Oracle® Cluster File System

Installation Notes

Release 1.0 for Red Hat Linux Advanced Server 2.1

November 2002

Part No. B10499-01

The purpose of this document is to provide step-by-step instructions on how to install Oracle Cluster File System (OCFS) on Red Hat Advanced Server 2.1.

These instructions assume that you have installed on your system Red Hat Advanced Server 2.1 and Oracle9*i* Release 2 (9.2.0.2.x) with Real Application Clusters (RAC). You must ensure that the Oracle distribution is installed on a local filesystem.

Read this document carefully before starting to ensure that Linux OCFS installation and configuration is successful.

1 Support

Note that OCFS is currently supported only on Red Hat Advanced Server 2.1 using SMP or Enterprise kernel, and without additional modifications or patches, except for the one provided by Oracle. If you modify the kernel, then Oracle Corporation cannot support it.

2 Where to obtain OCFS for Linux

Download OCFS for Linux in compiled form at the following Web site:

`http://otn.oracle.com/tech/linux/content.html`

You must download the following two RPM packages:

- `ocfs-support-1.0-1.i686.rpm`
- `ocfs-tools-1.0-1.i686.rpm`

In addition, you must download the kernel module RPM, `ocfs-2.4.9-3`*`typeversion`*`.rpm`, where the variable *`typeversion`* stands for the type and version of the kernel you are using.

Use the following command to find out which Red Hat kernel version is installed on your system:

```
uname -a
```

The alphanumeric identifier at the end of the kernel name indicates the kernel version you are running. Download the kernel module that matches your kernel version.

For example, if the kernel name returned with the `uname` command ends with with `-e.3smp`, then you would download the following kernel module (note text in bold font):

ocfs-2.4.9**`-e.3-smp`**-1.0-1.i686.rpm

Follow the instructions in the README file on how to build the module. Ensure that you use the SMP or Enterprise kernel shipped with Red Hat Advanced Server 2.1 without any patches or customization. If you modify the kernel, then Oracle Corporation cannot support it.

3 Installation

Complete the following procedure to prepare the environment to run OCFS. Note that you must perform all steps as `root`, and that each step must be performed on all nodes of the cluster.

1. Install RPM files.

   First Install the support RPM file, `ocfs-support-1.0.-1.i686.rpm`, and then install the tools RPM file, and the correct kernel module RPM file for your system.

   To install the files, enter the following command:

   ```
   rpm -i ocfs_rpm_package
   ```

   where the variable *`ocfs_rpm_package`* is the name of the rpm package that you are installing.

   For example, to install the kernel module RPM file for an e.3 enterprise kernel, you would enter the following command:

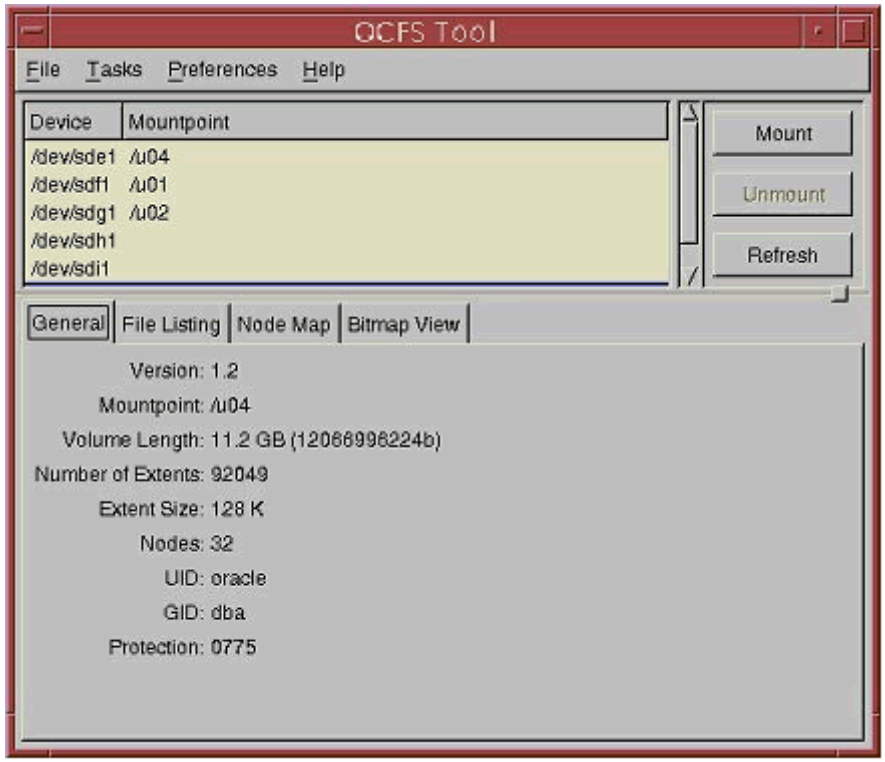
   ```
   rpm -i ocfs-2.4.9-e.3-enterprise-1.0-1.i686.rpm
   ```

