Quick Installation Guide

Oracle 10g RAC Release 2 on IBM 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 :

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Update</th>
<th>Who</th>
<th>Validated by</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>January 2006</td>
<td>- Creation</td>
<td>Frederic Michiara</td>
<td>Fabienne Lepetit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Thierry Plumeau</td>
<td>Paul Bramy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Alain Roy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Marc Tu-Duy-Kiem</td>
</tr>
</tbody>
</table>

Contributors :

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

Contact :

- 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 Recommendations, hints and tips ................................................................. 61
8.2 GPFS implementation ....................................................................................... 63
  8.2.1 Network Infrastructure for GPFS implementation, wih 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 CU to Determine if Installation Prerequisites are Complete ............... 75
12 Install the Cluster 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 usefulness 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 usefulness 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 usefulness 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 usefulness 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
16.4.3 Configure Oracle Enterprise Manager

17 Appendix B: examples of configuration files
17.1 Network
17.2 listener.ora and tnsnames.ora Configuration example
17.2.1 Listener.ora Files
17.2.2 Tnsnames.ora Files (Server Side) for ASMB Database
17.2.3 Tnsnames.ora Files (Server Side) for GPFS Database
17.2.4 Tnsnames.ora Files (Client Side) for ASMB Database
17.2.5 Tnsnames.ora Files (Client Side) for GPFS Database

18 Appendix C: Oracle technical notes
18.1 CRS and 10g Real Application Clusters
18.2 About RAC
18.3 About CRS
18.4 About VIP
18.5 About manual database creation
18.6 About Grid Control
18.7 About TAF
18.8 About Adding/Removing Node
18.9 About ASM
18.10 Metalink note to use in case of problem with CRS

19 Appendix F: Filesets to be installed on the machines of the cluster
19.1 RSCT 2.3.4 (provided with AIX 5.3) for all implementation
19.2 For GPFS implementation
1 THE AIM OF THIS DOCUMENT

This document is written to provide help installing Oracle 10g 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)

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

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 types 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)

<table>
<thead>
<tr>
<th>Eth0 (Public Network)</th>
<th>RAC Node 1</th>
<th>RAC Node 2</th>
<th>RAC Node 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eth1 (Private Network) RAC Interconnect</td>
<td>CRS</td>
<td>CRS</td>
<td>CRS</td>
</tr>
<tr>
<td></td>
<td>RAC</td>
<td>RAC</td>
<td>RAC</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>Listener</td>
<td>Listener</td>
</tr>
<tr>
<td></td>
<td>ASM Inst</td>
<td>ASM Inst</td>
<td>ASM Inst</td>
</tr>
<tr>
<td></td>
<td>DB Inst</td>
<td>DB Inst</td>
<td>DB Inst</td>
</tr>
</tbody>
</table>

Oracle Storage Management Solution:

- No need for HACMP
- No need for GPFS
- Only Oracle databases files (datafiles, redo logs, archive logs, flash recovery area, ...) are stored on the disks managed by Oracle ASM. No binaries.
- CRS files (OCR and Voting) are placed on raw disks.

3.1.2 IBM GPFS (General Parallel File System)

<table>
<thead>
<tr>
<th>Eth0 (Public Network)</th>
<th>RAC Node 1</th>
<th>RAC Node 2</th>
<th>RAC Node 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eth1 (Private Network) RAC Interconnect Eth2 (Private Network) GPFS Interconnect</td>
<td>GPFS</td>
<td>GPFS</td>
<td>GPFS</td>
</tr>
<tr>
<td></td>
<td>CRS</td>
<td>CRS</td>
<td>CRS</td>
</tr>
<tr>
<td></td>
<td>RAC</td>
<td>RAC</td>
<td>RAC</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>Listener</td>
<td>Listener</td>
</tr>
<tr>
<td></td>
<td>DB Inst</td>
<td>DB Inst</td>
<td>DB Inst</td>
</tr>
</tbody>
</table>

IBM Cluster Files System Solution:

- No need for HACMP
- All files (Oracle database files, CRS files with OCR and Voting, any binaries, any files) can be stored on IBM GPFS.
### 3.1.3 Concurrent Raw Devices (HACMP)

**IBM HACMP Solution:**

- HACMP is ONLY mandatory to achieve concurrent access (read/write) to shared raw devices.
- Only Oracle databases files (datafiles, redo logs, spfiles, CRS files with OCR and Voting) are stored on the concurrent raw devices. No archive logs, No binaries.

### 3.1.4 NAS (Network Appliance Filer)

**Network Appliance Solution:**

- No need for HACMP
- All files (Oracle database files, CRS files with OCR and Voting, any binaries, any files) can be stored on NAS filer.
- Specific Cookbook for Net App NAS storage in progress.

### 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 below:

![Diagram of hardware architecture](image-url)
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
- 8GB / 21B
System memory:
- (Std/Max)
Internal storage:
- (Std/Max)
Performance (rPerf range)***:
- 72.8GB / 28.1TB (using optional I/O drawers)
- 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](http://metalink.oracle.com). This network can also be used for GPFSS.

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 bandwith 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 ➔

Oracle Storage Compatibility Program Member List ➔

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 allows SAN access failover, and load balancing across SAN Fiber Channel attachments.

### 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

**MPIO for AIX 5.3 download page:**


**Installing the filesets:**

- `smitty install`
- `Install and Update Software`
- `Install Software`
  
  - * INPUT device / directory for software
  - `[/mydir_with_my_filesets]`
  - `SOFTWARE to install`  
    
    - 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.

- `* 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 installation success and the installation summary message:

Install devices.sddpcm.53

Select:
- devices.sddpcm.53
- ALL

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 files sets? yes
  + DETAILED output? no
  + Process multiple volumes? yes
  + ACCEPT new license agreements? no
  + Preview new LICENSE agreements? no

Installation Summary

<table>
<thead>
<tr>
<th>Name</th>
<th>Level</th>
<th>Part</th>
<th>Event</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>devices.fcp.disk.ibm.mpio.r</td>
<td>1.0.0.0</td>
<td>USR</td>
<td>APPLY</td>
<td>SUCCESS</td>
</tr>
</tbody>
</table>

Now you need to reboot all AIX nodes !!!
Installing Oracle 10g RAC Release 2 on IBM pSeries with AIX 5L

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 access 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

RAC Installation Steps (Start)

- Hardware and Network Requirements (Chap.)
- Software Requirements (Chap.)
- Users and Groups on Cluster Nodes (Chap.)
- Configure Kernel Parameters and Shell Limits on Cluster Nodes (Chap.)
- Configure Network Parameters on Cluster Nodes (Chap.)
- Configure Oracle environment on Cluster Nodes (Chap.)
- Synchronize the System Time on Cluster Nodes (Chap.)

Choose Storage option (Chap.)

- ASM (Chap.)
- GPFS (Chap.)
- Raw Devices (Chap.)

Create LUN’s or attached Disks for OCR and voting Disks (Chap.)

- Configure Disks For OCR/Voting (Chap.)
- Install and configure GPFS (Chap.)
- Install and configure HACMP (Chap.)

Create Raws Devices for OCR/Voting disks (Chap.)

Cluster Ready Services Installation (Chap.)

Oracle 10g software Installation with RAC option & VIP configuration (Chap.)

Database Creation (Chap.)

RAC Installation Steps (END)
### 5.2 Database Creation Steps

<table>
<thead>
<tr>
<th>Database Creation (Start)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose Storage option (Chap.)</td>
</tr>
<tr>
<td>ASM (Chap.)</td>
</tr>
<tr>
<td>GPFS (Chap.)</td>
</tr>
<tr>
<td>Raw Devices (Chap.)</td>
</tr>
<tr>
<td>Create LUN’s or attached Disks for Databases (Chap.)</td>
</tr>
<tr>
<td>Configure ASM Disks For database files (Chap.)</td>
</tr>
<tr>
<td>Configure GPFS for database files (Chap.)</td>
</tr>
<tr>
<td>Configure Raw devices for database files (Chap.)</td>
</tr>
<tr>
<td>Set Oracle:dba as ASM Disks/LUN’s owner, and necessary permissions</td>
</tr>
<tr>
<td>Set Oracle:dba as GPFS Disks/LUN’s owner, and necessary permissions</td>
</tr>
<tr>
<td>Set Oracle:dba as Raw Devices owner, and necessary permissions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Choose Database Creation Method (Chap.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual Database Creation</td>
</tr>
<tr>
<td>Database Creation using DBCA assistant</td>
</tr>
<tr>
<td>Run DBCA as Oracle user</td>
</tr>
<tr>
<td>Select all RAC nodes</td>
</tr>
<tr>
<td>Register instances and database in Oracle Cluster Registry (Using srvctl)</td>
</tr>
<tr>
<td>Select Storage option:</td>
</tr>
<tr>
<td>• ASM</td>
</tr>
<tr>
<td>• Raw Devices</td>
</tr>
<tr>
<td>• Cluster File System</td>
</tr>
<tr>
<td>Run NETCA as oracle user: To configure listener and tnsnames.ora</td>
</tr>
<tr>
<td>And answer necessary questions</td>
</tr>
</tbody>
</table>

Database Creation (END)
### 6 CHECK LIST TO USE AND FOLLOW

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

<table>
<thead>
<tr>
<th>Operations</th>
<th>Done on each node: Yes/No?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Node 1</td>
</tr>
<tr>
<td>1. Check the Hardware Requirements</td>
<td></td>
</tr>
<tr>
<td>2. Check the Network Requirements</td>
<td></td>
</tr>
<tr>
<td>3. Check the Software Requirements</td>
<td></td>
</tr>
<tr>
<td>4. Create Required UNIX Groups and User</td>
<td></td>
</tr>
<tr>
<td>5. Configure Kernel Parameters and Shell Limits</td>
<td></td>
</tr>
<tr>
<td>6. Identify Required Software Directories</td>
<td></td>
</tr>
<tr>
<td>7. Identify or Create an Oracle Base Directory</td>
<td></td>
</tr>
<tr>
<td>8. Create the CRS Home Directory</td>
<td></td>
</tr>
<tr>
<td>9. Choose a Storage Option for Oracle CRS, Database, and Recovery Files</td>
<td></td>
</tr>
<tr>
<td>11. Create Directories for Oracle CRS, Database, or Recovery Files</td>
<td></td>
</tr>
<tr>
<td>12. Configure Disks for Automatic Storage Management</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Or Configure Raw Devices And Verify the Cluster Software Configuration</td>
<td></td>
</tr>
<tr>
<td>Or Configure GPFS for Shared Files System</td>
<td></td>
</tr>
<tr>
<td>➔ And ACCESS Disk Storage concurrently from all nodes participating to the RAC cluster</td>
<td></td>
</tr>
<tr>
<td>13. Synchronize the System Time on Cluster Nodes</td>
<td></td>
</tr>
<tr>
<td>14. Stop Existing Oracle Processes</td>
<td></td>
</tr>
<tr>
<td>15. Configure the oracle User's Environment</td>
<td></td>
</tr>
</tbody>
</table>
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** $= 2 \times$ 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](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/](http://www-912.ibm.com/eserver/support/fixes/)

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

```
# lslpp -I 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 filessets.

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:

```
# /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/](http://www-912.ibm.com/eserver/support/fixes/)

