



Oracle10g™ Real Application Clusters Release 2 Install with Power Linux® (RHEL AS4 U3) and NetApp Storage

**Niranjan Mohapatra, Network Appliance, Inc.
January 2007 | TR-3527-0107**

Table of Contents

| | |
|---|-----------|
| 1. Introduction | 3 |
| 2. Assumptions | 3 |
| 3. The Server/System Environment | 3 |
| 4. Requirements | 4 |
| 4.1 Hardware Used for Tests..... | 4 |
| 4.2 Software Used..... | 4 |
| 5. Setup for NetApp Storage | 4 |
| 6. Patches, Environment, and OS Settings | 5 |
| 6.1 Patches..... | 5 |
| 6.2 OS Settings..... | 6 |
| 7. Preinstall Setup Tasks (Cluster Nodes) | 7 |
| 8. Installation Procedure | 9 |
| 8.1 Preparing to Install the Oracle RAC 10g on Cluster Nodes..... | 9 |
| 8.2 Installing the Oracle RAC 10g Cluster Ready Services (CRS)..... | 10 |
| 8.3 Installing Oracle RAC 10g software..... | 25 |
| 9. Appendix | 32 |
| 10. Acknowledgements | 32 |
| 11. Disclaimer | 32 |

1. Introduction

This technical report covers the installation of NetApp storage in NFS environment for Power Linux on IBM pSeries servers. The servers are running RHEL AS 4 U3 (Power Linux) operating system. The database is Oracle10g Real Application Clusters release 2 (hereafter referred to as Oracle10gR2 RAC) with Oracle Cluster Ready Services (hereafter referred to as Oracle CRS). This is now a certified configuration and, as such, the components presented in this paper have to be used in the same combination to gain support from all parties involved. The only exception to this is the application of certain patches (as defined and required by all the vendors in this configuration). This document will also cover the patches and recommendations for running Oracle10gR2 RAC on NetApp Storage in NFS environment.

2. Assumptions

We assume that readers are familiar with Oracle10g release 2 RAC and the operation of Network Appliance™ storage systems. We also assume that readers are familiar with the operation of the Power Linux environment and installation of Oracle® patches and any relevant Power Linux rpms. It is also important to be familiar with all networking terminology and implementations.

3. The Server/System Environment

The configuration presented in this document is based on the Oracle10gR2 RAC certification environment specified by Oracle and Network Appliance.

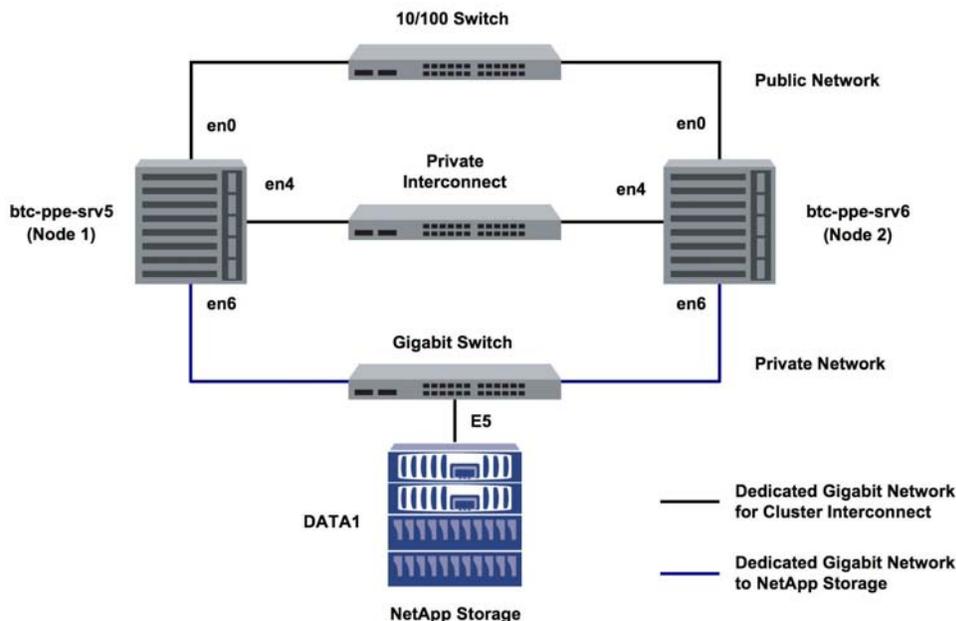


Figure 1) Oracle10gR2 RAC on IBM pSeries PowerPC servers with Network Appliance storage.

Figure 1 illustrates a typical configuration of Oracle10gR2 RAC with NetApp storage and IBM pSeries PowerPC servers running RHEL AS4 U3. This is a scalable configuration and allows users to scale horizontally and internally in terms of processor, memory, and storage.

As shown in the network diagram, we recommend that you dedicate a private network connection between the Oracle10gR2 RAC servers and the NetApp storage. This is accomplished using a dedicated gigabit network

(with a gigabit switch) to the NetApp storage. A dedicated network connection is beneficial for the following reasons:

- In an Oracle10g R2 RAC environment, it is important to eliminate any contentions and latencies
- Providing a separate network ensures security

The cluster interconnect is an essential part for Oracle database clusters. Along with cache fusion, it is also used to monitor the heartbeat of the servers in the existing cluster group. This is a typical configuration that can be deployed in a customer's environment.

4. Requirements

4.1 Hardware Used for Tests

Cluster Nodes

- Two IBM pSeries 520—Model 9111-520 64-bit PowerPC Servers
- One 4-Port 10/100 Base-TX Ethernet PCI Adapter
- One 10/100/1000 Base-T Ethernet PCI Adapter (for private interconnect)
- One 10/100/1000 Base-T Ethernet PCI Adapter (connected to NetApp storage)

Storage Infrastructure

- One Network Appliance FAS2xx/F7xx/F8xx/FASF9x/FAS30xx system with Data ONTAP® 7.2 or later
- One gigabit switch with at least four ports.
- One gigabit NIC in the system
- One or more disk shelves, based on the disk space requirements

4.2 Software Used

For both nodes in the participating cluster unless specified otherwise:

- RHEL AS4 (Update 3) Power Linux
- Oracle10g release 2 (10.2.0.1), with Real Application Clusters license

5. Setup for NetApp Storage

For more information, refer to the Network Appliance installation and setup guides at <http://now.netapp.com>.

1. Please configure a NetApp storage system running Data ONTAP 7.2 and with NFS and SnapRestore® license keys.
2. Create and export volumes for storing Oracle database files on the storage:
 - A. Create three volumes on the storage (Data1) as listed below:

```
orahome Shared Oracle and CRS Home(Binaries)
oradata Oracle datafiles and control files
ora10g CRS files
oralogs database logs, a copy of control file and archive logs.
```

To create volumes, use the following command at the NetApp storage console:

```
Data1> vol create oradata 14
```

Note: We created volume oradata with 14 disks and volumes oralogs and orahome with 8 disks each. You can create your volumes based on your workload needs.

Edit the /etc/exports file on NetApp storage (Data1) and add the following entries to that file:

```
/vol/orahome -anon=0
/vol/oradata -anon=0
/vol/oralogs -anon=0
/vol/ora10g -anon=0
```

B. Execute the following command at the storage system console:

```
Data1> exportfs -a
```

Note: NetApp recommends using flexible volumes in your database environment. NetApp FlexVol® technology pools storage resources automatically and enables you to create multiple flexible volumes on a large pool of disks. This flexibility means you can simplify operations, gain maximum spindle utilization and efficiency, and make changes quickly and seamlessly.

The database volume layout discussed in this document was defined for certification purposes and your setup may vary depending upon requirements. For database layout best practices on NetApp storage, please refer to NetApp Technical Report 3411 www.netapp.com/library/tr/3411.pdf.

6. Patches, Environment, and OS Settings

6.1 Patches

Before your Oracle10gR2 RAC install, the following rpms need to be applied on IBM servers. Some of these rpms may already be applied to your system. Please verify if they already exist before applying them.

To determine whether the required rpms are already installed and committed, enter a command similar to the following:

```
# rpm -qa | grep compat
```

If a Patches is not installed and committed, then install it. Here is a list of required Patches:

- binutils-2.15.92.0.2.13
- compat-libstdc++-33-3.2.3-47.3 (ppc64)
- make-3.80-5 (ppc)
- gcc-3.4.3-22.1(ppc)
- gcc-ppc32-3.4.3-22.1(ppc)
- gcc-c++-3.4.3-22.1(ppc)
- gcc-c++-ppc32-3.4.3-22.1(ppc)
- glibc-2.3.4-2.9 (ppc64)
- glibc-2.3.4-2.9(ppc)
- libaio-0.3.103-3(ppc)
- libaio-0.3.103-3(ppc64)
- libaio-devel-0.3.103-3(ppc)
- libaio-devel-0.3.103-3(ppc64)

- libgcc-3.4.3-22.1(ppc)
- libgcc-3.4.3-22.1(ppc64)
- libstdc++-3.4.3-22.1(ppc)
- libstdc++-devel-3.4.3-22.1(ppc64)
- perl-5.8.5-12.1
- tcl-8.4.7-2
- unzip-5.51-7
- zip-2.3-27
- tar-1.14-4

The IBM XL C/C++ Advanced Edition V7.0.1 for Linux Runtime Environment Component

If the IBM XL C/C++ Advanced Edition V7.0.1 for Linux on POWER compiler is installed, then the IBM XL C/C++ Advanced Edition V7.0.1 for Linux Runtime Environment Component will be automatically installed with the compiler. If the IBM XL C/C++ Advanced Edition V7.0 for Linux on POWER compiler is not installed, then the IBM XL C/C++ Advanced Edition V7.0.1 (or higher) for Linux Runtime Environment Component must be installed and can be downloaded for free without any license requirement from:

<http://www-1.ibm.com/support/docview.wss?uid=swg24007906>

Follow the instructions on the web page to identify the correct files for RHEL 4 "rhel4". Both component files must be downloaded for Oracle:

1. The RTE component: vacpp.rte.70.rhel4.dec2005.update.tar.gz
2. The XL Optimization Libraries: vac.lib.70.rhel4.tar

Java for Oracle JDBC/OCI Drivers

IBM Java 1.4.2 64-bit (SR1a) or higher.

IBMJava2-142-ppc64-SDK-1.4.2-1.0

From: http://www-128.ibm.com/developerworks/java/jdk/linux140/older_download.html

6.2 OS Settings

On RedHat systems, the default ulimits for individual users are set in /etc/security/limits.conf. As a root user, add the following entries using root users:

```
# Oracle specific settings
oracle soft nofile 4096
oracle hard nofile 65536
oracle soft nproc 2047
oracle hard nproc 16384
oracle soft memlock 3145728
oracle hard memlock 3145728
```

This needs to be done on all nodes of the cluster. A server reboot might be necessary to activate updated limits. After you modify the settings, “ulimit -a” command should display the following:

```
# ulimit -a
core file size          (blocks, -c) 0
data seg size          (kbytes, -d) unlimited
file size              (blocks, -f) unlimited
max locked memory      (kbytes, -l) unlimited
max memory size        (kbytes, -m) unlimited
open files              (-n) 1024
pipe size              (512 bytes, -p) 8
stack size             (kbytes, -s) unlimited
cpu time               (seconds, -t) unlimited
max user processes     (-u) 15168
virtual memory         (kbytes, -v) unlimited
```

Verify the same for the Oracle user.

