Oracle Database 10g Real Application Clusters R2 (RAC10g R2) on HP-UX Installation Cookbook

Authors: *Rebecca Schlecht (HP), Rainer Marekwia (Oracle)*
EMEA HP/Oracle Cooperative Technology Center (CTC)
http://www.hporaclectc.com
Date: *26th January 2007*

This document is based on our experiences, it is not an official HP/Oracle documentation. We're constantly updating this installation cookbook, therefore please check for the latest version of this cookbook on our HP/Oracle CTC web page at https://www.hporaclectc.com/cug/assets/10gR2RAChp.htm (pdf version).

If you have any comments or suggestions, please send us an email with your feedback! In case of issues during your installation, please also report this problem to HP and/or Oracle support.

Contents:

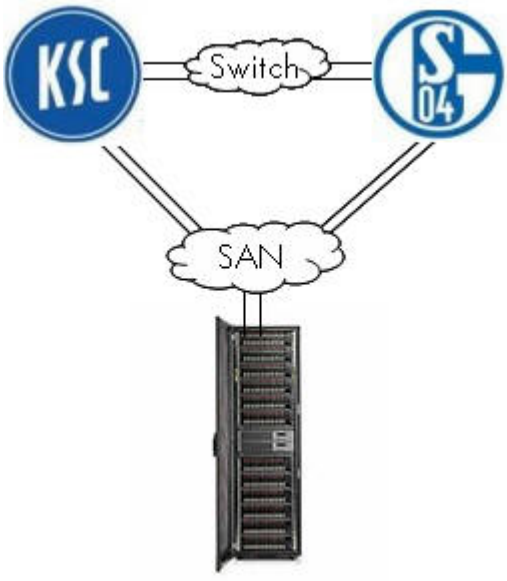
---

1. Aim of this document

This document is intended to provide help installing Oracle Real Application Clusters 10g Release 2 on HP servers running HP-UX operating system. This paper covers both Integrity and PA-RISC platform.

All information here is based on practical experiences.

All described scenarios are based on a 2 node cluster, node1 referred to as 'ksc' and node2 as 'schalke'.



In this paper, we use the following logic:

`ksc# <command>`	= command needs to be issued as root from node ksc
`schalke$ <command>`	= command needs to be issued as oracle from node schalke
`ksc/schalke# <command>`	= command needs to be issued as root from both nodes ksc + schalke

and so on.

This document should be used in conjunction with the following Oracle documentation:

B25292-02 Oracle® Database Release Notes 10g Release 2 (10.2) for HP-UX Itanium (pdf html)
B19067-04 Oracle® Database Release Notes 10g Release 2 (10.2) for HP-UX PA-RISC (64-Bit) (pdf html)
B14202-04 Oracle Clusterware and Oracle Real Application Clusters Installation and Configuration Guide for hp HP-U

It also includes material from HP Serviceguard + RAC10g papers written by ACSL labs which are available HP internally at http://haweb.cup.hp.com/ATC/Web/Whitepapers/default.htm.

2. Key New Features for RAC10g on HP-UX

Oracle Clusterware

New with RAC 10g, Oracle includes its own Clusterware and package management solution with the database product. The Oracle Clusterware consists of

- Oracle Cluster Synchronization Services (CSS) to provide cluster management functionality
- Oracle Cluster Ready Services (CRS) support services and workload management and help to maintain the continuous availability of the services. CRS also manages resources such as the virtual IP (VIP) address for the node and the global services daemon.
- Event Management (EVM) publishes events generated by CRS

This Oracle Clusterware is available on all various Oracle RAC platforms and based on the HP TruCluster product which Oracle licensed a couple of years ago.

Customers can now deploy Oracle RAC clusters without any additional 3rd party clusterware products such as SG/SGeRAC. However, customers might want to continue to use SG/SGeRAC for the cluster management (e.g. to make your complete cluster high available including 3rd party application, interconnect, etc.). In this case Oracle Clusterware interacts with the SG/SGeRAC to coordinate cluster membership information.

New Features for Oracle Clusterware with RAC 10g R2:

- Oracle 10g R2 comes with new Cluster Verification Utility that you can use to check whether or not your cluster is properly configured, to avoid installation failures, and to avoid database creation failures.
- With 10g R2, Oracle Clusterware provides the possibility to mirror the Oracle Cluster Registry (OCR) file, enhancing cluster reliability.
- With 10g R2, CSS has been modified to allow you to configure multiple voting disks. In RAC10g R1, you could configure only one voting disk. By enabling multiple voting disk configuration, the redundant voting disks allow you to configure a RAC database with multiple voting disks on independent shared physical disks.
- With Oracle 10g R2, in addition, while continuing to be required for RAC databases, Oracle Clusterware is also available for use with single-instance databases and applications that you deploy on clusters. The API libraries required for use with single-instance databases are provided with the Oracle Client installation media.

Oracle Automatic Storage Management

Oracle Automatic Storage Management (ASM) is a new feature that has be introduced in Oracle Database 10g to simplify the storage of Oracle data. ASM virtualizes the database storage into disk groups. The DBA is able to manage a small set of disk groups and ASM automates the placement of the database files within those disk groups.

In summary ASM does provide the following functionality:

- Manages groups of disks, called disk groups.
- Provides three mirroring options for protection against disk failure: none, two-way, and three-way mirroring.
- Spreads data evenly across all available storage resources to optimize performance and utilization.
- Enables the DBA to change the storage configuration without having to take the database offline.
- Automatically rebalances files across the disk group after disks have been added or dropped.

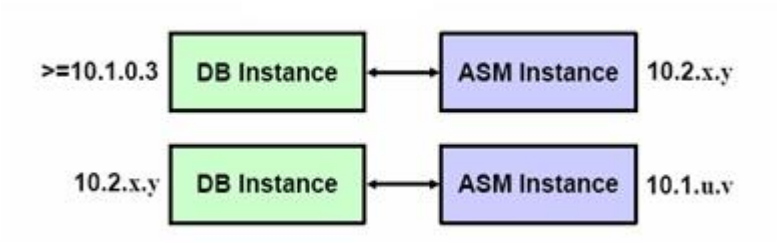
New Features for Oracle ASM with 10g R2:

- ASM Command-Line Utility for ASM file administration:

```
$ asmcmd help
```

- Oracle 10g R2 supports installation of Automatic Storage Management in a separate ASM home directory.
- Supports interoperability for all versions of ASM and Database instances starting

with RAC10g R1. This allows the ASM instance and DB instance to be upgraded independently:



- ASM migration utility with Enterprise Manager Grid Control GUI

HP Serviceguard Cluster File System for Oracle RAC

In September 2005, HP announced the availability of the new HP Serviceguard Storage Management Suite that offers enhanced database, cluster, and performance management capabilities for HP-UX 11i environments by integrating HP Serviceguard and Symantec VERITAS Storage Foundation. This new product suite is ideally suited to customers who need the highest levels of availability and superior Oracle database performance or who have an application that would benefit from a clustered file system.

The HP Serviceguard Cluster File System for Oracle RAC Suite includes the following technologies from Symantec VERITAS Storage Foundation:

- Cluster File System (CFS)—provides excellent I/O performance and simplifies the installation and ongoing management of a RAC database
- Advanced volume management and file system (AVMFS) capabilities—offers dynamic multipathing, database tablespace growth, and hot relocation of failed redundant storage. It also provides a variety of online options, including storage reconfiguration and volume and file system creation and resizing.
- Oracle Disk Manager (ODM)—delivers almost raw performance running direct I/O by caching frequently accessed data
- Quality of storage service (QoSS)—enables administrators to set policies that segment company data based on various characteristics and assign the data to appropriate classes of storage over time
- FlashSnap—helps database administrators easily establish a database clone, a duplicate database on a secondary host for off-host processing

This HP Serviceguard Storage Management Suite is offered and supported directly from HP for a single point of contact for all your support needs.

HP Product Number: T2777BA (HP Serviceguard CFS for RAC LTU).

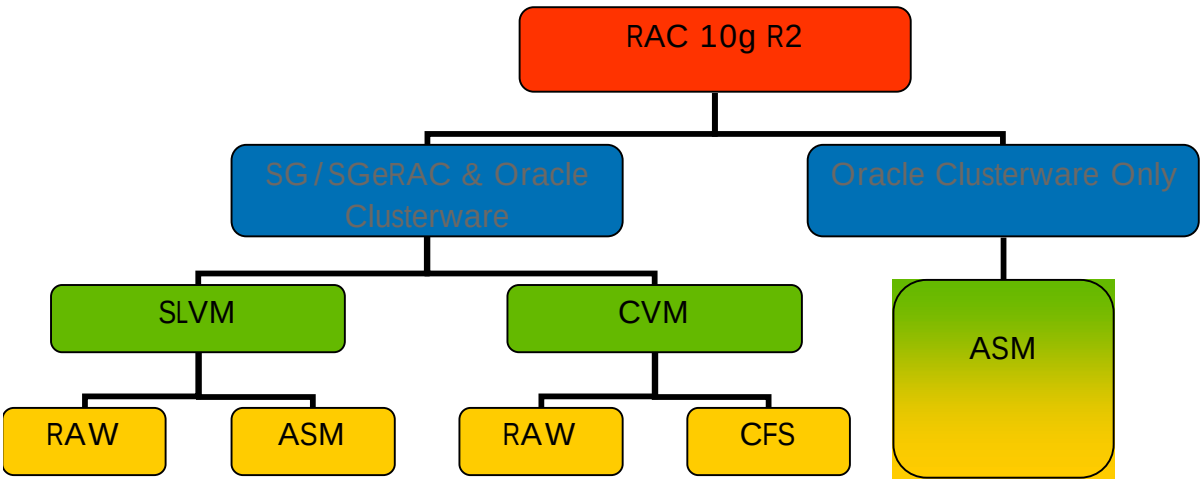
3. Supported Configurations with RAC10g on HP-UX

Customers do have a variety of choices with regards to the installation and set-up of Oracle Real Application Clusters 10g on the HP-UX platform.

First customers need to make a decision with regards to the underlying cluster software. Customers have the possibility to deploy their RAC cluster only with Oracle Clusterware. Alternatively, customers might want to continue to use HP Serviceguard & HP Serviceguard Extension for RAC (SGeRAC) for the cluster management. In this case Oracle’s CSS interacts with HP SG/SGeRAC to coordinate cluster membership information.

For storage management, customers have the choice to use Oracle ASM, HP's Cluster File Systems or RAW Devices.
Please note, for RAC with Standard Edition installations, Oracle mandates that the Oracle data must be placed under ASM control.

The figure below illustrates the supported configurations with Oracle RAC10gR2 on HP-UX.



The following table shows the storage options supported for storing Oracle Clusterware files, Oracle database files, and Oracle database recovery files. Oracle database files include data files, control files, redo log files, the server parameter file, and the password file. Oracle Clusterware files include Oracle Cluster Registry (OCR) and Voting disk. Oracle Recovery files include archive log files.

Storage Option	Clusterware	Database	Recovery
Automatic Storage Management	No	Yes	Yes
Shared raw logical volumes (requires SGeRAC)	Yes	Yes	No
Shared raw disk devices as presented to hosts	Yes	Yes	No
Shared raw partitions (only HP Integrity, no PA-Risc)	Yes	Yes	No
CFS	Yes	Yes	Yes

4. General System Installation Requirements

4.1 Hardware Requirements

- at least 1GB of physical RAM. Use the following command to verify the amount of memory installed on your system:

  ```
  # /usr/contrib/bin/machinfo | grep -i Memory
  ```
  or
  ```
  # /usr/sbin/dmesg | grep "Physical:"
  ```

- Swap space equivalent to the multiple of the available RAM, as indicated here:
  - If RAM between 1GB and 2GB, then swap space required is 1.5 times the size of RAM
  - If RAM > 2GB, then swap space required is equal to the size of RAM

Use the following command to determine the amount of swap space installed on your system:

```
# /usr/sbin/swapinfo -a
```

- 400 MB of disk space in the /tmp directory. To determine the amount of disk space available in the /tmp directory, enter the following command:

  ```
  # bdf /tmp
  ```

  If there is less than 400 MB of disk space available in the /tmp directory extend the file system or set the TEMP and TMPDIR environment variables when setting the oracle user's environment. This environment variables can be used to override /tmp.:

  ```
  $ export TEMP=/directory
  $ export TMPDIR=/directory
  ```

- 4 GB of disk space for the Oracle software. You can determine the amount of free disk space on the system using

  ```
  # bdf -k
  ```

- 1.2 GB of disk space for a preconfigured database that uses file system storage (optional)
- Operating System: HP-UX 11.23 (Itanium2), 11.23 (PA-RISC), 11.11 (PA-RISC). To determine if you have a 64-bit configuration enter the following command:

  ```
  # /bin/getconf KERNEL_BITS
  ```

  To determine which version of HP-UX is installed, enter the following command:

  ```
  # uname -a
  ```

- Asnyc I/O is required for Oracle on RAW devices and configured on HP-UX 11.23 by default. You can check if you have the following file:

  ```
  # ll /dev/async
  # crw-rw-rw- 1 bin bin 101 0x000000 Jun 9 09:38 /dev/async
  ```

- If you want to use Oracle on RAW devices and Async I/O is not configured, then
  - Create the /dev/async character device

    ```
    # /sbin/mknod /dev/async c 101 0x0
    # chown oracle:dba /dev/async
    # chmod 660 /dev/async
    ```

  - Configure the async driver in the kernel using SAM

    ```
    => Kernel Configuration
    => Kernel
    => the driver is called 'asyncdsk'
    Generate new kernel
    Reboot
    ```

  - Set HP-UX kernel parameter `max_async_ports` using SAM. `max_async_ports` limits the maximum number of processes that can concurrently use `/dev/async`. Set this parameter to the sum of 'processes' from init.ora + number of background processes. If `max_async_ports` is reached, subsequent processes will use synchronous i/o.
  - Set HP-UX kernel parameter `aio_max_ops` using SAM. `aio_max_ops` limits the maximum number of asynchronous i/o operations that can be queued at any time. Set this parameter to the default value (2048), and monitor over time using glance
- For PL/SQL native compilation, Pro*C/C++, Oracle Call Interface, Oracle C++ Call Interface, Oracle XML Developer's Kit (XDK):
  - HP-UX 11i v2 (11.23):
    - HP C/ANSI C Compiler (A.06.00): C-ANSI-C
    - HP aC++ Compiler (C.06.00): ACXX

  To determine the version, enter the following command:

  ```
  # cc -V
  ```

- To allow you to successfully relink Oracle products after installing this software, please ensure that the following symbolic links have been created (HP Doc-Id KBRC00003627):

```
# cd /usr/lib
# ln -s /usr/lib/libX11.3 libX11.sl
# ln -s /usr/lib/libXIE.2 libXIE.sl
# ln -s /usr/lib/libXext.3 libXext.sl
# ln -s /usr/lib/libXhp11.3 libXhp11.sl
# ln -s /usr/lib/libXi.3 libXi.sl
# ln -s /usr/lib/libXm.4 libXm.sl
# ln -s /usr/lib/libXp.2 libXp.sl
# ln -s /usr/lib/libXt.3 libXt.sl
# ln -s /usr/lib/libXtst.2 libXtst.sl
```

- Ensure that each member node of the cluster is set (as closely as possible) to the same date and time. Oracle strongly recommends using the Network Time Protocol feature of most operating systems for this purpose, with all nodes using the same reference Network Time Protocol server.

4.2 Network Requirements

You need the following IP addresses per node to build a RAC10g cluster:

- Public interface that will be used for client communication
- Virtual IP address (VIP) that will be bind by Oracle Clusterware to the public interface *(Why having this VIP? Well, clients will use this VIP addresses/names to access the RAC database. If a node or interconnect fails, then the affected VIP is relocated to the surviving instance, enabling fast notification of the failure to the clients connecting through that VIP -> prevents TCP/IP timeout!)*
- Private interface that will be used for inter-cluster traffic. There are four major categories of inter-cluster traffic:
  - SG-HB= Serviceguard heartbeat and communications traffic. This is supported over single or multiple subnet networks.
  - CSS-HB = Oracle CSS heartbeat traffic and communications traffic for Oracle Clusterware. CSS-HB uses a single logical connection over a single subnet network.
  - RAC-IC = RAC instance peer to peer traffic and communications for Global Cache Service (GCS) and Global Enqueue Service (GES), formally Cache Fusion (CF) and Distributed Lock Manager (DLM).
  - GAB/LLT (only when using CFS/CVM) = Symantec cluster heartbeat and communications traffic. GAB/LLT communicates over link level protocol (DLPI) and is supported over Serviceguard heartbeat subnet networks, including primary and standby links. GAB/LLT is not supported over APA or virtual LANs (VLAN).

When configuring these networks, please consider:

- The public and private interface names associated with the network adapters for each network should be the same on all nodes, e.g. lan0 for private interconnect and lan1 for public interconnect. If this is not the case, you can use the ioinit command to map the LAN interfaces to new device instances:
  - Write down the hardware path that you want to use:

```
# lanscan
Hardware Station Crd Hdw Net-Interface NM MAC HP-DLPI DLPI
Path Address In# State NamePPA ID Type Support Mjr#
1/0/8/1/0/6/0 0x000F203C346C 1 UP lan1 snap1 1 ETHER Yes 119
1/0/10/1/0 0x00306EF48297 2 UP lan2 snap2 2 ETHER Yes 119C
```

- Create a new ascii file with the following syntax:

```
Hardware_Path Device_Group New_Device_Instance_Number
```

Example:

```
# vi newio
1/0/8/1/0/6/0 lan 8
1/0/10/1/0 lan 9
```

Please note that you have to choose a device instance number that is currently not in use.