2. Using the utility `ocfstool`, generate the file `/etc/ocfs.conf`. Start up `ocfstool` as shown in the following example:

```
# DISPLAY=:0.0
# export DISPLAY
# /usr/bin/ocfstool
```

3. The *OCFS Tool* window appears (Figure 1). Click in the window to make it active, and either type CTRL-G, or from the menu select **tasks**, and then select the option **Generate Config**.

Figure 1 OCFS Tool Window



4. The OCFS Generate Config window opens. Check the values displayed in the window to confirm that they are correct, and then click the **OK** button (Figure 2). Based on information gathered from your installation, the ocfstool utility will generate the file `/etc/ocfs.conf`.

Figure 2 OCFS Generate Config Window

5. After the generation is completed, open the file /etc/ocfs.conf in a text file tool. It should look like the following example:

```
#
# ocfs config
# Ensure this file exists in /etc#

         node_name = serena
         node_number =
         ip_address = 130.35.149.132
         ip_port = 7000
         guid = 93DBFD4CDE57335D244500D0B7E4D526
```

6. To have the module `ocfs.o` loaded on startup, create the script `/etc/init.d/dbora`, as shown in the following example. In addition, "APPENDIX: Scripts Reference" lists the script `/var/opt/oracle/soft_startup.sh`.

```
#!/bin/sh
#
# This script will load the ocfs module, mount the ocfs filesystems,
start the Oracle Cluster
# Manager and the GSD.
#

echo "Loading OCFS Module"
/usr/sbin/load_ocfs
echo "Mounting OCFS FS"
/bin/mount -a -t ocfs

ORACLE_HOME=/usr/oracle/product/9.2.0.1
export ORACLE_HOME
/var/opt/oracle/soft_startup.sh # start the Oracle Cluster Manager
#
# Starting GSD
```

```
#
su - oracle -c "${ORACLE_HOME}/bin/gsdctl start"
#
```

7. Using the utility `fdisk`, partition the disk to create the OCFS filesystem according to your needs. Oracle Corporation recommends that you partition your system in accordance with Oracle Optimal Flexible Architecture (OFA) standards.

8. Once the partitions are created, use the following command to create the mount points for the OCFS filesystem:

   ```
   mkdir -p /u01 /u02 /u03
   ```

   Make a note of these mount points, as you will need them in step 11.