Kernel Settings

Add the following parameters for the shared memory and semaphores to the “/etc/sysctl.conf “ file using root user:

```
kernel.shmall = 2097152
kernel.shmmax = 2147483648
kernel.shmmni = 4096
kernel.sem = 250 32000 100 141
fs.file-max = 65536
net.ipv4.ip_local_port_range = 1024 65000
net.core.rmem_default = 1048576
net.core.wmem_default = 262144
net.core.rmem_max = 1048576
net.core.wmem_max = 262144
```

7. Preinstall Setup Tasks (Cluster Nodes)

This information is intended for both nodes in the participating cluster unless specified otherwise.

1. Please have two IBM pSeries PowerPC servers ready with the latest recommended patches and OS settings as discussed in Section 6.
2. Install/configure NI in the cluster nodes (three per node).
 - a. Public IP: As indicated by name.
 - b. Private interconnects: Connect one gigabit NIC back to back to the other node for cluster interconnects.
 - c. Server connection to NetApp storage: Connect one gigabit NIC to the gigabit switch, which will connect to the gigabit NIC on the NetApp storage.
3. Configure the network interfaces on each node.

- a. Configure the three network interfaces as indicated below.

```
# btc-ppe-srv5 (Host 1)
en0 - ip: 10.73.68.155, netmask 255.255.254.0
en6 - ip: 10.73.69.155, netmask: 255.255.255.0
en4 - ip: 192.168.73.1, netmask: 255.255.255.0
# btc-ppe-srv6 (Host 2)
en0 - ip: 10.73.68.156, netmask: 255.255.255.0
en6 - ip: 10.73.69.156, netmask: 255.255.255.0
en4 - ip: 192.168.73.2, netmask: 255.255.255.0
```

Where:

Interface `en0` is the public ip for each node.

Interface `en6` on both cluster nodes is connected to the gigabit switch for storage I/O.

Interface `en4` on both cluster nodes is connected back to back for cluster private interconnects.

- b. Update the `/etc/hosts` file on the cluster nodes and add entries for public, private, and VIP addresses. Please note that in addition to the preconfigured public and private network, Oracle Database 10g requires additional IP addresses that will be mapped to the public address as virtual IPs (VIPs). If a node fails when an application or user makes a connection using a VIP, the Oracle clusterware will transfer the VIP address to another surviving instance. You should add the VIP to the `/etc/hosts` file on all nodes in the cluster as well as all nodes accessing the database.

A sample for `/etc/hosts` entries

```
# Internet Address Hostname # Comments
10.73.68.155 btc-ppe-srv5 btc-ppe-srv5.btcppe.netapp.com
10.73.69.155 btc-ppe-srv5-en6
192.168.73.1 btc-ppe-srv5-i btc-ppe-srv5-i.btcppe.netapp.com
10.73.68.195 btc-ppe-srv5-v btc-ppe-srv5-v.btcppe.netapp.com
10.73.68.156 btc-ppe-srv6 btc-ppe-srv6.btcppe.netapp.com
10.73.69.156 btc-ppe-srv6-en6
192.168.73.2 btc-ppe-srv6-i btc-ppe-srv6-i.btcppe.netapp.com
10.73.68.196 btc-ppe-srv6-v btc-ppe-srv6-v.btcppe.netapp.com
10.73.69.105 data1
```

- c. Ensure the connectivity of each interface via the ping command. (interconnects, public IPs, and storage)
- d. Create NFS mount points and mount the volumes with the following mount options on all the cluster nodes. As a root user, update the `/etc/fstab` file on all server nodes and add the following entries:

```
Datal:/vol/oradata /oradata nfs
rw,bg,hard,rsize=32768,wsize=32768,vers=3,proto=tcp,actimeo=0,
nointr,suid,timeo=600
```

```
Datal:/vol/orahome /orahome nfs
rw,bg,hard,rsize=32768,wsize=32768,vers=3,proto=tcp,actimeo=0,
nointr,suid,timeo=600
```

```
Datal:/vol/oralogs /oralogs nfs
rw,bg,hard,rsize=32768,wsize=32768,vers=3,proto=tcp,actimeo=0,
nointr,suid,timeo=600
```

```
Data1:/vol/ora10g /ora10g nfs
hard,proto=tcp,vers=3,suid,nointr,rw,bg,rsize=32768,wsiz=32768,noac,
timeo=600
```

Where:

Data1 is the name of the NetApp storage system.

oradata, oralogs, orahome, and ora10g are the mount points on the cluster nodes.

ora10g is just a separate mount point for CRS files. CRS files (cluster registry file and voting disk file) can reside in the same /ora10g volume but must be mounted with “noac” mount option. “intr” mount option is required in failure scenarios in which CRS has to evict a node.

During the Oracle CRS install, please make sure to indicate a path starting with /ora10g directory when prompted for the Cluster registry file(ocr) and voting disk(css) file location.

Note: Oracle install will fail if you dynamically mount the NFS volumes without adding entries in /etc/fstab file.

- e. Create the following mount points on all cluster nodes:

```
#mkdir /oradata
#mkdir /oralogs
#mkdir /orahome
#mkdir /ora10g
```

Mount exported volumes on the mount points created above on all the cluster nodes. It is always a good idea to verify mount options by mount command on each node. After the NFS volumes are mounted, change the ownership of these mounted volumes to Oracle user.

8. Installation Procedure

For an Oracle10gR2 RAC install RHEL Power Linux, please refer to Oracle install documentation (part # B14203-05). The link below provides instructions for Oracle10gR2 RAC install on RHEL Power Linux:

8.1 Preparing to Install the Oracle RAC 10g on Cluster Nodes

1. This document assumes the Oracle user account and the group to be oracle and dba, respectively, on both cluster nodes. The user ID and group name for the oracle account should be the same on both cluster nodes. A sample oracle user .bash_profile file is provided in the appendix. Make sure the user profile file exports at least ORACLE_BASE, ORACLE_PRODUCT, ORACLE_HOME, ORACLE_SID, and PATH entries.
2. Grant appropriate permissions to the oracle user on all shared mounted volumes, /oradata, /orahome, and /oralogs as indicated below:

```
#chown -R oracle:dba /oradata
#chmod -R 755 /oradata
```

Repeat the same for orahome and oralogs volumes.

3. Set up account equivalence between the cluster nodes for the oracle user account. Add the following entries to the /etc/hosts.equiv file on all cluster nodes:

```
btc-ppe-srv5 oracle
btc-ppe-srv6 oracle
btc-ppe-srv5-i oracle
btc-ppe-srv6-i oracle
```

4. Test the oracle account equivalence using a remote shell utility such as rsh after logging in as the oracle user from both cluster nodes.

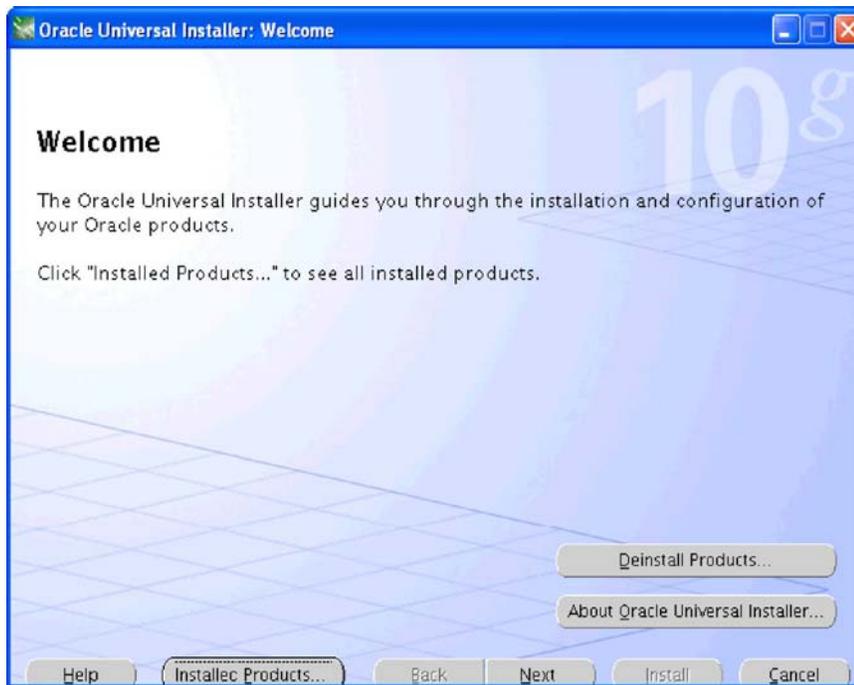
```
btc-ppe-srv5:
#su - oracle
$rsh btc-ppe-srv5 pwd
$rsh btc-ppe-srv6 pwd
$rsh btc-ppe-srv5-i pwd
$rsh btc-ppe-srv6-i pwd

btc-ppe-srv6:
#su - oracle
$rsh btc-ppe-srv5 pwd
$rsh btc-ppe-srv6 pwd
$rsh btc-ppe-srv5-i pwd
$rsh btc-ppe-srv6-i pwd
```

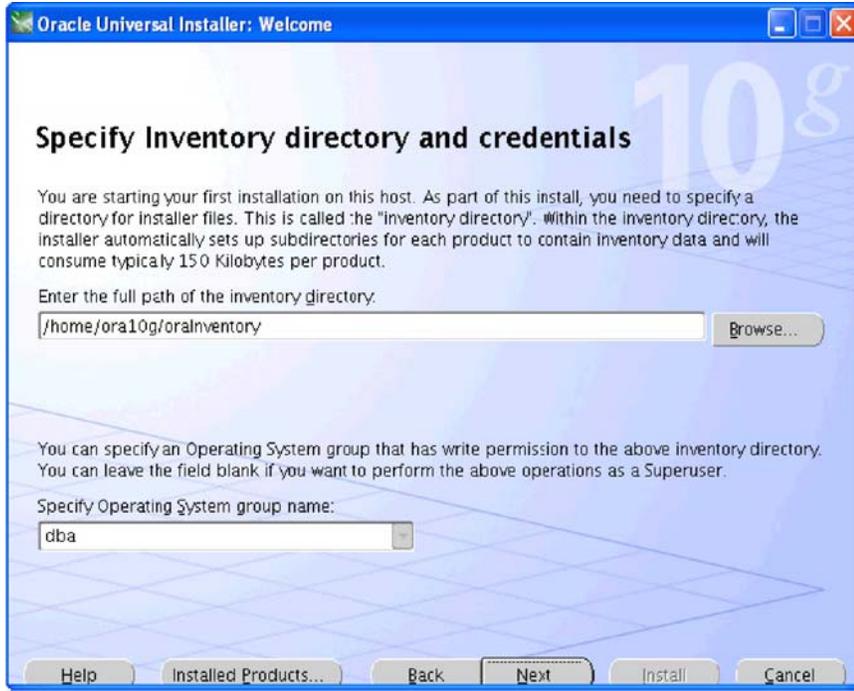
8.2 Installing the Oracle RAC 10g Cluster Ready Services (CRS)

Prior to installing the Oracle10g database, Oracle Cluster Ready Services (CRS) must be installed, configured, and started. Refer to Oracle Real Application Clusters Installation and Configuration Guide 10g release 2 (10.2.0.1) for UNIX[®] Systems at <http://otn.oracle.com/docs/content.html> for more information on installing Oracle Cluster Ready Services on Linux. This section briefly describes the procedures for using the Oracle Universal Installer (OUI) to install CRS. Note that the CRS home that you identify in this phase of the installation is only for CRS software; this home cannot be the same home as the Oracle10g RAC database home. In short, ORACLE_HOME and CRS HOME must be different locations.

1. Run the runInstaller command from the /crs subdirectory on the Oracle Cluster Ready Services release 21 (10.2.0.1) CD-ROM. This is a separate CD that contains the Cluster Ready Services software. This document assumes that the OUI is started from node 1 (btc-ppe-srv5). When the OUI displays the Welcome page, click Next.

