

COOKBOOK



Quick Installation Guide

Oracle 10g RAC Release 2 on IBM SERVER pSeries Running AIX 5L With SAN Storage

**Version 1.0
January 2006**

**European ORACLE / IBM
Joint Solutions Center**

IBM – Thierry Plumeau
European – Oracle/IBM Joint Solutions Center

Oracle - Frederic Michiara
European – Oracle/IBM Joint Solutions Center

**This document is based on our experiences.
 This is not an official (Oracle or IBM) documentation.
 This document will 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
1.00	January 2006	Creation	Frederic Michiara Thierry Plumeau	Fabienne Lepetit Paul Bramey Alain Roy Marc Tu-Duy-Kiem

Contributors :

- o **Thierry Plumeau : IBM France** (European ORACLE/IBM Joint Solutions Center)
- o **Paul Bramey : Oracle Corp.** (European ORACLE/IBM Joint Solutions Center)
- o **Alain Roy : IBM France** (European ORACLE/IBM Joint Solutions Center)
- o **Fabienne Lepetit : Oracle France** (European ORACLE/IBM Joint Solutions Center)
- o **Marc Tu-Duy-Kiem : IBM France** (European ORACLE/IBM Joint Solutions Center)
- o **Frederic Michiara : Oracle France** (European ORACLE/IBM Joint Solutions Center)
- o **Stephan Navaro : IBM France**

Contact :

- o European **ORACLE / IBM** Joint Solutions Center → oraclibm@fr.ibm.com

1	The aim of this document	6
2	What's new with RAC implementation on AIX5L	8
3	Ways to implement 10g RAC on AIX5L.....	9
3.1	10g RAC over Local Area Network.....	9
3.1.1	ASM (Automated Storage Management)	9
3.1.2	IBM GPFS (General Parallel File System).....	9
3.1.3	Concurrent Raw Devices (HACMP)	10
3.1.4	NAS (Network Appliance Filer).....	10
3.2	10g RAC over MAN (Metropolitan Area Network)	10
4	Hardware architecture	11
4.1	IBM pSeries servers.....	12
4.2	OPERATING SYSTEM	12
4.3	Network Infrastructure.....	13
4.3.1	Network Architecture Examples.....	13
4.3.2	EtherChannel	13
4.3.3	AIX Virtual I/O Network.....	13
4.4	AIX Virtual I/O disks	14
4.5	SAN Storage	14
4.5.1	IBM.....	14
4.5.1.1	IBM Multi-Pathing :	15
4.5.2	EMC	17
4.5.3	HITACHI.....	17
4.5.4	Others	17
5	Installation steps	18
5.1	RAC installation steps.....	18
5.2	Database creation steps	19
6	CHECK LIST to use and follow	20
7	Preparing the system	21
7.1	Hardware requirements	21
7.2	Aix Software requirements.....	22
7.2.1	Filesets / APAR's Requirements for 10gRAC R2 / ASM (NO HACMP)	23
7.2.2	Filesets / APAR's Requirements for 10gRAC R2 / GPFS (NO HACMP)	24
7.3	Oracle Software Requirements.....	25
7.4	Users and Groups.....	26
7.5	Configure Kernel Parameters and Shell Limits.....	27
7.5.1	Configure Shell Limits.....	27
7.5.2	Configure System Configuration Parameters	28
7.5.3	Configure Network Tuning Parameters	28
7.6	NETWORK CONFIGURATION	30
7.6.1	Possible network configuration layout of our pSeries RAC cluster :	30
7.6.1.1	Network implementation with 2 Nodes	30
7.6.1.2	Network implementation with 3 Nodes, and more.....	30
7.6.1.3	Network implementation with IBM GPFS, 2 Nodes.....	31
7.6.1.4	Network implementation with IBM GPFS, 3 Nodes and more	31
7.6.2	Configuration steps to follow.....	32
7.6.2.1	Network Card Identification	32
7.6.2.2	Nodes Name and IP Identification.....	32
7.6.2.3	Host file Setup	33
7.6.2.4	Defining Default gateway on public interface	33
7.6.2.5	User Equivalence Setup	34
7.7	ORACLE environment setup.....	40
7.8	Local disk for Oracle code (Oracle Clusterware and RAC software)	40
8	Choose a Storage Option for Oracle CRS, Database, and Recovery Files	41
8.1	ASM Implementation.....	42
8.1.1	LUN's creation	43
8.1.2	Register LUN's at AIX level	43
8.1.3	Preparing Raw Disks for CRS/Voting disks.....	44
8.1.4	Preparing Disks for ASM	52

8.1.5	Recommandations, hints and tips.....	61
8.2	GPFS implementation	63
8.2.1	Network Infrastructure for GPFS implementation, with 2 nodes	64
8.2.2	GPFS New Single Node Quorum Support (Tiebreaker Disk).....	64
8.2.3	LUN's creation	65
8.2.4	Register LUN's at AIX level	65
8.2.5	Configure RSH for GPFS with root user.....	66
8.2.6	INSTALLING GPFS	66
8.2.7	Creating a 2 nodes GPFS cluster using tie-breaker disk and a GPFS filesystem for datafiles.....	66
8.2.8	Creating a GPFS filesystem	68
8.2.9	Tuning suggestions.....	71
9	Synchronize the System Time on Cluster Nodes.....	72
10	Important TIPS for Oracle software and patches Installation Instructions.....	73
10.1	10g Installation on Aix 5.3, Failed with Checking operating system version must be 5200	73
10.2	10g recommended steps before installation and applying any Patch Set on AIX	74
11	Cluster VeriFication utility	75
11.1	Understanding and Using Cluster Verification Utility.....	75
11.2	Using CVU to Determine if Installation Prerequisites are Complete	75
12	Install the Custer Ready Services (CRS)	86
12.1	CRS Post-Installation Tasks.....	104
12.1.1	Update Oracle user .profile.....	104
12.1.2	Cluster Ready Services Health Check	105
12.1.3	Interconnect Network configuration Checkup.....	107
12.1.4	Oracle CLuster Registry content Check and Backup	108
12.2	Some usefull commands	109
12.3	Accessing CRS logs	109
12.4	What Has Been Done ?.....	109
12.5	VIP and CRS TroubleShouting.....	109
12.6	How to clean a failed CRS installation	110
13	Install the Oracle 10g software.....	111
13.1	10G Software Post-Installation task	118
13.1.1	Oracle User .profile update.....	118
13.1.2	Cluster Ready Services Health Check	119
13.2	What has been done ?	119
14	oracle listener configuration.....	120
14.1	Listener Post-Configuration task	120
14.2	What has been done ?	120
15	ASM Configuration (ASM Implementation Only).....	121
15.1	ASM Post-Configuration task	126
15.1.1	Cluster Ready Services Health Check	126
15.1.2	Oracle CLuster Registry content Check and Backup	127
15.2	Some usefull commands	129
15.3	What has been done ?	129
16	CREATING THE DATABASE USING DBCA.....	130
16.1	Database Creation On ASM.....	130
16.1.1	Database Post-Configuration task.....	138
16.1.1.1	Cluster Ready Services Health Check	138
16.1.1.2	Oracle CLuster Registry content Check and Backup.....	139
16.1.1.3	Some usefull commands	140
16.1.1.4	What has been done ?	140
16.2	Database creation On GPFS.....	141
16.2.1	Database Post-Configuration task.....	148
16.2.1.1	Cluster Ready Services Health Check	148
16.2.1.2	Oracle CLuster Registry content Check and Backup.....	149
16.2.1.3	Some usefull commands	150
16.2.1.4	What has been done ?	150
16.3	Checking DB CONSOLE :	151
16.4	Manual Database Creation.....	152
16.4.1	Tips to create a database in 10g Real Application Clusters.....	152

16.4.2	Configure listener.ora / sqlnet.ora / tnsnames.ora.....	154
16.4.3	Configure Oracle Enterprise Manager.....	154
17	Appendix B : examples of configuration files	155
17.1	Network.....	155
17.2	listener.ora and tnsnames.ora Configuration example	156
17.2.1	Listener.ora Files	156
17.2.2	Tnsnames.ora Files (Server Side) for ASMDB Database	157
17.2.3	Tnsnames.ora Files (Server Side) for GPFS Database	158
17.2.4	Tnsnames.ora Files (Client Side) for ASMDB Database.....	159
17.2.5	Tnsnames.ora Files (Client Side) for GPFS Database.....	159
18	Appendix C : Oracle technical notes.....	160
18.1	CRS and 10g Real Application Clusters	160
18.2	About RAC	170
18.3	About CRS	170
18.4	About VIP	170
18.5	About manual database cration	171
18.6	About Grid Control	171
18.7	About TAF	171
18.8	About Adding/Removing Node	171
18.9	About ASM	172
18.10	Metalink note to use in case of problem with CRS	172
19	Appendix F : Filesets to be installed on the machines of the cluster.	173
19.1	RSCT 2.3.4 (provided with AIX 5.3) for all implementation.....	173
19.2	For GPFS implementation.....	173

1 THE AIM OF THIS DOCUMENT

This document is written to provide help installing Oracle10g Real Application Clusters (10.2) release 2 on IBM pSeries servers with AIX 5L.

We will describe step by step two different architecture :

- Oracle CRS (Cluster Ready Service) on raw disks and database on ASM (Automatic Storage Management),
- Oracle CRS Data and database on GPFS (General Parallel File System),

As ASM and GPFS are mostly used in today customers architectures, "Concurrent Raw devices using IBM HACMP" will not be covered in this 10gRAC R2 cookbook, please read our 10gRAC R1 cookbook if you need to implement it.

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 10g Release 2 (10.2) for AIX http://download-uk.oracle.com/docs/cd/B19306_01/install.102/b14201/toc.htm	OTN	B14201
Oracle Clusterware and Oracle Real Application Clusters Administration and Deployment Guide - 10g Release 2 (10.2)	OTN	B14197-03
Oracle Database Release Notes – 10g Release 2 (10.2) for AIX 5L Based Systems (64bits)	OTN	B19074-03
Oracle Universal Installer Concepts Guide Release 2.2	OTN	A96697-01
Minimum Software Versions and Patches Required to Support Oracle Products on IBM pSeries	Oracle Metalink	Note:282036.1
IBM General Parallel File System (GPFS) and Oracle RAC on AIX 5L and IBM eServer pSeries	Oracle Metalink	Note:302806.1
GPFS for AIX5L – AIX Clusters Concepts, Planning and Installation Guide	IBM	GA22-7895-01
GPFS for AIX5L – AIX Clusters Administration and Programming Reference	IBM	SA22-7896-01
GPFS on AIX Clusters – High Performance File System – Administration Simplified	IBM	SG24-6035-00

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 EMEA 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://www-1.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. We want our technical papers to be as helpful as possible.
Please send us your comments about this document to the European EMEA Oracle/IBM Joint Solutions Center.

Use our email address :

oraclibm@fr.ibm.com

Or our phone number :

+33 (0)4 67 34 67 49

2 WHAT'S NEW WITH RAC IMPLEMENTATION ON AIX5L

With Oracle 10gRAC, HACMP is not any more necessary as clusterware software.

Oracle provide with 10gRAC its own clusterware, the CRS (Oracle Cluster Ready Service).

With 10gRAC, Oracle Clusterware is mandatory as clusterware.

Possible ways to implement 10gRAC on AIX :

1/ 10gRAC with OCR (Oracle Cluster Registry) disk(s), and Voting (Heartbeat) disk(s) on raw disks, and database on Oracle ASM (Automated Storage Management)

- NO HACMP
- NO GPFS
- ASM provided with oracle software

2/ 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)

- NO HACMP
- GPFS required

3/ 10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s), and database on concurrent raw devices with HACMP

- HACMP (volume group resource) mandatory for concurrent access to raw devices
- GPFS required

4/ 10gRAC with OCR (Oracle Cluster Registry) disk(s), Voting (Heartbeat) disk(s) and database on Network Appliance Filers

- HACMP mandatory for concurrent access to raw devices
- GPFS required

5/ Veritas Foudation, etc ...

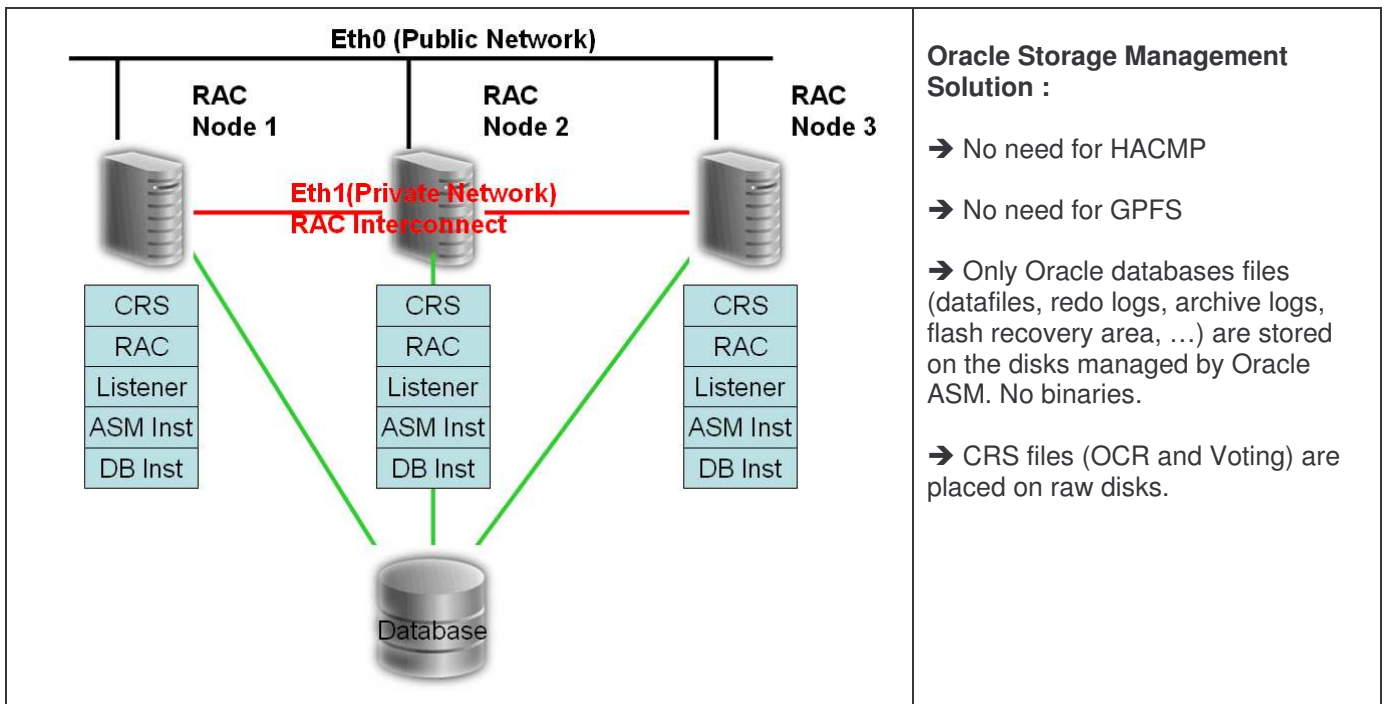
3 WAYS TO IMPLEMENT 10G RAC ON AIX5L

There is two type of infrastructure to implement RAC :

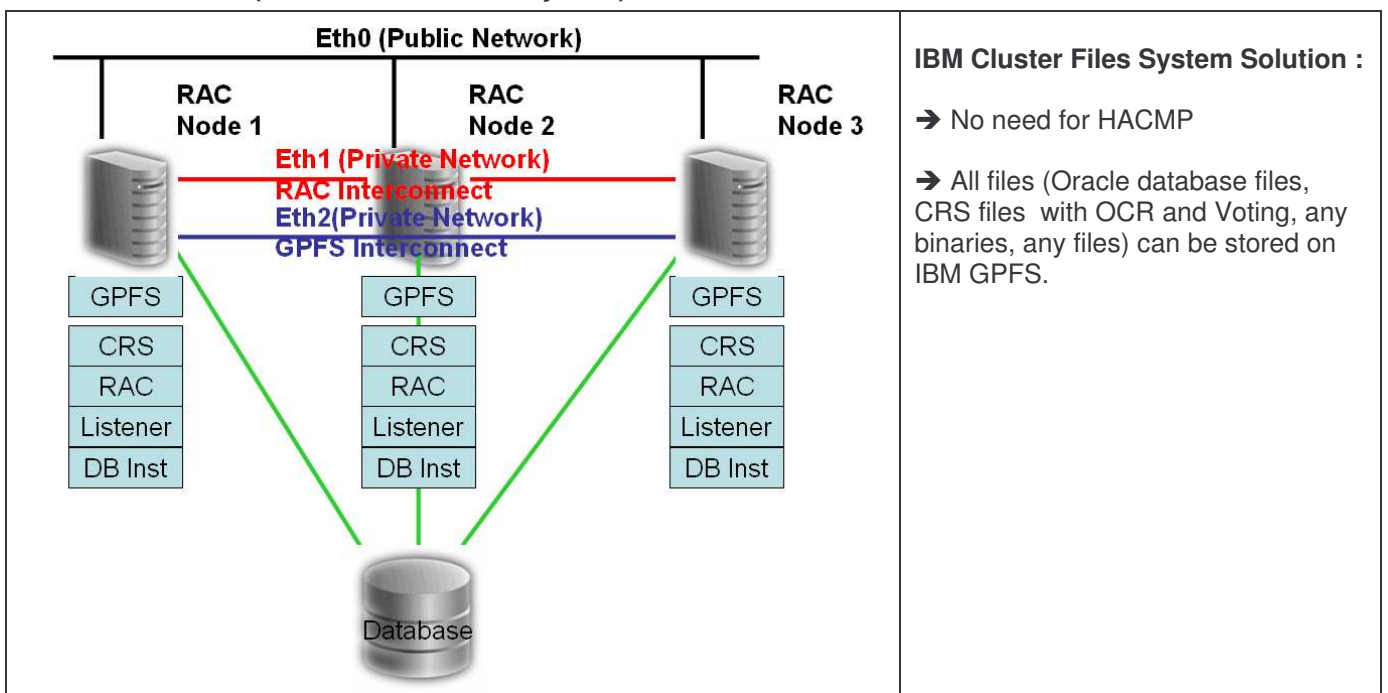
- **Over LAN (Local Area Network)** →
 - 2 to n pSeries, 1 storage
 - 2 to n pSeries, 2 storages
- **Over MAN (Metropolitan Network)** →
 - 2 to n pSeries, 2 storage with DWDM technology

3.1 10G RAC OVER LOCAL AREA NETWORK

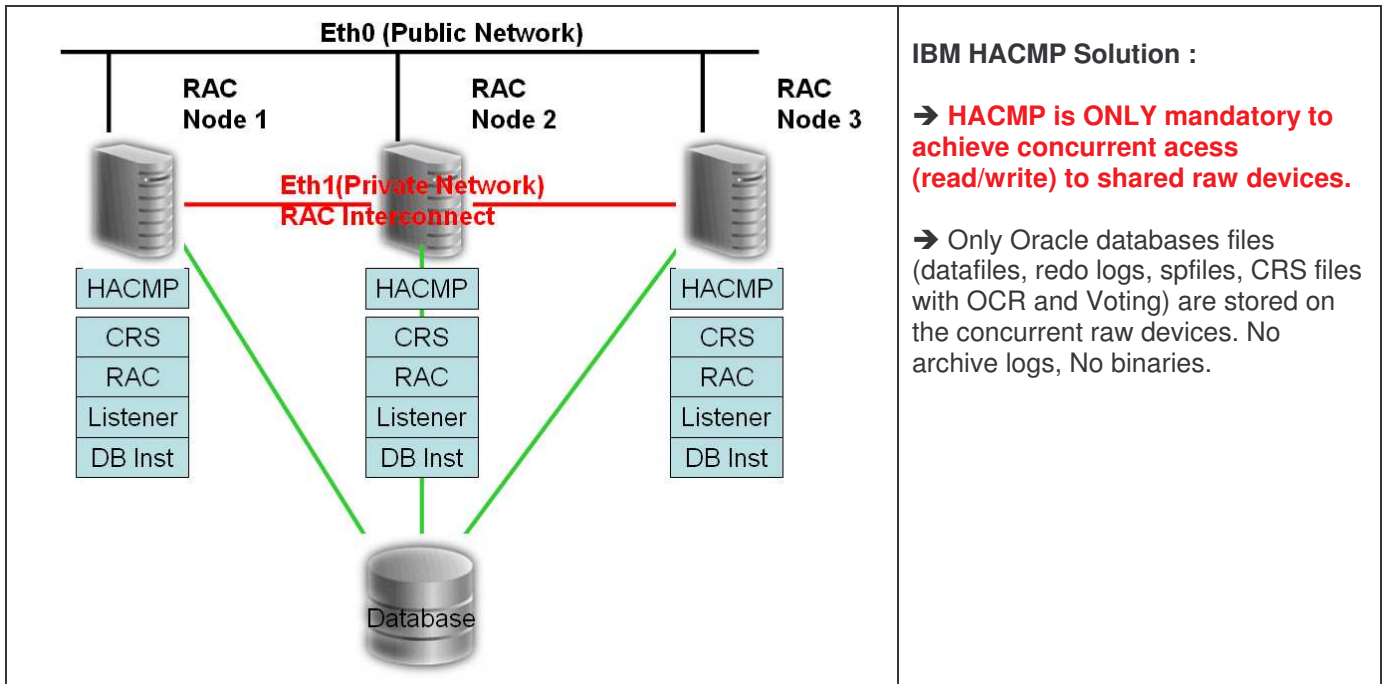
3.1.1 ASM (Automated Storage Management)



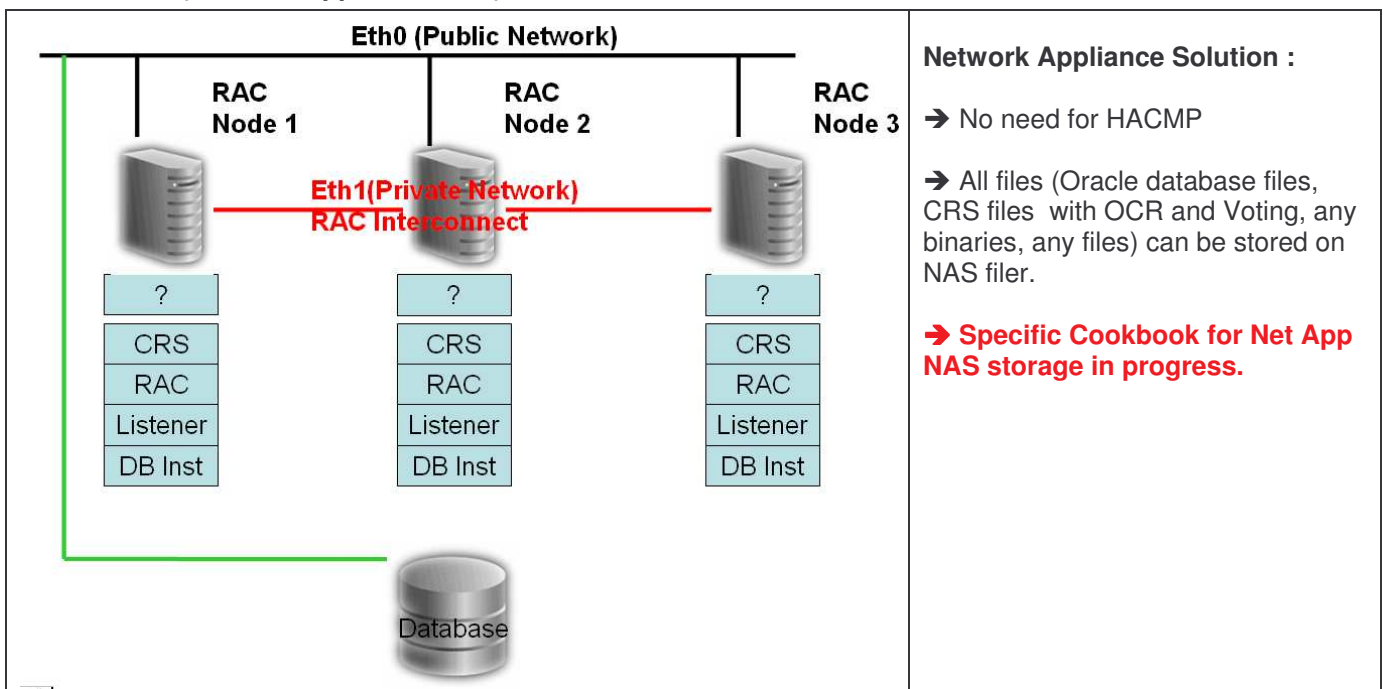
3.1.2 IBM GPFS (General Parallel File System)



3.1.3 Concurrent Raw Devices (HACMP)



3.1.4 NAS (Network Appliance Filer)



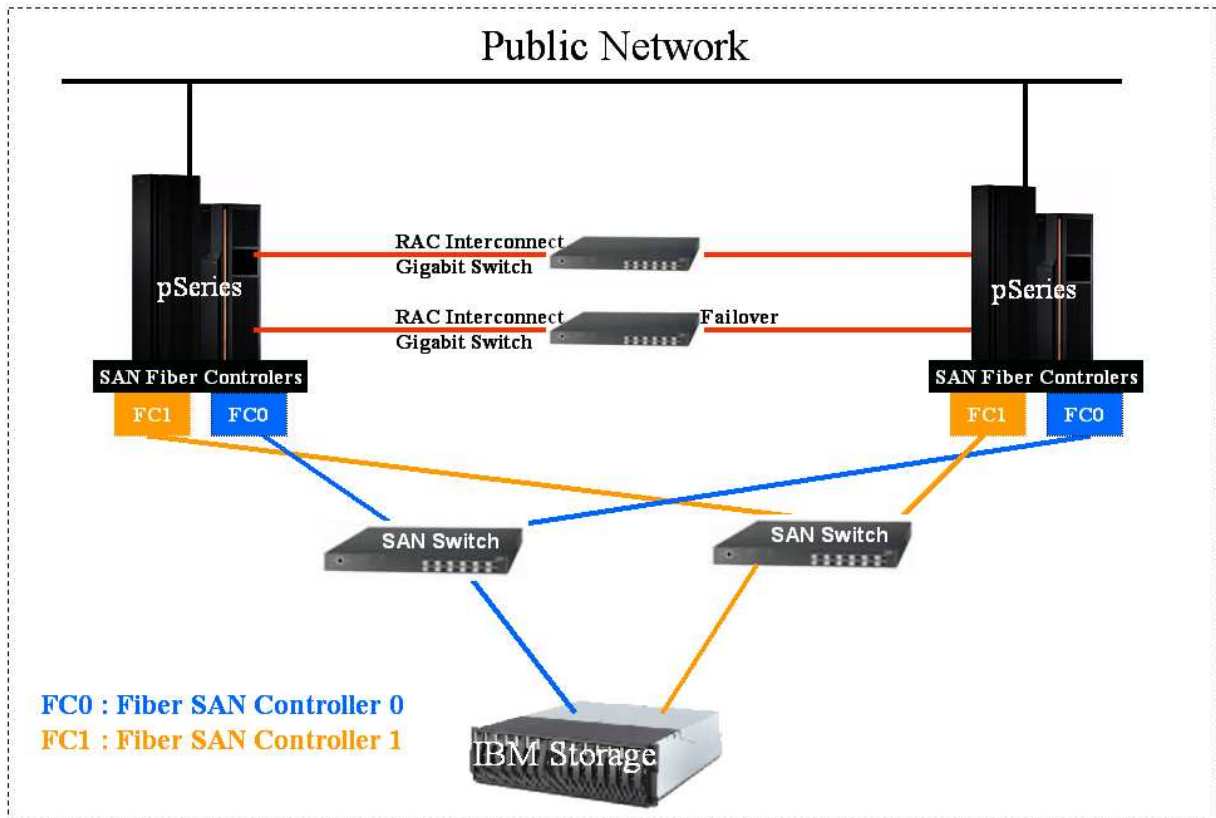
3.2 10G RAC OVER MAN (METROPOLITAN AREA NETWORK)

Please contact us to discuss about such architecture.

4 HARDWARE ARCHITECTURE

For our infrastructure, we used a cluster which is composed of three partitions (IBM LPAR) on an IBM @server pSeries 595 using AIX 5L.

BUT in the real world, to achieve true high availability it's necessary to have at least two IBM pSeries as show bellow :



4.1 IBM pSERIES SERVERS

This is the IBM pSeries server we used for our installation

IBM eServer p5 595
(Frame-mount)

- ➔ Product details
- ➔ Browse and buy



The **@server@** p5 595 server uses fifth-generation 64-bit IBM POWER5™ processors in up to 64-way symmetric multiprocessing configurations. p5-595 is IBM's most powerful UNIX®/Linux® system with the power, flexibility, scalability and security needed to run mission-critical enterprise applications.

Processor:	POWER5 16-way to 64-way
Clock rates (Min/Max)	1.65GHz / 1.90GHz
System memory (Std/Max)	8GB / 2TB
Internal storage (Std/Max)	72.8GB / 28.1TB (using optional I/O drawers)
Performance (rPerf range)****	80.86 to 306.21

<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.

4.2 OPERATING SYSTEM

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

➔ Check following chapters 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 resources 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

4.3 NETWORK INFRASTRUCTURE

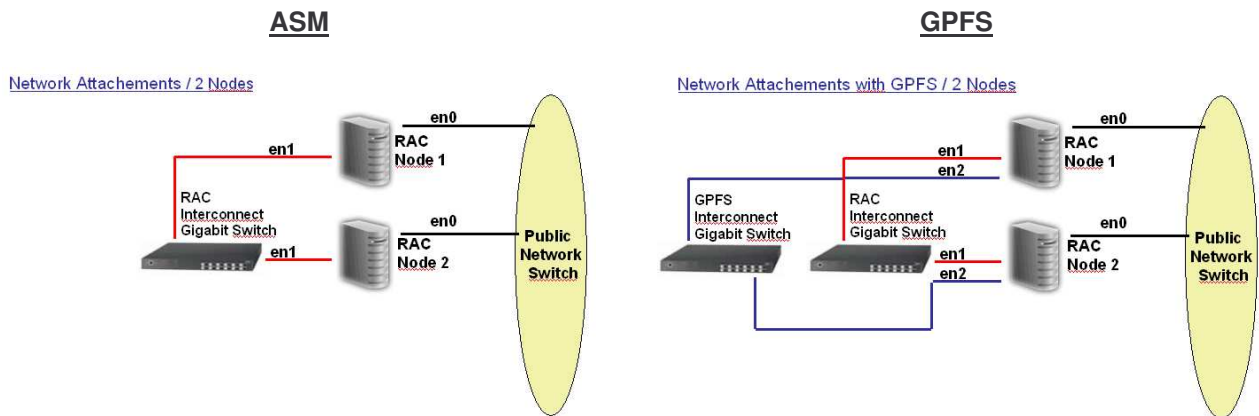
A private network (for instance a gigabit ethernet network, using a gigabit switch to link each cluster nodes) is designed only for Oracle interconnect use (cache fusion between instances). This dedicated network is mandatory.

✚ 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>). This network can also be used for GPFS.

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

4.3.1 Network Architecture Examples



- 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.

4.3.2 EtherChannel

EtherChannel is supported to be used with RAC on AIX, EtherChannel provide network bandwidth aggregation and failover. This is implemented at AIX level and fully transparent to oracle RAC.

4.3.3 AIX Virtual I/O Network

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

4.4 AIX VIRTUAL I/O DISKS

You can use virtual I/O disks for :

- 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....)

4.5 SAN STORAGE

Oracle Storage Compatibility Program →

<http://www.oracle.com/technology//deploy/availability/htdocs/oscp.html>

Oracle Storage Compatibility Program Member List →

<http://www.oracle.com/technology//deploy/availability/htdocs/vendors.html>

4.5.1 IBM

IBM TotalStorage products for IBM pSeries



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

At this date, IBM Storage DS300 and DS400 are not supported with 10gRAC. New micro code has been released, and it's planned to be tested in 2006 with 10gRAC.

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

4.5.1.1 IBM Multi-Pathing :

To implement multi-path I/O, you should implement IBM MPIO.

This feature allow SAN access failover, and load balancing accros SAN Fiber Channel attachements.

