

11G R1 RAC - X86-64 - RHEL5 - Installation Check list

Important Documents

[11g Clusterware Installation Guide for Linux Operating System](#)
[11g Storage Administration Guide](#)
[11g Real Application Clusters Installation Guide For Unix And Linux](#)
[11g Release Notes for Linux Operating System](#)
[Rac Assurance Team - Starter Kit](#)

Disclaimer

This document summarizes the main list of tasks that need to be completed, in order to facilitate a successful Oracle Real Application clusters installation. This document is not meant to be a comprehensive set of instructions to complete a Rac Installation, rather it is to be used as a companion to the Installation documents listed above. This document only lists the tasks to be completed and does not go into how to complete the tasks. The detailed instructions are in the documents listed above.

Hardware Requirements

Procure Certified Hardware

Confirm with the hardware vendor (server and disk) that the combination of hardware you have procured, is certified to run with Oracle RAC (ie the vendor has tested the configuration and have successful production implementations). Components include but are not limited to

- Server make and model
- Type of HBA
- HBA driver to be used
- O/S kernel version to be used
- Multipathing driver to be used.

You can find a list of Configurations that Oracle Corporation has validated, at [Oracle Validated Configurations](#)

Check [Rac technology Matrix for linux Platforms](#) for information regarding technologies supported to be used with Oracle Rac.

Make sure that you have redundant network interface cards to be used for the private nic and the public nic.

The Private interconnect should be connected using using a switch (Crossover cable is not supported). It is preferable to use a dedicated switch for this purpose. If you share a switch with other nodes, then you should create a VLAN on the switch including the ports used by the private interconnects.

Storage Requirements	You can choose to use a Storage Area Network , or Network Attached Storage as Shared storage to place your Database files on. When using a SAN, you can choose use the Fibre Channel Protocol or iScsi.
Memory Requirements	Atleast 1gb of physical memory on each server.
Host bus adapters	Ensure that you have atleast 2 HBA's on each node. Using multipathing software like Powerpath, HDLM to configure the HBA's.
Network Requirements	
Network interface requirements	<p>Each node must have atleast 2 network interfaces.</p> <p>We recommend having atleast 4 network interfaces. 2 for the private network and 2 for the public network.</p> <p>The NIC's used for the private network should be atleast GigE</p>
Nic Bonding	<p>If you have more than one network interface coard for each type of network interface (private and public), then bond a pair of interfaces together.</p> <p>For setting up bonding please refer to the following metalink notes</p> <p>298891.1</p>
Interface Names	<p>The public interface names associated with the network adapters for each network must be the same on all nodes, and the private interface names associated with the network adaptors should be the same on all nodes.</p> <p>For example: With a two-node cluster, you cannot configure network adapters on node1 with eth0 as the public interface, but on node2 have eth1 as the public interface. Public interface names must be the same, so you must configure eth0 as public on both nodes. You should configure the private interfaces on the same network adapters as well. If eth1 is the private interface for node1, then eth1 should be the private interface for node2.</p>
Network Protocol (Public Nic)	For the public network, each network adapter must support TCP/IP.
Network Protocol (Private Nic)	For the private network, the interconnect must support the user datagram protocol (UDP) using high-speed network

	adapters and switches that support TCP/IP (Gigabit Ethernet or better is a must).
Public IP address	<p>Assign One IP address with an associated host name (or network name) registered in the DNS for the public interface. Record the host name and IP address in the system hosts file also, /etc/hosts.</p> <p>Each node needs a public IP address</p>
Virtual IP address	<p>Assign One virtual IP (VIP) address with an associated host name registered in a DNS. Record the host name and VIP address in the system hosts file, /etc/hosts.</p> <p>Each node needs a virtual IP address.</p> <p>Select an address for your VIP that meets the following requirements:</p> <ul style="list-style-type: none"> - The IP address and host name are currently unused (it can be registered in a DNS, but should not be accessible by a ping command) - The VIP is on the same subnet as your public interface
Private IP address	<p>Configure the private interconnect network interface cards to have a private node name and a private IP address.</p> <p>This can be an interface bonded using IPMP.</p> <p>Choose a private IP address that is in the address range 10.*.* or 192.168.*.*</p> <p>The private IP address should be on a separate subnet than the public IP address.</p> <p>Oracle strongly recommends using a physically separate, private network (You can use a vlan on a shared switch if you must). You should ensure that the private IP addresses are reachable only by the cluster member nodes.</p>
Operating System Requirements	
Install Redhat Enterprise Linux 5 (Or Oracle Enterprise Linux)	Ensure that you are installing Redhat Enterprise Linux Update 5 (Best if you can be on the latest update). Install the 64 Bit, SMP. Version of the Operating System. All the

