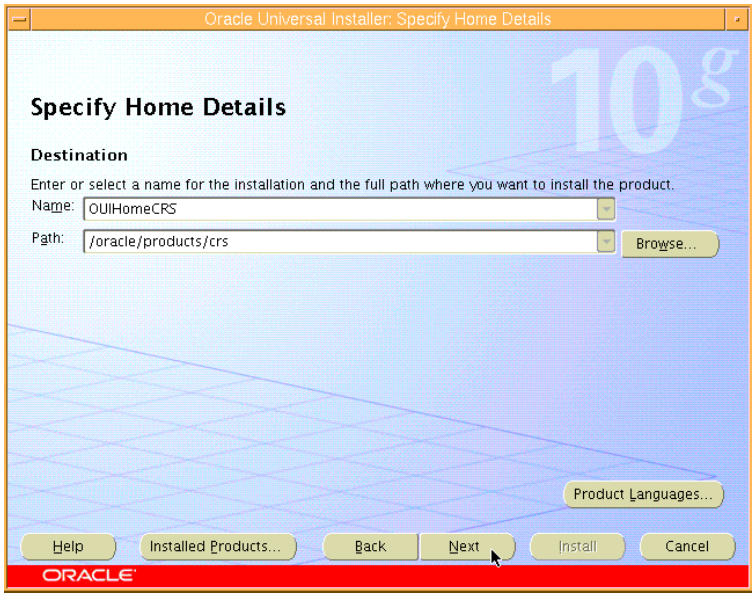


The OUI ask the user to run **orainstRoot.sh** in a separate window, if it’s the first Oracle product install on this machine. This script creates the file /etc/oraInst.loc, which is used by OUI for the list of installed products.

Connect as root on node 1, and run the orainstRoot.sh located in $ORACLE_BASE/oraInventory

This will change permissions, and group name to dba on the /etc/oraInst.loc file.

Then click Continue ...

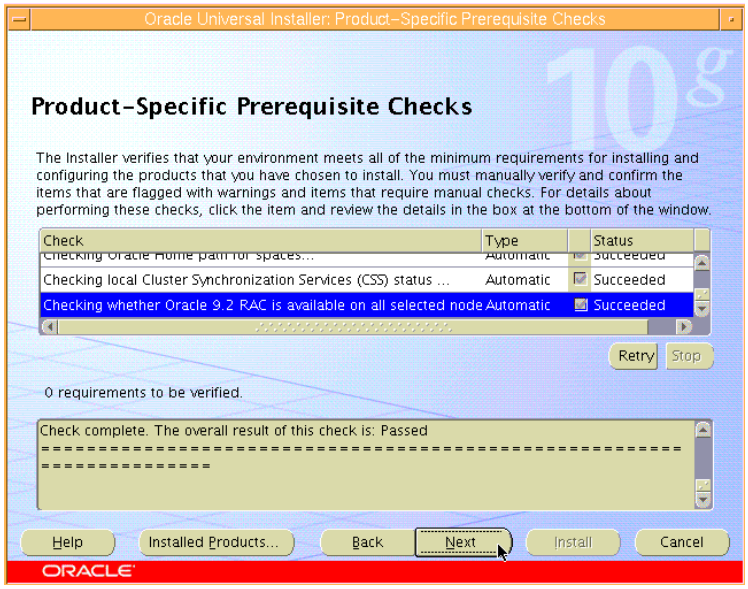


Specify file locations

Specify an ORACLE_HOME name and destination directory for the CRS installation.

The destination directory should be inside the $ORACLE_BASE

Then click Next ...



Product-Specific Prerequisite Checks :

Don't be afraid by the warning, this is just some AIX filesets missing which are not necessary.

Then click Next ...

Details of the prerequisite checks done by runInstaller

Checking operating system requirements ...
Expected result: One of 5200.004,5300.002
Actual Result: 5300.002
Check complete. The overall result of this check is: Passed
==========

Checking operating system package requirements ...
Checking for bos.adt.base(0.0); found bos.adt.base(5.3.0.51). Passed
Checking for bos.adt.lib(0.0); found bos.adt.lib(5.3.0.50). Passed
Checking for bos.adt.libm(0.0); found bos.adt.libm(5.3.0.40). Passed
Checking for bos.perf.libperfstat(0.0); found bos.perf.libperfstat(5.3.0.50). Passed
Checking for bos.perf.perfstat(0.0); found bos.perf.perfstat(5.3.0.50). Passed
Checking for bos.perf.proctools(0.0); found bos.perf.proctools(5.3.0.50). Passed
Check complete. The overall result of this check is: Passed
==========

Checking recommended operating system patches
Checking for IY59386(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY60930(bos.mp,5.3.0.1); found (bos.mp,5.3.0.54). Passed
Checking for IY60930(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp64,5.3.0.20); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp,5.3.0.20); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp,5.3.0.22); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp64,5.3.0.22); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.acct,5.3.0.1); found (bos.acct,5.3.0.51). Passed
Checking for IY58143(bos.adt.include,5.3.0.1); found (bos.adt.include,5.3.0.53). Passed
Checking for IY58143(bos.adt.libm,5.3.0.1); found (bos.adt.libm,5.3.0.40). Passed
Checking for IY58143(bos.adt.prof,5.3.0.1); found (bos.adt.prof,5.3.0.53). Passed
Checking for IY58143(bos.alt_disk_install.rte,5.3.0.1); found (bos.alt_disk_install.rte,5.3.0.51). Passed
Checking for IY58143(bos.cifs_fs.rte,5.3.0.1); found (bos.cifs_fs.rte,5.3.0.50). Passed
Checking for IY58143(bos.diag.com,5.3.0.1); found (bos.diag.com,5.3.0.51). Passed
Checking for IY58143(bos.perf.libperfstat,5.3.0.1); found (bos.perf.libperfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.perfstat,5.3.0.1); found (bos.perf.perfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.tools,5.3.0.1); found (bos.perf.tools,5.3.0.52). Passed
Checking for IY58143(bos.rte.boot,5.3.0.1); found (bos.rte.boot,5.3.0.51). Passed
Checking for IY58143(bos.rte.archive,5.3.0.1); found (bos.rte.archive,5.3.0.51). Passed
Checking for IY58143(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY58143(bos.rte.control,5.3.0.1); found (bos.rte.control,5.3.0.50). Passed
Checking for IY58143(bos.rte.filesystem,5.3.0.1); found (bos.rte.filesystem,5.3.0.51). Passed
Checking for IY58143(bos.rte.install,5.3.0.1); found (bos.rte.install,5.3.0.54). Passed
Checking for IY58143(bos.rte.libc,5.3.0.1); found (bos.rte.libc,5.3.0.53). Passed
Checking for IY58143(bos.rte.lvm,5.3.0.1); found (bos.rte.lvm,5.3.0.53). Passed
Checking for IY58143(bos.rte.man,5.3.0.1); found (bos.rte.man,5.3.0.50). Passed
Checking for IY58143(bos.rte.methods,5.3.0.1); found (bos.rte.methods,5.3.0.51). Passed
Checking for IY58143(bos.rte.security,5.3.0.1); found (bos.rte.security,5.3.0.53). Passed
Checking for IY58143(bos.rte.serv_aid,5.3.0.1); found (bos.rte.serv_aid,5.3.0.52). Passed
Check complete. The overall result of this check is: Passed
==========

Checking for Oracle Home incompatibilities ....
Actual Result: NEW_HOME
Check complete. The overall result of this check is: Passed
==========

Checking Oracle Home path for spaces...
Check complete. The overall result of this check is: Passed
==========

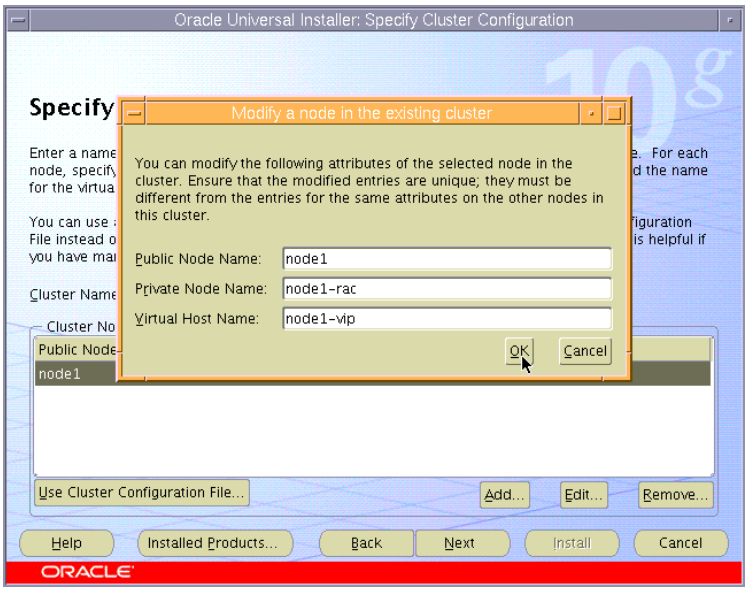
Checking local Cluster Synchronization Services (CSS) status ...
Check complete. The overall result of this check is: Passed
==========

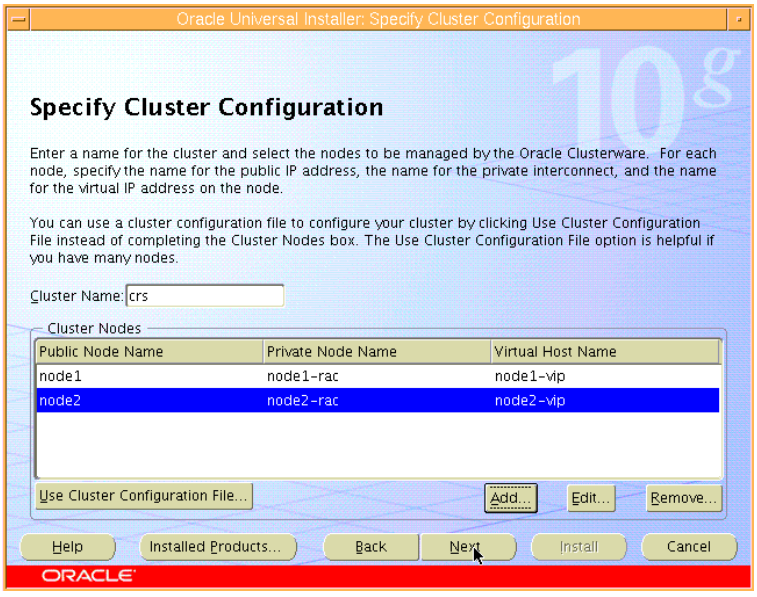
Checking whether Oracle 9.2 RAC is available on all selected nodes
Check complete. The overall result of this check is: Passed
==========

Specify Cluster Configuration :

Just to remember !!! **Public, Private, and Virtual Host Name layout**

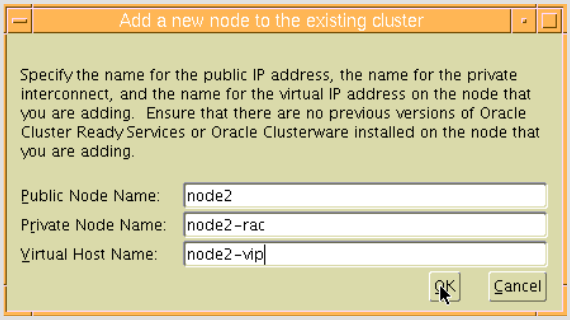
Public		VIP		RAC Interconnect (Private Network)	
en0				en1	
Node Name	IP	Node Name	IP	Node Name	IP
node1	10.3.25.81	node1-vip	10.3.25.181	node1-rac	10.10.25.81
node2	10. 3.25.82	node2-vip	10. 3.25.182	node1-rac	10.10.25.82





Specify a cluster name, "**crs**" in our example.

Click "**Add**" to **enter each node,**



And, specify one by one

- **Public Node Name,**

  Public correspond to IP address linked to connected to the public network

- **Private Node Name**

  Private correspond to IP address linked to the RAC interconnect.

- **Virtual Host Name**

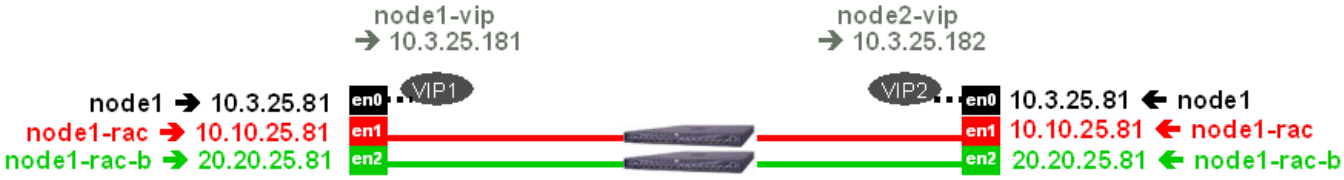
  Virtual Host Name correspond to IP address linked to the VIP.

Then click Next ...

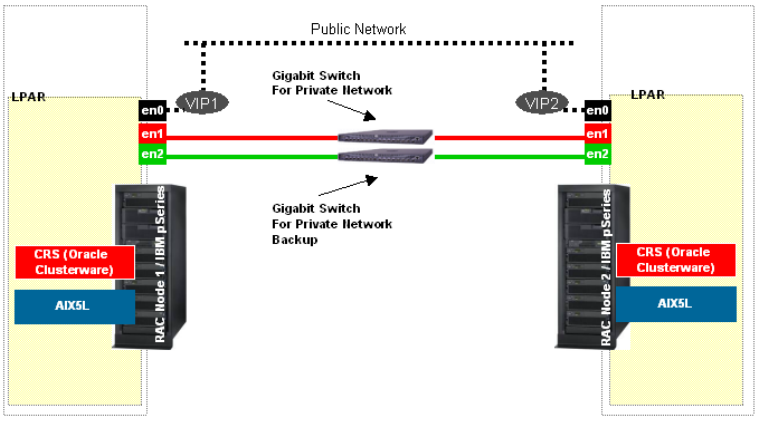
If you have problems at this stage when clicking on Next (error messages)

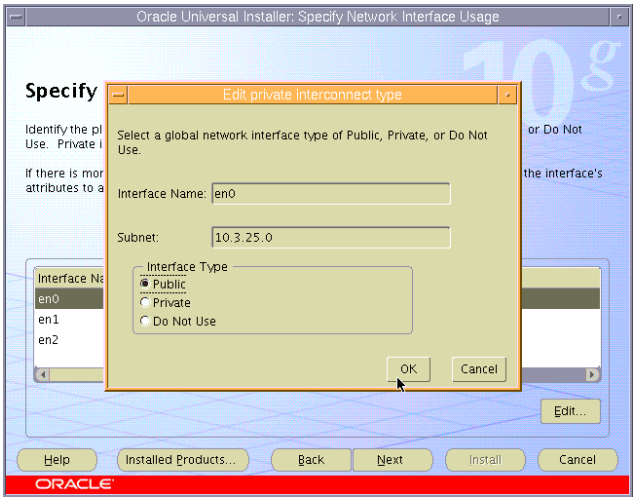
➔ **Check your network configuration, and User equivalence configurations on all nodes.**

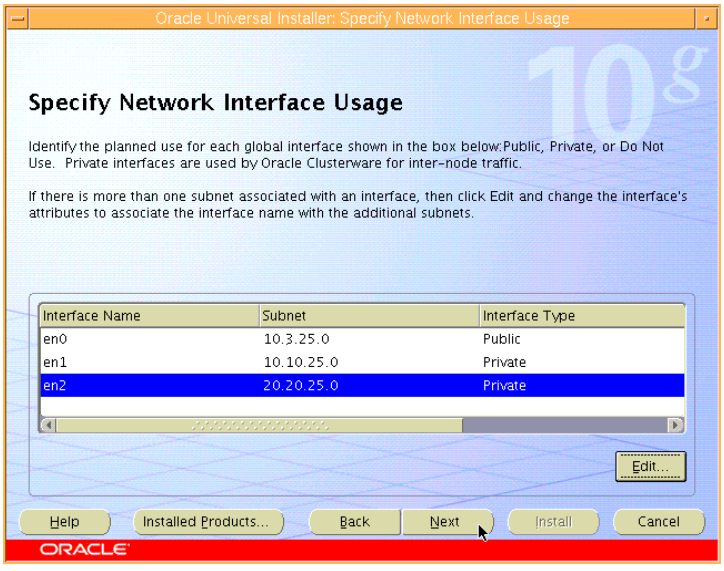
Just to remember !!!



Private Interconnect Enforcement :

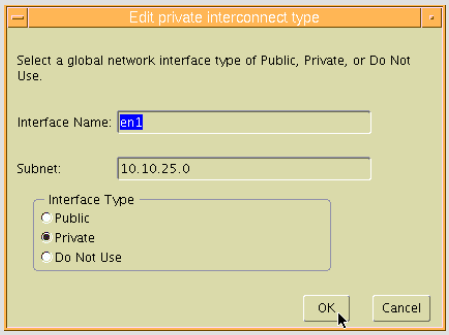


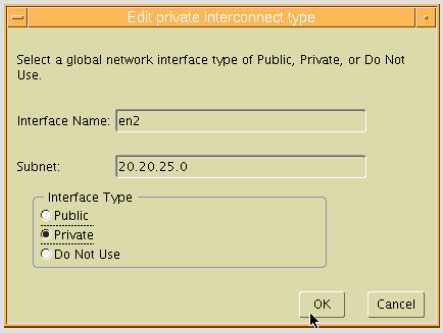




For each entry (en0,en1,en2),

Click "Edit" to Specify "Interface Type" for each network card correspond to the public network, and which one correspond to the private network (RAC Interconnect).





In our example, with or without RAC Interconnect backup implementation :

- **en0 (10.3.25.0)**

➔ must exist as "Public" on each node.

- **en1 (10.10.25.0)**

➔ must exist as "Private" on each node.

- **en2 (20.20.25.0)**

➔ **"**Do Not Use", OR "Private" for second RAC Interconnect (Backup)

- **An other en? …**

➔ Do Not Use

Then click Next …

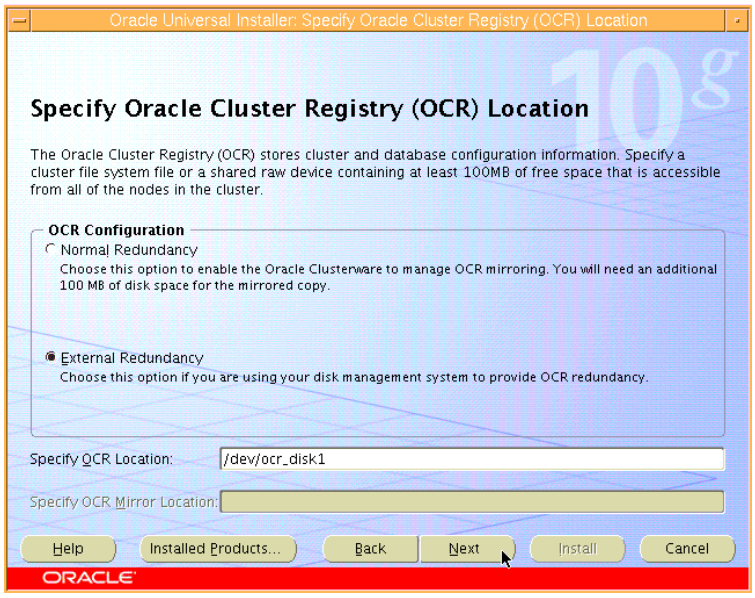
At this stage, you must have already configured the shared storage. At least for :

No Oracle Clusterware redundancy mechanism for OCR and Voting disks :	**OR**	**Oracle Clusterware redundancy mechanism for OCR and Voting disks :**
• 1 Oracle Cluster Registry Disk • 1 Voting Disk		• 2 Oracle Cluster Registry Disks (OCR Mirroring) • 3 Voting Disks (Voting Redundancy)

In our example, We will implement 2 Oracle Cluster Registry Disk, and 3 Voting Disks.

You must know where to install these 2 or 5 files to be reachable by all nodes participating to the RAC cluster.

With ASM implementation, You have to place these files on raw disks, using virtual devices :



Specify OCR Configuration, by selecting :

- **External Redundancy** (No OCR mirroring by Oracle Clusterware, should be provided by others options, disks management, etc …)

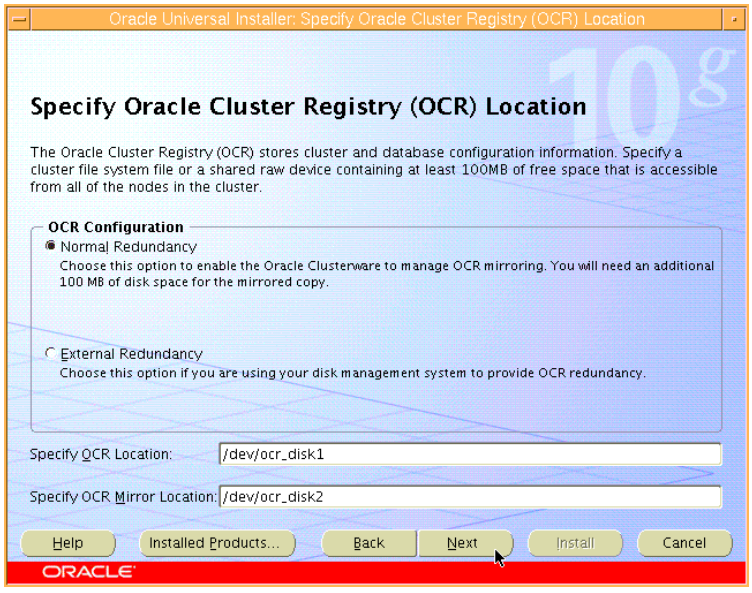
  If "External Redundancy" selected, specify raw disk location as follow :
  - /dev/ocr_disk1

OR

- **Normal Redundancy** (OCR mirroring by Oracle Clusterware)

  If "NORMAL Redundancy" selected, specify raw disk location as follow :
  - /dev/ocr_disk1
  - /dev/ocr_disk2

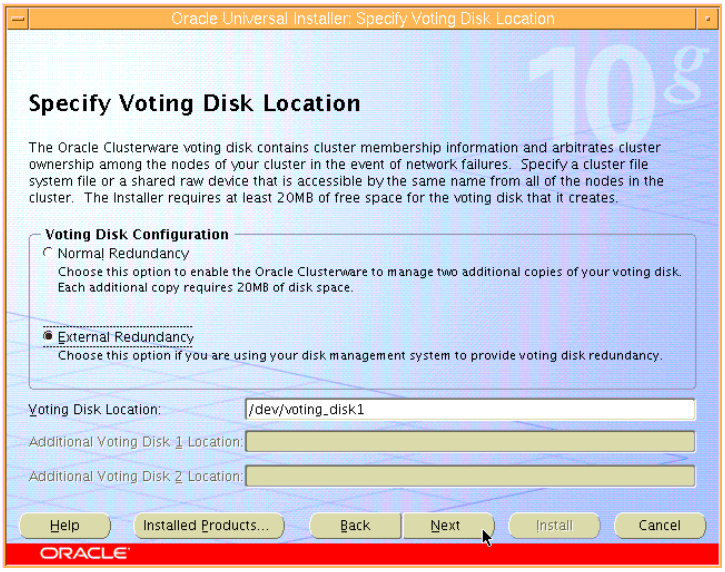
In our case, we will specify **"NORMAL Redundancy"** as described bellow :



Specify the OCR location : this must be a shared location on the shared storage reachable from all nodes.
And you must have the read/wright permissions on this shared location from all nodes.

Then click Next ...

If problems happens at this stage, do verifiy that location specified does exist, and is reachable from each AIX node, with right read/write access, and user/group owner.



Specify Voting Disk Configuration, by selecting :

- **External Redundancy** (No Voting copies managed by Oracle Clusterware, should be provided by others options, disks management, etc …)

  If "External Redundancy" selected, specify raw disk location as follow :

  - /dev/voting_disk1

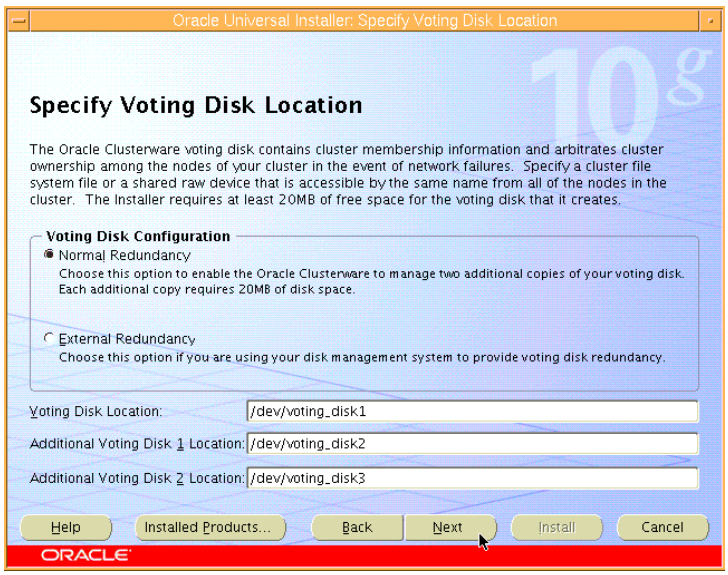
OR

- **Normal Redundancy** (Voting copies managed by Oracle Clusterware)

  If "NORMAL Redundancy" selected, specify raw disk location as follow :

  - /dev/voting_disk1
  - /dev/voting_disk2
  - /dev/voting_disk3

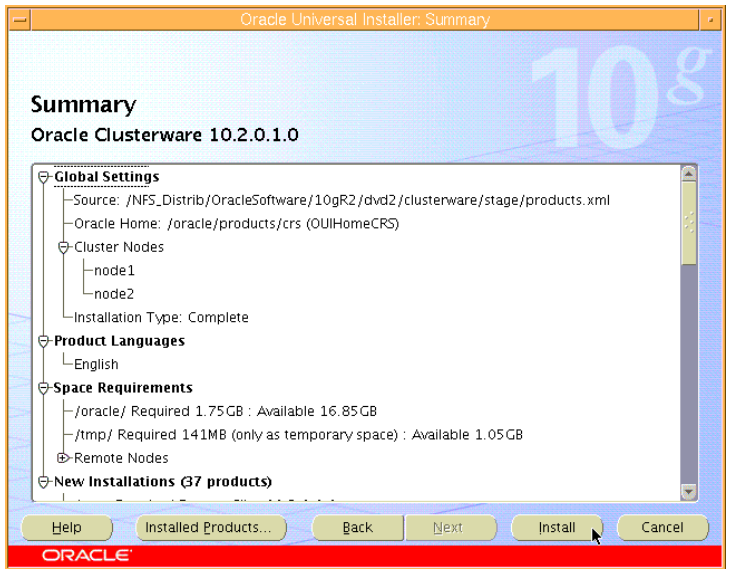
In our case, we will specify **"Normal Redundancy"** as described bellow :



Specify the Voting Disk location, this must be a shared location on the shared storage reachable from all nodes.
And you must have the read/wright permissions on this shared location from all nodes.

Then click Next ...

If problems happens at this stage, do verifiy that location specified does exist, and is reachable from each AIX node, with right read/write access, and user/group owner.

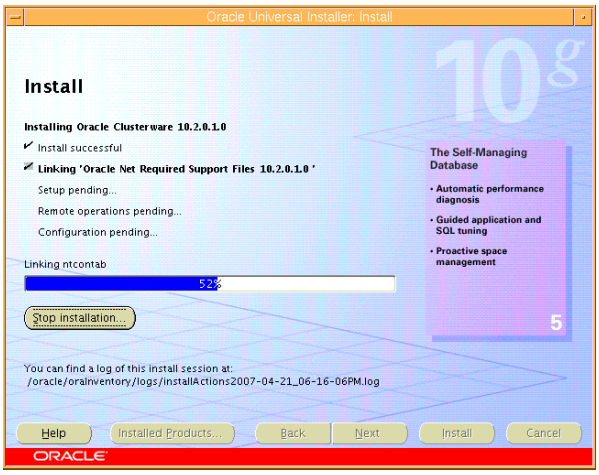


Summary :

Check Cluster Nodes and Remote Nodes lists.

The OUI will install the Oracle CRS software on to the local node, and then copy this information to the other selected nodes.

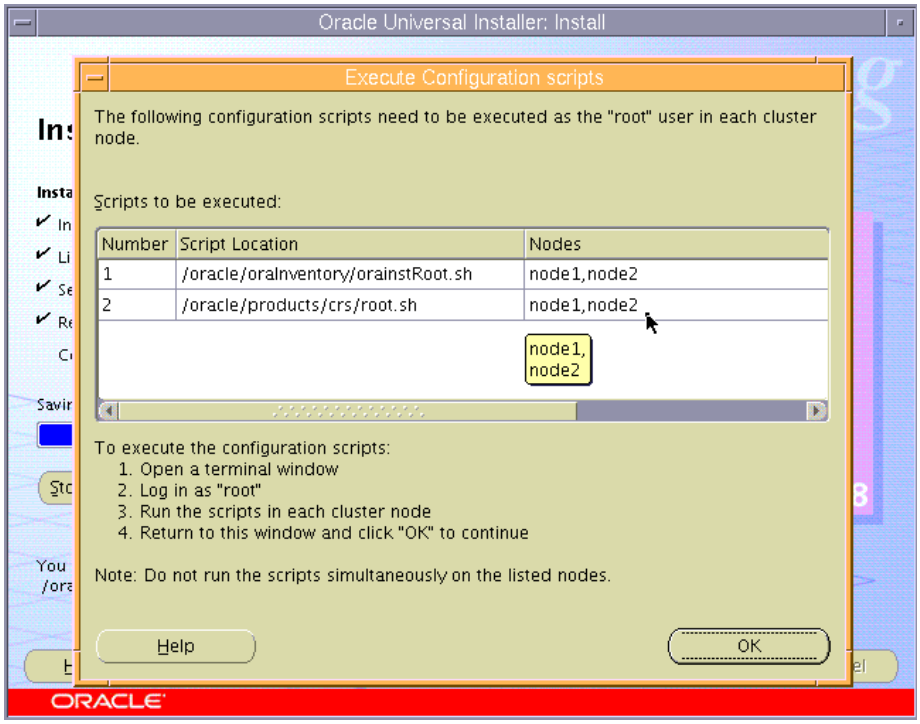
Then click Install ...



Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the 2 others selected nodes.

Just wait for the next screen ...



Execute Configuration Scripts :

KEEP THIS WINDOWS OPEN

AND Do execute scripts in the following order, waiting for each to succeed before running the next one !!!

AS root :

1 On node1, Execute orainstRoot.sh

2 On node2, Execute orainstRoot.sh

3 On node1, Execute root.sh

4 On node2, Execute root.sh

orainstRoot.sh :

Execute the orainstRoot.sh on all nodes.
The file is located in $ORACLE_BASE/oraInventory (OraInventory home) on each nodes

On node1 as root, execute ./orainstRoot.sh

```
{node1:root}/ oracle /oraInventory ->./orainstRoot.sh
Changing permissions of /oracle/oraInventory to 775.
Changing groupname of /oracle/oraInventory to dba.
The execution of the script is complete
{node1:root}/oracle/oraInventory ->
```

THEN On node2 as root, execute ./orainstRoot.sh

```
{node2:root}/oracle/oraInventory ->./orainstRoot.sh
Changing permissions of /oracle/oraInventory to 775.
Changing groupname of /oracle/oraInventory to dba.
The execution of the script is complete
{node2:root}/oracle/oraInventory ->
```

Before running root.sh script on each node, please check the following :

Check if your public network card (en0 in our case) is a standard network adapter, or a virtual network ethernet.

Issue the following command as root :

You should get the following if en0 is a normal network interface on each node (node1, node2) :

{node1:root} / -> **entstat –d en0**

```
------------------------------------------------------------
ETHERNET STATISTICS (en0) :
Device Type: 2-Port Gigabit Ethernet-SX PCI-X Adapter (14108802)
Hardware Address: 00:09:6b:ee:61:fc
Elapsed Time: 0 days 20 hours 3 minutes 49 seconds

Transmit Statistics:                          Receive Statistics:
--------------------                          -------------------
Packets: 0                                    Packets: 0
Bytes: 0                                      Bytes: 0
Interrupts: 0                                 Interrupts: 0
Transmit Errors: 0                            Receive Errors: 0
Packets Dropped: 0                            Packets Dropped: 0
                                              Bad Packets: 0
Max Packets on S/W Transmit Queue: 1
S/W Transmit Queue Overflow: 0
Current S/W+H/W Transmit Queue Length: 1

Broadcast Packets: 0                          Broadcast Packets: 0
Multicast Packets: 0                          Multicast Packets: 0
No Carrier Sense: 0                           CRC Errors: 0
DMA Underrun: 0                               DMA Overrun: 0
Lost CTS Errors: 0                            Alignment Errors: 0
Max Collision Errors: 0                       No Resource Errors: 0
Late Collision Errors: 0                      Receive Collision Errors: 0
Deferred: 0                                   Packet Too Short Errors: 0
SQE Test: 0                                   Packet Too Long Errors: 0
Timeout Errors: 0                             Packets Discarded by Adapter: 0
Single Collision Count: 0                     Receiver Start Count: 0
Multiple Collision Count: 0
Current HW Transmit Queue Length: 1

General Statistics:
-------------------
No mbuf Errors: 0
Adapter Reset Count: 0
Adapter Data Rate: 2000
Driver Flags: Up Broadcast Simplex
         Limbo 64BitSupport ChecksumOffload
         PrivateSegment LargeSend DataRateSet

2-Port Gigabit Ethernet-SX PCI-X Adapter (14108802) Specific Statistics:
--------------------------------------------------------------------
Link Status : Up
Media Speed Selected: Auto negotiation
Media Speed Running: Unknown
PCI Mode: PCI-X (100-133)
PCI Bus Width: 64-bit
Latency Timer: 144
Cache Line Size: 128
Jumbo Frames: Disabled
TCP Segmentation Offload: Enabled
TCP Segmentation Offload Packets Transmitted: 0
TCP Segmentation Offload Packet Errors: 0
Transmit and Receive Flow Control Status: Disabled
Transmit and Receive Flow Control Threshold (High): 45056
Transmit and Receive Flow Control Threshold (Low): 24576
Transmit and Receive Storage Allocation (TX/RX): 16/48
```

You should get the following if **en0** (public network interface) is a normal network interface :

IF so, go to the next step, executing the root.sh script

{node1:root}/ ->
entstat -d en0 | grep -iE ".*link.*status.*:.*up.*"
Link Status : Up
{node1:root}/ ->

OR

{node1:root} / -> **entstat –d en0**

You should get the following if en0 is a normal network interface :

```
ETHERNET STATISTICS (en0) :
Device Type: Virtual I/O Ethernet Adapter (l-lan)
Hardware Address: ee:51:60:00:10:02
Elapsed Time: 0 days 20 hours 3 minutes 24 seconds

Transmit Statistics:                          Receive Statistics:
--------------------                          -------------------
Packets: 136156                               Packets: 319492
Bytes: 19505561                               Bytes: 142069339
Interrupts: 0                                 Interrupts: 285222
Transmit Errors: 0                            Receive Errors: 0
Packets Dropped: 0                            Packets Dropped: 0
                                              Bad Packets: 0
Max Packets on S/W Transmit Queue: 0
S/W Transmit Queue Overflow: 0
Current S/W+H/W Transmit Queue Length: 0

Broadcast Packets: 208                        Broadcast Packets: 241831
Multicast Packets: 2                          Multicast Packets: 0
No Carrier Sense: 0                           CRC Errors: 0
DMA Underrun: 0                               DMA Overrun: 0
Lost CTS Errors: 0                            Alignment Errors: 0
Max Collision Errors: 0                       No Resource Errors: 0
Late Collision Errors: 0                      Receive Collision Errors: 0
Deferred: 0                                   Packet Too Short Errors: 0
SQE Test: 0                                   Packet Too Long Errors: 0
Timeout Errors: 0                             Packets Discarded by Adapter: 0
Single Collision Count: 0                     Receiver Start Count: 0
Multiple Collision Count: 0
Current HW Transmit Queue Length: 0

General Statistics:
-------------------
No mbuf Errors: 0
Adapter Reset Count: 0
Adapter Data Rate: 20000
Driver Flags: Up Broadcast Running
         Simplex 64BitSupport DataRateSet

Virtual I/O Ethernet Adapter (l-lan) Specific Statistics:
---------------------------------------------------------
RQ Length: 4481
No Copy Buffers: 0
Trunk Adapter: False
Filter MCast Mode: False
Filters: 255  Enabled: 1  Queued: 0  Overflow: 0
LAN State: Operational

Buffers        Reg  Alloc    Min     Max   MaxA  LowReg
 tiny          512    512    512    2048    512     509
 small         512    512    512    2048    553     502
 medium        128    128    128     256    128     128
 large          24     24     24      64     24      24
 huge           24     24     24      64     24      24
```

OR You should get the following if en0 is a virtual network interface :

IF so, do modifiy the ragvip script on each node before running the root.sh script, as described bellow.

➔ Bug 4437469: RACGVIP NOT WORKING ON SERVER WITH ETHERNET VIRTUALISATION This is only required if using AIX Virtual Interfaces for the Oracle Database 10.2.0.1 RAC public network

➔ It's Fixed with patchset 10.2.0.2

entstat -d en0 | grep -iE ".*lan state:.*operational.*"

{node1:root}/ ->
entstat –d en0 | grep -iE ".*lan State:.*operational.*"
LAN State: Operational
{node1:root}/ ->

in the $CRS_HOME/bin/racgvip script, you will have to modify the foolowing :

$ENTSTAT -d $_IF | $GREP -iEq ".*link.*status.*:.*up.*"
To be replaced by
$ENTSTAT -d $_IF | $GREP -iEq ".*lan state:.*operational.*"

At this stage, you should execute "root.sh" script :

Start with node 1 and wait for the result before executing on node 2.

This file is located in $ORACLE_BASE/crs directory on each node ("/oracle/products/crs" in our case).

ONLY For information, The root.sh script is executing two sub scripts, and one is the rootconfig.sh script which has interresting information to have a look at :

DO not modify the file !!!

What does the rootconfig.sh script executed by root.sh :

```
# rootconfig.sh for Oracle CRS homes
#
#    This is run once per node during the Oracle CRS install.
#    This script does the following:
#    1) Stop if any GSDs are running from 9.x oracle homes
#    2) Initialize new OCR device or upgrade the existing OCR device
#    3) Setup OCR for running CRS stack
#    4) Copy the CRS init script to init.d for init process to start
#    5) Start the CRS stack
#    6) Configure NodeApps if CRS is up and running on all nodes
```

Variables used by root.sh script,
the values are the result of your inputs in the Oracle Clusterware Universal Installer.

You can check the values to see if there are OK.

```
SILENT=false
ORA_CRS_HOME=/oracle/products/crs
CRS_ORACLE_OWNER=oracle
CRS_DBA_GROUP=dba
CRS_VNDR_CLUSTER=false
CRS_OCR_LOCATIONS=/dev/ocr_disk1,/dev/ocr_disk2
CRS_CLUSTER_NAME=crs
CRS_HOST_NAME_LIST=node1,1,node2,2
CRS_NODE_NAME_LIST=node1,1,node2,2
CRS_PRIVATE_NAME_LIST=node1-rac,1,node2-rac,2
CRS_LANGUAGE_ID='AMERICAN_AMERICA.WE8ISO8859P1'
CRS_VOTING_DISKS=/dev/voting_disk1,/dev/voting_disk2,/dev/voting_disk3
CRS_NODELIST=node1,node2
CRS_NODEVIPS='node1/node1-vip/255.255.255.0/en0,node2/node2-vip/255.255.255.0/en0'
```

FIRST On node1

As root, Execute /oracle/crs/root.sh

When finished, CSS deamon should be active on node 1.

Check for line "CSS is active on these nodes.

- **node1**

```
{node1:root}/oracle/products/crs -> ./root.sh
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
Checking to see if Oracle CRS stack is already configured

Setting the permissions on OCR backup directory
Setting up NS directories
Oracle Cluster Registry configuration upgraded successfully
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 1: node1 node1-rac node1
node 2: node2 node2-rac node2
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
Now formatting voting device: /dev/voting_disk1
Now formatting voting device: /dev/voting_disk2
Now formatting voting device: /dev/voting_disk3
Format of 3 voting devices complete.
Startup will be queued to init within 30 seconds.
Adding daemons to inittab
Adding daemons to inittab
Expecting the CRS daemons to be up within 600 seconds.
CSS is active on these nodes.
        node1
CSS is inactive on these nodes.
        node2
Local node checking complete.
Run root.sh on remaining nodes to start CRS daemons.
{node1:root}/oracle/products/crs ->
```

Don't worry about "WARNING: directory '/oracle/products' is not owned by root"
"WARNING: directory '/oracle' is not owned by root"

➔ This is just a message to forget …

IF CSS is not active at the end of the root.sh script :

➔ Check your network , shared disks configuration, and owner and access permissions (read/write) on OCR and Voting disks from each participating node. And execute again the root.sh script on node having the problem.

➔If CCS start on one node, but not on the others, Check shared Disks (OCR/Voting) for concurrent read/write access from all nodes, using unix dd command.

➔ If ASM or GPFS is implemented with HACMP installed and configured for other purposes then having database on concurrent raw devices, You must declare disks ressources in HACMP to be able to start the CRS (CSS).

➔ If ASM or GPFS is implemented, and HACMP is installed but not used at all, THEN remove HACMP or declare disks ressources in HACMP to be able to start the CRS (CSS).

➚ The IBM AIX clustering layer, HACMP filesets, MUST NOT be installed if you've chosen an implementation without HACMP. If this layer is implemented for other purpose, disks ressources necessary to install and run CRS data will have to be part of an HACMP volume group resource.

If you have previously installed HACMP, you must remove :

- HACMP filesets (cluster.es.*)
- rsct.hacmp.rte
- rsct.compat.basic.hacmp.rte
- rsct.compat.clients.hacmp.rte

If you did run a first installation of the Oracle Clusterware (CRS) with HACMP installed,
➔Check if /opt/ORCLcluster directory does exist and if so, remove it on all nodes.

TO BE ABLE TO RUN AGAIN the root.sh script on the node, you must :

Either

- Clean the failed CRS installation, and start again the CRS installation procedure.

➔ Metalink Note 239998.1 - 10g RAC: How to Clean Up After a Failed CRS Install
➔ Only Supported method by Oracle.

OR

- **Do the following just to find out and solve the problem** without installing again at each try, **And when solved, follow again the Metalink Note 239998.1** - 10g RAC: How to Clean Up After a Failed CRS Install to clean properly the system, and start again the installation as supported by oracle.

As root user on each node :

$CRS_HOME/bin/crsctl stop crs (to clean any remaining crs deamons)
rmitab h1 ➔ this will remove oracle CRS entry in the /etc/inittab
rmitab h2
rmitab h3
rm –Rf /opt/ORCL*
kill remaining process from output : ps-ef|grep crs and ps-ef|grep d.bin
rm –R /etc/oracle/*

For all OCR and Voting Disks :	Change owner, group (oracle:dba) and permission (660) for /dev/ocr_disk1 and /dev/voting_disk1 on each node of the cluster on node1 and node2 … Erase OCR and Voting diks content ➔ Format (Zeroing) on the disks from one node : dd if=/dev/zero of=/dev/ocr_disk1 bs=8192 count=25000 & dd if=/dev/zero of=/dev/voting_disk1 bs=8192 count=25000 &

THEN On node2

As root, Execute /oracle/crs/root.sh

When finished, CSS deamon should be active on node 1, 2.

You should have the following final result :

CSS is active on these nodes.
node1
node2

If CSS is not active on all nodes, or on one of the nodes, this means that you could have a problem with the network configuration, or the shared disks configuration for accessing OCR and Voting Disks.

➔ Check your network , shared disks configuration, and owner and access permissions (read/write) on OCR and Voting disks from each participating node. And execute again the root.sh script on node having the problem.

Check also as oracle user the following command from each node :

{node1:root}/ ->su - oracle
{node1:oracle}/oracle/crs/bin ->/oracle/crs/bin/**olsnodes**
node1
node2
{node1:oracle}/ oracle /crs/bin ->
rsh node2
{node2:oracle}/ oracle ->**olsnodes**
node1
node2
{node2:oracle}/ oracle ->

THEN On node2

As root, Execute /oracle/crs/root.sh

When finished, CSS deamon should be active on node 1 and 2.

Check for line "CSS is active on these nodes.

- **node1**
- **node2**

```
{node2:root}/oracle/products/crs -> ./root.sh
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
Checking to see if Oracle CRS stack is already configured

Setting the permissions on OCR backup directory
Setting up NS directories
Oracle Cluster Registry configuration upgraded successfully
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
clscfg: EXISTING configuration version 3 detected.
clscfg: version 3 is 10G Release 2.
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 1: node1 node1-rac node1
node 2: node2 node2-rac node2
clscfg: Arguments check out successfully.

NO KEYS WERE WRITTEN. Supply -force parameter to override.
-force is destructive and will destroy any previous cluster
configuration.
Oracle Cluster Registry for cluster has already been initialized
Startup will be queued to init within 30 seconds.
Adding daemons to inittab
Adding daemons to inittab
Expecting the CRS daemons to be up within 600 seconds.
CSS is active on these nodes.
        node1
        node2
CSS is active on all nodes.
Oracle CRS stack installed and running under init(1M)
Running vipca(silent) for configuring nodeapps

Creating VIP application resource on (2) nodes...
Creating GSD application resource on (2) nodes...
Creating ONS application resource on (2) nodes...
Starting VIP application resource on (2) nodes...
Starting GSD application resource on (2) nodes...
Starting ONS application resource on (2) nodes...


Done.
{node2:root}/oracle/products/crs ->
```

Don't worry about "WARNING: directory '/oracle/products' is not owned by root"
"WARNING: directory '/oracle' is not owned by root"

➔ This is just a message to forget …

On the second node, at he end of the root.sh script :

You should have the following lines :

Running vipca(silent) for configuring nodeapps

Creating VIP application resource on (2) nodes...
Creating GSD application resource on (2) nodes...
Creating ONS application resource on (2) nodes...
Starting VIP application resource on (2) nodes...
Starting GSD application resource on (2) nodes...
Starting ONS application resource on (2) nodes...

IF NOT, Check for the line :

"en0 is not public. Public interfaces should be used to configure virtual Ips",

And read next page to understand and solve it !!!

THEN RUN VIPCA manualy as explained on next page !!!

At this stage, if you get "en0 is not public. Public interfaces should be used to configure virtual Ips", Following execution of root.sh on last node, node2 in our case, THEN read the following :

VIP should have been configured in silent mode with the root.sh scripts executed on node2

➔ THIS IS NOT THE CASE

Note 316583.1 – VIPCA FAILS COMPLAINING THAT INTERFACE IS NOT PUBLIC

Symptoms
During CRS install while running root.sh, The following messages are displayed
Oracle CRS stack installed and running under init(1M)
Running vipca(silent) for configuring nodeapps
The given interface(s), "**en0" is not public. Public interfaces should be used to configure virtual IPs.**

Cause
When verifying the IP addresses, VIP uses calls to determine if a IP address is valid or not. In this case, VIP finds that the IPs are non routable (For example IP addresses like 192.168.* and 10.10.*.)
Oracle is aware that the IP's can be made public but since mostly such IP's are used for Private, it display this error message.

Solution
The workaround is to re-run vipca manually as root
#./vipca
or add the VIP using srvctl add nodeapps

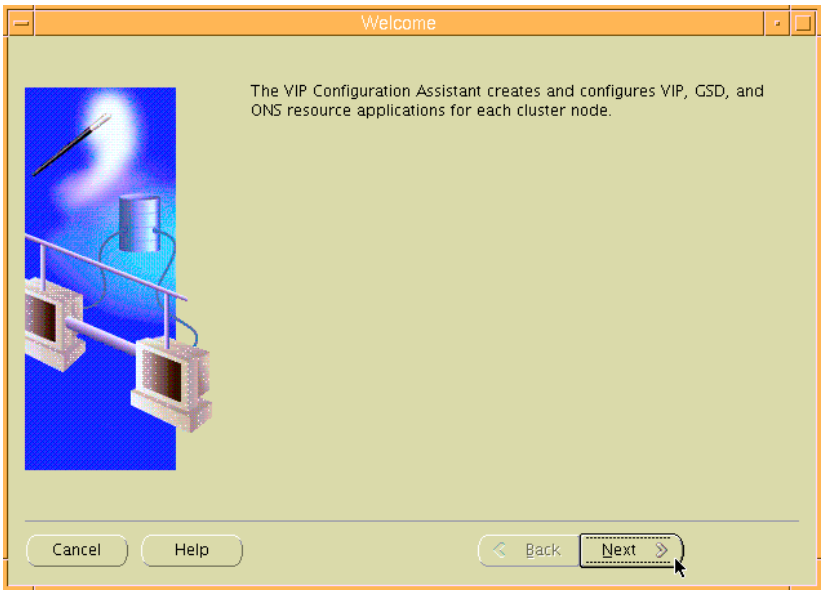
➔ YOU MUST CONFIGURE VIP by running vipca script as root user.

logon as root on second node

{node1:root}/ ->**export DISPLAY**
{node1:root}/ ->**cd $CRS_HOME/bin**
{node1:root}/ ->**./vipca**

On first or second node as root user, you must setup the DISPLAY

before running the vipca script located in /oracle/crs/bin

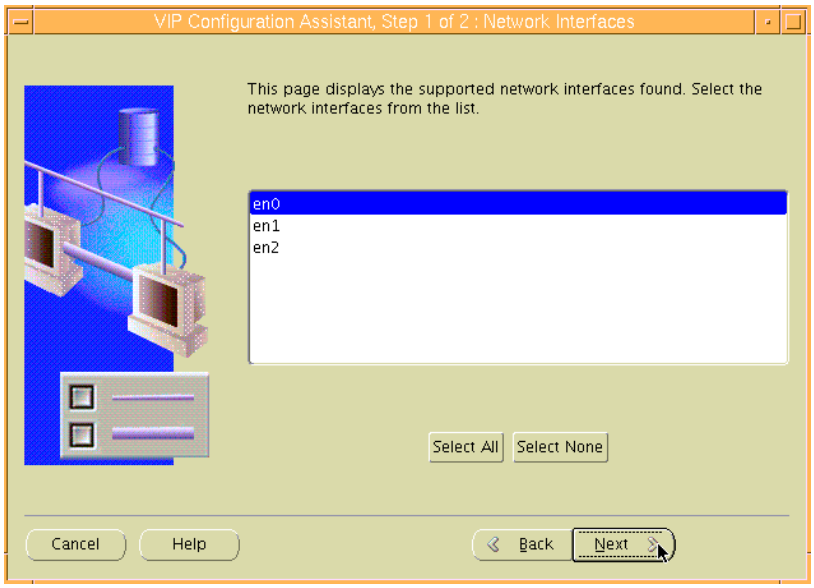


The VIP "Welcome" graphical screen will appear at the end of the root.sh script

Then click Next ...

Just to remember !!! **Public, Private, and Virtual Host Name layout**

	Public	VIP	RAC Interconnect	RAC Interconnect *Backup*
Network card on each node	en0	en0	en1	en2



1 of 2 : Select one and only one network interface.

Select the network interface corresponding to the Public Network Remember that each public network card on each node must have the same name, "**en0**" for example in our case.

en1 is the RAC Interconnect, or private network.

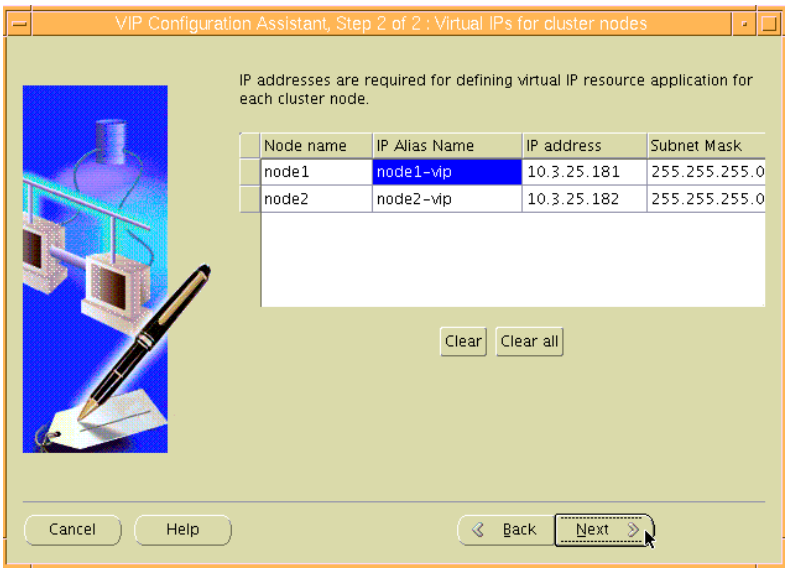
Please check with "ifconfig –a" on each node as root.

Select "en0" in our case

Then click Next ...