4.5.1.1.1 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>

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

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

Commands
to check disks :

```
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 :

```
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
```

4.5.2 EMC

With EMC, please refer to EMC to see supported storage with RAC.
For multi-pathing, EMC PowerPatch should be used.

4.5.3 HITACHI

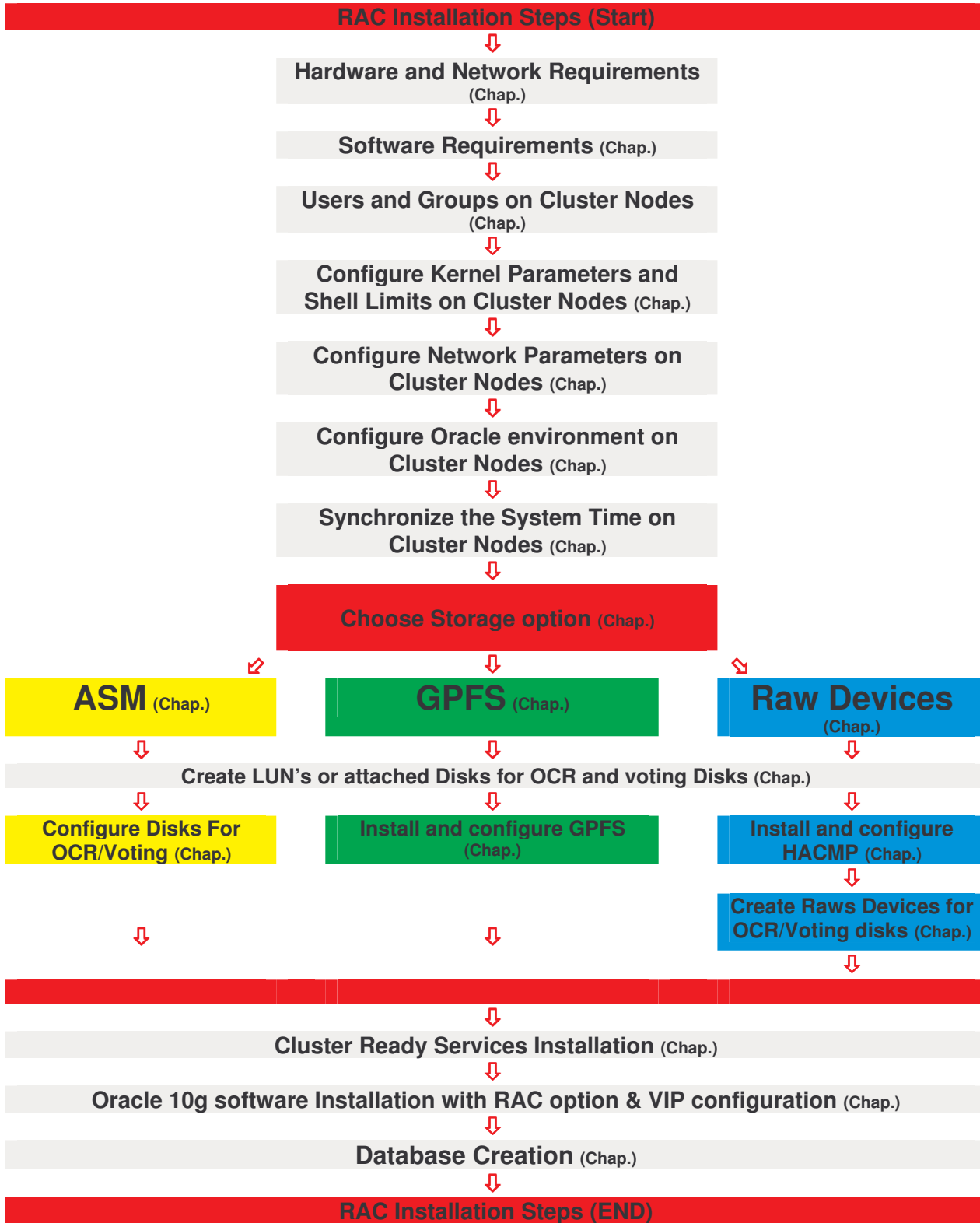
With Hitachi, please refer to Hitachi to see which HDS storage is supported with RAC.

4.5.4 Others

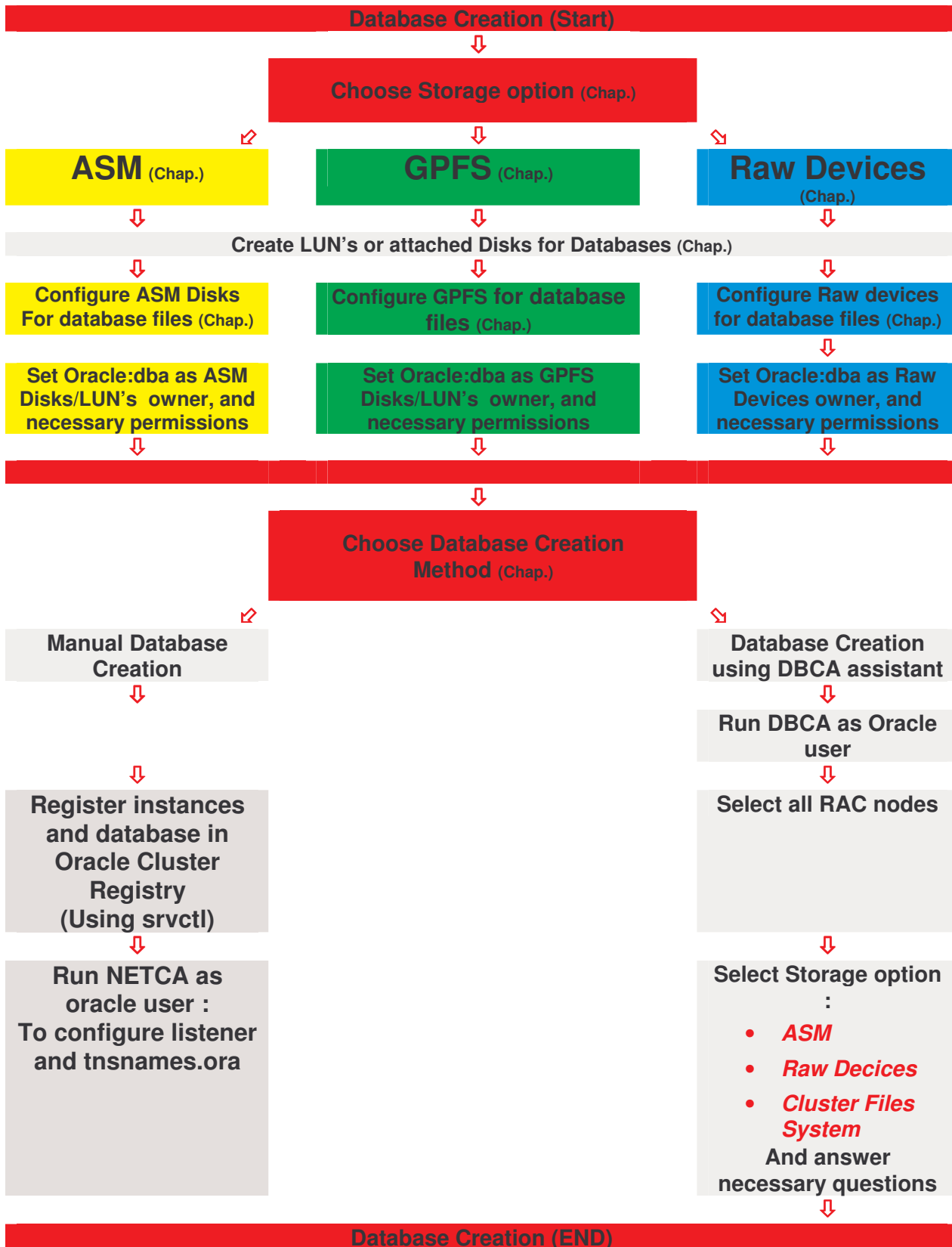
For most of the storage solutions, please contact the providing company for supported configuration, as read/write concurrent acces from all RAC nodes must be possible to implement a RAC solution. That means possibility to setup the disk reserve_policy to no_reserve or equivalent.

5 INSTALLATION STEPS

5.1 RAC INSTALLATION STEPS



5.2 DATABASE CREATION STEPS



6 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	<ul style="list-style-type: none"> • Check the Hardware Requirements 						
2	<ul style="list-style-type: none"> • Check the Network Requirements 						
3	<ul style="list-style-type: none"> • Check the Software Requirements 						
4	<ul style="list-style-type: none"> • Create Required UNIX Groups and User 						
5	<ul style="list-style-type: none"> • Configure Kernel Parameters and Shell Limits 						
6	<ul style="list-style-type: none"> • Identify Required Software Directories 						
7	<ul style="list-style-type: none"> • Identify or Create an Oracle Base Directory 						
8	<ul style="list-style-type: none"> • Create the CRS Home Directory 						
9	<ul style="list-style-type: none"> • Choose a Storage Option for Oracle CRS, Database, and Recovery Files 						
11	<ul style="list-style-type: none"> • Create Directories for Oracle CRS, Database, or Recovery Files 						
12	<ul style="list-style-type: none"> • Configure Disks for Automatic Storage Management • Or Configure Raw Devices And Verify the Cluster Software Configuration • Or Configure GPFS for Shared Files System <p>→ And ACCESS Disk Storage concurrently from all nodes participating to the RAC cluster</p>						
13	<ul style="list-style-type: none"> • Synchronize the System Time on Cluster Nodes 						
14	<ul style="list-style-type: none"> • Stop Existing Oracle Processes 						
15	<ul style="list-style-type: none"> • Configure the oracle User's Environment 						

7 PREPARING THE SYSTEM

7.1 HARDWARE REQUIREMENTS

Requirements to meet on ALL nodes !!!

- **RAM** \geq 512 MB minimum
 - ➔ Command to check the physical memory :
`lsattr -El sys0 -a realmem`
- **Internal disk** \geq 6 GB for the oracle code
- **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`

7.2 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 that the required software and patches are installed on the system.

AIX release supported with Oracle 10g RAC R2

- **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 :

```
# oslevel -r
```

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/>

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

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

```
# lspp -l bos.adt.base bos.adt.lib bos.adt.libm bos.perf.perfstat \
bos.perf.libperfstat bos.perf.proctools rsct.basic.rte
```

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

Depending on the products that you intend to install, verify that the following patches are installed on the system. The procedure following the table describes how to check these requirements.

- 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:

```
# /usr/sbin/instfix -i -k "IY43980 IY44810 IY45462 IY45707 IY46214 IY46605 \
IY48525 IY51801 IY56024"
```

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

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

```
# lspp -l -B U489726 U485561 ...
```

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

7.2.1 Filesets / APAR's Requirements for 10gRAC R2 / ASM (NO HACMP)
AIX filesets & 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
<p>Filesets</p> <ul style="list-style-type: none"> • 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 • xIC.aix50.rte 7.0.0.4 • xIC.rte 7.0.0.1 <p>Specific Filesets</p> <p>For EMC Symmetrix :</p> <ul style="list-style-type: none"> • EMC.Symmetrix.aix.rte.5.2.0.1 <p>For EMC CLARiiON :</p> <ul style="list-style-type: none"> • EMC.CLAR11ON.fcp.rte.5.2.0.1 	<p>Filesets</p> <ul style="list-style-type: none"> • 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 • xIC.aix50.rte 7.0.0.4 • xIC.rte 7.0.0.1 <p>Specific Filesets</p> <p>For EMC Symmetrix :</p> <ul style="list-style-type: none"> • EMC.Symmetrix.aix.rte.5.2.0.1 <p>For EMC CLARiiON :</p> <ul style="list-style-type: none"> • EMC.CLAR11ON.fcp.rte.5.2.0.1
<p>APAR's</p> <p>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.2 ML 4:</p> <ul style="list-style-type: none"> • 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. <p>JDK :</p> <ul style="list-style-type: none"> • 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) 	<p>APAR's</p> <p>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.3 ML02:</p> <ul style="list-style-type: none"> • 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. <p>JDK :</p> <ul style="list-style-type: none"> • 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)

7.2.2 Filesets / APAR's Requirements for 10gRAC R2 / GPFS (NO HACMP)
AIX filesets & AIX Patches (APAR) required on ALL nodes for 10gRAC R2 implementation over IBM GPFS !!!

(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
<p>Filesets</p> <ul style="list-style-type: none"> • 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 • xIC.aix50.rte 7.0.0.4 • xIC.rte 7.0.0.1 <p>Specific Filesets</p> <p>For EMC Symmetrix :</p> <ul style="list-style-type: none"> • EMC.Symmetrix.aix.rte.5.2.0.1 <p>For EMC CLARiiON :</p> <ul style="list-style-type: none"> • EMC.CLAR11ON.fcp.rte.5.2.0.1 	<p>Filesets</p> <ul style="list-style-type: none"> • 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 • xIC.aix50.rte 7.0.0.4 • xIC.rte 7.0.0.1 <p>Specific Filesets</p> <p>For EMC Symmetrix :</p> <ul style="list-style-type: none"> • EMC.Symmetrix.aix.rte.5.2.0.1 <p>For EMC CLARiiON :</p> <ul style="list-style-type: none"> • EMC.CLAR11ON.fcp.rte.5.2.0.1
<p>APAR's</p> <p>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.2 ML 4:</p> <ul style="list-style-type: none"> • 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. <p>JDK :</p> <ul style="list-style-type: none"> • 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) <p>APARs required for GPFS v2.3:</p> <ul style="list-style-type: none"> • IY63969: mandatory service for GPFS v2.3 for AIX • IY69911: corrupted kernel using direct I/O across GPFS block boundary • IY70276: GPFS returns out of space when there is still space in the filesystem • IY70277: receive enospc on mkfile in a GPFS file system when file system not full • IY74097 (efix): corrupt archive library after adding members 	<p>APAR's</p> <p>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.3 ML02:</p> <ul style="list-style-type: none"> • 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. <p>JDK :</p> <ul style="list-style-type: none"> • 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) <p>APARs required for GPFS v2.3:</p> <ul style="list-style-type: none"> • IY63969: mandatory service for GPFS v2.3 for AIX • IY69911: corrupted kernel using direct I/O across GPFS block boundary • IY70276: GPFS returns out of space when there is still space in the filesystem • IY70277: receive enospc on mkfile in a GPFS file system when file system not full • IY74097 (efix): corrupt archive library after adding members

7.3 ORACLE SOFTWARE REQUIREMENTS

Link to download code from <http://otn.oracle.com>

<http://www.oracle.com/technology/software/products/database/oracle10g/htdocs/10201aixsoft.html>

Oracle CD's needed for the RAC installation

Oracle Database 10g Release 2 (10.2.0.1.0) Enterprise/Standard Edition for AIX5L
[10gr2 aix5l64 database.cpio.gz](#) (1,268,576,110 bytes) (cksum - 3772623559)

Oracle Clusterware Release 2 (10.2.0.1.0)
[10gr2 aix5l64 cluster.cpio.gz](#) (732,798,920 bytes) (cksum - 1766193337)

Optional CD's

Oracle Database 10g Companion CD Release 2 (10.2.0.1.0)
[10gr2 aix5l64 companion.cpio.gz](#) (1,332,394,319 bytes) (cksum - 303439033)

Not used in the RAC installation.

[Oracle HTML DB v2.0](#) (56,104,916 bytes) - latest standalone version of HTML DB
New (13-Sep-05)

Oracle Database 10g Client Release 2 (10.2.0.1.0)
[10gr2 aix5l64 client.cpio.gz](#) (1,053,970,850 bytes) (cksum - 2541472317)

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 10g documentation can be found [here](#).
4. Review the certification matrix for this product [here](#).

7.4 USERS AND GROUPS

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

Be sure that all the groups and user numbers (203, 204 and 205 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)

```
dba!:204:oracle
oinstall!:205:oracle
```

Check if there is some AIX default limitations (especially on the file size) :

```
File size limitation → ulimit -f
All limitations → ulimit -a
```

See also the file `/etc/security/limits` which shows the limits for each user. The default standard applies to all new user to be created. This file can be modified by root with `vi`.

The default limits should be set to unlimited, except for core (e.g. `-1` in the file `/etc/security/limits`). To turn some user limitation to unlimited, use `smit users`.

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

```
passwd oracle
```

7.5 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.

7.5.1 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.

7.5.2 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.

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.

7.5.3 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 view the current value specified for these parameters, and to change them if necessary,

follow these steps :

1. To check the current values of the network tuning parameters, enter commands similar to the following: `# /usr/sbin/no -a | more`
2. 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
```

3. 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=2*655360
  /usr/sbin/no -o ipqmaxlen=512
fi
```

By adding these lines to the `/etc/rc.net` file, the values persist when the system restarts.

ELSE

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=2*655360
```

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.

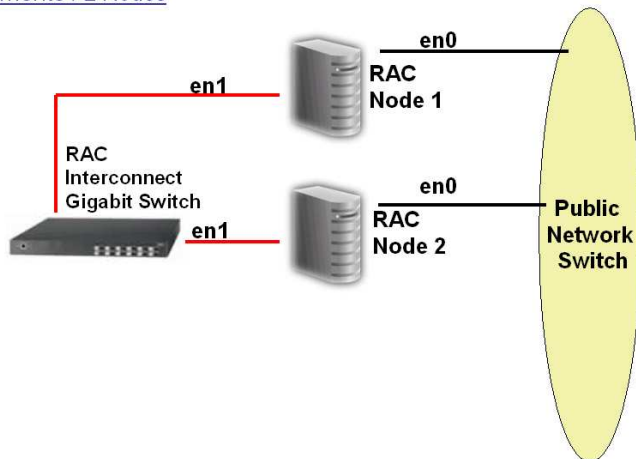
7.6 NETWORK CONFIGURATION

7.6.1 Possible network configuration layout of our pSeries RAC cluster :

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

7.6.1.1 Network implementation with 2 Nodes

Network Attachements / 2 Nodes

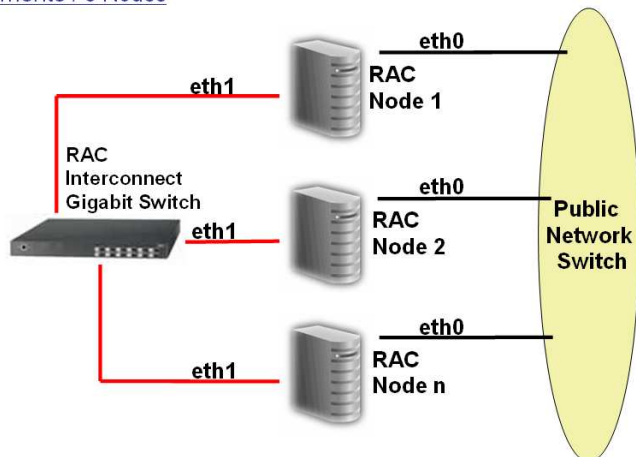


**Network cards for public network must have same name on each participating node in the RAC cluster.*

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

7.6.1.2 Network implementation with 3 Nodes, and more

Network Attachements / 3 Nodes

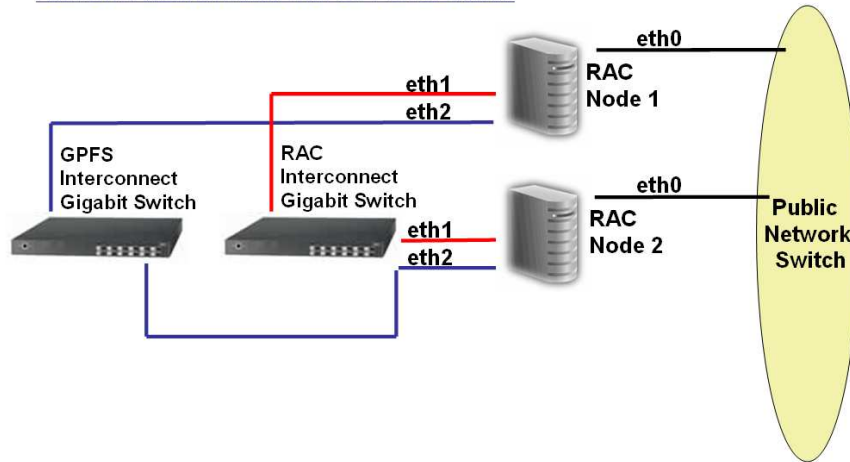


**Network cards for public network must have same name on each participating node in the RAC cluster.*

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

7.6.1.3 Network implementation with IBM GPFS, 2 Nodes

Network Attachements with GPFS / 2 Nodes

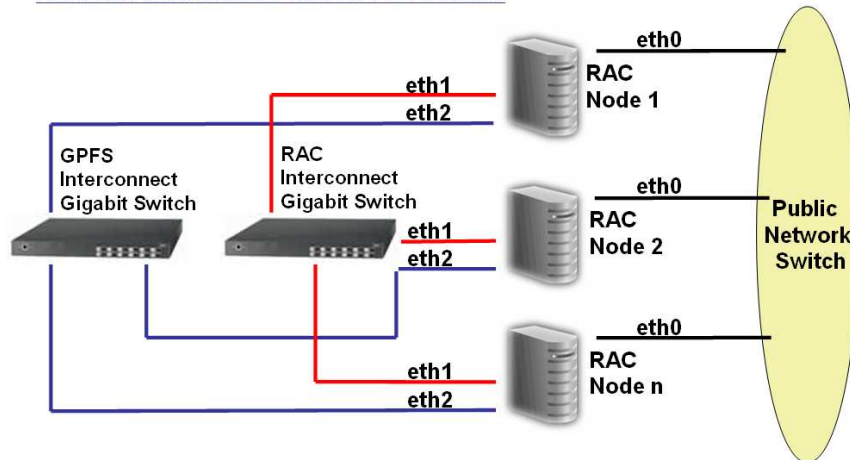


•Network cards for public network must have same name on each participating node in the RAC cluster.

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

7.6.1.4 Network implementation with IBM GPFS, 3 Nodes and more

Network Attachements with GPFS / 3 Nodes



•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.

7.6.2 Configuration steps to follow

7.6.2.1 Network Card Identification

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 0xfffff00 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 0xfffff00 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 0xfffff00 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}/ ->
```

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

	Public	VIP	RAC Interconnect	GPFS Interconnect <i>(If GPFS used)</i>
Network card On each node	en0	en0	en1	en2

7.6.2.2 Nodes Name and IP Identification

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

```
{node1:root}/ ->hostname {node2:root}/ ->hostname {node3:root}/ ->hostname
node1 node2 node3
```

✦ 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 ✦

Public		VIP		RAC Interconnect (Private Network)		GPFS Interconnect <i>(If GPFS used)</i>	
en0		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_gpfs	20.20.25.81
node2	10.3.25.82	node2_vip	10.3.25.182	node1_rac	10.10.25.82	node1_gpfs	20.20.25.82
node3	10.3.25.83	node3_vip	10.3.25.183	node1_rac	10.10.25.83	node1_gpfs	20.20.25.83

7.6.2.3 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
10.3.25.83      node3

# Virtual IP address
10.3.25.181     node1_vip
10.3.25.182     node2_vip
10.3.25.183     node3_vip

# Interconnect RAC
10.10.25.81     node1_rac
10.10.25.82     node2_rac
10.10.25.83     node3_rac

# Interconnect GPFS (If GPFS used)
20.20.25.81     node1_gpfs
20.20.25.82     node2_gpfs
20.20.25.83     node3_gpfs
```

NOTA :

RAC & GPFS Interconnect could be split in two part :

**One for RAC interconnect, and
One for GPFS interconnect**

OR

Same Interconnect link could also be used for RAC and GPFS

7.6.2.4 Defining Default gateway on public interface

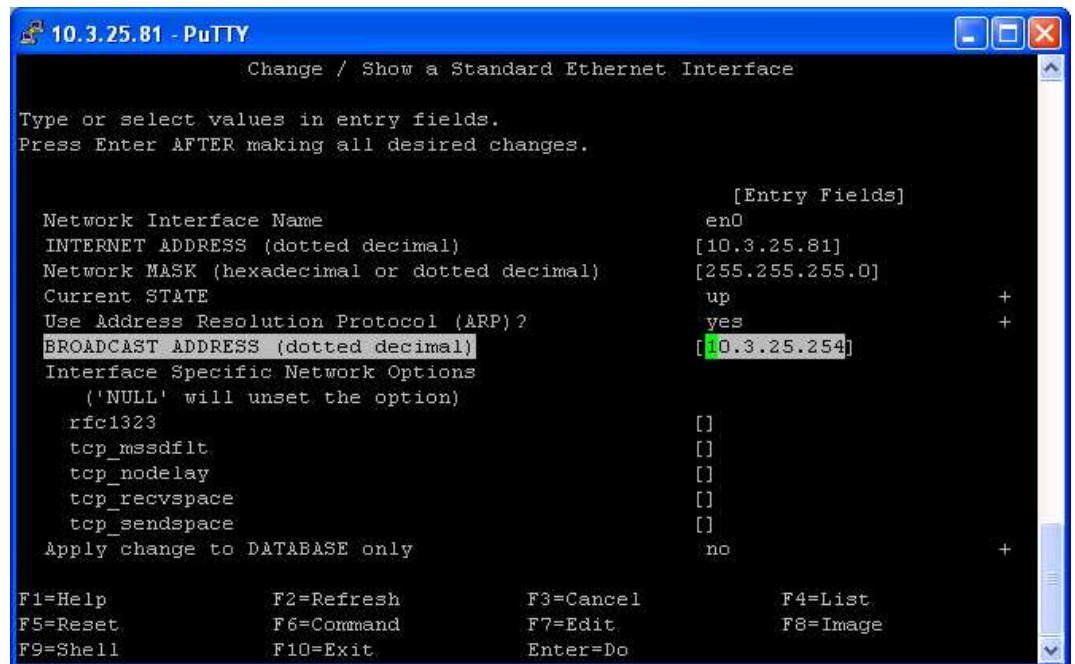
Using smitty, do set the default gateway on public network interface, en0 in our case, and on all nodes :

Smitty chinnet

→ select en0

Add entry and validated to get :

en0 changed



7.6.2.5 User Equivalence Setup

Before you install and use 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 and root user.
`node1` must ssh or rsh to `node2`, as oracle and root user.
`node2` must ssh or rsh to `node1`, as oracle and root user.
`node2` must ssh or rsh to `node2`, as oracle and root user.

7.6.2.5.1 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
node3          root
node1_gpfs    root (mandatory for GPFS implementation)
node2_gpfs    root (mandatory for GPFS implementation)
node3_gpfs    root (mandatory for GPFS implementation)
node1          oracle
node2          oracle
node3          oracle
```

\$HOME/.rhosts

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

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

```
node1:root-> su - root ↵
node1:root/> cd ↵
node1:root-> pg$HOME/.rhosts ↵
node1          root
node2          root
node3          root
(mandatory for GPFS implementation)
node1_gpfs    root
node2_gpfs    root
node3_gpfs    root
node1          oracle
node2          oracle
node3          oracle
```

```
node1:root-> su - oracle ↵
node1:oracle-/home/oracle> cd ↵
pg $HOME/.rhosts ↵
node1          root
node2          root
node3          root
(mandatory for GPFS implementation)
node1_gpfs    root
node2_gpfs    root
node3_gpfs    root
node1          oracle
node2          oracle
node3          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
node2 :oracle-/home/oracle> dfsvdc<sdv<vd<d 2006
```

7.6.2.5.2 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:

Note:

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

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
```

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
```


7.7 ORACLE ENVIRONMENT SETUP

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

⚡ To be done on each node.

```
export ORACLE_BASE=/oh10g
export AIXTHREAD_SCOPE=S           (S for system-wide thread scope)
export TEMP=/tmp
export TMP=/tmp
export TMPDIR=/tmp
umask 022
```

7.8 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 :

- 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-/> mkgv -f -y'oraclevg' -s'32' hdisk1
```

On node 2 ...

Create a 6GB file system /oh10g in the previous volume group (large file enabled) :

On node 1 ...

```
node1:root-/> crfs -v jfs2 -a bf=true -g'oraclevg' -a size='8388608' -m'/oh10g' -A'yes' -p'rw' -t'no' -a nbpi='8192' -a ag='64'
```

```
node1:root-/> mount /oh10g
node1:root-/> chown oracle:dba /oh10g
```

On node 2 ...

The Oracle code can also be located on shared concurrent disks on a GPFS file-system, if IBM GPFS is used.

8 CHOOSE A STORAGE OPTION FOR ORACLE CRS, DATABASE, AND RECOVERY FILES

The following table shows the storage options supported for storing Oracle Cluster Ready Services (CRS) files, Oracle database files, and Oracle database recovery files. Oracle database files include datafiles, control files, redo log files, the server parameter file, and the password file. Oracle CRS files include the Oracle Cluster Registry (OCR) and the CRS voting disk.

For all installations, you must choose the storage option that you want to use for Oracle CRS files and Oracle database files. If you want to enable automated backups during the installation, you must also choose the storage option that you want to use for recovery files (the flash recovery area).

Note: For the most up-to-date information about supported storage options for RAC installations, refer to the Certify pages on the OracleMetaLink Web site:

<http://metalink.oracle.com>

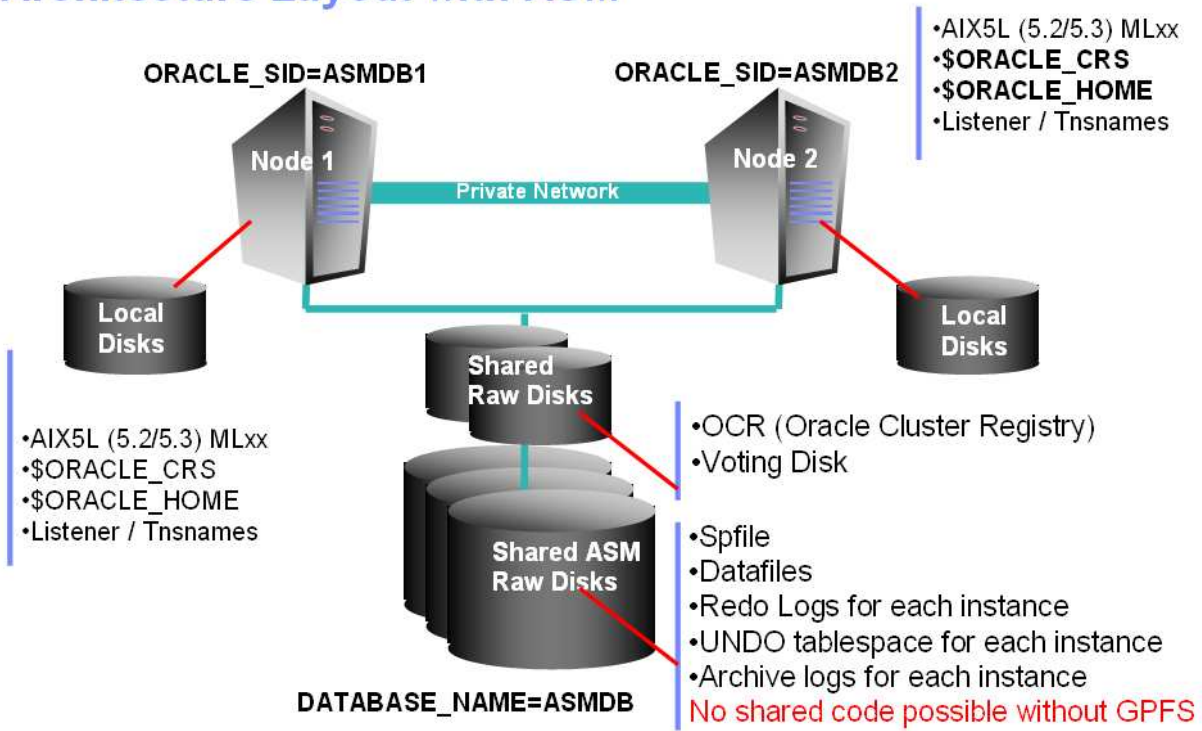
Storage Option	File Types Supported			
	OCR and Voting Disk	Oracle Software	Database	Recovery
Automatic Storage Management	No	No	Yes	Yes
IBM General Parallel File System (GPFS)	Yes	Yes	Yes	Yes
Local Storage	No	Yes	No	No
Raw Logical volumes Managed by HACMP	Yes	No	Yes	Yes

Use the following guidelines when choosing the storage options that you want to use for each file type:

- You can choose any combination of the supported storage options for each file type as long as you satisfy any requirements listed for the chosen storage options.
- For Standard Edition installations, ASM is the only supported storage option for database or recovery files.
- Automatic Storage Management cannot be used to store Oracle CRS files, because these files must be accessible before any Oracle instance starts.
- If you are not using HACMP, you cannot use shared raw logical volumes for CRS or database file storage.
- For information about how to configure disk storage before you start the installation, refer to one of the following sections depending on your choice:
 - To use **ASM** for database or recovery file storage, refer to chapters 8.1 to prepare RAW disks for OCR/Voting disks and disks for ASM implementation..
 - To use **GPFS cluster file system** for Oracle CRS, database, or recovery file storage, refer to chapter 8.2 for GPFS implementation to prepare cluster files system for OCR/Voting disks (seen as files).
 - **To use concurrent raw devices** implementation, please read 10gRAC R1 cookbook for details.

8.1 ASM IMPLEMENTATION

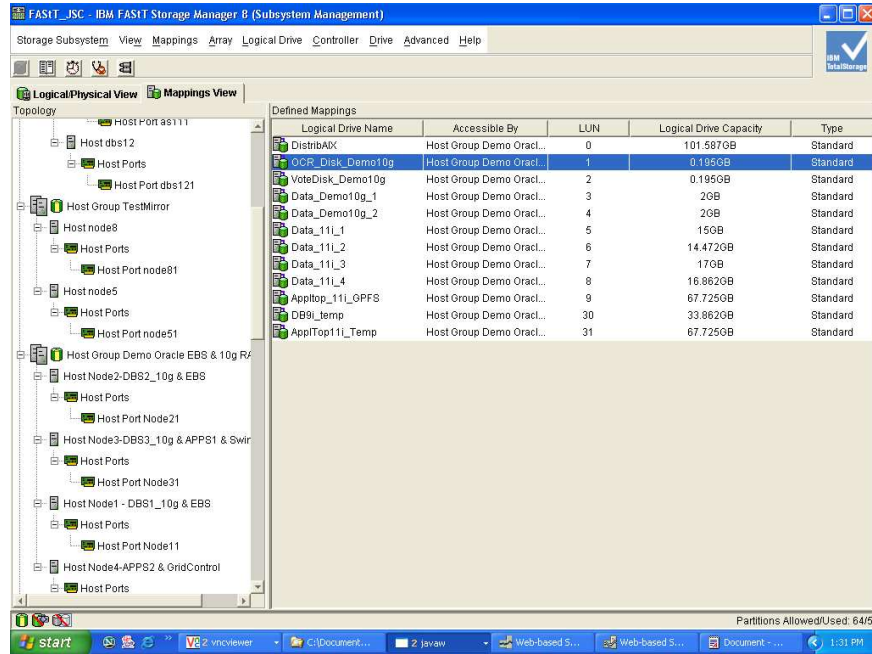
Architecture Layout with ASM



8.1.1 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 disk and Voting disk have ids 1 and 2. These ID's will help us on to identify which hdisk will be used.

Disks	LUN's ID Number	LUN's Size
OCR	L1	200 MB
Voting	L2	200 MB
Disk 1 for ASM	L3	n GB
Disk 2 for ASM	L4	n GB
Disk 3 for ASM	L5	n GB
.....	...	
Disk 8 for ASM	LA (meaning Lun 10)	n GB
Disk 9 for ASM	LB (meaning Lun 11)	n GB

8.1.2 Register LUN's at AIX level

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

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.

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

8.1.3 Preparing Raw Disks for CRS/Voting disks

We know the two LUN's available for OCR and Voting disks are L1 and L2.

Disks	LUN's ID Number
OCR	L1
Voting	L2

Identify the two 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	L1			
Voting	L2			

There are two methods to identify the corresponding hdisks :

- Identify LUN ID assign to hdisk, using “**lscfg -l hdisk?**” command
- Identify hdisks by assigning momentarily a PVID to each hdisk not having one

We strongly recommend to use the first method, to avoid playing with PVID.

IF using first method :

→ **DO MAKE shure to release PVID after hdisks identifications**

IF using second method :

→ **DO MAKE shure that no PVID are already assigned to these hdisks**

Get the list of hdisk available on each node :

On node 1 ...

```
node1:root-> Isdev -Ccdisk
hdisk0 Available      Virtual SCSI Disk Drive
hdisk1 Available 01-08-02 3552 (500) Disk Array Device
hdisk2 Available 01-08-02 3552 (500) Disk Array Device
hdisk3 Available 01-08-02 3552 (500) Disk Array Device
hdisk4 Available 01-08-02 3552 (500) Disk Array Device
hdisk5 Available 01-08-02 3552 (500) Disk Array Device
hdisk6 Available 01-08-02 3552 (500) Disk Array Device
hdisk7 Available 01-08-02 3552 (500) Disk Array Device
hdisk8 Available 01-08-02 3552 (500) Disk Array Device
hdisk9 Available 01-08-02 3552 (500) Disk Array Device
hdisk10 Available 01-08-02 3552 (500) Disk Array Device
hdisk11 Available 01-08-02 3552 (500) Disk Array Device
hdisk12 Available 01-08-02 3552 (500) Disk Array Device
hdisk13 Available 01-08-02 3552 (500) Disk Array Device
```

Get the list of hdisk available on each node :

On node 2 ...

```
Node2:root-> Isdev -Ccdisk
hdisk0 Available      Virtual SCSI Disk Drive
hdisk1 Available 01-08-02 3552 (500) Disk Array Device
hdisk2 Available 01-08-02 3552 (500) Disk Array Device
hdisk3 Available 01-08-02 3552 (500) Disk Array Device
hdisk4 Available 01-08-02 3552 (500) Disk Array Device
hdisk5 Available 01-08-02 3552 (500) Disk Array Device
hdisk6 Available 01-08-02 3552 (500) Disk Array Device
hdisk7 Available 01-08-02 3552 (500) Disk Array Device
hdisk8 Available 01-08-02 3552 (500) Disk Array Device
hdisk9 Available 01-08-02 3552 (500) Disk Array Device
hdisk10 Available 01-08-02 3552 (500) Disk Array Device
hdisk11 Available 01-08-02 3552 (500) Disk Array Device
hdisk12 Available 01-08-02 3552 (500) Disk Array Device
hdisk13 Available 01-08-02 3552 (500) Disk Array Device
```

Using First method :

Get the List of the hdisks on node1, and node2 :

List available hdisks on each node :

```

On node 1 ...
node1:root-/# lspv
hdisk0 0033c670e214eac5 rootvg active
hdisk1 none None
hdisk2 none None
hdisk3 none None
.....
→ NO PVID are assigned apart for the rootvg hdisk

On node 2 ...
node2:root-/# lspv
hdisk0 0033c670e214eac5 rootvg active
hdisk1 none None
hdisk2 none None
hdisk3 none None
hdisk4 none None
.....
→ NO PVID are assigned apart for the rootvg hdisk
    
```

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

Identify LUN ID assign to hdisk, using "iscfg -l hdisk?" command

```

On node 1 ...
node1:root-/# iscfg -vl hdisk2
hdisk2 U1.9-P1-I1/Q1-W200300A0B80C5404-L1000000000000 3552 (500) Disk Array Device

node1:root-/# iscfg -vl hdisk3
hdisk3 U1.9-P1-I1/Q1-W200300A0B80C5404-L2000000000000 3552 (500) Disk Array Device
    
```

THEN, We get the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node Corresponding hdisk
OCR	L1	hdisk2		hdisk?
Voting	L2	hdisk3		hdisk?

→ No need to assign PVID when using this method.

Using Iscfg 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 "iscfg -l hdisk?" command

```

On node 2 ...
node2:root-/# iscfg -vl hdisk3
hdisk2 U1.9-P1-I1/Q1-W200300A0B80C5404-L1000000000000 3552 (500) Disk Array Device

node2:root-/# iscfg -vl hdisk4
hdisk3 U1.9-P1-I1/Q1-W200300A0B80C5404-L2000000000000 3552 (500) Disk Array Device
    
```

THEN, We get the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node Corresponding hdisk
OCR	L1	hdisk2	hdisk3	hdisk?
Voting	L2	hdisk3	hdisk4	hdisk?

→ No need to assign PVID when using this method.

Using Second method :

Identify hdisks by assigning **momently** a PVID (Physical Volume ID) to each hdisk not having one

List available hdisks on each node :

<p>On node 1 ...</p> <pre>node1:root-> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 none None hdisk3 none None hdisk4 none None hdisk5 none None -> NO PVID are assigned apart for the rootvg hdisk</pre>	<p>On node 2 ...</p> <pre>node2:root-> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 none None hdisk3 none None hdisk4 none None hdisk5 none None -> NO PVID are assigned apart for the rootvg hdisk</pre>
--	--

If these hdisks do not have a PVID assign one to it as follow :

<p>Assign PVID to hdisk</p> <p>On node 1 ...</p> <pre>node1:root-> chdev -l hdisk2 -a pv=yes node1:root-> chdev -l hdisk3 -a pv=yes</pre>	<p>On node 2 ...</p> <pre>node2:root-> chdev -l hdisk2 -a pv=yes node2:root-> chdev -l hdisk3 -a pv=yes</pre>
--	--

And map PVID number from both nodes to identify which hdisk is mapped to the same LUN :

<p>List available hdisks with assigned PVID on each node :</p> <p>AND identify corresponding hdisks on each node !!!</p>	<p>On node 1 ...</p> <pre>node1:root-> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 0033c670b0f3ee84 None hdisk3 0033c670b0f3fb87 None hdisk4 none None hdisk5 none None -> PVID are assigned for posible hdisk to be used as ASM Disks</pre>	<p>On node 2 ...</p> <pre>node2:root-> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 0033c670b0f3ee82 None hdisk3 0033c670b0f3fb84 None hdisk4 0033c670b0f3fb87 None hdisk5 none None hdisk6 none None -> PVID are assigned for posible hdisk to be used as ASM Disks</pre>
---	--	---

⚡ Be carefull hdisk4 on node1 is not necessary hdisk4 on node2.

THEN We get the following table :

Disks	LUN's ID Number	Node 1 Corresponding hdisk	Node 2 Corresponding hdisk	Node Corresponding hdisk
OCR	L1	hdisk2	hdisk3	hdisk?
Voting	L2	hdisk3	hdisk4	hdisk?

-> PVID are assigned when using this method, THEN we need to clear PVID before going further.

