

Technical report:

Oracle RAC 10*g* Release 2 on SUSE Enterprise Linux 10 and IBM System Storage N series

Installation guide

Document NS3594-0

April 11, 2008



Table of contents

Abstract	3
Introduction	3
Assumptions	3
System environment	3
Requirements	
Hardware	
Software	5
Setup for IBM N series storage	6
Operating-system configuration	7
Patches	
Operating-system settings	8
Kernel settings	8
Preinstallation setup tasks (cluster nodes)	9
Installation procedure	11
Preparing to install Oracle RAC 10g R2 on cluster nodes	
Installing Oracle RAC 10g R2 CRS	13
Installing Oracle RAC 10g R2 software	33
Trademarks and special notices	42



Abstract

Oracle on Linux is the leading enterprise combination, delivering highly performing enterprise database environments. Oracle Real Application Clusters 10g Release 2 with Oracle Cluster Ready Services on a 64-bit SUSE Enterprise Linux 10 operating system is certified in a network file system environment with an IBM System Storage N series. This technical report covers the certified installation of Oracle on Linux in an IBM N series NFS environment.

Introduction

This technical report covers the installation of IBM® System Storage™ N series in a network file system (NFS) environment of servers running the 64-bit SUSE Enterprise Linux® 10 operating system (OS) and the database Oracle Real Application Clusters (RAC) 10*g* Release 2 (R2) with Oracle Cluster Ready Services (CRS). The components presented in this paper must be used in the same combination to gain support from all parties. The only exception to this is the application of certain patches, as defined and required by vendors. This document also covers the patches and recommendations for running Oracle RAC 10*g* R2 on IBM N series storage in an NFS environment.

Assumptions

This document assumes that readers are familiar with Oracle RAC 10*g* R2 and IBM N series storage systems. It also assumes that readers are familiar with the operation of the SUSE Enterprise Linux 10 environment and installation of Oracle Database patches and any relevant Linux RPMs. It is also important to be familiar with all networking terminology and implementations.

System environment

The configuration presented in this document is based on the Oracle RAC 10*g* R2 certification environment specified in Figure 1.



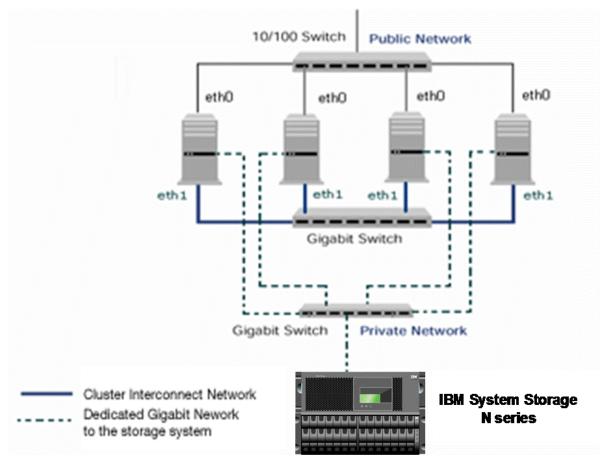


Figure 1) Oracle RAC 10g R2 on 64-bit SUSE Enterprise Linux 10 servers with IBM N series storage.

Figure 1 illustrates a typical configuration of Oracle RAC 10*g* R2 with IBM N series storage and 64-bit servers running SUSE10. This is a scalable configuration that allows users to scale horizontally and internally in terms of processor, memory, and storage.

As shown in Figure 1, N series recommends that you dedicate a private network connection between the Oracle RAC 10*g* R2 servers and the N series storage. This is accomplished by using a dedicated gigabit network (with a gigabit switch) to the N series storage. A dedicated network connection is recommended as a best practice for the following reasons:

- In an Oracle RAC 10*g* R2 environment, it is important to eliminate any contentions and latencies.
- Providing a separate network ensures security.

The cluster interconnect is an essential part of 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.



Requirements

Hardware

Cluster nodes:

- Four IBM xSeries® Servers—Model System X-3650 64-bit 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 IBM N series storage).

Storage infrastructure:

- One IBM System Storage N series system with IBM System Storage N series 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.

Software

For all nodes in the participating cluster, unless specified otherwise:

- SUSE Enterprise Linux 10 64-bit operating system
- Oracle RAC 10g R2 (10.2.0.1), with RAC license
- Oracle 10*g* R2 patch 3.



Setup for IBM N series storage

- 1.Configure an IBM N series storage system running Data ONTAP 7.22 or later and with NFS and IBM System Storage N series with SnapRestore license keys.
- 2. Create and export volumes for storing Oracle database files on the storage:
 - a. Create four volumes on the storage (Storage1):

orahome Shared Oracle and CRS Home (binaries)
oradata Oracle data files and control files
oralog CRS files
oralogs Database logs, a copy of control file and archive logs

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

```
Storagel> 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.

Add the following entries to the /etc/exports file on N series storage (Storage1):

```
/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:

```
Storage1> exportfs -a
```

Note: It is recommended that you use flexible volumes in your database environment. IBM System Storage N series with FlexVol® technology pools storage resources automatically and enables you to create multiple flexible volumes on a large pool of disks. This flexibility means that 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; your setup may vary depending on your requirements.



Operating-system configuration

Patches

Before you install Oracle RAC 10g R2 , the following RPMs must be applied on the servers. Some of these RPMs may have already been applied to your system. Be sure to verify whether 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
```

Here is a list of required patches. If any of the patches are not installed or committed, install them.