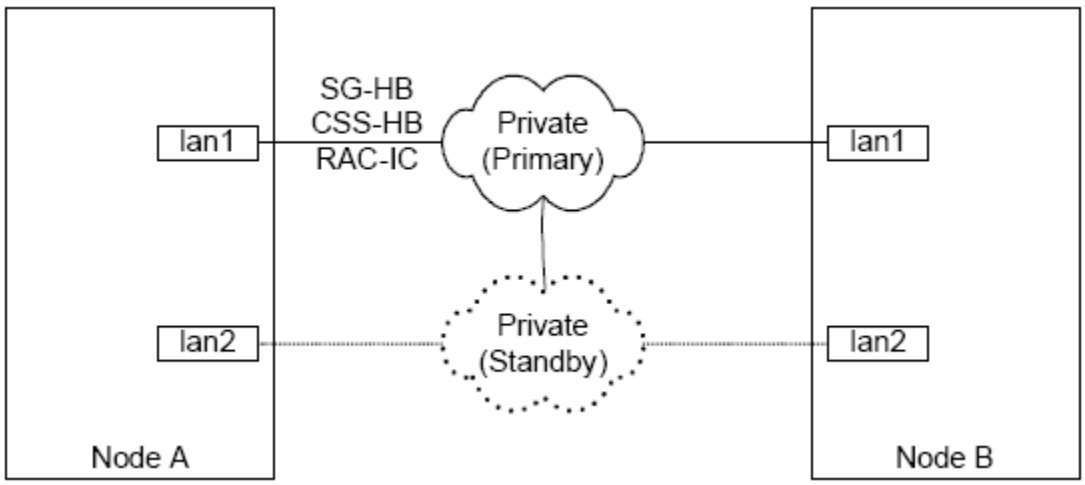
- Activate this configuration with the following command (-r option will issue a reboot):

```
# ioinit -f /root/newio -r
```

- When the system is up again, check new configuration:

```
# lanscan
Hardware Station Crd Hdw Net-Interface NM MAC HP-DLPI DLPI
Path Address In# State NamePPA ID Type Support Mjr#
1/0/8/1/0/6/0 0x000F203C346C 1 UP lan8 snap8 1 ETHER Yes 119
1/0/10/1/0 0x00306EF48297 2 UP lan9 snap9 2 ETHER Yes 119
```

- For the public network,
  - each network adapter must support TCP/IP.
- For the private network,
  - this must be configured in the /etc/hosts file on each node to associate private network names with private IP addresses.
  - the interconnect must support UDP as this is the default interconnect protocol for cache fusion, and TCP is the interconnect protocol for Oracle Clusterware.
  - Gigabit Ethernet or better is recommended, Hyperfabric is not supported any longer!
  - Crossover cables are not supported for the cluster interconnect; switch is mandatory for production implementation, even for only 2 nodes architecture.
  - It is preferred to have all interconnect traffics (SG-HB, CSS-HB, RAC-IC, opt. GAB/LLT) for cluster communications to go on a single heartbeat network that is redundant so that Serviceguard will monitor the network and resolve interconnect failures by cluster reconfiguration:

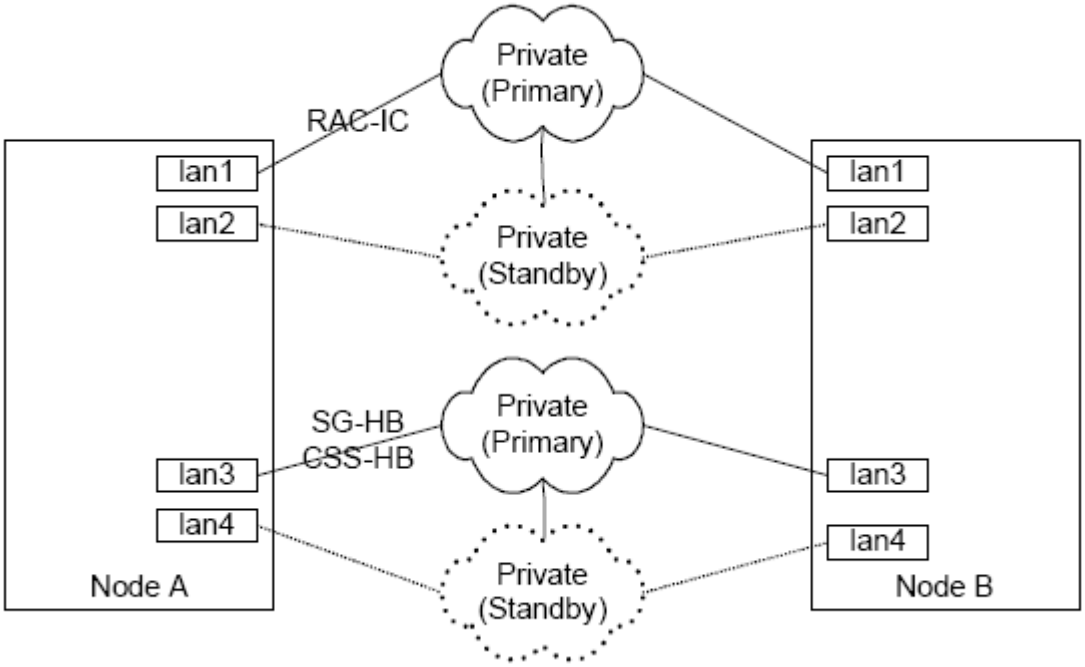


As illustrated in this picture, the primary and standby pair protects against single failure. Serviceguard monitors the network and performs local LAN failover if the primary fails. The local LAN failover is transparent to CSS-HB and RAC-IC. When both primary and standby fails, Serviceguard resolves the interconnect failure by performing a cluster reconfiguration. After Serviceguard completes its reconfiguration, SGeRAC notifies CSS and CSS updates RAC.

  - Please note that CSS-HB timeout default is 30sec for clusters without Serviceguard and 600 for clusters with Serviceguard. This ensures that Serviceguard will be first to recognize any failures and to initiate cluster reformation activities. (See Oracle Metalink

Note 294430.1 "CSS Timeout Computation in RAC 10g (10g Release 1 and 10g Release 2)")

- However, in some cases it might not be possible to place all interconnect traffic on the same network. For example if RAC-IC traffic is very high, so that a separate network for RAC-IC may be needed.



As illustrated in this picture, each primary and standby pair protects against single failure. The SG-HB and CSS-HB are placed on the same private network so that all heartbeat traffic remains on the same network. SG-HB is used to resolve interconnect failure where both primary and standby failed of the heartbeat network. Where there is a concern when both lan1 and lan2 failed, Serviceguard supports multiple standby adapters to increase availability. Additionally, Serviceguard packages can be configured with a subnet dependency on the RAC-IC network so that if both lan1 and lan2 failed, the Serviceguard package can request halting the RAC instance on the node where the interconnect failure is detected.

- For cluster without HP Serviceguard, you can use HP Auto Port Aggregation (APA) to increase reliability for public and private network adapters.

- For the virtual IP (VIP) address,
  - this must be on the same subnet as the public interface
  - this must be registered in DNS or maintained in /etc/hosts with the associated network name.
  - this Oracle VIP feature works at a low level with the device files for the network interface cards, and as a result might clash with any other SG Relocatable IP addresses also configured for the same public NIC. Therefore, it has not been supported to configure the public NIC used for Oracle VIP also for any other SG Relocatable IP address.
    - This issue has been addressed with Oracle bug fix #4699597 which ensures that Oracle VIP starts with logical interface number 801 (ie. lan1:801) so that there will not be any conflict with SG's Relocatable IP's.
    - This Oracle bug fix #4699597 is already available for 10.2.0.2 HP-UX Integrity and will be available for PA-RISC with 10.2.0.3.
  - See Oracle Metalink Note 296874.1 "Configuring the HP-UX Operating System for the Oracle 10g VIP")

Useful network commands:

```
# lanscan              # Determines the number of LAN interfaces on each node
# netstat -in          # Displays information for all network interfaces such
as IP address, state, etc.
# ifconfig lanX        # Displays current configuration for a specific
interface
(Config File: /etc/rc.config.d/netconf)
```

4.3 Required HP-UX Patches

HP-UX 11.23 (Integrity & PA-RISC):

- HP-UX B.11.23.0409 or later
- Patch Bundle for HP-UX 11i V2: BUNDLE11i_B.11.23.0409.3 (Note: patch bundle BUNDLE11i_B.11.23.0408.1 (Aug/2004) is a prerequisite for installing BUNDLE11i_B.11.23.0409.3)
- Quality Pack Bundle:
  - Latest patch bundle: Quality Pack Patches for HP-UX 11i v2, May 2005
- HP-UX 11.23 Patches:
  - PHCO_32426 Reboot(1M) cumulative patch
  - PHCO_34208 11.23 cumulative SAM patch [replaces PHCO_31820]
  - PHCO_34195 11.23 kernel configuration commands patch [replaces PHCO_33385]
  - PHCO_35048 11.23 libsec cumulative patch
  - PHKL_32646 wsio.h header file patch
  - PHKL_33025 11.23 file system tunables cumulative patch
  - PHKL_34907 Message Signaled Interrupts (MSI and MSI-X) [replaces PHKL_32632,PHKL_33807,PHKL_34430]
  - PHKL_34479 WSIO (IO) subsystem MSI/MSI-X/WC Patch [replaces PHKL_32645]
  - PHKL_35229 VM Copy on write data corruption fix [replaces PHKL_33552,PHKL_33563,PHKL_34596]
  - PHNE_35182 11.23 cumulative ARPA Transport patch
  - PHSS_34859 11.23 Integrity Unwind Library [replaces PHSS_31851,PHSS_34043]
  - PHSS_34858 11.23 linker + fdp cumulative patch [replaces PHSS_34040,PHSS_33275,PHSS_31849,PHSS_34440]
  - PHSS_34444 11.23 assembler patch [replaces PHSS_31850,PHSS_34044]
  - PHSS_34445 11.23 milli cumulative patch [replaces PHSS_31854,PHSS_34045]
  - PHSS_34853 11.23 Math Library Cumulative Patch [replaces PHSS_33276,PHSS_34042]
  - PHNE_35182 11.23 cumulative ARPA Transport patch [replaces PHNE_34671]
- ANSI + C++ patches:
  - PHSS_32511 11.23 HP aC++ Compiler (A.03.63)
  - PHSS_32512 11.23 ANSI C compiler B.11.11.12 cumulative patch
  - PHSS_32513 11.23 +O4/PBO Compiler B.11.11.12 cumulative patch
  - PHSS_35055 11.23 aC++ Runtime [replaces PHSS_31855,PHSS_34041,PHSS_31852]
- JDK patches:
  - PHCO_34944 11.23 pthread library cumulative patch [replaces PHCO_31553,PHCO_33675,PHCO_34718]
  - PHSS_35045 11.23 Aries cumulative patch [replaces PHSS_32213,PHSS_34201]
  - PHKL_31500 11.23 Sept04 base patch
  - check http://www.hp.com/products1/unix/java/patches/index.html for additional patches that may be required by JDK.

- Serviceguard 11.17 and OS Patches (optional, only if you want to use Serviceguard):
  - PHCO_32426 11.23 reboot(1M) cumulative patch
  - PHCO_35048 11.23 libsec cumulative patch [replaces PHCO_34740]
  - PHSS_33838 11.23 Serviceguard eRAC A.11.17.00
  - PHSS_33839 11.23 COM B.04.00.00
  - PHSS_35371 11.23 Serviceguard A.11.17.00 [replaces PHSS_33840]
  - PHKL_34213 11.23 vPars CPU migr, cumulative shutdown patch
  - PHKL_35420 11.23 Overtemp shutdown / Serviceguard failover

- LVM patches:
  - PHCO_35063 11.23 LVM commands patch; required patch to enable the Single Node Online Volume Reconfiguration (SNOR) functionality [replaces PHCO_34036,PHCO_34421]
  - PHKL_34094 LVM Cumulative Patch [replaces PHKL_34094]

- CFS/CVM/VxVM 4.1 patches:
  - PHCO_33080 11.23 VERITAS Enterprise Administrator Srvc Patch [replaces PHCO_33080]
  - PHCO_33081 11.23 VERITAS Enterprise Administrator Patch
  - PHCO_33082 11.23 VERITAS Enterprise Administrator Srvc Patch
  - PHCO_33522 11.23 VxFS Manpage Cumulative patch 1 SMS Bundle
  - PHCO_33691 11.23 FS Mgmt Srvc Provider Patch 1 SMS Bundle
  - PHCO_35431 11.23 VxFS 4.1 Command Cumulative patch 4 [replaces PHCO_34273]
  - PHCO_35476 VxVM 4.1 Command Patch 03 [replaces PHCO_33509, PHCO_34811]
  - PHCO_35518 11.23 VERITAS VM Provider 4.1 Patch 03 [replaces PHCO_34038, PHCO_35465]
  - PHKL_33510 11.23 VxVM 4.1 Kernel Patch 01 SMS Bundle [replaces PHKL_33510]
  - PHKL_33566 11.23 GLM Kernel cumulative patch 1 SMS Bundle
  - PHKL_33620 11.23 GMS Kernel cumulative patch 1 SMS Bundle
  - PHKL_35229 11.23 VM mmap(2), madvise(2) and msync(2) fix [replaces PHKL_34596]
  - PHKL_35334 11.23 ODM Kernel cumulative patch 2 SMS Bundle [replaces PHKL_34475]
  - PHKL_35430 11.23 VxFS 4.1 Kernel Cumulative patch 5 [replaces PHKL_34274, PHKL_35042]
  - PHKL_35477 11.23 VxVM 4.1 Kernel Patch 03 [replaces PHKL_34812]
  - PHKL_34741 11.23 VxFEN Kernel cumulative patch 1 SMS Bundle (Required to support 8 node clusters with CVM 4.1 or CFS 4.1)
  - PHNE_34664 11.23 GAB cumulative patch 2 SMS Bundle [replaces PHNE_33612]
  - PHNE_33723 11.23 LLT Command cumulative patch 1 SMS Bundle
  - PHNE_35353 11.23 LLT Kernel cumulative patch 3 SMS Bundle [replaces PHNE_33611, PHNE_34569]

- C and C++ patches for PL/SQL native compilation, Pro*C/C++, Oracle Call Interface, Oracle C++ Call Interface, Oracle XML Developer's Kit (XDK):
  - PHSS_33277 11.23 HP C Compiler (A.06.02)
  - PHSS_33278 11.23 aC++ Compiler (A.06.02)
  - PHSS_33279 11.23 u2comp/be/plugin library patch

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

- HP provides patch bundles at http://www.software.hp.com/SUPPORT_PLUS
- To determine whether the HP-UX 11i Quality Pack is installed:

```
# /usr/sbin/swlist -l bundle | grep GOLD
```

- Individual patches can be downloaded from http://itresourcecenter.hp.com/
- To determine which operating system patches are installed, enter the following command:

```
# /usr/sbin/swlist -l patch
```

- To determine if a specific operating system patch has been installed, enter the following command:

```
# /usr/sbin/swlist -l patch <patch_number>
```

- To determine which operating system bundles are installed, enter the following command:

```
# /usr/sbin/swlist -l bundle
```

4.4 Kernel Parameter Settings

Verify that the kernel parameters shown in the following table are set either to the formula shown, or to values greater than or equal to the recommended value shown. If the current value for any parameter is higher than the value listed in this table, do not change the value of that parameter. Please check also our HP-UX kernel configuration for Oracle databases for more details and for the latest recommendations.

You can modify the kernel settings either by using SAM or by using the kctune command line utility (kmtune on PA-RISC).

`# kctune > /tmp/kctune.log` (lists all current kernel settings)

`# kctune tunable>=value` The tunable's value will be set to value, unless it is already greater

`# kctune -D > /tmp/kctune.log` (Restricts output to only those parameters which have changes being held until next boot)

Parameter	Recommended Formula or Value
nproc	4096
ksi_alloc_max	(nproc*8)
max_thread_proc	1024
maxdsiz	1073741824 (1 GB)
maxdsiz_64bit	2147483648 (2 GB)
maxssiz	134217728 (128 MB)
maxssiz_64bit	1073741824 (1 GB)
maxswapchunks or swchunk (not used >= HP-UX 11iv2)	16384
maxuprc	((nproc*9)/10)
msgmap	(msgmni+2)
msgmni	nproc
msgseg	(nproc*4); at least 32767
msgtql	nproc
ncsize	(ninode+vx_ncsize); for >=HP-UX 11.23 use (ninode+1024)
nfile	(15*nproc+2048); for Oracle installations with a high number of data files this might not be enough, then use (number od Oracle processes)*(number of Oracle data files) + 2048
nflocks	nproc
ninode	(8*nproc+2048)
nkthread	(((nproc*7)/4)+16)
semmap	(semmni+2)
semmni	(nproc*2)
semmns	(semmni*2)
semmnu	(nproc-4)
semvmx	32767
shmmax	The size of physical memory 1073741824, whichever is greater. **Note:** To avoid performance degradation, the value should be greater than or equal to the size of the SGA.
shmmni	512
shmseg	120

swchunk	4096 (up to 65536 for large RAM)
vps_ceiling	64 (up to 16384 = 16MB for large SGA)

5. Create the Oracle User

- Log in as the root user
- Create database groups on each node. The group ids must be unique. The id used here are just examples, you can use any group id not used on any of the cluster nodes.
  - the OSDBA group, typically dba:
    ```
    ksc/schalke# /usr/sbin/groupadd -g 201 dba
    ```
  - the optional ORAINVENTORY group, typically oinstall; this group owns the Oracle inventory, which is a catalog of all Oracle software installed on the system.
    ```
    ksc/schalke# /usr/sbin/groupadd -g 200 oinstall
    ```
- Create the Oracle software user on each node. The user id must be unique. The user id used below is just an example, you can use any id not used on any of the cluster nodes.

  ```
  ksc# /usr/sbin/useradd -u 200 -g oinstall -G dba,oper oracle
  ```
- Check User:

  ```
  ksc# id oracle
  uid=203(oracle) gid=103(oinstall) groups=101(dba),104(oper)
  ```
- Create HOME directory for Oracle user

  ```
  ksc/schalke# mkdir /home/oracle
  ksc/schalke# chown oracle:oinstall /home/oracle
  ```
- Change Password on each node:

  ```
  ksc/schalke# passwd oracle
  ```
- Remote copy (rcp) needs to be enabled for both the root + oracle accounts on all nodes to allow remote copy of cluster configuration files. Include the following lines in the .rhosts file in root's home directory :

  ```
  # .rhosts file in $HOME of root
  ksc              root
  ksc.domain       root
  schalke          root
  schalke.domain   root
  ```

  ```
  # .rhosts file in $HOME of oracle
  ksc              oracle
  ksc.domain       oracle
  schalke          oracle
  schalke.domain   oracle
  ```

  Note: rcp only works if for the respective user a password has been set (root and oracle). You can test whether it is working with:

  ```
  ksc# remsh schalke ll
  ksc# remsh ksc ll
  schalke# remsh schalke ll
  schalke# remsh ksc ll

  ksc$ remsh schalke ll
  ksc$ remsh ksc ll
  schalke$ remsh schalke ll
  schalke$ remsh ksc ll
  ```

6. Oracle RAC 10g Cluster Preparation Steps

The cluster configuration steps vary depending on the chosen RAC 10g cluster model. Therefore,

we have split this section in respective sub chapters. Please follow the instructions that apply to your chosen deployment model.

6.1 RAC 10g with HP Serviceguard Cluster File System for RAC

In this example we create three cluster file systems for

- /cfs/oraclu: Oracle Clusterware Files 300MB
- /cfs/orabin Oracle binaries 10 GB
- /cfs/oradata: Oracle database files 10 GB

- For the cluster lock, you can either use a lock disk or a quorum server. Here, we do describe the steps to set-up a lock disk. This is done from node ksc:

```
ksc# mkdir /dev/vglock
ksc# mknod /dev/vglock/group c 64 0x020000   # If minor number 0x020000 is already in use, please use a free number!!
ksc# pvcreate -f /dev/rdsk/c6t0d1
Physical volume "/dev/rdsk/c6t0d1" has been successfully created.
ksc# vgcreate /dev/vglock /dev/dsk/c6t0d1
Volume group "/dev/vglock" has been successfully created.
Volume Group configuration for /dev/vglock has been saved in /etc/lvmconf/vglock.conf
```

- Check Volume Group definition on ksc:

```
ksc# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c3t0d0s2
/dev/vglock
/dev/dsk/c6t0d1
```

- Export the volume group to mapfile and copy this to node schalke

```
ksc# vgchange -a n /dev/vglock
Volume group "/dev/vglock" has been successfully changed.
ksc# vgexport -v -p -s -m /etc/cmcluster/vglockmap vglock
Beginning the export process on Volume Group "/dev/vglock".
/dev/dsk/c6t0d1
ksc# rcp /etc/cmcluster/vglockmap schalke:/etc/cmcluster
```

- Import the volume group definition on node schalke

```
schalke# mkdir /dev/vglock
schalke# mknod /dev/vglock/group c 64 0x020000   (Note: The minor number has to be the same as on node ksc)
schalke# vgimport -v -s -m /etc/cmcluster/vglockmap vglock
Beginning the import process on Volume Group "/dev/vglock".
Volume group "/dev/vglock" has been successfully created.
```

- Create the SG cluster config file from ksc:

```
ksc# cmquerycl -v -n ksc -n schalke -C RACCFS.asc
```

- Edit the cluster configuration file

  Make the necessary changes to this file for your cluster. For example, change the Cluster Name, and adjust the heartbeat interval and node timeout to prevent unexpected failovers. Also, ensure to have the right lan interfaces configured for the SG heartbeat according to chapter 4.2.

