

Oracle 10g RAC Configuration on VMware ESX Server

This document describes the steps to configure an Oracle 10g RAC environment on Windows running on VMware ESX Server. By using VMware's advanced technology, virtual hardware becomes standardized enabling the ability to deploy Oracle RAC environments across a wealth of various hardware platforms without having to make modifications to the design. This configuration suggests using 3 separate LUNs for the Oracle Cluster Ready Services (CRS) and the Automatic Storage Management (ASM) shared storage. This is not a requirement, rather, it is a suggested configuration. This document assumes that the administrator implementing the solution has a familiarity with the concepts and administration of VMware ESX Server 2.x and VMware VirtualCenter 1.x. The scope of this document does not cover the following topics: ESX performance enhancements, Virtual node performance, or Oracle software configuration.

Physical ESX Host Configuration

1. If you already have a Public VMFS LUN created for hosting your virtual disk files (.vmdk files), please skip the remainder of step 1 and continue with step 2. Create a LUN and allocate it to all of your ESX servers' respective HBA WWPNs. To connect to the allocated SAN disk use the following step:
 - a. The ESX server can rescan its HBA for newly allocated disk without disturbing the other instances that are running on top of the ESX host. Once the LUNs have created/masked to the WWPN of each ESX host, login to the VMware Management User Interface (VMware MUI) and navigate to Options | Storage Management.
 - b. Click on "Rescan SAN" (Note: this may take 2 attempts of rescanning before the SAN disk is presented. **If you are using 2 HBAs, you will still only see one path to the SAN disk. This is due in part to VMware using an Active/Passive strategy for their FC HBA redundancy.*
 - c. Create the new disk as "Public" VMFS volume will be needed to store the virtual disk files (.vmdk files) that contain the operating system & associated applications. **Public VMFS is the default configuration when creating VMFS filesystems.*
2. Create another 3 LUNs. These new LUNs will not be formatted by the ESX server and will be used by the virtual nodes as RAW disk. Use the same steps as in step 1a & 1b to have the ESX server recognize this new disk.
3. Create another new LUN and configure it as a Shared VMFS. Use the same steps as in step 1a & 1b to have the ESX server recognize this new disk. This Shared VMFS will be the storage point for all the Raw Device Mappings (RDMs) that you will create in the next step. **A RDM is simply a file that acts as a pointer to a raw-attached LUN. It masquerades as a .vmdk file and takes on the characteristics of the underlying raw disk.*
4. In order for VMware ESX to be able to pass the RAW disk LUNs through the virtualization kernel (vmkernel) to the virtual nodes, the RDMs are required for the ESX host to locate the disk and pass through I/O to and from the virtual nodes. There are several steps that are required for this configuration:
 - a. To create a RDM you need to determine the physical name of the adapter that the ESX host has assigned to each LUN. You can find the physical names of the adapter by either:

- i. Browsing the Storage Management section of the VMware MUI.
- ii. By logging into the service console (command line) and viewing the contents of /vmfs/. Ex. [root@localhost vmfs]# ls /vmfs/vmhba*.

**The format of the physical adapter name should be: vmhbaN:X:Y*

Where *N* is the number of the adapter; *X* is the target id; *Y* is the LUN number.

- b. Once determining the physical name of each LUN, you can create a RDM. Each RDM should be placed in the Shared VMFS that you created in step 3. The reason for using a Shared VMFS is so that vmkernel allows the virtual nodes to access a physical SCSI disk at the same time without causing conflicts or SCSI locking. If you have not already done so, login to the ESX server as root or with root equivalent privileges. Once logged in, issue the following command once for each of 3 LUN devices(case sensitive):

```
vmkfstools -r vmhbaN:X:Y:0 /vmfs/SHAREDVMFS/myRDMname.vmdk
```

**The 0 in the physical name "vmhbaN:X:Y:0" (which represents the partition number) is not assumed by the vmkfstools -r command, and is therefore required.*

Virtual Node Configuration

1. Deploy 2 standard imaged windows instances using VirtualCenter.*
 - a. Place a node on each of the two (2) ESX hosts. Configure each node to store its first hard disk (vmdk file that will contain the OS and applications) on the Public VMFS.

** If necessary, add additional disk for binaries, logs, etc. As an aside, additional disk should be converted to dynamic disk to prevent vmdk disks from growing to an unmanageable quantity. A Windows dynamic configured disk on VMware ESX gives an administrator the ability to extend the .vmdk file and then extend the windows volume contained there within.*
2. Once each node is deployed, you will need to make some configuration changes to the hardware profile of each virtual node prior to configuring the Operating System.