**THEN
Clear
PVID
for each
Disks,**

**On each
node !!!**

On node 1 ...

```
node1:root-/> chdev -l hdisk2 -a pv=clear
node1:root-/> chdev -l hdisk3 -a pv=clear
node1:root-/> lspv
hdisk0 0033c670e214eac5 rootvg active
hdisk1 none None
hdisk2 none None
hdisk3 none None
hdisk4 none None
hdisk5 none None
..... → NO PVID assigned apart for the rootvg hdisk, PVID cleared for hdisk2 and hdisk3.
```

On node 2 ...

```
Node2:root-/> chdev -l hdisk3 -a pv=clear
Node2:root-/> chdev -l hdisk4 -a pv=clear
Node2:root-/> lspv
hdisk0 0033c670e214eac5 rootvg active
hdisk1 none None
hdisk2 none None
hdisk3 none None
hdisk4 none None
hdisk5 none None
..... → NO PVID assigned apart for the rootvg hdisk, PVID cleared for hdisk3 and hdisk4.
```

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

Example for one hdisk :

Issue the command “lsattr -E -l hdisk2” to visualize all attributes for hdisk4
Or only lsattr -E -l hdisk2 | grep reserve

```
$ lsattr -E -l hdisk2 | grep reserve
reserve_policy single_path                Reserve Policy                True
```

- 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`

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-/> chdev -l hdisk2 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk3 -a reserve_policy=no_reserve
```

On node 2 ...

```
node1:root-/> rsh node2
node2:root-/> chdev -l hdisk3 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk4 -a reserve_policy=no_reserve
.....
```

Example for one hdisk :

Issue the command “lsattr -E -l hdisk2 |grep reserve” to visualize modified attributes for hdisk2

```
$ lsattr -E -l hdisk2 | grep reserve
reserve_policy no_reserve                Reserve Policy                True
```

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	L1	/dev/ocr_disk	hdisk2			hdisk3		
Voting	L2	/dev/vote_disk	hdisk3			hdisk4		

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

On node1 :

```
node1:root-> ls -l /dev/*hdisk2
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk2
crw----- 1 root system 20, 4 Oct 18 17:52 /dev/rhdisk2
node1:root-> ls -l /dev/*hdisk3
brw----- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk3
crw----- 1 root system 20, 5 Oct 18 17:52 /dev/rhdisk3
```

On node2 :

```
node1:root-> rsh node2
node2:root-> ls -l /dev/*hdisk3
brw----- 1 root system 12, 9 Oct 18 16:22 /dev/hdisk3
crw----- 1 root system 12, 9 Oct 18 16:22 /dev/rhdisk3
node2:root-> ls -l /dev/*hdisk4
brw----- 1 root system 12, 8 Oct 18 16:22 /dev/hdisk4
crw----- 1 root system 12, 8 Oct 18 16:22 /dev/rhdisk4
```

Disks	LUN's ID Number	Device Name	Node 1 Corresponding hdisk	Major Num.	Minor Num.	Node 2 Corresponding hdisk	Major Num.	Minor Num.
OCR	L1	<u>/dev/ocr_disk</u>	hdisk2	20	4	hdisk3	12	9
Voting	L2	<u>/dev/vote_disk</u>	hdisk3	20	5	hdisk4	12	8

THEN create a device on each node called /dev/ocr_disk and /dev/vote_d with the right major and minor number :

```
node1:root-> mknod /dev/ocr_disk c 20 4
node1:root-> mknod /dev/vote_disk c 20 5

node1:root-> rsh node2
node2:root-> mknod /dev/ocr_disk c 12,9
node2:root-> mknod /dev/vote_disk c 12,10
```

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/vote_disk
node1:root-> chmod 660 /dev/ocr_disk
node1:root-> chmod 660 /dev/vote_disk

node1:root-> ls -l /dev/* | grep "20, 4"
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk2
brw----- 1 root system 20, 4 Oct 18 17:52 /devr/hdisk2
crw-rw---- 1 oracle dba 20, 4 Oct 19 13:27 /dev/ocr_disk

node1:root-> ls -l /dev/* |grep "20, 5"
brw----- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk3
crw----- 1 root system 20, 5 Oct 18 17:52 /dev/rhdisk3
crw-rw---- 1 oracle dba 20, 5 Oct 19 13:28 /dev/vote_disk
```

On node2 ...

```
node1:root-> rsh node2
node2:root-> chown oracle.dba /dev/ocr_disk
node2:root-> chown oracle.dba /dev/vote_disk
node2:root-> chmod 660 /dev/ocr_disk
node2:root-> chmod 660 /dev/vote_disk

node2:root-> ls -l /dev/* | grep "20, 4"
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk3
brw----- 1 root system 20, 4 Oct 18 17:52 /devr/hdisk3
crw-rw---- 1 oracle dba 20, 4 Oct 19 13:27 /dev/ocr_disk

node2:root-> ls -l /dev/* |grep "20, 5"
brw----- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk4
crw----- 1 root system 20, 5 Oct 18 17:52 /dev/rhdisk4
crw-rw---- 1 oracle dba 20, 5 Oct 19 13:28 /dev/vote_disk
```

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

```
dd if=/dev/zero of=/dev/ocr_disk bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/vote_disk bs=8192 count=25000 &
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,

From node1 :

```
dd if=/dev/zero of=/dev/ocr_disk bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/vote_disk bs=8192 count=25000 &
etc ...
```

At the same time,

From node2 :

```
dd if=/dev/zero of=/dev/ocr_disk bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/vote_disk bs=8192 count=25000 &
etc ...
```

8.1.4 Preparing Disks for ASM

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

Disks	LUN's Number
Disk 1 for ASM	L3
Disk 2 for ASM	L4
Disk 3 for ASM	L5
Disk 4 for ASM	L6
Disk 5 for ASM	L7
Disk 6 for ASM	L8
Disk 7 for ASM	L9
Disk 8 for ASM	LA (meaning Lun 10)
Disk 9 for ASM	LB (meaning Lun 11)
Disk 10 for ASM	LC (meaning Lun 12)
...	...

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 Number	Node 1	Node 2	Node
Disk 1 for ASM	L3			
Disk 2 for ASM	L4			
Disk 3 for ASM	L5			
Disk 4 for ASM	L6			
Disk 5 for ASM	L7			
Disk 6 for ASM	L8			
Disk 7 for ASM	L9			
Disk 8 for ASM	LA			
Disk 9 for ASM	LB			
Disk 10 for ASM	LC			
...

There are two methods to identify the corresponding hdisks :

- Identify LUN ID assign to hdisk, using “**lscfg -l hdisk?**” command
- Identify hdisks by assigning momentarily a PVID to each hdisk not having one

We strongly recommend to use the first method, to avoid playing with PVID.

IF using first method :

→ **DO MAKE shure** to release PVID after hdisks identifications

IF using second method :

→ **DO MAKE shure** that no PVID are already assigned to these hdisks

!!! IN ANY case, DON'T assign PVID to hdisks already used by ASM !!!

Using First method :

Identify LUN ID assign to hdisk, using "lscfg -l hdisk?" command

⚡ Be carefull hdisk4 on node1 is not necessary hdisk4 on node2.

On node 1 ...

```
node1:root-> lscfg -vl hdisk4
hdisk4      U1.9-P1-I1/Q1-W200300A0B80C5404-L300000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk5
hdisk5      U1.9-P1-I1/Q1-W200300A0B80C5404-L400000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk6
hdisk6      U1.9-P1-I1/Q1-W200300A0B80C5404-L500000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk7
hdisk7      U1.9-P1-I1/Q1-W200300A0B80C5404-L600000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk8
hdisk8      U1.9-P1-I1/Q1-W200300A0B80C5404-L700000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk9
hdisk9      U1.9-P1-I1/Q1-W200300A0B80C5404-L800000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk10
hdisk10     U1.9-P1-I1/Q1-W200300A0B80C5404-L900000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk11
hdisk11     U1.9-P1-I1/Q1-W200300A0B80C5404-LA000000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk12
hdisk12     U1.9-P1-I1/Q1-W200300A0B80C5404-LB000000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk13
hdisk13     U1.9-P1-I1/Q1-W200300A0B80C5404-LC000000000000 3552 (500) Disk Array Device
```

On node 2 ...

```
node1:root-> rsh node2
node2:root-> lscfg -vl hdisk5
hdisk5      U1.9-P1-I1/Q1-W200300A0B80C5404-L300000000000 3552 (500) Disk Array Device
node2:root-> lscfg -vl hdisk6
hdisk6      U1.9-P1-I1/Q1-W200300A0B80C5404-L400000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk7
hdisk7      U1.9-P1-I1/Q1-W200300A0B80C5404-L500000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk8
hdisk8      U1.9-P1-I1/Q1-W200300A0B80C5404-L600000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk9
hdisk9      U1.9-P1-I1/Q1-W200300A0B80C5404-L700000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk10
hdisk10     U1.9-P1-I1/Q1-W200300A0B80C5404-L800000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk11
hdisk11     U1.9-P1-I1/Q1-W200300A0B80C5404-L900000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk12
hdisk12     U1.9-P1-I1/Q1-W200300A0B80C5404-LA000000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk13
hdisk13     U1.9-P1-I1/Q1-W200300A0B80C5404-LB000000000000 3552 (500) Disk Array Device
node1:root-> lscfg -vl hdisk14
hdisk14     U1.9-P1-I1/Q1-W200300A0B80C5404-LC000000000000 3552 (500) Disk Array Device
```

THEN

We get the following table :

Disks	LUN's Number	Node 1	Node 2	Node
Disk 1 for ASM	L3	hdisk4	hdisk5	hdisk?
Disk 2 for ASM	L4	hdisk5	hdisk6	hdisk?
Disk 3 for ASM	L5	hdisk6	hdisk7	hdisk?
Disk 4 for ASM	L6	hdisk7	hdisk8	hdisk?
Disk 5 for ASM	L7	hdisk8	hdisk9	hdisk?
Disk 6 for ASM	L8	hdisk9	hdisk10	hdisk?
Disk 7 for ASM	L9	hdisk10	hdisk11	hdisk?
Disk 8 for ASM	LA	hdisk11	hdisk12	hdisk?
Disk 9 for ASM	LB	hdisk12	hdisk13	hdisk?
Disk 10 for ASM	LC	hdisk13	Hdisk14	hdisk?
...

➔ No need to assign PVID when using this method.

Using Second method :

Identify hdisks by assigning **momently** a PVID (Physical Volume ID) to each hdisk not having one

List available hdisks on each node :

<p>On node 1 ...</p> <pre>node1:root-/> lspv hdisk0 0033c670e214eac5 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 → NO PVID are assigned apart for the rootvg hdisk</pre>	<p>On node 2 ...</p> <pre>node2:root-/> lspv hdisk0 0033c670e214eac5 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 hdisk14 none None → NO PVID are assigned apart for the rootvg hdisk</pre>
---	---

If these hdisks do not have a PVID assign one to it as follow, and map PVID number from both nodes to identify which hdisk is mapped to the same.

<p>On node 1 ...</p> <pre>node1:root-/> chdev -l hdisk4 -a pv=yes node1:root-/> chdev -l hdisk5 -a pv=yes node1:root-/> chdev -l hdisk6 -a pv=yes node1:root-/> chdev -l hdisk7 -a pv=yes node1:root-/> chdev -l hdisk8 -a pv=yes node1:root-/> chdev -l hdisk9 -a pv=yes node1:root-/> chdev -l hdisk10 -a pv=yes node1:root-/> chdev -l hdisk11 -a pv=yes node1:root-/> chdev -l hdisk12 -a pv=yes node1:root-/> chdev -l hdisk13 -a pv=yes</pre>	<p>On node 2 ...</p> <pre>node2:root-/> chdev -l hdisk5 -a pv=yes node2:root-/> chdev -l hdisk6 -a pv=yes node2:root-/> chdev -l hdisk7 -a pv=yes node2:root-/> chdev -l hdisk8 -a pv=yes node2:root-/> chdev -l hdisk9 -a pv=yes node2:root-/> chdev -l hdisk10 -a pv=yes node2:root-/> chdev -l hdisk11 -a pv=yes node2:root-/> chdev -l hdisk12 -a pv=yes node2:root-/> chdev -l hdisk13 -a pv=yes node2:root-/> chdev -l hdisk14 -a pv=yes</pre>
--	---

List available hdisks with assigned PVID on each node :

<p>On node 1 ...</p> <pre>node1:root-/> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 none None hdisk3 none None hdisk4 0033c670e214eac9 None hdisk5 0033c670e224eac5 None hdisk6 0033c670e234eac5 None hdisk7 0033c670e244eac5 None hdisk8 0033c670e254eac5 None hdisk9 0033c670e264eac5 None hdisk10 0033c670e274eac5 None hdisk11 0033c670e284eac5 None hdisk12 0033c670e294eac5 None hdisk13 0033c670e314eac5 None → PVID are assigned for possible hdisk to be used as ASM Disks</pre>	<p>On node 2 ...</p> <pre>node2:root-/> lspv hdisk0 0033c670e214eac5 rootvg active hdisk1 none None hdisk2 none None hdisk3 none None hdisk4 none None hdisk5 0033c670e214eac9 None hdisk6 0033c670e224eac5 None hdisk7 0033c670e234eac5 None hdisk8 0033c670e244eac5 None hdisk9 0033c670e254eac5 None hdisk10 0033c670e264eac5 None hdisk11 0033c670e274eac5 None hdisk12 0033c670e284eac5 None hdisk13 0033c670e294eac5 None hdisk14 0033c670e314eac5 None → PVID are assigned for possible hdisk to be used as ASM Disks</pre>
--	--

AND identify corresponding hdisks on each node !!!

⚡ Be carefull hdisk4 on node1 is not necessary hdisk4 on node2.

THEN We get the following table :

Disks	LUN's Number	Node 1	Node 2	Node ...
Disk 1 for ASM	L3	hdisk4	hdisk5	hdisk?
Disk 2 for ASM	L4	hdisk5	hdisk6	hdisk?
Disk 3 for ASM	L5	hdisk6	hdisk7	hdisk?
Disk 4 for ASM	L6	hdisk7	hdisk8	hdisk?
Disk 5 for ASM	L7	hdisk8	hdisk9	hdisk?
Disk 6 for ASM	L8	hdisk9	hdisk10	hdisk?
Disk 7 for ASM	L9	hdisk10	hdisk11	hdisk?
Disk 8 for ASM	LA	hdisk11	hdisk12	hdisk?
Disk 9 for ASM	LB	hdisk12	hdisk13	hdisk?
Disk 10 for ASM	LC	hdisk13	Hdisk14	hdisk?
...

→ PVID are assigned when using this method, THEN we need to clear PVID before going further.

THEN Clear PVID for each Disks

On node 1 ...

```
node1:root-> chdev -l hdisk4 -a pv=clear
node1:root-> chdev -l hdisk5 -a pv=clear
node1:root-> chdev -l hdisk6 -a pv=clear
node1:root-> chdev -l hdisk7 -a pv=clear
node1:root-> chdev -l hdisk8 -a pv=clear
node1:root-> chdev -l hdisk9 -a pv=clear
node1:root-> chdev -l hdisk10 -a pv=clear
node1:root-> chdev -l hdisk11 -a pv=clear
node1:root-> chdev -l hdisk12 -a pv=clear
node1:root-> chdev -l hdisk13 -a pv=clear
```

```
node1:root-> lspv
hdisk0      0033c670e214eac5      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
.....
```

→ NO PVID assigned apart for the rootvg hdisk, PVID cleared for hdisks to be used for ASM.

On node 2 ...

```
node2:root-> chdev -l hdisk4 -a pv=clear
node2:root-> chdev -l hdisk5 -a pv=clear
node2:root-> chdev -l hdisk6 -a pv=clear
node2:root-> chdev -l hdisk7 -a pv=clear
node2:root-> chdev -l hdisk8 -a pv=clear
node2:root-> chdev -l hdisk9 -a pv=clear
node2:root-> chdev -l hdisk10 -a pv=clear
node2:root-> chdev -l hdisk11 -a pv=clear
node2:root-> chdev -l hdisk12 -a pv=clear
node2:root-> chdev -l hdisk13 -a pv=clear
```

```
node2:root-> lspv
hdisk0      0033c670e214eac5      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
hdisk14     none                    None
.....
```

→ NO PVID assigned apart for the rootvg hdisk, PVID cleared for hdisks to be used for ASM.

Setup reserve_policy on each hdisk, on each node :

Example for one hdisk :

Issue the command “lsattr –E –l hdisk4” to visualize all attributes for hdisk4

Or only lsattr –E –l hdisk4 | grep reserve

```
$ lsattr -E -l hdisk4 | grep reserve
reserve_policy single_path          Reserve Policy          True
```

- 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

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-/> chdev -l hdisk4 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk5 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk6 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk7 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk8 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk9 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk10 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk11 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk12 -a reserve_policy=no_reserve
node1:root-/> chdev -l hdisk13 -a reserve_policy=no_reserve
```

On node 2 ...

```
node2:root-/> chdev -l hdisk5 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk6 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk7 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk8 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk9 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk10 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk11 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk12 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk13 -a reserve_policy=no_reserve
node2:root-/> chdev -l hdisk14 -a reserve_policy=no_reserve
```

Example for one hdisk :

Issue the command “lsattr –E –l hdisk4 |grep reserve” to visualize modified attributes for hdisk4

```
$ lsattr -E -l hdisk4 | grep reserve
reserve_policy no_reserve          Reserve Policy          True
```

As described before, hdisks might have different names from one node to another for example hdisk4 on node1 might be hdisk5 on node2, as shown in following table :

Disks	LUN's Number	Node 1	Node 2	
Disk 1 for ASM	L3	hdisk4	hdisk5	
Disk 2 for ASM	L4	hdisk5	hdisk6	
Disk 3 for ASM	L5	hdisk6	hdisk7	
Disk 4 for ASM	L6	hdisk7	hdisk8	
Disk 5 for ASM	L7	hdisk8	hdisk9	
Disk 6 for ASM	L8	hdisk9	hdisk10	
Disk 7 for ASM	L9	hdisk10	hdisk11	
Disk 8 for ASM	LA	hdisk11	hdisk12	
Disk 9 for ASM	LB	hdisk12	hdisk13	
Disk 10 for ASM	LC	hdisk13	Hdisk14	
...	

→ THEN, we need to implement devices to have same naming

OR

hdisks might have same names from one node to another for example hdisk4 on node1 might be hdisk4 on node2, and so on as shown in the following table :

Disks	LUN's Number	Node 1	Node 2	
Disk 1 for ASM	L3	hdisk4	hdisk4	
Disk 2 for ASM	L4	hdisk5	hdisk5	
Disk 3 for ASM	L5	hdisk6	hdisk6	
Disk 4 for ASM	L6	hdisk7	hdisk7	
Disk 5 for ASM	L7	hdisk8	hdisk8	
Disk 6 for ASM	L8	hdisk9	hdisk9	
Disk 7 for ASM	L9	hdisk10	hdisk10	
Disk 8 for ASM	LA	hdisk11	hdisk11	
Disk 9 for ASM	LB	hdisk12	hdisk12	
Disk 10 for ASM	LC	hdisk13	hdisk13	
...	

→ THEN, we can implement devices to have same other naming
OR use hdisk as it is

Making devices (if wanted) :

Disks	LUN's Number	Device Name	Node 1	Major Num.	Minor Num.	Node 2	Major Num.	Minor Num.
Disk 1 for ASM	L3	/dev/asm_disk1	hdisk4			hdisk5		
Disk 2 for ASM	L4	/dev/asm_disk2	hdisk5			hdisk6		
Disk 3 for ASM	L5	/dev/asm_disk3	hdisk6			hdisk7		
Disk 4 for ASM	L6	/dev/asm_disk4	hdisk7			hdisk8		
Disk 5 for ASM	L7	/dev/asm_disk5	hdisk8			hdisk9		
Disk 6 for ASM	L8	/dev/asm_disk6	hdisk9			hdisk10		
Disk 7 for ASM	L9	/dev/asm_disk7	hdisk10			hdisk11		
Disk 8 for ASM	LA	/dev/asm_disk8	hdisk11			hdisk12		
Disk 9 for ASM	LB	/dev/asm_disk9	hdisk12			hdisk13		
Disk 10 for ASM	LC	/dev/asm_disk10	hdisk13			Hdisk14		
...		

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

On node1 :

```
node1:root-> ls -l /dev/*hdisk4
brw----- 1 root system 20, 3 Oct 18 17:52 /dev/hdisk4
crw----- 1 root system 20, 3 Oct 18 17:52 /dev/rhdisk4
node1:root-> ls -l /dev/*hdisk5
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk5
crw----- 1 root system 20, 4 Oct 18 17:52 /dev/rhdisk5
.....
```

On node2 :

```
node1:root-> rsh node2
node2:root-> ls -l /dev/*hdisk5
brw----- 1 root system 12, 9 Oct 18 16:22 /dev/hdisk5
crw----- 1 root system 12, 9 Oct 18 16:22 /dev/rhdisk5
node2:root-> ls -l /dev/*hdisk6
brw----- 1 root system 12, 8 Oct 18 16:22 /dev/hdisk6
crw----- 1 root system 12, 8 Oct 18 16:22 /dev/rhdisk6
.....
```

Disks	LUN's Number	Device Name	Node 1	Major Num.	Minor Num.	Node 2	Major Num.	Minor Num.
Disk 1 for ASM	L3	/dev/asm_disk1	hdisk4	20	3	hdisk5	12	9
Disk 2 for ASM	L4	/dev/asm_disk2	hdisk5	20	4	hdisk6	12	8
Disk 3 for ASM	L5	/dev/asm_disk3	hdisk6	hdisk7
Disk 4 for ASM	L6	/dev/asm_disk4	hdisk7	hdisk8
Disk 5 for ASM	L7	/dev/asm_disk5	hdisk8	hdisk9
Disk 6 for ASM	L8	/dev/asm_disk6	hdisk9	hdisk10
Disk 7 for ASM	L9	/dev/asm_disk7	hdisk10	hdisk11
Disk 8 for ASM	LA	/dev/asm_disk8	hdisk11	hdisk12
Disk 9 for ASM	LB	/dev/asm_disk9	hdisk12	hdisk13
Disk 10 for ASM	LC	/dev/asm_disk10	hdisk13	Hdisk14
...

THEN create a device on each node called /dev/asm_disk1, /dev/asm_disk2, /dev/asm_.... with the right major and minor number :

```
node1:root-> mknod /dev/asm_disk1 c 20 3
node1:root-> mknod /dev/asm_disk2 c 20 4
.....
node1:root-> rsh node2
node2:root-> mknod /dev/asm_disk1 c 12,9
node2:root-> mknod /dev/asm_disk1 c 12,8
.....
```

If you're using devices (/dev/asm_disk1, ...) DO the following :

THEN Change owner, group and permission for :

**/dev/asm_disk1
/dev/asm_disk2
/dev/asm_disk...**

on each node of the cluster :

On node1 ...

```
node1:root-/> chown oracle.dba /dev/asm_disk1
node1:root-/> chown oracle.dba /dev/asm_disk2
node1:root-/> chmod 660 /dev/asm_disk1
node1:root-/> chmod 660 /dev/asm_disk2
```

```
node1:root-/> ls -l /dev/* | grep "20, 3"
brw----- 1 root system 20, 3 Oct 18 17:52 /dev/hdisk2
brw----- 1 root system 20, 3 Oct 18 17:52 /dev/rhdisk2
crw-rw---- 1 oracle dba 20, 3 Oct 19 13:27 /dev/asm_disk1
```

```
node1:root-/> ls -l /dev/* | grep "20, 4"
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk3
brw----- 1 root system 20, 4 Oct 18 17:52 /dev/rhdisk3
crw-rw---- 1 oracle dba 20, 4 Oct 19 13:28 /dev/asm_disk2
```

.... Do the same for other devices

On node2 ...

```
node1:root-/> rsh node2
node2:root-/> chown oracle.dba /dev/asm_disk1
node2:root-/> chown oracle.dba /dev/asm_disk2
node2:root-/> chmod 660 /dev/asm_disk1
node2:root-/> chmod 660 /dev/asm_disk2
```

```
node2:root-/> ls -l /dev/* | grep "12, 9"
brw----- 1 root system 12, 9 Oct 18 17:52 /dev/hdisk3
brw----- 1 root system 12, 9 Oct 18 17:52 /dev/rhdisk3
crw-rw---- 1 oracle dba 12, 9 Oct 19 13:27 /dev/asm_disk1
```

```
node2:root-/> ls -l /dev/* | grep "12, 8"
brw----- 1 root system 12, 8 Oct 18 17:52 /dev/hdisk4
brw----- 1 root system 12, 8 Oct 18 17:52 /dev/rhdisk4
crw-rw---- 1 oracle dba 12, 8 Oct 19 13:28 /dev/asm_disk2
```

.... Do the same for other devices

Zeroing devices If you're using devices (/dev/asm_disk1, ...) DO the following :

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

```
dd if=/dev/zero of=/dev/asm_disk1 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/asm_disk2 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/asm_disk3 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/asm_disk4 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/asm_disk5 bs=8192 count=25000 &
etc ...
```

If you're not using devices (/dev/asm disk1, ...) DO the following :

Change owner, group and permission for raw disks to be used as ASM disks on each node of the cluster :

On node 1 ...

```
node1:root-/> chown oracle.dba /dev/rhdisk4
node1:root-/> chown oracle.dba /dev/rhdisk5
node1:root-/> chown oracle.dba /dev/rhdisk6
node1:root-/> chown oracle.dba /dev/rhdisk7
node1:root-/> chown oracle.dba /dev/rhdisk8
node1:root-/> chown oracle.dba /dev/rhdisk9
node1:root-/> chown oracle.dba /dev/rhdisk10
node1:root-/> chown oracle.dba /dev/rhdisk11
node1:root-/> chown oracle.dba /dev/rhdisk12
node1:root-/> chown oracle.dba /dev/rhdisk13
node1:root-/> chmod 660 /dev/rhdisk4
node1:root-/> chmod 660 /dev/rhdisk5
node1:root-/> chmod 660 /dev/rhdisk6
node1:root-/> chmod 660 /dev/rhdisk7
node1:root-/> chmod 660 /dev/rhdisk8
node1:root-/> chmod 660 /dev/rhdisk9
node1:root-/> chmod 660 /dev/rhdisk10
node1:root-/> chmod 660 /dev/rhdisk11
node1:root-/> chmod 660 /dev/rhdisk12
node1:root-/> chmod 660 /dev/rhdisk13
```

On node 2 ...

```
node2:root-/> chown oracle.dba /dev/rhdisk5
node2:root-/> chown oracle.dba /dev/rhdisk6
node2:root-/> chown oracle.dba /dev/rhdisk7
node2:root-/> chown oracle.dba /dev/rhdisk8
node2:root-/> chown oracle.dba /dev/rhdisk9
node2:root-/> chown oracle.dba /dev/rhdisk10
node2:root-/> chown oracle.dba /dev/rhdisk11
node2:root-/> chown oracle.dba /dev/rhdisk12
node2:root-/> chown oracle.dba /dev/rhdisk13
node2:root-/> chown oracle.dba /dev/rhdisk14
node2:root-/> chmod 660 /dev/rhdisk5
node2:root-/> chmod 660 /dev/rhdisk6
node2:root-/> chmod 660 /dev/rhdisk7
node2:root-/> chmod 660 /dev/rhdisk8
node2:root-/> chmod 660 /dev/rhdisk9
node2:root-/> chmod 660 /dev/rhdisk10
node2:root-/> chmod 660 /dev/rhdisk11
node2:root-/> chmod 660 /dev/rhdisk12
node2:root-/> chmod 660 /dev/rhdisk13
node2:root-/> chmod 660 /dev/rhdisk14
```

Zeroing hdisk. If you're not using devices (/dev/asm disk1, ...) DO the following :

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

```
dd if=/dev/zero of=/dev/rhdisk4 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/rhdisk5 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
dd if=/dev/zero of=/dev/rhdisk7 bs=8192 count=25000 &
etc....
```


8.1.5 Recommendations, hints and tips

!!! IN ANY case, DON'T assign PVID to hdisks already used by ASM !!!

How to identify hdisks used by ASM

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

```
{node1:oracle}/oh10g/crs/bin ->ls -la /dev/rhdisk* | grep oracle
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk4
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk5
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk6
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk7
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk8
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk9
crw-rw---- 1 oracle dba      20, 4 Oct 18 17:52 /dev/rhdisk10
.....
```

And using AIX command :

Example with hdisk 7 :

```
{node1:oracle}/oh10g/crs/bin ->lquerypv -h /dev/rhdisk7|grep ORCLDISK
```

```
{node1:root}/ ->lquerypv -h /dev/rhdisk7|grep ORCLDISK
00000020 4F52434C 4449534B 00000000 00000000 |ORCLDISK.....|
```

OR

```
{node1:root}/ ->lquerypv -h /dev/rhdisk7
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.I..|
000000D0 01F588CA 150BA800 02001000 00100000 |.....|
000000E0 0001BC80 00001400 00000002 00000001 |.....|
000000F0 00000002 00000002 00000000 00000000 |.....|
{node1:root}/ ->
```

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 ...)

NON used ASM disks will give the following display :

```
{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 |.....|
```

ORCLDISK standing for oracle ASM disk

ASMDB_FLASHRECOVERY 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 ...)

```
{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 |.....|
```

How to free hdisk at AIX level, when hdisk has been removed at oracle level, from oracle ASM Disk Group ?

You must reset the hdisk header having the ASM stamp :

```
{node1:root}/ ->dd if=/dev/zero of=/dev/rhdisk7 bs=8192 count=25000 &
25000+0 records in.
25000+0 records out.
```

THEN query on the hdisk header will return nothing :

```
{node1:root}/ ->lquerypv -h /dev/rhdisk7
{node1:root}/ ->
```

8.2 GPFS IMPLEMENTATION

GPFS is IBM's high-performance parallel, scalable file system for IBM UNIX clusters capable of supporting multi-terabytes of storage within a single file system.

GPFS is a shared-disk file system where every cluster node can have parallel, concurrent read/write access to the same file. It is designed to provide high-performance I/O by "striping" data across multiple disks accessed from multiple servers. GPFS provides high availability through logging and replication, and can be configured for automatic failover from both disk and node malfunctions.

GPFS can be used for all components of an Oracle Database 10g RAC configuration including :

- the shared CRS Home,
- Oracle Home,
- OCR disk,
- voting disk,
- and the Oracle data and log files.

GPFS can also be used to complete the Oracle Automatic Storage Management (ASM) feature in Oracle Database 10g; managing the shared :

- CRS Home,
- Oracle Home,
- OCR disk
- and voting disk

while ASM manages the Oracle data and log files.

GPFS 2.1 and 2.2 were previously approved for Oracle RAC but GPFS 2.3 now offers several new key features including :

- Support for AIX 5L v5.3.
- **Single-node quorum with tie-breaker disks.**
- Single GPFS cluster type.
- More disaster recovery options.

For latest IBM GPFS information, see: http://publib.boulder.ibm.com/clresctr/library/gpfs_fags.html

To have the latest information regarding GPFS and Oracle, see following Metalink Note:

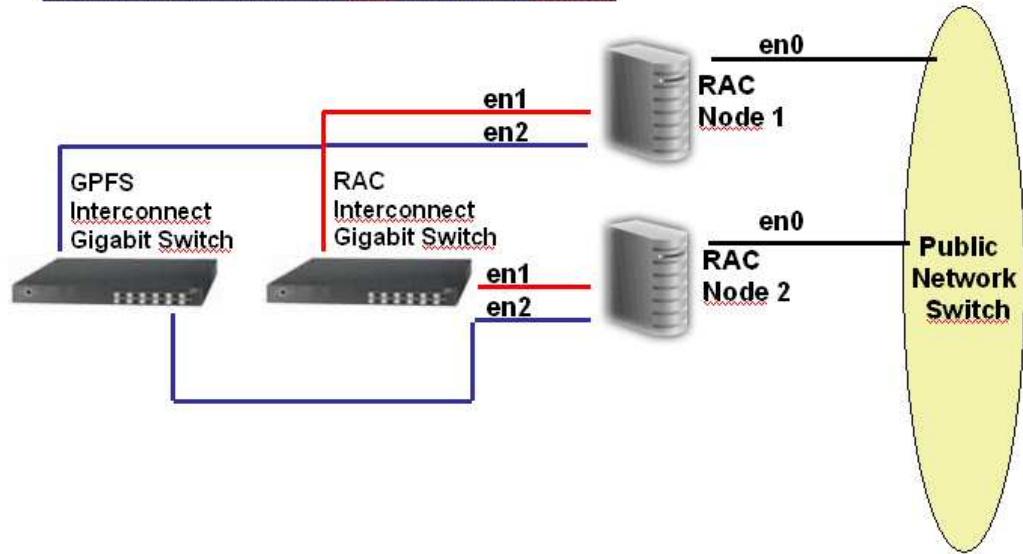
Doc ID: Note:302806.1

[IBM General Parallel File System \(GPFS\) and Oracle RAC on AIX 5L and IBM eSer...](#)
IBM General Parallel File System (**GPFS**) and Oracle RAC on AIX 5L and IBM eServer pSeries

Last Revision Date: 10-JAN-2006

8.2.1 Network Infrastructure for GPFS implementation, with 2 nodes

Network Attachments with GPFS / 2 Nodes



8.2.2 GPFS New Single Node Quorum Support (Tiebreaker Disk)

New single-node quorum support in GPFS 2.3 provides 2 node Oracle High Availability for all disk subsystems :

- Using new quorum type of “node quorum with tiebreaker disks” with 1 or 3 tie-breaker disks.
- Not dependant on storage architecture - designed to work with all storage.
- See [GPFS FAQ](#) (Q1 in “Disk Specific Questions) for currently verified storage and a storage support statement.

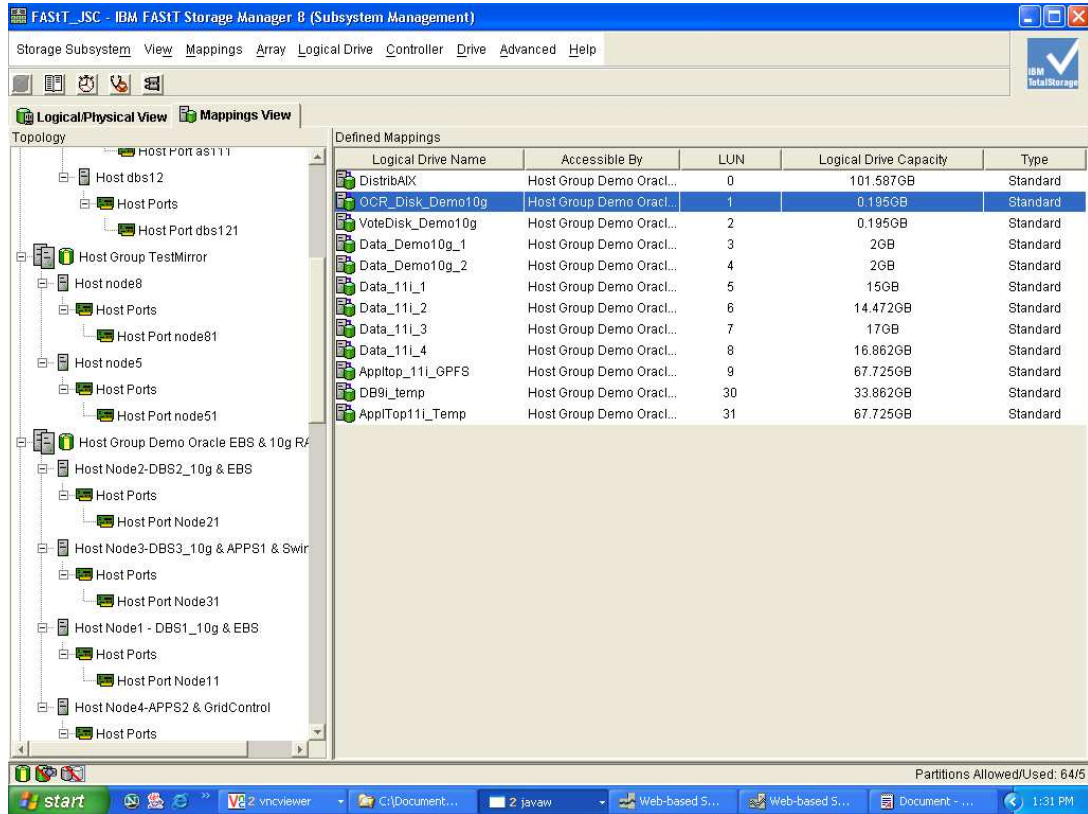
Implementing Disk tiebreaker with a 2 AIX nodes GPFS quorum avoid loosing all the GPFS cluster, when loosing 1 out of 2 AIX nodes. Even with 1 AIX node in the GPFS quorum, the GPFS cluster will stay available to the remaining AIX node, and then to Oracle RAC.

→ No more need for a third AIX node in the GPFS Quorum, when implementing a 2 Oracle RAC nodes cluster.

8.2.3 LUN's creation

Using the storage administration console, you have to create :

- LUN's for OCR and Voting Disks to be used in the GPFS cluster
- LUN's for disks to be used in the GPFS cluster



The following screen shows the LUN mapping for nodes used in our cluster. The LUN for OCR disk and Voting disk have ids 1 and 2. These ids will help us on to identify which hdisk will be used.

Disks	LUN's ID Number	LUN's Size
OCR	L1	200 MB
Voting	L2	200 MB
DB files	L3	n GB
DB files	L4	n GB
DB files	L5	n GB
.....	...	
DB files	LA (meaning Lun 10)	n GB
DB files	LB (meaning Lun 11)	n GB

8.2.4 Register LUN's at AIX level

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.