- Check the cluster configuration:

```
ksc# cmcheckconf -v -C RACCFS.asc
```

- Create the binary configuration file and distribute the cluster configuration to all the nodes in the cluster:

```
ksc # cmapplyconf -v -C RACCFS.asc
```

(Note: the cluster is not started until you run cmrunnode on each node or cmruncl.)

- Start and check status of cluster

```
ksc# cmruncl -v
Waiting for cluster to form ..... done
Cluster successfully formed.
Check the syslog files on all nodes in the cluster to verify that no warnings occurred during startup.
```

```
ksc# cmviewcl

CLUSTER        STATUS
RACCFS         up

  NODE           STATUS       STATE
  ksc            up           running
  schalke        up           running
```

- Disable automatic volume group activation on all cluster nodes by setting AUTO_VG_ACTIVATE to 0 in file /etc/lvmrc. This ensures that shared volume group vglock is not automatically activated at system boot time. In case you need to have any other volume groups activated, you need to explicitly list them at the customized volume group activation section.

- Initialize VxVM on both nodes:

```
ksc# vxinstall

VxVM uses license keys to control access. If you have not yet installed
a VxVM license key on your system, you will need to do so if you want
to use the full functionality of the product.

Licensing information:
System host ID: 3999750283
Host type: ia64 hp server rx4640

Are you prepared to enter a license key [y,n,q] (default: n) n

Do you want to use enclosure based names for all disks ?
[y,n,q,?] (default: n) n
Populating VxVM DMP device directories ....

V-5-1-0 vxvm:vxconfigd: NOTICE: Generating /etc/vx/array.info
-
The Volume Daemon has been enabled for transactions.
Starting the relocation daemon, vxrelocd.
Starting the cache deamon, vxcached.
Starting the diskgroup config backup deamon, vxconfigbackupd.

Do you want to setup a system wide default disk group?
[y,n,q,?] (default: y) n
```

schalke# vxinstall (same options as for ksc)

- Create CFS package

ksc# cfscluster config -t 900 -s (if it does not work, look at /etc/cmcluster/cfs/SG-CFS-pkg.log)

```
CVM is now configured
Starting CVM...
It might take a few minutes to complete
VxVM vxconfigd NOTICE V-5-1-7900 CVM_VOLD_CONFIG command received
VxVM vxconfigd NOTICE V-5-1-7899 CVM_VOLD_CHANGE command received
VxVM vxconfigd NOTICE V-5-1-7961 establishing cluster for seqno = 0x10f9d07.
VxVM vxconfigd NOTICE V-5-1-8059 master: cluster startup
VxVM vxconfigd NOTICE V-5-1-8061 master: no joiners
VxVM vxconfigd NOTICE V-5-1-4123 cluster established successfully
VxVM vxconfigd NOTICE V-5-1-7899 CVM_VOLD_CHANGE command received
VxVM vxconfigd NOTICE V-5-1-7961 establishing cluster for seqno = 0x10f9d08.
VxVM vxconfigd NOTICE V-5-1-8062 master: not a cluster startup
VxVM vxconfigd NOTICE V-5-1-3765 master: cluster join complete for node 1
VxVM vxconfigd NOTICE V-5-1-4123 cluster established successfully
CVM is up and running
```

- Check CFS status:

```
ksc# cfscluster status
Node             : ksc
Cluster Manager : up
CVM state : up (MASTER)
```

```
MOUNT POINT TYPE SHARED VOLUME DISK GROUP STATUS

Node             : schalke
Cluster Manager : up
CVM state : up
MOUNT POINT TYPE SHARED VOLUME DISK GROUP STATUS
```

- Check SG-CFS-pkg:

```
ksc# cmviewcl -v

....

MULTI_NODE_PACKAGES

  PACKAGE        STATUS       STATE        AUTO_RUN     SYSTEM
  SG-CFS-pkg     up           running      enabled      yes

    NODE_NAME      STATUS       SWITCHING
    ksc            up           enabled

      Script_Parameters:
      ITEM       STATUS   MAX_RESTARTS  RESTARTS   NAME
      Service    up       0             0          SG-CFS-vxconfigd
      Service    up       5             0          SG-CFS-sgcvmd
      Service    up       5             0          SG-CFS-vxfsckd
      Service    up       0             0          SG-CFS-cmvxd
      Service    up       0             0          SG-CFS-cmvxpingd

    NODE_NAME      STATUS       SWITCHING
    schalke        up           enabled

      Script_Parameters:
      ITEM       STATUS   MAX_RESTARTS  RESTARTS   NAME
      Service    up       0             0          SG-CFS-vxconfigd
      Service    up       5             0          SG-CFS-sgcvmd
      Service    up       5             0          SG-CFS-vxfsckd
      Service    up       0             0          SG-CFS-cmvxd
      Service    up       0             0          SG-CFS-cmvxpingd
```

- List path type and states for disks:

```
ksc# vxdisk list (DEVICE       TYPE            DISK         GROUP        STATUS
c2t1d0       auto:none       -            -            online invalid
c3t0d0s2     auto:LVM        -            -            LVM
c6t0d1       auto:LVM        -            -            LVM
c6t0d2       auto:none       -            -            online invalid
c6t0d3       auto:none       -            -            online invalid
c6t0d4       auto:none       -            -            online invalid
```

- Create disk groups for RAC:

```
ksc# /etc/vx/bin/vxdisksetup -i c6t0d2
ksc# vxdg -s init dgrac c6t0d2
```
(use the -s option to specify shared mode)

```
ksc# vxdg -g dgrac adddisk c6t0d3
```
(optional, only when you want to add more disks to a disk group)

Please note that his needs to be done from master node. Check for master/slave using

```
ksc# cfsdgadm display -v
Node Name : ksc (MASTER)
Node Name : schalke
```

- List again path type and states for disks:

```
ksc# vxdisk list
DEVICE       TYPE            DISK         GROUP        STATUS
c2t1d0       auto:none       -            -            online invalid
c3t0d0s2     auto:LVM        -            -            LVM
c6t0d1       auto:LVM        -            -            LVM
c6t0d2       auto:cdsdisk    c6t0d2       dgrac        online shared
c6t0d3       auto:cdsdisk    c6t0d3       dgrac        online shared
c6t0d4       auto:none       -            -            online invalid
```

- Generate the SG-CFS-DG package:

```
ksc# cfsdgadm add dgrac all=sw
Package name "SG-CFS-DG-1" is generated to control the resource
Shared disk group "dgrac" is associated with the cluster
```

- Activate SG-CFS-DG package:

```
ksc# cfsdgadm activate dgrac
```

- Check SG-CFS-DG package:

```
ksc# cmviewcl -v
...
MULTI_NODE_PACKAGES

  PACKAGE        STATUS       STATE        AUTO_RUN     SYSTEM
  SG-CFS-pkg     up           running      enabled      yes

    NODE_NAME      STATUS       SWITCHING
    ksc            up           enabled
...
    NODE_NAME      STATUS       SWITCHING
    schalke           up           enabled
...
  PACKAGE        STATUS       STATE        AUTO_RUN     SYSTEM
  SG-CFS-DG-1    up           running      enabled      no

    NODE_NAME      STATUS       STATE        SWITCHING
    ksc            up           running      enabled

      Dependency_Parameters:
      DEPENDENCY_NAME         SATISFIED
      SG-CFS-pkg              yes

    NODE_NAME      STATUS       STATE        SWITCHING
    schalke        up           running      enabled

      Dependency_Parameters:
      DEPENDENCY_NAME         SATISFIED
      SG-CFS-pkg              yes
```

- Create volumes, file systems and mount point for CFS from VxVM master node:

```
ksc# vxassist -g dgrac make vol1 300M
ksc# vxassist -g dgrac make vol2 10240M
ksc# vxassist -g dgrac make vol3 10240M

ksc# newfs -F vxfs /dev/vx/rdsk/dgrac/vol1
    version 6 layout
    307200 sectors, 307200 blocks of size 1024, log size 1024 blocks
    largefiles supported
ksc# newfs -F vxfs /dev/vx/rdsk/dgrac/vol2
     version 6 layout
    10485760 sectors, 10485760 blocks of size 1024, log size 16384 blocks
    largefiles supported
ksc# newfs -F vxfs /dev/vx/rdsk/dgrac/vol3
    version 6 layout
    10485760 sectors, 10485760 blocks of size 1024, log size 16384 blocks
    largefiles supported

ksc# cfsmntadm add dgrac vol1 /cfs/oraclu all=rw
Package name "SG-CFS-MP-1" is generated to control the resource
Mount point "/cfs/oraclu" is associated with the cluster
ksc# cfsmntadm add dgrac vol2 /cfs/orabin all=rw
Package name "SG-CFS-MP-2" is generated to control the resource
Mount point "/cfs/orabin" is associated with the cluster
ksc# cfsmntadm add dgrac vol3 /cfs/oradata all=rw
Package name "SG-CFS-MP-3" is generated to control the resource
Mount point "/cfs/oradata" is associated with the cluster
```

- Mounting Cluster Filesystems

```
ksc# cfsmount /cfs/oraclu
ksc# cfsmount /cfs/orabin
ksc# cfsmount /cfs/oradata
```

- Check CFS mountpoints:

```
ksc# bdf
Filesystem          kbytes    used   avail %used Mounted on
/dev/vg00/lvol3    8192000 1672312 6468768   21% /
/dev/vg00/lvol1     622592  221592  397896   36% /stand
/dev/vg00/lvol7    8192000 2281776 5864152   28% /var
/dev/vg00/lvol8    1032192   20421  948597    2% /var/opt/perf
/dev/vg00/lvol6    8749056 2958760 5745072   34% /usr
```

```
/dev/vg00/lvol5     4096000   16920 4047216     0% /tmp
/dev/vg00/lvol4     22528000 3704248 18676712   17% /opt
/dev/odm                  0       0       0    0% /dev/odm
/dev/vx/dsk/dgrac/vol1
                     307200    1802  286318    1% /cfs/oraclu
/dev/vx/dsk/dgrac/vol2
                    10485760   19651 9811985     0% /cfs/orabin
/dev/vx/dsk/dgrac/vol3
                    10485760   19651 9811985     0% /cfs/oradata
```

- Check SG cluster configuration:

```
ksc# cmviewcl
CLUSTER        STATUS
RACCFS         up

  NODE           STATUS       STATE
  ksc            up           running
  schalke        up           running

MULTI_NODE_PACKAGES

  PACKAGE        STATUS       STATE        AUTO_RUN     SYSTEM
  SG-CFS-pkg     up           running      enabled      yes
  SG-CFS-DG-1    up           running      enabled      no
  SG-CFS-MP-1    up           running      enabled      no
  SG-CFS-MP-2    up           running      enabled      no
  SG-CFS-MP-3    up           running      enabled      no
```

6.2 RAC 10g with RAW over SLVM

6.2.1 SLVM Configuration

To use shared raw logical volumes, HP Serviceguard Extensions for RAC must be installed on all cluster nodes.

For a basic database configuration with SLVM, the following shared logical volumes are required. Note that in this scenario, only one SLVM volume group is used for both Oracle Clusterware and database files. In cluster environments with more than one RAC database, it is recommended to have separate SLVM volume groups for Oracle Clusterware and for each RAC database.

Create a Raw Device for:	**File Size:**	**Sample Name:** *<dbname> should be replaced with your database name.*	**Comments:**
OCR (Oracle Cluster Repository)	108 MB	raw_ora_ocr_108m	You need to create this raw logical volume only once on the cluster. If you create more than one database on the cluster, they all share the same OCR.
Oracle Voting disk	28 MB	raw_ora_vote_28m	You need to create this raw logical volume only once on the cluster. If you create more than one database on the cluster, they all share the same Oracle voting disk.
SYSTEM tablespace	508 MB	raw_<dbname>_system_508m	
SYSAUX tablespace	300 + (Number of instances * 250)	raw_<dbname>_sysaux_808m	New system-managed tablespace that contains performance data and combines content that was stored in different tablespaces (some of which are no longer required) in earlier releases. This is a required tablespace for which you must plan disk space.
One Undo tablespace	508 MB	raw_<dbname>_undotbs*n*_508m	One tablespace for each

per instance			instance, where *n* is the number of the instance
EXAMPLE tablespace	168 MB	raw_<dbname>_example_168m	
USERS tablespace	128 MB	raw_<dbname>_users_128m	
Two ONLINE Redo log files per instance	128 MB	raw_<dbname>_redo*nm*_128m	n is instance number and m the log number
First and second control file	118 MB	raw_<dbname>_control[1|2] _118m	
TEMP tablespace	258 MB	raw_<dbname>_temp_258m	
Server parameter file (SPFILE):	5 MB	raw_<dbname>_spfile_raw_5m	
Password file	5 MB	raw_<dbname>_pwdfile_5m	

- Disks need to be properly initialized before being added into volume groups. Do the following step for all the disks (LUNs) you want to configure for your RAC volume group(s) from node ksc:

```
ksc# pvcreate –f /dev/rdsk/cxtydz ( where x=instance, y=target, and z=unit)
```

- Create the volume group directory with the character special file called group:

```
ksc# mkdir /dev/vg_rac
ksc# mknod /dev/vg_rac/group c 64 0x060000
```

  Note: <0x060000> is the minor number in this example. This minor number for the group file must be unique among all the volume groups on the system.

- Create VG (optionally using PV-LINKs) and extend the volume group:

```
ksc# vgcreate /dev/vg_rac /dev/dsk/c0t1d0 /dev/dsk/c1t0d0 (primary path ...
secondary path)
ksc# vgextend /dev/vg_rac /dev/dsk/c1t0d1 /dev/dsk/c0t1d1
```

  Continue with vgextend until you have included all the needed disks for the volume group(s).

- Create logical volumes as shown in the table above for the RAC database with the command

```
ksc# lvcreate –i 10 –I 1024 –L 100 –n Name /dev/vg_rac
-i: number of disks to stripe across
-I: stripe size in kilobytes
-L: size of logical volume in MB
```

- Check to see if your volume groups are properly created and available:

```
ksc# strings /etc/lvmtab
ksc# vgdisplay –v /dev/vg_rac
```

- Export the volume group:
  - De-activate the volume group:

    ```
    ksc# vgchange –a n /dev/vg_rac
    ```

  - Create the volume group map file:

    ```
    ksc# vgexport –v –p –s –m mapfile /dev/vg_rac
    ```

  - Copy the mapfile to all the nodes in the cluster:

    ```
    ksc# rcp mapfile schalke:/tmp/scripts
    ```

- Import the volume group on the second node in the cluster
  - Create a volume group directory with the character special file called group:

    ```
    schalke# mkdir /dev/vg_rac
    schalke# mknod /dev/vg_rac/group c 64 0x060000
    ```

    Note: The minor number has to be the same as on the other node.

  - Import the volume group:

    ```
    schalke# vgimport –v –s –m /tmp/scripts/mapfile /dev/vg_rac
    ```

    Note: The minor number has to be the same as on the other node.

- Check to see if devices are imported:

```
schalke# strings /etc/lvmtab
```

- Disable automatic volume group activation on all cluster nodes by setting AUTO_VG_ACTIVATE to 0 in file /etc/lvmrc. This ensures that shared volume group vg_rac is not automatically activated at system boot time. In case you need to have any other volume groups activated, you need to explicitly list them at the customized volume group activation section.
- It is recommended best practice to create symbolic links for each of these raw files on all systems of your RAC cluster.

```
ksc/schalke# cd /oracle/RAC/ (directory where you want to have the links)
ksc/schalke# ln -s /dev/vg_rac/raw_<dbname>_system_508 system
ksc/schalke# ln -s /dev/vg_rac/raw_<dbname>_users_128m user
```

  etc.
- Change the permissions of the database volume group vg_rac to 777, and change the permissions of all raw logical volumes to 660 and the owner to oracle:oinstall.

```
ksc/schalke# chmod 777 /dev/vg_rac
ksc/schalke# chmod 660 /dev/vg_rac/r*
ksc/schalke# chown oracle:dba /dev/vg_rac/r*
```

- Change the permissions of the OCR logical volumes:

```
ksc/schalke# chown root:oinstall /dev/vg_rac/raw_ora_ocr_108m
ksc/schalke# chmod 640 /dev/vg_rac/raw_ora_ocr_108m
```

- Optional: To enable Database Configuration Assistant (DBCA) later to identify the appropriate raw device for each database file, you must create a raw device mapping file, as follows:
  - Set the ORACLE_BASE environment variable :

```
ksc/schalke$ export ORACLE_BASE=/opt/oracle/product
```

  - Create a database file subdirectory under the Oracle base directory and set the appropriate owner, group, and permissions on it:

```
ksc/schalke# mkdir -p $ORACLE_BASE/oradata/<dbname>
ksc/schalke# chown -R oracle:oinstall $ORACLE_BASE/oradata
ksc/schalke# chmod -R 775 $ORACLE_BASE/oradata
```

  - Change directory to the $ORACLE_BASE/oradata/dbname directory.
  - Enter a command similar to the following to create a text file that you can use to create the raw device mapping file:

```
ksc# find /dev/vg_rac -user oracle -name 'raw*' -print > dbname_raw.conf
```

  - Create the dbname_raw.conf file that looks similar to the following:

```
system=/dev/vg_rac/raw_<dbname>_system_508m
sysaux=/dev/vg_rac/raw_<dbname>_sysaux_808m
example=/dev/vg_rac/raw_<dbname>_example_168m
users=/dev/vg_rac/raw_<dbname>_users_128m
temp=/dev/vg_rac/raw_<dbname>_temp_258m
undotbs1=/dev/vg_rac/raw_<dbname>_undotbs1_508m
undotbs2=/dev/vg_rac/raw_<dbname>_undotbs2_508m
redo1_1=/dev/vg_rac/raw_<dbname>_redo11_128m
redo1_2=/dev/vg_rac/raw_<dbname>_redo12_128m
redo2_1=/dev/vg_rac/raw_<dbname>_redo21_128m
redo2_2=/dev/vg_rac/raw_<dbname>_redo22_128m
control1=/dev/vg_rac/raw_<dbname>_control1_118m
control2=/dev/vg_rac/raw_<dbname>_control2_118m
spfile=/dev/vg_rac/rraw_<dbname>_spfile_5m
pwdfile=/dev/vg_rac/raw_<dbname>_pwdfile_5m
```

  - When you are configuring the Oracle user's environment later in this chapter, set the DBCA_RAW_CONFIG environment variable to specify the full path to this file:

```
ksc$ export DBCA_RAW_CONFIG=$ORACLE_BASE/oradata/dbname/dbname_raw.conf
```

6.2.2 SG/SGeRAC Configuration

After SLVM set-up, you can now start the Serviceguard cluster configuration.

In general, you can configure your Serviceguard cluster using lock disk or quorum server. We describe here the cluster lock disk set-up. Since we have already configured one volume group for the entire RAC cluster vg_rac (see last chapter 5.2.1), we use vg_rac for the lock volume as well.