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

```
# lslpp -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/](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.)

<table>
<thead>
<tr>
<th>AIX 5L version 5.2, Maintenance Level 04 or later</th>
<th>AIX 5L version 5.3, Maintenance Level 02 or later</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Filesets</strong></td>
<td><strong>Filesets</strong></td>
</tr>
<tr>
<td>• bos.adt.base</td>
<td>• bos.adt.base</td>
</tr>
<tr>
<td>• bos.adt.lib</td>
<td>• bos.adt.lib</td>
</tr>
<tr>
<td>• bos.adt.libm</td>
<td>• bos.adt.libm</td>
</tr>
<tr>
<td>• bos.perf.libperfstat</td>
<td>• bos.perf.libperfstat</td>
</tr>
<tr>
<td>• bos.perf.perfstat</td>
<td>• bos.perf.perfstat</td>
</tr>
<tr>
<td>• bos.perf.proctools</td>
<td>• bos.perf.proctools</td>
</tr>
<tr>
<td>• rsct.basic.rte</td>
<td>• rsct.basic.rte</td>
</tr>
<tr>
<td>• rsct.compat.clients.rte</td>
<td>• rsct.compat.clients.rte</td>
</tr>
<tr>
<td>• xIC.ai50.rte 7.0.0.4</td>
<td>• xIC.ai50.rte 7.0.0.4</td>
</tr>
<tr>
<td>• xIC.rte 7.0.0.1</td>
<td>• xIC.rte 7.0.0.1</td>
</tr>
<tr>
<td><strong>Specific Filesets</strong></td>
<td><strong>Specific Filesets</strong></td>
</tr>
<tr>
<td><strong>For EMC Symmetrix</strong></td>
<td><strong>For EMC Symmetrix</strong></td>
</tr>
<tr>
<td>• EMC.Symmetrix.aix.rte.5.2.0.1</td>
<td>• EMC.Symmetrix.aix.rte.5.2.0.1</td>
</tr>
<tr>
<td><strong>For EMC CLARiON</strong></td>
<td><strong>For EMC CLARiON</strong></td>
</tr>
<tr>
<td>• EMC.CLARiON.fcp.rte.5.2.0.1</td>
<td>• EMC.CLARiON.fcp.rte.5.2.0.1</td>
</tr>
</tbody>
</table>

**APAR’s**

**Authorized Problem Analysis Reports (APARs) for AIX 5L v5.2 ML 4:**
- IY63133: large percentage of CPU time spent in ldata_balance routine
- IY63366: disym returns null even for valid symbol in AIX520 ML-4
- IY64691: chv -b can cause corruption and crash
- IY64737: aio can hang in knotunlock
- IY65001: mklvcopy on a striped lv is failing to update lvcb
- IY64978: deadlock with concurrent renaming and unlinking under JFS
- IY65001
- IY75901, if IY69518 was previously installed.
- IY70029: “CORRUPTION FROM SIMULTANEOUS CIO WRITES WITH O_DSYNC ON JFS2” If using the IBM Journal File System Version 2 (JFS2) for Oracle Database files.
- **JDK:**
  - IY58350 Patch for SDK 1.3.1.16 (32-bit)
  - IY63533 Patch for SDK 1.4.2.1 (64-bit)
  - IY65305 Patch for SDK 1.4.2.2 (32-bit)

**APAR’s**

**Authorized Problem Analysis Reports (APARs) for AIX 5L v5.3 ML02:**
- IY68989: "WRITE TO MMAPPED SPACE HANGS"
- IY68874: An application that is run with mandatory large page data (LDR_CNTRL=LARGE_PAGE_DATA=M) may core-dump on the AIX 5.3 64-bit kernel in a LPAR environment.
- IY70031: “CORRUPTION FROM SIMULTANEOUS CIO WRITES WITH O_DSYNC ON JFS2” If using the IBM Journal File System Version 2 (JFS2) for Oracle Database files.
- **JDK:**
  - 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 Files / APAR’s Requirements for 10gRAC R2 / GPFS (NO HACMP)

AIX files & 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.)

<table>
<thead>
<tr>
<th>AIX 5L version 5.2, Maintenance Level 04 or later</th>
<th>AIX 5L version 5.3, Maintenance Level 02 or later</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Filessets</strong></td>
<td><strong>Filessets</strong></td>
</tr>
<tr>
<td>• bos.adt.base</td>
<td>• bos.adt.base</td>
</tr>
<tr>
<td>• bos.adt.lib</td>
<td>• bos.adt.lib</td>
</tr>
<tr>
<td>• bos.adt.libm</td>
<td>• bos.adt.libm</td>
</tr>
<tr>
<td>• bos.perf.libperfs tat</td>
<td>• bos.perf.libperfs tat</td>
</tr>
<tr>
<td>• bos.perf.perfs tat</td>
<td>• bos.perf.perfs tat</td>
</tr>
<tr>
<td>• bos.perf.proctools</td>
<td>• bos.perf.proctools</td>
</tr>
<tr>
<td>• rsct.basic.rte</td>
<td>• rsct.basic.rte</td>
</tr>
<tr>
<td>• rsct.compat.clients.rte</td>
<td>• rsct.compat.clients.rte</td>
</tr>
<tr>
<td>• xlC.aix50.rte 7.0.0.4</td>
<td>• xlC.aix50.rte 7.0.0.4</td>
</tr>
<tr>
<td>• xlC.rte 7.0.0.1</td>
<td>• xlC.rte 7.0.0.1</td>
</tr>
<tr>
<td><strong>Specific Filesets</strong></td>
<td><strong>Specific Filesets</strong></td>
</tr>
<tr>
<td>For EMC Symmetrix :</td>
<td>For EMC Symmetrix :</td>
</tr>
<tr>
<td>• EMC.Symmetrix.aix.rte.5.2.0.1</td>
<td>• EMC.Symmetrix.aix.rte.5.2.0.1</td>
</tr>
<tr>
<td>For EMC CLARiON :</td>
<td>For EMC CLARiON :</td>
</tr>
<tr>
<td>• EMC.CLAR11ON.fcp.rte.5.2.0.1</td>
<td>• EMC.CLAR11ON.fcp.rte.5.2.0.1</td>
</tr>
<tr>
<td><strong>APAR’s</strong></td>
<td><strong>APAR’s</strong></td>
</tr>
<tr>
<td><strong>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.2 ML 4:</strong></td>
<td><strong>Authorized Problem Analysis Reports (APARs) for AIX 5L v5.3 ML02:</strong></td>
</tr>
</tbody>
</table>
| • IY63133: large percentage of CPU time spent in ldata_balance routine | • IY68989: "WRITE TO MMAPPED SPACE HANGS"
| • IY63366: disym returns null even for valid symbol in AIX520 ML-4 | • 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.
| • IY64691: chwg -b can cause corruption and crash | • 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.
| • IY64737: AIO can hang in knutunlock | • JDK: |
| • IY65001: mklvcopy on a striped lv is failing to update lvcb | • IYS8350 Patch for SDK 1.3.1.16 (32-bit)
| • IY64978: deadlock with concurrent renaming and unlinking under JFS | • IY63533 Patch for SDK 1.4.2.1 (64-bit)
| • IY65001 | • IY65305 Patch for SDK 1.4.2.2 (32-bit)
| • IY75901, if IY69518 was previously installed. | **APARs required for GPFS v2.3:** |
| • 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. | • IY63969: mandatory service for GPFS v2.3 for AIX
| **JDK:** | • IY69911: corrupted kernel using direct I/O across GPFS block boundary
| • IY58350 Patch for SDK 1.3.1.16 (32-bit) | • IY70276: GPFS returns out of space when there is still space in the filesystem
| • IY63533 Patch for SDK 1.4.2.1 (64-bit) | • IY70277: receive enospc on mkfile in a GPFS file system when file system not full
| • IY65305 Patch for SDK 1.4.2.2 (32-bit) | • IY74097 (efix): corrupt archive library after adding members
| **APARs required for GPFS v2.3:** | **APARs required for GPFS v2.3:** |
| • IY63969: mandatory service for GPFS v2.3 for AIX | • IY63969: mandatory service for GPFS v2.3 for AIX
| • IY69911: corrupted kernel using direct I/O across GPFS block boundary | • 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 | • 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 | • 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

<table>
<thead>
<tr>
<th>Oracle CD's needed for the RAC installation</th>
<th>Optional CD's Not used in the RAC installation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle Database 10g Release 2 (10.2.0.1.0) Enterprise/Standard Edition for AIX5L 10gr2_aix5i64_database.cpio.gz (1,268,576,110 bytes) (cksum - 3772623559)</td>
<td>Oracle Database 10g Companion CD Release 2 (10.2.0.1.0) 10gr2_aix5i64_companion.cpio.gz (1,332,394,319 bytes) (cksum - 303439033)</td>
</tr>
<tr>
<td>Oracle Clusterware Release 2 (10.2.0.1.0) 10gr2_aix5i64_cluster.cpio.gz (732,798,920 bytes) (cksum - 1766193337)</td>
<td><strong>Oracle HTML DB v2.0</strong> (56,104,916 bytes) - latest standalone version of HTML DB New (13-Sep-05)</td>
</tr>
<tr>
<td></td>
<td>Oracle Database 10g Client Release 2 (10.2.0.1.0) 10gr2_aix5i64_client.cpio.gz (1,053,970,850 bytes) (cksum - 2541472317)</td>
</tr>
</tbody>
</table>

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](http://otn.oracle.com).
4. Review the certification matrix for this product [here](http://otn.oracle.com).
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.

<table>
<thead>
<tr>
<th>Shell Limit (As Shown in smit)</th>
<th>Recommended Value for oracle user</th>
<th>Recommended Value for root user</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft FILE size</td>
<td>-1 (Unlimited)</td>
<td>-1 (Unlimited)</td>
</tr>
<tr>
<td>Soft CPU time</td>
<td>-1 (Unlimited)</td>
<td>-1 (Unlimited)</td>
</tr>
<tr>
<td>Note: This is the default value.</td>
<td></td>
<td>Note: This is the default value.</td>
</tr>
<tr>
<td>Soft DATA segment</td>
<td>-1 (Unlimited)</td>
<td>-1 (Unlimited)</td>
</tr>
<tr>
<td>Soft STACK size</td>
<td>-1 (Unlimited)</td>
<td>-1 (Unlimited)</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Network Tuning Parameter</th>
<th>Recommended Value on all nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipqmaxlen</td>
<td>512</td>
</tr>
<tr>
<td>rfc1323</td>
<td>1</td>
</tr>
<tr>
<td>sb_max</td>
<td>1310720</td>
</tr>
<tr>
<td>tcp_recvspace</td>
<td>65536</td>
</tr>
<tr>
<td>tcp_sendspace</td>
<td>65536</td>
</tr>
<tr>
<td>Udp_recvspace</td>
<td>655360</td>
</tr>
<tr>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>Udp_sendspace</td>
<td>65536</td>
</tr>
<tr>
<td></td>
<td>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</td>
</tr>
</tbody>
</table>
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: 
   
   ```bash
   # /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:

   ```bash
   # /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:

   ```bash
   # /usr/sbin/no -o parameter_name=value
   ```

   For example:

   ```bash
   # /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:

   ```bash
   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:

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

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

   ```bash
   /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 Attachments / 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 Attachments / 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 Attachments 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 Attachments 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,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,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
inetc 10.10.25.81 netmask 0xfffff00 broadcast 10.10.25.255
tcp_sendspace 131072 tcp_recvspace 65536

en2:
flags=1e080863,80<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT,CHECKSUM_OFFLOAD,CHAIN>
inetc 20.20.25.81 netmask 0xfffff00 broadcast 20.20.25.255
tcp_sendspace 131072 tcp_recvspace 65536

lo0: flags=1e08084b<UP,BROADCAST,LOOPBACK,RUNNING,SIMPLEX,MULTICAST,GROUPRT,64BIT>
inetc 127.0.0.1 netmask 0xfffff0000 broadcast 127.255.255.255
inet6 ::1/128

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:

<table>
<thead>
<tr>
<th>Network card</th>
<th>Public</th>
<th>VIP</th>
<th>RAC Interconnect</th>
<th>GPFS Interconnect (If GPFS used)</th>
</tr>
</thead>
<tbody>
<tr>
<td>On each node</td>
<td>en0</td>
<td>en0</td>
<td>en1</td>
<td>en2</td>
</tr>
</tbody>
</table>

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
```

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

```
<table>
<thead>
<tr>
<th>Public</th>
<th>VIP</th>
<th>RAC Interconnect (Private Network)</th>
<th>GPFS Interconnect (If GPFS used)</th>
</tr>
</thead>
<tbody>
<tr>
<td>en0</td>
<td>en0</td>
<td>en1</td>
<td>en2</td>
</tr>
<tr>
<td>Node Name</td>
<td>IP</td>
<td>Node Name</td>
<td>IP</td>
</tr>
<tr>
<td>node1</td>
<td>10.3.25.81</td>
<td>node1_vip</td>
<td>10.3.25.181</td>
</tr>
<tr>
<td>node2</td>
<td>10.3.25.82</td>
<td>node2_vip</td>
<td>10.3.25.182</td>
</tr>
<tr>
<td>node3</td>
<td>10.3.25.83</td>
<td>node3_vip</td>
<td>10.3.25.183</td>
</tr>
</tbody>
</table>
```
### 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_gfps
20.20.25.82   node2_gfps
20.20.25.83   node3_gfps
```

**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:

1. **Smitty chinet**
   - select en0
2. **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**

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

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

**$HOME/.rhosts**

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

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

```
node1:root-> su - oracle
```

You should have the entries on each node for:

- node1
- node2
- node3
- node1_gdfs
- node2_gdfs
- node3_gdfs
- node1
- node2
- node3
- /etc/hosts
- /etc/hosts.equiv
- $HOME/.rhosts

**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-> rsh node1 date
```

```
node2:oracle-> rsh node1 date
dirsvdv<sdv<v<sdv<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's password:
   [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.
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 output, or ask questions, are modified so that they do not...
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 ORACLE_HOME=/oh10g
- export AIXTHREAD_SCOPE=S
- export TEMP=/tmp
- export TMP=/tmp
- export TMPDIR=/tmp
- umask 022

(S for system-wide thread scope)

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.