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

8.2.5 Configure RSH for GPFS with root user

See precedent chapter 7 for how to setup RSH.

On the GPFS node name corresponding to the GPFS interconnect network :

node1_gpfs must ssh or rsh to node1_gpfs, as root user.
node1_gpfs must ssh or rsh to node2_gpfs, as root user.
node2_gpfs must ssh or rsh to node1_gpfs, as root user.
node2_gpfs must ssh or rsh to node2_gpfs, as root user.

8.2.6 INSTALLING GPFS

On the directory containing all the filesets,

check if the hidden file .toc exists.

If not, when located on this directory, make inutoc .

If you are using a cdrom as source for this install, this file always exists on the installation media.

smit install specify the directory containing the filesets

Type F4 to list the filesets to install, rather than choosing "all latest".

To be sure of what you are doing, you can use the field *preview* only before proceeding to the real install.

Filesets gpfs.xx.xxx

See Appendix H : *Filesets to be installed on the machines of the cluster*, paragraph "GPFS" to check that all the necessary filesets have been installed.

This appendix provides the result of the command :

lsipp -L | grep gpfs

Post install task

Add GPFS binaries in root's path.

```
export PATH=$PATH:/usr/lpp/mmfs/bin
```

At this point, all the commands are only issued once, on one of the node part of the peer domain.

8.2.7 Creating a 2 nodes GPFS cluster using tie-breaker disk and a GPFS filesystem for datafiles

- Create the GPFS nodefile which will contain IP hostnames of GPFS interconnect network.

Example:

```
node1:~> cat /var/mmfs/etc/node.list
node1_gpfs:quorum
node2_gpfs:quorum
```

- Create the GPFS cluster

```
node1:~> mmcrcluster -p node1_gpfs -s node2_gpfs -n /var/mmfs/etc/node.list -C Demo10g -A
```

You can check the GPFS creation with the `mmlscluster` and the cluster configuration with `mmlsconfig`. In our example you see :

```
node1:~# mmlscluster

GPFS cluster information
=====
GPFS cluster name:      Demo10g.node1_gpfs
GPFS cluster id:       1446809217114175696
GPFS UID domain:       Demo10g.node1_gpfs
Remote shell command:  /usr/bin/rsh
Remote file copy command: /usr/bin/rcp

GPFS cluster configuration servers:
-----
Primary server:  node1_gpfs
Secondary server: node2_gpfs

Node number Node name IP address Full node name Remarks
-----
1 node1_gpfs 20.20.25.81 node1_gpfs quorum node
2 node2_gpfs 20.20.25.82 node2_gpfs quorum node
```

```
node1:~# mmlsconfig

Configuration data for cluster Demo10g.node1_gpfs:
-----
clusterName Demo10g.node1_gpfs
clusterId 1446809217114175696
clusterType lc
multinode yes
autoload yes
useDiskLease yes
maxFeatureLevelAllowed 809

File systems in cluster Demo10g.node1_gpfs:
-----
(none)
```

- Create a “disk descriptor file” containing the disks devices that will be used as a tie-breaker disk
Example :

```
node1:~# cat /var/mmfs/etc/tie.disk

hdisk13:node1_gpfs:node2_gpfs::tie
```

- Create the network shared disk with mmcrnsd command :

```
Node1: /> mmcrnsd -F /var/mmfs/etc/tie.disk
```

- Designate the tie-breaker disk

```
node1: /> mmchconfig tiebreakerDisks="tie"
Verifying GPFS is stopped on all nodes ...
mmchconfig: Command successfully
mmchconfig: 6027-1371 Propagating the changes to all affected nodes.
This is an asynchronous process.
```

```
node1: /> mmlsconfig
```

```
Configuration data for cluster Demo10g.node1_gpfs:
```

```
-----
clusterName Demo10g.node1_gpfs
clusterId 1446809217114175696
clusterType lc
multinode yes
autoload yes
useDiskLease yes
maxFeatureLevelAllowed 809
tiebreakerDisks tie
File systems in cluster Demo10g.node1_gpfs:
-----
(none)
```

- Start GPFS

```
node1: /> mmstartup -a
```

```
Frid Jan 27 10:38:26 NPT 2006: 6027-1642 mmstartup: Starting GPFS ...
```

8.2.8 Creating a GPFS filesystem

- Create a "disk descriptor file" containing the disks devices that will be shared in the GPFS cluster
Example :

```
node1: /> cat /var/mmfs/etc/data.disk
```

```
hdisk10:node1_gpfs:node2_gpfs::dataAndMetadata:nsd1
hdisk11:node1_gpfs:node2_gpfs::dataAndMetadata:nsd2
hdisk12:node1_gpfs:node2_gpfs::dataAndMetadata:nsd3
```

- Create the network shared disk with mmcrnsd command :

```
Node1: /> mmcrnsd -F /var/mmfs/etc/data.disk
```


You can list the nsd that were created by command `mmlsnsd` :

```
node1:~> mmlsnsd
```

File system	Disk name	Primary node	Backup node
(free disk) nsd1	nsd1	node1_gpfs	node2_gpfs
(free disk) nsd2	nsd2	node1_gpfs	node2_gpfs
(free disk) nsd3	nsd3	node1_gpfs	node2_gpfs
(free disk) tie	tie	node1_gpfs	node2_gpfs

- Create the file-system with `mmcrfs`

```
node1:~> mmcrfs /data_gpfs /dev/data_gpfs -F /var/mmfs/etc/data.disk -A yes -B 1024k -n 8
```

- F <list of disks and descriptors> modified from `mmcrfv`
- A Auto mount the file system at GPFS startup
- B Block size and Stripe size
- n estimated number of nodes
- N Number of I-Nodes

```
node1:~> mmlsnsd
```

File system	Disk name	Primary node	Backup node
data_gpfs nsd1	nsd1	node1_gpfs	node2_gpfs
data_gpfs nsd2	nsd2	node1_gpfs	node2_gpfs
data_gpfs nsd3	nsd3	node1_gpfs	node2_gpfs
(free disk) tie	tie	node1_gpfs	node2_gpfs

```
node1:~> mmlconfig
```

Configuration data for cluster `Demo10g.node1_gpfs`:

```
-----
clusterName Demo10g.node1_gpfs
clusterId 1446809217114175696
clusterType lc
multinode yes
autoload yes
useDiskLease yes
maxFeatureLevelAllowed 809
tiebreakerDisks tie
```

File systems in cluster `Demo10g.node1_gpfs`:

```
-----
/dev/data_gpfs
```

```
node1:~> grep -p /data_gpfs /etc/filesystems
```

```
/data_gpfs:  
dev = /dev/data_gpfs  
vfs = mmfs  
nodename = -  
mount = mmfs  
type = mmfs  
account = false  
options = rw,mtime,atime,dev=data_gpfs
```

- Mount the newly created file-system :

```
node1: /> mount /data_gpfs
```

- Set the permissions and the ownership of the file-systems

```
node1: /> chown oracle:dba /data_gpfs  
node1: /> chmod go+rw /data_gpfs
```

Note :

- GPFS block size should be around (256K, 512K, 1024K)
- GPFS block size is the file system stripe size.
- Not that important for regular database I/O since Direct I/O is used.
- Very important for operations that increase data file size.

8.2.9 Tuning suggestions

- **aioservers** : AIO should be enabled. The general rule for heavy I/O environments is to initially set the “*maxservers*” value to at least 10*(number of disks accessed asynchronously) and the “*minservers*” value can be set to half of the “*maxservers*” value. For an 8 disk GPFS file system, an example of the command to modify would be:

```
node1:./> chdev -l aio0 -a maxservers='80'
node1:./> chdev -l aio0 -a minservers='40'
```

See the “Tuning Asynchronous Disk I/O” section in the “**AIX 5L Version 5.1 Performance Management Guide**” for more discussion and details.

- **pagepool** : the pagepool is used to cache user data and indirect blocks. It is the GPFS pagepool mechanism that allows GPFS to implement read as well as write requests asynchronously. Increasing the size of pagepool increases the amount of GPFS buffer cache pinned memory available. The default value for pagepool is 20 MB and the maximum allowable value is 8GB. Applications that may benefit from a larger pagepool (compared to the default) potentially include those that either reuse data, those that have a random I/O pattern, and/or those that have a higher per client performance requirement. The size of the pagepool will depend on the “working set” of I/O data that needs to be cached. For instance to change pagepool to 100 MB:

```
node1:./> mmchconfig pagepool=100M

node1:./> mmlsconfig
Configuration data for nodeset MyCluster:
-----
pagepool 100M
dataStructureDump /tmp/mmfs
multinode yes
autoload yes
useSingleNodeQuorum no
wait4RVSD no
comm_protocol TCP
clusterType hacmp
group Gpfs.set1
recgroup GpfsRec.MyCluster
File systems in nodeset MyCluster:
-----
/oradata_gpfs
```

- **ipqmaxlen** : the ipqmaxlen network option controls the number of incoming packets that can exist on the IP interrupt queue. Since both GPFS and IBM Virtual Shared Disk use IP, the default value of 128 is often insufficient. This is especially important if your virtual shared disks are configured over IP. The recommended setting is 512.

```
node1:./> no -a ipqmaxlen=512
node1:./> rsh node2 no -a ipqmaxlen=512
```

9 SYNCHRONIZE THE SYSTEM TIME ON CLUSTER NODES

⚡ 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`.

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

1. Enter the following command to create required files, if necessary:

```
# 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:

```
# Sample NTP Configuration file

# Specify the IP Addresses of three clock server systems.

server ip_address1
server ip_address2
server ip_address3

# 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 synchronize 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**.

10 IMPORTANT TIPS FOR ORACLE SOFTWARE AND PATCHES INSTALLATION INSTRUCTIONS

10.1 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

@ (AuthWiz 2.0) Created from [SR 4219914.995](#).

@ [Click here](#) to edit in wizard.

Content Type: TEXT/X-HTML

Creation Date: 13-DEC-2004

Last Revision Date: 02-MAR-2005

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.

10.2 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.

11 CLUSTER VERIFICATION UTILITY

11.1 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

11.2 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 successfully 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 successfully 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
10g 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.

MAKE SURE You have "unzip" tool or symbolic link to unzip in /usr/bin on both node

**On node1
Execute the
following script :**

```

root@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_node1_root.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 "root"

Node Name	Comment
node2	passed
node1	passed

Result: User equivalence check passed for user "root".

→ this is ssh or rsh tests for root user ...

From node1 to node2
From node1 to node1
From node2 to node1
From node2 to node2

Checking administrative privileges...

Check: Existence of user "root"

Node Name	User Exists	Comment
node1	yes	passed
node2	yes	passed

Result: User existence check passed for "root".

→ this is user and group existence tests on node1

Check: Existence of group "oinstall"

Node Name	Status	Group ID
node1	exists	501
node2	exists	501

Result: Group existence check passed for "oinstall".

The following message is not a big issue :

Result: Group existence check failed for "oinstall".

→ Just create a oinstall group at AIX level, using smitty as root

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.

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 message 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 below :

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.x 192.168.x.x 10.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.

Check: System architecture
Node Name  Available           Required           Comment
-----
node1      powerpc             powerpc           passed
node2      powerpc             powerpc           passed
Result: System architecture check passed.

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.

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.

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.

Check: Free disk space in "/oh10g" 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.

```

- ➔ Kernel version test ...
Detecting AIX release
- ➔ System architecture test ...
Detecting Power processor
- ➔ Checking memory requirements ...
- ➔ Checking swap space requirements ...
- ➔ Checking /tmp free space requirements ...
- ➔ 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 as explained in the cookbook (chapter 7).

CVU is testing existence of all prerequisites needed for all implementations as :

- RAC implementation
- ASM implementation
- GPFS implementation
- HACMP implementation

Check: Package existence for "xIC.aix50.rte:7.0.0.4"

Node Name	Status	Comment
node1	xIC.aix50.rte:7.0.0.4	passed
node2	xIC.aix50.rte:7.0.0.4	passed

Result: Package existence check passed for "xIC.aix50.rte:7.0.0.4".

Check: Package existence for "xIC.rte:7.0.0.1"

Node Name	Status	Comment
node1	xIC.rte:7.0.0.1	passed
node2	xIC.rte:7.0.0.1	passed

Result: Package existence check passed for "xIC.rte:7.0.0.1".

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.base	IY63969:gpfs.docs.data	IY63969:gpfs.msg.en_US IY63969 passed
node2	IY63969:gpfs.base	IY63969:gpfs.docs.data	IY63969: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.mpl	IY61034:bos.mp64	IY61034 passed
node2	IY61034:bos.mpl	IY61034:bos.mp64	IY61034 passed

Result: Operating system patch check passed for "IY61034".

Check: Operating system patch for "IY61770"

Node Name	Applied	Required	Comment
node1	IY61770:rsct.basic.rtel	IY61770:rsct.core.errm	IY61770:rsct.core.hostrm
	IY61770:rsct.core.rmcl	IY61770:rsct.core.sec	
	IY61770:rsct.core.sensor	IY61770:rsct.core.utils	IY61770 passed
Node2	IY61770:rsct.basic.rtel	IY61770:rsct.core.errm	IY61770:rsct.core.hostrm
	IY61770:rsct.core.rmcl	IY61770:rsct.core.sec	
	IY61770:rsct.core.sensor	IY61770: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.prof	IY62191:bos.rte.libpthread	IY62191 passed
node2	IY62191:bos.adt.prof	IY62191:bos.rte.libpthread	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 "xlftrt:9.1"

Node Name	Status	Comment
node1	xlftrt:8.1.1.4	failed
node2	xlftrt:8.1.1.4	failed

Result: Package existence check failed for "xlftrt: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".

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".

Check: Operating system patch for "IY60930"

Node Name	Applied	Required	Comment
node1	IY60930:bos.mpl	IY60930:bos.mp64	passed
node2	IY60930:bos.mpl	IY60930:bos.mp64	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.rtelY58143:bos.acctIY58143:bos.adt.includelY58143:bos.adt.libmlY58143:bos.adt.proflY58143:bos.alt_disk_install.rtelY58143:bos.diag.com			
IY58143:bos.mp64IY58143:bos.mplIY58143:bos.net.ewlm.rtelY58143:bos.net.ipsec.keymgtIY58143:bos.net.ipsec.rtelY58143: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.controllIY58143:bos.rte.filesystemIY58143:bos.rte.installIY58143:bos.rte.libcIY58143:bos.rte.lvmlIY58143:bos.rte.manIY58143:bos.rte.methodsIY58143:bos.rte.securityIY58143:bos.rte.serv_aidIY58143:bos.sysmgt.nim.clientIY58143:bos.sysmgt.quotallY58143:bos.sysmgt.serv_aid			
IY58143:bos.sysmgt.sysbrIY58143:devices.chrp.base.rtelY58143:devices.chrp.pci.rtelY58143:devices.chrp.vdevice.rte			
IY58143:devices.common.IBM.atm.rtelY58143:devices.common.IBM.ethernet.rtelY58143:devices.common.IBM.fc.rte			
IY58143:devices.common.IBM.fda.diagIY58143:devices.common.IBM.mpio.rtelY58143:devices.fcp.disk.rte			
IY58143:devices.pci.00100f00.rtelY58143:devices.pci.14100401.diagIY58143:devices.pci.14103302.rte			
IY58143:devices.pci.14106602.rtelY58143:devices.pci.14106902.rtelY58143:devices.pci.14107802.rte			
IY58143:devices.pci.1410ff01.rtelY58143:devices.pci.22106474.rtelY58143:devices.pci.33103500.rte			
IY58143:devices.pci.4f111100.comIY58143:devices.pci.77101223.comIY58143:devices.pci.99172704.rte			
IY58143:devices.pci.c1110358.rtelY58143:devices.pci.df1000f7.comIY58143:devices.pci.df1000f7.diag			
IY58143:devices.pci.df1000fa.rtelY58143:devices.pci.e414a816.rtelY58143:devices.scsi.disk.rte			
IY58143:devices.vdevice.IBM.lan.rtelY58143:devices.vdevice.IBM.vscsi.rtelY58143:devices.vdevice.hvterm1.rte			
IY58143:devices.vtdev.scsi.rtelY58143:sysmgt.websm.appslY58143:sysmgt.websm.frameworkIY58143:sysmgt.websm.rtelY58143:sysmgt.websm.webaccess	IY58143		passed
Node2			
IY58143:X11.Dt.libIY58143:X11.base.rtelY58143:bos.acctIY58143:bos.adt.includelY58143:bos.adt.libmlY58143:bos.adt.proflY58143:bos.alt_disk_install.rtelY58143:bos.diag.com			
IY58143:bos.mp64IY58143:bos.mplIY58143:bos.net.ewlm.rtelY58143:bos.net.ipsec.keymgtIY58143:bos.net.ipsec.rtelY58143: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.controllIY58143:bos.rte.filesystemIY58143:bos.rte.installIY58143:bos.rte.libcIY58143:bos.rte.lvmlIY58143:bos.rte.manIY58143:bos.rte.methodsIY58143:bos.rte.securityIY58143:bos.rte.serv_aidIY58143:bos.sysmgt.nim.clientIY58143:bos.sysmgt.quotallY58143:bos.sysmgt.serv_aid			
IY58143:bos.sysmgt.sysbrIY58143:devices.chrp.base.rtelY58143:devices.chrp.pci.rtelY58143:devices.chrp.vdevice.rte			
IY58143:devices.common.IBM.atm.rtelY58143:devices.common.IBM.ethernet.rtelY58143:devices.common.IBM.fc.rte			
IY58143:devices.common.IBM.fda.diagIY58143:devices.common.IBM.mpio.rtelY58143:devices.fcp.disk.rte			
IY58143:devices.pci.00100f00.rtelY58143:devices.pci.14100401.diagIY58143:devices.pci.14103302.rte			
IY58143:devices.pci.14106602.rtelY58143:devices.pci.14106902.rtelY58143:devices.pci.14107802.rte			
IY58143:devices.pci.1410ff01.rtelY58143:devices.pci.22106474.rtelY58143:devices.pci.33103500.rte			
IY58143:devices.pci.4f111100.comIY58143:devices.pci.77101223.comIY58143:devices.pci.99172704.rte			
IY58143:devices.pci.c1110358.rtelY58143:devices.pci.df1000f7.comIY58143:devices.pci.df1000f7.diag			
IY58143:devices.pci.df1000fa.rtelY58143:devices.pci.e414a816.rtelY58143:devices.scsi.disk.rte			
IY58143:devices.vdevice.IBM.l-			

```
lan.rtelY58143:devices.vdevice.IBM.vscsi.rtelY58143:devices.vdevice.hvterm1.rte
Y58143:devices.vtdev.scsi.rtelY58143:sysmgt.websm.appslY58143:sysmgt.websm.framework
Y58143:sysmgt.websm.rtelY58143:sysmgt.websm.webaccess Y58143 passed
Result: Operating system patch check passed for "Y58143".
```

Check: Operating system patch for "Y66513"

Node Name	Applied	Required	Comment
node1	IY66513:bos.mplY66513:bos.mp64	IY66513	passed
node2	IY66513:bos.mplY66513:bos.mp64	IY66513	passed

Result: Operating system patch check passed for "Y66513".

Check: Operating system patch for "Y70159"

Node Name	Applied	Required	Comment
node1	IY70159:bos.mplY70159:bos.mp64	IY70159	passed
node2	IY70159:bos.mplY70159:bos.mp64	IY70159	passed

Result: Operating system patch check passed for "Y70159".

Check: Operating system patch for "Y59386"

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 "Y59386".

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".

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".

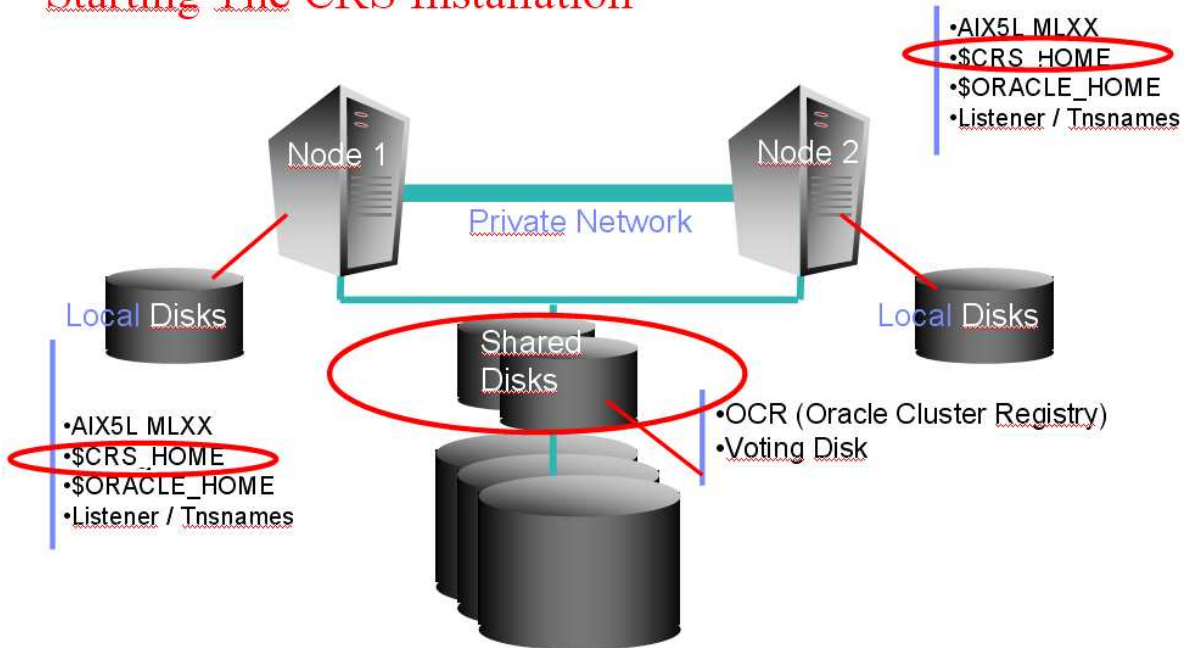
System requirement failed for 'crs'

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.

12 INSTALL THE CUSTER READY SERVICES (CRS)

Starting The CRS 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 ⚡

In -s /usr/sbin/lsattr /etc/lsattr

“/etc/lsattr” is used in vip check action

You can create a staging area. The name of the subdirectories is in the format "Disk1"
As root user, execute :

```
xhost +
```

⚡ Run the AIX command `"/usr/sbin/slibclean"` as "root" to clean all unreferenced libraries from memory

Login as oracle and follow the procedure hereunder...

⚡ Setup and export your DISPLAY, TMP and TEMP variables

- export DISPLAY=.....
- export TMP=/tmp With /tmp or other destination having enough free space, about 500Mb
- export TEMP=/tmp
- 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>.../crs/Disk1/runInstaller  
Or execute : ./crs/Disk1/runInstaller -ignoreSysPrereqs
```

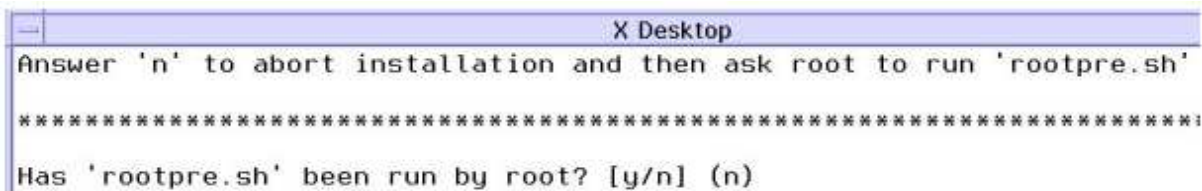
OUI (*Oracle Universal Installer*) chek the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To chek AIX maintenance 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.
```

OUI (*Oracle Universal Installer*) asks you to run `rootpre.sh` as root.

At the OUI Welcome screen, click Next.



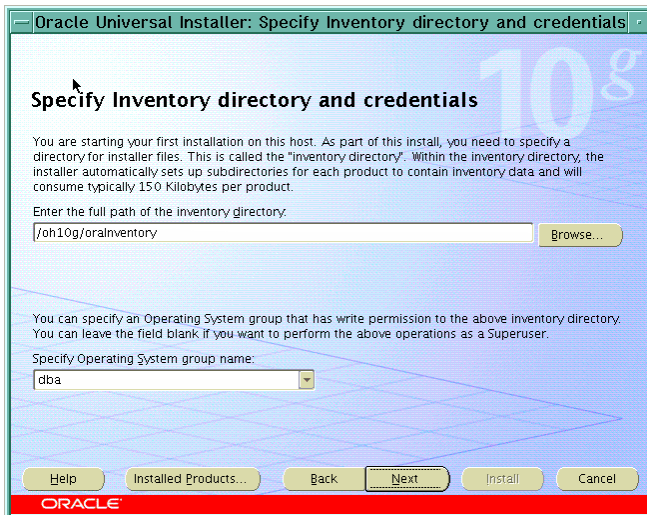
```
X Desktop  
Answer 'n' to abort installation and then ask root to run 'rootpre.sh'  
*****  
Has 'rootpre.sh' been run by root? [y/n] (n)
```

⚡ 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

Just click Next ...

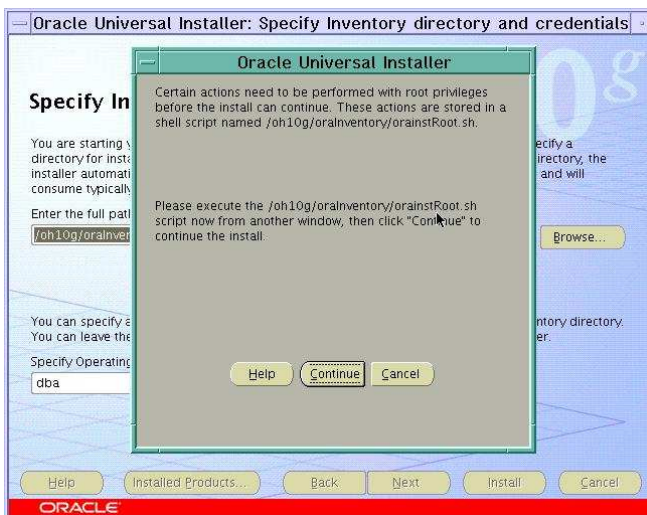


Specify where you want to create the inventory directory.

By default it will be created in the \$ORACLE_BASE

Operating system group name should be set as dba

Then click Next ...

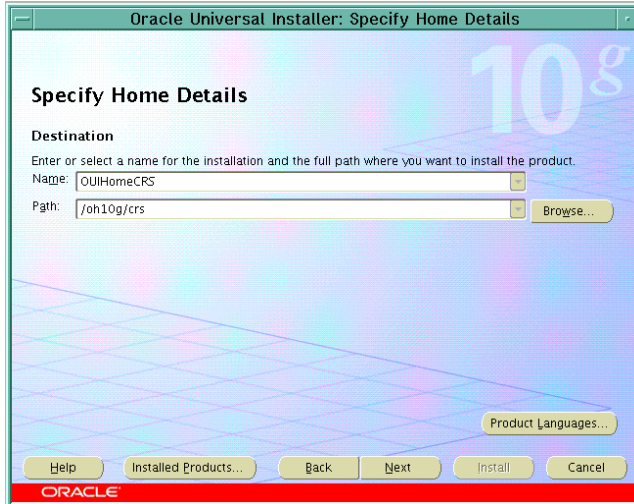


The OUI ask the user to run **orainstRoot.sh** in a separate window, if it's the first Oracle product install on this machine. This script creates the file /etc/orainst.loc, which is used by OUI for the list of installed products.

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 ...

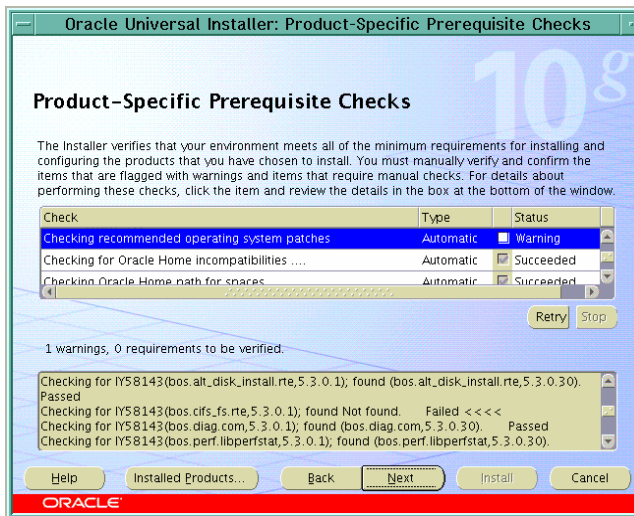


Specify file locations

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.

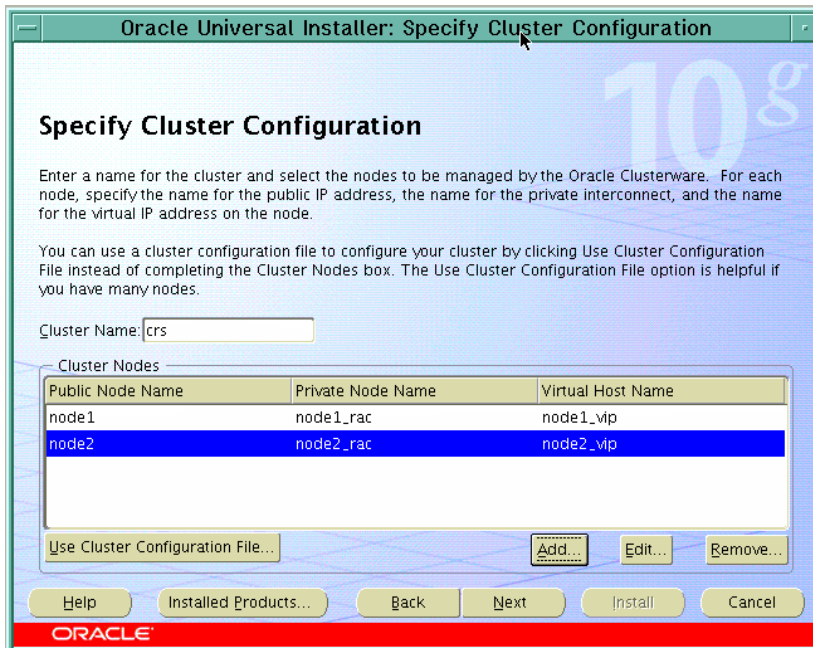
Then click Next ...

➤ Specify Cluster Configuration :

Just to remember !!!

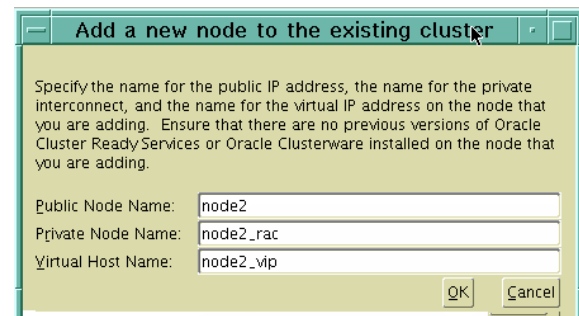
Public, Private, and Virtual Host Name layout

Public		VIP		RAC Interconnect (Private Network)	
en0		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)

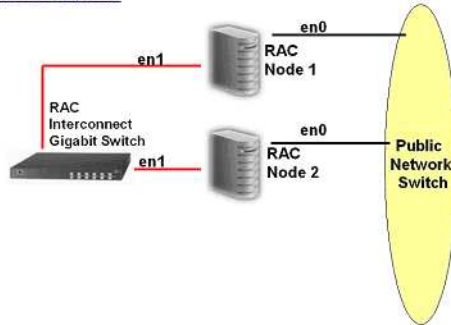
➔ Check your network configuration, and User equivalence configurations on all nodes.

Private Interconnect Enforcement :

Just to remember !!!

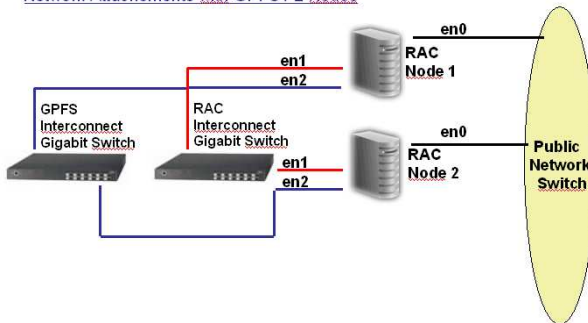
- 1st case – ASM – Network Layout

Network Attachments / 2 Nodes



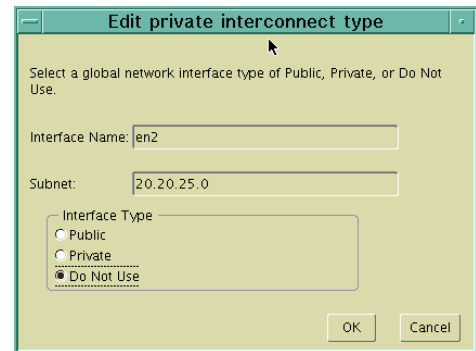
- 2nd case – GPFS – Network Layout

Network Attachments with GPFS / 2 Nodes



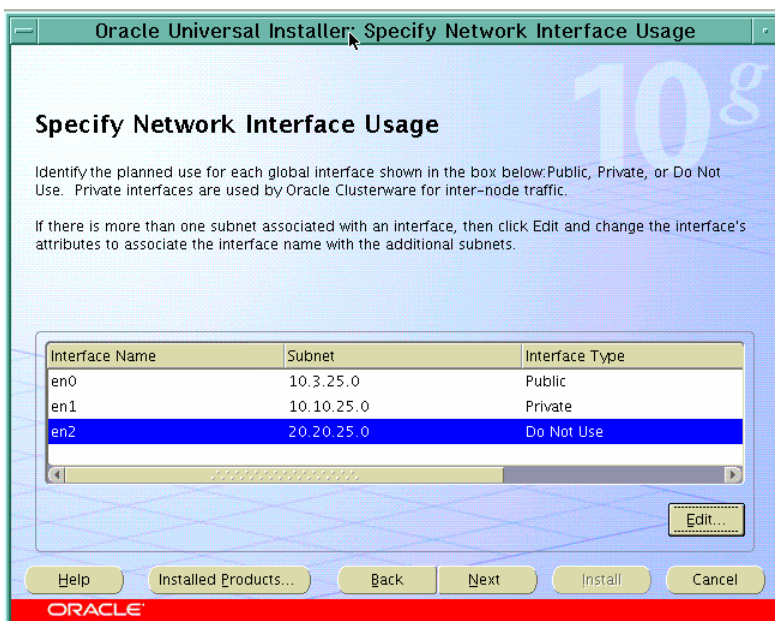
For each entry (en0,en1,en2),

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).



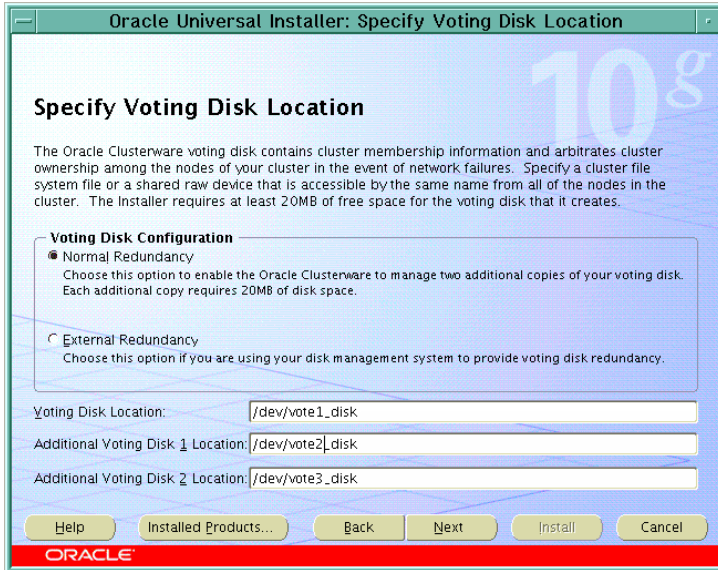
In our example , with or without GPFS implementation :

- 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
- An other en? ...
 - ➔ Do Not Use



Then click Next ...