Just to remember !!! **Public, Private, and Virtual Host Name layout**

Public		VIP		RAC Interconnect (Private Network)	
en0				en1	
Node Name	IP	Node Name	IP	Node Name	IP
node1	10.3.25.81	node1-vip	10.3.25.181	node1-rac	10.10.25.81
node2	10. 3.25.82	node2-vip	10. 3.25.182	node1-rac	10.10.25.82

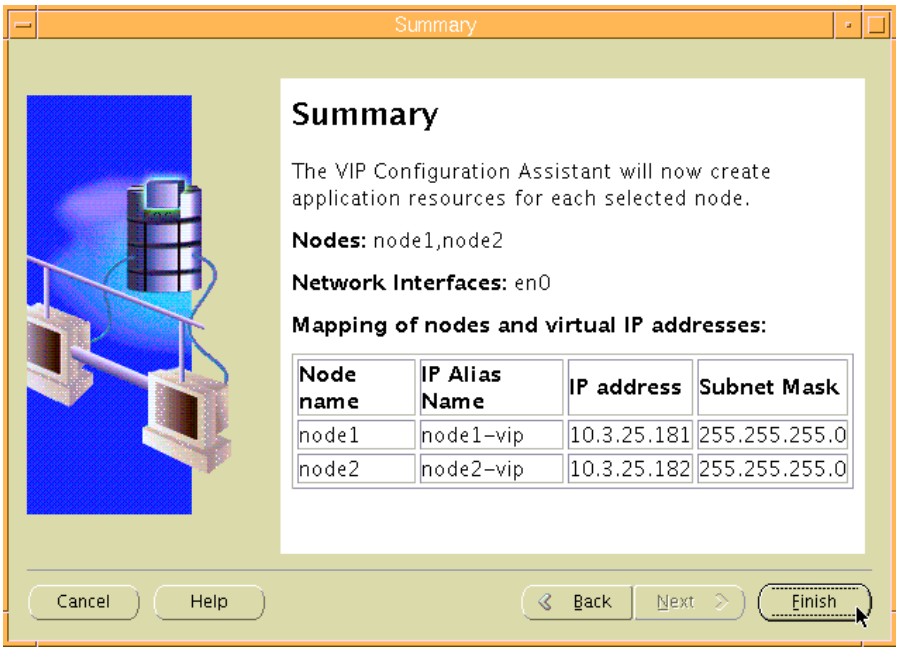


2 of 2 :

In the Virtual IPs for cluster nodes screen, you must provide the VIP node name for node1 and stroke the TAB key to automatically feel the rest.

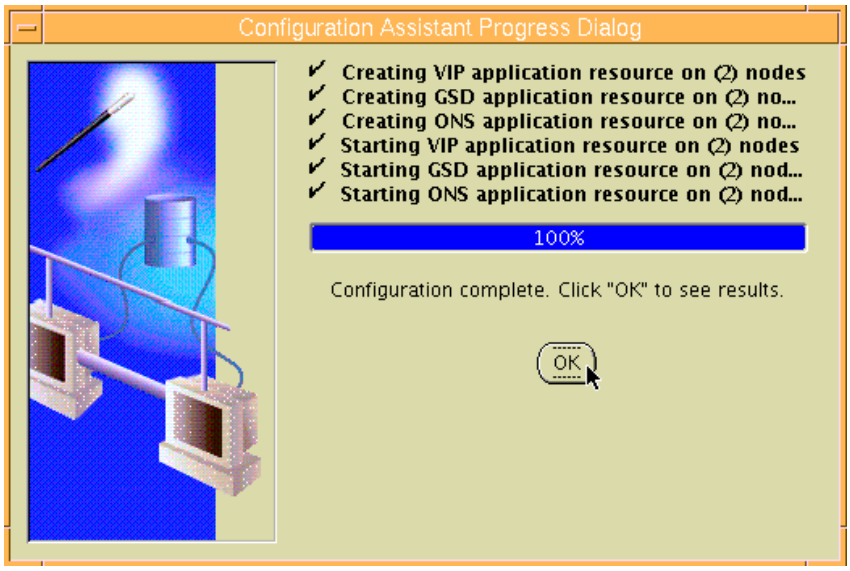
Check validity of the entries before proceeding.

Then click Next ...



The Summary screen will appear, please validate the entries, or go back to modify.

Then click Finish ...



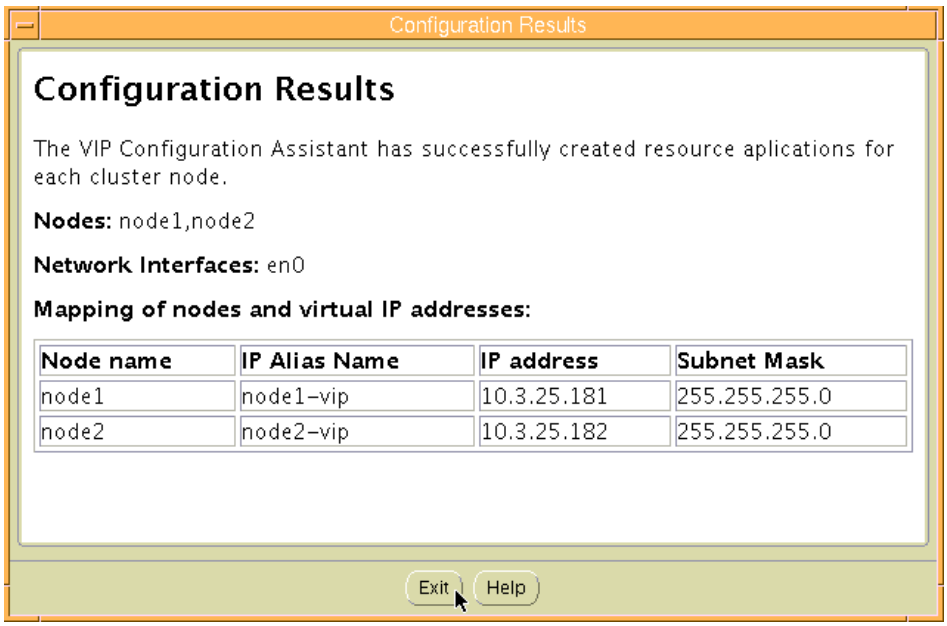
The VIP configuration Assistant will proceed with creation, configuration and startup of all application resources on all selected nodes.

VIP, GSD and ONS will be the application resources to be created.

Wait while progressing ...

If you don't get any errors, you'll be prompted to click OK as the configuration is 100% completed.

Then click OK ...



Check the Configuration results.

Then click Exit ...

Using "ifconfig –a" on each node,
check that each network card configured for Public network is mapping a virtual IP.

On node 1 :

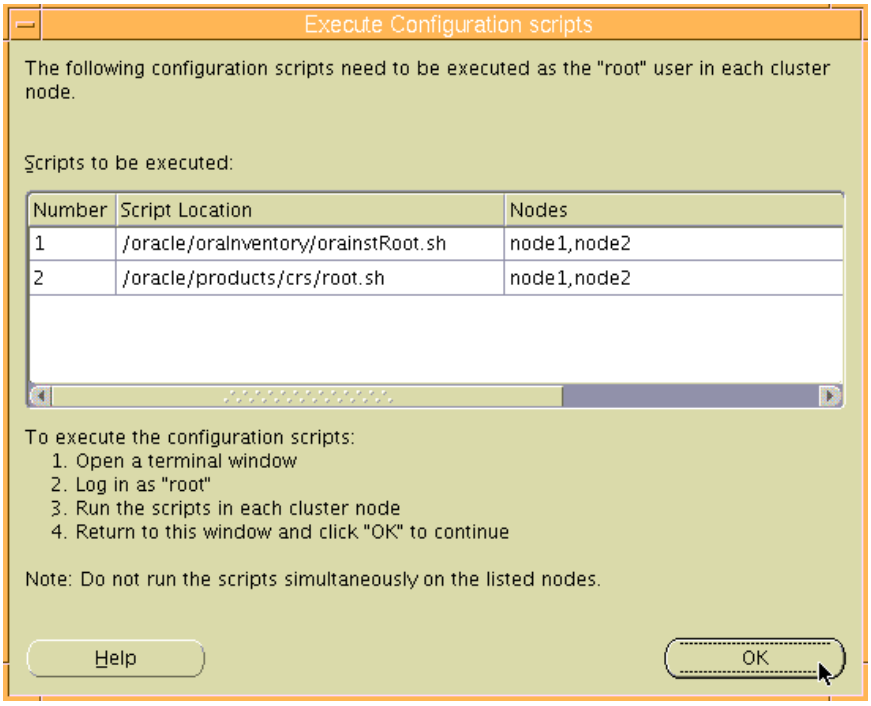
```
{node1:root}/ ->ifconfig -a
en0: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.3.25.81 netmask 0xffffff00 broadcast 10.3.25.255
        inet 10.3.25.181 netmask 0xffffff00 broadcast 10.3.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en1: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.10.25.81 netmask 0xffffff00 broadcast 10.10.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en2: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 20.20.25.81 netmask 0xffffff00 broadcast 20.20.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
lo0: flags=e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT>
        inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
        inet6 ::1/0
         tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1
{node1:root}/ ->
```

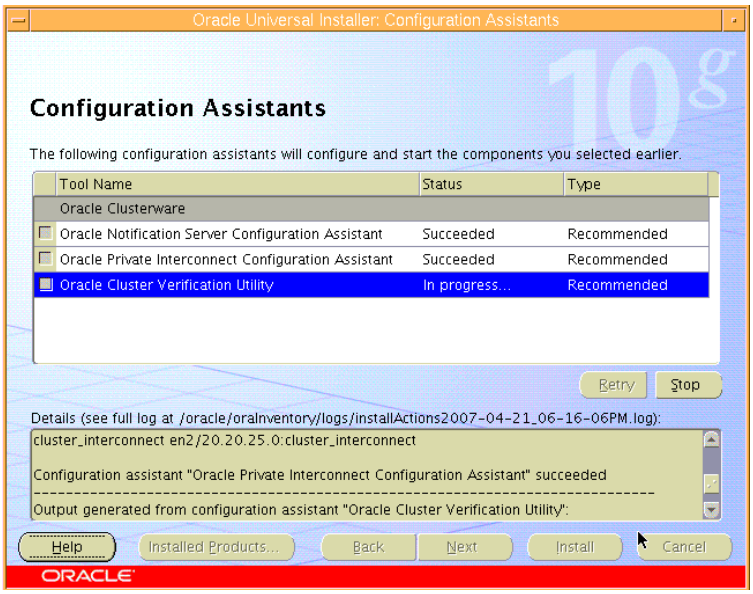
On node 2 :

```
{node2:root}/ ->ifconfig -a
en0: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.3.25.82 netmask 0xffffff00 broadcast 10.3.25.255
        inet 10.3.25.182 netmask 0xffffff00 broadcast 10.3.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en1: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 10.10.25.82 netmask 0xffffff00 broadcast 10.10.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
en2: flags=1e080863,80<UP,BROADCAST,NOTRAILERS,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
        inet 20.20.25.82 netmask 0xffffff00 broadcast 20.20.25.255
         tcp_sendspace 131072 tcp_recvspace 65536
lo0: flags=e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT>
        inet 127.0.0.1 netmask 0xff000000 broadcast 127.255.255.255
        inet6 ::1/0
         tcp_sendspace 131072 tcp_recvspace 131072 rfc1323 1
{node2:root}/ ->
```

Coming back to this previous screen.

Just click OK to continue ...





Configuration Assistants :

3 configuration assistants will be automatically executed.

- "Oracle notification Server Configuration Assistant"
- "Oracle Private Interconnect Assistant"
- "Oracle Cluster Verification Utility"

If successful, Next screen "End of Installation" will appear automatically !!!

If not Check for the result to be successful.

Then click Next ...

If all or parts of assistants are failed or not executed, check for problems in log /oracle/oraInventory/logs/installActions?????.log (as shown on the runInstaller window) and solve them.

Check also /oracle/crs/cfgtoollogs/configToolFailedCommands

```
{node1:oracle}/oracle/products/crs/cfgtoollogs -> cat configToolAllCommands
# Copyright (c) 1999, 2005, Oracle. All rights reserved.
/oracle/crs/bin/racgons add_config node1:6200  node2:6200
/oracle/crs/bin/oifcfg setif -global  en0/10.3.25.0:public
en1/10.10.25.0:cluster_interconnect
/oracle/crs/bin/cluvfy stage -post crsinst -n node1,node2
{node1:oracle}/oracle/products/crs/cfgtoollogs ->
```

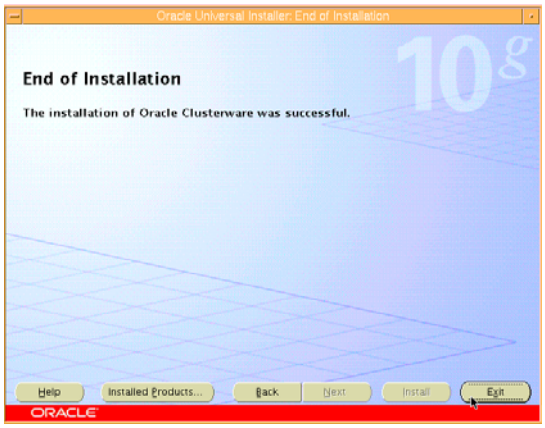
If you miss those steps or closed the previous screen of the runInstaller, you will have to run them manually before moving to the next step. Just adapt the lines with your own setting (node names, public/private network).

On one node as oracle user :

For "Oracle notification Server Configuration Assistant" ➔
/oracle/home10g/crs/bin/**racgons add_config node1:6200 node2:6200**

For "Oracle Private Interconnect Assistant" ➔
/oracle/home10g/crs/bin/**oifcfg setif -global en0/10.3.25.0:public en1/10.25.25.0:cluster_interconnect**

For "Oracle Cluster Verification Utility" ➔
/oracle/home10g/crs/bin/**cluvfy stage -post crsinst -n node1,node2**



End of Installation

Then click Exit ...

14.2 CRS Post-Installation Tasks

14.2.1 Update Oracle user .profile

To be done on each node.

Oracle environment : vi **$HOME/.profile** file in Oracle's home directory

Add the entries in bold blue color

```
export ORACLE_BASE=/oracle
export AIXTHREAD_SCOPE=S        (S for system-wide thread scope)
export TEMP=/tmp
export TMP=/tmp
export TMPDIR=/tmp
umask 022
export ORA_CRS_HOME=$ORACLE_BASE/products/crs
export CRS_HOME=$ORA_CRS_HOME
export LD_LIBRARY_PATH=:$ORA_CRS_HOME/lib:$ORA_CRS_HOME/lib32
export LD_LIBRARY_PATH
export LIBPATH=$LD_LIBRARY_PATH
export PATH=$ORA_CRS_HOME/bin:$PATH
```

Do disconnect from oracle user, and reconnect to load modified $HOME/.profile

14.2.2 Verify parameter CSS misscount

MISSCOUNT DEFINITION AND DEFAULT VALUES

The CSS misscount parameter represents the maximum time, in seconds, that a heartbeat can be missed before entering into a cluster reconfiguration to evict the node. The following are the default values for the misscount parameter and their respective versions when using Oracle Clusterware*:

10gR1 & 10gR2:

Linux	60 Seconds
Unix	30 Seconds
VMS	30 Seconds
Windows	30 Seconds

*CSS misscount default value when using vendor (non-Oracle) clusterware is 600 seconds. This is to allow the vendor clusterware ample time to resolve any possible split brain scenarios.

Subject: **CSS Timeout Computation in RAC 10g (10g Release 1 and 10g Release 2)**
Doc ID: **Note:294430.1**

Check css misscount

```
{node1:root}/oracle/products/crs -> crsctl get css misscount
Configuration parameter misscount is not defined.

➔ we should have a defined value
```

Check css disktimeout

```
{node1:root}/oracle/products/crs -> crsctl get css disktimeout
200
```

Check css reboottime

```
{node1:root}/oracle/products/crs -> crsctl get css reboottime
3
```

To compute the right values, do read metalink note 294430.1, and use following note to change the value :
Subject: **10g RAC: Steps To Increase CSS Misscount, Reboottime and Disktimeout** Doc ID: **Note:284752.1**

Set css misscount

And check …

```
Keep only one node up and running, stop the others

Backup the content of your OCR !!!

Do modify the CSS parameters with the crsctl command as root user

{node1:root}/oracle/products/crs -> crsctl set css misscount 30
Configuration parameter misscount is now set to 30.

{node1:root}/oracle/products/crs -> crsctl get css misscount
30

Restart all other nodes !!!
```

14.2.3 Cluster Ready Services Health Check

Check CRS processes on each nodes :

```
{node1:root}/oracle/products/crs -> ps -ef|grep d.bin
  oracle  901166 1073328   5 12:27:51      -  0:06 /oracle/products/crs/bin/ocssd.bin
    root 1204320 1212584   1 13:10:41  pts/0  0:00 grep d.bin
  oracle 1216656       1   0 12:27:16      -  0:00 /oracle/products/crs/bin/evmd.bin
    root 1220722  950452  11 12:27:49      -  0:22 /oracle/products/crs/bin/crsd.bin reboot
{node1:root}/oracle/products/crs/log/node1 ->
```

You have completed the CRS install. Now you want to verify if the install is valid.
To Ensure that the CRS install on all the nodes is valid, the following should be checked on all the nodes.

1. **Ensure that you have successfully completed running root.sh on all nodes during the install.** (Please do not re-run root.sh, this is very dangerous and might corrupt your installation, The object of this step is to only confirm if the root.sh was run successfully after the install

2. **Run the command $ORA_CRS_HOME/bin/crs_stat.** Please ensure that this command does not error out but dumps the information for each resource. It does not matter what CRS stat returns for each resource. If the crs_stat exits after printing information about each resource then it means that the CRSD daemon is up and the client crs_stat utility can communicate with it.

   - This will also indicate that the CRSD can read the OCR.
   - If the crs_stat errors out with CRS-0202: No resources are registered, Then this means that there are no resources registered, and at this stage you missed the VIP configuration. This is not an error but is mostly because at this stage you missed the VIP configuration.
   - If the crs_stat errors out with CRS-0184: Cannot communicate with the CRS daemon, Then this means the CRS deamons are not started.

Execute

crs_stat –t

on one node as oracle user :

```
{node2:oracle}/oracle ->crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node2:oracle}/oracle ->
```

3. **Run the command $ORA_CRS_HOME/bin/olsnodes**. This should return all the nodes of the cluster. Successful run of this command would mean that the css is up and running. Also the CSS from each node can talk to the CSS of other nodes.

Execute

olsnodes

on both node as oracle user :

```
{node1:root}/ -> su - oracle
{node1:oracle}/->/oracle/products/crs/bin/olsnodes
node1
node2
{node1:oracle -> rsh node2
{node2:oracle}/oracle ->/oracle/products/crs/bin/olsnodes
node1
node2
{node2:oracle}/oracle ->
```

4. Output of crsctl check crs / cssd / crsd / evmd returns "…. daemon appears healthy"

CRS health check

```
{node1:oracle}/oracle/products/crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oracle/products /crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmd health check

```
{node1:oracle}/oracle//products crs/bin ->crsctl check  cssd
CSS appears healthy
{node1:oracle}/oracle/products/crs/bin ->crsctl check  crsd
CRS appears healthy
{node1:oracle}/oracle/products/crs/bin ->crsctl check  evmd
EVM appears healthy

{node2:oracle}/oracle/crs/bin ->crsctl check  cssd
CSS appears healthy
{node2:oracle}/oracle/crs/bin ->crsctl check  crsd
CRS appears healthy
{node2:oracle}/oracle/crs/bin ->crsctl check  evmd
EVM appears healthy
```

CRS software version query

```
{node1:oracle}/oracle/products/crs/bin ->crsctl query crs activeversion
CRS software version on node [node1] is [10.2.0.1.0]
{node1:oracle}/oracle/products/crs/bin ->

{node2:oracle}/oracle/products/crs/bin ->crsctl query crs activeversion
CRS software version on node [node2] is [10.2.0.1.0]
{node2:oracle}/oracle/products/crs/bin ->
```

14.2.4 Interconnect Network configuration Checkup

After CRS installation is completed, verify that the public and cluster interconnect have been set to the desired values by entering the following commands as root:

Note : oifcfg is found in the <CRS HOME>/bin/oifcfg

oifcfg getif

Subject: **How to Change Interconnect/Public Interface IP Subnet in a 10g Cluster** Doc ID: **Note:283684.1**

This command should return values for global "public" and global "cluster_interconnect"; for example:

```
{node1:oracle}/oracle -> oifcfg getif
en0  10.3.25.0  global  public
en1  10.10.25.0  global  cluster_interconnect
{node1:oracle}/oracle ->
```

If the command does not return a value for global cluster_interconnect, enter the following commands:

```
{node1:oracle}/oracle ->oifcfg delif -global
# oifcfg setif -global <interface name>/<subnet>:public
# oifcfg setif –global <interface name>/<subnet>:cluster_interconnect
```

For example:

```
{node1:oracle}/oracle -> oifcfg delif -global
{node1:oracle}/oracle -> oifcfg setif -global en0/10.3.25.0.0:public
{node1:oracle}/oracle -> oifcfg setif -global en1/10.10.25.0.0:cluster_interconnect
```

Add a RAC interconnect backup :

```
{node1:oracle}/oracle -> oifcfg setif -global en2/20.20.25.0.0:cluster_interconnect
```

Enter the following command to verify the new settings:

```
# oifcfg getif

{node1:oracle}/oracle -> oifcfg getif
en0  10.3.25.0  global  public
en1  10.10.25.0  global  cluster_interconnect
en2  20.20.25.0  global  cluster_interconnect
{node1:oracle}/oracle ->
```

- **If necessary and only for troubleshooting purpose**, disable the automatic reboot of AIX nodes when node fail to communicate with CRS daemons, or fail to access OCR and Voting disk.

Subject: **10g RAC: Stopping Reboot Loops When CRS Problems Occur** Doc ID: **Note:239989.1**
Subject: **10g RAC: Troubleshooting CRS Reboots** Doc ID: **Note:265769.1**

If one node crashed after running dbca, netca tools, with CRS codedump and Authentication OSD error, check "crsd.log" file for missing $CRS_HOME/crs/crs/auth directory.

➔ THEN you need to re-create manually the missing directory, create the auth directory with correct owner, group, and permission using following metalink note :
Subject: **Crs Crashed With Authentication Osd Error** Doc ID: **Note:358400.1**

14.2.5 Oracle CLuster Registry content Check and Backup

Check Oracle Cluster Registry Integrity

As oracle user, Execute ocrcheck

```
{node1:oracle}/oracle/products/crs/bin ->ocrcheck
Status of Oracle Cluster Registry is as follows :
         Version                  :          2
         Total space (kbytes)     :     306972
         Used space (kbytes)      :       2976
         Available space (kbytes) :     303996
         ID                       : 1512272896
         Device/File Name         : /dev/ocr_disk1
                                    Device/File integrity check succeeded
         Device/File Name         : /dev/ocr_disk2
                                    Device/File integrity check succeeded

         Cluster registry integrity check succeeded

{node1:oracle}/oracle/products/crs/bin ->
```

Check Voting Disks

As oracle user, Execute :

```
{node1:oracle}/oracle -> crsctl query  css votedisk
 0.     0    /dev/voting_disk1
 1.     0    /dev/voting_disk2
 2.     0    /dev/voting_disk3

located 3 votedisk(s).
{node1:oracle}/oracle ->
```

AS root user :
Export Oracle Cluster Registry content

```
{node1:oracle}/oracle/products/crs/bin ->su
root's Password:
{node1:root}/oracle/products/crs/bin ->ocrconfig -export /oracle/ocr_export.dmp1 -s online
{node1:root}/oracle/products/crs/bin ->ls -la /oracle/*.dmp
-rw-r--r--   1 root     system       106420 Jan 30 18:30 /oracle/ocr_export.dmp
{node1:root}/oracle/products/crs/bin ->
```

➔ **you must not edit/modify this exported file**

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oracle/products/crs/bin ->ocrconfig -showbackup

node1     2006/01/30 16:03:03     /oracle/products/crs/cdata/crs

node1     2006/01/30 12:03:00     /oracle/products/crs/cdata/crs

node1     2006/01/30 08:02:59     /oracle/products/crs/cdata/crs

node1     2006/01/29 00:02:51     /oracle/products/crs/cdata/crs

node1     2006/01/25 13:02:10     /oracle/products/crs/cdata/crs

{node1:oracle}/oracle/products /crs/bin ->
```

14.3 SOME USEFULL COMMANDS

AS root user

Command to start/stop the CRS deamons :

To start the CRS :

```
{node1:root}/oracle -> crsctl start crs
Attempting to start CRS stack
The CRS stack will be started shortly
{node1:root}/oracle ->
```

To stop the CRS :

```
{node1:oracle}/oracle -> su
root's Password:
{node1:root}/oracle -> crsctl stop crs
Stopping resources. This could take several minutes.
Successfully stopped CRS resources.
Stopping CSSD.
Shutting down CSS daemon.
Shutdown request successfully issued.
{node1:root}/oracle ->
```

All crsctl command available :

```
{node1:root}/oracle -> crsctl
Usage: crsctl check  crs          - checks the viability of the CRS stack
       crsctl check  cssd         - checks the viability of CSS
       crsctl check  crsd         - checks the viability of CRS
       crsctl check  evmd         - checks the viability of EVM
       crsctl set    css <parameter> <value> - sets a parameter override
       crsctl get    css <parameter> - gets the value of a CSS parameter
       crsctl unset  css <parameter> - sets CSS parameter to its default
       crsctl query  css votedisk    - lists the voting disks used by CSS
       crsctl add    css votedisk <path> - adds a new voting disk
       crsctl delete css votedisk <path> - removes a voting disk
       crsctl enable  crs    - enables startup for all CRS daemons
       crsctl disable crs    - disables startup for all CRS daemons
       crsctl start crs  - starts all CRS daemons.
       crsctl stop  crs  - stops all CRS daemons. Stops CRS resources in case of cluster.
       crsctl start resources  - starts CRS resources.
       crsctl stop resources  - stops  CRS resources.
       crsctl debug statedump evm  - dumps state info for evm objects
       crsctl debug statedump crs  - dumps state info for crs objects
       crsctl debug statedump css  - dumps state info for css objects
       crsctl debug log css [module:level]{,module:level} ...
                            - Turns on debugging for CSS
       crsctl debug trace css - dumps CSS in-memory tracing cache
       crsctl debug log crs [module:level]{,module:level} ...
                            - Turns on debugging for CRS
       crsctl debug trace crs - dumps CRS in-memory tracing cache
       crsctl debug log evm [module:level]{,module:level} ...
                            - Turns on debugging for EVM
       crsctl debug trace evm - dumps EVM in-memory tracing cache
       crsctl debug log res <resname:level> turns on debugging for resources
       crsctl query crs softwareversion [<nodename>] - lists the version of CRS software installed
       crsctl query crs activeversion - lists the CRS software operating version
       crsctl lsmodules css - lists the CSS modules that can be used for debugging
       crsctl lsmodules crs - lists the CRS modules that can be used for debugging
       crsctl lsmodules evm - lists the EVM modules that can be used for debugging

 If necesary any of these commands can be run with additional tracing by
 adding a "trace" argument at the very front.
 Example: crsctl trace check css
{node1:root}/oracle ->
```

14.4 Accessing CRS logs

To view CRS logs

cd /oracle/products/crs/log/*nodename*/….

In our case nodename will be node1 for CRS logs on node1
cd /oracle/products/crs/log/node1

And nodename will be node2 for CRS logs on node2
cd /oracle/products/crs/log/node2

Contents example of ../crs/log/node1 with node1 :

```
{node1:root}/oracle/products/crs/log/node1 -> ls -la
total 256
drwxr-xr-t   8 root     dba             256 Mar 12 16:21 .
drwxr-xr-x   4 oracle   dba             256 Mar 12 16:21 ..
drwxr-x---   2 oracle   dba             256 Mar 12 16:21 admin
-rw-rw-r--   1 root     dba           24441 Apr 17 12:27 alertnode1.log
drwxr-x---   2 oracle   dba           98304 Apr 17 12:27 client
drwxr-x---   2 root     dba             256 Apr 16 13:56 crsd
drwxr-x---   4 oracle   dba             256 Mar 22 21:56 cssd
drwxr-x---   2 oracle   dba             256 Mar 12 16:22 evmd
drwxrwxr-t   5 oracle   dba            4096 Apr 16 12:48 racg
{node1:root}/oracle/products/crs/log/node1 ->
```

Look at Metalink Note :

Subject: **Oracle Clusterware consolidated logging in 10gR2** Doc ID: **Note:331168.1**

Extract from the note :

Oracle Clusterware consolidated logging in 10gR2

- CRS logs are in $ORA_CRS_HOME/log/<hostname>/crsd/
- CSS logs are in $ORA_CRS_HOME/log/<hostname>/cssd/
- EVM logs are in $ORA_CRS_HOME/log/<hostname>/evmd $ORA_CRS_HOME/evm/log/
- Resource specific logs are in $ORA_CRS_HOME/log/<hostname>/racg and the $ORACLE_HOME/log/<hostname>/racg
- RVM logs are in $ORA_CRS_HOME/log/<hostname>/client and the $ORACLE_HOME/log/<hostname>/client
- Cluster Network Communication logs are in the $ORA_CRS_HOME/log directory
- OPMN logs are in the $ORA_CRS_HOME/opmn/logs
- New in 10g Release 2 is a alert<nodename>.log present in the $ORA_CRS_HOME/log/<hostname>

Look at following Metalink Note to get all logs necessary for needed support diagnostics :

Subject: **CRS 10g R2 Diagnostic Collection Guide** **Doc ID**: **Note:330358.1**

14.5 CLUSTERWARE BASIC TESTING

Prior to go further, you should test how the clusterware is behaving at simple tests to validate your oracle clusterware installation.

Action to be done :	What should happen !!!!

(1) Node 1 and Node 2 are UP and Running	Before reboot of node1 :
	
	`{node2:oracle}/oracle ->crs_stat -t`
	`Name Type Target State Host`
	`------------------------------------------------------------`
	`ora.node1.gsd application ONLINE ONLINE node1`
	`ora.node1.ons application ONLINE ONLINE node1`
	`ora.node1.vip application ONLINE ONLINE node1`
	`ora.node2.gsd application ONLINE ONLINE node2`
	`ora.node2.ons application ONLINE ONLINE node2`
	`ora.node2.vip application ONLINE ONLINE node2`
	`{node2:oracle}/oracle ->`

➔ Reboot node1	While node1 is rebooting :
	Check on node2, and VIP from node1 should appear while node1 is out of order.
	
	`{node2:oracle}/oracle ->crs_stat -t`
	`Name Type Target State Host`
	`------------------------------------------------------------`
	`ora.node1.gsd application ONLINE OFFLINE`
	`ora.node1.ons application ONLINE OFFLINE`
	`ora.node1.vip application ONLINE ONLINE node2`
	`ora.node2.gsd application ONLINE ONLINE node2`
	`ora.node2.ons application ONLINE ONLINE node2`
	`ora.node2.vip application ONLINE ONLINE node2`
	`{node2:oracle}/oracle ->`
	
	When node1 is back with CRS up and running, the VIP will come back to its original position on node1.
	
	After reboot of node1 :
	
	`{node2:oracle}/oracle ->crs_stat -t`
	`Name Type Target State Host`
	`------------------------------------------------------------`
	`ora.node1.gsd application ONLINE ONLINE node1`
	`ora.node1.ons application ONLINE ONLINE node1`
	`ora.node1.vip application ONLINE ONLINE node1`
	`ora.node2.gsd application ONLINE ONLINE node2`
	`ora.node2.ons application ONLINE ONLINE node2`
	`ora.node2.vip application ONLINE ONLINE node2`
	`{node2:oracle}/oracle ->`

(2) Node 1 and Node 2 are UP and Running **➔ Reboot node2**	Check on node1, and VIP from node2 should appear while node2 is out of order. When node2 is back with CRS up and running, the VIP will come back to its original position on node2.
(2) Node 1 and Node 2 are UP and Running **➔ Reboot node1 and node2**	After reboot, both nodes will come back with CRS up and running, with VIP from both nodes on their respective positions, VIP (node1) on node1 and VIP (node2) on node2.

14.6 What Has Been Done ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**

14.7 VIP and CRS TroubleShouting

If problems occurs with VIP configuration assistant, please use the metalink notes specified in this chapter.

Metalink Note 296856.1- Configuring the IBM AIX 5L Operating System for the Oracle 10g VIP

Metalink Note 294336.1- Changing the check interval for the Oracle 10g VIP

Metalink Note 276434.1- Modifying the VIP of a Cluster Node

```
srvctl modify nodeapps -n <node_name> [-o <oracle_home>] [-A <new_vip_address>]

Options Description:
-n <node name> Node name.
-o <oracle home> Oracle home for the cluster database.
-A <new vip address> The node level VIP address (<name|ip>/netmask[/if1[|if2|...]]).

An example of the 'modify nodeapps' command is as follows:

$ srvctl stop nodeapps -n node1
$ srvctl modify nodeapps -n node1 -A 10.3.25.181/255.255.255.0/en0
$ srvctl start nodeapps -n node1

Note:   This command should be run as root.
```

Metalink Note 298895.1- Modifying the default gateway address used by the Oracle 10g VIP

Metalink Note 264847.1- How to Configure Virtual IPs for 10g RAC

How to delete VIP IP allias on public network card, if they are persistents even after the CRS shutdown :

Example for our case :

On node1 as root ➔ **Ifconfig en0 delete 10.3.25.181**
On node2 as root ➔ **Ifconfig en0 delete 10.3.25.182**

14.8 How to clean a failed CRS installation

Metalink Note 239998.1 - 10g RAC: How to Clean Up After a Failed CRS Install

On both Nodes :

As root user on each node :

$ORA_CRS_HOME/bin/crsctl stop crs (to clean any remaining crs deamons)
rmitab h1 ➔ this will remove oracle CRS entry in the /etc/inittab
rmitab h2
rmitab h3
rm –Rf /opt/ORCL*
kill remaining process from output : ps-ef|grep crs and ps-ef|grep d.bin
rm –R /etc/oracle/*

You should now remove the CRS installation, 2 options :

1/ **You want to keep the oraInventory** as its used for other oracle products which are installed and used, THEN run runInstaller as oracle user to uninstall the CRS installation. When done remove the content of the CRS directory on both nodes : **rm –Rf /oracle/crs/***

OR

2/ **You don't care about the oraInventory, and there's no other oracle products installed** on the nodes, THEN remove the full content of the $ORACLE_BASE including the oraInventory directory : **rm –Rf /oracle/***

Change owner, group and permission for /dev/ocr_disk and /dev/vote_disk on each node of the cluster :

On node1 …

```
node1:root-/> chown oracle.dba /dev/ocr_disk*
node1:root-/> chown oracle.dba /dev/voting_disk*
node1:root-/> chmod 660 /dev/ocr_disk*
node1:root-/> chmod 660 /dev/voting_disk*
```

On node2 …

```
node1:root-/> rsh node2
node2:root-/> chown oracle.dba /dev/ocr_disk1*
node2:root-/> chown oracle.dba /dev/voting_disk*
node2:root-/> chmod 660 /dev/ocr_disk*
node2:root-/> chmod 660 /dev/voting_disk*
```

To erase ALL OCR and Voting diks content ➔ Format (Zeroing) on the disks from one node :

```
for i in 1 2
do
        dd if=/dev/zero of=/dev/ocr_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
……..
```

```
for i in 1 2 3
do
        dd if=/dev/zero of=/dev/voting_disk$i bs=8192 count=25000 &
done
25000+0 records in.
25000+0 records out.
     ……..
```

14.9 10.2.0.3 CRS PATCHSET UPDATE

On each node :

Stop the nodeapps (VIP, GSD, ONS, …)

Stop all node applications on all nodes. To stop node applications running on a node, enter the following command where *node* is the name of the node where the applications are running:

As oracle user, execute :

$ *oracle_home*/bin/srvctl stop nodeapps -n *node*

```
{node1:root}/oracle -> srvctl stop nodeapps –n node1
{node1:root}/oracle ->
{node1:root}/oracle -> srvctl stop nodeapps –n node2
{node1:root}/oracle ->
```

On each node :

Shut down the Oracle Clusterware process by entering the following command on all nodes as the root user:

As root user, execute :

$ crs_home/bin/crsctl stop crs

```
{node1:root}/oracle -> crsctl stop crs
Stopping resources.
Successfully stopped CRS resources
Stopping CSSD.
Shutting down CSS daemon.
Shutdown request successfully issued.
{node1:root}/oracle ->

{node2:root}/oracle -> crsctl stop crs
Stopping resources.
Successfully stopped CRS resources
Stopping CSSD.
Shutting down CSS daemon.
Shutdown request successfully issued.
{node2:root}/oracle ->
```

On each node :

Then running “crs_stat –t” as root on node1 should return the following result :

```
{node1:root}/oracle -> crs_stat -t
CRS-0184: Cannot communicate with the CRS daemon.
```

Back Up the System, Oracle recommends that you create a backup of the Oracle Clusterware 10*g* installation before you install the patch set, and a backup of the OCR content.

On each node :

As root, execute preupdate.sh script from $ORA_CRS_HOME/install

```
# This is run once per node when upgrading or patching CRS homes.
# This script will stop the CRS stack running on the node.
# The script will also change ownership of CRS home and root owned files  to Oracle CRS user.
#
# USAGE: preupdate.sh -crshome <Oracle CRS home location> -crsuser <CRS user name> [-noshutdown]

{node1:root}/oracle/products/crs/install -> ./preupdate.sh -crshome
/oracle/products/crs -crsuser oracle
Shutting down Oracle Cluster Ready Services (CRS):
Stopping resources.
Error while stopping resources. Possible cause: CRSD is down.
Stopping CSSD.
Unable to communicate with the CSS daemon.
Shutdown has begun. The daemons should exit soon.
Checking to see if Oracle CRS stack is down...
Oracle CRS stack is down now.
{node1:root}/oracle/products/crs/install ->

{node2:root}/oracle/products/crs/install -> ./preupdate.sh -crshome
/oracle/products/crs -crsuser oracle
...
```

On each node :

As root on all nodes :

ln -s /usr/sbin/sync /usr/bin/sync

```
{node1:root}/oracle/products -> ln -s /usr/sbin/sync /usr/bin/sync
{node1:root}/oracle/products ->

{node2:root}/oracle/products -> ln -s /usr/sbin/sync /usr/bin/sync
{node2:root}/oracle/products ->
```

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node1:root}/oracle/products ->

{node2:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products ->
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

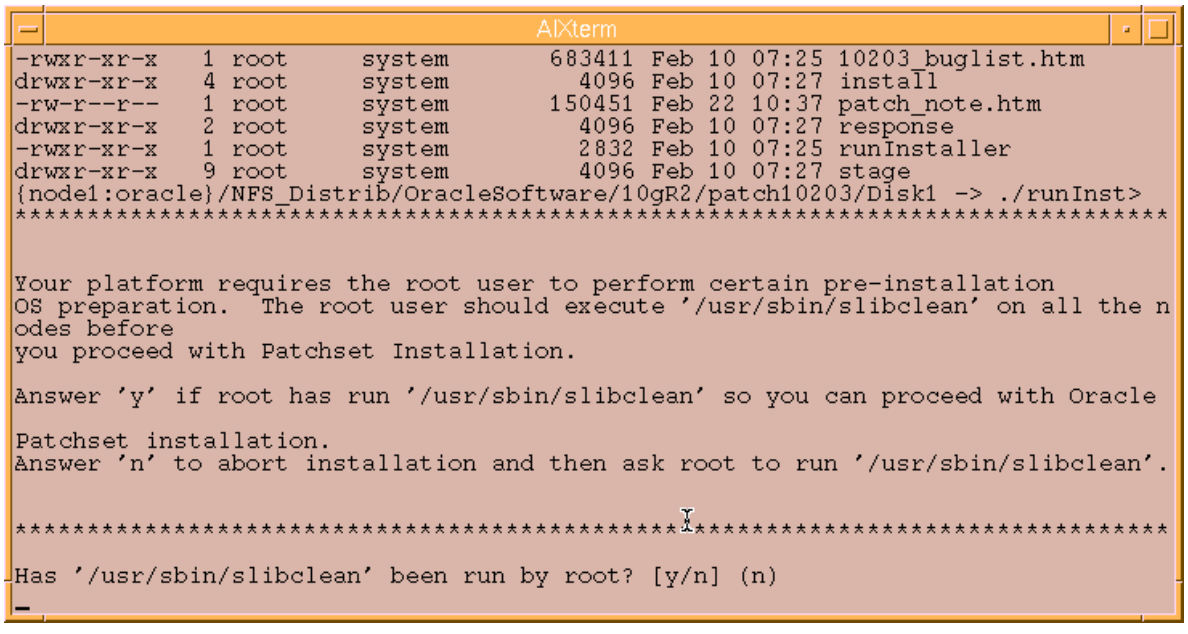
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

Or execute : ./runInstaller -ignoreSysPrereqs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

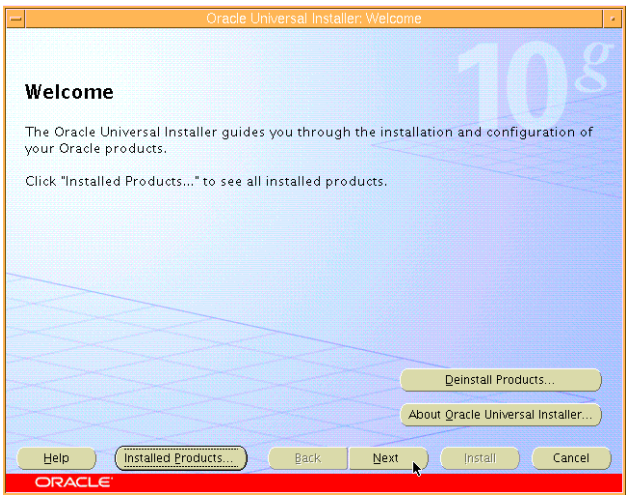
To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
All filesets for 5300-04_AIX_ML were found.
All filesets for 5300-05_AIX_ML were found.



At the OUI Welcome screen

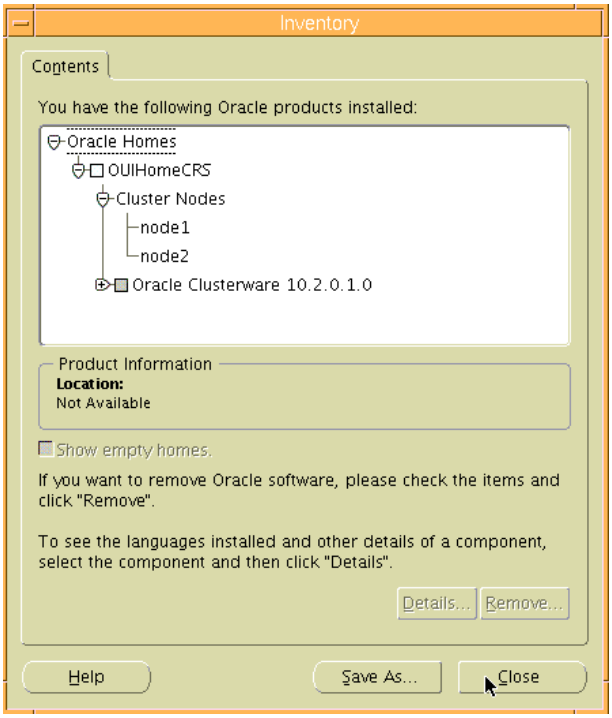
Enter "Yes" if /usr/sbin/slibclean has been executed on both node as root.

Just click Next ...



At the OUI Welcome screen

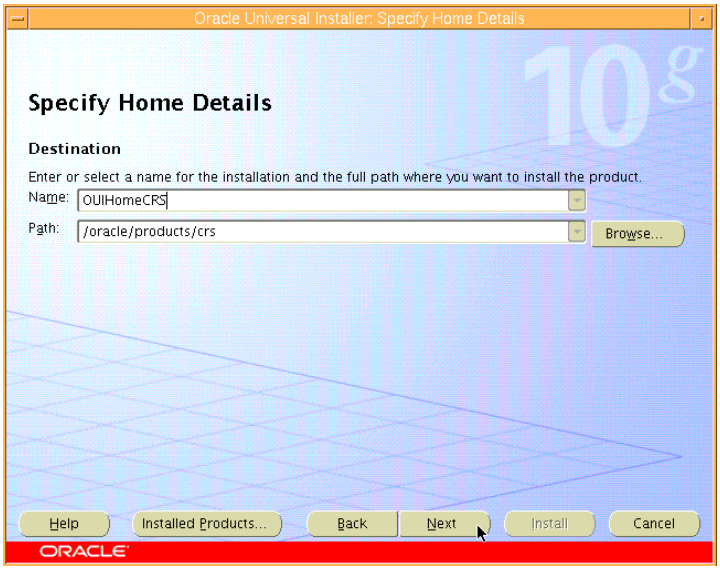
Just click Next ...



At the OUI Welcome screen

You can check what is already installed by selecting “Installed Products”

Just click Close and Next ...

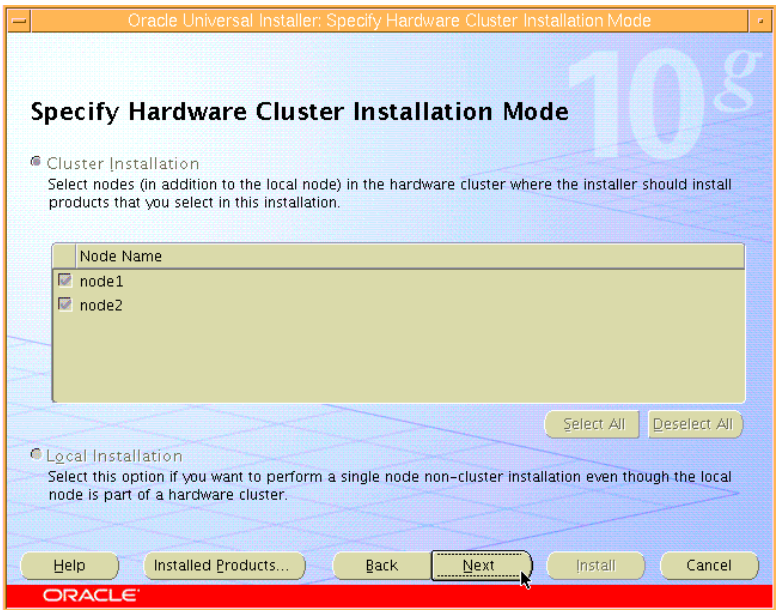


Select the ORACLE_HOME corresponding to the CRS_HOME Installation directory !!!

For our case :

CRS_Home
/oracle/products/crs

Just click Next ...

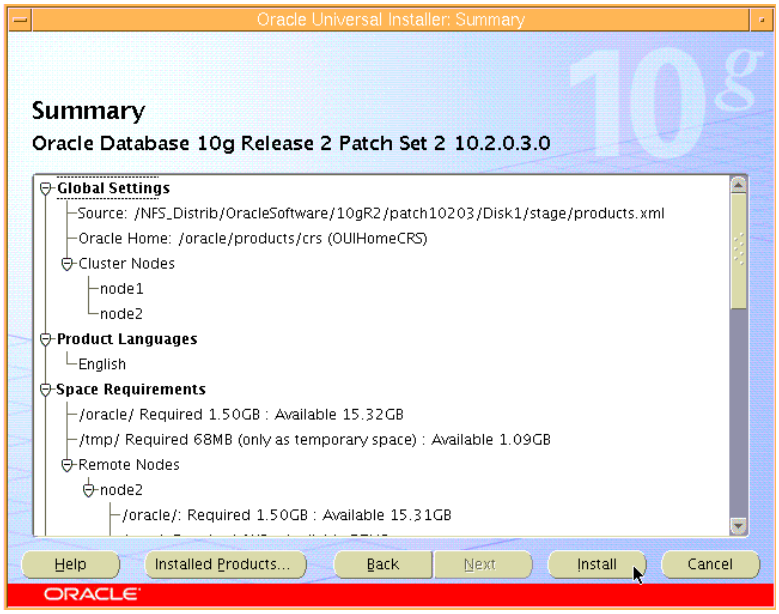


At this stage :

No choice available for local Installation.

Cluster Installation with all nodes selected by default.

Just click Next ...

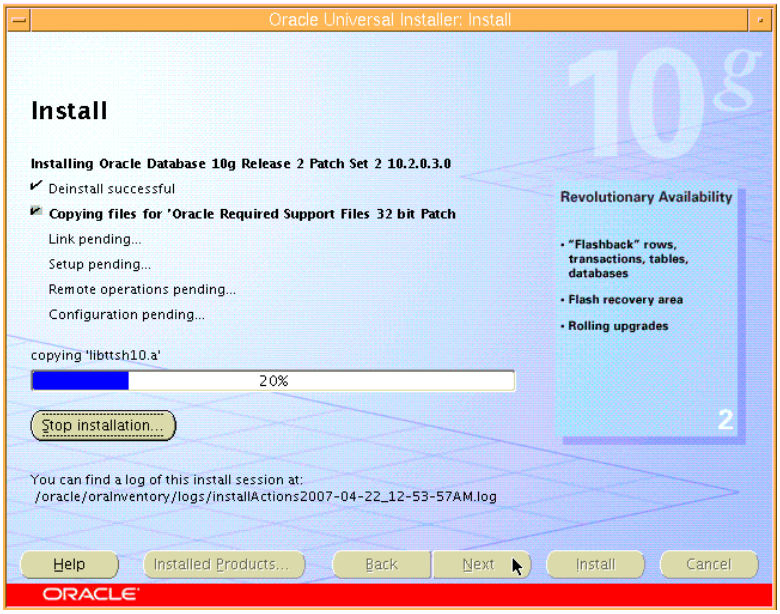


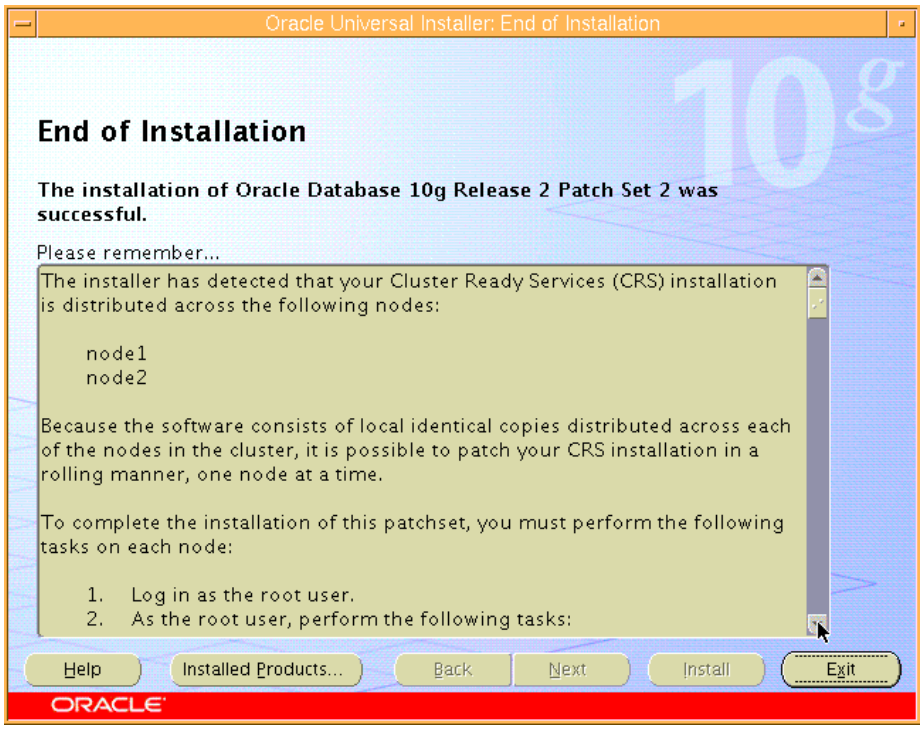
Summary :

Check for

- Oracle Home Path (CRS_HOME)
- Cluster nodes list
- Remote Nodes list

Just click Install, installation will process until next screen ...





End of Installation :

Scroll down and see below ...

The installer has detected that your Cluster Ready Services (CRS) installation is distributed across the following nodes:

node1
node2

Because the software consists of local identical copies distributed across each of the nodes in the cluster, it is possible to patch your CRS installation in a rolling manner, one node at a time.

To complete the installation of this patchset, you must perform the following tasks on each node:

1. Log in as the root user.
2. As the root user, perform the following tasks:

a. Shutdown the CRS daemons by issuing the following command:
/oracle/products/crs/bin/crsctl stop crs

b. Run the shell script located at:
/oracle/products/crs/install/root102.sh
This script will automatically start the CRS daemons on the patched node upon completion.