- Activate the lock disk on the configuration node ONLY. Lock volume can only be activated on the node where the cmapplyconf command is issued so that the lock disk can be initialized accordingly.
  ```
  ksc# vgchange -a y /dev/vg_rac
  ```
- Create a cluster configuration template:
  ```
  ksc# cmquerycl –n ksc –n schalke –v –C /etc/cmcluster/rac.asc
  ```
- Edit the cluster configuration file (rac.asc).
  Make the necessary changes to this file for your cluster. For example, change the Cluster Name, adjust the heartbeat interval and node timeout to prevent unexpected failovers due to DLM traffic. Configure all shared volume groups that you are using for RAC, including the volume group that contains the Oracle CRS files using the parameter OPS_VOLUME_GROUP at the bottom of the file. Also, ensure to have the right lan interfaces configured for the SG heartbeat according to chapter 4.2.
- Check the cluster configuration:
  ```
  ksc# cmcheckconf -v -C rac.asc
  ```
- Create the binary configuration file and distribute the cluster configuration to all the nodes in the cluster:
  ```
  ksc# cmapplyconf -v -C rac.asc
  ```
  Note: the cluster is not started until you run cmrunnode on each node or cmruncl.
- De-activate the lock disk on the configuration node after cmapplyconf
  ```
  ksc# vgchange -a n /dev/vg_rac
  ```
- Start the cluster and view it to be sure its up and running. See the next section for instructions on starting and stopping the cluster.

How to start up the cluster:

- Start the cluster from any node in the cluster
  ```
  ksc# cmruncl -v
  ```
  Or, on each node
  ```
  ksc/schalke# cmrunnode -v
  ```
- Make all RAC volume groups and Cluster Lock volume groups sharable and cluster aware (not packages) from the cluster configuration node. This has to be done only once.
  ```
  ksc# vgchange -S y -c y /dev/vg_rac
  ```
- Then on all the nodes, activate the volume group in shared mode in the cluster. This has to be done each time when you start the cluster.
  ```
  ksc# vgchange -a s /dev/vg_rac
  ```
- Check the cluster status:
  ```
  ksc# cmviewcl –v
  ```

How to shut down the cluster (not needed here):

- Shut down the RAC instances (if up and running)
- On all the nodes, deactivate the volume group in shared mode in the cluster:
  ```
  ksc# vgchange –a n /dev/vg_rac
  ```

- Halt the cluster from any node in the cluster

```
ksc# cmhaltcl –v
```

- Check the cluster status:

```
ksc# cmviewcl –v
```

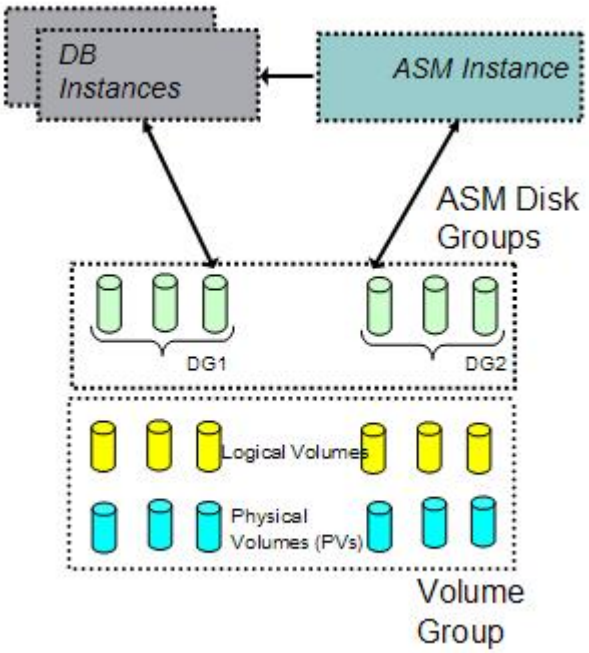
6.3 RAC 10g with ASM over SLVM

To use shared raw logical volumes, HP Serviceguard Extensions for RAC must be installed on all cluster nodes.

6.3.1 SLVM Configuration

Before continuing, check the following ASM-over-SLVM configuration guidelines:

- organize the disks/LUNs to be used by ASM into LVM volume groups (VGs)
- ensure that there are multiple paths to each disk, by configuring PV Links or disk level multipathing
- for each physical volume (PV), configure a logical volume (LV) using up all available space on that PV
- the ASM logical volumes should not be striped or mirrored, should not span multiple PVs, and should not share a PV with LVs corresponding to other disk group members as ASM provides these features and SLVM supplies only the missing functionality (chiefly multipathing)



- on each LV, set an I/O timeout equal to (# of PV Links) *(PV timeout)
- export the VG across the cluster and mark it shared

For a ASM database configuration on top of SLVM, you need shared logical volumes for the two Oracle Clusterware files OCR and Voting plus shared logical volumes for Oracle ASM.

Create a Raw Device for:	**File Size:**	**Sample Name:** *<dbname> should be replaced with your database name.*	**Comments:**
OCR (Oracle Cluster	108 MB	raw_ora_ocr*n*_108m	With RAC10g R2, Oracle lets you

Registry) [1/2]			have 2 redundant copies for OCR. In this case you need two shared logical volumes. n = 1 or 2. For HA reasons, they should not be on same set of disks.
Oracle CRS voting disk [1/3/..]	28 MB	raw_ora_vote*n*_28m	With RAC10g R2, Oracle is lets you have 3+ redundant copies of Voting. In this case you need 3+ shared logical volumes. n = 1 or 3 or 5 .... For HA reasons, they should not be on same set of disks.
ASM Volume #1 .. n	10GB	raw_ora_asm*n*_10g	

This ASM-over-SLVM configuration enables the HP-UX devices used for disk group members to have the same names on all nodes, easing ASM configuration.

In this example, ASM disk group using disks /dev/dsk/c9t0d1 and /dev/dsk/c9t0d2; alternate paths /dev/dsk/c10t0d1 and /dev/dsk/c10t0d2.

- Disks need to be properly initialized before being added into volume groups. Do the following step for all the disks (LUNs) you want to configure for your RAC volume group(s) from node ksc:

```
ksc# pvcreate –f /dev/rdsk/c9t0d1
ksc# pvcreate –f /dev/rdsk/c9t0d2
```

- Create the volume group directory with the character special file called group:

```
ksc# mkdir /dev/vgasm
ksc# mknod /dev/vgasm/group c 64 0x060000
```

  Note: <0x060000> is the minor number in this example. This minor number for the group file must be unique among all the volume groups on the system.

- Create VG (optionally using PV-LINKs) and extend the volume group:

```
ksc# vgcreate /dev/vgasm /dev/dsk/c9t0d1 /dev/dsk/c10t0d1 (primary path ...
secondary path)
ksc# vgextend /dev/vgasm /dev/dsk/c10t0d2 /dev/dsk/c9t0d2
```

- Create zero length LVs for each of the physical volumes:

```
ksc# lvcreate -n raw_ora_asm1_10g vgasm
ksc# lvcreate -n raw_ora_asm2_10g vgasm
```

- Ensure each LV will be contiguous and stay on one PV:

```
ksc# lvchange –C y /dev/vgasm/raw_ora_asm1_10g
ksc# lvchange –C y /dev/vgasm/raw_ora_asm2_10g
```

- Extend each LV to the full length allowed by the corresponding PV, in this case 2900 extents:

```
ksc# lvextend -l 2900 /dev/vgasm/raw_ora_asm1_10g /dev/dsk/c9t0d1
ksc# lvextend -l 2900 /dev/vgasm/raw_ora_asm2_10g /dev/dsk/c9t0d2
```

- Configure LV level timeouts, otherwise a single PV failure could result in a database hang. Here we assume a PV timeout of 30 seconds. Since there are 2 paths to each disk, the LV timeout is 60 seconds:

```
ksc# lvchange -t 60 /dev/vgasm/raw_ora_asm1_10g
ksc# lvchange -t 60 /dev/vgasm/raw_ora_asm2_10g
```

- Null out the initial part of each LV to ensure ASM accepts the # LV as an ASM disk group member (see Oracle Metalink Note 268481.1)

```
ksc# dd if=/dev/zero of=/dev/vgasm/raw_ora_asm1_10g bs=8192 count=12800
```

```
ksc# dd if=/dev/zero of=/dev/vgasm/raw_ora_asm2_10g bs=8192 count=12800
```

- Check to see if your volume groups are properly created and available:

```
ksc# strings /etc/lvmtab
ksc# vgdisplay –v /dev/vg_rac
```

- Export the volume group:
  - De-activate the volume group:

```
ksc# vgchange –a n /dev/vgasm
```

  - Create the volume group map file:

```
ksc# vgexport –v –p –s –m vgasm.map /dev/vgasm
```

  - Copy the mapfile to all the nodes in the cluster:

```
ksc# rcp vgasm.map schalke:/tmp/scripts
```

- Import the volume group on the second node in the cluster
  - Create a volume group directory with the character special file called group:

```
schalke# mkdir /dev/vgasm
schalke# mknod /dev/vgasm/group c 64 0x060000
```

    Note: The minor number has to be the same as on the other node.
  - Import the volume group:

```
schalke# vgimport –v –s –m /tmp/scripts/vgasm.map /dev/vgasm
```

    Note: The minor number has to be the same as on the other node.
  - Check to see if devices are imported:

```
schalke# strings /etc/lvmtab
```

- Disable automatic volume group activation on all cluster nodes by setting AUTO_VG_ACTIVATE to 0 in file /etc/lvmrc. This ensures that shared volume group vgasm is not automatically activated at system boot time. In case you need to have any other volume groups activated, you need to explicitly list them at the customized volume group activation section.

6.3.2 SG/SGeRAC Configuration

After SLVM set-up, you can now start the Serviceguard cluster configuration.

In general, you can configure your Serviceguard cluster using lock disk or quorum server. We describe here the cluster lock disk set-up. Since we have already configured one volume group for the RAC cluster vgasm (see last chapter 5.3.1), we use vgasm for the lock volume as well.

- Activate the lock disk on the configuration node ONLY. Lock volume can only be activated on the node where the cmapplyconf command is issued so that the lock disk can be initialized accordingly.

```
ksc# vgchange -a y /dev/vgasm
```

- Create a cluster configuration template:

```
ksc# cmquerycl –n ksc –n schalke –v –C /etc/cmcluster/rac.asc
```

- Edit the cluster configuration file (rac.asc).
  Make the necessary changes to this file for your cluster. For example, change the Cluster Name, adjust the heartbeat interval and node timeout to prevent unexpected failovers due to RAC traffic. Configure all shared volume groups that you are using for RAC, including the volume group that contains the Oracle Clusterware files using the parameter OPS_VOLUME_GROUP at the bottom of the file. Also, ensure to have the right lan interfaces configured for the SG heartbeat according to chapter 4.2.
- Check the cluster configuration:

```
ksc# cmcheckconf -v -C rac.asc
```

- Create the binary configuration file and distribute the cluster configuration to all the nodes in the cluster:

  ```
  ksc# cmapplyconf -v -C rac.asc
  ```

  Note: the cluster is not started until you run cmrunnode on each node or cmruncl.
- De-activate the lock disk on the configuration node after cmapplyconf

  ```
  ksc# vgchange -a n /dev/vgasm
  ```
- Start the cluster and view it to be sure its up and running. See the next section for instructions on starting and stopping the cluster.

How to start up the cluster:

- Start the cluster from any node in the cluster

  ```
  ksc# cmruncl -v
  ```

  Or, on each node

  ```
  ksc# cmrunnode -v
  ```
- Make all RAC volume groups and Cluster Lock volume groups sharable and cluster aware (not packages) from the cluster configuration node. This has to be done only once.

  ```
  ksc# vgchange -S y -c y /dev/vgasm
  ```
- Then on all the nodes, activate the volume group in shared mode in the cluster. This has to be done each time when you start the cluster.

  ```
  ksc# vgchange -a s /dev/vgasm
  ```
- Check the cluster status:

  ```
  ksc# cmviewcl –v
  ```

How to shut down the cluster (not needed here):

- Shut down the RAC instances (if up and running)
- On all the nodes, deactivate the volume group in shared mode in the cluster:

  ```
  ksc# vgchange –a n /dev/vgasm
  ```
- Halt the cluster from any node in the cluster

  ```
  ksc# cmhaltcl –v
  ```
- Check the cluster status:

  ```
  ksc# cmviewcl –v
  ```

6.4 RAC 10g with ASM

For Oracle RAC10g on HP-UX with ASM, please note:

- As said before (chapter 2), you cannot use Automatic Storage Management to store Oracle Clusterware files (OCR + Voting). This is because they must be accessible before Oracle ASM starts.
- As this deployment option is not using HP Serviceguard Extension for RAC, you cannot configure shared logical volumes (Shared Logical Volumer Manager is a feature of SGeRAC).
- Only one ASM instance is required per node. So you might have multiple databases, but they will share the same single ASM instance.
- The following files can be placed in an ASM disk group: DATAFILE, CONTROLFILE, REDOLOG, ARCHIVELOG and SPFILE. You cannot put any other files such as Oracle

binaries, or the two Oracle Clusterware files (OCR & Voting) into an ASM disk group.

- For Oracle RAC with Standard Edition installations, ASM is the only supported storage option for database or recovery files.
- You do not have to use the same storage mechanism for database files and recovery files. You can use raw devices for database files and ASM for recovery files if you choose.
- For RAC installations, if you choose to enable automated backups, you must choose ASM for recovery file storage.
- All of the devices in an ASM disk group should be the same size and have the same performance characteristics.
- For RAC installations, you must add additional disk space for the ASM metadata. You can use the following formula to calculate the additional disk space requirements (in MB: 15 + (2 * number_of_disks) + (126 * number_of_ASM_instances)
  For example, for a four-node RAC installation, using three disks in a high redundancy disk group, you require an additional 525 MB of disk space: 15 + (2 * 3) + (126 * 4) = 525
- Choose the redundancy level for the ASM disk group(s). The redundancy level that you choose for the ASM disk group determines how ASM mirrors files in the disk group and determines the number of disks and amount of disk space that you require, as follows:
  - External redundancy: An external redundancy disk group requires a minimum of one disk device. Typically you choose this redundancy level if you have an intelligent subsystem such as an HP StorageWorks EVA or HP StorageWorks XP.
  - Normal redundancy: In a normal redundancy disk group, ASM uses two-way mirroring by default, to increase performance and reliability. A normal redundancy disk group requires a minimum of two disk devices (or two failure groups).
  - High redundancy: In a high redundancy disk group, ASM uses three-way mirroring to increase performance and provide the highest level of reliability. A high redundancy disk group requires a minimum of three disk devices (or three failure groups).

Raw Disk for:	File Size:	Comments:
OCR (Oracle Cluster Registry) [1/2]	108 MB	With RAC10g R2, Oracle lets you have 2 redundant copies for OCR. In this case you need two shared logical volumes. n = 1 or 2. For HA reasons, they should not be on same set of disks.
Oracle CRS voting disk [1/3/..]	28 MB	With RAC10g R2, Oracle is lets you have 3+ redundant copies of Voting. In this case you need 3+ shared logical volumes. n = 1 or 3 or 5 .... For HA reasons, they should not be on same set of disks.
ASM Disk #1 .. n	10GB	Disks 1 .. n

To configure raw disk devices / partitions for database file storage, follow the following steps:

- To make sure that the disks are available, enter the following command on every node:

```
ksc/schalke# /usr/sbin/ioscan -funCdisk
```

```
The output from this command is similar to the following:

Class    I   H/W Path       Driver      S/W State      H/W Type     Description
=============================================================================
disk     4   255/255/0/0.0  sdisk       CLAIMED        DEVICE       HSV100 HP
                            /dev/dsk/c8t0d0 /dev/rdsk/c8t0d0
disk     5   255/255/0/0.1  sdisk       CLAIMED         DEVICE      HSV100 HP
                            /dev/dsk/c8t0d1 /dev/rdsk/c8t0d1
```

  This command displays information about each disk attached to the system, including the block device name (/dev/dsk/cxtydz) and the character raw device name (/dev/rdsk/cxtydz).

- If the ioscan command does not display device name information for a device that you want to use, enter the following command to install the special device files for any new devices:

  ksc/schalke# insf -e (please note, this command does reset the permissions to root for already existing device files, e.g. ASM disks!!)

- For each disk that you want to use, enter the following command on any node to verify that it is not already part of an LVM volume group:

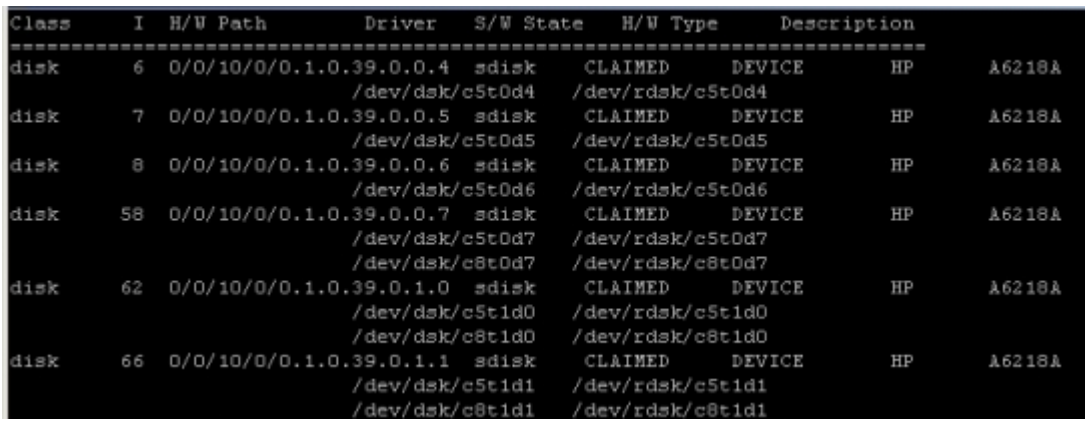
  ksc# pvdisplay /dev/dsk/cxtydz

  If this command displays volume group information, the disk is already part of a volume group. The disks that you choose must not be part of an LVM volume group.