	nodes in the cluster should have the exact same version of the kernel (Verify this by running <code>uname -r</code>)
Swap Space	Equal to physical memory if physical memory < 8gb. .75 x physical memory if physical memory > 8gb
Temp Space	Atleast 400mb of space in /tmp, preferably atleast 2gb
Synchronize system clocks	<p>Implement NTP on all nodes (www.ntp.org).</p> <p>After implementing, check to make sure that the times on all nodes are within seconds of each other.</p> <p>If you are running on vmware (not supported), you will have to implement vmtools and sync up the time on the vmware hosts.</p>
Check package versions	<p>Ensure (on all nodes) that the versions of the following rpm's are equal or higher.</p> <p>You can also use the package named oracle-validated, if you are using OEL5 to setup your packages.</p> <p>binutils-2.17.50.0.6 compat-libstdc++-33-3.2.3 compat-libstdc++-33-3.2.3 (32 bit) elfutils-libelf-0.125 elfutils-libelf-devel-0.125 gcc-4.1.1 gcc-c++-4.1.1 glibc-2.5-12 glibc-2.5-12 (32 bit) glibc-common-2.5 glibc-devel-2.5 glibc-devel-2.5-12 (32 bit) libaio-0.3.106 libaio-0.3.106 (32 bit) libaio-devel-0.3.106 libgcc-4.1.1 libgcc-4.1.1 (32 bit) libstdc++-4.1.1 libstdc++-4.1.1 (32 bit) libstdc++-devel 4.1.1 make-3.81 sysstat-7.0.0</p>
Disable SELinux	Run <code>/usr/bin/system-config-securitylevel</code>

Disable Host firewall	Run <code>/usr/bin/system-config-securitylevel</code>
Synchronize system clocks	<p>Implement NTP on all nodes (www.ntp.org).</p> <p>After implementing, check to make sure that the times on all nodes are within seconds of each other.</p> <p>If you are running on vmware (not supported), you will have to implement vmtools and sync up the time on the vmware hosts.</p>
Set kernel parameters	<p>add to <code>/etc/sysctl.conf</code></p> <pre>kernel.shmall=3279547 kernel.shmmax=3700000000 kernel.sem=250 32000 100 142 fs.file-max=327679 net.ipv4.ip_local_port_range=1024 65000 net.core.rmem_default=4194304 net.core.rmem_max=4194304 net.core.wmem_default=262144 net.core.wmem_max=262144 net.ipv4.tcp_rmem=4194304 4194304 4194304 net.ipv4.tcp_wmem=262144 262144 262144</pre> <p><code># /sbin/chkconfig boot.sysctl on</code></p> <p>use a text editor to change the <code>/etc/boot.sysctl</code> parameter <code>RUN_PARALLEL</code> flag from yes to no.</p> <p>Find out the gid of your dba group (Let us say it was 501) <code>echo 501 > /proc/sys/vm/hugetlb_shm_group</code></p> <p>add following line to <code>/etc/sysctl.conf</code> <code>vm.hugetlb_shm_group=501</code></p> <p>add to <code>/etc/security/limits.conf</code></p> <pre>oracle soft nofile 131072 oracle hard nofile 131072 oracle soft nproc 131072 oracle hard nproc 131072</pre> <p>as root run <code>sysctl -p</code>, so these values take effect.</p> <p>Add following lines to <code>/etc/pam.d/login</code> <code>session required pam_limits.so</code></p>

<p><u>Ensure that the hangcheck-timer module is loaded at system startup.</u></p>	<p>Metalink Note 726833.1</p> <p>lsmod grep -i hang (if this does not show you one line, then hangcheck timer is not loaded). If it is not loaded proceed to the next step.</p> <p>as root run /sbin/insmod /lib/modules/2.6.9-22.0.1.ELsmp/kernel/drivers/char/hangcheck-timer.ko hangcheck_tick=1 hangcheck_margin=10 hangcheck_reboot=1</p> <p>also add the command to /etc/rc.d/rc.local</p> <p>in the above insmod command, substitute 2.6.9-22.0.1.ELsmp, with your exact kernel version.</p>
<p>Configure a public network interface on each node.</p>	<p>This interface may be bonded (Not required). Make sure that the interface has the same interface name on each node (eg: bond0). Make sure that this interface has a valid routable ip address (Address OTHER THAN (10.*,172.16.*,192.168.*) (RFC 1597)). This node name and ip address should be added to the dns.</p>
<p>Configure a private network interface on each node.</p>	<p>The only traffic that traverses this network should be the communication between the two nodes. They should be connected through a network switch. This network interface should not be pingable from other nodes (ie the only nodes that should be able to ping the private interconnects should be the nodes in the cluster themselves) You can give them a non-routable (10.*, 172.16.*, 192.168.)(RFC 1597) address. Make sure that the private network interface on each node has the same interface name (eg: eth1). This node name and ip address should be added to /etc/hosts on all nodes.</p>
<p>Assign VIP's</p>	<p>One virtual Ip address should be allocated to each node. This should be a IP address which is assigned but not tied to any network interface. The configuration of the virtual interface should be entered in the /etc/hosts of the cluster nodes and the DNS. Eg if the node name is racsrv1, you can name the virtual interface racsrv1-v, and assign it an IP address. But do not configure any physical interface on the server to use that IP address (Double check by running ifconfig -a and checking the output to make sure that the</p>