9. Start the utility `ocfstool`:

   ```
   # DISPLAY=:0.0
   # export DISPLAY
   # /usr/bin/ocfstool
   ```

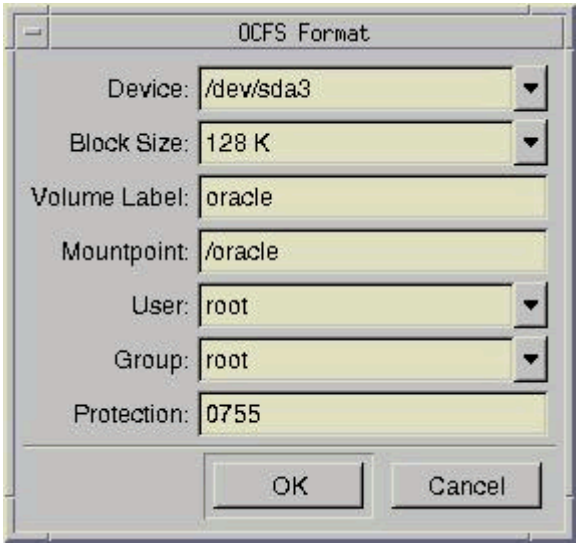
10. The *OCFS Tool* window appears, as shown in Figure 1. Click in the window to make it active, and either type CTRL-F, or from the menu select **tasks**, and select the option **Format**.

11. The *OCFS Format* window appears (Figure 3). The values in the text boxes are used to format the partitions and mount the filesystems.

    Fill the text field boxes according to the specifications for your system. The block size setting must be a multiple of the Oracle block size. Oracle Corporation recommends that you do not change the default block size, which is set to 128. Set the value for the text field **User** to `oracle` and the value for the text field **Group** to `dba`. Set the values for the text field **Volume label** and **Mountpoint** to the values you set in step 8, and then click the **OK** button. Formatting begins.

    The amount of time it will take to format and mount partitions will depend on the speed of your system disk drives and CPU.

Figure 3 OCFS Format Window



If you prefer to format OCFS using the command line, then use the following code example as a model:

```
# mkfs.ocfs  -F -b 128 -L /u04 -m /u04 -u 1001 -g 1001 -p 0775
/dev/sde1
```

where the syntax for ocfstool is:

```
mkfs.ocfs -b block-size [-C] [-F] [-g gid] -L volume-label
     -m mount-path [-n] [-p permissions] [-u uid] [-v] [-V] device
```

with the following options:

- -b Block size in kilo bytes
- -C Clear all data blocks
- -F Force format existing OCFS volume
- -g GID for the root directory
- -L Volume label
- -m Path where this device will be mounted
- -n Query only
- -p Permissions for the root directory
- -q quiet execution

- -u UID for the root directory
- -V Print version and exit

Note: After the partition is properly formatted, you must mount partitions initially one by one. When you mount each node for the first time, no other node should be attempting to mount the file systems.

OCFS requires this procedure for the initial mount to allow OCFS to initialize properly the filesystem.

To perform an individual mount, use the following mount command syntax:

mount -t ocfs /dev/ *device/mountpoint*

---

12. **OPTIONAL:** To mount the filesystem automatically on startup, add lines similar to the following to the `/etc/fstab` file:

```
/dev/sdf1       /u01      ocfs    uid=1001,gid=1001
/dev/sdg1       /u02      ocfs    uid=1001,gid=1001
/dev/sdh1       /u03      ocfs    uid=1001,gid=1001
```

In the following code example, the lines you need to add are shown in bold typeface, and placed in the proper location in the `etc/fstab` file. Ensure that you mount the OCFS filesystem in sequence, node after node, and wait for each mount to complete before starting the mount on the next node.

```
LABEL=/         /         ext3    defaults        1 1
LABEL=/boot     /boot     ext3    defaults        1 2
none            /dev/pts       devpts  gid=5,mode=620  0 0
LABEL=/oracle   /oracle   ext3    defaults        1 2
none            /proc     proc    defaults        0 0
none            /dev/shm   tmpfs   defaults        0 0
LABEL=/tmp      /tmp       ext3    defaults        1 2
LABEL=/usr      /usr       ext3    defaults        1 2
LABEL=/var      /var       ext3    defaults        1 2
/dev/sdb2       swap      swap    defaults        0 0
/dev/sdb3       swap      swap    defaults        0 0
/dev/sdc1       swap      swap    defaults        0 0
/dev/sdd1       swap      swap    defaults        0 0
/dev/fd0        /mnt/floppy    auto    noauto,owner,kudzu 0 0
/dev/sdf1       /u01      ocfs    uid=1001,gid=1001
/dev/sdg1       /u02      ocfs    uid=1001,gid=1001
/dev/sdh1       /u03      ocfs    uid=1001,gid=1001
```

See Also: "Known Limitations" for further information on the `/etc/fstab` file.