- Please note that the device paths for Oracle Clusterware and ASM disks must be the same from both systems. If they are not the same use the following command to map them to a new virtual device name:

  ```
  #mksf -C disk -H <hardware path> -I 62 <new virtual device name>
  #mksf -C disk -H <hardware path> -I 62 -r <new virtual device name>
  Example:
  #mksf -C disk -H 0/0/10/0/0.1.0.39.0.1.0 -I 62 /dev/dsk/c8t1d0
  #mksf -C disk -H 0/0/10/0/0.1.0.39.0.1.0 -I 62 -r /dev/rdsk/c8t1d0
  ```

  As you can see at the following output of the **ioscan** command, now multiple device names are mapped to the same hardware path.



```
Class     I  H/W Path        Driver   S/W State   H/W Type     Description
=============================================================================
disk      6  0/0/10/0/0.1.0.39.0.0.4  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t0d4   /dev/rdsk/c5t0d4
disk      7  0/0/10/0/0.1.0.39.0.0.5  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t0d5   /dev/rdsk/c5t0d5
disk      8  0/0/10/0/0.1.0.39.0.0.6  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t0d6   /dev/rdsk/c5t0d6
disk     58  0/0/10/0/0.1.0.39.0.0.7  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t0d7   /dev/rdsk/c5t0d7
                           /dev/dsk/c8t0d7   /dev/rdsk/c8t0d7
disk     62  0/0/10/0/0.1.0.39.0.1.0  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t1d0   /dev/rdsk/c5t1d0
                           /dev/dsk/c8t1d0   /dev/rdsk/c8t1d0
disk     66  0/0/10/0/0.1.0.39.0.1.1  sdisk    CLAIMED     DEVICE       HP       A6218A
                           /dev/dsk/c5t1d1   /dev/rdsk/c5t1d1
                           /dev/dsk/c8t1d1   /dev/rdsk/c8t1d1
```

- If you want to partition one physical raw disk for OCR and Voting, then you can use the idisk command provided by HP-UX Integrity (cannot be used for PA systems):
  - create a text file on one node

    ksc# vi /tmp/parfile

    ```
    2                       # number of partitions
    EFI 500MB               # size of 1st partition, this standard EFI partition can be used for
    any data
    HPUX 100%               # size of next partition, here we give it all the remaining space
    ```

    The comments here are added only for documentation purpose, using them will lead to an error in the next step.
  - create the two partitions using idisk on the node chosen in the step before

    ksc# idisk -f /tmp/parfile -w /dev/rdsk/c8t0d0

  - Install the special device files for any new disk devices on all nodes:

    ksc/schalke# insf -e -C disk

  - Check on all nodes, that you have now the partitions using the following command:

    ksc/schalke# idisk /dev/rdsk/c8t0d0

    and

    ksc/schalke# /usr/sbin/ioscan -funCdisk

    The output from this command is similar to the following:

    ```
    Class   I   H/W Path       Driver      S/W State      H/W Type     Description
    ============================================================================
    ```

```
disk    4   255/255/0/0.0  sdisk       CLAIMED        DEVICE       HSV100 HP
                           /dev/dsk/c8t0d0 /dev/rdsk/c8t0d0
                           /dev/dsk/c8t0d0 /dev/rdsk/c8t0d0s1
                           /dev/dsk/c8t0d0 /dev/rdsk/c8t0d0s2
```

and

```
ksc/schalke# diskinfo /dev/rdsk/c8t0d0s1
SCSI describe of /dev/rdsk/c8t0d0s1:
vendor: HP
product id: HSV100
type: direct access
size: 512000 Kbytes
bytes per sector: 512
# diskinfo /dev/rdsk/c8t0d0s2
SCSI describe of /dev/rdsk/c8t0d0s2:
vendor: HP
product id: HSV100
type: direct access
size: 536541 Kbytes
bytes per sector: 512
```

- Modify the owner, group, and permissions on the character raw device files on all nodes:
  - OCR:

```
ksc/schalke# chown root:oinstall /dev/rdsk/c8t0d0s1
ksc/schalke# chmod 640 /dev/rdsk/c8t0d0s1
```

  - ASM & Voting disks:

```
ksc/schalke# chown oracle:dba /dev/rdsk/c8t0d0s2
ksc/schalke# chmod 660 /dev/rdsk/c8t0d0s2
```

Optional: ASM Failure Groups:

Oracle lets you configure so-called failure groups for the ASM disk group devices. If you intend to use a normal or high redundancy disk group, you can further protect your database against hardware failure by associating a set of disk devices in a custom failure group. By default, each device comprises its own failure group. However, if two disk devices in a normal redundancy disk group are attached to the same SCSI controller, the disk group becomes unavailable if the controller fails. The controller in this example is a single point of failure. To avoid failures of this type, you could use two SCSI controllers, each with two disks, and define a failure group for the disks attached to each controller. This configuration would enable the disk group to tolerate the failure of one SCSI controller.

- Please note that you cannot create ASM failure groups using DBCA but you have to manually create them by connecting to one ASM instance and using the following sql commands:

```
$ export ORACLE_SID=+ASM1
$ sqlplus / as sysdba
SQL> startup nomount
SQL> create diskgroup DG1 normal redundancy
2 FAILGROUP FG1 DISK '/dev/rdsk/c5t2d0' name c5t2d0,
3 '/dev/rdsk/c5t3d0' name c5t3d0
4 FAILGROUP FG2 DISK '/dev/rdsk/c4t2d0' name c4t2d0,
5 '/dev/rdsk/c4t3d0' name c4t3d0;
DISKGROUP CREATED
SQL> shutdown immediate;
```

Useful ASM v$ views commands:

View	ASM Instance	DB Instance
V$ASM_CLIENT	Shows each database instance using an ASM disk group	Shows the ASM instance if
V$ASM_DISK	Shows disk discovered by the ASM instance, including disks which are not part of any disk group.	Shows a row for each disk
V$ASM_DISKGROUP	Shows disk groups discovered by the ASM instance.	Shows each disk group mo
V$ASM_FILE	Displays all files for each ASM disk group	Returns no rows

7. Preparation for Oracle Software Installation

The Oracle Database 10*g* installation requires you to perform a two-phase process in which you run the Oracle Universal Installer (OUI) twice. The first phase installs Oracle Clusterware (10.2.0.2) and the second phase installs the Oracle Database 10*g* software with RAC. Note that the ORACLE_HOME that you use in phase one is a home for the CRS software which must be different from the ORACLE_HOME that you use in phase two for the installation of the Oracle database software with RAC components.

In case that you have downloaded the software you might have the following files:

- 10gr2_clusterware_hpi.zip Oracle Clusterware
- 10gr2_database_hpi.zip Oracle Database Software

You can unpack the software with the following commands as root user:

```
ksc# /usr/local/bin/unzip 10gr2_clusterware_hpi.zip
```

7.1 Prepare HP-UX Systems for Oracle software installation

- The HP scheduling policy called SCHED_NOAGE enhances Oracle's performance by scheduling Oracle processes so that they do not increase or decrease in priority, or become preempted. On HP, most processes use a time sharing scheduling policy. Time sharing can have detrimental effects on Oracle performance by descheduling an Oracle process during critical operations, for example, holding a latch. HP has a modified scheduling policy, referred to as SCHED_NOAGE, that specifically addresses this issue.

  The RTSCHED and RTPRIO privileges grant Oracle the ability to change its process scheduling policy to SCHED_NOAGE and also tell Oracle what priority level it should use when setting the policy. The MLOCK privilege grants Oracle the ability to execute asynch I/Os through the HP asynch driver. Without this privilege, Oracle9i generates trace files with the following error message: "Ioctl ASYNCH_CONFIG error, errno = 1".

  As root, do the following:

  - If it does not already exist, create the /etc/privgroup file. Add the following line to the file:

    ```
    dba MLOCK RTSCHED RTPRIO
    ```

  - Use the following command syntax to assign these privileges:

    ```
    ksc/schalke# setprivgrp -f /etc/privgroup
    ```

- Create the /var/opt/oracle directory and make it owned by the oracle account. After installation, this directory will contain a few small text files that briefly describe the Oracle software installations and databases on the server. These commands will create the directory and give it appropriate permissions:

  ```
  ksc/schalke# mkdir /var/opt/oracle
  ksc/schalke# chown oracle:oinstall /var/opt/oracle
  ksc/schalke# chmod 755 /var/opt/oracle
  ```

- Create the following Oracle directories:

  - Local Home directory:

    Oracle Clusterware:

    ```
    ksc/schalke# mkdir -p /opt/oracle/product/CRS
    ```

    Oracle RAC:

    ```
    ksc/schalke# mkdir -p /opt/oracle/product/RAC10g
    ksc/schalke# chown -R oracle:oinstall /opt/oracle
    ```

```
ksc/schalke# chmod -R 775 /opt/oracle
```

- Shared CFS directory (commands only from one node):

  Oracle Clusterware:
  ```
  ksc# mkdir -p /cfs/orabin/product/CRS
  ```
  Oracle RAC:
  ```
  ksc# mkdir -p /cfs/orabin/product/RAC10g
  ksc# chown -R oracle:oinstall /cfs/orabin
  ksc# chmod -R 775 /cfs/orabin
  ```

  Oracle Cluster Files:
  ```
  ksc# mkdir -p /cfs/oraclu/OCR
  ksc# mkdir -p /cfs/oraclu/VOTE

  ksc# chown -R oracle:oinstall /cfs/oraclu
  ksc# chmod -R 775 /cfs/oraclu
  ```

  Oracle Database Files:
  ```
  ksc# chown -R oracle:oinstall /cfs/oradata
  ksc# chmod -R 755 /cfs/oradata
  ```

  From each node:
  ```
  ksc/schalke# chmod -R 755 /cfs
  ```

- Set Oracle environment variables by adding an entry similar to the following example to each user startup .profile file for the Bourne or Korn shells, or .login file for the C shell.

```
# @(#) $Revision: 72.2 $
# Default user .profile file (/usr/bin/sh initialization).

# Set up the terminal:
if [ "$TERM" = "" ]
then
eval ` tset -s -Q -m ':?hp' `
else
eval ` tset -s -Q `
fi
stty erase "^H" kill "^U" intr "^C" eof "^D"
stty hupcl ixon ixoff
tabs

# Set up the search paths:
PATH=$PATH:.

# Set up the shell environment:
set -u
trap "echo 'logout'" 0

# Set up the shell variables:
EDITOR=vi
export EDITOR
export PS1=`whoami`@`hostname`\['$ORACLE_SID'\]':$PWD$ '
REMOTEHOST=$(who -muR | awk '{print $NF}')
export DISPLAY=${REMOTEHOST%%:0.0}:0.0

# Oracle Environment
export ORACLE_BASE=/opt/oracle/product
export ORACLE_HOME=$ORACLE_BASE/RAC10g
export ORA_CRS_HOME=$ORACLE_BASE/CRS
export ORACLE_SID=<SID>
export ORACLE_TERM=xterm
export LD_LIBRARY_PATH=$ORACLE_HOME/lib:/lib:/usr/lib:$ORACLE_HOME/rdbms/lib
export PATH=$PATH:$ORACLE_HOME/bin:$ORA_CRS_HOME/bin
export CLASSPATH=$ORACLE_HOME/JRE:$ORACLE_HOME/jlib:$ORACLE_HOME/rdbms/jlib/
$CLASSPATH:$ORACLE_HOME/network/jlib
print ' '
print '$ORACLE_SID: '$ORACLE_SID
print '$ORACLE_HOME: '$ORACLE_HOME
print '$ORA_CRS_HOME: '$ORA_CRS_HOME
print ' '

# ALIAS
alias psg="ps -ef | grep"
alias lla="ll -rta"
alias sq="ied sqlplus '/as sysdba'"
alias oh="cd $ORACLE_HOME"
alias ohbin="cd $ORACLE_HOME/bin"
alias crs="cd $ORA_CRS_HOME"
alias crsbin="cd $ORA_CRS_HOME/bin"
```

7.2 Check Cluster Configuration with Cluster Verification Utility

Cluster Verification Utility (Cluvfy) is a new cluster utility introduced with Oracle Clusterware 10g Release 2. The wide domain of deployment of Cluvfy ranges from initial hardware setup through fully operational cluster for RAC deployment and covers all the intermediate stages of installation and configuration of various components

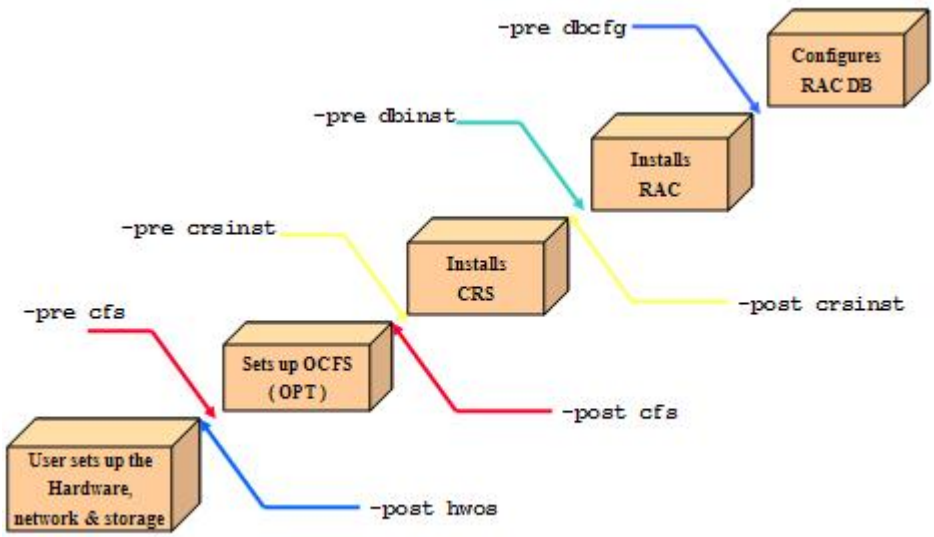
With Cluvfy, you can either

- check the status for a specific component

```
Valid components are:
        nodereach : checks reachability between nodes
        nodecon   : checks node connectivity
        cfs       : checks CFS integrity
        ssa       : checks shared storage accessibility
        space     : checks space availability
        sys       : checks minimum system requirements
        clu       : checks cluster integrity
        clumgr    : checks cluster manager integrity
        ocr       : checks OCR integrity
        crs       : checks CRS integrity
        nodeapp   : checks node applications existence
        admprv    : checks administrative privileges
        peer      : compares properties with peers
```

or

- check the status of your cluster/systems at a specific point (= stage) during your RAC installation. The following picture shows the different stages that can be queried with cluvfy:



The Cluvfy command line utility can be found at the Oracle Clusterware staging are at Clusterware/cluvfy/runcluvfy.sh.

```
md02:root:clusterware/cluvfy ./runcluvfy.sh


USAGE:
cluvfy [ -help ]
cluvfy stage { -list | -help }
cluvfy stage {-pre|-post} <stage-name> <stage-specific options>  [-verbose]
cluvfy comp  { -list | -help }
cluvfy comp  <component-name> <component-specific options>  [-verbose]
```

- Example1: Checking Network Connectivity among all cluster nodes:

```
ksc$ <OraStage>/clusterware/cluvfy/runcluvfy.sh comp nodecon -n ksc,schalke [-verbose]
```

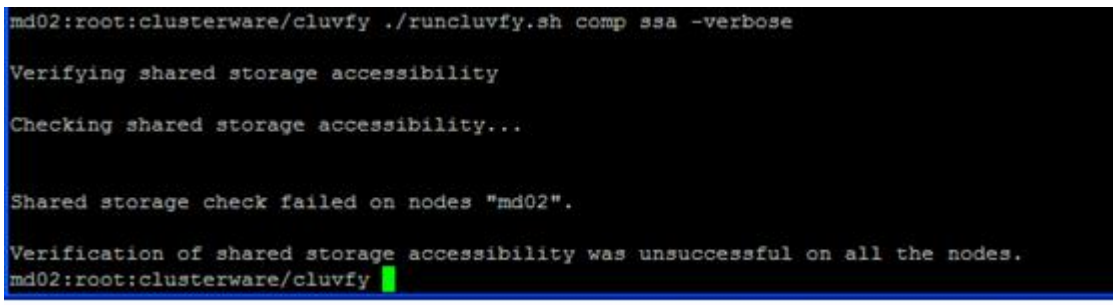
- Example 2: Performing post-checks for hardware and operating system setup

```
ksc$ <OraStage>/clusterware/cluvfy/runcluvfy.sh stage -post hwos -n ksc,schalke [-verbose]
```

- Example 3: Performing Performing pre-checks for cluster services setup

```
ksc$ <OraStage>/clusterware/cluvfy/runcluvfy.sh stage -pre crsinst -n ksc,schalke [-verbose]
```

Note: Current release of cluvfy is not working for shared storage accessibility check on HP-UX. So, this kind of error message are an expected behavior.

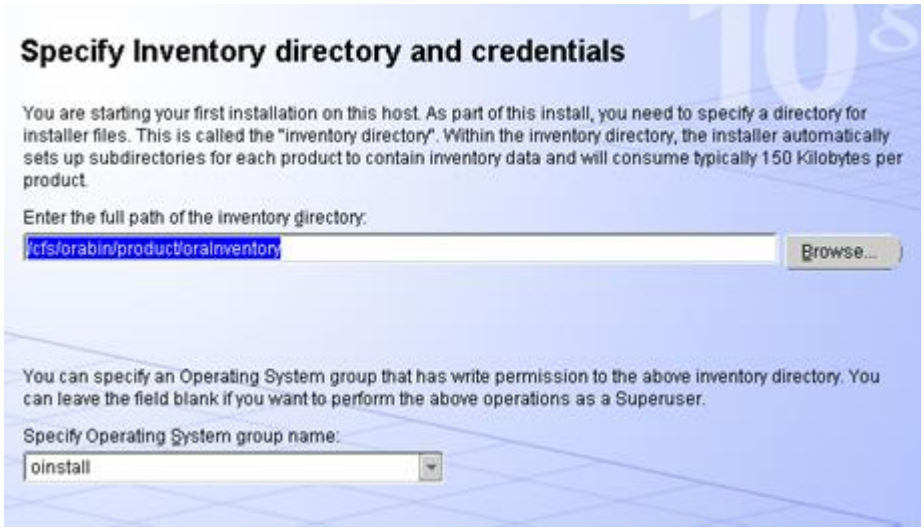


8. Install Oracle Clusterware

This section describes the procedures for using the Oracle Universal Installer (OUI) to install Oracle Clusterware.

Before you install Oracle Clusterware, you must choose the storage option that you want to use for the two Oracle Cluster Files OCR and Voting disk. Again, you cannot use ASM to store these files, because they must be accessible before any Oracle instance starts. If you are not using SGeRAC, you must use raw partitions to store these two files. You cannot use shared raw logical volumes to store these files without SGeRAC.

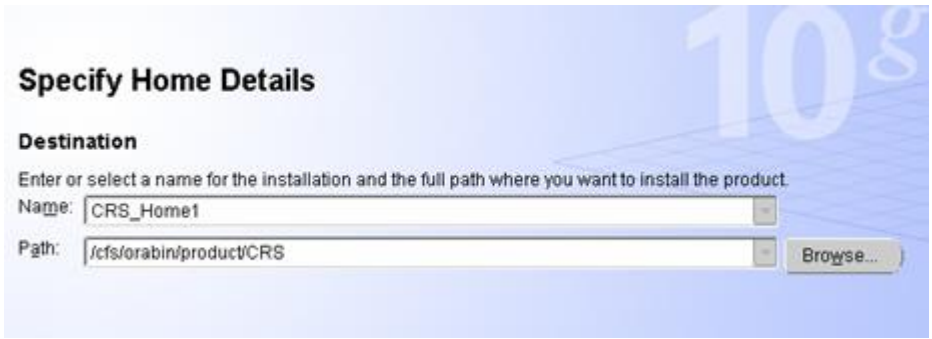
1:	If you are installing Oracle Clusterware on a node that already has a single-instance Oracle Database 10g installation, stop the existing ASM instances and Cluster Synchronization Services (CSS) daemon and use the script $ORACLE_HOME/bin/localconfig delete in the home that is running CSS to reset the OCR configuration information.
2:	Login as Oracle User and set the ORACLE_HOME environment variable to the Oracle Clusterware Home directory. Then start the Oracle Universal Installer from Disk1 by issuing the command `$ ./runInstaller &` Ensure that you have the DISPLAY set.
3:	At the OUI Welcome screen, click Next.
4:	If you are performing this installation in an environment in which you have never installed Oracle database software then the OUI displays the **Specify Inventory Directory and Credentials** page.