On node 1...

```
node1:root-> lsdev -C disk | grep SCSI
hdisk0 Available Virtual SCSI Disk Drive
```

On node 2...

```
node1:root-> mkvg -f -y'oraclevg' -s'32' hdisk1
```

To list the internal disks:

Create a volume group called oraclevg:

On node 1...

```
node1:root-> mkvg -f -y'oraclevg' -s'32' hdisk1
```

On node 2...

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

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

On node 2 ...

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

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

<table>
<thead>
<tr>
<th>Storage Option</th>
<th>OCR and Voting Disk</th>
<th>Oracle Software</th>
<th>Database</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic Storage Management</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IBM General Parallel File System (GPFS)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Local Storage</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Raw Logical volumes Managed by HACMP</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

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

ORACLE_SID=ASMDB1

Node 1

- Local Disks
- AIX5L (5.2/5.3) MLxx
- $ORACLE_HOME
- $ORACLE_CRS
- Listener / Tnsnames

- Shared Raw Disks
- OCR (Oracle Cluster Registry)
- Voting Disk
- Spfile
- Datafiles

- Shared ASM Raw Disks
- Redo Logs for each instance
- UNDO tablespace for each instance
- Archive logs for each instance

DATABASE_NAME=ASMDB

No shared code possible without GPFS

ORACLE_SID=ASMDB2

Node 2

- Local Disks
- AIX5L (5.2/5.3) MLxx
- $ORACLE_HOME
- $ORACLE_CRS
- Listener / Tnsnames
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.

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
<th>LUN’s Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>200 MB</td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>200 MB</td>
</tr>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td>n GB</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td>n GB</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td>n GB</td>
</tr>
<tr>
<td></td>
<td>……</td>
<td>……</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA (meaning Lun 10)</td>
<td>n GB</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB (meaning Lun 11)</td>
<td>n GB</td>
</tr>
</tbody>
</table>

8.1.2 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 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.

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
<th>Node 1 Corresponding hdisk</th>
<th>Node 2 Corresponding hdisk</th>
<th>Node …. Corresponding hdisk</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are two methods to identify the corresponding hdisks:

- Identify LUN ID assign to hdisk, using “Iscfg –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 ...**

```bash
node1:root-/ > lsdev -Cc disk
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
```

**On node 2 ...**

```bash
Node2:root-/ > lsdev -Cc disk
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:

<table>
<thead>
<tr>
<th>On node 1 ...</th>
<th>On node 2 ...</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>lspv</code></td>
<td><code>lspv</code></td>
</tr>
<tr>
<td><code>hdisk0</code> 10033c670e214eac5 rootvg active`</td>
<td><code>hdisk0</code> 0033c670e214eac5 rootvg active`</td>
</tr>
<tr>
<td><code>hdisk1</code> none</td>
<td><code>hdisk1</code> none</td>
</tr>
<tr>
<td><code>hdisk2</code> none</td>
<td><code>hdisk2</code> none</td>
</tr>
<tr>
<td><code>hdisk3</code> none</td>
<td><code>hdisk3</code> none</td>
</tr>
</tbody>
</table>

⇒ NO PVID are assigned apart for the rootvg hdisk

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

Identify LUN ID assign to hdisk, using “`lscfg -vl hdisk?`” command

<table>
<thead>
<tr>
<th>On node 1 ...</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>lsconf -vl hdisk2</code></td>
</tr>
<tr>
<td><code>hdisk2</code> U1.9-P1-11/Q1-W200300A0B80C5404-L100000000000000 3552 (500 Disk Array Device`</td>
</tr>
<tr>
<td><code>lsconf -vl hdisk3</code></td>
</tr>
<tr>
<td><code>hdisk3</code> U1.9-P1-11/Q1-W200300A0B80C5404-L200000000000000 3552 (500 Disk Array Device`</td>
</tr>
</tbody>
</table>

Then, We get the following table:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
<th>Node 1 Corresponding hdisk</th>
<th>Node 2 Corresponding hdisk</th>
<th>Node .... Corresponding hdisk</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>hdisk2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>hdisk3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⇒ No need to assign PVID when using this method.

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

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

Identify LUN ID assign to hdisk, using “`lscfg -vl hdisk?`” command

<table>
<thead>
<tr>
<th>On node 2 ...</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>lsconf -vl hdisk3</code></td>
</tr>
<tr>
<td><code>hdisk3</code> U1.9-P1-11/Q1-W200300A0B80C5404-L100000000000000 3552 (500 Disk Array Device`</td>
</tr>
<tr>
<td><code>lsconf -vl hdisk4</code></td>
</tr>
<tr>
<td><code>hdisk4</code> U1.9-P1-11/Q1-W200300A0B80C5404-L200000000000000 3552 (500 Disk Array Device`</td>
</tr>
</tbody>
</table>

Then, We get the following table:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
<th>Node 1 Corresponding hdisk</th>
<th>Node 2 Corresponding hdisk</th>
<th>Node .... Corresponding hdisk</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>hdisk2</td>
<td>hdisk3</td>
<td></td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>hdisk3</td>
<td>hdisk4</td>
<td></td>
</tr>
</tbody>
</table>

⇒ No need to assign PVID when using this method.
Using Second method:

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

<table>
<thead>
<tr>
<th>List available hdisk on each node:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>On node 1</strong>...</td>
</tr>
<tr>
<td><code>node1</code> root:/&gt; <code>lpv</code></td>
</tr>
<tr>
<td>hdisk0 0033c670e214eac5  rootvg  active</td>
</tr>
<tr>
<td>hdisk1 none  None</td>
</tr>
<tr>
<td>hdisk2 none  None</td>
</tr>
<tr>
<td>hdisk3 none  None</td>
</tr>
<tr>
<td>hdisk4 none  None</td>
</tr>
<tr>
<td>hdisk5 none  None</td>
</tr>
<tr>
<td>......</td>
</tr>
<tr>
<td>➔ NO PVID are assigned apart for the rootvg hdisk</td>
</tr>
</tbody>
</table>

| **On node 2**...                   |
| `node2` root:/> `lpv`              |
| 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 |

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

<table>
<thead>
<tr>
<th>Assign PVID to hdisk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>On node 1</strong>...</td>
</tr>
<tr>
<td><code>node1</code> root:/&gt; <code>chdev - l hdisk -a pv=yes</code></td>
</tr>
<tr>
<td><code>node1</code> root:/&gt; <code>chdev - l hdisk3 -a pv=yes</code></td>
</tr>
</tbody>
</table>

| **On node 2**...                   |
| `node2` root:/> `chdev - l hdisk2 -a pv=yes` |
| `node2` root:/> `chdev - l hdisk3 -a pv=yes` |

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

<table>
<thead>
<tr>
<th>List available hdisk with assigned PVID on each node:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>On node 1</strong>...</td>
</tr>
<tr>
<td><code>node1</code> root:/&gt; <code>lpv</code></td>
</tr>
<tr>
<td>hdisk0 0033c670e214eac5  rootvg  active</td>
</tr>
<tr>
<td>hdisk1 none  None</td>
</tr>
<tr>
<td>hdisk2 0033c670b0f3e84  None</td>
</tr>
<tr>
<td>hdisk3 0033c670b0f3b87  None</td>
</tr>
<tr>
<td>hdisk4 none  None</td>
</tr>
<tr>
<td>hdisk5 none  None</td>
</tr>
<tr>
<td>......</td>
</tr>
<tr>
<td>➔ PVID are assigned for possible hdisk to be used as ASM Disks</td>
</tr>
</tbody>
</table>

| **On node 2**...                                       |
| `node2` root:/> `lpv`                                  |
| hdisk0 0033c670e214eac5  rootvg  active                |
| hdisk1 none  None                                      |
| hdisk2 0033c670b0f3e82  None                           |
| hdisk3 0033c670b0f3b84  None                           |
| hdisk4 none  None                                      |
| hdisk5 none  None                                      |
| ......                                                 |
| ➔ PVID are assigned for possible hdisk to be used as ASM Disks |

⚠️ Be careful! hdisk4 on node1 is not necessary hdisk4 on node2.

THEN We get the following table:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN's ID Number</th>
<th>Node 1 Corresponding hdisk</th>
<th>Node 2 Corresponding hdisk</th>
<th>Node .... Corresponding hdisk</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>hdisk2</td>
<td>hdisk3</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>hdisk3</td>
<td>hdisk4</td>
<td>hdisk?</td>
</tr>
</tbody>
</table>

g ➔ PVID are assigned when using this method, THEN we need to clear PVID before going further.
**On node 1 ...**

node1:root/> chdev -l hdisk2 -a pv=clear
node1:root/> chdev -l hdisk3 -a pv=clear
node1:root/> lsvg

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/> lsvg

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 –I hdisk2” to visualize all attributes for hdisk4
Or only lsattr –E –I hdisk2 | grep reserve

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

- On IBM storage (ESS, FasTt, DSXXX) : Change the “reserve_policy” attribute to “no_reserve”
  ```
  chdev -I hdisk? -a reserve_policy=no_reserve
  ```

- On EMC storage : Change the “reserve_lock” attribute to “no”
  ```
  chdev -I 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 -I hdisk2 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk3 -a reserve_policy=no_reserve
```

On node 2 ...
```
node1:root/> rsh node2
node2:root/> chdev -I hdisk3 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk4 -a reserve_policy=no_reserve
```

Example for one hdisk:

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

```
$ lsattr -E -I 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...

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN's ID Number</th>
<th>Device Name</th>
<th>Node 1 Corresponding hdisk</th>
<th>Major Num.</th>
<th>Minor Num.</th>
<th>Node 2 Corresponding hdisk</th>
<th>Major Num.</th>
<th>Minor Num.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>/dev/ocr_disk</td>
<td>hdisk2</td>
<td></td>
<td></td>
<td>hdisk3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>/dev/vote_disk</td>
<td>hdisk3</td>
<td></td>
<td></td>
<td>hdisk4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

**On node1:**

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

**On node2:**

```bash
node2:root/> ls -l /dev/*hdisk
brw------- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk3
crw-------- 1 root system 20, 5 Oct 18 17:52 /dev/rhdisk3
```

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

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

```bash
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** ...

```bash
chown oracle.db /dev/ocr_disk
chown oracle.db /dev/vote_disk
chmod 660 /dev/ocr_disk
chmod 660 /dev/vote_disk
```

```bash
l $ ls -l /dev/* | grep "20,.4"
```

```bash
brw------- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk2
crw-rw--- 1 oracle dba 20, 4 Oct 19 13:27 /dev/ocr_disk
```

```bash
l $ ls -l /dev/* | grep "20,.5"
```

```bash
brw------- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk3
crw-rw--- 1 oracle dba 20, 5 Oct 19 13:28 /dev/vote_disk
```

**On node2** ...