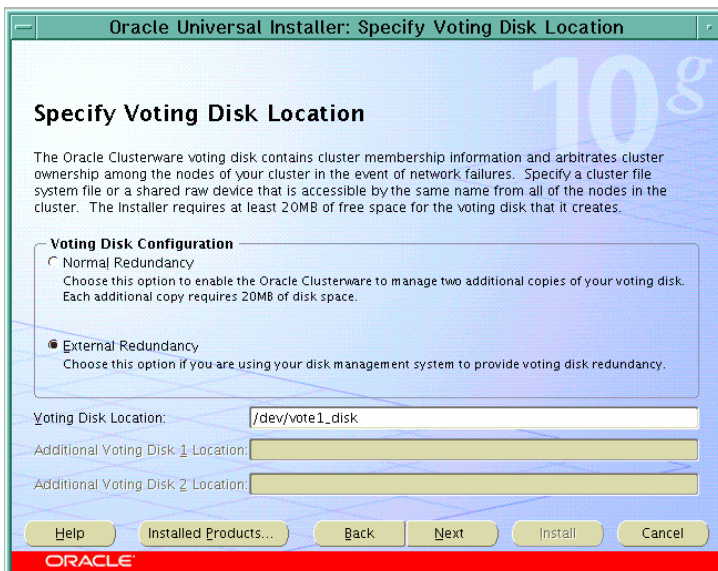
Voting Disk :



Specify Voting Disk Configuration, by selecting :

- **Normal Redundancy** (Voting copies managed by Oracle Clusterware)
 - **External Redundancy** (No Voting copies managed by Oracle Clusterware, should be provided by others options, disks management, etc ...)
- **If ASM implementation is used, and NORMAL Redundancy selected,** specify raw disk location as follow :
 - /dev/vote1_disk
 - /dev/vote2_disk
 - /dev/vote3_disk
 - **If GPFS implementation is used, and NORMAL Redundancy selected** specify files system location as follow:
 - /CRS1/vote1_disk
 - /CRS2/vote2_disk
 - /CRS2/vote3_disk

In our case, we will specify “External Redundancy” as described bellow :

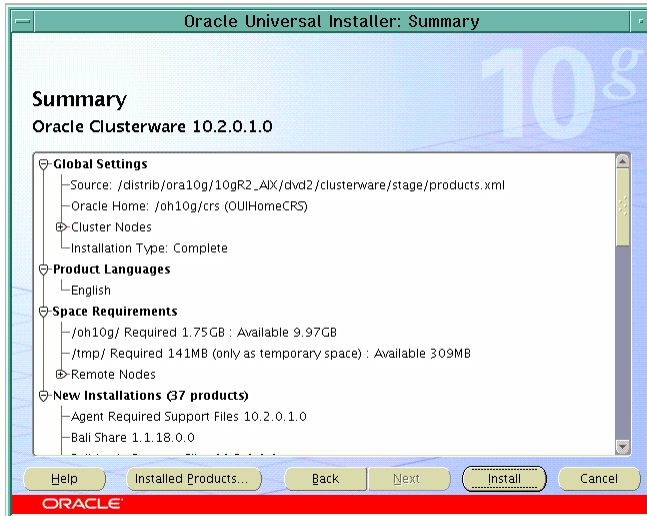


Specify the Voting Disk location, this must be a shared location on the shared storage reachable from all nodes.
And you must have the read/wright permissions on this shared location from all nodes.

- If ASM implementation is used, specify raw disk location
- If GPFS implementation is used, specify files system location

Then click Next ...

If problems happens at this stage, do verify that location specified does exist, and is reachable from each AIX node, with right read/write access, and user/group owner.

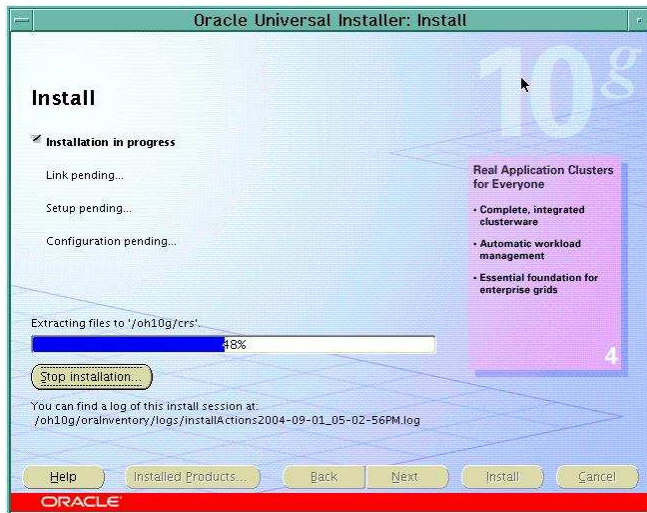


Summary :

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 ...

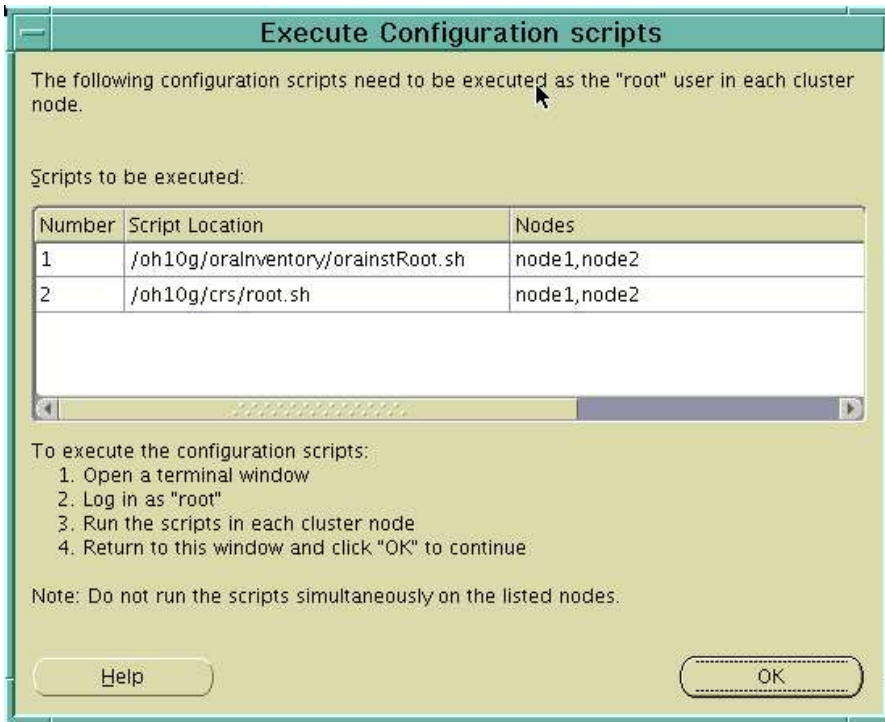


Install :

The Oracle Universal Installer will proceed the installation on the first node, then will copy automatically the code on the 2 others selected nodes.

During the copy, you may be prompt to enter the source location for Disk2.

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 rainstRoot.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

THEN On Node2 as root, execute ./orainstRoot.sh

```
{node1:root}/oh10g/orainventory ->./orainstRoot.sh
Changing permissions of /oh10g/orainventory to 775.
Changing groupname of /oh10g/orainventory to dba.
The execution of the script is complete
{node1:root}/oh10g/orainventory ->
```

```
{node2:root}/oh10g/orainventory ->./orainstRoot.sh
Changing permissions of /oh10g/orainventory to 775.
Changing groupname of /oh10g/orainventory to dba.
The execution of the script is complete
{node2:root}/oh10g/orainventory ->
```

Just before running the root.sh script on each node, do log on as oracle on one node, and execute the following oracle command :

crs_stat -t

```
{node1:root}/ ->su - oracle
{node1:root}/ ce /oh10g/crs/bin
{node1:oracle}/oh10g/crs/bin ->crs_stat -t
CRS-0202: No resources are registered
{node1:oracle}/oh10g/crs/bin ->
```

➔ This is Normal as no ressources are registered at CRS level for this moment.

At this stage, you should execute “root.sh” scripts :

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.

ONLY For information, The root.sh script is executing two sub scripts, and one is the rootconfig.sh script which has interesting information to have a look at :

DO not modify the file !!!

What does the rootconfig.sh script executed by root.sh :

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.

```
# 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

SILENT=false
ORA_CRS_HOME=/oh10g/crs
CRS_ORACLE_OWNER=oracle
CRS_DBA_GROUP=dba
CRS_VNDR_CLUSTER=false
CRS_OCR_LOCATIONS=/dev/ocr1_disk
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/vote1_disk
CRS_NODELIST=node1,node2
CRS_NODEVIPS='node1/node1_vip/255.255.255.0/en0,node2/node2_vip/255.255.255.0/en0'
```

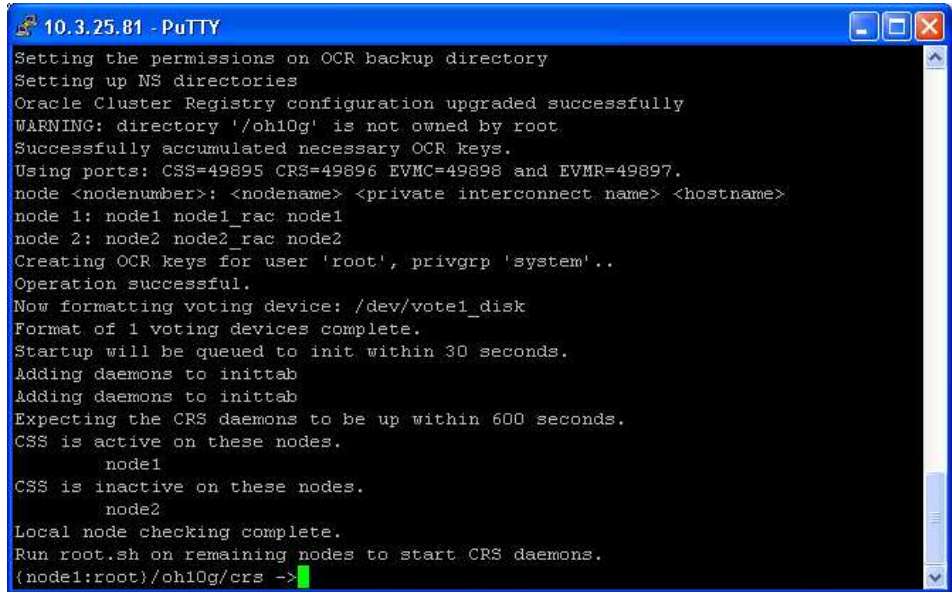

⚡ FIRST On node1

As root, Execute
/oh10g/crs/root.sh

When finished, CSS daemon should be active on node 1.

Check for line "CSS is active on these nodes.
node1

node2



```

10.3.25.81 - PuTTY
Setting the permissions on OCR backup directory
Setting up NS directories
Oracle Cluster Registry configuration upgraded successfully
WARNING: directory '/oh10g' 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/vote1_disk
Format of 1 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)/oh10g/crs ->
  
```

⚡ THEN On node2

As root, Execute
/oh10g/crs/root.sh

When finished, CSS daemon 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}/oh10g/crs/bin ->/oh10g/crs/bin/olsnodes
node1
node2
{node1:oracle}/oh10g/crs/bin ->
rsh node2
{node2:oracle}/oh10g ->olsnodes
node1
node2
{node2:oracle}/oh10g ->
  
```

⚡ On the second node, at the end of the root.sh script :

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 !!!**

At this stage :

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.

```

aixterm
dm          lib32          perl          wwq
emd         md             plsql        xdk
encryption mesg            precomp
has         mgw            racg
{node1:root}/oh10g/db10g -> ./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
ORACLE_OWNER= oracle
ORACLE_HOME= /oh10g/db10g

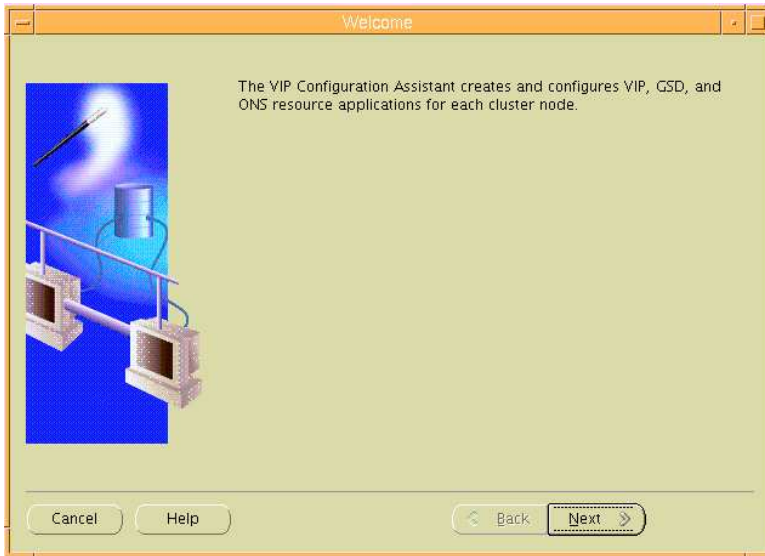
Enter the full pathname of the local bin directory: [/usr/local/bin]:
Copying dbhome to /usr/local/bin ...
Copying oraenv to /usr/local/bin ...
Copying coraenv to /usr/local/bin ...

Creating /etc/oratab file...
Adding entry to /etc/oratab file...
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.
  
```

⚡ On first or second node as root user, you must setup the DISPLAY

before running the vipca script located in /oh10g/crs/bin

Wait for Next Screen ...



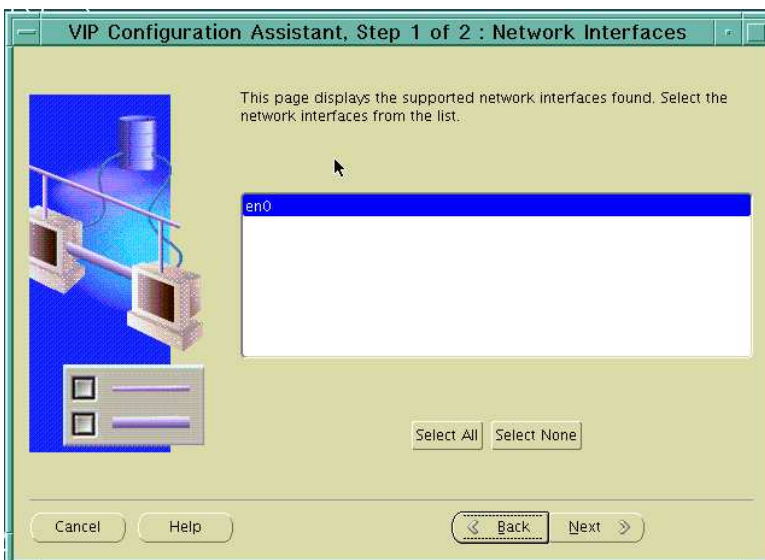
The VIP “Welcome” graphical screen will appear at the end of the root.sh script

Then click Next ...

Just to remember !!!

Public, Private, and Virtual Host Name layout

	Public	VIP	RAC Interconnect	GPFS Interconnect <i>(If GPFS used)</i>
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.

en0 is the RAC Interconnect, or private network.

Please check with “ifconfig -a” on each node as root.

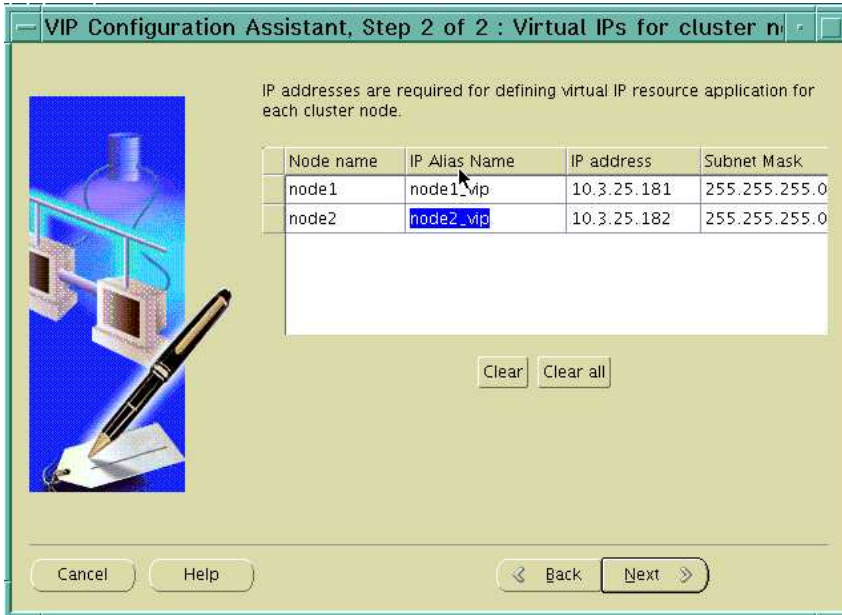
Select “en0” in our case

Then click Next ...

Just to remember !!!

Public, Private, and Virtual Host Name layout

Public		VIP		RAC Interconnect (Private Network)	
en0		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

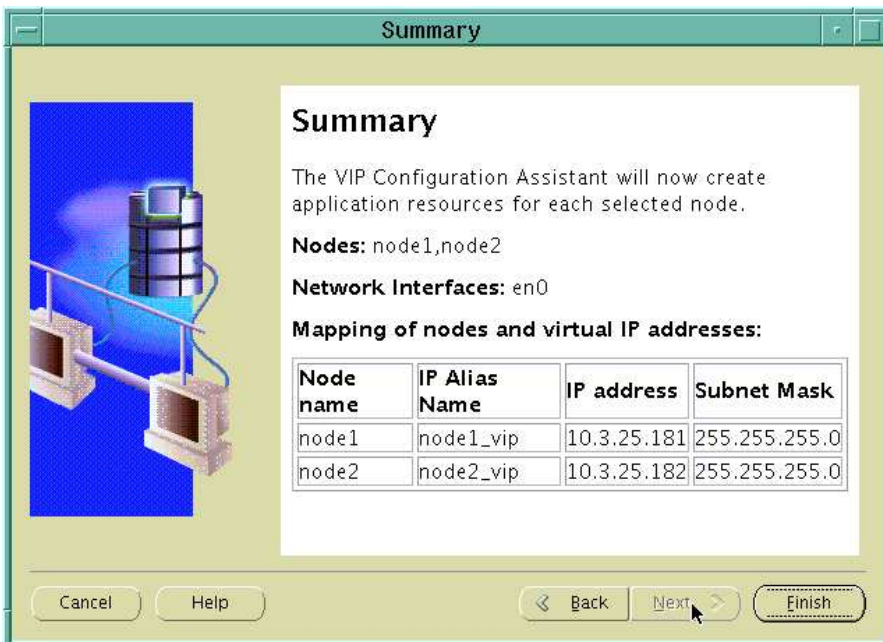


2 of 2 :

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.

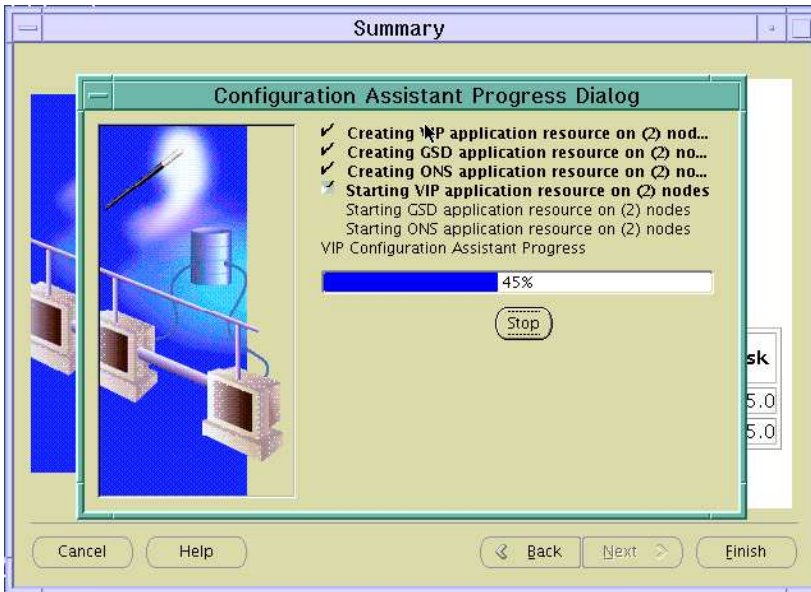
Check validity of the entries before proceeding.

Then click Next ...



The Summary screen will appear, please validate the entries, or go back to modify.

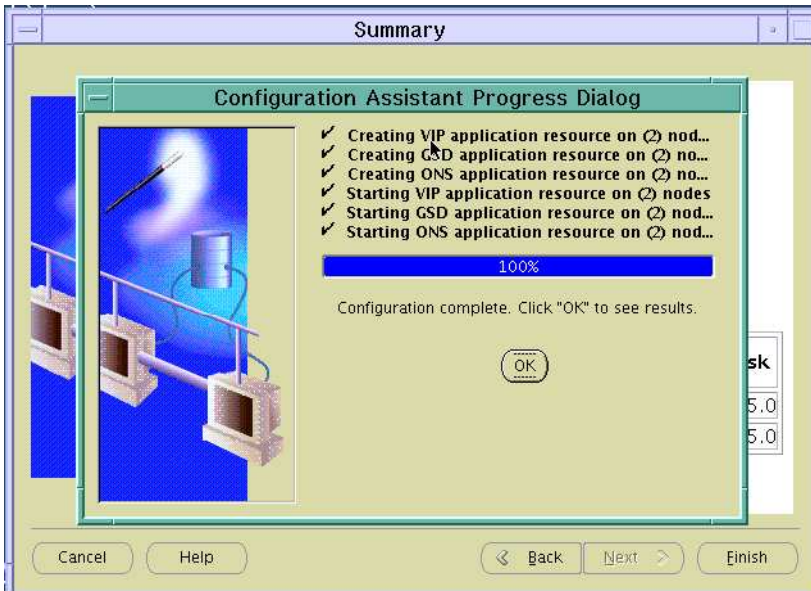
Then click Finish ...



The VIP configuration Assistant will proceed with creation, configuration and startup of all application resources on all selected nodes.

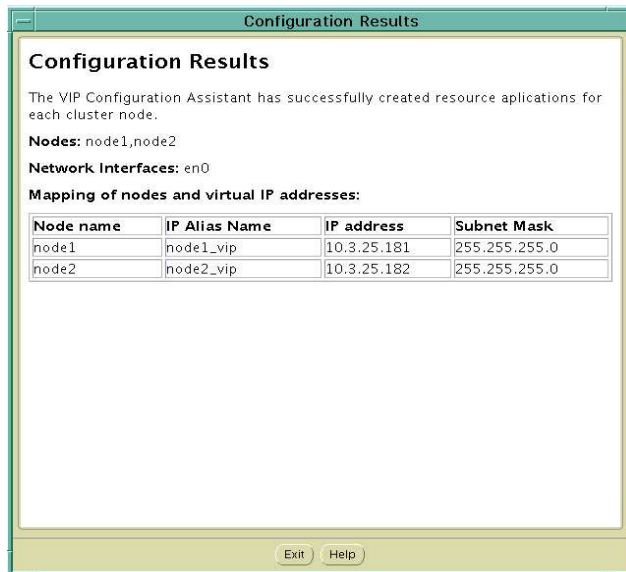
VIP, GSD and ONS will be the application resources to be created.

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 ...

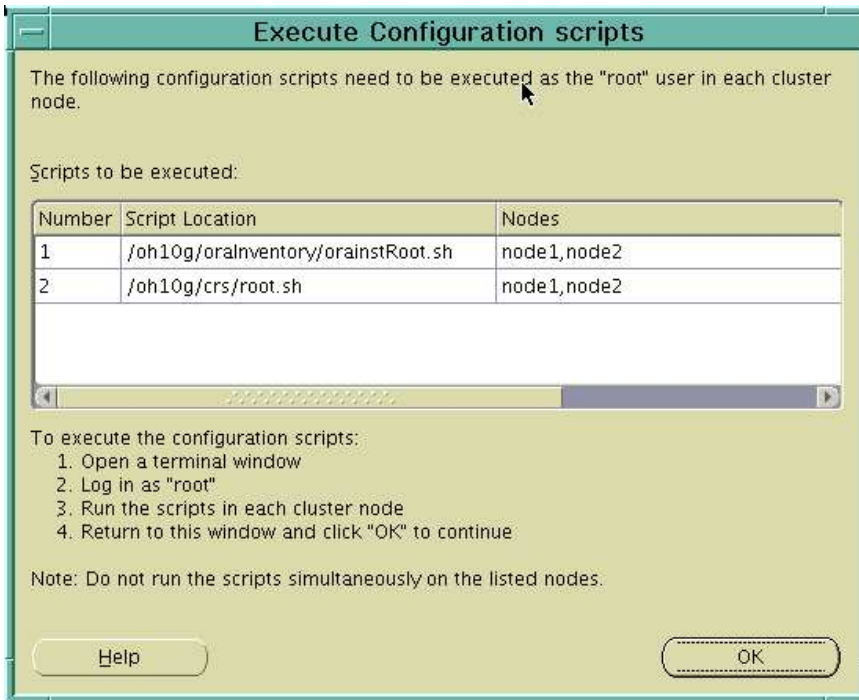
⚡ Using “ifconfig -a” on each node, check that each network card configured for Public network is mapping a virtual IP.

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 0xfffff00 broadcast 10.3.25.255
inet 10.3.25.181 netmask 0xfffff00 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 0xfffff00 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 0xfffff00 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}/ ->
```

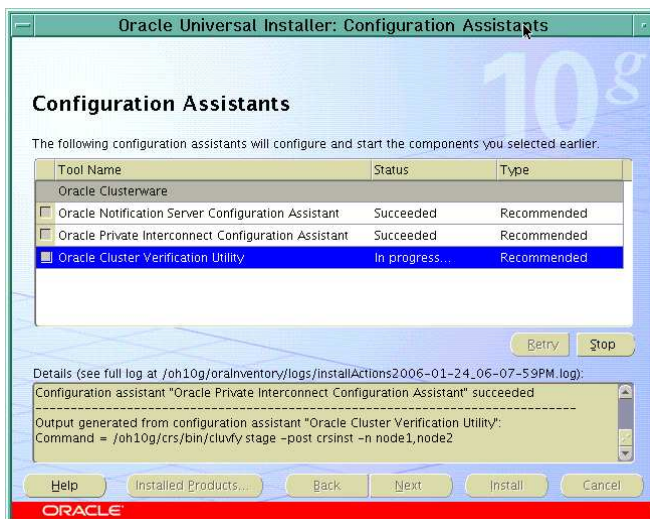
On node 2 :

```
{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 0xfffff00 broadcast 10.3.25.255
inet 10.3.25.182 netmask 0xfffff00 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 0xfffff00 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 0xfffff00 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 ...

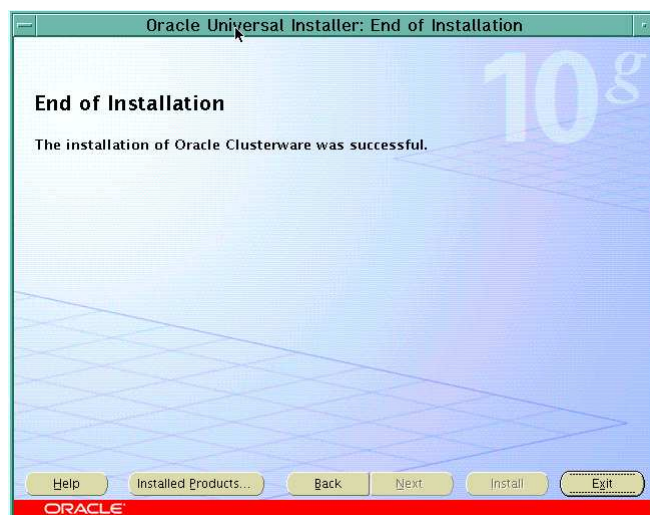


Configuration Assistants :

3 configuration assistants will be automatically executed.

Check for the result to be successful.

Then click Next ...



End of Installation

Then click Exit ...

12.1 CRS POST-INSTALLATION TASKS

12.1.1 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=/oh10g
export AIXTHREAD_SCOPE=S           (S for system-wide thread scope)
export TEMP=/tmp
export TMP=/tmp
export TMPDIR=/tmp
umask 022
export CRS_HOME=$ORACLE_BASE/crs
export ORACLE_CRS_HOME=$ORACLE_BASE/crs
export LD_LIBRARY_PATH=$CRS_HOME/lib:$CRS_HOME/lib32
export PATH=$CRS_HOME/bin:$PATH
```

Do disconnect from oracle user, and reconnect to load modified **\$HOME/.profile**

12.1.2 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}/ ->
```

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.

Execute

crs_stat -t

on one node
as oracle user :

```
{node1:oracle}/oh10g ->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
{node1:oracle}/oh10g ->
```

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:oracle}/oh10g ->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
{node1:oracle}/oh10g ->
```

4. Output of crsctl check crs / cssd / crsd / evmd returns "... daemon appears healthy"

CRS health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

```
{node2:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmd health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

```
{node2:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

CRS software version query

```
{node1:oracle}/oh10g/crs/bin ->crsctl query crs softwareversion
CRS software version on node [node1] is [10.2.0.1.0]
```

```
{node1:oracle}/oh10g/crs/bin ->crsctl query crs softwareversion node2
CRS software version on node [node2] is [10.2.0.1.0]
{node1:oracle}/oh10g/crs/bin ->
```

12.1.3 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
```

This command should return values for global “public” and global “cluster_interconnect”; for example:

```
{node1:oracle}/oh10g ->oifcfg getif
en0 10.3.25.0 global public
en1 10.10.25.0 global cluster_interconnect
{node1:oracle}/oh10g ->
```

If the command does not return a value for global cluster_interconnect, enter the following commands:

```
# oifcfg delif -global
# oifcfg setif -global <interface name>/<subnet>:public
# oifcfg setif -global <interface name>/<subnet>:cluster_interconnect
```

For example:

```
# oifcfg delif -global
# oifcfg setif -global en0/10.3.25.0.0:public
# oifcfg setif -global en1/10.10.25.0.0:cluster_interconnect
```

Enter the following command to verify the new settings:

```
# oifcfg getif
```

- **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.

12.1.4 Oracle CLuster Registry content Check and Backup

Check Oracle Cluster Registry Integrity

As oracle user,
Execute ocrcheck

```
{node1:oracle}/oh10g/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 not configured
```

Cluster registry integrity check succeeded

```
{node1:oracle}/oh10g/crs/bin ->
```

AS root user :

Export Oracle Cluster Registry content

```
{node1:oracle}/oh10g/crs/bin ->su
root's Password:
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export.dmp1 -s online
{node1:oracle}/oh10g/crs/bin ->ls -la /oh10g/*.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp
{node1:oracle}/oh10g/crs/bin ->
```

➔ you can/must not edit/modify this exported file

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oh10g/crs/bin ->ocrconfig -showbackup
node1 2006/01/30 16:03:03 /oh10g/crs/cdata/crs
node1 2006/01/30 12:03:00 /oh10g/crs/cdata/crs
node1 2006/01/30 08:02:59 /oh10g/crs/cdata/crs
node1 2006/01/29 00:02:51 /oh10g/crs/cdata/crs
node1 2006/01/25 13:02:10 /oh10g/crs/cdata/crs
{node1:oracle}/oh10g/crs/bin ->
```

12.2 SOME USEFULL COMMANDS

AS <u>root</u> user	<code>/oh10g/crs/bin/crsctl start crs</code>	to start the CRS
Command to start/stop the CRS deamons :	<code>/oh10g/crs/bin/crsctl stop crs</code>	to stop the CRS

12.3 ACCESSING CRS LOGS

To view CRS logs

`cd /oh10g/crs/nodename/....`

In our case nodename will be node1 for CRS logs on node1
And nodename will be node2 for CRS logs on node2

12.4 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.

12.5 VIP AND CRS TROUBLESHOOTING

⚡ 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](#)

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**

Example for our case :

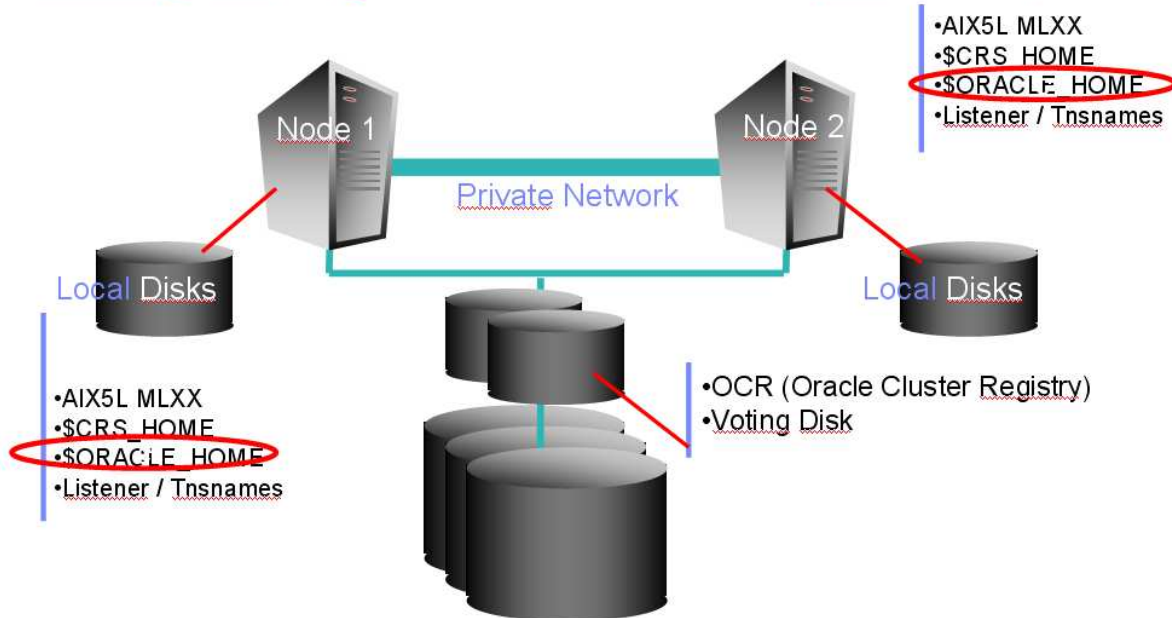
On node1 as root → **`ifconfig en0 delete 10.3.25.181`**
On node2 as root → **`ifconfig en0 delete 10.3.25.182`**

12.6 HOW TO CLEAN A FAILED CRS INSTALLATION

Metalink Note 239998.1 - [10g RAC: How to Clean Up After a Failed CRS Install](#)

13 INSTALL THE ORACLE 10G SOFTWARE

Starting The 10g Software Installation (Incl. RAC)



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"

As root user, execute :

```
xhost +
```

⚡ Run the AIX command **"/usr/sbin/slibclean"** as **"root"** to clean all unreferenced libraries from memory !!!

Login as oracle and follow the procedure hereunder...

⚡ Setup and export your DISPLAY, TMP and TEMP variables

- export DISPLAY=.....
- export TMP=/tmp With /tmp or other destination having enough free space, about 500Mb
- export TEMP=/tmp
- 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`

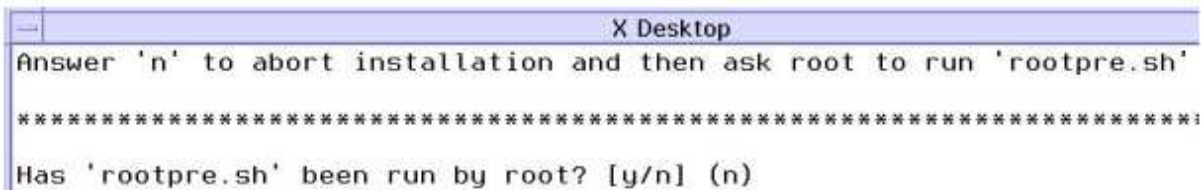
OUI (Oracle Universal Installer) check the operating system requirements for AIX5L 5.3.0.0. If AIX maintenance level 1, 2, 3 are installed, the installer will notice (no further actions) and will go to the next step.

To check AIX maintenance level installed on each node :

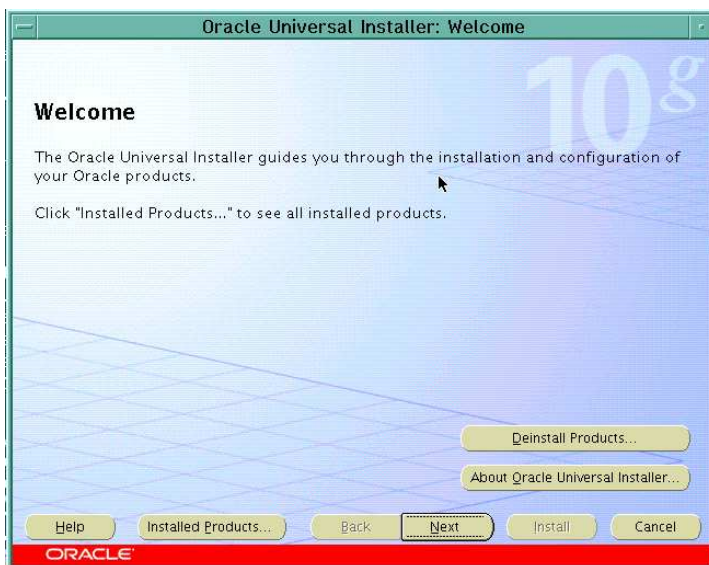
-> `instfix -i|grep ML`

```
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.
```

OUI (Oracle Universal Installer) asks you to run rootpre.sh as root. At the OUI Welcome screen, click Next.

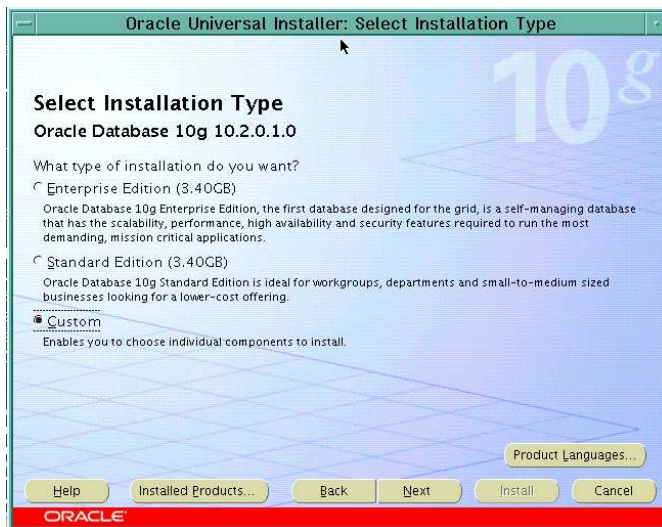


⚡ Make sure to execute rootpre.sh on each node (Should be already done with the CRS Installation), before you click to the next step



At the OUI Welcome screen

Just click Next ...