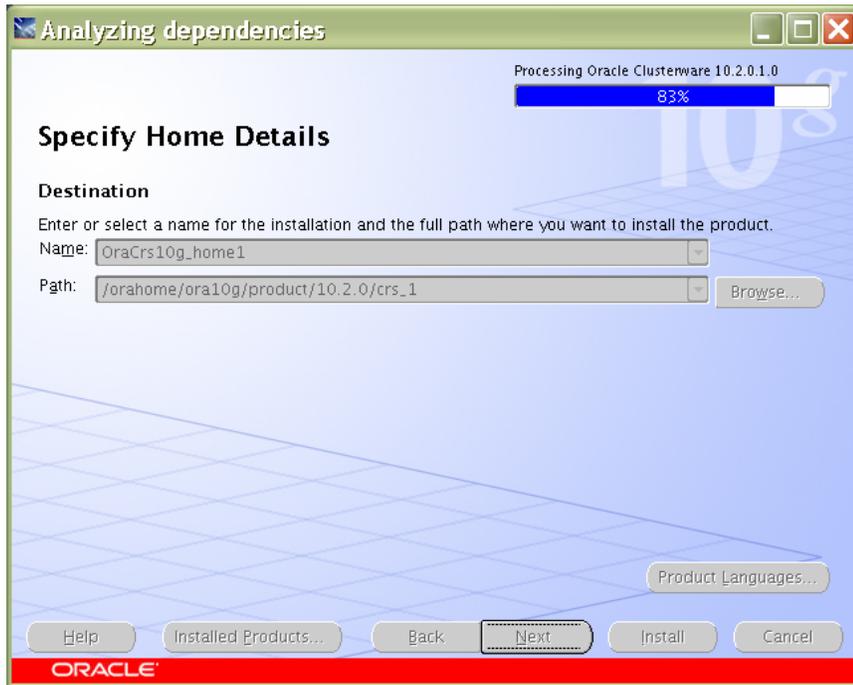


2. On the “Specify Inventory...” page, enter a nonshared location for Oracle Inventory. This is the only part of Oracle10g that should not be shared. For this test, we used /home/oracle/orainventory for the Oracle Inventory information. Click Next.

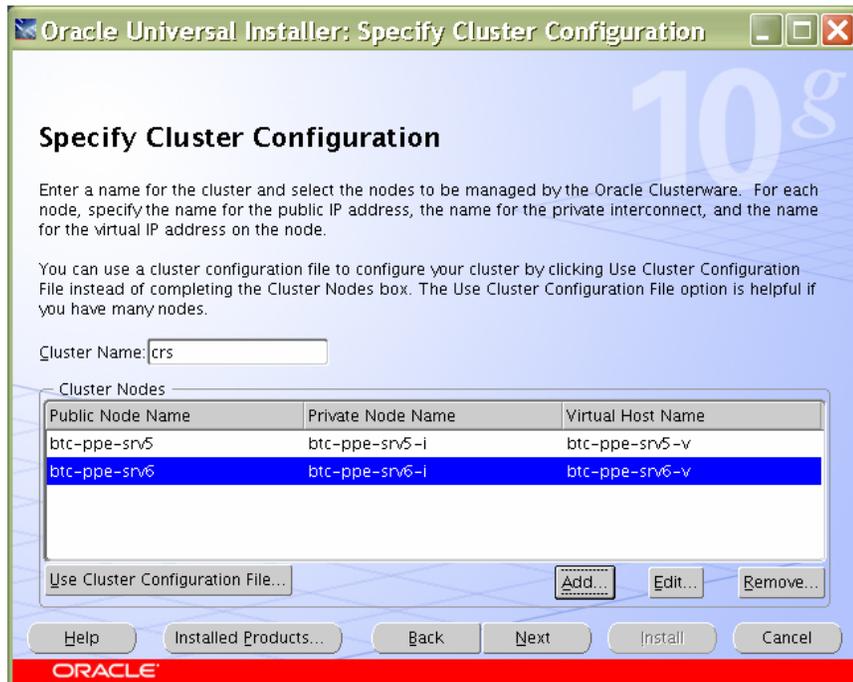


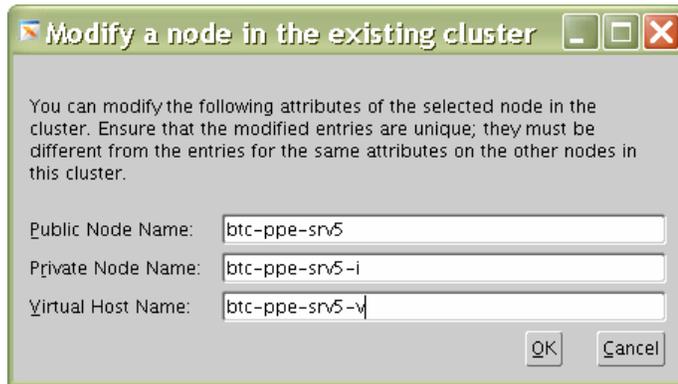
3. The Specify File Locations page contains predetermined information for the source of the installation files and the target destination information. Specify the destination path for the shared CRS home. The path should be on a shared file system and different from \$ORACLE_HOME. In this exercise, the shared CRS home was /orahome/ora10g/product/10.2.0/crs_1.



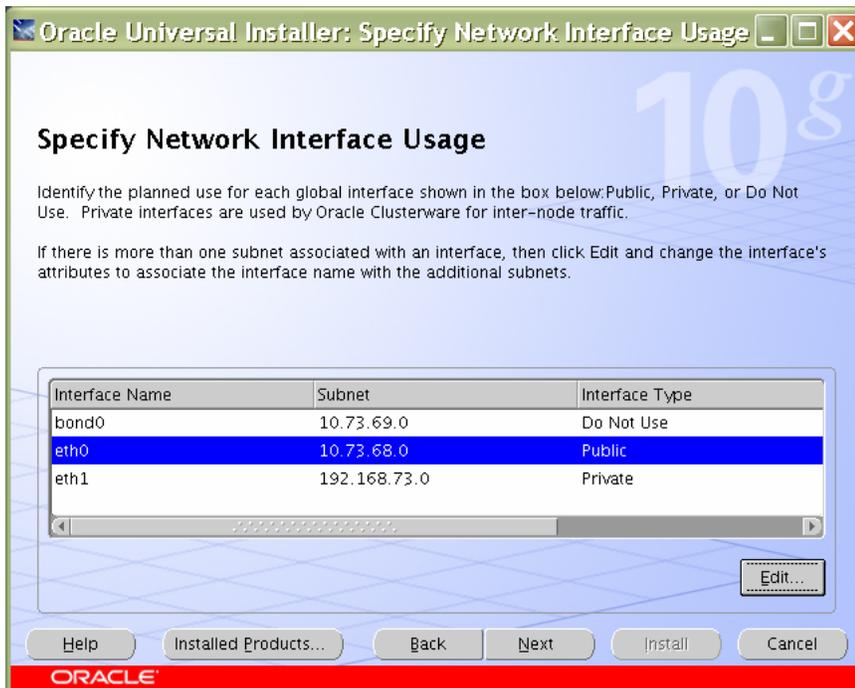


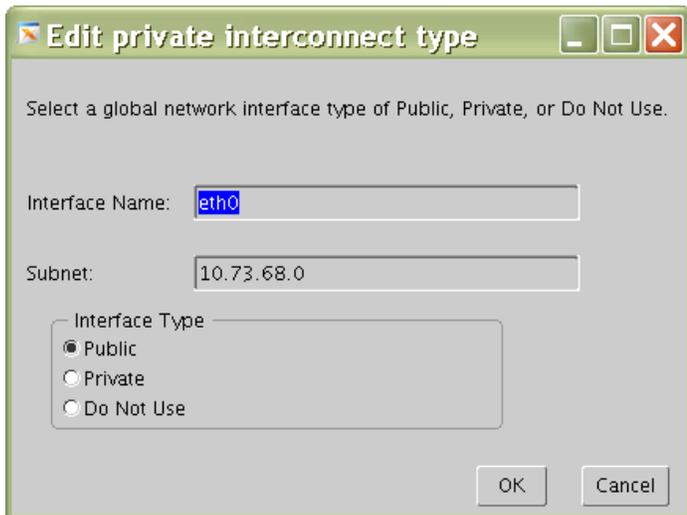
4. On the next screen, specify the cluster name, public names (hostnames), private names and virtual hostnames to be used for the cluster interconnect. In our case, the public names are btc-ppe-srv5 and btc-ppe-srv6, the private names are btc-ppe-srv5-i and btc-ppe-srv6-i and the Virtual hostnames are btc-ppe-srv5-v and btc-ppe-srv6-v. Click next after adding all the nodes name.



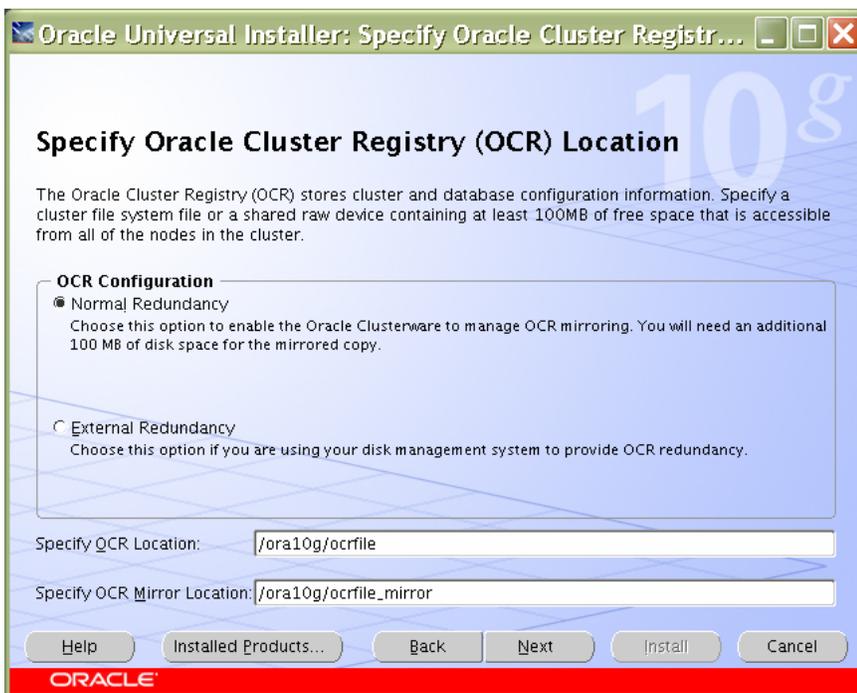


- On the Network Interface Usage page, specify the private network to be used for the cluster interconnect. This is a very important step. Do not leave it set to the default, which is Do Not Use. In this case, eth1(btc-ppe-srv5-i) was used as the private interconnect and eth0(btc-ppe-srv5) was used as public interface. Select the interface and click the edit button to modify it. Click Next.