```bash
rsh node2
```

```bash
chown oracle.db /dev/ocr_disk
chown oracle.db /dev/vote_disk
chmod 660 /dev/ocr_disk
chmod 660 /dev/vote_disk
```

```bash
ls -l /dev/* | grep "20,.4"
```

```bash
brw------- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk3
crw-rw--- 1 oracle dba 20, 4 Oct 19 13:27 /dev/ocr_disk
```

```bash
ls -l /dev/* | grep "20,.5"
```

```bash
brw------- 1 root system 20, 5 Oct 18 17:52 /dev/hdisk4
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:

```bash
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:**

```bash
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 &
e tc ...
```

**At the same time,**

**From node2:**

```bash
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 &
e tc ...
```
8.1.4 Preparing Disks for ASM

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

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA (meaning Lun 10)</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB (meaning Lun 11)</td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC (meaning Lun 12)</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s Number</th>
<th>Node 1</th>
<th>Node 2</th>
<th>Node ....</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

There are two methods to identify the corresponding hdisks:

- Identify LUN ID assign to hdisk, using “lscfg –I 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 –vl hdisk?” command

Be careful! hdisk4 on node1 is not necessary hdisk4 on node2.

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s Number</th>
<th>Node 1</th>
<th>Node 2</th>
<th>Node ....</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td>hdisk4</td>
<td>hdisk5</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td>hdisk5</td>
<td>hdisk6</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td>hdisk6</td>
<td>hdisk7</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
<td>hdisk7</td>
<td>hdisk8</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
<td>hdisk8</td>
<td>hdisk9</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
<td>hdisk9</td>
<td>hdisk10</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
<td>hdisk10</td>
<td>hdisk11</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA</td>
<td>hdisk11</td>
<td>hdisk12</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB</td>
<td>hdisk12</td>
<td>hdisk13</td>
<td>hdisk?</td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC</td>
<td>hdisk13</td>
<td>hdisk14</td>
<td>hdisk?</td>
</tr>
</tbody>
</table>

Then we get the following table:

⇒ No need to assign PVID when using this method.
Using Second method:

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

<table>
<thead>
<tr>
<th>On node 1 ...</th>
<th>On node 2 ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>_node1.root-/&gt; _lspv _hdisk_0 0033c670e214ec5 rootvg active _hdisk_1 none None _hdisk_2 none None _hdisk_3 none None _hdisk_4 none None _hdisk_5 none None _hdisk_6 none None _hdisk_7 none None _hdisk_8 none None _hdisk_9 none None _hdisk_10 none None _hdisk_11 none None _hdisk_12 none None _hdisk_13 none None</td>
<td>_node2.root-/&gt; _lspv _hdisk_0 0033c670e214ec5 rootvg active _hdisk_1 none None _hdisk_2 none None _hdisk_3 none None _hdisk_4 none None _hdisk_5 none None _hdisk_6 none None _hdisk_7 none None _hdisk_8 none None _hdisk_9 none None _hdisk_10 none None _hdisk_11 none None _hdisk_12 none None _hdisk_13 none None _hdisk_14 none None</td>
</tr>
<tr>
<td>_node1.root-/&gt; _chdev -l _hdisk_4 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_5 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_6 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_7 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_8 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_9 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_10 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_11 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_12 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_13 -a pv=yes</td>
<td>_node2.root-/&gt; _chdev -l _hdisk_5 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_6 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_7 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_8 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_9 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_10 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_11 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_12 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_13 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_14 -a pv=yes</td>
</tr>
</tbody>
</table>

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.

| List available hdisks with assigned PVID on each node: | List available hdisks on each node: |

<table>
<thead>
<tr>
<th>On node 1 ...</th>
<th>On node 2 ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>_node1.root-/&gt; _lspv _hdisk_0 0033c670e214ec9 rootvg active _hdisk_5 0033c670e224ec5 None _hdisk_6 0033c670e234ec5 None _hdisk_7 0033c670e244ec5 None _hdisk_8 0033c670e254ec5 None _hdisk_9 0033c670e264ec5 None</td>
<td>_node2.root-/&gt; _lspv _hdisk_0 0033c670e214ec9 rootvg active _hdisk_5 0033c670e224ec5 None _hdisk_6 0033c670e234ec5 None _hdisk_7 0033c670e244ec5 None _hdisk_8 0033c670e254ec5 None _hdisk_9 0033c670e264ec5 None</td>
</tr>
<tr>
<td>_node1.root-/&gt; _chdev -l _hdisk_4 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_5 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_6 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_7 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_8 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_9 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_10 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_11 -a pv=yes _node1.root-/&gt; _chdev -l _hdisk_12 -a pv=yes</td>
<td>_node2.root-/&gt; _chdev -l _hdisk_5 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_6 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_7 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_8 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_9 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_10 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_11 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_12 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_13 -a pv=yes _node2.root-/&gt; _chdev -l _hdisk_14 -a pv=yes</td>
</tr>
</tbody>
</table>

\_node2.root\-/> \_lspv
\_hdisk\_0 0033c670e214ec9 rootvg active
\_hdisk\_1 none None
\_hdisk\_2 none None
\_hdisk\_3 none None
\_hdisk\_4 none None
\_hdisk\_5 0033c670e214ec9 None
\_hdisk\_6 0033c670e224ec5 None
\_hdisk\_7 0033c670e234ec5 None
\_hdisk\_8 0033c670e244ec5 None
\_hdisk\_9 0033c670e254ec5 None
\_hdisk\_10 0033c670e264ec5 None
\_hdisk\_11 0033c670e274ec5 None
\_hdisk\_12 0033c670e284ec5 None
\_hdisk\_13 0033c670e314ec5 None

\_node2.root\-/> \_chdev -l \_hdisk\_5 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_6 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_7 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_8 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_9 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_10 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_11 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_12 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_13 -a pv=yes
\_node2.root\-/> \_chdev -l \_hdisk\_14 -a pv=yes

Thus:

- \_PVID are assigned for possible hdisk to be used as ASM Disks

AND identify corresponding hdisks on each node:!!!

Be careful: hdisk4 on node1 is not necessary, hdisk4 on node2.

\_NO PVID are assigned apart for the rootvg hdisk

\_NO PVID are assigned apart for the rootvg hdisk

\_PVID are assigned for possible hdisk to be used as ASM Disks
THEN We get the following table:

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

¬ 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

... 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, DSXXX) : Change the “reserve_policy” attribute to “no_reserve”
  ```
  chdev -I hdisk? -a reserve_policy=no_reserve
  ```

- On EMC storage : Change the “reserve_lock” attribute to “no”
  ```
  chdev -I 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 -I hdisk4 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk5 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk6 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk7 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk8 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk9 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk10 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk11 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk12 -a reserve_policy=no_reserve
node1:root/> chdev -I hdisk13 -a reserve_policy=no_reserve
```

```
On node 2 ...

node2:root/> chdev -I hdisk5 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk6 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk7 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk8 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk9 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk10 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk11 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk12 -a reserve_policy=no_reserve
node2:root/> chdev -I hdisk13 -a reserve_policy=no_reserve
node2:root/> chdev -I 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, hdisk might have different names from one node to another for example hdisk4 on node1 might be hdisk5 on node2, as shown in following table:

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s Number</th>
<th>Node 1</th>
<th>Node 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td>hdisk4</td>
<td>hdisk5</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td>hdisk5</td>
<td>hdisk6</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td>hdisk6</td>
<td>hdisk7</td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
<td>hdisk7</td>
<td>hdisk8</td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
<td>hdisk8</td>
<td>hdisk9</td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
<td>hdisk9</td>
<td>hdisk10</td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
<td>hdisk10</td>
<td>hdisk11</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA</td>
<td>hdisk11</td>
<td>hdisk12</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB</td>
<td>hdisk12</td>
<td>hdisk13</td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC</td>
<td>hdisk13</td>
<td>hdisk14</td>
</tr>
</tbody>
</table>

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

OR

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

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s Number</th>
<th>Node 1</th>
<th>Node 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td>hdisk4</td>
<td>hdisk4</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td>hdisk5</td>
<td>hdisk5</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td>hdisk6</td>
<td>hdisk6</td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
<td>hdisk7</td>
<td>hdisk7</td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
<td>hdisk8</td>
<td>hdisk8</td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
<td>hdisk9</td>
<td>hdisk9</td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
<td>hdisk10</td>
<td>hdisk10</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA</td>
<td>hdisk11</td>
<td>hdisk11</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB</td>
<td>hdisk12</td>
<td>hdisk12</td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC</td>
<td>hdisk13</td>
<td>hdisk13</td>
</tr>
</tbody>
</table>

→ THEN, we can implement devices to have same other naming

OR use hdisk as it is....
### Making devices (if wanted):

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk 1 for ASM</td>
<td>L3</td>
<td>/dev/asm_disk1</td>
<td>hdisk4</td>
<td>20</td>
<td>3</td>
<td>hdisk5</td>
<td>12</td>
</tr>
<tr>
<td>Disk 2 for ASM</td>
<td>L4</td>
<td>/dev/asm_disk2</td>
<td>hdisk5</td>
<td>20</td>
<td>4</td>
<td>hdisk6</td>
<td>12</td>
</tr>
<tr>
<td>Disk 3 for ASM</td>
<td>L5</td>
<td>/dev/asm_disk3</td>
<td>hdisk6</td>
<td>...</td>
<td>...</td>
<td>hdisk7</td>
<td>...</td>
</tr>
<tr>
<td>Disk 4 for ASM</td>
<td>L6</td>
<td>/dev/asm_disk4</td>
<td>hdisk7</td>
<td>...</td>
<td>...</td>
<td>hdisk8</td>
<td>...</td>
</tr>
<tr>
<td>Disk 5 for ASM</td>
<td>L7</td>
<td>/dev/asm_disk5</td>
<td>hdisk8</td>
<td>...</td>
<td>...</td>
<td>hdisk9</td>
<td>...</td>
</tr>
<tr>
<td>Disk 6 for ASM</td>
<td>L8</td>
<td>/dev/asm_disk6</td>
<td>hdisk9</td>
<td>...</td>
<td>...</td>
<td>hdisk10</td>
<td>...</td>
</tr>
<tr>
<td>Disk 7 for ASM</td>
<td>L9</td>
<td>/dev/asm_disk7</td>
<td>hdisk10</td>
<td>...</td>
<td>...</td>
<td>hdisk11</td>
<td>...</td>
</tr>
<tr>
<td>Disk 8 for ASM</td>
<td>LA</td>
<td>/dev/asm_disk8</td>
<td>hdisk11</td>
<td>...</td>
<td>...</td>
<td>hdisk12</td>
<td>...</td>
</tr>
<tr>
<td>Disk 9 for ASM</td>
<td>LB</td>
<td>/dev/asm_disk9</td>
<td>hdisk12</td>
<td>...</td>
<td>...</td>
<td>hdisk13</td>
<td>...</td>
</tr>
<tr>
<td>Disk 10 for ASM</td>
<td>LC</td>
<td>/dev/asm_disk10</td>
<td>hdisk13</td>
<td>...</td>
<td>...</td>
<td>Hdisk14</td>
<td>...</td>
</tr>
</tbody>
</table>

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

**On node 1:**

node1:root/> ls -l /dev/*hdisk*

brw------- 1 root system 20, 3 Oct 18 17:52 /dev/hdisk4
crw------- 1 root system 20, 3 Oct 18 17:52 /dev/rhdisk4

**On node 2:**

node2:root/> ls -l /dev/*hdisk*

brw------- 1 root system 20, 4 Oct 18 17:52 /dev/hdisk5
crw------- 1 root system 20, 4 Oct 18 17:52 /dev/rhdisk5

**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

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:

```
/disk1
/disk2
/disk...
```

on each node of the cluster:

```
On node1 ...
```

```
node1:root/> chown oracle.db /dev/asm_disk1
node1:root/> chmod 660 /dev/asm_disk1
```

```
node1:root/> ls -l /dev/* | grep "20, 3"
```

```
brw------- 1 root system  20, 3 Oct 18 17:52 /dev/hdisk2
```

```
crw------- 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
```

```
crw------- 1 oracle dba   20, 4 Oct 19 13:28 /dev/asm_disk2
```

```
node2:root/> rsh node2
```

```
node2:root/> chown oracle.db /dev/asm_disk1
node2:root/> chmod 660 /dev/asm_disk1
```

```
node2:root/> ls -l /dev/* | grep "12, 9"
```

```
brw------- 1 root system  12, 9 Oct 18 17:52 /dev/hdisk3
```

```
crw------- 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
```

```
crw------- 1 oracle dba   12, 8 Oct 19 13:28 /dev/asm_disk2
```

```
node1:root/> rsh node2
```

```
node1:root/> chown oracle.db /dev/asm_disk2
node1: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
```

```
crw------- 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
```

```
crw------- 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 oracledba /dev/rhdisks
node1:root/> chown oracledba /dev/rhdisks5
node1:root/> chown oracledba /dev/rhdisks6
node1:root/> chown oracledba /dev/rhdisks7
node1:root/> chown oracledba /dev/rhdisks8
node1:root/> chown oracledba /dev/rhdisks9
node1:root/> chown oracledba /dev/rhdisks10
node1:root/> chown oracledba /dev/rhdisks11
node1:root/> chown oracledba /dev/rhdisks12
node1:root/> chown oracledba /dev/rhdisks13
```

```
node1:root/> chmod 660 /dev/rhdisks
node1:root/> chmod 660 /dev/rhdisks5
node1:root/> chmod 660 /dev/rhdisks6
node1:root/> chmod 660 /dev/rhdisks7
node1:root/> chmod 660 /dev/rhdisks8
node1:root/> chmod 660 /dev/rhdisks9
node1:root/> chmod 660 /dev/rhdisks10
node1:root/> chmod 660 /dev/rhdisks11
node1:root/> chmod 660 /dev/rhdisks12
node1:root/> chmod 660 /dev/rhdisks13
node1:root/> chmod 660 /dev/rhdisks14
```

**On node 2 …**

```
node2:root/> chown oracledba /dev/rhdisks5
node2:root/> chown oracledba /dev/rhdisks6
node2:root/> chown oracledba /dev/rhdisks7
node2:root/> chown oracledba /dev/rhdisks8
node2:root/> chown oracledba /dev/rhdisks9
node2:root/> chown oracledba /dev/rhdisks10
node2:root/> chown oracledba /dev/rhdisks11
node2:root/> chown oracledba /dev/rhdisks12
node2:root/> chown oracledba /dev/rhdisks13
node2:root/> chown oracledba /dev/rhdisks14
```

```
node2:root/> chmod 660 /dev/rhdisks5
node2:root/> chmod 660 /dev/rhdisks6
node2:root/> chmod 660 /dev/rhdisks7
node2:root/> chmod 660 /dev/rhdisks8
node2:root/> chmod 660 /dev/rhdisks9
node2:root/> chmod 660 /dev/rhdisks10
node2:root/> chmod 660 /dev/rhdisks11
node2:root/> chmod 660 /dev/rhdisks12
node2:root/> chmod 660 /dev/rhdisks13
node2:root/> chmod 660 /dev/rhdisks14
```

**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/rhdisks bs=8192 count=25000 &
  25000+0 records in.
  25000+0 records out.
dd if=/dev/zero of=/dev/rhdisks5 bs=8192 count=25000 &
  25000+0 records in.
  25000+0 records out.
dd if=/dev/zero of=/dev/rhdisks7 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

<table>
<thead>
<tr>
<th>crw-rw----</th>
<th>oracle</th>
<th>dba</th>
<th>20, 4 Oct 18 17:52 /dev/rhdisk4</th>
</tr>
</thead>
<tbody>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk5</td>
</tr>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk6</td>
</tr>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk7</td>
</tr>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk8</td>
</tr>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk9</td>
</tr>
<tr>
<td>crw-rw----</td>
<td>oracle</td>
<td>dba</td>
<td>20, 4 Oct 18 17:52 /dev/rhdisk10</td>
</tr>
</tbody>
</table>

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

OR

(node1:root)/ -> lquerypv -h /dev/rhdisk7

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 ...)