13. Reboot the system. Upon reboot, the module should be loaded and the filesystems properly mounted.

4 Operating System Setup Tasks

Compete the following operating system configuration tasks.

4.1 System Parameter Configuration

You must change some of the system parameters to accommodate Oracle9i RAC and OCFS. Use the script `/etc/init.d/rhas_ossetup.sh` as provided in the following code example to perform this configuration. Using this script ensures that your system is correctly configured, and will help to avoid problems.

```
#!/bin/sh
#
# /etc/init.d/rhas_ossetup.sh
#
# This script will set the system parameter for use with Oracle9i RAC
and OCFS.
#
echo "65536 " > /proc/sys/fs/file-max
echo "2147483648" > /proc/sys/kernel/shmmax
echo "4096" > /proc/sys/kernel/shmmni
echo "2097152" > /proc/sys/kernel/shmall
echo 1024 65000 > /proc/sys/net/ipv4/ip_local_port_range
echo "1276 2552 3828 " > /proc/sys/vm/freepages
ulimit -u 16384
echo "100 32000 100 100" > /proc/sys/kernel/sem
ulimit -n 65536
```

Note that the settings in the preceding code example are valid for a cycle only, meaning, it is automatically reset to its original values upon reboot. To make the process automatic during the startup of the system, enter the following commands as `root`:

```
# ln -s /etc/init.d/rhas_ossetup.sh /etc/rc5.d/S77rhas_ossetup
# ln -s /etc/init.d/rhas_ossetup.sh /etc/rc3.d/S77rhas_ossetup
```

4.2 Swap Memory Partition Configuration

You must allocate at least 8 GB to the swap partition. As root, use the command `swapon -s` to verify that you have enough disk space allocated.

If you require more disk space, use the command `swapon -a`. Note that you can create a swap partition with a maximum size of 2Gb.

To have the swap automatically set on startup, add lines similar to the following to the `/etc/fstab` file:

```
/dev/sdb2 swap swap defaults 0 0
/dev/sdb3 swap swap defaults 0 0
/dev/sdc1 swap swap defaults 0 0
/dev/sdd1 swap swap defaults 0 0
```

In the following code example, the lines you need to add are shown in bold typeface, and placed in the proper location in the `etc/fstab` file.

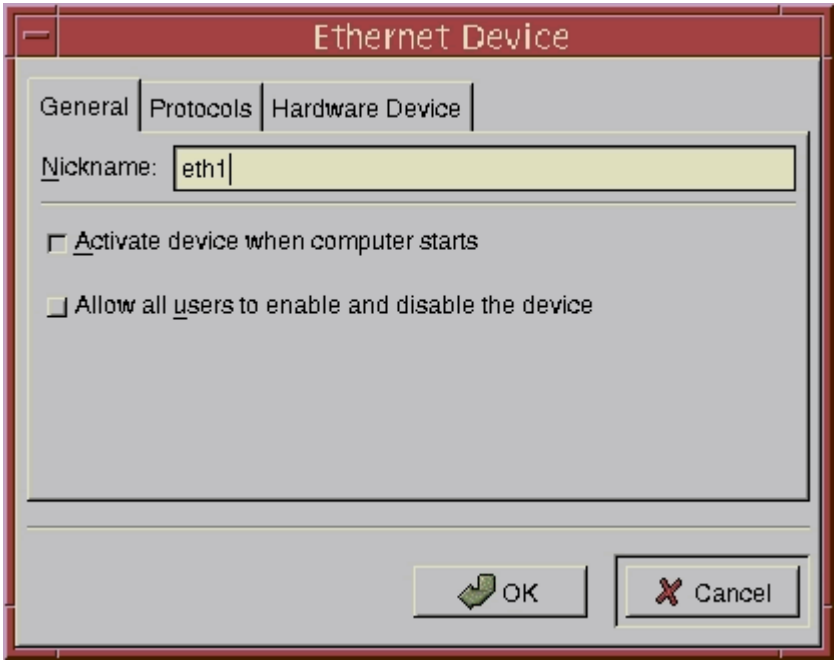
```
LABEL=/       /              ext3    defaults        1 1
LABEL=/boot   /boot          ext3    defaults        1 2
none          /dev/pts       devpts  gid=5,mode=620  0 0
LABEL=/oracle  /oracle       ext3    defaults        1 2
none          /proc          proc    defaults        0 0
none          /dev/shm       tmpfs   defaults        0 0
LABEL=/tmp    /tmp           ext3    defaults        1 2
LABEL=/usr    /usr           ext3    defaults        1 2
LABEL=/var     /var          ext3    defaults        1 2
/dev/sdb2      swap          swap    defaults        0 0
/dev/sdb3      swap          swap    defaults        0 0
/dev/sdc1      swap          swap    defaults        0 0
/dev/sdd1      swap          swap    defaults        0 0
/dev/fd0       /mnt/floppy   auto    noauto,owner,kudzu 0 0
/dev/sdf1      /u01          ocfs    uid=1001,gid=1001
/dev/sdg1      /u02          ocfs    uid=1001,gid=1001
```