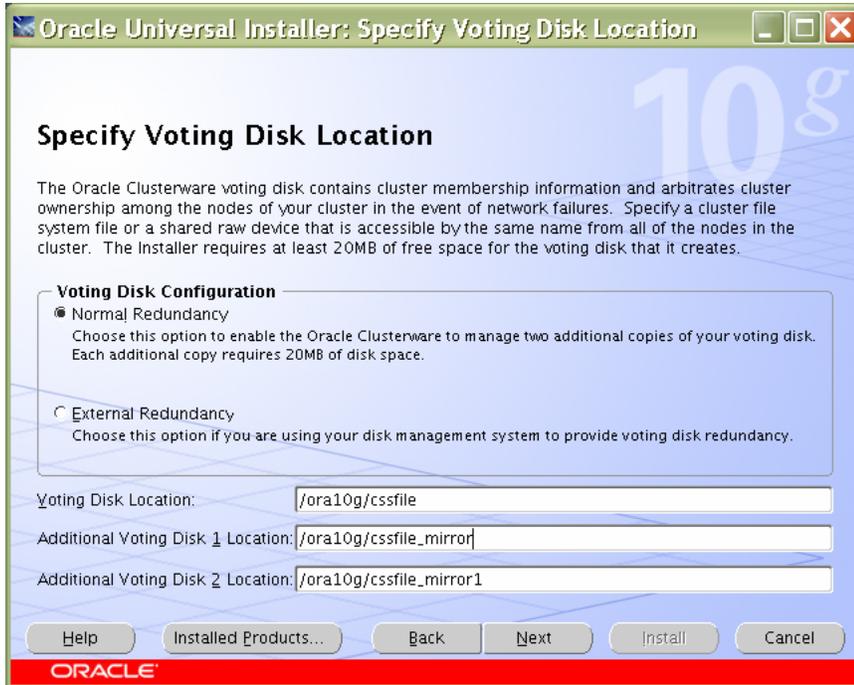




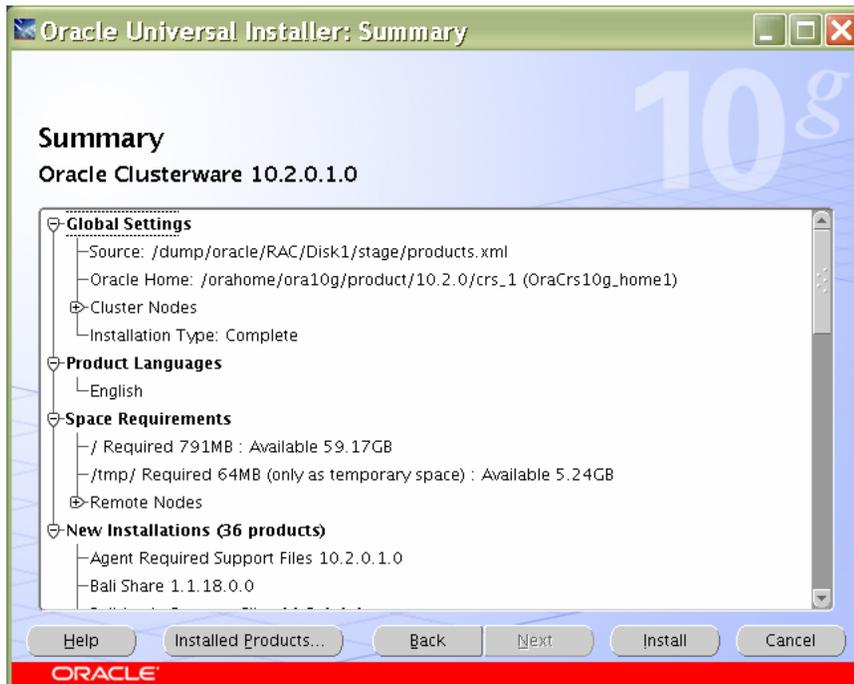
6. On the Oracle Cluster Registry page, specify the OCR (Oracle Cluster Registry) file. Make sure to specify the full path to a shared location along with the name of the file. Do same for mirror file if you want normal redundancy. In our case, we used /ora10g/ocrfile and /ora10g/ocrfile_mirror. Click Next.

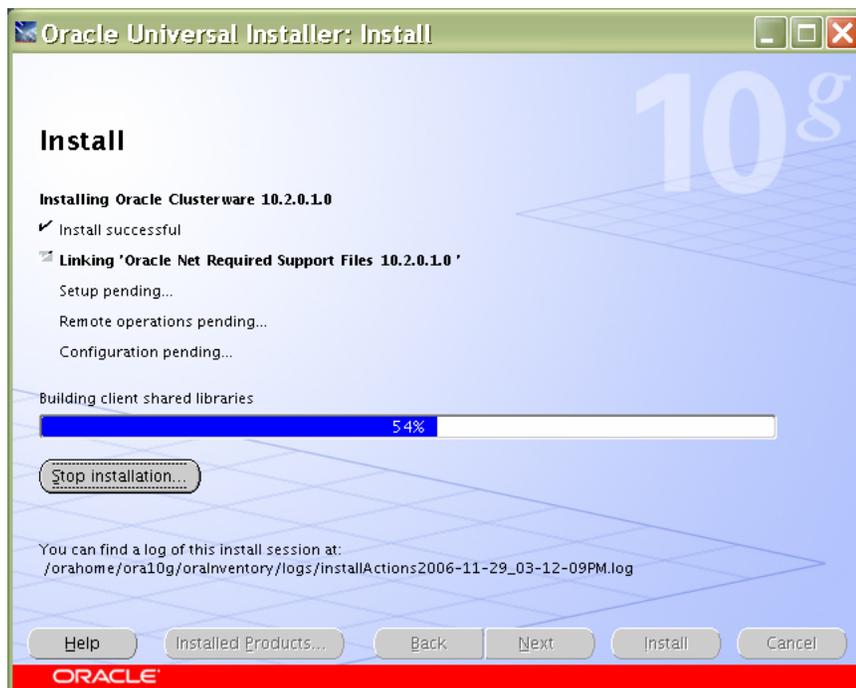
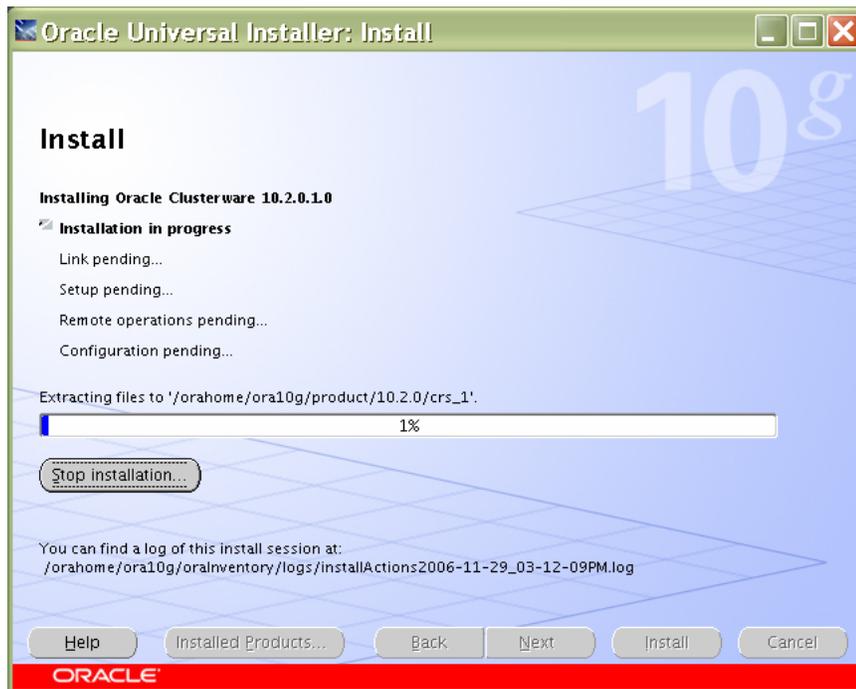


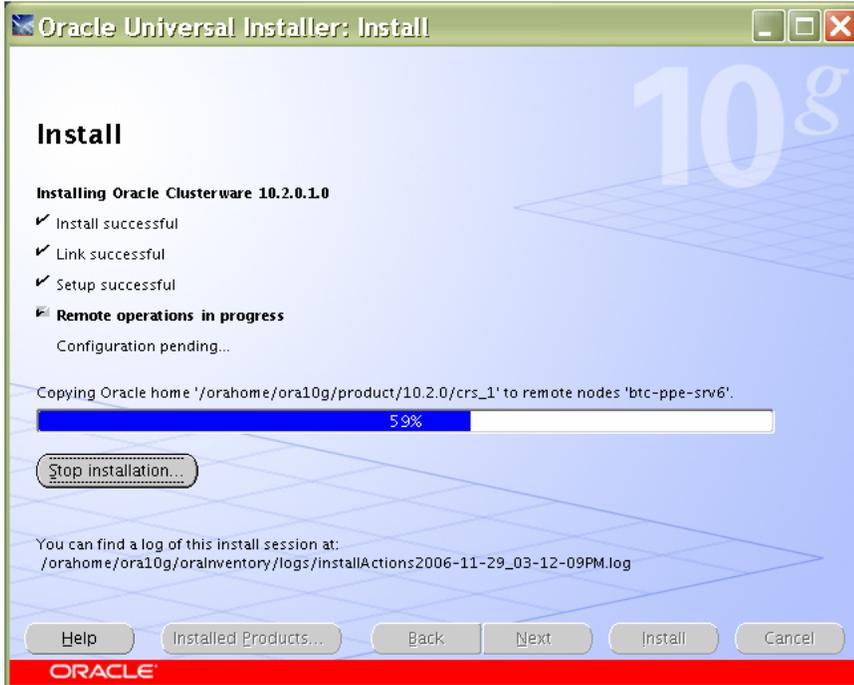
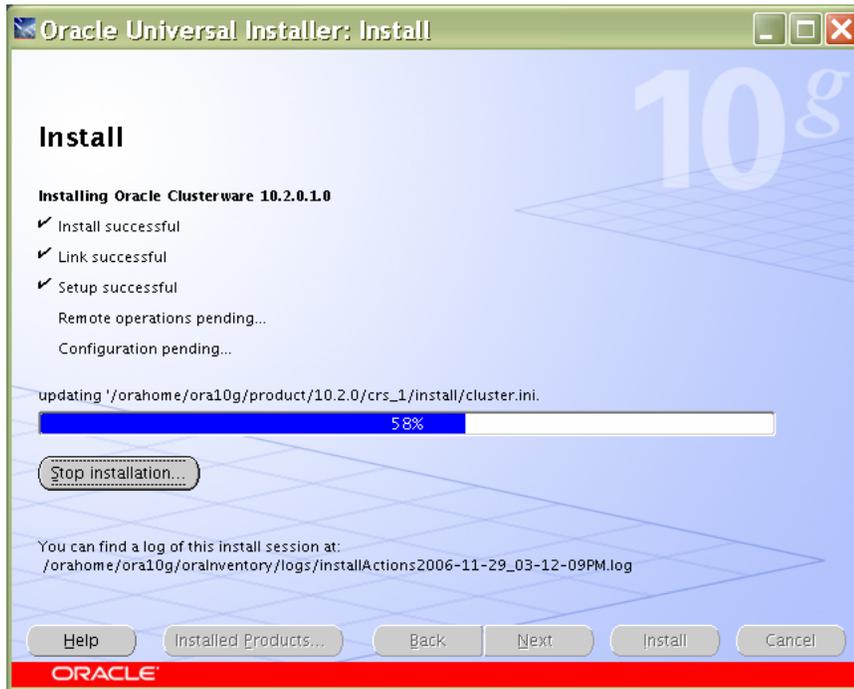
7. On the Voting Disk page, specify the CSS (Cluster Synchronization Services) voting disk file location. We used /ora10g/cssfile for CSS services. In case of normal redundancy specify the path along with name. Click Next.