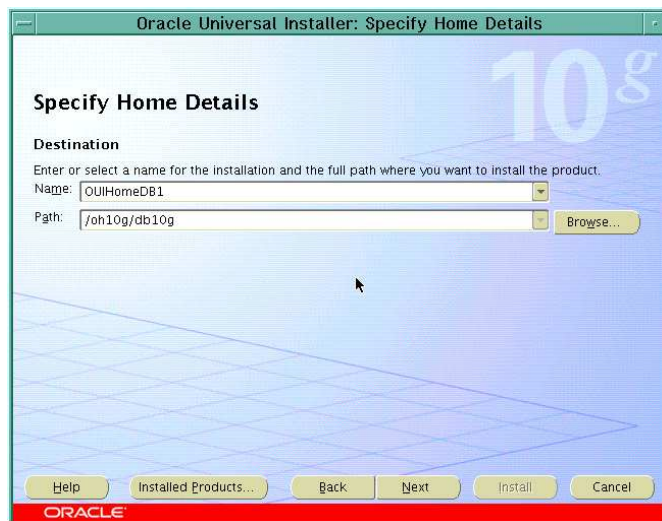


Select the installation type :

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 :

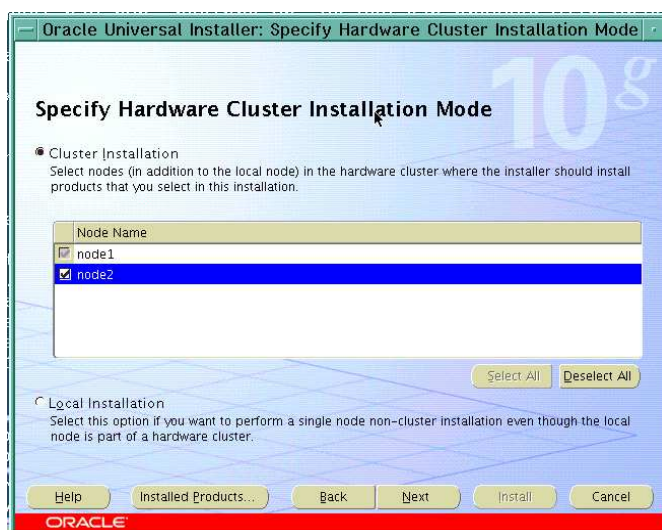
Do not change the Source field

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.

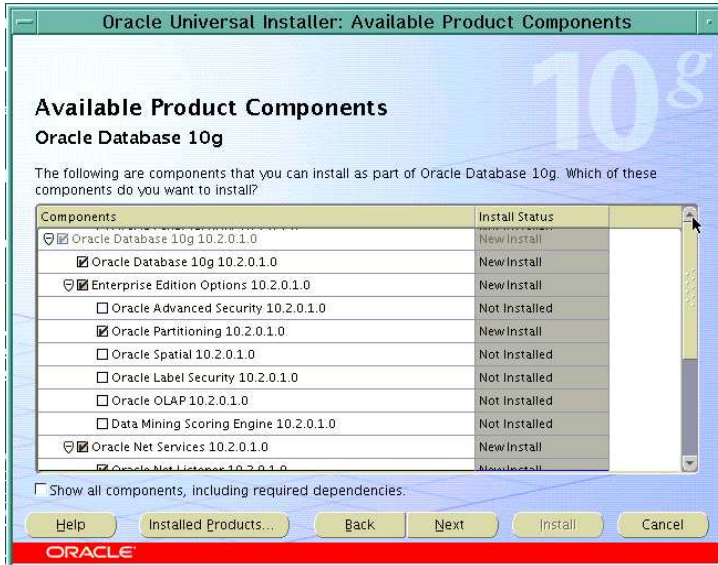


Specify Hardware Cluster Installation Mode :

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 ...

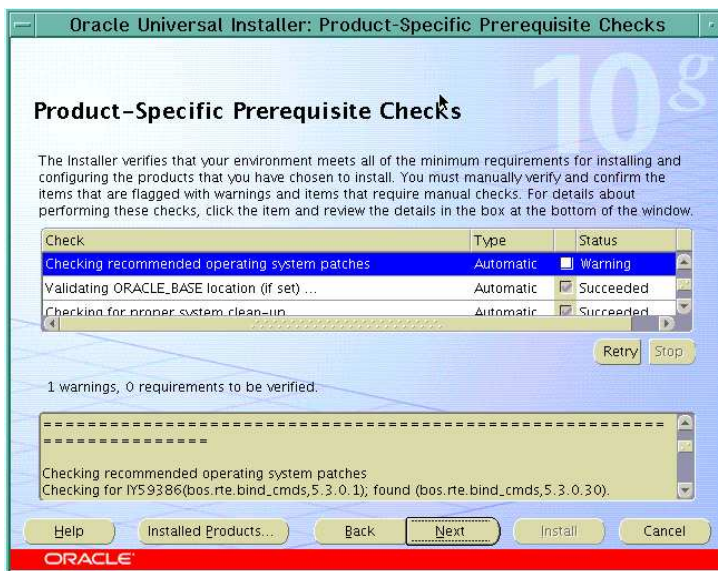


Available Product Components :

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.

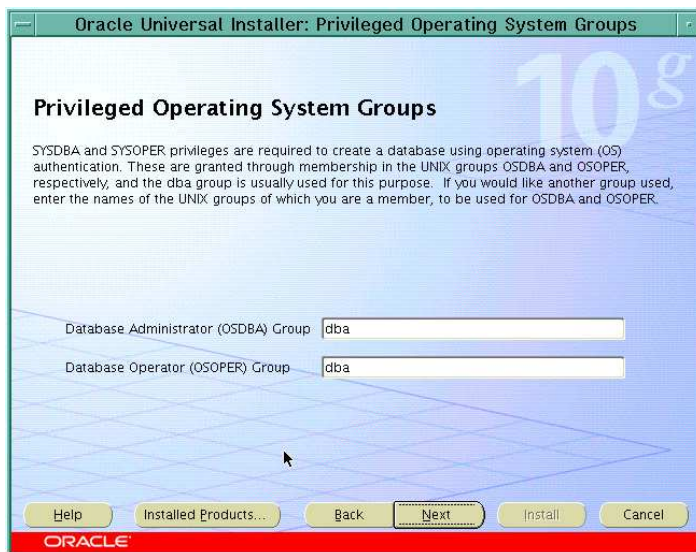
Then click Next ...



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 ...



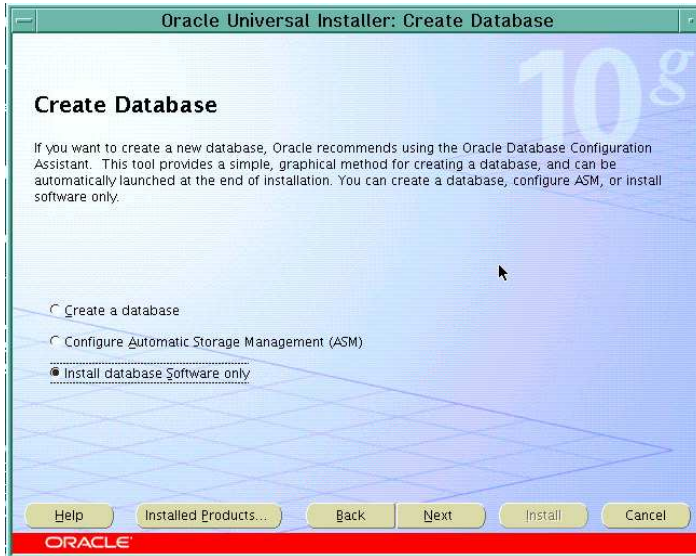
Privileged Operating Systems Groups :

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 “No”, we don't want to create a database at this stage.

Then click Next ...



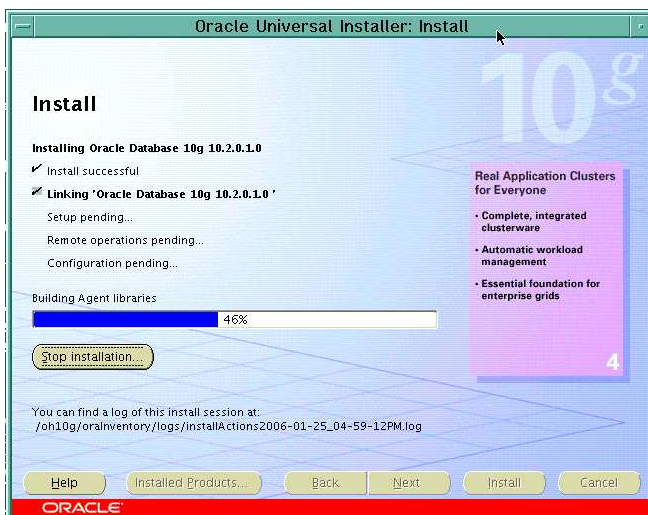
Summary :

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 ...



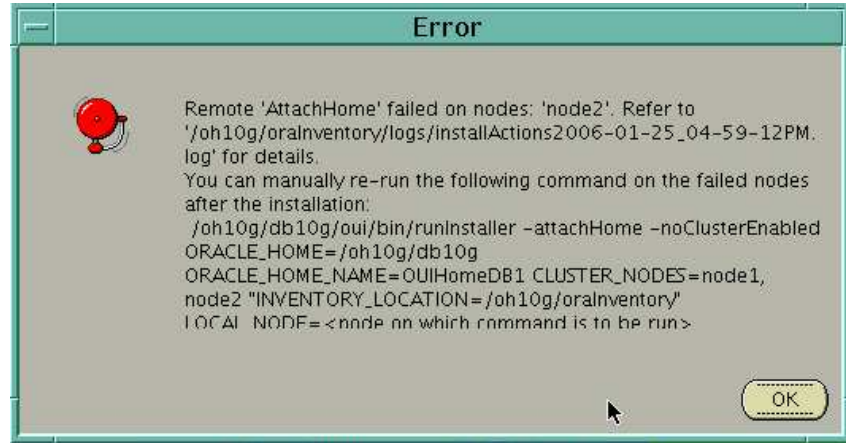
Install :

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 ...

During the copy, you may be prompt to enter the source location for Disk2 and Disk3.

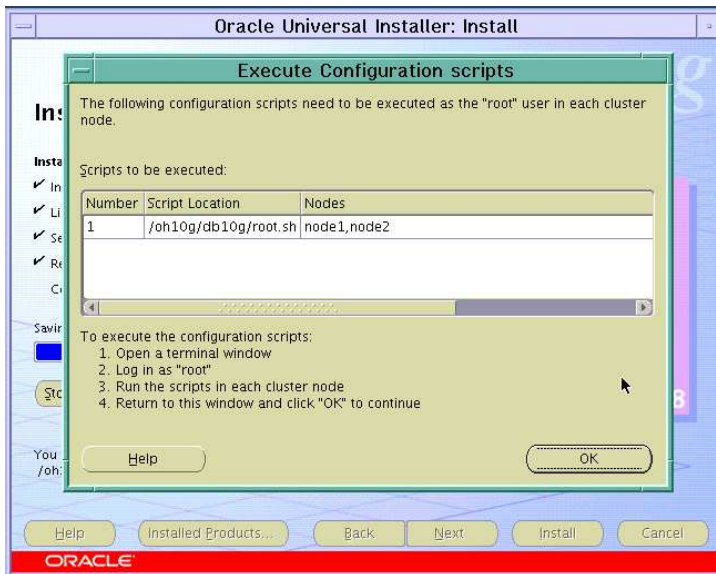
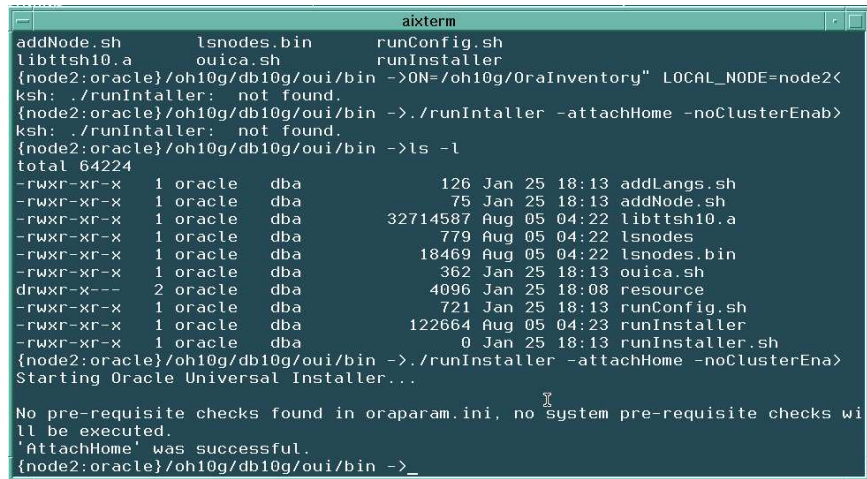
⚡ If you get this following error message →



Do execute the script below :

```
./runInstaller -attachHome -noClusterEnabled ORACLE_HOME=/oh10g/db10g
ORACLE_HOME_NAME=OUIHomeDB1 CLUSTER_NODES=node1,node2
"INVENTORY_LOCATION=/oh10g/Orainventory" LOCAL_NODE=node2
```

And check the result to be succeed !!



Setup Privileges :

At this stage, run the "root.sh" script on each selected nodes

Execute the root.sh on each node ...


```

129.35.161.127 - PuTTY
(node1:root)/ ->cd /oh10g/db10g
(node1:root)/oh10g/db10g ->./root.sh
Running Oracle10 root.sh script...

The following environment variables are set as:
ORACLE_OWNER= oracle
ORACLE_HOME= /oh10g/db10g

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]: y
Copying dbhome to /usr/local/bin ...
The file "oraenv" already exists in /usr/local/bin. Overwrite it? (y/n) [n]: y
Copying oraenv to /usr/local/bin ...
The file "coraenv" already exists in /usr/local/bin. Overwrite it? (y/n) [n]: y
Copying coraenv to /usr/local/bin ...

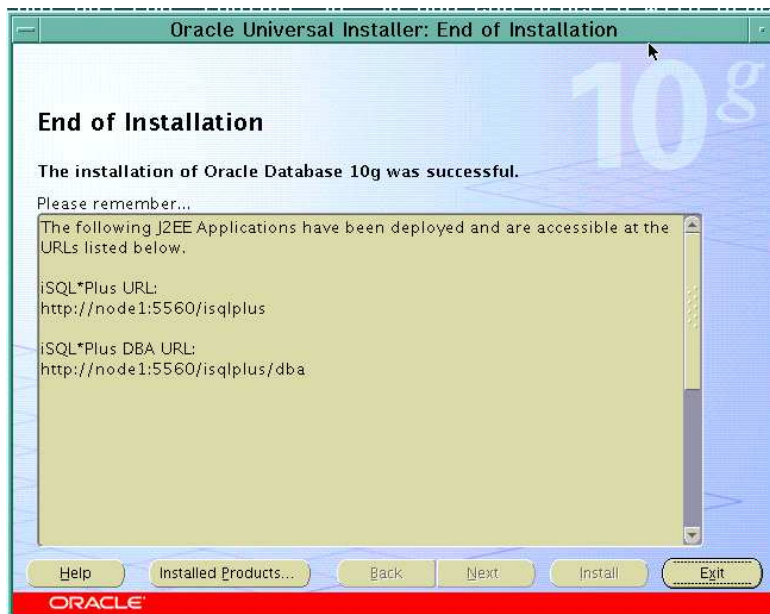
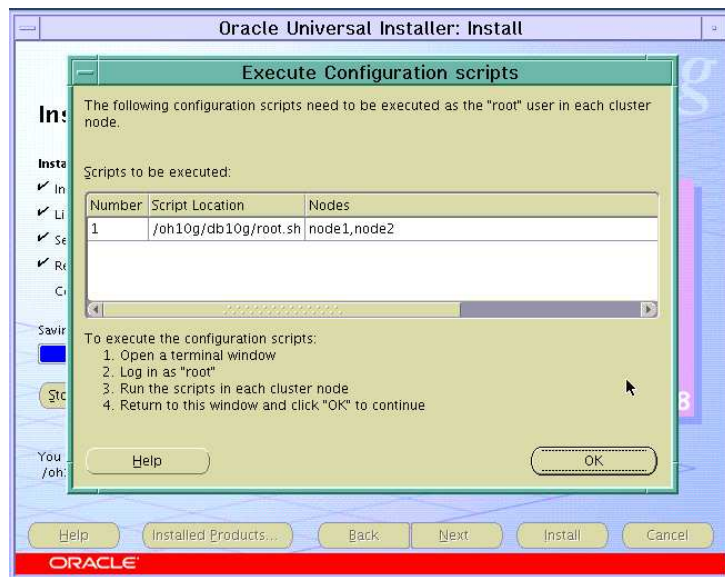
Creating /etc/oratab file...
Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.

(node1:root)/oh10g/db10g ->

```

On first node,
Then second node ...

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 ...

13.1 10G SOFTWARE POST-INSTALLATION TASK

13.1.1 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=/oh10g
export AIXTHREAD_SCOPE=S
export TMP=/tmp
export TEMP=/tmp
export TMPDIR=/tmp
umask 022
export CRS_HOME=$ORACLE_BASE/crs
export ORACLE_CRS_HOME=$ORACLE_BASE/crs

export ORACLE_HOME=$ORACLE_BASE/db10g
export LD_LIBRARY_PATH=$ORACLE_HOME/lib:$ CRS_HOME/lib:$ORACLE_HOME/lib32:$ CRS_HOME/lib32
export PATH=$ORACLE_HOME/bin:$CRS_HOME/bin:$PATH
```

Do disconnect from oracle user, and reconnect to load modified **\$HOME/.profile**

13.1.2 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}/oh10g ->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
{node1:oracle}/oh10g ->
```

CR health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmhealth check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

13.2 WHAT HAS BEEN DONE ?

At this stage :

- The Oracle 10g Software with the Real Application Cluster option is installed.

14 ORACLE LISTENER CONFIGURATION

Connect as Oracle unix user from first node, and **setup your DISPLAY**
Execute **netca &** to lanch the database configuration assistant

`./netca`



Select "Perform Typical Configuration"

Then click Next ...

14.1 LISTENER POST-CONFIGURATION TASK

Execute

`crs_stat -t`

on one node
as oracle user :

```
{node1:oracle}/oh10g ->crs_stat -t
Name      Type      Target State  Host
-----
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....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}/oh10g ->
```

AS root user :

Export Oracle
Cluster Registry
content

```
{node1:oracle}/oh10g/crs/bin ->su
root's Password:
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export2.dmp -s online
{node1:oracle}/oh10g/crs/bin ->ls -la /oh10g/*.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp
{node1:oracle}/oh10g/crs/bin ->
```

➔ you can/must not edit/modify this exported file

14.2 WHAT HAS BEEN DONE ?

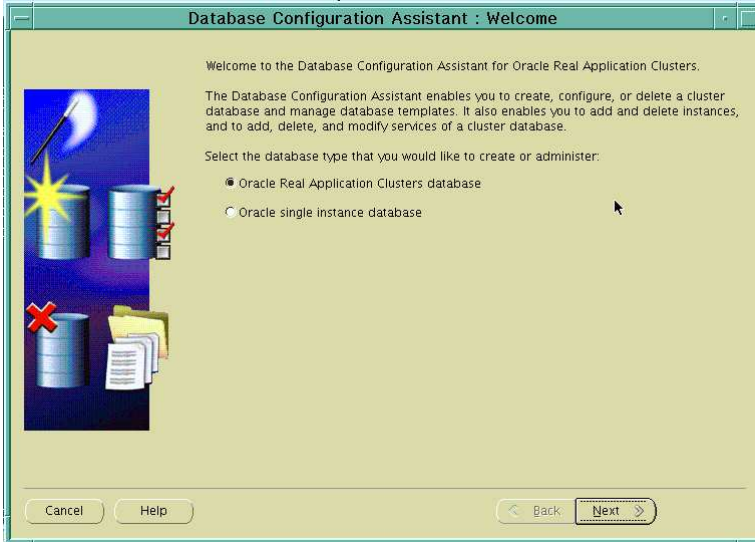
At this stage :

- The listeners for both nodes are configured and started.

15 ASM CONFIGURATION (ASM IMPLEMENTATION ONLY)

Connect as Oracle unix user from first node, and **setup your DISPLAY**
Execute **dbca &** to lanch the database configuration assistant

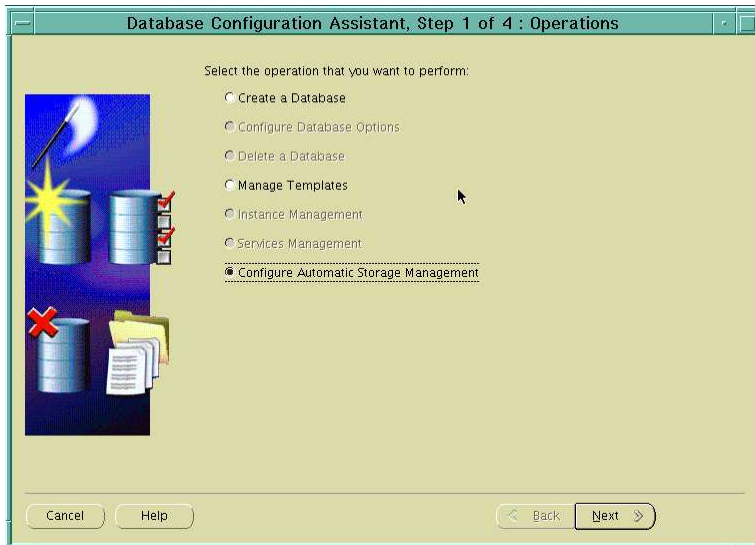
```
./dbca
```



DBCA Welcome Screen

Select the "Oracle Real Application Cluster Database" option.

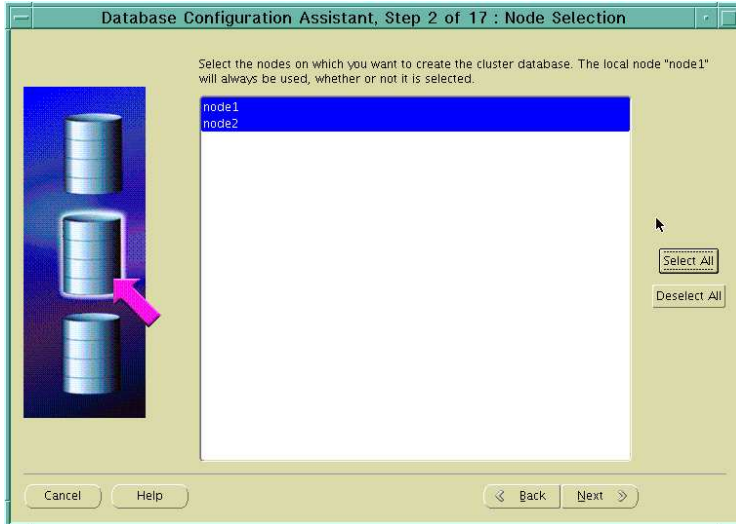
Then click Next ...



Operations :

Select the "**Configure Automatic Storage Management**" option.

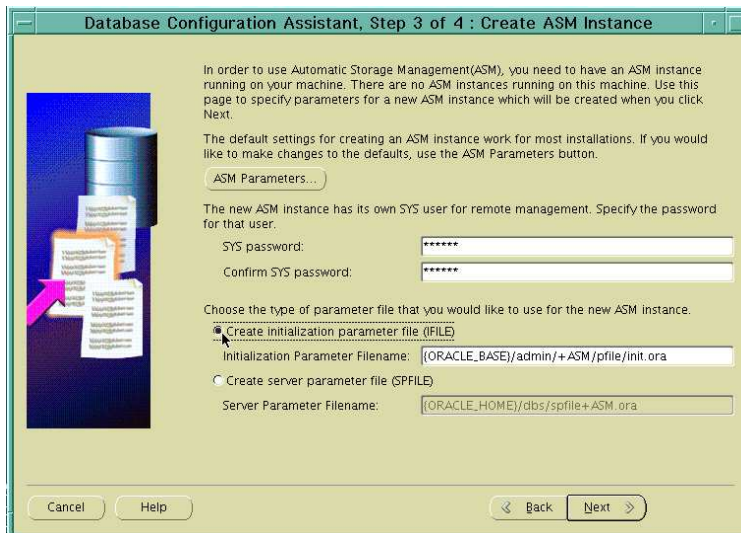
Then click Next ...



Node Selection :

Make sure to select all RAC nodes.

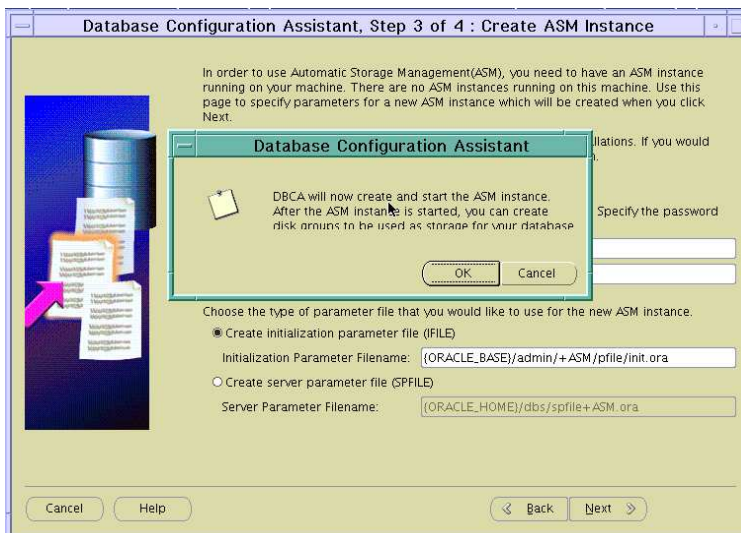
Then click Next ...



Create ASM Instance :

- **Specify SYS password for the ASM Instance**
- **Choose the type of parameter file "Create initialization parameter file (IFILE)"**

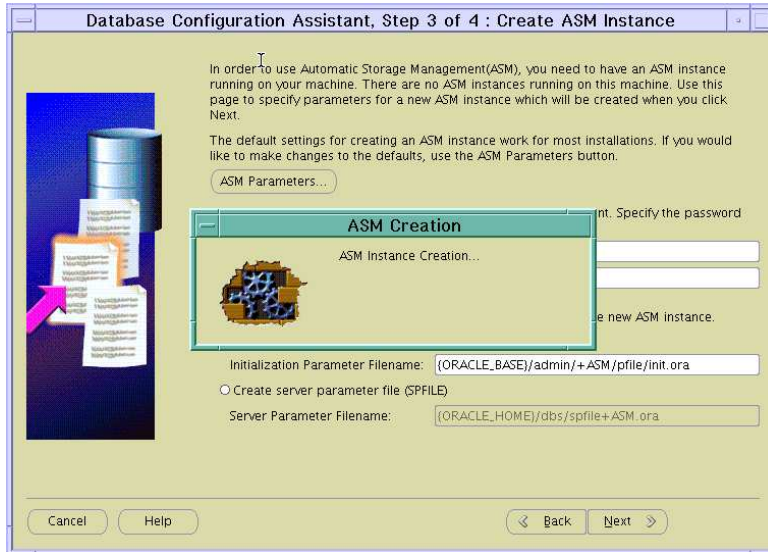
Then click Next ...



Create ASM Instance :

The assistant will create the ASM instance on each node.

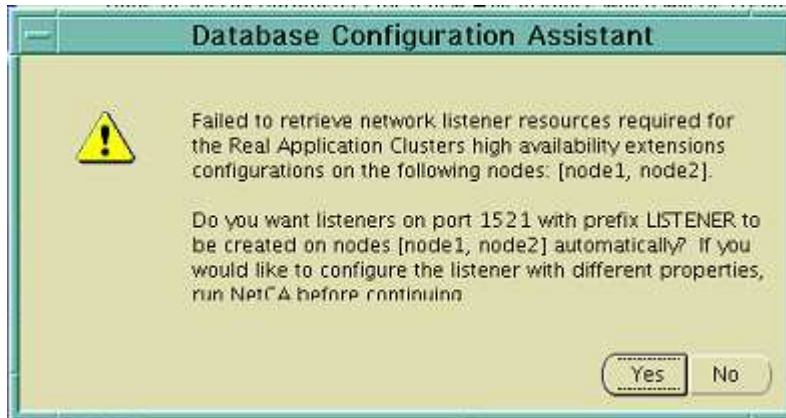
Then click OK..



Create ASM Instance :

Assistant is creating the ASM instance on each node.

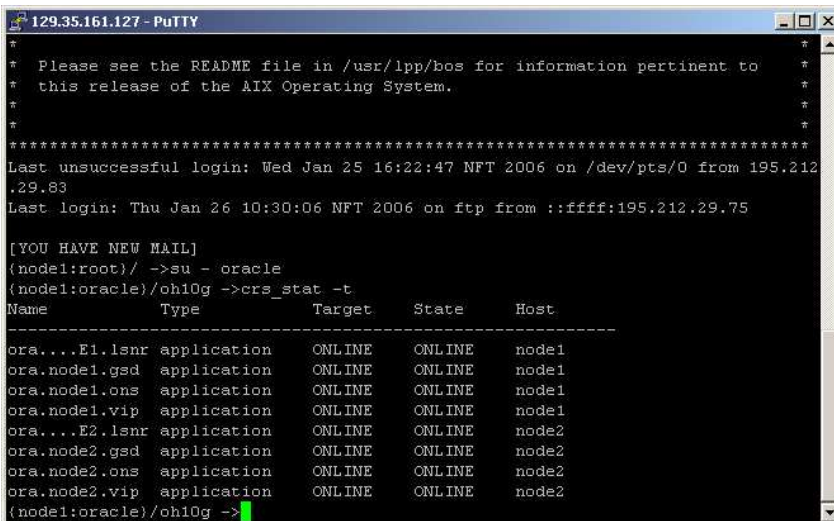
Just wait ...



Create ASM Instance :

This message will appears if Oracle listeners are not configured on all nodes.

IF SO → Just Click YES to create oracle listeners on all nodes ...

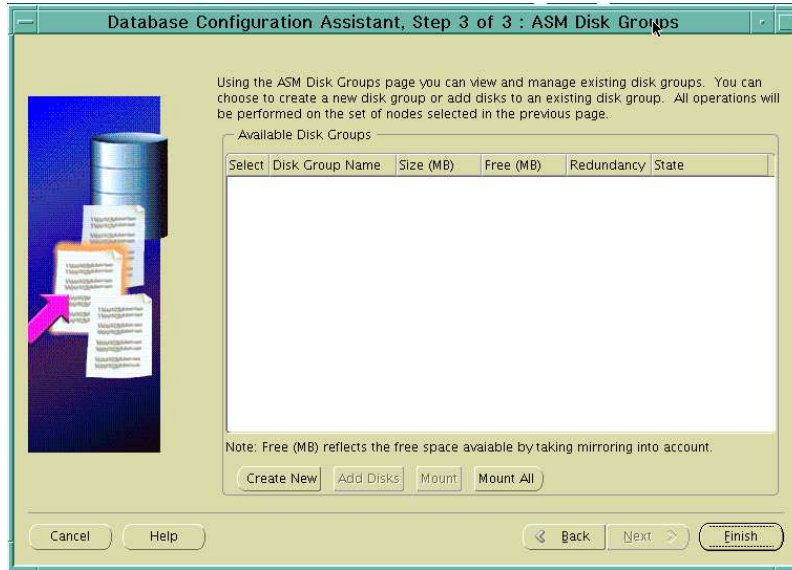


When assistant has finished to create the listeners :

Do OPEN an AIX terminal as oracle, and check with :

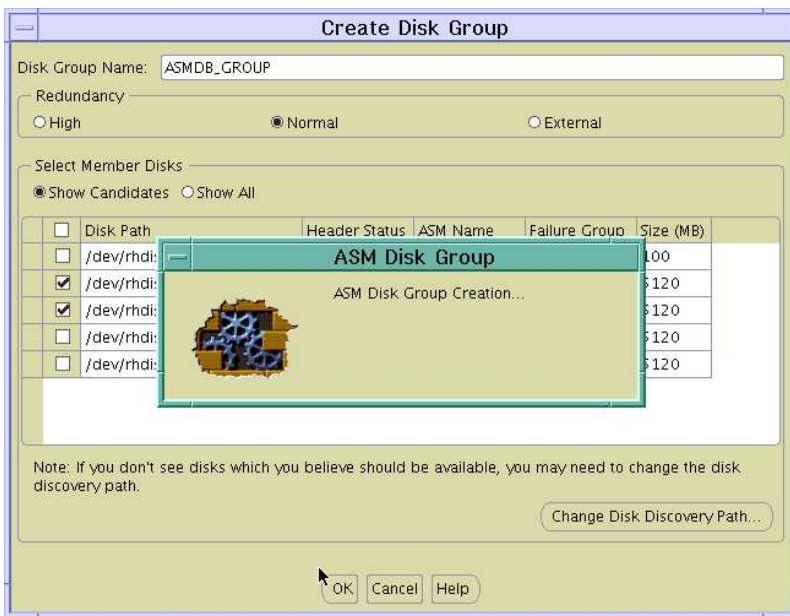
crs_stat -t

To see that listeners have been created on both nodes.



ASM Disk Groups :

Click on **“Create New”** ...



Create Disk Group :

Select **Show Candidates**

Specify Disk Group Name :

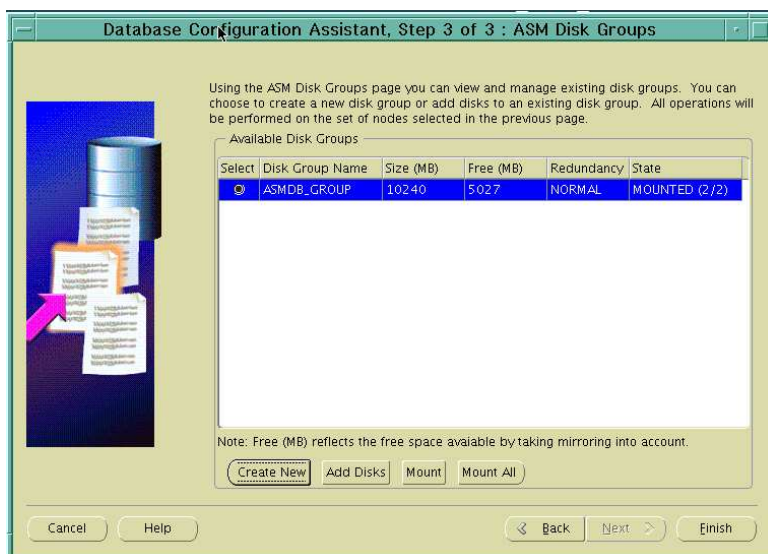
For our example, we specified **ASMDB_GROUP**

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...

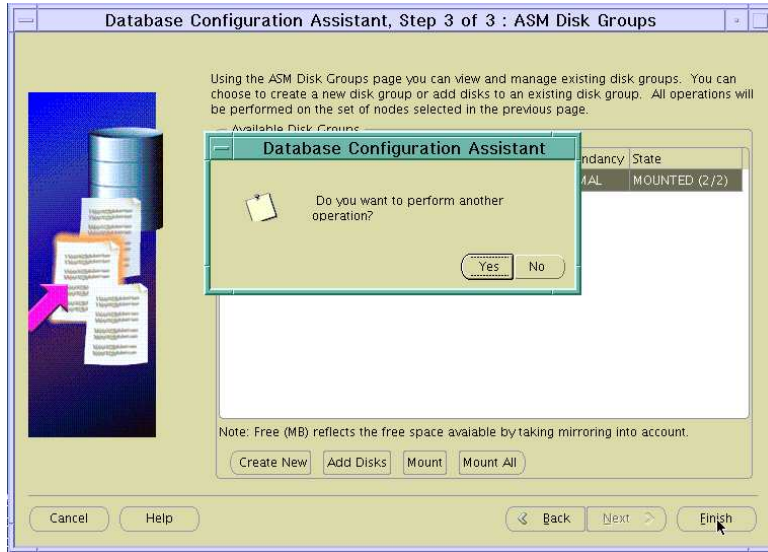


ASM Disk Groups :

ASMDB_GROUP Disk Group created

SELECT **“Create New”** for new Disk Group.

OR **“Finish”** to exit



Click “Yes” to perform an other operation

OR “No” to exit ...