	the IP address is not in use).
Create the dba group	<p>Create the group named dba with the same gid on each server.</p> <p>If you choose you could create separate groups for oinstall,dba,oper,asm.</p>
Create Oracle Software Owner	Create the linux user oracle, on each node and assign it to the group dba. Setup bash to be the default login shell. The users should have the same uid on every node.
Check the user “nobody”	Make sure that the user named “nobody” exists on the system
Check /etc/hosts	<p>Make sure that, in your /etc/hosts you have entries for each node in your cluster for</p> <ul style="list-style-type: none"> - Your public node name, public node name.domainname - Your private node name,private node name.domainname - Your vip node name, vip node name.domainname. <p>The /etc/hosts on each node should have entries for all nodes in the cluster, ie it should have the consolidated node list.</p> <p>It has to be there in /etc/hosts (In addition to the public node namd and vip node name being in the DNS)</p>
Configure SSH	<p>Configure SSH on the cluster nodes as per instructions in Metalink Note 300548.1 .</p> <p>Make sure that you can ssh from each node in the cluster to every other node in the cluster, using the hostname , hostname.domainname,hostname-priv,hostname-priv.domainname, without being prompted for the password. Also ensure that you do not have any banners showing when you ssh.</p>
Allocate a local disk mount point for the CRS_HOME	<p>This mount point should be different from that allocated for ORACLE_HOME and ASM_HOME (eg: /u01/crs). Allocate atleast 1gb of space for this mount point.</p> <p>The oracle user should have read write access to this directory</p>
Allocate a local disk mount	This mount point should be different from that allocated for

<p>point for the ASM_HOME</p>	<p>ORACLE_HOME and CRS_HOME (eg: /u01/asm). Allocate atleast 3gb of space for this mount point.</p> <p>The oracle user should have read write access to this directory</p>
<p>Allocate a local disk mount point for the ORACLE_HOME</p>	<p>This mount point should be different from that allocated for the CRS_HOME (eg: /u01/app). Allocate atleast 3Gb of space for this directory.</p> <p>The oracle user should have read write access to this directory</p>
<p>Allocate 3 devices for the voting disk (This assumes an 11g clusterware installation) (11g clusterware can use block devices for the voting disk) (10g clusterware requires raw devices)</p>	<p>This device should preferably mirrored at the hardware level. Size 280Mb. Configure this device by creating a partition on the lun, that skips the first 1Mb on the LUN. For best availability, keep these three luns on physically separate disks.</p> <p>The devices can be emc powerpath devices.</p> <p>The device names that point to the same lun, should be exactly the same on both nodes.</p> <p>Eg: if /dev/rdisk/xxdisk1 points to lun1 on node 1 then /dev/rdisk/xxdisk1 has to point to lun1 on the second node.</p> <p>After creating the devices run the following command to make absolutely sure that you have 280Mb of free space.</p> <pre>dd if=/dev/zero of=/dev/emcpowerb1 bs=+1M count=280 (Substitute emcpowerb1 with your device name)</pre> <p>Change the ownership and permissions of the disks</p> <pre>chown root:dba /dev/emcpowerb1 chmod 660 /dev/emcpowerb1</pre> <p>Add this permission change to /etc/rc.d/rc.local to persist this permission settings on reboots.</p> <p>You can also use device mapper, to configure multipathing.</p>
<p>Allocate 2 devices for the</p>	<p>This device should preferably be mirrored at</p>

<p>cluster registry (This assumes an 11g clusterware installation) (11g clusterware can use block devices for the voting disk) (10g clusterware requires raw devices)</p>	<p>the hardware level. Size 280 Mb. Configure this device by creating a partition on the lun, that skips the first 1Mb on the LUN. Device should preferably mirrored at the hardware level.</p> <p>The devices can be emc powerpath devices.</p> <p>Eg: if /dev/emcpowerb1 points to lun1 on node 1 then /dev/emcpowerb1 has to point to lun1 on the second node.</p> <p>After creating the devices run the following command to make absolutely sure that you have 100Mb of free space.</p> <pre>dd if=/dev/zero of=/dev/emcpowerb1 bs=+1M count=280 (Substitute raw1 with your rawdevice name)</pre> <p>Change the ownership and permissions on the devices</p> <pre>chown oracle:dba /dev/emcpowerb1 chmod 660 /dev/emcpowerb1</pre> <p>Add this permission change to /etc/rc.d/rc.local to persist this permission settings on reboots.</p> <p>You can also use device mapper, to configure multipathing.</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Ocfs2 Requirements