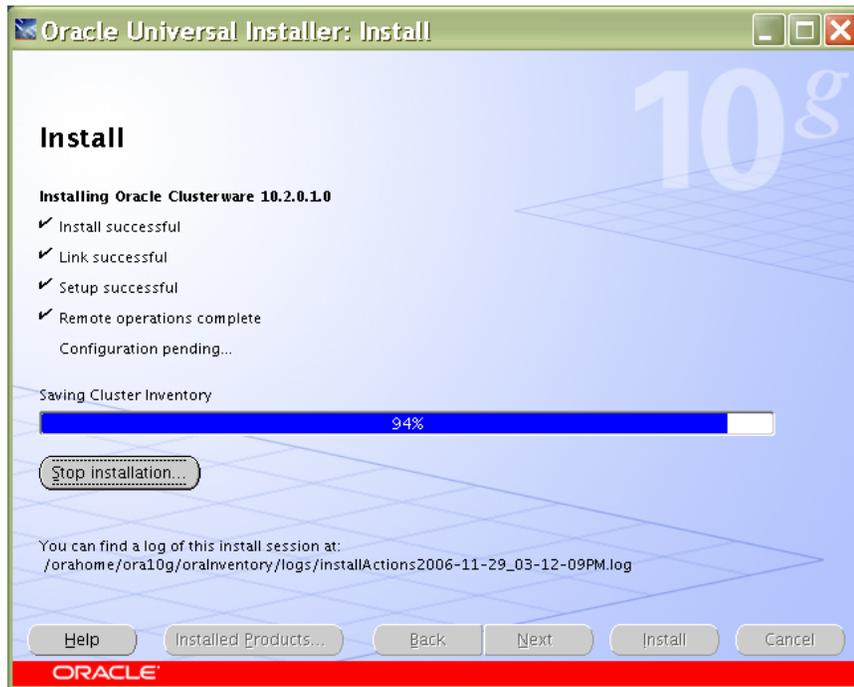


8. Click Install to continue CRS Installation

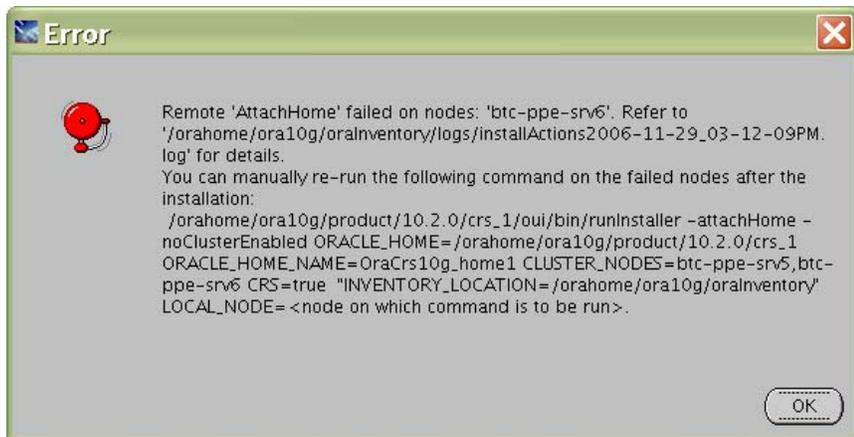




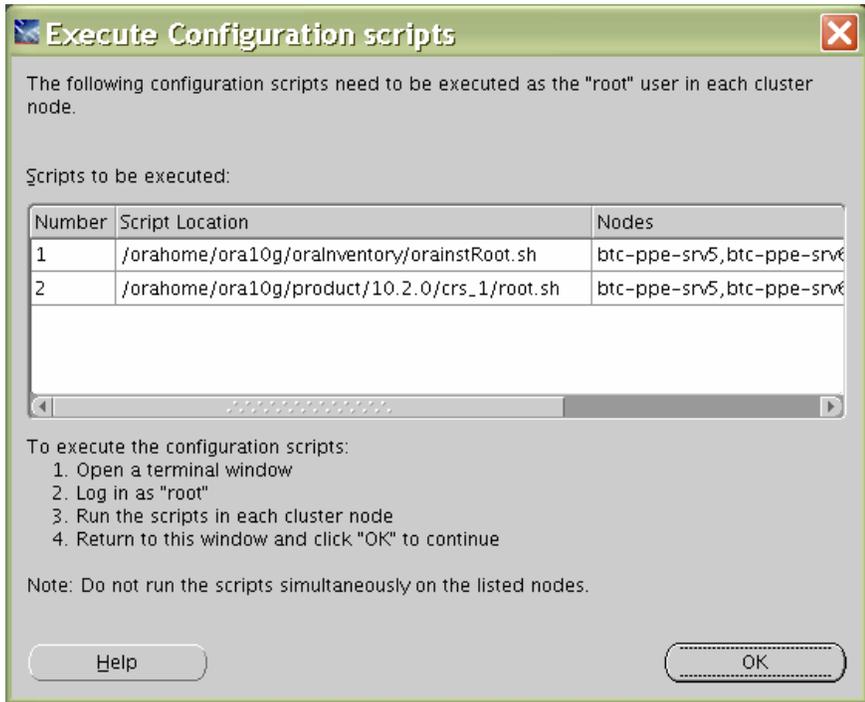




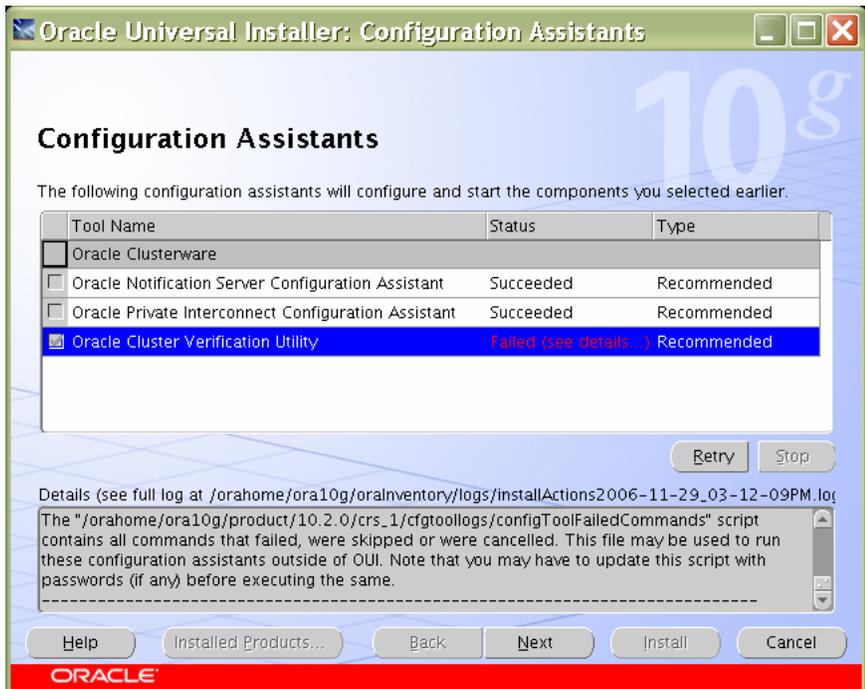
9. During Configuration you may get following error. Click OK to continue.

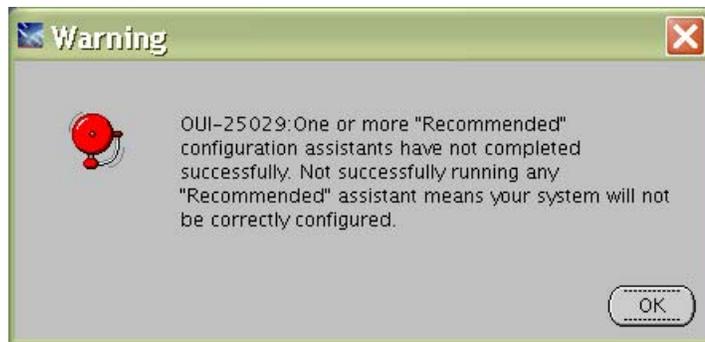
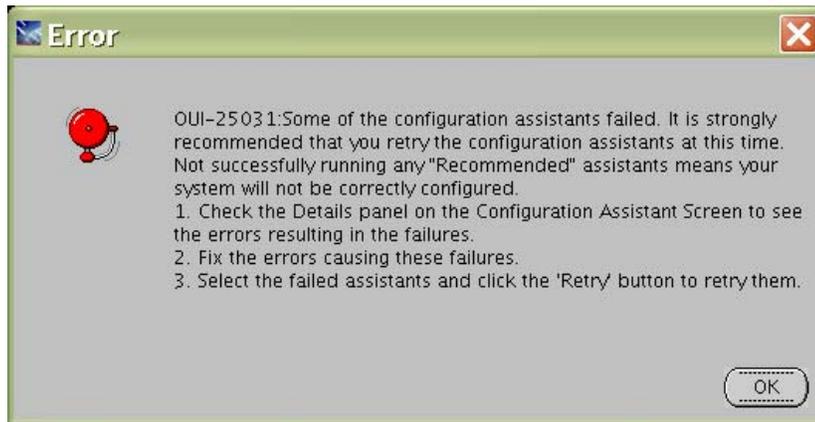


10. Run the following script as root user starting from primary node.



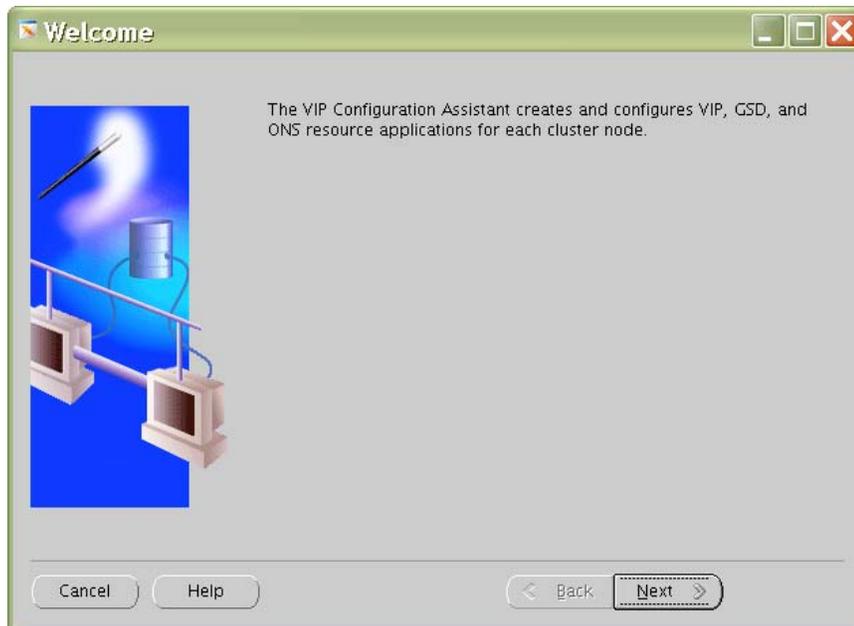
11. In the configuration assistant window you may get some warnings. Click OK to continue.