15.1 ASM POST-CONFIGURATION TASK

15.1.1 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}/oh10g ->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}/oh10g ->
```

CR health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmhealth
check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```


Check ASM deamons

For example on node1 :

```
{node1:oracle}/oh10g/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/db10g/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
```

15.1.2 Oracle CLuster Registry content Check and Backup
Check Oracle Cluster Registry Integrity
**As oracle user,
Execute ocrcheck**

```
{node1:oracle}/oh10g/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 not configured
```

Cluster registry integrity check succeeded

```
{node1:oracle}/oh10g/crs/bin ->
```

AS root user :
Export Oracle Cluster Registry content

```
{node1:oracle}/oh10g/crs/bin ->su
root's Password:
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export3.dmp -s online
{node1:oracle}/oh10g/crs/bin ->ls -la /oh10g/*.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp
{node1:oracle}/oh10g/crs/bin ->
```

➔ you can/must not edit/modify this exported file

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oh10g/crs/bin ->ocrconfig -showbackup

node1 2006/01/30 16:03:03 /oh10g/crs/cdata/crs
node1 2006/01/30 12:03:00 /oh10g/crs/cdata/crs
```

```
node1 2006/01/30 08:02:59 /oh10g/crs/cdata/crs
node1 2006/01/29 00:02:51 /oh10g/crs/cdata/crs
node1 2006/01/25 13:02:10 /oh10g/crs/cdata/crs
{node1:oracle}/oh10g/crs/bin ->
```


15.2 SOME USEFULL COMMANDS

AS oracle user

Command to start/stop the ASM Instances :

```
For node 1
srvctl start asm -n node1    to start the ASM instance
srvctl stop asm -n node1    to stop the ASM instance

For node 2
srvctl start asm -n node2    to start the ASM instance
srvctl stop asm -n node2    to stop the ASM instance
```

To acces an ASM instance with sqlplus

```
From node1

set ORACLE_SID=+ASM1
sqlplus /nolog
connect /as sysdba
show sga
.....

From node2

set ORACLE_SID=+ASM2
sqlplus /nolog
connect /as sysdba
show sga
.....
```

15.3 WHAT HAS BEEN DONE ?

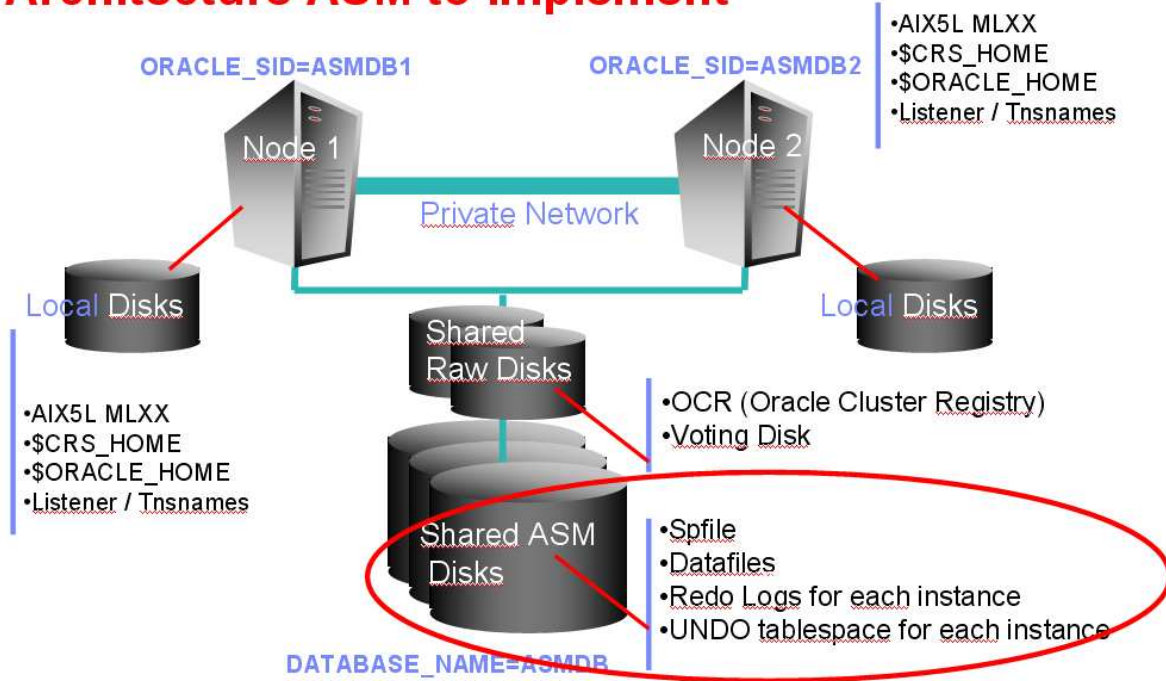
At this stage :

- ASM has been configured on both nodes
- The listeners are configured and started.

16 CREATING THE DATABASE USING DBCA

16.1 DATABASE CREATION ON ASM

Architecture ASM to implement



Connect as Oracle unix user from first node, and **setup your DISPLAY**
Execute **dbca &** to launch the database configuration assistant

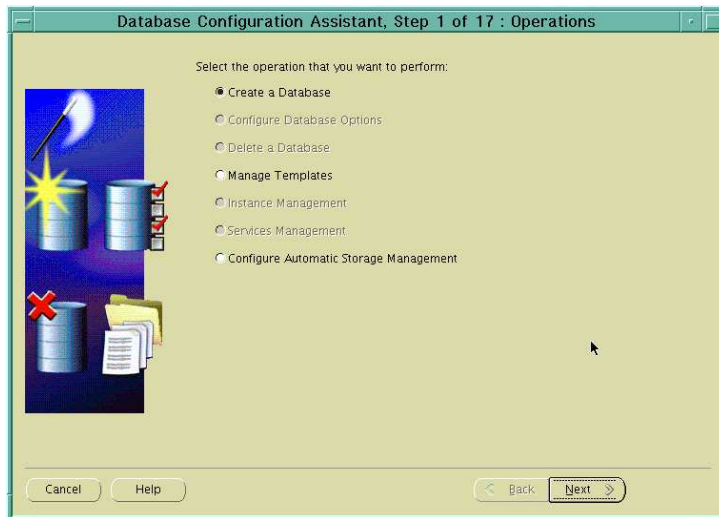
```
./dbca
```



DBCA Welcome Screen :

Select the "Oracle Real Application Cluster Database" option.

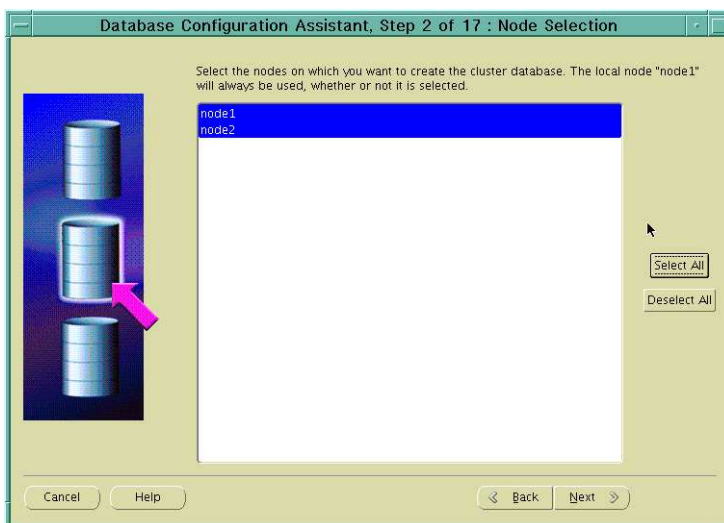
Then click Next ...



Operations :

Select the "Create a Database" option.

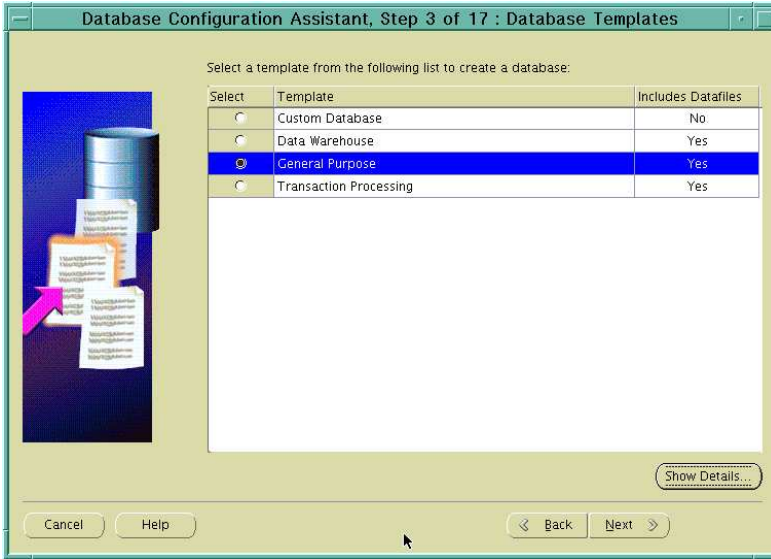
Then click Next ...



Node Selection :

Make sure to select all RAC nodes.

Then click Next ...

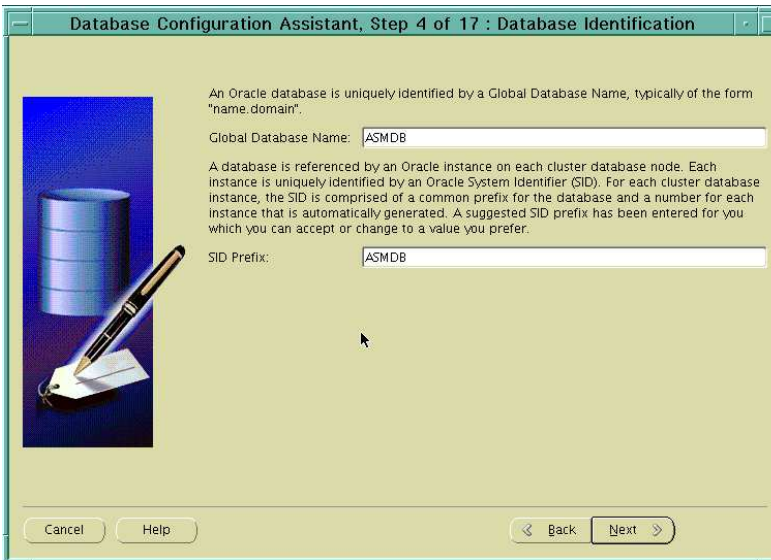


Database Templates :

Select **“General Purpose”**

Or “Custom Database” if you want to generate the creation scripts.

Then click Next ...

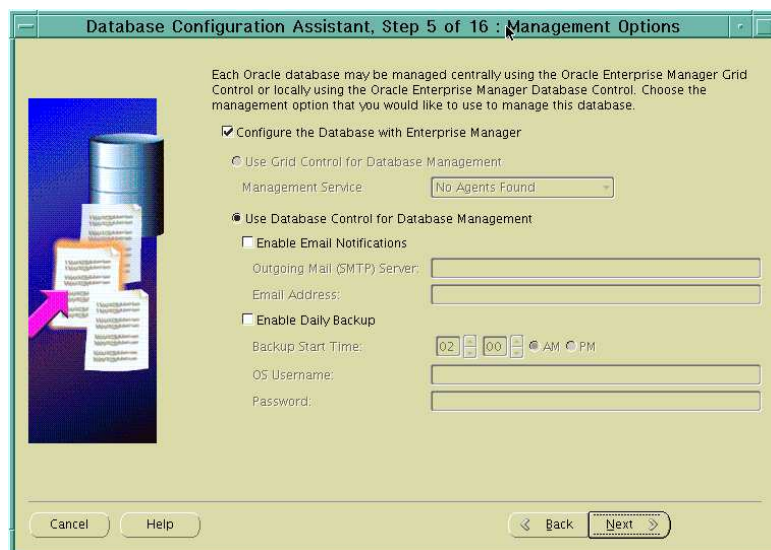


Database Identification :

Specify the **“Global Database Name”**

The “SID Prefix” will be automatically updated. (by default it is the Global Database Name)

Then click Next ...

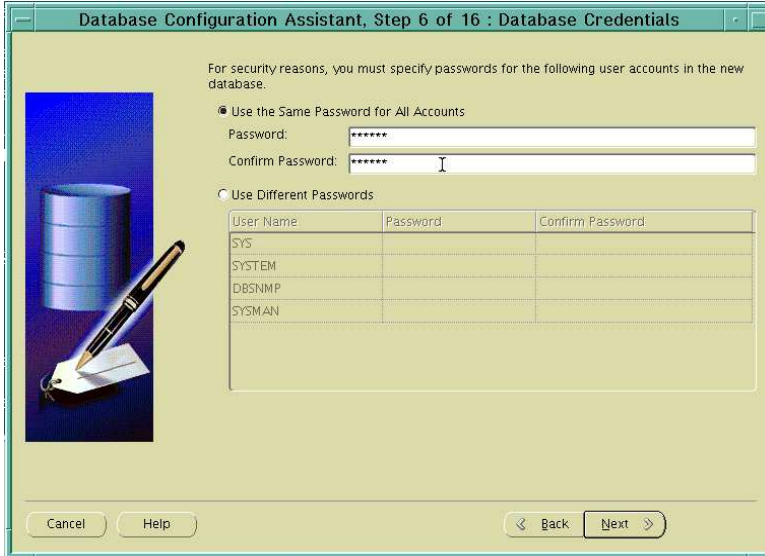


Management Options :

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)

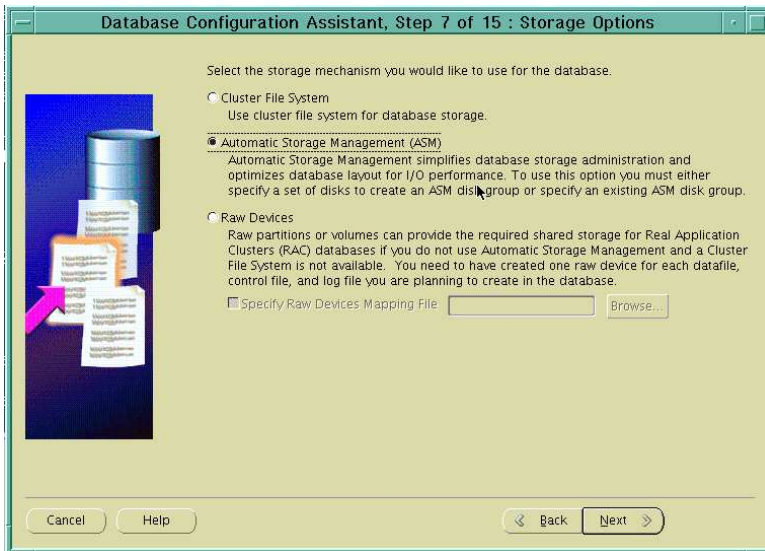
Then click Next ...



Database Credentials :

Specify same password for all administrator users, or specify individual password for each user.

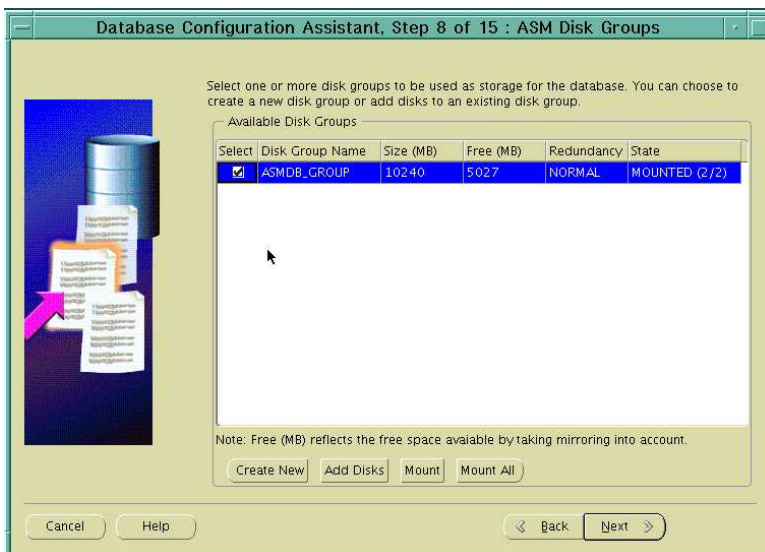
Then click Next ...



Storage Options :

Choose Automatic Storage Management (ASM)

Then click Next ...

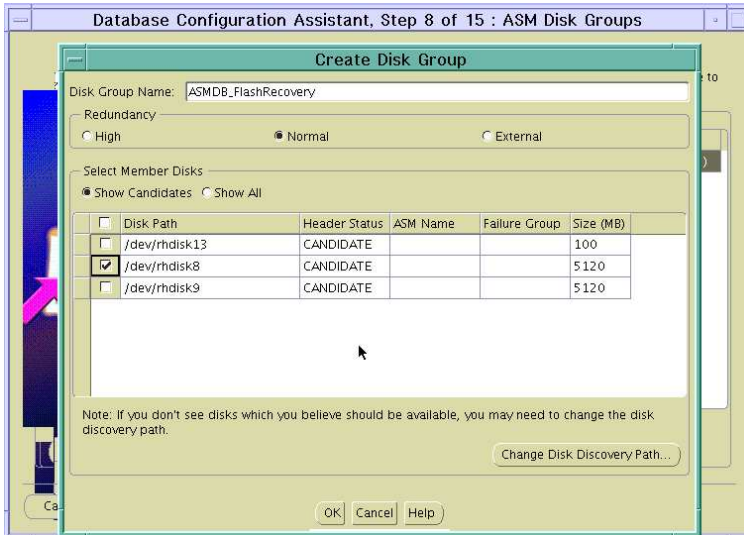


Create and Select Disk Group, or Select existing Disk Group :

Now the ASM Disks Group is created

Select your DiskGroup to be used to store the database to create.

Then click Ok ...

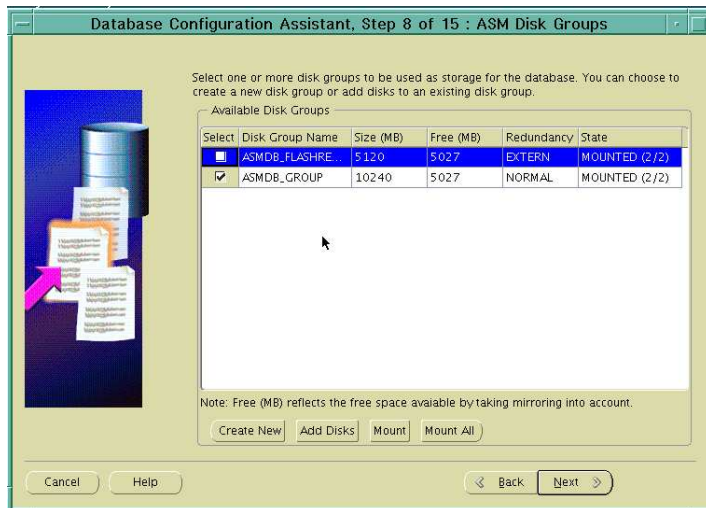


Create Disk Group

Specify ASM DiskGroup Name

Select disks from the candidates disks for the Flash Recovery Area Disk Group

Then click Ok ...

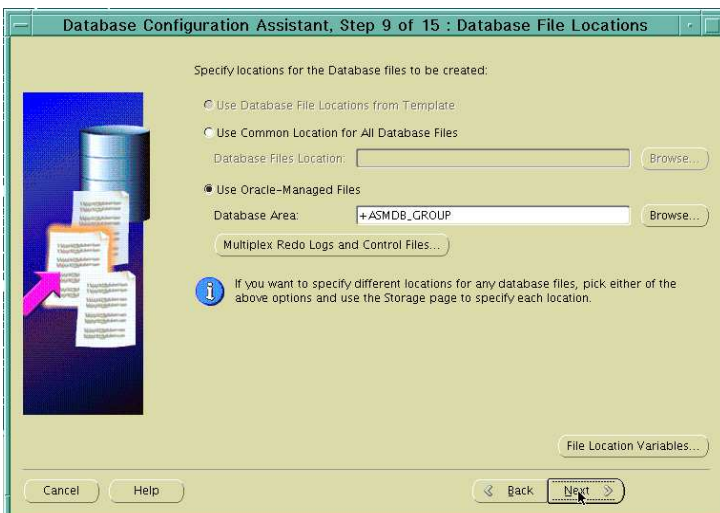


Create Disk Group

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 ...

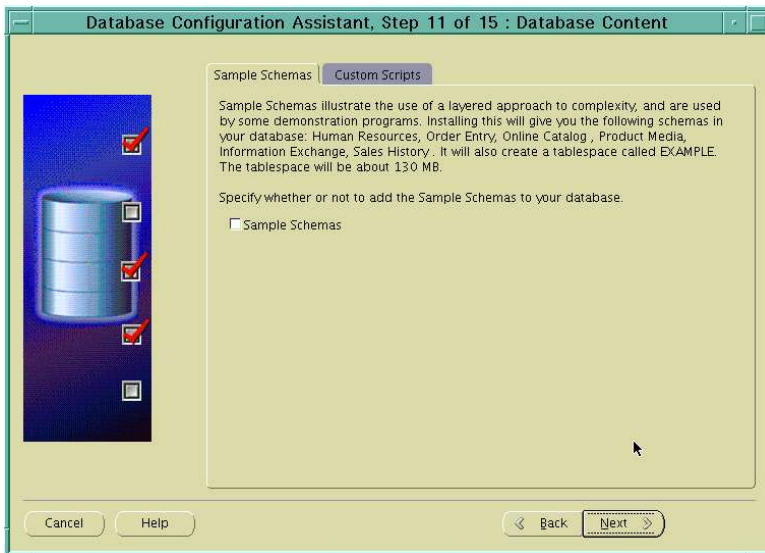


Recovery Configuration :

Select Use Oracle-Managed Files

AND Select DiskGroup to use for the Flash Recovery Area.

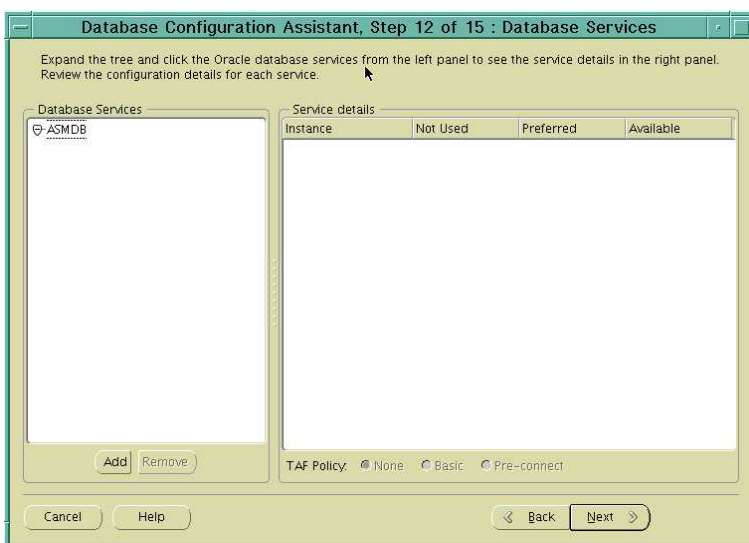
Then click Next ...



Database Content :

Select the options needed

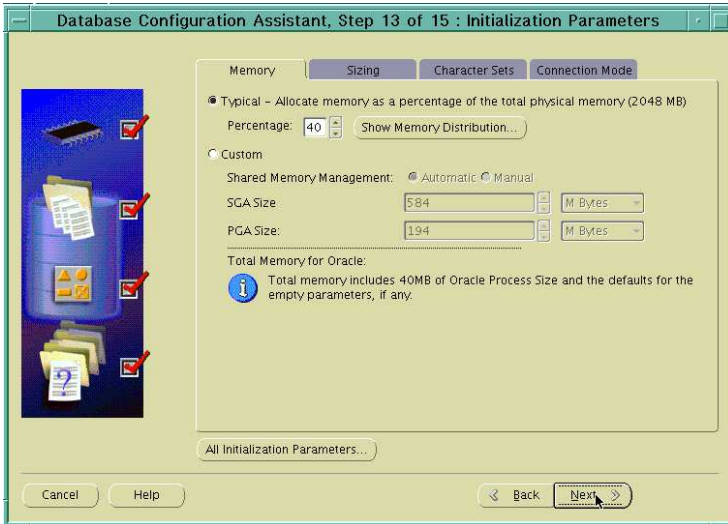
Then click Next ...



Database Services and TAF (Transaction Application Failover) policy :

Configure if necessary !!!

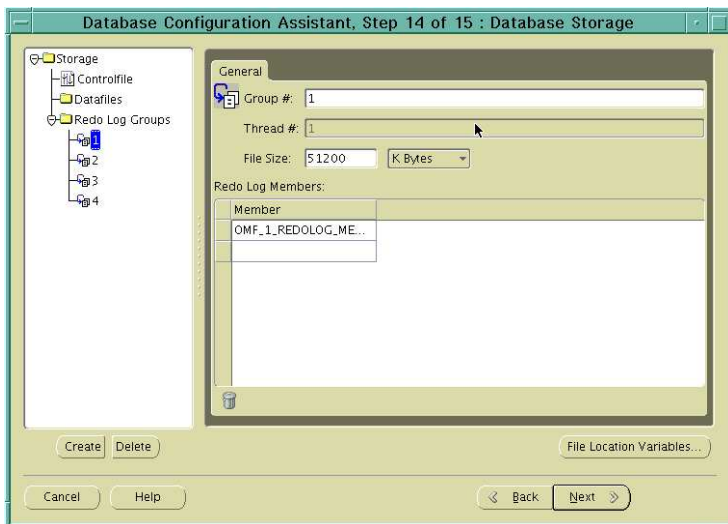
Then click Next ...



Initialization Parameters :

Select the parameters needed

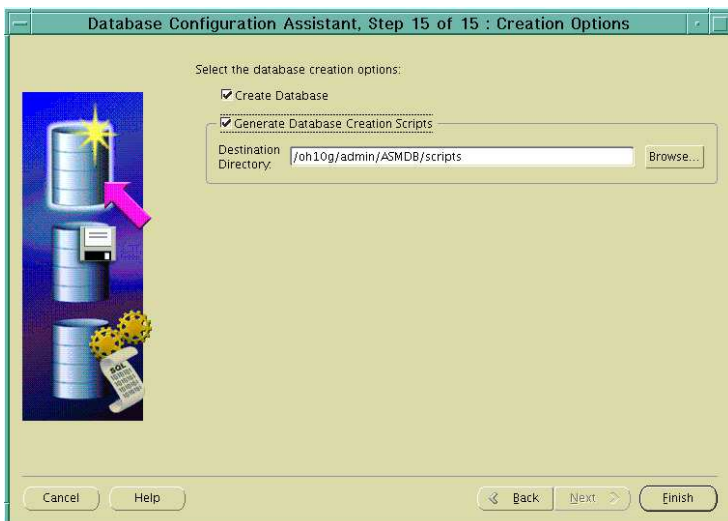
Then click Next ...



Database Storage :

Check the datafiles organization

Then click Next ...

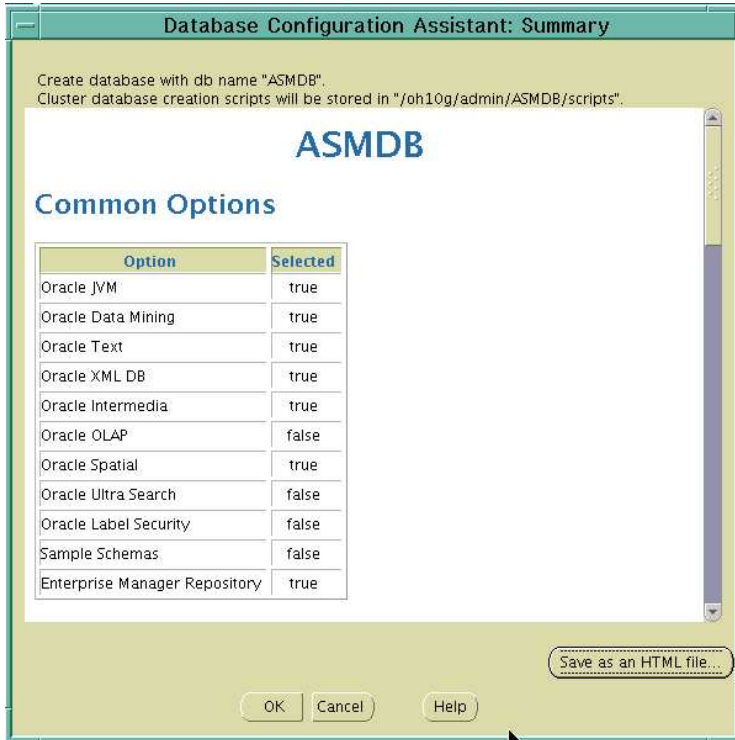


Creation Options :

Select the options needed

- Create Database
- Generate Database Creation Scripts

Then click Finish ...

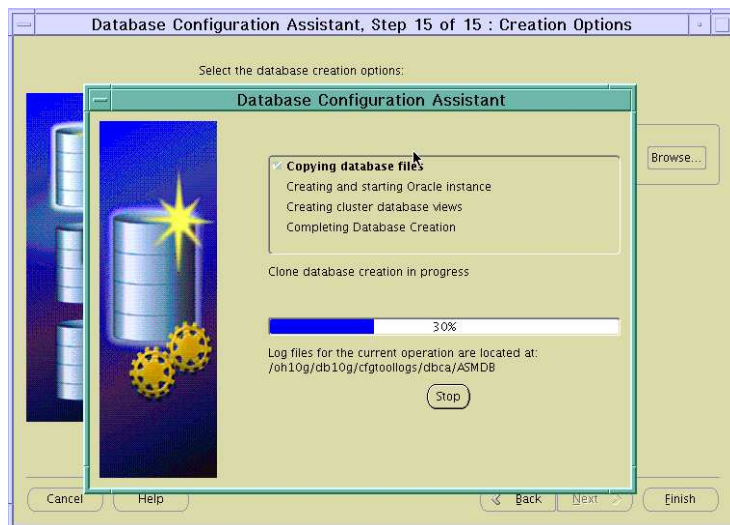


Summary :

Check the description

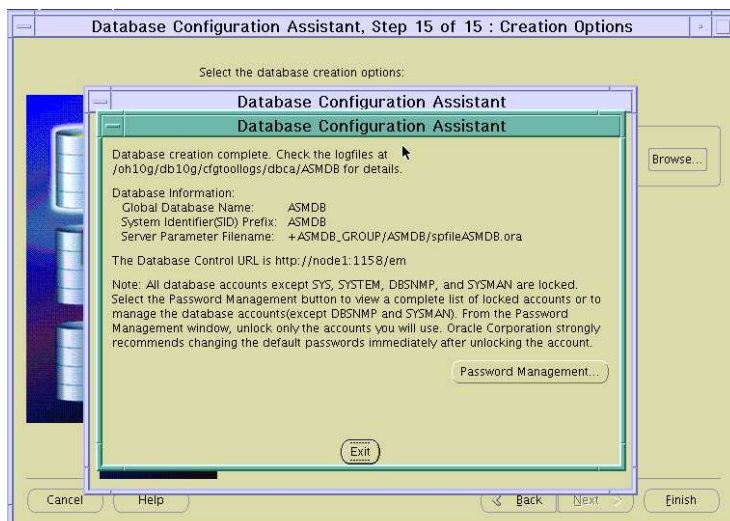
Save the HTML summary file if needed

Then click Ok ...



Database creation on progress :

Just wait while processing ...



Passwords Management

Enter in password management, if you need to change password, and unlock some user accounts that are locked by default (for security purpose).

Then click Exit ...

16.1.1 Database Post-Configuration task

16.1.1.1 Cluster Ready Services Health Check

Execute

crs_stat -t

**on one node
as oracle user :**

```
{node1:oracle}/oh10g ->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}/oh10g ->
```

CR health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

```
{node2:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

**cssd, crsd, evmhealth
check**

```
{node1:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

```
{node2:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

16.1.1.2 Oracle CLuster Registry content Check and Backup

Check Oracle Cluster Registry Integrity

As oracle user, Execute ocrcheck

```
{node1:oracle}/oh10g/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 not configured
```

Cluster registry integrity check succeeded

```
{node1:oracle}/oh10g/crs/bin ->
```

AS root user :

Export Oracle Cluster Registry content

```
{node1:oracle}/oh10g/crs/bin ->su
root's Password:
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export4.dmp -s online
{node1:oracle}/oh10g/crs/bin ->ls -la /oh10g/*.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp
{node1:oracle}/oh10g/crs/bin ->
```

➔ you can/must not edit/modify this exported file

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oh10g/crs/bin ->ocrconfig -showbackup
node1 2006/01/30 16:03:03 /oh10g/crs/cdata/crs
node1 2006/01/30 12:03:00 /oh10g/crs/cdata/crs
node1 2006/01/30 08:02:59 /oh10g/crs/cdata/crs
node1 2006/01/29 00:02:51 /oh10g/crs/cdata/crs
node1 2006/01/25 13:02:10 /oh10g/crs/cdata/crs
{node1:oracle}/oh10g/crs/bin ->
```


16.1.1.3 Some usefull commands

AS oracle user

Command to start/stop the Database, and Databases Instances :

```

From any nodes
srvctl start database -d ASMDB      to start the Database instance
srvctl stop database -d ASMDB       to stop the Database instance

For instance 1
srvctl start instance -d ASMDB -i ASMDB1  to start the DB.. instance
srvctl stop instance -d ASMDB -i ASMDB1  to stop the DB.. instance

For instance 2
srvctl start instance -d ASMDB -i ASMDB2  to start the DB.. instance
srvctl stop instance -d ASMDB -i ASMDB2  to stop the DB.. instance
    
```

To acces an ASM instance with sqlplus

```

From node1

export ORACLE_SID=+ASM1
sqlplus /nolog
connect /as sysdba
show sga
.....
    
```

To acces a Database instance stored in ASM with sqlplus

```

From node1

export ORACLE_SID=ASMDB1
sqlplus /nolog
connect /as sysdba
show sga
.....
    
```

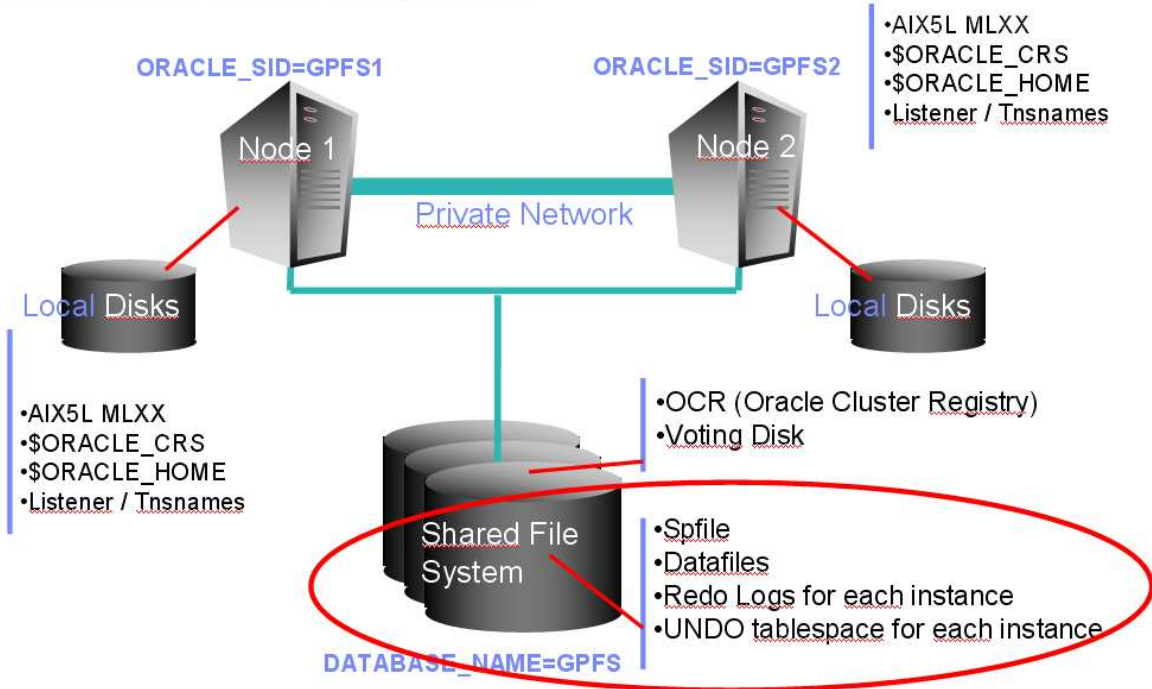
16.1.1.4 What has been done ?

At this stage :

- The Oracle Cluster Database (ASMDB) is created on Shared ASM Disks Group with
 - 1 database,
 - 2 databases instances (1 on each nodes,ASMDB1 and ASMDB2),

16.2 DATABASE CREATION ON GPFS

Architecture GPFS to implement



Connect as Oracle unix user from first node, and **setup your DISPLAY**
Execute **dbca &** to launch the database configuration assistant

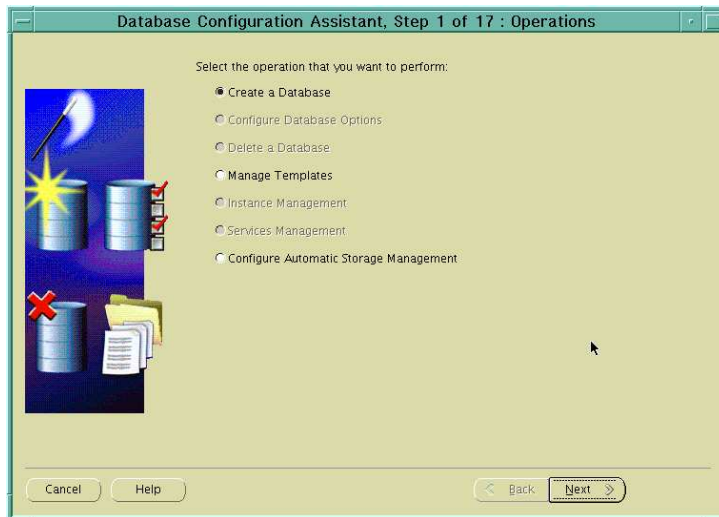
```
./dbca
```



DBCA Welcome Screen :

Select the **“Oracle Real Application Cluster Database”** option.