```
binutils
    compat-libstdc++
    gcc
    gcc-c++
    gcc-objc
    glib
    glib2
    glibc
    glibc-devel
    libaio
   libaio-devel
    libgcc
   libgcj
   libgcj-devel
    libobic
   libstdc++
    libstdc++-devel
    openmotif
    perl
    tcl
    unzip
    zip
    tar
Java™ for Oracle (JDBC/OCI Drivers)
    IBM Java 1.4.2. 64-bit Standard Edition or later
    Java-1_4_2-ibm-1.4.2.s4-23.4
```



Operating-system settings

On SUSE systems, the default ulimits for individual users are set in /etc/security/limits.conf. As a root user, add following entries in /etc/security/limits.conf to specify oracle user's limits.

```
# 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 must be done on all nodes of the cluster. A server reboot is required to activate updated limits. After you modify the settings, the 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, -1) 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 above setting for the oracle user.
```

Kernel settings

As a root user, add the following parameters for the shared memory and semaphores to the /etc/sysctl.conf file.

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

The oracle user account used to install Oracle 10.2.0.1 should not have the Oracle installation-related variables set by default. For example, ORACLE_HOME, PATH, and LD_LIBRARY_PATH should not be set to include Oracle binaries in .profile file, .login file and /etc/profile.d.

Also, no LD_ASSUME_KERNEL value should be used with SUSE Enterprise Linux 10.



Preinstallation setup tasks (cluster nodes)

This information is intended for all nodes that are participating in the cluster, unless specified otherwise.

- 1. Have all the SUSE Enterprise Linux 10 64-bit servers ready with the latest recommended patches and OS settings, as discussed in the section, *Operating System Configuration*.
- 2.Install and configure NICs in the cluster nodes (three per node).
 - a. Public IP: As indicated by name.
 - Private interconnects: Connect one gigabit NIC back to back to the other node for cluster interconnects.
 - c. Server connection to N series storage: Connect one gigabit NIC to the gigabit switch, which will connect to the gigabit NIC on the N series storage.
- 3. Configure the network interfaces on each node.
 - a. Configure the three network interfaces as indicated:

```
# btc-srv1 (Host 1)
eth0 - IP: 10.73.68.162, Netmask: 255.255.254.0
eth1 - IP: 192.168.73.55, Netmask: 255.255.255.0
eth2 - IP: 10.73.70.162, Netmask: 255.255.255.0
# btc-srv2 (Host 2)
eth0 - IP: 10.73.68.164, Netmask: 255.255.254.0
eth1 - IP: 192.168.73.56, Netmask: 255.255.255.0
eth2 - IP: 10.73.70.164, Netmask: 255.255.255.0
# btc-srv3 (Host 3)
eth0 - IP: 10.73.68.161, Netmask: 255.255.254.0
eth1 - IP: 192.168.73.53, Netmask: 255.255.255.0
eth2 - IP: 10.73.70.161, Netmask: 255.255.255.0
# btc-srv4 (Host 4)
eth0 - IP: 10.73.68.166,
                          Netmask: 255.255.254.0
eth1 - IP: 192.168.73.54, Netmask: 255.255.255.0
eth2 - IP: 10.73.70.166,
                          Netmask: 255.255.255.0
```

Where:

Interface eth0 is the public IP for each node.

Interface eth1 on all cluster nodes is connected to a switch for cluster private interconnects.

Interface eth2 on all cluster nodes is connected to the gigabit switch for storage I/O.



b. Update the /etc/hosts file on the cluster nodes and add entries for public, private, and VIP addresses. Note that in addition to the preconfigured public and private network, Oracle Database 10g requires additional IP addresses that are 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 transfers the VIP address to another surviving node. You should add the VIP to the /etc/hosts file on all nodes in the cluster as well as to all nodes accessing the database.

A sample for /etc/hosts entries:

T		
Internet Address	Hostname Com	mments
10.73.68.162	btc-srv1	btc-srv1.btcppe.ibmn.com
192.168.73.55	btc-srv1-i	
10.73.68.204	btc-srv1-v	btc-srv1-v.btcppe.ibmn.com
10.73.70.162	btc-srv1-stor	
10.73.68.164	btc-srv2	btc-srv2.btcppe.ibmn.com
192.168.73.56	btc-srv2-i	
10.73.68.205	btc-srv2-v	btc-srv2-v.btcppe.ibmn.com
10.73.70.164	btc-srv2-stor	
10.73.68.161	btc-srv3	btc-srv3.btcppe.ibmn.com
192.168.73.53	btc-srv3-i	
10.73.68.202	btc-srv3-v	btc-srv3-v.btcppe.ibmn.com
10.73.70.161	btc-srv3-stor	
10.73.68.166	btc-srv4	btc-srv4.btcppe.ibmn.com
192.168.73.54	btc-srv4-i	
10.73.68.203	btc-srv4-v	btc-srv4-v.btcppe.ibmn.com
10.73.70.166	btc-srv4-stor	
10.73.70.113	Storage1	

- 4. Ensure the connectivity of each interface by using the ping command (interconnects, public IPs, and storage IPs).
- 5.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:

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



Where:

Storage1 is the name of the IBM N series storage system.

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

orahome Shared Oracle and CRS Home (binaries)

oradata Oracle data files and control files

oral0q CRS files

oralogs database logs, a copy of control file and archive logs.

orallog is just a separate mount point for CRS files. CRS files (cluster registry file and voting disk file) can reside in the same /orallog volume but must be mounted with the noac mount option.

During the Oracle CRS installation, be sure to indicate a path starting with /oral0g directory when prompted for the cluster registry file (OCR) and voting disk (Cluster Synchronization Services, or CSS) file location.

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

6. Create the following mount points on all cluster nodes:

```
#mkdir /oradata
#mkdir /oralogs
#mkdir /orahome
#mkdir /oralog
```

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.

Installation procedure

This section provides instructions for installing Oracle RAC 10g R2 install on SUSE Enterprise Linux 10.

Preparing to install Oracle RAC 10g R2 on cluster nodes

- 1. This document assumes the oracle user account and the group to be oracle and dba, respectively, on all cluster nodes. The user ID and group name for the oracle account should be the same on all cluster nodes.
- 2. Grant appropriate permissions to the oracle user on all shared mounted volumes, /oradata, /orahome, and /oralogs:

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

Repeat this step for the orahome, orallog, and orallogs 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-srv1 oracle
btc-srv2 oracle
btc-srv3 oracle
```



btc-srv4 oracle

btc-srv1-i oracle btc-srv2-i oracle btc-srv3-i oracle btc-srv4-i oracle

4. Log in as the oracle user from all cluster nodes and then test the oracle account equivalence using a remote shell utility such as rsh.

btc-srv1#su – oracle \$rsh btc-srv1 pwd \$rsh btc-srv2 pwd \$rsh btc-srv3 pwd \$rsh btc-srv4 pwd

btc-srv2#su – oracle \$rsh btc-srv1 pwd \$rsh btc-srv2 pwd \$rsh btc-srv3 pwd \$rsh btc-srv4 pwd

btc-srv3#su – oracle \$rsh btc-srv1 pwd \$rsh btc-srv2 pwd \$rsh btc-srv3 pwd \$rsh btc-srv4 pwd

btc-srv4#su – oracle \$rsh btc-srv1 pwd \$rsh btc-srv2 pwd \$rsh btc-srv3 pwd \$rsh btc-srv4 pwd



Installing Oracle RAC 10 *g* **R2 CRS**

Before installing Oracle RAC 10*g* R2, Oracle CRS must be installed, configured, and started. Refer to the appropriate Oracle RAC 10*g* R2 Installation and Configuration Guide. This section briefly describes the procedures for using the Oracle Universal Installer (OUI) to install CRS.

Note: 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 Oracle RAC 10g database home. That is, ORACLE_HOME and CRS_HOME must be different locations. For the SUSE Enterprise Linux 10 platform, the OUI must be invoked with the -ignoresysPreregs option. Otherwise the installation of Oracle 10.2.0.1 will fail while checking the OS version, because the Oracle 10.2.0.1 CD doesn't list SUSE10 in the oraparam.ini file.

 Run the ./runInstaller -ignoreSysPrereqs command from the /crs subdirectory on the Oracle CRS Release 2 (10.2.0.1) CD-ROM or from the staging area where you have dumped the CRS software. This is a separate CD that contains the CRS software. This document assumes that the OUI is started from node 1 (btc-srv1). When the OUI displays the Welcome page, click Next.



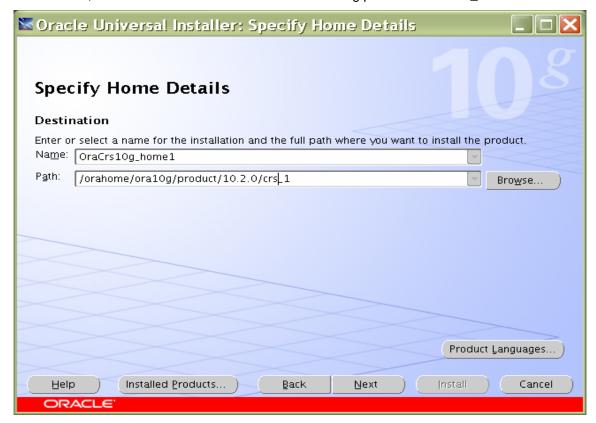


2.On the Specify Inventory Directory and Credentials page, enter a nonshared location for Oracle Inventory. This is the only part of Oracle10*g* that should not be shared. For this test, we used /home/oracle/oraInventory for the Oracle Inventory information. Click Next.



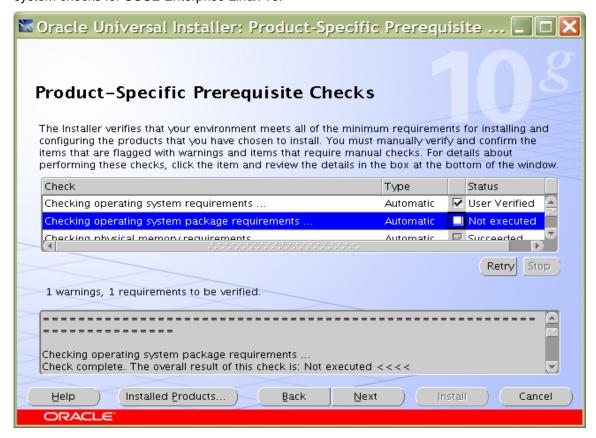


3. The Specify Home Details page contains predetermined information for the source of the installation files and the target destination information. Specify the destination path for the shared CRS home. The path should be on a shared file system and different from \$ORACLE_HOME. In this exercise, the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