<p>Ocfs2</p>	<p>OCFS2 can be downloaded from http://oss.oracle.com/projects/ocfs2/files/RedHat/RHEL5/</p> <p>Check if the following packages exist</p> <pre>ocfs2-tools ocfs2console ocfs2-tools-debuginfo ocfs2console</pre> <p>If they are not available they can be downloaded from http://oss.oracle.com/projects/ocfs2-tools/files/RedHat/RHEL5/</p> <p>The users guide can be found at http://oss.oracle.com/projects/ocfs2/documentation/v1.4/</p>
--------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ASM Requirements

<p>Download asmlib</p>	<p>Download the asm rpm's from http://www.oracle.com/technology/software/tech/linux/asmlib/rhel5.html</p> <p>oracleasm-support-2.1.3-1.el5.x86_64.rpm oracleasmlib-2.0.4-1.el5.x86_64.rpm</p> <p>Download also the smp version of oracleasm-xxxx-smp-xxxx.rpm (The version that matches your kernel version).</p>
<p>Allocate ASM disks for your ASM Data disk group</p>	<p>Take the total amount of space you require to place your database files (data, index, system, undo , temp, redo) and divide this up to create as many luns as possible. Rule of thumb is to present at least 4 luns and have the number of luns as close to the number of spindles you are allocating.</p> <p>Present these luns to the operating system.</p> <p>Eg: if you need to present 500gb of space then create and present 10 luns of 50gb each.</p> <p>These luns can be emc powerpath devices.</p> <p>Eg: if /dev/emcpowerb1 points to lun1 on node 1 then /dev/emcpowerb1 has to point to lun1 on the second node.</p> <p>After the powerpath devices are configured, use fdisk to create a partition of the disk, skipping the first 1mb of the disk. This new partition is what ASM will use.</p> <p>Change the ownership and permissions on the devices</p> <pre>chown oracle:dba /dev/emcpowerb1 chmod 660 /dev/emcpowerb1</pre> <p>Add this permission change to /etc/rc.d/rc.local to persist this permission settings on reboots.</p>
<p>Create Disks for your Flash Recovery Area ASM disk group</p>	<p>Take the total amount of space you require to place your Flash recovery area and divide this up to create as many luns as possible.</p> <p>Rule of thumb is to present at least 4 luns and have the number of luns as close to the number of spindles you are allocating.</p>

	<p>Present these luns to the operating system.</p> <p>Eg: if you need to present 500gb of space then create and present 10 luns of 50gb each.</p> <p>These luns can be emc powerpath devices.</p> <p>Eg: if /dev/emcpowerb1 points to lun1 on node 1 then /dev/emcpowerb1 has to point to lun1 on the second node.</p> <p>After the powerpath devices are configured, use fdisk to create a partition of the disk, skipping the first 1mb of the disk. This new partition is what ASM will use.</p> <p>Change the ownership and permissions on the devices</p> <pre>chown oracle:dba /dev/emcpowerb1 chmod 660 /dev/emcpowerb1</pre>
Download Oracle Software	<p>http://www.oracle.com/technology/software/products/database/oracle11g/111060_linux8664soft.html</p> <p>http://download.oracle.com/otn/linux/oracle11g/linux.x64_11gR1_clusterware.zip</p> <p>http://download.oracle.com/otn/linux/oracle11g/linux.x64_11gR1_database_1013.zip</p>
Download Oracle patches for the linux x86-64 platform.	11.1.0.7 6890831
Cluster Verification Utility	
Download cluster verification utility (x86-64).	<p>http://www.oracle.com/technology/products/database/clustering/cvu/cvu_download_homepage.html</p> <p>You will have to set the environment variables CV_JDKHOME and CVU_HOME before running cluvfy.</p>
Run CLUVFY	
Run the following cluvfy commands and verify the output	<pre>cluvfy stage -post hwos -n node1,node2 -verbose cluvfy stage -pre crsinst -n node1,node2 -orinv dba</pre>

	substitute node1,node2 with your nodenames.
Miscellaneous	
X-windows client	Make sure that, on the desktop from which you are going to login to the servers and perform the installation, that you have a good xwindows client like exceed. (Open source clients like xming and/or xmanager will not work). If you are using reflections, then make sure that you have the latest version of reflections.
Orion	If you want to do some i/o testing you can download oracle's free tool orion from http://download.oracle.com/otn/utilities_drivers/orion/orion_linux_em64t.gz The instructions for using it can be found at http://download.oracle.com/otn/utilities_drivers/orion/Orion_Users_Guide.pdf