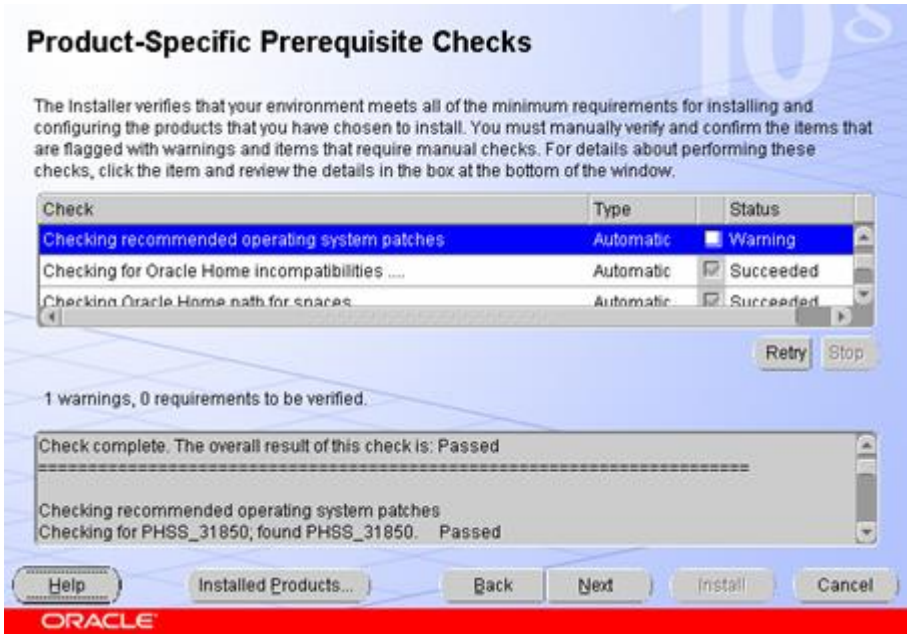
Enter the inventory location and oinstall as the UNIX group name information into the Specify Inventory Directory and Credentials page, click Next.

5: The **Specify Home Details** Page lets you enter the Oracle Clusterware home name and its location in the target destination.

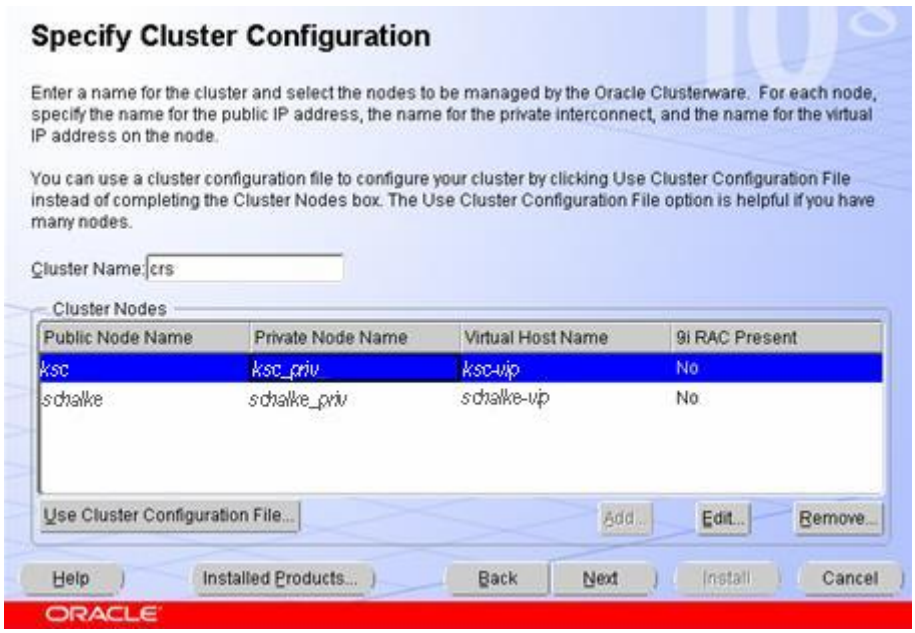
Note that the Oracle Clusterware home that you identify in this phase of the installation is only for Oracle Clusterware software; this home cannot be the same home as the home that you will use in phase two to install the Oracle Database 10g software with RAC.



6: Next, the Product-Specific Prerequisite Check screen comes up. The installer verifies that your environment meets all minumun requirements for installing and configuring Oracle Clusterware. Actually, it uses the Oracle Verification Cluster Utility (Cluvfy). Most probably you'll see a warning at step "Checking recommended operating system patches" as some patches already got replaced by newer ones.



7: In the next **Cluster Configuration** Screen you can specify the cluster name as well as the node information. If HP Serviceguard is running, then you' see these SG cluster configuration. Otherwise, you must select the nodes on which to install Oracle Clusterware. The private node name is used by Oracle for RAC Cache Fusion processing. You need to configure the private node name in the /etc/hosts file of each node in the cluster.

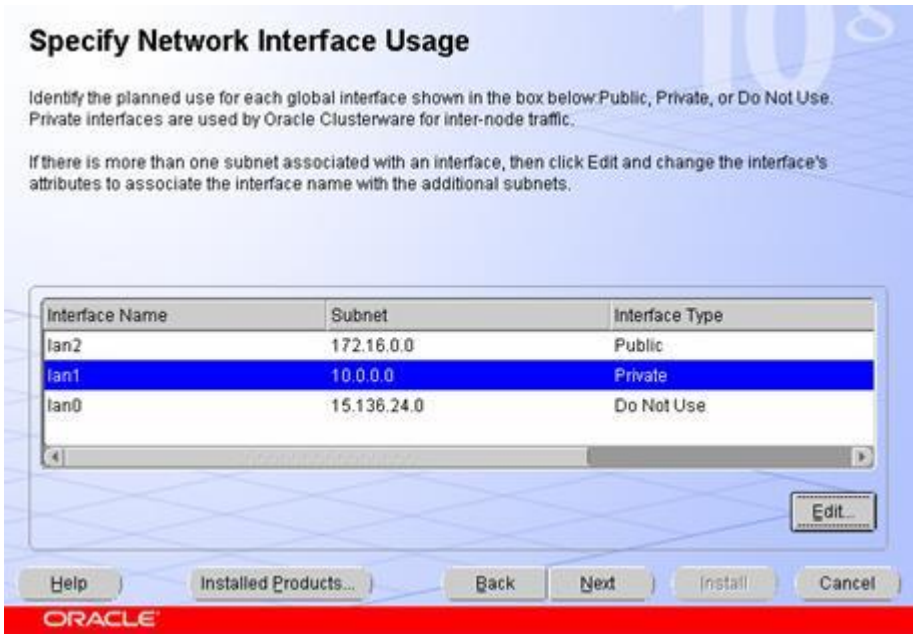


Please note that the interface names associated with the network adapters for each network must be the same on all nodes, e.g. lan0 for private interconnect and lan1 for public interconnect.

Note: in case you have in your /etc/hosts file first full qualified hostname with domain, then you need to give here also this full qualified name or change order in /etc/hosts:

```
172.16.22.41 ksc      ksc.sss.bbn.hp.com
172.16.22.42 schalke      schalke.sss.bbn.hp.com
172.16.22.43 ksc-vip  ksc-vip.sss.bbn.hp.com
172.16.22.44 schalke-vip  schalke-vip.sss.bbn.hp.com
10.0.0.1     ksc_priv
10.0.0.2     schalke_priv
```

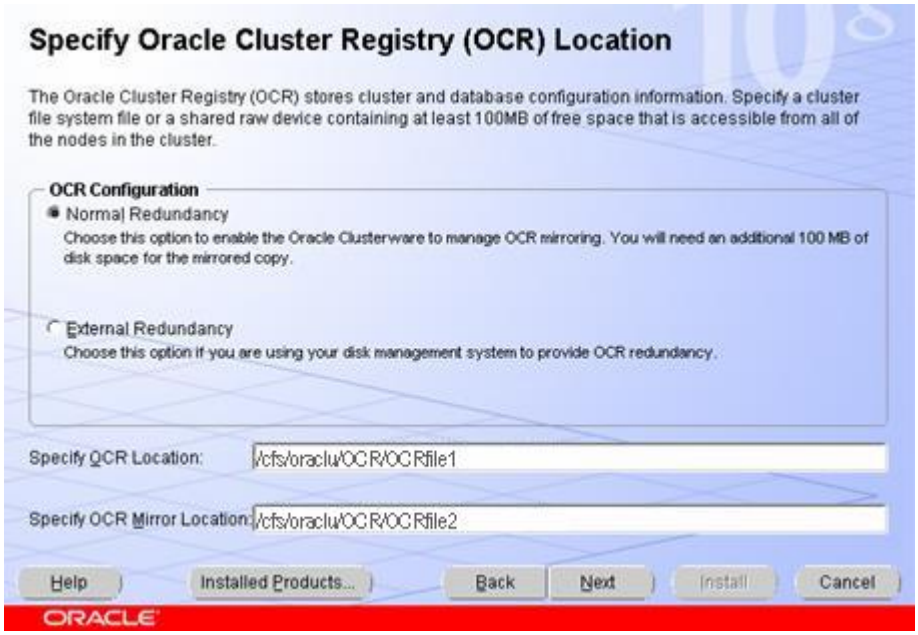
8: In the **Specify Network Interface** page the OUI displays a list of cluster-wide interfaces. If necessary, click edit to change the classification of the interfaces as Public, Private, or Do Not Use. You must classify at least one interconnect as Public and one as Private.



9: When you click Next, the OUI will look for the Oracle Cluster Registry file ocr.loc in the /var/opt/oracle directory. If the ocr.loc file already exists, and if the ocr.loc file has a valid entry for the Oracle Cluster Registry (OCR) location, then the Voting Disk Location page appears and you should proceed to Step 11. Otherwise, the **Oracle Cluster Registry Location** page appears.

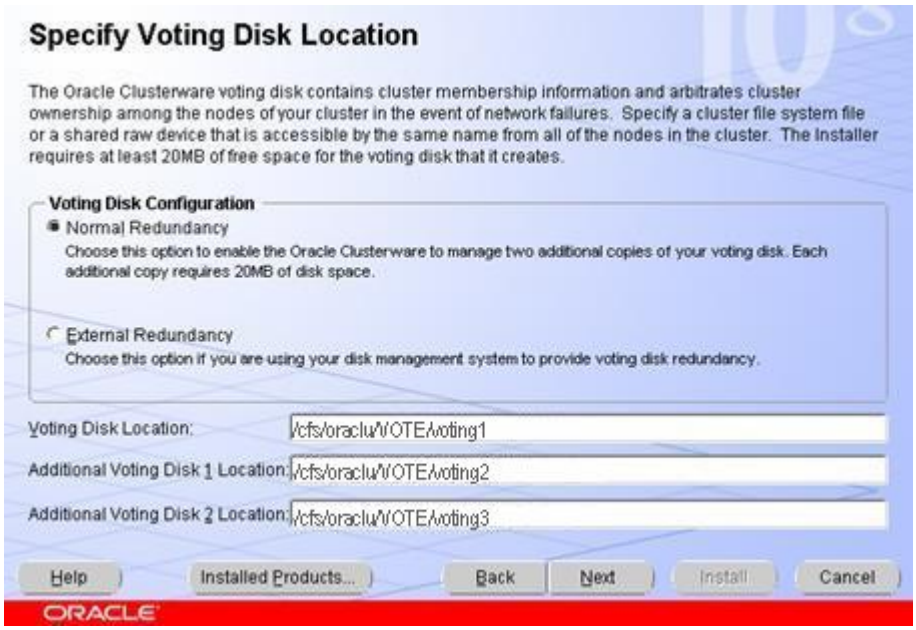
Enter a the complete path for the Oracle Cluster Registry file (not only directory but also including filename). Depending on your chosen deployment model, this might be a CFS location, a shared raw volume or a shared disk (/dev/**r**dsk/cxtxdx).

New with 10g R2, you can let Oracle manage redundancy for this OCR file. In this case, you need to give 2 OCR locations. Assuming the file system has redundancy, e.g. disk array LUNs or CVM mirroring, use of External Redundancy is sufficient and no need for Oracle Clusterware to manage redundancy. Besides, please ensure to place the OCRs on the different file systems for HA reasons.

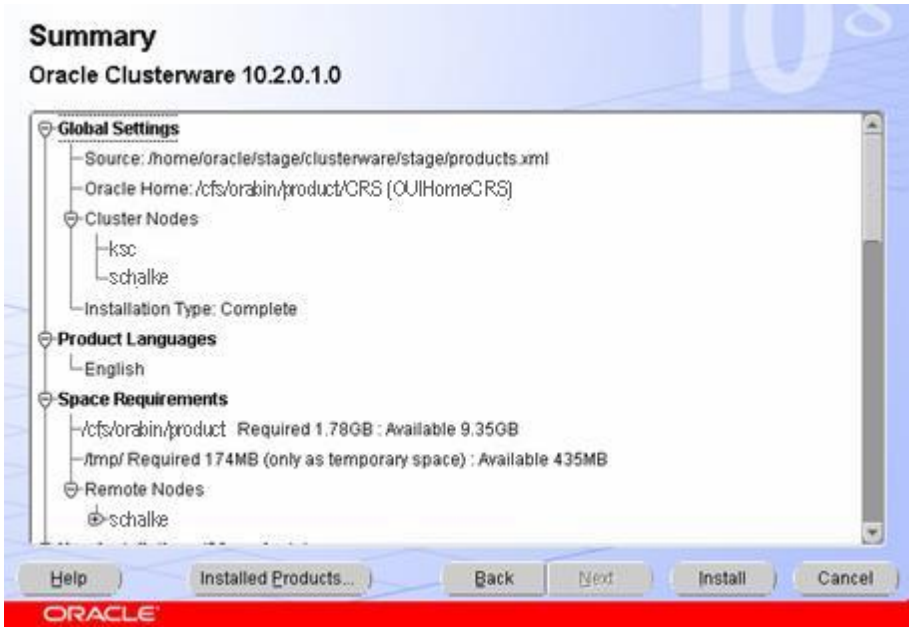


10: On the **Voting Disk Page**, enter a complete path and file name for the file in which you want to store the voting disk. Depending on your chosen deployment model, this might be a CFS location, a shared raw volume or a shared disk (/dev/**r**dsk/cxtxdx).

New with 10g R2, you can let Oracle manage redundancy for the Oracle Voting Disk file. In this case, you need to give 3 locations. Assuming the file system has redundancy, e.g. disk array LUNs or CVM mirroring, use of External Redundancy is sufficient and no need for Oracle Clusterware to manage redundancy. Besides, please ensure to place the Voting Disk files on different file systems for HA reasons.

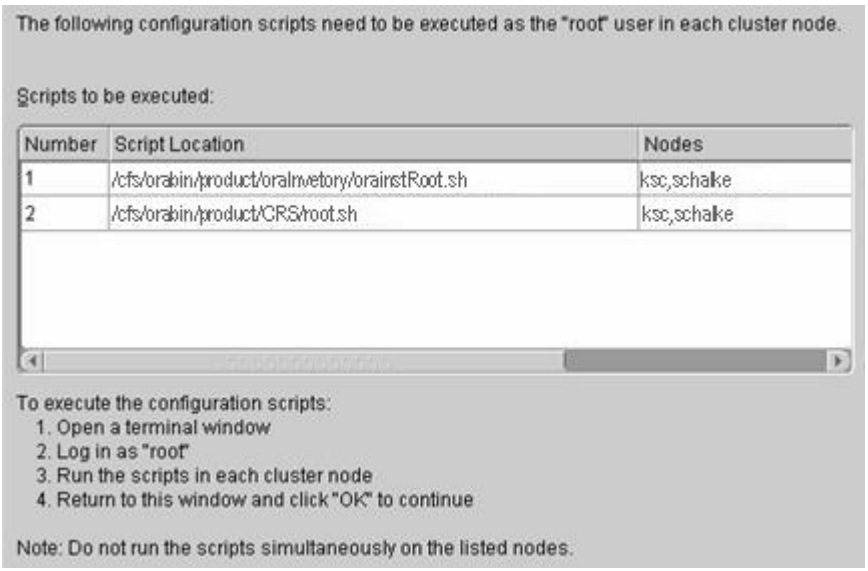


11: Next, Oracle displays a Summary page. Verify that the OUI should install the components shown on the **Summary** page and click Install.



During the installation, the OUI first copies software to the local node and then copies the software to the remote nodes.

12: Then the OUI displays the following windows indicating that you must run the two scripts orainstRoot.sh and root.sh on all nodes.



The scripts root.sh prepares OCR and Voting Disk and starts the Oracle Clusterware. Only start another session of root.sh on another node after the previous root.sh execution completes; do not execute root.sh on more than one node at a time.

```
ksc:root:oracle/product# /cfs/orabin/product/CRS/root.sh
WARNING: directory '/cfs/orabin/product' is not owned by root
WARNING: directory '/cfs/orabin' is not owned by root
WARNING: directory '/cfs' is not owned by root
Checking to see if Oracle CRS stack is already configured
Checking to see if any 9i GSD is up

Setting the permissions on OCR backup directory
Setting up NS directories
Oracle Cluster Registry configuration upgraded successfully
WARNING: directory '/cfs/orabin/product' is not owned by root
WARNING: directory '/cfs/orabin' is not owned by root
WARNING: directory '/cfs' is not owned by root
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 1: ksc ksc_priv ksc
```

```
node 2: schalke schalke_priv schalke
Creating OCR keys for user 'root', privgrp 'sys'..
Operation successful.
Now formatting voting device: /cfs/oraclu/VOTE/voting1
Now formatting voting device: /cfs/oraclu/VOTE/voting2
Now formatting voting device: /cfs/oraclu/VOTE/voting3
Format of 3 voting devices complete.
Startup will be queued to init within 30 seconds.
Adding daemons to inittab
Expecting the CRS daemons to be up within 600 seconds.
CSS is active on these nodes.
        ksc
CSS is inactive on these nodes.
        schalke
Local node checking complete.
Run root.sh on remaining nodes to start CRS daemons.
ksc:root:oracle/product#

schalke:root-/opt/oracle/product # /opt/oracle/product/CRS/root.sh
WARNING: directory '/cfs/orabin/product' is not owned by root
WARNING: directory '/cfs/orabin' is not owned by root
WARNING: directory '/cfs' is not owned by root
Checking to see if Oracle CRS stack is already configured
Checking to see if any 9i GSD is up

Setting the permissions on OCR backup directory
Setting up NS directories
Oracle Cluster Registry configuration upgraded successfully
WARNING: directory '/cfs/orabin/product' is not owned by root
WARNING: directory '/cfs/orabin' is not owned by root
WARNING: directory '/cfs' is not owned by root
clscfg: EXISTING configuration version 3 detected.
clscfg: version 3 is 10G Release 2.
Successfully accumulated necessary OCR keys.
Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.
node <nodenumber>: <nodename> <private interconnect name> <hostname>
node 1: ksc ksc_priv ksc
node 2: schalke schalke_priv schalke
clscfg: Arguments check out successfully.

NO KEYS WERE WRITTEN. Supply -force parameter to override.
-force is destructive and will destroy any previous cluster
configuration.
Oracle Cluster Registry for cluster has already been initialized
Startup will be queued to init within 30 seconds.
Adding daemons to inittab
Expecting the CRS daemons to be up within 600 seconds.
CSS is active on these nodes.
ksc
schalke
CSS is active on all nodes.
Waiting for the Oracle CRSD and EVMD to start
Oracle CRS stack installed and running under init(1M)
Running vipca(silent) for configuring nodeapps

Creating VIP application resource on (2) nodes ...
Creating GSD application resource on (2) nodes ...
Creating ONS application resource on (2) nodes ...
Starting VIP application resource on (2) nodes ...
Starting GSD application resource on (2) nodes ...
Starting ONS application resource on (2) nodes ...


Done.
schalke:root-/opt/oracle/product #
```

As highlighted in red, with R2 Oracle now configures the NodeApps already at the end of the last root.sh script execution in silent mode.