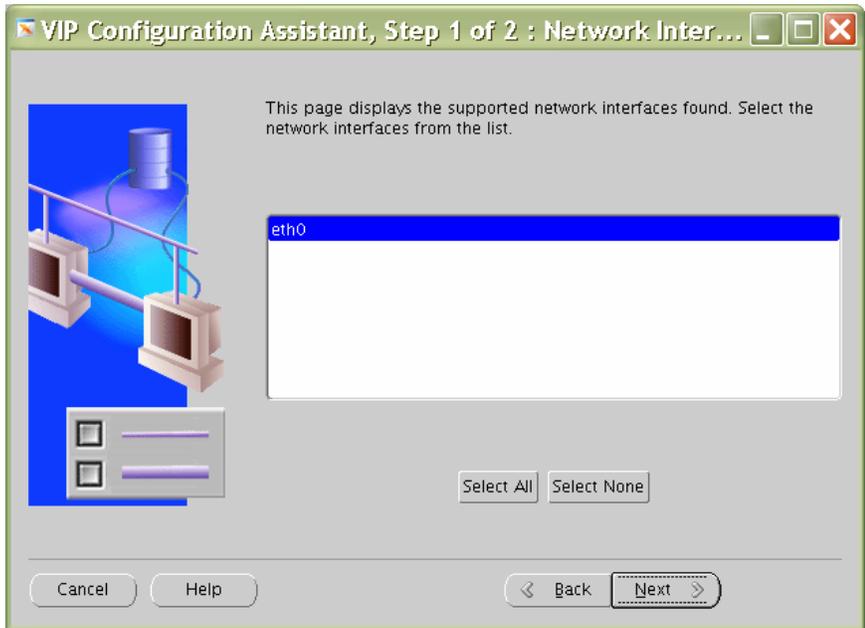


12. Run vipca utility from \$ORA_CRS_HOME/bin directory as root user on Master Node (btc-ppe-srv5). Click next.

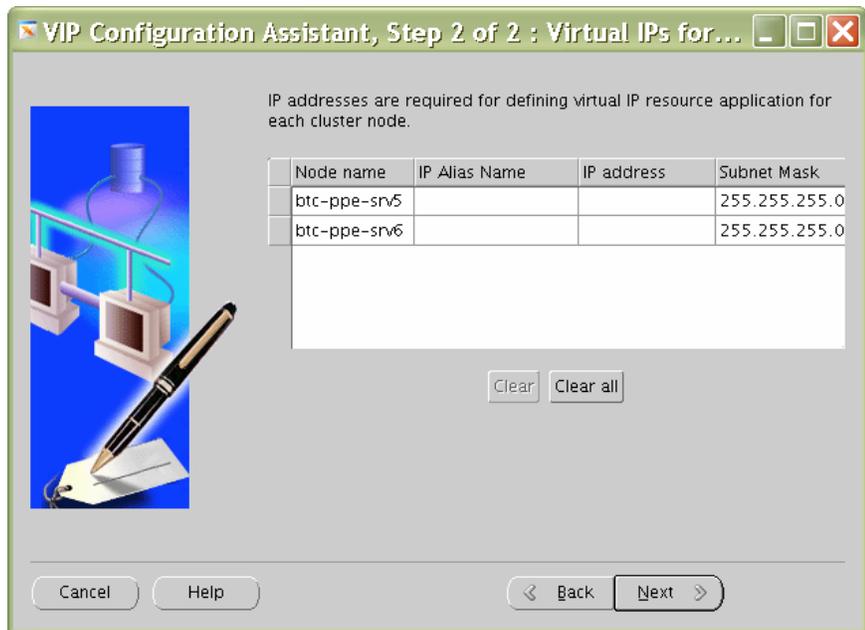
```
# ./vipca
```

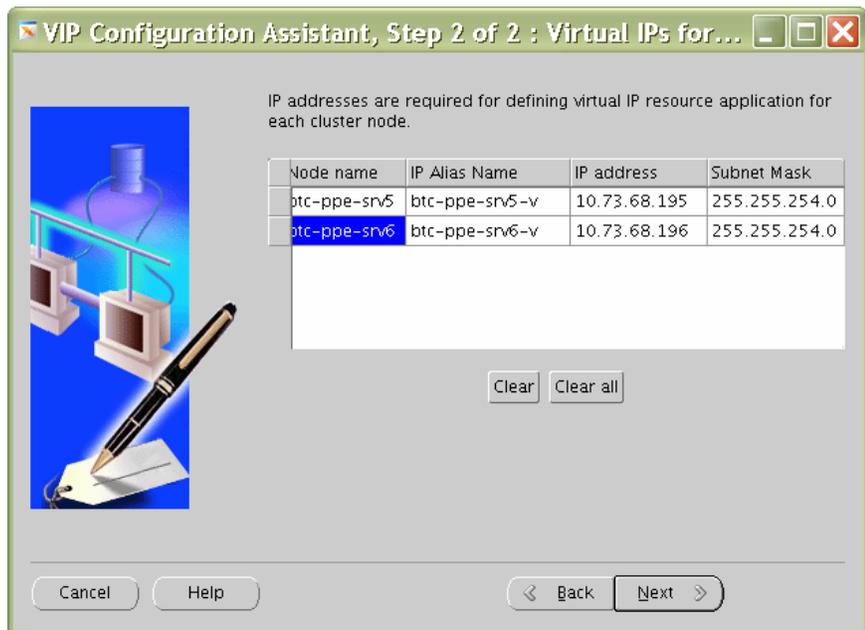


- Select the Public Interface. Click Next.

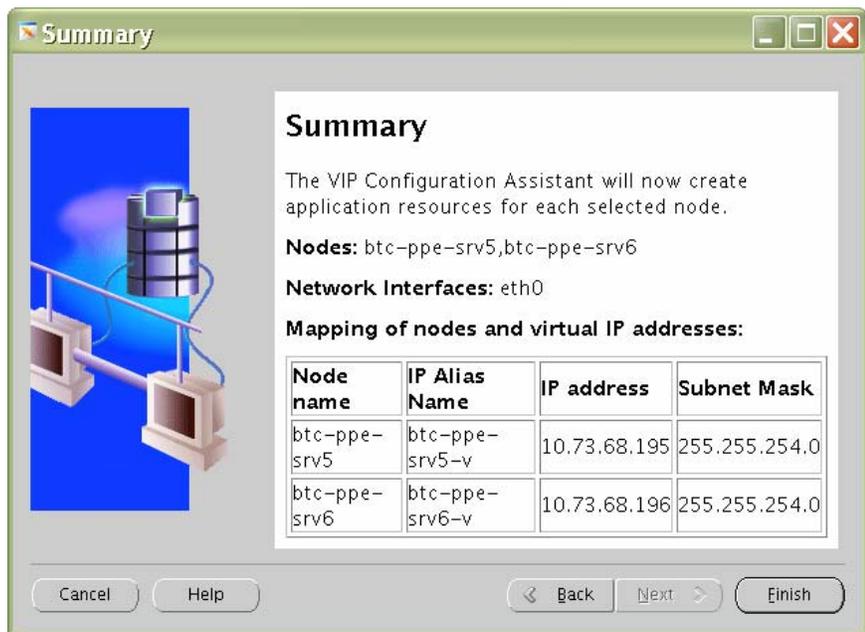


- Specify Virtual IP address and Subnet mask of each node. Click Next.





15. Click Finish to continue vipca.



Configuration Assistant Progress Dialog



- ✓ **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

VIP Configuration Assistant Progress

0%

Stop

Configuration Assistant Progress Dialog



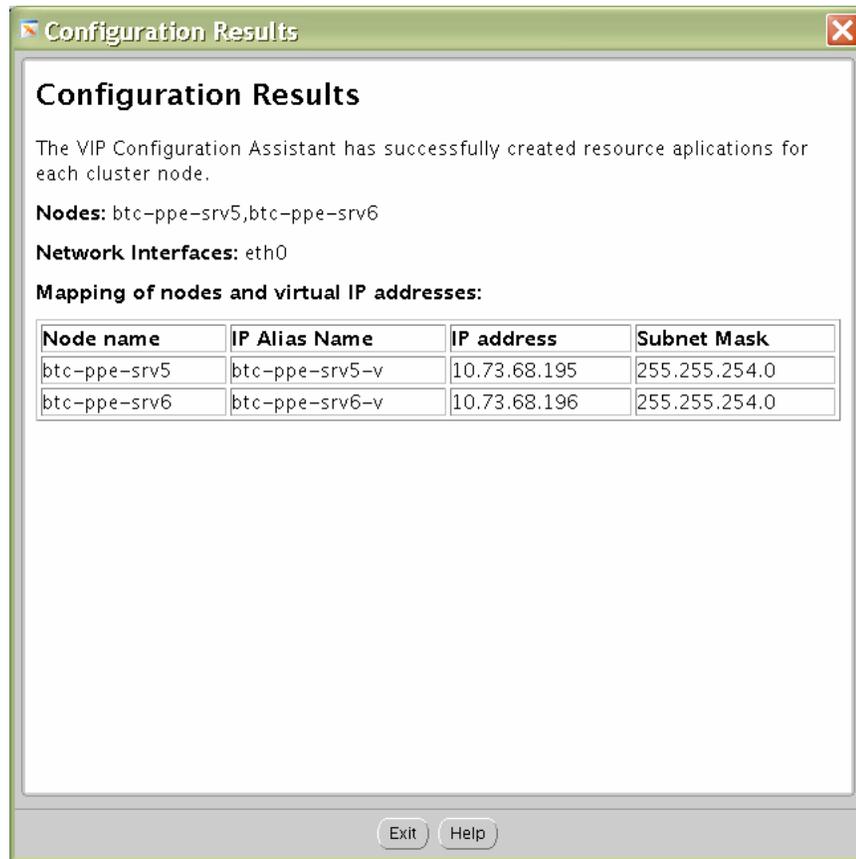
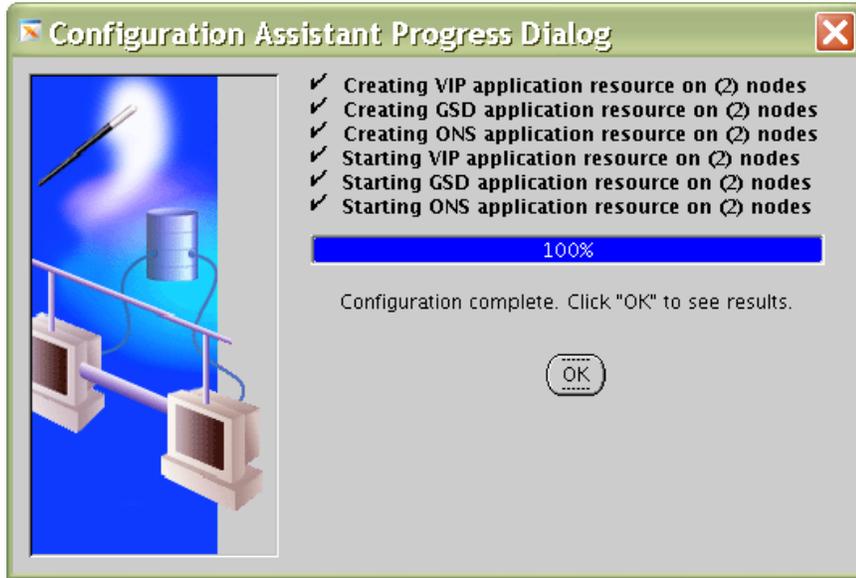
- ✓ **Creating VIP application resource on (2) nodes**
- ✓ **Creating GSD application resource on (2) nodes**
- ✓ **Creating ONS application resource on (2) nodes**
 - Starting VIP application resource on (2) nodes
 - Starting GSD application resource on (2) nodes
 - Starting ONS application resource on (2) nodes

VIP Configuration Assistant Progress

35%

Stop

16. Click OK and then exit to finish VIPCA.



17. You can verify your CRS installation by executing the `olsnodes` command from the CRS Home/bin directory. The `olsnodes` command syntax is:

```
olsnodes [-n] [-l] [-v] [-g]
```