3. After completing this procedure, proceed to the next node and repeat.

Execute Configuration Scripts :

AS root :

Execute root102.sh on each node, one after the other, as showed on next page …

This script is located in the ORA_CRS_HOME/install directory

For our case, this script is located in the CRS_Home /oracle/products/crs

Just click Exit ...

```
node1@root: /oracle/products/crs/install # ./root102.sh
Creating pre-patch directory for saving pre-patch clusterware files
Completed patching clusterware files to /oracle/products/crs
Relinking some shared libraries.
ld: 0711-783 WARNING: TOC overflow. TOC size: 67184     Maximum size: 65536
        Extra instructions are being generated for each reference to a TOC
        symbol if the symbol is in the TOC overflow area.
ld: 0711-783 WARNING: TOC overflow. TOC size: 67184     Maximum size: 65536
        Extra instructions are being generated for each reference to a TOC
        symbol if the symbol is in the TOC overflow area.
Relinking of patched files is complete.
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
Preparing to recopy patched init and RC scripts.
Recopying init and RC scripts.
Startup will be queued to init within 30 seconds.
Starting up the CRS daemons.
Adding daemons to inittab
Waiting for the patched CRS daemons to start.
  This may take a while on some systems.
.
10203 patch successfully applied.
clscfg: EXISTING configuration version 3 detected.
clscfg: version 3 is 10G Release 2.
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 1: node1 node1-rac node1
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
clscfg -upgrade completed successfully
node1@root: /oracle/products/crs/install #
```

First, execute root102.sh as root on node1

Just forget about the message : **"Id: 0711-783 WARNING: TOC overflow."**

When done on node 1, run the script on node2. Not to be done at same time…

```
node2@root: /oracle/products/crs/install # ./root102.sh
Creating pre-patch directory for saving pre-patch clusterware files
Completed patching clusterware files to /oracle/products/crs
Relinking some shared libraries.
ld: 0711-783 WARNING: TOC overflow. TOC size: 67184     Maximum size: 65536
        Extra instructions are being generated for each reference to a TOC
        symbol if the symbol is in the TOC overflow area.
ld: 0711-783 WARNING: TOC overflow. TOC size: 67184     Maximum size: 65536
        Extra instructions are being generated for each reference to a TOC
        symbol if the symbol is in the TOC overflow area.
Relinking of patched files is complete.
WARNING: directory '/oracle/products' is not owned by root
WARNING: directory '/oracle' is not owned by root
Preparing to recopy patched init and RC scripts.
Recopying init and RC scripts.
Startup will be queued to init within 30 seconds.
Starting up the CRS daemons.
Adding daemons to inittab
Waiting for the patched CRS daemons to start.
  This may take a while on some systems.
.
10203 patch successfully applied.
clscfg: EXISTING configuration version 3 detected.
clscfg: version 3 is 10G Release 2.
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 2: node2 node2-rac node2
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
clscfg -upgrade completed successfully
node2@root: /oracle/products/crs/install #
```

THEN execute root102.sh as root on node2

Just forget about the message : **"Id: 0711-783 WARNING: TOC overflow."**

When done on node 2.

THEN click EXIT on previous screen.

CRS software version query

```
{node1:oracle}/oracle/products/crs/bin ->crsctl query crs activeversion
CRS software version on node [node1] is [10.2.0.3.0]

{node2:oracle}/oracle/products/crs/bin ->crsctl query crs activeversion
CRS software version on node [node2] is [10.2.0.3.0]
{node2:oracle}/oracle/products/crs/bin ->
```

Apply Metalink Note 413077.1 !!!!! for usrvip script …
Subject: **%ORA_CRS_HOME% NOT REPLACED IN IN USRVIP SCRIPT INSTALLING 10.2.0.3** Doc ID: **Note:413077.1**

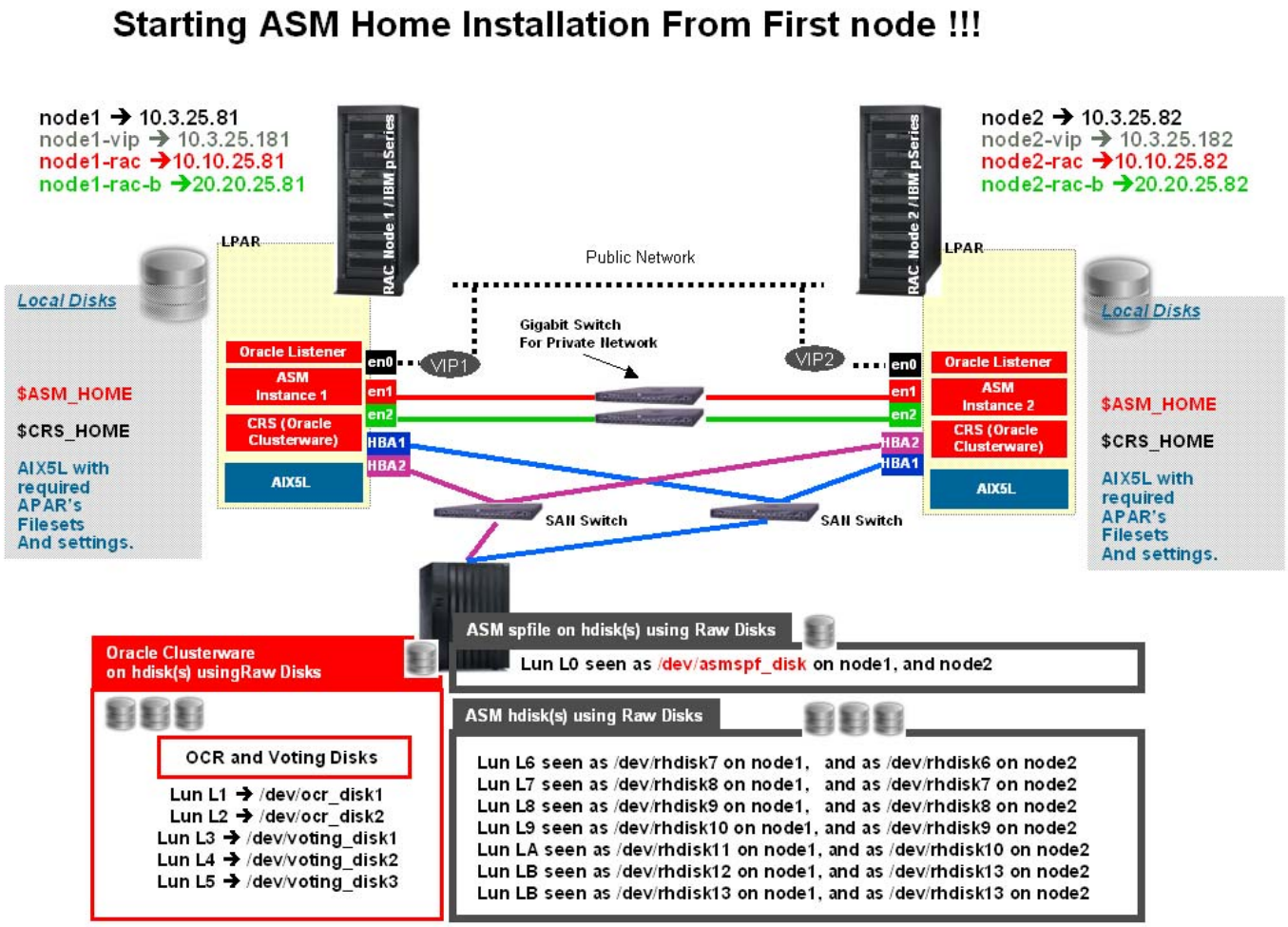
Apply Patch 5679560 **:**
Subject: **Known RAC/CRS issues After Applying 10.2.0.3 Patchset** Doc ID: **Note:406965.1**

14.10 WHAT HAS BEEN DONE ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**
- **10.2.0.3 patchset is applied for CRS home**

15 IMPLEMENTING ASM



Subject: **New Feature on ASM (Automatic Storage Manager).** Doc ID: **Note:249992.1**

15.1 ASM Installation and ASM instance creation

Oracle ASM installation just have to be done only starting from one node. Once the first node is installed, Oracle OUI automatically starts the copy of the mandatory files on the second node, using **rcp** command. This step could last long, depending on the network speed (one hour…), without any message. So, don't think the OUI is stalled, and look at the network traffic before canceling the installation !
You can also create a staging area. The name of the subdirectories is in the format "Disk1" to "Disk3"

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products -> /usr/sbin/slibclean
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

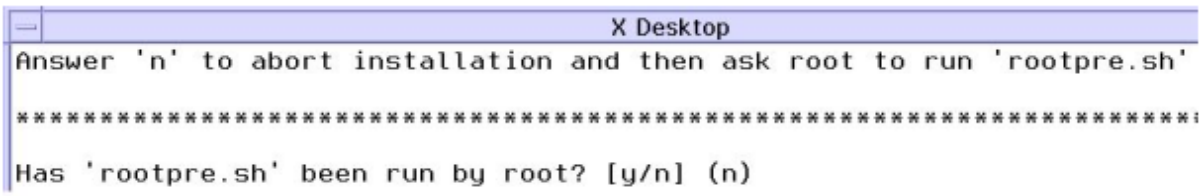
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

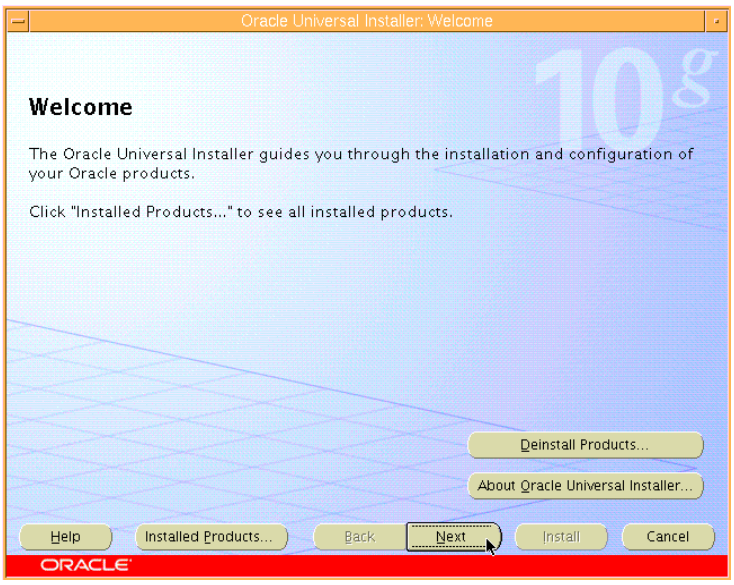
Or execute : ./runInstaller -ignoreSysPreregs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
…

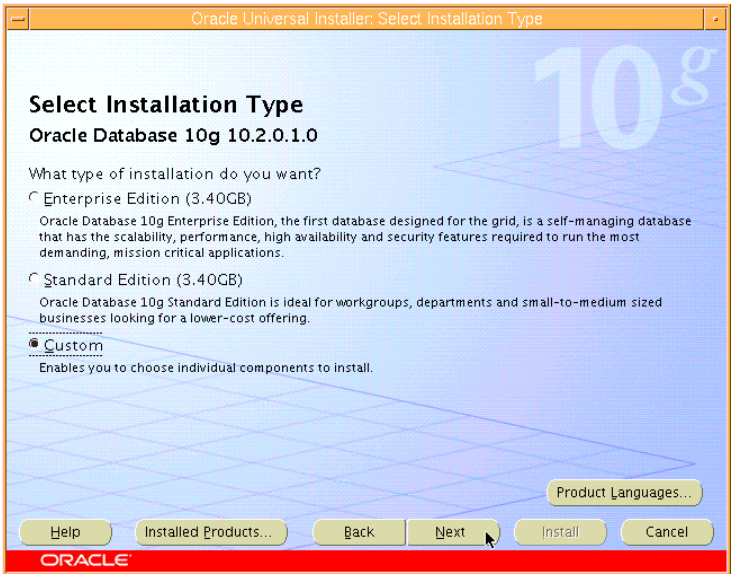


Should be already done with the CRS Installation), Answer "y" for next step



At the OUI Welcome screen

Just click Next ...

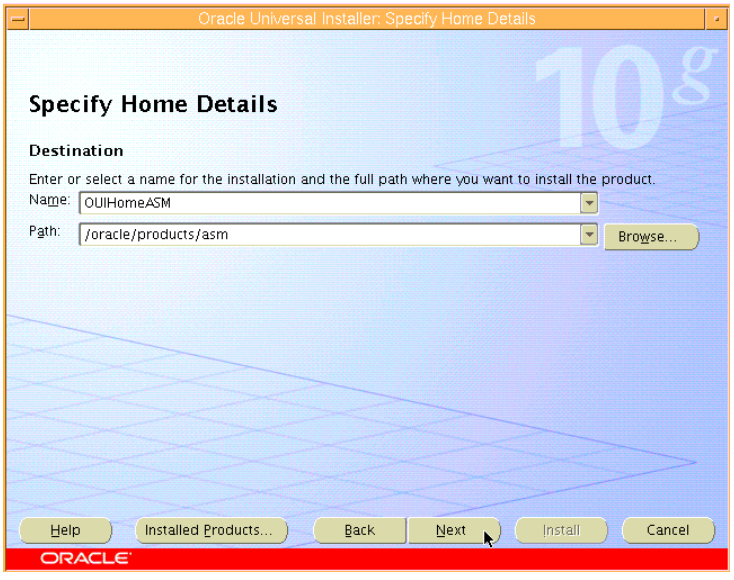


Select the installation type :

You have the option to choose Enterprise, Standard Edition, or Custom to proceed.

Choose the "Custom" option to avoid creating a database by default.

Then click Next ...



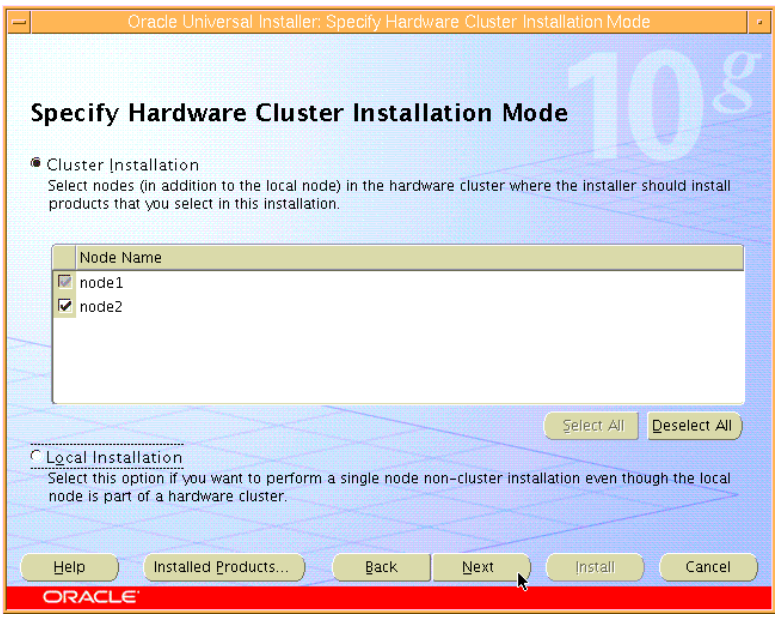
Specify File Locations :
Do not change the Source field

Specify a different ORACLE_HOME Name with its own directory for the Oracle software installation.

This ORACLE_HOME must be different then the CRS ORACLE_HOME.

Then click Next ...

If you don't see the following screen with Node selection, it might be that your CRS is down on one or all nodes. ➔ Please check if CRS is up and running on all nodes.

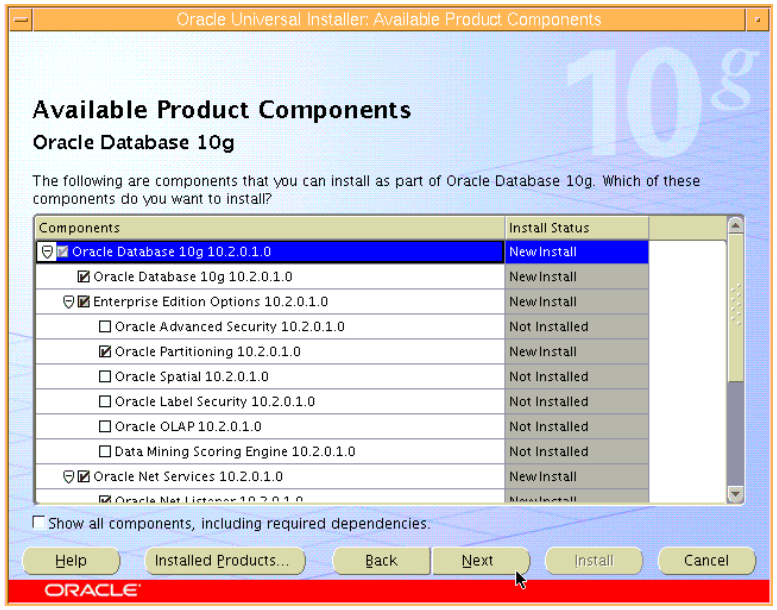


Specify Hardware Cluster Installation Mode :

Select **Cluster Installation**

AND the other nodes on to which the Oracle RDBMS software will be installed. It is not necessary to select the node on which the OUI is currently running. Click Next.

Then click Next ...

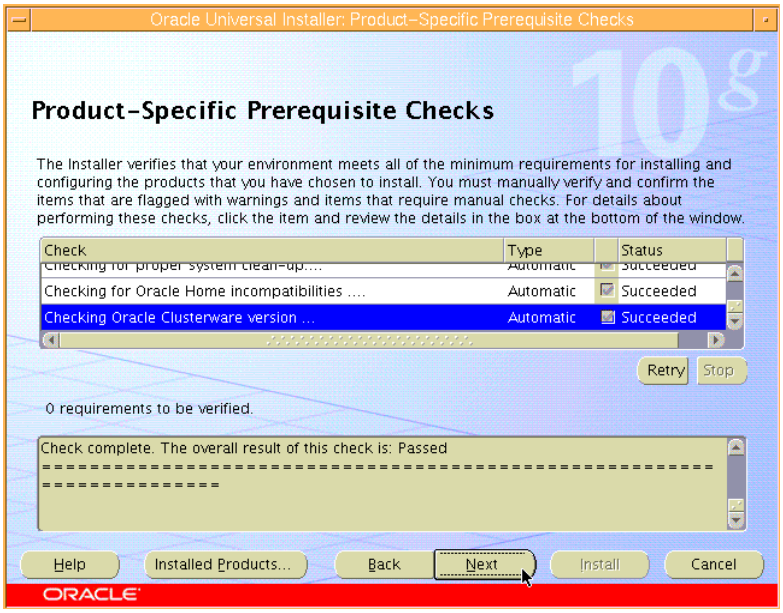


Available Product Components :

Select the product components for Oracle Database 10g that you want to install.

INFO : Compared to 10gRAC R1 installation, there is no “Real Application Cluster” option to select.

Then click Next ...



The installer will check some product-specific Prerequisite.

Don’t take care of the lines with checking at status “Not executed”, These are just warnings because AIX maintenance level might be higher then 5300, which is the case in our example (ML03).

Then click Next ...

Details of the prerequisite checks done by runInstaller

Checking operating system requirements ...
Expected result: One of 5200.004,5300.002
Actual Result: 5300.002
Check complete. The overall result of this check is: Passed
==========

Checking operating system package requirements ...
Checking for bos.adt.base(0.0); found bos.adt.base(5.3.0.51). Passed
Checking for bos.adt.lib(0.0); found bos.adt.lib(5.3.0.50). Passed
Checking for bos.adt.libm(0.0); found bos.adt.libm(5.3.0.40). Passed
Checking for bos.perf.libperfstat(0.0); found bos.perf.libperfstat(5.3.0.50). Passed
Checking for bos.perf.perfstat(0.0); found bos.perf.perfstat(5.3.0.50). Passed
Checking for bos.perf.proctools(0.0); found bos.perf.proctools(5.3.0.50). Passed
Check complete. The overall result of this check is: Passed
==========

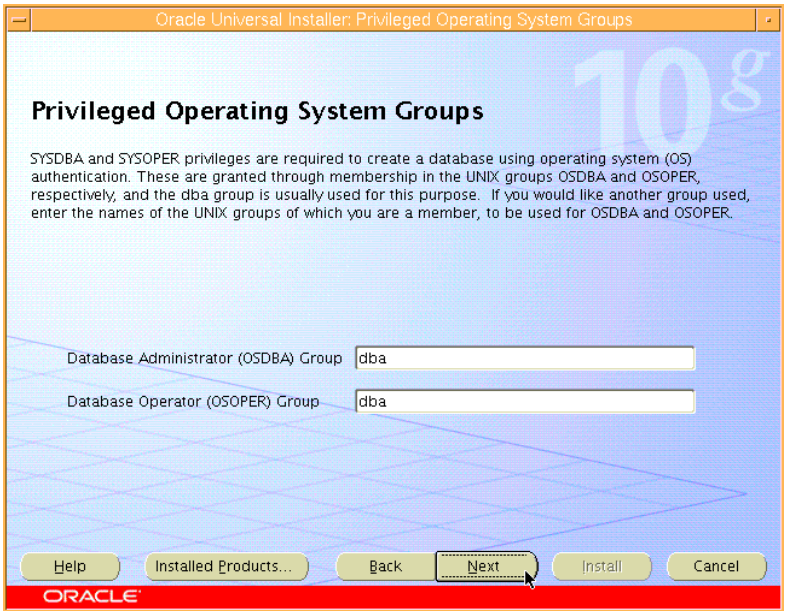
Checking recommended operating system patches
Checking for IY59386(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY60930(bos.mp,5.3.0.1); found (bos.mp,5.3.0.54). Passed
Checking for IY60930(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp64,5.3.0.20); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp,5.3.0.20); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp,5.3.0.22); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp64,5.3.0.22); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.acct,5.3.0.1); found (bos.acct,5.3.0.51). Passed
Checking for IY58143(bos.adt.include,5.3.0.1); found (bos.adt.include,5.3.0.53). Passed
Checking for IY58143(bos.adt.libm,5.3.0.1); found (bos.adt.libm,5.3.0.40). Passed
Checking for IY58143(bos.adt.prof,5.3.0.1); found (bos.adt.prof,5.3.0.53). Passed
Checking for IY58143(bos.alt_disk_install.rte,5.3.0.1); found (bos.alt_disk_install.rte,5.3.0.51). Passed
Checking for IY58143(bos.cifs_fs.rte,5.3.0.1); found (bos.cifs_fs.rte,5.3.0.50). Passed
Checking for IY58143(bos.diag.com,5.3.0.1); found (bos.diag.com,5.3.0.51). Passed
Checking for IY58143(bos.perf.libperfstat,5.3.0.1); found (bos.perf.libperfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.perfstat,5.3.0.1); found (bos.perf.perfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.tools,5.3.0.1); found (bos.perf.tools,5.3.0.52). Passed
Checking for IY58143(bos.rte.boot,5.3.0.1); found (bos.rte.boot,5.3.0.51). Passed
Checking for IY58143(bos.rte.archive,5.3.0.1); found (bos.rte.archive,5.3.0.51). Passed
Checking for IY58143(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY58143(bos.rte.control,5.3.0.1); found (bos.rte.control,5.3.0.50). Passed
Checking for IY58143(bos.rte.filesystem,5.3.0.1); found (bos.rte.filesystem,5.3.0.51). Passed
Checking for IY58143(bos.rte.install,5.3.0.1); found (bos.rte.install,5.3.0.54). Passed
Checking for IY58143(bos.rte.libc,5.3.0.1); found (bos.rte.libc,5.3.0.53). Passed
Checking for IY58143(bos.rte.lvm,5.3.0.1); found (bos.rte.lvm,5.3.0.53). Passed
Checking for IY58143(bos.rte.man,5.3.0.1); found (bos.rte.man,5.3.0.50). Passed
Checking for IY58143(bos.rte.methods,5.3.0.1); found (bos.rte.methods,5.3.0.51). Passed
Checking for IY58143(bos.rte.security,5.3.0.1); found (bos.rte.security,5.3.0.53). Passed
Checking for IY58143(bos.rte.serv_aid,5.3.0.1); found (bos.rte.serv_aid,5.3.0.52). Passed
Check complete. The overall result of this check is: Passed
==========

Validating ORACLE_BASE location (if set) ...
Check complete. The overall result of this check is: Passed
==========

Checking for proper system clean-up....
Check complete. The overall result of this check is: Passed
==========

Checking for Oracle Home incompatibilities ....
Actual Result: NEW_HOME
Check complete. The overall result of this check is: Passed
==========

Checking Oracle Clusterware version ...
Check complete. The overall result of this check is: Passed
==========

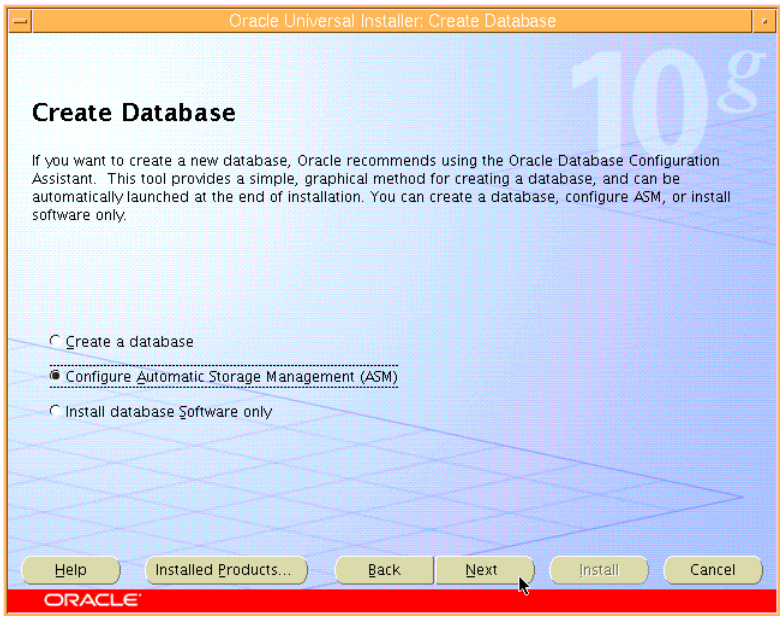


Privileged Operating Systems Groups :

Verify the UNIX primary group name of the user which controls the installation of the Oracle10g software.
(Use unix command id to find out)

And specify the Privileged Operating System Groups to the value found.
In our example, this must be “dba” (Primary group of unix oracle user) to be set for both entries.

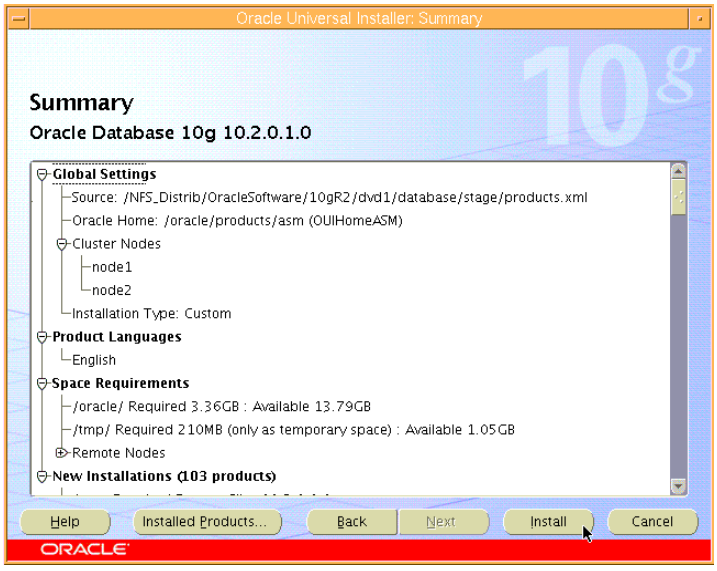
Then click Next ...



Create Database :

Choose “Configure Automatic Storage Management (ASM)”, we want to install the ASM software in its own ORACLE_HOME at this stage.

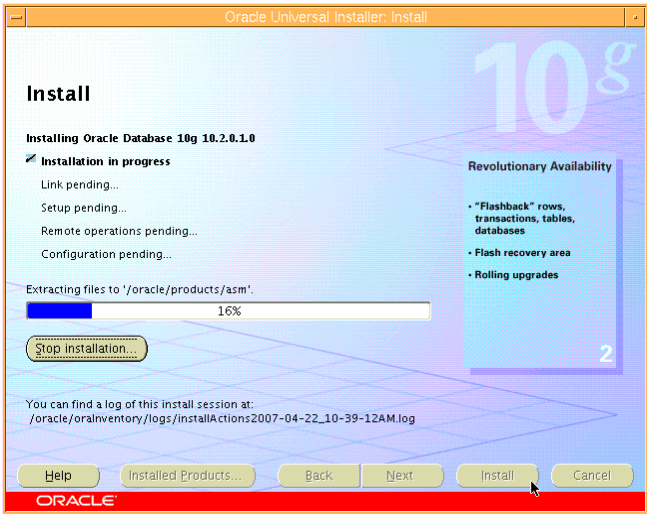
Then click Next ...



Summary :

The Summary screen will be presented. Confirm that the RAC database software and other selected options will be installed.
Check Cluster Nodes and Remote Nodes lists.
The OUI will install the Oracle 10g software on to the local node, and then copy this information to the other selected nodes.

Then click Install ...



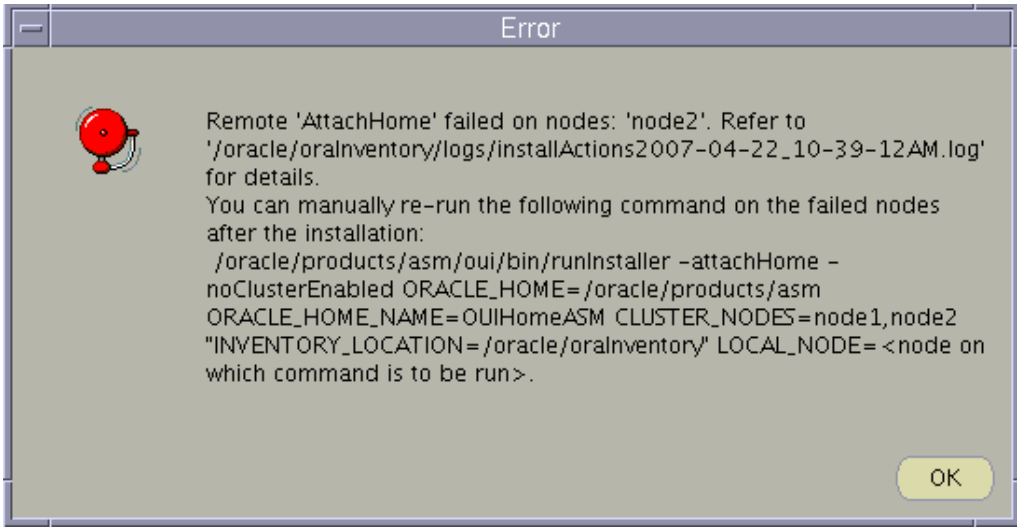
Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the others selected nodes.

Just wait for the next screen ...

This screen may appears, if so just run the specified command on the specified node as oracle user.

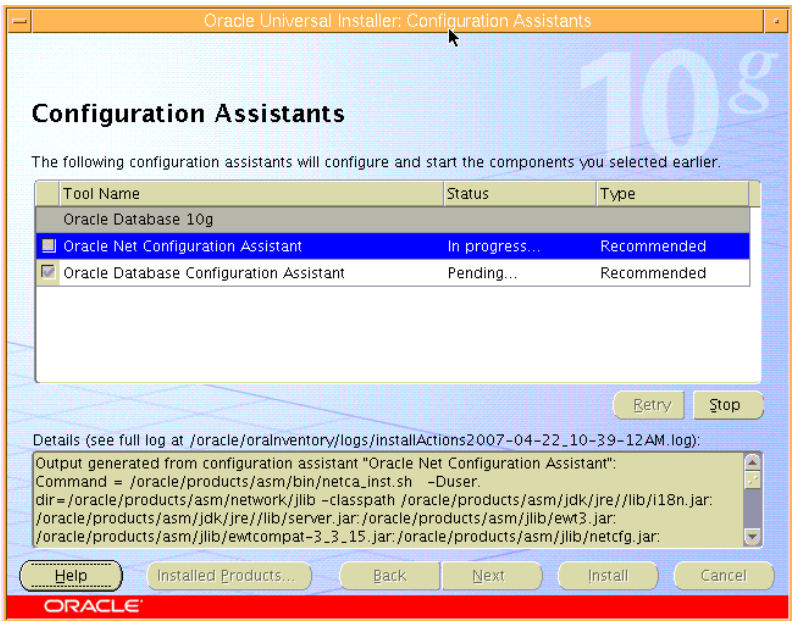
If you get this following error message ➔



Do execute the script bellow :

From node2 :

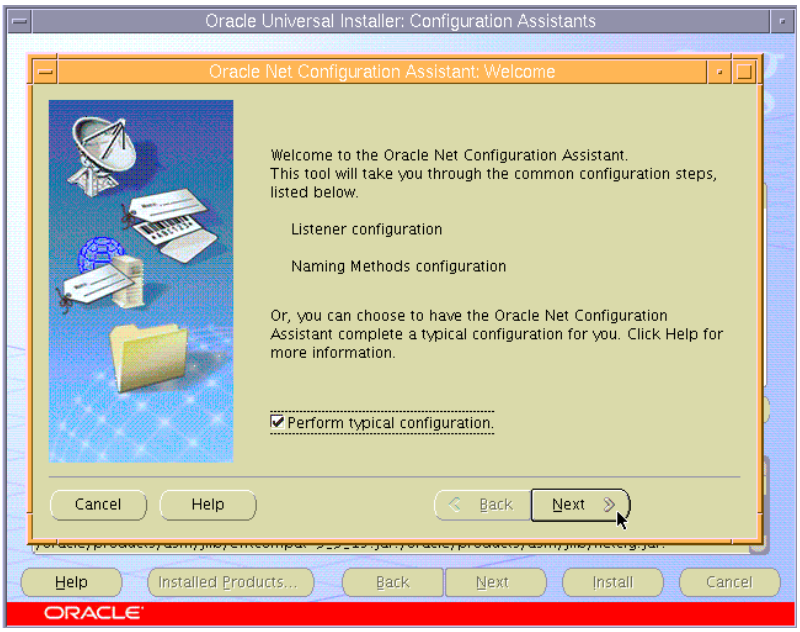
```
{node2:root}/oracle/products/rdbms -> su – oracle
{node2:oracle}/oracle -> /oracle/products/asm/oui/bin/runInstaller –attachHome –
noClusterEnabled ORACLE_HOME=/oracle/products/asm ORACLE_HOME_NAME=OUIHomeASM
CLUSTER_NODES=node1,node2 “INVENTORY_LOCATION=/oracle/oraInventory” LOCAL_NODE=node2
Starting Oracle Universal Installer…

No pre-requisite checks found in oraparam.ini, no system pre-requisite checks will be
executed.
The inventory pointer is located at /etc/oraInst.loc
The inventory is located at /oracle/oraInventory
‘AttachHome’ was successful.
{node2:oracle}/oracle ->
```



Configuration Assistants will be executed to :

- **Configure a listener in a cluster mode**
- **Configure and start the ASM instances, and create the ASM diskgroups.**

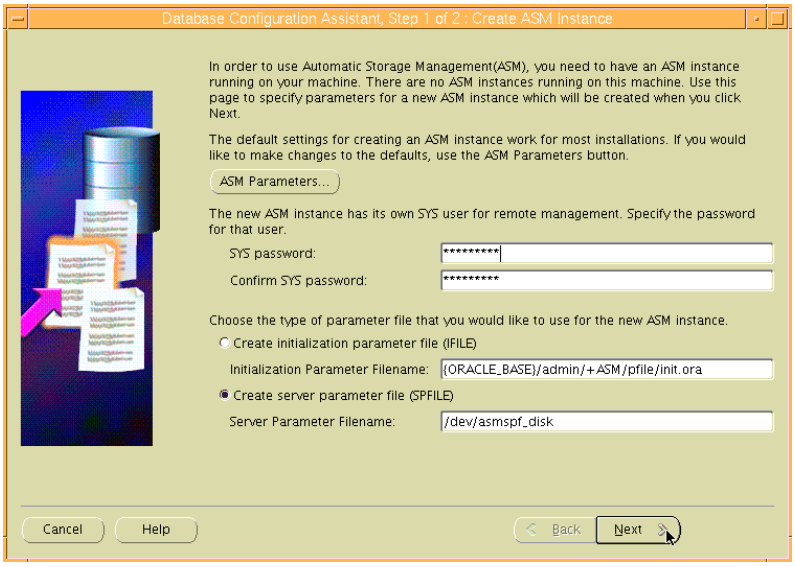
Just wait the few next screens …



Select "Perform Typical Configuration" For cluster listener configuration.

Listener port will be 1521 by default, it can be changed later thru netca.

Then click Next ...



Create ASM Instance :

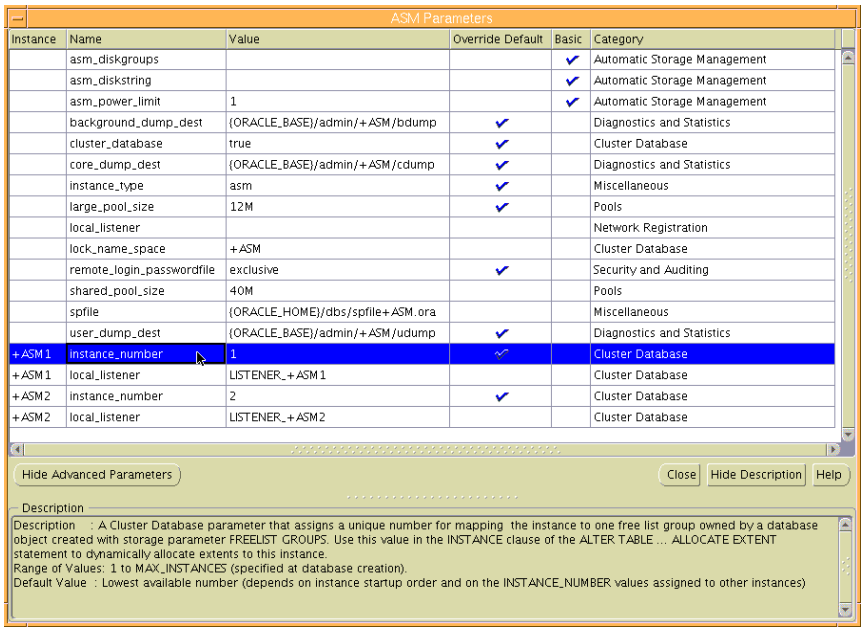
- **Specify SYS password for the ASM Instance**
- **Choose the type of parameter file "Create initialization parameter file (IFILE)"**

Click on "ASM Parameters" to see default parameters for ASM (See below).

Then click Next ...

Clicking on "ASM Parameters" will show default parameters for ASM, and

Click on "Show Advanced Parameters" to see screen as follow :



Details of the ASM parameters :

ASM_DISKGROUPS

Description: This value is the list of the Disk Group names to be mounted by the ASM at startup or when ALTER DISKGROUP ALL MOUNT command is used.

ASM_DISKSTRING

Description: A comma separated list of paths used by the ASM to limit the set of disks considered for discovery when a new disk is added to a Disk Group. The disk string should match the path of the disk, not the directory containing the disk. For example: /dev/rdsk/*.

ASM_POWER_LIMIT

Description: This value is the maximum power on the ASM instance for disk rebalancing.
Range of Values: 1 to 11
Default Value: 1

BACKGROUND_DUMP_DEST

Description : Specifies the pathname (directory or disc) where trace files are written for the back ground processes (LGWR, DBW n, and so on) during Oracle operations. It also defines the location of the database alert file which logs significant events and messages.
Range of Values: Any valid directory name.
Default Value : ORACLE_HOME/rdbms/log (operating system dependent)

CLUSTER_DATABASE

Description : Set CLUSTER_DATABASE to TRUE to enable Real Application Clusters option.
Range of Values: TRUE | FALSE
Default Value : FALSE

CORE_DUMP_DEST

Description : The directory name specifying the core dump location (for UNIX).
Range of Values: Any valid directory name.
Default Value : ORACLE_HOME/dbs

LARGE_POOL_SIZE

Description : Specifies the size of the large pool allocation heap, which is used by Shared Server for session memory, parallel execution for message buffers, and RMAN backup and recovery for disk I/O buffers.
Range of Values: 600K (minimum); >= 20000M (maximum is operating system specific).
Default Value : 0, unless parallel execution or DBWR_IO_SLAVES are configured

LARGE_POOL_SIZE

Description : Specifies the size of the large pool allocation heap, which is used by Shared Server for session memory, parallel execution for message buffers, and RMAN backup and recovery for disk I/O buffers.
Range of Values: 600K (minimum); >= 20000M (maximum is operating system specific).
Default Value : 0, unless parallel execution or DBWR_IO_SLAVES are configured

LOCAL_LISTENER

Description : A Oracle Net address list which identifies database instances on the same machine as the Oracle Net listeners. Each instance and dispatcher registers with the listener to enable client connections. This parameter overrides MTS_LISTENER_ADDRESS and MTS_MULTIPLE_LISTENERS parameters that were obsolete as 8.1.
Range of Values: A valid Oracle Net address list.
Default Value : (ADDRESS_LIST=(Address=(Protocol=TCP)(Host=localhost)(Port=1521)) (Address=(Protocol=IPC)(Key=DBname)))

LOCK_NAME_SPACE

Description : Specifies the name space that the distributed lock manager (DLM) uses to generate lock names. This might need to be set if there is a standby or clone database with the same database name on the same cluster.
Range of Values: Eight characters maximum, no special characters allowed.
Default Value : NULL

REMOTE_LOGIN_PASSWORDFILE

Description : Specifies whether passwords for privileged users are checked by the operating system or a file. With NONE, Oracle ignores the password file. With EXCLUSIVE, each privileged user is authenticated using the database's password file. With SHARED, many databases share the SYS and INTERNAL password file users.
Range of Values: NONE | SHARED | EXCLUSIVE
Default Value : NONE

SHARED_POOL_SIZE

Description : Specifies the size of the shared pool in bytes. The shared pool contains objects such as shared cursors, stored procedures, control structures, and Parallel Execution message buffers. Larger values can improve performance in multi-user systems.
Range of Values: 300 Kbytes - operating system dependent.
Default Value : If 64 bit, 64MB, else 16MB

SPFILE

Description: Specifies the name of the current server parameter file in use.
Range of Values: static parameter
Default Value: The SPFILE parameter can be defined in a client side PFILE to indicate the name of the server parameter file to use. When the default server parameter file is used by the server, the value of SPFILE will be internally set by the server.

USER_DUMP_DEST

Description : Specifies the pathname for a directory where the server will write debugging trace files on behalf of a user process. For example, this directory might be set to C:/ORACLE/UTRC on NT; to /oracle/utrc on UNIX; or to DISK$UR3:[ORACLE.UTRC] on VMS.
Range of Values: A valid local pathname, directory, or disk.
Default Value : Operating system dependent

INSTANCE_NUMBER

Description : A Cluster Database parameter that assigns a unique number for mapping the instance to one free list group owned by a database object created with storage parameter FREELIST GROUPS. Use this value in the INSTANCE clause of the ALTER TABLE ... ALLOCATE EXTENT statement to dynamically allocate extents to this instance.
Range of Values: 1 to MAX_INSTANCES (specified at database creation).
Default Value : Lowest available number (depends on instance startup order and on the INSTANCE_NUMBER values assigned to other instances)

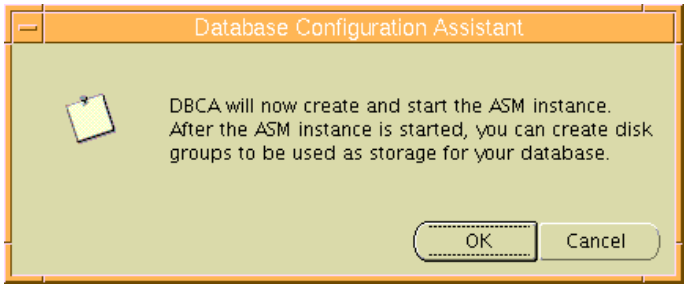
Please read following Metalink note :

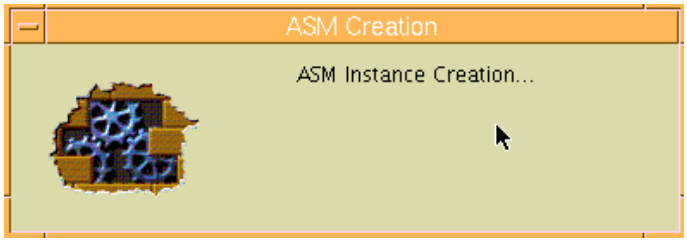
Subject: **SGA sizing for ASM instances and databases that use ASM** Doc ID: **Note:282777.1**

Create ASM Instance :

The assistant will create the ASM instance on each node.

Then click OK..

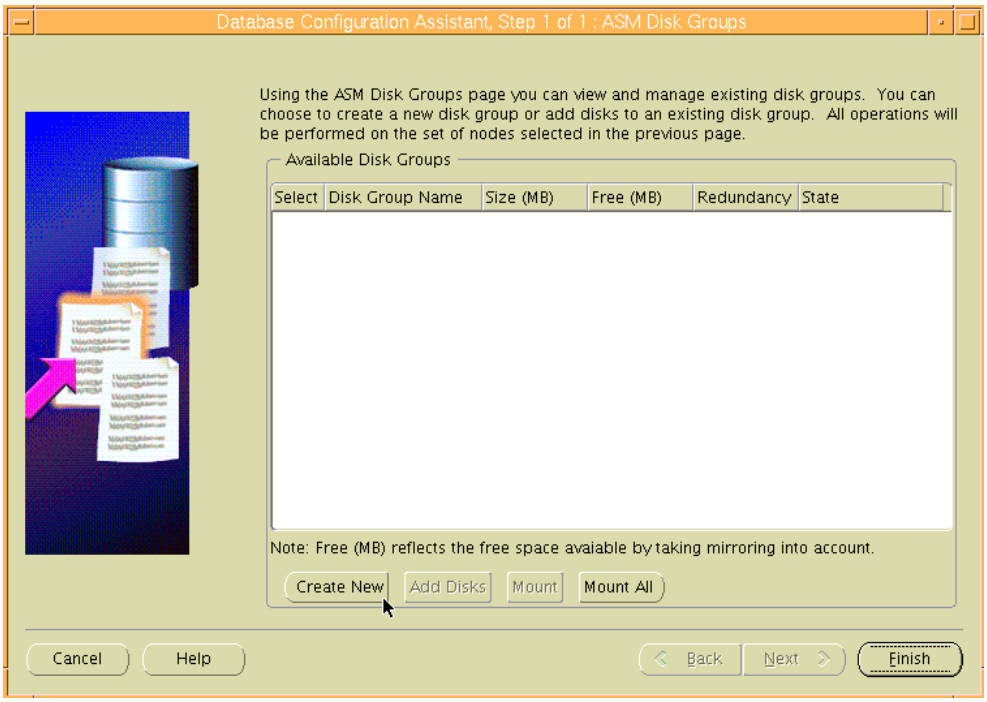




ASM Disk Groups :

Click on "Create New"

And see next screen ...



Create Disk Group :

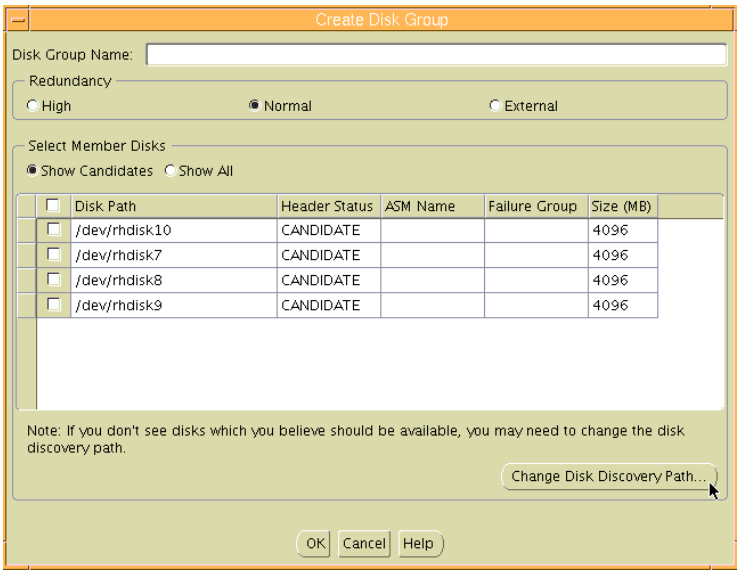
Select Show Candidates
Specify Disk Group Name :

Select Redundancy

- High (ASM mirroring, 3 copies)
- Normal (ASM mirroring, 2 copies)
- External (no ASM mirroring)

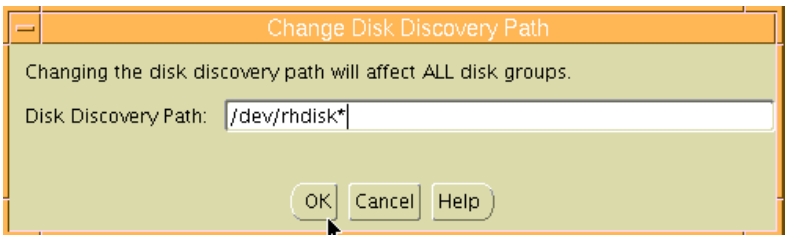
Then **Select Disks**

Then click OK to create ASM Diskgroup....

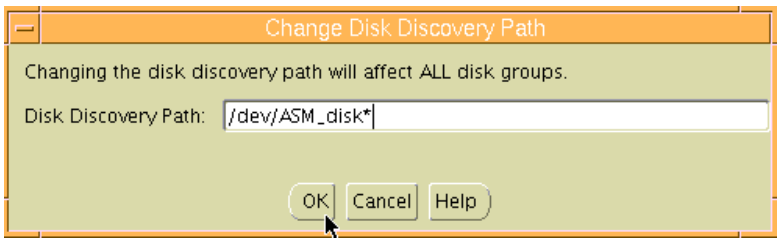


If you don't see any candidate disks, do click on "Change Disk Discovery Path", and change to the right path.

If you still don't see any candidates disks, do check the disks preparation.



OR



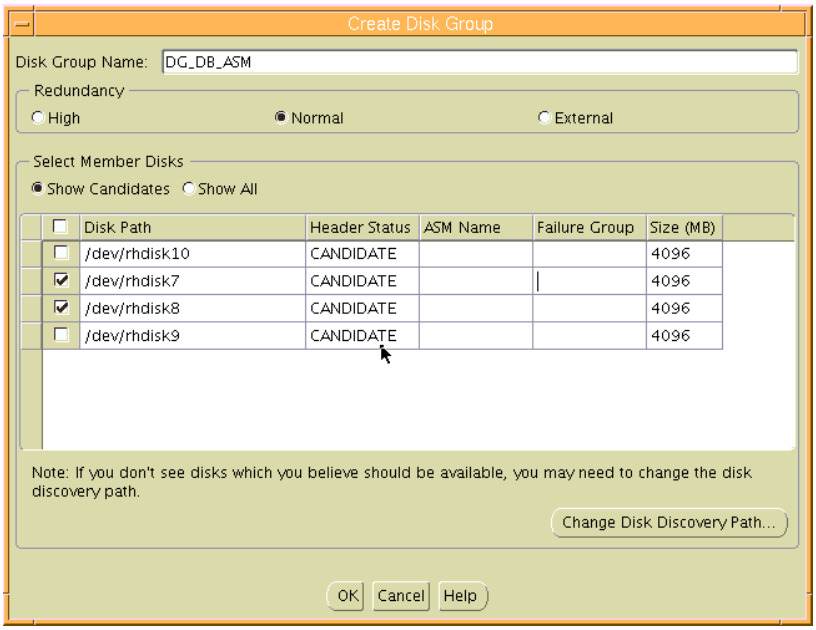
For our example, we specified DG_DB_ASM

Just select the disks you want in the Diskgroup.

If "Normal Redundancy" selected, ASM mirroring will be activated, meaning each database block will have a copy.

In this case, not using "Failure Group" will means that ASM will decide on each disks are going the copies.

Available free space will be divided by 2.



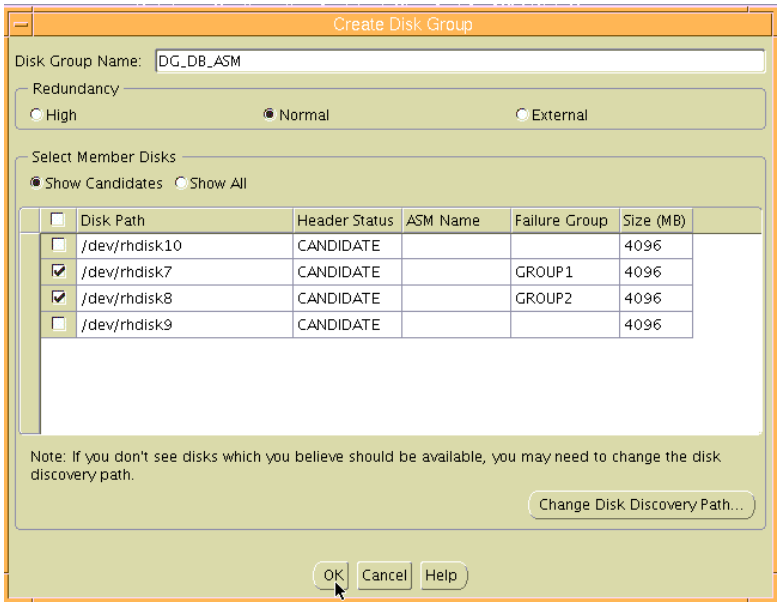
Thru SQLPlus, it will be the following syntax :

```
CREATE DISKGROUP DG_DB_ASM
NORMAL REDUNDANCY
DISK
'/dev/rhdisk7',
'/dev/rhdisk8';
```

OTHER Case : Normal Redundancy and Failure Group

- **If "Normal Redundancy" selected, ASM mirroring will be activated, meaning each database block will have a copy.**
- **If using "Failure Group", setting name of each group. This will ensure that each database block will have its original and copy on different Failure Group.**

Available free space will be divided by 2.