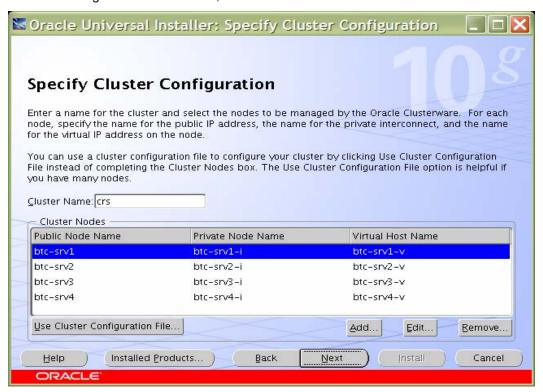


4.On the Product-Specific Prerequisite Checks page, select the check boxes to change the status to User Defined, because the Oracle 10.2.0.1 CD does not by default execute all the operating system checks for SUSE Enterprise Linux 10.



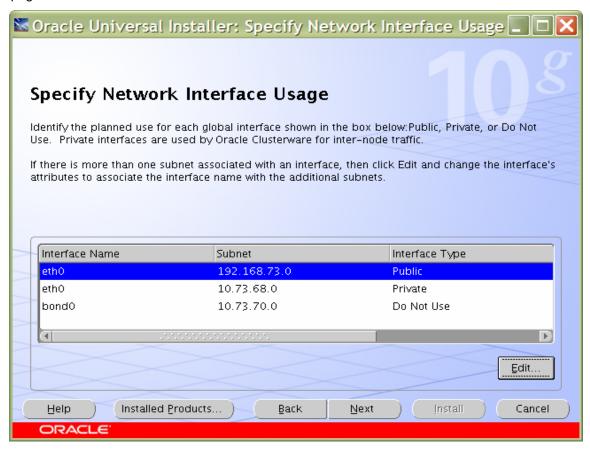


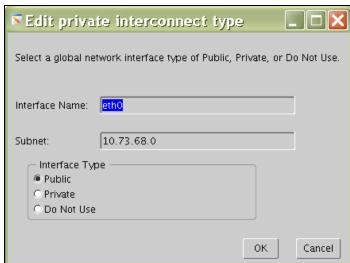
5.On the Specify Cluster Configuration page, specify the cluster name, public names (hostnames), private names, and virtual hostnames to be used for the RAC cluster. In this case, the public names are btc-srv1, btc-srv2, btc-srv3, and btc-srv4; the private names are btc-srv1-i, btc-srv2-i, btc-srv3-i, and btc-srv4-i; and the virtual hostnames are btc-srv1-v, btc-srv2-v, btc-srv3-v, and btc-srv4-v. After adding all the node names, click Next.





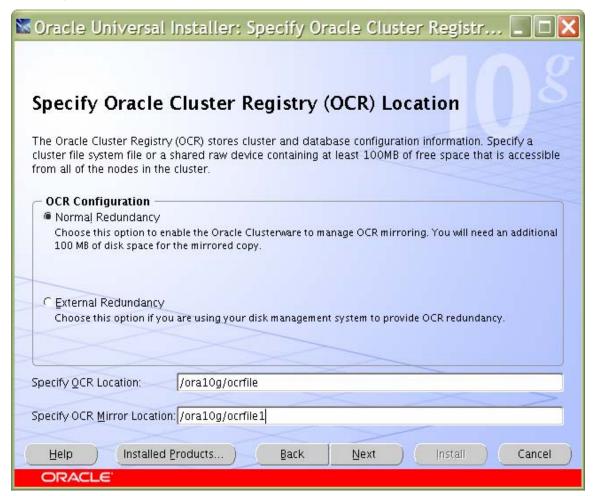
6.On the Specify Network Interface Usage page, specify the private network to be used for the cluster interconnect. This is a very important step. Do not leave it set to the default, which is Do Not Use. In this case, eth1 (btc-srv1-i) was used as the private interconnect and eth0 (btc-srv1) was used as the public interface. Select the interface and click the Edit button in the Edit Private Interconnect dialog box to modify it. Click OK to return to the Specify Network Interface Usage page. Click Next.





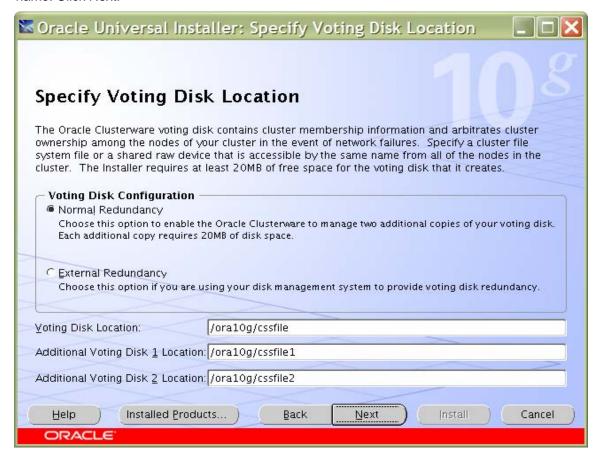


7.On the Specify Oracle Cluster Register (OCR) Location page, specify the path for the OCR file. Be sure to specify the full path to a shared location, along with the name of the file. If you want normal redundancy, do the same for a mirror file. In this case, we used /oral0g/ocrfile and /oral0g/ocrfilel. Click Next.