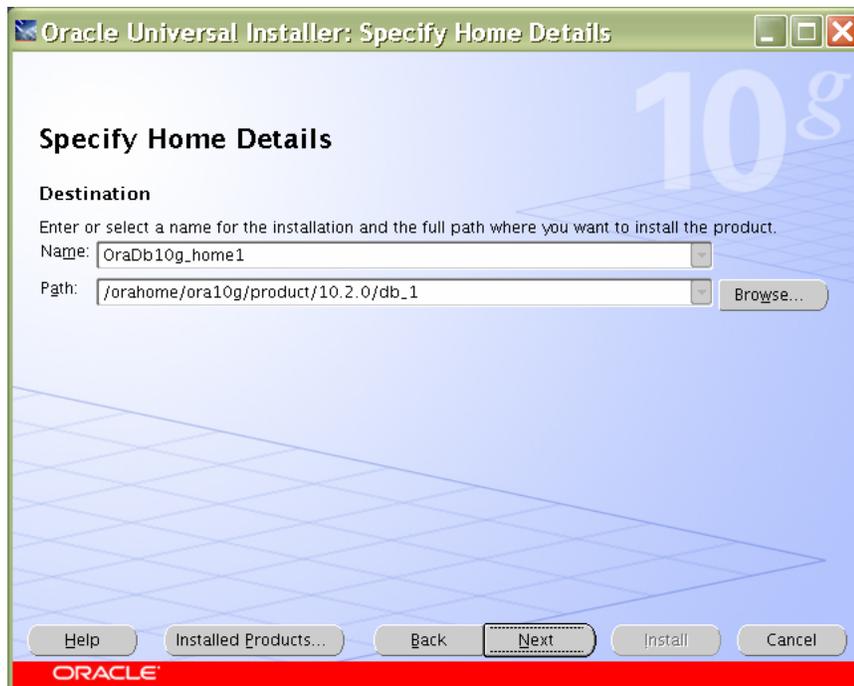
Where:

- n displays the member number with the member name
- l displays the local node name
- v activates verbose mode
- g activates logging

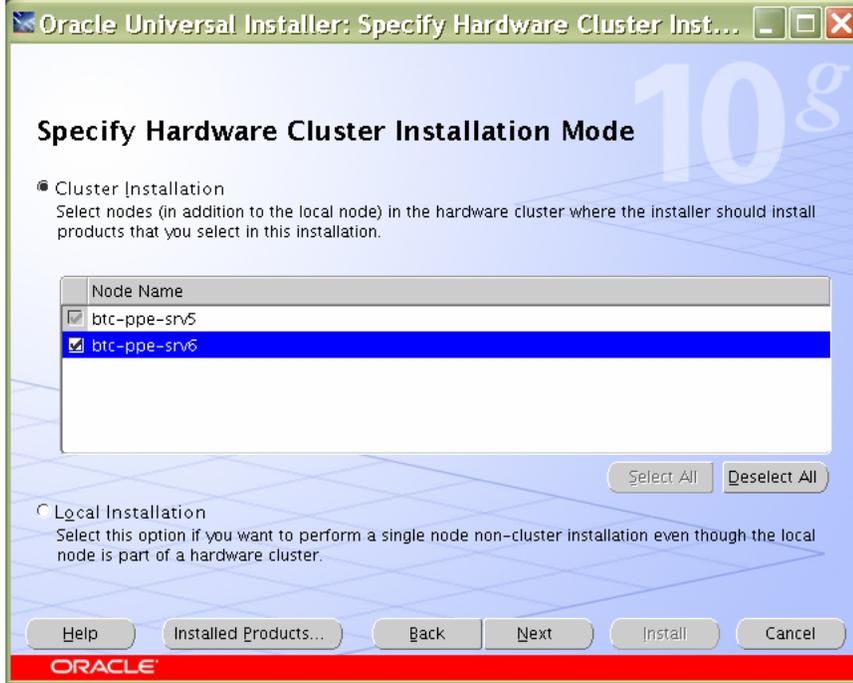
The output from this command should be a listing of the nodes on which CRS was installed.

8.3 Installing Oracle RAC 10g software

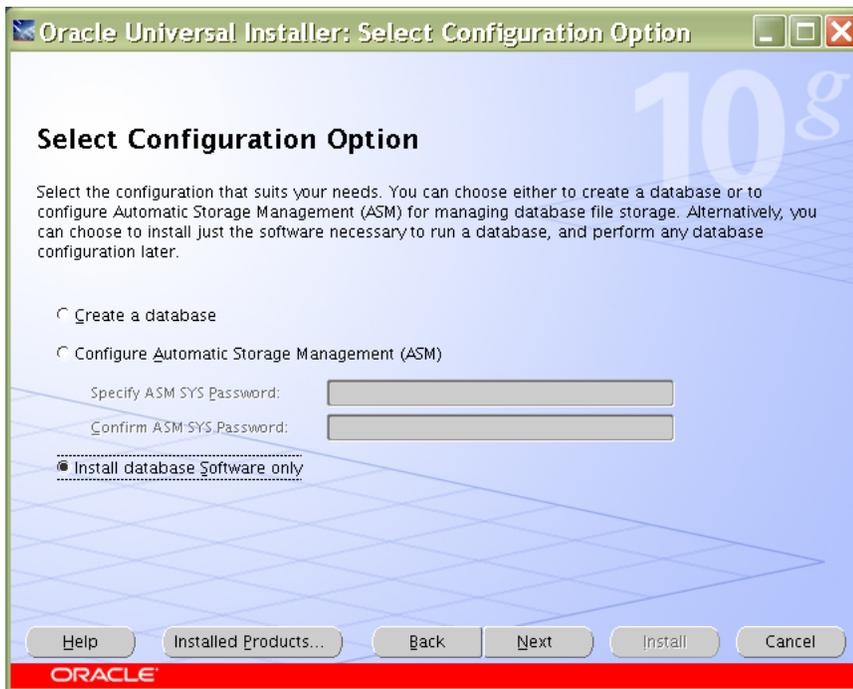
1. After making sure that Oracle Cluster Ready Services have started on the cluster nodes, start `runInstaller` from Disk1 of the Oracle10g release 2 CDs.
2. On the Specify File Locations screen, enter the destination path for the shared `ORACLE_HOME`. This should be a different location than the shared CRS Home. For this exercise, the shared `ORACLE_HOME` was `/orahome/ora10g/product/10.2.0/db_1`.

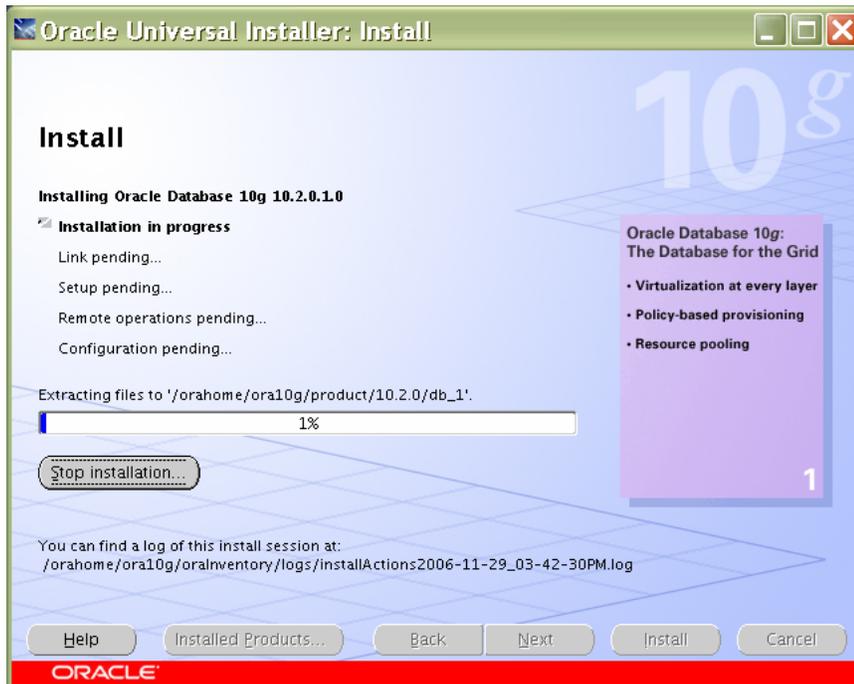
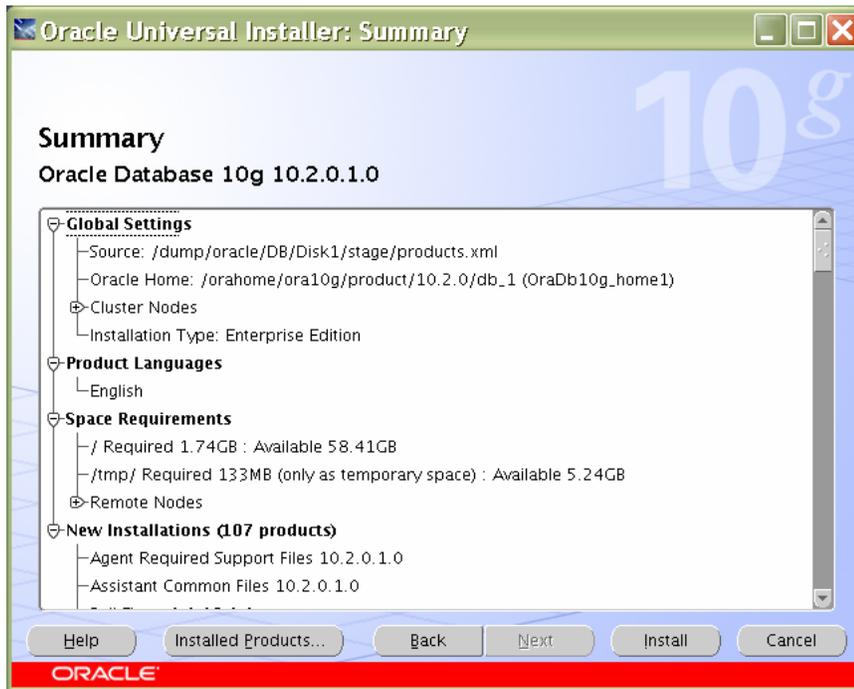


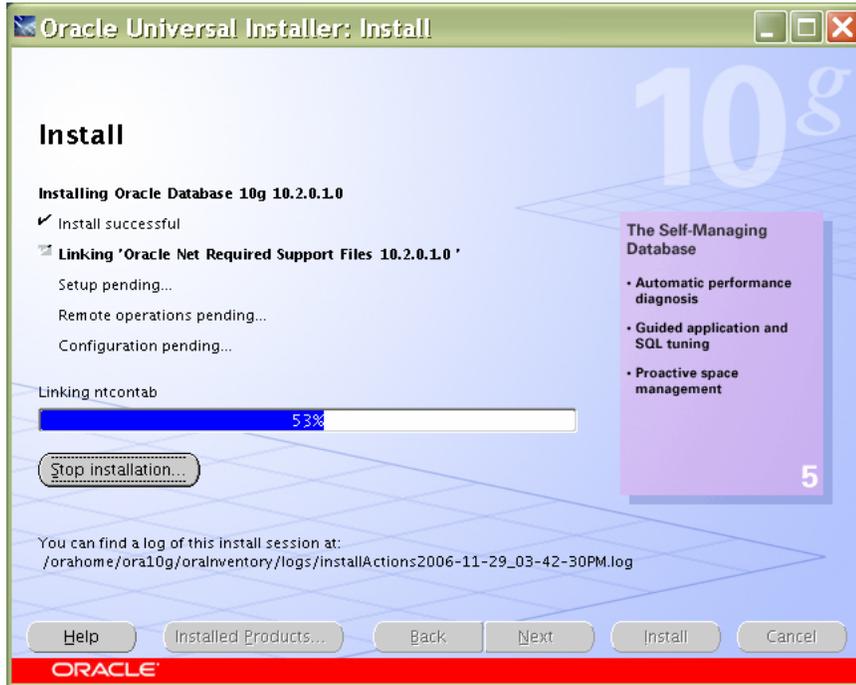
3. On the next screen, select Cluster Installation and choose all the nodes in the cluster. For our exercise, the two cluster nodes were `btc-ppe-srv5` and `btc-ppe-srv6`. Click Next. (Note: If the nodes are not displayed in the cluster node selection, then Oracle Cluster Ready Services are not configured or started on those cluster nodes.)