4.3 Network Adapter Configuration

You must have the network consistently coming up during reboot. To ensure that all Network adapters will be automatically enabled and in the correct order, complete the following tasks:

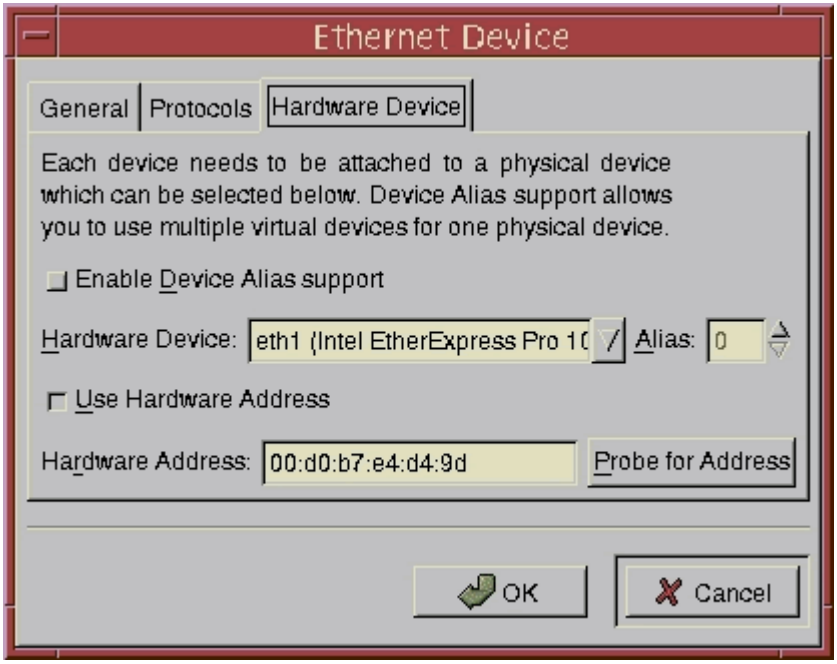
1. Ensure that you have the DISPLAY variable set to an authorized X Client, and launch the program `/usr/sbin/redhat-config-network`.
2. The Ethernet Device window opens (Figure 4). Click the checkbox for the option **Activate device when computer starts**, and click the **OK** button.

Figure 4 Ethernet Device Window: General Tab



3. Click the tab **Hardware Devices** (Figure 5). Click the checkbox for the option **Use Hardware Address option** box, and click the **Probe for Address** button to populate the field **Hardware Address**. Click the **OK** button to save the changes.

Figure 5 Ethernet Device Window: Hardware Device Tab



4. Ensure that the public and private node names of all member nodes in the Oracle Real Application Cluster are listed in the file `/etc/hosts`.

5 Oracle Patch Installation

Install the patches required for Oracle9*i* Release 2 (9.2.0.2).

6 Known Limitations

As is always the case when you modify or upgrade the Oracle database, ensure that you make backups of your database files before beginning this procedure.