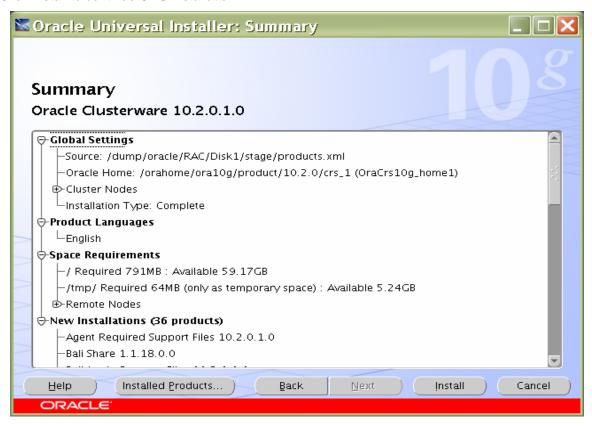


8.On the Specify Voting Disk Location page, specify the CSS voting disk file location. We used /oral0g/cssfile for CSS services. For normal redundancy, specify the path along with the name. Click Next.

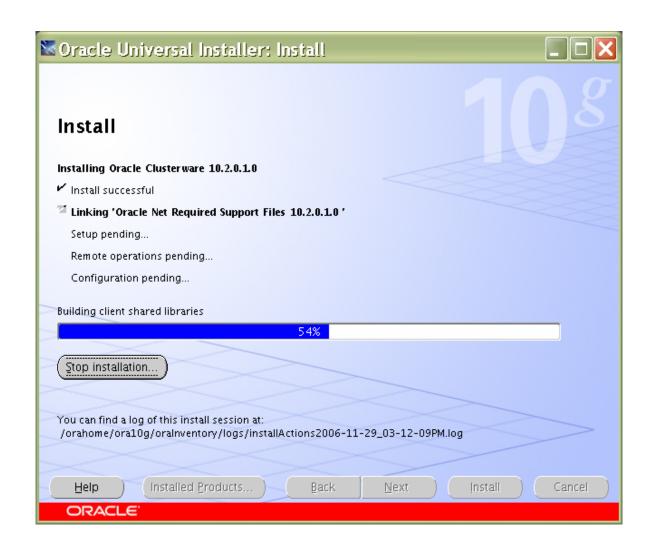




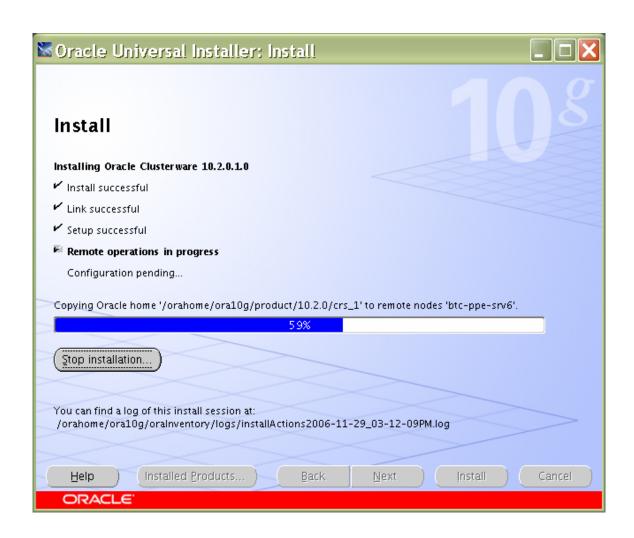
9. Click Install to continue CRS Installation.