Thru SQLPlus, it will be the following syntax :

```
CREATE DISKGROUP DG_DB_ASM
NORMAL REDUNDANCY
FAILGROUP GROUP1 DISK
'/dev/rhdisk7',
FAILGROUP GROUP2 DISK
'/dev/rhdisk8';
```

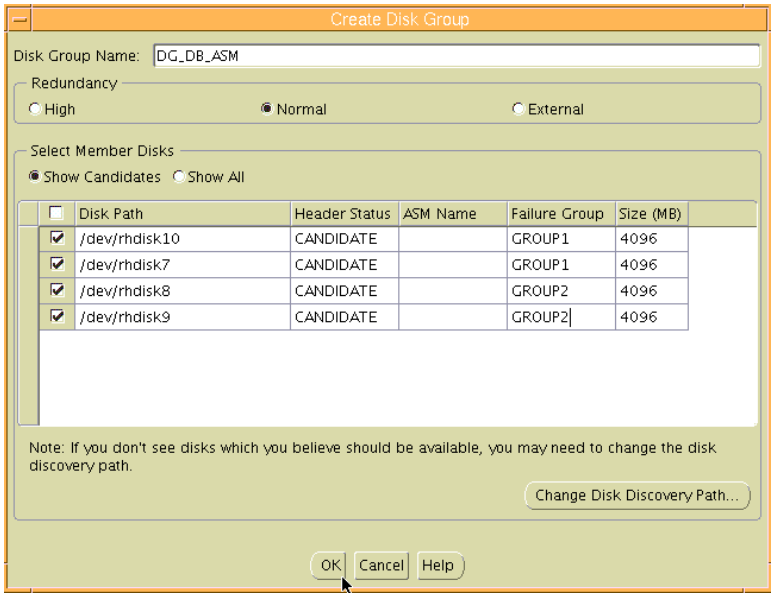
OTHER Case : High Redundancy and Failure Group

- **If "High Redundancy" selected, ASM mirroring will be activated, meaning each database block will have 2 copies.**
- **If using "Failure Group", setting name of each group, we'll need 3 failure groups, GROUP1, GROUP2 and GROUP3. This will ensure that each database block will have its original and copies on different Failure Group.**

Available free space will be divided by about 3

OTHER Case : Normal Redundancy and Failure Group

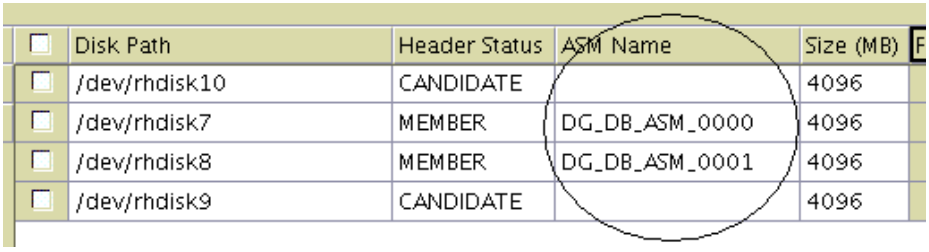
- **In this case, each failure group will have 2 disks each.**



Thru SQLPlus, it will be the following syntax :

```
CREATE DISKGROUP DG_DB_ASM
NORMAL REDUNDANCY
FAILGROUP GROUP1 DISK
'/dev/rhdisk10',
'/dev/rhdisk7',
FAILGROUP GROUP2 DISK
'/dev/rhdisk8',
'/dev/rhdisk9';
```

NEVER set any name in the field "ASM Name", ASM will generate the names.

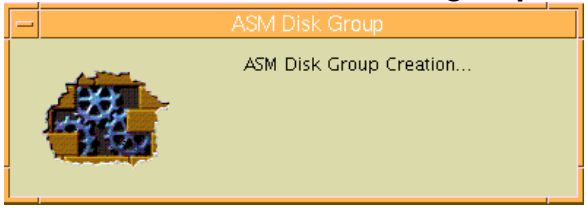


Disk Path	Header Status	ASM Name	Size (MB)
/dev/rhdisk10	CANDIDATE		4096
/dev/rhdisk7	MEMBER	DG_DB_ASM_0000	4096
/dev/rhdisk8	MEMBER	DG_DB_ASM_0001	4096
/dev/rhdisk9	CANDIDATE		4096

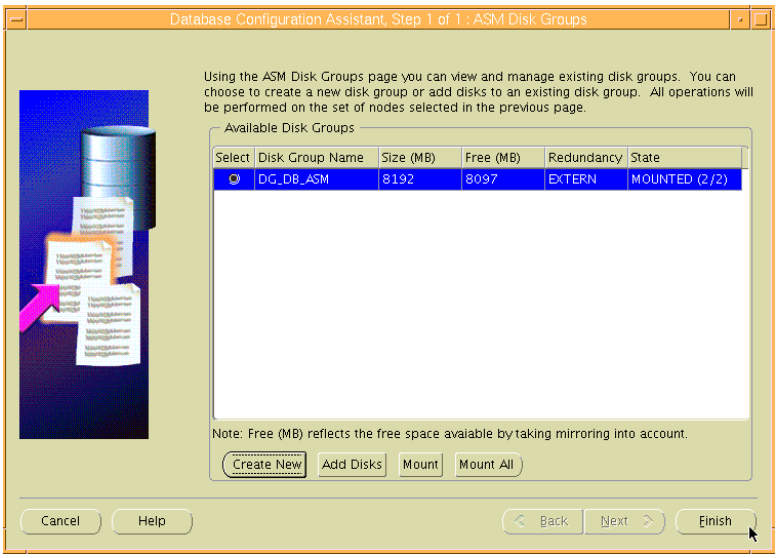
Do choose your Diskgroup configuration and click "OK" to create the diskgroup !!!

Create Disk Group :

The assistant will create a Disk Group named **DG_DB_ASM**.



ASM Disk Groups :



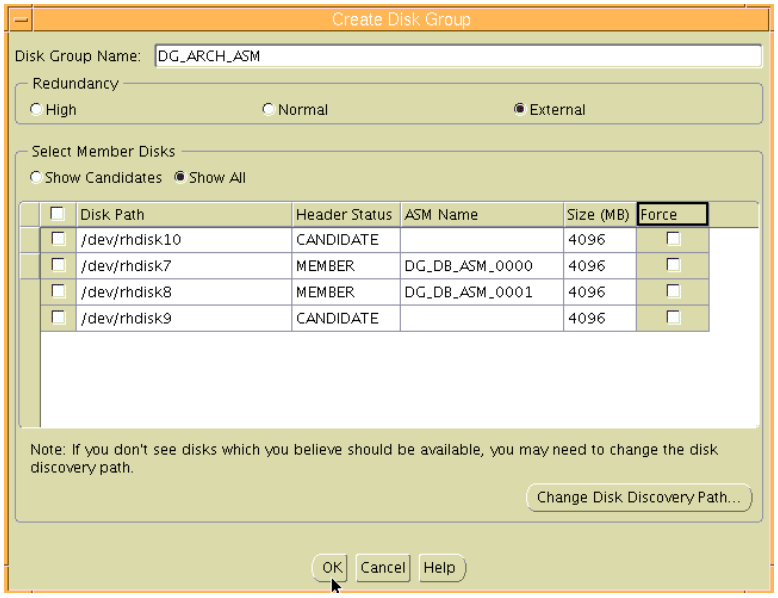
DG_DB_ASM Disk Group created

SELECT "**Create New**" for new Disk Group.

OR "**Finish**" to exit

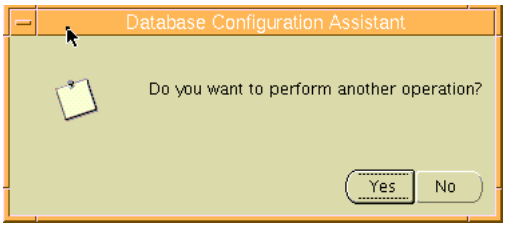
If you click on "add Disk", you can add more disks in the created diskgroup.

And see the "ASM name" generated by ASM for the disks.

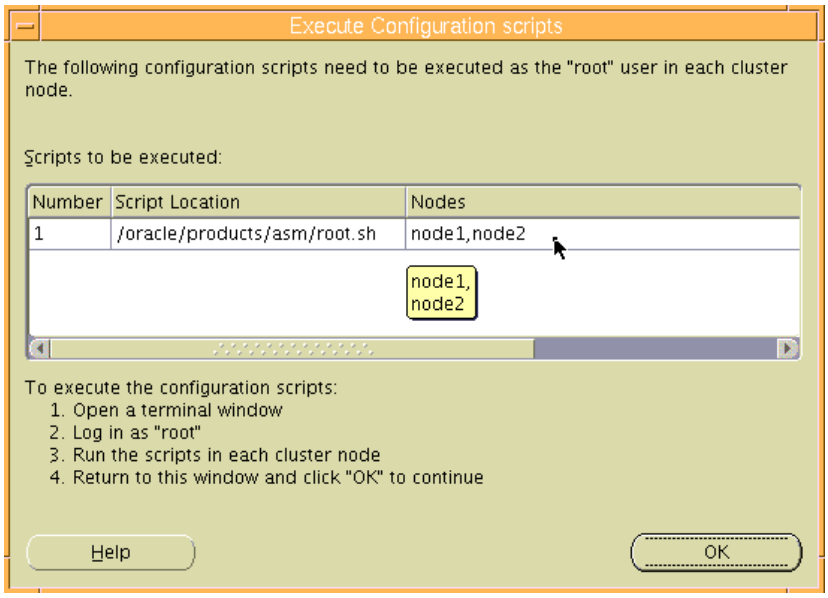


Coming back to :

Click on "Create New" to create other diskgroup like for redo and archive logs
Or click on "Finish" to exit



Click on "NO" to carry on the Install process …



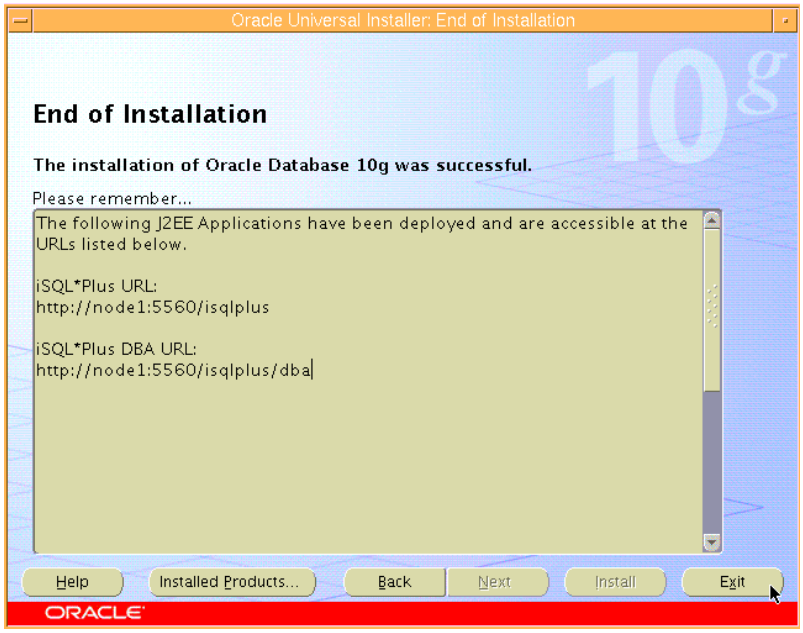
Execute Configuration Scripts will pop-up :

AS root, execute root.sh on each node.

For our case, this script is located in the /oracle/products/asm

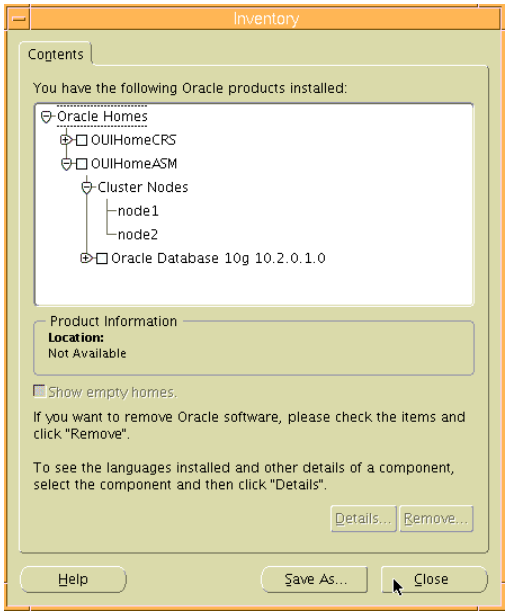
Just click OK ...

```
{node1:root}/oracle/products/asm -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/asm

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node1:root}/oracle/products/asm ->

------------------------------------------------------------

{node2:root}/oracle/products/asm -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/asm

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node2:root}/oracle/products/asm ->
```



End of Installation :

Click on “Installed Products” to chek with the oraInventory that ASM is installed on each node.



Just click on Close, then Exit ...

Subject: **How to cleanup ASM installation (RAC and Non-RAC)** Doc ID: **Note:311350.1**

Subject: **Backing Up an ASM Instance** Doc ID: **Note:333257.1**

Subject: **How to Re-configure Asm Disk Group?** Doc ID: **Note:331661.1**

Subject: **Assigning a Physical Volume ID (PVID) To An Existing ASM Disk Corrupts the ASM Disk Header** Doc ID: **Note:353761.1**

Subject: **How To Reclaim Asm Disk Space?** Doc ID: **Note:351866.1**

Subject: **Re-creating ASM Instances and Diskgroups** Doc ID: **Note:268481.1**

Subject: **How To Connect To Asm Instance Remotely** Doc ID: **Note:340277.1**

15.2 ASM Post-Configuration task

15.2.1 Cluster Ready Services Health Check

Check CRS processes on each nodes :

```
{node1:root}/ ->ps -ef|grep d.bin
    root  291030       1  99   Jan 25      - 747:31 /oh10g/crs/bin/crsd.bin reboot
  oracle  352346       1   0   Jan 25      -  0:18 /oh10g/crs/bin/evmd.bin
  oracle  401630  389342   0   Jan 25      -  2:35 /oh10g/crs/bin/ocssd.bin
    root  966694 1196094   0 17:35:57  pts/3  0:00 grep d.bin
{node1:root}/ ->
```

Execute

crs_stat –t

on one node as oracle user :

```
{node1:oracle}/oracle ->crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora....SM1.asm application    ONLINE    ONLINE    node1
ora....E1.lsnr application    ONLINE    ONLINE    node1
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora....SM2.asm application    ONLINE    ONLINE    node2
ora....E2.lsnr application    ONLINE    ONLINE    node2
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node1:oracle}/oracle ->
```

CR health check

```
{node1:oracle}/oracle/crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oracle/crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmchealth check

```
{node1:oracle}/oracle/crs/bin ->crsctl check  cssd
CSS appears healthy
{node1:oracle}/oracle/crs/bin ->crsctl check  crsd
CRS appears healthy
{node1:oracle}/oracle/crs/bin ->crsctl check  evmd
EVM appears healthy

{node2:oracle}/oracle/crs/bin ->crsctl check  cssd
CSS appears healthy
{node2:oracle}/oracle/crs/bin ->crsctl check  crsd
CRS appears healthy
{node2:oracle}/oracle/crs/bin ->crsctl check  evmd
EVM appears healthy
```

Check ASM deamons

For example on node1 :

```
{node1:oracle}/oracle/crs/bin ->ps -ef|grep ASM
  oracle  860214       1   0   Feb 06      -  0:09 asm_lck0_+ASM1
  oracle  925704       1   1   Feb 06      - 42:28
/oracle/products/asm/bin/racgimon daemon ora.fin1.ASM1.asm
  oracle  929828       1   0   Feb 06      -  0:12 asm_pmon_+ASM1
  oracle  938186       1   0   Feb 06      -  4:10 asm_diag_+ASM1
  oracle  942284       1   0   Feb 06      -  1:44 asm_lmon_+ASM1
  oracle  946382       1   0   Feb 06      -  0:54 asm_lmd0_+ASM1
  oracle  950480       1   0   Feb 06      -  1:11 asm_lms0_+ASM1
  oracle  954578       1   0   Feb 06      -  0:02 asm_mman_+ASM1
  oracle  962774       1   0   Feb 06      -  0:03 asm_dbw0_+ASM1
  oracle  966872       1   0   Feb 06      -  0:04 asm_lgwr_+ASM1
  oracle  970970       1   0   Feb 06      -  0:29 asm_ckpt_+ASM1
  oracle  975068       1   0   Feb 06      -  0:17 asm_smon_+ASM1
  oracle  979166       1   0   Feb 06      -  0:03 asm_rbal_+ASM1
  oracle 1089556       1   0   Feb 06      -  0:05 oracle+ASM1
(DESCRIPTION=(LOCAL=YES)(ADDRESS=(PROTOCOL=beq)))
  oracle 1224790       1   0   Feb 06      -  0:05 oracle+ASM1
(DESCRIPTION=(LOCAL=YES)(ADDRESS=(PROTOCOL=beq)))
    root 1286218 1265672   0 12:29:51  pts/0  0:00 grep ASM
```

{**node1**:root}/oracle/products/asm/dbs -> **cat init+ASM1.ora**
SPFILE='/dev/asmspf_disk'
{node1:root}/oracle/products/asm/dbs ->
{**node2**:root}/oracle/products/asm/dbs -> **cat init+ASM2.ora**
SPFILE='/dev/asmspf_disk'
{node2:root}/oracle/products/asm/dbs ->

Check ASM spfile or initfile content

For example on node1 :

```
##############################################################################
# Copyright (c) 1991, 2001, 2002 by Oracle Corporation
##############################################################################

###########################################
# Cluster Database
###########################################
cluster_database=true

###########################################
# Diagnostics and Statistics
###########################################
background_dump_dest=/oracle/admin/+ASM/bdump
core_dump_dest=/oracle/admin/+ASM/cdump
user_dump_dest=/oracle/admin/+ASM/udump

###########################################
# Miscellaneous
###########################################
instance_type=asm

###########################################
# Pools
###########################################
large_pool_size=12M

###########################################
# Security and Auditing
###########################################
remote_login_passwordfile=exclusive

asm_diskgroups='DG_DB_ASM'

+ASM2.instance_number=2
+ASM1.instance_number=1
```

15.2.2 Oracle User .profile update

To be done on each node.

Oracle environment : vi **$HOME/.profile** file in Oracle's home directory

Add the entries in bold blue color

```
export ORACLE_BASE=/oracle
export AIXTHREAD_SCOPE=S
export TMP=/tmp
export TEMP=/tmp
export TMPDIR=/tmp
umask 022
export ORA_CRS_HOME=$ORACLE_BASE/products/crs
export CRS_HOME=$ ORA_CRS_HOME
export ORA_ASM_HOME=$ORACLE_BASE/products/asm
export TNS_ADMIN=$ORA_ASM_HOME/network/admin
export LD_LIBRARY_PATH=$ORA_ASM_HOME/lib:$ORA_CRS_HOME/lib:$ORA_ASM_HOME/lib32:$ORA_CRS_HOME/lib32
export LIBPATH=$LD_LIBRARY_PATH
export PATH=$ORA_ASM_HOME/bin:$ORA_CRS_HOME/bin:$PATH
```

Do disconnect from oracle user, and reconnect to load modified $HOME/.profile

15.3 WHAT HAS BEEN DONE ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**
- **10.2.0.3 patchset is applied for CRS home**
- **ASM Home is installed**
- **ASM instance is created and started**
- **ASM Diskgroup is created**

15.4 INSTALL 10.2.0.3 PATCHSET TO ASM HOME

On each node :

Stop the listener from ASM_HOME

Stop listener on all nodes. To stop listener running on a node, enter the following command where *node* is the name of the node where the applications are running:

As oracle user, execute :

$ *oracle_home*/bin/srvctl stop listener -n *node* –l *listener_name*

```
{node1:root}/oracle -> srvctl stop listener –n node1
{node1:root}/oracle ->
{node1:root}/oracle -> srvctl stop listener –n node2
{node1:root}/oracle ->
```

On each node :

Stop the ASM instance

Stop ASM instance on all nodes. To stop ASM instance running on a node, enter the following command where *node* is the name of the node where the applications are running:

As oracle user, execute :

$ *oracle_home*/bin/srvctl stop asm -n *node*

```
{node1:root}/oracle -> srvctl stop asm –n node1
{node1:root}/oracle ->
{node1:root}/oracle -> srvctl stop asm –n node2
{node1:root}/oracle ->
```

On each node :

Then running "crs_stat –t" as oracle, or root on node1 should return the following result :

```
{node1:root}/oracle -> crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora....SM1.asm application    OFFLINE   OFFLINE
ora....E1.lsnr application    OFFLINE   OFFLINE
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora....SM2.asm application    OFFLINE   OFFLINE
ora....E2.lsnr application    OFFLINE   OFFLINE
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node1:oracle}/oracle ->
```

Back Up the System, Oracle recommends that you create a backup of the Oracle Clusterware 10*g* installation before you install the patch set, and a backup of the OCR content.

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node1:root}/oracle/products ->

{node2:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products ->
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

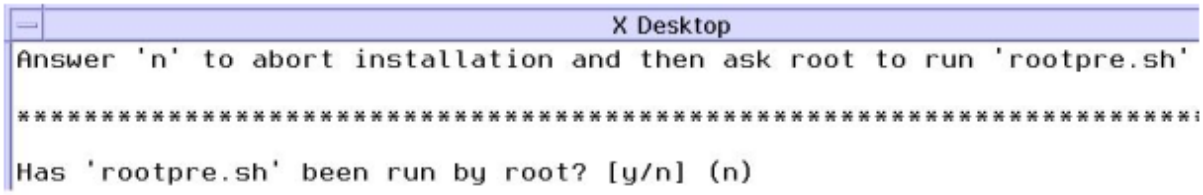
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

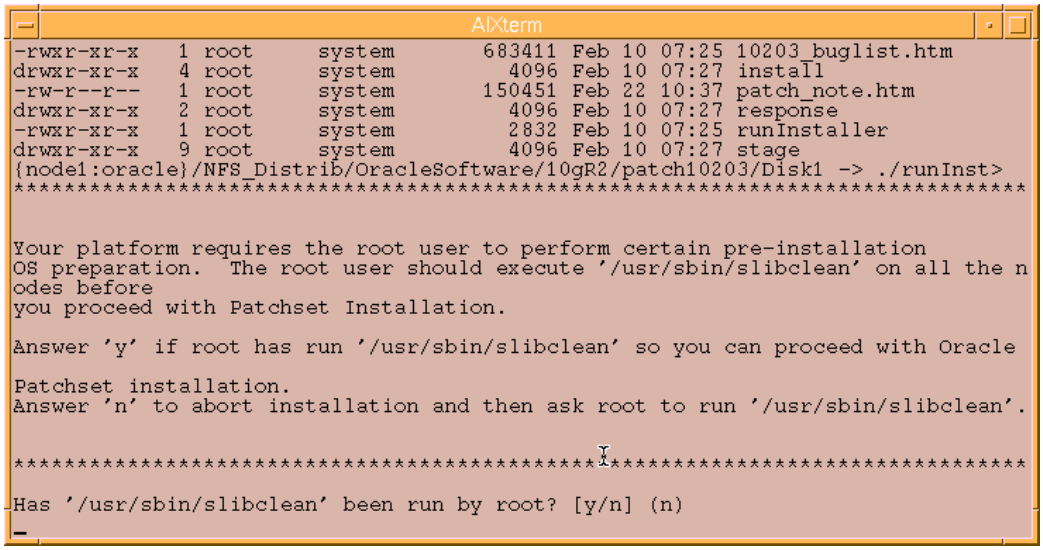
Or execute : ./runInstaller -ignoreSysPrereqs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
All filesets for 5300-04_AIX_ML were found.
All filesets for 5300-05_AIX_ML were found.



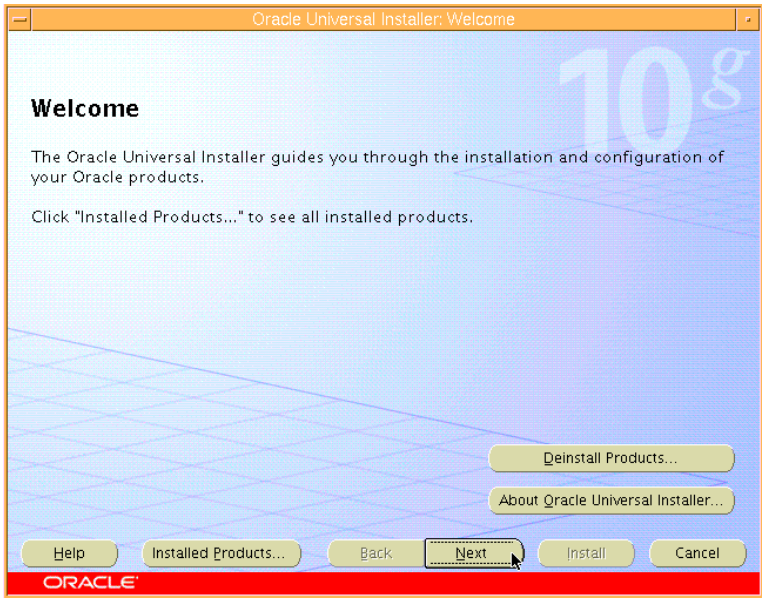
Should be already done with the CRS Installation), Answer “y” for next step



At the OUI Welcome screen

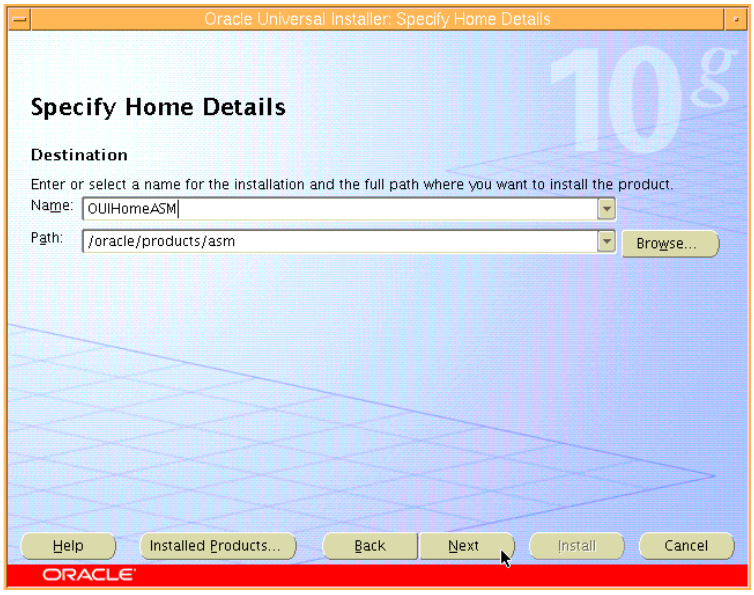
Enter "Yes" if /usr/sbin/slibclean has been executed on both node as root.

Just click Next ...



At the OUI Welcome screen

Just click Next ...



Select the ORACLE_HOME corresponding to the ASM_HOME Installation directory !!!

For our case :

OUIHomeASM
/oracle/asm

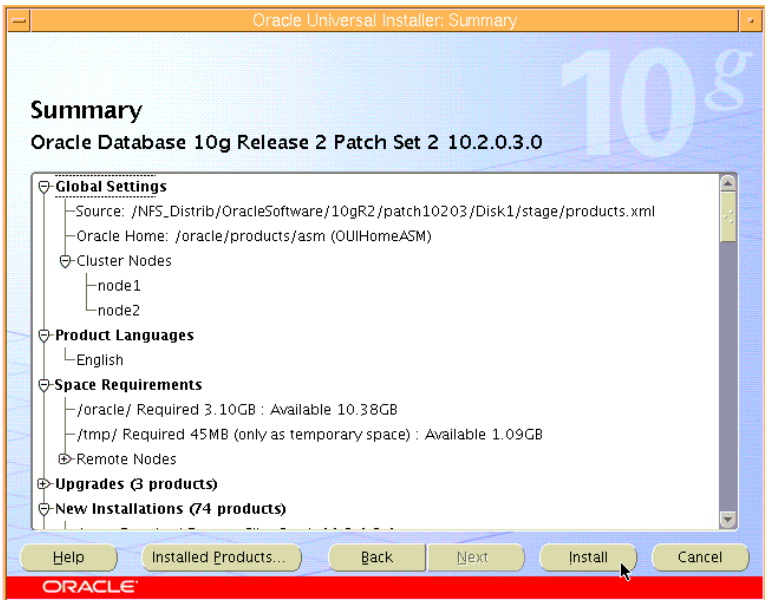
Just click Next ...

At this stage :

No choice available for local Installation.

Cluster Installation with all nodes selected by default.

Just click Next ...

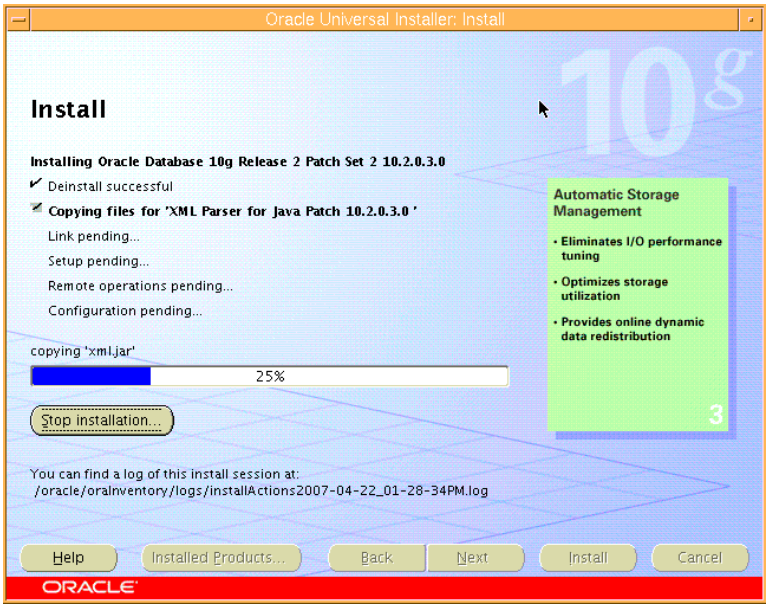


Summary :

The Summary screen will be presented.

Check Cluster Nodes and Remote Nodes lists.
The OUI will install the Oracle 10g software on to the local node, and then copy this information to the other selected nodes.

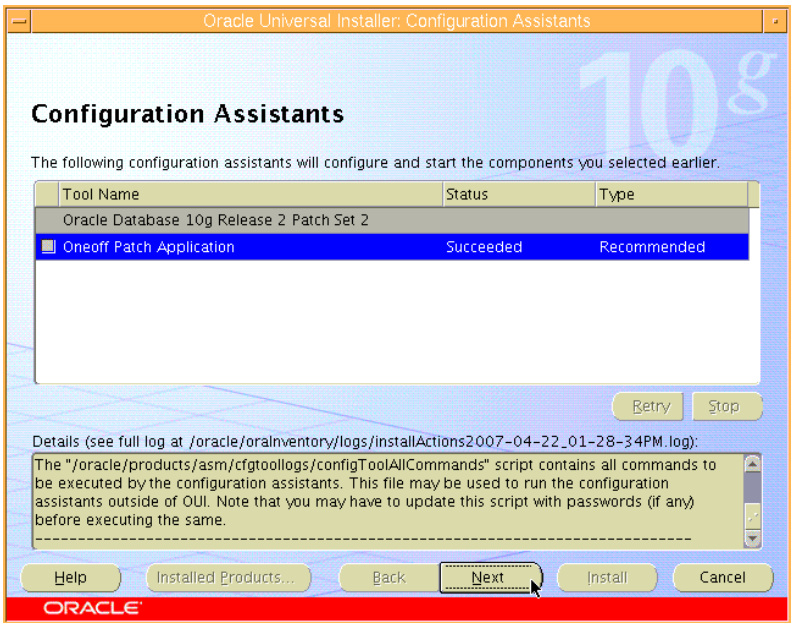
Then click Install ...



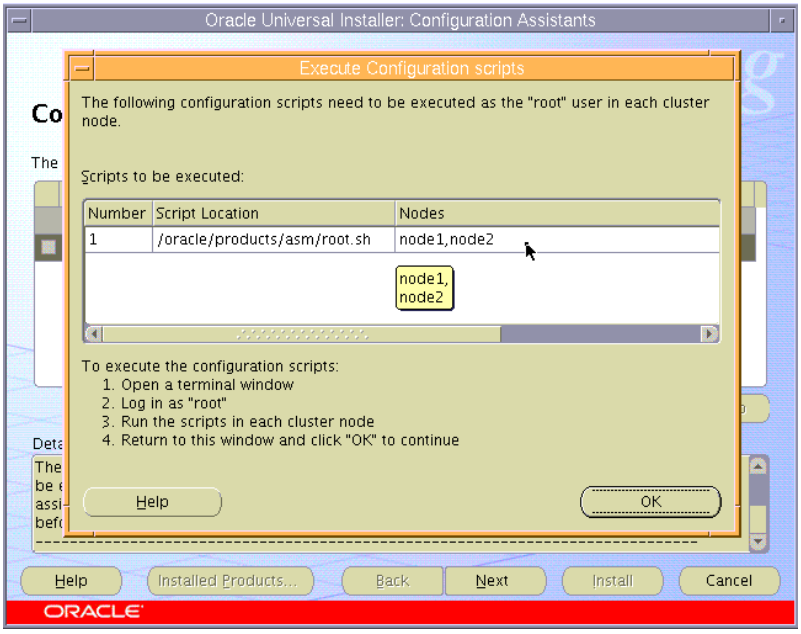
Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the others selected nodes.

Just wait for the next screen ...



Configuration Assistants will be executed, and if no problems appears next windows will be displayed …



Execute Configuration Scripts will pop-up :

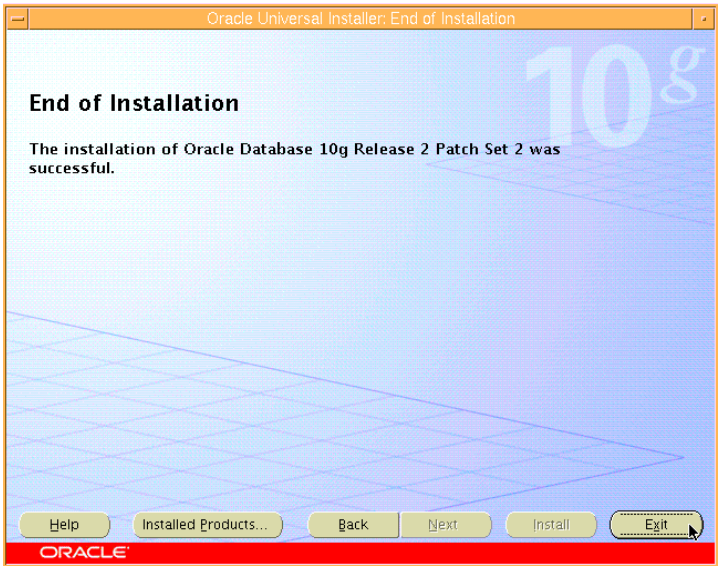
AS root, execute root.sh on each node.

For our case, this script is located in the /oracle/products/asm

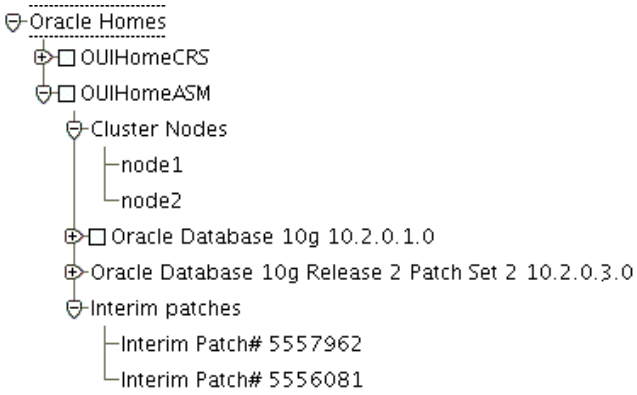
Just click OK ...

```
{node1:root}/oracle/products/asm -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/asm

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node1:root}/oracle/products/asm ->

------------------------------------------------------------------

{node2:root}/oracle/products/asm -> ./root.sh
Running Oracle10 root.sh script...
```

```
The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/asm

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node2:root}/oracle/products/asm ->
```



Enf of Installation :



Just click EXIT ...

Subject: **How to upgrade an ASM Instance From 10.2.0.1 To 10.2.0.2, or 10.2.0.3 ?** Doc ID: **Note:370355.1**
➔ **No need to update an ASM instance as it will be necessary/done for an Oracle database instance.**

Apply metalink note to remove ASM dependancy to VIP:

Subject: **Changes in Oracle Clusterware after applying 10.2.0.3 Patchset** Doc ID: **Note:401783.1**

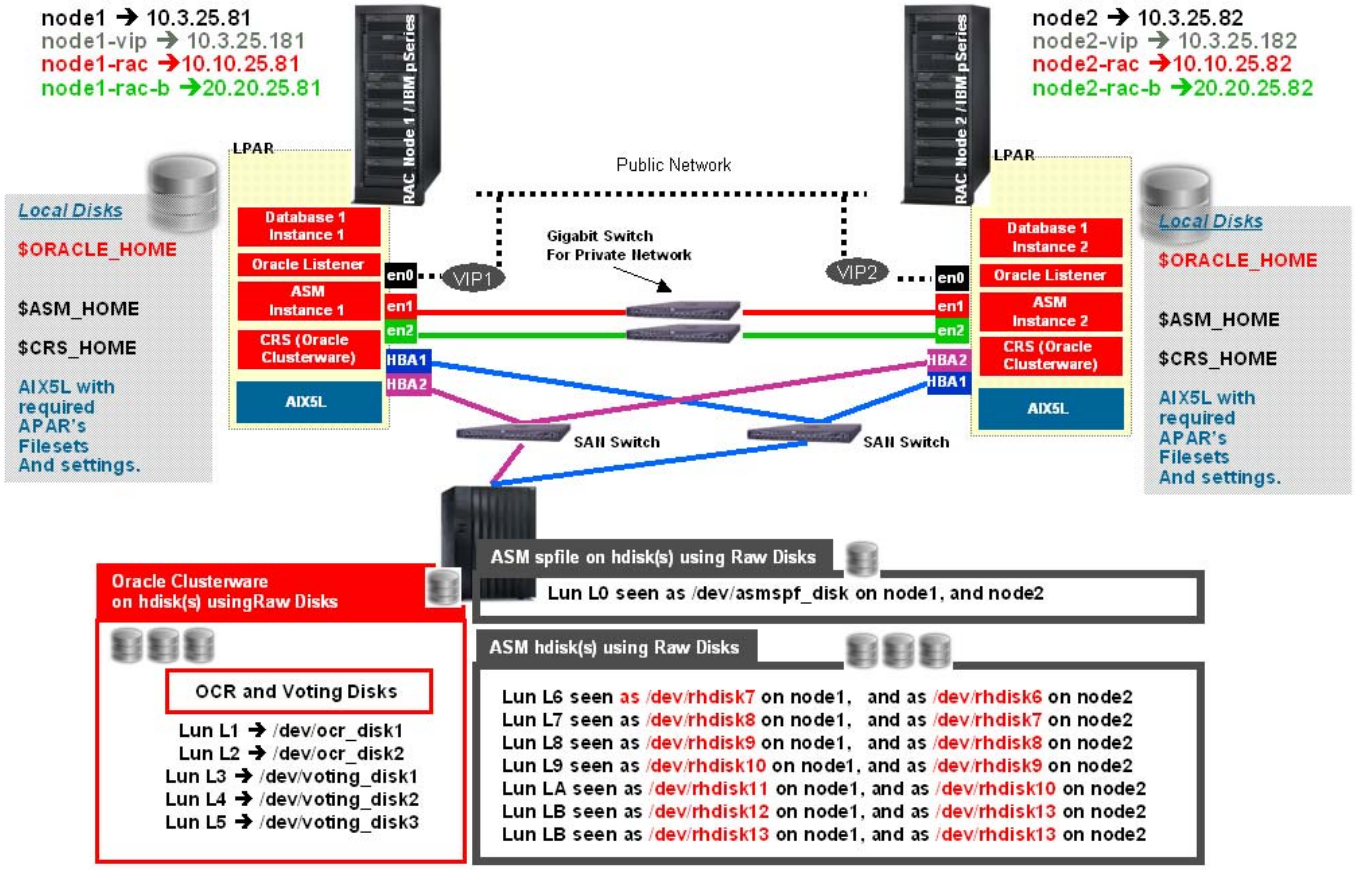
15.5 WHAT HAS BEEN DONE ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**
- **10.2.0.3 patchset is applied for CRS home**
- **ASM Home is installed**
- **ASM instance is created and started**
- **ASM Diskgroup is created**
- **10.2.0.3 pachset is applied for ASM home**

16 INSTALLING ORACLE 10G R2 SOFTWARE

Starting DB Home Installation From First node !!!



16.1 10G R2 RDBMS INSTALLATION

Oracle RAC option installation just have to be done only starting from one node. Once the first node is installed, Oracle OUI automatically starts the copy of the mandatory files on the second node, using **rcp** command. This step could last long, depending on the network speed (one hour…), without any message. So, don't think the OUI is stalled, and look at the network traffic before canceling the installation !
You can also create a staging area. The name of the subdirectories is in the format "Disk1" to "Disk3"

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products -> /usr/sbin/slibclean
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

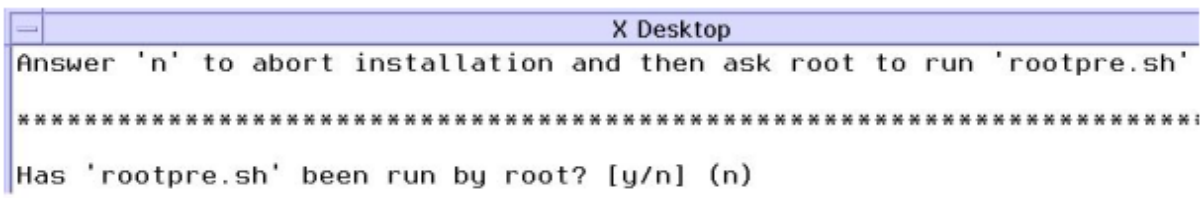
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

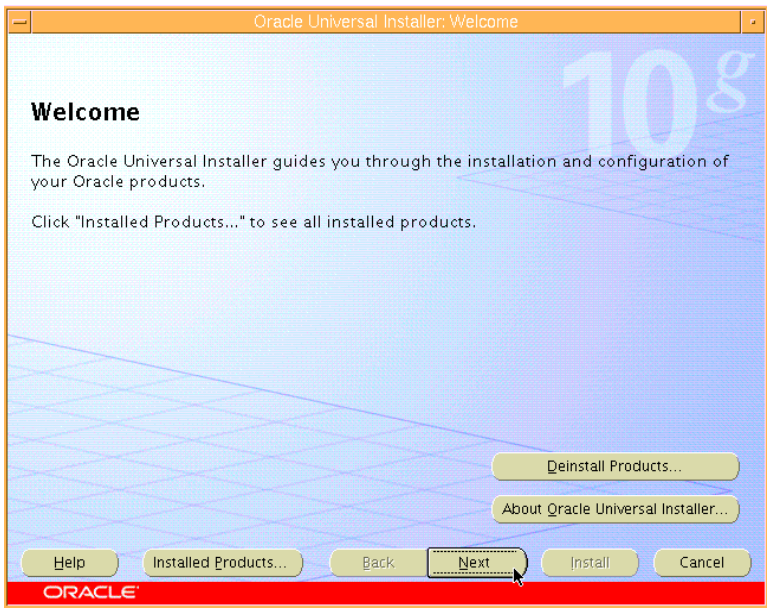
Or execute : ./runInstaller -ignoreSysPreregs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
…

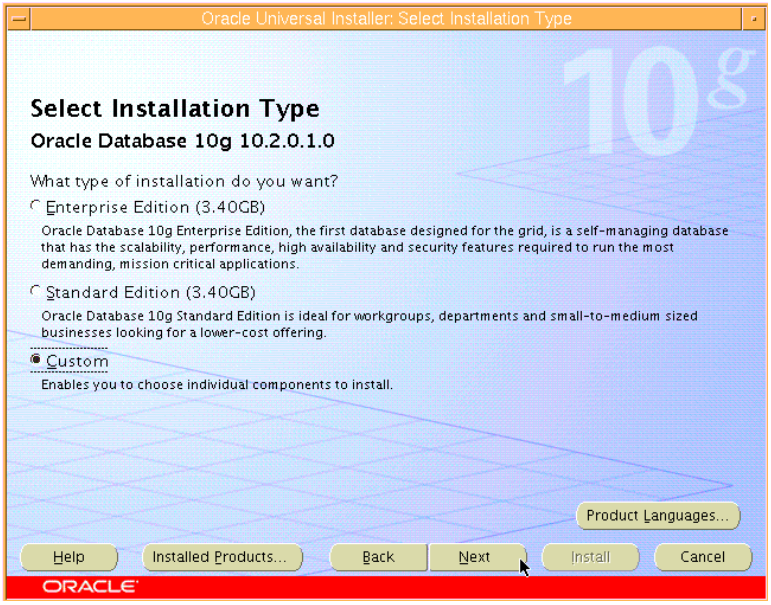


Should be already done with the CRS Installation), Answer "y" for next step



At the OUI Welcome screen

Just click Next ...

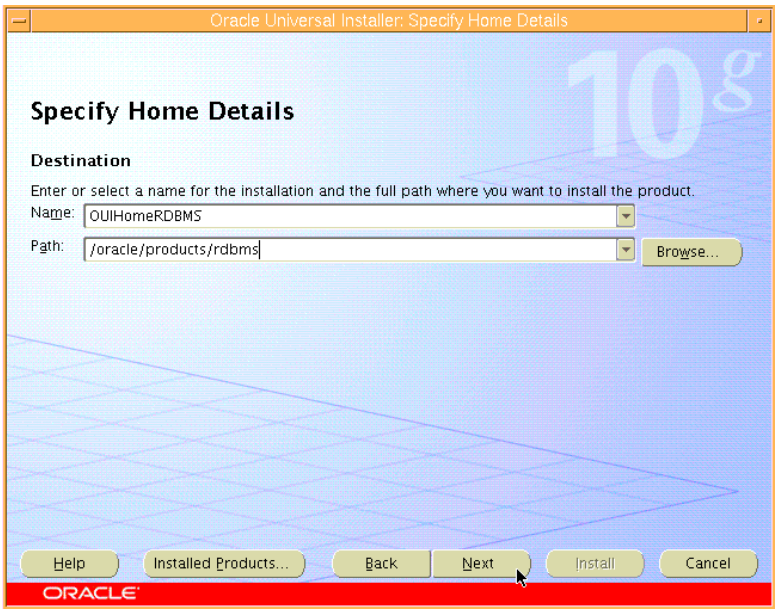


Select the installation type :

You have the option to choose Enterprise, Standard Edition, or Custom to proceed.

Choose the "Custom" option to avoid creating a database by default.

Then click Next ...



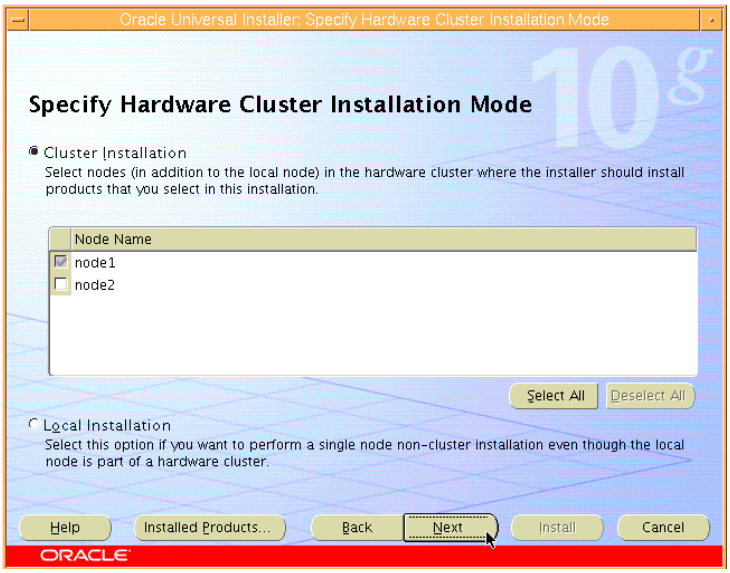
Specify File Locations :
Do not change the Source field

Specify a different ORACLE_HOME Name with its own directory for the Oracle software installation.

This ORACLE_HOME must be different then the CRS and ASM ORACLE_HOME.

Then click Next ...

If you don't see the following screen with Node selection, it might be that your CRS is down on one or all nodes. ➔ Please check if CRS is up and running on all nodes.

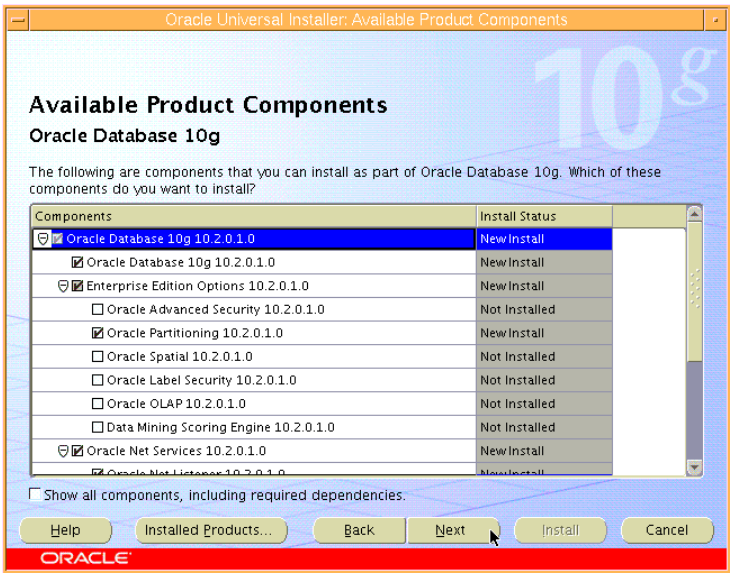


Specify Hardware Cluster Installation Mode :

Select **Cluster Installation**

AND the other nodes on to which the Oracle RDBMS software will be installed. It is not necessary to select the node on which the OUI is currently running. Click Next.

Then click Next ...

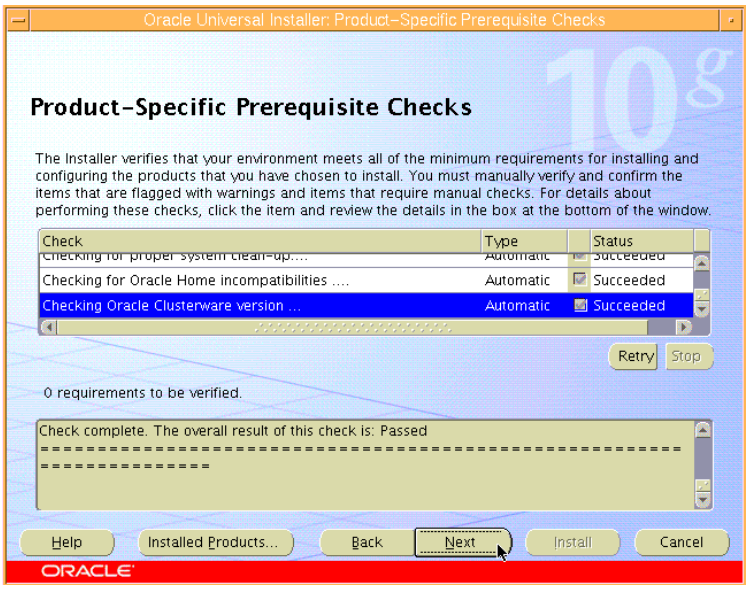


Available Product Components :

Select the product components for Oracle Database 10g that you want to install.

INFO : Compared to 10gRAC R1 installation, there is no "Real Application Cluster" option to select.

Then click Next ...



The installer will check some product-specific Prerequisite.

Don't take care of the lines with checking at status "Not executed", These are just warnings because AIX maintenance level might be higher then 5300, which is the case in our example (ML03).

Then click Next ...

Details of the prerequisite checks done by runInstaller

Checking operating system requirements ...
Expected result: One of 5200.004,5300.002
Actual Result: 5300.002
Check complete. The overall result of this check is: Passed
==========

Checking operating system package requirements ...
Checking for bos.adt.base(0.0); found bos.adt.base(5.3.0.51). Passed
Checking for bos.adt.lib(0.0); found bos.adt.lib(5.3.0.50). Passed
Checking for bos.adt.libm(0.0); found bos.adt.libm(5.3.0.40). Passed
Checking for bos.perf.libperfstat(0.0); found bos.perf.libperfstat(5.3.0.50). Passed
Checking for bos.perf.perfstat(0.0); found bos.perf.perfstat(5.3.0.50). Passed
Checking for bos.perf.proctools(0.0); found bos.perf.proctools(5.3.0.50). Passed
Check complete. The overall result of this check is: Passed
==========

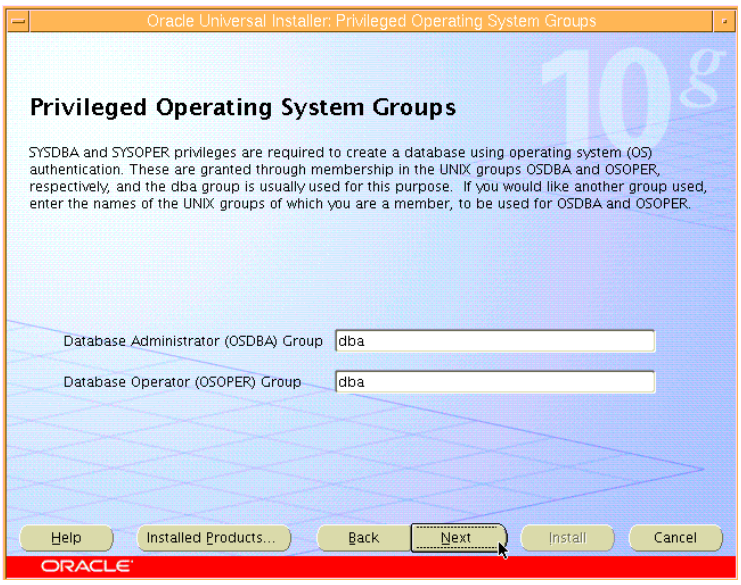
Checking recommended operating system patches
Checking for IY59386(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY60930(bos.mp,5.3.0.1); found (bos.mp,5.3.0.54). Passed
Checking for IY60930(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp64,5.3.0.20); found (bos.mp64,5.3.0.54). Passed
Checking for IY66513(bos.mp,5.3.0.20); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp,5.3.0.22); found (bos.mp,5.3.0.54). Passed
Checking for IY70159(bos.mp64,5.3.0.22); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.mp64,5.3.0.1); found (bos.mp64,5.3.0.54). Passed
Checking for IY58143(bos.acct,5.3.0.1); found (bos.acct,5.3.0.51). Passed
Checking for IY58143(bos.adt.include,5.3.0.1); found (bos.adt.include,5.3.0.53). Passed
Checking for IY58143(bos.adt.libm,5.3.0.1); found (bos.adt.libm,5.3.0.40). Passed
Checking for IY58143(bos.adt.prof,5.3.0.1); found (bos.adt.prof,5.3.0.53). Passed
Checking for IY58143(bos.alt_disk_install.rte,5.3.0.1); found (bos.alt_disk_install.rte,5.3.0.51). Passed
Checking for IY58143(bos.cifs_fs.rte,5.3.0.1); found (bos.cifs_fs.rte,5.3.0.50). Passed
Checking for IY58143(bos.diag.com,5.3.0.1); found (bos.diag.com,5.3.0.51). Passed
Checking for IY58143(bos.perf.libperfstat,5.3.0.1); found (bos.perf.libperfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.perfstat,5.3.0.1); found (bos.perf.perfstat,5.3.0.50). Passed
Checking for IY58143(bos.perf.tools,5.3.0.1); found (bos.perf.tools,5.3.0.52). Passed
Checking for IY58143(bos.rte.boot,5.3.0.1); found (bos.rte.boot,5.3.0.51). Passed
Checking for IY58143(bos.rte.archive,5.3.0.1); found (bos.rte.archive,5.3.0.51). Passed
Checking for IY58143(bos.rte.bind_cmds,5.3.0.1); found (bos.rte.bind_cmds,5.3.0.51). Passed
Checking for IY58143(bos.rte.control,5.3.0.1); found (bos.rte.control,5.3.0.50). Passed
Checking for IY58143(bos.rte.filesystem,5.3.0.1); found (bos.rte.filesystem,5.3.0.51). Passed
Checking for IY58143(bos.rte.install,5.3.0.1); found (bos.rte.install,5.3.0.54). Passed
Checking for IY58143(bos.rte.libc,5.3.0.1); found (bos.rte.libc,5.3.0.53). Passed
Checking for IY58143(bos.rte.lvm,5.3.0.1); found (bos.rte.lvm,5.3.0.53). Passed
Checking for IY58143(bos.rte.man,5.3.0.1); found (bos.rte.man,5.3.0.50). Passed
Checking for IY58143(bos.rte.methods,5.3.0.1); found (bos.rte.methods,5.3.0.51). Passed
Checking for IY58143(bos.rte.security,5.3.0.1); found (bos.rte.security,5.3.0.53). Passed
Checking for IY58143(bos.rte.serv_aid,5.3.0.1); found (bos.rte.serv_aid,5.3.0.52). Passed
Check complete. The overall result of this check is: Passed
==========

Validating ORACLE_BASE location (if set) ...
Check complete. The overall result of this check is: Passed
==========

Checking for proper system clean-up....
Check complete. The overall result of this check is: Passed
==========

Checking for Oracle Home incompatibilities ....
Actual Result: NEW_HOME
Check complete. The overall result of this check is: Passed
==========

Checking Oracle Clusterware version ...
Check complete. The overall result of this check is: Passed
==========

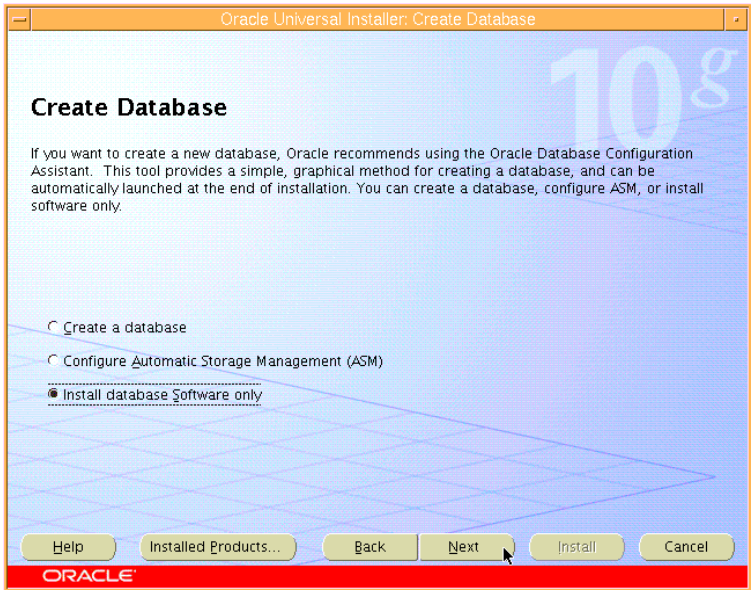


Privileged Operating Systems Groups :

Verify the UNIX primary group name of the user which controls the installation of the Oracle10g software.
(Use unix command id to find out)

And specify the Privileged Operating System Groups to the value found.
In our example, this must be “dba” (Primary group of unix oracle user) to be set for both entries.

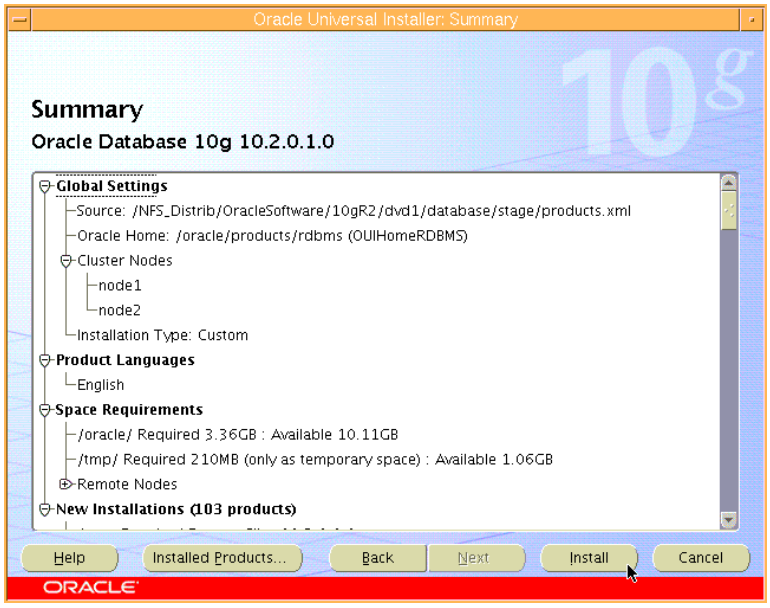
Then click Next ...



Create Database :

Choose “Install database Software only”, we don’t want to create a database at this stage.

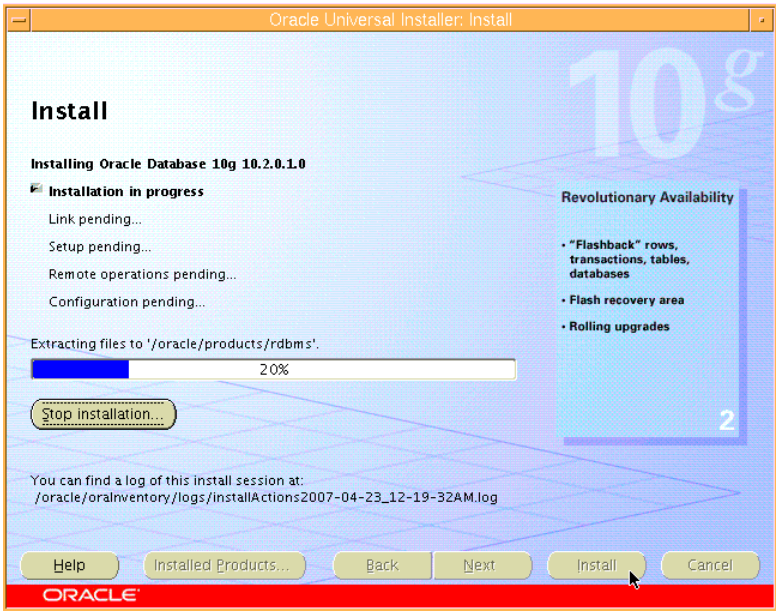
Then click Next ...



Summary :

The Summary screen will be presented. Confirm that the RAC database software and other selected options will be installed.
Check Cluster Nodes and Remote Nodes lists.
The OUI will install the Oracle 10g software on to the local node, and then copy this information to the other selected nodes.

Then click Install ...



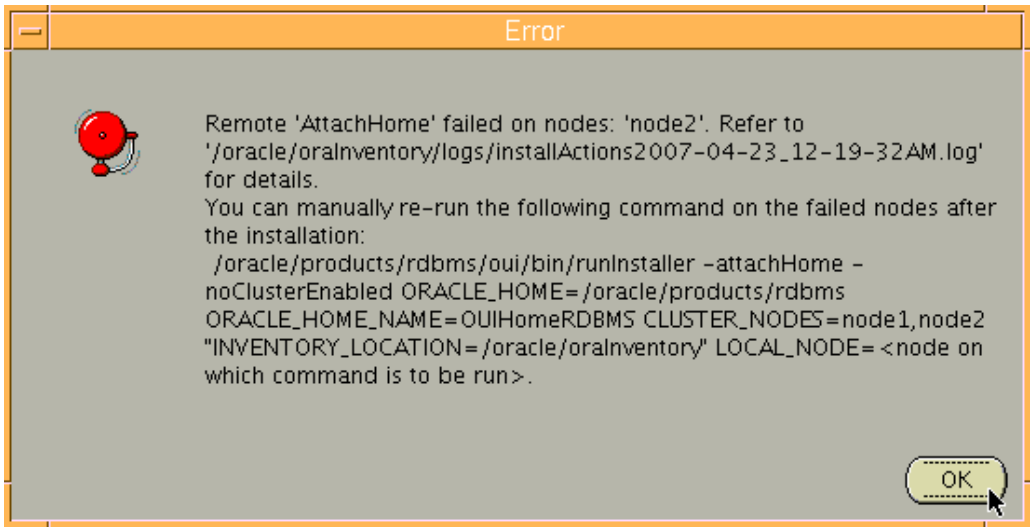
Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the others selected nodes.

Just wait for the next screen ...

This screen may appears, if so just run the specified command on the specified node as oracle user.

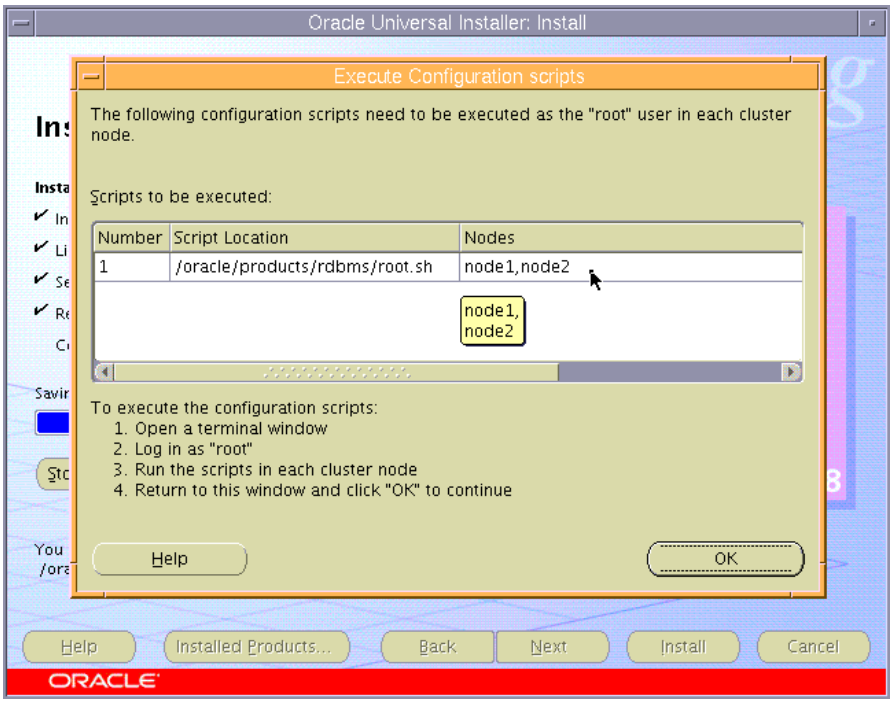
If you get this following error message ➔



Do execute the script bellow :

From node2 :