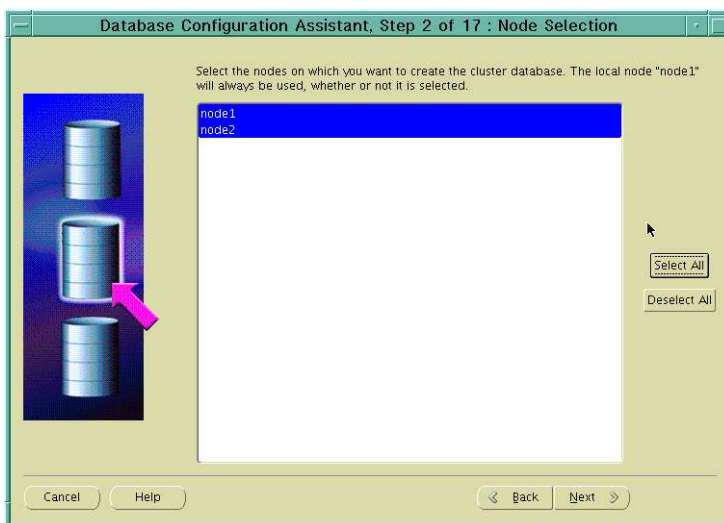
Then click **Next ...**



Operations :

Select the **“Create a Database”** option.

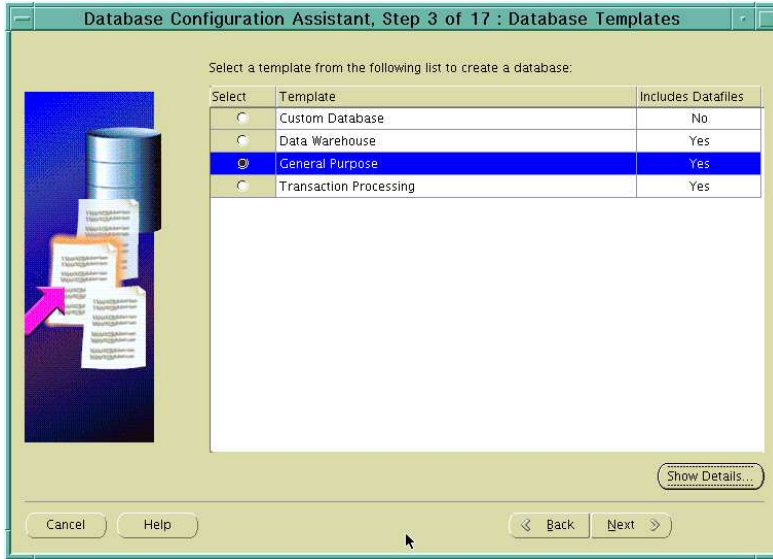
Then click **Next ...**



Node Selection :

Make sure to select all RAC nodes.

Then click **Next ...**

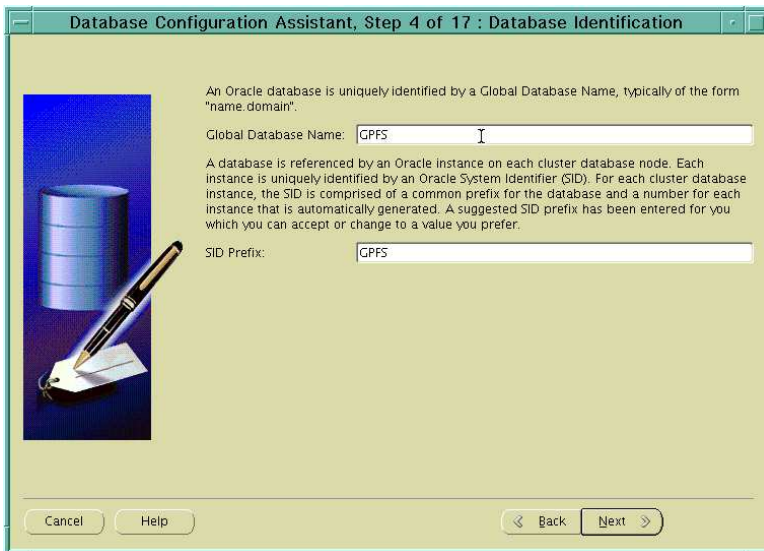


Database Templates :

Select **“General Purpose”**

Or “Custom Database” if you want to generate the creation scripts.

Then click Next ...

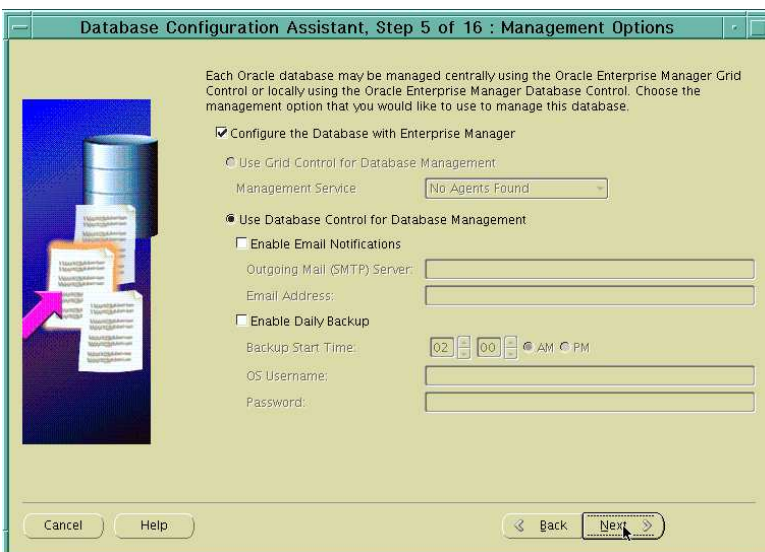


Database Identification :

Specify the **“Global Database Name”**

The “SID Prefix” will be automatically updated. (by default it is the Global Database Name)

Then click Next ...

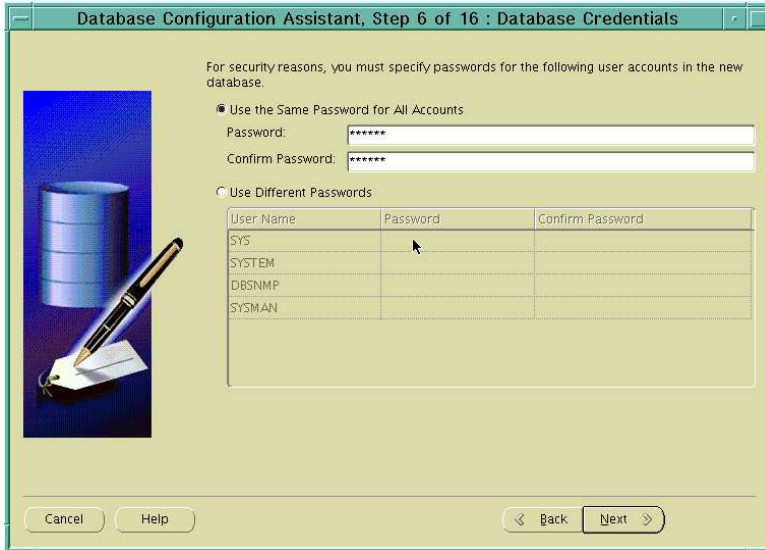


Management Options :

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)

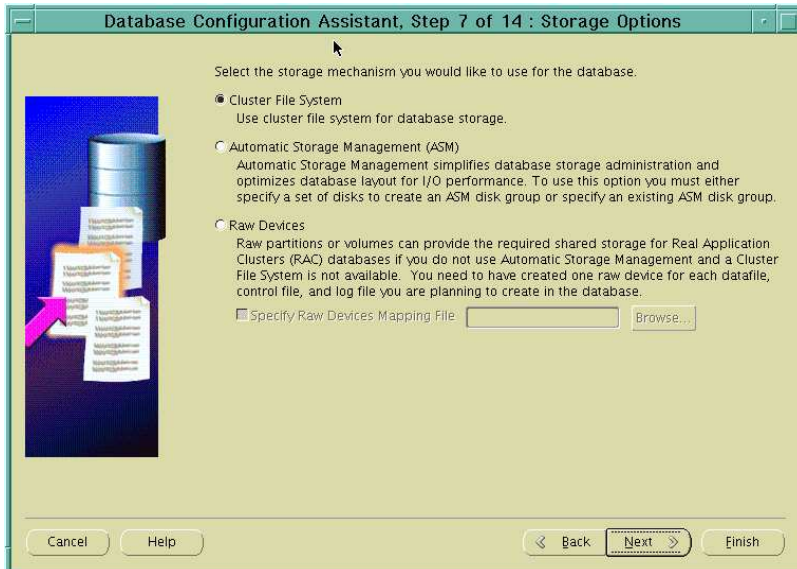
Then click Next ...



Database Credentials :

Specify same password for all administrator users, or specify individual password for each user.

Then click Next ...



Storage Options :

Choose Cluster File System

Then click Next ...

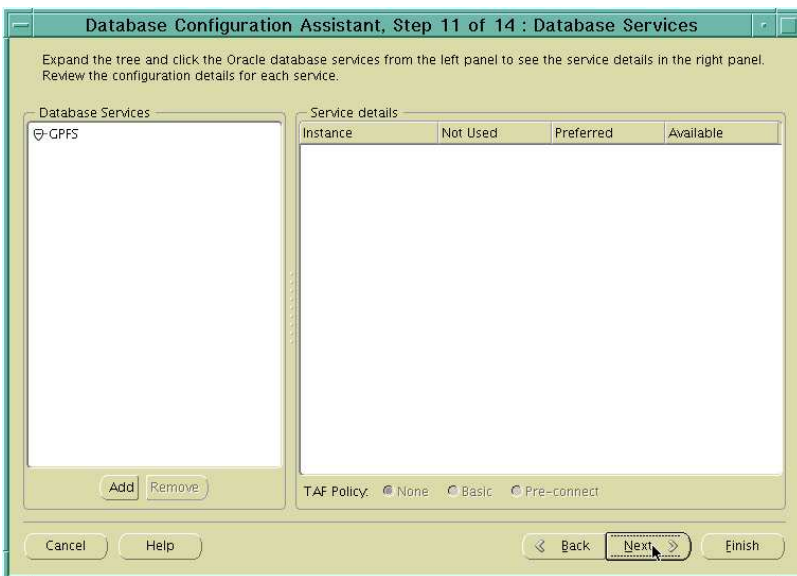
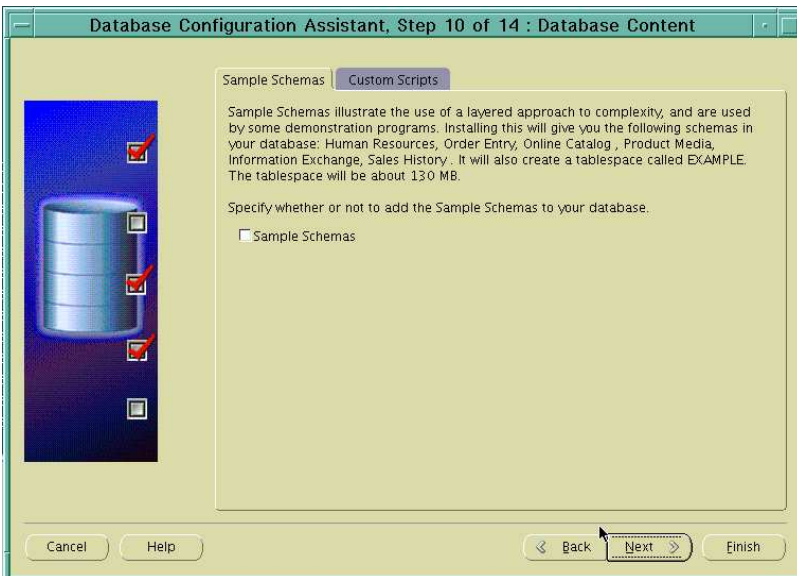
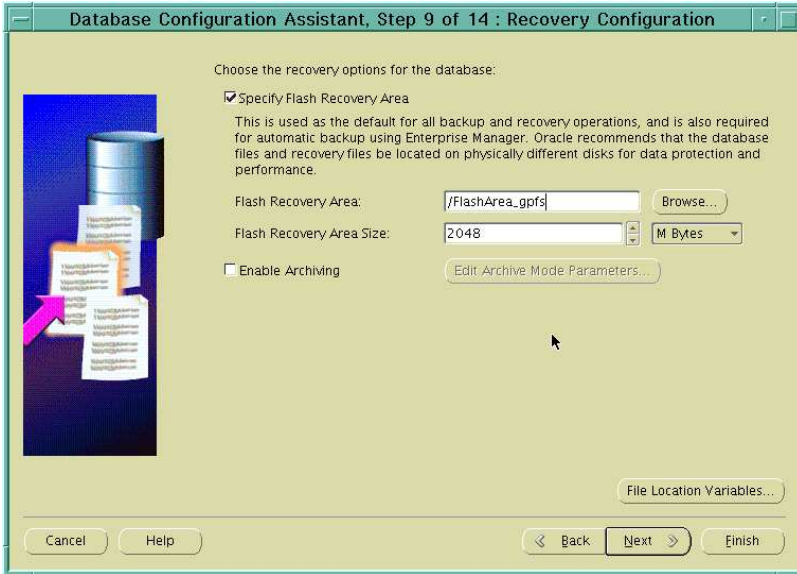


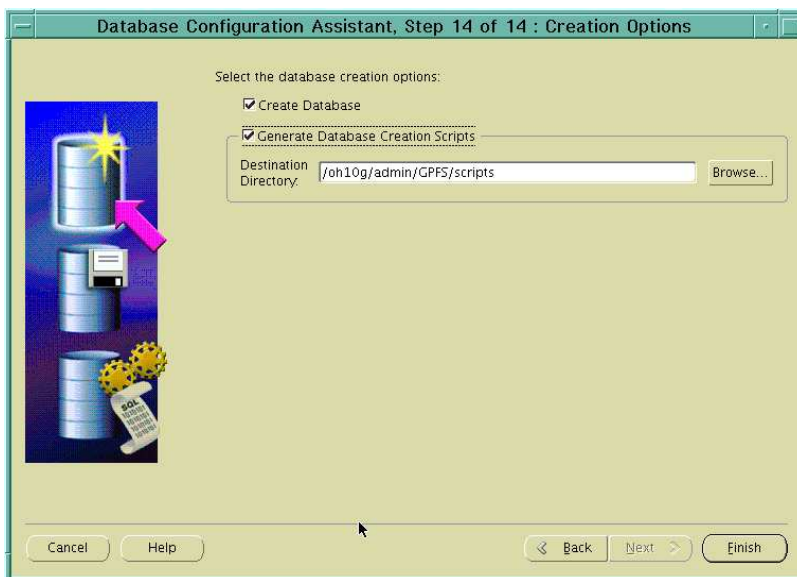
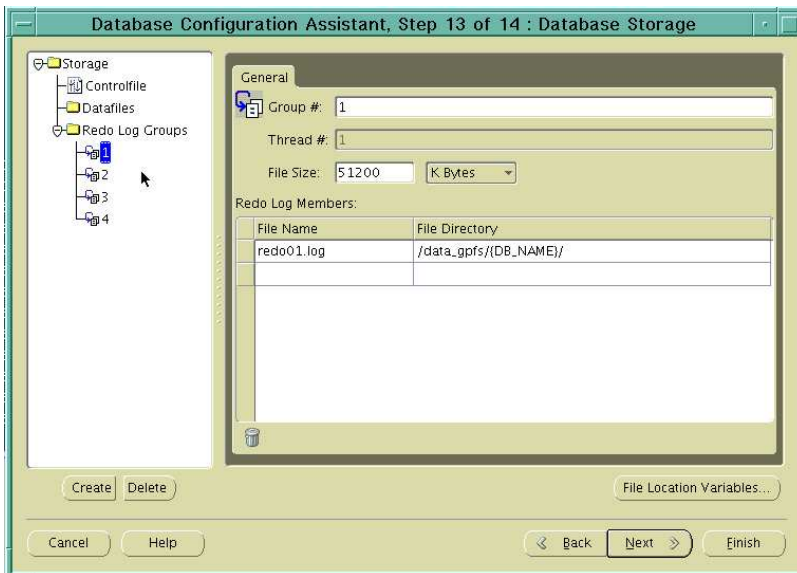
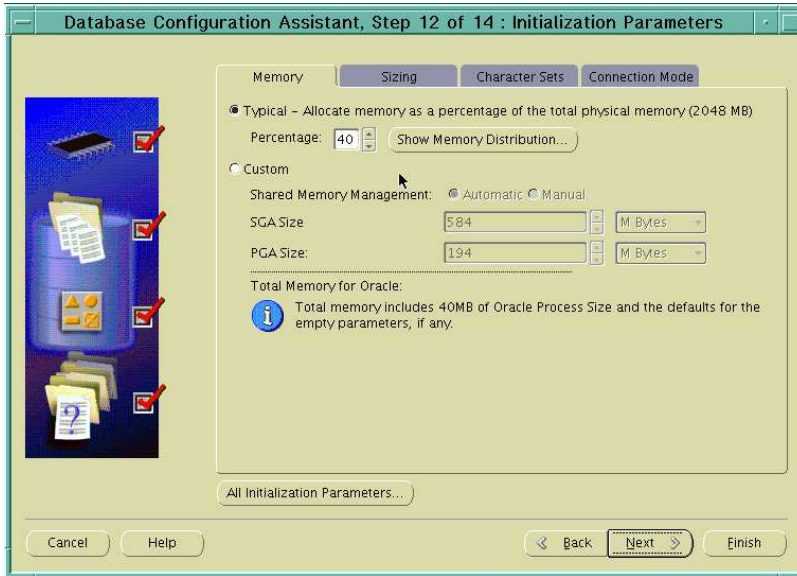
Database File Locations :

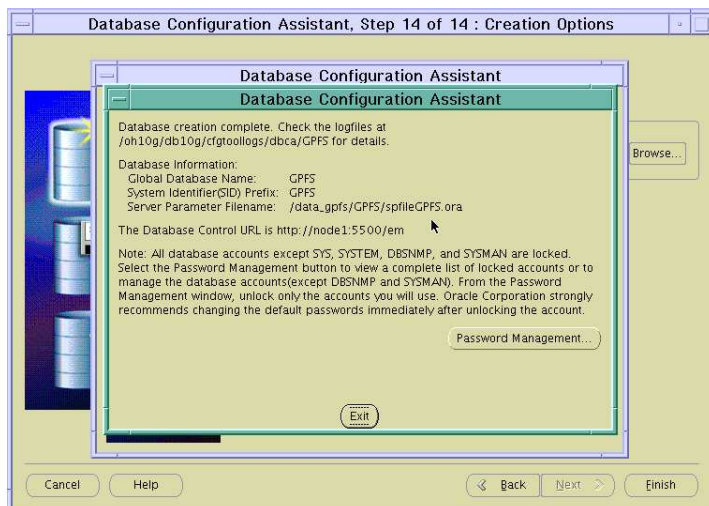
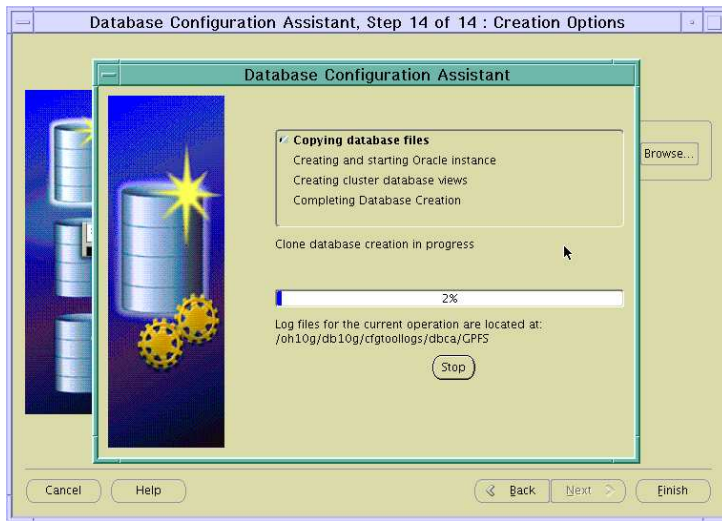
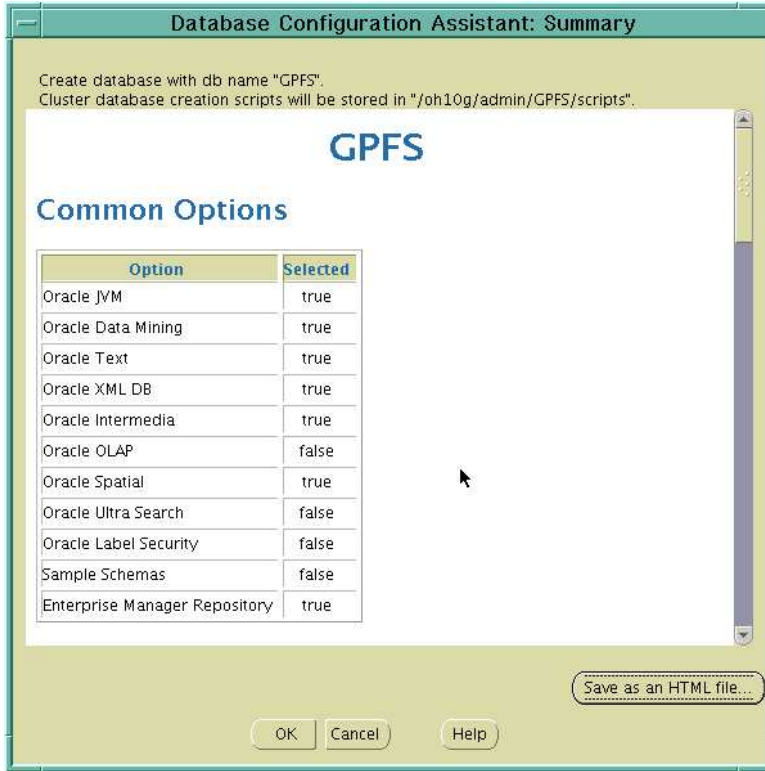
Select "Use Oracle-Managed Files", or "Use Common Location for All Databases Files"

AND Select DiskGroup to use for the Database Files.

Then click Next ...







16.2.1 Database Post-Configuration task

16.2.1.1 Cluster Ready Services Health Check

Execute

`crs_stat -t`

on one node
as oracle user :

```
{node1:oracle}/oh10g ->crs_stat -t
Name          Type          Target State  Host
-----
ora....S1.inst application ONLINE ONLINE node1
ora....S2.inst application ONLINE ONLINE node2
ora.GPFS.db   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....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}/oh10g ->
```

CR health check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

cssd, crsd, evmhealth
check

```
{node1:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node1:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy

{node2:oracle}/oh10g/crs/bin ->crsctl check cssd
CSS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check crsd
CRS appears healthy
{node2:oracle}/oh10g/crs/bin ->crsctl check evmd
EVM appears healthy
```

16.2.1.2 Oracle CLuster Registry content Check and Backup

Check Oracle Cluster Registry Integrity

**As oracle user,
Execute ocrcheck**

```
{node1:oracle}/oh10g/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 not configured

Cluster registry integrity check succeeded

{node1:oracle}/oh10g/crs/bin ->
```

AS root user :

Export Oracle Cluster Registry content

```
{node1:oracle}/oh10g/crs/bin ->su
root's Password:
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export5.dmp -s online
{node1:oracle}/oh10g/crs/bin ->ls -la /oh10g/*.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp
{node1:oracle}/oh10g/crs/bin ->
```

➔ you can/must not edit/modify this exported file

View OCR automatic periodic backup managed by Oracle Clusterware

```
{node1:oracle}/oh10g/crs/bin ->ocrconfig -showbackup
node1 2006/01/30 16:03:03 /oh10g/crs/cdata/crs
node1 2006/01/30 12:03:00 /oh10g/crs/cdata/crs
node1 2006/01/30 08:02:59 /oh10g/crs/cdata/crs
node1 2006/01/29 00:02:51 /oh10g/crs/cdata/crs
node1 2006/01/25 13:02:10 /oh10g/crs/cdata/crs
{node1:oracle}/oh10g/crs/bin ->
```

16.2.1.3 Some usefull commands

AS oracle user
 Command to start/stop the Database, and Databases Instances :

```

From any nodes
srvctl start database -d GPFS           to start the Database instance
srvctl stop database -d GPFS          to stop the Database instance

For instance 1
srvctl start instance -d GPFS -i GPFS1  to start the DB.. instance
srvctl stop instance -d GPFS -i GPFS1  to stop the DB.. instance

For instance 2
srvctl start instance -d GPFS -i GPFS2  to start the DB.. instance
srvctl stop instance -d GPFS -i GPFS2  to stop the DB.. instance
  
```

To acces a Database instance stored on GPFS with sqlplus

```

From node1

export ORACLE_SID=GPFS1
sqlplus /nolog
connect /as sysdba
show sga
.....
  
```

16.2.1.4 What has been done ?

At this stage :

- The Oracle Cluster Database (GPFS) is created on Shared ASM Disks Group with
 - 1 database,
 - 2 databases instances (1 on each nodes, GPFS1 and GPFS2).

16.3 CHECKING DB CONSOLE :

With 10gRAC R2, dbconsole is only configured on first node at database creation.

Commands to stop/start DBCONSOLE :

START	→ emctl start dbconsole
STOP	→ emctl stop dbconsole
STATUS	→ emctl status dbconsole

FOR RAC Node1 DO

BEGIN

As oracle user on node 1 :

Export DISPLAY=??????

Export ORACLE_SID=GPFS1

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://node1:5500/em>
using **sys** as user, connected **as sysdba**, with its password

IF dbconsole started and reachable with <http://node1:5500/em>

THEN dbconsole is OK on Node 1

ELSE See metalink note for DBConsole troubleshooting !!!

ELSE 1/ access dbconsole using Internet browser : <http://node1:5500/em>
using **sys** as user, connected **as sysdba**, with its password

IF dbconsole is reachable with <http://node1:5500/em>

THEN dbconsole is OK on Node 1

ELSE See metalink note for DBConsole troubleshooting !!!

END

16.4 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.

16.4.1 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)

16.4.2 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**)).

16.4.3 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 db -p <db_name> -o <ORACLE_HOME path>
srvctl add instance -p <db_name> -i <SID1> -n <node1>
srvctl add instance -p <db_name> -i <SID2> -n <node1>
```

17 APPENDIX B : EXAMPLES OF CONFIGURATION FILES

17.1 NETWORK

This appendix provides examples of the configuration files that are mentioned in the document.

/etc/hosts (on all nodes)

```
# Public Network
10.3.25.81          node1
10.3.25.82          node2
10.3.25.83          node3

# Virtual IP address
10.3.25.181        node1_vip
10.3.25.182        node2_vip
10.3.25.183        node3_vip

# Interconnect RAC
10.10.25.81        node1_rac
10.10.25.81        node2_rac
10.10.25.81        node3_rac

# Interconnect GPFS
20.20.25.81        node1_gpfs
20.20.25.81        node2_gpfs
20.20.25.81        node3_gpfs
```

/etc/hosts.equiv (on all nodes)

```
node1              root
node2              root
node3              root
node1_gpfs         root
node2_gpfs         root
node3_gpfs         root
node1              oracle
node2              oracle
node3              oracle
```

.rhosts : In the root's and oracle's home directory, put the list of machines.
\$HOME/.rhosts

```
node1              root
node2              root
node3              root
node1_gpfs         root
node2_gpfs         root
node3_gpfs         root
node1              oracle
node2              oracle
node3              oracle
```

17.2 LISTENER.ORA AND TNSNAMES.ORA CONFIGURATION EXAMPLE

17.2.1 Listener.ora Files

`$ORACLE_HOME/network/admin/listener.ora`

On node1 : # listener.ora.node1 Network Configuration File: /oh10g/db10g/network/admin/listener.ora.node1
Generated by Oracle configuration tools.

```

LISTENER_NODE1 =
(DESCRIPTION_LIST =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = node1_vip)(PORT = 1521)(IP = FIRST))
    )
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = 10.3.25.81)(PORT = 1521)(IP = FIRST))
    )
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC))
    )
  )
)

```

`$ORACLE_HOME/network/admin/listener.ora`

On node2 : # listener.ora.node2 Network Configuration File: /oh10g/db10g/network/admin/listener.ora.node2
Generated by Oracle configuration tools.

```

LISTENER_NODE2 =
(DESCRIPTION_LIST =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = node2_vip)(PORT = 1521)(IP = FIRST))
    )
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = 10.3.25.81)(PORT = 1521)(IP = FIRST))
    )
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC))
    )
  )
)

```


17.2.2 Tnsnames.ora Files (Server Side) for ASMDB Database

`$ORACLE_HOME/network/admin/tnsnames.ora` implementing TAF

On all RAC nodes, for ASMDB Database :

```
# tnsnames.ora Network Configuration File: /oh10g/db10g/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))
  )

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)
    )
  )
```

17.2.3 Tnsnames.ora Files (Server Side) for GPFS Database

`$ORACLE_HOME/network/admin/tnsnames.ora` implementing TAF

On all RAC
nodes, for
GPFS
Database :

```
# tnsnames.ora Network Configuration File: /oh10g/db10g/network/admin/tnsnames.ora
# Generated by Oracle configuration tools.

GPFS2 =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2_vip)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = GPFS)
      (INSTANCE_NAME = GPFS2)
    )
  )

GPFS1 =
  (DESCRIPTION =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1_vip)(PORT = 1521))
    (CONNECT_DATA =
      (SERVER = DEDICATED)
      (SERVICE_NAME = GPFS)
      (INSTANCE_NAME = GPFS1)
    )
  )

GPFS =
  (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 = GPFS)
    )
  )

LISTENERS_GPFS =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1_vip)(PORT = 1521))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node2_vip)(PORT = 1521))
  )
```

17.2.4 Tnsnames.ora Files (Client Side) for ASMDB Database

[\\$ORACLE_HOME/network/admin/tnsnames.ora](#) implementing TAF

On client side for ASMDB Database :

ASMDB
Connect String for :
→no failover
→load balancing

ASMDB_FAILOVER
Connect String for :
→failover
→load balancing

```
# tnsnames.ora Network Configuration File: /oh10g/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 =
  (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)
      )
    )
  )
)
```

17.2.5 Tnsnames.ora Files (Client Side) for GPFS Database

[\\$ORACLE_HOME/network/admin/tnsnames.ora](#) implementing TAF

On client side for ASMDB Database :

ASMDB
Connect String for :
→no failover
→load balancing

ASMDB_FAILOVER
Connect String for :
→failover
→load balancing

```
# tnsnames.ora Network Configuration File: /oh10g/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 =
  (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)
      )
    )
  )
)
```

18 APPENDIX C : ORACLE TECHNICAL NOTES

This appendix provides some useful notes coming from Oracle support. These notes can be found in Metalink.

18.1 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 engineers.

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) InterfaceFor 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 clusteware
- 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 occurred.

`$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.racl.gsd
TYPE=application
TARGET=ONLINE
STATE=ONLINE
```

```
NAME=ora.racl.oem
TYPE=application
TARGET=ONLINE
STATE=ONLINE
```

```
NAME=ora.racl.ons
TYPE=application
TARGET=ONLINE
STATE=ONLINE
```

```
NAME=ora.racl.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_CRN_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_CRN_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:

```
[opcbso11]/u01/home/usupport> ./crsstat
HA Resource
-----
ora.V10SN.V10SN1.inst      ONLINE      ONLINE on opcbso11
ora.V10SN.V10SN2.inst      ONLINE      ONLINE on opcbso12
ora.V10SN.db               ONLINE      ONLINE on opcbso12
ora.opcbso11.ASM1.asm       ONLINE      ONLINE on opcbso11
ora.opcbso11.LISTENER_OPCBSOL1.lsnr  ONLINE      ONLINE on opcbso11
ora.opcbso11.gsd           ONLINE      ONLINE on opcbso11
ora.opcbso11.ons           ONLINE      ONLINE on opcbso11
ora.opcbso11.vip           ONLINE      ONLINE on opcbso11
ora.opcbso12.ASM2.asm       ONLINE      ONLINE on opcbso12
ora.opcbso12.LISTENER_OPCBSOL2.lsnr  ONLINE      ONLINE on opcbso12
ora.opcbso12.gsd           ONLINE      ONLINE on opcbso12
ora.opcbso12.ons           ONLINE      ONLINE on opcbso12
ora.opcbso12.vip           ONLINE      ONLINE on opcbso12
```

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 interfacel[ |
        host interface2 |..]>][,] [<logical host name>]/<VIP address>/<net mask>
        [/<host interfacel[ | 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

18.2 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...](#)

18.3 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](#)

18.4 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](#)

18.5 ABOUT MANUAL DATABASE CRATION ...

240052.1 - [10g Manual Database Creation in Oracle \(Single Instance and RAC\)](#)

18.6 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](#)

18.7 ABOUT TAF ...

271297.1 - [Troubleshooting TAF Issues in 10g RAC](#)

18.8 ABOUT ADDING/REMOVING NODE ...

269320.1 - [Removing a Node from a 10g RAC Cluster](#)

270512.1 - [Adding a Node to a 10g RAC Cluster](#)

18.9 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](#)

18.10 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](#)

19 APPENDIX F : FILESETS TO BE INSTALLED ON THE MACHINES OF THE CLUSTER.

This appendix provides the list of filesets which must be installed depending on the implementation. Chosen.

19.1 RSCT 2.3.4 (PROVIDED WITH AIX 5.3) FOR ALL IMPLEMENTATION

rsct.basic.rte	2.4.3.0	C	F	RSCT Basic Function
rsct.compat.basic.rte	2.4.3.0	C	F	RSCT Event Management Basic
rsct.compat.clients.rte	2.4.3.0	C	F	RSCT Event Management Client
rsct.core.auditrm	2.4.3.0	C	F	RSCT Audit Log Resource
rsct.core.errm	2.4.3.0	C	F	RSCT Event Response Resource
rsct.core.fsrn	2.4.3.0	C	F	RSCT File System Resource
rsct.core.gui	2.4.3.0	C	F	RSCT Graphical User Interface
rsct.core.hostrm	2.4.3.0	C	F	RSCT Host Resource Manager
rsct.core.lprm	2.4.3.0	C	F	RSCT Least Privilege Resource
rsct.core.rmc	2.4.3.0	C	F	RSCT Resource Monitoring and
rsct.core.sec	2.4.3.0	C	F	RSCT Security
rsct.core.sensorm	2.4.3.0	C	F	RSCT Sensor Resource Manager
rsct.core.sr	2.4.3.0	C	F	RSCT Registry
rsct.core.utils	2.4.3.0	C	F	RSCT Utilities
rsct.lapi.msg.EN_US.rte	2.4.0.0	C	F	RSCT LAPI Msgs - U.S. English
rsct.lapi.msg.en_US.rte	2.4.0.0	C	F	RSCT LAPI Msgs - U.S. English
rsct.lapi.nam	2.4.2.0	C	F	RSCT LAPI NAM Files
rsct.lapi.rte	2.4.2.0	C	F	RSCT LAPI Runtime
rsct.lapi.samp	2.4.2.0	C	F	RSCT LAPI Samples
rsct.msg.EN_US.core.auditrm				
rsct.msg.EN_US.core.errm	2.4.0.0	C	F	RSCT Event Response RM Msgs -
rsct.msg.EN_US.core.fsrn	2.4.0.0	C	F	RSCT File System RM Msgs -
rsct.msg.EN_US.core.gui	2.4.0.0	C	F	RSCT GUI Msgs - U.S. English
rsct.msg.EN_US.core.hostrm				
rsct.msg.EN_US.core.lprm	2.4.0.0	C	F	RSCT LPRM Msgs - U.S. English
rsct.msg.EN_US.core.rmc	2.4.0.0	C	F	RSCT RMC Msgs - U.S. English
rsct.msg.EN_US.core.sec	2.4.0.0	C	F	RSCT Security Msgs - U.S.
rsct.msg.EN_US.core.sensorm				
rsct.msg.EN_US.core.sr	2.4.0.0	C	F	RSCT Registry Msgs - U.S.
rsct.msg.EN_US.core.utils	2.4.0.0	C	F	RSCT Utilities Msgs - U.S.
rsct.msg.en_US.core.auditrm				
rsct.msg.en_US.core.errm	2.4.0.0	C	F	RSCT Event Response RM Msgs -
rsct.msg.en_US.core.fsrn	2.4.0.0	C	F	RSCT File System RM Msgs -
rsct.msg.en_US.core.gui	2.4.0.0	C	F	RSCT GUI Msgs - U.S. English
rsct.msg.en_US.core.gui.com				
rsct.msg.en_US.core.hostrm				
rsct.msg.en_US.core.lprm	2.4.0.0	C	F	RSCT LPRM Msgs - U.S. English

rsct.msg.en_US.core.rmc	2.4.0.0	C	F	RSCT RMC Msgs - U.S. English
rsct.msg.en_US.core.rmc.com				
rsct.msg.en_US.core.sec	2.4.0.0	C	F	RSCT Security Msgs - U.S.
rsct.msg.en_US.core.sensorrm				
rsct.msg.en_US.core.sr	2.4.0.0	C	F	RSCT Registry Msgs - U.S.
rsct.msg.en_US.core.utils	2.4.0.0	C	F	RSCT Utilities Msgs - U.S.

19.2 FOR GPFS IMPLEMENTATION

gpfs.base	2.3.0.5	C	F	GPFS File Manager
gpfs.docs.data	2.3.0.5	C	F	GPFS Server Manpages and
gpfs.msg.en_US	2.3.0.4	C	F	GPFS Server Messages - U.S.