**For the purposes of this document, it is assumed that you are using VirtualCenter to configure your virtual nodes.*

 - a. Add a second NIC to each virtual machine. Connect each NIC on each node to a virtual switch that has connectivity to a public interface. The ESX hosts cannot communicate with one another when using a private virtual switch.
 - b. Add each of the 3 RAW disks (RDM files) to each of the instances. When adding the disk make the following selections:
 - i. In the "Add Hardware" wizard, select "Hard Disk" and click "Next"
 - ii. In the "Select a disk" dialogue select "System LUN" and click "Next"
 - iii. In the "Select and Configure a Raw LUN" dialogue, select one of the physical LUN names from the "Target LUN" drop-down and choose the "Physical" option from the "Compatibility" area. Click Finish.

- iv. Once you return to the properties of the virtual machine, select the newly attached physical hard disk and change the "Virtual device node" selection to a separate SCSI controller and SCSI ID (SCSI1:0).
- v. From the "Compatibility mode" section located above the "Virtual device node" section, ensure "Physical" is selected. Click OK for all your configuration changes to take effect.
- vi. Return to the properties of the virtual machine and ensure all configuration changes have taken place.
- vii. Select the newly created "SCSI Controller 1" from the list of devices and change the selection in the "SCSI Bus Sharing" drop-down from "None" to "Physical". This enables the virtual SCSI Controller (SCSI1) to share each of the attached devices with other virtual SCSI controllers on other physical hardware. For all changes to take effect, click OK.
- viii. Repeat steps 6b.i - 6b.vi for each of the other 2 RAW disks (RDM files). These RAW devices should also be placed on SCSI1:1 and SCSI1:2, respectively. Be sure to mimic the same configuration settings for both virtual nodes.

Virtual Nodes' Windows OS Configuration

1. Power on the first of the virtual nodes, leaving the second node powered off.
2. After logging into the virtual node, ensure that the VMware tools are installed to the latest version by running the installer from VirtualCenter.
3. Configure each of the network adapters; one adapter for a public interface, and one adapter for the internal communications.
4. Configure the 3 SCSI attached RAW disks in preparation for Oracle RAC. Perform the following steps to disable write caching on all the shared disks that you intend to use for your database files. You must do this from both nodes in your cluster:
 - a. Navigate to the Windows Device Manager and click on "Disk drives".
 - b. Expand the Disk drives hive and double-click the first drive listed.
 - c. Under the "Disk Properties" tab for the selected drive, uncheck the option that enables the write cache.
5. Repeat steps 4a – 4c for all the disks within the "Disk drives" hive.
6. Cluster Ready Services (CRS) provides overall management of the cluster activities. CRS requires two key files that must be located in logical drives on the shared disks: one for a Voting Disk and one for the Oracle Cluster Registry. Complete the following steps to configure both of these required logical drives:
 - a. From one of the existing nodes of the cluster, navigate to "Computer Management" and select "Disk Management".
 - b. Identify a shared disk that does not contain a primary partition and has free space available in one or more extended partitions.
 - c. Right click inside an unused part of an extended partition and choose "Create Logical Drive".
 - d. From the Disk configuration wizard, enter "20" (MB) as the size that you want for the voting disk logical drive.
 - e. Choose the option "Do not assign a drive letter" and then choose the option "Do not format this partition". Click "Finish" on the last page of the wizard.
 - f. Repeat Steps 6a – 6e on a second shared disk. Create a second logical drive of 100 MB for the Oracle Cluster Registry.

7. Check the two nodes in the cluster to ensure that the partitions are visible on both of them and to ensure that the Oracle partitions DO NOT have drive letters assigned.
8. You need to configure disk storage for use with Automatic Storage Management (ASM). ASM storage consists of one or more disk groups, each of which can span multiple disks. You need to create two or more logical drives, each on a different shared disk, for a total of at least 7 GB. To prepare each logical drive, perform these steps:
 - a. From one of the existing nodes of the cluster, navigate to "Computer Management" and select "Disk Management".
 - b. Choose the remaining shared disk and right click inside of the disk area.
 - c. Choose "Create Logical Drive" a wizard presents pages for configuring the logical drive.
 - d. Enter the size that you want for the partition. All of the partitions that you create for your ASM disk groups should include as much of the free space on the disk as possible. Additionally, however, the partitions should all be the same size.
 - e. Choose the option "Do not assign a drive letter" and then choose the option "Do not format this partition". Click "Finish" on the last page of the wizard.
9. Check the two nodes in the cluster to ensure that the partitions are visible on both of them and to ensure that the Oracle partitions DO NOT have drive letters assigned.