13: Next, the **Configurations Assistants** screen comes up. OUI runs the Oracle Notification Server Configuration Assistant, Oracle Private Interconnect Configuration Assistant, and Cluster Verification Utility. These programs run without user intervention.

14: When the OUI displays the **End of Installation** page, click Exit to exit the Installer.

15: Verify your CRS installation by executing the `olsnodes` command from the

	$ORA_CRS_HOME/bin directory:
	`# olsnodes -n`
	`ksc 1`
	`schalke 2`

16:	Now you should see the following processes running: - oprocd -- Process monitor for the cluster. Note that this process will only appear on platforms that do not use HP Serviceguard with CSS. - evmd -- Event manager daemon that starts the racgevt process to manage callouts. - ocssd -- Manages cluster node membership and runs as oracle user; failure of this process results in cluster restart. - crsd -- Performs high availability recovery and management operations such as maintaining the OCR. Also manages application resources and runs as root user and restarts automatically upon failure. You can check whether the Oracle processes evmd, occsd, and crsd are running by issuing the following command. **`# ps -ef | grep d.bin`**
	At this point, you have completed phase one, the installation of Cluster Ready Services 

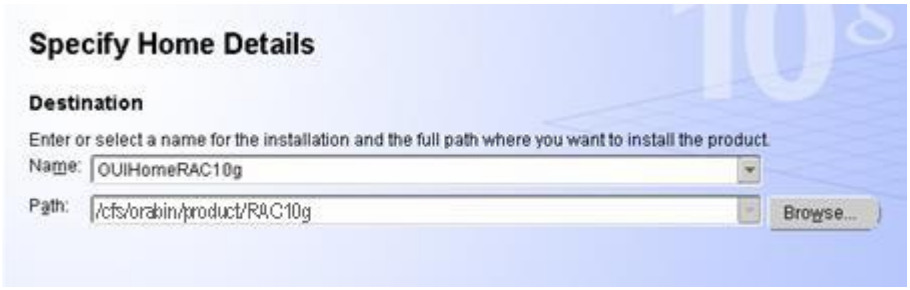
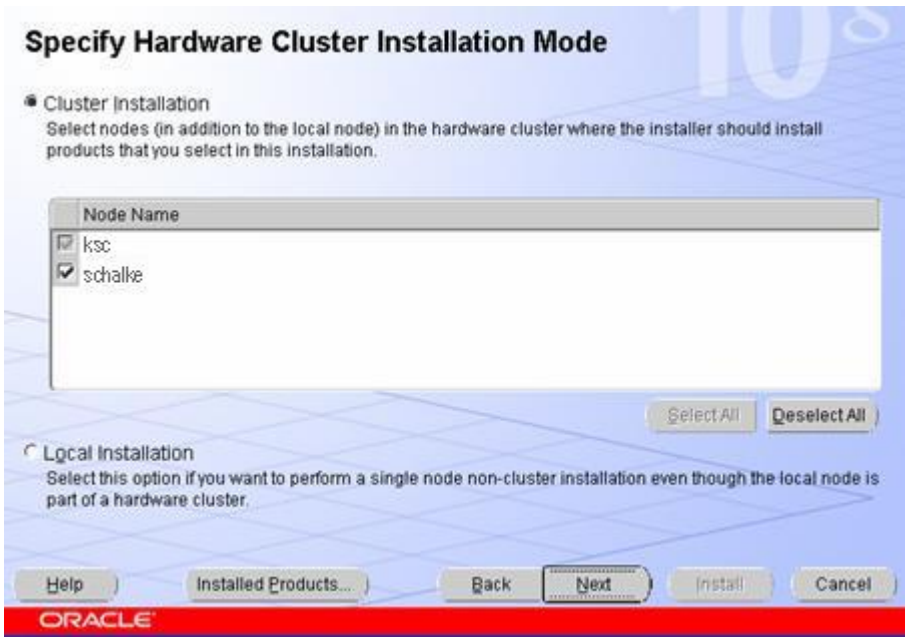
	Please note that Oracle added the following three lines to the automatic startup file /etc/inittab.
	
	`h1:3:respawn:/etc/init.d/init.evmd run >/dev/null 2>&1 </dev/null`
	`h2:3:respawn:/etc/init.d/init.cssd fatal >/dev/null 2>&1 </dev/null`
	`h3:3:respawn:/etc/init.d/init.crsd run >/dev/null 2>&1 </dev/null`
	Oracle Support recommends NEVER modifying these entries in the inittab or modifying the init scripts unless you use this method to stop a reboot loop or are given explicit instructions from Oracle support.

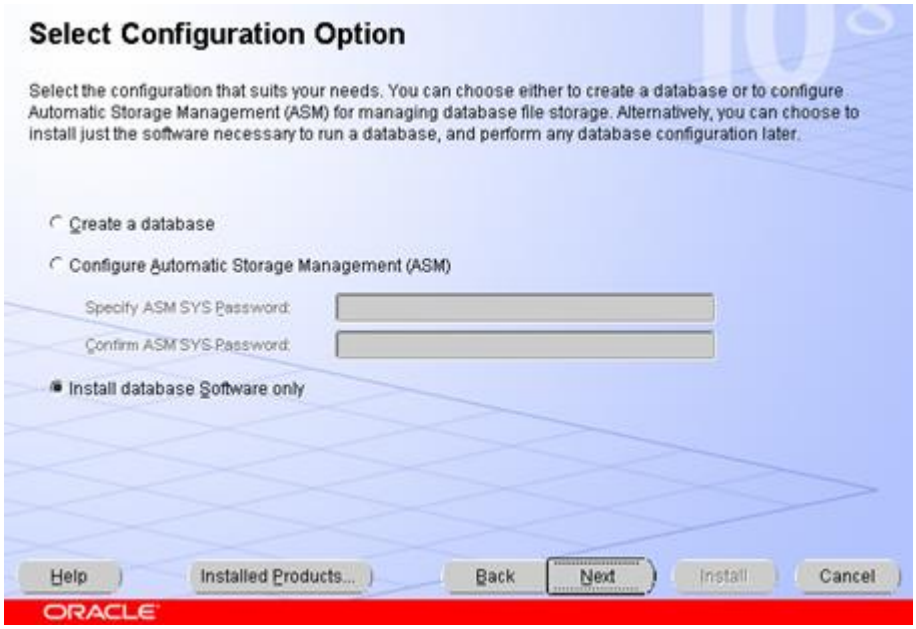
	To ensure that the Oracle Clusterware install on all the nodes is valid, the following should be checked on all the nodes: - **`$ $ORA_CRS_HOME/bin/crsctl check css`** `CSS daemon appears healthy`

9. Installation of Oracle Database RAC 10*g* R2

This part describes phase two of the installation procedures for installing the Oracle Database 10g with Real Application Clusters (RAC).

1:	Login as Oracle User and set the ORACLE_HOME environment variable to the Oracle Home directory. Then start the Oracle Universal Installer from Disk1 by issuing the command `$ ./runInstaller &` Ensure that you have the DISPLAY set.
2:	When the OUI displays the Welcome page, click Next, and the OUI displays the **Specify File Locations** page. The Oracle home name and path that you use in this step must be different from the home that you used during the Oracle Clusterware installation in phase

	one. 
3:	On the **Specify Hardware Cluster Installation Mode** page, select an installation mode. The Cluster Installation mode is selected by default when the OUI detects that you are performing this installation on a cluster.  When you click Next on the Specify Hardware Cluster Installation page, the OUI verifies that the Oracle home directory is writable on the remote nodes and that the remote nodes are operating.
4:	Next, the Product-Specific Prerequisite Check screen comes up. The installer verifies that your environment meets all minumun requirements for installing and configuring a RAC10g database. Actually, it uses the Oracle Verification Cluster Utility (Cluvfy). Most probably you'll see a warning at step "Checking recommended operating system patches" as some patches already got replaced by newer ones.
5:	On the **Select Configuration Option** page you can choose to either create a database, to configure Oracle ASM or to perform a software only installation. New with R2, you can install ASM into an own ORACLE_HOME to be decoupled from the database binaries. If you would like to do this, you need to select Oracle ASM. Please note that in this case the Oracle listener will be registered in CRS with the ORACLE_HOME of ASM which you need to manually change later to the database ORACLE_HOME.

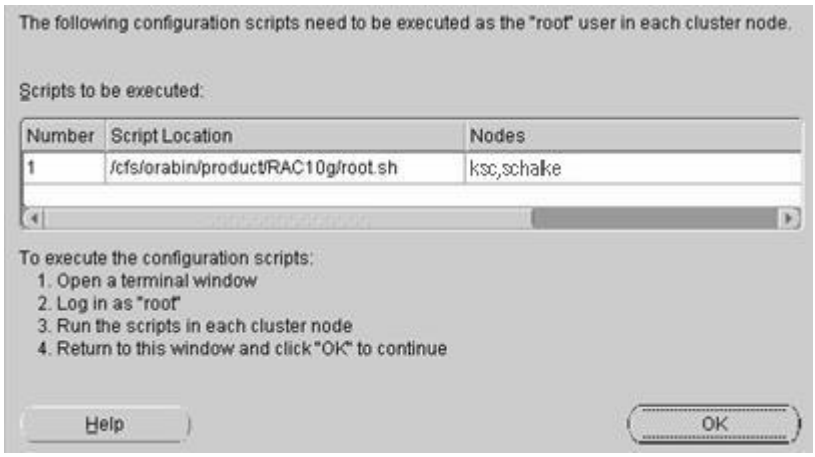


Here we recommend only to the software and not to create a starter database. We will create a database later with the Database Configuration Assistant.

6: The **Summary Page** displays the software components that the OUI will install and the space available in the Oracle home with a list of the nodes that are part of the installation session. Verify the details about the installation that appear on the Summary page and click Install or click Back to revise your installation.

During the installation, the OUI copies software to the local node and then copies the software to the remote nodes.

7: Then, OUI prompts you to run the root.sh script on all the selected nodes.



8. The OUI displays the **End of Installation** page, click Exit to exit the Installer.

9: You can check the installation with the command OCR commands $ORA_CRS_HOME/bin/ocrdump, $ORA_CRS_HOME/bin/ocrcheck, $ORA_CRS_HOME/bin/crs_stat. The crs_stat command will provide a description of the Oracle environment available in the cluster.

`# crs-stat -t` gives you a more compact output.

In addition we would recommend to copy the sample 10g CRS resource status query script from the Oracle Metalink Note:259301.1:

```
#!/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=/sbin/awk                           # if not available use /usr/bin/awk
ORA_CRS_HOME=/opt/oracle/product/CRS

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

# Table body:
$ORA_CRS_HOME/bin/crs_stat $QSTAT | $AWK \
'BEGIN { FS="="; state = 0; }
$1~/NAME/ && $2~/'$RSC_KEY'/ {appname = $2; state=1};
state == 0 {next;}
$1~/TARGET/ && state == 1 {apptarget = $2; state=2;}
$1~/STATE/ && state == 2 {appstate = $2; state=3;}
state == 3 {printf "%-45s %-10s %-18s\n", appname, apptarget, appstate; state=0;}'
```

10: Oracle Disk Manager (ODM) Configuration: ODM is only required when using Oracle RAC with CFS and SGeRAC.

Currently, there is a reported Oracle bug #5103839 with DBCA after enabling ODM (linking the ODM library). The workaround is to create the database first and then link ODM.

- Check that the VRTSodm package is installed.

```
# swlist VRTSodm
# VRTSodm 4.1 VERITAS Oracle Disk Manager
VRTSodm.ODM-KRN 4.1 VERITAS ODM kernel files
VRTSodm.ODM-MAN 4.1 VERITAS ODM manual pages
VRTSodm.ODM-RUN 4.1 VERITAS ODM commands
```

- Check libodm.sl

```
# ll /opt/VRTSodm/lib/libodm.sl
-r-xr-xr-x 1 root sys 78176 May 20 2005 /opt/VRTSodm/lib/libodm.sl
```

- Configure Oracle to use ODM: you need to link the Oracle Disk Manager library into ORACLE_HOME for Oracle 10g (as oracle user):
  - For Integrity systems:

```
$ rm ${ORACLE_HOME}/lib/libodm10.so
$ ln -s /opt/VRTSodm/lib/libodm.sl ${ORACLE_HOME}/lib/libodm10.so
```

  - For PA-Risc systems:

```
$ rm ${ORACLE_HOME}/lib/libodm10.sl
$ ln -s /opt/VRTSodm/lib/libodm.sl ${ORACLE_HOME}/lib/libodm10.sl
```

- Configure Oracle to Stop using ODM Library:
  - For Integrity systems:

```
$ rm ${ORACLE_HOME}/lib/libodm10.so  # this only removes the symbolic link
to /opt/VRTSodm/lib/libodm.sl
$ ln -s ${ORACLE_HOME}/lib/libodmd10.so ${ORACLE_HOME}/lib/libodm10.so
```

  - For PA-Risc systems:

```
$ rm ${ORACLE_HOME}/lib/libodm10.sl # this only removes the symbolic link
to /opt/VRTSodm/lib/libodm.sl
$ ln -s ${ORACLE_HOME}/lib/libodmd10.sl ${ORACLE_HOME}/lib/libodm10.sl
```

10. Configure the Oracle Listeners

First we recommend to configure the Oracle Listener using the Oracle Net Configuration Assistant:

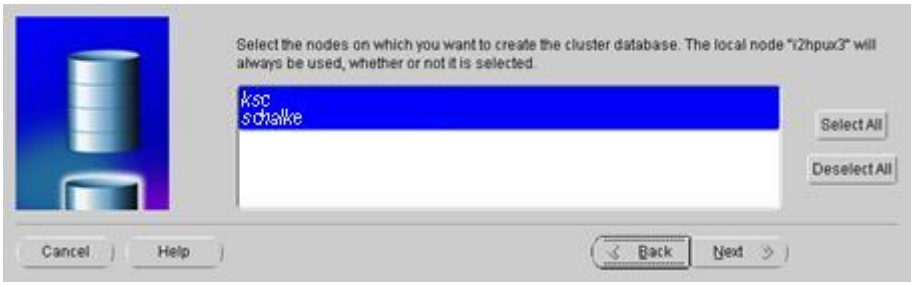
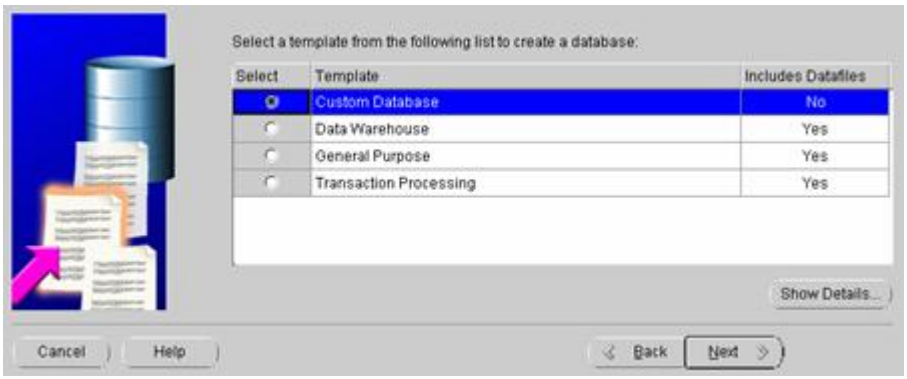
1:	Connect as oracle user and start the Oracle Net Configuration Assistant by issuing the command `$ netca &` Ensure that you have the DISPLAY set.
2:	Select 'Cluster Configuration' and click Next.
3:	The next screen lets you select the nodes for which to configure the Oracle listener. Select all nodes, and click Next.
4:	At the next page, select 'Listener configuration'
5:	Select 'Add', and click Next.
6:	Keep default name 'Listener', and click Next.
7:	Keep 'TCP' as selected protocol, and click Next.
8:	Keep standard protocol '1521', and click Next.
9:	Say 'no', when you get asked to configure additional listener and Exit NetCa.

10:	You can verify the Listener set-up through $ORA_CRS_HOME/bin/crs_stat (or the script you downloaded at chapter 9, step 9):
	
	`$ /opt/oracle/product/CRS/bin$ ./crs_great`
	`HA Resource Target State`
	`----------- ------ -----`
	`ora.ksc.LISTENER_KSC.lsnr ONLINE ONLINE on ksc`
	`ora.ksc.gsd ONLINE ONLINE on ksc`
	`ora.ksc.ons ONLINE ONLINE on ksc`
	`ora.ksc.vip ONLINE ONLINE on ksc`
	`ora.schalke.LISTENER_SCHALKE.lsnr ONLINE ONLINE on schalke`
	`ora.schalke.gsd ONLINE ONLINE on schalke`
	`ora.schalke.ons ONLINE ONLINE on schalke`
	`ora.schalke.vip ONLINE ONLINE on schalke`
	
	The listener names have now the syntax 'LISTENER_<nodename>':
	
	`$ lsnrctl status LISTENER_SCHALKE`

11:	As you can see in the listener configuration file $ORACLE_HOME/network/admin/listener.ora, the Listener process is using the Oracle VIP address.

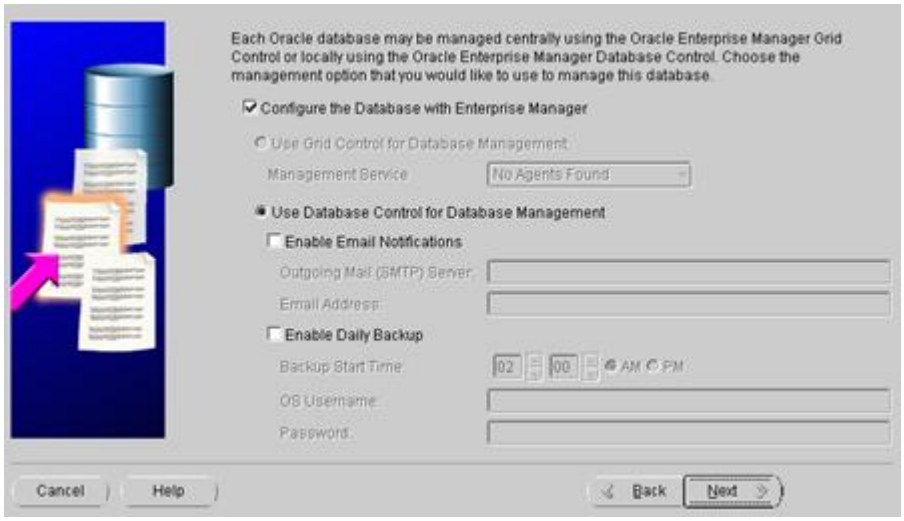
11. Create a RAC DB on CFS using Database Configuration

Assistant

1:	Connect as oracle user and start the Database Configuration Assistant by issuing the command `$ dbca &` Ensure that you have the DISPLAY set.
2:	The first page that the DBCA displays is the **Welcome** page for RAC. The DBCA displays this RAC-specific Welcome page only if the Oracle home from which it is invoked was cluster installed. If the DBCA does not display this Welcome page for RAC, then the DBCA was unable to detect whether the Oracle home is cluster installed. Select Real Application Clusters database, click Next.
3:	At the **Configure Database Options** page select 'Create a database' and click Next.
4:	At the **Node Selection** page select the nodes that you want to configure as members of your cluster database, then click Next. 
5:	The templates on the **Database Templates** page are Custom Database, Transaction Processing, Data Warehouse, and General Purpose. The Custom Database template does not include datafiles or options specially configured for a particular type of application. Select the General Purpose template for your cluster database, and click Next.
6:	At the **Database Identification** page enter the global database name and the Oracle system identifier (SID) prefix for your cluster database and click Next.  The SID prefix must begin with an alphabetical character and contain no more than 5 characters on UNIX-based systems.
7:	On the **Management Options** page, you can choose to manage your database with Enterprise Manager. On UNIX-based systems only, you can also choose either the Grid Control or Database Control option if you select Enterprise Manager database management. If you select Enterprise Manager with the Grid Control option and DBCA discovers agents running on the local node, then you can select the preferred agent from a

list.