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

NON used ASM disks will give the following display:

```
(node1.root)/ ->querypv -h /dev/rhdisk2
00000000 00000000 00000000 00000000 00000000 .........
00000010 00000000 00000000 00000000 00000000 .........
00000020 00000000 00000000 00000000 00000000 .........
00000030 00000000 00000000 00000000 00000000 .........
00000040 00000000 00000000 00000000 00000000 .........
00000050 00000000 00000000 00000000 00000000 .........
00000060 00000000 00000000 00000000 00000000 .........
00000070 00000000 00000000 00000000 00000000 .........
00000080 00000000 00000000 00000000 00000000 .........
00000090 00000000 00000000 00000000 00000000 .........
000000A0 00000000 00000000 00000000 00000000 .........
000000B0 00000000 00000000 00000000 00000000 .........
000000C0 00000000 00000000 00000000 00000000 .........
000000D0 00000000 00000000 00000000 00000000 .........
000000E0 00000000 00000000 00000000 00000000 .........
000000F0 00000000 00000000 00000000 00000000 .........
```

```
(node1.root)/ ->querypv -h /dev/rhdisk8
00000000 00820101 00000000 80000000 DEC8B940 .........@
00000010 00000000 00000000 00000000 00000000 .........|
00000020 4F52434C 44495348 00000000 00000000 .........|
00000030 00000000 00000000 00000000 00000000 .........|
00000040 0A100000 00000000 41534D44 425F464C .........|ASMDB_FL|
00000050 41534852 45534F56 4552595F 00000030 .........|ASHRECOVERY_0000|
00000060 00000000 00000000 41534D44 425F464C .........|ASMDB_FL|
00000070 41534852 45534F56 45525900 00000000 .........|ASHRECOVERY.....|
00000080 00000000 00000000 41534D44 425F464C .........|ASMDB_FL|
00000090 41534852 45534F56 4552595F 00000030 .........|ASHRECOVERY_0000|
000000A0 00000000 00000000 00000000 00000000 .........|
000000B0 00000000 00000000 00000000 00000000 .........|
000000C0 00000000 00000000 01F5874E 4E47F800 .........|NNG..|
000000D0 01F588CA 14FABC00 02001000 00100000 .........|
000000E0 0018CB0 00014000 00000002 00000001 .........|
000000F0 00000002 00000002 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 ...)

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)/ ->querypv -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/c1resctr/library/gpfs_faqs.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 eServer pSeries
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 Attachements 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.

<table>
<thead>
<tr>
<th>Disks</th>
<th>LUN’s ID Number</th>
<th>LUN’s Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCR</td>
<td>L1</td>
<td>200 MB</td>
</tr>
<tr>
<td>Voting</td>
<td>L2</td>
<td>200 MB</td>
</tr>
<tr>
<td>DB files</td>
<td>L3</td>
<td>n GB</td>
</tr>
<tr>
<td>DB files</td>
<td>L4</td>
<td>n GB</td>
</tr>
<tr>
<td>DB files</td>
<td>L5</td>
<td>n GB</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>DB files</td>
<td>LA (meaning Lun 10)</td>
<td>n GB</td>
</tr>
<tr>
<td>DB files</td>
<td>LB (meaning Lun 11)</td>
<td>n GB</td>
</tr>
</tbody>
</table>

### 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_gpf cors ssh or rsh to node1_gpf, as root user.
node1_gpf must ssh or rsh to node2_gpf, as root user.
node2_gpf must ssh or rsh to node1_gpf, as root user.
node2_gpf must ssh or rsh to node2_gpf, 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.

Fil sesets gpf s.xx.xxx

See Appendix H: Files 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:

ls lpp -L | grep gpf s
Post install task

Add GPFS binaries in root’s path.

export PATH=SPATH:/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_gpf:quorum
node2_gpf:quorum

- Create the GPFS cluster

node1:> mmrcrcluster -p node1_gpf -s node2_gpf -n /var/mmfs/etc/node.list -C Demo10g -A
You can check the GPFS creation with the mmlscluter and the cluster configuration with mmlsconfig. In our example you see:

```
node1:/> mmlscluter

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
```
Installing Oracle 10g RAC Release 2 on IBM pSeries with AIX 5L

- 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 NFT 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
```

<table>
<thead>
<tr>
<th>File system</th>
<th>Disk name</th>
<th>Primary node</th>
<th>Backup node</th>
</tr>
</thead>
<tbody>
<tr>
<td>(free disk)</td>
<td>nsd1</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>(free disk)</td>
<td>nsd2</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>(free disk)</td>
<td>nsd3</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>(free disk)</td>
<td>tie</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
</tbody>
</table>

- 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 mmcrlv
-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
```

<table>
<thead>
<tr>
<th>File system</th>
<th>Disk name</th>
<th>Primary node</th>
<th>Backup node</th>
</tr>
</thead>
<tbody>
<tr>
<td>data_gpfS</td>
<td>nsd1</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>data_gpfS</td>
<td>nsd2</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>data_gpfS</td>
<td>nsd3</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
<tr>
<td>(free disk)</td>
<td>tie</td>
<td>node1_gpfS</td>
<td>node2_gpfS</td>
</tr>
</tbody>
</table>

```
node1:~/> mmnconfig
```

Configuration data for cluster Demo10g.node1_gpfS:

```
clansterName Demo10g.node1_gpfS
clansterId 1446809217114175696
clansterType lc
multinode yes
autol0ad yes
useDiskLease yes
maxFeatureLevelAllowed 809
tiebreakerDisks tie
```

File systems in cluster Demo10g.node1_gpfS:
```
/dev/data_gpfS
```

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

Installing Oracle 10g RAC Release 2 on IBM server pSeries with AIX 5L
/data_gpf:
dev = /dev/data_gpf
vfs = mmfs
nodename = -
mount = mmfs
type = mmfs
account = false
options = rw,mtime,atime,dev=data_gpf

- Mount the newly created file-system:

    node1:/> mount /data_gpf

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

    node1:/> chown oracle:dba /data_gpf
    node1:/> chmod go+rw /data_gpf

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

- **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

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

2. Using any text editor, edit the /etc/ntp.conf file:
   
   ```bash
   # vi /etc/ntp.conf
   ```

3. Add entries similar to the following to the file:
   
   ```bash
   # 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:
   
   ```bash
   # /usr/bin/smitty xntpd
   ```

   b. Choose Start Using the xntpd Subsystem, then choose BOTH.
10 IMPORTANT TIPS FOR ORACLE SOFTWARE AND PATCHS 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.

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#1026198

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

```bash
.../clusterware/Disk1/rootpre/rootpre.sh
```

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

On node1

```bash
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

```bash
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 storagID_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"

<table>
<thead>
<tr>
<th>Destination Node</th>
<th>Reachable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>yes</td>
</tr>
<tr>
<td>node2</td>
<td>yes</td>
</tr>
</tbody>
</table>

Result: Node reachability check passed from node "node1".

Checking user equivalence...

Check: User equivalence for user "root"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node2</td>
<td>passed</td>
</tr>
<tr>
<td>node1</td>
<td>passed</td>
</tr>
</tbody>
</table>

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

Checking administrative privileges...

Check: Existence of user "root"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>User Exists</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>yes</td>
<td>passed</td>
</tr>
<tr>
<td>node2</td>
<td>yes</td>
<td>passed</td>
</tr>
</tbody>
</table>

Result: User existence check passed for "root".

Check: Existence of group "oinstall"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Group ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>exists</td>
<td>501</td>
</tr>
<tr>
<td>node2</td>
<td>exists</td>
<td>501</td>
</tr>
</tbody>
</table>

Result: Group existence check passed for "oinstall".

Check: Membership of user "oracle" in group "oinstall" [as Primary]

<table>
<thead>
<tr>
<th>Node Name</th>
<th>User Exists</th>
<th>Group Exists</th>
<th>User in Group</th>
<th>Primary</th>
</tr>
</thead>
<tbody>
<tr>
<td>tbas11b</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>tbas11a</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Result: Membership check for user "oracle" in group "oinstall" [as Primary] failed.

Administrative privileges check failed

⇒ this is equivalent to hostname ping …

⇒ 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

⇒ this is user and group existence tests on node1

The following message is not a big issue :

Result: Group existence check failed for "oinstall".

⇒ Just create a oinstall group at AIX level, using smitty as root
At this stage, node connectivity is checked !!!

Checking node connectivity...

Interface information for node “node1”

<table>
<thead>
<tr>
<th>Interface Name</th>
<th>IP Address</th>
<th>Subnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>en0</td>
<td>10.3.25.81</td>
<td>10.3.25.0</td>
</tr>
<tr>
<td>en1</td>
<td>10.10.25.81</td>
<td>10.10.25.0</td>
</tr>
<tr>
<td>en2</td>
<td>20.20.25.81</td>
<td>20.20.25.0</td>
</tr>
</tbody>
</table>

Interface information for node “node2”

<table>
<thead>
<tr>
<th>Interface Name</th>
<th>IP Address</th>
<th>Subnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>en0</td>
<td>10.3.25.82</td>
<td>10.3.25.0</td>
</tr>
<tr>
<td>en1</td>
<td>10.10.25.82</td>
<td>10.10.25.0</td>
</tr>
<tr>
<td>en2</td>
<td>20.20.25.82</td>
<td>20.20.25.0</td>
</tr>
</tbody>
</table>

Check: Node connectivity of subnet “10.3.25.0”

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
<th>Connected?</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1:en0</td>
<td>node2:en0</td>
<td>yes</td>
</tr>
</tbody>
</table>

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”

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
<th>Connected?</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1:en1</td>
<td>node2:en1</td>
<td>yes</td>
</tr>
</tbody>
</table>

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”

<table>
<thead>
<tr>
<th>Source</th>
<th>Destination</th>
<th>Connected?</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1:en2</td>
<td>node2:en2</td>
<td>yes</td>
</tr>
</tbody>
</table>

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.

ERROR: Could not find a suitable set of interfaces for VIPs.

This is due to a CVU issue as explain bellow:

Metalink Node ID 338924.1

CLUUFY Fails With Error: Could not find a suitable set of interfaces for VIPs

Per BUG:4437727, cluufy 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 x.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
def2 | AIX 5.3 | AIX 5.2 | passed
Result: Kernel version check passed.

Check: System architecture
Node Name | Available | Required | Comment
--- | --- | --- | ---
node1 | powerpc | powerpc | passed
def2 | powerpc | powerpc | passed
Result: System architecture check passed.

Check: Total memory
Node Name | Available | Required | Comment
--- | --- | --- | ---
node1 | 2GB (2097152KB) | 512MB (524288KB) | passed
def2 | 2GB (2097152KB) | 512MB (524288KB) | passed
Result: Total memory check passed.

Check: Swap space
Node Name | Available | Required | Comment
--- | --- | --- | ---
node1 | 1024GB (1073741824KB) | 1GB (1048576KB) | passed
def2 | 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
def2 | 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
def2 | 4.63GB (4860044KB) | 4GB (4194304KB) | passed
Result: Free disk space check passed.

At this stage, node system requirements for crs is checked !!!