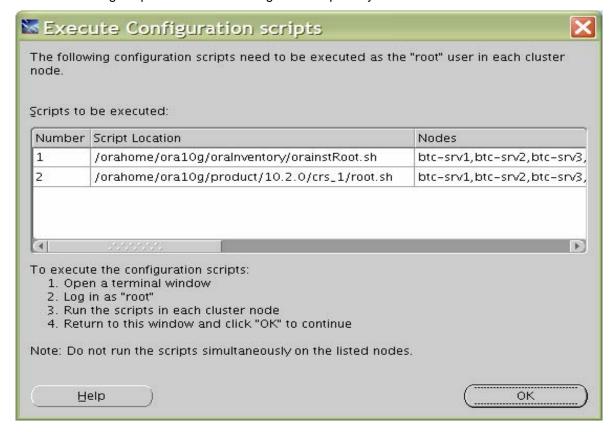






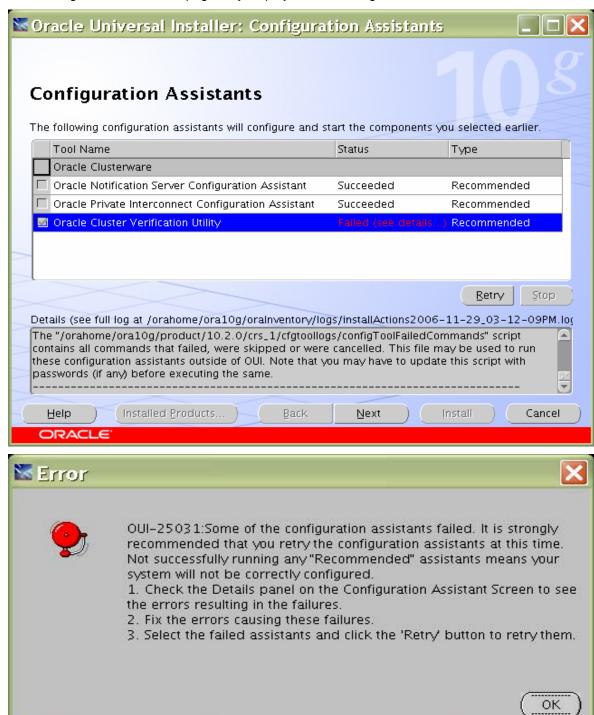


10. Run the following script as root user starting from the primary node.

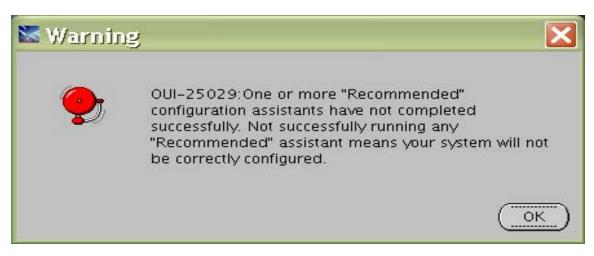




11. The Configuration Assistants page may display some warnings. Click OK to continue.







12. Before running vipca, the VIP Configuration Assistant, you must edit the vipca file in the CRS bin directory to prevent the following error:

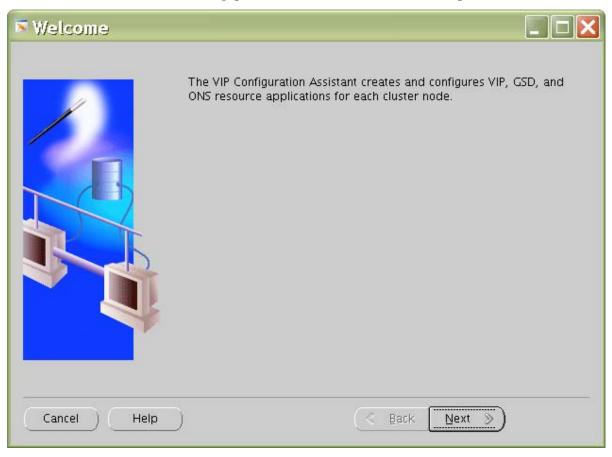
Oracle CRS stack installed and running under init(1M)
Running vipca (silent) for configuring nodeapps
/home/oracle/crs/oracle/product/10/crs/jdk/jre//bin/java: error while
loading shared libraries: libpthread.so.0: cannot open shared object
file: No such file or directory

Edit the vipca file in the \$ORA_CRS_HOME/bin directory (that is, /orahome/oral0g/product/10.2.0/crs_1/bin) to undo the setting of LD_ASSUME_KERNEL. After the IF statements, around line 120, add an unset command to ensure that LD ASSUME KERNEL is *not* set:



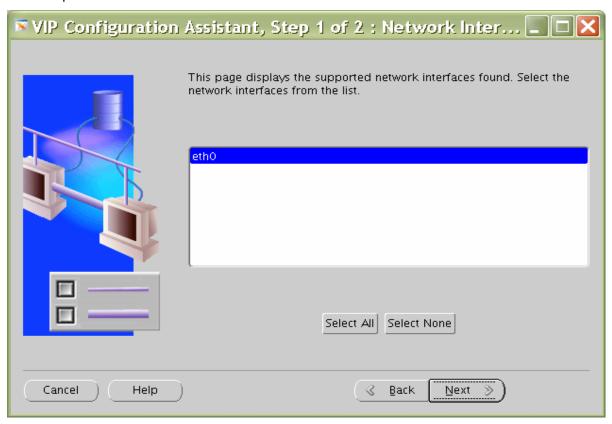
Then run vipca from the \$ORA_CRS_HOME/bin directory as root user on master node (btc-srv1). Click next.