The following known limitations apply for this release:

- Oracle Cluster File System supports only datafiles and archive log files. You must use RMAN to perform hot backups.

7 APPENDIX: Scripts Reference

The following section provides full listings of scripts for reference.

- /var/opt/oracle/soft_startup.sh:

```
#!/bin/bash
#
# Dependencies : ORACLE_HOME environment variable set.
#
# Macro definition
#

LSMOD=/sbin/lsmod
INSMOD=/sbin/insmod
RMMOD=/sbin/rmmod
GREP=/bin/grep
#

SOFT_MARGIN=60;
SOFT_NOBOOT=1;
NOWAYOUT=0;

#
# Check for required environment variable ORACLE_HOME
#
check_env()
{
   if [ "X${ORACLE_HOME}" == "X" ]; then
      echo "ORACLE_HOME variable must be set before running this
script.";
      exit 1;
   fi;
   LD_LIBRARY_PATH=${ORACLE_HOME}/lib:/lib:/usr/lib
   PATH=${ORACLE_HOME}/bin:${ORACLE_HOME}/oracm/bin:${PATH}
   export LD_LIBRARY_PATH PATH;
}

#
# Unload softdog function
#
soft_unload()
 {
 if [ ${3} -eq 0 ]; then
   ${RMMOD} softdog;
   RC=$?;
   if [ ${RC} -ne 0 ]; then
      echo "Failed when unloading softdog module rc=$?";
      exit ${RC};
   fi;
   echo "Softdog module unloaded successfuly.";
else
   echo "Cannot unload softdog module.";
   echo "Softdog: Resource or device busy";
```

```
   exit 1;
fi;
}

#
# Load the softdog module
#
soft_load()
{
#   ${INSMOD} softdog nowayout=${NOWAYOUT} soft_noboot=${SOFT_
NOBOOT}
soft_margin=${SOFT_MARGIN};
   ${INSMOD} softdog soft_noboot=${SOFT_NOBOOT} soft_margin=${SOFT_
MARGIN};
   RC=$?;
   if [ ${RC} -ne 0 ]; then
      echo "Failed to load softdog module rc=${RC}";
      exit ${RC};
   fi;
}

#
# Check if it is running as root
#
USER=‘whoami‘;
if [ "${USER}" != "root" ]; then
   echo "Needs to run a root.";
   exit 1;
fi;
 #

# Check environment
#
check_env;

#
# Check for the softdog module.
#
soft_line=‘${LSMOD} | ${GREP} softdog‘;
if [ $? -eq 0 ]; then
   echo "Unloading softdog module";
   soft_unload ${soft_line};
fi;

#
echo "Loading softdog module ";
soft_load;

#
# Start up watchdogd and oracm
```

```
#
${ORACLE_HOME}/oracm/bin/watchdogd -d /dev/null -l 0 -m 10000
sleep 5
${ORACLE_HOME}/oracm/bin/oracm /a:0 < /dev/null > ${ORACLE_
HOME}/oracm/log/cm.out 2>&1 &
sleep 10
```

8 Documentation Accessibility

Our goal is to make Oracle products, services, and supporting documentation accessible, with good usability, to the disabled community. To that end, our documentation includes features that make information available to users of assistive technology. This documentation is available in HTML format, and contains markup to facilitate access by the disabled community. Standards will continue to evolve over time, and Oracle Corporation is actively engaged with other market-leading technology vendors to address technical obstacles so that our documentation can be accessible to all of our customers. For additional information, visit the Oracle Accessibility Program Web site at `http://www.oracle.com/accessibility/`.

Accessibility of Code Examples in Documentation JAWS, a Windows screen reader, may not always correctly read the code examples in this document. The conventions for writing code require that closing braces should appear on an otherwise empty line; however, JAWS may not always read a line of text that consists solely of a bracket or brace.

Accessibility of Links to External Web Sites in Documentation This documentation may contain links to Web sites of other companies or organizations that Oracle Corporation does not own or control. Oracle Corporation neither evaluates nor makes any representations regarding the accessibility of these Web sites.