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

Installing Oracle 10g RAC Release 2 on IBM pSeries with AIX 5L

Page 79/174
### Check: Operating system patch for "IY63969"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY63969:gpfs.base</td>
<td>IY63969</td>
</tr>
<tr>
<td>node2</td>
<td>IY63969:gpfs.base</td>
<td>IY63969</td>
</tr>
</tbody>
</table>

Result: Package existence check passed for "gpfs.base:2.3.0.3".

### Check: Operating system patch for "IY69911"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY69911:gpfs.base</td>
<td>IY69911</td>
</tr>
<tr>
<td>node2</td>
<td>IY69911:gpfs.base</td>
<td>IY69911</td>
</tr>
</tbody>
</table>

Result: Package existence check passed for "cluster.license:5.2.0.0".

### Check: Operating system patch for "IY60759"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY60759</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>IY60759</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Package existence check failed for "cluster.license:5.2.0.0".

### Check: Operating system patch for "IY601034"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY601034:bos.mp1</td>
<td>IY601034</td>
</tr>
<tr>
<td>node2</td>
<td>IY601034:bos.mp1</td>
<td>IY601034</td>
</tr>
</tbody>
</table>

Result: Package existence check passed for "cluster.license:5.2.0.0".
<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY61770:rsct.basic.rteIY61770:rsct.core.errmIY61770:rsct.core.hostmIY61770:rsct.core.rmcIY61770:rsct.core.secY61770:rsct.core.sensorrmIY61770:rsct.core.utilsIY61770</td>
<td>passed</td>
<td></td>
</tr>
<tr>
<td>node2</td>
<td>IY61770:rsct.basic.rteIY61770:rsct.core.errmIY61770:rsct.core.hostmIY61770:rsct.core.rmcIY61770:rsct.core.secY61770:rsct.core.sensorrmIY61770:rsct.core.utilsIY61770</td>
<td>passed</td>
<td></td>
</tr>
</tbody>
</table>

**Result:** Operating system patch check passed for "IY61770".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "ElectricFence-2.2.2-1:2.2.2".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>xfrte:8.1.1.4</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>xfrte:8.1.1.4</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "xfrte:9.1".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "gdb-6.0-1:6.0".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "make-3.80-1:3.80".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "freeware.gnu.tar.rte:1.13.0.0".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "Java14_64.sdk:1.4.2.1".

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

**Result:** Package existence check failed for "Java131.rte.bin:1.3.1.16".
### Installing Oracle 10g RAC Release 2 on IBM Node 1

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>Java14.sdk:1.4.2.10</td>
<td></td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>Java14.sdk:1.4.2.10</td>
<td></td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Package existence check failed for "Java14.sdk:1.4.2.2".

### Check: Operating system patch for "IY65305"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY65305:Java14.sdk</td>
<td>IY65305</td>
<td>passed</td>
</tr>
<tr>
<td>node2</td>
<td>IY65305:Java14.sdk</td>
<td>IY65305</td>
<td>passed</td>
</tr>
</tbody>
</table>

Result: Operating system patch check passed for "IY65305".

### Check: Operating system patch for "IY58350"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>unknown</td>
<td>IY58350</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>unknown</td>
<td>IY58350</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Operating system patch check passed for "IY58350".

### Check: Operating system patch for "IY63533"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>unknown</td>
<td>IY63533</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>unknown</td>
<td>IY63533</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Operating system patch check failed for "IY63533".

### Check: Package existence for "mqm.server.rte:5.3"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Package existence check failed for "mqm.server.rte:5.3".

### Check: Package existence for "mqm.client.rte:5.3"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Package existence check failed for "mqm.client.rte:5.3".

### Check: Package existence for "sna.rte:6.1.0.4"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>missing</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>missing</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Package existence check failed for "sna.rte:6.1.0.4".

### Check: Package existence for "bos.net.tcp.server"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>bos.net.tcp.server:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>node2</td>
<td>bos.net.tcp.server:5.3.0.30</td>
<td>passed</td>
</tr>
</tbody>
</table>

Result: Package existence check passed for "bos.net.tcp.server".

### Check: Operating system patch for "IY44599"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>unknown</td>
<td>IY44599</td>
<td>failed</td>
</tr>
<tr>
<td>node2</td>
<td>unknown</td>
<td>IY44599</td>
<td>failed</td>
</tr>
</tbody>
</table>

Result: Operating system patch check failed for "IY44599".

### Check: Operating system patch for "IY60930"

<table>
<thead>
<tr>
<th>Node Name</th>
<th>Applied</th>
<th>Required</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>IY60930:os.mpi</td>
<td>IY60930</td>
<td>passed</td>
</tr>
<tr>
<td>node2</td>
<td>IY60930:os.mpi</td>
<td>IY60930</td>
<td>passed</td>
</tr>
</tbody>
</table>

Result: Operating system patch check passed for "IY60930".
<table>
<thead>
<tr>
<th>Node Name</th>
<th>Status</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>IY6513</td>
<td>bos.adt.base:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY6513</td>
<td>bos.adt.base:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY70159</td>
<td>bos.rte.bind_cmds: IY9386</td>
<td>passed</td>
</tr>
<tr>
<td>IY70159</td>
<td>bos.rte.bind_cmds: IY9386</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.adt.lib:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.adt.lib:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.adt.lib:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.adt.lib:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.perf.libperfstat:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.perf.libperfstat:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.perf.libperfstat:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>bos.perf.libperfstat:5.3.0.30</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>rsct.basic.rte:2.4.3.0</td>
<td>passed</td>
</tr>
<tr>
<td>IY59386</td>
<td>rsct.basic.rte:2.4.3.0</td>
<td>passed</td>
</tr>
</tbody>
</table>
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 ⚠️

```
ln -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
· 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 -igrep 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 ...
Specifying Cluster Configuration:

Just to remember !!!

Public, Private, and Virtual Host Name layout

<table>
<thead>
<tr>
<th>Node Name</th>
<th>IP</th>
<th>Node Name</th>
<th>IP</th>
<th>RAC Interconnect (Private Network)</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>10.3.25.81</td>
<td>node1_vip</td>
<td>10.3.25.181</td>
<td>Node1_rac</td>
</tr>
<tr>
<td>node2</td>
<td>10.3.25.82</td>
<td>node2_vip</td>
<td>10.3.25.182</td>
<td>Node1_rac</td>
</tr>
</tbody>
</table>

To 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

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 ...
At this stage, you must have already configured the shared storage. At least for:

- 1 Oracle Cluster Registry Disk  
- 1 Voting Disk  
- 2 Oracle Cluster Registry Disks (OCR Mirroring)  
- 3 Voting Disks (Voting Redundancy)

In our example, We will implement 1 Oracle Cluster Registry Disk, and 1 Voting Disk

You must know where to install these 2 or 5 files to be reachable by all nodes participating to the RAC cluster.

You have 2 implementations choices:

- Raw Disks with Oracle ASM (Automated Storage Management)
- Cluster files system with IBM GPFS

**Oracle Cluster Registry:**

Specify OCR Configuration, by selecting:

- **Normal Redundancy** (OCR mirroring by Oracle Clusterware)
- **External Redundancy** (No OCR mirroring 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/ocr1_disk
- /dev/ocr2_disk

If GPFS implementation is used, and NORMAL Redundancy selected, specify files system location as follow:

- /CRS1/ocr1_disk
- /CRS2/ocr2_disk

Specify the OCR location, this must be a shared location on the shared storage reachable from all nodes. And you must have the read/write 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.
In our case, we will specify “External Redundancy” as described below:

Specify Voting Disk Location

The Oracle Clusterware voting disk contains cluster membership information and arbitrates cluster ownership among the nodes of your cluster in the event of network failures. Specify a cluster file system file or a shared raw device that is accessible by the same name from all of the nodes in the cluster. The installer requires at least 20MB of free space for the voting disk that it creates.

- Normal Redundancy
  - Voting Disk Configuration
    - Voting Disk: /dev/vote1_disk
    - Additional Voting Disk Location: /dev/vote2_disk
    - Additional Voting Disk 2 Location: /dev/vote3_disk
  - If ASM implementation is used, specify raw disk location:
    - /dev/vote1_disk
    - /dev/vote2_disk
    - /dev/vote3_disk

- External Redundancy
  - Voting Disk Configuration
    - External Redundancy: /dev/vote1_disk
    - Additional Voting Disk Location: /dev/vote2_disk
    - Additional Voting Disk 2 Location: /dev/vote3_disk
  - If GPFS implementation is used, specify files system location:
    - /CRS1/vote1_disk
    - /CRS2/vote2_disk
    - /CRS2/vote3_disk

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/write 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 raiinstRoot.sh

3. On node1
   - Execute root.sh

4. On node2
   - Execute root.sh

**orainstRoot.sh:**

Execute the orainstRoot.sh on all nodes.
The file is located in $ORACLE_BASE/oraInventory (OraInventory home) on each nodes

**On Node1 as root, execute ./orainstRoot.sh**

[node1:root]/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 - >

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

[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:

```bash
{node1:root}/ - > su - oracle
{node1:root}/ ce /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 !!!

```
# rootconfig.sh for Oracle CRS homes
#
# This is run once per node during the Oracle CRS install.
# This script does the following:
# 1) Stop if any GSDs are running from 9.x oracle homes
# 2) Initialize new OCR device or upgrade the existing OCR device
# 3) Setup OCR for running CRS stack
# 4) Copy the CRS init script to init.d for init process to start
# 5) Start the CRS stack
# 6) Configure NodeApps if CRS is up and running on all nodes
```

Variables used by root.sh script, the values are the result of your inputs in the Oracle Clusterware Universal Installer.

You can check the values to see if there are OK.

```
SILENT=false
ORA_CRS_HOME=/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 deamon should be active on node 1.

Check for line “CSS is active on these nodes.

node1
node2

THEN On node2

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

When finished, CSS deamon should be active on node 1, 2.

You should have the following final result:

CSS is active on these nodes.

node1
node2

If CSS is not active on all nodes, or on one of the nodes, this means that you could have a problem with the network configuration, or the shared disks configuration for accessing OCR and Voting Disks.

- Check your network, shared disks configuration, and owner and access permissions (read/write) on OCR and Voting disks from each participating node. And execute again the root.sh script on node having the problem.

Check also as oracle user the following command from each node:

```
{node1:root}/ -> su - oracle
{node1:oracle}/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 he 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:

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.

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

<table>
<thead>
<tr>
<th></th>
<th>Public</th>
<th>VIP</th>
<th>RAC Interconnect</th>
<th>GPFS Interconnect (If GPFS used)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network card</td>
<td>en0</td>
<td>en0</td>
<td>en1</td>
<td>en2</td>
</tr>
<tr>
<td>On each node</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 ...

ORACLE
Just to remember !!!

Public, Private, and Virtual Host Name layout

<table>
<thead>
<tr>
<th>Public</th>
<th>VIP</th>
<th>RAC Interconnect (Private Network)</th>
</tr>
</thead>
<tbody>
<tr>
<td>en0</td>
<td>en0</td>
<td>en1</td>
</tr>
<tr>
<td>Node Name</td>
<td>IP</td>
<td>Node Name</td>
</tr>
<tr>
<td>node1</td>
<td>10.3.25.81</td>
<td>node1_vip</td>
</tr>
<tr>
<td>node2</td>
<td>10.3.25.82</td>
<td>node2_vip</td>
</tr>
</tbody>
</table>

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 fill 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=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 10.3.25.81 netmask 0xffffff00 broadcast 10.3.25.255
inet 10.3.25.181 netmask 0xffffff00 broadcast 10.3.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
en1:
flags=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 10.10.25.81 netmask 0xffffff00 broadcast 10.10.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
en2:
flags=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 20.20.25.81 netmask 0xffffff00 broadcast 20.20.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
lo0: flags=00001000<UP,BROADCAST,LOOPBACK,Running,SIMPLEX,NETWORK,GROUPRT,64BIT>
inet 127.0.0.1 netmask 0xffff0000 broadcast 127.255.255.255
inet6::1/128
```

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

**On node 2:**

```
(node2:root)/ -> ifconfig -a
```

```
en0:
flags=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 10.3.25.82 netmask 0xffffff00 broadcast 10.3.25.255
inet 10.3.25.182 netmask 0xffffff00 broadcast 10.3.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
en1:
flags=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 10.10.25.82 netmask 0xffffff00 broadcast 10.10.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
en2:
flags=10080863,80<UP,BROADCAST,NOTRAILERS,Running,SIMPLEX,MULTICAST,GROUPRT,64BIT,checksum_offload,CHAIN>
inet 20.20.25.82 netmask 0xffffff00 broadcast 20.20.25.255
tcp_sendspace 131072 tcp_recvspace 65536
```