btc-srv1:/orahome/ora10g/product/10.2.0/crs_1/bin/#. /vipca



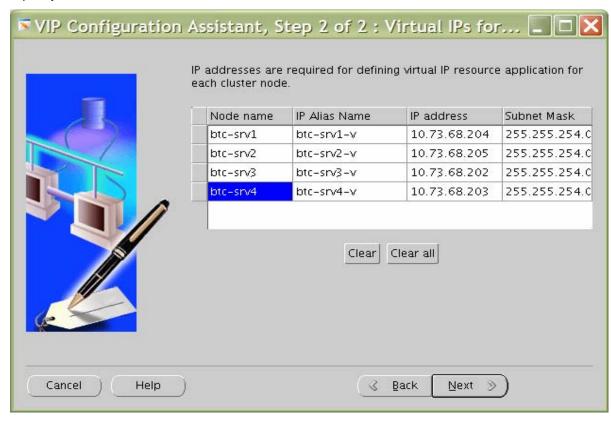


13. Select a public interface. Click Next.



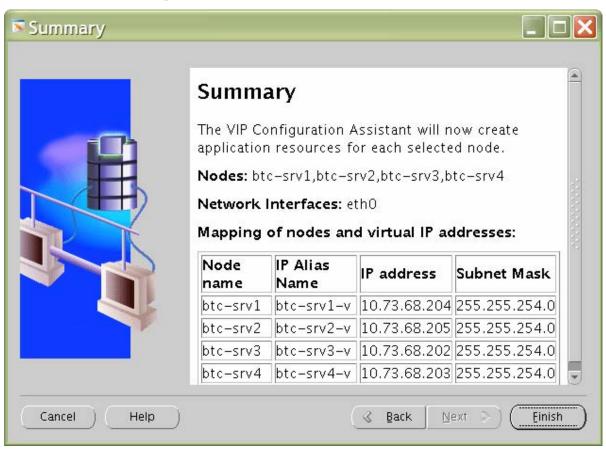


14. Specify the virtual IP address and subnet mask of each node. Click Next.



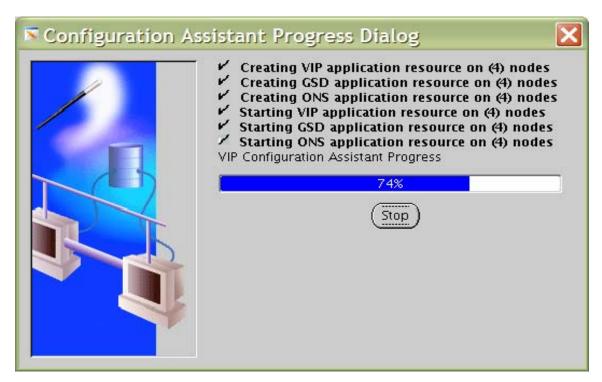


15. Click Finish to continue vipca.

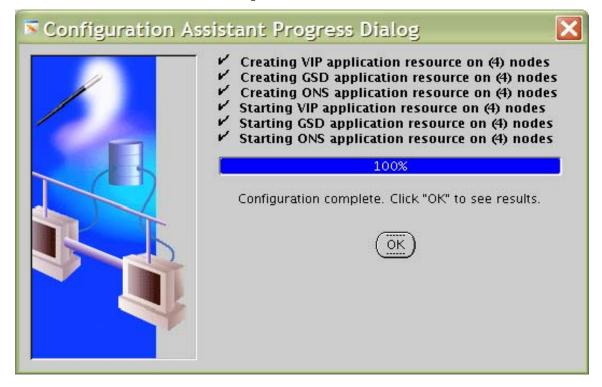




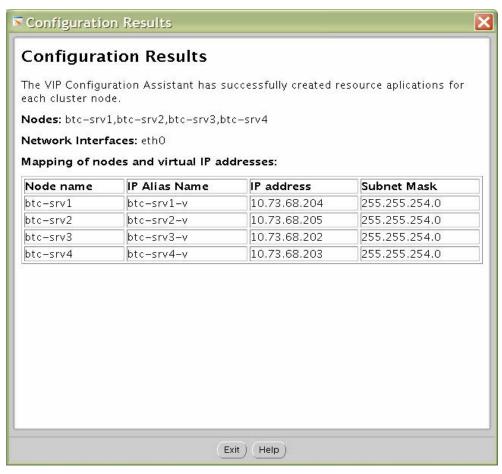




16. Click OK and then click Exit to finish vipca.







17. To verify the CRS installation, execute the olsnodes command from the \$CRS_HOME/bin directory. The olsnodes command syntax is:

Where:

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

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



Installing Oracle RAC 10g **R2 software**

 After making sure that Oracle CRS has started on the cluster nodes, start runInstaller from Disk1 of the Oracle RAC 10g R2 CDs or the staging area where you have dumped the Oracle RAC 10g R2 software.





2. On the Select Installation Type page, select Enterprise Edition and click Next.





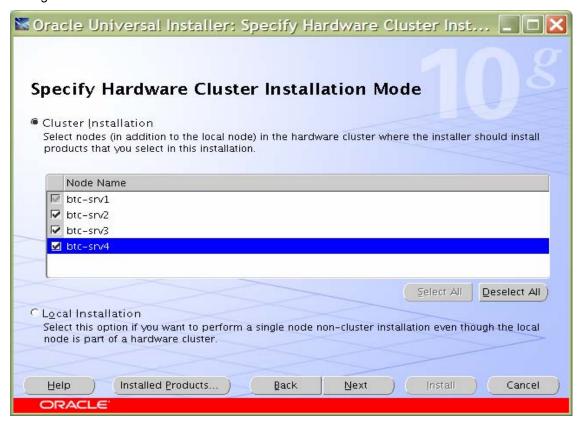
3. On the Specify Home Details page, enter the destination path for the shared ORACLE_HOME. This should be a different location than the shared CRS Home. For this exercise, the shared ORACLE_HOME was /orahome/ora10g/product/10.2.0/db_1.





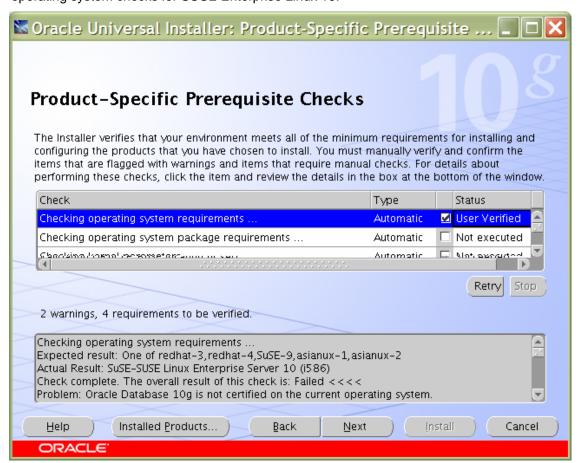
4. On the next page, select Cluster Installation and then select all the nodes in the cluster. For this exercise, four cluster nodes were selected: btc-srv1, btc-srv2, btc-srv3, and btc-srv4. Click Next.