Here, do the following selection, and click Next.

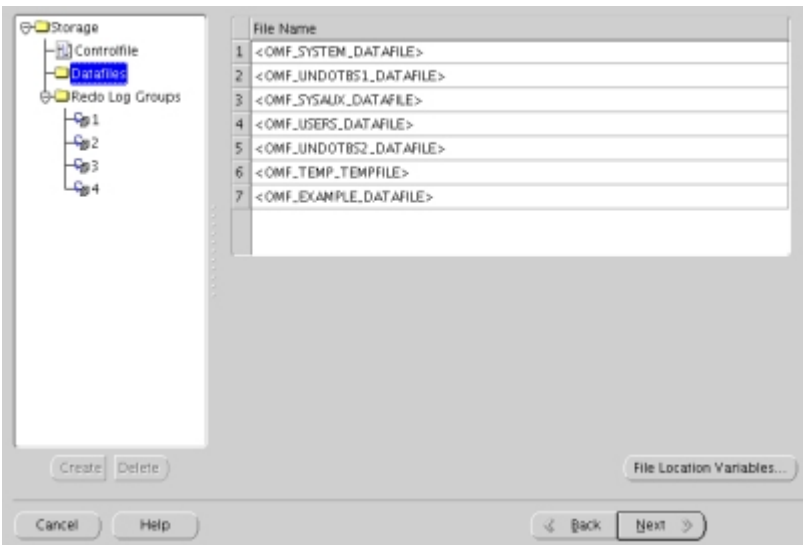


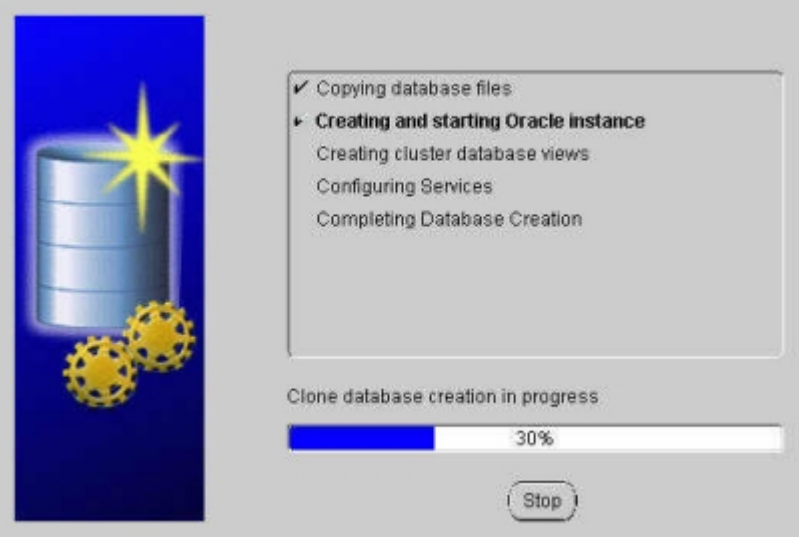
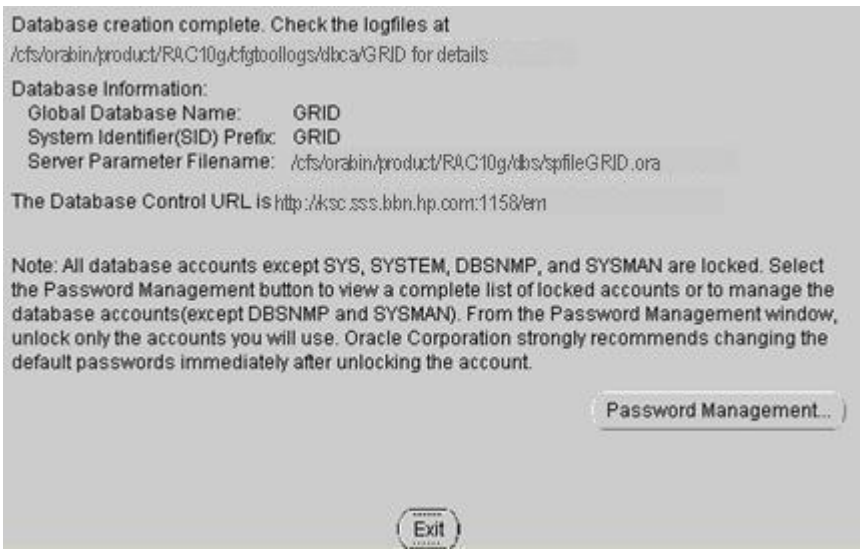
8: The the **Database Credentials** page you can enter the passwords for your database. You can enter the same or different passwords for the users SYS and SYSTEM, plus DBSNMP and SYSMAN if you selected Enterprise Manager on the Management Options page.

9: At the **Network Configuration** page select 'Register this database with all the listeners'.

10: At the **Storage Options** page you can select a storage type for the database. Please select the storage option that applies to your chosen deployment model.

Here, we show an installation for HP SG CFS for RAC.

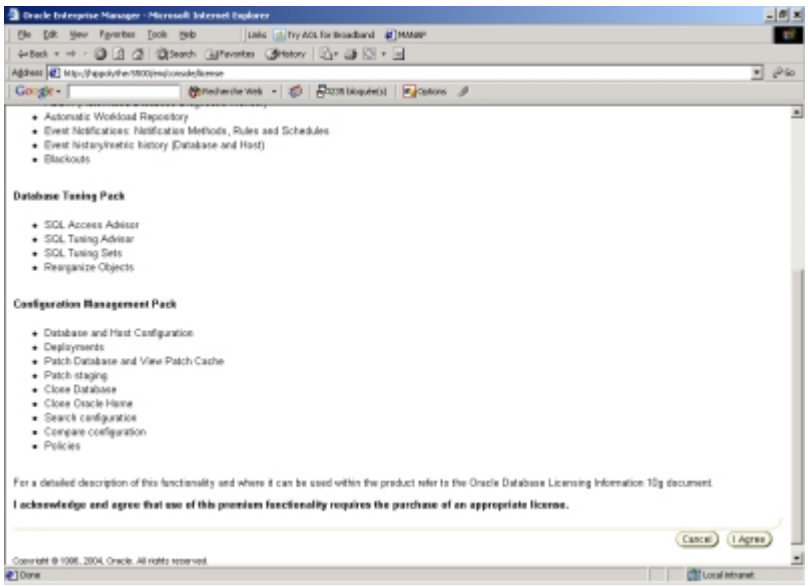
11:	On the **Recovery Configuration** page, you can specify the Flash Recovery Area and Enable Archiving. To use a flash recovery area, Oracle recommends that you create two separate disk groups: one for the database area and one for the recovery area.
12:	On the **Database Content** page, you can select whether to include the sample schemas in your database and to run custom scripts as part of the database creation processing.
13:	To create services on the **Database Services** page, expand the Services tree. Oracle displays the global database name (here GRID) in the top left-hand corner of the page. Select the global database name and click Add to add services to the database (here GRID1 and GRID2).  On the right-hand you see the service preferences for each service. Change the instance preference (Not Used, Preferred, or Available) and TAF policies for the service as needed. Repeat this procedure for each service and when you are done configuring services for your database, click Next.
14:	At the **Initialization Parameters** page you can review and adjust the initialization parameters.
15:	If you selected a preconfigured database template, such as the General Purpose template, then the DBCA displays the control files, datafiles, and redo logs on the **Database Storage** page. Select the folder and the file name underneath the folder to edit the file name.  After you complete your entries on the Database Storage page, click Next.
16:	On the **Creation Options** page, select Create Database and click Finish. You can also

	save this configuration as a database template file and use this template later to create a database.
17:	Review the **Summary** dialog information and click OK to create the database.   Change the db account password if necessary and click Exit.
18:	Congratulations ... you have now your RAC database configured :-)
19:	You can check the installation with the command $ORA_CRS_HOME/bin/crs_stat. This command will provide a description of the Oracle environment available in the cluster.

12. Oracle Enterprise Manager 10g Database Control

When you are installing the database software, the OUI also installs the software for Oracle Enterprise Manager Database Control and integrates this tool into your cluster environment. Once installed, Enterprise Manager Database Control is fully configured and operational for RAC. You can also install Enterprise Manager Grid Control onto other client machines outside your cluster to monitor multiple RAC and single-instance Oracle database environments.

1:	Start the DBConsole agent on one of the cluster nodes as Oracle user:

	$ emctl start dbconsole (other options: emctl start| stop| status| setpasswd dbconsole)
2:	To connect to the Oracle Enterprise Manager Database Control (default port 5500) open the following URL in our web browser: **http://<node1a>:1158/em**
3.	Log-in as sys and sysdba profile.
4.	Accept the licensing. 
4.	Now you will get to the OEM Database Control Home Page.

13. Implementation of SG Packages Framework for RAC 10g

HP offers an Oracle Integration Framework which contains a set of templates that can be used to help integrate Oracle RAC10g R2 with SGeRAC version 11.17.

This framework provides a uniform, easy-to-mange and intuitive method to

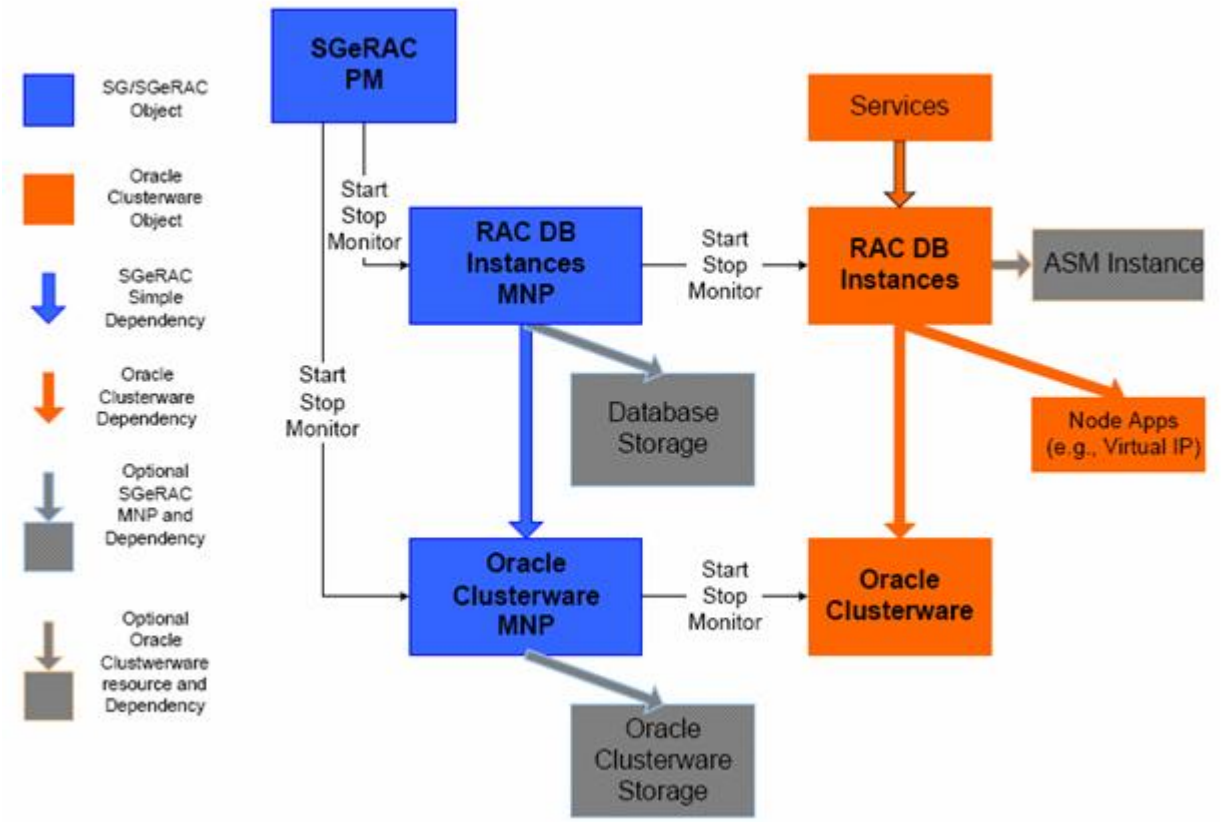
- co-ordinate the operation of this combined software stack
- manage all the storage options supported – SLVM, CFS/CVM and ASM over SLVM

The included templates are based on the following two new Serviceguard features introduced with 11.17:

- Multi-Node Package (MNP):
  - SG Package manager enhanced to allow a package to run on more than one node at once.
    - Each instance of a multi-node package behaves as a normal package but does not failover.
    - Overall package status is a summary of the status of package instances.
  - Commands enhanced to manage multi-node packages
    - cmrunpkg/cmhaltpkg enhanced to allow multiple –n parameters to be specified.
    - cmviewcl adds new display format for multi-node packages
- Simple Dependency

- Enforce package start/halt ordering
- If A depends on B and B fails, A is halted

This integration between SGeRAC and RAC10g can be configured as shown here:



The use of the framework is restricted to a maximum SGeRAC cluster size of 8 nodes.

The framework may not be used in Extended Cluster, Metrocluster or Continentalclusters configurations.

13.1 Configuration of Framework

- Assume SGeRAC, Oracle clusterware and RAC database(s) are already installed and configured
- Configure Oracle clusterware and the RAC database instances to be started on demand, instead of being automatically started:
  - Disable Oracle clusterware startup on boot
  - Disable auto-start of RAC db instances by CRS
  - See Oracle Metalink Note:298073.1
- Halt the databases and Oracle Clusterware
- Configure the framework, using the script bundle delivered by HP's ACSL organization and follow instructions in the README file in the script bundle.
  - This scripts can be downloaded from the HP internal web page at http://haweb.cup.hp.com/ATC/Web/Whitepapers/content/rac10g_fwk_020606.cpio. A more detailed white paper can be found at http://haweb.cup.hp.com/ATC/Web/Whitepapers/content/10gRAC_Framework.pdf.
  - In the package directory, copy the framework files for Oracle Clusterware or the RAC instance
  - Configure the RAC10g specific logic of the MNP in the framework configuration file.
  - Distribute the framework files to the other nodes of the MNP using remote copy commands.

- In the package directory, create SGeRAC package templates using cmmakepkg, and configure the SGeRAC specific part of the MNP in the package configuration and control script files
- Apply the package configuration using the cmcheckconf - cmapplyconf sequence.

14. Tips & Tricks

Oracle Clusterware:

- CRS and 10g Real Application Clusters; Oracle Metalink Note: 259301.1
- How to start the 10g CRS ClusterWare; Oracle Metalink Note Note:309542.1
- How to Clean Up After a Failed CRS Install; Oracle Metalink Note:239998.1
- How to Stop the Cluster Ready Services (CRS); Oracle Metalink Note:263897.1
- Stopping Reboot Loops When CRS Problems Occur; Oracle Metalink Note: 239989.1
- Troubleshooting CRS Reboots; Oracle Metalink Note:265769.1
- CRS 10g Diagnostic Collection Guide; Oracle Metalink Note:272332.1
- What Are The Default Settings For MISSCOUNT In 10g RAC ?, Oracle Metalink Note 300063.1
- CSS Timeout Computation in 10g RAC 10.1.0.3; Oracle Metalink Note:294430.1
- HOW TO REMOVE CRS AUTO START AND RESTART FOR A RAC INSTANCE; Oracle Metalink Note:298073.1

VIPs / Interconnect / Public Interface:

- Configuring the HP-UX Operating System for the Oracle 10g VIP; Oracle Metalink Note:296874.1
- How to Configure Virtual IPs for 10g RAC; Oracle Metalink Note:264847.1
- How to change VIP and VIP/Hostname in 10g ; Oracle Metalink Note:271121.1
- Modifying the VIP of a Cluster Node; Oracle Metalink Note:276434.1
- How to Change Interconnect/Public Interface IP Subnet in a 10g Cluster; Oracle Metalink Note:283684.1
- Troubleshooting TAF Issues in 10g RAC; Oracle Metalink Note:271297.1
- Oracle 10g VIP (Virtual IP) changes in Oracle 10g 10.1.0.4; Oracle Metalink Note:296878.1

OCR / Voting:

- How to Restore a Lost Voting Disk in 10g; Oracle Metalink Note:279793.1
- Repairing or Restoring an Inconsistent OCR in RAC; Oracle Metalink Note:268937.1

ASM:

- ASM Instance Shuts Down Cleanly On Its Own; Oracle Metalink Note:277274.1

Migration:

- How to migrate from 9iRAC to RAC10; CTC Technical Paper

Adding/Removing Nodes:

- Adding a Node to a 10g RAC Cluster; Oracle Metalink Note: 270512.1
- Removing a Node from a 10g RAC Cluster; Oracle Metalink Note:269320.1

15. Known Issues & Bug Fixes

We highly recommend to install Oracle Patch Set 10.2.0.2 as this proved to be much more stable.

Oracle VIP and HP SG Relocatable IP using same LAN interface - Oracle Bug #4699597

- The Oracle VIP feature works at a low level with the device files for the network interface cards, and as a result might clash with any other SG Relocatable IP addresses also configured for the same public NIC. Therefore, it has not been supported to configure the public NIC used for Oracle VIP also for any other SG Relocatable IP address.
  - This issue has been addressed with Oracle bug fix #4699597 which ensures that Oracle VIP starts with logical interface number 801 (ie. lan1:801) so that there will not be any conflict with SG's Relocatable IP's.
  - This Oracle bug fix #4699597 is already available for 10.2.0.2 HP-UX Integrity and will be available for PA-RISC with 10.2.0.3.
- See Oracle Metalink Note 296874.1 "Configuring the HP-UX Operating System for the Oracle 10g VIP"

DB Creation fails with ODM library - Oracle Bug #5103839

- Currently, there is a reported Oracle bug #5103839 with DBCA after enabling ODM (linking the ODM library). The workaround is to create the database first and then link ODM (see chapter 9 step 10).

CSS does not write any dump in case of system reboot

- PHKL_34374 when used with 10.2.0.2 patch set for 11.23 will result in a TOC (with dump) instead of a reboot.
- PHKL_34374 needs to be installed before the Oracle patchset otherwise it will default to reboot.
- Note that it does not fix any problem but provides a dump when Oracle's clusterware detects an issue.