```
lo0: flags=00001000<UP,BROADCAST,LOOPBACK,Running,SIMPLEX,NETWORK,GROUPRT,64BIT>
inet 127.0.0.1 netmask 0xffff0000 broadcast 127.255.255.255
inet6::1/128
```

```
{node2:root}/ ->
```
**Install Oracle 10g RAC Release 2 on IBM eServer pSeries with AIX 5L**

**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
export TEMP=/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/crs_d.bin...<snip>...
root  906694 1196094 0 17:35:57 pts/3 0:36 grep d.bin
```

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]
```

```
(node2: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

{o1/o10g/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

AS root user:

Export Oracle Cluster Registry content

{o1/o10g/crs/bin ->}su
root's Password:
{o1/o10g/crs/bin ->}ocrconfig -export /o10g/ocr_export.dmp1 -s online
{o1/o10g/crs/bin ->}ls -la /o10g/ocr_export.dmp
-rw-r--r-- 1 root system 106420 Jan 30 18:30 /o10g/ocr_export.dmp
{o1/o10g/crs/bin ->}

you can/must not edit/modify this exported file

View OCR automatic periodic backup managed by Oracle Clusterware

{o1/o10g/crs/bin ->}ocrconfig -showbackup
node1 2006/01/30 16:03:03 /o10g/crs/cdata/crs
node1 2006/01/30 12:03:00 /o10g/crs/cdata/crs
node1 2006/01/30 08:02:59 /o10g/crs/cdata/crs
node1 2006/01/29 00:02:51 /o10g/crs/cdata/crs
node1 2006/01/25 13:02:10 /o10g/crs/cdata/crs
{o1/o10g/crs/bin ->}
12.2 SOME USEFULL COMMANDS

As root user

Command to start/stop the CRS deamons:

```
/start /oh10g/crs/bin/crsctl start crs
to start the CRS
/stop /oh10g/crs/bin/crsctl stop crs
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 alias 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
- export TEMP=/tmp
- export TMPDIR=/tmp
```

With /tmp or other destination having enough free space, about 500Mb
IF AIX5L release 5.3 is used, do modify the file oraparam.ini, and cluster.ini in Disk1/installer

update entries AIX5200 to AIX5300 on both files, and execute:

	$/<cdrom_mount_point>/runInstaller

Or execute: ./runInstaller -ignoreSysPrereqs

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

To chek AIX maintenance level installed on each node:

	-> instfix -ilgrep ML

All files for 5.3.0.0_AIX_ML were found.
All files for 5300-01_AIX_ML were found.
All files for 5300-02_AIX_ML were found.
All files 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 than 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/Oralinventory" 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 ...
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

```bash
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     - 147:31 /oh10g/crs/bin/crsd.bin reboot
 oracle 352946  1  0 Jan  25     - 0:18 /oh10g/crs/bin/evmd.bin
 oracle 451600 303342  0 Jan  25     - 2:15 /oh10g/crs/bin/csdd.bin
 root  906694 196094  0 17:55:57 pts/3 0:30 grep d.bin
```

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
```

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, evmcheck 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
```

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 launch the database configuration assistant

![](image)

Select “Perform Typical Configuration”

Then click Next ...

14.1 LISTENER POST-CONFIGURATION TASK

Execute

```bash
(node1:oracle)/oh10g ->crs_stat -t
```

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Target</th>
<th>State</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora....E1.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.gsd application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....E2.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.gsd application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
</tbody>
</table>

AS root user:

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

root's Password:

```bash
{node1:oracle}/oh10g/crs/bin ->ocrconfig -export /oh10g/ocr_export2.dmp -s online
```

- `rw-r--r-- 1 root system 106420 Jan 30 18:30 /oh10g/ocr_export.dmp`

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 launch 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 appear 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  221090  1  59 Jan 25 - 7:47:31 /oh10g/crs/bin/crsd.bin crsbin
oracle  352345  1  0 Jan 25 - 2:35:18 /oh10g/crs/bin/crsd.bin
node1  3966094 1196094 0 10:35:57 pts/3 0:00 gkeep d.bin.
```

**Execute**

```
(node1:oracle)/oh10g ->crs_stat -t
```

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Target</th>
<th>State</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora....SM1.asm application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....E11snr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.gsd application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
</tbody>
</table>

**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, evmcheck 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
```
Check ASM daemons

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
```
<table>
<thead>
<tr>
<th>Node</th>
<th>Date/Time</th>
<th>Directory Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>node1</td>
<td>2006/01/30 08:02:59</td>
<td>/oh10g/crs/cdata/crs</td>
</tr>
<tr>
<td>node1</td>
<td>2006/01/29 00:02:51</td>
<td>/oh10g/crs/cdata/crs</td>
</tr>
<tr>
<td>node1</td>
<td>2006/01/25 13:02:10</td>
<td>/oh10g/crs/cdata/crs</td>
</tr>
<tr>
<td>node1:oracle</td>
<td></td>
<td>(node1:oracle)/oh10g/crs/bin -&gt;</td>
</tr>
</tbody>
</table>
15.2 SOME USEFULL COMMANDS

AS oracle user

Command to start/stop the ASM Instances:

For node 1

```bash
srvtl start asm –n node1
dsrtl stop asm –n node1
```
to start the ASM instance
to stop the ASM instance

For node 2

```bash
srvtl start asm –n node2
dsrtl stop asm –n node2
```
to start the ASM instance
to stop the ASM instance

To access an ASM instance with sqlplus

From node 1

```
sqlplus /nolog
connect /as sysdba
show sga
```

From node 2

```
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**

```
ORACLE_SID=ASMDB1

Node 1

Local Disks

• AIX5L MLXX
• $CRS_HOME
• $ORACLE_HOME
• Listener / Tnsnames

Shared Raw Disks

• OCR (Oracle Cluster Registry)
• Voting Disk

Shared ASM Disks

• Sppile
• Datafiles
• Redo Logs for each instance
• UNDO tablespace for each instance

Node 2

Local Disks

ORACLE_SID=ASMDB2

Private Network
```

DATABASE_NAME=ASMDB

```
Connect as Oracle unix user from first node, and **setup your DISPLAY**

Execute `dbca &` to launch the database configuration assistant

---

**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 Configuration Assistant, Step 6 of 16: Database Credentials

For security reasons, you must specify passwords for the following user accounts in the new database:

- Use the Same Password for All Accounts:
  - Password: [ ]
  - Confirm Password: [ ]

- Use Different Passwords:

<table>
<thead>
<tr>
<th>User Name</th>
<th>Password</th>
<th>Confirm Password</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYSTEM</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>DEFINIT</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>SYSAUX</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

Then click Next ...

Database Credentials:

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

Then click Next ...

Database Configuration Assistant, Step 7 of 15: Storage Options

Select the storage mechanism you would like to use for the database:

- Cluster File System
- Automatic Storage Management (ASM)

Automatic Storage Management (ASM) automatically manages database storage, and optimizes database layout for IO performance. To use this option you must either specify a set of disks to create an ASM disk group or specify an existing ASM disk group.

Storage Options:

- Choose Automatic Storage Management (ASM)

Then click Next ...

Database Configuration Assistant, Step 8 of 15: ASM Disk Groups

Select one or more disk groups to be used as storage for the database. You can choose to create a new disk group or select an existing disk group.

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**

```bash
(node1:oracle)/oh10g -gtcrs_stat -t
```

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Target</th>
<th>State</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora....B1.inst application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....B2.inst application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.ASMDB.db application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....SM1.asm application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....E1.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.gsd application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....SM2.asm application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora....E2.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.gsd application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
</tbody>
</table>

**CR health check**

```bash
(node1:oracle)/oh10g/crs/bin -gtcrsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy

(node2:oracle)/oh10g/crs/bin -gtcrsctl check crs
CSS appears healthy
CRS appears healthy
EVM appears healthy
```

**cssd, crsd, evmcheck health check**

```bash
(node1:oracle)/oh10g/crs/bin -gtcrsctl check cssd
CSS appears healthy

(node1:oracle)/oh10g/crs/bin -gtcrsctl check crsd
CRS appears healthy

(node1:oracle)/oh10g/crs/bin -gtcrsctl check evmd
EVM appears healthy

(node2:oracle)/oh10g/crs/bin -gtcrsctl check cssd
CSS appears healthy

(node2:oracle)/oh10g/crs/bin -gtcrsctl check crsd
CRS appears healthy

(node2:oracle)/oh10g/crs/bin -gtcrsctl 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_export4.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 useful commands

AS oracle user

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

From any nodes

```
svrctl start database -d ASMDB
svrctl stop database -d ASMDB
```

to start the Database instance
to stop the Database instance

For instance 1

```
svrctl start instance -d ASMDB -i ASMDB1
svrctl stop instance -d ASMDB -i ASMDB1
```

to start the DB... instance
to stop the DB... instance

For instance 2

```
svrctl start instance -d ASMDB -i ASMDB2
svrctl stop instance -d ASMDB -i ASMDB2
```

to start the DB... instance
to stop the DB... instance

To access an ASM instance with sqlplus

From node 1

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

To access a Database instance stored in ASM with sqlplus

From node 1

```
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 node, ASMDB1 and ASMDB2),
16.2 DATABASE CREATION ON GPFS

Architecture GPFS to implement

- **ORACLE_SID=GPFS1**
  - Node 1
  - Local Disks
  - • AIX5L MLXX
  - • SORACLE_CRS
  - • SORACLE_HOME
  - • Listener / Tnsnames

- **ORACLE_SID=GPFS2**
  - Node 2
  - Local Disks
  - • OCR (Oracle Cluster Registry)
  - • Voting Disk

- Shared File System
  - • Spfile
  - • Datafiles
  - • Redo Logs for each instance
  - • UNDO tablespace for each instance

- **DATABASE_NAME=GPFS**
  - Private Network

- • AIX5L MLXX
- • SORACLE_CRS
- • SORACLE_HOME
- • Listener / Tnsnames
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 ...
Installing Oracle 10g RAC Release 2 on IBM pSeries with AIX 5L
Database Configuration Assistant, Step 12 of 14: Initialization Parameters

Memory Settings:
- Default: 30%
- Manual: Enter Custom values

Shared Memory Management:
- Automatic
- Manual

SGA Size: 256 MB
PGA Size: 124 MB

Total Memory for Oracle:
- Includes 4GB of Process Size and the defaults for the empty parameters, if any.

Database Configuration Assistant, Step 13 of 14: Database Storage

File System Details:
- Group: 1
- File Size: 200,000,000 bytes
- File Name: redo01 redo02 redo03 redo04
- File Directory: /data/phys/OBNAME/

Database Configuration Assistant, Step 14 of 14: Creation Options

Creation Options:
- Create Database
- Generate Database Creation Scripts

Directory: /home/tom/asmenv/OBHomes

IBM
Installing Oracle 10g RAC Release 2 on IBM eServer pSeries with AIX SL
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:

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Target</th>
<th>State</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>ora....S1.inst application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....S2.inst application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.GPFS.db application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....E1.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora.node1.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node1</td>
<td></td>
</tr>
<tr>
<td>ora....E2.lsnr application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.ons application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
<tr>
<td>ora.node2.vip application</td>
<td>ONLINE</td>
<td>ONLINE</td>
<td>node2</td>
<td></td>
</tr>
</tbody>
</table>

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, evmcheckhealth 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

**AS root user:**

Export Oracle Cluster Registry content

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

You can/must not edit/modify this exported file

**View OCR automatic periodic backup managed by Oracle Clusterware**

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

<table>
<thead>
<tr>
<th>START</th>
<th>STOP</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>➔ emctl start dbconsole</td>
<td>➔ emctl stop dbconsole</td>
<td>➔ emctl status dbconsole</td>
</tr>
</tbody>
</table>

FOR RAC Node1

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

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
Install Oracle 10g RAC Release 2 on IBM pSeries with AIX 5L

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/rpdb1/user_default_1.dbf' size 2000M REUSE segment space management auto
SYSAUX DATAFILE '/u01/oracle/rpdb1/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:
1> CATALOG.SQL--creates the views of data dictionary tables and the
dynamic performance views
2> CATPROC.SQL--establishes the usage of PL/SQL functionality and
creates many of the PL/SQL Oracle supplied packages
3> 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)

```plaintext
# 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)

```plaintext
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:

```bash
# 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:

```bash
# 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

# 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:

```sql
# 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_GPFSS =
  (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 =

(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 =

(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
Subject: CRS and 10g Real Application Clusters
Type: BULLETIN
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) Interface
For more information on this, see Note 264847.1.

- The root.sh script at the end of the CRS installation starts the CRS stack.
If your CRS stack does not start, see Note 240001.1.