Note: If the nodes are not displayed in the cluster node selection, then Oracle CRS is not configured or started on those cluster nodes.



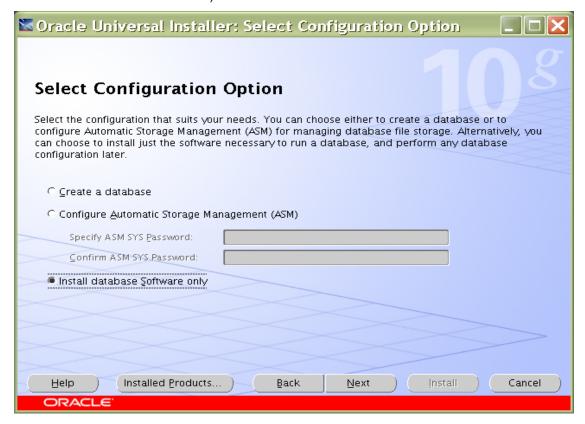


 On the Product-Specific Prerequisite Checks page, select the check boxes to change the status to User Defined, because the Oracle 10.2.0.1 CD does not by default execute all the operating system checks for SUSE Enterprise Linux 10.



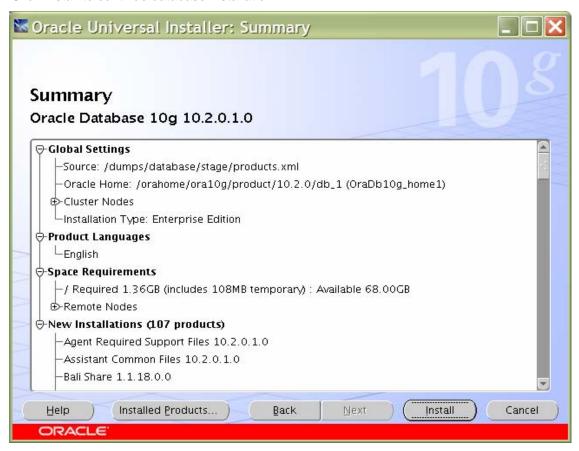


6. On the Select Configuration Option page, select Install Database Software Only. (We used DBCA to create a database later.) Click Next.



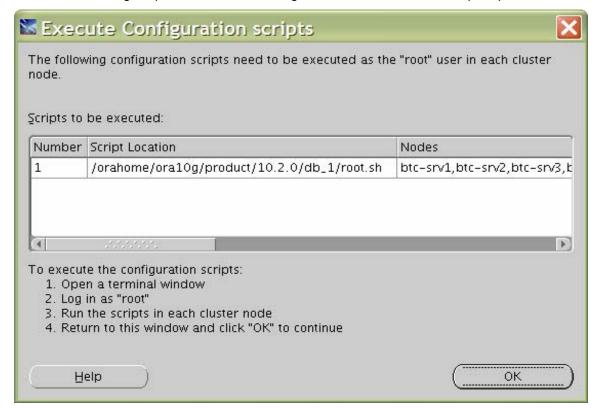


7. Click Install to continue database installation



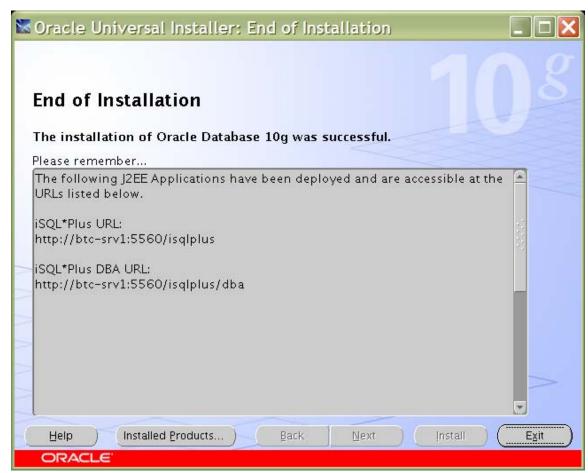


8. Run the following scripts as root user, starting from the master node when prompted.





9. Click Exit to finish the database installation.



The srvctl binary file must be modified in order to prevent following error:

/orahome/ora10g/product/10.2.0/db_1/jdk/jre/bin/java: error while loading shared libraries: libpthread.so.0: cannot open shared object file: No such file or directory

Edit the srvctl file in both \$ORA_CRS_HOME/bin and \$ORACLE_HOME/bin directories to unset the LD_ASSUME_KERNEL by adding one line, around line 180:

Note: Install Oracle RAC 10*g* R2 Patch 3 on both CRS_HOME and ORACLE_HOME using OUI. For details about the patch installation, refer to the Oracle *Patch Installation Guide*.



Trademarks and special notices

© International Business Machines 1994-2008. IBM, the IBM logo, System Storage, xSeries and other referenced IBM products and services are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. All rights reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

FlexVol, Data ONTAP, Network Appliance, the Network Appliance logo and SnapRestore are trademarks or registered trademarks of Network Appliance, Inc., in the U.S. and other countries.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.