4. For installation type, select Enterprise Edition and click Next.
5. On the Select Database Configuration page, choose the “Do not create a starter database” option. We used dbca to create a database later. Click Next to continue further.

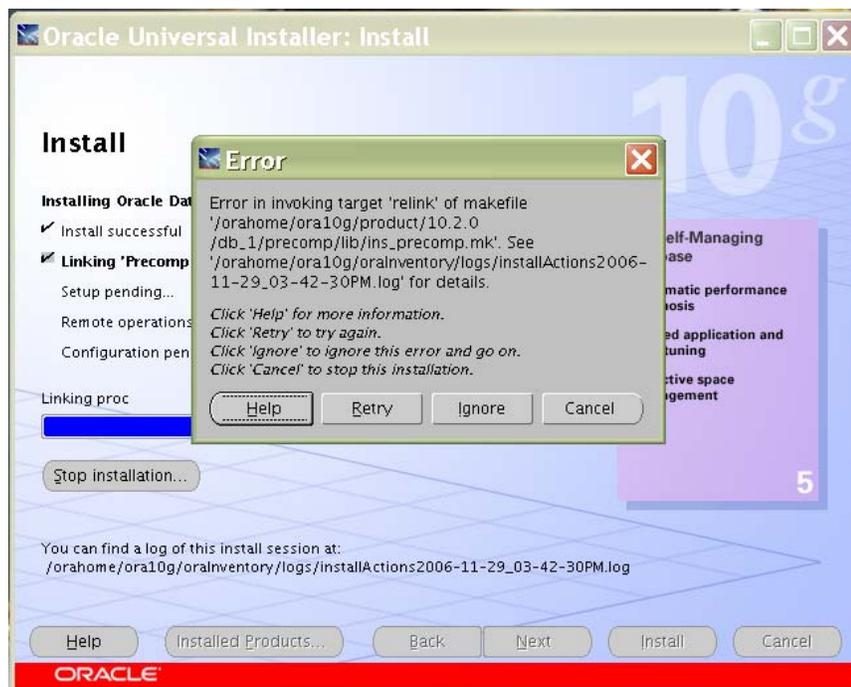


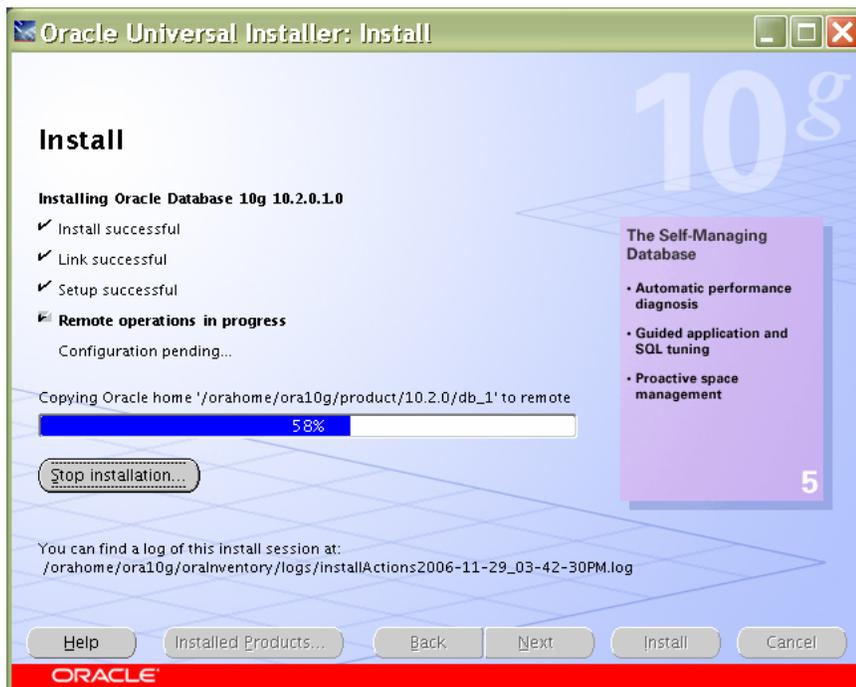
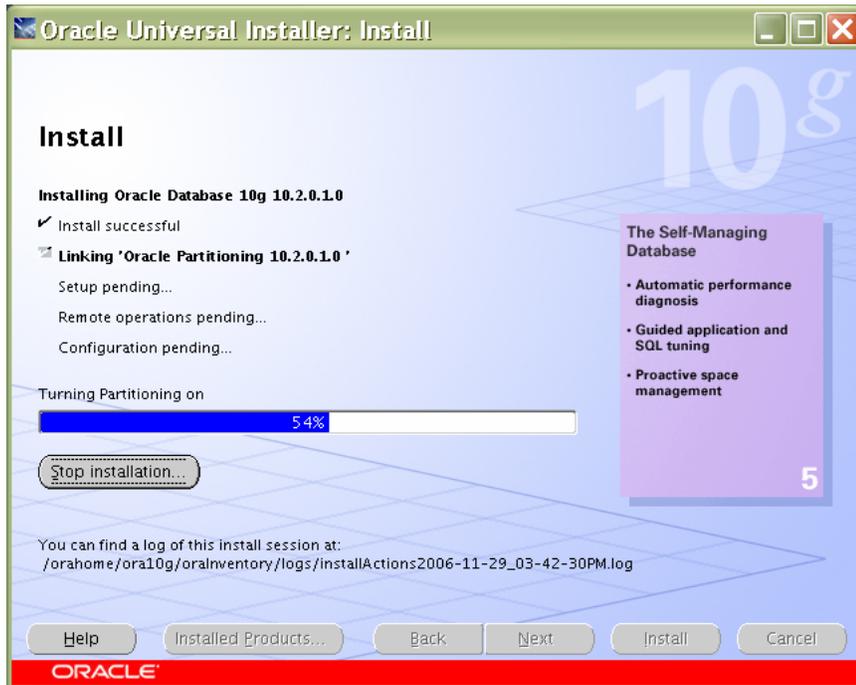




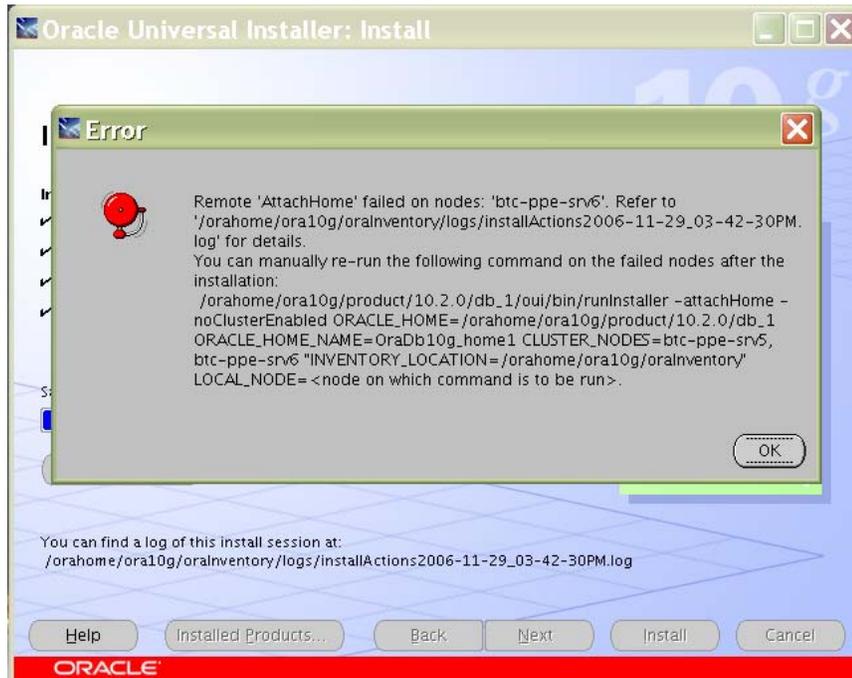
6. Please apply patch **4767801** during the linking phase of database Installation. Refer Metalink Notes: Note:341507.1. After patching click on retry.

http://download-west.oracle.com/docs/cd/B19306_01/install.102/b14203/prelinux.htm#sthref133

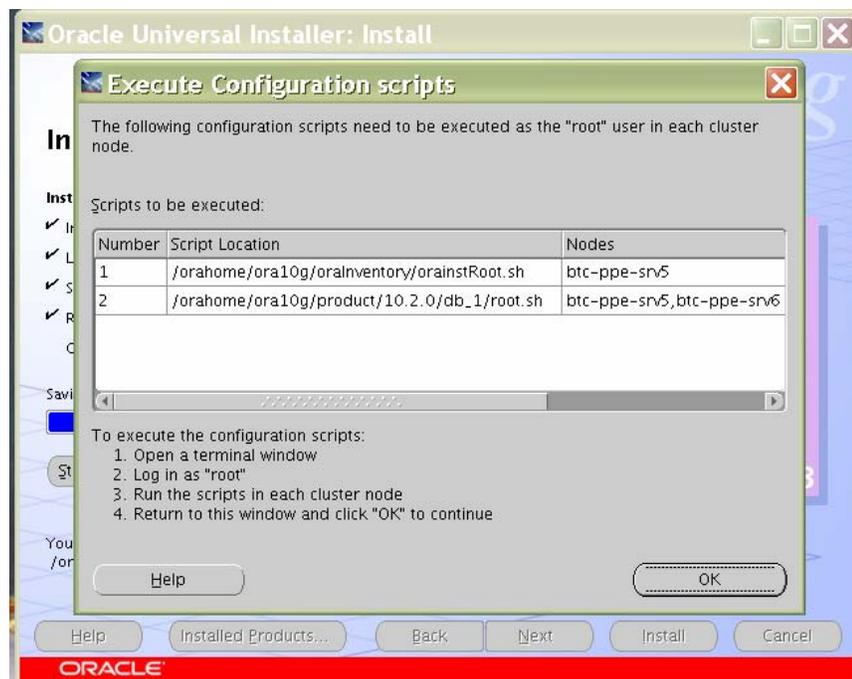




7. Click OK to continue.



- Run the following scripts as root user starting from master node when prompted.



- Click exit to finish the Database Installation.



9. Appendix

Sample `.bash_profile` file for the Oracle user:

```
export ORACLE_BASE=/orahome/ora10g;
export ORACLE_PRODUCT=$ORACLE_BASE/product;
export ORACLE_HOME=$ORACLE_PRODUCT/10.2.0/db_1;
export ORACLE_CRS=$ORACLE_PRODUCT/10.2.0/crs_1;
export ORACLE_SID=oce;
export
LD_LIBRARY_PATH=$ORACLE_HOME/lib:$ORACLE_CRS/lib:$ORACLE_HOME/lib32:$LD_LIBRARY
_PATH;
export LIBPATH=$ORACLE_HOME/lib:$ORACLE_CRS/lib:$ORACLE_HOME/lib32:$LIBPATH
export
PATH=$PATH:$ORACLE_HOME/bin:$ORACLE_HOME:$ORACLE_CRS/bin:/usr/java14/bin;
```

10. Acknowledgements

The author would like to thank the following individuals for their contribution to the certification process and technical report:

Daniel Morgan, Shanthi Adloori, Uday Shet, and Vasu Subbiah
Network Appliance Inc.

11. Disclaimer

Each environment has its own specific set of requirements and no guarantees can be given that the results presented in this report will work as expected on other platforms. This paper should assist in the research and troubleshooting that may be required in a particular case and serve as a checklist of items to be aware of. Please forward any errors, omissions, differences, new discoveries, and comments about this paper to niranjan.mohapatra@netapp.com.