```
{node2:root}/oracle/products/rdbms -> su – oracle
{node2:oracle}/oracle -> /oracle/products/asm/oui/bin/runInstaller –attachHome –
noClusterEnabled ORACLE_HOME=/oracle/products/rdbms ORACLE_HOME_NAME=OUIHomeRDBMS
CLUSTER_NODES=node1,node2 “INVENTORY_LOCATION=/oracle/oraInventory” LOCAL_NODE=node2
Starting Oracle Universal Installer…

No pre-requisite checks found in oraparam.ini, no system pre-requisite checks will be
executed.
The inventory pointer is located at /etc/oraInst.loc
The inventory is located at /oracle/oraInventory
‘AttachHome’ was successful.
{node2:oracle}/oracle ->
```



Execute Configuration Scripts will pop-up :

AS root, execute root.sh on each node.

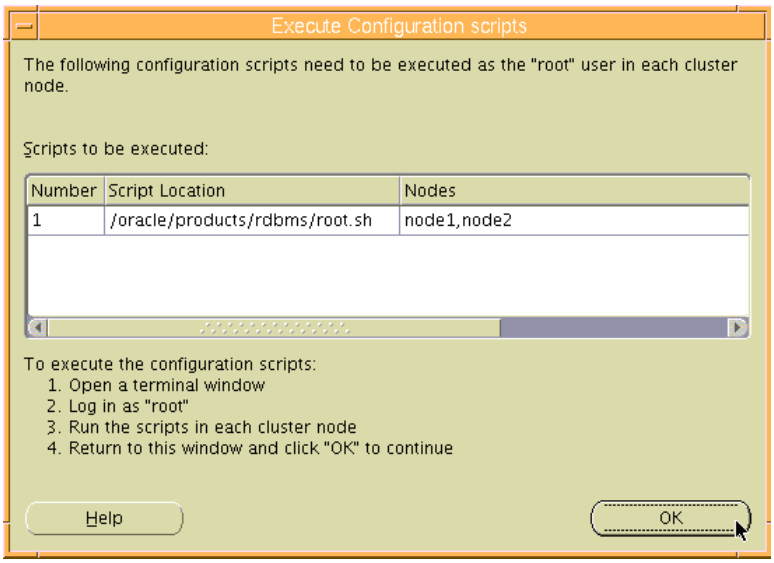
For our case, this script is located in the /oracle/products/asm

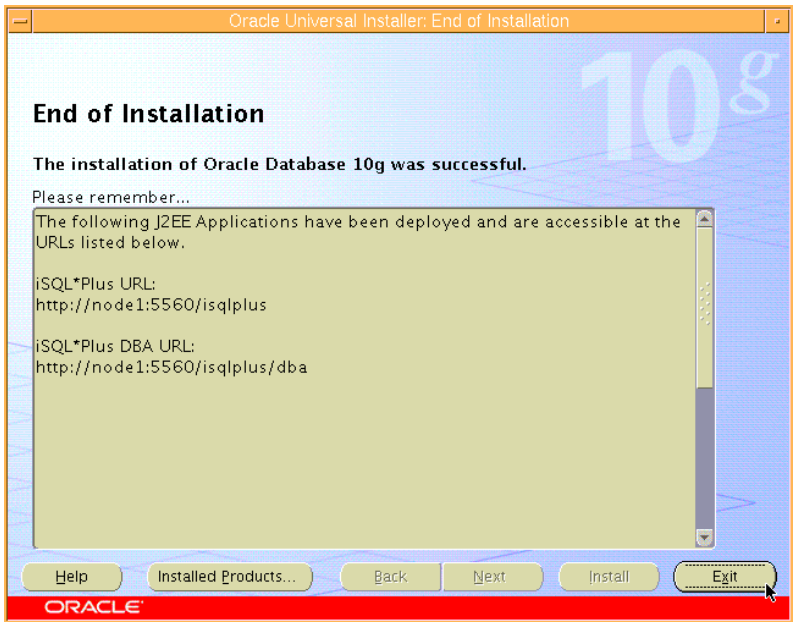
Just click OK ...

```
{node1:root}/oracle/products/rdbms -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/rdbms

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node1:root}/oracle/products/rdbms ->

------------------------------------------------------------------

{node2:root}/oracle/products/rdbms -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/rdbms

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node2:root}/oracle/products/rdbms ->
```

Coming back to this previous screen,

Just click OK



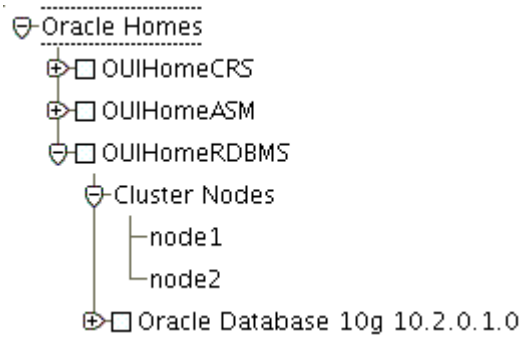


End of Installation :

This screen will automatically appear.

Check that it is successful and write down the URL list of the J2EE applications that have been deployed (isqlplus, …).

Then click Exit ...



16.2 10G SOFTWARE POST-INSTALLATION TASK

16.2.1 Symbolic links creation for listener.ora, tnsnames.ora and sqlnet.ora

As oracle user on each node :

```
{node1:oracle}/ -> ln -s /oracle/products/asm/network/admin/listener.ora
/oracle/products/rdbms/network/admin/listener.ora
{node1:oracle}/ -> ln -s /oracle/products/asm/network/admin/tnsnames.ora
/oracle/products/rdbms/network/admin/tnsnames.ora
{node1:oracle}/ -> ln -s /oracle/products/asm/network/admin/sqlnet.ora
/oracle/products/rdbms/network/admin/sqlnet.ora

{node1:oracle}/ -> ls –la /oracle/products/rdbms/network/admin/*.ora
lrwxrwxrwx   1 oracle   dba             47 Apr 23 10:19 listener.ora ->
/oracle/products/asm/network/admin/listener.ora
lrwxrwxrwx   1 oracle   dba             47 Apr 23 10:19 tnsnames.ora ->
/oracle/products/asm/network/admin/tnsnames.ora
lrwxrwxrwx   1 oracle   dba             47 Apr 23 10:19 sqlnet.ora ->
/oracle/products/asm/network/admin/sqlnet.ora
{node1:oracle}/ ->
```

➔ Doing this will avoid dealing with many listener.ora, tnsnames.ora and sqlnet.ora, as setting TNS_ADMIN variable will not be enough as some oracle assistants tools will not take consideration of this variable.

16.2.2 Oracle User .profile update

To be done on each node.

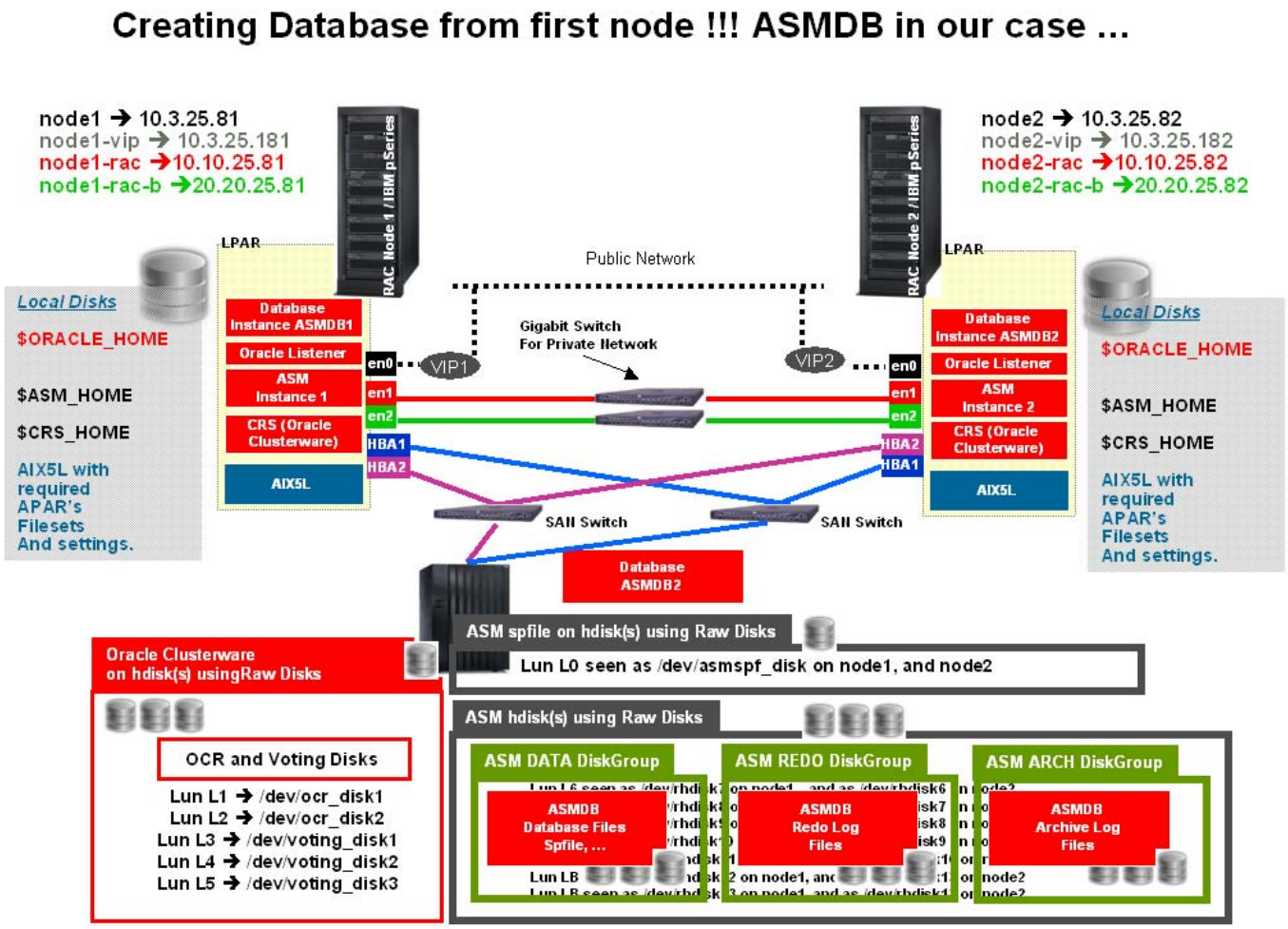
Oracle environment : vi **$HOME/.profile** file in Oracle's home directory

Add the entries in bold blue color

```
export ORACLE_BASE=/oracle
export AIXTHREAD_SCOPE=S
export TMP=/tmp
export TEMP=/tmp
export TMPDIR=/tmp
umask 022
export ORA_CRS_HOME=$ORACLE_BASE/products/crs
export CRS_HOME=$ORA_CRS_HOME
export ORA_ASM_HOME=$ORACLE_BASE/products/asm
export TNS_ADMIN=$ORA_ASM_HOME/network/admin

export ORACLE_HOME=$ORACLE_BASE/products/rdbms
export LD_LIBRARY_PATH=$ORACLE_HOME/lib:$CRS_HOME/lib:$ORACLE_HOME/lib32:$ CRS_HOME/lib32
export LIBPATH=$LD_LIBRARY_PATH
export PATH=$ORACLE_HOME/bin:$ORA_CRS_HOME/bin:$PATH
```

Do disconnect from oracle user, and reconnect to load modified $HOME/.profile

16.3 Database Creation On ASM

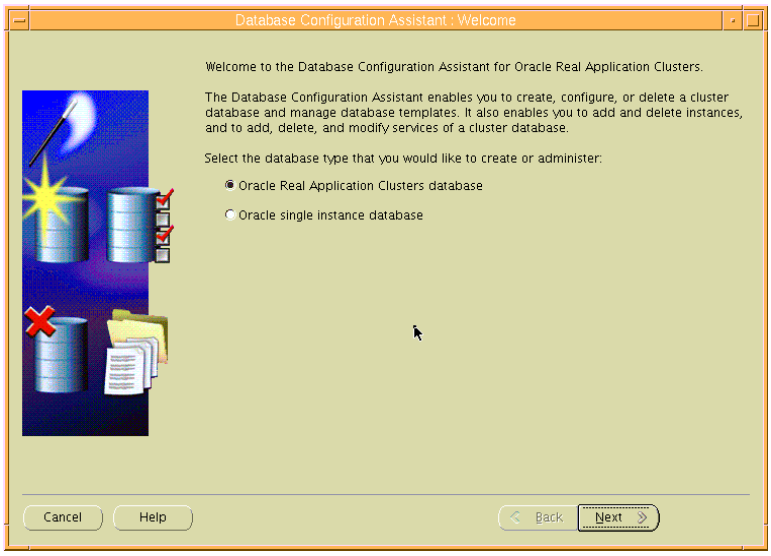


Connect as Oracle unix user from first node, and **setup your DISPLAY**

Execute **dbca &** to lanch the database configuration assistant

From node1 :

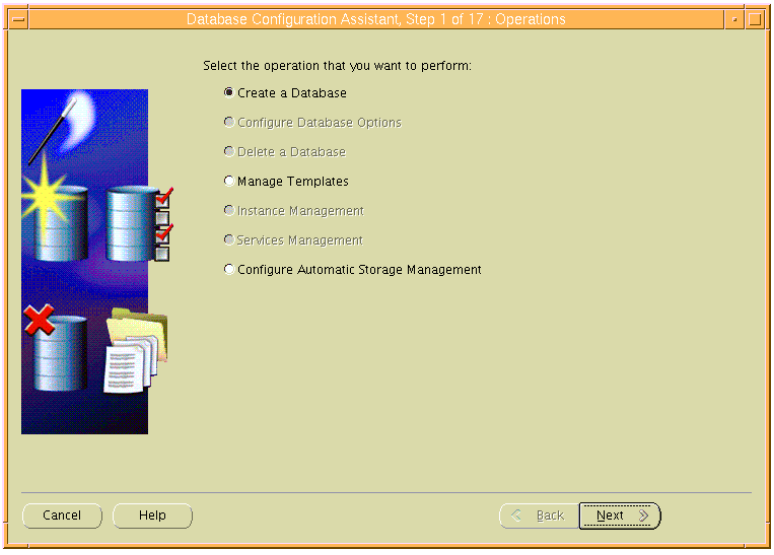
{**node1**:oracle}/ -> **export ORACLE_HOME=/oracle/products/rdbms**
{**node1**:oracle}/ -> **export ORACLE_SID=**
{**node1**:oracle}/ -> **cd $ORACLE_HOME/bin**
{**node1**:oracle}/oracle/products/rdbms ->**./dbca**



DBCA Welcome Screen :

Select the “Oracle Real Application Cluster Database” option.

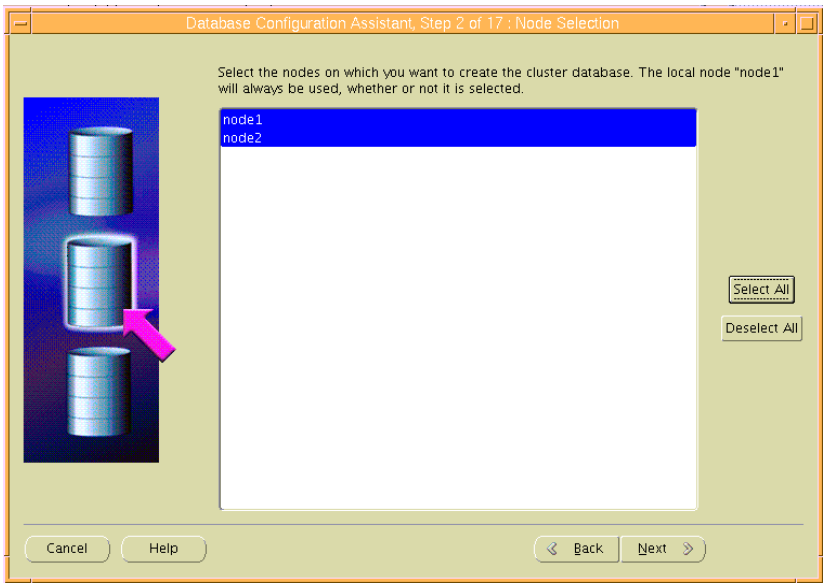
Then click Next ...



Operations :

Select the “Create a Database” option.

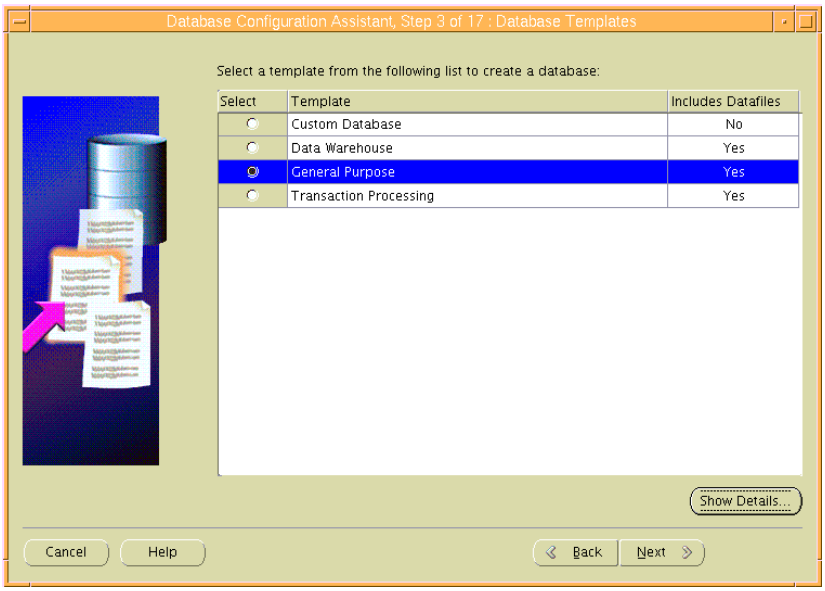
Then click Next ...



Node Selection :

Make sure to select all RAC nodes.

Then click Next ...

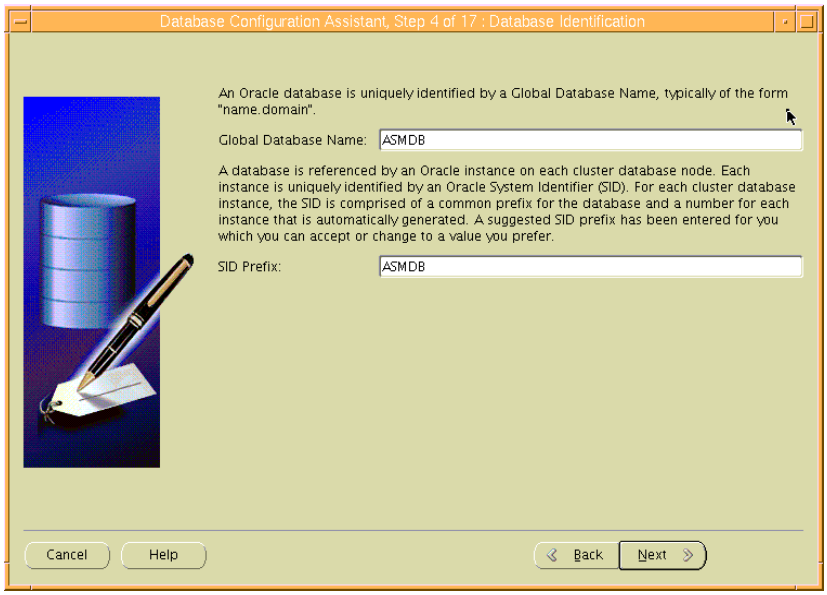


Database Templates :

Select "General Purpose"

Or "Custom Database" if you want to generate the creation scripts.

Then click Next ...

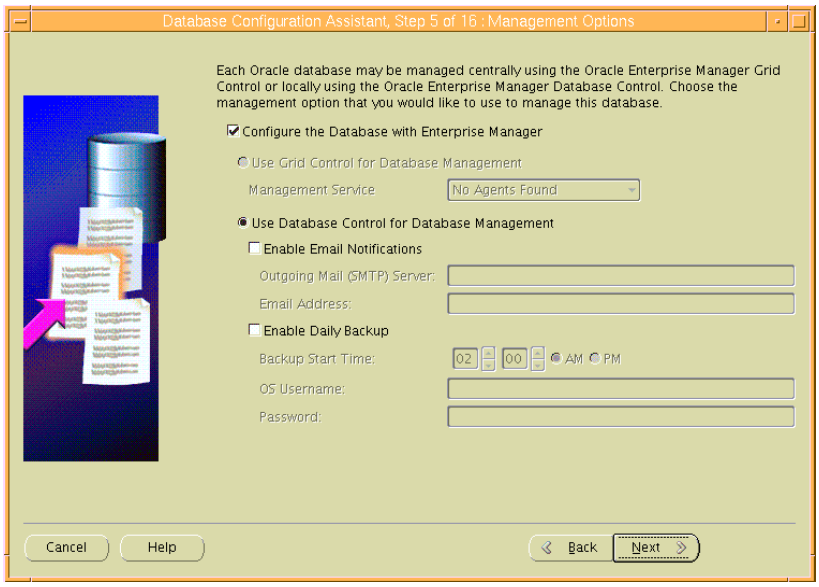


Database Identification :

Specify the "Global Database Name"

The "SID Prefix" will be automatically updated. (by default it is the Global Database Name)

Then click Next ...

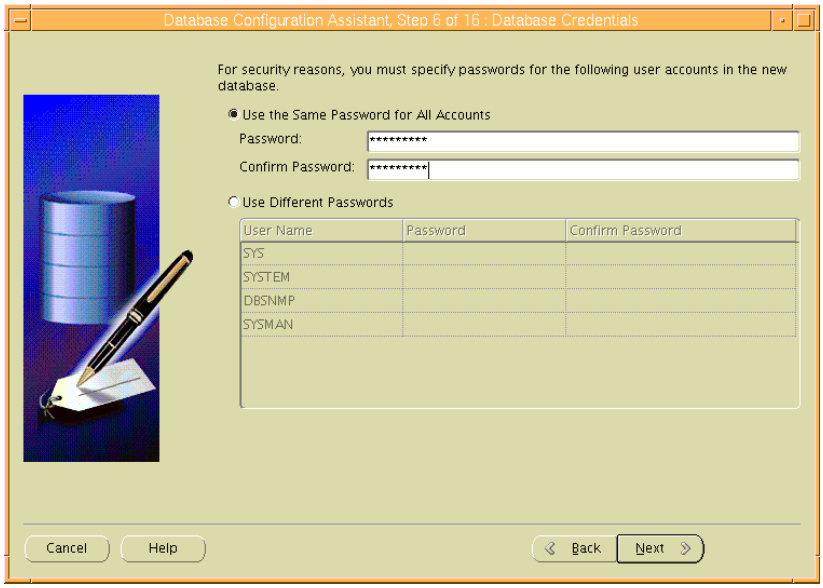


Management Options :

Check "Configure the database with Enterprise Manager" if you want to use the Database Control (local administration).

Or Don't check if you plan to administrate the database using the Grid Control (global network administration)

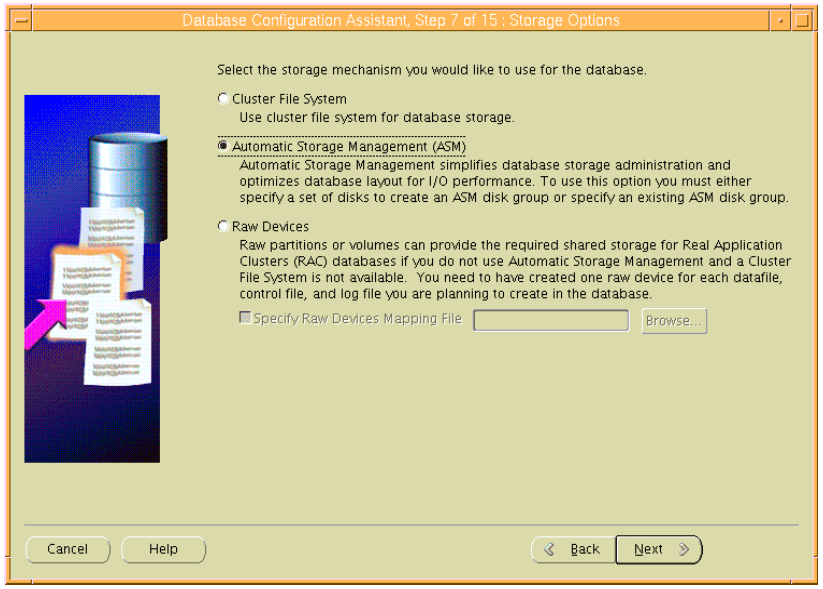
Then click Next ...



Database Credentials :

Specify same password for all administrator users,
or specify individual password for each user.

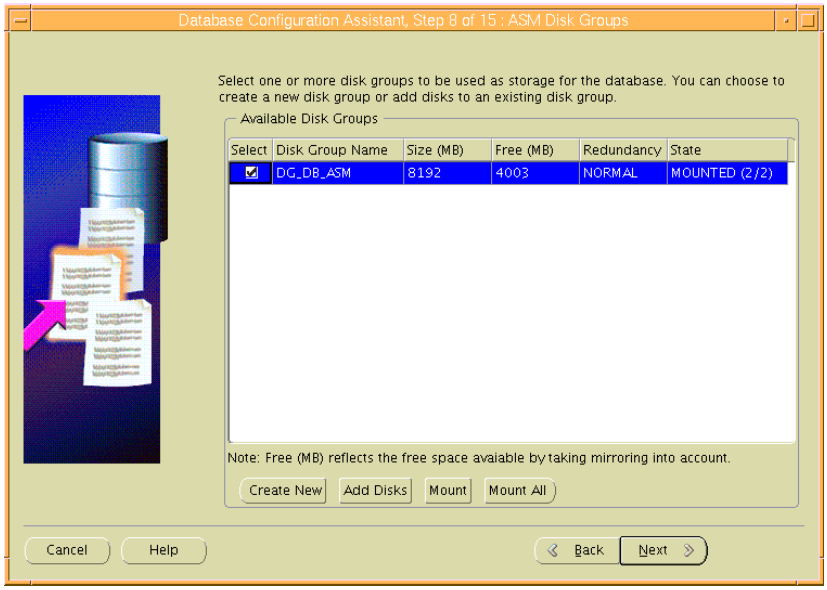
Then click Next ...



Storage Options :

Choose Automatic Storage Management (ASM)

Then click Next ...

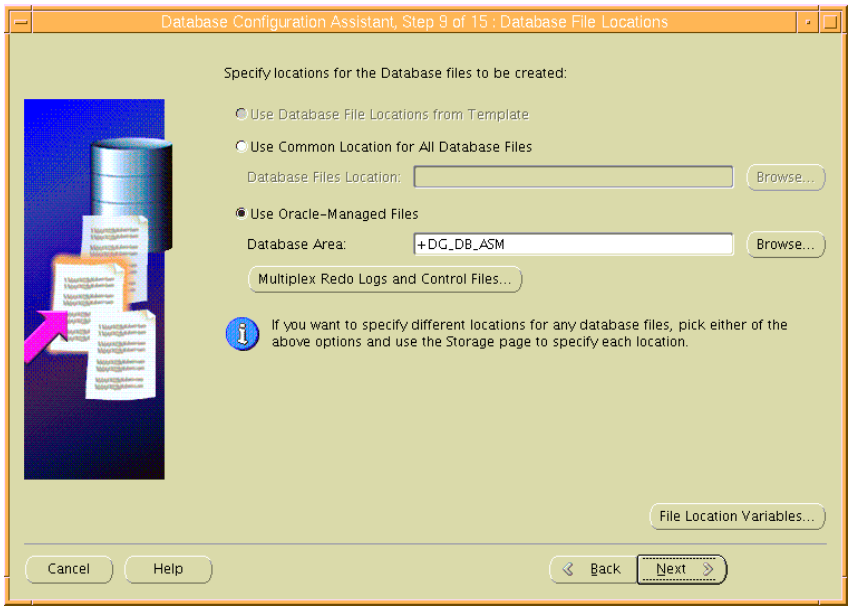


Create Disk Group

Now the ASM Disks Group is created

Select DiskGroup to be used !!!

Then click Ok ...

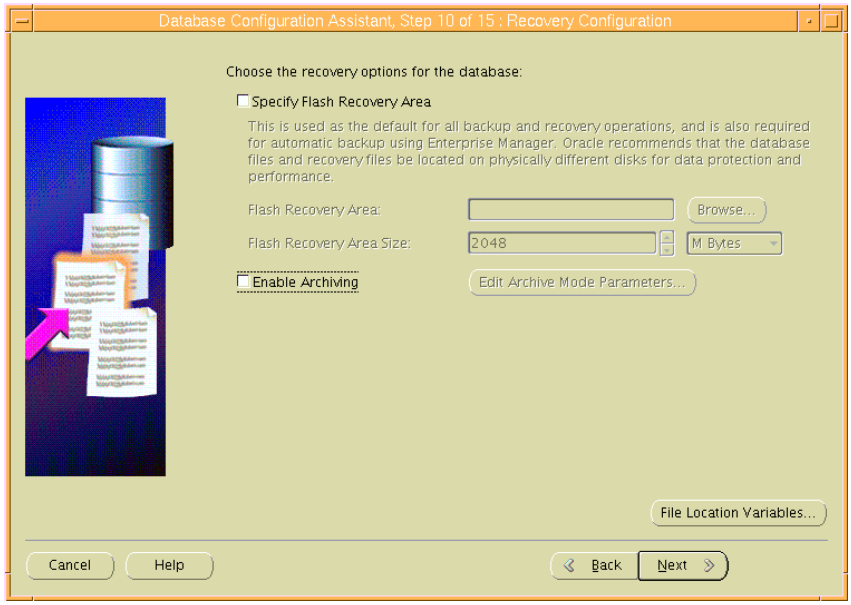


Database File Locations :

Select Use Oracle-Managed Files

AND Select DiskGroup to use for the Database Files.

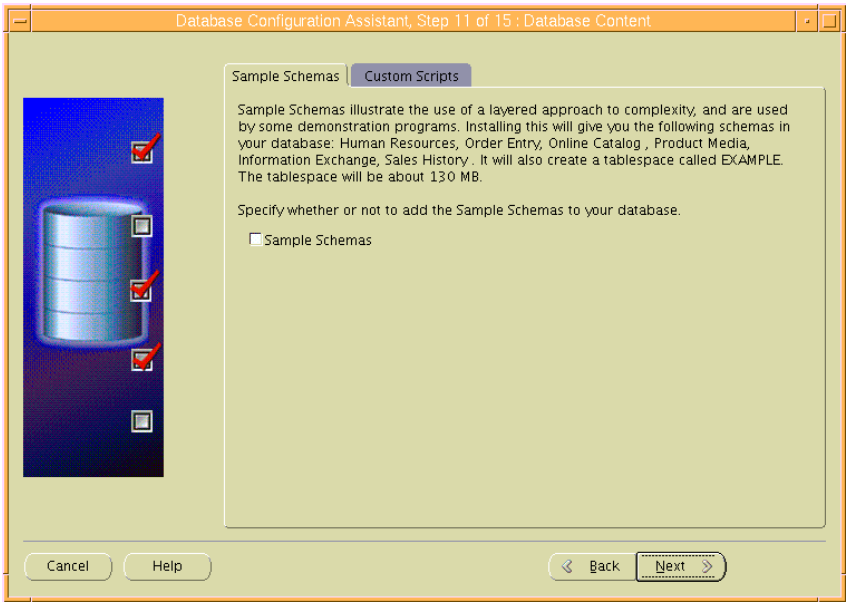
Then click Next ...



Recovery Configuration :

Select the DiskGroup to be used for the Flash Recovery Area.

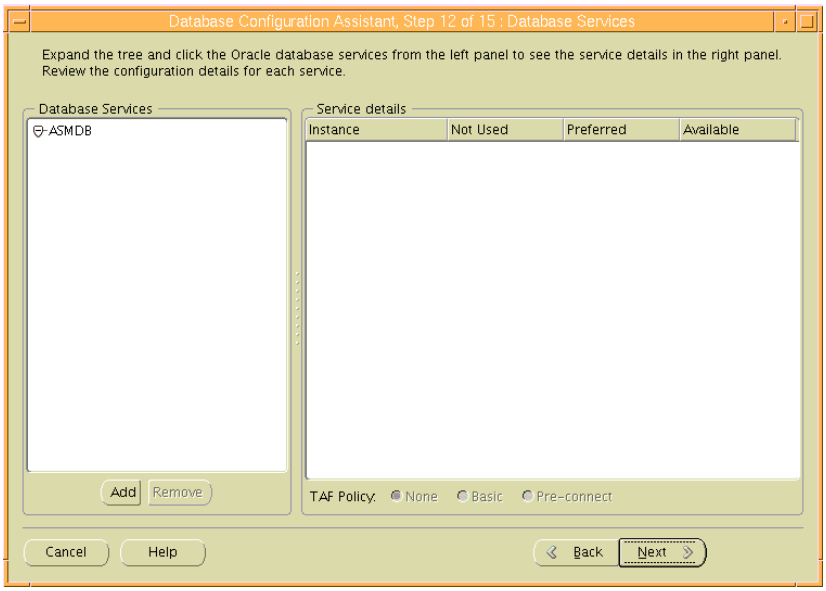
Then click Next ...



Database Content :

Select the options needed

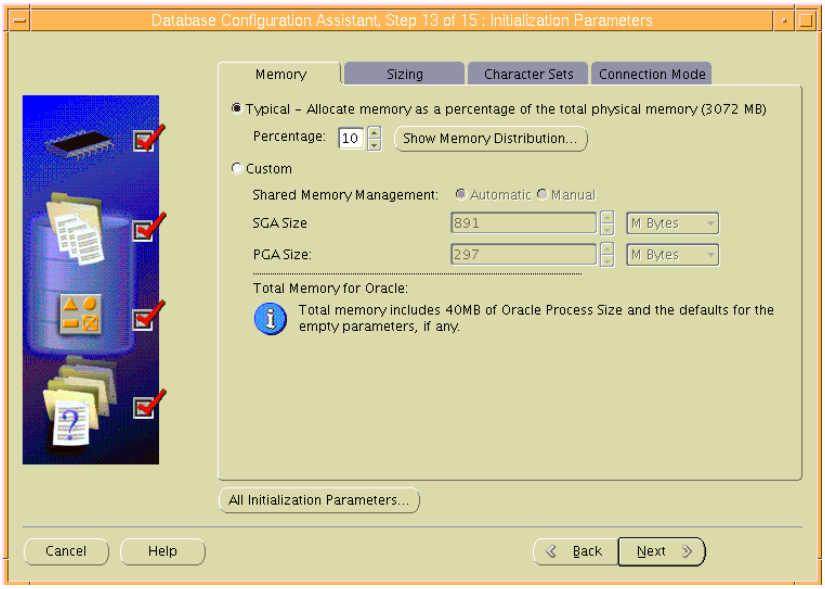
Then click Next ...



Database Services and TAF (Transaction Application Failover) policy :

Configure if necessary !!!

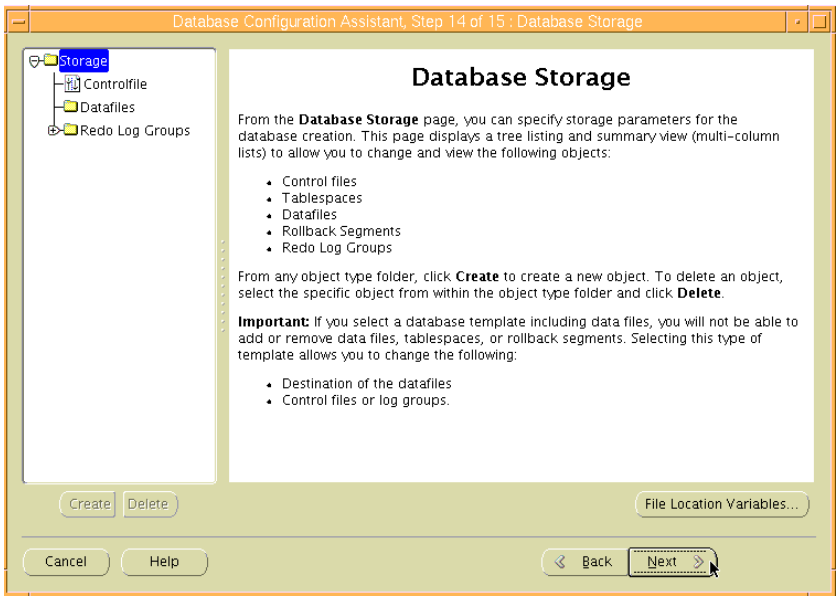
Then click Next ...



Initialization Parameters :

Select the parameters needed

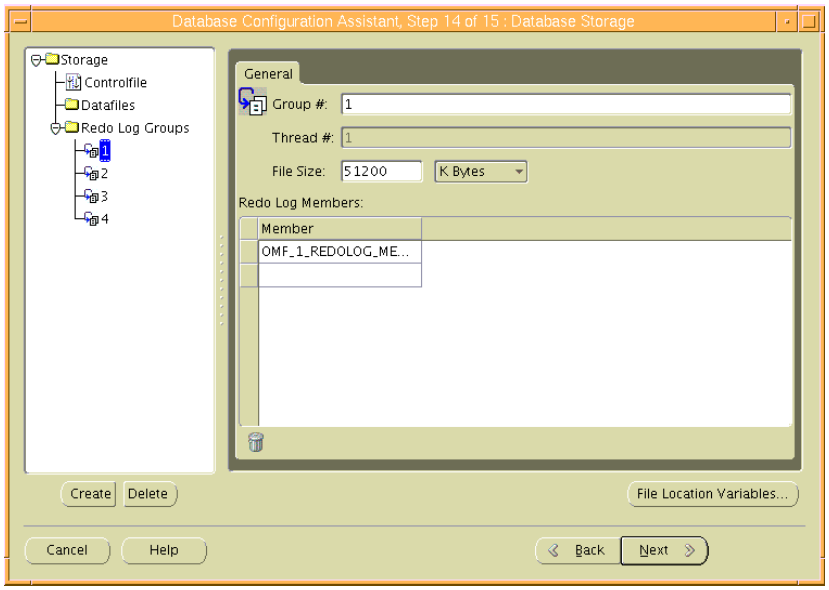
Then click Next ...

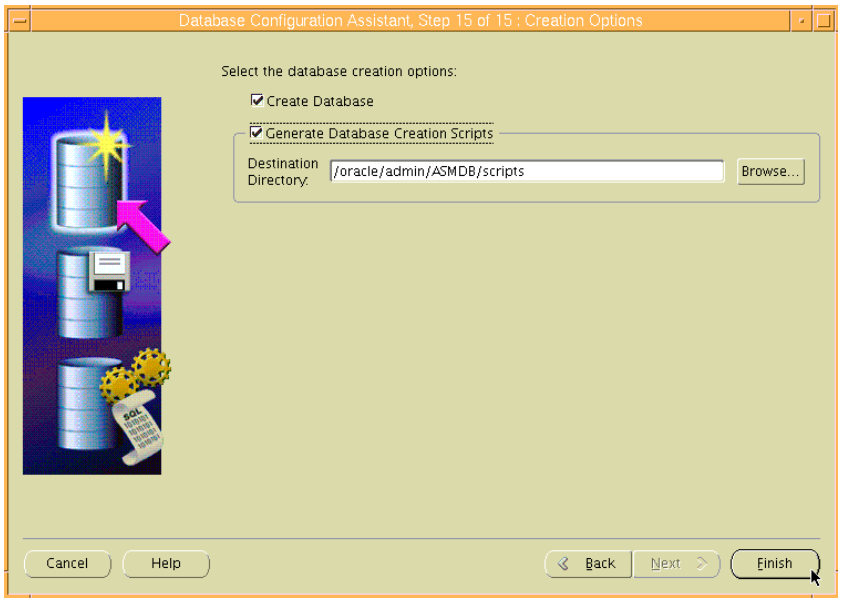


Database Storage :

Check the datafiles organization

Then click Next ...



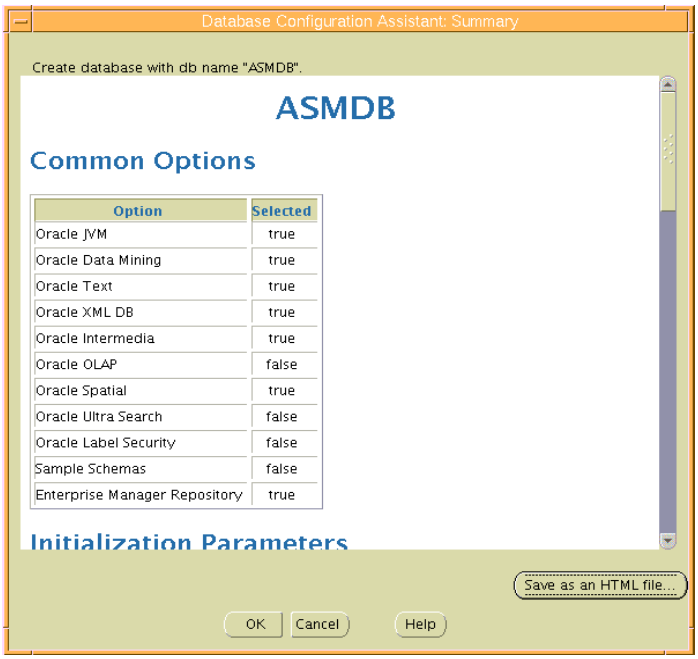


Creation Options :

Select the options needed

- Create Database
- Generate Database Creation Scripts

Then click Finish ...

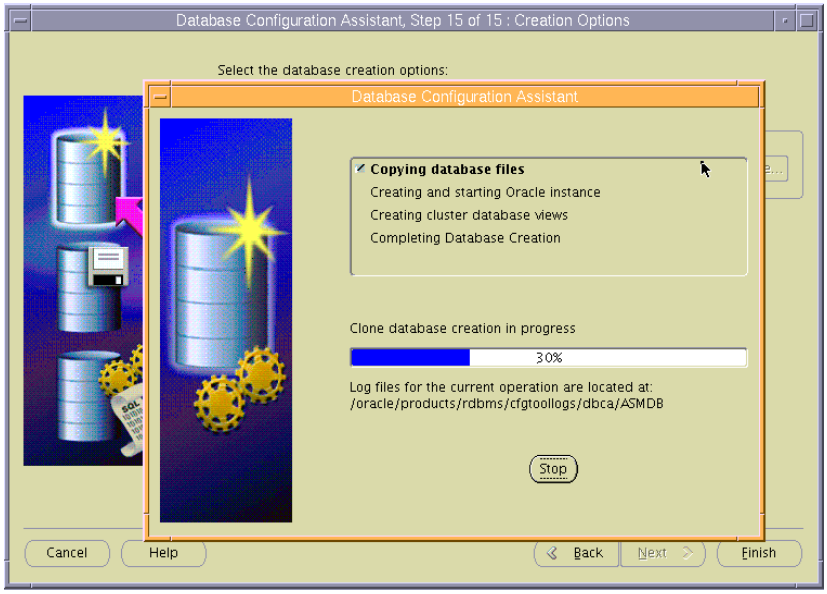


Summary :

Check the description

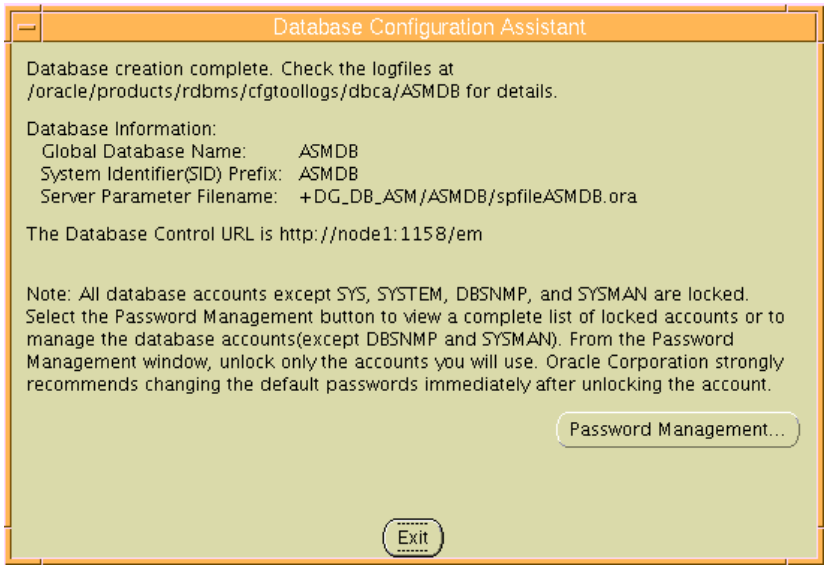
Save the HTML summary file if needed

Then click Ok ...



Database creation on progress :

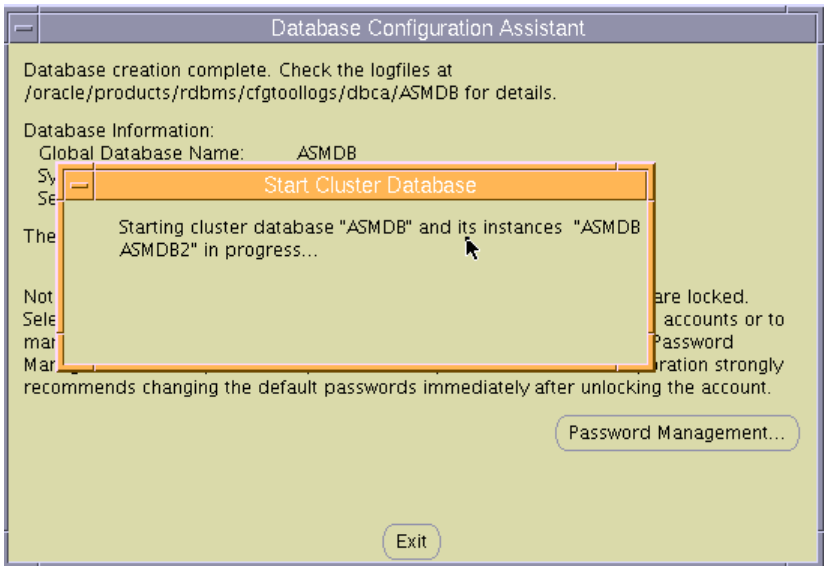
Just wait while processing ...



Passwords Management

Enter in password management, if you need to change pasword, and unlock some user accounts that are locked by default (for security purpose).

Then click Exit ...



Starting cluster database

Check ASMDB spfile or initfile content

For example on node1 :