- Only one set of CRS daemons can be running per RAC node.
- On Unix, the CRS stack is run from entries in /etc/inittab with "respawn".
- If there is a network split (nodes loose communication with each other). One or more nodes may reboot automatically to prevent data corruption.
- The supported method to start CRS is booting the machine. MANUAL STARTUP OF THE CRS STACK IS NOT SUPPORTED UNTIL 10.1.0.4 OR HIGHER.
- The supported method to stop is shutdown the machine or use "init.crs stop".
- Killing CRS daemons is not supported unless you are removing the CRS installation via Note 239998.1 because flag files can become mismatched.
- For maintenance, go to single user mode at the OS.

Once the stack is started, you should be able to see all of the daemon processes with a ps -ef command:

```
[rac1]/u01/home/beta> ps -ef | grep crs

oracle  1363  999  0 11:23:21 ?  0:00 /u01/crs_home/bin/evmlogger.bin -o /u01
oracle  999  1  0 11:21:39 ?  0:01 /u01/crs_home/bin/evmd.bin
root    1003  1  0 11:21:39 ?  0:01 /u01/crs_home/bin/crsd.bin
oracle  1002  1  0 11:21:39 ?  0:01 /u01/crs_home/bin/ocssd.bin
```

CRS DAEMON FUNCTIONALITY
------------------------

Here is a short description of each of the CRS daemon processes:

**CRSD:**
- Engine for HA operation
- Manages 'application resources'
- Starts, stops, and fails 'application resources' over
- Spawns separate 'actions' to start/stop/check application resources
- Maintains configuration profiles in the OCR (Oracle Configuration Repository)
- Stores current known state in the OCR.
- Runs as root
- Is restarted automatically on failure

**OCSSD:**
- OCSSD is part of RAC and Single Instance with ASM
- Provides access to node membership
- Provides group services
- Provides basic cluster locking
- Integrates with existing vendor clusteware, when present
- Can also runs without integration to vendor clustware
- Runs as Oracle.
- Failure exit causes machine reboot.
- --- This is a feature to prevent data corruption in event of a split brain.

**EVMD:**
- Generates events when things happen
- Spawns a permanent child evmlogger
- Evmlogger, on demand, spawns children
- Scans callout directory and invokes callouts.
- Runs as Oracle.
- Restarted automatically on failure

CRS LOG DIRECTORIES
----------------------
When troubleshooting CRS problems, it is important to review the directories under the CRS Home.

$ORACLE_HOME/crs/log - This directory includes traces for CRS resources that are joining, leaving, restarting, and relocating as identified by CRS.

$ORACLE_HOME/crs/init - Any core dumps for the crsd.bin daemon should be written here. **Note 1812.1** could be used to debug these.

$ORACLE_HOME/crs/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.

$ORACLE_HOME/crs/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.

$ORACLE_HOME/evm/log - Log files for the evm and evmlogger daemons. Not used as often for debugging as the CRS and CSS directories.

$ORACLE_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.

$ORACLE_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 $ORACLE_HOME/bin
./crs_stat

NAME=ora.rac1.gsd
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.oem
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.ons
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac1.vip
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.gsd
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.oem
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.ons
TYPE=application
TARGET=ONLINE
STATE=ONLINE

NAME=ora.rac2.vip
TYPE=application
TARGET=ONLINE
STATE=ONLINE
There is also a script available to view CRS resources in a format that is easier to read. Just create a shell script with:

```
#!/usr/bin/ksh
#
# Sample 10g CRS resource status query script
#
# Description:
# - Returns formatted version of crs_stat -t, in tabular format, with the complete rsc names and filtering keywords
# - The argument, $RSC_KEY, is optional and if passed to the script, will limit the output to HA resources whose names match $RSC_KEY.
# Requirements:
# - $ORA_CRS_HOME should be set in your environment

RSC_KEY=$1
QSTAT=-u
AWK=/usr/xpg4/bin/awk  # if not available use /usr/bin/awk

# Table header:echo ""
$AWK \n'BEGIN { printf "%-45s %-10s %-18s\n", "HA Resource", "Target", "State"; printf "%-45s %-10s %-18s\n", "---", "-----", "-----"; }'

# Table body:
$ORA_CRS_HOME/bin/crs_stat $QSTAT | $AWK \n'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;}'

Example output:

[opcbsoll]/u01/home/usupport> ./crsstat
HA Resource Target State
---------- ------ ------
oracle.V10SN.V10SN1.insta ONLINE ONLINE on opcbsoll
oracle.V10SN.V10SN2.insta ONLINE ONLINE on opcbsoll
oracle.V10SN.db ONLINE ONLINE on opcbsoll
oracleopcbsoll.ASM1.asm ONLINE ONLINE on opcbsoll
oracleopcbsoll.LISTENER_OPCBSOLL.lsnr ONLINE ONLINE on opcbsoll
oracleopcbsoll.gsd ONLINE ONLINE on opcbsoll
oracleopcbsoll.ons ONLINE ONLINE on opcbsoll
oracleopcbsoll.vip ONLINE ONLINE on opcbsoll
oracleopcbsoll2.ASM2.asm ONLINE ONLINE on opcbsoll
oracleopcbsoll2.LISTENER_OPCBSOLL2.lsnr ONLINE ONLINE on opcbsoll
oracleopcbsoll2.gsd ONLINE ONLINE on opcbsoll
oracleopcbsoll2.ons ONLINE ONLINE on opcbsoll
oracleopcbsoll2.vip ONLINE ONLINE on opcbsoll
```
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
INSTALLED ORACLE 10 RAC RELEASE 2 ON IBM

**STOP CRS RESOURCES**

```
'rvctl stop database -d <database-name> [-o <stop-options>]
[-c <connect-string> | -q]
rvctl stop instance -d <database-name> -i <instance-name> [,<instance-name-list>]
[-o <stop-options>][-c <connect-string> | -q]
rvctl stop service -d <database-name> [-s <service-name>[,<service-name-list>]]
[-i <instance-name>][-c <connect-string> | -q] [-f]
rvctl stop nodeapps -n <node-name>
rvctl 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**

```
rvctl 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>]
rvctl add instance -d <name> -i <inst_name> -n <node_name>
rvtctl add service -d <name> -s <service_name> -r <preferred_list>
[-a <available_list>] [-P <TAF_policy>] [-u]
rvtctl add nodeapps -n <node_name> -o <oracle_home>
[-A <name|ip|netmask[/if1[|if2|...]]]
rvtctl add asm -n <node_name> -i <asm_inst_name> -o <oracle_home>
```

**OPTIONS:**

- **-A** vip range, node, and database, address specification. The format of address string is:
  
  [logical host name]/<VIP address>/<net mask[/<host interface1 | host interface2 ...]>[:<logical host name][/<VIP address>/<net mask[/<host interface1 | host interface2 ...]>]]
- **-a** for services, list of available instances, this list cannot include preferred instances
- **-m** domain name with the format “us.mydomain.com”
- **-n** node name that will support one or more instances
- **-o** $ORACLE_HOME to locate Oracle binaries
- **-P** for services, TAF preconnect policy - NONE, PRECONNECT
- **-r** for services, list of preferred instances, this list cannot include available instances.
- **-s** spfile name
- **-u** updates the preferred or available list for the service to support the specified instance. Only one instance may be specified with the -u switch. Instances that already support the service should not be included.

**EXAMPLES:**

Add a new node:
rvctl add nodeapps -n myclust-1 -o $ORACLE_HOME -A

Installing Oracle 10g RAC Release 2 on IBM server pSeries with AIX 5L
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

```
svrctl modify database -d <name> [-n <db_name>] [-o <ohome>] [-m <domain>]
[-p <spfile>] [-r {PRIMARY | PHYSICAL_STANDBY | LOGICAL_STANDBY}]
[-s <start_options>]
svrctl modify instance -d <database-name> -i <instance-name> -n <node-name>
svrctl modify instance -d <name> -i <inst_name> [-s <asm_inst_name>] [-r]
svrctl modify service -d <database-name> -s <service_name> -i <instance-name>
-t <instance-name> [-f]
svrctl modify service -d <database-name> -s <service_name> -i <instance-name>
-r [-f]
svrctl 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.
```
svrctl modify instance -d ORACLE -n myclust-4
```
Modify a service to execute on another node.
```
svrctl modify service -d ORACLE -s HOT_BATCH -i RAC01 -t RAC02
```
Modify an instance to be a preferred instance for a service.
```
svrctl modify service -d ORACLE -s HOT_BATCH -i RAC02 -r
```

RELOCATE SERVICES

```
svrctl relocate service -d <database-name> -s <service-name> [-i <instance-name>]
-t <instance-name> [-f]
```

EXAMPLES:

Relocate a service from one instance to another
```
svrctl 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
# 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. Choose one:

## 19.1 RSCT 2.3.4 (Provided with AIX 5.3) for All Implementation

<table>
<thead>
<tr>
<th>Fileset</th>
<th>Version</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rsct.basic.rte</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Basic Function</td>
</tr>
<tr>
<td>rsct.compat.basic.rte</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Event Management Basic</td>
</tr>
<tr>
<td>rsct.compat.clients.rte</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Event Management Client</td>
</tr>
<tr>
<td>rsct.core.auditrm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Audit Log Resource</td>
</tr>
<tr>
<td>rsct.core.erm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Event Response Resource</td>
</tr>
<tr>
<td>rsct.core.fsm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT File System Resource</td>
</tr>
<tr>
<td>rsct.core.gui</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Graphical User Interface</td>
</tr>
<tr>
<td>rsct.core.hostrm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Host Resource Manager</td>
</tr>
<tr>
<td>rsct.core.lprm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Least Privilege Resource</td>
</tr>
<tr>
<td>rsct.core.rmc</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Resource Monitoring and</td>
</tr>
<tr>
<td>rsct.core.sec</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Security</td>
</tr>
<tr>
<td>rsct.core.sensorrrm</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Sensor Resource Manager</td>
</tr>
<tr>
<td>rsct.core.utils</td>
<td>2.4.3.0</td>
<td>C</td>
<td>RSCT Registry</td>
</tr>
<tr>
<td>rsct.lapi msg. EN_US.rte</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT LAPI Msgs - U.S. English</td>
</tr>
<tr>
<td>rsct.lapi msg. en_US.rte</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT LAPI Msgs - U.S. English</td>
</tr>
<tr>
<td>rsct.lapi nam</td>
<td>2.4.2.0</td>
<td>C</td>
<td>RSCT LAPI NAM Files</td>
</tr>
<tr>
<td>rsct.lapi rt</td>
<td>2.4.2.0</td>
<td>C</td>
<td>RSCT LAPI Runtime</td>
</tr>
<tr>
<td>rsct.lapi samp</td>
<td>2.4.2.0</td>
<td>C</td>
<td>RSCT LAPI Samples</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.auditrm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.erm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT Event Response RM Msgs -</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.fsm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT File System RM Msgs -</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.gui</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT GUI Msgs - U.S.</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.hostrm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.lprm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT LPRM Msgs - U.S. English</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.rmc</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT RMC Msgs - U.S. English</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.sec</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT Security Msgs - U.S.</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.sensorrrm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.sr</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT Registry Msgs - U.S.</td>
</tr>
<tr>
<td>rsct.msg. EN_US.core.utils</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT Utilities Msgs - U.S.</td>
</tr>
<tr>
<td>rsct.msg.en_US.core.auditrm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg.en_US.core.erm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT Event Response RM Msgs -</td>
</tr>
<tr>
<td>rsct.msg.en_US.core.fsm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT File System RM Msgs -</td>
</tr>
<tr>
<td>rsct.msg.en_US.core.gui</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT GUI Msgs - U.S. English</td>
</tr>
<tr>
<td>rsct.msg.en_US.core.com</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg.en_US.core.hostrm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rsct.msg.en_US.core.lprm</td>
<td>2.4.0.0</td>
<td>C</td>
<td>RSCT LPRM Msgs - U.S. English</td>
</tr>
</tbody>
</table>
### 19.2 For GPFS implementation

<table>
<thead>
<tr>
<th>Package</th>
<th>Version</th>
<th>Country</th>
<th>Font</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gpfs.base</td>
<td>2.3.0.5</td>
<td>C</td>
<td>F</td>
<td>GPFS File Manager</td>
</tr>
<tr>
<td>gpfs.docs.data</td>
<td>2.3.0.5</td>
<td>C</td>
<td>F</td>
<td>GPFS Server Manpages and</td>
</tr>
<tr>
<td>gpfs.msg.en_US</td>
<td>2.3.0.4</td>
<td>C</td>
<td>F</td>
<td>GPFS Server Messages - U.S.</td>
</tr>
</tbody>
</table>