```
ASMDB1.__db_cache_size=46137344
ASMDB2.__db_cache_size=58720256
ASMDB2.__java_pool_size=4194304
ASMDB1.__java_pool_size=4194304
ASMDB2.__large_pool_size=4194304
ASMDB1.__large_pool_size=4194304
ASMDB1.__shared_pool_size=104857600
ASMDB2.__shared_pool_size=92274688
ASMDB2.__streams_pool_size=0
ASMDB1.__streams_pool_size=0
*.audit_file_dest='/oracle/admin/ASMDB/adump'
*.background_dump_dest='/oracle/admin/ASMDB/bdump'
*.cluster_database_instances=2
*.cluster_database=true
*.compatible='10.2.0.3.0'
*.control_files='+DG_DB_ASM/asmdb/controlfile/current.262.617810231'
*.core_dump_dest='/oracle/admin/ASMDB/cdump'
*.db_block_size=8192
*.db_create_file_dest='+DG_DB_ASM'
*.db_domain=''
*.db_file_multiblock_read_count=16
*.db_name='ASMDB'
*.dispatchers='(PROTOCOL=TCP) (SERVICE=ASMDBXDB)'
ASMDB1.instance_number=1
ASMDB2.instance_number=2
*.job_queue_processes=10
*.open_cursors=300
*.pga_aggregate_target=16777216
*.processes=150
*.remote_listener='LISTENERS_ASMDB'
*.remote_login_passwordfile='exclusive'
*.sga_target=167772160
ASMDB2.thread=2
ASMDB1.thread=1
*.undo_management='AUTO'
ASMDB1.undo_tablespace='UNDOTBS1'
ASMDB2.undo_tablespace='UNDOTBS2'
*.user_dump_dest='/oracle/admin/ASMDB/udump'
```

16.4 Database Post-Configuration task

16.4.1 Cluster Ready Services Health Check

Execute

crs_stat –t

on one node as oracle user :

```
{node1:oracle}/oracle ->crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora....B1.inst application    ONLINE    ONLINE    node1
ora....B2.inst application    ONLINE    ONLINE    node2
ora.ASMDB.db   application    ONLINE    ONLINE    node1
ora....SM1.asm application    ONLINE    ONLINE    node1
ora....E1.lsnr application    ONLINE    ONLINE    node1
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora....SM2.asm application    ONLINE    ONLINE    node2
ora....E2.lsnr application    ONLINE    ONLINE    node2
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node1:oracle}/oracle ->
```

CR health check

```
{node1:oracle}/oracle/crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oracle/crs/bin ->crsctl check  crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

16.5 What has been done ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**
- **10.2.0.3 patchset is applied for CRS home**
- **ASM Home is installed**
- **ASM instance is created and started**
- **ASM Diskgroup is created**
- **10.2.0.3 pachset is applied for ASM home**
- **RDBMS home is installed**
- **One DB has been created**

16.6 Install 10.2.0.3 Patchset to RDBMS Software

Stop database ASMDB on all nodes, and any listener started from the DB ORACLE_HOME.

On each node :

Stop any extra created listener linked to the rdbms ORACLE_HOME

Stop listener on all nodes. To stop listener running on a node, enter the following command where *node* is the name of the node where the applications are running:

As oracle user, execute :

$ *oracle_home*/bin/srvctl stop listener -n *node*

Get the right listener name linked to the rdbms ORACLE_HOME, if any created.

```
{node1:root}/oracle -> srvctl stop listener... -n node1
{node1:root}/oracle ->
{node1:root}/oracle -> srvctl stop listener... -n node2
{node1:root}/oracle ->
```

On each node :

Stop the Database instances

Stop DATABASE instances on all nodes. To stop DATABASE instances running on a node, enter the following command where *node* is the name of the node where the applications are running:

As oracle user, execute :

$ *oracle_home*/bin/srvctl stop database –d *databaseName*
The command will stop all instances from this database.

```
{node1:root}/oracle -> srvctl stop database -d ASMDB
{node1:root}/oracle ->
```

On each node :

Then running "crs_stat –t" as oracle, or root on node1 should return the following result :

```
{node1:root}/oracle -> crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora....B1.inst application    OFFLINE   OFFLINE
ora....B2.inst application    OFFLINE   OFFLINE
ora.ASMDB.db   application    OFFLINE   OFFLINE
ora....SM1.asm application    ONLINE    ONLINE    node1
ora....E1.lsnr application    ONLINE    ONLINE    node1
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora....SM2.asm application    ONLINE    ONLINE    node2
ora....E2.lsnr application    ONLINE    ONLINE    node2
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node1:oracle}/oracle ->
```

Back Up the System, Oracle recommends that you create a backup of the Oracle Clusterware 10*g* installation before you install the patch set, and a backup of the OCR content.

On each node :

Run the AIX command "/usr/sbin/slibclean" as "root" to clean all unreferenced libraries from memory !!!

```
{node1:root}/oracle/products -> /usr/sbin/slibclean
{node1:root}/oracle/products ->

{node2:root}/oracle/products -> /usr/sbin/slibclean
{node2:root}/oracle/products ->
```

From first node As root user, execute :

Under VNC Client session, or other graphical interface, execute :

```
{node1:root}xhost +
access control disabled, clients can connect from any hosts
{node1:root}
```

Login as oracle and follow the procedure hereunder…

Setup and export your DISPLAY, TMP and TEMP variables

With /tmp or other destination having enough free space, about 500Mb on each node.

```
{node1:oracle}/ -> export DISPLAY=node1:1
{node1:oracle}/ -> export TMP=/tmp
{node1:oracle}/ -> export TEMP=/tmp
{node1:oracle}/ -> export TMPDIR=/tmp
```

IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

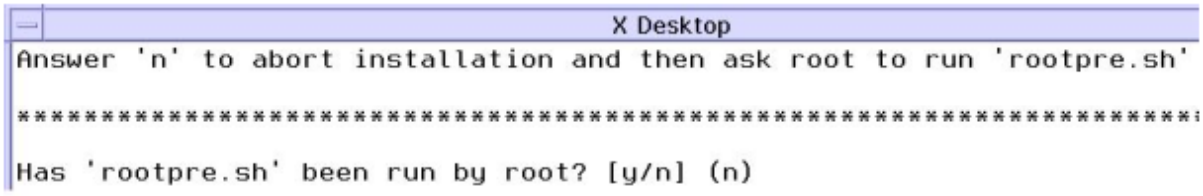
update entries AIX5200 to AIX5300 on both files, and execute :

$/<cdrom_mount_point>/runInstaller

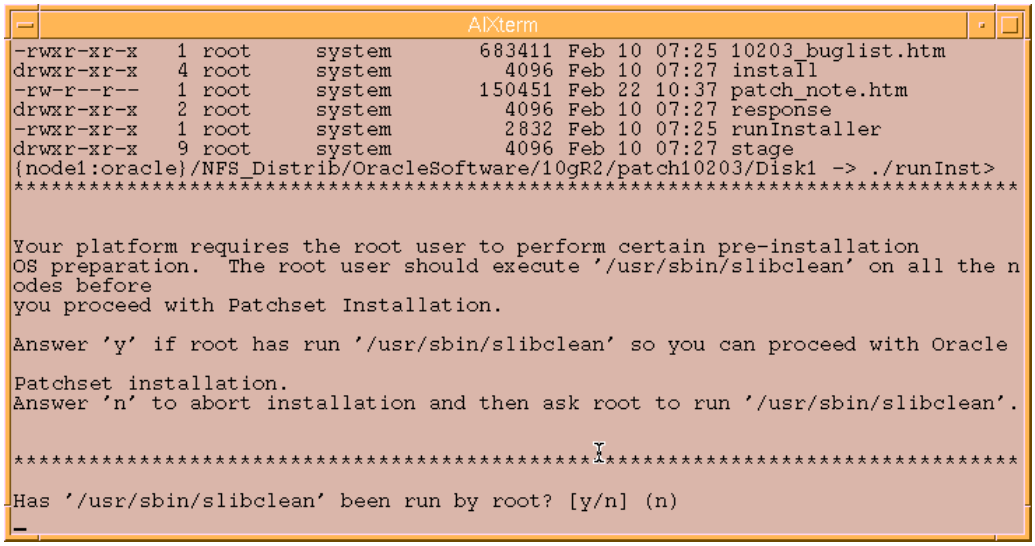
Or execute : ./runInstaller -ignoreSysPreregs

OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX mantenance level installed on each node :
-> **instfix –i|grep ML**
All filesets for 5.3.0.0_AIX_ML were found.
All filesets for 5300-01_AIX_ML were found.
All filesets for 5300-02_AIX_ML were found.
All filesets for 5300-03_AIX_ML were found.
All filesets for 5300-04_AIX_ML were found.
All filesets for 5300-05_AIX_ML were found.



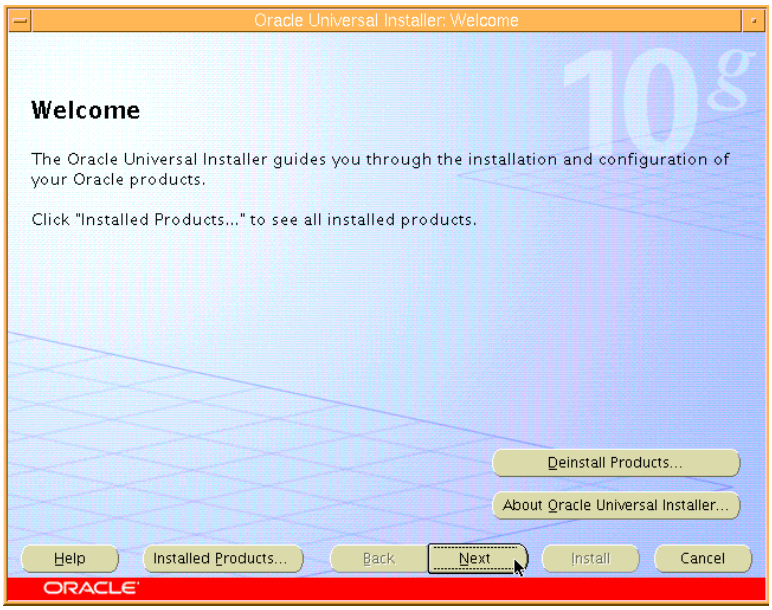
Should be already done with the CRS Installation), Answer “y” for next step



At the OUI Welcome screen

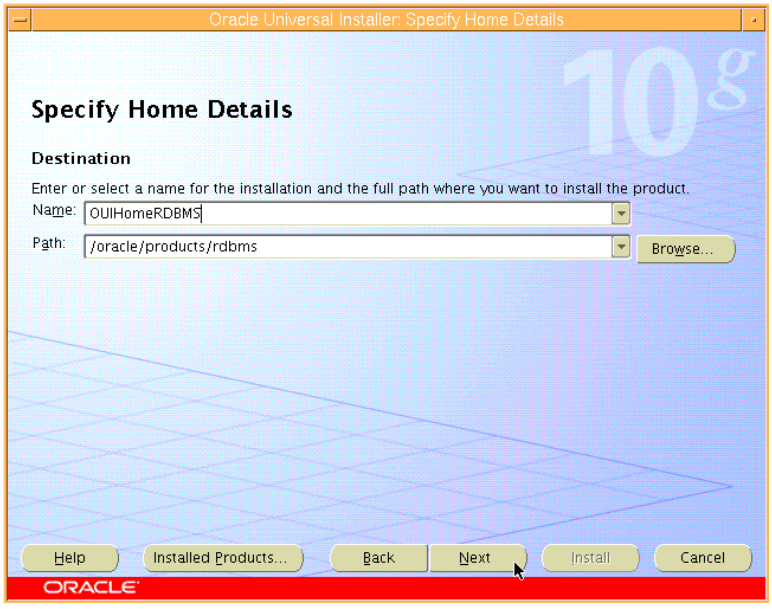
Enter "Yes" if /usr/sbin/slibclean has been executed on both node as root.

Just click Next ...



At the OUI Welcome screen

Just click Next ...

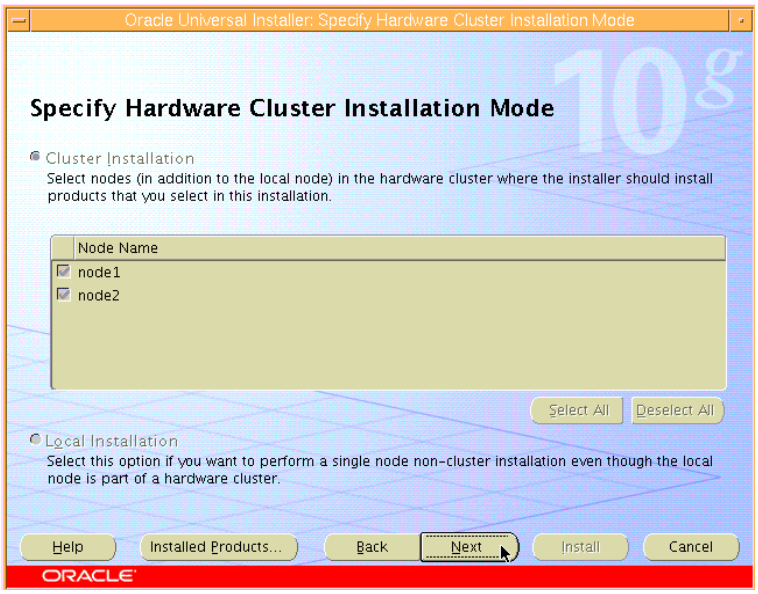


Select the ORACLE_HOME corresponding to the RDBMS_HOME Installation directory !!!

For our case :

OUIHomeRDBMS
/oracle/products/RDBMS

Just click Next ...

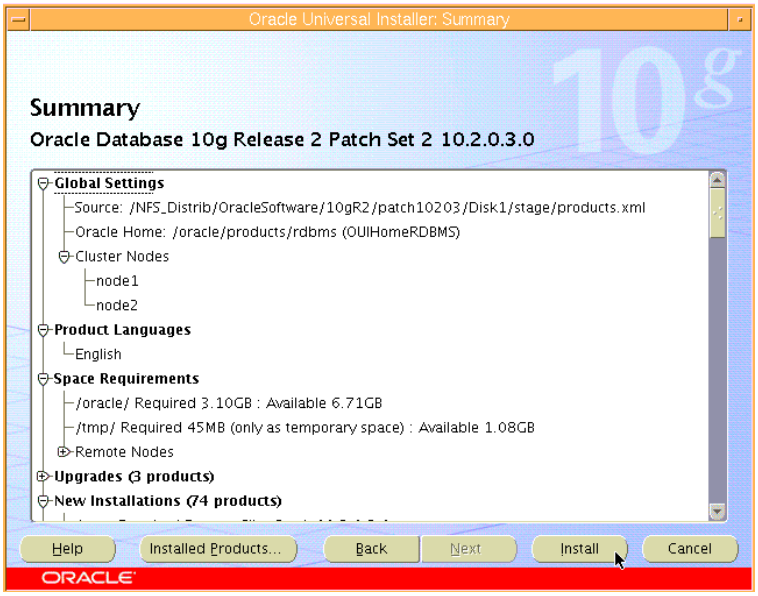


At this stage :

No choice available for local Installation.

Cluster Installation with all nodes selected by default.

Just click Next ...

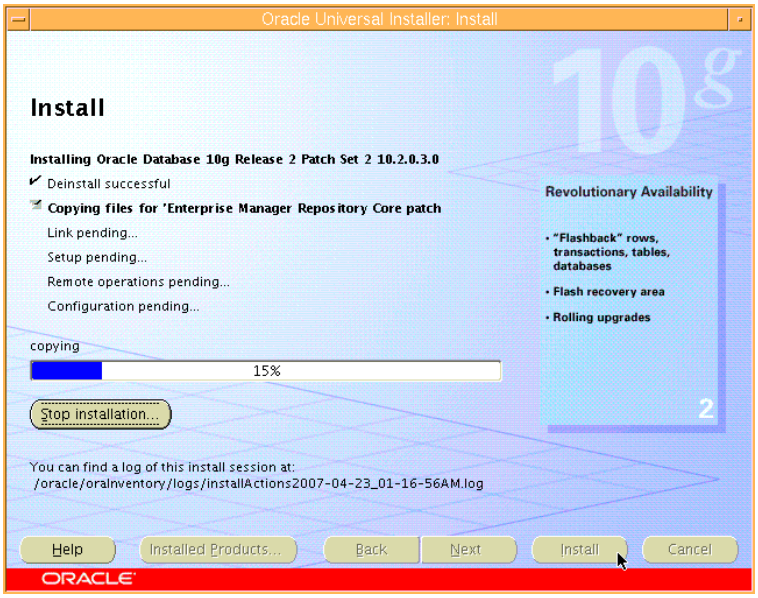


Summary :

The Summary screen will be presented.

Check Cluster Nodes and Remote Nodes lists.
The OUI will install the Oracle 10g 10.2.0.3 software update on to the local node, and then copy this information to the other selected nodes.

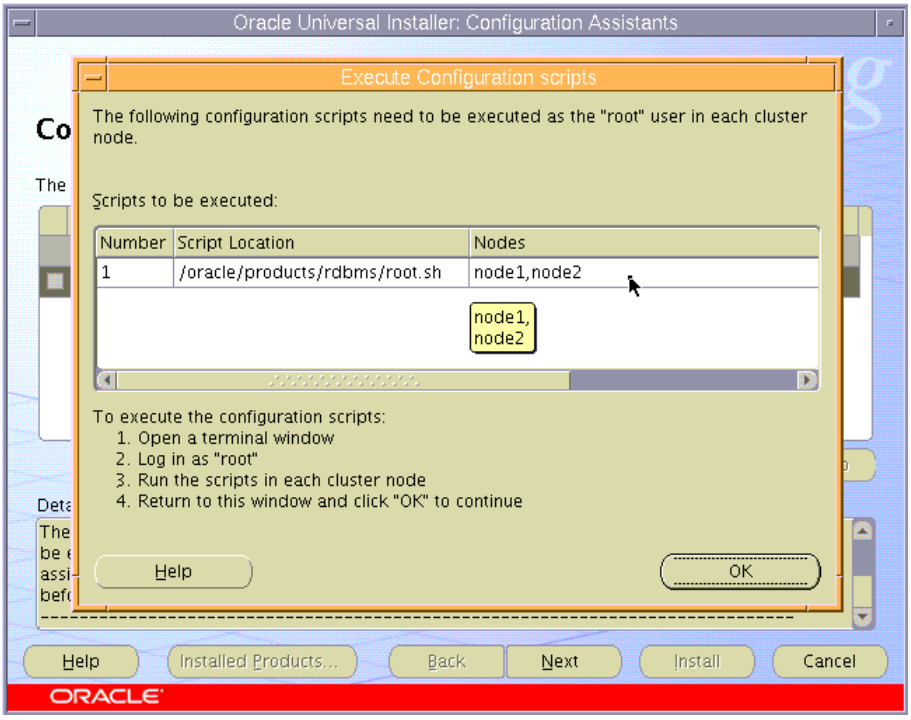
Then click Install ...



Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the others selected nodes.

Just wait for the next screen ...



Execute Configuration Scripts will pop-up :

AS root, execute root.sh on each node.

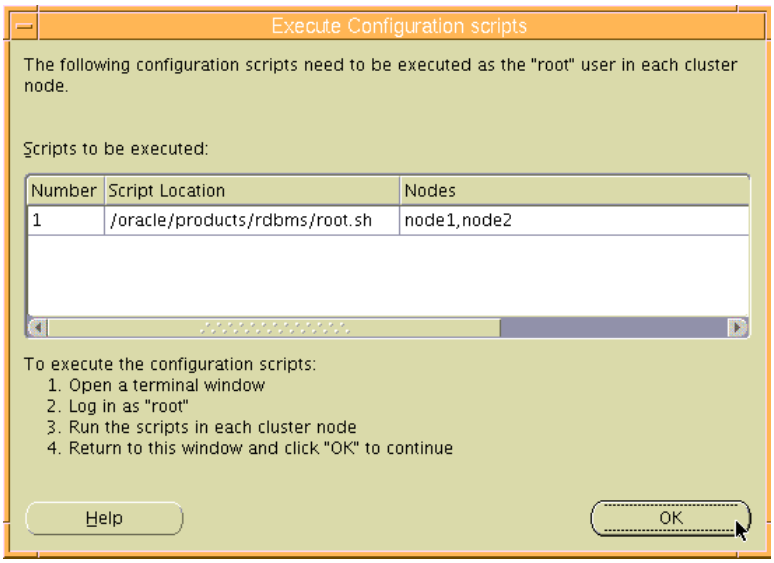
For our case, this script is located in the /oracle/products/asm

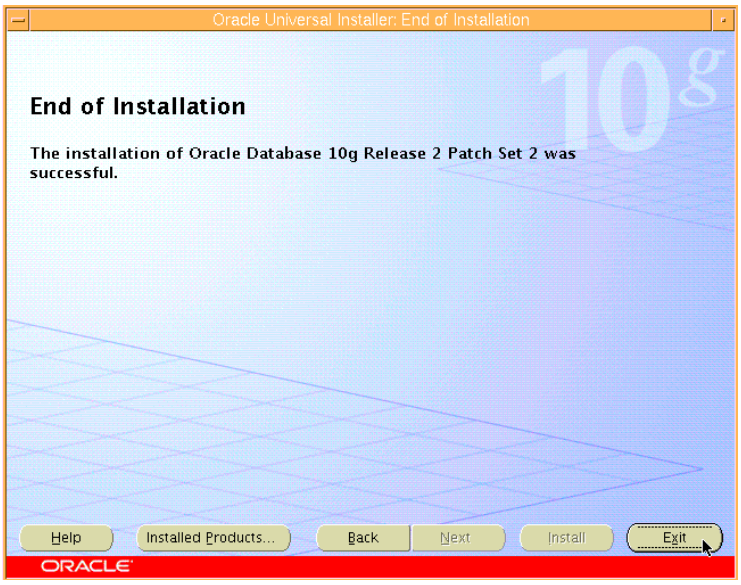
Just click OK ...

```
{node1:root}/oracle/products/rdbms -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/rdbms

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node1:root}/oracle/products/rdbms ->

------------------------------------------------------------

{node2:root}/oracle/products/rdbms -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /oracle/products/rdbms

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

{node2:root}/oracle/products/rdbms ->
```

Coming back to this previous screen,

Just click OK





End of Installation :

This screen will automatically appear.

Check that it is successful and write down the URL list of the J2EE applications that have been deployed (isqlplus, …).

Then click Exit ...

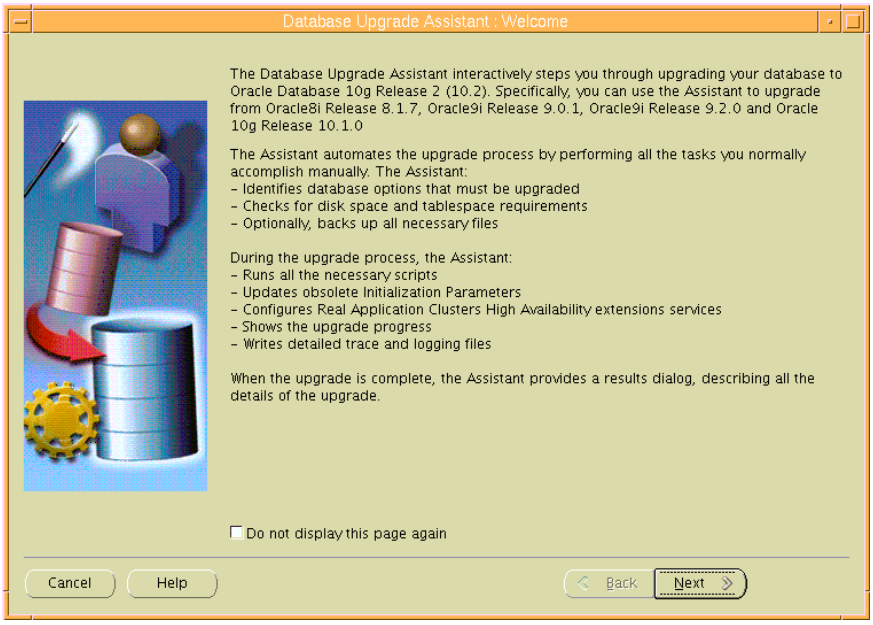
16.7 UPGRADING THE EXISTING DATABASE TO 10.2.0.3

Connect as Oracle unix user from first node, and **setup your DISPLAY**

Execute **dbua &** to lanch the database configuration assistant

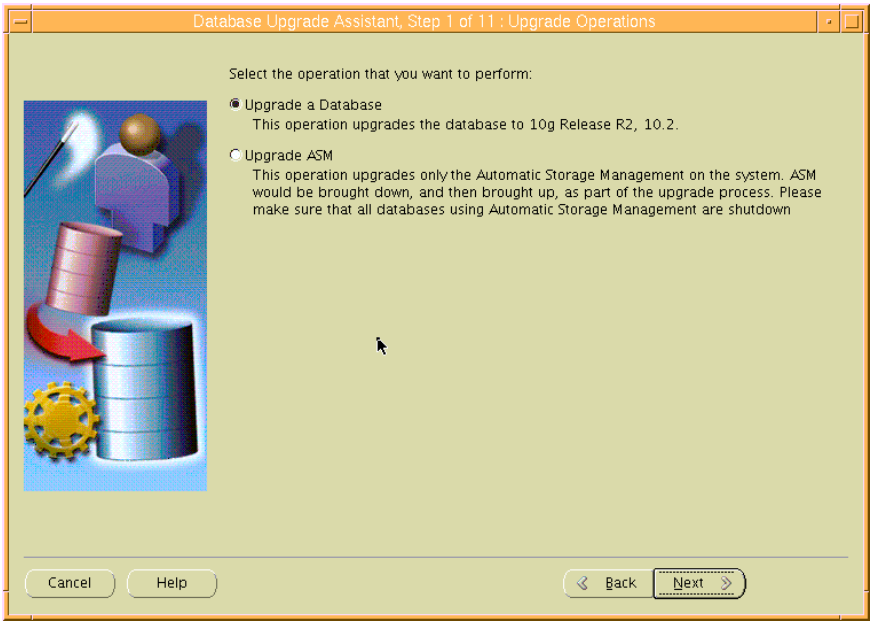
From node1 :

```
export ORACLE_HOME=/oracle/products/rdbms
export ORACLE_SID=ASMDB1
cd $ORACLE_HOME/bin
./dbua
```



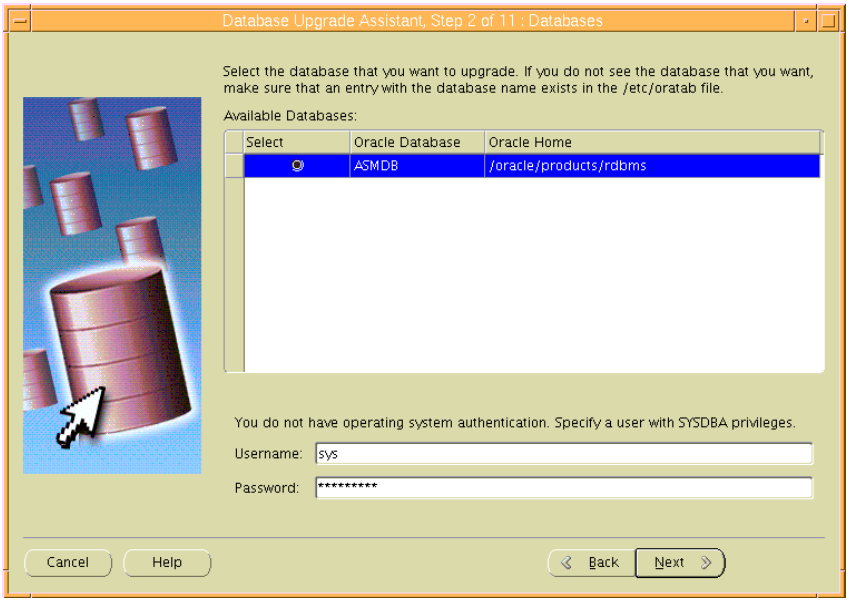
At the DBUA Welcome screen

Just click Next ...



Select to Upgrade a database

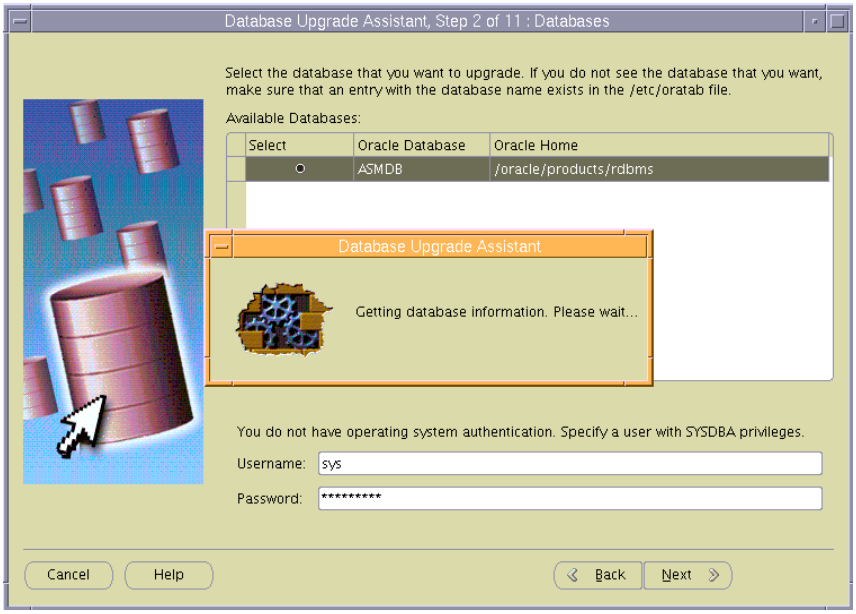
Just click Next ...



At this stage :

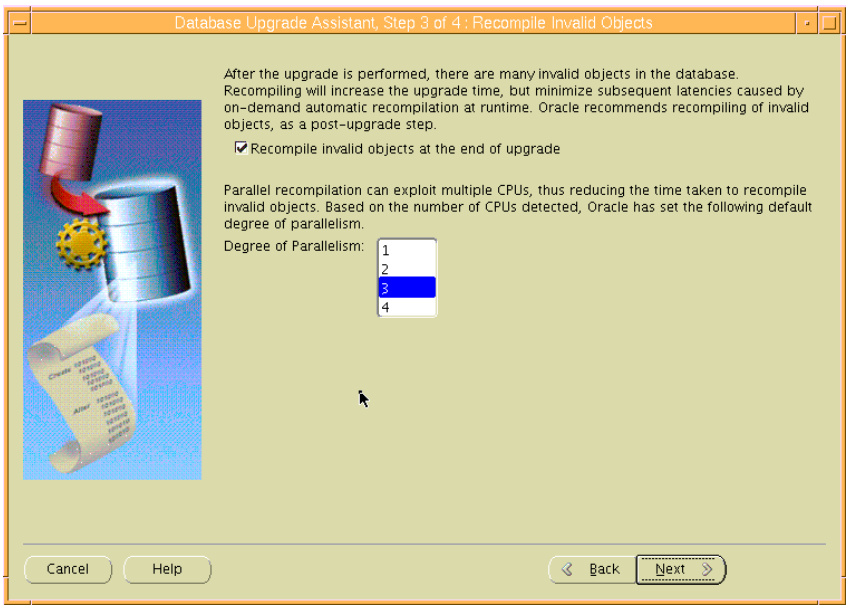
Choose which database you want to migrate, in our case ASMDB
and
provide the password for sys

Just click Next ...



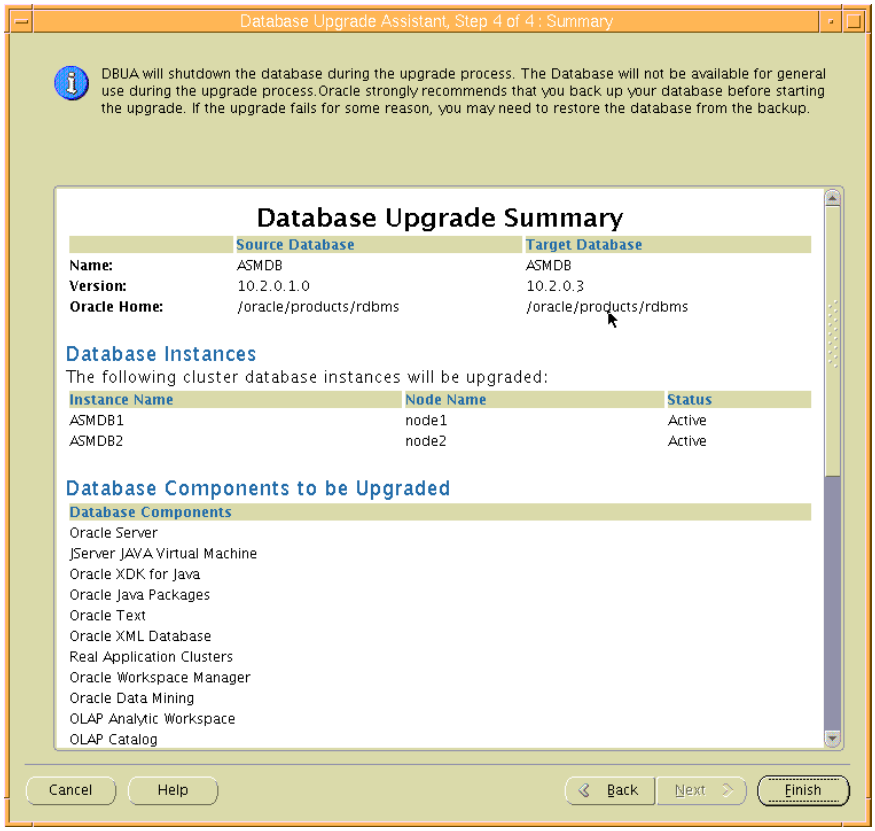
At this stage :

Wait until next screen !



At this stage :

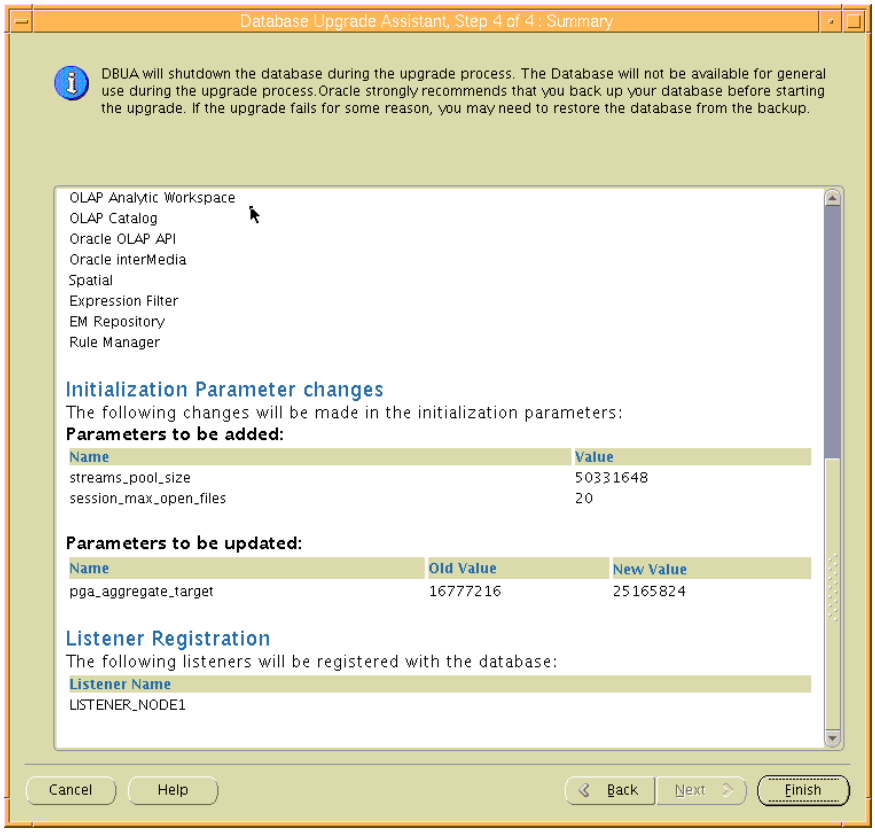
Just click Next ...



Database Upgrade Summary

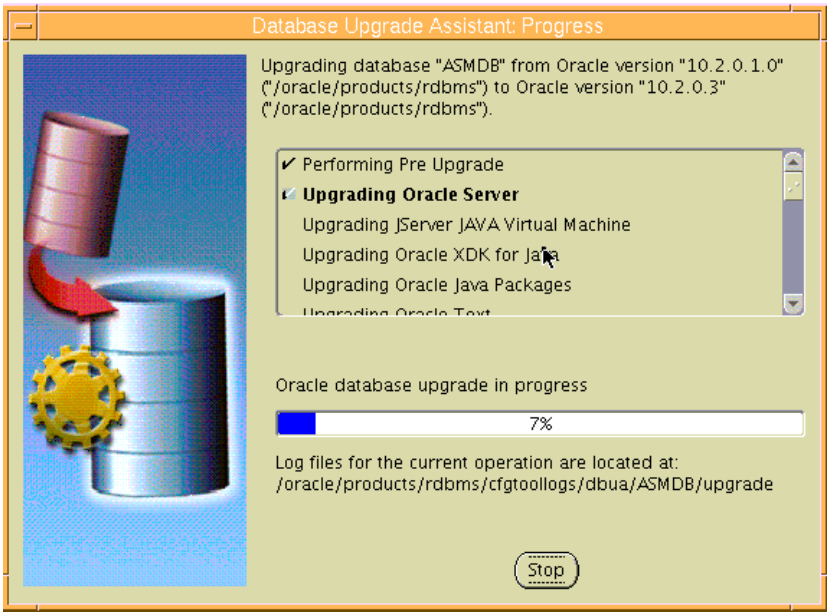
Read the summary, before going forward.

Con't…



Database Upgrade Summary

Just click Finish to start the Database update from 10.2.0.1 to 10.2.0.3 ...



DB migration on progress :

Just wait while processing ...

16.8 WHAT HAS BEEN DONE ?

At this stage :

- **The Oracle Cluster Registry and Voting Disk are created and configured**
- **The Oracle Cluster Ready Services is installed, and started on all nodes.**
- **The VIP (Virtual IP), GSD and ONS application resources are configured on all nodes.**
- **10.2.0.3 patchset is applied for CRS home**
- **ASM Home is installed**
- **ASM instance is created and started**
- **ASM Diskgroup is created**
- **10.2.0.3 pachset is applied for ASM home**
- **RDBMS home is installed**
- **One DB has been created**
- **10.2.0.3 pachset is applied for RDBMS home**
- **This existing DB has been migrated to 10.2.0.3**

17 CREATING ORACLE SERVICES

17.1 CREATION THRU SRVCTL COMMAND

NEVER Create a service with same name as database !!!

Subject: **Cannot manage service with srvctl, when it's created with same name as Database: PRKO-2120**
Doc ID: **Note:362645.1**

Details of command to add oracle services thru srvctl :

srvctl add service -d <name> -s <service_name> -r <preferred_list> [-a <available_list>] [-P <TAF_policy>] [-u]

- **-d** Database name
- **-s** Service name
- **-a** for services, list of available instances, this list cannot include preferred instances
- **-P** for services, TAF preconnect policy – NONE, BASIC, PRECONNECT
- **-r** for services, list of preferred instances, this list cannot include available instances.
- **-u** updates the preferred or available list for the service to support the specified instance. Only one instance may be specified with the
- **-u** switch. Instances that already support the service should not be included.

Examples for a 4 RAC nodes Cluster.
With a cluster database named ORA.
4 instances named ORA1, ORA2, ORA3 and ORA4.

Add a STD_BATCH service to an existing database with preferred instances (-r) and available instances (-a).
Use basic failover to the available instances.
srvctl add service -d RAC -s STD_BATCH -r ORA1,ORA2 -a ORA3,ORA4

Add a STD_BATCH service to an existing database with preferred instances in list one and available instances in list two. Use preconnect at the available instances.
srvctl add service -d ORACLE -s STD_BATCH -r ORA1,ORA2 -a ORA3,ORA4 -P PRECONNECT

In our case, we want to :

Add an OLTP service to an existing ASMDB database
with ASMDB1 and ASMDB2 as preferred instances (-r)
Using basic failover to the available instances.

As oracle user, from one node :

{**node1**:oracle}/oracle ->**srvctl add service -d ASMDB -s OLTP -r ASMDB1,ASMDB2**

Add a BATCH service to an existing ASMDB database
with ASMDB2 as preferred instances (-r)
and ASMDB1 as available instances (-a).
Using basic failover to the available instances.

As oracle user, from one node :

{**node1**:oracle}/oracle ->**srvctl add service -d ASMDB -s BATCH -r ASMDB2 -a ASMDB1**

```
{node1:oracle}/oracle -> crsstat.sh
HA Resource                                   Target     State
-----------                                   ------     -----
ora.ASMDB.ASMDB1.inst                         ONLINE     ONLINE on node1
ora.ASMDB.ASMDB2.inst                         ONLINE     ONLINE on node2
ora.ASMDB.BATCH.ASMDB2.srv                    ONLINE     ONLINE on node2
ora.ASMDB.BATCH.cs                            ONLINE     ONLINE on node2
ora.ASMDB.OLTP.ASMDB1.srv                     ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB2.srv                     ONLINE     ONLINE on node2
ora.ASMDB.OLTP.cs                             ONLINE     ONLINE on node2
ora.ASMDB.db                                  ONLINE     ONLINE on node1
ora.node1.ASM1.asm                            ONLINE     ONLINE on node1
ora.node1.LISTENER_NODE1.lsnr                 ONLINE     ONLINE on node1
ora.node1.gsd                                 ONLINE     ONLINE on node1
ora.node1.ons                                 ONLINE     ONLINE on node1
ora.node1.vip                                 ONLINE     ONLINE on node1
ora.node2.ASM2.asm                            ONLINE     ONLINE on node2
ora.node2.LISTENER_NODE2.lsnr                 ONLINE     ONLINE on node2
ora.node2.gsd                                 ONLINE     ONLINE on node2
ora.node2.ons                                 ONLINE     ONLINE on node2
ora.node2.vip                                 ONLINE     ONLINE on node2
{node1:oracle}/oracle ->
```

Entries added by the creation of Oracle services at Oracle Clusterware level :

Using crsstat.sh shell script.

```
BATCH =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = BATCH)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )

OLTP =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = OLTP)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )
```

Entries added by the creation of Oracle services in the tnsnames.ora :

For each services created thru dba, grid control, dbconsole or srvctl, one entry is added in the tnsnames.ora file from each node.

All entries are configured for failover (TAF), and load balancing if prefered on at least 2 nodes.

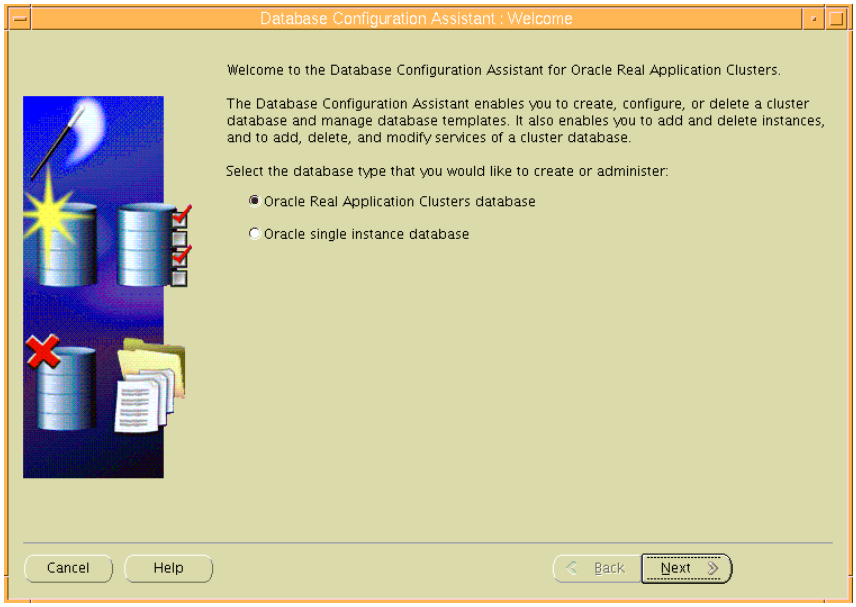
17.2 CREATION THRU DBCA TOOL

In our case, thru dbca we want to :

Add a DISCO service to an existing ASMDB database
with ASMDB1 as preferred instances (-r)
and ASMDB2 as available instances (-a).
Using basic failover to the available instances.

As oracle user, from one node :

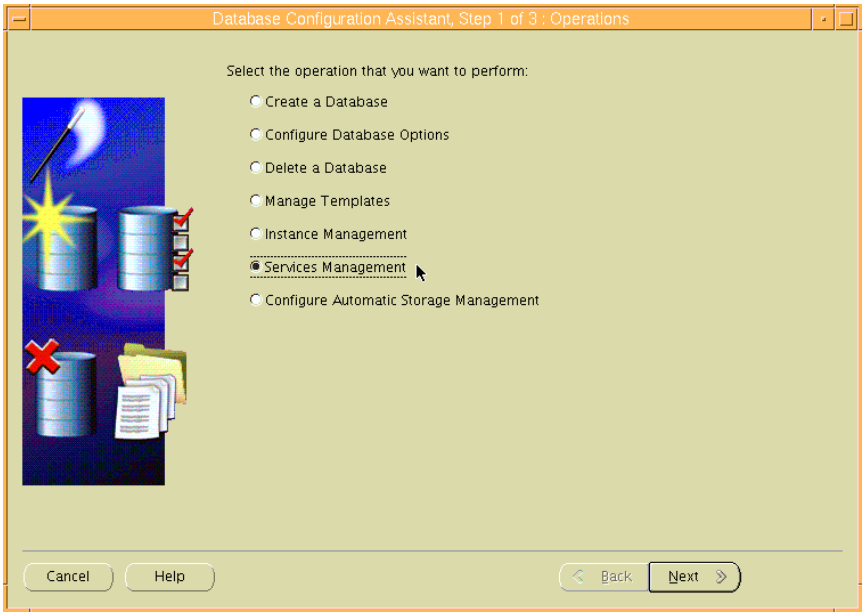
{**node1**:oracle}/oracle -> **export DISPLAY=node1 :1**
{**node1**:oracle}/oracle -> **export ORACLE_HOME=/oracle/products/rdbms**
{**node1**:oracle}/oracle -> **export ORACLE_SID=ASMDB1**
{**node1**:oracle}/oracle -> **cd $ORACLE_HOME/bin**
{**node1**:oracle}/oracle/products/rdbms/bin -> **dbca &**



At the DBCA "Welcome" screen :

Select "Oracle Real Application Clusters database"

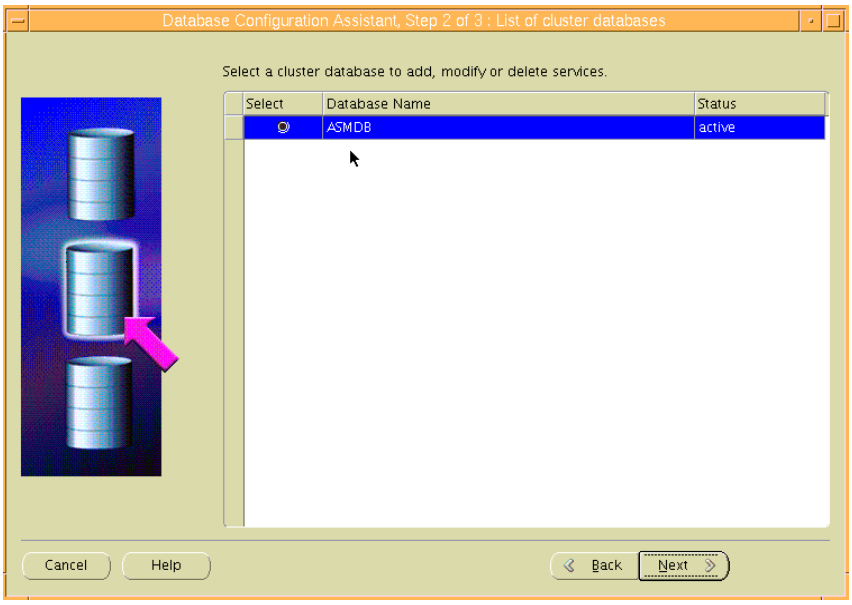
AND click Next ...



At the "Operations" screen :

Select "Services Management"

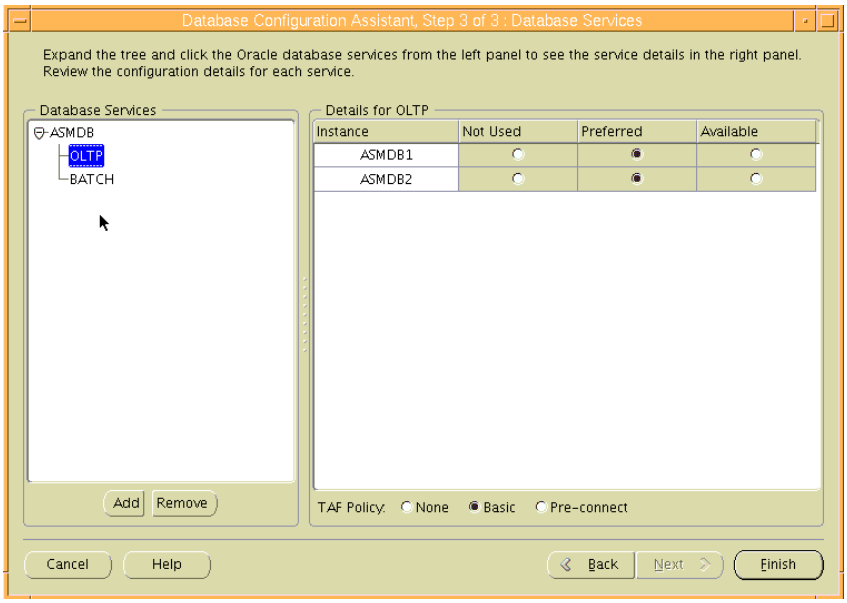
AND click Next ...



At the "list of cluster databases" screen :

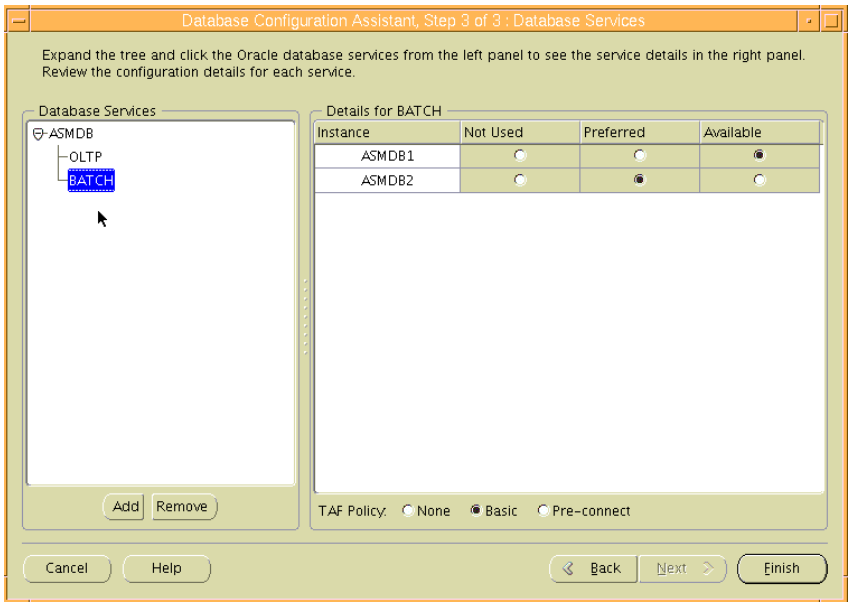
Select the database to create/delete/modify services associated to it.
Select "ASMDB" for our case.

AND click Next ...



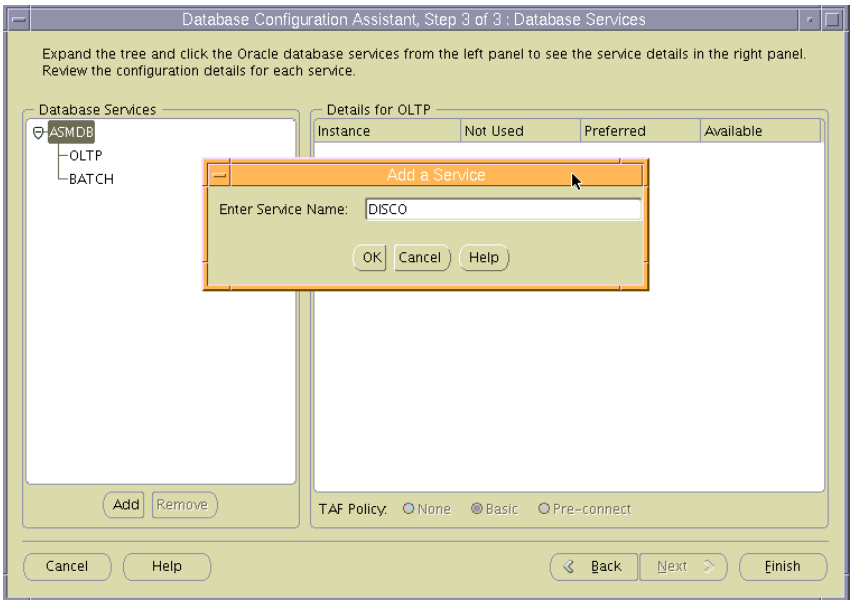
At the "Database Services" screen :

You can view the already defined OLTP service with ASMDB1 and ASMDB2 instances as preferred, with basic TAF.



At the "Database Services" screen :

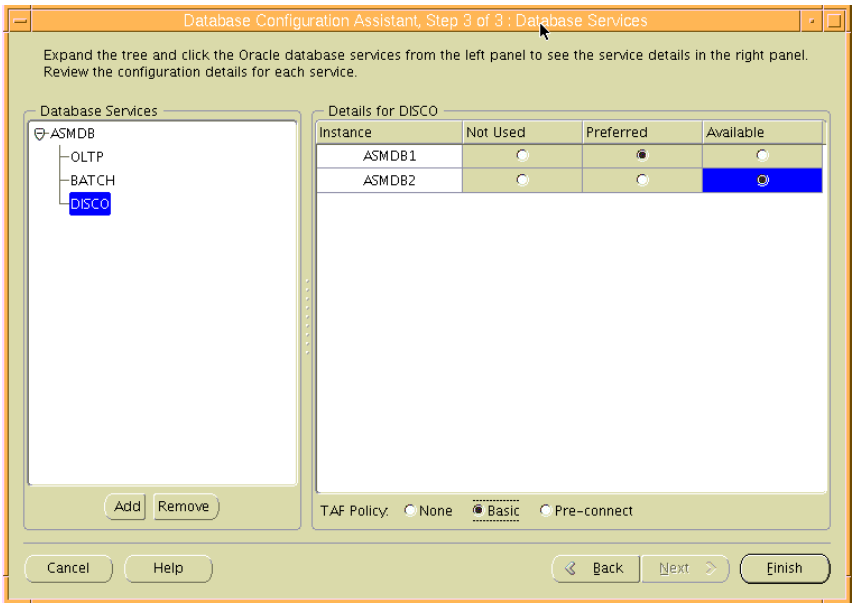
You can view the already defined BATCH service with ASMDB1 as available instance and ASMDB2 as preferred instance, with basic TAF if both instances are switched to preferred.



At the "Database Services" screen :

Click on "Add" to create new service.
Then enter service name.
"DISCO" in our case.

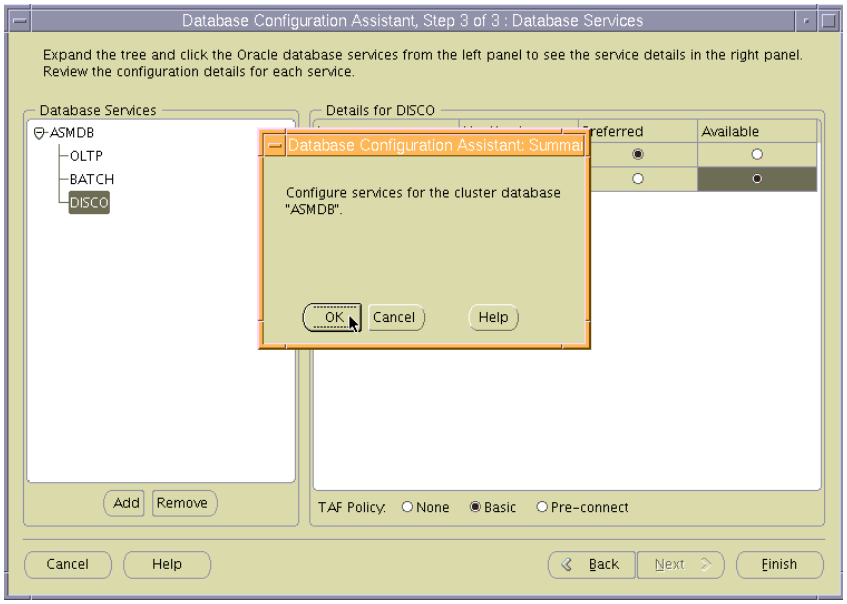
AND click OK ...



At the "Database Services" screen :

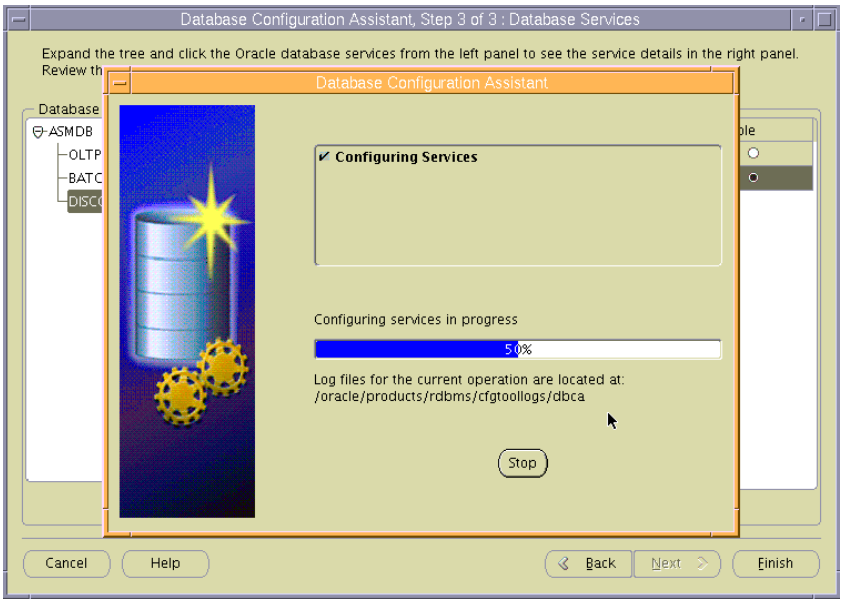
Set ASMDB1 as preferred instance.
Set ASMDB2 as available instance.
Set TAF Policy as "Basic", it could be also "None" as only one instance is preferred at a time.

AND click Finish ...



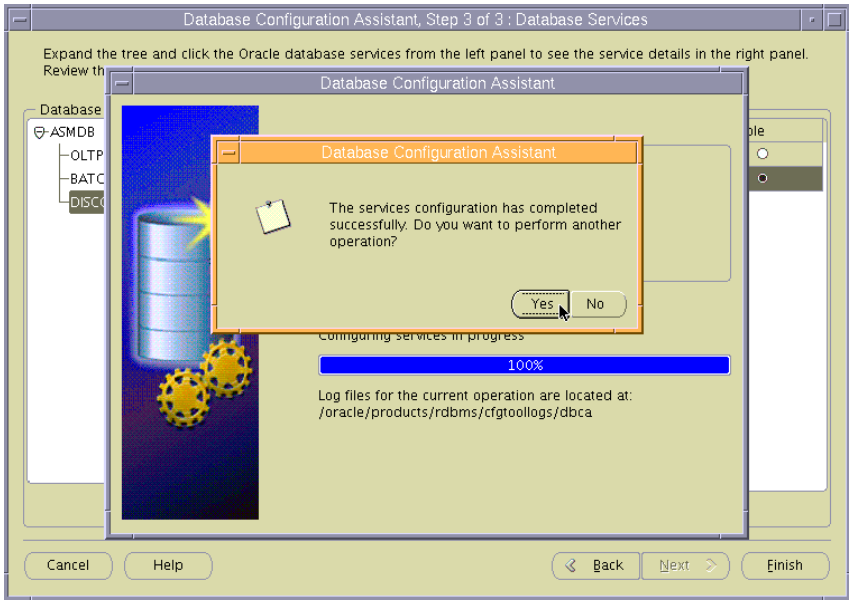
At the "Database Configuration Assistant Summary" screen :

Click "OK" to create the DISCO service.



At this stage :

Just wait while processing ...



At this stage, the service is created :

Just click “NO” to exit ...

```
{node1:oracle}/oracle -> crsstat.sh
HA Resource                                   Target     State
-----------                                   ------     -----
ora.ASMDB.ASMDB1.inst                         ONLINE     ONLINE on node1
ora.ASMDB.ASMDB2.inst                         ONLINE     ONLINE on node2
ora.ASMDB.BATCH.ASMDB2.srv                    ONLINE     ONLINE on node2
ora.ASMDB.BATCH.cs                            ONLINE     ONLINE on node2
ora.ASMDB.DISCO.ASMDB1.srv                    ONLINE     ONLINE on node1
ora.ASMDB.DISCO.cs                            ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB1.srv                     ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB2.srv                     ONLINE     ONLINE on node2
ora.ASMDB.OLTP.cs                             ONLINE     ONLINE on node2
ora.ASMDB.db                                  ONLINE     ONLINE on node1
ora.node1.ASM1.asm                            ONLINE     ONLINE on node1
ora.node1.LISTENER_NODE1.lsnr                 ONLINE     ONLINE on node1
ora.node1.gsd                                 ONLINE     ONLINE on node1
ora.node1.ons                                 ONLINE     ONLINE on node1
ora.node1.vip                                 ONLINE     ONLINE on node1
ora.node2.ASM2.asm                            ONLINE     ONLINE on node2
ora.node2.LISTENER_NODE2.lsnr                 ONLINE     ONLINE on node2
ora.node2.gsd                                 ONLINE     ONLINE on node2
ora.node2.ons                                 ONLINE     ONLINE on node2
ora.node2.vip                                 ONLINE     ONLINE on node2
{node1:oracle}/oracle ->
```

Entries added by the creation of Oracle services at Oracle Clusterware level :

Using crsstat.sh shell script.

17.3 CHECKING SERVICES

Entries added by the creation of Oracle services at Oracle Clusterware level :

Using crs_stat –t

```
{node1:oracle}/oracle -> crs_stat -t
Name           Type           Target    State     Host
------------------------------------------------------------
ora....B1.inst application    ONLINE    ONLINE    node1
ora....B2.inst application    ONLINE    ONLINE    node2
ora....DB2.srv application    ONLINE    ONLINE    node2
ora....ATCH.cs application    ONLINE    ONLINE    node2
ora....DB1.srv application    ONLINE    ONLINE    node1
ora....ISCO.cs application    ONLINE    ONLINE    node1
ora....DB1.srv application    ONLINE    ONLINE    node1
ora....DB2.srv application    ONLINE    ONLINE    node2
ora....OLTP.cs application    ONLINE    ONLINE    node2
ora.ASMDB.db   application    ONLINE    ONLINE    node1
ora....SM1.asm application    ONLINE    ONLINE    node1
ora....E1.lsnr application    ONLINE    ONLINE    node1
ora.node1.gsd  application    ONLINE    ONLINE    node1
ora.node1.ons  application    ONLINE    ONLINE    node1
ora.node1.vip  application    ONLINE    ONLINE    node1
ora....SM2.asm application    ONLINE    ONLINE    node2
ora....E2.lsnr application    ONLINE    ONLINE    node2
ora.node2.gsd  application    ONLINE    ONLINE    node2
ora.node2.ons  application    ONLINE    ONLINE    node2
ora.node2.vip  application    ONLINE    ONLINE    node2
{node1:oracle}/oracle ->
```

To get the Services definition using srvctl :

```
{node1:oracle}/oracle -> srvctl config service -d ASMDB -s OLTP
OLTP PREF: ASMDB1 ASMDB2 AVAIL:

{node1:oracle}/oracle -> srvctl config service -d ASMDB -s BATCH
BATCH PREF: ASMDB2 AVAIL: ASMDB1

{node1:oracle}/oracle -> srvctl config service -d ASMDB -s DISCO
DISCO PREF: ASMDB1 AVAIL: ASMDB2
```

How to start a service :

```
{node1:oracle}/oracle -> srvctl start service -d ASMDB -s OLTP
```

How to stop a service :

```
{node1:oracle}/oracle -> srvctl stop service -d ASMDB -s OLTP
```

How to relocate a service from one node to an other :

```
{node1:oracle}/oracle -> crs_stat | grep BATCH
ora.ASMDB.BATCH.ASMDB2.srv
ora.ASMDB.BATCH.cs
{node1:oracle}/oracle ->

➔ BATCH service is running on ASMDB2.

Let relocate the service on ASMDB1 :
{node1:oracle}/oracle -> srvctl relocate service –d ASMDB –s BATCH
```

-f : force

```
{node1:oracle}/oracle -> crs_relocate ora.ASMDB.BATCH.ASMDB2.srv –c node2
{node1:oracle}/oracle -> crs_relocate ora.ASMDB.BATCH.ASMDB2.srv –c node2 -f
```

```
BATCH =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = BATCH)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )

DISCO =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = DISCO)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
 )

OLTP =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = OLTP)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )
```

Entries added by the creation of Oracle services in the tnsnames.ora :

For each services created thru dba, grid control, dbconsole or srvctl, one entry is added in the tnsnames.ora file from each node.

All entries are configured for failover (TAF), and load balancing if prefered on at least 2 nodes.

At database level, you should see :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> show parameter service

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
service_names                        string      ASMDB, OLTP, DISCO
SQL>
```

On node1, we'll see OLTP and DISCO as they are prefered on ASMDB1.
BUT not BATCH as it's only available on ASMDB1.

From node2 :

```
{node2:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node2:oracle}/oracle -> export ORACLE_SID=ASMDB2
{node2:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> show parameter service

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
service_names                        string      ASMDB, OLTP, BATCH
SQL>
```

On node2, we'll see OLTP and BATCH as they are prefered on ASMDB2.
BUT not DISCO as it's only available on ASMDB2.

If service_names not set, empty or not well set, you should issue the following commands :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> ALTER SYSTEM SET service_names='ASMDB','OLTP','DISCO' SCOPE=BOTH SID='ASMDB1';
System Altered.
SQL> ALTER SYSTEM SET service_names='ASMDB','OLTP','BATCH' SCOPE=BOTH SID='ASMDB2';
System Altered.
SQL>
```

Before failure of one node, crs_stat –t will show :

```
{node1:oracle}/oracle -> crsstat.sh
HA Resource                                   Target     State
-----------                                   ------     -----
ora.ASMDB.ASMDB1.inst                         ONLINE     ONLINE on node1
ora.ASMDB.ASMDB2.inst                         ONLINE     ONLINE on node2
ora.ASMDB.BATCH.ASMDB2.srv                    ONLINE     ONLINE on node2
ora.ASMDB.BATCH.cs                            ONLINE     ONLINE on node2
ora.ASMDB.DISCO.ASMDB1.srv                    ONLINE     ONLINE on node1
ora.ASMDB.DISCO.cs                            ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB1.srv                     ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB2.srv                     ONLINE     ONLINE on node2
ora.ASMDB.OLTP.cs                             ONLINE     ONLINE on node2
ora.ASMDB.db                                  ONLINE     ONLINE on node1
ora.node1.ASM1.asm                            ONLINE     ONLINE on node1
ora.node1.LISTENER_NODE1.lsnr                 ONLINE     ONLINE on node1
ora.node1.gsd                                 ONLINE     ONLINE on node1
ora.node1.ons                                 ONLINE     ONLINE on node1
ora.node1.vip                                 ONLINE     ONLINE on node1
ora.node2.ASM2.asm                            ONLINE     ONLINE on node2
ora.node2.LISTENER_NODE2.lsnr                 ONLINE     ONLINE on node2
ora.node2.gsd                                 ONLINE     ONLINE on node2
ora.node2.ons                                 ONLINE     ONLINE on node2
ora.node2.vip                                 ONLINE     ONLINE on node2
{node1:oracle}/oracle ->
```

If node2 fails, or reboot for any reason :

What should we see on node1 ?

After failure : On node1, AS BATCH was prefered on ASMDB2, and available on ASMDB1, THEN after node2 failure, BATCH service will be swithed to prefered on ASMDB1.

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> show parameter service

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
service_names                        string      ASMDB, OLTP, DISCO, BATCH
SQL>
```

After failure of node2, crs_stat –t will show :

- **VIP from node2 swith to node1.**
- **ONS, GSD, Listener, ASM2 instance, and ASDB2 instance are switch to OFFLINE state.**
- **OLTP service is switch to OFFLINE state on node2. But still ONLINE state on node1**
- **BATCH service is switch to ONLINE state from node2 to node1. THEN BATCH service is still available thru node1.**

```
{node1:oracle}/oracle -> crsstat.sh
HA Resource                                   Target     State
-----------                                   ------     -----
ora.ASMDB.ASMDB1.inst                         ONLINE     ONLINE on node1
ora.ASMDB.ASMDB2.inst                         ONLINE     OFFLINE
ora.ASMDB.BATCH.ASMDB2.srv                    ONLINE     ONLINE on node1
ora.ASMDB.BATCH.cs                            ONLINE     ONLINE on node1
ora.ASMDB.DISCO.ASMDB1.srv                    ONLINE     ONLINE on node1
ora.ASMDB.DISCO.cs                            ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB1.srv                     ONLINE     ONLINE on node1
ora.ASMDB.OLTP.ASMDB2.srv                     ONLINE     OFFLINE
ora.ASMDB.OLTP.cs                             ONLINE     ONLINE on node2
ora.ASMDB.db                                  ONLINE     ONLINE on node1
ora.node1.ASM1.asm                            ONLINE     ONLINE on node1
ora.node1.LISTENER_NODE1.lsnr                 ONLINE     ONLINE on node1
ora.node1.gsd                                 ONLINE     ONLINE on node1
ora.node1.ons                                 ONLINE     ONLINE on node1
ora.node1.vip                                 ONLINE     ONLINE on node1
ora.node2.ASM2.asm                            ONLINE     OFFLINE
ora.node2.LISTENER_NODE2.lsnr                 ONLINE     OFFLINE
ora.node2.gsd                                 ONLINE     OFFLINE
ora.node2.ons                                 ONLINE     OFFLINE
ora.node2.vip                                 ONLINE     ONLINE on node1
{node1:oracle}/oracle ->
```

Check that services are registrered with command :

lsnrctl status ....

On node1 with listener_node1

On node2 with listener_node2

```
{node1:oracle}/oracle -> lsnrctl status listener_node1

LSNRCTL for IBM/AIX RISC System/6000: Version 10.2.0.3.0 - Production on 22-MAR-2007
16:27:21

Copyright (c) 1991, 2006, Oracle.  All rights reserved.

Connecting to (DESCRIPTION=(ADDRESS=(PROTOCOL=IPC)(KEY=EXTPROC)))
STATUS of the LISTENER
------------------------
Alias                     LISTENER_NODE1
Version                   TNSLSNR for IBM/AIX RISC System/6000: Version 10.2.0.3.0 -
Production
Start Date                20-MAR-2007 16:08:24
Uptime                    2 days 0 hr. 18 min. 59 sec
Trace Level               off
Security                  ON: Local OS Authentication
SNMP                      ON
Listener Parameter File   /oracle/products/asm/network/admin/listener.ora
Listener Log File         /oracle/products/asm/network/log/listener_node1.log
Listening Endpoints Summary...
  (DESCRIPTION=(ADDRESS=(PROTOCOL=ipc)(KEY=EXTPROC)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.181)(PORT=1521)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.81)(PORT=1521)))
Services Summary...
Service "+ASM" has 1 instance(s).
  Instance "+ASM1", status BLOCKED, has 1 handler(s) for this service...
Service "+ASM_XPT" has 1 instance(s).
  Instance "+ASM1", status BLOCKED, has 1 handler(s) for this service...
Service "ASMDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "ASMDBXDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "ASMDB_XPT" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "BATCH" has 1 instance(s).
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "DISCO" has 1 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
Service "OLTP" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "PLSExtProc" has 1 instance(s).
  Instance "PLSExtProc", status UNKNOWN, has 1 handler(s) for this service...
The command completed successfully
{node1:oracle}/oracle ->
```

```
Service "BATCH" has 1 instance(s).
 Instance "ASMDB2", status READY, has 1 handler(s) for this service...
```

This is normal as BATCH service is

- Available on ASMDB1 instance, node1
- Prefered on ASMDB2 instance, node2

```
Service "DISCO" has 1 instance(s).
 Instance "ASMDB1", status READY, has 2 handler(s) for this service...
```

This is normal as DISCO service is

- Prefered on ASMDB1 instance, node1
- Available on ASMDB2 instance, node2

```
Service "OLTP" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for
this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for
this service...
```

This is normal as OLTP service is

- Prefered on ASMDB1 instance, node1
- Prefered on ASMDB2 instance , node2

18 TAF CONFIGURATION

$ORACLE_HOME/network/admin/listener.ora
Which will be in /oracle/products/asm/network/admin

Checking content of listener.ora on node1 :

Listener is listening on node1-vip and on node1 (10.3.25.81).

Users will connect thru node1-vip

DBA's can connect thru node1 for administration purpose.

IPC not necessary.

```
{node1:oracle}/oracle/products/asm/network/admin -> cat listener.ora
# listener.ora.node1 Network Configuration File:
/oracle/products/asm/network/admin/listener.ora.node1
# Generated by Oracle configuration tools.

SID_LIST_LISTENER_NODE1 =
  (SID_LIST =
    (SID_DESC =
      (SID_NAME = PLSExtProc)
      (ORACLE_HOME = /oracle/products/asm)
      (PROGRAM = extproc)
    )
  )

LISTENER_NODE1 =
  (DESCRIPTION_LIST =
    (DESCRIPTION =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC))
      (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521)(IP = FIRST))
      (ADDRESS = (PROTOCOL = TCP)(HOST = 10.3.25.81)(PORT = 1521)(IP = FIRST))
    )
  )
{node1:oracle}/oracle/products/asm/network/admin ->
```

$ORACLE_HOME/network/admin/listener.ora
Which will be in /oracle/products/asm/network/admin

Checking content of listener.ora on node2 :

Listener is listening on node2-vip and on node2 (10.3.25.82).

Users will connect thru node1-vip

DBA's can connect thru node2 for administration purpose.

IPC not necessary.

```
{node2:oracle}/oracle/products/asm/network/admin -> cat listener.ora
# listener.ora.node2 Network Configuration File:
/oracle/products/asm/network/admin/listener.ora.node2
# Generated by Oracle configuration tools.

SID_LIST_LISTENER_NODE1 =
  (SID_LIST =
    (SID_DESC =
      (SID_NAME = PLSExtProc)
      (ORACLE_HOME = /oracle/products/asm)
      (PROGRAM = extproc)
    )
  )

LISTENER_NODE1 =
  (DESCRIPTION_LIST =
    (DESCRIPTION =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC))
      (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521)(IP = FIRST))
      (ADDRESS = (PROTOCOL = TCP)(HOST = 10.3.25.82)(PORT = 1521)(IP = FIRST))
    )
  )
{node2:oracle}/oracle/products/asm/network/admin ->
```

$ORACLE_HOME/network/admin/tnsnames.ora implementing TAF
Which will be in /oracle/products/asm/network/admin

On all RAC nodes, for ASMDB Database, and associated oracle services :

```
{node1:oracle}/oracle/products/asm/network/admin -> cat tnsnames.ora
# tnsnames.ora Network Configuration File:
/oracle/products/asm/network/admin/tnsnames.ora
# Generated by Oracle configuration tools.

LISTENERS_ASMDB =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
  )

LISTENER_ASMDB1 =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
  )

LISTENER_ASMDB2 =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
  )

ASMDB2 =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = ASMDB)
      (INSTANCE_NAME = ASMDB2)
    )
  )

ASMDB1 =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = ASMDB)
      (INSTANCE_NAME = ASMDB1)
    )
  )

ASMDB =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = ASMDB)
    )
  )
```

$ORACLE_HOME/network/admin/tnsnames.ora implementing TAF

Continued ….

```
DISCO =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = DISCO)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )

BATCH =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = BATCH)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )

OLTP =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2-vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = OLTP)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 180)
        (DELAY = 5)
      )
    )
  )

EXTPROC_CONNECTION_DATA =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC1))
    )
    (CONNECT_DATA =
      (SID = PLSExtProc)
      (PRESENTATION = RO)
    )
  )

{node1:oracle}/oracle/products/asm/network/admin ->
```

At database level, you should see :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> show parameter service

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
service_names                        string      ASMDB, OLTP, DISCO
SQL>
SQL> show parameter listener

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
remote_listener                      string      LISTENERS_ASMDB
local_listener                       string      LISTENER_ASMDB1
```

From node2 :

```
{node2:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node2:oracle}/oracle -> export ORACLE_SID=ASMDB2
{node2:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> show parameter service

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
service_names                        string      ASMDB, OLTP, BATCH
SQL>
SQL> show parameter listener

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
remote_listener                      string      LISTENERS_ASMDB
local_listener                       string      LISTENER_ASMDB2
```

- remote_listener from node1, and node2 MUST BE THE SAME, and ENTRIES MUST BE PRESENT in the tnsnames.ora from each node.
- local_listener from node1, and node2 are differents, and ENTRIES MUST BE PRESENT in the tnsnames.ora from each node.
- local_listener from node1, and node2 are not the ones defined in the listener.ora files from each node.

If remote_listener, and local_listener not set, empty or not well set, you should issue the following commands :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle -> sqlplus /nolog

SQL*Plus: Release 10.2.0.3.0 - Production on Fri Apr 6 12:50:59 2007

Copyright (c) 1982, 2005, Oracle.  All Rights Reserved.

SQL> connect /as sysdba
Connected.
SQL> ALTER SYSTEM SET remote_listener='LISTENERS_ASMDB' SCOPE=BOTH SID='*';
System Altered.
SQL> ALTER SYSTEM SET local_listener='LISTENER_ASMDB1' SCOPE=BOTH SID='ASMDB1';
System Altered.
SQL> ALTER SYSTEM SET local_listener='LISTENER_ASMDB2' SCOPE=BOTH SID='ASMDB2';
System Altered.
SQL>
```

Check using lsnrctl status listener_node1 :

```
{node1:oracle}/oracle -> lsnrctl status listener_node1

LSNRCTL for IBM/AIX RISC System/6000: Version 10.2.0.3.0 - Production on 22-MAR-2007 16:27:21

Copyright (c) 1991, 2006, Oracle.  All rights reserved.

Connecting to (DESCRIPTION=(ADDRESS=(PROTOCOL=IPC)(KEY=EXTPROC)))
STATUS of the LISTENER
------------------------
Alias                     LISTENER_NODE1
Version                   TNSLSNR for IBM/AIX RISC System/6000: Version 10.2.0.3.0 - Production
Start Date                20-MAR-2007 16:08:24
Uptime                    2 days 0 hr. 18 min. 59 sec
Trace Level               off
Security                  ON: Local OS Authentication
SNMP                      ON
Listener Parameter File   /oracle/products/asm/network/admin/listener.ora
Listener Log File         /oracle/products/asm/network/log/listener_node1.log
Listening Endpoints Summary...
  (DESCRIPTION=(ADDRESS=(PROTOCOL=ipc)(KEY=EXTPROC)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.181)(PORT=1521)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.81)(PORT=1521)))
Services Summary...
Service "+ASM" has 1 instance(s).
  Instance "+ASM1", status BLOCKED, has 1 handler(s) for this service...
Service "+ASM_XPT" has 1 instance(s).
  Instance "+ASM1", status BLOCKED, has 1 handler(s) for this service...
Service "ASMDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "ASMDBXDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "ASMDB_XPT" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "BATCH" has 1 instance(s).
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "DISCO" has 1 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
Service "OLTP" has 2 instance(s).
  Instance "ASMDB1", status READY, has 2 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "PLSExtProc" has 1 instance(s).
  Instance "PLSExtProc", status UNKNOWN, has 1 handler(s) for this service...
The command completed successfully
{node1:oracle}/oracle ->
```

Check using lsnrctl status listener_node2 :

```
{node2:oracle}/oracle -> lsnrctl status listener_node2

LSNRCTL for IBM/AIX RISC System/6000: Version 10.2.0.3.0 - Production on 22-MAR-2007 16:28:13

Copyright (c) 1991, 2006, Oracle.  All rights reserved.

Connecting to (ADDRESS=(PROTOCOL=tcp)(HOST=)(PORT=1521))
STATUS of the LISTENER
------------------------
Alias                     LISTENER_NODE2
Version                   TNSLSNR for IBM/AIX RISC System/6000: Version 10.2.0.3.0 - Production
Start Date                19-MAR-2007 18:40:46
Uptime                    2 days 21 hr. 47 min. 27 sec
Trace Level               off
Security                  ON: Local OS Authentication
SNMP                      ON
Listener Parameter File   /oracle/products/asm/network/admin/listener.ora
Listener Log File         /oracle/products/asm/network/log/listener_node2.log
Listening Endpoints Summary...
  (DESCRIPTION=(ADDRESS=(PROTOCOL=ipc)(KEY=EXTPROC)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.182)(PORT=1521)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.3.25.82)(PORT=1521)))
Services Summary...
Service "+ASM" has 1 instance(s).
  Instance "+ASM2", status BLOCKED, has 1 handler(s) for this service...
Service "+ASM_XPT" has 1 instance(s).
  Instance "+ASM2", status BLOCKED, has 1 handler(s) for this service...
Service "ASMDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 2 handler(s) for this service...
Service "ASMDBXDB" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 1 handler(s) for this service...
Service "ASMDB_XPT" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 2 handler(s) for this service...
Service "BATCH" has 1 instance(s).
  Instance "ASMDB2", status READY, has 2 handler(s) for this service...
Service "DISCO" has 1 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
Service "OLTP" has 2 instance(s).
  Instance "ASMDB1", status READY, has 1 handler(s) for this service...
  Instance "ASMDB2", status READY, has 2 handler(s) for this service...
Service "PLSExtProc" has 1 instance(s).
  Instance "PLSExtProc", status UNKNOWN, has 1 handler(s) for this service...
The command completed successfully
{node2:oracle}/oracle ->
```

$ORACLE_HOME/network/admin/tnsnames.ora implementing TAF
Which will be in /oracle/products/asm/network/admin

On client side for ASMDB Database :

ASMDB Connect String for :
➔no failover
➔load balancing

```
# tnsnames.ora Network Configuration File:
/oracle/client10g/network/admin/tnsnames.ora
# Generated by Oracle configuration tools.

ASMDB =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1_vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2_vip)(PORT = 1521))
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = ASMDB)
    )
  )
```

ASMDB_FAILOVER Connect String for :
➔failover
➔load balancing

```
ASMDB_FAILOVER =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = node1_vip)(PORT = 1521))
      (ADDRESS = (PROTOCOL = TCP)(HOST = node2_vip)(PORT = 1521))
    )
    (LOAD_BALANCE = yes)
    (CONNECT_DATA =
      (SERVICE_NAME = ASMDB)
      (FAILOVER_MODE =
        (TYPE = SELECT)
        (METHOD = BASIC)
        (RETRIES = 20)
        (DELAY = 60)
      )
    )
  )
```

19 ASM ADVANCED TOOLS

Access to document :

- **«Commanding ASM »** on http://www.oracle.com/oramag
  http://www.oracle.com/technology/oramag/oracle/06-mar/o26asm.html
  Access, transfer, and administer ASM files without SQL commands.
  By Arup Nanda

- **«Add Storage, Not Projects »** on http://www.oracle.com/oramag
  http://www.oracle.com/technology/oramag/oracle/04-may/o34tech_management.html
  Automatic Storage Management lets DBAs do their jobs and keep their weekends.
  By Jonathan Gennick

19.1 ASMCMD

To access the content of the ASM Disks Groups :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > help
        asmcmd [-p] [command]

        The environment variables ORACLE_HOME and ORACLE_SID determine the
        instance to which the program connects, and ASMCMD establishes a
        bequeath connection to it, in the same manner as a SQLPLUS / AS
        SYSDBA.  The user must be a member of the SYSDBA group.

        Specifying the -p option allows the current directory to be displayed
        in the command prompt, like so:

        ASMCMD [+DATAFILE/ORCL/CONTROLFILE] >

        [command] specifies one of the following commands, along with its
        parameters.

        Type "help [command]" to get help on a specific ASMCMD command.

        commands:
        --------
        cd
        du
        find
        help
        ls
        lsct
        lsdg
        mkalias
        mkdir
        pwd
        rm
        rmalias
      ASMCMD [+] >
```

Subject: **ASMCMD - ASM command line utility** Doc ID: **Note:332180.1**

Using ASMCD "lsct" command :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > help lsct
        lsct [-H] [group]

        List all clients and their attributes from V$ASM_CLIENT.  If group is
        specified, then return only information on that group.

        The -H flag suppresses the column headers from the output.
ASMCMD [+] >
ASMCMD [+] > lsct
DB_Name   Status        Software_Version  Compatible_version  Instance_Name
ASMDB     CONNECTED           10.2.0.3.0          10.2.0.3.0  ASMDB1
ASMCMD [+] >
```

Using ASMCD "lsdg" command :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > help lsdg
        lsdg [-H] [group]

        List all diskgroups and their attributes from V$ASM_DISKGROUP.  If
        [group] is specified, then return only information on that group.  The
        command also informs the user if a rebalance is currently under way
        for a diskgroup.

        The -H flag suppresses the column headers from the output.

        Not all possible disk group attributes are included.  To view the
        complete set of attributes for a disk group, query the V$ASM_DISKGROUP
        view.
ASMCMD [+] >
ASMCMD [+] > lsgd
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      8192     5114                0
5114              0  DG_DB_ASM/
ASMCMD [+] >
ASMCMD [+] >
```

Using ASMCD "du" command :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > help du
        du [-H] [dir]

        Display total space used for files located recursively under [dir],
        similar to "du -s" under UNIX; default is the current directory.  Two
        values are returned, both in units of megabytes.  The first value does
        not take into account mirroring of the diskgroup while the second does.
        For instance, if a file occupies 100 MB of space, then it actually
        takes up 200 MB of space on a normal redundancy diskgroup and 300 MB
        of space on a high redundancy diskgroup.

        [dir] can also contain wildcards.

        The -H flag suppresses the column headers from the output.
ASMCMD [+] >
```

```
ASMCMD [+] > du
Used_MB      Mirror_used_MB
   2982                2982
ASMCMD [+] >
```

Using ASMCD "ls" command :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > help ls
        ls [-lsdrtLaH] [name]

        List [name] or its contents alphabetically if [name] refers to a
        directory.  [name] can contain the wildcard "*" and is the current
        directory if unspecified.  Directory names in the display list
        have the "/" suffix to clarify their identity.  The first two optional
        flags specify how much information is displayed for each file, in the
        following manner:

        (no flag)        V$ASM_ALIAS.NAME

         -l              V$ASM_ALIAS.NAME, V$ASM_ALIAS.SYSTEM_CREATED;
                         V$ASM_FILE.TYPE, V$ASM_FILE.REDUNDANCY,
                         V$ASM_FILE.STRIPED, V$ASM_FILE.MODIFICATION_DATE

         -s              V$ASM_ALIAS.NAME;
                         V$ASM_FILE.BLOCK_SIZE, V$ASM_FILE.BLOCKS,
                         V$ASM_FILE.BYTES, V$ASM_FILE.SPACE

        If the user specifies both flags, then the command shows an union of
        their attributes, with duplicates removed.

        If an entry in the list is an user-defined alias or a directory,
        then -l displays only the V$ASM_ALIAS attributes, and -s shows only
        the alias name and its size, which is zero because it is negligible.
        Moreover, the displayed name contains a suffix that is in the form of
        an arrow pointing to the absolute path of the system-created filename
        it references:

        t_db1.f => +diskgroupName/DBName/DATAFILE/SYSTEM.256.1

        See the -L option below for an exception to this rule.

        The remaining flags have the following meanings:

         -d             If an argument is a directory,  list only its name
                        (not its contents).

         -r             Reverse the sorting order.

         -t             Sort by time stamp (latest first) instead of by name.

         -L             If an argument is an user alias, display information on
                        the file it references.

         -a             If an argument is a system-created filename, show the
                        location of its user-defined alias, if any.

         -H             Suppress the column header information, so that
                        scripting is easier.

        Note that "ls +" would return information on all diskgroups, including
        whether they are mounted.

        Not all possible file attributes or disk group attributes are included.
        To view the complete set of attributes for a file or a disk group,
        query the V$ASM_FILE and V$ASM_DISKGROUP views.

ASMCMD [+] >
```

```
ASMCMD [+] > ls
DG_DB_ASM/
ASMCMD [+] >
ASMCMD [+] > ls -la
State    Type    Rebal  Unbal  Name
MOUNTED  EXTERN  N      N      DG_DB_ASM/
ASMCMD [+] >
```

Using ASMCD "ls" and 'cd' commands :

From node1 :

```
{node1:oracle}/oracle -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle -> export ORACLE_SID=+ASM1
{node1:oracle}/oracle -> asmcmd -p
ASMCMD [+] > ls
DG_DB_ASM/
ASMCMD [+] > cd DG_DB_ASM/
ASMCMD [+DG_DB_ASM] > ls
ASMDB/
ASMCMD [+DG_DB_ASM] > cd ASMDB/
ASMCMD [+DG_DB_ASM/ASMDB] > ls
CONTROLFILE/
DATAFILE/
ONLINELOG/
PARAMETERFILE/
TEMPFILE/
spfileASMDB.ora
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la
Type           Redund  Striped  Time             Sys  Name
                                                Y    CONTROLFILE/
                                                Y    DATAFILE/
                                                Y    ONLINELOG/
                                                Y    PARAMETERFILE/
                                                Y    TEMPFILE/
                                                N    spfileASMDB.ora =>
+DG_DB_ASM/ASMDB/PARAMETERFILE/spfile.287.621487507
ASMCMD [+DG_DB_ASM/ASMDB] >
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la CONTROLFILE/
Type         Redund  Striped  Time             Sys  Name
CONTROLFILE  UNPROT  FINE     MAY 02 02:00:00  Y    none => Current.260.620649571
CONTROLFILE  UNPROT  FINE     MAY 02 09:00:00  Y    none => Current.267.621486441
CONTROLFILE  UNPROT  FINE     MAY 02 09:00:00  Y    none => Current.268.621486441
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la DATAFILE/
Type      Redund  Striped  Time             Sys  Name
DATAFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => SYSAUX.257.620649481
DATAFILE  UNPROT  COARSE   MAY 02 14:00:00  Y    none => SYSAUX.277.621486497
DATAFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => SYSTEM.256.620649481
DATAFILE  UNPROT  COARSE   MAY 02 09:00:00  Y    none => SYSTEM.275.621486459
DATAFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => UNDOTBS1.258.620649483
DATAFILE  UNPROT  COARSE   MAY 02 09:00:00  Y    none => UNDOTBS1.276.621486487
DATAFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => UNDOTBS2.264.620649617
DATAFILE  UNPROT  COARSE   MAY 02 09:00:00  Y    none => UNDOTBS2.279.621486511
DATAFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => USERS.259.620649483
DATAFILE  UNPROT  COARSE   MAY 02 09:00:00  Y    none => USERS.280.621486521
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la ONLINELOG/
Type       Redund  Striped  Time             Sys  Name
ONLINELOG  UNPROT  FINE     MAY 02 02:00:00  Y    none => group_1.261.620649577
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_1.269.621486443
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_1.270.621486443
ONLINELOG  UNPROT  FINE     MAY 02 02:00:00  Y    none => group_2.262.620649579
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_2.271.621486445
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_2.272.621486447
ONLINELOG  UNPROT  FINE     MAY 02 02:00:00  Y    none => group_3.265.620649661
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_3.281.621487491
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_3.282.621487493
ONLINELOG  UNPROT  FINE     MAY 02 02:00:00  Y    none => group_4.266.620649661
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_4.283.621487493
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_4.284.621487495
```

```
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_5.273.621486451
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_5.274.621486453
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_6.285.621487497
ONLINELOG  UNPROT  FINE     MAY 02 09:00:00  Y    none => group_6.286.621487501
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la PARAMETERFILE/
Type           Redund  Striped  Time             Sys  Name
PARAMETERFILE  UNPROT  COARSE   MAY 02 16:00:00  Y
+DG_DB_ASM/ASMDB/spfileASMDB.ora => spfile.287.621487507
ASMCMD [+DG_DB_ASM/ASMDB] > ls -la TEMPFILE/
Type      Redund  Striped  Time             Sys  Name
TEMPFILE  UNPROT  COARSE   MAY 02 02:00:00  Y    none => TEMP.263.620649593
TEMPFILE  UNPROT  COARSE   MAY 02 09:00:00  Y    none => TEMP.278.621486505
ASMCMD [+DG_DB_ASM/ASMDB] >
```

19.2 FTP AND HTTP ACCESS

Subject: **How to configure XDB for using ftp and http protocols with ASM** Doc ID: **Note:357714.1**

Subject: **/Sys/Asm path is not visible in XDB Repository** Doc ID: **Note:368454.1**

Subject: **How to Deinstall and Reinstall XML Database (XDB)** Doc ID: **Note:243554.1**

20 SOME USEFULL COMMANDS

Command to start/stop the Database, and Databases Instances :

From any node :

For database

{node1:oracle}/oracle -> **srvctl start database –d ASMDB** to start the Database instance
{node1:oracle}/oracle -> **srvctl stop database –d ASMDB** to stop the Database instance

For instance 1

{node1:oracle}/oracle -> **srvctl start instance –d ASMDB –i ASMDB1** to start the DB.. instance
{node1:oracle}/oracle -> **srvctl stop instance –d ASMDB –i ASMDB1** to stop the DB.. instance

For instance 2
{node1:oracle}/oracle -> **srvctl start instance –d ASMDB –i ASMDB2** to start the DB.. instance
{node1:oracle}/oracle -> **srvctl stop instance –d ASMDB –i ASMDB2** to stop the DB.. instance

To access an ASM instance with sqlplus

From node1 :

{node1:oracle}/oracle -> **set ORACLE_SID=+ASM1**
{node1:oracle}/oracle -> **sqlplus /nolog**
{node1:oracle}/oracle -> **connect /as sysdba**
show sga
……..

From node2 :

{node2:oracle}/oracle -> **set ORACLE_SID=+ASM2**
{node2:oracle}/oracle -> **sqlplus /nolog**
{node2:oracle}/oracle -> **connect /as sysdba**
show sga
……..

To acces a Database instance stored in ASM with sqlplus

From node1 :

{node1:oracle}/oracle -> **export ORACLE_SID=ASMDB1**
{node1:oracle}/oracle -> **sqlplus /nolog**
{node1:oracle}/oracle -> **connect /as sysdba**
show sga
…….

20.1 Oracle CLuster Registry content Check and Backup

Check Oracle Cluster Registry Integrity

As oracle user, Execute ocrcheck

```
{node1:oracle}/oracle/crs/bin ->ocrcheck
Status of Oracle Cluster Registry is as follows :
        Version                  :          2
        Total space (kbytes)     :     204712
        Used space (kbytes)      :       4624
        Available space (kbytes) :     200088
        ID                       :  668290851
        Device/File Name         : /dev/ocr1_disk
                                   Device/File integrity check succeeded

        Device/File Name         : /dev/ocr2_disk
                                   Device/File integrity check succeeded

        Cluster registry integrity check succeeded

{node1:oracle}/oracle/crs/bin ->
```

AS root user :

Export Oracle Cluster Registry content

```
{node1:oracle}/oracle/crs/bin ->su
root's Password:
{node1:root}/oracle/crs/bin ->ocrconfig -export /oracle/ocr_export4.dmp -s online
{node1:root}/oracle/crs/bin ->ls -la /oracle/*.dmp
-rw-r--r--   1 root     system       106420 Jan 30 18:30 /oracle/ocr_export.dmp
{node1:oracle}/oracle/crs/bin ->
```

➔ **you must not edit/modify this exported file**

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oracle/crs/bin ->ocrconfig -showbackup

node1     2006/01/30 16:03:03     /oracle/crs/cdata/crs

node1     2006/01/30 12:03:00     /oracle/crs/cdata/crs

node1     2006/01/30 08:02:59     /oracle/crs/cdata/crs

node1     2006/01/29 00:02:51     /oracle/crs/cdata/crs

node1     2006/01/25 13:02:10     /oracle/crs/cdata/crs

{node1:oracle}/oracle/crs/bin ->
```

20.2 CHECKING DB CONSOLE

With 10gRAC R2, dbconsole is only configured on first nore at database creation.
When using DBCONSOLE, each database will have it's own dbconsole !!!

Look at Metalink Note :

Subject: **How to manage DB Control 10.2 for RAC Database with emca** Doc ID: **Note:395162.1**

How to start/stop the dbconsole agent :

➔ When using DBCONSOLE, the oracle agent start and stop automatically with the dbconsole. If needed, it's still possible to start/stop the agent using the emctl tool as follow :

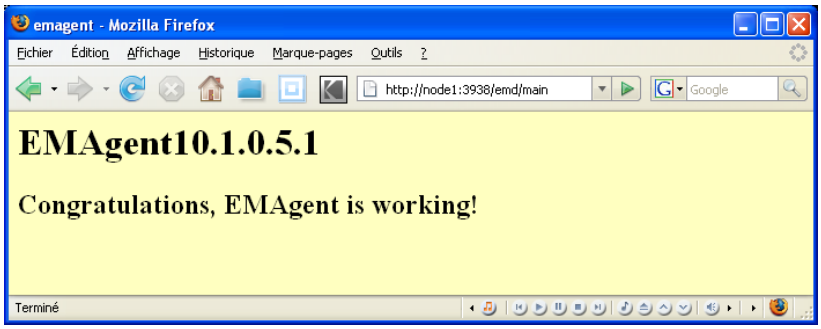
{**node1**:oracle}/oracle/products/rdbms -> **export ORACLE_HOME=/oracle/products/rdbms**
{node1:oracle}/oracle/products/rdbms -> **export ORACLE_SID=ASMDB1**
{node1:oracle}/oracle/products/rdbms -> **emctl start agent**
OR
{node1:oracle}/oracle/products/rdbms -> **emctl stop agent**

Check status of dbconsole agent for node1

{**node1**:oracle}/oracle/products/rdbms -> **export ORACLE_HOME=/oracle/products/rdbms**
{node1:oracle}/oracle/products/rdbms -> **export ORACLE_SID=ASMDB1**
{node1:oracle}/oracle/products/rdbms -> **emctl status agent**

```
Oracle Enterprise Manager 10g Database Control Release 10.2.0.3.0
Copyright (c) 1996, 2006 Oracle Corporation.  All rights reserved.
------------------------------------------------------------
Agent Version     : 10.1.0.5.1
OMS Version       : 10.1.0.5.0
Protocol Version  : 10.1.0.2.0
Agent Home        : /oracle/products/rdbms/node1_ASMDB1
Agent binaries    : /oracle/products/rdbms
Agent Process ID  : 1507562
Parent Process ID : 1437786
Agent URL         : http://node1:3938/emd/main
Started at        : 2007-04-23 10:44:38
Started by user   : oracle
Last Reload       : 2007-04-23 10:44:38
Last successful upload                       : 2007-04-23 17:32:23
Total Megabytes of XML files uploaded so far :     7.64
Number of XML files pending upload           :       0
Size of XML files pending upload(MB)         :     0.00
Available disk space on upload filesystem    :    37.78%
------------------------------------------------------------
Agent is Running and Ready
```
{node1:oracle}/oracle/products/rdbms ->

Check status of EM agent for node1, and node2 :



Commands to start DBCONSOLE :

```
{node1:oracle}/oracle/products/rdbms -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle/products/rdbms -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle/products/rdbms -> emctl start dbconsole
Oracle Enterprise Manager 10g Database Control Release 10.2.0.3.0
Copyright (c) 1996, 2006 Oracle Corporation.  All rights reserved.
http://node1:1158/em/console/aboutApplication
Starting Oracle Enterprise Manager 10g Database Control ..... started.
------------------------------------------------------------------
Logs are generated in directory /oracle/products/rdbms/node1_ASMDB1/sysman/log
{node1:oracle}/oracle/products/rdbms ->
```

Commands to stop DBCONSOLE :

```
{node1:oracle}/oracle/products/rdbms -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle/products/rdbms -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle/products/rdbms -> emctl stop dbconsole
Oracle Enterprise Manager 10g Database Control Release 10.2.0.3.0
Copyright (c) 1996, 2006 Oracle Corporation.  All rights reserved.
http://node1:1158/em/console/aboutApplication
Stopping Oracle Enterprise Manager 10g Database Control ...
 ...  Stopped.
{node1:oracle}/oracle/products/rdbms ->
```

Check status of dbconsole

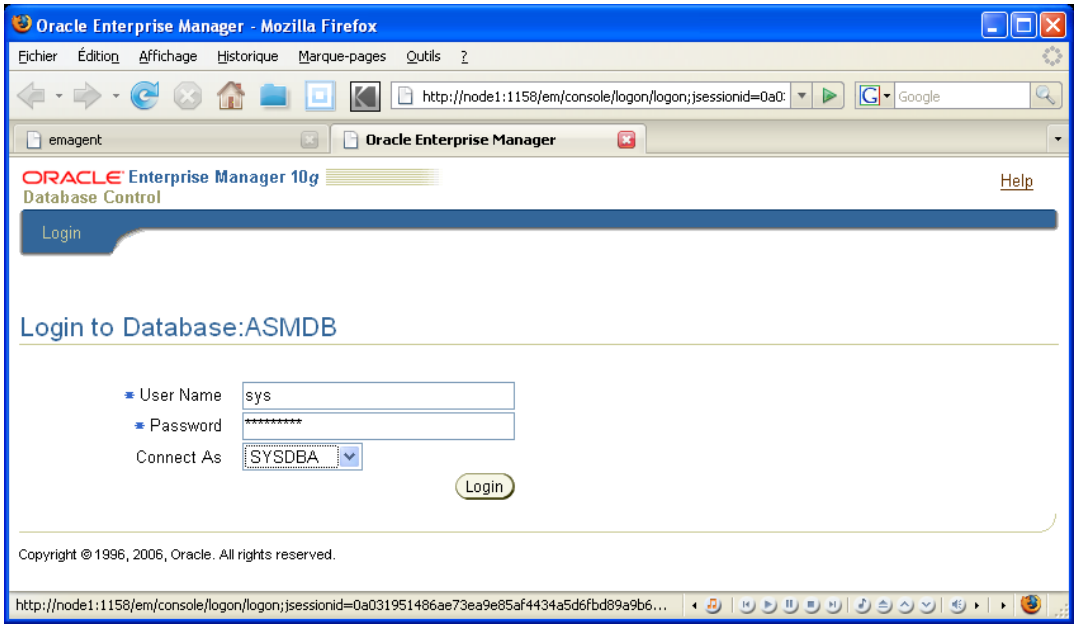
```
{node1:oracle}/oracle/products/rdbms -> export ORACLE_HOME=/oracle/products/rdbms
{node1:oracle}/oracle/products/rdbms -> export ORACLE_SID=ASMDB1
{node1:oracle}/oracle/products/rdbms -> emctl status dbconsole
Oracle Enterprise Manager 10g Database Control Release 10.2.0.3.0
Copyright (c) 1996, 2006 Oracle Corporation.  All rights reserved.
http://node1:1158/em/console/aboutApplication
Oracle Enterprise Manager 10g is running.
------------------------------------------------------------------
Logs are generated in directory /oracle/products/rdbms/node1_ASMDB1/sysman/log
{node1:oracle}/oracle/products/rdbms ->
```

Update your local host file with entries for node1 and node2.

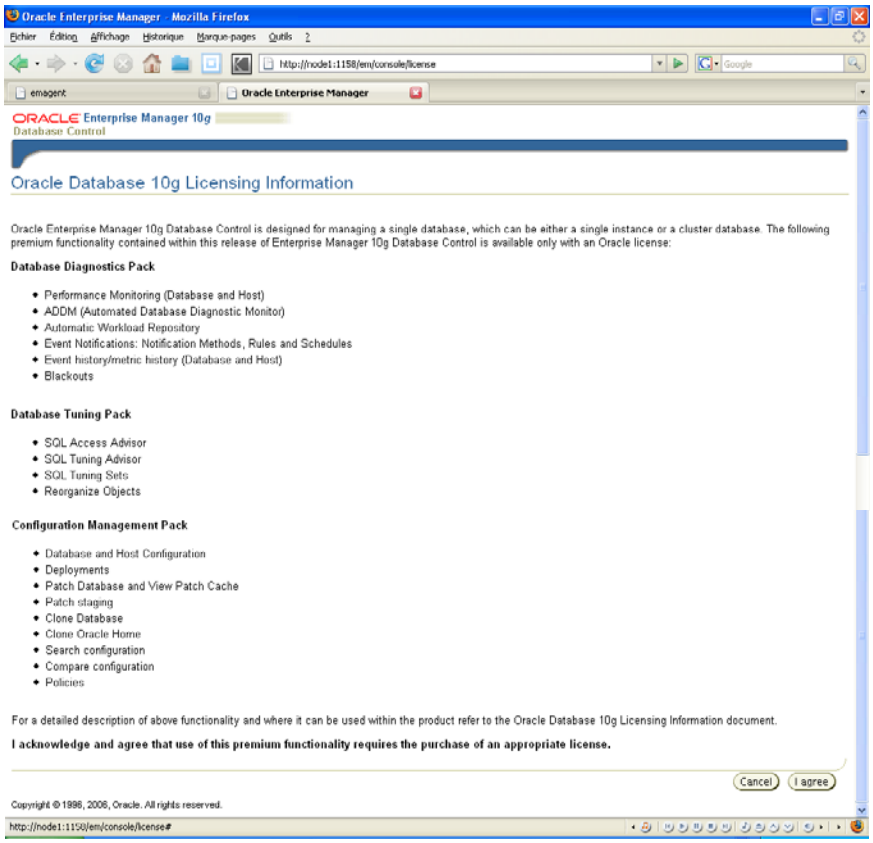
AND

Access DBCONSOLE thru http://node1:1158/em

Connect with oracle database “sys” user as “SYSDBA”



Just click on "I agree" to accept the Oracle Database 10g Licensing Information.



Look at Metalink Note to discover dbconsole, or get help to solve any issues :

Subject: **How to configure dbconsole to display information from the ASM DISKGROUPS** Doc ID: **Note:329581.1**

Subject: **Em Dbconsole 10.1.0.X Does Not Discover Rac Database As A Cluster** Doc ID: **Note:334546.1**

Subject: **DBConsole Shows Everything Down** Doc ID: **Note:332865.1**

Subject: **How To Config Dbconsole (10.1.x or 10.2) EMCA With Another Hostname** Doc ID: **Note:336017.1**

Subject: **How To Change The DB Console Language To English** Doc ID: **Note:370178.1**

Subject: **Dbconsole Fails After A Physical Move Of Machine To New Domain/Hostnam** Doc ID: **Note:401943.1**

Subject: **How to Troubleshoot Failed Login Attempts to DB Control** Doc ID: **Note:404820.1**

Subject: **How to manage DB Control 10.2 for RAC Database with emca** Doc ID: **Note:395162.1**

Subject: **How To Drop, Create And Recreate DB Control In A 10g Database** Doc ID: **Note:278100.1**

Subject: **Overview Of The EMCA Commands Available for DB Control 10.2 Installations** Doc ID: **Note:330130.1**

Subject: **How to change the password of the 10g database user dbsnmp** Doc ID: **Note:259387.1**

Subject: **EM 10.2 DBConsole Displaying Wrong Old Information** Doc ID: **Note:336179.1**

Subject: **How To Configure A Listener Target From Grid Control 10g** Doc ID: **Note:427422.1**

Testing if dbconsole is working or not !!!

For each RAC node :

As oracle user on node 1 :

Export DISPLAY=??????

Export ORACLE_SID=ASMDB1

Check if DBConsole is running : emctl status dbconsole

IF dbconsole not running

THEN 1/ execute : emctl start dbconsole
2/ access dbconsole using Internet browser : http://node1:1158/em using sys as user, connected as sysdba, with its password

IF dbconsole started and reachable with http://node1:1158/em

THEN dbconsole is OK on Node 1

ELSE See metalink note for DBConsole troubleshooting !!!

ELSE 1/ access dbconsole using Internet browser : http://node1:1158/em using sys as user, connected as sysdba, with its password

IF dbconsole is reachable with http://node1:1158/em

THEN dbconsole is OK on Node 1

ELSE See metalink note for DBConsole troubleshooting !!!

END

Troubleshooting Tips

The "dbconsole" may not start after doing an install through the Oracle Enterprise Manager (OEM) and selecting a database.

The solution is to:

Edit the file: "${ORACLE_HOME}/<hostname>_${ORACLE_SID}/sysman/config/emd.properties"

Locate the entry where the EMD_URL is set.

This entry should have the format:

EMD_URL=http://<hostname>:%EM_SERVLET_PORT%/emd/main

If you see the string: %EM_SERVLET_PORT% in the entry, then replace the complete string with an unused port number that is not defined in the "/etc/services" file. If this string is missing and no port number is in its place, then insert an unused port number that is not defined in the "/etc/services" file in between the "http://<hostname>:" and the "/emd/main" strings.

Use the command "emctl start dbconsole" to start the dbconsole after making this change.

For example:

EMD_URL=http://myhostname.us.oracle.com:5505/emd/main

20.3 MOVING FROM DBCONSOLE TO GRID CONTROL

You must either use DBCONSOLE, or GRID CONTROL, not both at the same time !!!

To move from Locally Managed (DBConsole) to Centrally Managed (Grid Control) :

- You must have an existing GRID Control, or install a new Grid Control on aseparate LPAR, or server. Grid Control is available on AIX5L, but could be installed on any supported operating system.
- You must install the Oracle Grid Agent on each RAC node AIX LPAR, with same unix user as Oracle Clusterware owner, and in the same oraInventory.
- **THEN you must follow this metalink note :**

Subject: **How to change a 10.2.0.x Database from Locally Managed to Centrally Managed** Doc ID: **Note:400476.1**

20.4 MANUAL DATABASE CREATION

Doc ID:	Note:240052.1	**Content Type:**	TEXT/PLAIN
Subject:	10g Manual Database Creation in Oracle (Single Instance and RAC)	**Creation Date:**	30-MAY-2003
Type:	BULLETIN	**Last Revision Date:**	07-DEC-2004
Status:	PUBLISHED		

PURPOSE

The purpose of this bulletin is to give an example of a manual database creation in 10g.
MANDATORY
==========
FROM 10g RELEASE WE HAVE SYSAUX TABLESPACE MANDATORY FOR STATISTICS WORKLOAD REPOSITORY FACILITIES (SWRF)
GOOD PRACTICE
==========
CREATE DEFAULT TABLESPACE, WHILE CREATING THE DATABASE. So whenever DBA will create a new user it will,
use the default permanent tablespace, unless DBA is mentioning the DEFAULT TABLESPACE clause while creating the user.
To use default tablespace option, it is mandatory to use the init.ora parameter
"Compatible must be >=10.0"

SCOPE & APPLICATION

Oracle recommends using the Database Configuration Assistant (DBCA) to create your database.
These steps are available for DBAs who want to manually create a
10g database either in single instance or Real Application Clusters mode.

20.4.1 Tips to create a database in 10g Real Application Clusters.

==========

Manual Database Creation steps for Real Application Clusters

==========

Here are the steps to be followed to create a Real Application Clusters database:

1. **Make a init<SID>.ora in your $ORACLE_HOME/dbs directory.** On Windows this file is in $ORACLE_HOME\database. To simplify, you can copy init.ora to init<SID>.ora and modify the file. Remember that your control file must be pointing to a pre-existing raw device or cluster file system location.
   *** Path names, file names, and sizes will need to be modified
   Example parameter settings for the first instance:

   Cluster-Wide Parameters for Database "RAC":

   db_block_size=8192
   db_cache_size=52428800
   background_dump_dest=/u01/32bit/app/oracle/product/9.0.1/rdbms/log
   core_dump_dest=/u01/32bit/app/oracle/product/9.0.1/rdbms/log
   user_dump_dest=/u01/32bit/app/oracle/product/9.0.1/rdbms/log
   timed_statistics=TRUE
   control_files=("/dev/RAC/control_01.ctl", "/dev/RAC/control_02.ctl")
   db_name=RAC
   shared_pool_size=52428800
   sort_area_size=524288
   undo_management=AUTO

cluster_database=true
cluster_database_instances=2
remote_listener=LISTENERS_RAC

Instance Specific Parameters for Instance "RAC1":

instance_name=RAC1
instance_number=1
local_listener=LISTENER_RAC1
thread=1
undo_tablespace=UNDOTBS

* The local_listener parameter requires that you first add the listener address to the TNSNAMES.ORA - remember to do so on both Node 1 and Node 2.

** You can also use an spfile as described in Note 136327.1.

2. Run the following sqlplus command to connect to the database:

sqlplus '/ as sysdba'

3. Startup up the database in NOMOUNT mode:

SQL> startup nomount

4. Create the Database (All raw devices must be pre-created) :
*** Path names, file names, and sizes will need to be modified

```
CREATE DATABASE <db_name>
CONTROLFILE REUSE
MAXDATAFILES 254
MAXINSTANCES 32
MAXLOGHISTORY 100
MAXLOGMEMBERS 5
MAXLOGFILES 64
DATAFILE '/dev/RAC/system_01_400.dbf' SIZE 900M segment space management auto
REUSE AUTOEXTEND ON NEXT  10240K MAXSIZE UNLIMITED
UNDO TABLESPACE "UNDOTBS" DATAFILE
'/dev/RAC/undotbs_01_210.dbf' SIZE 200M REUSE
DEFAULT TABLESPACE USER_DEFAULT DATAFILE
'/u01/oracle/rbdb1/user_default_1.dbf' size 2000M REUSE segment space management auto
SYSAUX DATAFILE '/u01/oracle/rbdb1/sysaux_1.dbf' size
500M REUSE segment space management auto
CHARACTER SET US7ASCII
LOGFILE GROUP 1 ('/dev/RAC/redo1_01_100.dbf') SIZE 100M REUSE,
GROUP 2 ('/dev/RAC/redo1_02_100.dbf') SIZE 100M REUSE;
```

5. Create a Temporary Tablespace:
*** Path names, file names, and sizes will need to be modified

```
CREATE TEMPORARY TABLESPACE "TEMP" TEMPFILE
'/dev/RAC/temp_01_50.dbf' SIZE 40M REUSE
```

6. Create a 2nd Undo Tablespace:
*** Path names, file names, and sizes will need to be modified

```
CREATE UNDO TABLESPACE "UNDOTBS2" DATAFILE
'/dev/RAC/undotbs_02_210.dbf' SIZE 200M REUSE
NEXT  5120K MAXSIZE UNLIMITED;
```

7. Run the necessary scripts to build views, synonyms, etc.:

The primary scripts that you must run are:
i> CATALOG.SQL--creates the views of data dictionary tables and the dynamic performance views
ii> CATPROC.SQL--establishes the usage of PL/SQL functionality and creates many of the PL/SQL Oracle supplied packages
iii> CATPARR.SQL--creates RAC specific views

8. Edit init<SID>.ora and set appropriate values for the 2nd instance on the 2nd Node:
*** Names may need to be modified

```
instance_name=RAC2
instance_number=2
local_listener=LISTENER_RAC2
thread=2
undo_tablespace=UNDOTBS2
```

9. From the first instance, run the following command:
*** Path names, file names, and sizes will need to be modified

```
alter database
add logfile thread 2
  group 3 ('/dev/RAC/redo2_01_100.dbf') size 100M,
  group 4 ('/dev/RAC/redo2_02_100.dbf') size 100M;
alter database enable public thread 2;
```

10. Start the second Instance. (Assuming that your cluster configuration is up and running)

20.4.2 Configure listener.ora / sqlnet.ora / tnsnames.ora

Use **netca** and/or **netmgr** to check the configuration of the listener and configure Oracle Net services (by default the Net service may be equal to the global database name (see instance parameter **service_names**).

20.4.3 Configure Oracle Enterprise Manager

Then start the OEM agent:
$agentctl start

- **Check /etc/oratab**

  The file should contain a reference to the database name, not to the instance name.
  The last field should always be “N” on a RAC environment to avoid 2 instances of the same name to be started.

- **Register the database with srvctl (this should not be necessary if the database was not created by DBCA)**

  srvctl add database –d <db_name> –o <ORACLE_HOME path>
  srvctl add instance –d <db_name> –i <SID1> –n <node1>
  srvctl add instance –d <db_name> –i <SID2> –n <node1>

21 APPENDIX A : ORACLE / IBM TECHNICAL DOCUMENTS

• **Technical Documents on Oracle Real Application Cluster :**

http://www.oracle.com/technology/products/database/clustering/index.html

•Technical Documents on Oracle Real Application Cluster :

•Note:282036.1 – Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries
•Note:302806.1 – IBM General Parallel File System (GPFS) and Oracle RAC on AIX 5L and IBM eServer pSeries
•Note:341507.1 – Oracle Products on Linux on IBM POWER
•Note:4044741 – Status of Certification of Oracle Clusterware with HACMP 5.3 & 5.4

•JSC Cookbooks (available at http://www.oracleracsig.org) :

•9i RAC Release 2 / AIX / HACMP / GPFS
•10g RAC Release 1 / AIX / GPFS / ASM / Concurrent Raw Devices on IBM SAN Storage
•10g RAC Release 2 / AIX / GPFS / ASM on IBM SAN Storage

•IBM Tech documents (http://www.ibm.com) :

•Implementing 10gRAC with ASM on AIX5L
•Oracle 10g RAC on AIX with Veritas Storage Foundation for Oracle RAC
•Oracle 9i RAC on AIX with VERITAS SFRAC
•Oracle's licensing in a multi-core environment
•Oracle Architecture and Tuning on AIX
•Diagnosing Oracle® Database Performance on AIX® Using IBM® NMON and Oracle Statspack Reports
•Observations Using SMT on an Oracle 9i OLTP Workload
•Oracle 9i & 10g on IBM AIX5L: Tips & Considerations
•Impact of Advanced Power Virtualization Features on Performance Characteristics of a Java® / C++ / Oracle® based Application
•Performance Impact of Upgrading a Data Warehouse Application from AIX 5.2 to AIX 5.3 with SMT
•Simultaneous Multi-threading (SMT) on IBM POWER5: Performance improvement on commercial OLTP workload using Oracle database version 10g

22 APPENDIX B : ORACLE TECHNICAL NOTES

This appendix provides some useful notes coming from Oracle support. These notes can be found in Metalink.

22.1 CRS AND 10G REAL APPLICATION CLUSTERS

Doc ID:	Note:259301.1	**Content Type:**	TEXT/X-HTML
Subject:	CRS and 10g Real Application Clusters	**Creation Date:**	05-DEC-2003
Type:	BULLETIN	**Last Revision Date:**	15-MAR-2005
Status:	PUBLISHED		

PURPOSE

This document is to provide additional information on CRS (Cluster Ready Services) in 10g Real Application Clusters.

SCOPE & APPLICATION

This document is intended for RAC Database Administrators and Oracle support enginneers.

CRS and 10g REAL APPLICATION CLUSTERS

CRS (Cluster Ready Services) is a new feature for 10g Real Application Clusters that provides a standard cluster interface on all platforms and performs new high availability operations not available in previous versions.

CRS KEY FACTS

Prior to installing CRS and 10g RAC, there are some key points to remember about CRS and 10g RAC:

- CRS is REQUIRED to be installed and running prior to installing 10g RAC.

- CRS can either run on top of the vendor clusterware (such as Sun Cluster, HP Serviceguard, IBM HACMP, TruCluster, Veritas Cluster, Fujitsu Primecluster, etc...) or can run without the vendor clusterware. The vendor clusterware was required in 9i RAC but is optional in 10g RAC.

- The CRS HOME and ORACLE_HOME must be installed in DIFFERENT locations.

- Shared Location(s) or devices for the Voting File and OCR (Oracle Configuration Repository) file must be available PRIOR to installing CRS. The voting file should be at least 20MB and the OCR file should be at least 100MB.

- CRS and RAC require that the following network interfaces be configured prior to installing CRS or RAC:
  - Public Interface
  - Private Interface
  - Virtual (Public) Interface
  For more information on this, see Note 264847.1.

- The root.sh script at the end of the CRS installation starts the CRS stack. If your CRS stack does not start, see Note 240001.1.

- Only one set of CRS daemons can be running per RAC node.

- On Unix, the CRS stack is run from entries in /etc/inittab with "respawn".

- If there is a network split (nodes loose communication with each other). One or more nodes may reboot automatically to prevent data corruption.

- The supported method to start CRS is booting the machine. MANUAL STARTUP OF THE CRS STACK IS NOT SUPPORTED UNTIL 10.1.0.4 OR HIGHER.
- The supported method to stop is shutdown the machine or use "init.crs stop".

- Killing CRS daemons is not supported unless you are removing the CRS installation via Note 239998.1 because flag files can become mismatched.

- For maintenance, go to single user mode at the OS.

Once the stack is started, you should be able to see all of the daemon processes with a ps -ef command:

```
[rac1]/u01/home/beta> ps -ef | grep crs

 oracle  1363   999  0 11:23:21 ?  0:00 /u01/crs_home/bin/evmlogger.bin -o /u01
 oracle   999     1  0 11:21:39 ?  0:01 /u01/crs_home/bin/evmd.bin
 root    1003     1  0 11:21:39 ?  0:01 /u01/crs_home/bin/crsd.bin
 oracle  1002     1  0 11:21:39 ?  0:01 /u01/crs_home/bin/ocssd.bin
```

CRS DAEMON FUNCTIONALITY

Here is a short description of each of the CRS daemon processes:

CRSD:
- Engine for HA operation
- Manages 'application resources'
- Starts, stops, and fails 'application resources' over
- Spawns separate 'actions' to start/stop/check application resources
- Maintains configuration profiles in the OCR (Oracle Configuration Repository)
- Stores current known state in the OCR.
- Runs as root
- Is restarted automatically on failure

OCSSD:
- OCSSD is part of RAC and Single Instance with ASM
- Provides access to node membership
- Provides group services
- Provides basic cluster locking
- Integrates with existing vendor clusteware, when present
- Can also runs without integration to vendor clustware
- Runs as Oracle.
- Failure exit causes machine reboot.
--- This is a feature to prevent data corruption in event of a split brain.

EVMD:
- Generates events when things happen
- Spawns a permanent child evmlogger
- Evmlogger, on demand, spawns children
- Scans callout directory and invokes callouts.
- Runs as Oracle.
- Restarted automatically on failure

CRS LOG DIRECTORIES

When troubleshooting CRS problems, it is important to review the directories under the CRS Home.

$ORA_CRS_HOME/crs/log - This directory includes traces for CRS resources that are joining, leaving, restarting, and relocating as identified by CRS.

$ORA_CRS_HOME/crs/init - Any core dumps for the crsd.bin daemon should be written here. Note 1812.1 could be used to debug these.

$ORA_CRS_HOME/css/log - The css logs indicate all actions such as reconfigurations, missed checkins , connects, and disconnects from the client CSS listener . In some cases the logger logs messages with the category of (auth.crit) for the reboots done by oracle. This could be used for checking the exact time when the reboot occured.

$ORA_CRS_HOME/css/init - Core dumps from the ocssd primarily and the pid for the css daemon whose death is treated as fatal are located here. If there are abnormal restarts for css then the core files will have the formats of core.<pid>. Note 1812.1 could be used to debug these.

$ORA_CRS_HOME/evm/log - Log files for the evm and evmlogger daemons. Not used as often for debugging as the CRS and CSS directories.

$ORA_CRS_HOME/evm/init - Pid and lock files for EVM. Core files for EVM should also be written here. Note 1812.1 could be used to debug these.

$ORA_CRS_HOME/srvm/log - Log files for OCR.

STATUS FOR CRS RESOURCES

After installing RAC and running the VIPCA (Virtual IP Configuration Assistant) launched with the RAC root.sh, you should be able to see all of your CRS resources with crs_stat. Example:

```
cd $ORA_CRS_HOME/bin
./crs_stat

NAME=ora.rac1.gsd
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.oem
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.ons
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.vip
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.gsd
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.oem
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.ons
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.vip
TYPE=application
TARGET=ONLINE
STATE=ONLINE
```

There is also a script available to view CRS resources in a format that is easier to read. Just create a shell script with:

---------- **Begin Shell Script** ----------

```
#!/usr/bin/ksh
#
# Sample 10g CRS resource status query script
#
# Description:
#    - Returns formatted version of crs_stat -t, in tabular
#      format, with the complete rsc names and filtering keywords
#   - The argument, $RSC_KEY, is optional and if passed to the script, will
#     limit the output to HA resources whose names match $RSC_KEY.
# Requirements:
#   - $ORA_CRS_HOME should be set in your environment

RSC_KEY=$1
QSTAT=-u
AWK=/usr/xpg4/bin/awk    # if not available use /usr/bin/awk

# Table header:echo ""
$AWK \
  'BEGIN {printf "%-45s %-10s %-18s\n", "HA Resource", "Target", "State";
          printf "%-45s %-10s %-18s\n", "-----------", "------", "-----";}'

# Table body:
$ORA_CRS_HOME/bin/crs_stat $QSTAT | $AWK \
 'BEGIN { FS="="; state = 0; }
  $1~/NAME/ && $2~/'$RSC_KEY'/ {appname = $2; state=1};
  state == 0 {next;}
  $1~/TARGET/ && state == 1 {apptarget = $2; state=2;}
  $1~/STATE/ && state == 2 {appstate = $2; state=3;}
  state == 3 {printf "%-45s %-10s %-18s\n", appname, apptarget, appstate; state=0;}'
```

---------- **End Shell Script** ----------

Example output:

```
[opcbsol1]/u01/home/usupport> ./crsstat
HA Resource                                   Target     State
-----------                                   ------     -----
ora.V10SN.V10SN1.inst                         ONLINE     ONLINE on opcbsol1
ora.V10SN.V10SN2.inst                         ONLINE     ONLINE on opcbsol2
ora.V10SN.db                                  ONLINE     ONLINE on opcbsol2
ora.opcbsol1.ASM1.asm                         ONLINE     ONLINE on opcbsol1
ora.opcbsol1.LISTENER_OPCBSOL1.lsnr           ONLINE     ONLINE on opcbsol1
ora.opcbsol1.gsd                              ONLINE     ONLINE on opcbsol1
ora.opcbsol1.ons                              ONLINE     ONLINE on opcbsol1
ora.opcbsol1.vip                              ONLINE     ONLINE on opcbsol1
ora.opcbsol2.ASM2.asm                         ONLINE     ONLINE on opcbsol2
ora.opcbsol2.LISTENER_OPCBSOL2.lsnr           ONLINE     ONLINE on opcbsol2
ora.opcbsol2.gsd                              ONLINE     ONLINE on opcbsol2
ora.opcbsol2.ons                              ONLINE     ONLINE on opcbsol2
ora.opcbsol2.vip                              ONLINE     ONLINE on opcbsol2
```

CRS RESOURCE ADMINISTRATION

You can use srvctl to manage these resources. Below are syntax and examples.

CRS RESOURCE STATUS

```
srvctl status database -d <database-name> [-f] [-v] [-S <level>]
srvctl status instance -d <database-name> -i <instance-name> >[,<instance-name-list>]
       [-f] [-v] [-S <level>]
srvctl status service -d <database-name> -s <service-name>[,<service-name-list>]
       [-f] [-v] [-S <level>]
srvctl status nodeapps [-n <node-name>]
srvctl status asm -n <node_name>
```

EXAMPLES:

```
Status of the database, all instances and all services.
        srvctl status database -d ORACLE -v
Status of named instances with their current services.
        srvctl status instance  -d ORACLE -i RAC01, RAC02 -v
Status of a named services.
        srvctl status service -d ORACLE -s ERP  -v
Status of all nodes supporting database applications.
        srvctl status node
```

START CRS RESOURCES

```
srvctl start database -d <database-name> [-o < start-options>]
       [-c <connect-string> | -q]
srvctl start instance -d <database-name> -i <instance-name>
       [,<instance-name-list>] [-o <start-options>] [-c <connect-string> | -q]
srvctl start service -d <database-name> [-s <service-name>[,<service-name-list>]]
       [-i <instance-name>]  [-o <start-options>] [-c <connect-string> | -q]
srvctl start nodeapps -n <node-name>
srvctl start asm -n <node_name> [-i <asm_inst_name>] [-o <start_options>]
```

EXAMPLES:

```
Start the database with all enabled instances.
        srvctl start database -d ORACLE
Start named instances.
        srvctl start instance  -d ORACLE -i RAC03, RAC04
Start named services.  Dependent instances are started as needed.
        srvctl start service -d ORACLE -s CRM
Start a service at the named instance.
        srvctl start  service -d ORACLE -s CRM -i RAC04
Start node applications.
        srvctl start  nodeapps -n myclust-4
```

STOP CRS RESOURCES

```
srvctl stop database -d <database-name> [-o <stop-options>]
       [-c <connect-string> | -q]
srvctl stop instance -d <database-name> -i <instance-name> [,<instance-name-list>]
       [-o <stop-options>][-c <connect-string> | -q]
srvctl stop service -d <database-name> [-s <service-name>[,<service-name-list>]]
       [-i <instance-name>][-c <connect-string> | -q] [-f]
srvctl stop nodeapps -n <node-name>
srvctl stop asm -n <node_name> [-i <asm_inst_name>] [-o <start_options>]
```

EXAMPLES:

```
Stop the database, all instances and all services.
        srvctl stop database -d ORACLE
Stop named instances, first relocating all existing services.
        srvctl stop instance  -d ORACLE -i RAC03,RAC04
Stop the service.
        srvctl stop service -d ORACLE -s CRM
Stop the service at the named instances.
        srvctl stop  service -d ORACLE -s CRM -i RAC04
Stop node applications.  Note that instances and services also stop.
        srvctl stop  nodeapps -n myclust-4
```

ADD CRS RESOURCES

```
srvctl add database -d <name> -o <oracle_home> [-m <domain_name>] [-p <spfile>]
       [-A <name|ip>/netmask] [-r {PRIMARY | PHYSICAL_STANDBY | LOGICAL_STANDBY}]
       [-s <start_options>] [-n <db_name>]
srvctl add instance -d <name> -i <inst_name> -n <node_name>
srvctl add service -d <name> -s <service_name> -r <preferred_list>
       [-a <available_list>] [-P <TAF_policy>] [-u]
srvctl add nodeapps -n <node_name> -o <oracle_home>
       [-A <name|ip>/netmask[/if1[|if2|...]]]
srvctl add asm -n <node_name> -i <asm_inst_name> -o <oracle_home>
```

OPTIONS:

```
-A      vip range, node, and database, address specification. The format of
        address string is:
        [<logical host name>]/<VIP address>/<net mask>[/<host interface1[ |
        host interface2 |..]>] [,] [<logical host name>]/<VIP address>/<net mask>
        [/<host interface1[ | host interface2 |..]>]
-a      for services, list of available instances, this list cannot include
        preferred instances
-m      domain name with the format “us.mydomain.com”
-n      node name that will support one or more instances
-o      $ORACLE_HOME to locate Oracle binaries
-P      for services, TAF preconnect policy - NONE, PRECONNECT
-r      for services, list of preferred instances, this list cannot include
        available instances.
-s      spfile name
-u      updates the preferred or available list for the service to support the
        specified instance. Only one instance may be specified with the -u
        switch.  Instances that already support the service should not be
        included.
```

EXAMPLES:

```
Add a new node:
        srvctl add nodeapps -n myclust-1 -o $ORACLE_HOME  –A
        139.184.201.1/255.255.255.0/hme0
Add a new database.
        srvctl add  database  -d ORACLE -o $ORACLE_HOME
Add named instances to an existing database.
        srvctl add instance -d ORACLE -i RAC01 -n myclust-1
        srvctl add instance -d ORACLE -i RAC02 -n myclust-2
        srvctl add instance -d ORACLE -i RAC03 -n myclust-3
Add a service to an existing database with preferred instances (-r) and
available instances (-a).  Use basic failover to the available instances.
        srvctl add service -d ORACLE -s STD_BATCH -r RAC01,RAC02 -a RAC03,RAC04
Add a service to an existing database with preferred instances in list one and
available instances in list two. Use preconnect at the available instances.
        srvctl add service -d ORACLE -s STD_BATCH -r RAC01,RAC02 -a RAC03,RAC04  -P
PRECONNECT
```

REMOVE CRS RESOURCES

```
srvctl remove database -d <database-name>
srvctl remove instance  -d <database-name> [-i <instance-name>]
srvctl remove service -d <database-name> -s <service-name> [-i <instance-name>]
srvctl remove nodeapps -n <node-name>
```

EXAMPLES:

```
Remove the applications for a database.
        srvctl remove database  -d ORACLE
Remove the applications for named instances of an existing database.
        srvctl remove instance -d ORACLE -i  RAC03
        srvctl remove instance -d ORACLE -i  RAC04
Remove the service.
        srvctl remove  service -d ORACLE -s STD_BATCH
Remove the service from the instances.
        srvctl remove  service  -d ORACLE -s STD_BATCH -i RAC03,RAC04
Remove all node applications from a node.
        srvctl remove  nodeapps -n myclust-4
```

MODIFY CRS RESOURCES

```
srvctl modify database -d <name> [-n <db_name] [-o <ohome>] [-m <domain>]
       [-p <spfile>]  [-r {PRIMARY | PHYSICAL_STANDBY | LOGICAL_STANDBY}]
       [-s <start_options>]
srvctl modify instance -d <database-name> -i <instance-name> -n <node-name>
srvctl modify instance -d <name> -i <inst_name> {-s <asm_inst_name> | -r}
srvctl modify service -d <database-name> -s <service_name> -i <instance-name>
       -t <instance-name> [-f]
srvctl modify service -d <database-name> -s <service_name> -i <instance-name>
       -r  [-f]
srvctl modify nodeapps -n <node-name> [-A <address-description> ] [-x]
```

OPTIONS:

```
-i <instance-name> -t <instance-name>  the instance name (-i) is replaced by the
   instance name (-t)
-i <instance-name> -r the named instance is modified to be a preferred instance
-A address-list for VIP application, at node level
-s <asm_inst_name> add or remove ASM dependency
```

EXAMPLES:

```
Modify an instance to execute on another node.
        srvctl modify  instance  -d ORACLE  -n myclust-4
Modify a service to execute on another node.
        srvctl modify service -d ORACLE  -s HOT_BATCH -i  RAC01 -t RAC02
Modify an instance to be a preferred instance for a service.
        srvctl modify service -d ORACLE  -s HOT_BATCH -i  RAC02 –r
```

RELOCATE SERVICES

```
srvctl relocate service -d <database-name> -s <service-name> [-i <instance-name >]-
t<instance-name > [-f]
```

EXAMPLES:

```
Relocate a service from one instance to another
        srvctl relocate  service -d ORACLE -s CRM -i RAC04 -t RAC01
```

ENABLE CRS RESOURCES (The resource may be up or down to use this function)

```
srvctl enable database -d <database-name>
srvctl enable instance -d <database-name> -i <instance-name> [,<instance-name-list>]
srvctl enable service -d <database-name> -s <service-name>] [, <service-name-list>]
[-i <instance-name>]
```

EXAMPLES:

```
Enable the database.
        srvctl enable database -d ORACLE
Enable the named instances.
        srvctl enable instance  -d ORACLE -i RAC01, RAC02
Enable the service.
        srvctl enable  service -d ORACLE -s ERP,CRM
Enable the service at the named instance.
        srvctl enable  service -d ORACLE -s CRM -i RAC03
```

DISABLE CRS RESOURCES (The resource must be down to use this function)

```
srvctl disable database -d <database-name>
srvctl disable instance -d <database-name> -i <instance-name> [,<instance-name-list>]
srvctl disable service -d <database-name> -s <service-name>] [,<service-name-list>]
[-i <instance-name>]
```

EXAMPLES:

```
Disable the database globally.
        srvctl disable database -d ORACLE
Disable the named instances.
        srvctl disable instance  -d ORACLE -i RAC01, RAC02
Disable the service globally.
        srvctl disable  service -d ORACLE -s ERP,CRM
Disable the service at the named instance.
        srvctl disable  service -d ORACLE  -s CRM -i RAC03,RAC04
```

For more information on this see the Oracle10g Real Application Clusters Administrator's Guide - Appendix B

RELATED DOCUMENTS

Oracle10g Real Application Clusters Installation and Configuration
Oracle10g Real Application Clusters Administrator's Guide

22.2 About RAC ...

282036.1 - Minimum software versions and patches required to Support Oracle Products on ...
283743.1 - Pre-Install checks for 10g RDBMS on AIX
220970.1 - RAC: Frequently Asked Questions
183408.1 - Raw Devices and Cluster Filesystems With Real Application Clusters
293750.1 - 10g Installation on Aix 5.3, Failed with Checking operating system version mu...

22.3 About CRS ...

263897.1 - 10G: How to Stop the Cluster Ready Services (CRS)
295871.1 - How to verify if CRS install is Valid
265769.1 - 10g RAC: Troubleshooting CRS Reboots
259301.1 - CRS and 10g Real Application Clusters
268937.1 - Repairing or Restoring an Inconsistent OCR in RAC
293819.1 - Placement of voting and OCR disk files in 10gRAC
239998.1 - 10g RAC: How to Clean Up After a Failed CRS Install
272332.1 - CRS 10g Diagnostic Collection Guide
279793.1 - How to Restore a Lost Voting Disk in 10g
239989.1 - 10g RAC: Stopping Reboot Loops When CRS Problems Occur
298073.1 - HOW TO REMOVE CRS AUTO START AND RESTART FOR A RAC INSTANCE
298069.1 - HOW TO REMOVE CRS AUTO START AND RESTART FOR A RAC INSTANCE
284949.1 - CRS Home Is Only Partially Copied to Remote Node
285046.1 - How to recreate ONS,GSD,VIP deleted from ocr by crs_unregister

22.4 About VIP ...

296856.1 - Configuring the IBM AIX 5L Operating System for the Oracle 10g VIP
294336.1 - Changing the check interval for the Oracle 10g VIP
276434.1 - Modifying the VIP of a Cluster Node
298895.1 - Modifying the default gateway address used by the Oracle 10g VIP
264847.1 - How to Configure Virtual IPs for 10g RAC

22.5 About manual database cration ...

240052.1 - 10g Manual Database Creation in Oracle (Single Instance and RAC)

22.6 About Grid Control ...

284707.1 - Enterprise Manager Grid Control 10.1.0.3.0 Release Notes

277420.1 - EM 10G Grid Control Preinstall Steps for AIX 5.2

22.7 About TAF ...

271297.1 - Troubleshooting TAF Issues in 10g RAC

22.8 About Adding/Removing Node ...

269320.1 - Removing a Node from a 10g RAC Cluster

270512.1 - Adding a Node to a 10g RAC Cluster

22.9 About ASM ...

243245.1 - 10G New Storage Features and Enhancements

268481.1 - Re-creating ASM Instances and Diskgroups

282777.1 - SGA sizing for ASM instances and databases that use ASM

274738.1 - Creating an ASM-enabled Database

249992.1 - New Feature on ASM (Automatic Storage Manager).

252219.1 - Steps To Migrate Database From Non-ASM to ASM And Vice-Versa

293234.1 - How To Move Archive Files from ASM

270066.1 - Manage ASM instance-creating diskgroup,adding/dropping/resizing disks.

300472.1 - How To Delete Archive Log Files Out Of +Asm?

265633.1 - ASM Technical Best Practices http://metalink.oracle.com/metalink/plsql/docs/ASM.pdf

For full article, download Automatic Storage Management (154K/pdf)

294869.1 - Oracle ASM and Multi-Pathing Technologies

22.10 METALINK NOTE TO USE IN CASE OF PROBLEM WITH CRS ...

263897.1 - 10G: How to Stop the Cluster Ready Services (CRS)
295871.1 - How to verify if CRS install is Valid
265769.1 - 10g RAC: Troubleshooting CRS Reboots
259301.1 - CRS and 10g Real Application Clusters
268937.1 - Repairing or Restoring an Inconsistent OCR in RAC
293819.1 - Placement of voting and OCR disk files in 10gRAC
239998.1 - 10g RAC: How to Clean Up After a Failed CRS Install
272332.1 - CRS 10g Diagnostic Collection Guide
239989.1 - 10g RAC: Stopping Reboot Loops When CRS Problems Occur
298073.1 - HOW TO REMOVE CRS AUTO START AND RESTART FOR A RAC INSTANCE
298069.1 - HOW TO REMOVE CRS AUTO START AND RESTART FOR A RAC INSTANCE
284949.1 - CRS Home Is Only Partially Copied to Remote Node

23 APPENDIX C : USEFULL COMMANDS

crsctl

Usage:

```
       crsctl check  crs          - checks the viability of the CRS stack
       crsctl check  cssd         - checks the viability of CSS
       crsctl check  crsd         - checks the viability of CRS
       crsctl check  evmd         - checks the viability of EVM
       crsctl set    css <parameter> <value> - sets a parameter override
       crsctl get    css <parameter> - gets the value of a CSS parameter
       crsctl unset  css <parameter> - sets CSS parameter to its default
       crsctl query  css votedisk    - lists the voting disks used by CSS
       crsctl add    css votedisk <path> - adds a new voting disk
       crsctl delete css votedisk <path> - removes a voting disk
       crsctl enable  crs    - enables startup for all CRS daemons
       crsctl disable crs    - disables startup for all CRS daemons
       crsctl start crs  - starts all CRS daemons.
       crsctl stop  crs  - stops all CRS daemons. Stops CRS resources in case of
cluster.
       crsctl start resources  - starts CRS resources.
       crsctl stop resources  - stops  CRS resources.
       crsctl debug statedump evm  - dumps state info for evm objects
       crsctl debug statedump crs  - dumps state info for crs objects
       crsctl debug statedump css  - dumps state info for css objects
       crsctl debug log css [module:level]{,module:level} ...
                           - Turns on debugging for CSS
       crsctl debug trace css - dumps CSS in-memory tracing cache
       crsctl debug log crs [module:level]{,module:level} ...
                           - Turns on debugging for CRS
       crsctl debug trace crs - dumps CRS in-memory tracing cache
       crsctl debug log evm [module:level]{,module:level} ...
                           - Turns on debugging for EVM
       crsctl debug trace evm - dumps EVM in-memory tracing cache
       crsctl debug log res <resname:level> turns on debugging for resources
       crsctl query crs softwareversion [<nodename>] - lists the version of CRS
software installed
       crsctl query crs activeversion - lists the CRS software operating version
       crsctl lsmodules css - lists the CSS modules that can be used for debugging
       crsctl lsmodules crs - lists the CRS modules that can be used for debugging
       crsctl lsmodules evm - lists the EVM modules that can be used for debugging

 If necesary any of these commands can be run with additional tracing by
 adding a "trace" argument at the very front.
 Example: crsctl trace check css
```

ocrconfig

```
Name:
        ocrconfig - Configuration tool for Oracle Cluster Registry.

Synopsis:
        ocrconfig [option]
        option:
                -export <filename> [-s online]      - Export cluster register contents to a file
                -import <filename>                  - Import cluster registry contents from a
file
                -upgrade [<user> [<group>]]
                                                    - Upgrade cluster registry from previous
version
                -downgrade [-version <version string>]
                                                    - Downgrade cluster registry to the
specified version
                -backuploc <dirname>                - Configure periodic backup location
                -showbackup                         - Show backup information
                -restore <filename>                 - Restore from physical backup
                -replace ocr|ocrmirror [<filename>] - Add/replace/remove a OCR device/file
                -overwrite                          - Overwrite OCR configuration on disk
                -repair ocr|ocrmirror <filename>    - Repair local OCR configuration
                -help                               - Print out this help information

Note:
        A log file will be created in
        $ORACLE_HOME/log/<hostname>/client/ocrconfig_<pid>.log. Please ensure
        you have file creation privileges in the above directory before
        running this tool.
```

ocrdump

```
ocrdump -help
Name:
        ocrdump - Dump contents of Oracle Cluster Registry to a file.
Synopsis:
        ocrdump [<filename>|-stdout] [-backupfile <backupfilename>] [-keyname <keyname>] [-xml]
[-noheader]
Description:
        Default filename is OCRDUMPFILE. Examples are:
        prompt> ocrdump
        writes cluster registry contents to OCRDUMPFILE in the current directory
        prompt> ocrdump MYFILE
        writes cluster registry contents to MYFILE in the current directory
        prompt> ocrdump -stdout -keyname SYSTEM
        writes the subtree of SYSTEM in the cluster registry to stdout
        prompt> ocrdump -stdout -xml
        writes cluster registry contents to stdout in xml format
Notes:
        The header information will be retrieved based on best effort basis.
        A log file will be created in
     ORACLE_HOME/log/<hostname>/client/ocrdump_<pid>.log.
        Make sure you have file
        creation privileges in the above directory before
        running this tool.
```

srvctl

```
Usage:
srvctl <command> <object> [<options>]
    command:
enable|disable|start|stop|relocate|status|add|remove|modify|getenv|setenv|unsetenv|config
    objects: database|instance|service|nodeapps|asm|listener
For detailed help on each command and object and its options use:
    srvctl <command> <object> -h
```

crs_stat –help

```
Usage:  crs_stat [resource_name [...]] [-v] [-l] [-q] [-c cluster_member]
        crs_stat [resource_name [...]] -t [-v] [-q] [-c cluster_member]
        crs_stat -p [resource_name [...]] [-q]
        crs_stat [-a] application -g
        crs_stat [-a] application -r [-c cluster_member]
        crs_stat -f [resource_name [...]] [-q] [-c cluster_member]
        crs_stat -ls [resource_name [...]] [-q]
```

Cluvfy

```
Usage:
       cluvfy [ -help ]
       cluvfy stage { -list | -help }
       cluvfy stage {-pre|-post} <stage-name> <stage-specific options>  [-verbose]
       cluvfy comp  { -list | -help }
       cluvfy comp  <component-name> <component-specific options>  [-verbose]
```

cluvfy stage –list

```
Usage:
       cluvfy stage {-pre|-post} <stage-name> <stage-specific options>  [-verbose]

       Valid stage options and stage names are:
              -post hwos    :  post-check for hardware and operating system
              -pre  cfs     :  pre-check for CFS setup
              -post cfs     :  post-check for CFS setup
              -pre  crsinst :  pre-check for CRS installation
              -post crsinst :  post-check for CRS installation
              -pre  dbinst  :  pre-check for database installation
              -pre  dbcfg   :  pre-check for database configuration
```

cluvfy stage -help

```
Usage:
       cluvfy stage {-pre|-post} <stage-name> <stage-specific options>  [-verbose]

       SYNTAX (for Stages):
       cluvfy stage -post hwos -n <node_list>  [ -s <storageID_list>  ]  [-verbose]
       cluvfy stage -pre cfs -n <node_list> -s <storageID_list>  [-verbose]
       cluvfy stage -post cfs -n <node_list>  -f <file_system>   [-verbose]
       cluvfy stage -pre crsinst -n <node_list>   [-r { 10gR1 | rdbms } ]
                                [ -c <ocr_location> ]  [ -q <voting_disk> ]
                                [ -osdba <osdba_group> ]
                                [ -orainv <orainventory_group> ]  [-verbose]
       cluvfy stage -post crsinst -n <node_list>  [-verbose]
       cluvfy stage -pre dbinst -n <node_list>   [-r { 10gR1 | rdbms } ]
                               [ -osdba <osdba_group> ]  [-verbose]
       cluvfy stage -pre dbcfg -n <node_list>  -d <oracle_home>  [-verbose]
```

Cluvfy comp -list

Usage:

```
        cluvfy comp  <component-name> <component-specific options>  [-verbose]

        Valid components are:
                nodereach : checks reachability between nodes
                nodecon   : checks node connectivity
                cfs       : checks CFS integrity
                ssa       : checks shared storage accessibility
                space     : checks space availability
                sys       : checks minimum system requirements
                clu       : checks cluster integrity
                clumgr    : checks cluster manager integrity
                ocr       : checks OCR integrity
                crs       : checks CRS integrity
                nodeapp   : checks node applications existence
                admprv    : checks administrative privileges
                peer      : compares properties with peers
```

Cluvfy comp -help

Usage :

```
cluvfy comp  <component-name> <component-specific options>  [-verbose]

SYNTAX (for Components):
cluvfy comp nodereach -n <node_list>  [ -srcnode <node>  ]  [-verbose]
cluvfy comp nodecon -n <node_list>  [ -i <interface_list>  ]  [-verbose]
cluvfy comp cfs  [ -n <node_list> ] -f <file_system>   [-verbose]
cluvfy comp ssa  [ -n <node_list> ]  [ -s <storageID_list> ]  [-verbose]
cluvfy comp space  [ -n <node_list> ] -l <storage_location>
                -z <disk_space> {B|K|M|G}  [-verbose]
cluvfy comp sys  [ -n <node_list> ] -p { crs | database }  [-r { 10gR1 | rdbms } ]
                 [ -osdba <osdba_group> ]  [ -orainv <orainventory_group> ]  [-
verbose]
cluvfy comp clu  [ -n <node_list> ]  [-verbose]
cluvfy comp clumgr  [ -n <node_list> ]  [-verbose]
cluvfy comp ocr  [ -n <node_list> ]  [-verbose]
cluvfy comp crs  [ -n <node_list> ]  [-verbose]
cluvfy comp nodeapp  [ -n <node_list> ]  [-verbose]
cluvfy comp admprv  [ -n <node_list> ]  [-verbose]
                -o user_equiv  [-sshonly]
                -o crs_inst  [-orainv <orainventory_group> ]
                -o db_inst  [-osdba <osdba_group> ]
                -o db_config  -d <oracle_home>
cluvfy comp peer [ -refnode <node>  ] -n <node_list>   [-r { 10gR1 | rdbms } ]
                [ -orainv <orainventory_group> ] [ -osdba <osdba_group> ]  [-verbose]
```

24 APPENDIX D : EMPTY TABLES TO USE FOR INSTALLATION

For Network Layout

	Public	VIP (Same as Public)	RAC Interconnect	RAC Interconnect Backup
Network card On each node	en….	en….	en….	en….

For Nodes Name and IP Identification

Public		VIP		RAC Interconnect (Private Network)		RAC Interconnect Backup	
Network card On each node							
en….		en… (same as public)		en…		en…	
Node Name	IP	Node Name	IP	Node Name	IP	Node Name	IP

For ASM Disks implementation

Disks	LUN's Number	Device Name	Node 1	Major Num.	Minor Num.	Node 2	Major Num.	Minor Num.
Disk for OCR 1								
Disk for OCR 2								
Disk for Voting 1								
Disk for Voting 2								
Disk for Voting 3								
Disk 1 for ASM								
Disk 2 for ASM								
Disk 3 for ASM								
Disk 4 for ASM								
Disk 5 for ASM								
Disk 6 for ASM								
Disk 7 for ASM								
Disk 8 for ASM								
Disk 9 for ASM								
Disk 10 for ASM								
…	…		…			…		